Applications of Random Matrix Theory on Fiber Optical Communications

A. Moustakas

A. Karadimitrakis (Athens)
P. Vivo (KCL)
R. Couillet (Paris-Central)
L. Sanguinetti (Pisa)
A. Muller (Huawei)
H. Hafermann (Huawei)

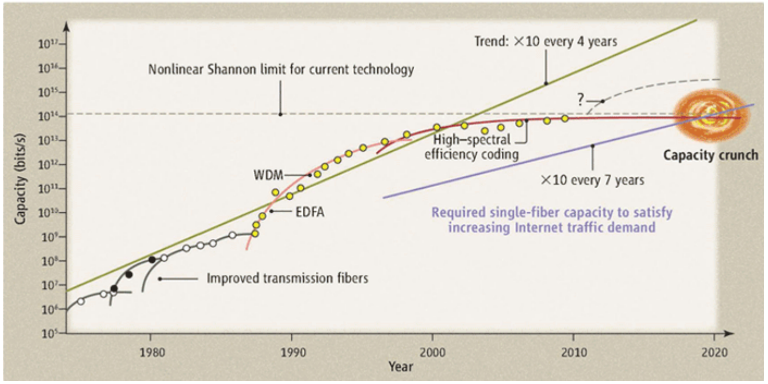
Introduction

- <u>Need for high speed infrastructure</u>

Bandwidth (BW) hungry technologies emerging (~2 dB increase per year)

-Wired: IPTV, telepresence,

online gaming, live streaming etc.



➡ - “Capacity- crunch” eminent

– <u>Solutions so far</u>
  - Optical Networks (WDM-DWDM) have exploited many degrees of freedom (BW, available power, polarization diversity) except one: spatial.

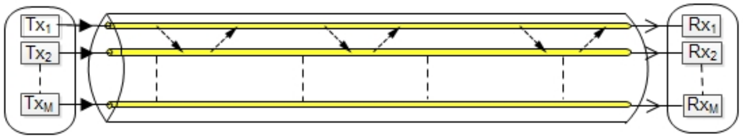
– Soon the online devices will exceed the number of global population!

– Use of optical Space Division Multiplexing (SDM) is compelling.
  - Ideally: without changing the already infrastructure…

Introduction

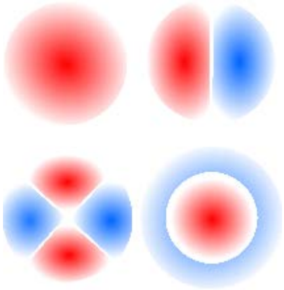
- <u>What is Optical MIMO?</u>

  Just like in wireless domain, in optical we can use N parallel transmission paths to greatly enhance the capacity of the system.



  – First paper on Optical MIMO: H. R. Stuart, "Dispersive multiplexing in multimode optical fiber," Science **289**(5477), 281–283 (2000)

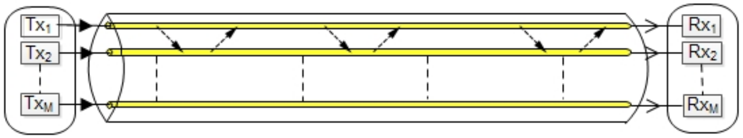
  – Multi-Mode Fibers (MMF): Use multiple modes to carry information



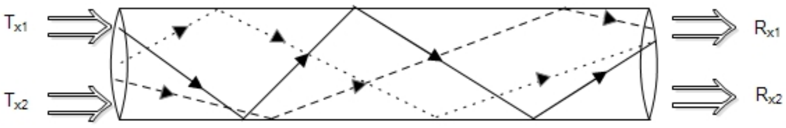
  – Multi-Core Fibers (MCF): Utilize different optical paths of different cores in the same fiber (within the same cladding)

Introduction

Problem: Crosstalk phenomenon arises



Crosstalking between adjacent cores in MCF



Light beam scattering resulting in crosstalking in MMF

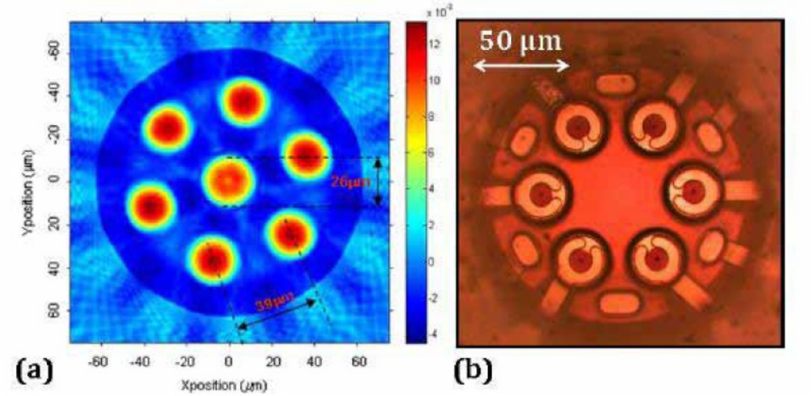
Due to :

– Extensive fiber length
– Bending of fiber
– Limited area with multiple power distributions
– Light beam scattering
– Non-linearities

Introduction

Problem: Crosstalk phenomenon

- Two Approaches:
  - Fight it



Section of 7-core optical fiber

Introduction

Problem: Crosstalk phenomenon

- Two Approaches:
  - Take advantage of it

"Classic" MIMO techniques required but with some twists:

- Low power constraints to avoid non-linear behavior.
- Optical channel matrix just a subset of a unitary matrix.

System Model

Consider a single-segment N-channel lossless optical fiber system: $N_t \leq N$ transmitting channels excited, $N_r \leq N$ receiving channels coherently. The $2N \times 2N$ scattering matrix is

$$\mathbf{S} = \begin{bmatrix} \mathbf{r_t} & \mathbf{t} \\ \mathbf{t}^{\mathbf{T}} & \mathbf{r_r} \end{bmatrix} \quad (\mathbf{S}=\mathbf{S}^{\mathrm{T}})$$

Only $\mathbf{t}$ (Haar-distributed $\mathbf{t}^{\dagger}\mathbf{t} = \mathbf{t}\mathbf{t}^{\dagger} = \mathbf{I}_{\mathrm{N}}$) sub-matrix is of interest ($\mathbf{r}$ is ~0).

Generally N_r , $N_t < N$:

- Other channels may be used from different, parallel transceivers
- Modelling of loss: additional energy lost during propagation

System Model

Only **t** (Haar-distributed $\mathbf{t}^\dagger\mathbf{t} = \mathbf{t}\mathbf{t}^\dagger = \mathbf{I}_N$) sub-matrix is of interest (**r** is ~0).

$$\mathbf{t} = \underbrace{}_{N_r}\overbrace{\begin{bmatrix} t_{11} & \dots & t_{1N} \\ & \dots & \\ t_{N1} & \dots & t_{NN} \end{bmatrix}}^{N_t}$$

Define N_t x N_r matrix **U** as

$$\mathbf{U} = \mathbf{P}_{N_t}{}^T \mathbf{S} \mathbf{P}_{N_r}$$

- where P projection operator:

$$\mathbf{P}_{N_t} = \begin{bmatrix} \mathbf{I}_{N_t} \\ \mathbf{0} \end{bmatrix} \begin{matrix} \}\ N_t \\ \}\ N\text{-}N_t \end{matrix}$$

System Model

Channel Equation:

$$\mathbf{y} = \mathbf{U}\mathbf{x} + \mathbf{z}$$

– Assume no differential delays between channels (*frequency flat fading*)the mutual information is

$$I_N(\mathbf{U}) = \log\det\left(\mathbf{I} + \rho\mathbf{U}^{\dagger}\mathbf{U}\right) = \sum_{k=1}^{N_t}\log\left(1 + \rho\lambda_k\right)$$

- Gaussian noise **z**
- Receiver knows the channel (pilot)
- Transmitter does not know the channel
- w.l.o.g. $N_t \leq N_r$

Information Metrics

– Outage Probability:

$$P_{out}(r) = \text{Prob}\left(I_N < N_t r\right) = E_{\mathbf{U}}\left[\Theta\left(N_t r - I_N(\mathbf{U})\right)\right]$$

- Optimal: Assumes infinite codewords

What is the price of finite codelengths?

– Gallager error bound for *M*-length code:

$$P_{err}(r) < E_{\mathbf{U}}\left[\exp\left[M \max_{0\le k\le 1}\left(N_t kr - k \log\det\left[\mathbf{I} + \frac{\rho}{1+k}\mathbf{U}^{\dagger}\mathbf{U}\right]\right)\right]\right]$$

Coulomb – Gas Analogy

- Joint probability distribution of eigenvalues of $\mathbf{U}^\dagger\mathbf{U}$

$$P(\lambda_1, \lambda_2, \ldots, \lambda_{N_t}) \propto \exp\left[N_0 \sum_k \log(1-\lambda_k) + (N_r - N_t)\sum_k \log\lambda_k + 2\sum_{k>m}\log|\lambda_k - \lambda_m|\right]$$

  - $N_0 = N - N_1 - N_2 \geq 0$

- Exponent is energy of point charges repelling logarithmically in the presence of external field

$$P \propto e^{-N_t^2 S_0[p]}$$

- Large N: charges coalesce to density

$$S_0 = \int dx\ p(x) V_{eff}(x) - \int\int dxdy\ p(y)\,p(x)\log|x-y|$$

$$V_{eff}(x) = -n\log(1-x) - (\beta-1)\log x$$

  - where $n = \dfrac{N_0}{N_t}$ $\quad \beta = \dfrac{N_r}{N_t}$

- Minimizing (convex) S_0 w.r.t p(x) gives the “Marcenko-Pastur” distribution

$$p_{MP}(x) = \frac{\sqrt{(x-a)(b-x)}}{2\pi x(1-x)} \qquad a,b = \left(\frac{\sqrt{n+1} \pm \sqrt{\beta(n+\beta)}}{n+\beta+1}\right)^2$$

Outage Probability

- We need tails of distribution: Optical Comms operate at $p_{out} \approx 10^{-8}$

$$P_{out}(r) = \mathrm{Prob}\left(I_N < N_t r\right) = E_{\mathbf{U}}\left[\Theta\left(N_t r - I_N(\mathbf{U})\right)\right]$$

  - Fourier transform (or use large deviations arguments)

$$S \to S_0 - k\int dx \;\; p(x)\left(\log(1+\rho x) - r\right)$$

$$V_{eff}(x) \to V_{eff}(x) - k\log\left(1+\rho x\right)$$

- k plays role of strength of logarithmic attraction/repulsion at $x_0 = -\rho^{-1}$
  - k>0 shifts charge density to larger values (R>Rerg), k<0 to smaller ones (R<Rerg)
- Minimizing S w.r.t p(x) gives the *generalized* Marcenko Pastur distribution
- Equivalent to balancing forces on the charge located at *x*:

$$2P\int \frac{p(x')}{x-x'}dx = \frac{n}{1-x} - \frac{\beta-1}{x} - \frac{k\rho}{1+k\rho}$$

Generalized MP equation

- Use Tricomi theorem to calculate p(x)
- Obtain closed form expr. for energy S[p]
- E.g. $\beta>1$; n>0

$$p(x)=\frac{\sqrt{(b-x)(x-a)}}{2\pi x(1+\rho x)}\left(\frac{n(1+\rho)}{(1-x)\sqrt{(1-a)(1-b)}}+\frac{\beta-1}{x\sqrt{ab}}\right)$$

  - a, b, k calculated from

$$p(b)=p(a)=0$$
$$r=\int_a^b dx\, p(x)\log(1+\rho x)$$
$$1=\int_a^b dx\, p(x)$$

Generalized MP equation

- In general

	S_{0b}	S_{ab}	S_{01}	S_{a1}
	a=0 b<1	a>0 b<1	a=0 b=1	a>0 b=1
n=0; β=1	$r<r_{c1}$	--	$r_{c1} < r < r_{c2}$	$r>r_{c2}$
n>0; β=1	$r<r_{c3}$	$r>r_{c3}$	--	--
n=0; β>1	--	$r<r_{c4}$	--	$r>r_{c4}$
n>0; β>1	--	all r	--	--

- Phase transitions (a=0 to a>0) etc are third order
  - discontinuous $S'''(r_c)$
  - Relation to Tracy-Widom (?)

Distribution Density of r

Finally

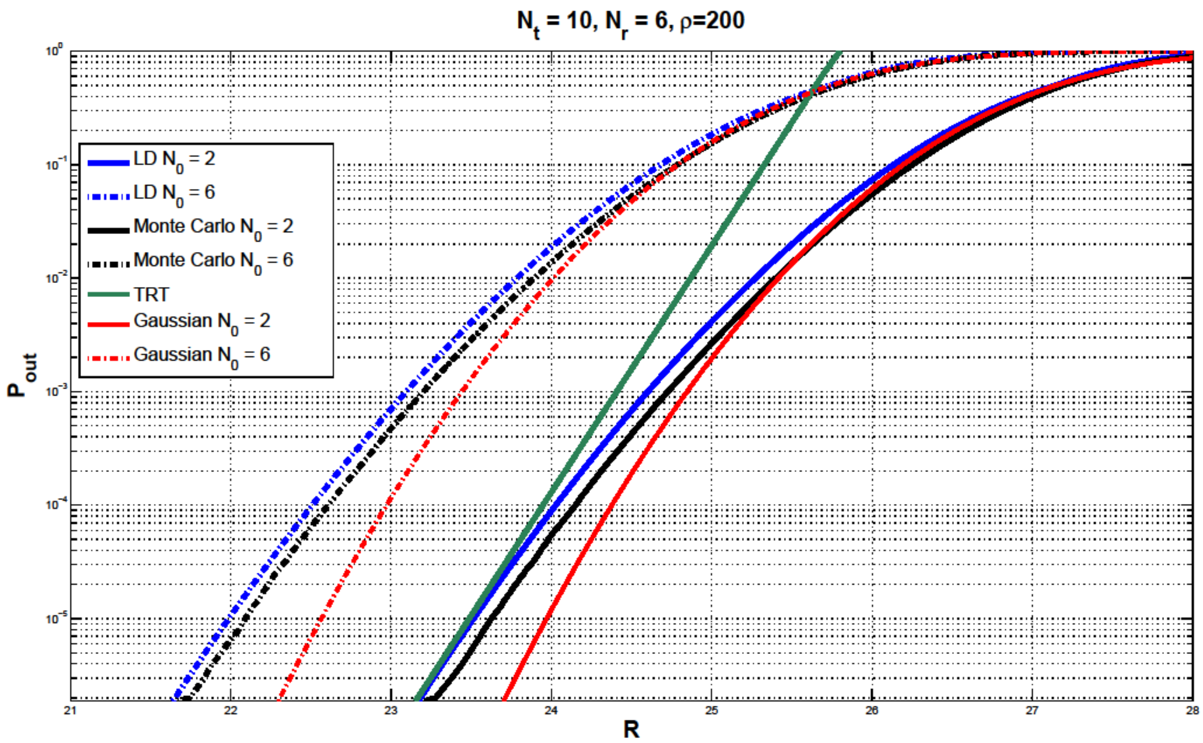
$$f(r) \approx N_t \frac{e^{-N_t^2\left[S(r)-S(r_{erg})\right]}}{\sqrt{2\pi v_{erg}}}$$

– where

$$v_{erg} = \frac{1}{S''(r_{erg})} = \log\left[\frac{\left(\sqrt{1+\rho b_0} - \sqrt{1+\rho a_0}\right)^2}{4\sqrt{1+\rho a_0}\sqrt{1+\rho b_0}}\right]$$

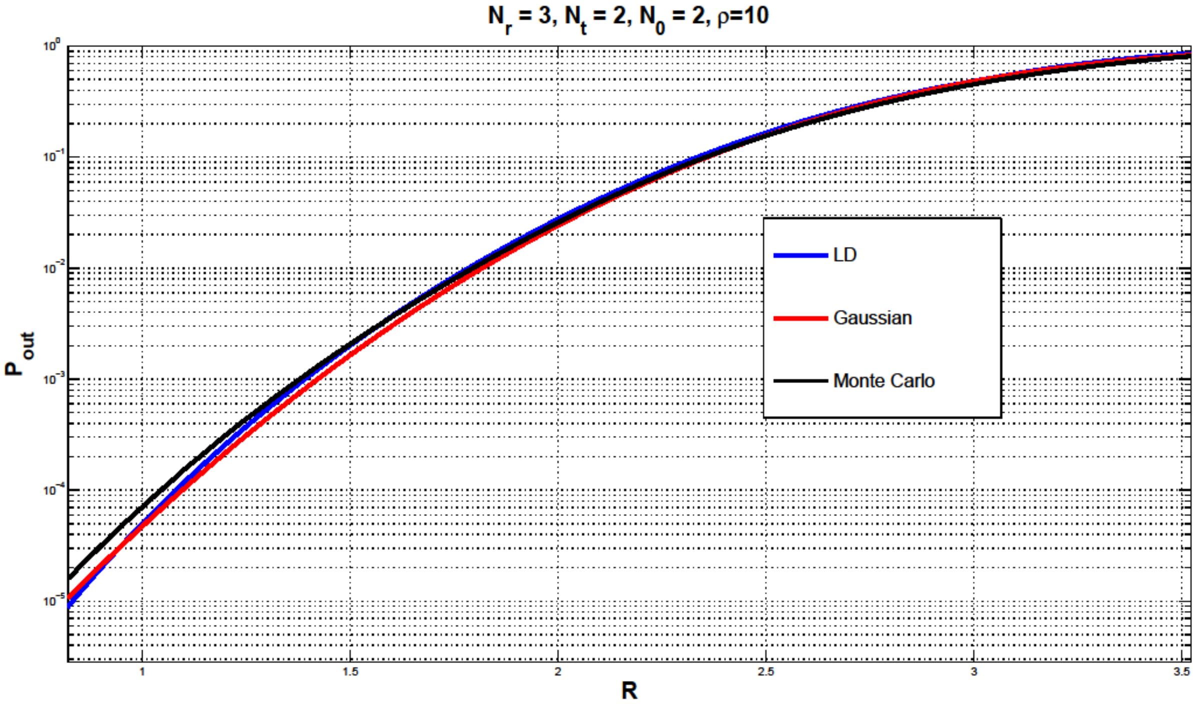
is the variance at the peak of the distribution.

Numerical Simulations ($\beta > 1$ and $n_0 > 0$)



The LD approach demonstrates better behavior, following Monte Carlo.

Numerical Simulations ($\beta > 1$ and $n_0 > 0$)



The LD approach demonstrates better behavior, following Monte Carlo.

For small values of N_r, N_t and N_0, the discrepancy is minimal

Finite Block-Length Error Probability

$$P_{err}(r) < E_{\mathbf{U}}\left[\exp\left[M\max_{0\le k\le 1}\left(N_t kr - k\sum_j \log\left(1+\frac{\rho}{1+k}\lambda_j\right)\right)\right]\right]$$

- Here we need

$$S \to S_0 + k\alpha\int dx\ \ p(x)\left[\log\left(1+\frac{\rho}{1+k}x\right)-r\right]$$

$$V_{eff}(x) \to V_{eff}(x) + k\alpha\log\left(1+\frac{\rho}{1+k}x\right)$$

  – where $\alpha = \dfrac{\mathrm{M}}{\mathrm{N}_t}$

- Now k is bounded in [0,1]
- Also when k<1

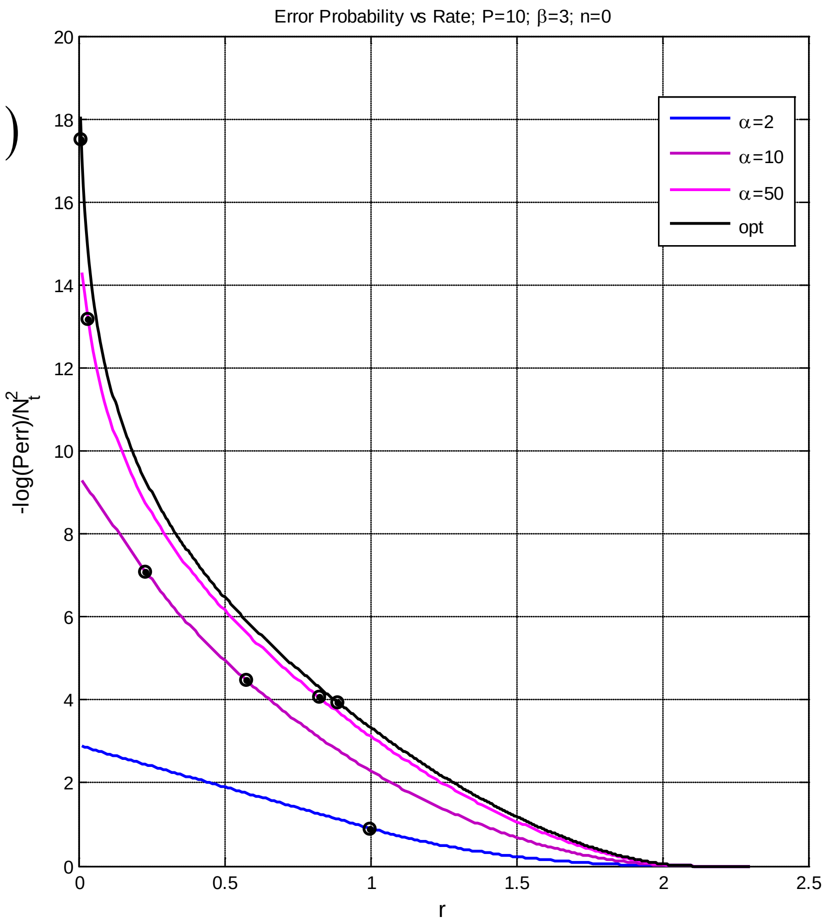
$$r = \int_a^b dx p(x)\left[\log\left(1+\frac{\rho}{1+k}x\right)+\frac{k}{1+k}\frac{\rho x}{1+\rho x}\right]$$

  – otherwise no constraint

Numerical Simulations (β > 1 and $n_0 = 0$)

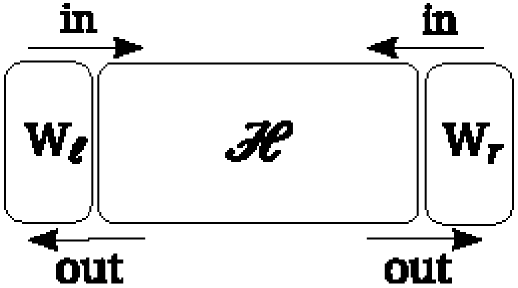
There are two types of phase transition points here:

- r_{c1}: At k=1 point – discontinuous $S''(r_{c1})$
- r_{c2}: Discontinuous $S'''(r_{c2})$



More Realistic Channel

- Chaotic cavity picture:



$$\mathbf{S} = \mathbf{I} - 2\pi i \mathbf{W}^{+}\left(H + i\pi \mathbf{W}\mathbf{W}^{+}\right)^{-1}\mathbf{W}$$

- Assuming very low backscattering at edges we obtain

$$\mathbf{U} = \mathbf{P}_r^{+}\mathbf{S}\mathbf{P}_t = -c\mathbf{P}_r^{+}\left(\mathbf{H}_0 + \gamma\mathbf{G} + i\mathbf{\Gamma}\right)^{-1}\mathbf{P}_t$$

  - Deterministic (mode energy) $\mathbf{H}_0$
  - Random (complex Gaussian) $\mathbf{G}$
  - Diagonal loss matrix $\mathbf{\Gamma}$

More Realistic Channel

- Mutual Information metric:

$$I(\mathbf{G}) = \log\det\left[\mathbf{I} + \rho \mathbf{P}_r^{+}\left(\mathbf{H}_0 + \gamma\mathbf{G} + i\mathbf{\Gamma}\right)^{-1}\mathbf{P}_t\mathbf{P}_t^{+}\left(\mathbf{H}_0 + \gamma\mathbf{G} - i\mathbf{\Gamma}\right)^{-1}\mathbf{P}_r\right]$$

- I – Distribution (mean-variance) can be obtained using replica theory

$$E\left[I(\mathbf{G})\right] = I_1 - I_2$$

$$I_1 = \log\det\left[(\rho+\delta+\gamma t)(1+\gamma r) + \left(\mathbf{H}_0 - \gamma p\right)^2\right] - N\left(tr - p^2\right)$$

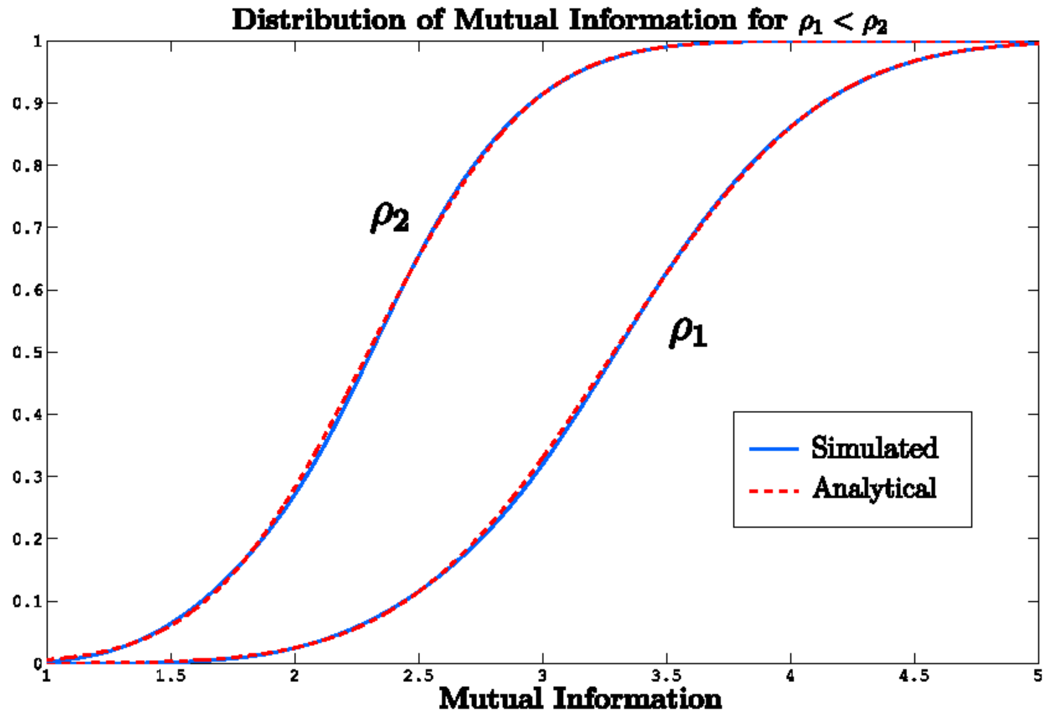
$$r = \frac{1}{N}Tr\left[\frac{\gamma(1+\gamma r)}{(\rho+\delta+\gamma t)(1+\gamma r) + \left(\mathbf{H}_0 - \gamma p\right)^2}\right]$$

$$p = \frac{1}{N}Tr\left[\frac{\gamma(\mathbf{H}_0 - \gamma p)}{(\rho+\delta+\gamma t)(1+\gamma r) + \left(\mathbf{H}_0 - \gamma p\right)^2}\right]$$

$$t = \frac{1}{N}Tr\left[\frac{\gamma(\rho+\delta+\gamma t)}{(\rho+\delta+\gamma t)(1+\gamma r) + \left(\mathbf{H}_0 - \gamma p\right)^2}\right]$$

  – I_2 same with ρ=0 – also expressions for variance

Numerical Simulations



Conclusions

- MIMO: promising idea in Optical Communications

- In this work:
  - Simple model for Optical MIMO channel
  - Large Deviation Approach provides tails for MIMO mutual information
  - Method provides metric for outage throughput and finite blocklength error

- Many issues still open:
  - Channel modeling still at its infancy
  - Transmitter/Receiver Architectures
  - Multiple fiber segments
  - Nonlinearities: Signal becomes interference
  - …