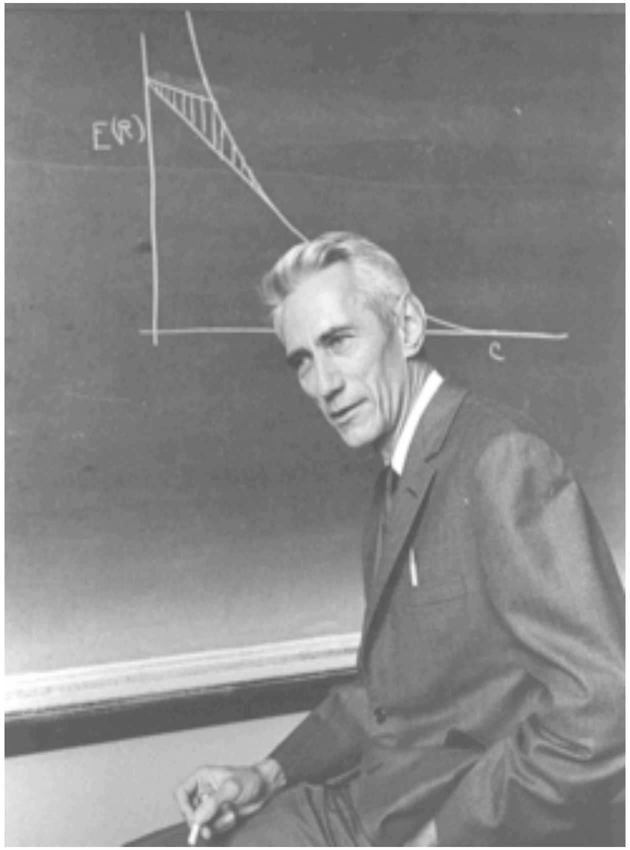
Bounds on the Capacity of Optical Fiber Channels

Gerhard Kramer (TUM)
based on joint work with René-Jean Essiambre (Bell Labs),
Mansoor Yousefi (Telecom Paristech) and Frank Kschischang (Univ. Toronto)

IEEE ITSoc Distinguished Lecture
Linköping University, Sweden
June 2, 2016

Unterstützt von / Supported by
Alexander von Humboldt
Stiftung / Foundation



Claude Elwood Shannon
Apr 30, 1916 – Feb 24, 2001

1) Introduction and Preliminaries

Main Message

An upper bound on the spectral efficiency of a standard optical fiber model

$$\eta \leq \log(1 + SNR) \quad \text{[bits/sec/Hz]}$$

- this is the first upper bound on a “full” model;
- the bound is tight at low SNR;
- the bound is likely extremely loose at high SNR;
  but it’s better than an upper bound of ∞

Preliminaries: Information Theory

- Entropy and Mutual Information: consider random variables **X,Y** with joint distribution $P_{X,Y}(.)$

$$H(X) = E\left[-\log P_X(X)\right] \overset{\text{discrete X}}{=} \sum_{a:P_X(a)\neq 0} -P_X(a)\log P_X(a)$$

$$H(X|Y) = E\left[-\log P_{X|Y}(X|Y)\right]$$

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

- Note: if X is a continuous random variable (e.g., Gaussian) then one calls H(X) a differential entropy and (usually) uses the notation h(X)

Gaussian Random Variables

- Example: Suppose **X,Y** are jointly Gaussian. Then we have

$$h(X) = \frac{1}{2}\log\left(2\pi e\, Var[X]\right)$$

$$h(X|Y) = \frac{1}{2}\log\left(2\pi e\, Var[X]\left(1-\rho^2\right)\right)$$

$$I(X;Y) = \frac{1}{2}\log\left(\frac{1}{1-\rho^2}\right)$$

- Var[X] is the variance of X
- ρ is the correlation coefficient of X and Y; instructive cases: $\rho = 0, \pm 1$

Capacity

- Capacity C of a channel $P_{Y|X}(.)$ is the maximum I(**X**;**Y**) under constraints put on **X**,**Y**
- Example: real-alphabet additive white Gaussian noise (AWGN) channel

$$Y = X + Z$$

with Var[Z]=N and an input power constraint $E[X^2] \leq P$ has

$$I(X;Y) \leq C = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

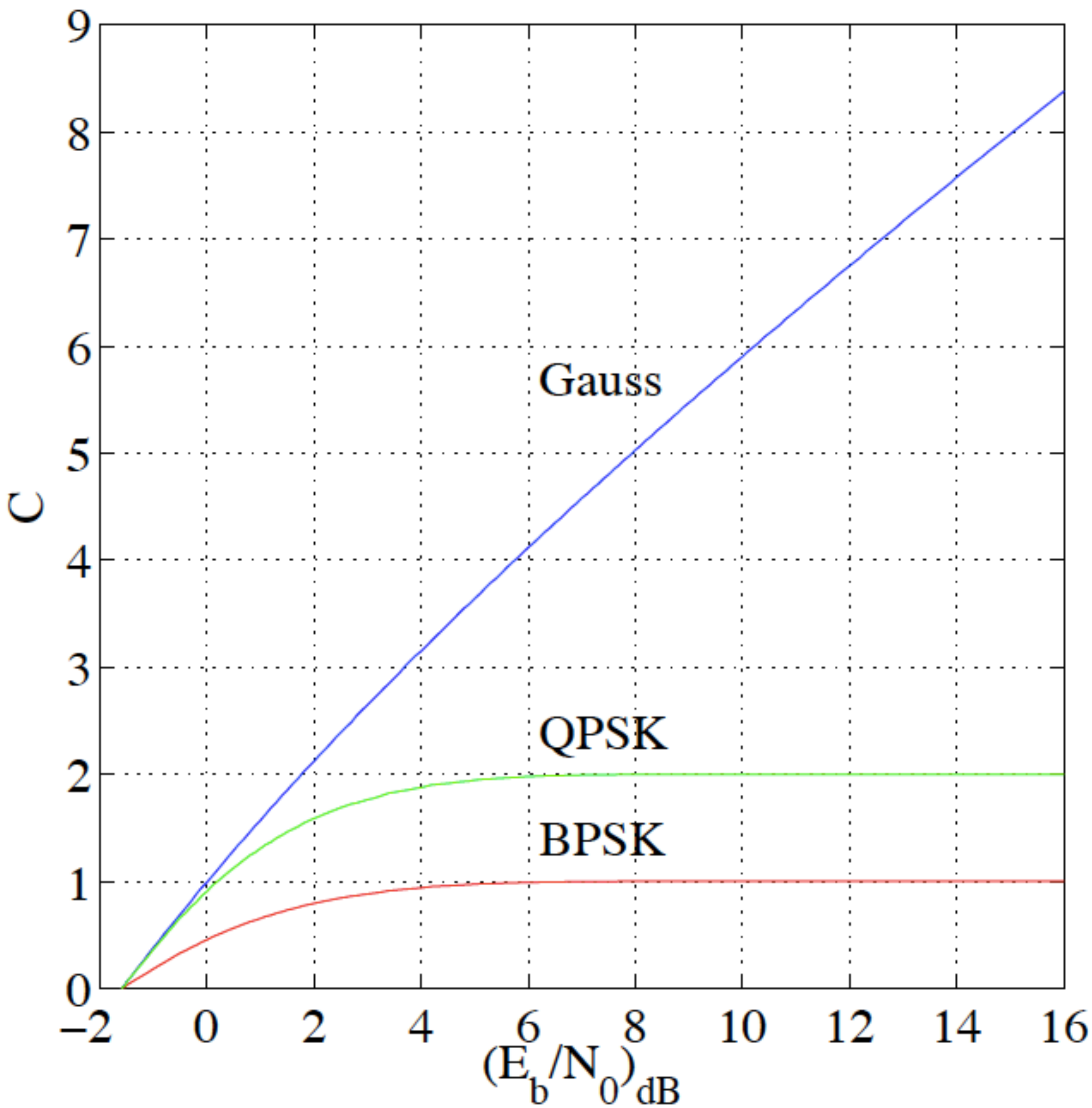
- Complex alphabet AWGN channels: C = log(1+P/N)
- For complex alphabets, N is usually taken as N_0W where N_0 is the (one-sided) noise PSD and W is the bandwidth
- Spectral efficiency is η =C if one uses sinc-pulses of bandwidth W

Capacity (continued)

Notes
- x-axis: energy per information bit
- y-axis: capacity in bits per symbol

- energy efficiency: slope of capacity
- two regimes:
1) energy efficient
2) high-rate

- parallel channels (multi-mode) let one increase energy efficiency and rate



2) Fiber Communication

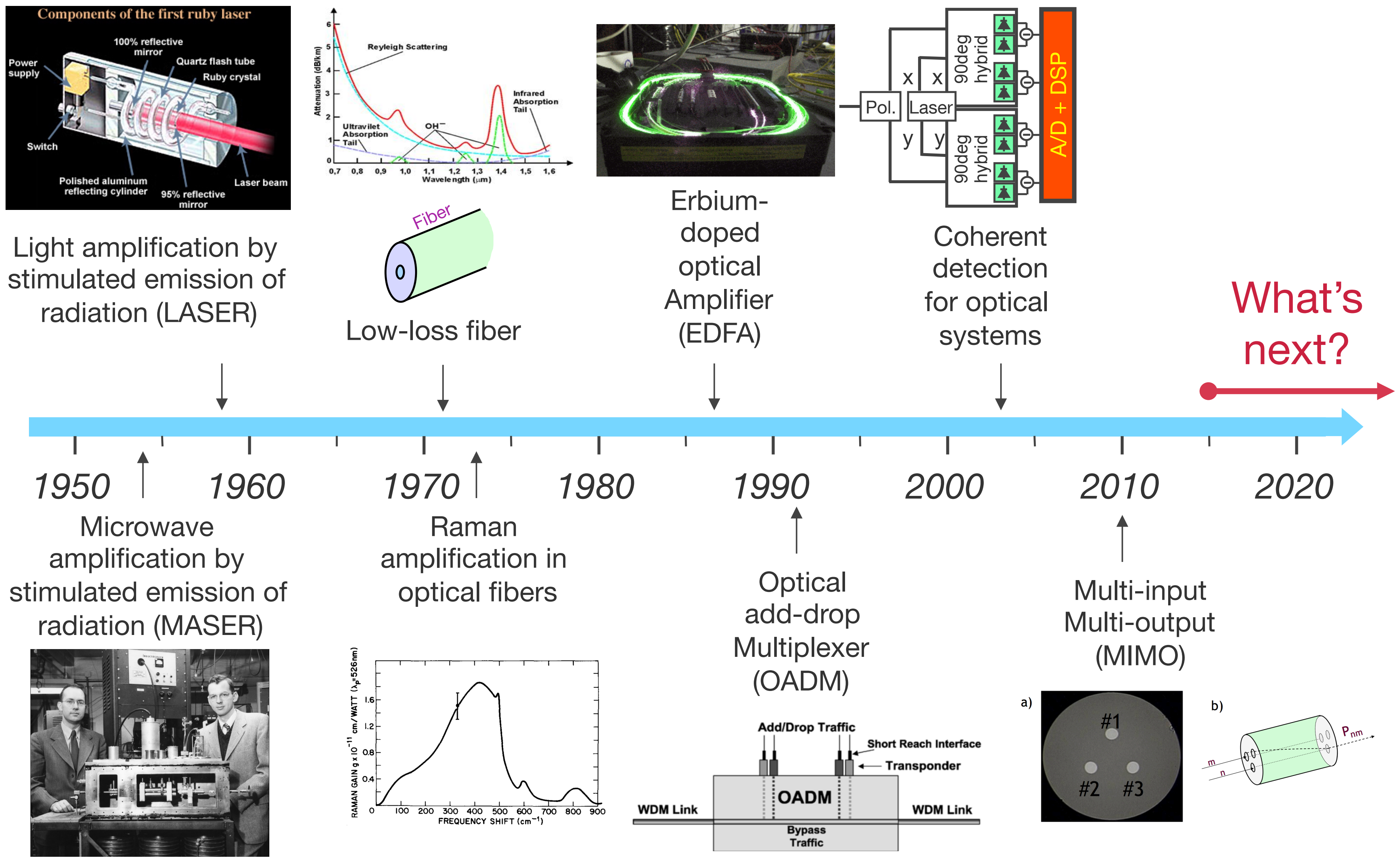
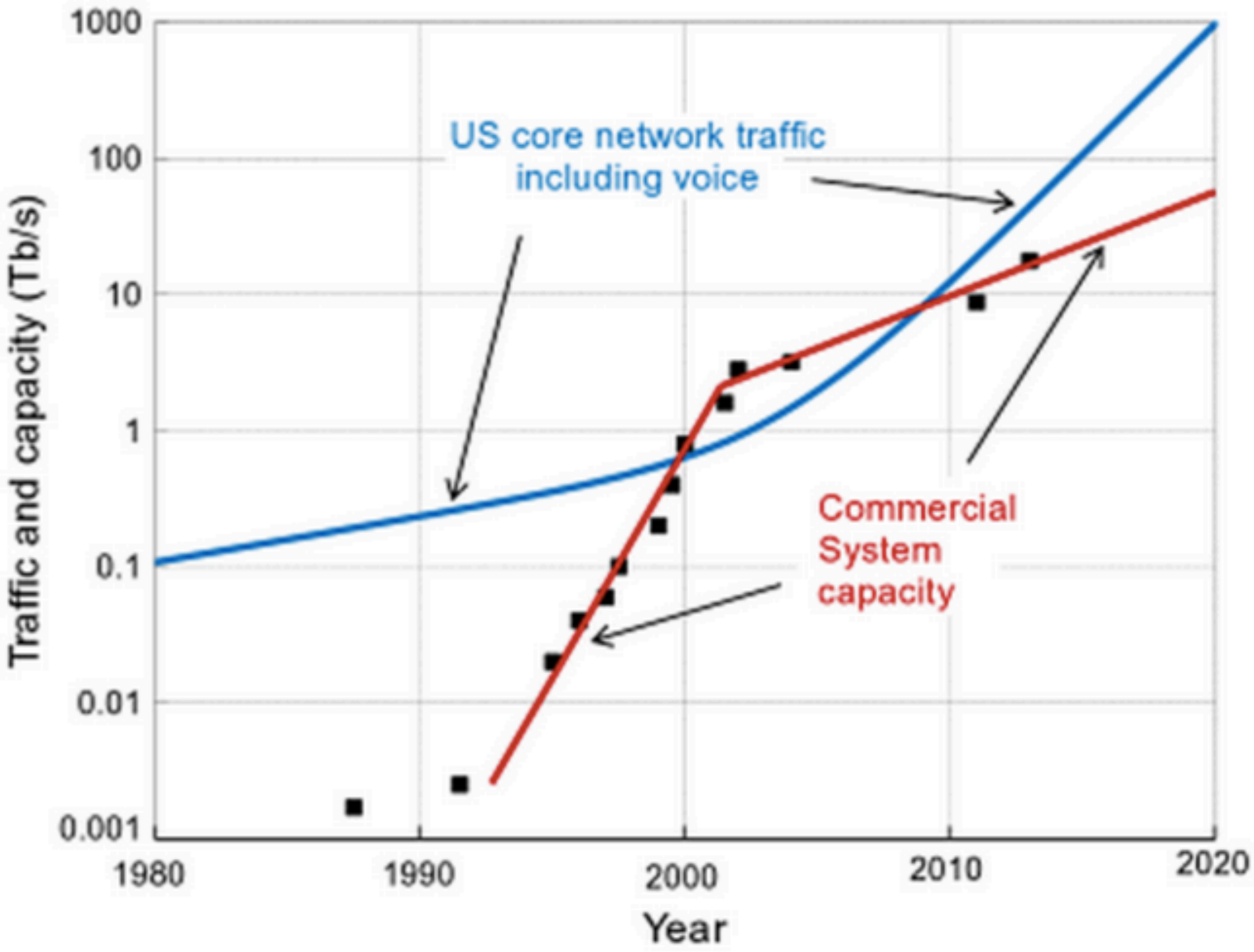


Figure courtesy of R.-J. Essiambre

Demand and Supply

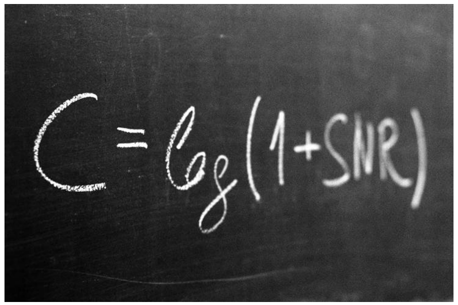


Data traffic is growing at a faster pace than fiber capacity

Figure courtesy of R.-J. Essiambre

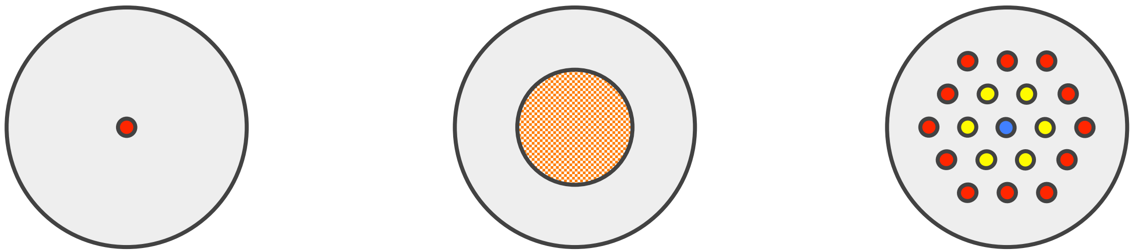
Questions

- There is much **dark** (unused) fiber, so what's the problem? \$\$
- Cost increases in proportion to the data rate.
  Example: Want 100 Mbps rather current 10 Mbps?
  Pay 10x the money.
- This may prevent further innovation
- So what should we do?
  Improve the networks (perhaps the first step)
  and (ultimately) the fiber channel capacity



3) Fiber Channel(s)

- Single-Mode Fiber (SMF): a small core that carries one mode of light
- Here one mode has 2 complex dimensions: two polarizations
- Theory papers often consider one complex dimension;
  the general case is interesting too of course (see below)
- In fact, a **hot** topic in the fiber community is MIMO fiber



Single-mode fiber (SMF) | Multi-mode fiber (MMF) | Multi-core fiber (MCF)

SMF Pulse Propagation Equation

- Maxwell's equations and low-order approximations* result in a generalized nonlinear Schrödinger equation (GNSE):

Linear | Nonlinear

$$\frac{\partial E}{\partial z} = -\frac{\alpha}{2}E - \frac{i}{2}\beta_2\frac{\partial^2 E}{\partial T^2} + \frac{1}{6}\beta_3\frac{\partial^3 E}{\partial T^3} + i\gamma\,|E|^2 E + n$$

Distance Evolution — Fiber Loss/Gain — Dispersion — Dispersion Slope — Kerr Nonlinearity — Noise (Gaussian, Bandlimited)

- E : Electromagnetic field, function of z and T
- z: Distance
- T : Retarded time $t-\beta_1 z$
- α : Fiber loss coefficient (~ 3 dB/15 km)
- β_1 : Inverse of group velocity
- β_2 : Fiber group velocity dispersion
- β_3 : Fiber dispersion slope (include if β_2 small)
- γ : Fiber nonlinear parameter $(n_2\ \omega)/(c\ A_{eff})$

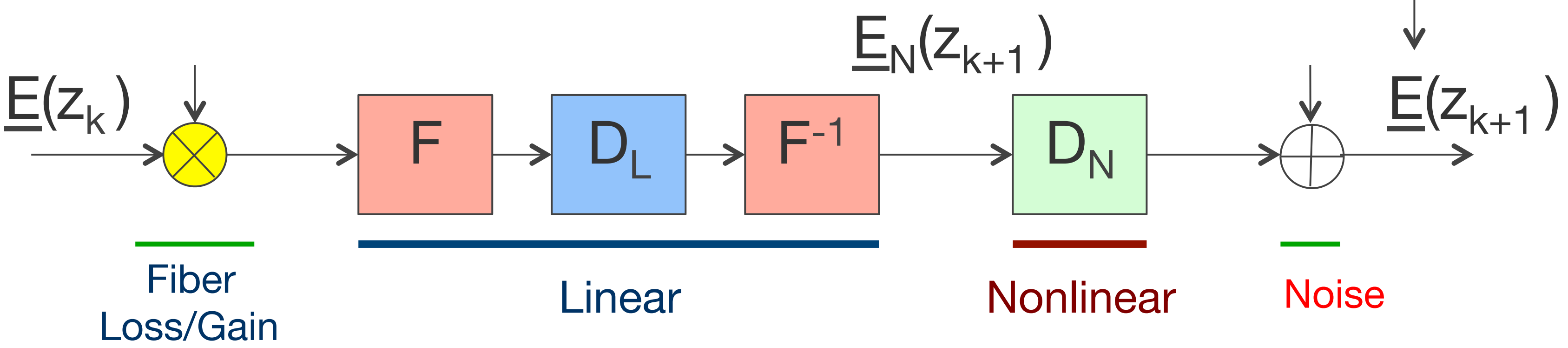
- n_2 : Fiber nonlinear coefficient
- ω : Angular frequency
- c : Speed of light
- A_{eff} : Fiber effective area

Figure courtesy of R.-J. Essiambre

*See Ch. 2 in G.P. Agrawal, "Nonlinear Fiber Optics", 3rd ed., 2001

- To simulate, split the fiber length z^* into K small steps (Δz) and the time T into **L** small steps (Δt)
- Split-step Fourier method at distance z_k, $k=0,1,\ldots,K$

time signal:
vector with **L** entries



- Ideal Raman amplification: removes the loss and adds noise
- F = Fourier transform
- D_L = diagonal matrix with fixed entries of unit amplitude (all-pass filter)
- D_N = diagonal matrix with unit amplitude entries; the (ℓ,ℓ)-entry phase shift is proportional to the magnitude-squared of the ℓ^{th} entry of $\underline{E}_N(z_{k+1})$

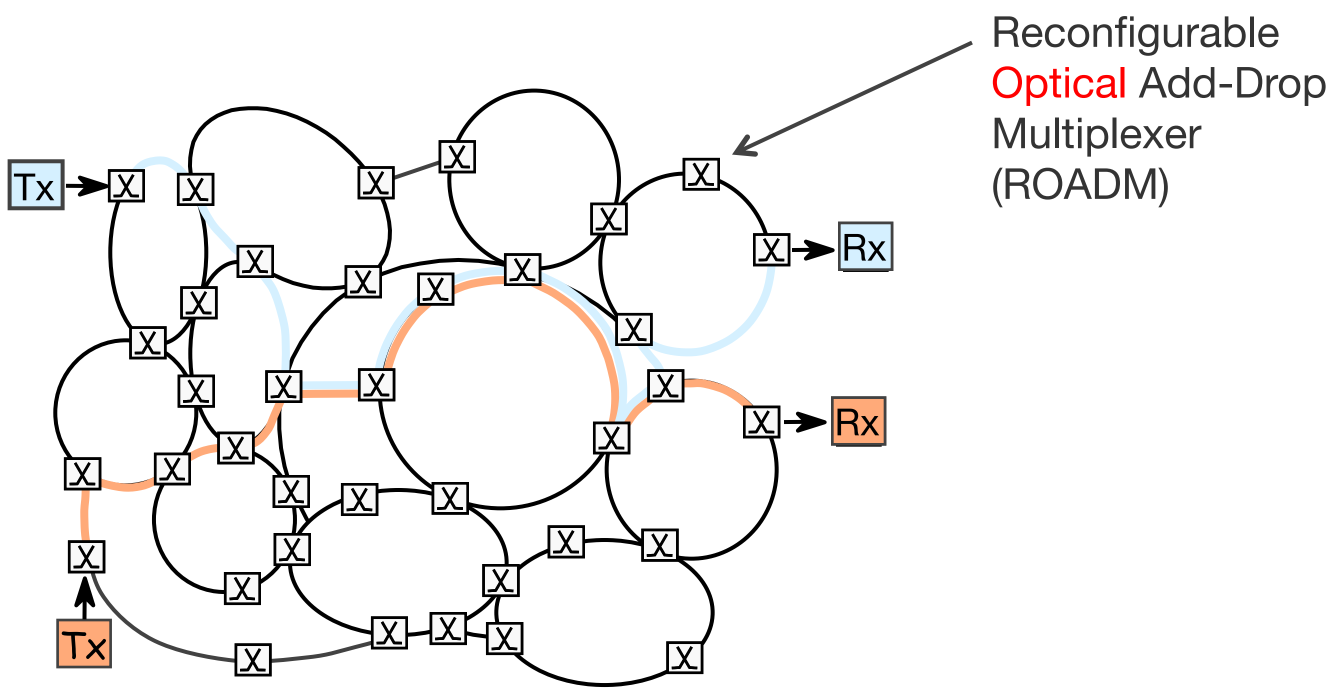
4) Lower Bounds on Spectral Efficiency

- No analytic lower bounds exist for “full” models. All lower bounds are based on simulations or approximate models
- What quantity should we study?
  1) Capacity Problem: ∞ bandwidth so ∞ capacity!?
  2) Spectral Efficiency, i.e., capacity per Hertz
     Critique: why is Fourier bandwidth* (Hz) a good currency?
     Shouldn’t we use Shannon bandwidth* (# dimensions)?
  3) Capacity of realistic fiber
     i.e., use realistic loss/dispersion/nonlinearity vs. frequency
     Problem: seems very difficult to analyze
- We study spectral efficiency for ideal Raman amplification, but the capacity with realistic loss functions is ultimately most interesting

* Terminology borrowed from J.L. Massey (1995)

Fiber Network Model

- Of course, capacity depends strongly on the model under study
- Optically-routed fiber-optic network model:



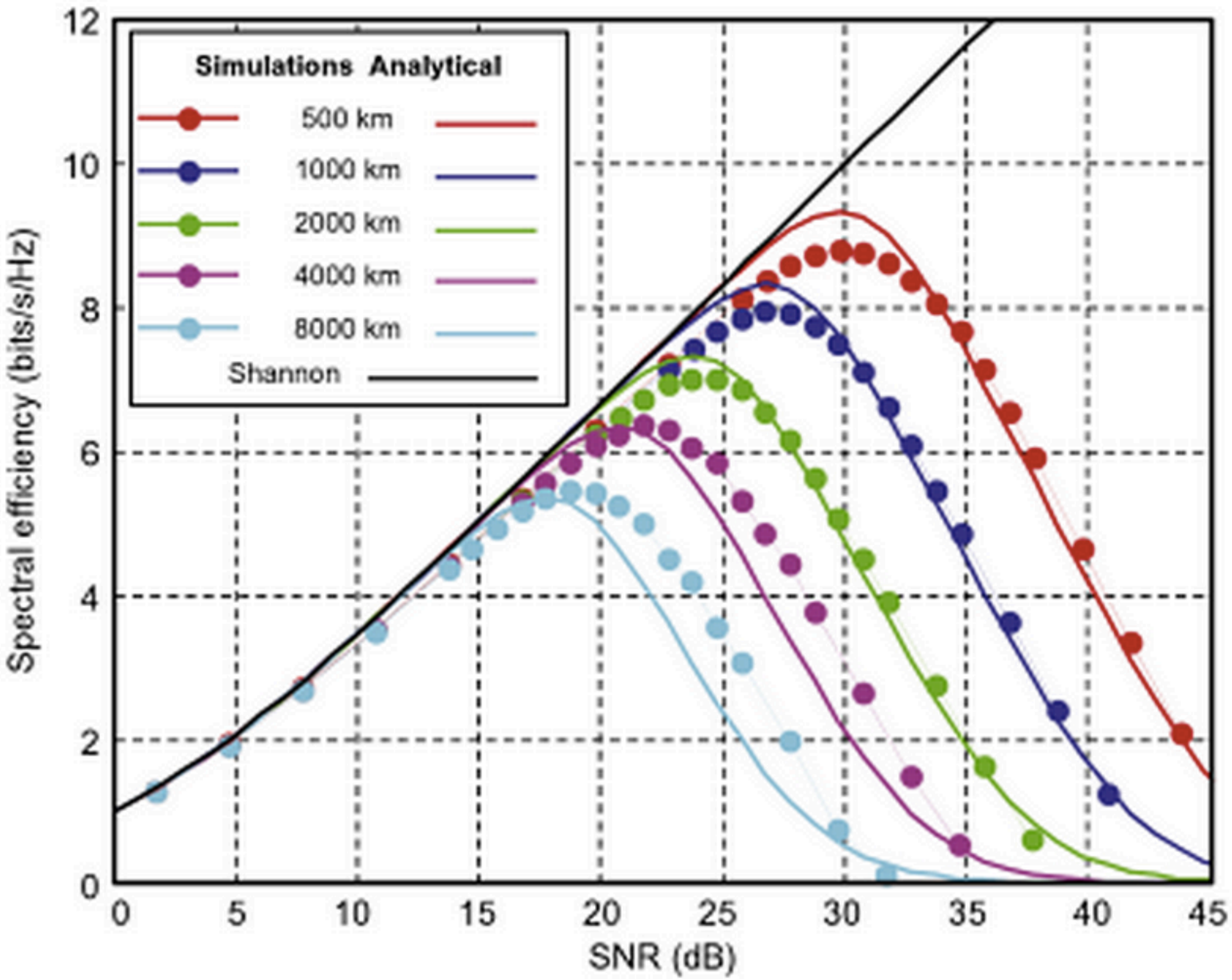
- WDM signals interfere due to fiber nonlinearities
- Signals co-propagate in a network environment

Figure courtesy of R.-J. Essiambre

Computed Lower Bounds

Notes

- curves for a WDM network with filters and per-channel receivers
- η may decrease with launch power
- low SNR: channel is almost linear
- new lower bounds by several groups (2013-)



R.-J. Essiambre, et al., "Capacity limits of optical fiber networks," IEEE/OSA J. Lightwave Technology, Feb. 2010.

5) Upper Bound

But First More IT Preliminaries

- Consider a complex column vector $\underline{X} = \underline{X}_c + j\,\underline{X}_s$ with covariance and pseudo-covariance matrices

$$\mathbf{Q}_{\underline{X}} = E\left[\left(\underline{X} - E[\underline{X}]\right)\left(\underline{X} - E[\underline{X}]\right)^{\dagger}\right]$$

$$\tilde{\mathbf{Q}}_{\underline{X}} = E\left[\left(\underline{X} - E[\underline{X}]\right)\left(\underline{X} - E[\underline{X}]\right)^{T}\right]$$

- For interest: $\underline{X}$ is called proper if its pseudo-covariance matrix is **0**

- Example: Consider a complex, zero-mean, scalar $X = X_c + j\,X_s$.

  X is proper if $E[X_c^2]=E[X_s^2]$ and $E[X_cX_s]=0$.

  Note: circularly symmetric X are proper, but proper X are not

  necessarily circularly symmetric (e.g. QAM signal sets)

Maximum Entropy

- Maximum Entropy: consider the correlation matrix $\mathbf{R}_{\underline{X}}=E[\underline{X}\,\underline{X}^{\dagger}]$ where $\underline{X}$ has L entries. Then

$$h(\underline{X}) \le \log\left[(\pi e)^{L} \det \mathbf{R}_{\underline{X}}\right]$$

with equality if and only if $\underline{X}$ is Gaussian and proper (or circularly symmetric)

- For a complex square matrix **M** we have

$$h(\mathbf{M}\,\underline{X}) = h(\underline{X}) + 2\log\left|\det(\mathbf{M})\right|$$

In particular, if **M** is unitary then $h(\mathbf{M}\,\underline{X}) = h(\underline{X})$

Entropy Power Inequality

- Entropy Power:

$$V(\underline{X}) = e^{h(\underline{X})/L} / (\pi e)$$

- Entropy Power Inequality: for independent $\underline{X}$ and $\underline{Y}$ we have

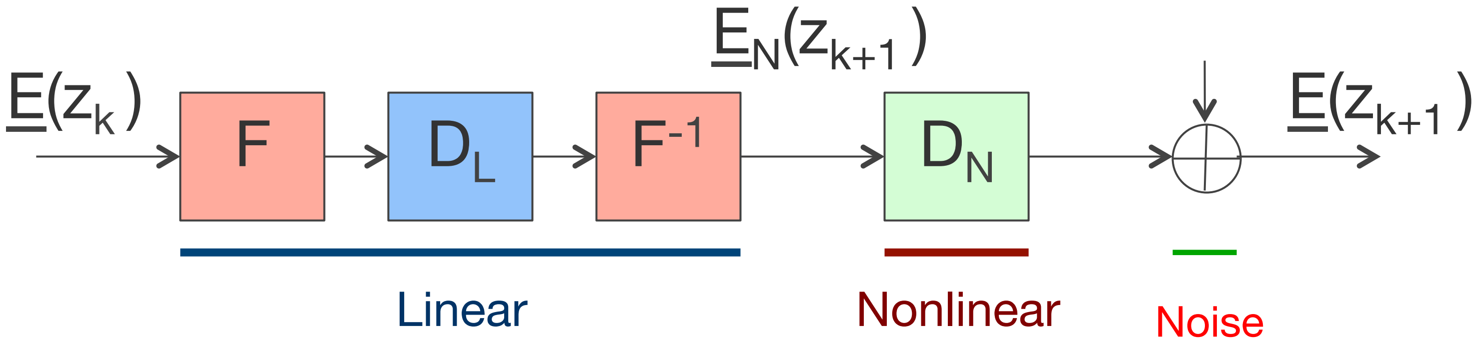
$$\boxed{V(\underline{X}+\underline{Y}) \geq V(\underline{X}) + V(\underline{Y})}$$

- Conditional version: for conditionally independent $\underline{X}$ and $\underline{Y}$ we have

$$V(\underline{X}|\underline{U}) = e^{h(\underline{X}|\underline{U})/L} / (\pi e)$$

$$V(\underline{X}+\underline{Y}|\underline{U}) \geq V(\underline{X}|\underline{U}) + V(\underline{Y}|\underline{U})$$

Energy and Entropy Conservation



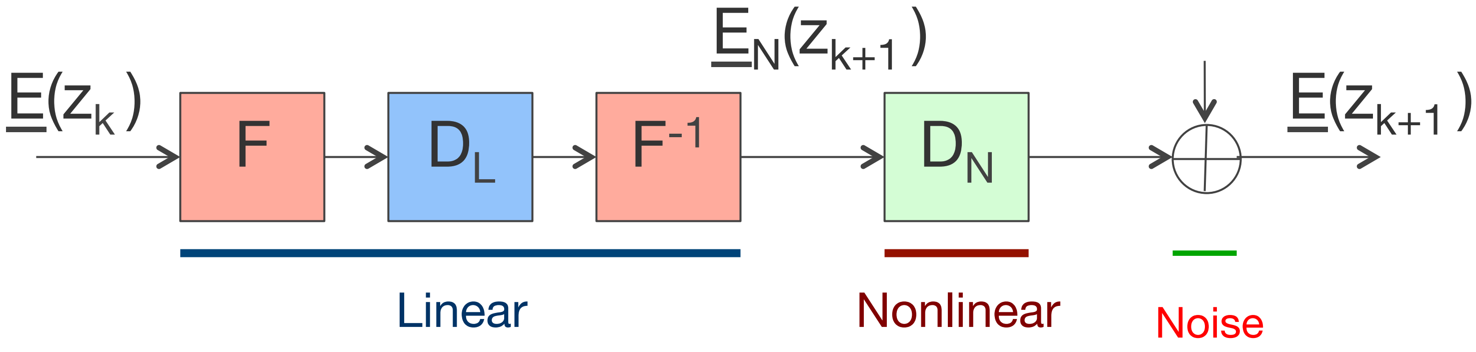
Main Observations

- The linear step conserves energy and entropy
- The non-linear step also conserves energy and entropy **(the key result)**

$$h\left(|a|e^{j\arg(a)+jf(|a|)}\right) = h\left(|a|,\arg(a)+f(|a|)\right) + \mathrm{E}\left[\log|a|\right]$$

$$= \underbrace{h\left(|a|\right) + h\left(\arg(a)+f(|a|)\,\big|\,|a|\right)}_{h\left(|a|,\arg(a)\right)} + \mathrm{E}\left[\log|a|\right] = h(a)$$

Energy Recursion



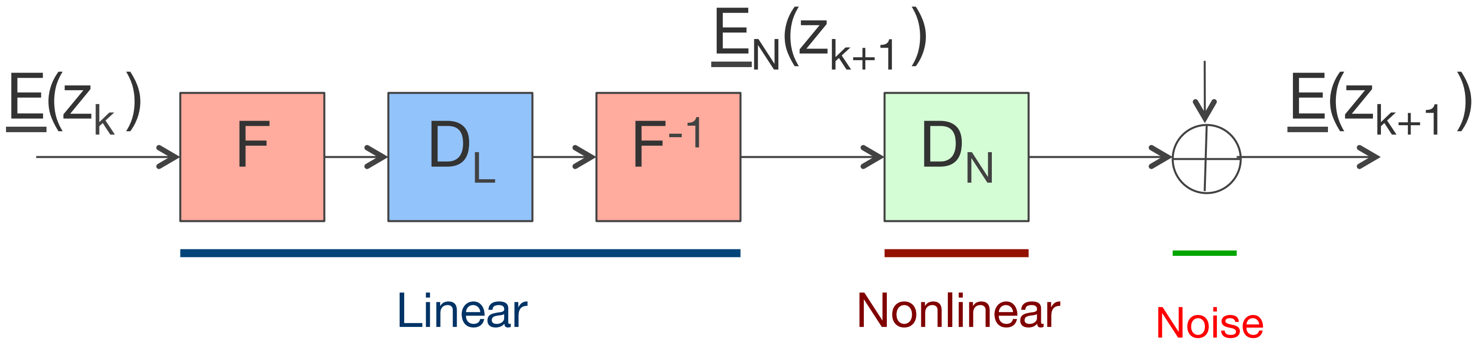
- Energy after K steps: $\text{Energy}_{\text{Launch}} + KN$. We thus have:

$$h(\underline{E}(z_K)) \le \log\left[(\pi e)^L \det\left(\mathbf{R}(\underline{E}(z_K))\right)\right] \quad \ldots \text{ maximum entropy}$$

$$\le \sum_{i=1}^{L} \log\left[\pi e\, R_{i,i}(\underline{E}(z_K))\right] \quad \ldots \text{ Hadamard's inequality}$$

$$\le L \cdot \log\left[\pi e\left(Energy_{\text{Launch}} + KN\right)/L\right] \quad \ldots \text{ Jensen's inequality}$$

Entropy Recursion



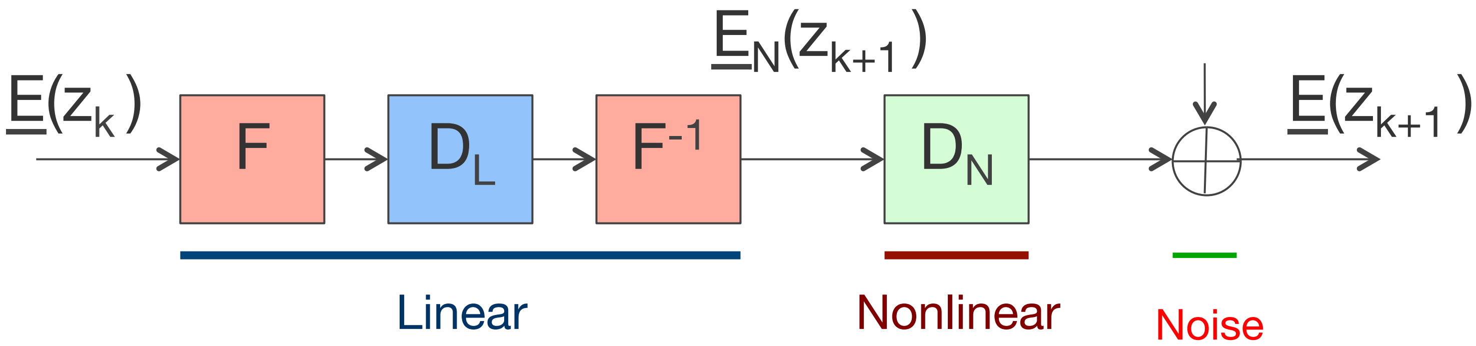
- Entropy recursion:

$$V\left(\underline{E}(z_{k+1})\middle|\underline{E}(z_0)\right) \geq V\left(\underline{E}(z_k)\middle|\underline{E}(z_0)\right) + N/L$$

- We thus have:

$$V\left(\underline{E}(z_K)\middle|\underline{E}(z_0)\right) \geq KN/L$$

$$\text{or} \quad h\left(\underline{E}(z_K)\middle|\underline{E}(z_0)\right) \geq L\log(\pi e\, KN/L)$$



So for every step we have:

- Signal energy grows by the noise variance: can upper bound h($\underline{E}(z_K)$)
- Entropy power grows by at least the noise variance: can lower bound h($\underline{E}(z_K) \mid \underline{E}(z_0)$)
- Result*:

$$I\left(\underline{E}(z_0);\underline{E}(z_K)\right) = h\left(\underline{E}(z_K)\right) - h\left(\underline{E}(z_K)\middle|\underline{E}(z_0)\right)$$

$$\leq L \cdot \log\left(1+SNR\right)$$

*SNR = receiver signal-to-noise ratio

$$\Rightarrow \quad \frac{1}{L} I\left(\underline{E}(z_0);\underline{E}(z_K)\right) \leq \log\left(1+SNR\right)$$

- Let B = 1/Δt be the "bandwidth" of the simulation
- So L = T/Δt = TB is the time-bandwidth product
- The spectral efficiency is thus bounded by

$$\eta \leq \log\left(1+SNR\right) \quad \text{[bits/sec/Hz]}$$

6) Discussion

$$\eta \leq \log(1 + SNR) \quad \text{[bits/sec/Hz]}$$

Q1: Why normalize by the simulation bandwidth B?
The “real” bandwidth W can be smaller.

A1: B can be chosen (this is even desirable) as the smallest bandwidth for which simulations give accurate results

Q2: What about capacity?

A2: Any real fiber has a maximal bandwidth B_{max}.
A capacity upper bound follows by multiplying B_{max} by log(1+SNR)

Discussion (continued)

$$\eta \leq \log\left(1 + SNR\right) \quad \text{[bits/sec/Hz]}$$

Q3: What about MIMO fiber?
A3: If energy and entropy are preserved by the linear and non-linear steps, and the noise is AWGN then the bound remains valid per mode

Q4: What about frequency-dependent (or mode-dependent) loss?

A4: Open research!

7) Conclusions

1) Spectral efficiency of (an idealized model of) SMF with linear polarization is ≤ log(1+SNR)
2) Many extensions are possible:
   - lumped amplification, 3^{rd}-order dispersion, delayed Kerr effect
   - uniform loss, linear filters (for capacity results)
   - MIMO fiber (MMF or MCF)
3) More difficult:
   - better bounds and understanding at high SNR
   - frequency-dependent loss, dispersion, non-linearity
4) Multi-user information theory for fiber should be developed