# Recent lower bounds on the capacity of fiber-optical links

**Erik Agrell[1], Tobias A. Eriksson[2], Naga V. Irukulapati[1], and Kamran Keykhosravi[1]**

[1]Department of Signals and Systems
[2]Department of Microtechnology and Nanoscience
Chalmers University of Technology, Sweden

*Munich Workshop on Information Theory of Optical Fiber, Munich, Germany, Dec. 7, 2015*

CHALMERS

FORCE
FIBER-OPTIC COMMUNICATIONS RESEARCH CENTER

Vetenskapsrådet

# Contents



*in collaboration with*

Alex Alvarado
Peter A. Andrekson
Giuseppe Durisi
Tobias Fehenberger
Norbert Hanik
Pontus Johannisson
Magnus Karlsson
Marco Secondini
Henk Wymeersch

*whose contributions are gratefully acknowledged*

# 1. Lower bounds from experiments

$$I(\boldsymbol{X};\boldsymbol{Y}) = \mathbb{E}\left[\log \frac{p(\boldsymbol{Y}|\boldsymbol{X})}{p(\boldsymbol{Y})}\right] \geq \mathbb{E}\left[\log \frac{q(\boldsymbol{Y}|\boldsymbol{X})}{q(\boldsymbol{Y})}\right]$$

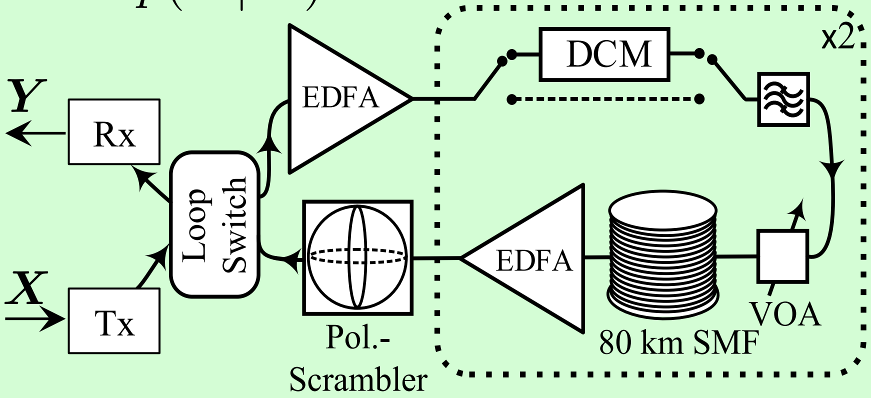
*Mutual information* (MI), cannot be calculated because $p(\boldsymbol{Y}|\boldsymbol{X})$ is unknown

*Achievable information rate* (AIR), calculated by Monte Carlo estimation from experimental data

**Tobias A. Eriksson *et al.***

## Recirculating loop experiment

$\Rightarrow p(\boldsymbol{Y}|\boldsymbol{X})$




## Mismatched Decoding [1]

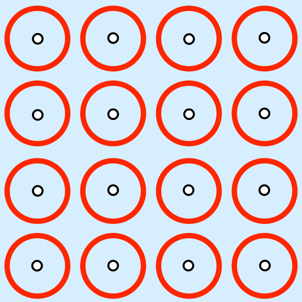
- Lower-bounding the MI by assuming an *auxiliary channel* $q(\boldsymbol{Y}|\boldsymbol{X})$
- Acheivable by a receiver designed for $q(\boldsymbol{Y}|\boldsymbol{X})$
- This study: comparison of different assumptions for $q(\boldsymbol{Y}|\boldsymbol{X})$
- Symbol-by-symbol receivers only

[1] Arnold *et al.*, *TIT*, 2006

# Channel distribution models

## 2D independent, identically distributed Gaussian (2D-iidG)

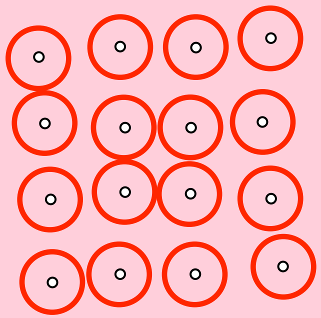
- 2D Gaussian $q(\boldsymbol{Y}|\boldsymbol{X})$
- Diagonal covariance
- Fixed means = $\boldsymbol{X}$



- The baseline $q(\boldsymbol{Y}|\boldsymbol{X})$, typically assumed in today's receivers

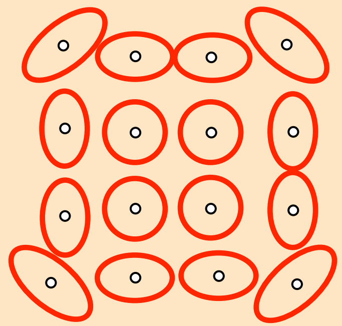
## 4D independent, identically distributed Gaussian (4D-iidG)

- 4D Gaussian $q(\boldsymbol{Y}|\boldsymbol{X})$
- Diagonal covariance
- Optimized mean for each $\boldsymbol{X}$



## 4D correlated Gaussian (4D-CG)

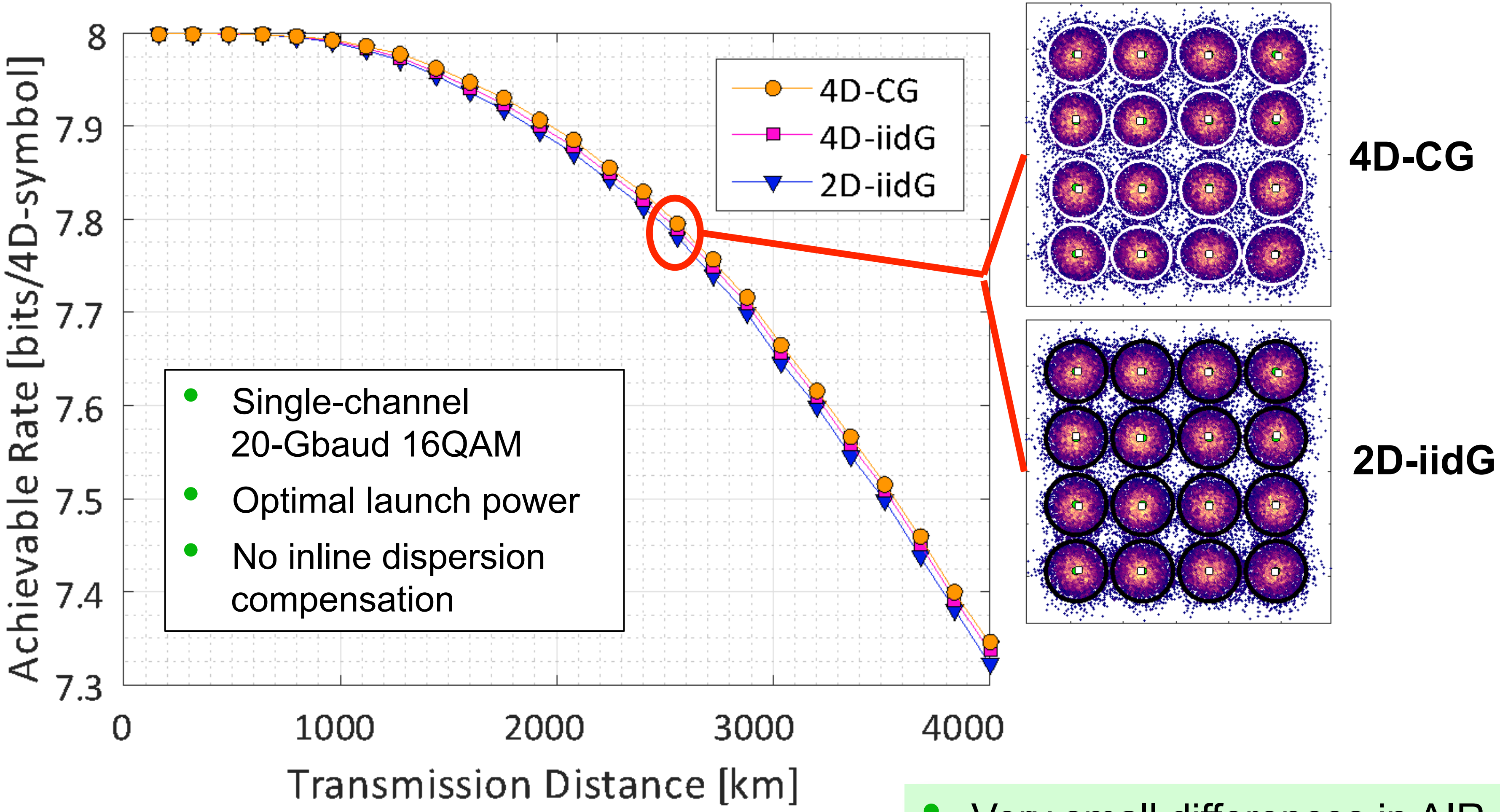
- 4D Gaussian $q(\boldsymbol{Y}|\boldsymbol{X})$
- Optimized covariance for each $\boldsymbol{X}$
- Optimized mean for each $\boldsymbol{X}$



○ = covariance

∘ = mean

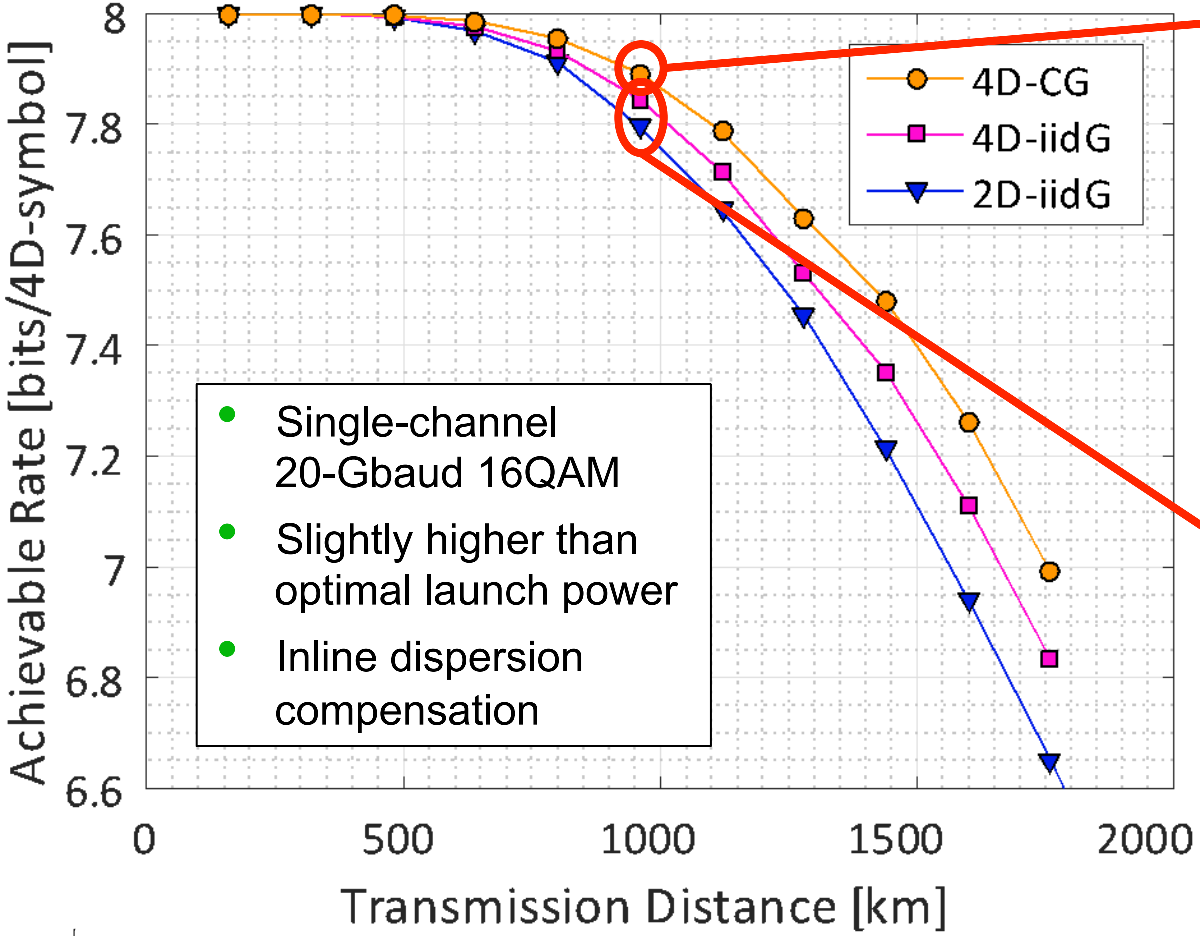
# Results without dispersion compensation



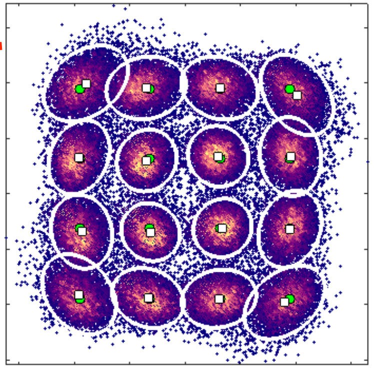

- Single-channel 20-Gbaud 16QAM
- Optimal launch power
- No inline dispersion compensation

- Very small differences in AIR
- Keep on using 2D-iidG!

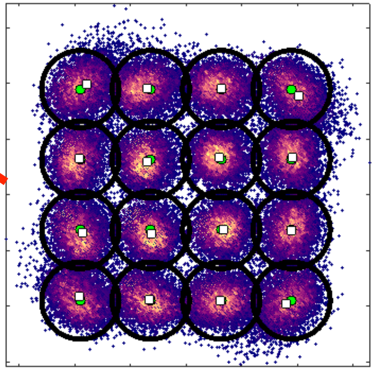
# *Results with dispersion compensation*




- Single-channel 20-Gbaud 16QAM
- Slightly higher than optimal launch power
- Inline dispersion compensation



4D-CG



2D-iidG

- Significant gains in AIR by optimizing 4D means
- Even larger gains if covariances are optimized

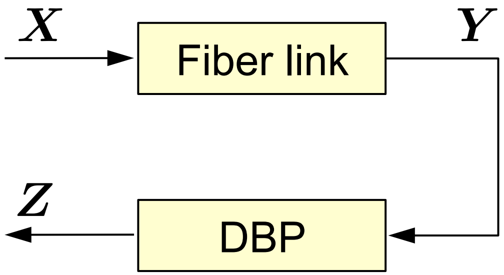
For details, see [2, 3].

Also, extensions to other distributions, other links, other baudrates, and WDM

[2] Eriksson *et al.*, *ECOC* 2015
[3] Eriksson *et al.*, *ArXiv*, Dec. 2015

# 2. Lower bounds from stochastic backpropagation

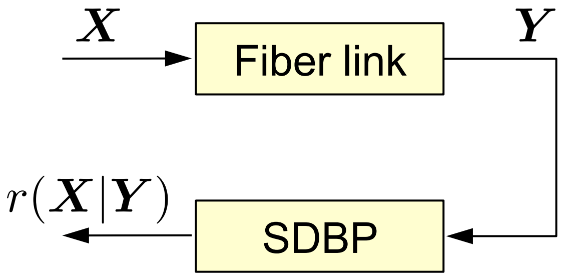
*Baseline:*
Digital backpropagation (DBP)




$$I(\boldsymbol{X};\boldsymbol{Y}) \geq I(\boldsymbol{X};\boldsymbol{Z})$$

Naga V. Irukulapati *et al.*

*Our approach:*
Stochastic DBP (SDBP)




**SDBP** [4, 5]:

- Returns a *distribution* of $\boldsymbol{X}$, not a single estimate
- Builds on factor graphs and the sum-product algorithm
- Marginalizes out unobserved variables to get $r(\boldsymbol{X}|\boldsymbol{Y})$
- Is used to model channel memory

Can SDBP increase the AIR over DBP?

[4] Irukulapati *et al., TCOM 2014*
[5] Irukulapati *et al., JLT 2015*

# *AIR computation*

Mismatched decoding using an auxiliary channel $q$:

$$I(\boldsymbol{X};\boldsymbol{Y}) = \mathbb{E}\left[\log \frac{p(\boldsymbol{Y}|\boldsymbol{X})}{p(\boldsymbol{Y})}\right] \geq \mathbb{E}\left[\log \frac{q(\boldsymbol{Y}|\boldsymbol{X})}{q(\boldsymbol{Y})}\right]$$

- Achievable using a receiver based on $q$ [1]
- Computing $q$ for a given $r$ is hard

Mismatched decoding using a *reverse auxiliary channel* $r$ :

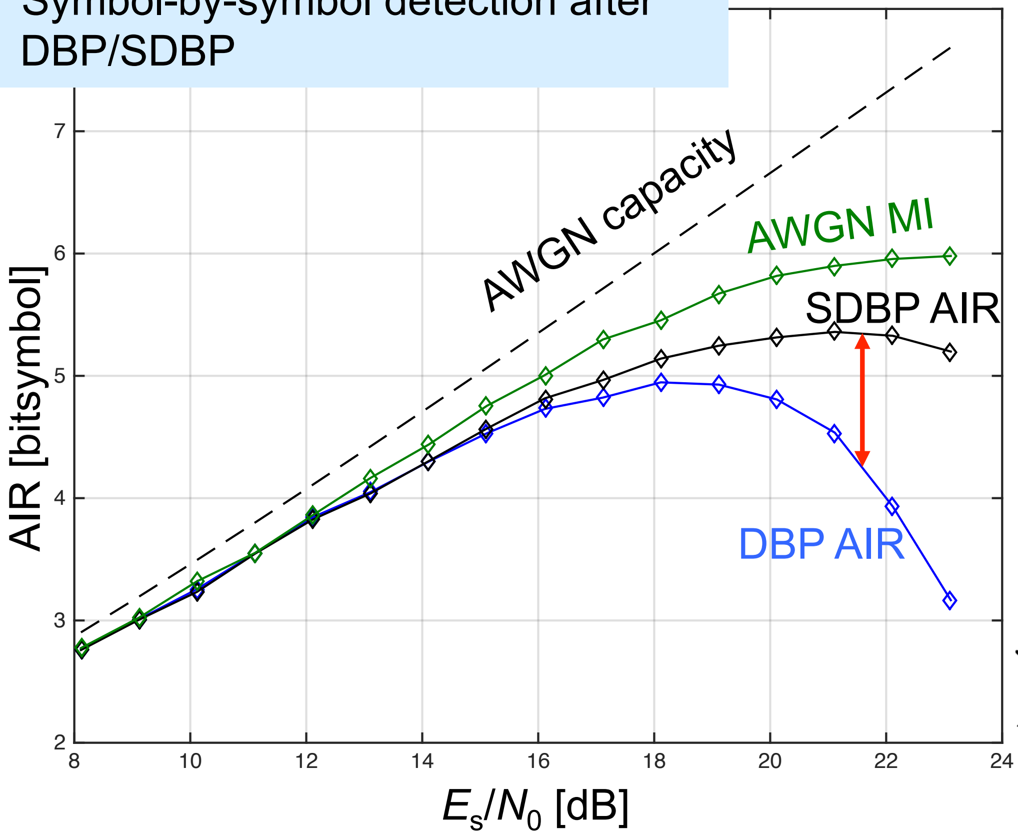
$$I(\boldsymbol{X};\boldsymbol{Y}) = \mathbb{E}\left[\log \frac{p(\boldsymbol{X}|\boldsymbol{Y})}{p(\boldsymbol{X})}\right] \geq \mathbb{E}\left[\log \frac{r(\boldsymbol{X}|\boldsymbol{Y})}{p(\boldsymbol{X})}\right]$$

- Knowledge of $q$ is not necessary for achievability
- Achievable using a receiver based on $r$ [6, 7]
- In this work, $r$ is obtained from SDBP

[1] Arnold *et al.*, *TIT,* 2006
[6] Ganti *et al.*, *TIT,* 2000
[7] Sadeghi *et al.*, *TIT,* 2009

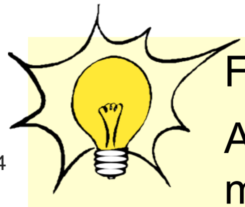
# Results

- DBP/SDBP operates on a *block* of symbols $\boldsymbol{Y}$
- Symbol-by-symbol detection after DBP/SDBP




- Single-channel, single-polarization 28-Gbaud 64-QAM
- 30 spans of 100 km fiber
- Inline dispersion compensation
- Split-step Fourier simulations

*SDBP gives much higher AIR than DBP in the nonlinear regime*



For details, see [8].

Also, extensions to other modulation formats, baudrates, and link setups

[8] Irukulapati *et al.*, in preparation

# 3. Lower bounds for the finite-memory GN model

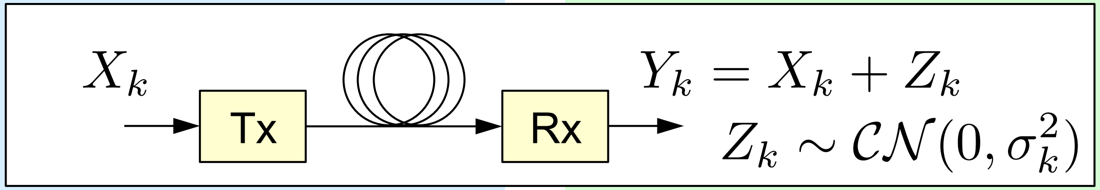
**The (regular) GN model** [9], [10]:

- Large dispersion, weak nonlinearity
- WDM (under some conditions also valid for single channel)
- iid input

⇒ the nonlinearity behaves as additive Gaussian noise

**The finite-memory GN (FMGN) model:**

- Similar to the GN model for iid input, more accurate for non-iid
- More complex—output depends on a sequence of inputs




$$\sigma_k^2 = \sigma_0^2 + \eta \mathbb{E}[|X_k|^2]^3$$

(assumed independent of $k$)

$$\sigma_k^2 = \sigma_0^2 + \eta \left( \frac{1}{2N+1} \sum_{i=k-N}^{k+N} |X_k|^2 \right)^3$$

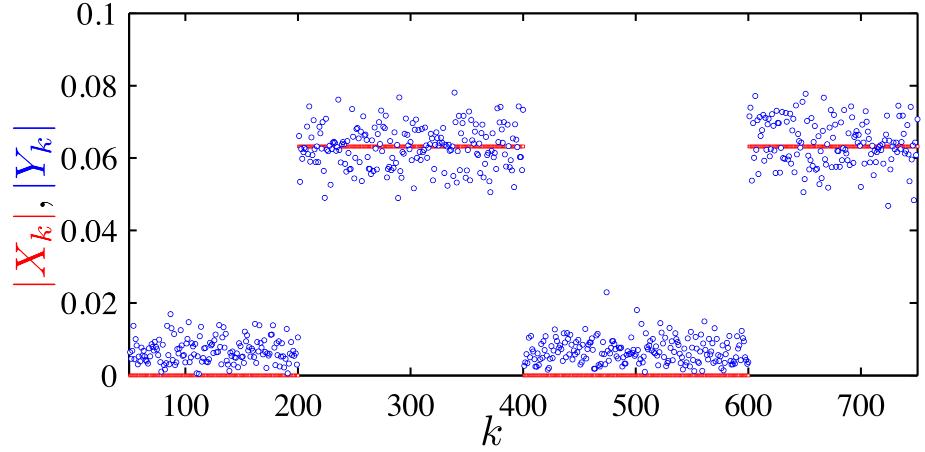
(depends on $k$)

[9] Splett *et al.*, *ECOC* 1993
[10] Poggiolini *et al.*, *JLT* 2014

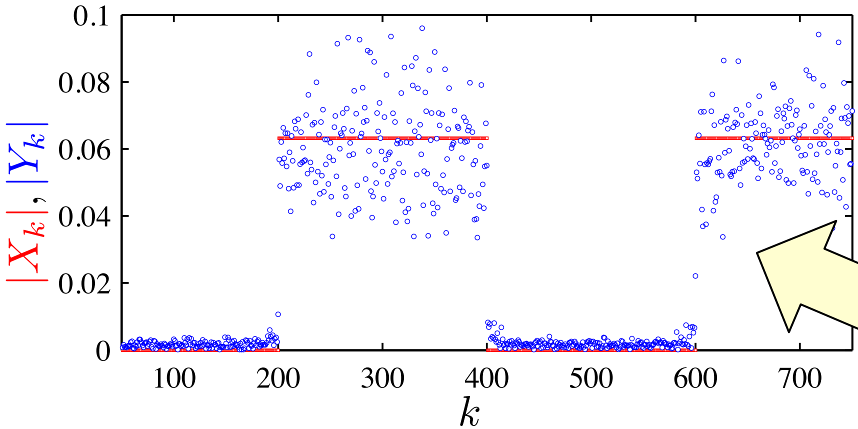
# Model validation

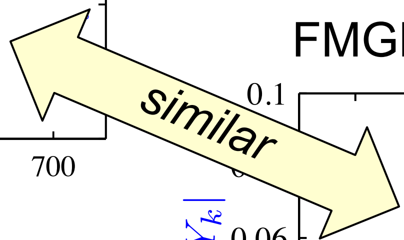
- Pulsed QPSK transmission: 200 symbols at 6 mW, then 200 at 0 mW, etc.

GN model:

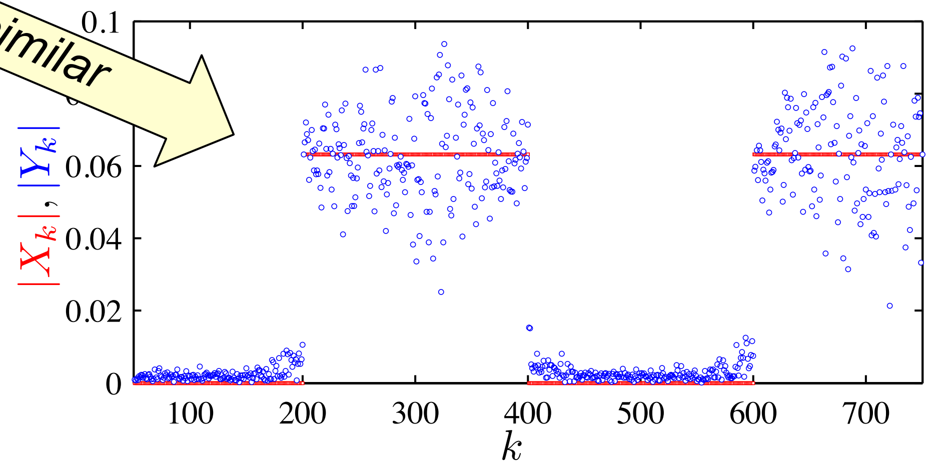



SSFM simulations:



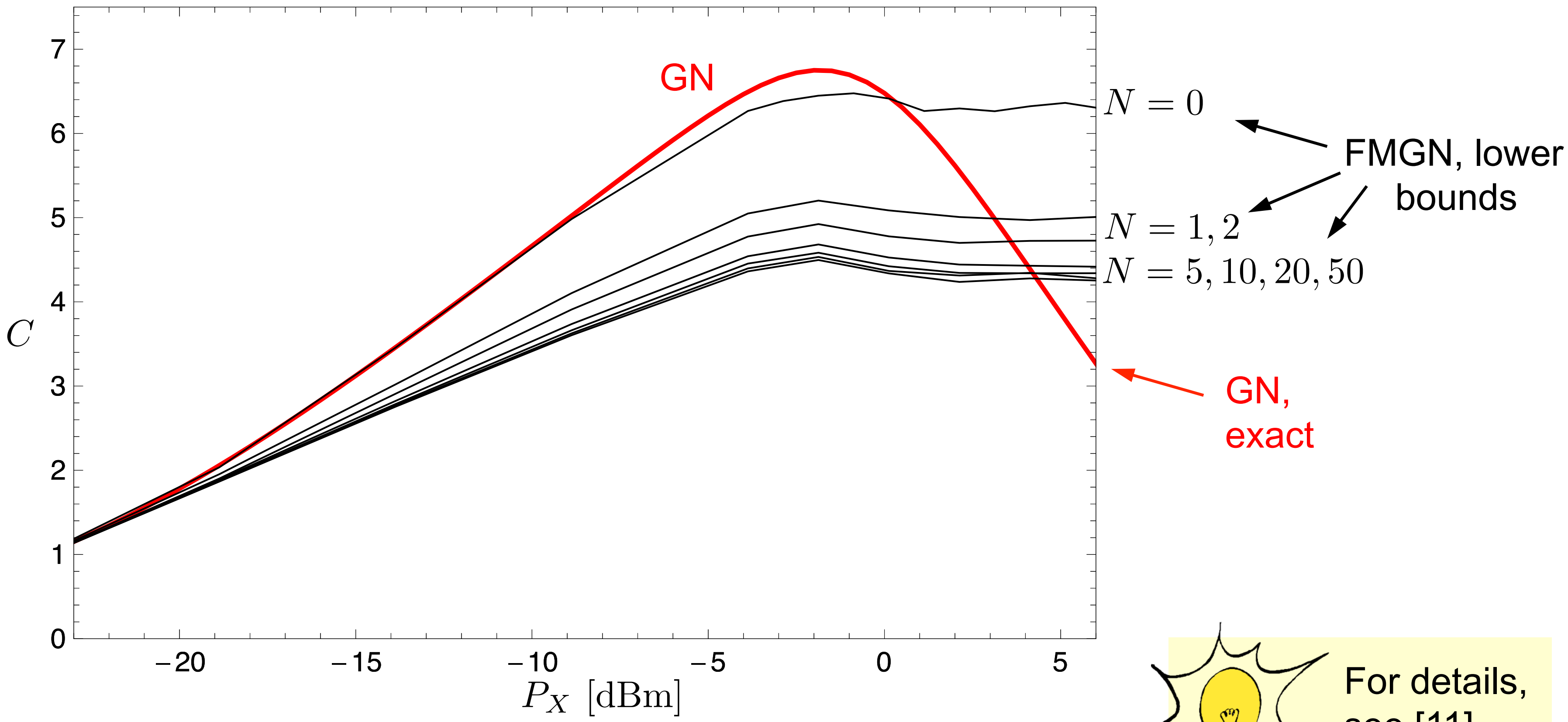




FMGN model ($N = 60$):




- Both models are accurate for iid input
- The FMGN model also allows time-varying noise power

# Channel capacity results




*Capacity estimates based on the GN model are too pessimistic*

For details, see [11].

[11] Agrell *et al.*, *JLT* 2014

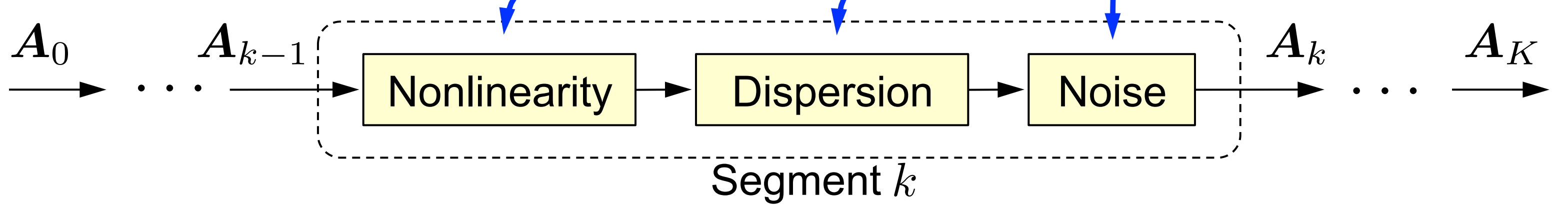
# 4. Lower bound for the split-step Fourier model


**Kamran Keykhosravi *et al.***

The nonlinear Schrödinger equation:

$$\frac{\partial A}{\partial t} = j\gamma |A|^2 A - j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + N$$

The split-step Fourier method:




$A(z,t)$ is sampled in time and space:

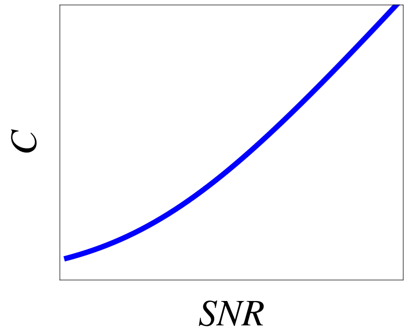
$$\boldsymbol{A}_k = [A_{k,0}, \ldots, A_{k,L-1}]$$
$$A_{k,l} = A(k\Delta z, l\Delta t)$$

- Often used for simulations
- Used as a channel model in [12]

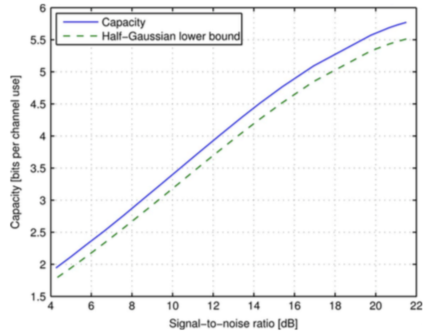
[12] Kramer *et al.*, *ITW* 2015

# Previous capacity results

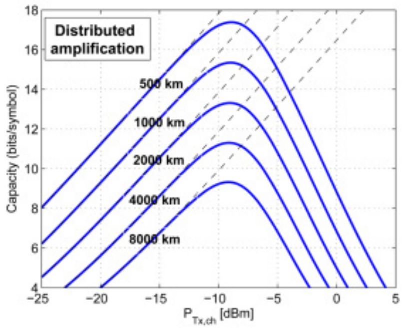
- *Ignoring nonlinearity,* the capacity is $C = \log(1 + SNR)$

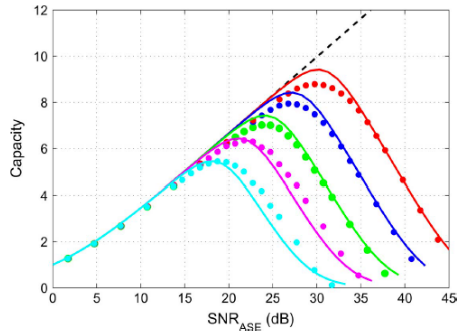


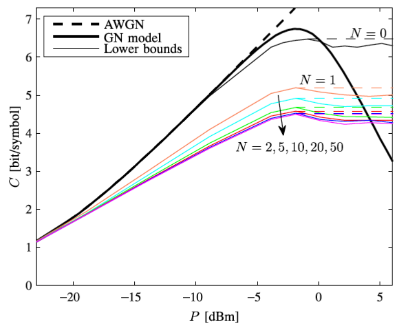

- *Ignoring dispersion,* the capacity grows as $C \sim (1/2) \log SNR$ [13, 14]




- Considering nonlinearity *and* distortion, all known lower bounds either saturate or fall to zero at high power [15, 16, 11]



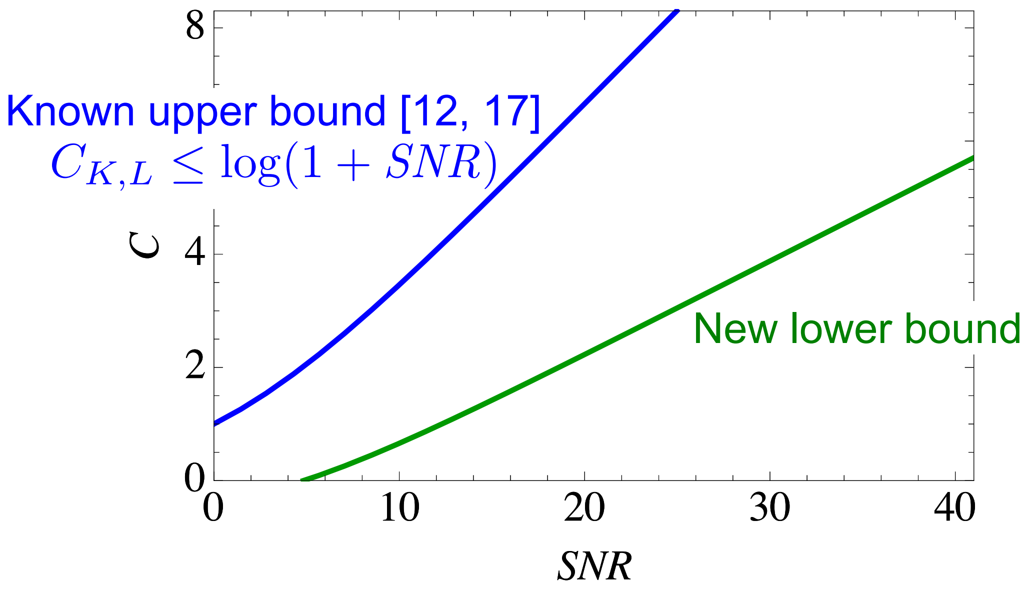

whereas the only known upper bound grows unboundedly [12, 17]

[13] Turitsyn *et al., PRL* 2003
[14] Yousefi and Kschischang, *TIT* 2011
[11] Agrell *et al., JLT* 2014
[15] Bosco *et al., OE* 2012
[16] Mecozzi and Essiambre, *JLT* 2012
[17] Yousefi *et al., CWIT* 2015

# A lower bound for the split-step Fourier channel

The capacity of the split-step Fourier channel with $K$ segments and $L$ time slots is defined as

$$C_{K,L} = \frac{1}{L} \sup I(\boldsymbol{A}_K; \boldsymbol{A}_0)$$

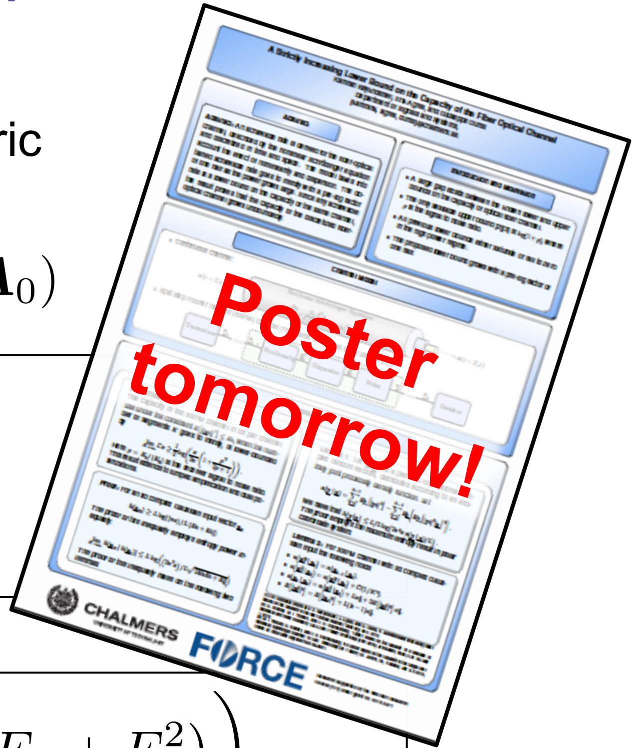



*Theorem:* If $\mathbb{E}[\|\boldsymbol{A}_0\|^2] \leq E_0$, $E_\mathrm{n}$ is the noise energy, and $SNR = E_0/E_\mathrm{n}$, then

$$\lim_{K\to\infty} C_{K,L} \geq \frac{1}{2}\log\left(\frac{e}{2\pi}\left(1 + \frac{SNR^2}{1+2SNR}\right)\right) \sim \frac{1}{2}\log SNR$$

[12] Kramer *et al.*, *ITW* 2015
[17] Yousefi *et al.*, *CWIT* 2015

# Sketch of the proof




Let the input $\boldsymbol{A}_0$ be iid complex circularly symmetric Gaussian and lower-bound

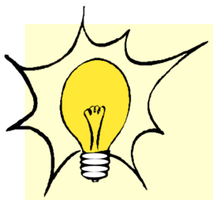
$$I(\boldsymbol{A}_K;\boldsymbol{A}_0) = h(\boldsymbol{A}_K) - h(\boldsymbol{A}_K|\boldsymbol{A}_0)$$

1. $$h(\boldsymbol{A}_K) \geq L\log\left(\frac{\pi e}{L}(E_0+E_{\mathrm{n}})\right)$$

   Proved via the entropy power inequality

2. $$\lim_{K\to\infty} h(\boldsymbol{A}_K|\boldsymbol{A}_0) \leq \frac{L}{2}\log\left(\frac{2\pi^3 e}{L^2}\left(2E_0E_{\mathrm{n}}+E_{\mathrm{n}}^2\right)\right)$$

   Proved via the maximum entropy result in polar coordinates



For details, see [18, 19].

Also, extensions to dual polarization and lumped amplification

[18] Keykhosravi *et al.*, *MIO* 2015, poster
[19] Keykhosravi *et al.*, *ArXiv,* Dec. 2015

# Conclusions

1. Assuming a 4D Gaussian auxiliary channel can increase the AIR over the conventional 2D assumption
2. Accounting for channel memory using SDBP can increase the AIR over DBP
3. The capacity of the finite-memory GN is much higher than that of the conventional GN model at high power
4. The capacity of the split-step Fourier channel grows unboundedly with power

CHALMERS

FORCE
FIBER-OPTIC COMMUNICATIONS RESEARCH CENTER

## Complete references

[1] D. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, “Simulation-based computation of information rates for channels with memory,” *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, 2006.

[2] T. A. Eriksson, T. Fehenberger, N. Hanik, P. A. Andrekson, M. Karlsson, and E. Agrell, “Four-dimensional estimates of mutual information in coherent optical communication experiments,” in *Proc. Eur. Conf. Opt. Commun. (ECOC),* 2015, paper We.4.6.5.

[3] T. A. Eriksson, T. Fehenberger, N. Hanik, P. A. Andrekson, M. Karlsson, and E. Agrell, “Impact of 4D channel distribution on the achievable rates in coherent optical communication experiments,” http://arxiv.org, Dec. 2015.

[4] N. V. Irukulapati, H. Wymeersch, P. Johannisson, E. Agrell, “Stochastic digital backpropagation,” *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3956–3968, 2014.

[5] N. V. Irukulapati, D. Marsella, P. Johannisson, E. Agrell, M. Secondini, H. Wymeersch, “Stochastic digital backpropagation with residual memory compensation,” *J. Lightw. Technol.*, vol. 33, no. 24, pp. 1–7, 2015.

[6] A. Ganti, A. Lapidoth, and I. 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, 2000.

[7] P. Sadeghi, P. O. Vontobel, and R. Shams, “Optimization of information rate upper and lower bounds for channels with memory,” *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 663–688, 2009.

[8] N. V. Irukulapati *et al.,* in preparation.

[9] A. Splett, C. Kurtzke, and K. Petermann, “Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities,” in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Montreux, Switzerland, Sept. 1993.

[10] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, “The GN-model of fiber non-linear propagation and its applications,” *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[11] Erik Agrell, Alex Alvarado, Giuseppe Durisi, and Magnus Karlsson, “Capacity of a nonlinear optical channel with finite memory,” *J. Lightw. Technol.,* vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

[12] G. Kramer, M. I. Yousefi, and F. R. Kschischang, “Upper bound on the capacity of a cascade of nonlinear and noisy channels,” in *Proc. IEEE Inf. Theory Workshop (ITW),* Jerusalem, Israel, Apr.–May, 2015.

[13] K. S. Turitsyn, S. A. Derevyanko, I. V. Yurkevich, and S. K. Turitsyn, “Information capacity of optical fiber channels with zero average dispersion,” *Phys. Rev. Lett.,* vol. 91, no. 20, pp. 203 901-1–203 901-4 Nov. 2003.

[14] M. I. Yousefi and F. R. Kschischang, “On the per-sample capacity of nondispersive optical fibers,” *IEEE Trans. Inf. Theory,* vol. 57, no. 11, pp. 7522–7541, Nov. 2011.

[15] G. Bosco, P. Poggiolini, A. Carena, V. Curri, and F. Forghieri, “Analytical results on channel capacity in uncompensated optical links with coherent detection: Erratum,” *Optics Express,* vol. 20, no. 17, pp. 19610–19611, Aug. 2012.

[16] A. Mecozzi and R.-J. Essiambre, “Nonlinear Shannon limit in pseudolinear coherent systems,” *J. Lightw. Technol.,* vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[17] M. I. Yousefi, G. Kramer, and F. R. Kschischang, “An upper bound on the capacity of the single-user nonlinear Schrödinger channel,” in *Can. Workshop Inf. Theory (CWIT),* St. John’s, NL, Canada, July 2015; online at http://arxiv.org/abs/1502.06455.

[18] K. Keykhosravi, E. Agrell, and G. Durisi, “A strictly increasing lower bound on the capacity of the fiber optical channel,” *Munich Workshop on Inform. Theory of Opt. Fiber (MIO)*, Munich, Germany, poster, Dec. 2015.

[19] K. Keykhosravi, E. Agrell, and G. Durisi, “A monotonically increasing lower bound on the capacity of the fiber-optical channel,” http://arxiv.org, Dec. 2015.