AAAS
Research
Volume 2021, Article ID 9780760, 9 pages
https://doi.org/10.34133/2021/9780760

Research
A SCIENCE PARTNER JOURNAL

Research Article

Quantifying Information via Shannon Entropy in Spatially Structured Optical Beams

Maria Solyanik-Gorgone,[1] **Jiachi Ye**,[1] **Mario Miscuglio**,[1] **Andrei Afanasev**,[2] **Alan E. Willner**,[3] **and Volker J. Sorger**[1]

[1]*Department of Electrical and Computer Engineering, The George Washington University, Washington, DC 20052, USA*
[2]*Department of Physics, The George Washington University, Washington, DC 20052, USA*
[3]*Department of Electrical Engineering at University of Southern California, Los Angeles, California 90089, USA*

Correspondence should be addressed to Volker J. Sorger; sorger@gwu.edu

Received 24 May 2021; Accepted 14 September 2021; Published 11 November 2021

Copyright © 2021 Maria Solyanik-Gorgone et al. Exclusive Licensee Science and Technology Review Publishing House. Distributed under a Creative Commons Attribution License (CC BY 4.0).

While information is ubiquitously generated, shared, and analyzed in a modern-day life, there is still some controversy around the ways to assess the amount and quality of information inside a noisy optical channel. A number of theoretical approaches based on, e.g., conditional Shannon entropy and Fisher information have been developed, along with some experimental validations. Some of these approaches are limited to a certain alphabet, while others tend to fall short when considering optical beams with a nontrivial structure, such as Hermite-Gauss, Laguerre-Gauss, and other modes with a nontrivial structure. Here, we propose a new definition of the classical Shannon information via the Wigner distribution function, while respecting the Heisenberg inequality. Following this definition, we calculate the amount of information in Gaussian, Hermite-Gaussian, and Laguerre-Gaussian laser modes in juxtaposition and experimentally validate it by reconstruction of the Wigner distribution function from the intensity distribution of structured laser beams. We experimentally demonstrate the technique that allows to infer field structure of the laser beams in singular optics to assess the amount of contained information. Given the generality, this approach of defining information via analyzing the beam complexity is applicable to laser modes of any topology that can be described by well-behaved functions. Classical Shannon information, defined in this way, is detached from a particular alphabet, i.e., communication scheme, and scales with the structural complexity of the system. Such a synergy between the Wigner distribution function encompassing the information in both real and reciprocal space and information being a measure of disorder can contribute into future coherent detection algorithms and remote sensing.

1. Introduction

An electromagnetic field is a fundamental physical carrier of information. It is capable of reliably transmitting a modulated signal and collecting information about the propagation channel itself. With relevance to this IT-driven age, the two longstanding goals in information processing are (i) achieving higher channel capacity (i.e., throughput) and (ii) (pre)processing of collected information. However, the fundamental challenge of rigorous qualification and quantification of information in EM waves still remains a topic of debate including both physical and even semi-philosophical notions. Information theory stands out from most of other approaches in physics. Being a higher level of abstraction, it focuses on a configuration of a system under consideration in the context of its prehistory, similarly to thermodynamics, without attachment to a particular class of objects under study in its axiomatics. Out of the broad scope of studies where information theory has a potential to contribute, at this point, we exemplary explore its applications to electrical engineering and signal processing.

One way of defining information is associating it with the presence of *distinctive features*. For instance, human speech can carry up to 2^{14} distinctive sounds (due to 14 binary distinctive features, e.g., [1]), only a small subset of which is realized in any particular known language. The amount of information a human can transmit per unit sentence containing a fixed number of words is indirectly correlated with the amount of distinctive features the language can handle (i.e., language capacity). Translating this into

optics, it has been understood that a monochromatic plane-wave photon in free space can carry a rather limited amount of distinctive features (i.e., polarization and wavelength), providing up to one bit of information per photon. Naturally, if one generates photons with up to one bit of information, one is also able to detect only 1-bit information per unit carrier. To use an analogy, consider "a marine biologist casting a fishing net with two inches wide meshes for exploring the life on the ocean, naturally one should not be surprised finding only sea-creature larger than two inches long" [2]. This stumbling block has been shifted with the seminal work by Allen et al. [3] where it was experimentally confirmed that laser beams are capable of carrying a well-defined orbital angular momentum (OAM). An ability of such laser modes to carry theoretically unbounded amount of information per photon, e.g., [4], dramatically expands the EM-field's "language capacity."

There are several approaches to assess a signal's information capacity developed in modern information theory, e.g., [5, 6]. In many cases, information is defined with respect to a particular alphabet, giving up the generality offered by statistics in the foundation of information theory. Several groups used conditional information approach to quantify the signal capacity [7]. Here, we introduce the concept of expressing information as a measure of structure in a physical system by applying the Shannon information theory to singular optical beams. We discuss how the Wigner distribution function (WDF) can be taken as a corresponding probability distribution function accounting for partial *quantumness* of a shaped photon source. The important synergy between a comprehensive description of physical systems in their phase space, delivered by the WDF, and the generalized axiomatics of the information theory has the potential to conduce a cumulative approach to high-information-density telecommunications and adaptive signal processing techniques. We validate this theoretical framework by experimentally showing how increased structural complexity of wavefront-shaped optical beams, such as Hermite-Gauss (HG) modes and optical vortices [8, 9], can be analyzed using wavefront sensors.

2. Results

2.1. The WDF and Classical Information in Optics. The WDF belongs to the generalized Cohen's class of dual-domain distributions. It is simultaneously the most complete analytical description of an optical beam, and an observable that can be experimentally measured. It provides access to the spatial beam profile and its Fourier transform. The WDF in one spatial dimension can be defined as

$$W(x,k_x)=\frac{1}{2\pi}\int dy\, u\left(x+\frac{y}{2}\right)u^*\left(x-\frac{y}{2}\right)e^{-ik_xy}, \tag{1}$$

where x is the variable in the coordinate space and k is the corresponding coordinate in reciprocal space. The following integrals have a probabilistic interpretation:

$$\begin{aligned}|u(k)|^2&=\int dx W(x,k),\\ |u(x)|^2&=\int dk\, W(x,k),\\ U_{\text{tot}}&=\int dxdk\, W(x,k),\end{aligned} \tag{2}$$

where $|u(k)|^2$ is the momentum distribution, $|u(x)|^2$ is the intensity distribution, and U_{tot} is the total energy of the incoming signal. For a fully coherent light source, the WDF's Fourier-transformed function $\Gamma(x)=u(x+a)u^*(x-a)$ is known as the mutual intensity used in wavefront sensing for turbulence analysis and adaptive detection techniques. For brevity, a list of useful optical properties of the WDF can be found elsewhere, e.g., [10, 11].

Now, let us consider a Gaussian beam, expressed in the following form [8]:

$$u_G(\rho)=\frac{A}{\mathrm{w}}e^{-\left(\rho^2/\mathrm{w}^2\right)}\,e^{i\tilde{\mathbf{k}}\rho^2/2R}, \tag{3}$$

where $\rho>0$ is the position vector in the beam profile $\rho=\sqrt{x^2+y^2}$, $\mathrm{w}=\mathrm{w}(z)$ is a beam waist, $R=R(z)$ is the radius of curvature of the beam's wavefront, and A is the normalization constant. The wave-number $\tilde{\mathbf{k}}=\{\tilde{k}_x,\tilde{k}_y,\tilde{k}_z\}$ is distinctive from the Fourier transform parameter $k=\{k_x,k_y,k_z\}$ in (1). The corresponding 1D WDF is [12]

$$W^{(G)}(\tilde{x},\kappa_x)=\frac{A}{\sqrt{2\pi}}e^{-\tilde{x}^2-\kappa_x^2}, \tag{4}$$

which is plotted in position-momentum space $\{x,k_x\}$ in Figure 1(a). Here, $\bar{x}=\sqrt{2}x/\mathrm{w}$ and κ_x is the redefined momentum:

$$\kappa_x=\frac{\mathrm{w}}{\sqrt{2}}\left(k_x-\frac{\tilde{k}_x x}{R}\right). \tag{5}$$

This Wigner distribution is properly normalized, delivering the beam intensity distribution when integrated over the entire momentum space.

Classical Shannon information, see [13], represents the amount of structure in the corresponding system. Its analog for an optical mode, characterised by its WDF, is

$$S=-\iint_{R^2}d\mathbf{r}\,d\mathbf{k}\,W(\mathbf{r},\mathbf{k})\cdot\ln\left[W(\mathbf{r},\mathbf{k})\right], \tag{6}$$

where the WDF is a scalar function of position $\mathbf{r}$ and momentum $\mathbf{k}$ vectors. Similar definitions of information applied to characterizing optical fields, while having been suggested in the field of optics earlier, e.g., [12, 14], have not been explicitly applied to topological optical beams in the way introduced here, to the best of the authors' knowledge. The main problem with this definition, Equation (6), is that the WDF is not a positive-semidefinite function and, hence, does not represent

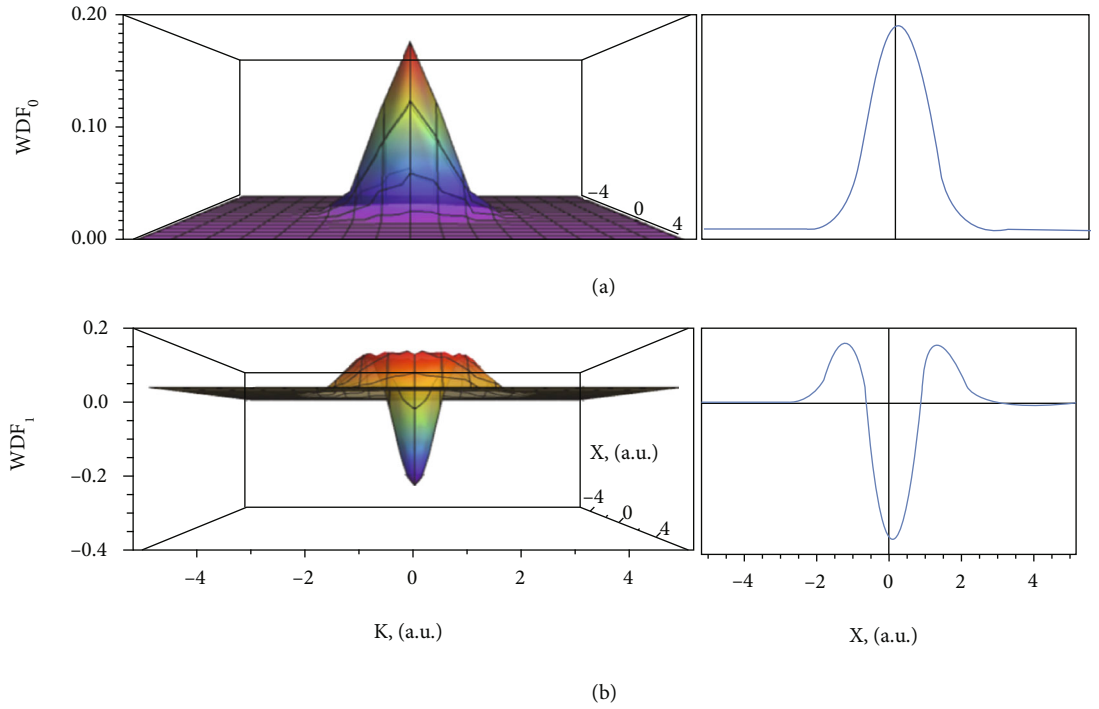


FIGURE 1: The Wigner distribution functions (WDFs) of one-dimensional (1D) Hermite-Gauss (HG) mode of zero-order HG_0 (Gaussian) and first-order HG_1 as functions of κ_x-wave vector (5), and $\bar{x}$-coordinate in the beam's transverse plane. One can see that in the case of the Gaussian mode, the WDF is positive, while for the first-order HG mode, there is a negative contribution in the near-zero region of the phase space.

a proper distribution. Negativity of the WDF can be interpreted as a marker of a phase-space interference and nonclassicality [15, 16]. However, in quantum description, interference is associated with a local violation of Heisenberg inequality, e.g., [17]. For these reasons, we seek a definition that would comply with the quantum nature of light.

Here, we propose to replace the regular product in the definition of information (6) with the Groenewold associative product [18].

$$\star = e^{i\hbar\left(\overleftarrow{\partial_x}\overrightarrow{\partial_p} - \overleftarrow{\partial_p}\overrightarrow{\partial_x}\right)/2}, \tag{7}$$

and hereby call it Shannon-Groenewold information:

$$\tilde{S} = -\frac{1}{(\pi\hbar)^2}\iint_{\mathbb{R}^2} d\mathbf{r}\, d\mathbf{k}\, W(\mathbf{r},\mathbf{k}) \star \ln\left[W(\mathbf{r},\mathbf{k})\right], \tag{8}$$

where $W(\mathbf{r}, k)$ is the four-dimensional (4D) WDF (see [19]). A manifestation of this definition is the Weyl quantization of a classical observable in phase space [20]. It respects the Heisenberg uncertainty principle and is normalized to the phase space volume, which is one of the main advantages of this definition of information over Shannon information Equation (6), defined via the WDF. As an entropic observable, it measures the structure present in the optical field analogous to [21], as opposed to the earlier attempts to utilize the conditional information [7], that is, by design, less fundamental due to being alphabet specific. In the context of Information Theory, it may be interpreted as a measure of the amount of *classical* information that can be extracted from the laser mode if all the quantum uncertainty is removed by an appropriate experiment.

Using the Shannon-Groenewold information (8), we obtain the following expression for the information in the Gaussian beam:

$$\tilde{S}_{1D} \simeq -U_{\text{tot}}\left[\ln\left(\frac{U_{\text{tot}}}{\pi}\right) - 1\right], \tag{9}$$

$$\tilde{S}_{2D} \simeq -U_{\text{tot}}^2\left[\ln\left(\frac{U_{\text{tot}}^2}{\pi^2}\right) - 2\right], \tag{10}$$

where $U_{\text{tot}} = A\sqrt{\pi/2}$ is the total EM energy of the beam per spatial degree of freedom. We assume that the radius of curvature of the beam wavefront is always larger than the physical dimensions of the beam spot size: $R \gg \{x, y\}$. It is important to note that for the Gaussian mode, both definitions (6) and (8) produce the same result. This is what one would expect since the WDF of pure Gaussian sources is positive-semidefinite and can be interpreted as a classical proper distribution function [22]. The fact that information is energy-dependent comes as no surprise if one remembers the important synergy between information and entropy in classical Shannon theory of information (see [13]). Even though entropy in information theory and thermodynamic entropy are not exactly identical, one would expect similar behavior when it comes to the fundamental laws of physics

[23, 24]. In this context, total energy of a laser mode is in equivalence with total internal energy of a thermodynamic system.

2.2. Higher-Order Modes and Shannon-Groenewold Information. The HG mode is a higher-order solution of the Gaussian family of beam-like solutions of the paraxial equation. It carries a nontrivial beam geometry with singularities, stable on propagation [25] due to being topologically protected. Hence, it is only logical to expect that the amount of information for this family of modes is higher than for the zero-order Gauss modes $\tilde{S}_G \lesssim \tilde{S}_{HG}$ [26]. Let us investigate this statement next.

The general solution of the 1D paraxial wave equation in cylindrical coordinates is given by

$$u_{HG}(x,z) = \sqrt{\frac{A}{\mathrm{w}}} H_m\left(\frac{\sqrt{2}x}{\mathrm{w}}\right) e^{-\left(x^2/\mathrm{w}^2\right)} e^{i\tilde{k}_x x^2/2R}. \tag{11}$$

The normalization is consistent with the one in Equation (3). The WDF for this mode is known, e.g., [27]. In our case, to keep the normalization consistent, we obtain

$$W_m^{(HG)}(x,k_x) = \frac{A\sqrt{\pi}}{(2\pi)^{3/2}} e^{-\bar{x}^2} \int_{-\infty}^{\infty} d\bar{y}\, e^{-i\kappa_x \tilde{y} - \left(\bar{y}^2/4\right)} H_m\left(\bar{x} + \frac{\bar{y}}{2}\right) H_m\left(\bar{x} - \frac{\bar{y}}{2}\right). \tag{12}$$

As expected, for $m = 0$, the WDF of HG mode is exactly equal to the Gaussian WDF (Equation (4)) and so is the corresponding information in Equation (9). Then, the first two higher-order WDFs have "elegant" analytical expressions:

$$W_1^{(HG)}(x,k_x) = \frac{2A}{\sqrt{2\pi}} e^{-\bar{x}^2-\kappa_x^2} H_1\left(\bar{x}^2 + \kappa_x^2 - \frac{1}{2}\right), \tag{13}$$

$$W_2^{(HG)}(x,k_x) = \frac{4A}{\sqrt{2\pi}} e^{-\bar{x}^2-\kappa_x^2} H_2\left(\bar{x}^2 + \kappa_x^2 - 1\right). \tag{14}$$

The WDF $W_1^{(HG)}(x,k_x)$ is plotted in Figure 1(b). One can see explicitly the negative contributions, in the 2nd-order mode in opposition to the positive-definite 0th-order Gauss beam. In a similar manner, one can work out higher-order modes using the integral form in Equation (12).

Next, we explore these WDF expressions further. We find that these WDFs take negative values in the central region of the phase space (Figure 1). The fundamental statement of classical information theory that "the more we know about a system's parameter space, the less is its uncertainty" is inevitably broken in the quantum context when considering correlated (conjugate) variables, such as position x and momentum p. If such "quantumness" is present in the corresponding PDF, it pushes the distribution into the negative domain [16] (Figure 1(b)). By respecting Weyl-Wigner quantization in the definition of information (8), we work around the WDF not being a well-defined PDF from the statistics point of view. Groenewold's product, introduced instead of conventional multiplication, enforces the negative regions to be integrated out, in compliance with the Heisenberg uncertainty relation. These lead to real and positive-valued, monotonously increasing with energy information:

$$\tilde{S}_{1(1\mathrm{D})}^{(HG)} \simeq -2U_{\mathrm{tot}}\left[\ln\left(\frac{2U_{\mathrm{tot}}}{\pi}\right) - 3\right], \tag{15}$$

$$\tilde{S}_{2(1\mathrm{D})}^{(HG)} \simeq -8U_{\mathrm{tot}}\left[\ln\left(\frac{4U_{\mathrm{tot}}}{\pi}\right) - 5\right]. \tag{16}$$

As HG modes form a complete orthonormal set, they can be used as an expansion basis [28]. Hence, this approach can be straightforwardly applied to OAM modes, such as Laguerre-Gauss (LG), e.g., [8]. Let us express the LG mode as follows:

$$u_{LG}(\rho,z) = \frac{A}{w}\left(\frac{\sqrt{2}\rho}{w}\right)^{|\ell|} L_p^{|\ell|}\left(\frac{2\rho^2}{w^2}\right) e^{-\left(\rho^2/w^2\right)} e^{ik\rho^2/2R} e^{i\ell\phi}, \tag{17}$$

where ℓ is the vorticity of the twisted mode and ϕ is the azimuthal angle in cylindrical coordinates, where the remaining parameters follow the definitions of Gauss (3) and HG (11) modes. Supplying the results from (12), we can express LG modes in terms of HG modes with a straightforward calculation; for instance, for the WDF of the 2D LG mode with $p = 0$ and $\ell = 1$,

$$W_{1(2\mathrm{D})}^{(\mathrm{LG})}(\bar{\rho};\kappa) = \frac{A^2}{2\pi} e^{-\bar{\rho}^2-\kappa^2}\left(\bar{\rho}^2 + \kappa^2 - 1\right), \tag{18}$$

where p and ℓ are correspondingly the order and degree numbers of the generalized Laguerre polynomial $L_p^{\ell}(\cdot)$; $\bar{\rho} = \sqrt{\bar{x}^2 + \bar{y}^2}$. The corresponding entropy can be calculated in a similar manner:

$$\tilde{S}_{1(2D)}^{(\mathrm{LG})} \simeq -2U_{\mathrm{tot}}^2\left(\ln\left(\frac{U_{\mathrm{tot}}^2}{\pi^2}\right) - 4\right). \tag{19}$$

The classical assessment of the amount of order in the optical mode is clearly increasing with the increasing complexity of the beam profile (see Figure 2), as one would expect from general considerations. It is important to recall that in the context of physical meaning of Shannon unconditional information, the WDF is normalised to U_{tot}—total energy. Consequently, constant A is bounded from above by $\sqrt{2/\pi}$.

The similarity sign in the expressions for information (structural complexity) of the considered above modes (9), (15), (16), and (19) is due to the presence of terms containing an odd logarithmic integral of the form:

$$\int_0^{\infty} dr\, r\, \ln\left(2r^2 - 1\right) I_0(\alpha r) \longrightarrow 0, \tag{20}$$

where I_0 is a modified Bessel function of the first kind. Numerical estimations show a tendency for these terms to go to zero; however, mathematically rigorous study of their absolute convergence has not been performed.

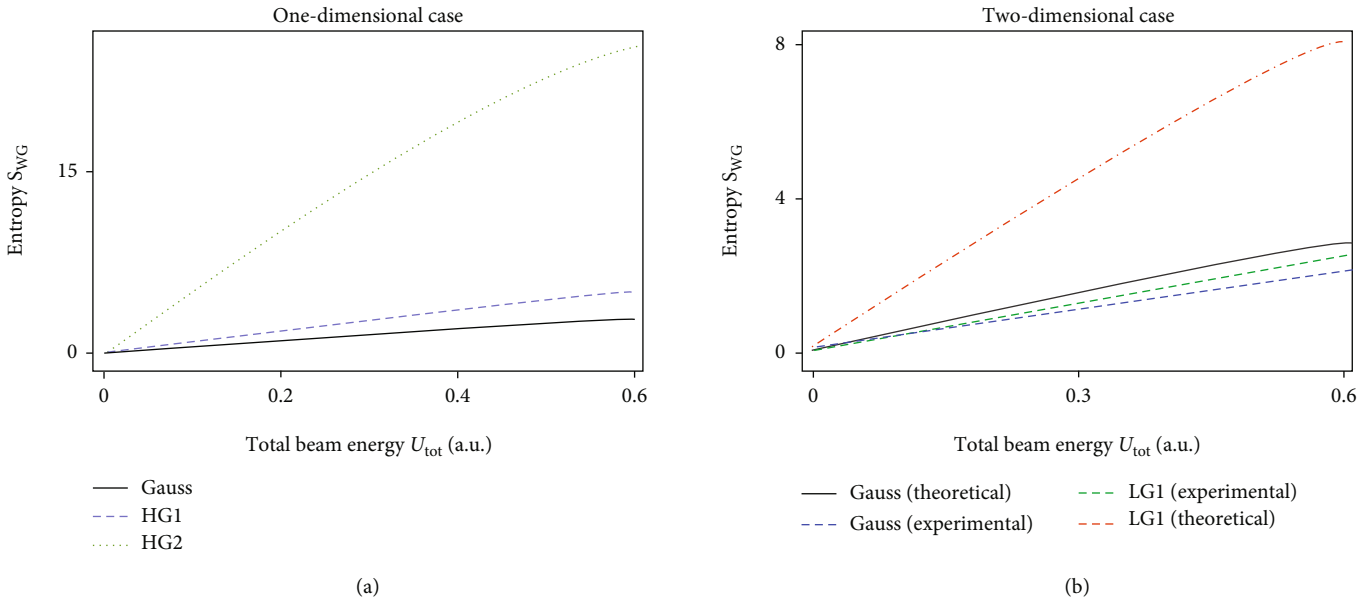


FIGURE 2: The Shannon-Groenewold information S_{WG} as a function of total energy U_{tot} (8) for one-dimensional (a) HG_0/Gauss mode (black), HG_1 (dashed-yellow), and HG_2 (dotted-green), and two-dimensional (b) theoretically predicted (black) and experimental (dashed-blue) Gauss mode, experimentally measured (long-dashed-green) and theoretically predicted (dot-dashed-red) LG_1 with the corresponding error bands resulting from the errors on fit parameters shown in light-red; see Section 3 for details. One notices the overall tendency for the amount of information to increase with the growth of the overall complexity of the corresponding optical signal. Also, the drop in the amount of information inferred from the experiment as compared to the theoretical curve is attributed to the SLM's beam conversion efficiency and expected information loss during the propagation in free space.

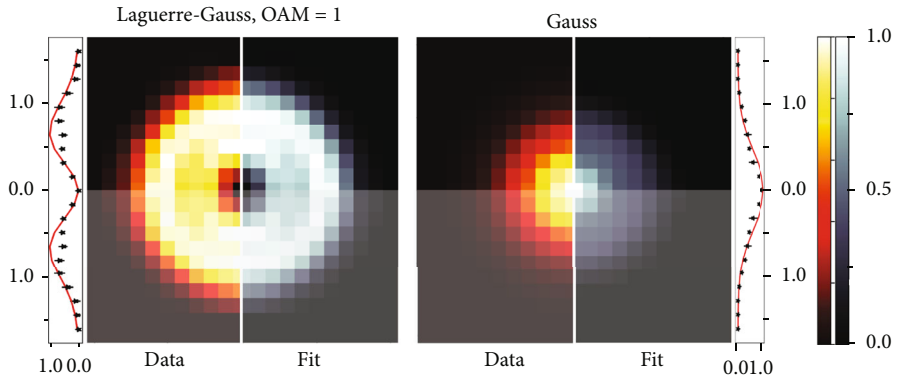


FIGURE 3: On the way to measuring the amount of information in optical beams. The intensity distributions of the EM beams are captured by a SHS. The algorithm for simulated data sets and the fits is based on Equations (21) and (22). The experimental data are averaged over 4 Laguerre-Gaussian LG_1 samples (left) and 10 Gaussian samples (right) with the SHS's sampling rate of 18 fps (black-starred scatter plot) and over the four quadrants in the beam intensity profile. In 2D color maps, measurement data is shown in gray scale, simulation is depicted in orange-red scale, and shaded regions are symmetry-based extrapolations. The measurement fits are shown in solid-red curves with the Gaussian model (right) using the WDF in Equation (4), and Laguerre-Gauss (LG) model (left) in Equation (18). The measured data are shown as black stars with the corresponding error bars. All the presented plots depict normalized intensity $I \in [0, 1]$.

With this theoretical framework, we performed an experiment to explore the possibility of obtaining the amount of information by measuring the structure in optical laser beams of various topology. Provided the wavefront reading was obtained from an off-the-shelf Shack-Hartmann sensor, we assessed the amount of information in several fundamental laser modes (see Figure 3).

3. Materials and Methods

Among the tools of adaptive optics, Shack-Hartman sensors (SHS) [29] occupy a unique place as a fast, affordable, and compact off-the-shelf tool for simultaneous intensity and angular distribution measurements. Advanced techniques for SHS state tomography [30] and WDF reconstruction [31]

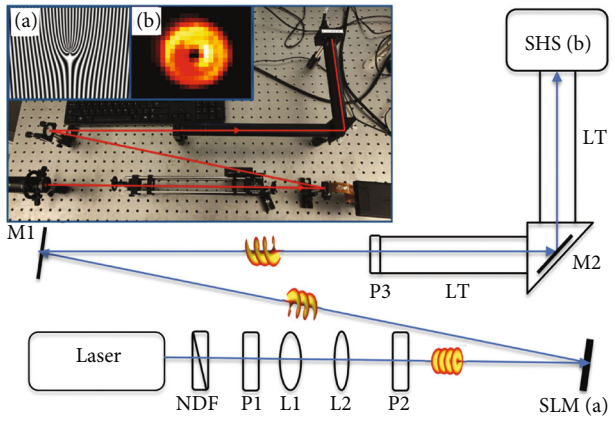


FIGURE 4: The schematic of the experimental setup, where a spatially filtered Gauss-like 635 nm, 4 mW (Thorlabs LDM635) laser beam is reflected off the spatial light modulator (SLM) onto the mirror M1, into the boxing of the Shack-Hartmann sensor (SHS) (Thorlabs: WFS20-7AR). The device is optimized to operate with ordinary Gauss-like signals. The SLM can be set to a mirror regime or to generate an orbital angular momentum (OAM) beam of Laguerre-Gauss-like profile. The OAM beam generation is accomplished by an SLM loaded with a computer-generated diffraction pattern with a fork dislocation (i). In that case, the measured intensity distribution of an OAM beam (ii) has a typical doughnut-like structure. L1 and L2 are the lenses of the beam expander; P1, P2, and P3 are the pinholes; M1 and M2 are the directing mirrors; and LT are the boxing elements of the SHS.

have been suggested alongside with conventional aberration correction techniques. Measurements of the wavefront distortions in EM beams with nontrivial topology are also of interest for both communication and sensing purposes, e.g., [32].

For the aim of this work, we use the SHS to reconstruct the WDF for the purpose of discriminating between the modes, assessing the beam quality and ultimately the amount of information in a beam. While a SHS is utilized generally to facilitate beam alignment and assess distortions within an optical channel by calculating the higher-order Zernike moments, here, we omit the SHS's wavefront data output and only consider the raw data of the intensity distribution in the superpixel array of the SHS's camera (Figure 4). We compare the results to the modelled intensity distribution based on the theoretical WDF calculation, whose approximation can be modelled as [31]

$$I(\mathbf{r}) = \frac{1}{\lambda f} \sum_{\substack{\ell=-L \\ m=-M}}^{L,M} \mathrm{SWDF}[W_b, W_a]\left(r'_{\ell,m}, u'_{\ell,m}\right) \mathrm{rect}\left(r'_{\ell,m}\right). \tag{21}$$

The smooth WDF is defined as follows [33]:

$$\mathrm{SWDF}[W_b, W_a]\left(r'_{\ell,m}, k'_{\ell,m}\right) = \iint d^2R\, d^2U\, W_b(R,U)\, W_a\left(R - r'_{\ell,m}, U - u'_{\ell,m}\right), \tag{22}$$

where the coordinate shift is defined as

$$r'_{\ell,m} = \{R_x - \ell w, R_y - mw\},$$
$$u'_{\ell,m} = \left\{U_x - \frac{x - \ell w}{\lambda f}, U_y - \frac{y - mw}{\lambda f}\right\}. \tag{23}$$

The functions W_b and W_a are the WDF of the incoming signal, e.g., (4), (13), (14) or (18), and the transmission function of a single lens aperture correspondingly. The parameters f (focal length) and w(width of a single lens in a lenslet array) are the parameters specific to the detector and define the angles in the local wavefront of the field.

Using this model, we first simulated the synthetic data sets for Gauss, HG, and LG modes. In the model, we considered the WDF W_b in the following general form:

$$W(\tilde{x}, \kappa_x) = \frac{A}{\sqrt{2\pi}} e^{-\tilde{x}^2 - \kappa_x^2} P\left(\tilde{x}^2 + k_x^2\right), \tag{24}$$

where $P(\alpha)$ is the polynomial of α. We started by testing two cases, namely, a Gauss mode (4) and a LG of order 1 (18). The polynomial fit in Equation (24) was taken to be

$$P\left(\tilde{x}^2 + k_x^2\right) = a + b\left(\tilde{x}^2 + \kappa_x^2\right), \tag{25}$$

with a and b being the fitting parameters (Figure 3). One can see that when $a = \text{const}$ and $b = 0$, the fit corresponds to a Gaussian profile (4), and when $a/b = 2$, the model includes LG_1-like distributions (18).

To assess the quality of the model, besides estimating the χ^2 per each fit, we run a simulation with 1000 fits to synthetic data with Equation (21), obtaining a histogram of the deviation between the supplied fitting parameters and those from the best fit (Figure 5(b)). The resulting data show that the parameters are centered near the "true" values, showing the satisfactory quality of the fit.

The inferred fundamental amount of information carried in an experimentally measured photon beam appears to be lower than the theoretically predicted for the ideal LG_1 mode (see Figure 2). The resulting entropy as a function of beam energy is extremely sensitive to the fit parameters a and b (25).

Since we did not consider the wavefront readings of the SHS, this model does not harvest the information stored in the reciprocal domain at this point, but rather infers it from the intensity distribution. This capability is to be explored and utilized in our future research. However, even at this conceptual level, the fit already can discriminate between the two modes. Hence, it provides an experimental estimate for the WDF (Figure 5(a)) supplied, however a "good" guess about the possible wavefront shape of the incoming signal. Due to intensity-only detection, the central region of the LG beam, generated in the experiment, is left out. Hence, the setup is inherently classical, in compliance with the definition of information, used here.

The quality of the fit, alongside with the accuracy of the numerical integration algorithm, also depends on the

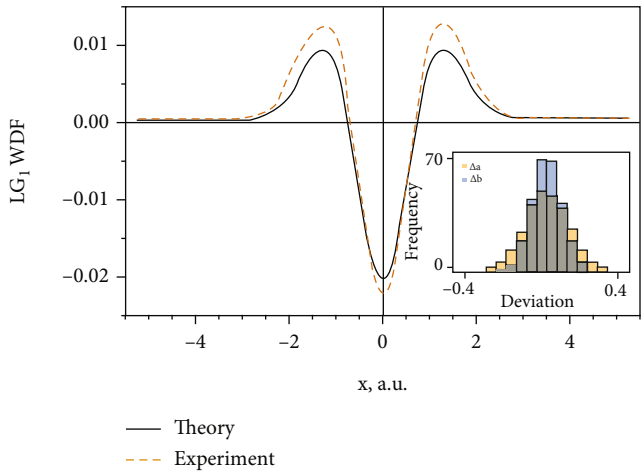


FIGURE 5: Comparing the experiment and the theoretical prediction for the WDF of the LG mode of order $\ell = 1$: (a) the WDF of the ideal LG_1 mode (solid black) and of the measured in an experiment laser beam (dashed yellow) with the corresponding error band; the frequency histogram of the deviation of the fitting parameters a (yellow) and b (blue) in the inset at the right-bottom, Equation (25), supplied to the model (21), and resulted from the fitting procedure: 1000 synthetic data sets have been generated and fitted with the model, mentioned above; for each run, the deviation between the supplied and fitted parameters has been calculated.

detector modelling scheme, (21) and (22), which in this case has been fairly generic. One of the crucial assumptions that has been made is the plane-wave approximation. This approximation, generally speaking, is too brave for the case of topological beams. Another unaccounted source of discrepancies is the optical cross-talk due to SHS's architecture that has been extensively discussed before on the level of mathematical modelling [31]. Hence, this model, thought already fruitful, has a great potential for improvement.

These results were used to assess the information stored in a physical channel and to compare them to theoretical curves (Figure 2). As for applications, the full 2D model can also provide information about the medium the beam interacted with that can be useful in remote sensing. Based on the results of Section 2.2, in the course of future research, we expect topological beams to outperform the modes with planar phase structure for two main reasons: (1) greater library of nontrivial signatures in the original beam profile; (2) reported robustness and self-healing properties of vortex modes.

4. Discussion

Interestingly, while information technologies have seen outstanding progress over the last century, which lead to the digital revolution and created flourishing businesses, the field of information theory has remained in a shade. We believe that a universal technique to assess both the quality and quantity of information in a received signal, if provided, could become a conceptually novel tool to physicists and engineers alike. The approach described in this work is by far not the first attempt, neither is it the most general. However, in this approach, classical information does not require an early choice of a communication scheme (i.e., alphabet). It is rather based on a fundamental assessment of an optical system's capability to carry information, based on its overall complexity. The WDF is uniquely used here as a probability density function for Shannon information in optics. The constraints of probability theory on the definition of information and of quantum mechanics on conjugate observables are satisfied working around the properties of the WDF—a *pseudo*-probability distribution. We foresee the relevance of this formalism in the context of recent developments for (i) free-space information-processing optics [34]; (ii) integrated photonics-based information processing [35] such as neural network-based accelerators [36] and photonic tensor cores [37]; (iii) adaptive sensing [38]; and (iv) analog optical and photonic processors [39–41]. As the data compression coefficient is naturally bounded by Shannon information, carried by the beam [42], this work indirectly points towards higher information capacity in beams with a nontrivial structure, like HG, LG, and Bessel-Gauss modes [43, 44].

Due to the WDF's relation to the EM-field correlation function, we foresee our approach to be extremely useful in adaptive optics. The reconstruction algorithm, when fully developed, has the potential to characterize the effects of decoherence in turbulent media, the 2D ambiguity function, and time-resolved frequency distribution, alongside with commonly available corrections for aberration, astigmatism, peak valley, and rms deformation provided by the SHS measurements. The WDF formalism uniquely gives access to such characteristics as mutual intensity of stochastic wave fields, which is of high importance when describing partially coherent sources.

In perspective, as the demand on high-speed data transfer and streaming grows exponentially, ADSL and fiber-to-home technologies are less and less likely to satisfy even an average consumer's data hunger, not to mention business and government agency calls. These, together with the recent advents in optical processing [34], micro- and nanofabrication [45], and OAM communications [46] put forward the mid-20th century's excitement around free-space communications in a new light. The new-generation free-space links will require coherent detection techniques to realise their potential to the fullest. For instance, the idea of using classical information as a quantifier for the amount of information processing that can be done on a given metasurface has already been implemented [47]. This approach elegantly gauges a specific pattern on the surface of a metastructure (lens) with the corresponding distribution of electromagnetic radiation in the far field [48]. With this progress, we believe that our approach may result in a better understanding of which types of measurements and device architectures are needed to efficiently mine information from a free-space link.

Data Availability

The data used in this research have been acquired manually in the OPEN Lab facility, the George Washington University, and can be shared by request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

M.S., J.Y., M.M., and V.J.S. acknowledge support from the Office of Naval Research (N00014-19-1-2595). A.A. acknowledges support from the U.S. Army Research Office Grant W911NF-19-1-0022.

Supplementary Materials

The authors provide supplementary information that offers details on some of the derivations and the theoretical approach presented in this paper. *(Supplementary Materials)*

References

[1] R. Jakobson, C. G. Fant, and M. Halle, *Preliminaries to Speech Analysis: the Distinctive Features and Their Correlates*, MIT press, 1951.

[2] A. S. Eddington, *The Philosophy of Physical Science: Tarner Lectures 1938*, CUP Archive, 1939.

[3] L. Allen, M. Beijersbergen, R. Spreeuw, and J. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Physical Review A*, vol. 45, no. 11, pp. 8185–8189, 1992.

[4] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the orbital angular momentum states of photons," *Nature*, vol. 412, no. 6844, pp. 313–316, 2001.

[5] T. M. Cover, *Elements of Information Theory*, John Wiley & Sons, 1999.

[6] V. Vedral, *Introduction to Quantum Information Science*, Oxford University Press on Demand, 2006.

[7] S. Restuccia, D. Giovannini, G. Gibson, and M. Padgett, "Comparing the information capacity of Laguerre-Gaussian and hermite-Gaussian modal sets in a finite-aperture system," *Optics express*, vol. 24, no. 24, pp. 27127–27136, 2016.

[8] E. J. Galvez, "Gaussian beams in the optics course," *American Journal of Physics*, vol. 74, no. 4, pp. 355–361, 2006.

[9] P. Vaity and L. Rusch, "Perfect vortex beam: Fourier transformation of a Bessel beam," *Optics Letters*, vol. 40, no. 4, pp. 597–600, 2015.

[10] M. J. Bastiaans, *Wigner distribution in optics*, Citeseer, 2009.

[11] T. Alieva, A. Cámara, and M. J. Bastiaans, *Wigner distribution moments for beam characterization*, CRC Press, 2012.

[12] M. J. Bastiaans, *Application of the Wigner distribution function in optics*, Elsevier Science, 1997.

[13] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[14] A. H. Nuttall, *Naval underwater system center new London lab*, 1990.

[15] D. Dragoman, "Applications of the Wigner distribution function in signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 10, 2005.

[16] A. Kenfack and K. yczkowski, "Negativity of the Wigner function as an indicator of nonclassicality," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, no. 10, pp. 396–404, 2004.

[17] T. Claasen and W. Mecklenbrauker, "Time-frequency signal analysis by means of the wigner distribution," in *ICASSP '81. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 69–72, Atlanta, GA,USA, 1981.

[18] H. J. Groenewold, "On the principles of elementary quantum mechanics," in *On the Principles of Elementary Quantum Mechanics*, pp. 1–56, Springer, Dordrecht, 1946.

[19] T. Curtright and C. Zachos, "Negative probability and uncertainty relations," *Modern Physics Letters A*, vol. 16, no. 37, pp. 2381–2385, 2001.

[20] S. W. McDonald, "Phase-space representations of wave equations with applications to the eikonal approximation for short-wavelength waves," *Physics Reports*, vol. 158, no. 6, pp. 337–416, 1988.

[21] A. I. Khinchin, "The entropy concept in probability theory," *Uspekhi Matematicheskikh Nauk*, vol. 8, pp. 3–20, 1953.

[22] D. Dragoman, "Phase-space interferences as the source of negative values of the Wigner distribution function," *Journal of the Optical Society of America A*, vol. 17, no. 12, pp. 2481–2485, 2000.

[23] Y. L. Klimontovich, "Information concerning the states of open systems," *Physica Scripta*, vol. 58, no. 6, pp. 549–555, 1998.

[24] M. P. Mueller, "Law without law: from observer states to physics via algorithmic information theory," *Quantum*, vol. 4, p. 301, 2020.

[25] G. Gbur, *Singular optics*, The Optics Encyclopedia, 2015.

[26] A. Kumar, S. Prabhakar, P. Vaity, and R. Singh, "Information content of optical vortex fields," *Optics Letters*, vol. 36, no. 7, pp. 1161–1163, 2011.

[27] R. Simon and G. Agarwal, "Wigner representation of Laguerre-Gaussian beams," *Optics Letters*, vol. 25, no. 18, pp. 1313–1315, 2000.

[28] I. Kimel and L. R. Elias, "Relations between hermite and Laguerre Gaussian modes," *IEEE Journal of Quantum Electronics*, vol. 29, no. 9, pp. 2562–2567, 1993.

[29] B. C. Platt and R. Shack, "History and principles of Shack-Hartmann wavefront sensing," *Journal of Refractive Surgery*, vol. 17, no. 5, pp. S573–S577, 2001.

[30] B. Stoklasa, L. Motka, J. Rehacek, Z. Hradil, and L. Sánchez-Soto, "Wavefront sensing reveals optical coherence," *Nature Communications*, vol. 5, no. 1, pp. 1–7, 2014.

[31] L. Tian, Z. Zhang, J. C. Petruccelli, and G. Barbastathis, "Wigner function measurement using a lenslet array," *Optics Express*, vol. 21, no. 9, pp. 10511–10525, 2013.

[32] S. Zhao, J. Leach, L. Gong, J. Ding, and B. Zheng, "Aberration corrections for free-space optical communications in atmosphere turbulence using orbital angular momentum states," *Optics Express*, vol. 20, no. 1, pp. 452–461, 2012.

[33] Z. Zhang and M. Levoy, "Wigner distributions and how they relate to the light field," in *2009 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–10, San Francisco, CA, USA, 2009.

[34] M. Miscuglio, Z. Hu, S. Li et al., "Massively parallel amplitude-only Fourier neural network," *Optica*, vol. 7, no. 12, pp. 1812–1819, 2020.

[35] T. Ferreira de Lima, A. N. Tait, A. Mehrabian et al., "Primer on silicon neuromorphic photonic processors: architecture and compiler," *Nanophotonics*, vol. 9, no. 13, pp. 4055–4073, 2020.

[36] J. Peng, Y. Alkabani, S. Sun, V. J. Sorger, and T. El-Ghazawi, "Dnnara: a deep neural network accelerator using residue

arithmetic and integrated photonics," in *49th International Conference on Parallel Processing-ICPP*, pp. 1–11, Edmonton, Canada, 2020.

[37] M. Miscuglio and V. J. Sorger, "Photonic tensor cores for machine learning," *Applied Physics Reviews*, vol. 7, no. 3, 2020.

[38] Y. Ren, G. Xie, H. Huang et al., "Adaptive-optics-based simultaneous pre- and post-turbulence compensation of multiple orbital-angular-momentum beams in a bidirectional free-space optical link," *Optica*, vol. 1, no. 6, pp. 376–382, 2014.

[39] J. Anderson, E. Kayraklioglu, S. Sun et al., "Roc: a reconfigurable optical computer for simulating physical processes," *ACM Transactions on Parallel Computing (TOPC)*, vol. 7, no. 1, pp. 1–29, 2020.

[40] S. Sun, M. Miscuglio, X. Ma et al., "Induced homomorphism: Kirchho's law in photonics," *Nanophotonics*, vol. 10, no. 6, pp. 1711–1721, 2020.

[41] M. Miscuglio, Y. Gui, X. Ma et al., "Analog computing with metatronic circuits," https://arxiv.org/abs/2007.05380.

[42] J. Preskill, *Lecture Notes for Physics 229: Quantum Information and Computation*, 1998, California Institute of Technology.

[43] G. J. Chaitin, "Information, randomness & incompleteness: papers on algorithmic information theory," *World Scientific*, vol. 8, 1990.

[44] F. E. Ruiz, P. S. Pérez, and B. I. Bonev, *Information Theory in Computer Vision and Pattern Recognition*, 2009, Springer Science & Business Media.

[45] Z. Ji, W. Liu, S. Krylyuk et al., "Photocurrent detection of the orbital angular momentum of light," *Science*, vol. 368, no. 6492, pp. 763–767, 2020.

[46] L. Li, H. Song, R. Zhang et al., "Increasing system tolerance to turbulence in a 100-gbit/s QPSK free-space optical link using both mode and space diversity," *Optics Communications*, vol. 480, article 126488, 2021.

[47] T.-J. Cui, S. Liu, and L.-L. Li, "Information entropy of coding metasurface," *Light: Science & Applications*, vol. 5, no. 11, article e16172, 2016.

[48] H. Wu, G. D. Bai, S. Liu et al., "Information theory of metasurfaces," *National Science Review*, vol. 7, no. 3, pp. 561–571, 2020.