ARTICLES
https://doi.org/10.1038/s41566-019-0501-8
nature photonics

Thermodynamic theory of highly multimoded nonlinear optical systems

Fan O. Wu, Absar U. Hassan and Demetrios N. Christodoulides*

Lately, there has been a resurgence of interest in nonlinear multimode optical systems. The sheer complexity emerging from the presence of a multitude of nonlinearly interacting modes has led not only to new opportunities in observing a host of novel optical effects but also to new theoretical challenges in understanding their collective dynamics. Here, we present a consistent thermodynamical framework capable of describing in a universal fashion the exceedingly intricate behaviour of such photonic configurations. In this respect, we derive new equations of state and show that both the 'internal energy' and optical power always flow in accord to the second law of thermodynamics. The laws governing isentropic processes are derived and the prospect for realizing Carnot-like cycles is presented. In addition to shedding light on fundamental issues, our work may pave the way towards a new generation of high-power multimode optical structures and could have ramifications in other disciplines, such as Bose–Einstein condensates and optomechanics.

Recent years have witnessed a strong comeback of multimode fibre technologies, largely in anticipation of high-speed communication systems that benefit from space division multiplexing[1–6]. These activities have, in turn, incited a renewed attention in the nonlinear properties of such many-mode structures[7–9] aimed at establishing new platforms for high-power fibre-based light sources[10]. During the course of this effort, a number of intriguing processes have been observed that have no counterpart whatsoever in single-mode settings. These include, for example, geometric parametric instabilities[11–14], spatiotemporal mode-locking[10], efficient supercontinuum generation[12,15], and the formation of multimode solitons along with a novel class of Cherenkov dispersive wave lines[16,17], to mention a few. In the same vein, in three independent studies, a peculiar effect was found to take place in such nonlinear multimode environments whereby the optical power gradually flowed towards the lowest group of modes[12,18,19]. This beam self-cleanup mechanism, which so far remains poorly understood, seems to result from the conservative component of the Kerr nonlinearity, while having no ties to any stimulated Raman and/or Brillouin effects. Even though one can in principle address this perplexing behaviour by resorting to global or multimode-nonlinear wave solvers[20], it is still impossible to either predict or decipher the convoluted response of such heavily multimoded systems. For example, just setting up a multimode propagation code for individually tracking M modes will first require a computation of M different dispersion curves, M self-phase modulation coefficients, and asymptotically M^2 cross-phase modulation constants and M^4 four-wave mixing products[7,21]—something that is virtually unattainable, especially in systems involving thousands of modes. More importantly, such approaches lend no insight as to how this energy exchange between modes transpires or how it can be harnessed to one's advantage. Clearly of interest will be to develop an appropriate formalism capable of providing the laws that dictate the collective dynamics of such optical multimoded nonlinear configurations. To some extent, this calls for a theory akin to that of thermodynamics or statistical mechanics that are known to serve as powerful tools in understanding the macroscopic properties of various phases of matter[22,23]. More striking is the fact that such descriptions can capture the physics of these many-body systems (even when involving Avogadro-like numbers), often without delving into the underlying nature of particle–particle interactions. In this regard, non-equilibrium kinetic formulations based on optical wave turbulence theories have been put forward to understand such behaviours in a number of settings[24,25]. Yet, an equilibrium thermodynamic theory capable of describing such process in heavily nonlinear multimode optical structures is still lacking.

Here, we show that under thermal equilibrium conditions, the nonlinear evolution dynamics in conservative optical arrangements with a large but finite number of modes can be rigorously described through a comprehensive thermodynamic formulation. These systems come with their own laws. The results derived here are universal in the sense that they apply to both continuous and discrete systems that evolve in either space or time—irrespective of the specific type of nonlinearity involved. During the process of thermalization, the total entropy always increases in such a way that the 'internal energy' flows from a hotter to a colder subsystem while any exchange of optical power is driven by the difference in chemical potentials. In this respect, we derive a new set of equations of state and we cast the fundamental thermodynamic equation of entropy in terms of the extensive variables associated with the internal energy, number of modes and optical power. Once the eigenspectrum of a specific system is known, one can then uniquely predict its equilibrium state from these three conserved quantities. In addition, the invariants governing isentropic compressions or expansions are presented. Finally, we discuss the possibility for negative temperatures—an equilibrium regime that happens to be completely opposite to that of beam self-cleaning.

Theoretical formulation

To illustrate our approach, let us consider an arbitrary nonlinear multimode optical waveguide supporting a finite number of M bound states—all propagating along the axial direction z. In general, this configuration can be continuous in nature[21], with an elevated refractive index profile $n(x, y)$, or discrete, like, for example a multicore optical fibre or waveguide array[26,27]. Each mode i is associated with a particular propagation constant and an orthonormal eigenfunction

CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL, USA. *e-mail: demetri@creol.ucf.edu

$|\psi_i\rangle$ as obtained from the pertinent eigenvalue problem. The distribution of the normalized propagation eigenvalues ε_i constitutes the eigenspectrum of the system (see Supplementary Information for all normalizations used in this work). In all occasions, the optical power $\mathcal{P}$ propagating in this system is conserved—as expected under continuous-wave or broad pulse excitation conditions[12,18,19]. An additional invariant is the system's Hamiltonian, which comprises linear and nonlinear components, that is, $H=H_{\mathrm{L}}+H_{\mathrm{NL}}$ (see Supplementary Information). As in recent experiments, we will assume that the power levels in this nonlinear multimode arrangement are relatively low and hence the Hamiltonian is heavily dominated by the linear contribution, $H\simeq H_{\mathrm{L}}$. As we will see, the role of nonlinearity is to allow for a random power exchange among the various modes. To some extent, this is analogous to a 'diluted gas of particles' whose internal energy is dominated by its kinetic part while the corresponding intermolecular potential energy can be neglected—even though it is responsible for thermalization through particle collisions[22].

At this point, the conserved internal energy U of this optical multimode system is defined as $U=-H_{\mathrm{L}}$, where the expectation value of the linear Hamiltonian operator is given by $H_{\mathrm{L}}=\langle\Psi\,|\hat{H}_{\mathrm{L}}|\,\Psi\rangle=\sum_{i=1}^{M}\varepsilon_i\,|c_i|^2$ with $\hat{H}_{\mathrm{L}}$ being the Hamiltonian operator describing the linear evolution of the state vector $|\Psi\rangle$ via $i\mathrm{d}|\Psi\rangle/\mathrm{d}z=-\hat{H}_{\mathrm{L}}|\Psi\rangle$ (Supplementary Information). In this latter expression, $|c_i|^2$ represent modal occupancy coefficients and are related to the total power via $\mathcal{P}=\sum_{i=1}^{M}|c_i|^2$. Note that the two independent variables U and $\mathcal{P}$ are completely determined by initial excitation conditions, that is, $\mathcal{P}=\sum_{i=1}^{M}|c_{i0}|^2$ and $U=-\sum_{i=1}^{M}\varepsilon_i\,|c_{i0}|^2$ where $c_{i0}=\langle\psi_i|\Psi_0\rangle$ stands for the complex coefficients resulting from the projection of the input field $|\Psi_0\rangle$ on the respective modes $|\psi_i\rangle$, right at the input. While the modal occupancies $|c_i|^2$ can vary significantly during propagation as a result of weak nonlinearity—a necessary ingredient for thermalization—they are always reshuffled in a manner that U and $\mathcal{P}$ remain invariant. A similar discussion holds for conservative nonlinear optical multimode-cavity configurations[28,29] that evolve in time, where in this case, the eigenvalues ε_i are now expressed in the frequency domain while the stored energy E plays the role of U.

In laying down our formulation, we assume that the non-integrability of the underlying nonlinear interactions enables the system to behave ergodically, thus allowing it to explore its constant energy (U) and power ($\mathcal{P}$) manifolds or all its accessible microstates in a fair manner—covering it uniformly with respect to the microcanonical probabilities[30]. Within this isolated microcanonical ensemble, we may then pose the following question: If the total optical power is subdivided into a very large number of indistinguishable packets (each carrying the same infinitesimal amount of power), in how many ways can one distribute them into M distinct modes—subject to the constraint imposed by U and $\mathcal{P}$ being constant? Given that, under these conditions, the number of such packets per state is exceedingly high, maximization of entropy directly leads to a subcase of the Bose–Einstein distribution—the so-called Rayleigh–Jeans (RJ) distribution (see Methods). In this case, after thermalization, the expectation values of the mode occupancies can be obtained from $|c_i|^2=-T/(\varepsilon_i+\mu)$, where the optical temperature T and chemical potential μ result from the Lagrange multipliers associated with the constraints imposed by U and $\mathcal{P}$, respectively. Here, the dimensionless quantities T and μ represent intensive properties of this system and have nothing to do with the actual thermal environment the optical multimode arrangement is embedded in. Along these lines, the entropy of the system can now be obtained from Boltzmann's expression ($S=\ln W$, where S is the entropy and W is number of accessible microstates), which leads to (see Methods)

$$S=\sum_{i=1}^{M}\ln\,|c_i|^2 \tag{1}$$

Interestingly, the entropy in equation (1) can in addition describe non-equilibrium processes and provides a basis in establishing a particular form of Boltzmann's H function[31]. By adopting non-equilibrium formulations, previous studies have considered multiwave mixing kinetics in extended nonlinear systems, ranging from $\chi^{(2)}$ materials to saturable and non-local media[24,32–36]. In contrast, thermal equilibrium (leading to an RJ distribution) can only be reached in non-extended nonlinear optical arrangements (having a finite number of modes), as long as the underlying nonlinear mechanisms promote dynamical chaos so as to establish ergodicity. In this respect, chaos is expected to take place in the system whenever the additional invariants needed for integrability (besides the Hamiltonian and the norm) are absent. The RJ distribution can result under general conditions[37–39] (see also Methods), an aspect that was also highlighted in previous studies[33,40]. We note that the nature of the equilibrium thermodynamical problem at hand is such that it requires a modal formulation, as opposed to a local description of amplitude-phase statistics at distinct sites[41–45].

Equation of state

From the RJ distribution and the two system invariants U and $\mathcal{P}$, we can then derive the following equation of state (see Methods)

$$U-\mu\mathcal{P}=MT \tag{2}$$

which relates the two intensive variables T and μ to the three extensive quantities U, M and $\mathcal{P}$. To some extent, this equation is analogous to that of an ideal gas, for example, $pV=Nk_{\mathrm{B}}T$ (ref. [22]). Equation (2) can be used to uniquely determine both the temperature and chemical potential at thermal equilibrium, once the internal energy and power are specified at the input of a nonlinear optical system with M modes (see Methods). Figure 1a shows a schematic of an optical multimode fibre involving $M\approx480$ modes, when excited at a normalized power $\mathcal{P}=42.7$ and internal energy $U=3.2\times10^3$, as dictated by the input power distribution among modes (Fig. 1b). For these initial conditions, our theoretical model predicts at steady state, a temperature $T=5.37$ and a chemical potential $\mu=-136.2$. Numerical simulations performed on this same microcanonical ensemble system are in excellent agreement with these predictions, once thermalization is attained (that is, the entropy S is maximized). Note that in this case, most of the power eventually finds its way to the lowest group of modes (Fig. 1b)—a behaviour consistent with the process of beam self-cleaning observed in experiments. A similar scenario was previously considered in ref. [40], where the use of a finite number of modes was crucial in regularizing the ultraviolet catastrophe that could result from the RJ distribution. In actual systems, this thermalization can be accelerated as a result of external perturbations like, for example, small random mode coupling effects, intensity and phase fluctuations in laser emission, fibre vibrations and so on[46]. We next carry out simulations on a discrete (multicore) nonlinear optical waveguide array, in this case a Lieb lattice (Fig. 1c), which is known to exhibit a massive degree of degeneracy in its flat band, where $\varepsilon_i=0$. For the U and $\mathcal{P}$ input parameters used in this example (as specified in Fig. 1d), we now theoretically expect a negative temperature $T=-0.76$ with a chemical potential $\mu=6.6$. Again, the numerically obtained final power distribution among modes, after thermalization, is in full agreement with that anticipated from theory—only this time the power tends to flow towards the highest group of modes (Fig. 1d), in direct contrast to beam self-cleaning. Negative temperatures are common to systems with a finite number of modes, when the internal energy exceeds the mean value of the eigenvalue spectrum. Such configurations are known to be 'hotter than hot'[47–49], that is, the internal energy should always flow from a negative temperature region to a positive one. Similar results are also obtained for a three-dimensional array of optical cavities (coupled resonator optical waveguides, CROWs) where thermalization takes place instead

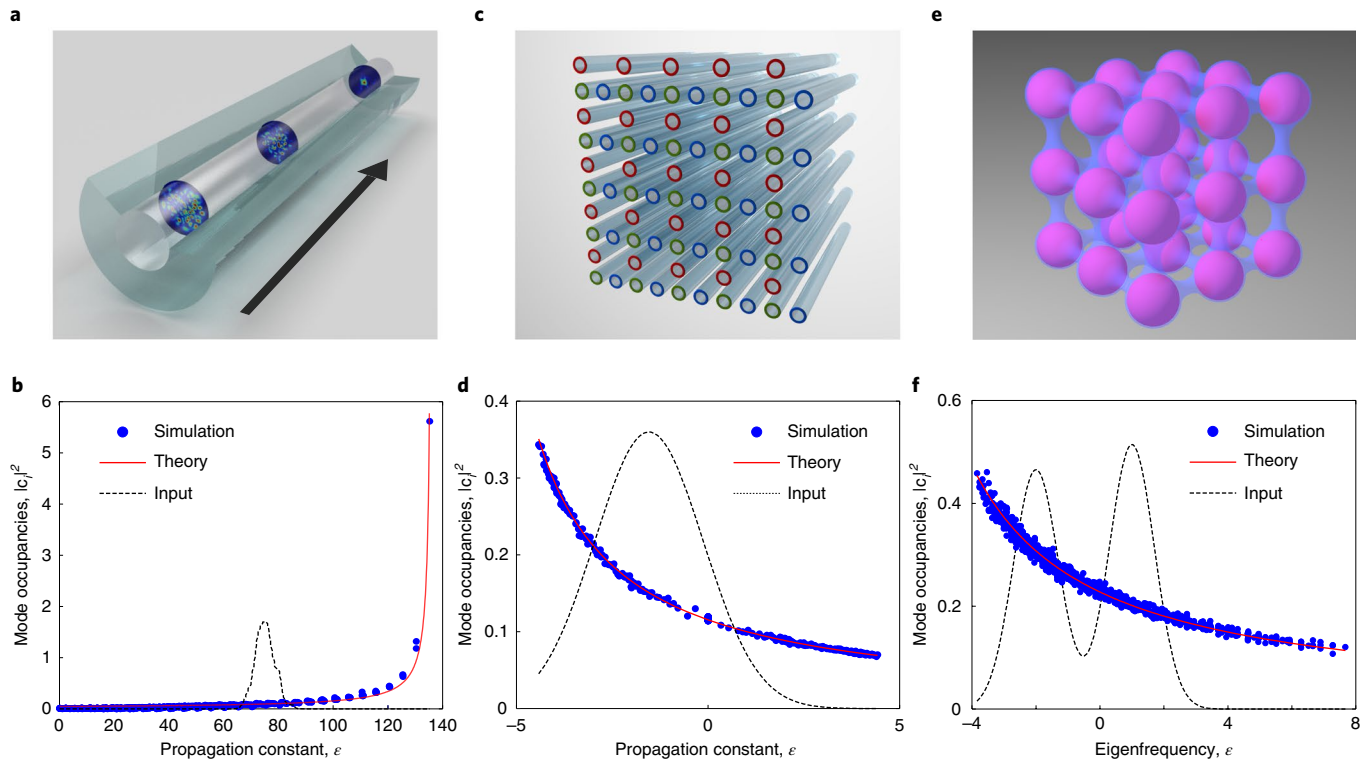

Fig. 1 | Thermalization dynamics in nonlinear multimode waveguides and cavity structures. a, A schematic of a graded-index multimode fibre system indicating the onset of thermalization, manifested as beam self-cleaning during propagation. **b**, The resulting RJ distributions after thermalization versus normalized propagation constants. In this case, the final optical temperature is positive, $T=5.37$, indicating a pronounced population in the lower-order modes. **c**, A nonlinear multicore Lieb waveguide lattice. **d**, Corresponding modal distributions at thermal equilibrium. Here, the final temperature is negative, $T=-0.76$, thus promoting significant population in higher-order modes. The heavy degeneracy in the flat band of the Lieb lattice does not affect the thermalization process. **e**, A three-dimensional coupled cavity nonlinear system. **f**, Resulting modal distributions after thermalization in time in a configuration involving 1,000 resonators. The final temperature and chemical potential at equilibrium are $T=1.75$ and $\mu=-7.7$. In all figures, the dashed curves represent the modal excitation distribution at the input, with random initial phases. For more information, see Supplementary Information.

in time (Fig. 1e,f). In all our simulations, we make sure that the linear part of the Hamiltonian remains invariant during evolution and hence the system behaves in a quasi-linear manner for the particular power levels used. In this regime, any soliton formation (representing a phase transition) is inhibited.

Extensivity of entropy and Euler equation

We next express the entropy S of a microcanonical optical system as a function of three extensive variables, $S=S(U,M,\mathcal{P})$, in a way similar to that employed in standard thermodynamics[22], where now the number of modes plays the role of volume and the optical power is analogous to the number of particles involved. As opposed to alternative formulations provided in previous studies, we here establish the fundamental thermodynamic equation $S(U,M,\mathcal{P})$ in terms of all these three conserved quantities. In fact, failure to do so, violates the very extensivity of the entropy itself. To justify this argument, let us double, for example, the input power and energy $(\mathcal{P},U)$, as well as the number of modes M in a system. From the equation of state, equation (2), one quickly concludes that in this case, the temperature and chemical potential remain the same as before enlarging this arrangement. We emphasize that in doubling the number of modes, the eigenspectrum distribution should remain invariant, in other words, the 'material composition' of the system should not change (see Methods). Under these conditions, each eigenvalue ε_i splits into a closely spaced doublet, and hence from $S=\sum \ln |c_i|^2$, we find that $S\to 2S$. This directly implies that for an arbitrary proportionality factor λ, $S(\lambda U,\lambda M,\lambda\mathcal{P})=\lambda S(U,M,\mathcal{P})$—hence guaranteeing the extensivity of the entropy with respect to $(U,M,\mathcal{P})$, as required by a self-consistent thermodynamic theory. From here, one can obtain the following conjugate intensive variables: temperature, T, chemical potential, μ, and 'pressure', $\hat{p}$, via

$$\frac{1}{T}=\frac{\partial S}{\partial U},\ \frac{\mu}{T}=-\frac{\partial S}{\partial \mathcal{P}},\ \frac{\hat{p}}{T}=\frac{\partial S}{\partial M} \tag{3}$$

We note that the aforementioned definitions for $1/T$ and μ/T are congruent with the equation of state equation (2) (see Methods). Meanwhile, it can be readily shown that $\partial S/\partial S=\hat{p}/T=(S/M)-1$. By using the fact that the entropy S is a homogeneous function of $(U,M,\mathcal{P})$, the Euler equation (see Methods) leads to a second equation of state that now relates seven thermodynamic variables in this nonlinear multimode optical system, that is

$$ST=U-\mu\mathcal{P}+\hat{p}M \tag{4}$$

Thermodynamic forces

We now illustrate the versatility of the formalism developed above in predicting the thermodynamic behaviour of complex, nonlinear heavily multimoded optical arrangements. Figure 2 illustrates two square lattice systems each involving 20×10 sites. The array on the left is excited with a left-hand circular polarization where the input power and energy are $\mathcal{P}_{\mathrm{LHP}}=41.5$ and $U_{\mathrm{LHP}}=15.4$. In contrast, the lattice on the right is excited with a right-hand circular

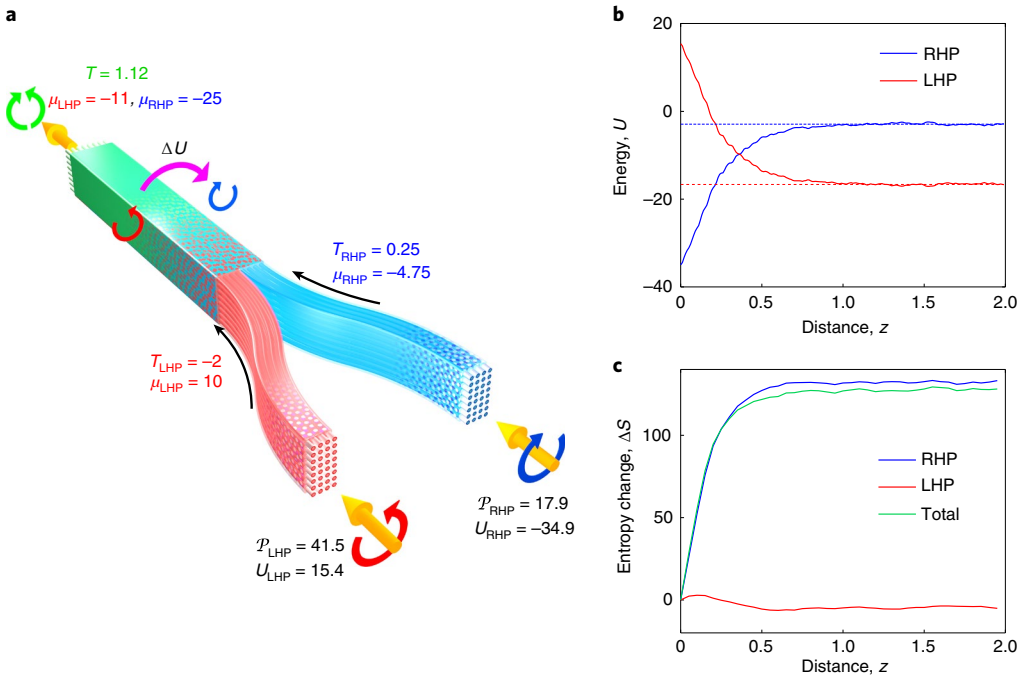

Fig. 2 | Thermalization in a canonical-like optical multimode nonlinear configuration involving two circular polarizations. a, A schematic of two separate waveguide arrays, each with 200 sites. The lattice on the left is excited with a left-circular polarization (LHP) while the one on the right with a right-handed polarization (RHP). The corresponding initial power and internal energies are listed in the figure. Once the two subsystems reach thermal equilibrium, they merge together. **b**, After merging, the two species (LHP and RHP) are now allowed to exchange energy ΔU, from hot (red) to cold (blue), as also indicated in **a**. Eventually the two circular polarizations reach a common temperature $T=1.12$, albeit with different chemical potentials, as predicted by theory and in agreement with numerical simulations. Dashed lines indicate the theoretical predictions for the thermalized energy in each of the two species. **c**, Corresponding changes in entropy for the two polarizations in the combined lattice, given that $S_T = S_{LHP} + S_{RHP}$.

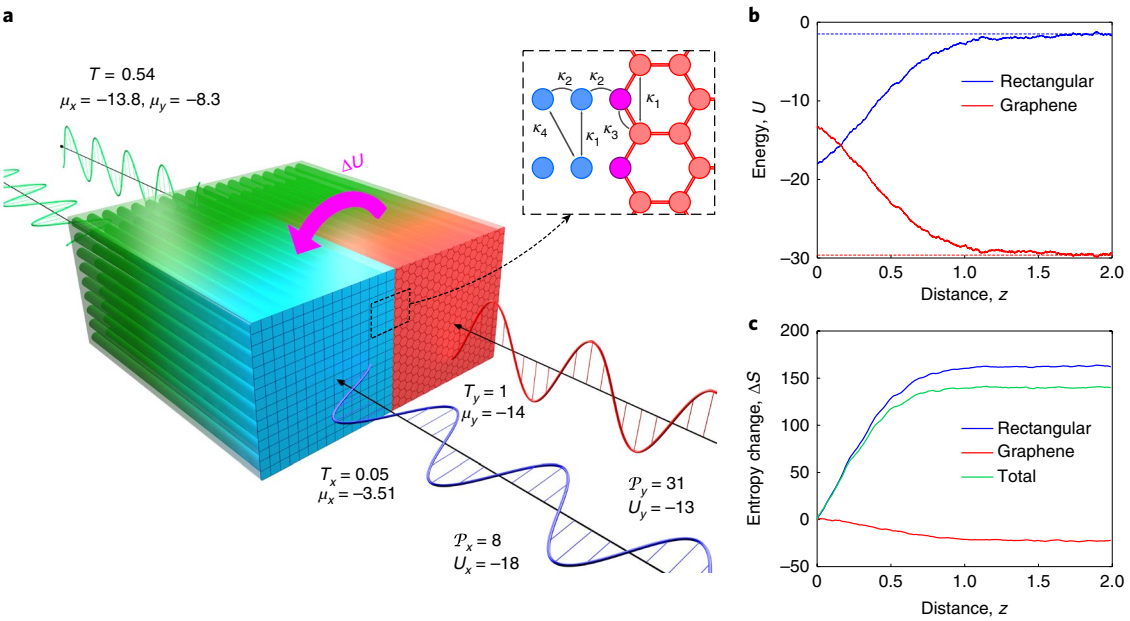

Fig. 3 | Thermalization of two different optical waveguide lattices in thermal contact. a, An optical graphene-like array (right) with 420 sites is brought in contact with a 200-site rectangular lattice (left) as shown in the inset, after the two lattices separately reach thermal equilibrium. The graphene lattice is excited with $\hat{y}$-polarized light and the rectangular one with $\hat{x}$. The thermal contact layer (shown in purple) allows for energy exchange ΔU via polarization cross-phase modulation while preventing any power transfer between the two lattices. The two subsystems, set initially at $(T_y, \mu_y) = (1, -14)$ and $(T_x, \mu_x) = (0.05, -3.51)$, eventually reach a common temperature $T=0.54$. Unlike in Fig. 2, where the two gases share the same vessel, here energy transfer takes place through a single diathermic layer. **b**, Evolution of energy transfer as a function of distance. Dashed lines show the predictions for energy in each of the two species at equilibrium. **c**, Corresponding entropy changes in the two lattices where $S_T = S_x + S_y$.

polarization with $\mathcal{P}_{RHP}=17.9$ and $U_{RHP}=-34.9$. Numerical simulations show that these two subsystems finally reach thermal equilibrium at $(T_{LHP}, \mu_{LHP})=(-2, 10)$ and $(T_{RHP}, \mu_{RHP})=(0.25, -4.75)$, in full accord with theory. Eventually, they are brought in contact, thus forming a regular square lattice with 400 sites. In this new canonical-like ensemble, the total energy $U_T=U_{LHP}+U_{RHP}$ is conserved, while the respective power components $\mathcal{P}_{LHP}$ and $\mathcal{P}_{RHP}$ in each polarization remain invariant. After merging, the two polarizations start to exchange energy dU through cross-phase modulation (Supplementary Information), where the change in the total entropy (S_T) is governed by the second law of thermodynamics, $dS_T=(T_{LHP}^{-1}-T_{RHP}^{-1})dU_{LHP}\geq 0$, (see Methods). This implies that the internal energy always flows from hotter to colder objects, in this case from the left-circular polarization states to the right. Once the composite system is thermalized at equilibrium, the two polarizations attain the same temperature $T=1.12$ while the chemical potentials settle to different values $\mu_{LHP}=-11$ and $\mu_{RHP}=-25$ (since no power exchange occurs), as predicted by theory and confirmed by simulations. The situation here is analogous to that of an ideal gas involving two species of particles (such as oxygen and nitrogen), set at different initial temperatures and then allowed to mix in the same vessel—thus reaching the same final temperature but different chemical potentials. In all cases, maximization of the entropy indicates optical thermal equilibrium. Figure 3 depicts another situation where two different array subsystems or 'solids' having different band structures are brought into thermal contact. The rectangular lattice on the left is excited with $\hat{x}$-polarized light and is at thermal equilibrium $(T_x, \mu_x)=(0.05, -3.51)$. Meanwhile, light in the right (graphene) lattice is $\hat{y}$-polarized and at steady state reaches $(T_y, \mu_y)=(1, -14)$. The thermal contact layer in between allows the two linear polarizations to locally interact via cross-phase modulation without again exchanging power, that is, $\mathcal{P}_x$ and $\mathcal{P}_y$ remain invariant. For a possible design, see Supplementary Methods and Supplementary Fig. 2. Unlike the previous situation, the two polarizations are confined in their respective arrays, only exchanging energy U through the thermally permeable wall. In other words, this is reminiscent of two different solids, initially kept at different temperatures and then brought together. Again, in this case, the two subsystems reach the same temperature but different chemical potentials since there is no power transfer.

Figure 4 shows the possibility of two different optical nonlinear multimoded subsystems (a Lieb and a square lattice) exchanging both energy and power, while the total energy $U_T=U_x+U_y$ and total power $\mathcal{P}_T=\mathcal{P}_x+\mathcal{P}_y$ are conserved. The Lieb subsystem on the left has 300 sites and is excited with $\hat{x}$-polarized light while the one on the right involves 400 modes and carries the $\hat{y}$ polarization. This grand canonical-like ensemble is analogous to that of two solids or gases that simultaneously allow both heat and particle transfer. A possible design for such an optically permeable wall that allows in addition, four-wave mixing mediated power exchange between the $\hat{x}$ and $\hat{y}$ polarizations, is provided in Supplementary Methods and Supplementary Fig. 3. As before, cross-phase modulation is responsible for energy exchange. Before merging together, the subsystems are at thermal equilibrium having $(T_x, \mu_x)=-0.17, 4.21)$ and $(T_y, \mu_y)=(0.065, -4.956)$. Once in contact, computer simulations show that they now attain not only the same temperature $T=-0.42$ but also the same chemical potential $\mu=6.46$, in excellent agreement with results anticipated from the theoretical formalism developed above. Again, the response of the combined system is driven by the second law of thermodynamics, since now $dS_T=(T_x^{-1}-T_y^{-1})dU_x+(\mu_y T_y^{-1}-\mu_x T_x^{-1})d\mathcal{P}_x\geq 0$. This latter equation directly implies that, in addition to having energy transfer from a hot to a cold object, the optical power will flow between the two polarizations in this multimode arrangement towards the subsystem with a lower chemical potential until thermal equilibrium is reached.

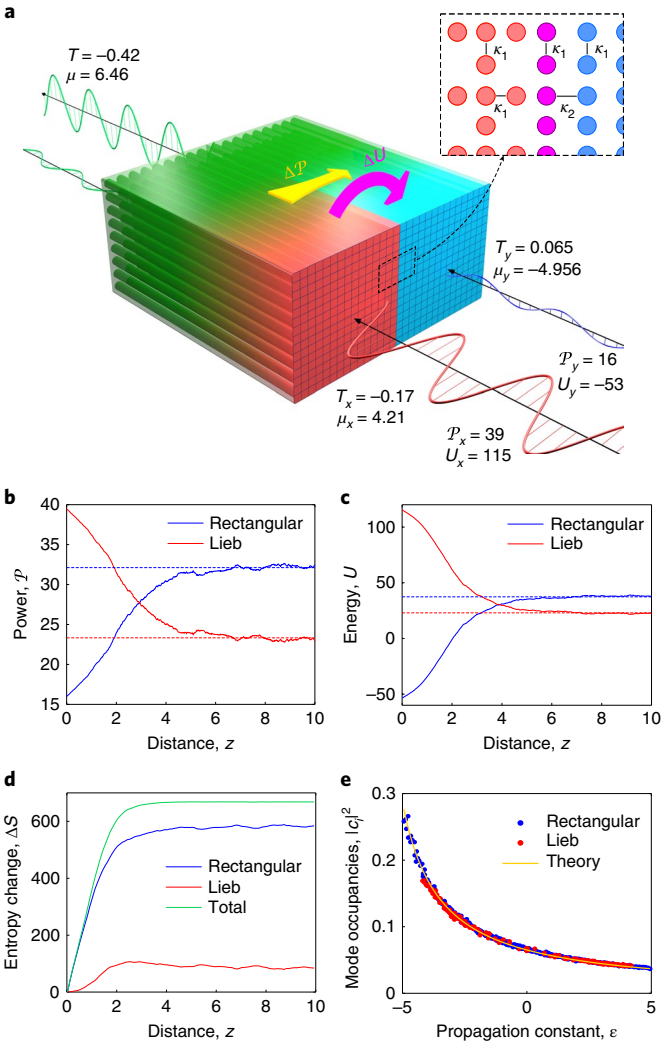

Fig. 4 | Thermalization in a grand canonical-like configuration involving a Lieb and a rectangular nonlinear optical lattice. a, As in Fig. 3, the two different arrays are excited with different polarizations and are brought together after each one of them reaches thermal equilibrium for the parameters provided in the figure. In addition to allowing energy transfer ΔU, the contact layer is now designed to permit power exchange $\Delta\mathcal{P}$ between the two polarizations. **b**,**c**, Evolution of optical power (**b**) and internal energy (**c**) during propagation. In this case, both the optical power and internal energy settle down to the predicted values (dashed lines). **d**, Corresponding entropic evolution for the entire system and the two subsystems. **e**, Resulting RJ distribution when the entire system finally thermalizes at the same negative temperature $T=-0.42$, as predicted by theory and confirmed numerically. In this case, the two species reach the same chemical potential as well.

Isentropic expansions/compressions and Carnot-like cycles

We now consider the interesting possibility of an all-optical Carnot-like cycle mediated by successive adiabatic or isentropic ($dS=0$) expansions and compressions, taking place between a cold and a hot subsystem. An example of a nonlinear multimode array undergoing an adiabatic expansion is shown in Fig. 5a. During this process, the number of modes remains the same while the discrete Hamiltonian energy U decreases during expansion as a result of the reduction in coupling strengths (see Methods). Similarly, during compression the energy increases. Under adiabatic conditions, the mode occupancies $|c_i|^2$ remain invariant and as a result the process

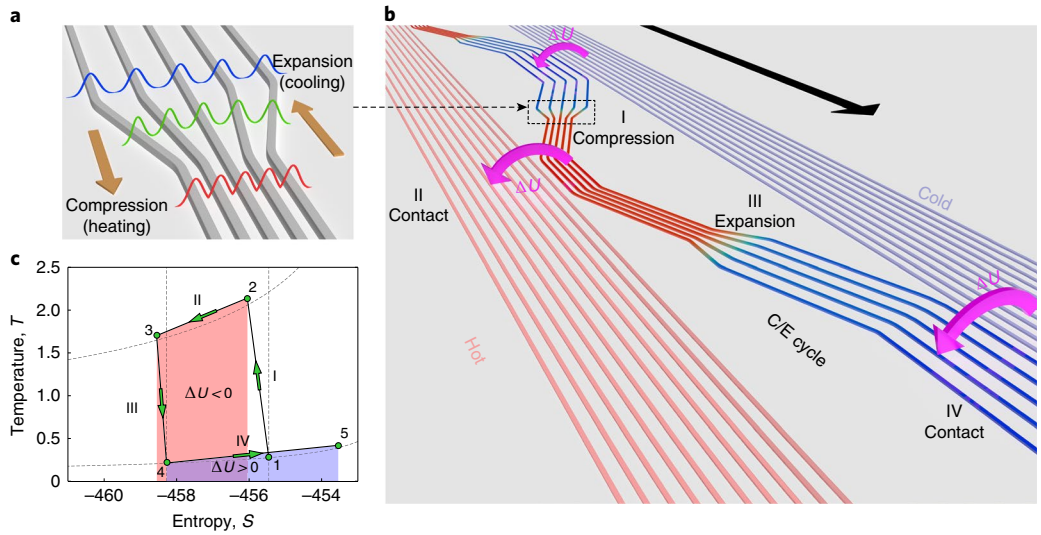

Fig. 5 | All-optical refrigeration via Carnot cycles. a, During adiabatic isentropic compressions, or expansions, the optical 'gas' heats up or cools down, respectively. This is achieved by altering the internal energy U through globally changing the coupling strengths between the individual elements comprising a multimode nonlinear optical array. **b**, Successive compression and expansion (C/E) cycles are used to achieve refrigeration in a particular subsystem (cold subsystem). In this configuration, three array subsystems are involved that are cyclically brought in thermal contact as shown in Fig. 3, allowing energy transfer ΔU from hot to cold. The array in the middle acts as a refrigerant by undergoing a sequence of successive expansions and compressions. After expanding, its temperature drops below that of the cold optical array subsystem on the right, thus extracting energy from it and hence further cooling it down during contact. This excess energy is then passed on to the hotter subsystem on the left (during contact) after the array in the middle goes through an isentropic compression—during which its temperature exceeds that of the hot array. **c**, The temperature–entropy Carnot cycle corresponding to these four events. Dashed lines indicate the ideal Carnot cycle for this configuration. The numbered green dots represent sequential end points in the Carnot cycle in the order they occurred.

is isentropic, given that $S = \sum \ln |c_i|^2$. From here, one can formally prove that the following laws hold during isentropic compressions or expansions:

$$\frac{\mu}{T} = \text{constant} \tag{5a}$$

$$\frac{U}{T} = \text{constant} \tag{5b}$$

Interestingly, equation (5a) is fully analogous to that expected from an ideal monoatomic gas undergoing isentropic transformations. Meanwhile, equation (5b) indicates that as the energy goes up or down, so does the temperature in an analogous fashion. In other words, during compression the temperature increases while it decreases during expansion. These effects will now be utilized to implement an all-optical 'refrigeration' cycle whereby energy is extracted from a cold multimode lattice (right) and transferred to a hotter system (left)—see Fig. 5b. The optical multimode system in between acts as a 'refrigerant' whereby the temperature first falls after expansion, thus inducing energy extraction from the cold object on contact. After it gets detached from the right arrangement, it undergoes a compression (thus increasing its temperature) to deliver the extracted energy to the hot system on the left. In essence, the three subsystems in Fig. 5b are allowed to exchange energy via contact layers but not optical power, as in the canonical-like ensemble arrangement presented in Fig. 3. The Carnot-like cycle corresponding to this process is depicted in Fig. 5c. Such optical refrigeration schemes can be judiciously deployed in high-power multimode fibre sources where the aim is to eventually drive the optical power into the lower group of modes via cooling—thus producing a high-quality output beam that is nearly free of speckle.

Conclusions

We have developed a thermodynamical formalism that can be utilized in a versatile manner to not only explain but also predict the complex response of nonlinear heavily multimoded optical systems. The thermodynamic laws derived here are universal. This theoretical framework can be used to devise novel all-optical techniques through which the equilibrium modal distribution (temperature, chemical potential) in a particular subsystem can be controlled at will. The principles derived here are not exclusive to optical structures such as waveguides or cavities but can also be employed in a straightforward manner in numerous other bosonic arrangements.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41566-019-0501-8.

Received: 29 November 2018; Accepted: 30 June 2019;
Published online: 19 August 2019

References

1. Richardson, D., Fini, J. & Nelson, L. Space-division multiplexing in optical fibres. *Nat. Photon.* **7**, 354–362 (2013).
2. Li, G., Bai, N., Zhao, N. & Xia, C. Space-division multiplexing: the next frontier in optical communication. *Adv. Opt. Photon.* **6**, 413–487 (2014).
3. Ryf, R. et al. Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 × 6 MIMO processing. *J. Lightwave Technol.* **30**, 521–531 (2012).
4. Fan, S. & Kahn, J. M. Principal modes in multimode waveguides. *Opt. Lett.* **30**, 135–137 (2005).
5. Xiong, W. et al. Complete polarization control in multimode fibers with polarization and mode coupling. *Light Sci. Appl.* **7**, 54 (2018).
6. Ambichl, P. et al. Super- and anti-principal-modes in multimode waveguides. *Phys. Rev. X* **7**, 041053 (2017).

7. Poletti, F. & Horak, P. Description of ultrashort pulse propagation in multimode optical fibers. *J. Opt. Soc. Am. B* **25**, 1645–1654 (2008).
8. Mafi, A. Pulse propagation in a short nonlinear graded-index multimode optical fiber. *J. Lightwave Technol.* **30**, 2803–2811 (2012).
9. Pourbeyram, H., Agrawal, G. P. & Mafi, A. Stimulated Raman scattering cascade spanning the wavelength range of 523 to 1750 nm using a graded-index multimode optical fiber. *Appl. Phys. Lett.* **102**, 201107 (2013).
10. Wright, L. G., Christodoulides, D. N. & Wise, F. W. Spatiotemporal mode-locking in multimode fiber lasers. *Science* **358**, 94–97 (2017).
11. Longhi, S. Modulational instability and space time dynamics in nonlinear parabolic-index optical fibers. *Opt. Lett.* **28**, 2363–2365 (2003).
12. Lopez-Galmiche, G. et al. Visible supercontinuum generation in a graded index multimode fiber pumped at 1064 nm. *Opt. Lett.* **41**, 2553–2556 (2016).
13. Krupa, K. et al. Observation of geometric parametric instability induced by the periodic spatial self-imaging of multimode waves. *Phys. Rev. Lett.* **116**, 183901 (2016).
14. Lopez-Aviles, H. E. et al. A systematic analysis of parametric instabilities in nonlinear parabolic multimode fibers. *APL Photon.* **4**, 022803 (2019).
15. Krupa, K. et al. Spatiotemporal characterization of supercontinuum extending from the visible to the mid-infrared in a multimode graded-index optical fiber. *Opt. Lett.* **41**, 5785–5788 (2016).
16. Renninger, W. H. & Wise, F. W. Optical solitons in graded-index multimode fibres. *Nat. Commun.* **4**, 1719 (2013).
17. Wright, L. G., Christodoulides, D. N. & Wise, F. W. Controllable spatiotemporal nonlinear effects in multimode fibres. *Nat. Photon.* **9**, 306–310 (2015).
18. Liu, Z., Wright, L. G., Christodoulides, D. N. & Wise, F. W. Kerr self-cleaning of femtosecond-pulsed beams in graded-index multimode fiber. *Opt. Lett.* **41**, 3675–3678 (2016).
19. Krupa, K. et al. Spatial beam self-cleaning in multimode fibres. *Nat. Photon.* **11**, 237–241 (2017).
20. Kolesik, M. & Moloney, J. V. Nonlinear optical pulse propagation simulation: from Maxwell's to unidirectional equations. *Phys. Rev. E* **70**, 036604 (2004).
21. Agrawal, G. P. *Nonlinear Fiber Optics* 5th edn (Springer, 2000).
22. Pathria, R. K. & Beale, P. D. *Statistical Mechanics* 3rd edn (Elsevier Science, 2011).
23. Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957).
24. Picozzi, A. et al. Optical wave turbulence: towards a unified nonequilibrium thermodynamic formulation of statistical nonlinear optics. *Phys. Rep.* **542**, 1–132 (2014).
25. Chiocchetta, A., Larré, P. É. & Carusotto, I. Thermalization and Bose–Einstein condensation of quantum light in bulk nonlinear media. *Europhys. Lett.* **115**, 24002 (2016).
26. Stone, J. M., Yu, F. & Knight, J. C. Highly birefringent 98-core fiber. *Opt. Lett.* **39**, 4568–4570 (2014).
27. Christodoulides, D. N., Lederer, F. & Silberberg, Y. Discretizing light behaviour in linear and nonlinear waveguide lattices. *Nature* **424**, 817–823 (2003).
28. Yariv, A., Xu, Y., Lee, R. K. & Scherer, A. Coupled-resonator optical waveguide: a proposal and analysis. *Opt. Lett.* **24**, 711–713 (1999).
29. Vahala, K. J. Optical microcavities. *Nature* **424**, 839–846 (2003).
30. Rigol, M., Dunjko, V. & Olshanii, M. Thermalization and its mechanism for generic isolated quantum systems. *Nature* **452**, 854–858 (2008).
31. Zakharov, V. E., L'vov, V. S. & Falkovich, G. *Kolmogorov Spectra of Turbulence I: Wave Turbulence* (Springer Science & Business Media, 2012).
32. Dyachenko, S., Newell, A. C., Pushkarev, A. & Zakharov, V. E. Optical turbulence: weak turbulence, condensates and collapsing filaments in the nonlinear Schrödinger equation. *Physica D* **57**, 96–160 (1992).
33. Picozzi, A. Towards a nonequilibrium thermodynamic description of incoherent nonlinear optics. *Opt. Express* **15**, 9063–9083 (2007).
34. Lagrange, S., Jauslin, H. R. & Picozzi, A. Thermalization of the dispersive three-wave interaction. *Europhys. Lett.* **79**, 64001 (2007).
35. Picozzi, A. & Rica, S. Condensation of classical optical waves beyond the cubic nonlinear Schrödinger equation. *Opt. Commun.* **285**, 5440–5448 (2012).
36. Sun, C. et al. Observation of the kinetic condensation of classical waves. *Nat. Phys.* **8**, 470–474 (2012).
37. Davis, M., Morgan, S. & Burnett, K. Simulations of Bose fields at finite temperature. *Phys. Rev. Lett.* **87**, 160402 (2001).
38. Davis, M., Morgan, S. & Burnett, K. Simulations of thermal Bose fields in the classical limit. *Phys. Rev. A* **66**, 053618 (2002).
39. Klaers, J., Schmitt, J., Vewinger, F. & Weitz, M. Bose–Einstein condensation of photons in an optical microcavity. *Nature* **468**, 545–548 (2010).
40. Aschieri, P., Garnier, J., Michel, C., Doya, V. & Picozzi, A. Condensation and thermalization of classsical optical waves in a waveguide. *Phys. Rev. A* **83**, 033838 (2011).
41. Rasmussen, K. Ø., Cretegny, T., Kevrekidis, P. G. & Grønbech-Jensen, N. Statistical mechanics of a discrete nonlinear system. *Phys. Rev. Lett.* **84**, 3740–3743 (2000).
42. Silberberg, Y., Lahini, Y., Bromberg, Y., Small, E. & Morandotti, R. Universal correlations in a nonlinear periodic 1D system. *Phys. Rev. Lett.* **102**, 233904 (2009).
43. Kottos, T. & Shapiro, B. Thermalization of strongly disordered nonlinear chains. *Phys. Rev. E* **83**, 062103 (2011).
44. Rumpf, B. Transition behavior of the discrete nonlinear Schrödinger equation. *Phys. Rev. E* **77**, 036606 (2008).
45. Derevyanko, S. A. Thermalized polarization dynamics of a discrete optical-waveguide system with four-wave mixing. *Phys. Rev. A* **88**, 033851 (2013).
46. Fusaro, A., Garnier, J., Krupa, K., Millot, G. & Picozzi, A. Dramatic acceleration of wave condensation mediated by disorder in multimode fibers. *Phys. Rev. Lett.* **122**, 123902 (2019).
47. Purcell, E. M. & Pound, R. V. A nuclear spin system at negative temperature. *Phys. Rev.* **81**, 279–280 (1951).
48. Ramsey, N. F. Thermodynamics and statistical mechanics at negative absolute temperatures. *Phys. Rev.* **103**, 20–28 (1956).
49. Braun, S. et al. Negative absolute temperature for motional degrees of freedom. *Science* **339**, 52–55 (2013).

Acknowledgements

We thank G. Tan and Y. Huang for computing support. We acknowledge financial support by the Office of Naval Research (ONR) (MURI N00014-17-1-2588 and N00014-18-1-2347), the National Science Foundation (NSF) (EECS-1711230) and the Qatar National Research Fund (NPRP9-020-1-006).

Author contributions

All authors contributed extensively to the work presented in this paper. F.O.W., A.U.H. and D.N.C. developed the theoretical formalism. F.O.W. performed the calculation. F.O.W., A.U.H. and D.N.C. wrote the paper. D.N.C. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41566-019-0501-8.

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

Correspondence and requests for materials should be addressed to D.N.C.

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

Methods

Mode occupancy distribution. We list the number of ways (W) in which one can distribute N indistinguishable packets of power (or 'photons' at a specific wavelength) in M distinct optical modes having energy levels ε_i (propagation constants), each associated with a degeneracy g_i (g_i could also represent the number of clustered sublevels[22]—see Supplementary Fig. 4). To do so, we assign to each ε_i level group or cell, n_i discrete power packets that are meant to be distributed over g_i distinguishable compartments. From here one finds that $W_i=(n_i+g_i-1)!/[n_i!(g_i-1)!]$. Hence, the total number of ways W can be obtained from

$$W(n_1,n_2,n_3,\ldots)=\prod_i^M \frac{(n_i+g_i-1)!}{n_i!\,(g_i-1)!}$$

The entropy of the system is given by $S_N=\ln W$. Given that in actual settings $n_i>>1$, and that $\ln n!=n\ln n-n$ (Stirling approximation), we find

$$S_N=\sum_i^M (n_i+g_i-1)\ln(n_i+g_i-1)-(n_i+g_i-1)-n_i\ln n_i+n_i-\ln(g_i-1)!$$

The entropy must be then maximized under the following two constraints:

$$N=\sum_i^M n_i$$

$$E=-\sum_i^M n_i\varepsilon_i$$

where E represents the total 'internal energy' in the system and N the total number of power packets or 'particles'. Extremization by means of two Lagrange multipliers α,β leads to a Bose–Einstein distribution

$$\frac{\partial}{\partial n_i}\left(S_N+\alpha\sum_i^M n_i+\beta\sum_i^M \varepsilon_i n_i\right)=0$$

$$\frac{n_i}{g_i}=\frac{1}{e^{-\alpha-\beta\varepsilon_i}-1}$$

Keeping in mind that even in heavily multimoded nonlinear optical systems, the number of power (or energy) packets is much larger than the number of available modes in a cluster, $n_i>>g_i$, we conclude that

$$e^{-\alpha-\beta\varepsilon_i}\to 1$$

$$e^{-\alpha-\beta\varepsilon_i}-1\simeq-\alpha-\beta\varepsilon_i$$

In doing so, the Bose–Einstein distribution readily reduces to an RJ distribution, that is

$$\frac{n_i}{g_i}=-\frac{1}{\alpha+\beta\varepsilon_i} \tag{6}$$

We would like to mention that in an optical waveguide multimode system, the eigenenergies are positioned according to $\varepsilon_1\le\varepsilon_2\le\varepsilon_3\ldots\le\varepsilon_M$, where ε_M represents the energy of the lowest-order mode (ground state) while ε_1 that of the highest. The reason for this particular choice of sequence is that we measure these propagation constants from the cladding region, in which case the ground state has the highest eigenvalue. This situation is reversed in nonlinear multimode cavity systems.

Optical entropy. The optical entropy can be directly obtained from the equation:

$$S_N=\sum_i^M (n_i+g_i-1)\ln(n_i+g_i-1)-(n_i+g_i-1)-n_i\ln n_i+n_i-\ln(g_i-1)!$$

derived above. Given that in a nonlinear multimoded optical system, $n_i/g_i>>1$, that is, each mode is very highly populated, one can then use a Taylor expansion and omit small terms:

$$\begin{aligned}S_N &= \sum_i^M (n_i+g_i-1)\ln\left[n_i\left(1+\frac{g_i-1}{n_i}\right)\right]-g_i+1-n_i\ln n_i-\ln(g_i-1)!\\ &= \sum_i^M (n_i+g_i-1)\ln n_i+(n_i+g_i-1)\left(\frac{g_i-1}{n_i}\right)-g_i+1-n_i\ln n_i\\ &\quad -\ln(g_i-1)!\end{aligned}$$

From here, since $n_i>>g_i$, one quickly finds that $S_N\cong\sum_i^M g_i\ln n_i$, or more conveniently the optical entropy can now be written as:

$$S_N=\sum_i^M \ln n_i$$

where in obtaining our last result we have set $g_i=1$ since the degeneracies were just used for clustering purposes. Moreover, M now represents the total number of modes. Interestingly, this same entropy also applies in other settings like, for example, image restoration[50]. In an optical system, the power in each mode $|c_i|^2$ is proportional to the number of infinitesimal power packets n_i in each ε_i level. The proportionality factor in this case can be selected in such a way that $n_i=n_c|c_i|^2$, where n_c represents the number of power packets per unit of normalized power/energy. Hence, the entropy can now be viewed as the sum of two components

$$S_N=\sum_{i=1}^M \ln(n_c)+\sum_{i=1}^M \ln|c_i|^2$$

Since n_c is a constant, the first part in the above equation represents a floor (a reference point) in the entropy while the second one denotes a more relevant entropic component that responds to nonlinear mode mixing. Hence from this point on, we write

$$S=\sum_{i=1}^M \ln\,|c_i|^2 \tag{7}$$

By adopting the more conventional definitions for temperature and chemical potential, that is, $\alpha=\mu(Tn_c)^{-1}$ and $\beta=(Tn_c)^{-1}$, equation (6) reduces to

$$|c_i|^2=-\frac{T}{\varepsilon_i+\mu} \tag{8}$$

Finally, by again using $n_i=n_c|c_i|^2$, the optical observables associated with the total power $\mathcal{P}$ and internal energy U (Supplementary Information) can be expressed as follows

$$U=-\sum_i^M \varepsilon_i\,|c_i|^2 \tag{9}$$

$$\mathcal{P}=\sum_i^M |c_i|^2 \tag{10}$$

Derivation of the first equation of state. After manipulating the expressions for the optical internal energy and power, we obtain

$$\frac{U}{T}-\frac{\mu}{T}\mathcal{P}=\sum_i^M\left[\left(\frac{\varepsilon_i}{\varepsilon_i+\mu}\right)-\left(\frac{-\mu}{\varepsilon_i+\mu}\right)\right]=\sum_i^M 1=M$$

Hence

$$U-\mu\mathcal{P}=MT \tag{11}$$

Predicting T and μ of a microcanonical ensemble at equilibrium. The temperature T and chemical potential μ can be directly determined from either equation (9) or equation (10) by invoking the first equation of state equation (11) to eliminate one variable among T and μ. For example, using the equilibrium modal distribution in equation (8) and given that from equation (11), $\mu=\mathcal{P}^{-1}(U-MT)$, we find that

$$\mathcal{P}=\sum_i^M -\frac{T}{\varepsilon_i+\mu}=\sum_i^M -\frac{T}{\varepsilon_i+\mathcal{P}^{-1}(U-MT)}$$

For a given input power $\mathcal{P}$ and internal energy U, and once the eigenspectrum ε_i for this M-mode system is known, the only unknown variable T can be uniquely determined by solving the equation above. Note that the only acceptable solution for T is the one that keeps each occupancy $|c_i|^2$ positive (equation (8)). From here, one can then obtain the chemical potential μ through equation (11).

Extensivity of the optical entropy. As previously indicated, at equilibrium, the entropy of a given multimode optical system is written as

$$S(U,M,\mathcal{P})=\sum_i^M \ln\,|c_i|^2=\sum_i^M \ln\left(-\frac{T}{\varepsilon_i+\mu}\right)$$

Let us now, for example, double the optical arrangement $M \to 2M$ in such a way that the structure of the system, or the profile of the density of states, remains invariant (see Supplementary Fig. 5). At the same time, we double the input optical power and internal energy: $U \to 2U, \mathcal{P} \to 2\mathcal{P}$. From equation (11), we find that T and μ still remain the same since $2U-2\mu\mathcal{P}=2MT$. The same conclusion can be reached by directly solving equations (9) and (10).

Since each energy level ε_i now splits into two closely spaced energy levels ε_{i1} and ε_{i2}, where $\varepsilon_{i1} \approx \varepsilon_{i2} \approx \varepsilon_i$, the entropy now becomes

$$S(2U,2M,2\mathcal{P}) = \sum_i^M \ln\left(-\frac{T}{\mu+\varepsilon_{i1}}\right) + \ln\left(-\frac{T}{\mu+\varepsilon_{i2}}\right)$$

Hence

$$S(2U,2M,2\mathcal{P}) = 2\sum_i^M \ln\left(-\frac{T}{\mu+\varepsilon_i}\right) = 2S(U,M,\mathcal{P})$$

In more general terms, $S(\lambda U,\lambda M,\lambda\mathcal{P}) = \lambda S(U,M,\mathcal{P})$

Thermodynamic driving forces of temperature, chemical potential and pressure. We next show that the temperature can be self-consistently obtained from the fundamental equation of thermodynamics via:

$$\frac{\partial S}{\partial U} = \frac{1}{T} \tag{12}$$

Since $\mu = \mathcal{P}^{-1}(U-MT)$, from equations (7) and (8) we obtain

$$S = \sum_i^M \ln\left(\frac{-T}{\varepsilon_i+\mu}\right) = \sum_i^M \ln\left(\frac{T}{-\varepsilon_i-\frac{1}{\mathcal{P}}(U-MT)}\right)$$

$$\left.\frac{\partial S(U,M,\mathcal{P};T)}{\partial U}\right|_{P,M} = \frac{\partial S}{\partial U} + \frac{\partial S}{\partial T}\frac{\partial T}{\partial U}$$

It can be directly seen that $\frac{\partial S}{\partial T}=0$, since

$$\frac{\partial S}{\partial T} = \sum_i^M \frac{1}{T} - \frac{M}{\mathcal{P}T}\sum_i^M \frac{T}{-\varepsilon_i-\mu} = \frac{M}{T} - \frac{M}{T} = 0$$

This leaves us with

$$\left.\frac{\partial S(U,M,\mathcal{P};T)}{\partial U}\right|_{P,M} = \frac{\partial S}{\partial U} = \sum_i^M \frac{\frac{1}{\mathcal{P}}}{-\varepsilon_i-\frac{1}{\mathcal{P}}(U-MT)} = \frac{1}{T}\frac{1}{\mathcal{P}}\sum_i^M |c_i|^2$$

$$\Rightarrow \left.\frac{\partial S}{\partial U}\right|_{\mathcal{P},M} = \frac{1}{T}$$

We next prove that the chemical potential can be formally defined according to

$$\frac{\partial S}{\partial \mathcal{P}} = -\frac{\mu}{T} \tag{13}$$

in a manner consistent with the first equation of state equation (11). If we instead write

$$T = \frac{U-\mu\mathcal{P}}{M}$$

and substitute it into equations (7) and (8), one obtains:

$$S = \sum_i^M \ln\left[\frac{U-\mu\mathcal{P}}{M(-\varepsilon_i-\mu)}\right]$$

$$\left.\frac{\partial S(U,M,\mathcal{P};\mu)}{\partial \mathcal{P}}\right|_{U,M} = \frac{\partial S}{\partial \mathcal{P}} + \frac{\partial S}{\partial \mu}\frac{\partial \mu}{\partial \mathcal{P}}$$

Similar to the last case, it can be directly shown that $\frac{\partial S}{\partial \mu}=0$, since

$$\frac{\partial S}{\partial \mu} = \sum_i^M \frac{-\mathcal{P}}{U-\mu\mathcal{P}} - \sum_i^M \frac{-M}{M(-\varepsilon_i-\mu)}$$

$$\frac{\partial S}{\partial \mu} = \frac{-M\mathcal{P}}{U-\mu\mathcal{P}} + \frac{1}{T}\sum_i^M \frac{T}{-\varepsilon_i-\mu}$$

$$\frac{\partial S}{\partial \mu} = -\frac{\mathcal{P}}{T} + \frac{\mathcal{P}}{T} = 0$$

We are finally left with

$$\left.\frac{\partial S(U,M,\mathcal{P},\mu)}{\partial P}\right|_{U,M} = \frac{\partial S}{\partial \mathcal{P}} = \sum_i^M \frac{-\mu}{U-\mu\mathcal{P}} = \frac{-\mu M}{U-\mu\mathcal{P}}$$

$$\Rightarrow \left.\frac{\partial S}{\partial \mathcal{P}}\right|_{U,M} = -\frac{\mu}{T}$$

We next define the third intensive variable $\hat{p}$ ('optical pressure') through

$$\frac{\partial S}{\partial M} = \frac{\hat{p}}{T} \tag{14}$$

In essence, the variable M plays the role of volume in conventional thermodynamics. As indicated in the text, any expansion in the number of modes M should be carried out in a way that leaves the structure of the system the same, that is, the profile of the density of states remains invariant (see Supplementary Fig. 5). To obtain the term $\partial S/\partial M$, we first define a density of states $D(\varepsilon) = \mathrm{d}M/\mathrm{d}\varepsilon$, providing the number of modes $D(\varepsilon)\mathrm{d}\varepsilon$ per unit interval $\mathrm{d}\varepsilon$. Hence, the total number of modes in the original system is given by $\int D(\varepsilon)\mathrm{d}\varepsilon = M_0$. We now change in a self-similar manner the number of modes in this system in a way that preserves the profile of the density of states. In this case, the new density of states $\widetilde{D}(\varepsilon)$ can be obtained simply by: $\widetilde{D}(\varepsilon) = VD(\varepsilon)$, where V is a scale factor indicating the fractional change in 'volume' or the number of modes. From equation (8), we can rewrite the expressions for $U, \mathcal{P}$ and S as follows

$$U = \int \frac{\varepsilon T}{\varepsilon+\mu} VD(\varepsilon)\mathrm{d}\varepsilon$$

$$\mathcal{P} = -\int \frac{T}{\varepsilon+\mu} VD(\varepsilon)\mathrm{d}\varepsilon$$

$$S = \int \ln\left(-\frac{T}{\varepsilon+\mu}\right) VD(\varepsilon)\mathrm{d}\varepsilon$$

The first equation of state equation (11) leads to:

$$U-\mu\mathcal{P} = T\int VD(\varepsilon)\mathrm{d}\varepsilon = VM_0T = MT$$

If we rewrite this equation as $\mu = (U-VM_0T)/\mathcal{P}$ and substitute into the expression for entropy, we get:

$$S = \int \ln\left[\frac{T}{-\varepsilon-\frac{1}{\mathcal{P}}(U-VM_0T)}\right] VD(\varepsilon)\mathrm{d}\varepsilon$$

Upon differentiating with respect to V one finds:

$$\left.\frac{\partial S(U,V,\mathcal{P};T)}{\partial V}\right|_{\mathcal{P},U} = \frac{\partial S}{\partial V} + \frac{\partial S}{\partial T}\frac{\partial T}{\partial V}$$

since $\partial S/\partial T = 0$ (as in the case of equation (12). From here one finds that:

$$\left.\frac{\partial S}{\partial V}\right|_{\mathcal{P},U} = \frac{\hat{p}}{T} = \frac{S}{V} - M_0$$

By setting the arbitrary reference point M_0 to be 1, hence allowing $V \to M$, we then obtain

$$\left.\frac{\partial S}{\partial M}\right|_{\mathcal{P},U} = \frac{\hat{p}}{T} = \frac{S}{M} - 1 \tag{15}$$

Euler equation and second equation of state. The extensive nature of entropy enables us to write:

$$S(\lambda U,\lambda M,\lambda\mathcal{P}) = \lambda S(U,M,\mathcal{P})$$

Differentiation on both sides of this last relation with respect to λ yields an expression for S in terms of all the variables involved in the multimoded system, better known as the Euler equation:

$$\frac{\partial S(\lambda U,\ldots)}{\partial(\lambda U)}U + \frac{\partial S(\lambda U,\ldots)}{\partial(\lambda M)}M + \frac{\partial S(\lambda U,\ldots)}{\partial(\lambda\mathcal{P})}\mathcal{P} = S(U,\mathcal{P},M)$$

$$\frac{U}{T}+\frac{\hat{p}}{T}M-\frac{\mu}{T}\mathcal{P}=S \tag{16}$$

where in deriving equation (16), we used equations (12)–(14). The Euler equation can also be consistently obtained from equation (11) and equation (15). The extensivity of entropy with respect to U, M and $\mathcal{P}$ is also apparent from equation (16).

Direction of energy flow between two multimoded systems in thermal contact. When two systems, both having initially reached thermal equilibrium independently, are allowed to exchange energy but not power (in a canonical-like ensemble), the second law of thermodynamics demands that the total entropy $S_T=S_1+S_2$, should never decrease, that is, $dS_T=dS_1+dS_2\geq 0$. Since, the power and volume associated with each subsystem is constant, the change in entropy depends only on energy exchange, $dS=\frac{1}{T}dU$. Hence

$$dS_T=\frac{1}{T_1}dU_1+\frac{1}{T_2}dU_2\geq 0$$

Since the total internal energy is conserved ($U_T=U_1+U_2$ and $dU_T=0$), $dU_1=-dU_2$. Therefore

$$\left(\frac{1}{T_1}-\frac{1}{T_2}\right)dU_1\geq 0$$

The last expression implies that the energy flows towards the first subsystem ($dU_1>0$) only when $(T_1^{-1}-T_2^{-1})>0$. In other words, if the temperatures of the two subsystems have the same sign, energy flows from the subsystem with a higher temperature to the one with the lower temperature; if the temperatures have a different sign, energy flows from the subsystem with a negative temperature to the subsystem with a positive temperature. The latter unconventional behaviour is depicted in Fig. 2a in the main text.

Direction of power and energy flow between two multimoded systems in diffusive contact. When two systems, which have initially reached thermal equilibrium independently, are allowed to exchange both energy and power (as in a grand canonical-like ensemble), the second law of thermodynamics now demands that

$$dS_T=\frac{1}{T_1}dU_1+\frac{1}{T_2}dU_2-\frac{\mu_1}{T_1}d\mathcal{P}_1-\frac{\mu_2}{T_2}d\mathcal{P}_2\geq 0$$

Because of the two conservation laws: $dU_1=-dU_2$, $d\mathcal{P}_1=-d\mathcal{P}_2$. Thus

$$\left(\frac{1}{T_1}-\frac{1}{T_2}\right)dU_1+\left(\frac{\mu_2}{T_2}-\frac{\mu_1}{T_1}\right)d\mathcal{P}_1\geq 0$$

As before, energy flows from a hotter system to a colder system. If the temperature is common to both subsystems ($T_1=T_2$), power will always flow from the subsystem with a higher chemical potential to that with a lower chemical potential, if the temperature is positive, and vice versa if it is negative, always in such a way that $(\mu_2-\mu_1)T^{-1}d\mathcal{P}_1>0$. Once equilibrium is reached and the chemical potentials reach the same value, exchange of power ceases.

Graphical way of predicting the final temperature of canonical-like ensembles consisting of several subsystems. Given a certain level of optical power $\mathcal{P}$, the $U(T)$ relation of any subsystem (associated with a set of propagation eigenvalues ε_i) can be obtained individually, as done in the microcanonical ensemble described earlier in the Methods. Consider a canonical system consisting of two subsystems A and B, each having its own $U(T)$ relation, $U_A(T_A)$ and $U_B(T_B)$. Since in a canonical ensemble, at equilibrium, all of the subsystems reach the same temperature through energy exchange, while the total energy of the combined system $U_0=U_A+U_B$ is conserved, we have:

$$U_A(T)=U_0-U_B(T)$$

From here, the common equilibrium temperature T of the subsystems can be graphically obtained through finding the intersection of the curve $U_A(T)$ with $(U_0-U_B(T))$. Supplementary Fig. 6 illustrates this graphical technique, as used to predict the final temperature associated with simulations in Figs. 2 and 3.

Isentropic invariants. During an adiabatic process, the mode occupancies $|c_i|^2$ remain the same, making it isentropic since $S=\sum \ln |c_i|^2$. Meanwhile, during this same process, the eigenenergies ε_i are expected to change through a common proportionality factor and so does the internal energy U. Consequently, the temperature and chemical potential are expected to adiabatically change as well. If initially, the system is at equilibrium, the RJ distribution requires $|c_i|^2=-T_0/(\varepsilon_i+\mu_0)$, where T_0 and μ_0 represent the initial temperature and chemical potential, respectively. We now assume that during the isentropic process, $\varepsilon_i\rightarrow\lambda\varepsilon_i$, $T_0\rightarrow T_1$ and $\mu_0\rightarrow\mu_1$, in a way that the RJ distribution is respected, that is, $|c_i|^2=-T_1/(\lambda\varepsilon_i+\mu_1)$. Since the mode occupancies $|c_i|^2$ are invariant during the process, we find that

$$\frac{\varepsilon_i}{T_0}+\frac{\mu_0}{T_0}=\frac{\lambda\varepsilon_i}{T_1}+\frac{\mu_1}{T_1}, \text{ for all possible } i$$

If we write $\mu/T=\alpha$, then for two different arbitrary eigenstates, $i=k$, m, we obtain

$$\varepsilon_k\left(\frac{1}{T_0}-\frac{\lambda}{T_1}\right)=\Delta\alpha$$

$$\varepsilon_m\left(\frac{1}{T_0}-\frac{\lambda}{T_1}\right)=\Delta\alpha$$

where $\Delta\alpha=\alpha_1-\alpha_0$. If we assume that $\Delta\alpha\neq 0$, the above equations imply that for any k and m, $\varepsilon_k=\varepsilon_m$, which is by itself a contradiction. The only way to avoid this contradiction is to set $\Delta\alpha=0$, which shows that during an isentropic process

$\mu/T=$ constant

For $\Delta\alpha=0$, the above equations suggest that $T_1=\lambda T_0$ and $\mu_1=\lambda\mu_0$. In addition, since

$$U_1=-\sum_i^M \lambda\varepsilon_i\,|c_i|^2=\lambda U_0$$

we find that $U/T=$ constant

Terminology. Throughout this paper, we loosely use the term 'canonical-like ensemble' to describe two different subsystems capable of exchanging internal energy U but no optical power $\mathcal{P}$ through a diathermic wall. This does not necessarily mean that the subsystem is placed in contact with a thermal bath at constant temperature. Meanwhile, we use the term 'grand canonical-like ensemble' to refer to the case where the two subsystems can also in addition exchange optical power via a diathermic permeable wall.

Simulations. Due to the chaotic nature of nonlinear interactions, the modal occupancies are fluctuating. Given that thermodynamics deals with the averaged values for these occupancies, it is important to describe how these averages are obtained in carrying out numerical simulations. In general, the averaging process can be conducted over ensembles or over the propagation distance (or time). Both approaches are expected to be in agreement as long as the system is ergodic. A combination of these two schemes can also be employed to reduce the computation time. In all the simulations carried out in this paper, we used a window size of ~100 units of normalized propagation distance (or time) to track the evolution of S, and equilibrium is determined once the entropy is maximized. The averaged values of modal occupancies at equilibrium are then determined by considering the last 20% of the data. In contrast, one can also carry out ensemble averaging by conducting a number of simulations with random initial phases while exciting the same modal group (that also happens to preserve the values of the extensive variables U and $\mathcal{P}$). We would like to mention that at equilibrium, the intensive parameters T and μ are universally determined only by the system's three invariants $\mathcal{P}$, M and U (or E when considering cavities), regardless of the specific modal groups initially excited.

In Fig. 1a,b, we used a chaotic Pullen–Edmonds potential with a refractive index profile $n(x,y)=n_0[1-\Delta(X^2+Y^2+X^2Y^2)/R_0^2]$ truncated at the cladding index $n(X,Y)=n_{clad}=1.446$, where the core radius is $R_0=25\,\mu$m, the maximum core index is $n_0=1.46$ and the relative index difference is $\Delta=0.01$. The simulation is conducted by solving 20 ensembles associated with the (two + one)-dimensional scalar nonlinear Schrödinger equation (see Supplementary Information) with a Kerr nonlinear coefficient $n_2=3.2\times10^{-20}\,m^2\,W^{-1}$. In contrast, the evolution of the optical field in the waveguide lattices discussed throughout the paper is governed by a discrete nonlinear Schrödinger equation (see Supplementary Information). In Fig. 1c,d, the Lieb lattice has normalized coupling coefficients $\kappa_x=2$, $\kappa_y=1$ with no detuning between sites and no cross-coupling along the diagonal. In Fig. 1e,f, the coupled resonator optical waveguide system is a $10\times10\times10$ structure with a cubic unit cell. All couplings along the edges are $\kappa_1=1$ and along the diagonals of each face are $\kappa_2=\sqrt{0.03}$. In this temporal case, the equilibrium modal occupancies follow instead the slightly modified RJ distribution $|c_i|^2=T/(\varepsilon_i-\mu)$ to be consistent with the thermodynamic driving forces mentioned earlier. In Fig. 2, the coupling between all adjacent sites is $\kappa_1=1$ and on the diagonal it is $\kappa_2=\sqrt{0.03}$. When the two lattices are brought together, the combined system retains the same geometric structure. In this case, the normalized nonlinear coefficient was $\gamma=0.1$ and the normalized propagation distance is scaled in units of 10^4 coupling lengths. The data presented correspond to an ensemble average over 30 samples. In Fig. 3, the values of coupling strengths indicated in the inset are $\kappa_1=0.5\sqrt{4/5}$, $\kappa_2=1$, $\kappa_3=\sqrt{4/5}$, $\kappa_4=\sqrt{0.03}$ and five ensembles were used in the averaging process. Here, the normalized propagation distance is scaled in units

of 10^5 coupling lengths. The coupling values used in Fig. 4 are $\kappa_1 = 1.5$ and $\kappa_2 = 1$. The plotted data correspond to an ensemble average over 20 samples. The distance axes are scaled in units of 10^4 coupling lengths. In all of our simulations associated with discrete systems, the normalized power per mode $(\mathcal{P}/M)$ is kept ~0.1 so as to allow the systems to evolve in a quasi-linear fashion.

Finally, the lattice subsystems in Fig. 5 are shown in one dimension to illustrate the concept, while in actual simulations we used a two-dimensional 10×20 square/rectangular structure for each lattice, with no diagonal-couplings. The hot and cold arrays have $\kappa_x = \kappa_y = 1$, while the coupling strengths in the compression and expansion array are $(\kappa_x, \kappa_y) = \lambda(1, \sqrt{5/4})$ where the coupling factor λ varies as $2 \rightarrow 0.25$ during expansion and vice versa during compression. The optical power in the left, middle and right arrays are $\mathcal{P} = 27$, 21 and 7.6, respectively. The data points 1–5 indicated by green dots were obtained when the array in the middle was detached from either of the two sublattices on the sides (see Supplementary Fig. 7). The dashed vertical lines in Fig. 5c represent the ideal isentropic processes. The upper and lower dashed curves were obtained by finding the T–S curve of the array in the middle when λ was set equal to 2 and 0.25, respectively.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

50. Kikuchi, R. & Soffer, B. H. Maximum entropy image restoration. I. The entropy expression. *J. Opt. Soc. Am.* **67**, 1656–1665 (1977).