Article

Spin squeezing of 10^{11} atoms by prediction and retrodiction measurements

https://doi.org/10.1038/s41586-020-2243-7

Received: 11 July 2019

Accepted: 26 February 2020

Published online: 13 May 2020

Check for updates

Han Bao[1], Junlei Duan[1], Shenchao Jin[1], Xingda Lu[1], Pengxiong Li[1], Weizhi Qu[1], Mingfeng Wang[1,2], Irina Novikova[3], Eugeniy E. Mikhailov[3], Kai-Feng Zhao[4], Klaus Mølmer[5]✉, Heng Shen[6,7]✉ & Yanhong Xiao[1,6]✉

The measurement sensitivity of quantum probes using N uncorrelated particles is restricted by the standard quantum limit[1], which is proportional to $1/\sqrt{N}$. This limit, however, can be overcome by exploiting quantum entangled states, such as spin-squeezed states[2]. Here we report the measurement-based generation of a quantum state that exceeds the standard quantum limit for probing the collective spin of 10^{11} rubidium atoms contained in a macroscopic vapour cell. The state is prepared and verified by sequences of stroboscopic quantum non-demolition (QND) measurements. We then apply the theory of past quantum states[3,4] to obtain spin state information from the outcomes of both earlier and later QND measurements. Rather than establishing a physically squeezed state in the laboratory, the past quantum state represents the combined system information from these prediction and retrodiction measurements. This information is equivalent to a noise reduction of 5.6 decibels and a metrologically relevant squeezing of 4.5 decibels relative to the coherent spin state. The past quantum state yields tighter constraints on the spin component than those obtained by conventional QND measurements. Our measurement uses 1,000 times more atoms than previous squeezing experiments[5–10], with a corresponding angular variance of the squeezed collective spin of 4.6×10^{-13} radians squared. Although this work is rooted in the foundational theory of quantum measurements, it may find practical use in quantum metrology and quantum parameter estimation, as we demonstrate by applying our protocol to quantum enhanced atomic magnetometry.

Measurements constitute the foundations of physical science. The aim of high-precision metrology is to reduce uncertainties and draw as accurate conclusions as possible from measurement data[1]. Quantum systems are described by wave functions or density matrices, which yield probabilistic measurement outcomes. For a continuously monitored system, the well established theory of quantum trajectories employs stochastic master equations to describe the evolution with time of the density matrix $\rho(t)$, which is governed by the system Hamiltonian, dissipation, and effects associated with the measurements[2]. For Gaussian states and operations, the theory is simplified to equations for mean values and covariances, equivalent to classical Kalman filter theory[11].

By knowing the value of $\rho(t)$, we can predict the outcome of a subsequent measurement on the system, and if QND probing has led to a state with reduced uncertainty on a specific observable, we may thus make an improved prediction of the subsequent measurement. Also, later measurements will have outcomes correlated with the present and previous ones; in the same way that daily life experience teaches us about past events and facts, one may ask if it is possible in a quantum experiment to obtain more knowledge about a quantum state by using both earlier and later observations on a system. Such retrodiction was initially introduced in the context of pre- and post-selection under projective measurements[12] and in the theory of weak value measurements[13], whereas the idea of a complete description of a quantum system at any time during a sequence of measurements[14] has found a general dynamical formulation in the so-called past quantum state (PQS)[3,4]. The PQS provides the probability distribution of the outcome of any general measurement on a quantum system at time t, conditioned on our knowledge about the system that is obtained by measurements performed both before and after t. The PQS has been demonstrated to yield better predictions than the usual conditional density matrix in trajectory simulations of the photon number evolution in a cavity[15], the excitation and emission dynamics of a superconducting qubit[16] and the motional state of a mechanical oscillator[17].

Here we show that the PQS elements of the quantum trajectory description could further improve already precise measurements with vapour cells for magnetometry[18–20], fundamental symmetry tests[21,22] and gravitational-wave detection[23]. In particular, we show that for a metrologically relevant macroscopic atomic spin system, the standard quantum limit determined by the atom projection noise can be surpassed by conditioning the measurement result on previous and

[1]Department of Physics, State Key Laboratory of Surface Physics and Key Laboratory of Micro and Nano Photonic Structures, Ministry of Education, Fudan University, Shanghai, China. [2]Department of Physics, Wenzhou University, Zhejiang, China. [3]Department of Physics, College of William and Mary, Williamsburg, VA, USA. [4]Applied Ion Beam Physics Laboratory, Key Laboratory of the Ministry of Education, and Institute of Modern Physics, Fudan University, Shanghai, China. [5]Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark. [6]State Key Laboratory of Quantum Optics and Quantum Optics Devices, Shanxi University, Taiyuan, China. [7]Clarendon Laboratory, University of Oxford, Oxford, UK. ✉e-mail: moelmer@phys.au.dk; heng.shen@physics.ox.ac.uk; yxiao@fudan.edu.cn

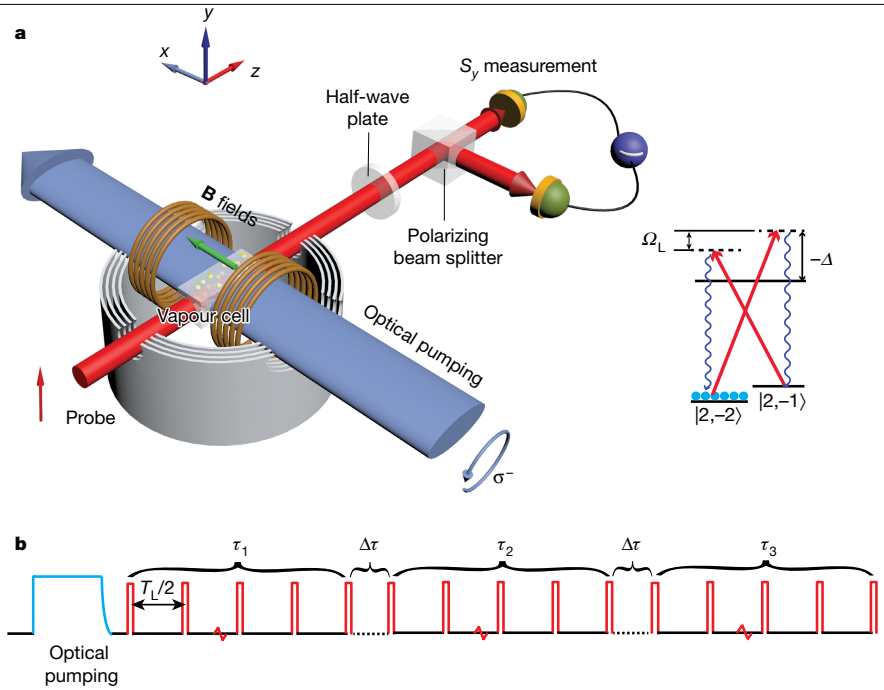


Fig. 1 | Experimental setup. a, Schematic of the setup. A paraffin-coated 20 mm × 7 mm × 7 mm rectangular vapour cell at 53 °C resides inside a four-layer magnetic shielding to screen the ambient magnetic field. The CSS is created by optical pumping, with a pump laser tuned to the Rb D1 transition $5S_{1/2}$ $F = 2 \rightarrow 5P_{1/2} F' = 2$ and a repump laser stabilized to the Rb D2 transition $5S_{1/2}$ $F = 1 \rightarrow 5P_{3/2} F' = 2$, both with σ⁻ circular polarization along the x direction. A magnetic field (along the x direction) of 0.71 G is applied to induce a ground-state Zeeman splitting (that is, a Larmor frequency of $\Omega_L \approx 2\pi \times 500$ kHz) and to hold the collective spin. A linearly polarized laser beam, which is blue-detuned by 2.1 GHz from the $5S_{1/2}, F = 2 \rightarrow 5P_{3/2}, F' = 3$ transition of the D2 line and propagates in the z direction, probes the quantum fluctuations of the spin. The Stokes component S_y is measured using a balanced polarimetry scheme and detected at the Larmor frequency Ω_L by a lock-in amplifier. **b**, Pulse sequence. The pump lasers prepare the atoms in the CSS and are then turned off adiabatically (see Methods). They are followed by the stroboscopic probe pulses, which are spaced by half the Larmor period, $T_L/2$. The first part (pulse duration τ_1) of the probe, called squeezing pulse, creates entanglement between S_y and J_z. J_z is squeezed through the detection of S_y, and the second part (pulse duration τ_2), called the verification pulse, verifies the squeezing. The state is further probed (squeezed) for a duration of τ_3. The time $\Delta\tau = 0.3$ ms between the three probe periods is to avoid interference from the lock-in amplifier.

later measurements on the system. The incorporation of later measurements supplements the well established measurement-based entanglement generation protocol[5–10,24] and provides further information about measurement outcomes at intermediate times. The combined information from prior and posterior measurements on the collective spin of $N_{at} = 1.87 \times 10^{11}$ hot atoms in a vapour cell is equivalent to a noise reduction of 5.6 dB and a spin squeezing of 4.5 dB using the Wineland criterion, and corresponds to an angular spin variance of 4.6×10^{-13} rad^2. In the following, we refer to this noise reduction as 'squeezing', but we recall that we are referring to the squeezing of an outcome probability distribution, not of a physical state.

Consider a collective atomic spin given by the sum of the total angular momenta of individual atoms, $\hat{J}_i = \sum_k \hat{j}_i^k$, with $i = x, y, z$. The macroscopic spin orientation J_x is along the applied bias magnetic field **B**, and the collective spin components $\hat{J}_{y,z}$ oscillate in the laboratory frame at the Larmor frequency Ω_L. In the rotating frame, they obey the commutation relation $[\hat{J}_y, \hat{J}_z] = iJ_x$ ($\hbar = 1$; $\hbar$, reduced Planck constant).

The QND measurement of the collective atomic spin is realized by coupling the atomic ensemble to a light beam with the off-resonant Faraday interaction described in equation (1), such that a direct measurement on the transmitted field provides information about the atomic spin[10,25]:

$$\hat{H}_{int} = \frac{\sqrt{2}\kappa}{\sqrt{N_{ph}N_{at}}} \hat{J}_z \hat{S}_z \quad (1)$$

Here N_{ph} is the number of photons in a pulse of duration τ and N_{at} is the atom number. $\hat{S}_z$ is the Stokes operator of the probe light, relating to the photon number difference between σ⁺ and σ⁻ polarization. The coupling constant $\kappa^2 \propto d_0\eta \propto N_{ph}N_{at}$ characterizes the measurement strength in QND detection, with d_0 the resonant optical depth and η the atomic depumping rate causing decay of the collective spin.

We use a ^{87}Rb ensemble of 10^{11} atoms contained in a paraffin-coated vapour cell[26], as shown in Fig. 1. The coating provides a spin-protecting environment, enabling high-performance optical pumping and allowing the long spin coherence time to reduce the information loss due to decoherence. The atoms are initially prepared in the state $5S_{1/2}$ $|F = 2, m_F = -2\rangle$ (defined by the quantization axis x) by optical pumping propagating along the x direction parallel to the **B** field. We achieve up to 97.9% polarization of the spins, resulting in a 6% increase of the measured variance compared to the fully polarized coherent spin state (CSS). The quantum fluctuations of the spin are probed by a linearly polarized off-resonant D2 laser beam propagating in the z direction. The projection noise limit is calibrated by measuring the noise of the collective spin of the unpolarized sample, which is 1.25 times that of the CSS (see Methods). The QND measurement of the spin component $\hat{J}_z$ is achieved by implementing the stroboscopic quantum back-action evasion protocol[10] (that is, modulating the measurement intensity at twice the Larmor frequency with an optimal duty factor of 14%).

To describe the atomic system and its collective spin fluctuations during the optical probing, we apply the general quantum theory of measurements. To account for a quantum state conditioned on both prior and posterior probing of a quantum system, we consider a system subject to three subsequent measurement processes. Each measurement (i) is described as a general positive-operator-valued

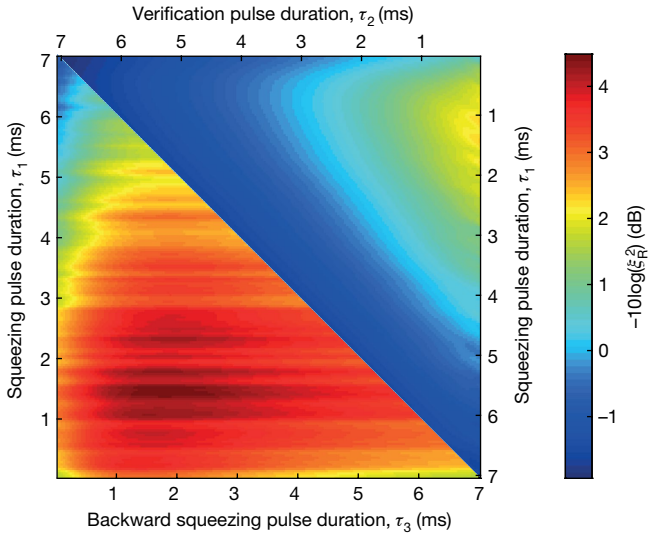


Fig. 2 | Experiment results. The lower diagonal shows the degree of spin squeezing (see colour bar) of the three-pulse scheme for various time durations of the first and third pulses. The duration of the second probe pulse is 0.037 ms. Better squeezing is observed for a shorter verification pulse τ_2, which minimizes the disturbance of the state prepared during the first pulse. The squeezing reaches its maximum value of 4.5 dB at $\tau_1 = 1.4$ ms and $\tau_3 = 1.7$ ms, as an optimal balance between the increased atom–light coupling strength with the higher photon number and the spin decoherence due to spontaneous emission. The upper diagonal shows the spin squeezing detected when using the traditional squeezing and verification two-pulse scheme as a function of τ_1 and τ_2. The best squeezing here is 2.3 dB. The probe laser has an average power of 1.18 mW in both experiments. ξ_R^2 is the squeezing parameter according to the Wineland criterion (see Supplementary Information).

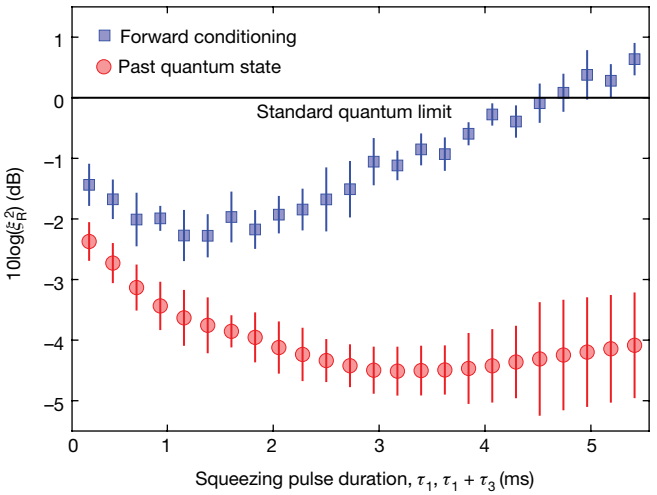


Fig. 3 | Squeezing versus total squeezing pulse duration in two- and three-pulse schemes. The horizontal axis shows τ_1 for the two-pulse scheme (forward conditioning) and $\tau_1 + \tau_3$ for the three-pulse scheme (PQS protocol). τ_2 is 0.037 ms for both curves. The attainable squeezing for the three-pulse scheme is greater and has a better long-time behaviour than the two-pulse scheme. The error bars (1 s.d.) are derived from 10 identical experiments, each consisting of 10,000 repetitions of the pulse sequence shown in Fig. 1b.

measurement (POVM) with a set of operators $\{\hat{\Omega}_m^{(i)}\}$ associated with the measurement outcome m and fulfilling $\sum_m \hat{\Omega}_m^{(i)\dagger}\hat{\Omega}_m^{(i)} = \hat{\mathbb{I}}$, where $\hat{\mathbb{I}}$ is the identity matrix. For a system represented by the density matrix ρ at the time of a measurement, the probability of measuring outcome m is

$$\Pr^{(i)}(m) = \mathrm{tr}\left(\hat{\Omega}_m^{(i)}\rho\hat{\Omega}_m^{(i)\dagger}\right) \tag{2}$$

and the resulting conditional state reads

$$\rho|_m = \frac{\hat{\Omega}_m^{(i)}\rho\hat{\Omega}_m^{(i)\dagger}}{\Pr^{(i)}(m)} \tag{3}$$

Assuming no further dynamics between the measurements, we can evaluate the joint probability that three subsequent measurements, described by $\{\hat{\Omega}_m^{(i)}\}$, yield outcomes m_1, m_2 and m_3 as

$$\Pr(m_1, m_2, m_3) = \mathrm{tr}\left(\hat{\Omega}_{m_3}^{(3)}\hat{\Omega}_{m_2}^{(2)}\hat{\Omega}_{m_1}^{(1)}\rho\hat{\Omega}_{m_1}^{(1)\dagger}\hat{\Omega}_{m_2}^{(2)\dagger}\hat{\Omega}_{m_3}^{(3)\dagger}\right) \tag{4}$$

This equation can be factored into: (i) the probability of obtaining the first outcome, m_1, (ii) the probability of obtaining outcome m_2 in the state conditioned on the first outcome, and (iii) the probability of obtaining outcome m_3 in the state conditioned on the first two outcomes. This is equivalent to the conventional evolution of quantum trajectories, where the quantum state—and hence the probability of a measurement outcome—depends on previous measurements. However, the joint probability distribution (4) also permits evaluation of the probability of, for example, the second measurement, conditioned on the outcome of the first and the last one

$$\Pr(m_2|m_1, m_3) = \Pr(m_1, m_2, m_3)/\sum_{m_2'} \Pr(m_1, m_2', m_3) \tag{5}$$

where m_1 and m_3 are fixed to the observed values and the denominator is merely a normalization factor.

Using equation (4) and the cyclic permutation property of the trace, we can write this probability as[3]

$$\Pr_{\mathrm{p}}(m_2, t) = \frac{\mathrm{tr}\left(\hat{\Omega}_{m_2}^{(2)}\rho|_{m_1}\hat{\Omega}_{m_2}^{(2)\dagger}E|_{m_3}\right)}{\sum_{m'}\mathrm{tr}\left(\hat{\Omega}_{m'}^{(2)}\rho|_{m_1}\hat{\Omega}_{m'}^{(2)\dagger}E|_{m_3}\right)} \tag{6}$$

where $\rho|_{m_1}$ is the state conditioned on the first measurement (see equation (3)) and $E|_{m_3} = \hat{\Omega}_{m_3}^{(3)\dagger}\hat{\Omega}_{m_3}^{(3)}$.

We observe that the conventional expression for the outcome probabilities in equation (2) depending only on the density matrix $\rho|_{m_1}$, conditioned on the prior evolution, is supplemented with the operator $E|_{m_3}$, which depends only on the later measurement outcomes. The same formalism applies to cases with continuous sequences of measurements occurring simultaneously with Hamiltonian and dissipative evolution. Examples of how the operators $\rho(t)$ and $E(t)$ evolve to time t from the initial and final time, respectively, are given in ref. [3].

The specific form of the POVM operators and their action on the quantum states in our experiments can be derived explicitly in a simplified form because our system dynamics is restricted to Gaussian states. This follows from the Holstein–Primakoff transformation that maps the spin operators perpendicular to the large mean spin on the canonical position and momentum operators, $\hat{x}_{\mathrm{A}} = \hat{J}_y/\sqrt{|\langle J_x\rangle|}$ and $\hat{p}_{\mathrm{A}} = \hat{J}_z/\sqrt{|\langle J_x\rangle|}$. The CSS with all atoms in $|F, m_F = -F\rangle$, characterized by $\mathrm{Var}(\hat{J}_y) = \mathrm{Var}(\hat{J}_z) = J_x/2 = N_{\mathrm{at}}F/2$, is equivalent to the Gaussian ground state of a harmonic oscillator, and an excitation with the ladder operator $\hat{b}^\dagger$ corresponds to a quantum of excitation distributed symmetrically among all atoms[10]. Similar canonical operators, $\hat{x}_{\mathrm{L}} = \hat{S}_y/\sqrt{|\langle S_x\rangle|}$ and $\hat{p}_{\mathrm{L}} = \hat{S}_z/\sqrt{|\langle S_x\rangle|}$, and Gaussian states describe the probe field degrees of freedom (see Supplementary Information).

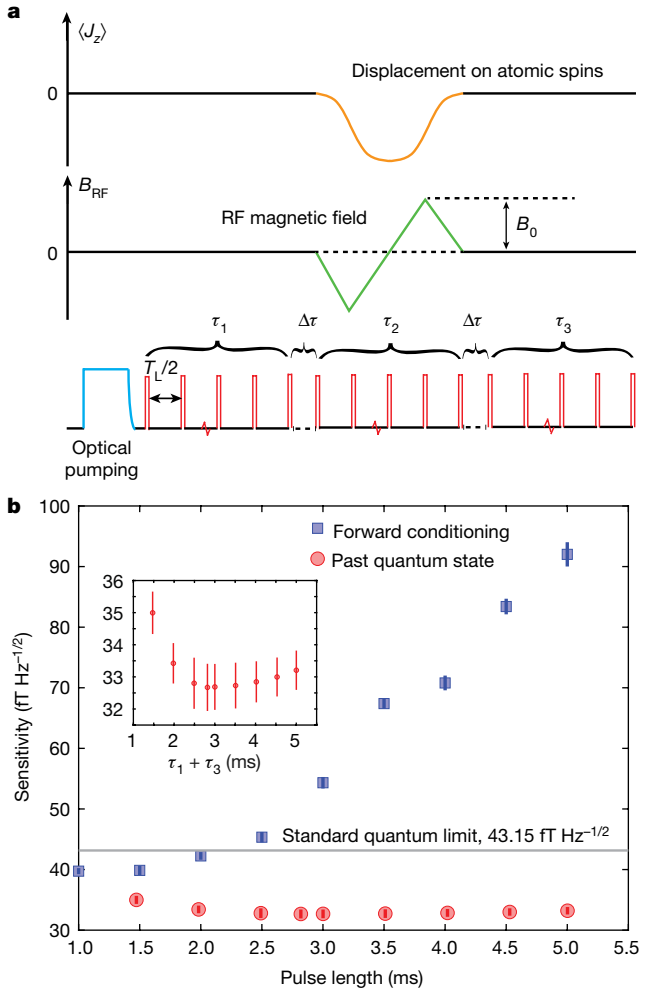


Fig. 4 | PQS-enhanced magnetometry. a, Pulse sequence. An RF magnetic field pulse oscillating at the Larmor frequency is switched on during the second probe sequence τ_2 in the direction orthogonal to the static **B** field. The amplitude of the RF field, B_{RF}, is modulated as a zero-area two-triangle profile. **b**, Sensitivity of the two- and three-pulse schemes as a function of the duration of the squeezing pulses. The horizontal axis shows τ_1 for the two-pulse scheme (forward conditioning) and $\tau_1+\tau_3$ for the three-pulse scheme (PQS protocol). τ_2 = 1 ms for both curves. Similar to the squeezing results in Fig. 3, the sensitivity of the three-pulse scheme is superior to that of the two-pulse scheme and has a better long-time behaviour. The error bars (1 s.d.) are derived from five identical experiments, each consisting of 2,000 repetitions. The grey line represents the sensitivity imposed by the standard quantum limit in our system. The inset magnifies the sensitivity scale for the PQS results.

In Supplementary Information we show that the measurement operator $\hat{\Omega}_m$ in equation (2) acting on the atomic state upon readout of the value m of the field quadrature $\hat{x}_L$ is given by $\hat{\Omega}_m = \int \psi_{\hat{x}_L}(m-\kappa a)|a\rangle\langle a|_{\hat{p}_A} da$, where $\psi_{\hat{x}_L}(m) = \frac{1}{\pi^{1/4}}\exp(-\frac{m^2}{2})$ characterizes the quadrature distribution of the input coherent state of the probe laser beam.

For two successive QND measurements with coupling strengths κ_1 and κ_2, the POVM formalism shows that the second outcome is governed by a Gaussian distribution with a mean value conditioned on the first outcome (see Supplementary Information)

$$\Pr(m_2|m_1) = \frac{1}{\sqrt{\pi}\sigma}\exp\left[-\frac{\left(m_2 - \frac{\kappa_2 m_1 \kappa_1}{1+\kappa_1^2}\right)^2}{2\sigma^2}\right] \quad (7)$$

Here the variance $\sigma^2 = \frac{1}{2} + \frac{1}{2}\frac{\kappa_2^2}{1+\kappa_1^2}$ is composed of a contribution of 1/2 from the light shot noise and a contribution from the atomic spin, which is reduced by the conditional spin squeezing by the first measurement with strength κ_1.

If the spin oscillator is further probed by a third QND pulse with coupling strength κ_3 and measurement outcome m_3, the conditional probability for the outcome of the middle measurement is obtained as

$$\Pr(m_2|m_1, m_3) = \frac{1}{\sqrt{\pi}\sigma_p}\exp\left\{-\frac{\left[m_2 - \frac{\kappa_2(m_1\kappa_1 + m_3\kappa_3)}{1+\kappa_1^2+\kappa_3^2}\right]^2}{2\sigma_p^2}\right\} \quad (8)$$

The past probability yields a Gaussian distribution with variance $\sigma_p^2 = \frac{1}{2} + \frac{1}{2}\frac{\kappa_2^2}{1+\kappa_1^2+\kappa_3^2}$. The reduction by $1+\kappa_1^2+\kappa_3^2$ shows that the incorporation of the information from later measurements has a similar effect as increasing the coupling strength of the first probing from κ_1^2 to $\kappa_1^2+\kappa_3^2$.

Experimentally, for the normal two-pulse scheme of forward-conditioning QND, we achieve the best spin squeezing of 2.3 ± 0.2 dB (Fig. 2, upper diagonal) according to the Wineland criterion[27] for τ_1 = 1.23 ms and a conditional noise reduction of about 4.3 dB, in good agreement with the theoretical prediction (see Supplementary Information). In stark contrast, as predicted by equation (8), for the three-pulse scheme that extracts the full information from the full measurement record using the PQS, we observe an improved conditional noise reduction of about 5.6 dB and spin squeezing of 4.5 ± 0.40 dB (Fig. 2, lower diagonal) according to the Wineland criterion for τ_1 = 1.4 ms and τ_3 = 1.7 ms.

The main reason that the probing before and after the verification pulse sequence yields stronger squeezing than an initial longer probing sequence is the decoherence of the spins. First, owing to decoherence, the spin squeezing is gradually lost, and measurement results obtained during the early stages of the squeezing (first) pulse sequence will be less correlated with the spin ensemble at the time of the verification (second) pulse. If we instead postpone these measurements to occur in the third pulse sequence immediately after the verification pulse, the correlations will be stronger, that is, the conditional variance will be lower. Secondly, the large average spin component J_x is reduced during probing, weakening the squeezing according to the Wineland criterion. With retrodicted squeezing, the spin variance is measured relative to the mean spin at the time of the verification pulse, which has not yet suffered the reduction due to the third pulse sequence.

As shown in Fig. 3, even if we keep the total duration of the squeezing equal for both schemes, the squeezing that is attainable when using the information obtained both before and after the second pulse is better than that achieved when using only the information before the second pulse.

Although retrodiction is not a state preparation method for spin squeezing, it provides metrological advantage, as demonstrated by radio-frequency (RF) magnetometry (Fig. 4). The pulse sequence is the same as that shown in Fig. 1, but a magnetic field pulse is applied during the second pulse τ_2 to generate a temporary offset of the spin component J_z. For our proof-of-principle demonstration, this field oscillates at the Larmor frequency and follows a time-varying profile with a known shape but unknown amplitude. The procedure is outlined in Methods and summarized as follows: the value of the atomic observable p_A is retrodicted in each experiment to a certain conditional mean value and a definite variance before and after the applied magnetic field. The m_2 readout signal thus reports directly a noisy estimate of the applied field pulse, as demonstrated by the results presented in Fig. 4b. We find that the PQS protocol gives a better sensitivity than the forward conditioning protocol for the same total duration of the full pulse sequence. Notably, as expected, the sensitivity of the

three-pulse scheme experiences no substantial influence of the spin decoherence that occurs during the last detection pulse. Given $\tau_2 = 1$ ms, the best sensitivity achieved via the PQS protocol is $B_{RF}\sqrt{\tau_2}/\mathrm{SNR} = 32.67 \pm 0.73$ fT $\mathrm{Hz}^{-1/2}$, where the signal-to-noise ratio SNR is the ratio of the mean to the standard deviation of the data obtained for $B_0 \approx 1$ pT (Fig. 4a) applied during τ_2 (ref. [19]). We note that our analysis is simplified here by the QND character of the probing, whereas applications in which the non-unitary measurement back-action is interspersed with unitary rotation of the spin ellipse[28,29] can also be handled by the more complete PQS analysis with Gaussian states[4].

This work introduces a higher limit on the size (in terms of the number of spins) that a physical system can have while still being subjected to measurements at the quantum limit. Further improvement of the squeezing is possible by realizing a multiple light-pass scheme[30,31] to enhance the coupling strength and incorporate unconditional spin squeezing. Atoms constitute ideal high-sensitivity probes for a number of physical phenomena[21,22], and our retrodiction procedure may affect the practical applications of quantum sensors. In particular, the retrodicted evolution of physical systems may offer insight and allow precision estimation of time-dependent perturbations[32] that are applicable, for example, to force sensing with mechanical oscillators[23,33].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2243-7.

1. Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. Photon.* **5**, 222–229 (2011).
2. Wiseman. H. M & Milburn, G. *Quantum Measurement and Control* (Cambridge Univ. Press, 2010).
3. Gammelmark, S., Julsgaard, B. & Mølmer, K. Past quantum states of a monitored system. *Phys. Rev. Lett.* **111**, 160401 (2013).
4. Zhang, J. & Mølmer, K. Prediction and retrodiction with continuously monitored Gaussian states. *Phys. Rev. A* **96**, 062131 (2017).
5. Hosten, O., Engelsen, N. J., Krishnakumar, R. & Kasevich, M. A. Measurement noise 100 times lower than the quantum-projection limit using entangled atoms. *Nature* **529**, 505–508 (2016).
6. Cox, K. C., Greve, G. P., Weiner, J. M. & Thompson, J. K. Deterministic squeezed states with collective measurements and feedback. *Phys. Rev. Lett.* **116**, 093602 (2016).
7. Appel, J. et al. Mesoscopic atomic entanglement for precision measurements beyond the standard quantum limit. *Proc. Natl Acad. Sci. USA* **106**, 10960–10965 (2009).
8. 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).
9. Chaudhury, S. et al. Quantum control of the hyperfine spin of a Cs atom ensemble. *Phys. Rev. Lett.* **99**, 163002 (2007).
10. Vasilakis, G. et al. Generation of a squeezed state of an oscillator by stroboscopic back-action-evading measurement. *Nat. Phys.* **11**, 389–392 (2015).
11. Mølmer, K. & Madsen, L. B. Estimation of a classical parameter with Gaussian probes: magnetometry with collective atomic spins. *Phys. Rev. A* **70**, 052102 (2004).
12. Aharonov, Y., Albert, D. Z. & Vaidman, L. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.* **60**, 1351 (1988).
13. Aharonov, Y. & Vaidman, L. Properties of a quantum system during the time interval between two measurements. *Phys. Rev. A* **41**, 11–20 (1990).
14. Aharonov, Y. & Vaidman, L. Complete description of a quantum system at a given time. *J. Phys. A* **24**, 2315–2328 (1991).
15. Rybarczyk, T. et al. Forward-backward analysis of the photon-number evolution in a cavity. *Phys. Rev. A* **91**, 062116 (2015).
16. Tan, D., Weber, S. J., Siddiqi, I., Mølmer, K. & Murch, K. W. Prediction and retrodiction for a continuously monitored superconducting qubit. *Phys. Rev. Lett.* **114**, 090403 (2015).
17. Rossi, M., Mason, D., Chen, J. & Schliesser, A. Observing and verifying the quantum trajectory of a mechanical resonator. *Phys. Rev. Lett.* **123**, 163601 (2019).
18. Shah, V., Vasilakis, G. & Romalis, M. V. High bandwidth atomic magnetometery with continuous quantum nondemolition measurements. *Phys. Rev. Lett.* **104**, 013601 (2010).
19. Wasilewski, W. et al. Quantum noise limited and entanglement-assisted magnetometry. *Phys. Rev. Lett.* **104**, 133601 (2010).
20. Martin Ciurana, F., Colangelo, G., Slodička, L., Sewell, R. J. & Mitchell, M. W. Entanglement-enhanced radio-frequency field detection and waveform sensing. *Phys. Rev. Lett.* **119**, 043603 (2017).
21. Smiciklas, M., Brown, J. M., Cheuk, L. W., Smullin, S. J. & Romalis, M. V. New test of local Lorentz invariance using a ^{21}Ne-Rb-K comagnetometer. *Phys. Rev. Lett.* **107**, 171604 (2011).
22. Bear, D., Stoner, R. E., Walsworth, R. L., Kostelecky, V. A. & Lane, C. D. Limit on Lorentz and CPT violation of the neutron using a two-species noble-gas maser. *Phys. Rev. Lett.* **85**, 5038 (2000).
23. Khalili, F. Ya. & Polzik, E. S. Overcoming the standard quantum limit in gravitational wave detectors using spin systems with a negative effective mass. *Phys. Rev. Lett.* **121**, 031101 (2018).
24. Kong, J., Jiménez-Martínez, J., Troullinou, C., Lucivero, V. G. & Mitchell, M. W. Measurement-induced nonlocal entanglement in a hot, strongly-interacting atomic system. Preprint at http://arXiv.org/quant-ph/1804.07818 (2018).
25. Hammerer, K., Sørensen, A. S. & Polzik, E. S. Quantum interface between light and atomic ensembles. *Rev. Mod. Phys.* **82**, 1041–1093 (2010).
26. Balabas, M. V., Karaulanov, T., Ledbetter, M. P. & Budker, D. Polarized alkali-metal vapor with minute-long transverse spin-relaxation time. *Phys. Rev. Lett.* **105**, 070801 (2010).
27. Wineland, D. J., Bollinger, J. J., Itano, W. M. & Heinzen, D. J. Squeezed atomic states and projection noise in spectroscopy. *Phys. Rev. A* **50**, 67 (1994).
28. Borregaard, J. & Sorensen, A. S. Near-Heisenberg-limited atomic clocks in the presence of decoherence. *Phys. Rev. Lett.* **111**, 090801 (2013).
29. Braverman, B. et al. Near-unitary spin squeezing in Yb-171. *Phys. Rev. Lett.* **122**, 223203 (2019).
30. Wang, M. F. et al. Two-axis-twisting spin squeezing by multipass quantum erasure. *Phys. Rev. A* **96**, 013823 (2017).
31. Takeuchi, M. et al. Spin squeezing via one-axis twisting with coherent light. *Phys. Rev. Lett.* **94**, 023003 (2005).
32. Tsang, M. Time-symmetric quantum theory of smoothing. *Phys. Rev. Lett.* **102**, 250403 (2009).
33. Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Experimental setup and characterization

From a technical perspective, the experimental realization of large-scale spin squeezing is challenging because classical noise amplitudes typically scale as the atom number N_{at} and dominate over that of the atom projection noise that is proportional to $\sqrt{N_{at}}$. Also, for large atomic ensembles it is difficult to achieve a uniform atom–light coupling across the entire ensemble, which is required for state preparation, manipulation and detection.

Meanwhile, strict orthogonality is required between the polarized spin and the wave vector of the probe field to avoid influence of the large polarized spin component in the y–z plane on the quantum noise measurement. In the alignment optimization, we used the intensity-modulated pump field as in a Bell–Bloom magnetometer configuration[34], and we found that an adiabatic turn-off of the pump pulse was necessary to minimize classical noise (see below).

Preparation and characterization of the atomic state. A d.c. magnetic field in the x direction creates the Zeeman splitting. Circularly polarized optical pumping and repumping beams along the x direction prepare the highly oriented spin states, which is crucial for the interface between light and atoms. As shown in the inset of Extended Data Fig. 1b, the pump and repump lasers are tuned to the ^{87}Rb D1 and D2 transitions, respectively.

Number of atoms in the vapour cell. To determine the number of atoms in the vapour cell, a Faraday rotation measurement is employed. A linearly polarized probe light travels through the atomic ensemble in the x direction. The almost fully oriented spins along the probe propagation direction cause the polarization of the probe light to rotate with Faraday angle θ. Assuming that the ensemble is fully polarized ($j_x = 2$), the number of atoms N_{at} can be estimated from θ by[35]

$$N_{at} = -\frac{32\pi\theta V\Delta}{a_1(\Delta)\Gamma\lambda^2 j_x l_c} \tag{9}$$

where l_c is the path length in the x direction, V is the volume of the cell, $\lambda = 780$ nm is the wavelength of the probe light and $\Gamma = 2\pi \times 6.067$ MHz is the full-width at half-maximum linewidth of the excited state. The vector polarizability a_1 is given in Supplementary Information as a function of the laser detuning Δ.

Atomic population. We use the magneto-optical resonance signal (MORS) method to characterize the atomic polarization[36]. In the experiment, a d.c. magnetic field B_x induces the Larmor precession at $\Omega_L = g_F\mu_F B_x/\hbar$ and a quadratic Zeeman splitting. A short RF magnetic field pulse at frequency Ω_L along the z direction is applied at the end of the optical pumping pulse to excite a $\Delta m = 1$ coherence between the magnetic sublevels. The subsequent spin evolution is measured through Faraday interaction with a weak linearly polarized probe beam propagating in the z direction. In Extended Data Fig. 1, the spin evolution after the short RF pulse and the corresponding Fourier transformation are plotted. The spin polarization is estimated to be 97.9% by fitting the experimental data to the model of ref. [36].

As a result of the imperfect optical pumping, a small fraction of the atoms remain in the $F = 1$ manifold. This amount can also be estimated using the MORS method, with the laser tuned close to the D2 $F = 1 \rightarrow F'$ transitions. The RF pulse excites $\Delta m = 1$ coherences in the $F = 2$ and $F = 1$ manifolds. The frequency of the $\Delta m = 1$ coherence for $F = 1$ is about 0.4% higher than that for $F = 2$, so we can distinguish them in the frequency domain, and we estimate the population in the $F = 1$ manifold to be less than 5% under the application of optical pumping, causing negligible effects in noise calibration.

The effective coupling strength $\tilde{\kappa}^2$ is calibrated by measuring the spin noise of the unpolarized atomic ensemble with equal population on all $F = 1$ and $F = 2$ ground states. The measured spin noise of the unpolarized sample is 1.25 times that of the CSS for the following reasons. The atoms in both the unpolarized state and the CSS are uncorrelated, so

$$\mathrm{Var}(\hat{J}_z) = \sum_{i=1}^{N_{at}} \mathrm{Var}\left(\hat{J}_z^i\right) \tag{10}$$

In the CSS, $\mathrm{Var}(\hat{J}_z) = \mathrm{Var}(\hat{J}_y) = \frac{j_x}{2} = 1$, whereas in the unpolarized state the spin is symmetric, which means $\mathrm{Var}(\hat{J}_z) = \langle \hat{J}_z^2 \rangle = \langle \hat{J}_x^2 \rangle = \langle \hat{J}_y^2 \rangle = \frac{F(F+1)}{3} = 2$ for $F = 2$. When all sublevels, including three $F = 1$ states that are not observed in the measurement, have the same population, the contribution of the five $F = 2$ sublevels to the observed noise is $2 \times \frac{5}{8} N_{at} = \frac{5}{4} N_{at}$. Whereas for the CSS, the observed noise should be $1 \times N_{at}$.

In our experiment, we use light to measure the spin noise. Thus, the total noise includes the light shot noise and spin noise. So we have

$$\tilde{\kappa}^2 = \left[\frac{\mathrm{Var}\left(\hat{S}_y^{\mathrm{thermal}}\right)}{\mathrm{Var}\left(\hat{S}_y^{\mathrm{light}}\right)} - 1 \right] \times 0.8 \tag{11}$$

Here $\hat{S}_y^{\mathrm{thermal}}$ and $\hat{S}_y^{\mathrm{light}}$ are the Stokes components acquired when measuring the unpolarized spin noise and photon shot noise, respectively.

When measuring the photon shot noise, the Larmor frequency is tuned far away from the lock-in detection bandwidth by changing the d.c. magnetic field, ruling out the noise contribution from spin noise. Extended Data Fig. 2 shows the dependence of photon shot noise on the input probe power, and the linearity demonstrates the behaviour of the photon shot noise limit, because for the coherent state of light the variances of $\hat{S}_y$ and $\hat{S}_z$ should satisfy $\mathrm{Var}(\hat{S}_y) = \mathrm{Var}(\hat{S}_z) = \frac{S_x}{2}$.

QND character of the measurement. In Extended Data Fig. 3, the coupling strength $\tilde{\kappa}^2$ and the atomic noise variance in the state prepared by optical pumping are plotted as functions of the atomic number. The observed linear scaling of spin noise power indicates a quantum limited performance and the QND character of the measurement. The atom number is independently measured by the off-resonant Faraday rotation, which gives an optical depth of about 70 at the operation temperature of 53.5 °C. This temperature was chosen as a trade-off to maximize the size of the atomic ensemble, prevent degradation of the paraffin coating, reduce the spin exchange process at higher temperature and attain high spin orientation.

Adiabatic turn-off of the pump fields. Even after fine-tuning the alignment of the optical pumping beams with the magnetic field, a small residual π-polarization component persists when viewing in the x-quantization basis, which, together with the σ^- component, creates unwanted ground-state coherence (associated with a superposition state $|F = 2, m_F = -2\rangle + \varepsilon|F = 2, m_F = -1\rangle$ where $\varepsilon \ll 1$) via two-photon processes, creating additional classical spin components $J_{y,z}$. Furthermore, an abrupt turn-off of the pump fields can excite more coherence owing to its broader Fourier spectrum. However, such unwanted coherence can be eliminated by slowly turning off the pump lasers as the parasitic superposition state adiabatically evolves to $|F = 2, m_F = -2\rangle$.

PQS-enhanced magnetometry

In this section we outline how the collective spin squeezing and the retrodicted spin uncertainty may benefit practical precision measurements. We consider the application of a time-dependent RF magnetic field with a slowly varying envelope of the form $B_{RF} = B_0 f(t)$, which causes a temporary displacement of the spin observable

$\langle J_z \rangle \propto F(t) \equiv \int^t f(\tau)\mathrm{d}\tau$. We assume that the shape of $f(t)$ is known and that it is completed with $F(t) = 0$ before the last measurements so that the subsequent m_3 measurements carry no information about B_0.

The perturbation coincides in time with the m_2 probe sequence, which hence yields a record of data proportional to the time-dependent offset of the spin $\propto F(t)B_0$ shown in the upper panel of Fig. 4a. The ith coherent probe pulse undergoes a coherent displacement by $\tilde{\kappa}_2[\hat{p}_A + F(t_i)B_0]$, and by subtracting the expectation value $\tilde{\kappa}_2\langle\hat{p}_A\rangle$ of the unperturbed atomic spin, which is inferred from the density matrix or PQS conditioned by the measured signal, we obtain a noisy estimate of $\tilde{\kappa}_2 F(t_i)B_0$. All field measurements are subject to a Gaussian error with a variance comprised of the measurement photon shot noise and $\tilde{\kappa}_2^2$ times the variance of $\hat{p}_A$. We have verified that the m_1 measurements, alone and in conjunction with the m_3 measurements, yield the mean value and variance of the m_2 measurements of $\hat{p}_A$ according to equations (7) and (8). These equations thus constitute the basis for estimating the RF magnetic field amplitude B_0. For simplicity, we disregard the measurement back-action of the individual (weak) m_2 pulses and hence treat their combined effect as an effective QND measurement of $\hat{p}_A$, including a time-weighted (equal weighting, for simplicity) integral of the displacement $F(t_i)B_0$. Subtracting in each run of the experiment the conditional mean spin given by equation (7) or (8) thus provides an estimate of B_0. The uncertainty in the B_0 measurement (determining the magnetometer sensitivity) is composed of the shot noise contributions and the spin variance, σ^2, given by equation (7) or (8). It is clear that the measurement uncertainty is reduced when we apply the PQS results, where the spin variance takes the smallest value.

Retrodiction is thus beneficial when measuring an RF magnetic field with zero mean amplitude. This inspires echo-type experiments in which, for example, B_{RF} is stable and lasts for τ_2, but at time $\tau_2/2$ one applies a very short π pulse so the displacement caused by B_{RF} is reversed and the final displacement is zero. Similar to our experimental study, using a third probe pulse for retrodiction will improve the measurement of B_{RF}. Other time-dependent signals, including noisy signals with known governing statistics, may be inferred from the more elaborate time-dependent PQS theory, which may hence apply to many naturally occurring physical situations[37].

In addition, we note that the length of τ_2 is a trade-off between two factors: on the one hand, increasing τ_2 will enhance sensitivity; on the other hand, when τ_2 is comparable to the entanglement lifetime of ~1 ms, the conditioning protocol (both forward and especially backward) does not help. In other words, our protocols are good for measurements of relatively fast profile changes (of the RF amplitude) owing to the finite entanglement lifetime. This is also the case for other squeezing-enhanced metrology applications[38–41].

RF magnetic field detection and calibration. In the RF atomic–optical magnetometry, a polarized spin ensemble is prepared by optical pumping in the presence of a static magnetic field, which determines the atomic Larmor frequency. A transverse RF magnetic field $B_{RF}e^{i\Omega_L t}$ at the Larmor frequency causes the spin ensemble to precess and the angle of precession is proportional to the RF magnetic field. The spin dynamics are monitored with a weak off-resonant linearly polarized probe beam. As the probe beam travels through the atomic vapour, its plane of polarization rotates by an angle proportional to the spin component along the propagation direction according to the Faraday effect.

The Stokes component $\hat{S}_y$ carrying the transverse spin information can be measured in a balanced polarimetry scheme in the ±45° basis. The signal at the Larmor frequency $\hat{S}_{y,c}$ is extracted[35] with a lock-in amplifier (Zurich Instrument). Here the subscript 'c' indicates 'cosine', the in-phase quadrature of the lock-in amplifier output. The sensitivity to the RF magnetic field is given by[42,43] $B_{sen} = B_{min}\sqrt{T}$ (where T is the measurement time) with the minimal detectable field $B_{min} = B_{RF}/\mathrm{SNR}$. In practice, the signal-to-noise ratio SNR in our magnetometer is defined as

$$\mathrm{SNR} = \frac{|\langle\hat{S}_{y,c}\rangle|}{\sqrt{\mathrm{Var}(\hat{S}_{y,c})}} \tag{12}$$

Experimentally, a pair of Helmholtz coils oriented along the z axis generates a RF magnetic field along the z axis, perpendicular to the main spin along the x direction. The pulse sequence employed in our PQS-enhanced magnetometry is schematically shown in Fig. 4a.

In the protocol of PQS-enhanced magnetometry, the denominator in equation (12) is replaced by $\sqrt{\mathrm{Var}(m_2|m_1, m_3)}$, with $\mathrm{Var}(m_2|m_1, m_3)$ the variance of m_2 conditioned on the measurements before and after τ_2, that is, m_1 and m_3. Here, m_2 (that is, $\hat{S}_{y,c}$) is the sum of all the data points obtained during τ_2 in one sequence. In our demonstration RF-field measurement of a triangularly shaped RF profile, $B_0 = \max(B_{RF})$ is the height of the triangle. To compare the sensitivity with other magnetometers, we use the following definition of the aforementioned sensitivity $B_{sen} = B_0\sqrt{\tau_2}/\mathrm{SNR}$.

To calibrate the RF coil, a pickup coil with $N_\omega = 30$ turns of copper wire and 8.35 mm diameter is employed, located at the position of the Rb cell, along the axis of the Helmholtz coils. The oscillating magnetic field creates a flux through the pickup coil that generates an electromotive force. When applying a sinusoidal magnetic field of frequency ω and amplitude B_{RF}, the current through the pickup coil can be found from measuring the voltage amplitude U_ω across the measurement resistor R_m (ref. [44]). Then we have the relation between B_{RF} and U_ω

$$|B_{RF}| = \frac{|1 + Z_{coil}/R_m||U_\omega|}{N_\omega A_{coil}\omega} \tag{13}$$

where A_{coil} is the cross-sectional area of the pickup coil. Its impedance is $Z_{coil} \approx i\omega L$ at frequency Ω_L with inductance $L \approx 30$ μH, because the resistance of the coil $R = 1.9\ \Omega$ is much smaller than ωL at the frequency at which we usually operate ($2\pi \times 500$ kHz). We use a spectrum analyser to read out the response generated in the pickup coil. The voltage is read out over the resistance $R_m = 50\ \Omega$. The measured amplitude of the voltage is $U_\omega = \sqrt{2}\,U_{rms}$ where U_{rms} is the root-mean-square voltage. The measurement result is shown in Extended Data Fig. 4b, which indicates that the pickup coil's voltage is in good linear relation with the voltage output of the signal generator (Agilent E8257D).

Extended Data Figure 4b seems to indicate that, combined with equation (13), we may get a relation between the RF field B_{RF} (seen by the atom) and the signal generator's output. However, as shown in Extended Data Fig. 4b, this calibration can only be done for relatively large RF output from the signal generator, owing to excess electrical noises dominating the small electromotive-force voltage on the pickup coil. In practice, we applied a smaller magnetic field on our atoms, which could not be directly measured via the pickup coil. The possible solution is the following. Given that $B_{RF} \propto U_{app}$, where U_{app} is the applied voltage on the RF coil, we may extrapolate B_{RF} for the lower RF output range from the magnetic field amplitude measured in the higher RF output range with the pickup coil.

To prove that such extrapolation to the low range of RF output in Extended Data Fig. 4b is valid, we use the atoms to measure the RF field B_{RF} in this range, which however still partially overlaps with the range of Extended Data Fig. 4b. Indeed, we found that the atoms are much more sensitive than the pickup coil. For very small RF output from the signal generator, B_{RF} can be measured by the displacement of atomic spins but not by the pickup coil. Extended Data Fig. 4a presents the results of the magnetic field B_{RF} calibration performed by monitoring the displacement of atomic spin J_z. The setup is the same as that used in the magnetic field detection experiment. We vary the peak amplitude of the RF magnetic field and record the mean value of the sum

of data points during the second sequence, that is, the mean value of m_2. As illustrated in Extended Data Fig. 4a, the linearity of the mean value versus the RF peak amplitude is good, indicating the validity of the linearity of the RF signal generator's output reading U_{set} and the response of atomic spins, further enabling the extrapolation that we use, $B_{RF} \propto U_{app} \propto U_{set}$.

Based on the aforementioned observation, we can obtain the relation between the applied magnetic field and the output of the signal generator as

$$B_0(T) = 9.686 \times 10^{-8} \times 10^{P_{set}/20} \quad (14)$$

where P_{set} (in units of dB m) is the set output power reading of the RF signal generator. Through this calibration, we get the peak amplitude of the applied RF magnetic field in the magnetic field detection experiment, which is about 1.00 pT ($P_{set} = -97$ dB m on our signal generator).

Data availability

The datasets generated and analysed during this study are available from the corresponding authors upon reasonable request.

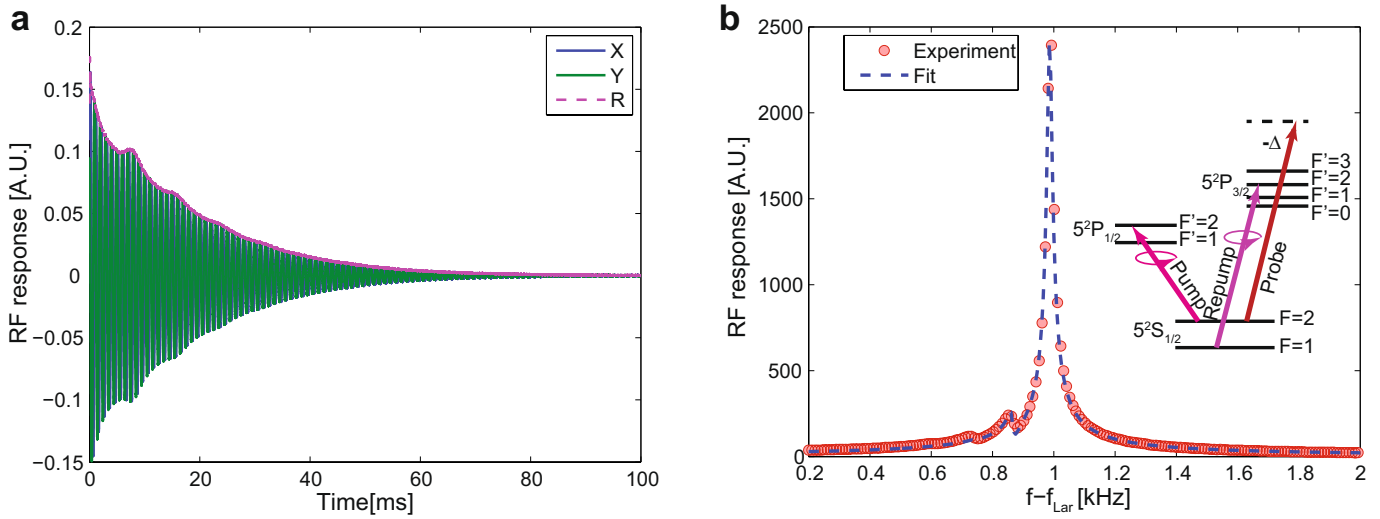
34. Bell, W. E. & Bloom, A. L. Optically driven spin precession. *Phys. Rev. Lett.* **6**, 280–281 (1961).
35. Shen, H. *Spin Squeezing and Entanglement with Room Temperature Atoms for Quantum Sensing and Communication*. PhD thesis, Univ. of Copenhagen (2015).
36. Julsgaard, B., Sherson, J., Sørensen, J. L. & Polzik, E. S. Characterizing the spin state of an atomic ensemble using the magneto-optical resonance method. *J. Opt. B* **6**, 5–14 (2004).
37. Yonezawa, H. et al. Quantum-enhanced optical-phase tracking. *Science* **337**, 1514–1517 (2012).
38. Huelga, S. F. et al. Improvement of frequency standards with quantum entanglement. *Phys. Rev. Lett.* **79**, 3865–3868 (1997).
39. André, A., Sørensen, A. S. & Lukin, M. D. Stability of atomic clocks based on entangled atoms. *Phys. Rev. Lett.* **92**, 230801 (2004).
40. Auzinsh, M. et al. Can a quantum nondemolition measurement improve the sensitivity of an atomic magnetometer? *Phys. Rev. Lett.* **93**, 173002 (2004).
41. Braverman, B., Kawasaki, A. & Vuletić, V. Impact of non-unitary spin squeezing on atomic clock performance. *New J. Phys.* **20**, 103019 (2018).
42. Budker, D. & Kimball, D. F. J. *Optical Magnetometry* (Cambridge Univ. Press, 2013).
43. Kominis, I. K., Kornack, T. W., Allred, J. C. & Romalis, M. V. A subfemtotesla multichannel atomic magnetometer. *Nature* **422**, 596 (2003).
44. Jensen, K. *Quantum Information, Entanglement and Magnetometry with Macroscopic Gas Samples and Non-Classical Light*. PhD thesis, Univ. of Copenhagen (2011).

Acknowledgements We thank M. Balabas for assistance in the vapour cell fabrication and V. Vuletić for discussions. This work is supported by the National Key Research Program of China under grants 2016YFA0302000 and 2017YFA0304204, and the NNSFC under grants 61675047 and 91636107. K.M. acknowledges support from the Villum Foundation. H.S. acknowledges financial support from a UK Royal Society Newton International Fellowship (NF170876).

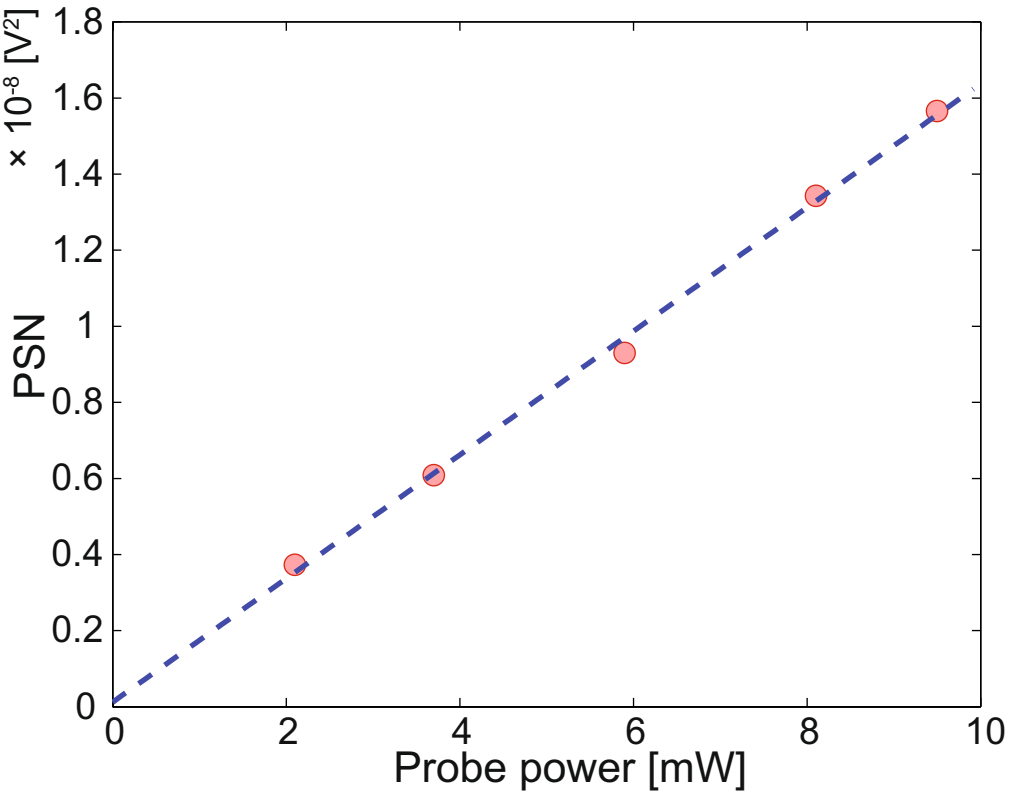
Author contributions K.M., H.S. and Y.X. conceived the idea. H.B., J.D., S.J., X.L., P.L., I.N., E.E.M., H.S. and Y.X. designed the experiment, performed the measurements and analysed the data together with all other authors. K.-F.Z. helped with the fabrication and characterization of vapour cells. H.B., M.W. and H.S. carried out the theoretical analysis under K.M.'s supervision. H.B., H.S., K.M. and Y.X. wrote the manuscript with contributions from all other authors. H.S. and Y.X. supervised the project.

Competing interests The authors declare no competing interests.

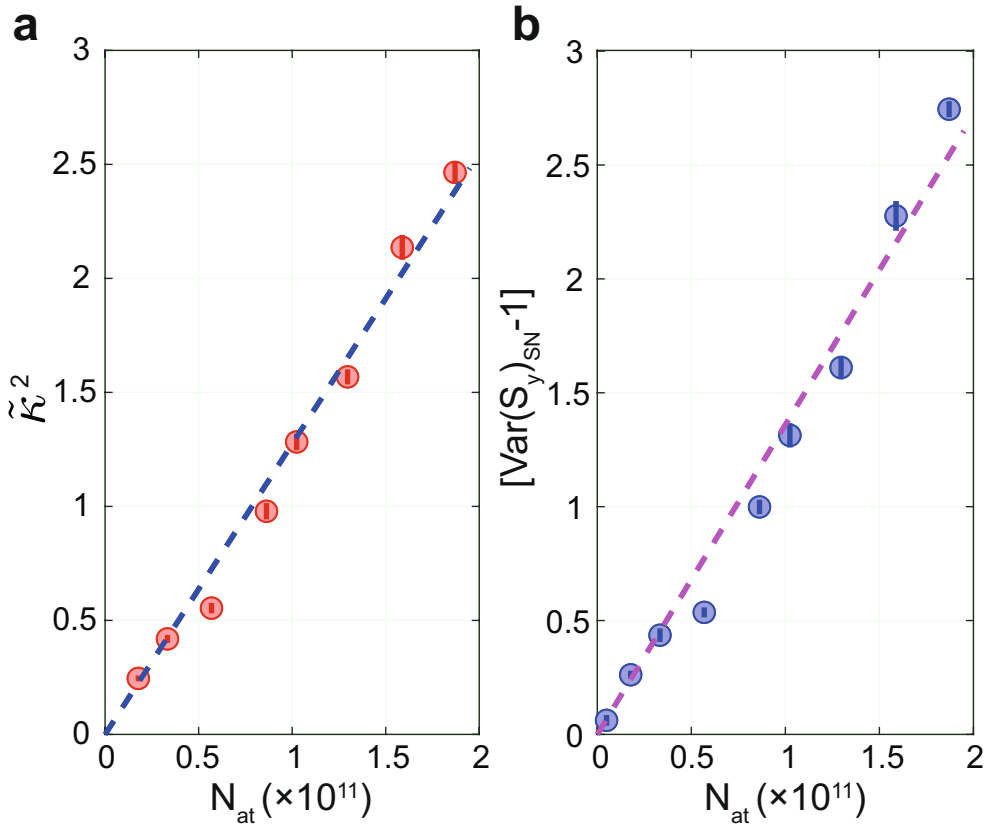
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2243-7.
Correspondence and requests for materials should be addressed to K.M., H.S. or Y.X.
Peer review information *Nature* thanks Julian Martinez-Rincon and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



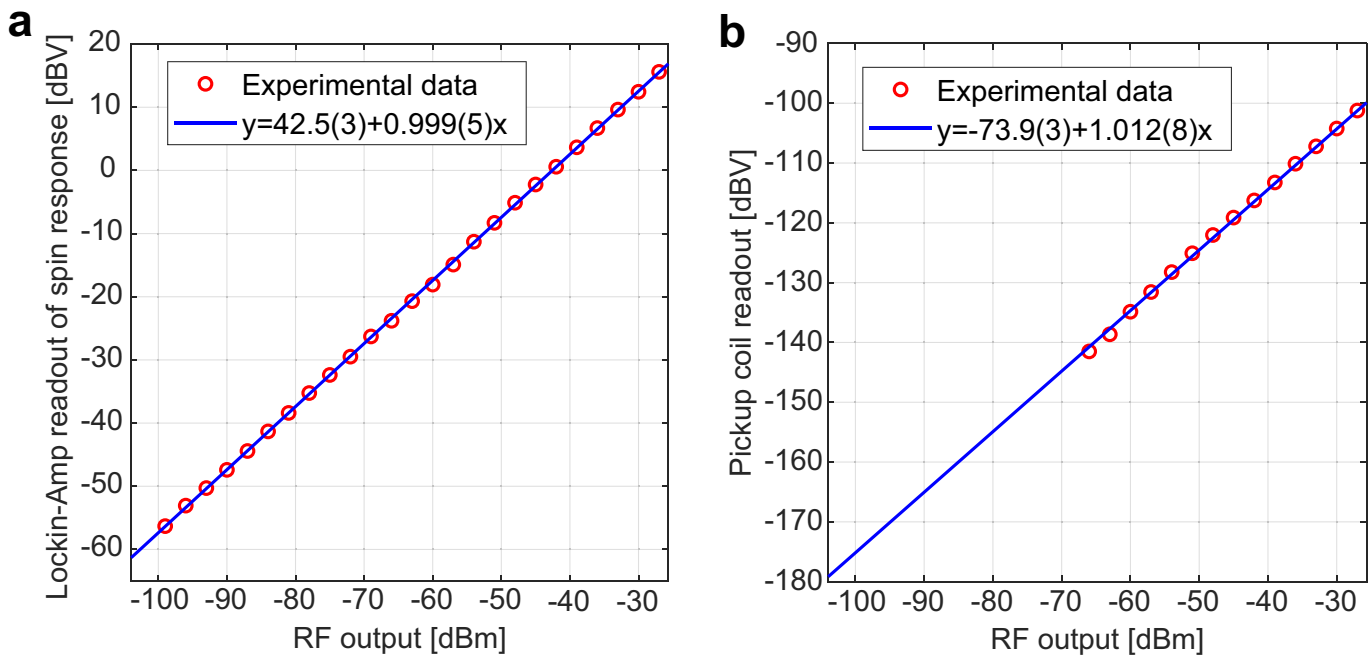
Extended Data Fig. 1 | Magneto-optical resonance signal. a, Spin response to an RF pulse. X and Y are the outputs of the lock-in amplifier, with a π/2 phase difference between them. $R=\sqrt{X^2+Y^2}$ is the demodulated amplitude. **b**, The associated Fourier transformation of the spin response signal. f_{Lar} is the centre frequency for demodulation, with the subscript 'Lar' representing 'Larmor frequency'. f is the actual frequency of the signal before demodulation. $f-f_{\mathrm{Lar}}$ represents the frequency of the signal after demodulation, that is, at the lock-in amplifier output. Inset, energy levels of ^{87}Rb. All the atoms are pumped into the $F=2$, $m_F=-2$ state, so that they are oriented along x. The magnetic field leads to a splitting of the magnetic sublevels by the Larmor frequency Ω_{L}. A.U., arbitrary units.



Extended Data Fig. 2 | Measured photon shot noise with different probe powers. Red circles are experimental data and the dashed line represents the linear fit of the data. PSN, photon shot noise.



Extended Data Fig. 3 | Spin noise versus atom number. a, Effective coupling constant $\tilde{\kappa}^2$ as a function of the number of atoms. The values of $\tilde{\kappa}^2$ are derived from the spin noise of the thermal state. **b**, Spin noise of prepared CSS versus the number of atoms. The observed linear dependence proves that technical noise is mostly suppressed and the measured spin noise is at the projection noise limit (PNL).



Extended Data Fig. 4 | Calibration of the applied RF magnetic field. **a**, Calibration using the displacement of atomic spins. **b**, Calibration using a small pickup coil. The amplitude of the RF output in our detection experiment is −97 dB m, which lies at the bottom left of the figure. In both curves, a slope near 1 indicates a good linear relation.