Photon Information Efficient Communication Through Atmospheric Turbulence—Part II: Bounds on Ergodic Classical and Private Capacities

Nivedita Chandrasekaran, Jeffrey H. Shapiro, and Ligong Wang

***Abstract*—Vacuum-propagation optical communication with high photon efficiency (many bits/photon) *and* high spectral efficiency (many bits/s·Hz) requires operation in the near-field power transfer regime with a large number of spatial modes. For terrestrial propagation paths, however, the effects of atmospheric turbulence must be factored into the photon and spectral efficiency assessments. In Part I of this study [N. Chandrasekaran and J. H. Shapiro, "Photon Information Efficient Communication Through Atmospheric Turbulence—Part I: Channel Model and Propagation Statistics," *J. Lightw. Technol.*, vol. 32, no. 6, pp. 1075–1087, Mar. 2014], modal-transmissivity statistics were derived for the turbulent channel that depend solely on the mutual coherence function of the atmospheric Green's function, and these bounds were evaluated for ∼200 spatial-mode systems whose transmitters used either focused-beam (FB), Hermite-Gaussian (HG), or Laguerre–Gaussian (LG) modes and whose receivers either did or did not employ adaptive optics. This Part II paper derives upper and lower bounds for the ergodic Holevo capacities of classical and private information transmission over the multiple spatial-mode turbulent channel that can be evaluated from Part I's transmissivity statistics. Also included are bounds on the ergodic capacity for on–off keying encoding and direct detection. It is shown that: 1) adaptive optics are not necessary to realize high photon information efficiency and high spectral efficiency simultaneously; 2) an FB-mode system with perfect adaptive optics outperforms its HG-mode and LG-mode counterparts; and 3) the converse is true when adaptive optics are not employed.**

***Index Terms*—Atmospheric turbulence, ergodic capacity, free-space optical communications, photon efficiency, private capacity, spectral efficiency.**

I. Introduction

ACHIEVING a photon information efficiency (PIE) of 10 bits/detected-photon *and* a spectral efficiency (SE) of 5 bits/s·Hz is impossible with a single spatial-mode optical communication link [1]. This is so even though high PIE can be obtained with pulse-position modulation (PPM) and direct detection (DD) [2], and high SE can be achieved with quadrature amplitude modulation and coherent detection [3]. In particular, PPM's high PIE comes at the cost of low SE, and the PIE of coherent detection systems cannot exceed 3 bits/detected-photon. Achieving high PIE and high SE in the same vacuum-propagation system is possible, but it requires operation in the near-field power transfer regime with 100's to 1000's of high-transmissivity spatial modes [1],[1] depending on whether communication is near the ultimate Holevo limit (100's of modes needed) or the Shannon limit for on–off keying (OOK) with DD (1000's of modes needed). For practical transmitter and receiver pupils, such large numbers of high-transmissivity spatial modes dictate that paths lengths of 1 to 10 km may be typical, implying that operation will very likely be through the atmosphere, rather than through vacuum. It has long been understood that the presence of clouds or fog along the propagation path will preclude high data-rate (Gbps) free-space optical communication [7]. Thus, limiting our quest for high-rate, high-PIE, high-SE free-space optical communication to clear-weather conditions, the vacuum-propagation capacities from [1] must be modified to account for the effects of atmospheric turbulence, viz., the random refractive-index fluctuations associated with turbulent mixing of air parcels with ∼1 K temperature differences [8]. That task is the principal goal of our two-part study that consists of the propagation statistics, developed in Part I [9], and the capacity bounds reported herein.

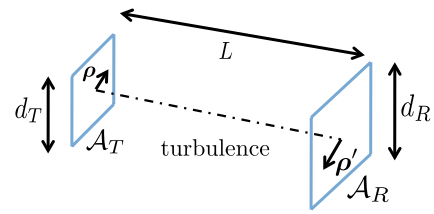

Fig. 1. Atmospheric-channel propagation geometry.

Vacuum propagation at wavelength λ over an L m line-of-sight path between coaxial square transmitter and receiver pupils of diameters d_T and d_R, respectively, has a near-field power transfer regime, determined by the condition $D_f \equiv (d_T d_R/\lambda L)^2 \gg 1$, with approximately D_f near-unity transmissivity spatial modes for each polarization [10]. When the same propagation geometry entails propagation through atmospheric turbulence, as shown in Fig. 1, there is still a near-field power transfer regime [11], but its modal eigenfunctions and power

Manuscript received July 14, 2013; revised November 11, 2013 and December 21, 2013; accepted December 23, 2013. Date of publication January 1, 2014; date of current version January 24, 2014. This work was supported by the NSF IGERT program Interdisciplinary Quantum Information Science and Engineering, the DARPA Information in a Photon Program, and the Office of Naval Research Basic Research Challenge Program.

The authors are with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: nivedita@mit.edu; jhs@mit.edu; wlg@mit.edu).

Digital Object Identifier 10.1109/JLT.2013.2296853

[1]What we choose to call the SE is termed the aggregate SE in research done on multi-core fiber communications [4]– [6].

transmissivities are, in general, stochastic. As opposed to the turbulent channel's far-field power transfer regime, for which there is an enormous corpus of propagation statistics, see, e.g., [12] and [13], there is a dearth of near-field results relevant to the determining the PIE versus SE tradeoff for free-space optical communication through turbulence.

In our Part I paper [9], we studied multiple spatial-mode propagation through atmospheric turbulence in the near-field power transfer regime using the channel geometry shown in Fig. 1. We assumed the transmitter employed a fixed set of M orthonormal spatial modes and we considered two receivers. One receiver used ideal (full-wave) adaptive optics to extract M orthonormal spatial modes spanning the vector space of spatial patterns that the transmitter modes create in the receiver's entrance pupil; the other extracted a fixed set of M orthonormal spatial modes from that pupil. We showed that the average modal transmissivities, for operation with and without adaptive optics, majorized the power-transfer eigenvalues of their respective average channel kernels. Because those eigenvalues can be found from the mutual coherence function for the atmospheric Green's function that appears in the extended Huygens–Fresnel principle, their behavior can be found in weak-perturbation through saturated-scintillation propagation conditions. In Part I we presented eigenvalue results for $\sim$200 spatial-mode systems that employed focused-beam (FB), Hermite–Gaussian (HG), or Laguerre–Gaussian (LG) modes. It is these results that provide the foundation for the present paper.

Our first task, in what follows, will be to develop upper and lower bounds on the ergodic Holevo capacities for classical and private information transmission of multiple spatial-mode systems for the turbulent channel that operate with or without adaptive optics. For both systems we shall assume that the receiver *and* the transmitter know the instantaneous channel matrix, and that they use this information to achieve their respective ergodic classical and private capacities. Our capacity bounds, derived in Section II for classical communication and in Section III for private communication, bracket the ultimate limits of multiple-spatial mode optical communication through turbulence. Approaching those limits, however, requires joint-detection receivers whose explicit implementations are as yet unknown, and are the subject of ongoing research [14], [15]. So, to set performance limits for a system we know how to build, Section II will also bound the multiple spatial-mode ergodic capacity for classical information transmission of a coherent-state (laser light) system that uses OOK and DD. In Section IV we instantiate the preceding bounds using the power-transmissivity statistics for FB, HG, and LG modes that we reported in [9]. As shown in that prior study, the unitary relationship between the HG and LG mode sets implies that their PIE versus SE behaviors are identical. Section IV will show that links using FB modes outperform those using HG and LG modes when ideal adaptive optics are employed, whereas the converse is true when adaptive optics are not employed. Section V concludes with a summary of our results and suggests avenues for future work.

Before embarking on the derivations of our capacity bounds, there is a point of comparison to be made. There is a wealth of published work on the capacity of multiple-input, multiple-output (MIMO) wireless communications that are related to what we will develop in that both are concerned with random fading, see [16] for early seminal results. Likewise, there have been studies of MIMO operation over optical fiber using coherent detection, see, e.g., [17]. Our study is distinguished from the literature in both such areas because we are concerned with quantum noise, not the additive Gaussian noise that they consider, and near-field propagation through turbulence results in a fading model—treated in our Part I paper—that is very different from those encountered in wireless and fiber MIMO communications.

II. Ergodic Capacity Bounds

In this section we will establish upper and lower bounds on the ergodic Holevo capacity and the OOK-DD capacity for transmitting classical information over a multiple spatial-mode bosonic (optical communication) channel with random modal power transmissivities. We begin our development with a review of the corresponding capacities for the single-mode pure-loss bosonic channel.

A. Channel Capacities for Single-Mode Pure-Loss Bosonic Channels

The pure-loss bosonic channel is one in which photons are lost en route from the transmitter (Alice) to the receiver (Bob), e.g., from diffraction, and the accompanying noise injection is at the minimum (vacuum-state) level necessary to preserve free-field commutator brackets [18]. Such a channel can be modeled as a beam splitter governed by the following Heisenberg evolution:

$$\hat{b} = \sqrt{\mu}\,\hat{a} + \sqrt{1-\mu}\,\hat{v} \tag{1}$$

where $\hat{a}$ is the annihilation operator for Alice's transmitted mode, $\hat{v}$ is the annihilation operator for a vacuum-state noise mode, μ (satisfying $0 \le \mu \le 1$) is the channel's transmissivity, and $\hat{b}$ is the annihilation operator for Bob's received mode. When the average photon number of Alice's transmitter is constrained to $\langle \hat{a}^\dagger \hat{a} \rangle \le N_T$, the Holevo capacity (in bits per channel use) of this channel is [18]–[20][2]

$$g_{\rm hol}(\mu N_T) = (\mu N_T + 1)\log_2(\mu N_T + 1) - \mu N_T \log_2(\mu N_T). \tag{2}$$

It is known that the pure-loss capacity is achieved with coherent-state encoding, and capacity-achieving polar codes have been found for this channel [14], but so far no explicit realizations are known for receivers that can approach this performance limit. Thus, for the pure-loss channel, we introduce the capacity of a coherent-state system that uses OOK modulation and DD under the same average photon-number constraint. The OOK-DD capacity of the pure-loss bosonic channel (in bits per channel use) is

$$g_{\rm ook}(\mu N_T) = \max_{0 \le p \le 1} f_{\rm ook}(\mu N_T, p) \tag{3}$$

[2]Note that [19], [20] predate the definition of Holevo capacity, and so do not deal with the issue of additivity, a point that [18] addressed.

where

$$f_{\rm ook}(x,p) \equiv H_B[p(1-e^{-x/p})] - pH_B(1-e^{-x/p}) \quad (4)$$

with p being the probability that an "on" pulse (of average photon number N_T/p) is sent, and

$$H_B(q) \equiv -q\log_2(q) - (1-q)\log_2(1-q) \quad (5)$$

is the binary entropy function (in bits). Unfortunately, $g_{\rm ook}(\cdot)$ does not have a closed form, which somewhat complicates bounding the ergodic classical capacity for its M spatial-mode extension. Thus, in Section II-B, we will need the following three lemmas about the behavior of $f_{\rm ook}(x,p)$, whose proofs are given in the Appendix, to derive bounds on OOK-DD's ergodic capacity.

Lemma A.1: For a fixed nonzero p, $f_{\rm ook}(x,p)$ is monotonically increasing in x.

Lemma A.2: For a fixed nonzero p, $f_{\rm ook}(x,p)$ is concave in x.

Lemma A.3: When evaluating

$$\max_{\substack{N_1,N_2\\ N_1+N_2=N_T\\ p_1,p_2:\,0\le p_1,p_2\le 1}} \sum_{m=1}^{2} \mu_m f_{\rm ook}(N_m,p_m) \quad (6)$$

the condition $\mu_1 \ge \mu_2$ implies that $N_1^* \ge N_2^*$, where N_1^* and N_2^* achieve the maximum in (6).

B. Multiple-Mode Ergodic Capacity Bounds

The capacity, in bits per channel use, of a deterministic, pure-loss, M spatial-mode optical channel whose transmitter is power-limited and therefore constrained to use at most N_T photons on average is[3]

$$C = \max_{\boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g(\mu_m N_m) \quad (7)$$

where $\boldsymbol{N} = [\,N_1 \quad N_2 \quad \cdots \quad N_M\,]$ is the vector of average photon numbers used for the M modes,[4] the $\{\mu_m\}$ are the modal transmissivities (arranged in nonincreasing order), and $g(\mu N)$ is the channel capacity—Holevo capacity if $g(\cdot) = g_{\rm hol}(\cdot)$, and OOK-DD capacity if $g(\cdot) = g_{\rm ook}(\cdot)$—of a single mode channel with transmissivity μ and average photon-number N. We are interested in the capacities of M spatial-mode communication through the turbulent channel when the transmitter and receiver have and use complete knowledge of the instantaneous channel matrix subject to the constraint that the transmitter employs a fixed set of spatial modes. Here the relevant quantity is the ergodic capacity

$$\langle C\rangle = \left\langle \max_{\boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g\left(\mu_m N_m\right) \right\rangle, \quad (8)$$

[3]We are assuming quasimonochromatic light, for which an average power constraint is equivalent to an average photon number constraint. This constraint is necessary because without it the capacity would be infinite.

[4]The $\{N_m\}$ and N_T need not take integer values, because they are *average* photon numbers.

where throughout this section $\mu_m = \mu_m^{\rm ad}$ ($\mu_m^{\rm non}$) for the system that does (does not) employ adaptive optics.

In Part I, we showed that the turbulent channel's modal transmissivities, $\{\mu_m\}$, are the power-transfer eigenvalues of propagation kernels that depend on the transmitter mode set, the instantaneous atmospheric Green's function, and whether adaptive optics are or are not employed at the receiver. However, their full statistics in the near-field power transfer regime are not available. Indeed, even obtaining their average values, $\{\langle\mu_m\rangle\}$, would appear to require Monte Carlo simulation. Therefore, in Part I we focused on the power-transfer eigenvalues, $\{\gamma_m\}$ arranged in nonincreasing order, of the average propagation kernels for operation with and without adaptive optics. Their evaluation only requires the mutual coherence function of the atmospheric Green's function, and they are majorized by the $\{\langle\mu_m\rangle\}$. Thus, here in Part II, we will first derive capacity bounds in terms of the $\{\mu_m\}$, and then use majorization to weaken them to bounds in terms of $\{\gamma_m\}$ that can be evaluated.

Upper bounds on the ergodic Holevo and OOK-DD capacities are easily obtained. If the single-mode capacities $g_{\rm hol}(\mu N)$ and $g_{\rm ook}(\mu N)$ are monotonically increasing in their eigenvalue argument μ, then setting all M power-transfer eigenvalues to unity gives us the multiple spatial-mode upper bound

$$\langle C\rangle \le M g(N_T/M), \quad (9)$$

which applies to adaptive and non-adaptive receivers. That $g_{\rm hol}(\mu N)$ is monotonically increasing with increasing μ is simple to demonstrate: its derivative with respect to μ is $N\log_2(1+1/\mu N)$ which is positive for $\mu N > 0$. Showing that $g_{\rm ook}(\mu N)$ is increasing in μ is best accomplished by considering $g_{\rm ook}(\mu_1 N) - g_{\rm ook}(\mu_2, N)$ for $\mu_1 \ge \mu_2$. Let p_1^* and p_2^* be the p values that yield $g_{\rm ook}(\mu_1 N)$ and $g_{\rm ook}(\mu_2 N)$, respectively, in Eq. (3). We then have that

$$\begin{aligned} g_{\rm ook}(\mu_1 N) - g_{\rm ook}(\mu_2 N) &= f_{\rm ook}(\mu_1 N, p_1^*) - f_{\rm ook}(\mu_2 N, p_2^*) \\ &\ge f_{\rm ook}(\mu_1 N, p_2^*) - f_{\rm ook}(\mu_2 N, p_2^*) \\ &\ge 0 \end{aligned} \quad (10)$$

where the second inequality follows from Lemma A.1. Therefore, $\langle C\rangle_{\rm hol} \le M g_{\rm hol}(N_T/M)$ and $\langle C\rangle_{\rm ook} \le M g_{\rm ook}(N_T/M)$. We do not, however, expect these bounds to be tight, especially for the non-adaptive receiver, so we turn our attention now to the more difficult—and more interesting—task of deriving lower bounds on the ergodic Holevo and OOK-DD capacities for adaptive and non-adaptive operation.

The maximum of an average cannot exceed the average of a maximum. So, we have

$$\langle C\rangle \ge \max_{\boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} \langle g\left(\mu_m N_m\right)\rangle. \quad (11)$$

For both Holevo and OOK-DD capacities, we claim that

$$\langle C\rangle \ge \max_{\boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} \langle\mu_m\rangle g(N_m) \quad (12)$$

i.e., these capacities can be bounded from below using knowledge of the $\{\langle\mu_m\rangle\}$. This is the first step toward bounds that can

be evaluated using the $\{\gamma_m\}$, whose FB-mode, HG-mode, and LG-mode values we found in Part I.

Proving (12) for $\langle C\rangle_{\rm hol}$ is simple. The single-mode Holevo capacity $g_{\rm hol}(\mu N)$ is concave in μ for fixed N: its second derivative with respect to μ is $-N/\mu(\mu N+1)\ln(2)$, which is negative for $\mu, N > 0$. Coupled with the fact that $g_{\rm hol}(0) = 0$, it follows that $\langle g_{\rm hol}(\mu N)\rangle \geq \langle\mu\rangle g_{\rm hol}(N)$, which allows us to arrive at (12) for $\langle C\rangle_{\rm hol}$.

Proving (12) for $\langle C\rangle_{\rm ook}$ requires more care and is done with the following argument:

$$\begin{aligned}\langle C\rangle_{\rm ook} &\geq \max_{\boldsymbol{N}:\sum_{m=1}^{M} N_m = N_T} \times \sum_{m=1}^{M}\left\langle \max_{p_m:0\leq p_m\leq 1} f_{\rm ook}\left(\mu_m N_m, p_m\right)\right\rangle \quad (13)\\ &\geq \max_{\substack{\boldsymbol{p}:\,\boldsymbol{0}\leq\boldsymbol{p}\leq\boldsymbol{1}\\ \boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T}} \sum_{m=1}^{M}\left\langle f_{\rm ook}\left(\mu_m N_m, p_m\right)\right\rangle \quad (14)\\ &\geq \max_{\substack{\boldsymbol{p}:\,\boldsymbol{0}\leq\boldsymbol{p}\leq\boldsymbol{1}\\ \boldsymbol{N}:\,\sum_{m=1}^{M} N_m = N_T}} \sum_{m=1}^{M}\langle\mu_m\rangle f_{\rm ook}\left(N_m, p_m\right) \quad (15)\\ &= \max_{\boldsymbol{N}:\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M}\langle\mu_m\rangle g_{\rm ook}\left(N_m\right) \quad (16)\end{aligned}$$

where $\boldsymbol{p}:\boldsymbol{0}\leq\boldsymbol{p}\leq\boldsymbol{1}$ denotes the probability vector $[\,p_1 \quad p_2 \quad \cdots \quad p_M\,]$. Inequality (14) follows because the maximum of an average cannot exceed the average of a maximum. Inequality (15) follows from the concavity of $f_{\rm ook}\left(\mu N, p\right)$ in μ for fixed N and p, which is a consequence of Lemma A.2, and the fact that $f_{\rm ook}\left(0, p\right) = 0$.

Our next task is to use majorization to weaken the lower bound in (12) to

$$\langle C\rangle \geq \max_{\boldsymbol{N}:\sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M}\gamma_m g(N_m) \quad (17)$$

which we will be able to evaluate with the $\{\gamma_m\}$ results from the Part I paper. To prove this result, it suffices to demonstrate that, for $\boldsymbol{N}$ achieving the maximum in (17),

$$\sum_{m=1}^{M}\left(\langle\mu_m\rangle - \gamma_m\right) g\left(N_m\right) \geq 0. \quad (18)$$

Rearranging terms in (18) we can write

$$\begin{aligned}&\sum_{m=1}^{M}(\langle\mu_m\rangle - \gamma_m) g\left(N_m\right)\\ &= \left[\sum_{m=1}^{M}(\langle\mu_m\rangle - \gamma_m)\right] g(N_M)\\ &\quad + \left[\sum_{m=1}^{M-1}(\langle\mu_m\rangle - \gamma_m)\right][g(N_{M-1}) - g(N_M)]\\ &\quad + \cdots + \left[\sum_{m=1}^{M-k}(\langle\mu_m\rangle - \gamma_m)\right][g(N_{M-k}) - g(N_{M-k+1})]\\ &\quad + \cdots + (\langle\mu_1\rangle - \gamma_1)[g(N_1) - g(N_2)]. \quad (19)\end{aligned}$$

With this rearrangement, proving (18) is quite simple for the ergodic Holevo capacity. We know that $g_{\rm hol}(\cdot)$ is a monotonically increasing, non-negative function of its argument. Lagrange-multiplier optimization yields the choice of $\boldsymbol{N}$, which we shall call $\boldsymbol{N}^*$, that achieves the bound given in (17) for $\langle C\rangle_{\rm hol}$. The $\{N_m^*\}$ are

$$N_m^* = \frac{1}{\exp(\beta/\gamma_m) - 1} \quad (20)$$

where β is chosen so that $\sum_{m=1}^{M} N_m^* = N_T$, and $\gamma_m = \gamma_m^{\rm ad}$ ($\gamma_m^{\rm non}$) when the receiver uses (does not use) adaptive optics. Because the $\{\gamma_m\}$ are nonincreasing with increasing m, the elements of $\boldsymbol{N}^*$ also have this property. Therefore, in (19), we have $g_{\rm hol}(N_M) \geq 0$ and $g_{\rm hol}(N_{M-k}) - g_{\rm hol}(N_{M-k+1}) \geq 0$ for $1 \leq k \leq M-1$. The inequality in (18) then follows immediately from the $\{\gamma_m\}$ being majorized by the $\{\langle\mu_m\rangle\}$, and so we obtain the lower bounds on the ergodic Holevo capacities for adaptive and non-adaptive operation that can be computed from available knowledge of the atmospheric Green's function's mutual coherence function.

Proving (17) for an M-spatial mode OOK-DD system is accomplished with Lemma A.3 and a proof by contradiction. Suppose that the $\boldsymbol{N}$ which maximizes the bound in (17) has $N_i > N_j$ for some $i > j$ for which, by our eigenvalue ordering, we know that $\gamma_i \leq \gamma_j$. Lemma A.3 implies that

$$\max_{N_i', N_j' \geq 0, N_i' + N_j' = N_i + N_j} \left[\gamma_i g_{\rm ook}(N_i') + \gamma_j g_{\rm ook}(N_j')\right] \quad (21)$$

occurs for $N_i' \leq N_j'$, contradicting our supposition that N_i, N_j are elements of the optimum $\boldsymbol{N}$. Because this argument applies for all $i > j$, the $\boldsymbol{N}$ that maximizes the bound in (17) must have nonincreasing N_m with increasing m. Inequality (19) then completes the proof of (17) for $\langle C\rangle_{\rm ook}$.

III. Ergodic Private Capacity Bounds

To establish bounds on the turbulent channel's private capacity, we begin in Section III-A with a review of the single-mode private capacity for the pure-loss channel.

A. Private Capacity for the Single-Mode Pure-Loss Bosonic Wiretap Channel

The pure-loss bosonic wiretap channel between transmitter Alice, intended receiver Bob, and eavesdropper Eve can be modeled as a beam splitter with transmissivity μ and reflectivity $1-\mu$ in which Bob observes the annihilation operator $\hat{b}$ of the transmitted light and Eve gets the annihilation operator $\hat{e}$ of the reflected light:

$$\hat{b} = \sqrt{\mu}\,\hat{a} + \sqrt{1-\mu}\,\hat{v} \quad (22)$$

$$\hat{e} = \sqrt{1-\mu}\,\hat{a} - \sqrt{\mu}\,\hat{v} \quad (23)$$

where $\hat{a}$ is Alice's annihilation operator and $\hat{v}$ is the annihilation operator for a vacuum-state noise mode. Note that this channel model is a worst-case scenario in that Eve receives all the photons that do not reach Bob. Smith [21] proved that the private capacity of the pure-loss bosonic wiretap channel equals that channel's quantum capacity, and Wolf *et al.* [22] derived that quantum capacity. From these works we have that Holevo private capacity for our wiretap channel, when Alice's transmitter is constrained to $\langle \hat{a}^\dagger \hat{a} \rangle \leq N_T$, is

$$g_{\text{hol}}^{\text{priv}}(\mu, N_T) = \begin{cases} g_{\text{hol}}(\mu N_T) - g_{\text{hol}}((1-\mu)N_T), & \mu > \dfrac{1}{2} \\ 0, & \mu \leq \dfrac{1}{2} \end{cases} \tag{24}$$

and it is achieved with coherent-state encoding.

B. Multiple-Mode Private Channel Capacity Bounds

The channel model for the M-mode, pure-loss bosonic wiretap channel—in which Alice's, Bob's, and Eve's modes have column vectors of annihilation operators $\hat{\boldsymbol{a}} = [\,\hat{a}_1 \quad \hat{a}_2 \quad \cdots \quad \hat{a}_M\,]^T$, $\hat{\boldsymbol{b}} = [\,\hat{b}_1 \quad \hat{b}_2 \quad \cdots \quad \hat{b}_K\,]^T$, and $\hat{\boldsymbol{e}} = [\,\hat{e}_1 \quad \hat{e}_2 \quad \cdots \quad \hat{e}_L\,]^T$, respectively, and the annihilation operators of the required vacuum-state noise modes are $\hat{\boldsymbol{v}} = [\,\hat{v}_1 \quad \hat{v}_2 \quad \cdots \quad \hat{v}_{K+L-M}\,]^T$—can be written as a multiple-mode beam splitter relation:

$$\begin{bmatrix} \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{e}} \end{bmatrix} = \mathbf{H} \begin{bmatrix} \hat{\boldsymbol{a}} \\ \hat{\boldsymbol{v}} \end{bmatrix}. \tag{25}$$

Here, the channel matrix $\mathbf{H}$ is a $(K+L) \times (K+L)$ unitary matrix that has the block form

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{ab} & \mathbf{H}_{vb} \\ \mathbf{H}_{ae} & \mathbf{H}_{ve} \end{bmatrix} \tag{26}$$

where the submatrix $\mathbf{H}_{ab}$ holds the couplings from Alice's modes to Bob's, and similarly for the other submatrices.

Before addressing the turbulent channel, we will first prove that the private capacity for the channel governed by Eq. (25) equals the private capacity of M independent pure-loss bosonic wiretap channels, whose transmissivities are the eigenvalues of $\mathbf{H}_{ab}^\dagger \mathbf{H}_{ab}$. We begin by noting that $\mathbf{H}^\dagger \mathbf{H}$ equals the $(K+L) \times (K+L)$ identity matrix, $\mathbf{1}_{(K+L)}$ so that $\mathbf{H}_{ab}^\dagger \mathbf{H}_{ab} + \mathbf{H}_{ae}^\dagger \mathbf{H}_{ae} = \mathbf{1}_M$. Therefore, $\mathbf{H}_{ab}$ and $\mathbf{H}_{ae}$ are simultaneously diagonalizable. Thus, there is an $M \times M$ unitary matrix, $\mathbf{V}$, and a diagonal matrix of eigenvalues, $\boldsymbol{\Lambda}_{ab} = \text{diag}(\mu_1, \mu_2, \ldots, \mu_M)$, such that

$$\mathbf{V}^\dagger \mathbf{H}_{ab}^\dagger \mathbf{H}_{ab} \mathbf{V} = \boldsymbol{\Lambda}_{ab} \tag{27}$$

$$\mathbf{V}^\dagger \mathbf{H}_{ae}^\dagger \mathbf{H}_{ae} \mathbf{V} = \boldsymbol{\Lambda}_{ae} = \mathbf{1}_M - \boldsymbol{\Lambda}_{ab}. \tag{28}$$

Without loss of generality, we will assume that $\mathbf{V}$ has been chosen such that the $\{\mu_m\}$ are in nonincreasing order.

Equations (27) and (28) lead to the singular-value decompositions

$$\mathbf{H}_{ab} = \mathbf{U}_{ab} \mathbf{S}_{ab} \mathbf{V}^\dagger \tag{29}$$

$$\mathbf{H}_{ae} = \mathbf{U}_{ae} \mathbf{S}_{ae} \mathbf{V}^\dagger \tag{30}$$

where $\mathbf{U}_{ab}$ and $\mathbf{U}_{ae}$ are $K \times K$ and $L \times L$ unitary matrices. The $K \times M$ and $L \times M$ matrices $\mathbf{S}_{ab}$ and $\mathbf{S}_{ae}$ are diagonal, and their nonzero elements are the nonzero elements of $\{\sqrt{\mu_1}, \sqrt{\mu_2}, \ldots, \sqrt{\mu_M}\}$ and $\{\sqrt{1-\mu_1}, \sqrt{1-\mu_2}, \ldots, \sqrt{1-\mu_M}\}$, respectively. These decompositions show that intermodal interference in the Eq. (25) channel model can be totally eliminated by Alice's applying $\mathbf{V}$—which can be realized as an $M \times M$ beam splitter—to her signal modes before transmitting them, and Bob and Eve's applying $\mathbf{U}_{ab}^\dagger$ and $\mathbf{U}_{ae}^\dagger$, respectively, to the modes they receive from Alice. This interference elimination proves that the private capacity of the multiple-mode channel specified by (25) equals the private capacity of M independent, single-mode pure-loss wiretap channels whose transmissivities $\boldsymbol{\mu} = [\,\mu_1 \quad \mu_2 \quad \cdots \quad \mu_M\,]$ are the eigenvalues of the channel kernel $\mathbf{H}_{ab}^\dagger \mathbf{H}_{ab}$.

The instantaneous private capacity C_{priv} of the M independent channels is additive, because each individual channel is stochastically degraded [21]. Therefore, implementing separate encoding and decoding schemes for each independent channel is optimal. Under the average photon-number constraint $\sum_{m=1}^{M} \langle \hat{a}_m^\dagger \hat{a}_m \rangle \leq N_T$, we have that the instantaneous private capacity of the Eq. (25) channel is

$$C_{\text{priv}}(\boldsymbol{\mu}, N_T) = \max_{\boldsymbol{N}: \sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g_{\text{hol}}^{\text{priv}}(\mu_m, N_m) \tag{31}$$

assuming that Alice and Bob both know the channel matrix $\mathbf{H}_{ab}$.

Now, if we allow the channel matrix $\mathbf{H}_{ab}$ to be random, as occurs for propagation through atmospheric turbulence, but assume that Alice and Bob have and use complete knowledge of that matrix, we can say that their ergodic Holevo private capacity is

$$\langle C_{\text{priv}}(\boldsymbol{\mu}, N_T) \rangle = \left\langle \max_{\boldsymbol{N}: \sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g_{\text{hol}}^{\text{priv}}(\mu_m, N_m) \right\rangle. \tag{32}$$

To link up with the propagation statistics from the Part I paper, we need to bound this ergodic private capacity using the $\{\gamma_m\}$, i.e., the eigenvalues of the average propagation kernel, $\langle \mathbf{H}_{ab}^\dagger \mathbf{H}_{ab} \rangle$. The following lemmas, which are proved in the Appendix, will do the trick.

Lemma A.4: For a fixed $N > 0$, the single-mode private capacity $g_{\text{hol}}^{\text{priv}}(\mu, N)$ is convex in μ.

Lemma A.5: For a fixed $N_T > 0$, the multi-mode private capacity $C_{\text{priv}}(\boldsymbol{\mu}, N_T)$ is both convex and Schur-convex in $\boldsymbol{\mu}$.

We establish an upper bound on $\langle C_{\text{priv}}(\boldsymbol{\mu}, N_T) \rangle$ in the following manner. Applying Lemma A.4, together with $g_{\text{hol}}^{\text{priv}}(0, N_m) = 0$, we have that

$$g_{\text{hol}}^{\text{priv}}(\mu_m, N_m) \leq \mu_m g_{\text{hol}}^{\text{priv}}(1, N_m) = \mu_m g_{\text{hol}}(N_m). \tag{33}$$

Thus, the instantaneous private capacity is bounded from above as follows,

$$C_{\text{priv}}(\boldsymbol{\mu}, N) \leq \sum_{m=1}^{M} \mu_m g_{\text{hol}}(N_m) \tag{34}$$

$$\leq \left(\sum_{m=1}^{M} \mu_m\right) g_{\rm hol}\left(\frac{\sum_{m=1}^{M} \mu_m N_m}{\sum_{m=1}^{M} \mu_m}\right) \quad (35)$$

$$\leq \tau g_{\rm hol}(N_T/\tau) \quad (36)$$

where $\tau \equiv \mathrm{tr}(\mathrm{H}_{\rm ab}^{\dagger}\mathrm{H}_{\rm ab}) = \sum_{\rm m=1}^{\rm M} \mu_{\rm m}$. Inequality (35) follows from the concavity of $g_{\rm hol}(\cdot)$, after which (36) follows from the fact that $g_{\rm hol}(\cdot)$ is monotonically increasing. To convert this result into an upper bound on $\langle C_{\rm priv}\rangle$, we note that $\tau g_{\rm hol}(N_T/\tau)$ is concave in τ, hence

$$\langle C_{\rm priv}(\boldsymbol{\mu}, N_T)\rangle \leq \tau_{\rm avg} g_{\rm hol}(N_T/\tau_{\rm avg}) \quad (37)$$

with

$$\tau_{\rm avg} \equiv \mathrm{tr}\Big(\langle \mathrm{H}_{\rm ab}^{\dagger}\mathrm{H}_{\rm ab}\rangle\Big) = \sum_{\rm m=1}^{\rm M} \gamma_{\rm m}. \quad (38)$$

To establish a lower bound on $\langle C_{\rm priv}\rangle$, we note that the eigenvalues, $\boldsymbol{\mu}$, of $\mathbf{H}_{ab}^{\dagger}\mathbf{H}_{ab}$ majorize the diagonal elements, $\boldsymbol{\lambda} = [\,\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_M\,]$, of that matrix when both are arranged in nonincreasing order. Combining this fact with Lemma A.5, we have that

$$\langle C_{\rm priv}(\boldsymbol{\mu}, N_T)\rangle \geq \left\langle \max_{\boldsymbol{N}:\, \sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g_{\rm hol}^{\rm priv}\,(\lambda_m\,, N_m)\right\rangle. \quad (39)$$

Now, because $C_{\rm priv}(\boldsymbol{\mu}, N)$ is convex in $\boldsymbol{\mu}$, we get another lower bound by replacing $\boldsymbol{\lambda}$ with the diagonal elements $\boldsymbol{\kappa} = [\,\kappa_1 \quad \kappa_2 \quad \cdots \quad \kappa_M\,]$ of the average channel kernel $\langle \mathbf{H}_{ab}^{\dagger}\mathbf{H}_{ab}\rangle$:

$$\langle C_{\rm priv}(\boldsymbol{\mu}, N_T)\rangle \geq \max_{\boldsymbol{N}:\, \sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g_{\rm hol}^{\rm priv}\,(\kappa_m\,, N_m)\,. \quad (40)$$

As noted below Eqs. (29) and (30), Alice and Bob can employ basis rotations on their modes without affecting their private capacity. Hence we can get the tightest lower bound obtainable from (40), by choosing, for each atmospheric state, bases that make $\boldsymbol{\lambda}$ the eigenvalues of $\mathbf{H}_{ab}^{\dagger}\mathbf{H}_{ab}$. This choice makes $\boldsymbol{\kappa}$ the eigenvalues, $\boldsymbol{\gamma}$, of $\langle \mathbf{H}_{ab}^{\dagger}\mathbf{H}_{ab}\rangle$, and we have

$$\langle C_{\rm priv}(\boldsymbol{\mu}, N_T)\rangle \geq \max_{\boldsymbol{N}:\, \sum_{m=1}^{M} N_m = N_T} \sum_{m=1}^{M} g_{\rm hol}^{\rm priv}\,(\gamma_m\,, N_m)\,. \quad (41)$$

This lower bound can be evaluated—for transmitters that use FB, HG, or LG modes, and receivers that do or do not use adaptive optics—from the results we presented in the Part I paper.[5]

IV. Results for FB, HG, and LG Modes

We are now reaching the culmination of this two-part study. Specifically, we are now equipped to bound the PIE versus SE behaviors obtained—with and without adaptive optics—for multiple spatial-mode optical communication through atmospheric turbulence. We will do so for transmitters that use ~200 FB, HG, or LG modes. We proved, in [9], that the HG and LG modes have identical eigenspectra, implying that their respective ergodic Holevo capacities, their ergodic OOK-DD capacities, and their ergodic Holevo private capacities also coincide. The numerical results that follow assume the propagation conditions from the Part I paper, namely, the Fig. 1 propagation geometry with $d_T = d_R = 17.6$ cm, and operation at $1.55\,\mu$m wavelength over an $L = 1$ km propagation path through uniformly-distributed Kolmogorov-spectrum turbulence. The vacuum-propagation Fresnel-number product for this scenario is $D_f = 400$. By choosing to use only ~200 spatial modes we thus afford ourselves some protection against the beam spread, angular spread, and scintillation incurred in propagation through turbulence. Three turbulence strengths will be considered: $C_n^2 = 5\times 10^{-15}\ \mathrm{m}^{-2/3}$ (mild turbulence), $C_n^2 = 5\times 10^{-14}\ \mathrm{m}^{-2/3}$ (moderate turbulence), and $C_n^2 = 5\times 10^{-13}\ \mathrm{m}^{-2/3}$ (strong turbulence).[6]

In Fig. 2 we have plotted PIE versus SE behaviors associated with the ergodic Holevo capacity upper bound (9) and its lower bound (17). FB-mode results with and without adaptive optics appear in Fig. 2(a), while the corresponding results for HG/LG modes appear in Fig. 2(b). We see from Fig. 2(a) that the lower bounds and upper bounds of FB systems that use adaptive optics in mild or moderate turbulence are nearly coincident with each other and with the vacuum-propagation curve; Fig. 2(b) shows that the same situation prevails for the HG and LG modes. Evidently, perfect full-wave adaptive optics enables near-field communication with considerable protection against the ill-effects of atmospheric turbulence. Indeed, these systems can deliver 10 bits/detected-photon at 5 bits/s·Hz performance with ~200 spatial modes at the ergodic Holevo limit with aperture efficiency $M/D_f \approx 0.5$.[7] In strong turbulence, the lower bounds on PIE versus SE for the FB and the HG/LG systems using adaptive optics drop well below their respective upper bounds. Nevertheless, they still indicate that multiple bits/detected-photon at multiple bits/s·Hz are possible. The same is true, but to a significantly lesser extent, for operation without adaptive optics. Those systems, however, will have to cope with appreciable crosstalk between the spatial patterns generated in the receiver pupil, where [9] indicates how much crosstalk might be present.[8]

It is interesting to note that systems employing perfect adaptive optics with FB modes have higher ergodic Holevo capacities—at least as measured by our lower bounds—than the corresponding systems using HG or LG spatial modes. This behavior seems to imply that, after propagation through

[5]The photon allocation $\boldsymbol{N}$ that maximizes the bound in (41) must have nonincreasing N_m with increasing m. This can be shown using the same techniques we used to prove Lemma A.3.

[6]As shown in the Part I paper for the eigenspectra, our capacity bounds for the preceding channel parameters apply to any uniform-turbulence, square-pupil scenario that has $d_T = d_R$, $D_f = 400$, and the same Rytov-theory logamplitude variance, $\sigma_\chi^2 = 0.124\,k^{7/6}C_n^2 L^{11/6}$, where $k = 2\pi/\lambda$ is the wave number at wavelength λ.

[7]It behooves us to note that the curves in Fig. 2 do *not* account for the photons needed by the adaptive-optics system's wavefront sensor. The same omission will apply to Figs. 3 and 4. We will comment further on this issue in Section V.

[8]It must also be said that when our lower bounds lie well below the corresponding upper bound—as is the case for strong turbulence—we cannot be confident of their tightness. Nevertheless, these lower bounds do guarantee that high PIE and high SE can be achieved simultaneously.

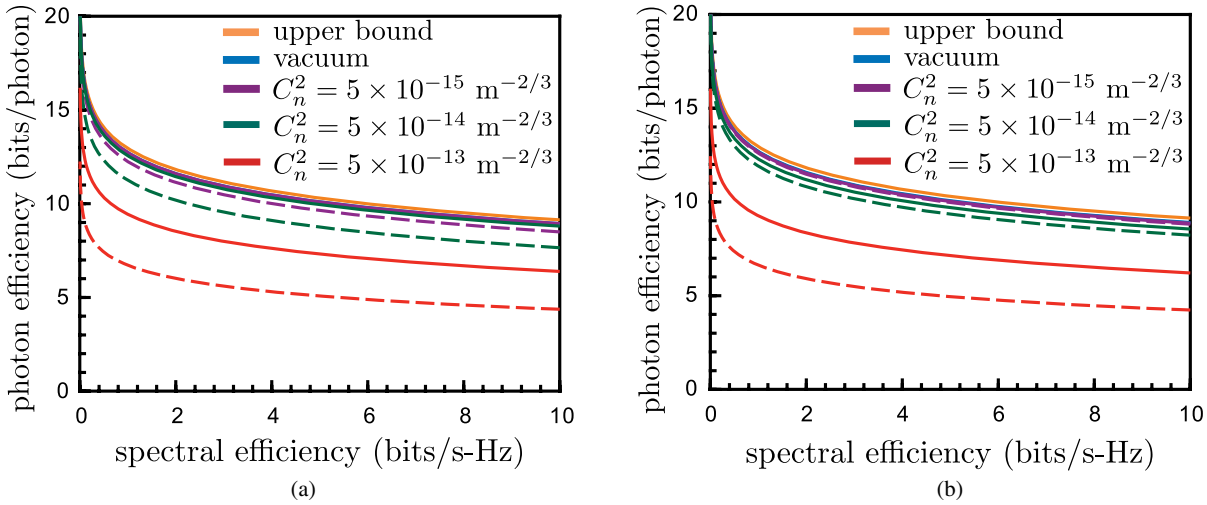

Fig. 2. Upper and lower bounds on PIE, as functions of SE, of the ergodic Holevo capacities for: (a) $M = 225$ FB-mode systems; and (b) $M = 231$ HG or LG systems. These results assume square-pupil transmitters and receivers with identical $d = 17.6$ cm side lengths, and operation at $\lambda = 1.55\,\mu$m wavelength over an $L = 1$ km propagation path. Solid lines are for systems with perfect adaptive optics, while dashed lines are for systems without adaptive optics.

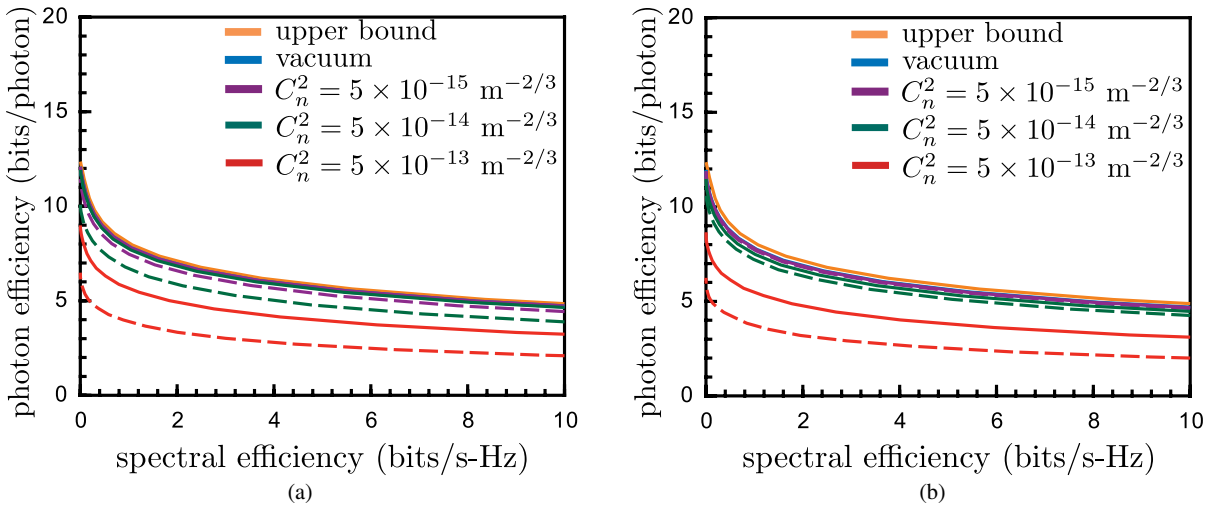

Fig. 3. Upper and lower bounds on PIE, as functions of SE, of the ergodic OOK-DD capacities for: (a) $M = 225$ FB-mode systems; and (b) $M = 231$ HG or LG systems. These results assume square-pupil transmitters and receivers with identical $d = 17.6$ cm side lengths, and operation at $\lambda = 1.55\,\mu$m wavelength over an $L = 1$ km propagation path. Solid lines are for systems with perfect adaptive optics, while dashed lines are for systems without adaptive optics.

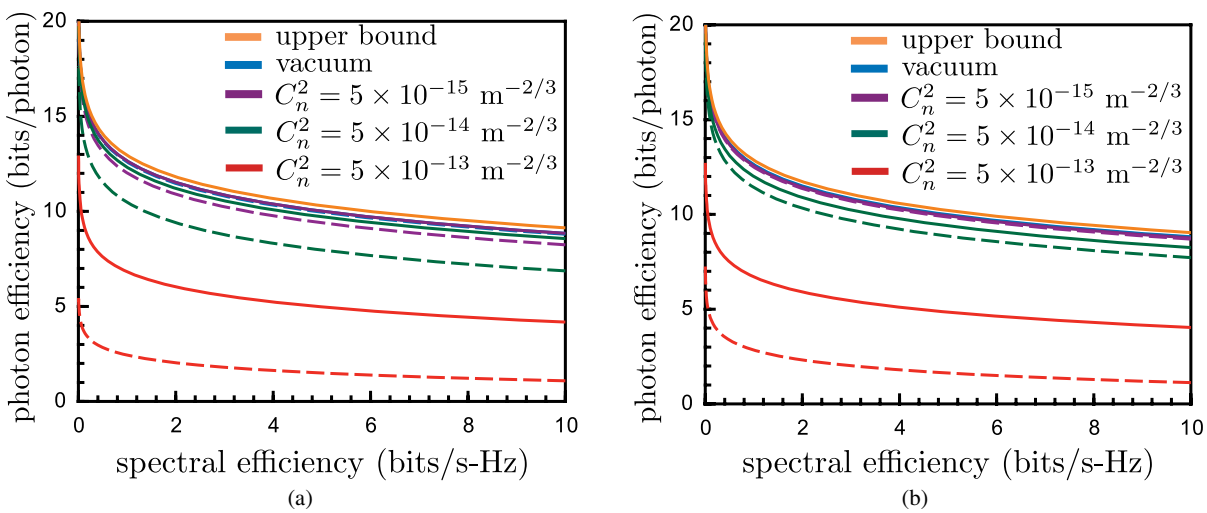

Fig. 4. Upper and lower bounds on PIE, as functions of SE, of the ergodic Holevo private capacities for: (a) $M = 225$ FB-mode systems; and (b) $M = 231$ HG or LG systems. These results assume square-pupil transmitters and receivers with identical $d = 17.6$ cm side lengths, and operation at $\lambda = 1.55\,\mu$m wavelength over an $L = 1$ km propagation path. Solid lines are for systems with perfect adaptive optics, while dashed lines are for systems without adaptive optics.

turbulence, FB modes have more robust power transfer to the receiver pupil than do the HG/LG modes. Conversely, when adaptive optics are not employed at the receiver, utilizing HG or LG spatial modes results in higher performance that what is obtained with FB modes, indicating that their vacuum-propagation output-mode patterns are better preserved, after propagation through turbulence, than those of the FB modes.

Fig. 3 shows PIE versus SE results for OOK DD systems. In comparison with the corresponding Holevo-capacity results from Fig. 2, the OOK-DD systems fall well short of the latter's PIE values. This is hardly surprising. Achieving 10 bits/detected-photon at 5 bits/s·Hz with OOK-DD requires 4500 high-transmissivity spatial modes for vacuum propagation, whereas 189 spatial modes will suffice at the Holevo limit. The general characteristics seen in Fig. 2 are still present here for OOK-DD: 1) with adaptive optics there is near-immunity to moderate turbulence; 2) FB modes outperform HG/LG modes when adaptive optics are used; and 3) HG/LG modes outperform FB modes when adaptive optics are not used.

Finally, in Fig. 4, we show the PIE versus SE behavior of the ergodic Holevo private capacities of FB and HG/LG systems with and without adaptive optics. The general trends we saw in Figs. 2 and 3 are still evident, but the effect of turbulence is considerably more severe for private capacity—especially when adaptive optics are not employed—than was the case earlier. This is due to the fact that modes whose transmissivity falls below 1/2 cannot transmit information securely because of our assumption that Eve receives all the light that does not reach Bob.

The overarching conclusion to be drawn from the cases we have evaluated in this section is that optical communication with high PIE *and* high SE—for both classical communication and private communication—*can* be delivered through moderate atmospheric turbulence with the appropriate multiple spatial-mode systems.

V. Conclusion

Multiple-spatial mode operation is necessary for free-space optical communication that realizes high PIE and high SE simultaneously. To gauge the impact of atmospheric turbulence on the PIE versus SE behavior that can be obtained with such systems, we derived upper and lower bounds on the ergodic Holevo capacities for classical and private communication and the ergodic OOK DD capacity. Fixed transmitter mode sets were assumed with receivers that either did or did not employ adaptive optics. Our bounds depend on the eigenspectra of the average channel kernels for adaptive and non-adaptive operation. Using our eigenspectra calculations from [9] we evaluated the PIE bounds, as functions of SE, for representative systems that used ~200 FB, HG, or LG modes in mild, moderate, or strong turbulence. We found that little or no ergodic capacity was lost when ideal full-wave adaptive optics were used in mild or moderate turbulence. More pronounced ergodic-capacity loss was incurred, especially for non-adaptive systems, in strong turbulence. Private capacity was more susceptible to turbulence degradation, because spatial modes whose transmissivities do not exceed 1/2 have zero private capacity when the eavesdropper receives all the light that does not reach the intended receiver.

There are several important avenues to pursue as follow-ons to the current capacity work. First, there is bounding the ergodic private capacity for the OOK-DD system, a problem made difficult by the analytic intractability of OOK-DD's single-mode private capacity. Second, there is the issue of full-wave tracking, as opposed to the phase-only compensation that almost all adaptive-optics systems employ. If we can obtain the eigenspectra of the average channel kernel for a phase-compensated system, then we should be able to bound its PIE versus SE behavior using the results derived herein. Finally, an architecture and photon-tracking budget must also be established for full-wave or phase-only adaptive optics for our multi-spatial-mode systems. For Gbps single-mode communication through turbulence between stationary terminals, the ~millisecond-duration atmospheric coherence time allows high signal-to-noise ratio to be maintained in the tracking loop for each coherence area in the receiver's entrance pupil while bleeding off only a small fraction of the incoming light. If, however, every spatial mode needs its own tracker, then the photon budget could become prohibitive, even with the large disparity between the atmospheric coherence time and the bit duration of Gbps communication. In this regard the recent experimental demonstration that phase-only wavefront sensing on the fundamental Gaussian mode may suffice for phase-compensation of the OAM-carrying LG modes with Gaussian radial distributions is quite important [23].

Appendix

Lemma A.1: For a fixed nonzero p, $f_{\rm ook}(x,p)$ is monotonically increasing in x.

Proof: Set $\varepsilon = e^{-x/p}$. We have that

$$\frac{\partial f_{\rm ook}(x,p)}{\partial x} = \frac{\varepsilon}{\ln(2)} \ln\left(\frac{1-p+\varepsilon p}{\varepsilon p}\right) \geq 0 \tag{42}$$

because $(1-p+\varepsilon p)/\varepsilon p \geq 1$ for positive p, which proves that $f_{\rm ook}(x,p)$ is monotonically increasing in x. ■

Lemma A.2: For a fixed nonzero p, $f_{\rm ook}(x,p)$ is concave in x.

Proof: Again set $\varepsilon = e^{-x/p}$. We have that

$$\frac{\partial^2 f_{\rm ook}(x,p)}{\partial x^2} = \frac{\varepsilon}{p\ln(2)}\left(\ln\left(\frac{\varepsilon p}{1-p+\varepsilon p}\right) + \frac{1-p}{1-p+\varepsilon p}\right) \tag{43}$$

$$\leq 0 \tag{44}$$

for positive p because $\ln(u) \leq u-1$, which proves that $f_{\rm ook}(x,p)$ is concave in x. ■

Lemma A.3: When evaluating

$$\max_{\substack{N_1,N_2 \\ N_1+N_2=N_T \\ p_1,p_2\,:\,0\leq p_1,p_2\leq 1}} \sum_{m=1}^{2} \mu_m f_{\rm ook}(N_m,p_m) \tag{45}$$

the condition $\mu_1 \geq \mu_2$ implies that $N_1^* \geq N_2^*$, where N_1^* and N_2^* achieve the maximum in (45).

Proof: We analyze the difference between the values of the objective function when $N_1 \geq N_2$ and $N_1 < N_2$ in order to see which case yields the greater value in (45) when $\mu_1 \geq \mu_2$. First suppose that $N_1 \geq N_2 = N_T - N_1$, and let $p^*_{N_1}$ and $p^*_{N_T - N_1}$ maximize $f_{\rm ook}(N_1, p)$ and $f_{\rm ook}(N_T - N_1, p)$, respectively. We have that

$$\max_{p_1, p_2 \,:\, 0 \leq p_1, p_2 \leq 1} [\mu_1 f_{\rm ook}(N_1, p_1) + \mu_2 f_{\rm ook}(N_T - N_1, p_2)] - \max_{p_1, p_2 \,:\, 0 \leq p_1, p_2 \leq 1} [\mu_1 f_{\rm ook}(N_T - N_1, p_1) + \mu_2 f_{\rm ook}(N_1, p_2)] = (\mu_1 - \mu_2)[f_{\rm ook}(N_1, p^*_{N_1}) - f_{\rm ook}(N_T - N_1, p^*_{N_T - N_1})] \tag{46}$$

$$\geq (\mu_1 - \mu_2)[f_{\rm ook}(N_1, p^*_{N_T - N_1}) - f_{\rm ook}(N_T - N_1, p^*_{N_T - N_1})] \tag{47}$$

$$\geq 0. \tag{48}$$

Inequality (47) is due to $\mu_1 \geq \mu_2$ and $p^*_{N_1}$'s maximizing $f_{\rm ook}(N_1, p)$, and inequality (48) follows because $N_1 \geq N_T - N_1$ and $f_{\rm ook}(x, p)$ is monotonically increasing in x for positive p.

Now, consider the case in which $N_2 > N_1 = N_T - N_2$, and let $p^*_{N_2}$ and $p^*_{N_T - N_2}$ maximize $f_{\rm ook}(N_2, p)$ and $f_{\rm ook}(N_T - N_2, p)$, respectively. Here we have that

$$\max_{p_1, p_2 \,:\, 0 \leq p_1, p_2 \leq 1} [\mu_1 f_{\rm ook}(N_T - N_2, p_1) + \mu_2 f_{\rm ook}(N_2, p_2)] - \max_{p_1, p_2 \,:\, 0 \leq p_1, p_2 \leq 1} [\mu_1 f_{\rm ook}(N_2, p_1) + \mu_2 f_{\rm ook}(N_T - N_2, p_2)] = (\mu_1 - \mu_2)[f_{\rm ook}(N_T - N_2, p^*_{N_T - N_2}) - f_{\rm ook}(N_2, p^*_{N_2})] \tag{49}$$

$$\leq (\mu_1 - \mu_2)[f_{\rm ook}(N_T - N_2, p^*_{N_T - N_2}) - f_{\rm ook}(N_2, p^*_{N_T - N_2})] \tag{50}$$

$$< 0. \tag{51}$$

Inequality (50) is due to $\mu_1 \geq \mu_2$ and $p^*_{N_2}$'s maximizing $f_{\rm ook}(N_2, p)$, and inequality (51) follows because $N_2 \geq N_T - N_2$ and $f_{\rm ook}(x, p)$ is monotonically increasing in x for positive p.

Combining inequalities (48) and (51) proves that for $\mu_1 \geq \mu_2$ and any non-negative N_a, N_b satisfying $N_a \geq N_b$ and $N_a + N_b = N_T$, (45) takes on a greater or equal value when $N_1 = N_a$ and $N_2 = N_b$ then it does when $N_1 = N_b$ and $N_2 = N_a$, thus completing the proof. ■

Lemma A.4: For fixed $N > 0$, the single-mode private capacity $g^{\rm priv}_{\rm hol}(\mu, N)$ is convex in μ.

Proof: We know that $g^{\rm priv}_{\rm hol}(\mu, N) = 0$ for $\mu \in [0, 1/2]$, so it is convex in that interval. For $\mu \in [1/2, 1]$, we have that

$$\frac{\partial^2 g^{\rm priv}_{\rm hol}(\mu, N)}{\partial \mu^2} = \frac{N}{(1 + (1-\mu)N)(1-\mu)} - \frac{N}{(1 + \mu N)\mu} \tag{52}$$

$$\geq 0 \tag{53}$$

which shows that $g^{\rm priv}_{\rm hol}(\mu, N)$ is also convex for $\mu \in [1/2, 1]$ thus completing the proof. ■

Lemma A.5: For a fixed $N_T > 0$, the multi-mode private capacity $C_{\rm priv}(\boldsymbol{\mu}, N_T)$ is both convex and Schur-convex in $\boldsymbol{\mu}$.

Proof: Because $C_{\rm priv}(\boldsymbol{\mu}, N_T)$ is symmetric in the elements of $\boldsymbol{\mu}$, its convexity will imply its Schur-convexity [24]. To prove that $C_{\rm priv}(\boldsymbol{\mu}, N)$ is convex in $\boldsymbol{\mu}$, consider any two vectors $\boldsymbol{\mu}^a$ and $\boldsymbol{\mu}^b$, whose elements are nonincreasing and lie in the interval $[0, 1]$. Now let $\boldsymbol{\mu}^c = p\boldsymbol{\mu}^a + (1-p)\boldsymbol{\mu}^b$ for arbitrary $p \in [0, 1]$. Defining $\boldsymbol{N}^c$ to be the modal photon allocation that achieves $C_{\rm priv}(\boldsymbol{\mu}_c, N_T)$, we have that

$$C_{\rm priv}(\boldsymbol{\mu}^c, N_T) = \sum_{m=1}^{M} g^{\rm priv}_{\rm hol}(\mu^c_m, N^c_m) \tag{54}$$

$$\leq \sum_{m=1}^{M} \left[p g^{\rm priv}_{\rm hol}(\mu^a_m, N^c_m) + (1-p) g^{\rm priv}_{\rm hol}(\mu^b_m, N^c_m)\right] \tag{55}$$

$$\leq p C_{\rm priv}(\boldsymbol{\mu}^a, N_T) + (1-p) C_{\rm priv}(\boldsymbol{\mu}^b, N_T). \tag{56}$$

Inequality (55) follows from the convexity of $g^{\rm priv}_{\rm hol}(\mu, N)$ in μ. Inequality (56), which proves that $C_{\rm priv}(\boldsymbol{\mu}, N_T)$ is convex (and hence Schur convex) in $\boldsymbol{\mu}$, follows from the suboptimality of the photon modal allocation $\boldsymbol{N}^c$ for transmissivity vectors $\boldsymbol{\mu}^a$ and $\boldsymbol{\mu}^b$.

References

[1] S. Guha, Z. Dutton, and J. H. Shapiro, "On quantum limit of optical communications: Concatenated codes and joint detection receivers," in *Dig. IEEE Int. Symp. Inform. Theory*, 2011, pp. 274–278.

[2] B. S. Robinson, A. J. Kerman, E. A. Dauler, R. O. Barron, D. O. Caplan, M. L. Stevens, J. J. Carney, S. A. Hamilton, J. K. W. Yang, and K. K. Berggren, "781 Mbit/s photon-counting optical communications using a superconducting nanowire detector," *Opt. Lett.*, vol. 31, no. 4, pp. 444–446, Feb. 15, 2006.

[3] G. Li, "Recent advances in coherent optical communication," *Adv. Opt. Photon.*, vol. 1, no. 2, pp. 279–307, Feb. 2009.

[4] B. Zhu, T. F. Taunay, M. Fishteyn, X. Liu, S. Chandrasekhar, M. F. Yan, J. M. Fini, E. M. Monberg, and F. V. Dimarcello, "112-Tb/s space-division multiplexed DWDM transmission with 14-b/s/Hz aggregate spectral efficiency over a 76.8-km seven-core fiber," *Opt. Exp.,*vol. 19, no. 17, pp. 16665–16671, Aug. 15, 2011.

[5] B. Zhu, J. M. Fini, M. F. Yan, X. Liu, S. Chandrasekhar, T. F. Taunay, M. Fishteyn, E. M. Monberg, and F. V. Dimarcello, "High-capacity space-division-multiplexed DWDM transmission using multicore fiber," *J. Lightw. Technol*, vol. 30, no. 4, pp. 486–492, Feb. 15, 2012.

[6] J. Sakaguchi, B. J. Puttnam, W. Klaus, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, K. Imamura, H. Inaba, K. Mukasa, R. Sugizaka, T. Kobayashi, and M. Watanabe, "305 Tb/s space division multiplexed transmission using homogeneous 19-core fiber," *J. Lightw. Technol*, vol. 31, no. 4, pp. 554–562, Feb. 15, 2013.

[7] R. S. Kennedy, "Communication through optical scattering channels: An introduction," *Proc. IEEE*, vol. 58, no. 10, pp. 1651–1665, Oct. 1970.

[8] J. H. Shapiro, "Imaging and optical communication through atmospheric turbulence," in *Laser Beam Propagation in the Atmosphere*, J. W. Strohbehn, Ed. Berlin, Germany: Springer-Verlag, 1978, ch. 6.

[9] N. Chandrasekaran and J. H. Shapiro, "Photon information efficient communication through atmospheric turbulence—Part I: Channel model and propagation statistics," *J. Lightw. Technol.*, vol. 32, no. 6, pp. 1075–1087, Mar. 2014.

[10] D. Slepian, "Analytical solution to two apodization problems," *J. Opt. Soc. Amer.*, vol. 55, no. 9, pp. 1110–1114, Sep. 1965.

[11] J. H. Shapiro, "Normal-mode approach to wave propagation in the turbulent atmosphere," *Appl. Opt.*, vol. 13, no. 11, pp. 2614–2619, Nov. 1974.

[12] L. C. Andrews and R. L. Phillips, *Laser Beam Scintillation with Applications*. Bellingham, WA, USA: SPIE, 2001.
[13] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*, 2nd ed. Bellingham, WA, USA: SPIE, 2005.
[14] M. M. Wilde, S. Guha, S.-H. Tan, and S. Lloyd, "Explicit capacity-achieving receivers for optical communication and quantum reading," in *Dig. IEEE Int. Symp. Inform. Theory*, 2012 pp. 551–555.
[15] S. Dolinar, B. I. Erkmen, B. Moision, K. Birnbaum, and D. Divsalar, "The ultimate limits of optical communication efficiency with photon-counting receivers," in *Dig. IEEE Int. Symp. Inform. Theory*, 2012, pp. 541–545.
[16] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
[17] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Exp.*, vol. 19, no. 17, pp. 16680–16696, Aug. 15, 2011.
[18] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," *Phys. Rev. Lett.*, vol. 92, no. 2, art. 027902, Jan. 16, 2004.
[19] J. P. Gordon, "Quantum effects in communication systems," *Proc. IRE*, vol. 50, no. 9, pp. 1898–1908, Sep. 1962.
[20] J. P. Gordon, "Quantum electronics and coherent light," in *Proc. Int. School Phys. Enrico Fermi*, 1964, pp. 156–181.
[21] G. Smith, "Private classical capacity with a symmetric side channel and its application to quantum cryptography," *Phys. Rev. A*, vol. 78, no. 2, art. 022306, Aug. 2008.
[22] M. M. Wolf, D. Pérez-García, and G. Giedke, "Quantum capacities of bosonic channels," *Phys. Rev. Lett.*, vol. 98, no. 13, art. 130501, Mar. 30, 2007.
[23] Y. Ren, H. Huang, G. Xie, N. Ahmed, Y. Yan, B. I. Erkmen, N. Chandrasekaran, M. P. J. Lavery, N. K. Steinhoff, M. Tur, S. Dolinar, M. Neifeld, M. J. Padgett, R .W. Boyd, J. H. Shapiro, and A. E. Willner, "Atmospheric turbulence effects on the performance of a free space optical link employing orbital angular momentum multiplexing," *Opt. Lett.*, vol. 38, no. 20, pp. 4062–4065, Oct. 2013.
[24] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York, NY, USA: Academic, 1979.

Nivedita Chandrasekaran (S'13) received the S.B. degrees in electrical engineering and physics, and the M.Eng. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2008, 2009, and 2009, respectively. She is currently working toward the Ph.D. degree in the Department of Electrical Engineering and Computer Science, MIT. Her research interests include quantum and optical communication through turbulence.

Jeffrey H. Shapiro (S'67–M'70–SM'84–F'95–LF'12) received the S.B., S.M., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1967, 1968, 1969, and 1970, respectively. As a graduate student, he was a National Science Foundation Fellow, a Teaching Assistant, and a Fannie and John Hertz Foundation Fellow. His doctoral research was a theoretical study of adaptive techniques for improved optical communication through atmospheric turbulence. From 1970 to 1973, he was an Assistant Professor of Electrical Sciences and Applied Physics at Case Western Reserve University, Cleveland, OH, USA. From 1973 to 1985, he was an Associate Professor of Electrical Engineering at MIT, where he was promoted to a Professor of Electrical Engineering in 1985, he was an Associate Head of the Department of Electrical Engineering and Computer Science from 1989 to 1999, and the Director of the Research Laboratory of Electronics from 2001 to 2011. In 1999, he became the Julius A. Stratton Professor of Electrical Engineering. His current research interests include the application of communication theory to optical systems. He is best known for his work on the generation, detection, and application of squeezed-state light beams, but he has also published extensively in the areas of atmospheric optical communication, coherent laser radar, and quantum information science. He is a Fellow of the Optical Society of America, the American Physical Society, the Institute of Physics, and SPIE (the International Society for Optical Engineering). In 2008, he was a corecipient of the Quantum Electronics Award from the IEEE Photonics Society (formerly known as IEEE Lasers and Electro-Optics Society), and also received the Quantum Communication Award for Theoretical Research from Tamagawa University, Tokyo, Japan. He has been an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and the *Journal of the Optical Society of America*. He was the Principal Organizer of the Sixth International Conference on Quantum Communication, Measurement, and Computing in 2002.

Ligong Wang (S'08–M'12) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004, and the M.Sc. and Dr. Sc. degrees in electrical engineering from ETH Zurich, Switzerland, in 2006 and 2011, respectively. He is currently a Postdoctoral Associate in the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, MA, USA. His research interests include classical and quantum information theory and optical communication.