Probabilistic Shaping for Direct-Detection Optical Systems

Joseph M. Kahn and Ethan M. Liang

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

Abstract

- We study probabilistic shaping for direct-detection systems that modulate the intensity or Stokes vector and are limited by thermal or amplifier noise, obtaining analytical formulas for the optimal (non-Gaussian) input distributions and corresponding shaping gains.

- To download updated slides and the papers cited:

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

Acknowledgments

We gratefully acknowledge the support of

- Meta (previously Facebook)
- Analog Devices (previously Maxim Integrated)
- Marvell Technology (previously InPhi Corporation)
- LG Electronics
- Hewlett Packard
- University of California MICRO Program
- National Science Foundation
  - Grant DGE-1656518
  - Grant ECS-0335013
  - Grant ECS-9710065

Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Outline

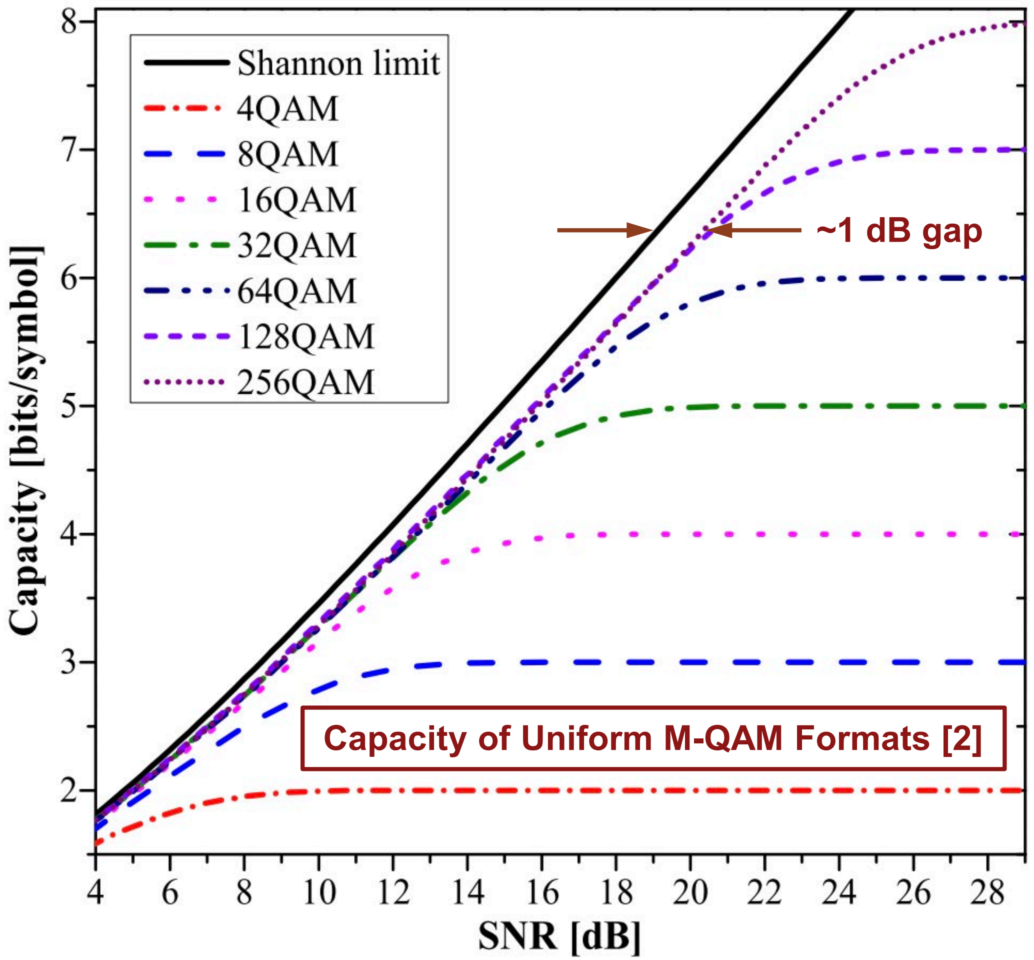
- Introduction
- Standard Coherent Detection
  - Shannon's Method — **1948**
  - Forney's Method — **1989**
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise — **1999**
  - Standard Direct Detection, Amplifier Noise — **2008**
  - Stokes Vector Detection, Thermal or Amplifier Noise — **2023**
  - Kramers Kronig Detection, Amplifier Noise — **2021**
- Discussion

Outline

- Introduction
  - Probabilistic Shaping
  - Assumptions and Performance Metrics
- Standard Coherent Detection
- Direct Detection Methods
- Discussion

Shaping for Coherent AWGN Channels

- Relevant for coherent detection in electrical or optical systems with additive Gaussian noise.
  - Includes amplifier noise or local oscillator shot noise.
- Capacity-achieving input distribution is Gaussian [1] (when modulation order not constrained).
- Uniform signaling does not achieve capacity. Subject to ~1 dB performance gap.
- Probabilistically shaped coded modulation
  - Provides a method to close the performance gap.
  - Enables rate-adaptive transmission without changing modulation order or codeword block length.



1. C. Shannon, Bell System Technical Journal **27.3** (1948).
2. Z. Qu and I. B. Djordjevic, IEEE Access **7** (2019).

Shaping for Direct-Detection Channels

- Would like to rigorously derive optimal input distribution and capacity for various direct-detection channels in optical communications.
- These cannot be computed exactly for several important direct-detection channels:
  - Standard intensity modulation / direct detection (SD) channel.
  - Stokes vector receiver (SVR) channel.
- A method developed by Forney and Wei [1] for coherent AWGN channels has been adapted to direct-detection channels. This allows us to compute the following, at least for high SNR:
  - Analytical input distributions that approach capacity.
  - Achievable gains compared to uniform input distributions.

the main subject of this talk

1. G. Forney and L. Wei, IEEE J. Sel. Areas Commun. **7** (1989).

Local Oscillator-Based vs. Direct Detection

Local Oscillator-Based Detection

- A received signal is mixed with light from a local oscillator laser to
  - Downconvert it to baseband.
  - Effectively amplify the signal.
  - Provide a phase reference.

Direct Detection

- No local oscillator laser is employed at the receiver.
- Detection may be aided by
  - Transmitting an unmodulated carrier with the modulated signal.
  - Mixing the received signal with a (potentially delayed or phase-shifted) copy of itself.

- *Local oscillator-based detection is not necessarily coherent.*
- *Direct detection is not necessarily noncoherent.*

Detection Methods in Optical Communications

LO-Based Detection Method	Direct Detection Method	Detects	Classification	Signal Dimensions (in 2 Polarizations)
Envelope detection	Standard direct detection	Intensity	Noncoherent	1 or 2
Delay-and-multiply detection	Delay interferometer detection	Differential phase shift	Differentially coherent	1 or 2
Polarization shift keying detection	3-D Stokes vector detection	Stokes parameters	Noncoherent + differentially coherent hybrid	3
Polarization shift keying & delay-and-multiply detection	4-D Stokes vector detection	Stokes parameters & differential phase shift	Noncoherent + differentially coherent hybrid	4
Standard coherent detection	Kramers Kronig detection	Field quadratures	Coherent	4

methods considered here

Adapted from:

E. Ip, A. P. T. Lau, D. J. F. Barros and J. M. Kahn, Optics Express **16** (2008).

Assumptions and Performance Metrics in this Tutorial

Constraints

- Average optical power is constrained.
  - Short-reach: eye safety.
  - Long-haul: fiber nonlinearity, amplifier pump power.
- Peak optical power not constrained.
  - All systems have peak power constraints. Fundamental studies start by considering average power constraint.
- Encoding and decoding complexity not constrained.

Impairments

- Consider one dominant additive noise, either:
  - Thermal noise: added to detected photo current.
  - Optical amplifier noise: added to optical electric field.
- Impairments not considered:
  - Inter-symbol interference, fiber nonlinearity, etc.

Performance Metrics

- Spectral efficiency measured in b/symbol, not b/s/Hz.
- Standard intensity modulation / direct detection systems
  - Intensity signals are non-negative.
  - There exist no band-limited, non-negative root-Nyquist pulses [1].
  - Standard direct detection systems cannot be strictly bandlimited while using matched filtering.
  - Rectangular pulses with 5-pole Bessel filters can achieve ~0.8 symbol/s/Hz.
- 3-D Stokes, Kramers Kronig, standard coherent systems
  - Signals can be bipolar.
  - Root-Nyquist pulses can approach 1 symbol/s/Hz.

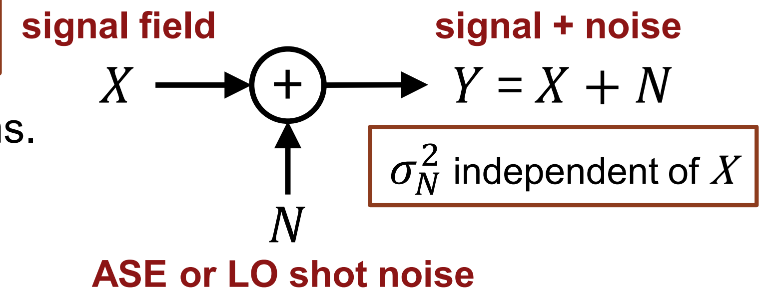
1. S. Hranilovic, IEEE Trans. Commun. **55** (2007).

Outline

- Introduction
- Standard Coherent Detection
  - Shannon’s Method
  - Forney’s Method
  - Implementation
- Direct Detection Methods
- Discussion

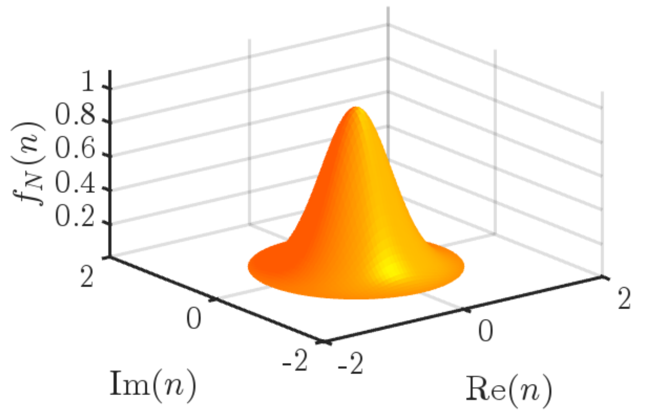
Standard Coherent: Shannon's Method (1/2)

- The method finds an optimal input distribution maximizing the mutual information subject to an average power constraint.
- The standard coherent channel is relevant for coherent electrical or optical systems. Consider one polarization for simplicity.



$$Y = X + N, \qquad X, N, Y \in \mathbb{C}, \qquad X, N \text{ independent}$$

$$N \text{ is Gaussian:} \quad f_N(n) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\|n\|_2^2\right)$$



- Optimization problem

$$\max_{f_X(x)} I(X;Y) = \max_{f_X(x)} h(Y) - h(Y|X)$$

$$= \max_{f_X(x)} h(Y) - h(N)$$

$$\text{s.t.} \qquad \int_{x \in \mathcal{X}} |x|^2 \, f_X(x)\, dx \le \bar{P}$$

C. Shannon, Bell System Technical Journal **27.3** (1948).

Standard Coherent: Shannon's Method (2/2)

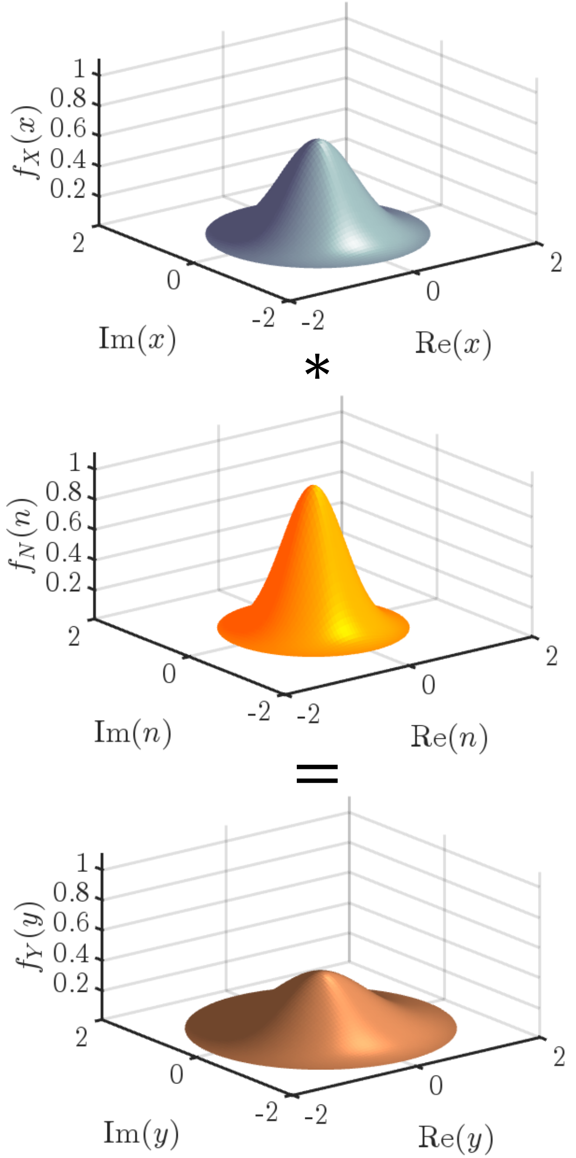
- By variational calculus, the maximum-entropy input distribution subject to an L^2 norm constraint is

$$f_X(x) = \frac{1}{2\pi\bar{P}} \exp\left(-\frac{1}{2\bar{P}} \|x\|_2^2\right) \Leftrightarrow X \sim \mathcal{CN}(0, \bar{P})$$

- When $X \sim \mathcal{CN}(0, \bar{P})$, then $Y \sim \mathcal{CN}(0, \bar{P} + 2\sigma^2)$.
  - $f_Y(y)$ is the maximum entropy distribution subject to an L^2 norm constraint.
  - Therefore, $X \sim \mathcal{CN}(0, \bar{P})$ is capacity achieving.

- Assuming an input distribution is $X \sim \mathcal{CN}(0, \bar{P})$, the capacity (in bits/symbol) is

$$C = h(Y) - h(N) = \log_2\left(\frac{\pi e(\bar{P} + 2\sigma^2)}{\pi e 2\sigma^2}\right) = \log_2\left(1 + \frac{\bar{P}}{2\sigma^2}\right)$$



C. Shannon, Bell System Technical Journal **27.3** (1948).

Outline

- Introduction
- Standard Coherent Detection
  - Shannon's Method
  - Forney's Method
  - Implementation
- Direct Detection Methods
- Discussion

Standard Coherent: Forney's Method

- The method assumes capacity-approaching, high-dimensional lattice codes. It finds an optimal bounding region minimizing the average power of the enclosed signal points while maintaining constant minimum distance.
- Forney's original method is applicable to coherent electrical or optical systems.
- Forney's original work addresses both lattice codes and bounding regions. We focus on the bounding regions.
- Constellation figure of merit for a constellation C (assuming all points in C are equally probable):

$$\mathrm{CFM}(C) \triangleq \frac{d_{\min}^2(C)}{\bar{P}(C)} = \frac{\text{squared minimum Euclidean distance}}{\text{average power}}$$

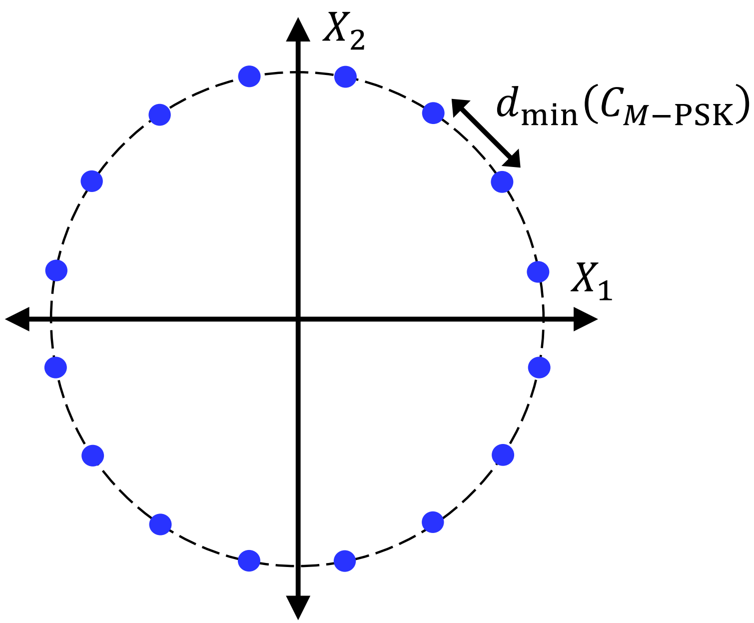
- Variables implicit in the CFM:
  - N: number of dimensions of each signal point $X \in C$
  - $|C|$: total number of messages in C

N and $|C|$ must be equated when comparing CFMs of different constellations.

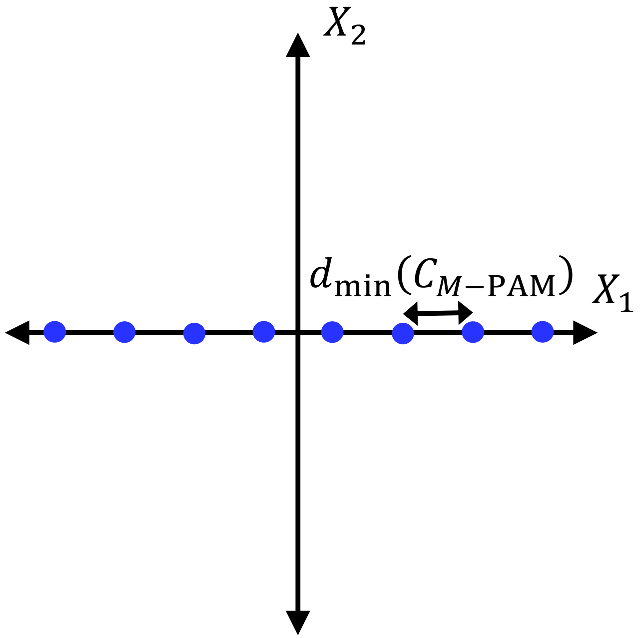
G. Forney and L. Wei, IEEE J. Sel. Areas Commun. **7** (1989).

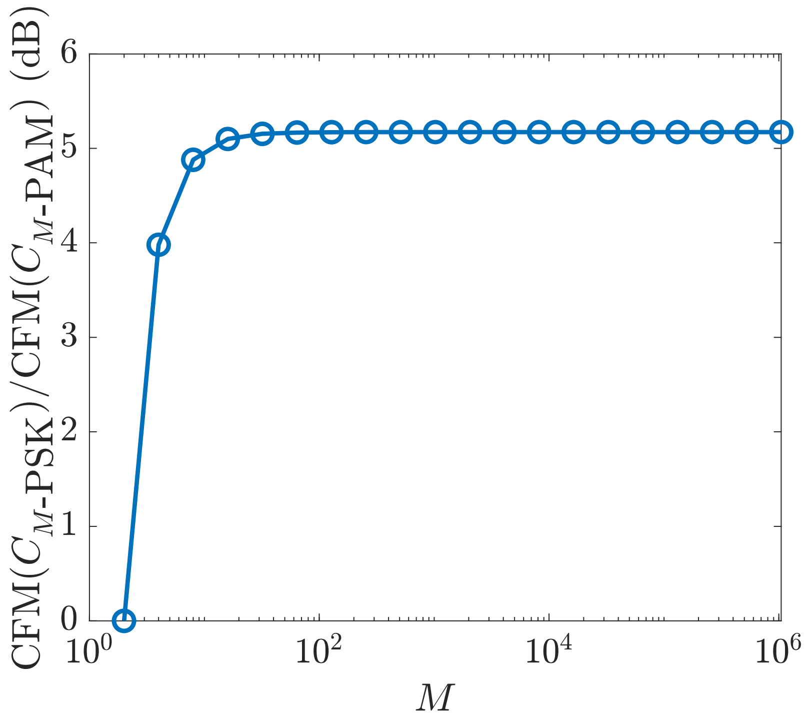
Constellation Figure of Merit Example

$C_{M-\mathrm{PSK}}$: $M = 16$



$C_{M-\mathrm{PAM}}$: $M = 8$





$$\mathrm{CFM}(C_{M-\mathrm{PSK}}) = \frac{d_{\min}^2(C_{M-\mathrm{PSK}})}{\bar{P}(C_{M-\mathrm{PSK}})} = 4\sin^2\left(\frac{\pi}{M}\right)$$

$$\mathrm{CFM}(C_{M-\mathrm{PAM}}) = \frac{d_{\min}^2(C_{M-\mathrm{PAM}})}{\bar{P}(C_{M-\mathrm{PAM}})} = \frac{12}{M^2-1}$$

$$\lim_{M\to\infty} \frac{\mathrm{CFM}(C_{M-\mathrm{PSK}})}{\mathrm{CFM}(C_{M-\mathrm{PAM}})} = \frac{\pi^2}{3} \approx 5.17 \text{ dB}$$

Lattice

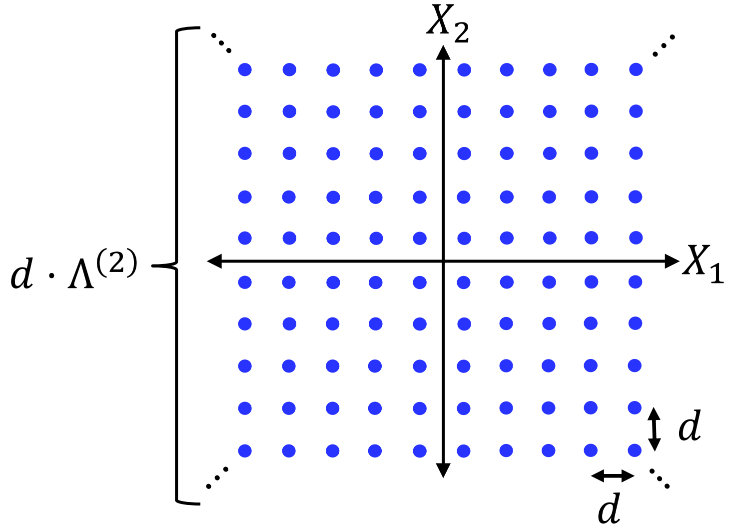
- All lattices are assumed to be symmetric about the origin.

Real Lattice

- N-D lattice $\Lambda^{(\mathrm{N})}$ is N-fold Cartesian product of 1-D lattice $\Lambda^{(1)}$:

$$\Lambda^{(1)} \triangleq \left\{\dots,-\frac{5}{2},-\frac{3}{2},-\frac{1}{2},\frac{1}{2},\frac{3}{2},\frac{5}{2},\dots\right\}$$

- The lattice $d\cdot\Lambda^{(\mathrm{N})}$ has a minimum Euclidean distance d.

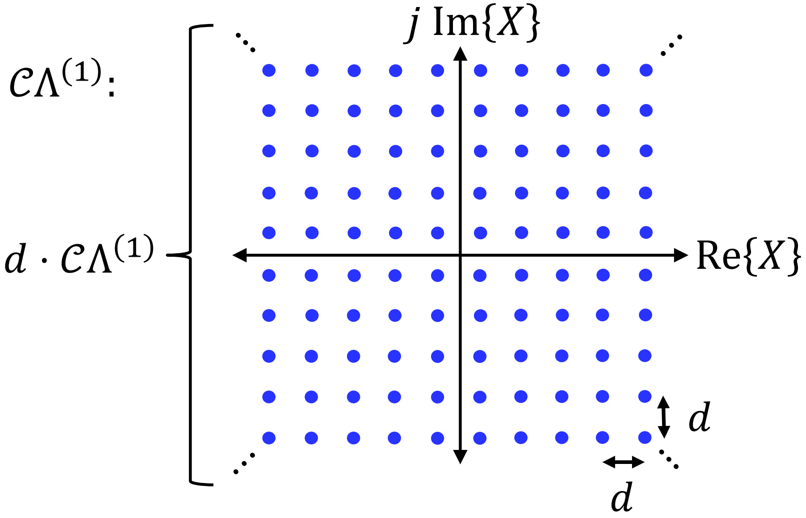


Complex Lattice

- N-D complex lattice $\mathcal{C}\Lambda^{(\mathrm{N})}$ is N-fold Cartesian product of 1-D complex lattice $\mathcal{C}\Lambda^{(1)}$:

$$\mathcal{C}\Lambda^{(1)} \triangleq \left\{\begin{matrix} \vdots \\ \dots,-\frac{3}{2}-j\frac{1}{2},-\frac{1}{2}-j\frac{1}{2},\frac{1}{2}-j\frac{1}{2},\frac{3}{2}-j\frac{1}{2},\dots \\ \dots,-\frac{3}{2}+j\frac{1}{2},-\frac{1}{2}+j\frac{1}{2},\frac{1}{2}+j\frac{1}{2},\frac{3}{2}+j\frac{1}{2},\dots \\ \vdots \end{matrix}\right\}$$

- The lattice $d\cdot\mathcal{C}\Lambda^{(\mathrm{N})}$ has a minimum Euclidean distance d.



Bounding Region

Reference Bounding Region

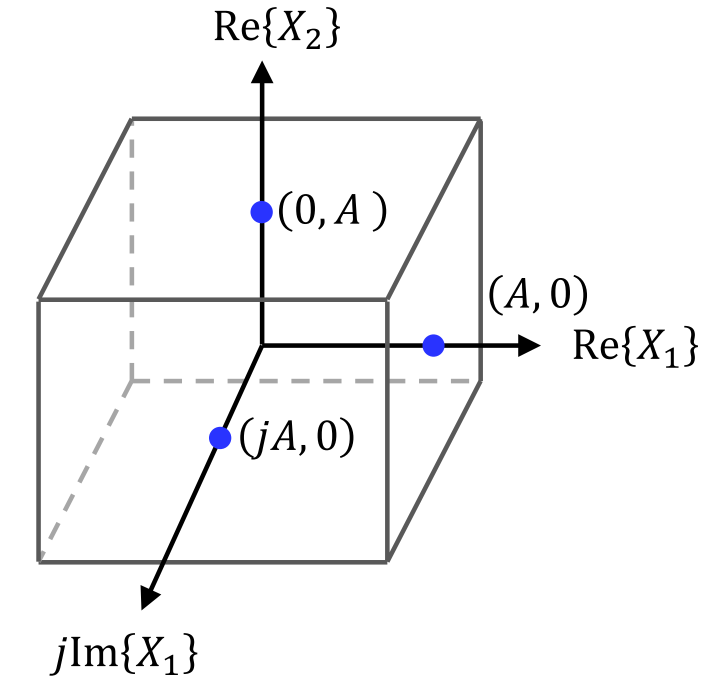
- The reference bounding region is an N-D hypercube:

$$\mathcal{R}_{\mathrm{ref}}(N,A)=\left\{X\in\mathbb{C}^N \,\middle|\, \max_{i\in\{1,2,\ldots,N\}}\{|\mathrm{Re}\{X_i\}|,|\mathrm{Im}\{X_i\}|\}=A\right\}$$

- The set of points enclosed by $\mathcal{R}_{\mathrm{ref}}(N,A)$ is:

$$\mathbb{R}_{\mathrm{ref}}(N,A)=\left\{X=(x_1,x_2,\ldots,x_N)\,\middle|\,\|X\|_\infty\le A\right\}$$

- The volume of the region enclosed by $\mathcal{R}_{\mathrm{ref}}(N,A)$ is $\mathcal{V}_{\mathrm{ref}}(N,A)$.



Optimal Bounding Region

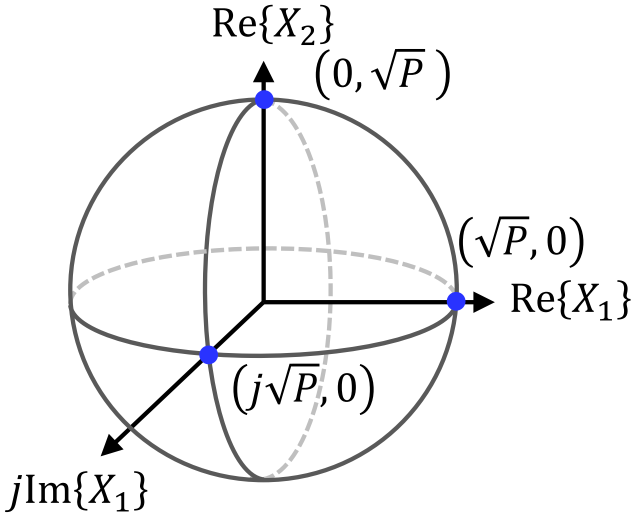
- The optimal bounding region minimizes the average physical energy of the enclosed signal points subject to a constraint on the number of enclosed points.
- The optimal bounding region for maximum physical energy P is:

$$\mathcal{R}_{\mathrm{shaped}}(N,P)=\left\{X\in\mathbb{C}^N\,\middle|\,\|X\|_2^2=P\right\}$$

- The set of points enclosed by the bounding region $\mathcal{R}_{\mathrm{shaped}}(N,P)$ is:

$$\mathbb{R}_{\mathrm{shaped}}(N,P)=\left\{X\in\mathbb{C}^N\,\middle|\,\|X\|_2^2\le P\right\}$$

- The volume of the region enclosed by $\mathcal{R}_{\mathrm{shaped}}(N,P)$ is $\mathcal{V}_{\mathrm{shaped}}(N,P)$.



Lattice Code Construction

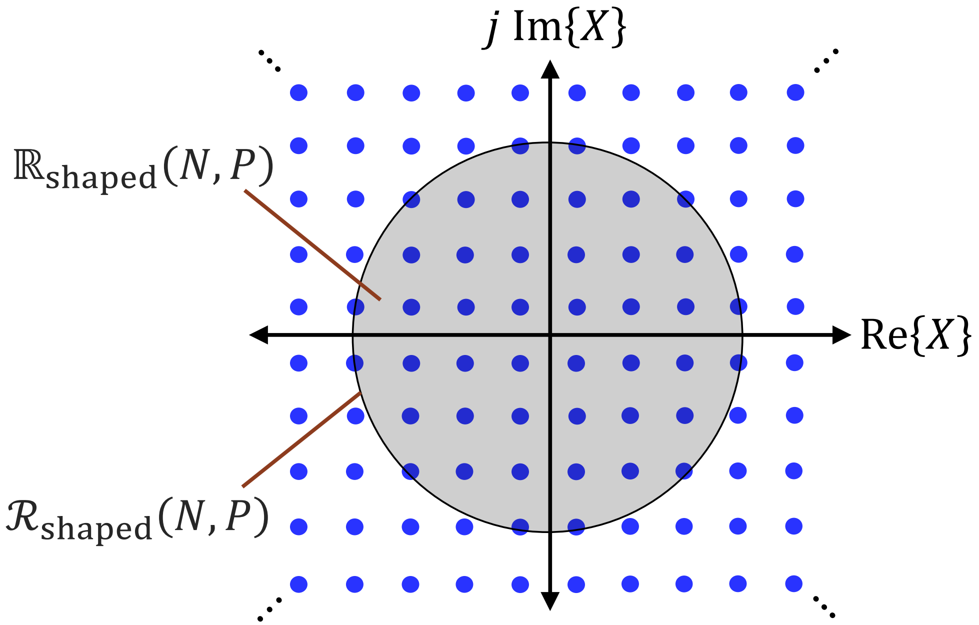
- Consider $d \cdot \Lambda^{(\mathrm{N})}$, which is an N-D lattice with minimum Euclidean distance d.
- C_{shaped} and C_{ref} are obtained as:

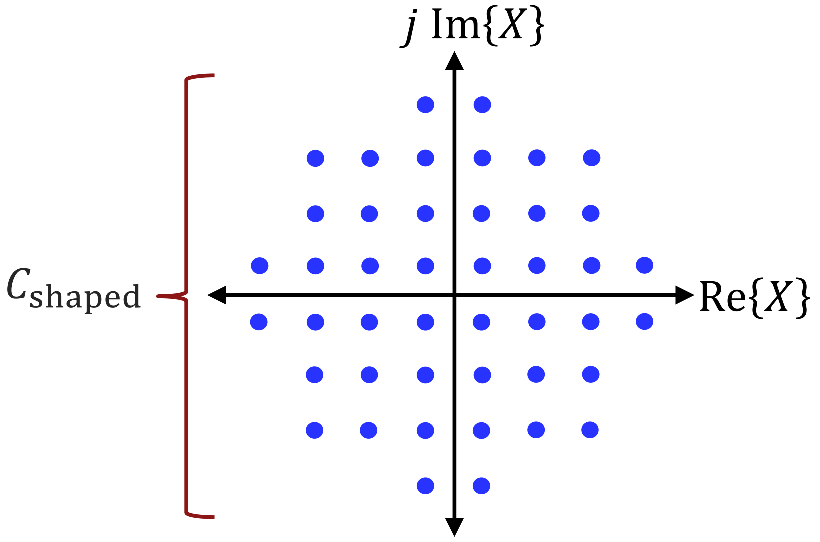
$$C_{\text{shaped}} = \mathbb{R}_{\text{shaped}}(N, P) \cap d \cdot \mathcal{C}\Lambda^{(\mathrm{N})}$$

$$C_{\text{ref}} = \mathbb{R}_{\text{ref}}(N, A) \cap d \cdot \mathcal{C}\Lambda^{(\mathrm{N})}$$

- The intersection selects signal points from $d \cdot \Lambda^{(\mathrm{N})}$ inside $\mathbb{R}_{\text{shaped}}(N, P)$ and $\mathbb{R}_{\text{ref}}(N, A)$ to construct C_{shaped} and C_{ref}, respectively.
- The minimum distances of C_{shaped} and C_{ref} are equal:

$$d_{\min}(C_{\text{shaped}}) = d_{\min}(C_{\text{ref}})$$





Continuous Approximation

- The continuous approximation replaces the lattice by a continuum. This becomes increasing accurate as $|C_{\text{shaped}}|$ or $|C_{\text{ref}}| \to \infty$.
- $|C_{\text{shaped}}|$ and $|C_{\text{ref}}|$ are replaced by $\mathcal{V}_{\text{shaped}}(N,P)$ and $\mathcal{V}_{\text{ref}}(N,A)$.
- The average power expressions $\bar{P}(C)$ are expressed as integrals:

$$\bar{P}(C_{\text{shaped}}) = \frac{1}{|C_{\text{shaped}}|} \sum_{X \in C_{\text{shaped}}} \|X\|_2^2 \cong \frac{1}{\mathcal{V}_{\text{shaped}}(N,P)} \int_{X \in \mathbb{R}_{\text{shaped}}(N,P)} \|X\|_2^2 dX$$

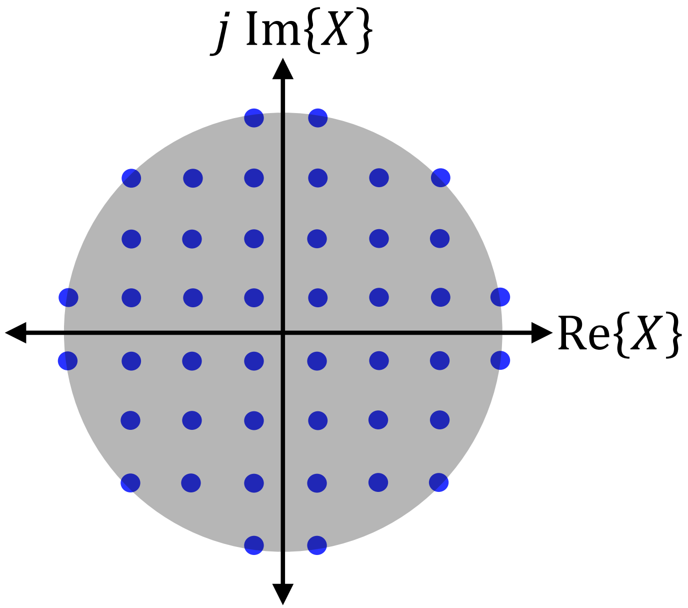
$$\bar{P}(C_{\text{ref}}) = \frac{1}{|C_{\text{ref}}|} \sum_{X \in C_{\text{ref}}} \|X\|_2^2 \cong \frac{1}{\mathcal{V}_{\text{ref}}(N,A)} \int_{X \in \mathbb{R}_{\text{ref}}(N,A)} \|X\|_2^2 dX$$

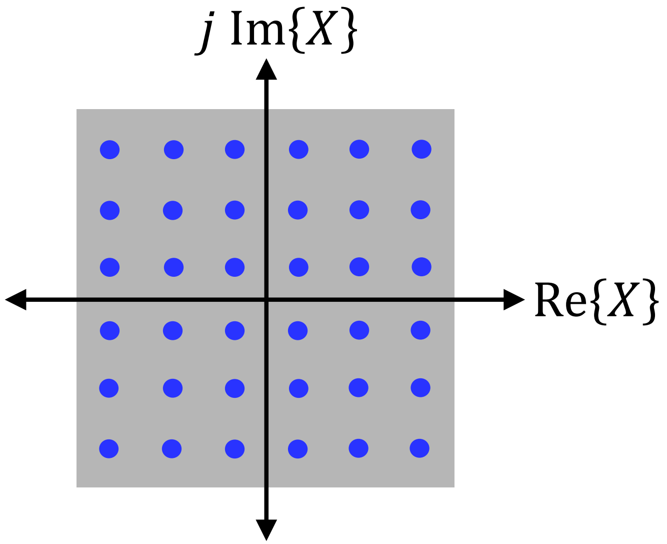
- Enclosed volumes:

$$\mathcal{V}_{\text{shaped}}^{(\text{SC})}(N,P) = \frac{(\pi P)^N}{N!}, \qquad \mathcal{V}_{\text{ref}}^{(\text{SC})}(N,A) = (4A^2)^N$$

- Average powers:

$$\bar{P}_{\text{shaped}}^{(\text{SC})}(N,P) = \frac{P}{N+1}, \qquad \bar{P}_{\text{ref}}^{(\text{SC})}(A) = \frac{2}{3}A^2$$





Induced Marginal Distribution (1/3)

- The shaping region $\mathbb{R}_{\text{shaped}}(N,P)$ can be partitioned into disjoint subsets by conditioning on a value in a basic dimension:

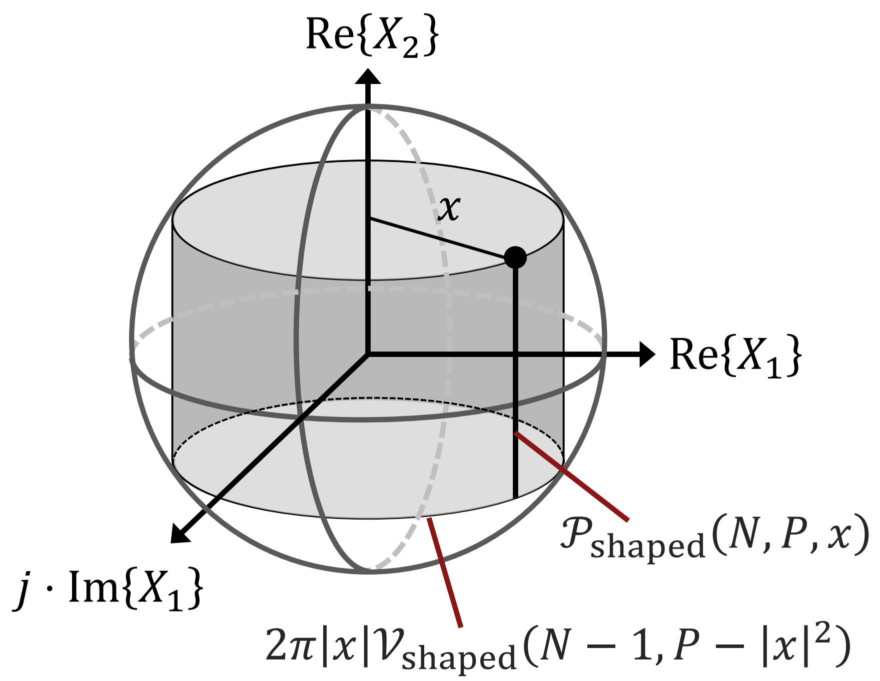
$$\mathcal{P}_{\text{shaped}}(N,P,x) \triangleq \left\{ \boldsymbol{X} \in \mathbb{R}_{\text{shaped}}(N,P) \,\middle|\, X_1 = x \right\}$$

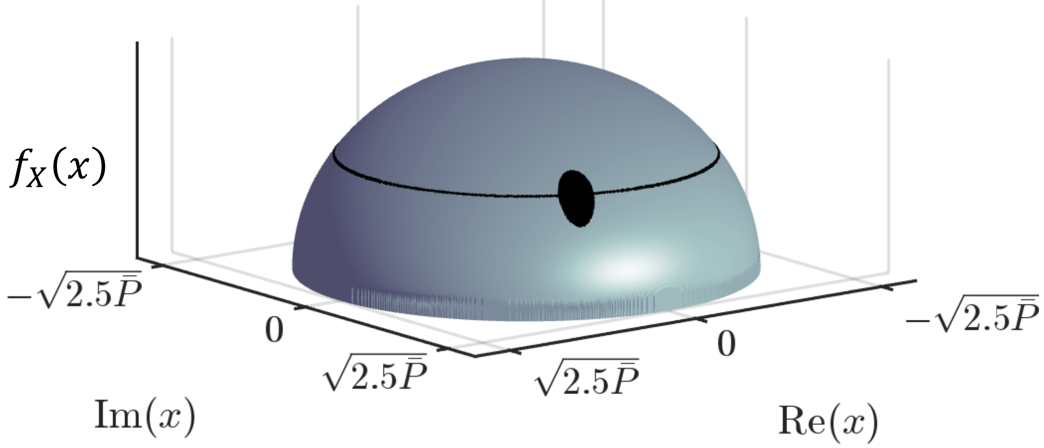
- The volume of $\mathcal{P}_{\text{shaped}}(N,P,x)$ is

$$\mathcal{V}_{\text{shaped}}(N-1, P-|x|^2)$$

- The induced probability density is proportional to the volume of the partition in $N-1$ dimensions:

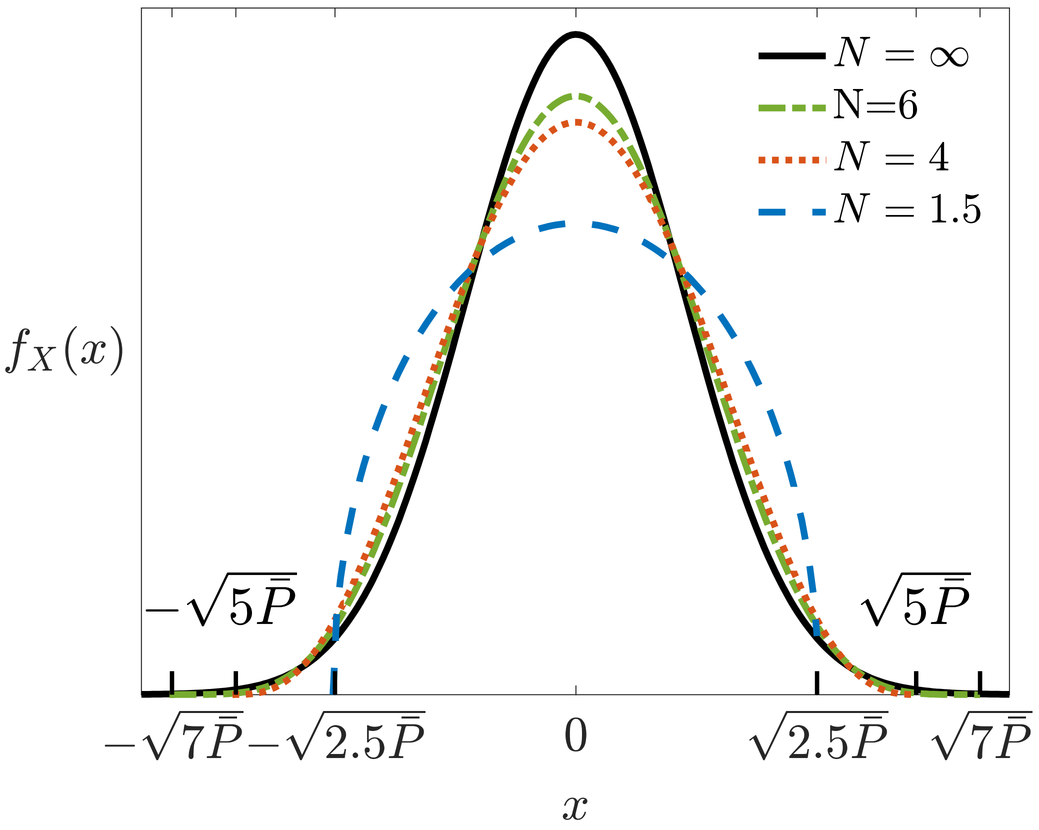
$$f_X(x) \propto \mathcal{V}_{\text{shaped}}(N-1, P-|x|^2)$$

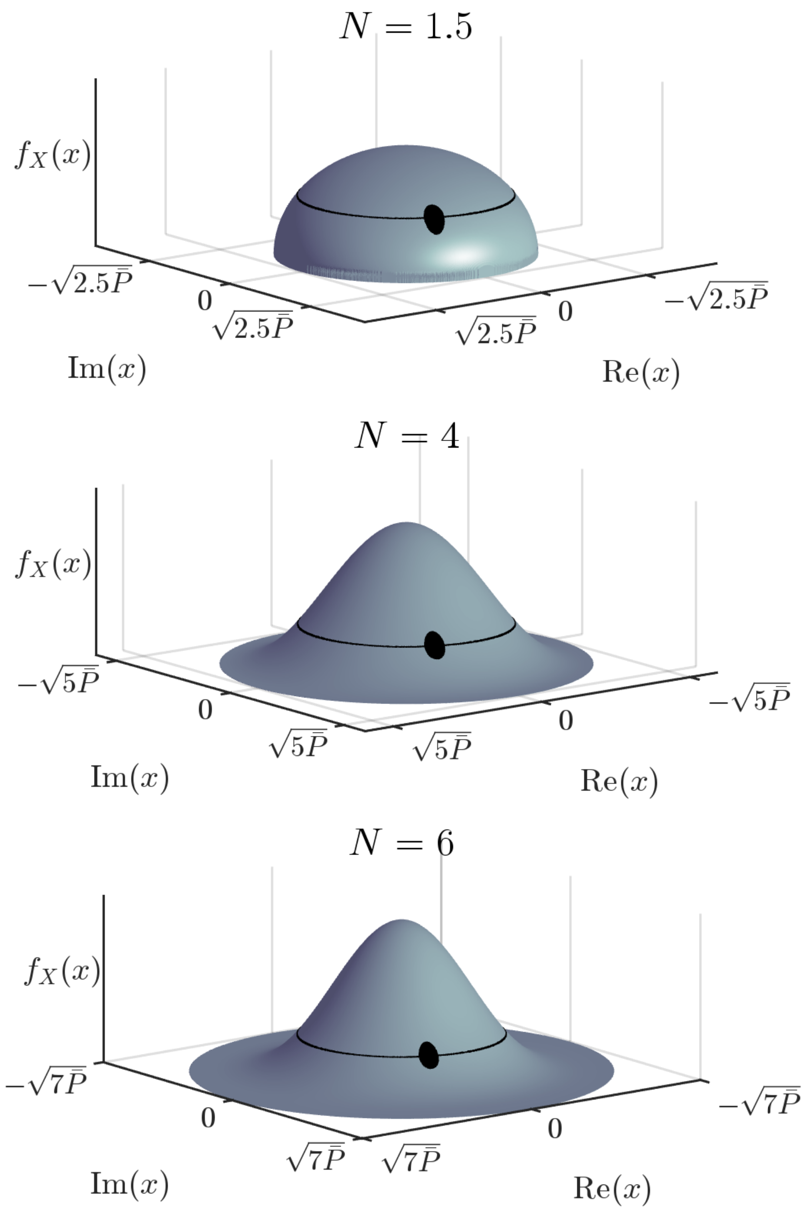




Induced Marginal Distribution (2/3)

Induced distribution in one real dimension for various N





- Peak-to-average power ratio:

$$P = (N+1)\bar{P}^{(\mathrm{SC})}_{\mathrm{shaped}}(N,P)$$

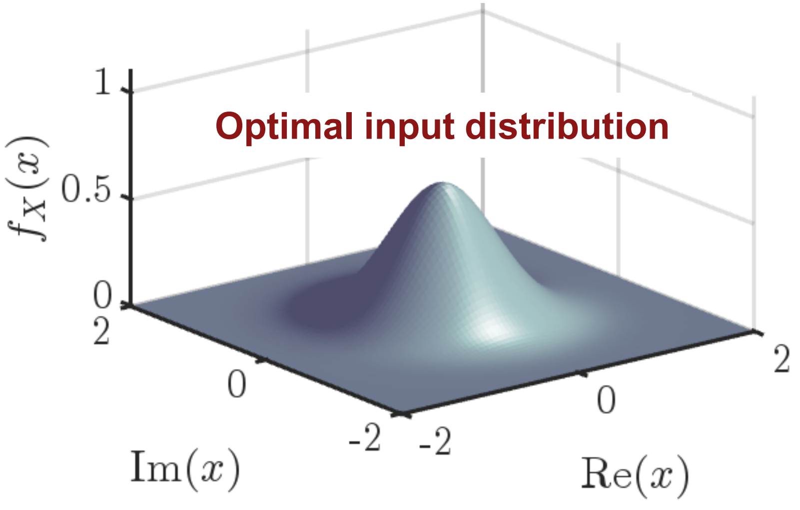
Induced Marginal Distribution (3/3)

Optimal Induced Distribution

$$f_X(x) \propto \mathcal{V}_{\text{shaped}}^{(\text{SC})}(N-1, P-|x|^2) = \left(1 - \frac{|x|^2}{(N+1)\bar{P}_{\text{shaped}}^{(\text{SC})}(N,P)}\right)^{N-1}$$

$$\Rightarrow \lim_{N\to\infty} f_X(x) \propto \exp(-|x|^2/\beta)$$

- The optimal input distribution is Gaussian in complex electric field.

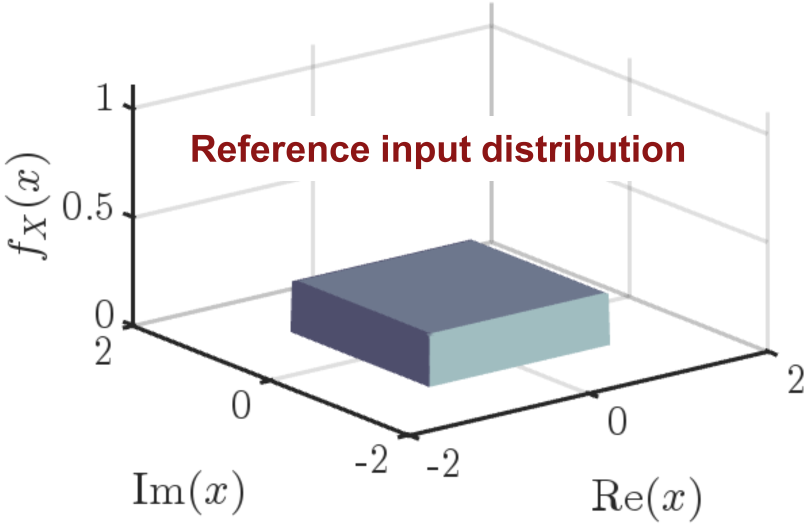


Shape Gain

- Equating volumes $\mathcal{V}_{\text{shaped}}^{(\text{SC})}(N,P) = \mathcal{V}_{\text{ref}}^{(\text{SC})}(N,P)$, the shape gain is

$$\gamma_s = \lim_{N\to\infty} \frac{\bar{P}_{\text{ref}}^{(\text{SC})}(A)}{\bar{P}_{\text{shaped}}^{(\text{SC})}(N,P)} = \frac{\pi e}{6} \approx 1.53 \text{ dB}$$

- The Gaussian distribution improves energy efficiency by ~1.5 dB.



Outline

- Introduction
- Standard Coherent Detection
  - Shannon's Method
  - Forney's Method
  - Implementation
- Direct Detection Methods
- Discussion

Standard Coherent: Implementation (1/2)

Discrete Input Distributions

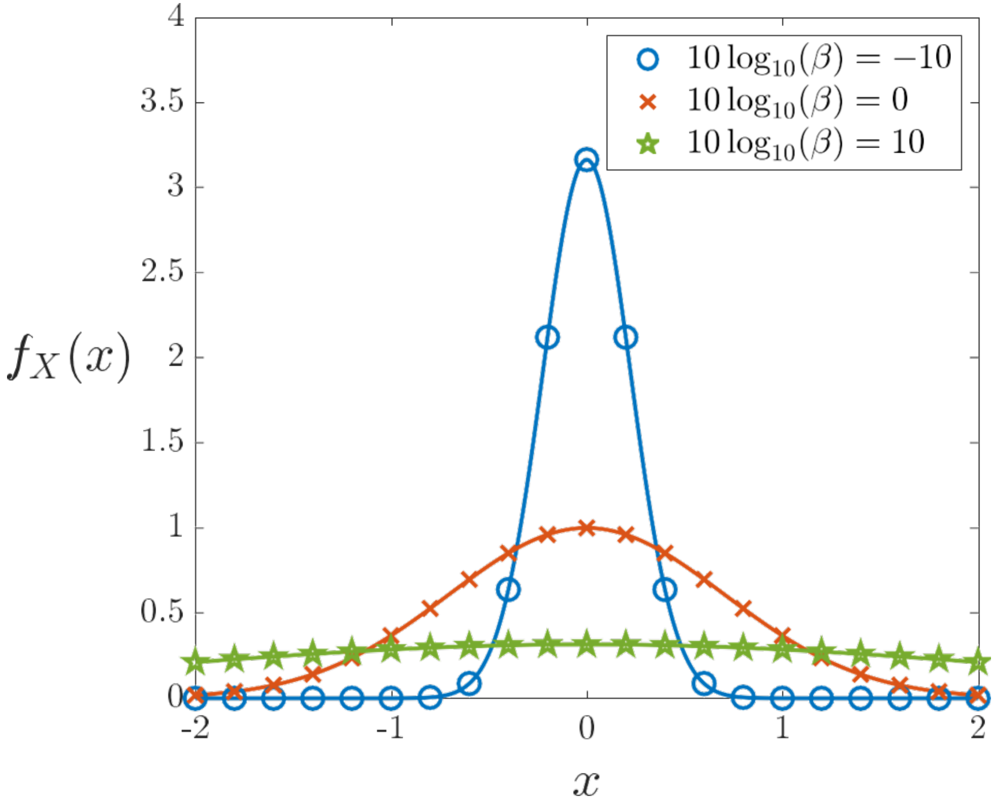
- Discrete constellations must be used in practice.
- Can numerically compute the optimal discrete distribution using Blahut-Arimoto algorithm [1], [2].
- Can employ Maxwell-Boltzmann distribution [3] for discrete x:

$$P_X(x) \triangleq \exp(-\|x\|^2/\beta)\,/Z(\beta), \qquad Z(\beta) \triangleq \sum_{x\in C} \exp(-\|x\|^2/\beta)$$

- Parameterized by a shaping parameter $\beta \geq 0$:
  - β controls the tradeoff between energy efficiency and bit rate.
  - $\beta \to \infty$ as $\mathrm{SNR} \to \infty$ (approaches a uniform distribution).
- MB distribution usually close to optimal at high SNR.
- N-D MB distribution is separable:

$$P_{\mathbf{X}}(\mathbf{x}) = \frac{1}{Z(\beta)} \exp(-\|\mathbf{x}\|^2/\beta) = \frac{1}{Z(\beta)} \prod_{i=1}^{N} \exp\left(-x_i^2/\beta\right)$$

21-point MB distribution for three values of β



1. S. Arimoto, IEEE Trans. Inf. Theory **18** (1972).
2. R. Blahut, IEEE Trans. Inf. Theory **18** (1972).
3. F. Kschischang and S. Pasupathy, IEEE Trans. Inf. Theory **9** (1993).

Standard Coherent: Implementation (2/2)

Probabilistic Amplitude Shaping

- PAS architecture combines a distribution matcher (DM) with forward error correction (FEC).
- Advantages
  - Capacity achieving and can be integrated with existing FEC schemes.
  - Enables rate adaptation without changing the FEC block length.
- Drawbacks
  - Requires a probability distribution that can be partitioned into pairs of points with equal probabilities.

Distribution Matcher

- Constant composition distribution matcher [1].
  - Maps to fixed-composition sequences.
- Shell shaping [2].
  - Maps to sequences with a maximum physical energy.
  - Can have superior performance when the bit sequence is short.
  - Can have high complexity, especially for long block lengths.

1. G. Böcherer, *et al.*, IEEE Trans. Commun. **63** (2015).

2. A. Amari, *et al.*, J. Lightw. Technol. **37** (2019).

Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise
  - Standard Direct Detection, Amplifier Noise
  - Stokes Vector Detection, Thermal or Amplifier Noise
  - Kramers Kronig Detection, Amplifier Noise
- Discussion

Transition to Direct Detection

Shannon's Method

- Optimization problems can be stated, but resulting Lagrangians cannot be solved analytically.

Forney's Method

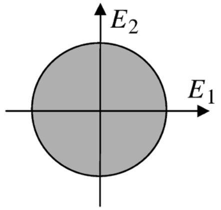
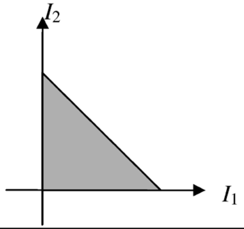
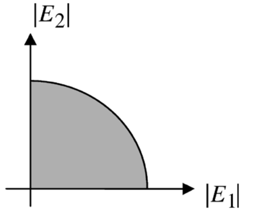
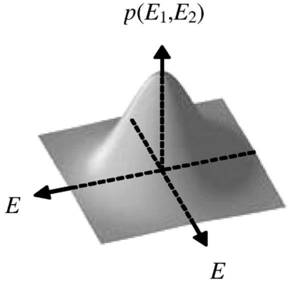
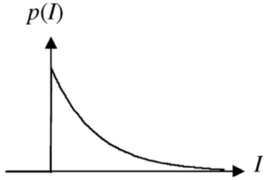
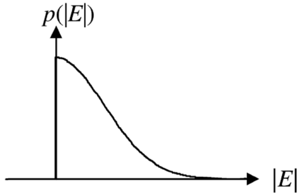
- Forney's original method for standard coherent detection used a uniform lattice in electric field coordinates.
- The binary error probability between any two N-D signal points depends only on the Euclidean distance between them.

Natural Coordinates

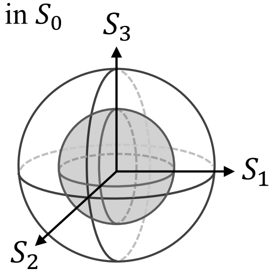
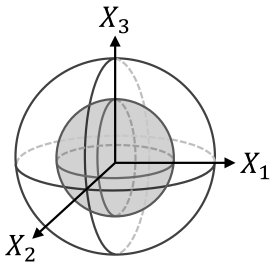
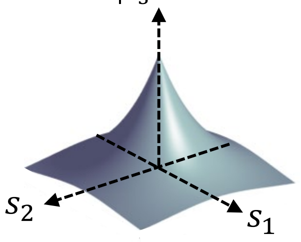
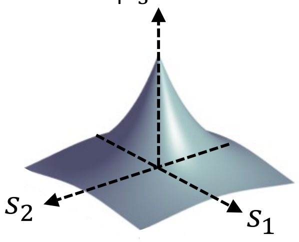
- Enable extension of Forney's method to direct-detection optical channels [1], [2].
- Define the uniform lattice in natural coordinates, in which the binary error probability is a decreasing function of Euclidean distance, at least asymptotically at high SNR.
- The natural coordinates depend on
  - Detection method.
  - Noise distribution.
  - Decision rule.

1. D. Shiu and J. M. Kahn, IEEE Trans. Inf. Theory **45** (1999).
2. W. Mao and J. M. Kahn, IEEE Trans. Commun. **56** (2008).

Shaping Results Summary (1/2)

Detection Method & Dominant Noise	Coherent Detection, Amplifier or LO Shot Noise	Noncoherent Detection, Thermal Noise	Noncoherent Detection, Amplifier or LO Shot Noise
Natural Coordinates	2-D electric field $E = \{E_1, E_2\}$	1-D electric field intensity $I = \lvert E\rvert^2$	1-D electric field magnitude $\lvert E\rvert$
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	$p(E) = \frac{1}{\pi P}\exp\left(-\frac{\lvert E\rvert^2}{P}\right)$, $E = (E_1, E_2)$ 	$p(I) = \frac{1}{P}\exp\left(-\frac{I}{P}\right)$, $I \geq 0$ 	$p(\lvert E\rvert) = \sqrt{\frac{2}{\pi P}}\exp\left(-\frac{\lvert E\rvert^2}{2P}\right)$, $\lvert E\rvert \geq 0$ 
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

Shaping Results Summary (2/2)

Detection Method & Dominant Noise	Stokes Vector Receiver, Thermal Noise	Stokes Vector Receiver, Amplifier Noise
Natural Coordinates	3-D Stokes vector $\mathbf{S} = \{S_1, S_2, S_3\}$	3-D scaled Stokes vector $\mathbf{X} = \{X_1, X_2, X_3\} = \{S_1, S_2, S_3\}/\sqrt{S_0}$
Optimal Shaping Region	Non-negative orthant bounded by N-simplex in S_0 	N-sphere centered at the origin 
Induced Optimal Signaling Distribution in Constituent Constellation	$p(\mathbf{S}) \propto \exp(-S_0/\beta)$ $f_{\mathbf{S}\|S_3}(\mathbf{s}\|0)$ 	$p(\mathbf{S}) \propto \exp(-S_0/\beta)$ $f_{\mathbf{S}\|S_3}(\mathbf{s}\|0)$ 
Ultimate Shape Gain	$\sim 0.961\sqrt[3]{\pi}e/3 \approx 1.056$ dB	$\pi e/6 \approx 1.53$ dB
Optimal Distribution	3-D exponential	3-D exponential
Key Works	H. Jia, E. Liang and J. M. Kahn, (2023).	H. Jia, E. Liang and J. M. Kahn, (2023).

Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise
  - Standard Direct Detection, Amplifier Noise
  - Stokes Vector Detection, Thermal or Amplifier Noise
  - Kramers Kronig Detection, Amplifier Noise
- Discussion

Standard Direct Detection, Thermal Noise (1/3)

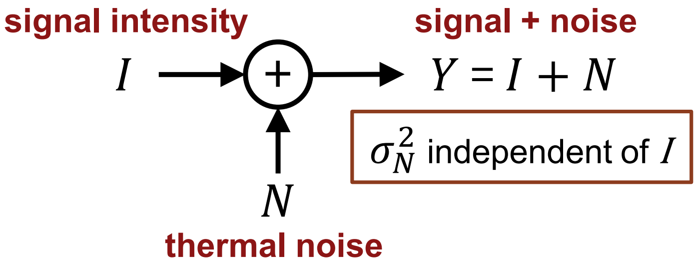
Determine natural coordinates

$$Y = I + N, \qquad I \in \mathbb{R}_+, \qquad Y, N \in \mathbb{R}, \qquad N \sim \mathcal{N}(0, \sigma_{\text{th}}^2)$$

- For intensities $I_1, I_2 \in \mathbb{R}_+$, binary error probability for an ML detector is

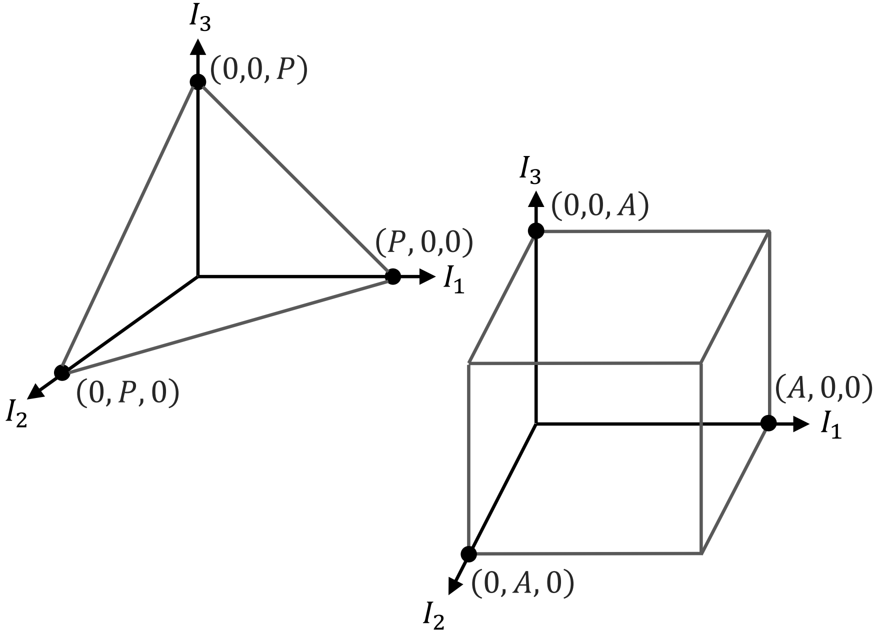
$$P_e = Q\left(\frac{\|I_1 - I_2\|_2}{2\sigma_{\text{th}}}\right).$$

- The natural coordinates are electric field intensities.



Find optimal and reference bounding regions

$$\mathcal{R}_{\text{shaped}}^{(\text{SD,th})}(N, P) = \{\mathbf{I} \in \mathbb{R}_+^N \mid \|\mathbf{I}\|_1 = P\}, \qquad \mathcal{R}_{\text{ref}}^{(\text{SD,th})}(N, P) = \{\mathbf{I} \in \mathbb{R}_+^N \mid \|\mathbf{I}\|_\infty = A\}$$



Compute enclosed volumes

$$\mathcal{V}_{\text{shaped}}^{(\text{SD,th})}(N, P) = \frac{P^N}{N!}, \qquad \mathcal{V}_{\text{ref}}^{(\text{SD,th})}(N, A) = A^N$$

Compute average powers

$$\bar{P}_{\text{shaped}}^{(\text{SD,th})}(N, P) = \frac{P}{N+1}, \qquad \bar{P}_{\text{ref}}^{(\text{SD,th})}(A) = \frac{A}{2}$$

D. Shiu and J. M. Kahn, IEEE Trans. Inf. Theory **45** (1999).

Standard Direct Detection, Thermal Noise (2/3)

Compute shape gain (for equal volumes)

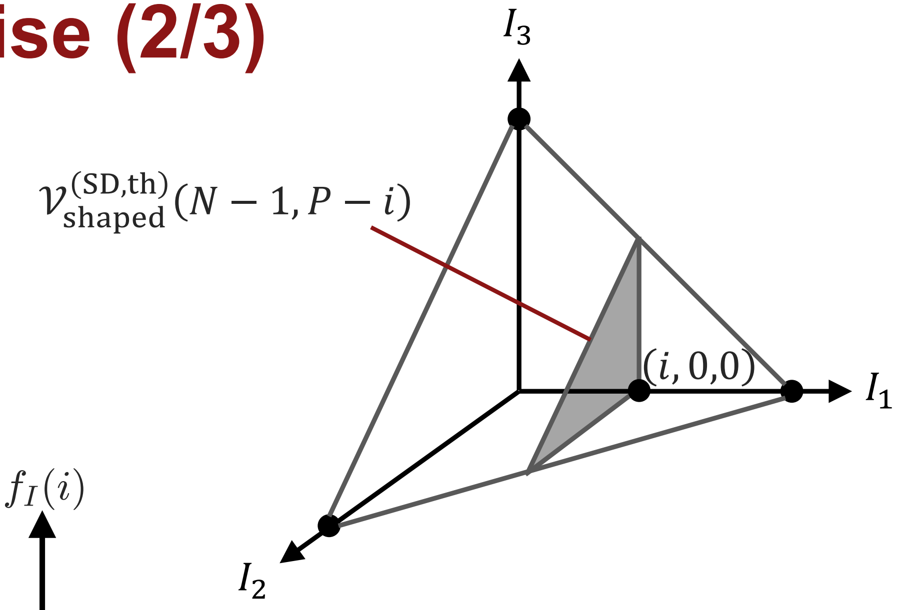
$$\gamma_s = \lim_{N\to\infty} \frac{\bar{P}_{\text{ref}}^{(\text{SD,th})}(A)}{\bar{P}_{\text{shaped}}^{(\text{SD,th})}(N,P)} = \frac{e}{2} \approx 1.33\ \text{dB}$$

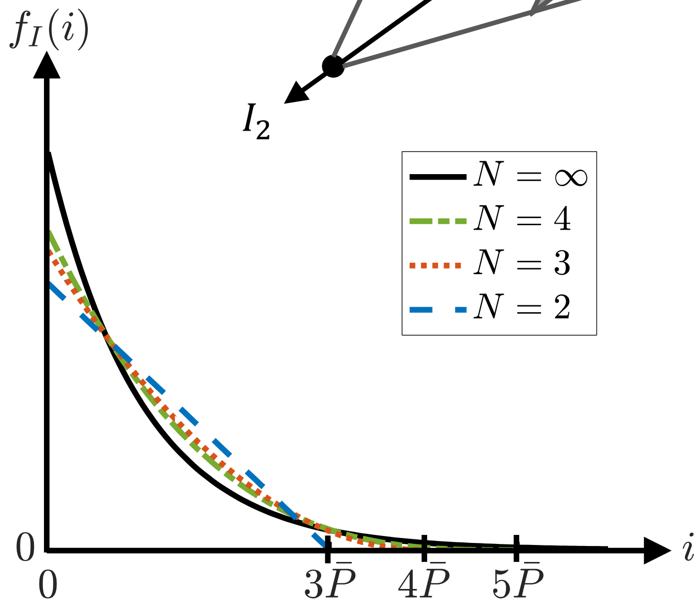
Compute induced distribution in one basic dimension

$$f_I(i) \propto \mathcal{V}_{\text{shaped}}^{(\text{SD,th})}(N-1, P-i)$$

$$\Rightarrow \lim_{N\to\infty} f_I(i) \propto \exp(-i/\beta), \qquad i \geq 0$$

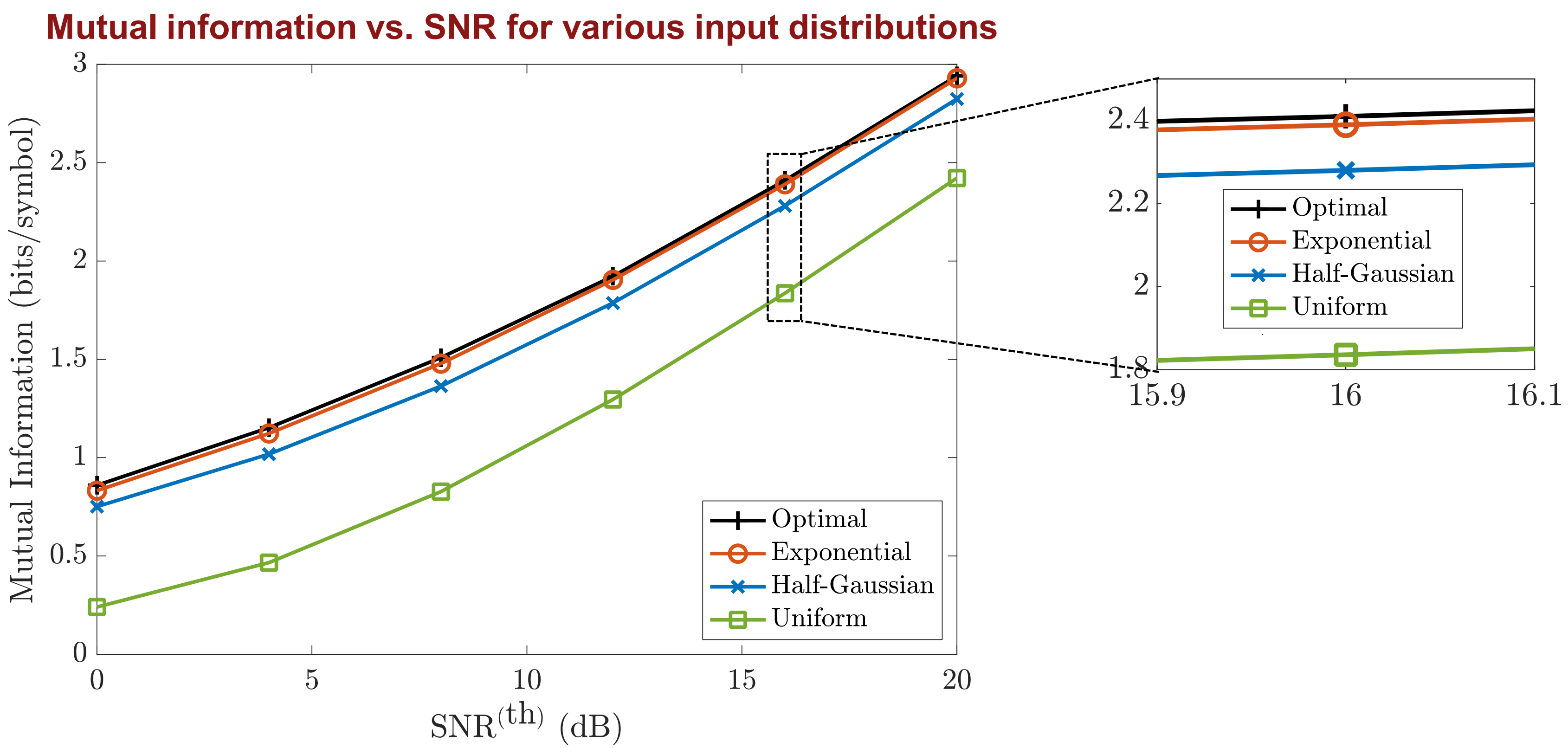
- Optimal input distribution is an exponential in electric field intensity.





D. Shiu and J. M. Kahn, IEEE Trans. Inf. Theory **45** (1999).

Standard Direct Detection, Thermal Noise (3/3)



- Optimal distribution computed using Blahut-Arimoto algorithm with many intensity levels.
- Exponential outperforms half-Gaussian.
- Optical power gains (half the $\mathrm{SNR}^{(\mathrm{th})}$ gains) substantially exceed $\gamma_s = 1.33$ dB at finite $\mathrm{SNR}^{(\mathrm{th})}$.

D. Shiu and J. M. Kahn,
IEEE Trans. Inf. Theory **45** (1999).

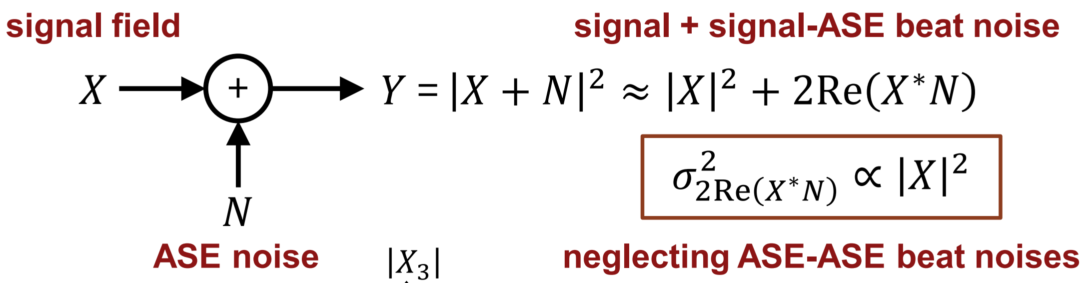
Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise
  - Standard Direct Detection, Amplifier Noise
  - Stokes Vector Detection, Thermal or Amplifier Noise
  - Kramers Kronig Detection, Amplifier Noise
- Discussion

Standard Direct Detection, Amplifier Noise (1/3)

Determine natural coordinates

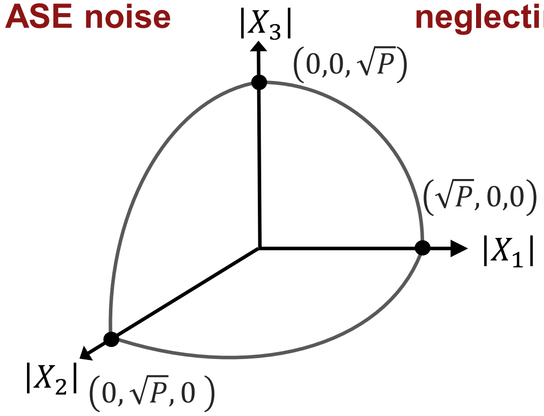
$$Y = |X+N|^2, \qquad X, N \in \mathbb{C}, \qquad Y \in \mathbb{R}_+, \qquad N \sim \mathcal{CN}\left(0, 2\sigma_{\mathrm{amp}}^2\right)$$

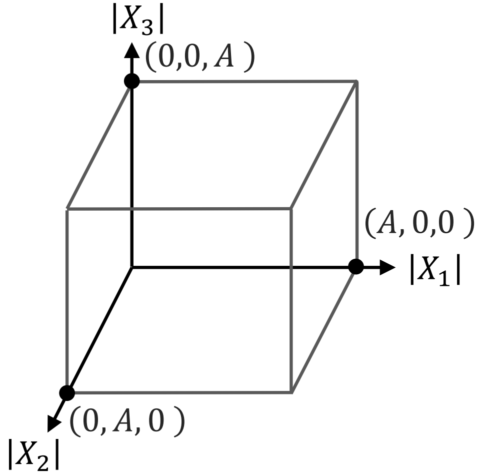


- Approximate decision threshold for field magnitudes $|X_1|, |X_2|$:

$$Y_D \approx \frac{|X_1|^2|X_2| + |X_1||X_2|^2}{|X_1| + |X_2|} = |X_1||X_2|$$

- Using Y_D, the approximate binary error probability is $P_e \approx Q\left(\frac{\| |X_1| - |X_2| \|_2}{2\sigma_{\mathrm{amp}}}\right)$.
- The natural coordinates are electric field magnitudes.
- Can assume $X = |X| \in \mathbb{R}_+$ without loss of generality.





Find optimal and reference bounding regions

$$\mathcal{R}^{(\mathrm{SD,amp})}_{\mathrm{shaped}}(N,P) = \{\mathbf{X} \in \mathbb{R}^N_+ \mid \|\mathbf{X}\|_2^2 = P\}, \qquad \mathcal{R}^{(\mathrm{SD,amp})}_{\mathrm{ref}}(N,P) = \{\mathbf{X} \in \mathbb{R}^N_+ \mid \|\mathbf{X}\|_\infty = A\}$$

Compute enclosed volumes

$$\mathcal{V}^{(\mathrm{SD,amp})}_{\mathrm{shaped}}(N,P) = \frac{1}{\Gamma\left(\frac{N}{2}+1\right)} \frac{(\pi P)^{N/2}}{2^N}, \qquad \mathcal{V}^{(\mathrm{SD,amp})}_{\mathrm{ref}}(N,A) = A^N$$

W. Mao and J. M. Kahn, IEEE Trans. Commun. **56** (2008).

Standard Direct Detection, Amplifier Noise (2/3)

Compute average powers

$$\bar{P}_{\text{shaped}}^{(\text{SD,th})}(N,P) = \frac{P}{N+2}, \qquad \bar{P}_{\text{ref}}^{(\text{SD,th})}(A) = \frac{A^2}{3}$$

Compute shape gain (for equal volumes)

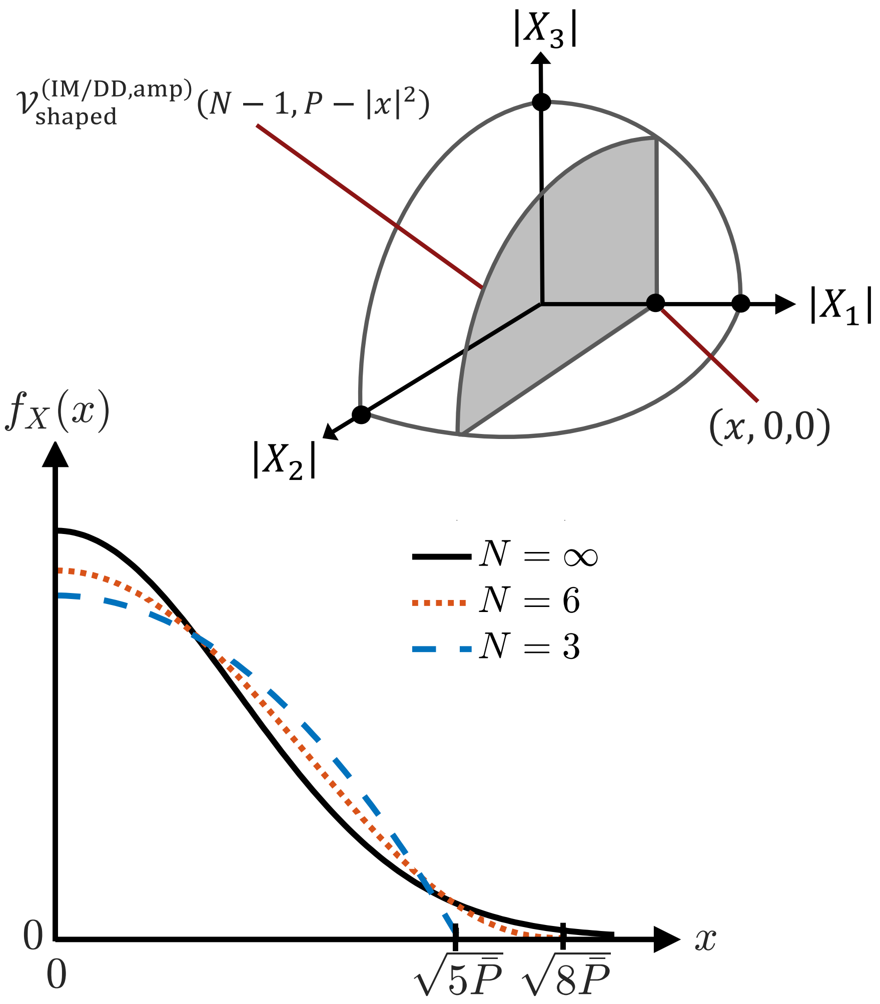
$$\gamma_s = \lim_{N\to\infty} \frac{\bar{P}_{\text{ref}}^{(\text{SD,amp})}(A)}{\bar{P}_{\text{shaped}}^{(\text{SD,amp})}(N,P)} = \frac{\pi e}{6} \approx 1.53 \text{ dB}$$

Compute induced distribution in one basic dimension

$$f_X(x) \propto \mathcal{V}_{\text{shaped}}^{(\text{SD,amp})}(N-1, P-|x|^2)$$

$$\Rightarrow \lim_{N\to\infty} f_X(x) \propto \exp(-x^2/\beta), \qquad x \geq 0$$

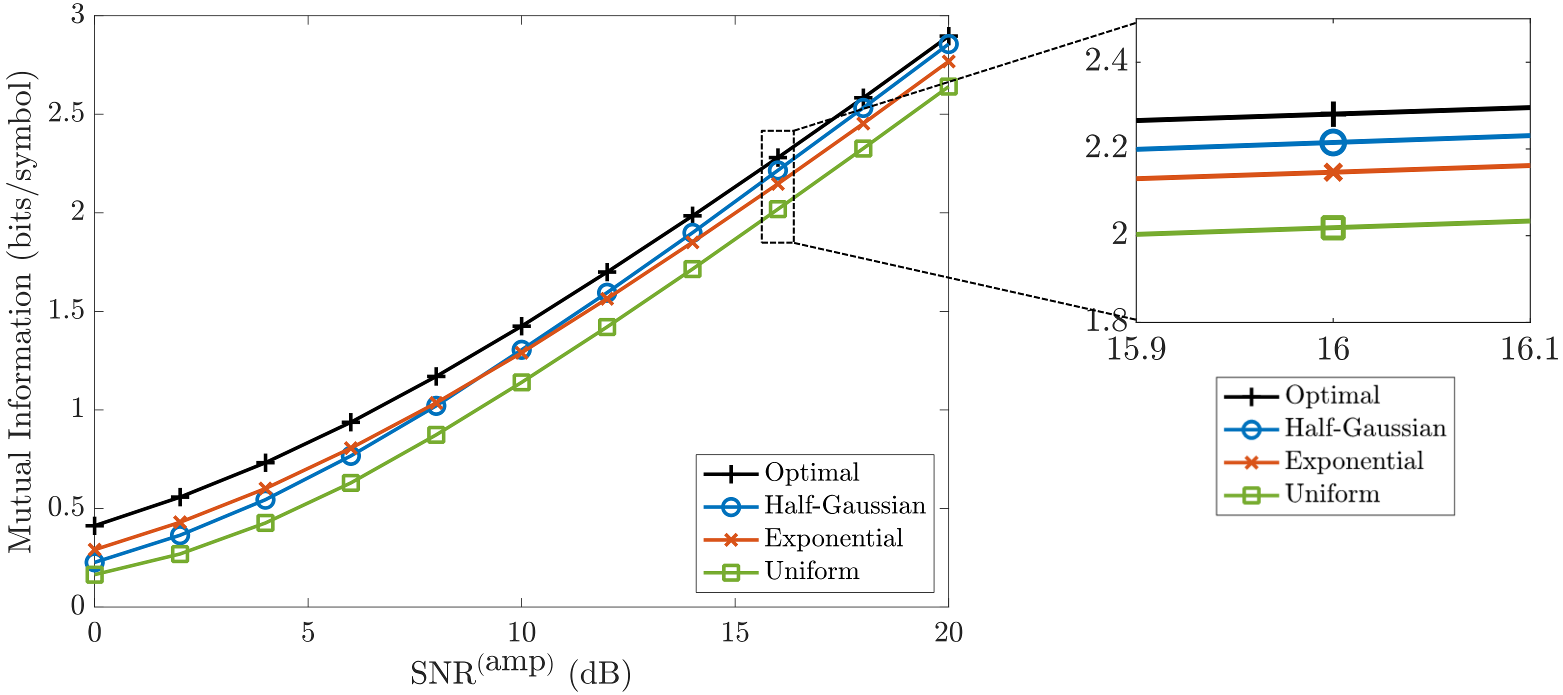
- Optimal input distribution is half-Gaussian in electric field magnitude.



W. Mao and J. M. Kahn, IEEE Trans. Commun. **56** (2008).

Standard Direct Detection, Amplifier Noise (3/3)

Mutual information vs. SNR for various input distributions



- Optimal distribution computed using Blahut-Arimoto algorithm with many intensity levels [2].

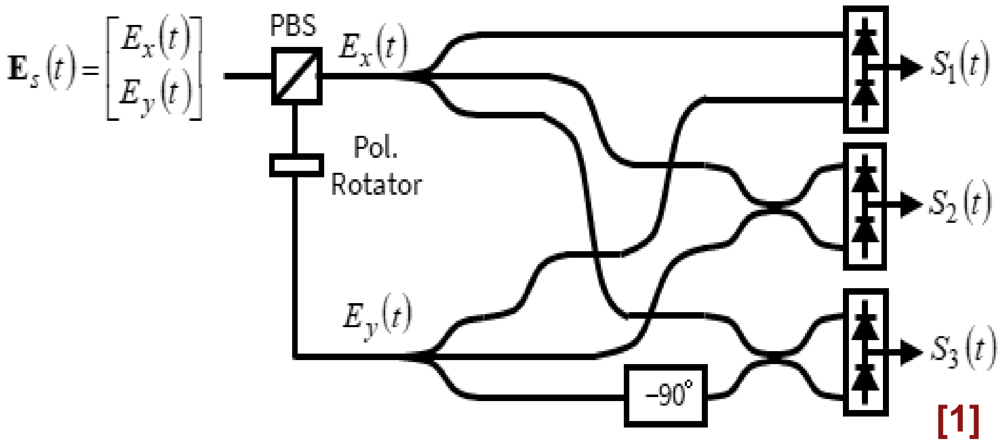
- Half-Gaussian outperforms exponential at high SNR, where signal-ASE beat noise is dominant.

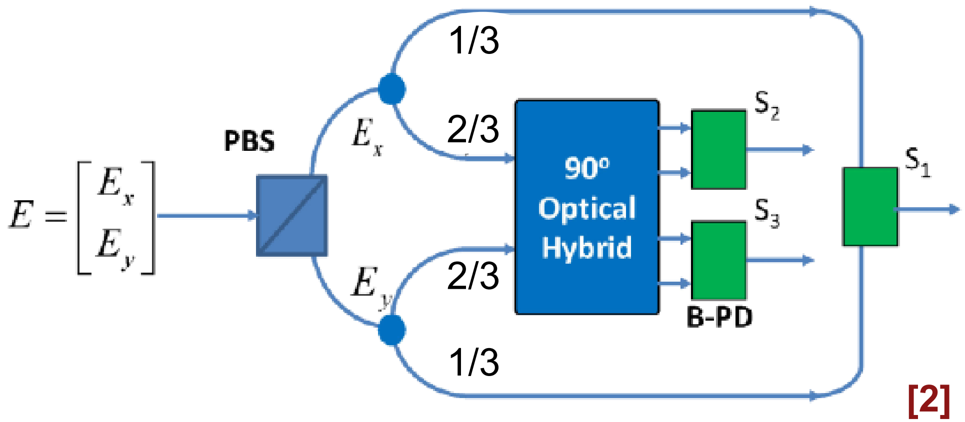
1. W. Mao and J. M. Kahn, IEEE Trans. Commun. **56** (2008).
2. K.-P. Ho, IEEE Photon. Technol. Lett. **17** (2005).

Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise
  - Standard Direct Detection, Amplifier Noise
  - Stokes Vector Detection, Thermal or Amplifier Noise
  - Kramers Kronig Detection, Amplifier Noise
- Discussion

Stokes Vector Receiver





Stokes vector detection

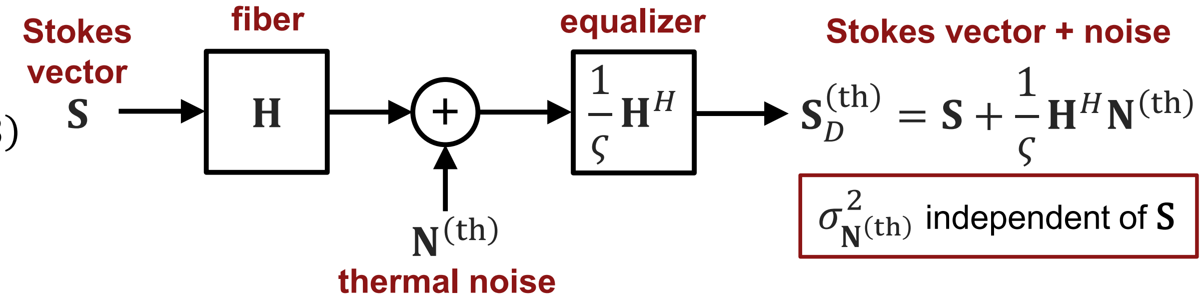
- Hybrid between noncoherent and differentially coherent detection.

Canonical configurations shown

- Directly yield three Stokes parameters with identical signal gains and noise distributions.
- Minimize number of electrical outputs and analog-digital converters.

1. E. Ip, A. P. T. Lau, D. J. F. Barros and J. M. Kahn, Optics Express **16** (2008).
2. W. Shieh, H. Khodakarami and D. Che, APL Photonics **1** (2016).

Stokes Vector Receiver: Thermal Noise (1/3)



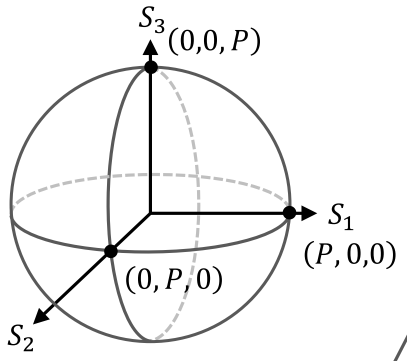
Determine natural coordinates

$$\mathbf{S}_D^{(\mathrm{th})} = \mathbf{S} + \frac{1}{\varsigma}\mathbf{H}^H\mathbf{N}^{(\mathrm{th})},\quad \mathbf{S}, \mathbf{S}_D^{(\mathrm{th})}, \mathbf{N}^{(\mathrm{th})} \in \mathbb{R}^3,\quad \mathbf{N}^{(\mathrm{th})} \sim \mathcal{N}\left(0, \sigma_{th}^2\mathbf{I}_3\right),\quad \mathbf{H} \in \mathrm{O}(3)$$

- For Stoke vectors $\mathbf{S}_1, \mathbf{S}_2$, binary error probability for ML detector:

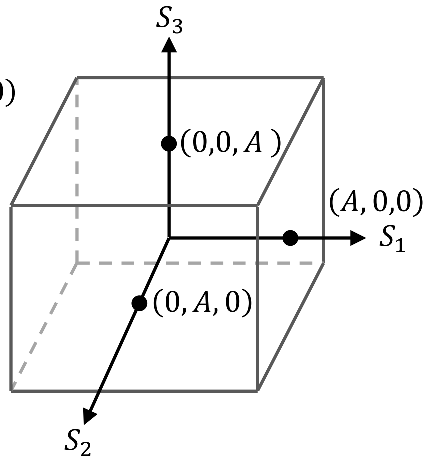
$$P_e = Q\left(\frac{\|\mathbf{S}_1 - \mathbf{S}_2\|_2}{2\sigma_{\mathrm{th}}}\right).$$

- The natural coordinates are the Stokes vector space $\{S_1, S_2, S_3\}$.



Find optimal and reference bounding regions

$$\mathcal{R}_{\mathrm{shaped}}^{(\mathrm{SVR,th})}(N,P) = \left\{\mathbf{S} \in \mathbb{R}^{3N} \,\middle|\, \textstyle\sum_{i=1}^{N}\|\mathbf{S}_i\|_2 = P\right\},\qquad \mathcal{R}_{\mathrm{ref}}^{(\mathrm{SVR,th})}(N,P) = \left\{\mathbf{S} \in \mathbb{R}^{3N} \mid \|\mathbf{S}\|_\infty = A\right\}$$



Compute enclosed volumes

$$\mathcal{V}_{\mathrm{shaped}}^{(\mathrm{SVR,th})}(N,P) = \frac{1}{\Gamma(3N+1)}(8\pi)^N P^{3N},\qquad \mathcal{V}_{\mathrm{ref}}^{(\mathrm{SVR,th})}(N,A) = (2A)^{3N}$$

Compute average powers

$$\bar{P}_{\mathrm{shaped}}^{(\mathrm{SVR,th})}(N,P) = \frac{3}{3N+1}P,\qquad \bar{P}_{\mathrm{ref}}^{(\mathrm{SVR,th})}(A) \approx 0.961A \leq A$$

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

Stokes Vector Receiver: Thermal Noise (2/3)

Compute shape gain (for equal volumes)

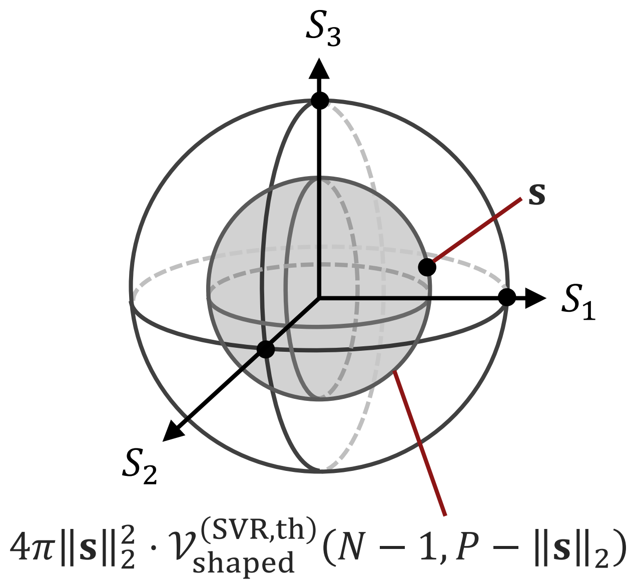
$$\gamma_s = \lim_{N\to\infty} \frac{\bar{P}_{\text{ref}}^{(\text{SVR,th})}(A)}{\bar{P}_{\text{shaped}}^{(\text{SVR,th})}(N,P)} = 0.961\frac{e}{3}\sqrt[3]{\pi} \approx 1.056\text{ dB} \leq \frac{e}{3}\sqrt[3]{\pi} \approx 1.23\text{ dB}$$

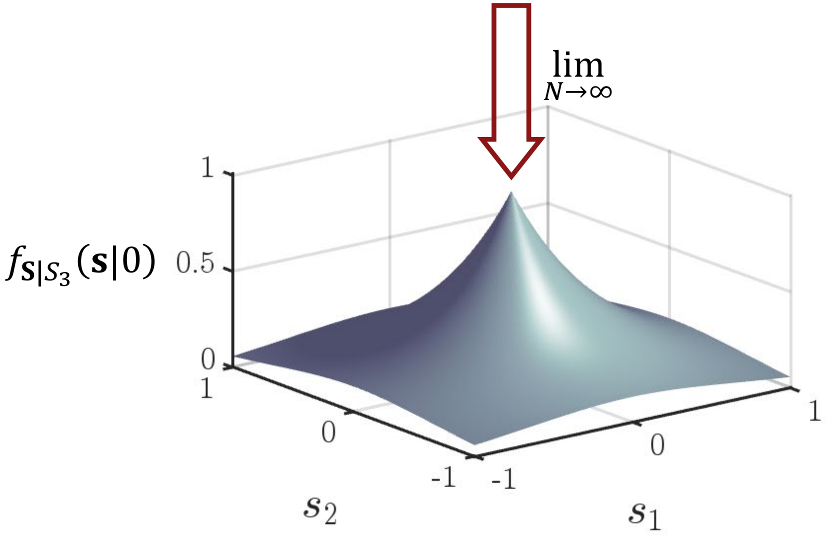
Compute induced distribution in one basic dimension

$$f_{\mathbf{S}}(\mathbf{s}) \propto \mathcal{V}_{\text{shaped}}^{(\text{SVR,th})}(N-1, P-\|\mathbf{s}\|_2)$$

$$\Rightarrow \lim_{N\to\infty} f_{\mathbf{S}}(\mathbf{s}) \propto \exp(-\|\mathbf{s}\|_2/\beta) \propto \exp(-s_0/\beta)$$

- Optimal 3-D input distribution is an exponential in the L^2 norm of the Stokes vector.

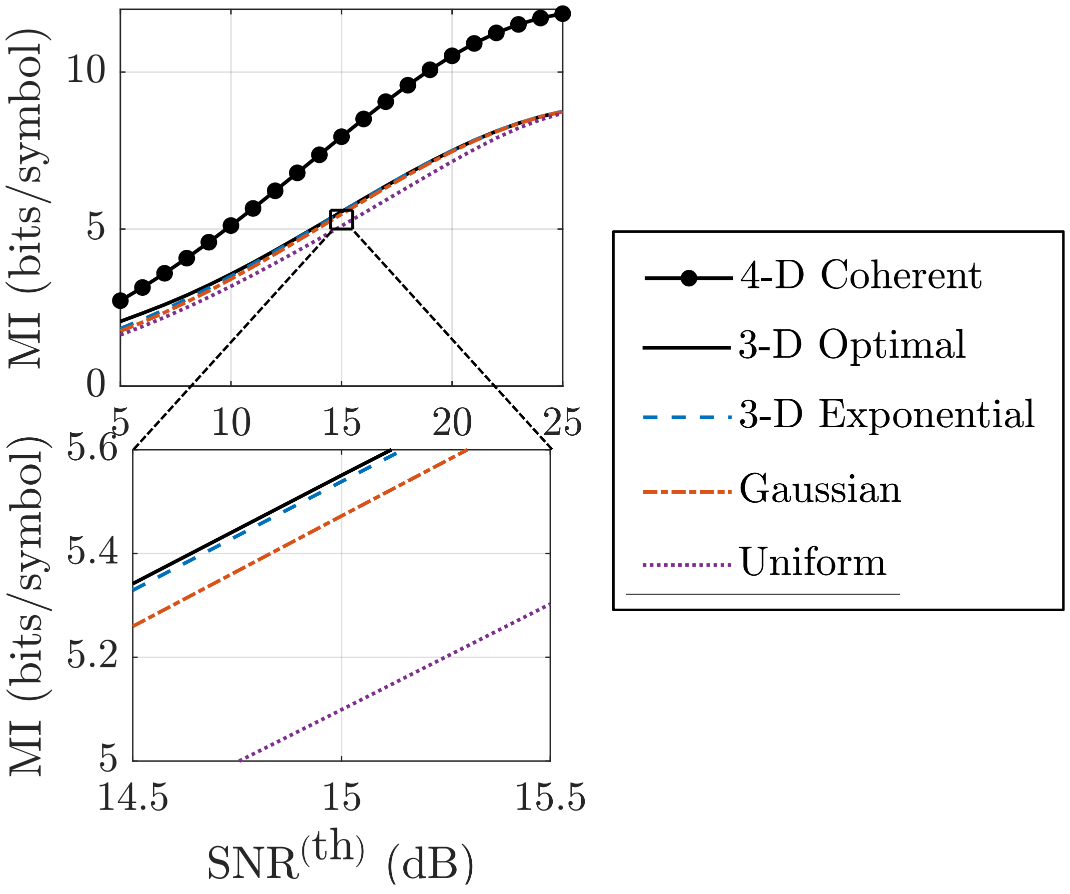




H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

Stokes Vector Receiver: Thermal Noise (3/3)

Mutual information vs. SNR for various input distributions



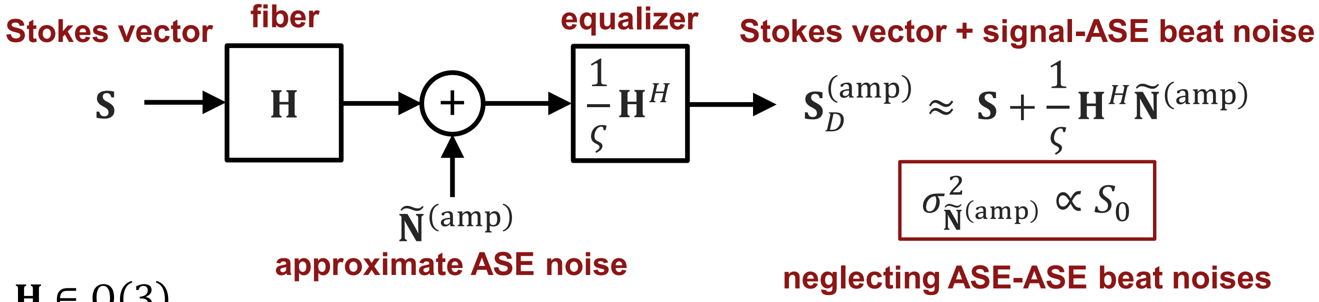
Using Stokes vector receiver unless noted otherwise:

- **4-D Coherent:** standard coherent receiver, input optimized using Blahut-Arimoto (BA) algorithm on dense 4-D lattice.
- **3-D Optimal**: input optimized using BA algorithm on dense 3-D lattice.
- **3-D Exponential**: 3-D exponential $\exp(-\|\mathbf{s}\|_2/\beta)$ with numerical optimization of parameter β.
- **Gaussian**: 3-D Gaussian $\exp(-\|\mathbf{s}\|_2^2/\beta)$ with numerical optimization of parameter β.
- **Uniform**: independent uniform distribution on each Stokes parameter.

- 3-D exponential outperforms Gaussian and performs close to 3-D optimal.

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

SVR: Amplifier Noise-Limited Regime (1/3)



Determine natural coordinates

- Ignoring spontaneous-spontaneous beat noise:

$$\mathbf{S}_D^{(\text{amp})} = \mathbf{S} + \frac{1}{\varsigma}\mathbf{H}^H\widetilde{\mathbf{N}}^{(\text{amp})},$$

$$\mathbf{S}, \mathbf{S}_D^{(\text{amp})}, \widetilde{\mathbf{N}}^{(\text{amp})} \in \mathbb{R}^3, \qquad \widetilde{\mathbf{N}}^{(\text{amp})}\big|\mathbf{S} \sim \mathcal{N}\left(0, 4\sigma^2_{\text{amp}}\sqrt{\varsigma}S_0\mathbf{I}_3\right), \qquad \mathbf{H} \in O(3)$$

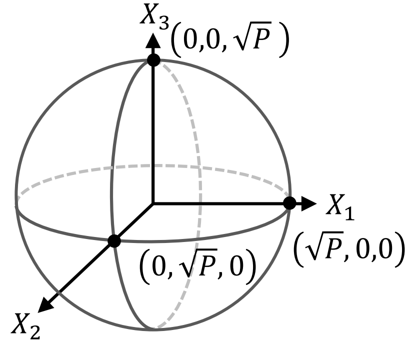
- Scaled Stokes vector:

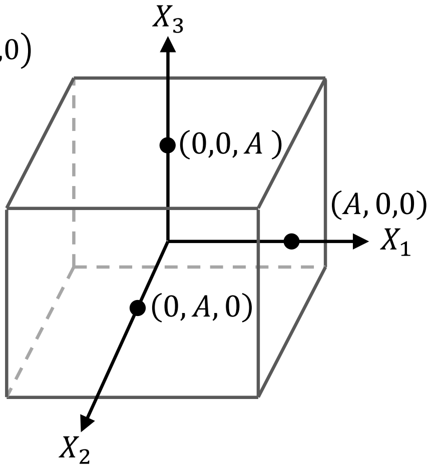
$$[X_1, X_2, X_3]^T = [S_1, S_2, S_3]^T/\sqrt{S_0}$$

- For Stokes vectors $\mathbf{S}_1, \mathbf{S}_2$, binary error probability for approximate ML detector:

$$P_e = Q\left(\frac{\|\mathbf{S}_1 - \mathbf{S}_2\|_2}{2\sigma_{\text{amp}}\left(\sqrt{\mathbf{S}_{1,0}} + \sqrt{\mathbf{S}_{2,0}}\right)}\right) \approx Q\left(\frac{\|\mathbf{X}_1 - \mathbf{X}_2\|_2}{4\sigma_{\text{amp}}}\right).$$

- The natural coordinates are the scaled Stokes vector space $\{X_1, X_2, X_3\}$.





Find optimal and reference bounding regions

$$\mathcal{R}_{\text{shaped}}^{(\text{SVR,amp})}(N,P) = \{\mathbf{X} \in \mathbb{R}^{3N} \,|\, \|\mathbf{X}\|_2^2 = P\}, \qquad \mathcal{R}_{\text{ref}}^{(\text{SVR,amp})}(N,P) = \left\{\mathbf{X} \in \mathbb{R}^{3N} \,\middle|\, \|\mathbf{X}\|_\infty = A\right\}$$

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

SVR: Amplifier Noise-Limited Regime (2/3)

Compute enclosed volumes

$$\mathcal{V}_{\text{shaped}}^{(\text{SVR,amp})}(N,P) = \frac{\pi^{3N/2}}{\Gamma(3N/2+1)}(P)^{3N/2}, \qquad \mathcal{V}_{\text{ref}}^{(\text{SVR,amp})}(N,A) = (2A)^{3N}$$

Compute average powers

$$\bar{P}_{\text{shaped}}^{(\text{SVR,amp})}(N,P) = \frac{3}{3N+2}P, \qquad \bar{P}_{\text{ref}}^{(\text{SVR,amp})}(A) = A^2$$

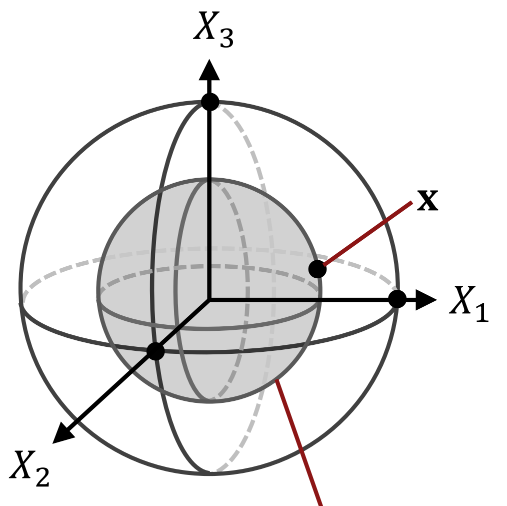
Compute shape gain (for equal volumes)

$$\gamma_s = \lim_{N\to\infty} \frac{\bar{P}_{\text{ref}}^{(\text{SVR,amp})}(A)}{\bar{P}_{\text{shaped}}^{(\text{SVR,amp})}(N,P)} = \frac{\pi e}{6} \approx 1.53 \text{ dB}$$

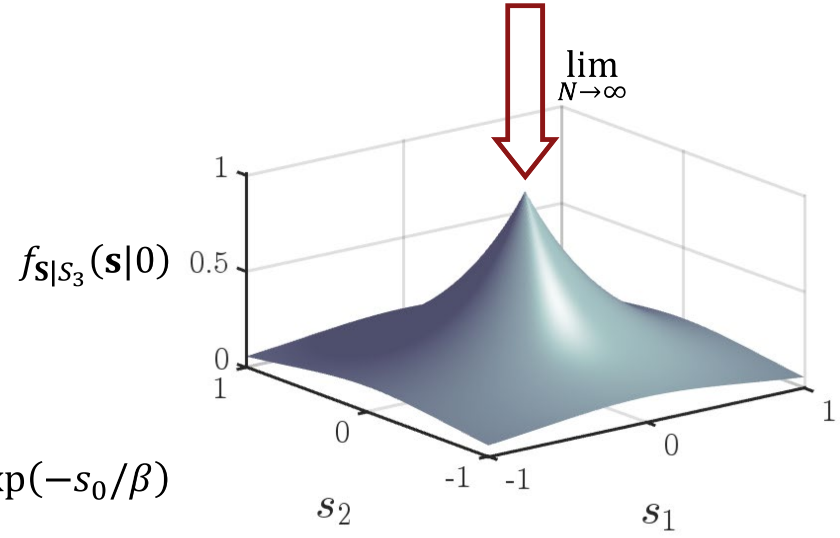
Compute induced distribution in one basic dimension

$$f_{\mathbf{X}}(\boldsymbol{x}) \propto \mathcal{V}_{\text{shaped}}^{(\text{SVR,amp})}(N-1, P-\|\boldsymbol{x}\|_2^2)$$

$$\lim_{N\to\infty} f_{\mathbf{X}}(\mathbf{x}) \propto \exp(-\|\boldsymbol{x}\|_2^2/\beta) \Rightarrow \lim_{N\to\infty} f_{\mathbf{S}}(\mathbf{s}) \propto \frac{1}{\|\mathbf{s}\|_2^{3/2}}\exp(-\|\mathbf{s}\|_2/\beta) \approx \exp(-\|\mathbf{s}\|_2/\beta) = \exp(-s_0/\beta)$$



$$4\pi\|\mathbf{x}\|_2^2 \cdot \mathcal{V}_{\text{shaped}}^{(\text{SVR,th})}(N-1, P-\|\mathbf{x}\|_2^2)$$

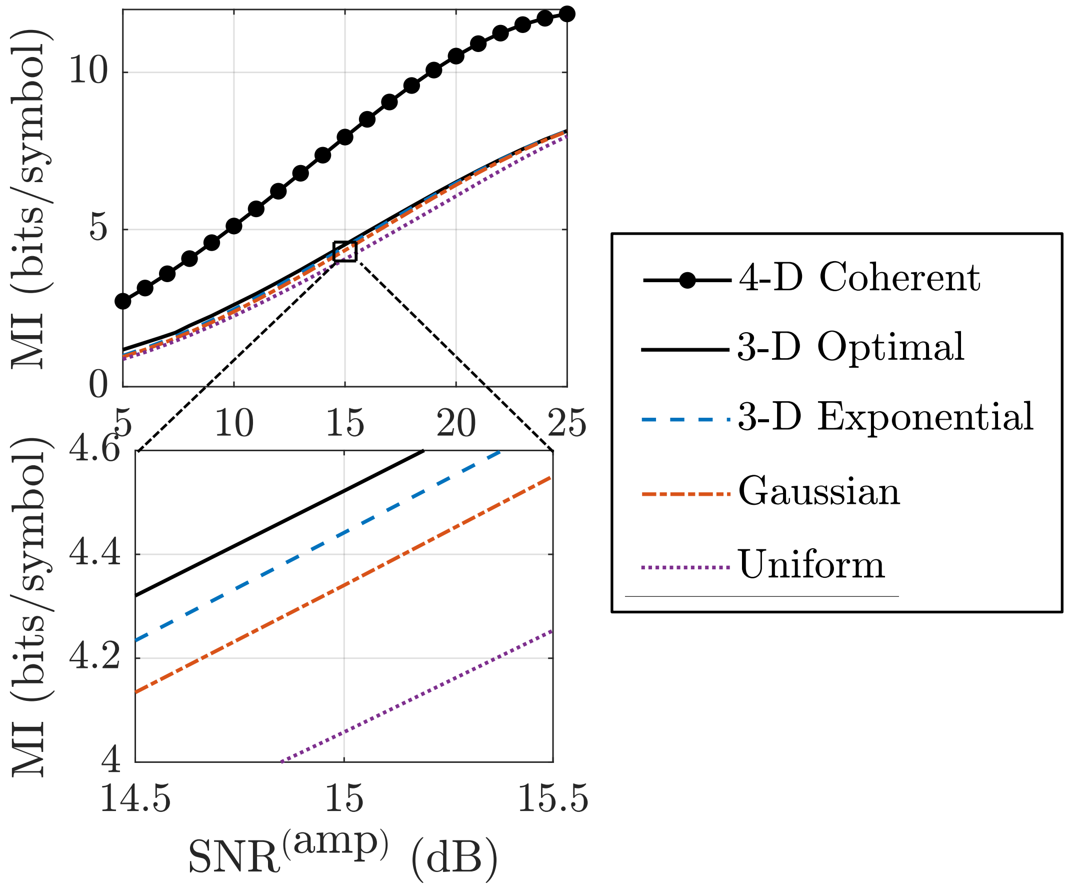


- At high SNR, optimal 3-D input distribution is approximately exponential in L^2 norm of Stokes vector.

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

SVR: Amplifier Noise-Limited Regime (3/3)

Mutual information vs. SNR for various input distributions

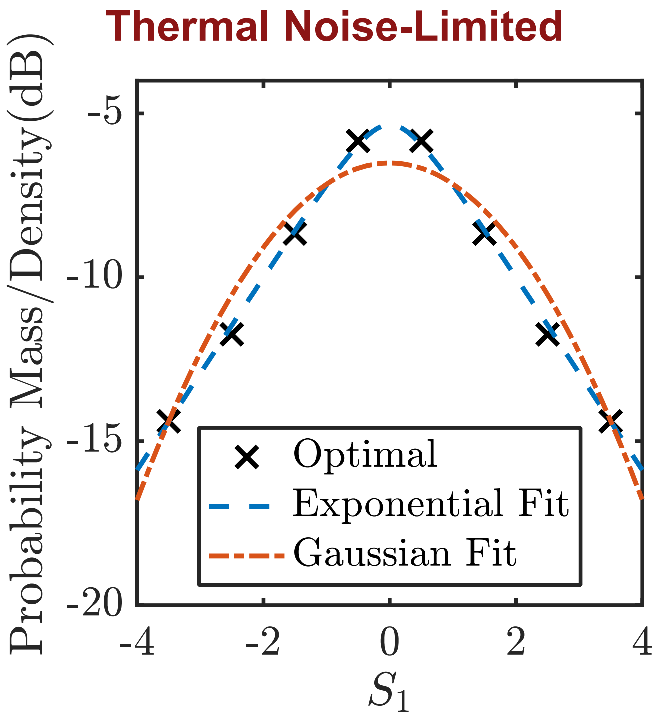


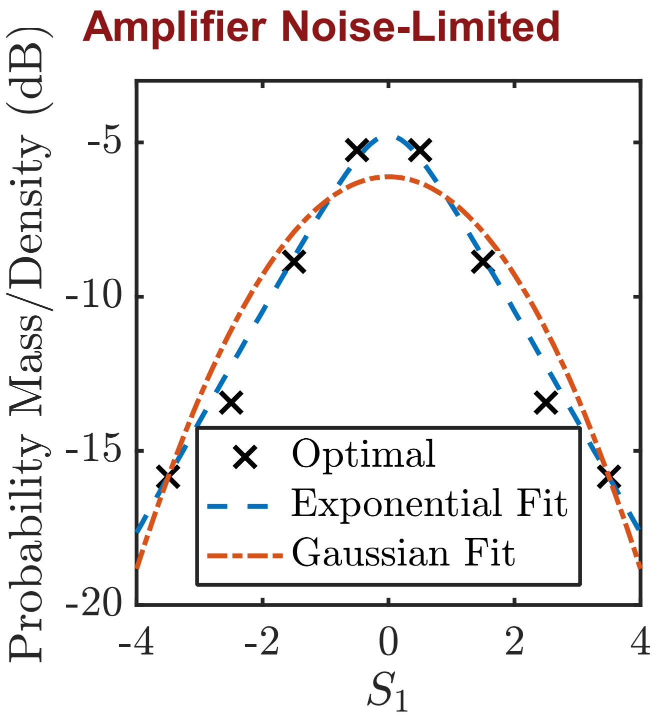
Using Stokes vector receiver unless noted otherwise:

- Performance evaluation includes ASE-ASE beat noises.
- **4-D Coherent:** standard coherent receiver, input optimized using Blahut-Arimoto (BA) algorithm on dense 4-D lattice.
- **3-D Optimal**: input optimized using BA algorithm on dense 3-D lattice.
- **3-D Exponential**: 3-D exponential $\exp(-\|\mathbf{s}\|_2/\beta)$ with numerical optimization of parameter β.
- **Gaussian**: 3-D Gaussian $\exp(-\|\mathbf{s}\|_2^2/\beta)$ with numerical optimization of parameter β.
- **Uniform**: independent uniform distribution on each Stokes parameter.
- 3-D exponential outperforms Gaussian and performs close to 3-D optimal.

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

SVR: Exponential vs. Gaussian Fit in Both Noise Regimes



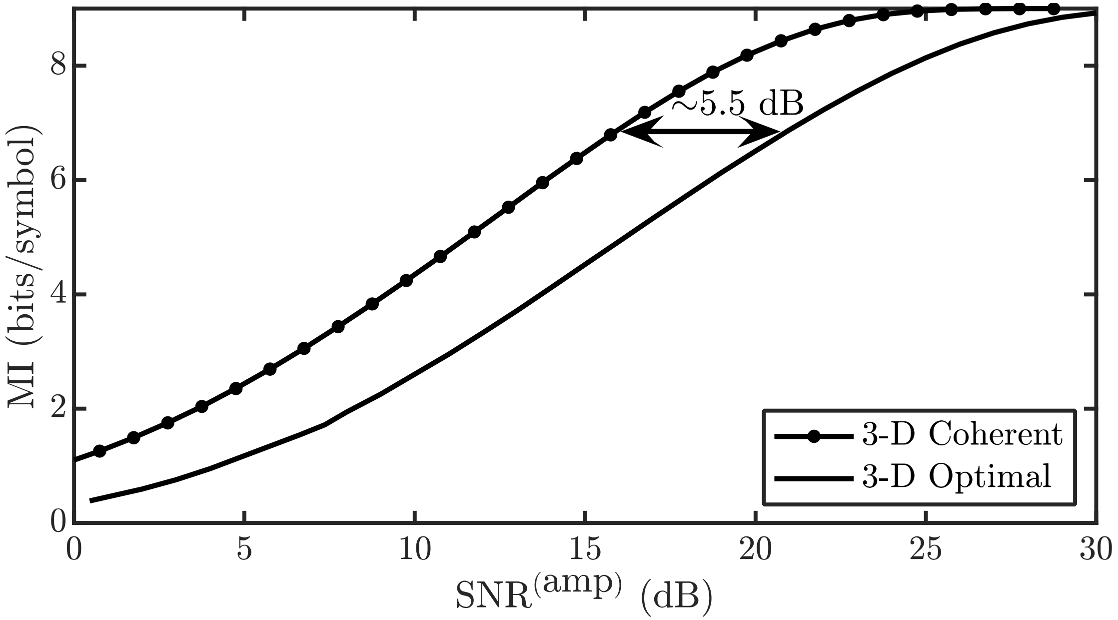


- **Optimal:** $8 \times 8 \times 8$ constellation optimized using BA algorithm.
- **Exponential and Gaussian:** fitted to minimize L^1 norm of error.
- Exponential fits better than Gaussian in both noise regimes.

- $\text{SNR} = 20$ dB
- $S_2 = S_3 = 0.5$

H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

SVR vs. 3-D Coherent in Amplifier Noise-Limited Regime



- **3-D Coherent**: standard coherent using three of four available dimensions, i.e., setting $E_{2Q} = 0$.
- **3-D Optimal**: SVR with input distribution designed using Blahut-Arimoto algorithm.
- Adjust signal powers and noise variances to correspond to equal optical SNRs [1].

Origin of gap: noise enhancement

- In coherent receiver, each component of photocurrent is corrupted only by noise in one polarization and one quadrature.
- In Stokes receiver, each component of photocurrent is corrupted by noises in both polarizations and quadratures.

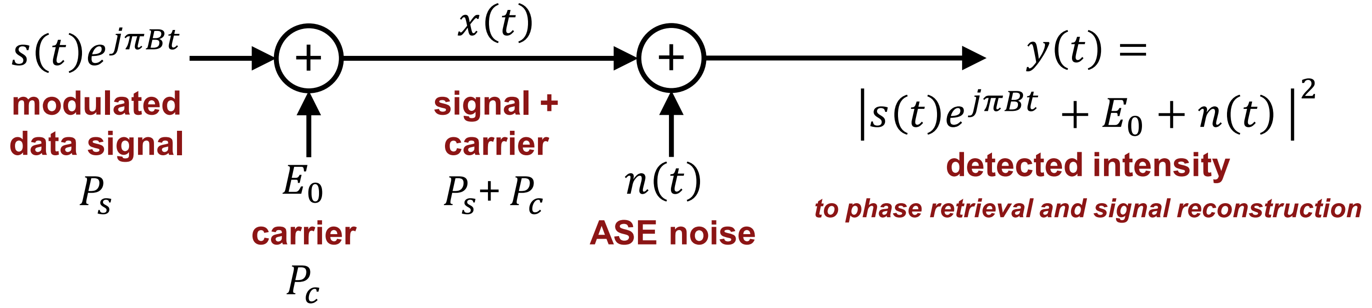
- Can analytically upper-bound the performance gap by 6 dB [1].

1. H. Jia, E. Liang and J. M. Kahn, J. Lightw. Technol. **15** (2023).

Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
  - Standard Direct Detection, Thermal Noise
  - Standard Direct Detection, Amplifier Noise
  - Stokes Vector Detection, Thermal or Amplifier Noise
  - Kramers Kronig Detection, Amplifier Noise
- Discussion

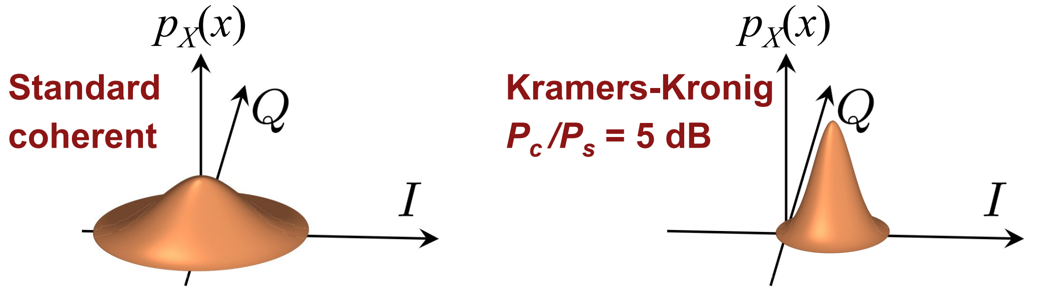
Kramers-Kronig Detection, Amplifier Noise-Limited Regime



- Transmitted signal is $x(t) = s(t)e^{j\pi Bt} + E_0$.
- Modulated data $s(t)e^{j\pi Bt}$ and carrier E_0 do not overlap, so total power is $P_s + P_c$. Assume P_c is sufficient to avoid phase retrieval errors.
- Carrier E_0 conveys no information and $e^{j\pi Bt}$ preserves information. Optimal distribution for $s(t)$ satisfies

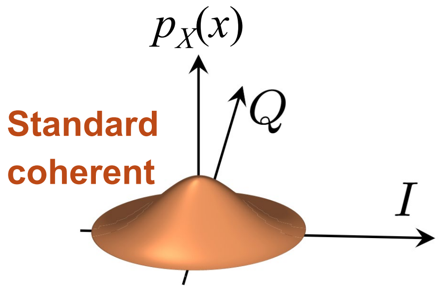
$$\max_{p_S(s)} I(S,Y) \quad \text{s.t.} \quad P_s \leq P_{\max} - P_c.$$

- Optimal $p_S(s)$ is complex circular Gaussian. Optimal $p_X(x)$ is shifted complex circular Gaussian.
- At each SNR, can optimize P_c/P_s numerically to maximize $I(X,Y)$ while achieving low phase error.

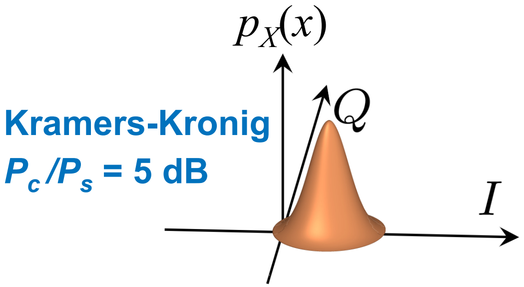


E. Chou, H. Srinivas and J. M. Kahn, J. Lightw. Technol. **40** (2022).

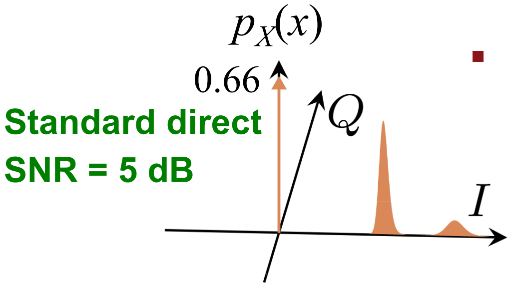
KK vs. Standard Coherent or Direct Detection, Amplifier Noise



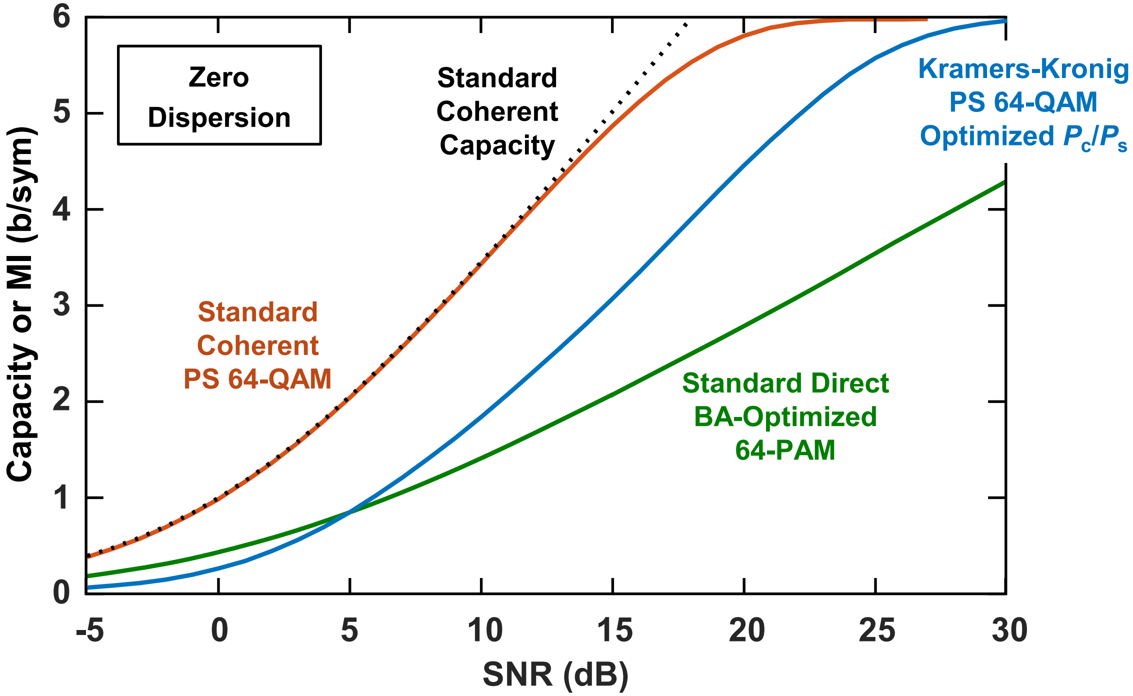
- Provides 2 real dimensions.
- Outperforms KK by 5 to 6 dB.



- Provides ~1.8 real dimensions; must increase P_c/P_s as increase SNR.
- Cannot use zero intensity level.
- Outperforms SD for SNR > 5 dB.



- Provides 1 real dimension.
- Can use zero intensity level, which has least noise.
- Outperforms KK for SNR < 5 dB.

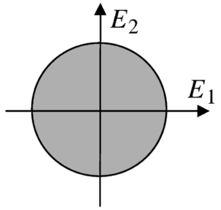
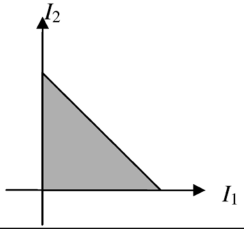
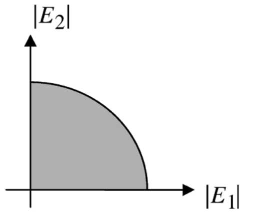
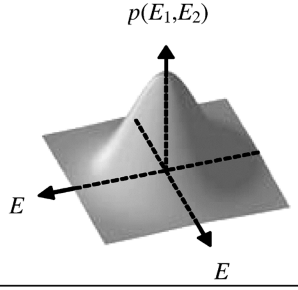
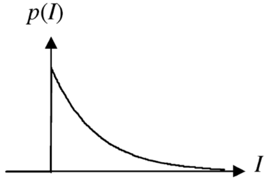
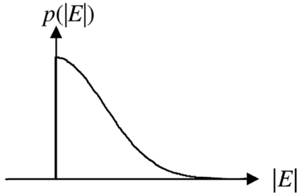


E. Chou, H. Srinivas and J. M. Kahn, J. Lightw. Technol. **40** (2022).

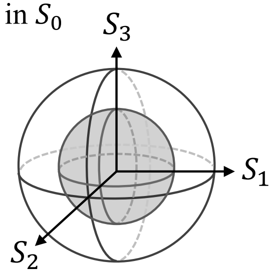
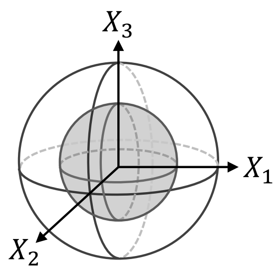
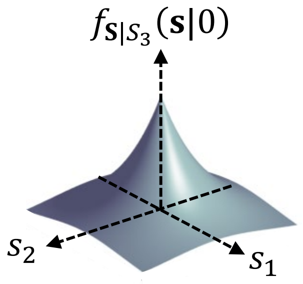
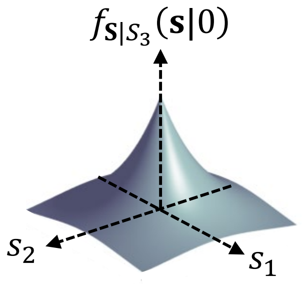
Outline

- Introduction
- Standard Coherent Detection
- Direct Detection Methods
- Discussion
  - Summary of Direct Detection Shaping Results
  - Implementation of Shaping with Direct Detection
  - References

Shaping Results Summary (1/2)

Detection Method & Dominant Noise	Coherent Detection, Amplifier or LO Shot Noise	Noncoherent Detection, Thermal Noise	Noncoherent Detection, Amplifier or LO Shot Noise
Natural Coordinates	2-D electric field $E = \{E_1, E_2\}$	1-D electric field intensity $I = \lvert E\rvert^2$	1-D electric field magnitude $\lvert E\rvert$
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	$p(E) = \frac{1}{\pi P}\exp\left(-\frac{\lvert E\rvert^2}{P}\right)$, $E = (E_1, E_2)$ 	$p(I) = \frac{1}{P}\exp\left(-\frac{I}{P}\right)$, $I \geq 0$ 	$p(\lvert E\rvert) = \sqrt{\frac{2}{\pi P}}\exp\left(-\frac{\lvert E\rvert^2}{2P}\right)$, $\lvert E\rvert \geq 0$ 
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

Shaping Results Summary (2/2)

<table>
<tr><td>Detection Method & Dominant Noise</td><td>Stokes Vector Receiver, Thermal Noise</td><td>Stokes Vector Receiver, Amplifier Noise</td></tr>
<tr><td>Natural Coordinates</td><td>3-D Stokes vector
$\mathbf{S} = \{S_1, S_2, S_3\}$</td><td>3-D scaled Stokes vector
$\mathbf{X} = \{X_1, X_2, X_3\} = \{S_1, S_2, S_3\}/\sqrt{S_0}$</td></tr>
<tr><td>Optimal Shaping Region</td><td>Non-negative orthant bounded by N-simplex in S_0

</td><td>N-sphere centered at the origin

</td></tr>
<tr><td>Induced Optimal Signaling Distribution in Constituent Constellation</td><td>$p(\mathbf{S}) \propto \exp(-S_0/\beta)$

</td><td>$p(\mathbf{S}) \propto \exp(-S_0/\beta)$

</td></tr>
<tr><td>Ultimate Shape Gain</td><td>$\sim 0.961\sqrt[3]{\pi e}/3 \approx 1.056$ dB</td><td>$\pi e/6 \approx 1.53$ dB</td></tr>
<tr><td>Optimal Distribution</td><td>3-D exponential</td><td>3-D exponential</td></tr>
<tr><td>Key Works</td><td>H. Jia, E. Liang and J. M. Kahn, (2023).</td><td>H. Jia, E. Liang and J. M. Kahn, (2023).</td></tr>
</table>

Implementation of Shaping with Direct Detection (1/2)

Obtaining a finite discrete distribution

- Sample the continuous distribution and adjust the shaping parameter (typically β) to maximize MI.

  or

- Use Blahut-Arimoto algorithm to optimize the discrete distribution.

Implementation of Shaping with Direct Detection (2/2)

		Losslessly Separable into 1-D Distributions		Not Losslessly Separable into 1-D Distributions
Symmetric Distribution	**1**	Kramers Kronig (Gaussian)	**3**	SVR Thermal (3-D exponential) SVR Amplifier (3-D exponential)
Asymmetric Distribution	**2**	SD Thermal (1-D exponential) SD Amplifier (1-D half-Gaussian)		

1. Use PAS with CCDM (or other DM)

- Separate into independent 1-D distributions and use a separate DM for each 1-D distribution.

2. Use a suboptimal solution

- **PAS / CCDM (or other DM):** modify target distribution to achieve pairwise symmetry [1].

or

- **Non-PAS scheme:** see [2] for an overview.

3. Use PAS with CCDM (or other DM)

- Directly implement one DM for the 3-D distribution (lossless).

or

- Separate into independent 1-D distributions and use a separate DM for each 1-D distribution (lossy).

1. Z. He, *et al.*, Opt. Express **27** (2019).
2. G. Böcherer, *et al.*, J. Lightw. Technol. **37** (2019).

Thank You!

Key References for This Talk (1/2)

Shaping Fundamentals

1. G. D. Forney and L.-F. Wei, "Multidimensional constellations. I. Introduction, figures of merit, and generalized cross constellations," IEEE Journal on Selected Areas in Communications, vol. 7, no. 6, pp. 877-892, Aug. 1989.
2. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," IEEE Transactions on Information Theory, vol. 39, no. 3, pp. 913-929, May 1993.
3. S. Arimoto, “An algorithm for computing the capacity of arbitrary discrete memoryless channels,” IEEE Transactions on Information Theory, vol. 18, no. 1, pp. 14–20, Jan. 1972.
4. R. Blahut, "Computation of channel capacity and rate-distortion functions," IEEE Transactions on Information Theory, vol. 18, no. 4, pp. 460-473, July 1972.

Shaping for Standard Direct Detection

5. Da-Shan Shiu and J. M. Kahn, "Shaping and nonequiprobable signaling for intensity-modulated signals," IEEE Transactions on Information Theory, vol. 45, no. 7, pp. 2661-2668, Nov. 1999.
6. W. Mao and J. M. Kahn, "Lattice codes for amplified direct-detection optical systems," IEEE Transactions on Communications, vol. 56, no. 7, pp. 1137-1145, July 2008.

Key References for This Talk (2/2)

Shaping for Stokes Vector Detection

7. W. Shieh, H. Khodakarami, and D. Che, “Polarization diversity and modulation for high-speed optical communications: architectures and capacity,” APL Photonics 1, 1 (4): 040801, July 2016.
8. H. Jia, E. Liang and J. M. Kahn, "Optimal Shaping for the Stokes Vector Receiver," Journal of Lightwave Technology, vol. 41, no. 22, pp. 6884-6897, 15 Nov.15, 2023.

Shaping for Kramers-Kronig Detection

9. E. S. Chou, H. Srinivas and J. M. Kahn, "Phase Retrieval-Based Coherent Receivers: Signal Design and Degrees of Freedom," Journal of Lightwave Technology, vol. 40, no. 5, pp. 1296-1307, 1 March1, 2022.

Other Key References

Shaping for Standard Coherent Detection

10. G. Böcherer, P. Schulte and F. Steiner, "Probabilistic Shaping and Forward Error Correction for Fiber-Optic Communication Systems," Journal of Lightwave Technology, vol. 37, no. 2, pp. 230-244, 15 Jan.15, 2019.
11. J. Cho and P. J. Winzer, "Probabilistic Constellation Shaping for Optical Fiber Communications," Journal of Lightwave Technology, vol. 37, no. 6, pp. 1590-1607, 15 March15, 2019.
12. A. Amari et al., "Introducing Enumerative Sphere Shaping for Optical Communication Systems With Short Blocklengths," Journal of Lightwave Technology, vol. 37, no. 23, pp. 5926-5936, 1 Dec.1, 2019.
13. J. Cho, "Probabilistic Constellation Shaping: An Implementation Perspective," 2022 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2022, pp. 1-39.

Shaping for Standard Direct Detection

14. D. Che, J. Cho and X. Chen, "Does Probabilistic Constellation Shaping Benefit IM-DD Systems Without Optical Amplifiers?," Journal of Lightwave Technology, vol. 39, no. 15, pp. 4997-5007, Aug.1, 2021.1109/JLT.2021.3083530.