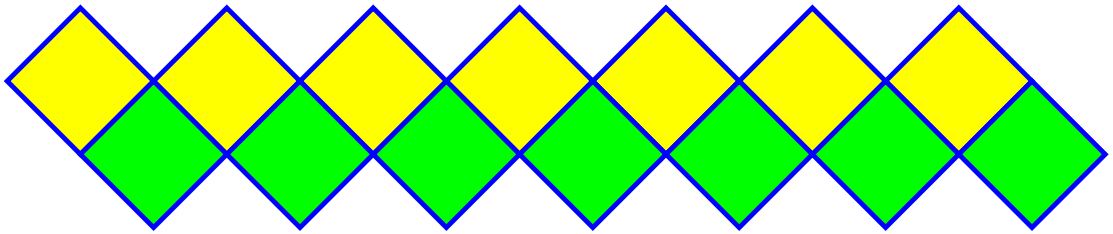
Staircase Codes

Error-correction for High-Speed Fiber-Optic Channels

Frank R. Kschischang
Dept. of Electrical & Computer Engineering
University of Toronto



Talk at Delft University of Technology, Delft, The Netherlands

April 23rd, 2013

Acknowledgements

Joint work with:

Benjamin P. Smith
University of Toronto
Andrew Hunt and **John Lodge**
Communications Research Centre, Ottawa
Arash Farhood
Cortina Systems Inc., Sunnyvale/Ottawa

Thank you to Drs. Nader Alagha and Fernando Kuipers for the invitation!

Fiber-Optic Communication Systems

Physics: Enabling Technologies

1. Low-loss optical fiber (∼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

Fiber-Optic Communication Systems: Challenges

Reliability

$P_e < 10^{-15}$

Speed

100 Gb/s per-channel data rates

Non-Linearity

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

Outline of Talk

Part I

Binary Error-Correcting Codes for High-Speed Communications

- Syndrome-based interative decoding
- Performance optimization for product-like codes
- Staircase codes
- FPGA-based simulation results
- Analytical error-floor predictions

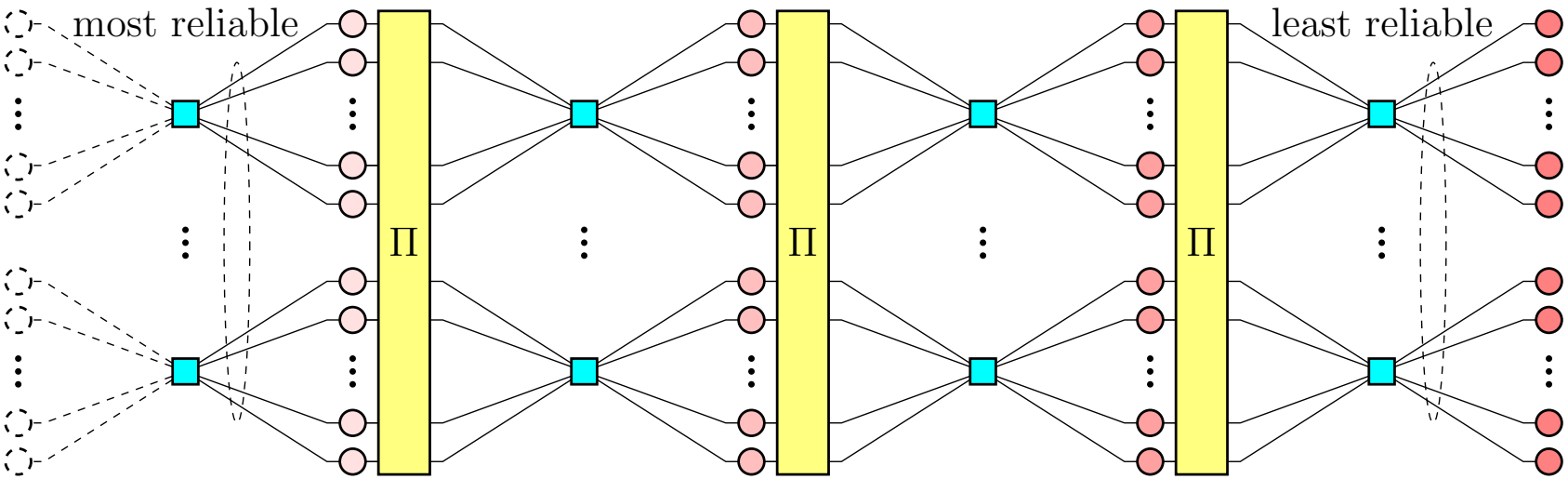
Part II

Spectrally-Efficient Fiber-Optic Communications

- Memoryless capacity estimates, with and without digital backpropagation
- Pragmatic coded modulation via staircase codes

Part I

Binary Error-Correcting Codes for High-Speed Communications



Outline

This part 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

This part 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. Generalizations and conclusions

Existing Solutions

Performance Measure

Net Coding Gain

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

$$Q = \sqrt{2}\mathrm{erfc}^{-1}\left(2 \cdot \mathrm{BER}\right).$$

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

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

and the Net Coding Gain (NCG) is

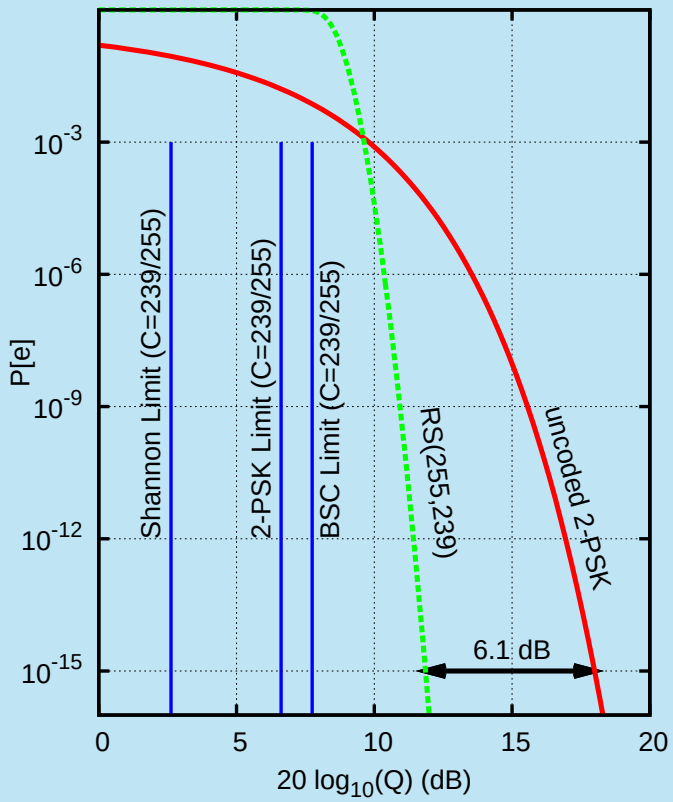
$$\mathrm{NCG\ (in\ dB)} = \mathrm{CG} + 10\log_{10}(\mathrm{R}),$$

where R is the rate of the code.

ITU-T G.975 (1996)

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

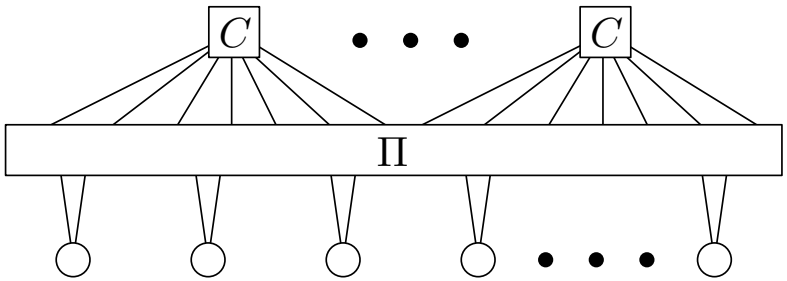


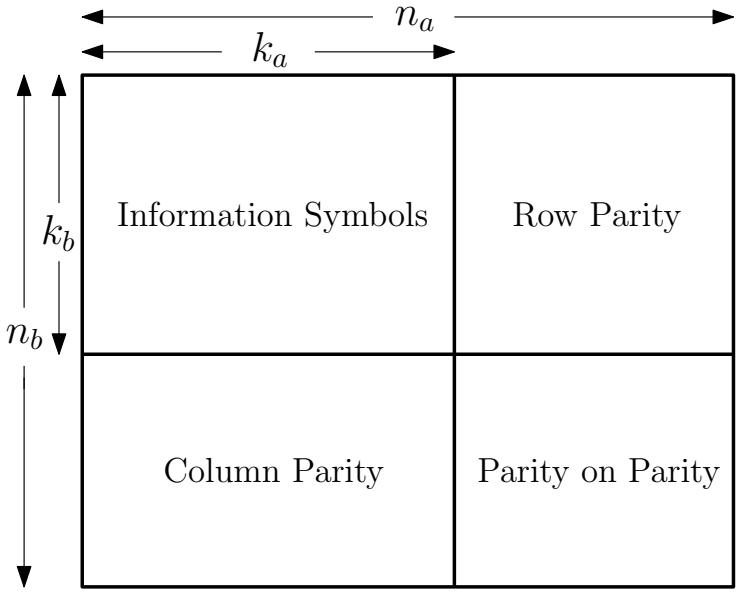
Constraint: Rate and framing structure **fixed** for future generations

ITU-T G.975.1 (2004)

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 @ $7 \cdot 10^{-14}$	$\sim$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

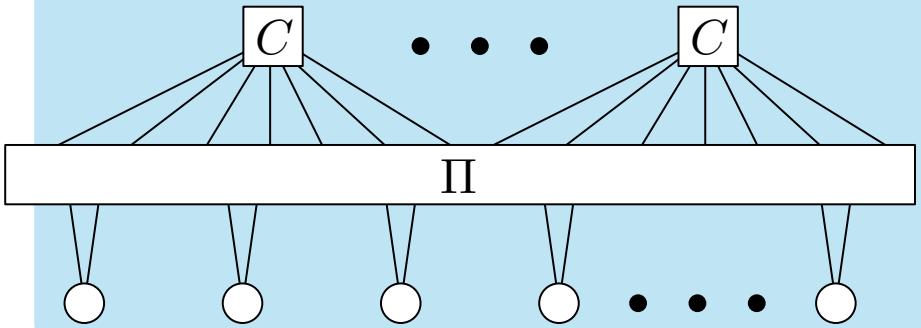
Low Implementation Complexity

Dataflow considerations at 100Gb/s

Implementation Considerations

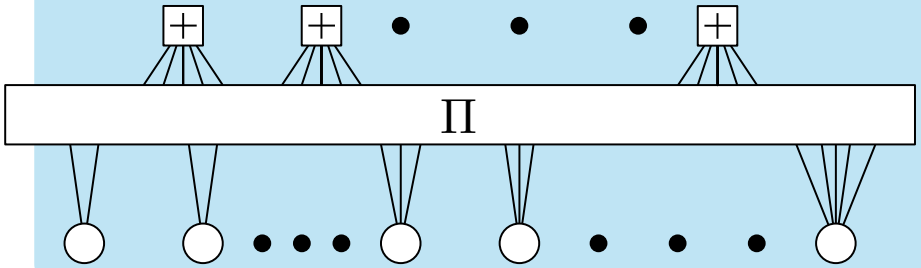
Hardware Considerations: Product vs. LDPC

Product Code



- Algebraic component codes
- Syndrome-based decoding

LDPC Code



- SPC component codes
- Belief propagation decoding

Syndromes

The codewords of a linear (n, k) block code C are the solutions of a homogeneous system of simultaneous linear equations, i.e.,

$$\mathbf{v} \in C \Leftrightarrow \mathbf{v}H^\top = \mathbf{0}.$$

Let $\mathbf{v} \in C$ be sent and let $\mathbf{r}$ be received. Define the *error pattern* $\mathbf{e}$ as $\mathbf{e} = \mathbf{r} - \mathbf{v}$ so that

$$\mathbf{r} = \mathbf{v} + \mathbf{e}.$$

Then the *syndrome* corresponding to $\mathbf{r}$ is

$$\mathbf{s} = \mathbf{r}H^\top = (\mathbf{v} + \mathbf{e})H^\top = \underbrace{\mathbf{v}H^\top}_{\mathbf{0}} + \mathbf{e}H^\top = \mathbf{e}H^\top.$$

An error-correcting decoder must infer the most likely error pattern from the syndrome.

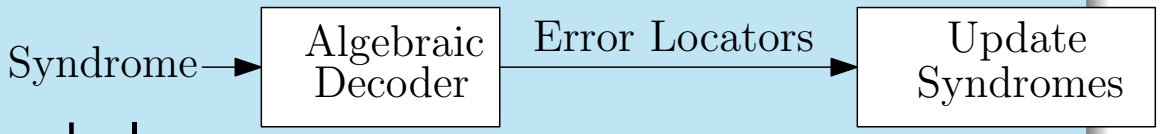
Syndrome-based Decoding of Product Codes

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**

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

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



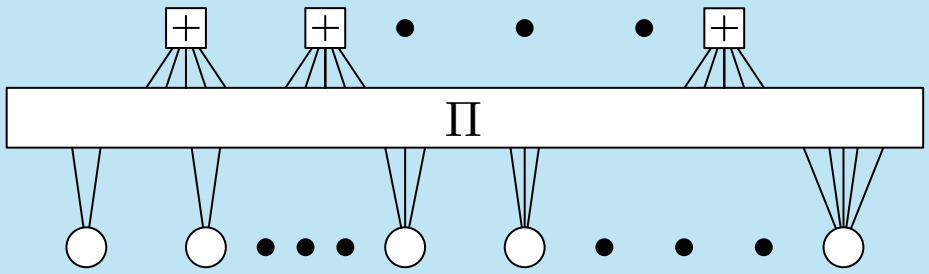
Total Dataflow

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!

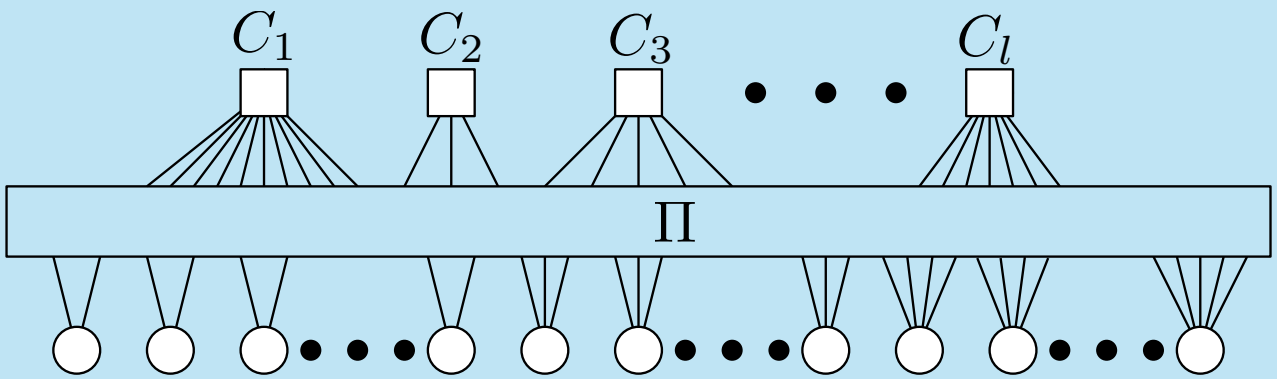
Dataflow Comparison

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



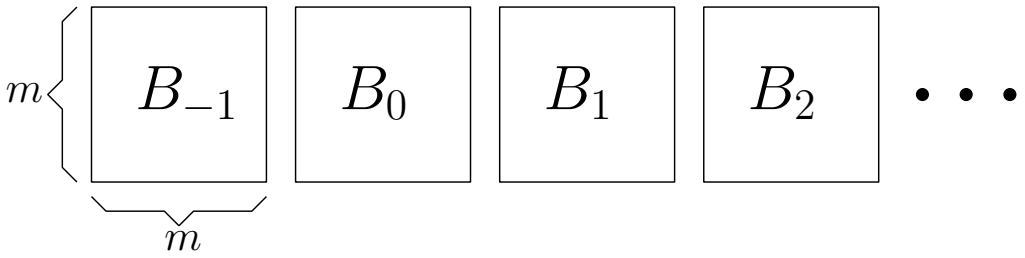
Graph Optimization

Degrees of Freedom

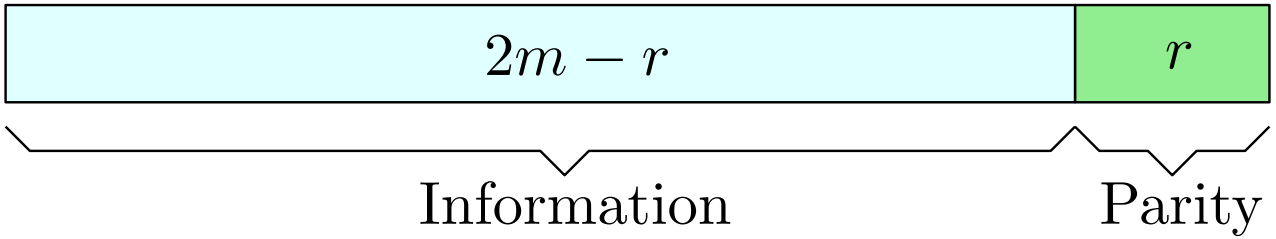
- Mixture of component codes (e.g., Hamming, BCH)
- Multi-edge-type structures

Staircase Codes: Construction

Consider a sequence of m-by-m matrices B_i



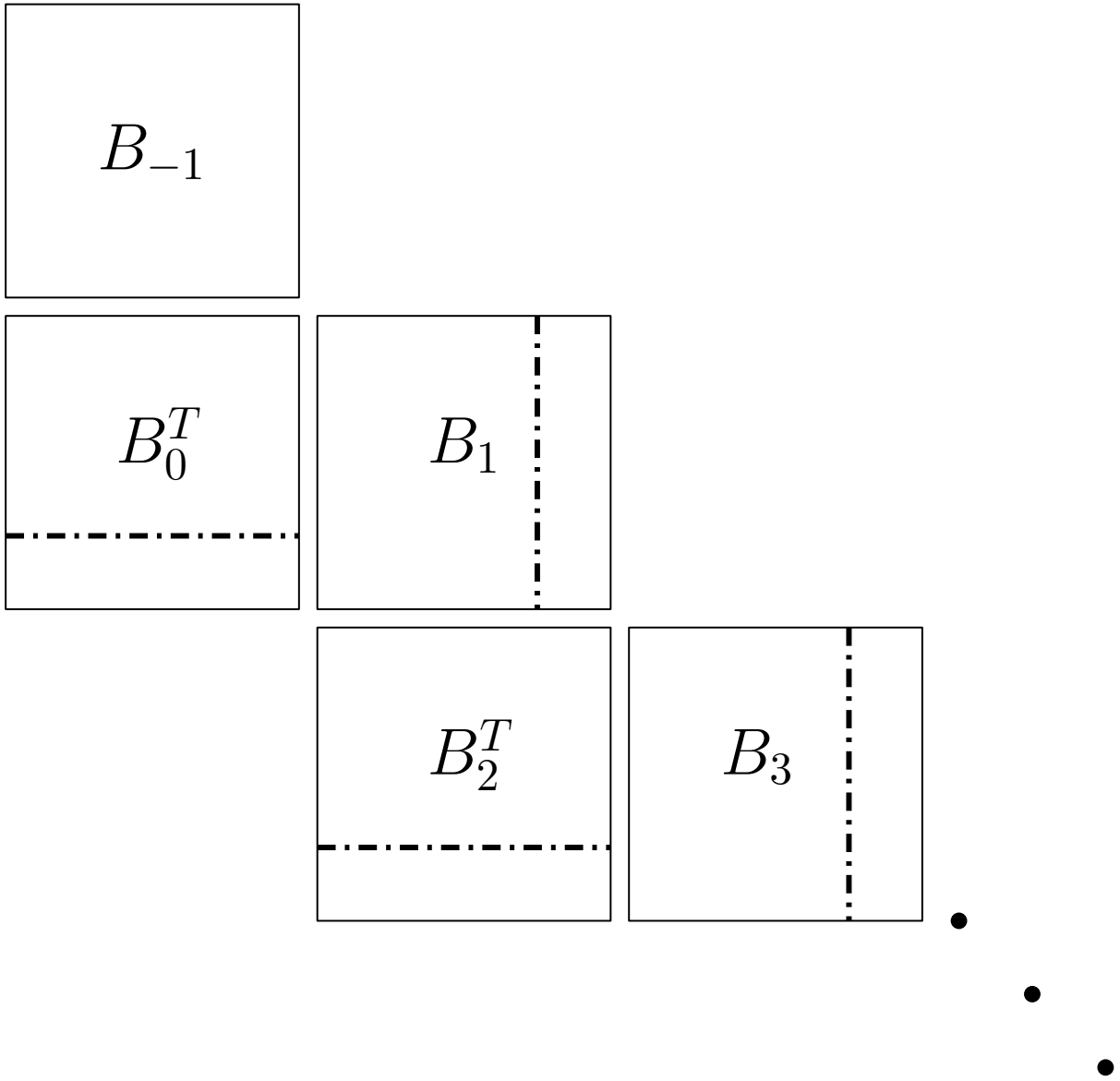
and a linear, systematic, $(n = 2m, k = 2m - r)$ component code C



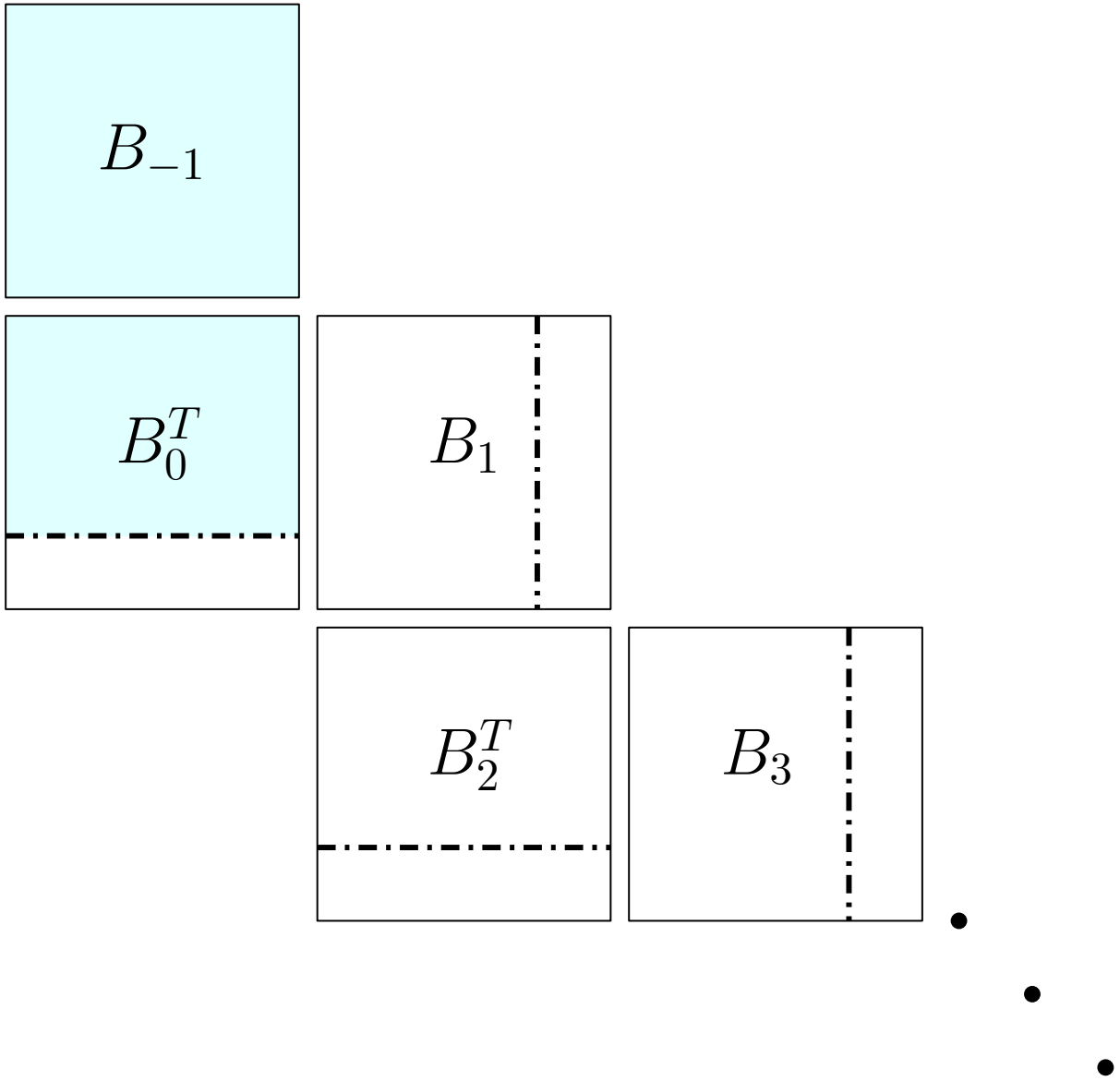
Encoding Rule

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

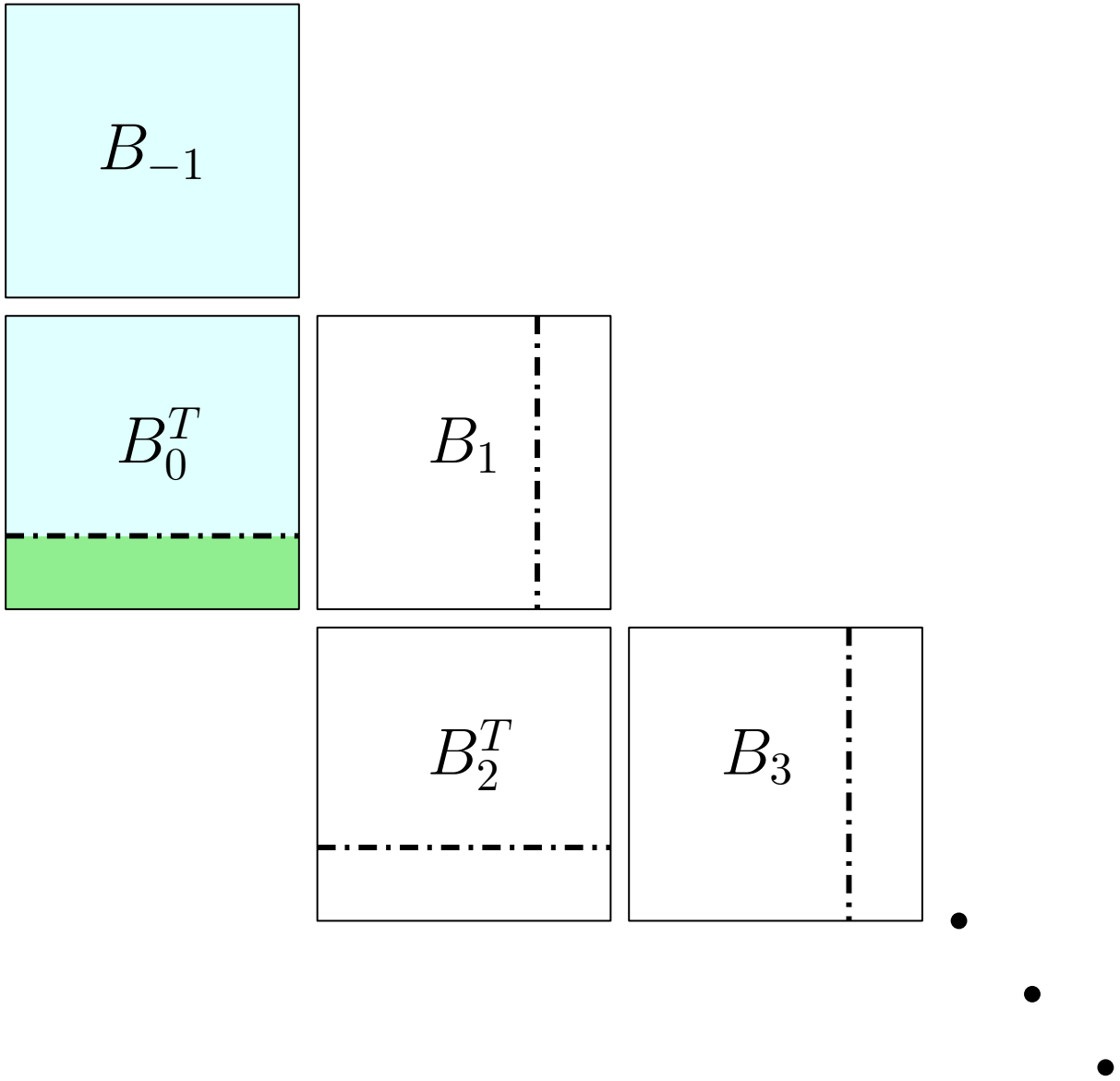
Staircase Codes: Construction



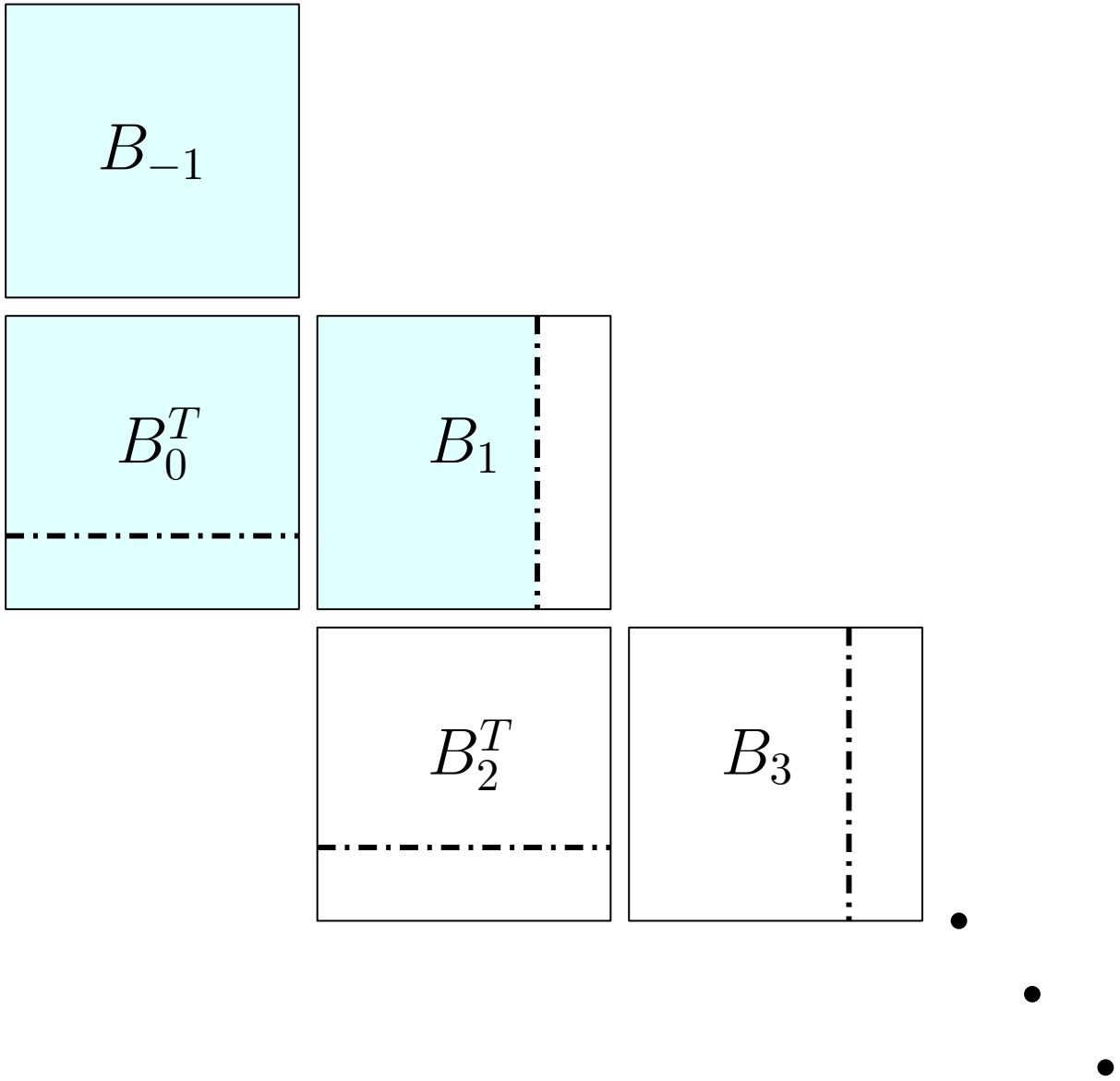
Staircase Codes: Construction



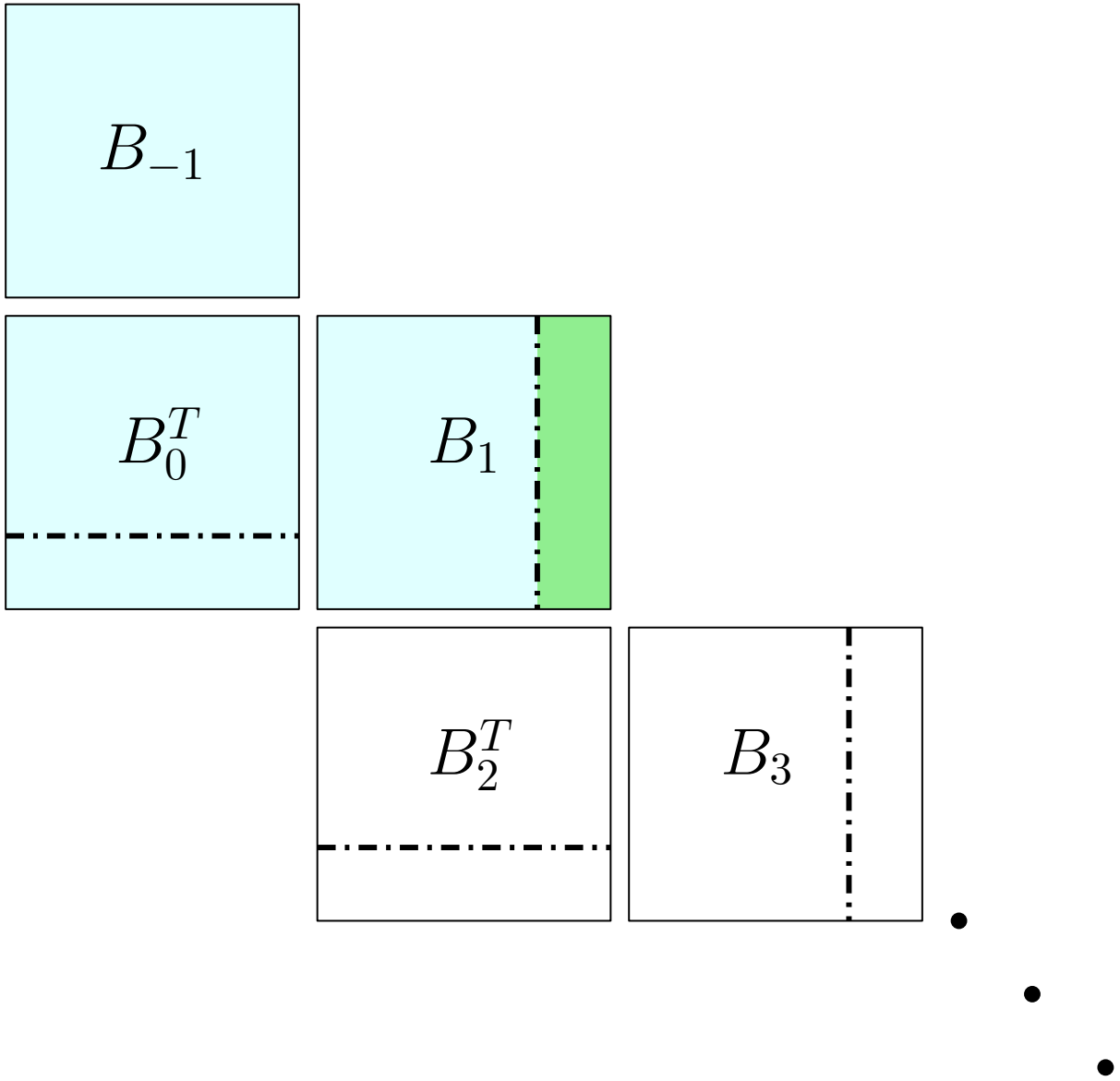
Staircase Codes: Construction



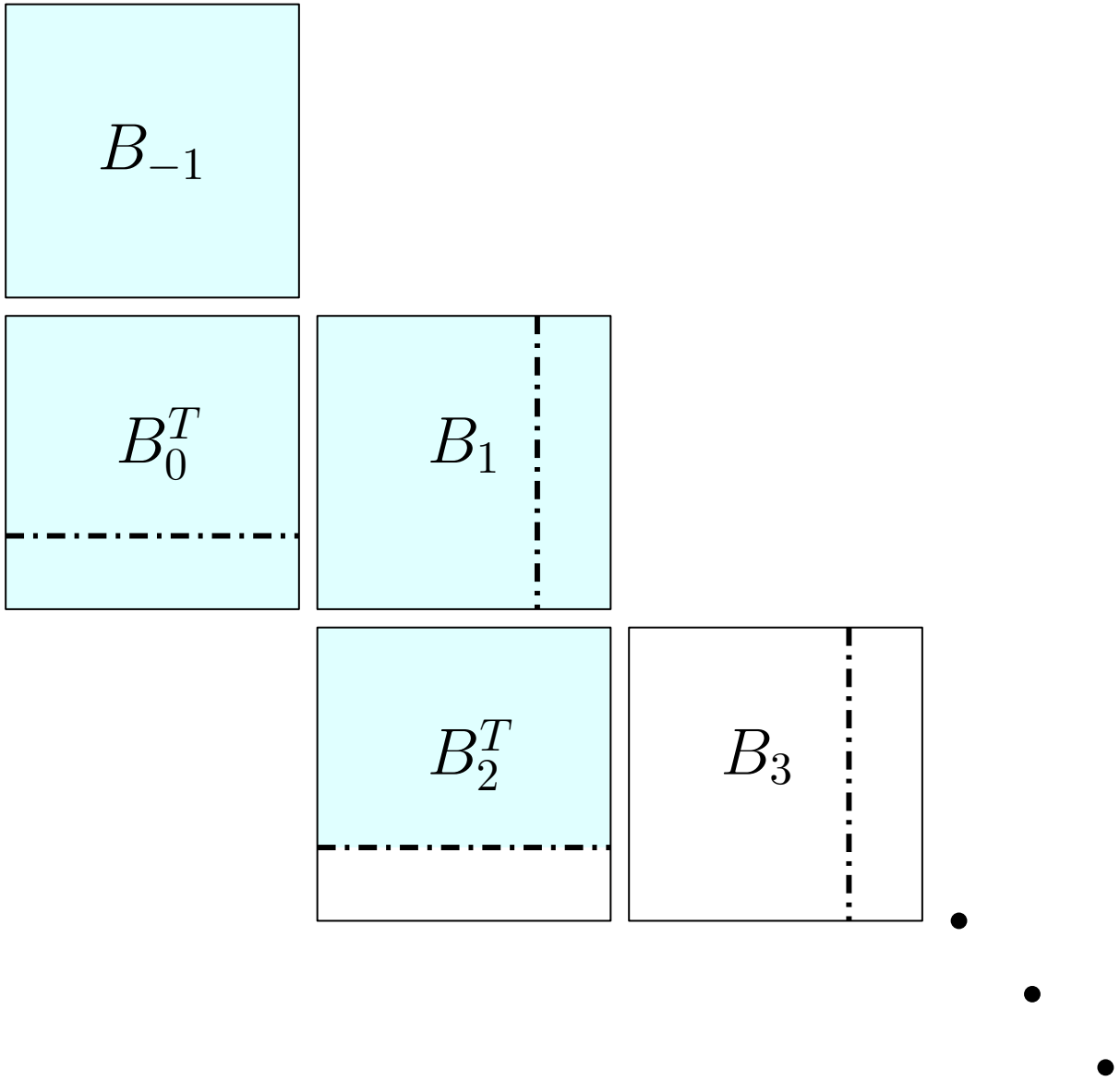
Staircase Codes: Construction



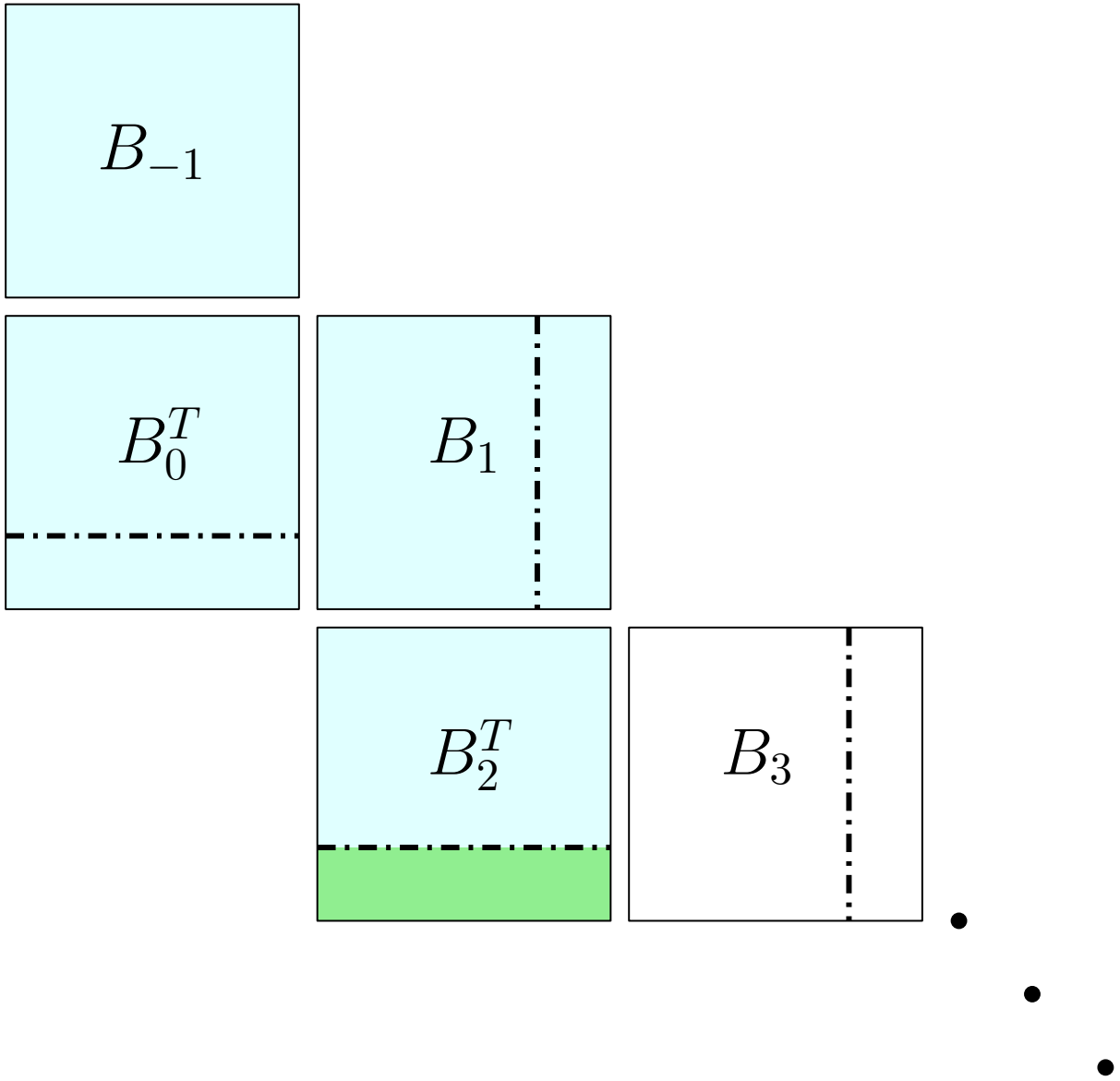
Staircase Codes: Construction



Staircase Codes: Construction

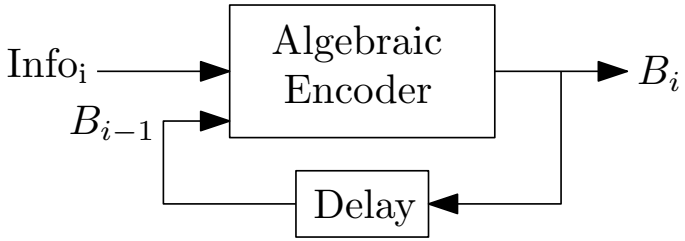


Staircase Codes: Construction

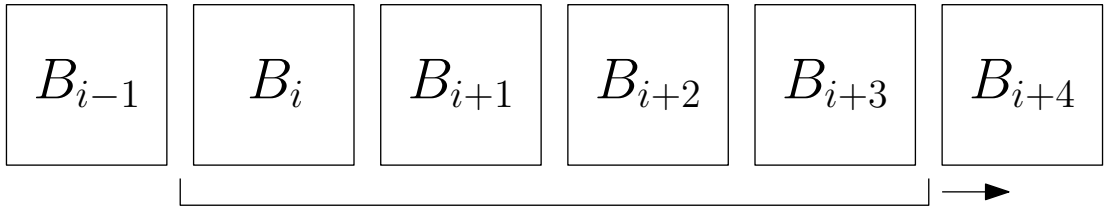


Staircase Codes: Properties

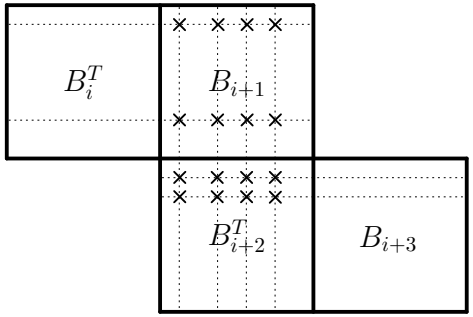
- Hybridization of recursive convolution coding and block coding
  - Recurrent Codes of Wyner-Ash (1963)



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



Staircase Codes: Properties



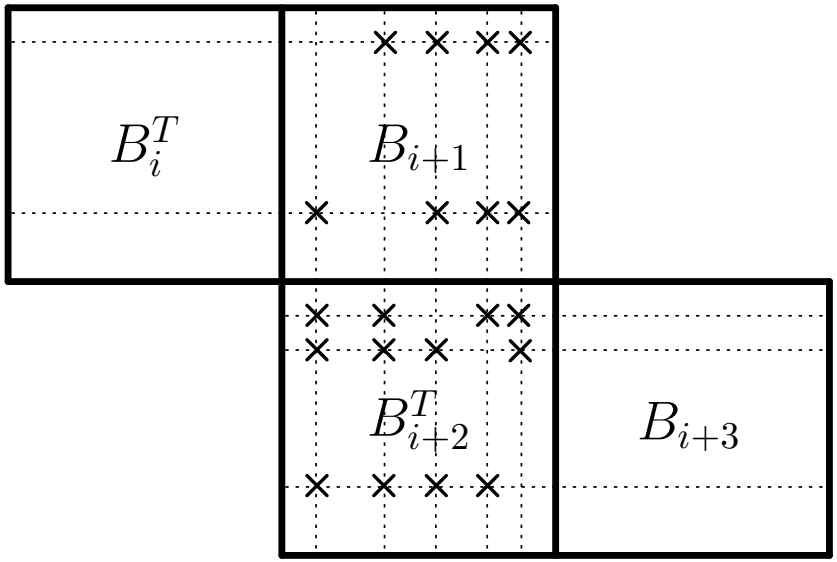
The multiplicity of (minimal) stalls of size $(t+1) \times (t+1)$ is

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

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

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

General Stall Patterns

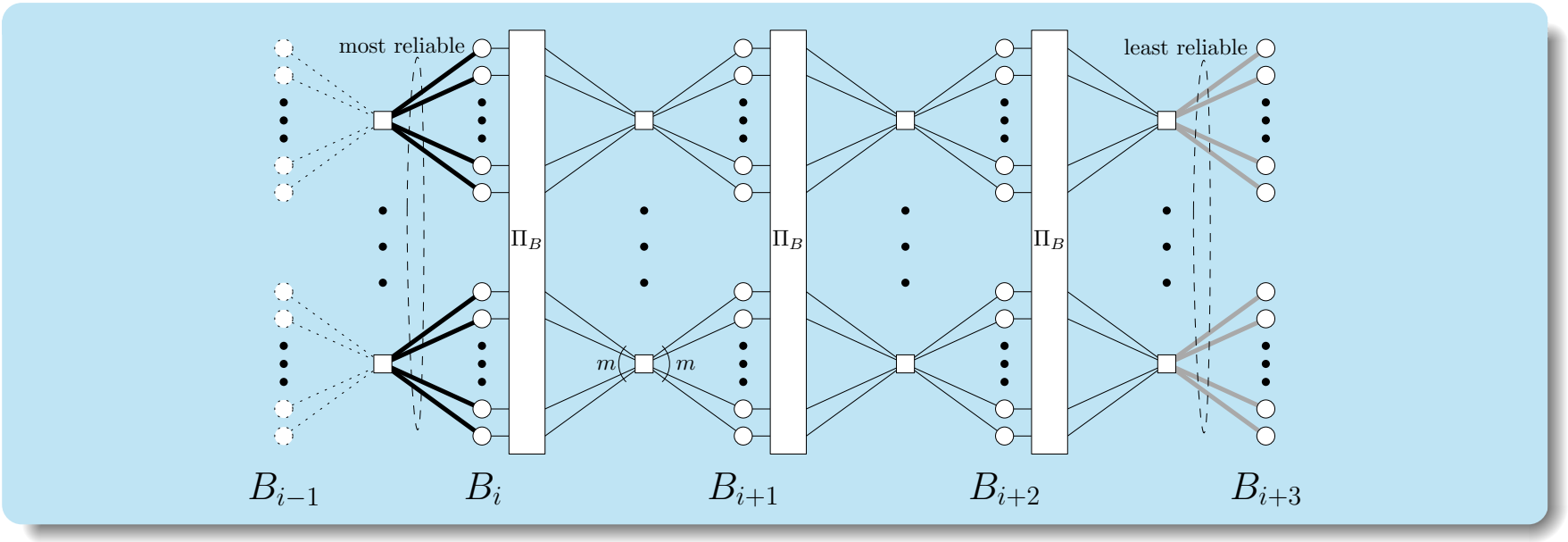


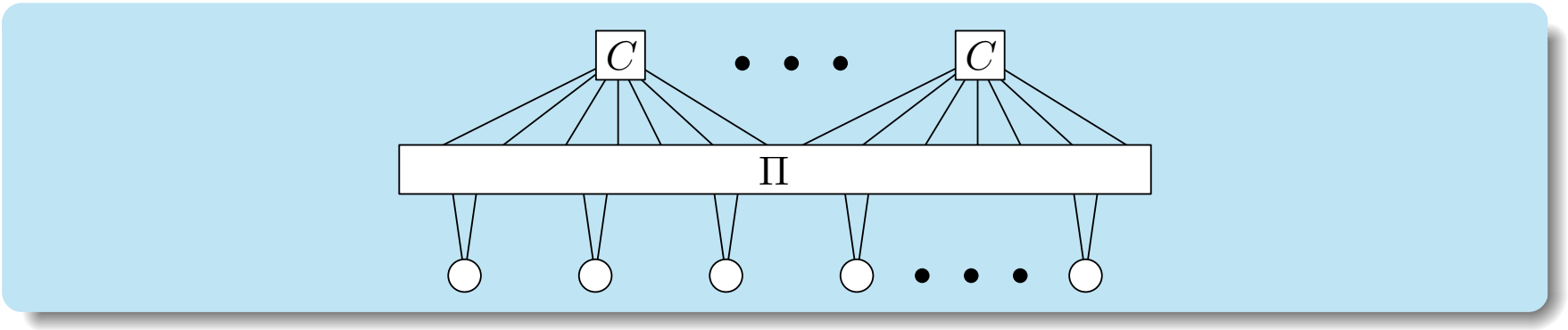
(5,5)-stall

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

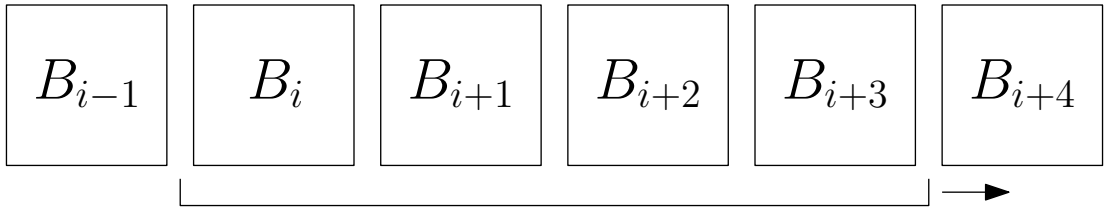
Contribution of (K, L)-stalls, $p = 4.8 \times 10^{-3}$.

Multi-Edge-Type Representation

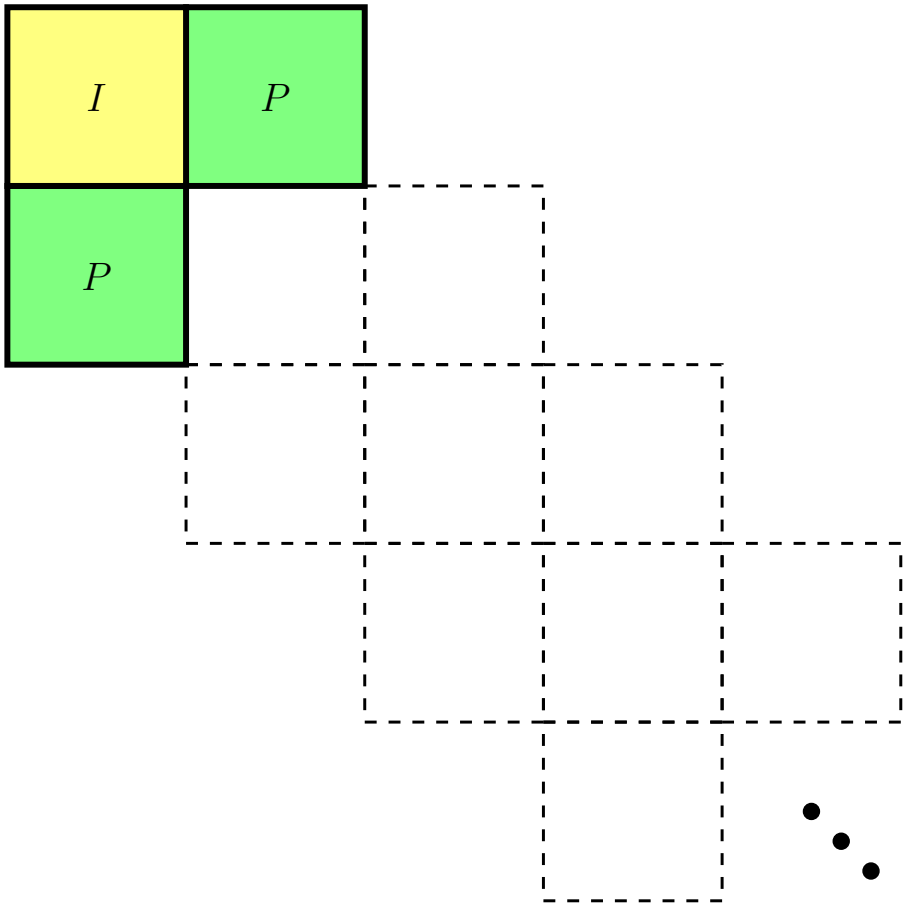




Movie Time

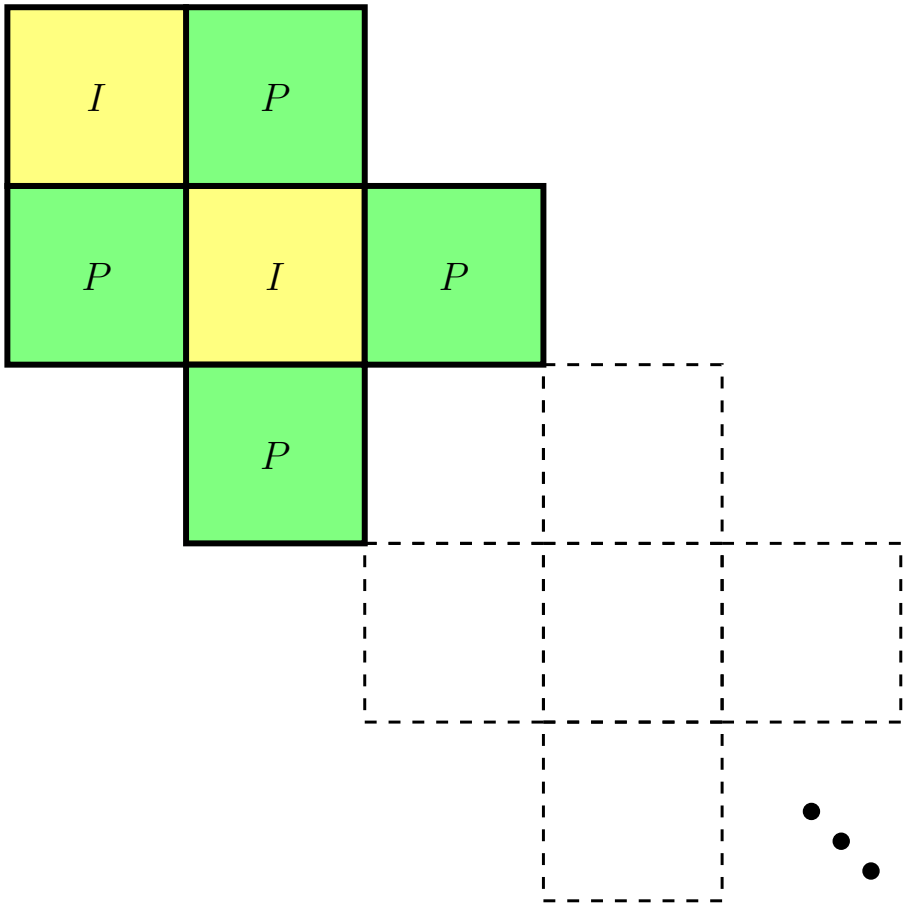


Braided Block Codes: Construction



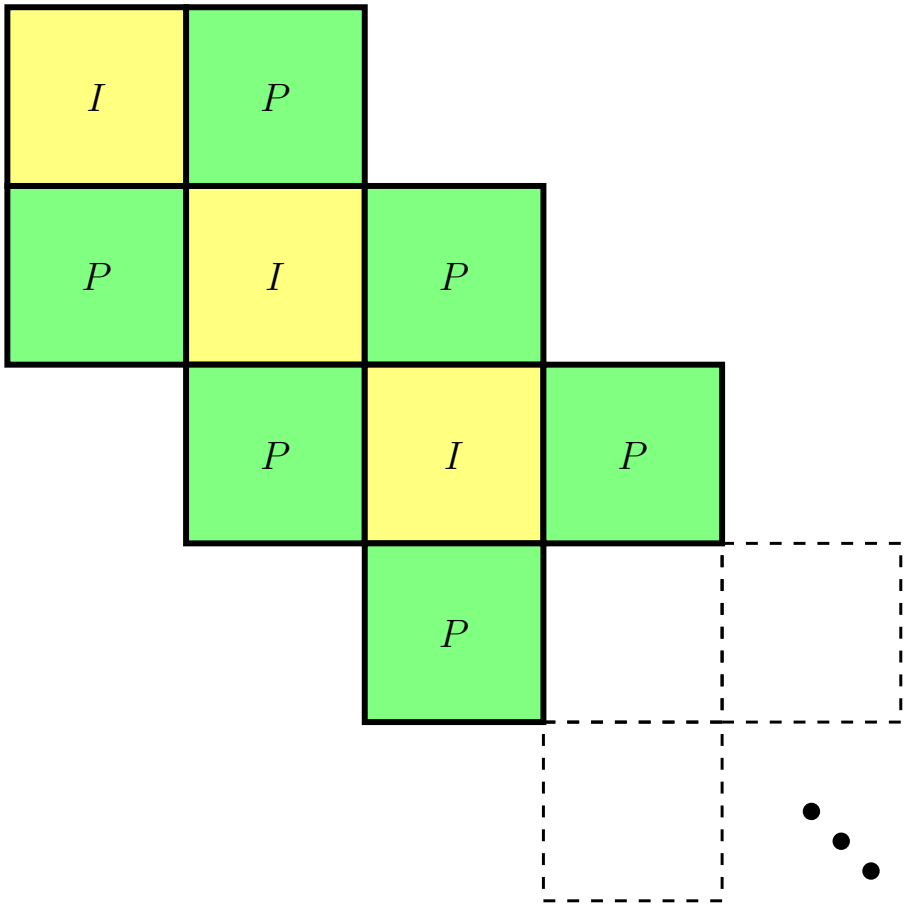
Felström, Truhachev, Lentmaier, Zigangirov, "Braided Block Codes," *IEEE Trans. on Info. Theory*, 2009.

Braided Block Codes: Construction



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

Braided Block Codes: Construction



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

Braided Block Codes: Performance

Our experiments show:

- For high-rate (239/255), braided block codes yield approximately the same performance as the corresponding product code
- "Multi-edge gain" diminished, since fewer bits (previous parity only) serve as "clean"

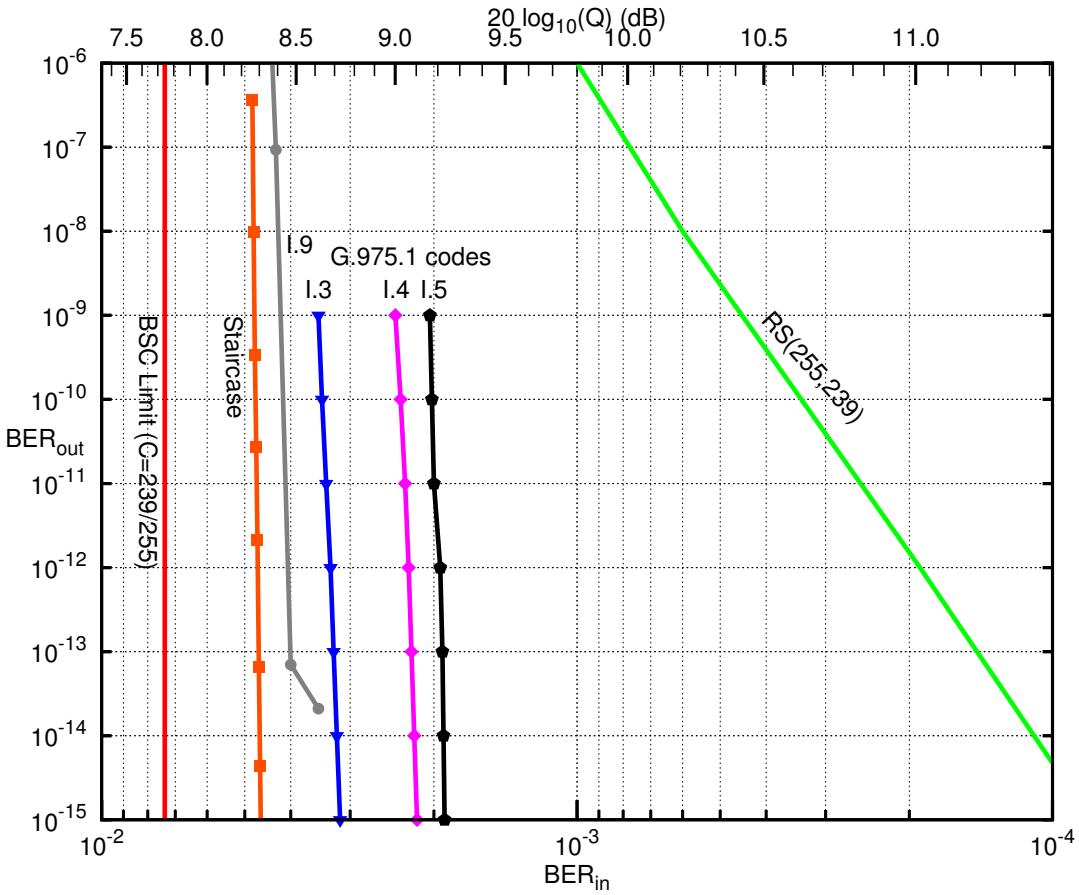
But:

- Spatially-coupled generalized LDPC ensembles with iterative hard-decision decoding can approach capacity at high rates: Y.-Y. Jian, H. Pfister, K. Narayanan, "Approaching Capacity at High Rates with Iterative Hard-decision Decoding," *Proc. IEEE Int. Symp. Inform. Theory*, July, 2012.

FPGA-based Simulation Results

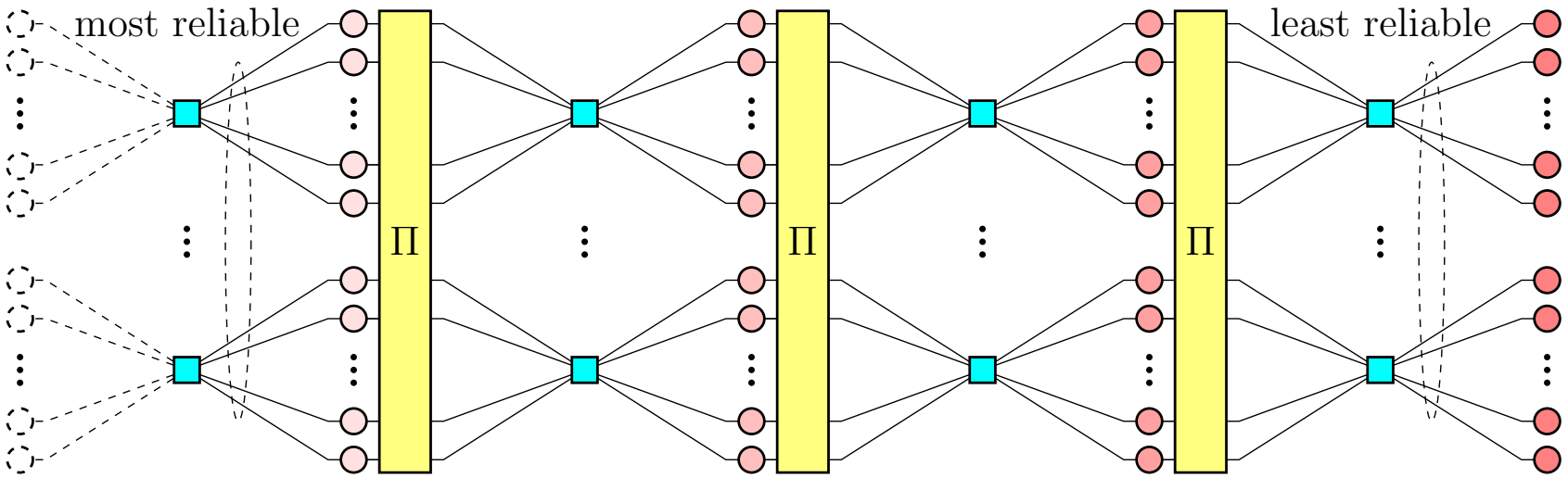
Code Parameters

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



Part II

Spectrally-Efficient Fiber-Optic Communications



The Kerr Effect

Kerr electro-optic effect (DC Kerr effect)

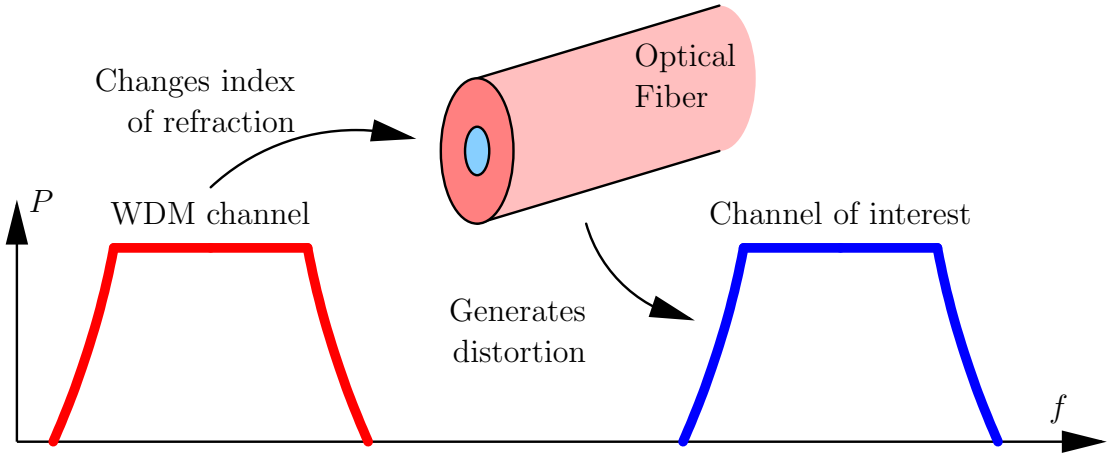
- An effect discovered by John Kerr in 1875
- It produces a change of refractive index in the direction parallel to the externally applied electric field
- The change of index is proportional to the *square* of the magnitude of the external field

John Kerr
(1824-1907)

The Kerr Effect in Optical Fibers

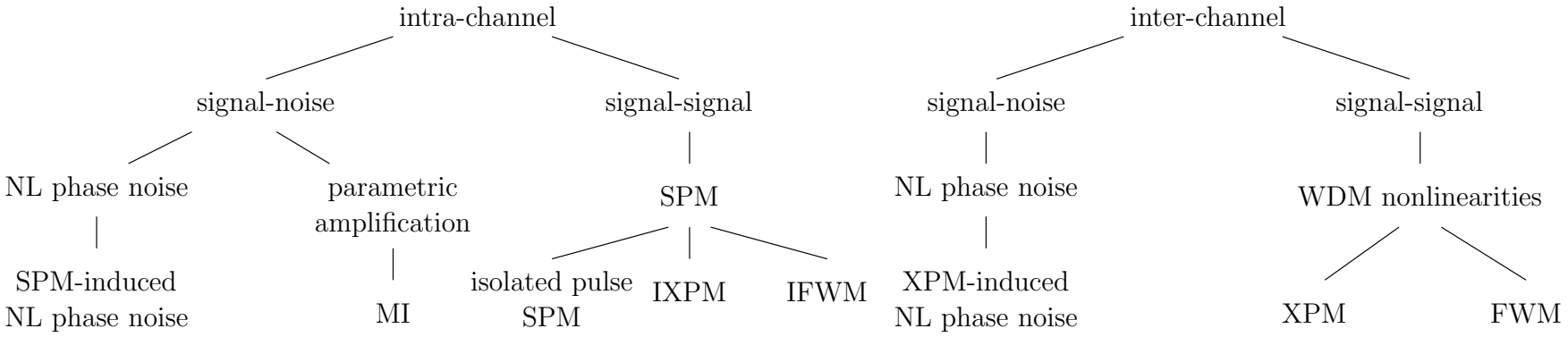
Optical Kerr Effect (or AC Kerr effect)

- 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



A signal in a certain frequency band can distort the signal in a different frequency band *without spectral overlap*

Nonlinear Effects in Fibers



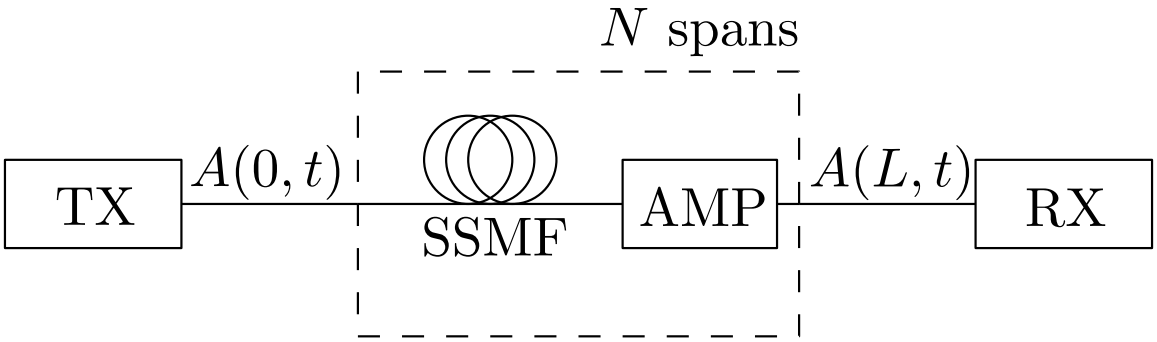
NL=nonlinear; SPM=self-phase modulation; MI=modulation instability; XPM=cross-phase modulation; IXPM=intrachannel XPM
FWM=four-wave mixing; IFWM=intrachannel FWM; WDM=wavelength division multiplexing

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

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

Evolution of $A(z,t)$

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

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

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

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

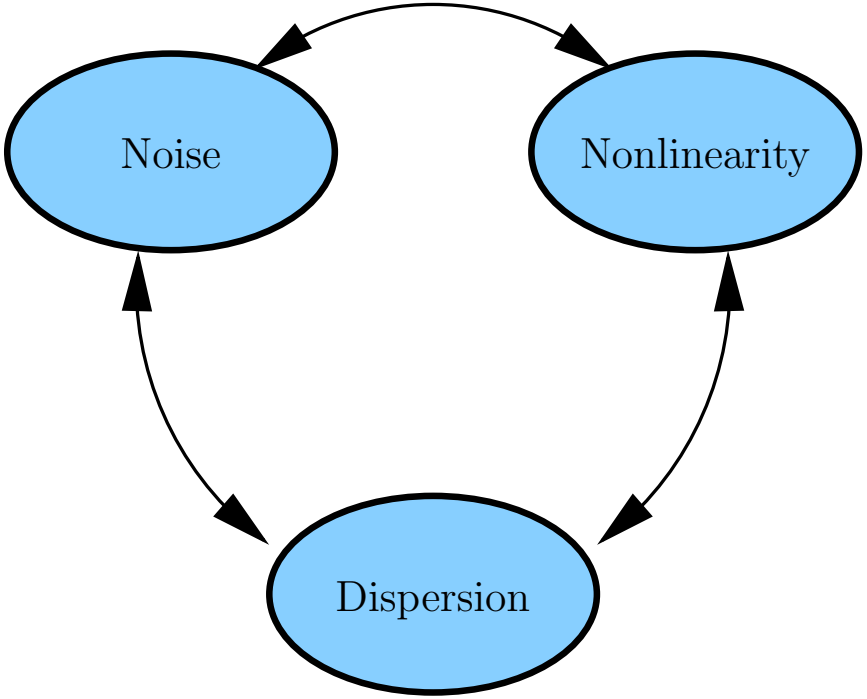
$$\mathcal{E}\left[n(z,t)n^\star(z',t')\right] = \alpha h\nu_s K_T \delta(z-z', t-t'),$$

System Parameters

Second-order dispersion β_2	-21.668 $\mathrm{ps}^2/\mathrm{km}$
Loss α	4.605×10^{-5} m^{-1}
Nonlinear coefficient γ	1.27 $\mathrm{W}^{-1}\mathrm{km}^{-1}$
Center carrier frequency ν_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.

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

Nonlinear Step

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

$$\frac{\partial A}{\partial z} = j\gamma |A|^2 A,$$

with solution

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

Linear Step

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

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

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

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2} A - \frac{j\beta_2}{2} \frac{\partial^2 A}{\partial t^2},$$

which can be efficiently solved in the frequency domain. Defining

$$A(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{A}(z, \omega) \exp(j\omega t) d\omega,$$

it can be shown that

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

Split-Step Propagator

Putting this together, we have

$$A(z_0 + h, t) = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\hat{A}(z_0, t)\right\} \exp\left(\left(j\frac{\beta_2}{2}\omega^2 - \frac{\alpha}{2}\right) h\right)\right\},$$

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

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

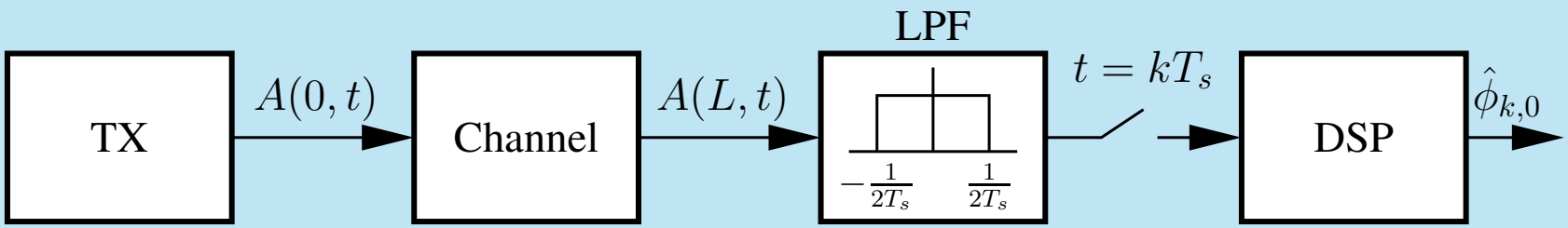
Digital Backpropagation

- 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

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 I signal:

$$X_I(t) = \sum_{k=-\infty}^{\infty} \frac{\phi_{k,I}}{\sqrt{T_s}} \mathrm{sinc}\left(\frac{t - kT_s}{T_s}\right),$$

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

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

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

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

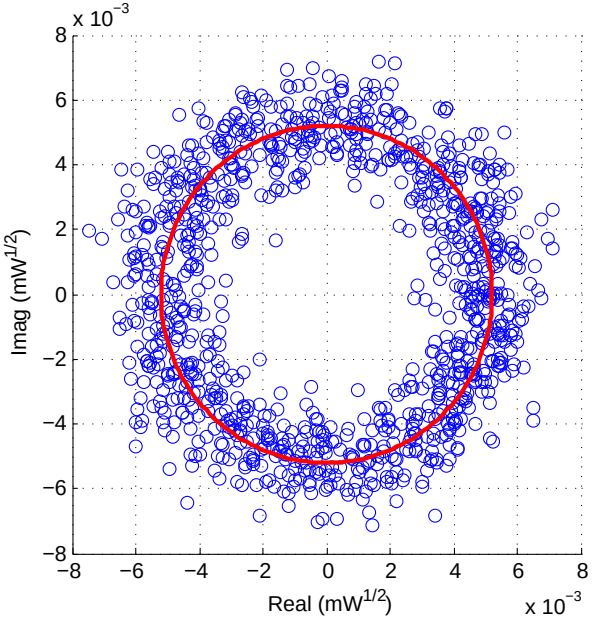
$$A(z=0,t) = \sum_{k=-\infty}^{\infty} \sum_{l=-B}^{B} \frac{\phi_{k,l}}{\sqrt{T_s}} \mathrm{sinc}\left(\frac{t-kT_s}{T_s}\right) e^{j2\pi lt/T_s}.$$

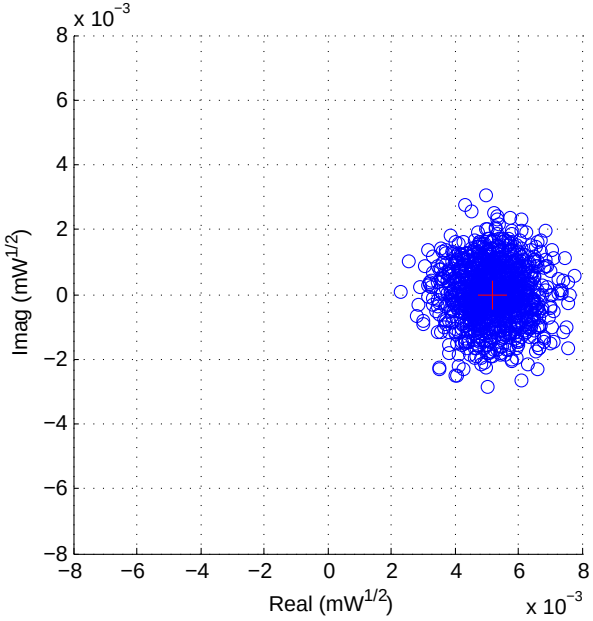
Towards a Probability Model

Back-rotation:

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

where Φ_{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 μ_i and covariance matrix Ω_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

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

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

Capacity Estimation

The mutual information of the (assumed) memoryless channel is

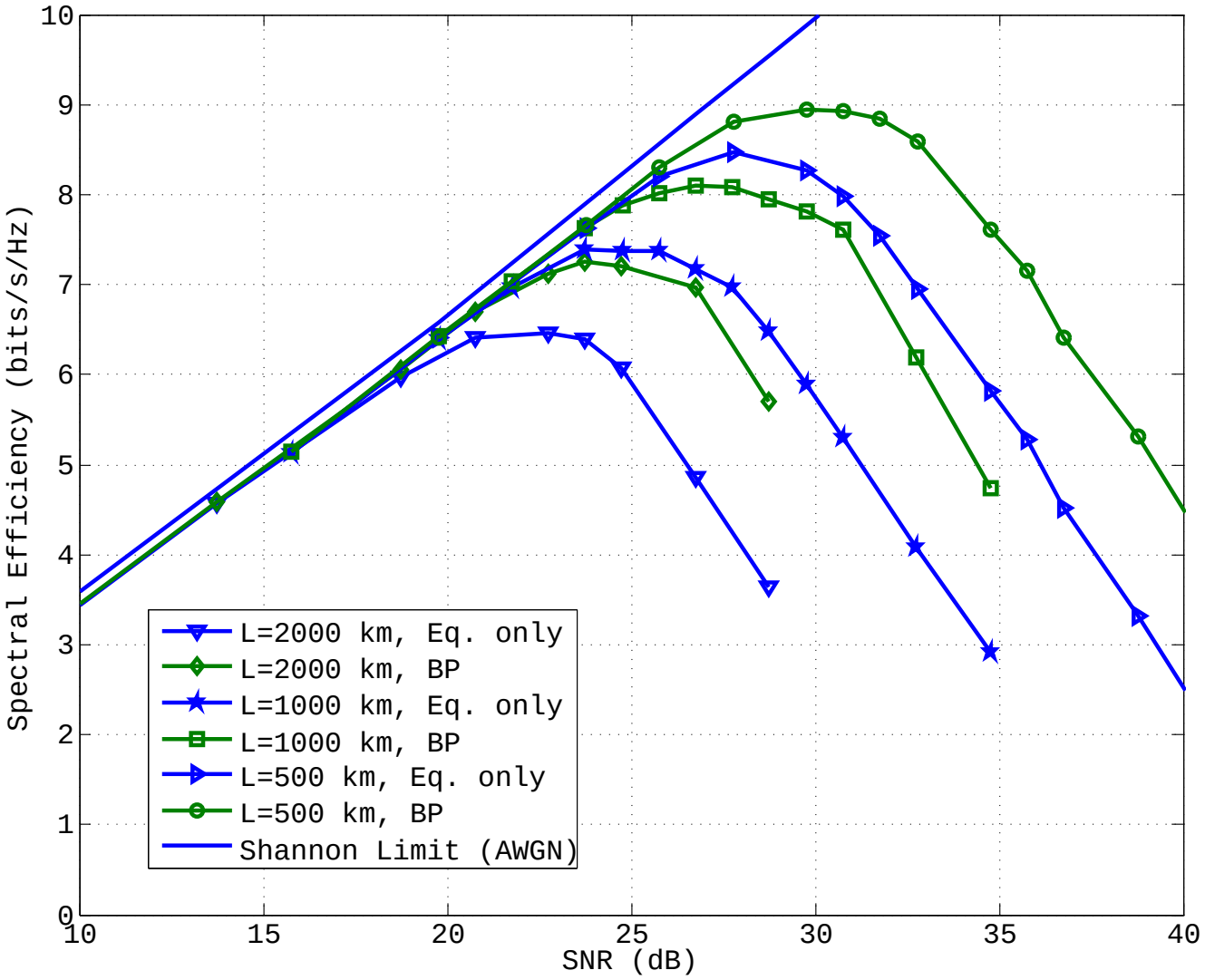
$$I(X;Y) = \int\int f(x,y) \log_2 \frac{f(y|x)}{f(y)} \, dx \, dy,$$

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	$2B + 1 = 5$

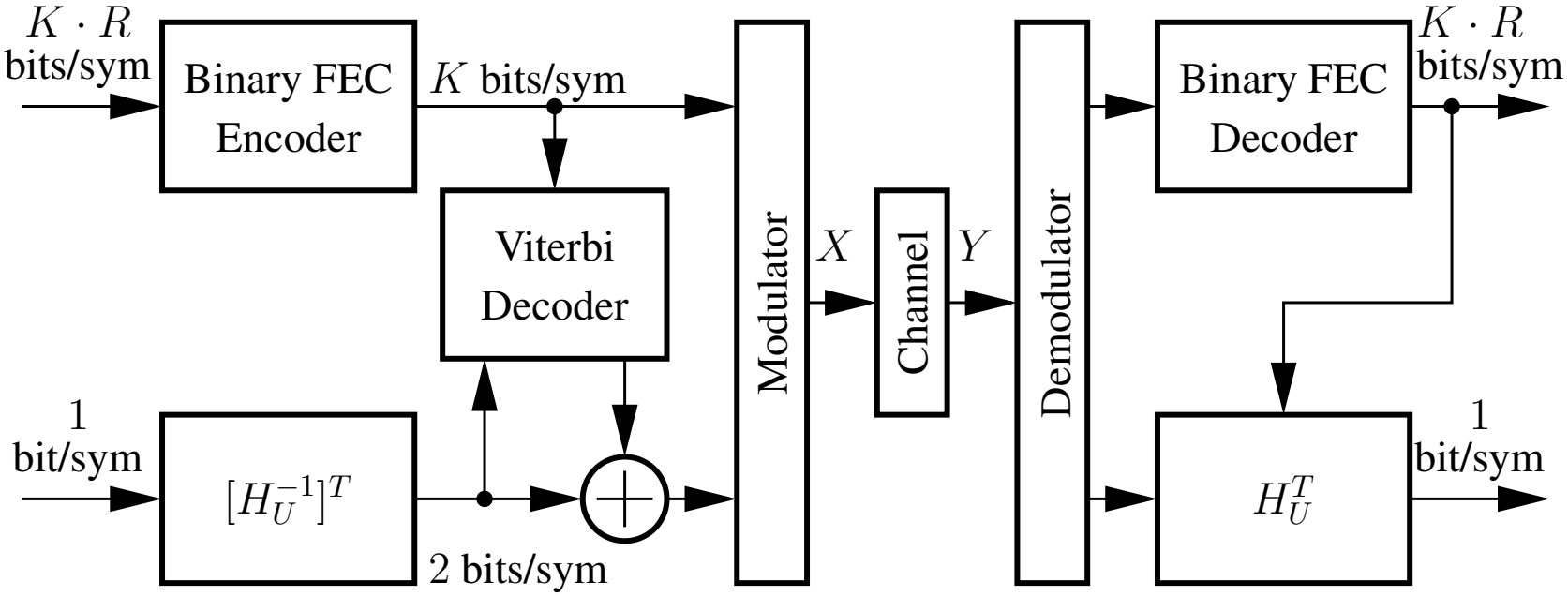
Achievable Rates from Memoryless Capacity Estimate



(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

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

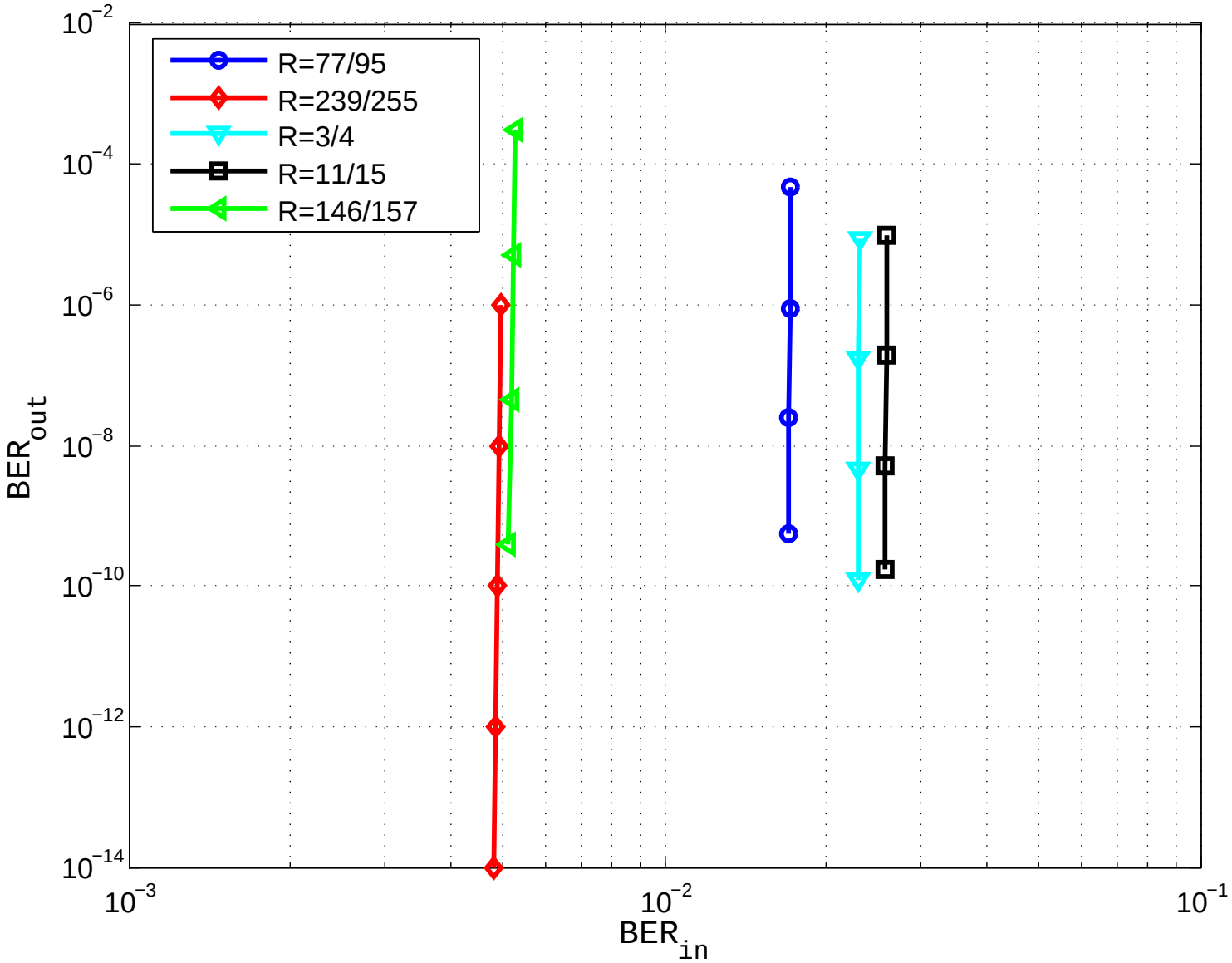
Achievable Rates

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

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

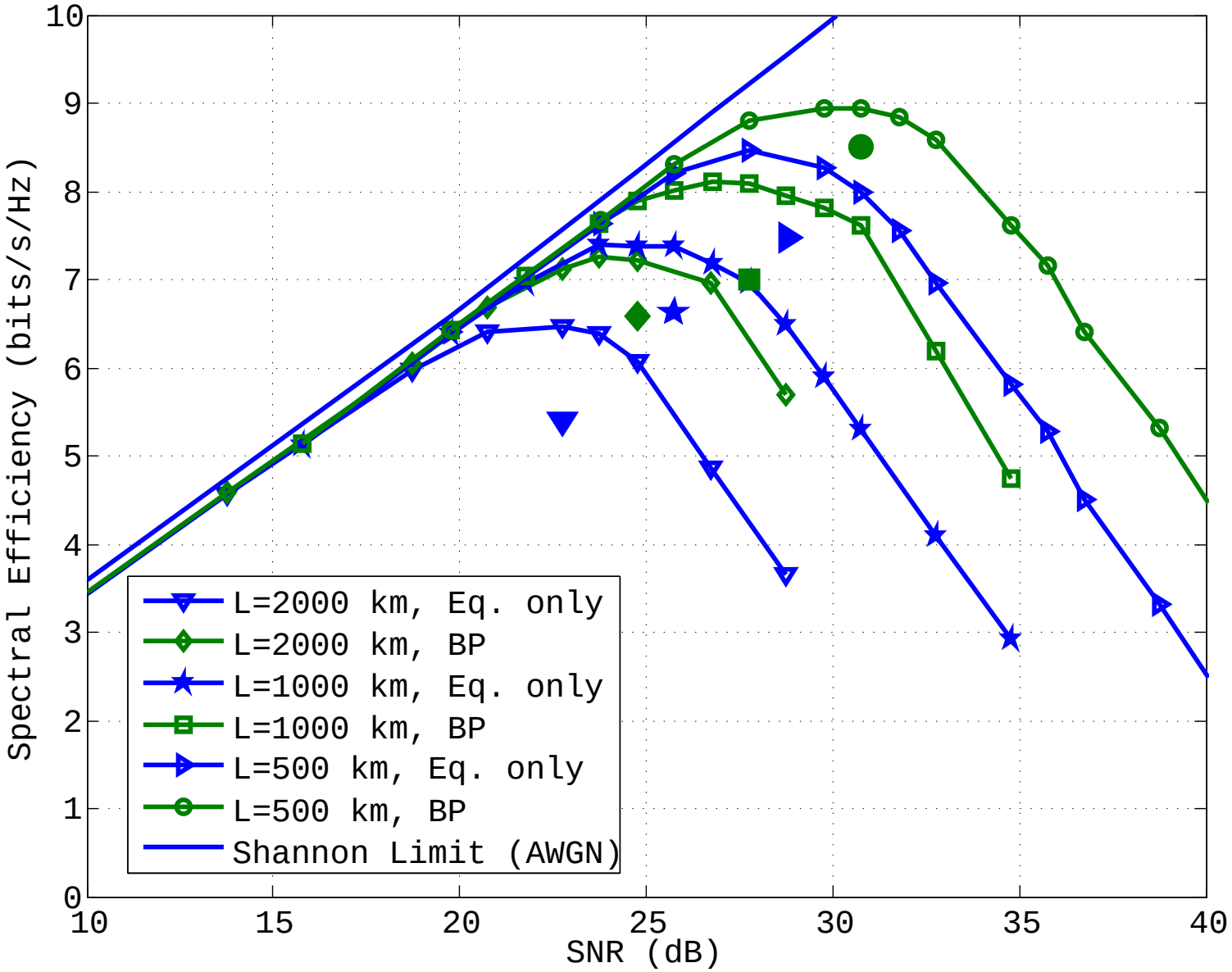
Staircase code design

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. (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

Conclusions

- At rate 239/255, staircase codes provide best-in-class 9.41 dB NCG
- For high-spectral efficiency modulation, a pragmatic coding approach using staircase codes yields performance with 0.6 bits/s/Hz (at $L = 2000$ km) with practical decoding complexity

Future Directions

- Can these capacity estimates be improved?
- Can soft-decisions be incorporated while retaining low complexity?
- Are there (well-principled) alternatives to digital backpropagation?

Future Directions

- Can these capacity estimates be improved?
- Can soft-decisions be incorporated while retaining low complexity?
- Are there (well-principled) alternatives to digital backpropagation?

Thank you for your attention!