

Entanglement with negative Wigner function of almost 3,000 atoms heralded by one photon

Robert McConnell^{1*}, Hao Zhang^{1*}, Jiazhong Hu¹, Senka Ćuk^{1,2} & Vladan Vuletić¹

Quantum-mechanically correlated (entangled) states of many particles are of interest in quantum information, quantum computing and quantum metrology. Metrologically useful entangled states of large atomic ensembles have been experimentally realized^{1–10}, but these states display Gaussian spin distribution functions with a non-negative Wigner quasiprobability distribution function. Non-Gaussian entangled states have been produced in small ensembles of ions^{11,12}, and very recently in large atomic ensembles^{13–15}. Here we generate entanglement in a large atomic ensemble via an interaction with a very weak laser pulse; remarkably, the detection of a single photon prepares several thousand atoms in an entangled state. We reconstruct a negative-valued Wigner function—an important hallmark of non-classicality—and verify an entanglement depth (the minimum number of mutually entangled atoms) of $2,910 \pm 190$ out of 3,100 atoms. Attaining such a negative Wigner function and the mutual entanglement of virtually all atoms is unprecedented for an ensemble containing more than a few particles. Although the achieved purity of the state is slightly below the threshold for entanglement-induced metrological gain, further technical improvement should allow the generation of states that surpass this threshold, and of more complex Schrödinger cat states for quantum metrology and information processing. More generally, our results demonstrate the power of heralded methods for entanglement generation, and illustrate how the information contained in a single photon can drastically alter the quantum state of a large system.

Entanglement is now recognized as a resource for secure communication, quantum information processing, and precision measurements. An important goal is the creation of entangled states of many-particle systems while retaining the ability to characterize the quantum state and validate entanglement. Entanglement can be verified in a variety of ways; one of the strictest criteria is a negative-valued Wigner function^{16,17}, which necessarily implies that the entangled state has a non-Gaussian wavefunction. To date, the metrologically useful spin-squeezed states^{1–10} have been produced in large ensembles. These states have Gaussian spin distributions and therefore can largely be modelled as systems with a classical source of spin noise, where quantum mechanics enters only to set the amount of Gaussian noise. Non-Gaussian states with a negative Wigner function are however manifestly non-classical, since the Wigner function as a quasiprobability function must remain non-negative in the classical realm. Whereas before this work a negative Wigner function had not been attained for atomic ensembles, in the optical domain, a negative-valued Wigner function has very recently been measured for states with up to 110 microwave photons¹⁸. Another entanglement measure is the entanglement depth¹⁹, that is, the minimum number of atoms that are demonstrably, but possibly weakly, entangled with one another. This parameter quantifies how widely shared among the particles an entangled state is. For a state of an ensemble characterized by collective measurements, the entanglement depth depends sensitively on the proximity of the state to the ideal symmetric subspace of all particles. The largest entanglement depth verified previously has been 170 out of 2,300

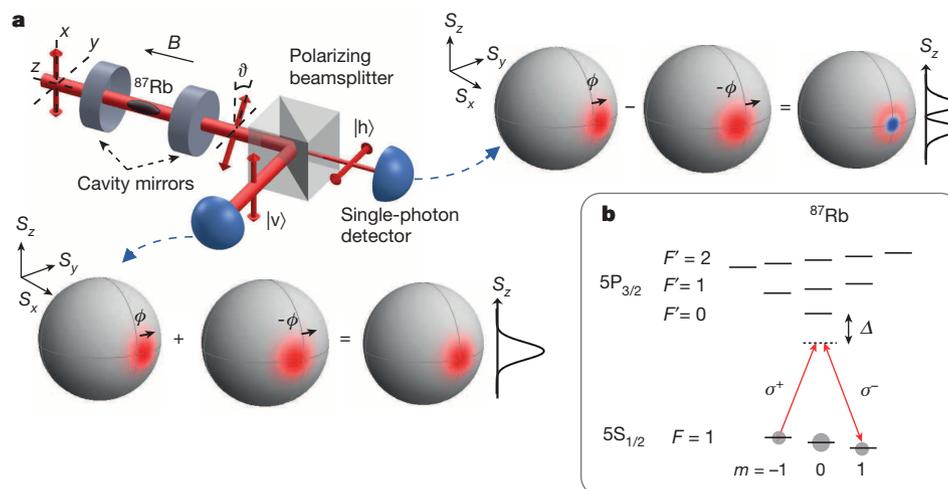


Figure 1 | Scheme for heralded entanglement generation in a large atomic ensemble by single-photon detection. **a**, Incident vertically polarized light (top left) experiences weak polarization rotation ϑ due to atomic quantum noise, and the detection of a horizontally polarized transmitted photon $|h\rangle$ heralds an entangled state of collective atomic spin. An optical resonator formed by two mirrors enhances the polarization rotation and the heralding probability. The Bloch spheres show the ideal Wigner distribution functions for

the collective spin S upon registering the corresponding single-photon detection events. **b**, Atoms of ^{87}Rb in the $5S_{1/2}$, $F=1$ hyperfine manifold are coupled to the excited $5P_{3/2}$ manifold via linearly polarized light, decomposed into two circular polarization components $|\sigma^\pm\rangle$ (indicated by red arrows) that interact with the atomic ground-state populations. The outgoing polarization state of the light reflects the quantum fluctuations between the $|5S_{1/2} F=1, m = \pm 1\rangle$ magnetic sublevels. See Methods for details.

¹Department of Physics, MIT-Harvard Center for Ultracold Atoms, and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ²Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia.

*These authors contributed equally to this work.

atoms for a spin-squeezed state⁶, and very recently 13 out of 41 atoms for a non-Gaussian state¹³.

Here we generate entanglement in a large atomic ensemble by detecting a single photon that has interacted with the ensemble²⁰. An incident vertically polarized photon experiences a weak random polarization rotation associated with the quantum noise of the collective atomic spin. The detection of a horizontally polarized emerging photon then heralds a non-Gaussian entangled state of collective atomic spin (Fig. 1) with a negative-valued Wigner function of -0.36 ± 0.08 , and an entanglement depth of 90% of our ensemble containing several thousand atoms.

The pertinent atom–light interaction is enhanced by an optical cavity, into which we load $N_a = 3,100 \pm 300$ laser-cooled ⁸⁷Rb atoms (Fig. 1a). The atoms are prepared in the $5S_{1/2}$, $F = 1$ hyperfine manifold, such that each atom i can be associated with a spin f_i , and the ensemble with a collective-spin vector $S = \sum_i f_i$. After polarizing the ensemble ($S_z \approx S = F N_a$) by optical pumping, the collective spin state is rotated onto the \hat{x} axis by means of a radio-frequency $\pi/2$ pulse. This (unentangled) initial state, which is centred about $S_z = 0$ with a variance $(\Delta S_z)^2 = S/2$, is known as a coherent spin state (CSS). In our experiment, the atoms are non-uniformly coupled to the optical mode used for state preparation and detection, but the relevant concepts can be generalized to this situation, as discussed in Methods.

Probe light resonant with a cavity mode and detuned from the ⁸⁷Rb D_2 transition is polarization-analysed upon transmission through the cavity. The vertical polarization state of each photon in the incident laser pulse $|v\rangle = (|\sigma^+\rangle + |\sigma^-\rangle)/\sqrt{2}$ can be decomposed into two circular polarization components $|\sigma^\pm\rangle$ that produce opposite differential light shifts between the atomic magnetic sublevels ($m = \pm 1$). Hence a $|\sigma^\pm\rangle$ photon causes a precession of the collective spin vector S in the x – y plane by a small angle $\pm\phi$ (see Methods), and we denote the corresponding slightly displaced CSS by $|\pm\phi\rangle$. Then the combined state of the atom–light system after the passage of one photon can be written as²⁰

$$|\psi\rangle \propto |\sigma^+\rangle|+\phi\rangle + |\sigma^-\rangle|-\phi\rangle \quad (1)$$

Conversely, atoms in the states $|m = \pm 1\rangle$ cause different phase shifts on the σ^\pm photons, resulting in a net rotation of the photon linear polarization if the states $|m = \pm 1\rangle$ are not equally populated. Then the atomic quantum fluctuations between $|m = \pm 1\rangle$ in the CSS randomly rotate the polarization of the input photons $|v\rangle$, giving rise to a non-zero probability $\propto \phi^2$ for an incident $|v\rangle$ photon to emerge in the polarization $|h\rangle = (|\sigma^+ - \sigma^-\rangle)/\sqrt{2}$, orthogonal to its input polarization. The detection of such a ‘heralding’ photon projects the atomic state onto $\langle h|\psi\rangle \propto |\phi\rangle - |-\phi\rangle$, which is not a CSS, but an entangled state of collective spin, namely, the first excited Dicke state²¹ $|\psi_1\rangle$ along \hat{x} (Fig. 1a). In contrast, if the photon is detected in its original polarization $|v\rangle$, the atomic state is projected onto $\langle v|\psi\rangle \propto |\phi\rangle + |-\phi\rangle$, a state slightly spin squeezed¹ and essentially identical to the input CSS. Thus the entangled atomic state $|\psi_1\rangle$ is post-selected by the detection of the heralding photon $|h\rangle$.

From a different perspective, the entangled state is generated by a single-photon measurement event. The incident photon undergoes Faraday rotation by an angle ϑ proportional to the collective spin along the cavity axis, S_z , that exhibits quantum fluctuations around $\langle S_z \rangle = 0$. Since detection of the outgoing photon in $|h\rangle$ is only possible if $S_z \neq 0$, such detection excludes values of S_z near 0 from the spin distribution²⁰, and biases the collective spin towards larger values of $|S_z|$. This creates a ‘hole’ in the atomic distribution near $S_z = 0$, as seen in Fig. 1a.

The mean photon number in the incident laser pulse $k \approx 210$ is chosen such that the probability for one photon to emerge in heralding polarization $|h\rangle$ is $p \approx 0.05 \ll 1$. This ensures a very small probability $\propto p^2$ for producing a different entangled state $|\psi_2\rangle$ heralded by two photons²⁰, a state which, owing to our photon detection efficiency of $q = 0.3 < 1$, we would (mostly) mistake for $|\psi_1\rangle$. This admixture of $|\psi_2\rangle$ to the heralded state is suppressed by a factor of $3p(1 - q) \approx 0.1$. Further state imperfection arises from false heralding events due to residual polarization impurity of the probe beam (independent of the atoms)

of $\sim 3 \times 10^{-5} = 0.1p/k$, adding an admixture of about 10% of the CSS to the heralded state.

In order to reconstruct the collective-spin state generated by the heralding event, we rotate the atomic state after the heralding process by an angle $\beta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$ about the \hat{x} axis before measuring S_z . (Thus $\beta = 0$ corresponds to measuring S_z , $\beta = \pi/2$ corresponds to S_y , and so on). The measurement is performed by applying a stronger light pulse in the same polarization-optimized set-up used for heralding. As the Faraday rotation angle $\vartheta \ll 1$ is proportional to S_z , and the probability

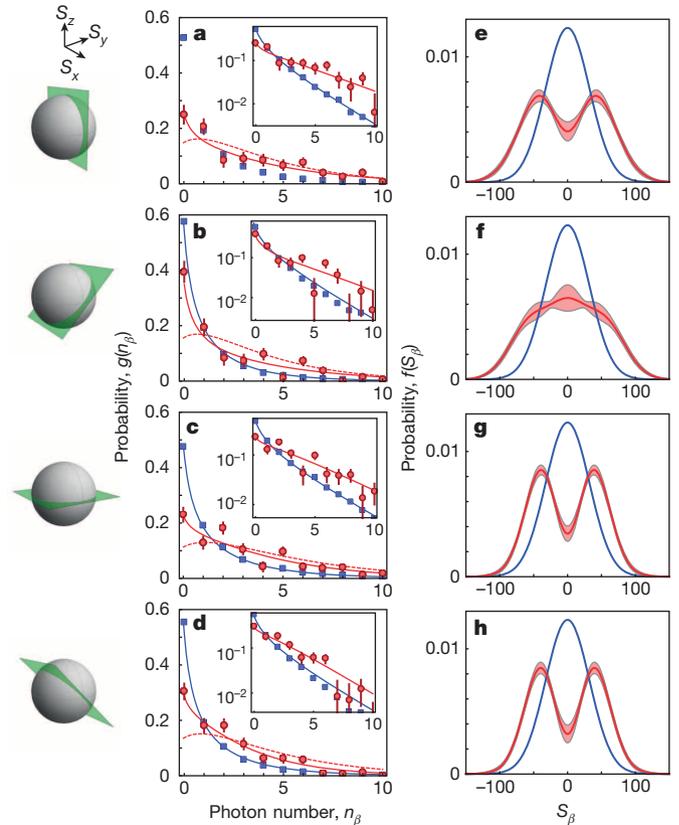


Figure 2 | Collective-spin distribution of atomic state heralded by one photon. **a–d**, Measured photon distributions $g(n_\beta)$ for no heralding photon detected (blue squares), and for one heralding photon detected (red circles), for rotation angles β of 0 (**a**), $\pi/4$ (**b**), $\pi/2$ (**c**) and $3\pi/4$ (**d**). Inset, logarithmic representations of the same data. In the ideal case, the ratio for the heralded state and the CSS is $\langle n_\beta \rangle_{\text{her}} / \langle n_\beta \rangle_{\text{CSS}} = \langle S_\beta^2 \rangle_{\text{her}} / \langle S_\beta^2 \rangle_{\text{CSS}} = 3$ for any angle β , and we measure $\langle n_\beta \rangle_{\text{her}} / \langle n_\beta \rangle_{\text{CSS}} = \{2.7 \pm 0.2, 2.2 \pm 0.2, 2.4 \pm 0.2, 2.1 \pm 0.1\}$ for $\beta = \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$. For each β , the blue and red data sets represent approximately 1.5×10^4 and 200 experiments, respectively. The solid blue and the dashed red curves are predictions without any free parameters, calculated from first principles and the separately measured atom number, for the CSS and the perfect first Dicke state, respectively. The solid red line corresponds to the simultaneous fit to all measurement angles β , that is, the reconstructed density matrix. Error bars, 1 s.d. The planes through the Bloch spheres indicate the measurement direction for the collective spin as specified by the angle β . **e–h**, Reconstructed collective spin distributions of the heralded state (red) for rotation angles β of 0 (**e**), $\pi/4$ (**f**), $\pi/2$ (**g**) and $3\pi/4$ (**h**). The spin distributions of the CSS (blue) are for reference. The horizontal axis S_β is expressed in terms of the effective atom number⁴ $N = (2/3)N_a = 2,100$, obtained by weighting each atom with its coupling strength to the standing-wave probe field inside the cavity, such that the experimentally measured spin fluctuation $(\Delta S_\beta)^2$ of the CSS via its interaction with the probe light satisfies the standard relation $(\Delta S_\beta)^2 = S/2 = NF/2$ for spin F atoms (see Methods). The shaded area indicates the statistical uncertainty of 1 s.d. The spin distribution in **f** shows no ‘hole’ in the middle owing to the lower quality of data for this measurement run, $\beta = \pi/4$.

for detecting $|h\rangle$ -polarized photons is proportional to \mathcal{I}^2 , the measured probability distribution of $|h\rangle$ photon number, $g(n_\beta)$, reflects the probability distribution of S_β^2 . Figure 2a–d shows that a single heralding photon substantially changes the spin distribution towards larger values of $\langle S_\beta^2 \rangle$. We further verify that the heralded state remains (nearly) spin polarized with a contrast of $\mathcal{C} = 0.99_{-0.02}^{+0.01}$, the same as for the CSS within error bars (Fig. 3a).

From the photon distributions $g(n_\beta)$ we can reconstruct the density matrix ρ_{mn} in the Dicke state basis²¹ along \hat{x} , where $|n=0\rangle$ denotes the CSS along \hat{x} , $|n=1\rangle$ the first Dicke state, $|n=2\rangle$ the second Dicke state, and so on. From the density matrix we obtain the Wigner function $W(\theta, \phi)$ on the Bloch sphere²² (Fig. 3). To accurately determine the Wigner function value on the axis, $W(\theta = \frac{\pi}{2}, \phi = 0) = \sum_n (-1)^n \rho_{nn}$, which depends only on the population terms ρ_{nn} , we average the photon distributions $g(n_\beta)$ over four angles β and thereby reduce the fitting parameters to just ρ_{nm} , $n \leq 4$. This is equivalent to constructing a rotationally symmetric Wigner function from the angle-averaged marginal distribution¹⁷. We obtain $\rho_{00} = 0.32 \pm 0.03$, $\rho_{11} = 0.66 \pm 0.04$ with negligible higher-order population terms, giving $W(\frac{\pi}{2}, 0) = -0.36 \pm 0.08$, to be compared to $W(\frac{\pi}{2}, 0) = -1$ for the perfect first Dicke state.

We can also fit the density matrix including the coherence terms simultaneously to $g(n_\beta)$ for all four angles β , without angle-averaging. Since the photon distributions $g(n_\beta)$ depend only on S_β^2 , they determine only the even terms of the density matrix, that is, ρ_{mn} where $m+n$ is even, and contain no information about the odd terms. If we calculate $W(\frac{\pi}{2}, 0)$ from the density matrix without angle-averaging, we find $W(\frac{\pi}{2}, 0) = -0.27 \pm 0.08$, within error bars consistent with the angle-averaged value. In order to display the Wigner function, we bound the odd terms ($m+n$ odd) by verifying that the heralding process does not

displace the state relative to the CSS (see Methods). Therefore we set the odd terms to zero, and display the resulting density matrix and corresponding Wigner function in Fig. 3b–d. The spin distributions $f(S_\beta)$ obtained from this density matrix are shown in Fig. 2e–h.

In order to quantify the minimum number of mutually entangled atoms, we use a criterion derived in ref. 13 that establishes entanglement depth as a function of the populations ρ_{00} and ρ_{11} . From this criterion, generalized to the case of non-uniform coupling to the measurement light field (see Methods), we deduce an average entanglement depth of $N_a = 2,910 \pm 190$ out of $N_a = 3,100$ atoms (Fig. 3e) using the angle-averaged density matrix. Our results represent the first (to our knowledge) experimental verification of the mutual entanglement shared by virtually all atoms in an ensemble that contains more than a few particles.

The above results demonstrate that even with limited resources, that is, weak atom–photon coupling, heralding schemes can be used to boost the effective interaction strength by a large factor, enabling the production of highly entangled states^{20,23}. Furthermore, by repeated trials and feedback the entanglement generation can be made quasi-deterministic^{24,25}. Our approach is related to other heralded schemes for quantum communication^{24–27} and entangled-state preparation^{28–30}, and it would be interesting to generalize the present analysis to infer characteristics of the atomic state from the measured optical signals in those experiments. We note that the same first Dicke state was created in an ensemble of up to 41 atoms with a scheme that uses many heralding photons in a strongly coupled atom–cavity system¹³. In our system, the maximum atom number of $\sim 3,000$ is set by the accuracy of the spin rotation, and could be increased by two orders of magnitude by better magnetic-field control¹⁰. The state purity ρ_{11} can probably be further improved by reducing the heralding probability, and a value of $\rho_{11} > 0.73$ would be required for the Fisher information¹⁴ to exceed that of the CSS, and enable metrological gain of up to 3 dB. The detection of two or more photons prepares Schrödinger cat states²⁰ of the atomic ensemble with

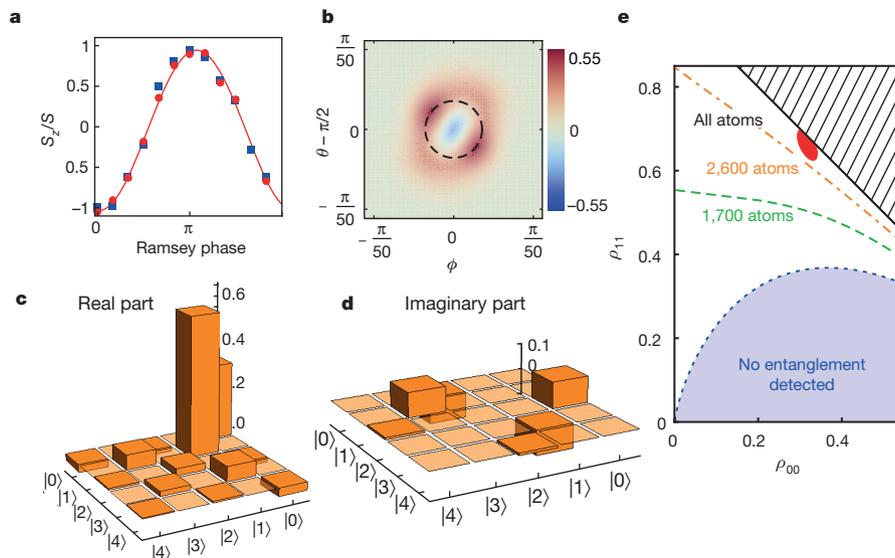


Figure 3 | Reconstruction of the heralded many-atom entangled state.

a, Normalized spin component S_z/S measured in a Ramsey sequence, as a function of the phase of the second Ramsey $\pi/2$ pulse, for the CSS (blue squares) and the heralded state (red circles). The fit (red line) shows a contrast of $0.99_{-0.02}^{+0.01}$ for the heralded state, within error bars the same as the contrast 0.995 ± 0.004 of the CSS. The negligible contrast reduction is expected given that we send only 210 photons into the system at large detuning from atomic resonance. **b**, Reconstructed Wigner function $W(\theta, \phi)$ for the heralded state on the Bloch sphere²² with a radius given by the effective atom number $N = 2,100$. θ is the polar angle with respect to \hat{z} and ϕ is the azimuthal angle with respect to \hat{x} . The first excited Dicke state and the CSS have $W(\frac{\pi}{2}, 0) = -1$ and $W(\frac{\pi}{2}, 0) = 1$, respectively. To provide a reference scale for the size of the

negative region, the black dashed line is the contour at which the CSS has a Wigner function value equal to $1/e$. **c, d**, Real and imaginary parts, respectively, of the reconstructed density matrix elements, in the Dicke state basis along \hat{x} , for the heralded state. **e**, Entanglement depth criterion¹³ for the heralded state, plotted in terms of density matrix elements ρ_{00} and ρ_{11} . The red shaded region represents the 1 s.d. confidence region for the heralded state. Lines (orange, green and blue, labelled with k) represent boundaries for k -particle entanglement in terms of atom number N_a ; a state with ρ_{11} greater than such a boundary displays at least k -particle entanglement. States falling within the blue shaded region are not probably entangled by the used criterion. The hatched area indicates the unphysical region where the density matrix trace would exceed unity.

more metrological gain. We expect that heralded methods can generate a variety of nearly pure, complex, strongly entangled states that are not accessible by any other means at the present state of quantum technology.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 October 2014; accepted 27 January 2015.

- Kitagawa, M. & Ueda, M. Squeezed spin states. *Phys. Rev. A* **47**, 5138–5143 (1993).
- Appel, J. *et al.* Mesoscopic atomic entanglement for precision measurements beyond the standard quantum limit. *Proc. Natl Acad. Sci. USA* **106**, 10960–10965 (2009).
- Takano, T., Tanaka, S.-I.-R., Namiki, R. & Takahashi, Y. Manipulation of nonclassical atomic spin states. *Phys. Rev. Lett.* **104**, 013602 (2010).
- Schleier-Smith, M. H., Leroux, I. D. & Vuletić, V. States of an ensemble of two-level atoms with reduced quantum uncertainty. *Phys. Rev. Lett.* **104**, 073604 (2010).
- Leroux, I. D., Schleier-Smith, M. H. & Vuletić, V. Implementation of cavity squeezing of a collective atomic spin. *Phys. Rev. Lett.* **104**, 073602 (2010).
- Gross, C., Zibold, T., Nicklas, E., Estève, J. & Oberthaler, M. K. Nonlinear atom interferometer surpasses classical precision limit. *Nature* **464**, 1165–1169 (2010).
- Riedel, M. F. *et al.* Atom-chip-based generation of entanglement for quantum metrology. *Nature* **464**, 1170–1173 (2010).
- Hamley, C. D., Gerving, C. S., Hoang, T. M., Bookjans, E. M. & Chapman, M. S. Spin-nematic squeezed vacuum in a quantum gas. *Nature Phys.* **8**, 305–308 (2012).
- Sewell, R. J. *et al.* Magnetic sensitivity beyond the projection noise limit by spin squeezing. *Phys. Rev. Lett.* **109**, 253605 (2012).
- Bohnet, J. G. *et al.* Reduced spin measurement back-action for a phase sensitivity ten times beyond the standard quantum limit. *Nature Photon.* **8**, 731–736 (2014).
- Leibfried, D. *et al.* Creation of a six-atom ‘Schrödinger cat’ state. *Nature* **438**, 639–642 (2005).
- Monz, T. *et al.* 14-qubit entanglement: creation and coherence. *Phys. Rev. Lett.* **106**, 130506 (2011).
- Haas, F., Volz, J., Gehr, R., Reichel, J. & Estève, J. Entangled states of more than 40 atoms in an optical fiber cavity. *Science* **344**, 180–183 (2014).
- Strobel, H. *et al.* Fisher information and entanglement of non-Gaussian spin states. *Science* **345**, 424–427 (2014).
- Lücke, B. *et al.* Detecting multiparticle entanglement of Dicke states. *Phys. Rev. Lett.* **112**, 155304 (2014).
- Leibfried, D. *et al.* Experimental determination of the motional quantum state of a trapped atom. *Phys. Rev. Lett.* **77**, 4281–4285 (1996).
- Lvovsky, A. I. *et al.* Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87**, 050402 (2001).
- Vlastakis, B. *et al.* Deterministically encoding quantum information using 100-photon Schrödinger cat states. *Science* **342**, 607–610 (2013).
- Sørensen, A. S. & Mølmer, K. Entanglement and extreme spin squeezing. *Phys. Rev. Lett.* **86**, 4431–4434 (2001).
- McConnell, R. *et al.* Generating entangled spin states for quantum metrology by single-photon detection. *Phys. Rev. A* **88**, 063802 (2013).
- Arecchi, F. T., Courtens, E., Gilmore, R. & Thomas, H. Atomic coherent states in quantum optics. *Phys. Rev. A* **6**, 2211–2237 (1972).
- Dowling, J. P., Agarwal, G. S. & Schleich, W. P. Wigner distribution of a general angular-momentum state: applications to a collection of two-level atoms. *Phys. Rev. A* **49**, 4101–4109 (1994).
- Agarwal, G. S., Loughovski, P. & Walther, H. Multiparticle entanglement and the Schrödinger cat state using ground-state coherences. *J. Mod. Opt.* **52**, 1397–1404 (2005).
- Duan, L. M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001).
- Matsukevich, D. N. *et al.* Deterministic single photons via conditional quantum evolution. *Phys. Rev. Lett.* **97**, 013601 (2006).
- Kuzmich, A. *et al.* Generation of nonclassical photon pairs for scalable quantum communication with atomic ensembles. *Nature* **423**, 731–734 (2003).
- Simon, J., Tanji, H., Thompson, J. K. & Vuletić, V. Interfacing collective atomic excitations and single photons. *Phys. Rev. Lett.* **98**, 183601 (2007).
- Choi, K. S., Goban, A., Papp, S. B., van Enk, S. J. & Kimble, H. J. Entanglement of spin waves among four quantum memories. *Nature* **468**, 412–416 (2010).
- Christensen, S. L. *et al.* Towards quantum state tomography of a single polariton state of an atomic ensemble. *New J. Phys.* **15**, 015002 (2013).
- Christensen, S. L. *et al.* Quantum interference of a single spin excitation with a macroscopic atomic ensemble. *Phys. Rev. A* **89**, 033801 (2014).

Acknowledgements We thank M. H. Schleier-Smith, E. S. Polzik and S. L. Christensen for discussions. This work was supported by the NSF, DARPA (QUASAR), and a MURI grant through AFOSR. S.C. acknowledges support from the Ministry of Education, Science and Technological Development of the Republic of Serbia, through grant numbers III45016 and OI171038.

Author Contributions The experiment and analysis were carried out by R.M., H.Z., J.H. and S.C.; V.V. supervised the work; all authors discussed the results and contributed to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.V. (vuletic@mit.edu).

METHODS

Photon polarization rotation by atomic spin. Probe laser light red-detuned by $\Delta_0/(2\pi) = -200$ MHz from the ^{87}Rb transition $5^2S_{1/2}, F = 1$ to $5^2P_{3/2}, F' = 0$ is sent through an optical cavity containing the atomic ensemble. We first consider the case where all the atoms are coupled with equal strength to the probe light. For detuning Δ much larger than the excited state linewidth $\Gamma/(2\pi) = 6.1$ MHz, the excited state manifold can be adiabatically eliminated. The vector component of the a.c. Stark shift is described by the Hamiltonian

$$\frac{H}{\hbar} = \frac{g^2}{\Delta} J_z S_z \quad (2)$$

where $J_z = \frac{1}{2}(a_+^\dagger a_+ - a_-^\dagger a_-)$, with a_\pm the annihilation operators for photons with σ^\pm circular polarizations. Here $2g$ is the effective single-photon Rabi frequency taking into account the multiple transitions from $5^2S_{1/2}, F = 1$ to $5^2P_{3/2}, F' = 0, 1, 2$, given by

$$g^2 = (g_{1,1}^{0,0})^2 + (g_{1,1}^{1,0})^2 + (g_{1,1}^{2,0})^2 - (g_{1,1}^{2,2})^2 \quad (3)$$

where $2g_{F,m}^{F',m'}$ is the single-photon Rabi frequency between the ground state $|F = 1, m\rangle$ and the excited state $|F', m'\rangle$. As Δ_0 is comparable to the hyperfine splittings of the $5^2P_{3/2}$ excited states, the interaction strength g^2/Δ is given by

$$\frac{g^2}{\Delta} = \frac{(g_{1,1}^{0,0})^2}{\Delta_0} + \frac{(g_{1,1}^{1,0})^2}{\Delta_0 - \Delta_1} + \frac{(g_{1,1}^{2,0})^2}{\Delta_0 - \Delta_1 - \Delta_2} - \frac{(g_{1,1}^{2,2})^2}{\Delta_0 - \Delta_1 - \Delta_2} \quad (4)$$

where $\Delta_1/(2\pi) = 72$ MHz is the hyperfine splitting between the $F' = 0$ and $F' = 1$ manifolds, $\Delta_2/(2\pi) = 157$ MHz between $F' = 1$ and $F' = 2$, and $\Delta/(2\pi) = -150$ MHz is the effective detuning when $\Delta_0/(2\pi) = -200$ MHz. The value g^2/Δ for our experiment is $2\pi \times 0.7$ kHz.

This vector shift (equation (2)) gives rise to a J_z -dependent Larmor precession of the atomic collective spin S in the x - y plane. Consider one $|\sigma^\pm\rangle$ photon passing through the optical cavity and causing the atomic spin to precess by phase $\pm\phi$. The characteristic atom-photon interaction time is $2/\kappa$, where κ is the cavity linewidth, therefore the atomic phase is given by^{20,31} $\phi = g^2/(\Delta\kappa) = \eta_v\Gamma/(4\Delta)$, where the cavity cooperativity $\eta_v = 4g^2/(\kappa\Gamma) = 0.07$. Another way to think of the Hamiltonian (equation (2)) is that the atomic spin component S_z causes different phase shifts on the photon σ^+ and σ^- components, resulting in a rotation of the linear polarization of the light. The polarization rotation angle $\vartheta = (g^2/\Delta)(S_z/2)(2/\kappa) = \phi S_z$.

In general, the incident light can introduce Raman transitions between different magnetic levels in the $F = 1$ ground state manifold. We apply a bias magnetic field of 4.7 G along the cavity axis to introduce a Zeeman shift between the magnetic levels, so that the Raman coupling is off-resonant. The Larmor frequency is $\omega_L/(2\pi) = 3.3$ MHz, larger than the cavity linewidth $\kappa/(2\pi) = 1.0$ MHz, so that the Raman coupling can be neglected. There is also an unimportant scalar light shift, as well as a tensor light shift that gives rise to squeezing that is negligible for our experimental conditions.

Experimental details. We load an ensemble of ^{87}Rb atoms, cooled to $T = 50$ μK , into a medium-finesse optical cavity (cavity finesse $\mathcal{F} = 5,600$, linewidth $\kappa/(2\pi) = 1.0$ MHz, cooperativity $\eta_0 = 0.2$ at an antinode on a transition with unity oscillator strength). The atoms are confined on the cavity axis by a far-detuned optical dipole trap at 852 nm with trap depth $U/\hbar = 20$ MHz. Characteristics of the optical cavity at the 780 nm probe laser wavelength and the 852 nm trap laser wavelength are summarized in Extended Data Table 1. One Glan-Taylor polarizing beamsplitter (Thorlabs GT5) purifies the polarization of probe light entering the cavity, while a second polarizing beamsplitter after the cavity allows us to measure the rotation of the probe light due to the atomic projection noise. Two Single Photon Counting Modules (SPCMs, models SPCM-AQRH-14-FC and SPCM-AQR-12-FC) are placed at the transmitting and reflecting ports of the polarizing beamsplitter to detect the photons. Owing to the fibre coupling and finite SPCM detection efficiency at 780 nm, the overall quantum efficiency of the detection process is $q = 0.3$.

Definition of effective atom number. Atoms are optically confined at the antinodes of the 852 nm trap laser standing wave. The 780 nm probe light in the cavity forms a standing wave that is incommensurate with the trap standing wave. Consequently, the atoms experience spatially varying couplings to the probe light and rotate the probe photon polarization by different amounts. For an atom at position z on the cavity axis, the cooperativity is $\eta(z) = \eta_v \sin^2(kz)$. When N_a atoms are prepared in a CSS, the atomic projection noise gives rise to fluctuations of the photon polarization rotation. The measured variance of the polarization rotation is proportional to $\frac{N_a}{2} \langle \eta^2(z) \rangle$ where averaging is performed over the position z . This variance differs by a factor of order unity from that of a CSS consisting of N_a atoms uniformly coupled to the light. As described in a previous paper⁴, we introduce the effective atom number N and the effective cavity cooperativity η to satisfy two conditions: that the experimentally measured variance equals that of N uniformly coupled

atoms, $\frac{N_a}{2} \langle \eta^2(z) \rangle = \frac{N}{2} \eta^2$, and that the total amount of interaction between the atomic ensemble and the probe light is the same, that is, $N_a \langle \eta(z) \rangle = N\eta$. To satisfy these two conditions we define the effective atom number $N = \frac{2}{3} N_a$ and the effective cavity cooperativity $\eta = \frac{3}{4} \eta_v$. This re-scaling allows direct comparison with the well-known expressions for the uniformly coupled CSS.

As in the main text and the rest of Methods, S_z refers to the collective spin of an ensemble containing N effective atoms, and therefore the atomic spin precession phase for each transmitting cavity photon is given by $\phi = \eta\Gamma/(4\Delta) = (3/4)\eta_v\Gamma/(4\Delta)$. Note that this value $\eta = 0.05 < 1$ corresponds to the weak atom-cavity coupling regime. For our parameters, $\phi = 5 \times 10^{-4} \ll \phi_{\text{CSS}} = 1.5 \times 10^{-2}$ where $\phi_{\text{CSS}} = \sqrt{1/(2S)}$ is the angular r.m.s. width of the CSS.

Choice of the heralding photon number. The heralding light must be weak enough that it does not introduce substantial decoherence of the desired atomic state. The fundamental shot noise between the σ^+ and σ^- circular polarization components of the heralding light gives rise to phase broadening of the atomic state, which limits the purity of the heralded entangled state. To measure the phase broadening, heralding light pulses with variable photon number are sent into the cavity, and the variance ΔS_z^2 is measured by applying a radio-frequency $\pi/2$ pulse to rotate the atomic state about the \hat{x} direction before measuring ΔS_z^2 . Extended Data Fig. 1 shows the measured atomic state variance ΔS_z^2 as a function of the photon number in the heralding light, in agreement with the predicted linear dependence. The heralding photon number is thus chosen to be ~ 210 , with corresponding herald detection probability $pq = 1.5\%$, to give fairly small phase broadening. Lower heralding photon number results in a purer heralded state, but at the expense of a lower heralding and state generation probability.

Relation between $f(S_\beta)$ and $g(n_\beta)$. To measure the atomic state spin distribution, measurement light with the same polarization $|\nu\rangle$ as the heralding light is sent through the atoms, and the number of photons with the orthogonal polarization $|\hbar\rangle$ is measured. The measurement light contains a large number of input photons $n_{\text{in}} = 1.7 \times 10^4$ to perform destructive measurements with good signal-to-noise ratio. The photon polarization is rotated by a small angle $\vartheta = \phi S_z$ and the probability for each photon to emerge in $|\hbar\rangle$ is ϑ^2 . For a given number of input photons n_{in} , the average number of detected photons with $|\hbar\rangle$ polarization is $\langle n \rangle = qn_{\text{in}}(\phi S_z)^2$, where q is the overall quantum efficiency. Therefore, a spin distribution $f(S_z)$ is mapped to a measured photon distribution $g(n)$. For a given S_z , the detected photons follow a Poisson distribution with mean number $\langle n \rangle$, and the probability to measure exactly n photons is given by

$$P(n, S_z) = \exp[-qn_{\text{in}}(\phi S_z)^2] \frac{[qn_{\text{in}}(\phi S_z)^2]^n}{n!} \quad (5)$$

For an atomic state with the spin distribution $f(S_z)$, the photon distribution $g(n)$ is given by

$$g(n) = \sum_{S_z} f(S_z) P(n, S_z) = \sum_{S_z} f(S_z) \exp[-qn_{\text{in}}(\phi S_z)^2] \frac{[qn_{\text{in}}(\phi S_z)^2]^n}{n!} \quad (6)$$

In order to measure the spin along a general direction, the atomic spin is rotated by an angle β with a radio-frequency pulse before detection. Replacing S_z by S_β in equation (6), we write the relation between the spin distribution $f(S_\beta)$ and the measured photon distribution $g(n_\beta)$ as

$$g(n_\beta) = \sum_{S_\beta} f(S_\beta) P(n_\beta, S_\beta) = \sum_{S_\beta} f(S_\beta) \exp[-qn_{\text{in}}(\phi S_\beta)^2] \frac{[qn_{\text{in}}(\phi S_\beta)^2]^{n_\beta}}{n_\beta!} \quad (7)$$

Choice of the measurement photon number. The measurement photon number is chosen to optimize the readout quality. Extended Data Fig. 2 illustrates the dependence of readout on the input measurement photon number n_{in} by showing how the reconstructed distributions $f(S_z)$ change as n_{in} is varied (the method of reconstruction is discussed later). When the photon number is small, there is large detection noise due to photon shot noise, reflected as the large error band. With increasing photon number, the photon scattering by atoms into free space increases and the atomic state is more strongly perturbed, therefore the ‘dip’ at $S_z = 0$ becomes less distinct. To balance these two competing effects, the optimized atomic-state-measurement photon number is set to 1.7×10^4 .

Subtracting background photon counts. Owing to the residual polarization impurity of the measurement light, there are a small number of background photon counts even when there are no atoms. The background counts account for about 4% of the photon signal of the heralded state. We independently measure the background photon distribution and subtract it from the directly measured atomic signal to obtain $g(n_\beta)$. If we were not to correct for these background counts, we would overestimate the density matrix population ρ_{11} by 10%.

Reconstruction of the density matrix. Using the measured photon distributions $g(n_\beta)$ for all four angles $\beta = 0, \pi/4, \pi/2, 3\pi/4$, the density matrix ρ of the heralded state can be reconstructed.

As the entangled state maintains $0.99^{+0.01}_{-0.02}$ contrast, the magnitude of the total spin $S \approx N$ and we can express the density matrix in the basis of Dicke states $|m\rangle_x$ along the \hat{x} direction

$$\rho = \rho_{00}|0\rangle_x \langle 0|_x + \rho_{11}|1\rangle_x \langle 1|_x + \rho_{01}|0\rangle_x \langle 1|_x + \rho_{10}|1\rangle_x \langle 0|_x + \rho_{22}|2\rangle_x \langle 2|_x + \rho_{02}|0\rangle_x \langle 2|_x + \rho_{20}|2\rangle_x \langle 0|_x + \dots \quad (8)$$

The spin distribution $f(S_\beta)$ can be written as a function of atom number N and the density matrix elements ρ_{00}, ρ_{11} , and so on:

$$\begin{aligned} f(S_\beta, \rho, N) &= \langle S_\beta | \rho | S_\beta \rangle \\ &= \rho_{00} G(0, S_\beta) G^*(0, S_\beta) + \rho_{11} G(1, S_\beta) G^*(1, S_\beta) + \rho_{01} G(0, S_\beta) G^*(1, S_\beta) \\ &\quad + \rho_{10} G(1, S_\beta) G^*(0, S_\beta) + \rho_{22} G(2, S_\beta) G^*(2, S_\beta) \\ &\quad + \rho_{02} G(0, S_\beta) G^*(2, S_\beta) + \rho_{20} G(2, S_\beta) G^*(0, S_\beta) + \dots \end{aligned} \quad (9)$$

Here $G(m, S_\beta) = \langle S_\beta | m \rangle_x$ is the wavefunction of Dicke state $|m\rangle_x$ in the representation of spin component S_β and is given by

$$G(m, S_\beta, N) = \frac{1}{\sqrt{2^m m!}} \left(\frac{1}{\pi N} \right)^{1/4} e^{im\beta - S_\beta^2 / (2N)} H_m \left(\sqrt{\frac{1}{N}} S_\beta \right) \quad (10)$$

where $H_m(x)$ is the m th order Hermite polynomial and N is the atom number. Using equation (7), we write the theoretically predicted photon distribution $g_{\text{th}}(n_\beta)$ as a function of the density matrix ρ , atom number N and input photon number n_{in}

$$\begin{aligned} g_{\text{th}}(n_\beta, \rho, N, n_{\text{in}}) &= \sum_{S_\beta} f_{\text{th}}(S_\beta, \rho, N) P(n_\beta, S_\beta) \\ &= \sum_{S_\beta} f_{\text{th}}(S_\beta, \rho, N) \exp \left[-qn_{\text{in}}(S_\beta \phi)^2 \right] \frac{[qn_{\text{in}}(S_\beta \phi)^2]^{n_\beta}}{n_\beta!}. \end{aligned} \quad (11)$$

We independently measure the input photon number n_{in} and find the atom number N by fitting the photon distributions of the CSS, whose only non-zero density matrix element is $\rho_{00} = 1$. The fitted atom numbers N for different angles β agree within 15% with the values independently measured from the shift of the cavity resonance. We then use the density matrix ρ of the heralded state as the only free parameter, to fit the theoretical distributions $g_{\text{th}}(n_\beta)$ to the measured photon distributions $g(n_\beta)$ along all four angles β . We do this by minimizing the least squares deviation D weighted by the error σ_g of $g(n_\beta)$, given by

$$D = \sum_{\beta} \sum_{n \geq 0} \left[\frac{g_{\text{th}}(n_\beta, \rho) - g(n_\beta, \rho)}{\sigma_g} \right]^2 \quad (12)$$

Since the photon distributions $g(n_\beta)$ measure S_β^2 , we can obtain the even terms of the density matrix (ρ_{mm} where $m + n$ is even) and our measurements are not sensitive to the odd terms. Because the overall heralding probability is $pq = 1.5\%$, the higher-order Dicke state components are exponentially suppressed. We fit the density matrix up to Dicke state $|4\rangle_x$. The fitted values $\rho_{22} = 0.03 \pm 0.02$, $\rho_{33} = 0.02 \pm 0.01$, $\rho_{44} = 0.01 \pm 0.01$ agree with the theoretical expectation²⁰ for our system.

From the fitted density matrix ρ (with coherence terms) we obtain the spin distributions $f(S_\beta)$ using equation (9) for different angles β , as shown in Fig. 2e–h.

To reconstruct the Wigner function for the spin state on the Bloch sphere^{20,22}, we convert ρ from the Dicke state basis into the spherical harmonic basis and obtain the normalized Wigner function according to

$$W(\theta, \phi) = \frac{1}{\sqrt{2S/\pi}} \sum_{k=0}^N \sum_{q=-k}^k \rho_{kq} Y_{kq}(\theta, \phi) \quad (13)$$

where the terms ρ_{kq} represent the density elements in the spherical harmonic basis and $Y_{kq}(\theta, \phi)$ are the spherical harmonics, with θ, ϕ being the polar and azimuthal angles on the Bloch sphere, respectively. The normalization factor $\sqrt{2S/\pi}$ is chosen such that the CSS has $W(\frac{\pi}{2}, 0) = 1$. Note that, in the limit of large atom number, this normalization also means that the pure first excited Dicke state has $W(\frac{\pi}{2}, 0) = -1$, and generally the value of the Wigner function on the \hat{x} axis depends only on the populations ρ_{mm} such that $W(\theta = \frac{\pi}{2}, \phi = 0) = \sum_n (-1)^n \rho_{nn}$.

Measurement of mean value of S_z . The measured photon distributions $g(n_\beta)$ do not give information about the density matrix odd terms (ρ_{mn} where $m + n$ is odd). In order to bound the odd terms, we verify that the heralding process does not displace the produced heralded state relative to the CSS. This is accomplished by

performing a measurement with a probe beam polarized at 45° relative to $|v\rangle$, such that the difference between the measured $|h\rangle$ and $|v\rangle$ photon numbers is proportional to S_z . We find a heralding-light-induced shift $\delta\langle S_z \rangle = -0.2 \pm 1.6$, consistent with zero, and very small compared to the CSS r.m.s. width $(\Delta S_z)_{\text{CSS}} \approx 30$. Therefore we set the odd terms of the density matrix to zero in Fig. 3b–d.

Entanglement depth for finite contrast. Entanglement depth is defined as the minimum number of entangled particles in an ensemble. A fully separable pure state can be written as $|\varphi\rangle = |\varphi_1\rangle \otimes \dots \otimes |\varphi_N\rangle$, where N is the atom number. A pure k -producible state can be written as $|\varphi\rangle = |\varphi_1^{1\dots k_1}\rangle \otimes \dots \otimes |\varphi_M^{1\dots k_M}\rangle$, where $k_1, \dots, k_M \leq k$, $k_1 + \dots + k_M = N$. If a state cannot be written as a pure $(k-1)$ -producible state or a mixed state of $(k-1)$ -producible states, then it has entanglement depth of at least k .

We slightly generalize the entanglement criterion derived in ref. 13 to take into account the finite contrast C of the collective atomic spin in our experiment. The derivation in ref. 13 considers the case in the fully symmetric Dicke subspace of N atoms, and finds that for a k -producible state the maximum population of the first Dicke state $\rho_{11}(P_1)$ as a function of the CSS population $\rho_{00}(P_0)$ is

$$\max_{P_0} P_1 = \frac{P_0}{N} \max \left[\sqrt{k} \max_{\prod_{i=1}^{M-1} a_i = x} F_{M-1}(a_1, \dots, a_{M-1}) + \sqrt{k} F_1(\sqrt{P_0}/x) \right]^2 \quad (14)$$

Here $M = [N/k]$, $k' = N - k(M-1)$, and $F_n(a_1, \dots, a_n) = \sum_{i=1}^n \frac{\sqrt{1-a_i^2}}{a_i}$. Equation (14) is generally not a concave function of P_0 . In order to obtain the upper bound for mixed states, denote the concave hull of the right side of equation (14) as $\mathcal{B}(P_0, k, N)$. We define $\mathcal{B}(P_0, k, N) = B(P_0, k, N)/N$. Note that when $N_1 < N_2$, $B(P_0, k, N_1) \leq B(P_0, k, N_2)$.

The heralded state we produce does not necessarily retain perfect contrast, so the state can be a mixture of different total spins $S = N, N-1, \dots, N(1-\epsilon)$, with $\epsilon \approx 1\%$. The contrast loss is mainly caused by the decoherence between $F=1$ magnetic sublevels, and the free space scattering of the heralding light by the atoms. We decompose the density matrix ρ into the total spin basis

$$\rho = \sum_{i=0}^{\epsilon N} w_i \rho_{N-i} \quad (15)$$

Here ρ_{N-i} is the density matrix in the subspace of total spin $S = N - i$, w_i is the weight for each ρ_{N-i} and $\sum w_i = 1$. For each ρ_{N-i}

$$\mathcal{B}(P_0, k, N-i) = B(P_0, N-i, k, N-i) / (N-i) \leq B(P_0, N-i, k, N) / (N-i) \quad (16)$$

Here $P_{0, N-i}$ is the probability for the state to be found in the ground state in the subspace of total spin $N - i$.

Measurements of the spin distributions do not allow us to determine the total spin of the system at single-atom resolution. We define populations of the CSS and the first Dicke state by

$$P_0 = \sum_{i=0}^{\epsilon N} w_i P_{0, N-i} \quad (17)$$

$$P_1 = \sum_{i=0}^{\epsilon N} w_i P_{1, N-i} \quad (18)$$

The upper bound of P_1 is given by

$$\max_{P_0} P_1 \leq \sum_{i=0}^{\epsilon N} w_i \max_{P_{0, N-i}} P_{1, N-i} \leq \sum_{i=0}^{\epsilon N} w_i \mathcal{B}(P_0, N-i, k, N-i) / (N-i) \quad (19)$$

Using equation (16) and the fact that $B(P_0, k, N)$ is a concave function of P_0 we have

$$\begin{aligned} \max_{P_0} P_1 &\leq \sum_{i=0}^{\epsilon N} w_i \mathcal{B}(P_0, N-i, k, N) / (N - \epsilon N) \\ &\leq \frac{1}{(1-\epsilon)N} \mathcal{B} \left(\sum_{i=0}^{\epsilon N} w_i P_{0, N-i}, k, N \right) \\ &= \frac{1}{C} \mathcal{B}(P_0, k, N). \end{aligned} \quad (20)$$

Here C is the contrast of the collective spin. Comparing to ref. 13, the result is modified by a factor $1/C$. In our experiment, $C = 0.99^{+0.01}_{-0.02}$, so the effects of finite contrast on entanglement depth are minimal.

Entanglement depth in terms of the actual atom number. In the experiment the atoms have spatially varying coupling to the probe light. However, the criterion in ref. 13 is derived for the case where atoms are equally coupled to the light. Here we

generalize the entanglement criterion to our experimental conditions and prove that the sample-averaged fractional entanglement depth for the ensemble containing 3,100 actual non-uniformly coupled atoms is the same as that of 2,100 uniformly coupled effective atoms. Consider an ensemble of N_a actual atoms where each atom j has spin component $f_{z,j}$ and cooperativity η_j . The effective total spin of the ensemble is S_z and the effective cooperativity is η , so that

$$S_z \eta = \sum_{j=1}^{N_a} f_{z,j} \times \eta_j \quad (21)$$

As mentioned in the main text, the ideal heralded state $|\psi_1\rangle$ (the first Dicke state of non-uniformly coupled atoms) is the destructive interference of two slightly displaced CSSs $|\pm\phi\rangle$ and can be written as

$$\begin{aligned} |\psi_1\rangle &\propto |\phi\rangle - |-\phi\rangle \\ &= \left[e^{iS_z \eta \Gamma / (4d)} - e^{-iS_z \eta \Gamma / (4d)} \right] |\psi_0\rangle \\ &= \left[e^{i\Gamma / (4d) \sum_{j=1}^{N_a} f_{z,j} \eta_j} - e^{-i\Gamma / (4d) \sum_{j=1}^{N_a} f_{z,j} \eta_j} \right] |\psi_0\rangle \end{aligned} \quad (22)$$

where $|\psi_0\rangle$ is the initial CSS along \hat{x} . By expanding the exponent to first order and using $f_z = (f_{+,x} - f_{-,x}) / (2i)$, we get

$$|\psi_1\rangle = \left(\sum_{j=1}^{N_a} \eta_j^2 \right)^{-1/2} \sum_{j=1}^{N_a} \eta_j \left[\prod_{j' \neq j} |0_{j'}\rangle_x \right] \otimes |1_j\rangle_x \quad (23)$$

where $|0_j\rangle_x$ and $|1_j\rangle_x$ are the single-particle spin eigenstates along \hat{x} of the atom j . For a fully separable state $|\varphi\rangle = \prod_{j=1}^{N_a} (\alpha_j |0_j\rangle_x + \beta_j |1_j\rangle_x + \dots)$ the population $P_1 = |\langle \varphi | \psi_1 \rangle|^2$ is given by

$$P_1 = \left(\sum_{j=1}^{N_a} \eta_j^2 \right)^{-1} \left| \sum_{j=1}^{N_a} \eta_j \beta_j \prod_{j' \neq j} \alpha_{j'} \right|^2 \quad (24)$$

The expression for P_1 is similar to that in ref. 13 and differs by the additional weight factor η_j . When the real atom number $N_a \gg 1$, the upper bound of P_1 for the fully separable state $|\varphi\rangle$, $\mathcal{B}(P_0, N_a)$, as a function of the population $P_0 = |\langle \varphi | \psi_0 \rangle|^2$, is the same as ref. 13, and independent of N_a .

Next consider a state which can be factorized into two subsets $|\varphi\rangle = \left| \varphi_1^{1\dots k_1} \right\rangle \otimes \left| \varphi_M^{1\dots k_2} \right\rangle$ where $k_1 + k_2 = N_a$. Each $|\varphi_{i=1,2}\rangle$ can be expanded as

$$|\varphi_i\rangle = a_i |\psi_0^{k_i}\rangle + b_i |\psi_1^{k_i}\rangle + \dots \quad (25)$$

where $|\psi_0^{k_i}\rangle$ is the CSS containing k_i atoms, and $|\psi_1^{k_i}\rangle$ is given by equation (23) with N_a replaced by k_i . The populations $P_0 = |\langle \varphi | \psi_0 \rangle|^2$ and $P_1 = |\langle \varphi | \psi_1 \rangle|^2$ are given by

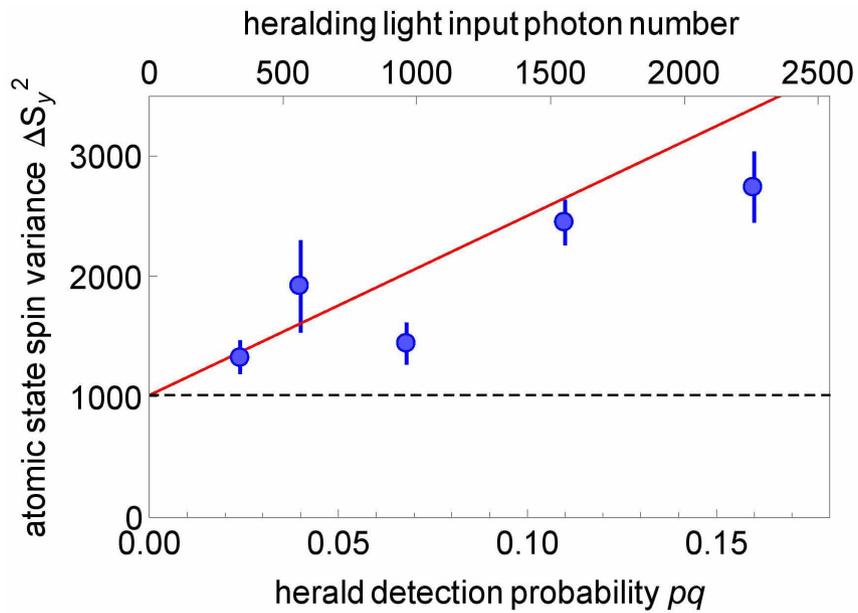
$$\begin{aligned} P_0 &= |a_1|^2 |a_2|^2, \\ P_1 &= \left(\sum_{j=1}^{N_a} \eta_j^2 \right)^{-1} \left| a_2 b_1 \sqrt{\sum_{j=1}^{k_1} \eta_j^2} + a_1 b_2 \sqrt{\sum_{j=k_1+1}^{N_a} \eta_j^2} \right|^2 \end{aligned} \quad (26)$$

The expression for P_1 recovers that of ref. 13 when $\eta_j = 1$. When k_1, k_2 and N_a are large, we take the ensemble averages $\sum_{j=1}^{k_1} \eta_j^2 = k_1 \langle \eta^2 \rangle$, $\sum_{j=k_1+1}^{N_a} \eta_j^2 = k_2 \langle \eta^2 \rangle$ and $\sum_{j=1}^{N_a} \eta_j^2 = N_a \langle \eta^2 \rangle$. Therefore the bound of P_1 in equation (26), $\mathcal{B}(P_0, k_1, k_2, N_a)$, is the same as $\mathcal{B}(P_0, k, N)$ for uniformly coupled atoms when $k_a / N_a = k / N$. This proves that the average fractional entanglement depth for the ensemble containing 3,100 actual non-uniformly coupled atoms is the same as that of 2,100 uniformly coupled effective atoms, thus in our system a minimum of 1,970 out of 2,100 effective atoms or 2,910 out of 3,100 real atoms are mutually entangled.

It might seem as if the addition of $N_w \gg N$ weakly coupled atoms (coupling strength η_w) to the system would increase the entanglement depth without having physical consequences as long as $N_w \eta_w^2 \ll N \eta^2$. However in this case the uncertainty $\Delta \mathcal{N}'$ on the entanglement depth also increases, given by $\frac{\Delta \mathcal{N}'}{N_w} = \frac{\Delta \mathcal{N}}{N} \frac{N \eta^2}{N_w \eta_w^2} \gg \frac{\Delta \mathcal{N}}{N}$, so as to be consistent with the entanglement depth \mathcal{N} before adding the weakly coupled atoms. Atoms that do not change the observed spin distribution have no effect on the entanglement depth.

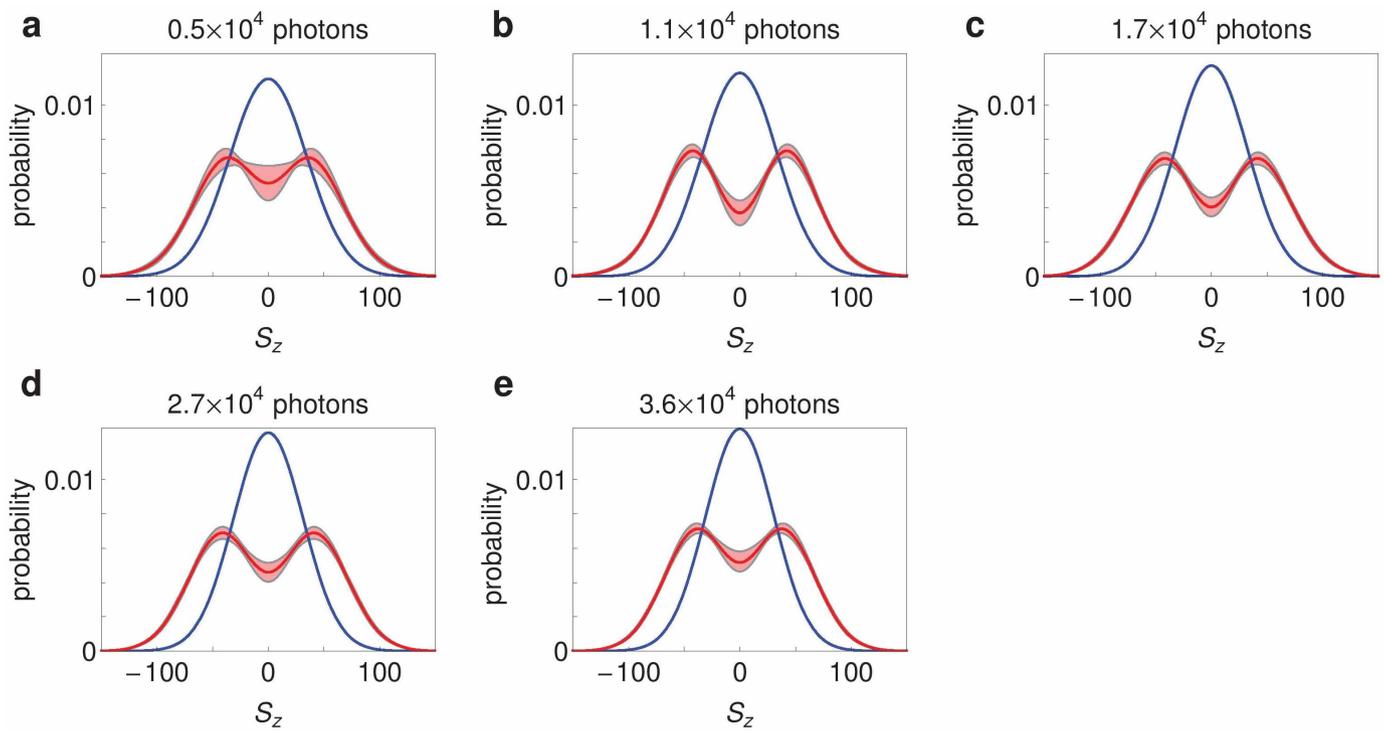
Sample size. No statistical methods were used to predetermine sample size in the above.

31. Tanji-Suzuki, H. *et al.* Interaction between atomic ensembles and optical resonators: classical description. *Adv. At. Mol. Opt. Phys.* **60**, 201–237 (2011).



Extended Data Figure 1 | The measured atomic state spin variance, ΔS_y^2 , as a function of the heralding light photon number and corresponding probability pq of detecting one photon. The solid red line is the prediction for

ΔS_y^2 broadened by the photon shot noise of the heralding light. The dashed black line shows the CSS variance for 2,030 $F = 1$ effective atoms used in this measurement. Error bars, 1 s.d.



Extended Data Figure 2 | Dependence of the reconstructed distribution of collective spin S_z on the measurement photon number. This dependence is illustrated by reconstructed spin distributions for photon numbers 0.5×10^4

(a), 1.1×10^4 (b), 1.7×10^4 (c), 2.7×10^4 (d) and 3.6×10^4 (e). Blue lines correspond to the CSS and red lines correspond to the heralded states. The shaded area indicates an uncertainty of 1 s.d.

Extended Data Table 1 | Resonator parameters

Parameter		$\lambda=780\text{nm}$	$\lambda=852\text{nm}$
Mirror separation	L	26.62(1) mm	
Mirror curvature radius	R	25.04(2) mm	
Free spectral range	$\omega_{FSR}/(2\pi)$	5632.0(2) MHz	
Transverse mode spacing	$\omega_{\lambda}/(2\pi)$	226.3(3) MHz	
Linewidth	$\kappa_{\lambda}/(2\pi)$	1.01(3) MHz	135(2) kHz
Finesse	F_{λ}	$5.6(2) \times 10^3$	$4.2(1) \times 10^4$
Mode waist	w_{λ}	56.9(4) μm	59.5(5) μm
Antinode cooperativity	$\eta_{0,\lambda}$	0.203(7)	1.65(4)

The mode waists are calculated at the position of the atoms. Outside this table, all resonator values refer to the probe wavelength $\lambda = 780$ nm.