Rate and Power Scaling of Space-Division Multiplexing via Nonlinear Perturbation

Francisco Javier García-Gómez and Gerhard Kramer, *Fellow, IEEE*

***Abstract*—A nonlinear perturbation model for fiber-optic channels is used to analyze achievable rates of combined space-division multiplexing (SDM) and wavelength-division multiplexing (WDM). The analysis shows that, for a large number S of SDM channels, the power of the nonlinear interference is at most quadratic in S and cubic in the launch power. The achievable rates scale at least as the cubic root of S and linearly with the optimized launch power for large S. The expressions are verified numerically for systems with up to 16 SDM channels and 5 WDM channels.**

***Index Terms*—Achievable rate, space-division multiplexing, optical fiber, regular perturbation.**

I. Introduction

Space-division multiplexing (SDM) can significantly increase the information rates of fiber optic channels [1]. Experiments with multi-mode and multi-core fiber reaching up to 10 petabits/s have been reported [2], [3]. An upper bound [4], [5] and several lower bounds [6]–[13] on the achievable rates have been derived in the literature. The lower bounds are usually obtained by using auxiliary (or mismatched) channel models and simulation. The fiber nonlinearity causes cross-phase modulation that limits the lower bounds [6]. Moreover, the nonlinear coupling between SDM channels [14], [15] lowers the peak rate per channel as the number of spatial modes increases [12], [13].

This paper addresses the rate and power scaling in terms of the number of SDM channels. The motivation is that, in analogy to massive multi-input multi-output (massive MIMO) for wireless channels [16], one can anticipate that massive SDM offers advantages in addition to large capacity, namely small energy consumption per information bit and simplified modulation and detection.

To assess performance, we use the regular perturbation (RP) analysis of [13] to derive analytic expressions for the nonlinear interference (NLI) power of SDM systems. The results show that the NLI power is cubic in the launch power and at most quadratic in the number of SDM channels. Upon optimizing the launch power, the information rate grows asymptotically linearly with the total launch power and with at least the third root of the number of SDM channels. Simulations with up to 16 SDM channels show that the achievable rates match the analytic expressions reasonably well.

Date of current version January 7, 2022. This work was supported by the German Research Foundation (DFG) under Grants KR 3517/8-1 and 3517/8-2. (*Corresponding author: Francisco Javier García-Gómez.*)

The authors are with the Institute for Communications Engineering, Technical University of Munich, 80333 Munich, Germany (e-mail: javier.garcia@tum.de; gerhard.kramer@tum.de).

II. Regular Perturbation for SDM

Consider S SDM channels for which the propagating signal vector is

$$\mathbf{u}(z,t) = \left(u^{[1]}(z,t), \ldots, u^{[S]}(z,t)\right) \tag{1}$$

where $u^{[s]}(z,t)$ is the complex-alphabet signal of SDM channel s at distance z and time t. Unlike [13], we here count each polarization as a different SDM channel, so for example a dual-polarization system has $S = 2$. The SDM propagation equations in the strong-coupling regime from [14], [15] are

$$\frac{\partial \mathbf{u}}{\partial z} = -j\frac{\beta_2}{2}\frac{\partial^2 \mathbf{u}}{\partial t^2} + j\gamma\kappa g(z)\left\|\mathbf{u}\right\|^2\mathbf{u} + \frac{\mathbf{n}}{\sqrt{g(z)}} \tag{2}$$

where β_2 is the dispersion coefficient, γ is the nonlinear coefficient, κ is a factor related to the strength of the coupling between the SDM channels [15, Eq. (39)], and $g(z)$ includes the loss and amplification profile. In the absence of nonlinearity, the entries $n^{[s]}(\mathcal{L},t)$ of the noise vector $\mathbf{n}(\mathcal{L},t)$ are independent and identically distributed (i.i.d.) circularly-symmetric complex Gaussian (CSCG) processes with autocorrelation function (ACF)

$$\left\langle N^{[s]}(\mathcal{L},t)N^{[s],*}(\mathcal{L},t')\right\rangle = N_{\text{ASE}}\mathcal{B}_{\text{ASE}}\,\text{sinc}\left(\mathcal{B}_{\text{ASE}}\left(t-t'\right)\right) \tag{3}$$

where $\mathcal{B}_{\text{ASE}}$ is the receiver bandwidth and N_{ASE} is the noise spectral density at the receiver.

We study wavelength-division multiplexing (WDM) with center angular frequencies Ω_c indexed by an integer c. The channel of interest (COI) $c = 0$ has $\Omega_0 = 0$. The symbols $x_m^{[s]}$ and $b_{c,m}^{[s]}$, $c \neq 0$, represent the modulation values transmitted over SDM channel s at time m and over the wavelengths of the COI and WDM channel c, respectively. All SDM and WDM channels use the same unit-energy pulse shape $s(t)$ and all SDM channels of the same WDM channel have the same delay τ_c. If the receiver uses chromatic dispersion compensation (CDC) and a sampled matched filter, the first-order RP analysis [13] of (2) gives the discrete-time model

$$y_m^{[s]} = x_m^{[s]} + w_m^{[s]} + \Delta x_m^{[s]} \tag{4}$$

where the noise samples $w_m^{[s]}$ are i.i.d. CSCG and the NLI is

$$\begin{aligned}\Delta x_m^{[s]} = {} & j\sum_{s'}\sum_{k,k'}^{n} S_{n,k,k'}x_{n+m}^{[s]}x_{k+m}^{[s']}x_{k'+m}^{[s'],*} \\ & + j\sum_{s'}\sum_{c\neq 0}\sum_{k,k'}^{n}\left(1+\delta_{s,s'}\right)D_{c;n,k,k'}x_{n+m}^{[s]}b_{c,k+m}^{[s']}b_{c,k'+m}^{[s'],*} \\ & + j\sum_{r\neq s}\sum_{c\neq 0}\sum_{k,k'}^{n} D_{c;n,k,k'}x_{n+m}^{[r]}b_{c,k+m}^{[s]}b_{c,k'+m}^{[r],*}\end{aligned} \tag{5}$$

where $\delta_{s,s'}$ is 1 if $s = s'$ and 0 otherwise. The coupling coefficients are

$$S_{n,k,k'} = D_{0;n,k,k'} \tag{6}$$

$$D_{c;n,k,k'} = \gamma\kappa \int_0^{\mathcal{L}} \mathrm{d}z\; g(z) \int_{-\infty}^{\infty} \mathrm{d}t\; s^*(z,t)s(z,t-nT)$$
$$s(z,t-\tau_c-kT+\beta_2\Omega_c)s^*(z,t-\tau_c-k'T+\beta_2\Omega_c z). \tag{7}$$

Expression (5) is from equations (12) and (31) of [13] that we write compactly by interpreting different polarizations as different SDM channels. If one uses digital back-propagation (DBP) of the COI instead of CDC then the self-phase modulation (SPM) terms with $S_{n,k,k'}$ are not present.

III. Achievable Rate of the RP Model

A. NLI Power

Suppose $x_m^{[s]}$ and $b_{c,m}^{[s]}$ are zero-mean, i.i.d., proper, and have (normalized) moments

$$E = \left\langle \left| X_m^{[s]} \right|^2 \right\rangle = \left\langle \left| B_{c,m}^{[s]} \right|^2 \right\rangle \tag{8}$$

$$\nu_4 = \frac{\left\langle \left| X_m^{[s]} \right|^4 \right\rangle}{E^2} = \frac{\left\langle \left| B_{c,m}^{[s]} \right|^4 \right\rangle}{E^2} \tag{9}$$

$$\nu_6 = \frac{\left\langle \left| X_m^{[s]} \right|^6 \right\rangle}{E^3}. \tag{10}$$

We have $E = \mathcal{P}T$ where $\mathcal{P}$ is the transmit power in one wavelength and one SDM channel and T is the symbol period. The variance of the random noise samples $W_m^{[s]}$ in (4) is thus N_{ASE} [13].

We now take a different approach from [13] and treat the conditional mean of $\Delta X_m^{[s]}$ as phase noise while the other NLI terms are treated as additive interference. The desired mean conditioned on $x_m^{[s]}$ is

$$\left\langle \Delta X_m^{[s]} \middle| x_m^{[s]} \right\rangle = j x_m^{[s]} \left(\theta + \tilde{\theta}_m^{[s]} \right) \tag{11}$$

where θ and $\tilde{\theta}_m^{[s]}$ are real and given by

$$\theta = E(1+S) \left(\sum_k S_{0,k,k} + \sum_{c\neq 0} \sum_k D_{c;0,k,k} \right) \tag{12}$$

$$\tilde{\theta}_m^{[s]} = \left(\left| x_m^{[s]} \right|^2 - 2E \right) S_{0,0,0}. \tag{13}$$

Define

$$\Delta\tilde{x}_m^{[s]} = \Delta x_m^{[s]} - j x_m^{[s]} \theta \tag{14}$$

and substitute (14) in (4) to write the RP approximation

$$y_m^{[s]} = x_m^{[s]} + j x_m^{[s]}\theta + \Delta\tilde{x}_m^{[s]} + w_m^{[s]}$$
$$\approx x_m^{[s]} e^{j\theta} + \Delta\tilde{x}_m^{[s]} + w_m^{[s]}. \tag{15}$$

The receiver can thus remove θ by multiplying by $e^{-j\theta}$. The mean of $\Delta\tilde{X}_m^{[s]}$ is zero and the variance is

$$\left\langle \left| \Delta\tilde{X}_m^{[s]} \right|^2 \right\rangle = \left\langle \left| \Delta X_m^{[s]} - j X_m^{[s]} \theta \right|^2 \right\rangle$$
$$= \left\langle \left| \Delta X_m^{[s]} \right|^2 \right\rangle - 2\theta \Im \left(\left\langle X_m^{[s],*} \Delta X_m^{[s]} \right\rangle \right) + E\theta^2$$
$$= \left\langle \left| \Delta X_m^{[s]} \right|^2 \right\rangle - E\theta^2 - 2\theta E^2 \left(\nu_4 - 2\right) S_{0,0,0} \tag{16}$$

where we used the theorem of total expectation to write

$$\left\langle X_m^{[s],*} \cdot \Delta X_m^{[s]} \right\rangle = \left\langle \left\langle X_m^{[s],*} \cdot \Delta X_m^{[s]} \middle| X_m^{[s]} \right\rangle_{X_m^{[s]}} \right\rangle$$
$$\overset{\text{(a)}}{=} \left\langle X_m^{[s],*} \left(j X_m^{[s]} \theta + j X_m^{[s]} \tilde{\theta}_m^{[s]} \right) \right\rangle$$
$$= jE\theta + jE^2 \left(\nu_4 - 2\right) S_{0,0,0} \tag{17}$$

where $\langle\cdot\rangle_X$ is a conditional expectation random variable that is a function of X and where step (a) follows by (11).

Using (5), (12), and (13) in (16), and $E = \mathcal{P}T$, the variance of the NLI after a lengthy calculation is

$$\left\langle \left| \Delta\tilde{X}_m^{[s]} \right|^2 \right\rangle = \left(\eta_2 S^2 + \eta_1 S + \eta_0 \right) E\mathcal{P}^2 \tag{18}$$

for coefficients η_2, η_1, η_0 that do not depend on $\mathcal{P}$ or S. We compute

$$\eta_2 = T^2 \sum_{n\neq 0} \left| \sum_k S_{n,k,k} + \sum_{c\neq 0} \sum_k D_{c;n,k,k} \right|^2. \tag{19}$$

The coefficients η_1 and η_0 have very long expressions. We approximate them by neglecting all NLI coefficients except $S_{n,k} \triangleq S_{n,k,n+k}$ and $D_{c;n,k} \triangleq D_{c;n,k,n+k}$, which dominate the others [7]. The result is

$$\eta_1 = T^2 \left[\sum_{n,k} |S_{n,k}|^2 + (\nu_4 - 2) \sum_k |S_{0,k}|^2 \right.$$
$$\left. + 2 \sum_{c\neq 0} \sum_{n,k} |D_{c;n,k}|^2 + (\nu_4 - 2) \sum_{c\neq 0} \sum_k |D_{c;0,k}|^2 \right] \tag{20}$$

$$\eta_0 = T^2 \left[\sum_{n,k} |S_{n,k}|^2 + 3(\nu_4 - 2) \sum_k |S_{0,k}|^2 + \right.$$
$$+ (\nu_4 - 2) \sum_n |S_{n,n}|^2 + (\nu_6 - 9\nu_4 + 12) |S_{0,0}|^2$$
$$\left. + 2 \sum_{c\neq 0} \sum_{n,k} |D_{c;n,k}|^2 + 3(\nu_4 - 2) \sum_{c\neq 0} \sum_k |D_{c;0,k}|^2 \right]. \tag{21}$$

Note that $\eta_1 = \eta_0$ if the input is CSCG, i.e., $\nu_4 = 2$ and $\nu_6 = 6$. We have $|\eta_2| \ll |\eta_1|$ because (19) does not include any $S_{n,k,n+k}$ or $D_{c;n,k,n+k}$. This means that the quadratic

behavior of the NLI power (18) is relevant only for very large S. From (7) we have

$$\sum_{k=-\infty}^{\infty} D_{c;n,k,k} = \kappa\gamma \int_0^{\mathcal{L}} \mathrm{d}z\; g(z) \int_{-\infty}^{\infty} \mathrm{d}t\; s^*(z,t)s(z,t-nT) \cdot \sum_{k=-\infty}^{\infty} |s(z,t-\tau_c-kT+\beta_2\Omega_c z)|^2 . \quad (22)$$

For the sinc pulse $s(t) = \sqrt{\mathcal{B}}\,\mathrm{sinc}(\mathcal{B}t)$ with $T = 1/\mathcal{B}$, the second line of (22) is constant, and the integral over time is zero for $n \neq 0$, yielding $\eta_2 = 0$.

B. CDC Rates

We use a memoryless auxiliary channel model [17]:

$$y_m^{[s]} = x_m^{[s]} e^{j\theta} + z_m^{[s]} \quad (23)$$

where the $Z_m^{[s]}$ are i.i.d. CSCG random variables with variance

$$\sigma_Z^2 = N_{\text{ASE}} + \left\langle \left| \Delta\tilde{X}_m^{[s]} \right|^2 \right\rangle . \quad (24)$$

Let $\mathbf{x}$ and $\mathbf{y}$ be two sequences of $S{\cdot}M$ transmitted and received symbols with M symbols in each SDM channel. A lower bound on the capacity (per SDM symbol) of (15) based on the model (23) is

$$\begin{aligned} R(S,\mathcal{P}) &= \frac{-\left\langle \log_2 q_{\mathbf{Y}}(\mathbf{Y}) \right\rangle + \left\langle \log_2 q_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X}) \right\rangle}{M} \\ &= -\frac{1}{M}\sum_{s,m} \left\langle \log_2 \frac{1}{\pi(E+\sigma_Z^2)} e^{-\frac{\left|Y_m^{[s]}\right|^2}{E+\sigma_Z^2}} \right\rangle \\ &+ \frac{1}{M}\sum_{s,m} \left\langle \log_2 \frac{1}{\pi\sigma_Z^2} e^{-\frac{\left|Y_m^{[s]} - X_m^{[s]} e^{j\theta}\right|^2}{\sigma_Z^2}} \right\rangle \\ &= S\log_2\frac{E+\sigma_Z^2}{\sigma_Z^2} + \frac{\sum_s \sum_m \left\langle \left|Y_m^{[s]}\right|^2 \right\rangle}{M(E+\sigma_Z^2)} - \frac{\sum_s \sum_m \left\langle \left|Y_m^{[s]} - X_m^{[s]} e^{j\theta}\right|^2 \right\rangle}{M\sigma_Z^2} \\ &\overset{(a)}{=} S\log_2\left(1+\frac{E}{\sigma_Z^2}\right). \end{aligned} \quad (25)$$

where $q_{\mathbf{Y}|\mathbf{X}}$ and $q_{\mathbf{Y}}$ are the auxiliary channel and output distributions. In (a) we used the following equality derived from (14) and (17) for CSCG inputs:

$$\left\langle X_m^{[s],*} \Delta\tilde{X}_m^{[s]} \right\rangle = jE^2(\nu_4 - 2)S_{0,0} = 0. \quad (26)$$

Substituting (18) and (24) in (25) yields

$$R(S,\mathcal{P}) = S\log_2\left(1 + \frac{\mathcal{P}}{\frac{N_{\text{ASE}}}{T} + (\eta_2 S^2 + \eta_1 S + \eta_0)\mathcal{P}^3}\right). \quad (27)$$

The rate (27) is a lower bound on the capacity under the RP approximations and the approximations (20)-(21). The rate (27) is not necessarily a lower bound on the true capacity but we will see in Sec. IV that the predictions are reasonable. The per-channel power and total power that maximize (27) are

$$\mathcal{P}_{\text{opt}}(S) = \left(\frac{N_{\text{ASE}}/T}{2(\eta_2 S^2 + \eta_1 S + \eta_0)}\right)^{1/3} \quad (28)$$

$$\mathcal{P}_{\text{tot}}(S) = S\mathcal{P}_{\text{opt}}(S) \quad (29)$$

and the maximum rate is

$$R(S) = S\log_2\left(1 + \frac{1}{3}\left(\frac{4T^2}{N_{\text{ASE}}^2(\eta_2 S^2 + \eta_1 S + \eta_0)}\right)^{1/3}\right). \quad (30)$$

C. DBP Rates

For DBP all terms with $S_{n,k,k'}$ in (19)-(21) vanish. Moreover, if we keep only the $D_{c;0,k,k}$ coefficients in (20) and (21) then we obtain

$$\eta_1 = \nu_4 r_\phi \quad (31)$$

$$\eta_0 = (3\nu_4 - 4)\, r_\phi \quad (32)$$

where

$$r_\phi = T^2 \sum_{c\neq 0}\sum_k |D_{c;0,k,k}|^2 \approx \gamma^2\mathcal{L}T \sum_{c\neq 0} \frac{1}{|\beta_2\Omega_c|}. \quad (33)$$

The approximation in (33) is valid for large accumulated dispersion [18]. The rate (27) becomes

$$\log_2\left(1 + \frac{\mathcal{P}}{\frac{N_{\text{ASE}}}{T} + (\eta_2 S^2 + [\nu_4(S+3) - 4]\, r_\phi)\mathcal{P}^3}\right). \quad (34)$$

For example, for $S = 1$ we can neglect η_2 and obtain the same rate as in [7, Eq. (116)]:

$$\log_2\left(1 + \frac{\mathcal{P}}{\frac{N_{\text{ASE}}}{T} + 4\gamma^2\mathcal{L}T(\nu_4 - 1)\mathcal{P}^3 \sum_{c\neq 0}\frac{1}{|\beta_2\Omega_c|}}\right). \quad (35)$$

D. Asymptotic Analysis

For $\eta_2 > 0$ and $S \to \infty$ we have

$$R(S) \sim \frac{\log_2 e}{3}\sqrt[3]{\frac{4T^2}{N_{\text{ASE}}^2\eta_2}}\, S^{1/3} \sim \frac{2T\log_2 e}{3N_{\text{ASE}}}\mathcal{P}_{\text{tot}}(S) \quad (36)$$

where $f(S) \sim g(S)$ means that $\lim_{S\to\infty}[f(S) - g(S)] = 0$. The total capacity of the massive SDM system thus grows linearly with the total power or, equivalently, with $S^{1/3}$. When using sinc pulses we have $\eta_2 = 0$ and the asymptotics are

$$R(S) \sim \frac{\log_2 e}{3}\sqrt[3]{\frac{4T^2}{N_{\text{ASE}}^2\eta_1}}\, S^{2/3} \sim \frac{2T\log_2 e}{3N_{\text{ASE}}}\mathcal{P}_{\text{tot}}(S). \quad (37)$$

The following upper bound on (30) is linear in S:

$$R(S) \leq S\log_2\left(1 + \frac{1}{3}\left(\frac{4T^2}{N_{\text{ASE}}^2(\eta_2 + \eta_1 + \eta_0)}\right)^{1/3}\right). \quad (38)$$

This upper bound is tight for small S and large signal-to-interference-and-noise ratio. Fig. 1 shows $R(S)$, the asymptotes (36) and (37) and the upper bound (38) for a 1000-km link using CDC with the parameters in Table I and two different pulses: a sinc pulse and a root raised cosine (RRC) pulse with roll-off factor 0.1. The values of η_2, η_1 and η_0 for this system are given in Table II. We see that the growth of $R(S)$ is close to linear in S for small and intermediate S and reduces to $S^{2/3}$ (sinc) or $S^{1/3}$ (RRC) for large S.

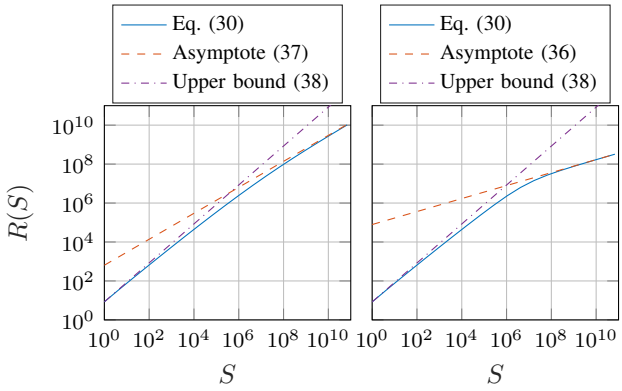

Fig. 1. SDM rates for the COI. Left plot: sinc pulse. Right plot: RRC pulse.

TABLE I
SYSTEM PARAMETERS

Parameter	Symbol	Value
Dispersion coefficient	β_2	-25.25 ps^2/km
Nonlinear coefficient	γ	0.835 W^{-1}km^{-1}
Noise spectral density	N_{ASE}	$5.902 \cdot 10^{-18}$ W/Hz
Channel bandwidth	$\mathcal{B}$	50 GHz
Channel spacing	$\Omega^{(1)}/(2\pi)$	50 GHz
RRC roll-off factor		0.1

TABLE II
NLI-SDM COEFFICIENTS OF (27) FOR THE SYSTEM IN TABLE I

Coeff. in Watt^{-2}	Sinc pulse		RRC pulse	
	DBP	CDC	DBP	CDC
η_2	0	0	$1.46 \cdot 10^{-5}$	0.0104
$\eta_1 = \eta_0$	10190	19288	11134	22170

Fig. 2 plots $R(S)$ against the total launch power $\mathcal{P}_{\text{tot}}(S)$. The asymptotic linear growth of $R(S)$ with power appears at large S while for smaller S the growth seems to be faster.

Fig. 3 plots the coefficients η_2, η_1 and η_0 as a function of the channel bandwidth $\mathcal{B}$. The coefficients scale approximately as $1/\mathcal{B}^2$ due to the averaging effect that reduces the NLI when the accumulated dispersion is long.

IV. NUMERICAL RESULTS

We simulated systems with $\mathcal{L} = 1000$ km and the parameters shown in Table I. There are 5 WDM channels and the COI is the center channel. The transmitters use sinc pulses and i.i.d. CSCG symbols with the optimal launch power (29). Table II lists the coefficients η_2, η_1, η_0 that were computed using (19)-(21). The receiver isolates the COI using a bandpass filter and then applies either CDC or joint DBP of all SDM channels, followed by a matched filter and sampler.

We compute $R(S)$ by using one of two auxiliary channel models: an additive white Gaussian noise (AWGN) model and an (impractical) receiver that uses particle filtering and the Markov phase noise model in [11]. The receiver processes each SDM channel separately and ignores the memory of the additive NLI, i.e., there is no whitening filter as in [11]. We compare the simulated rates with the rates computed with $R(S, \mathcal{P})$ in (27) and with the upper bound from [4], [5]:

$$R(S, \mathcal{P}) \leq S \log_2 \left(1 + \frac{\mathcal{P}}{N_{\text{ASE}}/T}\right). \tag{39}$$

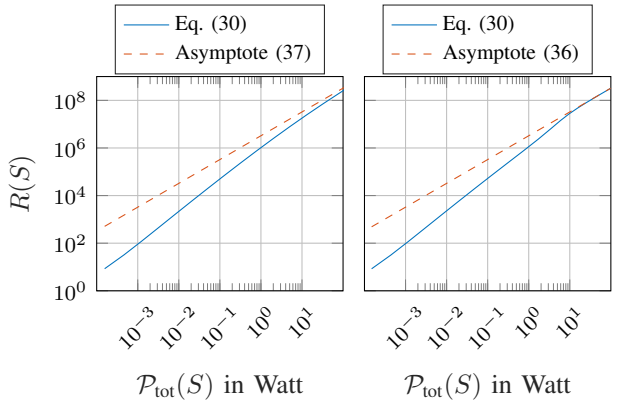

Fig. 2. SDM rates vs. total launch power for the COI. Left plot: sinc pulse. Right plot: RRC pulse.

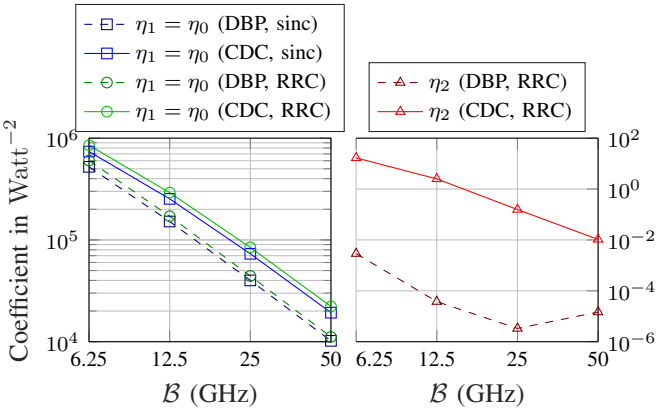

Fig. 3. Bandwidth dependence of η_2, η_1, η_0 in (27) for sinc and RRC pulses. Note that $\eta_2 = 0$ for the sinc pulse.

The results are shown in Fig. 4. Subfigures (a) and (b) show the rates as a function of S and $\mathcal{P}_{\text{tot}}(S)$. The rate scaling for these small S is faster than the asymptotic $S^{2/3}$ and $\mathcal{P}_{\text{tot}}(S)$. Subfigures (c) and (d) show the normalized rate $R(S)/S$ and the optimal power $\mathcal{P}_{\text{opt}}(S)$ as a function of S. Both should scale as $S^{-1/3}$ for large S.

Observe that the $R(S)$ in Fig. 4 are closer to the rates of the Markov phase noise model than the rates of the AWGN model. For CDC the AWGN model does not reach $R(S, \mathcal{P})$. This may be due to the model mismatch, or because (27) is an achievable rate of the RP approximation (15) and not necessarily of the optical channel. In the simulations, the peak rate of DBP is approximately 1.04 times[1] larger than that of CDC while the power (28) is 1.24 times larger. The larger power factor makes the CDC curves lie above the DBP curves in subfigure (b).

V. CONCLUSION

The RP analysis for optical fiber channels in [13] was used to show that the power of the NLI is at most quadratic in the number of SDM channels and cubic in the launch power. The RP rates grow linearly with the total launch power and at least with the cubic root of the number of SDM channels.

[1]This factor ranged from 1.036 to 1.041 in the simulations.

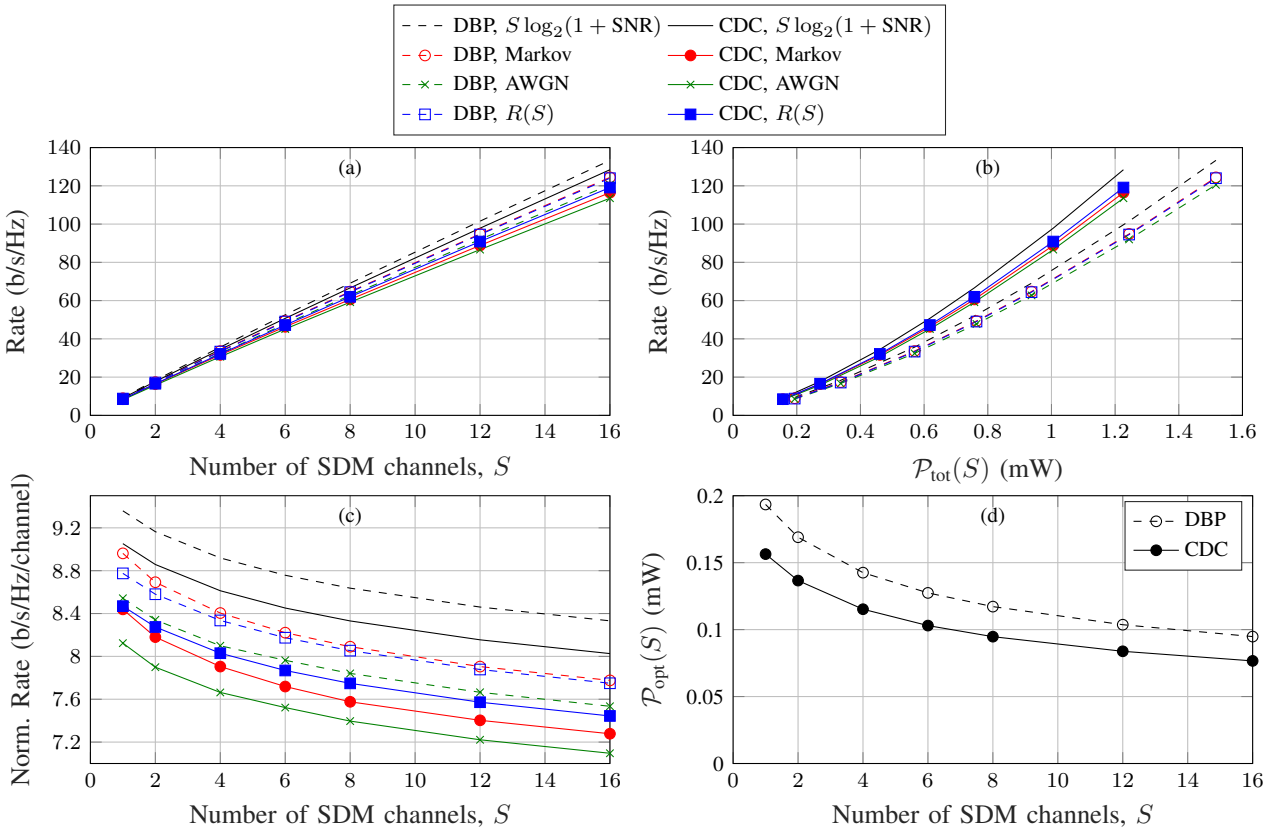

Fig. 4. SDM rates and optimized launch powers for sinc pulses and the COI.

The RP rates are close to the rates achieved by simulated SDM systems with up to 16 SDM channels.

References

[1] B. J. Puttnam, G. Rademacher, and R. S. Luís, "Space-division multiplexing for optical fiber communications," *Optica*, vol. 8, Sep 2021.

[2] D. Soma, Y. Wakayama, S. Beppu, S. Sumita, T. Tsuritani, T. Hayashi, T. Nagashima, M. Suzuki, M. Yoshida, K. Kasai, M. Nakazawa, H. Takahashi, K. Igarashi, I. Morita, and M. Suzuki, "10.16-Peta-b/s dense SDM/WDM transmission over 6-mode 19-core fiber across the C+L band," *Journal of Lightwave Technology*, vol. 36, no. 6, pp. 1362–1368, 2018.

[3] G. Rademacher, B. J. Puttnam, R. S. Luís *et al.*, "10.66 Peta-bit/s transmission over a 38-core-three-mode fiber," in *Optical Fiber Commun. Conf. (OFC)*, 2020, pp. 1–3.

[4] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *IEEE Inf. Theory Workshop*, April 2015, pp. 1–4.

[5] M. I. Yousefi, G. Kramer, and F. R. Kschischang, "Upper bound on the capacity of the nonlinear Schrödinger channel," in *IEEE Can. Workshop Inf. Theory*, July 2015, pp. 22–26.

[6] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb 2010.

[7] A. Mecozzi and R. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 2011–2024, June 2012.

[8] P. Poggiolini, G. Bosco, A. Carena *et al.*, "The GN-model of fiber nonlinear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb 2014.

[9] A. Carena, G. Bosco, V. Curri *et al.*, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun 2014.

[10] M. Secondini, E. Agrell, E. Forestieri *et al.*, "Nonlinearity mitigation in WDM systems: Models, strategies, and achievable rates," *J. Lightw. Technol.*, vol. 37, no. 10, pp. 2270–2283, May 2019.

[11] F. J. García-Gómez and G. Kramer, "Mismatched models to lower bound the capacity of optical fiber channels," *J. Lightw. Technol.*, vol. 38, no. 24, pp. 6779–6787, 2020.

[12] ——, "Mismatched models to lower bound the capacity of dual-polarization optical fiber channels," *J. Lightw. Technol.*, vol. 39, no. 11, pp. 3390–3399, 2021.

[13] ——, "Regular perturbation and achievable rates of space-division multiplexed optical channels," in *Int. Symp. Wireless Commun. Systems (ISWCS)*, 2021, pp. 1–5.

[14] A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multi-mode fibers in the strong coupling regime," *Opt. Express*, vol. 20, no. 11, pp. 11 673–11 678, May 2012.

[15] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: Generalization of the Manakov equations," *J. Lightw. Technol.*, vol. 31, no. 3, pp. 398–406, 2013.

[16] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, Nov. 2010.

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

[18] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25 685–25 699, Nov 2013.