Phase Retrieval-Based Coherent Receivers

Elaine S. Chou, Hrishikesh Srinivas and Joseph M. Kahn

Department of Electrical Engineering
E. L. Ginzton Laboratory
Stanford University

ee.stanford.edu/~jmk

OFC 2021 • Paper Th5F.1 • June 10, 2021

Research supported by

- Maxim Integrated
- Inphi Corporation (now Marvell Technology, Inc.)

Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Downconversion and detection

- **Downconversion**: extracting electrical photocurrents from optical fields.
- **Detection**: extracting data from electrical photocurrents.

Downconversion using a local oscillator	Downconversion without a local oscillator
Photocurrents are amplified by strong LO.	Photocurrents are not amplified by LO.
Photocurrents are linear in signal fields. After downconversion, can electrically: • Select a desired wavelength channel. • Demultiplex dual-polarization signals. • Detect dual-quadrature signals. • Compensate chromatic dispersion.	Photocurrents are quadratic in signal fields. Before downconversion, must optically: • Select a desired wavelength channel. • Demultiplex dual-polarization signals. If a carrier is transmitted with signal, can electrically: • Detect dual-quadrature signals. • Compensate chromatic dispersion.

Downconversion and detection methods

Detection Method	Noncoherent	Differentially Coherent	Hybrid Non-/Diff. Coherent	Coherent
Measures	Energies in signal dimension(s)	Phase differences between signal dimensions	Energies and phase differences	Field quadratures in signal dimensions
Modulation Methods	OOK, PAM, OFDM, WD-FSK	DPSK, CPFSK	PolSK, etc.	QAM, PSK, OFDM, etc.
Signal Dimensions in Two Polarizations	Up to 2	Up to 2	Up to 4	Up to 4
LO-Based Downconversion	Envelope detection	Delay-and-multiply detection	PolSK detection	Standard coherent detection
LO-Free Downconversion	Standard direct detection	Delay interferometer + direct detection	Stokes vector detection	Kramers-Kronig detection

This study

This study

Compares three detection methods

- Standard coherent, Kramers-Kronig, standard direct.

Scenario

- Amplifier noise dominant. Single wavelength. Single polarization (for now).
- No constraints on complexity (for now). Numerically approximate ideal continuous-time waveforms and signal processing operations.
- Scenario is generous to KK, which needs amplifiers to compete and has high complexity.

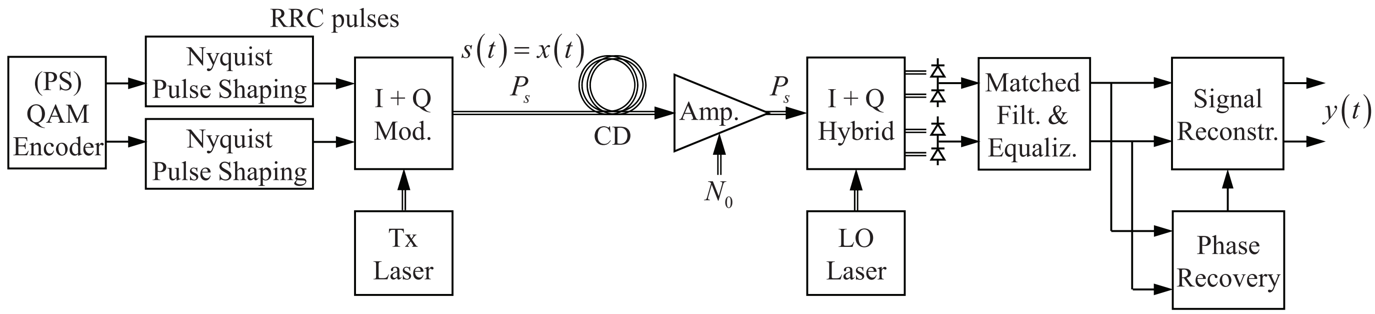
Aspects studied

- Mutual information (b/symbol) and effective degrees of freedom (dimensions/symbol).
- Transmitted signal distribution: probabilistically shaped or mutual information-maximizing.
- Joint optimization of carrier-to-signal power ratio and probabilistic shaping.
- Impact of chromatic dispersion.

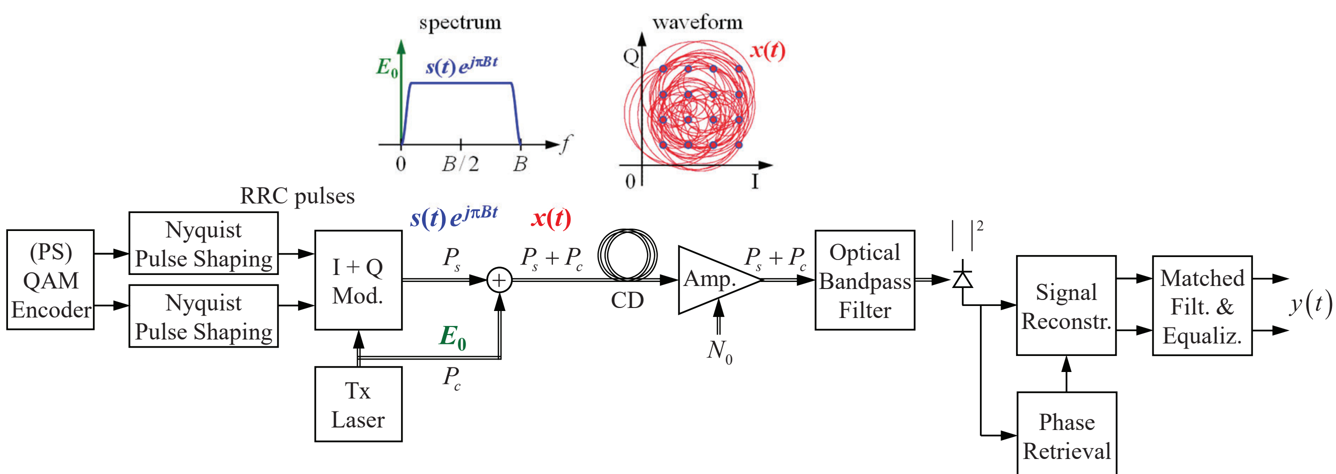
Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Standard coherent detection (one polarization)



Kramers-Kronig detection (one polarization)



- Neglecting dispersion and noise, received field is

$$x(t) = E_0 + s(t)e^{j\pi Bt}$$

and received intensity is

$$i(t) = |x(t)|^2.$$

- Assume $x(t)$ is single-sideband and does not encircle origin.

- Can recover phase as

$$\varphi(t) = H\left\{\log\sqrt{i(t)}\right\}$$

and recover signal as

$$s(t) = \left(\sqrt{i(t)}e^{j\varphi(t)} - E_0\right)e^{-j\pi Bt}.$$

A. Mecozzi, C. Antonelli and M. Shtaif, *Optica* **3** (2016).

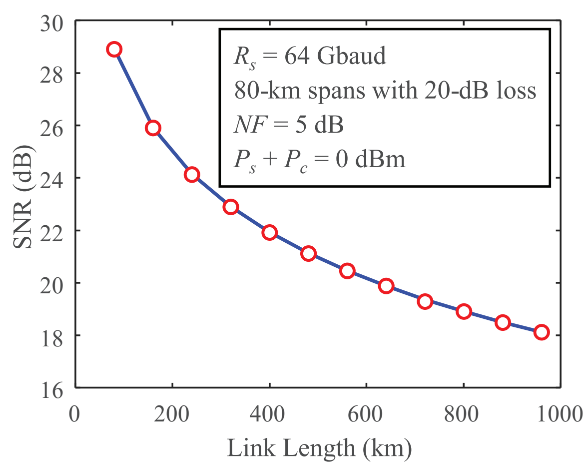
Simulation parameters for SC and KK detection

Parameter	Symbol	Value	Comment
Symbol rate	R_s	64 Gbaud	
Oversampling rate	r_{os}	64	Generous to KK, which requires higher rates than SC or SD.
QAM order	M	16, 64	
Excess bandwidth	β	0.01	Optimization of β yields small improvements.
Carrier-to-signal power ratio	$CSPR$	optimized	For KK only.
Probabilistic shaping parameter	λ	optimized	Used in Maxwell-Boltzmann distribution.
Chromatic dispersion	D	17.4 ps/nm·km	Standard single-mode fiber at 1550 nm.

Signal-to-noise ratio

- Signal-to-noise ratio in bandwidth equal to symbol rate in one polarization:

$$SNR = \frac{P_s + P_c}{N_0 R_s} = \frac{P_s}{N_0 R_s}\left(1 + \frac{P_c}{P_s}\right) = SNR_{signal}\left(1 + CSPR\right)$$



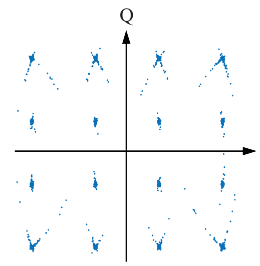
- Carrier-to-signal power ratio:

$$CSPR = \frac{P_c}{P_s}$$

Quantifying system performance

- Transmit a dense constellation. From received samples, estimate mutual information per symbol:

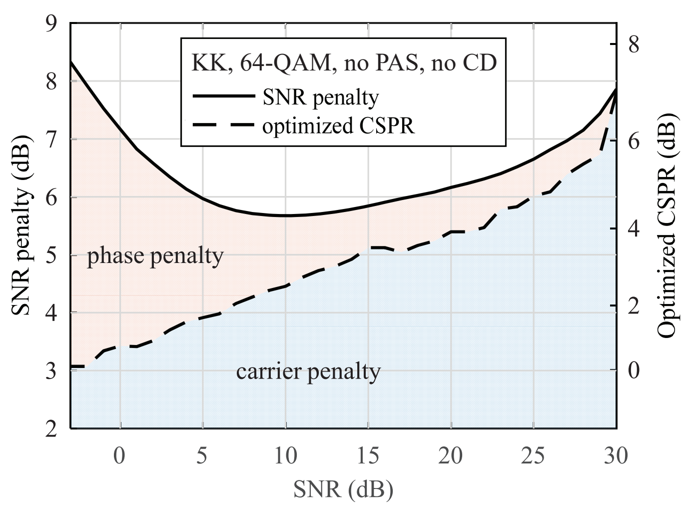
$$I(X;Y) = \mathbf{E}_{X,Y}\left[\log_2 \frac{f_{Y|X}(Y|X)}{f_Y(Y)}\right] = \sum_{x\in\mathcal{X}} p_X(x) \int_{\mathbb{C}} f_{Y|X}(y|x) \log_2 \frac{f_{Y|X}(y|x)}{f_Y(y)} dy$$



Ignore memory, so may underestimate mutual information very slightly.

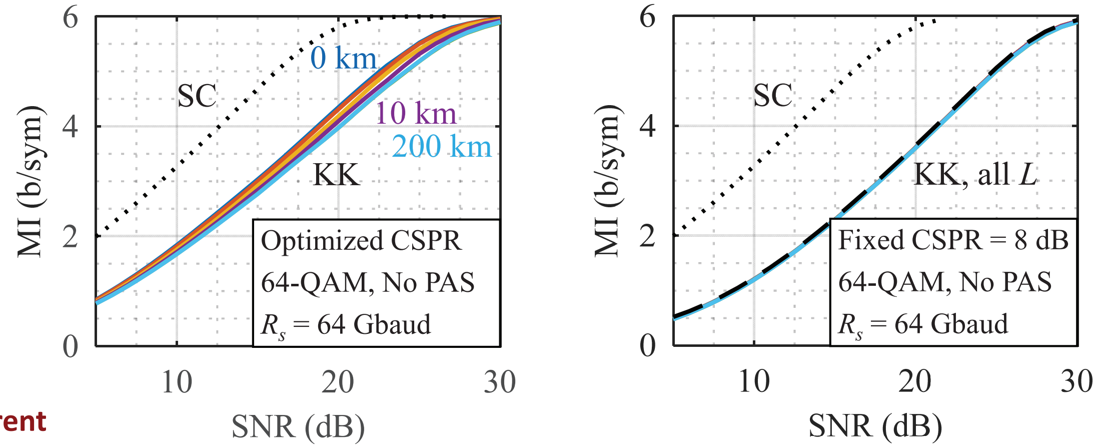
- Do not assume Gaussian noise *a priori*, since phase retrieval errors are non-Gaussian.
- Mutual information is a more fundamental metric than generalized mutual information. Both MI and GMI track achievable rates for bit-interleaved coded modulation at code rates ≥ 0.7.
- Quantify mutual information per symbol (b/symbol) instead of spectral efficiency (b/s/Hz). Generous to KK, which requires guard bands to allow optical demultiplexing of WDM channels (so does standard direct detection).

CSPR optimization for KK (no PAS, no CD)



- CSPR too low: phase retrieval error penalty dominates.
- CSPR too high: carrier power penalty dominates.
- CSPR optimal: the two penalties are balanced. As SNR increases, optimized CSPR increases to improve phase-retrieval accuracy.

Impact of chromatic dispersion (no PAS)



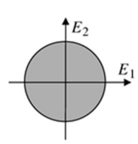
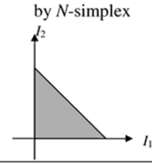
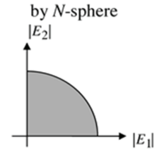
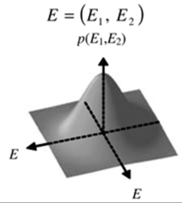
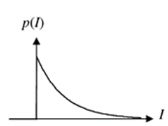
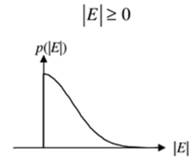
Standard coherent

- Negligible penalty from CD.

Kramers-Kronig

- Fixed CSPR > 7 dB: CD causes penalty < 0.1 dB.
- Optimized CSPR: CD increases peak-to-average ratio, requiring higher CSPR, causing penalty up to 1.7 dB. Most of the penalty is incurred in the first 10 km.

Canonical shaping problems in optical communications

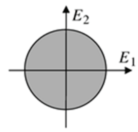
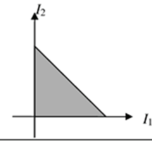
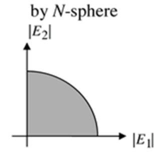
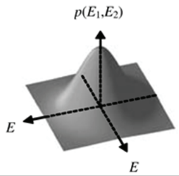
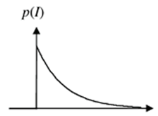
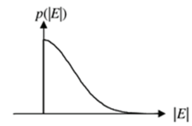
Detection Method & Dominant Noise	Coherent Detection, Amplifier or LO Shot Noise	Noncoherent Detection, Thermal Noise	Noncoherent Detection, Amplifier or LO Shot Noise
Constituent Constellation	2-D constellation with electric fields $E_i, i = 1, 2$ as coordinates	1-D constellation with field intensity $I = \|E\|^2$ as coordinate	1-D constellation with field magnitude $\|E\|$ as coordinate
Optimal Shaping Region	N-sphere centered at the origin 	Nonnegative orthant bounded by N-simplex 	Nonnegative orthant bounded by N-sphere 
Induced Optimal Signaling Distribution in Constituent Constellation	$p(E) = \frac{1}{\pi P}\exp\left(-\frac{\|E\|^2}{P}\right)$, $E = (E_1, E_2)$ 	$p(I) = \frac{1}{P}\exp\left(-\frac{I}{P}\right)$, $I \geq 0$ 	$p(\|E\|) = \sqrt{\frac{2}{\pi P}}\exp\left(-\frac{\|E\|^2}{2P}\right)$, $\|E\| \geq 0$ 
Ultimate Shape Gain	$\pi e/6 = 1.53$ dB	$e/2 = 1.33$ dB	$\pi e/6 = 1.53$ dB
Optimal Distribution	Complex circular Gaussian	Exponential	Half-Gaussian
Key Works	Forney *et al.*, 1984-89 Calderbank & Ozarow, 1990 Kschischang & Pasupathy, 1993	Shiu & Kahn, 1999 Hranilovic & Kschischang, 2003	Mao & Kahn, 2008

First shaping. Used here for SC and KK.

Discussed here for SD.

Adapted from
W. Mao and J. M. Kahn,
Trans. Comm. **56** (2008).

Canonical shaping problems in optical communications

Detection Method & Dominant Noise	Coherent Detection, Amplifier or LO Shot Noise	Noncoherent Detection, Thermal Noise	Noncoherent Detection, Amplifier or LO Shot Noise
Constituent Constellation	2-D constellation with electric fields $E_i, i = 1, 2$ as coordinates	1-D constellation with field intensity $I = \|E\|^2$ as coordinate	1-D constellation with field magnitude $\|E\|$ as coordinate
Optimal Shaping Region	N-sphere centered at the origin 	Nonnegative orthant bounded by N-simplex 	Nonnegative orthant bounded by N-sphere 
Induced Optimal Signaling Distribution in Constituent Constellation	$p(E) = \frac{1}{\pi P}\exp\left(-\frac{\|E\|^2}{P}\right)$, $E = (E_1, E_2)$ 	$p(I) = \frac{1}{P}\exp\left(-\frac{I}{P}\right)$, $I \geq 0$ 	$p(\|E\|) = \sqrt{\frac{2}{\pi P}}\exp\left(-\frac{\|E\|^2}{2P}\right)$, $\|E\| \geq 0$ 
Ultimate Shape Gain	$\pi e/6 = 1.53$ dB	$e/2 = 1.33$ dB	$\pi e/6 = 1.53$ dB
Optimal Distribution	Complex circular Gaussian	Exponential	Half-Gaussian
Key Works	Forney *et al.*, 1984-89 Calderbank & Ozarow, 1990 Kschischang & Pasupathy, 1993	Shiu & Kahn, 1999 Hranilovic & Kschischang, 2003	Mao & Kahn, 2008

First shaping in optical communications. Relevant for intra-data center, access and free-space links.

Adapted from
W. Mao and J. M. Kahn,
Trans. Comm. **56** (2008).

Capacity-achieving distribution and PAS for KK detection

- In Kramers-Kronig detection, the total signal is

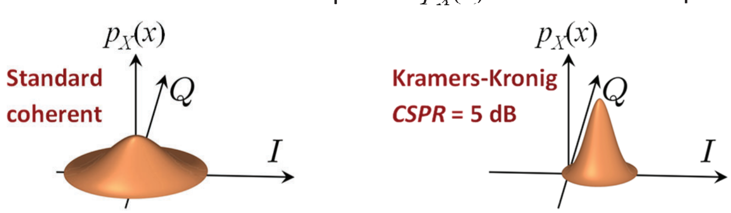
$$x(t) = s(t)e^{j\pi Bt} + E_0 .$$

The modulated portion $s(t)e^{j\pi Bt}$ and the carrier E_0 do not overlap, so the total power is $P_s + P_c$.

- The carrier E_0 conveys no information and $e^{j\pi Bt}$ preserves information. The optimal distribution for $s(t)$ satisfies

$$\max_{p_S(s)} I(S;Y) \quad \text{s. t.} \quad P_s \leq P_{max} - P_c .$$

- The optimal $p_S(s)$ is a complex circular Gaussian. The optimal $p_X(x)$ is a shifted complex circular Gaussian.



- Approximate the Gaussian by a Maxwell-Boltzmann*

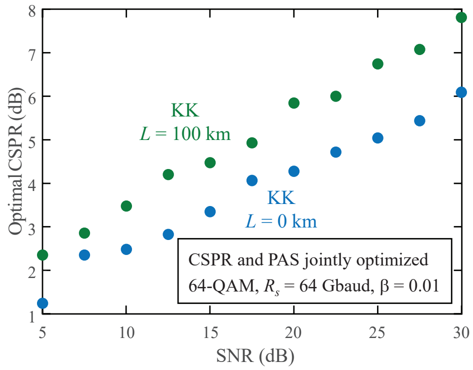
$$p_S(s) = \exp\left(-\lambda|s|^2\right) \Big/ \sum_{i=1}^{M} \exp\left(-\lambda|s_i|^2\right)$$

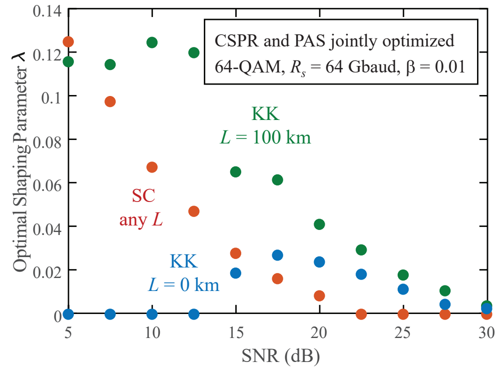
- For KK, numerically optimize λ jointly with CSPR and other parameters. For SC, numerically optimize λ.

*F. R. Kschischang and S. Pasupathy, *IEEE Trans. Info. Thy.* **39** (1993).

Joint optimization of CSPR and PAS for KK detection

- We jointly optimized PAS parameter λ, $CSPR$ and excess bandwidth β. We fix $\beta = 0.01$ in this talk.





Key principles

- PAS is helpful when SNR is too low to support full entropy of uniform constellation.
- Both PAS and CD increase peak-to-average ratio of signal, necessitating higher CSPR.

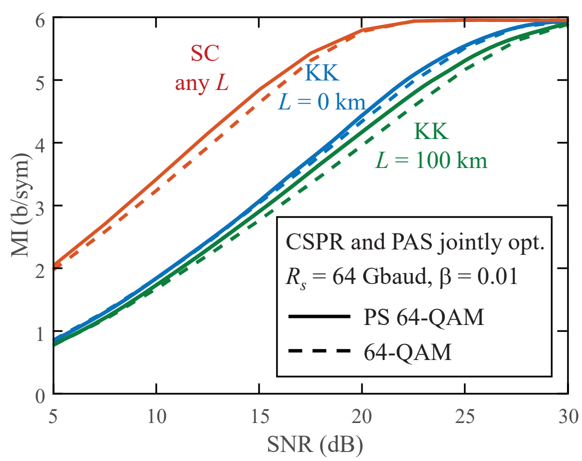
No chromatic dispersion

- Low SNR: KK uses weaker PAS than SC. High SNR: KK uses stronger PAS than SC.

High chromatic dispersion

- Low to medium SNR: KK uses stronger PAS than SC.

Mutual information using jointly optimized PAS and CSPR



No chromatic dispersion

- KK benefits significantly from PAS only for SNR ≥ 15 dB.
- KK incurs SNR penalty compared to SC of ~6 dB (with or without PAS).

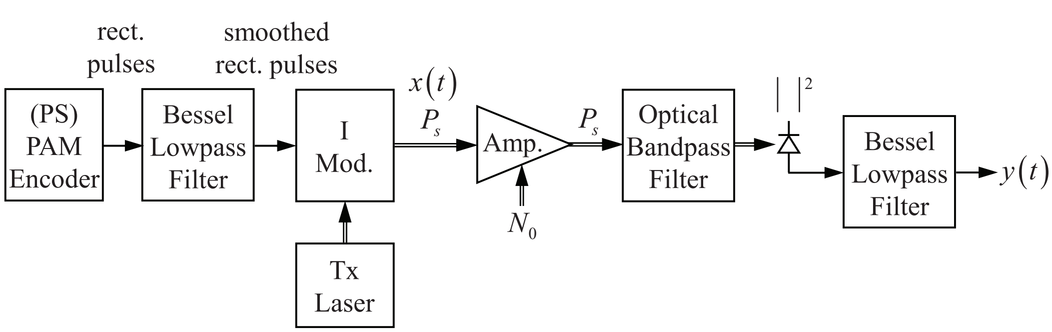
High chromatic dispersion

- KK benefits from PAS down to lower SNR values.
- KK incurs additional SNR penalty compared to SC up to 1.1 dB (with PAS) or 1.7 dB (without PAS).

Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Standard direct detection (one polarization)



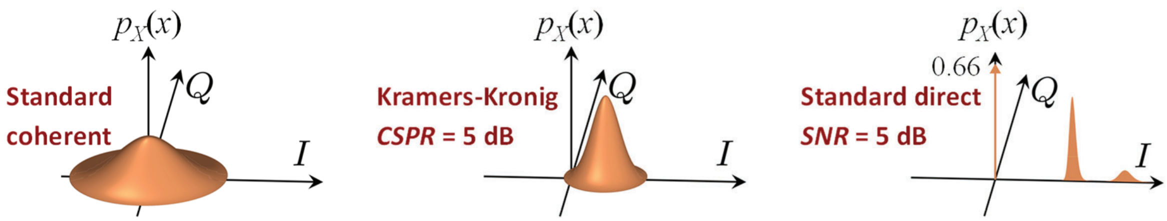
Optimization of standard direct detection

Coherent detection with LO shot noise or amplifier noise

- Capacity-achieving distribution (analytical) is Gaussian. [1]
- Optimal shaping distribution (analytical, high SNR) is also Gaussian. [2]

Noncoherent detection with amplifier noise

- Capacity-achieving distribution is obtained numerically using non-central χ^2 with two degrees of freedom.[3] Includes discrete component at zero intensity and continuous component at nonzero intensity.
- Optimal shaping distribution (analytical, high SNR) is half-Gaussian.[4] It approaches capacity only at high SNR ≥ 20 dB.



1. C. E. Shannon, *Bell Syst. Tech. J.* **27** (1948).
2. G. D. Forney *et al.*, *IEEE J. Sel. Areas Comm.* 2 (1984).
3. K.-P. Ho, *Photon. Technol. Lett*. **17** (2005).
4. W. Mao and J. M. Kahn, *IEEE Trans. Comm*. **56** (2008).

Quantifying performance of standard direct detection

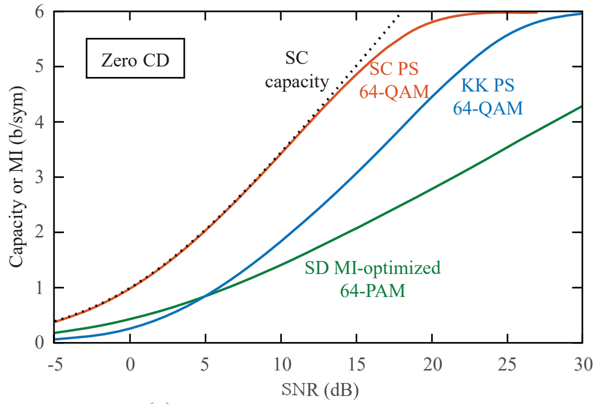
- Numerically compute capacity-achieving distribution using non-central χ^2 with 2 degrees of freedom.[1] Transmit a dense PAM constellation probabilistically shaped by the capacity-achieving distribution.
- Estimate mutual information by binning samples into histograms and estimating conditional entropy.
- Non-negative band-limited root-Nyquist pulses do not exist.[2] This complicates representing a continuous-time system by a discrete memoryless channel.
- Simulate a continuous-time system using rectangular pulses filtered by a five-pole Bessel lowpass filter. Optimize the cutoff frequency:

$$R_s \quad \text{at} \quad \text{SNR} = -5 \text{ dB} \qquad\qquad 1.6\, R_s \quad \text{at} \quad \text{SNR} = 25 \text{ dB}$$

- The continuous-time system has an average mutual information loss of about 5% compared to discrete memoryless channel at same SNR.

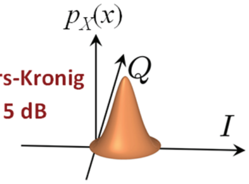
1. K.-P. Ho, *Photon. Technol. Lett*. **17** (2005).
2. S. Hranilovic, *IEEE Trans. Comm*. **55** (2007).

Comparing Kramers-Kronig to standard direct detection

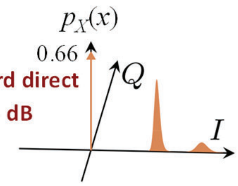


- Standard direct detection outperforms Kramers-Kronig detection for SNR < 5 dB.

Kramers-Kronig
***CSPR* = 5 dB**



Standard direct
***SNR* = 5 dB**



- Provides nearly 2 real dimensions.
- Cannot use zero intensity level.
- Wins at high SNR.

- Provides 1 real dimension.
- Can use zero intensity level, which has least noise.
- Wins at low SNR.

Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Effective degrees of freedom

- The number of complex dimensions actively conveying information can be estimated by*

$$EDOF(SNR) = \frac{d}{d\delta} MI\left(2^{\delta} \cdot SNR\right)\Big|_{\delta=0}$$

Low SNR: $EDOF$ limited by available power.

High SNR: $EDOF$ limited by available dimensions.

Standard coherent detection

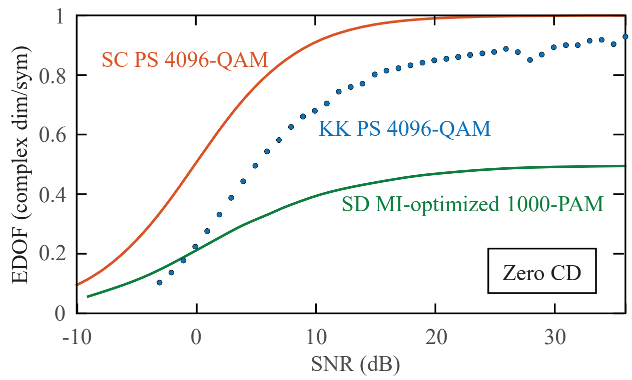
- $EDOF \leq 1$ (2 real dimensions).

Kramers-Kronig detection

- $EDOF \leq 0.9$ (1.8 real dimensions) over the SNR range studied.

Standard direct detection

- $EDOF \leq 1/2$ (1 real dimension).



* D.-S. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, *IEEE Trans. Comm.* **48** (2000).

Outline

- Introduction
- Kramers-Kronig vs. standard coherent detection
- Kramers-Kronig vs. standard direct detection
- Effective degrees of freedom
- Discussion and conclusion

Key findings

- KK detection is coherent at high SNR: yields nearly 2 real dimensions per symbol.
- KK detection benefits from joint optimization of CSPR and PAS.

Kramers-Kronig vs. standard coherent detection

Zero CD

- KK incurs SNR penalty of ~6 dB compared to SC.
- Optimization of PAS for KK different from that for SC.

High CD

- KK incurs additional SNR penalty of 1 to 1.7 dB compared to SC.
- Optimization of PAS for KK similar to that for SC.

Kramers-Kronig vs. standard direct detection

- KK outperforms SD at high SNR.
  SD outperforms KK at SNR < 5 dB.

Dual-polarization systems

- A coherent receiver – standard or KK – can electrically demultiplex dual-polarization signals.
- Problem for KK: if multiplex two independent minimum-phase signals and transmit through fiber, cannot guarantee received signal is minimum-phase.

Example:

$$\begin{pmatrix} E_{r,x} \\ E_{r,y} \end{pmatrix} = \underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}}_{\text{Unitary Jones matrix}} \begin{pmatrix} E_{t,x} \\ E_{t,y} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \boxed{E_{t,x} - E_{t,y}} \\ E_{t,x} + E_{t,y} \end{pmatrix}$$

Carrier cancelled out.
Not minimum-phase.

- Solutions:
  1. Optically demultiplex dual-polarization signals before downconversion.
  2. Add the carrier at the receiver (a "weak LO"), not the transmitter.
- A "weak LO" requires accurate frequency stabilization, much like a conventional LO.

Oversampling rates

Standard coherent detection

- Fundamentally requires $r_{os} \geq 1$.
- Obtain excellent performance using $r_{os} \geq 8/7$ or so.

Kramers-Kronig detection

- Near-ideal performance requires:

  Sampling with $r_{os} = 2$ (squaring).

  Upsampling to $r_{os} = 6$ (nonlinear operations).
- These requirements may be prohibitive for high-speed links.

To learn more

These slides and JLT paper on this study (when paper accepted)

ee.stanford.edu/~jmk/research/smfcom.html#dcs

Early papers on probabilistic shaping in optical communications

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 45, NO. 7, NOVEMBER 1999

Shaping and Nonequiprobable Signaling for Intensity-Modulated Signals

Da-shan Shiu, *Student Member, IEEE*, and Joseph M. Kahn, *Senior Member, IEEE*

ee.stanford.edu/~jmk/pubs/im.shaping.pdf

IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 56, NO. 7, JULY 2008

Lattice Codes for Amplified Direct-Detection Optical Systems

Wei Mao and Joseph M. Kahn, *Fellow, IEEE*

ee.stanford.edu/~jmk/pubs/shaping.amp.dir.det.pdf