Experimental implementation of a Raman-assisted six-quanta process

S.O. Mundhada,1,∗ A. Grimm,1 J. Venkatraman,1 Z.K. Minev,1 S. Touzard,1 N.E. Frattini,1 V.V. Sivak,1 K. Sliwa,1,† P. Reinhold,1 S. Shankar,1 M. Mirrahimi,2 and M.H. Devoret1,†

1Department of Applied Physics, Yale University, New Haven, CT 06511.
2QUANTIC team, INRIA de Paris, 2 Rue Simone Iff, 75012 Paris, France

(Dated: November 19, 2018)

Fault tolerant quantum information processing requires specific nonlinear interactions acting within the Hilbert space of the physical system that implements a logical qubit. The required order of nonlinearity is often not directly available in the natural interactions of the system. Here, we experimentally demonstrate a route to obtain higher-order nonlinearities by combining more easily available lower-order nonlinear processes, using a generalization of the Raman transitions. In particular, we demonstrate a Raman-assisted transformation of four photons of a high-Q superconducting cavity into two excitations of a superconducting transmon mode and vice versa. The resulting six-quanta process is obtained by cascading two fourth-order nonlinear processes through a virtual state. This process is a key step towards hardware efficient quantum error correction using Schrödinger cat-states.

Many theoretical proposals have shown that quantum information can be protected against errors in a hardware efficient manner by creating a decoherence protected stabilized manifold using continuous driven-dissipative processes [1–11]. Experimentally, in the domain of circuit quantum electrodynamics (QED), nonlinear interactions arising from Josephson junction circuits have played a key role in the stabilization of a single state [12] or a manifold of quantum states [13, 14]. In particular, a stabilized manifold spanned by four coherent states of a harmonic oscillator has been proposed for the implementation of a hardware efficient logical qubit [6, 15]. This stabilization requires a nonlinear driven-dissipative process to force the harmonic oscillator to gain and lose photons in sets of four, resulting in the protection of the coherent states against dephasing errors. Combining such stabilization with correction against photon loss errors using quantum nondemolition parity measurements [16–19] results in complete first-order quantum error correction (QEC).

One approach to engineer the four-photon driven-dissipative process has been proposed in [20]. The idea is to implement a six-quanta process that exchanges four photons of a high-Q cavity mode $a$ (destruction operator $a$) with two excitations of a transmon mode $b$ (destruction operator $b$) and vice versa, corresponding to an effective interaction given by $a^4b^2 + a^2b^4$ (see Fig. 1a). Adding a two-excitation drive and dissipation on the transmon by employing a combination of techniques demonstrated in [13, 21], will then result in a four-photon driven-dissipative process on the cavity. The implementation of $a^4b^2 + a^2b^4$ interaction requires a Raman-assisted cascading of two four-wave mixing interactions, each of which exchanges two cavity photons with a virtual excitation in the transmon mode and a pump photon, and vice versa [13].

Raman transitions using linear processes [22, Ch. 6] or a combination of one linear and one nonlinear process [23] have been previously demonstrated. The $a^4b^2 + a^2b^4$ interaction, however, requires the cascading of two nonlinear multi-quanta processes. In our experiment we show that not only the Raman-assisted cascading of nonlinear processes is feasible, but also the magnitude of the effective interaction can be made much larger than the damping rates of the high-Q modes, hence, generating a useful interaction for QEC. In principle, the same driven-dissipative process could instead be realized by using a six-wave mixing term in the Josephson cosine potential, addressed using an off-resonant pump. The currently achievable magnitude of higher-order mixing ($\geq$ four-wave), obtained from expanding the Josephson cosine potential, is small compared to the dissipation rates of the system and other spurious terms present in the Hamiltonian [24, Sec. 1C]. Hence, Raman-assisted virtual cascading of low-order mixing processes is essential for enhancing the strength of the desired four-photon driven-dissipative process for hardware efficient QEC.

From a wider perspective, while three- and four-wave mixing processes have played a key role in circuit QED applications [25–31], many proposals will benefit from increasingly higher-order nonlinear interactions [8, 11, 32]. Hence, the deeper motivation of our experiment is to demonstrate how to engineer higher-order nonlinear interactions from readily available lower-order mixing processes. As shown in [24, Sec. 1A], it is possible to cascade any two processes through a virtual state as long as the commutator of the operators that describe the processes is the operator describing the desired higher-order process. Therefore, such cascading could be useful for the broader field of quantum optics and quantum control. Moreover, the possibility of cascading indicates that advanced techniques like GRAPE (gradient-ascent pulse engineering) [33, 34] could utilize pulses addressing nonlinear processes to gain additional control knobs over the system, thus potentially increasing the speed and fidelity of the engineered unitary operations.

In order to demonstrate the feasibility of cascading
nonlinear processes through virtual states, our experiment focuses on the Raman-assisted \(|g4⟩ \leftrightarrow |f0⟩\) transition as explained in Fig. 1b (see figure caption for explanation). This transition is a precursor to the aforementioned \(a^4b^2 + a^2b^4\) process which requires the \(|g,n⟩ \leftrightarrow |f,n-4⟩\) transitions to all occur simultaneously. As shown in Fig. 1b, the system is initialized in the \(|f0⟩\) state. The two pumped processes, one connecting \(|f0⟩\) to a virtual state close to \(|e2⟩\) with the rate \(g1\) and the other one connecting the virtual state to \(|g4⟩\) with the rate \(g2\), are nonlinear four-wave mixing processes. The frequencies of the two pumps involved, as inferred from Fig. 1c, are

\[
\begin{align*}
\omega_{p1} &= 2\tilde{\omega}_a - \tilde{\omega}_b + \chi_{bb} - 2\chi_{ab} + \Delta, \\
\omega_{p2} &= 2\tilde{\omega}_a - \tilde{\omega}_b + 2\chi_{ab} - \Delta,
\end{align*}
\]

where \(\tilde{\omega}_a/b\) are the Stark shifted frequencies of the cavity and the transmon mode in presence of the pumps, \(\chi_{ab}\) is the cross-Kerr and \(\chi_{bb}\) is the self-Kerr of the transmon mode. The effective Hamiltonian of the system to second-order in the rotating wave approximation (RWA) [35] is

\[
H_{\text{eff}} = g_{4ph} (|g4⟩⟨f0| + |f0⟩⟨g4|),
\]

where \(g_{4ph}\) is the magnitude of the cascaded process, given by

\[
g_{4ph} = \sqrt{4g_1g_2} \left( \frac{1}{\Delta} - \frac{1}{\chi_{bb} - 4\chi_{ab} + \Delta} \right). \quad (3)
\]

For the effective Hamiltonian to be valid, one has to choose the parameters such that \(|g1,2| \ll \Delta\), since, as is ubiquitous in Raman transitions, the leakage rate to the intermediate state (\(|e2⟩\) in our case) is directly proportional to the ratios \(|2\Delta|^2\). Detailed derivation and discussion of the effective Hamiltonian is included in [24, Sec. 1B].

The experimental setup for testing our transition requires (i) a high-Q resonator, (ii) a transmon mode for the conversion process, (iii) the ability to couple pumps strongly with the conversion transmon while maintaining the quality factor of the system and (iv) a second transmon mode to perform Wigner tomography [36] of the resonator. The high-Q storage resonator (\(T_1 = 76\,\mu s\) is realized as a high purity aluminum, \(\lambda/4\)-type, post-cavity [37] with frequency \(\omega_0/2\pi = 8.03\,\text{GHz}\) (see Fig. 1c). The resonator is dispersively coupled to two transmons as shown in Fig. 1c. The transmon in the conversion arm has a resonance frequency \(\omega_0/2\pi = 5.78\,\text{GHz}\), anharmonicity \(\chi_{ab}/2\pi = 122.6\,\text{MHz}\) and a cross-Kerr of \(\chi_{bb}/2\pi = 7.4\,\text{MHz}\) with the high-Q resonator. The \(T_1\) and \(T_2\) of the conversion transmon are 50\,\mu s and 7.6\,\mu s respectively. The second transmon is employed to perform Wigner tomography on the storage resonator and has a cross-Kerr of 1.1\,MHz with it. Both transmons are coupled to low-Q resonators through which we perform single-shot measurements of the transmon state (see [24, Sec. II A] for remaining system parameters). In the case of the conversion transmon, the measurement distinguishes, in single-shot, between the first three states \(|g⟩, |e⟩\) and \(|f⟩\). The enclosure of the high-Q resonator acts as a rectangular waveguide high-pass filter with a cutoff at \(\sim 9.5\,\text{GHz}\). Since the two pump frequencies, \(\omega_{p1}/2\pi = 10.397\,\text{GHz}\) and \(\omega_{p2}/2\pi = 10.294\,\text{GHz}\), are above the cutoff, they are applied through the strongly coupled (waveguide mode \(Q \leq 100\)) pin at the top. The high-Q resonator and the
transmon modes are below the cutoff and are thus protected from relaxing through this pin.

In order to locate the correct pump frequencies for the transition of interest, we use the pulse sequence shown in Fig. 2a. The system is initialized in \( |f0\rangle \) and the two pumps are applied for a variable period of time. The pump frequencies are swept such that the frequency difference is maintained constant at \( \omega_{p1} - \omega_{p2} = \chi_{ab} - 4\chi_{ab} + 2\Delta \). We choose \( \Delta/2\pi = 5.1 \text{MHz} \) and \( g_{1,2}/2\pi \sim 0.5 \text{MHz} \). The rising and falling edges of the pump pulses are smoothed using a hyperbolic tangent function with a smoothing time of 192 ns. These parameters are empirically optimized to reduce the leakage to the \( |e2\rangle \) state while achieving a \( g_{4ph} \) that is an order of magnitude faster than the decoherence rates of the system. The resulting cavity state is characterized by applying a photon-number selective \( \pi \)-pulse [38] on the

tomography transmon. The pulse has a gaussian envelope of width \( \sigma_{sel} = 480 \text{ns} \) (total length \( 4\sigma_{sel} \)), resulting in a pulse bandwidth of \( \sim 332 \text{kHz} \), which is less than the cross-Kerr between the tomography transmon and the high-Q resonator. As a result the tomography transmon is excited only when the storage resonator is in \( |0\rangle \). Finally, the state of the tomography transmon is measured. An optional single-shot measurement of the conversion transmon can also be performed as indicated by dashed green measurement pulse in Fig. 2a.

The outcome of the described measurement is shown in Fig. 2b. The population fraction of the Fock state \( |0\rangle \) is plotted as a function of the duration for which the pump pulses are applied and the detuning of the first pump \( \omega_{p1} \) from the \( |f0\rangle \leftrightarrow |e2\rangle \) transition. The data displays Rabi oscillations arising from two processes. The one on the left occurs when pump 1 is resonant with

\[ \chi_{ab} - 4\chi_{ab} + 2\Delta \]

Resonance of the system. The system is initialized in \( |f0\rangle \) by using \( \pi \)-pulses on \( |g\rangle \leftrightarrow |e\rangle \) and \( |e\rangle \leftrightarrow |f\rangle \) transitions. Following this, the two pumps are applied with varying frequency and duration. Finally a measurement of storage resonator population and, optionally, the conversion transmon state is performed. The frequency difference of the two pumps is maintained constant at \( \chi_{ab} - 4\chi_{ab} + 2\Delta \). (b) Rabi oscillations in the population of Fock state \( |0\rangle \) (\( p_0 \), colorbar). The x-axis shows the detuning of pump 1 from the \( |f0\rangle \leftrightarrow |e2\rangle \) transition, the y-axis shows the duration for which the two pumps are applied. The frequency landscape above the data explains the origin of the two chevron like features.

In order to locate the correct pump frequencies for the transition of interest, we use the pulse sequence shown in Fig. 2a. The system is initialized in \( |f0\rangle \) and the two pumps are applied for a variable period of time. The pump frequencies are swept such that the frequency difference is maintained constant at \( \omega_{p1} - \omega_{p2} = \chi_{ab} - 4\chi_{ab} + 2\Delta \). We choose \( \Delta/2\pi = 5.1 \text{MHz} \) and \( g_{1,2}/2\pi = 0.5 \text{MHz} \). The rising and falling edges of the pump pulses are smoothed using a hyperbolic tangent function with a smoothing time of 192 ns. These parameters are empirically optimized to reduce the leakage to the \( |e2\rangle \) state while achieving a \( g_{4ph} \) that is an order of magnitude faster than the decoherence rates of the system. The resulting cavity state is characterized by applying a photon-number selective \( \pi \)-pulse [38] on the
FIG. 4. Conditional Wigner tomography of the storage resonator after a quarter period of the $|f0\rangle \leftrightarrow |g4\rangle$ oscillation. After quarter period of $|f0\rangle \leftrightarrow |g4\rangle$ oscillation the system ends up in $(|f0\rangle + |g4\rangle)/\sqrt{2}$. (a) and (b) show experimental and theoretical plots of the Wigner function of the storage resonator after post-selecting the conversion mode in the $|g\rangle$ and $|e\rangle$ states. This leaves the storage resonator in Fock states $|4\rangle$, $|0\rangle$ respectively. (c) Wigner function of the cavity after photon-number selective $\pi$-pulses from $|f0\rangle$ to $|e0\rangle$ and $|e0\rangle$ to $|g0\rangle$ (indicated by $U_{\text{sel}}$) and post-selecting the conversion transmon in $|g\rangle$. The Wigner function shows that the storage resonator is in a coherent superposition of $(|0\rangle + |4\rangle)/\sqrt{2}$. (d), (e) and (f) are the Wigner functions of theoretically expected states corresponding to the operations in (a), (b) and (c) respectively.

the $|f0\rangle \leftrightarrow |e2\rangle$ transition. The one on the right corresponds to the two pumps being equally detuned from the $|f0\rangle \leftrightarrow |g2\rangle$ and $|e2\rangle \leftrightarrow |g4\rangle$ transitions. This is the Raman-assisted $|f0\rangle \leftrightarrow |g4\rangle$ transition of interest. The resulting chevron pattern for this transition is narrower since the cascaded transition occurs at a slower rate than the $|f0\rangle \leftrightarrow |g2\rangle$ transition. From the frequency of the oscillations we extract $g_{4ph}/2\pi = 0.32$ MHz. In separate experiments, we accurately characterize the pump strengths $g_1/2\pi = 0.53$ MHz and $g_2/2\pi = 0.48$ MHz by measuring the Stark shifts of the conversion transmon when the pumps are resonant with the $|f0\rangle \leftrightarrow |g4\rangle$ transition and applied one at a time [24, Sec. III B]. This eliminates any frequency dependent attenuation of pump strengths. The value of $g_{4ph}/2\pi$ calculated by plugging these parameters in Eq. (3) is 0.33 MHz, in close agreement with the measured value.

Having found the desired $|f0\rangle \leftrightarrow |g4\rangle$ process, we fix our pump frequencies to be resonant with this transition and proceed to characterize the populations of different Fock states of the storage resonator. These are obtained by varying the frequency at which the photon-number selective pulse on the tomography transmon is applied. The excited state population of the tomography transmon is plotted as a function of frequency of the selective pulse and the duration of pump pulses in Fig. 3a. The population fractions of various Fock states are inferred by taking cross-sections at the resonance frequency of the tomography transmon conditioned on the number of photons in the high-Q resonator. It is clear from the data that the resonator oscillates between $|0\rangle$ and $|4\rangle$ with some leakage to $|2\rangle$ due to the finite detuning $\Delta$ from $|e2\rangle$ (see the $\omega_{T2}/2$ lines). In the future iterations of our experiment, this leakage can be minimized by increasing the detuning and making the pulses more adiabatic, albeit at the cost of making the overall process slower. It is also possible to use pulse shaping techniques like stimulated Raman adiabatic passage (STIRAP) [22, Ch. 6.2.3] to implement this transition without any leakage. The population appearing in $|1\rangle$ and $|3\rangle$ is due to finite energy relaxation time of the resonator mode. Plots in the first column of Fig. 3b show the evolution of the $|0\rangle$, $|2\rangle$ and $|4\rangle$ state populations of the storage resonator and the $|f\rangle$, $|e\rangle$, $|g\rangle$ state populations of the conversion transmon as a function of time. The conversion transmon populations are measured independently using the dashed-green measurement pulse shown in Fig. 2a. The respective populations oscillate in phase with each other as expected. The amplitude of the oscillations is limited by the $T_2$ of the conversion qubit and the contrast of the two measurements. We are also able to resolve an envelope of fast oscillations in the populations of $|e\rangle$, $|g\rangle$ and $|2\rangle$, $|4\rangle$ states. These occur at the rate of detuning $\Delta$ and are to be expected from a Raman process. The plots in the second column of Fig. 3b show numerical data obtained from simulating Lindblad master equation of the system [24, Sec. IV]. The contrast of the simulation is scaled by the measurement contrast of the experimental system. The rate of the oscillations match well for experimental and simulation results and even the fast oscillations in the data are reproduced.

Finally, we demonstrate that the oscillations are coherent by showing entanglement between conversion transmon and storage resonator. We stop the oscillations after a quarter of a period (372 ns), thus preparing dominantly $(|f0\rangle + |g4\rangle)/\sqrt{2}$, and perform a Wigner tomography of the resonator, conditioned on conversion transmon states. As expected, the resonator ends up in Fock state $|4\rangle$ when the conversion transmon is post-selected in $|g\rangle$ with $|f\rangle$ as shown in Fig. 4a (4b). Moreover, applying a photon number selective $f \rightarrow g$ pulse on the conversion transmon, conditioned on zero photons in the storage resonator, disentangles the transmon from the resonator, leaving the system in $|g\rangle \otimes (|0\rangle + |4\rangle)/\sqrt{2}$. The Wigner function of the resonator after post-selecting the conversion transmon in $|g\rangle$, shown in Fig. 4c, depicts a $|0\rangle + |4\rangle)/\sqrt{2}$ state, thus proving that the oscillations are coherent.

In conclusion, we have shown that nonlinear processes can be cascaded, through a virtual state, to engineer a higher-order nonlinear Hamiltonians. This constitutes an important step towards implementing a hardware efficient logical qubit. In the future, we aim to generalize the
\( |g4⟩ \leftrightarrow |f0⟩ \) oscillations to a complete \( a^4 |f⟩⟨g| + a^† 4 |g⟩⟨f| \) Hamiltonian, by making all of the \( |gn⟩ \leftrightarrow |f(n−4)⟩ \) oscillations resonant simultaneously. This would require making the strength of the pumped processes \( g_{1,2} \) to be higher than the cross-Kerr terms \( \chi_{ab} \) between the cavity and the conversion transmon. Such pump strengths are not achievable in our current system due to spurious transitions induced by strong pump strengths, similar to those seen in [39, 40]. There has been a proposal to accomplish an increased tolerance for the pump strengths by shunting the transmon with a linear inductor [41]. Moreover, the use of flux-biased circuits to cancel cross-Kerr between two modes has also been proposed [42]. We intend to leverage these developments to implement the complete drive-dissipative protocol. Additionally, our experiment also paves the way for stabilizing the Barut-Girardello coherent states in two coupled cavities [11, 43] that enable completely autonomous QEC.

The authors acknowledge support from ARO grant number W911NF-14-1-0011 and W911NF-18-1-0212. Use of fabrication facilities was supported by the Yale Institute for Nanoscience and Quantum Engineering(YINQE) and the Yale SEAS cleanroom.

---

* Electronic address: shantam.mundhada@yale.edu
† Present address: Quantum Circuits Inc., New Haven, CT 06511
‡ Electronic address: michel.devoret@yale.edu

[22] D. A. Steck, Quantum and Atom Optics (Open publication, 2007).


Supplemental Materials for “Experimental implementation of a Raman-assisted six-quanta process”

I. THEORY

Here, we discuss the necessary conditions for Raman-assisted virtual cascading and present the calculations pertaining to the $|f0⟩ ↔ |g4⟩$ oscillations demonstrated in the experiment. Additionally, we compare the magnitude of this process with the estimated magnitude of currently achievable six-wave mixing processes.

A. Designing a Raman-assisted higher-order process

In this subsection we use the expressions for second-order rotating wave approximation (RWA) [S1] to obtain some pointers towards designing Raman-assisted higher-order processes. Consider a Hamiltonian in an interaction picture with respect to the diagonal part, given by

$$\frac{H_1(t)}{\hbar} = \frac{H_c}{\hbar} + g_1 e^{i\Delta t} A_1 + g_1^* e^{-i\Delta t} A_1^\dagger + g_2 e^{-i\Delta t} A_2 + g_2^* e^{i\Delta t} A_2^\dagger.$$  \hspace{1cm} (S1)

Here, $H_c$ is time-independent part of $H_1(t)$ and $A_1, A_2$ are operators describing off-diagonal interactions available in the system. In the given rotating frame, the two processes are detuned by $+\Delta$ and $-\Delta$ respectively. The effective Hamiltonian to the second-order in RWA is given by

$$H_{\text{RWA}} = \overline{H}_I(t) - i \int_0^T (H_I(t) - \overline{H}_I(t)) \, dt \left( H_I(t) - \overline{H}_I(t) \right)$$  \hspace{1cm} (S2)

where $\overline{H}(t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T H(t) \, dt$. Applying this to the Hamiltonian in Eq. S1 we obtain

$$\frac{H_{\text{eff}}}{\hbar} = \frac{H_c}{\hbar} + \frac{|g_1|^2}{\Delta} [A_1, A_1^\dagger] - \frac{|g_2|^2}{\Delta} [A_2, A_2^\dagger] + \frac{g_1 g_2}{\Delta} [A_1, A_2] - \frac{g_2^* g_1^*}{\Delta} [A_1^\dagger, A_2^\dagger].$$  \hspace{1cm} (S3)

This reveals effective interactions given by the commutation relations of the operators $A_1$ and $A_2$. Therefore, in order to design a Raman-assisted higher-order process we have to use the following two principles.

- Select the lower-order processes such that one of their commutation relation is a non-zero operator describing the required higher-order process.
- Design the lower-order processes to be oscillating with equal and opposite frequencies $\Delta$ so that their product survives the second order RWA.

Additionally, one can also engineer the time independent part $H_c$ of Eq. (S1) such that unwanted higher-order terms in $H_{\text{eff}}$ are canceled. Another issue to keep in mind is the validity of second-order RWA. Eq. S3 is a good approximation only when $\frac{|g_1|^2}{\Delta}, \frac{|g_2|^2}{\Delta} \ll 1$. In general, along with the resonant interactions given in $H_{\text{eff}}$, the two individual processes described by $A_1, A_2$ are also off-resonantly enabled, leading to leakages corresponding to $A_1, A_2$ transitions. These leakages can be minimized by selecting smaller values of $\frac{|g_1|^2}{\Delta}$ albeit at the cost of slowing the desired effective process as well.

B. Calculations for $|f0⟩ ↔ |g4⟩$ process

The calculation for a general process described by $a^d|f⟩⟨g| + a^{14}|g⟩⟨f|$ is done in [S2]. As mentioned in the main text, this process requires all $|gn⟩ ↔ |fn - 4⟩$ transitions to be resonant simultaneously. However, when the cross-Kerr between the cavity mode and the transmon mode is large compared to the strengths of the individual four-wave mixing processes $|g_{1,2}⟩$, as is the case in our current system, the resonant frequencies of $|gn⟩ ↔ |fn - 4⟩$ transitions depends on $n$. Hence, the process that we have demonstrated is photon-number selective. Here we present the pertinent calculations where we ignore the presence of the tomography transmon. Following the analysis in [S2] the Hamiltonian of the system in a displaced frame where the drives have been absorbed in the cosine potential is

$$\frac{H}{\hbar} = \omega_a a^\dagger a + \omega_b b^\dagger b - \frac{E_J}{\hbar} \left[ \cos (\Phi(t)) + \frac{\Phi^2(t)}{2!} \right].$$  \hspace{1cm} (S4)
Here $\Phi(t)$ is the phase across the Josephson junction of the conversion transmon given by

$$
\Phi(t) = \phi_a (a + a^\dagger) + \phi_b (b + b^\dagger) + \phi_b \sum_{k=1,2} \xi_k \exp(-i\omega_{pk}t) + \xi_k^* \exp(i\omega_{pk}t),
$$

and $E_J$ is the Josephson energy. The ratios $\phi_{a,b}$ are the dimensionless inductive participation amplitudes of the cavity and the transmon modes in the junction with typical magnitudes of $\phi_{a}^2 \sim 10^{-3}$ and $\phi_{b}^2 \sim 10^{-1}$ respectively. The two applied pump drives are at the frequencies $\omega_{pk}$ defined as

$$
\omega_{p1} = 2\omega_a - \omega_b + \chi_{bb} - 2\chi_{ab} + \Delta + \delta
$$
$$
\omega_{p2} = 2\omega_a - \omega_b + 2\chi_{ab} - \Delta + \delta.
$$

The frequencies $\tilde{\omega}_a, \tilde{\omega}_b$ are the Stark shifted frequencies of modes $a$ and $b$ respectively. The shift $\delta$ is added to both pump frequencies in order to account for higher-order frequency shifts in the system. Additionally, for the purpose of this calculation, we have assumed that the pumps only couple to mode $b$. This assumption does not lead to any loss of generality since the coupling to mode $a$ can be effectively absorbed in the time dependent part of $\Phi(t)$ with a slight modification to $\xi_k$.

Expanding the Hamiltonian in Eq. S4 to the fourth order in cosine expansion we get

$$
H_{\text{sys}} = \frac{H_0}{\hbar} + \left( g_1 e^{-i\omega_{p1}t} + g_2 e^{-i\omega_{p2}t} \right) a^\dagger b a^2 b^\dagger + \left( g_1 e^{i\omega_{p1}t} + g_2 e^{i\omega_{p2}t} \right) a^2 b^\dagger b. \tag{S5}
$$

where

$$
\frac{H_0}{\hbar} = \tilde{\omega}_a a^\dagger a + \tilde{\omega}_b b^\dagger b - \chi_{ab} a^\dagger a b^\dagger b - \frac{\chi_{aa}}{2} a^2 - \frac{\chi_{bb}}{2} b^2 b^\dagger
$$

and

$$
g_{1,2} = -\frac{E_J \phi_{a,b}^2}{2} \xi_{1,2} = -\frac{\chi_{ab}}{2}. \tag{S6}
$$

In the expression for the Hamiltonian we have preemptively ignored the fast rotating terms in the frame of $H_0$. The expressions for the Stark shifted frequencies of the modes are

$$
\tilde{\omega}_a = \omega_a - (|\xi_1|^2 + |\xi_2|^2) \chi_{ab}
$$
$$
\tilde{\omega}_b = \omega_b - 2 (|\xi_1|^2 + |\xi_2|^2) \chi_{bb} \tag{S7}
$$

where $\omega_{a,b}$ are the bare frequencies. Going into the rotating frame with respect to $H_0/\hbar + \chi_{aa} a^2 - \delta b^\dagger b$ we get the Hamiltonian in the interaction picture

$$
\frac{H_I}{\hbar} = -6\chi_{aa} |g4\rangle\langle g4| - (\chi_{aa} + \delta) |e2\rangle\langle e2| - 2\delta |f0\rangle\langle f0| + \sqrt{4} \left[ g_1 \exp(i\Delta t) + g_2 \exp(-i(\chi_{bb} - 4\chi_{ab} + \Delta)t) \right] |f0\rangle\langle e2| + \text{h.c.}
$$
$$
+ \sqrt{12} \left[ g_1 \exp(i(\chi_{bb} - 4\chi_{ab} + \Delta)t) + g_2 \exp(-i\Delta t) \right] |e2\rangle\langle g4| + \text{h.c.}, \tag{S8}
$$

where h.c. indicates Hermitian conjugate. Comparing this expression with the expression given in Eq. (S1), we can infer that the first row is the time independent part $H_I$ and, and $|f0\rangle\langle e2|, |e2\rangle\langle g4|$ are the individual $A_1, A_2$ processes in Eq. (S1). The other terms in the Hamiltonian do not contribute after second order RWA and hence are ignored. Finally, performing the RWA as specified in Sec. I A, we get the effective Hamiltonian

$$
\frac{H_{\text{eff}}}{\hbar} = g_{4ph} \left( |g4\rangle\langle f0| + |f0\rangle\langle g4| \right)
$$
$$
+ \frac{12 |g_2|^2}{\Delta} \chi_{bb} - 4\chi_{ab} + \Delta + 6\chi_{aa} - \frac{12 |g_1|^2}{\Delta} |g4\rangle\langle g4|
$$
$$
- \frac{12 |g_2|^2 - 4 |g_1|^2}{\Delta} \chi_{bb} - 4\chi_{ab} + \Delta + \chi_{aa} + \delta |e2\rangle\langle e2|
$$
$$
+ \frac{4 |g_2|^2}{\Delta} - \frac{4 |g_2|^2 - 4 |g_1|^2}{\Delta} \chi_{bb} - 4\chi_{ab} + \Delta - 2\delta |f0\rangle\langle f0|. \tag{S9}
$$
The first row in the above equation is the desired effective interaction. The magnitude of \( g_{4ph} \) has been quoted in Eq. 3 of the main text. The other terms in the Hamiltonian are higher-order frequency shifts introduced by the pumps. We compensate for these shifts in the experiment by sweeping the pump frequencies while keeping the difference between them constant at \( \chi_{bb} - 4\chi_{ab} + 2\Delta \). This common shift of pump frequencies given by \( \delta \), amounts to

\[
\delta = 3\chi_{aa} + \left( \frac{2|g_1|^2 - 6|g_2|^2}{\Delta} + \frac{6|g_1|^2 - 2|g_2|^2}{\chi_{bb} - 4\chi_{ab} + \Delta} \right). 
\]

(S10)

It can be seen that for this value of delta, the higher-order frequency shifts introduced in \(|f0\rangle\) and \(|g4\rangle\) states are equal, thus making the \(|f0\rangle \leftrightarrow |g4\rangle\) transition resonant.

C. Comparison with the magnitude of six-wave mixing process

As mentioned in the main text, the four-photon driven-dissipative process required to stabilize a manifold of four-component Schrödinger cat states can, in principle, be implemented in two distinct ways. The first way is through Raman-assisted cascading, which is the topic of exploration for our letter. The other way is using the six-wave mixing capabilities of a Josephson junction. The idea is to exchange four photons of a storage resonator with a single excitation of a Josephson junction mode such as transmon, SQUID, SNAIL [S3] etc., accompanied by a release of pump-photon, and vice versa; a five-quanta process. In this section we compare the estimated magnitude of this six-wave mixing process with that of the Raman-assisted cascading.

The magnitude of the six-wave mixing process can be estimated by expanding the cosine potential in Eq. (S4) to the sixth-order. The expression for the rate of this interaction is

\[
g_{\text{6-wave}} = \frac{E_t}{24\hbar} \phi_0^4 \phi_0^2 \xi_0 = \frac{\phi_0^2}{24} \chi_{ab} \xi,
\]

(S11)

where \( \xi_0 \) is the strength of the pump addressing the five-quantum process. On the other hand, using the expressions in [S2], one gets the rate of the Raman-assisted \( a^4b^4 + a^4b^2 \) process as

\[
g_{\text{Raman}} = \frac{\chi_{ab} \xi}{20} \left( 1 - \frac{5\chi_{ab} \xi}{\chi_{bb} - 4\chi_{ab} + 2\Delta} \right).
\]

(S12)

Here we have substituted \( g_1 = g_2 = \chi_{ab} \xi/2 \) and \( \Delta = 10g_{1,2} = 5\chi_{ab} \xi \). This maintains \( \Delta \gg g_{1,2} \) for the RWA to be valid. In order to estimate the relative strength of the processes, we use \( \xi_0 = \xi_1 = \xi_2 \equiv 0.2 \) and \( \phi_0^2 \equiv 0.002 \) obtained by using the parameters of our system as a guide. These are representative of the typical parameters achievable in resonators coupled to transmon modes. The pump strengths, though can be increased, are limited in high-Q transmon-resonator systems by the chaotic behavior that has been observed at high pump powers [S4, S5]. Using Eq. (S11) and Eq. (S12), we estimate that the Raman-assisted process will be stronger than the six-wave mixing process by a factor of \( \approx 110 \) which is two-orders of magnitude. Additionally, assuming the validity of these rate expressions at higher pump-powers, the Raman transition dominates the six-wave mixing process till the pump strength \( \xi \geq 600 \). In reality the expressions shown here break down at such high pump strengths [S6] and, in high-Q devices, these regimes have not been experimentally achieved yet.

II. EXPERIMENT

Further details on the system parameters, experimental setup and measurement protocols involved are presented in this section.

A. System parameters

Here we give the detailed parameters of our experimental system. As mentioned in the main text, we have a high-Q storage resonator coupled to two transmon modes which are the conversion transmon and the tomography transmon. Each transmon is in turn coupled to one low-Q resonators which facilitates single-shot heterodyne measurement of the transmon. Hence, in total, we have five crucial modes in our system. Table I specifies their frequencies, coherence times and coupling strengths with each other (off-diagonal \( \chi \)s in the table). The self-Kerr and \( T_2 \) are measured and specified only for the high-Q modes which are the storage resonator and the two transmon modes. In particular it is noteworthy that the modes for which the cross-Kerr are listed as NA, are indeed isolated from each other physically and hence, have negligible cross-Kerr between them.
TABLE I. System parameters

<table>
<thead>
<tr>
<th></th>
<th>Storage resonator</th>
<th>Conversion transmon</th>
<th>Conversion resonator</th>
<th>Tomography transmon</th>
<th>Tomography resonator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency f</td>
<td>8.03 GHz</td>
<td>5.78 GHz</td>
<td>9.93 GHz</td>
<td>6.36 GHz</td>
<td>7.53 GHz</td>
</tr>
<tr>
<td>T1</td>
<td>72 µs</td>
<td>50 µs</td>
<td>38 µs</td>
<td>8.8 µs</td>
<td>0.9 MHz</td>
</tr>
<tr>
<td>T2</td>
<td>56 µs</td>
<td>7.6 µs</td>
<td>8.8 µs</td>
<td>264 MHz</td>
<td>0.9 MHz</td>
</tr>
<tr>
<td>χ self</td>
<td>122 kHz</td>
<td>122 MHz</td>
<td>264 MHz</td>
<td>0.38 MHz</td>
<td></td>
</tr>
<tr>
<td>χ</td>
<td>7.4 MHz</td>
<td>5.7 MHz</td>
<td>1.1 MHz</td>
<td>0.9 MHz</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>72 µs</td>
<td>50 µs</td>
<td>38 µs</td>
<td>8.8 µs</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>56 µs</td>
<td>7.6 µs</td>
<td>8.8 µs</td>
<td>264 MHz</td>
<td></td>
</tr>
<tr>
<td>χ self</td>
<td>122 kHz</td>
<td>122 MHz</td>
<td>264 MHz</td>
<td>0.38 MHz</td>
<td></td>
</tr>
<tr>
<td>χ</td>
<td>7.4 MHz</td>
<td>5.7 MHz</td>
<td>1.1 MHz</td>
<td>0.9 MHz</td>
<td></td>
</tr>
</tbody>
</table>

B. Measurement setup

The principles of our measurement setup are similar to those shown in [S7]. A detailed wiring diagram has been shown in Fig. S1. The upper half contains the room temperature wiring (above 300 K dashed line) of the experiment and the lower half shows the wiring inside the dilution refrigerator. As mentioned in the main text, we have two transmon qubits and the ability to perform single-shot measurement on both qubits. The low-Q resonator coupled to the conversion transmon has a frequency of 9.93 GHz as mentioned in table I. This is above the cutoff frequency of the waveguide enclosure and hence, this mode couples to the transmission line through the strongly coupled pin situated at the top of the waveguide. This coupling pin serves the dual purpose of measurement pin for the conversion transmon as well as the pin through which the off-resonant pumps are applied. Moreover, the coupling pin only addresses the waveguide mode with the polarization along the length of the pin. Hence, the applied pumps only couple to the conversion transmon while leaving the tomography transmon unperturbed. All the tones applied on this pin are combined using a directional coupler and routed to the coupling pin using a circulator. The directional coupler also sends most of the pump signal back to room temperature, hence effectively attenuating the pump tones without heating up the base plate of the dilution refrigerator. The circulator directs the reflected signal from the waveguide pin towards a Josephson parametric converter (JPC) which amplifies the signal at the conversion resonator frequency and sends it back to room temperature via circulator, isolators and a high electron mobility transistor (HEMT) amplifier placed at 4K. The coupling pin situated close to the conversion arm is weakly coupled to the system and is used to drive the conversion transmon. The pin situated on the tomography arm, however, is strongly coupled to the tomography resonator and is used for three purposes. Firstly, it is used to readout the tomography resonator in reflection. The signal is routed using two circulators to a SNAIL parametric amplifier (SPA) [S8] and the amplified signal is routed through the circulator towards the output chain. The other two purposes of the tomography arm coupling pin are to address the tomography qubit as well as the storage resonator. In fact, the relaxation time of the storage resonator is limited because of the coupling to the environment via this pin. At room temperature, we have five generators to address the system and two more for powering the amplifiers. The generators addressing the conversion resonator and the storage resonator are also combined to produce a tone close to the frequencies of the pumps thus phase locking the two modes with the pumps. The other three generators are used to address the pumping resonator, the tomography qubit and the tomography resonator.

III. METHODS

A. Sample fabrication

All the modes of the system are simulated using ANSYS HFSS and the Hamiltonian of the system is inferred using energy participation ratio black-box quantization technique [S9]. The cavity enclosure is machined into a single block of high purity aluminum in order to make a seamless re-entrant cavity [S10]. The transmons are fabricated
as Al/AlO$_x$/Al Josephson junctions on a c-plane double-side polished sapphire wafer using bridge-free electron beam lithography [S11]. The low-Q resonators are realized as stripline $\lambda/2$ resonators defined lithographically. The coupling pins shown in the Fig. 1d and Fig. S1 are coaxial couplers whose coupling strength is tuned by adjusting the length of their exposed pin.

B. Pump strength calibration

Accurate measurement of the rates $g_1$ and $g_2$ is necessary for comparing our experimental data for the Raman-assisted $|g4\rangle \leftrightarrow |f0\rangle$ oscillations with the theoretical predictions (Sec. I B) and the simulation results (Sec. IV). The rates $g_{1,2}$ are related to the pump strengths ($\xi_{1,2}$) as shown by Eq. (S6). However, the pumps experience frequency dependent coupling strengths and attenuation of the input lines, changing $\xi_{1,2}$ as a function of pump frequency. In order to eliminate this frequency dependence, we calibrate the pump strengths by measuring the Stark shift of the transmon mode caused by each individual pumps (while the other pump is off), when the pump frequencies are same as those used for addressing the $|g4\rangle \leftrightarrow |f0\rangle$ transition. Using Eq. (S7) we can relate the measured Stark shift to the pump strengths $\xi_{1,2}$ and by extension calibrate $g_{1,2}$ using Eq. (S6). The Stark shift due to pump 1 and pump 2 come out to be 5.15 MHz and 4.26 MHz respectively. This results in $g_1/2\pi = 0.53$ MHz and $g_2/2\pi = 0.48$ MHz.

C. Wigner tomography

The Wigner tomography of the storage resonator is performed in a similar manner to [S12, S13]. After preparing the storage resonator in the desired state, we apply a displacement pulse on the storage resonator, displacing it by $\beta$.

In this section we give the details of the simulation results presented in Fig. 3. The results are obtained by simulating the Lindblad master equation given by

$$\dot{\rho} = -\frac{i}{\hbar} [H_{sys}, \rho] + \kappa_a D[a] \rho + \kappa_b D[b] \rho + \Gamma_{\phi,b} D[a^\dagger a] \rho + \Gamma_{\phi,b} D[b^\dagger b] \rho.$$  

(S13)

Here $H_{sys}$ is the Hamiltonian of the system as quoted in Eq. S5. We also take into account the relaxation rates $\kappa_{a,b} = \frac{1}{\tau_{a,b}}$ and the dephasing of the modes $\Gamma_{\phi,a/b} = \frac{1}{\tau_{\phi,a/b}}$. The magnitude of all the quantities appearing in the Hamiltonian has been quoted in the main text except for the value of the global frequency shift $\delta$ which is found by using Eq. S10. From the resulting density matrix we find the populations of the various Fock states by tracing out the transmon (b) and also the $|g\rangle, |e\rangle, |f\rangle$ state populations of the transmon by tracing out the resonator (a). These populations are then scaled by the measurement contrast of the conversion transmon and the tomography transmon since the Fock-state populations are measured via the latter. The results of these simulations are plotted in Fig. 3 and they compare well with the experimental results.


FIG. S1. Wiring diagram.