Superradiance-mediated photon storage for broadband quantum memory

Anindya Rastogi,1 Erhan Saglamyurek,1,2 Taras Hrushevskyi,1 and Lindsay J. LeBlanc1,

1Department of Physics, University of Alberta, Edmonton AB T6G 2E1, Canada
2Department of Physics and Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada

Superradiance, characterized by the collective, coherent emission of light from an excited ensemble of emitters, generates photonic signals on timescales faster than the natural lifetime of an individual atom. The rapid exchange of coherence between atomic emitters and photonic fields in the superradiant regime enables a fast, broadband quantum memory. We demonstrate this superradiance memory mechanism in an ensemble of cold rubidium atoms and verify that this protocol is suitable for pulses on timescales shorter than the atoms' natural lifetime. Our simulations show that the superradiance memory protocol yields the highest bandwidth storage among protocols in the same system. These high-bandwidth quantum memories provide unique opportunities for fast processing of optical and microwave photonic signals, with applications in large-scale quantum communication and quantum computing technologies.

Photonic emission from a collection of identical excited atoms under superradiant conditions is very different from that of a single-atom in free-space [1]. Comprehensive studies of superradiance [2–4] have led to the observation of various quantum optics phenomena like quantum beats [5], collective Lamb shifts [6, 7], and novel cavity QED [8]. Initial studies focused on dense ensembles with sizes smaller than the excitation wavelength, whereas recent ones show that superradiance is also observable in large and dilute systems [9, 10]. Large atomic ensembles, providing a high degree of experimental control, are especially amenable to superradiant effects: when an ensemble with moderate optical density \((d > 1)\) is coherently excited by a temporally short (broadband) and weak probe, a subsequent radiation-burst is emitted along the forward direction of the incident pulse [11–15]. The characteristic decay time of this emission is inversely proportional to the medium's optical depth and, for \(d \gg 1\), can be much shorter than the spontaneous-emission lifetime of a single atom [5, 9, 10, 16–20]. This optical-depth dependence of decay time is the hallmark of superradiant emission, different from the free-induction-decay (FID) of atomic dipoles among inhomogeneously broadened emitters [21, 22].

Optical quantum memories [23–25] can take advantage of this rapid superradiant emission for broadband operation [21, 26, 27]. Here, we demonstrate a spinwave memory based on superradiance [4, 28] in a cloud of laser-cooled \(^{87}\)Rb atoms featuring a homogeneously-broadened optical transition. The superradiance (SR) mediated memory requires an initial excitation of the polarization coherence, followed by a second (fast) storage step, realizing a distinct memory mechanism that is inherently broadband and, in our experiments, is demonstrated without contribution from physical processes associated with other memory protocols. We find that the SR memory offers the most relaxed optical depth requirement for broadband signals in systems where the absorption linewidths are much narrower than the signal bandwidth to be stored. The inherently fast storage capability of this memory paves the way for bandwidth compatibility with conventional single-photon sources and high-speed quantum networks [29, 30].

To understand the SR memory, consider an \(N\)-atom ensemble whose energy levels form a \(\Lambda\)-configuration, where two ground-state spin levels \(|g\rangle\) and \(|s\rangle\) are optically coupled to an excited level \(|e\rangle\). We assume that all atoms initially populate \(|g\rangle\), and that both \(|e\rangle \leftrightarrow \{ |g\rangle, |s\rangle\} \) transitions are homogeneously-broadened, with Lorentzian lineshapes of characteristic width \(\Gamma\) and optical decoherence rates \(\gamma = \Gamma/2\). A weak probe field (the optical signal to be stored) is resonant with the \(|g\rangle \leftrightarrow |e\rangle\) transition, and a strong control field (to initiate probe storage or recall) with Rabi frequency \(\Omega_c(t)\) drives the \(|e\rangle \leftrightarrow |s\rangle\) transition. Resonant interactions between the optical fields and ensemble are described by the coupled Maxwell-Bloch equations in terms of coherences in the photonic (\(\hat{E}\)), polarization (\(\hat{P}\)), and spinwave (\(\hat{S}\)) modes [27, 31–34]. For the input probe, we consider an exponentially rising temporal profile of the form \(I_P(t) = I_0 e^{(t-\tau)/\Gamma} u(\tau - t)\) [\(\tau\) is the abrupt switch-off time]: characteristic duration \(T_P \ll 1/\gamma\) (Fig. 1); and bandwidth \(B = 0.157/T_P\) [34], such that memory operation is broadband with \(2\pi B \gg \Gamma\).

The SR memory protocol (Fig. 1) proceeds via three stages: absorption, writing, and retrieval, which follow the general principles of “fast” lambda-type storage and readout [27, 31], and can also be extended to the “ladder”-type atomic configurations [35, 36]. During absorption, the short incident probe builds up polarization coherence \(|P(z,t)|^2\) (Fig. 1c) across the entire ensemble [34] on a timescale much shorter than the excited-state decoherence, which can therefore be neglected \((e^{-\gamma T_P} \approx 1)\). Atomic polarization is maximized at the conclusion of the input pulse, leading to a subsequent superradiant reemission (Fig. 1a). Regardless of the line broadening mechanism (homogeneous or inhomogeneous), such reemission occurs naturally whenever a broadband pulse is absorbed by a spectral feature with linewidth narrower than the bandwidth of the incident
pulse [21]. This absorption regime differs from that of photon-echo processes, where the probe bandwidths must be smaller than or comparable to the width of the inhomogeneously broadened emitters [37].

In the writing stage, storage is achieved by converting the built-up polarization into a collective spin excitation via a control field (Fig. 1a) before superradiant emission proceeds. For efficient $P \rightarrow S$ mapping, (i) the control duration ($T_C$) must be much shorter than the superradiant decay time ($T_C \ll T_{SR}$) ensuring minimal leakage, and (ii) the control pulse-area must be $\pi$, ensuring maximum transfer [21, 27, 34, 38]. Like all spin-wave memories, the initial coherence remains stored up to a time limited by the spin-wave decoherence. Finally, retrieval is implemented by re-applying the $\pi$-pulse to map the coherence from $S \rightarrow \hat{P} \rightarrow \hat{E}$, recovering the probe into the output photonic mode (Fig. 1b, d). Bidirectional emission, as observed in some superradiance experiments [39–41], is not a factor in this memory system. The phase pattern of excited dipoles is imposed by the interference between probe and control fields, and phase matching conditions ensure that the retrieved probe is emitted into the desired mode.

To experimentally demonstrate the SR memory, we use an ensemble of $N_A \approx 2 \times 10^8$ laser-cooled $^{87}$Rb atoms. With $1/\gamma^2$ Gaussian diameters of 2.5 mm, 3.6 mm, and 4 mm, the cloud has volume density $7 \times 10^9$ cm$^{-3}$ and peak optical depth $d = 9$ (along the longest direction). We form the Λ-configuration on the “D2” line, with levels $|g\rangle = |5S_1/2, F = 1\rangle$, $|s\rangle = |5S_1/2, F = 2\rangle$, and $|e\rangle = |5P_{3/2}, F' = 2\rangle$. The probe and control fields, resonant with the $|g\rangle \leftrightarrow |e\rangle$ and $|s\rangle \leftrightarrow |e\rangle$ transitions, are derived from two independent, phase-locked lasers, and temporally shaped via electro-optic (EOM) and acousto-optic (AOM) modulators, respectively (Fig. 2a). Depending on mean photon number per probe pulse ($\bar{n}_{in}$), the control field is oriented either at $\theta = 5^\circ$ (high $\bar{n}_{in} \sim 10^3$) or $\theta = 50^\circ$ (single-photon-level probe with $\bar{n}_{in} < 1$, where the large angle permits spatial filtering of scattered control noise [42]) relative to the probe axis. The output probe is detected along the forward direction using a single-photon detector (SPD) and a time-to-digital counter (TDC). An experimental sequence begins with ensemble preparation, followed by 1000 measurement-and-detection events performed over 1 ms.

To begin, we characterize the superradiant behaviour of our cold Rb ensemble by sending short ($\bar{T}_p = 20$ ns $< 1/\gamma = 54$ ns) probe pulses with increasing-exponential profiles and sharp switch-off (Fig. 2b). The probe is almost completely absorbed up to the switch-off time, and subsequently reemitted along the forward direction, as superradiance with characteristic time $T_{SR} = 8$ ns. The observed superradiant decay time is shorter than both the input duration and the spontaneous emission lifetime [5, 9, 10, 43–45], by factors of 2.5 and 3.5, respectively. A linewidth ($\Gamma$) measurement of the probe transition confirms no inhomogeneous broadening [46, 47], which together with the probe bandwidth $2\pi B > \Gamma$, shows the conditions for superradiance are satisfied (Fig. 2b, inset).

Next, we verify the superradiant nature of reemission by measuring its decay time as a function of optical depth (Fig. 2c). We control the optical depth between $d = 1.5$ to 9 by varying the atomic cloud’s time-of-flight spatial expansion before measurement. For each $d$, we extract the $1/e$ decay time ($T_{SR}$), shown in Fig. 2c inset, with times from $T_{SR} = (8.0 \pm 0.1)$ ns to $T_{SR} = (18.8 \pm 0.8)$ ns, confirming emission’s superradiant behaviour: the enhancement of the decay rate with respect to optical depth is the characteristic signature of superradiance [5, 9, 10, 43–45]. This inset also shows SR emission efficiency: the higher the $d$, the greater the energy contained within the emitted signal and the shorter the emission timescale, suggesting that high memory efficiency requires both large optical depth (leading to an efficient polarization build-up) and fast writing.

We demonstrate the distinct operation of the SR spin-wave memory by storing an exponentially rising 10-ns probe (containing $\bar{n}_{in} \approx 1.5 \times 10^3$ photons), and re-
Figure 2. Demonstration of SR memory in cold $^{87}$Rb atoms. (a) Schematic of the setup. FC: fiber coupler, AOM: acousto-opto-modulator, EOM: electro-optic-modulator, SPD: single-photon-detector, TDC: time-to-digital converter, BD: beam dump. (b) Superradiant emission following an exponential probe. Inset: probe spectral profile (solid red) with bandwidth $B = 1.23 \Gamma/2\pi$ for linewidth $\Gamma = 2\pi \times 6$ MHz (shaded blue). (c) Decay of the emission intensity (normalized to its own maximum) for optical depth 9 (red circles), 5 (blue squares), and 3.5 (green diamonds), with solid lines fit to the data. Black dashed line shows spontaneous-emission decay (independent of $d$). Inset: superradiant decay time $T_{SR}$ and emission efficiency versus optical depth. Solid black line is fit of measured $T_{SR}$ values using Eq. (1), which matches the probe duration for optimal storage via the SR protocol. (d) Storage and forward retrieval of a $T_p = 10$ ns, $B = 12.7$ MHz probe containing $\tilde{n}_0 > 1$. Inset: SR emission without control. (e) Variation of memory efficiency with storage time [34]. (f) Storage and recall at the single-photon level with $\tilde{n}_0 = 0.18$. Inset: memory-retrieved signal intensity relative to noise measured in four configurations (other colors) [34].

Achieving this efficiency requires optimization of both probe and control fields. An optimized probe is characterized by an exponentially rising temporal shape with a duration that matches the superradiant decay of our system, which makes it broadband, since $B \times 1/T_p \approx 1/T_{SR} > \Gamma$. In our demonstration, both the probe shape and duration [$T_p = 10$ ns $\approx T_{SR} = 8$ ns] adequately fulfill these optimal conditions [34], but the efficiency is limited by non-optimized control: even at our maximum power, the shortest $\pi$-pulse we can generate exceeds $T_{SR}$, violating the condition $T_C \ll T_{SR}$. As a result, a significant fraction of the polarization coherence $\hat{P}$ is lost via superradiant emission before it can be mapped to the spin-wave $\hat{S}$. This loss could be eliminated by using a nanosecond-long control, which would require two orders of magnitude more intensity for the $\pi$-pulse. Still, in the absence of such intensity, the fast-writing process is evident in these experiments. First, the temporal separation between probe and write-control ensures the storage mechanisms from other protocols are suppressed [34]; second, when the system is prepared at a lower optical depth $(d = 6)$ and consequently longer $T_{SR}$ (11 ns), the maximum power acts over this longer $T_{SR}$ and transfers a greater proportion of $\hat{P}$ to $\hat{S}$, increasing the memory efficiency (from 3% to 5%) [34]. This counterintuitive increase in efficiency for lower $d$ under non-optimal conditions is characteristic of the SR mechanism, and would not be found in alternate memory mechanisms. Additionally, we find that under similar conditions, the efficiency for longer probe pulses (lower bandwidths) decreases, indicating that SR memory operates best for shorter pulses, up to where $T_p$ is comparable to $T_{SR}$ [34].

Next, we investigate the compatibility of SR memory with quantum signals by operating with probe pulses at the single-photon level. In particular, we measure noise from extraneously added photons during the memory operation and use this to predict fidelity for quantum state storage. We measure an unconditional noise probability...
of \((2.1 \pm 0.2) \times 10^{-4}\) [34], which corresponds to a stored-and-recalled signal-to-noise ratio (SNR) of \(12.4 \pm 1.2\) for a 10-ns probe with \(n_{\text{in}} = 0.18\) (Fig. 2f). If quantum states were encoded in these pulses, this SNR indicates a maximum quantum storage fidelity of \(F = 1 - 1/\text{SNR} = 0.92 \pm 0.09\). Experimentally, memory efficiency remains constant across different values of \(n_{\text{in}}\), yielding a linear increase in SNR [34]. Importantly, SNR is limited only by noise from scattered control beams and not by any physical process linked to memory operation, suggesting that the SR protocol will reliably store quantum signals, such as those encoded in space, polarization, or time bins.

With these proof-of-concept results at hand, we consider future implementations by examining optimality conditions for the SR protocol. Independent of any memory approach, the “optimality” criterion dictates that the maximum achievable memory efficiency \(\eta_{\text{opt}}\) is universal and depends only on the medium’s optical depth [27, 31], and can be satisfied by implementing protocol-specific optimizations of both the probe and the control, under the backward-recall configuration.

Probe optimization in the SR protocol relies on both maximizing the polarization build-up and ensuring its proper spatial-distribution \(P(z)\) in the absorption stage, which leads to an optimal spin-wave mode in the subsequent writing and readout stages. This is achieved by an input temporal profile that matches the time reversed replica of the superradiantly emitted pulse, both in shape and duration [22, 26, 48–50]. For a Lorentzian spectral feature (as in our cold Rb system, Fig. 2b, inset), the optimal shape of this probe is an exponentially-rising envelope with its duration given by [21, 38]

\[
T_P^{\text{opt}} \approx \frac{1}{\Gamma} \left( \frac{1}{1 + d/4} \right),
\]

which corresponds to the system’s \(T_{\text{SR}}\) emission time, as verified numerically and experimentally (solid fit in Fig. 2c, inset). For optical depths \(d > 1\), the pulse time is shorter than spontaneous emission lifetime, and thus the probe bandwidth exceeds \(\Gamma\), rendering this protocol inherently broadband.

Control optimization requires write-and-readout \(\pi\)-pulses with durations \(T_C\) much shorter than \(T_{\text{SR}}\). For square control pulses, numerical simulations verify that \(T_C^{\text{opt}} \approx T_{\text{SR}}/10 \approx T_P^{\text{opt}}/10\) is sufficient for optimal efficiency, in turn requiring \(\Omega_C \gg \gamma(1 + d)\) for fast writing and retrieval [27, 31]. For the Gaussian controls used in our experiments (with \(d = 9\)), this translates to a peak Rabi frequency \(\Omega_C^{\text{opt}} = 130\gamma\), well above our experimental \(\Omega_C^{\text{exp}} = 4.8\gamma\), accounting for the memory inefficiency.

Under optimized probe-and-control conditions, we investigate the optical-depth dependence of probe-duration (bandwidth) in terms of the adiabaticity parameter \((T_P^{\text{opt}}d\gamma)\). Using Eq. (1) and for \(d \gg 1\), \(T_P^{\text{opt}}d\gamma \leq 2\), which represents the “fast” operating regime opposite to that of adiabatic memories, which instead are characterized by \(T_P^{\text{opt}}d\gamma \gg 1\) [27, 31, 33]. Clearly, SR memory falls into the class of non-adiabatic memories [37, 51], which are inherently suitable for optimal storage of broadband signals \(2\pi B > \Gamma\). We note, however, that in homogeneously-broadened media, the bandwidth of the SR control fields is ultimately limited by the spacing between ground levels. For example, storage of a one-nanosecond probe would require control fields with GHz spectral width, comparable to the ground-state hyperfine splitting of alkali atoms. Such broad control fields may off-resonantly excite atoms in \(|g\rangle\), creating spurious spin-waves that can add photonic noise at the memory output via four-wave mixing (FWM) [25, 52–55].

Finally, we make a performance comparison (in terms of efficiency, optical depth, bandwidth, and control power) between SR and other memory approaches, including Autler-Townes splitting (ATS) [42, 56–58] and electromagnetically induced transparency (EIT) [59–61], which are examples of non-adiabatic and adiabatic protocols, respectively. We base this comparison on typ-
ical adiabaticity parameters, corresponding to $T_P d\gamma = 2$ (SR); $T_P d\gamma = 14$ (ATS); and $T_P d\gamma = 60$ (EIT), for a broadband, exponential probe with $T_P \ll 1/\gamma$. For a given bandwidth (probe duration), the optical depth required for optimal SR efficiency is 7 times lower than in an ATS memory and 30 times lower than in an EIT memory (Fig. 3a). Equivalently, for a given optical depth, the bandwidth that can be optimally stored using the SR protocol is higher than the corresponding bandwidth in the ATS and EIT protocols by the same factors, showing that SR protocol is the fastest among all protocols that are suitable for homogeneously-broadened transitions (Fig. 3b). Furthermore, Fig. 3c shows scaling of the optimal peak control intensity ($\propto \Omega_C^2$) as a function of probe bandwidth $B$, with $\Omega_{ATS} = 2\pi(2B)$; $\Omega_{EIT} = 2\pi(6.6B)$; and $\Omega_{SR} = 2\pi(31.2B)$, implying orders of magnitude larger control-power is required for the SR protocol. In sum, SR memory operates optimally at substantially lower optical depths, but its demand on control-power is significantly larger. A protocol’s resource dependence (on $d$ and $\Omega_C$) plays a crucial role in determining the impact of fundamental noise processes on the storage fidelity. In particular, photonic four-wave-mixing (FWM) noise bears an exponential and quadratic scaling to the optical depth and control power, respectively [54, 62, 63], and as such, the FWM strength associated with an optimal SR memory can be a factor of 3-7 times lower than that for an adiabatic memory [34, 42].

In conclusion, we experimentally demonstrated a broadband spin-wave memory based on the superradiance effect in a cold Rb gas. This memory approach offers the shortest pulse storage in systems with transition linewidths narrower than the signal bandwidth. Besides conventional platforms like atomic gases and solid state systems, a high-performance SR memory could be implemented using Bose-Einstein condensates in the originally conceived superradiant regime with large density and dimensions comparable to the excitation wavelength [1], and other interesting effects may emerge by considering subwavelength structured arrays when applied to memory [64, 65]. Beyond broadband quantum memories, our results may find use in the realizations of fast and efficient heralded single photon sources [38], atom-based optical processing [66], and superradiance-mediated dipole blockade effects [67, 68].

This work was supported by the University of Alberta; the Natural Sciences and Engineering Research Council, Canada (Grants No. RGPIN-04523-16, No. RGPIN-2014-06618, and No. CREATE-495446-17); the Alberta Quantum Major Innovation Fund; Alberta Innovates; and the Canada Research Chairs (CRC) Program. As researchers at the University of Alberta, we acknowledge that we are located on Treaty 6 territory, and that we respect the histories, languages, and cultures of First Nations, Métis, Inuit, and all First Peoples of Canada, whose presence continues to enrich our vibrant comm-

*m Corresponding authors: lindsay.leblanc@ualberta.ca, rastogi1@ualberta.ca*

[1] R. H. Dicke, Phys. Rev. **93**, 99 (1954).
[2] A. V. Andreev, V. I. Emel’yanov, and Y. A. Il’inskii, Sov. Phys. - Uspekhii **23**, 493 (1980).
[3] M. Gross and S. Haroche, Phys. Rep. **93**, 301 (1982).
[4] K. Cong, Q. Zhang, Y. Wang, G. T. Noe, A. Belyanin, and J. Kono, J. Opt. Soc. Am. B **33**, C80 (2016).
[5] H. S. Han, A. Lee, K. Sinha, F. K. Fatemi, and S. L. Rolston, Phys. Rev. Lett. **127**, 073604 (2021).
[6] M. O. Scully, Phys. Rev. Lett. **102**, 143601 (2009).
[7] A. A. Svidzinsky and M. O. Scully, Opt. Commun. **282**, 2894 (2009).
[8] M. O. Scully and A. A. Svidzinsky, Science **325**, 1510 (2009).
[9] M. O. Aratújo, I. Krešić, R. Kaiser, and W. Guerin, Phys. Rev. Lett. **117**, 73002 (2016).
[10] S. J. Roof, K. J. Kemp, M. D. Havey, and I. M. Sokolov, Phys. Rev. Lett. **117**, 73003 (2016).
[11] C. C. Kwong, T. Yang, M. S. Pramod, K. Pandey, D. Delande, R. Pierrat, and D. Wilkowski, Phys. Rev. Lett. **113**, 233601 (2014).
[12] C. C. Kwong, T. Yang, D. Delande, R. Pierrat, and D. Wilkowski, Phys. Rev. Lett. **115**, 223601 (2015).
[13] R. Pierrat, M. Chalony, D. Delande, and D. Wilkowski, Front. Opt. **2012/Laser Sci. XXVII 84**, 1LTu1I.3 (2012).
[14] N. Dudovich, D. Oron, and Y. Silberberg, Phys. Rev. Lett. **88**, 4 (2002).
[15] L. S. Costanzo, A. S. Coelho, D. Pellegrino, M. S. Mendes, L. Acioli, K. N. Cassemiro, D. Felinto, A. Zavatta, and M. Bellini, Phys. Rev. Lett. **116**, 023602 (2016).
[16] M. O. Scully, E. S. Fry, C. H. Ooi, and K. Wódkiewicz, Phys. Rev. Lett. **96**, 10501 (2006).
[17] A. A. Svidzinsky, J. T. Chang, and M. O. Scully, Phys. Rev. Lett. **100**, 160504 (2008).
[18] T. Bienaimé, M. Petruzzo, D. Bigerni, N. Piovella, and R. Kaiser, J. Mod. Opt. **58**, 1942 (2011).
[19] R. T. Sutherland and F. Robicheaux, Phys. Rev. A **93**, 23407 (2016).
[20] S. L. Bromley, B. Zhu, M. Bishop, X. Zhang, T. Bothwell, J. Schachenmayer, T. L. Nicholson, R. Kaiser, S. F. Yelin, M. D. Lukin, A. M. Rey, and J. Ye, Nat. Commun. **7**, 1 (2016).
[21] V. C. Vivoli, N. Sangouard, M. Afzelius, and N. Gisin, New J. Phys. **15**, 95012 (2013).
[22] A. Walther, A. Amari, S. Kröll, and A. Kalachev, Physical Review A - Atomic, Molecular, and Optical Physics **80**, 1 (2009).
[23] A. I. Lvovsky, B. C. Sanders, and W. Tittel, Nat. Photonics **3**, 706 (2009).
[24] K. Hammerer, J. Sherson, B. Julsgaard, J. I. Cirac, and E. S. Polzik, Quantum Inf. with Contin. Var. Atoms Light **82**, 513 (2007).
[25] K. Heshami, D. G. England, P. C. Humphreys, P. J. Bystard, V. M. Acosta, J. Nunn, and B. J. Sussman, J. Mod. Opt. **63**, 2005 (2016).
[26] A. Kalachev, Phys. Rev. A **76**, 043812 (2007).
The term “superradiance” may refer both to the case of externally imposed dipoles (as in this work) and the case of what is also commonly referred to as “superfluorescence,” in which dipoles self-align after a spontaneous emission event. The emission properties in both cases are the same.
Supplemental Material for “Superradiance-mediated photonic storage for broadband quantum memory”

SI. Evolution of coherences in superradiance protocol

The operation of superradiance protocol (Fig. 1a-d) is based on resonant interaction between an optically dense atomic ensemble and two electromagnetic fields, the “probe” and “control”. In the main text, we discuss its implementation in a Λ-type atomic system in terms of the exchange of coherences between the photonic ($|E(z,t)|^2$) and the atomic ($|P(z,t)|^2$, $|S(z,t)|^2$) modes, using the Maxwell-Bloch equations

$$\partial_z \hat{E}(z,t) = i \frac{g \sqrt{N}}{c} \hat{P}(z,t), \quad \text{(S1)}$$
$$\partial_t \hat{P}(z,t) = -\gamma \hat{P}(z,t) + ig \sqrt{N} \hat{E}(z,t) + i \frac{1}{2} \Omega_c(t) \hat{S}(z,t), \quad \text{(S2)}$$
$$\partial_t \hat{S}(z,t) = i \frac{1}{2} \Omega_c^*(t) \hat{P}(z,t), \quad \text{(S3)}$$

where we consider the 1D-propagation of optical fields along the medium from $z = \{0, L\}$. The photonic coherence $\hat{E}(z,t)$ describes the spatial and temporal variation of the probe field, while $\hat{P}(z,t)$ is the macroscopic polarization coherence of the $|g\rangle \leftrightarrow |e\rangle$ transition and $\hat{S}(z,t)$ is the spin-wave atomic coherence of the $|e\rangle \leftrightarrow |s\rangle$ transition. The atom-probe coupling strength, in terms of optical depth, is $g \sqrt{N} = \sqrt{cd \gamma / 2L}$, and we assume negligible relaxation between the levels $|g\rangle$ and $|s\rangle$.

Following the mathematical framework of [21, 38], we obtain analytic expressions for $P(z,t)$ and $S(z,t)$ during the absorption, writing, and retrieval stages of the protocol. Further, since Maxwell-Bloch equations are applicable to both weak classical fields (inducing macroscopic atomic polarization) and single-photon fields (inducing a single excitation in the ensemble), in the following analyses, we drop the operator notation from the corresponding variables, treating them as fields and macroscopic observables.

A. Absorption stage: generation of atomic polarization

To understand how the atomic coherence is initially established, we examine the temporal and spatial profiles of the macroscopic polarization $P(z,t)$ as it is generated in the medium during the absorption of input probe. In the absence of a control field ($\Omega_c = 0$), equations (S1)-(S3) reduce to

$$\partial_z E(z,t) = \frac{ig \sqrt{N}}{c} P(z,t), \quad \text{(S4)}$$
$$\partial_t P(z,t) = -\gamma P(z,t) + ig \sqrt{N} E(z,t), \quad \text{(S5)}$$

with initial conditions given by the input photonic mode’s shape: $E(z=0,t) = E_{in}(t) = E_p e^{i/T} u(-t)$. Here, $E_{in}(t)$ is the electric field of the input probe, which has an amplitude $E_p$ and a $1/e$ duration of $T$ and $u(-t)$ is the Heaviside unit step function. As discussed in the main text, this exponentially-rising temporal shape optimises both the amplitude as well as the spatial distribution of the polarization build-up, since it corresponds to the time-reversed impulse-response of the cold atomic ensemble which is characterized by a Lorentzian lineshape in this study.

With the above definition of the input pulse and setting $t = 0$ as the time at which the input is suddenly shut off, we consider the scenario in which absorption and the associated polarization build-up occurs for times $t \leq 0^-$, and for which subsequent re-emission or storage can occur at times $t \geq 0^+$. We assume that the probe field maintains its temporal shape during propagation through the medium ($tT \gg L$) and that it is only changed by an attenuation of its amplitude. We also note that in the main text, $T_p$ refers to the $1/e$ duration of probe intensity, such that $T_p = T/2$.

Next, to study the temporal dynamics of the polarization, we convert Eqs. (S4)-(S5) from the spatial to the complex spatial-frequency domain by taking the Laplace transform $z \rightarrow u$ and denoting the transformed functions with “bars”: $\bar{X}(u,t) = \mathcal{L}[X(u,t)] = \int dz e^{-uz} X(z,t)$. In the transformed basis, the equivalents of (S4) and (S5) are

$$\bar{E}(u,t) = \frac{ig \sqrt{N}}{cu} \bar{P}(u,t) + \frac{1}{u} \bar{E}_{in}(t) \quad \text{(S6)}$$
$$\bar{P}(u,t) = -\gamma \bar{P}(u,t) + ig \sqrt{N} \bar{E}(u,t). \quad \text{(S7)}$$

Inserting (S6) into (S7), we obtain a first-order linear differential equation in $\bar{P}(u,t)$

$$\bar{P}(u,t) + AP(u,t) = \frac{ig \sqrt{N}}{u} \bar{E}_{in}(t), \quad \text{(S8)}$$

where

$$A \equiv \gamma + \frac{g^2 N}{cu} = \gamma \left(1 + \frac{d}{2Lu} \right) \quad \text{(S9)}$$

and $d = g^2 NL / 2 \gamma c$ is the on-resonant optical depth. Solving (S8) analytically, we find

$$\bar{P}(u,t) = \bar{P}(u,0) e^{-At} + \frac{ig \sqrt{NT}}{u(1+AT)} \bar{E}_{in}(t).$$
Since polarization is induced only by the probe field, and no atoms are excited initially, \( \hat{P}(u, 0) = 0 \) and
\[
\hat{P}(u, t) = \frac{ig\sqrt{NT}}{u(1 + AT)} \mathcal{E}_{\text{in}}(t). \tag{S10}
\]
This expression describes generation of the polarization that is in phase with the input photonic field for a given spatial position \( z \rightarrow u \) (e.g., see Fig. 1c for the polarization profile at \( z = 0 \)). While deriving (S10), we have assumed that the probe switch-off at time \( t = 0 \) is instantaneous (i.e., the finite fall-time is neglected) such that the peak value of \( \hat{P} \) occurs at \( \hat{P}(u, t = 0) \), when the probe amplitude is maximized.

Next, we consider the spatial variation of the polarization as the probe propagates through the atomic medium: we expect that the polarization, because it is instantaneous (i.e., the finite fall-time is neglected) such that the peak value of \( \hat{P} \) occurs at \( \hat{P}(u, t = 0) \), the time when the write control is turned on
\[
\hat{P}(z, \omega) = \frac{1}{\omega + \gamma} \left[ ig\sqrt{N} \hat{E}(z, \omega) + P(z, t = -\infty) \right].
\tag{S12}
\]
Since the medium does not contain any initial polarization, \( P(z, -\infty) = 0 \) and equation (S12) can be substituted into (S11) to obtain:
\[
\partial_z \tilde{E}(z, \omega) = \frac{g^2 N}{c(\omega + \gamma)} \tilde{E}(z, \omega),
\tag{S13}
\]
yielding
\[
\tilde{E}(z, \omega) = e^{-g^2 Nz/(\omega + \gamma)} \tilde{E}(z = 0, \omega)
\tag{S14}
\]
Equation (S13) represents the exponential decay in the probe-field amplitude along the propagation direction. Inserting it into (S12) gives the spatially decaying polarization profile
\[
\tilde{P}(z, \omega) = \frac{ig\sqrt{N}}{\omega + \gamma} e^{-g^2 Nz/(\omega + \gamma)} \tilde{E}(z = 0, \omega)
\tag{S14}
\]
Thus, at a fixed position \( z \) within the medium, the polarization amplitude grows in time, following the envelope of the probe field (S10), whereas at any fixed instant in time, the polarization amplitude exhibits an exponential decay along the medium length (S14).

### B. Writing stage: Polarization-to-spin-wave conversion

During absorption, the input-probe field imprints its phase pattern onto the atomic ensemble in the form of a polarization build-up over the entire probe duration. By applying a "write" control field immediately after the probe is absorbed (and when the photonic coherence has been completely converted to the polarization coherence), this atomic polarization is transferred into a long-lived spin-wave coherence of the atomic ensemble (Fig. 1a,c). In our formalism, absorption is maximized at \( t = 0 \) with the value
\[
\hat{P}(u, 0) = \frac{ig\sqrt{NT}}{u(1 + AT)} \mathcal{E}_{\text{p}}, \tag{S15}
\]
which serves as the source term upon which the write control acts. The transformation between the polarization and spin coherences is governed by the Maxwell-Bloch equations (S1)-(S3), written again here for convenience:
\[
\partial_z E(z, t) = \frac{ig\sqrt{N}}{c} P(z, t),
\tag{S16}
\]
\[
\partial_t P(z, t) = -\gamma P(z, t) + ig\sqrt{N} E(z, t) + \frac{i}{2} \Omega_C(t) S(z, t), \tag{S17}
\]
\[
\partial_t S(z, t) = \frac{i}{2} \Omega_C(t) P(z, t) \tag{S18}
\]
where we have assumed \( \Omega_C \) to be real and ignored any spatial depletion of the control intensity (i.e. \( \Omega_C(z, t) = \Omega_C(0, t) \)). We consider a square control pulse
\[
\Omega_C(t) = \Omega_c, \quad 0 \leq t \leq T_c. \tag{S19}
\]
As the \( P \rightarrow S \) mapping occurs over a time-interval given by the control duration, \( T_c \) must be much shorter than the superradiant emission time \( T_{\text{SR}} \), or else a significant portion of the built-up polarization is lost via superradiant emission. Taking the Laplace transform of (S16)-(S18) from \( z \rightarrow u \):
\[
\tilde{E}(u, t) = \frac{ig\sqrt{N}}{cu} \tilde{P}(u, t) + \frac{1}{u} \hat{E}(z = 0, t)
\tag{S20}
\]
\[
\partial_t \tilde{P}(u, t) = -\gamma \tilde{P}(u, t) + ig\sqrt{N} \tilde{E}(u, t) + \frac{i}{2} \Omega_C(t) \tilde{S}(u, t) \tag{S21}
\]
\[
\partial_t \tilde{S}(u, t) = \frac{i}{2} \Omega_C \tilde{P}(u, t), \tag{S22}
\]
Since no photonic field \( E \) is present in the medium at the time when the write control is turned on \(( t = 0^+ \) ), \( E(z = 0, t) = 0 \). Next, substituting (S20) and (S22) into the time-derivative of (S21), we obtain a second-order linear differential equation in \( \hat{P}(u, t) \)
\[
\hat{P}(u, t) + A\hat{P}(u, t) + B \hat{P}(u, t) = 0, \tag{S23} 
\]
where
\[
A = \frac{\gamma}{u(1 + AT)} + \frac{ig\sqrt{N}}{cu}, \quad B = \frac{ig\sqrt{N}}{cu}. 
\]
where $A = \gamma + g^2 N/\epsilon_0 u = \gamma(1 + d/2L \epsilon_0 u)$ and $B = \Omega_c^2/4$. The roots of the characteristic equation in (S23), can be real or complex depending on the sign of the discriminant $D = A^2 - 4B$ which is determined by the relative strengths of $A$ and $B$ coefficients. For the superradiance protocol the condition for fast storage and retrieval requires a strong driving control such that
\[ \Omega_c \gg \gamma (1 + d), \]
implying $\Omega_c \gg \gamma (1 + d/2L \epsilon_0 u)$ and thus, $4B \gg A^2$. Roots of the characteristic equation are thus complex $(-A/2 \pm i\Omega_c/2)$ giving a solution of the form
\[ \bar{P}(u, t) = e^{-At/2} \left[ C(u) \cos \left( \frac{\Omega_c t}{2} \right) + D(u) \sin \left( \frac{\Omega_c t}{2} \right) \right] \]
for $t \geq 0$.

Applying the initial condition at $t = 0$ (S15), the coefficient $C(u) = \bar{P}(u, t = 0)$. Also, the fact that $\bar{P}(u, t = T_C)$ must ideally be 0 by the end of the control pulse suggests $D(u) = 0$. The expression for $\bar{P}(u, t)$ during the writing stage is thus,
\[ \bar{P}(u, 0 \leq t \leq T_c) = e^{-At/2} \cos \left( \frac{\Omega_c t}{2} \right) \bar{P}(u, t = 0). \] (S24)

As can be seen, for $\Omega_c T_C = \pi$ (i.e., a $\pi$-control-pulse with duration $T_C$), no polarization remains in the medium by the end of control duration (i.e., $\bar{P}(u, t = T_C) = 0$) and is completely converted into the spin-wave $\bar{S}(u, t = T_C) = \frac{\Omega_c}{2} \int_0^{T_C} \bar{P}(u, t') dt'$:
\[ \bar{S}(u, T_C) = \frac{i\Omega_c}{A^2 + \Omega_c^2} \bar{P}(u, 0) [\Omega_c e^{-AT_C/2} + A] \] (S25)

Now, since we are in the regime of $\Omega_c \gg A$: $\pi/T_C \gg A$ or $AT_C/2 \ll 1$, giving $e^{-AT_C/2} \approx 1$. Equation (S25) thus simplifies to:
\[ \bar{S}(u, T_C) = i\bar{P}(u, t = 0), \] (S26)
indicating that a lossless conversion of the polarization coherence into the spin-wave coherence can be achieved in the fast storage regime using a $\pi$-control-pulse.

C. Converting the stored spin-wave back into polarization during retrieval

In previous subsection, we saw that a fast $\pi$-pulse can quickly act on the atomic polarization to convert it into the spin-wave with a unity-transfer-efficiency, thus effective storage of the input photonic coherence. The probe field remains stored as the spin-wave up to a time, referred to as the memory lifetime ($T_{Mem}$), governed by the spin-wave decoherence processes. Re-applying the $\pi$ pulse after a desired storage interval ($T_S < T_{Mem}$), initiates the reversed mapping of $\bar{S}$ into $\bar{P}$ which then couples to the output photonic field (Fig. 1b,d). Assuming the absence of any decoherence mechanism that would otherwise degrade the spin-wave amplitude, we have the initial condition as $\bar{S}(u, \tau) = \bar{S}(u, T_C)$ where $\bar{S}(u, T_C)$ is the spin-wave generated by the end of the writing stage and $\tau = T_C + T_S$ with all times measured relative to the the end of absorption process ($t = 0$). The control field for readout is defined as
\[ \Omega_c(t) = \Omega_c, \quad \tau \leq t \leq \tau + T_c, \] (S27)
so that the spin-to-polarization transfer occurs within the short duration of $T_c$. Taking the time derivative of (S22) and then inserting (S20) and (S21), we get a second order differential equation in $\bar{S}(u, t)$
\[ \ddot{\bar{S}}(u, t) + A\dot{\bar{S}}(u, t) + B\bar{S}(u, t) = 0, \] (S28)
where we used the fact that no photonic field is initially present at the time of readout; the $A$ and $B$ coefficients are defined as before. The condition for fast retrieval yields $4B \gg A^2$, giving a solution of the form
\[ \bar{S}(u, \tau \leq t \leq \tau + T_c) = e^{-At/2} \cos \left( \frac{\Omega_c t}{2} \right) \bar{S}(u, t = T_C). \] (S29)

Again for a control pulse such that $\Omega_c T_C = \pi$ during the interval $t = T_C$, the spin-wave in (S29) goes to 0 and is completely converted into the polarization coherence using $\bar{P}(u, t) = \left(2/\Omega_c\right) \partial_t \bar{S}(u, t)$:
\[ \bar{P}(u, t) = \frac{ie^{-At/2}}{\Omega_c} \left[ \Omega_c \sin \left( \frac{\Omega_c t}{2} \right) + A \cos \left( \frac{\Omega_c t}{2} \right) \right] \times \bar{S}(u, t = T_C), \]
which, for a $\pi$-pulse (and $e^{-AT_C/2} \approx 1$) simplifies to
\[ \bar{P}(u, \tau + T_c) = i\bar{S}(u, t = T_C), \] (S30)
indicating a lossless conversion of spin-wave into polarization at the time of readout.

SII. Spectral bandwidth of input probe

We derive the Fourier-limited relationship between the rise time $T_P$ and the bandwidth at full-width-half-maximum $B$ of an excitation probe that has an exponentially rising temporal intensity profile of the form:
\[ I(t) = e^{t/T_F} u(-t) + e^{-t/T_F} u(t), \] (S31)
where $T_P$ and $T_F$ are, respectively the 1/e rise and fall times of the probe intensity, such that ideally $T_F \ll T_P$. The FWHM duration of such a pulse is given by
\[ T_{\text{FWHM}} = (T_P + T_F) \ln(2). \] We determine the spectral width of the probe in terms of the 3 dB frequency range \( B = \omega_{\text{dB}}/2\pi \) using the Fourier transform of the probe electric field \( E(t) = e^{i2\omega t}u(-t) + e^{-i2\omega t}u(t) \), given by

\[
E(\omega) = \frac{1}{2\pi} \left[ \int_{-\infty}^{0} e^{i\frac{\omega}{T_P}t} + \int_{0}^{+\infty} e^{-i\frac{\omega}{T_P}t} \right] dt
= \frac{T_P + T_F}{\pi} \left[ \frac{1}{(1 + 4T_P\omega^2 - 2i(T_P - T_F)\omega)} \right]
\]

The intensity spectral profile, computed using \( I(\omega) = E^*(\omega)E(\omega) \) is

\[
I(\omega) = \left( \frac{T_P + T_F}{\pi} \right)^2 \frac{1}{(1 + 4T_P\omega^2)^2 + 4(T_P - T_F)^2\omega^2}
\]

which can then be numerically plotted to estimate \( \omega_{\text{dB}} \). As an example, for numerical simulations in Figs. 1 and 3 of the main text, we chose \( T_F = \frac{1}{10} T_P \). Substituting in (S33), we obtain \( \omega_{\text{dB}}T_{\text{FWHM}} = 0.75 \) or \( B = 0.157/T_P \) with \( T_{\text{FWHM}} = 0.76T_P \). In our experiments, the fall time \( T_F^{\text{sp}} \) is limited by the electronic response of the arbitrary waveform generator driving the probe-shaping EOM, and was measured to be \( \approx 6 \) ns. Therefore, probe pulses with \( T_P = 20 \) ns and \( T_F^{\text{sp}} = \frac{1}{2} T_P \) (Fig. 2 b,c in main text) correspond to \( B = 0.147/T_P = 7.4 \) MHz which is equivalent to \( 1.23\Gamma/2\pi \) for Rb atoms. Similarly, probe pulses with \( T_P = 10 \) ns (Fig. 2d,e,f) correspond to \( B = 0.13/T_P = 12.7 \) MHz = \( 2.12\Gamma/2\pi \).

### A. Cold Atom Preparation

An ensemble of laser-cooled \(^{87}\text{Rb}\) atoms serves as the medium for either probing the superradiant emission (Fig. 2b,c) or for storing short optical pulses via the superradiance (SR) protocol (Fig. 2d,e,f). The atoms are loaded from a background atomic vapour into a 3-dimensional retro-reflected magneto-optical-trap (MOT) operating on the D2 line. The MOT repump laser is frequency locked to the \(|F = 1\rangle \rightarrow |F' = 2\rangle \) transition using saturated absorption spectroscopy, at a detuning of -80 MHz. The repump beam is upshifted by +80 MHz using an acousto-optic-modulator (AOM) to make it resonant with the repump transition. The cooling (trapping) laser is frequency stabilized with respect to the repump laser using the offset beatnote locking technique such that it is detuned by \( \approx -180 \) MHz from the \(|F = 2\rangle \rightarrow |F' = 3\rangle \) transition. The cooling beam is passed through a double-pass AOM stage which upshifts the frequency by +160 MHz, making it red-detuned from the cooling transition effectively by \( \approx 3.5\Gamma \). The 4 seconds of MOT cooling is followed by 5 ms of polarization-gradient optical molasses, at the end of which about \( N_A = (2.0 \pm 0.5) \times 10^8 \) Rb atoms are prepared in \(|F = 2\rangle \). Time-of-flight (TOF) absorption-imaging measurements yield an elliptical-shaped cloud with \( 1/e^2 \) Gaussian diameters of 2.5 mm and 3.6 mm along the radial direction, and a cloud temperature of \( \approx 35 \) \( \mu \)K. A 3.5 ms optical-pumping stage follows molasses, wherein the repump beam is turned off but the cooling beam is kept on to off-resonantly transfer atoms from \(|F = 2\rangle \) into the \(|F = 1\rangle \) state, which serves as the initial population level \(|g\rangle \) for the input probe field. Absorption-and-reemission measurements of short, resonant, and weak optical probes yield an ensemble optical depth of \( \delta = 9 \pm 1 \). For datasets like the one in Fig. 2c, a variable time-of-flight spatial expansion stage was introduced after the optical pumping stage, to controllably vary this optical depth.

### B. Generation of probe and control fields

In our experiments, the probe and the control fields resonant with \(|F = 1\rangle \rightarrow |F' = 2\rangle \) and \(|F = 2\rangle \rightarrow |F' = 2\rangle \) transitions are derived from the MOT repump and cooling lasers respectively. To generate the probe pulse at the desired frequency, the CW beam from the repump laser is both gated and frequency-shifted by -80 MHz using an AOM, such that the gated pulse is -160 MHz away from the probe transition. This gated pulse (200 ns wide) is subsequently passed through a double-pass AOM which imparts a frequency shift of +160 MHz, thereby producing a resonant probe pulse. This resonant probe is then passed through an electro-optic intensity modulator (EOM) which shapes the gated pulse into an exponentially-rising temporal envelope with the desired rise time \( (T_P) \). The CW cooling-laser-beam is passed through an AOM that generates Gaussian write and read temporal control pulses \( (T_C = 20 \) ns at FWHM), and imparts an additional frequency shift of -85 MHz making them resonant with the control transition. The RF oscillators driving the gated-probe and control AOMs are amplitude controlled using one two-channel arbitrary waveform generator (AWG), while the probe-intensity EOM is driven directly from a separate AWG. The center frequency of probe pulse is set by adjusting the carrier frequency of the probe-double-pass AOM, while that of control pulse is set by changing the beatnote offset frequency between the repump and cooling lasers. AWGs are also used to adjust the relative timing positions between the probe and control fields.

After setting the amplitude (power), frequency, and timing of probe and control pulses, both beams are launched separately into the MOT chamber with the control field oriented at an angle (\( \theta \)) of either 5\(^\circ\) or 50\(^\circ\) with respect to the probe-axis. The probe is focused at the center of the atomic cloud with a \( 1/e^2 \) Gaussian dia-
mater of 70 μm, while the control is collimated at 2.5 mm diameter. The initially linear polarization of each field is adjusted using λ/4-waveplates to maximise the atom-light interaction strength.

C. Measurement and detection

The useful signals at the output of the probe arm include the input probe photons (in the absence of both atoms and control), superradiantly emitted probe photons (with atoms but without control), and stored-and-retrieved photons (with both atoms and control). We detect these signals using time-resolved photon counting measurements, enabled by a single-photon-detector (SPD) and a time-to-digital converter (TDC), in the form of histograms of the detected photon counts with respect to their arrival times. The SPD-TDC combination provided both a better temporal resolution as well as a lower electronic noise-floor for our detection signals compared to a photodiode + oscilloscope pair.

In a single experimental cycle, after preparing the cold ensemble, we perform measurement-and-detection during a 1 ms window by sending a sequence of 1000 probe pulses (also 1000 write-read control pulses when implementing memory) with a 1 μs-separation between consecutive pulses. The measurement-and-detection stage is followed by a 200 ms ‘wait-time’ before the next experimental cycle begins. By repeating several such prepare-and-measure cycles (n_cyc), signal integration over a large number of trials is obtained (n_cyc × 10^3).

The probe pulses injected into the atomic cloud either contain large input mean photon numbers per pulse (i.e., \( n_{\text{in}} \sim 10^3 \)) or are at the single-photon-level where \( n_{\text{in}} \) can be much less than one. In both cases, we estimate \( n_{\text{in}} \) by measuring the number of probe photons detected (\( N_D \)) with respect to the number of probe pulses generated over the same time-interval (\( N_P = n_{\text{cyc}} \times 10^3 \)) and in the absence of atoms, such that \( n_{\text{in}} = (1/\eta_T) \times (N_D/N_P) \), where \( \eta_T \approx 0.25 \) is the overall detection efficiency including the end-to-end transmission efficiency of the optical setup, fiber-coupling efficiency, and SPD efficiency. For measurements with \( n_{\text{in}} \gg 1 \), a series of neutral-density-filters providing up to 38 dB optical-attenuation are placed along the output-path of the probe arm. This ensures that in spite of large \( n_{\text{in}} \), the effective number of photons reaching the SPD remain below 1 so that the detector does not go into the saturation mode.

Superradiance memory using \( n_{\text{in}} \gg 1 \) probe is implemented at the \( \theta = 5^\circ \) setting whereas the single-photon-level memory measurements (\( n_{\text{in}} < 1 \)) are taken at the \( \theta = 50^\circ \) setting. Limited by the power in our setup, the shortest write-and-read control pulses we could generate with a pulse area close to \( \pi \) correspond to a Gaussian temporal shape with an FWHM duration of \( T_C = 20 \) ns and a peak power of 8 mW.

D. Probe absorption-linewidth and peak optical-depth

To characterize the \( |5S_{1/2}, F = 1 \rangle \leftrightarrow |5P_{3/2}, F' = 2 \rangle \) probe transition we perform spectral measurements of the absorption linewidth (\( \Gamma_{\text{exp}} \)) and the resonant optical depth (\( d_{\text{res}} \)) for our Rb ensemble. Comparison of \( \Gamma_{\text{exp}} \) with respect to the natural linewidth would indicate whether there are additional broadening mechanism (eg: Doppler broadening) in our system. We send a sequence of 1000 square probe pulses, of duration 400 ns, at a fixed detuning (\( \delta \)) and detect the probe transmission counts both without \( |I_{\text{in}}(\delta)\rangle \) and with \( |I_{\text{out}}(\delta)\rangle \) atoms. By performing such sequence over many cycles, reliable statistics can be obtained. The relatively long duration of the probe (\( T_P = 3.7/\gamma \)) ensures no reemission occurs after absorption. These measurements are performed for values of \( \delta \) between -19 MHz to + 19 MHz with the optical depth in each case computed as \( d(\delta) = -\ln(I_{\text{out}}/I_{\text{in}}) \).

By fitting a Lorentzian to the measured \( d \) values

\[
\frac{d(\delta)}{1 + 4(\delta/\Gamma)^2} = \frac{d_{\text{res}}}{1 + 4(\delta/\Gamma)^2},
\]

we extract the absorption linewidth \( \Gamma_{\text{exp}} = 2\pi \times (6.2 \pm 0.1 \text{ MHz}) \), and the peak optical depth \( d_{\text{res}} = 7 \pm 0.2 \) (Fig. S1a). We note that the maximum measured OD from our data is limited to \( \approx 5.6 \) and hence less than what the fit in (S34) gives. We attribute this to the limited dynamic range of the SPD and thus exclude the data points near \( \delta = 0 \) while performing the fit. The estimated value of \( \Gamma_{\text{exp}} \) agrees very well with the natural linewidth, confirming negligible effect of any inhomogeneous broadening and that the condition \( 2\pi B > \Gamma_{\text{exp}} \) is satisfied for observing superradiant emission, as in Fig. 2b.

Next, using the value of \( \Gamma_{\text{exp}} \), we simulate our measurements for the absorption-and-reemission of short probe pulses, shown in Fig. S1b,c for \( T_P = 30 \) ns (0.55/\( \gamma \)) and 20 ns (0.37/\( \gamma \)). A good agreement between experiments and simulations is obtained for a peak optical depth \( d_{\text{res}} = 9 \pm 1 \) which is 25% higher than that obtained from spectral measurements. Since, the spectral
measurements are carried out with pulses that are an order-of-magnitude larger, depletion of the atomic cloud during each pulse sequence can account for the observed discrepancy. Hence, the peak optical depth of our ensemble is closer to what we expect from the time-domain absorption measurements.

E. Operational SR regime in our experiment

The SR protocol benefits from the generation of a collective polarization coherence at timescales shorter than the $T_2 = 1/\gamma$ relaxation rate of the system and is inherently suited for storing pulses with duration $T_P < 1/\gamma$. In this superradiant regime, an optimized probe duration is close to the superradiant emission time of the system i.e., $T_P^{\text{opt}} \approx T_{\text{SR}}$, where $T_{\text{SR}}$ is determined from the system parameters $(d, \gamma)$. Figure S2 shows numerically calculated efficiencies ($\eta$) relative to the optimal efficiency ($\eta^{\text{forward}}$) for storage followed by forward-recall of an exponentially rising input probe of varying duration (bandwidth) and an abrupt switch-off ($T_F = 0.1T_P$) under our experimental configuration of $d = 9$, $\gamma = 2\pi \times 3$ MHz, that yield $T_{\text{SR}} = 8$ ns. The simulations consider square-shaped $\pi$-control pulses for writing and readout with duration $T_C = 2T_{\text{SR}}$. The shaded region depicts that for a given optical-depth, only a small range of probe durations around $T_{\text{SR}}$ will yield optimal efficiency. The reduction in efficiency at lower values of $T_P$ is due to inadequate optical-depth resulting in reduced probe absorption whereas at higher values, the drop is due to dephasing in the atomic dipoles ($\propto e^{-\gamma T_P}$) during absorption.

Also shown are the experimentally measured efficiencies ($\eta^{\text{exp}}$) for probe pulses with $1/e$ rise times of 10 ns, 20 ns, and 30 ns, each having a finite fall time of $\approx 6$ ns, using our non-optimal and Gaussian-shaped control with $T_C = 20$ ns. Although, the $\eta^{\text{exp}}$ values are much smaller than the expected $\eta^{\text{forward}}$ due to the reasons discussed in the main text, they do follow the trend that the SR protocol readily favours the storage of shorter pulses and also confirm that our chosen parameters for the probe pulse to be stored lie in the good operational regime of the SR protocol at the optical depth that we have.

F. Optimizing the SR memory in our experiment

Our experimental SR memory efficiency is primarily limited by the non-optimal writing stage: at $d = 9$, the superradiant emission in our system drains the excited state within $T_{\text{SR}} = 8$ ns whereas the control-induced polarization-to-spin-wave transfer occurs over an interval six times longer, given by the full duration of our experimental write control $(2.25 \times T_C = 45$ ns). A comparison of probe’s reemission profiles both in the absence and presence of the control field (Fig. 2d) suggests that a large fraction of the polarization coherence, accumulated during absorption, is lost via superradiant emission even when the control is present (referred to as “leakage”). Faster writing requires shorter control duration and in turn more power to maintain the $\pi$-area. Limited by the control power, another way to alleviate this leakage is to extend the superradiant emission by lowering the system’s optical depth. Figure S3 illustrates the storage-and-retrieval of a 10 ns probe at optical depths of 9 and 6, where at the lower $d$, an increased superradiant emission time enables the control to transfer a greater fraction of the polarization coherence into spin-wave, thereby yielding a higher efficiency, even at the expense of increased probe transmission. Figure S2 shows similar behaviour is observed for other probe durations. However, this optim-
ization is limited by the trade-off between the transmission and leakage losses: lowering the $d$ further makes the probe transmission dominant and reduces the efficiency.

G. Estimation of memory lifetime

Figure 2e in the main text shows the measured superradiance memory efficiency as a function of the storage time i.e., the time interval between the write and read control pulses. The dataset was obtained for the small-angle ($\theta = 5^\circ$) probe-control separation, wherein short exponential probe pulses of 10 ns duration were stored and then retrieved at later time intervals ranging between 200 ns to 6.2 $\mu$s. We recorded a memory efficiency of 3% for the storage duration of 200 ns. The efficiencies (black circles) at other storage times are normalized with respect to the above value. The degradation in the memory efficiency follows an exponential decay (solid red line) yielding with a $1/e$ memory lifetime of $4.2 \pm 0.3 \mu$s. This memory lifetime is limited by the dephasing of the collective spinwave excitation, induced by the residual / ambient magnetic fields. For the single-photon-level measurements done at the wider probe-control separation (50$^\circ$), the spinwave dephasing induced by the random atomic motion dominates since it is an increasing function of $\theta$ and limits the memory lifetime in our setup to $1.2 \pm 0.2 \mu$s.

H. Single-photon-level measurements

To implement storage and recall of single-photon-level probe pulses ($\bar{n}_{\text{in}} < 1$), we spatially filter out the noise from the scattered control photons leaking into the spatial mode of the memory retrieved signal. This is done by increasing the angle of the control beam with respect to the probe from 5$^\circ$ to 50$^\circ$ which increases the isolation from 80 dB to 120 dB (four-orders-of-magnitude extinction), whereas the memory efficiency reduces to 1.3% from the typical value of 3% for 200 ns storage time.

We define $\rho_S$ as the probability of detecting a ‘signal’ count from the retrieved probe as $\rho_S = (1/\eta_T) \times (N_S/N_{\text{Mem}})$, where $N_S$ is the total number of detection counts for the recalled probe after $N_{\text{Mem}}$ memory trials, computed over a time window $\Delta t$ centered around the readout time. Analogously, we can define the probability of detecting a ‘noise’ photon at the readout as $\rho_R = (1/\eta_T) \times (N_R/N_S)$ where $N_R$ is the number of noise detections (within the signal readout window $\Delta t$) in the absence of the input probe (i.e., $\bar{n}_{\text{in}} = 0$) after $N_S$ trials. Both $\rho_S$ and $\rho_R$ are independently measured detection probabilities, with the memory efficiency ($\eta_{\text{Mem}}$) and the ratio of the stored-and-recalled signal to noise (SNR) given as $\eta_{\text{Mem}} = \rho_S/\bar{n}_{\text{in}}$ and SNR $= \rho_S/\rho_N$.

Figure. 2f in the main article shows the storage and recall after 200 ns of a $T_p = 10$ ns probe at $\bar{n}_{\text{in}} = 0.18$, after $N_{\text{Mem}} = 10^6$ attempts giving $N_S = 645 \pm 25$ counts within a $\Delta t = 50$ ns window yielding $\rho_S = (2.6 \pm 0.1) \times 10^{-3}$ and $\eta_{\text{Mem}} = (1.3 \pm 0.1)\%$. The inset of the readout window shows the noise contribution relative to the recalled probe (signal) under four configurations: (i) ambient light + scattering from optics and atoms (both probe and control are off) ; (ii) leak from the input probe (control off); (iii) control without atoms and probe; and (iv) control and atoms without probe. Specifically, counts from (iii) and (iv) would indicate noise from photons that are added during the memory operation, such as from scattered control (iii) or any fundamental process (iv). We determine the unconditional noise probabilities of $(6 \pm 1) \times 10^{-5}$, $(7 \pm 1) \times 10^{-5}$, $(1.3 \pm 0.1) \times 10^{-4}$, and $(2.1 \pm 0.2) \times 10^{-4}$ respectively. The fact that both (iii) and (iv) configurations yielded similar $\rho_R$ indicates that the scattered control remains the dominant noise source in our implementation and more importantly, there is no measurable contribution from any photonic noise linked to the operation of the superradiance memory. With the above noise contributions (all included in (iv)), we obtain an SNR = $12.4 \pm 1.2$. Figure S4 shows SNR as a function of $\bar{n}_{\text{in}}$ after 200-ns storage: while the memory efficiency stays roughly the same, the SNR shows a linear dependence.

SIV. Scaling of the four-wave-mixing photonic noise in SR, ATS, and EIT memories

Photonic noise generated at the memory output via four-wave-mixing process is detrimental towards the reliable implementation of broadband spinwave storage-and-retrieval in $\Lambda$-type atomic media [52–54, 62]. FWM noise is an important concern in memory configurations involving backward retrieval of the signal probe which on
the one hand yield optimal memory efficiencies but on the other hand allow for perfect phase-matching among the optical fields involved in the FWM process [42]. In such a scenario, the impact of a phase-matched FWM noise corrupting the memory operation is characterized using a noise-strength parameter ($S_{\text{FWM}}$) that depends on the optical-depth ($d$) and control Rabi frequency ($\Omega_C$) as [54, 62, 63]:

$$S_{\text{FWM}} \propto \Omega_C^4 \left[ \sinh \left( \frac{\zeta \gamma_{eg}}{2 \Delta_{gs}} \right) \right]^2,$$

(S35)

where $(\zeta, \gamma_{eg}, \Delta_{gs})$ are parameters specific to the storage platform: $\zeta$ is the ratio of the control-induced couplings from $|s\rangle \rightarrow |e\rangle$ (desired) and $|g\rangle \rightarrow |e\rangle$ (unwanted); $\gamma_{eg}$ is the optical decoherence of the excited state ; and $\Delta_{gs}$ is the hyperfine splitting between the $|g\rangle$ and $|s\rangle$ spin levels. Since the $(d, \Omega_C)$ values required to optimally realize a memory protocol are proportional to the memory bandwidth $B$ along with proportionality constants specific to the dynamics of that protocol, $S_{\text{FWM}}$ depends very strongly on the memory bandwidth and the resource scaling of the implementing protocol. Figure S5 shows a comparison of the FWM noise strength among the SR, ATS, and EIT protocols for optimal broadband storage in a homogeneously-broadened $^{87}$Rb ensemble where $\gamma_{eg} = \Gamma_{eg}/2 = 2\pi \times 3 \text{ MHz}$, $\Delta_{gs} = 2\pi \times 6.83 \text{ GHz}$, and $\zeta = 1.33$. We calculate $S_{\text{FWM}}$ using equation (S35) for bandwidths $10\Gamma_{eg}/2\pi \leq B \leq 60\Gamma_{eg}/2\pi$ corresponding to exponential probes of duration $0.4 \text{ ns} \leq T_P \leq 2.6 \text{ ns}$. For each protocol, the optimal values of the peak optical-depth and peak control Rabi frequency associated with a given bandwidth are determined from the non-adiabatic / adiabatic operation conditions of that protocol (Fig. 3, main text). By combining those $d$ vs. $B$ and $\Omega_C$ vs. $B$ scaling in equation (S35) for the range of bandwidths under consideration, we find that the probability of FWM noise affecting the SR memory is 3-4 orders of magnitude higher than the ATS memory (due to highest control power requirement) while it is still a factor of 3-7 times lower than the EIT memory (due to lowest optical-depth requirement) i.e., $S_{\text{ATS}}^{\text{SR}} < S_{\text{FWM}}^{\text{SR}} < S_{\text{FWM}}^{\text{EIT}}$.

SV. Discerning storage based on SR and ATS protocols in non-optimal situations

Storage in both SR and ATS protocols relies on the excitation of the atomic polarization coherence ($P(z,t)$), but with one subtle difference: in the ATS memory [56, 57] the writing stage happens over the entire duration of the probe, in which the control field mediates both probe absorption (which generates $P(z,t)$) as well as the conversion from $P \rightarrow S$ at the same time. In contrast, the SR protocol has absorption and writing in two separate stages: the write control is applied only after the $P$ has maximised and conversion into $S$ happens only during the time the write control is on, which can be much shorter than the probe duration for optimal performance (Fig. 1, main text). This fundamental distinction leads to very different requirements for optimal implementation of the two memories such that the associated storage processes can never mix. However, in non-optimal SR implementations with control duration longer than the system’s superradiant decay: if the absorption and writing are not well-separated such that there is a partial or complete overlap between the probe and write control fields, the storage maybe contributed by processes related to SR and ATS protocols both. Larger the probe-control overlap, greater is the ATS contribution to this ‘mixed’
storage. To identify implementations where storage is mainly dominated by SR protocol’s mechanism, especially in a non-optimal setting, we define the following dimensionless character factors:

1. \( \mathcal{P} \) is the ratio of the polarization build-up in the presence of write control to that in its absence. In either case, the polarization build-up is computed over the entire probe duration (\( \tau_p \)) and across the entire medium length (\( L \)) as
   \[
   \frac{1}{\tau_p} \int_0^{\tau_p} \int_0^L |P(z,t)|^2 dz \, dt,
   \]
   where \( \tau_p = 3(T_P + T_F) \) for an exponential probe. Implementations for which \( \mathcal{P} \to 1 \) correspond to pure SR storage with no or negligible probe-control overlap such that the write control does not affect the generated polarization. As the probe-control overlap increases, the value of \( \mathcal{P} \) decreases due to its simultaneous mapping into the spin-wave, affecting storage via ATS protocol’s mechanism.

2. \( \mathcal{S} \) is the ratio of the spin-wave generated during the probe-control overlap time (\( \Delta t \)) to the total spinwave generated during writing (\( \tau_c \)). The overlap spinwave can be computed using
   \[
   \frac{1}{\Delta t} \int_0^{\Delta t} \int_0^L |S(z,t)|^2 dz \, dt,
   \]
   where for a Gaussian control with FWHM duration as \( T_C \), \( \tau_c = 2.25 T_C \).

   Memory implementations featuring \( \mathcal{S} \to 0 \) represent pure SR storage in that no or negligible spinwave is generated while the probe field is being absorbed.

In figure S6, we plot \( \mathcal{P} \) and \( \mathcal{S} \) character values for several memory implementations in our experiment, each associated with a storage mechanism that is based on either pure SR, or mixed SR + ATS, or pure ATS. The only parameter that we vary is the arrival-time of the write control after the input probe is applied, which we define as the time delay between their peak intensities. Other parameters are kept fixed and include: exponential probe with \( T_P = 10 \) ns and \( B = 12.7 \) MHz; Gaussian control with \( T_C = 20 \) ns and \( A_C = 0.85\pi; \ d = 9; \ \gamma = 2\pi.3 \) MHz.

Larger delay times correspond to a small probe-control overlap such that the spin-wave storage is achieved only from the mechanism of the SR protocol (\( \mathcal{P} \to 1, \mathcal{S} \to 0 \)). By shifting the write pulse closer to probe i.e., decreasing the arrival-time and increasing the overlap, storage is affected by the mechanisms of both ATS and SR protocols as evident from the decrease (increase) in \( \mathcal{P} (\mathcal{S}) \). Finally, the ATS mechanism dominates when the write-control and probe are on almost at the same time and the overlap is close to maximum. Our results indicate that under non-optimized control setting, the transition from one memory to the other is smooth and can occur by a simple change of the control field’s position relative to the probe. We also measure the experimental memory efficiency \( \eta^{\text{exp}} \) in each storage regime. Importantly, the (unoptimized) memory efficiency in the mixed regime can be higher than that in the true SR regime: for example, with 25 ns separation from the probe, the non-optimized control fields in our experiment yield an efficiency of 3% in the true SR regime; shifting the write-control towards the probe by only a few nanoseconds can easily double this efficiency, however, the storage mechanism does not remain pure-SR based. By continually shifting the write-control in the pursuit of higher memory efficiency, one can enter the regime of pure yet unoptimized ‘ATS’-based storage yielding efficiency upto 10%. We emphasize that in all cases, the efficiency remains less than what can be achieved using the optimal implementation of either protocol (having completely different probe-and-control requirements). This further stresses the need to operate a memory protocol in its native and optimal regime where the advantages associated with that protocol’s mechanism are best met.