MIMO capacities and outage probabilities in spatially multiplexed optical transport systems

Peter J. Winzer* and Gerard J. Foschini

Bell Labs, Alcatel-Lucent, 791 Holmdel-Keyport Rd., Holmdel, New Jersey 07733, USA

**peter.winzer@alcatel-lucent.com*

Abstract: With wavelength-division multiplexing (WDM) rapidly nearing its scalability limits, space-division multiplexing (SDM) seems the only option to further scale the capacity of optical transport networks. In order for SDM systems to continue the WDM trend of reducing energy and cost per bit with system capacity, integration will be key to SDM. Since integration is likely to introduce non-negligible crosstalk between multiple parallel transmission paths, multiple-input multiple output (MIMO) signal processing techniques will have to be used. In this paper, we discuss MIMO capacities in optical SDM systems, including related outage considerations which are an important part in the design of such systems. In order to achieve the low-outage standards required for optical transport networks, SDM transponders should be capable of individually addressing, and preferably MIMO processing all modes supported by the optical SDM waveguide. We then discuss the effect of distributed optical noise in MIMO SDM systems and focus on the impact of mode-dependent loss (MDL) on system capacity and system outage. Through extensive numerical simulations, we extract scaling rules for mode-average and mode-dependent loss and show that MIMO SDM systems composed of up to 128 segments and supporting up to 128 modes can tolerate up to 1 dB of per-segment MDL at 90% of the system's full capacity at an outage probability of 10^{-4}.

OCIS codes: (060.4510) Optical communications; (060.4230) Multiplexing.

References and links

1. R. W. Tkach, "Scaling optical communications for the next decade and beyond," Bell Labs Tech. J. **14**(4), 3–10 (2010).
2. J. Gray and P. Shenoy, "Rules of thumb in data engineering," Microsoft Research Technical Report MS-TR-99-100 (2000).
3. J. L. Hennessy and D. A. Patterson, *Computer Architectures: A Quantitative Approach* (Morgan Kaufmann, 2003).
4. `http://top500.org/lists/2010/06/performance_development`
5. F. B. Shepherd and P. J. Winzer, "Selective randomized load balancing and mesh networks with changing demands," J. Opt. Netw. **5**(5), 320–339 (2006).
6. P. J. Winzer, "Beyond 100G ethernet," IEEE Commun. Mag. **48**(7), 26–30 (2010).
7. R.-J.Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," J. Lightwave Technol. **28**(4), 662–701 (2010).
8. A. R. Chraplyvy, "The coming capacity crunch," European Conference on Optical Communication (ECOC'09), plenary talk (2009).
9. P. J. Winzer and R.-J. Essiambre, "Advanced optical modulation formats," in *Optical Fiber Telecommunications V B*, I. Kaminow, T. Li, and A. Willner (eds.), (Academic, 2008), ch. 2, pp. 23–94.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

10. P. J. Winzer, "Modulation and multiplexing in optical communication systems," IEEE-LEOS Newsletter, Feb. 2009, http://photonicssociety.org/newsletters/feb09/modulation.pdf.
11. P. J. Winzer, "Energy-efficient optical transport capacity scaling through spatial multiplexing," IEEE Photon. Technol. Lett. **23**(13), 851–853 (2011).
12. R. J. Essiambre, "Impact of fiber parameters on nonlinear fiber capacity," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), OTuJ1 (2011).
13. T. Morioka, "New generation optical infrastructure technologies: EXAT initiative towards 2020 and beyond," Proc. Optoelectronics and Communications Conference (OECC'09), FT4 (2009).
14. Y. Kokubun and M. Koshiba, "Novel multi-core fibers for mode division multiplexing: proposal and design principle," IEICE Electron. Express **6**(8), 522–528 (2009).
15. C. R. Doerr and T. F. Taunay, "Silicon photonics core-, wavelength-, and polarization-diversity receiver," IEEE Photon. Technol. Lett. **23**(9), 597–599 (2011).
16. P. J. Winzer, A. H. Gnauck, A. Konczykowska, F. Jorge, and J.-Y. Dupuy, "Penalties from in-band crosstalk for advanced optical modulation formats," Proc. European Conference on Optical Communication (ECOC'11), Tu.5.B.7 (2011).
17. G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," Bell Labs Tech. J. **1**(2), 41–59 (1996).
18. J. Sakaguchi, Y. Awaji, N. Wada, A. Kanno, T. Kawanishi, T. Hayashi, T. Taru, T. Kobayashi, and M. Watanabe, "109-Tb/s (7x97x172-Gb/s SDM/WDM/PDM) QPSK transmission through 16.8-km homogeneous multi-core fiber," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), PDPB6 (2011).
19. B. Zhu, T. F. Taunay, M. Fishteyn, X. Liu, S. Chandrasekhar, M. F. Yan, J. M. Fini, E. M. Monberg, F. V. Dimarcello, K. Abedin, P. W. Wisk, D. W. Peckham, and P. Dziedzic, "Space-, wavelength-, polarization-division multiplexed transmission of 56-Tb/s over a 76.8-km seven-core fiber," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), PDPB7 (2011).
20. R. Ryf, R.-J. Essiambre, S. Randel, A. H. Gnauck, P. J. Winzer, T. Hayashi, T. Taru, and T. Sasaki, "MIMO-based crosstalk suppression in spatially multiplexed 56-Gb/s PDM-QPSK signals in strongly-coupled 3-core fiber," accepted for publication in IEEE Photon. Technol. Lett. (2011).
21. R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle, "Space-division multiplexing over 10 km of three-mode fiber using coherent 6x6 MIMO processing," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), PDPB10 (2011).
22. M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Astruc, L. Provost, F. Cerou, and G. Charlet, "Transmission at 2x100Gb/s, over two modes of 40-km-long prototype few-mode fiber, using LCOS-based mode multiplexer and demultiplexer," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), PDPB9 (2011).
23. A. Li, A. Al Amin, X. Chen, and W. Shieh, "Reception of mode and polarization multiplexed 107-Gb/s CO-OFDM signal over a two-mode fiber," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), PDPB8 (2011).
24. S. Berdague and P. Facq, "Mode division multiplexing in optical fibers," Appl. Opt. **21**(11), 1950–1955 (1982).
25. H. R. Stuart, "Dispersive multiplexing in multimode optical fiber," Science **289**(5477), 281–283 (2000).
26. S. Murshid, B. Grossman, and P. Narakorn, "Spatial domain multiplexing: a new dimension in fiber optic multiplexing," Opt. Laser Technol. **40**(8), 1030–1036 (2008).
27. A. Tarighat, R. C. J. Hsu, A. R. Shah, A. H. Sayed, and B. Jalali,"Fundamentals and challenges of optical multiple-input-multiple-output multimode fiber links," IEEE Commun. Mag., 57–63 (2007).
28. S. Schoellmann, N. Schrammar, and W. Rosenkranz, "Experimental realisation of 3x3 MIMO system with mode group diversity multiplexing limited by modal noise," Proc. Optical Fiber Communications Conference (OFC/NFOEC'08), JWA68 (2008).
29. M. Nazarathy and A. Agmon, "Coherent transmission direct detection MIMO over short-range optical interconnects and passive optical networks," J. Lightwave Technol. **26**, 2037–2045 (2008).
30. B. Franz, D. Suikat, R. Dischler, F. Buchali, and H. Buelow, "High speed OFDM data transmission over 5 km GI-multimode fiber using spatial multiplexing with 2×4 MIMO processing," Proc. European Conference on Optical Communication (ECOC'10), Tu.3.C.4 (2010).
31. P. J. Winzer and G. J. Foschini, "Outage calculations for spatially multiplexed fiber links," Proc. Optical Fiber Communications Conference (OFC/NFOEC'11), OThO5 (2011).
32. C. Koebele, M. Salsi, G. Charlet, and S. Bigo, "Nonlinear effects in long-haul transmission over bimodal optical fibre," Proc. European Conference on Optical Communication (ECOC'10), Mo.2.C.6 (2010).
33. H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," IEEE Trans. Commun. **50**(2), 225–234 (2002).
34. A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—a key to Gigabit wireless," Proc. IEEE **92**(2), 198 –218 (2004).
35. H. Kogelnik, L. E. Nelson, and R. M. Jopson, "Polarization mode dispersion," in *Optical Fiber Telecommunications IV B*, I. P. Kaminow and T. Li (eds.), San Diego: Academic, ch. 15, 725–861 (2002).
36. M. Brodsky, N. J. Frigo, and M. Tur, "Polarization mode dispersion," in *Optical Fiber Telecommunications V A*,

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

I. P. Kaminow, T. Li, and A. E. Willner (eds.), (Academic, 2008), ch. 17, pp. 605–670.
37. C. Xie, "Polarization-mode-dispersion impairments in 112-Gb/s PDM-QPSK coherent systems," Proc. European Conference on Optical Communication (ECOC'10), Th.10.E.6 (2010).
38. F. Mezzadri, "How to generate random matrices from the classical compact groups," Notices of the AMS **54**, 592–604 (2007).
39. B. Wedding and C. N. Haslach, "Enhanced PMD mitigation by polarization scrambling and forward error correction," Proc. Optical Fiber Communication Conference (OFC'01), WAA1 (2001).
40. X. Liu, C. R. Giles, X. Wei, A. J. van Wijngaarden, Y.-H. Kao, C. Xie, L. Moller, and I. Kang, "Demonstration of broad-band PMD mitigation in the presence of PDL through distributed fast polarization scrambling and forward-error correction," IEEE Photon. Technol. Lett. **17**(5), 1109–1111 (2005).
41. M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," Opt. Express **16**(18), 13918–13932 (2008).
42. A. Nafta, E. Meron, and M. Shtaif, "Capacity limitations in fiber-optic communications systems as a result of polarization dependent loss," Opt. Lett. **34**(23), 3613–3615 (2009).
43. E. Meron, A. Andrusier, M. Feder, and M. Shtaif, "Use of space-time coding in coherent polarization-multiplexed systems suffering from polarization dependent loss," Opt. Lett. **35**(21), 3547–3549 (2010).
44. A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," IEEE Photon. Technol. Lett. **14**(3), 313–315 (2002).
45. Y. Fukada, "Probability density function of polarization dependent loss (PDL) in optical transmission system composed of passive devices and connecting fibers," J. Lightwave Technol. **20**(6), 953–964 (2002).
46. M. Yu, C. Kan, M. Lewis, and A. Sizmann, "Statistics of polarization-dependent loss, insertion loss, and signal power in optical communication systems," IEEE Photon. Technol. Lett. **14**(12), 1695–1697 (2002).
47. A. Mecozzi and M. Shtaif, "Signal-to-noise-ratio degradation caused by polarization-dependent loss and the effect of dynamic gain equalization," J. Lightwave Technol. **22**(8), 1856–1871 (2004).
48. L. E. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, and L. Rapp, "Statistics of polarization dependent loss in an installed long-haul WDM system," Opt. Express **19**(7), 6790–6796 (2011)
49. A. Steinkamp, S. Vorbeck, and E. I. Voges, "Polarization mode dispersion and polarization dependent loss in optical fiber systems," Proc. SPIE **5596**, 243–254 (2004).
50. A. El Amari, N. Gisin, B. Perny, H. Zbinden, and C. W. Zimmer, "Statistical prediction and experimental verification of concatenations of fiber optic components with polarization dependent loss," J. Lightwave Technol. **16**(3), 332–339 (1998).
51. P. J. Winzer, A. H. Gnauck, C. R. Doerr, M. Magarini, and L.L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM," J. Lightwave Technol. **28**(4), 547–556 (2010).
52. A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications* (Academic, 1979).
53. J. B. Lasserre, "A trace inequality for matrix product," IEEE Trans. Autom. Control **40**(8), 1500–1501 (1995).

1. Introduction

Over the past 20 years, the build-out of the Internet and of related bandwidth-hungry data services has fueled a consistent exponential growth in the demand for network transport bandwidth, at a rate of close to 60% (or $10\log_{10}(1.6) = 2dB$) per year [1], as visualized by the red circles in Fig. 1(a). (The fact that exponential traffic growth rates differ somewhat among carriers and geographical regions is not critical in this context; it merely changes the discussed time lines.) Future data-centric cloud services may accelerate these growth rates, as the network is increasingly taking on the role of a distributed computer interface, whose bandwidth scaling is tied to the exponential scaling of microprocessor performance at close to 90% (or 2.7 dB) per year by Amdahl's rule of thumb [2, 3]; the green squares in Fig. 1(a) illustrate the evolution of supercomputer processing power, measured in floating point operations per second (Flops) [4]. In addition, multi-media real-time (and hence non-cacheable) applications with large bandwidth demands, such as telepresence, are expected to gain importance, which puts an additional burden on the transport network. Furthermore, independent of machine- or user-generated traffic growth, data networks are confronted with the pressing need for a significant energy reduction. An efficient way to accomplish this task is to adopt network architectures that minimize the use of power-hungry packet routers by providing ample energy efficient optical transport bandwidth (see, e.g., Ref. [5]).

The explosive demand for network bandwidth has been economically enabled by wavelength-division multiplexing (WDM) technologies, researched and developed in the early

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

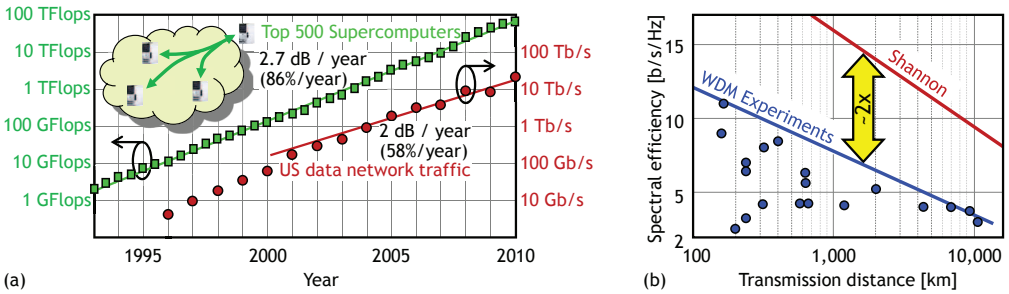

Fig. 1. (a) Traffic growth in US data networks [1] and evolution of supercomputer processing power [4], dictating machine-to-machine traffic growth. (b) Experimentally achieved WDM spectral efficiencies (blue) have approached the fundamental Shannon limit (red) to within a factor of 2 (Fig. after [11]).

1990s. Initially, WDM allowed optical network capacities to grow at a rate of ∼80% per year [1]. While this high growth rate initially exceeded the demand for network bandwidth and contributed to the telecom bubble around the year 2000 [1], both research and development have experienced a dramatic slow-down in WDM capacity growth to about 20% per year since 2002, despite the additional use of polarization-division multiplexing (PDM) [6]. Recent studies on the fundamental, nonlinear Shannon capacity limits of optical transport systems [7] explain this observation: Current experimental results have approached their fundamental limits to within a factor of ∼2, as visualized in Fig. 1(b) by the gap between the nonlinear Shannon limit of Ref. [7] (red) and recent experimental record results reported by research labs around the world (blue). Hence, the stagnation of achievable per-fiber WDM capacity growth is inevitable, and an optical transport "capacity crunch" [8] is imminent. (Note that at a traffic growth rate of 60% per year, a factor of 2 in capacity is reached within 18 months.) Critical for the business of carriers (as well as for the access fees charged to business, residential, and government customers), this implies that the per-bit cost and energy consumption of WDM data transport will not continue to drop as it has for the past 20 years to provide our society with ubiquitous data services.

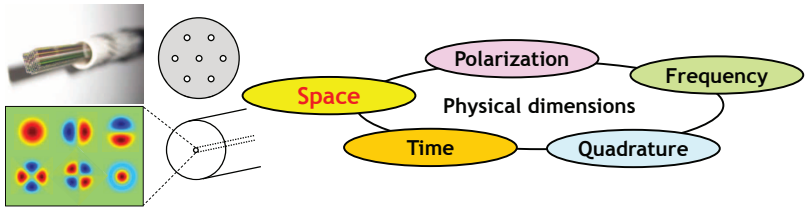

Fig. 2. Spatial multiplexing exploits the only known physical dimension that has not yet been used in optical transport systems. Implementations include fiber bundles, multi-core, and multi-mode fiber (Fig. after [10]).

Since WDM has almost reached its scalability limits, it is necessary to start intense research efforts on multiplexing strategies beyond WDM. Figure 2 shows all known physical multiplexing dimensions in an optical communication system [9, 10]. All but one of these dimensions are already being exploited in the latest generation of commercial optical communication systems: Time-division multiplexing (TDM) is used to aggregate traffic into high-speed bit and symbol streams; the two quadrature components (real and imaginary parts or amplitude and phase) of the complex optical field are being used to construct higher-order quadrature ampli-

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

tude modulation (QAM) formats; the frequency (or wavelength) dimension is used for WDM; more recently, PDM is employed to gain another factor of 2 in system capacity. *Space* is the only known physical dimension left to exploit in optical transport, which makes *spatial multiplexing* (or *space-division multiplexing*, SDM) the only known scheme that could allow for a substantial increase of optical transport capacities. Recent studies [11, 12] have shown that alternative single-lane capacity scaling approaches using high spectral efficiencies over low-loss and low-nonlinearity fibers are economically much less attractive.

An SDM system uses M parallel transmission paths per wavelength, and in its ideal version multiplies the available capacity of a given link by a factor of M. As indicated in Fig. 2, these parallel optical paths could be multiple single-mode fiber strands within a fiber cable, multiple cores within a multi-core fiber, or multiple modes within a few-mode or multi-mode waveguide. Importantly, like WDM (and unlike PDM), SDM is highly scalable since M can potentially be chosen very large. This puts high hopes on SDM as a method to satisfy the data traffic growth for the next decade and beyond [8, 13, 14].

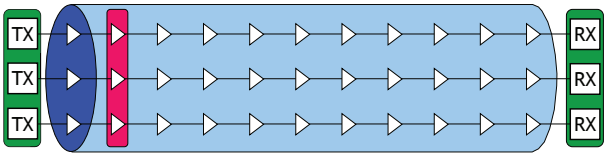

Fig. 3. Spatial multiplexing uses M parallel, integrated transmission paths to increase system capacity by a factor of M. Integration of system components is key for sustainable scalability.

Since SDM on the basis of M parallel independent optical line systems multiplies capacity but also cost and energy consumption by M, *integration* of transponders [15], optical amplifiers, and waveguides is essential to reduce the cost and energy per transmitted information bit, as shown in Fig. 3. An important impairment arising from integration is *crosstalk* between multiple SDM paths. If crosstalk rises to levels where it induces unacceptable transmission penalties [16], *multiple-input-multiple-output* (MIMO) techniques, originally developed for wireless systems [17], are used. Recently, several experimental demonstrations of SDM transport have been reported using both low-crosstalk multi-core fiber [18, 19] without MIMO processing and high-crosstalk multi-core [20] and multi-mode fiber [21–23] with MIMO processing.

Earlier work on mode-division multiplexing or mode-group multiplexing [24–30] mostly focused on average SDM capacity gains in multi-mode fiber, neglecting the high reliability requirements of optical transport systems. In this paper, which extends the results of [31], we present SDM based on two key premises that can potentially allow for low-outage optical transport capacity scaling:

- Transponders are able to coherently address a complete orthonormal set of modes whose propagation is supported by the transmission waveguide.
- MIMO processing is used to deal with modal crosstalk.

After describing the SDM system model in Sec. 2, we quantify in Sec. 3 the importance of being able to address all waveguide modes at the transponders for reliable system performance and propose methods for mitigating a potential mismatch of transponder and fiber mode count. In Sec. 4 we discuss distributed optical noise in optical MIMO systems. In Sec. 5 we focus on the impact of mode-dependent loss (MDL) on the outage behavior of SDM systems. Throughout this paper, we explore outage considerations as an integral part in the design of high-availability SDM transport systems.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

2. Space-division multiplexing system model

Neglecting inter- and intra-modal fiber nonlinearities [32], an SDM system can be represented by the linear matrix channel

$$\boldsymbol{y} = \sqrt{E_0}\sqrt{L}\,\mathrm{H}\,x + \boldsymbol{n}\,, \tag{1}$$

as visualized in Fig. 4(a). (Throughout this paper, we use **bold print** to indicate stochastic processes and random variables. Scalar and vector quantities appear in *italics*, and matrices are set in regular (not italicized) font.) The SDM waveguide supports a set of M orthogonal (spatial and polarization) propagation modes that may be subject to coupling, differential gain or loss, and differential delay. (Note that we use the term 'mode' to refer to a member of a complete orthonormal set of physical modes or linear combinations thereof, whose propagation is supported by the SDM waveguide; MIMO literature often uses the term 'mode' to refer to the eigenmodes of the channel matrix.) Throughout this paper, we ignore the effect of explicit differential modal delays, i.e., we deal with the two extreme cases of (*i*) no differential delays, leading to frequency-flat fading, and (*ii*) differential delays that are much larger than the inverse signal bandwidth, such that the signal samples the entire channel statistics over its bandwidth, which allows for capacity averaging [33, 34]. We assume that the transmitter is able to excite $M_T \leq M$ waveguide modes, and that the receiver is able to coherently extract $M_R \leq M$ modes. The average signal energy transmitted per symbol period and *per channel mode* is E_0, hence the total transmit energy across all modes is $M_T E_0$ per symbol and $\langle ||x||^2 \rangle = M_T$. Note that this definition differs from the constant total power constraint often used in wireless MIMO, motivated by the fact that due to fiber nonlinearity the optical power per mode is likely a more suitable parameter in fiber-optics than the average power across all modes [32]. Circularly symmetric complex Gaussian noise $\boldsymbol{n}$ with power spectral density N_0 per mode is added at the receiver. (The case of distributed noise loading in a system of concatenated waveguide sections is treated in Sec. 4.) The SDM waveguide is described by an $M \times M$ matrix $\widetilde{\mathrm{H}}$, which we normalize as $\widetilde{\mathrm{H}} = \sqrt{L}\,\mathrm{H}'$ by factoring out the mode-average net propagation loss (or gain) L,

$$L = \frac{1}{M}\mathrm{tr}\{\widetilde{\mathrm{H}}\widetilde{\mathrm{H}}^\dagger\} = \frac{1}{M}\sum_{i=1}^{M}\widetilde{\lambda}_i\,; \tag{2}$$

$\mathrm{tr}\{\cdot\}$ denotes the trace of a matrix, $\widetilde{\lambda}_i$ are the M eigenvalues of $\widetilde{\mathrm{H}}\widetilde{\mathrm{H}}^\dagger$, and $\widetilde{\mathrm{H}}^\dagger$ is the conjugate transpose of $\widetilde{\mathrm{H}}$. Apart from noise, the MIMO channel is then given by the $M_R \times M_T$ matrix H spanning the subspace of H' addressed by the transponders.

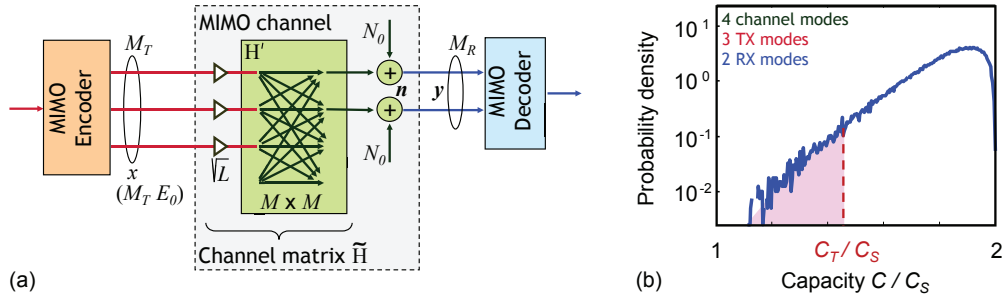

Fig. 4. (a) Basic MIMO system model for $M_T = 3$, $M_R = 2$, and $M = 4$. (b) Histogram of MIMO capacities $\boldsymbol{C}$, normalized to the single-mode capacity C_S. The shaded area represents the probability that $\boldsymbol{C}$ is smaller than C_T, leading to system outage if the system is designed to code for capacity C_T.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

Throughout this work, we assume that each individual instantiation H of the channel matrix is *known to the receiver* (e.g., through training sequences) but is *unknown to the transmitter* (e.g., due to the long roundtrip delays usually found in optical transport networks). In this case, the transmitter sends uncorrelated signals of equal power on all transmit modes and at best achieves the open-loop Bell Labs Layered Space Time (BLAST) capacity [17]

$$C = \sum_{i=1}^{r} \log_2 \left(1 + \lambda_i \frac{E_0 L}{N_0} \right) , \tag{3}$$

where λ_i are the $r \leq \min\{M_T, M_R\} \leq M$ non-zero eigenvalues of $\mathrm{HH}^\dagger$ and r is referred to as the rank of the MIMO channel. The term

$$\mathrm{SNR} = \frac{E_0 L}{N_0} \tag{4}$$

represents the mode-average signal-to-noise ratio (SNR) at the receiver when all M channel modes are excited and detected ($M_T = M_R = M$) or, equivalently, the SNR measured as the ratio of the total received signal power to the total received noise power for equal power launched into all modes. Since H is a random instantiation of an ensemble of possible channel matrices **H**, the capacity given by Eq. (3) is an instantiation of the ensemble of potential channel capacities $\boldsymbol{C}$. For frequency-flat fading (i.e., no differential modal delays), an instantiation of the channel capacity applies to the entire signal. In the presence of differential modal delays, the number of uncorrelated frequency bins within the signal bandwidth is approximately given by the ratio of the signal bandwidth to the channel's coherence bandwidth (which is inversely proportional to the channel's delay spread), and each independent signal frequency component (e.g., each independent subcarrier of an orthogonal frequency division multiplexed (OFDM) signal) experiences a different instantiation of $\boldsymbol{C}$; the applicable instantiation of the channel capacity is then found by averaging the capacity instantiations of the independent frequency subchannels [33, 34]. For very large differential modal delays, the signal ultimately sees the average channel capacity $\langle \boldsymbol{C} \rangle$.

3. Trade-off between system outage and individually addressable modes

3.1. Outage probabilities

Assuming that channel impairments randomly couple or attenuate modes, but not to an extent where modes are completely eliminated upon propagation, the rank of the MIMO channel is given by $r = \min\{M_T, M_R\} \leq M$. In the absence of mode-dependent loss, the maximum potential capacity of the under-addressed SDM waveguide is r times the single-mode capacity C_S,

$$\boldsymbol{C} \leq r\, C_S = r\, \log_2 (1 + \mathrm{SNR}) , \tag{5}$$

as shown in the Appendix. To exploit the full MIMO capacity of an M-mode waveguide, one should hence use M transmitters and M receivers that are able to individually excite and extract any suitable complete orthonormal basis of M channel modes, such as the set of guided linearly polarized (LP) modes in a fiber. This requirement is an extension of the familiar 2×2 PDM system, where both polarizations have to be individually addressed by transmitter and receiver in order to double the capacity of a single-polarization system. Using less than M transmitters underutilizes the channel. In addition, using less than M receivers can result in situations where the channel randomly couples transmit modes to propagation modes that are not extracted by the receiver. This, in turn, randomly reduces the capacity of the channel. With reference to Fig. 4(a), some random channel instantiation may couple 2 out of 3 transmit signals

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

to the 2 undetected receive modes, leaving only a single transmit mode for information transmission. This worst-case situation may occur very rarely, though, and designing for this worst case may result in unattractively low SDM capacities. Hence, one takes a different approach: Without knowledge of the channel instantiation, the transmitter is configured to code for a certain (fixed) capacity C_T, which is chosen larger than the worst-case capacity of the ensemble of random channel matrices. In fact, C_T is chosen such that the *probability* that a random channel instantiation only supports a capacity $C < C_T$ is acceptably small. These rare cases will then inherently lead to *system outage*. Hence, outage considerations are an integral part of SDM system design, similar to what is commonly done in systems suffering from polarization-mode dispersion (PMD) impairments [35–37].

In order to quantify system outage probabilities, we generated 100,000 random realizations of channel matrices $\mathbf{H}$ following Ref. [38], assuming a waveguide with uniformly distributed unitary coupling among all propagation modes (i.e., a waveguide without mode-dependent loss, resulting in $L = 1$ for all matrices). For the two-dimensional case, which represents two polarizations in a single transverse mode, we gathered strong numerical evidence that the resulting ensembles of unitary matrices yield uniform coverage of the Poincare sphere, independent of the input state of polarization. While uniform coverage of the Poincare sphere is a generally accepted (though not necessarily always applicable) model for the polarization rotation in fiber segments, its extension to the M-dimensional case is a pure assumption here. The stochastic nature of realistic SDM channel matrices is likely to depend on the underlying waveguide structure and is left as a topic for future research.

Figure 4(b) shows a histogram (based on 100,000 random matrix realizations) of the channel capacity $\boldsymbol{C}$ for the scenario of Fig. 4(a), where the channel supports $M = 4$ (spatial and polarization) modes, the transmitter is capable of addressing 3 channel modes, and the receiver is capable of extracting 2 modes. (Note that the number of spatial and polarization modes supported by optical waveguides does not take on arbitrary values, as is assumed here; for example, a step-index fiber supports $\{2, 6, 10, 12, 16, 20, \ldots\}$ modes, and hexagonally packed multi-core arrangements support $\{2, 14, 38, 74, \ldots\}$ modes.) Most of the time, capacities approaching $2C_S$ are attained, but sometimes it may happen that the capacity almost drops to C_S. If the transmitter codes for capacity C_T, the shaded area represents the probability P_{out} that the channel capacity will fall below C_T to result in system outage,

$$P_{\text{out}} = \int_0^{C_T} p_{\boldsymbol{C}}(C)\, \mathrm{d}C\,, \tag{6}$$

where $p_{\boldsymbol{C}}(C)$ is the probability density function (PDF) of the random channel capacity $\boldsymbol{C}$. In contrast to some wireless MIMO systems, fiber-optic transport networks require exceptionally low outage probabilities, typically below 10^{-5}, corresponding to 99.999% ('five nines') in system availability.

As discussed above, if differential group delays between modes are large, such that each signal frequency component experiences a different channel instantiation, the signal ultimately sees the average channel capacity $\langle \boldsymbol{C} \rangle$. In the idealized limiting case of a highly frequency selective channel, this average capacity is then guaranteed, i.e., if the transmitter codes for $C_T = \langle \boldsymbol{C} \rangle$, the receiver will always be able to extract C_T without experiencing outage.

3.2. Under-addressed SDM waveguides using fixed or dynamically switched mode sets

Figure 5(a) shows outage probability curves for $M = 4$ and fixed combinations of $M_T \times M_R$ addressable transmit $\times$ receive modes for an SNR of 20 dB. In contrast to wireless MIMO systems, the assumption of constant per-mode transmit power lets the SDM capacities of this

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

channel be symmetric in M_T and M_R. Deterministic, outage-free capacities (i.e., step-like outage curves) are observed when either M transmitters or M receivers are being used, with only the $M \times M$ system fully exhausting the available channel capacity for an M-fold capacity increase over a single-mode system. Capacity loss and system outage have to be accepted when choosing M_T or M_R less then M. Figure 5(b) illustrates the impact of the SNR on the outage probabilities. The symmetric transponder cases ($M_T = M_R = M_{TR}$) of Fig. 5(a) for a 20-dB SNR are redrawn (solid) together with the curves for an SNR of 10, 30, and 40 dB (dashed). At a fixed outage probability, higher SNRs result in higher capacity improvements over the equivalent single-mode system. Also shown in Fig. 5(b) as dotted lines are the average capacities $\langle \boldsymbol{C} \rangle$, as experienced on a highly frequency-selective channel, i.e., in the presence of differential modal delays that are much larger than the inverse signal bandwidth. At low outage probabilities, these capacities are generally superior to the capacities achieved over a frequency-flat channel.

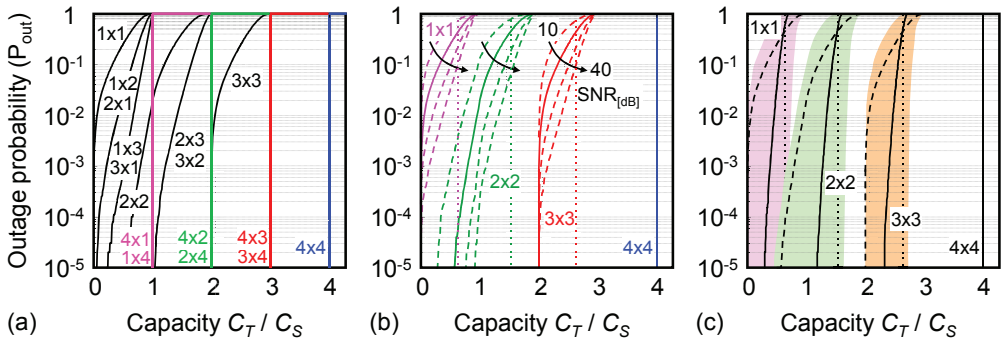

Fig. 5. (a) Achievable SDM capacities C_T normalized to the single-mode capacity C_S at a given outage probability for different combinations of transmit $\times$ receive modes on a waveguide supporting $M = 4$ modes. (b) Impact of mode-average SNR on the outage performance of an under-addressed 4-mode waveguide (dashed). Also shown are the step-like outage curves for a highly frequency-selective channel (dotted). (c) Outage performance of systems using more than one fixed (dashed) combination of mode sets (shaded areas); solid curves: dynamic switching among all possible mode set combinations; dotted lines: highly frequency-selective channel.

If the number of modes addressable by the SDM transponder is limited to less than the number of channel modes by the complexity of the required MIMO processing but *not* by the complexity of the required mode excitation optics, outage probabilities can be significantly improved over the above discussed case where a static, fixed combination of transmit and receive modes is being used. Since there are $\binom{M}{M_T}$ combinations of transmit modes and $\binom{M}{M_R}$ combinations of receive modes that can be used to transport SDM information, there are a total of $\binom{M}{M_T} \times \binom{M}{M_R}$ potential SDM capacities C for each of the random channel instantiations. The dashed and dotted lines in Fig. 5(c) represent the case of fixed mode selection in the frequency-flat and highly frequency-selective scenario of Fig. 5(a), respectively. The shaded areas give the ranges of simulated outage curves for the case where the *receiver* of a transponder supporting $M_T = M_R = M_{TR}$ modes always selects the best or always selects the worst set of receive modes on a 4-mode channel at 20-dB SNR, while the transmitter uses a fixed mode combination. In order to find and dynamically adapt to the best set of receive modes, a system could employ M_S *scouting modes*, whereby the transmitter codes information across M_{TR} modes, but at a code rate that allows the receiver to extract the full information from detecting only a set of $M_{TR} - M_S$ modes; the maximum possible capacity of the system is thus reduced to

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

$(M-M_S)C_S$. The M_S scouting modes are dynamically switched among $M-(M_{TR}-M_S)$ modes to sound their quality. If a scouting mode is found by the receiver to exhibit superior quality than any of the $M_{TR}-M_S$ modes currently used for information extraction, this scouting mode becomes a new information-bearing mode while the worst mode previously used for information extraction becomes a new scouting mode. In another, less sophisticated implementation, the transmitter and/or the receiver randomly switch among a substantial number of possible sets of waveguide modes at a rate that exceeds the burst error correction capability of the underlying error-correcting code, similar to what has been proposed to counteract penalties from PMD in single-mode fiber systems [39,40]. This results for each channel matrix in the average capacity of the chosen mode combinations. The solid lines in Fig. 5(c) give the outage probabilities for the case where the receiver switches among *all* possible mode combinations, dwelling for an equal time duration over each of the selections.

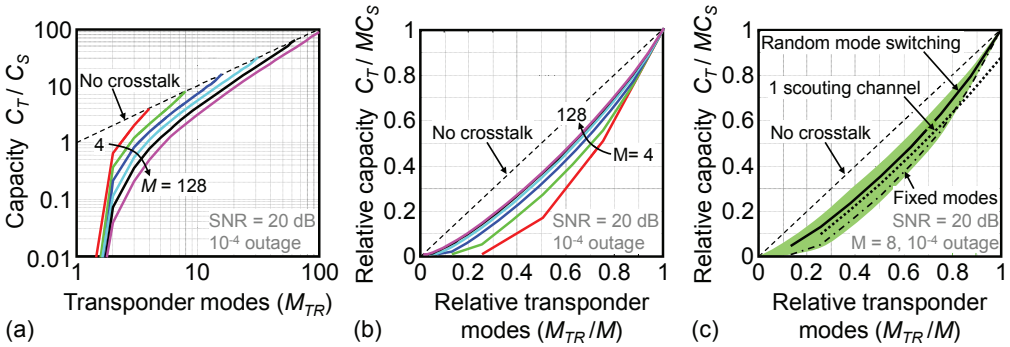

Fig. 6. (a) Achievable SDM capacity C_T relative to the single-mode capacity C_S at 10^{-4} outage and 20 dB SNR, as a function of the number of modes M_{TR} processed by a fixed mode set transponder. The number of modes M supported by the waveguide parameterizes the curves. (b) Normalized version of (a). (c) Capacity gain of dynamically mode set switching transponders on a waveguide supporting $M=8$ modes.

Figures 6(a) and 6(b) extend our considerations for the flat-fading, fixed mode set case to channels supporting M = 4, 8, 16, 32, 64, and 128 modes, assuming symmetric transponder pairs ($M_T=M_R=M_{TR}$) at 20 dB SNR. In (a) we show the achievable SDM capacity C_T in multiples of the single-mode capacity C_S at an outage probability of 10^{-4}. The dashed line $C_T=M_{TR}\,C_S$ represents crosstalk-free performance as an upper bound to any MIMO system, only to be reached in the unitary MIMO case if the transponders are able to address *all* channel modes. It can be seen that for a given transponder complexity (i.e., for fixed M_{TR}) it is generally preferable in terms of aggregate SDM capacity to use a waveguide that supports $M=M_{TR}$ modes. Deploying 'future-proof' waveguides whose mode count exceeds current transponder capabilities results in a significant start-up capacity penalty. For example, operating a waveguide supporting $M=32$ modes with transponders that only address $M_{TR}=8$ modes results in a capacity close to $4C_S$. This capacity can also be achieved using only *half* the amount of transponders (and consequently much less MIMO signal processing) on a waveguide supporting just 4 modes. Hence, one should always strive to match the number of coupled waveguide modes to the number of modes that the transponders can address and are able to MIMO-process. Figure 6(b) shows the above curves on a normalized scale, plotting the SDM capacity relative to its maximum value of $M\,C_S$ as a function of the fraction of waveguide modes addressed by the transponders. The crosstalk-free system corresponds to the straight dashed line. Interestingly, the MIMO curves converge to a common asymptote for large M.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

Figure 6(c) represents a flat-fading system with $M = 8$ waveguide modes and a receiver designed to detect all possible combinations of M_{TR} out of M modes at the receiver. The static case of using a fixed set of receive modes is redrawn from Fig. 6(b) as a thin dash-dotted line. The boundaries of the shaded area represent the two limiting cases where the system always uses the 'best' and the 'worst' possible mode set for each channel matrix instantiation. Randomly switching among all possible mode permutations for each channel matrix at a speed that allows for capacity averaging results in the solid curve. Using $M_S = 1$ out of M_{TR} addressed channels for scouting purposes results in the dotted line, with a maximum achievable capacity of $7/8\, C_S$ at $M_{TR}/M = 1$, since one channel is always used for scouting.

4. Distributed noise loading

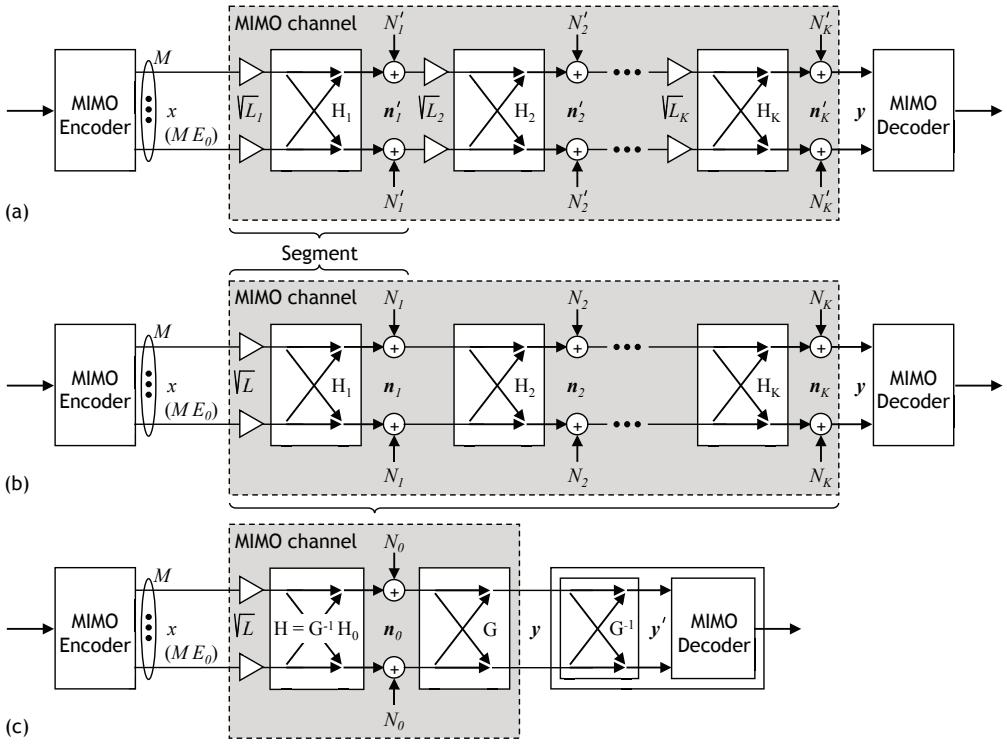

Fig. 7. System models for a MIMO channel with distributed noise.

With reference to Fig. 7(a), we now consider an $M \times M$ SDM system operating on a matched M-mode waveguide composed of K concatenated segments with segment matrices H_i. Since we only consider the $M \times M$ case here, we have $\mathrm{H}'_i = \mathrm{H}_i$. Noise power N'_i is added at the end of each segment. As before, we factor out the net mode-average gain or loss L_i according to Eq. (2). In the linear regime covered here, it is immaterial whether the mode-independent gain or loss within a segment physically occurs before or after the mode coupling described by the segment matrix. If a segment is identified with a net transparent amplification span, we have $L_i = 1$, and the physical segment loss is reflected in the amount of noise N'_i associated with optical amplification to reach per-span transparency. Note that in the presence of MDL, some modes may experience a net gain while others may experience a net loss. Consider, for example,

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

the simple 2×2 case with the two segment matrices

$$\mathrm{H}_1 = \begin{pmatrix} \sqrt{4/3} & 0 \\ 0 & \sqrt{2/3} \end{pmatrix} \quad \text{and} \quad \mathrm{H}_2 = \begin{pmatrix} \sqrt{2/3} & 0 \\ 0 & \sqrt{4/3} \end{pmatrix} . \tag{7}$$

The channel eigenvalues of both segments by themselves are $\frac{2}{3}$ and $\frac{4}{3}$, resulting in $L_{1,2} = \frac{1}{2}(\frac{2}{3}+\frac{4}{3}) = 1$, and consequently in $L = 1$. (Note that although $L_{1,2} = 1$, these channels are *not* entirely passive, as they amplify one of their modes but attenuate the other.) The concatenated channel has the twofold degenerate eigenvalue 8/9, resulting in $L_\Sigma = 8/9$, a mode-average *loss* of 0.5 dB. Concatenating the two matrices with a 90-degree rotation matrix in between results in the channel eigenvalues 4/9 and 16/9, and hence in $L_\Sigma = 10/9$, a net mode-average *gain* of 0.5 dB.

We next define an equivalent system according to Fig. 7(b) by replacing the individual segment losses by a lumped loss of $L = \prod_{i=1}^{K} L_i$ at the input to the channel and adjusting the noise powers at each segment such as to match the actual noise found in the original system, i.e., $N_i = N_i' \prod_{k=i+1}^{K} L_k$. Note that for an ensemble of random, non-unitary segment matrices, the overall net mode-average loss (or gain) L_Σ for each instantiation of the concatenated system will generally differ from our lumped loss variable L, with some channel instantiations showing more and some showing less net mode-average loss than L, as discussed above. However, as we will show below, for all cases of practical interest studied here, the *ensemble average* of the mode-average loss of the concatenated system is always closely approximated by L. We will use the equivalent system of Fig. 7(b) in all further considerations.

4.1. Unitary segments

If all noise sources are statistically independent, we find for the noise correlation matrix at the receiver:

$$\mathrm{R}_{\boldsymbol{n}} = \langle \boldsymbol{n}\boldsymbol{n}^\dagger \rangle = \left(N_1 \mathrm{H}_K \cdots \mathrm{H}_3 \mathrm{H}_2 \mathrm{H}_2^\dagger \mathrm{H}_3^\dagger \cdots \mathrm{H}_K^\dagger + N_2 \mathrm{H}_K \cdots \mathrm{H}_3 \mathrm{H}_3^\dagger \cdots \mathrm{H}_K^\dagger + \cdots + N_{K-1} \mathrm{H}_K \mathrm{H}_K^\dagger + N_K \mathrm{I}_M \right) , \tag{8}$$

where I_M is the $M\times M$ identity matrix. If only modal crosstalk occurs within each fiber segment (i.e., if the segments H_i are *unitary*), the correlation equals that of spatio-temporally white noise,

$$\mathrm{R}_{\boldsymbol{n}} = \sum_{i=1}^{K} N_i \mathrm{I}_M = N_0 \mathrm{I}_M . \tag{9}$$

Hence, a concatenation of unitary segments with distributed noise has the same capacity as a unitary channel that is noise-loaded at the receiver by the same amount of noise, $N_0 = \sum_{i=1}^{K} N_i$.

4.2. Non-unitary segments

If the segment matrices are not unitary, the channel can be written as

$$\boldsymbol{y} = \sqrt{E_0}\sqrt{L}\, \mathrm{H}_0\, x + \mathrm{G}\, \boldsymbol{n}_0 , \tag{10}$$

where $\mathrm{H}_0 = \mathrm{H}_K\, \mathrm{H}_{K-1}\, \ldots \mathrm{H}_2\, \mathrm{H}_1$ is the overall channel matrix for the signal, and the white noise $\boldsymbol{n}_0$ of power density N_0 is colored by a matrix G such that $\mathrm{R}_{\boldsymbol{n}} = \mathrm{G}\,\mathrm{G}^\dagger$, as shown in Fig. 7(c). Placing the deterministic matrix G^{-1} as a whitening filter inside the receiver leaves the MIMO capacity unchanged but produces the equivalent white-noise MIMO channel

$$\boldsymbol{y}' = \sqrt{E_0}\sqrt{L}\, \mathrm{G}^{-1}\mathrm{H}_0\, x + \boldsymbol{n}_0 , \tag{11}$$

as shown in Fig. 7(c); its open-loop BLAST capacity is given by Eq. (3) with $\mathrm{H} = \mathrm{G}^{-1}\mathrm{H}_0$. The matrix G can be found from the measured noise correlation matrix $\mathrm{R}_{\boldsymbol{n}}$ by decomposing

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

$\mathrm{R}_{\boldsymbol{n}}$ into the form $\mathrm{U}\,\Theta\,\mathrm{U}^{\dagger}$, where U is a unitary matrix and Θ is a diagonal matrix containing the eigenvalues θ_i of $\mathrm{R}_{\boldsymbol{n}}$. The matrix G is then given as $\mathrm{U}\,\Theta^{\frac{1}{2}}\,\mathrm{U}^{\dagger}$, where $\Theta^{\frac{1}{2}}$ is diagonal with elements $\sqrt{\theta_i}$, and G^{-1} is given as $\mathrm{U}\,\Theta^{-\frac{1}{2}}\,\mathrm{U}^{\dagger}$, where $\Theta^{-\frac{1}{2}}$ is diagonal with elements $1/\sqrt{\theta_i}$.

5. Impact of mode-dependent loss (MDL) on SDM capacity

Based on the above results, we next investigate MIMO capacity statistics for a channel with K segments and M modes, all of which can be addressed by transmitter and receiver. Each segment is composed of a random mode coupling element followed by a random MDL element. Without loss of generality, we assume the mode-average optical loss L_i for each individual segment to be unity. Hence, each segment matrix is composed of a random unitary matrix $\mathbf{U}_i$, generated as discussed in Sec. 3.1, followed by a random diagonal matrix $\mathbf{V}_i$ whose M real-valued, positive elements $\sqrt{\boldsymbol{\nu}_{kk}}$ satisfy $\sum_{k=1}^{M}\boldsymbol{\nu}_{kk}=M$ to ensure that $L_i=1$. For each MDL segment matrix $\mathbf{V}_i$, the numbers v_{kk} are drawn from a uniform distribution such that the ratio of maximum to minimum matrix element equals the pre-specified value for the per-segment MDL,

$$\mathrm{MDL}_S=\frac{\max\{v_{kk}\}}{\min\{v_{kk}\}}\,. \tag{12}$$

We assume that all K segments have equal per-segment MDLs. Both the mode-average loss $\boldsymbol{L_{\Sigma}}$ and the MDL of the concatenated system $\mathbf{MDL_{\Sigma}}$ are random variables. The former is given by the trace of $\mathbf{HH}^{\dagger}$ and the latter by the ratio of largest to smallest eigenvalue of $\mathbf{HH}^{\dagger}$. Two example segment matrices for $M=2$ and $\mathrm{MDL}_S=3$ dB are given in Eq. (7), with the two concatenation examples leading to MDL_{Σ} of 0 dB and 6 dB, respectively. For the 2-mode case, i.e., in the context of polarization-dependent loss (PDL) in single-mode fiber, analytical relations of various sorts have been derived to describe the relationship between mode-average loss and mode-dependent loss as well as their statistical properties (see, e.g., [41–50]. An analytical extension of the statistics of 2-mode parameters to M modes is beyond the scope of this paper; instead, we present in what follows the results of extensive numerical simulations, capturing ensembles of 100,000 random realizations of $K=\{2,4,8,16,32,64,128\}$ spans with $\mathrm{MDL}_S=\{0.01,0.1,0.2,0.5,1,2,5,10\}$ dB, and SDM waveguides supporting $M=\{2,4,8,16,32,64,128\}$ modes at SNR$=\{5,10,15,20,25,30,35,40\}$ dB.

5.1. Statistics of the concatenated system's mode-average loss and mode-dependent loss

For small segment MDLs ($\mathrm{MDL}_S\lesssim 1$ dB), we extract the following empirical expressions for mean and standard deviation of the (linearly) averaged concatenated mode-average loss $\boldsymbol{L_{\Sigma}}$:

$$\langle\boldsymbol{L_{\Sigma}}\rangle\approx 1\,, \tag{13}$$

$$\sigma_{\boldsymbol{L_{\Sigma}}}\approx\alpha\frac{K}{M}\,\mathrm{MDL}_S^2\,, \tag{14}$$

where α drops from approximately $6\cdot10^{-3}$ to $3\cdot10^{-3}$ as M increases from 4 to 128. Over a wide range of the studied parameters, the PDF of $\boldsymbol{L_{\Sigma}}$, expressed in decibels, is well fitted by a Gaussian distribution.

Again for $\mathrm{MDL}_S\lesssim 1$ dB, the concatenated system's average mode-dependent loss, expressed in decibels after linear averaging, is well approximated by

$$\langle\mathbf{MDL_{\Sigma}}\rangle_{[\mathrm{dB}]}\approx\sqrt{K}\,\mathrm{MDL}_{S\,[\mathrm{dB}]}\,. \tag{15}$$

The standard deviation of $\mathbf{MDL_{\Sigma}}$ obeys more complicated scaling rules with K, M, and MDL_S. As with the mode-average loss, the Gaussian distribution is a good fit to the distribution of $\mathbf{MDL}_{\boldsymbol{\Sigma}\,[\mathrm{dB}]}$.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

5.2. *Statistics of the concatenated system's MIMO capacity*

For small segment MDLs ($\mathrm{MDL}_S \lesssim 1$ dB), we extract the following empirical expressions for the normalized channel capacity $\boldsymbol{C_n} = \boldsymbol{C}/(M\,C_S)$:

$$\langle \boldsymbol{C_n} \rangle \approx \beta^{\,K\,\mathrm{MDL}^2_{S\,[\mathrm{dB}]}}\,, \tag{16}$$

$$\sigma_{\boldsymbol{C_n}} \approx \gamma \frac{\sqrt{K}}{\sqrt{M}}\,\mathrm{MDL}^2_{S\,[\mathrm{dB}]}\,, \tag{17}$$

with β increasing mildly from 0.9990 to 0.9997 as the SNR increases from 5 dB to 40 dB, and $\gamma \approx 4 \cdot 10^{-4}$. Gaussian distributions are again found to reasonably fit the distribution of $\boldsymbol{C}$ for $M \gtrsim 4$.

5.3. *SDM outage capacities of the concatenated transmission system*

Figure 8 shows the achievable MIMO capacity for the frequency-flat channel at a given outage probability, both for noise loading at the receiver (red, solid) and for distributed noise loading at each segment (blue, dashed) for $K = 64$ segments and $M = 16$ modes, using the numerically simulated (unfitted) data. The SNR used in the four subplots is 10, 20, 30, and 40 dB, and within each subplot the per-segment MDL is varied from 0.5 to 5 dB as indicated. As expected, the achievable MIMO capacity drops with increasing MDL. Distributed noise loading gives slightly better capacities than noise loading at the receiver, which is due to the reduced noise enhancement by receive-side (spatial) equalization. This is in analogy to the better performance of single-mode receivers when noise loading is performed prior to narrow bandpass filtering [51]: if an optical filter is placed after all noise has been added, the filter acts on signal *and* noise, which lets the receive-side equalizer restore approximately white noise conditions during signal equalization. On the other hand, an optical filter placed before any noise is added only affects the signal but not the noise, which results in severe noise enhancement upon signal equalization within the receiver. The outage-free capacities $\langle \boldsymbol{C} \rangle$ describing the highly frequency-selective channel are given in Fig. 8 by the dotted lines for MDL_S of 2 and 5 dB. The differences to the frequency-flat channel outages at low outage probabilities are negligibly small at lower MDL_S.

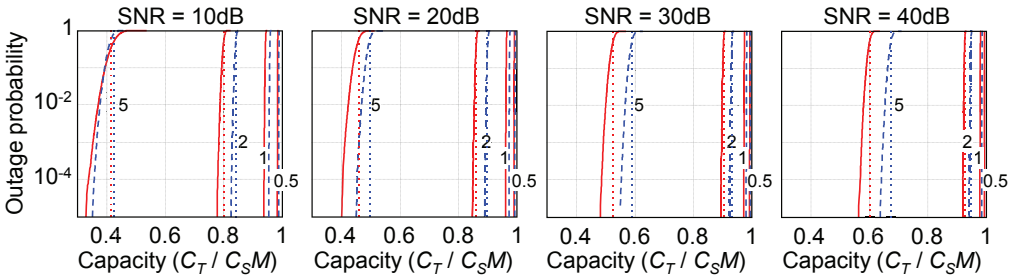

Fig. 8. Achievable SDM capacities C_T normalized to M times the single-mode capacity C_S at a given outage probability for $K = 64$ segments and $M = 16$ modes, with the per-segment MDL as a parameter. The SNR is varied from 10 to 40 dB among the four subplots. Solid red curves pertain to noise loading at the receiver, while dashed blue curves represent distributed noise loading. Dotted curves for MDL_S of 5 and 2 dB are the capacities of the highly frequency-selective channel.

Plotting the outage curves of Fig. 8 based on a Gaussian approximation of the PDF of the instantaneous channel capacity $\boldsymbol{C}$,

$$P_{\mathrm{out}}(C_T) = 0.5\,\mathrm{erfc}\left(\frac{\langle \boldsymbol{C} \rangle - C_T}{\sqrt{2}\sigma_{\boldsymbol{C}}}\right)\,, \tag{18}$$

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

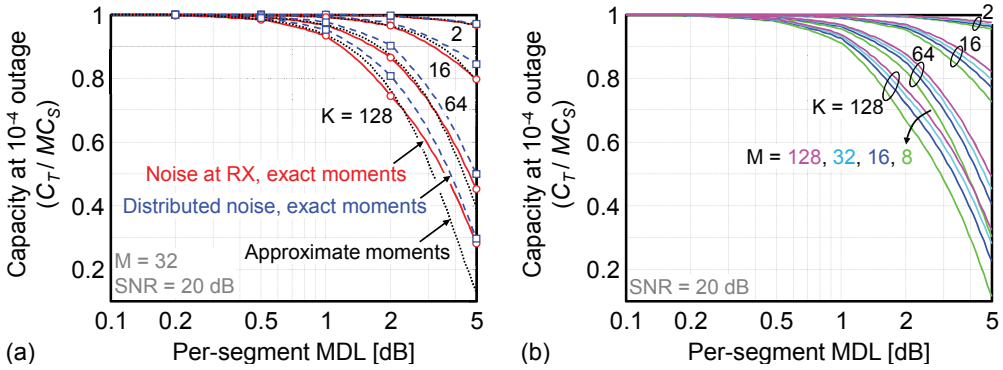

Fig. 9. System capacity at 10^{-4} outage probability as a function of per-segment MDL and number of segments K. (a) Systems with $M = 32$ modes and 20 dB SNR. Red circles and blue squares are numerically exact results for receive-side and distributed noise loading, respectively. Solid red and dashed blue lines represent the Gaussian approximation with numerically exact moments. Dotted black lines use approximate moments. (b) Systems with $M = \{8, 16, 32, 128\}$ modes and $K = \{2, 16, 64, 128\}$ segments, represented by the Gaussian approximations with exact moments for 20 dB SNR and noise loading at the receiver.

where erfc is the complementary error function, and using the numerically exact mean $\langle \boldsymbol{C} \rangle$ and standard deviation $\sigma_{\boldsymbol{C}}$ extracted from the simulated statistics of $\boldsymbol{C}$ results in curves that are indistinguishable from their numerically exact counterparts over a wide parameter range. This gives us confidence that Eq. (18) can be used to predict system performance at much lower outage probabilities than 10^{-4}, where full Monte Carlo simulations can be very time consuming.

Evaluating the MIMO capacity for the frequency-flat channel, normalized to M times the single-mode capacity, at an outage probability of 10^{-4} as a function of the per-segment MDL for K =2, 16, 64, and 128 segments results in the curves shown in Fig. 9. In (a), we focus on $M = 32$ modes and show the numerically exact outage calculations (symbols), both with noise loading at the receiver (red, circles) and with distributed noise loading (blue, squares). The solid red lines and the dashed blue lines give the corresponding analytical results following Eq. (18),

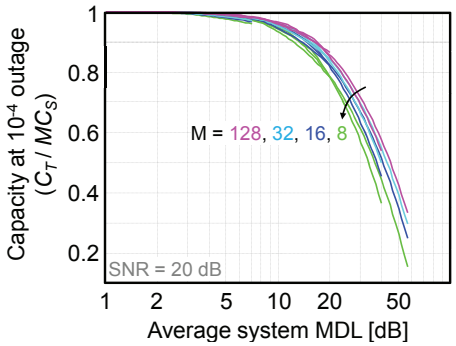

Fig. 10. System capacity at 10^{-4} outage probability as a function of the *aggregate* system MDL with the number of segments K and modes M as parameters.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

using the numerically exact values for mean and standard deviation of $\boldsymbol{C}$, which expectedly match the exact outage curves. The dotted black lines represent outage calculations based on Eq. (18) using the approximate moments of Eqs. (16) and (17). A reasonable agreement with the exact results is evident. As expected, the SDM channel capacity drops with the per-segment MDL as well as with the number of segments. For 128 segments, a per-segment MDL of 1 dB is tolerable to still achieve over 90% of the ideal channel capacity. Figure 9(b) includes both K and M as parameters. As expected from our scaling rules [Eqs. (16) and (17)], we observe only a mild dependence of the outage curves on M. Intuitively, system performance is slightly worse at lower mode count because fewer matrix elements need to be aligned in a 'bad' way to produce a 'bad' channel, which makes such adverse situations more likely for smaller M than for larger M.

To further test the scaling of SDM capacity with K, we plot the outage curves for the frequency-flat channel and distributed noise loading as a function of the *aggregate* average system MDL, $\langle \mathbf{MDL_\Sigma} \rangle_{[\mathrm{dB}]}$, Eq. (15), in Fig. 10. This lets all curves with equal M essentially collapse, as expected from the scaling relations [Eqs. (16) and (17)]. Aggregate average system MDLs of 10 dB are seen to be acceptable for less than a 10% hit in SDM capacity, impressively illustrating the robustness of optical MIMO to MDL effects.

6. Conclusions

Since wavelength-division multiplexing (WDM) is rapidly nearing its scalability limits, new optical multiplexing options must be explored. The only known physical dimension not yet exploited for optical data transport is *space*, which points at space-division multiplexing (SDM) as the only option to further scale the capacity of optical transport networks. The high levels of integration that are necessary to make SDM systems scale as economically attractive as WDM systems will likely introduce non-negligible crosstalk among multiple parallel transmission paths. Multiple-input-multiple-output (MIMO) signal processing techniques known from wireless communications provide the method of choice to deal with crosstalk in SDM systems. However, optical SDM systems operate under a different set of constraints than wireless MIMO systems. Among the most important differentiating factors between wireless MIMO and optical SDM are:

- Optical nonlinearities in the transmission fiber, which are likely to imply per-mode peak power limitations as opposed to a per-transmitter average power limitation;
- The existence of a finite number of modes supported by the optical transmission waveguide, all of which can be individually excited by the transmitter and coherently detected by the receiver;
- The distributed nature of optical noise, which in optical SDM is added throughout the system as opposed to being lumped at the receiver;
- Very high system reliability requirements (i.e., very low acceptable outage probabilities) typical of high-capacity optical transport networks; and
- Round-trip propagation delays that may exceed the channel dynamics and hence make feedback from receiver to transmitter difficult.

In this paper, we discussed MIMO capacities in optical SDM systems with an emphasis on system outage considerations. We pointed out that in order to achieve the high availability standards required for optical transport networks, SDM transponders should be capable of individually addressing, and preferably MIMO processing, all modes supported by the optical SDM

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011

waveguide. We discussed the impact of distributed optical noise loading and mode-dependent loss (MDL) on system capacity and system outage. Through extensive numerical simulations, we extracted scaling rules for mode-average and mode-dependent loss and motivated the Gaussian assumption for SDM capacity statistics, leading to a closed-form expression for the system's outage performance. We then showed that MIMO SDM systems composed of up to 128 segments and supporting up to 128 modes can tolerate up to about 1 dB of per-segment MDL at 90% of the system's full capacity at an outage probability of 10^{-4}. This corresponds to a total acceptable average system MDL of over 10 dB.

Appendix

In this appendix, we show that for *any* given channel matrix H with r channel eigenvalues λ_i, the potential open-loop BLAST capacity C satisfies

$$C \leq r \, \log_2 (1 + \mathrm{SNR}\, M/r) \; . \tag{19}$$

Further, if the SDM waveguide is represented by a unitary matrix and the MIMO channel is formed by any $M_R \times M_T$ submatrix thereof, we show that

$$C \leq r \log_2 (1 + \mathrm{SNR}) = r\, C_S \; . \tag{20}$$

To prove the above inequalities, we first recall that the SDM waveguide matrix H$'$ is normalized such that $\mathrm{tr}(\mathrm{H}'\mathrm{H}'^{\dagger}) = M$. The $M_R \times M_T$ MIMO channel matrix H, a submatrix of H$'$, can at best conserve the energy transmitted over the channel, i.e.,

$$\mathrm{tr}(\mathrm{H}\mathrm{H}^{\dagger}) = \sum_{i=1}^{r} \lambda_i = \mu \leq M \; . \tag{21}$$

Using this constraint, it can be shown by a simple optimization process that the MIMO capacity [Eq. (3)] is largest if all λ_i are equal, i.e., if $\lambda_i = \mu/r$ for $1 \leq i \leq r$, resulting in a potential MIMO capacity of

$$C = r \log_2 (1 + \mathrm{SNR}\, \mu/r) \; . \tag{22}$$

Since $\mu \leq M$, inequality [Eq. (19)] is proven.

If the SDM waveguide is represented by a $M \times M$ unitary matrix H$'$ = U, we find that $\mu \leq r = \min\{M_T, M_R\}$, which together with Eq. (22) proves inequality [Eq. (20)]. To see this, we write the MIMO channel matrix H, a submatrix of U, as $\mathrm{S}_R\, \mathrm{U}\, \mathrm{S}_T$, with S_R and S_T being $M \times M$ selector matrices that pick out the rows and columns of U interfacing to transmitter and receiver; S_R and S_T are diagonal matrices with, respectively, M_R and M_T ones at the desired locations on their diagonals, all remaining matrix elements being zero. With these definitions, we have

$$\begin{aligned} \mathrm{tr}\{\mathrm{H}\mathrm{H}^{\dagger}\} &= \mathrm{tr}\{\mathrm{S}_R\, \mathrm{U}\, \mathrm{S}_T\, \mathrm{S}_T\, \mathrm{U}^{\dagger} \mathrm{S}_R\} \\ &= \mathrm{tr}\{\mathrm{S}_R\, \mathrm{U}\, \mathrm{S}_T\, \mathrm{U}^{\dagger}\} = \mathrm{tr}\{\mathrm{S}_R\, \mathrm{D}\} \\ &\leq \sum_{i=1}^{M} \lambda_i(\mathrm{S}_R) \lambda_i(\mathrm{D}) = \min\{M_T, M_R\} \; . \end{aligned} \tag{23}$$

In going from the first to the second line we used the fact that $\mathrm{S}_T \mathrm{S}_T = \mathrm{S}_T$ and $\mathrm{tr}\{\mathrm{AB}\} = \mathrm{tr}\{\mathrm{BA}\}$ for arbitrary matrices A and B. We then identified $\mathrm{U}\, \mathrm{S}_T\, \mathrm{U}^{\dagger}$ as the eigendecomposition of some matrix D whose M_T non-zero eigenvalues $\lambda_i(D)$ are all unity. In going from the second to the third line, we applied a matrix theorem given in [52, 53].

Acknowledgment

We acknowledge valuable discussions with A. R. Chraplyvy, R. W. Tkach, H. Kogelnik, S. K. Korotky, R. Ryf, S. Randel, and D. Chizhik.

#148738 - $15.00 USD Received 6 Jun 2011; revised 4 Aug 2011; accepted 9 Aug 2011; published 15 Aug 2011