# Nonlinear bosonic Maxwell's demon by coupling to qubits

Check for updates

Atirach Ritboon [1,2] ✉ & Radim Filip [2]

The concept of Maxwell's demon, proposed in classical physics as a means to extract work from a thermodynamics system beyond the constraints set by the second law of thermodynamics, has since been extended to modern quantum physics. Realization of the quantum Maxwell's demon remains of actual interest given the potential of continuous-variable systems in quantum thermodynamics and current experimental opportunities. We propose a quantum Maxwell's demon method, based on a Jaynes-Cummings two-level system, for subtracting bosonic energy inferred from successful measurements of excited qubits after linear and nonlinear interactions. The effect of these subtractions can suppress the tails of bosonic noise better than the linear interactions alone. The system statistics reaches an out-of-equilibrium state, becoming much closer to Poissonian distributions as indicated by the mean-to-noise ratio. The inclusion of a few additional optimal nonlinear subtractions can improve the success rate to ten times higher than the linear scheme, making the method significantly more efficient in exciting hundreds of qubits.

Maxwell's demon, a thought experiment, demonstrates that if one could access information about the state of a system through a classical measurement, then one can exploit such information to gain mechanical work or energy from the system through classical control over it. This thought experiment leads to the generalization of the second law of thermodynamics by emphasizing the possibility of information-work conversion. It is one of the vital principles that rectify thermal fluctuations, without using strong nonlinearity, simply by measurement and classical control[1]. In classical Maxwell's demon, the measurement is ideally arbitrarily precise, and the back action on the system does not need to be considered, as a measured quantity of a system is treated as a hidden variable. It changes dramatically when the measurement cannot give full information about the system's state. The outcome of the measurement thus implies the new state of the system through the inference of the gained information. The different couplings to a probe and its subsequent measurement form new states of the system. Such events often turn out to be destructive; however, they sometimes can conditionally distill the system into a more useful resource[2,3], as broadly explored in quantum information, especially in resource theories. If the distillation fails, it can still be repeated continuously until obtaining the resource. In this way, the chance of failure, in principle, can be minimized to zero at the cost of the speed or protocol multiplexing. The initial thermal fluctuations are used at the output with a negligible probability. Practically, it does not significantly influence the generated resource. The quantum transformation to a better resource, therefore, depends on the coupling and

measurement back action[4–6]. First, we need to define free states, free couplings, and free measurements[7,8]. The free states are naturally thermal equilibrium states diagonal in the energy eigenbasis. The essential free unitary coupling of the system to the probe is then energy-conserving; the free probe energy measurement commutes with the probe's energy. Free-controlled operations are also energy-conserving couplings with an ancilla in thermal states that maximally have the energy of the input ones. It defines the most basic but nontrivial playground to explore and compare Maxwell demon methods mutually and also with other noise rectification strategies.

In the resource theory, the usefulness of a quantum state is characterized operationally by some groups of physical and implementable processes that cannot generate the given resource, such as local operation and classical communication (LOCC), identifying the entanglement resource[7,9]. Quantifiers that are monotonically non-increasing under such physical processes are called resource monotones. In the case of continuous-variable systems, one can consider state transformations under Gaussian thermal operation to identify resource monotones such as temperature-like quantities, generalizing the equilibrium temperature[10]. The work of Serafini et al.[11], on the other hand, provides a full characterization of possible single-mode transformations through Gaussian thermal operation leading to a no-go theorem, preventing lowering the entropy of a single-mode below the background with algorithmic cooling.

The quantum Maxwell demon is more involved and diverse for a bosonic system representing a single mode of photons, phonons, or other

[1]Research Unit in Energy Innovations and Modern Physics (EIMP), Thammasat University, Khlong Nueng, Khlong Luang, Pathum Thani 12120, Thailand.
[2]Department of Optics, Palacky University, 17. listopadu 12, 771 46 Olomouc, Czech Republic. ✉e-mail: ritboon@tu.ac.th

bosonic particles. Here, the simplest case of a free coupling is an energy-conserving beam-splitter type of resonant coupling. After this beam-splitter coupling, macroscopic measurement integrating energy already allows conditional manipulation with continuous energy statistics, for example, used in the work of Iskhakov et al.[12,13]. Microscopic single-quanta detection opens space for subtracting individual energy quanta conditionally[14–21], even for macroscopic thermal states, and charging the macroscopic battery by average energy[22,23].

However, for microscopic phononic states with few quanta on average, the statistics after subtraction become crucial for charging a microscopic battery. Such a battery is represented by a two-level system coupled to the phonons, light, or microwave fields. Multiple subtractions increase mean energy and reduce autocorrelation between quanta, causing them to be more statistically independent[17,24]. They mainly increase the mean-to-deviation ratio of the system's energy, which is essential for information theory and thermodynamics[25]. Moreover, as recently demonstrated, the correlations between two thermal baths allow Maxwell's demon-based protocol to extract more work[26], and measurement strategies have been used in quantum memristors[27].

In this work, we propose a nonlinear bosonic Maxwell's demon working at the quantum level through the simple and deterministic protocol, which is expected to be straightforwardly realized in various quantum platforms. We first investigate this deterministic Maxwell demon method for a broadly feasible energy-conserving coupling, a linear Jaynes–Cummings (JC) coupling[28], probing a bosonic system sequentially by two-level systems to reach an out-of-equilibrium state. We then prove that the output state can excite another two-level system better than any thermal state. Differently from photonic Maxwell's demon, we consider phononic systems represented, for example, by the extensively used mechanical modes of a single atom[29] or, recently, a macroscopic oscillator[30]. Alternatively, microwave superconducting experiments can also be considered for the experimental tests[31]. In these cases, usually, the mean thermal occupation per mode can be much higher than that of thermal light sources. Despite the low dimension of the probes, the deterministic linear subtraction increases both the mean-to-deviation ratio of energy and the probability of exciting two-level systems higher than that from thermal states. It proves the power of such operations beyond a conventional Fock state lowering[32]. The remaining noise can be further suppressed by a sharper measurement using still energy-conserving nonlinear JC coupling[33–35] available at trapped-ion platforms[36,37], cavity quantum electrodynamics[38], and superconducting circuits[39] mentioned above to perform a nonlinear subtraction of more quanta at once. Trilinear interactions, additionally, can also considered as alternative options[40–43].

We prove that optimally implementing nonlinear subtractions after the linear ones increases both the mean-to-deviation ratio of energy and the probability of excitation of atoms. Such improvement after a few nonlinear subtractions becomes distinctly significant, more than 10 times improved, for successfully exciting hundreds of qubits. It is, to our knowledge, the first example showing that two-quanta processes can bring the statistics of phononic mode closer to a Poissonian, without any classical external drive and intense nonlinear saturation, which is typical for such processes in laser[37]. It proves that Maxwell's demon, based on available nonlinear energy-conserving couplings, can open a new territory for quantum statistical and thermodynamical investigations.

## Results
### Overview

To understand the overall picture and procedure of this work, we devote this subsection to explain the overview of the proposed protocol employed to gradually shape the probability distribution of a harmonic oscillator close to a Poissonian for a better probability of exciting a two-level system, which is regarded as a quantum battery, via JC interaction. Figure 1 displays the overall processes of the scheme. A harmonic oscillator in thermal equilibrium with a thermal bath undergoes a linear excitation subtraction using a
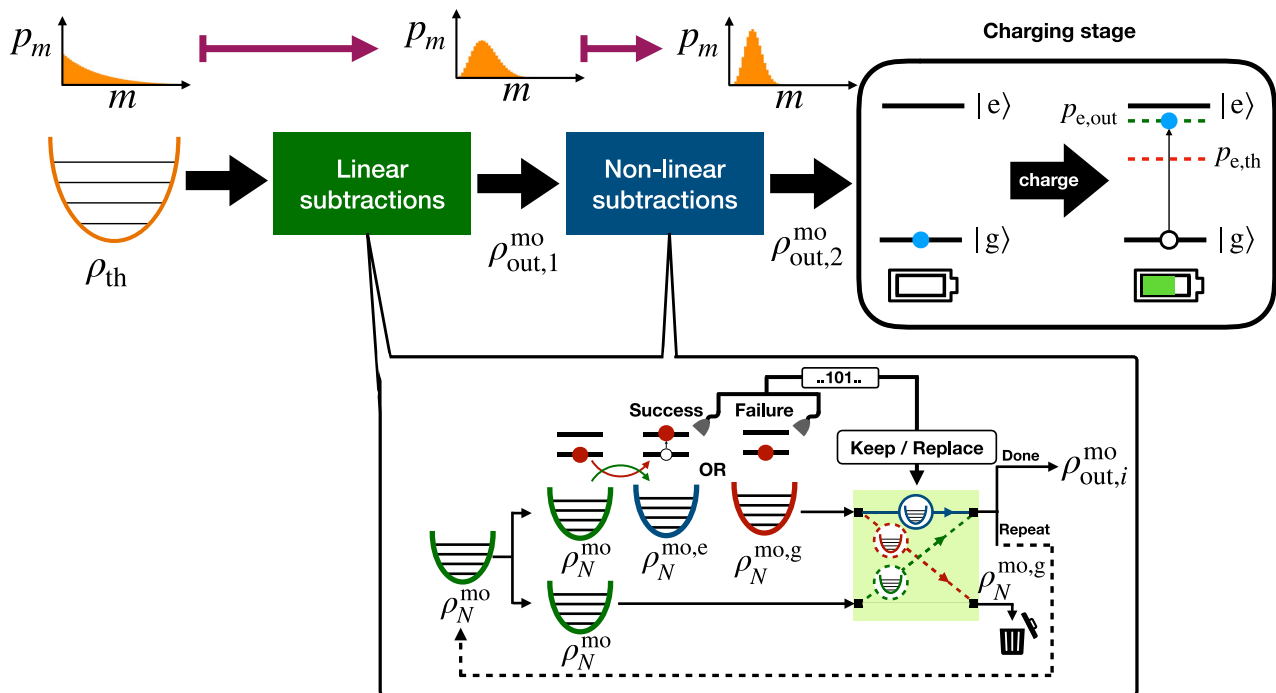


**Fig. 1 | The scheme of the overall protocol.** The diagram represents the overview of the deterministic nonlinear bosonic protocol for resonantly charging a quantum battery, represented by a two-level system. The protocol is divided into two parts: phonon subtractions and the two-level charging process. The excitation subtraction also consists of two parts: linear (green) and nonlinear (blue) subtractions using protocol II, illustrated in the bottom inset and described in detail in the latter part of the subsection "Linear subtraction" in Results. Several linear subtractions are performed to gradually shape the probability distribution of a harmonic oscillator, initially in a thermal state $\rho_{th}$, into a nearly Gaussian distribution before nonlinear subtractions take place to trim and squeeze the probability distribution further. Such harmonic oscillators, hence, are used to charge the quantum batteries, initially being set to be in the ground state $|g\rangle$, through Jaynes-Cummings interaction. The excitation probability $p_{e,out}$ is expected to exceed its corresponding thermal bound.

linear JC coupling for several times before further extracting its excitations through the nonlinear interaction so that its probability distribution becomes even more squeezed from both sides at low and high quanta. The output harmonic oscillator is then used to charge a two-level battery through the linear JC coupling to demonstrate that it outperforms the thermal bound, the maximum probability of exciting a qubit with a thermal state $\rho_{th}$, and the initial state it is associated with. This improvement originated from the successive change in the population distribution of the harmonic oscillators as a result of sequential subtractions of the oscillators' excitation. In the limit of a high initial mean excitation $\bar{n}$ of the oscillators, only linear subtraction cannot give the charging performance higher than the combination of both linear and nonlinear subtractions. For example, for $\bar{n} = 70$, using only a sequence of linear subtractions gives rise to the highest charging performance of $P_e = 0.9745$ after 42 subtractions, while using a shorter sequence of 15 linear and 5 nonlinear subtractions gives rise to the performance higher at $P_e = 0.9784$. This difference is more visible once we independently and completely excite a hundred atoms, as 42 linear subtractions give the probability of 0.076 compared to 0.11 using 15 linear and 5 nonlinear subtractions. The involvement of nonlinear subtractions allows a sufficiently high charging performance to be achieved with a significantly smaller number of required subtractions.

The lower inset depicts the excitation subtraction procedure of both linear and nonlinear subtractions. A harmonic oscillator in a motional state $\rho_N^{mo}$ resonantly interacts with a two-level system through a JC interaction, with the interaction Hamiltonian linear in oscillator variables, or a nonlinear JC interaction[34] having higher powers of the oscillator variables in the interaction Hamiltonian depending on which type of excitation subtraction being performed at that stage. After they interact for a strategically chosen period of time, the state of the qubit is measured. If the measurement results in the excited state, the subtraction is thus performed successfully, otherwise, it fails. The measured outcome then feeds forward to decide whether the harmonic oscillator could be kept or replaced by its previous successful version before repeating every step again in the next round of subtraction.

## Linear subtraction

The linear subtraction is performed using a resonant JC interaction whose Hamiltonian in the interaction picture can be written as

$$\hat{H}_{int} = \hbar\lambda\left(\hat{\sigma}_+ \hat{a} + \hat{\sigma}_- \hat{a}^\dagger\right), \tag{1}$$

where $\lambda$ is the coupling strength, $\hat{\sigma}_+$ ($\hat{\sigma}_-$) is the rising (lowering) operator of the qubit, and $\hat{a}^\dagger$ ($\hat{a}$) is the creation (annihilation) operator of the oscillators. We then can express the unitary operator associated with this coupling

running for an interval time $t$ as

$$\begin{aligned}\hat{U}_{JC}(t) =\ & |e\rangle\langle e| \cos(\lambda t\sqrt{\hat{n}+1}) + |g\rangle\langle g| \cos(\lambda t\sqrt{\hat{n}}) \\ & - i\hat{\sigma}_- \hat{a}^\dagger \frac{\sin(\lambda t\sqrt{\hat{n}+1})}{\sqrt{\hat{n}+1}} - i\hat{\sigma}_+ \hat{a} \frac{\sin(\lambda t\sqrt{\hat{n}})}{\sqrt{\hat{n}}},\end{aligned} \tag{2}$$

where $|g\rangle\langle g|$ and $|e\rangle\langle e|$ denote the projections on the ground and excited states of the qubits, respectively, and $\hat{n} = \hat{a}^\dagger\hat{a}$ is the number operator of the harmonic oscillators.

The first proposed protocol, denoted as protocol I, for linearly subtracting the motional excitations is schematically illustrated in Fig. 2. At the beginning, harmonic oscillators are in thermal equilibrium with a thermal bath at temperature $T$, while the probes, two-level systems, are prepared in the ground state $|g\rangle$. The initial composite state of this can be expressed as

$$\rho_0 = |g\rangle\langle g| \otimes \rho_{th}, \tag{3}$$

where $\rho_{th}$ denotes the state of a harmonics oscillator in thermal equilibrium with a mean number of excitations $\bar{n}$,

$$\rho_{th} = \sum_{m=0}^{\infty} \frac{\bar{n}^m}{(\bar{n}+1)^{m+1}} |m\rangle\langle m| \equiv \rho_0^{mo}, \tag{4}$$

and is also regarded as the initial motional state of the oscillators.

The mean number of motional excitations $\bar{n}$ is related to the temperature $T$ by $\bar{n} = (\exp(\hbar\omega/k_B T) - 1)^{-1}$, where $\omega$ is the angular frequency of the oscillators. The interaction between the qubits and the oscillators is run for the optimal time, $t_0^{op}$, chosen to maximize the probability of exciting the qubits, approximately related to $\bar{n}$ as $\lambda t_0^{op} \approx \pi/(2\sqrt{\bar{n}+1})$ (see the Methods section more details). Subsequently, the measurement in the eigenbasis $\{|g\rangle, |e\rangle\}$ on these qubits is performed. We then postselect only those oscillators with the probes in the excited state $|e\rangle$ to be used in the further steps of the protocol. The measurement and postselection project the state of the qubits and the harmonic oscillators onto

$$\begin{aligned}\rho_1'(t_0^{op}) =\ & \frac{1}{P_e^{(0)}(t_0^{op})} \left(\hat{\sigma}_+ \hat{a} \frac{\sin(\lambda t_0^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}} \rho_0 \frac{\sin(\lambda t_0^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}} \hat{a}^\dagger \hat{\sigma}_-\right) \\ =\ & |e\rangle\langle e| \otimes \sum_{m=0}^{\infty} \frac{\bar{n}^{m+1}}{(\bar{n}+1)^{m+2}} \frac{\sin^2(\lambda t_0^{op}\sqrt{m+1})}{P_e^{(0)}(t_0^{op})} |m\rangle\langle m| \\ =\ & |e\rangle\langle e| \otimes \rho_0^{mo,e},\end{aligned} \tag{5}$$



**Fig. 2 | The scheme of protocol I.** The diagram illustrates the proposed phonon subtraction protocol. In the $N$th round of subtraction, an oscillator in a state $\rho_N^{mo}$ couples with a two-level system at its ground state for the optimal interaction time $t_N^{op}$. Then a measurement in the energy basis of the qubit is performed. If the excited state $|e\rangle$ is obtained, we keep the oscillator (blue), as its energy is successfully subtracted, for the next round, otherwise replacing it with a new oscillator in the initial thermal state.
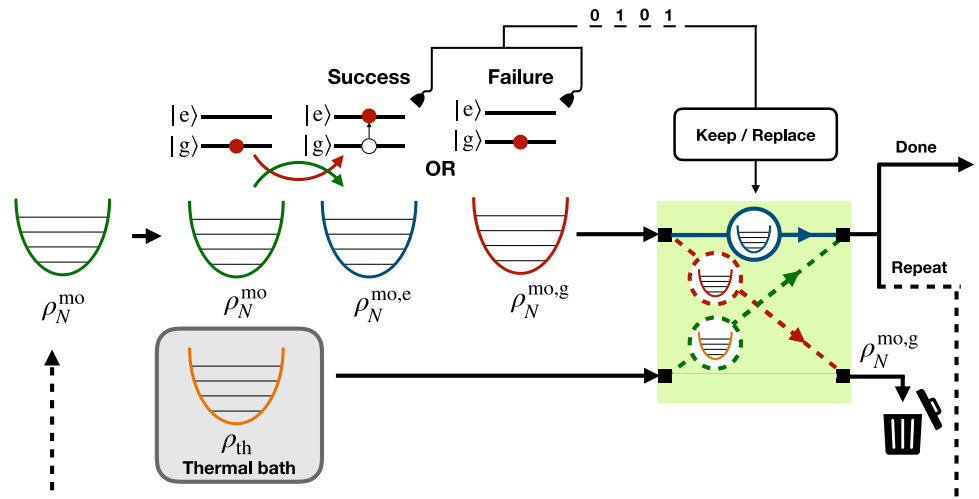
where $P_e^{(0)}(t_0^{op})$ is the probability of observing the excited state $|e\rangle$ at the optimal time, which acts as the normalization factor of the term in the bracket and $\rho_0^e$ is the state of the oscillators after the postselection. The qubits of those postselected systems are then reset back to their ground state $|g\rangle$ by the dissipation of their energy to the environment. We consider that the qubit dissipation is much slower than the JC interaction and, therefore, does not decohere the process. We mention here that the excited qubits are no longer coupled to harmonics oscillators after the measurement, but coupled with the radiation modes instead to erase them. The qubit thus relaxes its excitation to the cold optical environment by spontaneous photon emission. It is essential for ground qubit state cooling of trapped atoms (and superconducting qubits), reaching nearly 100%. We also would like to emphasize that, like others, Maxwell's demons, this proposed protocol does not invalidate the second law of thermodynamics as it requires a resource to reset the probe to its ground state[44].

After that, we replace those failed systems with new systems in the initial state $\rho_0$. All mentioned processes are then repeated again, but this time, the initial motional state of the ensemble for the new round has changed from $\rho_0^{mo}$ due to the measurement back action of the first subtraction. For the $N$th round of repeat-until-success subtraction by the JC interaction, the state of the ensemble can be expressed as

$$\rho_N^{mo} = \hat{a}\frac{\sin(\lambda t_{N-1}^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}}\rho_{N-1}^{mo}\frac{\sin(\lambda t_{N-1}^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}}\hat{a}^\dagger + (1 - P_e^{(N-1)}(t_{N-1}^{op}))\rho_0^{mo}, \quad (6)$$

where $\rho_{N-1}^{mo}$ is the achieved motional state of the previous subtraction, $P_e^{(N-1)}$ is the probability of getting the excited state $|e\rangle$ in the previous round, and $t_{N-1}^{op}$ is the interaction time that maximizes $P_e^{(N-1)}$. Note that the motional state in Eq. (6) when $N = 1$ differs from the state in Eq. (5) by additional terms associated with the repeat-until-success subtractions as studied by Marek et al.[45]. The probability $P_e^{(N)}$ of getting an excited state at the optimal interaction time $t_N^{op}$ can be written as

$$P_e^N(t_N^{op}) = \sum_{m=0}^{\infty} p_m^{(N)}\sin^2(\lambda t_N^{op}\sqrt{m}). \quad (7)$$

where the probability distribution $p_m^{(N)}$ of the harmonic oscillators in the state $\rho_N$ can be expressed as

$$p_m^{(N)} = p_{m+1}^{(N-1)}\sin^2(\lambda t_{N-1}^{op}\sqrt{m+1}) + (1 - P_e^{(N-1)}(t_{N-1}^{op}))p_m^{(0)}, \quad (8)$$

We denote $p_m^{(N-1)}$ to be the achieved probability distribution in the previous round with $p_m^{(0)} = \bar{n}^m/(\bar{n}+1)^{m+1}$, the initial probability distribution at thermal equilibrium. This equation describes how each subtraction gradually shapes the probability distribution of the oscillators in each round.

Let us consider the semiclassical case when the average excitation is very large, $\bar{n} \gg 1$. The square of the sine function, the first term in Eq. (8), acts as a population filter. After the first subtraction, $N = 1$, with the optimal time $\lambda t_0^{op} \approx \pi/(2\sqrt{\bar{n}+1})$, the probabilities $p_n^{(1)}$ with the value of $n$ very different from the initial average excitation $\bar{n}$ are suppressed as the values of $\sin^2(\pi\sqrt{m+1}/(2\sqrt{\bar{n}+1}))$ is considerably smaller than unity, while those probabilities $p_m^{(1)}$ with $m$ close to $\bar{n}$ dominate the new probability distribution. This filtering effect still holds true for the subsequent subtractions, and it also makes the optimal interaction times of several further rounds are approximately the same as the first one: $\lambda t_N^{op} \approx \lambda t_0^{op} \approx \pi/(2\sqrt{\bar{n}+1})$ for $N \ll \bar{n}$. As the probability of having the qubit in the excited state grows progressively with the number of performed subtractions, see the detailed discussion in subsection "Charging performance" in Results, the last term in Eq. (8) gradually becomes a smaller contribution. Each subtraction of the motional excitation, with optimized coupling, thus gradually modifies the probability distribution $p_m^{(N)}$ into a Gaussian distribution centered around $\bar{n}$. The center of the probability distribution becomes noticeably shifted toward the motional ground state, when the number of subtractions becomes comparable with the initial average excitation number $\bar{n}$, indicating that the average excitation number decreases slightly each time we perform a subtraction.

Protocol I can be further improved, if in the last step of each subtraction, instead of replacing the failed systems with systems in thermal equilibrium, $\rho_0^{mo}$, we replace them with the successfully achieved systems of the previous round. The diagram of the second protocol, named protocol II, is depicted in Fig. 3. The achieved state of the ensemble after the $N$th round of subtraction becomes

$$\rho_N^{mo} = \hat{a}\frac{\sin(\lambda t_{N-1}^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}}\rho_{N-1}^{mo}\frac{\sin(\lambda t_{N-1}^{op}\sqrt{\hat{n}})}{\sqrt{\hat{n}}}\hat{a}^\dagger + (1 - P_e^{(N-1)}(t_{N-1}^{op}))\rho_{N-1}^{mo}, \quad (9)$$

where the population distribution of the harmonic oscillators is modified as

$$p_m^{(N)} = p_{m+1}^{(N-1)}\sin^2(\lambda t_{N-1}^{op}\sqrt{m+1}) + (1 - P_e^{(N-1)}(t_{N-1}^{op}))p_m^{(N-1)}. \quad (10)$$

This modification can suppress both tails of the probability distribution faster and better than the protocol I at the cost of collecting and storing the outcomes of the previous steps. The nearly Gaussian statistics of phonons produced by protocols I and II are illustrated and compared in Fig. 4.

We note here that the required conditions to have the statistics of the population close to a Gaussian distribution are both sufficiently high initial excitation $\bar{n}$, i.e. high temperature and a sufficient number of subtractions by linear Jaynes-Cummings model with protocol II. A few rounds of such subtractions are insufficient to make the distribution symmetric. A low



**Fig. 3 | The scheme of protocol II.** The diagram of the improved protocol is similar to that of protocol I shown in Fig. 2, but in this case, the failed systems are replaced by the successful systems obtained in the previous round instead of using systems in the initial thermal state.
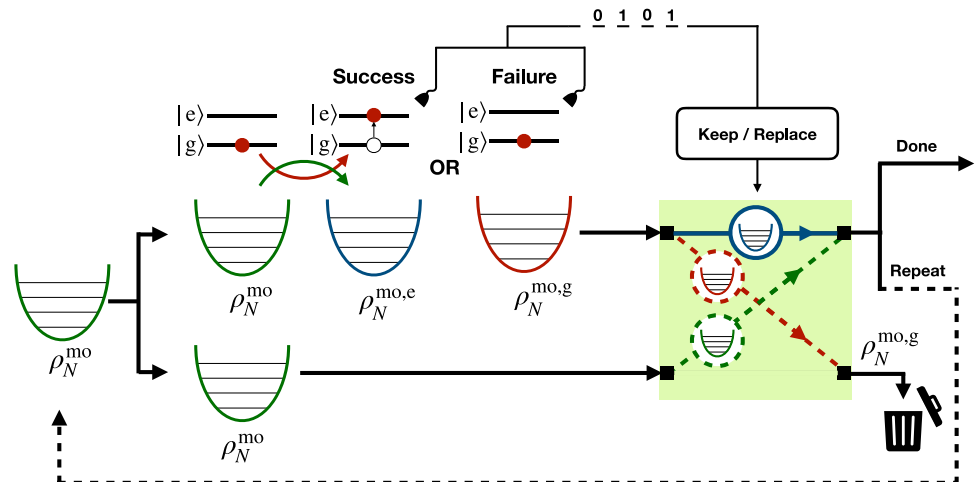
**Fig. 4 | The phonon statistics after linear subtractions.** The change in the population distributions after five linear subtractions using **a** protocol I and **b** protocol II is demonstrated. The black solid lines represent the initial population distribution of the oscillators being in thermal equilibrium with the mean excitation number $\bar{n} = 30$. Sequential subtractions gradually form a nearly Gaussian distribution probability distributions (orange), with its peak located slightly lower than the initial mean excitation $\bar{n} = 30$. The distribution obtained from protocol II, in (**b**), is noticeably narrower than that obtained from protocol I but has slightly lesser mean excitation, $\langle \hat{n} \rangle$. These two distributions are then compared to Poisson distributions (blue) with the same average phonon numbers. Their relevant information is also given, including their mean excitation $\langle \hat{n} \rangle$, their second-order correlation functions $g^2(0)$, their mean-to-deviation ratios of excitation $\mathcal{R}$, defined in Eq. (13), and their Fano factors $F$.
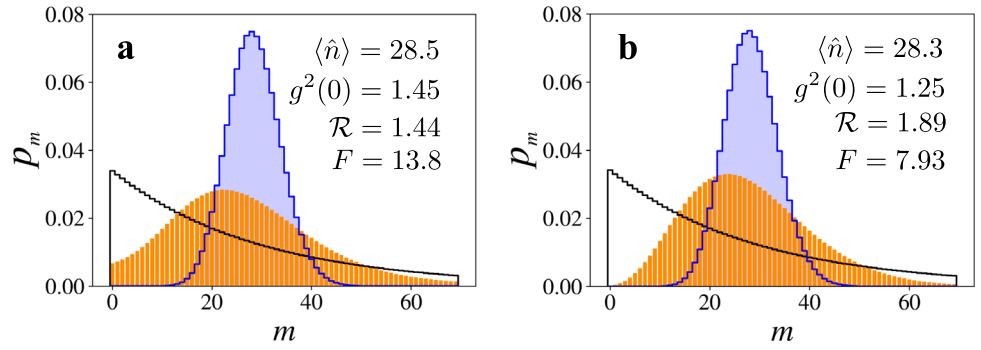
value $\bar{n}$, on the other hand, cannot allow enough rounds of linear subtraction, as they would deplete the excitation of the oscillator and bring it close to the ground state instead.

We assume that the thermalization time of the motional state is very long compared to the total time spent in all processes of protocols I and II so that the heat transferred from the thermal bath to the considered oscillators is very small and negligible. The thermalization effect, as a result, can be ignored.

**Charging performance**

A population inversion of qubits happens when the probability $P_e$ of finding the qubits in the excited state $|e\rangle$ exceeds the probability $P_g$ of finding them in the ground state $|g\rangle$, i.e., $P_e - P_g > 0$ or $P_e > 1/2$. We then devote this part to demonstrate and explain the performance of sequential linear subtractions. Let us first discuss the relation between the population distribution of a harmonic oscillator and the maximum excitation probability of a two-level system. For a qubit initially being in the ground state $|g\rangle$, the probability $P_e$ of getting the excited state $|e\rangle$ after it is coupled an oscillator via the JC interaction for $t$ is

$$P_e(t) = \sum_{m=0}^{\infty} p_m \sin^2(\lambda t \sqrt{m}), \qquad (11)$$

where $p_m$ is the population distribution of the oscillator. As an oscillator being in its motional ground state $|0\rangle$ cannot be coupled with a qubit in the ground state $|g\rangle$, the probability $P_e$ thus must be smaller than $1 - p_0$, where $p_0$ is the probability of finding the oscillator in its ground state. This means the desired probability distribution should have a small probability $p_0$. For a weak coupling case in which $\lambda t \ll 1$ so that $p_m \sin^2(\lambda t \sqrt{m}) \sim p_m(\lambda t)^2 m \ll 1$, the probability approaches a linear rule $(\lambda t)^2 \langle \hat{n} \rangle < 1$ and the statistics of the oscillator do not matter in the classical excitation limit. For stronger coupling, however, this simple approximation breaks. Each term in the summation oscillates in time differently depending on its index $m$ due to the sine function. Narrower probability distributions thus give a constructive summation of those terms in the summation, as they cause smaller mis-match between the oscillating time dependence of the dominant probabilities $p_m$ and provide a higher chance of getting the excited state $|e\rangle$. For example, the perfect scenario in which the excited state $|e\rangle$ is obtained via the JC interaction for certain is when the oscillator is in an arbitrary Fock state $|n\rangle$ as we can just choose the interaction time $t$ precisely to match $\lambda t \sqrt{n} = \pi/2$, leading to $P_e = 1$. Another factor to be considered for a Gaussian-like distribution, with probabilities $p_m$ falling shapely when being far from the peak, is the mean excitation number, $\langle \hat{n} \rangle$. If two distributions have an identical Gaussian-like shape but different mean excitation numbers, the one with a larger mean excitation can give larger $P_e$. A larger $\langle \hat{n} \rangle$ provides smaller mis-match between the oscillating time dependence $\lambda t \sqrt{m}$ of the sine function in the summation. When $m$ differs from the mean excitation number $\langle \hat{n} \rangle$ by $\Delta m$ such that $m = \langle \hat{n} \rangle + \Delta m$ and $|\Delta m| \ll \langle \hat{n} \rangle$, the oscillating time dependence of the probability $p_m$ can be approximated as

$$
\begin{aligned}
\lambda \sqrt{m} &= \lambda \sqrt{\langle \hat{n} \rangle + \Delta m} \\
&\approx \lambda \sqrt{\langle \hat{n} \rangle} + \frac{\lambda \Delta m}{2\sqrt{\langle \hat{n} \rangle}}.
\end{aligned}
\qquad (12)
$$

The difference in the time dependence of the dominant probabilities in the summation of Eq. (11) is, therefore, inversely proportional to $\sqrt{\langle \hat{n} \rangle}$. Of course, the mean excitation becomes irrelevant when it comes to the case of an excited Fock state, as demonstrated earlier that, with a single oscillating term in the summation, $P_e = 1$ can be obtained for certain regardless of the mean excitation number. However, we need to bear in mind that the statistics will immediately play a crucial role once there exists a small deviation from Fock states. From these discussed facts, among the parameters commonly used for analyzing the statistics of excitation, such as second-order correlation function, $g^2(0) = \langle \hat{a}^{\dagger 2} \hat{a}^2 \rangle / \langle \hat{a}^\dagger \hat{a} \rangle^2$, and Fano factors $F = \langle (\Delta \hat{n})^2 \rangle / \langle \hat{n} \rangle$, the appropriate parameter indicating the desirable phonon statistics, motivated by Eq. (12), should be the mean-to-deviation ratio (MDR) of the population, denoted by $\mathcal{R}$, which is defined as

$$\mathcal{R} = \frac{\langle \hat{n} \rangle}{\sqrt{\langle (\Delta \hat{n})^2 \rangle}} \qquad (13)$$

where $\langle (\Delta \hat{n})^2 \rangle = \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2$ represents the phonon fluctuation. It is more likely that the atom is excited better by a phonon with a greater value of $\mathcal{R}$.

Figure 5 shows the maximum probability of having the excited state $|e\rangle$ increases after each linear subtraction. This increase originates from the fact that the population distribution becomes narrower after each subtraction. However, for a small initial mean excitation number, around $\bar{n} \sim 1\text{–}2$, the subtractions do not always increase $P_e^{(N)}$ as the excitations of the oscillators are almost exhausted, i.e., most oscillators are in their motional ground state and no longer coupled with the two-level systems. As expected, since protocol II gives a smaller probability of being in the motional ground state, $p_0$, and a narrower probability distribution, it then gives higher probabilities $P_e^{(N)}$, for $N \geq 2$. From the figure, we can clearly see that the increase of $P_e^{(N)}$ gradually becomes saturated, as the value of $\Delta P_e^{(N)} = P_e^{(N)} - P_e^{(N-1)}$ becomes smaller. Further subtractions barely increase the excitation probability. It is obvious from the figure that the saturated value of $P_e^{(N)}$ obtained

**Fig. 5 | The relation of charging performance and the initial phonon number $\bar{n}$ of a thermal oscillator after subtractions.** The figure demonstrates the increase of the maximum probability $P_\text{e}^{(N)}$ of getting the excited state $|e\rangle$ through the JC coupling after the $N$th subtraction, when using **a** protocol I, depicted in Fig. 2, and **b** protocol II, shown in Fig. 3. The probability $P_\text{e}^{(N)}$ increases sharply with the initial mean excitation $\bar{n}$ and reaches its plateau around $\bar{n} \sim 5$. Protocol II apparently provides a better chance of exciting a two level-system compared to protocol I. The dotted lines are set to be at 0.5 to mark population inversion $P_\text{e}^{(N)} > 0.5$, while the dash-dotted lines display the excitation probability if the oscillator was in a coherent state. A thermal state with a very large mean excitation, $\bar{n} \gg 1$, can excite a qubit with probability of $P_\text{e} \approx 0.6411$.
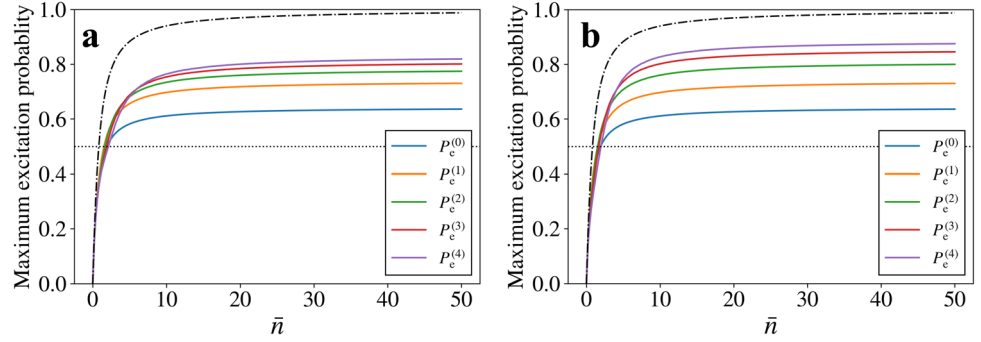
from protocol II is slightly greater than that from protocol I. This is because protocol II shapes the distribution in such a way that the probability $p_0$ and its neighborhood become very small, as shown in Fig. 4b, compared to the distribution obtained from protocol I, depicted in Fig. 4a.

For a large number $\bar{n}$, the performance of charging a two level system in an $N$th round of operation, for $N \leq 20$, can be roughly approximated as $P_\text{e,I}^{(N)} \approx 0.84 - 0.20e^{-0.5N} + 0.003N$, for protocol I, and $P_\text{e,II}^{(N)} \approx 0.93 - 0.29e^{-0.4N} + 0.002N$ for protocol II, where $P_\text{e,I}^{(N)}$ ($P_\text{e,III}^{(N)}$) is the charging performance in the $N$th round using protocol I (protocol II). The charging performance from both protocol in this case is displayed in Fig. 6. These approximations explicitly quantify how the performance of the proposed protocols improved with the number of operation rounds $N$, to portray the improvement of the performance in each round compared to the cost of resources using for running the protocols.

## Nonlinear subtraction

As we pointed out in Fig. 4, the distribution still has a long decaying tail for larger populations resembling the thermal statistics. To remove this limitation and shape the population distribution even faster and better, linear subtraction alone is no longer sufficient. From Fig. 6, the performance of the linear subtractions eventually will reach its saturation, but there is still a way to break through it by utilizing a nonlinear interaction whose interaction Hamiltonian is of the form,

$$\hat{H}_\text{non} = \hbar\lambda' \left( \hat{\sigma}_+ \hat{a}^2 + \hat{\sigma}_- \left(\hat{a}^\dagger\right)^2 \right), \tag{14}$$

where $\lambda'$ denotes the coupling strength of the interaction. This interaction Hamiltonian would have the same form as the Hamiltonian in Eq. (1) if the annihilation and creation operators in Eq. (1) were replaced by their squares, $\hat{a} \to \hat{a}^2$ and $\hat{a}^\dagger \to \left(\hat{a}^\dagger\right)^2$. The unitary operator describing the time evolution of this nonlinear coupling is given as

$$\hat{U}_\text{non}(t) = |e\rangle\langle e| \cos\left(\lambda' t \sqrt{(\hat{n}+1)(\hat{n}+2)}\right) + |g\rangle\langle g| \cos\left(\lambda' t \sqrt{\hat{n}(\hat{n}-1)}\right)$$
$$- i\sigma_- \left(\hat{a}^\dagger\right)^2 \frac{\sin\left(\lambda' t \sqrt{(\hat{n}+1)(\hat{n}+2)}\right)}{\sqrt{(\hat{n}+1)(\hat{n}+2)}}$$
$$- i\sigma_+ \hat{a}^2 \frac{\sin\left(\lambda' t \sqrt{\hat{n}(\hat{n}-1)}\right)}{\sqrt{\hat{n}(\hat{n}-1)}}. \tag{15}$$

As mentioned, nonlinear repeat-until-success subtractions are performed in the same way as linear subtractions explained in subsection "Linear subtraction" in Results, except this time, the employed interaction becomes nonlinear. However, unlike the linear case, a nonlinear subtraction cannot trim the tails of the population distribution desirably, see
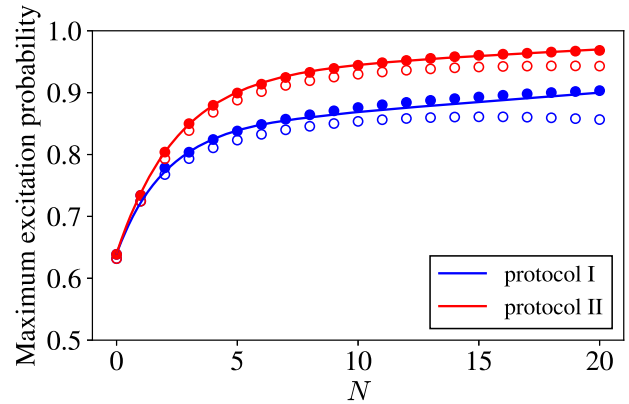


**Fig. 6 | The comparison of the charging performance using protocols I and II.** The charging performance using protocol I (blue) and protocol II (red) when the initial mean excitations are $\bar{n} = 30$ (empty circles) and $\bar{n} = 70$ (filled circles) after $N$ linear subtractions, with $N \leq 20$, is displayed. The solid lines represent the asymptotic approximations, when $\bar{n} \gg 1$. As these protocols are not yet different for $N < 2$, the blue circles, in this case, are placed behind the red circles.

Supplementary Note 1 for more details. Therefore, the prior population distribution should be similar to a Gaussian to some extent, and the probabilities $p_m$ associated with high-energy levels, $m \gg \bar{n}$, must be already very small. Otherwise, the nonlinear subtraction will form a ripple in the probability distribution of the harmonic oscillators, which is an undesired effect. Due to the normalization condition, the probability distribution with a ripple is more dispersed compared to those without it. The probability $P_\text{e}$, as a result, is not as large as it potentially should be.

Consequently, in order to use nonlinear subtractions properly, we have to perform several linear subtractions using protocol II, displayed in Fig. 3, so that the population distribution is modified close to a Gaussian with sufficiently short tails. After that, nonlinear subtractions using the procedure of protocol II can then be performed. The achieved state after a nonlinear subtraction can be expressed as

$$\rho^\text{f} = \hat{a}^2 S(\hat{n}) \rho^\text{i} S(\hat{n}) \left(\hat{a}^\dagger\right)^2 + \left(1 - P'_\text{e}(\tau^\text{op})\right)\rho^\text{i}, \tag{16}$$

with

$$S(\hat{n}) = \frac{\sin\left(\lambda' \tau^\text{op} \sqrt{\hat{n}(\hat{n}-1)}\right)}{\sqrt{\hat{n}(\hat{n}-1)}} \tag{17}$$

where $\rho^\text{i}$ is the former state of the oscillators before the nonlinear subtraction, $\tau^\text{op}$ is the interaction time that gives the first locally optimal probability of successful subtraction. The probability of successful subtraction, on the

other hand, reads

$$P'_e(t) = \sum_{m=2}^{\infty} p_m^i \sin^2\left(\lambda' t \sqrt{m(m-1)}\right), \quad (18)$$

where $p_m^i$ is the prior probability distribution of the harmonic oscillators, the diagonal elements of $\rho^i$. In contrast to linear subtractions, the suitably chosen interaction time for nonlinear subtractions is not the time that optimizes the probability $P'_e$. From the equation, it is easy to notice that the optimal interaction time is approximately at $\lambda' t = \pi/2$, resulting in $\sin^2\left(\lambda' t\sqrt{m(m-1)}\right) \approx 1$ for $m > 1$. However, with this interaction time, a nonlinear subtraction almost does not modified the probability distribution at all, as the sine function barely shows its influence. We instead need to choose the interaction time $\tau^{op}$ that gives the first locally optimal $P'_e$ such that $\lambda'\tau^{op} < \pi/2$. The interaction time $\tau^{op}$ for this case is approximately related to the mean phonon number $2\langle\hat{n}\rangle$ as

$$\lambda'\tau^{op} \approx \frac{\pi}{2\langle\hat{n}\rangle}. \quad (19)$$

This means we can run a nonlinear interaction even faster than a linear one. A nonlinear subtraction using protocol II manipulates the probability distribution as

$$p_m^f = p_{m+2}^i \sin^2\left(\lambda'\tau^{op}\sqrt{(m+2)(m+1)}\right) + (1 - P'_e(\tau^{op}))p_m^i. \quad (20)$$

Figure 7 compares the probabilities of success and their change after each subtraction of two different schemes: the scheme that employs only linear subtractions (the upper bar chart) and the scheme that uses linear subtractions followed by nonlinear subtractions (the lower bar chart) to boost the performance before charging the quantum battery. From the upper bar chart, when only linear subtractions are being used, the sequential increase of the probability of success gradually reaches its saturation.

Therefore, additional linear subtractions just barely improve the probability of success and the charging performance, denoted by the last green bar. The lower bar chart, on the other hand, demonstrates that the saturated performance can be further boosted with the help of nonlinear subtractions, whose probabilities of success are represented by the two red bars. The success probability of a nonlinear subtraction is noticeably lower than that of the previous linear subtractions, but it is still sufficiently large to make the nonlinear subtraction protocol practical. After six linear and two nonlinear subtractions, the charging performance or the probability of getting the excited state through the JC coupling, denoted by the green bar, becomes even larger than its previous version when all eight subtractions are linear. This increase lies in the change in the shape of the population distribution after the nonlinear subtractions. As shown in the inset of Fig. 7, the distribution becomes more squeezed, which is more desirable for exciting a two-level system.

We note here that a nonlinear subtraction cannot be used with the procedure of protocol I, as at the end of the protocol, the failed systems are replaced by systems in a thermal state, which makes high-energy populations not sufficiently small. As a result, nonlinear subtractions cause the population distribution to be even more dispersed and a small ripple in the distribution to form.

## Discussion

We previously compared the charging performance obtained from the two subtraction strategies: the linear-subtractions-only strategy and the combination of linear and nonlinear subtractions. It emerged that the latter provides a better charging performance compared to the first. The remaining question is whether nonlinear subtractions should be performed at an earlier stage at the very end of a subtraction sequence, or even something in between, to get the optimal charging performance. To answer this, we then examined different combinations of linear and nonlinear subtractions, where the coupling is optimized for each subtraction depending on the previous measurement outcomes. The result turns out to be that the later the nonlinear subtractions take place, the greater the



**Fig. 7 | The improved charging performance after using nonlinear subtractions. a** The bar chart represents the probabilities of success to obtain the excited state in $N$th subtractions using protocol II with an initial average phonon number of $\bar{n} = 30$. **b** On the other hand, the bar chart below also shows such probabilities, but this time, the 7th and 8th subtractions are performed by two consecutive nonlinear subtractions, with their probabilities of success denoted with the red bars. The actual probabilities of success, $P_e$, are written in white on these bars so that we can see their fractional differences. Above each bar, a probability distribution of the harmonic oscillator contributing to the probability of success is displayed together with the value $\mathcal{R}$ of the mean-to-deviation ratio (MRD) of phonons. The last green bars of both bar charts represent the probabilities of getting the excited state through the JC coupling after eight subtractions, which can be regarded as the charging performance. To compare the performance of both schemes, we mark the height of the upper green bar on the lower one with the white dashed line. In the inset, the probability distributions of the two schemes of subtraction are compared. The solid black line represents the distribution obtained from eight linear subtractions, while the orange histogram shows the distribution obtained when the two nonlinear subtractions are introduced. The latter is then compared with a Poissonian distribution of the same mean excitation $\langle\hat{n}\rangle$, displayed by the blue histogram.
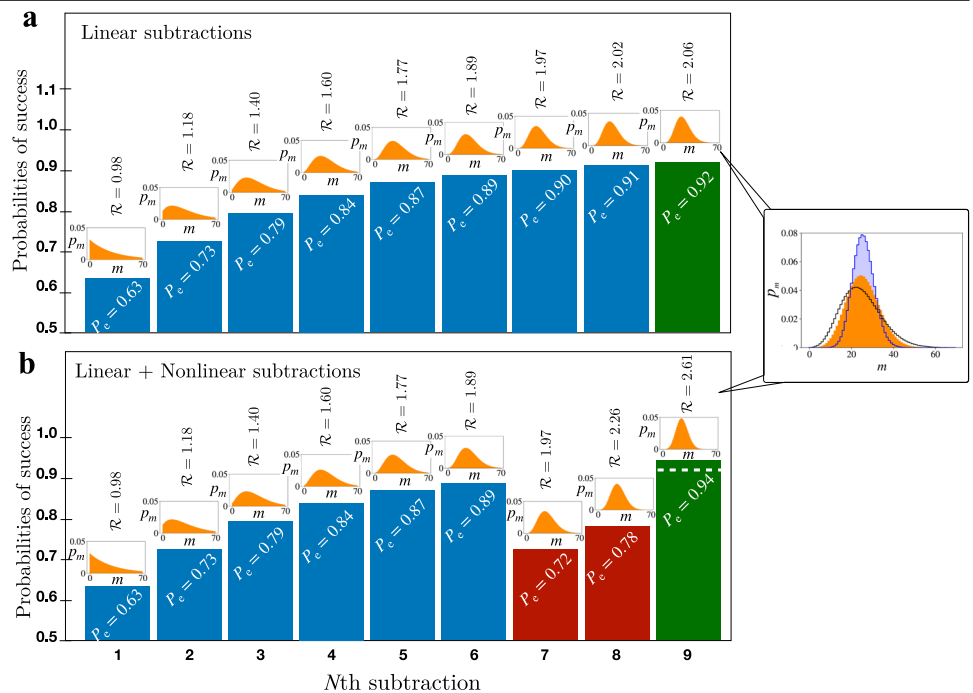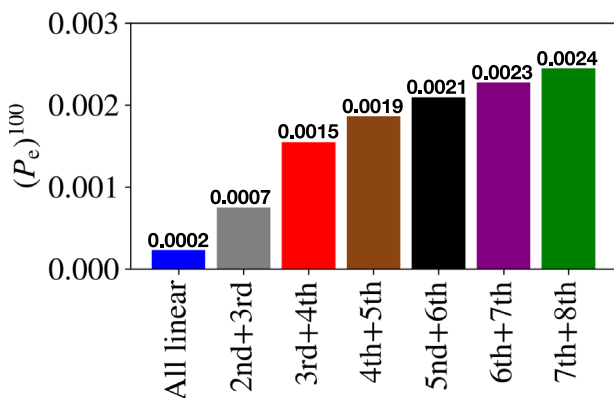
performance can be. The probabilities of successfully charging an individual battery associated with these strategies of subtraction may, at first, look insignificantly different, but at the scale of mass production the differences become eventually magnified. This fact is demonstrated in Fig. 8 by comparing the probabilities of successfully charging a hundred out of a hundred quantum batteries using different combinations of linear and nonlinear subtractions. To relate it with the result in subsection "Nonlinear subtraction" in Results, we consider only the cases in which eight subtractions, including six linear and two consecutive nonlinear subtractions, are performed before the charging stage and compare their charging performances to that of the linear-subtractions-only strategy. From the figure, the best performance is obtained if the last two subtractions are nonlinear. Its increase is even several times larger than the order of the excitation probability $(P_e)^{100}$ obtained through the only-linear-subtraction case. The underlying reason originates from the fact that nonlinear subtractions are better at squeezing the phonon population but improper for trimming its tails. Earlier use of nonlinear subtractions, as a result, reduces the mean phonon number of the end-product motional state and leads to a lower mean-to-deviation ratio $\mathcal{R}$. Several linear subtractions thus prepare a properly trimmed phonon distribution to be squeezed by the following nonlinear subtractions.

We add here that as the third law of quantum thermodynamics forbids us to prepare a perfectly pure state, an ideal projective measurement is thus impossible to achieve with a finite amount of time and resource[44]. This emphasizes the benefit of the nonlinear interaction as it helps us reduce the required number of subtractions and measurements to achieve a sufficiently high charging performance. For example, for $\bar{n} = 30$, we need 14 linear subtractions with protocol II to achieve a charging performance of $P_e = 0.94$, while this performance can be accomplished by 8 subtractions: six linear and two nonlinear subtractions, as depicted in Fig. 7. Moreover, in comparison with the previously proposed Maxwell's demons[46,47], where the measurement is made to ensure the energy transfer from cold to hot reservoirs, our protocols use measurement to conditionally change the state so that the population distributions are transformed into a less noisy out-of-equilibrium distribution sufficient to excite qubits.

## Conclusions

The idea of classical Maxwell's demon initiates the reformation of classical thermodynamics, generates a connection between information and



**Fig. 8 | The performance of charging a hundred qubits.** The figure compares the probability $(P_e)^{100}$ of successful charging a hundred qubits after eight sequential phonon subtractions on a thermal state with the initial phonon number of $\bar{n} = 30$, using different combinations of linear and nonlinear subtractions. The blue bar denotes the performance of charging after eight linear subtractions, while the others demonstrate the performance after six linear and two consecutive nonlinear subtractions in different arrangements of order. Under these bars, the labels $n$th $+(n+1)$th are used to identify the appearance order of the two nonlinear subtractions in the subtraction sequence. It is apparent that the performance of charging is improved better if we assign the nonlinear subtractions to perform last. The numbers above these bars are the actual values of such probabilities $(P_e)^{100}$.

thermodynamic work, and provides the fundamental idea of the conversion between these two quantities. With the information of a system obtained through measurements and precise control, the system can then be manipulated into an out-of-equilibrium state, and, in return, its energy can be extracted. The idea of such conversion is carried on to its quantum version with some fundamental differences. In contrast to the classical case, in which a measurement is treated to be arbitrarily sharp without any back action on the measured system, in the quantum domain, both measurement outcomes and their back actions unavoidably affect the way we conditionally control and manipulate the system in order to generate a useful resource. Each performed measurement not only extracts the system's information but also transforms its state accordingly.

We have proposed a simple but deterministic protocol to realize a bosonic Maxwell's demon at a quantum level, exploiting the free coupling between the mechanical modes of a single atom and its internal electronic state. A measurement of the qubit with its outcome implying absorption of phonons by the qubit is regarded as phonon subtraction. It is shown that linear subtractions from both protocols I and II transform the phonon state from an initial thermal state into an out-of-equilibrium state with a nearly Gaussian phonon distribution. This transformed motional state can eventually be used to charge a microscopic battery, another qubit, by exciting it through a linear JC coupling. The charging performance of such out-of-equilibrium states is higher than that of its initial thermal state, which can be indicated by its increased mean-to-deviation ratio, $\mathcal{R}$. The performance is enhanced each time a subtraction is performed but becomes saturated eventually. To break through this limitation, a nonlinear subtraction, using a nonlinear JC coupling to absorb more phonons at once, must be exploited using the procedure of subtraction protocol II. The nonlinear interaction can boost the charging performance further, at the cost of its speed in the repeat-until-success protocols. It can further squeeze the phonon distribution better than the linear version, which increases $\mathcal{R}$ as a result. Nonetheless, it still has a drawback as it cannot trim the tails of the phonon population properly, making it better used as a final performance booster. The involvement of nonlinear subtractions helps us reduce the number of required subtractions to achieve a sufficiently high charging performance. Although we use a trapped ion as an example quantum platform, this proposed protocol can also be realized easily in other platforms in which a nonlinear JC coupling is available, such as superconducting circuits and cavity quantum electrodynamics.

Using such states from nonlinear subtractions in parallel to fully and independently excite a hundred two-level systems can give more than ten times higher success rates than those from only linear ones. We, therefore, believe that nonlinear-based Maxwell's demon can potentially pave the way for a new area of theoretical and experimental research in quantum statistics and thermodynamics.

## Methods
### Optimal interaction times

In semi-classical treatment $\bar{n} \gg 1$, the probability of a qubit in its ground state being excited by interacting with a thermal oscillator having a large mean phonon number $\bar{n} \gg 1$ as can be approximated as[48]

$$P_e(t) \approx \lambda t \sqrt{\bar{n}} D(\lambda t \sqrt{\bar{n}}) \qquad (21)$$

where $\lambda$ is the coupling strength between the qubit and phonons, and $D(x)$ is Dawson's integral, defined by

$$D(x) = e^{-x^2} \int_0^x e^{x'^2} dx'. \qquad (22)$$

Before reaching the quantum-revival region, this semi-classical approximation agrees well with the calculation obtained from rigorous quantum treatment. The probability $P_e$ is, therefore, maximized when $\lambda t \sqrt{\bar{n}} \approx 1.502$.

On the other hand, a thermal oscillator with small mean phonon number, $\bar{n} \ll 1$ dominantly occupies in its ground and the first excited states, $|0\rangle$ and $|1\rangle$. The population of all other motional states becomes negligible compared to that of the two states. It is the phonon in the Fock state $|1\rangle$ that mainly excites a qubit. The probability of successful excitation of a qubit in this case then becomes

$$P_e(t) = \sum_{m=0}^{\infty} p_m \sin^2(\lambda t \sqrt{m}) \approx \bar{n} \sin^2(\lambda t \sqrt{1}) + \mathcal{O}(\bar{n}^2). \quad (23)$$

This imply that the maximum $P_e$ occurs when $\lambda t \approx \pi/2$. To compromise between these two extreme limits, $\bar{n} \gg 1$ and $\bar{n} \ll 1$, we then approximate the optimal interaction time as $\lambda t_0^{\text{op}} \approx \pi/(2\sqrt{\bar{n}+1})$, which approaches $\pi/2$ for small $\bar{n}$ and still well agrees with the semi-classical treatment for large $\bar{n}$. We note here that this approximation still hold true even in the intermediate limit of the average phonon number $\bar{n}$. For example, when $\bar{n} = 2$, the approximation of $\lambda t_0^{\text{op}} \approx \pi/(2\sqrt{\bar{n}+1})$ is differed from its actual value only by ~1%.

For the semi-classical case, $\bar{n} \gg 1$, after several linear subtractions, the population distribution of phonons is modified into nearly Gaussian with its peak centered around its average phonon number $\langle \hat{n} \rangle$ and can, therefore, be qualitatively approximated as a Gaussian distribution as

$$p_m \sim \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(m - \langle \hat{n} \rangle)^2}{2\sigma^2}\right), \quad (24)$$

where $\sigma^2$ represents the variance of the distribution. The optimal interaction times for linear subtractions for the approximated Gaussian distribution is then of the form $\lambda t_N^{\text{op}} \approx \pi/(2\sqrt{\langle \hat{n} \rangle + 1})$, which is similar to the previous result. On the other hand, for this nearly Gaussian distribution, the probability associated with a successful nonlinear subtraction becomes

$$P_e' \approx \frac{1 - e^{-2\lambda'^2 t^2 \sigma^2} \cos(2\lambda' t(\langle \hat{n} \rangle))}{2}. \quad (25)$$

With a sufficiently small variance $\sigma^2$, we, therefore, can qualitatively approximate the optimal interaction time for a nonlinear subtraction as $\lambda' \tau^{\text{op}} \approx \pi/2\langle \hat{n} \rangle$, as shown in Eq. (19).

## Data availability
The data supporting the findings of this study are available from the first author upon reasonable request.

## Code availability
The theoretical results of the manuscript are reproducible from the formulas and derivations given therein. Additional code is available from the first author upon reasonable request.

## References

1. Maruyama, K., Nori, F. & Vedral, V. Colloquium: the physics of Maxwell's demon and infomation. *Rev. Mod. Phys.* **81**, 1 (2009).
2. Bennett, C. H., Bernstein, H. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. *Phys. Rev. A* **53**, 2046–2052 (1996).
3. Kwiat, P. G., Barraza-Lopez, S., Stefanov, A. & Gisin, N. Experimental entanglement distillation and 'hidden' non-locality. *Nature* **409**, 1014 (2001).
4. Parrondo, J., Horowitz, J. & Sagawa, T. Thermodynamics of information. *Nat. Phys.* **11**, 131 (2015).
5. Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E. & Sano, M. Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. *Nat. Phys.* **6**, 988 (2010).
6. Debiossac, M., Grass, D., Alonso, J. J., Lutz, E. & Kiesel, N. Thermodynamics of continuous non-Markovian feedback control. *Nat. Commun.* **11**, 1360 (2020).
7. Chitambar, E. & Gour, G. Quantum resource theories. *Rev. Mod. Phys.* **91**, 025001 (2019).
8. Brandão, F. G. S. L., Horodecki, M., Oppenheim, J., Renes, J. M. & Spekkens, R. W. Resource theory of quantum states out of thermal equilibrium. *Phys. Rev. Lett.* **111**, 250404 (2013).
9. Datta, C., Kondra, T. V., Miller, M. & Streltsov, A. Catalysis of entanglement and other quantum resources. *Rep. Prog. Phys.* **86**, 116002 (2023).
10. Narasimharchar, V. et al. Thermaldynamics resources in continuous-variable quantum systems. *npj Quant. Inf.* **7**, 9 (2021).
11. Serafini, A. et al. Gaussian thermal operations and the limits of algorithmic cooling. *Phys. Rev. Lett.* **124**, 010602 (2020).
12. Iskhakov, T. S. et al. Heralded source of bright multi-mode mesoscopic sub-Poissonian light. *Opt. Lett.* **41**, 2149–2152 (2016).
13. Iskhakov, T. S., Usenko, V. C., Filip, R., Chekhova, M. V. & Leuchs, G. Low-noise macroscopic twin beams. *Phys. Rev. A* **93**, 043849 (2016).
14. Parigi, V., Zavatta, A., Kim, M. & Bellini, M. Probing quantum commutation rules by addition and subtraction of single photons to/from a light field. *Science* **317**, 1890 (2007).
15. Zavatta, A., Parigi, V., Kim, M. S. & Bellini, M. Subtracting photons from arbitrary light fields: experimental test of coherent state invariance by single-photon annihilation. *New J. Phys.* **10**, 123006 (2008).
16. Fedorov, I. A., Ulanov, A. E., Kurochkin, Y. V. & Lvovsky, A. I. Quantum vampire: collapse-free action at a distance by the photon annihilation operator. *Optica* **2**, 112 (2015).
17. Bogdanov, Y. I. et al. Multiphoton subtracted thermal states: description, preparation, and reconstruction. *Phys. Rev.* **96**, 063803 (2017).
18. Katamadze, K. G., Avosopiants, G. V., Bogdanov, Y. I. & Kulik, S. P. How quantum is the quantum vampire effect?: testing with thermal light. *Optica* **5**, 723 (2018).
19. Katamadze, K. G., Kovlakov, E. V., Avosopiants, G. V. & Kulik, S. P. Direct test of the quantum vampire's shadow absence with use of thermal light. *Opt. Lett.* **44**, 3286 (2019).
20. Katamadze, K. G., Avosopiants, G. V., Bogdanova, N. A., Bogdanov, Yu. I. & Kulik, S. P. Multimode thermal states with multiphoton subtraction: Study of the photon-number distribution in the selected subsystem. *Phys. Rev. A* **101**, 013811 (2020).
21. Enzian, G. et al. Single-phonon addition and subtraction to a mechanical thermal state. *Phys. Rev. Lett.* **126**, 033601 (2021).
22. Vidrighin, M. D. et al. Photonic Maxwell's demon. *Phys. Rev. Lett* **116**, 050401 (2016).
23. Shu, A., Dai, J. & Scarani, V. Power of an optical Maxwell's demon in the presence of photon-number correlations. *Phys. Rev. A* **95**, 022123 (2017).
24. Barnett, S. M., Ferenczi, G., Gilson, C. R. & Speirits, F. C. Statistics of photon-subtracted and photon-added states. *Phys. Rev. A* **98**, 013809 (2018).
25. Hloušek, J., Ježek, M. & Filip, R. Work and information from thermal states after subtraction of energy quanta. *Sci. Rep.* **7**, 13046 (2017).
26. Zanin, G. L. et al. Enhanced photonic Maxwell's demon with correlated baths. *Quantum* **6**, 810 (2022).
27. Spagnolo, M. et al. Experimental photonic quantum memristor. *Nat. Photon.* **16**, 318 (2022).
28. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281 (2003).
29. Chu, Y. et al. Creation and control of multi-phonon Fock states in a bulk acoustic-wave resonator. *Nature* **563**, 666 (2018).

30. Shore, B. W. & Knight, P. L. The Jaynes-Cummings model. *J. Mod. Opt.* **40**, 1195 (1993).
31. Fink, J. M. et al. Climbing the Jaynes-Cummings ladder and observing its nonlinearity in a cavity QED system. *Nature* **454**, 315–318 (2008).
32. Um, M. et al. Phonon arithmetic in a trapped ion system. *Nat. Commun.* **7**, 11410 (2016).
33. Sukumar, C. V. & Buck, B. Multi-phonon generalisation of the Jaynes-Cummings model. *Phys. Lett.* **83A**, 211 (1981).
34. Singh, S. Field statistics in some generalized JaynesCummings models. *Phys. Rev. A* **25**, 3206 (1982).
35. Villas-Boas, C. J. & Rossatto, D. Z. Multiphoton Jaynes-Cummings model: arbitrary rotations in fock space and quantum filters. *Phys. Rev. Lett.* **122**, 123604 (2019).
36. McCormick, K. C. et al. Quantum-enhanced sensing of a single-ion mechanical oscillator. *Nature* **572**, 86 (2019).
37. Scully, M.O. & Zubairy, M.S. *Quantum Optics* (Cambridge University Press, 1997).
38. Brune, M., Raimond, J. M., Goy, P., Davidovich, L. & Haroche, S. Realization of a two-photon maser oscillator. *Phys. Rev. Lett.* **59**, 1899 (1987).
39. Garziano, L. et al. Multiphoton quantum Rabi oscillations in ultrastrong cavity QED. *Phys. Rev. A* **92**, 063830 (2015).
40. Ding, S., Maslennikov, G., Hablützel, R., Loh, H. & Matsukevich, D. Quantum parametric oscillator with trapped ions. *Phys. Rev. Lett.* **119**, 150404 (2017).
41. Ding, S., Maslennikov, G., Hablützel, R. & Matsukevich, D. Cross-Kerr nonlinearity for phonon counting. *Phys. Rev. Lett.* **119**, 193602 (2017).
42. Maslennikov, G. et al. Quantum absorption refrigerator with trapped ions. *Nat. Commun.* **10**, 202 (2019).
43. Ding, S., Maslennikov, G., Hablützel, R. & Matsukevich, D. Quantum simulation with a trilinear Hamiltonian. *Phys. Rev. Lett.* **121**, 130502 (2018).
44. Guryanova, Y., Friis, N. & Huber, M. Ideal projective measurements have infinite resource costs. *Quantum* **5**, 222 (2020).
45. Marek, P., Provazník, J. & Filip, R. Loop-based subtraction of a single photon from a traveling beam of light. *Opt. Express* **26**, 29837 (2018).
46. Kieu, T. D. Quantum heat engines, the second law and Maxwell's demon. *Eur. Phys. J. D* **38**, 115 (2006).
47. Minganti, F., Macrí, V., Settineri, A., Savasta, S. & Nori, F. Dissipative state transfer and Maxwell's demon in single quantum trajectories: excitation transfer between two noninteracting qubits via unbalanced dissipation rates. *Phys. Rev. A* **103**, 052201 (2021).
48. Knight, P. L. & Radmore, P. M. Quantum revivals of a two-level system driven by chaotic radiation. *Phys. Lett* **90A**, 342 (1982).

## Acknowledgements

## Author contributions

R.F. conceived and supervised the project. A.R. and R.F. jointly build theory and analyze the results. A.R. performed numerical simulations and wrote the paper with feedback from R.F.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s42005-024-01546-4.

**Correspondence** and requests for materials should be addressed to Atirach Ritboon.

**Peer review information** *Communications Physics* thanks Andrew Hogan and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at http://www.nature.com/reprints

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