Approximate Message Passing for Short-Reach Optical Communication

Gerhard Kramer*
Institute for Communications Engineering (ICE)
Technical University of Munich (TUM)

Workshop on Information Theory for Future Networks (IT-FuN)
KIT, March 4, 2026

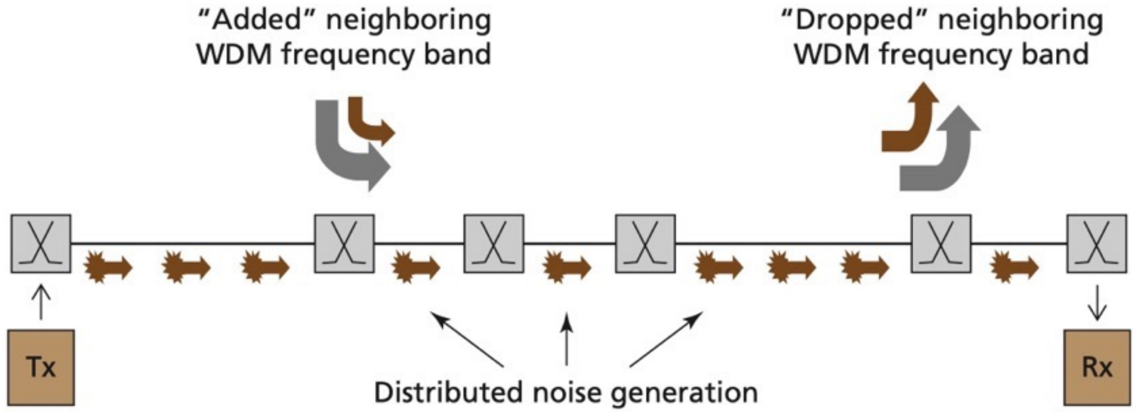
* contributions by
- **Daniel Plabst**, Tobias Prinz, Thomas Wiegart, Alex Jäger of TUM-ICE
- Mohamed Akrout, Amine Mezghani of University of Manitoba, Canada

Outline

1) Background
   - Channel model, information theory
   - Transceiver structure
2) Equalizer algorithms
   - Classic to modern with machine learning
   - Focus on AMP/EP* and compare rates/energies/ complexities
3) AIRs** from numerical experiments

* AMP = approximate message passing
EP = expectation propagation
** AIR = achievable information rate

1) Background

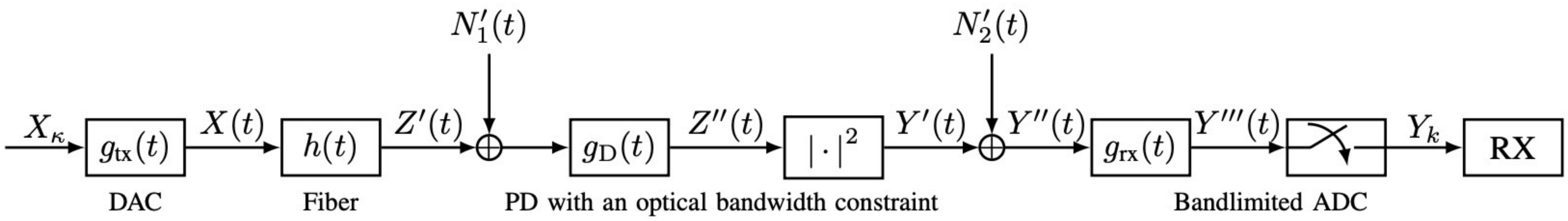


- Model: generalized NLSE with EM fields E(z,t) and n(z,t) at position z and retarded time T

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

- α, β_2, β_3, γ are fiber coefficients, n(z,T) is amplifier noise
- Notation: X(t) = E(0,t) and Y(t) = E(L,t) where L is fiber length
- Short reach: discard non-linearity, keep chromatic dispersion (CD)
- Use direct detection (DD): model as instantaneous squaring

Short-Reach Model



- Consider either optical $N_1'(t)$ or electrical $N_2'(t)$ noise
- DD doubles the bandwidth, so must double the sampling rate for sufficient statistics; this gives long memory under band limitations
- Theory [1] based on Fourier series: capacity is within 1 bpcu of coherent capacity for optically amplified links!
- Reason: the receiver can recover phase information
- This result motivated our work: Is the theory practical?
  And can one do better? The answer is yes to both questions.

[1] Mecozzi, Shtaif, IEEE/OSA JLT 2018

- Question: what reliabilities/rates/energies are possible?
- Digital approach:
  - discretize waveforms in time* and values** to get vectors $\underline{X}$ and $\underline{Y}$
  - use error-control codes to combat noise (including interference)
- Reliable communication is possible if the # of source bits is less than

$$I(\underline{X};\underline{Y}) = \mathrm{E}\left[\log \frac{P(\underline{X},\underline{Y})}{P(\underline{X})P(\underline{Y})}\right] = \begin{cases} H(\underline{Y}) - H(\underline{Y}|\underline{X}) \\ H(\underline{X}) - H(\underline{X}|\underline{Y}) \end{cases}$$

- Two Challenges:
  1) How to compute P($\underline{y}$|$\underline{x}$) and/or I($\underline{X}$;$\underline{Y}$)?
  2) How to build the transmitter and receiver?

* More generally: use projections (filter & sample is a projection)
** Quantization

Challenges

- Problem 1: approximate $P(\underline{y}|\underline{x})$ with a surrogate $Q(\underline{y}|\underline{x})$ used to decode. The AIR is:

$$I_Q(\underline{X};\underline{Y}) = E[\log Q(\underline{Y}|\underline{X}) / Q(\underline{Y})] \le I(\underline{X};\underline{Y})$$

Common target surrogate of equalization* is: $Y_k = X_k + N_k$ for all k

- Problem 2: need
  a) higher-order coded modulation and shaping
  b) multi-level encoding/decoding to approach $I_Q(\underline{X};\underline{Y})$
  c) an equalizer* that achieve high rates with reasonable complexity
     Example: for linear channels and loose delay constraints: use orthogonal frequency-division multiplexing (OFDM)

* Equalizer: a device that "reduces memory" in a "good way": little information loss and low complexity for both the equalizer and the encoder/decoder.

Equalizer Algorithms

2) Equalizer Algorithms

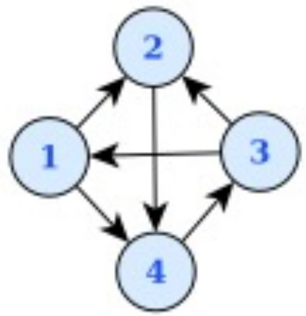
- Examples of classic soft-output equalizers:
  - Linear [1]: Wiener filter, simple, but cannot address nonlinearities
  - Forward-backward algorithm (FBA) [2]: optimal but too complex
  - Channel shortening & small FBA [3]: better, but interference-limited
  - Decision feedback: simple but requires reliable decisions
  - Volterra filters: poor complexity/performance tradeoff
- Modern equalizers based on machine learning perform better:
  - Gibbs sampling (GS) [4]: a Markov chain Monte Carlo method
  - Neural networks (NNs) [5]: use NNs that are FBA-based, recurrent, and time-varying
  - AMP / EP: our main focus, see recent arXiv paper

[1] D. Plabst, et al. IEEE Commun. Lett., 2020
[2] D. Plabst, et al. IEEE/OSA JLT, 2022
[3] T. Wettlin, et al. IEEE/OSA JLT 2020
[4] T. Prinz, et al. IEEE Trans. Commun. 2024
[5] D. Plabst, et al. IEEE Trans. Commun., 2025

Complexities

- Measure multiplications per symbol for M-ASK, n = # of symbols, K=channel memory, N_{iter} = # iterations
  - Linear: $\log_2 n + M$ where $\log_2 n$ follows by using FFT
  - FBA: M^K … exponential in K
  - Channel shortening & unit-memory FBA: $2 \cdot \log_2 n + M^2$
  - GS: $K^2 \cdot \log_2 M \cdot N_{iter} \cdot N_{par}$ with N_{par} parallel samplers
  - NNs: $L \cdot \ell^2$ with L layers of size ℓ nodes each
  - AMP: $N_{iter} \cdot \log_2 n + M$
- Only AMP achieves high rates with reasonable complexity for 4 km SSMF: (FBA,CS,GS,NNs) achieve at most (0.25,1,2,2) bpcu
- We normalize to multiplications per information bit (mpib) for the plots

Message Passing



AMP/EP

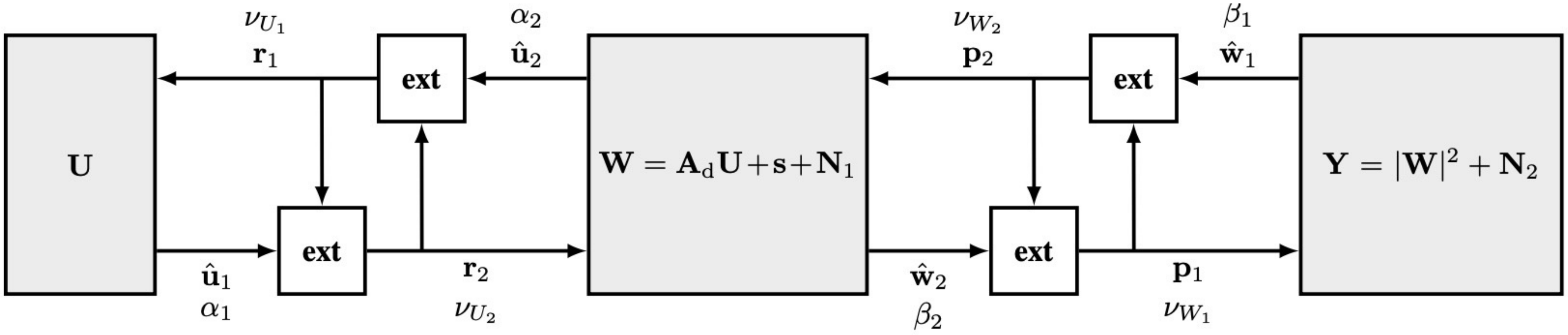
- Expectation Propagation (EP) approximates Belief Propagation
- We use EP/AMP with Gaussian messages (means and variances)
- System formula for short-reach links:

$$\mathbf{Y} = \big|\underbrace{\mathbf{A}_{\mathrm{d}}\mathbf{U} + \mathbf{S} + \mathbf{N}_1}_{:=\mathbf{W}}\big|^2 + \mathbf{N}_2$$

  where
  - **U** represents real data symbols (we consider one MZM only)
  - $\mathbf{A}_{\mathrm{d}}$ is a (complex convolution) matrix with the filter & CD responses
  - **S** is a function of the decoded symbols from previous SIC stages
  - $\mathbf{N}_1$ and $\mathbf{N}_2$ represent optical and electrical noise, respectively
  - the squaring represents the DD operation
- The vector **W** is the output of a linear channel

AMP/EP (2)

- System formula: $\mathbf{Y} = |\underbrace{\mathbf{A}_\mathrm{d}\mathbf{U} + \mathbf{S} + \mathbf{N}_1}_{:= \mathbf{W}}|^2 + \mathbf{N}_2$
- System graph:



- AMP/EP updates messages of vector means and scalar variances
- Outer boxes use memoryless non-linear operations
- Main simplification: the middle box computes Gaussian messages via linear minimum mean square error (LMMSE) estimation
- Most complexity is due to the convolution: use FFTs to simplify
- Terminology: the gray boxes are input, LMMSE, and output *denoisers*

AMP/EP (3)

Generalized vector AMP (GVAMP) algorithm from [1]

Algorithm 1 GVAMP for SIC stage ℓ

Initialization: extrinsic $(\mathbf{p}_1, \nu_{W_1})$.

repeat

▷ *Output Denoiser:*

Calculate $(\hat{\mathbf{w}}_1, \beta_1)$ (143)-(145).

Convert to extrinsic (75b).

▷ *LMMSE Denoiser:*

Calculate extrinsic $(\mathbf{r}_1, \nu_{U_1})$ (98)-(99).

▷ *Input Denoiser:*

Calculate $(\hat{\mathbf{u}}_1, \alpha_1)$ (64)-(65).

Convert to extrinsic (75a).

▷ *LMMSE Denoiser:*

Calculate $(\hat{\mathbf{w}}_2, \beta_2)$ (68)-(86).

Convert to extrinsic (81b).

until convergence

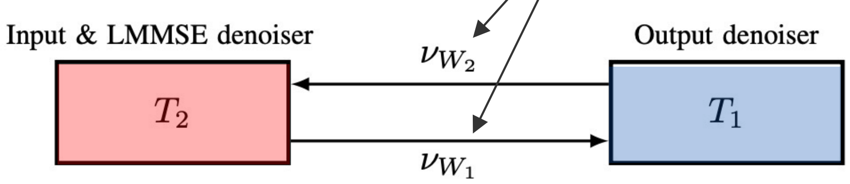
Compute decoding metrics $Q_{\ell,t}(v_{\ell,t})$, $t \in [\![N]\!]$ (84).

Return $Q_{\ell,t}(v_{\ell,t})$.

1^{st} half-iteration.

2^{nd} half-iteration.

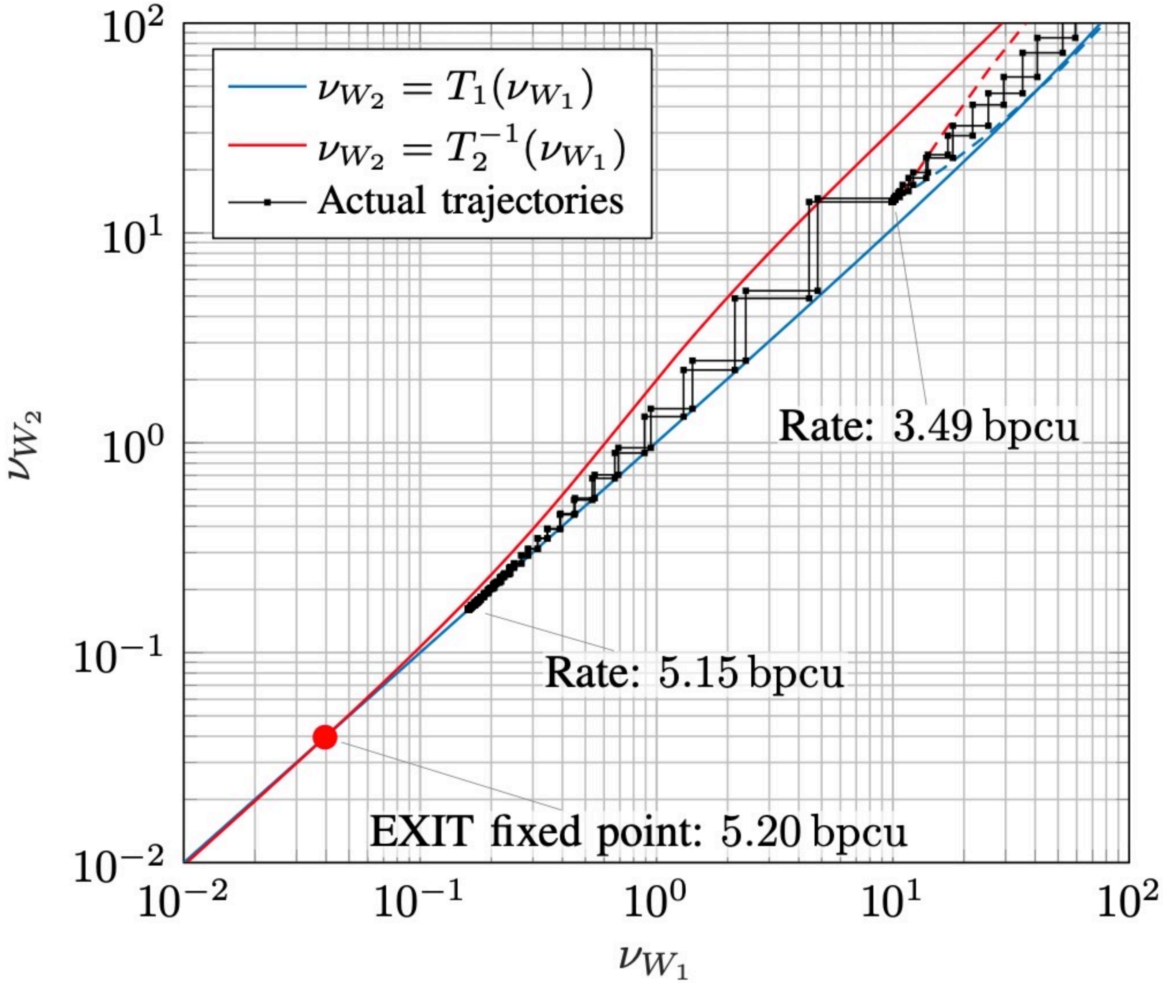
EXIT chart tracks variances

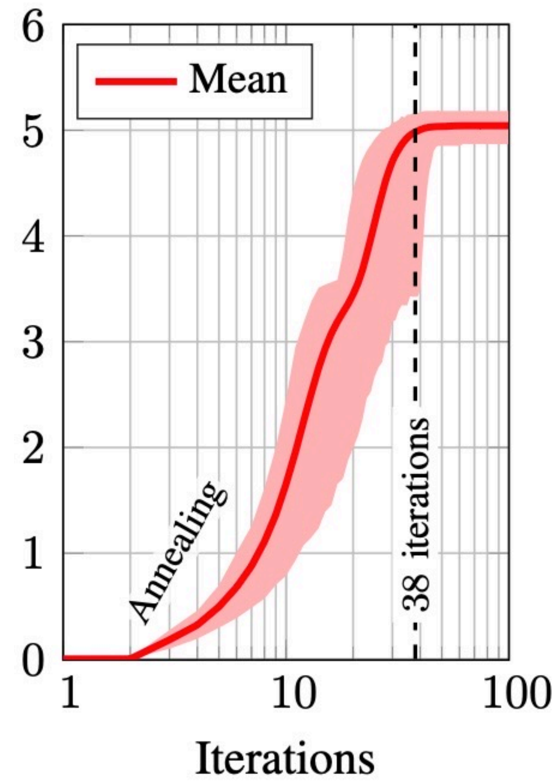


[1] D. Plabst, et al. arXiv, 2025

AMP/EP (4)

- Convergence can be tracked via EXIT charts:
  T_1 output denoiser, T_2 combined input & LMMSE denoiser
- Example: bipolar 64-ASK for 4 km SSMF achieves > 5 bits per channel use (bpcu) at 300 Gbaud (> 1.5 Tbit/s) with ≈ 38 iterations



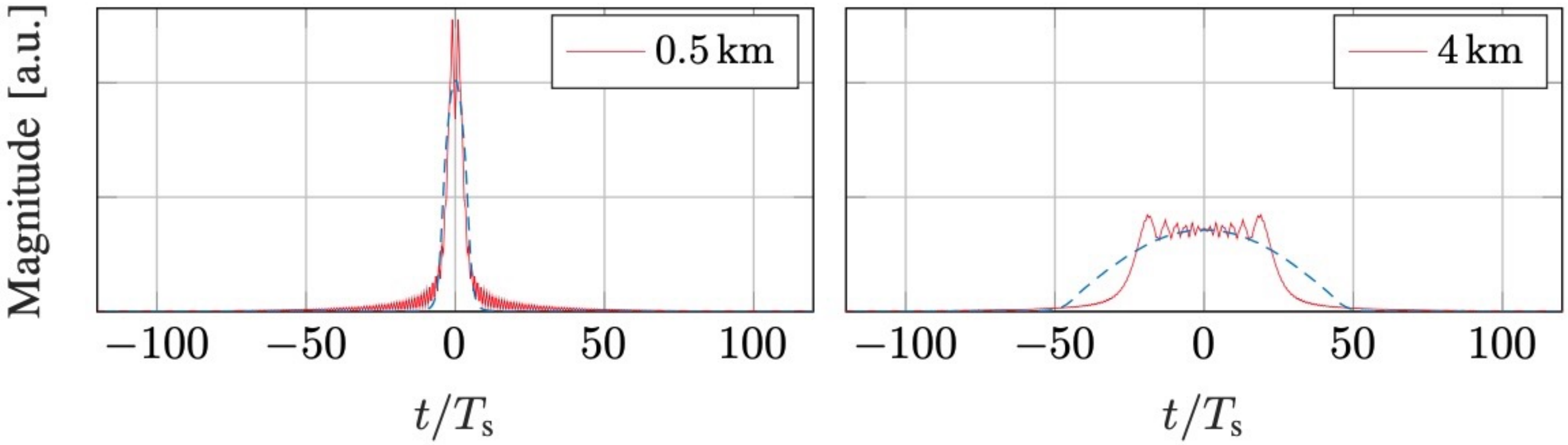


3) AIRs from Numerical Experiments

- Transmit 128 blocks, each having n=2048 input symbols
- System parameters: 4 km of SSMF in the C-band

Carrier wavelength	1550 nm (C-band)
Symbol rate	$B = 300\,\mathrm{GBd}$
Fiber length	4 km
Attenuation factor	0.2 dB/km (or $0.046\,\mathrm{km}^{-1}$)
Group velocity dispersion	$\beta_2 = -2.168 \times 10^{-23}\,\mathrm{s}^2/\mathrm{km}$
DAC	FD-RRC with $\alpha_{\mathrm{tx}} = 1\%$
TX bandwidth	$(1 + \alpha_{\mathrm{tx}})B$
SSMF response	$H(f) = \exp\left(-\mathrm{j}\,\beta_2/2\,\omega^2 L_{\mathrm{fib}}\right)$
Pre-DD complex AWGN	See (29), ACF (30a)
Post-DD real AWGN	See (29), ACF (30b)
ADC oversampling factor	$N_{\mathrm{os}} = 2$
Receive filter	$g_{\mathrm{rx}}(t) = B_{\mathrm{rx}}\,\mathrm{sinc}(B_{\mathrm{rx}} t)$
Precoder	Random orthogonal (FFT-based)

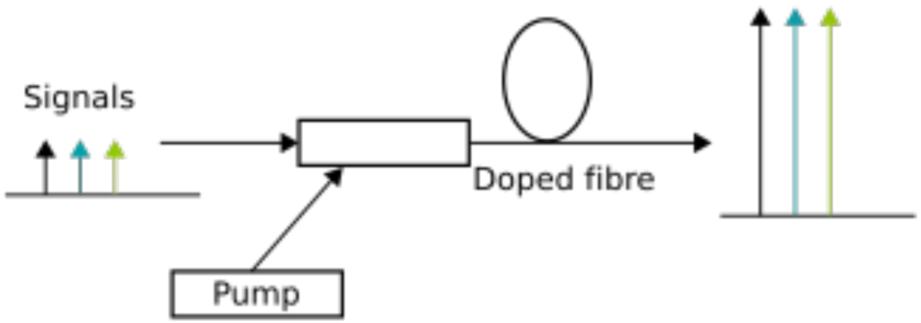
Channel Response & Modulation



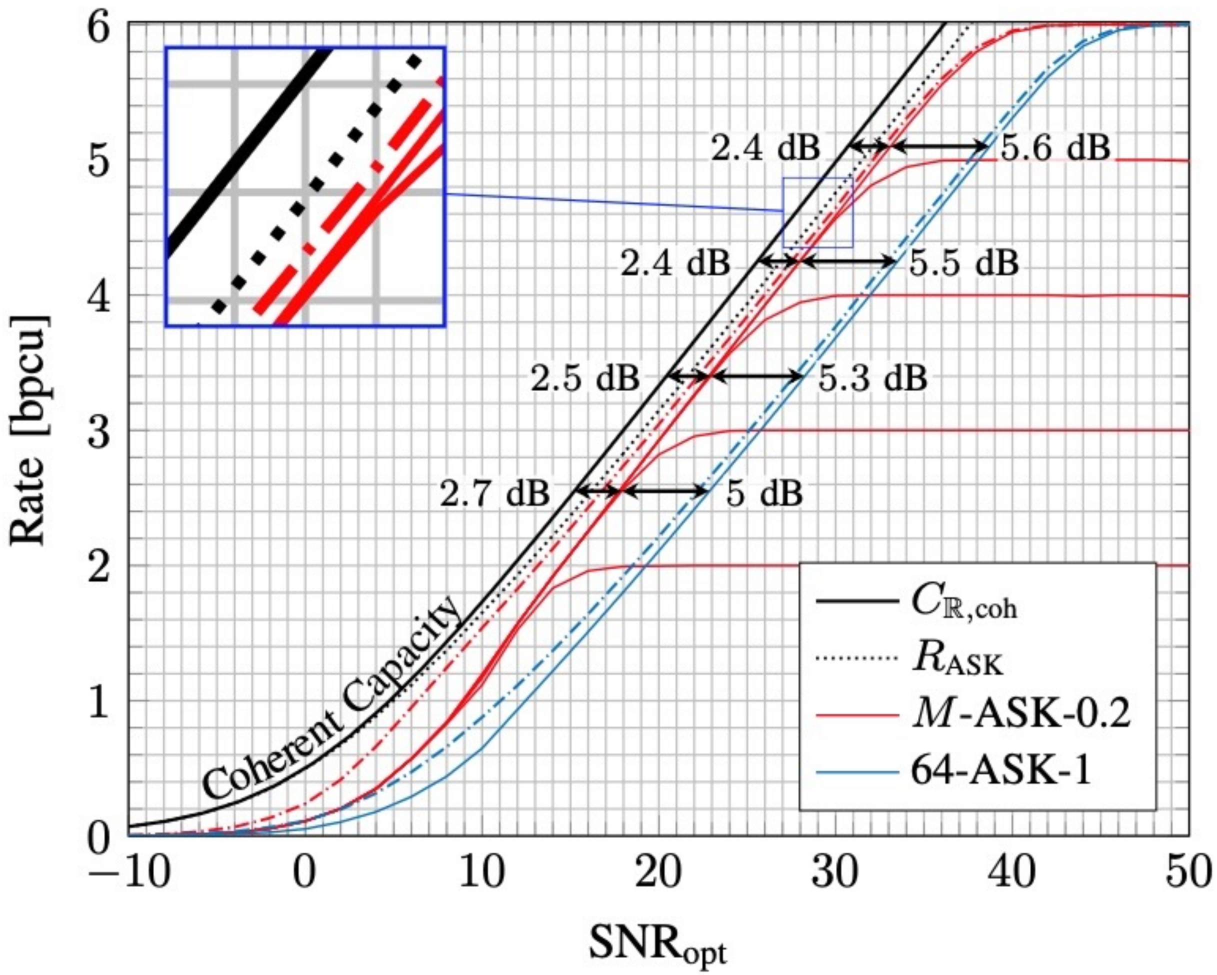
- Red & blue: filter & CD response magnitude for roll-offs 1% & 99%
- Approximate with 501 taps to capture >99.9% of its energy
- Real M-ASK-o alphabet {±1, ±3,…, ±(M-1)}/(M-1) + o and offsets o = 0.2, 0.25, 1
- Classic: o=0 is bipolar-ASK and o=1 is unipolar-ASK
- Number of encoding/decoding stages: S=1-4
- NNs: L=3-5 layers, ℓ=32-300 nodes/layer, see [1] for more information

[1] D. Plabst, et al. IEEE Trans. Commun., 2025

Optical Amplifiers

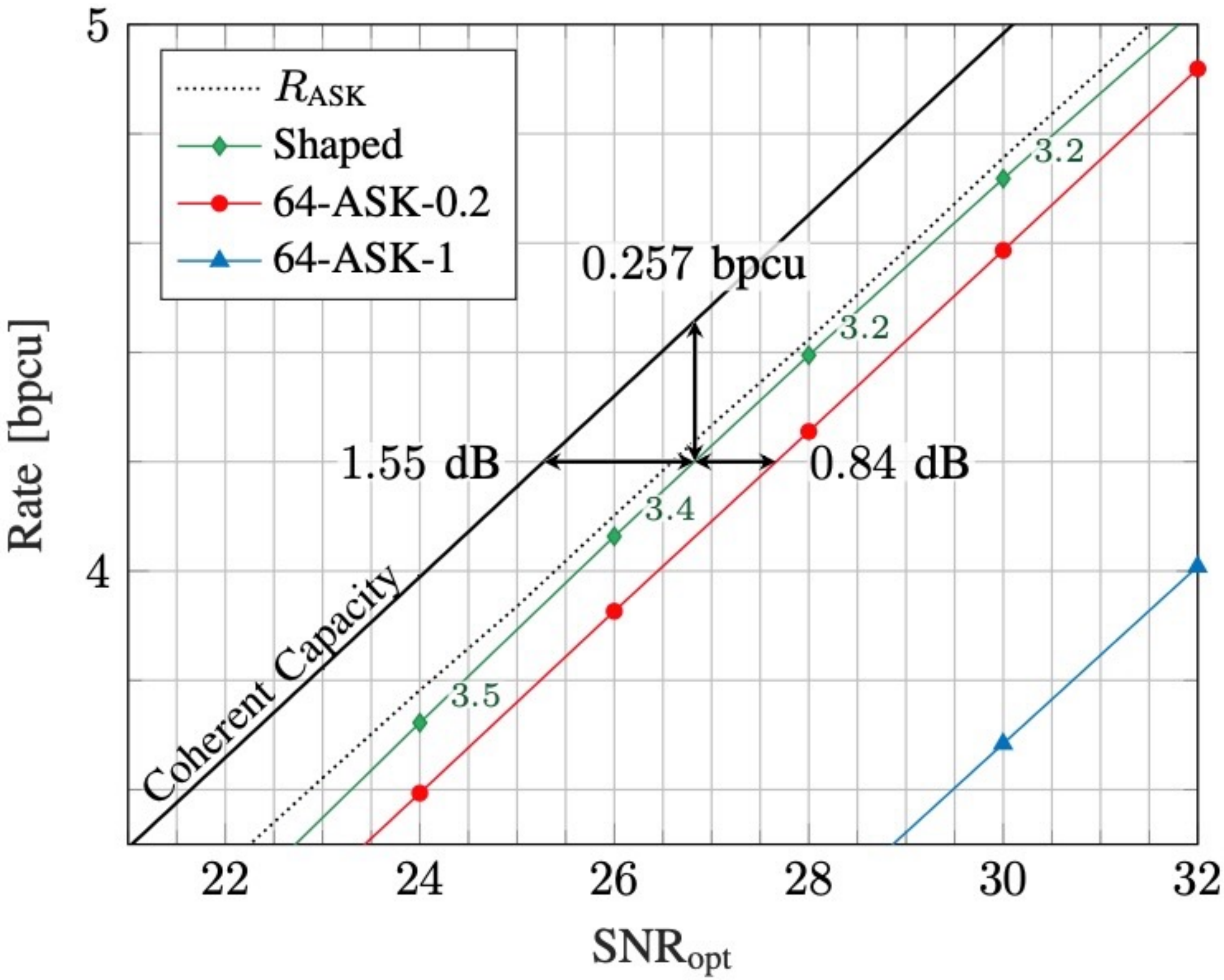


Optically Amplified Links



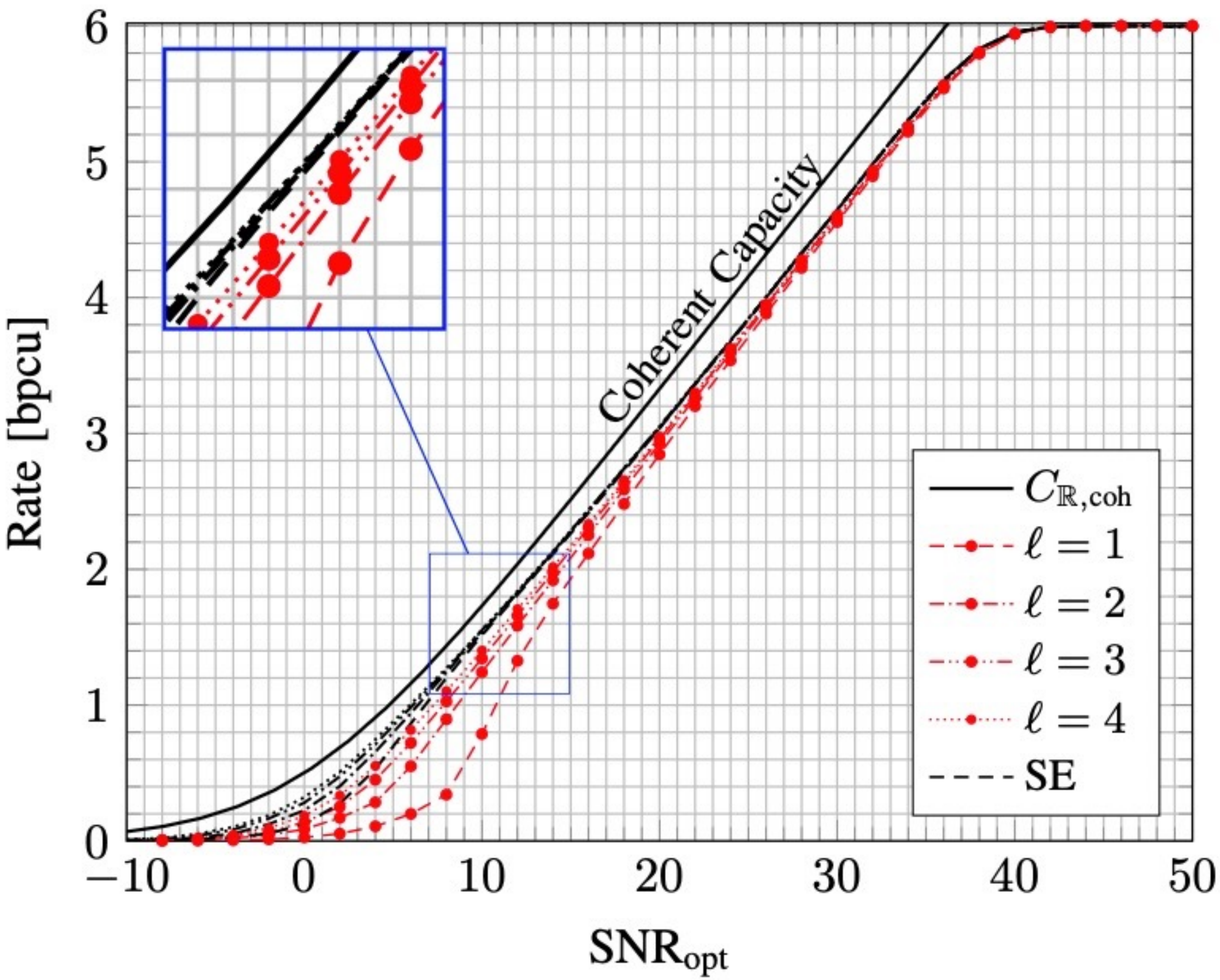
- S=4 stages, M=4,8,16,32,64, offsets o=0.2,1
- $C_{\mathbb{R},coh}$ is the coherent capacity with probabilistic shaping
- R_{ASK} is the equiprobable-ASK capacity with a coherent receiver
- Dash-dotted curves show the EXIT rate predictions
- Note the large gains of 5-6 dB using bipolar ASK

Optically Amplified Links: Add Shaping



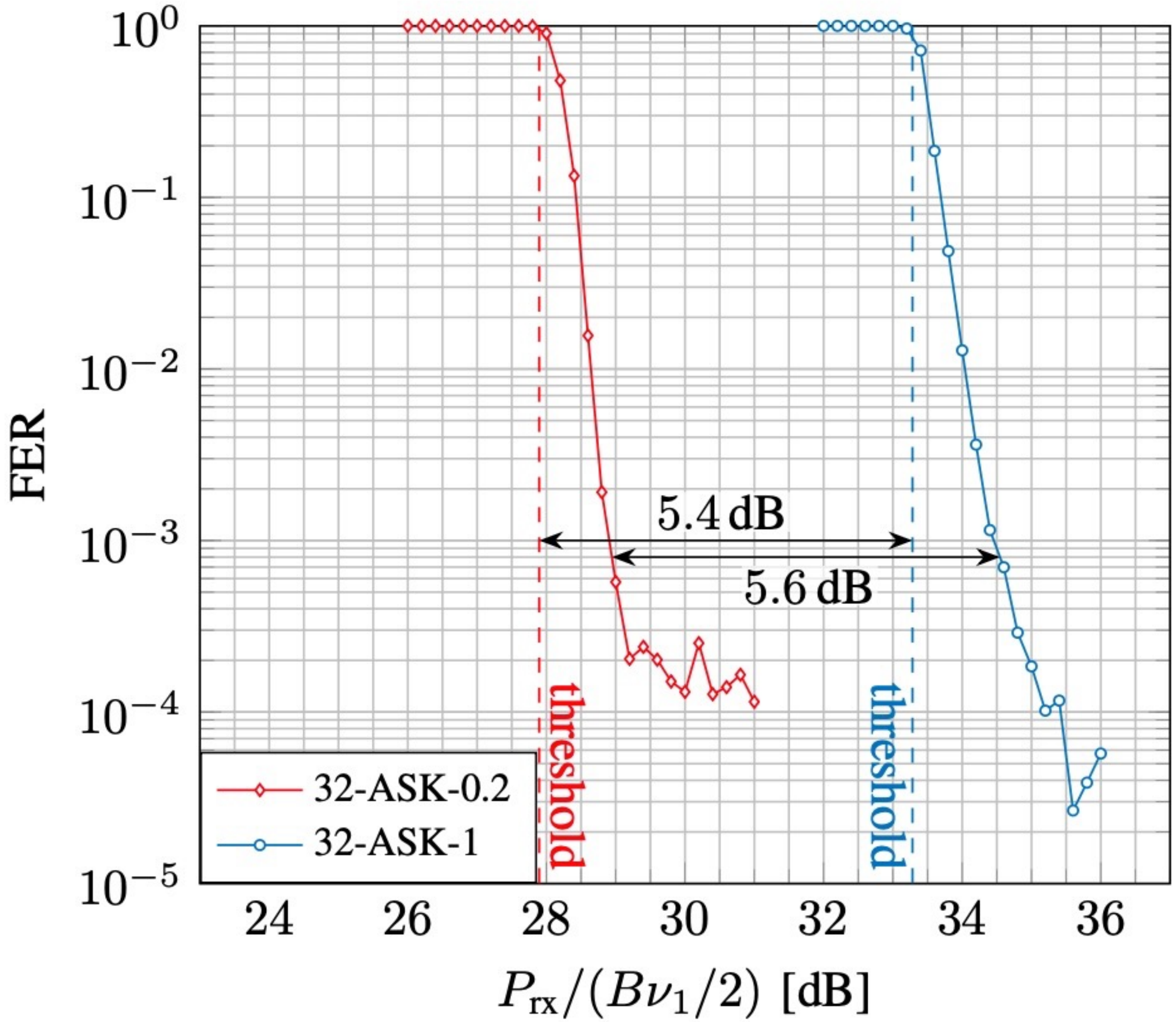
- S=4 stages, 64-ASK-o with o=0.2,0.25,1
- Gaussian-like shaping with offset o=0.25 gains >6 dB at high rates
- Rates are within 0.3 bpcu of the coherent capacity

Optically Amplified Links: SIC Gains



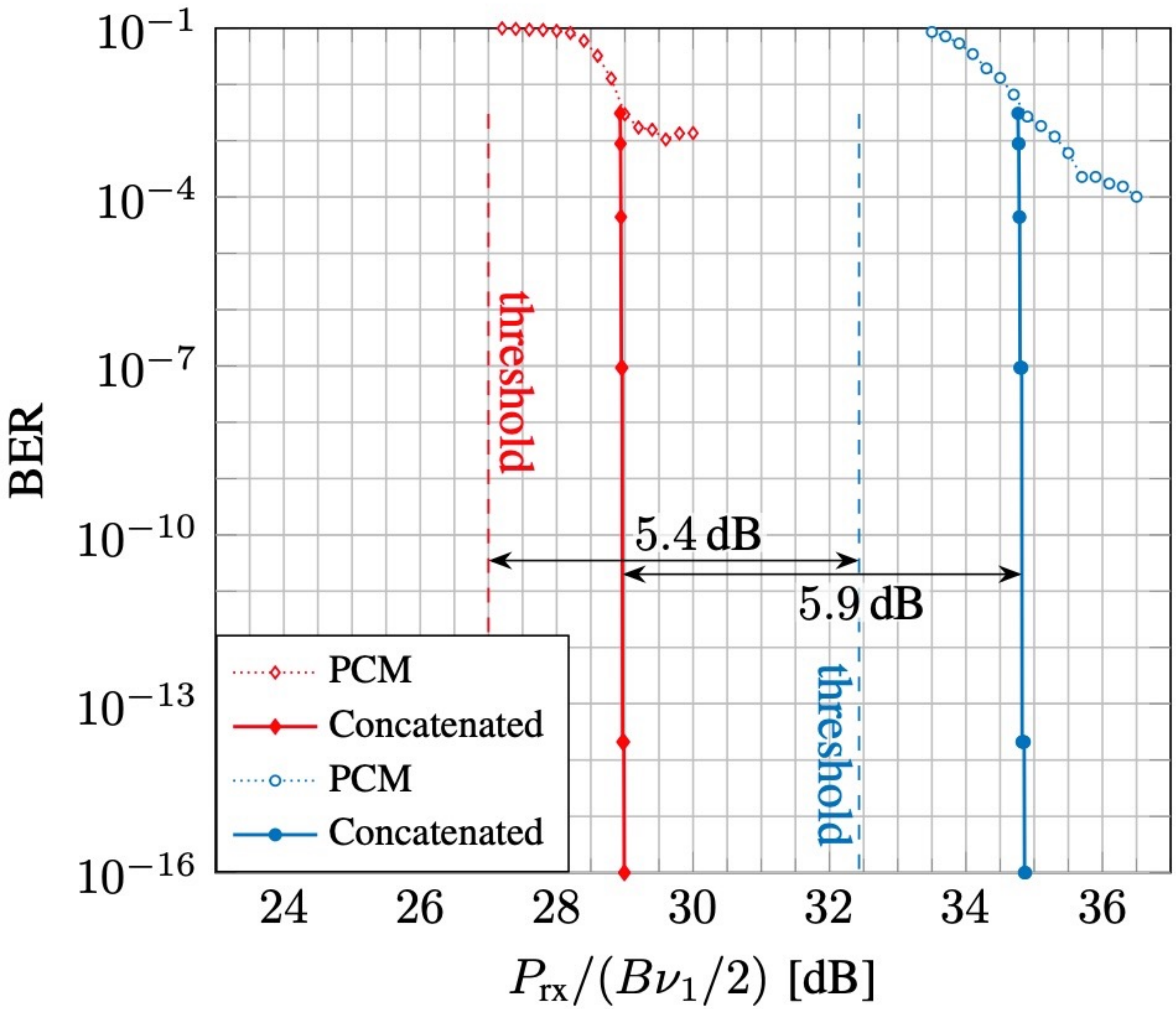
- S = 1 to 4 stages, 64-ASK-0.2
- Black dotted curves show the EXIT rate predictions
- Note that S=1 stage suffices for R > 3 bpcu

Polar Coded Modulation



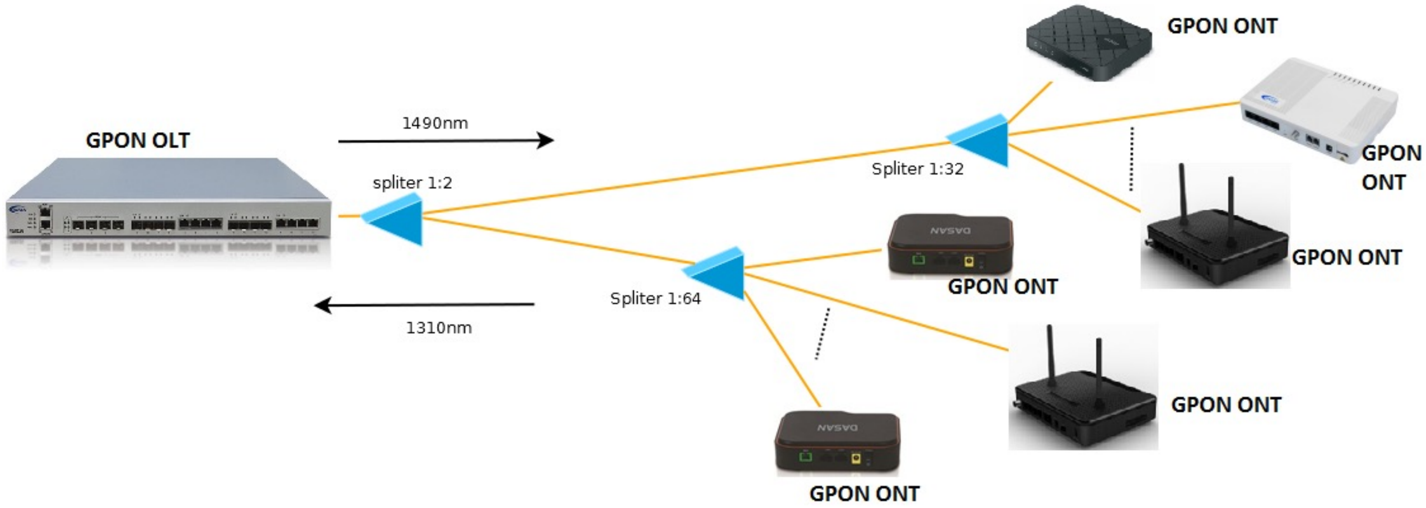
- Polar coded modulation with 5 bit levels, rate 4.2 bpcu
- Polar codes optimized with Monte Carlo methods
- 16-bit CRC, SCL decoding list size 32, list-passing across levels

Add Outer BCH-BCH Product Code

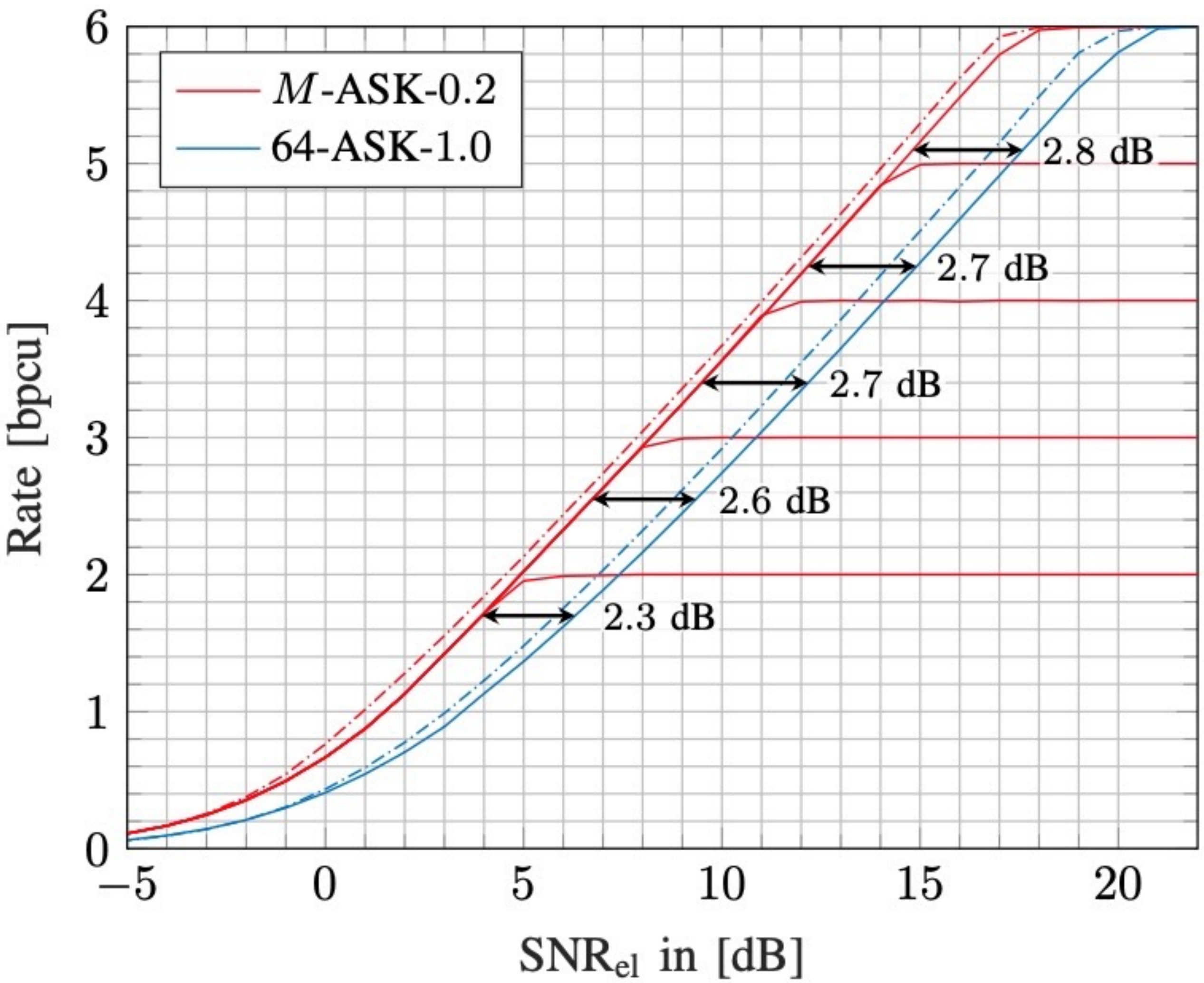


- ITU-T G.975.1 BCH-BCH outer code, hard decoding
- 6.9% overhead, giving overall rate 3.94 bpcu

PON

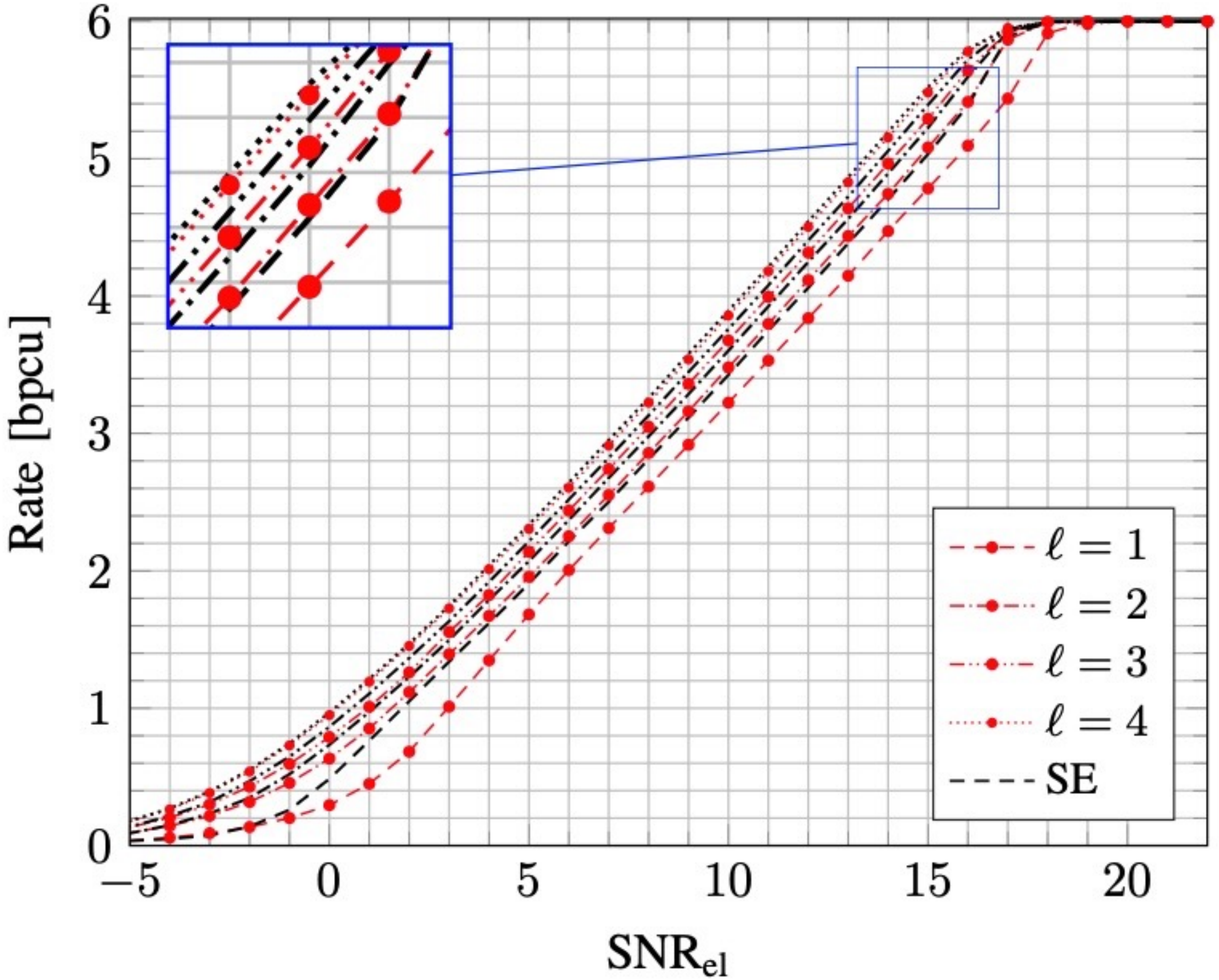

GPON OLT
1490nm
spliter 1:2
1310nm
Spliter 1:64
Spliter 1:32
GPON ONT
GPON
ONT
GPON ONT
GPON ONT
GPON ONT
GPON ONT

Unamplified Links



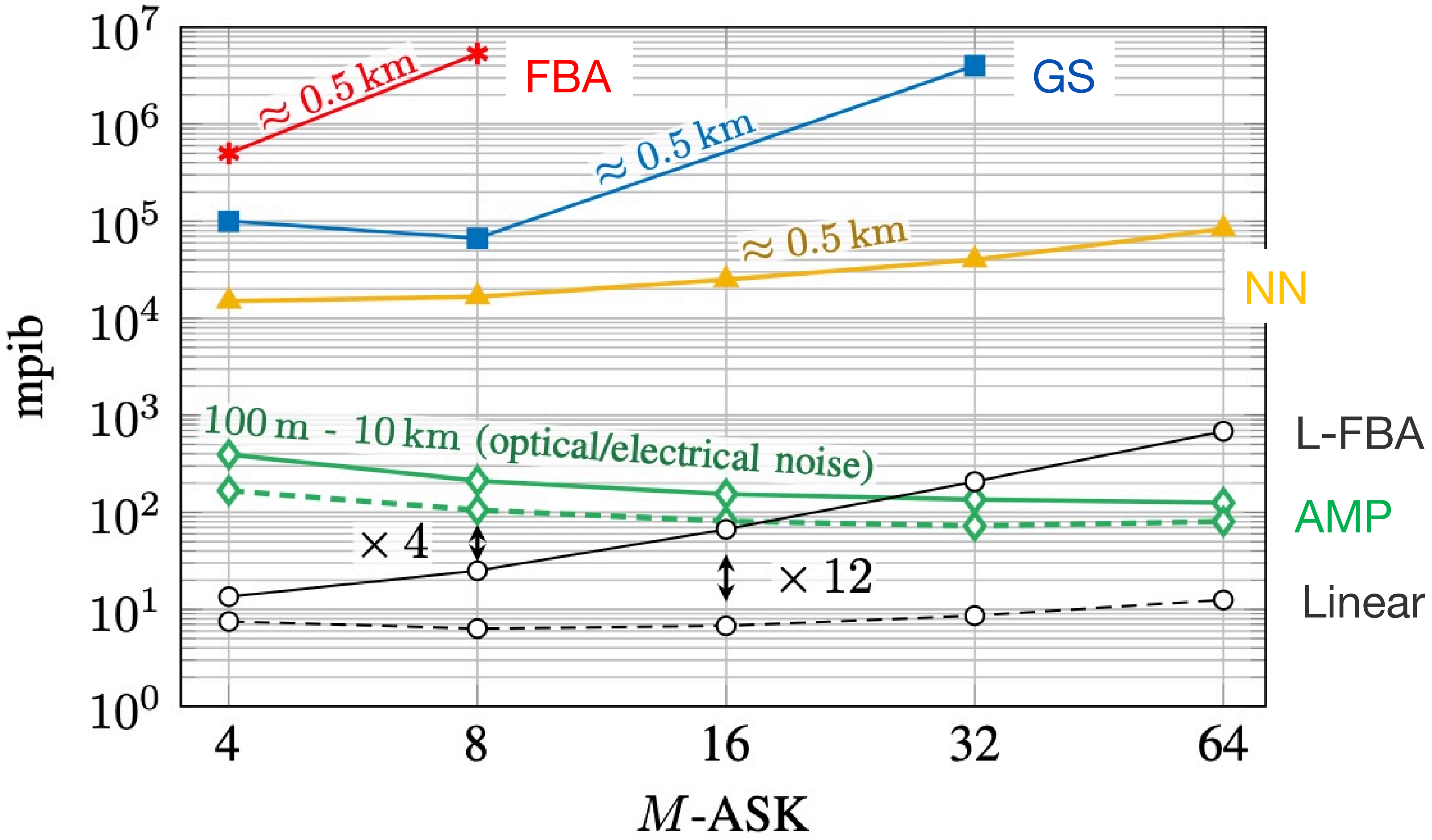
- S=4 stages, 64-ASK-o with o=0.2,1
- Dash-dotted curves show the EXIT rate predictions
- Note the large gains of 2-3 dB using bipolar ASK

Unamplified Links: SIC Gains



- S = 1 to 4 stages, 64-ASK-0.2
- Black dotted curves show the EXIT rate predictions
- Note: multi-stage encoding/decoding now gains 2-4 dB

Complexity vs. M



- mpib for S=1; the AMP curve is for optical amplification
- Note 1: NNs reduce complexity as compared to FBA and GS
- Note 2: only AMP achieves high rates with reasonable complexity beyond 1 km

Summary

Machine learning equalizers outperform classic equalizers:

- NNs reduce complexity as compared to FBA and GS
- AMP/EP outperforms NNs and other algorithms by a wide margin

Main Message

AMP/EP is the current best option for nonlinearities*

- High Rate: approach JDD rates
- Lowest Complexity: as compared to other algorithms
- Flexibility: multi-stage encoding/decoding uses classic coded modulation and fits well with multi-level coding

* also for nonlinear wireless and many other channels!