SUPPLEMENTARY MATERIAL

Atom-light & atom-field interactions

We define the collective spin operator $\hat{\mathbf{F}} \equiv \sum_i \hat{\mathbf{f}}^{(i)}$, where $\hat{\mathbf{f}}^{(i)}$ is the spin of the i'th atom. The collective spin obeys commutation relations $[\hat{F}_{\rm x}, \hat{F}_{\rm y}] = i\hat{F}_{\rm z}$. Probe pulses are described by the Stokes operator $\boldsymbol{S}$ defined as $\hat{S}_i \equiv \frac{1}{2}(\hat{a}_+^\dagger, \hat{a}_-^\dagger)\sigma_i(\hat{a}_+, \hat{a}_-)^T$, where the σ_i are the Pauli matrices and $\hat{a}_\pm$ are annihilation operators for $\sigma_\pm$ polarization, which obey $[\hat{S}_{\rm x}, \hat{S}_{\rm y}] = i\hat{S}_{\rm z}$ and cyclic permutations. The input pulses are fully $\hat{S}_{\rm x}$-polarized, i.e. with $\langle \hat{S}_{\rm x}\rangle = N_{\rm L}/2$, $\langle \hat{S}_{\rm y}\rangle = \langle \hat{S}_{\rm z}\rangle = 0$ and $\Delta^2 S_i = N_{\rm L}/4$, $i \in \{x, y, z\}$ where $N_{\rm L}$ is the number of photons in the pulse.

The atoms and light interact via an effective hamiltonian

$$\tau \hat{H}_{\rm eff} = G_1 \hat{S}_{\rm z}\hat{F}_{\rm z} + G_2(\hat{S}_{\rm x}\hat{J}_{\rm x} + \hat{S}_{\rm y}\hat{J}_{\rm y} + \hat{S}_0\hat{J}_m/\sqrt{3}) \qquad (1)$$

where G_1 and G_2 are coupling constants describing vector and tensor lights shifts, respectively, and τ is the pulse duration [36, 37]. The operators $\hat{J}_k \equiv \sum_i^{N_A} \hat{j}_i$ where $\hat{j}_x \equiv \hat{f}_x^2 - \hat{f}_y^2$ and $\hat{j}_y \equiv \hat{f}_x\hat{f}_y + \hat{f}_y\hat{f}_x$ describe single-atom Raman coherences, i.e., coherences between states with m_f different by 2, and $\hat{j}_m \equiv (3\hat{f}_z^2 - \hat{\mathbf{f}}^2)/\sqrt{3}$ describes the population difference between the $m_f = 0$ and $m_f = \pm 1$ magnetic sublevels.

The first term in Eq. (1) describes paramagnetic Faraday rotation: it rotates the polarization in the $\hat{S}_{\rm x}$, $\hat{S}_{\rm y}$ plane by an angle $\phi = G_1\hat{F}_{\rm z} \ll 1$, and leaves the atomic state unchanged, so that a measurement of $\hat{S}_{\rm y}^{(\rm out)}/\hat{S}_{\rm x}^{(\rm in)}$ indicates $\hat{F}_{\rm z}$ with a shot-noise-limited sensitivity of $\Delta\hat{F}_{\rm z} = \Delta\hat{S}_{\rm y}/G_1$. Acting alone, this describes a QND measurement of $\hat{F}_{\rm z}$, i.e., with no back-action on $\hat{F}_{\rm z}$. The second term, in contrast, leads to an optical rotation $\hat{S}_{\rm x} \to \hat{S}_{\rm z}$ (due to the birefringence of the atomic sample), and drives a rotation of the atomic spins in the $\hat{F}_{\rm z}$, $\hat{J}_y$ plane (alignment-to-orientation conversion) by an angle $\tan\theta = G_2\hat{S}_{\rm x}/2$ [21, 37]. This leads to a detected output

$$\hat{S}_{\rm y}^{(\rm out)} = \hat{S}_{\rm y}^{(\rm in)} + G_1\hat{S}_{\rm x}^{(\rm in)}(\hat{F}_{\rm z}^{(\rm in)} + \tan\theta \hat{J}_{\rm y}^{(\rm in)}). \qquad (2)$$

For the experiments described here $\theta \simeq 0.3$, and the $\tan\theta$ term can be safely ignored. The contribution of the remaining terms in Eq. (1) is negligible.

The atoms interact with the applied magnetic field via the hamiltonian

$$\hat{H}_{\rm mag} = -\gamma\hat{\mathbf{F}}\cdot\mathbf{B}. \qquad (3)$$

During a single probe-pulse the atomic spins rotate by an angle $\Theta = \gamma B\tau$, where $B = |\mathbf{B}|$. For our parameters $\Theta = 0.08$ radians, so we can neglect the rotation of the spins during the probe pulses.

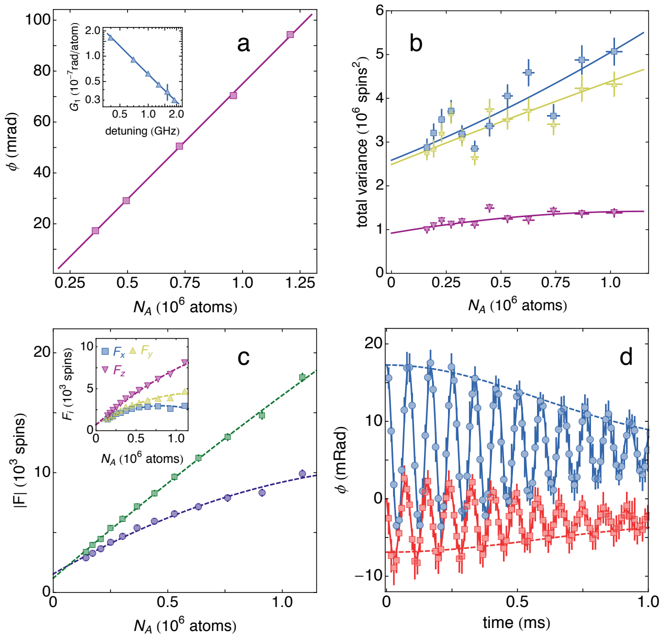

FIG. 1. (Color online) (a) Calibration of G_1 coupling constant. We correlate the observed rotation angle ϕ against an independent measurement of atom number N_A via absorption imaging. Inset: from a fit to G_1 vs. the probe detuning Δ we estimate the effective atom-light interaction area A and tensor light shift G_2. (b) Free induction decay (FID) measurement of the applied magnetic field using atoms as an in-situ vector magnetometer. Blue circles: input $\hat{F}_{\rm z}$-polarized atomic state. Blue circles: input $\hat{F}_{\rm y}$-polarized atomic state. Solid line: fit described by Eq. (4). Dashed line: gaussian envelope of FID signal. (c) Length of spin vector $|\hat{F}|$ detected by the first (blue circles) and second (green squares) measurement. Inset: length of individual spin components $\hat{F}_i$ detected by the first measurement. (d) Noise scaling of total variance $\mathcal{V}_p = {\rm Tr}(\Gamma_p)$ of the first two QND measurements, and conditional variance $\mathcal{V}_{2|1} = {\rm Tr}(\Gamma_{2|1})$. Blue squares: first measurement. Yellow triangles: second measurement. Purple inverted triangles: conditional variance.

Probe calibration

The light-atom coupling constant G_1 is calibrated by correlating the DANM signal with an independent count of the atom number via absorption imaging [12, 21, 41]. In Fig. 1(a) we show the calibration data. We find $G_1 = 9.0 \pm 0.1 \times 10^{-8}$ radians per spin at the detuning $\Delta = -700$ MHz. In the inset of Fig. 1(a) we plot G_1 vs. Δ. We fit this data to find the effective atom-light interaction area A [41], from which we estimate the tensor light shift $G_2 = -4.1^{+0.4}_{-0.5} \times 10^{-9}$ radians per spin at $\Delta = -700$ MHz.

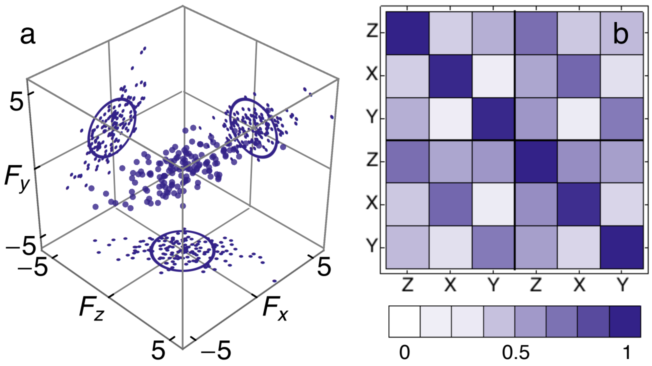

FIG. 2. (Color online) (a) Measured spin distribution (in units of 10^3 spins) of the input TSS following the state preparation procedure described in the main text. (b) Correlation matrix between two consecutive three-component collective spin measurements showing strong correlations between measurements of each spin component $\hat{F}_i$.

Noise scaling & Read-Out Noise

To estimate the atomic noise contribution to the observed total variance $\mathcal{V} = \mathrm{Tr}(\Gamma)$ of the QND measurements we fit the polynomial $\mathcal{V}(N_A) = \mathcal{V}_0 + 2N_A + cN_A^2$ to the measured data, and calculate $\widetilde{\mathcal{V}}_p = \mathcal{V}_p - \mathcal{V}_0$, subtracting the read-out noise $\mathcal{V}_0$ from the measured total variances. The data and resulting fits are shown in Fig. 1(b). The fit to the first (second) measurement yields $\mathcal{V}_0 = 2.59 \pm 0.08 \times 10^6$ $(2.49 \pm 0.08 \times 10^6)$ and $c = 4 \pm 2 \times 10^{-7}$ $(1 \pm 2 \times 10^{-7})$. We fit the polynomial $\mathcal{V}_{2|1}(N_A) = \mathcal{V}_0 + aN_A + cN_A^2$ to the measured conditional variance, giving $\mathcal{V}_0 = 9.2 \pm 0.8 \times 10^5$, $a = 0.9 \pm 3$ and $c = -4 \pm 2 \times 10^{-7}$, indicating the presence of some correlated technical noise in the detection system.

Residual polarization

We observe a small residual atomic polarization due to atoms that are not entangled in the mascroscopic singlet state. In Fig. 1(c) we plot the length of the spin vector $|\hat{F}|$ detected in the two measurements. With $N_A = 1.1 \times 10^6$ atoms, we observe a maximum $|F| = 13.3 \pm 0.2 \times 10^3$ $(18.3 \pm 0.2 \times 10^3)$ spins for the first (second) measurement, i.e. a residual polarization $|\hat{F}|/(fN_A) = 1.66 \pm 0.02 \times 10^{-3}$. In principle with these values we could achieve 20dB of spin squeezing, entangling up to 99% of the atoms in a macroscopic singlet, before back-action due to the spin uncertainty relations limits the achievable squeezing. This residual polarization could be removed by adding a feedback loop to the measurement sequence [40], which would produce an unconditionally squeezed macroscopic singlet centered at the origin.

Magnetic field calibration

We measure the applied magnetic field using the atoms as an in-situ vector magnetometer. Our technique is described in detail in Ref. [39]. We polarize the atoms via optically pumping along first $\hat{F}_z$ and then $\hat{F}_y$, and observe the free induction decay (FID) of the resulting Larmor precession using the Faraday rotation probe. We model density distribution along the length of the trap with a gaussian $A\exp(-(z-z_0)^2/2\sigma^2)$, with an RMS width $\sigma = 2.68 \pm 0.3$ mm. A typical density profile and gaussian fit is shown in Fig. 1(d). This leads to observed signals for the two input states

$$\theta(t) = \frac{G_1}{B^2}\begin{cases} \left(B_z^2 + \left(B_x^2 + B_y^2\right)\cos\omega\exp\left(-t^2/T_2^2\right)\right) F_z(0) \\ \left(B_y B_z\left(1 - \cos\omega\exp\left(-t^2/T_2^2\right)\right) + B_x B\sin\omega\exp\left(-t^2/T_2^2\right)\right) F_y(0) \end{cases} \tag{4}$$

where $\omega = \gamma B t$, $B = |\mathbf{B}|$, and $\gamma = \mu_B g_f/\hbar$ is the atomic gyromagnetic ratio. By fitting theses signals, we extract the vector field $\mathbf{B}$ and the FID transverse relaxation time $T_2 = 1/(\sigma\gamma\partial B/\partial z)$. For the data shown we find $B_x = 9.6 \pm 0.4$ mG, $B_y = 9.7 \pm 0.4$ mG, $B_z = 9.9 \pm 0.1$ mG and $T_2 = 745 \pm 45$ μs.

Input state

In Fig. 2(a) we plot the spin distribution $\mathbf{F}^{(1)}$ of the collective spin of a sample with $N_A = 1.4 \times 10^6$ atoms measured by the first three probe pulses. We measure an initial spin covariance matrix of

$$\Gamma_1 = \begin{pmatrix} 1.90 & 1.10 & 1.10 \\ 1.10 & 1.40 & 0.81 \\ 1.10 & 0.81 & 1.30 \end{pmatrix} \times 10^6\,\mathrm{spins}^2. \tag{5}$$

For comparison, an ideal TSS would have $\Gamma = \mathrm{diag}(0.93, 0.93, 0.93) \times 10^6$ spins^2 with the same number of atoms. The larger measured variances, and non-zero covariances, in Γ_1 indicate the presence of atomic technical noise due to imperfect state preparation and shot-to-shot fluctuations in the atom number and applied magnetic field.

Measurement correlations

In Fig. 2(b) we plot the correlations $\rho_{ij} \equiv \mathrm{cov}(\hat{F}_i, \hat{F}_j)/\Delta\hat{F}_i\Delta\hat{F}_j$ between the first six QND measurements. The off-diagonal elements indicate that successive measurements of the same spin component $\hat{F}_k$ are well correlated. This allows us to predict the outcome of the second measurements $F_k^{(2)}$ with a reduced conditional uncertainty. The residual correlation between measurements of different spin components is due to correlated technical noise in the atomic state preparation, and in the detection system.