This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

Secrecy Capacities in Space-Division Multiplexed Fiber Optic Communication Systems

Kyle Guan, *Member*, *IEEE*, Antonia M. Tulino, *Fellow*, *IEEE*, Peter J. Winzer, *Fellow*, *IEEE*, and Emina Soljanin, *Fellow*, *IEEE*
Bell Labs, Alcatel-Lucent, Holmdel, NJ, USA

***Abstract*—Space-division multiplexed (SDM) fiber optic transmission systems can not only increase system capacity, but also achieve physical-layer security against fiber tapping attacks. In this paper, we examine the information-theoretic security of optical multiple-input-multiple-output (MIMO) SDM by evaluating the trade-off between the achievable information rate and the confidentiality for different channel dynamics. In particular, we provide problem formulations for secure communication over these channels and study three types of secrecy capacities: guaranteed capacity, outage capacity, and average capacity, each serving as a performance metric for a coding strategy tailored to a specific type of MIMO-SDM channel. We also assess the impact of key system parameters, such as the number of modes, the mode-dependent loss (MDL), and the signal-to-noise ratio (SNR), on the various secrecy capacities. Our results indicate that, with a proper design of channel codes that balance information rate and security, an SDM system has the potential of offering confidential data transmission at a rate that could be orders of magnitude higher than what can be achieved through other means of encryption. Moreover, we show that MDL, unavoidably induced by fiber tapping, can allow information-theoretic security even if the SNR of the eavesdropper's receiver is better than that of the legitimate receiver.**

***Index Terms*—Space-division Multiplexing (SDM), multiple-input-multiple-output (MIMO) systems, secrecy capacity, mode-dependent loss (MDL)**

I. Introduction

A. Motivation

Fiber-optic communication systems are vulnerable to various types of physical-layer attacks [1]-[2]. A common attack is fiber tapping through a fiber bend – an eavesdropper with physical access to the fiber extracts a portion of the propagating signal by bending the fiber and detecting the evanescent optical field coupled out of the fiber at the bend. The fact that this type of eavesdropping in conventional optical fiber networks is relatively easy to implement and could remain largely unnoticed is a major concern for physical-layer security. Quantum key distribution (QKD) addresses this issue from a quantum mechanical angle by allowing an exchange of a fundamentally secure key between a transmitter and a receiver while at the same time providing intrusion detection [3]. However, the benefits of QKD are offset by stringent limitations in terms of data rate and reach (e.g., 1 Mb/s over 50 km of fiber [4]), as well as by severe problems arising from optical amplifier noise and interactions between classical communications and QKD signals on a common optical transmission infrastructure [5].

Recently, space-division multiplexing (SDM) has been proposed [6]-[9] and demonstrated [10]-[12] as a promising solution to overcome the looming capacity crunch in optical networks. Using a strand of parallel single-mode fibers, uncoupled or coupled cores of a multi-core fiber, or even individual modes of a few-mode fiber in combination with *multiple-input-multiple-output* (MIMO) digital signal processing, SDM can drastically increase system capacity and at the same time reduce cost and energy per bit via the integration of components. Independent of its network capacity benefits, MIMO-SDM also has the potential for physical-layer security in high-speed optical information transmission [13]. The notion of security in SDM systems is based on the fact that a fiber bend induced by an eavesdropper upon an SDM waveguide fundamentally perturbs the spatial information seen by the eavesdropper. This makes the eavesdropper's MIMO channel less favorably conditioned than that of the legitimate user, which in turn lets the eavesdropper extract a much reduced amount of information from the channel. Furthermore, the mode-dependent loss (MDL) induced onto the propagating signal allows the legitimate receiver to detect the presence of an eavesdropper[1].

Extending the results previously developed in [14], [15], we study the information-theoretically provable security of MIMO-SDM systems in this paper. Information-theoretic security is a widely accepted and fundamentally provable notion of secrecy in a classical communication system. When operating in information-theoretic secrecy, the *equivocation* (randomness) of the source, measured by its information entropy, is not reduced for the eavesdropper when observing the output of its wiretap channel. In other words, even though the eavesdropper observes the wiretap channel output, he or she is not better off than randomly guessing the data. The *secrecy capacity* then quantifies the maximum achievable information that can be transmitted to a legitimate receiver such that an eavesdropper cannot receive any useful information [16].

B. Main Contributions

The focus of this work is to provide an in-depth and comprehensive evaluation of secrecy capacities for various channel dynamics and system design objectives. In particular,

[1]This paper focuses on analyzing the secrecy capacity of MIMO-SDM systems. The aspect of detecting the presence of an eavesdropper based on MDL changes recorded at the legitimate receiver (e.g., by observing the MDL compensation of its adaptive MIMO equalizer) is left for future studies.

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

we adopt a system model that reflects the unique physical characteristics of the optical fiber MIMO-SDM channels and provide formulations and analyses for the following three types of secrecy capacities: *guaranteed secrecy capacity*, *outage secrecy capacity*, and *average secrecy capacity*. We also assess the impact of key system parameters on the secrecy capacity. Our results indicate that MIMO-SDM has the potential to provide provable physical-layer security at data rates orders of magnitude higher than what can be achieved through other encryption approaches. Moreover, as a result of the eavesdropper's MDL, information-theoretic security can still be achieved even if the SNR of the eavesdropping channel is better than that of the legitimate receiver. Since MIMO-SDM systems can provide Petabit/s capacities and beyond, enormous information-theoretically secret data rates (in the range of multi-Terabit/s) promise to be possible over a transmission system that is likely to be in widespread commercial use within the coming decade.

C. Related Work

The results from the vast body of research [16]-[26] in information theoretic security provide the theoretical foundation for this work. Wyner first introduced the concepts of equivocation and secrecy [16]. Cheong and Hellman studied the single-input-single-output (SISO) Gaussian channel and showed that the secrecy capacity is the difference between the capacities of main and eavesdropping channel [17]. This result implies that the capacity of the main channel needs to be larger than that of the eavesdropping channel, otherwise fundamentally secure communication is impossible. Barros and Rodrigues studied the outage secrecy capacity of SISO fading channels, where channel realizations were treated as random variables [18]. The security issues of MIMO systems have been extensively studied recently for wireless systems [19]-[26]. The security benefit of MIMO systems was first evaluated in [19] in terms of low probability of interception. The results in [20]-[23] provide the theoretical framework for analyzing the secrecy capacity of MIMO wire-tap channels with *channel state information (CSI) known at both transmitter and receiver*. The use of spatial diversity to improve the confidentiality of atmospheric MIMO free-space optical communication was studied in [24]. Recently, several papers [25], [26] started to look into the secrecy capacity of a wireless MIMO system without CSI at the transmitter.

The remainder of this paper is organized as follows. In Section II we provide an overview of fiber-optic MIMO-SDM systems and describe the system model for SDM waveguide tapping. In Section III we characterize the secrecy capacities associated with different channel dynamics. In Section IV we provide main results and discuss their implications. In Section V we summarize our main findings.

II. SDM Fiber/Waveguide Tapping Model

A. Optical MIMO-SDM Systems

To scale optical transport capabilities both economically and energy efficiently, SDM technologies need to leverage *integration* of system components among channels [27], [28]. Since integration generally comes at the expense of *crosstalk* among parallel paths, proper crosstalk management is an important aspect of SDM systems. In the high crosstalk regime, MIMO techniques [29], originally developed for wireless systems, can be used to mitigate crosstalk [9], [30], [31]. Compared with MIMO wireless systems, optical MIMO-SDM systems have some unique characteristics [9], [32]:

- Carrier-grade reliability of 99.999% is normally required, i.e., outage probabilities must be in the 10^{-5} range.
- The receiver-to-transmitter feedback delays in optical transport systems are comparable to the channel dynamics. Consequently, there is no possibility of receiver-transmitter feedback and hence no availability of CSI at the transmitter.
- Transmitters are able to excite and receivers are able to extract the full mode set supported by the SDM waveguide with generally low MDL. Hence, the MIMO-SDM channel can be approximated as a "perturbed unitary" channel.
- The transmit power is constrained for each mode individually by fiber nolinearities, as opposed to the total average power constraint that is normally used in studying wireless MIMO systems.
- SDM waveguides can be designed with certain goals in mind, and thus can be tailored to the needs of physical layer security.

Many of these characteristics will be factored into the models and formulations used in this paper.

B. System Model

Fig. 1 (a) illustrates the fiber-tapping model. A signal vector $\mathbf{x}$ sent by Alice, with x_i denoting the signal transmitted on the ith mode, is received by a legitimate receiver (Bob) as $\mathbf{y}$ and by an eavesdropper (Eve) as $\mathbf{y^e}$. The abstraction of the channel model is shown in Fig. 1 (b). Specifically, we consider an SDM system that supports a set of M orthogonal propagation modes, which are subject to coupling and MDL. Here we ignore the fiber inter- and intra-modal nonlinearities and model the SDM system as a linear matrix MIMO channel. That is, we use $M \times M$ (normalized) matrices $\mathbf{H}$ and $\mathbf{H^e}$ to represent the realizations of the legitimate (main) and eavesdropping channel, respectively[2]. Assuming that the noise generated within the receiver dominates, the received signals $\mathbf{y}$ and $\mathbf{y^e}$ are:

$$\mathbf{y} = \sqrt{E_0}\sqrt{L}\mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

$$\mathbf{y^e} = \sqrt{E_0}\sqrt{L^e}\mathbf{H^e}\mathbf{x} + \mathbf{n^e}, \quad (2)$$

where L and L^e are normalization factors, with $L = \mathrm{tr}\{\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger\}/M$ and $L^e = \mathrm{tr}\{\tilde{\mathbf{H}}^\mathbf{e}\tilde{\mathbf{H}}^{\mathbf{e}\dagger}\}/M$. L and L^e characterize the mode-average loss of the respective channels $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{H}}^\mathbf{e}$

[2]In this work, we only focus on the limiting cases of frequency-flat (rapidly-varying and slowly-varying) channels, without considering other intermediate cases. Thus our model leaves out the detailed frequency dependence of $\mathbf{H}$ and $\mathbf{H^e}$.

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

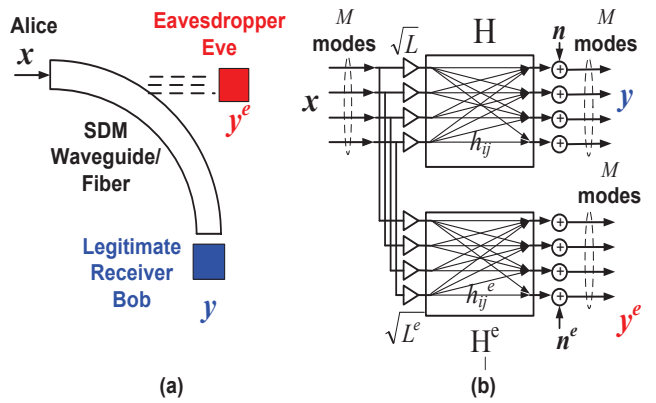

Fig. 1. (a) Fiber-tapping model; (b) A MIMO-SDM system that supports a set of M orthogonal propagation modes.

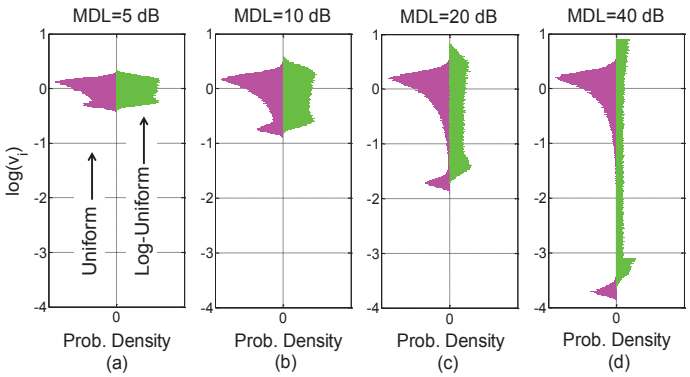

Fig. 2. The distributions of v_i for uniformly and log-uniformly distributed MDL models are shown on the left and right side of each subplot, respectively. Subplots (a)-(d) correspond to MDLs of 5, 10, 20, and 40 dB, respectively.

[9]. The distributions of channel noises $\mathbf{n}$ and $\mathbf{n^e}$ are circularly symmetric complex Gaussian with per-mode power spectral density N_0 and N_0^e for Bob and Eve, respectively.

As mentioned above, the receiver-to-transmitter feedback delays in optical transport systems (hundreds to thousands of fiber kilometers, corresponding to round-trip times of several microseconds, processing excluded) are comparable to the time constant expected for MIMO-SDM channel dynamics (with spectral content from the kHz up into the MHz regime [33][34]). As such, we make the assumption that CSI is not available at the transmitter. However, Bob can estimate the channel state of the MIMO-SDM system by using training symbols. We thus assume that the individual realization of $\mathbf{H}$ is known to Bob. Similarly, the channel realization of $\mathbf{H^e}$ is unknown to Alice. In our model, we grant Eve as much capability as possible, including her knowledge of the training symbols. Thus $\mathbf{H^e}$ can be estimated by and thus is known to Eve.

We consider a phenomenological channel model for the effect of fiber bending tapping [35]. Motivated by an eavesdropper's desire to couple as little light out of the SDM fiber as possible (to avoid being detected), we assume that the legitimate channel remains essentially unperturbed, apart from a unitary transform. That is, we model $\mathbf{H}$ as a random unitary matrix: $\mathbf{H} = \mathbf{U}$, with $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$; while Eve sees an MDL channel. Specifically, we model $\mathbf{H^e}$ as a concatenation of two operations – a rotation operation $\mathbf{U^e}$ followed by a scaling operation $\sqrt{\mathbf{V^e}}$. That is, $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$, where $\mathbf{U^e}$ is a random unitary matrix and $\mathbf{V^e}$ is a random diagonal matrix. We refer to $\mathbf{V^e}$ as the MDL matrix, with diagonal elements v_i, $i = 1, ..., M$. We specify the value of MDL as the ratio of maximum v_i to minimum v_i [9]. For $M > 2$, the MDL matrix is not uniquely defined by the MDL value. We therefore assume two statistical distribution models for the v_is [35]:

- *Uniformly distributed MDL* : The diagonal MDL matrix elements v_i are randomly drawn from a uniform distribution in $[\min\{v_i\}, \max\{v_i\}]$, subject to the specified MDL value of $\max\{v_i\}/\min\{v_i\}$ and the trace normalization $\sum_{i=1}^{M} v_i = M$.
- *Log-uniformly distributed MDL* : The diagonal MDL matrix elements, expressed in decibels as $10\log_{10}(v_i)$, are randomly drawn from a uniform distribution in $[\min\{10\log_{10}(v_i)\}, \max\{10\log_{10}(v_i)\}]$. The specified MDL value is expressed in decibels as $\text{MDL}_{\text{dB}} = 10\log_{10}(\max\{v_i\}/\min\{v_i\})$. Similarly, the v_is are subject to the (linear) trace normalization $\sum_{i=1}^{M} v_i = M$.

As an example, in Fig. 2 we plot the distribution of v_i (generated by 10^5 random realizations) for uniformly distributed and log-uniformly distributed MDL models with $M = 8$ and MDL=$\{5,\ 10,\ 20,\ 40\}$ dB. We observe that the differences between the two distribution models are more pronounced when MDL is large. The impact of these differences on secrecy capacity will be analyzed in Section IV. Other statistical models for MDL are discussed in [36], [37].

III. Secrecy Capacity of MIMO-SDM Systems

A. Secrecy Capacity Per Channel Realization

The results for the secrecy capacity of MIMO-SDM systems for given channel realizations $\mathbf{H}$ and $\mathbf{H^e}$ provide a baseline for characterizing the secrecy capacity of our optical MIMO-SDM systems without CSI at the transmitter. We use the framework developed in [20]-[23] and incorporate the physical characteristics of the optical MIMO-SDM channel into the constraints. As mentioned above, the power in optical SDM systems is constrained by fiber nonlinearities on a *per mode* basis. As such we set an upper bound $\overline{P}$ for the power of each mode individually. It follows directly from Eq. (9) in [23] that the secrecy capacity is achieved when $\mathbf{x}$ is a circularly symmetric complex Gaussian with zero mean and covariance $\mathbf{Q_x} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$, such that $\mathbb{E}[|x_i|^2] \le \overline{P}$. Furthermore, the secrecy capacity C_s is given by:

$$\begin{aligned} C_s = &\max_{\mathbf{Q_x}=\mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]} \log_2\left[\det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{Q_x}\mathbf{H}^\dagger)\right] \\ &- \log_2\left[\det(\mathbf{I} + \text{SNR}^e\mathbf{H^e}\mathbf{Q_x}\mathbf{H^e}^\dagger)\right] \\ &\text{s.t.}: \ \mathbb{E}[|x_i|^2] \le \overline{P},\ i \in 1, ..., M, \end{aligned} \tag{3}$$

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

$$v_1^* = \frac{(\mathrm{MDL}-1)^2 - M(\mathrm{MDL}^2 - 2\mathrm{SNR}^e\mathrm{MDL}+1) + \sqrt{4M^2\mathrm{SNR}^e\mathrm{MDL}(1+\mathrm{MDL})^2 + [M-1+\mathrm{MDL}(2+\mathrm{MDL}(M-1)-2\mathrm{SNR}^eM))^2]}}{2M\mathrm{SNR}^e\mathrm{MDL}(1+\mathrm{MDL})} \quad (7)$$

$$v_2^* = \mathrm{MDL}v_1^* \quad (8)$$

$$v_i^* = \frac{M - v_1^*(1+\mathrm{MDL})}{M-2},\ i = 3,4,...,M. \quad (9)$$

where $\mathrm{SNR} = LE_0/N_0$ and $\mathrm{SNR}^e = L^eE_0/N_0^e$ are the mode-averaged signal-to-noise ratios of the legitimate and eavesdropping channels, respectively; det denotes the determinant of a matrix. Further, according to our optical MIMO-SDM channel models ($\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$), we can reduce the solution space of $\mathbf{Q_x}$ by optimizing over only the diagonal covariance matrices $\mathbf{\Lambda_x}$ [14], [38]. It suffices to observe that since $\mathbf{Q_x}$ is non-negative definite, it can be decomposed (via a singular value decomposition) as $\mathbf{Q_x} = \mathbf{W}\mathbf{\Lambda_x}\mathbf{W}^\dagger$, where $\mathbf{W}$ is a unitary matrix and $\mathbf{\Lambda_x}$ is a diagonal matrix. $\widetilde{\mathbf{U}} = \mathbf{UW}$ and $\widetilde{\mathbf{U}}^\mathbf{e} = \mathbf{U^eW}$ are also unitary. Therefore, all the statistical characteristics of $\mathbf{Q_x}$ are fully captured by $\mathbf{\Lambda_x}$ as a result of the underlying channel model. We summarize this result as follows:

Result 1: For given channel realizations $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$, the secrecy capacity of the optical MIMO-SDM channel is achieved by a circularly symmetric complex Gaussian input, with zero-mean and diagonal covariance $\mathbf{\Lambda_x}$, such that $\mathbb{E}[|x_i|^2] \le \overline{P}$. The secrecy capacity is given by:

$$\begin{aligned} C_s = &\max_{\mathbf{\Lambda_x}} \log_2\left[\det(\mathbf{I} + \mathrm{SNR}\mathbf{U}\mathbf{\Lambda_x}\mathbf{U}^\dagger)\right] \\ &- \log_2\left[\det(\mathbf{I} + \mathrm{SNR}^e\sqrt{\mathbf{V^e}}\mathbf{U^e}\mathbf{\Lambda_x}\mathbf{U^{e\dagger}}\sqrt{\mathbf{V^e}})\right] \\ &\text{s.t.}: \ \mathbb{E}[|x_i|^2] < \overline{P},\ i \in 1,...,M. \end{aligned} \quad (10)$$

B. Guaranteed Secrecy Capacity

When the random channel realizations $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$ are unknown at the transmitter, we cannot predict which of the transmitting modes will be extracted stronger than others by the eavesdropper due to mode coupling within the SDM fiber. As a result, the secrecy capacity of the MIMO-SDM channel is randomly distributed, as illustrated by a numerically simulated histogram of the secrecy capacity C_s in Fig. 3 for the case of $M = 8$ modes, SNR=SNR^e=20 dB, and a uniformly distributed MDL model with an MDL vaule of 20 dB, based on 10^5 random realizations both for the legitimate and the eavesdropping channel. Here C_s is normalized to C_0, the capacity per mode of the legitimate channel, with $C_0 = \log_2(1+\mathrm{SNR})$. We observe a cutoff on the left side of the histogram. Transmission at a rate lower than this cutoff capacity (indicated by the black dashed line in Fig. 3) will be perfectly secure regardless of the channel realization. We refer to this cutoff capacity as *guaranteed secrecy capacity* and denote it as C_s^{grt}. In the following, we characterize the guaranteed secrecy capacity.

We consider only the uniformly distributed MDL model introduced in Section II.B. The analysis can be extended to the log-uniformly distributed model by a simple change of variables [15]. With an identical power allocation on all the modes ($\mathbf{P} = P_0\mathbf{I}$), the capacity of the legitimate channel is given by MC_0 due to our assumption of a unitary legitimate channel. The capacity of the eavesdropper's channel is:

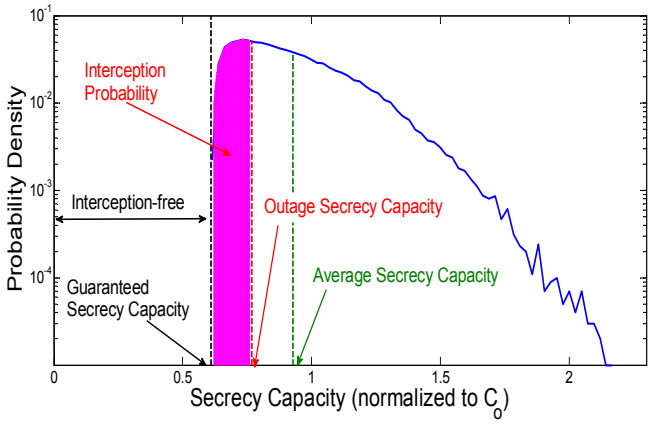

Fig. 3. Histogram of the secrecy capacities (normalized to C_0) generated from 10^5 random realizations, for the uniformly distributed MDL model.

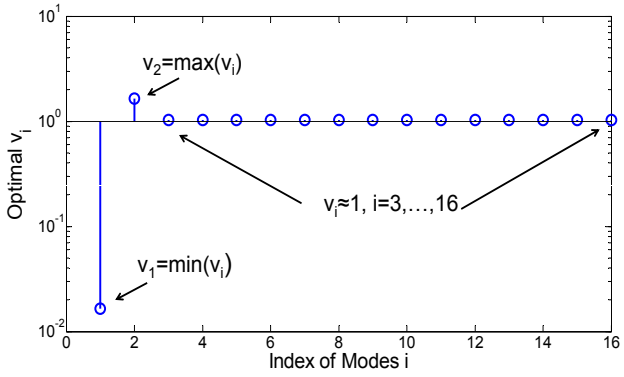

Fig. 4. The optimal values of v_is that achieve the guaranteed secrecy capacities, with $M = 16$, SNR=SNR^e=20 dB, and MDL=20 dB.

$$C_e = \sum_{i=1}^{M} \log_2(1 + \mathrm{SNR}^e v_i). \quad (5)$$

Note that for a given SNR^e and MDL, C_e depends only on the values of v_i. As such, we can obtain an upper bound of

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

C_e by maximizing over the choices of v_i:

$$\begin{aligned} \max_{v_i}: \quad & C_e = \sum_{i=1}^{M} \log_2(1+\text{SNR}^e v_i) \\ \text{s.t.}: \quad & \sum_{i=1}^{M} v_i = M, \\ & v_2 = \text{MDL} v_1, \\ & v_1 \le v_i \le v_2, \quad i = 3,4,...,M. \end{aligned} \tag{6}$$

Here, without loss of generality, we let $v_1 = \min\{v_i\}$ and $v_2 = \max\{v_i\}$. Since the objective function is concave and the constraints are linear and bounded, a unique maximum exists [39]. Let $C_e^{\max}$ be the maximum capacity achieved by the eavesdropper. We then have the guaranteed secrecy capacity as: $C_s^{grt} = MC_0 - C_e^{\max}$, when $MC_0 > C_e^{\max}$.

Using the technique of Lagrange multipliers, we can exactly solve for the optimal v_is (denoted as v_i^*) as given in (7) - (9). The maximum capacity of the eavesdropping channel is thus $\sum_{i=1}^{M} \log_2(1+\text{SNR}^e v_i^*)$. Fig. 4 shows an example of optimal v_is for the case of $M = 16$, MDL=20 dB, and SNR = SNR^e=20 dB.

The results in (7) - (9) are applicable to any SNR^e and MDL values. For high SNR^e (20 dB and above) and MDL in the range of 5 to 40 dB, we can approximate the optimal v_i to much simpler forms [15]:

$$\begin{aligned} v_1 &\approx \frac{2}{1+\text{MDL}}, \\ v_2 &\approx 2\frac{\text{MDL}}{1+\text{MDL}}, \\ v_i &\approx 1, \quad i = 3,4,...,M. \end{aligned} \tag{10}$$

We then obtain

$$\begin{aligned} C_s^{grt} \approx\ & 2C_0 - \log_2(1+\frac{2}{1+\text{MDL}}\text{SNR}^e) \\ & - \log_2(1+\frac{2\text{MDL}}{1+\text{MDL}}\text{SNR}^e) \\ & + (M-2)\log_2\frac{\text{SNR}}{\text{SNR}^e}. \end{aligned} \tag{11}$$

Equation (11) indicates that MDL, SNR, SNR^e, and the SNR-to-SNR^e ratio all contribute to the guaranteed secrecy capacity. However, only the contribution from the SNR-to-SNR^e ratio scales with the number of modes M, as shown by the last term of (11). If the legitimate and the eavesdropping channels have the same SNR, C_s^{grt} can be further simplified to:

$$\begin{aligned} C_s^{grt} \approx\ & 2C_0 - \log_2(1+\frac{2}{1+\text{MDL}}\text{SNR}) \\ & - \log_2(1+\frac{2\text{MDL}}{1+\text{MDL}}\text{SNR}). \end{aligned} \tag{12}$$

In this case, the guaranteed secrecy capacity is independent of the number of modes M, thus does not change as M increases. That is, designing the entire system based on the worst-case scenario could yield very poor scalability for the secrecy capacity. We also note that the guaranteed secrecy capacity derived for the uniformly distributed MDL model also applies to the log-uniformly distributed model [15]. In other words, the optimal v_is and the associated C_s^{grt} depend only on the MDL vaule and are independent of the MDL distributions.

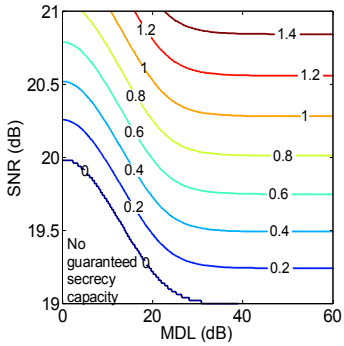

Fig. 5. C_s^{grt} as a function of MDL and SNR, with $M = 16$ and SNR$^e = 20$ dB.

It is also of interest to analyze the range of parameters under which a guaranteed secrecy capacity exists, i.e., $C_s^{grt} > 0$. When SNR $=$ SNR$^e \ge 20$ dB, it can be shown trivially that C_s^{grt}, as expressed in Eq. (12), is strictly larger than zero for MDL > 0 dB. For general M, MDL, SNR, and SNR^e, we numerically evaluate the region where $C_s^{grt} > 0$. For the case of $M = 16$ and SNR$^e = 20$ dB, Fig. 5 shows C_s^{grt} as a function of MDL and SNR in the form of a contour plot. Fig. 5 shows that C_s^{grt} increases with both MDL and SNR. More importantly, the contour line labeled with 0 marks the boundary of the MDL-SNR region where C_s^{grt} is strictly positive. In other words, points of MDL-SNR pairs that lie above this boundary line guarantee strictly positive secrecy capacities. We note that with sufficiently large MDL, guaranteed secrecy capacity exists even if the SNR^e is higher than SNR. This is in contrast to the result of the SISO Gaussian wiretap channel [17], for which the secrecy capacity is strictly positive only when the eavesdropping channel is SNR-degraded (SNR$^e <$ SNR). The impact of the eavesdropper's MDL on the security of MIMO-SDM systems will be discussed in detail in Section IV.

We summarize the main result on the guaranteed secrecy capacity as follows:

Result 2: With equal power allocation, the guaranteed secrecy capacity is given by:

$$C_s^{grt} = \begin{cases} MC_0 - C_e^{\max}, & \text{if } MC_0 > C_e^{\max}; \\ 0, & \text{otherwise,} \end{cases} \tag{13}$$

where $C_e^{\max}$ is obtained by solving (6).

C. Outage Secrecy Capacity

For slowly-varying frequency-flat channels, the channel characteristics are constant across the signal bandwidth and change very slowly over time compared to the block length of the underlying coding scheme (cf. Fig. 6 (a)). Since each coding block experiences its own random channel realization, some blocks will be perfectly decodable while others will not be. Thus, the system lends itself naturally to *outage* analyses.

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

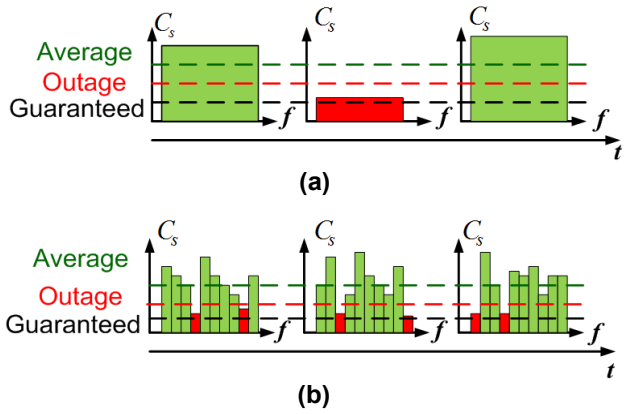

Fig. 6. (a) Visualization of slowly-varying temporal evolution of a frequency-flat channel; and (b) of a strongly frequency-selective channel. The green boxes indicate that the channel secrecy capacities are above the transmitted capacity, while the red boxes indicate that the secrecy capacities are below the transmitted capacity, where interception may in principle occur.

If the transmitter chooses to communicate at a rate R higher than the cutoff (as shown by the vertical and horizontal red dashed lines in Fig. 3 and Fig. 6 (a), respectively), there is a finite probability (as shown by the shaded area in Fig. 3) that the eavesdropper can, at least in principle, learn something about the secret information. We refer to this probability as the *interception probability* $p_{int}(R)$. When $\mathbf{H}$, $\mathbf{H^e}$, and $\mathbf{\Lambda_x}$ are given, we have

$$\begin{aligned} p_{int}(R) &= \mathbb{P}[\log_2\left(\det(\mathbf{I}+\text{SNR}\mathbf{H}\mathbf{\Lambda_x}\mathbf{H}^\dagger)\right) \\ &- \log_2\left[\det(\mathbf{I}+\text{SNR}^e\mathbf{H^e}\mathbf{\Lambda_x}\mathbf{H}^{e\dagger}) < R\right]. \end{aligned} \tag{14}$$

The interception probability can be minimized by choosing an optimal power allocation scheme (covariance matrices $\mathbf{\Lambda_x}$):

$$\begin{aligned} \min_{\mathbf{\Lambda_x}}: &\quad p_{int}(R) \\ \text{s.t.}: &\quad \mathbb{E}[|x_i|^2] < \overline{P}. \end{aligned} \tag{15}$$

The interception probability, as expressed in (15), is a very complicated function of $\mathbf{\Lambda_x}$ and R and is hard to solve analytically. In [14] we conducted extensive numerical evaluations of $p_{int}(R)$ as a function of different power allocation schemes and conjectured that the power allocation scheme in the form of $\mathbf{\Lambda_x} = \overline{P}\mathbf{I}$ is optimal. That is, given that Bob or Eve's CSI is unavailable for Alice (all modes are thus statistically equivalent) and the power is constrained on a per-mode basis, sending uncorrelated signals of equal power $\overline{P}$ on all modes is likely to incur the lowest interception probability.

With equal power allocation of $\overline{P}$ on all the modes and hence without making use of any CSI at the transmitter, the expression for the secrecy capacity (10) for given channel instantiations $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$ can be simplified to:

$$C_s = \sum_{i=1}^{M}[\log_2(1+\text{SNR}) - \log_2(1+\text{SNR}^e v_i)]. \tag{16}$$

We further normalize C_s by the capacity per mode of the legitimate channel C_0 and arrive at

$$\frac{C_s}{C_0} = M - \frac{\sum_{i=1}^{M}\log_2(1+\text{SNR}^e v_i)}{\log_2(1+\text{SNR})}. \tag{17}$$

This is the maximum rate, in units of the raw SDM channel capacity per mode, that can be transmitted in perfect information-theoretic secrecy over a particular MIMO-SDM channel instantiation. We refer to the maximum achievable secrecy rate such that the interception probability is less than ϵ as the *outage secrecy capacity* with interception probability ϵ [18]. That is,

$$p_{int}(C_s^{out}(\epsilon)) = \epsilon. \tag{18}$$

The operational meaning of the outage secrecy capacity lies in the trade-off between secrecy rate and interception probability: a higher secrecy rate can be achieved at the expense of a higher interception probability. We summarize the main result on the outage secrecy capacity as follows:

Result 3: For $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$, transmitting uncorrelated signals of equal power $\overline{P}$ on all the modes will incur the smallest interception probability. The interception probability for a secrecy rate R is given by:

$$\begin{aligned} p_{int}(R) &= \mathbb{P}[\sum_{i=1}^{M}[\log_2(1+\text{SNR}) \\ &- \log_2(1+\text{SNR}^e v_i)] < R]. \end{aligned} \tag{19}$$

For an interception probability ϵ, the outage secrecy capacity C_s^{out} can be obtained by solving:

$$\mathbb{P}[\sum_{i=1}^{M}[\log_2(1+\text{SNR}) - \log_2(1+\text{SNR}^e v_i)] < C_s^{out}] = \epsilon. \tag{20}$$

D. Average Secrecy Capacity

For rapidly-varying frequency-flat channels, the channel characteristics are constant across the signal bandwidth, but change rapidly over time. We can average many channel realizations by coding over long time intervals and a secrecy rate of $C_s^{avg} = \langle C_s \rangle$ can be achieved. We refer to this capacity as *average secrecy capacity*. Equivalently, the average secrecy capacity can be obtained for a frequency-selective channel, where the channel characteristics vary rapidly across the signal's bandwidth, as shown in Fig. 6 (b). If each signal frequency component experiences a different channel instantiation, the system ultimately sees the average capacity (as shown by the vertical and horizontal green dashed lines in Fig. 3 and Fig. 6 (b), respectively). In the idealized and limiting case of a highly frequency-selective channel, the average secrecy capacity can be guaranteed assuming the appropriate code is available. That is, if the transmitter codes for C_s^{avg}, the legitimate receiver can extract information at a rate C_s^{avg}, fundamentally without the possibility of being intercepted. Let R_s^{avg} denote the average secrecy rate. When $\mathbf{\Lambda_x}$ is given, R_s^{avg} can be expressed as:

$$\begin{aligned} R_s^{avg} &= \mathbb{E}[\log_2\left(\det(\mathbf{I}+\text{SNR}\mathbf{H}\mathbf{\Lambda_x}\mathbf{H}^\dagger)\right) \\ &- \log_2(\det(\mathbf{I}+\text{SNR}^e\mathbf{H^e}\mathbf{\Lambda_x}\mathbf{H}^{e\dagger}))], \end{aligned} \tag{21}$$

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

where the expectation is taken over the random realizations of $\mathbf{H}$ and $\mathbf{H^e}$. Further, by maximizing over the choices of $\mathbf{\Lambda_x}$, we can obtain the average secrecy capacity C_s^{avg}. That is,

$$\begin{aligned} \max_{\mathbf{\Lambda_x}}: \quad & R_s^{avg} \\ \text{s.t.}: \quad & \mathbb{E}[|x_i|^2] < \overline{P}. \end{aligned} \tag{22}$$

We show that $\mathbf{\Lambda_x} = \overline{P}\mathbf{I}$ is also a covariance matrix that achieves average secrecy capacity, by using the fundamental relationship between the mutual information and the minimum mean-square error (MMSE) in Gaussian channels [40], and the properties of Schur concavity [41]. The details of the proof are provided in the appendix. We summarize the main result on the average secrecy capacity as follows:

Result 4: For $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \sqrt{\mathbf{V^e}}\mathbf{U^e}$, transmitting uncorrelated signals of equal power $\overline{P}$ on all modes will achieve the average secrecy capacity, which is given by:

$$C_s^{avg} = \mathop{\mathbb{E}}_{v_i}[\sum_{i=1}^{M}[\log_2(1+\text{SNR}) - \log_2(1+\text{SNR}^e v_i)]]. \tag{23}$$

IV. Results and Discussion

The analyses in Section III show that the various definitions of secrecy capacities depend on several system parameters, such as the number of modes, the distribution of MDL matrices, and the SNRs of the legitimate and eavesdropping channels. In this section, we evaluate the impact of each of these system parameters. Since closed form expressions for both the outage and average secrecy capacity (as specified by Eq. (20) and (23), respectively) are difficult to obtain (except for the simple case of $M = 2$), we mostly resort to a numerical approach. That is, for a given set of parameters M, MDL, SNR and SNR^e, we first run a simulation that generates 10^5 random channel realizations of $\mathbf{H} = \mathbf{U}$ and $\mathbf{H^e} = \mathbf{U^e}\sqrt{\mathbf{V^e}}$. For each channel realization, we calculate the corresponding C_s by using (16). Based on the statistics generated by these instantiations, we then calculate outage (for a given p_{int} of 10^{-4}) or average secrecy capacity by using Eq. (20) or (23), respectively. To evaluate the guaranteed secrecy capacity for a given set of parameters M, MDL, SNR and SNR^e, we solve the optimization problems of (6) numerically to obtain C_s^{grt}.

A. Impact of the Number of Modes

In Fig. 7 we plot the average, outage (with $p_{int} = 10^{-4}$), and guaranteed secrecy capacities (normalized to C_0) as functions of the number of modes M for the case of MDL = $\{20, 50\}$ dB and SNR = SNR^e=20 dB. Here we use the uniformly distributed MDL model. We observe that, when the main and eavesdropping channels have the same SNR, the guaranteed secrecy capacity is independent of the number of modes, as indicated in Section III. B. In contrast to the guaranteed secrecy capacity, Fig. 7 shows that the average secrecy capacities increase proportional to M and the outage secrecy capacities increase approximately proportional to M when $M > 20$. Therefore, coding at the average capacity or allowing a small interception probability can take advantage of the scale-up of the number of modes of MIMO-SDM systems.

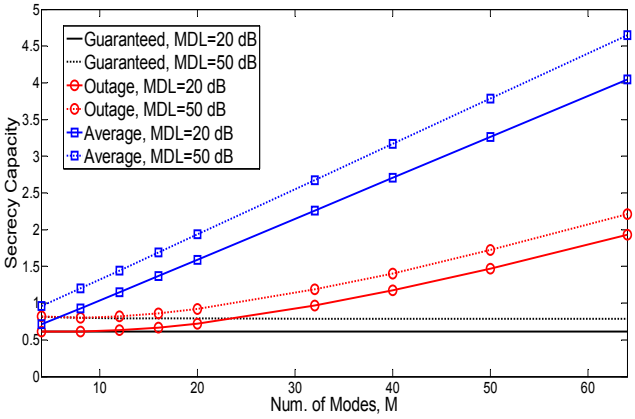

Fig. 7. Secrecy capacity (normalized to C_0) vs. number of modes M, with MDL=$\{20, 50\}$ dB and SNR = SNR^e = 20 dB, for the uniformly distributed MDL model.

B. Impact of the MDL Models

We now study the impact of the value of the eavesdropper's MDL. For the uniformly distributed MDL model, Fig. 8 shows the outage secrecy capacity (with $p_{int} = 10^{-4}$) and the average secrecy capacity, all normalized to C_0, as functions of MDL. We assume that both legitimate receiver and eavesdropper have the same SNR of 20 dB. For $M = 4$ and $M = 64$, the secrecy capacities quickly saturate as MDL increases beyond about 25 dB. In the context of designing a "secure" SDM waveguide, this result implies that it is sufficient to ensure that noticeable evanescent coupling at a fiber bend induces at least 25 dB of MDL. For $M = 4$, the outage and average secrecy capacities are fairly close, while for $M = 64$ the average secrecy capacities are much higher than outage secrecy capacities.

The results shown in Fig. 7 and Fig. 8 demonstrate the potential of SDM systems in supporting very high secure data rates. Taking an SDM system with $M = 4$ as an example, an MDL of 20dB can yield a secure data rate of approximately 60% of the per-mode capacity shown by Fig. 8. Even if the eavesdropper merely introduces 5 dB of MDL, we can still have about 7% of the per-mode capacity in secrecy, as indicated by the arrow in Fig. 8. Using just a single wavelength channel modulated at 100 Gb/s per spatial mode already yields secure data rates on the scale of Gb/s, which are orders of magnitude higher than what is achievable through QKD. The secrecy capacity can be further increased by using more spatial modes and rendering less favorable conditions for the eavesdropper.

We next evaluate the impact of MDL distribution models. Specifically, we compare the outage secrecy capacity and the average secrecy capacity generated by uniformly and log-uniformly distributed v_is of the MDL matrix $\mathbf{V}$. In Fig. 9 (a) and (b), we plot the capacities generated by log-uniformly and uniformly distributed models as a function of MDL, for

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

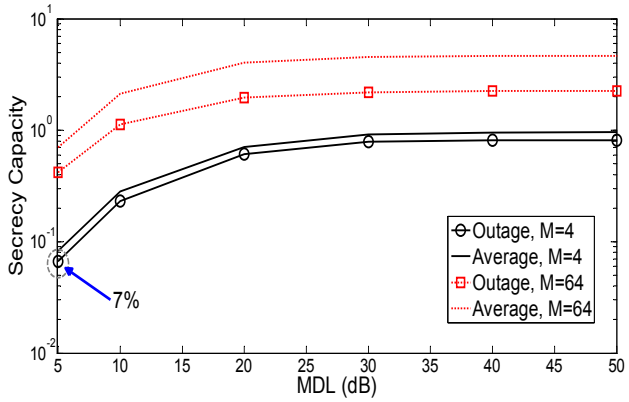

Fig. 8. Secrecy capacity (normalized to C_0) vs. MDL, SNR=SNR^e=20 dB. Uniformly distributed MDL model is used.

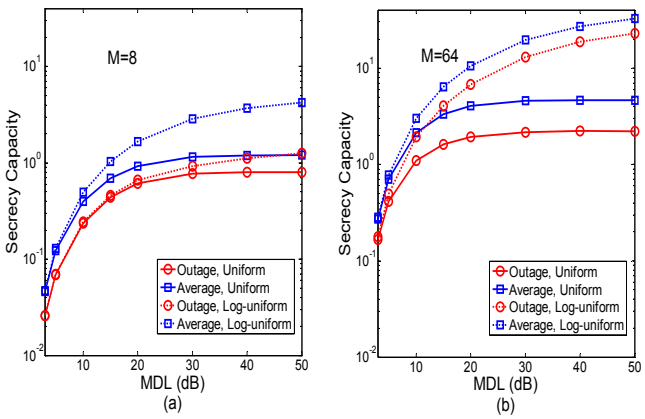

Fig. 9. Difference of secrecy capacities generated by log-uniformly and uniformly distributed models vs. MDL for $M = 8$ (a) and $M = 64$ (b).

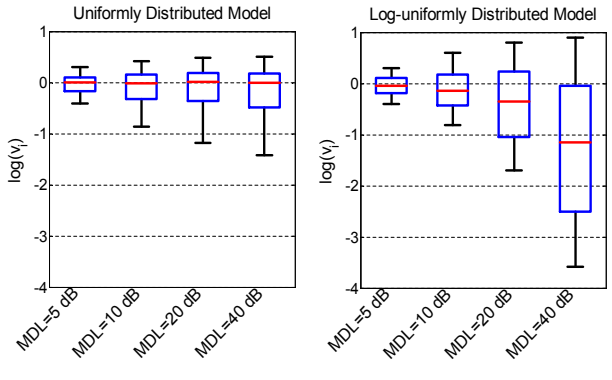

Fig. 10. Box plots for uniformly (left panel) and log-uniformly (right panel) distributed MDL models.

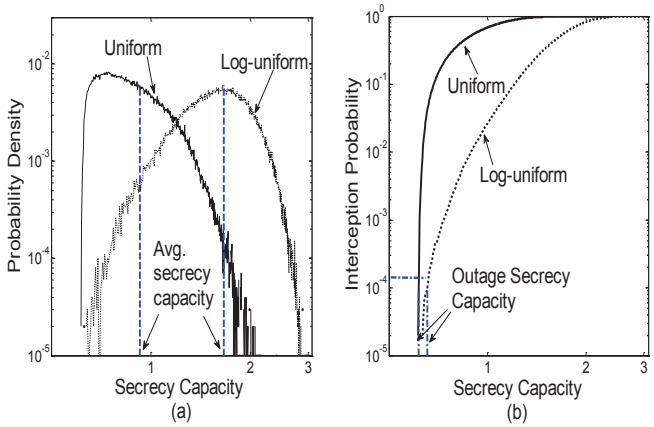

Fig. 11. (a) Histograms of secrecy capacities for uniformly and log-uniformly distributed models; (b) interception probability vs. secrecy capacity for uniformly and log-uniformly distributed models. The secrecy capacity is normalized to C_0. We assume $M = 8$, MDL = 20 dB, and SNR = SNR^e = 20 dB.

$M = 8$ and $M = 64$, respectively. Here, we assume that SNR = SNR^e = 20 dB. When MDL $\leq$ 10 dB, the differences between the two models are insignificant for both $M = 8$ and $M = 64$. However, for MDL $>$ 10 dB, the secrecy capacities of log-uniformly distributed MDL appear higher. In addition, the gap between the capacities generated by the two models increases with increasing MDL and M. To explain this, we show in Fig. 10 the distributions of v_i in the form of box plots for both uniform and log-uniform MDL models. The figure shows that for MDL $\leq$ 10 dB, the difference between the two distribution models, indicated by 25%, 50% (median), and 75% percentile, is less obvious. However, for MDL $\geq$ 20 dB the difference becomes significant. As the value of MDL increases, the median of v_is generated by the uniform MDL model stays almost unchanged at unity, Swhile the median (as well as 75% percentile) of v_is generated by the log-uniform MDL model decreases significantly. In other words, the random v_is generated by the log-uniform model are skewed towards much lower values for large MDL (e.g., MDL = 40 dB). To better understand how the difference between the two MDL models affects the statistical characteristics of the secrecy capacity, we compare the histograms of the secrecy capacities (Fig. 11 (a)) and interception probability vs. secrecy capacity curves (Fig. 11 (b)) for the two models. Here we assume $M = 8$, MDL = 20 dB, and SNR = SNR^e = 20 dB. As shown in the plots, the distribution generated by the log-uniform model is biased towards higher secrecy capacities. As such, the log-uniformly distributed model yields higher average and outage secrecy capacities.

C. Impact of the SNR and SNR^e

We first consider the case where SNR is higher than SNR^e. The guaranteed capacity increases with the number of modes M, as shown in Fig. 12. As the values of M and the SNR of the legitimate channel increase, the guaranteed secrecy capacity is dominated more by the contribution of the SNR differences between the main and the eavesdropping channel and less by the contribution from MDL (cf. (11)). Motivated by the possibility that the eavesdropper's receiver could be located very close to the transmitter and thus could experience higher SNR^e than the legitimate receiver, we evaluate the impact of an eavesdropper's higher SNR^e on the secrecy capacity. Fig. 13 and Fig. 14 illustrate the outage secrecy capacity (normalized

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

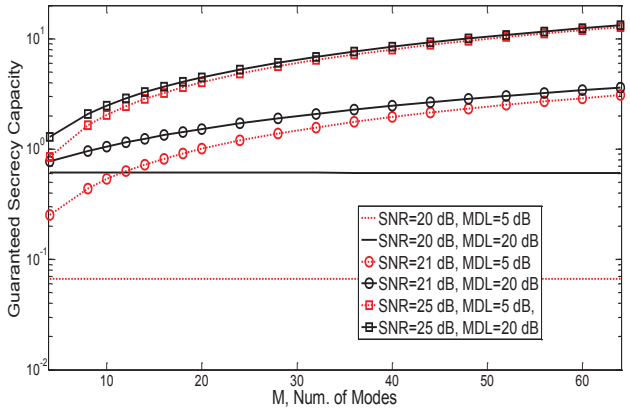

Fig. 12. Guaranteed secrecy capacities (normalized to C_0) vs. M for different transmitter SNRs, with SNR^e=20 dB.

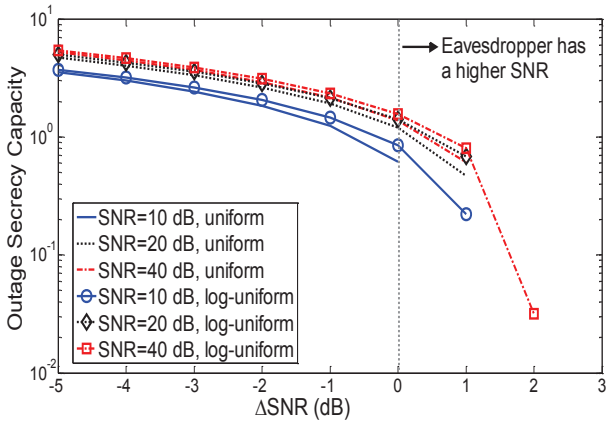

Fig. 13. Outage secrecy capacities (normalized to C_0) vs. ΔSNR for uniformly and log-uniformly distributed models, with $M = 8$ and MDL = 20 dB. The interception probability p_{int} is 10^{-3}. C_0 is evaluated at SNR $= 10$ dB.

to C_0) as a function of $\Delta\text{SNR} = \text{SNR}^e - \text{SNR}$ (in dB) for uniform and log-uniform channel models with MDL=20 dB and MDL=40 dB, respectively. In these two plots, the SDM system has 8 modes ($M = 8$) and the legitimate receiver has an SNR of $[10, 20, 40]$ dB. We set the interception probability p_{int} at 10^{-3} and evaluate C_0 at SNR $= 10$ dB. We observe that higher SNR and higher MDL in general enhance the security, even if SNR^e is higher than SNR. The figures also show that the log-uniform model results in better resilience against higher SNR^e than the uniform model does.

Fig. 15 and Fig. 16 illustrate the average secrecy capacity (normalized to C_0) as a function of ΔSNR, for the same sets of parameters used in Fig. 13 and Fig. 14, respectively. We observe a similar trend that higher SNRs and higher MDLs in general enhance the security. Consistent with the results obtained so far, the average secrecy capacity can tolerate much higher ΔSNR than the outage secrecy capacity does.

These results show that MDL plays an important role in enhancing the robustness of the security. With large MDLs, strictly positive secrecy capacity can still be achieved even if the eavesdropper has a higher SNR^e. In comparison, the

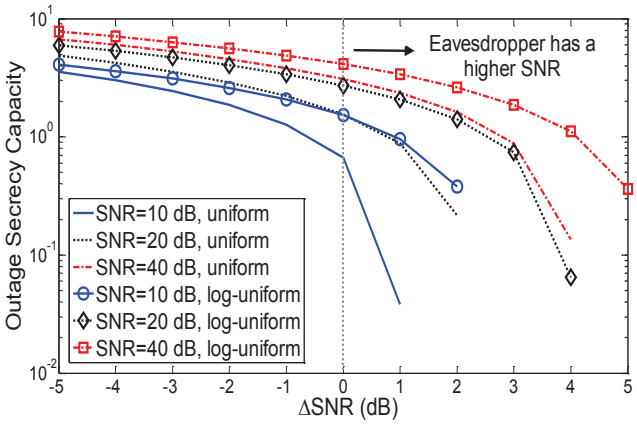

Fig. 14. Outage secrecy capacities (normalized to C_0) vs. ΔSNR for uniformly and log-uniformly distributed models, with $M = 8$ and MDL = 40 dB. The interception probability p_{int} is 10^{-3}. C_0 is evaluated at SNR $= 10$ dB.

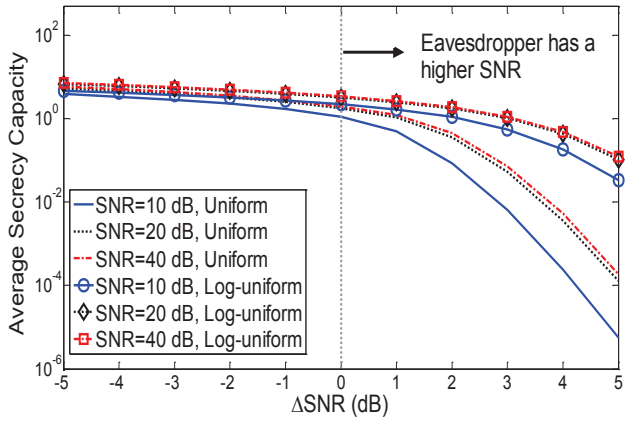

Fig. 15. Average secrecy capacities (normalized to C_0) vs. ΔSNR for uniformly and log-uniformly distributed models, with $M = 8$ and MDL = 20 dB. C_0 is evaluated at SNR $= 10$ dB.

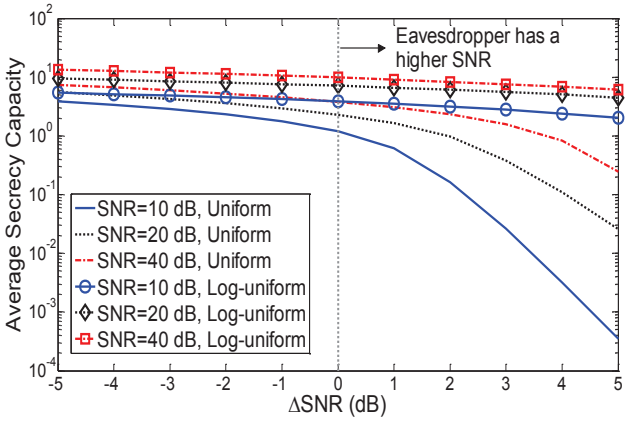

Fig. 16. Average secrecy capacities (normalized to C_0) vs. ΔSNR for uniformly and log-uniformly distributed models, with $M = 8$ and MDL = 40 dB. C_0 is evaluated at SNR $= 10$ dB.

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

secrecy capacity of a Gaussian wiretap channels is strictly positive only when the eavesdropping channel has a lower SNR [17].

V. Conclusions

Fiber wire-tapping by coupling spatial information out of a MIMO-SDM waveguide through bending leads to inherent changes in the spatial information content, for both eavesdropper and legitimate transmitter-receiver pair. As a result, the MIMO channel of the eavesdropper will be generally less favorably conditioned than that of the legitimate user; at the same time, a bend-induced MDL recorded at the legitimate receiver reveals the possible presence of the eavesdropper. As such, MIMO-SDM has the potential of providing a provably secure high-capacity medium of transmission. In this work, we evaluated the security benefits of MIMO-SDM by using equivocation as a measure of secrecy. Our results show that the secrecy capacity achieved by an SDM system can potentially be orders of magnitude higher than what can be offered by QKD. In addition, SDM systems can provide fundamental security even if an eavesdropper has a higher SNR than a legitimate receiver.

Evaluating the secrecy capacity only represents the first step towards designing a secure MIMO-SDM system. We note that the secrecy capacity, like the Shannon limit, provides an information theoretic limit without a prescription on the codes that may actually approach it. It is consequently also important to examine the practical implementation of secure MIMO-SDM systems, especially the design of secure modulation/coding schemes. In addition to equivocation, we also evaluate *rate-distortion* as an alternative and more practical metric for secrecy [42].

VI. Appendix

In this section, we prove that $\mathbf{Q_x} = \mathbf{I}$ maximizes the average secrecy rate $\mathbb{E}[I(x,y) - I(x,y^e)]$.

First, we show that we can greatly reduce the solution space of the optimization problem (of $\mathbb{E}[I(x,y) - I(x,y^e)]$) by considering only diagonal covariance matrices and using the properties of unitary matrices $\mathbf{\Lambda^Q}$ [38]. In other words, all the statistical characteristics of $\mathbf{Q_x}$ are preserved in $\mathbf{\Lambda^Q}$. As such, the received signals of the legitimate receiver and the eavesdropper can be equivalently expressed as:

$$\mathbf{y} = \sqrt{\gamma}\sqrt{\mathbf{\Lambda^Q}}\mathbf{x} + \mathbf{n}, \quad (24)$$

$$\mathbf{y}^e = \sqrt{\gamma^e}\sqrt{\mathbf{V^e}}\widetilde{\mathbf{U}}^{\mathbf{e}}\sqrt{\mathbf{\Lambda^Q}}\mathbf{x} + \mathbf{n^e}. \quad (25)$$

Next, we show that $\mathbb{E}[I(\mathbf{x},\mathbf{y}) - I(\mathbf{x},\mathbf{y}^e)]$ is an increasing function of λ_i^Q (diagonal elements of the covariance matrix $\mathbf{\Lambda^Q}$). Using the fundamental relationship between the mutual information and the minimum mean-square error in Gaussian channels [40], we have:

$$\nabla_{\mathbf{\Lambda^Q}} I(\mathbf{x},\mathbf{y}) = \gamma\left(\mathbf{I} + \gamma\ \mathbf{\Lambda^Q}\right)^{-1}, \quad (26)$$

$$\nabla_{\mathbf{\Lambda^Q}} I(\tilde{\mathbf{x}},\tilde{\mathbf{y}^e}) = \gamma^e\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}\dagger}\left(\mathbf{I} + \gamma^e\ \mathbf{\Lambda^Q}\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}\dagger}\right)^{-1} \quad (27)$$

Since $\tilde{\mathbf{U}}^{\mathbf{e}}$ is a Haar matrix, $\nabla_{\mathbf{\Lambda^Q}}\mathbb{E}[I(\tilde{\mathbf{x}},\tilde{\mathbf{y}^e})]$ is concave and symmetric with respect to $\mathbf{V^e}$, and thus Schur concave. Using the fact that [41, Theorem 4.3.26]:

$$[(\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V}\tilde{\mathbf{U}}^{\mathbf{e}\dagger})_{11},\ldots,(\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V}\tilde{\mathbf{U}}^{\mathbf{e}\dagger})_{MM}] \preceq [v_1,\ldots,v_M] \quad (28)$$

and Jensen's inequality, we arrive at:

$$\begin{aligned}
&\nabla_{\mathbf{\Lambda^Q}}\mathbb{E}[I(x,y) - I(x,y^e)] \\
=\ &\gamma\left(\mathbf{I} + \gamma\ \mathbf{\Lambda^Q}\right)^{-1} \\
&-\gamma^e\mathbb{E}\left[\tilde{\mathbf{U}}^{\mathbf{e}\dagger}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}}\left(\mathbf{I} + \gamma^e\ \mathbf{\Lambda^Q}\tilde{\mathbf{U}}^{\mathbf{e}\dagger}\mathbf{V}\tilde{\mathbf{U}}^{\mathbf{e}}\right)^{-1}\right] \\
\succeq\ &\gamma\left(\mathbf{I} + \gamma\ \mathbf{\Lambda^Q}\right)^{-1} \\
&-\gamma^e\ \mathbb{E}\left[\mathrm{diag}(\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}\dagger})\right]\left(\mathbf{I} + \gamma^e\ \mathbf{\Lambda^Q}\ \mathbb{E}\left[\mathrm{diag}(\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}\dagger})\right]\right)^{-1} \quad (29)
\end{aligned}$$

Finally, using the fact that:

$$\mathbb{E}[\tilde{\mathbf{U}}^{\mathbf{e}}\mathbf{V^e}\tilde{\mathbf{U}}^{\mathbf{e}\dagger}]) = \frac{1}{M}\mathbb{E}\left[\mathbf{I}\cdot\mathrm{tr}(\mathbf{V^e})\right] = \mathbf{I} \quad (30)$$

and substituting (30) into (29)), we have:

$$\nabla_{\mathbf{\Lambda^Q}}\mathbb{E}[I(x,y) - I(x,y^e)] \succeq \mathbf{0}, \quad (31)$$

which shows that $\mathbb{E}[I(x,y) - I(x,y^e)]$ is indeed an increasing function of λ_i^Qs.

The last step is straight forward. Since the power is constrained on a per mode basis ($0 \geq \lambda_i^Q \leq 1$) and $\mathbb{E}[I(x,y) - I(x,y^e)]$ is an increasing function of λ_i^Qs, it follows that $\lambda_i^Q = 1$ maximizes $\mathbb{E}[I(x,y) - I(x,y^e)]$. We thus prove that $\mathbf{Q_x} = \mathbf{\Lambda^Q} = \mathbf{I}$ achieves the average secrecy capacity.

References

[1] K. Shaneman and S. Gray, "Optical network security: technical analysis of fiber tapping mechanisms and methods for detection and prevention," in *Military Communications Conference*, Monterey, CA, 2004, vol. 2, pp.711-716.

[2] M. Medard, *et al.*, "Security issues in all-optical networks," *IEEE Netw.*, vol. 11, no. 3, pp. 42-48, May/June, 1997.

[3] V. Scarani, *et al.*,"The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, no. 3, pp. 1301-1350, July-Sep. 2009.

[4] A. R. Dixon, *et al.*, "Continuous operation of high bit rate quantum key distribution," *Appl. Phys. Lett.*. vol. 96, pp. 161102, 2010.

[5] N. A. Peters, *et al.*, "Quantum communications in reconfigurable optical networks: DWDM QKD through a ROADM," in *Proc. Opt. Fiber Commun. Conf.*, 2010, paper OTuk1.

[6] A. Chraplyvy, "The coming capacity crunch," plenary talk, in *Proc. Eur. Conf. Optical Commun.*, Vienna, Austria, 2009, pp.1.

[7] T. Morioka, "New generation optical infrastructure technologies: EXAT Initiative towards 2020 and beyond," in *Proc. OptoElectronics Communication Conf. (OECC)*, Hong Kong, 2009, pp. 1-2.

[8] P. J. Winzer, "Modulation and multiplexing in optical communication systems," *IEEE LEOS Newsletter*, pp. 4-10, Feb. 2009.

[9] 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, 2011.

[10] S. Chandrasekhar, *et al.*,"WDM/SDM transmission of 10 x 128-Gb/s PDM-QPSK over 2688-km 7-core fiber with a per-fiber net aggregate spectral-efficiency distance product of 40,320 km.b/s/Hz," *Proc. Eur. Conf. Optical Commun.*, 2011, Paper Th.13.C.4.

[11] H. Takahashi, *et al.*, "First demonstration of MC-EDFA-repeatered SDM transmission of 40 x 128-Gbit/s PDM-QPSK signals per core over 6,160-km 7-core MCF," in *Proc. Eur. Conf. Optical Commun.*, 2012, Paper Th3.C.3.

[12] S. Randel, *et al.*, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6MIMO Processing," *IEEE/OSA J. Lightw. Technol.*, vol. 30, no. 4, pp. 521-531, Feb. 2012

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2015.2405897, IEEE Transactions on Information Forensics and Security

[13] K. Guan, P. J. Winzer, and E. Soljanin, "Information-theoretic security in space-division multiplexed fiber optic networks," in *Proc. Eur. Conf. Optical Commun.*, 2012, paper Tu.3.C.4.

[14] K. Guan, *et al.*, "Physical layer security in space-division multiplexed fiber optic communications, " in *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2012, pp.654-658.

[15] K. Guan, P. J. Winzer, E. Soljanin, and A. M. Tulino, "On the secrecy capacity of the space-division multiplexed fiber optical communication system, " in *IEEE Conference on Communication and Network Security (CNS)*, Washington D.C., 2013, pp.207-214.

[16] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp.1355-1387, Oct. 1975 .

[17] S. Cheong and M.E. Hellman,"The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp.451-456, July 1978.

[18] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *International Symposium Information Theory (ISIT)*, Seattle, WA, 2006, pp.356-360.

[19] A. O. Hero, "Secure space-time communication," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3235-3249, Dec. 2003.

[20] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961-4972, Aug. 2011.

[21] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas-I: the MISOME wiretap channel, " *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088-3104, July 2010.

[22] A. Khisti and G. W. Wornell, "Secure Transmission with multiple antennas-II: the MIMOME wiretap channel," *IEEE. Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515-5532, Nov. 2010.

[23] R. Bustin, R. Liu, H. V. Poor, and S. Shamai, "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP J. Wireless Commun. Network.*, pp. 2602-2606, 2009.

[24] A. Pruyear and V. W. S. Chan, "Using spatial diversity to improve the confidentiality of atmospheric free space optical communication," in *IEEE Globecomm* , Houston, TX, 2011, pp. 1-6.

[25] N. Yang, *et al.*, "Transmit antenna selection for security enhancement in MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 144-154, Jan 2013.

[26] S. Lin and C. Lin "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," arXiv:1309.1516v1 [cs.IT], Sep. 2013.

[27] P. J. Winzer, "Spatial multiplexing: the next frontier in network capacity scaling," in *Proc. Eur. Conf. Optical Communication*, 2013, Paper We.1.D.1.

[28] P. J. Winzer, "Making spatial multiplexing a reality, *Nature Photonics*, vol. 8, pp. 345-348, 2014.

[29] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading evnvironment when using multi-element antennas," Bell Labs Tech. J. vol.1, no. 2, pp. 41-59, 1996.

[30] R. Ryf, *et al.*, "Analysis of mode-dependent gain in Raman amplified few-mode fiber," in *Proc. Opt. Fiber Commun. Conf.*, 2012, Paper PDP5C.2.

[31] S. Randel, *et al.*, "Adaptive MIMO signal processing for mode-division multiplexing," in *Proc. Opt. Fiber Commun. Conf.*, 2012, Paper PDP5C.5.

[32] S. Ö Arik, J. M. Kahn, and K.-P. Ho, "MIMO signal processing for mode-Division Multiplexing," *IEEE Signal Process. Mag.*, vol. 31, pp. 25-34, Mar. 2014.

[33] P. Krummrich, K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," in *Proc. Opt. Fiber Commun. Conf.*, 2004, Paper FI3.

[34] P. Krummrich, E.-D. Schmidt, W. Weiershausen, A. Mattheus, "Field trial results on statistics of fast polarization changes in long haul WDM transmission systems," *Proc. Opt. Fiber Commun. Conf.*, 2005, Paper OThT6.

[35] K. Guan, P. J. Winzer, and M. Shtaif, "BER performance of MDL-impaired MIMO-SDM systems with finite constellation inputs," *IEEE Photon. Technol. Lett*, vol.26, no. 12, pp. 1223-1226, June 2014.

[36] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: statistics and effect on mode-division multiplexing," *Opt. Exp.*, vol.19, no. 17, pp. 16612-16635, August 2011.

[37] C. Antonelli, A. Mecozzi, and M. Shtaif, and P. J. Winzer, "Stokes-space analysis of modal dispersion in fibers with multiple mode transmission," *Opt. Exp.*, vol. 20, no. 11, pp. 11718-11733, May 2012.

[38] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends In Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[39] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, Belmont, MA, 2003.

[40] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE. Trans. Inf. Theory*, Vol. 51, No. 4, pp. 1261-1283, Apr. 2005.

[41] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge University Press, 1991.

[42] K. Guan, P. J. Winzer, A. M. Tulino, and E. Soljanin, "An error probability approach for quantifying physical layer security of MIMO-SDM systems," *Pro. Eur. Conf. Optical Commun.*, 2014, Paper P.5.22.

1556-6013 (c) 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.