Efficient in-situ generation of photon-memory entanglement in a nonlinear cavity

Hoi-Kwan Lau,1,2 Hong Qiao,1 Aashish A. Clerk,1 and Tian Zhong1

1Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA
2Department of Physics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

(Dated: October 17, 2024)

Parametrically driving an optical cavity that simultaneously couples to an atomic ensemble quantum memory enables in-situ generation of multimode photon-memory entanglement. A high-rate bi-party photon-memory entanglement can be generated even after discarding one entangled optical mode. This protocol can be realized with existing technologies based on photonic resonators integrated with a rare-earth-ion doped quantum memory. The proposed scheme shows significant advantages in entanglement generation rates compared with prevailing quantum memory protocols and experiments, with theoretical Ebit rates of tens of MHz without fine-tuned operating conditions. Such a photon-memory entanglement source offers a versatile resource for quantum networking and interconnect applications.

Introduction.- Entanglement between itinerant optical photons and matter degrees of freedom is a quintessential ingredient for remote quantum interconnects, with imminent applications in quantum networks [1–3], quantum transduction [4, 5], quantum-enhanced telescopes [6], and distributed quantum computation [7–9]. Using Raman transitions, the well-known Duan-Lukin-Cirac-Zoller (DLCZ) protocol generates one entangled bit (Ebit) between two remote atomic ensembles, with built-in entanglement purification and noise resilience [10]. This protocol was extended to multimode operation by combining an EPR photon pair source with a broadband quantum memory realized via inhomogeneously broadened atomic ensembles in solids, leading to orders of magnitude speedup of the Ebit generation rate and reduced experimental complexities [11]. While these protocols are long established, experimental demonstrations of multimode entanglement between two remote atomic memories were only achieved very recently [12, 13]. Long-standing challenges hampering this scheme include photon loss and bandwidth mismatch at the interface between itinerant optical modes and the stationary memory degree of freedom. For instance, in the experiments employing separate photon sources and memories, the entanglement generation rates are limited even over a short fiber distance [12, 13]. One strategy to mitigate this loss at the light-matter interface is to use an atom-containing, high-finesse optical cavity with a fine-tuned impedance matching (i.e. unit coupling coopertivity) [14, 15]. However, such cavity-enhanced memory interfaces [16–19] would require a matching narrow-band photon source, which not only adds significant system complexities, but also constraints the overall entanglement generation bandwidth and throughput.

In this Letter we propose a scheme for efficient multimode entanglement generation between an optical photon and an atomic memory within a nonlinear cavity. By parametrically driving the cavity, we produce non-degenerate photon pairs (i.e. signal and idler photons), with one of the photons (i.e. signal) having spectral overlap with an inhomogeneously broadened atomic ensemble. By constructing an equivalent circuit representation of the steady-state input-output relations, we uncover the basic structure of the resulting multi-partite entanglement generated between the atomic and two photonic degrees of freedom. We find that our scheme is capable of generating photon-memory entanglement over a broad spectrum of modes, some of which can involve highly excited states. We quantify the performance of our system by the entanglement rate, which is exact beyond the weak excitation regime and upper-bounds the performance of any practical entanglement distribution protocol. Our scheme demonstrates robust and efficient entanglement creation without requiring cavity impedance matching, parameter fine tuning or strong coupling. Even with an imperfect intra-cavity photon-memory transfer and discarding the signal photons, we see no significant degradation of bipartite entanglement between the memory and the idler photons.

We further describe experimental realizations using existing technologies based on photonic microcavities coupled to ensembles of rare-earth ions doped in a crystalline substrate. The proposed system is capable of generating up to 50 MHz Ebits over a range of accessible experimental parameters. The high Ebit generation rate is result of a large number of frequency modes within the inhomogeneous linewidth of the memory, and the accurate accounting of entanglement in the strongly driven regime. Our proposal is distinct from the DLCZ protocol with its quasi-CW, multimode operation, while it eliminates the loss and mismatch challenges in pair-source based repeater protocols [11]. Furthermore, the intrinsic tri-partite nature of the entanglement - between two photonic and one memory degrees of freedom - opens future opportunities of generating and distributing multi-partite entanglement over a network for quantum secret sharing [20, 21], multi-partite teleportation [22], and distributed sensing [23].

Entanglement generation.- Our entanglement generation process can be modelled as two cavity modes, sig-
FIG. 1: (a) Generalized model of three-mode system: Entangled signal ($B_S$) and idler ($B_I$) are generated by parametric driving and signal is coupled to atomic memory ($\tilde{C}_M$) (b) Spectral slicing of inhomogeneously broadened atomic ensemble memory. (c) Equivalent circuit consists of one two-mode-squeezing gate sandwiched by two beam-splitters. The amount of output excitation is characterized by the TMS strength $\cosh^2{r}$. The second beam-splitter (BS2), which determines the amount of entanglement or photon going into memory, has a beam-splitting angle $\theta_2$. The first beam splitter (BS1), which has a beam-splitting angle $\theta_1$, is irrelevant for our situation of vacuum/ground state input.

nal ($\tilde{B}_S$) and idler ($\tilde{B}_I$), which are spectrally distinct, and which couple independently to a common waveguide with a coupling rate $\kappa$. The waveguide frequency mode, that is detuned by $\omega$ from the signal (idler) mode frequency, is represented by $\tilde{b}_{S\omega}$ ($\tilde{b}_{I\omega}$). The signal mode also resonantly couples to an atomic ensemble memory with a spectral distribution $\rho(\omega)$ (i.e. the signal cavity frequency coincides with the center of the atomic ensemble inhomogeneous distribution). In the rotating frame of the cavity modes the Hamiltonian of the system is

$$
\hat{H} = \xi (\hat{B}^\dagger_S \hat{B}^\dagger_I + \hat{B}_S \hat{B}_I) + \int \omega \hat{c}^\dagger_{M,\omega} \hat{c}_{M,\omega} d\omega + G \int \sqrt{\rho(\omega)} (\hat{B}^\dagger_S \hat{c}_{M,\omega} + \hat{c}^\dagger_{M,\omega} \hat{B}_S) d\omega ,
$$

where $G$ is the collective atom-photon coupling. In our protocol the device is driven by a constant intensity pump at the frequency $\nu_p = (\nu_S + \nu_I)/2$ from $t = t_o$ to $t = t_f$, this will induce a parametric drive with strength $\xi$ in the cavities. Generally, atoms with the frequencies within the ‘strip’ $\omega$ to $\omega + d\omega$ are not resolvable if $d\omega$ is much smaller than the resolvable bandwidth $\delta_r \equiv 1/(t_f - t_o)$. We consider the ensemble has a wide inhomogeneous broadening $\Gamma \gg \delta_r$, so the collective atomic excitation within each strip can be represented by the bosonic operator $\hat{c}^\dagger_{M,\omega}$, and the collective atom-photon coupling of each strip is scaled by $\sqrt{\rho(\omega)}$ [24]. We note that this representation is compatible with holeburning memory protocols such as atomic frequency combs [25], in which the comb spacing $\Delta$ cannot be resolved during the pump time, i.e. $\delta_r \gg \Delta$, and the ensemble can be effectively modelled with a modified atomic density.

For sufficiently long pump time, i.e. $\tau_{\text{pump}} \equiv t_f - t_o \gg 1/\kappa$, the output photonic modes and atoms can be characterized by the steady-state scattering relation

$$
\begin{align*}
\tilde{B}^\dagger_{I\omega} &= T_{II} \tilde{B}^\dagger_{I\omega} + T_{IS} \tilde{B}^\dagger_{S\omega} + T_{IM} \tilde{C}^\dagger_{M\omega} \\
\tilde{B}^\dagger_{S\omega} &= T_{SI} \tilde{B}^\dagger_{I\omega} + T_{SS} \tilde{B}^\dagger_{S\omega} + T_{SM} \tilde{C}^\dagger_{M\omega} \\
\tilde{C}^\dagger_{M\omega} &= T_{MI} \tilde{B}^\dagger_{I\omega} + T_{MS} \tilde{B}^\dagger_{S\omega} + T_{MM} \tilde{C}^\dagger_{M\omega} .
\end{align*}
$$

We note that the scattering amplitudes $T$ depend on detuning $\omega$ [24], which is omitted in the above equations only for brevity. The input/output operators of the photonic waveguide follow the standard definitions, $B_u^{\text{in/out}}(t) \equiv \sqrt{1/2\pi} \int \tilde{b}_{u\omega}(t_0/\tau_f) e^{-i\omega(t-t_0/\tau_f)} d\omega$, for $u = S, I$. Analogously, we can define the effective input and output operators of atoms as $\tilde{C}_M^{\text{in/out}}(t) \equiv \sqrt{1/2\pi} \int \tilde{c}_{M\omega}(t_0/\tau_f) e^{-i\omega(t-t_0/\tau_f)} d\omega$, which characterizes the transformation of the atomic state before and after interacting with the cavity.

This scattering relation shows that our system intrinsically generates tri-partite entanglement between each frequency mode of idler, signal, and memory. The equivalent circuit is depicted in Fig. 1(c), which involves two beam splitters (BS) with the same angle $\tan \theta_1 = \left| T_{MI}(\omega) / T_{SI}(\omega) \right| = \tan \theta_2 = \left| T_{M\omega}(\omega) / T_{SI}(\omega) \right|$, and a two-mode-squeezing (TMS) interaction with strength $\cosh^2{r} = |T_{II}(\omega)|^2$. We note that the mode transformation denoted by Eq. (2) applies to any input state, so our setup can also be applied to engineering logic gates between itinerant photons and quantum states stored in the memory [26]. For our current goal of in-situ generation of photon-memory entanglement, we assume both signal and idler inputs are vacuum and the memory is initialized in the ground state.

To quantify the amount of entanglement, we consider a Lorentzian spectrum, $\rho(\omega) = \frac{1}{2\pi} \frac{\Gamma}{\omega^2 + (\Gamma/2)^2}$, with an ensemble inhomogeneous broadening $\Gamma$. Based on the equivalent circuit in Fig. 1(c), one might guess that a complete transfer of the entangled signal photon to the memory necessarily requires a large effective beam splitting angle

$$
\tan^2\theta_2 = \left| T_{MI}(\omega) / T_{SI}(\omega) \right|^2 = \frac{C}{1 + 4\omega^2/\Gamma^2}.
$$

This would require in turn a large collective atomic ensemble-cavity cooperativity $C \equiv 4\Gamma^2/\kappa \gg 1$. We find however that this condition is neither necessary nor optimal for our memory. Heuristically, increasing $C$ enhances the effective damping of the cavity, which is harmful as
it suppresses photon generation due to the parametric drive.

Further, our analysis indicates that there is no special utility in achieving impedance matching (i.e. $C = 1$): we thus conclude that fine tuning to reach this matching is not required. To be more specific, we quantify the performance of our protocol by the entanglement rate $E_R$, which is the total amount of photon-memory entanglement generated in all frequency modes per unit time [27, 28]:

$$E_R \equiv \frac{\sum_\omega \mathcal{E}_F(\omega)}{\tau_{\text{pump}}} \approx \frac{1}{2\pi} \int \mathcal{E}_F(\omega) d\omega.$$  \hspace{1cm} (4)

We note that the entanglement between modes with different frequencies is negligible as long as $\tau_{\text{pump}} \gg 1/\kappa, 1/\Gamma$ [23]. We account for the entanglement of each frequency mode at $\omega$ by entanglement of formation $\mathcal{E}_F(\omega)$ because it has clear operational meaning and can be computed exactly for our output state that is Gaussian and balanced. Our choice of $E_R$ as entanglement quantifier is advantageous as it does more than account for the entanglement associated with single excitations: contributions from vacuum and highly excited states are also fully included. Furthermore, it represents the device’s maximum performance in entanglement generation for a fixed choice of system parameters, i.e. any post-selection or non-Gaussian process cannot improve the rate of entanglement generation.

Fig. 2(a) plots the memory-idler entanglement rate $E_R$ as a function of both the atomic ensemble linewidth and the ensemble-cavity coupling. For maximizing $E_R$, it is clear that impedance matching is neither necessary nor universally optimal. We also observe that around the regime of maximum performance, the value of $E_R$ is robust against order-unity changes in system parameters. This is ultimately because the memory captures the entangled photon directly from the signal cavity, a process that can occur efficiently without the fine tuning that would be required if instead photons had to be captured from a waveguide.

The above analysis considered memory-idler photon entanglement assuming that outgoing signal photons are discarded. One might worry that this discarding results in a significant loss of entanglement, especially since the effective beam splitting angle in Eq. (3) is not overwhelmingly large at the parameters of maximum performance in Fig. 2(a) (i.e. $C$ is of order unity). To investigate this concern, we compare in Fig. 2(b) the entanglement rate for two cases: (i) between memory and both photonic (signal and idler) outputs (computed using the entropy of the final memory state); (ii) between memory and idler output only, with the signal output discarded. Interestingly, the atom-photon entanglement rate is barely degraded when the drive is weak (i.e. $\xi \ll \kappa$), and shows only about a factor of two reduction when $\xi$ is comparable to $\kappa$ (instability occurs when $\xi \to \kappa/\sqrt{2}$ in our choice of parameters [24]). The intuition behind this effect is that the entanglement rate has significant contributions from off-resonant modes (i.e. $\omega \neq 0$). These modes are only weakly excited even when the resonant mode is driven close to instability. On the other hand, idler-memory entanglement would only be destroyed when a photon is detected in signal output. Since the signal output of off-resonant modes contain mainly vacuum, there is a high probability that the idler-memory entanglement is retained even if signal is monitored or discarded. Nevertheless, despite the weak excitation, the tri-partite entanglement between idler, signal and memory is genuine and can be harnessed as a resource for quantum information applications [20–23].

**Experimental scheme.**- The proposed in-situ entanglement generation can be realized experimentally with existing technologies. We design an integrated device that consists of a silicon micro-ring resonator on top of a $^{167}$Er$^{3+}$:Y$_2$SiO$_5$ crystalline substrate, as shown in Fig. 3. Entangled signal and idler photon pairs are generated by spontaneous four-wave-mixing (sFWM) in the silicon resonator with an estimated MHz generation rate at μW pump power [29, 30] [24]. The on-chip micron-scale resonator implementation is necessary for a sufficiently large (i.e. $\geq 1$ nm) free spectral range [24] so that the signal, idler and pump photons can be efficiently separated with dense wavelength-division-multiplexing (DWDM) filters. The signal cavity mode is resonant with the $^{167}$Er$^{3+}$ : YSO optical transition at 1539 nm, between the ground and excited-state $|m_s = 7/2\rangle_{e-g}$ hyperfine levels [31]. Assuming a doping concentration of 17 parts per million (ppm), the evanescent coupling between the transverse magnetic (TM) mode of the resonator and the Er ensembles leads to a collective coupling strength of

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig2.png}
\caption{(a) Normalized entanglement rate $E_F/\kappa$ as a function of the coupling $G$ and the atomic ensemble linewidth $\Gamma$ while keeping a weak parametric drive at $\xi = 0.1\kappa$. The cavity impedance matching condition is also plotted as in red curve. (b) Normalized entanglement rate $E_F/\kappa$ as a function of parametric drive $\xi/\kappa$. Here we assumed $2G=\Gamma=\kappa$, so $C = 1$. Orange (Blue) curve represents the entanglement between memory and idler output only (both signal and idler output).}
\end{figure}
The quantum memory protocol we adopt is based on AFC and the relevant \( ^{167}\text{Er}^{3+} \) energy levels. Right inset shows the simulation of transverse magnetic mode profile of Si waveguide on a \( ^{167}\text{Er}^{3+} : \text{YSO} \) substrate.

\[ G = 2\pi \times 0.173 \text{ GHz} \quad [24]. \]

The quantum memory protocol we adopt is based on atomic optical frequency combs (AFC), which have built-in multimode capability [25]. A long-lived AFC is first prepared by initializing and holeburning on the \( ^{167}\text{Er}^{3+} |m_s = 7/2\rangle \rightarrow |g\rangle \) transition prior to entanglement generation. Then a quasi-CW pump laser is turned on with a time duration \( \tau_{\text{pump}} \gg 1/\kappa \) to continuously generate entangled photons in idler and signal modes. If there is no memory, the output is a continuum of entangled photon with frequency centered at \( \nu_I \) and \( \nu_S \) respectively, which can also be viewed as a continuous stream of temporally (or time-energy) entangled photons with the duration of each defined by the cavity lifetime \( 1/\kappa \). In our scheme that involves additionally the memory, photons in the signal cavity mode are stored into the AFC with a storage time \( T_M = 1/\Delta \) where \( \Delta \) is the AFC comb spacing. We further require that the pump duration is shorter than the storage time, so that our ensemble description of the memory remains valid.

AFC memory storage efficiency in our scheme is determined by Eq. (3), in which the effective cooperativity should take into account the finesse of the AFC: \( F = \Delta/\gamma \), where \( \gamma \) is the spectral width of the comb tooth. For \( ^{167}\text{Er}:\text{YSO} \), a storage time of 1\( \mu \text{s} \) is achievable, with a AFC finesse \( F=3, \Delta=1 \text{ MHz} \), and \( \gamma \approx 2\pi \times 0.3 \text{ MHz} \) [32, 33]. Long-term storage is also feasible with spin-wave memory [34] using an additional hyperfine level [31]. While the signal photons are stored, the idler photons are propagating out of the cavity. The average idler photon number in each temporal mode is controlled by adjusting the pump strength \( \xi \). We stress that the entanglement generated in our scheme can be accounted for even when \( \xi \) is large.

**Entanglement rate.** - We first analyze the entanglement rate of our experimental setup. Remote entanglement generation with separate photon pair source and quantum memory has demonstrated heralded entanglement rate up to 15.6 kHz [12], which is mainly limited by a low photon pair generation rate, high loss and signal photon storage inefficiency. Our in-situ scheme can circumvent the limitations of interface loss and storage inefficiency through directly coupling to memory. Furthermore, the entanglement rate \( E_R \) takes into account the entanglement of highly excited modes, so the performance of our setup can be faithfully quantified even the cavity is strongly driven.

In Fig. 4(a) with a fixed coupling \( G^{\text{AFC}} = G/\sqrt{F} = 2\pi \times 0.1 \text{ GHz} \), we show the entanglement rate at varying parametric drive \( \xi \) and cavity decay rate \( \kappa \). The peak entanglement rate can reach as much as 85 MHz Ebits. For the example experimental parameters in Table. I, the entanglement rate at \( \xi = 2\pi \times 80 \text{ MHz} \) (white dot) is 79

**TABLE I: Example experimental parameters**

| \( \Gamma/2\pi \) | \( G^{\text{AFC}}/2\pi \) | \( \nu/2\pi \) | \( \xi/2\pi \) | \( F \) | \( T_M \) |
|---|---|---|---|---|---|
| 150 | 100 | 100 | 80 | 3 | 1 \( \mu \text{s} \) |

**FIG. 4:** (a) Continuous-variable (CV) entanglement rate \( E_R \) of generated photon-memory pair and (b) heralding efficiency \( \eta \) of memory excitation per detected idler photon for discrete-variable (DV) entanglement generation. The white dots indicate system parameters given in Table 1.
MHz. This high Ebit rate is achieved when the cavity is driven just below the instability threshold, that is the black region in Fig. 4. Although the resonant frequency mode will be highly excited near the threshold, the time-bin mode that is addressed by usual time-resolved detectors contains much less than one photon on average, due to the contribution of the off-resonant frequency modes which are weakly excited. This combination of in-situ generation and accounting the full structure of entanglement shows that our setup has the potential to achieve orders of magnitude higher rates than that of latest experiments in [12, 13].

**Heralding efficiency.** - The inherently continuous-variable entanglement in our scheme can also be a resource for discrete-variable encoding. One use case is heralded entanglement generation in a quantum repeater network, in which detection of an itinerant photon heralds the storage of a qubit in the memory. We here analyze the heralding efficiency as an important figure of merit in such context. Previously, Ref. [35] demonstrated heralded single-photon source with 83% heralding efficiency. However, this efficiency would be much lower if one needs to store the photon into a quantum memory. We here analyze the heralding efficiency as an important figure of merit in such context. Previously, Ref. [35] demonstrated heralded single-photon source with 83% heralding efficiency. However, this efficiency would be much lower if one needs to store the photon into a quantum memory (19% demonstrated in [12] and 14.3% AFC efficiency in a two-photon-detection heralded entanglement distribution experiment [13]). We thus define the heralding efficiency as rate of non-zero memory excitation to the output rate of non-zero idler photon,

\[
\eta = \frac{\int \left(1 - \frac{1}{1+|T_M(\omega)|^2} \right) d\omega}{\int \left(1 - \frac{1}{1+|T_M(\omega)|^2} \right) d\omega}. \tag{5}
\]

We plot the heralding efficiency as a function of κ and ξ in Fig. 4(b). With a \(G_{\text{AFC}} = 2\pi \times 0.1\) GHz, \(\Gamma = 2\pi \times 150\) MHz, \(\xi = 2\pi \times 10\) MHz, a heralding efficiency up to 99% can be achieved with a \(\kappa = 2\pi \times 3\) MHz. With more practical parameters in Table I, we obtain a heralding efficiency of 79%.

**Memory retrieval.** - The stored entanglement eventually needs to be retrieved. Here we outline, without loss of generality, two retrieval strategies. First, the entanglement in the memory can be retrieved by transducing to and readout through an auxiliary non-optical mode. For instance, coupling the atomic ensemble to a proximal superconducting resonator allows transduction of the memory excitation to microwave photons following the protocol in [36]. The entanglement retrieval efficiency in this case is determined by the transduction efficiency, which can theoretically reach unity [37]. The second strategy retrieves the entanglement through an optical mode - for convenience we consider through the same cavity mode in which entanglement was originally generated. A perfect retrieval is only achieved when all the memory spectral modes fulfils the impedance matching condition. While a standard impedance matched cavity only guarantees perfect retrieval for the memory frequency modes within a fraction \((\approx 1/3)\) of the cavity linewidth [15, 38], [38] proposed using dispersion compensation to enlarge this impedance matching bandwidth beyond the cavity \(\kappa\) towards the full ensemble inhomogeneous linewidth. In our scheme, we can realize a similar extended cavity impedance matching by dynamically switching the cavity to impedance matching point (e.g. via carrier injection to switch the Q) and rapidly sweeping the cavity resonance frequency across the ensemble inhomogeneous broadening. The integrated nanophotonic platform is amenable for such active cavity switching and fast chirping, which has been realized by incorporating high-speed field modulation (up to 10 GHz in LiNbO\(_3\) resonators [39]) or current injection [40] elements to the resonator. Finally, even without a perfect fidelity, the entanglement extracted from the memory, when optimized for a high throughput generation, is still a desirable resource for quantum network as imperfect entanglement would be distilled during the entanglement purification step [41–43] of quantum repeater protocols. Notably, using the same parameters in Table I and assuming no memory errors, we find that 81% of the generated entanglement can be retrieved [24], and this efficiency is further optimized when operating at the exact impedance matching condition at an expense of entanglement generation rate.

We acknowledge helpful discussions with Kurt Jacob. This work was funded by a National Science Foundation (NSF) Faculty Early Career Development Program (CAREER) Grant (No. 1944715), partially supported by the University of Chicago Materials Research Science and Engineering Center, which is funded by the National Science Foundation under award number DMR-2011854, and the NSF QLCI for Hybrid Quantum Architectures and Networks (NSF Grant No. 2016136). H.-K.L. acknowledges support from Canada Research Chairs (CRC-2020-00134).

[1] H. J. Kimble, *Nature* **453**, 1023 (2008).
[2] S. Wehner, D. Elkouss, and R. Hanson, *Science* **362**, eaam9288 (2018).
[3] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, *Phys. Rev. Lett.* **78**, 3221 (1997).
[4] N. J. Lambert, A. Rueda, F. Sedlmair, and H. G. L. Schwefel, *Advanced Quantum Technologies* **3**, 31900077 (2019).
[5] M. Mirhosseini, A. Sipahigil, M. Kalae, and O. Painter, *Nature* **588**, 599 (2020).
[6] D. Gottesman, T. Jennewein, and S. Croke, Physical review letters **109**, 070503 (2012).
[7] L. K. Grover, arXiv preprint quant-ph/9704012 (1997).
[8] L. Jiang, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, *Phys. Rev. A* **76**, 062323 (2007).
[9] C. Monroe, R. Rausendorf, A. Ruthven, K. R. Brown, P. Maunz, L.-M. Duan, and J. Kim, *Phys. Rev. A* **89**, 022317 (2014).
[10] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Nature 414, 413 (2001).
[11] C. Simon, H. de Riedmatten, M. Afzelius, N. Sangouard, H. Zbinden, and N. Gisin, Physical Review Letters 98, 190503 (2007).
[12] D. Lago-Rivera, S. Grandi, J. V. Rakonjac, A. Seri, and H. de Riedmatten, Nature 594, 37 (2021).
[13] X. Liu, J. Hu, Z.-F. Li, X. Li, P.-Y. Li, P.-J. Liang, Z.-Q. Zhou, C.-F. Li, and G.-C. Guo, Nature 594, 41 (2021).
[14] M. Afzelius and C. Simon, Physical Review A 82, 022310 (2010).
[15] S. A. Moiseev, S. N. Andrianov, and F. F. Gubaidullin, Physical Review A 82, 022311 (2010).
[16] M. Sabooni, Q. Li, S. Kröll, and L. Rippe, Phys. Rev. Lett. 110, 133604 (2013).
[17] P. Jobez, I. Usmani, N. Timoney, C. Laplane, N. Gisin, and M. Afzelius, New Journal of Physics 16, 083005 (2014).
[18] R. A. Akhmedzhanov, L. A. Gushchin, A. A. Kalachev, N. A. Nizov, V. A. Nizov, D. A. Sobgaya, and I. V. Zelesnkiy, Laser Physics Letters 13, 115203 (2016).
[19] M. M. Minnegaliev, K. I. Gerasimov, R. V. Urmancheev, A. M. Zheltikov, and S. A. Moiseev, Phys. Rev. B 103, 174110 (2021).
[20] T. Tyc and B. C. Sanders, Physical Review A 65, 042310 (2002).
[21] H. K. Lau and C. Weedbrook, Physical Review A 88, 042313 (2013).
[22] Y. Lian, C. Xie, and K. Peng, New Journal of Physics 9, 314 (2007).
[23] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Physical Review A 97, 032329 (2018).
[24] See the Supplemental Material for more details.
[25] M. Afzelius, C. Simon, H. De Riedmatten, and N. Gisin, Physical Review A 79, 052329 (2009).
[26] H.-K. Lau and A. A. Clerk, npj Quantum Information 5, 31 (2019).
[27] Z. Jia Deng, S. J. M. Habraken, and M. Marquardt, New Journal of Physics 18, 063022 (2016).
[28] C. Zhong, X. Han, H. X. Tang, and L. Jiang, Physical Review A 101, 032345 (2020).
[29] S. Azzini, D. Grassani, M. J. Strain, M. Sorel, L. G. Helt, J. E. Sipe, M. Liscidini, M. Galli, and D. Bajoni, Optics Express 20, 23100 (2012).
[30] W. C. Jiang, X. Lu, J. Zhang, O. Painter, and Q. Lin, Optics Express 23, 20884 (2015).
[31] M. Rančić, M. P. Hedges, R. L. Ahlefeldt, and M. J. Sellars, Nature Physics 14, 50 (2018).
[32] S. Yasui, M. Hiraishi, A. Ishizawa, H. Omi, R. Kaji, S. Adachi, and T. Tawara, Opt. Express 29, 27137 (2021).
[33] I. Craiciu, M. Lei, J. Rochman, J. M. Kindem, J. G. Bartholomew, E. Miyazono, T. Zhong, N. Sinclair, and A. Faraon, Phys. Rev. Applied 12, 024062 (2019).
[34] S. A. Moiseev and S. Kröll, Phys. Rev. Lett. 87, 173601 (2001).
[35] S. Ramelow, A. Mech, M. Giustina, S. Gröblacher, W. Wieczorek, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, A. Zeilinger, and R. Ursin, Optics Express 21, 6707 (2013).
[36] C. O’Brien, N. Lauk, S. Blum, G. Morigi, and M. Fleischhauer, Phys. Rev. Lett. 113, 063603 (2014).
[37] L. A. Williamson, Y.-H. Chen, and J. J. Longdell, Physical Review Letters 113, 203601 (2014).
[38] E. S. Moiseev, A. Tashchilina, S. A. Moiseev, and B. C. Sanders, New Journal of Physics 23, 063071 (2021).
[39] C. Wang, M. Zhang, B. Stern, M. Lipson, and M. Lončar, Opt. Express 26, 1547 (2018).
[40] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, Nature 435, 325 (2005).
[41] W. Dirr, H.-J. Briegel, J. I. Cirac, and P. Zoller, Phys. Rev. A 59, 169 (1999).
[42] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Physical Review Letters 77, 2818 (1996).
[43] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Physical Review Letters 76, 722 (1996).
[44] B. J. Dalton, S. M. Barnett, and B. M. Garraway, Physical Review A 64, 053813 (2001).
[45] D. Tamasceli, A. Smirne, S. F. Huelga, and M. B. Plenio, Physical Review Letters 120, 030402 (2018).
[46] W. S. Levine, Control system fundamentals, Electrical engineering handbook series (CRC Press, 2011).
[47] S. Tserkis and T. C. Ralph, Physical Review A 96, 062338 (2017).
[48] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Reviews of Modern Physics 84, 621.
[49] G. Adesso, Physical Review A 83, 024301 (2011), 1102.2355.
[50] D. L. McAnuslan, J. J. Longdell, and M. J. Sellars, Physical Review A 80, 062307 (2009).
Supplemental Material

Derivation of scattering relations

We consider a system consisting of two cavity modes, signal ($\hat{B}_S$) and idler ($\hat{B}_I$), of which their resonant frequencies are respectively $\nu_S$ and $\nu_I$. The cavity is driven by a pump with frequency $\nu_p = (\nu_S + \nu_I)/2$, which induces a parametric interaction on the modes due to cavity’s intrinsic non-linearity. The cavity couples to a waveguide. Since our operation time is much longer than the round-trip time of the optical cavity, i.e. $t_f - t_o \gg 1/|\nu_I - \nu_S|$, each cavity mode couples mainly to the waveguide frequency modes around its resonant frequency, and thus the waveguide can be modelled as two continuum of modes centered at $\nu_S$ and $\nu_I$. The signal mode is further coupled to an atomic ensemble memory that consists of $N$ atoms. In the rotating frame with respect to the cavity mode frequencies, the total Hamiltonian is given by

$$
\hat{H}_{\text{full}} = \xi(\hat{B}_S^\dagger \hat{B}_I^\dagger + \hat{B}_S \hat{B}_I) + \int \omega \hat{b}_{S \omega}^\dagger \hat{b}_{S \omega} d\omega + \int \omega \hat{b}_{I \omega}^\dagger \hat{b}_{I \omega} d\omega + \sum_n \omega_n \hat{\sigma}_+^{(n)} \hat{\sigma}_-^{(n)} + \sqrt{\kappa_S \over 2\pi} \int (\hat{B}_S^\dagger \hat{b}_{S \omega} + \hat{b}_{S \omega}^\dagger \hat{B}_S) d\omega + \sqrt{\kappa_I \over 2\pi} \int (\hat{B}_I^\dagger \hat{b}_{I \omega} + \hat{b}_{I \omega}^\dagger \hat{B}_I) d\omega + g \sum_n (\hat{B}_S^\dagger \hat{\sigma}_-^{(n)} + \hat{\sigma}_+^{(n)} \hat{B}_S) .
$$

(6)

where $\xi$ is the strength of the parametric coupling; $\hat{b}_{S \omega}$ and $\hat{b}_{I \omega}$ are the annihilation operators of the waveguide frequency modes that are $\omega$ detuned from the signal and idler cavity modes, respectively; $\omega_n \equiv \nu_n - \nu_S$ is the detuning between the energy level splitting $\nu_n$ of the $n^{th}$ atom and the signal mode frequency; $\hat{\sigma}_+^{(n)}$ and $\hat{\sigma}_-^{(n)}$ are respectively the lowering and raising operators of the $n$th atom; $\kappa_S$ and $\kappa_I$ are respectively the cavity coupling rates of the signal and idler cavity modes; the coupling strength between the atom and signal mode is assumed to homogeneous and is given by $g$. Because the cavity-waveguide coupling rate is assumed to be much weaker than the cavity mode frequency difference, the off-resonant waveguide modes are barely excited and so the integral domains of the detuning $\omega$ can be well approximated as from $-\infty$ to $\infty$. For simplicity, we have neglected all internal losses, but they can be straightforwardly incorporated and will be analyzed in future works.

We assume the spectrum of the atomic transition frequencies is sufficiently dense so that the spectral distribution can be approximated by a normalized continuous function $\rho(\omega)$, where $\int \rho(\omega) d\omega = 1$. During the long but finite operation time of our system, the resolvable frequency can be recognized as $\sim 1/(t_f - t_o)$, as such atoms cannot be distinguished if their energy difference $d\omega$ is much smaller, i.e. $d\omega \ll 1/(t_f - t_o)$. Thus the excitation of the $N \rho(\omega) d\omega$ atoms within the frequency strip ($\omega, \omega + d\omega$) can be considered a collective one. We define the collective atomic operator at detuning $\omega$ as

$$
\hat{c}_{M \omega} \equiv \frac{1}{\sqrt{N \rho(\omega) d\omega}} \sum_{n \in (\omega, \omega + d\omega)} \hat{\sigma}_-^{(n)},
$$

(7)

where the subscript $M$ denotes memory. The collective atomic operator can be approximated as a bosonic operator as it follows approximately the bosonic commutation relation in the low excitation regime, i.e. $[\hat{c}_{M \omega}, \hat{c}_{M \omega'}] \approx \frac{d\omega}{d\omega'} \approx \delta(\omega - \omega')$. We note that the last relation is a valid definition of Dirac delta function when $d\omega \to 0$. By using the collective atomic operators, the total Hamiltonian can be rewritten as

$$
H = \xi(\hat{B}_S^\dagger \hat{B}_I^\dagger + \hat{B}_S \hat{B}_I) + \int \omega \hat{b}_{S \omega}^\dagger \hat{b}_{S \omega} d\omega + \int \omega \hat{b}_{I \omega}^\dagger \hat{b}_{I \omega} d\omega + \int \omega \hat{c}_{M \omega}^\dagger \hat{c}_{M \omega} d\omega + \sqrt{\kappa_S \over 2\pi} \int (\hat{B}_S^\dagger \hat{b}_{S \omega} + \hat{b}_{S \omega}^\dagger \hat{B}_S) d\omega + \sqrt{\kappa_I \over 2\pi} \int (\hat{B}_I^\dagger \hat{b}_{I \omega} + \hat{b}_{I \omega}^\dagger \hat{B}_I) d\omega + G \int \sqrt{\rho(\omega)} (\hat{B}_S^\dagger \hat{c}_{M \omega} + \hat{c}_{M \omega}^\dagger \hat{B}_S) d\omega ,
$$

(8)

where $G \equiv \sqrt{N g}$ is the collectively enhanced atomic ensemble-cavity coupling. We note that the continuum approximation of atomic spectral density can describe the frequency comb if each frequency strip is wider than the spacing of the frequency comb, i.e. one can find a $d\omega$ such that $1/(t_f - t_o) \gg d\omega \gg \Delta$.

The dynamics of the system is governed by the Langevin equations:

$$
\dot{\hat{B}}_I = -\frac{\kappa_I}{2} \hat{B}_I - i\xi \hat{B}_S^\dagger - i\sqrt{\kappa_I} \hat{B}_I^{\text{in}}
$$

(9a)

$$
\dot{\hat{B}}_S = -\frac{\kappa_S}{2} \hat{B}_S - i\xi \hat{B}_I^\dagger - iG \int \sqrt{\rho(\omega)} \hat{c}_{M \omega} d\omega - i\sqrt{\kappa_S} \hat{B}_S^{\text{in}}
$$

(9b)

$$
\dot{\hat{c}}_{M \omega} = -i\omega \hat{c}_{M \omega} - iG \sqrt{\rho(\omega)} \hat{B}_S .
$$

(9c)
The input and output operators are defined as
\[ \hat{B}_\text{in}(t) \equiv \frac{1}{\sqrt{2\pi}} \int \hat{b}_{\omega}(t_o)e^{-i\omega(t-t_o)}d\omega \quad \text{and} \quad \hat{B}_\text{out}(t) \equiv \frac{1}{\sqrt{2\pi}} \int \hat{b}_{\omega}(t_f)e^{-i\omega(t-t_f)}d\omega, \]
for \( u = S, I \).

Our aim is to understand the entanglement generated between the signal and idler output fields and the atomic memory, this can be obtained from the scattering relation of the frequency modes. First, we perform the Fourier transform of Eq. (9b) and the complex conjugate of Eq. (9a),
\[
-i\omega \hat{B}_I^\dagger[\omega] = -\frac{K_I}{2} \hat{B}_I^\dagger[\omega] + i\xi \hat{B}_S[\omega] + i\sqrt{\kappa_I} \hat{B}_I^\dagger[\omega]
\]
\[
-i\omega \hat{B}_S[\omega] = -\frac{K_S}{2} \hat{B}_S[\omega] - i\xi \hat{B}_I^\dagger[\omega] - i\sqrt{\kappa_S} \hat{B}_S^\dagger[\omega] - iG \int \int \phi(\omega')\hat{c}_{M,\omega'}(t)e^{i\omega t}d\omega'dt.
\]  

(10a)

The input-output relation of the radiation modes remains in the standard form
\[
\hat{B}_\text{out}^\dagger[\omega] = \hat{B}_\text{in}^\dagger[\omega] + i\sqrt{\kappa_I} \hat{B}_I^\dagger[\omega]
\]
\[
\hat{B}_\text{out}[\omega] = \hat{B}_\text{in}[\omega] - i\sqrt{\kappa_S} \hat{B}_S[\omega].
\]  

(11a)

(11b)

We can obtain the analogous input-output relation for the atomic memory by integrating Eq. (9c) from the initial time \( t_o \) and to the final time \( t_f \), i.e.
\[
\hat{c}_{M,\omega}(t) = \hat{c}_{M,\omega}(t_o)e^{-i\omega(t-t_o)} - iG\sqrt{\rho(\omega)} \int_{t_o}^{t} \hat{B}_S(t')e^{-i\omega(t-t')}dt'
\]  

(12a)

\[
= \hat{c}_{M,\omega}(t_f)e^{-i\omega(t-t_f)} + iG\sqrt{\rho(\omega)} \int_{t}^{t_f} \hat{B}_S(t')e^{-i\omega(t-t')}dt'.
\]  

(12b)

Subtracting both equations, we have
\[
\hat{c}_{M,\omega}(t_f)e^{i\omega t_f} = \hat{c}_{M,\omega}(t_o)e^{i\omega t_o} - iG\sqrt{2\pi\rho(\omega)} \hat{B}_S[\omega].
\]  

(13)

This relation accounts for the transformation of collective atomic modes after interacting with the cavity. To match the standard definition of photonic input and output operators, we rescale the collective atomic operators as
\[
\hat{\tilde{C}}^\text{in}_M[\omega] = \sqrt{2\pi}\hat{c}_{M,\omega}(t_o)e^{i\omega t_o}, \quad \hat{\tilde{C}}^\text{out}_M[\omega] = \sqrt{2\pi}\hat{c}_{M,\omega}(t_f)e^{i\omega t_f}.
\]  

(14)

For an operation time much longer than the dynamical time scale of the system, we can set \( t_o \rightarrow -\infty \) and \( t_f \rightarrow \infty \), then the integral in Eq. (13) can be recognized as the Fourier transform of \( \hat{B}_S \). The input-output relation of the collective atomic modes can be written as
\[
\hat{\tilde{C}}^\text{out}_M[\omega] = \hat{\tilde{C}}^\text{in}_M[\omega] - iG\sqrt{2\pi\rho(\omega)} \hat{B}_S[\omega].
\]  

(15)

The remaining step to obtain the input-output relation is to evaluate the last integral in Eq. (10b). By substituting in Eq. (12a), we get
\[
-i\omega \hat{B}_S[\omega] = -\frac{K_S}{2} \hat{B}_S[\omega] - i\xi \hat{B}_I^\dagger[\omega] - i\sqrt{\kappa_S} \hat{B}_S^\dagger[\omega]
\]
\[
-iG \int \int \sqrt{\rho(\omega')}\hat{c}_{M,\omega'}(t_o)e^{-i\omega'(t-t_o)}e^{i\omega t}d\omega'dt = -G^2 \int \int \rho(\omega') \hat{B}_S(t')e^{-i\omega(t-t')}dt'e^{i\omega t}d\omega'dt.
\]

The first integral can be evaluated by first integrating \( t \):
\[
\int \int \sqrt{\rho(\omega')}\hat{c}_{M,\omega'}(t_o)e^{-i\omega'(t-t_o)}e^{i\omega t}d\omega'dt = \int \int \rho(\omega') \hat{\tilde{C}}^\text{in}_M[\omega] - 2\pi\delta(\omega - \omega') d\omega'
\]
\[
= \sqrt{2\pi\rho(\omega)} \hat{\tilde{C}}^\text{in}_M[\omega].
\]  

(17)

For the second integral, we first rewrite the integral domain of \( t' \) in terms of the Heaviside step function \( \Theta(t-t') \), and then change the integral of \( t \) to \( \tau \equiv t-t' \):
\[
\int \int \rho(\omega') \Theta(t-t') \hat{B}_S(t')e^{-i\omega'(t-t')}dt'e^{i\omega t}d\omega'dt = \int \int \rho(\omega') \int \Theta(\tau) \hat{B}_S(t')e^{-i\omega'(t-t')}e^{i\omega\tau}d\tau e^{i\omega t}d\omega'dt'
\]
\[
= \int \int \rho(\omega') \Theta(\tau) e^{-i\omega'(t-t')}e^{i\omega\tau}d\tau d\omega' \int e^{i\omega t} \hat{B}_S(t')dt'
\]
\[
= 2\pi \rho_+(\omega) \hat{B}_S[\omega],
\]  

(18)
where the positive-time frequency distribution $\rho_+(\omega)$ is defined as

$$
\rho_+(\omega) = \frac{1}{2\pi} \int \int \rho(\omega') \Theta(\tau) e^{-i\omega'\tau} e^{i\omega\tau} d\tau d\omega' = \frac{1}{2\pi} \int_{0}^{\infty} \int \rho(\omega') e^{-i(\omega'-\omega)\tau} d\omega' d\tau .
$$

(19)

Putting all these into Eq. (10b), we have

$$
-i \omega \hat{B}_S[\omega] = \left(-\frac{\kappa_S}{2} - G^2/2\pi \rho_+(\omega)\right) \hat{B}_S[\omega] - i\xi \hat{B}^\dagger_S[\omega] - i\sqrt{\kappa_S} \tilde{B}^\dagger_S[\omega] - iG\sqrt{2\pi \rho(\omega)} \tilde{C}^\dagger_M[\omega] .
$$

(20)

By combining Eqs. (10a), (20), (11), and (15), the input-output relation can be obtained as (i.e. Eq. (2) in the main text)

$$
\begin{align}
\hat{B}^\text{out}\dagger[\omega] &= T_{1I}(\omega) \hat{B}^\text{in}\dagger[\omega] + T_{1S}(\omega) \hat{B}^\text{in}\dagger_S[\omega] + T_{IM}(\omega) \tilde{C}^\dagger_M[\omega] \\
\hat{B}_S^\text{out}[\omega] &= T_{SI}(\omega) \hat{B}^\text{in}\dagger[\omega] + T_{SS}(\omega) \hat{B}^\text{in}\dagger_S[\omega] + T_{SM}(\omega) \tilde{C}^\dagger_M[\omega] \\
\tilde{C}_M^\text{out}[\omega] &= T_{MI}(\omega) \hat{B}^\text{in}\dagger[\omega] + T_{MS}(\omega) \hat{B}^\text{in}\dagger_S[\omega] + T_{MM}(\omega) \tilde{C}^\dagger_M[\omega] ,
\end{align}
$$

(21a)

(21b)

(21c)

or in the matrix form $\hat{B}^\text{out} = T \hat{B}^\text{in}$, where $\hat{B}^\text{in/out} = (\hat{B}^\text{in/out}\dagger[\omega] \hat{B}^\text{in/out}_S[\omega] \tilde{C}^\text{in/out}_M[\omega])^\text{T}$. The matrix $T$, which contains the scattering amplitudes, can be calculated by

$$
T(\omega) \equiv I_3 + iK^\text{T} (\omega) \left(\omega I_2 - \mathcal{H}(\omega)\right)^{-1} \mathcal{K}(\omega) ,
$$

(22)

where $I_k$ is the $k \times k$ identity matrix;

$$
\mathcal{H}(\omega) \equiv \begin{pmatrix} -i\kappa_I/2 & -\xi \\ -\kappa_S/2 - iG^2/2\pi \rho_+(\omega) & \xi \end{pmatrix} ; \quad \mathcal{K}(\omega) \equiv \begin{pmatrix} \sqrt{\kappa_S} & 0 \\ -i\sqrt{\kappa_S} & 0 \end{pmatrix} .
$$

(23)

System stability

The formalism outlined in the last section applies to any spectral distribution of the atomic ensemble. For simplicity, in this stability analysis we consider a Lorentzian spectrum, $\rho(\omega) = 1/(\pi \omega^2 + \Gamma/2)^2$, which is characterized by only the inhomogeneous broadening $\Gamma$. The merit of Lorentzian spectrum is that simple analytical results can be obtained for various properties of interest (e.g. Eq. (3) in the main text), which are generally challenging to obtain for a non-flat spectrum of atoms. This can provide us valuable physical insights regarding the properties of the system.

To be more specific, we recall that the dynamics of a quantum system coupling to a non-Markovian bath is equivalent to one that interacts with dissipative quasi-modes that couple to Markovian baths (see Refs. [44, 45]). The simplicity of a Lorentzian spectrum is that only one quasi-mode is needed to account for the non-Markovian dynamics. For our setup, we consider the following equivalent system, which instead of a Lorentzian atomic ensemble the signal mode is coupled to a resonant but damping quasi-mode, $\hat{a}_Q$, 

$$
H = \xi (\hat{B}^\dagger_S \hat{B}_I + \hat{B}_S \hat{B}_I) + G(\hat{B}^\dagger_S \hat{a}_Q + a^\dagger_Q \hat{B}_S) + \int \omega \hat{b}^\dagger_{S\omega} \hat{b}_{S\omega} d\omega + \int \omega \hat{b}^\dagger_{I\omega} \hat{b}_{I\omega} d\omega + \int \omega \hat{b}^\dagger_{Q\omega} \hat{b}_{Q\omega} d\omega \\
+ \sqrt{\kappa_S/2\pi} \int (\hat{B}^\dagger_S \hat{b}_{S\omega} + \hat{b}^\dagger_{S\omega} \hat{B}_S)d\omega + \sqrt{\kappa_I/2\pi} \int (\hat{B}^\dagger_I \hat{b}_{I\omega} + \hat{b}^\dagger_{I\omega} \hat{B}_I)d\omega + \sqrt{\Gamma/2\pi} \int \left(\hat{a}^\dagger_Q \hat{b}_{Q\omega} + \hat{b}^\dagger_{Q\omega} \hat{a}_Q\right) d\omega ,
$$

(24)

where signal mode and the quasi-mode are interacting through a beam-splitter interaction with strength $G$, and damping of quasi-mode is modelled as a coupling to a Markovian bath that consists of a continuum of bosonic modes at different detuning, $\hat{b}_{Q\omega}$. As we will see, the quasi-mode damping rate $\Gamma$ coincides with the inhomogeneous broadening in our setup.

The equation of motion of the cavity and quasi- modes are given by

$$
\begin{align}
\dot{\hat{B}}_I &= -\frac{\kappa_I}{2} \hat{B}_I - i\xi \hat{B}^\dagger_S - i\sqrt{\kappa_I} \hat{B}^\dagger_I \\
\dot{\hat{B}}_S &= -\frac{\kappa_S}{2} \hat{B}_S - i\xi \hat{B}^\dagger_I - iG \hat{a}_Q - i\sqrt{\kappa_S} \hat{B}^\dagger_S \\
\dot{\hat{a}}_Q &= -\frac{\Gamma}{2} \hat{a}_Q - iG \hat{B}_S - i\sqrt{\Gamma} \hat{B}^\dagger_Q ,
\end{align}
$$

(25a)

(25b)

(25c)
where \( \hat{B}_Q^\dagger(t_o)e^{-i\omega(t-t_o)}d\omega \). By integrating Eq. (25c) and substituting into Eq. (25b), we get

\[
\dot{\hat{B}}_S = -\frac{\kappa_S}{2} \hat{B}_S - i\kappa\hat{B}_S^\dagger - G^2 \int dt' e^{-\frac{\kappa}{2}(t-t')} \hat{B}_S(t')dt' - G\sqrt{T} \int dt' e^{-\frac{\kappa}{2}(t-t')} \hat{B}_Q^\dagger(t')dt' - i\sqrt{\kappa_S} \hat{B}_S^\dagger
\]

\[
= -\frac{\kappa_S}{2} \hat{B}_S - i\kappa\hat{B}_S^\dagger - G^2 \int \frac{\Gamma}{2\pi} \xi e^{-i\omega(t-t')} d\omega \hat{B}_S(t')dt' - G \int \frac{\Gamma}{2\pi} \frac{1}{-i\omega + \Gamma/2} \hat{B}_Q^\dagger(t_o)e^{-i\omega(t-t_o)}d\omega - i\sqrt{\kappa_S} \hat{B}_S^\dagger.
\]

(26)

On the other hand, in our setup we can derive the equation of motion of signal mode by integrating Eq. (9c) and substituting into Eq. (9b):

\[
\dot{\hat{B}}_S = -\frac{\kappa_S}{2} \hat{B}_S - i\kappa\hat{B}_S^\dagger - G^2 \int dt' \rho(\omega)e^{-i\omega(t-t')} \hat{B}_S(t')d\omega dt' - iG \int \bar{\rho}(\omega) e_M(t_o)e^{-i\omega(t-t_o)}d\omega - i\sqrt{\kappa_S} \hat{B}_S^\dagger.
\]

(28)

For our choice of Lorentzian distribution \( \rho(\omega) = \frac{1}{2\pi \omega^2 + (\Gamma/2)^2} \), it is obvious that Eqs. (28) and (27) are equivalent, up to a frequency dependent phase in each bath mode which does not affect the dynamics of \( \hat{B}_S \) (i.e. \( \sqrt{\frac{\Gamma}{2\pi} \frac{1}{-i\omega + \Gamma/2}} = i\sqrt{\rho(\omega)}e^{i\phi} \) for some phase factor \( \phi \)). Since Eqs. (9a) and (25a) are identical, we can conclude that the cavity mode dynamics is the same in our setup as the quasi-mode model.

As a result, the stability of our setup can be analyzed by studying the equivalent equation of motion (25). We first rewrite Eq. (25) in the matrix form

\[
\dot{B} = MB + B^{in}, \quad \text{where} \quad M \equiv \begin{pmatrix} -\kappa/2 & i\xi & 0 \\ -i\xi & -\kappa/2 & -iG \\ 0 & -iG & -\Gamma/2 \end{pmatrix},
\]

(29)

\( B \equiv (\hat{B}_I^\dagger \hat{B}_S \hat{a}_Q)^T \) and \( B^{in} \equiv (i\sqrt{\kappa_B^\dagger} - i\sqrt{\kappa_B^\dagger} - i\sqrt{\kappa_B^\dagger})^T \). We have assumed \( \kappa_S = \kappa_I \equiv \kappa \) for simplicity. The system is stable if and only if all eigenvalues of \( M \) have strictly negative real parts. This can be verified by checking the Routh-Hurwitz stability criterion of the eigenvalue equation of \( M \) [46]. We find that the system is stable when the parametric drive strength \( \xi \) obeys

\[
\frac{4\xi^2}{\kappa^2} < \min \left\{ C + 1, \left( \frac{C}{2} x + x + 1 \right) (x + 1) \right\},
\]

(30)

where \( C \equiv 4G^2/\kappa\Gamma \) is the atomic ensemble-cavity cooperativity and \( x \equiv \Gamma/\kappa \) is the ratio of inhomogeneous broadening to cavity damping rate.

**Entanglement rate**

We quantify the performance of our system by the rate of entanglement generated between the atomic ensemble memory and the output radiation, i.e. Eq. (4) in the main text. Here we outline the procedure of computing this quantity. We first recall that, according to the scattering relation Eq. (2) in the main text, the idler and signal frequency modes \( B_I^\text{out}[-\omega] \) and signal \( B_S^\text{out}[\omega] \) and the collective atomic mode \( \hat{C}_M^\text{out}[\omega] \) form a three-mode entangled state that is equivalent to a TMS state passing through a beam splitter. The scattering relation also tells us that frequency modes with different \( \omega \) do not couple to other others. As a result, the total state of the output radiation and atomic memory is the tensor product of the state of each frequency mode.

Secondly, because the entangled state of each frequency mode is Gaussian, the entanglement can be quantified by established continuous-variable entanglement measures. We employ the computable measure proposed in Ref. [47], which is generally the lower bound of entanglement of formation. Moreover, as we will see that the state of each frequency mode is balanced, this entanglement measure is exactly the entanglement of formation in our setup. We note that the entanglement rate calculated by our method is the maximum amount of entanglement generated per unit time by the system; this rate will not be increased by any manipulation of the output, e.g. post-processing or non-Gaussian operations.

We computed the entanglement rate for two cases: (i) between the atomic ensemble and all radiation outputs, i.e. both idler and signal, and (ii) between the spin ensemble and idler output only, i.e. signal output is discarded. In case (i), because the three-mode state is pure, the entanglement of formation coincides with the von Neumann entropy of the atoms:

\[
E(\omega) = S(\rho_\omega) = (\bar{n}(\omega) + 1) \log_2 (\bar{n}(\omega) + 1) - \bar{n}(\omega) \log_2 \bar{n}(\omega).
\]

(31)
$\bar{n}(\omega)$ is the mean excitation number of collective atomic mode at frequency $\omega$; this can be obtained from the scattering relation (Eq. (2) in the main text):

$$\bar{n}(\omega) = |T_{MI}(\omega)|^2 = \sin^2 \theta_2(\omega) \sinh^2 r(\omega) .$$  \hspace{1cm} (32)

The last expression relates to the effective TMS strength $r(\omega)$ and relevant BS angle $\theta_2(\omega)$ in the equivalent circuit model in Fig. 1(c).

In case (ii), we first construct the covariance matrix [48] of the two-mode state between idler mode $\hat{B}^\text{out}[-\omega]$ and collective atomic mode $\hat{C}^\text{out}_M[\omega]$. This can be obtained by using the input-output relation in main text Eq.(2), and the fact that the input photonic and initial atomic states are respectively vacuum and ground state, i.e. $\hat{B}^\text{in}_I|\text{vac}\rangle = \hat{B}^\text{in}_I|\text{vac}\rangle = \hat{C}^\text{in}_M|\text{gnd}\rangle = 0$. By expressing the covariance matrix in the standard form of two-mode Gaussian state [? ], we get

$$V = \begin{pmatrix}
\frac{a}{2} & 0 & \frac{b}{2} & 0 \\
0 & \frac{a}{2} & 0 & -\frac{b}{2} \\
\frac{b}{2} & 0 & \frac{c}{2} & 0 \\
0 & -\frac{b}{2} & 0 & \frac{c}{2}
\end{pmatrix}$$  \hspace{1cm} (33)

where the entries are given by

$$a \equiv 2(|T_{IS}(\omega)|^2 + |T_{IM}(\omega)|^2) + 1 = 2\sinh^2 r(\omega) + 1 \hspace{1cm} (34a)$$

$$b \equiv 2|T_{IM}(\omega)T^*_I(\omega)| = 2\sin \theta(\omega) \cosh r(\omega) \sinh r(\omega) \hspace{1cm} (34b)$$

$$c \equiv 2|T_{IM}(\omega)|^2 + 1 = 2\sin^2 \theta_2(\omega) \sinh^2 r(\omega) + 1 . \hspace{1cm} (34c)$$

This two-mode state is balanced in the sense that the anti-diagonal blocks are proportional to a Pauli $Z$ matrix. Ref. [47] shows that the entanglement of formation of such state is exactly given by

$$\mathcal{E}(\omega) = \mathcal{E}_F(\omega) = \left( \sin^2 r_0(\omega) + 1 \right) \log_2 \left( \sin^2 r_0(\omega) + 1 \right) - \sin^2 r_0(\omega) \log_2 \sinh^2 r_0(\omega) , \hspace{1cm} (35)$$

where the effective squeezing parameter $r_0$ can be obtained as

$$e^{2r_0(\omega)} = \frac{1 + \sin \theta_2(\omega) \tanh r(\omega)}{1 - \sin \theta_2(\omega) \tanh r(\omega)} . \hspace{1cm} (36)$$

As a reality check, we verify that the entanglement in case (i) is always higher than that in case (ii), because discarding a part must not increase entanglement. We first note that both Eqs. (31) and (35) have the form of entropy, so their magnitude can be compared using the argument $\bar{n}(\omega)$ and $\sinh^2 r_0(\omega)$. Then it can be easy to see that

$$\bar{n}(\omega) = \sin^2 \theta_2(\omega) \sinh^2 r(\omega) = \frac{\sin^2 \theta_2(\omega) \sinh^2 r(\omega)}{\cosh^2 r(\omega) - \sinh^2 r(\omega)} \geq \frac{\sin^2 \theta_2(\omega) \sin^2 r(\omega)}{\cosh^2 r(\omega) - \sin^2 \theta_2(\omega) \sinh^2 r(\omega)} = \sinh^2 r_0(\omega) , \hspace{1cm} (37)$$

and so $S(\rho_c(\omega)) \geq \mathcal{E}_F(\omega)$, as expected.

**Photon retrieval efficiency**

The entanglement generated between idler output and atomic memory can be retrieved by different means, for instance, by coupling the atomic ensemble to superconducting resonators and transducing to microwave photons. In this section, we discuss the efficiency of a particular way of retrieval that the stored quantum state is converted back to the optical signal photon output.

In the retrieval stage, the parametric drive will be switched off to avoid unwanted mixing with the atomic state that stores the entanglement. We consider a general situation where the inhomogeneously broadened quantum memory will be re-focused after certain storage time, e.g. due to implementation of atomic frequency combs. The stored quantum state will then be converted to photons in the signal cavity, and subsequently will leak to the signal waveguide. Such a process can be described by the scattering relation Eq. (2) in the main text with $\xi = 0$. We note that in employing these scattering relations, the initial state of the atoms (characterized by $\hat{C}^\text{in}_M[\omega]$) is no longer the ground state but rather the stored entangled state. To focus our investigation on the retrieval process, we assume memory does not decohere during the storage, so the $\hat{C}^\text{in}_M[\omega]$ in the retrieval stage can be chosen as the $\hat{C}^\text{out}_M[\omega]$ of the entanglement generation stage.
Since the idler is decoupled when $\xi = 0$, the steady-state scattering relation in the retrieval process involves only a mixing of signal waveguide mode $\hat{B}^n_S[\omega]$ and collective atomic mode $\hat{C}^n_M[\omega]$, both at the detuning $\omega$. The process is thus analogous to a beam splitter with a detuning dependent angle. The efficiency of retrieving the state of the collective atomic mode is thus determined by the scattering amplitude $T_{SM}$, i.e.

$$|T_{SM}(\omega)|^2 = \frac{G^2\Gamma\kappa}{(G^2 + \kappa\Gamma/4 - \omega^2)^2 + (\kappa/2 + \Gamma/2)^2\omega^2}.$$  

(38)

From this explicit expression, a perfect retrieval (i.e. $|T_{SM}|^2 = 1$) requires two conditions: first, the frequency modes have zero detuning $\omega = 0$, and second, the system parameters satisfy $G^2 = \kappa\Gamma/4$, which is the impedance matching condition. For our broadband quantum memory, perfect retrieval is impossible for all frequency modes with any set of system parameters.

Instead of looking for designs and controls to achieve broadband perfect retrieval, which are expected to be challenging to implement and outside the scope of the present work, we study how the imperfect retrieval affects the retrieved entanglement. In analogy to Fig. 2(a) in the main text, we show in Fig. 5 the entanglement rate between idler output and the retrieved signal output. This is computed with the overall scattering relation, which is obtained by cascading the scattering relation in Eq. (2) for the retrieval stage after that for the entanglement generation stage.

![Fig. 5: Normalized entanglement rate between the idler output at the entanglement generation stage and the signal output at the retrieval stage. The system parameters are identical to Fig. 2(a) in the main text. Dotted line represents the impedance matching condition.](image)

With the same generic, weakly driven (i.e. $\xi = 0.1\kappa$) system parameters as in the main text Fig. 2(a), Fig. 5 shows that the impedance matching condition is not optimal for retrieving entanglement from our broadband quantum memory. This is the consequence of two effects: first, as shown in the main text Fig. 2(a), the entanglement generation rate is not optimal at impedance matching; second, entanglement is generated across a spectrum of frequency modes, so the parameters that optimize the resonant mode retrieval do not generally optimize the collective retrieval efficiency of all frequency modes. On the other hand, the retrieved entanglement rate remains considerable for a wide range of system parameters when comparing to the entanglement generated in the atomic memory. As shown in Fig. 6, over 80% of entanglement can be retrieved in this generic example. This is because of the broadband nature of the system (i.e. characteristic frequency range is determined by the cavity-waveguide coupling $\kappa$ and inhomogeneous broadening $\Gamma$), and also the robustness of CV entanglement against loss [49].

We note that the high retrieval rate is not only achievable in the weak drive regime (i.e. $\xi \ll \Gamma, \kappa$). In Fig. 7, we also show the retrieved entanglement rate for a system with realistic parameters given in the main text Table I. By varying the cavity decay rate $\kappa$ and parametric drive strength $\xi$, our simulation shows that the retrieved entanglement rate can be higher than 60 MHz. The retrieval efficiency, which is defined as the ratio of the retrieved entanglement rate to the entanglement generation rate (i.e. Fig. 4(a) in the main text), is over 80% for a wide range of system parameters,
covering both the weak and strong driving regimes. This further supports that our in-situ quantum memory scheme allows high retrieval efficiency without precisely tuning to the impedance matching condition.

FIG. 7: (Left) Retrieved entanglement rate for a system with parameters given in the main text Table I. (Right) Retrieval efficiency of the same system.

**Tripartite entanglement**

Our system will generate a tri-partite entangled state between the respective frequency modes of idler output, signal output, and memory. The state is equivalent to the resultant state after a mode of a two-mode squeezed state interact with another mode in the ground state through a beam splitter. To verify this entangled state is truly tri-partite, we first note that the phases of the two-mode squeezing and beam splitter does not affect the entanglement, therefore we examine the case where both phases are zero for simplicity. The state will then be completely characterized by the two-mode squeezing strength $r$ and the beam-splitting angle $\theta$. The covariance matrix of a general state in this class
will take the form

\[
\begin{pmatrix}
a & 0 & b & 0 & c & 0 \\
0 & a & 0 & -b & 0 & -c \\
b & 0 & d & 0 & e & 0 \\
0 & -b & 0 & d & 0 & -e \\
c & 0 & e & 0 & f & 0 \\
0 & -c & 0 & -e & 0 & f \\
\end{pmatrix},
\]

(39)

where \( a \equiv \frac{1}{2} \cosh 2r \), \( b \equiv \cos \theta \cosh r \sinh r \), \( c \equiv -\sin \theta \cosh r \sinh r \), \( d \equiv \frac{1}{2}(\cos^2 \theta \cosh 2r + \sin^2 \theta) \), \( e \equiv -\sin \theta \cos \theta \sinh^2 r \), \( f \equiv d \equiv \frac{1}{2}(\sin^2 \theta \cosh 2r + \cos^2 \theta) \). A pure state is genuinely multi-partite entangled if every partition of the parties are entangled. In our case, there can be three different partitions: \{signal, idler + memory\}, \{idler, signal + memory\}, and \{memory, idler + signal\}. The entanglement of all three partitions can be verified by showing the reduced state of each mode to be mixed. It is easy to see that \( a > \frac{1}{2}, \ d > \frac{1}{2}, \text{ and } \ f > \frac{1}{2} \) for any non-trivial squeezing parameter \( r \) and beam-splitting angle \( \theta \). This implies the reduced state of all three modes are thermal state with non-zero excitation [48], and thus mixed.

Cavity design and experimental parameters

![Cavity design and experimental parameters](image)

FIG. 8: Experimental resonator design. (Left) Zero-dispersion wavelength for four-wave-mixing phase matching by changing the waveguide geometry. The designed waveguide thickness is 315 and the width is 650 nm. (Right) Spontaneous four-wave-mixing (sFWM) generation rate as a function of pump power in a Si ring resonator with radius of 5.7\( \mu \)m and Q of 10^6.

Here we show detailed calculations of the experimental configuration of a micro-ring resonator on a \( ^{167}\text{Er}^{3+}:\text{Y}_2\text{SiO}_5 \) crystal based on the spectroscopic properties of this quantum memory material [31]. The \( ^{167}\text{Er}^{3+} \) ions are evanescently coupled to the optical field in the silicon waveguide as shown in Fig. 3 and Fig. 8(Left) inset. The inset shows the mode profile of silicon waveguide on top of a \( ^{167}\text{Er}^{3+}:\text{YSO} \) substrate. From the waveguide transverse-magnetic (TM) mode simulation, we engineer the zero dispersion point at pump frequency so that the signal and idler photons will be phase matched at one free-spectral range (FSR) away from the pump frequency, as shown in Fig. 8(Left). In this nonlinear cavity, entangled signal and idler photon pairs can be generated in a spontaneous four-wave-mixing (sFWM) process with a low optical pumping power. According to [29], we estimate the sFWM generation rate as a function of pump power for our system in Fig. 8(Right),

\[
\xi_{\text{FWM}} = \left( \gamma P_{\text{p}} 2\pi \frac{2\pi R}{\nu_g} \right)^2 \left( \frac{Q
\nu_g}{\nu_p R} \right)^3 \frac{v_g}{4\pi R} P_{\text{pump}}^2 = 2 \left( \frac{\text{MHz}}{\mu\text{W}^2} \right) P_{\text{pump}}^2
\]

(40)

where \( \gamma = \frac{\omega_0}{\omega_{\text{eff}}} \). \( A_{\text{eff}} \) is the effective mode area, \( P_{\text{p}} \) is pump power, \( v_g \) is group velocity and \( R=5.7 \mu\text{m} \) is the design ring radius. We operate at a low pump power (<10\muW) for cryogenic experiment.

The generated signal photon is coupled to the 1539 nm optical transition of \( ^{167}\text{Er}^{3+} \) ensemble, transferring the photon-photon entanglement to photon-memory entanglement. The left inset of Fig. 3 and Fig. S5 shows \( ^{167}\text{Er}^{3+} \) hyperfine levels, their spacings and relative detunings between relevant optical transitions near 1539 nm. The ground
state and excited state are spitted into 16 hyperfine states denoted by nuclear spin quantum number \( m_s \). With 7 T magnetic field applied along D1 axis of \( ^{167}\text{Er}^{3+} \): YSO crystal, \( \Delta m_s = 0, \pm 1 \) transitions between the ground and excited hyperfine states become spectrally resolved from each other. In order to initialize the spin states with optical pumping and isolate a two level system for AFC quantum memory, both \( \Delta m_s = 0 \) (like-like) transitions and \( \Delta m_s = \pm 1 \) cross transitions should be spectrally separated by more than the cavity linewidth. The splitting between ground state \( m_s = 7/2 \) and \( m_s = 5/2 \) hyperfine levels at 7 T magnetic field is 796 MHz [31]. A typical Si ring resonator with a quality factor \( Q = 10^6 \) gives a cavity linewidth that is comparable to inhomogeneous broadening of the Er dopants \( \Gamma_{\text{inh}} = 150 \) MHz, and is much smaller than the \( ^{167}\text{Er}^{3+} \) ground state hyperfine splitting. This configuration allows us to couple signal photons dominantly to the \( m_s = 7/2 \) like-like transition (there is a partial overlap with the \( m_s = 5/2 \) like-like transition, which we discuss below) while efficient optical pumping via the \( \Delta m_s = +1 \) cross transition is still achievable with a more intense pump.

FIG. 9: Hyperfine levels and optical transitions of \( ^{167}\text{Er}^{3+} \): YSO with respect to the cavity photons. Left panel illustrates the proposed atomic frequency comb preparation procedure, which uses a cleaning laser resonant with the \( |5/2\rangle_e - |5/2\rangle_g \) transitions to deplete residual population in the \( |5/2\rangle_g \) ground level.

Ideally, we would like the signal photons predominantly coupled to the \( m_s = 7/2 \) like-like transition, and not to other adjacent optical transitions in \( ^{167}\text{Er}^{3+} \) for future spin-wave memory realization. Figure S5 shows in details the energy levels of \( ^{167}\text{Er}^{3+} \): YSO and their spacings in the optical spectrum. The immediately adjacent optical transition is the \( |5/2\rangle_e - |5/2\rangle_g \) transition, which is red-detuned by 93 MHz from the \( |7/2\rangle_e - |7/2\rangle_g \) AFC transition [31]. The next adjacent transition, \( |3/2\rangle_e - |3/2\rangle_g \), is 186 MHz red-detuned, which would have very small coupling to the signal photons given the coupling strength \( G = 100 \) MHz. To suppress potential coupling to the \( |5/2\rangle_e - |5/2\rangle_g \) transition during photon storage, we can implement a cleaning laser around resonance of the \( |5/2\rangle_e - |5/2\rangle_g \) transition while the AFC is created on the \( |7/2\rangle_e - |7/2\rangle_g \) transition. This procedure is illustrated in the left panel of Fig. S5, which would deplete most of the population in the \( |5/2\rangle_g \) level, thus suppresses coupling to the cavity photons. The residual atomic population will end up in the \( |3/2\rangle_g \) or other hyperfine ground states, which do not contribute to the photon storage. We should also note that for the optical two-level AFC storage, this requirement of only coupling to one hyperfine spin level is relaxed: any residual population in the \( |5/2\rangle_g \) level and coupling to it can be considered as an expansion of the AFC bandwidth at the lower frequency tail of the Er transition, which would not degrade the optical AFC memory performance.

The simulated mode volume of our designed ring resonator is \( V_{\text{YSO}} = \frac{f_{\text{YSO}} r_{\text{YSO}} |E(r)|^2}{\epsilon_{\text{max}} |E_{\text{max}}|^2} = 0.7 \mu m^3 \), which leads to \( 1 \times 10^5 \) ions in cavity with a 17 ppm doping concentration. With a single-ion cavity coupling

\[
g_0 = \mu \sqrt{\frac{\omega}{2 \hbar V_{\text{mode}} \epsilon_{\text{max}}}}
\]
and a transition dipole moment \( \mu = 2.07 \times 10^{-32} \text{Cm} \) [50], the collective coupling of the entire ensemble is

\[
\left( \int g_{\text{ion}}^{2}(\vec{r})\rho dV \right)^{1/2} = \left( \frac{\rho \omega \mu^{2}}{2\hbar \epsilon_{\text{Y SO}}} \left[ \vec{\mu} \cdot \vec{E}(\vec{r}) \right]^{2} \right)^{1/2} = 2\pi \times 0.173\text{GHz}. \tag{41}
\]

The collective cooperativity is calculated as

\[
C = \frac{|W(0)|}{\kappa/2} \tag{42}
\]

where \( W(\omega) = G^{2} \int \rho(\omega')d\omega' \) is the atomic absorption for Er ion inhomogeneous distribution \( \rho(\omega') \). When an AFC is prepared in the ensemble, the magnitude of \( G \) is reduced by square root of the AFC finesse, and the cooperativity is reduced by the AFC finesse. For a finesse of 3 in the proposed experiment, the effective cooperativity for a \( \kappa = 2\pi \times 100 \text{ MHz} \) is 6.