Experimental Demonstration of Memory-Enhanced Quantum Communication

Experimental demonstration of memory-enhanced quantum communication

M. K. Bhaskar[1,*], R. Riedinger[1,*], B. Machielse[1,*], D. S. Levonian[1,*], C. T. Nguyen[1,*],

E. N. Knall[2], H. Park[1,3], D. Englund[4], M. Lončar[2], D. D. Sukachev[1], M. D. Lukin[1,†]

[1] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*

[2] *John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

[3] *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA*

[4] *Research Laboratory of Electronics, MIT, Cambridge, MA 02139, USA*

The ability to communicate quantum information over long distances is of central importance in quantum science and engineering[1]. While some applications of quantum communication such as secure quantum key distribution (QKD)[2,3] are already being successfully deployed[4–7], their range is currently limited by photon losses and cannot be extended using straightforward measure-and-repeat strategies without compromising unconditional security[8]. Alternatively, quantum repeaters[9], which utilize intermediate quantum memory nodes and error correction techniques, can extend the range of quantum channels. However, their implementation remains an outstanding challenge[10–16], requiring a combination of efficient and high-fidelity quantum memories, gate operations, and measurements. Here we use a single solid-state spin memory integrated in a nanophotonic diamond resonator[17–19] to

implement asynchronous photonic Bell-state measurements, a key component of quantum repeaters. In a proof-of-principle experiment, we demonstrate high-fidelity operation that effectively enables quantum communication at a rate that surpasses the ideal loss-equivalent direct-transmission method while operating at megahertz clock speeds. These results represent a significant step towards practical quantum repeaters and large-scale quantum networks[?,20].

Efficient, long-lived quantum memory nodes are expected to play an essential role in extending the range of quantum communication[9], as they enable asynchronous quantum logic operations, such as Bell-state measurements (BSM), between optical photons. Such an asynchronous BSM is central to many quantum communication protocols, including the realization of scalable quantum repeaters[9] with multiple intermediate nodes. Its elementary operation can be understood by considering a specific implementation of quantum cryptography[21,22] illustrated in Fig. 1a. Here two remote communicating parties, Alice and Bob, try to agree on a key that is secure against potential eavesdroppers. They each send a randomly chosen photonic qubit $\{|\pm x\rangle, |\pm y\rangle\}$ encoded in one of two conjugate bases (X or Y) across a lossy channel to an untrusted central node (Charlie), who performs a BSM and reports the result over an authenticated public channel. After a number of iterations, Alice and Bob publicly reveal their choice of bases to obtain a correlated bit string (sifted key) from the cases when they used a compatible basis. A potentially secure key can subsequently be distilled provided the BSM error rate is low enough. While a photonic BSM can be implemented with linear optics and single photon detectors, in this "direct-transmission" approach, the BSM is only successful when photons from Alice and Bob arrive simultaneously. Thus, when Alice and Bob are separated by a lossy fiber with a total transmission probability $p_{\mathrm{A}\rightarrow\mathrm{B}} \ll 1$, Charlie measures photon coincidences with probability also limited by $p_{\mathrm{A}\to\mathrm{B}}$, leading to a fundamental bound[8] on the maximum possible distilled key rate of $R_{\max} = p_{\mathrm{A}\to\mathrm{B}}/2$ bits per channel use for an unbiased basis choice[4]. While linear optical techniques to circumvent this bound are now being actively explored[23], they offer only limited improvement and cannot be scaled beyond a single intermediate node. Alternatively, this bound can be surpassed using a quantum memory node at Charlie's location. In this approach, illustrated in Fig. 1b, the state of Alice's photon is stored in the heralded memory while awaiting receipt of Bob's photon over the lossy channel. Once the second photon arrives, a BSM between Alice's and Bob's qubits yields a distilled key rate that for an ideal memory scales as[24] $R_s \propto \sqrt{p_{\mathrm{A}\to\mathrm{B}}}$, potentially leading to substantial improvement over direct transmission.

This Letter describes the operation of a quantum node that enables BSM rates that exceed those of an ideal system based on linear optics. We focus on the demonstration and characterization of the BSM node, leaving the implementation of source-specific technical components of full-scale QKD systems, such as decoy states[25], basis biasing[26], a finite key error analysis[27], and a physical separation of Alice and Bob for future work. Our realization is based on a single silicon-vacancy (SiV) color-center integrated inside a diamond nanophotonic cavity[17–19] (Fig. 2a). Its key figure-of-merit, the cooperativity[13] C, describes the ratio of the interaction rate with individual cavity photons compared to all dissipation rates. A low mode volume ($0.5(\lambda/n)^3$), high quality factor (2×10^4), and nanoscale positioning of SiV centers enable an exceptional $C = 105 \pm 11$. Cavity photons at $737\,\mathrm{nm}$ are critically coupled to a waveguide and adiabatically transferred into a single-mode optical fiber[18] that is routed to superconducting nanowire single-photon detectors, yielding a full system detection efficiency of about 85% (Methods). The device is placed inside a dilution refrigerator, resulting in electronic spin quantum memory[19] time $T_2 > 0.2\,\mathrm{ms}$ at temperatures below $300\,\mathrm{mK}$.

The operating principle of the SiV-cavity based spin-photon interface is illustrated in Fig. 2. Spin dependent modulation of the cavity reflection at incident probe frequency f_0 (Fig. 2b) results in the direct observation of electron spin quantum jumps (Fig. 2c, inset), enabling nondestructive single-shot readout of the spin state (Fig. 2c) in $30\,\mu\mathrm{s}$ with fidelity $F = 0.9998^{+0.0002}_{-0.0003}$. Coherent control of the SiV spin qubit ($f_Q \approx 12\,\mathrm{GHz}$) is accomplished using microwave fields delivered via an on-chip gold coplanar waveguide[19]. We utilize both optical readout and microwave control to perform projective feedback-based initialization of the SiV spin into the $|{\downarrow}\rangle$ state with a fidelity of $F = 0.998 \pm 0.001$. Spin-dependent cavity reflection also enables quantum logic operations between an incoming photonic time-bin qubit, defined by a phase-coherent pair of attenuated laser pulses, and the spin memory[19,28]. We characterize this by using the protocol illustrated in Fig. 2d to generate the spin-photon entangled state $(|e\uparrow\rangle + |l\downarrow\rangle)/\sqrt{2}$ conditioned on successful reflection of an incoming single photon with overall heralding efficiency $\eta = 0.423 \pm 0.004$ (Methods). Here, $|e\rangle$ and $|l\rangle$ denote the presence of a photon in an early or late time-bin separated by $\delta t = 142\,\mathrm{ns}$ respectively. We characterize the entangled state by performing measurements in the joint spin-photon ZZ and XX bases (Fig. 2e), implementing local operations on the reflected photonic qubit with a time-delay interferometer (Fig. 2a, dashed box). By lowering the average number of photons $\langle n\rangle_m$ incident on the device during the SiV memory time, we reduce the possibility that an additional photon reaches the cavity without being subsequently detected, enabling high spin-photon gate fidelities for small $\langle n \rangle_m$ (Fig. 2f). For $\langle n \rangle_m = 0.002$ we measure a lower bound on the fidelity[19] of the spin-photon entangled state of $F \geq 0.944 \pm 0.008$, primarily limited by residual reflections from the $|{\downarrow}\rangle$ state.

This spin-photon logic gate can be directly used to herald the storage of an incoming photonic qubit by interferometrically measuring the reflected photon in the X basis[19]. To implement a memory-assisted BSM, we extend this protocol to accommodate a total of N photonic qubit time-bins within a single initialization of the memory (Fig. 3a). Each individual time-bin qubit is encoded in the relative amplitudes and phases of a pair of neighboring pulses separated by δt. Detection of a reflected photon heralds the arrival of the photonic qubit formed by the two interfering pulses without revealing its state[19]. Two such heralding events, combined with subsequent spin-state readout in the X basis, constitute a successful BSM on the incident photons. This can be understood without loss of generality by restricting input photonic states to be encoded in the relative phase ϕ between neighboring pulses with equal amplitude: $(|e\rangle + e^{i\phi}|l\rangle)/\sqrt{2}$ (Fig. 3b). Detection of the first reflected photon in the X basis teleports its quantum state onto the spin, resulting in the state $(|{\uparrow}\rangle + m_1 e^{i\phi_1}|{\downarrow}\rangle)/\sqrt{2}$, where $m_1 = \pm 1$ depending on which detector registers the photon[19]. Detection of a second photon at a later time within the electron spin T_2 results in the spin state $(|{\uparrow}\rangle + m_1 m_2 e^{i(\phi_1+\phi_2)}|{\downarrow}\rangle)/\sqrt{2}$. The phase of this spin state depends only on the sum of the incoming phases and the product of their detection outcomes, but not the individual phases themselves. As a result, if the photons were sent with phases that meet the condition $\phi_1 + \phi_2 \in \{0, \pi\}$, a final measurement of the spin in the X basis $(m_3 = \pm 1)$ completes an asynchronous BSM, distinguishing two of the four Bell-states based on the total parity $m_1 m_2 m_3 = \pm 1$ (Methods).

This approach can be directly applied to generate a correlated bit-string within the protocol illustrated in Fig. 1a. We analyze the system performance by characterizing the overall quantum-bit error rate (QBER)[4,21] for $N = 124$ photonic qubits per memory initialization. We use several random bit strings of incoming photons from $\{|\pm x\rangle, |\pm y\rangle\}$ and observe strong correlations between the resulting BSM outcome and the initial combination of input qubits for both bases (Fig. 3c). Using this method, we estimate the average QBER to be $E = 0.116 \pm 0.002$ for all combinations of random bit strings measured, significantly below the limit of $E_i = 0.146$, which could provide security against individual attacks[4] (note that the measured error rate is also well below the minimum average QBER[21] of $E_{lo} = 0.125$ achievable using a linear optics BSM with weak coherent pulse inputs, see Methods). In our experiment, the QBER is affected by technical imperfections in the preparation of random strings of photonic qubits. We find specific periodic patterns of photonic qubits to be less prone to these effects, resulting in a QBER as low as $E = 0.097 \pm 0.006$, which falls within the threshold corresponding to unconditional security[3] of $E_u = 0.110$ with a confidence level of 0.986 (Methods). We further verify security by testing the Bell-CHSH inequality[14] using input states from four different bases, each separated by an angle of $45°$ (Methods). We find that the correlations between input photons (Fig. 3d) violate the Bell-CHSH inequality $S_{\pm} \leq 2$, observing $S_+ = 2.21 \pm 0.04$ and $S_- = 2.19 \pm 0.04$ for positive and negative BSM parity results respectively. This result demonstrates that this device can be used for quantum communication that is secured by Bell's theorem.

Finally, we benchmark the performance of memory-assisted quantum communication. For each experiment, we model an effective channel loss by considering the mean photon number $\langle n \rangle_p$

incident on the device per photonic qubit. Assuming that Alice and Bob emit roughly one photon per qubit, this yields an effective channel transmission probability $p_{A\to B} = \langle n\rangle_p^2$, resulting in the maximal distilled key rate R_{max} per channel use for the direct transmission approach[21], given by the red line in Fig. 4. We emphasize that this is a theoretical upper bound for a linear optics based BSM, assuming ideal single-photon sources and detectors and balanced basis choices. The measured sifted key rates of the memory-based device are plotted as open circles in Fig. 4. Due to the high overall heralding efficiency and the large number of photonic qubits per memory time (up to $N = 504$), the memory-assisted sifted key rate exceeds the capability of a linear-optics based BSM device by a factor of 78.4 ± 0.7 at an effective channel loss of about 88 dB.

In practice, errors introduced by the quantum memory node could leak information to the environment, reducing the quality and potential security of the sifted key[3]. A shorter secure key can be recovered from a sifted key with finite QBER using classical error correction and privacy amplification techniques. The fraction of distilled bits r_s that can be secure against individual attacks rapidly diminishes[4] as the QBER approaches $E_i = 0.147$. For each value of the effective channel loss, we estimate the QBER and use it to compute r_s, enabling extraction of distilled key rates R_{S}, plotted in black in Fig. 4. Even after error-correction, we find that the memory-assisted distilled key rate outperforms the ideal limit for the corresponding direct-transmission implementation by a factor of up to $R_{\text{S}}/R_{\text{max}} = 4.1 \pm 0.5$ (± 0.1 systematic uncertainty, for $N = 124$). We further find that this rate also exceeds the fundamental bound on repeaterless communication[8] $R_{\text{S}} \leq 1.44 p_{\text{A}\to\text{B}}$ with a statistical confidence level of 99.2% ($^{+0.2\%}_{-0.3\%}$ systematic uncertainty, see Methods). Despite experimental overhead time associated with operating the device (T_R in Fig. 1b), the performance of the memory-assisted BSM node (for $N = 248$) is competitive with an ideal unassisted system running at a $4\,\mathrm{MHz}$ average clock rate (Methods).

These experiments demonstrate a form of quantum advantage allowed by memory-based communication nodes and represent a crucial step towards realizing functional quantum repeaters. Several important technical improvements will be necessary to apply this advance for practical long-distance quantum communication. First, this protocol must be implemented using truly independent, distant communicating parties. Additionally, frequency conversion from telecommunications wavelengths to $737\,\mathrm{nm}$, as well as low-loss optical elements used for routing photons to and from the memory node, will need to be incorporated. Finally, rapid generation of provably secure keys will require implementation of decoy-state protocols[25], biased bases[26], and finite-key error analyses[27], all compatible with the present approach. With these improvements, our approach is well-suited for deployment in real-world settings. It does not require phase stabilization of long-distance links and operates efficiently in the relevant regime of $p_{A\to B} \approx 70\,\mathrm{dB}$, corresponding to about $350\,\mathrm{km}$ of telecommunications fiber. Additionally, a single device can be used at the center of a star network topology[29], enabling quantum communication between several parties beyond the metropolitan scale. Furthermore, the present approach can be extended along several directions. The use of long-lived $^{13}\mathrm{C}$ nuclear spin qubits could eliminate the need to operate at low total $\langle n \rangle_m$ and would provide longer storage times, potentially enabling hundred-fold enhancement of BSM success rates[15,19]. Recently implemented strain-tuning capabilities[30] should allow for operation of many quantum nodes at a common network frequency. Unlike linear-optics based alternatives[23], the approach presented here can be extended to implement the full repeater protocol, enabling a polynomial scaling of the communication rate with distance[9]. Finally, the demonstrated multi-photon gate operations can also be adapted to engineer large cluster-states of entangled photons[31], which can be utilized for rapid quantum communication[32]. Implementation of these techniques could enable the realization and applications of scalable quantum networks[1] beyond QKD, ranging from non-local quantum metrology[?] to modular quantum computing architectures[20].

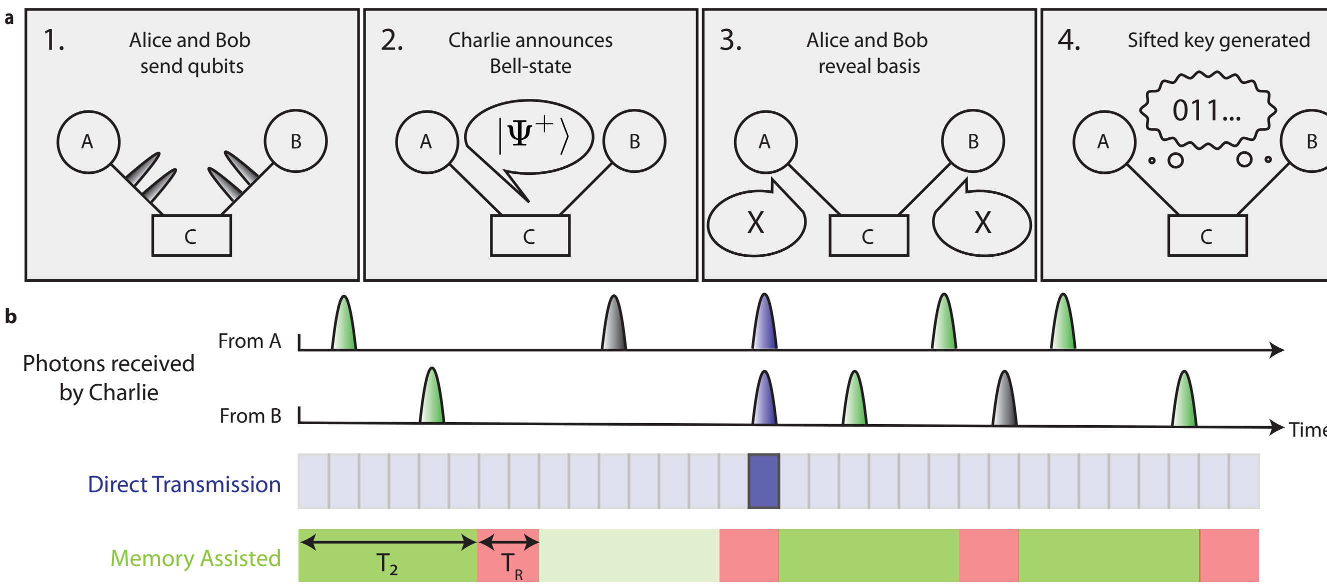


Figure 1: Concept of memory-enhanced quantum communication. a, Quantum communication protocol. Alice and Bob send qubits encoded in photons to a measurement device (Charlie) in between them. Charlie performs a BSM and announces the result. After verifying which rounds Alice and Bob sent qubits in compatible bases, a sifted key is generated. **b,** Illustration of memory-enhanced protocol. Photons arrive at Charlie from A and B at random times over a lossy channel, and are unlikely to arrive simultaneously (indicated in purple), leading to a low BSM success rate for direct transmission. Despite overhead time T_R associated with operating a quantum memory (red), a BSM can be performed between photons that arrive at Charlie within memory coherence time T_2, leading to higher success rates (green). BSM successes and failures are denoted by dark and light shaded windows respectively for both approaches.

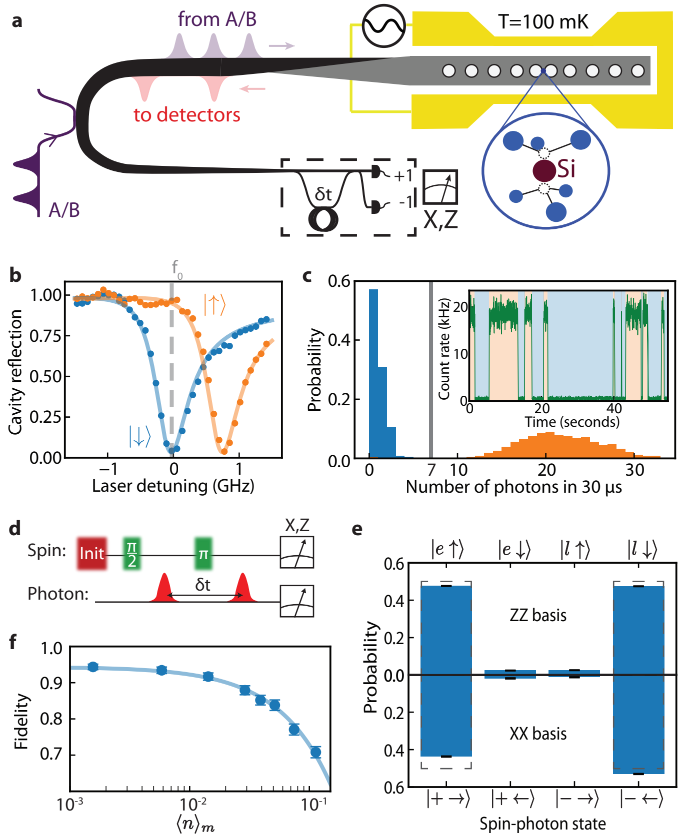


Figure 2: Realization of heralded spin-photon gate. a, Schematic of memory-assisted implementation of Charlie's measurement device. Weak pulses derived from a single laser simulate incoming photons from Alice and Bob (purple). Reflected photons (red) are detected in a heralding setup (dashed box). **b,** Reflection spectrum of memory node, showing spin-dependent device reflectivity. **c,** Histogram of detected photon numbers during a $30\,\mu$s laser pulse, enabling single-shot readout based on a threshold of 7 photons. (Inset) Electron spin quantum jumps under weak illumination. **d,** Schematic of spin-photon quantum logic operation used to generate and verify spin-photon entangled state. **e,** Characterization of resulting spin-photon correlations in the ZZ and XX bases. Dashed bars show ideal values. **f,** Measured spin-photon entanglement fidelity as a function of $\langle n \rangle_m$, the average incident photon number during each initialization of the memory.

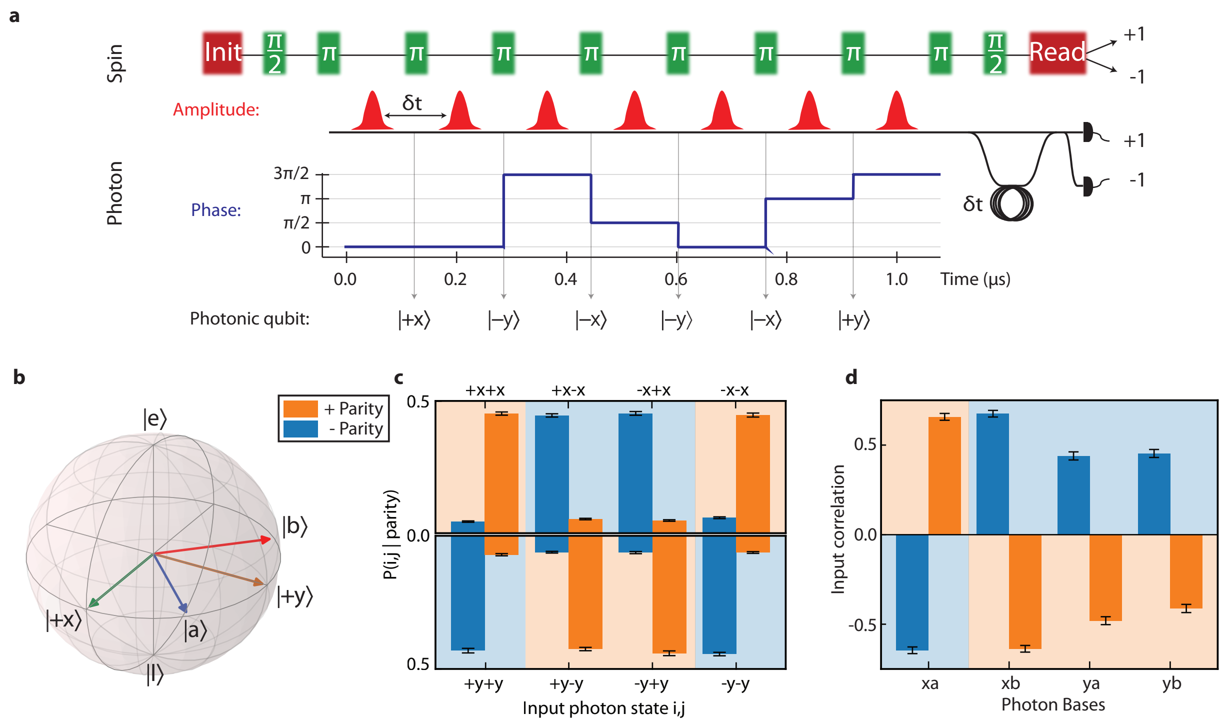


Figure 3: Asynchronous Bell-state measurements using quantum memory. a, Example sequence with $N = 6$ photonic qubits sent in a single memory time. Microwave π pulses (green) are interleaved with incoming optical pulses. Photons have fixed amplitude (red) and qubits are defined by the relative phases between subsequent pulses (blue). **b,** Bloch sphere representation of input photonic time-bin qubits used for characterization. **c,** Characterization of asynchronous BSM. Conditional probabilities for Alice and Bob to have sent input states (i, j) given a particular parity outcome for input states in the X (top) and Y (bottom) bases. **d,** Bell test using the CHSH inequality. Conditioned on the BSM outcome, the average correlation between input photons is plotted for each pair of bases used (Methods). Shaded backgrounds denote the expected parity.

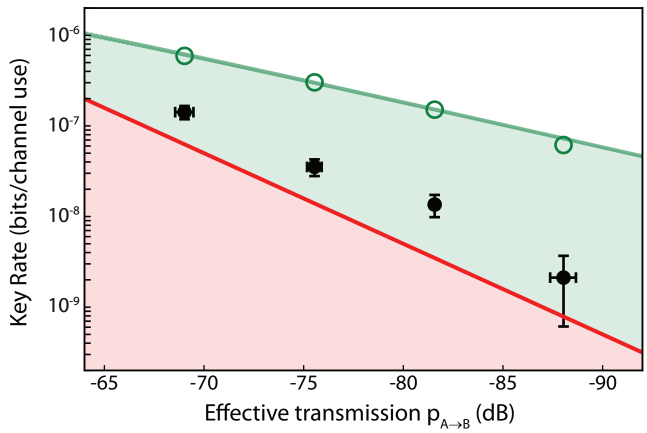


Figure 4: Performance of memory-assisted quantum communication. Log-log plot of key rate in bits per channel use versus effective channel transmission ($p_{\mathrm{A}\to\mathrm{B}} = \langle n \rangle_p^2$, where $\langle n \rangle_p$ is the average number of photons incident on the measurement device per photonic qubit). Red line: theoretical maximum for loss-equivalent direct transmission experiment. Green open circles: experimentally measured sifted key rate (green line is the expected rate). To ensure optimal operation of the memory, $\langle n \rangle_m = \langle n \rangle_p N \approx 0.02$ is kept constant (Methods). From left to right, points correspond to $N = \{60, 124, 248, 504\}$. Black filled circles: distilled key rates R_S using memory device. Vertical error bars are given by the 68% confidence interval and horizontal error bars represent the standard deviation of the systematic power fluctuations.

Methods

Experimental setup and device fabrication[18,30,33,34] for millikelvin nanophotonic cavity QED experiments with SiV centers are thoroughly described in a separate publication[35]. Measurements of the cavity QED parameters of the device used in the main text (Extended Data Fig. 1), microwave characterization of the spin qubit (Extended Data Fig. 2), and statistical techniques for estimating QBER are described in detail the Supplementary Methods.

Experimental details. An asynchronous BSM (Fig. 3a) relies on (1) precise timing of the arrival of optical pulses (corresponding to photonic qubits[36,37] from Alice and Bob) with microwave control pulses on the quantum memory and (2) interferometrically stable rotations on reflected time-bin qubits for successful heralding. In order to accomplish (1), all equipment used for generation of microwave and optical fields is synchronized by a single device (National Instriuments HSDIO, Extended Data Fig. 1a) with programming described in Extended Data Table 1-2.

In order to accomplish (2), we use a single, narrow linewidth ($< 50\,\mathrm{kHz}$) Ti:Sapphire laser (M Squared SolsTiS-2000-PSX-XF, Extended Data Fig. 1b) both for generating photonic qubits and locking the time-delay interferometer (TDI) used to herald their arrival. In the experiment, photonic qubits are reflected from the device, sent into the TDI, and detected on superconducting nanowire single photon detectors (SNSPD, Photon Spot). All detected photons are processed digitally on a field-programmable gate array (FPGA, Extended Data Fig. 1a), and the arrival times of these heralding signals are recorded on a time-tagger (TT, Extended Data Fig. 1a), and constitute one bit of information of the BSM (m_1 or m_2). At the end of the experiment, a $30\,\mu\mathrm{s}$ pulse from the readout path is reflected off the device, and photons are counted in order to determine the spin state (m_3) depending on the threshold shown in Fig. 2c.

To minimize thermal drift of the TDI, it is mounted to a thermally weighted aluminum breadboard, placed in a polyurethane foam-lined and sand filled briefcase, and secured with glue to ensure passive stability on the minute timescale. We halt the experiment and actively lock the interferometer to the sensitive Y-quadrature every $\sim 200\,\mathrm{ms}$ by changing the length of the roughly $28\,\mathrm{m}$ long ($142\,\mathrm{ns}$) delay line with a cylindrical piezo. In order to use the TDI for X-measurements of the reflected qubits, we apply a frequency shift of $1.8\,\mathrm{MHz}$ using the qubit AOM, which is $1/4$ of the free-spectral range of the TDI. Since the nanophotonic cavity, the TDI, and the SNSPDs are all polarization sensitive, we use various fiber-based polarization controllers (Extended Data Fig. 1b). All fibers in the network are covered with aluminum foil to prevent thermal polarization drifts. This results in an interference visibility of the TDI of $> 99\%$ that is stable for several days without any intervention with lab temperature and humidity variations of $\pm 1^\circ$ C and $\pm 5\%$ respectively.

In order to achieve high-fidelity operations we have to ensure that the laser frequency (which is not locked) is resonant with the SiV frequency f_0 (which is subject to the spectral diffusion[35]). To do that we implement a so-called preselection procedure, described in Extended Data Table 1-2 and Extended Data Fig. 1a. First, the SiV spin state is initialized by performing a projective measurement and applying microwave feedback. During each projective readout, the reflected counts are compared with two thresholds: a “readout” threshold of 7 photons (used only to record m_3), and a “status” threshold of 3 photons. The status trigger is used to prevent the experiment from running in cases when the laser is no longer on resonance with f_0, or if the SiV has ionized to an optically inactive charge state. The duty cycle of the status trigger is externally monitored and is used to temporarily abort the experiment and run an automated re-lock procedure that locates and sets the laser to the new frequency f_0, reinitailizing the SiV charge state with a $520\,\mathrm{nm}$ laser pulse if nececssary. This protocol enables fully automated operation at high fidelities (low QBER) for several days without human intervention.

Fiber network. The schematic of the fiber-network used to deliver optical pulses to and collect reflected photons from the nanophotonic memory device is shown in Extended Data Fig. 1b. Photons are routed through the lossy (1%) port of a 99:1 fiber beamsplitter (FBS) to the nanophotonic device. We note that for practical implementation of memory-assisted quantum communication, an efficient optical switch or circulator should be used instead. In this experiment, since we focus on benchmarking the performance of the memory device itself, the loss introduced by this beamsplitter is incorporated into the estimated channel loss. Reflected photons are collected and routed back through the efficient (99%) port of the FBS and are sent to the TDI in the heralding setup.

The outputs of the TDI are sent back into the dilution refrigerator and directly coupled to SNSPDs (PhotonSpot), which are mounted at the 1K stage and are coated with dielectrics to optimize detection efficiency exactly at $737\,\mathrm{nm}$. We estimate all losses that reduce the heralding efficiency η by two independent calibration methods. These rely on three calibrated photodetectors shown in Extended Data Fig. 1b (M1, M2, MC) and are described in detail in the Supplementary Methods, yielding consistent values of $\eta_1 = 0.425 \pm 0.008$ and $\eta_2 = 0.422 \pm 0.005$. This corresponds to a collection efficiency of reflected photons of $\approx 85\%$.

For values cited in the main text and data points presented in the figures, we use an average value of the heralding efficiency inferred from the two calibration techniques: $\eta = 0.423 \pm 0.004$. The residual uncertainty in the heralding efficiency results in a systematic uncertainty, which is explicitly mentioned in the main text where appropriate. We note that this heralding efficiency is consistent with the scaling of spin decoherence with the number of photons at the cavity $\langle n \rangle_m$. An example of this effect is shown in the red point in Extended Data Fig. 3e.

Theoretical description of asynchronous Bell state measurement. Due to the critical coupling of the nanocavity, the memory node only reflects photons when the SiV spin is in the state $|\uparrow\rangle$. The resulting correlations between the spin and the reflected photons can still be used to realize a BSM between two asynchronously arriving photonic time-bin qubits using an adaptation of the well known proposal of Duan and Kimble[28] for entangling a pair of photons incident on an atom-cavity system. As a result of the critical coupling, we only have access to two of the four Bell states at any time, with the inaccessible Bell states corresponding to photons being transmitted through the cavity (and thus lost from the detection path). Depending on whether there was an even or odd number of π-pulses on the spin between the arrival of the two heralded photons, we distinguish either the $\{|\Phi_\pm\rangle\}$ or $\{|\Psi_\pm\rangle\}$ states (defined below). For the sake of simplicity, we first describe the BSM for the case when the early time bin of Alice's and Bob's qubits both arrive after an even number of microwave π pulses after its initialization. Thereafter we generalize this result and describe the practical consequences for the quantum communication protocol.

The sequence begins with a $\pi/2$ microwave pulse, preparing the spin in the state $|\psi_i\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$. In the absence of a photon at the device, the subsequent microwave π-pulses, which follow an XY8-N type pattern, decouple the spin from the environment and at the end of the sequence should preserve the spin state $|\psi_i\rangle$. However, reflection of Alice's photonic qubit $|A\rangle = (|e\rangle + e^{i\phi_1}|l\rangle)/\sqrt{2}$ from the device results in the entangled spin-photon state $|\psi_A\rangle = (|\uparrow e\rangle + e^{i\phi_1}|\downarrow l\rangle)/\sqrt{2}$. The full system is in the state

$$|\psi_A\rangle = \frac{|{+x}\rangle\left(|\uparrow\rangle + e^{i\phi_1}|\downarrow\rangle\right) + |{-x}\rangle\left(|\uparrow\rangle - e^{i\phi_1}|\downarrow\rangle\right)}{2}. \tag{1}$$

Regardless of the input photon state, there is equal probability to measure the reflected photon to be $|\pm x\rangle$. Thus, measuring the photon in X basis (through the TDI) does not reveal the initial photon state. After this measurement, the initial state of the photon $|A\rangle$ is teleported onto the spin: $|\psi_{m_1}\rangle = (|\uparrow\rangle + m_1 e^{i\phi_1}|\downarrow\rangle)\sqrt{2}$, where $m_1 = \pm 1$ denotes the detection outcome of the TDI[19,38]. The quantum state of Alice's photon is now stored in the spin state, which is preserved by the dynamical decoupling sequence.

Reflection of the second photon $|B\rangle = (|e\rangle + e^{i\phi_2}|l\rangle)\sqrt{2}$ from Bob results in the spin-photon state $|\psi_{m_1,B}\rangle = (|\uparrow e\rangle + m_1 e^{i(\phi_1+\phi_2)}|\downarrow l\rangle)/\sqrt{2}$. This state now has a phase that depends on the initial states of both photons, enabling the photon-photon BSM measurements described below. Rewriting Bob's reflected photon in the X basis, the full system is in the state

$$|\psi_{m_1,B}\rangle = \frac{|{+x}\rangle\left(|\uparrow\rangle + m_1 e^{i(\phi_1+\phi_2)}|\downarrow\rangle\right) + |{-x}\rangle\left(|\uparrow\rangle - m_1 e^{i(\phi_1+\phi_2)}|\downarrow\rangle\right)}{2}. \tag{2}$$

The second measurement result m_2 once again contains no information about the initial state $|B\rangle$, yet heralds the final spin state $|\psi_{m_1,m_2}\rangle = (|\uparrow\rangle + m_1 m_2 e^{i(\phi_1+\phi_2)}|\downarrow\rangle)$ as described in the main text. When this state lies along the X axis of the Bloch sphere ($\phi_1 + \phi_2 = \{0, \pi\}$), the final result of the X basis measurement on the spin state m_3 has a deterministic outcome, dictated by all values of the parameters $\{\phi_1, \phi_2\}$ (known only to Alice and Bob) and $\{m_1, m_2\}$ (which are known to Charlie, but are completely random). Conversely, all information available to Charlie $\{m_1, m_2, m_3\}$ only contains information on the correlation between the photonic qubits, not on their individual states. The resulting truth table for different input states is given in Extended Data Table 3. For all input states, there is equal probability of measuring ± 1 for each individual measurement m_i. However, the overall parity of the three measurements $m_1 m_2 m_3$ depends on whether or not the input photons were the same, or opposite, for inputs $|A\rangle, |B\rangle \in |\pm x\rangle$ or $|\pm y\rangle$.

We now address the fact that the BSM distinguishes either between $\{|\Phi_\pm\rangle\}$ or $\{|\Psi_\pm\rangle\}$ if there was an even or odd number of microwave π pulses between incoming photons respectively. This effect arises because each π pulse in the dynamical decoupling sequence toggles an effective frame change: $Y \leftrightarrow -Y$. The impact on this frame change on the BSM can be seen by writing the pairs of Bell states ($|\Phi_\pm\rangle = (|ee\rangle \pm |ll\rangle)/\sqrt{2}$ and $|\Psi_\pm\rangle = (|el\rangle \pm |le\rangle)/\sqrt{2}$) in the X and Y bases, where we have

$$|\Phi_\pm\rangle^{(X)} = (|{+x}\rangle |{\pm x}\rangle + |{\mp x}\rangle |{-x}\rangle)/\sqrt{2} \tag{3}$$

$$|\Phi_\pm\rangle^{(Y)} = (|{+y}\rangle |{\mp y}\rangle + |{\pm y}\rangle |{-y}\rangle)/\sqrt{2} \tag{4}$$

$$|\Psi_\pm\rangle^{(X)} = (|{+x}\rangle |{\pm x}\rangle - |{\mp x}\rangle |{-x}\rangle)/\sqrt{2} \tag{5}$$

$$|\Psi_\pm\rangle^{(Y)} = i(|{+y}\rangle |{\pm y}\rangle - |{\mp y}\rangle |{-y}\rangle)/\sqrt{2} \tag{6}$$

For X basis inputs, as seen by Eq. 3 and 5, switching between $\{|\Phi_\pm\rangle\}$ and $\{|\Psi_\pm\rangle\}$ measurements does not affect the inferred correlation between input photons. For Y basis inputs however, this does result in an effective bit flip in the correlation outcome (see Eq. 4 and 6). In practice, Alice and Bob can keep track of each Y photon sent and apply a bit flip accordingly, as long as they have the appropriate timing information about MW pulses applied by Charlie. If Charlie does not give them the appropriate information, this will result in an increased QBER which can be detected.

As an additional remark, this scheme also works for pairs of photons that are not both in the X or Y basis but still satisfy the condition $\phi_1 + \phi_2 = 0$. For example, $|a\rangle$ and $|b\rangle$ from Fig. 3b satisfy this condition. In this case, adequate correlations can still be inferred about the input photons, although they were sent in different bases.

Finally, we would like to note that this asynchronous scheme for performing a BSM between two pulses has an important advantage over the synchronous, linear-optical implementation. In the case where Alice and Bob use attenuated laser pulses to encode photonic qubits, which is by far the most technologically simple implementation, there is an inherent QBER of 25% for photons sent in the X and Y bases. This is because the linear-optical implementation relies on the pulses to overlap on a beamsplitter and interfere via the Hong-ou-Mandel effect, which has a finite visibility of 50% for coherent pulses[21]. Intuitively, this finite error arises from a fundamental inability to distinguish between detection outcomes where two individual single-photons arrived from Alice and Bob versus a two-photon component from either Alice or Bob in the synchronous scheme. In the asynchronous scheme, since pulses do not arrive at the same time from Alice and Bob, this error is mitigated. As a result, for ideal quantum memory operation and sufficiently attenuated laser pulses, the ultimate limit to the QBER is zero.

Test of Bell-CHSH inequality. In order to perform a test of the Bell-CHSH inequality[39], we send input photons equally distributed from all states $\{|\pm x\rangle, |\pm y\rangle, |\pm a\rangle, |\pm b\rangle\}$ (Fig. 3b). We select for cases where two heralding events arise from input photons $\{A, B\} = \pm 1$ that are either $45°$ or $135°$ apart from one another. Conditioned on the parity outcome of the BSM (± 1), the Bell-CHSH inequality bounds the correlations between input photons as

$$S_{\pm} = |\langle A \cdot B\rangle_{xa} - \langle A \cdot B\rangle_{xb} - \langle A \cdot B\rangle_{ya} - \langle A \cdot B\rangle_{yb}| \leq 2, \tag{7}$$

where the subscripts denote the bases the photons were sent in. The values of each individual term in Eq. 7, denoted as "input correlations," are plotted in Fig. 3d for positive and negative parity outcomes.

Optimal parameters for asynchronous Bell state measurements. We minimize the experimentally extracted QBER for the asynchronous BSM to optimize the performance of the memory node. The first major factor contributing to QBER is the scattering of a third photon that is not detected, due to the finite heralding efficiency $\eta = 0.423 \pm 0.04$. This is shown in Fig. 2f, where the fidelity of the spin-photon entangled state diminishes for $\langle n\rangle_m \gtrsim 0.02$. At the same time, we would like to work at the maximum possible $\langle n\rangle_m$ in order to maximize the data rate to get enough statistics to extract QBER (and in the quantum communication setting, efficiently generate a key).

To increase the key generation rate per channel use, one can also fit many photonic qubits within each initialization of the memory. In practice, there are 2 physical constraints: (1) the bandwidth of the SiV-photon interface and (2) the coherence time of the memory. We find that one can satisfy (1) at a bandwidth of roughly $50\,\mathrm{MHz}$ with no measurable infidelity. For shorter optical pulses ($< 10\,\mathrm{ns}$), the spin-photon gate fidelity is reduced. In principle, the SiV-photon bandwidth can be increased by reducing the atom-cavity detuning (here $\sim 60\,\mathrm{GHz}$) at the expense of having to operate at higher magnetic fields where microwave qubit manipulation is not as convenient[35].

Even with just an XY8-1 decoupling sequence (number of π pulses $N_\pi = 8$), the coherence time of the SiV is longer than $200\,\mu$s (Extended Data Fig. 3c) and can be prolonged to the millisecond range with longer pulse sequences[19]. Unfortunately, to satisfy the bandwidth criteria (1) and to drive both hyperfine transitions (Extended Data Fig. 3a), we must use short ($32\,\mathrm{ns}$ long π pulses), which cause additional decoherence from ohmic heating[35] already at $N_\pi = 64$ (Extended Data Fig. 3e). Due to this we limit the pulse sequences to a maximum $N_\pi = 128$, and only use up to $\approx 20\,\mu$s of the memory time. One solution would be to switch to superconducting microwave delivery. Alternatively, one can use a larger value of τ to allow the device to cool down in between subsequent pulses[35] at the expense of having to stabilize a TDI of larger δt. Working at larger δt also enables temporal multiplexing by fitting multiple time-bin qubits per free-precession interval. In fact, with $2\tau = 142\,\mathrm{ns}$, even given constraint (1) and the finite π time, we can already fit up to 4 optical pulses per free-precession window, enabling a total number of photonic qubits of up to $N = 504$ for only $N_\pi = 128$.

In benchmarking the asynchronous BSM for quantum communication, we optimize the parameters $\langle n \rangle_m$ and N to maximize our enhancement over the direct transmission approach, which is a combination of both increasing N and reducing the QBER, since a large QBER results in a small distilled key fraction r_s. As described in the main text, the effective loss can be associated with $\langle n \rangle_p$, which is the average number of photons per photonic qubit arriving at the device, and is given straightforwardly by $\langle n \rangle_p = \langle n \rangle_m / N$. The most straightforward way to sweep the loss is to keep the experimental sequence the same (fixed N) and vary the overall power, which changes $\langle n \rangle_m$. The results of such a sweep are shown in Extended Data Fig. 5a, b. For larger $\langle n \rangle_m$ (corresponding to lower effective channel losses), the errors associated with scattering an additional photon reduce the performance of the memory device.

Due to these considerations, we work at roughly $\langle n \rangle_m \lesssim 0.02$ for experiments in the main text shown in Fig. 3 and 4, below which the performance does not improve significantly. At this value, we obtain BSM successes at a rate of roughly $0.1\,\mathrm{Hz}$. By fixing $\langle n \rangle_m$ and increasing N, we maintain a tolerable BSM success rate while increasing the effective channel loss. Eventually, as demonstrated in Extended Data Fig. 5c and in the high-loss data point in Fig. 4, effects associated with microwave heating result in errors that again diminish the performance of the memory node for large N. As such, we conclude that the optimal performance of our node occurs for $\langle n \rangle_m \sim 0.02$ and $N \approx 124$, corresponding to an effective channel loss of $69\,\mathrm{dB}$ between Alice and Bob, which is equivalent to roughly $350\,\mathrm{km}$ of telecommunications fiber.

We also find that the QBER and thus the performance of the communication link is limited by imperfect preparation of photonic qubits. Photonic qubits are defined by sending arbitrary phase patterns generated by the optical AWG to a phase modulator. For an example of such a pattern, see the blue curve in Fig. 3a. We use an imperfect pulse amplifier with finite bandwidth ($0.025 - 700$ MHz), and find that the DC component of these waveforms can result in error in photonic qubit preparation on the few $\%$ level. By using a tailored waveform of phases with smaller (or vanishing) DC component, we can reduce these errors. We run such an experiment during the test of the Bell-

CHSH inequality. We find that by evaluating BSM correlations from $|\pm a\rangle$ and $|\pm b\rangle$ inputs during this measurement, we estimate a QBER of 0.097 ± 0.006.

Finally, we obtain the effective clock-rate of the communication link by measuring the total number of photonic qubits sent over the course of an entire experiment. In practice, we record the number of channel uses, determined by the number of sync triggers recorded (see Extended Data Fig. 1a) as well as the number of qubits per sync trigger (N). We then divide this number by the total experimental time from start to finish ($\sim$ 1-2 days for most experimental runs), including all experimental downtime used to stabilize the interferometer, readout and initialize the SiV, and compensate for spectral diffusion and ionization. For $N = 248$, we extract a clock rate of $1.2\,\mathrm{MHz}$. As the distilled key rate in this configuration exceeds the conventional limit of $p/2$ by a factor of 3.8 ± 1.1, it is competitive with a standard linear-optics based system operating at $4.5^{+1.3}_{-1.2}$ MHz clock rate.

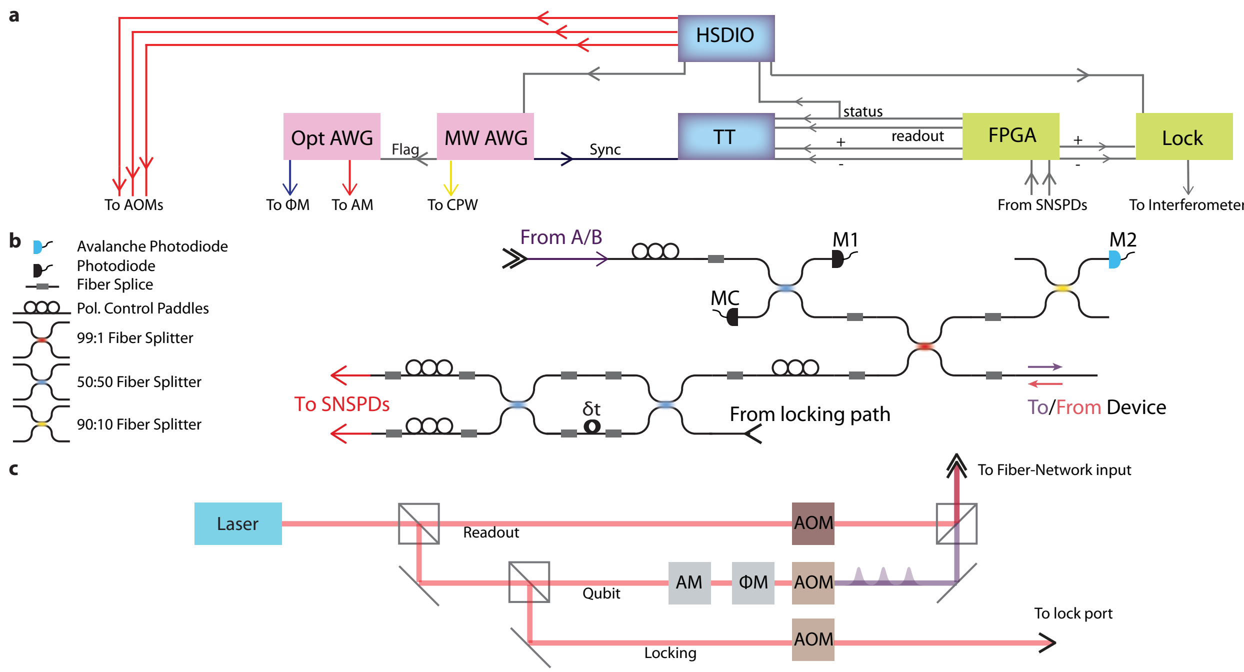
Performance of memory-assisted quantum communication. A single optical link can provide many channels, for example, by making use of different frequency, polarization, or temporal modes. To account for this, when comparing different systems, data rates can be defined on a per-channel-use basis. In a quantum communication setting, full usage of the communication channel between Alice and Bob means that both links from Alice and Bob to Charlie are in use simultaneously. For an asynchronous sequential measurement, typically only half of the channel is used at a time, for example from Alice to Charlie or Bob to Charlie. The other half can in principle be used for a different task when not in use. For example, the unused part of the channel could be routed to a secondary asynchronous BSM device. In our experiment, we can additionally define as a second normalization the rate per channel "occupancy", which accounts for the fact that only half the channel is used at any given time. The rate per channel occupancy is therefore half the rate per full channel use. For comparison, we typically operate at 1.2% channel use and 2.4% channel occupancy.

To characterize the optimal performance of the asynchronous Bell state measurement device, we operate it in the optimal regime determined above ($N = 124$, $\langle n \rangle_m \lesssim 0.02$). We note that the enhancement in the sifted key rate over direct transmission is given by

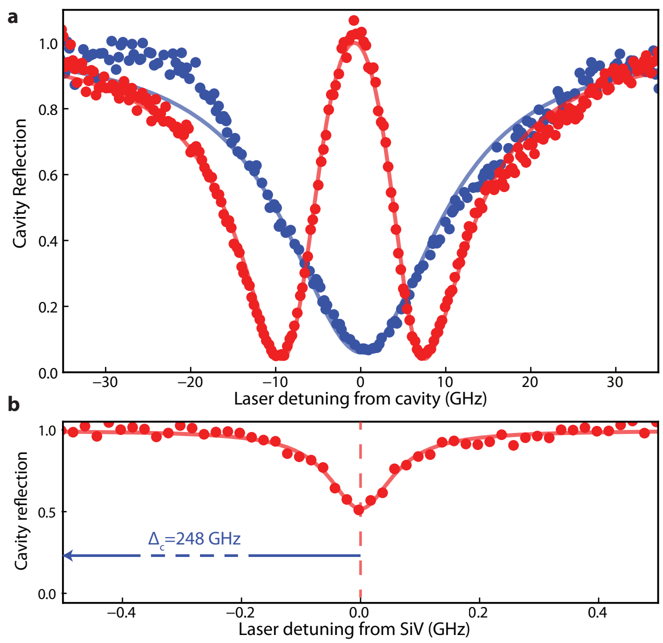
$$\frac{R}{R_{\text{max}}} = \eta^2 \frac{(N_\pi - 1)(N_\pi - 2)N_{\text{sub}}}{2N_\pi} \tag{8}$$

and is independent of $\langle n \rangle_m$ for a fixed number of microwave pulses N_π and optical pulses per microwave pulse N_{sub} and thus fixed $N = N_\pi N_{\text{sub}}$. For low $\langle n \rangle_m$, three photon events become negligible and therefore QBER saturates, such that the enhancement in the distilled key rate saturates as well (Extended Data Fig. 5a). We can therefore combine all data sets with fixed $N = 124$ below $\langle n \rangle_m \lesssim 0.02$ to characterize the average QBER of 0.116 ± 0.002 (Fig. 3c). The key rates cited in the main text relate to a data set in this series ($\langle n \rangle_m \approx 0.02$), with a QBER of 0.110 ± 0.004. A summary of key rates calculated on a per-channel use and per-channel occupancy basis, as well as comparisons of performance to an ideal linear-optics BSM and the repeaterless bound[8] are given in Extended Data Table 4.

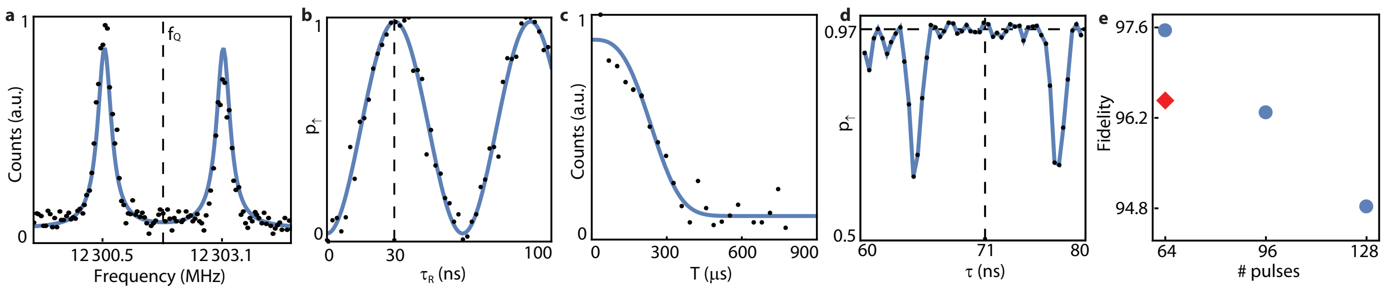
Furthermore, we extrapolate the performance of our memory node to include biased input bases from Alice and Bob. This technique enables a reduction of channel uses where Alice and Bob send photons in different bases, but is still compatible with secure key distribution[26], allowing for enhanced distilled key rates by at most a factor of 2. The extrapolated performance of our node for a bias of 99:1 is also displayed in Extended Data Table 4, as well as comparisons to the relevant bounds. We note that basis biasing does not affect the performance when comparing to the equivalent direct-transmission experiment, which is limited by $p_{A\to B}/2$ in the unbiased case and $p_{A\to B}$ in the biased case. However, using biased input bases does make the performance of the memory-assisted approach more competitive with the fixed repeaterless bound[8] of $1.44 p_{A\to B}$.



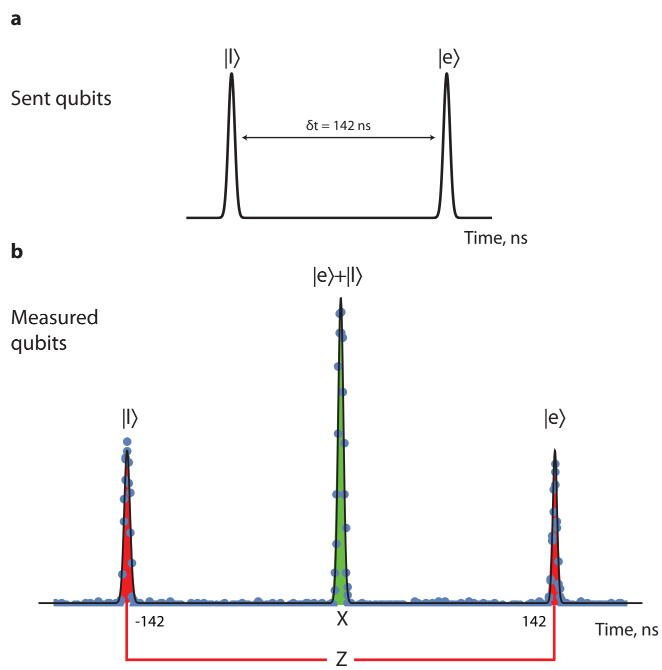
Extended Data Figure 1: Experimental schematic. a, Control flow of experiment. Opt (MW) AWG is a Tektronix AWG7122B $5\,\mathrm{GS/s}$ (Tektronix AWG70001a $50\,\mathrm{GS/s}$) arbitrary waveform generator used to generate photonic qubits (microwave control signals). All signals are recorded on a time-tagger (TT, PicoQuant HydraHarp 400). **b,** Fiber network used to deliver photons to and collect photons from the memory device, including elements for polarization control and diagnostic measurements of coupling efficiencies. **c,** Preparation of optical fields. The desired phase relation between lock and qubit paths is ensured by modulating AOMs using phase-locked RF sources with a precise $1.8\,\mathrm{MHz}$ frequency shift between them.



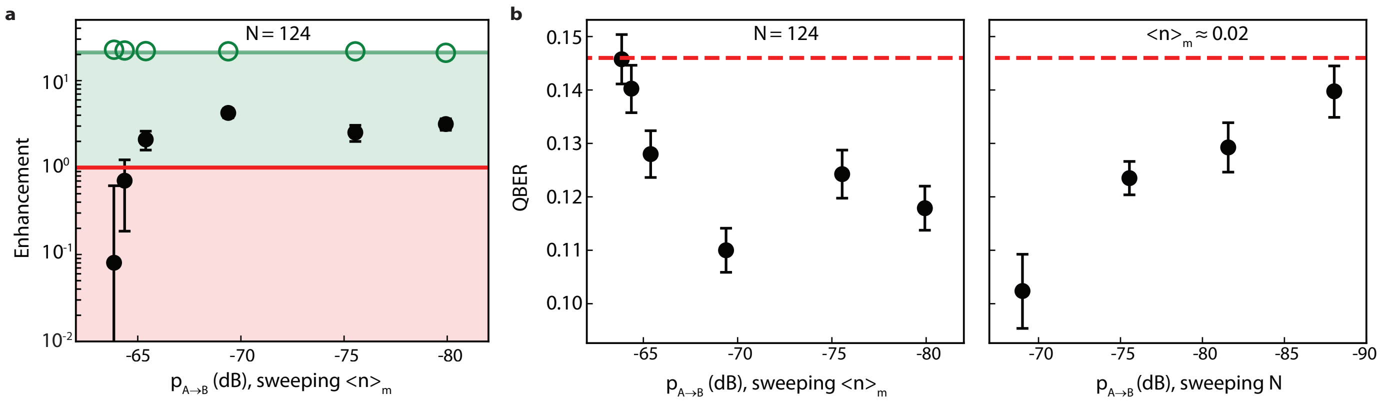
Extended Data Figure 2: Characterization of device cooperativity. a, Cavity reflection spectrum far-detuned (blue) and on resonance (red) with SiV center. Blue solid line is a fit to a Lorentzian, enabling extraction of linewidth $\kappa = 21.8\,\text{GHz}$. Red solid line is a fit to a model used to determine the single-photon Rabi frequency $g = 8.38 \pm 0.05$ GHz and shows the onset of a normal mode splitting. **b,** Measurement of SiV linewidth far detuned ($\Delta_c = 248\,\text{GHz}$) from cavity resonance. Red solid line is a fit to a Lorentzian, enabling extraction of natural linewidth $\gamma = 0.123\,\text{GHz}$.



Extended Data Figure 3: Microwave characterization of spin-coherence properties. a, ODMR spectrum of the qubit transition at $\sim 12\,\mathrm{GHz}$ split by coupling to a nearby $^{13}\mathrm{C}$. **b,** Rabi oscillations showing π time of $30\,\mathrm{ns}$. A π time of $32\,\mathrm{ns}$ is used for experiments in the main text. **c,** XY8-1 dynamical decoupling signal (unnormalized) as a function of total time T, showing coherence lasting on the several hundred μs timescale. **d,** XY8-8 dynamical decoupling signal (normalized) revealing region of high fidelity at relevant value of $2\tau = 142\,\mathrm{ns}$. **e,** Fidelity of spin state after dynamical decoupling sequence with varying number of π pulses (N_π), blue points. Red point (diamond) is under illumination with $\langle n \rangle_m = 0.02$.



Extended Data Figure 4: Measurements on a single time-bin qubit in Z and X bases. a, Example of optical pulses sent for example in the experiment described in Fig. 2d. **b,** Time trace of detected photons on + detector when pulses shown in (a) are sent directly into the TDI. The first and last peaks correspond to late and early photons taking the long and short paths of the TDI, which enable measurements in the Z basis $\{|e\rangle, |l\rangle\}$. The central bin corresponds to the late and early components overlapping and interfering constructively to come out of the + port, equivalent to a measurement of the time bin qubit in the $|+x\rangle$ state. A detection event in this same timing window on the - detector (not shown) would constitute a $|-x\rangle$ measurement. In this measurement, the TDI was left unlocked, so we observe no interference in the central window.



Extended Data Figure 5: Performance of memory-device versus of channel loss. a, Enhancement of memory-based approach compared to direct transmission approach, keeping $N = 124$ fixed and varying $\langle n \rangle_m$ in order to vary the effective channel transmission probability $p_{A\to B}$. At high $p_{A\to B}$ (larger $\langle n \rangle_m$), r_s approaches 0 due to increased QBER arising from undetected scattering of a third photon. **b,** (Left) Plot of QBER for same sweep of $\langle n \rangle_m$ shown in **a**. (Right) Plot of QBER while sweeping N in order to vary loss. These points correspond to the same data shown in Fig. 4. At lower $p_{A\to B}$ (larger N), microwave-induced heating-related dephasing leads to increased QBER.

Step	Process	Duration	Proceed to
1	Lock time-delay interferometer	$200\,\mathrm{ms}$	2
2	Readout SiV	$30\,\mu\mathrm{s}$	If status LOW: 4, else: 3
3	Apply microwave π pulse	$32\,\mathrm{ns}$	2
4	Run main experiment script	$\sim 200\,\mathrm{ms}$	1

Extended Data Table 1: High-level experimental sequence. This sequence is programmed into the HSDIO and uses feedback from the status trigger sent from the FPGA (see Extended Data Fig. 1a). Main experimental sequence is described in Extended Data Table 2. External software is also used to monitor the status trigger. If it is HI for $\gtrsim 2\,\mathrm{s}$, the software activates an automatic re-lock procedure which compensates for spectral diffusion and ionization of the SiV center (Methods).

Step	Process	Duration	Proceed to
1	Run sequence in Fig. 3a for a given N	$10 - 20\ \mu\mathrm{s}$	2
2	Readout SiV + report readout to TT	$30\ \mu\mathrm{s}$	If status LOW: 1, else: 3
3	Apply microwave π pulse	$32\ \mathrm{ns}$	4
4	Readout SiV	$30\ \mu\mathrm{s}$	If status LOW: 3, else: 1

Extended Data Table 2: Main experimental sequence for memory-enhanced quantum communication. This script is followed until step 1 is run a total of 4000 times, and then terminates and returns to step 1 of Extended Data Table 1. The longest step is the readout step, which is limited by the fact that we operate at a photon detection rate of $\sim 1\ \mathrm{MHz}$ to avoid saturation of the SNSPDs.

Alice	Bob	Parity	Bell state
$\lvert +x\rangle$	$\lvert +x\rangle$	$+1$	$\lvert \Phi_+\rangle$
$\lvert +x\rangle$	$\lvert -x\rangle$	-1	$\lvert \Phi_-\rangle$
$\lvert -x\rangle$	$\lvert +x\rangle$	-1	$\lvert \Phi_-\rangle$
$\lvert -x\rangle$	$\lvert -x\rangle$	$+1$	$\lvert \Phi_+\rangle$
$\lvert +y\rangle$	$\lvert +y\rangle$	-1	$\lvert \Phi_-\rangle$
$\lvert +y\rangle$	$\lvert -y\rangle$	$+1$	$\lvert \Phi_+\rangle$
$\lvert -y\rangle$	$\lvert +y\rangle$	$+1$	$\lvert \Phi_+\rangle$
$\lvert -y\rangle$	$\lvert -y\rangle$	-1	$\lvert \Phi_-\rangle$

Extended Data Table 3: Truth table of asynchronous BSM protocol, showing the parity (and BSM outcome) for each set of valid input states from Alice and Bob. In the case of Y basis inputs, Alice and Bob adjust the sign of their input state depending on whether it was commensurate with an even or odd numbered free-precession interval, based on timing information provided by Charlie (Methods).

	per channel occupancy	per channel occupancy	per channel use	per channel use
X:Y basis bias	$50:50$	$99:1$	$50:50$	$99:1$
Distilled key rate R $[10^{-7}]$	$1.19^{+0.14}_{-0.14}$	$2.33^{+0.28}_{-0.28}$	$2.37^{+0.29}_{-0.28}$	$4.66^{+0.56}_{-0.55}$
$R/R_{\text{max}}(\text{X:Y})$	$2.06^{+0.25}_{-0.25}$	$2.06^{+0.25}_{-0.25}$	$4.13^{+0.50}_{-0.49}$	$4.13^{+0.50}_{-0.49}$
$R/(1.44p_{A\to B})$	$0.71^{+0.09}_{-0.08}$	$1.40^{+0.17}_{-0.17}$	$1.43^{+0.17}_{-0.17}$	$2.80^{+0.34}_{-0.33}$
$1-$confidence level		$1.1^{+0.4}_{-0.3}\times 10^{-2}$	$8^{+3}_{-2}\times 10^{-3}$	$1.3^{+0.5}_{-0.3}\times 10^{-7}$

Extended Data Table 4: Quantum-memory-based advantage. Distilled key rates with the asynchronous BSM device and comparison to ideal direct communication implementations, based on the performance of our network node for $N = 124$ and $\langle n\rangle_m \sim 0.02$. Distillable key rates for $E = 0.110 \pm 0.004$ for unbiased and biased basis choice are expressed in a per-channel-occupancy and per-channel-use normalization (Methods). Enhancement is calculated versus the linear optics BSM limit ($R_{\text{max}}(50:50) = p_{A\to B}/2$ for unbiased bases, $R_{\text{max}}(99:1) = 0.98p_{A\to B}$ with biased bases) and versus the fundamental repeaterless channel capacity[8] ($1.44p_{A\to B}$). Confidence levels for surpassing the latter bound[8] are given in the final row.

1. Kimble, H. J. The quantum internet. *Nature* **453**, 1023 (2008). URL https://doi.org/10.1038/nature07127.

2. Bennett, C. H. & Brassard, G. Quantum cryptography: Public key distribution and coin tossing. *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* **175**, 8 (1984). URL https://doi.org/10.1016/j.tcs.2014.05.025.

3. Shor, P. W. & Preskill, J. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters* **85**, 441–444 (2000). URL https://doi.org/10.1103/PhysRevLett.85.441.

4. Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. *Reviews of Modern Physics* **74**, 145–195 (2002). URL https://doi.org/10.1103/RevModPhys.74.145.

5. Boaron, A. *et al.* Secure Quantum Key Distribution over 421 km of Optical Fiber. *Physical Review Letters* **121**, 1–4 (2018). URL https://doi.org/10.1103/PhysRevLett.121.190502.

6. Zhang, Q., Xu, F., Chen, Y.-A., Peng, C.-Z. & Pan, J.-W. Large scale quantum key distribution: challenges and solutions [Invited]. *Optics Express* **26**, 24260–24273 (2018). URL https://doi.org/10.1364/OE.26.024260.

7. Pirandola, S. *et al.* Advances in Quantum Cryptography (2019). URL http://arxiv.org/abs/1906.01645.

8. Pirandola, S., Laurenza, R., Ottaviani, C. & Banchi, L. Fundamental limits of repeaterless quantum communications. *Nature Communications* **8**, 15043 (2017). URL https://doi.org/10.1038/ncomms15043.

9. Briegel, H.-J., Dür, W., Cirac, J. I. & Zoller, P. Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication. *Physical Review Letters* **81**, 5932–5935 (1998). URL https://doi.org/10.1103/PhysRevLett.81.5932.

10. Chou, C.-W. *et al.* Functional Quantum Nodes for Entanglement Distribution over Scalable Quantum Networks. *Science* **316**, 1316 LP – 1320 (2007). URL https://doi.org/10.1126/science.1140300.

11. Yuan, Z.-S. *et al.* Experimental demonstration of a BDCZ quantum repeater node. *Nature* **454**, 1098 (2008). URL https://doi.org/10.1038/nature07241.

12. Gao, W. B., Fallahi, P., Togan, E., Miguel-Sanchez, J. & Imamoglu, A. Observation of entanglement between a quantum dot spin and a single photon. *Nature* **491**, 426 (2012). URL https://doi.org/10.1038/nature11573.

13. Reiserer, A. & Rempe, G. Cavity-based quantum networks with single atoms and optical photons. *Reviews of Modern Physics* **87**, 1379–1418 (2015). URL https://doi.org/10.1103/RevModPhys.87.1379.

14. Hensen, B. *et al.* Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682 (2015). URL https://doi.org/10.1038/nature15759.

15. Kalb, N. *et al.* Entanglement distillation between solid-state quantum network nodes. *Science* **356**, 928 LP – 932 (2017). URL https://doi.org/10.1126/science.aan0070.

16. Kaneda, F., Xu, F., Chapman, J. & Kwiat, P. G. Quantum-memory-assisted multi-photon generation for efficient quantum information processing. *Optica* **4**, 1034–1037 (2017). URL https://doi.org/10.0.5.84/OPTICA.4.001034.

17. Evans, R. E. *et al.* Photon-mediated interactions between quantum emitters in a diamond nanocavity. *Science* **362**, 662 LP – 665 (2018). URL https://doi.org/10.1126/science.aau4691.

18. Burek, M. J. *et al.* Fiber-Coupled Diamond Quantum Nanophotonic Interface. *Physical Review Applied* **8**, 24026 (2017). URL https://doi.org/10.1103/PhysRevApplied.8.024026.

19. Nguyen, C. T. *et al.* Quantum Network Nodes Based on Diamond Qubits with an Efficient Nanophotonic Interface. *Phys. Rev. Lett.* **123**, 183602 (2019). URL https://link.aps.org/doi/10.1103/PhysRevLett.123.183602.

20. Monroe, C. *et al.* Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects. *Physical Review A* **89**, 22317 (2014). URL https://doi.org/10.1103/PhysRevA.89.022317.

21. Lo, H.-K., Curty, M. & Qi, B. Measurement-Device-Independent Quantum Key Distribution. *Physical Review Letters* **108**, 130503 (2012). URL https://doi.org/10.1103/PhysRevLett.108.130503.

22. Braunstein, S. L. & Pirandola, S. Side-Channel-Free Quantum Key Distribution. *Physical Review Letters* **108**, 130502 (2012). URL https://doi.org/10.1103/PhysRevLett.108.130502.

23. Minder, M. *et al.* Experimental quantum key distribution beyond the repeaterless secret key capacity. *Nature Photonics* **13**, 334–338 (2019). URL https://doi.org/10.1038/s41566-019-0377-7.

24. Panayi, C., Razavi, M., Ma, X. & Lütkenhaus, N. Memory-assisted measurement-device-independent quantum key distribution. *New Journal of Physics* **16**, 43005 (2014). URL https://doi.org/10.1088/1367-2630/16/4/043005.

25. Lo, H.-K., Ma, X. & Chen, K. Decoy State Quantum Key Distribution. *Physical Review Letters* **94**, 230504 (2005). URL https://link.aps.org/doi/10.1103/PhysRevLett.94.230504.

26. Lo, H.-K., Chau, H. F. & Ardehali, M. Efficient Quantum Key Distribution Scheme and a Proof of Its Unconditional Security. *Journal of Cryptology* **18**, 133–165 (2005). URL https://doi.org/10.1007/s00145-004-0142-y.

27. Curty, M. *et al.* Finite-key analysis for measurement-device-independent quantum key distribution. *Nature Communications* **5**, 3732 (2014). URL https://doi.org/10.1038/ncomms4732.

28. Duan, L.-M. & Kimble, H. J. Scalable Photonic Quantum Computation through Cavity-Assisted Interactions. *Physical Review Letters* **92**, 127902 (2004). URL https://link.aps.org/doi/10.1103/PhysRevLett.92.127902.

29. Biham, E., Huttner, B. & Mor, T. Quantum cryptographic network based on quantum memories. *Physical Review A* **54**, 2651–2658 (1996). URL https://link.aps.org/doi/10.1103/PhysRevA.54.2651.

30. Machielse, B. *et al.* Quantum Interference of Electromechanically Stabilized Emitters in Nanophotonic Devices. *Physical Review X* **9**, 31022 (2019). URL https://link.aps.org/doi/10.1103/PhysRevX.9.031022.

31. Raussendorf, R. & Briegel, H. J. A One-Way Quantum Computer. *Physical Review Letters* **86**, 5188–5191 (2001). URL https://link.aps.org/doi/10.1103/PhysRevLett.86.5188.

32. Borregaard, J. *et al.* One-way quantum repeater based on near-deterministic photon-emitter interfaces 1–19 (2019). URL http://arxiv.org/abs/1907.05101.

33. Burek, M. J. *et al.* High quality-factor optical nanocavities in bulk single-crystal diamond. *Nature Communications* **5**, 5718 (2014). URL https://doi.org/10.1038/ncomms6718.

34. Atikian, H. A. *et al.* Freestanding nanostructures via reactive ion beam angled etching. *APL Photonics* **2**, 51301 (2017). URL https://doi.org/10.1063/1.4982603.

35. Nguyen, C. T. *et al.* An integrated nanophotonic quantum register based on silicon-vacancy spins in diamond. *Phys. Rev. B* **100**, 165428 (2019). URL `https://link.aps.org/doi/10.1103/PhysRevB.100.165428`.

36. de Riedmatten, H. *et al.* Tailoring photonic entanglement in high-dimensional Hilbert spaces. *Physical Review A* **69**, 50304 (2004). URL `https://doi.org/10.1103/PhysRevA.69.050304`.

37. Sasaki, T., Yamamoto, Y. & Koashi, M. Practical quantum key distribution protocol without monitoring signal disturbance. *Nature* **509**, 475 (2014). URL `https://doi.org/10.1038/nature13303`.

38. Kalb, N., Reiserer, A., Ritter, S. & Rempe, G. Heralded Storage of a Photonic Quantum Bit in a Single Atom. *Physical Review Letters* **114**, 220501 (2015). URL `https://doi.org/10.1103/PhysRevLett.114.220501`.

39. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters* **23**, 880–884 (1969). URL `https://doi.org/10.1103/PhysRevLett.23.880`.

Acknowledgements We thank Pavel Stroganov, Kristiaan de Greve, Johannes Borregaard, Eric Bersin, Benjamin Dixon, and Neil Sinclair for discussions, Vikas Anant from PhotonSpot for providing SNSPDs, and Jim MacArthur for assistance with electronics. This work was supported by the NSF, CUA, DoD/ARO DURIP, AFOSR MURI, ONR MURI, ARL, and a Vannevar Bush Faculty Fellowship. Devices were fabricated at Harvard CNS, NSF award no. 1541959. M. K. B. and D. S. L. acknowledge support from an

NDSEG Fellowship. R. R. acknowledges support from the Alexander von Humbolt Foundation. B. M. and E. N. K. acknowledge support from an NSF GRFP.

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to M. D. L. at lukin@physics.harvard.edu