

Timing jitter induced by stochastic baseline fluctuations in high-count-rate superconducting nanowire single-photon detectors

Dianpeng Wang^{1,2,3} You Xiao^{1,2} Jiamin Xiong^{1,2} Chenrui Wang^{1,2,3} Zhen Wan^{1,2} Hongxin Xu^{1,2,3}
Chaomeng Ding^{1,2,3} Jia Huang^{1,2} Lixing You^{1,2,3} and Hao Li^{1,2,4,*}

¹Shanghai Key Laboratory of Superconductor Integrated Circuit Technology,
Shanghai Institute of Microsystem and Information Technology, *Chinese Academy of Sciences*,
Shanghai 200050, China

²National Key Laboratory of Materials for Integrated Circuits,
Shanghai Institute of Microsystem and Information Technology, *Chinese Academy of Sciences*,
Shanghai 200050, China

³Center of Materials Science and Optoelectronics Engineering,
University of Chinese Academy of Sciences, Beijing 100049, China

⁴Shanghai Research Center for Quantum Sciences, Shanghai 201315, China

 (Received 14 May 2026; revised 10 July 2026; accepted 21 July 2026; published 18 August 2026)

Superconducting nanowire single-photon detectors (SNSPDs) have demonstrated timing jitter in the few-picosecond regime, yet their timing resolution deteriorates substantially under high-count-rate operation. Existing interpretations mainly attribute this degradation to deterministic waveform distortions, such as multiphoton responses and pulse pile-up; however, the experimentally observed jitter broadening at high count rates cannot be fully accounted for within this picture. Here, we show that stochastic baseline fluctuations arising from finite-memory readout dynamics constitute an intrinsic source of the count-rate-dependent timing jitter in SNSPD systems. For stochastically arriving photons, overlapping recovery responses accumulate in the readout chain and generate statistically fluctuating baselines, which are converted into timing uncertainty through threshold-based timing extraction. We develop a stochastic-process framework that quantitatively connects photon statistics, readout dynamics, and timing jitter. The framework predicts characteristic scaling behaviors, including a nonmonotonic dependence of baseline fluctuations under pulsed excitation with a maximum near half of the repetition frequency. These predictions are quantitatively verified through systematic variations of count rate, circuit time constant, and detector dynamical properties. Our results identify stochastic baseline dynamics as a fundamental mechanism limiting timing resolution in high-count-rate SNSPD operation and provide a general framework for optimizing finite-memory high-speed photon-counting systems.

DOI: [10.1103/8yf7-blyh](https://doi.org/10.1103/8yf7-blyh)

I. INTRODUCTION

Timing jitter is a key performance metric of single-photon detectors, which quantifies the temporal uncertainty in the photon arrival time recorded by the detector. Superconducting nanowire single-photon detectors (SNSPDs) [1] have achieved timing jitter in the few-picosecond regime [2–7], establishing them as one of the most precise photon-timing technologies currently available. Continuous advances in nanowire engineering,

device fabrication, and cryogenic readout techniques have progressively reduced the timing jitter of SNSPD systems. As the achievable timing resolution approaches the few-picosecond level, increasing attention has shifted toward system-level mechanisms that govern timing performance under practical operating conditions. In particular, maintaining low timing jitter at high count rates (tens of Mcps and above) has become a central challenge for high-speed photon-counting applications, including quantum information processing [8–17], deep-space optical communication [18–20], and single-photon ranging [5,21,22].

Experimentally, timing jitter is well known to increase substantially as the count rate rises [20,23–25]. Existing interpretations mainly attribute this degradation to deterministic waveform distortions, including multiphoton responses and pulse pile-up effects [23,26–28], which modify the pulse shape and alter threshold-crossing

*Contact author: lihao@mail.sim.ac.cn

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

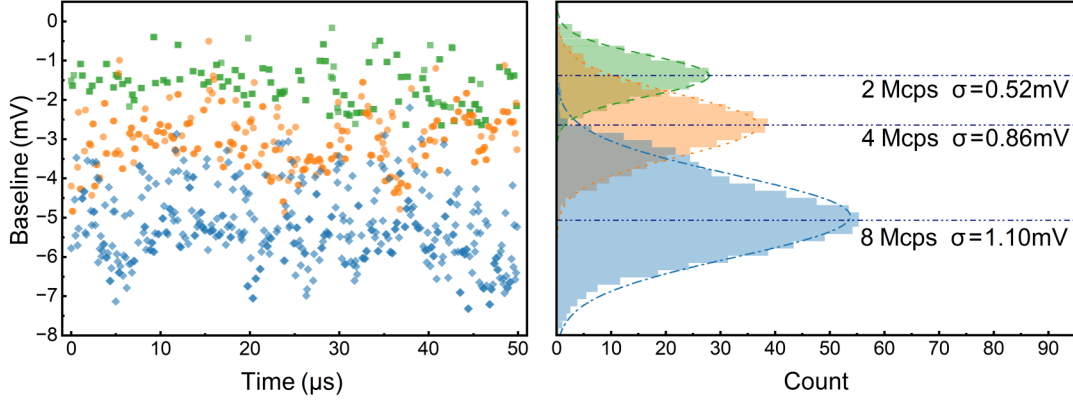


FIG. 1. Experimental evidence of stochastic baseline fluctuations induced by finite-memory readout dynamics in SNSPD systems. Left panel: instantaneous pre-event baseline values measured over a 50- μs acquisition window at different count rates, where each point corresponds to the baseline level immediately preceding a detection event. Right panel: corresponding statistical distributions obtained from the same baseline data. As the count rate increases, the baseline exhibits progressively stronger fluctuations accompanied by systematic broadening of the corresponding distributions. These observations directly reveal the stochastic accumulation of overlapping finite-memory responses in the readout chain under high-count-rate operation.

dynamics. These mechanisms have provided important insight into high-count-rate timing behavior. However, experimental observations consistently show that substantial timing broadening persists even when such deterministic effects are strongly suppressed, suggesting that an additional physical mechanism contributes to the observed degradation.

At the same time, the low-jitter SNSPD systems are commonly implemented using ac-coupled or effectively high-pass-filtered readout architectures to suppress low-frequency noise and stabilize broadband amplification [2–7]. Such readout schemes inherently possess finite-memory dynamics: each detection event produces a recovery response that persists over a finite temporal window. Under stochastic photon arrival, these residual responses accumulate statistically and continuously perturb the signal baseline. As the count rate increases, the baseline therefore evolves from a static reference level into a stochastic dynamical variable governed by the collective history of prior detection events.

Figure 1 experimentally illustrates this behavior. As the count rate increases, the pre-event baseline exhibits progressively stronger stochastic fluctuations, accompanied by a systematic broadening of the corresponding statistical distributions. These observations reveal the growth of baseline variance arising from the accumulation of overlapping finite-memory responses. Because threshold-based timing extraction depends sensitively on the instantaneous baseline state, such fluctuations are directly converted into timing uncertainty. Under high-count-rate operation, the resulting jitter broadening can become comparable to or even exceed the low-count-rate system timing jitter.

In this work, we identify stochastic baseline fluctuations as an intrinsic mechanism governing timing-jitter

broadening in high-count-rate SNSPD systems. We show that finite-memory dynamics in the readout chain transform stochastic photon arrivals into fluctuating baseline trajectories, which are subsequently converted into timing uncertainty through threshold-based timing extraction. Based on this physical picture, we develop a stochastic theoretical framework to describe the interplay among photon statistics, system memory, and timing performance under both continuous-wave and pulsed excitation. Systematic experiments performed across different count rates, recovery timescales, and detector dynamical regimes confirm the predicted scaling behavior and clarify the conditions under which stochastic baseline dynamics dominate the timing performance. These results establish a general physical framework for understanding timing limitations in high-speed photon-counting systems.

II. PHYSICAL PICTURE

The physical origin of the observed stochastic baseline fluctuations can be understood from the finite-memory response characteristics of the SNSPD readout chain, as illustrated in Fig. 2. In practical SNSPD systems, ac coupling or effectively high-pass-filtered amplification is widely employed to suppress low-frequency noise and stabilize broadband readout operation. Such readout architectures impose a charge-balancing constraint on the electrical response, such that each detection pulse is accompanied by a compensating baseline relaxation [29], as schematically illustrated in Fig. 2(a). Consequently, a detection event cannot produce a strictly localized temporal response but instead generates a residual perturbation that persists over a finite recovery time constant. Following a single detection event, the baseline relaxes gradually toward equilibrium over a characteristic time constant

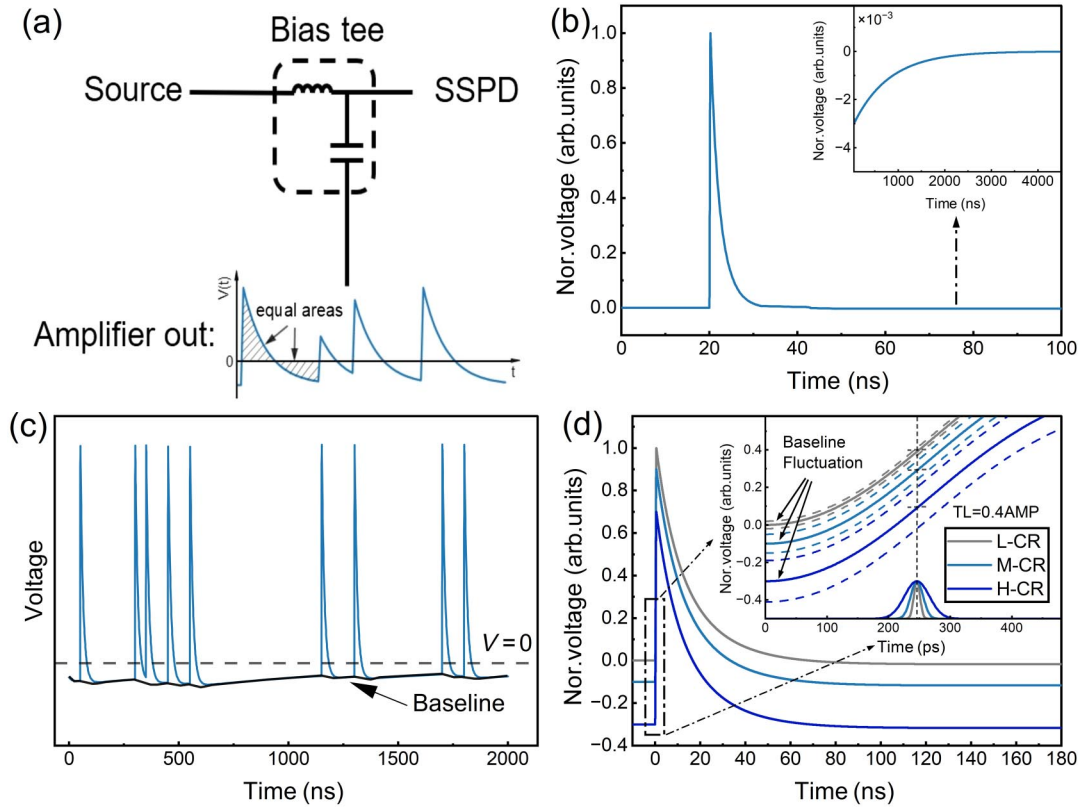


FIG. 2. Physical picture of stochastic baseline fluctuations and timing-jitter broadening in finite-memory SNSPD readout systems. (a) Schematic illustration of the finite-memory response introduced by an ac-coupled or effectively high-pass-filtered SNSPD readout chain. The bias tee is typically composed of a broadband inductor in the dc bias arm and a coupling capacitor in the rf output arm. Together with the input capacitance and input impedance of the subsequent broadband amplifier, the coupling capacitor determines the low-frequency cutoff and effective recovery timescale of the readout system. Each detection event therefore generates a compensating baseline displacement that relaxes over a finite temporal window. (b) Temporal evolution of the baseline following a single detection event. Owing to the readout-chain memory recovery time constant τ_{el} , the baseline does not return instantaneously to equilibrium but instead exhibits a gradual recovery governed by the readout dynamics. The inset shows the long-timescale recovery process following the SNSPD rising edge, illustrating the slowly relaxing baseline tail induced by the finite-memory response. (c) Under finite-count-rate operation, residual baseline perturbations from successive detection events overlap and accumulate statistically, driving the baseline into a fluctuating stochastic state. (d) Conversion of stochastic baseline fluctuations into timing jitter in threshold-based timing extraction. Variations in the instantaneous baseline modify the effective threshold-crossing condition and introduce additional timing uncertainty. The inset illustrates representative pulse waveforms under low-count-rate (L-CR), medium-count-rate (M-CR), and high-count-rate (H-CR) operation, showing progressively stronger baseline fluctuations that shift the apparent pulse onset and broaden the measured timing distribution under fixed-threshold discrimination.

determined by the effective recovery dynamics of the readout system, as illustrated in Fig. 2(b). This finite relaxation process introduces temporal memory into the system, causing the instantaneous baseline to depend on the cumulative history of preceding detection events rather than solely on the current pulse.

When the detector operates at a finite count rate, successive detection events occur within this recovery window. As illustrated schematically in Fig. 2(c), the residual responses associated with different events therefore accumulate and overlap in time. Under stochastic photon arrival, this accumulation process drives the baseline into a fluctuating dynamical state whose statistical properties are governed jointly by the photon-arrival statistics and the

finite-memory response of the readout chain. The baseline thus evolves from a static reference level into a stochastic variable with history-dependent fluctuations.

These stochastic baseline fluctuations directly affect timing extraction in threshold-discrimination-based SNSPD systems. Because the photon arrival time is determined by a fixed voltage threshold, fluctuations in the instantaneous baseline modify the effective threshold-crossing condition, thereby converting baseline-voltage fluctuations into additional timing uncertainty, as illustrated in Fig. 2(d). Unlike conventional multiphoton-response or pulse-pile-up mechanisms, which are mainly associated with waveform distortions caused by specific preceding detection events, which arise from deterministic

modifications of the detector waveform, the present mechanism originates from the statistical fluctuations of accumulated finite-memory responses over the memory time of the readout circuit.

This physical picture establishes the basis for a stochastic-process description of baseline dynamics in finite-memory SNSPD systems. In the following section, we develop a quantitative framework that connects photon-arrival statistics, readout response dynamics, and timing-jitter broadening under both continuous-wave and pulsed excitation conditions.

III. THEORY

To quantitatively describe stochastic baseline fluctuations in high-count-rate SNSPD systems, the readout dynamics are treated as a finite-memory linear system driven by stochastic photon arrivals. Under strongly attenuated optical excitation, photon detection events are well described by a Poisson point process [30,31] with an average count rate λ . In the experimentally relevant regime considered here, detection events are statistically independent, and the baseline dynamics are dominated by the linear superposition of residual recovery responses from preceding events.

Each detection event generates a finite-memory response in the readout baseline. The baseline dynamics can therefore be described as a stochastic shot-noise process formed by the superposition of all preceding responses. The single-event response kernel is described by

$$h(t) = \frac{Q}{\tau_{el}} e^{-\frac{t}{\tau_{el}}} u(t), \quad (1)$$

where Q is the effective pulse area associated with a single detection event, τ_{el} is the readout-chain memory recovery time constant, and $u(t)$ is the Heaviside step function.

For a sequence of photon-arrival times t_i , the instantaneous baseline is given by

$$B(t) = \sum_i h(t - t_i) \quad (2)$$

which describes a Poisson-driven finite-memory stochastic process.

Under Poisson excitation, the stationary baseline statistics can be derived analytically using standard shot-noise theory. The ensemble-averaged baseline is

$$E[B(t)] = \int_{-\infty}^t \lambda h(t - t') dt' = \lambda Q \quad (3)$$

while the baseline variance follows directly from Campbell's theorem [32]:

$$\text{Var}[B(t)] = \lambda \int_0^{+\infty} h^2(t) dt = \frac{\lambda Q^2}{2\tau_{el}} \quad (4)$$

leading to a baseline standard deviation

$$\sigma_B = Q \sqrt{\frac{\lambda}{2\tau_{el}}} \quad (5)$$

This scaling law indicates that baseline fluctuations increase with event rate and decrease with readout-chain memory recovery time constant, reflecting the intrinsic memory effect of the readout circuit.

For pulsed excitation, the stochastic process becomes discretized by the laser repetition period $T = 1/f_{\text{rep}}$. In this regime, photon arrivals follow a Bernoulli process with detection probability $p = \lambda/f_{\text{rep}}$ per excitation cycle. The resulting baseline evolution obeys a discrete stochastic recurrence relation (see Appendix B), yielding the steady-state fluctuation amplitude

$$\sigma_B = \frac{Q}{\tau_{el} \sqrt{e^{\frac{2T}{\tau_{el}}} - 1}} \sqrt{\frac{\lambda}{f_{\text{rep}}} \left(1 - \frac{\lambda}{f_{\text{rep}}}\right)} \quad (6)$$

In the limit $T \ll \tau_{el}$, this reduces to

$$\sigma_B = \frac{Q}{\sqrt{2}\tau_{el}} \sqrt{\lambda \left(1 - \frac{\lambda}{f_{\text{rep}}}\right)} \quad (7)$$

while in the limit $f_{\text{rep}} \rightarrow \infty$, the result recovers the continuous excitation case.

Notably, the theory predicts a characteristic nonmonotonic dependence of baseline fluctuations under pulsed excitation. The fluctuation amplitude reaches a maximum near $\lambda \approx f_{\text{rep}}/2$, reflecting the competition between stochastic event randomness and the increasing regularity of the excitation sequence at high detection probability. In the low-count-rate regime ($\lambda \ll f_{\text{rep}}$), detection events are sparse and residual responses overlap weakly, resulting in small baseline fluctuations. As λ increases, stochastic overlap between successive recovery responses enhances the baseline variance. In contrast, as $\lambda \rightarrow f_{\text{rep}}$, nearly every excitation pulse produces a detection event, causing the sequence to approach a quasi-deterministic periodic train and thereby suppressing stochastic fluctuations.

The theoretical predictions for continuous-wave and pulsed excitation are summarized in Fig. 3. Under continuous excitation, the baseline fluctuations increase monotonically with count rate due to the accumulation of finite-memory responses. Under pulsed excitation, however, the finite repetition constraint produces a distinct nonmonotonic dependence. As the repetition frequency increases, the pulsed-excitation results continuously converge toward the continuous excitation, demonstrating that both excitation regimes are naturally unified within the same stochastic framework.

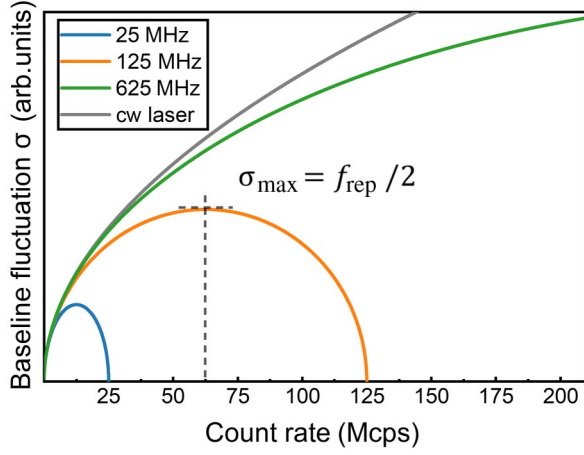


FIG. 3. Scaling behavior of stochastic baseline fluctuations under continuous-wave and pulsed excitation. Calculated standard deviation of the baseline fluctuations as a function of detection rate for continuous-wave and pulsed excitation conditions. Under continuous excitation, the baseline variance increases monotonically with count rate due to the statistical accumulation of overlapping finite-memory responses. Under pulsed excitation, the discrete repetition period modifies the fluctuation statistics and produces a nonmonotonic dependence with a maximum near one-half of the repetition frequency. As the repetition frequency increases, the pulsed-excitation results continuously approach the continuous excitation, demonstrating the unified stochastic description of both operating regimes.

Timing extraction in SNSPD systems is typically performed using fixed-threshold discrimination. Baseline fluctuations are therefore converted into timing uncertainty through the local pulse slope $S = dV/dt$ according to

$$\sigma_t = \frac{\sigma_B}{S}. \quad (8)$$

Thus, the scaling behavior of timing jitter directly inherits that of the baseline fluctuations, establishing a quantitative link between readout statistics and timing performance.

Under continuous excitation, using $Q = A\tau_{es}$ and $S = A/t_{rise}$, this leads to

$$\sigma_{t_{cw}} = \frac{t_{rise} \tau_{es}}{\sqrt{2\tau_{el}}} \sqrt{\lambda} \quad (9)$$

while under pulsed excitation,

$$\sigma_{t_{pulse}} = \frac{t_{rise} \tau_{es}}{\sqrt{2\tau_{el}}} \sqrt{\lambda \left(1 - \frac{\lambda}{f_{rep}}\right)}, \quad (10)$$

where t_{rise} is the rise time of the SNSPD pulse, τ_{es} is the electrical recovery time constant, τ_{el} is the readout-chain memory recovery time constant, A is the amplitude of the pulse, λ is the count rate, and f_{rep} is the repetition

frequency of the pulsed laser. These expressions provide a direct theoretical prediction for the count-rate dependence of timing jitter.

IV. EXPERIMENTAL VALIDATION

To quantitatively compare the theoretical predictions with experiment, the stochastic-process parameters are related to experimentally measurable pulse characteristics of the SNSPD readout system. For the extraction of theoretical prediction parameters, please see Appendix D.

A. Count-rate dependence under pulsed excitation

We first experimentally examine the predicted count-rate scaling of stochastic baseline fluctuations under pulsed excitation. The detector count rate is continuously tuned by varying the incident optical power while maintaining a fixed laser repetition frequency. Output waveforms are recorded using a high-bandwidth oscilloscope, from which baseline statistics are extracted, while count rate and timing jitter are independently measured using a time-to-digital converter (TDC). The experiment is performed using an SNSPD with an electrical recovery time constant τ_{es} of approximately 11.4 ns, operated at a bias current of 19 μ A. The detector is driven by a pulsed laser with a repetition frequency of 20 MHz [see Appendix A for more information about the experimental device and Appendix F for the verification of the 125-MHz pulsed laser and continuous-wave (cw) laser], and an external coupling capacitor of 10 nF is employed in the readout chain. The theoretical prediction shown in Fig. 4(a) is calculated directly from independently measured system parameters without adjustable fitting coefficients.

As shown in Fig. 4(a), the measured baseline fluctuations exhibit a pronounced nonmonotonic dependence on count rate. The fluctuation amplitude initially increases with count rate, reaches a maximum near half of the repetition frequency, and decreases at higher rates. The measurements quantitatively follow the theoretical prediction for pulsed stochastic excitation, confirming the finite-memory accumulation mechanism predicted by the stochastic framework. The slight discrepancy that the experimental measurements are marginally higher than the theoretical predictions arises from the inherent baseline noise of the readout system, which exhibits a standard deviation of approximately 0.45 mV. A more pronounced deviation is observed for the final data point at the highest count rate. Experimentally, the measured count rate reaches 20.06 MHz, slightly exceeding the 20 MHz laser repetition rate due to residual dark counts and a small timing mismatch between the laser synchronization signal and the TDC clock. In contrast, the theoretical model assumes an ideal steady-state condition in which every laser pulse generates exactly one detection event. The additional stochastic events (dark counts) perturb this

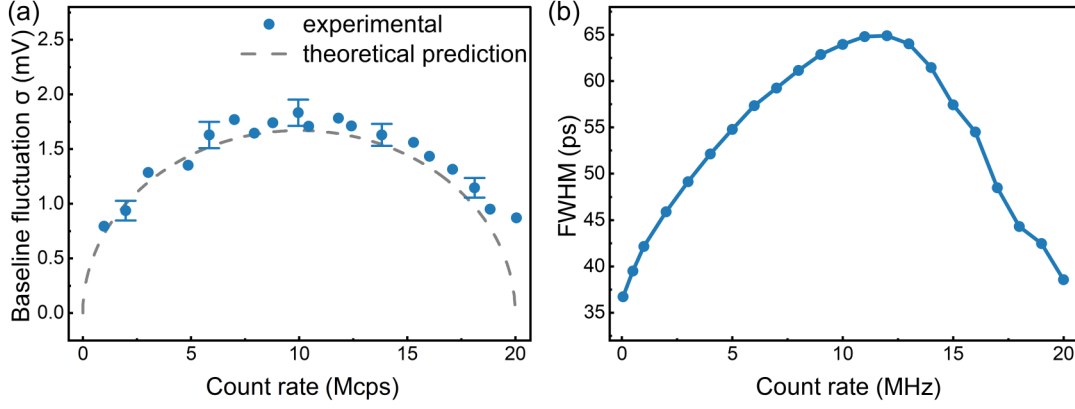


FIG. 4. Experimental validation of stochastic baseline fluctuations and timing-jitter broadening under pulsed excitation. Error bars denote the standard deviation obtained from repeated measurements ($N = 6$). For clarity, repeated measurements were performed only for five representative count rates covering the investigated range. (a) Measured standard deviation of the pre-event baseline as a function of count rate under pulsed excitation. The experimental results quantitatively reproduce the predicted nonmonotonic dependence, with a maximum near one-half of the laser repetition frequency. The dashed curve represents the theoretical prediction calculated from independently measured system parameters without adjustable fitting coefficients. (b) Measured timing-jitter full width at half maximum (FWHM) of the single-photon peak as a function of count rate. The single-photon contribution is extracted from the multipoint timing histogram using the procedure described in Appendix D. The timing jitter exhibits the same count-rate dependence as the baseline fluctuations, demonstrating the direct conversion of stochastic baseline fluctuations into timing uncertainty.

periodic steady state, leading to larger baseline fluctuations than predicted by the idealized framework.

Furthermore, the timing jitter extracted from the single-photon response peak (see Appendix E for the extraction procedure) is shown in Fig. 4(b). The single-photon component is analyzed separately in order to exclude additional timing broadening associated with multiphoton detection events [26,28]. The measured timing jitter follows the same count-rate dependence as the baseline fluctuations, indicating that stochastic baseline dynamics are converted into timing uncertainty through threshold-based timing extraction. We further extended the stochastic finite-memory framework to predict the overall timing jitter in Appendix G. This appendix demonstrates the framework's capacity to predict total system jitter and the limitation of the simplified model in linearly converting baseline fluctuations to timing jitter.

B. Influence of readout-chain memory recovery time constant

To isolate the role of finite-memory recovery dynamics, the readout-chain memory recovery time constant τ_{el} is systematically tuned by varying the coupling capacitance (the actual capacitors we use are 2.2, 4.7, 10, 47, and 100 nF) in the ac-coupled readout chain. The experiment is performed using an SNSPD with an electrical recovery time constant τ_{es} of approximately 11.4 ns, operated at a bias current of 19 μ A. The detector is driven by a pulsed laser with a repetition frequency of 20 MHz, and the counting rate of SNSPD is fixed at 10 Mcps. The theoretical prediction shown in Fig. 5(a) is calculated

directly from independently measured system parameters without adjustable fitting coefficients.

The experimentally observed dependence quantitatively follows the theoretical prediction, confirming that shorter readout-chain memory times reduce temporal averaging and thereby enhance stochastic baseline variance. The corresponding timing-jitter measurements are summarized in Fig. 5(b). The measured single-photon timing jitter exhibits the same dependence on the readout-chain memory recovery time constant as the baseline fluctuations, increasing systematically as τ_{el} decreases. The correlated evolution of baseline variance and timing jitter further supports the theoretical prediction that the count-rate-dependent timing degradation originates from stochastic baseline dynamics.

C. Influence of device dynamical properties

As shown in Fig. 6(a), devices with lower kinetic inductance exhibit faster electrical relaxation dynamics and correspondingly narrower output pulses. As a result, the time-integrated perturbation associated with each detection event is reduced, corresponding to a smaller pulse area Q in the stochastic theory.

The impact of this effect on stochastic baseline fluctuations is shown in Fig. 6(b). Devices with lower kinetic inductance exhibit systematically reduced baseline fluctuations across the entire count-rate range. This behavior confirms that stochastic baseline accumulation is governed by the strength of the perturbation introduced by each detection event.

For the device with the smallest kinetic inductance (220 nH), the baseline standard deviation exhibits only a

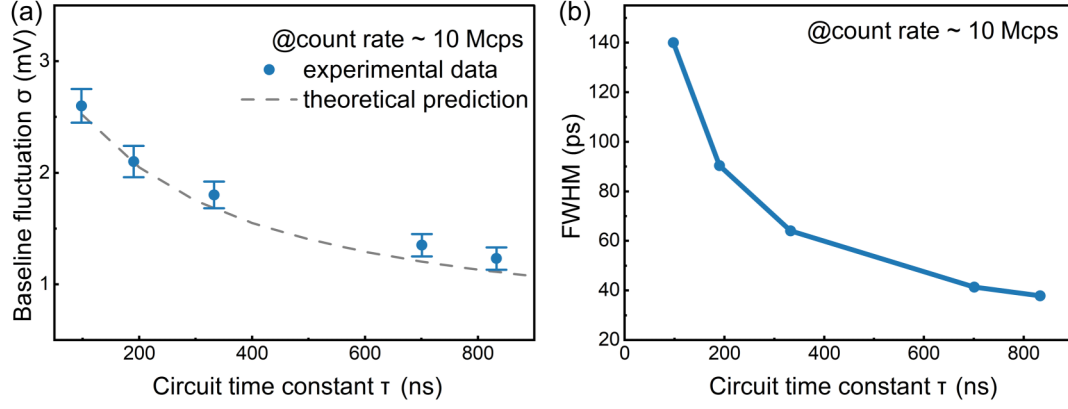


FIG. 5. Dependence of stochastic baseline fluctuations and timing jitter on the readout-chain memory recovery time constant of the readout circuit. All measurements are performed at a fixed count rate of 10 Mcps. Error bars denote the standard deviation obtained from repeated measurements ($N = 6$). (a) Measured standard deviation of the pre-event baseline as a function of the readout-chain memory recovery time constant τ_{el} . The experimental results quantitatively follow the theoretical prediction without adjustable fitting parameters. Shorter readout-chain memory recovery times produce larger baseline fluctuations due to reduced temporal averaging of residual responses. (b) Measured timing-jitter FWHM of the single-photon peak as a function of τ_{el} . The timing jitter follows the same dependence on recovery dynamics as the baseline fluctuations, demonstrating the direct conversion of stochastic baseline fluctuations into timing uncertainty.

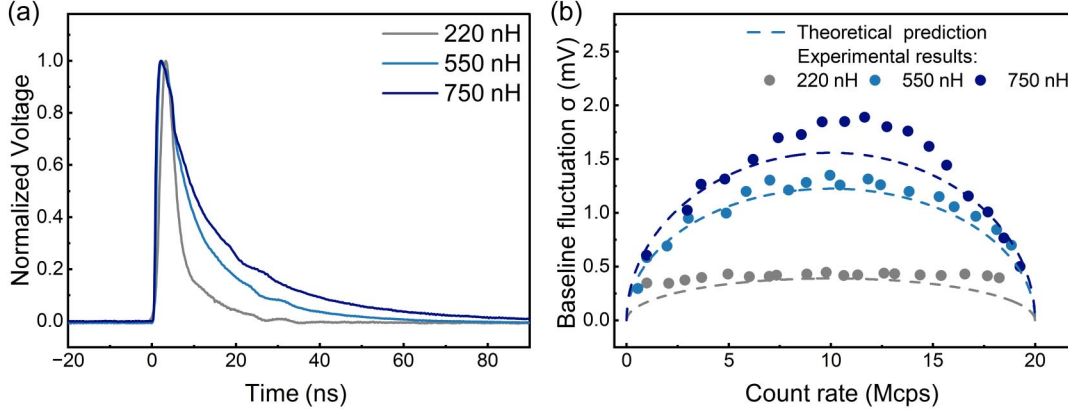


FIG. 6. Influence of detector dynamical properties on stochastic baseline fluctuations. (a) Normalized output pulse waveforms for SNSPD devices with different kinetic inductance values (220, 550, and 750 nH). Devices with lower kinetic inductance exhibit faster electrical relaxation and narrower pulse responses, corresponding to reduced time-integrated perturbation per detection event. (b) Measured baseline standard deviation as a function of count rate for devices with different kinetic inductance. Markers represent experimental data, and dashed curves represent theoretical predictions. Devices with lower kinetic inductance exhibit systematically reduced baseline fluctuations due to weaker single-event perturbations. For the lowest-inductance device, the baseline fluctuations approach the electronic noise floor, resulting in weak count-rate dependence. For the highest-inductance device, systematic deviation from the theoretical prediction emerges at high count rates due to incomplete detector recovery and additional correlations between successive detection events.

weak dependence on count rate. This indicates that the baseline fluctuation is masked by the system noise floor, placing the system in a noise-limited regime rather than a fluctuation-limited regime.

For the device with the largest kinetic inductance, systematic deviation from the theoretical prediction is observed at high count rates. This deviation originates from the breakdown of the independent-arrival approximation underlying the stochastic theory. In this regime, the detector electrical recovery time constant becomes comparable to the laser repetition period $T = 50$ ns, leading to

incomplete detector recovery and additional correlations between successive detection events. The observed deviation therefore marks the crossover from a readout-dominated stochastic regime to a detector-dynamics-dominated regime [26,27].

V. DISCUSSION

The stochastic framework developed in this work reveals a timing-jitter mechanism originating from the interplay between stochastic excitation and finite-memory

constraints imposed by the readout chain. In systems with a readout-chain memory recovery time constant τ_{el} , the baseline evolves into a stochastic dynamical variable driven by the accumulation of random detection events. Through threshold-based timing extraction, these baseline fluctuations are directly converted into timing uncertainty.

This mechanism is fundamentally distinct from previously identified effects such as multiphoton responses and pulse pile-up. Multiphoton effects are governed by photon-number statistics and primarily modify waveform shape, whereas pulse pile-up originates from deterministic temporal overlap between adjacent events. In contrast, the mechanism identified here arises from the stochastic accumulation of single-event responses in a finite-memory system, producing timing broadening even in the absence of strong multiphoton distortion or direct pulse overlap. We note that the baseline fluctuation investigated in this work and pulse pile-up both originate mathematically from the residual responses of previous detection events; however, they differ in their characteristic timescales and correction approaches. Pulse pile-up is a short-time-scale effect associated with the deterministic correlation between adjacent detection events, and its influence can be characterized and corrected by establishing a correction relationship based on the temporal intervals between events. In contrast, the baseline fluctuation is a long-time-scale stochastic statistical process resulting from the accumulation of residual responses from multiple previous events within the memory time of the readout circuit. Therefore, it cannot be effectively corrected by a deterministic correction curve based only on neighboring events.

The present theory considers the experimentally relevant regime in which the readout response is dominated by the readout-chain memory recovery time constant τ_{el} . In realistic SNSPD readout systems, frequency-dependent impedance and multiple intrinsic relaxation processes may produce deviations from an ideal single-exponential response. Furthermore, the present conversion from baseline fluctuation to timing jitter assumes that timing broadening arises solely from fluctuations of the pulse baseline at the discriminator threshold. In practice, the same finite-memory dynamics also modulate the instantaneous detector bias current through the ac-coupled readout circuit. Since the intrinsic timing jitter of an SNSPD depends on its bias current, these bias-current fluctuations introduce an additional contribution to the overall timing jitter that is not included in the present conversion model. Consequently, the predicted timing-jitter degradation should be regarded as a lower-bound estimate of the total impact of baseline fluctuations. Another limitation arises for detectors with relatively long electrical recovery times. In this case, pulse pile-up and finite-memory baseline fluctuations coexist over part of the count-rate range. When pulse pile-up becomes the dominant mechanism determining the pulse baseline, the experimentally

observed baseline fluctuations can exceed the values predicted by the present theory, which assumes that the baseline is governed solely by the accumulation of long-memory recovery responses. Nevertheless, by choosing experimental conditions that effectively separate pulse pile-up from finite-memory baseline fluctuations, the predicted scaling behavior is consistently reproduced across multiple independently controlled parameters. This agreement confirms that stochastic finite-memory baseline dynamics constitute the dominant physical origin of the observed timing broadening.

The scaling behavior $\sigma_{t_{\text{pulse}}} \propto t_{\text{rise}}\tau_{es}/\sqrt{\tau_{el}} \times \sqrt{\lambda(1 - \lambda/f_{\text{rep}})}$ predicted by the stochastic framework further clarifies the physical parameters governing high-count-rate timing performance. Neglecting numerical prefactors and focusing on the dominant scaling relations, the baseline-induced timing broadening increases approximately with the square root of count rate and with the effective temporal perturbation associated with each detection event while decreasing with the square root of the readout-chain memory recovery time constant. Physically, larger single-event pulse areas introduce stronger perturbations into the finite-memory readout chain, whereas faster signal rise slopes reduce the sensitivity of threshold-crossing extraction to baseline variations. At the same time, longer readout-chain memory recovery dynamics effectively average stochastic fluctuations over extended temporal windows, thereby suppressing baseline variance. These scaling relations provide a direct physical picture connecting detector pulse dynamics, readout memory effects, and timing performance in high-speed photon-counting systems. An alternative strategy for mitigating the finite-memory effect identified in this work is to further reduce the coupling capacitance, thereby shortening the readout-chain memory time constant τ_{el} and suppressing the accumulation of residual baseline perturbations. In principle, this approach could substantially mitigate the memory effect discussed here. However, excessively reducing the coupling capacitance may introduce waveform overshoot, leading to earlier latching and after-pulsing. Further investigation is therefore required to determine whether this trade-off can be effectively balanced in future high-speed SNSPD readout systems.

From a system-design perspective, the present results highlight an inherent trade-off in high-speed readout architectures. ac coupling is widely employed to suppress low-frequency noise and stabilize broadband amplifier operation, but it simultaneously introduces an effective high-pass response with finite-memory recovery dynamics. Although dc coupling can partially relax this constraint, it often introduces long-term baseline drift and reduced operating stability. Consequently, practical high-speed SNSPD systems frequently adopt hybrid readout architectures. For example, previously reported high-count-rate SNSPD systems [23] employ dc coupling in

the front-end stage while retaining ac-coupled amplification in subsequent stages, indicating that effective high-pass filtering remains difficult to eliminate entirely in realistic high-bandwidth implementations. The stochastic baseline fluctuations identified here should therefore be regarded as a generic feature of finite-memory high-speed photon-counting systems rather than an implementation-specific effect.

VI. CONCLUSIONS

In this work, we establish stochastic baseline fluctuations as an intrinsic mechanism of timing-jitter broadening in high-count-rate SNSPD systems. Under stochastic photon arrival, finite-memory responses in the readout chain accumulate statistically and drive the baseline into a fluctuating dynamical state, which is directly converted into timing uncertainty through threshold-based discrimination.

We develop a unified stochastic framework that quantitatively connects photon-arrival statistics, readout dynamics, and timing performance under both continuous-wave and pulsed excitation. The theory predicts characteristic scaling laws, including the nonmonotonic fluctuation behavior observed under pulsed excitation, and these predictions are quantitatively confirmed through systematic control of count rate, readout-chain memory recovery time constant τ_{el} , and detector dynamical properties.

These results establish stochastic baseline dynamics as a fundamental system-level contribution to timing broadening that is distinct from conventional multiphoton-response and pulse-pile-up mechanisms. More generally, the present work reveals that timing performance in high-speed photon-counting systems is fundamentally governed by the interplay between stochastic event statistics, finite system memory, and single-event response dynamics.

From a practical perspective, the present theory provides concrete design principles for high-speed, low-jitter photon-counting systems. In particular, suppressing low-frequency finite-memory responses, minimizing the time-integrated perturbation associated with individual detection events, and maximizing the signal slew rate near the discrimination threshold can systematically reduce baseline-induced timing broadening. These results provide a unified framework for the co-optimization of detector dynamics and readout electronics, offering a pathway toward next-generation high-throughput ultrafast photon-timing systems.

ACKNOWLEDGMENTS

This work was supported by the Quantum Science and Technology–National Science and Technology Major Project (No. 2023ZD0300100), and the National Natural Science Foundation of China (Grants No. U24A20320, No. 62401554).

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: EXPERIMENTAL SETUP AND SNSPD CHARACTERIZATION

The experimental setup is illustrated in Fig. 7(a), including laser excitation, detector biasing, cryogenic amplification, and signal readout. Optical excitation is provided by a 1550-nm femtosecond pulsed laser source (Carmel X-1550, repetition frequency 20 MHz). The optical power is adjusted using optical attenuators (Keysight 81570A) and a three-paddle polarization controller, calibrated by a power meter (Thorlabs PM100USB + S155C), before coupling into a self-aligned packaged SNSPD. For system detection efficiency measurements, the optical attenuation was adjusted to obtain an optical power of -108.92 dBm at the detector input, corresponding to an average photon flux of approximately 100 kcps at 1550 nm. The detector is current-biased through the dc port of a bias tee, while the output pulses are amplified using a custom cryogenic amplifier and recorded either by a high-bandwidth oscilloscope (LeCroy WavePro 804HD) or a TDC (Swabian Time Tagger Ultra). The SNSPD is operated at a base temperature of approximately 0.85 K. The timing jitter of the femtosecond laser source is negligible compared with the measured system jitter. The TDC introduces an intrinsic timing uncertainty of approximately 10 ps FWHM, which contributes in quadrature to the measured timing resolution. Figure 7(b) shows a scanning electron micrograph of the fabricated SNSPD. The detector consists of a double-layer NbN nanowire (NbN/SiO₂/NbN) [33, 34] integrated on a Si substrate with a high-reflectivity backside optical cavity [35]. The thicknesses of the NbN/SiO₂/NbN layers are 6/3/6 nm, respectively. The device has an active area of 12 μm , with a nanowire width of 85 nm and a pitch of 210 nm. The room-temperature sheet resistance of the double-layer nanowire is approximately 296 $\Omega/\square$, corresponding to an equivalent sheet resistance of 592 $\Omega/\square$ for a single NbN layer. The detector exhibits a switching current of approximately 26.8 μA and an electrical recovery time constant, τ_{es} , of approximately 11.4 ns. As shown in Fig. 7(e), the system detection efficiency reaches a plateau above a bias current of approximately 24.5 μA , indicating operation close to saturated internal detection efficiency. The detector achieves a maximum system detection efficiency of approximately 75% while maintaining a low dark count rate.

Representative pulse waveforms measured at different count rates are shown in Fig. 7(c). As the count rate increases, progressively enhanced baseline fluctuations are observed due to the accumulation of finite-memory

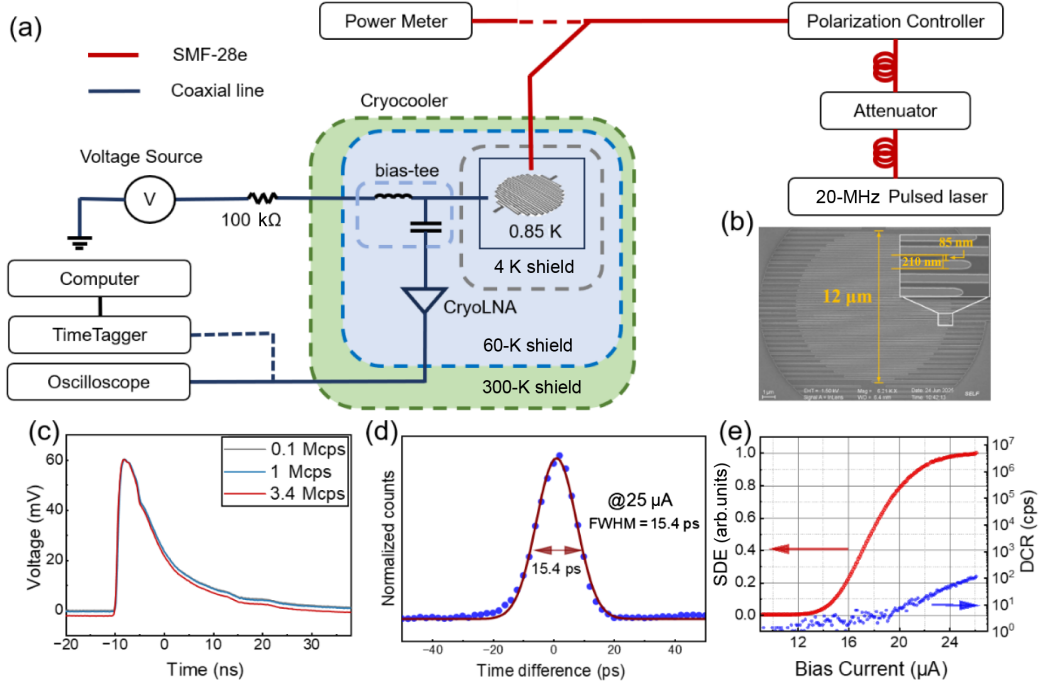


FIG. 7. (a) Experimental setup for SNSPD timing-jitter characterization, including optical excitation, detector biasing, cryogenic amplification, and electronic readout. (b) Scanning electron micrograph of the fabricated SNSPD device. (c) Measured pulse waveforms at different count rates, showing progressively enhanced baseline fluctuations at high count rates due to finite-memory accumulation. (d) Measured timing distribution at a bias current of 25 μA , yielding a system timing jitter of 15.4 ps FWHM. (e) Normalized system detection efficiency and dark count rate as a function of bias current. The device exhibits a maximum system detection efficiency of approximately 75%.

recovery responses. Figure 7(d) shows the measured timing distribution at a bias current of 25.0 μA and a count rate of 0.1 MHz, yielding a system timing jitter of 15.4 ps FWHM, which represents the detector performance under near-optimal operating conditions. In contrast, all high-count-rate measurements presented in the main text [Fig. 4(b)] were performed at a fixed bias current of 19.0 μA to avoid introducing bias current as an additional variable, since the ac-coupled readout shifts the detector operating point toward earlier latching at high count rates. Consequently, the minimum timing jitter reported in the main text is larger than the value shown here.

The normalized system detection efficiency and dark count rate as a function of bias current are presented in Fig. 7(e). The SNSPD used has an intrinsic detection efficiency close to saturation and a negligible dark count compared with the experimental photon count rate.

APPENDIX B: DISCRETE-TIME STOCHASTIC THEORY UNDER PULSED EXCITATION

In this appendix, we extend the stochastic finite-memory framework developed in the main text to the case of pulsed laser excitation. The system is naturally described in discrete time, where photon arrivals are synchronized with a repetition period $T = 1/f_{\text{rep}}$.

The impulse response of a single detection event is assumed to be unchanged from the continuous-time case and is given by

$$h(t) = \frac{Q}{\tau_{el}} e^{-\frac{t}{\tau_{el}}} u(t), \quad (\text{B1})$$

where Q characterizes the effective pulse area induced by a single event, τ_{el} is the readout-chain memory recovery time constant.

Under pulsed excitation, photon arrivals are described by a discrete Bernoulli process. In each repetition period, a detection event occurs with probability

$$p = \frac{\lambda}{f_{\text{rep}}} \quad (\text{B2})$$

and the random variable N_n describing the the n th period takes values

$$N_n = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1 - p. \end{cases} \quad (\text{B3})$$

The baseline is then described as the superposition of all past impulse responses,

$$B(t) = \sum_n N_n h(t - nT) \quad (\text{B4})$$

and the baseline immediately before each pulse is defined as

$$B_n = B(nT^-). \quad (\text{B5})$$

From the finite-memory property of the impulse response, the baseline evolution follows the discrete recursion relation:

$$B_{n+1} = \left(B_n + \frac{Q}{\tau_e} N_n \right) e^{-\frac{T}{\tau_{el}}}. \quad (\text{B6})$$

Defining $a = e^{-T/\tau_{el}}$, the mean baseline is obtained from the steady-state condition and takes the form:

$$\mu = a\mu + \frac{aQ}{\tau_{el}} p \xrightarrow{p=\frac{\lambda}{f_{\text{rep}}}} = \frac{Q}{\tau_{el}} \frac{\lambda}{f_{\text{rep}}(e^{\frac{T}{\tau_{el}}} - 1)}. \quad (\text{B7})$$

In the limit $T \ll \tau_{el}$, this expression reduces to

$$\mu = \lambda Q, \quad (\text{B8})$$

which recovers the continuous-time Poisson-driven result presented in the main text. To characterize fluctuations, we define the zero-mean variable

$$\tilde{B}_n = B_n - \mu. \quad (\text{B9})$$

Substituting into the recursion relation, we obtain

$$\tilde{B}_{n+1} = a\tilde{B}_n + \frac{aQ}{\tau_{el}} (N_n - p). \quad (\text{B10})$$

The variance evolution follows

$$\sigma_{n+1}^2 = a^2 \sigma_n^2 + \left(\frac{aQ}{\tau_{el}} \right)^2 \text{Var}(N_n). \quad (\text{B11})$$

Using $\text{Var}(N_n) = p(1-p)$, under steady-state conditions $\sigma_{n+1}^2 = \sigma_n^2 = \sigma_B^2$, yielding

$$\sigma_B^2 = \frac{a^2 Q^2 p(1-p)}{\tau_{el}^2 (1-a^2)}. \quad (\text{B12})$$

Substituting $a = e^{-T/\tau_{el}}$ and $p = \lambda/f_{\text{rep}}$, the baseline standard deviation becomes

$$\sigma_B = \frac{Q}{\tau_{el} \sqrt{e^{\frac{2T}{\tau_{el}}} - 1}} \sqrt{\frac{\lambda}{f_{\text{rep}}} \left(1 - \frac{\lambda}{f_{\text{rep}}} \right)}. \quad (\text{B13})$$

In the limit $T \ll \tau_{el}$, we use the approximation

$$e^{\frac{2T}{\tau_{el}}} - 1 \approx \frac{2T}{\tau_{el}}, \quad (\text{B14})$$

and the expression reduces to

$$\sigma_B \approx \frac{Q}{\sqrt{2\tau_{el}}} \sqrt{\lambda \left(1 - \frac{\lambda}{f_{\text{rep}}} \right)}. \quad (\text{B15})$$

Finally, in the limit $f_{\text{rep}} \rightarrow \infty$, the discrete excitation converges to the continuous Poisson limit, recovering

$$\sigma_B = Q \sqrt{\frac{\lambda}{2\tau_{el}}}, \quad (\text{B16})$$

which is consistent with the continuous-time result in the main text.

APPENDIX C: BASELINE DEFINITION AND EXTRACTION PROCEDURE

The baseline used in this work is defined as the instantaneous signal level immediately preceding each detection event, consistent with the definition adopted in the stochastic theory developed in the main text.

The experimental extraction procedure is illustrated in Fig. 8. Pulse peaks are first identified by locating waveform maxima exceeding a fixed discrimination threshold of 30 mV, as shown in Fig. 8(a). For each detected event, the pulse onset is determined using a slope-based criterion, defined as the first sample for which the rising slope

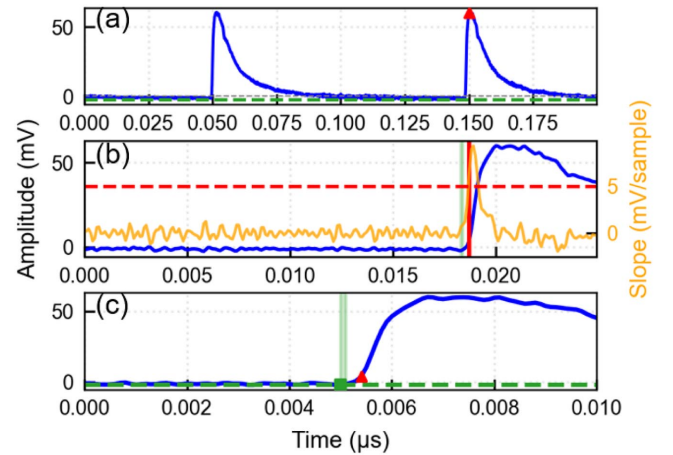


FIG. 8. Baseline definition and extraction procedure. (a) Pulse peak detection based on waveform maxima exceeding a 30-mV threshold. (b) Pulse onset determination using a slope criterion (first point exceeding 5 mV per sample). (c) Baseline assignment as the signal value 0.3–0.5 ns before the identified onset.

exceeds 5 mV per sample, as shown in Fig. 8(b). The baseline value is then assigned as the signal amplitude at a fixed temporal offset of 0.3–0.5 ns prior to the onset point, as shown in Fig. 8(c).

This procedure ensures a consistent estimation of the pre-event baseline under high-count-rate operation, where incomplete recovery of preceding pulses may otherwise bias baseline estimation if a broader temporal window is used. The resulting baseline definition is compatible with the theoretical framework, enabling quantitative comparison between measured fluctuations and theoretical predictions.

APPENDIX D: EXTRACTION OF THEORETICAL PREDICTION PARAMETERS

The baseline fluctuation model described by Eq. (B13) contains five parameters: the single-event pulse area Q , the readout-chain memory recovery time constant τ_{el} , the laser repetition period T , the repetition frequency f_{rep} , and the photon count rate λ .

Among these parameters, f_{rep} and λ are directly determined by the experimental conditions. In all measurements presented in Figs. 4(a), 5(a), and 6(b), a pulsed laser with a repetition frequency of $f_{\text{rep}} = 20$ MHz was used, corresponding to a pulse period of $T = 50$ ns. The photon count rate λ was treated as a variable for Figs. 4(a) and 6(b), while it was fixed at 10 Mcps for Fig. 5(a).

The single-event pulse area Q represents the integrated perturbation introduced into the ac-coupled readout circuit by one detection event. For an approximately exponential SNSPD pulse, it is given by

$$Q = \int V(t) dt \approx A\tau_{es}, \quad (\text{D1})$$

where A is the pulse amplitude and τ_{es} is the SNSPD electrical recovery time constant.

For the detector used in Figs. 4(a) and 5(a), the bias current was set to 19 μA , and the measured amplifier gain was 63. The extracted recovery time constant was $\tau_{es} = 11.4$ ns, resulting in $Q = 684$ mV ns. For the three detectors compared in Fig. 6(b), the extracted SNSPD recovery time constants were 4.40, 11.4, and 15.0 ns, and the corresponding values of A , obtained by direct

measurement, were 35, 44, and 44 mV, resulting in pulse areas of 154, 503, and 662 mV ns, respectively.

The readout-chain memory recovery time constant is determined by

$$\tau_{el} = R_{\text{load}} C_{\text{eff}}, \quad (\text{D2})$$

where $R_{\text{load}} = 50 \Omega$ and C_{eff} is the effective coupling capacitance. The effective capacitance includes both the intrinsic capacitance of the cryogenic amplifier (approximately 20 nF) and the external coupling capacitor. With an external capacitor of 10 nF, the effective capacitance is 6.67 nF, yielding $\tau_{el} = 333$ ns. For Fig. 5(a), external capacitors of 2.2, 4.7, 10, 47, and 100 nF were used, corresponding to τ_{el} values of 99.1, 190, 333, 701, and 833 ns, respectively. The parameters used in each theoretical calculation are summarized in Table 1.

APPENDIX E: EXTRACTION OF SINGLE-PHOTON TIMING JITTER

At high count rates, the timing histogram exhibits multiple peaks due to multiphoton detection events [26], as shown in Fig. 9. These peaks correspond to different photon-number responses and result in an overall broadening of the timing distribution. The rightmost peak is attributed to the single-photon response and is used for jitter extraction, as it corresponds to the latest detection events in the multiphoton hierarchy.

To isolate the single-photon contribution, the timing histogram is fitted using a multi-Gaussian model. Each peak is represented by a Gaussian function, and the overall distribution is obtained through least-squares minimization. The fitting stability is verified by the insensitivity of extracted peak parameters to variations in fitting range.

The timing jitter is defined as the FWHM of the rightmost Gaussian component. This definition removes multiphoton-induced broadening and provides a consistent metric for comparison with theoretical predictions.

APPENDIX F: VALIDATION UNDER DIFFERENT EXCITATION REGIMES

Figure 10 shows the measured baseline standard deviation as a function of count rate under 20 MHz pulsed

TABLE I. Summary of parameters used in theoretical calculations

Figures	f_{rep}	T	τ_{el}	Q	λ
Figure 4(a)	20 MHz	50 ns	333 ns	684 mV ns	Variable
Figure 5(a)	20 MHz	50 ns	99.1–833 ns	684 mV ns	10 Mcps
Figure 6(b)	20 MHz	50 ns	333 ns	154, 503, 662 mV ns	Variable

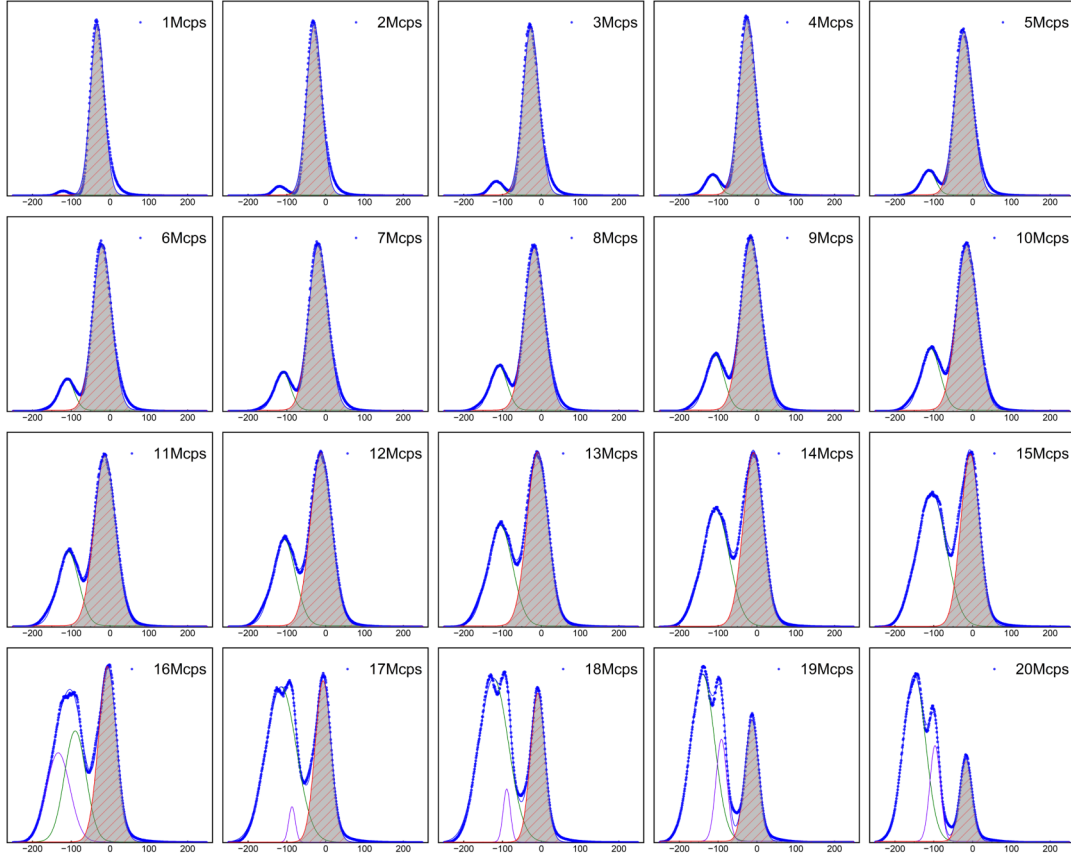


FIG. 9. Timing histogram under high-count-rate operation. Multiple peaks arising from multiphoton detection events are observed at high count rates [26, 28]. The rightmost peak corresponds to the single-photon response and is used for timing-jitter extraction. For all panels, the horizontal axis represents time (ps), and the vertical axis represents the number of counts.

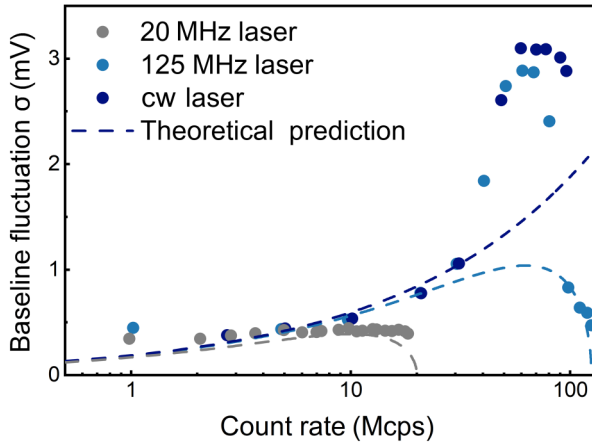


FIG. 10. Comparison of baseline-fluctuation scaling under different excitation conditions. Measured baseline standard deviation as a function of count rate under 20 MHz pulsed excitation, 125 MHz pulsed excitation, and cw operation. Experimental data are compared with theoretical predictions from the stochastic-process framework.

excitation, 125 MHz pulsed excitation, and cw operation. In the present experiments, a conventional SNSPD with a smaller active area (5 μm) and a shorter electrical recovery

time constant (approximately 4.4 ns) is employed in the measurements to mitigate pulse-overlap effects associated with the shorter repetition period. The detector bias current is maintained at 10 μA for all measurements.

As shown in Fig. 10, under 20 MHz pulsed excitation, the measured baseline fluctuations exhibit only weak dependence on count rate, indicating that the fluctuation amplitude is largely buried within the system noise floor. In contrast, under 125 MHz pulsed excitation and cw operation, the experimentally measured baseline fluctuations deviate upward from the ideal theoretical prediction at high count rates.

This deviation arises from the breakdown of the assumption of independent detection events in the stochastic framework. In this regime, the interevent interval becomes comparable to the detector electrical recovery time constant, leading to incomplete relaxation of the detector between successive events. Consequently, detector recovery dynamics introduce additional correlations that enhance baseline fluctuations beyond the readout-limited prediction.

These results indicate that the stochastic finite-memory framework captures the dominant scaling behavior of baseline fluctuations across different excitation conditions,

while also identifying the operating regime in which detector recovery dynamics become non-negligible.

APPENDIX G: OVERALL TIMING JITTER PREDICTION

The stochastic finite-memory framework developed in this work is further used to predict the overall timing jitter. Following the conversion model introduced in the main text, the baseline-induced timing jitter is obtained from the calculated baseline fluctuations and combined with the intrinsic detector jitter and readout-circuit jitter according to

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{intr}} + \sigma_{\text{noise}} + \sigma_t}, \quad (\text{G1})$$

where σ_{intr} is the intrinsic timing jitter of the SNSPD, σ_{noise} is the timing jitter introduced by the readout electronics, and σ_t is the baseline-induced timing jitter predicted from the stochastic finite-memory model. Following the simplified conversion introduced in the main text, σ_t is obtained from the calculated baseline fluctuations using the experimentally measured pulse slew rate, $S = 0.088$ mV/ps, directly extracted from the oscilloscope waveform.

Figure 11 compares the predicted overall timing jitter with the experimental measurements. The theoretical prediction success fully reproduces the overall evolution trend of the measured timing jitter over the entire count-rate range, demonstrating that the proposed stochastic finite-memory framework captures the dominant mechanism governing timing-jitter degradation under high-count-rate operation.

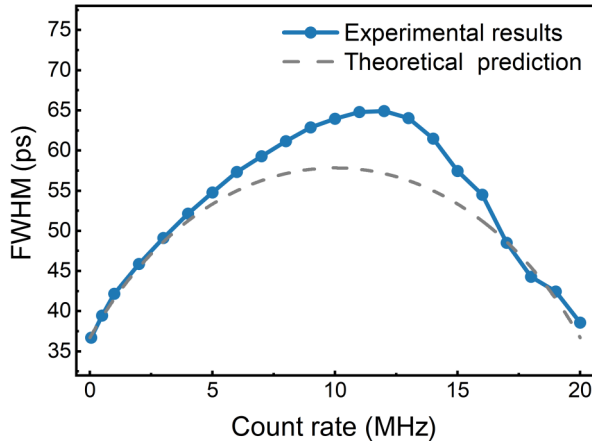


FIG. 11. Predicted and measured overall timing jitter as a function of count rate. The theoretical prediction is obtained by combining the baseline fluctuations calculated from the stochastic finite-memory model with the intrinsic detector jitter and readout-circuit jitter using the conversion model introduced in the main text.

The remaining quantitative discrepancy originates from the simplified conversion from baseline fluctuations to timing jitter rather than from the stochastic finite-memory framework itself. For simplicity, the present conversion considers only the direct influence of baseline fluctuations on the threshold-crossing position. In practice, baseline fluctuations also modulate the detector bias current through the ac-coupled readout circuit, thereby introducing an additional contribution to the intrinsic timing jitter. Quantitatively incorporating this contribution is beyond the scope of the present conversion model. In addition, the conversion adopts a linear leading-edge approximation with a constant slew rate, whereas the actual SNSPD leading edge is inherently nonlinear. Consequently, the present framework should be regarded as providing a reliable prediction of the overall timing-jitter evolution, while a more complete conversion model is required for quantitative agreement.

-
- [1] G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, C. Williams, and R. Sobolewski, Picosecond superconducting single-photon optical detector, *Appl. Phys. Lett.* **79**, 705 (2001).
 - [2] B. Korzh, Q.-Y. Zhao, J. P. Allmaras, S. Frasca, T. M. Autry, E. A. Bersin, A. D. Beyer, R. M. Briggs, B. Bumble, M. Colangelo *et al.*, Demonstration of sub-3 ps temporal resolution with a superconducting nanowire single-photon detector, *Nat. Photonics* **14**, 250 (2020).
 - [3] I. Esmail Zadeh, J. W. N. Los, R. B. M. Gourgues, J. Chang, A. W. Elshaari, J. R. Zichi, Y. J. van Staaden, J. P. E. Swens, N. Kalhor, A. Guardiani *et al.*, Efficient single-photon detection with 7.7 ps time resolution for photon-correlation measurements, *ACS Photonics* **7**, 1780 (2020).
 - [4] X.-Y. Zhang, W.-J. Zhang, H. Zhou, X.-F. Zhang, L.-X. You, H. Li, D.-H. Fan, Y.-M. Pan, H.-Q. Yu, L.-Y. Li *et al.*, NbN superconducting nanowire single-photon detector with 90.5% saturated system detection efficiency and 14.7 ps system jitter at 1550 nm wavelength, *IEEE J. Sel. Top. Quantum Electron.* **28**, 1 (2022).
 - [5] A. McCarthy, G. G. Taylor, J. Garcia-Armenta, B. Korzh, D. V. Morozov, A. D. Beyer, R. M. Briggs, J. P. Allmaras, B. Bumble, M. Colangelo *et al.*, High-resolution long-distance depth imaging LiDAR with ultra-low timing jitter superconducting nanowire single-photon detectors, *Optica* **12**, 168 (2025).
 - [6] I. Esmail Zadeh, J. W. N. Los, R. B. M. Gourgues, V. Steinmetz, G. Bulgarini, S. M. Dobrovolskiy, V. Zwiller, and S. N. Dorenbos, Single-photon detectors combining high efficiency, high detection rates, and ultra-high timing resolution, *APL Photonics* **2**, 111301 (2017).
 - [7] J. Wu, L. You, S. Chen, H. Li, Y. He, C. Lv, Z. Wang, and X. Xie, Improving the timing jitter of a superconducting nanowire single-photon detection system, *Appl. Opt.* **56**, 2195 (2017).

- [8] W. Li, L. Zhang, H. Tan, Y. Lu, S.-K. Liao, J. Huang, H. Li, Z. Wang, H.-K. Mao, B. Yan, Q. Li, Y. Liu, Q. Zhang, C.-Z. Peng, L. You, F. Xu, and J.-W. Pan, High-rate quantum key distribution exceeding 110 Mb s^{-1} , *Nat. Photonics* **17**, 416 (2023).
- [9] G.-W. Zhang, S.-T. Zheng, Y. Xiao, F.-X. Wang, W.-J. Ding, D.-P. Wang, P.-L. Hao, L. Zhang, J.-L. Chen, Y.-Y. Ding *et al.*, 5-GHz chip-based quantum key distribution with 1 Mbps secure key rate over 150 km, *Laser Photonics Rev.* **20**, e03250 (2026).
- [10] F. Gr nenfelder, A. Boaron, G. V. Resta, M. Perrenoud, D. Rusca, C. Barreiro, R. Houlmann, R. Sax, L. Stasi, S. El-Khoury *et al.*, Fast single-photon detectors and real-time key distillation enable high secret-key-rate quantum key distribution systems, *Nat. Photonics* **17**, 422 (2023).
- [11] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu *et al.*, Quantum computational advantage using photons, *Science* **370**, 1460 (2020).
- [12] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su *et al.*, Phase-programmable Gaussian boson sampling using stimulated squeezed light, *Phys. Rev. Lett.* **127**, 180502 (2021).
- [13] Y.-H. Deng, Y.-C. Gu, H.-L. Liu, S.-Q. Gong, H. Su, Z.-J. Zhang, H.-Y. Tang, M.-H. Jia, J.-M. Xu, M.-C. Chen *et al.*, Gaussian boson sampling with Pseudo-photon-number-resolving detectors and quantum computational advantage, *Phys. Rev. Lett.* **131**, 150601 (2023).
- [14] Z. Shu, Y. Li, X. Liu, H. Han, Y. Xiao, J. Huang, H. Xu, W. Zhang, L. You, Z. Wang, and H. Li, On-chip superconducting nanowire single-photon detectors integrated with pump rejection for entanglement characterization, *Photonics Res.* **13**, 1067 (2025).
- [15] C. Ding, X. Zhang, J. Xiong, Y. Xiao, T. Zhang, J. Huang, H. Xu, X. Liu, L. You, Z. Wang, and H. Li, Photon-number-resolving single-photon detector with a system detection efficiency of 98% and photon-number resolution of 32, *ACS Photonics* **12**, 4924 (2025).
- [16] Y. Xiao, S. Wei, J. Xu, R. Ma, X. Liu, X. Zhang, T. H. Tao, H. Li, Z. Wang, L. You, and Z. Wang, Superconducting single-photon spectrometer with 3D-printed photonic-crystal filters, *ACS Photonics* **9**, 3450 (2022).
- [17] J.-N. Zhang, R. Yang, X.-H. Li, C.-W. Sun, Y.-C. Liu, Y. Wei, J.-C. Duan, Z.-D. Xie, Y.-X. Gong, and S.-N. Zhu, Realization of a source-device-independent quantum random number generator secured by nonlocal dispersion cancellation, *Adv. Photonics* **5**, 036003 (2023).
- [18] F. I. Khatri, B. S. Robinson, M. D. Semprucci, and D. M. Boroson, Lunar laser communication demonstration operations architecture, *Acta Astronaut.* **111**, 77 (2015).
- [19] H. Ivanov, S. Mejri, A. Di Mira, K. J. Schulz, and C. Heese, Review of deep space optical communications, *Int. J. Satell. Commun. Networking* **43**, 193 (2025).
- [20] H. Hao, Q.-Y. Zhao, Y.-H. Huang, J. Deng, F. Yang, S.-Y. Ru, Z. Liu, C. Wan, H. Liu, Z.-J. Li, H.-B. Wang, X.-C. Tu, L.-B. Zhang, X.-Q. Jia, X.-L. Wu, J. Chen, L. Kang, and P.-H. Wu, A compact multi-pixel superconducting nanowire single-photon detector array supporting gigabit space-to-ground communications, *Light Sci. Appl.* **13**, 25 (2024).
- [21] H. Li, K. Zheng, R. Ge, L. Zhang, L. Zhang, W. He, B. Zhang, M. Wu, B. Wang, M. Mi, Y. Guan, J. Tan, H. Wang, Q. Chen, X. Tu, Q. Zhao, X. Jia, J. Chen, L. Kang, Q. Chen, and P. Wu, Noise-tolerant LiDAR approaching the standard quantum-limited precision, *Light Sci. Appl.* **14**, 138 (2025).
- [22] H. Zhou, Y.-H. He, C.-L. L , L.-X. You, Z.-H. Li, G. Wu, W.-J. Zhang, L. Zhang, X.-Y. Liu, X.-Y. Yang, and Z. Wang, Photon-counting chirped amplitude modulation lidar system using superconducting nanowire single-photon detector at 1550-nm wavelength*, *Chin. Phys. B* **27**, 018501 (2018).
- [23] I. Craiciu, B. Korzh, A. D. Beyer, A. Mueller, J. P. Allmaras, L. Narv ez, M. Spiropulu, B. Bumble, T. Lehner, E. E. Wollman, and M. D. Shaw, High-speed detection of 1550 nm single photons with superconducting nanowire detectors, *Optica* **10**, 183 (2023).
- [24] J.-X. Chen and H. Li, Research progress of high speed superconducting nanowire single-photon detector, *J. Funct. Mater. Devices* **31**, 1 (2025).
- [25] L. Stasi, T. Taher, G. V. Resta, H. Zbinden, R. Thew, and F. Bussi res, Enhanced detection rate and high photon-number efficiencies with a scalable parallel SNSPD, *ACS Photonics* **12**, 320 (2025).
- [26] D. Wang, Y. Xiao, Z. Wan, C. Ding, H. Xu, J. Huang, J. Xiong, L. You, and H. Li, Origins and correction of secondary peaks in timing jitter of high-speed superconducting nanowire single-photon detectors, *Appl. Phys. Lett.* **128**, 162601 (2026).
- [27] A. Mueller, E. E. Wollman, B. Korzh, A. D. Beyer, L. Narvaez, R. Rogalin, M. Spiropulu, and M. D. Shaw, Time-walk and jitter correction in SNSPDs at high count rates, *Appl. Phys. Lett.* **122**, 044001 (2023).
- [28] T. Schapeler, N. Lamberty, T. Hummel, F. Schlue, M. Stefszky, B. Brecht, C. Silberhorn, and T. J. Bartley, Electrical trace analysis of superconducting nanowire photon-number-resolving detectors, *Phys. Rev. Appl.* **22**, 014024 (2024).
- [29] G. F. Knoll, *Radiation Detection and Measurement* (John Wiley & Sons, New York, 2010).
- [30] R. N. Clark, B. Puzio, O. M. Green, S. T. Pradyumna, O. Trojak, A. Politi, and J. C. F. Matthews, Integrated photonics for continuous-variable quantum optics, *Nat. Photonics* **20**, 489 (2026).
- [31] A. M. Fox, *Quantum Optics: An Introduction* (Oxford University Press, New York, 2006).
- [32] N. Campbell, The study of discontinuous phenomena (1909).
- [33] P. Hu, H. Li, L. You, H. Wang, Y. Xiao, J. Huang, X. Yang, W. Zhang, Z. Wang, and X. Xie, Detecting single infrared photons toward optimal system detection efficiency, *Opt. Express* **28**, 36884 (2020).
- [34] J.-M. Xiong, W.-J. Zhang, G.-Z. Xu, L.-X. You, X.-Y. Zhang, L. Zhang, C.-J. Zhang, D.-H. Fan, Y.-Z. Wang, H. Li, and Z. Wang, Reducing current crowding in meander superconducting strip single-photon detectors by thickening bends, *Supercond. Sci. Technol.* **35**, 055015 (2022).
- [35] H. Xu, H. Han, Y. Xiao, J. Xiong, C. Ding, Z. Shu, Y. Li, X. Liu, L. You, Z. Wang, and H. Li, Impact of distributed Bragg reflectors on the intrinsic detection efficiency of superconducting nanowire single-photon detectors, *Superconductivity* **13**, 100152 (2025).