Achievable Rate Estimates from Fiber-Optic Transmission Experiment Measurements

Georg Böcherer

Institute for Communications Engineering
Technical University of Munich
georg.boecherer@tum.de

Sixth Van Der Meulen Seminar: Optical Communications
15. September 2016
Eindhoven University of Technology, The Netherlands

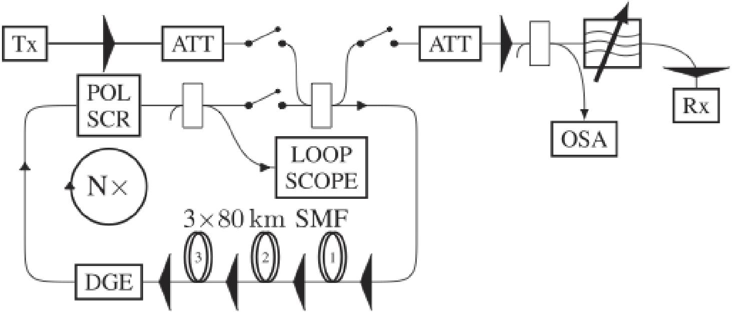
References

- [1] F. Buchali, G. Böcherer, W. Idler, L. Schmalen, P. Schulte, and F. Steiner, Experimental demonstration of capacity increase and rate-adaptation by probabilistically shaped 64-QAM, in European Conference on Optical Communication (ECOC), Valencia, Spain, 2015, paper **PDP3.4**.
- [2] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration, J. Lightw. Technol., Apr. 2016.

Outline

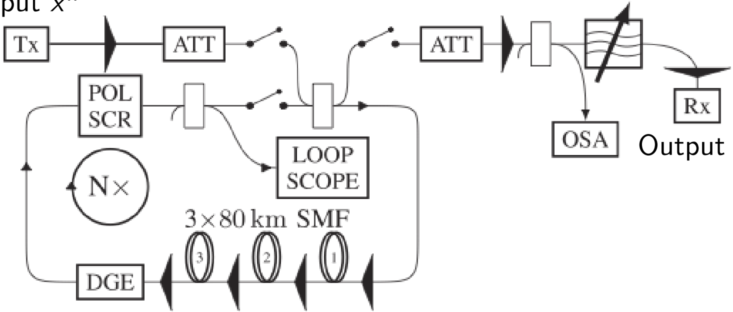
Outline

Experimental Setup



Experimental Setup: Interfaces

Input x^n

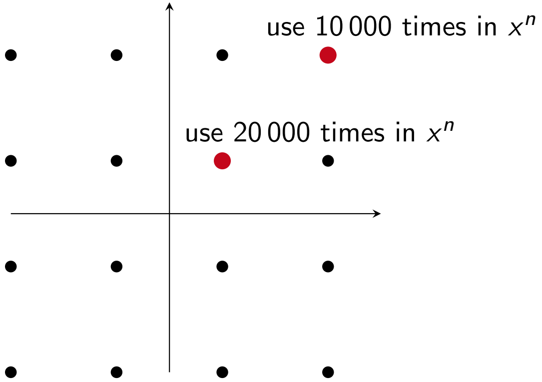


Output y^n

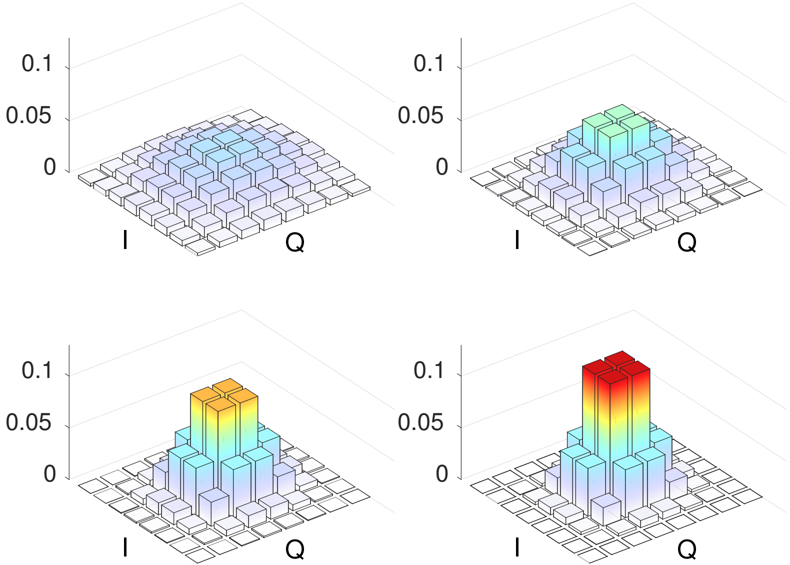
Input Sequences

We provide input sequences to the experimentalist:

- $n = 200\,000$ QAM symbols $x^n = x_1, x_2, \ldots, x_n$.
- **Probabilistic shaping:** use outer QAM symbols less often:



Example 64-QAM Distributions P_X



Measurement Campaign

- Measurement campaign with provided **input sequences** x^n.
- For each measurement, a noisy output sequences $y^n = y_1, y_2, \ldots, y_n$ is stored.
- We get a data set with the **noisy output sequences** y^n.

Outline

Reliable Transmission

- The receiver must recover input sequence x^n from output sequence y^n.
- An **achievable information rate** (AIR) indicates if recovery is possible.

Channel Coding Theorem (Achievability Part)

- Memoryless channel $p_{Y|X}$.
- Random code $\mathcal{C} = \{X^n(1), \ldots, X^n(2^{nR})\}$ with entries iid $\sim P_X$.
- Message $w \in \{1, 2, \ldots, 2^{nR}\}$
- ML decoder

$$\hat{W} = \operatorname*{argmax}_{w} p_{Y^n|X^n}(Y^n|X^n(w)) = \prod_{i=1}^{n} p_{Y|X}(Y_i|X_i(w))$$

- Error probability $\Pr(W \neq \hat{W}) \overset{n\to\infty}{\to} 0$ if

$$R < \mathrm{I}(X; Y).$$

Estimating Mutual Information

- Mutual information:

$$\mathrm{I}(X;Y) = \mathrm{E}\left[\log \frac{p_{Y|X}(Y|X)}{p_Y(Y)}\right].$$

- Calculation by Monte-Carlo simulation: Sample sequences x^n, y^n of n independent channel uses.
- Weak law of large numbers:

$$\mathrm{I}(X;Y) \approx \hat{\mathrm{I}}(x^n;y^n) := \frac{1}{n}\sum_{i=1}^{n} \log \frac{p_{Y|X}(y_i|x_i)}{p_Y(y_i)}.$$

Measurements

- $\hat{\mathrm{I}}$ can be calculated also when x^n, y^n are measurements of some channel.
- The memoryless channels $p_{Y|X}$ is now an auxiliary channel of our choice.

Auxiliary AWGN Channel

- Input distribution is $P_{X^n} = P_X^n$.
- Memoryless auxiliary output channel $p_{Y^n|X^n} = p_{Y|X}^n$
- Auxiliary I/O relation

$$Y = h \cdot X + Z$$

  with $Z \sim \mathcal{N}(0, \sigma^2)$.
- Choose h, σ^2 maximizing $\hat{\mathrm{I}}$.

What is the operational meaning of $\hat{\mathrm{I}}$ now?

Assumptions of channel coding theorem not fulfilled:

- We don't know the channel $p_{Y^n|X^n}$.
- The "true" channel very likely has memory.

Mismatched Decoding

- [3] G. Kaplan and S. Shamai (Shitz), Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment, AEÜ, vol. 47, no. 4, pp. 228–239, 1993.
- [4] A. Ganti, A. Lapidoth, and E. Telatar, Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit, IEEE Trans. Inf. Theory, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.
- [5] G. Böcherer, Achievable rates for shaped bit-metric decoding, arXiv preprint, 2015. [Online]. Available: http://arxiv.org/abs/1410.8075

LM Rate

- Random code ensemble $\sim P_{X^n}$.
- Auxiliary channel $q(\cdot|\cdot)$
- Decoder $\hat{W} = \operatorname{argmax}_w q(\tilde{Y}^n|X^n(w))$.
- Achievable rate

$$\mathsf{R}_{\mathsf{LM}} = \text{p-}\liminf_{n\to\infty} \underbrace{\frac{1}{n}\log\frac{q(\tilde{Y}^n|X^n)^s r(X^n)}{q_s(\tilde{Y}^n)}}_{=:\hat{\mathsf{R}}_{\mathsf{LM}}}$$

where

- Auxiliary output distribution $q_s(\cdot) = E[q(\cdot|X^n)^s r(X^n)]$
- $s \geq 0$.
- Function r: $\frac{1}{n}\log[r(X^n)] \overset{n\to\infty}{\to} E\{\frac{1}{n}\log[r(X^n)]\}$.

Mutual Information Estimate

- $\hat{\mathrm{I}} = \hat{\mathrm{R}}_{\mathrm{LM}}$ for $s = 1, r(x) = 1$, and

$$q = p^n_{Y|X}$$

$\Rightarrow$ $\hat{\mathrm{I}}$ is an achievable rate lower bound for a decoder assuming memoryless channel $q = p^n_{Y|X}$.

- Optimizing over s, r may improve the bound.

Discussion: Signal-to-Noise Ratio

- "Signal-to-noise ratio" is

$$\mathsf{SNR} = \frac{h^2 \, \mathsf{E}[|X|^2]}{\sigma^2}$$

- Depends on our model parameters h, σ^2.
- Can be very different from OSNR measured by the spectral analyzer.

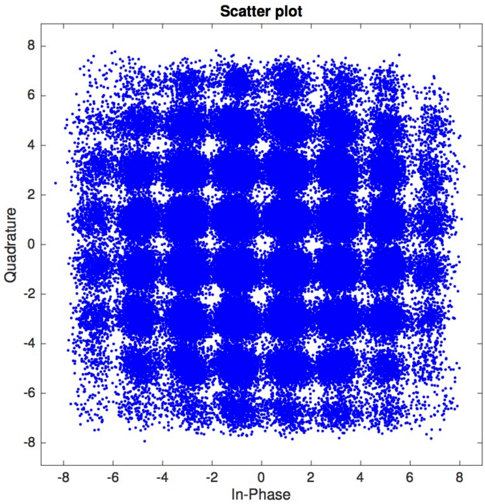
Discussion: Parameter Choice at the Receiver

- MMSE estimate:

$$h = \frac{\boldsymbol{x}\boldsymbol{y}^{\mathrm{H}}}{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}}, \quad \sigma^2 = \frac{1}{n}(\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}} - |h|^2\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}})$$

- Problem: the receiver does not know $x^{\boldsymbol{n}}$.

Discussion: Scatterplot



Discussion: Blind Estimation of h, σ^2

- **Blind:** use only y^n to optimize h.
- Approach:

$$\mathrm{D}(p_Y \| q_Y) \geq 0 \Rightarrow \mathrm{E}\left[\log_2 \frac{1}{q_Y(Y)}\right] \geq \mathrm{H}(Y).$$

$\Rightarrow$ minimize expectation over q_Y.

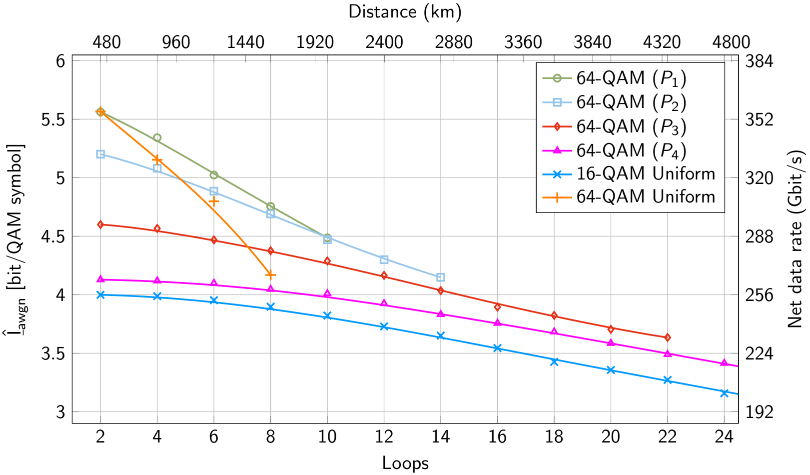
- Choose

$$q_{Y_h}(y) = \sum_{x \in \mathcal{X}} P_X(x) P_{Z_h}(y - h \cdot x)$$

where $Z_h \sim \mathcal{N}\left(0, \frac{\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}}{n} - |h|^2 \mathrm{Var}(X)\right)$,

$$h_{\text{blind}} = \underset{h}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{1}{q_{Y_h}(y_i)}$$

AIR Estimates[1]



[1] [2] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration, J. Lightw. Technol., vol. 34, no. 8, Apr. 2016.

Outline

Input Distribution P_X

- **Symmetry:**

$$P_X(x) = P_X(-x).$$

- **Amplitude-Sign Factorization:**

$$P_X(x) = P_A(|x|)P_S(\text{sign}(x))$$

where $A := |X|$ and $S := \text{sign}(X)$.

- **Uniform sign:**

$$P_S(-1) = P_S(1) = \frac{1}{2}.$$

Generation of Amplitude Sequence

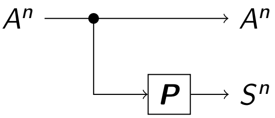
Constant Composition Distribution Matching (CCDM) [2]:

$$D^k \text{—} \boxed{\text{CCDM}} \text{—} A^n$$

- Data bits D_i iid Bernoulli(1/2).
- $A_i \sim P_A$.
- Rate is $k/n = \mathrm{H}(A)$.
- invertible: D^k can be recovered from A^n with zero error.

[2] [6] P. Schulte and G. Böcherer, Constant composition distribution matching,' IEEE Trans. Inf. Theory, vol. 62, no. 1, pp. 430–434, Jan. 2016.

Shaping and Channel Coding



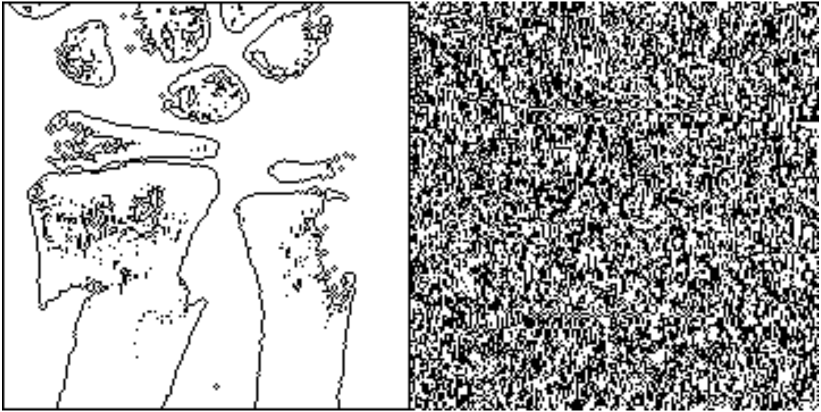
- Binary systematic rate $(m-1)/m$ generator matrix $\boldsymbol{G} = [\boldsymbol{I}|\boldsymbol{P}]$.
- Binary amplitude representation $\boldsymbol{b}(A_i) \in \{0,1\}^{m-1}$.
- Binary sign representation $b(S_i) \in \{0,1\}$.
- $b(S)^n = \boldsymbol{b}(A)^n \boldsymbol{P}$.
- **Assumption:** S_i is approximately uniformly distributed.

$$A_i S_i \sim P_X.$$

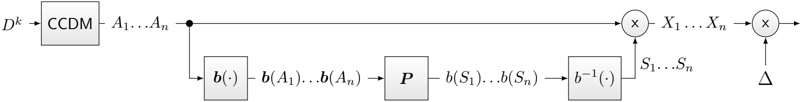
Uniform Check Bit Assumption: Example

DVB-S2 rate 1/2 LDPC code

- **Data:** empirical distribution $P_D(1) = 1 - P_D(0) = 0.1082$.
- **Check bits:** empirical dist. $P_R(1) = 1 - P_R(0) = 0.4970$.



Probabilistic Amplitude Shaping (PAS)[3]



$$P_{S_i \cdot A_i} = P_X$$

[3] [7] G. Böcherer, F. Steiner, and P. Schulte, Bandwidth efficient and rate-matched low-density parity-check coded modulation, IEEE Trans. Commun., Dec. 2015.

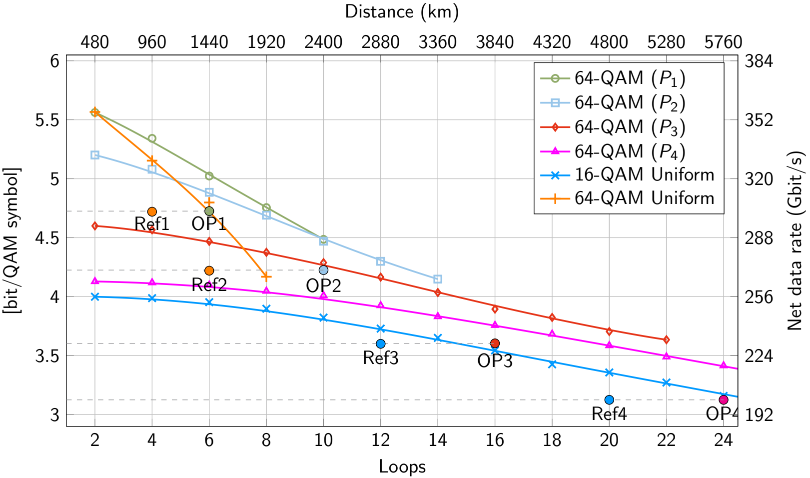
Receiver

- Binary label of X is $b(S)\boldsymbol{b}(A) =: B_1 B_2 \cdots B_m$
- The demapper calculates bitwise soft-information

$$L_j = \log \frac{P_{B_j}(0)}{P_{B_j}(1)} + \log \frac{p_{Y|B_j}(Y|0)}{p_{Y|B_j}(Y|1)}, \quad j = 1, \ldots, m.$$

- No iterative demapping.

Numerical Results: Operating Points FER = "0"



Conclusions

- Achievable Information Rates as interface between fiber-optic transmission experiment and coded modulation system design.
- Probabilistic shaping opportunities: reach & rate increase.
- Experimental work is rewarding.

References I

[1] F. Buchali, G. Böcherer, W. Idler, L. Schmalen, P. Schulte, and F. Steiner, "Experimental demonstration of capacity increase and rate-adaptation by probabilistically shaped 64-QAM," in *European Conference on Optical Communication (ECOC)*, Valencia, Spain, 2015, paper PDP3.4.

[2] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 8, Apr. 2016.

[3] G. Kaplan and S. Shamai (Shitz), "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEÜ*, vol. 47, no. 4, pp. 228–239, 1993.

[4] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[5] G. Böcherer, "Achievable rates for shaped bit-metric decoding," *arXiv preprint*, 2015. [Online]. Available: http://arxiv.org/abs/1410.8075

References II

[6] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 430–434, Jan. 2016.

[7] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.