Photon cooling: Linear versus nonlinear interactions

A. Hovhannisyan,[1] V. Stepanyan,[2] and A. E. Allahverdyan[3,*]
[1]*Institute of Applied Problems of Physics, Yerevan 0014, Armenia*
[2]*Physics Department, Yerevan State University, Yerevan 0025, Armenia*
[3]*Alikhanian National Laboratory, Yerevan Physics Institute, Yerevan 0036, Armenia*

(Received 27 March 2022; accepted 7 September 2022; published 16 September 2022)

Linear optics imposes a relation that is more general than the second law of thermodynamics: For modes undergoing a linear evolution, the full mean occupation number, i.e., the photon number for optical modes, does not decrease, provided the evolution starts from a (generalized) diagonal state. This relation connects to noise increasing (or heating) and is akin to the second law and holds for a wide set of initial states. Also, the Bose entropy of modes increases, though this relation imposes additional limitations on the initial states and on linear evolution. We show that heating can be reversed via nonlinear interactions between the modes. They can cool, i.e., decrease the full mean occupation number and the related noise, an equilibrium system of modes provided their frequencies are different. Such an effect cannot exist in energy cooling, where only a part of an equilibrium system is cooled. We describe the cooling setup via both efficiency and coefficient of performance and relate the cooling effect to the Manley-Rowe theorem in nonlinear optics.

DOI: 10.1103/PhysRevA.106.032214

I. INTRODUCTION

Cooling is needed for noise reduction and for capturing quantum degrees of freedom. It has been studied during the past 100 years in various setups [1–3]. Cooling processes are also fundamental for thermodynamics: They sharpen the understanding of the second law and are instrumental for the third law [4]. An interesting example of this is the laser cooling of solids via the anti-Stokes effect, which has both quantum and thermodynamic nature [1]. Much attention is currently devoted to cooling processes in quantum thermodynamics [5–20]. It is known that only a part of a thermally isolated (initially equilibrium) system can be cooled in terms of energy (or temperature), that cooling such systems costs high-grade energy (work) [hence the definition of the coefficient of performance (COP)], and that cooling is limited by energy spectra and complexity costs.

Here we consider bosonic (for clarity photonic) degrees of freedom (modes) and show that linear transformations (e.g., linear optics) always increase the full photon number of the system. This statement holds for a wide class of initial states. For such states, increasing the mean photon number relates to increasing the noise (heating). The heating is more general than the second law. To confirm this point, we studied the full Bose entropy of modes. This coarse-grained entropy is conditionally maximal at equilibrium and can change under a unitary evolution, in contrast to the fine-grained von Neumann entropy. We show that the Bose entropy can increase, but this relation (a formulation of the second law) demands additional limitations on both the initial states and linear evolution.

Heating can be reversed by nonlinear interactions. One can cool in this sense an initially *equilibrium* system, which consists of two or more modes. This is not possible for energy cooling, where, as demanded by the second law, only the subsystem's energy can be decreased (cooled). Our cooling setup is characterized by two efficiencylike parameters: the COP and the efficiency. The former refers to the energy costs of cooling, while the latter normalizes the cooling result over the total changes introduced in the system. Nonlinear interactions achieve cooling in near-resonance regimes, where there is an effective conservation law in the number of photons (Manley-Rowe theorem) [21,22]. Thus, this cooling scenario uncovers the thermodynamic role of nonlinear optical processes. We work in terms of photons, but our results hold for other bosons (e.g., phonons).

This paper is organized as follows. Section II shows that the mean boson (photon) number increases in linear evolution if the evolution starts from a certain class of generalized diagonal initial states. This class is sufficiently large and includes the usual diagonal states (in the Fock basis), independent states (over the modes), etc. This section also relates the increase of the mean number to noise and formulates this as a heating (no-cooling) principle for linear evolution. In Sec. III we study the Bose entropy for modes and explain under what additional restrictions (compared with the mean photon number increase) this entropy grows. This section also addresses the physical meaning of the Bose entropy. Section IV describes the optimal cooling setup for two modes and introduces the basic characteristics of cooling, viz., efficiency and the COP. This section emphasizes the key feature of this cooling setup, namely, the global cooling of an equilibrium system in terms of the mean photon number (and noise) is possible. Section V demonstrates that cooling is possible

*Corresponding author: armen.allahverdyan@gmail.com

 ©2022 American Physical Society

also via a feasible nonlinear two-mode interaction, works out a simple example of such interactions, and establishes the relation to the Manley-Rowe theorem, a known result in nonlinear physics. Section VI summarizes our results.

II. NO COOLING FOR LINEAR INTERACTIONS

A. A single bosonic mode

Linear processes, which are described with Hamiltonians quadratic in creation and annihilation operators, describe the lion's share of boson dynamics [23,24]. Consider the simplest example of such processes: a single mode that undergoes a linear evolution governed by a quadratic Hamiltonian. In the Heisenberg picture, the general form of this evolution connects the initial $a = a(0)$ and final $b = a(t)$ annihilation operators of the mode

$$b = Sa + Ra^\dagger + f, \tag{1}$$

where S, R, and f are complex c-numbers that characterize the evolution. The initial state (density matrix) ρ of the mode satisfies

$$\langle a\rangle \equiv \mathrm{tr}(a\rho) = \langle a^\dagger\rangle = 0. \tag{2}$$

The commutation relation $[b, b^\dagger] = [a, a^\dagger] = 1$ imposes $|S|^2 - |R|^2 = 1$ in (1). Then we get, from (1) and (2),

$$\langle b^\dagger b\rangle - \langle a^\dagger a\rangle = 2|R|^2\langle a^\dagger a\rangle + |R|^2 + |f(t)|^2 \geqslant 0, \tag{3}$$

i.e., the mean photon number difference defined on the left-hand side of (3) can only increase. In particular, this conclusion holds for linear amplifiers [23]. According to (1), also the dispersion of the photon number increases:

$$\langle (b^\dagger b)^2\rangle - \langle b^\dagger b\rangle^2 \geqslant \langle (a^\dagger a)^2\rangle - \langle a^\dagger a\rangle^2. \tag{4}$$

We emphasize that the analog of (1) and (3) for a fermion mode does not generally hold. One heuristic reason for this is that only for the bosonic mode can the mean (photon) number be arbitrary large.

B. Many modes, relations with noise, and heating

1. Linear Heisenberg evolution

Importantly, Eq. (3) extends to the completely general N-mode situation, where instead of (1) we write for initial $a_i = a_i(0)$ and final $b_i = a_i(t)$ Heisenberg operators

$$b_i = \sum_{j=1}^{N}(S_{ij}a_j + R_{ij}a_j^\dagger) + f_i, \quad i = 1, \ldots, N, \tag{5}$$

where S_{ij}, R_{ij}, and f_i are c-numbers [cf. (1)]. We write (5) in block-matrix form

$$\begin{pmatrix} b \\ b^\dagger \end{pmatrix} = E\begin{pmatrix} a \\ a^\dagger \end{pmatrix} + \begin{pmatrix} f \\ f^* \end{pmatrix}, \quad E = \begin{pmatrix} S & R \\ R^* & S^* \end{pmatrix}, \tag{6}$$

where $a = (a_1, \ldots, a_N)^T$; $a^\dagger$, b, $b^\dagger$, f, and f^* are N-columns; and T and $*$ denote transposition and complex conjugation, respectively. Below $^\dagger = {}^{*T}$ will denote Hermitian conjugation.

Now commutation relations $[b_i, b_j^\dagger] = [a_i, a_j^\dagger] = \delta_{ij}$, where δ_{ij} is the Kronecker delta, and $[b_i, b_k] = [a_i, a_k] = 0$ lead, from (6), to

$$SS^\dagger - RR^\dagger = I, \quad SR^T = RS^T, \tag{7}$$

respectively, where I is the $N \times N$ unit matrix. Equations (7) imply

$$E^{-1} = \begin{pmatrix} S^\dagger & -R^T \\ -R^\dagger & S^T \end{pmatrix}. \tag{8}$$

The reasoning that led to (7) is now applied to (8), since the same commutation relations hold. Then we get, in addition to (7), the new relations

$$S^\dagger S - R^T R^* = I, \quad S^\dagger R = R^T S^*. \tag{9}$$

2. Initial state

Now assume that the initial state ρ of N modes fulfills the two conditions

$$\langle a_j\rangle \equiv \mathrm{tr}(\rho a_j) = 0, \tag{10}$$

$$\langle a_i a_j\rangle \equiv \mathrm{tr}(\rho a_j a_i) = 0, \tag{11}$$

where $i, j = 1, \ldots, N$. Two interesting examples of (10) and (11) are as follows. First, Eq. (10) can refer to initially independent modes in states with $\langle a_i\rangle = 0$. Then (11) holds automatically due to the independence

$$\langle a_i a_j\rangle = \langle a_i\rangle\langle a_j\rangle = 0. \tag{12}$$

Second, we can consider diagonal states ρ_{diag} that read, in the Fock basis,

$$\rho_{\mathrm{diag}} = \sum_{\nu_1,\ldots,\nu_N=0}^{\infty} r_{\nu_1,\ldots,\nu_N}|\nu_1, \ldots, \nu_N\rangle\langle\nu_1, \ldots, \nu_N|, \tag{13}$$

$$a_i^\dagger a_i|\mu_1, \ldots, \mu_N\rangle = \mu_i|\mu_1, \ldots, \mu_N\rangle. \tag{14}$$

Equation (14) defines the Fock basis and (13) ensures the conditions (10) and (11). It should be clear that neither independence nor diagonality is necessary for the validity of (10) and (11); for example, a nondiagonal state maintaining (10) and (11) can be easily constructed starting from (13). To be concise, we will refer to the states ρ satisfying (10) and (11) as generalized diagonal states.

3. Increase of the mean photon number

Using (10) and (11) together with the first of Eqs. (9), we find that the change of the total occupation number is nonnegative,

$$\sum_{i=1}^{N}(\langle b_i^\dagger b_i\rangle - \langle a_i^\dagger a_i\rangle) = \sum_{i=1}^{N}|f_i|^2 + \sum_{i,j=1}^{N}|R_{ij}|^2 + 2\sum_{i=1}^{N}\mathrm{tr}(Y_i\rho Y_i^\dagger) \geqslant 0, \tag{15}$$

where we define

$$Y_i \equiv \sum_{k=1}^{N} R_{ik}^* a_k. \tag{16}$$

When deducing (15), the condition (10) is needed to nullify terms proportional to $f_i\langle a_k\rangle$ in (15), while (11) is needed to nullify terms proportional to $(R^\dagger S)_{kl}\langle a_k a_l\rangle$.

Equation (5) can describe absorption (attenuation) of photons from a few selection target modes, at the expense of their

overall increase. For the particular case of Gaussian initial states, Eq. (15) follows from the result of Ref. [25] on the maximal work. Thus, according to (15), the full mean photon number can only increase under linear evolution.

Where do these additional photons come from? Answering this question is contingent on realization of the linear transformation. For example, the genesis of additional photons is relatively clear when the increase of the mean number of photons is accompanied by an increase in the overall mean energy [cf. (3)]. This energy increase comes from external sources that realize the linear dynamics. In particular, this is the case when the N modes start their evolution from the overall vacuum state, because then the mean energy can only increase. More generally, the relation between the mean energy increase and the mean photon number increase in a linear dynamics is absent: The latter is more general than the former [see (31) for clarification]. In such cases the genesis of additional photons should be prescribed to the general fact that the mean photon number is not conserved within linear dynamics.

4. Noise increase and heating

We emphasize that (15) can be interpreted as uncertainty increase. To this end, let us note, for a mode with annihilation operator a, that $\langle a^\dagger a\rangle$ characterizes the dispersion $\langle\Delta a^2\rangle$ of a [23],

$$\langle\Delta a^2\rangle \equiv \tfrac{1}{2}\langle aa^\dagger + a^\dagger a\rangle - |\langle a\rangle|^2 = \langle a^\dagger a\rangle + \tfrac{1}{2} - |\langle a\rangle|^2 \quad (17)$$

$$= \langle x^2\rangle - \langle x\rangle^2 + \langle y^2\rangle - \langle y\rangle^2, \quad a = x + iy, \quad (18)$$

where $x = (a + a^\dagger)/2$ and y are Hermitian operators. Equation (17) is the definition of dispersion for non-Hermitian a, while (18) shows how it can be measured via its Hermitian components x and y. Note from (17) that for $\langle a\rangle = 0$, the dispersion $\langle\Delta a^2\rangle$ reduces to the mean photon number $\langle a^\dagger a\rangle$.[1]

For initial states (10) considered, we have $\langle a_i\rangle = \langle b_i\rangle = 0$ and then (15) and (18) imply that also the sum of uncertainties (17) increases together with the photon number

$$\sum_{i=1}^{N}\left(\langle\Delta b_i^2\rangle - \langle\Delta a_i^2\rangle\right) = \sum_{i=1}^{N}(\langle b_i^\dagger b_i\rangle - \langle a_i^\dagger a_i\rangle) \geqslant 0, \quad (19)$$

i.e., as the mean photon number rises, so does the total dispersion. Equation (19) holds due to initial conditions (10) and (11) and will be interpreted as heating. Likewise, the decrease of both quantities in (19), which is possible due to nonlinear interactions, will mean cooling (discussed below).

We close this section by stressing that the relation between $\sum_{i=1}^{N}(\langle\Delta b_i^2\rangle - \langle\Delta a_i^2\rangle)$ and $\sum_{i=1}^{N}(\langle b_i^\dagger b_i\rangle - \langle a_i^\dagger a_i\rangle)$ is not automatic. For example, the linear dynamics for a particular condition $f_i = 0$ in (5) will hold (15) for the condition (11) only, i.e., Eq. (10) is now not needed. If $\langle a_i\rangle \neq 0$, then generically also $\langle b_i\rangle \neq 0$ and $\sum_{i=1}^{N}(\langle\Delta b_i^2\rangle - \langle\Delta a_i^2\rangle) \geqslant 0$ does not hold, though $\sum_{i=1}^{N}(\langle b_i^\dagger b_i\rangle - \langle a_i^\dagger a_i\rangle) \geqslant 0$ still holds due to $f_i = 0$.

[1]This quantity also controls the shot noise in photodetection [24].

III. ENTROPIC FORMULATION OF THE SECOND LAW FOR BOSONS

A. When Bose entropy increases for a linear dynamics

1. Definition of Bose entropy

Equation (15) shows that for the initial conditions (10) and (11) the total mean number of photons can only increase. In the context of this unidirectional change it is natural to ask whether one can find a suitable entropy function that also increases under linear dynamics. As we show below, the answer to this question is positive provided the initial states and the type of the linear dynamics are restricted.

First of all, we need to define the entropy function: As always with the unitary dynamics the von Neumann entropy $-\mathrm{tr}(\rho\ln\rho)$ (with ρ the density matrix) is not suitable for defining the second law, since it is conserved. We need a more coarse-grained (i.e., less microscopic) definition of entropy. A good choice is the time-dependent Bose entropy

$$S(t) = \sum_{k=1}^{N} s(n_k(t)), \quad n_k(t) \equiv \langle a_k^\dagger(t)a_k(t)\rangle, \quad (20)$$

$$s(n_k) \equiv (1 + n_k)\ln(1 + n_k) - n_k\ln(n_k). \quad (21)$$

Equation (20) is deduced for an ideal Bose gas from the microcanonic distribution [22]. If $s(n_k)$ from (21) is maximized for a fixed mean energy $\hbar\omega_k n_k$ of the mode k with frequency ω_k, one obtains the thermal expression for the mean occupation (photon) number. Indeed, making the Lagrange function $s(n_k) - \beta\hbar\omega_k n_k$, where β is the Lagrange multiplier (inverse temperature), one obtains $n_k = (e^{\beta\hbar\omega_k} - 1)^{-1}$. Equation (20) also increases in time within kinetic equations for weakly interacting bosons (see [26] for a recent discussion).

2. Increase of Bose entropy

To study the behavior of S in time for our situation, we need to add an additional initial condition to (10) and (11),

$$\langle a_i^\dagger a_j\rangle = \delta_{ij}\langle a_i^\dagger a_i\rangle, \quad (22)$$

where (22) holds for the examples (12) and (13). Without (22), i.e., staying with (10) and (11) only, we cannot express $n_k(t)$ via $n_k(0)$. Together with (22) this task is possible from (5),

$$n_i(t) = \langle b_i^\dagger b_i\rangle = \sum_{k=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2)n_k(0) + \sum_{k=1}^{N}|R_{ik}|^2 + \sum_{i=1}^{N}|f_i|^2, \quad (23)$$

where (7) and (9) imply

$$\sum_{k=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2) = 1 + 2\sum_{k=1}^{N}|R_{ik}|^2 \geqslant 1, \quad (24)$$

$$\sum_{i=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2) = 1 + 2\sum_{i=1}^{N}|R_{ik}|^2 \geqslant 1. \quad (25)$$

Let us assume that [consistently with (24) and (25)] there exists a double stochastic matrix Θ_{ik}, i.e., a matrix holding

$$\Theta_{ik} \geqslant 0, \quad \sum_{i=1}^{N} \Theta_{ik} = 1, \quad \sum_{k=1}^{N} \Theta_{ik} = 1, \tag{26}$$

such that[2]

$$|S_{ik}|^2 + |R_{ik}|^2 \geqslant \Theta_{ik}. \tag{27}$$

Matrices $|S_{ik}|^2 + |R_{ik}|^2$ that satisfy (27) are called double superstochastic [27,28]. Once (26) and (27) are assumed, the derivation of the second law in the Bose-entropic formulation becomes straightforward from noting that $s(n_k)$ from (21) is a positive, increasing, and concave function:

$$\begin{aligned} S(t) &= \sum_{i=1}^{N} s(n_i(t)) \geqslant \sum_{i=1}^{N} s\left(\sum_{k=1}^{N} \Theta_{ik} n_k(0)\right) \\ &\geqslant \sum_{i,k=1}^{N} \Theta_{ik} s(n_k(0)) = \sum_{k=1}^{N} s(n_k(0)) = S(0). \end{aligned} \tag{28}$$

Thus the initial conditions (10), (11), and (22) and dynamic restriction (27) are sufficient for the second law (28).

3. Validity of inequality (27)

Note that (27) trivially holds for $|R_{ik}|^2 = 0$. We emphasize that (27) implies (24) and (25), but the converse does not hold. To avoid confusion note that $\sum_{k=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2) \leqslant 1$ and $\sum_{i=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2) \leqslant 1$ imply $|S_{ik}|^2 + |R_{ik}|^2 \leqslant \Theta_{ik}$ for some double stochastic matrix Θ_{ik} [27,28].

The inequality (27) holds for $N = 2$ [see Appendix A, which also discusses the simplest counterexample of (27) for $N = 3$]. A constructive necessary and sufficient condition for the validity of (27) was found in [29],

$$\sum_{i\in\mathcal{I},\, k\in\mathcal{J}} (|S_{ik}|^2 + |R_{ik}|^2) \geqslant |\mathcal{I}| + |\mathcal{J}| - N, \tag{29}$$

where (29) should hold for *all* subsets $\mathcal{I}$ and $\mathcal{J}$ of $\{1, \ldots, N\}$ and where $|\mathcal{I}|$ and $|\mathcal{J}|$ are the numbers of elements in $\mathcal{I}$ and $\mathcal{J}$, respectively. The conditions (29) are straightforward to check at least for not very large N. The physical meaning of (29) is that sufficiently small values of $|S_{ik}|^2 + |R_{ik}|^2$ are to be excluded.

More general (but less constructive) sufficient dynamical conditions for (28) can be stated as well. For example, whenever (27) does not hold, but still

$$\begin{aligned} &s\left(\sum_{k=1}^{N}(|S_{ik}|^2 + |R_{ik}|^2) n_k(0) + \sum_{k=1}^{N} |R_{ik}|^2 + \sum_{i=1}^{N} |f_i|^2\right) \\ &\quad\geqslant \sum_{k=1}^{N}(|S_{ik}|^2 - |R_{ik}|^2) s(n_k(0)) \end{aligned} \tag{30}$$

holds for all i, we sum both parts of (30) over i, employ (25), and find $S(t) \geqslant S(0)$.

B. Similarities and differences with the standard formulation of the second law

We found two unidirectional relations inherent in linear dynamics for bosons: The inequality (15) shows an increase of the mean photon number, while (28) is about the increase of the Bose entropy. It is useful to compare these relations with the standard (Thomson) formulation of the second law [30,31]: A unitary dynamics does not decrease the mean energy of a quantum system that started its evolution from a Gibbsian equilibrium (or at least passive) state. The unitary dynamics is realized via a time-dependent cyclically changing Hamiltonian; the cyclic condition is needed to ensure that the initial and final Hamiltonians are equal [30,31].

The following are the similarities.

(i) Equations (15) and (28) and Thomson's formulation refer to unidirectional changes inherent in a unitary evolution. All of them hold for specific initial states.

(ii) For the single-mode situation Eq. (15) [i.e., (3)] refers to basically the same quantity as the Thomson formulation, since the mean photon number is proportional to the mean energy.

(iii) Equation (15) relates to a noise increase [cf. (17) and (18)]. The same holds for the entropic formulation (28) that refers to the Bose entropy (20). Thomson's formulation has a similar bridge, since it also tells about the broadening of the energy distribution in the final state compared to the initial state. This broadening is quantified by the entropy of the energy probability distribution [30,31].

The following are the differences.

(a) The second law holds for any unitary evolution, while (15) is restricted to a linear evolution of boson modes. The inequality (28) assumes even more restriction [see (27) and (22)].

(b) The direct relation between the energy and photon number is broken for the multimode situation, i.e., the analog of (15) for energy does not hold: The mean energy change

$$\sum_{i=1}^{N} \omega_i(\langle b_i^\dagger b_i\rangle - \langle a_i^\dagger a_i\rangle) \tag{31}$$

need not have a definite sign for the initial conditions (10) and (11). For (31), the derivation that led to (15) breaks down at the point when after the summation over index i, one needs to employ the first of Eqs. (9). The same holds for (30): It does not apply to the mean energy. In other words, Eq. (30) states that the Bose entropy must increase without simultaneously increasing the mean energy (or at least keeping it constant).

(c) In the applicability domain the second law demands equilibrium (e.g., Gibbsian) or at least a passive initial state [30,31], while (10), (11), and (2) allow initial states that need not be equilibrium or passive [cf. (13)]. Recall that a passive state has a density matrix ρ that is a nonincreasing function of the Hamiltonian H [30,31]. For a (Gibbsian) equilibrium state this function is specific: $\rho = e^{-\beta H}/\mathrm{tr}e^{-\beta H}$, where $\beta > 0$ is the inverse temperature [30,31]. Thus, Eq. (15) is more general than the second law in the context of initial states, but at the

[2]For the validity of (28) we in fact need instead of (26) a seemingly weaker condition, where $\sum_{i=1}^{N} \Theta_{ik} = 1$ in (26) is replaced by $\sum_{i=1}^{N} \Theta_{ik} \geqslant 1$. However, this condition together with $\sum_{k=1}^{N} \Theta_{ik} = 1$ and $\Theta_{ik} \geqslant 0$ leads to $\sum_{i=1}^{N} \Theta_{ik} = 1$.

same time it is less general in the context of dynamics, as it is restricted to linear evolution. The inequality (28) assumes more restriction on the initial state [see (22)].

IV. COOLING TWO EQUILIBRIUM MODES

A. Setup

Once (19) is understood to define heating for linear dynamics with the initial conditions (10) and (11), it is natural to ask whether nonlinear processes can cool, i.e., decrease the initial number of photons. To facilitate the thermodynamic meaning of this question, we will consider two initially Gibbsian equilibrium bosonic modes at the same temperature T. Now a single equilibrium mode cannot be cooled by any unitary (generally nonlinear) operation, since the mean occupation number is proportional to the energy, and the energy decrease for such a situation is prohibited by the second law. However, two initially equilibrium modes at different frequencies can be cooled, in terms of the mean full occupation number, via specific nonlinear interactions. Hence, we will first determine the optimal cooling and then turn to a nonoptimal but feasible scenario from the viewpoint of experimentally realizable nonlinear interactions.

Consider the initial state of two modes with frequencies ω_1 and ω_2 at temperature T,

$$\rho = \xi \exp\left(-\beta \sum_{i=1}^{2} \omega_i \hat{n}_i\right), \quad \xi = (1 - e^{-\beta\omega_1})(1 - e^{-\beta\omega_2}), \tag{32}$$

$$\hat{n}_i \equiv a_i^\dagger a_i, \quad i = 1, 2, \quad \hat{n} \equiv \sum_{i=1}^{2} a_i^\dagger a_i, \tag{33}$$

where $\hbar = 1$, $\beta = 1/k_B T$, and $\hat{n}_i$ is the occupation number operator for each mode. The two-mode system undergoes a unitary process that aims at cooling:

$$\rho(t) = U \rho U^\dagger, \quad U U^\dagger = 1. \tag{34}$$

B. COP and efficiency

Besides targeting the mean occupation number, we characterize the cooling via two efficiencylike quantities. Since ρ in (32) is an equilibrium state, the final average energy found from (34) is larger than the initial one, which is the second law:

$$\sum_{i=1}^{2} \omega_i \Delta n_i \geqslant 0, \quad \Delta n_i \equiv \mathrm{tr}[\rho(U^\dagger \hat{n}_i U - \hat{n}_i)]. \tag{35}$$

Equation (35) defines the energy cost of cooling and it motivates the usual definition of the COP [20], where the achieved cooling $-\sum_{i=1}^{2} \Delta n_i > 0$ is divided over the energy cost $\sum_{i=1}^{2} \omega_i \Delta n_i$.

Let us define the frequency ratio as

$$\alpha \equiv \frac{\omega_2}{\omega_1} < 1. \tag{36}$$

We use the dimensionless COP conventionally defined as

$$K = -\frac{\Delta n_1 + \Delta n_2}{\Delta n_1 + \alpha \Delta n_2}, \tag{37}$$

where a larger K means, e.g., a better cooling with a smaller energy cost. In (37) we take $\alpha < 1$ without loss of generality. Hence, the fact of cooling $-\sum_{i=1}^{2} \Delta n_i > 0$ implies, via (35) and $\alpha < 1$,

$$0 \leqslant \alpha(-\Delta n_2) \leqslant \Delta n_1 \leqslant -\Delta n_2. \tag{38}$$

Now (38) motivates us to define $\Delta n_1 - \Delta n_2 = |\Delta n_1| + |\Delta n_2|$ as the total number of occupation changes introduced in the system. This is consistent with thinking about the cooling as photon conversion: Some amount of low-energy photons ($\Delta n_2 < 0$) transforms into a smaller amount of higher-energy photons ($\Delta n_1 > 0$). The sum of low-energy photons given and high-energy photons received will be the total number of occupation changes. Only a fraction η of those lead to cooling:

$$\eta = -\frac{\Delta n_1 + \Delta n_2}{\Delta n_1 - \Delta n_2}. \tag{39}$$

We call η the efficiency of cooling. It is similar to other quantum efficiencies employed in optics [3,24]. Using (35) and (38), we get a bound where temperatures are replaced by frequencies

$$\eta \leqslant \frac{\Delta n_1 + \Delta n_2}{\Delta n_2} \leqslant 1 - \frac{\min[\omega_1, \omega_2]}{\max[\omega_1, \omega_2]}, \tag{40}$$

i.e., cooling is impossible for $\omega_1 = \omega_2$. Note that (40) is more similar to the Otto efficiency than to the Carnot efficiency of heat engines [32].

C. Optimal cooling

Given (34) and (33), we look for the unitary which minimizes the mean of $\hat{n}$ in the final state:

$$U_{\mathrm{opt}} = \mathrm{argmin}_U[\mathrm{tr}(U \rho U^\dagger \hat{n})]. \tag{41}$$

Noting the eigenresolutions [cf. (32) and (33)]

$$\rho = \sum_{k=0}^{\infty} r_k |r_k\rangle\langle r_k|, \quad \hat{n} = \sum_{l=0}^{\infty} \nu_l |\nu_l\rangle\langle\nu_l|, \tag{42}$$

we get, from (34), (42), and (41),

$$\mathrm{tr}(U \rho U^\dagger \hat{n}) = \sum_{k,l=0}^{\infty} r_k \nu_l z_{kl}, \quad z_{kl} = |\langle\nu_l|U|r_k\rangle|^2, \tag{43}$$

where

$$\sum_k z_{kl} = \sum_l z_{kl} = 1, \tag{44}$$

i.e., z_{km} is a doubly stochastic matrix [cf. (26)]. Such matrices form a compact convex set with vertices being permutation matrices [28]. As (43) is linear over z_{km}, it reaches the minimum value on the vertices, i.e., on permutation matrices z_{kl}. This implies, from (43), that U_{opt} can be chosen as a permutation matrix.

Thus U_{opt} is a permutation matrix and its form is seen from (43) and (42):

$$\mathrm{min}_U[\mathrm{tr}(U \rho U^\dagger \hat{n})] = \sum_{k=0}^{\infty} \nu_k^\uparrow r_k^\downarrow, \tag{45}$$

$$\nu_1^\uparrow \leqslant \nu_2^\uparrow \leqslant \nu_3^\uparrow \leqslant \cdots, \quad r_1^\downarrow \geqslant r_2^\downarrow \geqslant r_3^\downarrow \geqslant \cdots. \tag{46}$$

In (46) [cf. (42)] the ordered (antiordered) eigenvalues of $\hat{n}$ (ρ) refer to the final state in (34). We visualize the orderings of eigenvalues in the initial state of (32) and (33),

$\hat{n}$	0	1	2	3	$\cdots$
	(0, 0)	(0, 1), (1, 0),	(0, 2), (1, 1), (2, 0)	(0, 3), (1, 2), (2, 1), (3, 0)	$\cdots$
ρ	1	y^{α}, y	$y^{2\alpha}, y^{\alpha+1}, y^2$	$y^{3\alpha}, y^{2\alpha+1}, y^{\alpha+2}, y^3$	$\cdots$

(47)

where $y \equiv e^{-\beta\omega_1}$. The first, second, and third rows in (47) show the eigenvalues of $\hat{n}$, $(\hat{n}_1, \hat{n}_2)$, and ρ, respectively, with the prefactor ξ omitted [cf. (32)]. The unitary process of (34) and (45) permutes the eigenvalues of ρ. Using (47), we calculate the averages of $\hat{n}$ and $\hat{n}_i = a_i^\dagger a_i$:

$$\begin{aligned}\langle\hat{n}\rangle &= \xi(1y^{\alpha} + 1y + 2y^{2\alpha} + 2y^{\alpha+1} + 2y^2 + \cdots),\\ \langle\hat{n}_1\rangle &= \xi(0y^{\alpha} + 1y + 0y^{2\alpha} + 1y^{\alpha+1} + 2y^2 + \cdots),\\ \langle\hat{n}_2\rangle &= \xi(1y^{\alpha} + 0y + 2y^{2\alpha} + 1y^{\alpha+1} + 0y^2 + \cdots).\end{aligned} \tag{48}$$

The eigenvalues of ρ in (47) are organized in columns. Whenever the maximal element $y^{k\alpha}$ of the kth column is larger than the minimal element y^l of the lth column ($l < k$), we interchange them and achieve some cooling. Formally, we should iterate until all elements in the third row are arranged in descending order [cf. (46)]. Thus the optimal cooling increases the probability of eigenstates of $\hat{n}$ with lower photon number. Note from (47) and (48) that we can interchange elements within each column without changing Δn.

When Δn is fixed, the descending order of ρ's eigenvalues in the final state yields simultaneously the minimum value of Δn_1 and the maximum value of Δn_2. This is because the eigenvalues of $\hat{n}_1$ ($\hat{n}_2$) in (47) are arranged in ascending (descending) order. Equations (37) and (39) show that thereby also η and K reach their maximal values at the optimal Δn. The rule (48) stays intact and can be used after permutations.

The exact calculation of (45) is out of reach, since ρ has an infinite number of eigenvalues. However, we can develop a useful bound for it by focusing on permutations between nearest-neighbor columns. We define, from (36),

$$m \equiv \left\lceil \frac{\alpha}{1-\alpha} \right\rceil, \tag{49}$$

where $\lceil c \rceil$ is the smallest integer greater than or equal to c. Looking at (47) we see that for $k \geqslant m$, the maximal element of the $(k+1)$th column is larger than the minimal element of the kth column. Permuting them will contribute to Δn calculated via (45). Likewise, for $k \geqslant m+2$, the next to maximal element of the $(k+1)$th column is larger than the next to minimal element of the kth column. To visualize this situation, consider a part of (47) between columns $m+p$ and $m+p+1$ ($p \geqslant 0$):

$\hat{n}$	$m+p$	$m+p+1$
ρ	$\ldots, y^{\alpha+m+p-1}, y^{m+p}$	$y^{(m+p+1)\alpha}, y^{(m+p)\alpha+1}, \ldots$

(50)

where we omitted the second row of (47). Continuing this logic, we see that a new permutation appears for each even p and that we can cover all nearest-neighbor permutations. Hence a bound [cf. (32) and (45)]

$$\begin{aligned}0 < -\Delta n_{\rm opt} &\equiv \sum_{k=0}^{\infty}(n_k r_k - n_k^{\uparrow} r_k^{\downarrow})\\ &\geqslant \xi \sum_{l=0}^{\infty} y^{l(\alpha+1)} \sum_{k=m}^{\infty}(y^{(k+1)\alpha} - y^k)\\ &= \frac{(1-y)y^{\alpha(m+1)} - (1-y^{\alpha})y^m}{1-y^{\alpha+1}}.\end{aligned} \tag{51}$$

According to (51), cooling is possible for any $0 \leqslant \alpha < 1$, i.e., Eq. (51) is positive and grows with $y = e^{-\beta\omega_1}$ changing from 0 (at $y = 0$) to $\frac{1-\alpha}{1+\alpha}$ at $y = 1$. Appendix B studies the optimal cooling numerically. In particular, it shows numerical plots for the optimal $K_{\rm opt}$ (COP) and $\eta_{\rm opt}$ (efficiency).

Now assume that m given by (49) satisfies $m \gg 1$. Then the bound (51) gets small and becomes nearly exact, since the relative error between $\Delta n_{\rm opt}$ and (51) scales as $O(y^{2m})$. This estimate follows from the contribution of next-nearest-neighbor permutations and is confirmed in Appendix C. We report here the limiting values of K and η only, which are obtained as described above [cf. (37), (39), and (40)]:

$$K_{\rm opt} \to \infty, \quad \eta_{\rm opt} \to 0 \quad \text{for } \alpha \to 1, \tag{52}$$

$$K_{\rm opt} \to \infty, \quad \eta_{\rm opt} \to 1 \quad \text{for } \alpha \to 0. \tag{53}$$

In (53) $\alpha = \omega_2/\omega_1 \to 0$ is understood in the sense of a large ω_1 and a small ω_2. It is also important to note that both Δn_1 and Δn_2 are functions of α and in the limit of $\alpha \to 1$ both tend to zero. In the second limits of (52) and (53) η coincides with the Otto bound. In both limits the energy costs of cooling are negligible: $K_{\rm opt} \to \infty$. In the more general case of large ω_1 and fixed ω_2, we study $\eta_{\rm opt}$ and $K_{\rm opt}$ in Appendix D.

V. FEASIBLE INTERACTION HAMILTONIAN FOR COOLING

How is a permutation unitary $U_{\rm opt}$ realized? This relates to one of the major questions of quantum control (see, e.g., [33]). Any Hamiltonian that is a polynomial of a fixed degree over a_1, $a_1^\dagger$, a_2, and $a_2^\dagger$ can be realized via sufficiently many linear operations plus a single-mode nonlinearity [34]. However, realizing the permutation $U_{\rm opt}$ should be difficult in practice, since it refers to a Hamiltonian that is highly nonlinear over a_1, $a_1^\dagger$, a_2, and $a_2^\dagger$.

Now we focus on a feasible nonlinear interaction and determine its cooling ability. The feasibility comes at a cost: Now cooling will be possible mostly next to nonlinear resonances $\omega_2 \gtrsim 2\omega_1$ or $2\omega_2 \lesssim \omega_1$. This will also connect to the Manley-Rowe theorem, a known relation of nonlinear optics [21,22].

The simplest χ^2 nonlinear interactions can be realized in an anisotropic (e.g., crystalline) medium. Here the medium polarization $\vec{P}$ is quadratic in the electric field $\vec{E}$ [3,22,35,36]: $\vec{P} = \chi^{(1)}\vec{E} + \vec{E}\chi^{(2)}\vec{E}$, where $\chi^{(1)}$ and $\chi^{(2)}$ are susceptibilities. Neglecting the polarization degree of freedom for the electric field, its quantum operator representation is $\vec{E} \to a^\dagger + a$ [3,36]. Hence, the nonlinear interaction can be written as

$$H_I = (a_1^\dagger + a_1)(a_2^\dagger + a_2)^2 + (a_1^\dagger + a_1)^2(a_2^\dagger + a_2), \tag{54}$$

with the full Hamiltonian of the system being

$$H = \omega_1 a_1^\dagger a_1 + \omega_2 a_2^\dagger a_2 + gH_I = H_0 + gH_I, \tag{55}$$

where g is the interaction constant.

Yet another scenario for (55) is realized in the optomechanics. In addition to its applications in quantum technologies [37], this field emerged as a potential basis for quantum gravity and foundations of quantum mechanics [38,39]. In the optomechanical setting, the interaction between a laser and a mechanical oscillator is such that the resonance frequency $\omega_1(x)$ of the laser depends on the position x of the mechanical oscillator. Hence their joint Hamiltonian reads $H = \omega_1(x)a_1^\dagger a_1 + \omega_2 a_2^\dagger a_2$ [37]. Here a_1 and a_2 are the annihilation operators for the laser and the mechanical oscillator, respectively. Keeping up to the linear term of the Taylor expansion of $\omega_1(x)$ and using $x = a_2^\dagger + a_2$, we get

$$H = \omega_1 a_1^\dagger a_1 + \omega_2 a_2^\dagger a_2 + (\partial_x \omega_1) a_1^\dagger a_1 (a_2^\dagger + a_2), \tag{56}$$

which closely relates to (55).

To employ (54) and (55) in (35) we introduce the free Heisenberg interaction Hamiltonian $H_I(t) = e^{iH_0 t} H_I e^{-iH_0 t}$ and represent $\rho(t) = e^{-itH}\rho e^{itH}$ in (34) via chronological exponent $\overleftarrow{\exp}$:

$$\rho(t) = e^{-iH_0 t}\tilde{U}\rho\,\tilde{U}^\dagger e^{iH_0 t}, \quad \tilde{U} = \overleftarrow{\exp}\left(-i\int_0^t ds\, gH_I(s)\right). \tag{57}$$

Now we expand $\tilde{U}$ into a Dyson series

$$\tilde{U} = 1 - ig\int_0^t ds\, H_I(s) - g^2\int_0^t ds_1 \int_0^{s_1} ds_2 H_I(s_1)H_I(s_2) + \cdots. \tag{58}$$

Using $e^{iH_0 s}a_k e^{-iH_0 s} = e^{-i\omega_k s}a_k$ $(k = 1, 2)$ in $H_I(t)$, we can show that the order of magnitude estimate of the kth term in (58) reads

$$g^k \Omega^{-k} \sin^k(\Omega t/2), \tag{59}$$

$$\Omega = \min[\omega_1, \omega_2, |2\omega_1 - \omega_2|, |2\omega_2 - \omega_1|]. \tag{60}$$

Thus, for a suitable g, ω_1 and ω_2 we can keep in (58) the first three terms. Within this weak-coupling approximation we calculate (35) in Appendix E, showing that sufficiently large cooling $\Delta n < 0$ is possible only for

$$\omega_2 \gtrsim 2\omega_1 \quad \text{or} \quad 2\omega_2 \lesssim \omega_1, \tag{61}$$

i.e., for two possible near-resonance conditions. Restricting ourselves to the latter case $\alpha \equiv \omega_2/\omega_1 \lesssim 0.5$, we note that terms $a_1 a_2^{\dagger 2} + a_1^\dagger a_2^2$ in (54) oscillate much slower than other terms. Hence, within the rotating-wave approximation we can take, in (54),

$$H_I \simeq \overline{H}_I \equiv a_1 a_2^{\dagger 2} + a_1^\dagger a_2^2. \tag{62}$$

The approximation is studied in Appendix E, where we also work out (54). Now $\overline{H}_I$ in (62) leads to an exact operator conservation

$$2\hat{n}_1 + \hat{n}_2 = \text{const}, \quad \hat{n}_k = a_k^\dagger a_k, \quad k = 1, 2. \tag{63}$$

This conservation is the Manley-Rowe theorem for the considered nonlinear system [21,22]. The theorem does not generally hold for the complete interaction Hamiltonian (54). However, the cooling necessitates $\alpha \lesssim 0.5$ (or $\alpha \gtrsim 2$) and is accompanied by an approximate conservation law (63) (or $\hat{n}_1 + 2\hat{n}_2 = \text{const}$). Using (62) and (63), we get from (58), (57), and (35), keeping there the first three terms only (the order of g^2),

$$\Delta n_1 = \frac{8g^2 \sin^2\left(\frac{(2\omega_2 - \omega_1)t}{2}\right)}{(2\omega_2 - \omega_1)^2} \frac{e^{\beta\omega_1} - e^{2\beta\omega_2}}{(e^{\beta\omega_1} - 1)(e^{\beta\omega_2} - 1)^2}, \tag{64}$$

$$\Delta n_2 = -2\Delta n_1, \quad \Delta n = -\Delta n_1. \tag{65}$$

Hence the cooling at $\alpha \lesssim 0.5$ is described via $\eta = \frac{1}{3}$ and $K = \frac{1}{1-2\alpha}$ [cf. (37) and (39)]. Once η is finite and K is large, we achieve cooling with a small energy cost.

Equation (64) shows that a sizable cooling is achieved for sufficiently long times, because $\sin^2(\frac{(2\omega_2 - \omega_1)t}{2})$ is maximized for $|2\omega_2 - \omega_1|t \sim \pi$, while $|2\omega_2 - \omega_1|$ is small [cf. (61)]. This relation resembles the third law for the ordinary (energy) cooling, though more efforts are needed for its systematic investigation; e.g., we need a more complete understanding of the evolution generated by (55).

VI. SUMMARY

Our starting point was that linear transformations on boson modes (linear optics) increase the overall mean photon number provided the initial state is (generalized) diagonal [see (10), (11), and (15)]. This unidirectional relation refers to the linear evolution, but applies for a wider set of initial states (10) and (11) than the second law does. Its similarities and differences with respect to the second law were discussed in Sec. III B. We formulated this relation in full generality. The literature was close to such a formulation several times [23,25]. Given that the lion's share of boson dynamics is linear, this general result will hold for a number of fields including optics and phononics. Importantly, we showed explicitly that the relation (15) connects to increasing the overall noise in the system (though its subsystems can get a noise reduction as happens, e.g., in squeezing [24]). Hence we interpret it as heating.

It is interesting to ask how specifically the increase (15) of the overall mean photon number for initial states (10) and (11) relates to the second law. To answer this question, we studied the behavior of the Bose entropy (20) for linear dynamics and for the same class of initial states (10) and (11). The Bose entropy is conditionally maximized at equilibrium and it can change during unitary evolution in contrast to the (fine-grained) von Neumann entropy. We showed in Sec. III that for a subclass of linear evolution the Bose entropy (20) increases

and this increase also demands more restricted initial states (10), (11), and (22) than the validity of (15). A precise definition of this subclass relates to certain nontrivial problems in linear algebra. We thus confirm that for linear evolution the increase (15) of the overall mean photon number is a more general unidirectional relation than the second law.

We showed that the inverse of the heating in terms of the mean photon number (i.e., cooling) is possible within nonlinear (intermode) interactions. The cooling interpretation is not arbitrary and is characterized by efficiency and coefficient of performance. The former holds Otto's bound of the heat-engine efficiency (i.e., Carnot efficiency with temperatures replaced by frequencies). For the COP we anticipated, but so far have not identified, a general relation similar to Carnot's bound for the refrigeration COP [20].

We studied feasible nonlinear processes (e.g., χ^2 [3,22,35,36]) on two modes with different frequencies ω_1 and ω_2. Then the cooling in terms of the mean photon number happens (mostly) in the vicinity of nonlinear resonances. We also studied the optimal cooling, which is possible for any $\omega_1 \neq \omega_2$ but is demanding from the viewpoint of dynamic realization.

ACKNOWLEDGMENTS

We are grateful to Karen Hovhannisyan for important remarks and to David Petrosyan for discussions. This work was supported by State Committee of Science of Armenia, Grant No. 20TTAT-QTa003.

APPENDIX A: EXAMPLES AND COUNTEREXAMPLES FOR INEQUALITY (26)

We are given any 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{A1}$$

with non-negative elements and

$$a+b \geqslant 1, \quad c+d \geqslant 1, \quad a+c \geqslant 1, \quad b+d \geqslant 1. \tag{A2}$$

We define

$$c = \min[a, b, c, d] \tag{A3}$$

and note that only $c < 1$ is nontrivial, since otherwise (26) holds for (A1) and any double stochastic matrix. Now

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \geqslant \begin{pmatrix} 1-c & c \\ c & 1-c \end{pmatrix} = \Theta, \tag{A4}$$

where the latter matrix is double stochastic. Thus (26) holds for $N = 2$.

The simplest counterexamples for (26) at $N = 3$ is the following matrix with non-negative elements:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}. \tag{A5}$$

Here $a_{21} = a_{31} = 0$. We assume

$$\sum_{i=1}^{3} a_{ik} \geqslant 1, \quad \sum_{k=1}^{3} a_{ik} \geqslant 1, \tag{A6}$$

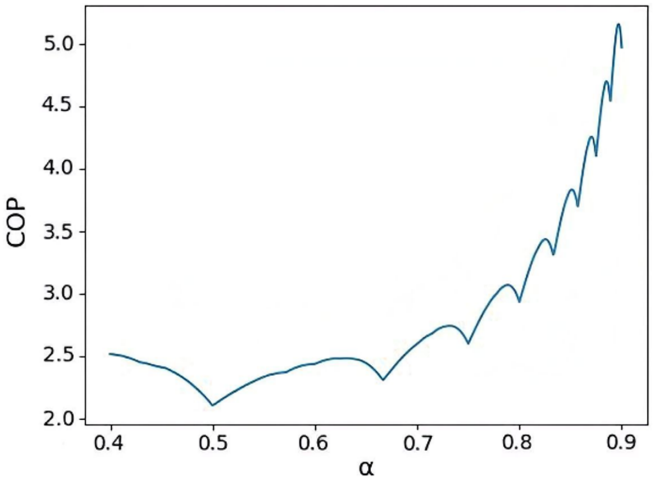

FIG. 1. Optimal COP K_{opt} versus $\alpha = \omega_2/\omega_1 < 1$ for the optimal cooling. Here $y^{\alpha} = e^{-\beta\omega_2} = 0.6$ and numerical calculations are done up to the block number 300 [see (47)].

and additionally

$$a_{22} + a_{32} < 1. \tag{A7}$$

If a double stochastic matrix holding (26) exists, then we have

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \geqslant \begin{pmatrix} 1 & 0 & 0 \\ 0 & \Theta_{22} & \Theta_{23} \\ 0 & \Theta_{32} & \Theta_{33} \end{pmatrix}. \tag{A8}$$

Now the latter matrix cannot be double stochastic and hold (26), as the condition (A7) is violated, if we take, in (29), $\mathcal{I} = 2, 3$ and $\mathcal{J} = 1, 2$.

APPENDIX B: NUMERICAL RESULTS FOR OPTIMAL COOLING

Recall our discussion after (47). There we explained that the optimal cooling, with respect to all involved quantities Δn_{opt} (photon number difference), K_{opt} (COP), and η_{opt} (efficiency), is achieved once all eigenvalues of the final density matrix are arranged in descending order [see the third row in (47)]. Numerically, this means that we need to take a sufficiently long but finite sequence of eigenvalues (starting from the largest one) and ensure that the results are stable with respect to the increasing length of this block.

Our numerical results are shown in Figs. 1–3. First, recall that in $K = -\frac{\Delta n_1 + \Delta n_2}{\Delta n_1 + \alpha \Delta n_2}$, the achieved photon number decrease $\Delta n = \Delta n_1 + \Delta n_2 < 0$ is divided over the dimensionless energy cost $\Delta n_1 + \alpha \Delta n_2$ [cf. (37)]. It is seen from Fig. 1 that K_{opt} as a function of α [cf. (36)] has (singular) local minima at points $\alpha = \frac{k}{k+1}$, where $k \in \mathbb{N}$ is an integer. We checked that these local minima of K come mostly from the singular behavior of the energy cost $\Delta n_1 + \alpha \Delta n_2$ (see Fig. 2). Now Δn (not shown in figures) shows weak singularities at those points $\alpha = \frac{k}{k+1}$, but these singularities are much weaker than those of the energy cost $\Delta n_1 + \alpha \Delta n_2$.

The origin of these singularities for K_{opt} (and $\Delta n_1 + \alpha \Delta n_2$) can be clarified as follows. Recall that [cf. (49)]

$$m \equiv \left\lceil \frac{\alpha}{1-\alpha} \right\rceil \tag{B1}$$

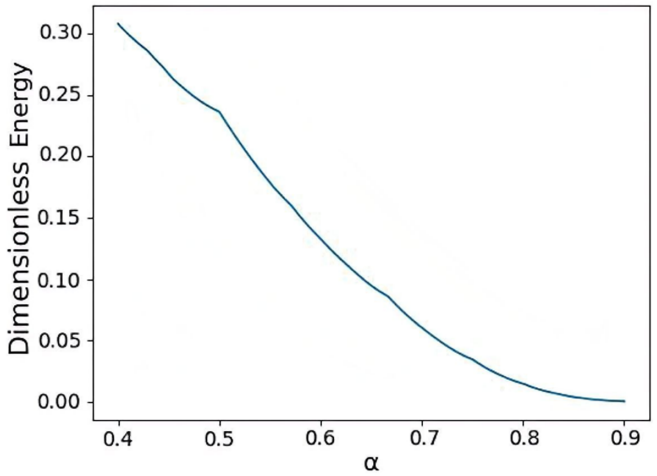

FIG. 2. Same as in Fig. 1, but for the optimal dimensionless energy cost $\Delta n_1 + \alpha \Delta n_2$ versus α defined via (36).

refers to the the group of eigenvalues of the initial state ρ starting from which the eigenvalues of ρ are not arranged in descending order [see the discussion after (47)]. At points $\alpha = \frac{k}{k+1}$ the index of the block from which the permutations start undergoes a jump discontinuity of increasing by one.

Figure 3 presents the numerical behavior of η_{opt} as a function of α. It is seen that η_{opt} also shows singularities at $\alpha = \frac{k}{k+1}$, though these singularities are weaker than those for K_{opt} (cf. Fig. 1). In particular, these singularities do not change the monotonic behavior of η_{opt} as a function of α.

APPENDIX C: ASYMPTOTIC RESULTS FOR OPTIMAL COOLING: THE LIMIT $\alpha \to 1$

Equation (51) provides the nearest-neighbor approximation for Δn. There we also indicated that (51) becomes close to its exact value whenever m defined via (B1) is sufficiently large or, equivalently, $\alpha \to 1$. The precision of this approximation relates to the necessity of next-nearest-neighbour permutations. The largest value of p in (50), where such permutations are necessary, can be estimated from the following

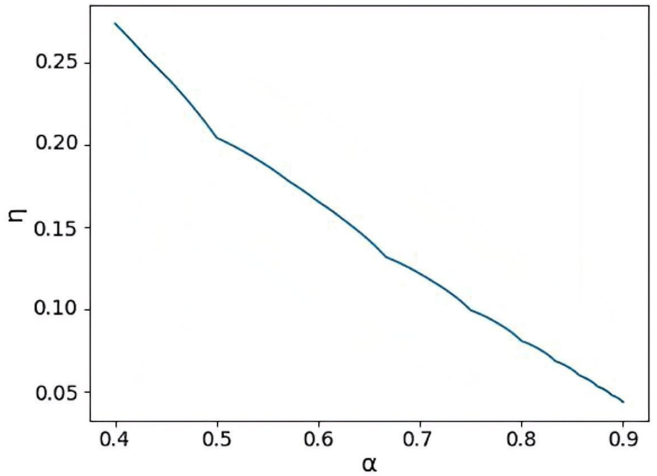

FIG. 3. Same as in Fig. 1, but for the efficiency η_{opt} versus α.

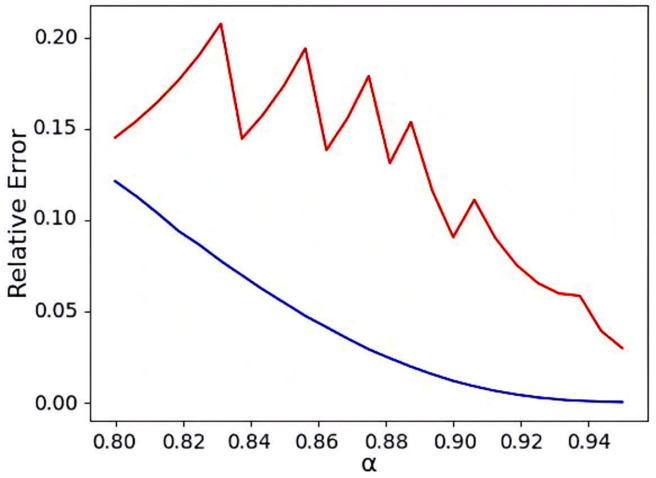

FIG. 4. Here Δn is calculated in a numerically exact way by directly arranging all eigenvalues of the final density matrix $\rho(t)$ in descending order [cf. (47)]. We take $y^\alpha = e^{-\beta\omega_2} = 0.8$ and numerical calculations are done up to block number 100, which is greater than $5m$ [cf. (B1)] for all those α values included in the graph. This quantity is denoted by Δn_{exact}. We denote by Δn_{NN} the nearest-neighbor approximation given by (51). The blue curve shows the relative error $\frac{|\Delta n_{\text{exact}} - \Delta n_{\text{NN}}|}{\Delta n_{\text{NN}}}$ and the red curve $|\frac{y^{2m(\alpha)}}{\Delta n_{\text{NN}}}|$. The red curve is kinked because $m(\alpha)$ is [see (B1)]. It can be seen that the relative error is well within the announced range $|\frac{O(y^{2m(\alpha)})}{\Delta n_{\text{NN}}}|$.

diagram:

$$\begin{array}{c|c|c|c} \hat{n} & 2m & 2m+1 & 2m+2 \\ \hline \rho & \ldots, y^{2m} & \cdots & y^{(2m+2)\alpha}, \ldots \end{array}. \tag{C1}$$

Now note from (B1) that $y^{2m} < y^{(2m+2)\alpha}$, i.e., a next-nearest-neighbor permutation is necessary. Hence the contribution from next-nearest-neighbor permutation scales as $O(y^{2m})$ and for $m \gg 1$ this is smaller than what was retained in (51). This estimate is crude, since it does not account for permutations that already occurred (within the nearest-neighbor approach) between columns $2m$ and $2m+1$. However, it is sufficient for our purposes. Indeed, Fig. 4 shows the relative error of numerically exact calculation of Δn and compares it with (51), showing that it is well within the above bound $O(y^{2m})$.

COP in the limit $\alpha \to 1$

To study the COP K, we can write the mean changes of $\hat{n}_1$ and $\hat{n}_2$ in the approximation of nearest-neighbor permutations [cf. (50)]

$$\Delta n_1 = \xi \sum_{i=\tilde{n}}^{\infty} (y^{\alpha(i+1)} - y^i) \sum_{j=0}^{\infty} y^{j(1+\alpha)} (i+j),$$

$$\Delta n_2 = -\xi \sum_{i=\tilde{n}}^{\infty} (y^{\alpha(i+1)} - y^i) \sum_{j=0}^{\infty} y^{j(1+\alpha)} (i+j+1), \tag{C2}$$

where $\xi = (1 - e^{-\beta\omega_1})(1 - e^{-\beta\omega_2})$ is the normalization factor [cf. (32)]. Note that to obtain (C2) we do not make any permutation within columns with the same eigenvalue of $\hat{n}$ [cf. (47) and (D4)]. Doing such permutations will make the estimates in (C2) closer to the minimal value of Δn_1 and the maximal

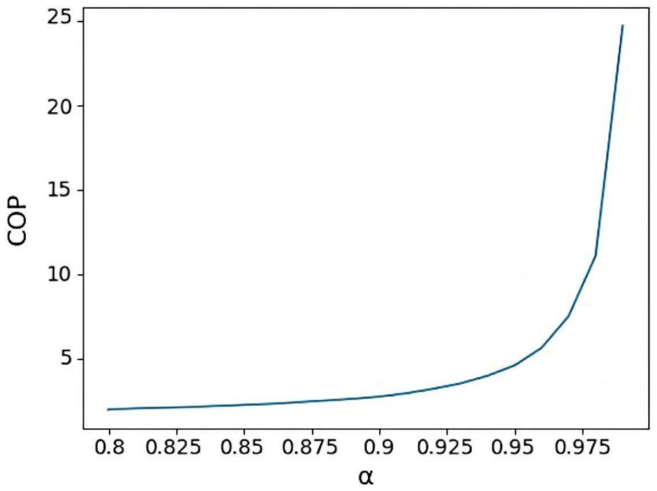

FIG. 5. Numerical results for the COP K. Here $y^{\alpha} = 0.8$ and numerical calculations are done up to block number $5m$ for all those α values included in the graph.

value of Δn_2 (both for a fixed Δn). Hence (C2) suffices for bounding K from below,

$$K \geqslant \frac{1}{(1-\alpha)(\frac{y^{1+\alpha}}{1-y^{1+\alpha}} + \frac{1}{1-y^{\alpha}})} \Rightarrow \lim_{\alpha\to 1} K \to \infty, \tag{C3}$$

which is also observed numerically; see Fig. 5.

APPENDIX D: ASYMPTOTIC RESULTS FOR OPTIMAL COOLING: THE LIMIT $\alpha \to 0$

1. Error estimation

For α finite and sufficiently close to 0, the action of an optimal unitary results in

N	0	1	$\cdots$	a'	$\cdots$
R	1	$y^{\alpha}, y^{2\alpha}$	$\cdots$	$\ldots, y^{m_1\alpha}, y, y^{(m_1+1)\alpha}, \ldots$	$\cdots$

$a'+i$	$\cdots$
$\ldots, y^{(m_2-1)\alpha}, y^2, y^{(m_2+1)\alpha} y^{(m_2+2)\alpha}, \ldots$	$\cdots$

(D1)

where $m_1 = \lfloor 1/\alpha \rfloor$, $m_2 = \lfloor 2/\alpha \rfloor \geqslant 2m_1$, and a' is determined from

$$a'(a'+1)/2 \leqslant m_1 \leqslant (a'+1)(a'+2)/2 \tag{D2}$$

and i from

$$(a'+i)(a'+i+1)/2 \leqslant m_2 \leqslant (a'+i+1)(a'+i+2)/2. \tag{D3}$$

We see that $i < 3a'$. Now we show that for the calculation of averages of photon numbers we can use

N	0	1	2	3	$\cdots$
R	1	$y^{\alpha}, y^{2\alpha}$	$y^{3\alpha}, y^{4\alpha}, y^{5\alpha}$	$y^{6\alpha}, y^{7\alpha}, y^{8\alpha}, y^{9\alpha}$	$\cdots$

(D4)

instead of (D1), as in the limit $\alpha \to 0$ corresponding error terms vanish. We denote by $n_1^{(0)}$ and $n_1^{(*)}$ the average $\hat{n}_1$ calculated with (D1) and (D4), respectively, and by Δn_{1e} the error term $n_1^{(0)} - n_1^{(*)}$. First, we write the contribution from the a'th block to the error term

$$\begin{aligned}\Delta n_{1e}^{a'} &= (y - y^{(m_1+1)\alpha})(m_1 + 1 - u) \\ &\quad + (y^{(m_1+1)\alpha} - y^{(m_1+2)\alpha})(m_1 + 2 - u) + \cdots \\ &\quad + (y^{(u+a'-1)\alpha} - y^{(u+a')\alpha})a',\end{aligned} \tag{D5}$$

where $u = a'(a'+1)/2$. The term $\Delta n_{1e}^{a'}$ can be estimated from above,

$$\begin{aligned}\Delta n_{1e}^{(a')} &\leqslant (y - y^{(m_1+1)\alpha})a' + (y^{(m_1+1)\alpha} - y^{(m_1+2)\alpha})a' \\ &\quad + \cdots + (y^{(u+a'-1)\alpha} - y^{(u+a')\alpha})a' \\ &= (y - y^{(u+a')\alpha})a'.\end{aligned} \tag{D6}$$

Similarly, one can estimate the contribution from the $(a'+1)$th block

$$\Delta n_{1e}^{(a'+1)} \leqslant (y^{(u+a')\alpha} - y^{(u+2a'+2)\alpha})(a'+1). \tag{D7}$$

Summing up all contributions, we get

$$\begin{aligned}\Delta n_{1e} &\leqslant (y + y^{(u+a')\alpha} + y^{(u+2a'+2)\alpha} + y^{(u+3a'+5)\alpha} + \cdots) \\ &\quad + [(y^2 - y^{(m_2+1)\alpha})(a'+i) + (y^3 - y^{(m_3+1)\alpha})(a'+i') \\ &\quad + (y^4 - y^{(m_4+1)\alpha})(a'+i'') + \cdots],\end{aligned} \tag{D8}$$

where $m_3 = \lfloor 3/\alpha \rfloor \geqslant 3m_1$ and $m_4 = \lfloor 4/\alpha \rfloor \geqslant 4m_1$. In (D8) i' and i'' are determined from conditions similar to (D3),

$$\begin{aligned}\frac{(a'+i')(a'+i'+1)}{2} &\leqslant m_3 \leqslant \frac{(a'+i'+1)(a'+i'+2)}{2}, \\ \frac{(a'+i'')(a'+i''+1)}{2} &\leqslant m_4 \leqslant \frac{(a'+i''+1)(a'+i''+2)}{2},\end{aligned} \tag{D9}$$

and result in $i' < 4a'$ and $i'' < 5a'$. Now we can estimate (D8) further,

$$\begin{aligned}\Delta n_{1e} &\leqslant (y + y^2 + y^3 + y^4 + \cdots) \\ &\quad + (y^2 3a' + y^3 4a' + y^4 5a' + \cdots) \\ &\leqslant \frac{y(1+y)}{1-y} + a' \frac{y \ln y}{(1-y)^2}.\end{aligned} \tag{D10}$$

Now note that in the limit $\alpha \to 0$, which is $\omega_1 \to \infty$, a' goes to infinity as $\sqrt{\frac{\omega_1}{\omega_2}}$. As $0 \leqslant \Delta n_{1e} \leqslant \frac{y(1+y)}{1-y} + a' \frac{y \ln y}{(1-y)^2}$ we conclude that (remember that $y = e^{-\omega_1 \beta}$)

$$\lim_{\alpha\to 0} \Delta n_{1e} \to 0. \tag{D11}$$

2. Asymptotic expressions and their integral representations

In our further calculations we use (D4). Using the same procedure as in (48), we find from (D4) the expressions for $\Delta n_{1,2}$ ($\epsilon = \alpha\beta\omega_1$),

$$\begin{aligned}\Delta n_1 &= \xi \sum_{a=0}^{\infty} e^{-\epsilon a(a+1)/2} \sum_{b=0}^{a} e^{-\epsilon b} b - \xi_1 \sum_a a e^{-\beta\omega_1 a} \\ &= \xi \sum_{a=0}^{\infty} \frac{e^{2\epsilon}}{(e^{\epsilon}-1)^2} e^{-(a/2+1)(a+1)\epsilon} [a(e^{-\epsilon}-1) + e^{a\epsilon} - 1] \\ &\quad - \frac{1}{e^{\beta\omega_1} - 1},\end{aligned}$$

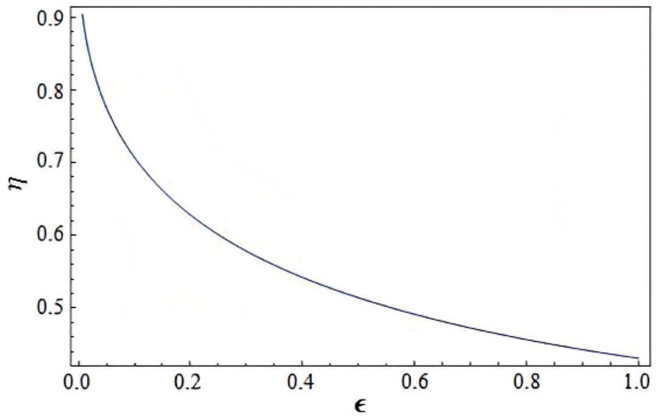

FIG. 6. Numerically calculated efficiency η versus ϵ in the limiting case $\omega_1 \gg \omega_2$ using (D14) and (D15). Here we set $\beta\omega_1 = 10$ and the smallest value of ϵ is 0.007.

$$\Delta n_2 = \xi \sum_{a=0}^{\infty} e^{-\epsilon a(a+1)/2} \sum_{b=0}^{a} e^{-\epsilon(a-b)} b - \xi_2 \sum_{a} a e^{-\epsilon a}$$
$$= \xi \sum_{a=0}^{\infty} \frac{e^{\epsilon}}{(e^{\epsilon}-1)^2} e^{-[a(a+1)/2]\epsilon} [a(e^{\epsilon}-1) + e^{-a\epsilon} - 1]$$
$$- \frac{1}{e^{\epsilon}-1}, \tag{D12}$$

where we define $\xi_1 = 1 - e^{-\beta\omega_1}$ and $\xi_2 = 1 - e^{-\beta\omega_2}$; hence $\xi = \xi_1\xi_2$. Before studying (D12) numerically, we apply Hubbard-Stratonovich transformation

$$e^{-(a^2/2)\epsilon} = \sqrt{\frac{1}{2\pi\epsilon}} \int_{-\infty}^{\infty} dv\, e^{-v^2/2\epsilon - iav} \tag{D13}$$

for faster and more accurate calculations:

$$\sum_{a=0}^{\infty} e^{-(a/2+1)(a+1)\epsilon} [a(e^{-\epsilon}-1) + e^{a\epsilon} - 1]$$
$$= \sqrt{\frac{1}{2\pi\epsilon}} e^{-\epsilon} \int_{-\infty}^{\infty} dv\, e^{-v^2/2\epsilon} \left((e^{-\epsilon}-1) \frac{e^{-[iv+(3/2)\epsilon]}}{(1-e^{-[iv+(3/2)\epsilon]})^2} \right.$$
$$\left. + \frac{1}{(1-e^{-[iv+(1/2)\epsilon]})} - \frac{1}{(1-e^{-[iv+(3/2)\epsilon]})} \right), \tag{D14}$$

$$\sum_{a=0}^{\infty} e^{-[a(a+1)/2]\epsilon} [a(e^{\epsilon}-1) + e^{-a\epsilon} - 1]$$
$$= \sqrt{\frac{1}{2\pi\epsilon}} \int_{-\infty}^{\infty} dv\, e^{-v^2/2\epsilon} \left((e^{\epsilon}-1) \frac{e^{-[iv+(1/2)\epsilon]}}{(1-e^{-[iv+(1/2)\epsilon]})^2} \right.$$
$$\left. + \frac{1}{(1-e^{-[iv+(3/2)\epsilon]})} - \frac{1}{(1-e^{-[iv+(1/2)\epsilon]})} \right). \tag{D15}$$

The results of numerical calculations for η and K are depicted in Figs. 6 and 7. As seen from the figures, $K \to \infty$ and $\eta \to 1$ in the limit $\epsilon \ll 1$. Below we show analytically that indeed K and η reach these limits.

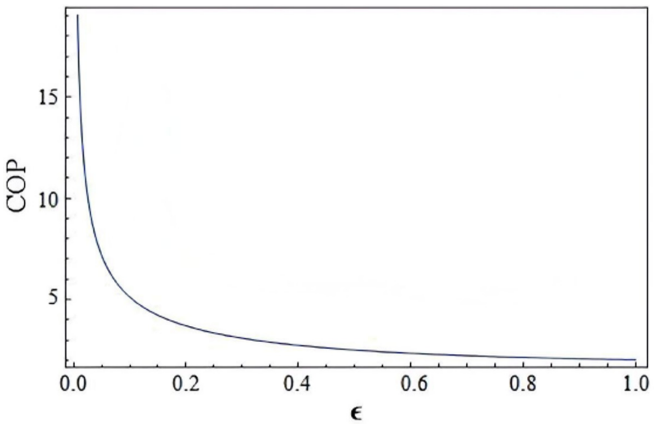

FIG. 7. Same as in Fig. 6 but for COP K.

3. Asymptotic results via the Euler-Maclaurin formula

To study the asymptotics of η and K in the limit $\epsilon \ll 1$ we apply the Euler-Maclaurin formula for the sums in (D12),

$$\sum_{a=0}^{\infty} \frac{e^{2\epsilon}}{(e^{\epsilon}-1)^2} e^{-(a/2+1)(a+1)\epsilon} [a(e^{-\epsilon}-1) + e^{a\epsilon} - 1]$$
$$\equiv \sum_{a=0}^{\infty} f_1(a) = \int_0^{\infty} dx\, f_1(x) + I_1 = S_1 + I_1, \tag{D16}$$

where

$$I_1 = \frac{f_1(0) + f_1(\infty)}{2}$$
$$+ \sum_{k=1}^{\lfloor p/2 \rfloor} \frac{B_{2k}}{(2k)!} \left[f_1^{(2k-1)}(\infty) - f_1^{(2k-1)}(0) \right] + R_p, \tag{D17}$$
$$R_p \leqslant \frac{2\zeta(p)}{(2\pi)^p} \int_0^{\infty} dx \left| f_1^{(p)}(x) \right|. \tag{D18}$$

Here B_{2k} are Bernoulli numbers, $\zeta(p)$ is the Riemann zeta function, and $f^{(p)}(x)$ is the pth-order differential. In (D16) p takes different integer values $p \geqslant 2$ and we use $p = 2$, because this is the simplest case amenable to estimates. Similarly,

$$\sum_{a=0}^{\infty} \frac{e^{\epsilon}}{(e^{\epsilon}-1)^2} e^{-[a(a+1)/2]\epsilon} [a(e^{\epsilon}-1) + e^{-a\epsilon} - 1]$$
$$\equiv \sum_{a=0}^{\infty} f_2(a)$$
$$= \int_0^{\infty} dx\, f_2(x) + I_2 = S_2 + I_2. \tag{D19}$$

The leading diverging terms in (D16) and (D19) when $\epsilon \to 0$ are S_1 and S_2 and we omit I_1 and I_2. Using (D16) and (D19) for the efficiency and COP, we get the relations

$$K \approx -\frac{\xi S_1 - n_{1i} + \xi S_2 - n_{2i}}{\xi S_1 - n_{1i} + \alpha(\xi S_2 - n_{2i})},$$
$$\eta \approx -\frac{\xi S_1 - n_{1i} + \xi S_2 - n_{2i}}{\xi S_1 - n_{1i} - \xi S_2 + n_{2i}}, \tag{D20}$$

where n_{1i} and n_{2i} are initial average occupation numbers. The limits $\lim_{\alpha\to 0}\xi S_{1,2}/n_{2i}$ can be studied analytically and we get

$$\lim_{\alpha\to 0}\frac{\xi S_{1,2}}{n_{2i}}=0.$$

Thus, for $K_{\rm opt}$ and $\eta_{\rm opt}$ we obtain

$$K_{\rm opt}\to\infty,\quad \eta_{\rm opt}\to 1. \tag{D21}$$

APPENDIX E: PERTURBATIVE TREATMENT OF THE FULL NONLINEAR HAMILTONIAN

Let us return to the full, i.e., without the rotating-wave approximation, nonlinear Hamiltonian given by (54).

See (55) for the complete Hamiltonian. Here we will employ (54) in the second order of Dyson's series given by (58) [see in this context (57)]. For simplicity we will scale out the factor β, i.e., we denote βg by g, $\beta\omega_{1,2}$ by $\omega_{1,2}$, and t/β by t.

Using (58) and (57), we get

$$\begin{aligned}\mathrm{tr}[\rho(t)\hat n-\rho(0)\hat n]=O(g^3)+g^2\\ \times\Bigg(\int_0^t ds\,H_I(s)\hat n\int_0^t ds\,H_I(s)\\ -\int_0^t ds_1\int_0^{s_1}ds_2 H_I(s_1)H_I(s_2)\hat n\\ -\hat n\int_0^t ds_1\int_0^{s_1}ds_2 H_I(s_2)H_I(s_1)\Bigg).\end{aligned} \tag{E1}$$

Formally, the same equation holds for $\hat n_k=a_k^\dagger a_k$, where $k=1,2$ and $\hat n=\hat n_1+\hat n_2$.

Substituting (54) into (E1), we get

$$\begin{aligned}\Delta n_k&=\mathrm{tr}[\rho(t)\hat n_k-\rho(0)\hat n_k]\\ &=g^2[A_k\Phi(\omega_1+2\omega_2)+B_k\Phi(\omega_1-2\omega_2)+C_k\Phi(\omega_1)]\\ &\quad+g^2[D_k\Phi(\omega_2+2\omega_1)+E_k\Phi(\omega_2-2\omega_1)+F_k\Phi(\omega_2)],\end{aligned} \tag{E2}$$

where $k=1,2$,

$$\Phi(x)\equiv\frac{4\sin^2(\frac{1}{2}xt)}{x^2}, \tag{E3}$$

$$\begin{aligned}A_1&=\frac{2(e^{\omega_1+2\omega_2}-1)}{(e^{\omega_1}-1)(e^{\omega_2}-1)^2},\quad A_2=\frac{4(e^{\omega_1+2\omega_2}-1)}{(e^{\omega_1}-1)(e^{\omega_2}-1)^2},\\ B_1&=\frac{2(e^{\omega_1}-e^{2\omega_2})}{(e^{\omega_1}-1)(e^{\omega_2}-1)^2},\quad B_2=\frac{-4(e^{\omega_1}-e^{2\omega_2})}{(e^{\omega_1}-1)(e^{\omega_2}-1)^2},\\ C_1&=\frac{4e^{\omega_2}}{(e^{\omega_2}-1)^2},\quad C_2=0.\end{aligned} \tag{E4}$$

Now D_1, E_1, and F_1 are obtained from A_2, B_2, and C_2, respectively, upon swapping ω_1 and ω_2. Likewise, D_2, E_2, and F_2 are obtained from A_1, B_1, and C_1, respectively, upon swapping ω_1 and ω_2.

For a representative pair of frequencies ω_1 and ω_2, Fig. 8 demonstrates to what extent $\Delta n=\Delta n_1+\Delta n_2$ calculated via (E2) predicts cooling, i.e., $\Delta n<0$. As stated in the main text, cooling happens in near-resonance conditions $\omega_2\gtrsim 2\omega_1$ or $2\omega_2\lesssim\omega_1$, which is seen in Fig. 8 (see also Fig. 9 for additional information).

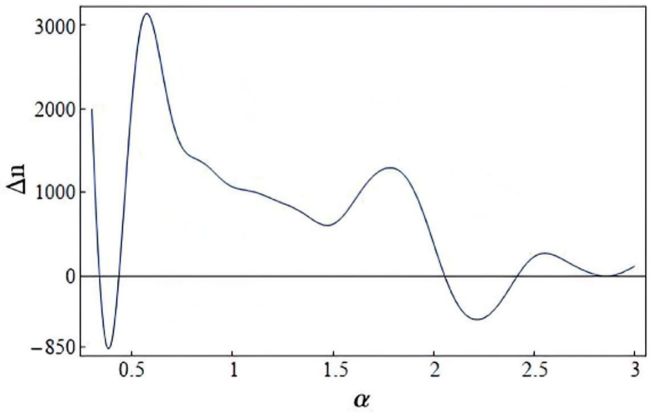

FIG. 8. Photon number difference Δn obtained from (E2) and (E4) for $\omega_1=0.35$ and $t=10\pi$, where $\alpha=\omega_2/\omega_1$. It can be seen that near the resonating frequencies $\alpha\lesssim 0.5$ and $\alpha\gtrsim 2$ the interaction Hamiltonian (54) results in cooling. We see that $\Delta n>0$ (no cooling) for other values of α.

Now the essence of the rotating-wave approximation in (E2) is that, e.g., for $|2\omega_2-\omega_1|\ll\min[\omega_1,\omega_2,|2\omega_1-\omega_2|]$, we can take $\Phi(\omega_1-2\omega_2)$ in (E2) much larger than other terms. This reverts to (65).

Estimation of the higher-order terms in Dyson's series

Using (58), we can show that the terms $O(g^l)$ in Dyson's series [cf. (E1)] are based on the structure

$$\prod_{i=1}^{k}g\int_0^{s_{i-1}}ds_i H_I(s_i)\rho(0)\times\prod_{i=1}^{k'}g\int_0^{s_{i-1}}ds_i H_I(s_{k'-i+1})\hat n, \tag{E5}$$

where $k+k'=l$ and $s_0=t$. To get from (E5) the term $O(g^l)$ in Dyson's series we should take the trace of (E5) and sum it as $\sum_{k=1,\,k'=1;k+k=l}^{l}$.

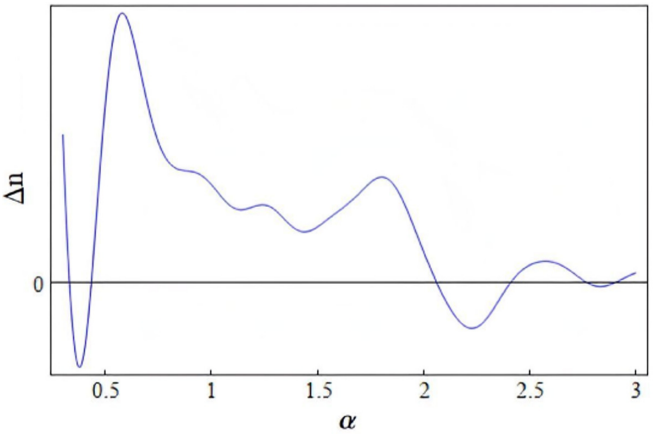

FIG. 9. Photon number difference Δn obtained from (E2) and (E4) for $\omega_1=0.6$ and $t=6\pi$, where $\alpha=\omega_2/\omega_1$. We see that at the nonresonant point $\alpha\approx 2.8$ there are small values of $\Delta n<0$ (cooling). Hence, although very small, cooling can also be achieved far from the resonant frequencies.

To study (E5), let us take its leftmost multiplier

$$\prod_{i=1}^{k} g \int_0^{s_{i-1}} ds_i H_I(s_i) = \sum_{\alpha_1=1}^{8} \cdots \sum_{\alpha_k=1}^{8} \prod_{i=1}^{k} g \int_0^{s_{i-1}} ds_i h_{\alpha_i}(s_i). \tag{E6}$$

Here $\{h_i\}_{i=1}^8$ is the set of all monomials in the interaction Hamiltonian (55):

$$\{h_i\}_{i=1}^8 = \{a_1 a_2^2, a_1^\dagger a_2^2, a_1 a_2^{\dagger 2}, a_1^\dagger a_2^{\dagger 2}, a_1^2 a_2, a_1^{\dagger 2} a_2, a_1^2 a_2^\dagger, a_1^{\dagger 2} a_2^\dagger\}. \tag{E7}$$

Let us also define the frequency set $\{W_i\}_{i=1}^8$,

$$\{W_i\}_{i=1}^8 = \{\omega_1 + 2\omega_2, -\omega_1 + 2\omega_2, \omega_1 - 2\omega_2, -\omega_1 - 2\omega_2, 2\omega_1 + \omega_2, -2\omega_1 + \omega_2, 2\omega_1 - \omega_2, -2\omega_1 - \omega_2\}. \tag{E8}$$

Keeping in mind the equation $a_{1,2}(s) = a_{1,2} e^{-is\omega_{1,2}}$, let us take one term from the sum (E6) corresponding to some $\alpha_1 \cdots \alpha_k$:

$$\prod_{i=1}^{k} g \int_0^{s_{i-1}} ds_i h_{\alpha_i}(s_i) = \prod_{i=1}^{k} g h_{\alpha_i} \times \int_0^{s_{i-1}} ds_i \exp(-is_i W_{\alpha_i}) = \prod_{i=1}^{k} g h_{\alpha_i} \frac{1}{k!} \prod_{i=1}^{k} \int_0^t ds_i \exp(-is_i W_{\alpha_i}). \tag{E9}$$

The last step uses the fact that we have $k!$ ways to order k different items and that after taking the operator part out of the integration we get integration of complex-valued functions which do not change with ordering. Similarly, for the rightmost multiplier of (E5),

$$\prod_{i=0}^{k'-1} g h_{\alpha'_{k'-i}} \frac{1}{k'!} \prod_{i=1}^{k'} \int_0^t ds_i \exp(-is_i W_{\alpha'_i}). \tag{E10}$$

Straightforward calculation shows that the integral terms in (E9) and (E10) result in

$$\prod_{i=1}^{k} \frac{1 - e^{-iW_{\alpha_i} t}}{iW_{\alpha_i}}, \quad \prod_{i=1}^{k'} \frac{1 - e^{-iW_{\alpha'_i} t}}{iW_{\alpha'_i}}. \tag{E11}$$

Now we can write (E5) as

$$\sum_{\alpha_1=1}^{8} \cdots \sum_{\alpha_k=1}^{8} \sum_{\alpha'_1=1}^{8} \cdots \sum_{\alpha'_{k'}=1}^{8} \frac{g^k}{k!} \frac{g^{k'}}{k'!} \times \left(\prod_{i=1}^{k} h_{\alpha_i} \prod_{i=0}^{k'-1} h_{\alpha'_{k'-i}} \times \prod_{i=1}^{k} \frac{1 - e^{-iW_{\alpha_i} t}}{iW_{\alpha_i}} \prod_{i=1}^{k'} \frac{1 - e^{-iW_{\alpha'_i} t}}{iW_{\alpha'_i}} \right) \tag{E12}$$

and the equation for $\Delta n_{1,2}$ will be

$$\Delta n_{1,2} = \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \delta^0_{k+k'} \sum_{\alpha_1=1}^{8} \cdots \sum_{\alpha_k=1}^{8} \sum_{\alpha'_1=1}^{8} \cdots \sum_{\alpha'_{k'}=1}^{8} \frac{g^k}{k!} \frac{g^{k'}}{k'!} \times \prod_{i=1}^{k} \frac{1 - e^{-iW_{\alpha_i} t}}{iW_{\alpha_i}} \prod_{i=1}^{k'} \frac{1 - e^{-iW_{\alpha'_i} t}}{iW_{\alpha'_i}} \times \mathrm{tr}\left(\prod_{i=1}^{k} h_{\alpha_i} \rho(0) \prod_{i=0}^{k'-1} h_{\alpha'_{k'-i}} \hat{n}_{1,2} \right). \tag{E13}$$

Here the sum $\sum_{\alpha_1=1}^8 \cdots \sum_{\alpha_k=1}^8 \sum_{\alpha'_1=1}^8 \cdots \sum_{\alpha'_{k'}=1}^8$ will have 8^l elements for any l, so the number of terms of $O(g^l)$ is $8^l(l+1)$. This may put doubt in the claim that the higher-order $O(g^l)$ terms of Δn can be neglected. However, we believe that it can be done because 8^l is a huge overestimation; for most $\alpha_1 \cdots \alpha_k, \alpha'_1 \cdots \alpha'_{k'}$ the trace

$$\mathrm{tr}\left(\prod_{i=1}^{k} h_{\alpha_i} \rho(0) \prod_{i=0}^{k'-1} h_{\alpha'_{k'-i}} \hat{n}_{1,2} \right) \tag{E14}$$

is zero. Moreover, direct algebraic calculation shows that (E14) is nonzero only if the operator

$$\Theta = \prod_{i=1}^{k} h_{\alpha_i} \prod_{i=0}^{k'-1} h_{\alpha'_{k'-i}} \tag{E15}$$

is Hermitian. For example, for $l = 2$, from 192 terms we get 18 nonzero terms.

[1] M. Sheik-Bahae and R. I. Epstein, Nat. Photon. **1**, 693 (2007).
[2] A. Abragam and M. Goldman, Rep. Prog. Phys. **41**, 395 (1978).
[3] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer Science + Business Media, New York, 2007).
[4] *A Critical Review of Thermodynamics*, edited by E. B. Stuart, A. J. Brainard, and B. Gal-Or (Mono Book Corporation, New York, 1970).
[5] G. Mahler, *Quantum Thermodynamic Processes: Energy and Information Flow at the Nanoscale* (CRC, Boca Raton, 2014).
[6] R. Silva, G. Manzano, P. Skrzypczyk, and N. Brunner, Phys. Rev. E **94**, 032120 (2016).
[7] H. Wilming and R. Gallego, Phys. Rev. X **7**, 041033 (2017).
[8] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, Phys. Rev. Lett. **123**, 170605 (2019).
[9] N. Freitas, R. Gallego, L. Masanes, and J. P. Paz, *Thermodynamics in the Quantum Regime* (Springer, Berlin, 2018), pp. 597–622.
[10] S. Raeisi, Phys. Rev. A **103**, 062424 (2021).

[11] D. Gelbwaser-Klimovsky, W. Niedenzu, and G. Kurizki, Adv. At. Mol. Opt. Phys. **64**, 329 (2015).
[12] R. Uzdin, A. Levy, and R. Kosloff, Phys. Rev. X **5**, 031044 (2015).
[13] P. Liuzzo-Scorpo, L. A. Correa, R. Schmidt, and G. Adesso, Entropy **18**, 48 (2016).
[14] R. Long and W. Liu, Physica A **443**, 14 (2016).
[15] J. Gonzalez-Ayala, A. Medina, J. M. M. Roco, and A. C. Hernández, Phys. Rev. E **97**, 022139 (2018).
[16] V. Singh, T. Pandit, and R. S. Johal, Phys. Rev. E **101**, 062121 (2020).
[17] S. Raeisi and M. Mosca, Phys. Rev. Lett. **114**, 100404 (2015).
[18] P. Taranto, F. Bakhshinezhad, A. Bluhm, R. Silva, N. Friis, M. P. Lock, G. Vitagliano, F. C. Binder, T. Debarba, E. Schwarzhans *et al.*, arXiv:2106.05151.
[19] A. E. Allahverdyan, K. V. Hovhannisyan, D. Janzing, and G. Mahler, Phys. Rev. E **84**, 041109 (2011).
[20] A. E. Allahverdyan, K. Hovhannisyan, and G. Mahler, Phys. Rev. E **81**, 051129 (2010).
[21] M. T. Weiss, Proc. IRE **45**, 1012 (1957).
[22] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media* (Elsevier, Amsterdam, 2013), Vol. 8.
[23] C. M. Caves, Phys. Rev. D **26**, 1817 (1982).
[24] J. Garrison and R. Chiao, *Quantum Optics* (Oxford University Press, Oxford, 2008).
[25] K. V. Hovhannisyan, F. Barra, and A. Imparato, Phys. Rev. Res. **2**, 033413 (2020).
[26] L. Madeira, A. D. García-Orozco, F. E. A. Dos Santos, and V. S. Bagnato, Entropy **22**, 956 (2020).
[27] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer Science + Business Media, New York, 2011).
[28] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications* (Academic, New York, 1979).
[29] A. B. Cruse, Linear Algebra Appl. **12**, 21 (1975).
[30] C. Lindblad, *Non-Equilibrium Entropy and Irreversibility* (Springer Science + Business Media, New York, 2001), Vol. 5.
[31] A. Allahverdyan and T. M. Nieuwenhuizen, Physica A **305**, 542 (2002).
[32] A. E. Allahverdyan, R. S. Johal, and G. Mahler, Phys. Rev. E **77**, 041118 (2008).
[33] R.-B. Wu, C. Brif, M. R. James, and H. Rabitz, Phys. Rev. A **91**, 042327 (2015).
[34] S. Lloyd and S. L. Braunstein, *Quantum Information with Continuous Variables* (Springer, Berlin, 1999), pp. 9–17.
[35] G. New, *Introduction to Nonlinear Optics* (Cambridge University Press, Cambridge, 2011).
[36] M. Hillery, Acta Phys. Slovaca **59**, 1 (2009).
[37] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, Rev. Mod. Phys. **86**, 1391 (2014).
[38] A. Belenchia, D. M. T. Benincasa, S. Liberati, F. Marin, F. Marino, and A. Ortolan, Phys. Rev. Lett. **116**, 161303 (2016).
[39] F. Armata, L. Latmiral, A. D. K. Plato, and M. S. Kim, Phys. Rev. A **96**, 043824 (2017).