INSTITUTE
OF COMMUNICATION,
INFORMATION
AND PERCEPTION
TECHNOLOGIES

Scuola Superiore
Sant'Anna

Achievable information rates in optical fiber communications

Marco Secondini

Acknowledgments:
Enrico Forestieri, Domenico Marsella
Erik Agrell

2015 Munich Workshop on Information Theory of Optical Fiber (MIO2015)
Munich, Germany, December 7-8, 2015

Outline

- Introduction
- A bit of information theory
- The optical fiber channel and its models
- Some numerical results and bounds
- Discussion and conclusions.

Outline

- **Introduction**
- A bit of information theory
- The optical fiber channel and its models
- Some numerical results and bounds
- Discussion and conclusions.

What we do know about the optical fiber channel …

At **low powers** (linear regime):

- We have an **explicit channel model**
- We know how to design systems that operate **close to channel capacity**
- Capacity increases with power as in the **AWGN channel**

At **high powers** (nonlinear regime):

- Signal propagation is governed by the **NLSE** (Manakov) equation
- Conventional systems reach an optimum operating point, after which their performance **decreases with power**
- It's been impossible (so far) to increase the information rate beyond a certain **limit** (nonlinear Shannon limit?)

… and what we don't know

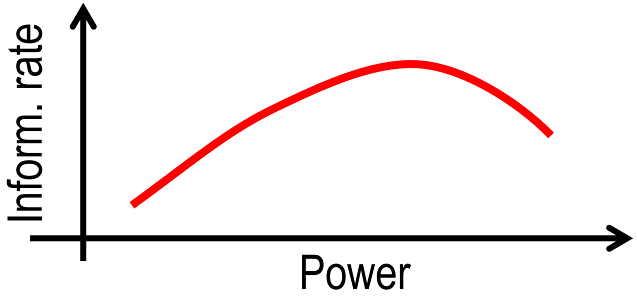
At **high powers** (nonlinear regime):

- We don't have an **explicit channel model**
- We don't know what is the **optimum detector**
- We don't know what is the **optimum input distribution**
- … **annel capacity**

Channel capacity: position of the scientific community

Pessimists ☹

Systems are substantially limited by the so-called **nonlinear Shannon limit**

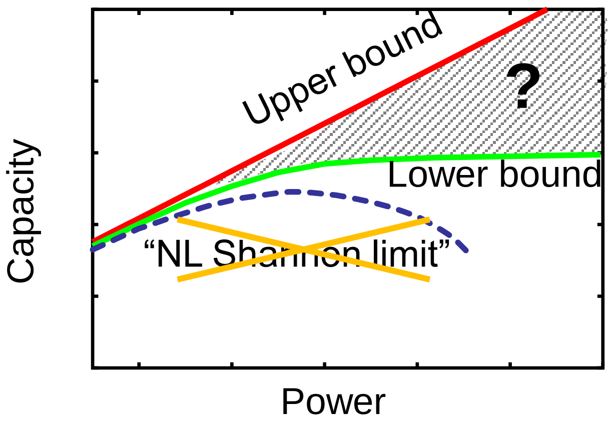


- A. Splett et al. "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," ECOC 1993
- P. P. Mitra et al. "Nonlinear limits to the information capacity of optical fiber communications," Nature 2001.
- R.-J. Essiambre et al. "Capacity limits of optical fiber networks," JLT 2010.
- G. Bosco et al. "Analytical results on channel capacity in uncompensated optical links with coherent detection," Opt. Exp. 2011.
- A. Mecozzi et al., "Nonlinear Shannon limit in pseudolinear coherent systems," JLT 2012.

Optimists ☺

Higher information rates can be achieved

- R. Dar et al. "New bounds on the capacity of the nonlinear fiber-optic channel, Opt. Lett. 2014.
- M. Secondini et al. "On XPM mitigation in WDM fiber-optic systems," PTL 2014.



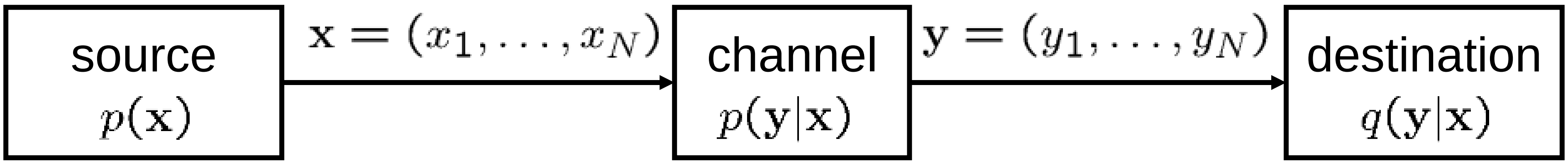
Channel capacity **might** be unbounded

- K. S. Turitsyn et al. "Information capacity of optical fiber channels with zero average dispersion," PRL 2003.
- E. Agrell et al. "Influence of behavioral models on multiuser channel capacity," JLT 2015.
- G. Kramer et al. "Upper bound on the capacity of a cascade of nonlinear and noisy channels," ITW 2015.

Outline

- Introduction: the capacity problem
- **A bit of information theory**
- The optical fiber channel and its models
- Some numerical results and bounds
- Discussion and conclusions.

Discrete-time channels and achievable rates



Modulation

Mismatched decoding

Achievable information rate [*]

$$I_q(\mathbf{X};\mathbf{Y}) \triangleq \lim_{N\to\infty} \frac{1}{N} E_p \left\{ \log_2 \frac{q_{\mathbf{y}|\mathbf{x}}(\mathbf{y} \mid \mathbf{x})}{\int p_{\mathbf{x}}(\mathbf{x}) q_{\mathbf{y}|\mathbf{x}}(\mathbf{y} \mid \mathbf{x}) d\mathbf{x}} \right\}$$

- Practical **lower bound** to average mutual information and capacity
- **Achievable** with given modulation and mismatched decoder
- **Easily evaluated** through numerical simulations
- No need to know the true channel law $p_{\mathbf{y}|\mathbf{x}}$

[*] D. M. Arnold et al. “Simulation-based computation of information rates for channels with memory,” *IEEE Trans. Inform. Theory*, v. 52, pp. 3498–3508, 2006.

Relation between AIR and channel capacity

- Capacity is obtained by maximizing AIR w.r.t. *p*(**x**) and *q*(**y**|**x**)

$$C = \max_{p_{\mathbf{x}}, q_{\mathbf{y}|\mathbf{x}}} I_q(\mathbf{X}; \mathbf{Y})$$

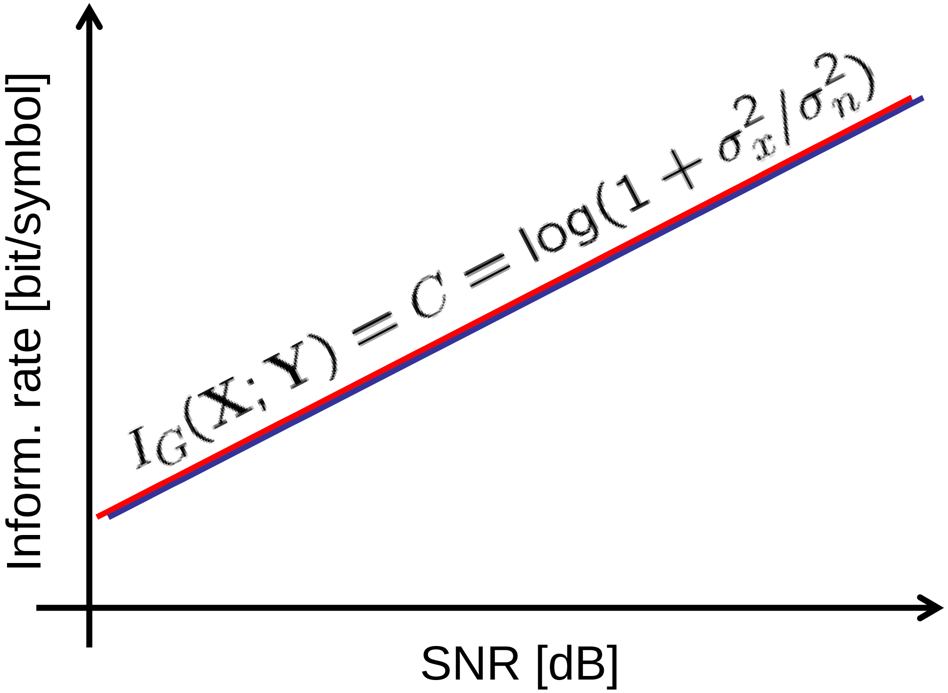
- A common capacity **lower bound** is the AIR with i.i.d. Gaussian inputs $p_{\mathbf{x}}$ and an assuming a $q_{\mathbf{y}|\mathbf{x}}$ matched to an AWGN channel with same input-output correlation

$$I_G(\mathbf{X}; \mathbf{Y}) = \log_2 \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sigma_y^2 - |\sigma_{xy}|^2} \leq C$$

- In general, the bound may be **loose**.

Examples: AWGN channel [*]

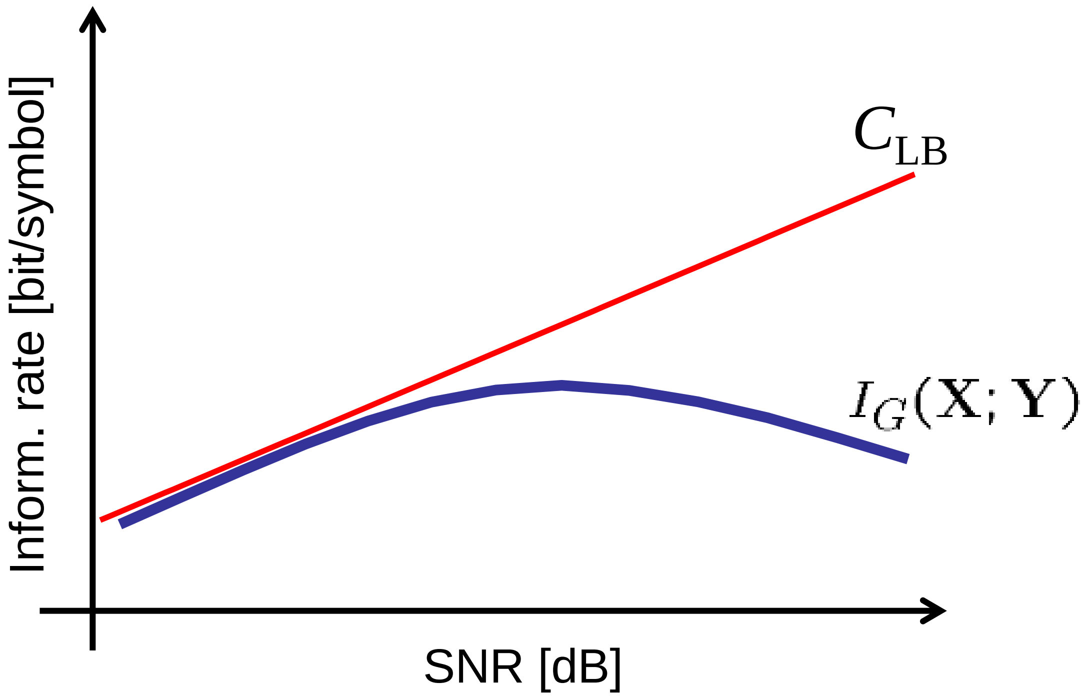
$$y_k = x_k + n_k$$



[*] C. E. Shannon, “A Mathematical Theory of Communication”, Bell Sys. Tech J., 1948

Examples: nonlinear phase-noise channel [*]

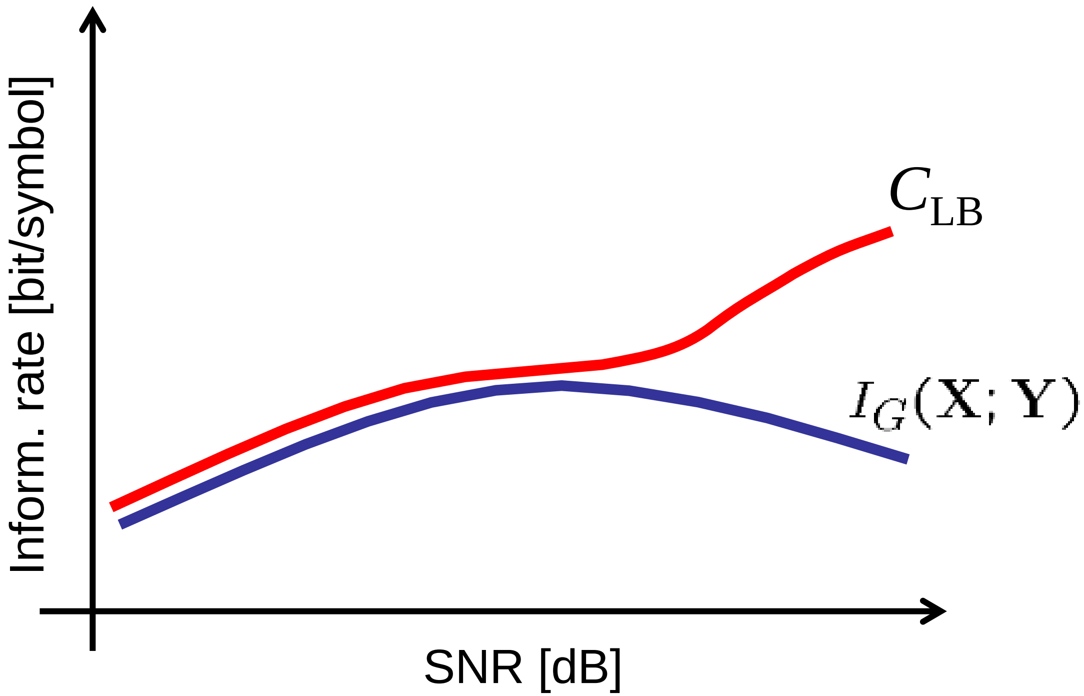
Limited by signal-noise interaction



[*] K. S. Turitsyn et al. “Information capacity of optical fiber channels with zero average dispersion,” PRL 2003.

Examples: rudimentary FWM channel [*]

Limited by (nonlinear) inter-channel interference
(all channels with same power and distribution)

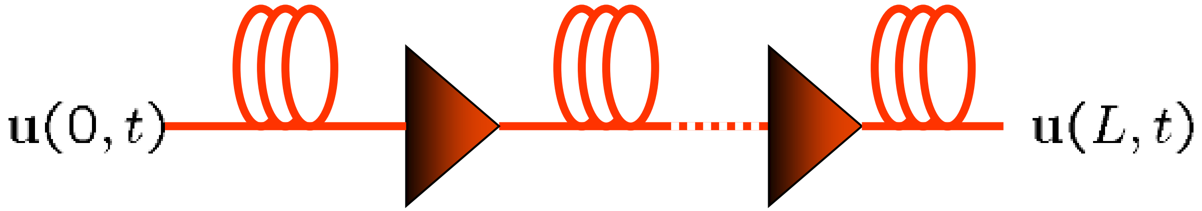


[*] E. Agrell et al. "Influence of behavioral models on multiuser channel capacity," JLT 2015.

Outline

- Introduction: the capacity problem
- A bit of information theory
- **The optical fiber channel and its models**
- Some numerical results and bounds
- Discussion and conclusions.

The fiber-optic waveform channel



Signal propagation is governed by the noisy and lossy **Manakov equation** (**nonlinear Schrödinger equation** (**NLSE**) for single polarization signals)

$$\frac{\partial \mathbf{u}}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{u}}{\partial t^2} - j\gamma a(z)|\mathbf{u}|^2\mathbf{u} + \mathbf{n}(z,t)$$

Dispersion

Attenuation/amplification

Nonlinearity

Noise

This equation defines an **implicit** model for a **waveform** channel

Solving the equation (40 years later)

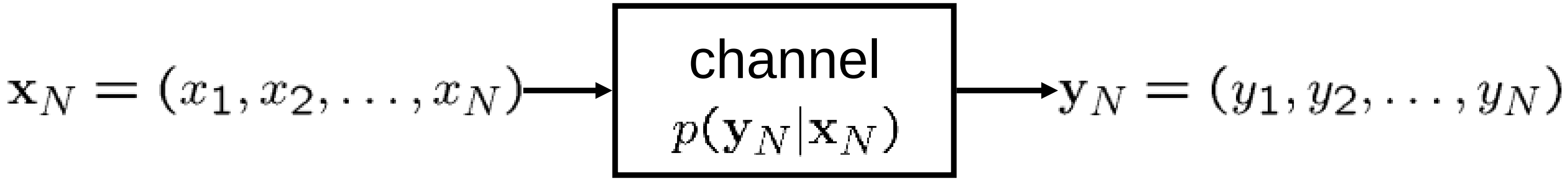
- The NLSE for the optical fiber (Hasegawa & Tappert, 1973)
- The split-step Fourier method (Hardin & Tappert, 1973)
- The inverse scattering transform (Zhakarov & Shabat, 1972)

↓ 40 years later

- ➢ Some refinements of the methods have been studied
- ➢ The SSFM is still the most used approach
- ➢ The IST is the hottest topic of the moment

- ➢ Perturbation methods to account for the presence of noise

Explicit versus *implicit* channel models



- **Implicit** model: allow to **draw samples** from p
- **Explicit** model: allow **to compute** p (analytically/easily)

<u>Approximated models</u>

– Gaussian noise model

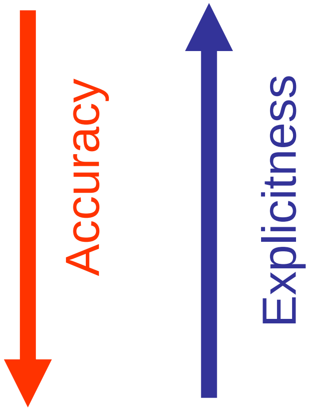
⋮

– Perturbation methods

– Nonlinear Fourier transform

⋮

– Split-step Fourier method



Perturbation methods

- Applied both to the **NLSE** and to the Zhakarov-Shabat system
- Used to model **inter-channel NL**, intra-channel NL, signal-noise interaction, …
- Regular, **logarithmic**, combined, …

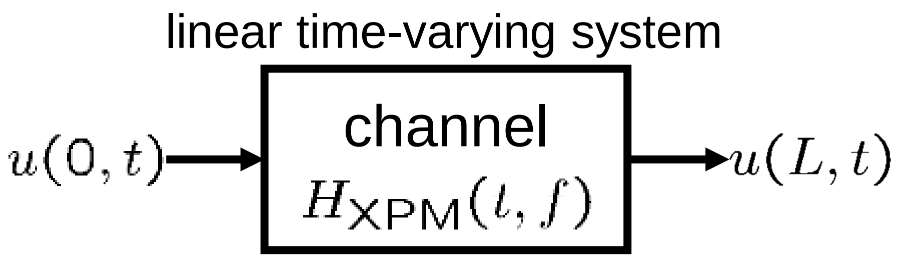
Inter-channel nonlinearity: a linear time-varying model…
… for a nonlinear time-invariant system

- Propagation in WDM systems (signal-noise interaction and FWM negligible)

$$\frac{\partial u}{\partial z} = j\frac{\beta_2}{2}\frac{\partial^2 u}{\partial t^2} - j\gamma g(z)(\cancel{|u|^2} + 2|w|^2)u$$

Get rid of it by single-channel **backpropagation**

- **Linear Schrödinger equation** with a time- and space-varying stochastic potential



- P. P. Mitra, J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications", *Nature*, 2001.
- M. Secondini, E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation", *PTL*, 2012
- R. Dar et al., "Time varying ISI model for nonlinear interference noise", *OFC*, 2014.

Frequency-resolved logarithmic perturbation model

Time-varying transfer function

$$u(0,t) \longrightarrow \boxed{H_{\mathsf{XPM}}(t,f) = e^{-j\theta(t,f)}} \longrightarrow u(L,t)$$

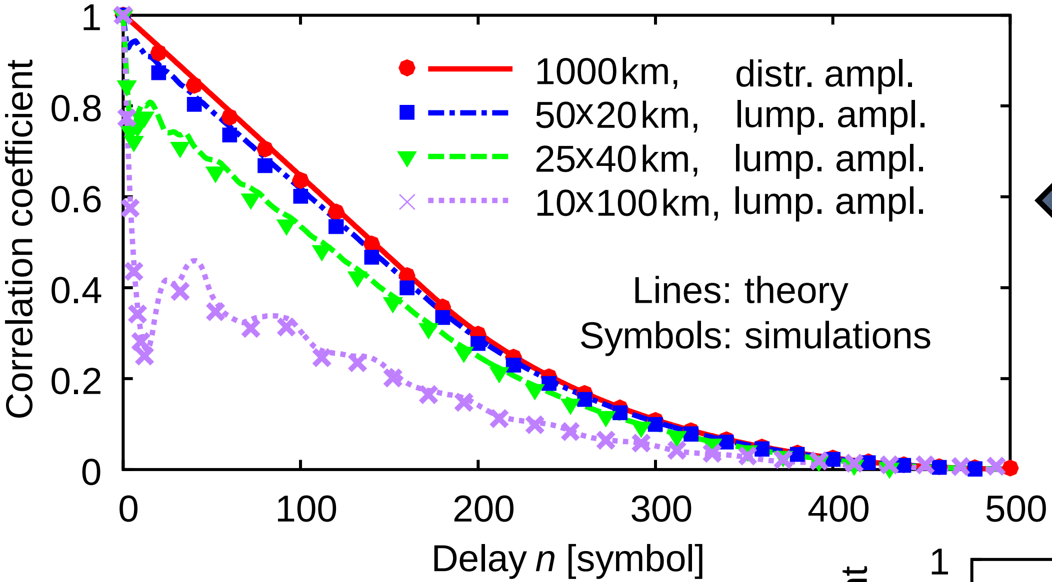
<u>XPM term</u> $\theta(f,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(f,\mu,\nu)W(\mu)W^{*}(\nu)e^{j2\pi(\mu-\nu)t}d\mu d\nu$

— depends on symbols transmitted by the other users (channels)

— shows significant **correlation** both in **time** and **frequency**

XPM causes linear ISI (with time-varying coefficients) and can be mitigated by an adaptive linear equalizer (Kalman algorithm)

Channel coherence

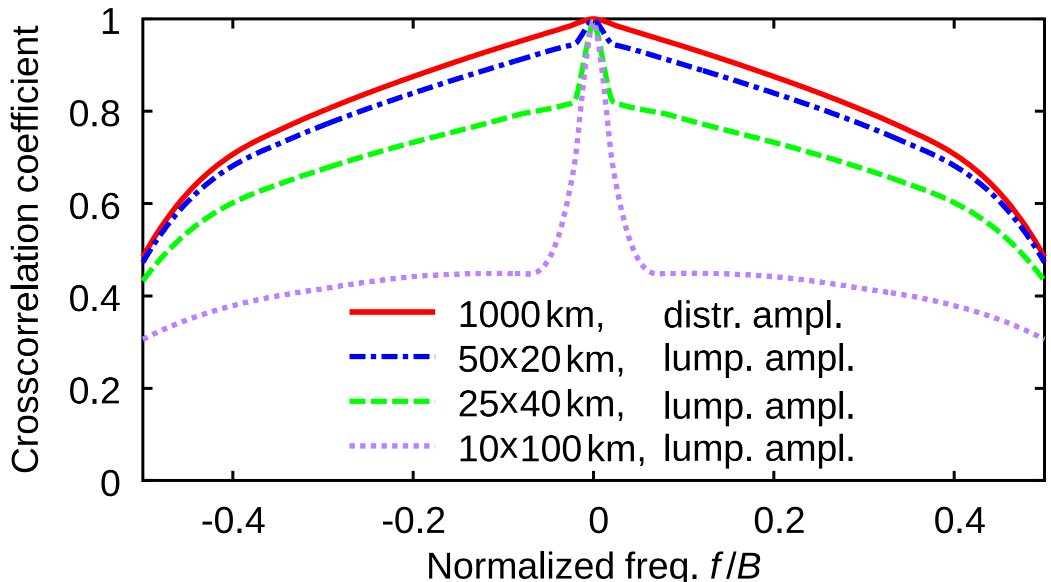


Coherence **time**

depends on amplification!

Coherence **bandwidth**

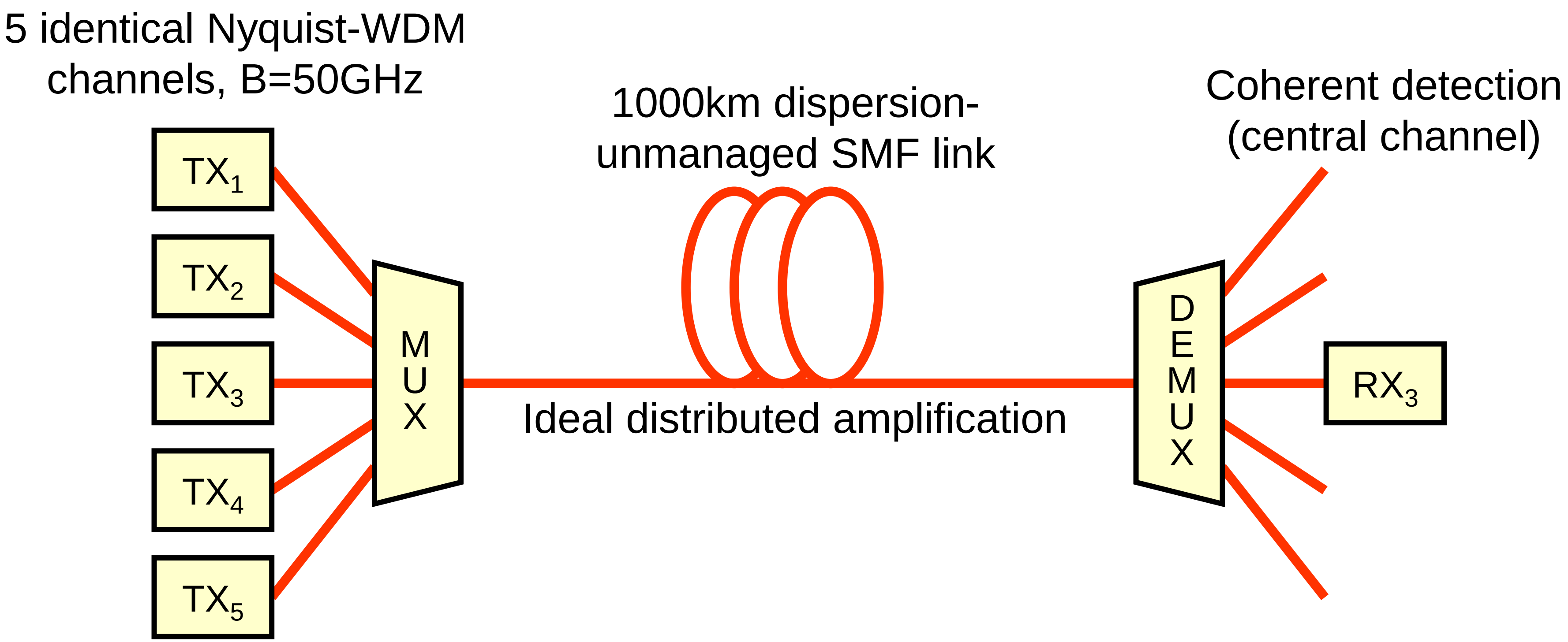
depends on amplification!



Outline

- Introduction: the capacity problem
- A bit of information theory
- The optical fiber channel and its models
- **Some numerical results and bounds**
- Discussion and conclusions.

Assumptions about the system



Computation of AIRs

Numerically
(MC simulations and SSFM)

$$I_q(\mathbf{X};\mathbf{Y}) \triangleq \lim_{N\to\infty} \frac{1}{N} E_p\left\{\log_2 \frac{q_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\mid\mathbf{x})}{\int p_{\mathbf{x}}(\mathbf{x})\, q_{\mathbf{y}|\mathbf{x}}(\mathbf{y}\mid\mathbf{x})\, d\mathbf{x}}\right\}$$

i.i.d. Gaussian

Different (mismatched)
approximated models

Capacity bounds

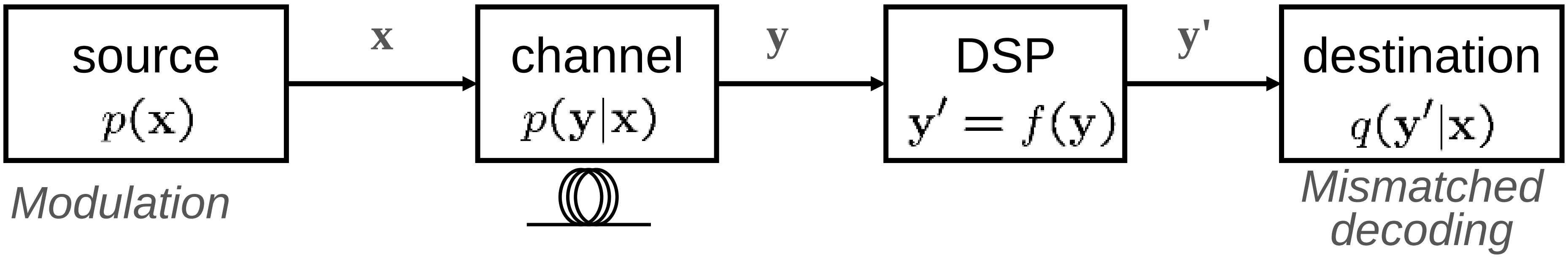
- **Capacity lower bound**
  - Capacity is a non-decreasing function of power [*]

- **Capacity upper bound**
  - NLSE and Manakov equation preserve energy and entropy [**]

[*] E. Agrell, “Conditions for a monotonic channel capacity,” TCOM 2015.

[**] G. Kramer et al. “Upper bound on the capacity of a cascade of nonlinear and noisy channels,” ITW 2015.

DSP and detection metric

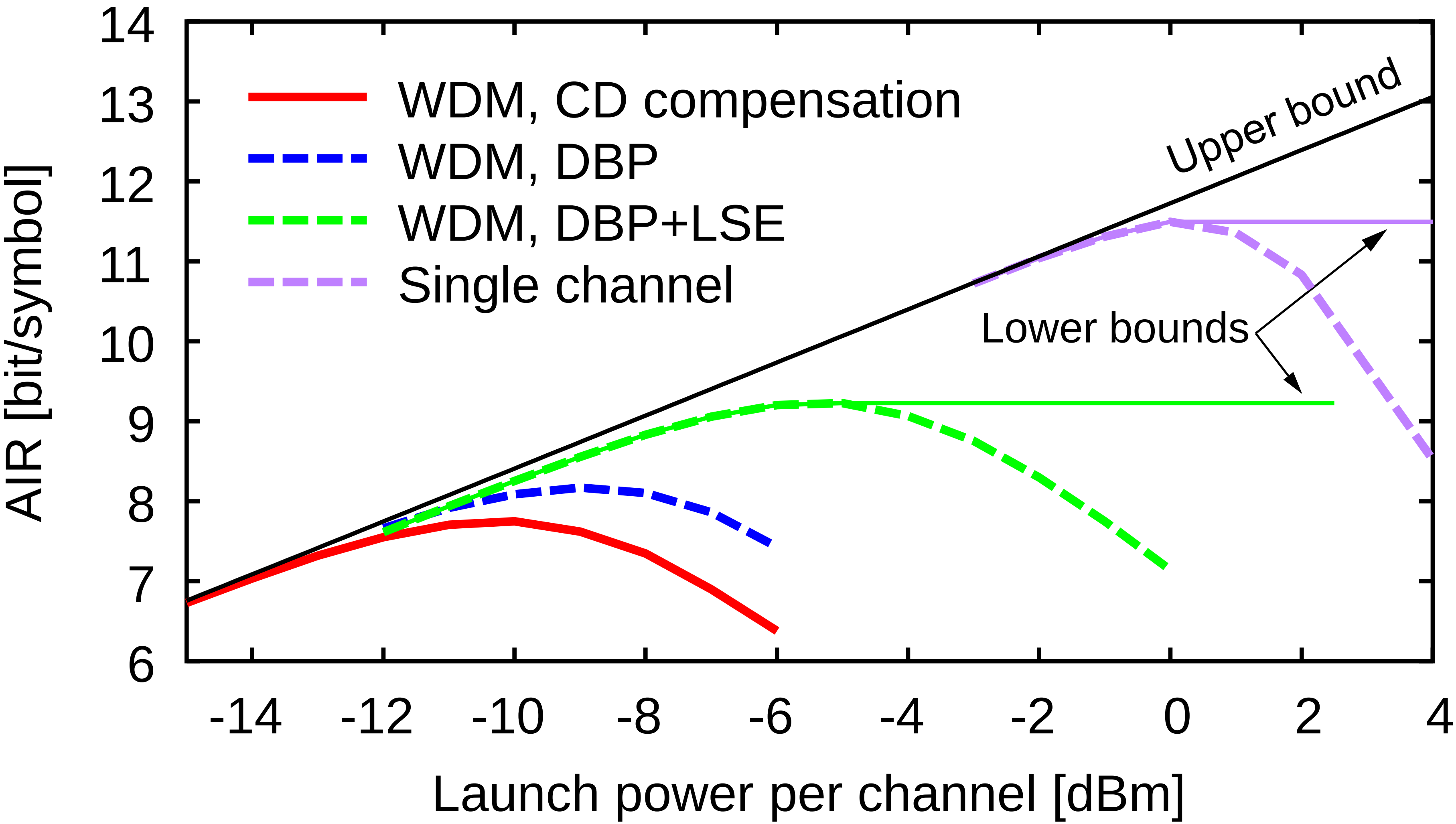


DSP does not change mutual information, but can increase AIR by reducing mismatch between channel and decoder

Different DSP for nonlinearity mitigation

- **Chromatic dispersion (CD) compensation**
  - Dispersion compensation + AWGN detector (i.e., matched to AWGN channel)
  - Optimum detector if the GN model is exact
- **Digital backpropagation (DBP)**
  - Usually based on the SSFM.
  - DBP removes deterministic single-channel nonlinearity
- **Least-square equalization (LSE)**
  - Inter-channel nonlinearity causes linear time-varying ISI (FRLP model)
  - Linear time-varying channel tracked and equalized by linear least-square equalizer

Single-polarization systems



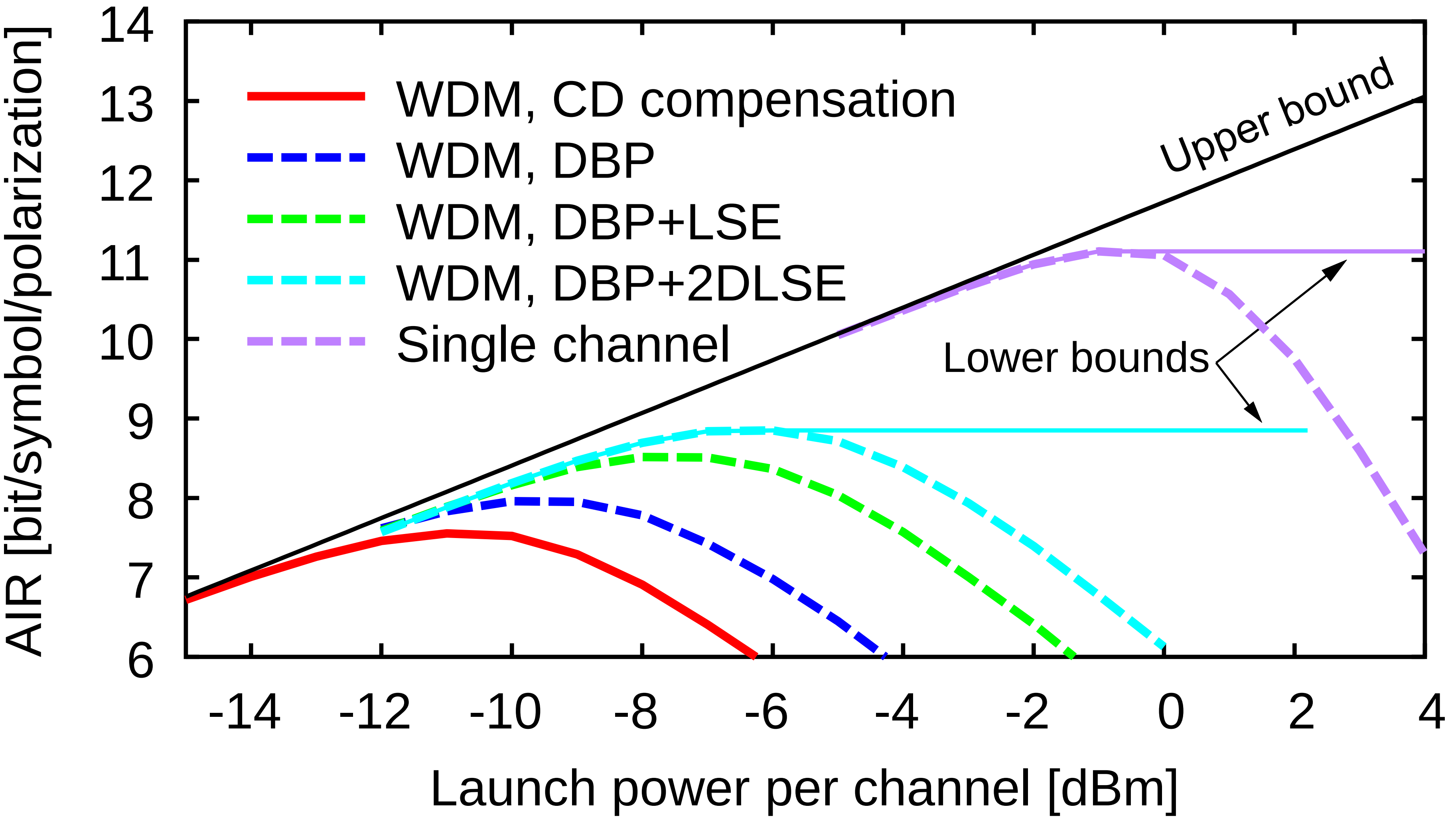
Some improvements

- Single-polarization systems are not efficient
- Least square equalization (LSE) is complicated
- Transmitted symbols (used for LSE) are not available
- Ideal distributed amplification is not practical
- Gains are too small

2D-LSE for polmux systems

- **Chromatic dispersion (CD) compensation**
  - Dispersion compensation + AWGN detector (i.e., matched to AWGN channel)
  - Optimum detector if the GN model is exact
- **Digital backpropagation (DBP)**
  - Usually based on the SSFM.
  - DBP removes deterministic single-channel nonlinearity
- **Least-square equalization (LSE)**
  - Inter-channel nonlinearity causes linear time-varying ISI (FRLP model)
  - Linear time-varying channel tracked and equalized by linear least-square equalizer
- **Two-dimensional least-square equalization (2D-LSE)**
  - Similar to LSE, but employing a two-dimensional equalizer
  - More suitable for Manakov equation

Polarization-multiplexed systems



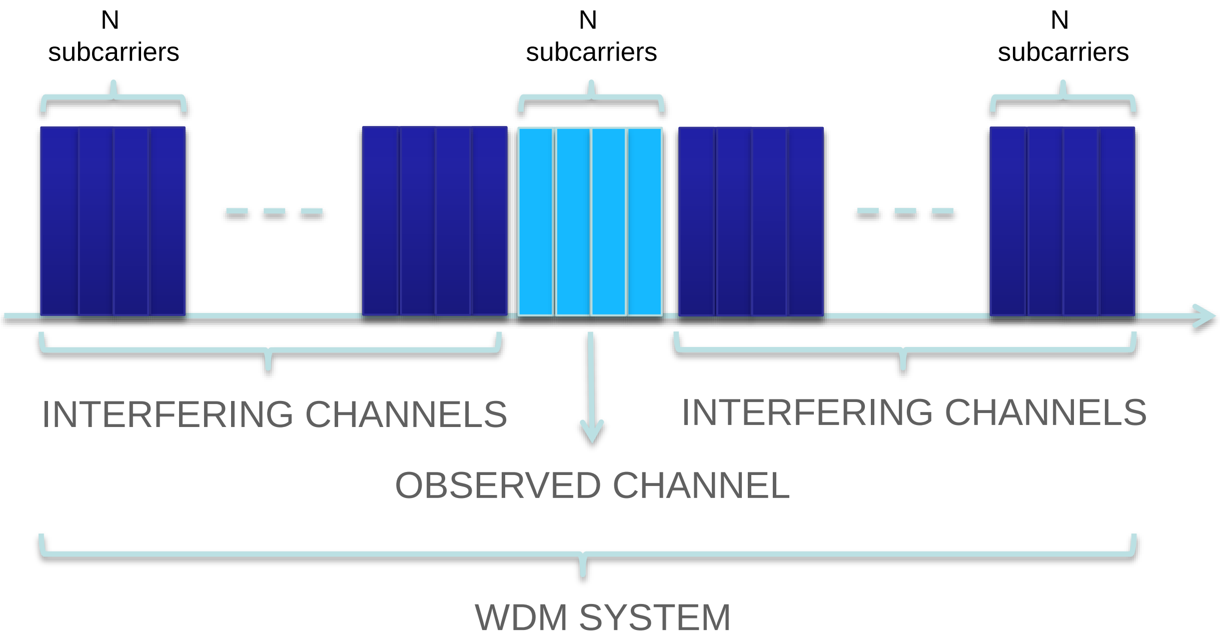
Some improvements

- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
- Transmitted symbols (used for LSE) are not available
- Ideal distributed amplification is not practical
- Gains are too small

Some improvements

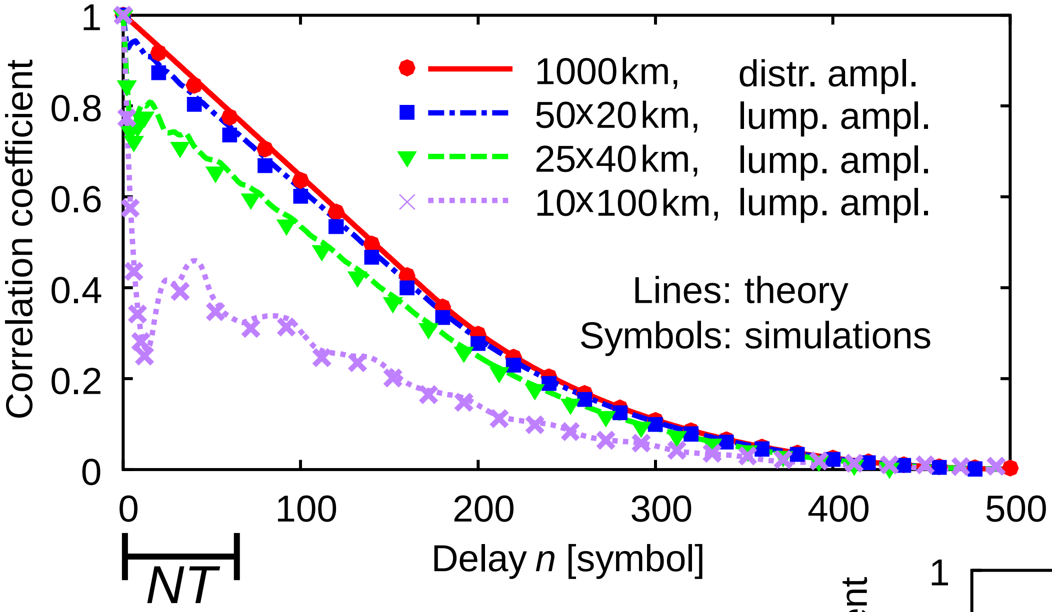
- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
- Transmitted symbols (used for LSE) are not available
- Ideal distributed amplification is not practical
- Gains are too small

Multicarrier modulation



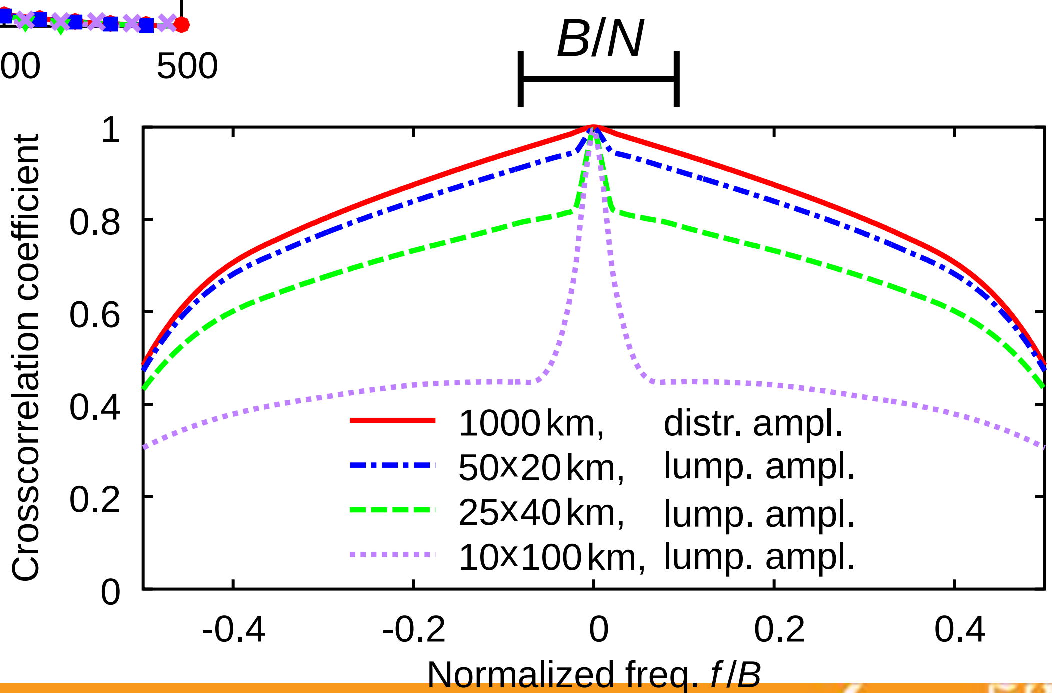
- Each subcarrier has a **narrower bandwidth** (divided by N)
- Each subcarrier has a **longer symbol time** (multiplied by N)

Multicarrier modulation: impact on channel coherence



Longer subchannel symbol time

Narrower subchannel bandwidth



Multi-carrier modulation: detection metric

$\theta(t,f)$ approximately constant:
— over each subband (more accurate for large N)
— during each symbol time (more accurate for small N)

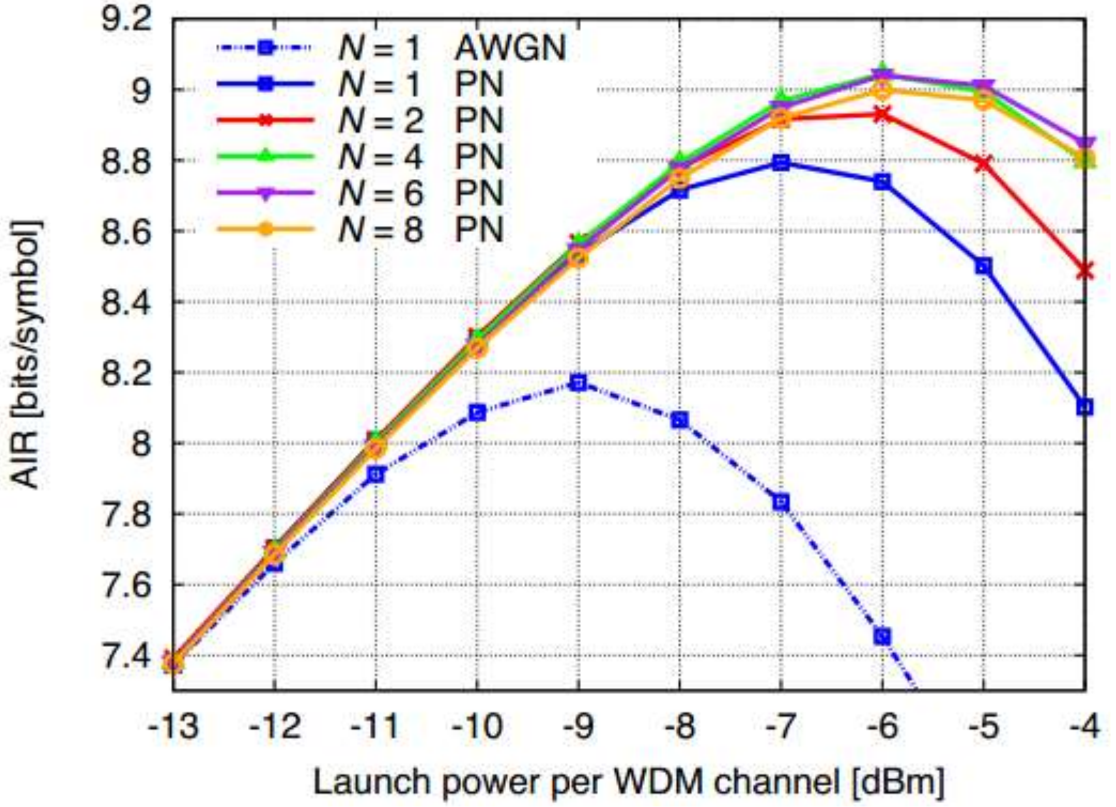
Each subchannel is independently detected
Detection metric $q(\mathbf{y}|\mathbf{x})$ is matched to the following approx. channel model

$$y_k = x_k e^{j\theta_k} + n_k$$

n_k → AWGN

θ_k → phase noise: wrapped AR process

Multi-carrier modulation: AIR



Some improvements

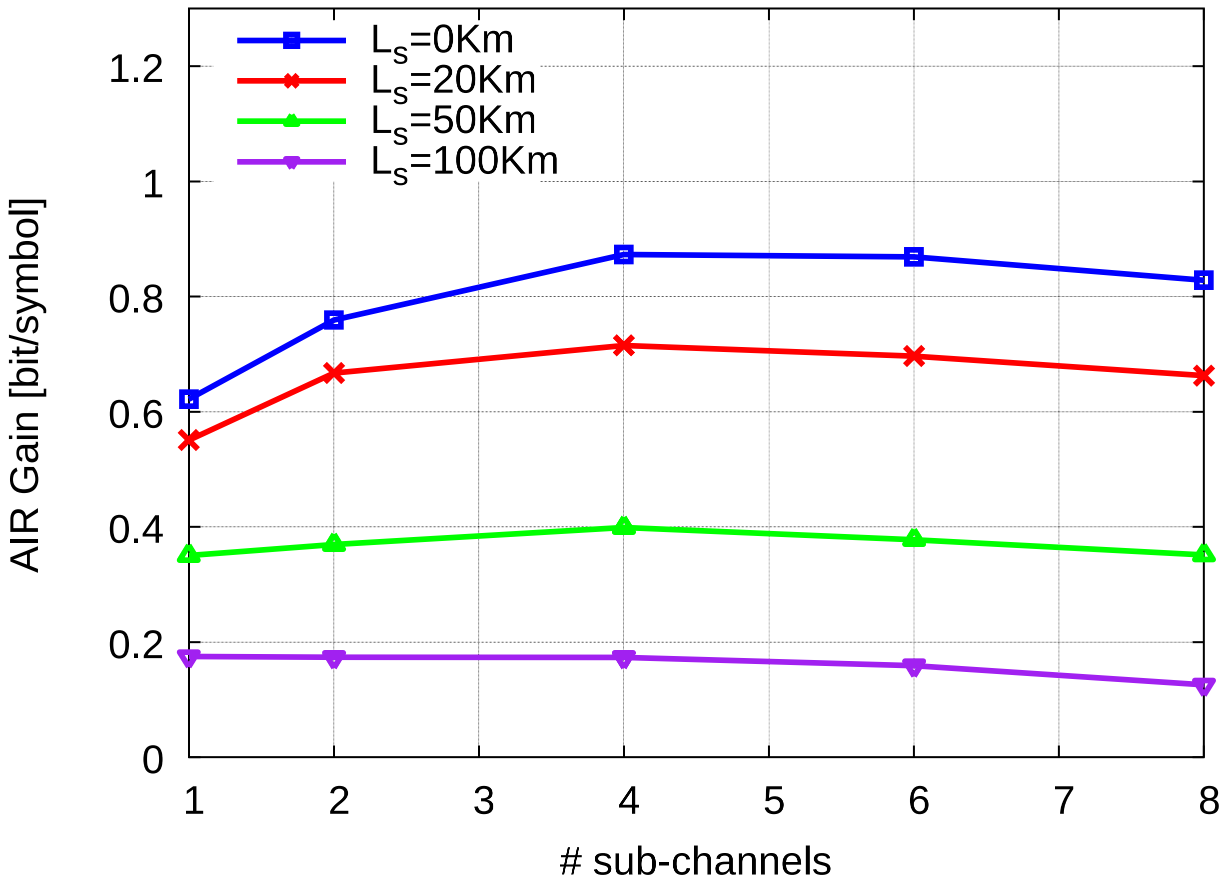
- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
  - ✓ Same gains with multi-carrier modulation and simpler detection
- Transmitted symbols (used for LSE) are not available
  - ✓ They are not needed in the multicarrier approach
- Ideal distributed amplification is not practical

- Gains are too small

Some improvements

- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
  - ✓ Same gains with multi-carrier modulation and simpler detection
- Transmitted symbols (used for LSE) are not available
  - ✓ They are not needed in the multicarrier approach
- Ideal distributed amplification is not practical
- Gains are too small

Max. AIR gain with distributed/lumped amplification



Some improvements

- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
  - ✓ Same gains with multi-carrier modulation and simpler detection
- Transmitted symbols (used for LSE) are not available
  - ✓ They are not needed in the multicarrier approach
- Ideal distributed amplification is not practical
  - ✓ Some gain also with lumped amplification, but it decreases with span length
- Gains are too small
  - ✓ We need to optimize the input distribution

Some improvements

- Single-polarization systems are not efficient
  - ✓ Similar gains can be achieved in polmux systems
- Least square equalization (LSE) is complicated
  - ✓ Same gains with multi-carrier modulation and simpler detection
- Transmitted symbols (used for LSE) are not available
  - ✓ They are not needed in the multicarrier approach
- Ideal distributed amplification is not practical
  - ✓ Some gain also with lumped amplification, but it decreases with span length
- Gains are too small
  - ✓ We need to optimize the input distribution

What is next?

- More accurate channel models $q_{\mathbf{y}|\mathbf{x}}$
- Optimization of the input distribution $p_{\mathbf{x}}$
- Don’t forget about complexity issues!

- Perturbation methods
- Nonlinear Fourier transform
- Particle filtering or other model-agnostic methods

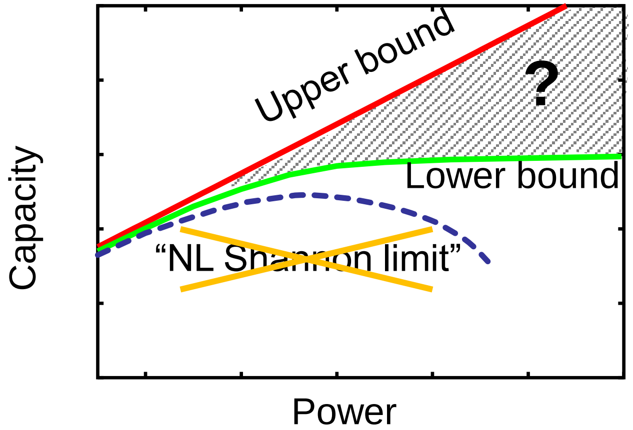
Capacity: final remarks

- ✓ Channel modeling is a crucial step for **nonlinearity mitigation** and **capacity evaluation**

- ✓ Improved detection strategies (based on more accurate models) allow to achieve **higher information rates** w.r.t. the so-called nonlinear Shannon limit

- ✓ A Gaussian input provides a loose bound to channel capacity at high powers, as it causes a highly detrimental nonlinear interference. Much more can be expected by **input optimization**.

➢ The **capacity** problem remains **open**.



thank you!

email: marco.secondini@sssup.it