Rapid Communications

Experimentally detecting a quantum change point via the Bayesian inference

Shang Yu,[1,2] Chang-Jiang Huang,[1,2] Jian-Shun Tang,[1,2,*] Zhih-Ahn Jia,[1,2] Yi-Tao Wang,[1,2] Zhi-Jin Ke,[1,2] Wei Liu,[1,2] Xiao Liu,[1,2] Zong-Quan Zhou,[1,2] Ze-Di Cheng,[1,2] Jin-Shi Xu,[1,2] Yu-Chun Wu,[1,2] Yuan-Yuan Zhao,[1,2] Guo-Yong Xiang,[1,2,†] Chuan-Feng Li,[1,2,‡] Guang-Can Guo,[1,2] Gael Sentís,[3,§] and Ramon Muñoz-Tapia[4,‖]

[1]*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui 230026, China*

[2]*Synergetic Innovation Center of Quantum Information & Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[3]*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, 57068 Siegen, Germany*

[4]*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*

(Received 7 December 2017; revised manuscript received 1 April 2018; published 15 October 2018)

Detecting a change point is a crucial task in statistics that has been recently extended to the quantum realm. A source state generator that emits a series of single photons in a default state suffers an alteration at some point and starts to emit photons in a mutated state. The problem consists in identifying the point where the change took place. In this Rapid Communication, we build a pseudo-on-demand single-photon source to prepare the photon sequences, and consider a learning agent that applies Bayesian inference on experimental data to solve this problem. This learning machine adjusts the measurement over each photon according to the past experimental results and finds the change position in an online fashion. Our results show that the local-detection success probability can be largely improved by using such a machine-learning technique. This protocol provides a tool for improvement in many applications where a sequence of identical quantum states is required.

DOI: 10.1103/PhysRevA.98.040301

Introduction. The change point problem is a crucial concept in statistics [1–3] that has been studied in many real-world situations, from stock markets [4] to protein folding [5]. One of the key goals in this field of research is to devise procedures that detect the exact point where a sudden change has occurred. This point could indicate, for example, the trigger of a financial crisis or a misfolded protein step [6]. Many settings can be formulated as a change point problem, however, we can understand it in the simplest terms as a heads or tails game. Alice sends a series of fair coins to Bob who can toss each of them. She then starts sending another type of coin and Bob's task is to recognize from his observations the most likely point where the switching occurred.

In recent works, this problem has been extended to the quantum realm [7–9]. Considering a quantum state generator that is supposed to emit a sequence of photons in a default state but suffers an uncontrolled alteration at some unknown point (e.g., a rotation of the polarization), the problem then is to determine where the change took place from measurements on the photons. A comparison of several methods which were proposed in Ref. [8] is shown in Fig. 1. The most general procedure consists in waiting until all the photons have reached the detector and measuring them at the very end [6,8]. An optimal global measurement naturally provides the best identification performance, but it is difficult to realize experimentally as it requires a quantum memory to store photons as well as collective quantum operations. On the other hand, it is much more feasible to measure the state of each photon as soon as it arrives (local measurement). The most basic local procedure consists of fixed measurements that unambiguously detect a mutated state with some probability, whereas the performance of this simple procedure is far below the optimal one. The Bayesian inference procedure, however, that adapts the local measurements according to the previous outcome, can greatly improve the success probability. Such a protocol can be viewed as a machine-learning (ML) mechanism, while many recent works have proved that a learning machine (either classical [8,10–15] or quantum [16–19]) can provide an efficient route to quantum characterization, verification, and validation. We have already benefited from using conventional machine learning to solve many quantum problems [10–15,20,21]. In particular, Bayesian inference as a classic technique in the ML domain has been discussed in Refs. [11,22–24] and applied to quantum tasks in many cases, such as quantum Hamiltonian learning [11], phase estimation [23–27], and several others [28,29] (spectral function reconstruction and Josephson oscillation characterization).

In this Rapid Communication, we experimentally detect a quantum change point in one sequence of photons using the Bayesian inference (BI) and basic local (BL) strategies, which are proposed theoretically in Ref. [8]. We compare the success probabilities between these two methods and with respect to the (theoretical) optimal global strategy. For the BI strategy, we build a learning agent (a programed computer) to guess the change point position. Once the photon arrives, we detect

*tjs@ustc.edu.cn
†gyxiang@ustc.edu.cn
‡cfli@ustc.edu.cn
§gael.sentis@uni-siegen.de
‖Ramon.Munoz@uab.cat

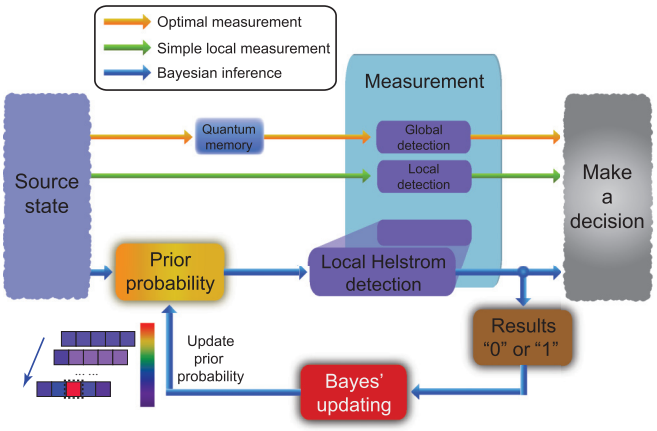

FIG. 1. Schematic diagram of the three different detection protocols. The optimal measurement is constituted by four steps: the source state preparation, a quantum memory for storing each photon, global detection, and decision making. For the simple local strategy, such as basic local (BL) detection, the quantum memory model is removed and global detection is replaced by a local detection. The measurement base is fixed in the BL method, but is changed in the Bayesian inference (BI) strategy. The Helstrom measurement in this method is decided by the prior probability obtained from the last step (see Appendix A). Meanwhile, the priors are also updated according to the measurement results using the Bayesian update rule. The change point is determined by choosing the position with the largest prior probability after the last step.

the result and update the priors for each hypothesis. Then the agent decides which measurement basis is used for the next photon. In other words, the detection basis is being learned as the protocol proceeds and a refined guess is provided at each step. The output guess is given after the last measurement. In contrast, the measurement basis of the BL strategy is fixed during the whole experiment, and the guessed change position is determined by the first conclusive detection of a mutated state (see Appendix A). In order to compare the performances of these methods, we conduct a total of 1000 experiments for a certain overlap between the default and mutated state (50 times for one possible change point, and there are 20 possible change points in our situation). Our results show that the learning agent provides a significant advantage over the BL detection and a performance very close to the optimal (global) one.

Brief introductions for the theoretical framework. First, we briefly review the BI strategy according to the theoretical framework proposed in Ref. [8]. Consider an on-demand single-photon sequence which contains n photons, a quantum change occurs at the kth position, and the corresponding source state can be expressed as

$$|\Psi_k\rangle = |H\rangle^{\otimes k-1}|\phi\rangle^{\otimes n-k+1}, \tag{1}$$

where $|H\rangle$ is the default state and $|\phi\rangle = c|H\rangle + s|V\rangle$ denotes the mutated state (without loss of generality, we set c to be real and positive, and both the default as well as mutated states are known to the detecting agent). At the sth step, the BI approach contains the following processes [8]: First, the learning agent calculates the measurement basis $\Pi_0^{(s)}$ according to the prior probability $\eta_k^{(s)} = p(k|r_1, r_2, \ldots, r_{s-1})$ [31], which represents the estimation of how likely the change point occurs at position k under the sth learning step; second, the agent performs a Helstrom measurement [32] and obtains the experimental result ($r_s = 0$ or 1); third, depending on the outcome, the priors are renewed according to the Bayes updating rule $\eta_k^{(s+1)} = p(r_s|k)\eta_k^{(s)}/\Sigma_{l=1}^{n} p(r_s|l)\eta_l^{(s)}$, where $p(r_s|k)$ denotes the condition probability which means how probable it is to have the observed result r_s given a hypothetical change point at position k. The procedure is repeated until the last photon is measured and the decision is given by the hypothesis with the highest updated prior. The detailed calculations are shown in Appendix A.

Experimental setup. The experimental setup is sketched in Fig. 2(a), which can be recognized as two parts: a source state generator, which is controlled by Alice, and a classical learning agent, which is owned by Bob. According to the above discussion, an on-demand single-photon source is required for the source state preparation. Here, we propose a technique to create the pseudo-on-demand single-photon source, which is helped with a chopper and postselection method, and the preparation requirements are met. First, the heralded single photon is generated by a type-II spontaneous parametric down-conversion (SPDC) process in beta-barium-borate (BBO) crystal. Then, the chopper driven by an arbitrary wave-form generator (AWG) is applied to separate the single-photon sequence into n equal time bins [see the second line in Fig. 2(b)], which determines the number of particles (n) in the source state and provides a time label for Alice to create the change point accurately. After that, the logic unit postselects the first effective event in each bin to be the detectable photons in the source state [30]. These procedures ensure that there must be one (and only one) photon in each bin that can be detected. This method provides a way to create a pseudo-on-demand single-photon source which can also be applied in other similar experiments when the genuine on-demand single-photon source technology is difficult to implement and not very mature.

While Alice begins to generate the source state, AWG will send a trigger signal [at 100-ms intervals, the first line of the time sequence diagram, as shown in Fig. 2(b)] to Bob and inform him the detection task should begin. The prepared pseudo-on-demand single photons then pass through a polarization beam splitter (PBS) to obtain the default states $|H\rangle$. An electro-optic modulator (EOM), in conjunction with half-wave plates (HWPs) and a quarter-wave plate (QWP), are used to change the default state $|H\rangle$ into a mutated state $|\phi\rangle$ at some certain point [see the third and fourth lines in Fig. 2(b)]. During the experiment, Alice can decide at which point (k) the mutation occurs by controlling the relative time delay of the signal's output from the AWG, and without loss of generality, we set $n = 20$ in all experiments (the corresponding chopper period is 5 ms). After the above procedures, the source state $|\Psi_k\rangle$ can be generated.

Once the trigger signal is received, Bob starts his measurement. At the first step ($s = 1$), he tunes the electronic-controlled HWP to the basis $\Pi_0^{(1)}$ which is calculated from the previous information on the computer. Then, a two-outcome single-shot measurement is performed on the first photon and

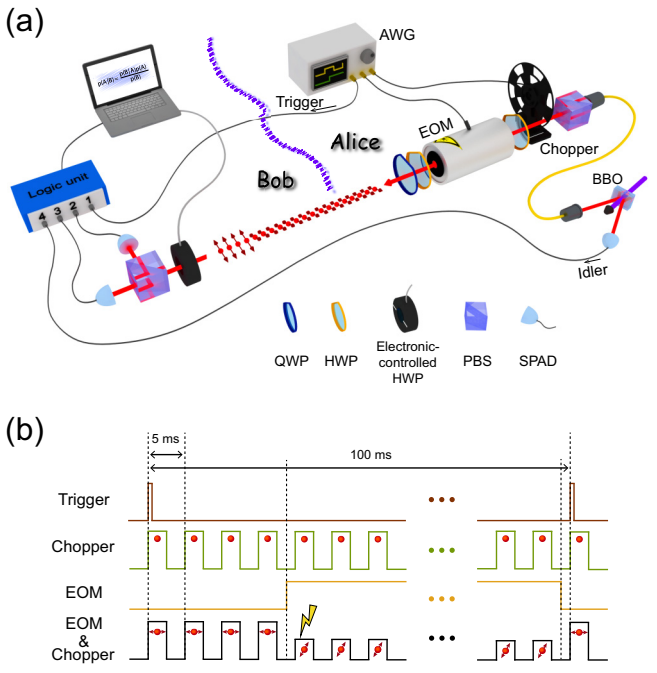

FIG. 2. (a) Experimental setup. This setup consists of two main parts: the source state preparation part at the Alice side and the detecting agent part controlled by Bob. Preparation part: The pseudo-on-demand single photon is generated by the SPDC process in BBO crystal associated with the postselection method and a chopper, which divides the photon sequence into n equal discrete time bins [30]. The photons then pass through a PBS to produce the default states $|H\rangle$ and an EOM after that creating the change point at some certain place in the sequence. Detection part: The measurement device consists of an electronic-controlled HWP, a PBS, and a classical learning agent (personal computer). The measurement basis realized by the PBS and the electronic-controlled HWP can respond rapidly when the learning agent calculates the new basis from the latest priors (see Appendix A). (b) The time sequence diagram. The first line is the trigger signal (generated by AWG), which has a 100-ms interval. The second line denotes the time bins created by the chopper. Each time bin has a 2.5-ms width and the red balls represent signal photons postselected by the logic unit [30]. The third and fourth lines are the EOM signal and the folding signal by EOM and chopper, respectively. In the picture, we show a case where the change point is set at the fifth photon, and the arrows on the red balls denote the photon polarization.

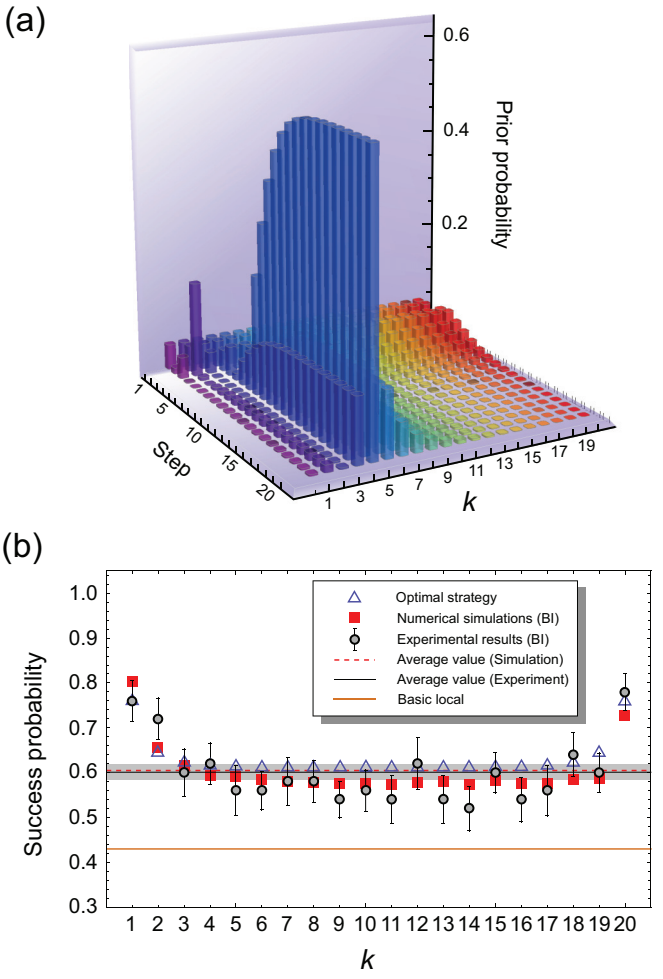

FIG. 3. (a) The results of prior probabilities at each step when $c^2 = 0.604$ and $k = 5$. From the first step to the last (21st) step, the priors fluctuate depend on the measurement results and begin to converge at the sixth step. At the last step, the highest column is located at $k = 5$, which corresponds to a prior probability of 0.569. This result corresponds to the correct change point in this experiment. (b) The relationship between success probability and change point position. Here, Alice sets $c^2 = 0.604$, but varies each change position. The purple triangles correspond to the success probabilities for the optimal (global) strategy [8], the red squares are the numerical simulation results for the BI strategy, and the circles are the actual experimental results. The simulation average value of the success probabilities are denoted by the dashed red line. The black solid line and the gray region are the average values and associated errors obtained from experiments, respectively. It is clear that the probabilities derived from the BI protocol are all beyond the BL one, denoted as the solid orange line.

detected by two single-photon avalanche diodes (SPADs). After that, the result ("0" or "1") [33] is sent to the computer to calculate the priors $\eta_k^{(2)}$ according to the Bayes' updating rules (the details are shown in Appendix A). A new measurement basis then can be determined for the next step. Until the last measurement is finished, Bob produces a guess $\hat{k}$ that maximizes $\eta_k^{(21)}$ for the change point. It worth noting that there is only one sequence of photons prepared by Alice in each experiment. Therefore, Bob is able to measure only one photon in each learning step but does not have any other copies of the photons that can detected. Namely, at the sth step in one learning process, once the sth photon is detected by Bob, he should respond to this result (completes the learning calculations and resets the measurement basis) before the next photon comes, i.e., less than the time interval of each bin (5 ms).

Results and discussion. An example of a single learning process is shown in Fig. 3(a). We choose a random value of the overlap, $c^2 = 0.604$, and a change point at position $k = 5$. The height of the columns in each row represents the priors for every hypothesis $k = 1, \ldots, 20$ after each measurement step. The initial priors are set to be uniform since Bob does not have prior knowledge about the position of the change

point. As the measurements proceed, the Bayesian inference method is able to learn the right position and to correct previous wrong guesses or mistakes caused by experimental noise [34]: We can find that the prior distribution begins to converge on the correct point after the sixth step, although an incorrect value occurred at the third step (the highest column occurs at $k = 2$, which is not the correct change point position). As can be readily seen in Fig. 3(a), the highest updated prior probability at the end of the process $\hat{k} = \arg\max_k\{\eta_k^{(21)}\}$ is precisely $\hat{k} = 5$.

Next, we analyze the detection performance for each position of the change point. We repeat the above experiment 50 times to gather statistics and compute the success probabilities for each source state $|\Psi_k\rangle$, $k = 1, 2, \ldots, 20$ (we fix the overlap to be the same as before, i.e., $c^2 = 0.604$). The conditional success probabilities as a function of k are shown in Fig. 3(b). The red squares represent the numerical results of a Monte Carlo simulation of the experiment, and the actual experimental data are shown as circles. Notice the good agreement between both values that coincides within the experimental error bars. We find that most change positions ($k = 3$–19) have constant success probabilities of approximately 0.58 without large fluctuations. We also observe that the first two and the last positions can be better detected than the rest (with a success probability larger than 0.7). This is an expected result, as change points occurring toward the beginning or end of the sequence are more distinguishable from their neighbors than those occurring at the bulk of the sequence. For comparison, we also show the optimal (global) conditional probabilities with purple triangles. One can also appreciate from Fig. 3(b) that there is a small but systematic difference between the local BI and the global protocol for most values of k. However, at the end points, this difference disappears and the BI protocol performs almost optimally [35]. Furthermore, the success probabilities produced by the classical learning agent all surpass the BL method (the solid orange line) including the errors. This advantage remains even for large n. More numerical simulation results are given in the Supplemental Material [30]. Here, the error bars are obtained by 100 Monte Carlo simulations (we use the same number of simulations below).

Finally, to assess the overall performance of each protocol, we conduct the aforementioned experiments and compute the corresponding average success probability for all possible change points and for each value of the overlap c^2, that we space in intervals of $\Delta c^2 = 0.05$. We show our results in Fig. 4. The red dots and orange triangles are the experimental results for the BI and BL strategies, respectively. The black dashed line represents the optimal measurement, and the purple solid line indicates the theoretical BL detection limit. It is clear from the results that the improvement over the BL strategy provided by a machine-learning-enhanced local procedure almost closes the gap with the optimal measurement (which is extremely hard to implement, especially for large n).

Although the theoretical optimal detection strategy establishes an ultimate performance bound for the task of identifying quantum change points, it had yet to be seen how close a realistic experimental implementation could get. We have demonstrated that BI is an easily implementable local strategy, within the reach of current experimental techniques, that performs quasioptimally. The identification of quantum change

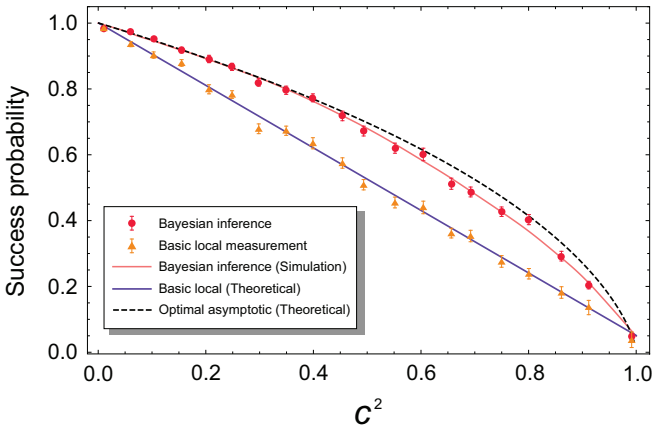

FIG. 4. Comparison of the BI, BL, and optimal measurement for every overlap. The red dots are the experimental results obtained using BI detection, and the solid light red line is the corresponding simulation value. The orange triangles are the results obtained from BL detection, and the solid purple line is its theoretical value. The success probability for the optimal measurement is shown as the dashed black line. Although the BI approach cannot surpass the optimal (global) detection, it nearly covers the gap between the BL and the optimal measurement.

points demonstrated in our experiment might be applicable in many practical situations, such as identifying domain walls in condensed-matter systems [36] and detecting the changes of fluorescence polarization in some biological processes [37]. Our work can also be extended to high-dimensional quantum states, mixed states, and multiple change points. In addition, while our experimental setup uses a classical learning agent, the BI approach is also implementable with quantum learning algorithms [38]. This will be studied in the future.

Conclusion. In summary, we report a demonstration of the quantum change point detection under the pseudo-on-demand single-photon sequences and based on the Bayesian inference strategy, which yields a higher success probability compared to other local methods. The BI approach contains a learning agent that can adjust the measurement basis at each step using the Bayes rule. This agent gives a very significant advantage that yields a performance very close to optimality and also provides the capacity to correct mistakes caused by experimental noise. The ability to accurately detect a quantum change point has an immediate impact on many quantum information tasks. Our learning method demonstrated that the protocol can be efficiently implemented to identify the position of a sudden change in the quantum settings where a sequence of identical quantum states is required.

Acknowledgments. This work is supported by the National Key Research and Development Program of China (No. 2016YFA0302700), the National Natural Science Foundation of China (Grants No. 11674304, No. 61327901, No. 61490711, No. 11774334, No. 11774335, No. 11474267, No. 11325419, No. 11574291, and No. 91321313), the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (Grant No. QYZDY-SSW-SLH003), Anhui Initiative in Quantum Information Technologies (AHY020100, AHY060300), the Youth Innovation Promotion Association of Chinese Academy of Sciences (Grant

No. 2017492), the Foundation for Scientific Instrument and Equipment Development of Chinese Academy of Sciences (No. YJKYYQ20170032), and the Fundamental Research Funds for the Central Universities (No. WK2470000026). R.M.T. acknowledges the Spanish MINECO through Contracts No. FIS2013-40627-P and No. FIS2016-80681-P. G.S. acknowledges financial support from the ERC Consolidator Grant No. 683107/TempoQ, and the DFG.

S.Y. and C.-J.H. contributed equally to this work.

APPENDIX A: THEORY AND CALCULATIONS

The source state prepared by Alice can be expressed as $|\Psi_k\rangle = |H\rangle^{\otimes k-1}|\phi\rangle^{\otimes n-k+1}$, where the k denotes the position of the change point and n is the number of the photons in the source state. Here, $|\phi\rangle = c|H\rangle + s|V\rangle$ ($s = \sqrt{1-c^2}$, and H, V represent the horizontal and vertical polarization, respectively). Without loss of generality, we choose c to be real and positive.

Basic local. For a chain of photons for which the polarization state begins with $|H\rangle$ and changes into the state $|\phi\rangle$ at an unknown point k, the simplest online strategy (namely, basic local) is to measure each photon in the basis $\{|H\rangle, |V\rangle\}$. The measurements are performed sequentially until the outcome $|\phi\rangle$ (i.e., get result "1") is obtained for the first time at the rth step. We are sure that the rth particle was in the state $|\phi\rangle$, which means that the change must have occurred at some position $k \leqslant r$. Then, our best guess for the change point is $k = r$. The success probability is $p = 1 - c^2 + \frac{c^2}{n}$.

We want to note here that the measurement basis need not be $\{|H\rangle, |V\rangle\}$. We also can set it as others, such as the Helstrom basis $\{\Pi_0, \Pi_1\}$ [32], which is given by the projectors onto the positive and negative parts of the spectrum of the Helstrom matrix $\Gamma = |H\rangle\langle H| - |\phi\rangle\langle\phi|$. However, it can be proven that the success probability of all the fixed-basis methods cannot surpass the learning-enhanced local strategy [39].

Bayesian inference. This online strategy can improve the success probability by adjusting the measurement basis according to the experimental result in each step. We build a classical learning agent to guess where the change point occurs. It starts with a uniform prior $p(k) = \frac{1}{n}$ about the hypothesis of the change point and updates the expectation as new data are obtained. In order to update the information at the sth step, the learning algorithm is designed as follows [8]:

(1) Find the probability $p_H^{(s)}$ (or $p_\phi^{(s)}$) of the most likely sequence that has the particle at position s being in the state $|H\rangle$ (or $|\phi\rangle$), which in general depends on all previous results $r_1, r_2, \ldots, r_{s-1}$ through the priors $\eta_k^{(s)} \equiv p(k|r_1, r_2, \ldots, r_{s-1})$:

$$p_H^{(s)} = \max_k \left\{\eta_k^{(s)}\right\}_{k=s+1}^{n},$$

$$p_\phi^{(s)} = \max_k \left\{\eta_k^{(s)}\right\}_{k=1}^{s}.$$

(2) Perform the Helstrom measurement [32] on the sth photon. The measurement basis $\{\Pi_0^{(s)}, \Pi_1^{(s)}\}$ is given by the projectors onto the positive and negative parts of the spectrum of the Helstrom matrix $\Gamma^{(s)} = p_H^{(s)}|H\rangle\langle H| - p_\phi^{(s)}|\phi\rangle\langle\phi|$

(3) After the sth measurement has been performed, the prior is updated in accordance with the measurement result, using Bayes' update rule,

$$\eta_k^{(s+1)} = \frac{p(r_s|k)\eta_k^{(s)}}{\sum_{l=1}^{n} p(r_s|l)\eta_l^{(s)}}.$$

(4) Go back to the first procedure for the next measurement.

After the last measurement, the agent updates the prior to $\eta_k^{(n+1)}$ and produces the guess $\hat{k}$ that maximizes $\eta_k^{(n+1)}$ for the change point [$\hat{k} = \mathrm{argmax}_k(\eta_k^{(n+1)})$]. If the guess is correct ($\hat{k} = k$), mark this experiment as successful. Otherwise, mark it as unsuccessful.

APPENDIX B: EXPERIMENTAL ERROR ANALYSIS

In our experiment, the experimental error is caused by several aspects, such as the background noise, the systematic error, and the statistical error. In Fig. 4, the success probabilities at $c^2 = 0.010$ for two strategies are $P_{\mathrm{BI}} = 0.984 \pm 0.002$ and $P_{\mathrm{BL}} = 0.982 \pm 0.004$, respectively.

Because of the imperfect experimental apparatuses, such as the EOM and the PBS, we cannot obtain an extremely high extinction ratio. In other words, we cannot create the ideal orthogonal state ($c^2 = 0$). In this case ($c^2 = 0.010$), the default state and the mutated state are nearly orthogonal, and the statistical error is minimized since all the experimental results in this case should give the correct guessing point. Therefore, the distances between the experimental data and the corresponding theoretical values ($P_{\mathrm{BI}}^{\mathrm{th}} = 0.995$ and $P_{\mathrm{BL}}^{\mathrm{th}} = 0.991$) are mainly caused by the background noise and the systematic error. The fluctuation of success probabilities that appears at other places is mainly caused by the statistical error. Since we run 50 experiments for each change point at one overlap, the statistical probability obtained by 50 experiments [Fig. 3(b)] and 1000 experiments (Fig. 4) will have a statistical fluctuation. We can find a relatively large fluctuation (a relatively large error bar) in Fig. 3(b) because there are just 50 experiments for each change point. However, in Fig. 4, the fluctuation (or error bars) becomes much smaller since they are calculated from a total of 1000 experiments (50 times for each change point and there are 20 possible change positions for each overlap). Although the number of experiments for each source state is not very large, the experimental results for different k (or c^2) matches the numerical results and shows the characteristic of the relationship between success probability and the change position (or overlap).

[1] M. Pollak, Ann. Stat. **13**, 206 (1985).

[2] M. Baseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application* (Prentice Hall, Upper Saddle River, NJ, 1993).

[3] J. Chen and A. K. Gupta, *Parametric Statistical Change Point Analysis*, 2nd ed. (Birkhäuser, Boston, 2012).

[4] J. Chen and A. K. Gupta, J. Am. Stat. Assoc. **92**, 438 (1997).

[5] M. Pirchi *et al.*, Nat. Commun. **2**, 493 (2011).

[6] L. Zyga, Taking statistics to the quantum domain (2016), https://phys.org/news/2016-11-statistics-quantum-domain.html.

[7] D. Akimoto and M. Hayashi, Phys. Rev. A **83**, 052328 (2011).

[8] G. Sentís, E. Bagan, J. Calsamiglia, G. Chiribella, and R. Muñoz-Tapia, Phys. Rev. Lett. **117**, 150502 (2016).

[9] G. Sentís, J. Calsamiglia, and R. Muñoz-Tapia, Phys. Rev. Lett. **119**, 140506 (2017).

[10] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[11] J. Wang *et al.*, Nat. Phys. **13**, 551 (2017).

[12] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

[13] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).

[14] S. Mavadia, V. Frey, J. Sastrawan, S. Dona, and M. J. Biercuk, Nat. Commun. **8**, 14106 (2017).

[15] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).

[16] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).

[17] N. Wiebe, D. Braun, and S. Lloyd, Phys. Rev. Lett. **109**, 050505 (2012).

[18] S. Lloyd, M. Mohseni, and P. Rebentrost, Nat. Phys. **10**, 631 (2014).

[19] P. Rebentrost, M. Mohseni, and S. Lloyd, Phys. Rev. Lett. **113**, 130503 (2014).

[20] K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. S. Kim, npj Quantum Inf. **3**, 36 (2017).

[21] D. Lu *et al.*, npj Quantum Inf. **3**, 45 (2017).

[22] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature (London) **549**, 195 (2017).

[23] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).

[24] N. Wiebe and C. Granade, Phys. Rev. Lett. **117**, 010503 (2016).

[25] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Nat. Photonics **5**, 43 (2011).

[26] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[27] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature (London) **450**, 393 (2007).

[28] Y. Burnier and A. Rothkopf, Phys. Rev. Lett. **111**, 182003 (2013).

[29] J. Javanainen and R. Rajapakse, Phys. Rev. A **92**, 023613 (2015).

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.98.040301 for the logic unit postselection, more numerical simulations details, and the analysis of advantage of Bayesian inference approach.

[31] $\Pi_0^{(s)} = |\lambda_0^{(s)}\rangle\langle\lambda_0^{(s)}|$ ($\Pi_1^{(s)} = |\lambda_1^{(s)}\rangle\langle\lambda_1^{(s)}|$) is calculated from the Helstrom matrix $\Gamma^{(s)} = p_H^{(s)}|H\rangle\langle H| - p_\phi^{(s)}|\phi\rangle\langle\phi|$, where $\lambda_0^{(s)}$ and $\lambda_1^{(s)}$ are the positive and negative parts of the spectrum of $\Gamma^{(s)}$. Here, $p_H^{(s)} = \max_k \{\eta_k^{(s)}\}_{k=s+1}^{n}$, and $p_\phi^{(s)} = \max_k \{\eta_k^{(s)}\}_{k=1}^{s}$. In the first step ($s = 1$), the prior probabilities are uniform, $\eta_k^{(1)} = 1/n$, so there is $p_H^{(s)} = p_\phi^{(s)}$ at this step. More details can be found in the Supplemental Material [30].

[32] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[33] Here, "0" represents the default state is detected and "1" means the mutated state is detected.

[34] This error correction ability will be shown when $c > 0$, i.e., under the *quantum* situation.

[35] The success probabilities when $k = 4, 12, 18$ also reach the optimal measurement values, but we conclude it is caused by the experimental errors [30], which means they are likely to fluctuate below the optimal value in the next experiment.

[36] W. Wang, F. Yang, C. Gao, J. Jia, G. D. Gu, and W. Wu, APL Mater. **3**, 083301 (2015).

[37] M. D Hall, A. Yasgar, T. Peryea, J. C Braisted, A. Jadhav, A. Simeonov, and N. P. Coussens, Methods Appl. Fluoresc. **4**, 022001 (2016).

[38] G. H. Low, T. J. Yoder, and I. L. Chuang, Phys. Rev. A **89**, 062315 (2014).

[39] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia (unpublished).