Coherent transmission channels as 4d rotations

Coherent transmission channels as 4d rotations

Magnus Karlsson (1), Cristian B. Czegledi (2), Erik Agrell (2)

(1) Department of Microtechnology and Nanoscience, (2) Department of Signals and Systems Chalmers University of Technology, SE-412 96 Göteborg, Sweden

email:magnus.karlsson@chalmers.se

Abstract: We present and discuss a real four-dimensional (4d) channel model for coherent links, and compare it with the more conventional complex two-dimensional (Jones matrix) model.

OCIS codes: (060.0060) Fiber optics and optical communications; (060.2330) Fiber optics communications

1. Introduction

The rise of coherent communications since the emergence of fast digital signal processing (DSP)-based coherent receivers [1] made transmission of quadrature phase-shift keying (QPSK) data in two polarizations, so-called polarization multiplexed QPSK, PM-QPSK, possible. This modulation format can be interpreted in different and seemingly equivalent ways, e.g., as four independent binary phase shift keying (BPSK) channels transmitted in parallel, or as two independent QPSK channels, or as a single 16-ary constellation, formed by the vertices of the four-dimensional (4d) hypercube [2, 3]. The latter of these descriptions is the richest, in the sense that it provides a full description of the transmitted constellation, and can account for all linear imperfections that intermixes the four quadratures (in-phase/quadrature (I/Q) component imbalance, polarization-dependent effects, phase noise, etc.). An optimum receiver will thus need to account for the whole 4d constellations rather than four 1d projections (as in the 4xBPSK model) or two 2d projections (as in the 2xQPSK model). Therefore it is often necessary, especially in the design of the receiver DSP, to have a full 4d channel model rather than a set of parallel 1d or 2d models.

There are still two major ways to express the discrete-time, linear channel model, even for the full 4d signal. The most common one is to use the 2d complex description, known in optics as the *Jones-vector* space. There the received signal amplitude is described as $\boldsymbol{e}_r = T\boldsymbol{e}_t + \boldsymbol{n}$, where $\boldsymbol{e}_t$ denotes the 2d transmitted complex vector (Jones vector), $\boldsymbol{e}_r$ is the received Jones vector [4], and $\boldsymbol{n}$ is a 2d vector of complex additive white Gaussian noise.[1] The transfer matrix (Jones matrix [4]) is denoted T and is a 2x2 matrix with complex elements. An alternative (but not fully equivalent, which is the point of this work) channel model would be to use a fully real, but 4d, description as $\vec{E}_r = R\vec{E}_t + \vec{N}$, where the vectors $\vec{E}_r, \vec{E}_t, \vec{N}$ are the corresponding real 4d vectors and R is a 4x4 transfer matrix. The 4d vectors $\vec{E}_r, \vec{E}_t$ consist of the I/Q components of the x/y-polarization components, so they contain exactly the same information as the Jones vectors $\boldsymbol{e}_r, \boldsymbol{e}_t$. The 4d channel model was used recently in, e.g., [5, 6]. The two channel models might seem equivalent, but somewhat surprisingly, the real 4d model is *richer* than the complex 2d model. This is evident just by comparing the number of elements in the transfer matrices; T has 4 complex, i.e., 8 real, parameters whereas R has 16 real parameters. If we limit ourselves to lossless transformations, i.e., where the total signal power $P = |\boldsymbol{e}_t|^2 = |\vec{E}_t|^2$ is unchanged by the transfer matrix, then one may show that T must be a *unitary* matrix, satisfying $T^\dagger T = I$, which leaves 4 free parameters, and R an *orthogonal* matrix, satisfying $R^t R = I$, which leaves 6 free parameters.

In this contribution, we discuss these differences between the channel models in some more detail, and explain why they arise and how they may possibly be utilized in improving existing channel models and coherent DSP. The paper is in part based on work presented in [7, 8], but we also describe some novel results and show how to create a random channel matrix that is useful when simulating coherent transmission systems and subsystems. For example, if we want to simulate a channel in a random phase and polarization state, what probability distribution shall the elements of T or R have, and how can the appropriate matrices be realized?

2. The Jones and 4d channel models

We limit ourselves to the power-preserving case, i.e., when T is unitary and R orthogonal. We will first describe the conventional Jones channel and its random properties, and then deal with the 4d model.

[1] The notation is the same as in [7], using upper-case arrowing for 4d column vectors, lower-case arrowing for 3d row vectors, and lower-case bold for 2d complex vectors. Arrowed greek letters denote 3d row vectors with matrices as components.

2.1. *The Jones channel model*

We can parametrize the 2x2 channel matrix T in terms of a common (for both polarizations) phase change ϕ and a polarization change described by the unitary matrix $U(\vec{a})$, parametrized by the real 3d vector $\vec{a} = (a_1, a_2, a_3) = a\hat{a}$ with length a and unit vector direction $\hat{a}$. Thus[2]

$$T = \exp[i\phi - i\vec{a}\cdot\vec{\sigma}] = \exp(i\phi)U(\vec{a}) = \exp(i\phi)[I\cos(a) - i\hat{a}\cdot\vec{\sigma}\sin(a)]. \tag{1}$$

We use here the matrix exponential function, formally defined for matrices via its Taylor series, but summed to closed form in the left-hand side of (1), as it allows for an elegant and compact connection between the two Jones and 4d models. The unit matrix is denoted I here and elsewhere, with size (2x2 or 4x4) evident from context. The notation $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3) = (\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & -i\\ i & 0\end{smallmatrix}\right))$ is used for the 3d vector formed by the three Pauli matrices. The inner product $\vec{a}\cdot\vec{\sigma}$ is a compact expression for an arbitrary linear combination of the Pauli matrices. In Stokes space, the corresponding polarization change would be described by a rotation of the Poincaré sphere by an angle $2a$ around the axis $\hat{a}$. This means that to cover all possible polarization changes, $a \in [0, \pi]$ and $\hat{a} \in S_3$, where S_3 denotes the surface of the unit 3d sphere. Another useful way of describing unitary matrices is via unit quaternions [9]. The quaternion corresponding to U is the unit-length 4d vector $\vec{U} = (\cos(a), \sin(a)\hat{a})^t$, which means that $\vec{U}$ lies on the surface S_4 of the 4d unit sphere. Note that this has nothing to do with the 4d channel model discussed below, it just so happens that every unitary 2x2 matrix corresponds to a point on the surface of the 4d unit sphere. The angular distance between two such vectors can be used as a convenient metric of the "closeness" of two unitary matrices. Furthermore, the four components of $\vec{U}$ can be identified as the real and imaginary parts of the first row of the elements of the matrix $U(\vec{a})$.

Now what if we want to describe a *uniformly random* channel matrix T, what does that mean? We mean that a given input signal vector $\boldsymbol{e}_t$ is transformed to any other signal vector $\boldsymbol{e}_r$ by a uniform probability. From symmetry arguments it can be shown that this is obtained for a unitary matrix U whose corresponding 4d vector $\vec{U}$ is isotropically distributed on S_4. Such a vector is straightforward to create by generating a real 4d vector $\vec{X}$ with 4 independent normal-distributed components, and then forming $\vec{U} = \vec{X}/|\vec{X}|$. The corresponding vector $\vec{a}$ is then obtained from $\vec{U} = (\cos(a), \sin(a)\hat{a})^t$. It can be shown that the resulting probability density function (pdf) for each component of $\vec{U}$ is $f_x(x) = (2/\pi)\sqrt{1-x^2}$, $x \in [-1, 1]$, and that the pdf for the rotation angle a is $f_a(a) = (2/\pi)\sin^2(a)$, $a \in [0, \pi]$. The pdf for ϕ is uniform in $[0, 2\pi]$.

2.2. *The 4d channel model*

The 4d rotations form the group of *orthogonal* 4d matrices. This group has six free parameters, which can be separated in two commuting subgroups (*right-* and *left-isoclinic*), with three parameters each. An isoclinic rotation rotates all vectors the same angle (which is not possible for 3d rotations). The two subgroups are useful here for two reasons: (i) the two groups commute, which facilitates analysis, and (ii) there is a direct mapping between the unitary Jones matrices $U(\vec{a})$ and the right-isoclinic rotations. To be specific, we can write the most general 4d rotation as

$$R(\vec{a}, \vec{b}) = \exp[-\vec{b}\cdot\vec{\lambda} - \vec{a}\cdot\vec{\rho}] = \exp[-\vec{b}\cdot\vec{\lambda}]\exp[-\vec{a}\cdot\vec{\rho}], \tag{2}$$

where $\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ and $\vec{\rho} = (\rho_1, \rho_2, \rho_3)$ are 3d vectors with suitably chosen 4x4 permutation matrices (defined in [7]) as components, i.e., they are "vectors of matrices", similar to $\vec{\sigma}$. The vectors $\vec{a}, \vec{b}$ are the rotation parameters. The special case $\vec{b} = (\phi, 0, 0)$ is

$$R = \exp[-(\phi, 0, 0)\cdot\vec{\lambda} - \vec{a}\cdot\vec{\rho}] = [I\cos(\phi) - \lambda_1\sin(\phi)][I\cos(a) - \hat{a}\cdot\vec{\rho}\sin(a)], \tag{3}$$

and it is the transformation equivalent to the more familiar 2d complex model (1). The direct similarity between (1) and (3) reveals that the above results regarding random transfer matrices holds also for the 4d matrix; pick $(\cos(a), \sin(a)\hat{a})$ uniformly from S_4 and ϕ uniformly in $[0, 2\pi]$.

It is instructive to express the resulting rotation angle for a signal vector $\vec{E}$ by the rotation (3) relative to its input. It gives the relevant movement in signal space from the channel, and can be used to estimate how much signals move relative to, e.g., decision thresholds or signals that propagate through a different channel matrix. The angular cosine between the vector and the rotated vector $\vec{E}^t R\vec{E} = P\cos(\gamma)$ can after some algebra be expressed as

$$P\cos(\gamma) = P\cos(\phi)\cos(a) + (\hat{a}\cdot\vec{e})\sin(\phi)\sin(a) \tag{4}$$

[2] Eq (1) corrects a sign error of the phase angle α_0 (here denoted ϕ) in the expression for the T matrix in row 1 of Table I in [7,8].

where $\vec{e} = \boldsymbol{e}^\dagger \vec{\sigma} \boldsymbol{e} = \vec{E}^t \lambda_1 \vec{\rho} \vec{E}$ is the *Stokes vector* corresponding to the 4d signal $\vec{E}$ and the Jones vector $\boldsymbol{e}$. The expression (4) can alternatively be obtained from the Jones vector model by taking the real part of the complex number $\boldsymbol{e}^\dagger T \boldsymbol{e}$, using T from (1). It is interesting to note some special cases of this result:

(i) The rotation angle lies between the extremes $\phi \pm a$, depending on the alignment of the Stokes vector $\vec{e}$ and the vector $\hat{a}$.

(ii) The pure polarization rotation ($\phi = 0$) causes a movement by an angle $\gamma = a$, *independently* of $\vec{E}$, which is counterintuitive as the corresponding Stokes vector $\vec{e}$ will move an angle between 0 and $2a$ depending on its alignment with $\hat{a}$. The resolution of this paradox lies in the fact that when the polarization (in terms of the Stokes vector) changes less, the signal acquires a phase shift that makes up for this difference, keeping γ the same for all input vectors. This is also, indeed, the definition of a pure isoclinic transformation. However, the general transformation, with arbitrary phase (ϕ) and polarization (a) shifts is *not* isoclinic.

3. The unphysical rotations

Clearly, the second and third components of $\vec{b}$ in (2) are not needed to describe the changes of the optical field during propagation. In fact, it can be shown [7] that rotations (2) where the second and third components of $\vec{b}$ are nonzero, are incompatible with the boson commutation relations, and thus cannot arise for passive optical devices. Nonetheless they can be synthesized in DSP. An example of such an "unphysical" rotation is, e.g., exchanging the real parts of the x- and y-polarization components while leaving the imaginary parts unaffected. Other examples involving modulation formats were discussed in [7], e.g., transforming single-polarization QPSK to dual-polarization BPSK. Also, completely novel realizations of known formats can be created that move modulation between polarization changes and phase changes.

The unphysical set of rotations illustrates that the four quadratures of the electromagnetic field are not allowed to be changed and routed in a completely arbitrary fashion. A well-known manifestation of this is the fact that a passive polarization beam splitter can be used to (losslessly) demultiplex the two polarizations, but a similar passive device to demultiplex the I and Q components from a single polarization component does not exist (even if designing *active* such subsystems is a vivid field of research [10–12]).

4. Summary

This paper compares and contrasts the conventional 2d complex (Jones matrix) channel model for coherent transmission links with the real 4d channel model, and presents the relevant random channels. The richer space of the 4d transformations can be useful for developing new transmitter and receiver algorithms for coherent links.

We gratefully acknowledge funding from the Swedish Research Council, and the Knut and Alice Wallenberg Foundation. We also acknowledge inspiring discussions with colleagues within the Chalmers FORCE center.

References

1. H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," Opt. Express **16**, 873–879 (2008).
2. S. Betti, F. Curti, G. De Marchis, and E. Iannone, "A novel multilevel coherent optical system: 4-quadrature signaling," J. Lightwave Tech. **9**, 514–523 (1991).
3. E. Agrell and M. Karlsson, "Power-efficient modulation mormats in moherent transmission systems," J. Lightwave Tech. **27**, 5115–5126 (2009).
4. R. C. Jones, "A new calculus for the treatment of optical systems I. Description and discussion of the calculus," J. Opt. Soc. Am. **31**, 488–493 (1941).
5. H. Louchet, K. Kuzmin, and A. Richter, "Joint carrier-phase and polarization rotation recovery for arbitrary signal constellations," IEEE Phot. Technol. Lett. **26**, 922–924 (2014).
6. C. B. Czegledi, E. Agrell, and M. Karlsson, "Symbol-by-symbol joint polarization and phase tracking in coherent receivers," in Proc. OFC/NFOEC, paper W1E (2015).
7. M. Karlsson, "Four-dimensional rotations in coherent optical communications," J. Lightwave Technol. **32**, 1246–1257 (2014).
8. M. Karlsson, "The connection between polarization calculus and four-dimensional rotations," arXiv:1303.1896 (2013).
9. M. Karlsson and M. Petersson, "Quaternion approach to PMD and PDL phenomena in optical fiber systems," J. Lightwave Tech. **22**, 1137–1146 (2004).
10. A. Lorences-Riesgo, L. Liu, S. L. I. Olsson, R. Malik, A. Kumpera, C. Lundström, S. Radic, M. Karlsson, and P. A. Andrekson, "Quadrature demultiplexing using a degenerate vector parametric amplifier," Opt. Express **22**, 29424–29434 (2014).
11. N.-K. Kjøller, M. Galili, K. Dalgaard, H. C. H. Mulvad, K. M. Røge, and L. K. Oxenløwe, "Quadrature decomposition by phase conjugation and projection in a polarizing beam splitter," in Proc. ECOC, paper P.3.8 (2014).
12. F. Parmigiani, R. Slavík, G. Hesketh, P. Petropoulos, and D. Richardson, "Quadrature decomposition of optical fields using two orthogonal phase sensitive amplifiers," in Proc. ECOC, paper Tu4.6.2 (2014).