Heisenberg-scaling measurement of the single-photon Kerr non-linearity using mixed states

Geng Chen,1,2 Nati Aharon,3 Yong-Nan Sun,1,2 Zi-Huai Zhang,1,2 Wen-Hao Zhang,1,2 De-Yong He,1,2 Jian-Shun Tang,1,2 Xiao-Ye Xu,1,2 Yaron Kedem,4 Chuan-Feng Li,1,2 and Guang-Can Guo1,2

1CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, People’s Republic of China
2Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, People’s Republic of China
3Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Givat Ram, Israel
4Department of Physics, Stockholm University, AlbaNova University Center, 106 91 Stockholm, Sweden

Improving the precision of measurements is a significant scientific challenge. The challenge is twofold: first, overcoming noise that limits the precision given a fixed amount of a resource, $N$, and second, improving the scaling of precision over the standard quantum limit (SQL), $1/\sqrt{N}$, and ultimately reaching a Heisenberg scaling (HS), $1/N$. Here we present and experimentally implement a new scheme for precision measurements. Our scheme is based on a probe in a mixed state with a large uncertainty, combined with a post-selection of an additional pure system, such that the precision of the estimated coupling strength between the probe and the system is enhanced. We performed a measurement of a single photon’s Kerr non-linearity with an HS, where an ultra-small Kerr phase of $\simeq 6 \times 10^{-6} \text{rad}$ was observed with unprecedented precision of $\simeq 3.6 \times 10^{-10} \text{rad}$. Moreover, our scheme utilizes an imaginary weak-value, the Kerr non-linearity results in a shift of the mean photon number of the probe, and hence, the scheme is robust to noise originating from the self-phase modulation.

Consider a physical process that is described by an interaction Hamiltonian $gH$, which depends linearly on a small parameter $g$ that we want to estimate. The precision of this estimation is ultimately limited by the Cramér-Rao bound, which implies that

$$\Delta g \geq \Delta g_{\text{min}} = \frac{1}{\sqrt{F(\rho)V}},$$

where $F(\rho)$ is the quantum Fisher information (QFI) of the final state, $\rho$, and $V$ is the number of times $H$ is used. For pure states the QFI is equal to $\Delta H^2$, where $\Delta H = \sqrt{\langle H^2 \rangle - \langle H \rangle^2}$ is the standard deviation of $H$ with respect to the initial state. Hence, by preparing an initial pure state with a large $\Delta H$, one may improve the precision. The interaction can involve a large number, $N$, of subsystems, and in case that there are no interactions between the subsystems, $H = \sum_{i=1}^{N} H_i$. The scaling of the precision with respect to $N$ is of special importance. If the initial state is such that the subsystems are uncorrelated, $\langle H_i H_j \rangle - \langle H_i \rangle \langle H_j \rangle \propto \delta_{ij}$, then $\Delta H \propto \sqrt{N}$ and the precision follows the SQL, in general. However, when the subsystems are initially entangled, it is possible to have $\Delta H \propto N$, which yields an HS if $\Delta H \propto N$. Thus, a significant amount of effort was put into generating highly entangled states, such as NOON states or squeezed states.

Generally, for mixed states the above reasoning does not hold; an initial mixed state with $\Delta H \propto N$, does not yield an HS. It has been shown, however, that for non-linear interactions the precision can be improved beyond the SQL with mixed states. Our scheme is directly focused on maximizing $\Delta H$ by introducing externally induced fluctuations to the initial probe state. Remarkably, for the metrological task of estimating a coupling strength between a probe and a pure quantum system, our scheme enables to utilize these classical fluctuations, i.e., mixed states, in order to improve the precision. In particular, in the measurement of a single photon’s Kerr non-linearity an HS is achieved. In order to see how to utilize these fluctuations, we use the formalism of weak measurements. Consider an interaction Hamiltonian

$$H = f(t)\hat{P}C,$$

where $\hat{P}$, which is related to a probe, and $\hat{C}$, which is related to a system, are both Hermitian operators and $f(t)$ is a coupling function with a finite support that satisfies $g = \int f(t)dt$. If the system is prepared in a state $|\psi\rangle$ before the interaction and post-selected later to a state $|\phi\rangle$, then $\langle \hat{P} \rangle$ will be modified according to $\langle \hat{P} \rangle \rightarrow \langle \hat{P} \rangle + \delta \hat{P}$, with

$$\delta \hat{P} = 2g (\Delta P)^2 \text{Im} C_w,$$

where $C_w = \langle \phi | \hat{C} | \psi \rangle$ is the weak value of $\hat{C}$, and $\Delta P$ is the standard deviation of $\hat{P}$. Hence, a measurement of $\langle \hat{P} \rangle$ yields a precision of $\Delta \hat{P} \propto \Delta \hat{P}^{-1}$. As we noted before, when $\hat{P}$ pertains to $N$ uncorrelated systems where $\Delta P \propto \sqrt{N}$, e.g., the coherent state, the SQL still applies. However, it was shown that $\delta \hat{P}$ holds even when $\Delta \hat{P}$ is due to classical fluctuations, and these fluctuations can improve the measurement precision by increasing $\Delta \hat{P}$. Our method is based on this idea, and by taking the limit of the largest possible fluctuations $\Delta P \sim \langle \hat{P} \rangle$, we obtain precision with an HS. We experimentally demonstrate our scheme by measuring a single photon’s Kerr non-linearity, achieving a robust HS that results in an improved precision compared to the standard method.

The general idea of our scheme is depicted in Fig. 1 by illustrating it in an optical phase diagram and comparing it to the known methods using coherent and squeezed states. The intuition arising from this picture, is that “stretching” the state can, in a particular case, have a similar effect as squeezing. Increasing the total uncertainty is possible when external noise/modulations are added.

Consider the setup shown in Fig. 1(c). The interaction is described by $U = e^{-i\hat{g}C\hat{t}}$, where $\hat{n}$ is the photon number operator on the probe and $\hat{C}$ is the “which path” operator acting on a

\[ \text{CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, People’s Republic of China} \]
\[ \text{Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, People’s Republic of China} \]
\[ \text{Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Givat Ram, Israel} \]
\[ \text{Department of Physics, Stockholm University, AlbaNova University Center, 106 91 Stockholm, Sweden} \]
FIG. 1. **Strategies for precision measurements.** (a) The standard method: The uncertainty of a coherent state $|\alpha\rangle$ implies a finite support on the phase diagram with a characteristic extent of $\Delta \text{Re}[\alpha] = \Delta \text{Im}[\alpha] = 1/2$, represented by a fading disc. An accumulation of a phase $g$ shifts the initial state, shown in red, to the final one, shown in blue, by a distance of $g |\alpha\rangle \propto \sqrt{N}$ and $N$ is the mean photon number of the probe. The precision scaling can be understood by considering the ratio between the uncertainty and the shift, which in this case is $\propto 1/\sqrt{N}$, corresponding to the SQL. (b) Squeezed states: reducing the uncertainty of some quadrature $\Delta X$ can improve the precision. However, due to the fundamental relation this implies an increase in the uncertainty of the conjugate quadrature $\Delta P$, which should be smaller than the average distance from the origin $\sqrt{N}$. When $\Delta X \propto \Delta P^{-1} \propto 1/\sqrt{N}$, the Heisenberg scaling is obtained. (c) The scheme: a single-photon goes through an interferometer, where in one arm it interacts with a coherent state and a phase is acquired, while in the other nothing happens. At the exit port the two paths interfere so the probability of the photon coming out there depends on the phase it acquired by the interaction, and thus, also on the number of photons in the probe. Post-selecting the probe pulses accordingly induces a shift in the average photon number, from which the interaction strength can be estimated. (d) The new method: a statistical ensemble of $|\alpha\rangle$ (only a few are drawn), showing the two paths for each member: Blue with the phase and red without. Due to the probability of postselection, $|\alpha\rangle \propto \sqrt{n}$, where $n$ can be fluctuated to be $\propto N$. The ratio between the uncertainty and the shift is $\propto 1/N$, resulting in the Heisenberg scaling.

Let us apply this method in the task of measuring the Kerr non-linearity of a single photon. A strong pulse (probe) and a single photon (system) overlap inside a fiber where the dependence of the refractive index on the intensity of light induces both a self-phase modulation (SPM) and a cross-phase modulation (XPM). The effective Hamiltonian is $[31] H = \tilde{g}_S \hat{n}^2 + \tilde{g} \hat{C} \hat{n}$, where $\hat{n}$ now refers to the photon number in the probe, $\hat{C}$ is the photon number in the system, and $\tilde{g}_S (\tilde{g})$ is the coupling due to the SPM (XPM). Integrating along the fiber yields the coefficient $g_S (g)$ for the SPM (XPM), such that the evolution is given by $U = e^{-ig_S \hat{n}^2 - ig \hat{n}}$. The XPM represents an interaction involving a single photon, and thus, measuring $g$ is highly important for many applications [32, 33]. An experiment to achieve this was recently performed by Matsuda et al. [25], using the standard approach as described above. The main limitation in their setup came from the additional noise introduced by the SPM, which is dominant when $N g_S \sim 1$. Using our scheme, as we show below, the SPM is insignificant since the intensity is measured instead of the phase.

We now show, in detail, how to measure $g$ using our method and analyze the resulting precision. The system photon is sent into an interferometer, with one arm containing the fiber, such that its initial state is $|\psi\rangle = (|1\rangle + |0\rangle) / \sqrt{2}$ where $|1\rangle$ and $|0\rangle$ are eigenstates of $\hat{C}$ with eigenvalues 0 and 1, respectively. The probe is in a coherent state $|\alpha\rangle$, but by modulating the particular form of the distribution (see supplementary information section II).

If the system is prepared in a state $|\psi\rangle$ and post-selected later to $|\phi\rangle$, one can approximate the impact of the interaction as $\langle \phi | e^{-igS \hat{n}^2} | \psi \rangle \approx e^{-igS}_\psi$. A coherent state would transform according to $|\alpha\rangle \rightarrow e^{ig\alpha} |\alpha\rangle$, which implies a tangential shift of $g |\alpha\rangle \text{Re} C_w$ and a radial shift of $g |\alpha\rangle \text{Im} C_w$, yielding a precision that is still limited by the SQL. In our method, we use a statistical mixture of number states or an ensemble of coherent states, as illustrated in Fig. 1(d). The probability of post-selection when the probe has $n$ photons is given by (see supplementary information section II)

$$p_{\phi|n} = |\langle \phi | \psi \rangle|^2 (1 + 2g \text{Im} C_w n + o(n)^2).$$

States with a higher photon number entail larger (smaller) probability if $\text{Im} C_w > 0$ ($\text{Im} C_w < 0$); therefore the photon-number distribution is shifted. This change in the average photon number is given by $\delta n = 2g (\Delta n)^2 \text{Im} C_w$, where $\Delta n$ is the standard deviation of the initial distribution. Obtaining $\delta n$ experimentally entails a measurement error $\propto \Delta n$, and thus, the estimation $g = \delta n / (2 (\Delta n)^2 \text{Im} C_w)$ yields a precision $\Delta g \propto 1/\Delta n$ (other contributions to the estimation error are $O(\delta n/\Delta n) \ll 1$). A coherent state $|\alpha\rangle$ has an uncertainty of $\Delta n = |\alpha| = \sqrt{N}$, for which this method yields the SQL. However, one can produce a distribution with a large deviation [30], $\Delta n \propto N$, where $N$ is the average photon number of the distribution, and achieve an HS. Note that the precise form of the distribution is insignificant for this result; it depends only on the mean value and the variance of the distribution (see supplementary information section II).

Let us consider a superposition of the two arm, i.e., not in an eigenstate of $\hat{C}$, the QFI of the joint system-probe state after the interaction is $\propto N^2$, where $N = \langle \hat{n} \rangle$ [27]. However, using a probe that is initially in a coherent state, a measurement of only its phase, or any other quadrature, cannot extract this information, since Heisenberg scaling arises only in the post-selection process itself [24, 27, 28]. In this experiment, we measured the photon number, for a given post-selection. When the probe is initially in a statistical ensemble with a wide distribution, i.e., the standard deviation is proportional to $N$, the classic Fisher information [29] for this choice is also $\propto N^2$, regardless of the single photon going through the interferometer. If the photon is in a superposition of the two arm, i.e., not in an eigenstate of $\hat{C}$, the QFI of the joint system-probe state after the interaction is $\propto N^2$, where $N = \langle \hat{n} \rangle$ [27]. However, using a probe that is initially in a coherent state, a measurement of only its phase, or any other quadrature, cannot extract this information, since Heisenberg scaling arises only in the post-selection process itself [24, 27, 28]. In this experiment, we measured the photon number, for a given post-selection. When the probe is initially in a statistical ensemble with a wide distribution, i.e., the standard deviation is proportional to $N$, the classic Fisher information [29] for this choice is also $\propto N^2$, regardless of the
power of the laser we obtain a distribution of \(\alpha\), which can be written as a mixed state, \(\rho_\alpha = \sum \alpha \rho_\alpha \langle \alpha | \langle \alpha \rangle\), where \(\rho_\alpha\) is the probability of having a coherent state \(|\alpha\rangle\). After the probe goes through the fiber, we measure its average photon number. However, only the trials when the system photon is found in a specific exit port are taken into account; we post-select the state of the system as \(|\psi\rangle = (|1\rangle - e^{-i\epsilon}|0\rangle)/\sqrt{2}\), with the post-selection parameter \(\epsilon \ll 1\), which is set by tuning the interferometer. The result of the measurement on the probe is given by

\[
\langle \hat{n} \rangle = \frac{Tr[\hat{n} \rho_i \hat{p}^W \hat{U}^\dagger \rho^W_i \hat{U} \rho_i]}{Tr[\hat{U} \rho_i \hat{p}^W \hat{U}^\dagger \rho^W_i \hat{U} \rho_i]},
\]

where \(\rho^W_i = |\psi\rangle \langle \psi|\) and \(\rho_i = |\phi\rangle \langle \phi|\) are the pre- and post-selected state, respectively. \(\rho^W_i\) is inserted to represent the post-selection and this is also the reason for the normalization denominator. Since \([\hat{U}, \hat{n}] = 0\), only the diagonal element \(\rho^W_i \hat{p}^{n_i} \rightarrow (n_i \rho^W_i)\) are significant, and we can replace the trace over the probe states with a sum over the photon number \(Tr_{\rho_i} \rightarrow \sum_i n_i \rho_i^{n_i} \hat{p}^{n_i}\) while replacing \(\hat{n} \rightarrow n\). The trace over the probe states can be approximated as

\[
Tr[\hat{U} \rho_i \hat{p}^W \hat{U}^\dagger \rho^W_i \hat{U} \rho_i] \approx \epsilon^2 (1 + 2n_\epsilon^2)\]

for \(n_\epsilon^2 \ll 1\). Therefore, the change in the average photon number is given by (see supplementary information section II)

\[
\delta n \simeq \frac{\sum_i n_i \rho_i^{n_i} \rho_i^{n_i} (1 + 2n_\epsilon^2)}{\sum_i n_i \rho_i^{n_i} (1 + 2n_\epsilon^2)} - N \simeq 2 \frac{\hat{g}}{\epsilon} (\Delta n)^2.
\]

Since \(C_\nu \simeq \frac{i}{\epsilon}\), Eq. 5 agrees with Eq. 2. In case that \(\Delta n \approx N\), we obtain an HS.

Moreover, due to the usage of an imaginary weak value, the interaction results in a shift of the average photon number rather than a phase shift, our scheme is robust to phase noise, and in particular, the SPM part is completely canceled.

Our method, and weak measurements in general, requires a post-selection, and for a large weak value, the post-selection is rare. This can diminish the precision due to a decrease in the number of successful post-selecting events \(n \sim \nu |\langle \phi |\psi\rangle|^2\). On the other hand, by calculating the Fisher information directly from Eq. 3, one obtains another amplification factor of \(\epsilon^{-2} \sim |\langle \phi |\psi\rangle|^{-2}\); therefore, when using the Cramér-Rao bound Eq. 1, the dependence on \(|\langle \phi |\psi\rangle|\) cancels.

The experimental setup is shown in Fig. 2. A photon pair is generated by spontaneous parametric down conversion. One photon is used for heralding and the other enters a polarization Sagnac interferometer (PSI), which contains a photonic crystal fiber (PCF). The single photon’s polarization is set as \(|V\rangle + |H\rangle)/\sqrt{2}\) (\(V\) and \(H\) represent the vertical and horizontal polarization respectively). After entering the PSI, the photon is in an equal superposition of clockwise and counter-clockwise propagation. Only the counter-clockwise component can interact with the probe pulse; hence, the system becomes \(|\psi\rangle = (|1\rangle_{\psi} + |0\rangle_{\phi})/\sqrt{2}\), where \(|0,1\rangle\) represents the interacting photon number. After the PSI the system is post-selected using its polarization. Faraday units cause the two components to have the same polarization inside the PCF. Preparation of the probe, which is a strong pulse, involves (i) modulating its intensity using an acoustic optical modulator (AOM), (ii) introducing delay using a translatable mirror and (iii) filtering the spectrum to prevent an overlap with the spectrum of the single photon. The probe then enters the PCF through a polarized beam splitter (PBS), where it overlaps with one component of the single photon and the interaction takes place. Upon exiting the PCF, through another PBS, the intensity of the probe is measured, which depends on the detection of both the heralding photon and the post-selected photon from the PSI. Separating the single photon from the strong pulse after the interaction is performed using both the polarization and spectrum degrees of freedom (see the Methods for more details).

We start by demonstrating the validity of Eq. 5 in our system by modifying the quantities on the right side: the interaction strength \(g\), the standard deviation \(\Delta \nu\) and the weak value \(i/\epsilon\), and by measuring \(\delta n\). Tuning \(g\) is performed by varying the temporal overlap between the probe and the single photon.

The standard deviation is controlled by changing the modulation amplitude \(D\) in the intensity of the probe (see the Methods for details). \(\epsilon\) is set by choosing the post-selected polarization state of the single photon exiting the PSI. In Fig. 3 we plot the normalized change in the photon number \(\delta \hat{n} = \delta n/N\) in a number of ways. The error bars are shown as the uncertainty.
in $\delta n$, which is written as $\sigma / \sqrt{N}$. Here, $\sigma$ is the standard deviation of measured $\delta n$ and $N$ is number of recorded probe pulses by FHO. The results demonstrate the ability to detect the interaction of the probe with a single photon.

We now experimentally demonstrate the precision of our method, and, in particular, we show how the precision scales with the average photon number of the probe. Theoretically, the precision can be obtained from Eq. 5, for example, by calculating the Fisher information (see supplementary information section I and II). Nonetheless, it is important to show that the precision can be reached in practice, for large values of the photon number, and that the method is indeed advantageous compared to the alternative methods. In order to quantify the precision, we study the dependence of the measured quantity on the varying estimated parameter $g$. Each value of $g$ is calibrated from Eq. 5, with $N = 9 \times 10^4$, $\varepsilon = 0.1$ and $\Delta n \approx 0.5N$, when $\nu \approx 9 \times 10^{-8}$. Taking into account of the uncertainty in the measurement of the probe, we can obtain the practical precision. The results, shown in Fig. 4 demonstrate an HS, up to values of $N \approx 10^5$. The ultimate precision of $\Delta g \approx 3.6 \times 10^{-10} \text{rad}$ is an improvement on a recent result for the same task [25]. It can be seen that the achieved precisions in our experiment are close to the bound (red line) set by the maximum QFI for mixed states, which can be calculated from Eq. S3 in supplementary information.

The method we presented requires $Ng \ll 1$, which implies that it cannot be a single-shot measurement. The information regarding $g$ can only be gathered from an ensemble that is large enough for the statistical distribution of the initial probe state to be meaningful. Nonetheless, at the most interesting scenario $g \to 0$, i.e., detecting the utmost miniscule effects, the scaling implies that when $g \to g = a^{-1}g$, modifying $N \to N = aN$ would maintain the same relative precision, for any number $a$.

In summary, in this work a new scheme for the metrological task of estimating a weak coupling strength between a pure quantum system and a probe was presented. We theoretically and experimentally demonstrated that mixed probe state, combined with a post-selection of the pure quantum system, can be utilized to improve the precision. Specifically, the extracted FI is close to and scales as the QFI. We performed a measurement of the Kerr non-linearity of a single photon with an HS. Moreover, because an imaginary weak value was employed, our measurement was robust to phase noise, and in
particular, to SPM noise. This enabled us to reach an unprecedented precision of $\approx 3.6 \times 10^{-10} \text{rad}$ in a measurement of an ultra-small Kerr phase of $\approx 6 \times 10^{-8} \text{rad}$. Enhancing the precision with weak measurement has been theoretically investigated in the context of metrology.\cite{28,37,39}, and some other works questioned this advantage considering the discarded resources.\cite{27,34}. In our scheme, the mixed states increase the variance of the Hamiltonian, and weak measurements enable the increased variance to improve the precision. This new technique further develops the theoretical framework of weak measurement. The maximal possible magnitude of fluctuations is limited by the experimental resources, which, in our case, is the photon number. The precision scales inversely with the magnitude of fluctuations. Thus, when the resource is scaled up, the precision improves towards the QFI limit; in our case reaching an HS. The fact that our scheme is based on the utilization of mixed states enables its practical scalability (up to the limits of the scheme). Hence, our work paves a new route for precision measurements, which can significantly modify the vast amount of effort devoted to this task.

**METHODS**

**Preparation of Heralded Single Photons (System) and Strong Pulses (Probe).** Single photons are generated by a non-degenerate spontaneous parametric down conversion process (SPDC). At first, 130 fs laser pulses centered at approximately 800 nm are up-converted to 400 nm by a second harmonic generation (SHG) process in a β-barium borate (BBO) crystal. Afterwards, a second BBO crystal is pumped by the 400 nm pulses to generate down-converted photon pairs. The cut angle of BBO crystal is designed to generate collinear 785 nm and 815 nm photon pairs. The photons propagate collinearly and are then separated by a dichroic mirror. The 785 nm photon is reflected into the interferometer as a system pulse while the 815 nm photon is transmitted and then detected by the first single photon detector to herald the 785 nm photons. The residual 800 nm laser pulse after the SHG process is attenuated as a probe pulse. A feedback control of the probe pulse is realized by a half wave plate (HWP) mounted in a motorized rotation stage, so that the power of the probe pulse is well stabilized. Before the probe pulse is coupled into the PCF, its amplitude, spectrum and time domain are tuned. The amplitude modulation is performed through an AOM placed at the confocal point of a doublet lens. The AOM is driven by an arbitrary wave generator (Tektronics AWG 3252). The driving frequency is 200 MHz to maximize the diffraction efficiency. The specific form of the driving function is not important and only the standard deviation affects. For convenience we apply a \textit{sine} type modulation on the driver described as $V = V_0(1 - D \sin(\omega t))$, the photon number fluctuates around the mean $N$ with a standard deviation decided by $D$. $V_0$ is selected where the diffraction efficiency is approximately half of the maximum and $N$ is determined by the incident power on AOM. The modulation frequency $\omega$ is fixed at 1KHz. The modulation depth $D$ can be varied from 0 to 1, as a result, $\Delta n$ can be tuned to expected values according to the records from FHO. The first diffraction order is isolated by a pin-hole and delayed to overlap with pump photons inside the PCF. This synchronization is realized by a silver mirror on a manual linear translation stage with the precision of $\approx 10$ femtosecond. The first spectrum filter (SF1) is an optical 4-f system including two transmitting gratings (1200 Grooves/mm) and a pair of lenses (300 mm focus length). By aligning a slit on the confocal plane, short wavelengths below 795 nm are filtered. Consequently the system and probe pulses can be separated in the spectrum, which is essential when implementing post-selection on system photons.

**Interaction in PSI.** The initial state of the system photons is prepared as $|\psi\rangle = (|V\rangle + |H\rangle)/\sqrt{2}$, where $H$ and $V$ represent the horizontally and vertically polarized components, respectively. These two components counter-propagate through the PSI, which contains an 8 m long PCF (NL-2.4-800, Blaze Photonics). The incident light is collected into the PCF by two Faraday units, each consisting of a 45° Faraday rotator and a HWP, cause the two components to have the same linear polarization in the PCF. Vertically polarized system photons are synchronized to overlap with the probe pulses. Two internal polarized beam splitters are used to allow the system photons to enter and exit the PCF.

**Detection apparatus.** After the probe pulses are separated from the system photons, they shine on a low-noise amplified photon detector and the waveform is sampled by a 12-bit (4096 level) full HD oscilloscope working in the external trigger mode. The system photon exits the PSI from another port and is then post-selected by a polarizer. The post-selection state $|\phi\rangle = (|V\rangle - e^{-i\theta}|H\rangle)/\sqrt{2}$ is set to be nearly orthogonal to $|\psi\rangle$. The second spectrum (SF2) filter contains a 4-f system similar to the first one, but here, only photons with wavelengths below 790 nm can pass. As a result, probe photons leaking out of the PSI are filtered. A subsequent 10 nm band pass filter centered at 785 nm and a short-wavelength pass filter cutoff at 790 nm reinforce this filtering. Coincidence signals are used to trigger the FHO and post-select the probe pulses. The final recording rate of post-selected probe pulses is mainly determined by the value of $\epsilon$. When $\epsilon$ equals 0.1, data were recorded for a total of 6 hours, and $\approx 220$ K probe pulses waveforms are recorded. The value of $n$ is given by the root mean square (RMS) of the recorded waveform. The uncertainty $\sigma/\sqrt{V}$ is also estimated as the standard error of these RMS values.

*email:yaron.kedem@fysik.su.se
email:cfli@ustc.edu.cn

[1] Helstrom, C. W. Quantum detection and estimation theory
Phys. Rev. Lett. 114, 210801 (2015)

Choi, Y. W., Lim, H. T., Ra, Y. S., and Kim, Y. H. assisted on operating the AWG and FHO. G.C., Y.K. and N.A. proposed the framework of the experiment and made the calculations. G.C. planned and designed the experiment. G.C. carried out the experiment assisted by Y.-N.S., X.-Y. X., Z.-H.Z. and W.-H.Z., whereas J.-S.T. designed the computer programs and D.-Y.H. assisted on operating the AWG and FHO. G.C., Y.K. and N.A analyzed the experimental results and wrote the manuscript. G.-C.G. and C.-F.L. supervised the project. All authors discussed the experimental procedures and results.

Acknowledgments This work was supported by the National Key Research and Development Program of China (No. 2017YFA0304100), the National Natural Science Foundation of China (Nos. 61327901, 61490711, 11654002, 60778057).
11325419, 61308010, 91536219), Key Research Program of Frontier Sciences, CAS (No. QYZDY-SSW-SLH003), the Fundamental Research Funds for the Central Universities (No. WK2470000026). N.A acknowledges the support of the Israel Science Foundation(grant no. 039-8823) and EU Project DI-ADEMS.