Multistability is an extraordinary nonlinear property of dynamical systems and can be explored to implement memory and switches. Here we experimentally realize the tristability in a three-mode cavity magnonic system with Kerr nonlinearity. The three stable states in the tristable region correspond to the stable solutions of the frequency shift of the cavity magnon polariton under specific driving conditions. We find that the system staying in which stable state depends on the history experienced by the system, and this state can be harnessed to store the history information. In our experiment, the memory time can reach as long as 5.11 s. Moreover, we demonstrate the ternary logic gate with good on-off characteristics using this multistable hybrid system. Our new findings pave a way towards cavity magnonics-based information storage and processing.

Introduction.—Light-matter interaction plays a crucial role in information processing, storage, and metrology (see, e.g., Refs. 1–5). Its utilization allows one to control, configure, and create new phases of matter or light signal. In the past few years, cavity magnonics [6–33], which is built on coherently [6–8] or dissipatively [23–26] coupled cavity photons and magnons, has increasingly demonstrated its unique advantages in both fundamental and application researches. The magnon mode can work as a transducer between microwave and optical photons, and the related studies are typically referred to as optomagnonics [17–19]. Coherent coupling and entanglement between the magnon mode and superconducting qubits are also experimentally realized [13, 28, 34]. These achievements show that the cavity magnonics has become a versatile platform for interdisciplinary studies, and also provides a useful building block for hybrid quantum systems [35–38].

In ferrimagnetic materials, such as the yttrium iron garnet (YIG), on account of the magnetocrystalline anisotropy, the magnon mode can equivalently act as a nonlinear resonator with Kerr-type nonlinearity [5, 21, 39]. When the magnon mode is driven by a microwave field, the excitation number of magnons increases. The Kerr effect induces a frequency shift and bistability can occur. An intriguing question is whether the extraordinary higher-order multistability [41–48] can be realized in the cavity magnonic system [49, 50]. Also, it is known that in a bistable or multistable system, the system staying in which stable state is related to the driven history [51–54], and it is important to characterize this memory effect.

Moreover, logic gates can be implemented by harnessing the stable states as logic states. Most of the current processing systems are based on the binary system, and the information is stored in “0” and “1” bits. The multivalued system can, however, be used to implement ternary, quaternary, or even higher-valued logic gates, significantly reducing the number of devices and the overall system complexity [55]. Furthermore, it has been shown that ternary logic gates working as elementary computing units can be more efficient in artificial intelligence simulations [56–58].

In this Letter, we report the first experimental observation of the multistability in a three-mode cavity magnonic system. This hybrid system is composed of two YIG spheres (1 mm in diameter) strongly coupled to a three-dimensional microwave cavity. Owing to the flexible tunability of the cavity magnonic system, we can continuously tune the system parameters to gradually turn the system from being bistable to multistable. Then we characterize the memory effect of the stable state in the tristable region. We find that the cavity magnonic system exhibits a remarkable long-time memory on the driven history that the system experienced. The memory time can approach several seconds, far exceeding the coherence time of the magnon mode. We further demonstrate the on-off switch of the frequency shifts, which reveals a well-repeatable operability. Finally, we build a ternary logic gate by utilizing three separated stable states as the logic output states, and the inputs correspond to zero, moderate, and high drive powers.

In a prior study on the paramagnetic spin ensemble, bistability with a long relaxation time was also observed at the temperature $\sim 25$ mK [59], which is achieved via the nonlinearity arising from the high magnon excitations in the magnon-photon coupling. However, one cannot use this nonlinearity for a ferrimagnetic or ferromagnetic spin ensemble, because less magnons can be excited by a drive field before the system enters the unstable regime referred to as the Suhl instability [60]. Thus, we harness the magnon Kerr effect in the YIG sphere. This involves a different mechanism for the nonlinearity. In addition, a very low temperature is needed for the paramagnetic spin ensemble [59], while multistability can be implemented at a high temperature in our hybrid system, since YIG has a Curie temperature of 560 K. This makes it possible for the
logic and switch devices built on cavity magnonics to work in a wide range of temperatures.

System and model.—The experimental setup is schematically shown in Fig. 1(a). Two 1 mm-diameter YIG spheres are placed in a microwave cavity with a dimension of \(44 \times 20 \times 6 \text{ mm}^3\). The cavity has three ports. Ports 1 and 2 are connected to the vector network analyzer (VNA) for probing the system, and port 3 is terminated with a loop antenna for loading the drive fields. Two microwave sources are combined at port 3 to jointly provide the drive tone, working as input A and input B in the logic gate operation. The YIG spheres are glued on the cavity walls where the magnetic-field antinodes occur for the cavity TE\(_{102}\) mode. Here the YIG sphere 1 is placed near port 3, which can be directly driven by the loop antenna. Because of the antenna-enhanced radiative damping, the damping rate of the Kittel mode \([61]\) in YIG sphere 1 (called magnon model 1) is much larger than that of the Kittel mode in YIG sphere 2 (called magnon model 2). For these two magnon modes and the cavity mode, their damping rates are measured to be \(\gamma_1/2\pi = 13.6 \text{ MHz}, \gamma_2/2\pi = 1.3 \text{ MHz},\) and \(\kappa_c/2\pi = 3.2 \text{ MHz}\). The system is placed in a uniform bias magnetic field \((\mathbf{B}_0)\), and two local magnetic fields \((\mathbf{B}_1\) and \(\mathbf{B}_2\)) are further applied to tune the frequencies of the magnon modes 1 and 2 separately.

When the magnon mode 1 is driven by a microwave field, the hybrid system can be described by

\[
H/\hbar = \omega_c a^\dagger a + \omega_1 b_1^\dagger b_1 + \omega_2 b_2^\dagger b_2 + K_1 b_1^\dagger b_1 b_1 + K_2 b_2^\dagger b_2 b_2 + g_1(a^\dagger b_1 + a b_1^\dagger) - ig_2 (a^\dagger b_2 - a b_2) - i \Omega_d (b_1 e^{i\omega_d t} - b_1^\dagger e^{-i\omega_d t}),
\]

where \(a^\dagger(a)\) is the creation (annihilation) operator of the cavity photon at frequency \(\omega_c\), \(b_1^\dagger(b_1)\) is the creation (annihilation) operator of the magnon mode 1(2) at frequency \(\omega_1(2)\), \(g_1(2)\) is the coupling strength between the magnon mode 1(2) and the cavity mode, \(\Omega_d(\omega_d)\) is the drive-field strength (frequency), and \(K_1(2)\) is the Kerr coefficient of the magnon mode 1(2). When the [100] crystal axis of the YIG sphere is aligned parallel to the bias magnetic field, the Kerr coefficient is positive \([5, 21]\). This coefficient can be tuned from positive to negative by adjusting the angle between the crystal axis and the bias field \([5, 62]\).

It is worth noting that \(g_1\) and \(g_2\) have a \(\pi/2\) phase difference due to the relative positions of the two YIG spheres in the cavity magnetic field as shown in Fig. 1(a). We get \(g_1/2\pi = 41.5 \text{ MHz}\) and \(g_2/2\pi = 50.5 \text{ MHz}\) by fitting the transmission spectrum. When a considerable number of magnons are excited, the Kerr term gives rise to a frequency shift of the magnon modes and cavity mode \([62]\),

\[
[\delta_e^2 + (\kappa_c + \eta_c)\gamma_2^2] \Lambda - g_1^2 \chi_1 = 0,
\]

\[
\{\delta_1 + \chi_1 - \eta_1 \delta_e^2 + [\gamma_1 + \eta_1 (\kappa_c + \eta_c) \gamma_2^2]\} \chi_1 = k P_d,
\]

\[
[\delta_2 + \chi_2 + \eta_2 \delta_e^2 + \gamma_2^2] \chi_2 - g_2^2 \xi \Lambda = 0,
\]

where \(\Lambda = 2 K_1 [A^2, \varepsilon = K_2/K_1, \delta_1(2) = \omega_1(2) - \omega_d, \delta_e = \omega_c - \omega_d, \delta_1 = \delta_1 - \eta_2 (\delta_2 + \chi_2), \eta_2 = g_2^2 / (\delta_2 + \chi_2)^2 + \gamma_2^2, \eta_1 = g_1^2 / [\delta_e^2 + (\kappa_c + \eta_c) \gamma_2^2],\) and \(k P_d = 2 K_1 |\Omega_d|^2\). By solving Eq. (2), we can obtain the frequency shifts \(\chi_1\) and \(\chi_2\). With these frequency shifts of the individual modes, it is easy to obtain the frequency shifts of the cavity magnon polaritons (CMPs) \([62]\). To find out the multistability, we need to solve the equations numerically to get the CMP frequency shifts at specific drive powers [Fig. 1(b)]. In this work, we focus on the upper-branch CMP with a frequency of \(\omega_{1P}\). In the experiment, \(K_1/2\pi = 0.0074 \text{ MHz}, K_2/2\pi = 0.0295 \text{ MHz},\) and the numbers of magnons excited are both less than 3% of the total number of spins in the two YIG spheres \([62]\).

Transition from bistability to multistability.—We tune the two magnon modes to \(\omega_1/2\pi = 10.212\) and \(\omega_2/2\pi = 10.198 \text{ GHz}\), being off resonance to the cavity mode TE\(_{102}\) \((\omega_c/2\pi = 10.080 \text{ GHz})\). In the case with positive Kerr coefficient, the system is found to reach the multistable regime only when the drive-field frequency detuning...
The hysteresis loop is observed at drive power forward and backward. In Fig. 2(a), only one χ shift area depicts the tristable region. The yellow star indicates the curves are numerical results calculated using Eq. (2). The gray area depicts the tristable region. The yellow star indicates the state in which we measure the memory effect.

\[ \delta_{UP} \equiv \omega_{UP} - \omega_d \] is negative. We monitor the frequency shift \( \chi_{UP} \) of the upper-branch CMP while sweeping the drive power forward and backward. In Fig. 2(a), only one hysteresis loop is observed at \( \delta_{UP}/2\pi = -2 \) MHz, corresponding to the bistability of the upper-branch CMP. As we tune \( \delta_{UP}/2\pi \) to \(-6\) MHz, we observe two separated hysteresis loops [Fig. 2(b)]. In this case, multistability still does not occur. We further tune \( \delta_{UP}/2\pi \) to \(-10\) MHz. Now, two hysteresis loops start to merge with each other, and the multistability is going to emerge [Fig. 2(c)]. When the detuning becomes \( \delta_{UP}/2\pi = -18 \) MHz, the frequency shift of the upper-branch CMP in Fig. 2(d) clearly indicates the occurrence of three stable states in the drive power range from 18.5 to 22 dBm (gray region). The sandwiched state in the tristable region is observed by sweeping the drive power backward when the upper-branch polariton frequency shift is initially prepared in the middle stable state (MSS), as depicted by the red triangles.

Theoretical curves are calculated using Eq. (2), which fit well with the experimental results. Multistability was first predicted for the cavity magnonic system in Ref. [49]. Our observations in Fig. 2 convincingly verify this prediction and reveal the excellent tunability of the cavity magnonic system.

**Long-time memory of the stable state.**—Multistability has great potential applications in memory and switches. In previous work, the spin memory of hundreds of exciton polaritons was reported [53]. In the multistable region, the system staying in which stable state depends on the history the system has experienced. In our system, after turning off the drive field, the higher-energy stable state will have its energy relaxed. Then the frequency shift of the polaritons will quickly vanish. If the microwave drive is turned off for a sufficiently long period, the system can no longer return to the higher-energy stable state when the microwave drive is turned on again. However, there is a critical time (termed as the memory time) for the relaxed system that can still recover to the higher-energy stable state when the drive field is applied again. In this work, we experimentally testify that the multistable state of the CMP has a considerably long memory time.

We investigate the memory time of the MSS in the tristable region indicated by the yellow star in Fig. 2(d). In Fig. 3(a), we first prepare the CMP in the higher-energy state with a certain frequency shift by applying a 26 dBm drive field. Meanwhile, we turn on the probe field to continuously monitor the frequency shift of the upper-branch CMP. Then we gradually sweep the drive power to 21 dBm. The system now enters the tristable region and reaches the state labeled by the yellow star. This state can be accessed by sweeping the drive power backward when the upper-branch polariton frequency shift. The frequency shift quickly vanishes.

**FIG. 2.** Frequency shift \( \chi_{UP} \) of the upper-branch CMP versus the drive power \( P_d \). The blue (green) triangles are the forward (backward) sweeping results. (a) \( \delta_{UP}/2\pi = -2 \) MHz, (b) \(-6\), (c) \(-10\), and (d) \(-18\) MHz. The red triangles are obtained by sweeping the drive power backward from 26 dBm, when the upper-branch CMP is in the middle stable state (MSS). The black curves are numerical results calculated using Eq. (2). The gray area depicts the tristable region. The yellow star indicates the state in which we measure the memory effect.

**FIG. 3.** (a) The memory time measurement. The system is first prepared in the MSS located in the tristable region. The drive field is turned off and on to find out the maximum memory time (5.11 s). The inset is the schematic of the drive-power sweeping procedure. (b) The observed frequency-shift relaxation and fitting. The fitting equation is \( \chi_{UP}/2\pi = 19.84 e^{-t/\Gamma_{UP}} \) MHz, which gives the relaxation rate \( \Gamma_{UP}/2\pi = 0.049 \) Hz. (c) A sequence of switches with different power-off times. From left to right, they are 3.60, 4.20, 4.80, and 5.20 s, respectively. For the power-off time of 5.20 s, we find that the system can no longer recover to the MSS.
After a period of time, we turn on the drive field again, and the system starts to revive and finally returns to the MSS. We term the time interval between turning off and on the drive field as the memory time. In the case of the drive power switching on and off at 21 dBm, the maximum memory time is found to be 5.11 s [62]. When the turning-off time exceeds 5.11 s, the CMP frequency shift cannot recover to the MSS again [see Fig. 3(c)].

The relaxation rate of the frequency shift reflects the time scale of the magnetocrystalline anisotropy energy relaxation [62]. In Fig. 3(b), the exponential relaxation of the CMP frequency shift is fitted and the relaxation rate is found to be $\Gamma_{\text{UP}}/2\pi = 0.049$ Hz. It shows that this energy relaxation process owns a rate far smaller than the linewidth of the magnon mode. Relaxation rates measured at various initial states are fitted [62], which are all close to 0.049 Hz.

**Ternary logic gate.**—Our tristable cavity magnonic system is a controllable multivalued system, which can be used to build ternary logic gates. As mentioned above, corresponding to a specific drive power, there is a frequency shift of the upper-branch CMP. The magnitude of these frequency shifts are determined by the theoretical multistability curve solved using Eq. (2). By suitably selecting three input powers, three highly distinguishable frequency shifts can be defined as the “0,” “1,” and “2” output states, as schematically shown in Fig. 1(b). The corresponding inputs 0, 1, and 2 are illustrated in Fig. 1(b) and the left panel of Fig. 4(a). The ternary logic gate we demonstrate is an OR gate, in which, as long as one of the two inputs is 2, the output is 2. If there is no strong drive power (2) input, but there are two or one moderate inputs 1, the output is 1. If both of the two inputs are 0, the output is 0, i.e., there is no CMP frequency shift.

In the case with the two drive fields applied simultaneously, to have the joint drive power equal to the larger input power of the two inputs, we add a phase shifter at one of the input channels. The phase shifter is carefully tuned and fixed to achieve a $2\pi/3$ phase delay between the two input channels [67], which makes the synthesized power of two 2 (1) inputs equal single 2 (1) input [see Fig. 4(a)]. For a 2 input plus a 1 input, the strong drive power is several times the moderate drive power. After the synthesis, the joint drive power can be approximately equal to the strong drive power. Our follow-up experiments have confirmed this.

In the experiment, we design a series of input A-input B combinations, including all the nine logic configurations listed in the truth table in Fig. 1(c). The input sequences are shown in the top panel of Fig. 4(b). The output of the ternary OR gate is displayed in the bottom panel of Fig. 4(b). We find that different input combinations give excellent distinguishable output states of 0, 1, and 2, which correspond to the zero, small, and large frequency shifts of the upper-branch CMP. These results unambiguously indicate the implementation of an OR gate.

**Conclusions.**—We have experimentally demonstrated the multistability in a three-mode cavity magnonic system. In the tristable region, we find the long-time memory effect of the middle stable state of the CMP frequency shift. The memory time can be as long as 5.11 s, which is millions of times the coherence time of the magnon mode and that of the cavity mode. The switch function of the memory is also characterized within the maximum memory time, which shows a good switchable feature. Utilizing the multivalued cavity magnonic system, we build a ternary logic OR gate, which exhibits highly distinguishable logic states. Our findings offer a novel way towards cavity magnonics-based memory and computing.

This work is supported by the National Natural Science Foundation of China (No. 11934010, No. U1801661, and No. 11774022), the National Key Research and Development Program of China (No. 2016YFA0301200), Zhejiang Province Program for Science and Technology (No. 2020C01019), and the Fundamental Research Funds for the Central Universities (No. 2021FZZX001-02). G.S.A. thanks AFOSR for support from AFOSR Award No. FA9550-20-1-0366.

*yipuwang@zju.edu.cn*
Communications 4, 2008 (2013).
[54] V. Kubytskyi, S.-A. Bihs, and P. Ben-Abdallah, Phys. Rev. Lett. 113, 074301 (2014).
[55] S. B. Jo, J. Kang, and J. H. Cho, Advanced Science 8, 2004216 (2021).
[56] S. Kak, arXiv preprint arXiv:1807.06419 (2018).
[57] L. Luo, Z. Dong, X. Hu, L. Wang, and S. Duan, International Journal of Bifurcation and Chaos 30, 2050222 (2020).
[58] In realistic problems, there are uncertainties and ambiguities, with things not simply black or white, and intermediate states can emerge. By utilizing the ternary device, one can record the intermediate state in only one physical element.
[59] A. Angerer, S. Putz, D. O. Krimmer, T. Astner, M. Zens, R. Glattauer, K. Streltsov, W. J. Munro, K. Nemoto, S. Rotter, J. Schmiedmayer, and J. Majer, Science Advances 3, e1701626 (2017).
[60] H. Suhl, J. Phys. Chem. Solids 1, 209 (1957)
[61] C. Kittel, Phys. Rev. 73, 155 (1948).
[62] See Supplemental Material at http:// for additional details of the theoretical derivations, measurement methods, data analysis, and the frequency-shift relaxation time, which includes Refs. [7],[8],[39],[5], and [63–66].
[63] S. Blundell, Magnetism in Condensed Matter (Oxford University Press, Oxford, 2001).
[64] D. D. Stancil and A. Prabhakar, Spin Waves: Theory and Applications (Springer, New York, 2009).
[65] O. O. Soykal and M. E. Flatte, Phys. Rev. Lett. 104, 077202 (2010).
[66] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
[67] For a cosine signal with amplitude $A$, the time average of the signal square corresponds to the mean power of the signal: $\int_0^{2\pi} (A \cos \omega t)^2 dt \bigg/ (2\pi) = \frac{A^2}{2}$, where $2\pi/\omega$ is the period of the signal. For two equal-amplitude cosine signals jointly supplied as the drive tone, i.e., $A \cos \omega t + A \cos (\omega t + \varphi)$, when $\varphi = 2\pi/3$, the mean power of the synthetic signal is still equal to $\frac{A^2}{2}$: $\int_0^{2\pi} [A \cos \omega t + A \cos (\omega t + \varphi)]^2 dt \bigg/ (2\pi) = \frac{A^2}{2}$. In the experiment, this can be easily realized by adding a phase shifter at one of the input channels and carefully tuning the channel phase delay to make $\varphi = 2\pi/3$. 
I. THE HAMILTONIAN OF THE CAVITY MAGNONIC SYSTEM

We study a three-mode cavity magnonic system consisting of two Kittel modes in two YIG spheres and a cavity mode. Among them, one of the magnon modes is driven by a microwave field. The Hamiltonian of the cavity magnonic system is

\[ H = H_c + \sum_{i=1}^{2} (H_{m,i} + H_{\text{int},i}) + H_d, \]  

(S1)

where \( H_c \) is the Hamiltonian of the cavity mode, \( H_{m,1(2)} \) is the Hamiltonian of the magnon mode 1(2), \( H_{\text{int},1(2)} \) is the interaction Hamiltonian between the magnon mode 1 (2) and the cavity mode, and \( H_d \) is the interaction Hamiltonian related to the drive field. In the experiment, the probe-field power is \(-20\) dBm, which is much smaller than the drive-field power of \(15 \sim 30\) dBm. Also, the probe field is loaded through the cavity port and the drive field is directly applied to the YIG sphere via the antenna, yielding the latter to have a much higher excitation efficiency on the magnons. Therefore, we ignore the magnons excited by the probe field in our experiment.

First, we focus on the Hamiltonian of the magnon mode. Under the bias magnetic field \( \mathbf{B}_0 \), the Hamiltonian reads [1, 2]

\[ H_m = -\int_{V_m} \mathbf{M} \cdot \mathbf{B}_0 d\tau - \frac{\mu_0}{2} \int_{V_m} \mathbf{M} \cdot \mathbf{H}_{\text{an}} d\tau, \]  

(S2)

where the first term represents the Zeeman energy and the second term is the magnetocrystalline anisotropy energy. In Eq. S2, \( V_m \) is the volume of the YIG sphere, \( \mathbf{M} = (M_x, M_y, M_z) \) is the magnetization of the YIG sphere, \( \mu_0 \) is the vacuum permeability, and \( \mathbf{H}_{\text{an}} \) is the anisotropic field due to the magnetocrystalline anisotropy in the YIG crystal.

We adopt the direction of the bias magnetic \( \mathbf{B}_0 \) as the \( z \) direction (\( \mathbf{B}_0 = B_0 \mathbf{e}_z \)). When the [100] crystal axis of the YIG sphere is aligned along the bias magnetic field, the anisotropic field is given by [3]

\[ \mathbf{H}_{\text{an}} = \frac{2K_{an}M_z}{\mu_0 M^2} \mathbf{e}_z, \]  

(S3)

where \( K_{an} \) is the first-order magnetocrystalline anisotropy constant. The Hamiltonian of the magnon mode is

\[ H_m = -B_0 M_z V_m - \frac{K_{an} M_z^2}{M^2} V_m. \]  

(S3)

Using the macrospin operator \( \mathbf{S} = \mathbf{M} V_m / \hbar \gamma \equiv (S_x, S_y, S_z) \) [4], where \( \gamma \) is the gyromagnetic ratio and \( \hbar \) is the reduced Planck constant, the Hamiltonian \( H_m \) can be written as

\[ H_m / \hbar = -\gamma B_0 S_z - \frac{hK_{an} \gamma^2}{M^2 V_m} S_z^2. \]  

(S4)

Here we can define the Kerr coefficient as

\[ K = -\frac{hK_{an} \gamma^2}{M^2 V_m}. \]  

(S5)

With the parameters \( M = 143.2 \) kA/m, \( K_{an} = -610 \) J/m\(^3\) [3], \( \gamma / 2\pi = 28 \) GHz/T, \( \hbar = 1.0546 \times 10^{-34} \) J \cdot s, and \( V_m = 4.19 \times 10^{-9} \) m\(^3\) for the YIG sphere of 1 mm in diameter, we obtain the Kerr coefficient \( K / 2\pi = 0.0295 \) nHz for the [100] crystal axis of the YIG sphere aligned along the bias magnetic field. By rotating the YIG sphere to tune the angle between the crystal axis and the bias magnetic field, the resulting Kerr coefficient can be changed from positive to negative [5].

To achieve the strong-coupling regime, we place the two YIG spheres at the antinodes of the cavity magnetic field \( h \mathbf{c} \) (cf. Fig. 1 in the main text). Compared with the cavity volume, the YIG spheres are relatively small, so we can assume that the cavity magnetic field is nearly uniform throughout each YIG sphere. Considering the different phases of the cavity magnetic field at the two YIG spheres, we can write the interaction Hamiltonians between the macrospins and the cavity mode as [2, 4]

\[ H_{\text{int},1} / \hbar = 2g_{s,1} S_{x,1} (a^\dagger + a) \]
\[ = g_{s,1} (S^+_1 + S^-_1) (a^\dagger + a), \]  

(S6)
\[ H_{\text{int},2}/\hbar = 2g_{s,2}S_{y,2}(a^\dagger + a) = -i g_{s,2}(S_2^+ - S_2^-)(a^\dagger + a), \] (S7)

where the macrospin of the YIG sphere 1(2) couples to the cavity magnetic field along the \( x \) (\( y \)) direction, and \( g_{s,1(2)} \) is the coupling strength between the macrospin 1(2) and the cavity mode. In the experiment, a drive field of frequency \( \omega_d \) is applied to directly pump the YIG sphere 1 along the \( y \) direction. The corresponding interaction Hamiltonian is

\[ H_d/h = -i \Omega_s(S_1^+ - S_1^-)(e^{i\omega_d t} + e^{-i\omega_d t}), \] (S8)

where \( \Omega_s \) denotes the coupling strength between the drive field and the macrospin of the YIG sphere 1. The drive magnetic field at port 3 is designed to be perpendicular to the cavity magnetic field. Also, the cavity microwave magnetic field will not pass through the loop of the antenna, which results in a nearly zero coupling between the antenna and the cavity, as shown in Fig. S1. Therefore, we neglect the drive-field effect on the cavity mode in the Hamiltonian.

Then, we harness the Holstein-Primakoff transformation [6]: \( S^+ = \sqrt{2S} - b^\dagger b_b \), \( S^- = b^\dagger \sqrt{2S} - b_b b \), and \( S_z = S - b^\dagger b \), where \( S \) is the total spin of the sample. The mean excitation numbers of magnons in two YIG spheres can be calculated using the frequency shift \( \chi_{1(2)} \) and Kerr coefficient \( K_{1(2)} \) via the relation \( \chi_{1(2)} = 2K_{1(2)} \langle b_{1(2)}^\dagger b_{1(2)} \rangle \) (see Eqs. S11-S16 for details). In our experiment, when the drive power is 30 dBm, \( \langle b_{1(2)}^\dagger b_{1(2)} \rangle = 3.26 \times 10^{17} \) and \( \langle b_{1(2)}^\dagger b_{1(2)} \rangle = 9.45 \times 10^{16} \). The excitation numbers are much smaller than the total spin \( S \) of a 1 mm-diameter YIG sphere, \( S = 5.52 \times 10^{18} \), calculated using \( S = \rho V_m s \), with \( \rho = 4.22 \times 10^{27} \text{ m}^{-3} \), \( V_m = \frac{4}{3}\pi R^3 \), \( R = 1 \text{ mm} \), and \( s = \frac{5}{2} \) [7, 8]. Therefore, we get \( \langle b_{1(2)}^\dagger b_{1(2)} \rangle/2S = 2.9\% \) and \( \langle b_{1(2)}^\dagger b_{1(2)} \rangle/2S = 0.86\% \). The proportions of the magnons excited with respect to the total spin number are both less than 3\%, and the higher-order terms can be neglected in the Holstein-Primakoff transformation. In this case, we have \( S^+ \approx b\sqrt{2S} \) and \( S^- \approx b^\dagger \sqrt{2S} \). Under the rotating-wave approximation (i.e., neglecting the fast oscillating terms), the total Hamiltonian of the cavity magnonic system becomes

\[ H/h = \omega_c a^\dagger a + \omega_1 b_{1(2)}^\dagger b_{1(2)} + \omega_2 b_{2(2)}^\dagger b_{2(2)} + K_1 b_{1(2)}^\dagger b_{1(2)} b_{1(2)}^\dagger b_{1(2)} + K_2 b_{2(2)}^\dagger b_{2(2)} b_{2(2)}^\dagger b_{2(2)} + g_1 (a^\dagger b_{1(2)} + ab_{1(2)}^\dagger) - ig_2 (a^\dagger b_{2(2)} - ab_{2(2)}^\dagger) - i\Omega_d (b_{1(2)} e^{i\omega_d t} - b_{1(2)}^\dagger e^{-i\omega_d t}), \] (S9)

where \( \omega_{1(2)} \) is the frequency of the magnon mode 1(2), \( g_{1(2)} = \sqrt{2S}g_{s,1(2)} \) is the coupling strength between the magnon mode 1(2) and the cavity mode, and \( \Omega_d = \sqrt{2S}\Omega_s \) is the drive-field Rabi frequency. In order to obtain a large multistable region at a low drive power, numerical simulations indicate that it is preferable to choose \( K_2 \) two times of \( K_1 \) in our setup. For the YIG sphere 2, its [100] crystal axis is set to be parallel to the bias magnetic field, so we get \( K_2/2\pi = 0.0295 \text{ nHz} \) and then set \( K_1 \) to be \( K_1/2\pi = 0.0074 \text{ nHz} \) by adjusting the orientation of the YIG sphere 1.

In the rotating reference frame with respect to the drive-field frequency, we obtain the following quantum Langevin equations:

\[ \dot{a} = -i(\delta_c - i\kappa_c) a - ig_1 b_1 - g_2 b_2, \]
\[ \dot{b}_1 = -i(\delta_1 + 2K_1 b_{1(2)}^\dagger b_{1(2)} + K_1 - i\gamma_1) b_1 - ig_1 a + \Omega_d, \]
\[ \dot{b}_2 = -i(\delta_2 + 2K_2 b_{2(2)}^\dagger b_{2(2)} + K_2 - i\gamma_2) b_2 + g_2 a, \] (S10)

where \( \delta_c = \omega_c - \omega_d \), and \( \delta_{1(2)} = \omega_{1(2)} - \omega_d \). Then, we write each of the operators \( a, b_1, \) and \( b_2 \) as the sum of the steady-state value and the fluctuation, i.e., \( a \equiv A + \delta a, b_1 \equiv B_1 + \delta b_1, \) and \( b_2 \equiv B_2 + \delta b_2 \). At the steady state, we can...
have the equations for $A$, $B_1$, and $B_2$:
\[
\begin{align*}
(\delta_c - i\kappa_c)A + g_1 B_1 - ig_2 B_2 &= 0, \\
(\delta_1 + \chi_1 - i\gamma_1)B_1 + g_1 A + i\Omega_d &= 0, \\
(\delta_2 + \chi_2 - i\gamma_2)B_2 + ig_2 A &= 0,
\end{align*}
\] (S11)
where we have used the mean-field approximation for the Kerr terms and the condition \(\langle b_1^{\dagger(2)} b_1^{\dagger(2)} \rangle \gg 1\). Here, \(\chi_{1(2)} = 2K_{1(2)} \langle b_1^{\dagger(2)} b_1^{\dagger(2)} \rangle\) is the frequency shift of the magnon mode 1(2) due to the Kerr nonlinearity. From the third equation in Eq. S11, we have
\[
B_2 = -i\eta_2(\delta_2 + \chi_2 + i\gamma_2)A/g_2,
\] (S12)
where \(\eta_2 = g_2^2/[(\delta_2 + \chi_2)^2 + \gamma_2^2]\). Substituting Eq. S12 into the first equation in Eq. S11, we have
\[
[\delta_c - \eta_2(\delta_2 + \chi_2) - i(\kappa_c + \eta_2 \gamma_2)]A + g_1 B_1 = 0.
\] (S13)
Then, we define \(\tilde{\delta}_c \equiv \delta_c - \eta_2(\delta_2 + \chi_2)\). Substitution of Eq. S13 into the second equation in Eq. S11 gives
\[
\left\{\delta_1 + \chi_1 - \eta_1 \tilde{\delta}_c - i[\gamma_1 + \eta_1(\kappa_c + \eta_2 \gamma_2)]\right\}B_1 + i\Omega_d = 0,
\] (S14)
where \(\eta_1 = g_1^2/[(\delta_1^2 + (\kappa_c + \eta_2 \gamma_2)]\). Multiplying Eqs. S12-S14 with their complex conjugates, we can obtain the following equations for the frequency shifts of the magnon modes:
\[
\begin{align*}
[\tilde{\delta}_c^2 + (\kappa_c + \eta_2 \gamma_2)]A - g_1^2 \chi_1 &= 0, \\
[(\delta_1 + \chi_1 - \eta_1 \tilde{\delta}_c)^2 + [\gamma_1 + \eta_1(\kappa_c + \eta_2 \gamma_2)]^2] \chi_1 - kP_d &= 0, \\
[(\delta_2 + \chi_2)^2 + \gamma_2^2] \chi_2 - g_2^2\varepsilon \Lambda &= 0,
\end{align*}
\] (S15)
where \(\Lambda = 2K_1|A|^2\), \(\varepsilon = K_2/K_1\), and \(kP_d = 2K_1|\Omega_d|^2\). This is Eq. (2) in the main text. It should be noted that \(k\) is the only fitting parameter in our experiment. It is related to the conversion efficiency of the drive power \(P_d\) to the magnon mode 1 via the Kerr effect of magnons. For a given drive power \(P_d\), we can solve Eq. S15 to obtain the numerical solution for the magnon-mode frequency shift \(\chi_{1(2)}\). With these frequency shifts included, we then obtain the effective Hamiltonian of the hybrid system,
\[
H/\hbar \approx \omega_c a^{\dagger} a + (\omega_1 + \chi_1)b_1^{\dagger} b_1 + (\omega_2 + \chi_2)b_2^{\dagger} b_2 + g_1(a^{\dagger} b_1 + a b_1^{\dagger}) - ig_2(a^{\dagger} b_2 - a b_2^{\dagger}).
\] (S16)
With given \(\chi_1\) and \(\chi_2\), this Hamiltonian has three eigenvalues, i.e., the frequencies of the lower-branch (\(\omega_{\text{LP}}\)), middle-branch (\(\omega_{\text{MP}}\)), and upper-branch (\(\omega_{\text{UP}}\)) cavity magnon polaritons (CMPs). We define the frequency shift of the upper-branch CMP as \(\chi_{\text{UP}} = \omega_{\text{UP}}(\chi_1, \chi_2) - \omega_{\text{UP}}(\chi_1 = 0, \chi_2 = 0)\). By numerically solving Eqs. S15 and S16, we can obtain the frequency shift \(\chi_{\text{UP}}\) of the upper-branch CMP. Under appropriate conditions, the frequency shift has five solutions, three of them are stable and other two are unstable. From the experimental results, we have the following parameters of the hybrid system: \(\omega_1/2\pi = 10.212\) GHz, \(\omega_2/2\pi = 10.198\) GHz, \(\omega_c/2\pi = 10.080\) GHz, \(g_1/2\pi = 41.5\) MHz, \(g_2/2\pi = 50.5\) MHz, \(\gamma_1/2\pi = 13.6\) MHz, \(\gamma_2/2\pi = 1.3\) MHz, \(\kappa_c/2\pi = 3.2\) MHz, \(K_1/2\pi = 0.0074\) nHz, and \(K_2/2\pi = 0.0295\) nHz. Via the multistability experiment, we obtain the parameter \(k = 20\) by fitting the multistability curves with the numerical results.

II. THE MEASUREMENT OF THE MULTISTABILITY

To get the coupling strengths, in the experiment, we first place YIG sphere 1 in the cavity and measure the transmission spectrum. The coupling strength between the Kittel mode in YIG sphere 1 and the cavity mode is determined at the resonance point, which is \(g_1/2\pi = 41.5\) MHz. Then we place YIG sphere 2 in the cavity without removing YIG sphere 1 out. The coupling strength between the Kittel mode in YIG sphere 2 and the cavity mode is obtained as \(g_2/2\pi = 50.5\) MHz by fitting the transmission curve shown in Fig. S2(b).

The transmission spectra of the cavity magnonic system without applying a drive field are measured and plotted versus both the magnet-coil current and the probe-field frequency, as shown in Fig. S2(a), where three branches of CMP occur. When the magnet-coil current is 5.37 A, the frequencies of the two magnon modes are \(\omega_1/2\pi = 10.212\) GHz and \(\omega_2/2\pi = 10.198\) GHz. In this case, the three branches of CMP are labeled in Fig. S2(b). The frequencies of the lower-branch, middle-branch, and upper-branch CMPs are denoted as \(\omega_{\text{LP}}, \omega_{\text{MP}}, \text{and} \omega_{\text{UP}}\), respectively. We apply the drive field on the upper-branch CMP with \(\delta_{\text{UP}}/2\pi = (\omega_{\text{UP}} - \omega_d)/2\pi = -18\) MHz and measure the transmission spectra by...
sweeping the drive-field power. The measured 2D contour plots are shown in Figs. S2(c)-S2(e), where the peaks in red represent the upper-branch CMP. The central frequencies of the peaks correspond to the frequency of the upper-branch polariton. We use the ‘max’ function in Matlab® to find the peak frequencies of the transmission spectra. Figures S2(c) and S2(d) are obtained by sweeping the drive power forward and backward, respectively. The frequency shifts and jumps of the upper-branch CMP are clearly observed. Figure S2(e) shows the result obtained by sweeping the drive power backward when the upper-branch CMP is initially prepared in the frequency-shift middle stable state (MSS). The tristable region can be found between the two vertical dashed lines. The color scale is the same for all graphs.

### III. THE FREQUENCY-SHIFT RELAXATION RATE

Following Eq. S3, the Hamiltonian for the magnetocrystalline anisotropy energy of the YIG sphere is

$$H_{an}/\hbar = -\frac{\hbar K_{an} \gamma^2}{M^2 V_m} S_z^2.$$  \hspace{1cm} (S17)

Using the Holstein-Primakoff transformation [6] $S_z = S - b^\dagger b$, we have

$$H_{an}/\hbar = -\frac{\hbar K_{an} \gamma^2}{M^2 V_m} (S^2 - 2Sb^\dagger b + b^\dagger bb^\dagger b).$$  \hspace{1cm} (S18)

In the low-excitation case with $\langle b^\dagger b \rangle \ll S$, $H_{an}$ can be approximated as

$$H_{an}/\hbar \approx -\frac{\hbar K_{an} \gamma^2}{M^2 V_m} (S^2 - 2Sb^\dagger b).$$  \hspace{1cm} (S19)
Without the magnon excitation, the magnetocrystalline anisotropy energy is $H_{an} / \hbar = -h K_{an} \gamma^2 S^2 / (M^2 V_m)$. Under the mean-field approximation, the variation of the magnetocrystalline anisotropy energy due to the magnon excitation can be written as

$$\Delta E_{an} / \hbar \approx \frac{2S \hbar K_{an} \gamma^2}{M^2 V_m} \langle b^\dagger b \rangle = 2SK \langle b^\dagger b \rangle,$$

(S20)

which is proportional to the magnon excitation number.

FIG. S3. Relaxation rate measurements for the CMP frequency shift. (a) The frequency shift of the upper-branch CMP is first tuned to one MSS where the drive power is 20.75 dBm. The green circles are the experimental results, and the black line is the exponential fitting curve. The relaxation rate is fitted to be $\Gamma_{UP}/2\pi = 0.051$ Hz. (b) The initial state is another MSS where the drive power is 20.5 dBm. The blue rhombuses are the experimental results. The relaxation rate is fitted to be $\Gamma_{UP}/2\pi = 0.046$ Hz.

In the experiment, when the drive field is turned off, the magnon excitation number will decrease and the magnetocrystalline anisotropy energy will relax. Since the frequency shift of the magnon mode is also proportional to the magnon excitation number, the relaxation rate of the frequency shift reflects the time scale of the magnetocrystalline anisotropy energy relaxation. In the main text, we have shown that when the upper-branch CMP is driven to the middle stable state (MSS) where the drive power is 21 dBm, the measured energy relaxation rate is 0.31 Hz. In Fig. S3, we measure the magnetocrystalline anisotropy energy relaxation after the upper-branch CMP is prepared in the MSS by using two different drive powers. For the MSS with the drive power of 20.75 dBm, the energy relaxation rate is 0.051 Hz. For the MSS with the drive power of 20.5 dBm, the energy relaxation rate is 0.046 Hz. As expected, the measured energy relaxation rates are very close to each other. These individual measurements reveal the intrinsic relaxation rate of the magnetocrystalline anisotropy energy.

FIG. S4. Relaxation of the upper-branch CMP frequency shift in the tristable region. The critical memory time is determined by the relaxation time of the frequency shift from A to B.
IV. RELATIONSHIP BETWEEN THE MAXIMUM MEMORY TIME AND THE RELAXATION TIME OF THE CMP FREQUENCY SHIFT

As shown in Fig. S4, the gray triangles are the frequency shift $\chi_{UP}$ of the upper-branch CMP at different drive powers. The data points are the same as in Fig. 2(d) in the main text. After the frequency shift is prepared to be at the point A, we turn off the drive field. The frequency shift first relaxes to B, and then to C (zero frequency shift) afterwards. When we turn on the drive field before the frequency shift goes down below B, the frequency shift can still go back to A. Therefore, the state A is revived. When we turn on the drive field after the frequency shift goes down below B, the frequency shift can only go back to B even if we turn on the drive field again. Then, the memory of the state A is lost.

The maximum memory time is determined by the relaxation time of the CMP frequency shift from A to B, as shown in Fig. S4, which is 5.11 s in the present case.

* yipuwang@zju.edu.cn
† jqyou@zju.edu.cn

[1] S. Blundell, *Magnetism in Condensed Matter* (Oxford University Press, Oxford, 2001).
[2] Y. P. Wang, G. Q. Zhang, D. Zhang, X. Q. Luo, W. Xiong, S. P. Wang, T. F. Li, C.-M. Hu and J. Q. You, Phys. Rev. B 94, 224410 (2016).
[3] D. D. Stancil and A. Prabhakar, *Spin Waves: Theory and Applications* (Springer, New York, 2009).
[4] O. O. Soykal and M. E. Flatte, Phys. Rev. Lett. 104, 077202 (2010).
[5] A. G. Gurevich and G. A. Melkov, *Magnetization Oscillations and Waves* (CRC, Boca Raton, FL, 1996), pp. 50.
[6] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
[7] Y. Tabuchi, S. Ishino, T. Ishikawa, R. Yamazaki, K. Usami, and Y. Nakamura, Phys. Rev. Lett. 113, 083603 (2014).
[8] X. Zhang, C.-L. Zou, L. Jiang, and H. X. Tang, Phys. Rev. Lett. 113, 156401 (2014).