Thermalization of orbital angular momentum beams in multimode optical fibers

E. V. Podivilov[1,2,†], F. Mangini[3,4,†,*], O. S. Sidelnikov[1], M. Ferraro[4], M. Gervaziev[1],

D. S. Kharenko[1,2], M. Zitelli[4], M. P. Fedoruk[1], S. A. Babin[1,2], and S. Wabnitz[1,4]

[1] *Novosibirsk State University, Novosibirsk 630090, Russia*
[2] *Institute of Automation and Electrometry SB RAS, 1 ac. Koptyug ave., Novosibirsk 630090, Russia*
[3] *Department of Information Engineering, University of Brescia, Via Branze 38, 25123 Brescia, Italy*
[4] *Department of Information Engineering, Electronics, and Telecommunications,*
Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy
Corresponding author: fabio.mangini@unibs.it and
† *These authors have contributed equally*
(Dated: December 28, 2021)

We present a general theory of thermalization of light in multimode optical fibers, including optical beams with nonzero orbital angular momentum or vortex beams. A generalized Rayleigh-Jeans distribution of asymptotic mode composition is obtained, based on the conservation of the angular momentum. We confirm our predictions by numerical simulations and experiments based on holographic mode decomposition of multimode beams. This establishes new constraints for the achievement of spatial beam self-cleaning, giving previously unforeseen insights into the underlying physical mechanisms.

Statistical physics has been traditionally and successfully employed to describe the average properties of a large ensemble of particles, whose interactions are governed by classical mechanics. This approach lies at the basis of thermodynamics, whose laws determine the macroscopic properties of matter, that evolve in a low-dimensional or reduced phase space. Subsequently, the thermodynamic approach has been extended to describe the statistical evolution of a large number of classical electromagnetic waves, analogously to bosonic systems, such as superconductors and superfluids [1–5].

Peculiar is the case of multimode optical fibers (MMFs), which are an excellent study case for classical wave condensation phenomena. Indeed, Bose-Einstein condensation of the fiber modes has been demonstrated in graded-index (GRIN) MMFs [6], and it can be theoretically described by a model based on a weak wave turbulence approach [7]. Whereas a general model of thermalization of light in multimode systems has been recently introduced, showing that the average number of photons in each mode of the fiber obeys a Rayleigh-Jeans (RJ) distribution [8]. Because of the role of high-order modes at the occurrence of thermal equilibrium in MMFs, thermalization of a multimode field is a more general situation than condensation [9].

On the other hand, experimental observations have revealed that, as the input power of a laser beam coupled into a graded-index (GRIN) MMF grows above a certain threshold, the intensity speckles generated by multimode interference may spontaneously reorganize into a bell-shaped beam, which approaches the fundamental mode of the fiber [10, 11]. This spatial self-organization effect is known as beam self-cleaning (BSC) [12], and it has similarities with wave condensation in hydrodynamic 2D turbulent systems [13]. Since its first demonstration, BSC has been extensively experimentally studied [13–20], in order to fully clarify its physical mechanism. All of the studies reported so far in the literature agree on the fact that modal four-wave-mixing (FWM) interactions are crucial for activating BSC. The condensation of energy in

the fundamental mode has been verified to obey the expected dependence on the initial degree of spatial correlation [14], or on the internal energy of the input beam with a fixed power value [6].

As a thermodynamic phenomenon, BSC can be seen as the tendency of the optical beam to experience an irreversible evolution towards a state of thermal equilibrium, which is established by conservation laws. Specifically, the total number of photons and the total momentum of motion must be simultaneously conserved [6, 9].

As a matter of fact, another quantity is conserved when a beam of light propagates in waveguide systems: its total orbital angular momentum (OAM). First introduced by Allen et al. in 1992 [21], interest in OAM beams has increased tremendously, thanks to their widespread potential for applications. These range from telecommunications [22] to quantum optics [23], holography [24] and optical tweezers [25]. To date, BSC has only been observed with laser beams that carry no OAM. In this work we extend current knowledge by describing, both theoretically and experimentally, the thermalization of OAM-carrying multimode beams in nonlinear optical fibers. We present a general theory of thermalization of light in a MMF, which takes into account the conservation of OAM. This permits to derive a generalized RJ distribution for the relative occupation of the fiber modes, which directly stems out of the conservation laws that rule FWM process in MMFs. Remarkably, our model shows that BSC can only be achieved by means of laser beams which do not carry any OAM. Theoretical predictions are then compared to numerical simulations, which turn out to be in excellent agreement. Finally, we carried out an experimental characterization of the thermal mode distribution at the fiber output, based on a holographic mode decomposition (MD) technique. The MD was performed by means of a phase-only spatial light modulator (SLM) [26]. In our implementation, OAM is imparted to an input Gaussian laser beam by means of properly adjusting its coupling condition into the fiber [27].

In our theoretical model, we do not use the conventional Laguerre-Gauss basis for decomposing the amplitude (A) of a light pulse in a GRIN MMFs. Conversely, we write A in terms of the real normalized *radial* OAM eigenfunctions $F_{\ell,m}$, so that

$$A(t,z,r,\phi) = \sum_{k,\ell,m} A_{\ell,m}(\omega_k,z) e^{i\omega_k t - i p_{\ell,m}(\omega_k)z + im\phi} F_{\ell,m}(r), \quad (1)$$

and $2\pi \int_0^\infty r dr (F_{\ell,m}(r))^2 = 1$. Here, $p_{\ell,m}(\omega_k)$ and $A_{\ell,m}(\omega_k,z)$ are the propagation constant and the slowly varying amplitude of a mode with radial index ℓ, azimuthal index m, and frequency $\omega_k = \omega_0 + 2\pi H k$, where ω_0 is the carrier frequency and H is the pulse repetition rate. In the Supplementary Notes we outline the transformation from the Laguerre-Gauss basis to the OAM basis.

Although BSC is a purely spatial effect, its high threshold powers have required the use of pulsed sources for its observation. Moreover, it has been shown that BSC is accompanied by significant temporal pulse reshaping [28], which may be associated with a transfer of disorder between spatial and temporal degrees of freedom [29]. For this reason, in our derivation, we keep explicit the dependence of the light amplitude on frequency ω_k.

Let us normalize $|A(t,z,r,\phi)|^2$ to the beam intensity, so that for each mode (ℓ,m) and frequency component ω_k we may introduce the average power $W_{k,\ell,m}(z) = |A_{\ell,m}(\omega_k,z)|^2$, the mode energy in a pulse $E_{k,\ell,m}(z) = W_{k,\ell,m}(z)/H$, the number of photons $N_{k,\ell,m}(z) = E_{k,\ell,m}(z)/\hbar\omega_k$, the longitudinal component of pulse momentum of motion $P_{k,\ell,m}(z) = p_{\ell,m}(\omega_k)N_{k,\ell,m}(z)$, and the OAM $M_{k,\ell,m}(z) = \hbar m N_{k,\ell,m}(z)$. There are four conservation laws in the FWM process. Specifically, the total energy of each pulse $E = \sum_{k,\ell,m} E_{k,\ell,m}(z) = const$, the number of photons in each pulse $N = \sum_{k,\ell,m} N_{k,\ell,m}(z) = const$, the longitudinal component of the momentum of motion of a pulse $P = \sum_{k,\ell,m} P_{k,\ell,m}(z) = const$, and the longitudinal component of the pulse OAM $M = \sum_{k,\ell,m} M_{k,\ell,m}(z) = const$. All of the conserved quantities (E,N,P,M) are fully defined by the injection conditions of the laser beam into the fiber. Whereas, during propagation, FWM leads to energy exchange between fiber modes, similarly to particle collisions in a gas, thus shuffling the values of $N_{\ell,m,k}(z)$. One could expect then that the photon system reaches thermodynamic equilibrium over a finite "time" (i.e., a finite distance z), owing to the well known Onsager's principle of detailed balance. When applied to mode interactions, this principle means that, at thermal equilibrium, each elementary FWM process leading to energy transfer into some modes is equally probable as its reverse process.

As it is well-known, at thermal equilibrium, which is reached after a certain propagation distance, say, z^*, the statistics of an ideal gas is described by the Boltzmann distribution. At $z > z^*$, this leads to a RJ distribution for the number of photons $N_{k,\ell,m} = \frac{T}{\hbar\omega'_{k,\ell,m} - \mu}$ occupying the mode (k,ℓ,m) in the coordinate system moving with the light pulse at speed V, and rotating with its angular velocity Ω [30]. Here, T is

the *statistical* temperature of photons in a light pulse (which is analogous to the temperature of electrons in particle accelerator electron beams), μ is the chemical potential, $\omega'_{k,\ell,m} = \gamma(\omega_k - V p_{\ell,m} - \Omega m)$, and $\gamma = 1/(1 - V^2/c^2)^{1/2} \simeq \sqrt{2}$ is a relativistic factor, c is the speed of light in vacuum. At thermal equilibrium, the number of photons occupying mode (ℓ,m) with frequency ω_k in laboratory system reads as

$$N_{k,\ell,m} = \frac{T/\sqrt{2}}{\hbar\omega_k - \frac{\mu}{\sqrt{2}} - \hbar V p_{\ell,m} - \hbar\Omega m}, \quad (2)$$

which is a generalized form of the RJ distribution. It is worth to noting that Eq.(2) can be equivalently derived starting from conservation laws, without recurring to the change of coordinate system (see Supplementary Materials).

FWM scattering of waves must obey the conservation laws of E, N, P and M, which lead to the following conditions

$$\omega_1 + \omega_2 = \omega_3 + \omega_4, \quad (3)$$
$$p_{\ell_1,m_1}(\omega_1) + p_{\ell_2,m_2}(\omega_2) = p_{\ell_3,m_3}(\omega_3) + p_{\ell_4,m_4}(\omega_4), \quad (4)$$
$$m_1 + m_2 = m_3 + m_4, \quad (5)$$

where ℓ_i, m_i and ω_i with $i = 1,2,3,4$ characterize each of the four waves. If any of Eqs.(3), (4) or (5) are not satisfied, the scattering process is forbidden, the ergodicity hypothesis fails, and the multimode optical system never reaches its thermodynamic equilibrium.

Let us consider the FWM of narrow spectrum beams, i.e., $|\omega_j - \omega_0| \ll \omega_0$. In this case, the mode propagation constants may be expanded as $p_{\ell,m}(\omega_j) = p_{\ell,m}(\omega_0) + p'_{\ell,m}(\omega_j - \omega_0) + 0.5 p''_{\ell,m}(\omega_j - \omega_0)^2$. According to Eqs.(3) and (4), FWM processes within a single transverse mode have a mismatch $\delta p_{\ell,m} = p''_{\ell,m}(\omega_1 - \omega_3)(\omega_1 - \omega_4)$. Now, the efficiency of FWM is strongly suppressed for mismatch values larger than the inverse of the nonlinear length. As a result, the nonlinear spectral broadening of a light beam is restricted at long propagation distances [31]. When the interaction of different transverse modes is involved, the FWM mismatch may be equal to zero only occasionally, that is, for just a few quartets of waves. As a result, thermalization broadening of wave spectra fails to occur. Therefore, in the following we shall limit our treatment to pulses with a narrow spectrum ($\omega_k \simeq \omega$). As a matter of fact, in our experiments we use relative narrow-band picosecond pulses, whose spectral broadening is negligible over distances as long as z^*, which turns out to be of a few meters.

In the special case of GRIN fibers, the mode propagation constants are equidistant $p_{\ell,m} = p_{0,0} - n(2\pi/L_B)$, where $n = 2\ell + |m|$ is dubbed quantum number, whereas L_B is the self-imaging distance. The condition of Eqs.(4) and (5) may be met for many quartets with $n_1 + n_2 = n_3 + n_4$ and $m_1 + m_2 = m_3 + m_4$ simultaneously. As a result, the ergodicity hypothesis is verified, and the equilibrium distribution (2) is achieved after a suitable nonlinear length: it can be written as

$$N_{\ell,m} = \frac{N_{0,0}}{1 - (2\pi V/L_B\tilde{\mu})(2\ell + |m|) + (\Omega/\tilde{\mu})m}, \quad (6)$$

where $\tilde{\mu} = \mu/\hbar\sqrt{2} + V p_{0,0} - \omega$. Note that the average power of transverse mode W_{lm} has the same distribution. In the Supplementary Notes, we show that the FWM process in a multimode fiber can be described in frame of a kinetic equation approach [7], for which the distribution (6) is found to be a stationary solution.

Importantly, Eq.(6) shows that the equilibrium RJ distribution is asymmetric with respect to $m = 0$, owing to the presence of the Lagrange's multiplier Ω, which is associated to the conservation of the total OAM (see Fig.1b, where $\Omega > 0$). Specifically, the frequencies $(2\pi V/L_B)$ and Ω must be comparable, in order to significantly modify the symmetry of the RJ distribution around $m = 0$. Whereas, if $\Omega = 0$, i.e., if the theory does not impose the conservation of the longitudinal OAM, then one recovers the conventional symmetric RJ distribution [9, 20], cfr. Fig.1a. Starting from this consideration, one can associate the presence of an OAM with the asymmetry of the mode distribution, i.e., with a non zero value of the average azimuthal index $\langle m \rangle \propto \Omega$, where

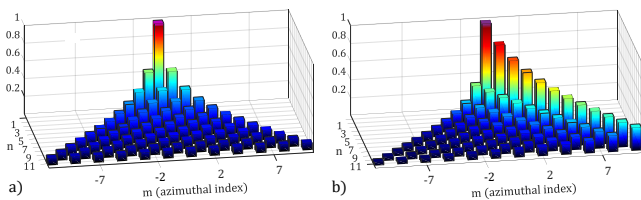$$\langle m \rangle = \sum_{\ell,m} m W_{\ell,m}. \tag{7}$$

FIG. 1. Equilibrium distribution for $\Omega/\tilde{\mu} = 0$ (a) and $\Omega/\tilde{\mu} = -0.75$ (b). In both plots, $N_{0,0} = 1$ and $2\pi V/L_B\tilde{\mu} = 1$.

The symmetry-breaking of the equilibrium distribution with respect to $m = 0$ means that, whenever thermalization without condensation ($\tilde{\mu} \neq 0$) occurs, a bell-shaped output beam profile cannot be obtained, unless $\langle m \rangle = \Omega = 0$, in spite of the fact that the fundamental mode is always the dominant mode in the output thermal distribution. Experimentally, this means that, in order to achieve BSC, one always need to inject a laser beam which does not carry OAM, e.g., on-axis Gaussian beams, coupled at the center of the fiber core. It is worth pointing out that BSC was earlier demonstrated by using different input laser coupling configurations. This is the case of Ref.[6], where a diffuser changes the spatial distribution of the input beam, and of Ref.[18], where fiber was tilted with respect to the laser direction. However, in both of these cases no OAM was carried by the input beam.

In order to seed an OAM, we chose a peculiar input beam, i.e., a Gaussian beam which is injected with a tilt angle ϑ and a transverse offset y_0 with respect to the fiber axis (see Fig.2a and b). Such injection condition leads to helical propagation of the laser beam inside the fiber core [27, 32]: its trajectory can be visualized by the naked eye by exploiting the luminescence of fiber defects [33]. The helical trajectory carries a

longitudinal OAM, which can be calculated as

$$\langle m \rangle_{the} = 2\pi \frac{y_0 \sin \vartheta}{\lambda}. \tag{8}$$

Interestingly, the magnitude of the input OAM can be tuned by acting on the injection offset. Specifically, the input OAM grows larger with y_0, and its sign can be flipped by injecting the laser at diametrically opposite points, thus reversing the helix chirality [27].

The theoretical value of $\langle m \rangle$ was used to verify the validity of our numerical and experimental mode truncation. In the Supplementary Materials, we show that, as long as the offset does not exceed a few microns, there is an excellent agreement between experimental and theoretical values of $\langle m \rangle$, which are calculated numerically with Eq.(7) or analytically via Eq.(8).

In Fig.2c we show the mode distribution at the input of the fiber, corresponding to the injection condition $\vartheta = 2°$ and $y_0 = -3$ µm. By applying Eq.(7), this corresponds to $\langle m \rangle = 0.75$ while $\langle n \rangle = \sum_{\ell,m}(2\ell + |m|)W_{\ell,m} = 2.23$.

In Fig.2d we display a convenient way for grouping the OAM modes, which will be used in the remainder of this work. In this way, we can emphasize the difference between groups of modes, represented by number g: they share the same quantum number n, but have different signs of m. Specifically, modes with odd n are grouped in three blocks, one for $m < 0$, one for $m > 0$, and one of $m = 0$. On the other hand, modes with even n miss the value $m = 0$, so they are grouped in two blocks only, one for $m < 0$ and one for $m > 0$.

In order to verify the validity of our theoretical predictions, we performed numerical simulations. Besides FWM, we also considered the effects of linear random mode coupling [34]. Further details about the numerical model are reported in the Supplementary Notes. We limited our simulation to include the 78 modes with the highest values of momentum, i.e., we only considered GRIN fiber modes with $n < 12$. This value represents a trade-off for maintaining the validity of the theoretical mode decomposition (the power fraction of the beam which is carried by modes with $n > 11$ is negligible for our injection conditions) on the one hand, and reducing the computation time of both numerical simulations and the experimental mode reconstruction algorithm, on the other hand.

We ran simulations for different values of the input peak power (W_p), in order to compare quasi-linear with highly nonlinear propagation regimes. In the latter ($W_p = 30$ kW), the mode distribution at the output of the fiber (shown in Fig.2e) turns out to be in excellent agreement with the generalized RJ distribution (6). For a clearer comparison, we report in Fig.2e the experimental (histogram) and fitting (cyan dots) values of the mode power fraction. Whereas in the linear regime (i.e., for $W_p = 0.1kW$), the lack of significant FWM interactions prevents mode thermalization: as a result, a fit with Eq.(6) fails (see Fig.2f). Details on simulation parameters are given in the Supplementary Notes.

The validity of theory and numerics was verified by the experiments based on the MD of OAM beams at output of GRIN MMFs. We used 1 ps laser pulses at 1030 nm, and a 2 m

to the equilibrium is distribution progressively reduced, when increasing W_p. This indicates that when enough power is provided, the FWM processes allow for reaching the ergodicity condition for the multimode system, hence its thermalization into an equilibrium distribution (6).
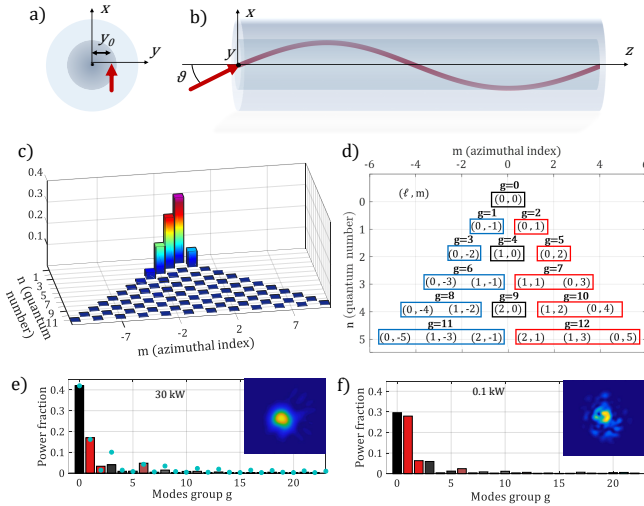

FIG. 2. a,b) Sketch of front and side views of injection conditions and helical propagation carrying positive OAM. c) Input mode distribution. d) Mode grouping by index g. e) Numerically simulated output mode distribution when $W_p = 30$ kW. Cyan dots in the 2D plot represent the values of mode power fraction, obtained by fitting experimental data with Eq.(6). f) Same as e) when $W_p = 0.1$ kW. Images in the inset of e) and f) represent output intensity profiles of the beams.

long 50/125 GRIN MMF. A full description of the set-up is reported in the Supplementary Notes. Here, we studied light thermalization, which is obtained by varying the input pulse peak power, with two different injection conditions. Specifically, in Fig.3, we report a MD analysis of the beam output profile for input beams carrying either positive (Fig.3a-c) or negative (Fig.3d-f) OAM. This was obtained by setting $y = +2$ μm or $y = -1$ μm, respectively. In the Supplementary Notes, we also report the limit case of $\langle m \rangle = 0$, which is achieved by injecting the laser with no offset with respect to the fiber axis.

In Fig.3a and d, we report histograms of the mode power fraction of the output beams, for several values of W_p. As it can be seen, the mode content changes when increasing W_p, eventually approaching an equilibrium distribution once overcoming the critical value for thermalization. One can appreciate that the distributions at $W_p = 17.6$ kW and $W_p = 26.5$ kW are quite similar. Whereas at lower powers, significantly different output mode contents are observed. In the inset of Fig.3a and d, we compare the measured output near field intensities (images in the left column) with the MD reconstructions (images in the right column). These images are impressively similar for all input power values, thus proving the accuracy of our MD method.

The cyan dots shown in the graphs of Fig.3a and d for the highest input power values provide the fitting of the experimental mode occupancy with Eq.(6). As it can be seen, a good agreement is found between the experimental mode power fractions and the prediction of the generalized RJ distribution. In Fig.3b and e, we show that the root-mean-square error (RMSE) of the observed mode occupancy with respect
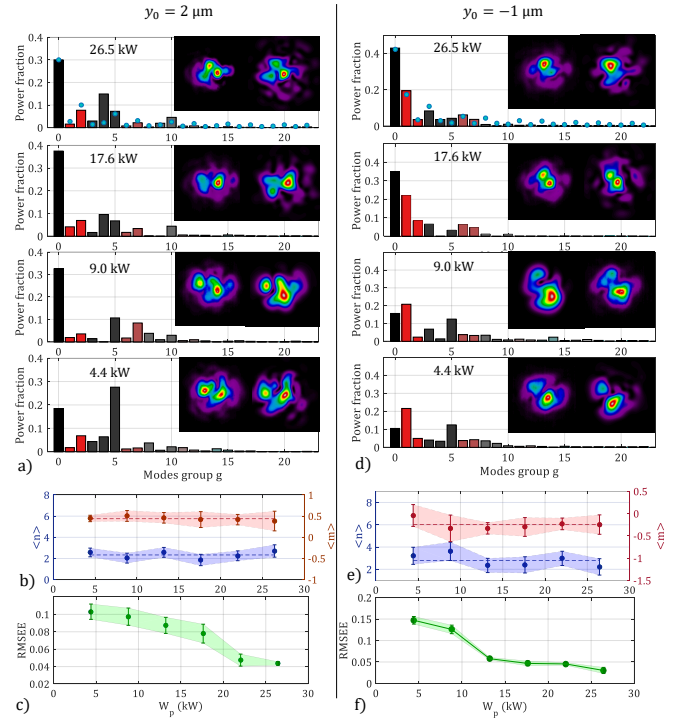
FIG. 3. Experimental results. a) MD of the output beam for different values of W_p, when $y_0 = +2$ μm.. The inset images are the measured output beam profiles (left) and their reconstructions (right). The blue dots are extracted by fitting experimental data with Eq.(6). b) Root-mean-square error of the experimental mode distributions with respect to the generalized RJ distribution (cyan dots in a) vs. input peak power. c) Conservation of $\langle n \rangle$ (blue) and $\langle m \rangle$ (red). The error bars are estimated by considering all of the reconstructions of the output beam near-field at each input power [9]. d-f) Same as a-c), when injecting the laser beam with an offset $y_0 = -1$ μm.

Finally, we proved the validity of the hypothesis behind our theoretical derivation, i.e. the conservation laws of E, N, P and M. As a matter of fact, our MD method does not allow for estimating $N_{\ell,m}$, since only averaged quantities can be extracted from camera pictures. Nevertheless, it is well known that the energy of each pulse E, and accordingly the photon number N, are conserved in our experimental conditions, since dissipative effects, of either linear or nonlinear origin, are negligible over a few meters of fiber for picosecond pulses of few tens of kW of peak power, at wavelengths around 1 μm. We experimentally verified the conservation of $\langle n \rangle$ and $\langle m \rangle$, which are related to the linear and angular momentum, respectively. As it can be seen in Fig.3c and f, when varying W_p, both quantities oscillate within the experimental error bars around a constant value. Specifically, we found that $\langle m \rangle \simeq 0.46$ for $y_0 = +2$ μm and $\langle m \rangle \simeq -0.24$ for $y_0 = -1$

μm. This is in agreement with theoretical expectations: Eq.(8) gives $\langle m \rangle = 0.44$ and $\langle m \rangle = -0.22$, respectively.

In conclusion, we derived a general theoretical description of light thermalization in multimode fibers. Theoretical predictions have been confirmed by numerical and experimental studies. Remarkably, we found that the thermalization of OAM beams in GRIN MMFs may occur without the generation of bell-shaped output beams. Our work sheds a new light on the nonlinear dynamics and manipulation of vortex optical beams by multimode fibers.

[1] S. Dyachenko, A. Newell, A. Pushkarev, and V. Zakharov, Physica D: Nonlinear Phenomena **57**, 96 (1992).

[2] C. Connaughton, C. Josserand, A. Picozzi, Y. Pomeau, and S. Rica, Phys. Rev. Lett. **95**, 263901 (2005).

[3] A. Picozzi, J. Garnier, T. Hansson, P. Suret, S. Randoux, G. Millot, and D. Christodoulides, Physics Reports **542**, 1 (2014), optical wave turbulence: Towards a unified nonequilibrium thermodynamic formulation of statistical nonlinear optics.

[4] C. Sun, S. Jia, C. Barsi, S. Rica, A. Picozzi, and J. W. Fleischer, Nature Physics **8**, 470 (2012).

[5] D. R. Tilley and J. Tilley, *Superfluidity and superconductivity* (Routledge, 2019).

[6] K. Baudin, A. Fusaro, K. Krupa, J. Garnier, S. Rica, G. Millot, and A. Picozzi, Phys. Rev. Lett. **125**, 244101 (2020).

[7] P. Aschieri, J. Garnier, C. Michel, V. Doya, and A. Picozzi, Phys. Rev. A **83**, 033838 (2011).

[8] F. O. Wu, A. U. Hassan, and D. N. Christodoulides, Nature Photonics **13**, 776 (2019).

[9] F. Mangini, M. Gervaziev, M. Ferraro, D. Kharenko, M. Zitelli, Y. Sun, V. Couderc, E. Podivilov, S. Babin, and S. Wabnitz, arXiv preprint arXiv:2111.08063 (2021).

[10] K. Krupa, A. Tonello, A. Barthélémy, V. Couderc, B. M. Shalaby, A. Bendahmane, G. Millot, and S. Wabnitz, Phys. Rev. Lett. **116**, 183901 (2016).

[11] K. Krupa, A. Tonello, A. Barthélémy, T. Mansuryan, V. Couderc, G. Millot, P. Grelu, D. Modotto, S. A. Babin, and S. Wabnitz, APL Photonics **4**, 110901 (2019).

[12] K. Krupa, A. Tonello, B. M. Shalaby, M. Fabert, A. Barthélémy, G. Millot, S. Wabnitz, and V. Couderc, Nat. Photon. **11**, 234 (2017).

[13] E. V. Podivilov, D. S. Kharenko, V. A. Gonta, K. Krupa, O. S. Sidelnikov, S. Turitsyn, M. P. Fedoruk, S. A. Babin, and S. Wabnitz, Phys. Rev. Lett. **122**, 103902 (2019).

[14] A. Fusaro, J. Garnier, K. Krupa, G. Millot, and A. Picozzi, Phys. Rev. Lett. **122**, 123902 (2019).

[15] E. Deliancourt, M. Fabert, A. Tonello, K. Krupa, A. Desfarges-Berthelemot, V. Kermene, G. Millot, A. Barthélémy, S. Wabnitz, and V. Couderc, OSA Continuum **2**, 1089 (2019).

[16] E. Deliancourt, M. Fabert, A. Tonello, K. Krupa, A. Desfarges-Berthelemot, V. Kermene, G. Millot, A. Barthélémy, S. Wab-nitz, and V. Couderc, Opt. Express **27**, 17311 (2019).

[17] Y. Leventoux, A. Parriaux, O. Sidelnikov, G. Granger, M. Jossent, L. Lavoute, D. Gaponov, M. Fabert, A. Tonello, K. Krupa, A. Desfarges-Berthelemot, V. Kermene, G. Millot, S. Février, S. Wabnitz, and V. Couderc, Opt. Express **28**, 14333 (2020).

[18] M. Fabert, M. Săpânțan, K. Krupa, A. Tonello, Y. Leventoux, S. Février, T. Mansuryan, A. Niang, B. Wetzel, G. Millot, S. Wabnitz, and V. Couderc, Scientific Reports **10**, 20481 (2020).

[19] M. Zitelli, M. Ferraro, F. Mangini, and S. Wabnitz, Photon. Res. **9**, 741 (2021).

[20] Y. Wu, H. Pourbeyram, D. N. Christodoulides, and F. W. Wise, Opt. Lett. **46**, 3312 (2021).

[21] L. Allen, M. W. Beijersbergen, R. Spreeuw, and J. Woerdman, Physical Review A **45**, 8185 (1992).

[22] J. Wang, J.-Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, *et al.*, Nature Photonics **6**, 488 (2012).

[23] R. Fickler, G. Campbell, B. Buchler, P. K. Lam, and A. Zeilinger, Proceedings of the National Academy of Sciences **113**, 13642 (2016).

[24] H. Oraizi and H. Emamian, Scientific Reports **10**, 1 (2020).

[25] X. Zhuang, Science **305**, 188 (2004).

[26] M. D. Gervaziev, I. Zhdanov, D. S. Kharenko, V. A. Gonta, V. M. Volosi, E. V. Podivilov, S. A. Babin, and S. Wabnitz, Laser Physics Letters **18**, 015101 (2020).

[27] F. Mangini, M. Ferraro, M. Zitelli, V. Kalashnikov, A. Niang, T. Mansuryan, F. Frezza, A. Tonello, V. Couderc, A. Aceves, and S. Wabnitz, Scientific reports **11**, 1 (2021).

[28] K. Krupa, A. Tonello, V. Couderc, A. Barthélémy, G. Millot, D. Modotto, and S. Wabnitz, Phys. Rev. A **97**, 043836 (2018).

[29] J. Lægsgaard, Opt. Lett. **43**, 2700 (2018).

[30] L. Landau and E. Lifshitz, in *Statistical Physics (Third Edition)*, edited by L. Landau and E. Lifshitz (Butterworth-Heinemann, Oxford, 1980) third edition ed., pp. 111–157.

[31] S. A. Babin, D. V. Churkin, A. E. Ismagulov, S. I. Kablukov, and E. V. Podivilov, J. Opt. Soc. Am. B **24**, 1729 (2007).

[32] F. Mangini, M. Ferraro, M. Zitelli, A. Niang, T. Mansuryan, A. Tonello, V. Couderc, A. De Luca, S. Babin, F. Frezza, *et al.*, Optics Letters **47**, 1 (2022).

[33] F. Mangini, M. Ferraro, M. Zitelli, A. Niang, A. Tonello, V. Couderc, and S. Wabnitz, Physical Review Applied **14**, 054063 (2020).

[34] O. S. Sidelnikov, E. V. Podivilov, M. P. Fedoruk, and S. Wabnitz, Optical Fiber Technology **53**, 101994 (2019).

Supplementary Notes

BASE TRANSFORMATION FROM LG TO OAM

Since the mode decomposition code operates on a Laguerre-Gauss mode basis, while the theory is expressed in a generic OAM basis, all experimentally obtained mode decomposition data had to be transformed into an appropriate basis. The total field in the therms of mode decomposition software is expressed as follow:

$$U = \sum_{l,m} N_m F_{l,|m|} (A_{l,m} \cos m\phi + A_{l,-m} \sin m\phi) \qquad (9)$$

where $F_{l,|m|}$ the function represents Laguerre-Gaussian (LG) modes with the (l,m) indexes, $A_{l,m}$ and $A_{l,-m}$ the measured values. N_m normalized coefficient that is equal to 1 for $m=0$ and $\sqrt{2}$ for $m \neq 0$ as normalization for $\sin(x)$ and $\cos(x)$ differs from $exp(ix)$. To simplify further derivations just consider U_l component of the field and omit the l in mode amplitudes:

$$U_m \propto N_m (A_m \cos m\phi + A_{-m} \sin m\phi). \qquad (10)$$

Let's express the sine and cosine function by complex exponent using Euler's formula $e^{ix} = \cos x + i \sin x$ for both m and $-m$ cases:

$$\begin{cases} \cos m\phi = (e^{im\phi} + e^{-im\phi})/2 \\ \sin m\phi = (e^{im\phi} - e^{-im\phi})/2. \end{cases} \qquad (11)$$

Substitute the last Eqs. (4) into (3) and let's group the coefficient before the same exponents:

$$U_m \propto \frac{N_m}{2}(A_m - iA_{-m})e^{im\phi} + \frac{N_m}{2}(A_m + iA_{-m})e^{-im\phi} =$$
$$= B_m e^{im\phi} + B_{-m} e^{-im\phi} \qquad (12)$$

where we denote B_m and B_{-m} as follow:

$$\begin{cases} B_{l,0} = A_l \\ B_{l,m} = (A_m - iA_{-m})/\sqrt{2} \\ B_{l,-m} = (A_m + iA_{-m})/\sqrt{2} \end{cases} \qquad (13)$$

In matrix form it will look like

$$\begin{pmatrix} B_{l,m} \\ B_{l,-m} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \begin{pmatrix} A_{l,m} \\ A_{l,-m}. \end{pmatrix} \qquad (14)$$

To calculate the angular momentum express $|B_{l,m}|^2$ through $A_{l,m}$ and $A_{l,-m}$ and take into account that $z = a + ib$ and $z - z^* = 2ib$:

$$2|B_{l,m}|^2 = |A_{l,m}|^2 + |A_{l,-m}|^2 + 2\Im(A_{l,m}^* A_{l,-m}) \qquad (15)$$
$$2|B_{l,-m}|^2 = |A_{l,m}|^2 + |A_{l,-m}|^2 - 2\Im(A_{l,m}^* A_{l,-m}).$$

DERIVATION OF THE EQUILIBRIUM DISTRIBUTION FROM CONSERVATION LAWS

We aim at studying the thermalization of a multimode optical system taking into account the conservation laws of the number of photons N, the momentum P and the *longitudinal* orbital angular momentum (OAM), which we dub M. For sake of simplicity, we are considering monochromatic waves, otherwise we should have added a fourth conservation law: that of the energy of the pulse, as in the main text.

Thermalization occurs when maximizing the entropy, which reads as

$$S = \ln \mathcal{W}, \qquad (16)$$

where $\mathcal{W}$ is the number of ways in which one can distribute N (indistinguishable) photons in distinct optical modes. Let us label with the index i all of the g_i modes having the momentum P_i and OAM M_i. Thus the parameter g_i is the mode degeneracy. Assigning to the i-th group of modes an occupancy N_i, we can write

$$\mathcal{W} = \prod_i \frac{(N_i + g_i - 1)!}{N_i!(g_i - 1)!}. \qquad (17)$$

Here and in the following, the index i runs over all of the non-degenerate guided modes, whose (finite) number is established by the fiber cut-off. The entropy becomes:

$$S = \sum_i (N_i + g_i - 1)\ln(N_i + g_i - 1) - N_i \ln N_i - \ln(g_i - 1)!, \qquad (18)$$

where we used the Stirling formula $\ln x! \simeq x \ln x$ for $x \gg 1$. The conservation laws for N, P and M respectively read:

$$N = \sum_i N_i, \qquad (19)$$

$$P = \sum_i P_i N_i, \qquad (20)$$

and

$$M = \sum_i M_i N_i. \qquad (21)$$

The conservation laws (19), (20) and (21) are the constraints for the maximization of S. We can introduce 3 Lagrange's multiplier α, β and γ in order to find the stationary points of S and we impose

$$\frac{\partial}{\partial N_i} \left[S + \alpha \left(N - \sum_i N_i \right) + \beta \left(P - \sum_i P_i N_i \right) + \gamma \left(M - \sum_i M_i N_i \right) \right] = 0. \qquad (22)$$

This gives:

$$\ln \left(\frac{N_i + g_i - 1}{N_i} \right) - \alpha - \beta P_i - \gamma M_i = 0 \qquad (23)$$

$$N_i = \frac{g_i - 1}{e^{\alpha + \beta P_i + \gamma M_i} - 1}, \tag{24}$$

which, imposing $g_i \gg 1$, yields the Bose-Einstein distribution:

$$N_i = \frac{g_i}{e^{\alpha + \beta P_i + \gamma M_i} - 1}. \tag{25}$$

Under the approximation $N_i \gg g_i$, which holds for highly multimode systems, the Bose-Eistein distribution boils down to

$$N_i = \frac{g_i}{\alpha + \beta P_i + \gamma M_i}. \tag{26}$$

Note that, so far, we have extended the derivation of Ref.[8] to the case of systems which conserve the longitudinal OAM. Now, we would like to apply this result to multimode fibers, where the modes are generally defined by two integer numbers (ℓ, m). Therefore, it is convenient to get rid of the degeneracy. To do so, we need to do the following substitution:

$$\sum_i g_i X_i = \sum_{\ell,m} X_{\ell,m}, \tag{27}$$

where the new sums run over all of the possible values of ℓ and m, independently of the mode degeneracy and X is a generic quantity which depends on the mode indexes. Accordingly, Eq.(19) becomes:

$$N = \sum_{\ell,m} \frac{1}{\alpha + \beta P_{\ell,m} + \gamma M_{\ell,m}}, \tag{28}$$

and we can identify the occupancy of the mode (ℓ, m) with

$$N_{\ell,m} = \frac{1}{\alpha + \beta P_{\ell,m} + \gamma M_{\ell,m}}, \tag{29}$$

so that, by exploiting the equalities (27) and (28), the conservation laws (19), (20) and (21) respectively become

$$N = \sum_{\ell,m} N_{\ell,m}, \tag{30}$$

$$P = \sum_{\ell,m} P_{\ell,m} N_{\ell,m}, \tag{31}$$

and

$$M = \sum_{\ell,m} M_{\ell,m} N_{\ell,m}. \tag{32}$$

So far, we have not chosen a basis for the mode representation. As we want to exploit the OAM of the field, it is convenient to choose a mode basis for which the *longitudinal* OAM operator

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \tag{33}$$

is diagonalized. Such a basis is a generalization of that of Laguerre-Gauss modes, and it is dubbed OAM basis. In the latter, the radial and azimuthal variables are separated, e.g., the mode amplitude $\psi_{\ell,m}$ can be written as

$$\psi_{\ell,m}(z, r, \phi) = A_{\ell,m}(z) e^{-i P_{\ell,m} z} \cdot F_{\ell,m}(r) \cdot e^{im\phi}, \tag{34}$$

where $2\pi \int dr \cdot r |F(r)|^2 = 1$. Therefore, each mode (ℓ, m) carries a longitudinal OAM $M_{\ell,m} = m\hbar$, as

$$\hat{L}_z \cdot \psi_{\ell,m} = m\hbar \cdot \psi_{\ell,m} \tag{35}$$

and the total longitudinal OAM reads as:

$$M = \hbar \sum_{\ell,m} m \cdot N_{\ell,m}. \tag{36}$$

Finally, with this basis we recover the generalized Rayleigh-Jeans distribution of our manuscript, after defining the following parameters:

$$\alpha = \frac{\hbar \omega \sqrt{2} - \mu}{T}, \tag{37}$$

$$\beta = -\frac{\hbar V \sqrt{2}}{T}, \tag{38}$$

$$\gamma = -\frac{\hbar \Omega \sqrt{2}}{T}. \tag{39}$$

KINETIC EQUATION

The FWM process in a multimode fiber can be described in frame of a kinetic equation approach [7]. For simplicity of notation, let us introduce a multi-index $\mathbf{q} = (\ell, m)$. The slowly variable mode amplitudes $A_{\mathbf{q}}(z) = A_{\ell,m}(z)$ obey the nonlinear evolution equation

$$\sum_{\mathbf{q1}} F_{\mathbf{q1}}(r) e^{im_1\phi - ip_{\mathbf{q1}}z} \frac{dA_{\mathbf{q1}}}{dz} =$$
$$= -i(\omega n_2/c) \sum_{\mathbf{q2},\mathbf{q3},\mathbf{q4}} e^{-i(m_2 - m_3 - m_4)\phi} e^{i(p_{\mathbf{q2}} - p_{\mathbf{q3}} - p_{\mathbf{q4}})z}$$
$$\cdot F_{\mathbf{q2}}(r) F_{\mathbf{q3}}(r) F_{\mathbf{q4}}(r) A_{\mathbf{q2}}^*(z) A_{\mathbf{q3}}(z) A_{\mathbf{q4}}(z) \tag{40}$$

where $n_2 \simeq 2 \cdot 10^{-16} \, sm^2/W$ is the coefficient of the nonlinear correction to the refractive index of silica glass $\delta n_{NL} = n_2 |A_{tot}|^2$, and c vacuum speed of light. By using the orthogonality of the radial eigenfunctions, we may multiply Eq.(40) by $F_{\mathbf{q1}}(r) \exp(-i(m_1\phi - p_{\mathbf{q1}}z))$, and then integrate over the fiber cross-section. As a result, we obtain the propagation equation for the light amplitude in the modal representation

$$\frac{dA_{\mathbf{q1}}}{dz} =$$
$$-i \sum_{\mathbf{q2},\mathbf{q3},\mathbf{q4}} U_{\mathbf{q1},\mathbf{q2},\mathbf{q3},\mathbf{q4}} A_{\mathbf{q2}}^*(z) A_{\mathbf{q3}}(z) A_{\mathbf{q4}}(z) e^{i(p_{\mathbf{q1}} + p_{\mathbf{q2}} - p_{\mathbf{q3}} - p_{\mathbf{q4}})z}$$
$$\cdot \Delta(m_1 + m_2 - m_3 - m_4), \tag{41}$$

where $\Delta(j)$ is the Kronecker's symbol, and

$$U_{\mathbf{q}1,\mathbf{q}2,\mathbf{q}3,\mathbf{q}4} = \frac{\omega n_2}{c}\int\limits_0^\infty rdr F_{\mathbf{q}1}(r)F_{\mathbf{q}2}(r)F_{\mathbf{q}3}(r)F_{\mathbf{q}4}(r)$$

is an overlap integral. Let us assume that, at some point z^* along the fiber, the optical system reaches its steady state. Namely, all light modes have a Gaussian and delta-correlated statistics $\langle A_{\mathbf{q}}(z^*)A_{\mathbf{q}'}(z^*)\rangle = \Delta(l-l')\Delta(m-m')I_{\mathbf{q}}(z^*)$, and the mode power $I_{\mathbf{q}}(z)$ remains a constant for $z > z^*$. Then, exactly at $z = z^*$, one obtains

$$\begin{aligned}
&\frac{dI_{\mathbf{q}1}}{dz}|_{z=z^*} = -2\mathrm{Re}\{\\
&i\sum_{\mathbf{q}2,\mathbf{q}3,\mathbf{q}4} U_{\mathbf{q}1,\mathbf{q}2,\mathbf{q}3,\mathbf{q}4}\langle A_{\mathbf{q}1}^*(z^*)A_{\mathbf{q}2}^*(z^*)A_{\mathbf{q}3}(z^*)A_{\mathbf{q}4}(z^*)\rangle\\
&\cdot e^{i(p_{\mathbf{q}1}+p_{\mathbf{q}2}-p_{\mathbf{q}3}-p_{\mathbf{q}4})z^*}\Delta(m_1+m_2-m_3-m_4))\} =\\
&= 2\mathrm{Im}\sum_{\mathbf{q}2}U_{\mathbf{q}1,\mathbf{q}2,\mathbf{q}1,\mathbf{q}2}2I_{\mathbf{q}1}(z^*)I_{\mathbf{q}2}(z^*) = 0. \quad (42)
\end{aligned}$$

From Eq.41, it follows that the amplitude of any mode $A_{\mathbf{q}j}(z)$ should change its value according to the equation

$$\begin{aligned}
&\delta A_{\mathbf{q}j}(z) = A_{\mathbf{q}j}(z) - A_{\mathbf{q}j}(z^*) =\\
&= -i\int\limits_{z^*}^z dz' \sum_{\mathbf{q}j',\mathbf{q}j'',\mathbf{q}j'''} U_{\mathbf{q}j,\mathbf{q}j',\mathbf{q}j'',\mathbf{q}j'''}A_{\mathbf{q}j'}^*(z')A_{\mathbf{q}j''}(z')A_{\mathbf{q}j'''}(z')\\
&\cdot e^{i(p_{\mathbf{q}j}+p_{\mathbf{q}j'}-p_{\mathbf{q}j''}-p_{\mathbf{q}j'''})z'}\Delta(m_1+m_2-m_3-m_4). \quad (43)
\end{aligned}$$

Let us assuming, that the length $z - z^*$ is smaller than the interaction distance $L_{NL} \simeq \lambda/2\pi\delta n_{NL}$, where $\delta n_{NL} \simeq n_2 W_p/S_{eff}$ - nonlinear correction to refractive index, and S_{eff} - effective area of light beam. Than Eq. (43) leads to a nonzero correction to the right-hand side (RHS) of Eq.(42), which is of the first order with respect to the small perturbation δA. One obtains

$$\begin{aligned}
&\frac{dI_{\mathbf{q}1}}{dz}|_{z>z^*} = -2\mathrm{Re}\{\\
&i\sum_{\mathbf{q}2,\mathbf{q}3,\mathbf{q}4} U_{\mathbf{q}1,\mathbf{q}2,\mathbf{q}3,\mathbf{q}4}\big(\langle\delta A_{\mathbf{q}1}^*(z)A_{\mathbf{q}2}^*(z^*)A_{\mathbf{q}3}(z^*)A_{\mathbf{q}4}(z^*)\rangle +\\
&\langle A_{\mathbf{q}1}^*(z^*)\delta A_{\mathbf{q}2}^*(z)A_{\mathbf{q}3}(z^*)A_{\mathbf{q}4}(z^*)\rangle +\\
&\langle A_{\mathbf{q}1}^*(z^*)A_{\mathbf{q}2}^*(z^*)\delta A_{\mathbf{q}3}(z)A_{\mathbf{q}4}(z^*)\rangle +\\
&\langle A_{\mathbf{q}1}^*(z^*)A_{\mathbf{q}2}^*(z^*)A_{\mathbf{q}3}(z^*)\delta A_{\mathbf{q}4}(z)\rangle\big)e^{i(p_{\mathbf{q}1}+p_{\mathbf{q}2}-p_{\mathbf{q}3}-p_{\mathbf{q}4})z^*}\\
&\cdot\Delta(m_1+m_2-m_3-m_4)\}. \quad (44)
\end{aligned}$$

By substituting here δA from Eq.(43), and averaging over the Gaussian statistics, one obtains the general wave kinetic equation for monochromatic light in a multimode optical fiber:

$$\begin{aligned}
&\frac{dI_{\mathbf{q}1}}{dz}|_{z>z^*} = 4\sum_{\mathbf{q}2,\mathbf{q}3,\mathbf{q}4}|U_{\mathbf{q}1,\mathbf{q}2,\mathbf{q}3,\mathbf{q}4}|^2 I_{\mathbf{q}1}(z^*)I_{\mathbf{q}2}(z^*)I_{\mathbf{q}3}(z^*)I_{\mathbf{q}4}(z^*)\\
&\cdot\left(I_{\mathbf{q}1}^{-1}(z^*)+I_{\mathbf{q}2}^{-1}(z^*)-I_{\mathbf{q}3}^{-1}(z^*)-I_{\mathbf{q}4}^{-1}(z^*)\right)\\
&\cdot\delta_{z,z*}(p_{\mathbf{q}1}+p_{\mathbf{q}2}-p_{\mathbf{q}3}-p_{\mathbf{q}4})\Delta(m_1+m_2-m_3-m_4) \quad (45)
\end{aligned}$$

where

$$\delta_{z,z*}(x) = \frac{\sin(x(z-z*))}{x(z-z*)}. \quad (46)$$

For a steady-state solution, the RHS of Eq.(45) must be equal to zero.

For GRIN multimode fibers, after averaging over the small self-imaging distance L_B, the function (46) reduces to $\delta_{z,z*}(p_{\mathbf{q}1}+p_{\mathbf{q}2}-p_{\mathbf{q}3}-p_{\mathbf{q}4}) = |z - z^*|\Delta(k1+k2-k3-k4)$ for $z - z^* \gg L_B \simeq 5$ mm. One could check that the thermal equilibrium distribution

$$I_{\ell,m}^{-1} = I_{0,0}^{-1}(1+a(2\ell+|m|)+bm) \quad (47)$$

is indeed a steady-state solution of the wave kinetic equation (45): the expression in parenthesis equals to zero, thanks to the delta-symbols. The constants $I_{0,0}$, a and b are defined by the three conservation laws (power, momentum and angular momentum) from their input values. The number of photons $N_{\ell,m}$ passing through the end of the fiber at time τ has the same distribution, i.e., $N_{\ell,m} = \tau I_{\ell,m}/\hbar\omega$.

On the other hand, for SI multimode fibers the function (46) only reduces to the Dirac's delta-function in the limit of an infinite fiber length, i.e., whenever $z - z^* \to \infty$. Formally, one could expect to observe a thermal distribution in the form $I_{\ell,m} = I_{0,0}/(1+ap_{\ell,m}+bm)$, but the mismatch $\delta p = p_{\ell1,m1}+p_{\ell2,m2}-p_{\ell3,m3}-p_{\ell4,m4} \neq 0$ for any combination of transverse mode indices, so that $I_{\ell,m}$ does not reach the steady-state value if the constant $a \neq 0$ (i.e., the expression among brackets in (46) does not turn to zero).

For small powers, the main process of energy exchange is random mode coupling. Momentum and angular momentum conservation laws are broken by this process. Random mode coupling can be described as the scattering of the modal amplitude $A_{\ell1,m1}$ on a noisy Bragg grating with a period $2\pi/(p_{\ell1,m1}-p_{\ell2,m2})$ into the modal amplitude $A_{\ell2,m2}$, and back. In this case, the thermal equilibrium distribution is the equidistribution of mode powers, $I_{\ell,m} = I_{0,0}$: hence both constants a and b must be equal to zero.

MODE TRUNCATION

Here, we discuss the agreement between both simulation and experiments with theoretical predictions, when considering only modes with $n < 12$. In Fig.4, we report as green circles the values of $\langle m \rangle$ and $\langle n \rangle$ at the fiber input, when setting $\vartheta = 2°$ while varying the injection offset y_0. The corresponding output values, calculated by numerical simulations, are shown by black squares. These are calculated by means of Eq.(7). Similarly, the values of $\langle m \rangle$ and $\langle n \rangle$ obtained by MD experiments are shown by red diamonds. Finally, the theoretical values calculated by Eq.(8) are shown by a solid line. As it can be seen, as long as the offset is smaller than a few microns, the MD analysis is in excellent agreement with the theory. To the contrary, when y_0 grows too large, a discrepancy between both numerics and experiments and the analytical prediction

is found. As a consequence, in order to re-establish a coherence between those, one would need to extend the number of modes which are considered in both simulations and experiments. This would considerably increase both computation time and MD reconstruction time in the experiments. More importantly, such high-order modes may in fact not be guided in the actual GRIN fiber. Therefore, in our work we limited to offset up to 2 μm of magnitude.
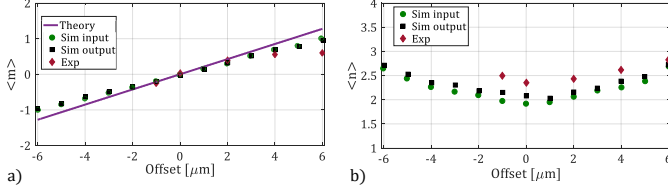

FIG. 4. a) Values of $\langle m \rangle$ as a function of the offset y_0. The solid line is calculated by Eq. (8); Green circles, black squares and red diamonds are calculated by Eq.(7) when considering the mode distribution at the fiber input, and at the fiber output obtained by numerical simulations and MD experiments, respectively. b) Same as a) for $\langle n \rangle$.

NUMERICAL METHODS

To simulate the spatial evolution of the modal amplitudes of a beam in a GRIN MMF, we numerically solved the following system of coupled-mode propagation equations [34]:

$$
\frac{\partial A_{\ell,m}}{\partial z} = \sum_{\ell1,m1} C_{m1,m}^{\ell1,\ell} A_{\ell1,m1} - \frac{\alpha}{2} A_{\ell,m} -
$$
$$
- i \frac{kn_2}{n_0} \sum_{m1,m2,m3} \sum_{\ell1,\ell2,\ell3} f_{m1,m2,m3,m}^{\ell1,\ell2,\ell3,\ell} A_{\ell1,m1}^* A_{\ell2,m2} A_{\ell3,m3}, \quad (48)
$$

where $A_{\ell,m}$ are the mode amplitudes obtained by projecting the field on the basis of Laguerre-Gaussian modes; $k = 2\pi n_0/\lambda$ is the wavenumber and n_0 is the core refractive index. The coefficients $f_{m1,m2,m3,m}^{\ell1,\ell2,\ell3,\ell}$ are obtained by calculating the overlap integrals of the spatial modes distributions. The coupled-mode equations take into account the presence of the Kerr effect and linear losses ($\alpha = 2.72$ dB/km). In addition, we also included terms describing random linear coupling between all spatial modes, owing to fiber imperfections, bends and stresses. Specifically, the coefficients $C_{m,\ell}^{m1,\ell1}$ are normally distributed random numbers with zero mean value and standard deviation $s = 2 \cdot 10^{-4}$. As done for the experimental mode decomposition, in the simulations we only consider the 78 modes whose mode number $n = 2\ell + |m| \leq 11$. Finally, in our simulations we consider a propagation distance of 10 m, whereas the spatial integration step was set to 0.01 mm.

EXPERIMENTAL SET-UP

The MD experimental setup that we used to study the OAM thermalization distribution in GRIN MMFs is shown in Fig.**??**. It consists of an ultra-short pulse laser system pumped by a femtosecond Yb-based laser (Lightconversion PHAROS-SP-HP), generating pulses with adjustable duration (by means of a dispersive pulse stretcher), at 100 kHz repetition rate and $\lambda = 1030$ nm, and with Gaussian beam shape (M^2=1.3). The pulse shape was measured by using an autocorrelator (APE PulseCheck type 2), resulting in a sech temporal shape with pulse width equal to 1 ps. As shown in Fig.**??**, the laser beam was injected by a positive lens (L_0) into the core of the GRIN fiber. The input diameter at $1/e^2$ of peak intensity was measured to be 30 μm. We employed 2 m long standard 50/125 GRIN fibers (GIF50E from Thorlabs), whose core radius, core refractive index along the axis, relative core-cladding index difference, numerical aperture and fundamental mode radius at $\lambda = 1030$ nm are $r_c = 25$ μm, $n_0 = 1.472$, $\Delta = 0.0103$, $NA = 0.2$ and $r_{0,0} = 6.33$ μm, respectively. The fiber is positioned on a 4-Axis micro-block stages, to vary the offset and choose a proper input tilt angle ϑ. The near-field profile at the fiber output is imaged onto an SLM (Hamamatsu LCOS- X15213) by means of two confocal lenses (L_1, with $f_1 = 2.75$ mm and L_2, with $f_2 = 400$ mm). Between those, we placed a bandpass filter (BPF, 1030 ± 5 nm), a half-wave plate ($\lambda/2$), and a polarizer (P). Our measurement system allows us to avoid the parasitic influence of nonlinear frequency conversions, e.g., provided by Raman scattering or geometric parametric instability [10], which is detrimental for our MD reconstruction algorithm. As a matter of fact, the phase pattern on the SLM for the profile reconstruction algorithm must be chosen at a given wavelength. We could also tune the intensity of the beam reaching the SLM by means of the $\lambda/2$. A flip mirror (FM) is used for imaging the near field profile at the fiber output facet onto an IR camera (NF, Gentec Beamage-4M-IR). Images acquired in this way were used as a reference, in order to check the quality of the reconstruction made by the MD algorithm. At last, a convex lens ($L_3 = 400$ mm) projects the field reflected by the SLM onto a second camera (FF camera). The lens is placed in the middle between the SLM and the camera, so that both these objects are at its focal distance. Finally, the beam average power both at the input and the output of the fiber was measured by a photo-diode power meter (Thorlabs).

ZERO ANGULAR MOMENTUM

Here, we report the particular case of beam propagating without carrying an OAM. In this case, the distribution 2 boils down to the RJ distribution. We consider the case of a meridional ray, which is achieved by injecting the laser beam without any offset with respect to the fiber axis ($y_0 = 0$) while keeping a tilted geometry ($\vartheta = 2°$). In Fig.**??**a, we report the output mode distribution when operating in the linear (bottom
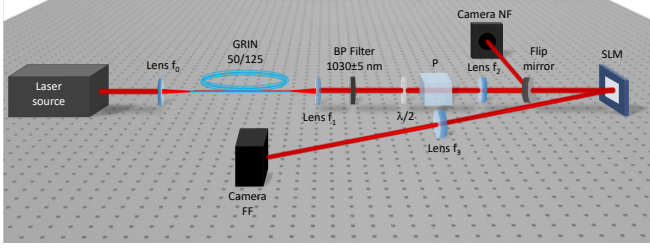
FIG. 5. Sketch of the experimental set-up to study the modes decomposition.

panel) or in the nonlinear (top panel) regime. As it can be seen, the mode content changes when varying W_p. Specifically, the fundamental mode becomes more populated when W_p grows larger. Eventually, the output distribution reaches an equilibrium, as for the cases considered in the main text. However, in this case, when the equilibrium is reached, we may remark the symmetry of the power fraction associated with opposite signs of the azimuthal index. In the inset of Fig.**??**a, we compare the measured output near field intensities (left) with the MD reconstructions (right). In Fig.3b, we show that the root-mean-square error (RMSE) of the observed mode occupancy with respect to the RJ distribution progressively reduces, when increasing W_p. By comparing Fig.3b with the experimental results in the main text, one can see that, in the abseence of OAM, thermodynamic equilibrium is reached for lower powers. Finally, in Fig.3c we verified the conservation of $\langle m \rangle$ and $\langle n \rangle$.
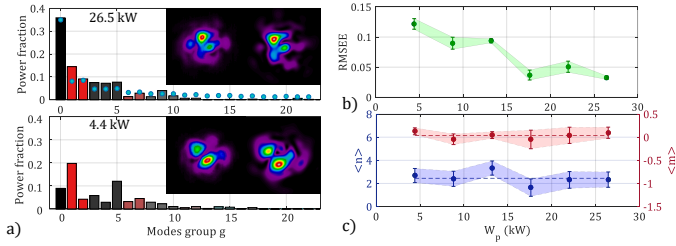

FIG. 6. Experimental results in the case of zero angular momentum. a) MD of the output beam when operating either in the linear (bottom panel) or in the nonlinear (top panel) regime. The inset images show the measured output beam profiles (left) and their reconstructions (right). The blue dots are extracted from the fit of the experimental data by Eq.(6). b) Root-mean-square error of the experimental mode distributions with respect to the RJ distribution (blue dots in a) vs. input peak power. c) Conservation of $\langle n \rangle$ (blue) and $\langle m \rangle$ (red).