A high-speed noise-free optical quantum memory

K. T. Kaczmarek[1,*], P. M. Ledingham[1], B. Brecht[1],
S. E. Thomas[1,2], G. S. Thekkadath[3], O. Lazo-Arjona[1],
J. H. D. Munns[1,2], E. Poem[1,4], A. Feizpour[1], D. J. Saunders[1],
J. Nunn[1], I. A. Walmsley[1]

[1]Clarendon Laboratory, University of Oxford, Parks Road, Oxford, OX1 3PU, UK.
[2]QOLS, Blackett Laboratory, Imperial College London, London SW7 2BW, UK.
[3]University of Ottawa, 25 Templeton St, Ottawa, K1N 6N5, Canada.
[4]Department of Physics of Complex Systems, Weizmann Institute of Science,
Rehovot 7610001, Israel.

E-mail: *k.kaczmarek1@physics.ox.ac.uk

Light is an ideal information carrier for quantum networks [1]: its quantum properties are not degraded by noise in ambient conditions and it has a large information capacity owing to its high bandwidth. The realisation of scalable high-speed photonic quantum networks, e.g. for quantum simulation and information processing, requires noise-free, low-latency, and technically simple quantum optical memories to overcome the exponentially poor scaling of current approaches based on linear optics. While low noise [2, 3, 4, 5, 6, 7], high speed [8], and operation in ambient conditions [9, 10, 11, 12] have been shown separately, to date no atomic system satisfies all these desiderata simultaneously. Here we introduce the off-resonant cascaded absorption (ORCA) memory, and demonstrate for the first time the storage and recall of GHz-bandwidth single photons with zero added noise in a warm caesium vapour. Because of its technical simplicity, its high speed and its low noise, ORCA provides a viable route towards next generation, low-latency quantum-network technologies.

The basic operating principle of a quantum optical memory is to prepare an empty storage state $|s\rangle$ and then map on-demand a quantum state of light to and from an atomic coherence between the ground state $|g\rangle$ and $|s\rangle$. With an additional intermediate atomic level $|e\rangle$ and an auxiliary control field this can be achieved - the input "signal" field drives the "populated transition" ($|g\rangle \leftrightarrow |e\rangle$), while the control field drives the "empty transition" ($|e\rangle \leftrightarrow |s\rangle$), as shown in Fig. 1. By using atomic ensembles, one increases the probability of a photon interacting with an atom, leading to strong light-matter coupling [13]. Traditionally the energy levels are arranged in a Λ configuration: two ground states coupled via an excited state, since the long ground state coherence time gives prospects for long-distance quantum communication applications [14]. Narrowband ($\sim$MHz), resonant Λ-type quantum optical memories based on electromagnetically-induced transparency (EIT) [2, 15], gradient echo memory (GEM) [10] and the full atomic frequency comb (AFC) [16] protocol have performed well, with high efficiency [17] and preservation of photon statistics [7] being demonstrated. However, their experimentally accessible narrow bandwidths combined with the additional need for preparation of the storage state (by means of optical pumping or spectral holeburning) limits their maximum clock-rate. Focussing on the requirements of local quantum networks, being able to operate at a high clock-rate for elementary building block synchronisation motivates the need for a broadband memory for short pulses of light [18].

The Λ configuration does allow the freedom of having the signal and control fields far-off single-photon resonance while remaining in two-photon resonance, as in the Raman memory scheme [19, 20]. In this protocol, the control creates a broadband "virtual" resonance that the signal couples to. Because the signal is far detuned from the "real" atomic resonance, there is very little linear absorption. Interestingly, this means the system can be reduced to a beamsplitter-like interaction between different travelling optical modes and a stationary matter mode, where the "reflectivity" of the

beamsplitter is determined by the storage and recall efficiencies [21, 22, 23]. The caveat is that, in the absence of atomic selection rules, the control can off-resonantly couple to the populated transition, inducing spontaneous Raman scattering and polluting the storage state with spurious atomic excitations [24]. The latter leads to four-wave mixing (FWM) noise in the same optical mode as the desired quantum signal when the memory is read out, deteriorating its photon statistics [8, 24]. Noise contamination of the signal, due to control field induced fluorescence or especially FWM in off-resonant broadband memories, remains one of the greatest challenges prohibiting network ready quantum memories.
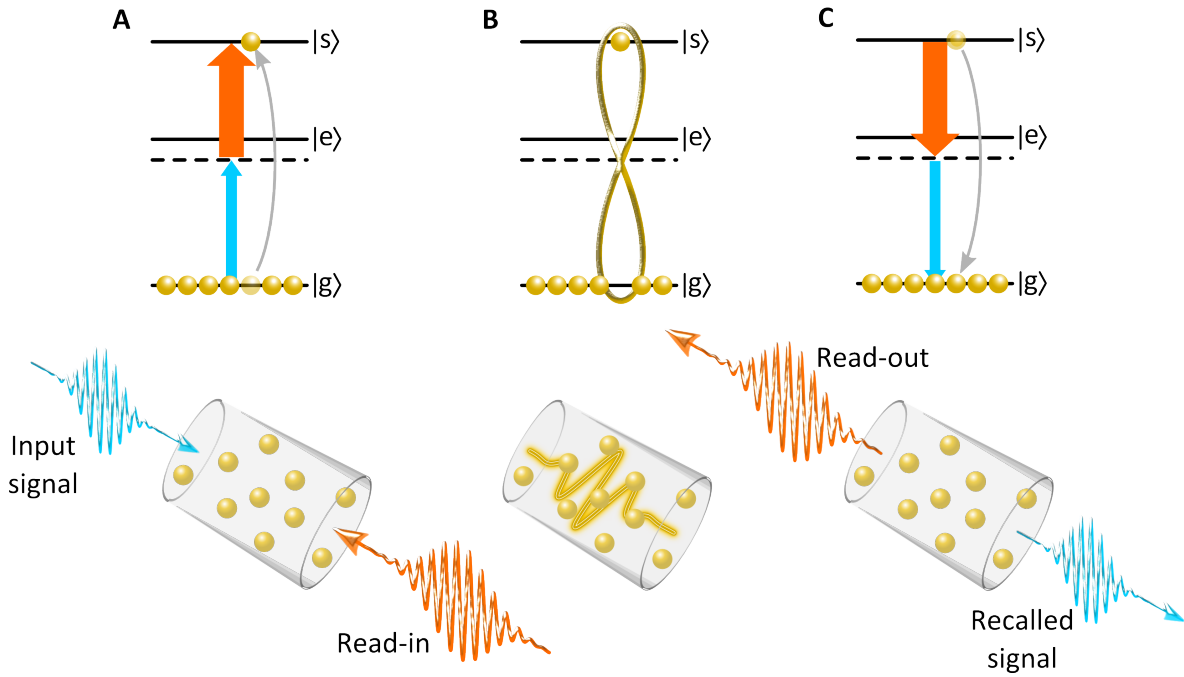
Figure 1. The ORCA protocol.
a, Storage: a weak "signal" field (blue) and strong "control" (orange) are overlapped counter-propagating in an atomic vapour. The broadband fields are on two-photon resonance with a doubly-excited storage state $|s\rangle$, while remaining far-detuned from the intermediate state $|e\rangle$. **b,** The storage process maps the signal field to an atomic coherence (dashed line) between the ground $|g\rangle$ and storage $|s\rangle$ states. Motion-induced dephasing of the coherence is reduced owing to the counter-propagating field configuration. **c,** Recall: re-applying the control pulse after the desired storage time leads to re-mapping of the atomic coherence back into an optical field and thus re-emission of the signal in the forward direction.

Here we introduce a new quantum memory scheme based on off-resonant cascaded absorption (ORCA) in a three-level ladder configuration that is optimised for warm atomic vapours (see Fig. 1). The operational principle is similar to the Raman memory: two broadband optical fields, signal and control, satisfy a two-photon resonance condition with a storage state, while individually being far-off resonance from their respective single-photon transitions. Unlike in Λ-type protocols, the ORCA memory

bandwidth is not in principle limited by the ground state splitting, and the counter-propagating beam geometry reduces motional-induced dephasing of the stored coherence (see Supplementary Information), giving the potential for functional time-bandwidth products, i.e. the number of clock-cycles over which a memory allows synchronisation [25].

In the new scheme, the storage state is a doubly-excited electronic state, which has no thermal excitations even at high temperatures. Therefore the protocol in principle requires no preparation of the atomic ensemble prior to storage, and there is no contamination of the recalled fields due to imperfect optical pumping. This points to the major benefit of ORCA storage in that it is *noise-free*. The signal and control wavelengths can be chosen such that the control photons are significantly detuned from the populated transition (THz detunings are readily available in the rich atomic structure of alkalis). This effectively eliminates any control field induced fluorescence noise. More importantly though, due to the ladder configuration, there is no scattering process that could populate the storage state, and so four-wave mixing noise is eliminated. Finally, efficient suppression of control field leakage on the output detection is readily achievable using off-the-shelf low-loss interference filters, and bolstered by the counter-propagating field configuration, in principle enabling *end-to-end* device efficiencies approaching the *internal* memory efficiency. The ORCA scheme is therefore a broadband and noise-free quantum optical memory protocol operating within a technically simple, room-temperature platform.

As a proof-of-principle demonstration, we implement the protocol with near-infrared light in warm caesium vapour (see Methods section for more details). We interface the ORCA memory with a 1 GHz bandwidth heralded single-photon source based on parametric down-conversion (PDC) [24], as shown in Fig. 2. In type-II PDC, a "pump" field is converted into orthogonally polarised pairs of photons by means of spontaneous scattering. In the weakly-pumped regime, when the pair production rate is low, the detection of a photon in one of the modes ("idler") heralds the presence of another ("signal"), which we attempt to store. We choose a storage time of 3.5 ns, as residual motion-induced dephasing limits the memory lifetime in the current implementation in Cs vapour to 5.4 ns (see Supplementary Information). Owing to the short lifetime, and without the need for time-consuming atomic state preparation, we are able to operate the experiment at the full 80 MHz repetition rate of our pump, as shown in Fig. 3a.

Fig. 3a shows a section of the photon arrival time trace with the control field off ("SIG") and on ("MEM"). We measure a memory efficiency (storage and recall) of $\eta = 14.6 \pm 1.9$ %. Higher efficiencies can be reached by stronger driving, increased atomic density, and appropriate shaping of the control pulse, as predicted by existing theory of quantum memories based on three-level systems [26, 27, 28]. Due to the absence of four-wave mixing gain [13] however, ORCA is in principle capable of reaching these high efficiencies without adding noise.

We benchmark the noise performance of the memory by evaluating $\mu_1 = \langle n^{\mathrm{noise}} \rangle / \eta$,
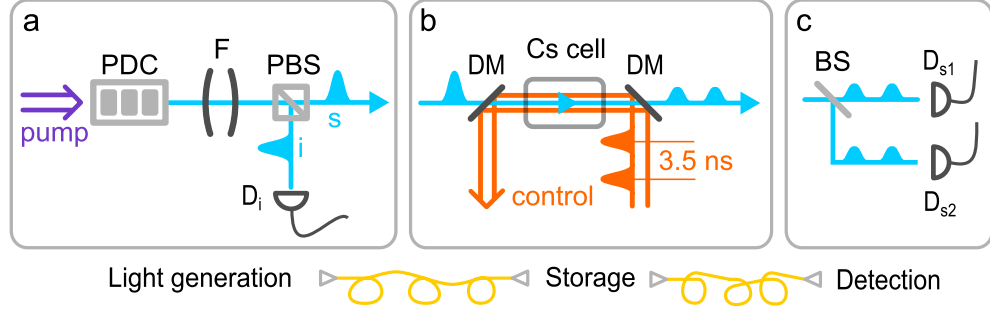
Figure 2. Experimental setup. a, Photon pairs at 852 nm are generated in parametric down-conversion (PDC) and consecutively filtered (F) to $\sim$ 1 GHz bandwidth. The signal (s) and idler (i) photons are separated on a polarizing beam splitter (PBS). While the idler photons are detected with a single-photon detector (D_i), the signal photons are sent to the memory. **b,** Dichroic mirrors (DM) are used to combine the signal field with a bright, counter-propagating control field at 917 nm (read-in and read-out) inside a caesium vapour cell. After the memory, the transmitted and recalled light is sent to the detection stage. **c,** At the detection stage, the signal photons are split on a balanced beam splitter (BS) and detected by two single-photon detectors (D_{s1} and D_{s2}) in a Hanbury-Brown-Twiss configuration.

i.e. the ratio of the average number of noise photons per control pulse $\langle n^{\mathrm{noise}} \rangle$ and η [29]. We find $\mu_1 \leq (39.4 \pm 0.2) \times 10^{-6}$, the lowest ever reported from an atom-based quantum memory. Note that the measured μ_1 is an upper bound, limited by the technical noise (dark counts) of our detectors.

To verify the quantum performance of the memory we measure the photon number statistics and correlations of the recalled signal, and compare it with the input. Fig. 3b shows the detected coincidence clicks between the detectors D_i & $D_{s1/2}$ at different times with the control off ("SIG") and on ("MEM"). First, we evaluate the cross-correlation function $g^{(1,1)}$ of signal and idler photons (see Methods section). Values of $g^{(1,1)} > 2$ signify non-classical correlations [30]. We measure $g^{(1,1)} = 131.3 \pm 0.2$ for the input signal field, and upon recall obtain $g^{(1,1)} = 120.0 \pm 0.1$, clearly exceeding the classical bound and demonstrating the preservation of non-classical correlations in ORCA. We attribute the reduction of the $g^{(1,1)}$ in the read-out to dark count contamination.

In order to determine whether ORCA preserves the quantum statistics of the input, we next evaluate the second-order heralded auto-correlation function $g_{\mathrm{h}}^{(2)}$ (see Methods section). A value of $g_{\mathrm{h}}^{(2)} < 1$ is a direct measure of quantum statistics, with $g_{\mathrm{h}}^{(2)} = 0$ corresponding to a perfect single photon [3]. We measure the $g_{\mathrm{h}}^{(2)}$ of our input signal (Fig. 3c) to be 0.020 ± 0.005, as expected from a weakly pumped PDC source [24]. Upon recall (Fig. 3d), we obtain $g_{\mathrm{h}}^{(2)} = 0.028 \pm 0.009$, the lowest ever measured from a room-temperature quantum memory. The agreement between input and output $g_{\mathrm{h}}^{(2)}$ provides further strong experimental confirmation that the atomic memory adds zero noise, similarly to a non-atomic memory based on electronically switched free-space cavities [31].

In conclusion, we have demonstrated the first noise-free broadband atomic quantum
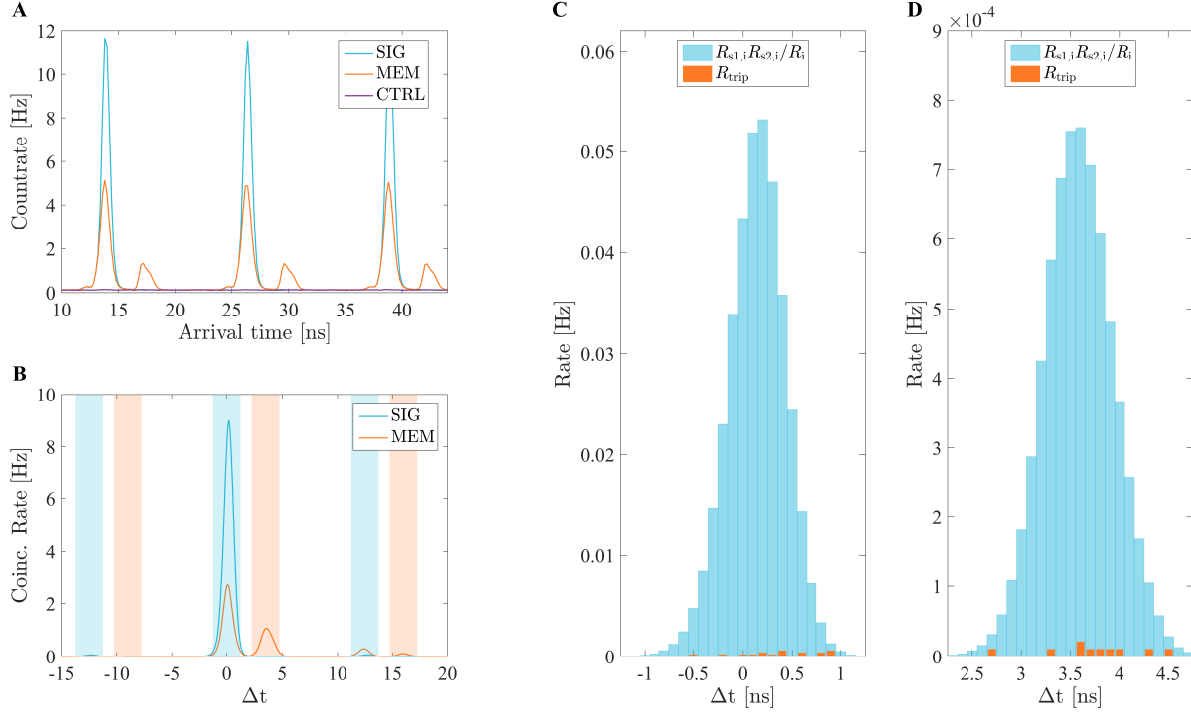
Figure 3. Storage of single photons using ORCA.
a, Histogram showing timing of total signal detector clicks with respect to external 1 MHz trigger. "SIG" is the signal field with the control field off, "MEM" is the signal field with the control on for a storage time of 3.5 ns. "CTRL" has the control field on, but with no input signal. Any photons detected in the "CTRL" configuration correspond to noise generated in the memory. **b,** Histogram of the time difference Δt between idler and total signal detector clicks, with the control off ("SIG") and on ("MEM"). The blue (orange) shaded areas correspond to the integration windows for the SIG (MEM) setting. The peaks at 12.5 and 16 ns in the "MEM" trace come from consecutive control pulses reading out residual coherence (see Supplementary Information). **c,** Input signal: orange - histogram of the time difference Δt between idler detector clicks, and signal detector coincidences, i.e. triple coincidence histogram R_{trip}; blue - product of the two-fold coincidences between the idler and two signal detectors, $R_{\mathrm{s1,i}}$ and $R_{\mathrm{s2,i}}$, normalized by the idler counts R_{i}, i.e. predicted triple coincidence histogram for coherent light of the same average photon rate as the PDC. The ratio between the two histograms corresponds to the measured heralded auto-correlation function $g_{\mathrm{h}}^{(2)}$. **d,** Same as **c**, but for the recalled signal.

memory protocol at room temperature, based on off-resonant cascaded absorption (ORCA), successfully storing and recalling a single photon from a warm vapour containing $\sim 10^{10}$ moving atoms interacting with $\sim 10^9$ control photons.

In the current experiment in Cs we achieved a time-bandwidth product of ~ 5. By using a different atomic species with smaller signal/control wavelength mismatch, the memory lifetime can be significantly extended. Recently, Finkelstein et al. [25] demonstrated a fast ladder memory (FLAME) – equivalent to ORCA when far-detuned – with classical signal pulses in rubidium vapour. Here, the near-degeneracy of the signal and control fields allowed close to Doppler-free operation, with a memory lifetime

~ 85 ns, granting Finkelstein et al. a time-bandwidth product of ~ 50.

With foreseeable improvements in memory efficiency and lifetime, ORCA will be capable of achieving the performance necessary for high clock-rate synchronisation of probabilistic operations in local quantum networks [18]. Furthermore, ORCA is compatible with waveguide structures [32, 33], enabling efficient interfacing with integrated optical circuits [34]. This will facilitate the creation of photonic quantum states of unprecedented scale, opening the way to a new regime of quantum simulation, computation and sensing.

Methods

Experiment

We use the Cs D2 line at 852 nm for our signal field, with $6S_{1/2}(F = 4)$ as the ground state $|g\rangle$ and the $6P_{3/2}(F = 3, 4, 5)$ manifold as the ORCA intermediate state $|e\rangle$. A strong 917 nm control field couples this signal to the storage state $|s\rangle$, i.e. the $6D_{5/2}(F = 2, 3, 4, 5, 6)$ manifold. This configuration can be reasonably treated as a three-level system under broadband excitation [35]. We detune both fields by 6 GHz from the intermediate state towards the ground state, enabling good coupling with negligible ($< 2\%$) linear absorption.

The generation of heralded single photons is achieved using type-II parametric down-conversion (PDC) in a periodically poled potassium titanyl phosphate waveguide. The source generates THz-bandwidth pairs of signal and idler photons, both of which are consequently filtered down to ~ 1 GHz bandwidth centred at the signal frequency using a series of Fabry-Perót etalons and grating filters [24]. We send the PDC idler to a single-photon silicon avalanche photodiode (APD) and the signal towards the memory. Our heralding efficiency before the memory is $\eta_{\text{herald}} \approx 5\%$. The source is pumped at a rate of 80 MHz from a frequency doubled actively mode-locked titanium sapphire laser, with a $\sim 0.8\%$ chance of producing a photon pair of the correct bandwidth per pump pulse.

The control field is derived from a second mode-locked titanium sapphire laser, locked-to-clock with the PDC pump. The laser bandwidth is ~ 0.9 GHz, corresponding to a pulse duration of ~ 500 ps. In order to investigate storage times < 12.5 ns, i.e. smaller than the time between consecutive pulses from the laser, we use an unbalanced Mach-Zehnder interferometer with a variable delay in one arm to split the control pulse train into two: read-in and read-out, and delay them with respect to each other. The read-in and read-out control pulse energies are 0.21(1) and 0.97(1) nJ, respectively.

We combine the signal and counter-propagating control fields on a dichroic mirror. Both beams are focused down to a ~ 300 μm waist and temporally overlapped inside a 72 mm -long uncoated glass Cs-133 vapour cell heated to $\sim 91^{o}$C.

After the signal field leaves the memory it is sent into a Hanbury-Brown-Twiss detection setup, composed of a balanced beamsplitter and two fibre-coupled single-photon silicon APDs connected to a time-to-digital converter (same as the idler detector), allowing us to reconstruct the quantum photon number statistics of the stored/recalled signal fields.

Photon statistics

By evaluating the $g^{(1,1)}$ cross-correlation function between the signal and idler (herald) pulses, we obtain a measure for the strength of the correlations between them. $g^{(1,1)}$ is defined as $p_{\text{si}}/p_{\text{s}}p_{\text{i}}$, where p_{si} is the probability for a signal-idler coincidence click, and $p_{\text{s(i)}}$ is the signal (idler) click probability. Values of $g^{(1,1)} > 2$ signify non-classical

correlations. To calculate $g^{(1,1)}$ from the measurements, we use

$$g^{(1,1)} = \frac{R_{s,i}}{R_s R_i} R_T,$$ (1)

where $R_{s,i}$ is the sum of D_i-D_{s1} and D_i-D_{s2} coincidences, R_T is the total number of trigger events during the whole measurement time, R_s is the sum of D_{s1} and D_{s2} clicks, and R_i is the number of D_i clicks.

The heralded auto-correlation is defined as $g_h^{(2)} = p_{(s1,s2|i)}/p_{(s1|i)}p_{(s2|i)}$. Here, $p_{(s1,s2|i)}$ is the conditional probability of detecting a coincidence between D_{s1} and D_{s2} given a click in D_i, and $p_{(s1|i)}$ ($p_{(s2|i)}$) is the probability to detect a click in D_{s1} (D_{s2}) given a click in D_i. Any $g_h^{(2)} < 1$ verifies non-classical photon-number statistics. We evaluate $g_h^{(2)}$ using

$$g_h^{(2)} = \frac{R_{trip}}{R_{s1,i} R_{s2,i}} R_i,$$ (2)

where R_{trip} is the number of triple coincidences between D_i, D_{s1}, and D_{s2}, R_i is the number of idler clicks, and $R_{s1(2),i}$ is the number of D_i-D_{s1} (D_i-D_{s2}) coincidences. More details can be found in the Supplementary Information.

Theoretical modeling

An ab-initio model using Maxwell-Bloch equations is used to characterize the memory performance. We obtain equations for the time evolution of velocity class ensembles $\hat{\rho}(v)$ coupled with signal and control fields. We consider 12 atomic states corresponding to the hyperfine levels of the $6S_{1/2}, 6P_{3/2}, 6D_{5/2}$ multiplets. The fields are coupled to the atomic states through a master equation with a Hamiltonian under electric dipole and rotating wave approximations, and in an appropriately chosen rotating frame, as well as Lindblad terms accounting for spontaneous decay5. The control field is taken to be strong enough to propagate without dispersion from the atomic vapour. The signal field is coupled to the total density matrix $\hat{\rho} = \int dv g(v)\hat{\rho}(v)$ (where $g(v)$ is a Maxwell-Boltzmann velocity distribution) through the source term of a wave equation under the slowly-varying envelope approximation.

We numerically solve these equations using experimental parameters with only electric dipole matrix elements as free parameters. Once a good fit is found for the memory efficiencies as a function of control pulse power, the same parameters are used throughout this work.

References

[1] H. J. Kimble. The quantum internet. *Nature*, 453(7198):1023–1030, 2008.

[2] T. Chanelière, D. N. Matsukevich, S. D. Jenkins, S. Y. Lan, T. A. B. Kennedy, and A. Kuzmich. Storage and retrieval of single photons transmitted between remote quantum memories. *Nature*, 438(7069):833, 2005.

[3] S. Zhou, S. Zhang, C. Liu, J. F. Chen, J. Wen, M. M. T. Loy, G. K. L. Wong, and S. Du. Optimal storage and retrieval of single-photon waveforms. *Optics express*, 20(22):24124–31, 2012.

[4] D. Maxwell, D. Szwer, D. Paredes-Barato, H. Busche, J. Pritchard, A. Gauguet, K. Weatherill, M. Jones, and C. Adams. Storage and Control of Optical Photons Using Rydberg Polaritons. *Physical Review Letters*, 110(10):103001, mar 2013.

[5] D. S. Ding, K. Wang, W. Zhang, S. Shi, M. X. Dong, Y. C. Yu, Z. Y. Zhou, B. S. Shi, and G. C. Guo. Entanglement between Lowly and Highly Lying Atomic Spin Waves. *Physical Review A*, 1:1–5, 2016.

[6] E. Distante, P. Farrera, A. Padrón-Brito, D. Paredes-Barato, G. Heinze, and H. de Riedmatten. Storing single photons emitted by a quantum memory on a highly excited Rydberg state. *Nature Communications*, 8:14072, 2017.

[7] A. Seri, A. Lenhard, D. Rieländer, M. Gündoğan, P. M. Ledingham, M. Mazzera, and H. de Riedmatten. Quantum correlations between single telecom photons and a multimode on-demand solid-state quantum memory. *Physical Review X*, 7(2):1–7, 2017.

[8] D. G. England, K. A. G. Fisher, J.-P. W. MacLean, P. J. Bustard, R. Lausten, K. J. Resch, and B. J. Sussman. Storage and Retrieval of THz-Bandwidth Single Photons Using a Room-Temperature Diamond Quantum Memory. *Physical Review Letters*, 114(5):053602, 2015.

[9] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurásek, and E. S. Polzik. Experimental demonstration of quantum memory for light. *Nature*, 432(7016):482–486, 2004.

[10] M. Hosseini, G. Campbell, B. M. Sparkes, P. K. Lam, and B. C. Buchler. Unconditional room-temperature quantum memory. *Nature Physics*, 7(10):794–798, 2011.

[11] F. Ripka, Y.-H. Chen, R. Löw, and T. Pfau. Rydberg polaritons in a thermal vapor. *Physical Review A*, 93(5):5, 2016.

[12] N. Šibalić, J. M. Kondo, C. S. Adams, and K. J. Weatherill. Dressed-state electromagnetically induced transparency for light storage in uniform-phase spin waves. *Physical Review A*, 94(3), 2016.

[13] S. E. Thomas, J. H. D. Munns, K. T. Kaczmarek, C. Qiu, B. Brecht, A. Feizpour, P. M. Ledingham, I. a. Walmsley, J. Nunn, and D. J. Saunders. High Efficiency Raman Memory by Suppressing Radiation Trapping. *New Journal of Physics*, 063034, 2016.

[14] F. Bussières, N. Sangouard, M. Afzelius, H. de Riedmatten, C. Simon, and W. Tittel. Prospective applications of optical quantum memories. *Journal of Modern Optics*, 60(18):1519–1537, 2013.

[15] M. D. Eisaman, A. André, F. Massou, M. Fleischhauer, A. S. Zibrov, and M. D. Lukin. Electromagnetically induced transparency with tunable single-photon pulses. *Nature*, 438(7069):837–41, 2005.

[16] P. Jobez, C. Laplane, N. Timoney, N. Gisin, A. Ferrier, P. Goldner, and M. Afzelius. Coherent Spin Control at the Quantum Level in an Ensemble-Based Optical Memory. *Physical Review Letters*, 114(23):1–5, 2015.

[17] M. P. Hedges, J. J. Longdell, Y. Li, and M. J. Sellars. Efficient quantum memory for light. *Nature*, 465(7301):1052–6, June 2010.

[18] J. Nunn, N. K. Langford, W. S. Kolthammer, T. F. M. Champion, M. R. Sprague, P. S. Michelberger, X.-M. Jin, D. G. England, and I. A. Walmsley. Enhancing Multiphoton Rates with Quantum Memories. *Physical Review Letters*, 110(13):133601, mar 2013.

[19] K. F. Reim, J. Nunn, V .O. Lorenz, and B. J. Sussman. Towards high-speed optical quantum memories. *Nature Photonics*, (March):8–11, 2010.

[20] R. Chrapkiewicz, M. Dąbrowski, and W. Wasilewski. High-Capacity Angularly Multiplexed

Holographic Memory Operating at the Single-Photon Level. *Physical Review Letters*, 118(6):063603, 2017.

[21] K. F. Reim, J. Nunn, X.-M. Jin, P. S. Michelberger, T. F. M. Champion, D. G. England, K. C. Lee, W. S. Kolthammer, N. K. Langford, and I. A. Walmsley. Multipulse Addressing of a Raman Quantum Memory: Configurable Beam Splitting and Efficient Readout. *Physical Review Letters*, 108(26):263602, jun 2012.

[22] G. T. Campbell, O. Pinel, M. Hosseini, T. C. Ralph, B. C. Buchler, and P. K. Lam. Configurable unitary transformations and linear logic gates using quantum memories. *Physical Review Letters*, 113(6):1–5, 2014.

[23] K. A. G. Fisher, D. G. England, J.-P. W. MacLean, P. J. Bustard, K. J. Resch, and B. J. Sussman. Frequency and bandwidth conversion of single photons in a room-temperature diamond quantum memory. *Nature communications*, 7:11200, 2016.

[24] P. S. Michelberger, T. F. M. Champion, M. R. Sprague, K. T. Kaczmarek, M. Barbieri, X. M. Jin, D. G. England, W. S. Kolthammer, D. J. Saunders, J. Nunn, and I. A. Walmsley. Interfacing GHz-bandwidth heralded single photons with a room-temperature Raman quantum memory. *New Journal of Physics*, 17(4):43006, 2015.

[25] R. Finkelstein, E. Poem, O. Michel, O. Lahad, and O. Firstenberg. Fast, noise-free memory for photon synchronization in room temperature. arXiv:1708.01919. (2017).

[26] J. Nunn, I. A. Walmsley, M. Raymer, K. Surmacz, F. Waldermann, Z. Wang, and D. Jaksch. Mapping broadband single-photon wave packets into an atomic memory. *Physical Review A*, 75(1):011401, jan 2007.

[27] I. Novikova, A. V. Gorshkov, D. F. Phillips, A. S. Sørensen, M. D. Lukin, and R. L. Walsworth. Optimal control of light pulse storage and retrieval. *Physical Review Letters*, 98(24):1–4, 2007.

[28] A. V. Gorshkov, A. André, M. Fleischhauer, A. S. Sørensen, and M. D. Lukin. Universal approach to optimal photon storage in atomic media. *Physical Review Letters*, 98(12):1–4, 2007.

[29] M. Gündoğan, P. M. Ledingham, K. Kutluer, M. Mazzera, and H. De Riedmatten. Solid State Spin-Wave Quantum Memory for Time-Bin Qubits. *Physical Review Letters*, 114(23):1–5, 2015.

[30] P. Farrera, N. Maring, B. Albrecht, G. Heinze, and H. de Riedmatten. Non-classical correlations between a C-band telecom photon and a stored spin-wave. *Optica*, 3(9):1019, 2016.

[31] F. Kaneda, B. G. Christensen, J. J. Wong, H. S. Park, K. T. McCusker, and P. G. Kwiat. Time-multiplexed heralded single-photon source. *Optica*, 2(12):1010, 2015.

[32] M. R. Sprague, P. S. Michelberger, T. F. M. Champion, D. G. England, J. Nunn, X.-M. Jin, W. S. Kolthammer, A. Abdolvand, P. St. J. Russell, and I. A. Walmsley. Broadband single-photon-level memory in a hollow-core photonic crystal fibre. *Nature Photonics*, 8(4):287–291, mar 2014.

[33] K. T. Kaczmarek, D. J. Saunders, M. R. Sprague, W. S. Kolthammer, A. Feizpour, P. M. Ledingham, B. Brecht, E. Poem, I. A. Walmsley, and J. Nunn. Ultrahigh and persistent optical depths of caesium in Kagomé-type hollow-core photonic crystal fibres. *Optics Letters*, 40(23):1–5, 2015.

[34] N. C. Harris, G. R. Steinbrecher, J. Mower, Y. Lahini, M. Prabhu, T. Baehr-Jones, M. Hochberg, S. Lloyd, and D. Englund. Bosonic transport simulations in a large-scale programmable nanophotonic processor. *Nature Photonics*, 11(7):447–452, 2017.

[35] B. Huber, T. Baluktsian, M. Schlagmüller, A. Kölle, H. Kübler, R. Löw, and T. Pfau. GHz Rabi Flopping to Rydberg States in Hot Atomic Vapor Cells. *Physical Review Letters*, 107(24):243001, December 2011.

[36] B. Zhao, Y.-A. Chen, X.-H. Bao, T. Strassel, C.-S. Chuu, X.-M. Jin, J. Schmiedmayer, Z.-S. Yuan, S. Chen, and J.-W. Pan. A millisecond quantum memory for scalable quantum networks. *Nature Physics*, 5(2):95–99, 2009.

Supplementary Information

S.1. Memory lifetime

There are three effects limiting the lifetime of this ORCA memory: spontaneous emission from the doubly excited atomic state, motion-induced dephasing due to the Doppler
5 effect, and destructive interference between hyperfine components of the excited state.

Motion-induced dephasing arises due to the Doppler effect and atomic motion in a warm ensemble. This is because the stored excitation is spread over atoms belonging to different velocity classes in the ensemble. Due to the Doppler effect, each velocity class experiences the signal and control fields shifted in frequency. As a consequence,
10 the phase of the stored coherence evolves at different rates in different velocity classes. The motion-induced dephasing lifetime is [36] $\tau_D = k_r v_s$, where $v_s = \sqrt{k_B T/m}$, and $k_r = \frac{2\pi}{\lambda_s} - \frac{2\pi}{\lambda_c}$; with T the temperature of the atomic vapour, m the mass of the atom, and $\lambda_{s/c}$ the wavelengths of the signal and control fields. In other words, the collective coherence will dephase at a rate proportional to the square root of the temperature, and
15 the wavenumber mismatch of the fields.

In the absence of optical pumping, the off-resonant two-photon excitation that stores the signal photon has contributions from all allowed paths connecting the $6S_{1/2}$ and $6D_{5/2}$ manifolds. The resulting excitation is spread across components of the $6D_{5/2}$ manifold. During storage time these components oscillate with different rates as given
20 by their energy differences, and at read-out time they can interfere destructively in the creation of the read-out photon. Optical pumping restricting the memory interaction to the hyperfine levels $F = 4 \rightarrow F = 5 \rightarrow F = 6$ reduces this effect.
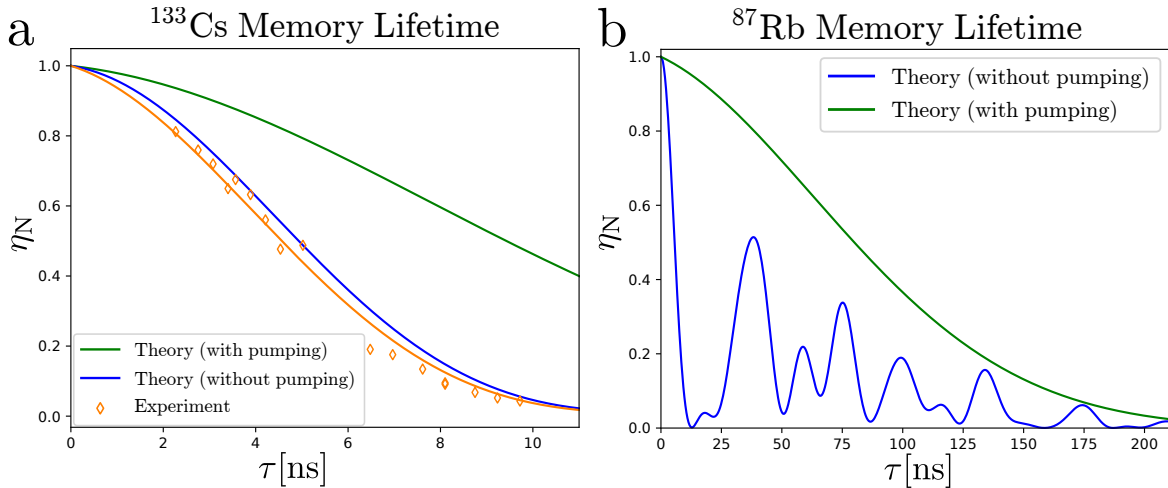
Figure S.1. a. The measured (normalised to $\tau = 0$) memory efficiency η_N (orange diamonds, experimental errors are smaller than the markers) versus storage time τ along with a theory fit of the memory lifetime curve (orange line) yielding a (1/e) lifetime of 5.5 ± 0.1 ns. Also shown is the predicted memory lifetime curves with (green) and without (blue line) optical pumping. **b**. Memory lifetimes predicted from theory for ^{87}Rb with (green) and without (blue line) optical pumping.

We determine the actual memory lifetime by measuring the memory efficiency for different storage times using a weak coherent state signal. In Fig. S.1a we show the measured normalized memory efficiency versus storage time (orange diamonds). Fitting our model to the data we obtain a $1/e$ lifetime of 5.5 ± 0.1 ns. Using the Maxwell-Bloch model directly we predict a memory lifetime of 5.9 ns, very close to the measured one.

While seemingly short, the memory lifetime can be extended through optical pumping to reduce the destructive interference of hyperfine components or by using a different atomic species with a smaller signal/control wavenumber mismatch. If we choose the dipole couplings in the model such that only the transition $F = 4 \rightarrow F = 5 \rightarrow F = 6$ is allowed (as with optical pumping), we obtain a memory lifetime of 11.5 ns (the green curve in figure S.1a). Furthermore, a simulation of the $5S_{1/2} \rightarrow 5P_{3/2} \rightarrow 5D_{5/2}$ cascade in ^{87}Rb (signal at 780 nm, control at 776 nm) yields a memory lifetime of 99 ns as shown by the green curve in S.1b.
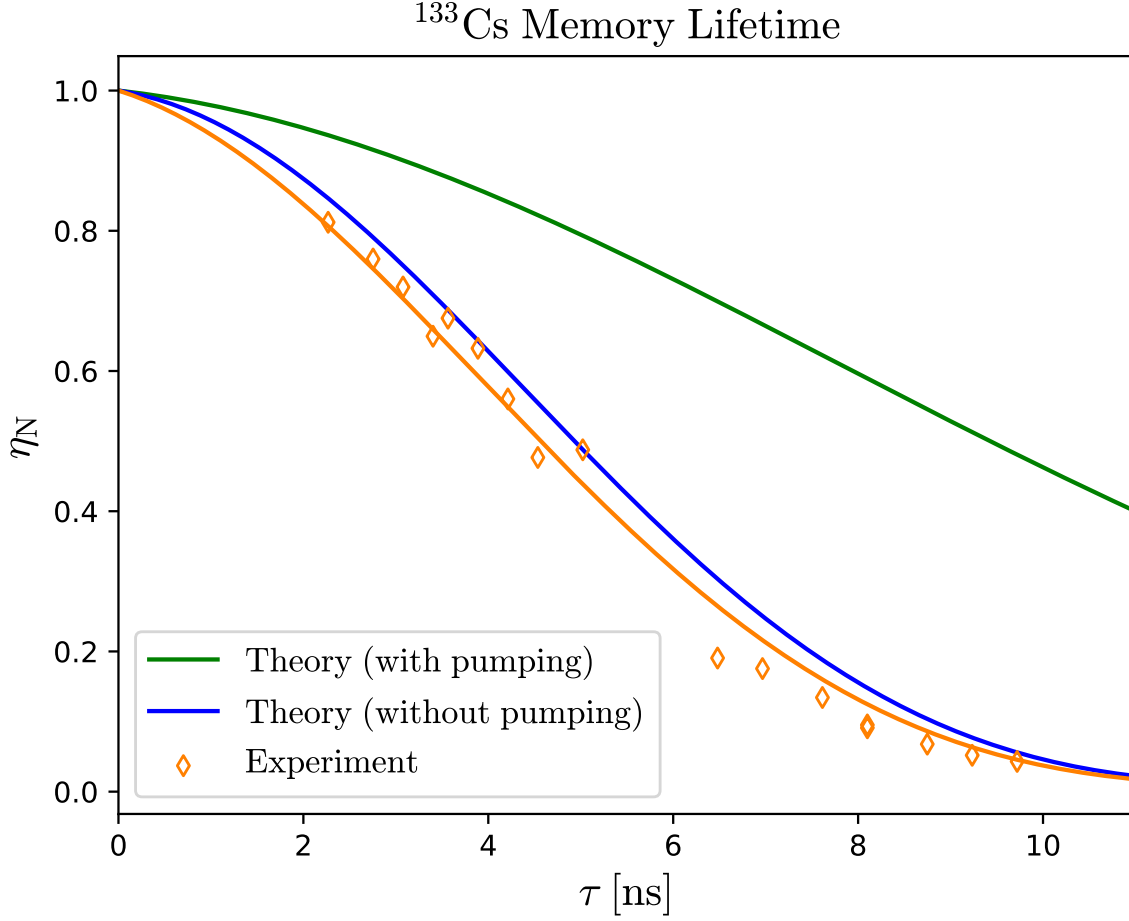
S.2. Experimental setup

Figure S.2 shows a schematic of the ORCA experimental setup. For the memory lifetime measurement (section S.1) we produce a weak coherent state signal (average photon number of ~ 2) by picking pulses using a fast Pockels cell (extinction 20,000:1) at a 1 MHz rate from a 80 MHz train of pulses generated by a ~ 330 ps actively mode-locked titanium sapphire (Ti:Sapph) laser operated at 852 nm and filtered by a Fabry-Pérot (FP) etalon down to 0.81 GHz bandwidth. Using a scanning FP etalon connected to a PC running LabVIEW, we reference-lock the Ti:Sapph's center frequency (via the voltage on a Gires-Tournois-Interferometer inside the laser cavity) to a continuous wave (CW) laser locked to the Cs D2 line via saturated absorption spectroscopy.

We generate the control field from a second ~ 500 ps actively mode-locked Ti:Sapph laser operated at 917 nm, with its center frequency locked using a wavelength meter, and its repetition rate locked to the signal Ti:Sapph using a commercial lock-to-clock (L2C) system. We use an unbalanced free-space Mach-Zehnder interferometer to split the 80 MHz pulse train into two, with a variable delay $< \sim 4$ ns between them, in order to investigate storage times < 12.5 ns. For storage times 6 ns $< \tau < 12.5$ ns we use the L2C electronics to change the timing between signal and control pulse trains such that read-in and read-out are switched. We also use the L2C to temporally overlap the signal and control pulses in the memory cell.

We combine the signal and control fields on a dichroic mirror, which - followed by a 10 nm bandpass filter centred at 850 nm - reduces control field leakage to the detectors from back-reflections by a factor of $\sim 10^9$. We focus signal and control beams down to a ~ 300 μm waist inside a 72 mm -long uncoated caesium borosilicate reference cell heated using a custom-made oven. We estimate the cell temperature to be $\sim 91^o$C by frequency scanning a weak CW probe laser over the Cs D2 line and fitting a Voigt profile to the measured atomic absorption line.

After the signal field passes through the memory and the filters, we send it into

Figure S.2. Schematic of the experimental setup for ORCA. **(a)** Signal generation stage. **(b)** Control generation stage. **(c)** Memory and detection stage. SHG - periodically poled potassium titanyl (ppKTP) bulk crystal; ppKTPw - ppKTP waveguide; DM - dichroic mirror; FP - Fabry-Pérot etalon; PBS - polarizing beamsplitter; HWP - half-wave plate; QWP - quarter-wave plate; APD - avalanche photodiode detector; BF - bandpass filter.

a Hanbury-Brown-Twiss setup, composed of a half-waveplate, polarising beamsplitter and two fibre-coupled single-photon avalanche photodiodes. The two signal and the idler avalanche photodiodes were connected to a time-to-digital converter. For the weak coherent state data and cross-correlation measurements, we add the counts on the two
5 signal detectors to estimate the total magnitude of the transmitted/recalled signal.

S.3. Photon source and setup losses

The generation of heralded single photons is achieved using type-II parametric down-conversion in a ppKTP waveguide. The waveguide, operated in a single-pass configuration and of length 20 mm, is pumped with pulses of approximately 270 ps
10 duration at a wavelength of 426 nm. This pump light is derived by doubling the above mentioned 852 nm Ti:Sapph laser via second harmonic generation in a separate 2 mm

long ppKTP crystal. With an incident average power near $700\,\mathrm{mW}$ at $852\,\mathrm{nm}$ at the crystal we arrive with $4\,\mathrm{mW}$ average power at $426\,\mathrm{nm}$ before the PDC waveguide. This light is then coupled to the waveguide with a total transmission of $<10\%$ including the loss at the in- and out-coupling lenses. We note that the waveguide is not single-mode for the pump wavelength and that the coupling is optimised to primarily excite the fundamental spatial mode, resulting in a low overall transmission. The generated frequency-degenerate but polarisation-orthogonal signal and idler modes have a bandwidth on the order of $1\,\mathrm{THz}$.

We characterise beam propagation transmission using an "alignment" mode, which is coupled to the fundamental mode of the waveguide, and thereby comparable to the signal and idler modes allowing for "classical" measurements to be made. These modes are then subject to frequency filtering. First, we apply coarse filtering using edge filters. Then, the modes are spatially separated via a PBS to then be coupled to their own single-mode fibre (SMF) with an efficiency of $(64 \pm 1)\%$ for the signal mode and $(53 \pm 2)\%$ for the idler mode. The modes are then out-coupled and recombined on a PBS forming a common spatial mode to then pass two etalons, one of $\mathrm{FSR} = 18.4\,\mathrm{GHz}$ and one of $103\,\mathrm{GHz}$, which gives an effective FSR of 1 THz (lowest common multiple). This is followed by a holographic volume Bragg grating with a width $\sim 100\,\mathrm{GHz}$. Using a narrowband ($\sim$ MHz) probe the measured width this filtering has is $1\,\mathrm{GHz}$ for both modes. Finally the modes are separated spatially again via a PBS, the idler coupled to an APD (efficiency $\eta \approx 50\%$, dark counts $= 163 \pm 1\,\mathrm{Hz}$) via a multimode fibre (total transmission from after waveguide to in front of idler detector is $\eta_{\mathrm{i,filt}} = (9.7 \pm 0.1)\,\%$), while the signal is coupled to a SMF to be out-coupled and steered to the memory (total transmission from after waveguide to after this SMF is $\eta_{\mathrm{s,filt}} = (12.8 \pm 0.3)\,\%$).

The filtered signal photon is now steered toward the memory. First there is an edge filter which is used to prevent the $917\,\mathrm{nm}$ control from backward-coupling toward the source which presents additional loss to the signal mode. Further, the caesium cell used is uncoated, adding more loss. After passing the cell, the signal mode is then separated from the control mode via a dichroic mirror and finally passes a bandpass filter about $852\,\mathrm{nm}$ before entering a Hanbury-Brown-Twiss set-up. The mode is spatially separated into two and coupled to two APDs ($\eta \approx 50\,\%$ dark counts $= 296 \pm 2\,\mathrm{Hz}$ and $\eta \approx 50\,\%$ dark counts $= 356 \pm 2\,\mathrm{Hz}$) via SMF. The total transmission from the source to in front of these detectors (averaging over the two SMF couplings) is $\eta_{\mathrm{s,total}} = 3.7 \pm 0.1\,\%$. That is to say, the photon undergoes an additional $\eta_{\mathrm{s,add}} = 30\,\%$ transmission from after the initial filtering stage.

For all results presented in this paper we operated with an average pump power of $4\,\mathrm{mW}$ in front of the waveguide in-coupling lens. Typically, we measure an idler (signal) count rate of around $30\,\mathrm{kHz}$ ($10\,\mathrm{kHz}$) for the case of no control pulses. The typical Klyshko efficiency η_k measured is $0.7\,\%$. This allows to calculate a heralding efficiency of $\eta_{\mathrm{herald}} = \eta_k / \eta_{\mathrm{det}} / \eta_{\mathrm{s,add}} = 4.7\,\%$, which is well above the μ_1 of the memory, as required for single-photon storage [29]. Finally, the heralding efficiency just after the waveguide is $\eta_{\mathrm{s,waveguide}} = \eta_k / \eta_{\mathrm{det}} / \eta_{\mathrm{s,total}} = 38\,\%$. The missing factor of 2.6 we

attribute to not measuring explicitly the loss inside the waveguide, the out-coupling loss from waveguide to free-space and the potential frequency mismatch of the etalon centroids between signal and idler.

S.4. Data acquisition and post-processing

During the measurements, the settings of three mechanical shutters which selectively blocked the read-in, read-out, and signal beams defined four different configurations (see Fig. S.3): memory measurements with all three shutters open ("MEM"); read-in measurements with signal and read-in shutters open, and read-out shutter closed ("RI"); signal measurements with signal shutter open and both read-in and read-out closed ("SIG"); and noise measurements with read-in and read-out shutters open and signal shutter closed ("CTRL"). A single measurement consisted of recording the number of detector counts registered in a period of 180 s in the "MEM" configuration, followed by recording the total counts over 10 s in the "RI", "SIG", and "CTRL" configurations. After completion of all four configurations, the corresponding data was written to disk and the measurement repeated. This mode of operation was chosen in order to mitigate the effect of slow drifts in the setup that arose from changes in the laboratory temperature and humidity.

Figure S.3. Schematic of the different measurement settings. Black rectangles signify a closed shutter. For more information see the text.

For each configuration in each measurement, we recorded arrival time histograms for the three detectors (D_i, D_{s1}, D_{s2}). These are histograms of firing times of the single-photon detectors with respect to a 1 MHz trigger signal derived from the Ti:Sapph recorded with the time-to-digital converter (TDC). We chose a time-bin width of 200 ps as a compromise between temporal resolution of the TDC and total number of time bins in the histogram. For data visualization, we added all arrival time histograms and normalised them to both the number of measurements (521) and the respective measurement duration (180 s for "MEM", 10 s else), obtaining a count rate per time bin in units of Hertz.

Fig. S.4 shows a section of the arrival time histograms. To reduce the impact of spurious noise counts (primarily from detector dark counts), we applied time gates to the recorded arrival time histograms and only kept events that lay within the time gates. The time gates for the read-in pulses (blue regions) are centred around the maxima of the individual read-in peaks and have a width of 2.5 ns, chosen such that the peaks were completely inside the gating region. Similar time gates were chosen for the recalled light, where the centre of these read-out gates (orange regions) was offset
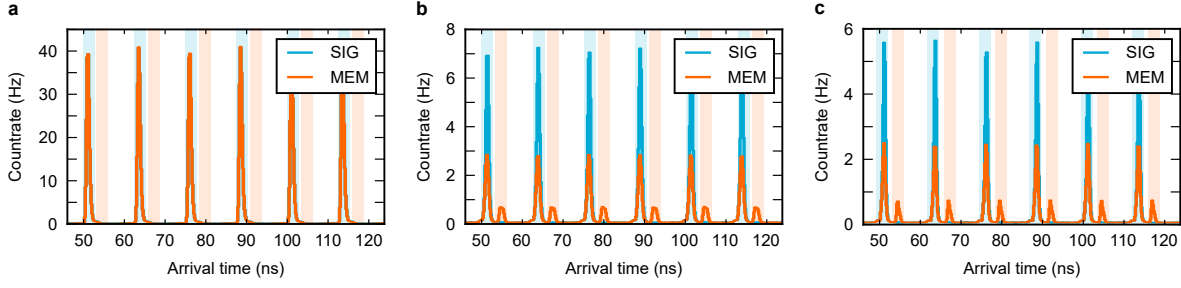
Figure S.4. Sections of the arrival time histograms with indicated time gates. a A section of the arrival time histogram for D_i We show the histograms for the "SIG" (blue trace) and "MEM" (orange trace) configuration. **b** A section of the arrival time histograms for detector D_{s1}. **c** The same for detector D_{s2}. For more details see the text.

from the corresponding read-in time gates by 3.5 ns, which was the storage time chosen for the experiment. By integrating the detection events over only the gate regions, we calculated the read-in, read-out, and total memory efficiencies stated in the main text.

In addition to the arrival time histograms, we also recorded coincidence histograms for signal-idler two-fold coincidences (D_i-D_{s1}, D_i-D_{s2}) and the three-fold coincidences (D_i-D_{s1}-D_{s2}). These are start-stop histograms, where the detection of an idler photon starts the measurement and the detection of a signal photon (the detection of a D_{s1}-D_{s2} coincidence) serves as the stop signal for the two-fold (three-fold) coincidence measurement. Here, the time-bin width of the TDC was chosen to be 100 ps to ensure that the temporal resolution of the measurement was not limited by the TDC, and the time gates had a width of 3.5 ns. Again, the data was post-processed for visualization similar to the arrival time histograms. The resulting coincidence traces are plotted in Fig. S.5. Note that the unconventional shape of the traces originates from the logarithmic scaling of the y axes.

The "SIG" traces show a dominant peak at a time difference of 0 ns, with subsequent smaller peaks at integer multiples of the laser repetition time of 12.5 ns. From this we calculate the $g^{(1,1)}$ signal-idler cross-correlation function. The results for the "SIG" configuration are summarized in the first row of Tab. S.1, where we find $g^{(1,1)} = 131(1)$ for a time difference of 0 ns and $g^{(1,1)} \approx 1$ for integer multiples of 12.5 ns. We also note that the values at the read-out times (3.5 ns offset from the 12.5 ns time slots) are meaningless, since there is no actual signal at the detectors.

Turning our attention to the "MEM" configuration (orange traces in Fig. S.5), we again find a dominating peak at a time difference of 0 ns with side peaks at integer multiples of 12.5 ns. In addition, we see peaks that are offset from the major peaks by 3.5 ns. These originate from coincidence events between idler photons and signal photons that have been stored in and recalled from the memory. We also note that the side peak at 12.5 ns is higher than the corresponding peak for the "SIG" configuration. The reason for this lies in the non-unity read-out efficiency of our memory. A stored photon is not necessarily read out after 3.5 ns, but can remain stored in the memory.
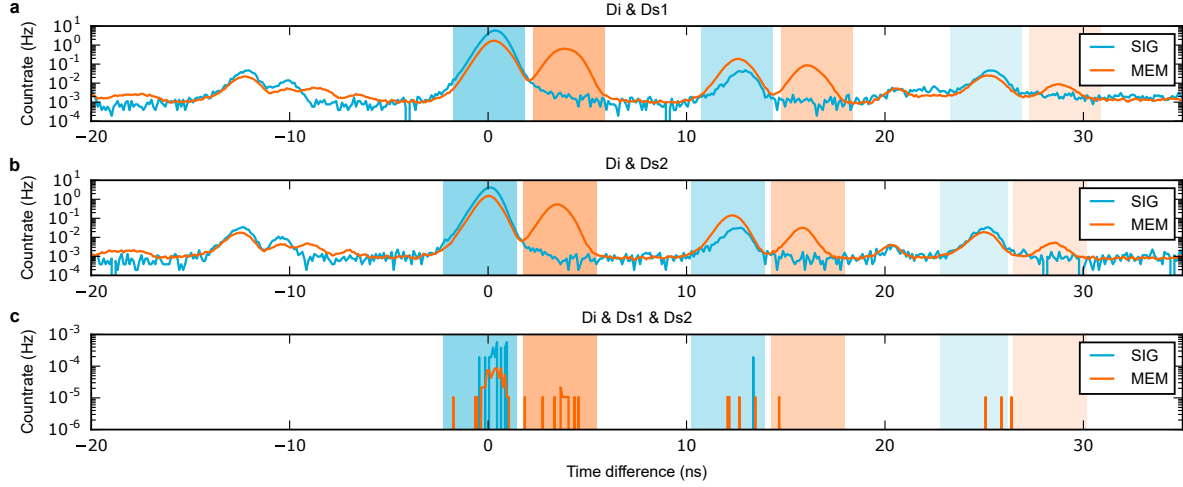
Figure S.5. Correlation histograms. a Accumulated normalised correlation histogram for two-fold coincidences between detectors D_i & D_{s1}, shown for the "SIG" (blue trace) and "MEM" (orange trace) configurations. Note the logarithmic scaling of the y axis. In the main text, we analyse the $g^{(1,1)}(0)$ cross-correlation for the initial time at 0 ns (read-in; blue-shaded) and 3.5 ns (read-out; orange-shaded). Successive read-in (read-out) time bins are indicated by shaded regions with decreasing saturation. **b** The same as in **a**, now however for two-fold coincidences between detectors D_i & D_{s2}. **c** Correlation histogram for three-fold coincidences between detectors D_i & D_{s1} & D_{s2}.

| | $g^{(1,1)}$ | | | | | |
	$\tau = 0$ ns	$\tau = 3.5$ ns	$\tau = 12.5$ ns	$\tau = 16$ ns	$\tau = 25$ ns	$\tau = 28.5$ ns
SIG	131(1)	0.203(1)	1.13(2)	0.0773(1)	1.23(2)	0.137(1)
MEM	80.5(1)	120(1)	9.74(2)	11.3(1)	1.78(2)	1.19(1)

Table S.1. Cross-correlation for successive read-outs. Calculating $g^{(1,1)}$ for higher-order read-outs at integer multiples of 12.5 ns (plus 3.5 ns for read-out pulses) yields the preservation of non-classical correlations by the memory up to around three times the memory lifetime.

Then, it can be read out by the next laser pulse arriving at 12.5 ns, and so on. To quantify this effect, we again evaluate $g^{(1,1)}$. The results are summarized in the second row of Tab. S.1. In this case, we find non-classical values for $g^{(1,1)}$ up to a time of 16 ns, which corresponds to around three times the lifetime of our memory. These results highlight the noise-free operation of ORCA: non-classical photon correlations are retained even after the memory efficiency has decayed to around 5% of its initial maximum value $(1/e^3)$.

Acknowledgements

Author Contributions

²⁵ K.T.K., A.F., E.P. and J.N. invented the protocol. K.T.K., G.S.T., O.L.-A., A.F. and J.N. developed the theory. K.T.K. designed the experiment. K.T.K., P.M.L., B.B., S.E.T., J.H.D.M. and D.J.S. built the setup. K.T.K. and P.M.L. conducted the experiment. B.B. wrote the data acquisition and analysis software. K.T.K., D.J.S., J.N. and I.A.W. supervised the work. All authors contributed to writing the manuscript.

³⁰ Author Information

Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.T.K. (k.kaczmarek1@physics.ox.ac.uk) or I.A.W. (i.walmsley1@physics.ox.ac.uk). K.T.K., A.F., E.P. and J.N. have jointly applied for a patent for the work presented in this ³⁵ paper.