Regular Perturbation and Achievable Rates of Space-Division Multiplexed Optical Channels

Francisco Javier García-Gómez and Gerhard Kramer

Abstract—Regular perturbation is applied to space-division multiplexing (SDM) on optical fibers and motivates a correlated rotation-and-additive noise (CRAN) model. For $S$ spatial modes, or $2S$ complex-alphabet channels, the model has $4S(S + 1)$ hidden independent real Gauss-Markov processes, of which $2S$ model phase noise, $2S(2S − 1)$ model spatial mode rotation, and $4S$ model additive noise. Achievable information rates of multi-carrier communication are computed by using particle filters. For $S = 2$ spatial modes with strong coupling and a 1000 km link, joint processing of the spatial modes gains 0.5 bits/s/Hz/channel of communication. See also [17], [18].

Index Terms—Capacity, multi-mode, optical fiber, space-division multiplexing.

I. INTRODUCTION

Space-division multiplexing (SDM) via multi-core and multi-mode transmission considerably increases the capacity of optical fibers [1]. Several experiments reaching 10 Pbit/s have been reported, e.g., see [2]–[4]. The analysis in [5], [6] generalizes to SDM and gives a spectral efficiency upper bound of $\log_2(1 + \text{SNR})$ bits/s/Hz/channel, where SNR is the signal-to-noise ratio without the fiber nonlinearity and “channel” refers to a complex-alphabet channel.

Various mismatched models have been developed to compute achievable rates for single and dual polarization (1-pol, 2-pol) transmission [7]–[9]. We focus on models based on regular perturbation (RP) [10]–[13] and logarithmic perturbation (LP) [14], [15]. A combined RP and LP model is motivated in [16], see also [17], [18].

We follow [19] and [20] and develop an analysis for SDM by using the coupling equations in [21], [22]. RP, and an approximation related to LP. The result is a correlated rotation-and-additive noise (CRAN) model with correlated phase noise, spatial mode rotations, and additive noise. We use the model to compute achievable rates for SDM channels.

Notation: We use similar notation as in [19] and refer to that paper for details. For instance, we write the Fourier transform of a function $u(t)$ as $\mathcal{F}(u(t))$ and the inverse Fourier transform of $U(\Omega)$ as $\mathcal{F}^{-1}(U(\Omega)) = \mathcal{F}^{-1}(U(\Omega))(t)$. For vectors, operators such as $\mathcal{F}$ and $\mathcal{F}^{-1}$ are applied entrywise.

II. SPACE-DIVISION MULTIPLEXING

Each spatial mode has two complex-alphabet channels with the same spatial field distribution but orthogonal polarization.

Date of current version September 28, 2021. This work was supported by the German Research Foundation (DFG) under Grants KK 35178/1-1 and 35178/2. (Corresponding author: Francisco Javier García-Gómez.)

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

Consider $S$ spatial modes, i.e., $2S$ complex-alphabet channels in total. The propagating signal is

$$\mathbf{u}(z, t) = (\mathbf{u}^1(z, t)^T \mathbf{u}^2(z, t)^T \cdots \mathbf{u}^S(z, t)^T)^T \quad (1)$$

where $\mathbf{u}^s(z, t) = (\mathbf{u}^{s1}(z, t), \mathbf{u}^{s2}(z, t))^T$ is the vector of signals of the $s$-th spatial mode, $s = 1, \ldots, S$. The variable $z$ represents distance and $t$ is time. We consider the following two propagation scenarios. The models assume that fiber birefringence changes randomly with $z$ [21].

Weak Coupling: the linear coupling among spatial modes is neglected, which is reasonable for multi-core transmission. The propagation equation for spatial mode $s$ is [21] Eq. (26)

$$\frac{\partial \mathbf{u}^s}{\partial z} = j\beta_0^s \mathbf{u}^s - \beta_1^s \frac{\partial \mathbf{u}^s}{\partial t} - j\frac{\beta_2^s}{2} \frac{\partial^2 \mathbf{u}^s}{\partial t^2} + \frac{n^s}{\sqrt{g_s(z)}} + j\gamma \left( f_{s, r} g_s(z) \left\| \mathbf{u}^r \right\|^2 + \sum_{r \neq s} f_{s, r} g_r(z) \left\| \mathbf{u}^r \right\|^2 \right) \mathbf{u}^s \quad (2)$$

where $\beta_0^s$, $\beta_1^s$, and $\beta_2^s$ are the coefficients of the Taylor expansion of the propagation constant $\beta(z)$ in the angular frequency $\Omega$, averaged over the two polarizations. The noise signals $n^s = (n^{s1}(z, t), n^{s2}(z, t))^T$ are independent Wiener processes in $z$ such that, in the absence of nonlinearity ($\gamma = 0$), the accumulated noise in a bandwidth of $B_{\text{ASE}}$ at $z = \mathcal{L}$ has autocorrelation function (ACF) $N_{\text{ASE}} B_{\text{ASE}} \text{sinc}^2(B_{\text{ASE}}(t - t'))$.

The nonlinear coupling coefficients $\gamma$ and $f_{s, r}$ are described in [21]. Note that we have absorbed the factors $(8/9)$ and $(4/3)$ from [21] Eq. (26)) into $f_{s, r}$. The functions $g_s(z)$ account for attenuation and amplification. Ideal distributed amplification (IDA) has $g_s(z) = 1$.

Strong Coupling: the linear coupling between modes is strong, i.e., the vector $\mathbf{u}$ is subject to random unitary transformations that change rapidly with $z$. This has an averaging effect that simplifies (2) to [21] Eq. (38)

$$\frac{\partial \mathbf{u}}{\partial z} = -j\beta_2 \frac{\partial^2 \mathbf{u}}{\partial t^2} + j\gamma g(z) \left\| \mathbf{u} \right\|^2 \mathbf{u} + \frac{n}{\sqrt{g(z)}} \quad (3)$$

where $n(z, t)$ has the same statistics as $(n^{11}(z, t), \ldots, n^{SS}(z, t))^T$, the average dispersion coefficient is $\beta_2 = (\sum_s \beta_2^s)/S$, and $\kappa$ is defined in [21].

We develop a SDM-CRAN model for weak coupling. The model for strong coupling is a special case with the parameters

$\beta_0^s = \beta_2^s = 0$, $\beta_2^s = \beta_2$ and $f_{s, r} = \kappa$.

1 This paper refers to mode-division multiplexing but equation (4) also applies to multi-core transmission under certain conditions [24].

Copyright (c) 2021 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
III. REGULAR PERTURBATION

We expand \( u^{[s]} \) in (2) in powers of a small perturbation \( \gamma \):

\[
\begin{align*}
\mathbf{u}^{[s]}(z, t) &= \mathbf{u}^{[s]}_0(z, t) + \gamma \Delta \mathbf{u}^{[s]}(z, t) + \mathcal{O}(\gamma^2).
\end{align*}
\]

(4)

The steps are the same as in [19], [20] and we arrive at

\[
\begin{align*}
\mathbf{u}^{[s]}_0(z, t) &= \mathbf{u}^{[s]}_{\text{IN}}(z, t) + \mathbf{u}^{[s]}_{\text{ASE}}(z, t)
\end{align*}
\]

(5)

where \( \mathbf{u}^{[s]}_{\text{IN}}(z, t) = D_z^{[s]} \mathbf{u}^{[s]}(0, t) \) and

\[
\begin{align*}
\Delta \mathbf{u}^{[s]}(z, t) &= j D_z^{[s]} \left[ \int_0^z D_z^{[s']} \left( f_{s,s}g_s(z') \right) \left\| \mathbf{u}^{[s]}_0(z', t) \right\|^2 \mathbf{u}^{[s]}_0(z', t) \right. \\
&+ \sum_{r \neq s} \int_0^z f_{s,r}g_r(z') \left\| \mathbf{u}^{[r]}_0(z', t) \right\|^2 \mathbf{u}^{[s]}_0(z', t) \right] dz \\
&\left. \right] - \sum_{r \neq s} \int_0^z f_{s,r}g_r(z') \left\| \mathbf{u}^{[r]}_0(z', t) \right\|^2 \mathbf{u}^{[s]}_0(z', t) \right] dz
\end{align*}
\]

(6)

where

\[
D_z^{[s]} u(t) = F^{-1} \left( \int e^{-j \frac{\gamma[s] - \gamma[s']}{\Omega} + \frac{j \gamma[s]}{\Omega^2}} z F(u(t)) \right). \quad (7)
\]

The effect of \( \gamma[s] \) is a z-dependent phase shift, and the effect of \( \gamma[s'] \) is a z-dependent delay. Similar to [20], the 2S entries of the S noise signals \( u_{\text{ASE}}^{[s]}(z, t) \) are independent circularly-symmetric complex Gaussian (CSCG) processes. At the end of the fiber (\( z = L \)), their ACF for \( \gamma = 0 \) is \( N_{\text{ASE}} B_{\text{ASE}} \text{Sinc}(B_{\text{ASE}}(t - t')) \).

We use pulse amplitude modulation (PAM) and wavelength division multiplexing (WDM). The indexes \( i \) of the WDM channels are in the set \( \{c_{\min}, \ldots, i, \ldots, c_{\max}\} \). The center angular frequency of channel \( e \) is \( \Omega_e \), and the WDM channel of interest (COI) has \( c = 0 \) and \( \Omega_0 = 0 \). The transmitted signal of spatial mode \( s \) is

\[
\begin{align*}
\mathbf{u}^{[s]}(0, t) &= \sum_{m=-\infty}^{\infty} \left( \mathbf{x}^{[m]}_s s(t - mT - \tau^{[s]}_m) \right) / \mathbf{x}^{[m]}_s s(t - mT - \tau^{[s]}_0) \\
&+ \sum_{c \neq 0} e^{j \Omega_c t} \sum_{k=-\infty}^{\infty} \left( \mathbf{b}^{[c]}_{s,k} s(t - kT - \tau^{[c]}_k) \right) / \mathbf{b}^{[c]}_{s,k} s(t - kT - \tau^{[c]}_0)
\end{align*}
\]

(8)

where \( T \) is the symbol period. The base pulse \( s(t) \) has most of its energy in the frequency band \( \Omega \leq \pi B \). The channels may have different delays \( \tau^{[s]}_m \) and \( \tau^{[c]}_k \). The transmit symbols of the COI (\( c = 0 \)) are \( \mathbf{x}^{[m]}_0 \) and \( \mathbf{b}^{[c]}_{0,k} \), and the transmit symbols of the interfering frequencies are \( \mathbf{b}^{[s,c]}_{0,k} \) and \( \mathbf{b}^{[c,k]}_{s,0} \). We consider symmetric symbol energies and fourth moments:

\[
\begin{align*}
E &= \left\langle |\mathbf{x}^{[m]}|^2 \right\rangle, \quad \text{all } s, m \\
E_c &= \left\langle |\mathbf{b}^{[c]}_{s,k}|^2 \right\rangle, \quad \text{all } s, k, c \neq 0 \\
Q_c &= \left\langle |\mathbf{b}^{[c]}_{s,k}|^4 \right\rangle, \quad \text{all } s, k, c \neq 0.
\end{align*}
\]

(9)

The receiver applies a band-pass filter \( h_B(t) \) followed by linear distortion compensation (LDC) or digital back-propagation (DBP), a matched filter, and a sampler:

\[
y^{[s]}_m = \int_{-\infty}^{\infty} s_g^* (t - mT) \left\{ D_{L,2} [h_B(t) \ast u^{[s]}(L, t)] \right\} \text{d}t. \quad (10)
\]

We focus on \( y^{[s]}_m \) to save space. Due to symmetry, the model for \( y^{[s]}_m \) is obtained by replacing all variables with a “bar” by the corresponding variables without a “bar”, and vice versa. We assume that \( h_B(t) \ast s(t) = s(t) \).

Substituting (10) into (10), and using (8), we obtain the RP model of the received symbols:

\[
y^{[s]}_m = x^{[s]} + u^{[s]}_m + \Delta x^{[s]}_m. \quad (11)
\]

As in [20], the noise \( u^{[s]}_m \) is independent and identically distributed (i.i.d.) complex Gaussian with variance \( N_{\text{ASE}} \). Also as in [20], we neglect signal-noise mixing, and the nonlinear interference (NLI) becomes

\[
\begin{align*}
\Delta x^{[s]}_m &= j \sum_{s', n,k} S_{n,k}^{[s,s']} x^{[s',n+m,k+m]} x^{[s]}_m \\
&+ j \sum_{s', n,k} S_{n,k}^{[s,s']} x^{[s',n+m,k+m]+} x^{[s]}_m \\
&+ j \sum_{s', n,k} S_{n,k}^{[s,s']} x^{[s',n+m,k+m]+} x^{[s]}_m \\
&+ j \sum_{s', n,k} S_{n,k}^{[s,s']} x^{[s',n+m,k+m]+} x^{[s]}_m.
\end{align*}
\]

(12)

The coefficients in (12) can be written in terms of

\[
A_{n,k}^{[s]}(t_1, t_2, t_3) = \gamma f_{s,s'} \int_{-\infty}^{\infty} dz g_s(z) \int_{-\infty}^{\infty} dt s^*(z, t)
\]

(13)

We choose \( s_0^{[s]} = 0 \). For \( r \neq s \), we have

\[
\begin{align*}
S_{n,k}^{[s,s']} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
S_{n,k}^{[s,s']} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
C_{n,k}^{[s,r]} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
C_{n,k}^{[s,r]} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
C_{n,k}^{[s,r]} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
C_{n,k}^{[s,r]} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]}) \\
D_{n,k}^{[s,r]} &= A_{n,k}^{[s]}(0, r_0^{[s]}, r_0^{[s]})
\end{align*}
\]

(14)

(15)

(16)

(17)

(18)

(19)

(20)

Using single-polarization DBP on the COI removes the terms with \( S_{n,k,k'}^{[s]} \) from (12). Using dual-polarization DBP on the COI removes the terms with \( S_{n,k,k'}^{[s]} \) or \( S_{n,k,k'}^{[s]} \). Using multimode DBP removes all terms with \( S_{n,k,k'}^{[s]} \) or \( S_{n,k,k'}^{[s]} \).
IV. SDM-CRAN MODEL

As in [20], we gather all terms from {12} that depend on the current symbols. Assuming multi-mode DBP, we write

\[
\begin{bmatrix}
\Delta x_m^{[1]} \\
\Delta x_m^{[S]}
\end{bmatrix} = J_m \begin{bmatrix}
\tau_m^{[1]} \\
\tau_m^{[S]}
\end{bmatrix} + \begin{bmatrix}
\tau_m^{[1]} \\
\tau_m^{[S]}
\end{bmatrix}
\]

(21)

where \( J_m \) is defined block-wise as

\[
J_m = \begin{bmatrix}
D_m^{[1,1]} & D_m^{[1,2]} & \cdots & D_m^{[1,S]} \\
B_m^{[2,1]} & D_m^{[2,2]} & \cdots & D_m^{[2,S]} \\
\vdots & \vdots & \ddots & \vdots \\
B_m^{[S,1]} & B_m^{[S,2]} & \cdots & D_m^{[S,S]}
\end{bmatrix}
\]

(22)

The diagonal blocks (of size \( 2 \times 2 \)) are

\[
D_m^{[s,s]} = \begin{bmatrix}
\theta_m^{[s]} \\
\psi_m^{[s]}
\end{bmatrix}
\begin{bmatrix}
\theta_m^{[s]} \\
\psi_m^{[s]}
\end{bmatrix}^T
\]

and the off-diagonal blocks are

\[
B_m^{[s,r]} = \begin{bmatrix}
\xi_m^{[s,r]} \\
\psi_m^{[s,r]}
\end{bmatrix}
\begin{bmatrix}
\xi_m^{[s,r]} \\
\psi_m^{[s,r]}
\end{bmatrix}^T
\]

(23)

(24)

where, for \( r \neq s \), we have

\[
\theta_m^{[s]} = \sum_{c,0} \sum_{k,0} C_{c,0,k,k}^s b_{c,k,m}^s b_{c,k,m}^{s,*}
\]

(25)

\[
\psi_m^{[s,s]} = \sum_{c,0} \sum_{k,0} \sum_{k',m} C_{c,0,k,k}^s \tau_{c,k,m}^{s,*} b_{c,k,m}^{s,*}
\]

(26)

\[
\xi_m^{[s,r]} = \sum_{c,0} \sum_{k,0} \sum_{k',m} D_{c,k,m}^{s,r} b_{c,k,m}^s b_{c,k,m}^{r,*}
\]

(27)

and \( \theta_m^{[s]} \), \( \psi_m^{[s,s]} \), and \( \xi_m^{[s,r]} \) are obtained by swapping \( b_{c,k,m}^s \) with \( b_{c,k,m}^{s,*} \) in (25) (27), swapping \( \tau_{c,k,m}^{s,*} \) with \( \tau_{c,k,m}^s \) in (14) (20), and shifting all delays such that \( \tau_{c,k,m}^s = 0 \). The residual NLI terms \( \psi_m^{[s]} \) and \( \tau_m^{[s]} \) are given by (12) without the first two lines (in the DBP case) and without the summands with \( n = 0 \).

We choose all pulses of the same frequency channel to be synchronized, i.e., we set \( \tau_{c,k,m}^s = \tau_m^s \) for all \( s \) and \( c \).

We thus have \( D_m^{[s,s]} = D_{c,k,m}^{s,s} = D_{c,k,m}^{s,s,*} \), and

\[
\psi_m^{[s,s]} = \tau_m^{s,*}, \quad \xi_m^{[s,r]} = \tau_m^{s,*}, \quad \xi_m^{[s,r]} = \tau_m^{r,s}.
\]

(28)

Since the \( \theta_m^{[s]} \) and \( \psi_m^{[s]} \) are real [18], the matrix \( J_m \) is Hermitian. Using (25) (27), it follows that all the diagonal terms \( \Theta_m^{[s]} \) and \( \Theta_m^{[s]} \) are correlated and all other crosscorrelations are 0.


dd

A. Mode Rotation Approximation

As in [20], we apply an LP-like approximation

\[
I + jJ_m = \exp(jJ_m).
\]

(29)

Using (21) and (11), we write the SDM-CRAN model as

\[
y_m = \exp(jJ_m)x_m + v_m + w_m
\]

(30)

where \( y_m = (y_m^{[1]}, y_m^{[2]}, \ldots, y_m^{[S]}, y_m^{[S]})^T \), and \( x_m, v_m, \) and \( w_m \) are defined similarly. The matrix \( \exp(jJ_m) \in \mathbb{C}^{2S \times 2S} \) is unitary. The components of the residual NLI \( v_m \) are uncorrelated and have memory in the time parameter \( m \), see [20]. We state results for strong coupling next.

B. Statistics of \( J_m \) for Strong Coupling

Strong coupling has \( f_{x,s'} = \kappa, \beta_2^s = \beta_2 \) and \( g_s(z) = g(z) \). Furthermore, all the coefficients of the same frequency channel \( c \), except for \( C_{c,n,k,k'}^s \), are equal and independent of \( s, s' \), and \( r \). We call these coefficients \( X_{c,n,k,k'} \).

\[
X_{c,n,k,k'} := \frac{C_{c,n,k,k'}^s}{2} = c_{c,n,k,k'}^s = \bar{c}_{c,n,k,k'}^s
\]

(31)

for all \( s, s' \) and \( r \neq s \). The means of the entries \( J_{i,k}[m] \) of the matrix \( J_m \) in (30) are

\[
\langle J_{i,i}[m] \rangle = (1 + 2S) \sum_{c,0} \sum_{k} X_{c,0,k,k'}
\]

(32)

\[
\langle J_{i,k}[m] \rangle = 0, \quad k \neq i.
\]

(33)

The auto- and cross-covariance functions of the \( J_{i,k}[m] \) can be written in terms of the following two functions:

\[
r[c,k,\ell] = \sum_{c,0} (Q_c - E_c^2) \sum_{k} X_{c,0,k,k} X_{c,0,k-k,\ell-\ell}
\]

(34)

\[
s[c,k,\ell] = \sum_{c,0} E_c^2 \sum_{k \neq k'} X_{c,0,k,k'} X_{c,0,k-k',\ell-\ell}
\]

(35)

If one defines

\[
\tau_{ik,k'}[c,\ell] = \langle J_{i,k}[m] J_{i',k'}[m+\ell] \rangle - \langle J_{i,k}[m] \rangle \langle J_{i',k'}[m+\ell] \rangle
\]

(36)

then the autocovariance functions are

\[
r_{iis}[|c,\ell|] = (3 + 2S) r[|c,\ell|], \quad k \neq i
\]

(37)

\[
r_{iik}[|c,\ell|] = s[|c,\ell|], \quad k \neq i
\]

(38)

and the other crosscorrelation functions of two diagonal elements is

\[
r_{iikk}[c,\ell] = (2 + 2S) r[|c,\ell|], \quad k \neq i
\]

(39)

and the other crosscorrelation functions are 0:

\[
r_{iik'|k}[c,\ell] = 0, \quad i' \neq k'
\]

(40)

\[
r_{iik'|k} = 0, \quad (i, k) \neq (i', k') \quad \text{and} \quad (i, k) \neq (k', i')
\]

(41)

where we recall that \( J_m \) is Hermitian. If the inputs \( B_c^{[s]} \) and \( D_{c,k}^{[s]} \) are CSCG, then we have \( r[|c,\ell|] = s[|c,\ell|] \). In the limit of large
accumulated dispersion, the approximations introduced in [23] apply (see [19], [20]):

\[
\langle J_{i,i}[m] \rangle \approx (1 + 2S) \gamma \kappa \frac{L}{T} \sum_{c \neq 0} E_c \tag{42}
\]

\[
r[\ell] \approx \frac{\gamma^2 \kappa^2 L}{T} \sum_{c \neq 0} \frac{Q_c - E^2_c}{|\beta_2 \Omega_c|} \left[ 1 - \frac{\ell T}{|\beta_2 \Omega_c|} \right]^+ \tag{43}
\]

\[
s[\ell] \approx \frac{\gamma^2 \kappa^2 L}{T} \sum_{c \neq 0} E^2_c \left[ 1 - \frac{\ell T}{|\beta_2 \Omega_c|} \right]^+ \tag{44}
\]

where one assumes that the contributions of the summands with \( X_{c;0,k,k} \) in (43) and (45) dominate, see [20, Eq. 2]. For example, in a 2-pol system (\( S = 1 \)), the ACF \( r[\ell] \) of the diagonal elements of \( J_m \) is \( 5r[\ell] \) (see (37)), with \( r[\ell] \) approximately given by (43). This matches [20, Eq. 75]).

V. SIMPLIFIED SDM-CRAN MODEL

In the following, the receiver removes the means \( \langle \Theta[m] \rangle \equiv \langle J_{2s-1,2s-1}[m] \rangle \) and \( \langle Z[m] \rangle \equiv \langle J_{2s,2s}[m] \rangle \) of the phase noise, and we abuse notation and write \( J_{i,j}[m] \) for the resulting zero-mean variables. Based on the statistical analysis of Section IV-B, we model the entries of the matrix \( J_m \) in (30) as

\[
j_{i,i}[m] = 2\phi_i[m] + \sum_{i' \neq i} \phi_{i'}[m] \tag{45}
\]

\[
j_{k,i}[m] = j_{i,k}[m] \tag{46}
\]

where the \( \phi_i[m] \), \( i = 1, \ldots, 2S \), are independent, real, Gaussian, memory-\( \mu \) Markov processes generated using the procedure in [20, Section IV-B] with autocovariance function \( r[\ell] \) in (44) for \( \ell \in \{-\mu, \ldots, \mu\} \). The \( J_{i,k}[m] \) for \( i < k \) are independent CSCG memory-\( \mu \) Markov processes generated using the same procedure with function \( s[\ell] \) in (35) for \( \ell \in \{-\mu, \ldots, \mu\} \). The number of hidden independent real Markov processes in \( J_m \) is thus \( 4S^2 \), of which \( 2S \) are the \( \Phi_i[m] \).

As in [19], [20], we combine the ASE noise and the residual NLI noise into one additive noise term \( z_m = v_m + w_m \). The simplified (mismatched) SDM-CRAN model is

\[
y_m = \exp(jJ_m)x_m + z_m. \tag{47}
\]

We model the entries \( Z_i[m] \) of \( z_m \) as independent CSCG processes with real ACFs

\[
r_{Z_i}[\ell] := \langle Z_i[m]Z_i^*[m+\ell] \rangle = N_{ASE} \delta[\ell] + \langle V_i[m]V_i^*[m+\ell] \rangle \tag{48}
\]

where \( v_i[m] \) is the \( i \)-th entry of \( v_m \). The ACF \( r_{Z_i}[\ell] \) has short memory and a similar shape as [19, Fig. 3].

VI. NUMERICAL RESULTS

A. Estimating Model Parameters

We use a training set to estimate the parameters of the model (47). Since the matrix \( \exp(jJ_m) \) in (47) is unitary, if we use i.i.d. Gaussian inputs \( x_m \) and neglect the small correlations in \( z_m \) then the distribution of \( ||Y_m||^2 \) given \( ||X_m|| \) is noncentral chi-squared with \( 4S^2 \) degrees of freedom, and independent for each \( m \). We thus estimate the noise variance \( \sigma_2^2 = r_z[0] \) as

\[
\sigma_2^2 = \arg \max_{\sigma^2} \log \left[ \frac{\exp^{-\frac{||Y_m||^2}{2\sigma^2}}}{\sigma^2} \right]. \tag{49}
\]

B. Achievable Rates

We simulated the strong coupling system (3) with \( S = 2 \) spatial modes and 5 WDM channels by using the split-step Fourier method. The system has IDA and the parameters in Table I. As in [20], we compare single- and multi-carrier systems, where the latter system has four subcarriers (4SC) of bandwidth 12.5 GHz each. All subcarriers have the same power, i.e., there is no frequency-dependent power allocation as in [19], [20]. The input symbols have a Gaussian density. The channel delays \( \tau_i \) are chosen randomly between \( -T/2 \) and \( T/2 \). The receiver applies a 50-GHz band-pass filter to isolate the COI, followed by joint DBP on all COI spatial modes and subcarriers, and then matched filtering and sampling. A training set of 24 sequences of 4092 symbols (20 sequences of 4 x 1023 symbols in the 4SC system) is used to estimate the SDM-CRAN model parameters.

We apply particle filtering on a test set of 120 sequences (100 sequences for 4SC) to compute achievable rates as described in [20, Sec. VI]. The mismatched output distribution is Gaussian [20, Eq. (91)]. The results are plotted in Fig. II that compares four receiver algorithms with successively more complex processing:

- separate processing of each of the 2S (complex-alphabet) channels using a memoryless mismatched model with i.i.d. phase-and-additive noise, see [19, Sec. VIII.A];
- separate 1-pol CPAN (1pCPAN) processing of each of the 2S channels with 2S particle filters, see [19];

\[\text{Parameter} \quad | \quad \text{Symbol} \quad | \quad \text{Value}
\]

| Parameter | Symbol | Value |
|----------|--------|-------|
| Dispersion coefficient | \( \beta_2 \) | \(-21.7 \ \text{pm}^2/\text{km}\) |
| Nonlinear coefficient | \( \gamma \) | \(1.27 \ \text{W}^{-1} \ \text{km}^{-1}\) |
| RX noise spectral density | \( N_{ASE} \) | \(5.902 \times 10^{-18} \ \text{W/Hz}\) |
| Number of spatial modes | \( S \) | 2 |
| WDM channel indexes | \( c_{\min}, c_{\max} \) | \(-2, 2\) |
| Transmitted pulse shape | \( s(\ell) \) | \(\text{sinc}\) |
| Channel bandwidth | \( B \) | 50 GHz |
| Channel spacing | \( \Omega(1)/(2\pi) \) | 50 GHz |
separate 2-pol CP AN (2pCP AN) processing of the spatial modes with $S$ particle filters, see [20];
• joint CRAN processing of all $2S$ channels with one particle filter that uses the mismatched model (45)-(48).

With respect to memoryless processing, the 1pCPAN model gains 0.29 bits/s/Hz/channel, the 2pCPAN model gains a further 0.07 bits/s/Hz/channel, and the SDM-CRAN model gains another 0.17 bits/s/Hz/channel. 4SC gains between 0.07 and 0.1 bits/s/Hz/channel with respect to single-carrier transmission. The rate gain from the peak of the lowermost curve to the peak of the uppermost curve is approximately 0.6 bits/s/Hz/channel. The rate gain from “Memoryless” to “SDM-CRAN” for either single- or multi-carrier transmission is 0.5 bits/s/Hz/channel. For 4SC, the power gain from the rate peak with memoryless processing to the SDM-CRAN curve at the same rate (7.79 bits/s/Hz/channel) is 1.4 dB.

VII. CONCLUSIONS

We extended the analysis of the CP AN model [19], [20] to SDM for weak and strong coupling. The SDM-CRAN model and a multi-carrier scheme were applied to strong coupling to obtain achievable rates for a $S = 2$ spatial mode system. The rates are 0.6 bits/s/Hz/channel larger than the rates for a single-carrier system with separate and memoryless processing per complex-alphabet channel.

We remark that computational complexity limits the numerical calculations to a small number $S$ of spatial modes. An important direction for future work is speeding up the calculations and designing simplified receivers that exploit the correlations predicted by the SDM-CRAN model in practical multi-mode systems. An interesting theory question is to extend the LP analyses of [14]–[16] to SDM.