SUPPLEMENTARY INFORMATION

doi:10.1038/nature12512

I. SUPPLEMENTARY DISCUSSION

A. Theoretical Framework

In this Section, we present the theoretical framework following the approach of Ref. [1]. To a good approximation, the atomic density is constant across the probe beam, and all experiments are done in the regime where the blockade radius is larger than the beam waist. For these reasons, a one-dimensional approximation holds. Along the propagation direction z, we consider a Gaussian atomic density $\rho(z) = \exp[-z^2/(2\sigma_{ax}^2)]$, normalized by the peak density ρ_0, with root-mean-square width σ_{ax}. We define the peak atom-photon coupling constant g_p via $g_p^2/(\Gamma c) = \rho_0\sigma_p = OD/(\sqrt{2\pi}\sigma_{ax})$, where σ_p is the resonant atomic cross-section and OD the resonant optical depth of the medium. We consider the evolution of the slowly varying operators $\hat{\mathcal{E}}^\dagger(z)$, $\hat{\mathcal{P}}^\dagger(z)$, and $\hat{\mathcal{S}}^\dagger(z)$ corresponding to the creation of a photon, an intermediate-state excitation ($|e\rangle$), and a Rydberg excitation ($|r\rangle$), respectively, at position z. These satisfy the same-time commutation relations $[\hat{\mathcal{E}}(z), \hat{\mathcal{E}}^\dagger(z')] = [\hat{\mathcal{P}}(z), \hat{\mathcal{P}}^\dagger(z')] = [\hat{\mathcal{S}}(z), \hat{\mathcal{S}}^\dagger(z')] = \delta(z-z')$. The Heisenberg equations of motion are then given by [2]

$$\partial_t\hat{\mathcal{E}}(z,t) = -c\partial_z\hat{\mathcal{E}}(z,t) + i\frac{g_p}{2}\sqrt{\rho(z)}\hat{\mathcal{P}}(z,t), \tag{1}$$

$$\partial_t\hat{\mathcal{P}}(z,t) = -\left(\frac{\Gamma}{2} - i(\Delta+\delta)\right)\hat{\mathcal{P}}(z,t) + i\frac{g_p}{2}\sqrt{\rho(z)}\hat{\mathcal{E}}(z,t) + i\frac{\Omega_c}{2}\hat{\mathcal{S}}(z,t), \tag{2}$$

$$\partial_t\hat{\mathcal{S}}(z,t) = -\left(\frac{\gamma_{gr}}{2} - i\delta\right)\hat{\mathcal{S}}(z,t) + i\frac{\Omega_c}{2}\hat{\mathcal{P}}(z,t) - i\int dz' V(z-z')\hat{\mathcal{S}}^\dagger(z',t)\hat{\mathcal{S}}(z',t)\hat{\mathcal{S}}(z,t), \tag{3}$$

where $V(z) = C_6/z^6$, $\Delta = \omega_{re} - \omega_{\text{control}}$ and $\delta = \omega_{\text{probe}} + \omega_{\text{control}} - \omega_{rg}$. Here ω_{probe} and ω_{control} are the probe and control frequencies, while ω_{re} and ω_{rg} are the $|e\rangle \to |r\rangle$ and $|g\rangle \to |r\rangle$ transition frequencies, respectively. The Langevin noise is omitted since it does not affect our calculations [3]. As in Ref. [1], the input is assumed to be a weak coherent state of the form

$$\exp\left[\alpha\int dz(\hat{\mathcal{E}}^\dagger(z) - \hat{\mathcal{E}}(z))\right]|0\rangle, \tag{4}$$

where, for simplicity, α is assumed to be real. The single-photon and two-photon probability amplitudes are defined as $E(z) = \langle 0|\,\hat{\mathcal{E}}(z)\,|\Psi\rangle$ and $EE(z_1,z_2) = \langle 0|\,\hat{\mathcal{E}}(z_1)\hat{\mathcal{E}}(z_2)\,|\Psi\rangle$, respectively, where $|\Psi\rangle$ is the wavefunction of the system while $|0\rangle$ is the vacuum state. We define

$$\psi(z_1,z_2) = \frac{EE(z_1,z_2)}{E(z_1)E(z_2)}, \tag{5}$$

which is unity in the absence of interactions, and

$$\psi(\tau) = \psi(z_1 = z_0 + c\tau, z_2 = z_0), \tag{6}$$

where z_0 is chosen to be outside of the medium: $z_0 \gg \sigma_{ax}$. Then $g^{(2)}(\tau) = |\psi(\tau)|^2$ and $\phi(\tau) = \arg\psi(\tau)$, provided that the single-photon component dominates the denominator of $g^{(2)}$. (This approximation may break down in the presence of strong linear absorption.)

In order to compute $\psi(\tau)$ numerically (solid lines in Figs. 2c and 2d, and solid blue lines in Fig. 3 in the main text), we follow the approach described in the Supplementary Information of Ref. [1]. In particular, the denominator in $\psi(\tau)$ can be easily found analytically using

$$E(z \geq z_0) = \alpha\exp\left[-\frac{\frac{\Gamma}{2}\left(\frac{\gamma_{gr}}{2} - i\delta\right)\frac{OD}{2}}{\left(\frac{\Omega_c}{2}\right)^2 + \left(\frac{\Gamma}{2} - i(\Delta+\delta)\right)\left(\frac{\gamma_{gr}}{2} - i\delta\right)}\right]. \tag{7}$$

This formula (divided by α) is used to obtain the blue and, for $\Omega_c = 0$, the gray dashed curves in Fig. 1c of the main text and in Fig. SI2. It also shows that peak linear transmission is at $\delta = -\frac{\Delta\gamma_{gr}}{\Gamma} + \mathcal{O}(\gamma_{gr}^2)$, while peak (Raman) absorption is at $\delta = \frac{\Omega_c^2}{4\Delta}\left(1 - \frac{\gamma_{gr}^2}{\Omega_c^2}\right)\left(1 + \frac{\Gamma\gamma_{gr}}{\Omega_c^2}\right) + \mathcal{O}(\Delta^{-3})$.

To obtain analytical insight into the physics underlying the numerator of $\psi(\tau)$, we approximate the medium as a homogeneous slab with length $L = 4.2\sigma_{ax}$ [1], *i.e.* we re-define $\rho(z)$ to be 1 in $[0, L]$ and 0 otherwise and rescale g_p^2 by $\sqrt{2\pi}\sigma_{ax}/L$. Furthermore, we take $\delta = \gamma_{gr} = 0$, so that, in particular, $E(z) = \alpha$. Then for $z_1, z_2 \in [0, L]$, repeating the same approximations as in Ref. [1], we obtain a Schroedinger-like equation for the two-photon probability amplitude

$$i\partial_R EE(R, r) = \left[-\frac{1}{2\tilde{m}(r)}\partial_r^2 + U(r)\right] EE(R, r), \tag{8}$$

where the spatially dependent mass $\tilde{m}(r)$ and potential $U(r)$ are given by

$$-\frac{1}{2\tilde{m}(r)} = \frac{4L}{OD}\left(\frac{2\Delta}{\Gamma} + i - \left(\frac{\Omega_c}{\Gamma}\right)^2 \mathcal{V}(r)\right), \tag{9}$$

$$U(r) = \frac{OD}{L}\mathcal{V}(r), \tag{10}$$

and

$$\mathcal{V}(r) = \frac{1}{\frac{2\Delta}{\Gamma} + i + 2\frac{r^6}{r_b^6}}. \tag{11}$$

Here $R = (z_1 + z_2)/2$, $r = z_1 - z_2$, and the resonant blockade radius is defined as $r_b = (2\Gamma C_6/\Omega_c^2)^{1/6}$. The initial condition is a uniform wavefunction $EE(z_1 = 0, z_2) = EE(z_1, z_2 = 0) = \alpha^2$.

By comparing the solutions of Eq. (8) with numerical simulations of the full dynamics, we find that for $\Delta \neq 0$, this equation does not approximate the full dynamics as well as it does for $\Delta = 0$, with the error in EE as large as $\sim 20\%$. However, we find that it still captures the main qualitative features of the two-photon evolution. In the regime $|\Delta| \gg \Gamma, \Omega_c$, an excellent agreement with the full dynamics can be achieved by keeping higher-order derivatives in the effective equation.

In the presence of nonzero Δ, the blockade radius is increased to $\bar{r}_B = r_b\left(\left(\frac{2\Delta}{\Gamma}\right)^2 + 1\right)^{1/12}$. In the limit of $|\Delta| \gg \Gamma$, it corresponds to the off-resonant blockade radius $r_B = (4\,|\Delta|\, C_6/\Omega_c^2)^{1/6}$, defined in the main text. Outside the blockade region, $\tilde{m}(r > \bar{r}_B)$ stems directly from the effective mass of a single dark-state polariton, which, in the limit $|\Delta| \gg \Gamma$, is given by [4–6]

$$m = \frac{2\hbar}{v_g}\tilde{m} = -\frac{1}{16\pi}\frac{c}{v_g}\frac{\lambda}{l_a}\frac{\Gamma}{\Delta}\frac{\hbar\omega}{c^2}. \tag{12}$$

For our parameters, at the center of the medium, it corresponds to $|m| \approx 10^3\,\frac{\hbar\omega}{c^2}$.

These considerations indicate that, for $\Delta > 0 (< 0)$ and $|\Delta| \gg \Gamma, \Omega_c$, we obtain a Schroedinger equation with negative (positive) mass and a potential barrier (dip) within the blockade radius. Note that, for the boundary value problem, the solution for a negative mass and a repulsive potential is formally equivalent to that for a positive mass and an attractive potential under the exchange $EE \to EE^*$. However, for $\Delta < 0$, $\mathcal{V}(r)$, and hence the potential, have a resonant Raman feature around the blockade radius, which breaks this symmetry.

To simplify the analysis, we make an additional approximation by assuming that the boundary conditions are $EE(R = 0, r) = EE(R, r = \pm\infty) = \alpha^2$. We find that this approximation is more forgiving than the approximations used in the derivation of Eq. (8). Dropping the r-dependent term in the effective mass (since it is typically small) and approximating the potential with a square well, we end up with a Schroedinger equation with a complex mass and a square-well potential with a complex amplitude, which can be solved directly (see, *e.g.*, the top theoretical curve in Fig. 1d of the main text). The analytical solution is further simplified if one approximates the square-well potential with a δ function of the same area. This is a reasonable approximation because the variations in $EE(R, r)$ occur at a

scale much larger than $\bar{r}_B$; for $|\Delta| \gg \Gamma$, when the mass and the potential are real, this follows from the fact that there is a single bound state and its extent is much larger than $\bar{r}_B$, as we will verify below. The mass and the potential then simplify to

$$\begin{aligned} -\frac{1}{2\tilde{m}} &= \frac{4L}{OD}\left(\frac{2\Delta}{\Gamma}+i\right), \\ U(r) &= c_0 2OD_B \frac{1}{\frac{2\Delta}{\Gamma}+i}\delta(r), \end{aligned} \tag{13}$$

where $OD_B = OD \times \bar{r}_B/L$ is the optical depth within a blockade radius and $c_0 = \frac{\pi}{2^{1/6}3} e^{i \arg\left(\frac{2\Delta}{\Gamma}+i\right)^{1/6}}$ is chosen to keep $\int dr U(r)$ unchanged under the approximation. For $|\Delta| \gg \Gamma$, $\arg(c_0) = 0$ and $\pi/6$ for $\Delta > 0$ and $\Delta < 0$, respectively, accounting for non-negligible Raman absorption for $\Delta < 0$ and capturing the asymmetry between positive and negative Δ in Figs. 3a and 3b of the main text. The resulting problem is equivalent to a free-particle Schroedinger equation on $R \in [0, L]$ and $r \in [0, \infty]$ with mixed boundary conditions at $r = 0$. Using Laplace transformation in R, we find

$$\psi(\tau = 0) = \psi(R = L, r = 0) = e^{u^2}\mathrm{erfc}(u), \tag{14}$$

where erfc is the complementary error function and

$$u = \frac{c_0\sqrt{OD}OD_B}{2\left(1 - i\frac{2\Delta}{\Gamma}\right)^{3/2}}. \tag{15}$$

This formula was used to make the solid black curves in Figs. 3a and 3b of the main text. At small u, we have $\psi(\tau = 0) - 1 = -\frac{2u}{\sqrt{\pi}} + O(u^2)$, which, for $|\Delta| \gg \Gamma$, gives $\arg[\psi(0) - 1] = \pm\pi/4 + \arg(c_0)$ for $\Delta \lessgtr 0$.

B. Bound state

Within the δ-function approximation and by further assuming that $\Delta \gg \Gamma$, we obtain a real (negative) mass and a real (positive) potential. To get insight into the role of the bound state, we solve for the dynamics in this case analytically:

$$\psi(R, r) = \psi_b(R, r) + \psi_s(R, r), \tag{16}$$

where ψ_b and ψ_s are the contributions of the bound state and the scattering states, respectively,

$$\psi_b(R, r) = 2e^{-\kappa|r|}e^{-i8\kappa^2 RL\Delta/(OD\Gamma)}, \tag{17}$$

$$\psi_s(R, r) = \int_0^\infty dk \frac{i(b_k^{-1} - 1)}{2\pi k}\left(e^{ik|r|} + b_k e^{-ik|r|}\right)e^{i8k^2 RL\Delta/(OD\Gamma)}, \tag{18}$$

$\kappa = c_0(OD\Gamma/\Delta)^2\bar{r}_B/(16L^2)$ and $b_k = (ik+\kappa)/(ik-\kappa)$. Taking $OD = 22$, $\Delta/(2\pi) = 14$ MHz, and $\Omega_c/(2\pi) = 10$ MHz, the condition $\kappa\bar{r}_B = c_0\left(OD_B\Gamma/\Delta\right)^2/16 \approx 1/14 \ll 1$ ensures that the extent of the bound state is indeed much wider than the blockade radius justifying the δ-function approximation. For the case of a square well, $\psi_b(0, r)$, $\psi_s(0, r)$, and $|\psi(L, r)|^2$ are shown as red, black, and solid blue curves in Fig. 1d in the main text. Within this solution, the observed bunching can be understood as resulting from the relative phase evolution between the bound and the scattering states. For the parameters given above, both terms in Eq. (16) contribute to the superpoissonian feature $|\psi|^2 > 0$ at $r = 0$. The bound state ψ_b acquires a phase and becomes the dominant contribution to the imaginary part of ψ. The superposition of scattering states ψ_s starts with a dip (because the bound-state contribution is subtracted), but its phase evolution quickly "fills in" the dip associated with the real part of ψ, while contributing very little to the imaginary part. A combination of both the real and imaginary parts of ψ results in the bunching feature of $|\psi|^2$. Therefore, consistent with a simple intuition, the superpoissonian $g^{(2)}(0)$ is indeed driven by the bound-state formation.

C. Relation between $\psi(\tau)$ and the solution to Eq. (8)

As in Ref. [1], under the approximation of a homogeneous medium, the intuition we have just developed for the region of the z_1-z_2 plane where both photons are inside the medium is not sufficient for computing $\psi(\tau \neq 0)$. Indeed, one has to use this solution to obtain the boundary condition to the problem in the region where one photon is inside the medium while the other photon is already outside. As in Eqs. (9-11) of the Supplementary Information of Ref. [1], the latter problem is equivalent to the retrieval from the medium of a spin-wave (defined along $z_1 = L$) corresponding to the second excitation (see, for example, Eq. (23) in Ref. [7]). While at $\Delta = 0$ and $OD \gg 1$, the retrieval can be described in terms of simple rescaling by the group velocity v_g (provided EIT bandwidth effects are ignored), at nonzero Δ the situation is more complicated [7]. At the same time, the theoretical prediction shown in Fig. 1d of the main text (obtained with $EE(R = L, v_g\tau)$ calculated from Eq. (8) using the boundary conditions $EE(R = 0, r) = EE(R, r = \pm\infty) = \alpha^2$) is in a good qualitative agreement with both the measured $\psi(\tau)$ and the full calculations described above. We emphasize that this comparison can only be interpreted qualitatively, especially at large photon separations, since it (1) assumes simple group velocity propagation and (2) corresponds to a read out of the second excitation along $R = L$ instead of $z_1 = L$. The comparison is, however, better than one might naively expect since the two approximations partially compensate for each other as they, respectively, underestimate and overestimate the mass-induced broadening of the second excitation.

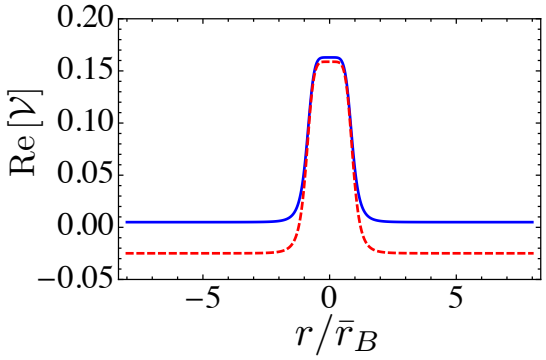
D. Engineering the two-photon potential with a non-zero Raman detuning

While we focused so far on the case of $\delta = \gamma_{gr} = 0$, let us briefly consider the case of nonzero δ and γ_{gr}. To get a qualitative insight, let us assume that δ and γ_{gr} are small enough that the approximations leading to Eq. (8) still hold. Then Eqs. (9,11) are modified to

$$-\frac{1}{2\tilde{m}(r)} = \frac{4L}{OD}\left(\frac{2(\Delta+2\delta)}{\Gamma} + i\left(1+\frac{\gamma_{gr}}{\Gamma}\right) - \left(\frac{\Omega_c}{\Gamma}\right)^2\left(1+\frac{2\delta+i\gamma_{gr}}{2(\Delta+\delta)+i\Gamma}\right)\mathcal{V}(r)\right), \tag{19}$$

$$\mathcal{V}(r) = \left(\frac{2(\Delta+2\delta)}{\Gamma} + i(1+\frac{\gamma_{gr}}{\Gamma}) + \frac{2i\Omega_c^2}{2\Gamma(\gamma_{gr}-i2\delta)+i\Omega_c^2(r_b/r)^6}\right)^{-1}. \tag{20}$$

Since $|\delta|, \frac{1}{2}\Omega_c \ll |\Delta|$, the mass is affected by changes in δ only weakly. At the same time, the potential $U(r)$ is affected substantially by changes in δ. Specifically, we observe (Fig. SI1) that moving towards (away from) the Raman absorption peak at $\delta \approx \Omega_c^2/(4\Delta)$ makes the well deeper (shallower) by shifting the $r \gg \bar{r}_B$ baseline, which implies tighter (weaker) binding of photons, and resulting in more (less) bunching, consistent with the experimental observations in Figs. 3c and 3d of the main text.



Supplementary Figure SI1: **The shape of the effective potential well.** For $\Delta/(2\pi) = 18$ MHz, $\Omega_c/(2\pi) = 10$ MHz, $\gamma_{gr}/(2\pi) = 500$ kHz, $\delta = 0$ (solid blue) and $\delta/(2\pi) = 250$ kHz (dashed red), we plot Re$[\mathcal{V}(r)]$. Moving from $\delta = 0$ towards the Raman resonance at $\delta = 1.3\ (2\pi)$ MHz $\approx \Omega_c^2/(4\Delta)$ makes the potential well deeper by shifting the $r \gg \bar{r}_B$ baseline.

II. SUPPLEMENTARY METHODS: MEASUREMENT OF THE NONLINEAR PHASE

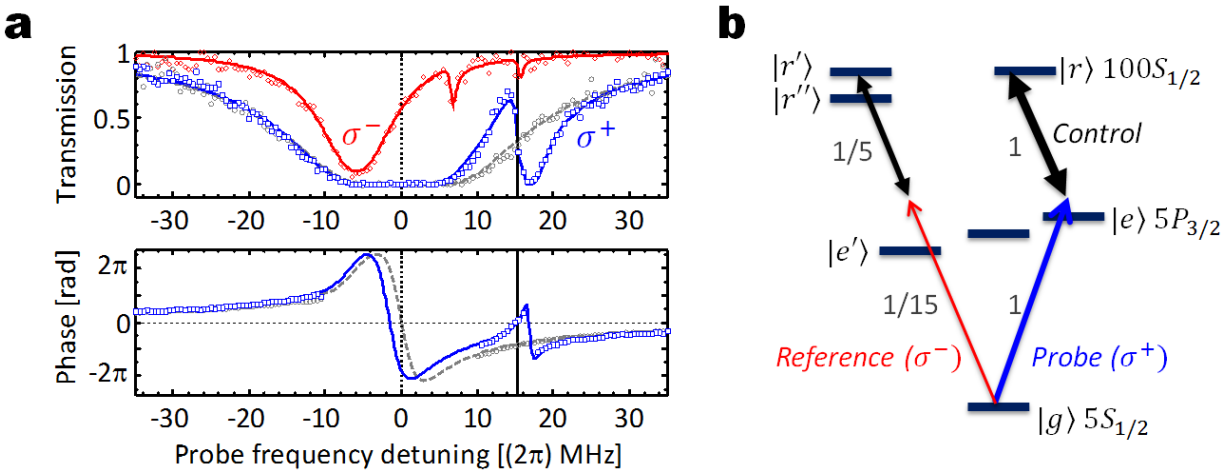
This section provides supplementary information to the Methods section. In order to measure the conditional phase acquired by two circularly polarized σ^+ photons, we effectively create an interferometer, where σ^- photons are used as a phase reference. The interaction of the σ^- field with the atomic medium is much weaker than that of the σ^+ field, as described in Fig. SI2. The incoming state is chosen to be an equal superposition of the two circular polarizations, $(|\sigma^+\rangle + |\sigma^-\rangle)/\sqrt{2}$. In the absence of decoherence, the outgoing one-photon state, detected at time t, is $|1\rangle_t = (\eta_+ |\sigma^+\rangle_t + \eta_- |\sigma^-\rangle_t)/\sqrt{2}$. Here, η_+ and η_- characterize the linear susceptibility of the medium (see Fig. SI2a), accounting for absorption and phase shift leading to polarization rotation. For two photons arriving at times t_1 and t_2 on two single-photon detectors, the corresponding outgoing state is:

$$|1,1\rangle_{t_1,t_2} = \left[\eta_+^2 \psi_{t_1,t_2} |\sigma^+\sigma^+\rangle_{t_1,t_2} + \eta_+\eta_- \chi_{t_1,t_2} \left(|\sigma^+\sigma^-\rangle_{t_1,t_2} + |\sigma^-\sigma^+\rangle_{t_1,t_2}\right) + \eta_-^2 \mu_{t_1,t_2} |\sigma^-\sigma^-\rangle_{t_1,t_2}\right]/2. \quad (21)$$

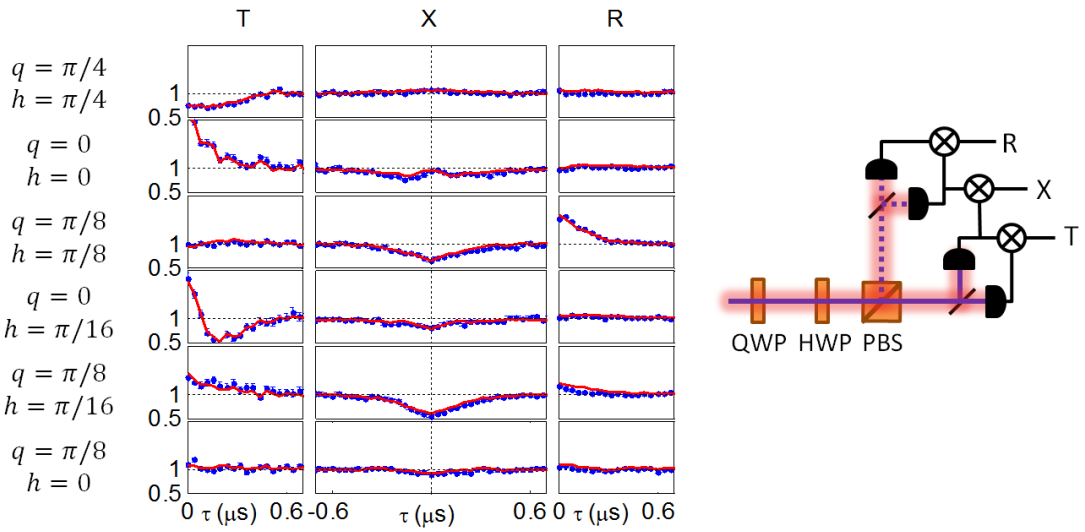
The photon-photon interactions are described by ψ_{t_1,t_2}, χ_{t_1,t_2}, and μ_{t_1,t_2}, which are chosen to be unity in the absence of nonlinear response. Here, the main quantity of interest is $\arg(\psi_{t_1,t_2})$, while the squared amplitude of ψ_{t_1,t_2} is equal to the normalized second-order correlation function of σ^+ photons $g^{(2)}_{++}(t_1,t_2)$.

In the presence of decoherence, the outgoing state of the photons must be described by density matrices, with $\rho^{(1)}(t)$ replacing $|1\rangle_t \langle 1|_t$, and $\rho(t_1,t_2)$ replacing $|1,1\rangle_{t_1,t_2} \langle 1,1|_{t_1,t_2}$. It is convenient to define the interaction matrix $\tilde{\rho}_{i,j}(t_1,t_2) = \rho_{i,j}(t_1,t_2)/[\rho^{(1)}(t_1) \otimes \rho^{(1)}(t_2)]_{i,j}$, with $\arg[\tilde{\rho}_{++,--}(t_1,t_2)] = \arg(\psi_{t_1,t_2}\mu^*_{t_1,t_2})$ being the nonlinear phase-shift of a $\sigma^+\sigma^+$ photon-pair with respect to that of the weakly interacting $\sigma^-\sigma^-$ pair.

The density matrices for the one-photon and two-photon states are measured by quantum state tomography [8],



Supplementary Figure SI2: σ^+ **and** σ^- **spectra. a,** Transmission (top) of the σ^+ polarization in the absence of control field (◯, gray dashed line) and of the σ^+ (□, blue line) and σ^- (◇, red line) with a control field red-detuned by $\Delta = 14$ (2π) MHz. The phase shift between the two circular polarizations (bottom) exhibits a difference of $\sim \pi \approx OD(\Gamma/\Delta)/4$ at EIT resonance (solid vertical line). **b,** Schematic representation of the atomic transitions. The atoms are initially pumped into the ground state $|g\rangle = |5S_{1/2}, F = 2, m_F = 2\rangle$ magnetic sublevel. The σ^+ and σ^- components of the incoming linearly polarized probe light respectively couple to the $|e\rangle = |5P_{3/2}, F = 3, m_F = 3\rangle$ and $|e'\rangle = |5P_{3/2}, F = 3, m_F = 1\rangle$ excited states. For our magnetic field, the Zeeman splitting between these levels is 6 (2π) MHz, comparable to their inverse lifetime $\Gamma = 6.1$ (2π) MHz. The coupling dipole matrix element for the σ^+ transition is larger than for the σ^- transition by a factor of $\sqrt{15}$. The σ^- polarized control field couples the stretched state $|e\rangle$ to the Rydberg state with maximal projections of the nucleus spin (m_I) and total electronic angular momentum (m_J), $|r\rangle = |100S_{1/2}, m_I = 3/2, m_J = 1/2\rangle$. In addition, it couples $|e'\rangle$ to $|r'\rangle = |100S_{1/2}, m_I = -1/2, m_J = 1/2\rangle$ and $|r''\rangle = |100S_{1/2}, m_I = 1/2, m_J = -1/2\rangle$ with a $\sqrt{5}$-times weaker resonant Rabi frequency. Because the magnetic dipole moment of the nuclear spin is negligible compared to that of the electron, the energies of the levels $|r\rangle$ and $|r'\rangle$ are equally shifted by the magnetic field. As a consequence, both σ^+ and σ^- photons propagate under Rydberg EIT conditions. Nevertheless, the combined effect of the Zeeman shift of the intermediate level and the reduced dipole coupling strongly suppresses the probability for σ^- photons to create or be affected by Rydberg blockade. As the σ^- probe photons interact only negligibly between themselves and weakly with σ^+ photons, they provide a reliable reference for measuring phase shifts of the σ^+ photons.



Supplementary Figure SI3: **Normalized photon-photon correlation functions in 6 polarization bases for $\Delta = 1.5\Gamma$. a,** Each setting of the quarter wave-plate (QWP) at an angle q and the half wave-plate (HWP) at an angle h (angles specified on the left) followed by a polarizing beam splitter determines a polarization basis for three $g^{(2)}_{\alpha\beta}(\tau)$ measurements (blue points): for two transmitted photons (T), for two reflected photons (R), and for one-transmitted-one-reflected (X). The bases are equivalent to those proposed in Ref. [10]. The 18 pair counts from 6 different bases are used to tomographically reconstruct the two-photon density matrix using the maximum-likelihood estimation. Together with the reconstructed one-photon density matrix (obtained from the single counts), one can calculate the reconstructed $g^{(2)}_{\alpha\beta}(\tau)$ (red line).

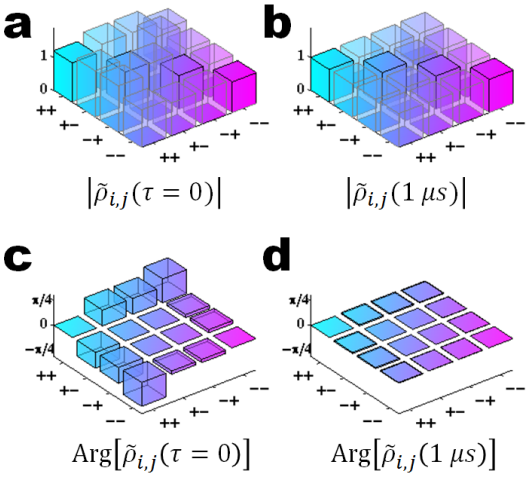
as described in Methods and Fig. SI3. The purity and the concurrence [9] in Fig. 4 of the main text are calculated directly from the two-photon density matrix. Figure SI3 also presents a comparison between $g^{(2)}_{\alpha\beta}(\tau)$ curves measured in different polarization bases (after rescaling for offset at large τ, see discussion below) and those calculated from the reconstructed density matrices after maximum-likelihood estimation. The resulting scaled density matrices are plotted in Fig. SI4 in the limits of proximal ($t_1 = t_2$) and distant (non-interacting) $|t_1 - t_2| = 1\ \mu s$ photons. As expected, at large time separation, the elements of the interaction matrix are all equal to unity. For photons exiting the medium simultaneously, we observe the existence of a large conditional phase-shift between $|\sigma^+\sigma^+\rangle$ and $|\sigma^-\sigma^-\rangle$. The small phase emerging between the $|\sigma^+\sigma^-\rangle$ and $|\sigma^-\sigma^-\rangle$ components underlines that photons with opposite polarizations interact very weakly.

At time seperations $\tau = |t_1 - t_2| \approx 2\ \mu$s, much larger than the typical width of the bunching feature, the measured $g^{(2)}_{++}$ functions exceed unity by about 0.2. Previous measurements of $g^{(2)}_{++}$ with 100 μs-long probe pulses [1] reveal that this deviation is composed of a local ($\tau \lesssim 20\ \mu$s) super-Poissonian feature and a global offset. As discussed in Ref. [1], the occasional decay of an atom in the Rydberg $100S_{1/2}$ state to another metastable Rydberg state creates a temporary EIT blockade in the medium; the resulting fluctuations in the transmission account for the τ-dependent increase in the correlation function. The τ-independent increase is attributed to drifts in the system much slower than the duration of each experimental cycle, with the dominant factor being the frequency drift of the probe laser.

Due to this offset, the measured $g^{(2)}_{++}(t_1, t_2)$ deviates from its ideal value $|\psi_{t_1,t_2}|^2$, where ψ_{t_1,t_2} is defined in Eq. (21). To compensate for this deviation, we model both mechanisms – the temporary local Rydberg blockades and the slow drift – as inhomogeneous fluctuations, and assume in general that, with probability p_i, the linear transmission of the medium changes to T_i. For an incoming photon rate n_0, the mean count rate at the output is $\langle n(t)\rangle = \sum_i p_i T_i n_0$. We further assume for both mechanisms that the change in the transmission does not affect the dynamics of the photon-photon interaction, provided that both photons are eventually transmitted. Under this assumption, the coincident count rate is $\langle n_1(t_1)\, n_2(t_2)\rangle = \sum_i p_i \left|T_i n_0 \psi(t_1, t_2)\right|^2$, and the resulting normalized correlation function is

$$g^{(2)}_{++}(t_1, t_2) = \frac{\langle n_1(t_1)\, n_2(t_2)\rangle}{\langle n_1(t_1)\rangle \langle n_2(t_2)\rangle} = \frac{\sum_i p_i T_i^2}{\left(\sum_i p_i T_i\right)^2} \left|\psi(t_1, t_2)\right|^2 . \tag{22}$$

Finally, assuming no photon-photon interaction ($\psi = 1$) at large enough time separation ($\tau \approx 2\ \mu$s), one finds that normalizing the measured $g^{(2)}_{++}(t_1, t_2)$ by its asymptotic value $g^{(2)}_{++}(\tau \approx 2\ \mu\text{s})$ yields $\left|\psi(t_1, t_2)\right|^2$.



Supplementary Figure SI4: **Tomographic reconstruction of the scaled interaction matrix $\tilde{\rho}$. a,b,c,d,** Amplitude and phase of the scaled two-photon density matrix $\tilde{\rho}_{i,j}(\tau) = \rho_{i,j}(\tau)/[\rho^{(1)} \otimes \rho^{(1)}]_{i,j}$ for two photons with time separation $\tau = 0$ **(a,c)** and $\tau = 1\ \mu s$ **(b,d)** at a detuning of $\Delta = 2.3\Gamma$. All $\tilde{\rho}_{i,j}(1\,\mu s) = 1$, as expected in the absence of nonlinearity. The bunching is evident by $\tilde{\rho}_{++,++} > 1$ **(a)**, while the nonlinear (conditional) phase shift is given by $\arg(\tilde{\rho}_{++,--}) \approx -\pi/4$ **(b)**.

[1] T. Peyronel, O. Firstenberg, Q.-Y. Liang, S. Hofferberth, A. V. Gorshkov, T. Pohl, M. D. Lukin, and V. Vuletic, Nature (London) **488**, 57 (2012).
[2] A. V. Gorshkov, J. Otterbach, M. Fleischhauer, T. Pohl, and M. D. Lukin, Phys. Rev. Lett. **107**, 133602 (2011).
[3] A. V. Gorshkov, R. Nath, and T. Pohl, Phys. Rev. Lett. **110**, 153601 (2013).
[4] F. E. Zimmer, A. André, M. D. Lukin, and M. Fleischhauer, Opt. Comm. **264**, 441 (2006).
[5] F. E. Zimmer, J. Otterbach, R. G. Unanyan, B. W. Shore, and M. Fleischhauer, Phys. Rev. A **77**, 063823 (2008).
[6] J. Otterbach, R. G. Unanyan, and M. Fleischhauer, Phys. Rev. Lett. **102**, 063602 (2009).
[7] A. V. Gorshkov, A. Andre, M. D. Lukin, and A. S. Sorensen, Phys. Rev. A **76**, 033805 (2007).
[8] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
[9] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).
[10] R. B. A. Adamson, L. K. Shalm, M. W. Mitchell, and A. M. Steinberg, Phys. Rev. Lett. **98**, 043601 (2007).