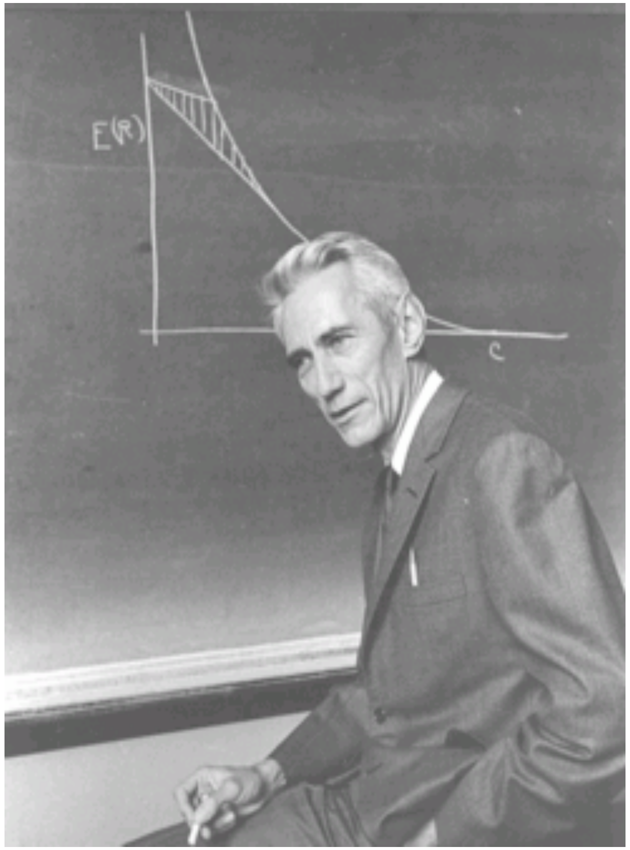
Models and Capacity Bounds for Optical Fiber Channels

Gerhard Kramer
contributions by H. Ghozlan, L. Barletta,
R.-J. Essiambre, M. Yousefi, F. Kschischang

IEEE Information Theory Society Distinguished Lecture
Imperial College, University of York, University of Cambridge
December 14-16, 2016

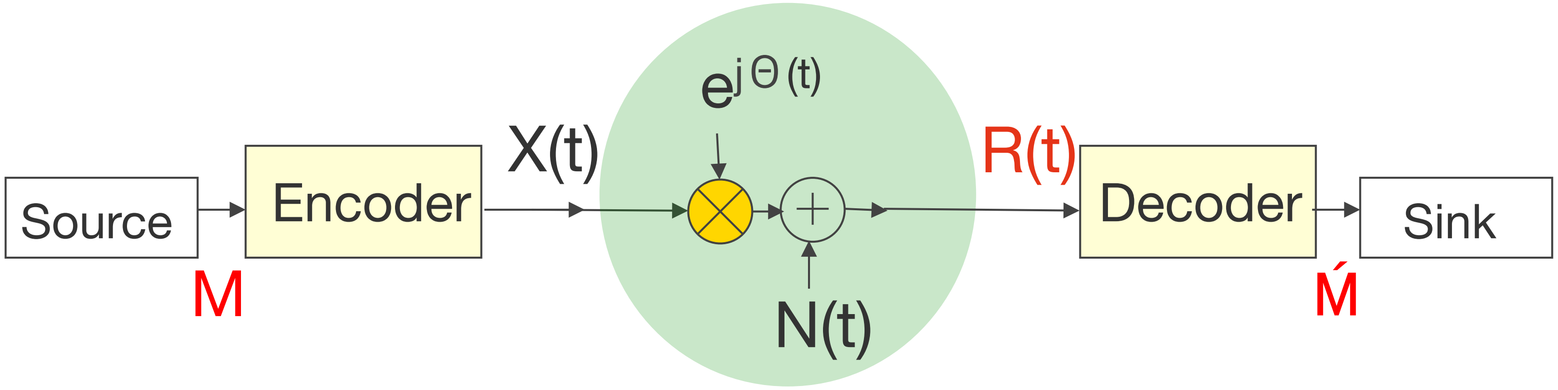
Unterstützt von / Supported by
Alexander von Humboldt
Stiftung/Foundation



Claude Elwood Shannon

Apr 30, 1916 – Feb 24, 2001

1) Phase Noise Models


Source

M

Encoder

$X(t)$

$e^{j\Theta(t)}$

$N(t)$

$R(t)$

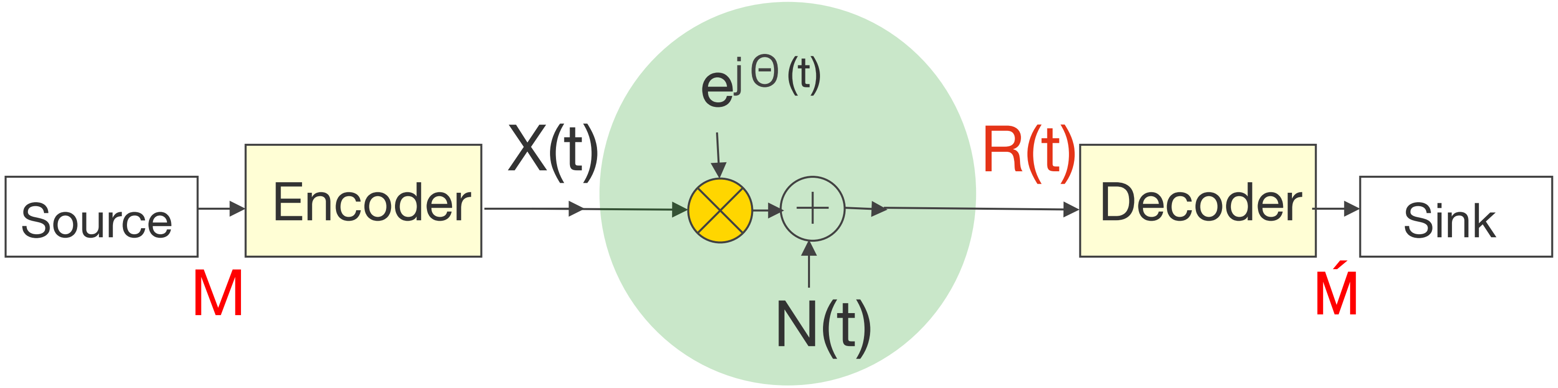
Decoder

$\acute{M}$

Sink

- Phase noise due to oscillator instability and channel non-linearities
- Phase noise statistics:
  - phase-locked loops (PLLs) residual noise: von Mises/Tikhonov distribution
  - satellite (DVB-S2): white Gaussian process filtered by IIR filters
  - fiber-optic lasers: Wiener process
  - Raman amplification: large bandwidth Gaussian process

White Phase Noise



- Simplified model (Barletta-Kramer, 2014)

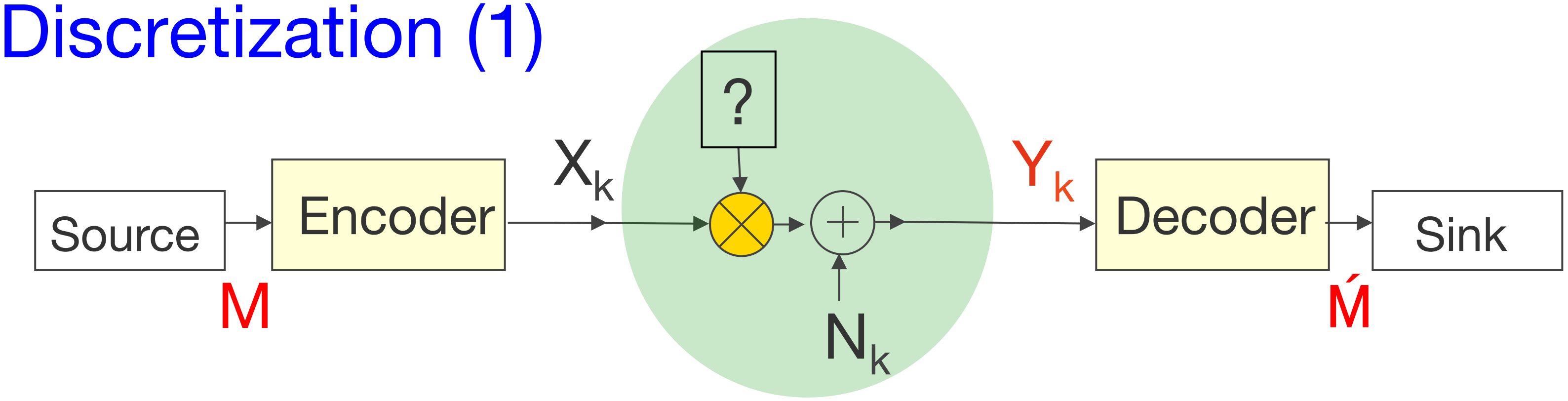
$$R(t) = X(t) \cdot e^{j\Theta(t)} + N(t)$$

$\Theta(t)$ is white* and $N(t)$ is white Gaussian* (both are idealizations)

- Motivation: phase noise bandwidth much larger than receiver bandwidth
- Mathematically: let $\{\phi_m(t)\}$ be an orthonormal basis of $L^2[0,T]$ and project $X(t)$, $N(t)$, and $R(t)$ onto the $\phi_m(t)$

* We use $E[\Theta(t)\Theta(t+\tau)] = \sigma^2\delta_\tau$ and $E[N(t)N(t+\tau)] = \sigma^2\delta(\tau)$

Discretization (1)



- X(t) and N(t):

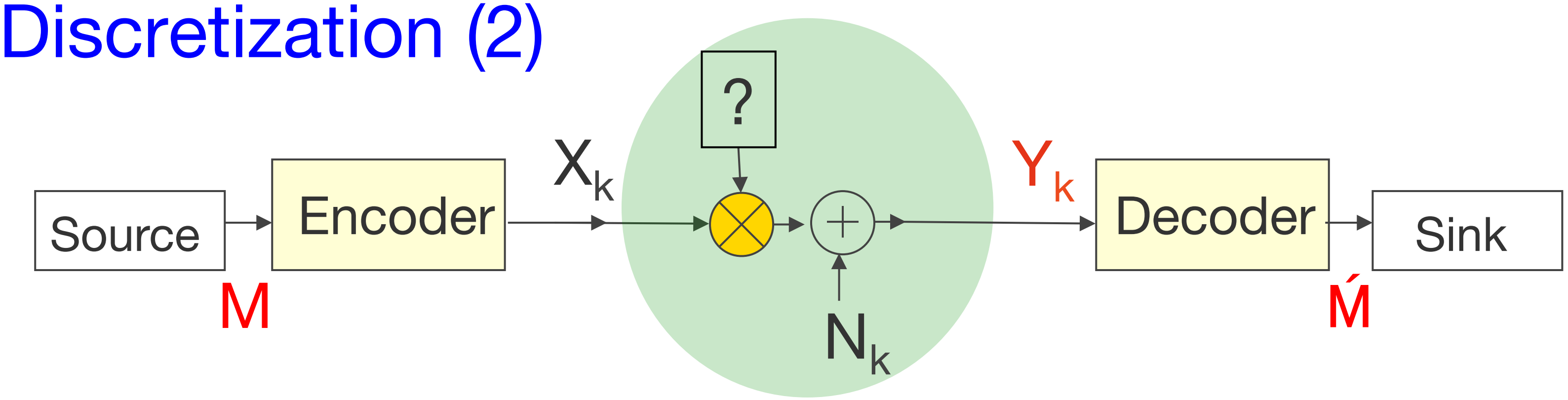
$$X(t)=\sum_{m=1}^{M}X_m\phi_m(t) \qquad N(t)=\sum_{m=1}^{\infty}N_m\phi_m(t)$$

- Receiver:

$$Y_k=\left\langle X(t)e^{j\Theta(t)}+N(t),\phi_k(t)\right\rangle$$

$$=\sum_{m=1}^{M}X_m\underbrace{\left\langle\phi_m(t)e^{j\Theta(t)},\phi_k(t)\right\rangle}_{\Phi_{m,k}}+N_k$$

Discretization (2)



- Samples:

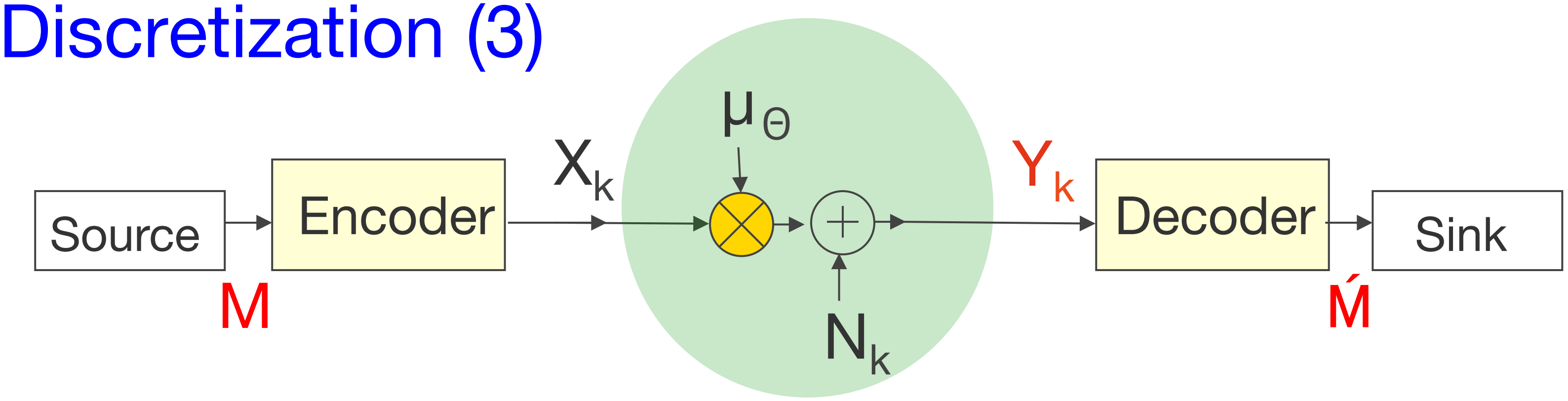
$$\Phi_{m,k} = \int_0^T \phi_m(t)\, e^{j\Theta(t)}\, \phi_k(t)^*\, dt$$

$$= \lim_{L\to\infty} \frac{1}{L} \sum_{i=1}^{L} \phi_m\left(t_i^{(L)}\right) e^{j\Theta\left(t_i^{(L)}\right)} \phi_k\left(t_i^{(L)}\right)^*, \quad t_i^{(L)} = \frac{(i-1)T}{L}$$

$$\overset{!}{=} \begin{cases} E\left[e^{j\Theta(t)}\right], & m = k \\ 0, & \text{else} \end{cases}$$

Barletta-Kramer, 2014:
Almost sure convergence for white phase noise with uncorrelated samples of process $\{e^{j\Theta(t)}\}$

Discretization (3)



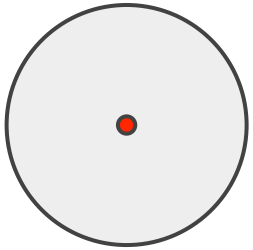
- Model*: $Y_k = X_k \cdot \underbrace{E\left[e^{j\Theta(t)}\right]}_{\mu_\Theta} + N_k$

  An AWGN channel with no phase noise and SNR penalty $|\mu_\Theta|^2$
- Penalty called spectral loss**: energy is not conserved.
  An interpretation: the “lost” power is spread across all frequencies.

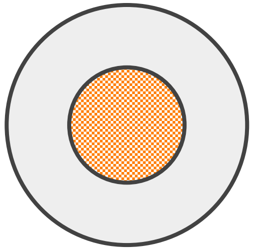
* Barletta-Kramer 2014, ** Goebel et al. 2011

2) Optical Fiber Channels

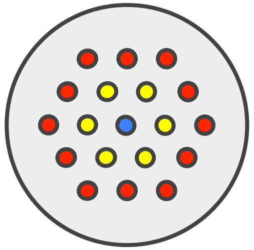
- Single-Mode Fiber (SMF): a small core that carries one mode of light
- Here one mode has 2 complex dimensions: two polarizations
- Theory papers often consider one polarization only; we do too
- A long-term research topic is multimode/multicore/MIMO fiber



Single-mode
fiber (SMF)



Multi-mode
fiber (MMF)



Multi-core
fiber (MCF)

Pulse Propagation Equation

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

$$\underbrace{\frac{\partial E}{\partial z}}_{\text{Distance Evolution}} = \overbrace{-\underbrace{\frac{\alpha}{2}E}_{\text{Fiber Loss/Gain}} - \underbrace{\frac{i}{2}\beta_2\frac{\partial^2 E}{\partial T^2}}_{\text{Dispersion}} + \underbrace{\frac{1}{6}\beta_3\frac{\partial^3 E}{\partial T^3}}_{\text{Dispersion Slope}}}^{\text{Linear}} + \overbrace{\underbrace{i\gamma\,|E|^2 E}_{\text{Kerr Nonlinearity}}}^{\text{Nonlinear}} + \underbrace{n}_{\text{Noise (Gaussian, Bandlimited)}}$$

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 $(n_2\ \omega)/(c\ A_{eff})$

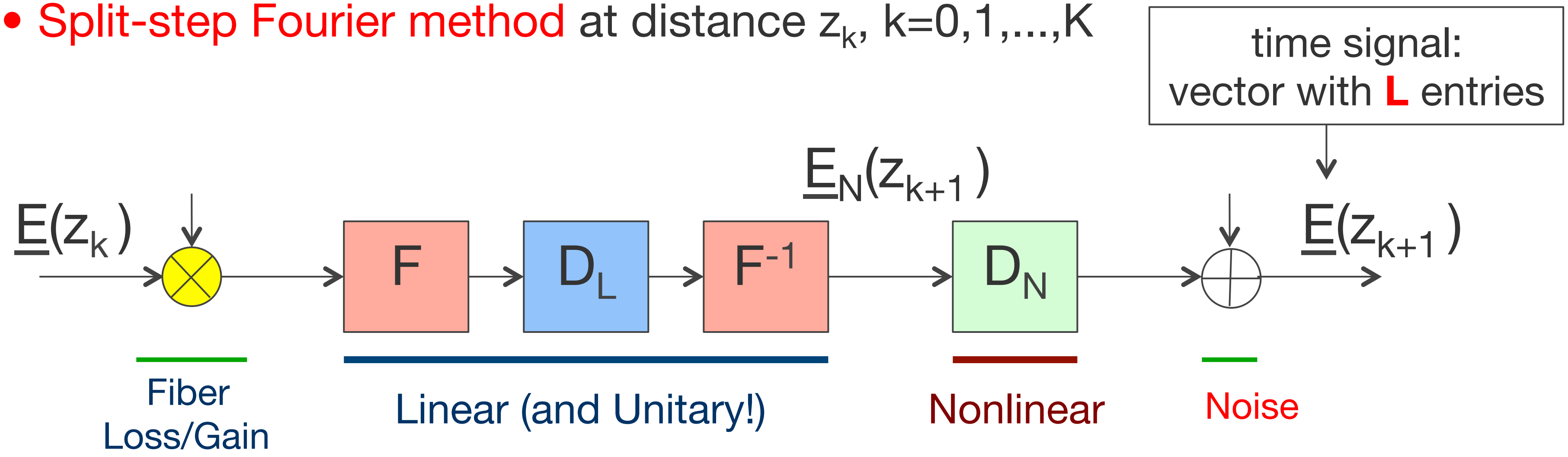
n_2 : Fiber nonlinear coefficient
ω : Angular frequency
c : Speed of light
A_{eff} : Fiber effective area

Figure courtesy of R.-J. Essiambre

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

Calculation/Simulation

- Split fiber into K pieces (Δz meters) and time into **L** steps (Δt seconds)
- Split-step Fourier method at distance z_k, k=0,1,...,K



- Ideal Raman amplification removes the loss and adds distributed 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 $\underline{E}_N(z_{k+1})$

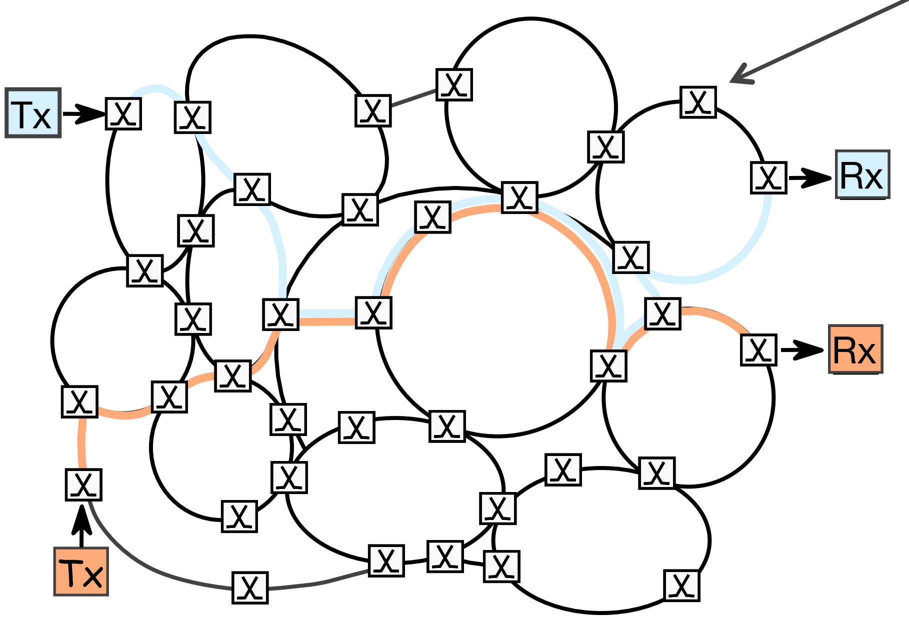
3) Fiber Capacity: Lower Bounds

- Existing bounds are based on simulations or approximate models
- What quantity should we study?
  1) Capacity of ∞ bandwidth model: ∞ 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: frequency dependent loss/dispersion/nonlinearity/amplification/filtering. Problem seems difficult
- We study spectral efficiency for ideal distributed amplification,
  but capacity with realistic fiber 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:



Reconfigurable Optical Add-Drop Multiplexer (ROADM)

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

Figure courtesy of R.-J. Essiambre

Lower Bounds for 2000 km

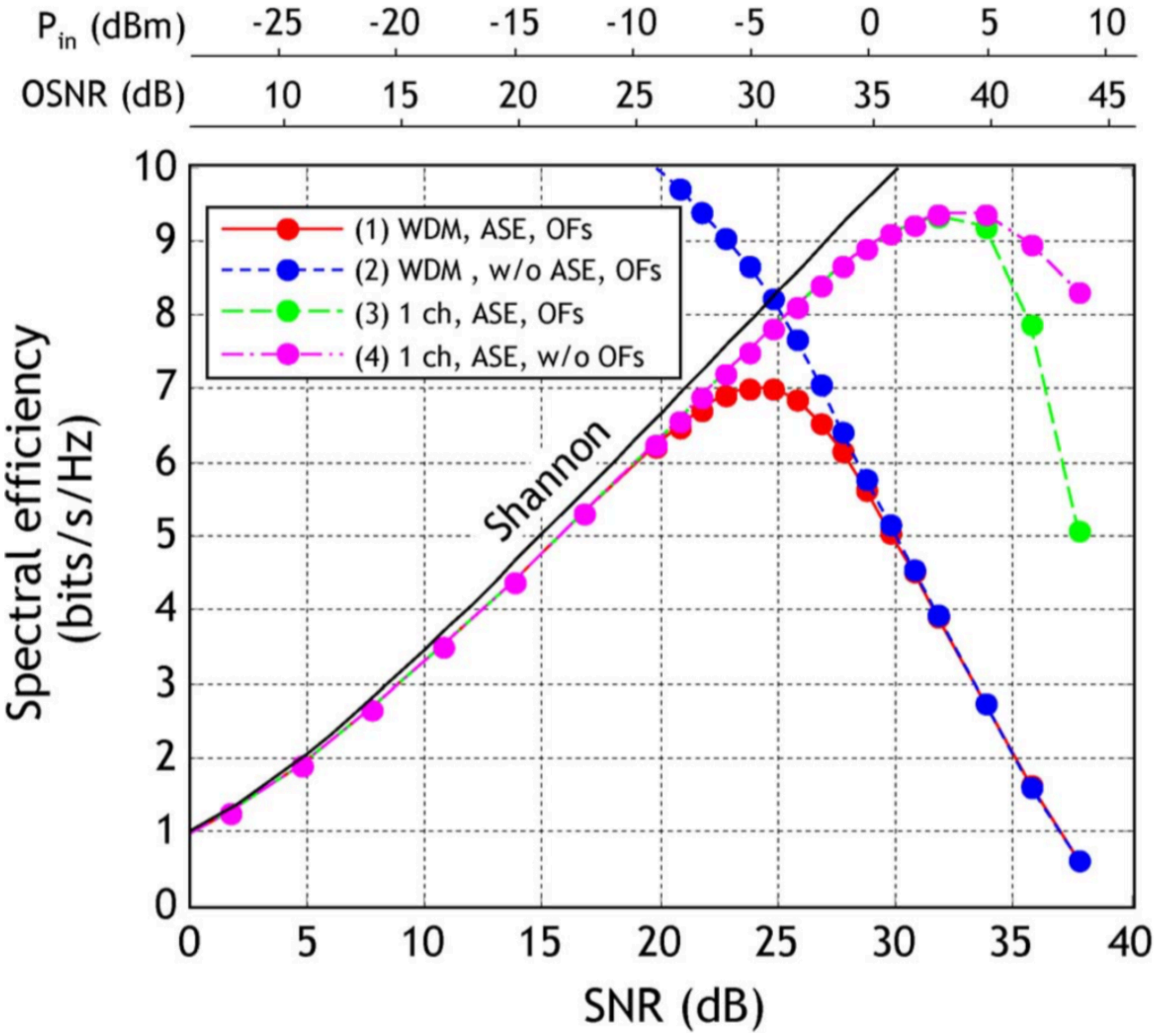
Notes

Curves:

- WDM network with *filters* and *per-channel* receivers (η may decrease with launch power)
- WDM w/o ASE
- 1 channel
- 1 channel w/o OFs

Interpretation:

- WDM: signal-signal interference limits rates; suppressing interference could increase C a little
- 1 channel: OFs limit more than signal-noise effects



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

4) Spectral Efficiency Upper Bound

An analytic upper bound: the spectral efficiency η of the GNSE satisfies

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

Features:

- the only existing analytic bound on η for this model
- the bound is tight at “low” SNR and good at “medium” SNR

Limitations:

- the bound is likely extremely loose at high SNR
- implicitly requires signals to be bandlimited with bandwidth at most the noise bandwidth*

---

* more on this later

Derivation: First Some Fun Facts

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

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

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

- Definition: $\underline{X}$ is circularly symmetric if $\underline{X}e^{j\Theta}$ has the same probability distribution as $\underline{X}$ for all real Θ
- Note: circular symmetry implies pseudo-covariance matrix is **0**

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

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

  with equality if and only if $\underline{X}$ is Gaussian and 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:

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

- Entropy Power Inequality: for independent X̲ and Y̲ we have

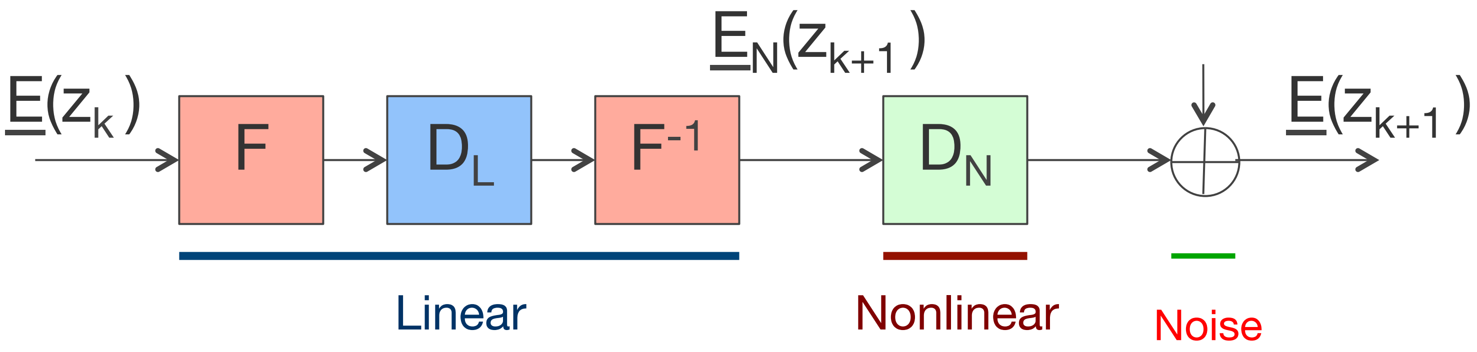
$$\boxed{V\left(\underline{X}+\underline{Y}\right)\geq V\left(\underline{X}\right)+V\left(\underline{Y}\right)}$$

- Conditional version: for conditionally independent X̲ and Y̲ we have

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

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

Energy and Entropy Conservation



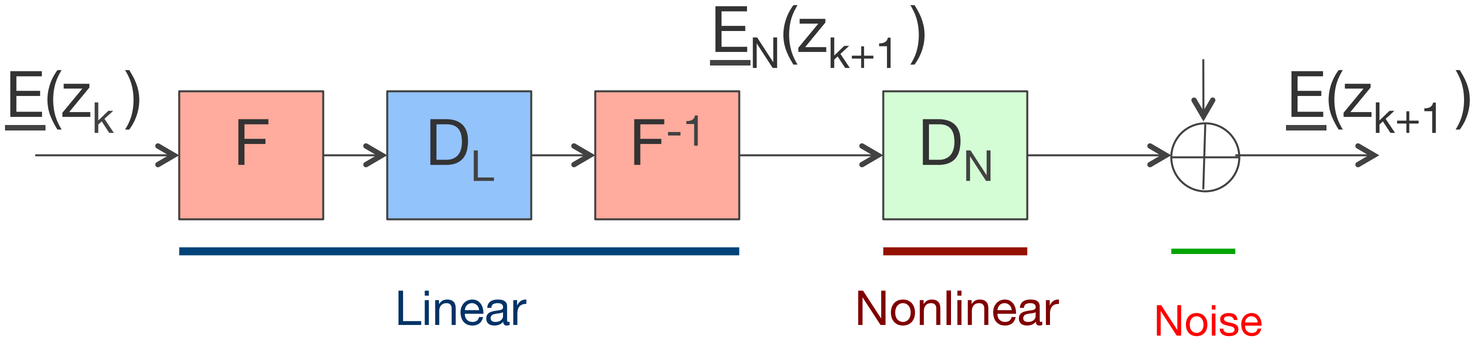
Main Observations

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

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

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

Energy Recursion



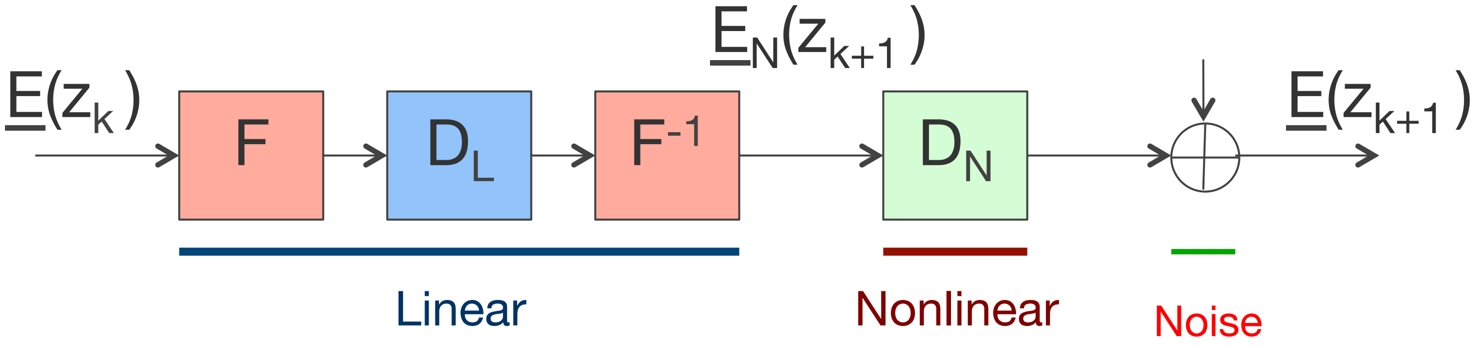
- Energy after K steps: $\text{Energy}_{\text{Launch}} + KN$. We thus have:

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

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

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

Entropy Recursion



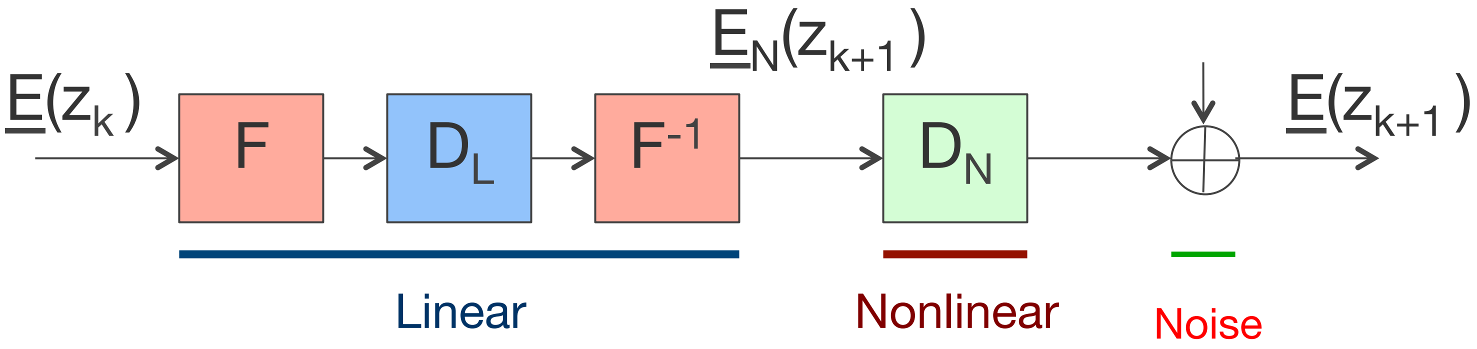
- Entropy recursion:

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

- 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

$$\Rightarrow \quad \frac{1}{L} I\left(\underline{E}(z_0);\underline{E}(z_K)\right) \le \log\left(1+SNR\right) \quad \text{[bits/entry]}$$

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

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

5) 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 is then B_{max} log(1+SNR).

Discussion (continued)

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

Q3: What about MIMO (multi-mode, multi-core) 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 freq.-dependent (or mode-dependent) noise and loss?

A4: Good question because

- the entropy power inequality blows up if there is no noise in some dimension (such as a given time or band)
- freq.-dependent loss complicates analysis (but can still get bounds)

6) Conclusions

1) Spectral efficiency of (an idealized model of) SMF with linear polarization is ≤ log(1+SNR)
2) Many extensions possible:
   - lumped amplification, 3rd-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 noise, loss, dispersion, non-linearity