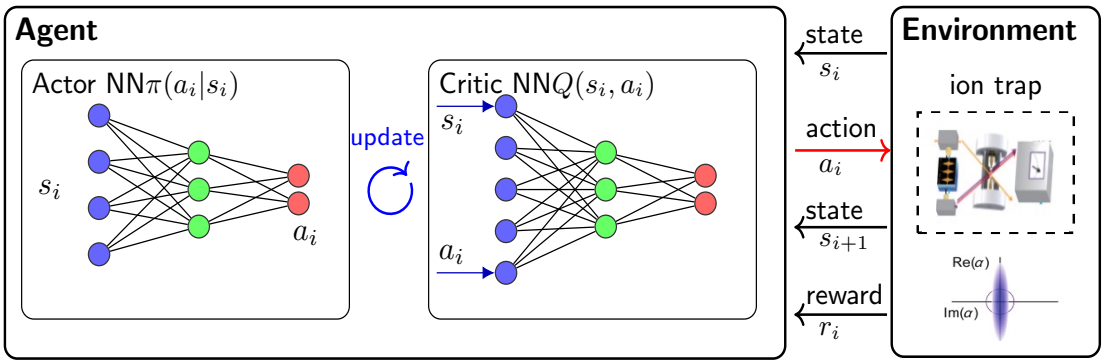
Supplemental Material for "Criticality-Enhanced Charging of Quantum Battery via Optimized Quantum Amplification"

Jiawei Zhang, Teng Liu, Mingshen Li, Qingshou Tan, Mang Feng, and Le Luo

Supplementary Note 1: Fast realization of squeezing and anti-squeezing using PPO reinforcement learning

In this work, we leverage reinforcement learning (RL) to accelerate both squeezed-state preparation and the subsequent anti-squeezing process directly within the framework of the standard quantum Rabi model described in the main text. Specifically, we implement the proximal policy optimization (PPO) algorithm [1], which is a widely adopted reinforcement learning method, particularly effective in optimizing policies within complex environments. The schematic representation of the PPO learning process is shown in Fig. S1. It improves upon traditional policy gradient methods by mitigating instability caused by large policy updates, imposing constraints on the step size during each update. Specifically, PPO introduces a "trust region" that limits deviations between the new and old policies through a carefully designed objective function. This function uses a clipping technique to prevent excessive updates, thereby balancing exploration and exploitation while enhancing learning stability and efficiency.



Supplementary Figure S1. Schematic representation of the PPO learning process. The PPO agent comprises two neural networks (NNs): the actor and the critic. At each time step i, the actor selects an action a_i from the current state s_i according to the policy $\pi(a_i|s_i)$. The environment, modeled as a trapped-ion system, evolves to a new state s_{i+1} and returns a reward r_i. This transition (s_i, a_i, r_i, s_{i+1}) is then used to update both the actor and critic networks, thereby improving the policy $\pi(a|s)$ and the value function $Q(s, a)$ for enhanced decision-making.

At the core of PPO is the objective function used for policy updates, which prevents large updates while optimizing expected rewards. The objective function is given by

$$L^{\mathrm{CLIP}}(\theta) = \mathbb{E}_t \left[\min \left(r_t(\theta)\hat{A}_t,\ \mathrm{clip}\left(r_t(\theta), 1-\epsilon, 1+\epsilon\right)\hat{A}_t \right) \right], \tag{S1}$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\mathrm{old}}}(a_t|s_t)}$ is the probability ratio between the new policy π_θ and the old policy $\pi_{\theta_{\mathrm{old}}}$, $\hat{A}_t$ is the advantage estimate at time step t, ϵ is a hyperparameter controlling the range of allowable updates, $\mathrm{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)$ restricts the probability ratio to the range $[1-\epsilon, 1+\epsilon]$, ensuring that updates remain close to the old policy. This objective function minimizes the original probability ratio $r_t(\theta)\hat{A}_t$ and its clipped version, ensuring that updates remain within the trust region. This approach enhances stability while allowing effective learning, with the clipping mechanism being a key feature of PPO, preventing large and destabilizing policy changes. In practice, an entropy bonus is often added to encourage exploration

$$L^{\mathrm{PPO}}(\theta) = L^{\mathrm{CLIP}}(\theta) + c_s S[\pi_\theta](s_t), \tag{S2}$$

where $S[\pi_\theta](s_t)$ is the entropy of the policy at state s_t, and c_s is a scaling factor that controls the weight of the entropy term. This encourages exploration by preventing premature convergence to suboptimal policies.

In our quantum battery charging and amplification protocol, reinforcement-learning control is employed in both stage I (squeezing preparation) and stage III (anti-squeezing). The total evolution time T_s is discretized into n

segments of equal duration $\delta t = T_s/n$, which define a sequence of control actions forming the RL trajectory. At each discrete time step t, the agent observes the system state $s_t = |\Psi(t)\rangle$, representing the instantaneous wavefunction of the ion-oscillator system, and selects an action $a_t \in \mathcal{A}$ according to the policy $\pi_\theta(a_t|s_t)$. The action determines the drive amplitude $\Omega_0(t)$, thereby updating the instantaneous coupling $g_c(t) = \eta\,\Omega_0(t)/\sqrt{\delta_b^2 - \delta_r^2}$. The environment (ion-oscillator system) then evolves unitarily under the effective Hamiltonian as $|\Psi(t+\delta t)\rangle = \exp\left(-i\hat{H}_{\text{eff}}\delta t\right)|\Psi(t)\rangle$, and returns a scalar reward that quantifies the progress toward the desired target state. The reward is defined through the instantaneous fidelity $F(t) = |\langle\Psi(t)\,|\,\Psi_{\text{target}}\rangle|^2$. During the squeezing stage, the target state is chosen as the squeezed vacuum $|\Psi_{\text{target}}\rangle = \hat{S}(r_{\text{target}})\,|0\rangle$, where r_{target} denotes the prescribed squeezing strength. In the anti-squeezing stage, the control objective is to restore the oscillator to the unsqueezed vacuum by setting $r_{\text{target}} = 0$, such that $|\Psi_{\text{target}}\rangle = |0\rangle$. The detailed learning process is illustrated in Fig. S1.

Through iterative PPO policy updates with entropy-regularized exploration, the agent learns an optimal sequence of control actions, corresponding to a nonadiabatic trajectory $g_c(t)$ that efficiently steers the system toward the target state. This training strategy consistently yields final preparation fidelities exceeding $F > 0.999$, demonstrating the effectiveness of the RL framework for fast and high-precision generation as well as reversal of strong squeezing. Here, we employed the Ray Python toolkit for training, with NN parameters set to Ray's default values, as summarized in Table I.

TABLE I. PPO algorithm hyperparameters

Hyperparameter	**Value**
Learning Parameters	
Learning rate (α)	1.0×10^{-4}
Training batch size	4000
Minibatch size	128
Discount factor (γ)	0.99
Policy Optimization	
Clipping parameter (ϵ)	0.3
GAE parameter (λ)	1.0
Value function loss coefficient	1.0
Entropy coefficient	0.0
KL divergence coefficient	0.2
Network Architecture	
Hidden layers	2×256
Activation function	ReLU
Weight initialization	Xavier
Training Setup	
Environment runners	2
Optimizer	Adam

Supplementary Note 2: Adiabatic preparation of squeezed and anti-squeezed states

For reference, we characterize the adiabatic approach to preparing the squeezed and anti-squeezed states required in our charging protocol. In the adiabatic case, we adopt a smooth sinusoidal ramp for the coupling strength,

$$g_c(t) = g_{\text{final}} \sin\left(\tfrac{\pi t}{2T_{\max}}\right), \tag{S3}$$

which suppresses diabatic transitions by ensuring slow variation throughout the evolution. Numerical simulations are performed using QuTiP with a Hilbert space truncation of $N = 100$.

The adiabatic preparation time is fundamentally constrained by the energy gap that closes as the system approaches the critical point. For squeezed state preparation, the minimal time required scales as $T \propto 1/(\gamma\epsilon(g))$ under the condition $\gamma \leq 0.1$, where $\epsilon(g)$ is the instantaneous energy gap. Consequently, as the target squeezing increases ($g_{\text{final}}^2 \to 1$), the required evolution time grows dramatically: for $g_{\text{final}}^2 = 0.99$, adiabatic preparation demands $T_{\max} = 100/\omega_0$; for $g_{\text{final}}^2 = 0.999$, this increases to $500/\omega_0$; and for $g_{\text{final}}^2 = 0.9997$, it reaches $1500/\omega_0$.

The anti-squeezing process (returning from a squeezed state to the vacuum) faces an even more severe time constraint. Starting from a highly squeezed initial state with an extremely small energy gap, the adiabatic condition becomes more stringent ($\gamma \leq 0.05$), leading to longer preparation times than for squeezing. Using a cosine ramp,

$$g_c(t) = g_{\text{init}} \cos\left(\tfrac{\pi t}{2T_{\max}}\right), \tag{S4}$$

the required durations scale accordingly: for initial $g^2_{\rm init} = 0.99$, $T_{\max} = 500/\omega_0$; for $g^2_{\rm init} = 0.999$, $T_{\max} = 1500/\omega_0$; and for $g^2_{\rm init} = 0.9997$, $T_{\max} = 2500/\omega_0$.

This dramatic increase in required time stems from the vanishing energy gap at criticality-a fundamental feature of second-order quantum phase transitions. As the system approaches the critical point, the timescale for adiabatic evolution diverges, making the preparation of highly squeezed states experimentally impractical. This limitation motivates the use of alternative control strategies, such as the reinforcement learning approach demonstrated in the main text, which bypasses the adiabatic constraint by exploiting nonadiabatic dynamics.

Supplementary Note 3: Comparative Analysis of Ergotropy Decay in Displaced vs. Squeezed Vacuum States

In this section, we present a theoretical analysis comparing two distinct energy storage strategies in quantum batteries: one utilizing displaced vacuum states and the other utilizing squeezed vacuum states. The primary motivation for transferring energy from squeezed to displaced states in practical implementations will be elucidated.

We consider a single bosonic mode quantum battery with Hamiltonian $\hat{H} = \omega_0(\hat{a}^\dagger\hat{a} + 1/2)$, which is weakly coupled to a thermal bath at inverse temperature β with dissipation rate γ. The system dynamics are governed by the Lindblad master equation:

$$\dot{\hat{\rho}} = -i[\hat{H}, \hat{\rho}] + \gamma(1+\bar{n})\mathcal{D}[\hat{a}]\hat{\rho} + \gamma\bar{n}\mathcal{D}[\hat{a}^\dagger]\hat{\rho}, \tag{S5}$$

where $\bar{n} = (e^{\beta\omega_0} - 1)^{-1}$ and $\mathcal{D}[\hat{L}]\hat{\rho} = \hat{L}\hat{\rho}\hat{L}^\dagger - \frac{1}{2}\{\hat{L}^\dagger\hat{L}, \hat{\rho}\}$.

To ensure a fair comparison, we consider initial states with equal ergotropy. Specifically: Displaced vacuum state: $\hat{\rho}_d(0) = \hat{D}(\alpha)|0\rangle\langle 0|\hat{D}^\dagger(\alpha)$, where $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$. Squeezed vacuum state: $\hat{\rho}_s(0) = \hat{S}(r)|0\rangle\langle 0|\hat{S}^\dagger(r)$, where $\hat{S}(r) = \exp\left[\frac{r}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)\right]$. The initial ergotropies are given by

$$\mathcal{E}_d(0) = \omega_0|\alpha|^2, \quad \mathcal{E}_s(0) = \omega_0 \sinh^2(r). \tag{S6}$$

Setting $\mathcal{E}_d(0) = \mathcal{E}_s(0)$ yields the relation $|\alpha|^2 = \sinh^2(r)$.

To quantitatively compare the discharge behavior of displaced and squeezed vacuum states, we follow the analytical framework established in Ref. [3]. We first introduce the thermal fluctuation strength, defined as $f(\beta) = \bar{n} + \frac{1}{2} = \frac{1}{2}\coth\left(\frac{\beta\omega_0}{2}\right)$, which characterizes the equilibrium thermal noise of the bath. For the initial squeezed vacuum state, the associated passive state is the vacuum itself, implying $f(\beta_\pi) = 1/2$. We also define the auxiliary function

$$\Delta_\beta(t) = f(\beta) + \left[\frac{1}{2} - f(\beta)\right] e^{-\gamma t}, \tag{S7}$$

which describes the relaxation of the thermal component in the absence of squeezing.

The displaced vacuum state evolves as a displaced thermal state with a time-dependent displacement amplitude

$$\alpha(t) = \alpha e^{-(i\omega_0 + \gamma/2)t}.$$

Consequently, its ergotropy decays exponentially according to

$$\mathcal{E}_d(t) = \omega_0 \sinh^2(r)\, e^{-\gamma t}. \tag{S8}$$

In contrast, the squeezed vacuum state exhibits a more complex evolution due to the nontrivial dynamics of its covariance matrix. The effective thermal fluctuation strength at time t is given by

$$f(\beta_t^s) = \sqrt{\Delta_\beta(t)^2 + 2f(\beta)e^{-2\gamma t}(e^{\gamma t} - 1)\sinh^2(r)}.$$

The corresponding ergotropy evolves as

$$\mathcal{E}_s(t) = \omega_0\left[\Delta_\beta(t) + e^{-\gamma t}\sinh^2(r) - f(\beta_t^s)\right]. \tag{S9}$$

To quantify the difference in discharge rates, we define

$$D(t) = \mathcal{E}_d(t) - \mathcal{E}_s(t) = \omega_0\left[f(\beta_t^s) - \Delta_\beta(t)\right]. \tag{S10}$$

From the explicit form of $f(\beta_t^s)$, we observe that for $t > 0$ and $r > 0$, $f(\beta_t^s) > \sqrt{\Delta_\beta(t)^2} = |\Delta_\beta(t)|$. Since $\Delta_\beta(t) > 0$ for all $t \geq 0$ (as $f(\beta) \geq 1/2$), it follows that $|\Delta_\beta(t)| = \Delta_\beta(t)$. Therefore, $f(\beta_t^s) > \Delta_\beta(t)$ for $t > 0$, $r > 0$, and consequently

$$D(t) = \omega_0 \left[f(\beta_t^s) - \Delta_\beta(t) \right] > 0 \quad \text{for } t > 0. \tag{S11}$$

This inequality confirms, as illustrated in Fig. S2, that for identical initial ergotropy, the displaced vacuum state retains more extractable work at any finite time than the squeezed vacuum state. The faster discharge of squeezed states originates from the non-monotonic evolution of their passive-state energy. Although the total energy of both states decays exponentially, the passive energy $E[\pi(t)] = \omega_0 f(\beta_t^s)$ for the squeezed state initially rises due to the interplay between squeezing and thermal noise. Because ergotropy is defined as $\mathcal{E} = E[\rho] - E[\pi]$, this transient increase in passive energy accelerates the depletion of extractable work, explaining why squeezed states discharge more rapidly than displaced states.

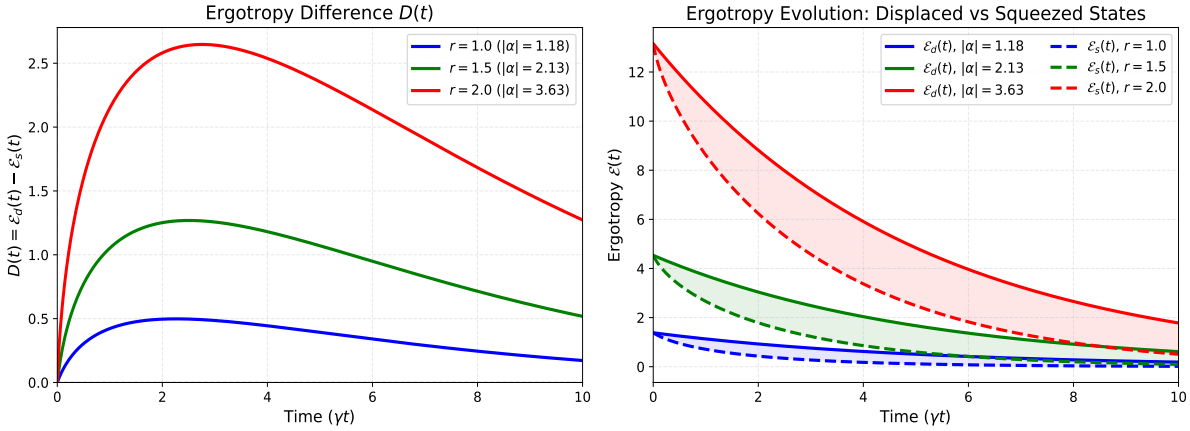


FIG. S2. Comparative of ergotropy decay in displaced vs. squeezed vacuum states. Here the parameters are given as $\gamma/\omega_0 = 0.2, \beta/\omega_0 = 0.5$.

Supplementary Note 4: Experimental details

The phonon heating rate of our system is approximately 61 quanta/s, measured using the following procedure: First, the ion is cooled down to an average phonon number of $\bar{n} = 0.04$ via Doppler cooling and sideband cooling. After a variable delay time to allow for heating, a blue-sideband transition is applied. The average phonon number is extracted by fitting the resulting blue-sideband excitation curve using the probability for the nth Fock state, which is $P(n) = \frac{1}{\bar{n}+1}\left(\frac{\bar{n}}{\bar{n}+1}\right)^n$. The heating rate is then determined from the slope of a linear fit of the average phonon number versus heating time. The experimental results are presented in Fig. S3(a,b).

To characterize the coherence of phononic states, we employ Ramsey interferometry. The Ramsey fringes are measured via spin population after two $\pi/2$ blue-sideband pulses with a variable delay τ. The fringes are then fitted with the function $Ae^{-t/T}\cos(2\pi\delta t) + C$, where A, T, and C are fitting parameters, and T represents the coherence time. The coherence time is estimated to be 5.18 ms, and the corresponding experimental data and fit are shown in Fig. S3(c).

Supplementary Note 5: The fitting details

To probe the probability distribution across the phonons of the ion's motional state, we apply a blue-sideband pulse to the final state. Using fluorescence detection, we measure the time-dependent population $P_{|e\rangle}(t)$ of the excited state $|e\rangle = |^2S_{1/2}, F = 1, m_F = 0\rangle$ under blue-sideband driving. The detection probability for the excited state is given by Ref. [2]

$$P_{|e\rangle}(t) = \frac{1}{2} - \sum_{n=0}^{\infty} P_n e^{-\gamma_n t} \cos\left(\Omega_{n+1,n} t\right), \tag{S12}$$

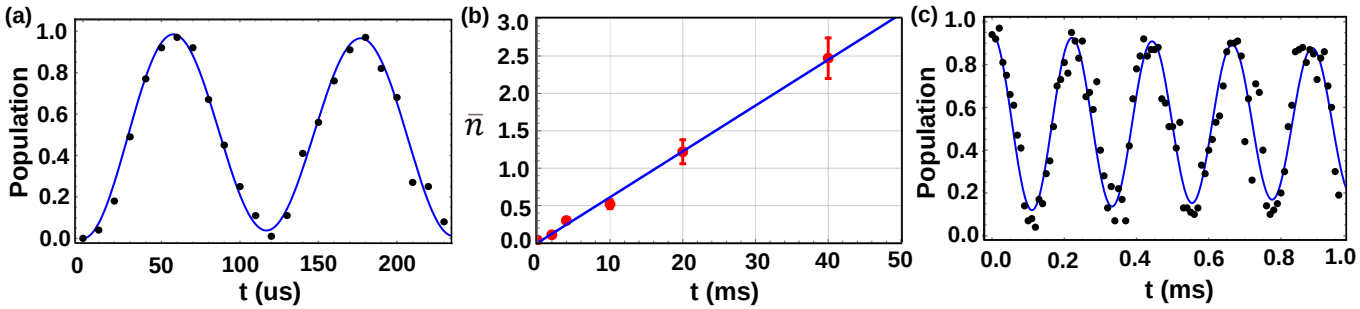


FIG. S3. Blue sideband Rabi oscillation, heating rate, and Ramsey. (a) The blue sideband Rabi oscillation used for average phonon number fitting. The black dots are the experimental data, and the blue line is the fitting curve. (b) The measurement of the heating rate. The red dots represent the experimental data with error bars, and the blue line is the linear fitting curve, whose slope is the heating rate. (c) The Ramsey used for coherence time fitting. The black dots represent the experimental data, and the blue line is the Ramsey fitting curve.

where P_n denotes the population in the n-th Fock state and $\Omega_{n+1,n}$ represents the Rabi frequency between the n-th and $(n+1)$-th Fock states. At low phonon numbers, $\Omega_{n+1,n}$ is well approximated by $\Omega_{n+1,n} \simeq \eta\Omega\sqrt{n+1}$, where $\Omega = 2\pi \times 55.7$ kHz is the carrier Rabi frequency on the ground state transition, and $\eta \approx 0.14$ is the Lamb–Dicke parameter under our experimental conditions. The terms γ_n are phenomenological decay rates that are assumed to follow $\gamma_n = \gamma\sqrt{n+1}$, reflecting their dependence on the phonon number n.

In Eq. (S12), distinct motional states correspond to different probability distributions P_n across Fock states. By fitting the measured blue-sideband spectrum, we can determine the population distribution of the motional state, from which parameters such as displacement and squeezing can be extracted.

By applying the displacement operator $\hat{D}(\alpha) = \exp\left(\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right)$ to the phononic ground state $|0\rangle$, a coherent state is generated, given by the expression

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle, \tag{S13}$$

where $\hat{a}^\dagger$ and $\hat{a}$ represent the creation and annihilation operators, respectively, and α denotes the complex displacement parameter. The resulting probability distribution across phonon number states is described by

$$P_n = |\langle n \mid \alpha\rangle|^2 = \frac{e^{-|\alpha|^2}|\alpha|^{2n}}{n!}. \tag{S14}$$

For a squeezed state $|\zeta\rangle$, where the uncertainty of one quadrature is reduced below the standard quantum limit of a coherent state, the state is given by

$$|\zeta\rangle = \frac{1}{\sqrt{\cosh r}}\sum_{n=0}^{\infty}\frac{e^{in\theta}(-\tanh r)^n\sqrt{(2n)!}}{2^n n!}|2n\rangle, \tag{S15}$$

where the squeezing parameter $\zeta = re^{i\theta}$ is complex, with r denoting the squeezing strength and θ the squeezing angle. The state exhibits a non-zero probability distribution solely over even-numbered Fock states, while the probability on odd-numbered states is identically zero,

$$P_{2n} = \frac{(\tanh r)^{2n}}{\cosh r}\frac{(2n)!}{\left(2^n n!\right)^2}. \tag{S16}$$

For a state that is first squeezed by parameter ζ and then displaced by α, the resulting population distribution over Fock states can be expressed as

$$P_n = |\langle n \mid \alpha, \zeta\rangle|^2 = \frac{\left(\frac{1}{2}\tanh r\right)^n}{n!\cosh r}\left|H_n\left(\frac{\alpha\cosh r + \alpha^* e^{i\theta}\sinh r}{\sqrt{e^{i\theta}\sinh(2r)}}\right)\right|^2 \exp\left[-|\alpha|^2 - \frac{1}{2}\left(\alpha^{*2}e^{i\theta} + \alpha^2 e^{-i\theta}\right)\tanh r\right]. \tag{S17}$$

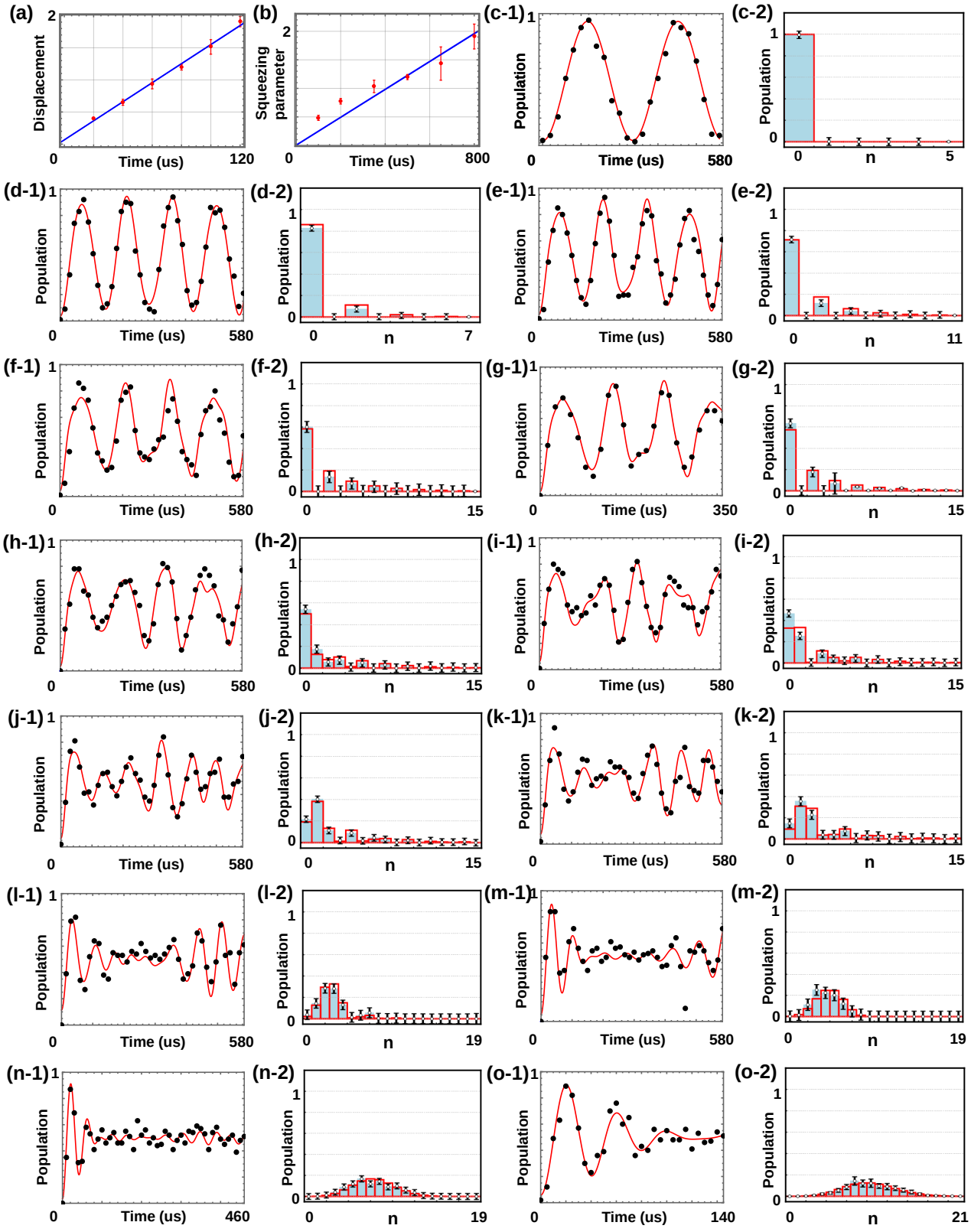


FIG. S4. Experimental results and fitting. (a) Variation of displacement amplitude with laser interaction duration. (b) Variation of squeezing strength with laser interaction duration. (c-1)-(o-1) Blue sideband oscillation curves and their corresponding fits at various nodes of the squeeze-displace-antisqueeze sequence. In each panel, experimental data are denoted by black points, while fitting curves are represented by red lines. (c-2)-(o-2) The fitted P_n distributions corresponding to the fitting curves in (c-1)-(o-1). In each panel, red hollow columns represent the theoretical populations of different Fock states, while blue solid columns indicate the fitted populations. Black open circles with error bars denote the corresponding standard errors.

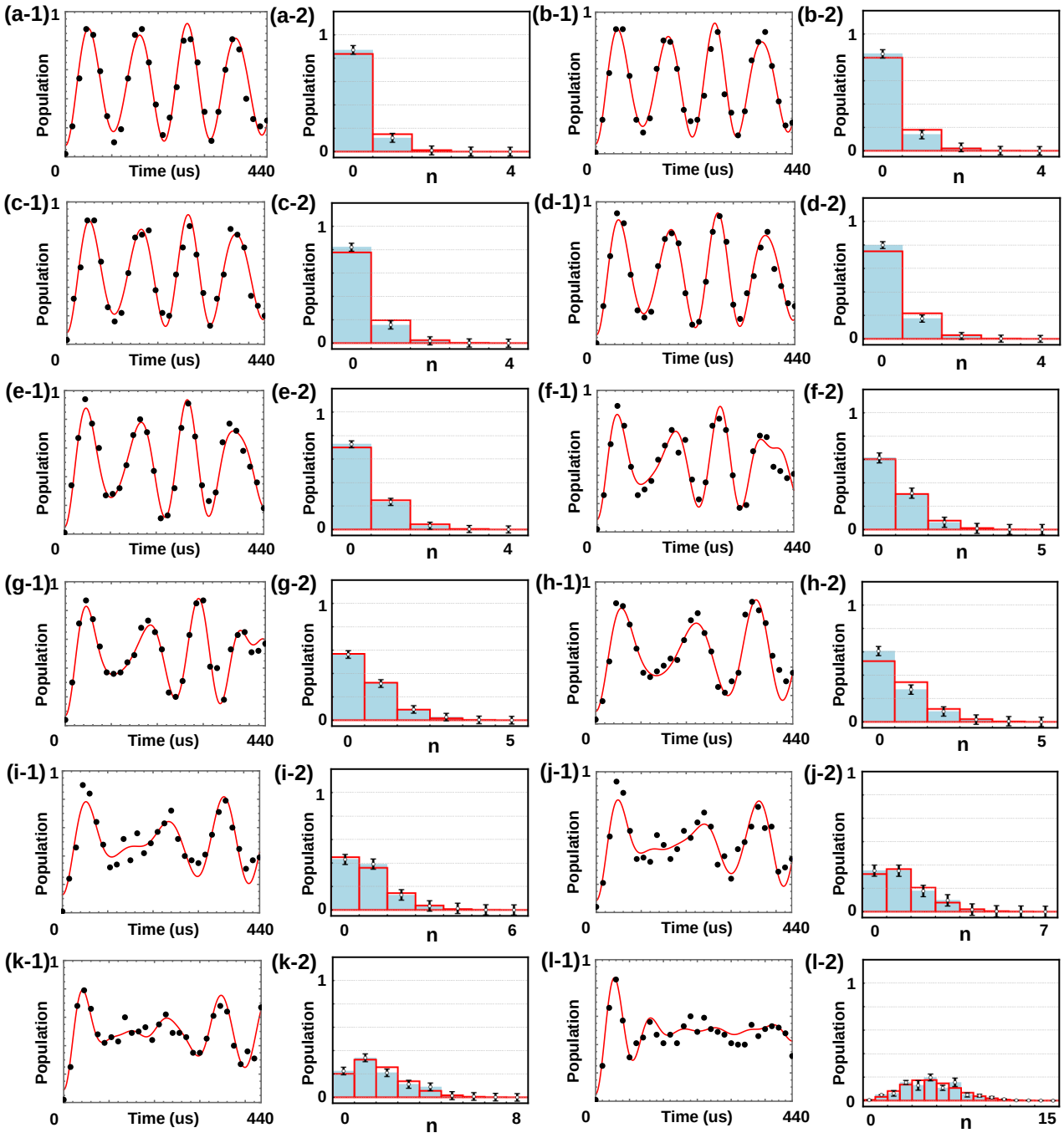


FIG. S5. Experimental results and fitting for displacement amplification. (a-1)-(l-1) Blue sideband oscillations and their corresponding fits under I-XII squeezing sequences and amplifications. The black points are experimental results, and the red lines are fitting curves. (a-2)-(l-2) The fitted P_n distributions (Q_n) corresponding to the fitting curves in (a-1)-(l-1). In each panel, red hollow columns represent the theoretical populations of different Fock states, while blue solid columns indicate the fitted populations. Black open circles with error bars denote the corresponding standard errors.

Using this method, we have first verified the linear time-dependence of both the displacement and squeezing magnitude by simultaneously applying the nearly resonant red and blue sidebands (Figs. S4(a) and (b)). We then measure the phonon number distribution P_n at 13 distinct time points (nodes) throughout the charging process, as illustrated in Fig. 2(a) of the main text. Figs. S4(c–g) show the blue-sideband evolution data (with fits) and the reconstructed distributions for five nodes during the squeezing stage. Similarly, Figs. S4(h–k) and (l–o) present the same for each of the four nodes during the displacement and anti-squeezing stages, respectively. The fits are

generated using P_n from Eq. (S16) (Figs. S4(c–g)), Eq. (S17) (Figs. S4(h–k)), and numerical simulation (Figs. S4(l–o)). Nonlinear fitting of the experimental data points is then performed in *Mathematica*. The fitting parameters for each Q_n (the fitted probability distribution) are constrained within ±30% of their theoretical values P_n. The maximum phonon number for fitting is truncated where the theoretical probability distribution falls below 1%, corresponding to $n = \{5, 7, 11, 15, 15, 15, 15, 15, 15, 19, 19, 19, 21\}$ for the respective cases. The error bars represent the standard errors σ_n associated with each fitted parameter Q_n. Notably, in the final dataset shown in Fig. S4, multicollinearity becomes particularly pronounced due to the large number of phonon energy levels included in the fitting under large displacement conditions. We employ a bootstrap fitting approach to evaluate the standard error for each Q_n.

Since our experiment does not capture the complete phase information (the off-diagonal elements of the density matrix) of the motional Fock states, we employ a fidelity evaluation method that relies solely on comparing the diagonal elements. Specifically, we reconstruct a diagonal density matrix as the experimental density matrix ρ_{exp} based on the fitted values of Q_n, and then compute the fidelity $F(\rho, \sigma) = \left(\text{tr}\sqrt{\sqrt{\rho_{\text{exp}}}\rho\sqrt{\rho_{\text{exp}}}}\right)^2$ in combination with the theoretical density matrix ρ, which retains only its principal diagonal. The error σ_F of the fidelity F originates from the error propagation of the standard errors σ_n of each Q_n involved in the calculation, according to the formula $\sigma_F = \sqrt{F\sum_n \frac{P_n}{Q_n}\sigma_n^2}$.

For the Wigner function calculation and plot in Fig. 2(b) of the main text, we employ an approach that combines fitting parameters with theoretical calculations. Based on the density matrix of the phononic state, $\hat{\rho} = |\psi\rangle\langle\psi| = \sum_{m,n}\rho_{m,n}|m\rangle\langle n|$, the value of the Wigner function at the phase space point (x, y) is

$$W(x,y) = \frac{2}{\pi}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\rho_{m,n}W_{m,n}(x,y), \tag{S18}$$

where $\rho_{m,n} = c_m c_n^*$, with m and n truncated at the maximum phonon number. The functional form of $W_{m,n}(x, y)$ is

$$W_{m,n}(x,y) = (-1)^m e^{-\left(x^2+y^2\right)}\sqrt{\frac{m!}{n!}}[\sqrt{2}(x+iy)]^{n-m}L_m^{(n-m)}\left(2\left(x^2+y^2\right)\right), \quad for \quad n \geq m \tag{S19}$$

and

$$W_{m,n}(x,y) = (-1)^n e^{-\left(x^2+y^2\right)}\sqrt{\frac{n!}{m!}}[\sqrt{2}(x-iy)]^{m-n}L_n^{(m-n)}\left(2\left(x^2+y^2\right)\right), \quad for \quad n < m, \tag{S20}$$

where $L_a^{(b)}(x)$ denotes the associated Laguerre polynomials. Due to the absence of phase information in the experimental data, we use QuTiP to theoretically calculate the phase information for the Fock states. This calculated phase information is then combined with the fitted Fock state probabilities Q_n (i.e., $\rho_{n,n}$) obtained from the experimental data shown in Fig. S4.

Supplementary Note 6: The concrete sequence and final experimental result versus different coupling strengths

We present the squeezing and anti-squeezing sequences used in the experiments in Table II and Table III, respectively. Each row in the tables represents one sequence, and each sequence is divided into several segments (denoted as seg.). The squeezing sequence consists of 10 segments, each with a duration of $0.4/\omega_0$, while the anti-squeezing sequence consists of 20 segments, each with a duration of $0.2/\omega_0$. The tables list the g^2 value corresponding to each segment. During the experiment, the corresponding sideband Rabi frequency $\Omega_{r(b)}$ is calculated using the formula $\Omega_{r(b)} = g\sqrt{\delta_b^2 - \delta_r^2}$, and the mapping between the theoretical sequence and the experimental optical intensity is achieved by calibrating the relationship between the Rabi frequency Ω and the amplitude of the AOM signal that controls the corresponding laser intensity.

Using an arbitrary waveform generator, we executed the above sequences. The experimental results and their corresponding fits are presented in Fig. S5. The final displacements α_f and state fidelities were measured to be [0.426±0.046, 0.470±0.040, 0.493±0.041, 0.600±0.061, 0.607±0.028, 0.745±0.043, 0.775±0.063, 0.769±0.047, 0.910±0.057, 1.100±0.106, 1.300±0.051, 2.042±0.206] and [0.998±0.078, 0.996±0.085, 0.997±0.072, 0.996±0.061, 0.998±0.067, 0.998±0.106, 0.998±0.089, 0.992±0.098, 0.998±0.126, 0.997±0.138, 0.991±0.106, 0.982±0.075], respectively.

TABLE II. The Squeezing Sequence

Sequence \ Segment (g^2)	1	2	3	4	5	6	7	8	9	10
I ($g_c^2 = 0.2$)	0.386	0.297	0.136	0.460	0.336	0.504	0.530	0.426	0.602	0.496
II ($g_c^2 = 0.5$)	0.386	0.297	0.136	0.460	0.336	0.504	0.530	0.426	0.602	0.496
III ($g_c^2 = 0.6$)	0.970	0.100	0.887	0.100	0.510	0.634	0.911	0.413	0.907	0.606
IV ($g_c^2 = 0.7$)	0.946	0.100	0.891	0.204	0.590	0.634	0.931	0.423	0.917	0.739
V ($g_c^2 = 0.8$)	0.990	0.100	0.990	0.100	0.721	0.897	0.674	0.739	0.825	0.806
VI ($g_c^2 = 0.9$)	0.950	0.172	0.990	0.990	0.065	0.990	0.786	0.801	0.990	0.810
VII ($g_c^2 = 0.92$)	0.620	0.950	0.160	0.980	0.980	0.400	0.980	0.600	0.940	0.950
VIII ($g_c^2 = 0.94$)	0.660	0.960	0.400	0.970	0.970	0.410	0.990	0.620	0.960	0.960
IX ($g_c^2 = 0.96$)	0.688	0.960	0.590	0.990	0.990	0.409	0.990	0.624	0.990	0.960
X ($g_c^2 = 0.98$)	0.988	0.990	0.990	0.990	0.990	0.409	0.990	0.624	0.990	0.980
XI ($g_c^2 = 0.99$)	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.250	0.990	0.990
XII ($g_c^2 = 0.999$)	1.999	2.000	0.937	1.189	0.900	1.100	1.000	0.200	0.941	0.999

TABLE III. The Anti-squeezing Sequence

Sequence \ Segment (g^2)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
I ($g_c^2 = 0.2$)	0.260	0.420	0.450	0.650	0.100	0.250	0.100	0.250	0.250	0.250	0.250	0.960	0.150	0.150	0.150	0.100	0.000	0.000	0.000	0.000
II ($g_c^2 = 0.5$)	0.100	0.163	0.100	0.100	0.990	0.424	0.100	0.100	0.355	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
III ($g_c^2 = 0.6$)	0.100	0.100	0.100	0.193	0.139	0.802	0.854	0.303	0.100	0.406	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IV ($g_c^2 = 0.7$)	0.392	0.214	0.100	0.764	0.575	0.990	0.100	0.627	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
V ($g_c^2 = 0.8$)	0.100	0.121	0.443	0.465	0.000	0.928	0.990	0.698	0.154	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
VI ($g_c^2 = 0.9$)	0.125	0.915	0.371	0.375	0.796	0.850	0.900	0.900	0.900	0.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
VII ($g_c^2 = 0.92$)	0.120	0.850	0.250	0.250	0.980	0.960	0.960	0.960	0.900	0.950	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
VIII ($g_c^2 = 0.94$)	0.120	0.890	0.400	0.650	0.980	0.960	0.960	0.940	0.900	0.950	0.500	0.450	0.450	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IX ($g_c^2 = 0.96$)	0.120	0.890	0.530	0.650	0.980	0.980	0.980	0.940	0.900	0.950	0.680	0.880	0.880	0.000	0.000	0.000	0.000	0.000	0.000	0.000
X ($g_c^2 = 0.98$)	0.100	0.980	0.308	0.765	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.000	0.000	0.000	0.000	0.000	0.000
XI ($g_c^2 = 0.99$)	0.715	0.471	0.492	0.782	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.000	0.000	0.000
XII ($g_c^2 = 0.999$)	0.100	0.100	0.840	0.919	2.000	0.451	1.152	2.000	0.802	0.100	2.000	1.340	0.100	2.000	0.811	1.494	0.100	2.000	2.000	0.100

References

[1] Schulman, J. et al. Proximal Policy Optimization Algorithms. arXiv:1707.06347 (2017).

[2] Burd, S. C. et al. Quantum amplification of mechanical oscillator motion. *Science* **364**, 1163 (2019).

[3] Medina, I. et al. Anomalous discharging of quantum batteries: The ergotropic Mpemba effect. *Phys. Rev. Lett.* **134**, 220402 (2025).