Photon number dependent afterpulsing in superconducting nanostrip single-photon detectors

Sebastian M. F. Raupach,[1,*] Mariia Sidorova [ORCID],[2,3] and Alexej D. Semenov[2]

[1]*Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany*
[2]*DLR Institute of Optical Systems, Rutherfordstrasse 2, 12489 Berlin, Germany*
[3]*Department of Physics, Humboldt-Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany*

Superconducting nanostrip single-photon detectors (SNSPDs) are widespread tools in photonic quantum technologies. Here, we study the afterpulsing phenomenon in commercial SNSPDs exhibiting significant levels of afterpulses even at low counting rates. We find different contributions, where the probability of an afterpulse is not a constant but depends on the mean number of photons per light pulse including mean numbers much less than 1. Our observations exclude the electrical circuit as the primary cause of the main contribution to the observed afterpulsing probability, which exhibits a strong dependence on mean photon number. We propose a phenomenological model that qualitatively explains our findings via the introduction of slowly relaxing "afterpulsing centers," storing the absorbed photons' energy for several tens of nanoseconds. We conjecture that two-level systems in amorphous materials are physical candidates for the role of such afterpulsing centers.

I. INTRODUCTION

Single-photon detectors [1] are pivotal components of photonic quantum technologies and applications such as quantum key distribution [2,3], photonic quantum computing [4,5], biophotonics [6,7], deep-space communication and astronomy [8–12], as well as particle physics [13] and are likely to become an important tool in time and frequency transfer and metrology [14]. While superconducting transition edge sensors play an important role in astronomy [10,12], particle physics [13], and photonic quantum computing [5], arguably the most widespread types of single-photon detector technologies are semiconductor-based avalanche diodes (SPADs) and superconducting nanostrips as single-photon detectors (SNSPDs). While SPADs allow for room-temperature operation and straightforward integration using standard processes of semiconductor technology, SNSPDs require cryogenic operating temperatures of around 3 K or less. SNSPDs, on the other hand, exhibit higher detection efficiencies, in particular in the telecom wavelength range and when embedded into an optical stack, as well as a fast response to incoming light pulses and short reset times of nanoseconds, routinely allowing for small timing jitter of the order of 10 ps and count rates of a few hundreds of megahertz [15–19].

Semiconductor-based single-photon detectors suffer from parasitic trapping and release of charge carriers following a detection event, leading to a temporary increase of dark counts, i.e., detection events in the absence of light but in the wake of another detection event. This behavior is termed afterpulsing. Afterpulsing in SPADs is typically handled by implementing a hold-off time of the order of microseconds in the detector's driving electronics, to allow for a decay of the parasitic release of charge carriers before the detector is armed again.

Afterpulsing in SNSPD is a subtle phenomenon that is usually assumed to be absent or simply ignored but is occasionally observed, e.g., in [20–24]. There, on a timescale of a few nanoseconds in parallel nanowires, it was attributed to a trade-off between the recovery of the bias current and the instantaneous critical current [21], or in single nanowires to technical sources, such as perturbations of the bias current due to reflections in the detector's readout electronics [20,22], or to an overshoot in bias current caused by a discharge current from capacitors in ac-coupled amplifiers after a detection event at high event rates [23,24].

Considering afterpulsing as a kind of excess noise, one would look for known sources destroying superconductivity locally or introducing electrical noise in superconducting devices. Two-level systems (TLSs) are rather ubiquitous objects residing in dielectric layers next to the superconductor such as substrates, intermediate layers as those from an optical stack, or oxide layers. In superconducting qubits and superconducting resonators, two-level systems are regarded as being the main factors in decoherence and losses, because of mutual interaction phenomena and dissipative phonon interaction [25]. Possible effects of two-level systems on the performance of SNSPDs do not yet seem to have found attention.

In this paper, we investigate afterpulsing behavior in an off-the-shelf commercial SNSPD. We put particular emphasis on the validity of common explanations for afterpulsing and the possible role of two-level systems. We find, that, while reflections and back-actions of the readout amplifier on the nanostrip may shape the probability distribution of afterpulses over time, they are not the primary cause of the observed afterpulsing phenomenon. In view of the timescales involved and of a significant dependence on the mean photon number
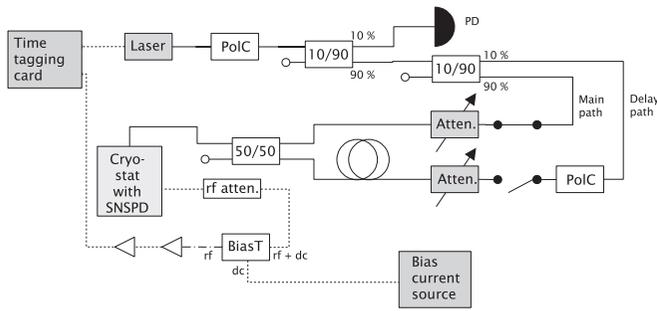
*sebastian.raupach@ptb.de

FIG. 1. Schematic sketch of the experimental setup. The optical setup is fiber based, where the laser emits light pulses with a duration of less than 100 ps. PolC: Manual polarization controller; Atten.: Variable optical attenuator; PD: Calibrated photodiode (dash-dotted line: Coaxial cables with variable total length).

per incoming light pulse, we find strong indications that some kind of two- or multilevel system may play a dominant role in the afterpulsing behavior of SNSPDs via storage and delayed release of the absorbed photons' energy. However, further work is required to elucidate the nature of these "afterpulsing centers" and their relation to previously described TLSs.

II. EXPERIMENTAL SETUP

A schematic sketch of our experimental setup is shown in Fig. 1. The fiber-coupled calibrated optical setup follows the "double attenuator" approach [26], here comprising a picosecond pulse laser, a calibrated monitor photodiode, and calibrated variable attenuators to provide light pulses with a duration of less than 100 ps at a nominal wavelength of 1548 nm and with known mean numbers of photons per pulse at the detector. The repetition rate of laser pulses is set by a waveform generator (not shown). The setup (and data taking/analysis) is very similar to the one described in [27], but here the setup includes manual polarization controllers to maximize the count rate prior to each measurement run due to the SNSPDs' polarization sensitivity. To also allow for a second, delayed light pulse at the SNSPD, we add a second light path (delay path). Each path can be individually attenuated, and opened or blocked using built-in shutters (switch symbols in Fig. 1). Unless stated otherwise, for the experiments described here only the "main" path is used. The SNSPD devices are mounted inside a dry vacuum cryostat system (Entropy GmbH) and kept at an ambient temperature of around 2.9 K [28].

The SNSPDs used in this work are two nominally identical, commercial devices (QuantumOpus, delivered in 2021) made of an amorphous, unknown material [29] and undisclosed geometry, where the material is likely to be a silicide or highly disordered nitride. As the manufacturer shared very limited or no technical information regarding these devices' layout and material that potentially might be relevant, these commercial devices essentially have to be treated as being black-box-like. The devices' pronounced polarization sensitivity indicates a meander design. We measured their critical temperature to be around 5.7 K, which is a typical value for amorphous materials. Measurement of the trigger-to-detection jitter of around 500 ps at a bias current of 11.24 μA [system detection

efficiency (SDE) around 20%] was largely limited by the time resolution of our time tagger (250 ps), which therefore constitutes an upper bound.

One of the two SNSPDs ("SNSPD 2") is read out using a cryogenic electronics board provided by the manufacturer and mounted on the 2.9 K stage of the cryostat; the other SNSPD ("SNSPD 1") is read out using commercial room-temperature electronics, consisting of a bias tee (Mini-Circuits ZFBT-282-1.5A+, 10 MHz to 2.8 GHz) and two amplifiers in series (Mini-Circuits ZX60-3018G-S+, 20 MHz to 3.0 GHz), as well as a 3-dB rf attenuator between bias tee and cryostat. The purpose of this attenuator is to serve as a well impedance matched room-temperature shunt resistor (Mini-Circuits VAT-3+, 50 Ω impedance), which we measured to have a throughput dc resistance of around 150 Ω [30]. With this setup (through the attenuator), and using a sourcemeter (Keithley 2450), we measured an apparent critical (switching) current of SNSPD 1 of at least 13.8 μA, while the intrinsic switching current of the device actually seems to be smaller [31]. For the cryogenic readout the observed SDE was around or larger than 85% for both devices; for the room-temperature readout the SDE over the range of currents used here varied between around 10% and around 65% (see Appendix B; for mean photon numbers larger than about 1 despite correcting for deadtime and taking Poisson statistics into account, we saw a systematic decrease in SDE, an observation reported also in [32] for detectors using a parallel-circuit layout). After verifying that both SNSPDs show comparable afterpulsing behavior including the dependence on mean photon number (see, e.g., Fig. 2), in view of the availability of technical information on the readout electronics required for modeling, subsequent analysis is done solely for results obtained with SNSPD 1 (i.e., using room-temperature electronics).

III. DATA ANALYSIS

With a time tagging card (TimeHarp 260; PicoQuant), each trigger pulse to the laser as well as each count event from the readout electronics (electric pulse) is registered and time tagged (nominal resolution 250 ps). Registered count events, hereafter also counts, originate either from photons absorbed by SNSPDs (light counts) or from intrinsic fluctuations in SNSPDs (dark counts), which are either conditional afterpulses or equilibrium dark counts. Count histograms have a total time axis equal to the period of laser pulses (interval), i.e., 50 μs for a repetition rate of 20 kHz. The repetition rates were intentionally chosen to be small to minimize afterpulse count "overspill" from one interval to another. We analyze the count histograms for each measurement run, e.g., for each mean photon number for a given bias current.

To visualize and quantify afterpulsing, we filter the histograms such that only those intervals are included that do contain a detection event (light count) within a certain time window following each trigger pulse to the laser. When selecting the window and deriving the histograms, we accounted for the broadening and shift of the time tags distribution for light counts with the current due to jitter [33]. As our data showed, a photon count may initiate more than one afterpulse. To distinguish between the first and secondary afterpulses, the histograms are further filtered to only include the first afterpulse following a detection event in the same interval.
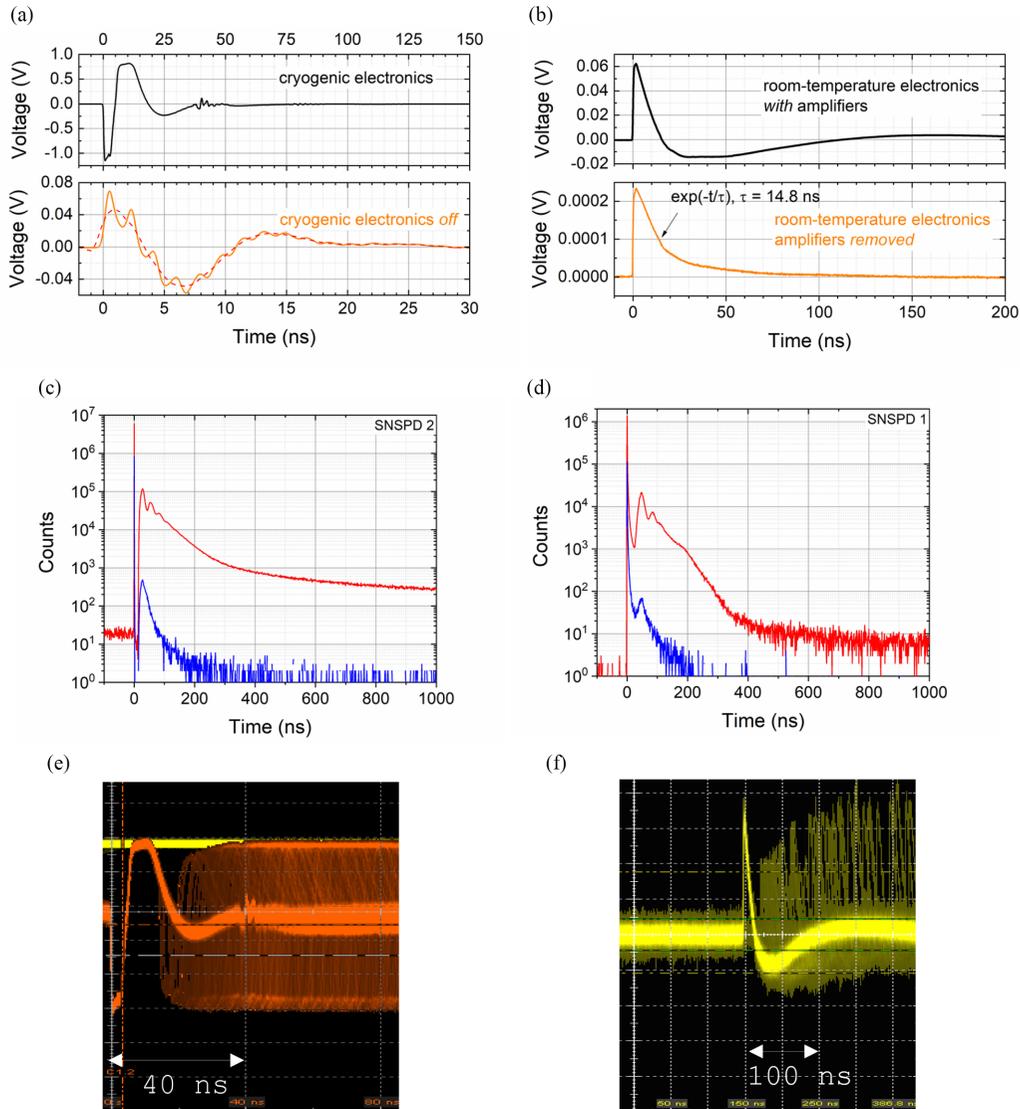
FIG. 2. Averaged time trace recorded at the output of (a) the cryogenic electronics of SNSPD 2 (dashed line in the bottom panel: Sliding average to guide the eye) and (b) at the output of the room-temperature electronics of SNSPD 1. Count histogram (1 ns bins) of (c) SNSPD 2 (300 s) and (d) SNSPD 1 (150 s) at a bias current of 11.24 μA, for mean photon numbers per incoming light pulse of 11/pulse (upper red curve) and 0.27/pulse (bottom blue curve). Zero of the time axis is set at the mean onset of light pulses. The repetition rate of light pulses was set to 20 kHz. Panels (e) and (f) show example screenshots of the oscilloscope in persistent mode showing the light detection pulse followed by afterpulses.

Finally, we calculate the distribution of the afterpulsing probability over time by dividing the number of afterpulses in each bin by the total number of intervals contributing to the respective histogram. The latter number equals the total number of detected laser pulses (light counts) in each measurement run.

IV. OBSERVATIONS

Figures 2(a) and 2(b) show (averaged) time traces of the SNSPDs' electrical pulses recorded with an oscilloscope (R&S RTO2044; nominal bandwidth 6 GHz). The output of the cryogenic electronics board [SNSPD 2, the upper trace in Fig. 2(a)] shows a sharp voltage peak followed by a broader over- and undershoot voltage. When the power supply to the board is turned off (the lower trace), a train of peaks becomes visible, that are superimposed onto the electrical pulse. The

period of this peak train corresponds to the expected round-trip time between the SNSPD and the cryogenic electronics board of 1.8 ns, suggesting that there are reflections between them due to some impedance mismatch. This was verified by changing the cable length and observing a corresponding increase in the period. These reflections affect the shape of the recorded traces also when the electronics board is powered.

The output of the room-temperature electronics [SNSPD 1, the upper trace in Fig. 2(b)] shows the typical steep leading edge followed by a slow decay and a voltage under- and overshoot. When removing the amplifiers [the lower trace in Fig. 2(b)], we do not observe any under- and overshoot; the pulse decays exponentially with a time constant of 14.8 ns.

Figures 2(c) and 2(d) demonstrate that regardless of the type of readout electronics used (with or without visible reflections in time traces), in both cases afterpulsing
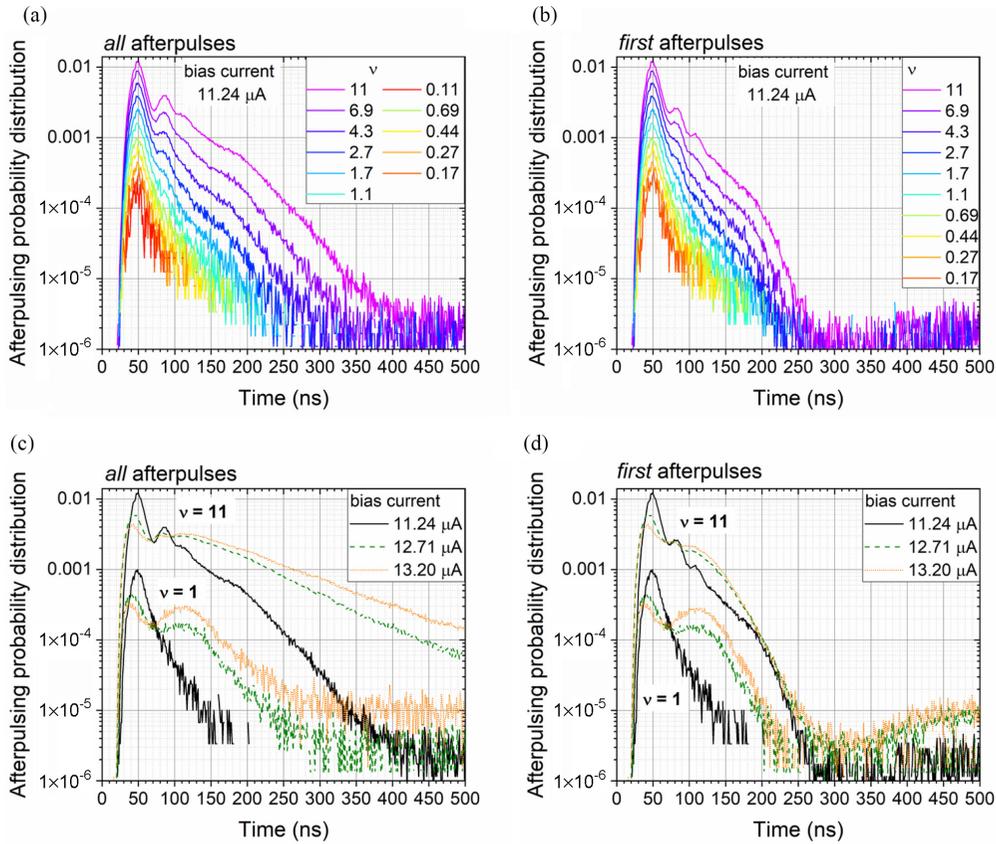
FIG. 3. Afterpulsing probability distributions for [(a),(c)] all afterpulses (i.e., including higher-order afterpulses) and [(b),(d)] first afterpulses. The mean number of photons ν per incoming light pulse as well as the bias currents are indicated.

histograms qualitatively exhibit the same shape, i.e., after the sharp peak corresponding to a light detection event (at zero time), there is an increase in count number due to afterpulsing events followed by a slowly decaying tail. The shape and the height of the afterpulsing count histogram as well as the length of its tail change significantly with the mean photon number of the incoming light pulses and with the bias current. The somewhat larger width of the photon peak as well as the faster decay of the afterpulsing tail in the case of SNSPD 1, operated with the 3 dB attenuator as a shunt, likely indicate a different (smaller) current effectively flowing through SNSPD 1 compared to SNSPD 2, which may be an effect related to the manufacturer's readout or subtle differences from fabrication. Furthermore, SNSPD2 exhibits a considerably longer afterpulsing "tail" compared to SNSPD1. To further investigate this, for SNSPD 1, we compared the decay of the long-term dark count level for different mean photon numbers and measurements taken at repetition rates of 20 and 40 kHz to the (constant) count level of "dark measurements" (i.e., shutter closed) routinely taken for each measurement run. Employing a large binning of 1 µs we found that the level depends significantly on the signal repetition rate (for identical mean photon number) and even for SNSPD 1 remained (very slightly) elevated for signal detection intervals up to about 1 ms. Panels (e) and (f) of Fig. 2 show examples of screenshots of the oscilloscope in persistent mode revealing pulses produced by light counts that are followed by afterpulses, where the early afterpulses exhibit smaller amplitudes.

Figure 3 displays how the distribution of the afterpulsing probability over time changes with the mean number of photons per incoming light pulse ν and with the bias current (note that here and in the following only data from SNSPD 1 are shown and discussed). Panels (a) and (c) correspond to distributions for all registered afterpulses (i.e., including higher-order afterpulses) and panels (b) and (d) for only first afterpulses. The width and the height of the afterpulsing probability distribution increases monotonically with mean photon number ν. For large ν, we find a sequence of arguably three peaks in the distribution of all afterpulses as well as in the distribution of only first afterpulses. These peaks are evidently not due to higher-order afterpulsing. The peaks are followed by a decay, which, as expected, is slower when higher-order afterpulses are included. For small $\nu < 1$, the distributions exhibit only one main peak and are the same for both all afterpulses and only first afterpulses.

As seen in Figs. 3(c) and 3(d), at different bias currents, the dependence of the afterpulsing probability distribution on ν remains, i.e., the height and the width of the distribution increase with ν. When comparing measurements taken at different signal repetition rates of 20 and 40 kHz (not shown here), we find the amplitude of the first afterpulsing peak to be independent of the repetition rate. However, one can see that currents affect the shape of the distributions, which might be the effect of electronics or detector layout [34].

In order to address the effect of electronics, we evaluated the probability distribution of the first afterpulse for fixed ν
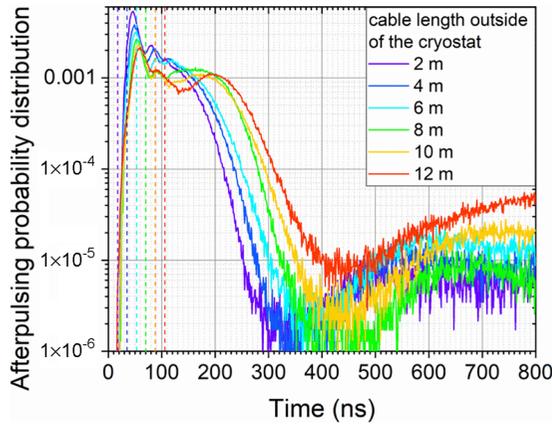
FIG. 4. Probability distributions of the first afterpulse for different cable lengths between SNSPDs and room-temperature amplifiers (see Fig. 1). The vertical dashed lines indicate a conservative estimate of the round-trip time between SNSPDs and amplifiers according to the different cable lengths at room temperature, and based on a measured signal propagation time of 4.4 ns/m.

FIG. 5. Count histograms (here 250 ps bins) of several "pump-and-probe" type measurements, where the initial light pulse is followed by a second delayed light pulse, for a bias current of 11.24 μA. Light-gray curves: Main path blocked, delay path open; red curves: Both paths open. The optical delay t_d is realized by using in the delay path optical fibers of different lengths l_d and their combinations ($l_{d,0}$, $l_{d,0} + 1.5$ m, $l_{d,0} + 2.5$ m, $l_{d,0} + 5.0$ m, $l_{d,0} + 7.5$ m, $l_{d,0} + 12.5$ m).

and bias current and for different cable lengths between the SNSPD and the room-temperature amplifier. We increased the cable length such that the round-trip time of any electrical pulse would be larger than 110 ns, which is far beyond the onset of afterpulsing and the maximum in its probability distribution. The signal propagation time per unit cable length of 4.4 ns/m was measured separately. The results are shown in Fig. 4.

In Fig. 4, vertical dashed lines mark the arrival times of first reflected electrical pulses at the detector for each particular cable length (color encoded). The first falling edge in the probability distribution at around 50 ns delay is not affected by the cable length. Contraryily, the second falling edge at around 200 ns moves to larger delay times as the cable length increases. However, the relative delay between this second falling edge and the arrival time of the first reflected pulse does not depend on the cable length, i.e., on the round-trip time. The second falling edge is followed by the period of the long-term tail of the afterpulsing distribution, which can be seen to be affected by the cable length as well.

At this point, we conclude that although reflections under certain conditions, e.g., for longer cable lengths, may alter the shape of the electrical pulse and also of the afterpulsing probability distribution, they are not the primary cause of the observed afterpulsing.

The other common explanation given in the literature for afterpulsing in a single SNSPD is discharge currents from ac-coupled amplifiers. These lead to a temporary increase in the current through the device that overshoots the bias current ("device current" I_d in [23], Fig. 1) and result in a corresponding temporary increase in the dark count rate. Since the amplitude of the electrical pulses produced by light and dark counts depends on the current through the SNSPD, in the presence of current overshoots one may expect an increase in the amplitude of afterpulses relative to that of the light detection pulses. In contrast to that, Fig. 2(f) shows that the amplitude of afterpulses is smaller and then recovers monotonically, which strongly questions overshoot currents as being the primary
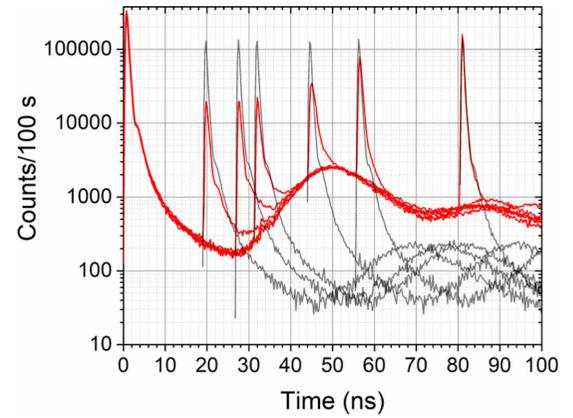
cause of afterpulsing. Actually, the model of [23] assumes high event rates of at least 10% of the inverse recovery time of the detector after a detection event. In our case, where the recovery time is less than 100 ns and the laser repetition rate is 20 kHz, we are several orders of magnitude below these rates.

To further address potential overshoot currents, we also perform a "pump-and-probe" experiment, where the initial light pulse (main path in Fig. 1) is followed by a second one (delay path in Fig. 1) at a variable delay. The results are shown in Fig. 5. Any overshoot in the otherwise exponential current return to the detector is expected to cause a nonmonotonic dependence of the relative detection probability of the second light pulse as a function of the delay time. These relative probabilities correspond to the height of the secondary histogram peaks shown as red curves (both optical paths are open) relative to the black curves (the main path is blocked, while the delay path is open). In contrast to expectations, we see a monotonic increase (recovery) in the detection probability of the second light pulse with the delay time. This holds for all bias currents used here. We should emphasize also that for this pump-and-probe experiment, the round-trip time of an electrical pulse between the amplifier and SNSPD was less than the onset time of afterpulsing (approximately 20 ns). We, therefore, exclude an effect of overshoot currents as well as of the first reflection on the main afterpulsing probability peak and the detection efficiency as a primary cause of afterpulsing. However, given the sparse sampling of delays in the pump-and-probe experiment (Fig. 5) one cannot exclude "ringing" in the return current overlaid on its exponential return, even though Fig. 2 does not indicate it.

V. DISCUSSION

We have shown in the previous section, that back action of the readout circuit is not the primary cause of the observed afterpulsing peak. Furthermore, it is distinct from the slow

decay of the dark count level in that it does not change when increasing the repetition rate.

Hence, this "primary" afterpulsing contribution appears to be a phenomenon distinct from "simple" fluctuations in current (or temperature), and essentially witness the ability to store and after tens of nanoseconds release the energy of photons being absorbed. Therefore the question remains open, as to this primary cause in view of its dependence on the mean photon number per incoming light pulse.

In constructing the microscopic model to follow further below, we concentrate on low mean photon numbers to minimize potential nonlinearities due to multiphoton events as well as a potential crosstalk from the afterpulsing's distribution long tail.

Here, we first note that the probability of the first afterpulse after a light count, which is the integral (sum) over time of a corresponding probability distribution [Fig. 3(b)] increases almost proportionally to the mean number of photons ν in the whole examined range of mean photon numbers [Figs. 6(b) and 8(a)]. Given the Poisson distribution of the probability $P_n(\nu) = e^{-\nu} \nu^n / n!$ that exactly n photons arrive at the detector within a random pulse, for the pulse sequence with $\nu \ll 1$ the majority of light pulses arriving at the detector will contain either one or no photons. Furthermore, pulses with one photon will dominate nonempty pulses. For $\nu = 0.1$ (lower edge of our experimental range), corresponding relative numbers are $P_{n \leqslant 1}(0.1) = 99\%$ and $P_1(0.1)/P_{n \geqslant 1}(0.1) = 90\%$. Assuming that for $\nu \ll 1$ afterpulsing as well as pulse detection are both linear responses, i.e., the probability of an afterpulse after a light count and the probability of pulse detection are constant, we come to the conclusion that the afterpulsing probability normalized to the number of detected pulses, $P_A(\nu)$, should saturate at a constant value. This definitely contradicts our experimental observation of the linear dependence of the normalized probability on ν at small mean photon numbers.

The reason for a $P_A(\nu) \propto \nu$ dependence at $\nu \ll 1$ could be either a significant nonzero correlation between subsequent light counts or nonlinearity of the afterpulsing response, i.e., the dependence of the probability of the first afterpulse on the number of photons per pulse. For our typical experimental conditions, the mean time interval between two subsequent nonempty pulses $(\nu F)^{-1}$ (F is the pulse repetition rate) is a few hundreds of microseconds. Although correlation spanning such time interval is feasible, e.g., via heating of the detector holder, it would cause a dependence of the afterpulsing probability on the pulse repetition rate that we did not observe. Invoking the latter nonlinearity, one has to adopt the discreteness of the afterpulsing response. This means that each absorbed photon may initiate an afterpulse independently on others. In order to implement a first approach for the explanation of our experimental data we extend the electrothermal SNSPD detection model [35–38] for the multiphoton case and introduce a phenomenological model of afterpulsing centers.

The extended electrothermal model relies on the assumption that each detected photon initiates a normal domain in the superconducting strip. Although these domains are thermally independent, their dynamics is correlated via the common current through the strip, which is defined by the readout circuit. Mathematical details and the major results are presented in Appendix A. The extended model showed that the lifetime of
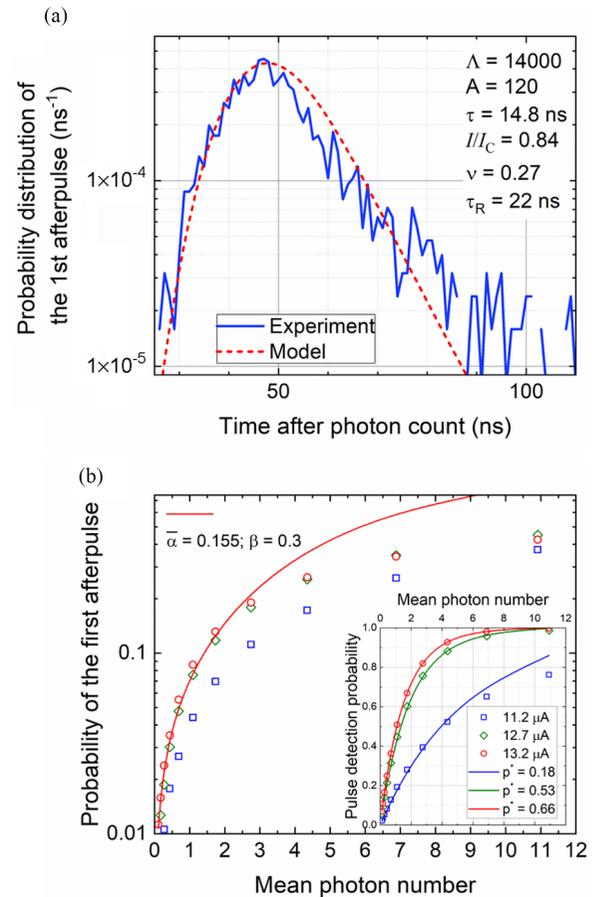

FIG. 6. (a) Probability distribution of the first afterpulse for $\nu = 0.3$ and the model fit (dashed line). (b) Probability of the first afterpulse as a function of the mean photon number per pulse (symbols) for bias currents 11.2 μA (square), 12.7 μA (diamond), and 13.2 μA (circle). The solid lines show a qualitative approximation for the bias current 13.2 μA assuming a $\bar{\alpha}(\nu)$ dependence in Eq. (1). Inset: Pulse detection probability as a function of mean photon number for the same three currents (legend) and the best fits with PDP(ν) [see Eq. (1)], applicable to small ν, where the detector is in the single-photon detection regime. Best-fit values of the probabilities p^* to detect a single photon are indicated in the legend.

domains is one order of magnitude less than the characteristic time of the exponential current return in the strip after a count event. Hence, although afterpulses can be initiated by each domain independently, afterpulsing is a retarded response requiring intermediate energy storage.

For the phenomenological ("microscopic") model of afterpulsing, we assume that each normal domain activates with the probability α of an "afterpulsing center," e.g., in the strip or in the substrate underneath. Each excited center stores a part of the photons' energy released in the form of a burst of thermal phonons. The centers decay into the nonexcited ground states with a relaxation time τ_R. When returning to its ground state, the center releases the stored energy and may initiate a dark count with the probability β. On this level, the shape of the probability distribution can be qualitatively reproduced with our phenomenological model and estimated microscopic parameters, where we note
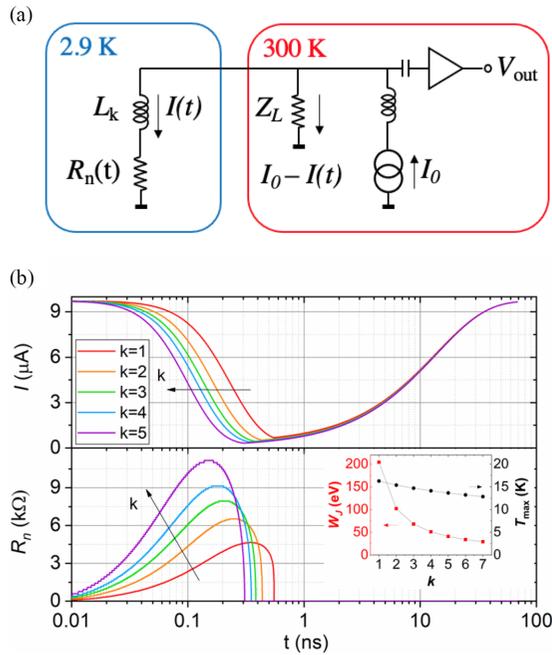
FIG. 7. (a) Equivalent electrical circuit. (b) Modeled current through the device and total resistance (of all domains) for various numbers of independent domains k. Inset: Maximum temperature and dissipated Joule emerge per domain vs the number of domains.
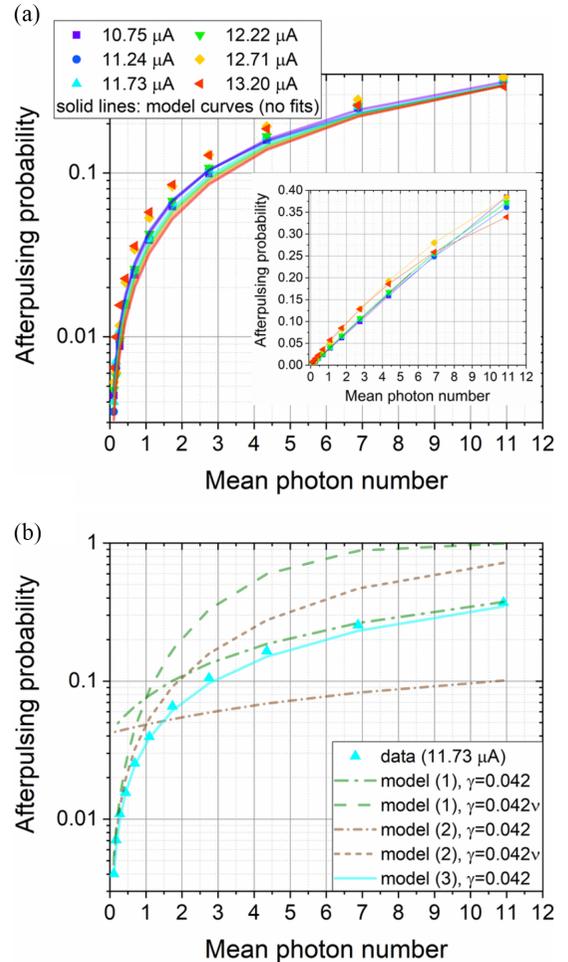

FIG. 8. Probability of observing at least one afterpulse after a detection event. Panel (a): Data for a wide range of bias currents and model curves for $\gamma = 0.042$ and a model assuming an "activation threshold" of one photon (see text) in a semilogarithmic presentation. Inset: Experimental data in a linear presentation. Panel (b): Experimental data for a bias current of 11.7 μA and model curves (see text).

that the dependence of the dark count rate in our SNSPD on the relative current $i = I/I_C$ is perfectly described in the framework of the modified theory of thermally activated phase slips [39] as $\gamma(i) = \Lambda \exp[-A(1-i)^{5/4}]$ with the critical current $I_C = 13.8$ μA, attempt rate $\Lambda = 14\,000$ ns^{-1}, and the activation exponent $A = 120$. Invoking exponential return of the current in the strip $i(t) = i_0[1 - \exp(-t/\tau)]$ with the characteristic time $\tau = 14.8$ ns, one can express the probability distribution, $PD(t)$, of the first afterpulse as $PD(t) \propto \gamma(i)\exp[-\gamma(i)t]\exp(-t/\tau_R)$. In Fig. 6(a), the dashed curve computed according to this expression with a relaxation time of $\tau_R = 22$ ns represents the best fit to the experimental probability distribution of the first afterpulse.

The probability, $P_m(p^*)$, that exactly $m \leqslant n$ photons from n arrived at the detector are detected follows the binomial distribution $P_m(p^*) = n!/[(n-m)!m!](p^*)^m(1-p^*)^{n-m}$ where p^* is the probability for a single photon to be detected. The probability, $P_k(\alpha)$, that exactly k centers are excited and the probability, $P_j(\beta)$, that exactly j centers initiated an afterpulse also follow binomial distributions. The probability for at least one afterpulse to appear, $P_A(\nu)$, is the sum over all $n, m, k, j \geqslant 1$ of the product of the Poisson $P_n(\nu)$ and all three binomial distributions. As predicted by the simple physical treatment above, for $\nu \ll 1$, the linear approximation [40] contradicts experimental data. Mathematical details of the linear approximation and a "black-box model" fitting all ν (under the assumption of an activation threshold of one photon) are presented in Appendix B.

Elaborating nonlinear approximation requires the dependence of the excitation probability α on m. For an arbitrary dependence, the analytic expression in the closed form is not available. We therefore approximated the sum over all indices

by introducing the mean value, $\bar{k} = np^*\bar{\alpha}$, for summations over m and k. With this approximation, the sum over j reduces to $1 - (1-\beta)^{\bar{k}}$ and the normalized probability of a first afterpulse becomes

$$
P_A(\nu) \approx \frac{\Sigma_{n \geqslant 1} P_n(\nu)[1 - (1-\beta)^{np^*\bar{\alpha}}]}{\Sigma_{n \geqslant 1} P_n(\nu)(1 - (1-p^*)^n)}
$$
$$
= \frac{1 - \exp\{\nu[(1-\beta)^{\bar{\alpha}p^*} - 1]\}}{1 - \exp\{-\nu p^*\}}, \tag{1}
$$

where $\bar{\alpha}$ is the weighted average of the excitation probability α over the range of detected photons per light pulse. Note that $p^* < p$ where p is the probability that a single photon is absorbed in the detector. If each absorbed photon may excite an afterpulsing center with the same probability α, p should be used instead of p^* in the numerator. The denominator represents the pulse detection probability PDP(ν), i.e., the probability to detect a light pulse from a random, Poisson-distributed sequence of pulses with the mean photon number pro pulse ν. We used this expression to fit at $\nu \ll 1$ the

experimental pulse detection probability defined as N/K where N is the number of detected pulses from a random sequence with K pulses. The fit returns the probabilities p^* to detect a single photon as a function of the bias current. Here they are 0.18, 0.53, and 0.66 for bias currents 11.2, 12.7, and 13.2 µA, respectively. The experimental data and fits are shown in the inset to Fig. 6(b).

For small ν, $P_A(\nu)$ from Eq. (1) approaches a constant nonzero value not depending on ν. To address experimental observation, we leave the question of whether and how α and also β may depend on the number of simultaneously detected photons to the microscopic picture of excited centers. For the model to agree qualitatively with the experimental data at $\nu < 1$, we let $\bar{\alpha}$ increase with the mean photon number as $\nu^{3/4}$. The solid line in Fig. 6(b) represents the best fit obtained in this way with Eq. (1) for $\nu < 1$ and the current 13.2 µA; values $\bar{\alpha}$ and β used for the fit are indicated in the legend in Fig. 6(b).

Below we specifically consider known two-level systems as potential physical candidates for our phenomenological afterpulsing centers. TLSs are well known to affect the performance of kinetic inductance detectors and destroy coherence of superconducting qubits. These findings rely on the cumulative effect of many TLSs on the macroscopic properties (e.g., resonance frequency) of microscopically large objects (e.g., strip line resonator). Here we tentatively are looking for a local microscopic action of just a single TLS on the superconducting current in the strip. Consider a TLS with a typical level separation [25] corresponding to a resonance frequency in the gigahertz range. Assuming, e.g., $\Omega = 20$ GHz (equivalent temperature 1 K), at our operating temperature of 2.9 K, approximately 40% of TLSs are expected to populate the excited state. We further assume that the energy $h\Omega$ released by a TLS is directly coupled [41] to the superconducting condensate in the effective fluctuation volume $\xi^2 d$ [39] where ξ is the superconducting coherence length and d is the film thickness. The kinetic energy of the condensate in this volume is $\varepsilon = L_{sq}I^2(\xi/w)^2/2$ where L_{sq} and w are the kinetic inductance of the square and the width of the strip. The expected relative change of the current caused by the energy released by TLS is $dI/I = h\Omega/4\varepsilon$, where h is the Planck constant. Using a typical coherence length of 5 nm for amorphous Si-based superconductors and the kinetic inductance of the detector (Appendix A) we obtain an approximate 10% increase in the current that drives the volume above the critical state.

While further experimental work as well as a detailed knowledge of the detector's layout and material is required to further elucidate the afterpulsing centers' nature, our estimate shows that a single TLS may in principle be capable of either causing a deterministic dark count or at least noticeably increasing the effectiveness of intrinsic thermal fluctuations, which are responsible for the background dark count rate [39]. The estimate of the TLS activation efficiency α is a more subtle problem. According to the electrothermal model, the temperature in the center of a normal domain increases up to 15 K (Appendix A), but the Joule energy dissipated per one normal domain decreases quickly with the increase of ν and increases with the current. Assuming that this Joule energy activates TLS results in a decrease in α as a function of ν. On the other hand, the energy released by photons absorbed

within the thermal healing length is proportional to the number of such photons. However, at this point identifying the afterpulsing centers with known TLSs described in the literature to explain the observed storage and release of the absorbed photons' energy has to remain a conjecture awaiting amendment, correction, or confirmation when further measurements and technical information on these and other devices become available.

VI. CONCLUSION

We have investigated afterpulsing in commercial SNSPDs that appear to be particularly sensitive to different sources of afterpulsing, possibly due to their specific layout or material. It appears likely that the detectors' are made of amorphous material such as silicides or highly disordered nitrides. We found that back-action of the electric readout circuit, which is the conventionally accepted cause of afterpulsing, is not sufficient to explain the observed afterpulsing features listed below. We explained these features invoking microscopic centers that capture and retard thermal energy released in the superconducting strip via photon absorption. Specifically, the features are as follows: (i) the afterpulsing probability grows monotonically with the mean number of photons per incoming light pulse in the range of mean numbers ν from 0.1 to 10 and this growth appears to only weakly depend on the detection efficiency over a wide range of bias currents; (ii) the relaxation time of microscopic centers responsible for afterpulsing is 22 ns as obtained from the best model fit of the afterpulsing probability distribution; (iii) while the SNSPD remains in the single-photon detection regime, excitation of afterpulsing centers is an essentially nonlinear process showing the dominance of multiphoton events in the probability of an afterpulse.

As a plausible candidate for the role of afterpulsing centers we bring forward two-level systems as a conjecture. The complete microscopic description of afterpulsing, however, requires detailed knowledge of the properties of TLSs or their possible alternatives for a given detector layout and material, which remains a challenge for future work.

We finally note that our observations underline the need for metrology in quantum technologies, i.e., an independent characterization of commercial "quantum devices" to avoid introducing unnoticed systematic errors into a measurement, e.g., when determining the efficiency of single-photon sources driven with cw excitation or performing assumedly uncorrelated measurements separated by tens or hundreds of nanoseconds or more. At the same time, given a quantitative microscopic understanding and prediction, the combination of detection and afterpulsing probabilities from a metrological viewpoint might open up exciting new avenues in the calibration of single-photon detectors.

APPENDIX A: NORMAL DOMAINS IN THE ELECTROTHERMAL MODEL

We derive microscopic inputs for the model from the measured parameters of SNSPD 1, using values of the relaxation time of the electron energy and of the diffusion coefficient typical for superconducting silicides [42] and numerically simulate the evolution of a normal domain and current in a superconducting strip using the electrothermal model:

$$c_e \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\lambda \frac{\partial T}{\partial x}\right) - K\left(T^q - T_{\text{bath}}^q\right) + \frac{I(t)^2 R_{sq}}{w^2 d},$$

$$L_k \frac{dI}{dt} + I(t)[R_n(t) + Z_L] = Z_L I_0, \qquad (A1)$$

where the solution of the thermal equation, i.e., the temperature distribution $T(x, t)$ along the strip of the width w and thickness d, is coupled to the circuit equation [circuit is shown in Fig. 7(a)] via the time-dependent current $I(t)$ through the strip. In Eqs. (A1), T_{bath} is the bath temperature, R_{sq} is the square resistance, c_e is the electron heat capacitance, L_k is the total kinetic inductance, and Z_0 is the load impedance. The thermal equation follows the treatment in [43], with the exponent $q = 3$. The electron thermal conductivity $\lambda = D c_e(T_C)$ as well as the effective thermal conductance $K = c_e(T_C)/[q\tau_E(T_C)T_C^{q-1}]$ are taken at the transition temperature, T_C, and are the same in the superconducting and in the normal state. Here D is the diffusion coefficient of electrons and τ_E is the relaxation time of the electron energy to the substrate. Equations (A1) were solved by the finite difference (forward Euler) method as an initial value problem. The size of the seed domain was defined via energy conservation, i.e., a photon with the energy 0.8 eV (1550 nm) heats up the electron and phonon subsystems (with energies defined by the Drude and Debye models) up to temperature T_C. For simplicity, the temperature-dependent resistance of the strip was modeled as a step function, $R(T) = 0$ below T_C and $R(T) = R_{sq}l/w$ above T_C. The transition between the superconducting and normal states was controlled using the Ginzburg-Landau depairing current $I_{\text{dep}}(T)$, i.e., if $I \geqslant I_{\text{dep}}(T)$, the state is normal, or otherwise superconducting with zero resistance.

To estimate the relevant parameters of the studied device, we measured critical temperature (5.7 K), device resistance in the normal state (2.3 MΩ), switching current (12.5 μA, accounting for the rf attenuator), and the electrical signal without amplifier [exponential decay with a time constant of 14.8 ns; the lower trace in Fig. 2(b)]. We assume that the device is a MoSi-based meander with $w = 100$ nm, $l = 700$ μm, $d = 5$ nm, $L_k = 670$ nH, $R_{sq} = 350$ Ω, $D = 0.5$ cm^2/s, and $\tau_E = 140$ ps. Here, we assumed that it has a circular layout with a filling factor of 50% and a diameter of 13 μm that is typically used for commercial fiber-coupled devices for telecommunication wavelengths. With these parameters,

we numerically solve Eqs. (A1) for a various number of independent domains created simultaneously Fig. 7(b), semilogarithmic scale]. As the domain evolves, the current through the device decreases being diverted into the circuit. When the domain disappears, the current reaches its minimum value and slowly recovers returning back to the strip. The inset to Fig. 7(b) shows the maximum temperature in the domain center vs the number of domains k, and the Joule energy dissipated in one domain $W_J = k^{-1} \int I(t)^2 R_n(t) dt$ vs the number of domains. We estimate that the total dissipated energy elevates the temperature of the device by $\Delta T = W_J k/(c_v V) = 5$ nK, where the phonon heat capacitance per chip volume at 2.9 K is about $c_v \approx 6$ J/K m^3 and the chip volume $V \approx 1.5 \times 10^{-9}$ m^3. Here we assumed that the substrate material is silicon with a typical thickness of about 300 μm and that the chip diameter is ≈2.5 mm (the studied device is coupled to the optical fiber as described in [44]).

From the above analysis, we conclude the following: (i) domain lifetimes do not exceed 1 ns, (ii) the current recovery time does not depend on the number of created domains, (iii) dissipated Joule energy per domain decreases with the number of domains as $W_J \propto 1/k$, (iv) its value is two orders of magnitude larger than the photon energy (1550 nm, 0.8 eV), and (v) it does not lead to any noticeable increase in the chip temperature.

APPENDIX B: "BLACK-BOX" MODEL

Here we introduce a phenomenological description of the experimental probability of observing at least one afterpulse treating the detector as a "black box." Figure 8 displays the afterpulsing probability given a detection event over a wide range of bias currents (from 10.75 to 13.20 μA) [panel (a)], correponding to detection efficiencies from ∼10% to ∼65%. The probabilities here were computed by integrating (summation) the corresponding probability distributions, after subtracting the respective distribution's long-term mean value as a background correction [45]. Over a wide range of currents the probabilities exhibit a very similar and almost linear dependence on the mean number of photons ν per incoming light pulse, in particular for bias currents of 11.73 μA and less. A simple linear fit to the data yields a mean slope γ_0 of around 0.035, which may be interpreted as the linear approximation to an "afterpulsing efficiency" γ for converting a photon into an afterpulse. The actual nonlinearity of the probability curve is clearly visible for the largest currents, but is present at all currents. Taking into account the bounding of the afterpulsing probability (first afterpulse) to values <1, the probability rather handwavingly may be approximated by $1 - \exp(-\gamma\nu)$, where a fit to the data yields a mean value of $\gamma \approx 0.042$. While this may serve as a rough back-of-the-envelope estimate of the afterpulsing efficiency, we now set out to systematically develop a phenomenological black-box model in detail.

As a first approach we straightforwardly assume that afterpulses are due to the detected fraction $p^*\nu$ of the incoming photons. In general, the probability of an afterpulse is then the probability of a detection event (given a Poissonian input distribution $\sum_n e^{-\nu}\nu^n/n!$, where the mean photon number is reduced to $p^*\nu$) times the probability that a photon

triggers an afterpulse with probability ("efficiency") γ. For a given photon number n in a pulse, this can be expressed by $1 - (1 - \gamma)^n$, where $(1 - \gamma)^n$ is the probability that a dark count (afterpulse) is initiated. In this black-box model we make no specific assumptions on the microscopic processes inside the SNSPD other than a mutual independence of each photon's effect. Multiplying the probabilities for a given n and summing over all n, after proper normalization yields the expression (model 1)

$$
\begin{aligned}
P_A(\nu, p^*, \gamma) &= \frac{\exp(-p^*\nu)}{1 - \exp(-p^*\nu)} \sum_{n=0}^{\infty} \frac{(p^*\nu)^n}{n!}[1 - (1 - \gamma)^n] \\
&= \frac{1 - \exp(-p^*\nu\gamma)}{1 - \exp(-p^*\nu)}.
\end{aligned} \tag{B1}
$$

The data for a bias current of 11.73 µA as well as the model curve (brown short-dash-dotted line; $\gamma = 0.042$; $p^*(\nu)$ calculated from the pulse detection efficiency) are shown in Fig. 8(b). We note that also for other values of γ, this model does not reproduce the observed shape of the probability of observing at least one afterpulse well. If instead assuming that *all* absorbed photons contribute to initiating an afterpulse, we obtain the following expression (here assuming an absorbance of 100%; model 2):

$$
\begin{aligned}
P_A(\nu, p^*, \gamma) &= \frac{\exp(-\nu)}{1 - \exp(-p^*\nu)} \sum_{n=0}^{\infty} \frac{\nu^n}{n!}[1 - (1 - p^*)^n] \\
&\quad \times [1 - (1 - \gamma)^n] \\
&= 1 - \exp(-\gamma\nu)\frac{1 - \exp[-p^*\nu(1 - \gamma)]}{1 - \exp(-p^*\nu)}. \tag{B2}
\end{aligned}
$$

The green (upper), dash-dotted line shows the model curve ($\gamma = 0.042$). While for large incoming mean photon numbers the model now approximates the data quite well, this is not the case for small mean photon numbers. Up to now we have assumed the "black-box afterpulsing efficiency" γ to be a constant. On the other hand, the energy released by a

single absorbed photon might not be sufficient to lead to the initiation of an afterpulse, which might require a "collaborative" behavior of more than one photon absorption. As our data are not taken with a single-photon source but with attenuated laser light, this assumption would introduce a dependence of the afterpulsing efficiency on the mean photon number at least for small ν, where photon number states with $n = 1$ become dominant over the fraction of all other states where $n > 1$. For illustration, we tentatively replace γ by $\gamma\nu$ in Eqs. (B1) and (B2), yielding the dashed model curves in panel (b), respectively. As can be seen from the figure, under this assumption both models describe the data considerably better for small mean photon numbers. In particular, the second model, where all absorbed photons contribute to the probability of observing at least one afterpulse, phenomenologically describes the data quite well under the assumption of a linear dependence on the mean photon number for small ν, and a transition region between a photon number dependent and independent afterpulsing efficiency at $\nu \approx 1$.

From these considerations, we now dispense again with the rather arbitrary direct linear dependence on ν, but only assume that all photons absorbed in the SNSPD may contribute to afterpulsing and that the process has an activation threshold corresponding to one photon. This yields (model 3)

$$
\begin{aligned}
P_A(\nu, p^*, \gamma) &= \frac{\exp(-\nu)}{1 - \exp(-p^*\nu)} \sum_{n=0}^{\infty} \frac{\nu^n}{n!}[(1 - (1 - p^*)^n] \\
&\quad \times [1 - (1 - \gamma)^{(n-1)}] \\
&= 1 - \frac{\exp(-\gamma\nu)}{1 - \gamma} \frac{1 - \exp[-p^*\nu(1 - \gamma)]}{1 - \exp(-p^*\nu)}. \tag{B3}
\end{aligned}
$$

The resulting model curves for $\gamma = 0.042$ are shown in Fig. 8(a). We find that even without fitting γ, for all ν model (3) qualitatively describes the experimental data quite well but systematically underestimates the probability of observing at least one afterpulse slightly. This underestimate increases with increasing p^*, implying that the overall afterpulsing efficiency γ actually increases with the bias current.

[1] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, Single-photon sources and detectors, Rev. Sci. Instrum. **82**, 071101 (2011).

[2] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li *et al.*, Satellite-to-ground quantum key distribution, Nature (London) **549**, 43 (2017).

[3] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020).

[4] J. M. Arrazola, V. Bergholm, K. Brádler, T. R. Bromley, M. J. Collins, I. Dhand, A. Fumagalli, T. Gerrits, A. Goussev, L. G. Helt *et al.*, Quantum circuits with many photons on a programmable nanophotonic chip, Nature (London) **591**, 54 (2021).

[5] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins *et al.*, Quantum computational advantage with

a programmable photonic processor, Nature (London) **606**, 75 (2022).

[6] K. Niwa, T. Numata, and D. Fukuda, Few-photon color imaging using energy-dispersive superconducting transition-edge sensor spectrometry, Sci. Rep. **7**, 45660 (2017).

[7] C. Bruschini, H. Homulle, I. M. Antolovic, and E. Charbon, Single-photon avalanche diode imagers in biophotonics: Review and outlook, Light Sci. Appl. **8**, 87 (2019).

[8] F. I. Khatri, B. S. Robinson, M. D. Semprucci, and D. M. Boroson, Lunar laser communication demonstration operations architecture, Acta Astronaut. **111**, 77 (2015).

[9] *DOCOMAS Deep Space Optical Communications Architecture Study—Executive Summary*, ESOC Contract No. 4000106720/12/F/MOS (2016).

[10] S. J. Smith, Multiabsorber transition-edge sensors for x-ray astronomy, J. Astron. Telesc. Instrum. Syst. **5**, 021008 (2019).

[11] Y.-H. Deng, H. Wang, X. Ding, Z. C. Duan, J. Qin, M. C. Chen, Y. He, Y. M. He, J. P. Li, Y. H. Li *et al.*, Quantum Interference

between Light Sources Separated by 150 Million Kilometers, Phys. Rev. Lett. **123**, 080401 (2019).

[12] J. W. Appel, C. L. Bennett, M. K. Brewer, R. Bustos, M. Chan, D. T. Chuss, J. Cleary, J. D. Couto, S. Dahal, R. Datta *et al.*, Calibration of transition-edge sensor (TES) bolometer arrays with application to CLASS, Astrophys. J. Suppl. Ser. **262**, 52 (2022).

[13] N. Bastidon, D. Horns, and A. Lindner, Characterization of a transition-edge sensor for the ALPS II experiment, arXiv:1509.02064.

[14] C. Spiess and F. Steinlechner, Clock synchronisation with pulsed single photon sources, arXiv:2212.12589.

[15] G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, and A. Dzardanov, Picosecond superconducting single-photon optical detector, Appl. Phys. Lett. **79**, 705 (2001).

[16] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, Superconducting nanowire single-photon detectors: Physics and applications, Supercond. Sci. Technol. **25**, 063001 (2012).

[17] M. Sidorova, A. Semenov, H.-W. Hübers, A. Kuzmin, S. Doerner, K. Ilin, M. Siegel, I. Charaev, and D. Vodolazov, Timing jitter in photon detection by straight superconducting nanowires: Effect of magnetic field and photon flux, Phys. Rev. B **98**, 134504 (2018).

[18] A. D. Semenov, Superconducting nanostrip single-photon detectors some fundamental aspects in detection mechanism, technology and performance, Supercond. Sci. Technol. **34**, 054002 (2021).

[19] J. Chang, J. W. N. Los, J. O. Tenorio-Pearl, N. Noordzij, R. Gourgues, A. Guardini, J. R. Zichi, S. F. Pereira, H. P. Urbach, V. Zwiller *et al.*, Detecting telecom single photons with 99.5% system detection efficiency and high time resolution, APL Photon. **6**, 036114 (2021).

[20] M. Fujiwara, A. Tanaka, S. Takahashi, K. Yoshino, Y. Nambu, A. Tajima, S. Miki, T. Yamashita, Z. Wang, A. Tomita *et al.*, Afterpulse-like phenomenon of superconducting single-photon detector in high speed quantum key distribution system, Opt. Express **19**, 19562 (2011).

[21] F. Marsili, F. Najafi, E. Dauler, R. J. Molnar, and K. K. Berggren, Afterpulsing and instability in superconducting nanowire avalanche photodetectors, Appl. Phys. Lett. **100**, 112601 (2012).

[22] V. Burenkov, H. Xu, B. Qi, R. H. Hadfield, and H.-K. Lo, Investigations of afterpulsing and detection efficiency recovery in superconducting nanowire single-photon detectors, J. Appl. Phys. **113**, 213102 (2013).

[23] A. J. Kerman, D. Rosenberg, R. J. Molnar, and E. A. Dauler, Readout of superconducting nanowire single-photon detectors at high count rates, J. Appl. Phys. **113**, 144511 (2013).

[24] S. Miki, M. Yabuno, T. Yamashite, and H. Terai, Stable, high-performance operation of a fiber-coupled superconducting nanowire avalanche photon detector, Opt. Express **25**, 6796 (2017).

[25] C. Müller, J. H. Cole, and J. Lisenfeld, Towards understanding two-level-systems in amorphous solids: Insights from quantum circuits, Rep. Prog. Phys. **82**, 124501 (2019).

[26] M. López, H. Hofer, and S. Kück, Detection efficiency calibration of single-photon silicon avalanche photodiodes traceable using double attenuator technique, J. Mod. Opt. **62**, 1732 (2015).

[27] S. M. F. Raupach, I. P. Degiovanni, H. Georgieva, A. Meda, H. Hofer, M. Gramegna, M. Genovese, S. Kück, and M. López, Detection rate dependence of the inherent detection efficiency in single-photon detectors based on avalanche diodes, Phys. Rev. A **105**, 042615 (2022).

[28] Indicating the name of any manufacturer or supplier by PTB or DLR in this scientific journal shall not be taken to be PTB's or DLR's endorsement of specific samples of products of the said manufacturer, or recommendation of the said supplier.

[29] C. Cahall, K. L. Nicoloch, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, Multi-photon detection using a conventional superconducting nanowire single-photon detector, Optica **4**, 1534 (2017).

[30] While an inspection of the overall layout of the cryogenic electronics board showed it not to match the information provided by the manufacturer, we were able to identify its shunt resistor (possibly preceded by an inductance) and to measure its dc resistance at cryogenic temperatures (66Ω).

[31] We also performed a measurement of the switching current by directly connecting the sourcemeter (via a *RC* low-pass filter of nom. 720 Hz on the current force line) to the detector in a four-wire measurement (both with the detector mount grounded to the cryostat and taking care to electrically isolate the detector from the cryostat ground). There we found a critical current (resistance step) at an input current of around 2.5 μA and thermal runaway setting in either directly (no *RC* filter) or at around 3.4 μA (with *RC* filter), as witnessed by the detector becoming fully normally conducting. Therefore, compared to other SNSPDs, the shunt resistance might play a more prominent role as a current divider to set the working point of the device given a certain input current.

[32] F. Marsili, F. Najafi, E. Dauler, F. Bellei, X. Hu, M. Csete, R. J. Molnar, and K. K. Berggren, Single-photon detectors based on ultranarrow superconducting nanowires, Nano Lett. **11**, 2048 (2011).

[33] M. Sidorova, A. Semenov, H.-W. Huebers, I. Charaev, A. Kuzmin, S. Doerner, and M. Siegel, Physical mechanisms of timing jitter in photon detection by current-carrying superconducting nanowires, Phys. Rev. B **96**, 184504 (2017).

[34] To address the effect of a nonconventional detector layout, e.g., a layout based on parallel nanowires [24], we performed preliminary simulations for this case neglecting quantization of magnetic flux. The simulations show a slight increase in the dark count level due to current imbalances in the parallel wires. Notably, the imbalance magnitude remained constant during recursive simulations, i.e., when the imbalanced currents derived from previous runs served as input values. As such, this leads to an elevated permanent level of afterpulsing that cannot explain the primary aftepulsing peak. The electrothermal model describes the evolution of a normal domain and currents in two parallel superconducting strips via solution of coupled differential equations, i.e., the heat balance equation in Eq. (A1) and an appropriate circuit equation; the model will be published elsewhere.

[35] A. D. Semenov, P. Haas, B. Günther, H.-W. Hübers, K. Il'in, M. Siegel, A. Kirste, J. Beyer, D. Drung, T. Schurig, and A. Smirnov, An energy-resolving superconducting nanowire photon counter, Supercond. Sci. Technol. **20**, 919 (2007).

[36] A. Semenov, P. Haas, H.-W. Hübers, K. Ilin, M. Siegel, A. Kirste, D. Drung, T. Schurig, and A. Engel, Intrinsic quantum efficiency and electro-thermal model of a superconducting nanowire single-photon detector, J. Mod. Opt. **56**, 345 (2009).

[37] A. J. Kerman, J. K. W. Yang, R. J. Molnar, E. A. Dauler, and K. K. Berggren, Electrothermal feedback in superconducting nanowire single-photon detectors, Phys. Rev. B **79**, 100509(R) (2009).

[38] J. K. Yang, A. J. Kerman, E. A. Dauler, V. Anant, K. M. Rosfjord, and K. K. Berggren, Modeling the electrical and thermal response of superconducting nanowire single-photon detectors, IEEE Trans. Appl. Supercond. **17**, 581 (2007).

[39] A. D. Semenov, M. Sidorova, M. A. Skvortsov, A. Kuzmin, K. Ilin, and M. Siegel, Local thermal fluctuations in current-carrying superconducting nanowires, Phys. Rev. B **102**, 184508 (2020).

[40] Linear means that (i) the mean number of microscopic centers activated by many absorbed photons is proportional to the number of absorbed photons and the proportionality coefficient, α, does not depend on the number of photons; (ii) the mean number of initiated afterpulses is proportional to the number of activated centers and the proportionality coefficient, β, does not depend on the number of activated centers.

[41] H. le Sueur, A. Svilans, N. Bourlet, A. Murani, L. Bergé, L. Dumoulin, and P. Joyez, Microscopic charged fluctuators as a limit to the coherence of disordered superconductor devices, arXiv:1810.12801.

[42] A. Banerjee, L. J. Baker, A. Doye, M. Nord, R. M. Heath, K. Erotokritou, D. Bosworth, Z. H. Barber, I. MacLaren, and R. H. Hadfield, Characterisation of amorphous molybdenum silicide (MoSi) superconducting thin films and nanowires, Supercond. Sci. Technol. **30**, 084010 (2017).

[43] M. Sidorova, A. D. Semenov, H. W. Huebers, S. Gyger, and S. Steinhauer, Phonon heat capacity and self-heating normal domains in NbTiN nanostrips, Supercond. Sci. Technol. **35**, 105005 (2022).

[44] A. J. Miller, A. E. Lita, B. Calkins, I. Vayshenker, S. M. Gruber, and S. W. Nam, Compact cryogenic self-aligning fiber-to-detector coupling with losses below one percent, Opt. Express **19**, 9102 (2011).

[45] We note that for the largest mean photon number and in particular for the largest currents of the long afterpulsing "tail" this subtraction may lead to a slight underestimate of the afterpulsing probability.