

# Digital Communication over Optical Fibers

**Frank R. Kschischang**

*Department of Electrical & Computer Engineering  
University of Toronto*

[frank@ece.utoronto.ca](mailto:frank@ece.utoronto.ca)

**2016 European School of Information Theory**

**Chalmers University of Technology**

**April 4th, 2016**

# Thank You To:

My collaborators:

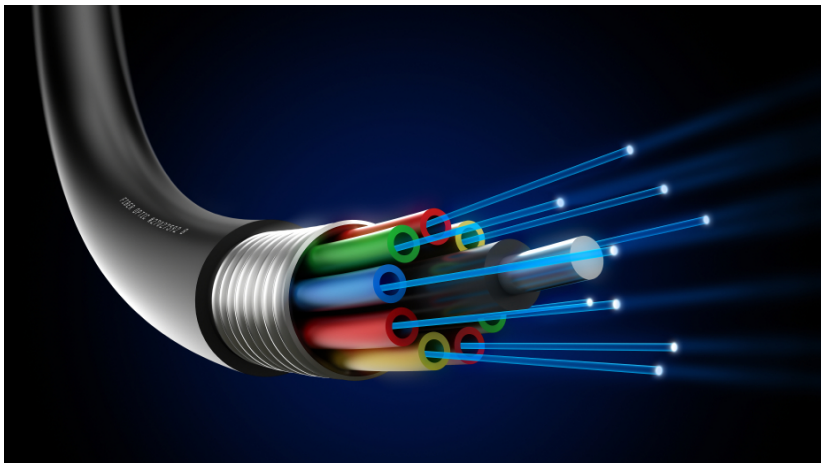
- Mansoor Yousefi (University of Toronto, Technical University of Munich, ParisTech)
- Gerhard Kramer (Technical University of Munich)
- Siddarth Hari (University of Toronto)
- Lei Zhang (University of Toronto)
- Alan P.-T. Lau (Hong Kong Polytechnical University)

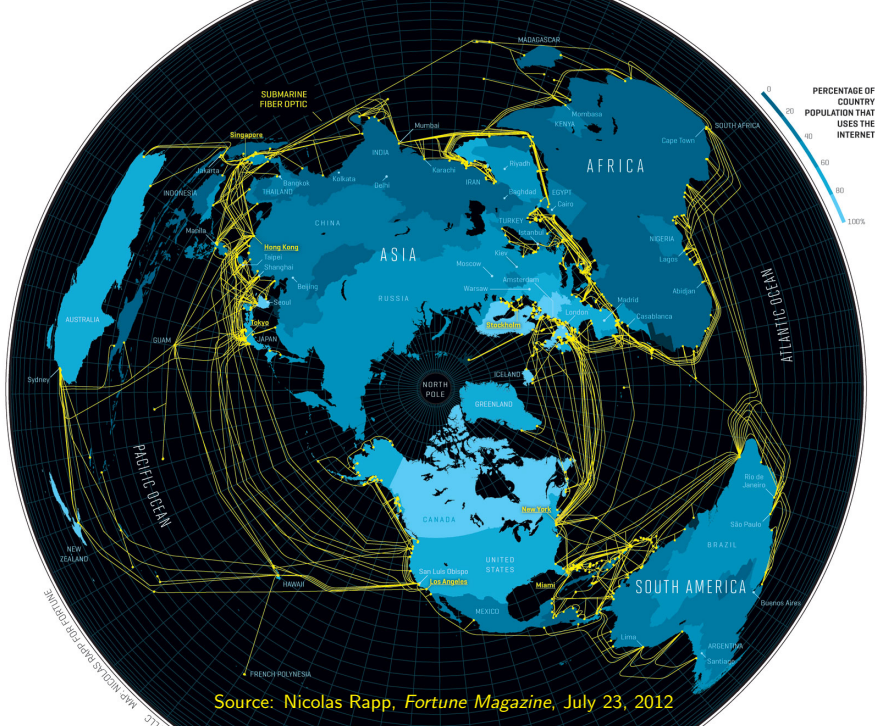
The organizers:

- Fredrik Bränström
- Giuseppe Durisi
- Alexandre Graell i Amat

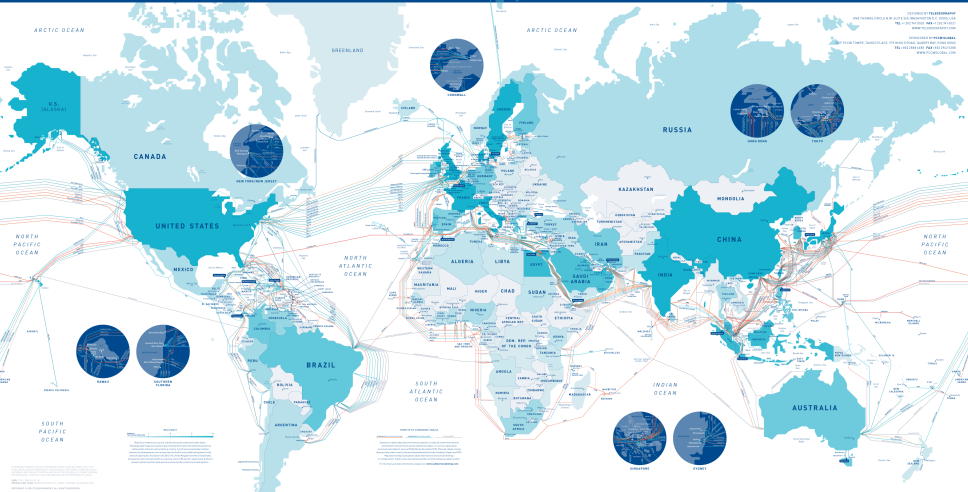


Illustrating the principle of total internal reflection.





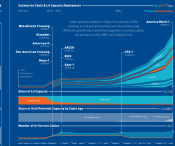
Source: Nicolas Rapp, *Fortune Magazine*, July 23, 2012



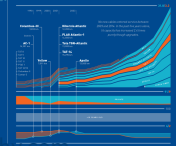
## MAJOR ROUTES



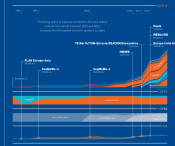
## U.S. - LATIN AMERICA



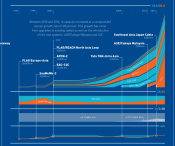
## TRANS-ATLANTIC



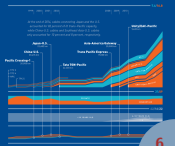
## EUROPE-ASIA



## INTRA-ASIA

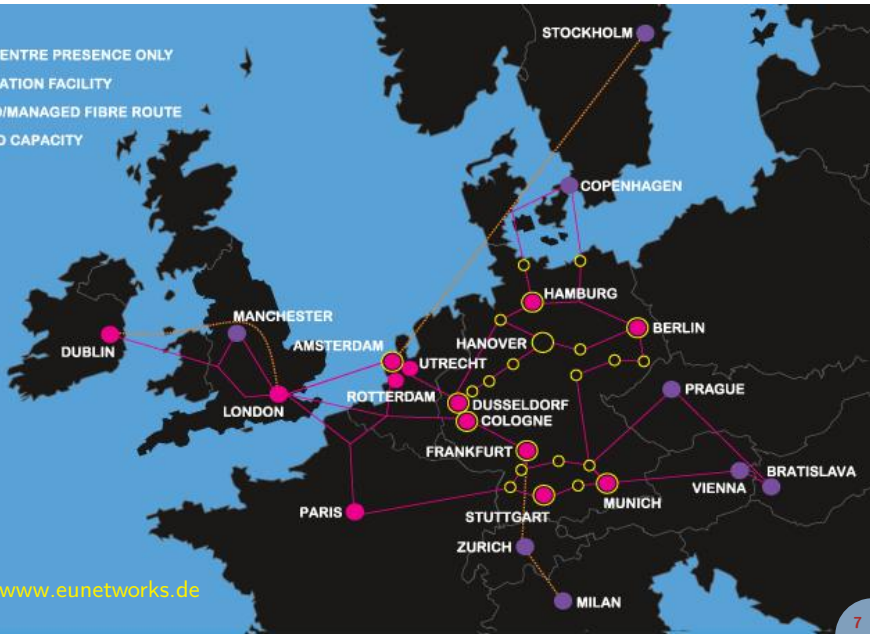


## TRANS-PACIFIC



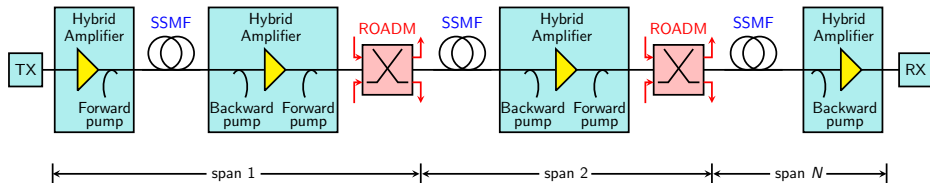
# euNetworks Optical Fiber Network Map

- MAN
- DATA CENTRE PRESENCE ONLY
- COLOCATION FACILITY
- OWNED/MANAGED FIBRE ROUTE
- LEASED CAPACITY



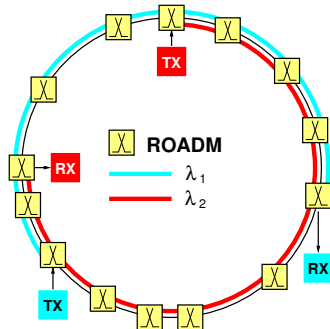
Source: [www.eunetworks.de](http://www.eunetworks.de)

# Fiber-Optic Communication Systems



## Physics: Enabling Technologies

- 1 Low-loss ( $\sim 0.2$  dB/km) optical fiber, huge bandwidth ( $\sim 54$  THz bandwidth)
- 2 Optical amplifiers
- 3 Laser transmitters and Mach-Zehnder modulators





# Fiber-Optic Communication Systems: Challenges

## Reliability

$$P_e < 10^{-15}$$

# Fiber-Optic Communication Systems: Challenges

## Reliability

$$P_e < 10^{-15}$$

## Speed

100 Gb/s per-channel data rates (or even more)

# Fiber-Optic Communication Systems: Challenges

## Reliability

$$P_e < 10^{-15}$$

## Speed

100 Gb/s per-channel data rates (or even more)

## Non-Linearity

The fiber-optic channel is **non-linear** in the input power!

# The Kerr Effect

## Kerr electro-optic effect (DC Kerr effect)

- An effect discovered by John Kerr in 1875
- It produces a change of refractive index in the direction parallel to the externally applied electric field
- The change of index is proportional to the *square* of the magnitude of the external field

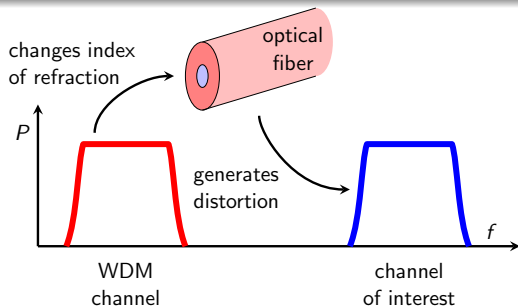


**John Kerr**  
(1824-1907)

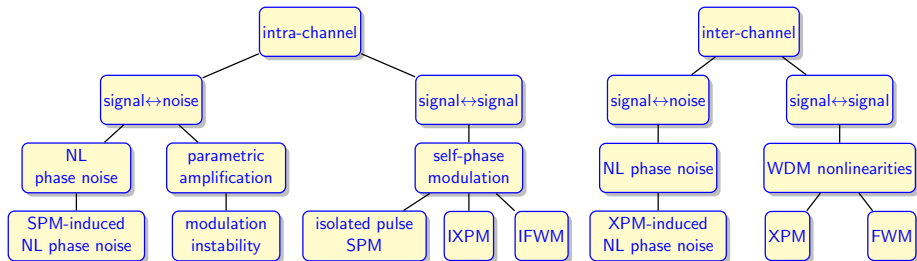
# The Kerr Effect in Optical Fibers

## Optical Kerr Effect (or AC Kerr effect)

- No externally applied electric field is necessary
- The signal light itself produces the electric field that changes the index of refraction of the material (fused silica)
- The change in index in turn changes the signal field
- The change in index of refraction is proportional to the square of the field magnitude



# Nonlinear Effects in Fibers



## Key

NL = nonlinear; SPM = self-phase modulation; (I)XPM = (intra-channel) cross-phase modulation; (I)FWM = (intra-channel) four-wave mixing; WDM = wavelength-division multiplexing

Courtesy of R. J. Essiambre

# Maxwell's Equations

## Maxwell's Equations

$$\text{Gauss's Law: } \nabla \cdot \mathbf{D} = \rho$$

$$\text{Gauss's Law for Magnetism: } \nabla \cdot \mathbf{B} = 0$$

$$\text{Maxwell-Faraday Equation: } \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\text{Ampère's Law: } \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

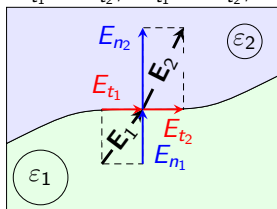
with  $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$  and  $\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})$ .

Tangential (t) and normal (n) components subject, at a material boundary, to:

$$E_{t_1} = E_{t_2}; H_{t_1} = H_{t_2}; D_{n_1} - D_{n_2} = \rho_S; B_{n_1} = B_{n_2}.$$

- $\epsilon_0 \approx \frac{1}{36\pi} 10^{-9}$  As/Vm: free space dielectric permittivity
- $\mu_0 = 4\pi \times 10^{-7}$  Vs/Am: free space magnetic permeability

- **E**: electric field (V/m)
- **H**: magnetic field (A/m)
- **D**: electric displacement (As/m<sup>2</sup>)
- **B**: magnetic induction (Vs/m<sup>2</sup>)
- **J**: current density (A/m<sup>2</sup>)
- **P**: material polarization (As/m<sup>2</sup>)
- **M**: material magnetization (A/m)
- $\rho$ : charge density (As/m<sup>3</sup>)
- $\epsilon_0 \mu_0 = \frac{1}{c^2}$ , where  $c \approx 3 \times 10^8$  m/s: free space speed of light



# Maxwell's Equations in Air and Glass

**Approximation 1:**  $\rho = 0$ ; there are no free charges.

**Approximation 2:**  $\mathbf{J} = 0$ ; there are no currents.

**Approximation 3:**  $\mathbf{M} = 0$ ; there is no magnetization.

## Maxwell's Equations after Approximations 1–3

$$\nabla \cdot \mathbf{D} = 0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}$$

with  $\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}$  and  $\mathbf{B} = \mu_0 \mathbf{H}$ .



# Material Polarization

Material polarization in a general medium is often written as

$$\mathbf{P} = \epsilon_0(X^{(1)}(\mathbf{E}) + X^{(2)}(\mathbf{E}, \mathbf{E}) + X^{(3)}(\mathbf{E}, \mathbf{E}, \mathbf{E}) + \dots)$$

where  $X^{(i)}$  are linear ( $i = 1$ ) and nonlinear ( $i > 1$ ) *electric susceptibility tensors* (multilinear maps).

We will assume that in glass fiber:

## Approximation 4:

The material polarization  $\mathbf{P}$  parallels the electric field  $\mathbf{E}$  (glass is a homogeneous medium).

## Approximation 5:

Changes in  $\mathbf{P}$  follow changes in  $\mathbf{E}$  instantaneously (i.e., quicker than any relevant time scale involved).

Assuming material homogeneity (Approximation 4),  $X^{(i)}$  is a constant, and we write

$$X^{(i)}(\mathbf{E}, \mathbf{E}, \dots, \mathbf{E}) = \chi^{(i)} \|\mathbf{E}\|^i \frac{\mathbf{E}}{\|\mathbf{E}\|} = \chi^{(i)} \|\mathbf{E}\|^{i-1} \mathbf{E}.$$

# Kerr Nonlinearity

Glass, like many materials, has an “inversion symmetry” (it’s mirror image has the same structure as itself)  $\Rightarrow$  all even-order nonlinear susceptibilities are zero:  $\chi^{(2)} = \chi^{(4)} = \dots = 0$ .

## Approximation 6:

Higher-order susceptibilities beyond  $\chi^{(3)}$  are negligible.

This gives

$$\mathbf{P} = \varepsilon_0(\chi^{(1)} + \chi^{(3)}\|\mathbf{E}\|^2)\mathbf{E}$$

and, therefore,

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P} = \varepsilon_0(1 + \chi^{(1)} + \chi^{(3)}\|\mathbf{E}\|^2)\mathbf{E}.$$

This is the Kerr nonlinearity!

# Towards the Nonlinear Schrödinger Equation

By considering  $\nabla \times \nabla \times \mathbf{E}$ , one can derive a *nonlinear wave equation*.

Then, after introducing

- a complex-envelope representation for a propagating electric field in the  $z$  direction;
- a co-propagating reference frame;
- wavelength-dependence of the index of refraction (dispersion) (up to second order);
- fiber loss;

one obtains a partial differential equation for the envelope  $Q(t, z)$  in the absence of noise:

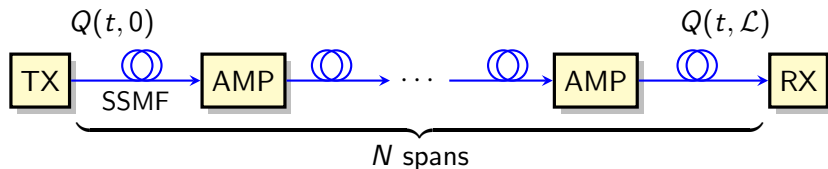
$$\frac{\partial}{\partial z} Q(t, z) + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} Q(t, z) - j\gamma |Q(t, z)|^2 Q(t, z) + \frac{\alpha}{2} Q(t, z) = 0.$$

# System Model

- coherent fiber-optic communication system
- standard-single-mode fiber
- ideal distributed Raman amplification

But, could also consider

- systems with inline dispersion-compensating fiber
- lumped amplification



# Generalized Nonlinear Schrödinger Equation

- $Q(t, z)$  is the **complex baseband representation** of the signal (the full field, representing co-propagating DWDM signals)
- **Transmitter** sends  $Q(t, 0)$
- **Receiver** gets  $Q(t, \mathcal{L})$ , where  $\mathcal{L}$  is the total system length

## Evolution of $Q(t, z)$

The generalized non-linear Schrödinger (GNLS) equation expresses the evolution of  $Q(t, z)$ :

$$\frac{\partial Q(t, z)}{\partial z} + \underbrace{\frac{j\beta_2}{2} \frac{\partial^2 Q(t, z)}{\partial t^2}}_{\text{dispersion}} - \underbrace{j\gamma |Q(t, z)|^2 Q(t, z)}_{\text{nonlinearity}} = \underbrace{n(t, z)}_{\text{noise}}$$

No loss term (since ideal distributed Raman amplification is assumed).

## System Parameters

$n(t, z)$  is a circularly symmetric complex Gaussian noise process with autocorrelation

$$\mathcal{E} [n(t, z)n^*(t', z')] = \alpha h \nu_s K_T \delta(t - t', z - z'),$$

where  $h$  is Planck's constant,  $\nu_s$  is the optical frequency, and  $K_T$  is the phonon occupancy factor.

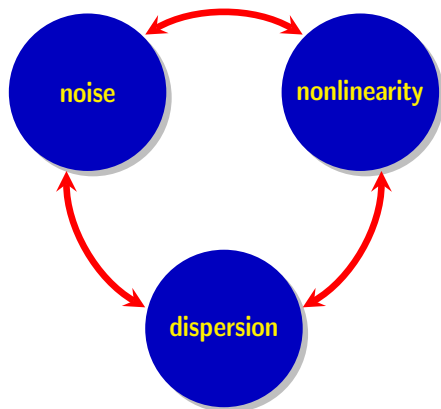
---

|                                   |  |
|-----------------------------------|--|
| Second-order dispersion $\beta_2$ | -21.668 ps <sup>2</sup> /km            |
| Loss $\alpha$                     | $4.605 \times 10^{-5}$ m <sup>-1</sup> |
| Nonlinear coefficient $\gamma$    | 1.27 W <sup>-1</sup> km <sup>-1</sup>  |
| Center carrier frequency $\nu_s$  | 193.41 THz                             |
| Phonon occupancy factor $K_T$     | 1.13                                   |

---

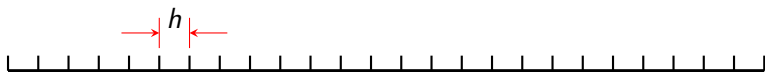
# Solving the GNLS Equation

Throughout propagation over an optical fiber, stochastic effects (noise), linear effects (dispersion) and nonlinear effects (Kerr nonlinearity) **interact**.



Even in the absence of noise, solving the GNLS equation requires **numerical techniques**.

# Split-Step Fourier Method



- divide fiber length into short segments
- consider each segment as the concatenation of (separable) nonlinear and linear transforms
- for distributed amplification, an additive noise is added after the linear step.

$$Q(t, z_0) \longrightarrow Q(t, z_0 + h) \quad \text{step size } h$$



## Nonlinear Step

In the absence of linear effects, the GNLS equation has the form

$$\frac{\partial Q(t, z)}{\partial z} = j\gamma |Q(t, z)|^2 Q(t, z),$$

with solution

$$Q(t, z_0 + h) = Q(t, z_0) \exp(j\gamma |Q(t, z_0)|^2 h).$$

### Nonlinear Step ...

... is best taken in the **time domain**.

## Linear Step

Now use the previous solution as input to the linear step, i.e., let

$$\hat{Q}(t, z_0) = Q(t, z_0) \exp(j\gamma |Q(t, z_0)|^2 h)$$

be the input to the linear step. The linear form of the GNLS equation is

$$\frac{\partial Q(t, z)}{\partial z} = -\frac{j\beta_2}{2} \frac{\partial^2 Q(t, z)}{\partial t^2}.$$

Defining

$$Q(t, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{Q}(\omega, z) \exp(j\omega t) d\omega, \text{ i.e., } Q(t, z) \xleftrightarrow{\mathcal{F}} \tilde{Q}(\omega, z),$$

it can be shown that

$$\tilde{Q}(\omega, z_0 + h) = \tilde{Q}(\omega, z_0) \exp\left(j\frac{\beta_2}{2} \omega^2 h\right).$$

### Linear Step ...

... is best taken in the **frequency domain**.

# Split-Step Propagator

Putting this together, we have

$$Q(t, z_0 + h) =$$

$$\underbrace{\mathcal{F}^{-1} \left\{ \mathcal{F} \left\{ \underbrace{Q(t, z_0) \exp(j\gamma |Q(t, z_0)|^2 h)}_{\text{nonlinear step}} \right\} \exp\left(j\frac{\beta_2}{2} \omega^2 h\right) \right\}}_{\text{linear step}}$$

where  $\mathcal{F}$  is the Fourier transform operator.

In practice:

- 1 extensive use is made of the FFT
- 2 step size can be adapted
- 3 linear and nonlinear steps can be “offset” by a half-step

## Local-error method

Given the signal  $q(t, z)$ , compute the signal at  $z + 2h$  in two ways: a “coarse” solution  $q_c$  using the symmetric split-step scheme once with a step-size  $2h$ , and a “fine” solution  $q_f$  using the symmetric split-step scheme twice with step-size  $h$ .

Since the error in a single step of the symmetric scheme is  $\mathcal{O}(h^3)$ , these solutions are related to the actual solution  $q_a$  as

$$q_c = q_a + \kappa(2h)^3 + \mathcal{O}(h^4),$$

$$q_f = q_a + 2\kappa h^3 + \mathcal{O}(h^4).$$

Using a linear combination of the coarse and fine solutions, you can get an approximate solution at  $z + 2h$  for which the leading error term is  $\mathcal{O}(h^4)$

$$q(t, z + 2h) = \frac{4}{3}q_f - \frac{1}{3}q_c = q_a + \mathcal{O}(h^4)$$

## Local-error method: adaptive step sizes

Define a *relative local error*

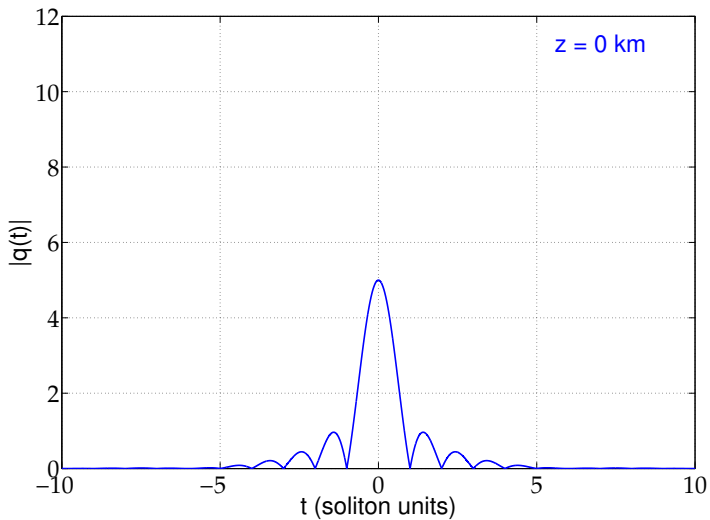
$$\delta = \frac{\|q_f - q_c\|}{\|q_f\|}$$

The step-size is adaptively chosen to keep the relative local error within a specified tolerance range ( $tol/2, tol$ )

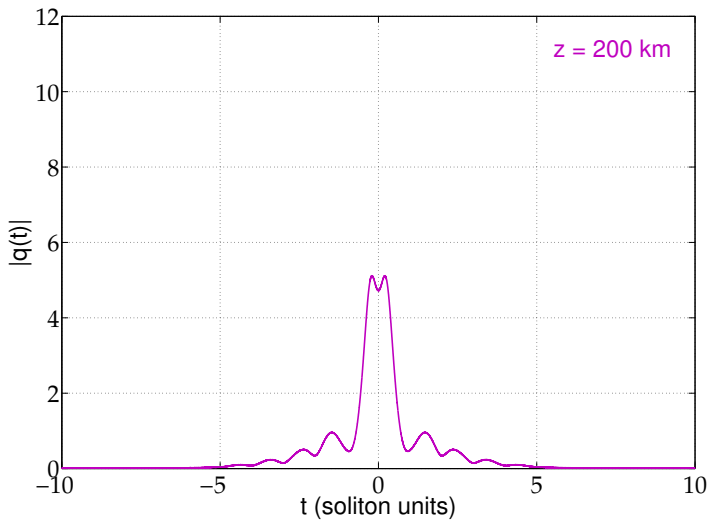
- If  $\delta < tol/2$ , then  $h = 2^{1/3}h$ ,
- else if  $\delta \in (tol/2, tol)$ ,  $h = h$ ,
- else if  $\delta \in (tol, 2tol)$ ,  $h = h/2^{1/3}$ ,
- else if  $\delta > 2tol$ , discard solution and  $h = h/2$ .

See: Sinkin, Holzlohner, Zweck, Menyuk, *J. Lightwave Tech.*, 2003.

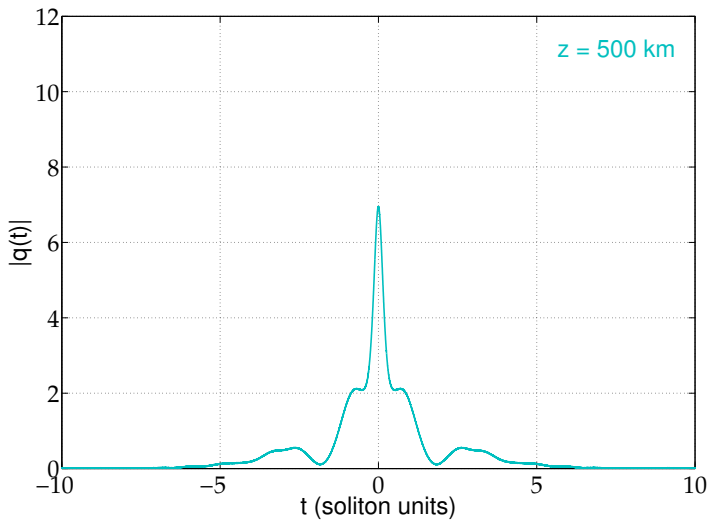
# Launching a Pulse



# Launching a Pulse

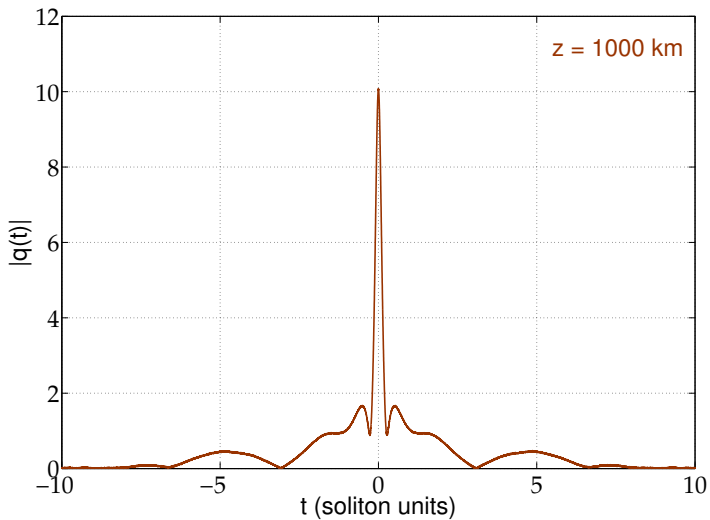


# Launching a Pulse





# Launching a Pulse



# An Upper Bound on Spectral Efficiency of Fiber Channels

(Joint work with Mansoor Yousefi and Gerhard Kramer)

Basic idea:

- take the split-step method as a channel model
- observe that each noiseless step is both energy and entropy preserving
- total power accumulated gives an upper bound on entropy of output
- entropy power inequality gives a lower bound on the entropy of a noise ball

# Maximum Entropy

Let  $\mathbf{X} \in \mathbb{C}^L$  have correlation matrix  $R_{\mathbf{X}} = E[\mathbf{X}\mathbf{X}^H]$ . Then

$$h(\mathbf{X}) \leq \log[(\pi e)^L \det R_{\mathbf{X}}],$$

with equality if and only if  $\mathbf{X}$  is Gaussian and circularly symmetric. If  $E[\mathbf{X}^H\mathbf{X}] = LP$  (per-sample energy  $P$ ), then

$$h(\mathbf{X}) \leq L \log(\pi e P)$$

with equality if and only if  $\mathbf{X}$  is iid circularly symmetric Gaussian. For any complex square matrix  $\mathbf{M}$ , we have

$$h(\mathbf{M}\mathbf{X}) = h(\mathbf{X}) + 2 \log |\det \mathbf{M}|.$$

In particular, if  $\mathbf{M}$  has unit determinant (e.g., if  $\mathbf{M}$  is unitary), then  $h(\mathbf{M}\mathbf{X}) = h(\mathbf{X})$ .

# Entropy Power Inequality

For  $\mathbf{X} \in \mathbb{C}^L$ , define the “entropy power”  $V(\mathbf{X})$  so that

$$h(\mathbf{X}) = L \log(\pi e V(\mathbf{X})).$$

## Entropy Power Inequality

For independent  $\mathbf{X}$  and  $\mathbf{Y}$  we have

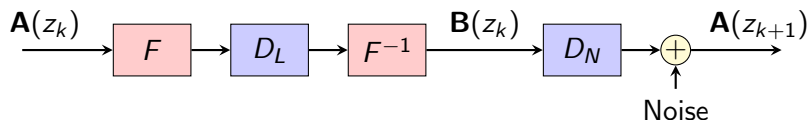
$$V(\mathbf{X} + \mathbf{Y}) \geq V(\mathbf{X}) + V(\mathbf{Y}).$$

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

$$\begin{aligned} h(\mathbf{X} | \mathbf{U}) &= L \log(\pi e V(\mathbf{X} | \mathbf{U})), \text{ and} \\ &= V(\mathbf{X} + \mathbf{Y} | \mathbf{U}) \geq V(\mathbf{X} | \mathbf{U}) + V(\mathbf{Y} | \mathbf{U}) \end{aligned}$$

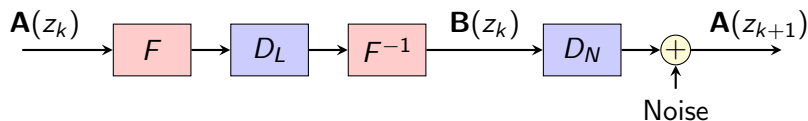
# Fiber-Channel via Split-Step

Split-step channel model: divide fiber into small steps.



- Ideal Raman amplification: removes loss but adds noise
- $F$  = Fourier transform
- $D_L$  = diagonal matrix with fixed entries of unit magnitude (all-pass filter)
- $D_N$  = diagonal matrix with unit amplitude entries; the  $(\ell, \ell)$ th-entry phase shift is proportional to the magnitude-squared of the  $\ell$ th entry of  $\mathbf{B}(z_k)$ .

# Main Observations



- The linear step conserves energy and entropy.

$$\mathbf{B}(z_k) = \underbrace{F^{-1} D_L F}_{\text{unitary}} \mathbf{A}(z_k)$$

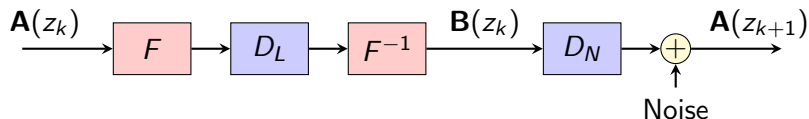
- The nonlinear step **also** conserves energy and entropy.

*Sketch:* Let  $(r, \theta)$ ,  $r \geq 0$ ,  $\theta \in [0, 2\pi)$ , be random variables, and let  $X = re^{j\theta}$ ; then  $h(X) = h(r, \theta) + E[\log(r)]$ .

Now, if  $Y = re^{j(\theta+g(r))}$  for some function  $g(\cdot)$ , then

$$\begin{aligned} h(Y) &= h(r, \theta+g(r)) + E[\log(r)] = h(r) + h(\theta+g(r) | r) + E[\log(r)] = \\ &= h(r) + h(\theta | r) + E[\log(r)] = h(X). \end{aligned}$$

## Total Energy after $K$ steps

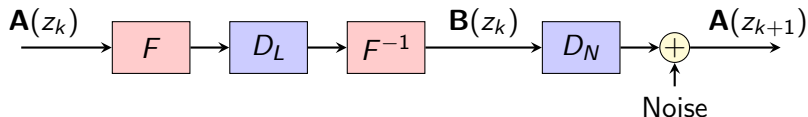


$$\begin{aligned}\text{Channel Output Energy} &= \text{Launch Energy} + \text{Noise Energy} \\ &= LP + LKN\end{aligned}$$

Thus:

$$h(\mathbf{A}(z_K)) \leq L \cdot \log(\pi e(P + KN))$$

# Equivocation Growth



By the EPI:

$$V(\mathbf{A}(z_{k+1}) | \mathbf{A}(z_0)) \geq V(\mathbf{A}(z_k)) + N,$$

thus

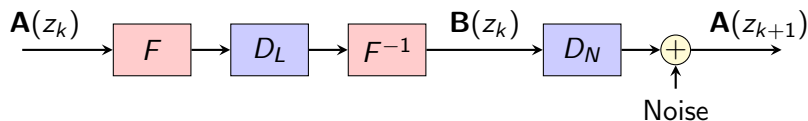
$$V(\mathbf{A}(z_K) | \mathbf{A}(z_0)) \geq KN$$

or

$$h(\mathbf{A}(z_K) | \mathbf{A}(z_0)) \geq L \log(\pi eKN).$$



# Result



Thus we get:

$$\begin{aligned} I(\mathbf{A}(z_0); \mathbf{A}(z_K)) &= h(\mathbf{A}(z_K)) - h(\mathbf{A}(z_K) | \mathbf{A}(z_0)) \\ &\leq L \log(\pi e(P + KN)) - L \log(\pi eKN) \\ &= L \log\left(1 + \frac{P}{KN}\right) \\ &= L \log(1 + \text{SNR}) \end{aligned}$$

and finally, per sample,

$$\frac{1}{L} I(\mathbf{A}(z_0); \mathbf{A}(z_K)) \leq \log(1 + \text{SNR}).$$

# Discussion

- Let  $B = 1/(\Delta t)$  be the “bandwidth” of the simulation
- Then  $L = T/(\Delta t) = TB$  is the time-bandwidth product
- The spectral efficiency is thus bounded by

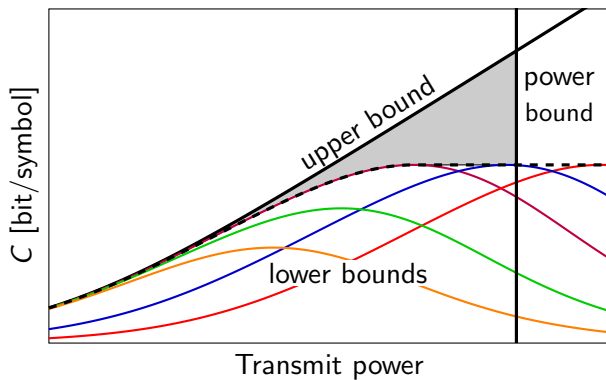
$$\eta \leq \log(1 + \text{SNR}) \text{ [bit/s/Hz]}$$

- Why normalize by the simulation bandwidth  $B$ ? The “real” bandwidth  $W$  can be smaller. (Answer:  $B$  can be chosen as the smallest bandwidth for which simulations give accurate results.)
- What about capacity? (Answer: any real fiber has a maximal bandwidth  $B_{\max}$ ; a capacity upper bound follows by multiplying  $B_{\max}$  by  $\log(1 + \text{SNR})$ .)

## Discussion (cont'd)

- What about MIMO fiber? (Answer: if energy and entropy are preserved by the linear and non-linear steps, and the noise is AWGN, then the above bound remains valid per mode.)
- What about lower bounds? (Answer: open research! Bounding bandwidth expansion is important.)

# Optical Fiber Capacity Bounds

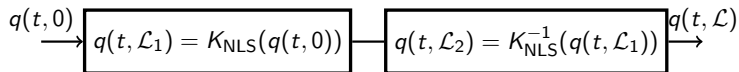


Upper bound:

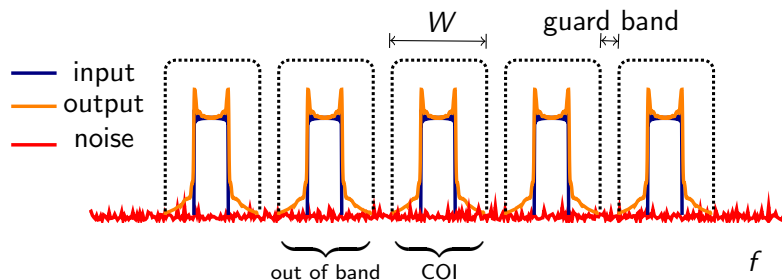
- 1 G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper Bound on the Capacity of a Cascade of Nonlinear and Noisy Channels," *Proc. IEEE Info. Theory Workshop*, Jerusalem, Israel, Apr. 2015.
- 2 M. I. Yousefi, G. Kramer, and F. R. Kschischang, "Upper Bound on the Capacity of the Nonlinear Schrödinger Channel," *Proc. 14th Canadian Workshop on Info. Theory*, St. John's, NL, Jul. 2015.

# Current Approaches

## ① Digital Backpropagation



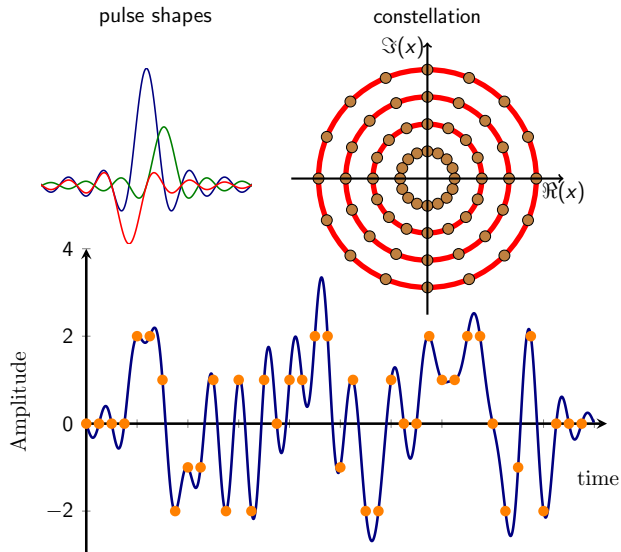
## ② Wavelength-division multiplexing (WDM)



# Digital Backpropagation

- Digital backpropagation = split-step Fourier method, using a negative step-size  $h$ , performed at the receiver
- Full compensation (involving multiple WDM channels) generally impossible, even in absence of noise (due to wavelength routing)
- Noise is neglected (cf. zero-forcing equalizer)
- Single-channel backpropagation typically performed, after extraction of desired channel using a filter

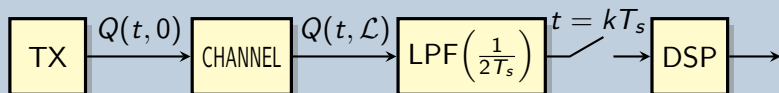
# Current Achievable Rates



# Capacity Estimation

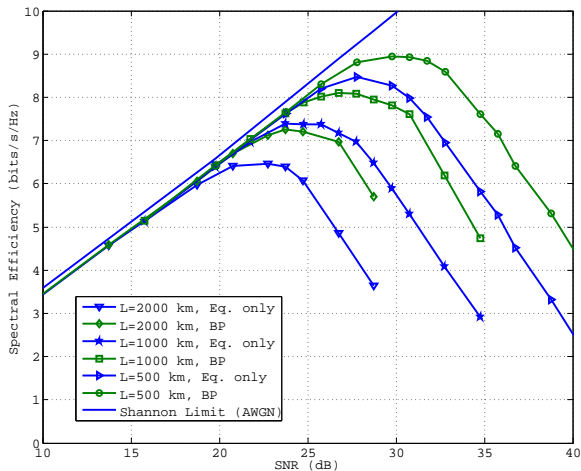
See: R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, pp. 662–701, Sept./Oct. 2010.

## System Model





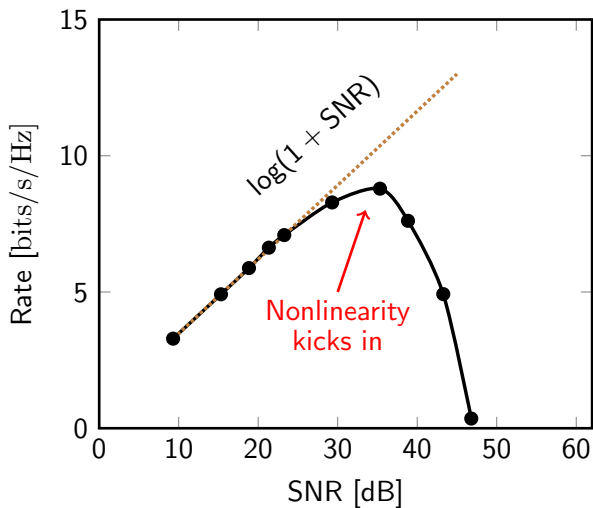
# Achievable Rates from Memoryless Capacity Estimate



(BP adds  
0.55 to 0.75  
bits/s/Hz  
relative to  
EQ)

B. P. Smith and F. R. Kschischang, *J. Lightwave Techn.*,  
vol. 30, pp. 2047–2053, 2012.

# Schematic Achievable Rate Curve



noise-limited

...

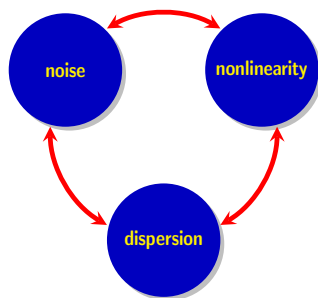
nonlinearity-limited

# Summary

Pulse propagation in optical fibers is well-modeled by the **stochastic nonlinear Schrödinger equation**:

$$\frac{\partial Q(t, z)}{\partial z} + \underbrace{\frac{j\beta_2}{2} \frac{\partial^2 Q(t, z)}{\partial t^2}}_{\text{dispersion}} - \underbrace{j\gamma |Q(t, z)|^2 Q(t, z)}_{\text{nonlinearity}} = \underbrace{n(t, z)}_{\text{noise}}$$

- $z$  is distance along the fiber; TX at  $z = 0$ , RX at  $z = \mathcal{L}$
- $t$  is retarded time, i.e.,  $t = \tau - \beta_1 z$  where  $\tau$  is ordinary time and  $\beta_1$  is a constant
- $Q(t, z)$  is the complex envelope of the propagating signal
- $\beta_2$  is the *chromatic dispersion coefficient*
- $\gamma$  is the *nonlinearity parameter*
- $V(t, z)$  is space-time white Gaussian noise



# Staircase Codes

This part of the talk is about ...

- FEC for the binary symmetric channel (BSC)
- After optical (and/or) electrical compensation, suitable as forward-error-correction (FEC) for 100Gb/s PD-QPSK systems *without* soft information

## Outline

- 1 Existing solutions (G.975, G.975.1)
- 2 Implementation considerations
- 3 Staircase codes
- 4 Ensemble analysis
- 5 Conclusions

## Net Coding Gain

Given a particular BER, we can obtain the corresponding  $Q$  via

$$Q = \sqrt{2} \operatorname{erfc}^{-1}(2 \cdot \text{BER}).$$

The coding gain (CG) of an error-correcting code is defined as

$$\text{CG (in dB)} = 20 \log_{10}(Q_{\text{out}}/Q_{\text{in}}),$$

and the Net Coding Gain (NCG) is

$$\text{NCG (in dB)} = \text{CG} + 10 \log_{10}(R),$$

where  $R$  is the rate of the code.

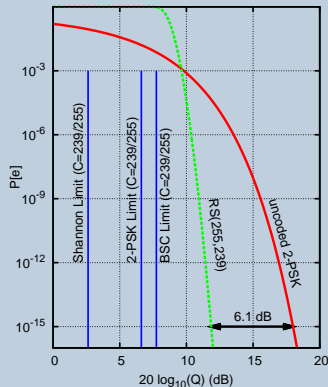
## Overhead

$$\text{Overhead} = \frac{\# \text{ redundant symbols}}{\# \text{ information symbols}} = \frac{n-k}{k} = \frac{1-R}{R}$$

# Existing Solutions

## Reed-Solomon (255,239) Code

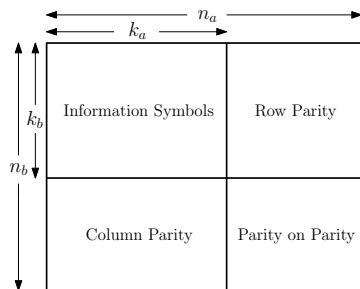
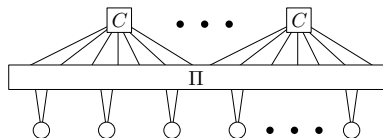
- Coding Symbols are bytes
- Depth-16 interleaving corrects (some) bursts up to 1024 bits
- Coding Gain = 6.1 dB
- Net Coding Gain = 5.8 dB
- Overhead =  $16/239 = 6.69\%$



Constraint: Rate and framing structure **fixed** for future generations

## Concatenated Codes

- Product-like codes with algebraic component codes





# ITU-T G.975.1 (2004)

| Code | NCG                          | Notes   |
|------|------------------------------|---|
| I.2  | 8.88 dB @ $10^{-15}$         | Outer RS, Inner CSOC                          |
| I.3  | 8.99 dB @ $10^{-15}$         | Outer BCH ( $t = 3$ ), Inner BCH ( $t = 10$ ) |
| I.4  | 8.67 dB @ $10^{-15}$         | Outer RS, Inner BCH ( $t = 8$ )               |
| I.5  | 8.5 dB @ $10^{-15}$          | Outer RS, Inner Product ( $t = 1$ )           |
| I.6  | 8.02 dB @ $10^{-15}$         | LDPC  |
| I.7  | 8.09 dB @ $10^{-15}$         | Outer BCH ( $t = 4$ ), Inner BCH ( $t = 11$ ) |
| I.8  | 8.00 dB @ $10^{-15}$         | RS(2720,2550)                                 |
| I.9  | 8.63 dB @ $7 \cdot 10^{-14}$ | ~Product BCH ( $t = 3$ ), Erasure Dec?        |

# Objectives

## Increased Net Coding Gain

NCG  $\Rightarrow$  Shannon Limit of BSC (9.97 dB at  $10^{-15}$ )

## Error Floor

- Error floor  $\ll 10^{-15}$
- Lower error floor  $\Rightarrow$  'Insurance' in presence of correlated errors

## Block Length

$n \approx 2 \cdot 10^6$  or less

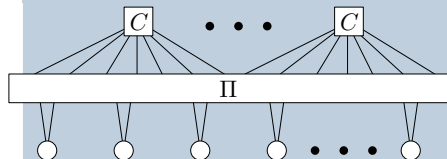
## Low Implementation Complexity

Dataflow considerations at 100Gb/s

## Implementation Considerations

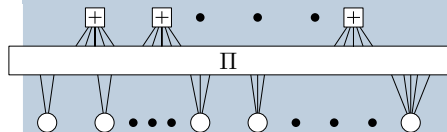
# Hardware Considerations: Product vs. LDPC

## Product Code



- Algebraic component codes
- Syndrome-based decoding

## LDPC Code



- SPC component codes
- Belief propagation decoding

# Syndrome-based Decoding of Product Codes

## Decoding an $(n, k)$ component codeword

- $n$  received symbols  $\Rightarrow n - k$  symbol syndrome
- $R = 239/255$ ,  $n \approx 1000$ ,  $n - k \approx 32$ 
  - For high-rate codes, syndromes provide **compression**

- $\sim 3$  decodings/component
- $\leq 96$  bits/decoding
- $\sim \frac{2}{1000}$  components/symbol

$\Rightarrow$  **0.768 bits/symbol**



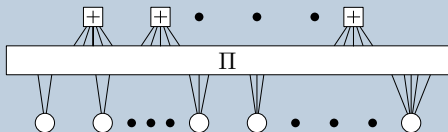
## Total Dataflow

At 100Gb/s, 76.8 Gb/s internal dataflow

# LDPC Belief Propagation Decoding

- $\sim 15$  iterations
- 2 messages/iteration·edge
- $\sim 5$  bits/message
- $\sim 3$  edges/symbol

$\Rightarrow$  **450 bits/symbol**



## Total Dataflow

At 100Gb/s, 45Tb/s internal dataflow!

## Dataflow Comparison

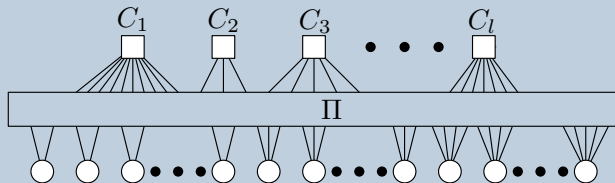
76.8 Gb/s  $\ll$  45 Tb/s

2–3 orders of magnitude (huge implementation challenge for soft message-passing LDPC decoders).

## Our Solution



# Coding with Algebraic Component Codes



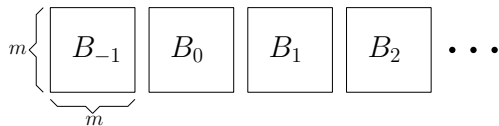
## Graph Optimization

Degrees of Freedom

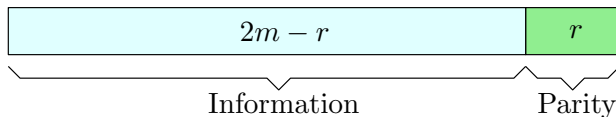
- Mixture of component codes (e.g., Hamming, BCH)
- Multi-edge-type structures

# Staircase Codes: Construction

Consider a sequence of  $m$ -by- $m$  matrices  $B_i$



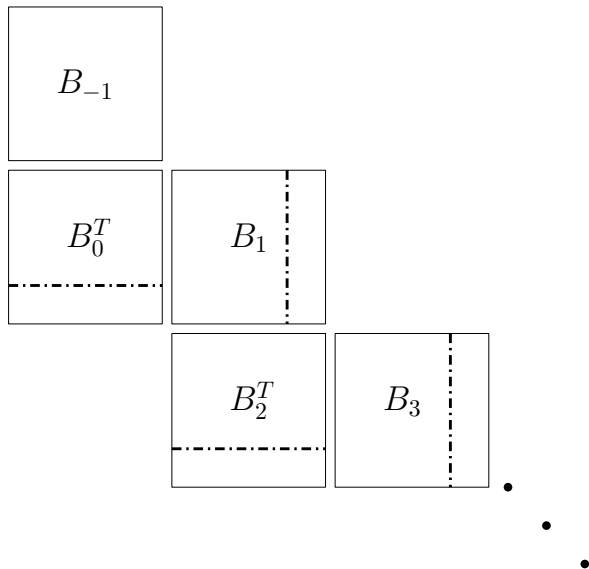
and a linear, systematic,  $(n = 2m, k = 2m - r)$  component code  $C$



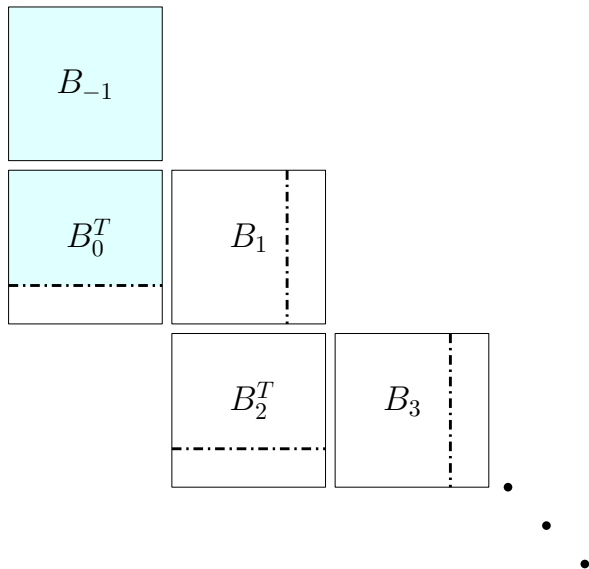
## Encoding Rule

$\forall i \geq 0$ , all rows of  $[B_{i-1}^T B_i]$  are codewords in  $C$

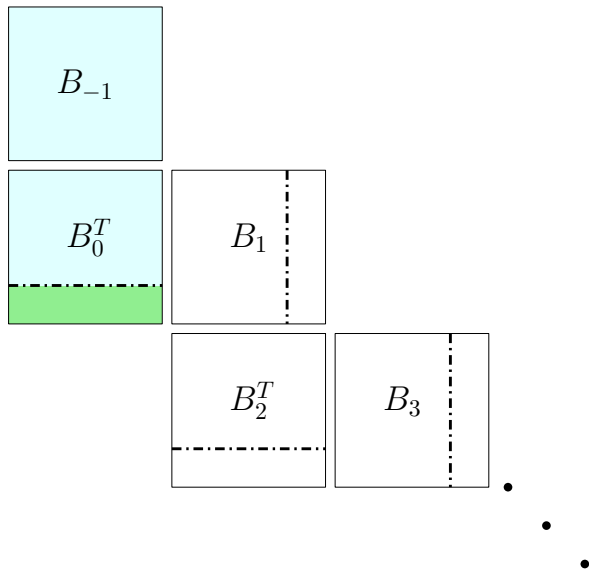
# Staircase Codes: Construction



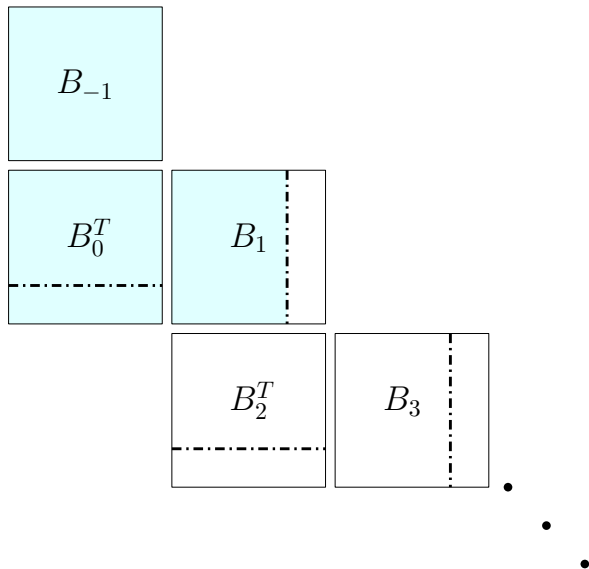
# Staircase Codes: Construction



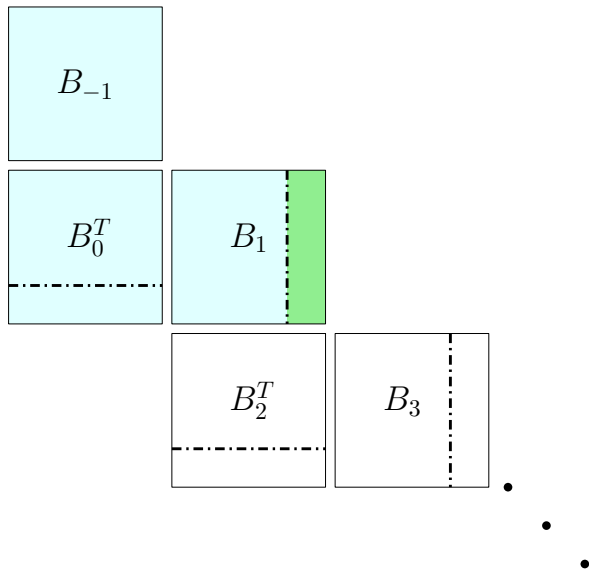
# Staircase Codes: Construction



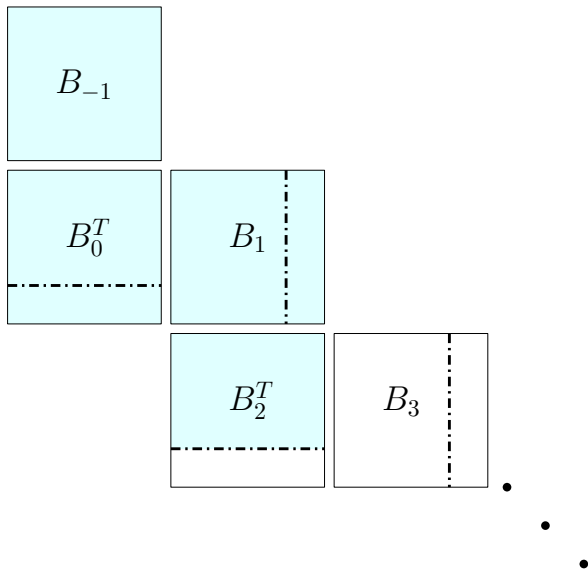
# Staircase Codes: Construction



# Staircase Codes: Construction

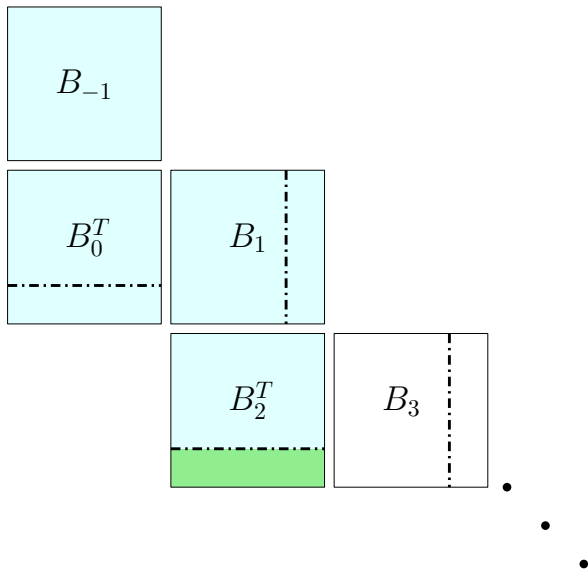


# Staircase Codes: Construction



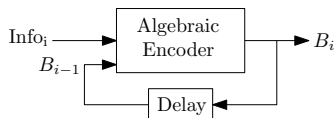


# Staircase Codes: Construction

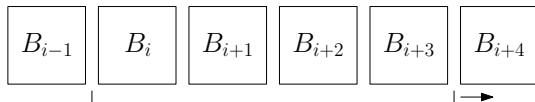


# Staircase Codes: Properties

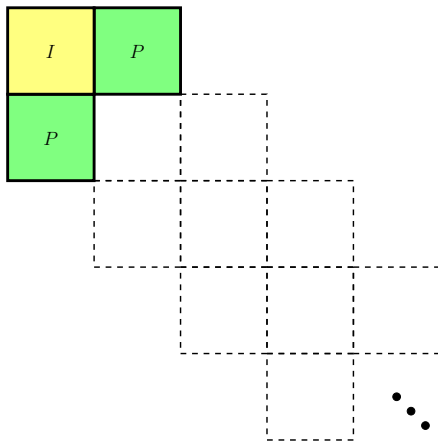
- Hybridization of recursive convolution coding and block coding
  - Recurrent Codes of Wyner-Ash (1963)



- Rate :  $R = 1 - r/m$
- Variable-latency (sliding-window) decoder

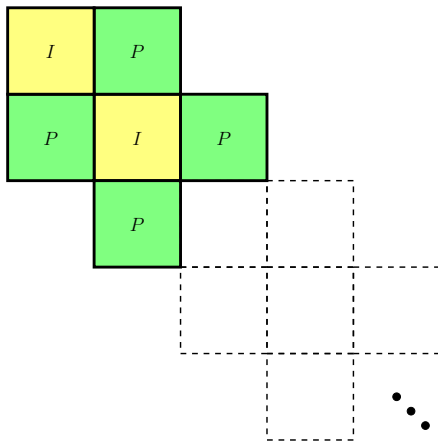


# Staircase Codes are a type of Braided Block Code



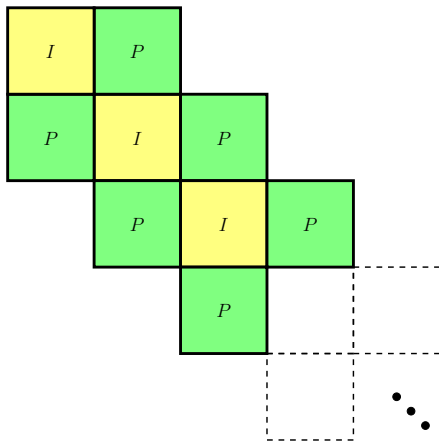
Felström, Truhachev, Lentmaier, Zigangirov, "Braided Block Codes," *IEEE Trans. on Info. Theory*, 2009.

# Staircase Codes are a type of Braided Block Code



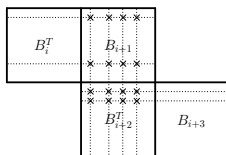
Felström, Truhachev, Lentmaier, Zigangirov, “Braided Block Codes,” *IEEE Trans. on Info. Theory*, 2009.

# Staircase Codes are a type of Braided Block Code



Felström, Truhachev, Lentmaier, Zigangirov, "Braided Block Codes," *IEEE Trans. on Info. Theory*, 2009.

# Staircase Codes: Properties



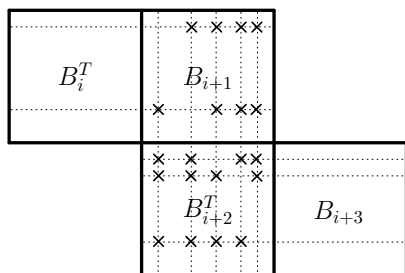
The multiplicity of (minimal) stalls of size  $(t + 1) \times (t + 1)$  is

$$K = \binom{m}{t+1} \cdot \sum_{j=1}^{t+1} \binom{m}{j} \binom{m}{t+1-j},$$

and the corresponding contribution to the error floor, for transmission over a binary symmetric channel with crossover probability  $p$ , is

$$\text{BER}_{\text{floor}} = K \cdot \frac{(t+1)^2}{m^2} \cdot p^{(t+1)^2}.$$

# General Stall Patterns

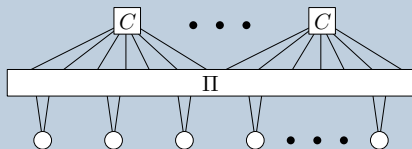
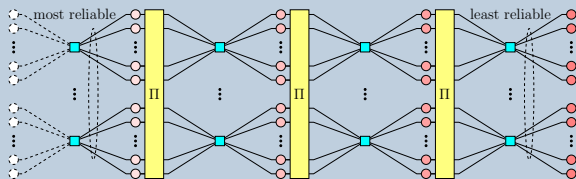


(5,5)-stall

| $K$ | $L$ | Contribution           |
|-----|-----|------------------------|
| 4   | 4   | $3.55 \times 10^{-21}$ |
| 4   | 5   | $7.81 \times 10^{-28}$ |
| 5   | 5   | $2.54 \times 10^{-22}$ |
| 5   | 6   | $2.21 \times 10^{-28}$ |
| 6   | 6   | $1.40 \times 10^{-23}$ |
| 6   | 7   | $1.49 \times 10^{-29}$ |
| 7   | 7   | $8.53 \times 10^{-25}$ |
| 7   | 8   | $1.83 \times 10^{-32}$ |

Contribution of  $(K, L)$ -stalls,  
 $p = 4.8 \times 10^{-3}$ .

# Graphical Representation

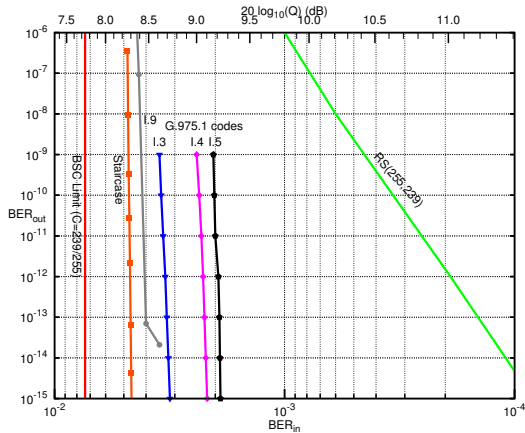




# FPGA-based Simulation Results

## Code Parameters

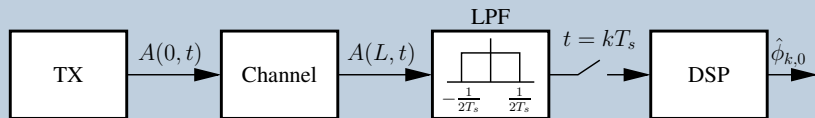
- $m = 510$ ,  $r = 32$ , triple-error-correcting BCH component code



# Capacity Estimation

See: R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, pp. 662–701, Sept./Oct. 2010.

## System Model



# Transmitted Signal

Channel  $l$  signal:

$$X_l(t) = \sum_{k=-\infty}^{\infty} \frac{\phi_{k,l}}{\sqrt{T_s}} \operatorname{sinc} \left( \frac{t - kT_s}{T_s} \right),$$

where  $\operatorname{sinc}(\theta) = \frac{\sin \pi \theta}{\pi \theta}$ .

$\phi_{k,l}$  are elements of a discrete-amplitude continuous-phase input constellation  $\mathcal{M}$ , i.e, for  $N$  rings,  $\theta \in [0, 2\pi)$ , and  $r \geq 0$ ,

$$\mathcal{M} = \{m \cdot r \exp(j\theta) \mid m \in \{1, 2, \dots, N\}\}.$$

Each ring is assumed equiprobable, and for a given ring, the phase distribution is uniform.

# Multi-channel systems

In the general case of a multi-channel system having  $2B + 1$  channels with a channel spacing  $1/T_s$  Hz, the input to the fiber has the form

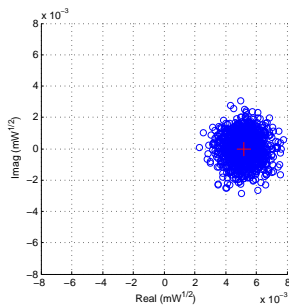
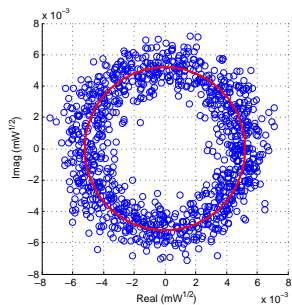
$$A(z = 0, t) = \sum_{k=-\infty}^{\infty} \sum_{l=-B}^B \frac{\phi_{k,l}}{\sqrt{T_s}} \operatorname{sinc} \left( \frac{t - kT_s}{T_s} \right) e^{j2\pi lt/T_s}.$$

# Towards a Probability Model

Back-rotation:

$$\tilde{\phi}_{k,l} = \hat{\phi}_{k,l} \exp(-j(\Phi_{\text{XPM}} + \angle\phi_{k,l})),$$

where  $\Phi_{\text{XPM}}$  is a constant (input-independent) phase rotation contributed by cross-phase modulation (XPM).



## Gaussian fitting

For each  $i$  and a fixed  $l$  (the channel of interest), we calculate the mean  $\mu_i$  and covariance matrix  $\Omega_i$  (of the real and imaginary components) of those  $\tilde{\phi}_{k,l}$  corresponding to the  $i$ -th ring, and model the distribution of those  $\tilde{\phi}_{k,l}$  by  $\mathcal{N}(\mu_i, \Omega_i)$ . From the rotational invariance of the channel, the channel is modeled as

$$f(y|x = r \cdot i \exp(j\phi)) \sim \mathcal{N}(\mu_i \exp(j\phi), \Omega_i),$$

where the (constant) phase rotation due to  $\Phi_{\text{XPM}}$  is ignored, since it can be canceled in the receiver.

# Capacity Estimation

The mutual information of the (assumed) memoryless channel is

$$I(X; Y) = \iint f(x, y) \log_2 \frac{f(y|x)}{f(y)} dx dy,$$

where  $f(x)$  represents the input distribution on  $\mathcal{M}$  with equiprobable rings and a uniform phase distribution, which provides an estimate of the capacity of an optically-routed fiber-optic communication system.

## Signaling Parameters

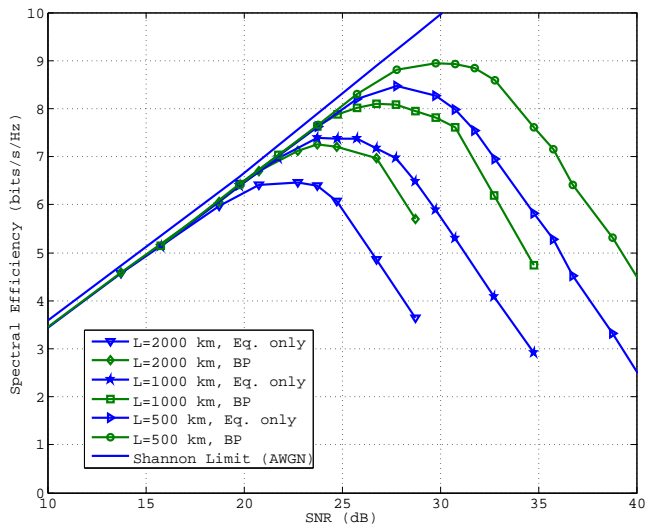
---

|                       |              |
|-----------------------|--------------|
| Baud rate $1/T_s$     | 100 GHz      |
| Channel bandwidth $W$ | 101 GHz      |
| Number of rings $N$   | 64           |
| Number of channels    | $2B + 1 = 5$ |

---



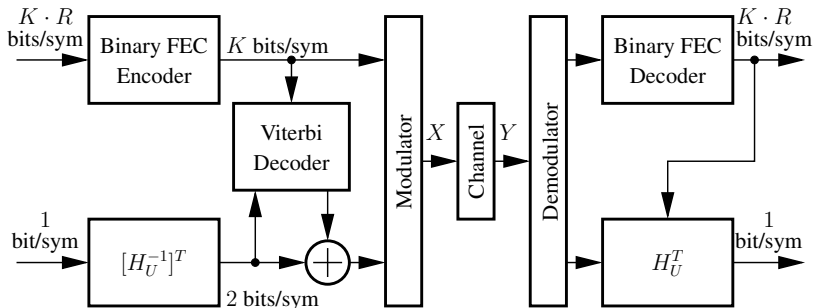
# Achievable Rates from Memoryless Capacity Estimate



(BP adds 0.55 to 0.75 bits/s/Hz relative to EQ)

# Pragmatic Coded Modulation via Staircase Codes

Approach: BICM + shaping, modulation  $2^{K+2}$ -QAM, hard-decisions



Syndrome-former matrix

$$H_U^T = [1 + D + D^2, 1 + D^2]^T.$$

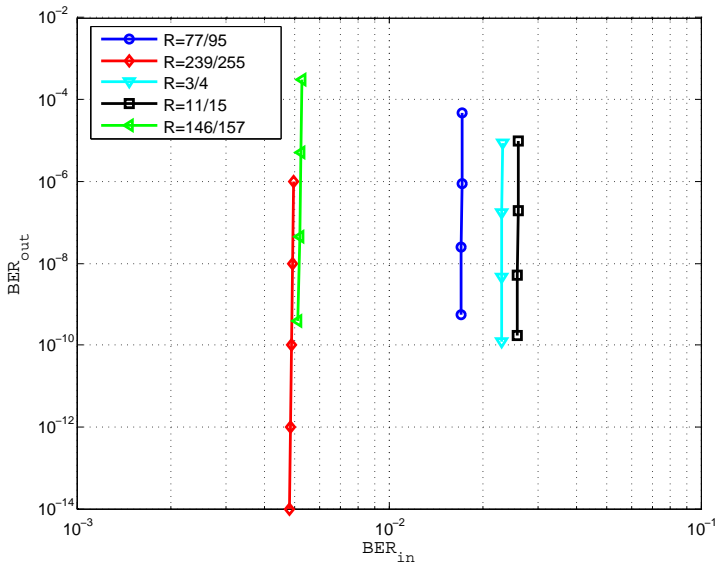
# Achievable Rates

| Fiber System      | $K$ | $p_{\text{avg}}$      | $P_{\text{in}}$<br>(dBm) | $I_P$<br>(bits/s/Hz) |
|-------------------|-----|-----------------------|--------------------------|----------------------|
| $L = 500$ km, EQ  | 8   | $1.61 \times 10^{-2}$ | -6                       | 8.05                 |
| $L = 500$ km, BP  | 8   | $3.52 \times 10^{-3}$ | -4                       | 8.73                 |
| $L = 1000$ km, EQ | 6   | $3.88 \times 10^{-3}$ | -6                       | 6.78                 |
| $L = 1000$ km, BP | 8   | $2.22 \times 10^{-2}$ | -4                       | 7.77                 |
| $L = 2000$ km, EQ | 6   | $2.52 \times 10^{-2}$ | -6                       | 5.98                 |
| $L = 2000$ km, BP | 6   | $5.16 \times 10^{-3}$ | -4                       | 6.72                 |

(These achieve within 0.4 to 0.6 bits/s/Hz of estimated channel capacity.)

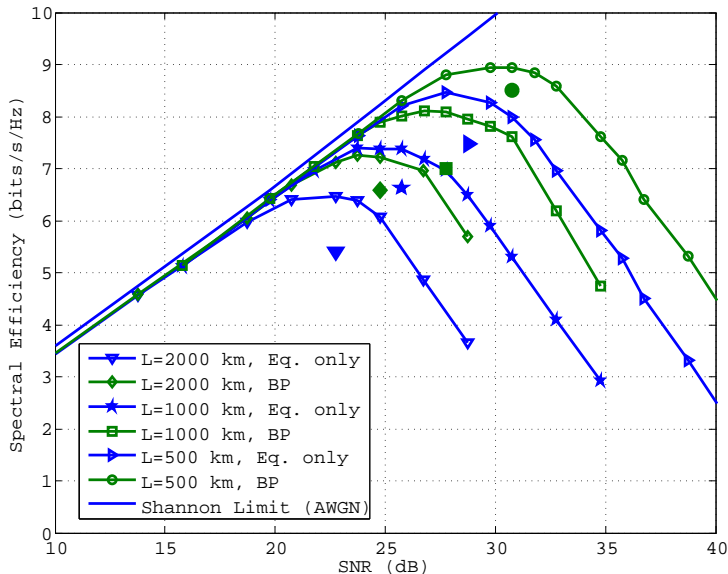
# Staircase code design

First, design a collection of staircase codes of various (appropriate) rates:



# Coded Modulation Performance

Then, simulate their performance on the actual channel:



# Code Parameters

| Fiber System      | $m$ | $t$ | $R$     | Spec. Eff.<br>(bits/s/Hz) |
|-------------------|-----|-----|---------|---------------------------|
| $L = 500$ km, EQ  | 190 | 4   | 77/95   | 7.48                      |
| $L = 500$ km, BP  | 255 | 3   | 239/255 | 8.50                      |
| $L = 1000$ km, EQ | 255 | 3   | 239/255 | 6.62                      |
| $L = 1000$ km, BP | 144 | 4   | 3/4     | 7.00                      |
| $L = 2000$ km, EQ | 120 | 4   | 11/15   | 5.40                      |
| $L = 2000$ km, BP | 628 | 4   | 146/157 | 6.58                      |

### **Central Question:**

Does fiber nonlinearity really place an upper limit on achievable spectral efficiency?

To try to answer this question, we have to understand the nonlinearity more deeply.

### **Spoiler:**

We will not give a satisfactory answer in this talk.

# Evolution Equations

An **evolution equation** (in 1+1 dimensions) is a partial differential equation for an unknown function  $q(t, z)$  of the form

$$\frac{\partial q}{\partial z} = K(q),$$

where  $K(q)$  is an expression involving only  $q$  and its derivatives with respect to  $t$ .

- Heat:  $K(q) = c^2 q_{tt}$
- NLSE:  $K(q) = -j (q_{tt}(t, z) + 2|q(t, z)|^2 q(t, z))$
- KdV:  $K(q) = q_{ttt}(t, z) + q(t, z)q_t(t, z)$

(Here **subscripts** denote partial derivatives, thus  $q_{tt} = \frac{\partial^2}{\partial t^2} q(t, z)$ .)



# Normalized NLS Equation

Changing variables (to so-called “soliton units”):

$$q = \frac{Q}{\sqrt{P}}, \quad z' = \frac{z}{\mathcal{L}}, \quad t' = \frac{t}{T_0},$$

with  $T_0 = \sqrt{\frac{|\beta_2|\mathcal{L}}{2}}$  and  $P = \frac{2}{\gamma\mathcal{L}}$  (and then dropping the “primes”) we get the:

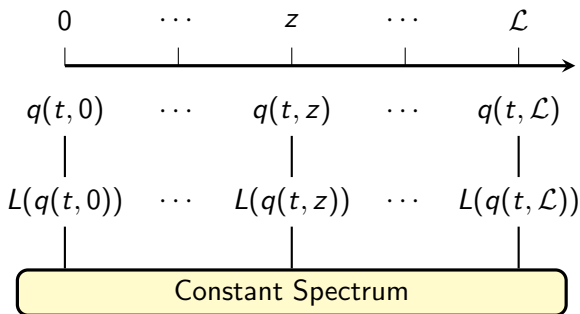
## Normalized NLS equation

$$jq_z(t, z) = q_{tt} + 2|q(t, z)|^2q(t, z) + v(t, z).$$

# Isospectral Flow

## A Key Idea

We seek an **invariant** under evolution (in the absence of noise), e.g., let  $L$  be a linear differential operator (depending on  $q(t, z)$ ). It **may** be possible to find an  $L$  whose (eigenvalue) *spectrum* remains constant, even as  $q$  evolves (in  $z$ ) according to some evolution equation.



## Example: Isospectral Families of Matrices

Let  $L(z)$  be a family of diagonalizable square matrices whose entries are functions of  $z$ . Clearly, the eigenvalues of these matrices in general depend on  $z$ .

For some families (**isospectral families**), it might be the case that while the entries of the matrix change with  $z$ , the eigenvalues remain constant.

Each member of an isospectral family is similar to a constant diagonal matrix  $\Lambda$ , i.e.,

$$L(z) = G(z)\Lambda G^{-1}(z)$$

for some similarity transformation  $G(z)$ .

# Isospectral Families of Operators

Let  $\mathcal{H}$  be a Hilbert space, and let  $L(z)$  be a family of diagonalizable bounded linear operators operating on  $\mathcal{H}$ , e.g.,

$$L(q(t, z)) = -\frac{\partial^2}{\partial t^2} + q(t, z).$$

If the eigenvalues of  $L(z)$  do not depend on  $z$ , then we refer to  $L(z)$  as an **isospectral family** of operators.

For each  $z$ ,  $L(z)$  is similar to a *multiplication operator*  $\Lambda$  (the operator equivalent of a diagonal matrix), i.e.,

$$L(z) = G(z)\Lambda G^{-1}(z),$$

for some operator  $G(z)$ .

The spectrum of an operator is defined as

$$\sigma(L) = \{\lambda \mid L - \lambda I \text{ is not invertible}\}.$$

Spectrum can be discrete (like matrices), continuous, residual, etc.

# The Lax Equation

We have  $L(z) = G(z)\Lambda G^{-1}(z)$ , where  $\Lambda$  does not depend on  $z$ . Assuming that  $L(z)$  varies smoothly with  $z$ , we can form

$$\begin{aligned}\frac{dL(z)}{dz} &= G' \Lambda G^{-1} + G \Lambda (-G^{-1} G' G^{-1}) \\ &= \underbrace{G' G^{-1}}_{M(z)} \underbrace{(G \Lambda G^{-1})}_{L(z)} - \underbrace{(G \Lambda G^{-1})}_{L(z)} \underbrace{G' G^{-1}}_{M(z)} \\ &= M(z)L(z) - L(z)M(z) = [M, L]\end{aligned}\tag{1}$$

where  $[M, L] \triangleq ML - LM$  is the *commutator bracket*.

In other words, every diagonalizable isospectral operator  $L(z)$  satisfies the differential equation (1).

The converse is also true.

# Lax Pairs

## Lemma

Let  $L(z)$  be a diagonalizable family of operators. Then  $L(z)$  is an isospectral family if and only if it satisfies

$$\frac{dL}{dz} = [M, L], \quad (2)$$

for some operator  $M$ , where  $[M, L] = ML - LM$ .

## Definition

The operators  $L$  and  $M$  satisfying (2) are called a *Lax Pair* (after Peter D. Lax, who introduced the concept [1968]).



Recall that  $L$  may depend on a function  $q(t, z)$  and so can  $M$ . The commutator bracket  $[M, L]$  can create nonlinear evolution equations for  $q(t, z)$  in the form

$$L_z = [L, M] \quad \Leftrightarrow \quad \frac{\partial q}{\partial z} = K(q),$$

where  $K(q)$  is some, in general nonlinear, function of  $q(t, z)$  and its time derivatives.

# KdV Equation

An example of this is the Korteweg-de Vries (KdV) equation (which arises in the evolution of waves on shallow water surfaces). Let  $q(t, z)$  be a real-valued function and choose

$$L = \frac{\partial^2}{\partial t^2} + \frac{1}{3}q,$$
$$M = 4\frac{\partial^3}{\partial t^3} + q_t + q\frac{\partial}{\partial t}.$$

The Lax equation  $L_z = [M, L]$  is easily simplified to

$$\frac{1}{3}q_z - \frac{1}{3}(q_{ttt} + qq_t) + (\text{some terms})\frac{\partial}{\partial t} \equiv \mathbf{0},$$

where  $\mathbf{0}$  is the zero operator. The zero-order term of this equation, which must be zero, produces the KdV equation  $q_z = q_{ttt} + qq_t$ .



## NLS equation

If  $q(t, z)$  varies based on the NLS equation

$$jq_z = q_{tt} + 2|q|^2q$$

then the spectrum of

$$L = j \begin{pmatrix} \frac{\partial}{\partial t} & -q(t, z) \\ -q^*(t, z) & -\frac{\partial}{\partial t} \end{pmatrix} \quad (3)$$

is preserved during the evolution (Zakharov-Shabat 1972).

In addition, the operator  $M$  is given by

$$M = \begin{pmatrix} 2j\lambda^2 - j|q(t, z)|^2 & -2\lambda q(t, z) - jq_t(t, z) \\ 2\lambda q^*(t, z) - jq_t^*(t, z) & -2j\lambda^2 + j|q(t, z)|^2 \end{pmatrix}$$

**Thus the NLS equation is indeed generated by a Lax pair!**

# Eigenvalues and Eigenvectors of $L$

The eigenvalues of the operator  $L$ , which are constant in an isospectral flow, are defined via

$$Lv = \lambda v \quad (4)$$

Taking the  $z$  derivative of (4) and using the Lax equation  $L_z = [M, L]$ , we can show that an *eigenvector* of  $L$  evolves according to the linear equation  $v_z = Mv$ .

Furthermore, we can re-write (4) as  $v_t = Pv$  for some operator  $P$ . Thus we get

## Evolution of Eigenvectors of $L$

The *eigenvectors* of  $L$  evolve according to the linear system

$$v_z = Mv \quad (5)$$

$$v_t = Pv \quad (6)$$

## Zero-Curvature Condition

Combining equations (5) and (6) by using the equality of mixed derivatives, i.e.,  $v_{tz} = v_{zt}$ , the Lax equation (2) is reduced to the

### Zero-curvature Condition

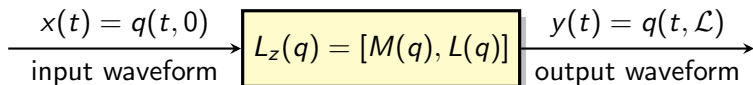
$$P_z - M_t + [P, M] = 0. \quad (7)$$

Note that the nonlinear equation derived from (7) results as a compatibility condition between the two linear equations (5) and (6). Thus certain nonlinear evolution equations possess a certain “hidden linearity.”

We can work with the Lax pair  $(L, M)$  or equivalently with the pair  $(P, M)$ .

# Integrable Systems

We refer to a system described by a Lax pair  $M, L$ , measured at distance  $\mathcal{L}$ , as the *integrable system*  $S = (L, M; \mathcal{L})$ .



## Definition (Lax Convolution)

The action of an integrable system  $S = (L, M; \mathcal{L})$  on the input  $q(t, 0)$  is called the *Lax convolution* of  $q$  with  $S$ . We write the system output as  $q(t, \mathcal{L}) = q(t, 0) * (L, M; \mathcal{L})$ .

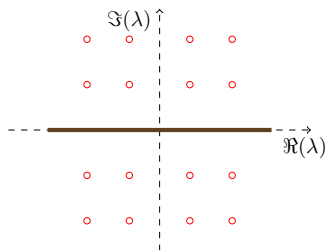
## Definition (Integrable communication channels)

A waveform communication channel  $C : x(t) \times v(t, z) \rightarrow y(t)$  with inputs  $x(t) \in L^1(\mathbb{R})$  and space-time noise  $v(t, z) \in L^2(\mathbb{R}, \mathbb{R}^+)$ , and output  $y(t) \in L^1(\mathbb{R})$ , is said to be *integrable* if the noise-free channel is an integrable system.

# Spectrum of $L$

The Zakharov-Shabat operator for the NLS equation has two types of spectra:

- the **discrete** (or point) spectrum, which occurs in  $\mathbb{C}^+$  and corresponds to solitons.
- The **continuous** spectrum, which in general includes the whole real line, corresponds to the non-solitonic (or radiation) component of the signal.



## Eigenspace Associated with $\lambda$

Associated with each point  $\lambda$  in the spectrum of  $L$  is a two-dimensional eigenspace in which eigenvectors  $v$  evolve in time according to

$$v_t = Pv = \begin{pmatrix} -j\lambda & q(t, z) \\ -q^*(t, z) & j\lambda \end{pmatrix} v.$$

Assuming  $q(t, z)$  is “pulse-like,” decaying to zero as  $|t| \rightarrow \infty$ , the  $P$  operator approaches a diagonal matrix, and the eigenvectors approach

$$v(t, \lambda) \rightarrow (\alpha e^{-j\lambda t}, \beta e^{j\lambda t})^T, \quad \alpha, \beta \in \mathbb{C}.$$

# Canonical Eigenvectors

For example, the “canonical” eigenvectors at  $t \rightarrow -\infty$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda t}, \quad \begin{pmatrix} -1 \\ 0 \end{pmatrix} e^{-j\lambda^* t}$$

are bounded for  $\lambda$  in the upper (lower) half-complex plane respectively.

We can propagate these vectors forward in time according to  $v_t = P v$ , let them “interact” with  $q(t, z)$ , and measure what is “scattered” at  $t = +\infty$  in the basis

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{j\lambda t}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{j\lambda^* t}$$

The corresponding coefficients  $a(\lambda)$  and  $b(\lambda)$  are called the **nonlinear Fourier coefficients**.

# Continuous and Discrete Spectra

It can be shown that these scattering projections are well defined

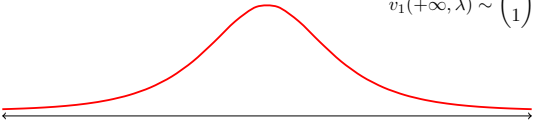
- at *all* points  $\lambda \in \mathbb{R}$ , the real line  $\Rightarrow$  **continuous spectrum**
- at those points  $\lambda \in \mathbb{C}^+$ , the upper-half complex plane, where  $a(\lambda) = 0$  (necessarily isolated points due to analyticity of  $a(\lambda)$ )  $\Rightarrow$  **discrete spectrum**



# Computing the Fourier coefficients

$$v_t = \begin{pmatrix} -j\lambda & q(t) \\ -q^*(t) & j\lambda \end{pmatrix} v$$

$$v_2(-\infty, \lambda) \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda t} \quad \rightsquigarrow \quad v_2(+\infty, \lambda)$$

$$v_1(+\infty, \lambda) \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{j\lambda t}$$


$$\tilde{v}_1(+\infty, \lambda^*) \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda t}$$

Since the nonlinear Fourier coefficients are time independent, simply compute them at  $t = +\infty$ , by projecting  $v^2(+\infty, \lambda)$  and  $\tilde{v}^2(+\infty, \lambda)$  onto the basis  $v^1(+\infty, \lambda)$  and  $\tilde{v}^1(+\infty, \lambda)$  to obtain

$$[v^2(+\infty, \lambda), \tilde{v}^2(+\infty, \lambda)] = [\tilde{v}^1(+\infty, \lambda), v^1(+\infty, \lambda)] S,$$

where

$$S = \begin{pmatrix} a(\lambda) & b^*(\lambda^*) \\ b(\lambda) & -a^*(\lambda^*) \end{pmatrix}.$$

# Nonlinear Fourier Transform

For the purpose of developing the inverse transform, it is sufficient to work with the ratios

$$\hat{q}(\lambda) = \frac{b(\lambda)}{a(\lambda)} \text{ and } \tilde{q}(\lambda_j) = \frac{b(\lambda_j)}{a'(\lambda_j)}.$$

# Nonlinear Fourier Transform

## Definition (Nonlinear Fourier transform)

Let  $q(t)$  be a sufficiently smooth function in  $L_1(\mathbb{R})$ . The nonlinear Fourier transform of  $q(t)$  with respect to a given Lax operator  $L$  consists of the continuous and discrete spectral functions  $\hat{q}(\lambda) : \mathbb{R} \mapsto \mathbb{C}$  and  $\tilde{q}(\lambda_j) : \mathbb{C}^+ \mapsto \mathbb{C}$  where

$$\hat{q}(\lambda) = \frac{b(\lambda)}{a(\lambda)}, \quad \tilde{q}(\lambda_j) = \frac{b(\lambda_j)}{a'(\lambda_j)}, j = 1, 2, \dots, N,$$

in which  $\lambda_j$  are the zeros of  $a(\lambda)$ . Here, the spectral coefficients  $a(\lambda)$  and  $b(\lambda)$  are given by

$$\begin{aligned} a(\lambda) &= \lim_{t \rightarrow \infty} v_1^2 e^{j\lambda t}, \\ b(\lambda) &= \lim_{t \rightarrow \infty} v_2^2 e^{-j\lambda t} \end{aligned}$$

# The (ordinary) Fourier Transform

- The nonlinear Fourier transform measures, for each fixed eigenvalue  $\lambda$ , the interaction of a propagating eigenvector (initialized as a complex exponential) with a waveform  $q(t)$ .
- The ordinary Fourier transform measures, for each fixed frequency  $\omega$ , the interaction of the complex exponential  $e^{j\omega t}$  with a waveform  $q(t)$  via the inner product

$$Q(\omega) = \langle q(t), e^{j\omega t} \rangle = \int_{-\infty}^{\infty} q(t) e^{-j\omega t} dt.$$

- An “eigenvector evolution” version of the ordinary Fourier transform is recovered by defining

$$v_t = \begin{pmatrix} 0 & q(t) \\ 0 & -j\omega \end{pmatrix} v,$$

giving

$$v = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \int_{-\infty}^t q(\tau) e^{-j\omega\tau} d\tau \\ e^{-j\omega t} \end{pmatrix}$$

## Example: NFT of a Rectangular Pulse

Consider the rectangular pulse

$$q(t) = \begin{cases} A, & t \in [t_1, t_2]; \\ 0, & \text{otherwise.} \end{cases}$$

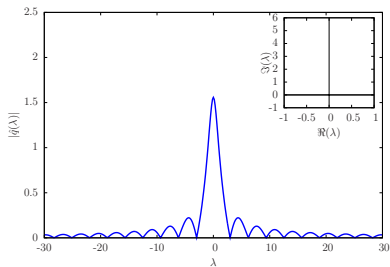
Let  $T = t_2 - t_1$  and  $T' = t_2 + t_1$ , and let  $\Delta = \sqrt{\lambda^2 + |A|^2}$ . After some work, we find that the zeros of  $a(\lambda)$  in  $\mathbb{C}^+$ , satisfy

$$j \tan(T \sqrt{|A|^2 + \lambda^2}) = \sqrt{1 + \frac{|A|^2}{\lambda^2}},$$

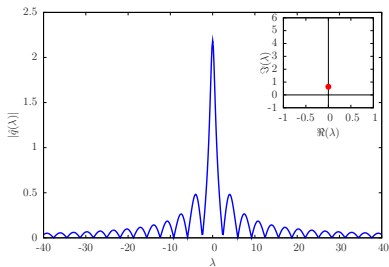
giving rise to the discrete spectrum. The continuous spectrum is given by

$$\hat{q}(\lambda) = \frac{A^*}{j\lambda} e^{-2j\lambda t_2} \left(1 - \frac{\Delta}{j\lambda} \cot(\Delta T)\right)^{-1} \rightarrow -A^* T e^{-\frac{t_1+t_2}{2} \lambda} \text{sinc}(2Tf).$$

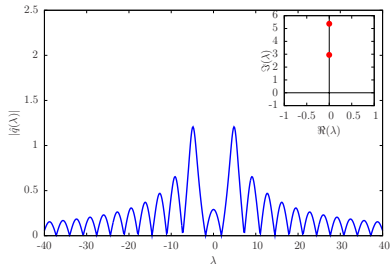
# Discrete and Continuous Spectra of Rectangular Pulse



$A = 1$



$A = 2$



# NFT Properties

Let  $q(t) \leftrightarrow (\widehat{q}(\lambda), \widetilde{q}(\lambda_k))$  be a nonlinear Fourier transform pair.

- (The ordinary Fourier transform as limit of the nonlinear Fourier transform): If  $\|q\|_{L_1} \ll 1$ , there is no discrete spectrum and  $\widehat{q}(\lambda) \rightarrow Q(\lambda)$ , where  $Q(\lambda)$  is the ordinary (linear) Fourier transform of  $-q^*(t)$

$$Q(\lambda) = - \int_{-\infty}^{\infty} q^*(t) e^{-2j\lambda t} dt.$$

- (Weak nonlinearity): If  $|a| \ll 1$ , then  $\widehat{aq}(\lambda) \approx a\widehat{q}(\lambda)$  and  $\widetilde{aq}(\lambda_k) \approx a\widetilde{q}(\lambda_k)$ . In general, however,  $\widehat{aq}(\lambda) \neq a\widehat{q}(\lambda)$  and  $\widetilde{aq}(\lambda_k) \neq a\widetilde{q}(\lambda_k)$ .
- (Constant phase change):  $\widehat{e^{j\phi}q(t)}(\lambda) = e^{j\phi}\widehat{q(t)}(\lambda)$  and  $\widetilde{e^{j\phi}q(t)}(\lambda_k) = e^{j\phi}\widetilde{q(t)}(\lambda_k)$ .

## Fourier Transform Properties (cont'd)

- (Time dilation):  $\widehat{q(\frac{t}{a})} = |\widehat{a}| \widehat{q(a\lambda)}$  and  $\widetilde{q(\frac{t}{a})} = |\widetilde{a}| \widetilde{q(a\lambda_k)}$ ;
- (Time shift):  $q(t - t_0) \leftrightarrow e^{-2j\lambda t_0} (\widehat{q}(\lambda), \widetilde{q}(\lambda_k))$ ;
- (Frequency shift):  $q(t)e^{-2j\omega t} \leftrightarrow (\widehat{q}(\lambda - \omega), \widetilde{q}(\lambda_k - \omega))$ ;
- (Parseval identity):  $\int_{-\infty}^{\infty} \|q(t)\|^2 dt = \widehat{E} + \widetilde{E}$ , where

$$\widehat{E} = \frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 + |\widehat{q}(\lambda)|^2) d\lambda, \quad \widetilde{E} = 4 \sum_{j=1}^N \Im(\lambda_j).$$

The quantities  $\widehat{E}$  and  $\widetilde{E}$  represent the energy contained in the continuous and discrete spectra, respectively.



# The Key NFT Property

- (Lax convolution): If  $q_2(t) = q_1(t) * (L, M; \mathcal{L})$ , then  $\hat{q}_2(\lambda) = c(\lambda, \mathcal{L})\hat{q}_1(\lambda)$  and  $\tilde{q}_2(\lambda_k) = c(\lambda, \mathcal{L})\tilde{q}_1(\lambda_k)$ .

For the NLS equation,  $c(\lambda, \mathcal{L}) = \exp(-4j\lambda^2\mathcal{L})$ .

## LTI systems:

$$y(t) = h(t) * x(t) + z(t) \Leftrightarrow Y(\omega) = H(\omega)X(\omega) + Z(\omega)$$

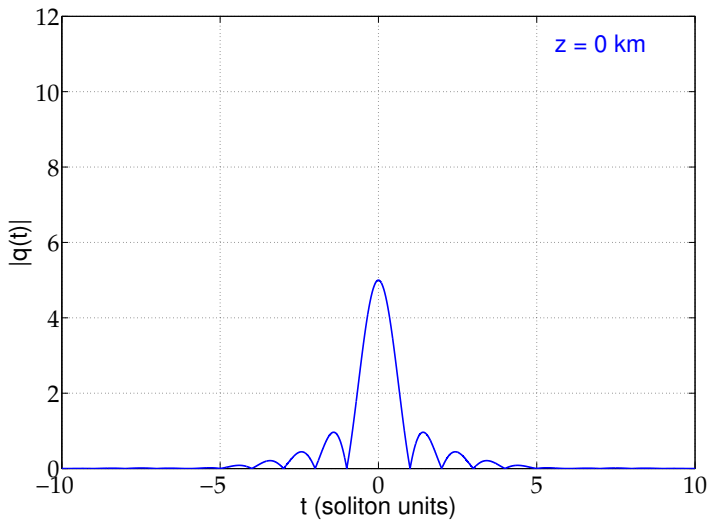
## Nonlinear Integrable systems:

$$y(t) = x(t) * (L, M; \mathcal{L}) + z(t) \Leftrightarrow \begin{cases} \hat{Y}(\lambda) = H(\lambda)\hat{X}(\lambda) + \hat{Z}(\lambda), \\ \tilde{Y}(\lambda_j) = H(\lambda_j)\tilde{X}(\lambda_j) + \tilde{Z}(\lambda_j) \end{cases}$$

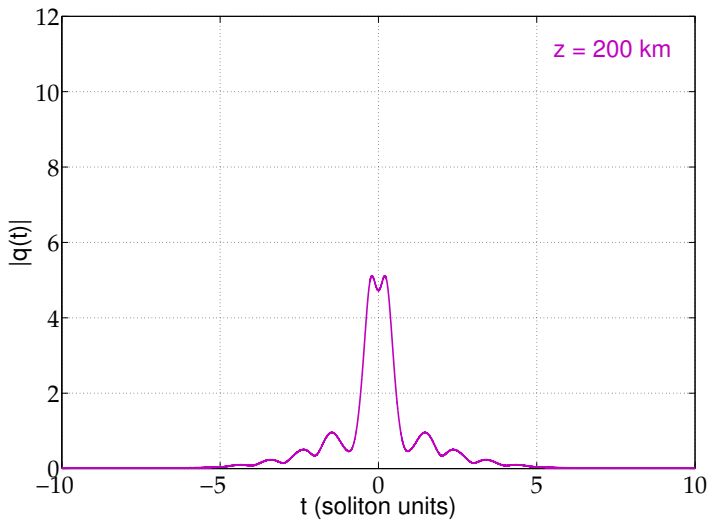
where the channel filter is  $H(\lambda) = e^{-4j\lambda^2 z}$ .

We see that the operation of the Lax convolution in the nonlinear Fourier domain is described by a simple multiplicative (diagonal) operator (i.e., filters), much in the same way that the ordinary Fourier transform maps  $x(t) * y(t)$  to  $X(\omega) \cdot Y(\omega)$ .

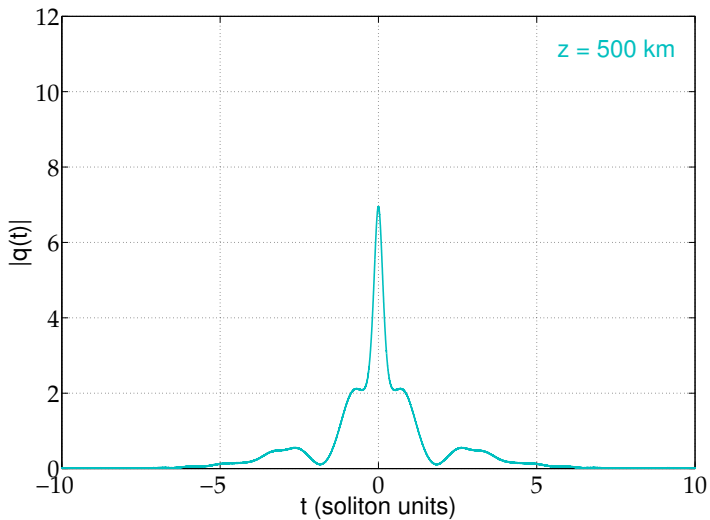
## Launching a Pulse (revisited)



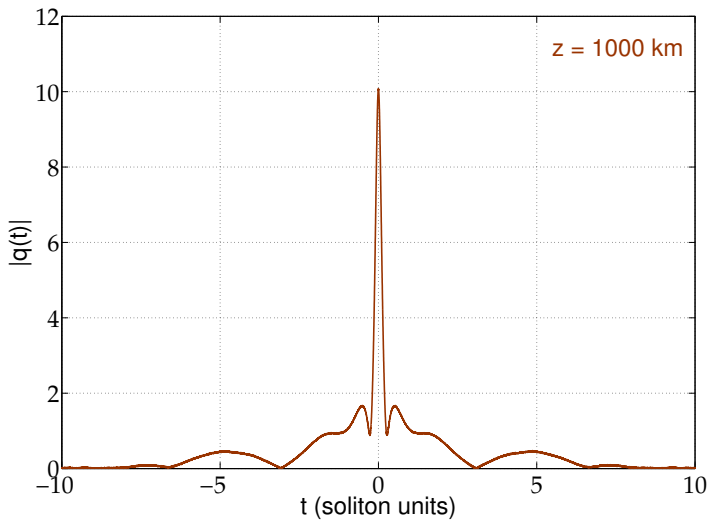
## Launching a Pulse (revisited)



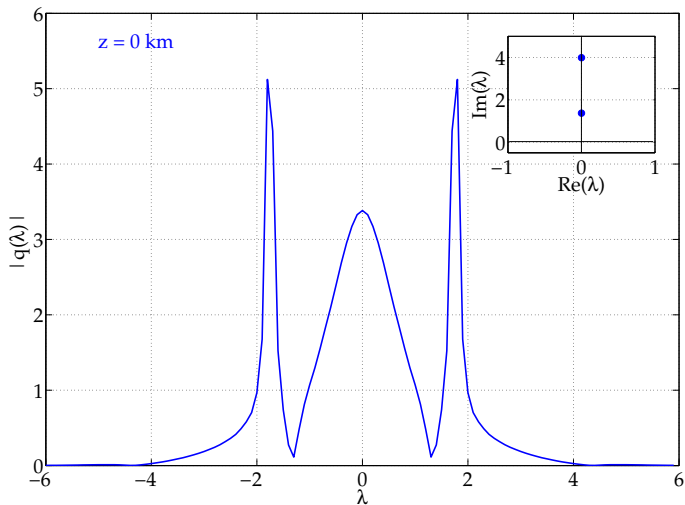
## Launching a Pulse (revisited)



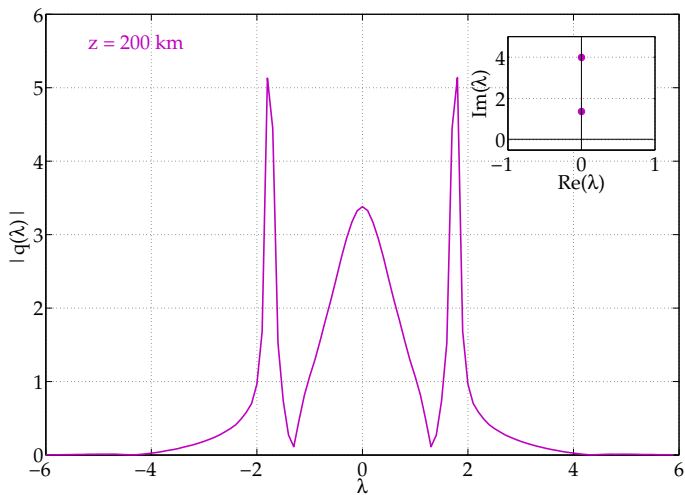
## Launching a Pulse (revisited)



# Launching a Pulse (revisited)

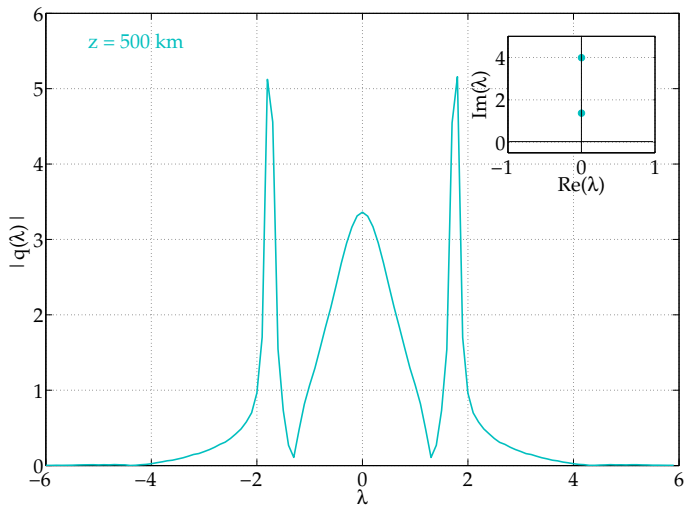


# Launching a Pulse (revisited)

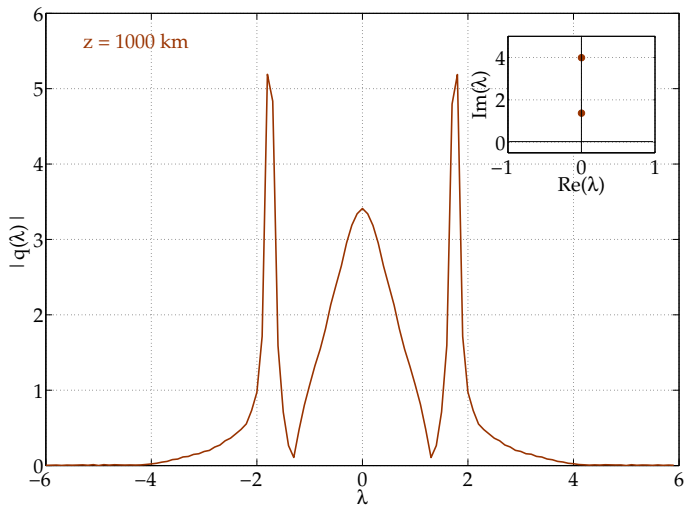




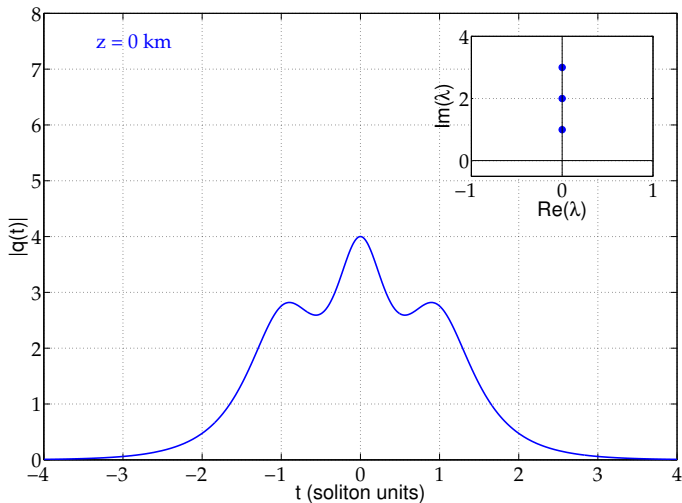
# Launching a Pulse (revisited)



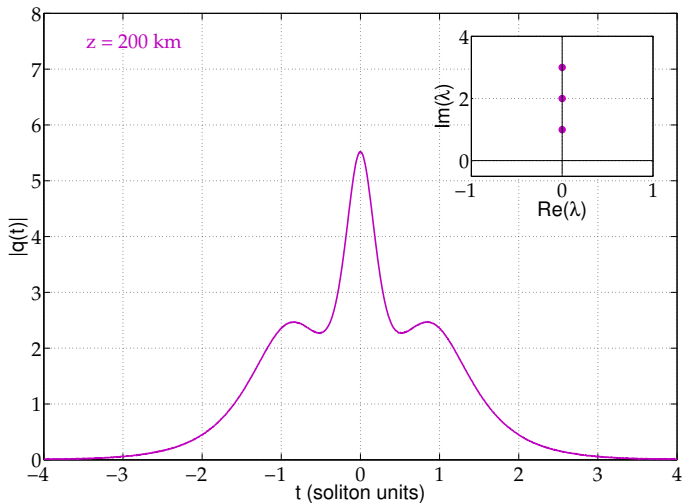
# Launching a Pulse (revisited)



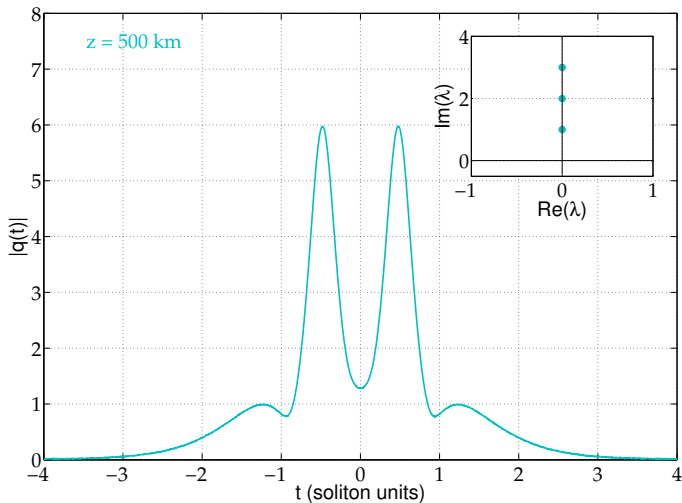
# Launching a Pulse (revisited)



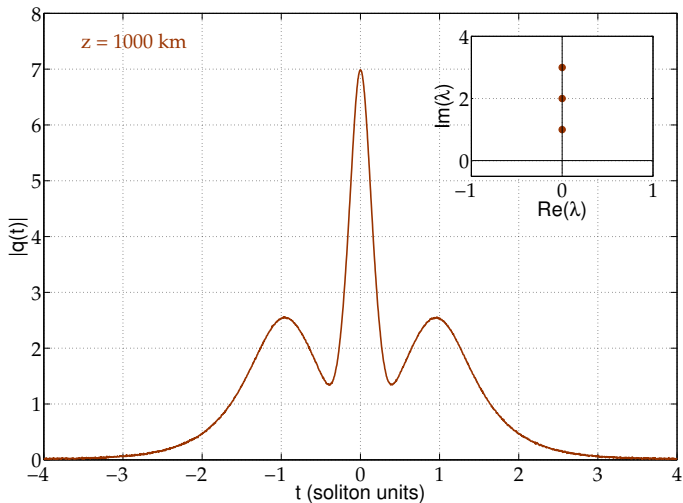
# Launching a Pulse (revisited)



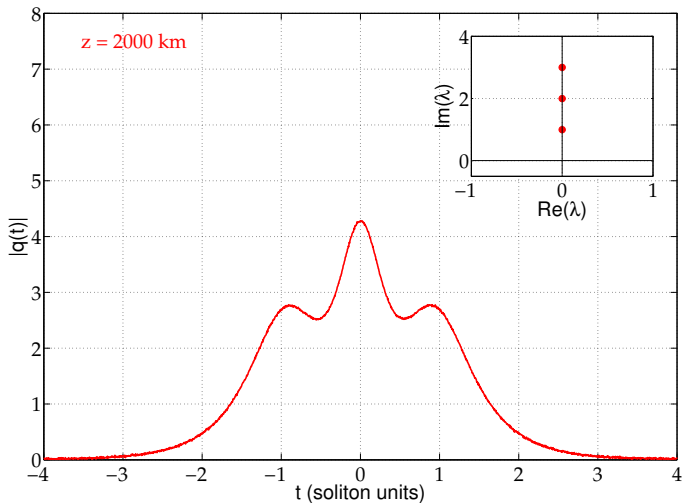
# Launching a Pulse (revisited)



# Launching a Pulse (revisited)

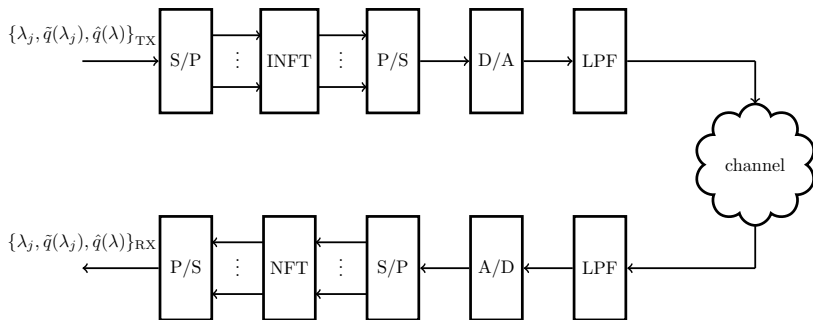


# Launching a Pulse (revisited)



# Nonlinear FDM

Immediate application: encode data in the nonlinear spectrum!

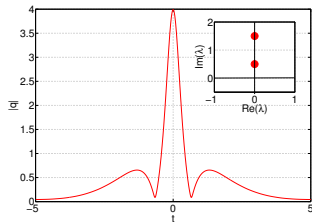
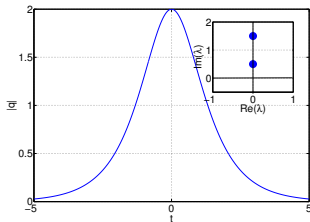


See also: A. Hasegawa and T. Nyu, "Eigenvalue Communication," *J. of Lightwave Technology*, vol. 11, pp. 395–399, March 1993.



# Multisoliton Communication

Multisolitons are signals whose NFT spectrum only contains the discrete component, i.e., the corresponding  $L$  operator has only discrete eigenvalues



$2\text{sech}(t)$  — Satsuma-Yajima 2-soliton

# Generation of Multisoliton Signals

- Inverse NFT problem: Given a set of points in  $\mathbb{C}^+$ , compute a signal  $q(t)$  such that the eigenvalues of its corresponding  $L$  operator are exactly these points
- Several approaches: Riemann–Hilbert system of linear equations, Hirota bilinearization scheme, Darboux transformation

## Darboux method (recursive construction)

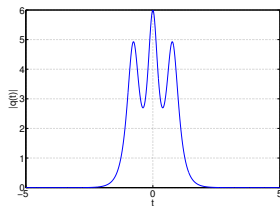
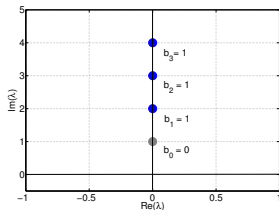
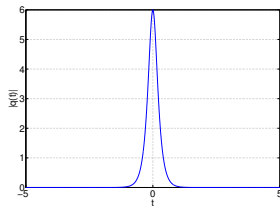
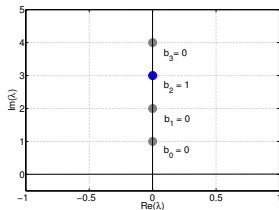
- Starting with a multisoliton signal  $q(t, \lambda_1, \lambda_2, \dots, \lambda_k)$ , and a solution,  $\phi(t, \lambda_{k+1})$ , of

$$v_t(t, \lambda) = \begin{pmatrix} -j\lambda & q(t, z) \\ -q^*(t, z) & j\lambda \end{pmatrix} v(t, \lambda),$$

$$\text{compute } \tilde{q} = q - 2j(\lambda_{k+1}^* - \lambda_{k+1}) \frac{\phi_2^* \phi_1}{|\phi_1|^2 + |\phi_2|^2}.$$

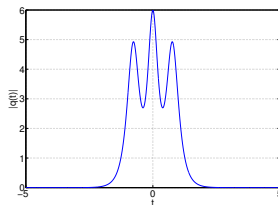
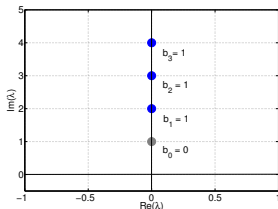
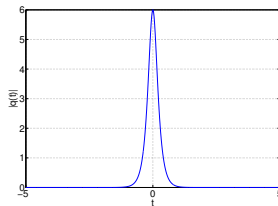
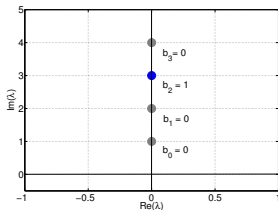
- $\tilde{q}$  has eigenvalues of  $q$  and new eigenvalue from  $\phi$

# On-Off Eigenvalue Encoder (OOK)



OOK signals for input bits  $[0\ 0\ 1\ 0]$  and  $[0\ 1\ 1\ 1]$

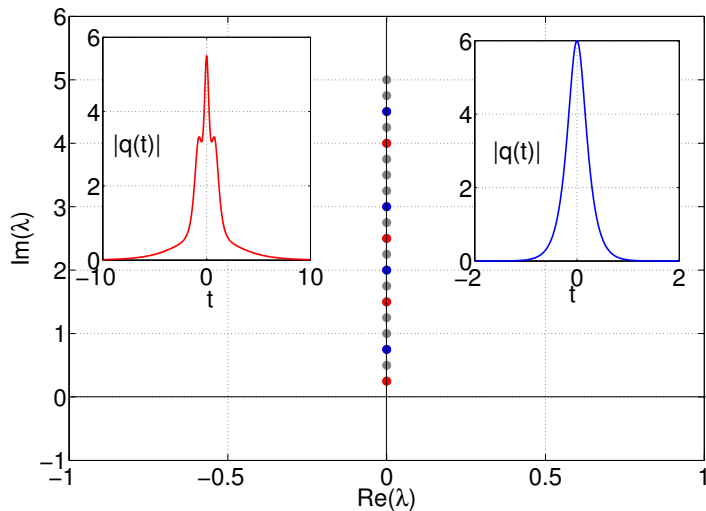
# On-Off Eigenvalue Encoder (OOK)



OOK signals for input bits  $[0\ 0\ 1\ 0]$  and  $[0\ 1\ 1\ 1]$

Problem: Spectral efficiency drops as the number of points is increased

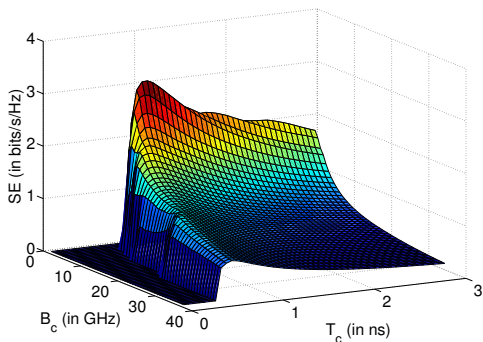
# Multi-eigenvalue position encoding



Multi-eigenvalue position encoding

# Multi-eigenvalue position encoding

- Generate  $\binom{N}{1} + \binom{N}{2} + \dots + \binom{N}{k}$  signals for a given grid of  $N$  points, and at most  $k$  eigenvalues in a transmit signal
- Select cut-off bandwidth and pulsewidth



Spectral efficiency as a function of the cutoff parameters ( $N=50$ ,  
 $\Delta = 0.1j$ ,  $k = 5$ )

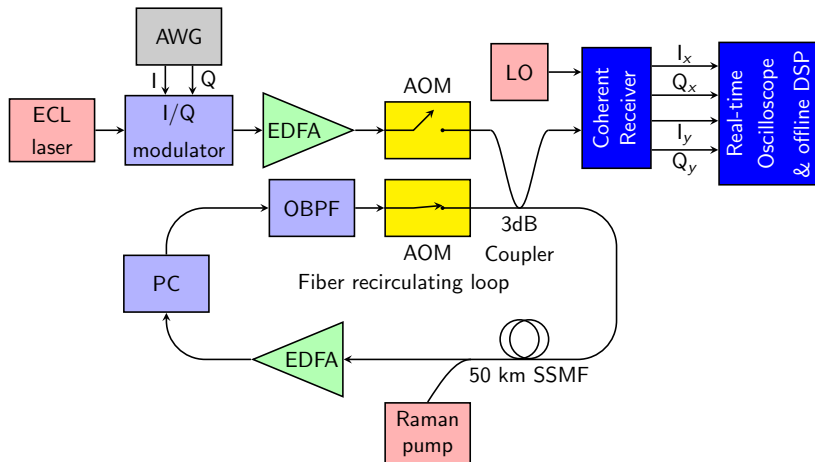
# Multi-eigenvalue position encoding

## Achievable spectral efficiencies

| $k$ | $\tilde{M}$ | $M$   | $T_c$   | $B_c$    | SE (in bits/s/Hz) |
|-----|-------------|-------|---------|----------|-------------------|
| 3   | 20,876      | 201   | 0.88 ns | 6.54 GHz | 1.34              |
| 4   | 251,176     | 804   | 0.88 ns | 5.08 GHz | 2.17              |
| 5   | 2,369,936   | 2,569 | 1 ns    | 3.64 GHz | 3.14              |

**Table :** Spectral efficiencies of the multi-eigenvalue position encoder ( $N = 50$  and  $\Delta = 0.1$ )

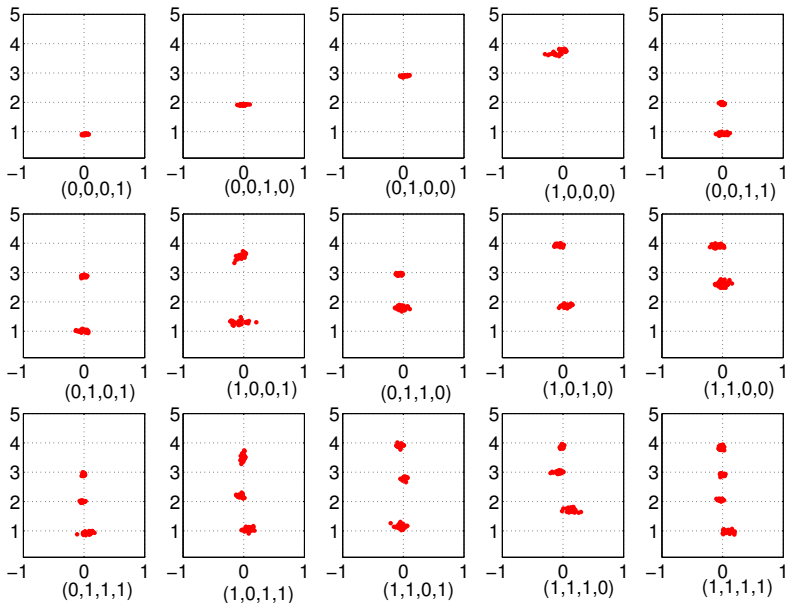
# Experimental Demonstration



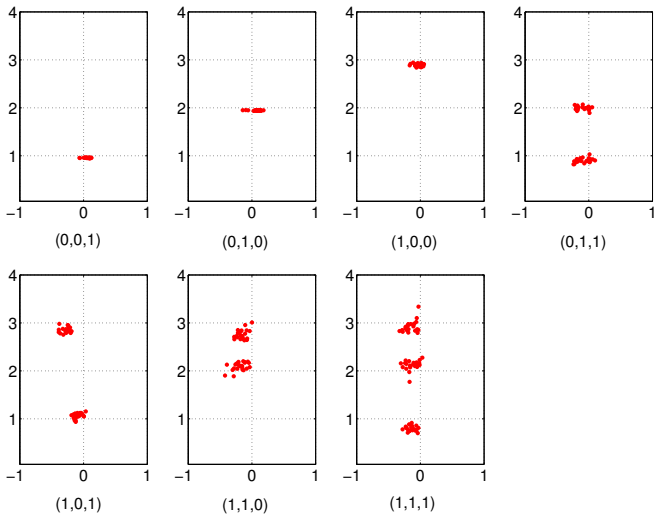
Z. Dong, S. Hari, T. Gui, K. Zhong, M. I. Yousefi, C. Lu, P.-K. A. Wai, F. R. Kschischang, and A. P.-T. Lau, "Nonlinear Frequency Division Multiplexed Transmissions based on NFT," *IEEE Photonics Techn. Letters*, vol. 27, no. 15, pp. 1621–1623, Aug. 2015.



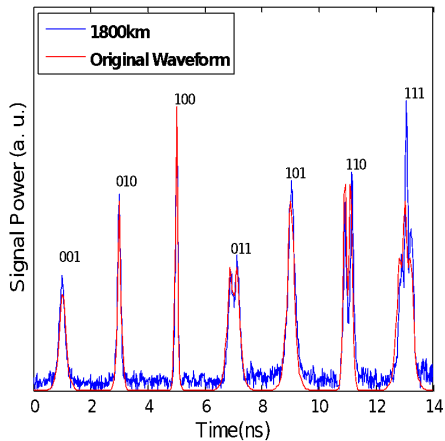
# 4 eigenvalues (300km)



# 3 eigenvalues (1800km)



# Received Signal (1800km)



### **Central Question:**

Does fiber nonlinearity really place an upper limit on achievable spectral efficiency?

To try to answer this question, we have to understand the nonlinearity more deeply.

### **Spoiler:**

We will not give a satisfactory answer in this talk, however . . .

# Conclusion

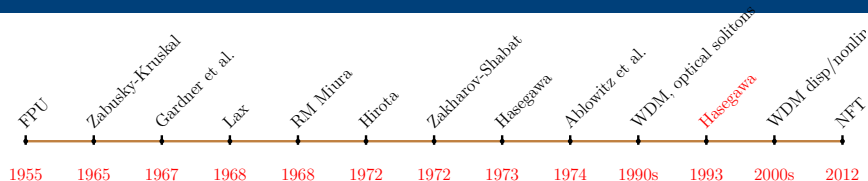
Pulse train transmission and WDM may not be the best transmission strategy for the nonlinear optical channel.

All deterministic distortions are zero for all users under NFDM:

- NFDM removes inter-channel interference (cross-talk) between users of a network sharing the same fiber channel;
- NFDM removes inter-symbol interference (ISI) (intra-channel interactions) for each user;
- with NFDM, information in each channel of interest can be conveniently read anywhere in a network without knowledge of the distance or any information about other users;
- spectral invariants are remarkably stable features of the NLS flow.

Exploiting the integrability of the NLS equation, NFDM modulates non-interacting degrees of freedom and thereby does not suffer from cross-talk, a major limiting factor for prior work.

# Conclusions



- The NFT can be used for transmission of information
- NFT solves at least two major problems
  - ① it removes ISI
  - ② spectral data appear to be more robust to noise compared with time-domain processing (e.g., digital backpropagation)
- **Is it better than existing methods?** We do not know yet. So far we can match existing methods. However, our insights suggest that the achievable rates for NFDM will trend upwards for increasing SNRs.
- disadvantages:
  - ① critically relies on integrability
  - ② it is hard to implement or sometimes to simulate.

Plenty of open problems:

- Entropy (rate) evolution (production) in NLS equation with additive noise
- Rigorous bounds on capacity
- Vector models (for polarization, multimode, multicore)
- Discrete models
- Development of fast algorithms
- ...

# References

- 1 A. Hasegawa and T. Nyu, "Eigenvalue Communication," *J. of Lightwave Technology*, vol. 11, pp. 395–399, March 1993.
- 2 R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, pp. 662–701, Sept./Oct. 2010.
- 3 M. I. Yousefi and F. R. Kschischang, "Information Transmission using the Nonlinear Fourier Transform, Part I: Mathematical Tools," "Part II: Numerical Methods," "Part III: Discrete Spectrum Modulation," *IEEE Trans. on Inform. Theory*, July 2014.
- 4 E. Meron, M. Feder and M. Shtaiif, "On the achievable communication rates of generalized soliton transmission systems," arXiv:1207.02, 2012.
- 5 P. Kazakopoulos and A. L. Moustakas, "Transmission of information via the non-linear Schrödinger equation: The random Gaussian input case," ArXiv:1210.794, 2013.
- 6 S. Wahls and H. V. Poor, "Fast numerical nonlinear Fourier transforms," to appear in *IT Transactions*, 2015.



## References (cont'd)

- 7 H. Bülow, "Experimental assessment of nonlinear Fourier transform based detection under fiber nonlinearity," ECOC 2014.
- 8 J. E. Prilepsky, et al., "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," PRL, Jul. 2014.
- 9 H. Terauchi and A. Manuta, "Eigenvalue modulated optical transmission system based on digital coherent technology," OECC, 2013.
- 10 H. Terauchi, Y. Matsuda, A. Toyota, and A. Maruta, "Noise tolerance of eigenvalue modulated optical transmission system based on digital coherent technology," OECC/ACOFT, 2014.
- 11 S. Wahls, S. T. Le, J. E. Prilepsky, H. V. Poor, S. K. Turitsyn, "Digital backpropagation in the nonlinear Fourier domain," SPAWC, Stockholm, Sweden, Jun. 2015.
- 12 V. Aref, H. Bülow, K. Schuh, W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," ECOC 2015, Sept. 2015.

## 2006 Steele Prize

For their discovery in the 1970s of the mathematical framework underlying the nonlinear Fourier transform, Clifford S. Gardner, John M. Greene, Martin D. Kruskal and Robert M. Miura received the prestigious 2006 Leroy P. Steele Prize for a Seminal Contribution to Research, awarded by the American Mathematical Society.

### **From the Steele Prize Announcement:**

Nonlinearity has undergone a revolution: from a nuisance to be eliminated, to a new tool to be exploited.

## No Time To Discuss:

- polarization (Manakov system)
- few-mode, multi-mode, multi-core fibers
- interference models (Extended GN, Kolmogorov-Zhakarov)
- shaping
- challenges in coded modulation design (decoding energy!)
- other approaches to nonlinearity mitigation (e.g., optical phase conjugation)

# Acknowledgment

This work was supported:

- by the **Natural Sciences and Engineering Research Council of Canada** through the Discovery Grants Program, and
- by the **Institute for Advanced Study, Technische Universität München**, through a Hans Fischer Senior Fellowship.

I am profoundly grateful for all support.

# Acknowledgment

This work was supported:

- by the **Natural Sciences and Engineering Research Council of Canada** through the Discovery Grants Program, and
- by the **Institute for Advanced Study, Technische Universität München**, through a Hans Fischer Senior Fellowship.

I am profoundly grateful for all support.

**Thank You!**