A scalable photonic computer solving the subset sum problem

Xiao-Yun Xu1,2, Xuan-Lun Huang1,2, Zhan-Ming Li1,2, Jun Gao1,2, Zhi-Qiang Jiao1,2, Yao Wang1,2, Ruo-Jing Ren1,2, H. P. Zhang3, Xian-Min Jin1,2,4*

The subset sum problem (SSP) is a typical nondeterministic-polynomial-time (NP)–complete problem that is hard to solve efficiently in time with conventional computers. Photons have the unique features of high propagation speed, strong robustness, and low detectable energy level and therefore can be promising candidates to meet the challenge. Here, we present a scalable chip built-in photonic computer to efficiently solve the SSP. We map the problem into a three-dimensional waveguide network through a femtosecond laser direct writing technique. We show that the photons sufficiently dissipate into the networks and search for solutions in parallel. In the case of successive primes, our approach exhibits a dominant superiority in time consumption even compared with supercomputers. Our results confirm the ability of light to realize computations intractable for conventional computers, and suggest the SSP as a good benchmarking platform for the race between photonic and conventional computers on the way toward “photonic supremacy.”

INTRODUCTION

NP-complete problems (1) are typically defined as the problems solvable in polynomial time on a nondeterministic Turing machine (NTM), which indicates that these problems are computationally hard on conventional electronic computers, a general type of deterministic Turing machines. The subset sum problem (SSP) with practical application in resource allocation (2) is a benchmark NP-complete problem (3), and its intractability has been harnessed in cryptosystems resistant to quantum attacks (4, 5). Given a finite set S of N integers, the SSP asks whether there is a subset of S whose sum is equal to the target T. Apparently, the number of subset grows exponentially with the problem size N, which leads to an exponential time scaling and thus strongly limits the size of the problem that can be tackled in reality.

Despite the immense difficulty, some researchers attempt to solve NP-complete problems in polynomial time with polynomial resource. A memcomputing machine (6, 7) as powerful as an NTM has been demonstrated, while the ambitious claim is not valid in a realistic environment with inevitable noise (8). Designs of an NTM, where the magical oracles (1, 9, 10) are realized by simultaneously exploring all computation paths, are proposed (11, 12). In the cost of space or material, parallel exploration provides an alternative to decrease time consumption. As time is irreversible, not reusable and completely out of our charge, it is reasonable to trade physical resources for it. Besides the above NTM proposals, similar measurements have been taken, for instance, the increasingly powerful electronic supercomputers with an integration of an increasing number of processors (13) and molecule-based computation using large quantities of DNAs or motor molecules (14–18). Furthermore, optimized algorithms are applied to specific instances (19–21).

Although improvements have been made, conventional electronic computers are ultimately limited by heat dissipation problem (16), which is also a possible limitation for memcomputing machines consisting of commercial electronic devices (8). The molecule-based computation is limited by the slow movement (16–18) or the long reaction process (14, 15). Quantum computation is still hindered by decoherence and scalability (22, 23). Other proposals are still in the stage of theory (11, 12, 24–26). However, we notice that photons have been extensively applied in proof-in-principle demonstrations of supercomputing (27) even without quantum speed-up, including an NP problem such as prime factorization (28) and NP-complete problems such as traveling salesman problem (29), Hamiltonian path problem (30–32), and dominating set problem (33). The #P-complete problem, boson sampling (34–39), other computational functions (40), and algorithms (41, 42) are also demonstrated in a photonic regime. The successful applications imply that photons are potential excellent candidates to solve the SSP.

Here, we present a photonic computer constructed with chips serving as processing units to solve the SSP in a physically scalable fashion. Like the current signal in electronic computers or the molecule in molecular computers, photons contained in the optical source are treated as individual computation carriers. They travel in chips along buried waveguide networks to perform parallel computations. The specific instances of the problem are successfully encoded into the networks according to particular rules. The existence of target sums is judged by the arrival of photons to the corresponding output ports of the networks. We further investigate its scalability and performance in time consumption, showing the photon-enabled advantages.

RESULTS

Configuration of the photonic computer for the SSP

The proposed photonic computer solving the SSP can be classified as a non–Von Neumann architecture (see the Supplementary Materials...
for its role in the evolution of computers). As shown in Fig. 1A, the photonic computer consists of an input unit, a processing unit, and an output unit. The input unit is used to generate horizontally polarized photons at 810 nm. Photons are then coupled into the processing unit to dissipate into the waveguide network to execute the computation task. After photons are emitted from the processing unit, the evolution results are read out by the output unit. Here, the processing unit is an analog to the central processing unit (CPU) of an electronic computer, playing a key role in the computation. In the following, we will discuss the design of the processing unit from mathematical and physical implementation aspects to illustrate its capability of solving the SSP.

The processing unit can be represented by an abstract network composed of nodes and lines, which is primarily based on the proposal of Nicolau et al. (16), while physical implementation has to be designed to fit integrated photonics. As the network for the specific instance {2, 5, 7, 9} in Fig. 1B shows, there are three different types of nodes representing split junctions, pass junctions, and converge junctions. Note that although the circular yellow nodes overlap with the hexagonal nodes in the abstract network, they are physically separate, as the waveguide network in Fig. 1A presents. Once the photons enter the network from the top node, the computation process is activated. The photons are split into two portions at hexagonal nodes (split junctions), traveling vertically and diagonally. When meeting the circular white nodes (pass junctions), the photons proceed along the original directions. Meanwhile, the circular yellow nodes (converge junction), located at the end of the diagonal routes that start from the former row of hexagonal nodes, are responsible for transferring photons from diagonal lines to vertical lines before the next splits.

The specific SSP is encoded into the network according to particular arithmetical and scalable rules: (i) The vertical distance (measured as the number of nodes) between two subsequent rows of hexagonal nodes is equal to the value of the element from the set {2, 5, 7, 9}, as denoted by the integers on the left. (ii) The diagonal routing leads to a horizontal displacement of photons, whose magnitude is also equal to the integer on the left. The diagonal movement of photons represents that the corresponding element is included into the summation. On the contrary, the vertical movement means that the element is excluded from the summation. (iii) The value of the ultimate sums is equivalent to the spatial position of the output signals, as denoted by the port numbers. For example, the path for port 14, highlighted by a translucent gray band, reveals that only elements 5 and 9 contribute to the subset sum 14. Owing to the vast parallelism, the photons arrive at the output ports with all possible subset sums generated.

We fabricate the processing unit in Corning Eagle XG glass with femtosecond laser direct writing technique (see Materials and Methods). The top left corner of the waveguide network in Fig. 1A and the abstract network in Fig. 1B are depicted in detail in Fig. 1 (C to F). As we can see, the split junction is realized by a modified three-dimensional beam splitter, where the two waveguides first couple evanescently (red segment) and then decouple with one of the waveguides climbing upward and the other proceeding along the initial direction. To avoid extra loss, a vertical decoupling distance of 25 μm is deliberately selected. Meanwhile, the coupling length and coupling distance are set to 1.8 mm and 10 μm, respectively, to achieve the desirable splitting ratio. As the intensity distribution in Fig. 1D reveals, the modified beam splitter is unbalanced, which is on the purpose of compensating the bending loss caused by the subsequent arc ¯m and arc ¯n.

The converge junction is almost the mirror-image split junction except with a different coupling length of 3.3 mm. The photons in path ¯g should be completely transferred to path ¯h in an ideal case, and the residual induced by the imperfect fabrication is minimized. The output intensity at port g is less than 3% of that at port h. The pass junction is implemented in the form of one waveguide crossing over the other at a decoupling vertical distance of 25 μm. As the output intensity in Fig. 1F reveals, the three-dimensional architecture ensures an excellent pass junction whose extinction ratio is around 24 dB. Supported by the powerful fabrication capability of femtosecond laser writing technique (43, 44), we are able to map the abstract network of the SSP into a three-dimensional photonic chip.

Experimental demonstration

We demonstrate the computation of the SSP at the specific cases of {3, 7, 11} and {3, 7, 9, 11}. As shown in Fig. 2 (A and C), the evolution results are read out from a one-shot image, where the photons appear in a line of spots. Every single spot is an accepted witness of the existence of the corresponding sum (denoted by the integer below the spot) if the experiments are trusted. Because the involved problem size is not too large, we check the reliability of our experimental results by enumeration and conclude that all the spots observed are supposed to appear and that none of the expected results is absent.

The reliability of our experiments is further investigated by a comprehensive analysis of the intensity distribution, as presented in Fig. 2 (B and D). We calculate the theoretical distribution through a lossless model consisting of balanced split junctions, perfect pass junctions, and ideal converge junctions. Therefore, the theoretical outcomes can be regarded as benchmarks of the SSPs. For the case of {3, 7, 11}, the theoretical result is either 0 or 0.125, while it is either 0 or 0.0625 in the case of {3, 7, 9, 11}. In this theoretical regime, zero intensity indicates that a sum does not exist; otherwise, it exists.

We apply a threshold to analyze the retrieved intensity for every output port. A valid appearance can be identified if the intensity goes beyond a reasonable threshold; otherwise, an absence can be confirmed (highlighted by solidus pattern). The tolerant intervals of the thresholds applicable in our experiment are presented with bands filled with a solidus in Fig. 2 (B and D), straightforward revealing the lower bounds and the upper bounds. Beneficial from the good signal-to-noise ratio obtained in our experiments, there is a wide tolerant band to accept a large range of thresholds, which implies the great accuracies of our experiments and verifies the feasibility of our approach.

Time-consumption budget

We find that the optical source launched into the photonic circuit has a notable influence on the performance of our photonic computer. Note that the photonic supremacy in time consumption over other schemes is achieved by classical light (a stream of photons), not quantum light. We obtain the same evolution results with both classical light and quantum light, and the heralded single-photon source fails to outperform the classical light. This phenomenon is attributed to the fact that a bunch of photons arrive together in the case of classical light, while heralded single-photon source only launches one photon at a time. Under these circumstances, it takes longer time with quantum light to collect enough signal photons to be distinguished from the environment, leading to a worse performance
Fig. 1. Schematic of the design and setup. (A) A power-adjustable and horizontally polarized optical source is guaranteed by the quarter-wave plate (QWP), half-wave plate (HWP), and polarization beam splitter (PBS) in the input unit. The photons at 810 nm are prepared and coupled into the network in the processing unit and then travel to generate all possible subset sums. The evolution results at the output ports are retrieved by the charge-coupled device (CCD) to testify the existence of the corresponding sums. (B) The abstract network for the specific instance {2, 5, 7, 9} is composed of three different kinds of nodes representing split junctions, pass junctions, and converge junctions. Split junctions (hexagonal nodes) divide the stream of photons into two portions. One portion moves vertically, and the other travels diagonally. Pass junctions (circular white nodes) allow the photons to proceed along their initial directions. Converge junctions (circular yellow nodes) play a role in transferring photons from diagonal lines to vertical lines. Although the circular yellow nodes overlap with the hexagonal nodes in the abstract network, they are physically separate, as shown in (A). Photons traveling diagonally from a split junction to the next split junction represent including an element into the summation. The value of the element is equal to the number of junctions between two subsequent rows of split junctions, as denoted by the integers on the left. The generated subset sums are equal to the spatial positions of the output signals, as the port numbers denote. (C) The x-y view of the top left corner of the waveguide network in (A) and the abstract network in (B) is composed of the three basic junctions whose x-z views are shown in (D) to (F). The split junction is realized by a modified three-dimensional beam splitter where a coupling distance of 10 μm, a coupling length of 1.8 mm, and a vertical decoupling distance of 25 μm are deliberately selected, leading to a desirable splitting ratio. The unbalanced output of split junctions, revealed by the intensity distribution in (D), is designed to compensate for the bending loss caused by the subsequent arc cm and arc mf in (C). The converge junction is almost a mirror-image split junction except with a different coupling length of 3.3 mm. The residual in port g is small enough to be ignored. A vertical decoupling distance of 25 μm guarantees an excellent pass junction whose extinction ratio is around 24 dB, as the intensity distribution in (F) presents.
than classical light and making it more challenging to surpass electronic computers (see the Supplementary Materials).

To show the photon-enabled advantages, we further investigate the time-consumption performance in the case of classical light. Here, the computing time is determined by the propagation speed of photons and the longest path in the waveguide network. Owing to the fast movement of flying photons and the compactness of the chip-based networks, it only takes the processing units a fraction of 1 ns to accomplish the computations in our experiments, which has already surpassed many representative electronic computers emerging in these decades (see the Supplementary Materials).

Furthermore, the potential of our approach is explored in the context of successive primes by comparing with other competitors (see Materials and Methods for time estimation of different approaches), as shown in Fig. 3. It is noticed that the photonic computer has a significant advantage over the molecular computation, which is attributed to the similar time scaling resulting from the similar configuration of the computing networks, and the superiority of photons in moving speed over molecules, i.e., ~2 × 10^11 mm/s for 810 nm photons in waveguides and ~5 × 10^3 mm/s for actin filaments (16). Although faster biological molecules are reported in a latest research (18), they are still on the long journey of chasing after photons.

The time consumption of representative electronic competitors with the conventional Von Neumann architecture, characterized by floating point operations per second (FLOPS) (45), is also presented. It is found that the photonic computer outperforms the state-of-the-art CPU (46) at a small size that is probably accessible in subsequent experimental demonstrations. Compared with the graphics processing unit (GPU) (47), the photonic computer exceeds it until N = 12. Apparently, it is increasingly challenging to beat an increasingly strong competitor. Nevertheless, the most powerful supercomputer (13), Summit, composed of an enormous number of CPUs and GPUs, can be also surpassed at a modest size of 28. Besides, the superiority of the photonic computer is reinforced with the growth of problem size, as the trend reveals. Even at a medium size, our approach consumes many orders of magnitude shorter computing time than the molecular and electronic rivals, exhibiting strong competitiveness in solving the SSP in the case of successive primes (see Materials and Methods for the speed-up of our photonic computer).

**DISCUSSION**

In summary, we demonstrate a photonic computer solving the SSP by mapping the problem into a waveguide network in a three-dimensional architecture. With the demonstrated standardized structure of basic junctions, regular configuration of the network, and the mature femtosecond laser writing technique, the SSP can be encoded into a physical network and conveniently solved in a scalable fashion. The computational power is further analyzed by investigating the time-consumption performance. The results suggest that, for successive primes, photonic computers are very likely to beat the most powerful supercomputer with a near-future accessible problem size.
Other performances, such as signal-to-noise ratio and Fisher information, are also discussed (see the Supplementary Materials).

The photon-enabled advantage in solving the SSP can be understood from the unique features of light. First, light is essentially a stream of photons, which can be sufficiently used to probe all the paths in parallel, by being dissipated into a large network with very small fraction of light in each path (can be down to single-photon level). Second, the ultimate speed of flying photons makes the evolution time very short in the designed structures, even for a large and complicated photonic network. Third, photons can be confined in a very limited space with the technique of integrated photonics, which is beneficial to both the computing speed and scalability. Last but not least, interference is a unique strength of photons, whereas we cannot see its contribution to the speed-up of the proposed photonic computer. Nevertheless, it can be potentially used to achieve a reconfigurable photonic computer for different SSPs in the future (see the Supplementary Materials).

Besides the fundamental interest of racing with conventional electronic computers, it would be more fascinating to map many real-life problems into the framework of solving the SSP, which may boost the building of this photonic computer toward industrialization. It is also possible, but still open, to solve other NP problems in this purpose-built photonic computer. In light of the fact that any NP problem can be reduced to an NP-complete problem efficiently (3), any NP problem is able to be mapped to the proposed network in principle. Therefore, a photonic solution of the SSP implies possible solutions of a wide range of NP problems. Moreover, photon-enabled unique feature may also show its strength in other new computing architectures (48, 49).

MATERIALS AND METHODS

Photonic chip fabrication
Waveguide networks in three-dimensional architecture were written by the femtosecond laser with a repetition rate of 1 MHz, a central wavelength of 513 nm, a pulse duration of 290 fs, and a pulse energy of 190 nJ. Before radiating into the borosilicate substrate at a depth of 170 μm, the laser beam was shaped by a cylindrical lens and then focused by a 100× objective with a numerical aperture of 0.7. During the fabrication, the translational stage moved in x, y, and z directions according to the user-defined program at a constant speed of 15 mm/s. The careful measurements and characterization on the geometric parameter dependence of the three types of junction, such as coupling length, coupling distance, decoupling distance, and curvature, were taken to optimize the performance to form the standard elements.

Estimation of computing time
For both molecular computation and our approach, the computing time was determined by the moving speed of computation carrier (i.e., molecules and photons) and the longest path in the network. For example, in the case of \( \{2, 5, 7, 9\} \) shown in Fig. 1, the longest path is the one linking to port 23, which represents the sum of all the elements in the set. According to the geometrical parameters and the scalable rules of our waveguide network, it is easy to calculate the length of the longest path. The propagating speed of photons is estimated on the basis of the refractive index of Corning Eagle XG (50) and the refractive index change induced by femtosecond laser writing (51). The structural parameters of molecular computation were derived from the experiment by Nicolau et al. (16). The faster molecules, actin filaments, were chosen to compare with our approach.

The running time taken by conventional electronic computers working in a brute-force mode, searching the entire solution space consisting of all possible subsets, to solve the SSP was estimated by multiplying FLOPS by the total number of arithmetic operations. The data of FLOPS adopted in our research are either the peak performance or theoretical performance of the corresponding electronic machine. Performance degradation (46) in practical scenario was neglected.

Speed-up of the photonic computer
Given a set of \( N \) elements, the number of subsets grows exponentially with \( N \). According to the definition of the SSP, it requires us to verify every possible subset. If we regard the verification of a subset as a subtask, the number of subtasks or the number of computation operations increases at an exponential rate.

For a conventional electronic computer working sequentially, all subtasks are executed in sequence. Therefore, the total computing time is equivalent to the product of the number of computation operations and the unit time taken by a single operation, growing at an exponential rate. For our photonic computer, which works in a parallel mode, all subtasks can be executed simultaneously. In our implementations, each subset is mapped to a path of the photonic circuits. With light beam (a stream of photons) being split and propagating along all possible paths, all subsets were verified at the same time. On such an occasion, the total computing time only depends on the verification of the largest subset.

Here, the verification of the largest subset corresponds to the movement of photons from the input port to the output port through the longest path. As a result, the computing time is equal to the traveling time of photons in the longest path, growing at a sub-exponential rate, which is slower than that of electronic computers. Moreover, as photons have an ultrahigh propagating speed and the integrated...
photic circuit has a compact structure, the computing process is further sped up.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/5/eaay5853/DC1

The evolution of computers and the role of non–Von Neumann architecture

The influence of optical source on time consumption

Time-consuming performance

Signal-to-noise ratio

Fisher information

The role of interference

Fig. S1. The role of non–Von Neumann architecture.

Fig. S2. Time-consuming performance.

Fig. S3. Signal-to-noise ratio.

**REFERENCES AND NOTES**


Acknowledgments

Funding: The authors thank Jian-Wei Pan for helpful discussions. This research was supported by the National Key R&D Program of China (2019YFA0308700, 2017YFA0303700), the National Natural Science Foundation of China (61734005, 11761141014, 11690033, and 11774222), the Science and Technology Commission of Shanghai Municipality (STCSM) (17JC1400403), the Shanghai Municipal Education Commission (SMEC) (2017-01-07-00-02- E00049), and the Program for Professor of Special Appointment at Shanghai Institutions of Higher Learning (Grant GZ2016004). X.-M.J. acknowledges additional support from a Shanghai talent program. Author contributions: X.-M.J. and H.P.Z. conceived the project. X.-M.J. supervised the project. X.-Y.X. and X.-M.J. designed the experiment. X.-Y.X. fabricated the photonic chips. X.-Y.X., X.-L.H., Z.-M.L., J.G., Z.-Q.J., Y.W., R.-J.R., and X.-M.J. performed the experiment and analyzed the data. X.-Y.X. and X.-M.J. wrote the paper with input from all the other authors. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 1 July 2019
Accepted 21 November 2019
Published 31 January 2020
10.1126/sciadv.aay5853

A scalable photonic computer solving the subset sum problem
Xiao-Yun Xu, Xuan-Lun Huang, Zhan-Ming Li, Jun Gao, Zhi-Qiang Jiao, Yao Wang, Ruo-Jing Ren, H. P. Zhang and Xian-Min Jin

Sci Adv 6 (5), eaay5853.
DOI: 10.1126/sciadv.aay5853

ARTICLE TOOLS  http://advances.sciencemag.org/content/6/5/eaay5853

SUPPLEMENTARY MATERIALS  http://advances.sciencemag.org/content/suppl/2020/01/27/6.5.eaay5853.DC1

REFERENCES  This article cites 39 articles, 5 of which you can access for free  http://advances.sciencemag.org/content/6/5/eaay5853#BIBL

PERMISSIONS  http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service