Towards the standardization of quantum state verification using optimal strategies

Xinhe Jiang,1,* Kun Wang,2,* Kaiyi Qian,1,* Zhaozhong Chen,1,* Zhiyu Chen,1
Liangliang Lu,1 Lijun Xia,1 Fangmin Song,1 Shining Zhu,1 and Xiaosong Ma1,†

1National Laboratory of Solid-state Microstructures, School of Physics,
Collaborative Innovation Center of Advanced Microstructures,
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology,
Nanjing University, Nanjing 210093, China
2Shenzhen Institute for Quantum Science and Engineering, Southern
University of Science and Technology, Shenzhen 518055, China
(Dated: February 4, 2020)

Quantum devices for generating entangled states have been extensively studied and widely used. As so, it becomes necessary to verify that these devices truly work reliably and efficiently as they are specified. Here, we experimentally realize the recently proposed two-qubit entangled state verification strategies using both local measurements (nonadaptive) and active feed-forward operations (adaptive) with a photonic platform. About 1650/268 number of copies (N) are required to achieve a 90% confidence to verify the target quantum state for nonadaptive/adaptive strategies. These optimal strategies provide the Heisenberg scaling of the infidelity $\epsilon$ as a function of $N$ ($\epsilon\sim N^r$) with the parameter $r = -1$, exceeding the standard quantum limit with $r = -0.5$. We experimentally obtain the scaling parameter of $r = -0.88\pm0.03$ and $-0.78\pm0.07$ for nonadaptive and adaptive strategies, respectively. Our experimental work could serve as a standardized procedure for the verification of quantum states.

* These authors contributed equally to this work.
† xiaosong.ma@nju.edu.cn
Quantum state plays an important role in quantum information processing. Quantum devices for creating quantum states are building blocks for quantum technology. Being able to verify these quantum states reliably and efficiently is an essential step towards practical applications of quantum devices. Typically, a quantum device is designed to output some desired state $\rho$, but the imperfection in the device’s construction and noise in the operations may result in the actual output being deviate from it to some random and unknown states $\sigma_i$. A standard way to distinguish these two cases is quantum state tomography. However, this method is both time-consuming and computationally challenging. Non-tomographic approaches have also been proposed to accomplish the task, yet these methods make some assumptions either on the quantum states or on the available operations. It is then natural to ask whether there exists an efficient non-tomographic approach to accomplish the task?

The answer is affirmative. In a recent work, Pallister et al. proposed an optimal strategy to verify non-maximally entangled two-qubit pure states under locally projective and nonadaptive measurements. The locality constraint induces only a constant-factor penalty over the nonlocal strategies. Since then, various works have been done along this line of research, targeting on different states and measurements. We also remark related works by Dimić et al. and Saggio et al., in which they developed a generic protocol for efficient entanglement detection using local measurements and detect entanglement with an exponentially growing confidence using only few copies of the quantum state.

In this work, we report an experimental two-qubit state verification procedure using both optimal nonadaptive (local measurements) and adaptive (active feed-forward operations) strategies with an optical setup. Compared with previous works merely on minimizing the number of measurement settings, we also minimize the number of copies (i.e. coincidence counts in our experiment) required to verify the quantum state generated by the quantum device. We perform two tasks – Task A and Task B. With Task A, we obtain a fitting infidelity and the number of copies required to verify the quantum state generated by the quantum device. We also minimize the number of copies versus the number of copies $N$ and the adaptive strategy to the two-qubit states. With our methods, we obtain a comprehensive judgement about the quantum state generated by a quantum device. Present experimental and data analysis workflow may be regarded as a standard procedure for quantum state verification.

**Results**

**Quantum state verification.** Consider a quantum device $D$ designed to produce the two-qubit pure state

$$|\Psi\rangle = \sin \theta |HH\rangle + \cos \theta |VV\rangle,$$

where $\theta \in [0, \pi/4]$. However, it might work incorrectly and actually outputs two-qubit fake states $\sigma_1, \sigma_2, \cdots, \sigma_N$ in $N$ runs. The goal of the verifier is to determine the fidelity threshold of these fake states to the target state with certain confidence. To realize the verification of our quantum device, we perform the following two tasks in our experiment (see Fig. 1):

**Task A:** Performing measurements according to verification strategy on the fake states copy by copy, and make statistics on the number of copies required before we find the first fail event. The concept of Task A is shown in Fig. 1b.

**Task B:** Performing a fixed number ($N$) of measurements according to verification strategy, and make statistics on the number of copies that pass the verification tests. The concept of Task B is shown in Fig. 1c.

**Task A** is based on the assumption that there exists some $\epsilon > 0$ for which the fidelity $\langle \Psi|\sigma_i|\Psi \rangle$ is either 1 or satisfies $\langle \Psi|\sigma_i|\Psi \rangle \leq 1 - \epsilon$ for all $i \in \{1, \cdots, N\}$ (see Fig. 1b). Our task is to determine which is the case for the quantum device. To achieve Task A, we perform binary-outcome measurements from a set of available projectors to test the state. Each binary-outcome measurement $\{M_l, 1 - M_l\}$ is specified by an operator $M_l$, corresponding to passing the test. For simplicity, we use $M_l$ to denote the corresponding binary measurement. This measurement is performed with probability $p_l$. We require the target state $|\Psi\rangle$ always passes the test, i.e. $M_l|\Psi\rangle = |\Psi\rangle$. In the bad case ($\langle \Psi|\sigma_i|\Psi \rangle \leq 1 - \epsilon$), the maximal probability that $\sigma_i$ can pass the test is given by,

$$\max_{\langle \Psi|\sigma_i|\Psi \rangle \leq 1 - \epsilon} \text{Tr}(\Omega \sigma_i) = 1 - [1 - \lambda_2(\Omega)] \epsilon := 1 - \Delta,$$

where $\Omega = \sum_l p_l M_l$ is called an strategy, $\Delta$ is the probability $\sigma_i$ fails a test and $\lambda_2(\Omega)$ is the second largest eigenvalue of $\Omega$. Whenever $\sigma_i$ fails the test, we know immediately that the device works incorrectly. After $N$ runs, $\sigma_i$ in the incorrect case can pass all these tests with probability being at most $[1 - (1 - \lambda_2(\Omega)) \epsilon]^{N}$. Hence to achieve confidence $1 - \delta$, it suffices to conduct $N$ number of measurements satisfying

$$N \geq \frac{\ln \delta}{\ln [1 - (1 - \lambda_2(\Omega)) \epsilon]} \approx \frac{1}{[1 - \lambda_2(\Omega)] \epsilon} \ln \frac{1}{\delta}.$$
From Eq. (3) we can see that an optimal strategy is obtained by minimizing the second largest eigenvalue $\lambda_2(\Omega)$, with respect to the set of available measurements. Pallister, Linden, and Montanaro\textsuperscript{28} proposed an optimal strategy for Task A, using only locally projective measurements. Since no classical communication is involved, this strategy (hereafter labelled as $\Omega_{\text{opt}}$) is non-adaptive. Wang and Hayashi\textsuperscript{21} later gave the optimal adaptive strategies using local operations and either one-way (hereafter labelled as $\Omega_{\text{opt}}^+$) or two-way classical communication. We refer to the Supplementary Note 1 and Note 2 for more details on these strategies.

In reality, quantum devices are never perfect. Another practical scenario is to conclude with high confidence that the fidelity of the output states are above or below a certain threshold. To be specific, we want to distinguish the following two cases:

Case 1: $\mathcal{D}$ works correctly – $\forall i, \langle \psi | \sigma_i | \psi \rangle > 1 - \epsilon$. In this case, we regard the device as “good”.

Case 2: $\mathcal{D}$ works incorrectly – $\forall i, \langle \psi | \sigma_i | \psi \rangle \leq 1 - \epsilon$. In this case, we regard the device as “bad”.

We call this Task B (see Fig. 1c), which is different from Task A, since the condition for ‘$\mathcal{D}$ working correctly’ is less restrictive compared with that of Task A. It turns out that the verification strategies proposed for Task A are readily applicable to Task B. Concretely, we perform the nonadaptive verification strategy $\Omega_{\text{opt}}$ sequentially in $N$ runs and count the number of passing events $m_{\text{pass}}$. Let $X_i$ be a binary variable corresponding to the event that $\sigma_i$ passes the test ($X_i = 1$) or not ($X_i = 0$). Thus we have $m_{\text{pass}} = \sum_{i=1}^{N} X_i$. Assume that the device is “good”, then from Eq. (2) we can derive that the passing probability of the generated states is no smaller than $1 - [1 - \lambda_2(\Omega_{\text{opt}})]\epsilon$. We refer to Lemma 3 in the Supplementary Note 3A for a proof. Thus the expectation of $X_i$ satisfies $\mathbb{E}[X_i] \geq 1 - (1 - \lambda_2(\Omega_{\text{opt}}))\epsilon \equiv \mu$. The independent, identically distribution (I.I.D.) assumption together with the law of large numbers then guarantee $m_{\text{pass}} \geq N\mu$, when $N$ is sufficiently large. Following the analysis method in refs.\textsuperscript{26,27,31}, we use the Chernoff bound to upper bound the probability that the device works incorrectly as

$$\delta \equiv e^{-N D(\frac{\mu}{N} || \mu)}, \quad (4)$$

where $D(x||y) := x \log_2 \frac{2^x}{2^y} + (1 - x) \log_2 \frac{2^{1-x}}{2^{1-y}}$ the Kullback-Leibler divergence. That is to say, we can conclude with confidence $\delta_{B1} = 1 - \delta$ that $\mathcal{D}$ belongs to Case 1. Conversely, if the device is “bad”, then using the same argument we can conclude with confidence $\delta_{B2} = 1 - \delta$ that $\mathcal{D}$ belongs to Case 2.

To perform Task B with the adaptive strategy $\Omega_{\text{opt}}^+$, we record the number of passing events $m_{\text{pass}} = \sum_{i=1}^{N} X_i$. If the device is “good”, the passing probability of the generated states is no smaller than $\mu_s \equiv 1 - [1 - \lambda_4(\Omega_{\text{opt}})]\epsilon$.
where $\lambda_4 (\Omega_{\text{opt}}^\dagger) = \sin^2 \theta / (1 + \cos^2 \theta)$ is the smallest eigenvalue of $\Omega_{\text{opt}}^\dagger$, as proved by Lemma 5 in Supplementary Note 3B. Similarly, the I.I.D. assumption along with the law of large numbers guarantee that $m_{\text{pass}} \geq N\mu_s$, when $N$ is sufficiently large. On the other hand, if the device is “bad”, we can prove that the passing probability of the generated states is no larger than $\mu_1 = 1 - [1 - \lambda_2 (\Omega_{\text{opt}}^\dagger)]\epsilon$, where $\lambda_2 (\Omega_{\text{opt}}^\dagger) = \cos^2 \theta / (1 + \cos^2 \theta)$, by Lemma 4 in Supplementary Note 3B. Again, the I.I.D. assumption and the law of large numbers guarantee that $m_{\text{pass}} \leq N\mu_t$, when $N$ is large enough. Therefore, we consider two regions regarding the value of $m_{\text{pass}}$ in the adaptive strategy, i.e. the region $m_{\text{pass}} \leq N\mu_s$ and the region $m_{\text{pass}} \geq N\mu_t$. In these regions, we can conclude with $\delta_{B1} = 1 - \delta_1 / \delta_{B2} = 1 - \delta_s$ that the device belongs to \textbf{Case 1/Case 2}. The expressions for $\delta_1$ and $\delta_s$ and all the details for applying adaptive strategy to \textbf{Task B} can be found in Supplementary Note 3B.

\textbf{Experiment setup and verification procedure.} Our two-qubit entangled state is generated based on a type-II spontaneous parametric down-conversion in a 20 mm-long periodically-poled potassium titanyl phosphate (PPKTP) crystal, embedded in a Sagnac interferometer\cite{32,33} (see Fig. 2). A continuous-wave external-cavity ultraviolet (UV) diode laser at 405 nm is used as the pump light. A half-wave plate (HWP1) and quarter-wave plate (QWP1) transform the linear polarized light into the appropriate elliptically polarized light to provide the power balance and phase control of the pump field. With an input pump power of $\sim 30$ mW, we typically obtain 120 kHz coincidence counts.

![Experimental setup for optimal verification of two-qubit quantum state](image)

The target state has the following form

$$|\psi\rangle = \sin \theta |HV\rangle + e^{i\phi} \cos \theta |VH\rangle,$$

where $\theta$ and $\phi$ represent amplitude and phase, respectively. This state is locally equivalent to $|\Psi\rangle$ in Eq. (1) by $U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$. By using Lemma 1 in Supplementary Note 1, the optimal strategy for verifying $|\psi\rangle$ is $\Omega_{\text{opt}} = U \Omega_{\text{opt}} U^\dagger$, where $\Omega_{\text{opt}}$ is the optimal strategy verifying $|\Psi\rangle$ in Eq. (1). In the Supplementary Note 2, we write down explicitly the optimal nonadaptive strategy\cite{18} and adaptive strategy\cite{21} for verifying $|\psi\rangle$.

In our experiment, we implement both the nonadaptive and adaptive measurements to realize the verification strategies. There are four settings $\{P_0, P_1, P_2, P_3\}$ for nonadaptive measurements\cite{18} while only three settings $\{\tilde{T}_0, \tilde{T}_1, \tilde{T}_2\}$ are required for the adaptive measurements\cite{21}. The exact form of these projectors is given in the Supplementary Note 2. Note that the measurements $P_0 = \tilde{T}_0 = |H\rangle\langle H| \otimes |V\rangle\langle V| + |V\rangle\langle V| \otimes |H\rangle\langle H|$ are the standard $\sigma_z$ base for both the nonadaptive and adaptive strategies, which are orthogonal and can be realized with a combination of QWP, HWP and polarization beam splitter (PBS). For adaptive measurements, the measurement bases $\tilde{v}_+ = e^{i\phi} \cos \theta |H\rangle + \sin \theta |V\rangle$ /
\( \tilde{w}_+ = e^{i\phi} \cos \theta |H\rangle - i \sin \theta |V\rangle \) and \( \tilde{w}_- = e^{i\phi} \cos \theta |H\rangle - \sin \theta |V\rangle / \tilde{w}_- = e^{i\phi} \cos \theta |H\rangle + i \sin \theta |V\rangle \) at Bob’s site are not orthogonal. We transmit the results of Alice’s measurements to Bob through classical communication channel, which is implemented by real-time feed-forward operation of the electro-optic modulators (EOMs). As shown in Fig. 2, we trigger two EOMs at Bob’s site to realize the adaptive measurements based on the results of Alice’s measurement.

If Alice’s outcome is \( |+\rangle = (|V\rangle + |H\rangle)/\sqrt{2} \) or \( |R\rangle = (|V\rangle + i|H\rangle)/\sqrt{2} \), EOM1 implements the required rotation and EOM2 is identity operation. Conversely, if Alice’s outcome is \( |-\rangle = (|V\rangle - |H\rangle)/\sqrt{2} \) or \( |L\rangle = (|V\rangle - i|H\rangle)/\sqrt{2} \), EOM2 will implement the required rotation and EOM1 is identity operation. Our verification procedure is the following.

1. Specifications of quantum device. We adjust the HWP1 and QWP1 of our Sagnac source to generate the desired state in Eq. (5), where the parameter \( \theta \) and we set \( \beta = \alpha = \alpha (\theta) = (1-\alpha(\theta))/3 \) with \( \alpha(\theta) = (2-\sin(2\theta))/(4+\sin(2\theta)) \). For the adaptive strategy, \( \{\tilde{T}_0, \tilde{T}_1, \tilde{T}_2\} \) projectors are generated according to the probabilities \( \beta(\theta), (1-\beta(\theta))/2, (1-\beta(\theta))/2 \). For Task A, we use CC to decide whether the outcome of each measurement is pass or fail for each \( \sigma_i \). The passing probabilities for the nonadaptive strategy can be respectively expressed as,

\[
P_0 : \frac{CC_{HV} + CC_{VH}}{CC_{HH} + CC_{HV} + CC_{VH} + CC_{VV}},
\]

\[
P_1 : \frac{CC_{\tilde{u}_i\tilde{v}_i} + CC_{\tilde{u}_i\tilde{v}_i}}{CC_{\tilde{u}_i\tilde{v}_i} + CC_{\tilde{u}_i\tilde{v}_i} + CC_{\tilde{u}_i\tilde{v}_i} + CC_{\tilde{u}_i\tilde{v}_i}}.
\]

where \( i = 1, 2, 3, \) and \( \tilde{u}_i/\tilde{u}_i \) and \( \tilde{v}_i/\tilde{v}_i \) are the orthogonal bases for each photon and their expressions are given in the Supplementary Note 2A. For \( P_0 \), if the individual CC is in \( CC_{HV} \) or \( CC_{VH} \), it indicates that \( \sigma_i \) passes the test and we set \( X_i = 1 \); otherwise, it fails to pass the test and we set \( X_i = 0 \). For \( P_i \), \( i = 1, 2, 3, \) if the individual CC is in \( CC_{\tilde{u}_i\tilde{v}_i}, CC_{\tilde{u}_i\tilde{v}_i} \) or \( CC_{\tilde{u}_i\tilde{v}_i} \), it indicates that \( \sigma_i \) passes the test and we set \( X_i = 1 \); otherwise, it fails to pass the test and we set \( X_i = 0 \). For the adaptive strategy, we set the value of the random variables \( X_i \) in a similar way.

We increase the number of copies \( (N) \) to decide the occurrence of the first failure for Task A and the frequency of passing events for Task B. From these data, we obtain the relationship of the confidence parameter \( \delta \), the infidelity parameter \( \epsilon \), and the number of copies \( N \). There are certain probabilities that the verifier fails for each measurement. In the worst case, the probability that the verifier fails to assert \( \sigma_i \) is given by \( 1 - \Delta_c \), where \( \Delta_c = 1 - \epsilon/(2+\sin^2 \theta) \) for nonadaptive strategy \( ^{18} \) and \( \Delta_c = 1 - \epsilon/(2-\sin^2 \theta) \) for adaptive strategy \( ^{21} \).

Results and analysis of two-qubit optimal verification. The target state to be verified is the general two-qubit state in Eq. (5), where the parameter \( \theta = k \pi/10 \) and \( \phi \) is optimized with maximum likelihood estimation method.

In this section, we present the results of \( k = 2 \) state (termed as k2, see Supplementary Note 2) as an example. The other states under verification are presented in Supplementary Note 2B. The theoretical target state is \( \theta = 0.6283 \) \( (k = 2) \). In experiment, we obtain \( |\tilde{u}\rangle = 0.5987 |HV\rangle + 0.8010 e^{i2.00344} |VH\rangle \) \( (\theta = 0.6419, \phi = 3.2034) \) as our target state to be verified. In order to realize the verification strategy, the projective measurement is performed sequentially by randomly choosing the projectors. We take 10000 rounds for a fixed 6000 number of copies.

Task A. According to this verification task, we make a statistical analysis on the number of measurements required for the first occurrence of failure. According to the geometric distribution, the probability that the \( n \)-th measurement (out of \( n \) measurements) is the first failure is

\[
\Pr(N_{first} = n) = (1 - \Delta_c)^{n-1} \Delta_c,
\]

where \( n = 1, 2, 3, \ldots \). We then obtain the cumulative probability

\[
\delta_A = \sum_{N_{first}=1}^{n_{exp}} \Pr(N_{first})
\]
which is the confidence of the device generating the target state $|\psi\rangle$. In Fig. 3a, we show the distribution of the number $N_{\text{first}}$ required before the first failure for the nonadaptive (Non) strategy. From the figure we can see that $N_{\text{first}}$ obeys the geometric distribution. We fit the distribution with the function in Eq. (8) and obtain an experimental infidelity $\epsilon_{\text{exp}}^{\text{Non}} = 0.0034(15)$, which is a quantitative estimation of the infidelity for the generated state. From the experimental statistics, we obtain the number $n_{\text{exp}}^{\text{Non}} = 1650$ required to achieve the 90% confidence (i.e. 90% cumulative probability for $N_{\text{first}} \leq n_{\text{exp}}^{\text{Non}}$) of judging the generated states to be the target state in the nonadaptive strategy.

![Fig. 3](image)

Fig. 3 The distribution of the number required before the first failure for the a nonadaptive and b adaptive strategy. From the statistics, we obtain the fitting infidelity of $\epsilon_{\text{exp}}^{\text{Non}} = 0.0034(15)$ and $\epsilon_{\text{exp}}^{\text{Adp}} = 0.0121(6)$. The numbers required to achieve a 90% confidence are $n_{\text{exp}}^{\text{Non}} = 1650$ and $n_{\text{exp}}^{\text{Adp}} = 268$, respectively.

The results for the adaptive (Adp) verification of Task A are shown in Fig. 3b. The experimental fitting infidelity for this distribution is $\epsilon_{\text{exp}}^{\text{Adp}} = 0.0121(6)$. The number required to achieve the same 90% confidence as the nonadaptive strategy is $n_{\text{exp}}^{\text{Adp}} = 268$. Note this nearly six (i.e. $n_{\text{exp}}^{\text{Non}} / n_{\text{exp}}^{\text{Adp}} \sim 6$) times difference of the experimental number required to obtain the 90% confidence is partially because the infidelity with adaptive strategy is approximately four times larger than the nonadaptive strategy. However, the number of copies required to achieve the same confidence by using the adaptive strategy is still about two times fewer than the nonadaptive strategy even if the infidelity of the generated states is the same, see the analysis presented in Supplementary Note 5. This indicates that the adaptive strategy requires a significantly lower number of copies to conclude the device output state $|\psi\rangle$ with 90% confidence compared with the nonadaptive one.

**Task B.** By using Task B, we can judge whether the device belongs to Case 1 or Case 2 with a certain confidence. For each case, we can reduce the parameter $\delta = 0.10$ by increasing the number of copies of the quantum state. Thus, the confidence $\delta_B = 1 - \delta$ to judge the device belongs to Case 1/Case 2 is obtained. The passing probability $m_{\text{pass}} / N$ can finally reach a stable value 0.9986±0.0002 after about 1000 number of copies (See Supplementary Note 6). This value is smaller than the desired passing probability $\mu$ when we choose the infidelity $\epsilon_{\text{min}}$ to be 0.001. In this situation, we conclude the state belongs to Case 2. Conversely, the stable value is larger than the desired passing probability $\mu$ when we choose the infidelity $\epsilon_{\text{max}}$ to be 0.006. In this situation, we conclude the state belongs to Case 1. In Fig. 4, we present the results for the verification of Task B. First, we show the the confidence parameter $\delta$ versus the number of copies for the nonadaptive strategy in Figs. 4a, b. With about 3088 copies of quantum state, the $\delta$ parameter reaches 0.10 for Case 2. This indicates that the device belongs to Case 1 with probability at most 0.10. In other words, there are at least 90% confidence that we can say the device is in ‘bad’ case after about 3088 measurements. Generally, more copies of quantum states are required to reach a level $\delta = 0.10$ for Case 1, because there are fewer portion for the number of passing events $m_{\text{pass}}$ to be chosen in the range of $\mu N$ to $N$. From Fig. 4b, we can see that it takes about 9011 copies of quantum state in order to reduce the parameter $\delta$ to be below 0.10. At this stage, we can say that the device belongs to Case 2 with probability at most 0.10. That is, there are at least 90% confidence that we can say the device is in ‘good’ case after about 9011 measurements.

Figure 4c, d are the results of adaptive strategy. For the adaptive strategy, the passing probability $m_{\text{pass}} / N$ finally reaches a stable value 0.9914±0.0005 (see Supplementary Note 6), which is smaller than the nonadaptive measurement due to the limited fidelity of the EOM’s modulation. Correspondingly, the infidelity parameter for the two cases are chosen to be $\epsilon_{\text{min}} = 0.008$ and $\epsilon_{\text{max}} = 0.017$, respectively. We can see from the figure that it takes about 5313 number of copies for $\delta_B$ to be decreased to 0.10 when choosing $\epsilon_{\text{min}}$, which indicates that the device belongs to Case 2 with at least 90% confidence after about 5313 measurements. On the other hand, about 11762 number of copies are needed for $\delta_B$ to be decreased to 0.10 when choosing $\epsilon_{\text{max}}$, which indicates that the device belongs to Case 1 with at least 90% confidence after about 11762 measurements. Note that the difference of nonadaptive and adaptive comes from the different descent speed of $\delta$ versus the number of copies $N$ which results from the differences in passing probabilities and the infidelity parameters. See Supplementary Note 6 for detail explanations.
From another perspective, we can fix $\delta$ and see how the parameter $\epsilon$ changes when increasing the number of copies. Figure 5 presents the variation of $\epsilon$ versus the number of copies in the log-log scale when we set the $\delta$ to be 0.10. At small number of copies, the infidelity is large and drops fast to a low level when the number of copies increases to be approximately 100. The decline becomes slow when the number of copies exceeds 100. It should be noted that the $\epsilon$ asymptotically tends to a value of 0.0036 (calculated by $1 - \Delta r = 0.9986$) and 0.012 (calculated by $1 - \Delta r = 0.9914$) for the nonadaptive and adaptive strategies, respectively. Therefore, we are still in the region of $m_{pass}/N \geq \mu$. We can also see that the scaling of $\epsilon$ versus $N$ is linear in the small number of copies region. We fit the data in the linear region with $\epsilon \sim N^r$ and obtain a slope $r \sim -0.88 \pm 0.03$ for nonadaptive strategy and $r \sim -0.78 \pm 0.07$ for adaptive strategy. This scaling exceeds the standard quantum limit $\epsilon \sim N^{-0.5}$ scaling$^{31,34}$ in physical parameter estimation. Thus, our method is better for estimating the infidelity parameter $\epsilon$ than the classical metrology. Note that $m_{pass}/N$ is a good estimation for our state fidelity. If the state fidelity increases, the slope of linear region will decreases to the Heisenberg limit $\epsilon \sim N^{-1}$ in quantum metrology (see Supplementary Note 6).
Discussion

The advantage of the optimal verification strategy lies in that it requires fewer number of measurement settings, and more importantly, the number of copies to estimate the quantum states generated by a quantum device. In standard quantum state tomography, the minimum number of settings required for a complete reconstruction of the density matrix is 3\(^n\), where \(n\) is the number of qubits. For two-qubit system, the standard tomography will cost nine settings whereas present verification strategy only needs four/three measurement settings for the nonadaptive/adaptive strategy. To quantitatively compare the verification strategy with the standard tomography, we show the scaling of the parameters \(\delta\) and \(\epsilon\) versus the number of copies \(N\) in Fig. 6. For each number of copies, the fidelity estimation \(F \pm \Delta F\) can be obtained by the standard quantum state tomography. The \(\delta\) of standard tomography is calculated by the confidence assuming normal distribution of the fidelity with mean \(F\) and standard deviation \(\Delta F\). The \(\epsilon\) of standard tomography is calculated by \(\epsilon = 1 - F\). The result of verification strategy is taken from the data in Figs. 4a and 5a for the nonadaptive strategy. For \(\delta\) versus \(N\), we fit the curve with equation \(\delta = e^{g N}\), where \(g\) is the scaling of \(\log(\delta)\) with \(N\). We obtain \(g_{\text{tomo}} = -6.84 \times 10^{-5}\) for the standard tomography and \(g_{\text{verif}} = -7.35 \times 10^{-4}\) for the verification strategy. This indicates that present verification strategy achieves better confidence than standard quantum state tomography given the same number of copies. For \(\epsilon\) versus \(N\), as shown in Fig. 6b, the standard tomography will finally reach a saturation value when increasing the number of copies. With the same number of copies \(N\), the verification strategy obtains a smaller \(\epsilon\), which indicates that the verification strategy can give a better estimation for the state fidelity than the standard quantum state tomography when small number of quantum states are available for a quantum device.

![Fig. 6](image)

The variation of \(\delta\) and \(\epsilon\) versus the number of copies \(N\) by using standard quantum state tomography (tomo) and present verification strategy (verif). For standard tomography, the fidelity \(F \pm \Delta F\) is first obtained from the reconstructed density matrix of each copy \(N\). Then confidence parameter \(\delta\) is then estimated by assuming normal distribution of the fidelity with mean \(F\) and standard deviation \(\Delta F\). The infidelity parameter \(\epsilon\) is estimated by \(\epsilon = 1 - F\). Note that the experimental data symbols shown in a looks like lines due to the dense data points.

Our work, including experiment, data processing and analysis framework, can be used as a standardized procedure for verifying quantum states. In Task A, we give an estimation of the infidelity parameter \(\epsilon_{\text{exp}}\) of the generated states and the confidence \(\delta_A\) to produce the target quantum state by detecting certain number of copies. With the \(\epsilon_{\text{exp}}\) obtained from Task A, we can choose \(\epsilon_{\text{max}}\) or \(\epsilon_{\text{min}}\) which divides our device to be Case 1 or Case 2. Task B is performed based on the chosen \(\epsilon_{\text{min}}\) and \(\epsilon_{\text{max}}\). We can have an estimation for the scaling of the confidence parameter \(\delta\) versus the number of copies \(N\) based on the analysis method of Task B. With a chosen \(\delta\), we can also have an estimation for the scaling of the infidelity parameter \(\epsilon\) versus \(N\). With these steps, we can have a comprehensive judgement about how well our device really works.

In summary, we report experimental demonstrations for the optimal two-qubit pure state verification strategy with and without adaptive measurements. We give a clear discrimination and comprehensive analysis for the quantum states generated by a quantum device. Two tasks are proposed for practical applications of the verification strategy. The variation of confidence and infidelity parameter with the number of copies for the generated quantum states are presented. The obtained experimental results are in good agreement with the theoretical predictions. Furthermore, our experimental framework offers a precise estimation on the reliability and stability of quantum devices. This ability enables our framework to serve as a standard tool for analysing quantum devices. Our experimental framework can also be extended to other platforms.

References
1. Nielsen, M. A. & Chuang, I. L. Quantum Computation and Quantum Information (Cambridge University Press, UK, 2010).
2. Paris, M. & Rehacek, J. Quantum state estimation, vol. 649 (Springer Science & Business Media, 2004).
3. Sugiyama, T., Turner, P. S. & Murao, M. Precision-guaranteed quantum tomography. Phys. Rev. Lett. 111, 160406 (2013).
4. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. Phys. Rev. Lett. 105, 150401 (2010).
5. Haah, J., Harrow, A. W., Ji, Z., Wu, X. & Yu, N. Sample-optimal tomography of quantum states. In Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing, STOC 2016, 913–925 (ACM, New York, NY, USA, 2016).
6. Donnell, R. & Wright, J. Efficient quantum tomography. In Proceedings of the 48th Annual ACM Symposium on Theory of Computing, STOC 2016, 899–912 (ACM, New York, NY, USA, 2016).
7. Donnell, R. & Wright, J. Efficient quantum tomography ii. In Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, 962–974 (ACM, New York, NY, USA, 2017).
8. Haffner, H. & et al. Scalable multipartite entanglement of trapped ions. Nature (London) 438, 643–646 (2005).
9. Carolan, J. & et al. On the experimental verification of quantum complexity in linear optics. Nature Photonics 8, 621–626 (2014).
10. Tóth, G. & Gühne, O. Detecting genuine multipartite entanglement with two local measurements. Phys. Rev. Lett. 94, 060501 (2005).
11. Flammia, S. T. & Liu, Y.-K. Direct fidelity estimation from few pauli measurements. Phys. Rev. Lett. 106, 230501 (2011).
12. da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. Phys. Rev. Lett. 107, 210404 (2011).
13. Aolita, L., Gogolin, C., Kliesch, M. & Eisert, J. Reliable quantum certification of photonic state preparations. Nature Communications 6, 8498 (2015).
14. Hayashi, M. & Morimae, T. Verifiable measurement-only blind quantum computing with stabilizer testing. Phys. Rev. Lett. 115, 220502 (2015).
15. McCutcheon, W. & et al. Experimental verification of multipartite entanglement in quantum networks. Nature Communications 7, 13251 (2016).
16. Takeuchi, Y. & Morimae, T. Verification of many-qubit states. Phys. Rev. X 8, 021060 (2018).
17. Bădescu, C., Donnell, R. & Wright, J. Quantum state certification. In Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, 503–514 (ACM, New York, NY, USA, 2019).
18. Pallister, S., Linden, N. & Montanaro, A. Optimal verification of entangled states with local measurements. Phys. Rev. Lett. 120, 170502 (2018).
19. Zhu, H. & Hayashi, M. Efficient verification of hypergraph states. Phys. Rev. Applied 12, 054047 (2019).
20. Zhu, H. & Hayashi, M. Efficient verification of pure quantum states in the adversarial scenario. arXiv:1909.01900 (2019).
21. Wang, K. & Hayashi, M. Optimal verification of two-qubit pure states. Phys. Rev. A 100, 032315 (2019).
22. Yu, X.-D., Shang, J. & Gühne, O. Optimal verification of general bipartite pure states. npj Quantum Information 5, 112 (2019).
23. Li, Z., Han, Y.-G. & Zhu, H. Efficient verification of bipartite pure states. Phys. Rev. A 100, 032316 (2019).
24. Liu, Y.-C., Yu, X.-D., Shang, J., Zhu, H. & Zhang, X. Efficient verification of dicke states. Phys. Rev. Applied 12, 044020 (2019).
25. Li, Z., Han, Y.-G. & Zhu, H. Optimal verification of greenberger-horne-zeilinger states. arXiv:1909.08979 (2019).
26. Dimić, A. & Dakić, B. Single-copy entanglement detection. npj Quantum Information 4, 11 (2018).
27. Saggio, V. & et al. Experimental few-copy multipartite entanglement detection. Nature Physics 15, 935–940 (2019).
28. Knips, L., Schwemmer, C., Klein, N., Wieśniak, M. & Weinfurter, H. Multiparticle entanglement detection with minimal effort. Phys. Rev. Lett. 117, 210504 (2016).
29. Bavaresco, J. & et al. Measurements in two bases are sufficient for certifying high-dimensional entanglement. Nature Physics 14, 1032 (2018).
30. Friis, N., Vitagliano, G., Malik, M. & Huber, M. Entanglement certification from theory to experiment. Nature Reviews Physics 1, 72–87 (2019).
31. Zhang, W.-H. & et al. Experimental optimal verification of entangled states using local measurements. arXiv:1905.12175 (2019).
32. Kim, T., Fiorentino, M. & Wong, F. N. C. Phase-stable source of polarization-entangled photons using a polarization sagnac interferometer. Phys. Rev. A 73, 012316 (2006).
33. Fedrizzi, A., Herbst, T., Poppe, A., Jennewein, T. & Zeilinger, A. A wavelength-tunable fiber-coupled source of narrowband entangled photons. Opt. Express 15, 15377–15386 (2007).
34. Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. Nature Photonics 5, 222–229 (2011).
35. Altepeter, J. B., Jeffery, E. R. & Kwiat, P. Photonic state tomography. vol. 52 of Advances In Atomic, Molecular, and Optical Physics, 105–159 (Academic Press, 2005).
of China (No. 3704), NSFC-BRICS (No. 61961146001), the Natural Science Foundation of Jiangsu Province (No. BK20170010), the Program for Innovative Talents and Entrepreneurs in Jiangsu and the Fundamental Research Funds for the Central Universities.

**Author contributions**

X.-H. J., K.-Y. Q., Z.-Z. C., Z.-Y. C. and X.-S. M. designed and performed the experiment. K. W. performed the theoretical analysis. X.-H. J., K.-Y. Q. analysed the data. X.-H. J., K. W. and X.-S. M. wrote the paper with input from all authors. All authors discussed the results and read the manuscript. F.-M. S., S.-N. Z. and X.-S. M. supervised the work.
Supplementary Information – Towards the standardization of quantum state verification using optimal strategies

Contents

1. Supplementary Note 1: The theory of optimal state verification 12
   A. Optimal nonadaptive strategy 12
   B. Optimal adaptive strategy using one-way classical communication 13
   C. Optimal verification of Bell state 13
   D. Optimal verification of a product state 14

2. Supplementary Note 2: The experiment of optimal state verification 14
   A. The verification strategies for experimentally generated state 14
   B. The experimental verification procedure 16

3. Supplementary Note 3: Applying the verification strategy to Task B 17
   A. Applying the nonadaptive strategy to Task B 18
   B. Applying the adaptive strategy to Task B 20

4. Supplementary Note 4: The experimental results for the three target states 21

5. Supplementary Note 5: Explanations for the results in Task A 22

6. Supplementary Note 6: Comparison of nonadaptive and adaptive in Task B 23
Supplementary Note 1: The theory of optimal state verification

First we argue that it suffices to derive optimal verification strategies for the state $|\Psi\rangle$ defined in Eq. (1) of the main text. This is because locally unitarily equivalent pure states have locally unitarily equivalent optimal verification strategies. This fact is proved in ref.\textsuperscript{S1} (Lemma 2) and we restate here for completeness.

**Lemma 1** (Lemma 2 in the Supplemental Material of ref.\textsuperscript{S1}). Given any two-qubit state $|\psi\rangle$ with optimal strategy $\Omega$, a locally unitarily equivalent state $(U \otimes V)|\psi\rangle$, where $U$ and $V$ are unitaries, has optimal strategy $(U \otimes V)\Omega(U \otimes V)^\dagger$.

### A. Optimal nonadaptive strategy

In this section, we briefly summarize the optimal nonadaptive strategy. According to ref.\textsuperscript{S1}, any optimal strategy for verifying state the state $|\Psi\rangle$ $(\theta \in (0, \pi/4))$ defined in Eq. (1) of the main text, that accepts $|\Psi\rangle$ with certainty and satisfies the properties of locality, projectivity, and trusty, can be expressed as a strategy involving the following four measurements,

$$\Omega_{opt} = \alpha(\theta)P_{Z Z}^+ + \frac{1 - \alpha(\theta)}{3} \sum_{k=1}^{3} [I - |u_k\rangle\langle u_k| \otimes |v_k\rangle\langle v_k|], \tag{S1}$$

where $\alpha(\theta) = (2 - \sin(2\theta))/(4 + \sin(2\theta))$, $P_{Z Z}^+ = |HH\rangle\langle HH| + |VV\rangle\langle VV|$ is the projector onto the positive eigenspace of the tensor product of Pauli matrix $Z \otimes Z$, and the states $\{|u_k\rangle\}$ and $\{|v_k\rangle\}$ are written explicitly in the following:

$$|u_1\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{\pi i/4}}{\sqrt{1 + \cot \theta}} |V\rangle, \tag{S2}$$

$$|v_1\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{\pi i/4}}{\sqrt{1 + \cot \theta}} |V\rangle, \tag{S3}$$

$$|u_2\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{\pi i/4}}{\sqrt{1 + \cot \theta}} |V\rangle, \tag{S4}$$

$$|v_2\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{i \pi}}{\sqrt{1 + \cot \theta}} |V\rangle, \tag{S5}$$

$$|u_3\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{1}{\sqrt{1 + \cot \theta}} |V\rangle, \tag{S6}$$

$$|v_3\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle - \frac{1}{\sqrt{1 + \cot \theta}} |V\rangle. \tag{S7}$$

Correspondingly, the second largest eigenvalue of $\Omega_{opt}$ is given by

$$\lambda_2(\Omega_{opt}) = \frac{2 + \sin(2\theta)}{4 + \sin(2\theta)}. \tag{S8}$$

Substitute this value into Eq. (3) of the main text, we find the optimal number of measurements required to verify $|\Psi\rangle$ with infidelity $\epsilon$ and confidence $1 - \delta$ satisfies

$$n_{opt} \approx \frac{1}{[1 - \lambda_2(\Omega_{opt})]\epsilon} \ln \frac{1}{\delta} = (2 + \sin \theta \cos \theta) \frac{1}{\epsilon} \ln \frac{1}{\delta}. \tag{S9}$$

For analysis, we consider the spectral decomposition of $\Omega_{opt}$. This decomposition is helpful for computing the passing probability of the states being verified. Let $|\Psi^-\rangle := \cos \theta |HH\rangle - \sin \theta |VV\rangle$. One can check that $\{|\Psi\rangle, |\Psi^-\rangle, |HV\rangle, |VH\rangle\}$ forms an orthonormal basis of a two-qubit space. The spectral decomposition of $\Omega_{opt}$ is

$$\Omega_{opt} = |\Psi\rangle\langle \Psi| + \lambda_2(\Omega_{opt}) \left( |\Psi^-\rangle\langle \Psi^-| + |VH\rangle\langle VH| + |HV\rangle\langle HV| \right), \tag{S10}$$

where $\lambda_2(\Omega_{opt})$ is given in Eq. (S8).
B. Optimal adaptive strategy using one-way classical communication

In this section, we briefly summarize the optimal adaptive strategy using one-way classical communication. According to ref.\textsuperscript{S2} any optimal strategy for verifying the state $|\Psi\rangle$ $(\theta \in (0, \pi/4))$ defined in Eq. (1) of the main text, that accepts $|\Psi\rangle$ and can be implemented by local measurements together with one-way classical communication, can be expressed as a strategy involving the following three measurements,

$$\Omega_{\text{opt}}^\rightarrow = \beta(\theta)P^+_{ZZ} + \frac{1 - \beta(\theta)}{2}T_1 + \frac{1 - \beta(\theta)}{2}T_2,$$

(S11)

where $\beta(\theta) = \cos^2 \theta/(1 + \cos^2 \theta)$ and

$$T_1 = |+\rangle\langle+| \otimes |v_+\rangle\langle v_+| + |-\rangle\langle-| \otimes |v_-\rangle\langle v_-|,$$

(S12)

$$T_2 = |R\rangle \langle R| \otimes |w_+\rangle\langle w_+| + |L\rangle \langle L| \otimes |w_-\rangle\langle w_-|,$$

(S13)

such that

$$|+\rangle = \frac{|V| + |H\rangle}{\sqrt{2}},$$

$$|-\rangle = \frac{|V\rangle - |H\rangle}{\sqrt{2}},$$

(S14)

$$|v_+\rangle = \cos \theta |V\rangle + \sin \theta |H\rangle,$$

$$|v_-\rangle = \cos \theta |V\rangle - \sin \theta |H\rangle,$$

(S15)

$$|R\rangle = \frac{|V\rangle + i|H\rangle}{\sqrt{2}},$$

$$|L\rangle = \frac{|V\rangle - i|H\rangle}{\sqrt{2}},$$

(S16)

$$|w_+\rangle = \cos \theta |V\rangle - i \sin \theta |H\rangle,$$

$$|w_-\rangle = \cos \theta |V\rangle + i \sin \theta |H\rangle.$$  

(S17)

Correspondingly, the second largest eigenvalue of $\Omega_{\text{opt}}^\rightarrow$ is given by

$$\lambda_2(\Omega_{\text{opt}}^\rightarrow) = \frac{\cos^2 \theta}{1 + \cos^2 \theta}.$$  

(S18)

Substitute this value into Eq. (3) of the main text, we find the optimal number of measurements required to verify $|\Psi\rangle$ with infidelity $\epsilon$ and confidence $1 - \delta$ satisfies

$$n_{\text{opt}}^\rightarrow \approx \frac{1}{\ln(1 - \lambda_2(\Omega_{\text{opt}}^\rightarrow))\epsilon} \ln \frac{1}{\delta} = (1 + \cos^2 \theta) \frac{1}{\epsilon} \ln \frac{1}{\delta}.$$  

(S19)

The spectral decomposition of $\Omega_{\text{opt}}^\rightarrow$ is given by

$$\Omega_{\text{opt}}^\rightarrow = |\Psi\rangle \langle\Psi| + \lambda_2(\Omega_{\text{opt}}^\rightarrow) (|\Psi^\perp\rangle \langle\Psi^\perp| + |HV\rangle \langle HV|) + \lambda_4(\Omega_{\text{opt}}^\rightarrow) |VH\rangle \langle VH|,$$  

(S20)

where $\lambda_2(\Omega_{\text{opt}}^\rightarrow)$ is the second largest eigenvalue given in Eq. (S18) and $\lambda_4(\Omega_{\text{opt}}^\rightarrow) = \sin^2 \theta/(1 + \cos^2 \theta)$ is the fourth (which is also the smallest) eigenvalue of $\Omega_{\text{opt}}^\rightarrow$.

C. Optimal verification of Bell state

As pointed out in ref.\textsuperscript{S1}, the above nonadaptive strategy is actually only optimal for $\theta \in (0, \pi/4) \cup (\pi/4, \pi/2)$. That is to say, it is no longer optimal for verifying the Bell state, for which $\theta = \pi/4$. In this section, we summary explicitly the optimal strategy for verifying the Bell state of the following form:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|HH\rangle + |VV\rangle).$$  

(S21)

According to Eq. (8) of ref.\textsuperscript{S1}, any optimal strategy for verifying the Bell state $|\Phi^+\rangle$ implementable by locally projective measurements, can be expressed as a strategy involving the following three measurements,

$$\Omega_{\text{opt,Bell}} = \frac{1}{3} (P^+_{XX} + P^+_{YY} + P^+_{ZZ}),$$  

(S22)

where $P^+_{XX}$ is the projector onto the positive eigenspace of the tensor product of Pauli matrix $X \otimes X$ and $P^+_{YY}$ is the projector onto the negative eigenspace of the tensor product of Pauli matrix $Y \otimes Y$, similarly defined as that of $P^+_{ZZ}$. We can check that $\lambda_2(\Omega_{\text{opt,Bell}}) = 1/3$ and thus the optimal number of measurements required to verify $|\Phi^+\rangle$ with infidelity $\epsilon$ and confidence $1 - \delta$ satisfies

$$n_{\text{opt,Bell}} \approx \frac{1}{\ln(1 - \lambda_2(\Omega_{\text{opt,Bell}}))\epsilon} \ln \frac{1}{\delta} = \frac{3}{2\epsilon} \ln \frac{1}{\delta},$$  

(S23)
D. Optimal verification of a product state

For the product state $|HV\rangle$ ($\theta = 0$ or $\pi/2$), the optimal strategy is provably given by $\Omega_{\text{opt.pd}} = |HV\rangle \langle HV|$. Obviously, $\lambda_2(\Omega_{\text{opt.pd}}) = 1$ and thus the optimal number of measurements required to verify $|HV\rangle$ with infidelity $\epsilon$ and confidence $1 - \delta$ satisfies

$$n_{\text{opt.pd}} \approx \frac{1}{|1 - \lambda_2(\Omega_{\text{opt.pd}})|\epsilon} \ln \frac{1}{\delta} = \frac{1}{\epsilon} \ln \frac{1}{\delta}.$$  \hfill (S24)

**Supplementary Note 2:** The experiment of optimal state verification

**A. The verification strategies for experimentally generated state**

In our experiment setting, the target state has the form $\psi(\theta, \phi) = \sin \theta |HV\rangle + e^{i\phi} \cos \theta |VH\rangle$, \hfill (S25)

where $\theta \in [0, \pi/4]$ and $\phi \in [0, 2\pi]$, which will be determined by the experimental data. We can see that $|\psi(\theta, \phi)\rangle$ is equivalent to $|\Psi\rangle$ defined in Eq. (1) of the main text through the following unitary operator:

$$U \equiv (1 \otimes \kappa) = \begin{pmatrix} 0 & e^{i\phi} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & e^{i\phi} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \sin \theta \\ 0 \\ 0 \\ \cos \theta \end{pmatrix} = \begin{pmatrix} 0 \\ \sin \theta \\ e^{i\phi} \cos \theta \\ 0 \end{pmatrix} = |\psi(\theta, \phi)\rangle.$$  \hfill (S26)

This is indeed the case since

$$(1 \otimes \kappa)|\Psi\rangle = \begin{pmatrix} 0 & e^{i\phi} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & e^{i\phi} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \sin \theta \\ 0 \\ 0 \\ \cos \theta \end{pmatrix} = \begin{pmatrix} 0 \\ \sin \theta \\ e^{i\phi} \cos \theta \\ 0 \end{pmatrix} = |\psi(\theta, \phi)\rangle.$$  \hfill (S27)

Then Lemma 1 together with the optimal strategies for verifying $|\Psi\rangle$ summarized in the last section yields the optimal strategies for verifying $|\psi(\theta, \phi)\rangle$. For completeness, we write down explicitly the measurements of these strategies.

**Optimal nonadaptive strategy for** $|\psi(\theta, \phi)\rangle$. By Lemma 1 and Eq. (S1), the optimal nonadaptive strategy for verifying $|\psi(\theta, \phi)\rangle$ has the following form,

$$\Omega'_{\text{opt}} = \mathbb{U} \Omega_{\text{opt}} \mathbb{U}^\dag = \alpha(\theta) P_0 + \frac{1 - \alpha(\theta)}{3} (P_1 + P_2 + P_3),$$  \hfill (S28)

where

$$P_0 = \mathbb{U} P_{ZZ}^+ \mathbb{U}^\dag = |H\rangle \langle H| \otimes |V\rangle \langle V| + |V\rangle \langle V| \otimes |H\rangle \langle H|,$$  \hfill (S29)

and $P_i$ for $i = 1, 2, 3$ satisfies $P_i = |\tilde{u}_i\rangle \langle \tilde{u}_i| \otimes |\tilde{v}_i\rangle \langle \tilde{v}_i|$ such that

$$|\tilde{u}_1\rangle = |u_1\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{i\phi}}{\sqrt{1 + \cot \theta}} |V\rangle,$$  \hfill (S30)

$$|\tilde{v}_1\rangle = \kappa |v_1\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |V\rangle + \frac{e^{i\phi}}{\sqrt{1 + \cot \theta}} |H\rangle,$$  \hfill (S31)

$$|\tilde{u}_2\rangle = |u_2\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{e^{i\phi}}{\sqrt{1 + \cot \theta}} |V\rangle,$$  \hfill (S32)

$$|\tilde{v}_2\rangle = \kappa |v_2\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |V\rangle + \frac{e^{i\phi}}{\sqrt{1 + \cot \theta}} |H\rangle,$$  \hfill (S33)

$$|\tilde{u}_3\rangle = |u_3\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |H\rangle + \frac{1}{\sqrt{1 + \cot \theta}} |V\rangle,$$  \hfill (S34)

$$|\tilde{v}_3\rangle = \kappa |v_3\rangle = \frac{1}{\sqrt{1 + \tan \theta}} |V\rangle + \frac{e^{i\phi}}{\sqrt{1 + \cot \theta}} |H\rangle.$$  \hfill (S35)
Optimal adaptive strategy using one-way classical communication for $|\psi(\theta, \phi)\rangle$. By Lemma 1 and Eq. (S11), the optimal adaptive strategy for verifying $|\psi(\theta, \phi)\rangle$, when one-way classical communication is allowed, has the following form,

$$\Omega_{\text{opt}}' = \beta(\theta)\tilde{T}_0 + \frac{1 - \beta(\theta)}{2} \tilde{T}_1 + \frac{1 - \beta(\theta)}{2} \tilde{T}_2,$$

(S36)

where

$$\tilde{T}_0 = UP_{ZZ}^+ U^\dagger = |H\rangle\langle H| \otimes |V\rangle\langle V| + |V\rangle\langle V| \otimes |H\rangle\langle H|,$$
(S37)

$$\tilde{T}_1 = UT_1 U^\dagger$$

$$= |+\rangle\langle +| \otimes \kappa |v_+\rangle\langle v_+| + |\rangle\langle -| \otimes \kappa |v_-\rangle\langle v_-|$$

$$= |+\rangle\langle +| \otimes \tilde{v}_+ \langle \tilde{v}_+| + |\rangle\langle -| \otimes \tilde{v}_- \langle \tilde{v}_-|,$$

(S38)

$$\tilde{T}_2 = UT_2 U^\dagger$$

$$= |R\rangle\langle R| \otimes \kappa |w_+\rangle\langle w_+| + |L\rangle\langle L| \otimes \kappa |w_-\rangle\langle w_-|$$

$$= |R\rangle\langle R| \otimes \tilde{w}_+ \langle \tilde{w}_+| + |L\rangle\langle L| \otimes \tilde{w}_- \langle \tilde{w}_-|,$$

(S39)

and

$$|+\rangle = \frac{|V\rangle + |H\rangle}{\sqrt{2}},$$

$$|-\rangle = \frac{|V\rangle - |H\rangle}{\sqrt{2}},$$

(S40)

$$|R\rangle = \frac{|V\rangle + i|H\rangle}{\sqrt{2}},$$

$$|L\rangle = \frac{|V\rangle - i|H\rangle}{\sqrt{2}},$$

(S41)

$$|v_+\rangle = \kappa |v_+\rangle = e^{i\phi} \cos \theta |H\rangle + \sin \theta |V\rangle,$$

$$|v_-\rangle = \kappa |v_-\rangle = e^{i\phi} \cos \theta |H\rangle - \sin \theta |V\rangle,$$

(S42)

$$|\tilde{v}_+\rangle = \kappa |v_+\rangle = e^{i\phi} \cos \theta |H\rangle + i \sin \theta |V\rangle,$$

$$|\tilde{v}_-\rangle = \kappa |v_-\rangle = e^{i\phi} \cos \theta |H\rangle - i \sin \theta |V\rangle.$$

(S43)

Experimental optimal verification of Bell state. Experimentally, our target Bell state has the following form

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}} (|HV\rangle - |VH\rangle),$$

(S44)

which is locally unitarily equivalent to the Bell state $|\Phi^+\rangle$ defined in Eq. (S21) through

$$(\mathbb{1} \otimes XZ)|\Phi^+\rangle = \frac{1}{\sqrt{2}} (\mathbb{1} \otimes XZ) (|HH\rangle + |VV\rangle) = \frac{1}{\sqrt{2}} (|HV\rangle - |VH\rangle) = |\Phi^-\rangle.$$

(S45)

By Lemma 1, the optimal adaptive strategy for verifying $|\Phi^-\rangle$ is given by

$$\Omega_{\text{opt,Bell}}' = \frac{1}{3} (M_1 + M_2 + M_3),$$

(S46)

where

$$M_1 = (\mathbb{1} \otimes XZ) P_{XX}^+ (\mathbb{1} \otimes XZ)^\dagger = |\rangle\langle +| \otimes |\rangle\langle +| + |\rangle\langle +| \otimes |\rangle\langle -| + |\rangle\langle -| \otimes |\rangle\langle -|,$$

(S47)

$$M_2 = (\mathbb{1} \otimes XZ) P_{YY}^+ (\mathbb{1} \otimes XZ)^\dagger = |L\rangle\langle L| \otimes |R\rangle\langle R| + |R\rangle\langle R| \otimes |L\rangle\langle L|,$$

(S48)

$$M_3 = (\mathbb{1} \otimes XZ) P_{ZZ}^+ (\mathbb{1} \otimes XZ)^\dagger = |H\rangle\langle H| \otimes |V\rangle\langle V| + |V\rangle\langle V| \otimes |H\rangle\langle H|.$$

(S49)

In the above derivation, we have used the following relations:

$$XZ|H\rangle = |V\rangle,$$

$$XZ|V\rangle = -|H\rangle,$$

(S50)

$$XZ|+\rangle = -|+\rangle,$$

$$XZ|-\rangle = |+\rangle,$$

(S51)

$$XZ|L\rangle = -i|L\rangle,$$

$$XZ|R\rangle = i|R\rangle.$$
B. The experimental verification procedure

We now describe the procedure on how to verify the states $|\psi(\theta, \phi)\rangle$ described above. The parameter $\theta$ and $\phi$ can be chosen to generate arbitrary pure states according to the angles of our wave plate (QWP1 and HWP1 in Fig. 2 of the main text). In order to tune the angle $\theta$ to generate several states between $0$ and $\pi/2$, we let $\theta = k \cdot \pi/10$ ($k = 1, 2, 3, 4$). For reference, we list the tomographic parameters for $k = 2$ (k2), maximally entangled state (Max) and product state (HV) used in this article in Supplementary Table 1.

**Supplementary Table 1** Target states verified in our experiment.

| $k$ | $k\pi/10$ | $\theta$ | $\phi$ | Fidelity | Passing Prob. | $m/N$ |
|-----|-----------|---------|-------|---------|---------------|-------|
| k2  | 0.6283    | 0.6194  | 3.2034| 0.9964+0.0002 | 0.9986+0.0002 |
| Max | $\pi/4$   | $\pi/4$ | $\pi$ | 0.9973+0.0002 | 0.9982+0.0002 |
| HV  | $\pi/2$   | $\pi/2$ | 0     | 0.9992+0.0001 | 0.9992+0.0001 |

**Determine the target state.** For our target state defined in Eq. (S25), $\theta$ and $\phi$ are two parameters to be set experimentally. For each $k$, we first calculate the ratio of HV component to VH component, i.e. $(\sin \theta / \cos \theta)^2$, in the target state. Based on this ratio, we rotate the angles of QWP1 and HWP1 of pump light to generate this target state. Then, the density matrix $\rho$ of the target state can be obtained by means of the maximum likelihood estimations. The overlap between $|\psi\rangle$ and $\rho$ of the target state can be obtained by means of the maximum likelihood estimations. If we want to perform from the two EOMs to compensate this extra phase. We calibrate the optical axis of these two EOMs and fix them.

**The projective measurements.** The nonadaptive strategy requires four projective measurements. From the expression of the four projectors (Eq. (S29-S35)), we can see that $P_0$ is the standard $\sigma_z$ projection while $P_i$ ($i = 1, 2, 3$) are three general projectors that are realized by rotating the angles of wave plates in the state analyser (see Fig. 2 in main text). The function of $P_i$ projectors is to transform the $|\vec{v}_i\rangle$ and $|\vec{v}_i\rangle$ states to the $|H\rangle$ state with combinations of QWP and HWP. Therefore, we treat the angles of QWP and HWP as two quantities to realize the transformation utilizing Jones matrix method. By solving the equations, we find the angles of wave plates that realizes the four projective measurements.

For the adaptive strategy, three measurements $\vec{T}_0$, $\vec{T}_1$ and $\vec{T}_2$ are required. From Eq. (S37-S43), we can see that $\vec{T}_0$ is the standard $\sigma_z$ projection. The $\vec{T}_1$ and $\vec{T}_2$ measurements requires classical communication to transmit the results of Alice’s measurements to Bob. Then Bob applies the corresponding measurements using the electro-optic modulators (EOMs) according to Alice’s results. For $\vec{T}_1$, Alice implements the $\{|+\rangle, |-\rangle\}$ orthogonal measurements while Bob implements the two measurements $|\vec{v}_+\rangle$ and $|\vec{v}_-\rangle$ are not orthogonal. For $\vec{T}_2$, Alice’s measurements $\{|R\rangle, |L\rangle\}$ are orthogonal whereas Bob need to apply the nonorthogonal measurements $|\vec{w}_+\rangle$ and $|\vec{w}_-\rangle$ accordingly. In the following, we describe how to realize these adaptive measurements by means of local operations and classical communication (LOCC) in real time.

In order to implement the $\{|\vec{v}_+\rangle, |\vec{v}_-\rangle\}$ and $\{|\vec{w}_+\rangle, |\vec{w}_-\rangle\}$ measurements at Bob’s site according to Alice’s measurement outcomes, we use two EOMs, as shown in Supplementary Figure 1. The $|\vec{v}_+\rangle$ and $|\vec{w}_+\rangle$ polarized states are rotated to $H$ polarization with EOM1, which finally exit from the transmission port of polarized beam splitter (PBS, the wave plates before PBS are rotated to $HV$ bases). Correspondingly, the $|\vec{v}_-\rangle$ and $|\vec{w}_-\rangle$ are transformed to the $V$ polarization with EOM2 and exit from the reflection port of PBS. We use the output of $|+\rangle/|R\rangle$ to trigger the EOM1 and the output of $|-\rangle/|L\rangle$ to trigger the EOM2, respectively. This design guarantees that only one EOM takes effect and the other executes the identity operation for each of the generated state. In Supplementary Table 2, we list the operations realized with these two EOMs. We can see that the adaptive measurements are genuinely realized with our setup.

There are two key ingredients for our adaptive measurement that distinguish it from the nonadaptive measurement. First, Alice produces binary outcomes for a given basis. In each run, only one of these two outcomes will be obtained. Bob’s measurements have to be adaptive according to Alice’s measurement results. Second, it doesn’t require fast changes among the $\vec{T}_0$, $\vec{T}_1$ and $\vec{T}_2$ measurements. This can be seen from the fact that $\{|\vec{w}_+\rangle, |\vec{w}_-\rangle\}$ have an extra $\pi/2$ phase compared with the $\{|\vec{v}_+\rangle, |\vec{v}_-\rangle\}$ bases (see Eqs. (S42,S43)). We add a QWP (see Supplementary Figure 1) before the two EOMs to compensate this extra phase. We calibrate the optical axis of these two EOMs and fix them.

If we want to perform from $\vec{T}_1$ to $\vec{T}_2$ measurement or vice versa, we add or remove this QWP. For $\vec{T}_0$ measurement, we turn off the drivers of these two EOMs and remove the QWP. In the single-photon experiment, we can optimize the modulation contrast of these two EOMs with the coincidence ratio of $CC_{+}\vec{v}_+/CC_{+}\vec{v}_+$ and $CC_{-}\vec{v}_-/CC_{-}\vec{v}_-$ as our optimization goals (see Supplementary Table 2).
Supplementary Figure 1 The adaptive measurements implemented with electro-optic modulator (EOM). Here we take $\tilde{T}_1$ measurement as an illustration. The output signal of $|+\rangle$ is used to trigger EOM1 to execute the $|\tilde{v}_+\rangle$ measurement, while the output signal of $|-\rangle$ is used to trigger EOM2 to execute the $|\tilde{v}_-\rangle$ measurement. Only one EOM is triggered at a time and the other EOM is an identity operation, which guarantees the adaptive measurements are realized in real time. A QWP before the two EOMs is used to compensate the extra phase from $\tilde{T}_1$ to $\tilde{T}_2$ measurement.

Supplementary Table 2 Adaptive measurements $\{\tilde{T}_0, \tilde{T}_1, \tilde{T}_2\}$ realized with two EOMs.

| Setting | Alice’s operations | Bob’s operations | Output | Modulation contrast* | Success probability |
|---------|-------------------|------------------|--------|----------------------|-------------------|
| $\tilde{T}_0$ | $H$ (EOM1) | $I$ (EOM2) | — | $CC_{HH}/CC_{HH}$ | $CC_{HH}+CC_{VV}$ |
| | $V$ (EOM1) | $I$ (EOM2) | — | $CC_{VH}/CC_{VH}$ | $CC_{VH}+CC_{HH}+CC_{VV}$ |
| $\tilde{T}_1$ | $+$ | $\tilde{v}_+$ (EOM1) | $I$ (EOM2) | $\tilde{v}_+ \rightarrow H$ | $CC_{+}+CC_{++}$ | $CC_{+}+CC_{++}$ |
| | $-$ | $\tilde{v}_-$ (EOM1) | $I$ (EOM2) | $\tilde{v}_- \rightarrow V$ | $CC_{-}+CC_{\mp}$ | $CC_{-}+CC_{\mp}$ |
| $\tilde{T}_2$ | $R$ | $\tilde{w}_+$ (EOM1) | $I$ (EOM2) | $\tilde{w}_+ \rightarrow H$ | $CC_{R}+CC_{\tilde{w}}$ | $CC_{R}+CC_{\tilde{w}}$ |
| | $L$ | $\tilde{w}_-$ (EOM1) | $I$ (EOM2) | $\tilde{w}_- \rightarrow V$ | $CC_{L}+CC_{\tilde{w}}$ | $CC_{L}+CC_{\tilde{w}}$ |

* $\tilde{v}_+/\tilde{v}_-$ and $\tilde{w}_+/\tilde{w}_-$ are the orthogonal states to $\tilde{v}_+/\tilde{v}_-$ and $\tilde{w}_-/\tilde{w}_-$, respectively.

Supplementary Note 3: Applying the verification strategy to Task B

In this section, we explain in detail how we use the verification strategy proposed in Supplementary Note 1 to execute Task B. Generally speaking, Task B is to distinguish whether the quantum device outputs states $\epsilon$-far from the target state $|\Psi\rangle$ or not. Let $\mathcal{S}$ be the set of states that are at least $\epsilon$-far always from $|\Psi\rangle$, i.e.

$$\mathcal{S} := \{\sigma_i : \langle \Psi | \sigma_i | \Psi \rangle \leq 1 - \epsilon\}, \quad (S53)$$

and $\overline{\mathcal{S}}$ to be the complement of $\mathcal{S}$, which is the set of states that are $\epsilon$-close to $|\Psi\rangle$. See Supplementary Figure 2 for illustration.

Before going to the details, we compare briefly the differences between Task A and Task B.

1. **Task A**\textsuperscript{81,82} considers the problem: Performing tests sequentially, how many tests are required before we identify a $\sigma_i$ that does not pass the test? Once we find a $\sigma_i$ not passing the test, we conclude that the device is bad. Each test is a Bernoulli trial with probability $\Delta_\epsilon$, determined by the chosen verification strategy. The expected number of tests required is given by $1/\Delta_\epsilon$. It should be stressed that this expectation is achieved with the assumption that all generated states are $\epsilon$-far from $|\Psi\rangle$. Once the generated states are much further from $|\Psi\rangle$ by $\epsilon$ (Case 2), then the observed number of tests will be less than $1/\Delta_\epsilon$.

2. **Task B**\textsuperscript{84} considers the problem: Performing $N$ tests, how many states $\sigma_i$ can pass these tests on average? If all generated states are exactly $\epsilon$-far from $|\Psi\rangle$, then the expected number of states passing the test is $N\mu$, where $\mu$ is the expected passing probability of each generated state, determined by the chosen verification strategy. If all generated states are much far from $|\Psi\rangle$ by $\epsilon$ (Case 2), then the expected number of states passing the test is less than $N\mu$. If all generated states are much close to $|\Psi\rangle$ by $\epsilon$ (Case 1), then the expected number of states passing the test is larger than $N\mu$. 

\[\]
Supplementary Figure 2 $S$ is the set of states that are at least $\epsilon$-far always from $|\Psi\rangle$, that is, states that are out of the brown circle. The states that are on the circle (the blue circle, $\langle \Psi | \psi \rangle = 1 - \epsilon$) are the most difficult to be distinguished from $|\Psi\rangle$, compared to states in $S$.

A. Applying the nonadaptive strategy to Task B

Here we illustrate how to use the nonadaptive verification strategy $\Omega_{\text{opt}}$ proposed in Supplementary Note 1 A to execute Task B. First of all, the following two lemmas are essential, which state that for arbitrary state belonging to $S$ ($S$), its probability of passing the test $\Omega_{\text{opt}}$ is upper (lower) bounded.

**Lemma 2.** For arbitrary state $\sigma \in S$, it can pass $\Omega_{\text{opt}}$ with probability no larger than $1 - |1 - \lambda_2(\Omega_{\text{opt}})|\epsilon$, where $\lambda_2(\Omega_{\text{opt}})$ is the second largest eigenvalue of $\Omega_{\text{opt}}$ given in Eq. (S8). The upper bound is achieved by states in $S$ that are exactly $\epsilon$-far from $|\Psi\rangle$.

Intuitively, the further a state $\sigma \in S$ is from $|\Psi\rangle$, the smaller the probability it passes the test $\Omega_{\text{opt}}$. Lemma 2 justifies this intuition and shows quantitatively that the passing probability for states in $S$ cannot be larger than $1 - |1 - \lambda_2(\Omega_{\text{opt}})|\epsilon$. However, we remark that it is possible that for some states in $S$ the passing probability is small.

**Proof.** From the spectral decomposition Eq. (S10) of $\Omega_{\text{opt}}$ and the fact that the off-diagonal parts of $\sigma$ do not affect the trace $\text{Tr}[\Omega_{\text{opt}} \sigma]$, we can assume without loss of generality that $\sigma$ has the form

$$\sigma = p_1|\Psi\rangle\langle\Psi| + p_2|\Psi^\perp\rangle\langle\Psi^\perp| + p_3|HV\rangle\langle HV| + p_4|VH\rangle\langle VH|,$$

where the last constraint $p_1 \leq 1 - \epsilon$ follows from the precondition that $\sigma \in S$. Then

$$\text{Tr}[\Omega_{\text{opt}} \sigma] = p_1 + \sum_{i=2}^{4} \lambda_ip_i \leq p_1 + (1 - p_1)\lambda_2 \leq (1 -\epsilon)(1 - \lambda_2) + \lambda_2 = 1 - (1 - \lambda_2)\epsilon,$$

where $\lambda_i$ is the $i$-th eigenvalue of $\Omega_{\text{opt}}$, the first inequality follows from $\lambda_2 \geq \lambda_3 \geq \lambda_4$ and the second inequality follows from $p_1 \leq 1 - \epsilon$.

**Lemma 3.** For arbitrary state $\sigma \in \overline{S}$, it can pass the nonadaptive strategy $\Omega_{\text{opt}}$ with probability no smaller than $1 - |1 - \lambda_2(\Omega_{\text{opt}})|\epsilon$. The lower bound is achieved by states in $S$ that are exactly $\epsilon$-far from $|\Psi\rangle$.

**Proof.** Following the proof for Lemma 2 and considering the spectral decomposition Eq. (S10) of $\Omega_{\text{opt}}$, we can assume without loss of generality that $\sigma$ has the form

$$\sigma = p_1|\Psi\rangle\langle\Psi| + p_2|\Psi^\perp\rangle\langle\Psi^\perp| + p_3|HV\rangle\langle HV| + p_4|VH\rangle\langle VH|,$$

where the last constraint $p_1 \geq 1 - \epsilon$ follows from the precondition that $\sigma \in \overline{S}$. Then

$$\text{Tr}[\Omega_{\text{opt}} \sigma] = p_1 + \sum_{i=2}^{4} \lambda_ip_i \geq p_1 - \lambda_2 \epsilon,$$

where $\lambda_i$ is the $i$-th eigenvalue of $\Omega_{\text{opt}}$, the first inequality follows from $\lambda_2 \geq \lambda_3 \geq \lambda_4$ and the second inequality follows from $p_1 \geq 1 - \epsilon$.
where the last constraint $p_1 \geq 1 - \epsilon$ follows from the precondition that $\sigma \notin S$. Then

$$\text{Tr} [\Omega_{\text{opt}} \sigma] = p_1 + \lambda_2(\Omega_{\text{opt}}) \sum_{i=2}^4 p_i$$

$$= p_1 + \lambda_2(\Omega_{\text{opt}})(1 - p_1)$$

$$\geq (1 - \epsilon)(1 - \lambda_2(\Omega_{\text{opt}})) + \lambda_2(\Omega_{\text{opt}})$$

$$= 1 - (1 - \lambda_2(\Omega_{\text{opt}}))\epsilon,$$

where the first equality follows from the spectral decomposition of Eq. (S10) and the inequality follows from $p_1 \geq 1 - \epsilon$.

\[\text{(S57)}\]

**Task B** is to distinguish among the following two cases for a given quantum device $D$:

**Case 1:** $\forall i, \sigma_i \in S$. That is, the device always generate states that are sufficient close to the target state $|\Psi\rangle$. If it is the case, we regard the device as “good”.

**Case 2:** $\forall i, \sigma_i \in S$. That is, the device always generate states that are sufficient far from the target state $|\Psi\rangle$. If it is the case, we regard the device as “bad”.

To execute **Task B**, we perform $N$ tests on the outputs of the device with $\Omega_{\text{opt}}$ and record the number of tests that pass the test as $m_{\text{pass}}$. What conclusion can we obtain from the relation between $m_{\text{pass}}$ and $N$? From Lemma 2, we know that if the device is bad (belonging to **Case 2**), $m_{\text{pass}}$ cannot be too large, since the passing probability of each generated state is upper bounded. Conversely, Lemma 3 guarantees that if the device is good (belonging to **Case 1**), $m_{\text{pass}}$ cannot be too small, since the passing probability of each generated state is lower bounded. Let’s then justify this intuition rigorously. We define the binary random variable $X_i$ to represent the event that whether state $\sigma_i$ passes the test or not. If it passes, we set $X_i = 1$; If it fails to pass the test, we set $X_i = 0$. After $N$ tests, we obtain a sequence of independently identically distribution (iid.) random variables $\{X_i\}_{i=1}^N$. Then $m_{\text{pass}} = \sum_{i=1}^N X_i$. Now let’s analyze the expectation of each $X_i$. Let $\mu := 1 - [1 - \lambda_2(\Omega_{\text{opt}})]\epsilon$. If the device were ‘good’ (it belongs to **Case 1**), then Lemma 3 implies that the expectation of $X_i$, denoted as $\mathbb{E}(X_i)$, shall satisfy $\mathbb{E}(X_i) \geq \mu$. The iid. assumption together with the law of large numbers guarantee $m_{\text{pass}} \geq N\mu$, when $N$ is sufficiently large. On the other hand, if the device were ‘bad’ (it belongs to **Case 2**), then Lemma 2 asserts that the expectation of $X_i$ shall satisfy $\mathbb{E}(X_i) \leq \mu$. Again, the iid. assumption together with the law of large numbers guarantee $m_{\text{pass}} \leq N\mu$, when $N$ is sufficiently large. That is to say, we consider two regions regarding the value of $m_{\text{pass}}$, the region that $m_{\text{pass}} \leq N\mu$ and the region that $m_{\text{pass}} \geq N\mu$. We refer to Supplementary Figure 3 for illustration.

**Supplementary Figure 3** The number of copies $m_{\text{pass}}$ that pass the nonadaptive strategy $\Omega_{\text{opt}}$ determines the quality of the device. If $m_{\text{pass}} \leq N\mu$, the device is very likely to be **Case 2 (Small Region)**. Conversely, the device is very likely to be **Case 1 (Large Region)** if $m_{\text{pass}} \geq N\mu$.

**Large Region:** $m_{\text{pass}} \geq N\mu$. In this region, the device belongs to **Case 1** with high probability, as only in this case $m_{\text{pass}}$ can be large. Following the analysis method in refs. S5–S7, we use the Chernoff bound to upper bound the probability that the device belongs to **Case 2** as

$$\delta \equiv e^{-N D(\frac{m_{\text{pass}}}{N} || \mu)},$$

where $D(x||y) := x \log_2 \frac{x}{y} + (1 - x) \log_2 \frac{1-x}{1-y}$ the Kullback-Leibler divergence. That is to say, if the device did belong to **Case 2**, the probability that $m_{\text{pass}}$ is larger than $N\mu$ decays exponentially as $m_{\text{pass}}$ becomes larger. In this case, we reach a conclusion:

*The device belongs to **Case 2** with probability at most $\delta$.*

Equivalently, we conclude that

*The device belongs to **Case 1** with probability at least $1 - \delta$.***
Lemma 5. For arbitrary state $\sigma$, there are exactly $X_m$ that pass the test as $m_{pass}$ becomes smaller. In this case, we reach a conclusion:

The device belongs to Case 1 with probability at most $\delta$.

Equivalently, we conclude that

The device belongs to Case 2 with probability at least $1 - \delta$.

B. Applying the adaptive strategy to Task B

The way that we use the adaptive verification strategy $\Omega^\rightarrow_{opt}$ proposed in Supplementary Note 1B to execute Task B is similar to that discussed in the previous section. The following two lemmas are required, which state that for arbitrary state belonging to $S$, its probability of passing the test $\Omega^\rightarrow_{opt}$ is upper (lower) bounded.

Lemma 4. For arbitrary state $\sigma \in S$, it can pass $\Omega^\rightarrow_{opt}$ with probability no larger than $1 - |1 - \lambda_2(\Omega^\rightarrow_{opt})|\epsilon$, where $\lambda_2(\Omega^\rightarrow_{opt})$ is the second largest eigenvalue of $\Omega^\rightarrow_{opt}$ given in Eq. (S18). The upper bound is achieved by states in $S$ that are exactly $\epsilon$-far from $|\Psi\rangle$.

The proof of Lemma 4 is the same as that of Lemma 2, so we omit the details.

Lemma 5. For arbitrary state $\sigma \in S$, it can pass the adaptive strategy $\Omega^\rightarrow_{opt}$ with probability no smaller than $1 - |1 - \lambda_4(\Omega^\rightarrow_{opt})|\epsilon$, where $\lambda_4(\Omega^\rightarrow_{opt}) = \sin^2 \theta/(1 + \cos^2 \theta)$ is the fourth (also the smallest) eigenvalue of $\Omega^\rightarrow_{opt}$. The lower bound is achieved by states in $S$ that are exactly $\epsilon$-far from $|\Psi\rangle$.

Proof. Following the proof for Lemma 3 and considering the spectral decomposition Eq. (S20) of $\Omega^\rightarrow_{opt}$, we can assume without loss of generality that $\sigma$ has the form

$$\sigma = p_1|\Psi\rangle\langle\Psi| + p_2|\Psi\rangle\langle\Psi| + p_3|VH\rangle + p_4|HV\rangle,$$

where the last constraint $p_1 \geq 1 - \epsilon$ follows from the precondition that $\sigma \in S$. Then

$$\text{Tr} \left[ \Omega^\rightarrow_{opt} \sigma \right] = p_1 + \sum_{i=2}^{4} \lambda_i(\Omega^\rightarrow_{opt}) p_i$$

$$\geq p_1 + \lambda_4(\Omega^\rightarrow_{opt}) \sum_{i=2}^{4} p_i$$

$$= p_1 + \lambda_4(\Omega^\rightarrow_{opt})(1 - p_1)$$

$$\geq (1 - \epsilon)(1 - \lambda_4(\Omega^\rightarrow_{opt})) + \lambda_4(\Omega^\rightarrow_{opt})$$

$$= 1 - (1 - \lambda_4(\Omega^\rightarrow_{opt}))\epsilon,$$

where the first inequality follows from the fact that $\lambda_4(\Omega^\rightarrow_{opt})$ is the smallest eigenvalue of $\Omega^\rightarrow_{opt}$, and the second inequality follows from $p_1 \geq 1 - \epsilon$.

To execute Task B, we perform $N$ tests on the outputs of the device with $\Omega^\rightarrow_{opt}$ and record the number of tests that pass the test as $m_{pass}$. Define the binary random variable $X_i$ to present the event that whether state $\sigma_i$ passes the test or not. If it passes, we set $X_i = 1$; If it fails to pass the test, we set $X_i = 0$. After $N$ tests, we obtain a sequence of iid. random variables $\{X_i\}_{i=1}^{N}$. Then $m_{pass} = \sum_{i=1}^{N} X_i$. We analyze the expectation of each $X_i$. Let $\mu_i := 1 - |1 - \lambda_2(\Omega^\rightarrow_{opt})|\epsilon$ and $\mu_s := 1 - |1 - \lambda_4(\Omega^\rightarrow_{opt})|\epsilon$. If the device were ‘good’ (it belongs to Case 1), then Lemma 5 implies that $\mathbb{E}(X_i) \geq \mu_s$. The iid. assumption together with the law of large numbers guarantee $m_{pass} \geq N\mu_s$, when
$N$ is sufficiently large. That is to say, if in practical it turns out that $m_{\text{pass}} \leq N\mu_s$, then we are pretty sure the device is bad. On the other hand, if the device were 'bad' (it belongs to Case 2), then Lemma 4 asserts that $E(X_i) \leq \mu_l$. Again, the iid. assumption together with the law of large numbers guarantee $m_{\text{pass}} \leq N\mu_l$, when $N$ is sufficiently large. That is, if in practical it turns out that $m_{\text{pass}} \geq N\mu_l$, then we are pretty sure the device is good. Thus, we shall consider two regions regarding the value of $m_{\text{pass}}$: the region that $m_{\text{pass}} \leq N\mu_s$ and the region that $m_{\text{pass}} \geq N\mu_l$. We refer to Supplementary Figure 4 for illustration.

**Supplementary Figure 4** The number of copies $m_{\text{pass}}$ that pass the adaptive strategy $\Omega_{\text{opt}} \rightarrow$ determines the quality of the device. If $m_{\text{pass}} \leq N\mu_s$, the device belongs very likely to Case 2 (Small Region). Conversely, the device belongs with high probability to Case 1 (Large Region) if $m_{\text{pass}} \geq N\mu_l$.

**Large Region:** $m_{\text{pass}} \geq N\mu_l$. In this region, the device belongs to Case 1 with high probability. Following the analysis method in refs. S5–S7, we use the Chernoff bound to upper bound the probability that the device belongs to Case 2 as

$$\delta_l \equiv e^{-N D\left(\frac{m_{\text{pass}}}{N} \parallel \mu_l\right)}.$$  

(S62)

That is to say, if the device did belong to Case 2, the probability that $m_{\text{pass}}$ is larger than $N\mu$ decays exponentially when $m_{\text{pass}}$ becomes large. In this case, we reach a conclusion:

*The device belongs to Case 2 with probability at most $\delta_l$.***

Equivalently, we conclude that

*The device belongs to Case 1 with probability at least $1 - \delta_l$.***

**Small Region:** $m_{\text{pass}} \leq N\mu_s$. The analysis for this region is almost the same as that of the large region. In this region, the device belongs to Case 2 with high probability, as only in this case $m_{\text{pass}}$ can be smaller than $N\mu_s$. Following the analysis method in refs. S5–S7, we use the Chernoff bound to upper bound the probability that the device belongs to Case 1 as

$$\delta_s \equiv e^{-N D\left(\frac{m_{\text{pass}}}{N} \parallel \mu_s\right)},$$  

(S63)

That is to say, if the device did belong to Case 1, the probability that $m_{\text{pass}}$ is less than $N\mu$ decays exponentially as $m_{\text{pass}}$ becomes smaller. In this case, we reach a conclusion:

*The device belongs to Case 1 with probability at most $\delta_s$.***

Equivalently, we conclude that

*The device belongs to Case 2 with probability at least $1 - \delta_s$.***

**Supplementary Note 4:** The experimental results for the three target states

In this section, we give the results for three target states, i.e. k2, Max and HV (see Supplementary Table 1). Note that the failing probability for each measurement is different for the three states S1:

$$k2 : \Delta_\epsilon = \frac{\epsilon}{2 + \sin \theta \cos \theta} \quad (S64a)$$  

$$\text{Max} : \Delta_\epsilon = \frac{2\epsilon}{3} \quad (S64b)$$  

$$\text{HV} : \Delta_\epsilon = \epsilon \quad (S64c)$$  

Here we discuss the results of the nonadaptive strategy. In Eq. (S64), $\epsilon$ can be replaced with $1 - F$, where $F$ is the fidelity for corresponding states. From the experimental data, the stable passing probability $m_{\text{pass}}/N$ can be
obtained, which are $0.9986 \pm 0.0002$, $0.9982 \pm 0.0002$ and $0.9992 \pm 0.0001$, respectively. With $\mu = 1 - \Delta, m_{\text{pass}}/N$, we have an estimation for the fidelities of these three states, which are $F_{k2} = 0.9964 \pm 0.0002$, $F_{\text{Max}} = 0.9973 \pm 0.0002$ and $F_{HV} = 0.9992 \pm 0.0001$. The variation of $\delta$ versus $N$ for the three target states is shown in Supplementary Figure 5 in the log scale. To make the comparison fairly, we take $\eta := |\mu - m_{\text{pass}}/N|$ to be same for the three states. In this condition, the slope of decline (i.e. $g \propto g \cdot N$) is $g_{HV} > g_{k2} > g_{\text{Max}}$, which indicates that the state with a larger passing probability will have a faster decline.

**Supplementary Figure 5** Experimental results on the variation of $\delta$ versus the number of copies for the three target states. Here the difference $\eta := |\mu - m_{\text{pass}}/N|$ is set the same for the three states. The small region (a) decrease faster than the large region (b), which indicates it is easier to verify Case 2 than Case 1. The state with a larger passing probability will have a faster slope of decline. Note that the experimental data symbols shown in the figure looks like lines due to the dense data points.

In Supplementary Figure 6, we present the results of $\epsilon$ versus $N$. We can see that $\epsilon$ first decreases linearly with $N$ and then approaches a asymptote in the log-log scale. The asymptotic values for these three states are the infidelities calculated from Eq. (S64), which are $0.0036 \pm 0.0002$, $0.0027 \pm 0.0002$ and $0.0008 \pm 0.0001$ for these three states. The linear scaling region is fitted with slopes of $-0.88 \pm 0.03$, $-0.87 \pm 0.10$ and $-0.99 \pm 0.09$ for $k2$, Max and HV states, respectively. The error bars are obtained by fitting different groups of $\epsilon$ versus $N$ when considering the errors of $\epsilon$ (shown in Fig. 5 of the main text).

**Supplementary Note 5: Explanations for the results in Task A**

In this section, we explain on the difference of experimental and theoretical improvements for the number of measurements required in Task A of main text. Consider a general state $\rho$ produced in the experiment that is exactly $\epsilon$-far
from $|\Psi\rangle$, which can be diagonalized in the bases \{ $|\Psi\rangle$, $|\Psi^\perp\rangle$, $|HV\rangle$, $|VH\rangle$ \} as:

$$
\rho = (1 - \epsilon)|\Psi\rangle\langle\Psi| + p_2|\Psi^\perp\rangle\langle\Psi^\perp| + p_3|HV\rangle\langleHV| + p_4|VH\rangle\langleVH|,
$$

(S65)

where the normalization condition requires $p_2 + p_3 + p_4 = \epsilon$. Given the spectral decomposition of $\Omega_{\text{opt}}$ in Eq. (S10), the passing probability of $\rho$ that passes the nonadaptive strategy $\Omega_{\text{opt}}$ can be expressed as

$$
\text{Tr} [\Omega_{\text{opt}}\rho] = 1 - \epsilon + \lambda_2(\Omega_{\text{opt}})(p_2 + p_3 + p_4)
= 1 - [1 - \lambda_2(\Omega_{\text{opt}})]\epsilon
= 1 - \frac{1}{2 + \sin \theta \cos \theta} \epsilon
= 1 - \Delta_\epsilon.
$$

We can see that the passing probability of adaptive is not only dependent on the $\epsilon$, but also dependent on $p_4$. Note that the number of measurements required to obtain a $1 - \delta$ confidence and infidelity $\epsilon$ is given by $n \propto \frac{1}{\Delta_\epsilon} \ln \frac{1}{\delta}$.

Supplementary Note 6: Comparison of nonadaptive and adaptive in Task B

In Supplementary Figure 7, we give the variation of experimental passing probability $m_{\text{pass}}/N$ versus the number of copies. For clearness, we also plot the expected passing probability $\mu = 1 - \Delta_\epsilon$ which is chosen by the verifier. The $\epsilon_{\text{min}}$ is adopted for Case 2 and the $\epsilon_{\text{max}}$ is adopted for Case 1. With enough number of copies of quantum states, the passing probability $m_{\text{pass}}/N$ will reach a stable value. It can be seen that the experimental passing probability is smaller than the expectation for Case 2 while it is larger for Case 1. The stable passing probability of adaptive 0.9914±0.0005 is smaller than the nonadaptive strategy 0.9986±0.0002 due to their different infidelities.

In the main text, we see that the scaling behaviour of parameters $\delta$ and $\epsilon$ versus number of copies $N$ is different for nonadaptive and adaptive strategies in Task B. Here, we give some analyses to explain.

For the variation of $\delta$ versus $N$, the speed of descent is determined by the difference between experimental passing probability $m_{\text{pass}}/N$ and expected passing probability $\mu = 1 - \Delta_\epsilon$. Because the nonadaptive and adaptive strategy have different $m_{\text{pass}}/N$ and $\mu$, the behaviour of $\delta$ versus $N$ is also different. To have a comprehensive understanding, we consider two situations. First, we assume they have the same $m_{\text{pass}}/N$ and the same $\epsilon$. The expected passing probabilities are $\mu_{\text{Non}} = 1 - \epsilon/(2 + \sin \theta \cos \theta)$ and $\mu_{\text{Adp}} = 1 - \epsilon/(2 - \sin^2 \theta)$ for nonadaptive and adaptive strategies, respectively. We can see that $\mu_{\text{Non}}$ is larger than $\mu_{\text{Adp}}$ under the same $\epsilon$. In Supplementary Figure 8, we plot the variation of $\delta$ versus $\mu$ using equation $\delta = \epsilon^{-N D(m_{\text{pass}}/N||\mu)}$, irrespective of the specific strategy. Therefore, the
curves of nonadaptive and adaptive strategy coincide with each other in this situation. The expected probabilities $\mu_{\text{Non}}$ and $\mu_{\text{Adp}}$ are located at the different positions of the horizontal axis, as shown in the inset of Supplementary Figure 8. Note that $\mu \leq m_{\text{pass}}/N$ is for Case 1 and $\mu \geq m_{\text{pass}}/N$ is for Case 2 region when we choose different $\mu$, i.e. $\epsilon = 1 - \Delta \epsilon_{\text{min}}/2$. From the figure we can see that the larger the difference between $\mu$ and $m_{\text{pass}}/N$, the smaller of the $\delta$. This indicates that $\delta$ decreases more quickly for adaptive strategy in the Case 1 region, whereas it decreases more quickly for nonadaptive strategy in the Case 2 region.

Supplementary Figure 8 The parameter $\delta$ changes with the anticipating passing probability $\mu$ when adopting the same $m_{\text{pass}}/N$. With the same $\epsilon$, adaptive will always have a smaller $\mu$ than nonadaptive. Therefore, the adaptive will decrease faster than nonadaptive due to a larger difference in the $\mu \leq m_{\text{pass}}/N$ region. In the $\mu \geq m_{\text{pass}}/N$ region, the trend is opposite.

Second, we consider the situation where the nonadaptive and adaptive strategies have different passing probabilities $m_{\text{pass}}/N$, as in our experiment. For comparison, we adopt $m_{\text{pass}}/N = 0.9986$ for nonadaptive and $m_{\text{pass}}/N = 0.9914$ for adaptive based on our experimental data. The variation of $\delta$ versus $\mu$ is shown in Supplementary Figure 9. We can see that the comparison of nonadaptive and adaptive depends on the choice of $\mu$. In the figure, we label the crosspoint of the two curves as $\mu_0$. If the $\epsilon$ is chosen such that both $\mu_{\text{Non}}$ and $\mu_{\text{Adp}}$ are smaller than $\mu_0$, the nonadaptive will drop faster than adaptive because nonadaptive has a smaller $\delta$. On the contrary, if both $\mu_{\text{Non}}$ and $\mu_{\text{Adp}}$ are larger than $\mu_0$, the adaptive will drop faster than nonadaptive due to the smaller $\delta$. If $\mu_{\text{Non}}$ and $\mu_{\text{Adp}}$ are located at two sides of $\mu_0$, the one which has a smaller $\delta$ in Supplementary Figure 9 will have a faster decline.
**Supplementary Figure 9** The parameter $\delta$ changes with the anticipating passing probability $\mu$ for different $m_{\text{pass}}/N$. In this situation, the one which adopts $\mu$ that results in a smaller $\delta$ will have a faster decline.

**Supplementary Figure 10** The slope of linear scaling versus the passing probability. The slope decreases with the increase of the passing probability. Our experimental fitting slopes for the adaptive and nonadaptive strategies in the linear region is shown as the black stars. The slope approaches the optimal -1 when the passing probability are close 1. Due to the smaller passing probability of adaptive compared with nonadaptive, the absolute value of slope for adaptive strategy is also smaller. When the passing probability is large enough, the two strategies almost have an equal linear scaling slope at the same passing probability.

For $\epsilon$ versus the number of copies $N$, the scaling is determined by the passing probability $m_{\text{pass}}/N$, i.e. the fidelity of our generated states. The larger the passing probability, the faster the infidelity parameter $\epsilon$ decreases with the number of copies. For both nonadaptive and adaptive, $\epsilon$ will finally approach a asymptotic value. In Supplementary Figure 10, we plot the fitting slope of the linear scaling region versus the passing probability for both the nonadaptive and adaptive strategies. We can see that the adaptive strategy will have an advantage compared with nonadaptive strategy at a small passing probability. However, the slope tends to the optimal value -1 when the passing probability is large enough. This indicates that the optimal scaling can only be obtained at a high fidelity for the generated states. In the region of high fidelity, there is minor differences for the nonadaptive and adaptive strategies if we obtain a same passing probability. However, the adaptive strategy has a smaller passing probability than the nonadaptive strategy for our experimental data. Therefore, the descent speed of adaptive strategy is slower than the nonadaptive strategy. This can be seen quantitatively from their fitting slopes $-0.78\pm0.07$ (Adp) and $-0.88\pm0.03$ (Non), shown as
stars in Supplementary Figure 10.

Supplementary References

[S1] Pallister, S., Linden, N. & Montanaro, A. Optimal verification of entangled states with local measurements. *Phys. Rev. Lett.* **120**, 170502 (2018).

[S2] Wang, K. & Hayashi, M. Optimal verification of two-qubit pure states. *Phys. Rev. A* **100**, 032315 (2019).

[S3] During the optimization, we use the program offered by Kwiat Quantum Information Group. The code can be downloaded from their website. Last accessed: 2018-05-10.

[S4] Yu, X.-D., Shang, J. & Gühne, O. Optimal verification of general bipartite pure states. *npj Quantum Information* **5**, 112 (2019).

[S5] Dimić, A. & Dakić, B. Single-copy entanglement detection. *npj Quantum Information* **4**, 11 (2018).

[S6] Saggio, V. et al. Experimental few-copy multipartite entanglement detection. *Nature Physics* **15**, 935–940 (2019).

[S7] Zhang, W.-H. et al. Experimental optimal verification of entangled states using local measurements. *arXiv:1905.12175* (2019).