Detecting spins with a microwave photon counter

Emanuele Albertinale\textsuperscript{1}, Léo Balembois\textsuperscript{1}, Eric Billault\textsuperscript{1}, Vishal Ranjan\textsuperscript{2}, Daniel Flanigan\textsuperscript{1}, Thomas Schenkel\textsuperscript{3}, Daniel Estève\textsuperscript{4}, Denis Vion\textsuperscript{1}, Patrice Bertet\textsuperscript{5}, Emmanuel Flurin\textsuperscript{6}

\textsuperscript{1}Université Paris-Saclay, CEA, CNRS, SPEC, 91191 Gif-sur-Yvette Cedex, France
\textsuperscript{2}National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, UK
\textsuperscript{3}Accelerator Technology and Applied Physics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

(Dated: February 3, 2021)

Quantum emitters respond to resonant illumination by radiating electromagnetic fields. A component of these fields is phase-coherent with the driving tone, while another one is incoherent, consisting of spontaneously emitted photons and forming the fluorescence signal. Atoms and molecules are routinely detected by their fluorescence at optical frequencies, with important applications in chemistry, biology, and medicine. It is only recently that the detection of single photons has been extended to the microwave range \cite{1, 2, 3, 4, 5, 6}, and circuit quantum electrodynamics \cite{9, 13–18}. Here, we demonstrate the detection of a small ensemble of donor spins in silicon by their fluorescence at microwave frequency and millikelvin temperatures. We enhance the spin radiative decay rate by coupling them to a high-quality-factor and small-mode-volume superconducting resonator \cite{8}, and we connect the device output to a newly-developed microwave single photon counter \cite{9} based on a superconducting qubit. We discuss the potential of fluorescence detection as a novel method for magnetic resonance spectroscopy of small numbers of spins.

Microwave measurements at the quantum limit have recently become possible thanks to the development of superconducting parametric amplifiers that linearly amplify a signal at cryogenic temperatures with minimal added noise \cite{10}. These advances enable efficient measurement of the field quadratures $X$ and $Y$ of a given microwave mode, as needed for qubit readout in circuit quantum electrodynamics \cite{11}. Even then, the signal-to-noise ratio remains ultimately limited by vacuum fluctuations enforced by Heisenberg uncertainty relations, imposing that the quadrature standard deviations satisfy $\delta X = \delta Y = 1/2$. As a result, a linear amplifier is not well suited for detecting fluorescence signals consisting of a few incoherent photons emitted randomly over many modes. In contrast, such signals are ideally detected by a photon counter: because it measures in the energy eigenbasis, it is in principle noiseless when the field is in vacuum and only clicks for photons incoming within its detection bandwidth \cite{12}.

Operational Single Microwave Photon Detectors (SM-PDs) have been developed only recently based on cavity and circuit quantum electrodynamics \cite{9, 13–18}. Here, we report the first use of such a SMPD for sensing applications, to detect spin fluorescence at microwave frequencies. Consider an ensemble of $N$ electron spins $1/2$, resonantly coupled with a spin-photon coupling constant $g_0$ to a resonator of frequency $\omega_0$ and linewidth $\kappa$ (Fig. 1). After being excited by a $\pi$ pulse, the spins will relax exponentially into their ground state with a characteristic time $T_1$. If their dominant relaxation channel is radiative (the so-called Purcell regime \cite{8}), they will do so by spontaneously emitting $N$ microwave photons at the Purcell rate $\Gamma_p = 4g_0^2/\kappa = T_1^{-1}$. Whereas a linear amplifier can only detect these incoherent photons as a slight increase of noise above the background \cite{19, 20}, a single-photon counter with sufficient bandwidth is expected to detect each of them as a click occurring at a random time and revealing an individual spin-flip event.

For our demonstration, we use the electronic spins of an ensemble of bismuth donors implanted about 100 nm below the surface of a silicon chip enriched in the nuclear-spin-free silicon 28 isotope. These spins couple magnetically to the inductor of a superconducting LC resonator with frequency $\omega_0/2\pi = 6.94$ GHz patterned in aluminium on the surface of the chip (see Fig. 1). Applying a static magnetic field $B_0 \approx 17$ mT parallel to the inductor tunes the lowest transition frequency of the bismuth donors in resonance with the resonator. At this field, the energy loss rate $\kappa_i$ in the resonator is 3.5 times higher than the energy leak rate $\kappa_e$ in the measuring line, yielding a total resonator bandwidth $\kappa/2\pi = 0.68$ MHz with $\kappa = \kappa_i + \kappa_e$. As the electron spin of donors in silicon have hour-long spin-lattice relaxation times at low temperatures \cite{21}, they easily reach the Purcell regime when coupled to micron-scale superconducting resonators \cite{8}. For our sample parameters, we measure a spin relaxation time $T_1 = 300 \pm 10$ ms (see Sec. 3 in Methods) dominated by the radiative contribution. In our experiment [see Fig 1(a)], the resonator coupled to the spins has a single input-output port connected through a circulator (and coaxial cables) to both the line used to drive the spins and to a SMPD.
Figure 1: **Principle of spin detection with a photon counter.** (a) Schematics of the experiment. Each spin in the ensemble is coupled with a strength $g_0$ to a resonator of angular frequency $\omega_0$ and internal loss rate $\kappa_i$, itself coupled with a rate $\kappa_c$ to an input-output line. This line allows to drive the spins with microwave pulses through a circulator and to route the photons emitted by the spins while relaxing radiatively towards a Single Microwave Photon Detector (SMPD). (b) Spin device schematics. The spins are bismuth donors implanted in a silicon substrate isotopically enriched in the $^{28}$Si isotope, on top of which a superconducting LC resonator is patterned. A magnetic field $B_0$ is applied parallel to the substrate to tune the lowest bismuth donor frequency in resonance with $\omega_0$ at $B_0 \sim 17$ mT. (c) SMPD device and operation schematics. The SMPD relies on a transmon qubit at frequency $\omega_q/2\pi = 6.13$ GHz, coupled to three ports, namely buffer, pump, and waste. The buffer port consists of a resonator at frequency $\omega_b(\phi)$ that can be tuned to $\omega_0$ by applying a flux $\phi$ to a SQUID loop inserted in it (see inset). The waste port consists of a resonator at a fixed frequency $\omega_w/2\pi = 7.63$ GHz. Incoming photons at the buffer are converted into excitation of the transmon qubit and into a photon in the waste resonator by a four-wave mixing process enabled by a pump tone sent via the pump port at frequency $\omega_p = \omega_q + \omega_w - \omega_b$. The SMPD operation consists of a three-step cycle of total duration $11.7 \mu$s which can be repeated continuously. The Reset (R) step is achieved by driving the waste port and the pump. The Detection step (D) is achieved by switching off the waste port, and keeping the pump on. The Measurement step (M) is performed by switching off the pump, and sending a microwave pulse on the buffer port to perform dispersive qubit readout. The SMPD efficiency $\eta_d$ and dark count rate $\alpha$ are indicated in the inset for relevant values of $\phi$ in the inset.

This SMPD (see Fig. 1c) consists of a superconducting circuit with a transmon qubit [22] of frequency $\omega_q$ capacitively coupled to two coplanar waveguide resonators: a 'buffer' resonator whose frequency $\omega_b$ can be tuned to $\omega_0$ by applying a magnetic flux to an embedded superconducting quantum interference device (SQUID) [23], and a 'waste' resonator with fixed frequency $\omega_w$. As described in Ref. [9], the detection of a photon (here at frequency $\omega_b$) relies on the irreversible excitation of the transmon when driven by a non-resonant pump tone at frequency $\omega_p = \omega_q + \omega_w - \omega_b$. The SMPD is cycled continuously, each cycle consisting of three steps (see Fig. 1c). First, a reset step (R), during which the qubit is set to its ground state by turning on the pump (violet pulse) while applying to the waste resonator a weak resonant coherent tone (green pulse). Second, a detection step (D) that
starts when the microwave at $\omega_m$ is switched off, while the pump is kept on: a photon possibly entering the buffer gets mixed with the pump through a four-wave mixing process that triggers both the excitation of the transmon and the creation of a photon in the waste; this photon is lost in the 50 $\Omega$ port of the waste, which guarantees the irreversibility of the detection and the mapping of the incoming photon into a transmon excitation. The third step (M) is the measurement of the transmon state using the dispersive shift [24] of the buffer resonator (orange pulse). At $\omega_b = \omega_m$, the probability to detect a click when one photon reaches the detector during the detection window $D$ (intrinsic photon detector efficiency) is measured to be $\eta = 0.53 \pm 0.1$ (mainly limited by the transmon energy relaxation, see Methods), and the rate of false positive detection, referred to as the the dark count rate, is $\alpha = 1.53$ clicks/ms. The detection duty cycle (step $D$ duration over total cycle duration) is $\eta_{\text{duty}} = 0.43$, with a complete cycle lasting 11.7 µs. Note that the detector evidently saturates for signals having more than 1 photon every 10 µs, approximately. The detector bandwidth $\Delta \omega/2\pi \approx 2$ MHz is larger than the spin resonator linewidth $\kappa/2\pi$, implying that there is no filtering of the photons emitted by the spins.

As a first experiment, we measure the spontaneous emission of the spin ensemble: a $\pi$-pulse inverts the spin population, so that $\langle S_x \rangle = N/2$, and $\langle S_z \rangle = \langle S_y \rangle = 0$, $S_{x,y,z} = \sum_i S_{x,y,z}^{(i)}$ being the sum of the individual dimensionless spin operators $S^{(i)}$ (Fig. 2a). Since $\langle S_{x,y} \rangle = 0$, the coherent part of the output field also satisfies $\langle X \rangle = \langle Y \rangle = 0$. On the other hand, by energy conservation, a flux of incoherent photons is emitted at a rate $(X^2+Y^2)-1/2$ proportional to $-\partial_t \langle S_x \rangle$ (see Fig. 2a), forming the spin fluorescence signal and triggering clicks in the SMPD. This signal decays back to 0 within the spin relaxation time $T_1$. Note that the $\pi$ pulse perturbs the SMPD during a dead-time of $\sim 200$ µs, after which it can be used normally.

One measurement record consists of $2 \times 10^5$ consecutive SMPD detection cycles, spanning a total measurement time of 2s. Each cycle yields one binary outcome $c(t_i)$, $t_i$ being the time around which cycle $i$ is centered. Figure 2a displays an example of a measurement record at early (0 s to 0.1 s) and at late times (1 s to 1.1 s): due to the emission by the spins, more counts are observed in the 0 s to 0.1 s interval than in 1 s to 1.1 s. Repeating the measurement 500 times and histogramming the number of counts, we obtain the average count rate $\langle \hat{c}(t_d) \rangle$ as a function of the delay $t_d$ after the $\pi$ pulse. Figure 2b shows this rate with and without $\pi$-pulse applied. Without pulse, a constant $\langle \hat{c}(t_d) \rangle = 1.53$ counts/ms is recorded, which corresponds to the dark count rate $\alpha$ of the SMPD. With $\pi$-pulse, $\langle \hat{c}(t_d) \rangle$ shows an excess of 0.85 counts/ms exponentially decaying towards 0, with a fitted time constant of 309 ms. Because this time is the same as the independently measured spin relaxation time $T_1$ (see Methods), we conclude that the SMPD detects the photons spontaneously emitted by the spins upon relaxation. Note that at short $t_d$ we moreover observe an extra excess rate of 0.3 counts/ms decaying with a 20 ms time constant (see inset of Fig. 2b), which we attribute to the radiative relaxation of spurious two-levels-systems. (c) Measured probability distribution of the number of counts $C$ integrated from 46.8 ms to 585 ms, obtained for 500 repetitions of the experiment when a $\pi$ pulse is either applied (magenta) or not (blue) to the spins. Solid lines represent Poissonian fits.

Figure 2: Detection of spin relaxation by photon counting. (a) Pulse sequence and spin dynamics. The applied pulse amplitude $A_p$, transverse and longitudinal magnetizations ($S_x$) and $\langle S_z \rangle$, output field quadrature ($X$), and photon count rate $\langle \hat{c} \rangle$ are shown as a function of the delay $t_d$ after the $\pi$ pulse. Solid curves are sketches of the expected dynamics. The transit time during the pulse, $\langle S_z \rangle$ quickly goes to 0 because of the spin ensemble inhomogeneous broadening; $\langle S_z \rangle$ on the other hand is inverted and relaxes towards equilibrium in a characteristic time $T_1 = 0.3$ s. During this time, $\langle X \rangle$ is also 0, whereas a flux of spontaneous photons (the spin fluorescence) $\langle X^2 + Y^2 \rangle - 1/2$ is emitted. The lower panel shows one typical experimental time trace of the detected clicks as a function of $t_d$. (b) Average count rate $\langle \hat{c}(t_d) \rangle$ measured in 19 ms time bins in the case where a $\pi$ pulse is (magenta) or is not (blue) applied to the spins. An exponential fit for $t_d > 40$ ms (solid line) leads to the characteristic time $T_1 = 300$ ms. The observed excess rate at short times $t_d < 50$ ms (inset) has been investigated and attributed to fast-relaxing two-levels-systems. (c) Measured probability distribution of the number of counts $C$ integrated from 46.8 ms to 585 ms, obtained for 500 repetitions of the experiment when a $\pi$ pulse is either applied (magenta) or not (blue) to the spins. Solid lines represent Poissonian fits.
Figure 3: Detection of spin echo by photon counting. (a) Pulse sequence and spin dynamics. A Hahn echo pulse sequence (solid orange line) consisting of a $\pi/2$ pulse, a delay $\tau$, and a $\pi$ pulse is applied to the spin ensemble, causing a revival of the spin transverse magnetisation $\langle S_z \rangle$ at time $2\tau$, and the emission of a coherent microwave echo on the $X$ quadrature. After the $\pi/2$ pulse, the spin population $\langle S_z \rangle$ slowly decays by spontaneous emission with a characteristic time $T_1$, with in addition a small drop during the echo. Being coherent, the echo can be detected by homodyne detection or by a SMPD. Solid curves are sketches of the expected dynamics. An example of an experimental sequence is shown in the bottom, with a click detected at the echo time. (b) Average number of counts $\langle c \rangle$ in 23 µs bins, as a function of the time $t$ from the beginning of the echo sequence, averaged over 83 sequences. Blue-shadowed areas represent the 200 µs blind time of the detector after each strong spin pulse. The increased count probability at $t_i = 2\tau$ is the spin-echo. (c) Average probability $p(c)$ of having one or no count in the time bin centred at echo time $t_i = 2\tau$ (orange) or in one of the subsequent bins $2\tau < t_i < 1$ ms (blue). The difference between click probabilities (dashed lines) lead to a signal $c_e$ of 0.3 photons re-emitted coherently.

and without $\pi$ pulse, an average of $\langle C(\pi) \rangle = 1050$ and $\langle C(0) \rangle = 900$ counts are detected, the difference defining the spin signal $C_{\text{spin}} = 150$ photons. The ratio of this signal to the total number of excited spins $N$ defines an overall detection efficiency $\eta = C_{\text{spin}}/N$. For a Poissonian distribution, one expects the width of $p(C)$ to be $\delta C = \sqrt{\alpha t_w}$ and $\delta C = \sqrt{\alpha t_w + \eta(1 - \eta)N}$ without and with $\pi$ pulse, respectively. In our case, because $\alpha t_w \gg \eta N$, both distributions have approximately the same width $\delta C \approx 30$ counts, dominated by the dark counts fluctuations contribution.

It is interesting to note that the signal-to-noise ratio $\eta N/\sqrt{\alpha t_w + \eta(1 - \eta)N}$ can in principle become arbitrar-
coherent detection.

In our demonstration of microwave photon echo detection, the echo duration is shorter than the detection cycle, and only one photon at most can be detected at the echo time $2\tau$; we thus center the detection step $D$ of the SMPD at $2\tau$. We also chose $\tau = 350\mu$s, larger than the detector dead time. A typical photo-counting trace is visible in Fig. 3a, showing in particular one click at the expected echo time. Repeating several echo sequences yields $\langle c(t) \rangle$ (see Fig. 3d), clearly showing an excess of counts for $t = 2\tau$.

The click probability histogram is shown in Fig. 3 at and out of the echo time. The average number of detected photons during the spin-echo, $c_{\text{echo}} = \langle c(2\tau) \rangle - \langle c(t_i > 2\tau) \rangle = 0.3$, is as expected much lower than $C_{\text{spin}}$, the number of photons detected in the spontaneous emission experiment of Fig. 2. The standard deviation $\delta c_{\text{echo}} = 0.46$ during the echo (Fig. 3) yields a signal-to-noise ratio $c_{\text{echo}}/\delta c_{\text{echo}} = 0.65$, significantly lower than the one obtained with the spontaneous emission method, although both measurements were performed with the same repetition time $\sim 2T_1$ and thus also the same initial spin polarization.

We finally demonstrate that SMPD detection can be used to perform usual spin characterisation measurements. First, spin spectroscopy is performed by varying the magnetic field $B_0$ around the resonance value and using the three different detection methods already mentioned: homodyne detection with echo, SMPD fluorescence detection, and SMPD detection of the echo. As seen in Fig. 4a, all three methods give similar spectra. Second, we observe Rabi rotations in the fluorescence signal. In Fig. 4b, the spin signal $C_{\text{spin}}$ is plotted as a function of the spin driving pulse duration $\tau$. Oscillations are observed, with a frequency linearly dependent on the pulse amplitude, reflecting the Rabi oscillations of $\langle S_z \rangle$ since $C_{\text{spin}} = \eta[1 + 2 \langle S_z \rangle]/2$. This Rabi oscillations can also be measured with the microwave photon echoes method, by varying the duration $\tau$ of the refocusing pulse; the SMPD signal $\langle c(2\tau) \rangle$ shows the expected Rabi oscillations (Fig. 4c). The oscillation contrast in Figs. 4b and c diminishes with pulse duration $\tau$ due to the spread of Rabi frequencies in the ensemble. This inhomogeneity has a different impact on the spontaneous emission signal and on the echo signal [30], which is quantitatively reproduced by simulations (see Methods), as seen in Figs. 4b-c. Finally, the spin coherence time is measured by microwave photon echo detection. In Fig. 4b, $\langle c(2\tau) \rangle$ is plotted as a function of $\tau$. An exponential fit to the data yields $T_2 = 2.7\text{ms}$, in agreement with the value measured using homodyne detection (Sup. Mat.). Overall, this demonstrates that SMPD detection can be used to perform standard ESR spectroscopy measurements.

Our results represent the first use of a SMPD for quantum sensing. Beyond the fundamental interest of such a proof of principle, we conclude by discussing the potential of SMPDs for spin detection. Whereas the SNR of echo detection by a SMPD (Fig. 3) and by homodyne detection are comparable, spin fluorescence detection by a SMPD (Fig. 2) on the other hand presents several features that make it a truly interesting method for ESR spectroscopy. Indeed, the SNR can reach much higher values than in echo-detection, since the signal (number of emitted photons) can be as high as the total number of excited spins, whereas the noise is entirely dominated by SMPD non-idealities, which are likely to be improved in future devices [31–33]. We therefore expect that the development of better SMPDs with lower dark count rates and higher efficiency will push further ultra-sensitive spin detection, possibly down to a single spin. This perspective is all the more interesting that the method applies equally well to spins with short coherence times such as encountered in real-world spin systems, making practical single-spin ESR spectroscopy a possible future perspec-
tive.

Acknowledgements

We acknowledge technical support from P. Sénat, D. Duet, P.-F. Orfila and S. Delprat, and are grateful for fruitful discussions within the Quantronics group. This project has received funding from the European Unions Horizon 2020 research and innovation program under Marie Sklodowska-Curie Grant Agreement No. 765267 (QuSCO). E.F. acknowledges support from the ANR Marie Sklodowska-Curie Grant Agreement No. 765267 (REIMIC project).

Author contributions

E.A., P.B. and E.F. designed the experiment. T.S. provided the bismuth-implanted isotopically purified silicon sample, on which V.R. fabricated the Al resonator. E.A. designed and fabricated the SMPD with the help of D.V. and E.F.. E.A., E.V., D.E. and E.F. contributed useful input to the manuscript.

D.F., D.V., D.E. and E.F. contributed useful input to the simulations. E.A., P.B. and E.F. wrote the manuscript. E.A., E.B. and V.R. performed the measurements, with help from L.B., D.F. and P.B. E.A., P.B. and E.F. analysed the data. E.A., E.B. and V.R. performed the simulations. E.A., P.B. and E.F. wrote the manuscript. D.F., D.V., D.E. and E.F. contributed useful input to the manuscript.

E. A., P.B. and E.F. designed the experiment. T.S. provided the bismuth-implanted isotopically purified silicon sample, on which V.R. fabricated the Al resonator. E.A. designed and fabricated the SMPD with the help of D.V. and E.F. E.A., E.V., E.F. performed the measurements, with help from L.B., D.F. and P.B. E.A., P.B. and E.F. analysed the data. E.A., E.B. and V.R. performed the simulations. E.A., P.B. and E.F. wrote the manuscript. D.F., D.V., D.E. and E.F. contributed useful input to the manuscript.
METHODS

1. SETUP

The experimental setup used to drive the spin ensemble and operate the SMPD, with six input lines (labeled 1-6) and one output line (labeled 7), is shown in Fig. 5. We first discuss the room-temperature part (microwave and dc signals generation), and then the low-temperature part (cabling inside the dilution refrigerator).

Room temperature setup

The room-temperature setup includes four microwave sources and two 4-channel arbitrary waveform generators (AWG 5014 from Tektronix). All microwave pulses needed in the experiment are generated by mixing the output of a source with 2 AWG channel outputs used to drive the I and Q ports of an I/Q mixer at an intermediate frequency indicated in Fig. 6. The pulses are used to drive the spins (at the spin resonator frequency \( \omega_0 \)) and to operate the SMPD.

SMPD operation requires:
- a dc flux-bias of the SQUID in the buffer resonator, in order to tune \( \omega_b \) in resonance with \( \omega_0 \). This is achieved with a dc current source (Yokogawa 7651) connected to an on-chip antenna near the SQUID (line 4 in Fig. 5).
- microwave pulses at the pump frequency \( \omega_p \) to satisfy the 4-wave mixing condition \( \omega_p = \omega_q + \omega_w - \omega_b \).
- microwave pulses to readout the qubit state via the qubit-state-dependent dispersive shift of the buffer resonator. They are at frequency \( \omega_b + \chi_{qb} \), the buffer resonator frequency with qubit in the \( e \) state.
- microwave pulses at the waste frequency \( \omega_w \) to reset the qubit.

Moreover, qubit readout pulses are amplified by a flux-pumped Josephson Parametric Amplifier (JPA) in degenerate mode [34]. The JPA needs dc flux biasing to adjust the JPA frequency; it is provided by an on-chip antenna near the JPA SQUID array, fed by a constant voltage source biasing a resistor at room-temperature (line 1 in Fig. 5). The JPA also requires flux-pumping to achieve gain on the signal-bearing quadrature. The pump tone is generated by frequency-doubling the same source used to generate the readout pulses (line 2 in Fig. 5), followed by mixing with an intermediate frequency (see Fig. 6). The relative phase between signal and pump is adjusted with a phase shifter for maximum gain on the signal-bearing quadrature.

The same source (Keysight MWG, shown in yellow in Fig. 5) is used for driving the spins, qubit state readout, JPA pumping, and as local oscillator for signal demodulation yielding the quadratures of the qubit readout pulses. Spin driving pulses and qubit state readout pulses are sent via the same line (line 3 in Fig. 5). Spin driving pulses require much larger powers than qubit readout pulses. Therefore, in the room-temperature setup, the line was split before recombination, and in one of the branches an amplifier was inserted in-between two microwave switches.

A second source (Vaunix Labbrick, shown in green in Fig. 5) is used for the qubit reset pulses at \( \omega_w \) (line 5). A third source (Keysight, shown in purple in Fig. 5) is used for SMPD pumping. Pump pulses are generated through I/Q mixing and amplification of the generator output. The signal is then passed through a 70 MHz band-pass filter to prevent spurious wave mixing caused by side-band resonances and LO leakage, before reaching the cryostat input on line 6. A fourth source (Vaunix Labbrick, shown in blue in Fig. 5) is used for SMPD tuning and characterization.

Low-temperature setup

Line 3 is heavily attenuated at low-temperatures in order to minimize spurious excitations of the transmon qubit in the SMPD and therefore dark counts (see Fig. 5). It is then connected to the spin resonator input via a double circulator. The reflected signal is routed by the same circulator towards the SMPD input (buffer resonator), and the signal reflected on the SMPD is finally routed towards the input of the JPA and the detection chain. Two double circulators isolate the SMPD from the JPA, to minimize noise reaching the SMPD and potentially causing spurious qubit excitations and dark counts. The JPA output (reflected signal) is routed to a High-Electron-Mobility-Transistor (HEMT) amplifier from Low-Noise Factory anchored at the 4K stage of the cryostat, and then to output line 7. Infrared filters are inserted on all the lines leading to the SMPD to minimize out-of-equilibrium quasi-particle generation leading to spurious qubit excitations and dark counts. To minimize heating of the low-temperature stage by the strong pump tone of the SMPD, the necessary attenuation of the pump line at 10 mK is achieved with a 20 dB directional coupler that routes most of the pump power towards the 100 mK stage where it is dissipated.

Using the same line both for spin excitation and SMPD readout raises potential issues that are now discussed. First, the spin excitation pulse also leads to a large field build-up in the buffer resonator (since \( \omega_b = \omega_0 \)), which excites the qubit and perturbs the proper functioning of the SMPD during a time that we quantify to be 200 \( \mu \)s (detector dead-time). Then, one may also wonder about spurious excitation of the spins caused by the repeated qubit readout pulses. This is avoided, because qubit readout is performed at \( \omega_b + \chi_{qb} \), which is thus shifted from \( \omega_0 \) by \( \chi_{qb}/2\pi = -3.5 \) MHz.
Figure 5: Schematic of the setup.
2. FABRICATION

Spin sample

The bismuth donors are implanted in a 700 nm epilayer of $^{28}$Si-enriched silicon. The implantation profile ranges from 50 to 150nm depth, with a peak concentration of 8 x 10$^{16}$ donors/cm$^3$ (see refs [28] [35] for more details). The spin resonator consists of an interdigitated capacitor, shunted by a 1μm-wide, 450μm-long inductive wire, with a design similar to the one used in [28]. It is deposited on top of the silicon sample by evaporation of a 50 nm thick aluminium film through a resist mask patterned using e-beam lithography, followed by liftoff. The chip is then placed inside a 3D copper cavity into which a pin protrudes, controlling the capacitive coupling to the input line [28]. A superconducting coil applies an in-plane magnetic field $B_0$ to tune the spin frequency.

Single microwave photon detector

The single photon detector circuit is based on the design by Lescanne et al. [9]. It is fabricated using wet etching of a 60nm aluminium layer evaporated on a high-resistivity intrinsic silicon substrate. Before metal deposition, the substrate is pre-cleaned with a SC1 process. The wafer is first immersed for 10 min at 80°C in a bath of 5 parts H$_2$O to 1 part H$_2$O$_2$ (30%) to 1 part NH$_4$OH (29%), then is immersed for 2 min in HF (5%) solution to remove the surface oxide. The substrate is then loaded in an electron-beam evaporator within 10 minutes, after which a 60 nm aluminium layer is deposited. Patterning of the circuit is achieved by electron beam lithography of a UV3 resist mask, followed by wet etching of the aluminium using a TMAH-based developer (Microposit CD26). The Josephson junctions are evaporated using the Dolan bridge technique and re-contacted to the main circuit through aluminium bandage patches [30]. Finally, circuit gaps are isotropically trenched with a SF6-based reactive ion etch, which has shown to decrease the internal losses of superconducting resonators [37] [38]. The resulting 10 mm by 3 mm chip is glued and wired to a Printed-Circuit-Board, placed in a copper box, magnetically shielded, and attached to the cold stage of the dilution refrigerator.

3. CHARACTERISATION

Electron spin resonance spectroscopy by homodyne measurements

Prior to the experiments reported in the main text, the spin ensemble is characterised by pulsed electron spin resonance spectroscopy, comparable to previous work [28]. Interestingly this can be done in the same cooldown as the SMPD measurements reported in the main text, because of the fact that the SMPD setup also includes a Josephson Parametric Amplifier (JPA) for qubit state readout. To switch from single photon detection to homodyne spin measurements, we simply tune the buffer resonator frequency $\omega_b$ at a frequency far from $\omega_0$, and tune the JPA at resonance with $\omega_0$. In that way, the spin-echo signal simply reflects off the SMPD without triggering any qubit excitation, and gets amplified by the JPA, exactly as was achieved in similar experiments [28]. Output signal demodulation then yields the spin-echo quadrature and its integral $A_e$.

We measure the spin relaxation time $T_1$ at $B_0 = 17$ mT with an inversion recovery sequence, in which a $\pi$ pulse is first applied, followed after a duration $\tau$ by a Hahn-echo detection sequence. The echo area $A_e$ is shown as a function of $\tau$ in Fig. [3], together with an exponential fit yielding $T_1 = 300 \pm 10$ ms. We also measure the spin coherence time by measuring the echo amplitude as a function of the delay $2\tau$ between the $\pi/2$ pulse and the echo (see Fig. [6]). An exponential fit yields $T_2 = 2.7$ ms. Rabi nutations are obtained by measuring the echo area $A_e$ as a function of the refocusing pulse amplitude $A$ (see Fig. [6]). Finally, the bismuth donor spin spectrum is obtained by recording the echo amplitude $A_e$ as a function of the field $B_0$ (see Fig. [3]).

The data in Fig. [6] are modelled using a simulation tool described elsewhere [30]. It computes the evolution of the spin ensemble under the application of driving pulses at the resonator input. The spread in spin Larmor frequency (due to strain-induced inhomogeneous broadening) and in spin-photon coupling (due to the spatial inhomogeneity of the $B_1$ field generated by the resonator) are taken into account by describing the spin as an ensemble of packets, with coupling constant density $\rho(g_0)$ and frequency density $\rho_{\text{spin}}(\omega)$. The evolution of each packet is computed independently under the drive pulses, and the echo response is obtained by summing the packet contributions. Purcell relaxation is also taken into account [30].

Here, we make two extra simplifying assumption. Because the inhomogeneous broadening is much larger than the cavity linewidth $\kappa$, and that the signal originates essentially from spins within this linewidth, we consider the spin density to be constant, $\rho_{\text{spin}}(\omega) = \rho_{\text{spin}}$. Moreover, we model the coupling constant inhomogeneity as a Gaussian centered on $g_0$ and width $\delta g_0$, $\rho(g_0) = \frac{1}{\delta g_0 \sqrt{2\pi}} e^{-\frac{(g_0-g_0)^2}{2\delta g_0^2}}$. We adjust the values of $g_0$ and $\delta g_0$ to get good agreement with the relaxation and Rabi nutation data in Fig. [3] and d, yielding $g_0/2\pi = 290$ Hz, and $\delta g_0/2\pi = 25$ Hz. Because the spin density $\rho_{\text{spin}}$ only rescales the signal amplitude in Fig. [6] its determination requires other measurements that are described below, enabling us to infer the number of excited spins and the
overall photon detection efficiency.

Single photon detector characterisation and tuning

The single microwave photon detector consists of a transmon qubit whose ground state and first excited state encode the detector click. The transmon frequency is $\omega_q/2\pi = 6.14\text{GHz}$ and its anharmonicity is $-200\text{MHz}$. It is capacitively coupled to a tunable buffer resonator (maximum frequency $\omega_b^{\text{max}}/2\pi = 7.09\text{GHz}$) and a waste resonator $(\omega_w/2\pi = 7.62\text{GHz})$ with dispersive shifts $\chi_{qb}/2\pi = -3.5\text{MHz}$ and $\chi_{qw}/2\pi = -8.1\text{MHz}$ respectively. The buffer resonator is coupled to the external microwave line via a capacitance (energy damping rate $\kappa_b = 13.7 \times 10^6 \text{s}^{-1}$ at the working point $\omega_b/2\pi = 6.946\text{GHz}$), while the waste resonator is coupled through a Purcell filter to a $50\ \Omega$-terminated line (energy damping rate $\kappa_w = 2.8 \times 10^6 \text{s}^{-1}$).

Detector tuning

In order to perform spin detection the SMPD must be tuned in resonance with the spin-emitted photons. This is achieved by changing the magnetic flux threading a superconducting quantum interference device (SQUID) embedded in the input resonator of the detector, which allows a tunability range of about $200\text{MHz}$ (see fig. 7a and Fig1 of the main text). The photon detector will be now characterised in the vicinity of this working point.
the ratio $p(\epsilon|0)/p(\epsilon|\pi)$ between the false positive and true positive, therefore optimising the dark count rate with respect to the efficiency. We measure a ground state fidelity of $1 - p(\epsilon|0) = 99.2\%$ and an excited state fidelity of $p(\epsilon|\pi) = 71\%$

Efficiency calibration

The first characterisation of the device as a photon detector consists in tuning the 4-wave mixing process by sending a weak coherent tone at the detector frequency while scanning the pump frequency and amplitude. When the 4-wave mixing matching condition is satisfied the qubit is left in its excited state, as shown in Fig. 8a. The chosen working point is the one maximising the probability of the qubit excitation while minimising the residual qubit excitation due to pump heating or spurious process. Note that the frequency drift for the matching condition corresponds to the qubit AC-stark shift induced by the pump tone increasing amplitude.

The detector efficiency is measured by sending a weak coherent tone onto the buffer resonator with $n$ photons on average, and measuring the probability $p_e$ of finding the qubit in its excited state as function of time $T$ after the reset sequence, showing out-of-equilibrium qubit excited population reaching thermal equilibrium on a timescale $\approx T_1$. Black dashed line at $t_D = 5\,\mu s$ marks the point at which detector is operated. (d) Measured (blue dots) and fitted (orange solid line) detector efficiency $\eta_d$ as a function of the duration of the detection step $t_D$. Fit model takes into account bandwidth-limited detection efficiency for short detection windows and $T_1$-decay effect for increasing $t_D$. Black dashed line at $t_D = 5\,\mu s$ marks the point at which detector is operated, to optimise the photo-detected echo signal.

Qubit readout

The qubit readout is based on standard dispersive readout through the buffer resonator. A tone is sent at the buffer resonance frequency pulled by the qubit dispersive shift $\omega_b + \chi_{qb}$ (thus avoiding to drive the spin resonator at $\omega_b = \omega_h$). The qubit state is encoded in the phase of the reflected tone which is subsequently amplified, demodulated and numerically integrated. The readout performances are evaluated by histogramming this reflected signal conditioned on the application of a $\pi$ pulse in Fig. 8b and c. We observe two Gaussian distributions separated by $10\sigma$ corresponding to the qubit being in the ground and excited state, the spurious tail between the Gaussian corresponds to relaxation events of the qubit during the readout time. A threshold enables the discrimination of the qubit state. It is chosen to minimise

Detector bandwidth

As long as the qubit lies in its ground state, the detector response (Fig. 7d) can be modelled by considering that the buffer and waste resonator are coupled with a constant $G = -\xi_p \sqrt{\chi_{qb} \chi_{qw}}$ due to the 4-wave parametric process involving the qubit, where $\xi_p$ is the pump amplitude in units of square root of photons and $\chi_{qb}(\chi_{qw})$ the dispersive coupling of the buffer (waste) resonator to the qubit. One can write down the system of coupled

\[
\begin{align*}
\text{(a) } & & \text{(b) } & & \text{(c) } & & \text{(d) }
\end{align*}
\]
equations for the buffer and waste intra-resonator fields $\alpha$ and $\beta$:

$$\dot{\alpha} = -i\delta_b \alpha - i\mathcal{G} \beta - \frac{\kappa_b}{2} + \sqrt{\kappa_b} \alpha_{in}$$

(1)

$$\dot{\beta} = -i\delta_w \beta - i\mathcal{G}^* \alpha - \frac{\kappa_w}{2} + \sqrt{\kappa_w} \beta_{in}$$

(2)

where $\delta_b$ and $\delta_w$ are the buffer and waste frequencies in the frame rotating at the probing frequency and $\alpha_{in}$, $\beta_{in}$ are the respective input field amplitudes. Now using the relation between the intra-resonator fields and the input and output flux $\sqrt{\kappa_b} \alpha = \alpha_{in} + \alpha_{out}$, from the equilibrium solution of the coupled system we can extract the transmission coefficient $|S_{21}|^2 = |\beta_{out}/\alpha_{in}|^2$. Assuming zero input flux on the waste this leads to:

$$|S_{21}|^2 = \left| \frac{2\xi \sqrt{\kappa_b \kappa_w \chi_b \chi_w}}{-4\delta_b \delta_w + 2i\delta_b \kappa_b + 2i\delta_w \kappa_b + \kappa_b \kappa_w + \chi_b \chi_w \xi_p^2} \right|^2$$

This expression can be directly related to the detector efficiency when varying the input photon frequency. Figure 7: show a fit of this expression to experimental data with only $\xi_p$ and a scale factor as free parameters. From the curve we extract a bandwidth of $\Delta_{det}/2\pi \approx 2.1$ MHz.

### Dark Counts

A key figure of merit of the detector is its dark count rate. We characterise this quantity by applying a reset pulse to the qubit through the waste resonator and by keeping the pump tone turned on while no photon pulse is sent to the buffer resonator. By varying the duration of the pump tone, we observe an increasing excited state population as shown in Fig 8. The residual qubit population rises with a slope of 1.7 ms$^{-1}$ from an initial value of $0.35 \times 10^{-2}$. The qubit reaches a finite population of $1.77 \times 10^{-2}$ after a few characteristic time $T_1 = 8.1$ $\mu$s. Note that the qubit is initialised well below its thermal population by the reset process. This finite population can be divided in distinct contributions. In the absence of the pump, we measure a residual excited state population of the qubit of $0.81 \times 10^{-2}$ tone which is attributed to out-of-equilibrium quasi-particles in the superconducting film [41]. By detuning the pump from the matching condition, the heating effect of the pump alone can be evaluated. We measure a negligible rise of the excited population, smaller than $10^{-3}$, compared to the population in absence of pumping. Therefore, most of the excess qubit population $\delta n_e \sim 0.8 \times 10^{-2}$ can be attributed to the finite temperature of the buffer line. Such a finite thermal occupancy $n_{th, buffer}$ triggers the detector over its full bandwidth $\Delta_{det}$ and leads to dark counts that are integrated over the qubit lifetime $T_1$. The expected rise of qubit population is thus given by $\delta n_e = n_{th, buffer} T_1 / \tau_{det}$ which gives a thermal occupancy of the buffer line of $n_{th, buffer} \sim 1.5 \times 10^{-4}$ that corresponds to a residual temperature for the microwave line of 27 mK.

### 4. ESTIMATION OF THE NUMBER OF SPINS

The goal of this section is to explain how we determine the number of spins $N$ excited in the spontaneous emission detection experiment, enabling us to quantify the overall photon detection efficiency $\eta$. We use two independent methods, using two different datasets. The first method was already used in previous work [35, 42]. It relies on measurements performed with homodyne detection, without any use of the SMPD. We measure a complete spin-echo sequence (including the
control pulses), and we use simulations to fit the data. The ratio of spin-echo to control pulse amplitude allows fitting the spin density $\rho_{\text{spin}}$ needed to account for the data as the only adjustable parameter.

Results are shown in Fig. 9a. Quantitative agreement is obtained for $\rho_{\text{spin}} = 14.6 \text{ spins kHz}^{-1}$. We can then determine $N$, the total number of spins excited by a $\pi$ pulse in the spontaneous emission detection experiment of Fig. 2 in the main text, by running a second dedicated simulation with spin density $\rho_{\text{spin}}$. We obtain $N = 14.6 \times 10^3$.

The second method uses a comparison between the two different types of measurements performed with the SMPD. On the one hand, the number of photons $C_{\text{spin}}$ detected in the spontaneous emission experiment scales linearly with $N$. On the other hand, the echo signal $c_e$ scales like $N^2$, as explained in the main text, indicative of the phase-coherent character of echo emission. Therefore, the ratio $C_{\text{spin}}/c_e$ can give access to $N$. Note that because both datasets are obtained with the same SMPD and setup, $C_{\text{spin}}/c_e$ is independent of $\eta$, as required for a proper evaluation of $N$.

To make this reasoning quantitative, we again resort to simulations of both the echo sequence and the $\pi$ pulse. From each simulation, we extract the number of spins involved by computing the change in the total magnetization $\langle S_z \rangle$. The ratio of these two numbers should be equal to the experimentally determined $C_{\text{spin}}/(\eta \text{duty} \cdot c_e)$ (the $\eta \text{duty}$ correction is due to the fact that echo detection is gated and therefore is insensitive to the detector duty cycle); we use the spin density $\rho_{\text{spin}}$ as the only adjustable parameter to reach the agreement.

For the experimental value of $C_{\text{spin}}$ we use the data shown in Fig. 2 of the main text. For $c_e$ we use different pulse parameters than shown in Fig. 3 of the main text (same parameters for the $\pi/2$ pulse, but lower amplitude and 5.5$\mu$s duration for the $\pi$ pulse), to get a lower value of $c_e$ and minimize the risk of SMPD saturation. The experimental ratio is then $C_{\text{spin}}/(\eta \text{duty} \cdot c_e) = 3 \cdot 10^3$. Figures Fig. 9b and 9c show the evolution of $\langle S_z \rangle$ at early times in the case of a Hahn echo sequence and of a $\pi$-pulse respectively, with the pulse parameters used in the experiment. The correct ratio is reproduced for $\rho_{\text{spin}} = 12 \text{ spins kHz}^{-1}$. This yields $N = 1.2 \times 10^4$.

The two methods are in agreement within an estimated uncertainty $\sigma_{\text{spin}} \approx 10^3$ on the number of spins taking part to the process. We take the average of the two values $N = (13.3 \pm 1.1) \times 10^3$ as the reference value for the efficiency estimation. From this, an overall collection efficiency $\eta = 1.1 \times 10^{-2}$ is obtained, as explained in the main text.