Modern Wireless Network Design Based on Constrained Capacity

April 20, 2005

Matthew Valenti
Assistant Professor
West Virginia University
Morgantown, WV 26506-6109
mvalenti@wvu.edu

Overview

- Key observations:
  - Capacity approaching binary codes are now practical.
  - M-ary modulation such as PSK, QAM, and FSK continue to be used.
  - Block space time coding is an effective way to modulate across multiple transmit antennas.
- Implications of these observations:
  - It makes sense to study point-to-point links in terms of the capacity under modulation constraints.
  - It is desirable to match binary codes with M-ary modulation.
- Overview of talk:
  - Capacity under modulation constraints.
  - Bit interleaved coded modulation (BICM).
  - BICM with iterative demodulation and EXIT charts.
  - Efficient cross-layer design of retransmission (MAC) and routing (network-layer) protocols.

Noisy Channel Coding Theorem

- Claude Shannon, “A mathematical theory of communication,” *Bell Systems Technical Journal*, 1948.
- Every channel has associated with it a ***capacity*** C.
  – Measured in bits per channel use (modulated symbol).
- The channel capacity is an upper bound on ***information rate*** r.
  – There exists a code of rate $r < C$ that achieves reliable communications.
    - Reliable means an arbitrarily small error probability.

Computing Channel Capacity

- The capacity is the ***mutual information*** between the channel's input X and output Y maximized over all possible input distributions:

$$C = \max_{p(x)} \{ I(X;Y) \}$$

$$= \max_{p(x)} \left\{ \int\int p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} dxdy \right\}$$

Capacity of AWGN with Unconstrained Input

- Consider an AWGN channel with 1-dimensional input:
  - y = x + n
  - where n is Gaussian with variance $N_o/2$
  - x is a signal with average energy (variance) E_s
- The capacity in this channel is:

$$C = \max_{p(x)}\{I(X;Y)\} = \frac{1}{2}\log_2\left(\frac{2E_s}{N_o}+1\right) = \frac{1}{2}\log_2\left(\frac{2rE_b}{N_o}+1\right)$$

  - where E_b is the energy per (information) bit.
- This capacity is achieved by a Gaussian input x.
  - This is not a practical modulation.

Capacity of AWGN with BPSK Constrained Input

- If we only consider antipodal (BPSK) modulation, then

$$X = \pm\sqrt{E_s}$$

- and the capacity is:

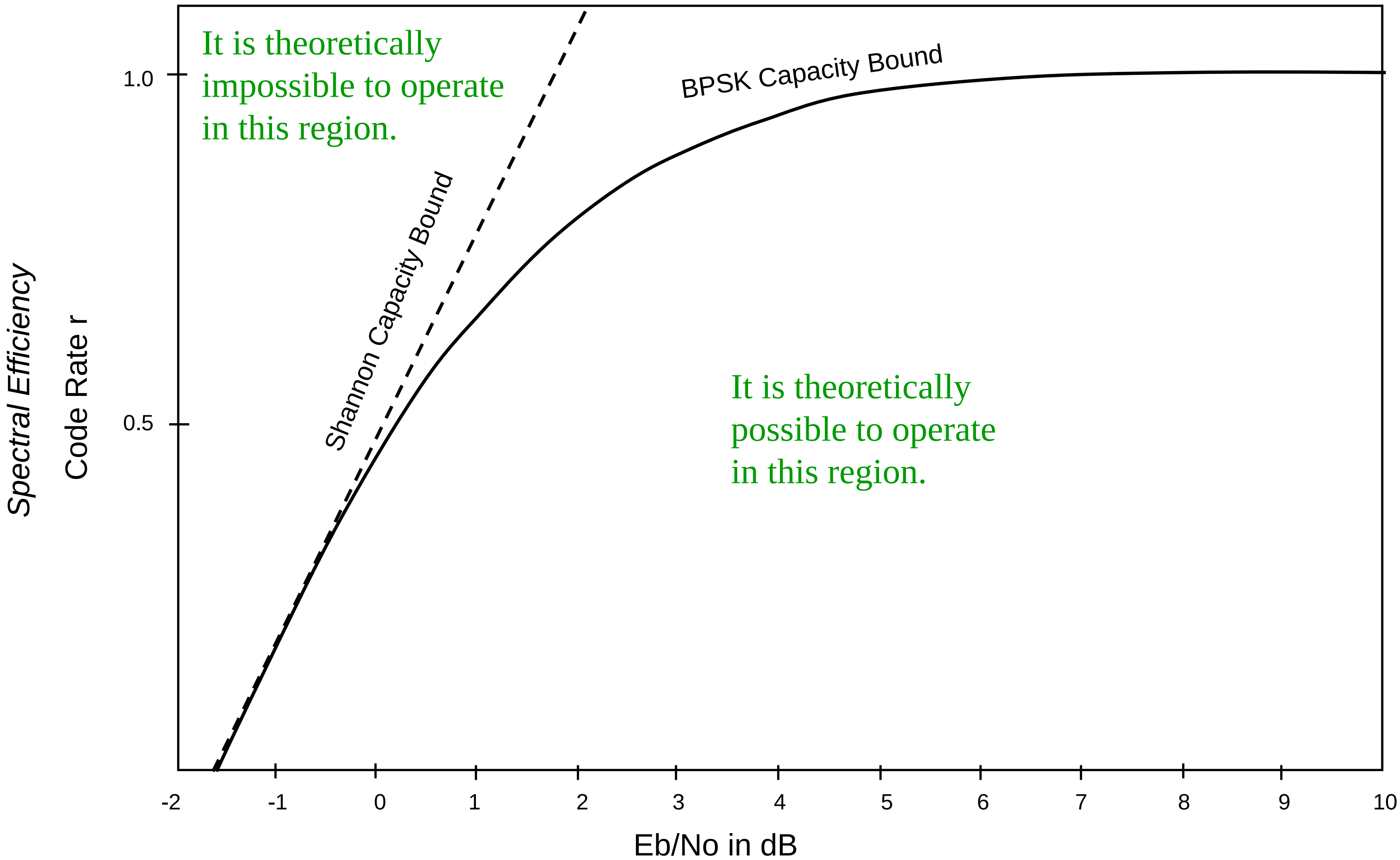
$$\begin{aligned} C &= \max_{p(x)} \{ I(X;Y) \} \\ &= I(X;Y)\big|_{p(x):p=1/2} \\ &= H(Y) - H(N) \\ &= \int_{-\infty}^{\infty} p(y)\log_2 p(y)dy - \frac{1}{2}\log_2(\pi e N_o) \end{aligned}$$

maximized when two signals are equally likely

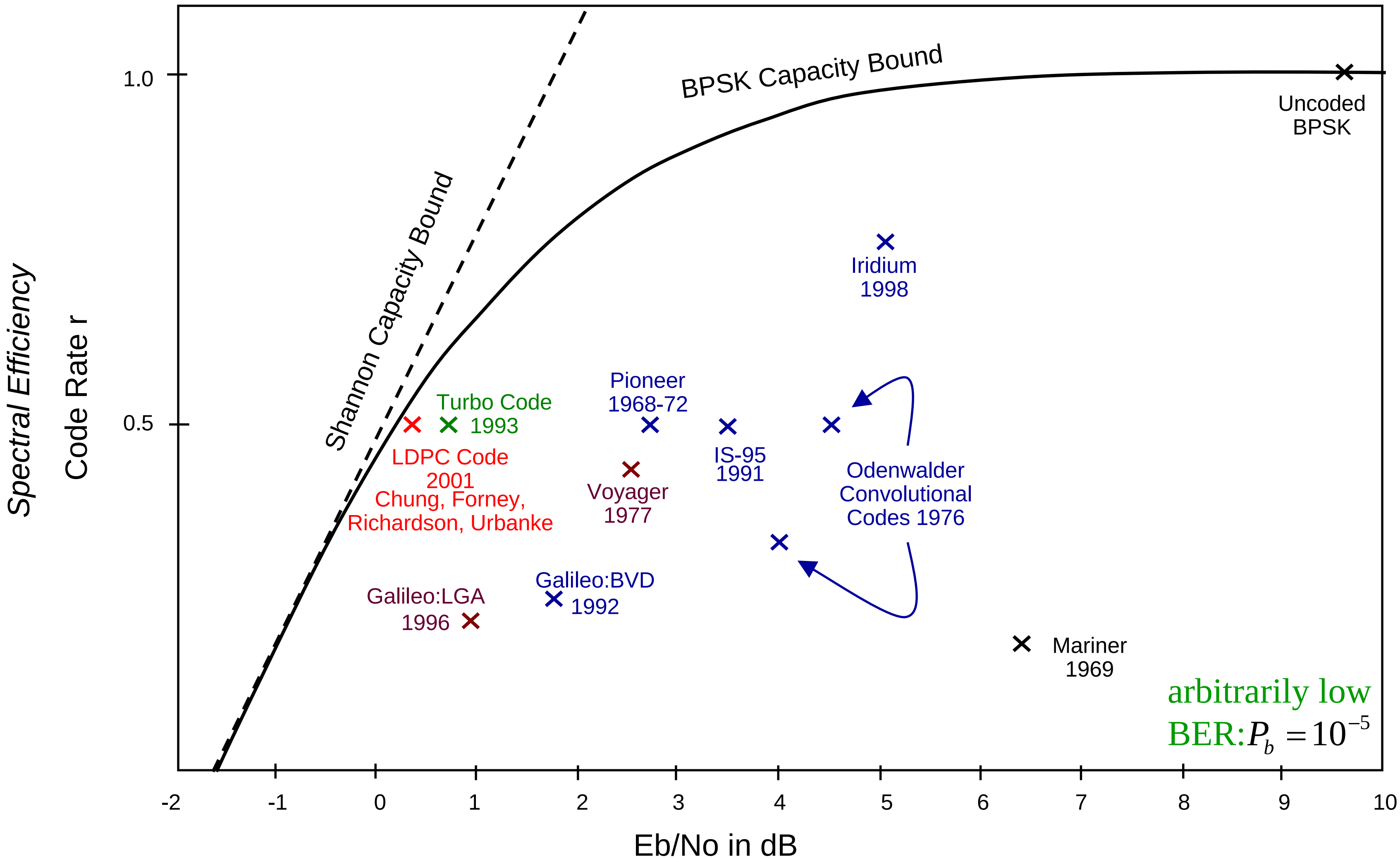
This term must be integrated numerically with

$$p_Y(y) = p_X(y) * p_N(y) = \int_{-\infty}^{\infty} p_X(\lambda) p_N(y-\lambda) d\lambda$$

Capacity of AWGN w/ 1-D Signaling



Power Efficiency of Standard Binary Channel Codes



M-ary modulation

- $\mu = \log_2 M$ bits are mapped to the symbol $\mathbf{x}_k$, which is chosen from the set $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\}$
  - The symbol is multidimensional.
  - 2-D Examples: QPSK, M-PSK, QAM
  - M-D Example: FSK, block space-time codes (BSTC)
- The signal $\mathbf{y} = \mathbf{x}_k + \mathbf{n}$ is received
  - More generally (BSTC), $\mathbf{Y} = \mathbf{HX} + \mathbf{N}$
- For each signal in S, the receiver computes $p(\mathbf{y}|\mathbf{x}_k)$
  - This function depends on the modulation, channel, and receiver.

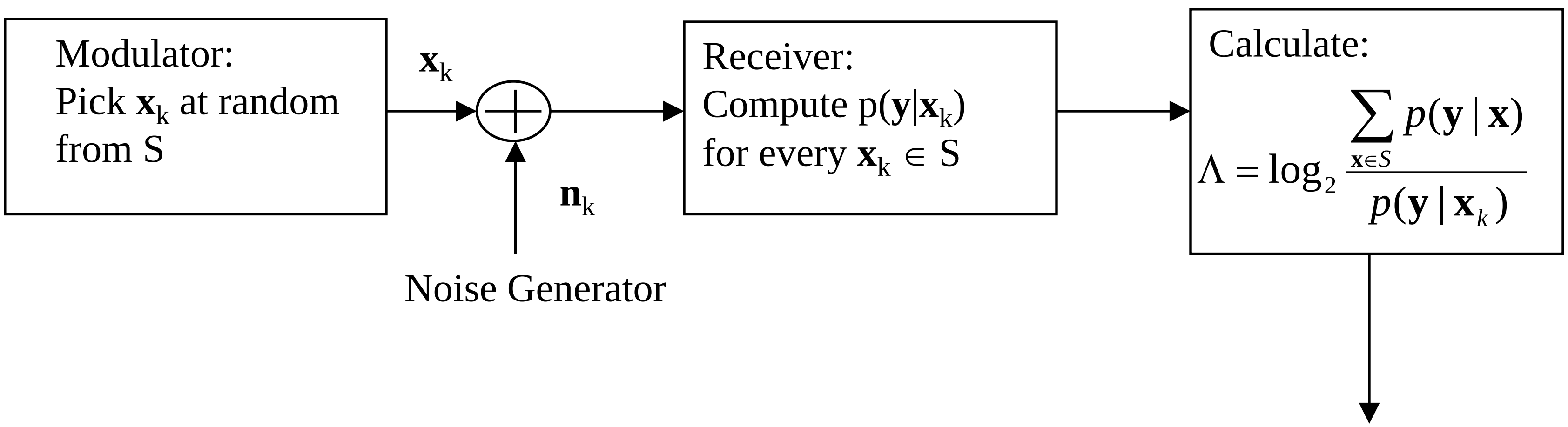
Monte Carlo Approach to Computing Modulation Constrained Capacity

- Suppose we want to compute capacity of M-ary modulation
  - In each case, we cannot control input distribution.
  - The capacity is merely the mutual information between channel input and output.
- The mutual information can be measured as the following expectation:

$$C = I(X;Y) = \mu - E\left[\log_2 \frac{\sum_{\mathbf{x}\in S} p(\mathbf{y}\mid\mathbf{x})}{p(\mathbf{y}\mid\mathbf{x}_k)}\right]$$

- This expectation can be obtained through Monte Carlo simulation.

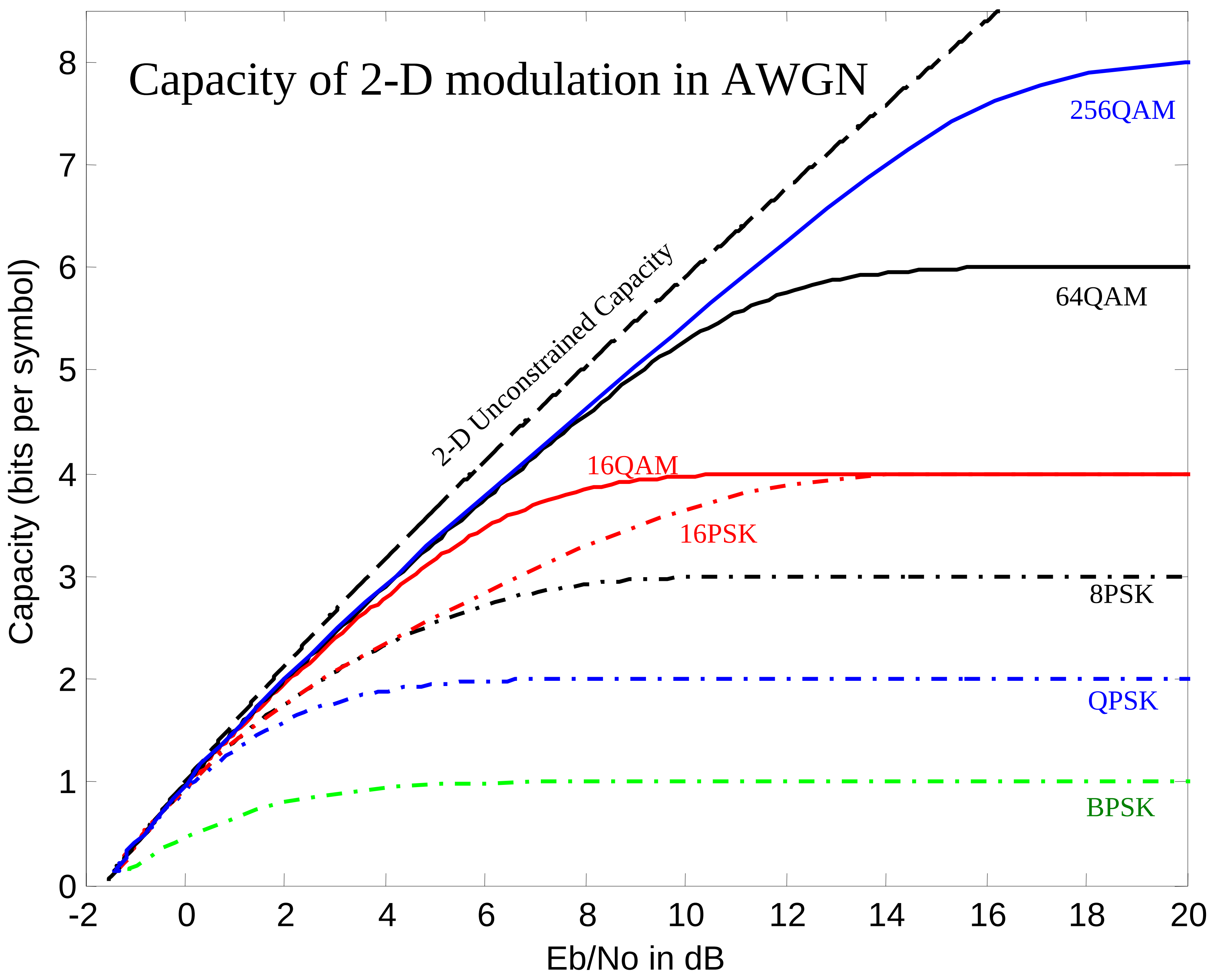
Simulation Block Diagram



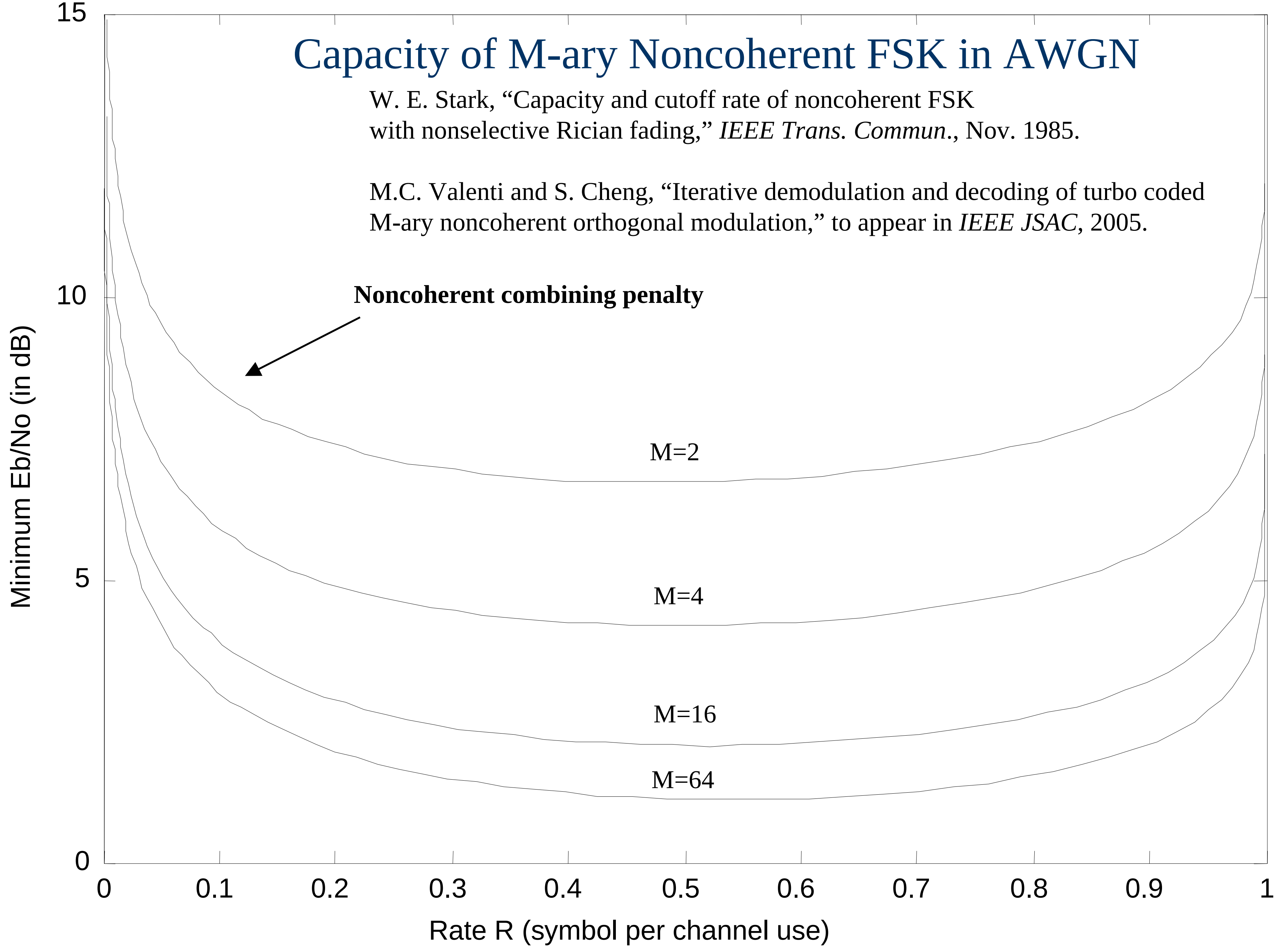
After running many trials, calculate:

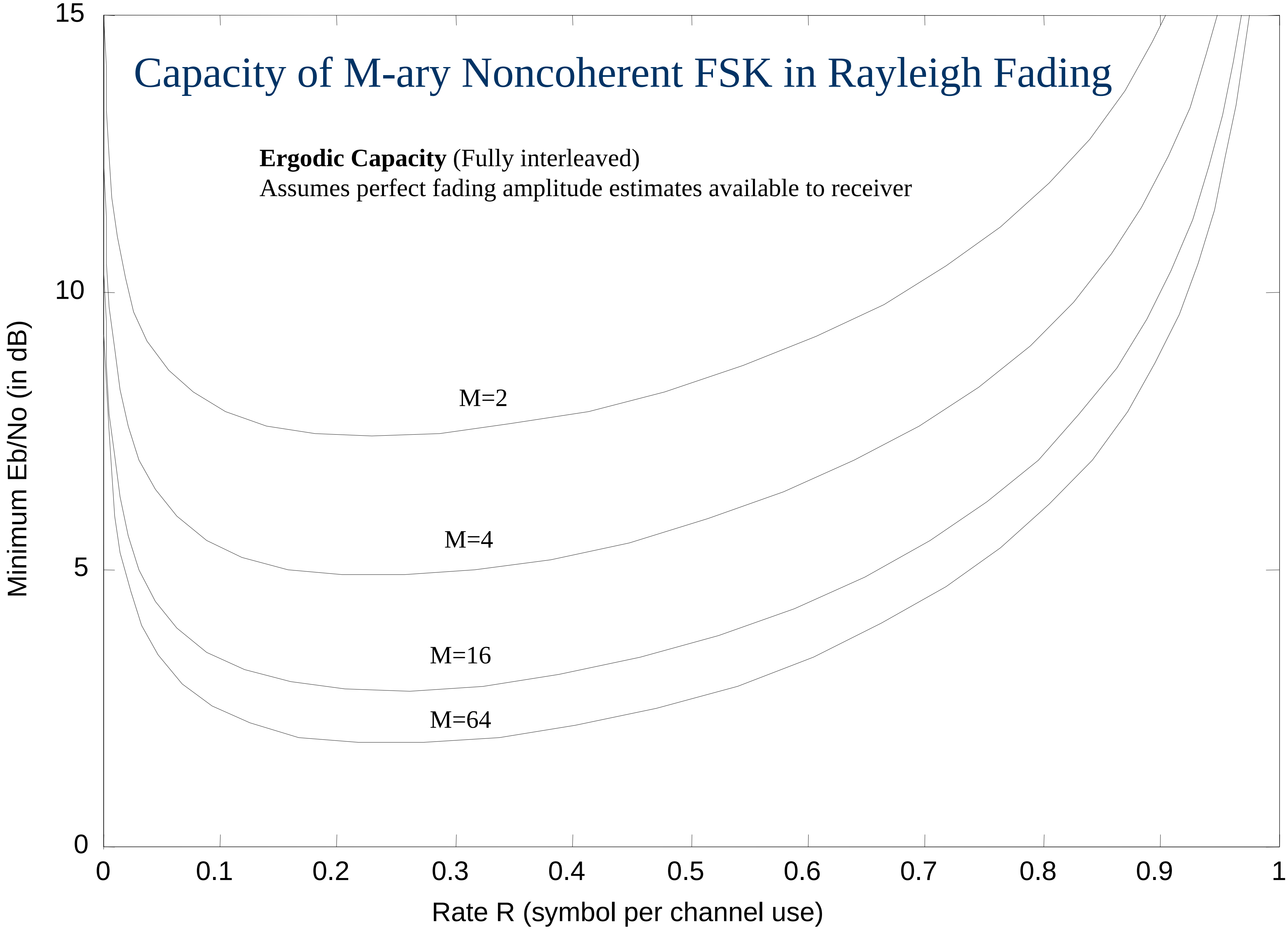
$$C = \mu - E[\Lambda]$$

Benefits of Monte Carlo approach:
-Allows high dimensional signals to be studied.
-Can determine performance in fading.
-Can study influence of receiver design.



Capacity of 2-D modulation in AWGN
2-D Unconstrained Capacity
256QAM
64QAM
16QAM
16PSK
8PSK
QPSK
BPSK
Capacity (bits per symbol)
Eb/No in dB
0
1
2
3
4
5
6
7
8
-2
0
2
4
6
8
10
12
14
16
18
20

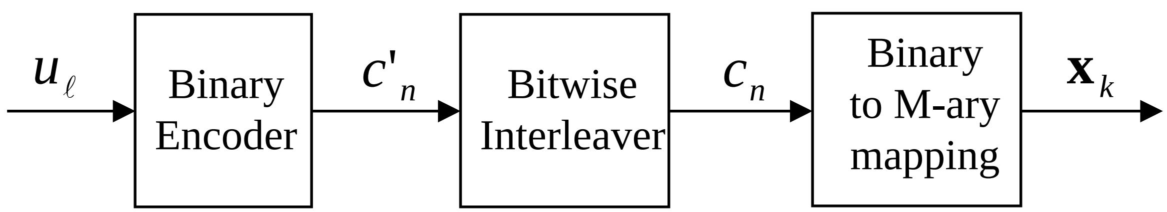

Capacity of M-ary Noncoherent FSK in AWGN
W. E. Stark, “Capacity and cutoff rate of noncoherent FSK with nonselective Rician fading,” IEEE Trans. Commun., Nov. 1985.
M.C. Valenti and S. Cheng, “Iterative demodulation and decoding of turbo coded M-ary noncoherent orthogonal modulation,” to appear in IEEE JSAC, 2005.
Noncoherent combining penalty
M=2
M=4
M=16
M=64
15
10
5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Minimum Eb/No (in dB)
Rate R (symbol per channel use)



Capacity of M-ary Noncoherent FSK in Rayleigh Fading
Ergodic Capacity (Fully interleaved)
Assumes perfect fading amplitude estimates available to receiver
M=2
M=4
M=16
M=64
15
10
5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Minimum Eb/No (in dB)
Rate R (symbol per channel use)

BICM

- Coded modulation (CM) is required to attain the aforementioned capacity.
  - Channel coding and modulation handled jointly.
  - e.g. trellis coded modulation (Ungerboeck); coset codes (Forney)
- Most off-the-shelf capacity approaching codes are binary.
- A pragmatic system would use a binary code followed by a bitwise interleaver and an M-ary modulator.
  - Bit Interleaved Coded Modulation (BICM); Caire 1998.



BICM Receiver

- Like the CM receiver, the BICM receiver calculates p(**y**|**x**$_k$) for each signal in S.
- Furthermore, the BICM receiver needs to calculate the log-likelihood ratio of each code bit:

$$\lambda_n = \log \frac{\sum_{\mathbf{x} \in S_n^{(1)}} p(\mathbf{y}|\mathbf{x})}{\sum_{\mathbf{x} \in S_n^{(0)}} p(\mathbf{y}|\mathbf{x})}$$

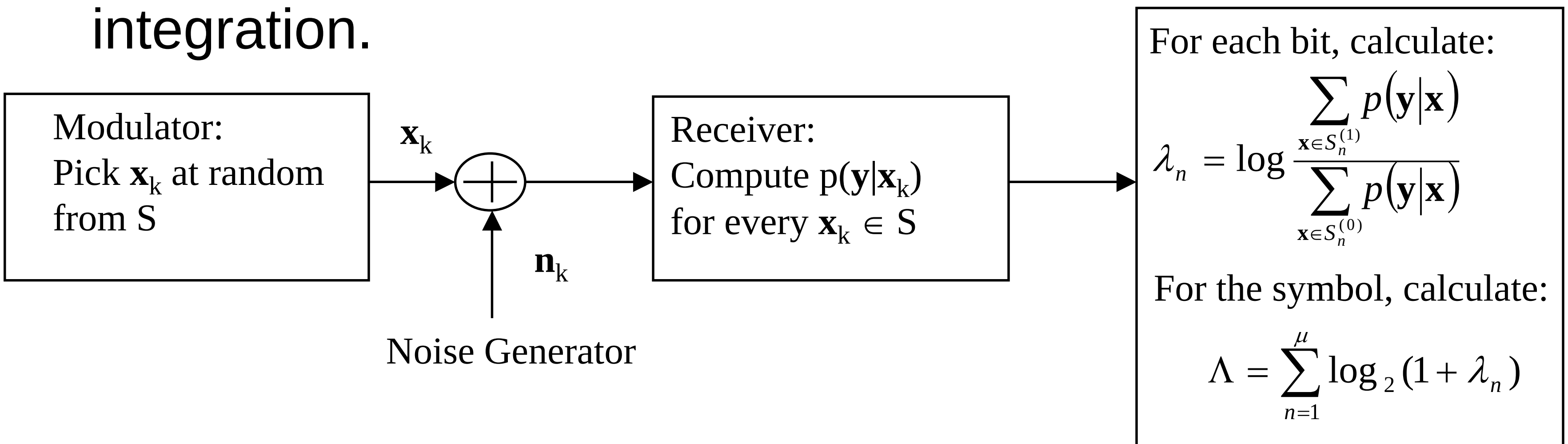
  - where $S_n^{(1)}$ represents the set of symbols whose nth bit is a 1.
  - and $S_n^{(0)}$ is the set of symbols whose nth bit is a 0.

BICM Capacity

- The BICM capacity is then [Caire 1998]:

$$C = I(X;Y) = \mu - E\left[\sum_{n=1}^{\mu} \log_2(1+\lambda_n)\right]$$

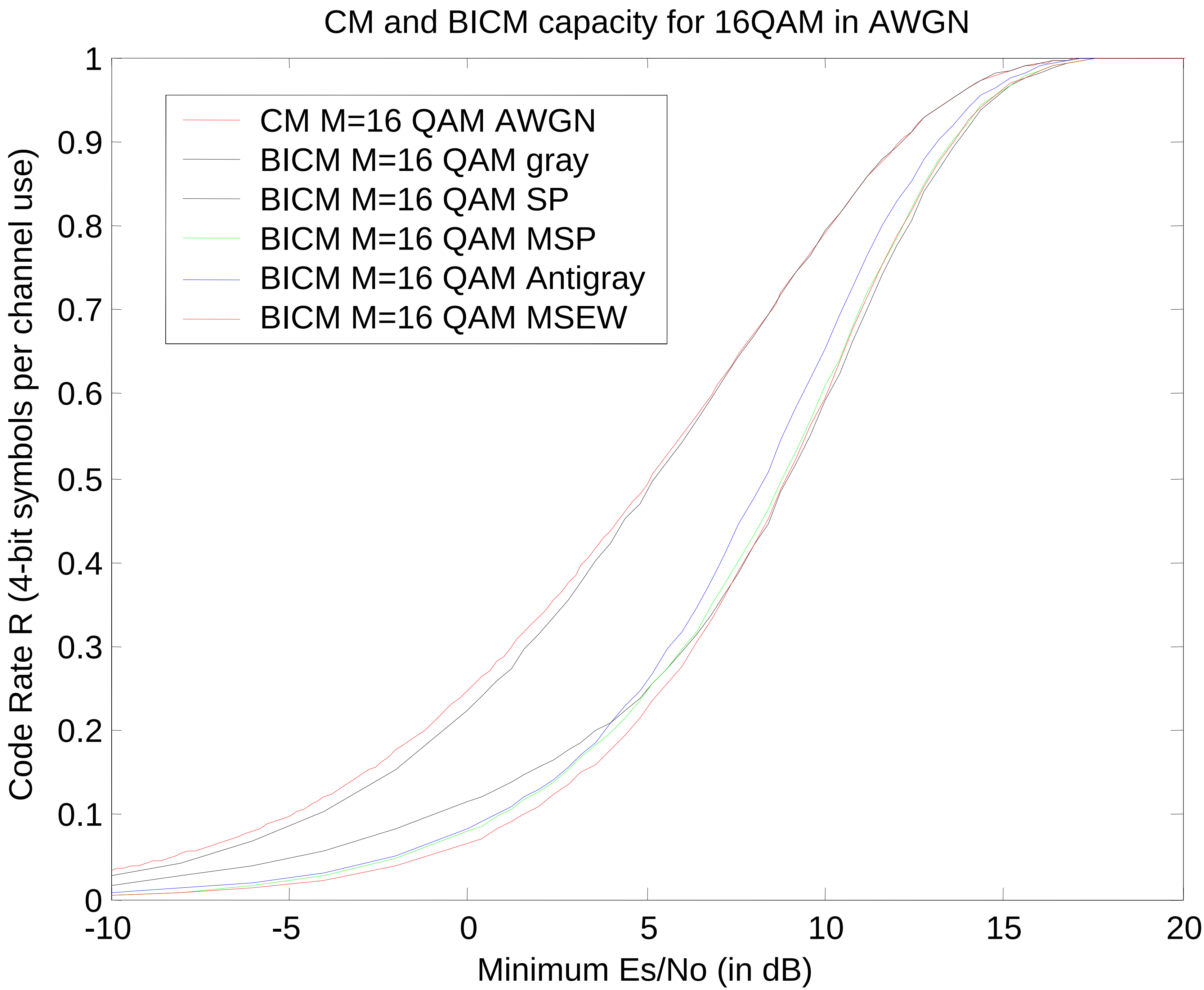
- As with CM, this can be computed using a Monte Carlo integration.



After running many trials, calculate:

$$C = \mu - E[\Lambda]$$

Unlike CM, the capacity of BICM depends on how bits are mapped to symbols



CM and BICM capacity for 16QAM in AWGN
CM M=16 QAM AWGN
BICM M=16 QAM gray
BICM M=16 QAM SP
BICM M=16 QAM MSP
BICM M=16 QAM Antigray
BICM M=16 QAM MSEW
Code Rate R (4-bit symbols per channel use)
Minimum Es/No (in dB)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-10
-5
0
5
10
15
20

BICM-ID

- The conventional BICM receiver assumes that all bits in a symbol are equally likely:

$$\lambda_n = \log \frac{\sum_{\mathbf{x} \in S_n^{(1)}} p(\mathbf{y}|\mathbf{x})}{\sum_{\mathbf{x} \in S_n^{(0)}} p(\mathbf{y}|\mathbf{x})}$$

- However, if the receiver has estimates of the bit probabilities, it can use this to weight the symbol likelihoods.

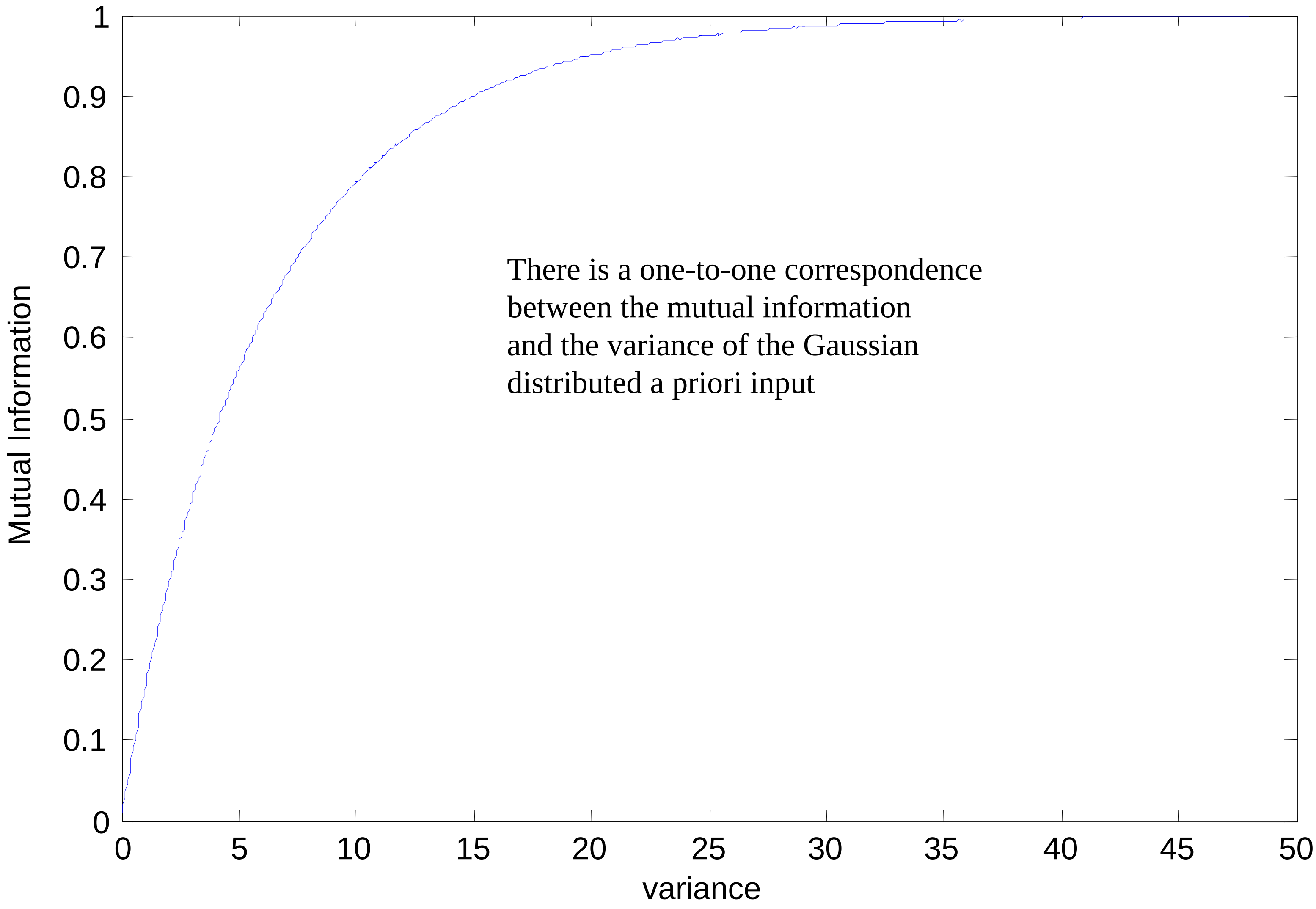
$$\lambda_n = \log \frac{\sum_{\mathbf{x} \in S_n^{(1)}} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x} \mid c_n = 1)}{\sum_{\mathbf{x} \in S_n^{(0)}} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x} \mid c_n = 0)}$$

- This information is obtained from decoder feedback.
  - Bit Interleaved Coded Modulation with Iterative Demodulation
  - Li and Ritcey 1999.

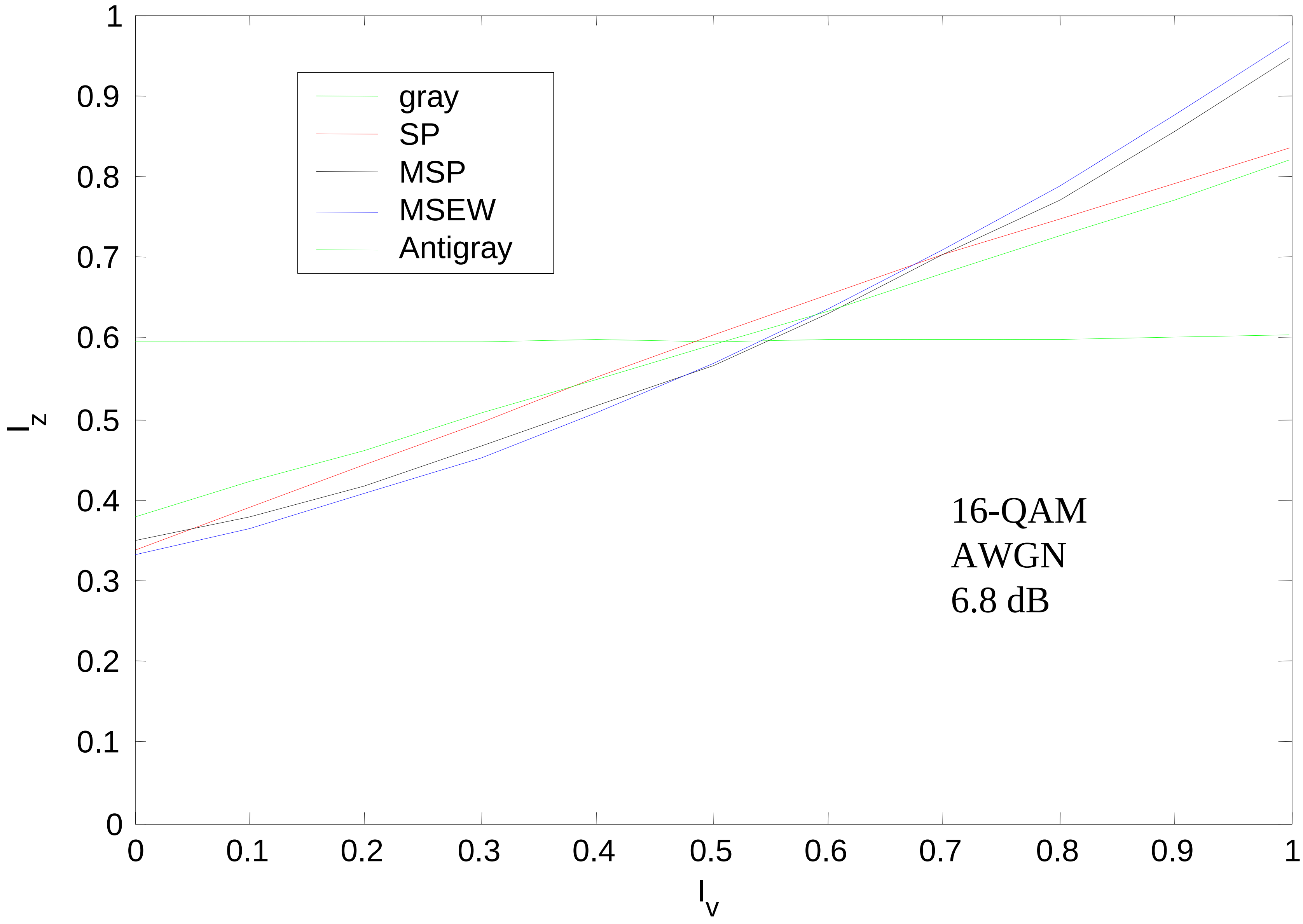
Mutual Information Transfer Chart

- Now consider a receiver that has a priori information about the code bits (from a soft output decoder).
- Assume the following:
  - The a priori information is in LLR form.
  - The a priori LLR's are Gaussian distributed.
  - The LLR's have mutual information I_v
- Then the mutual information I_z at the output of the receiver can be measured through Monte Carlo Integration.
  - I_z vs. I_v is the ***Mutual Information Transfer Characteristic.***
  - ten Brink 1999.

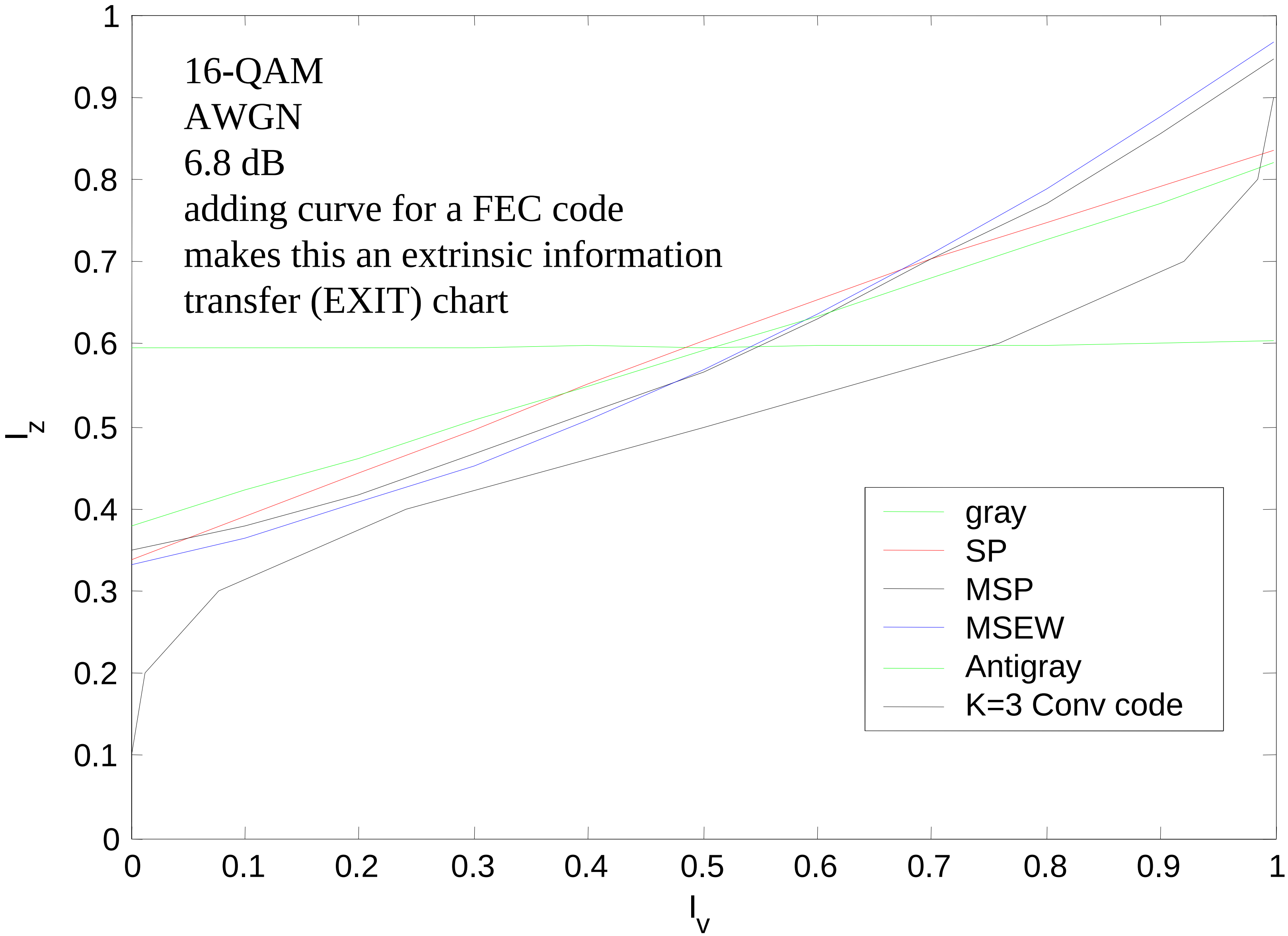
Generating Random a Priori Input



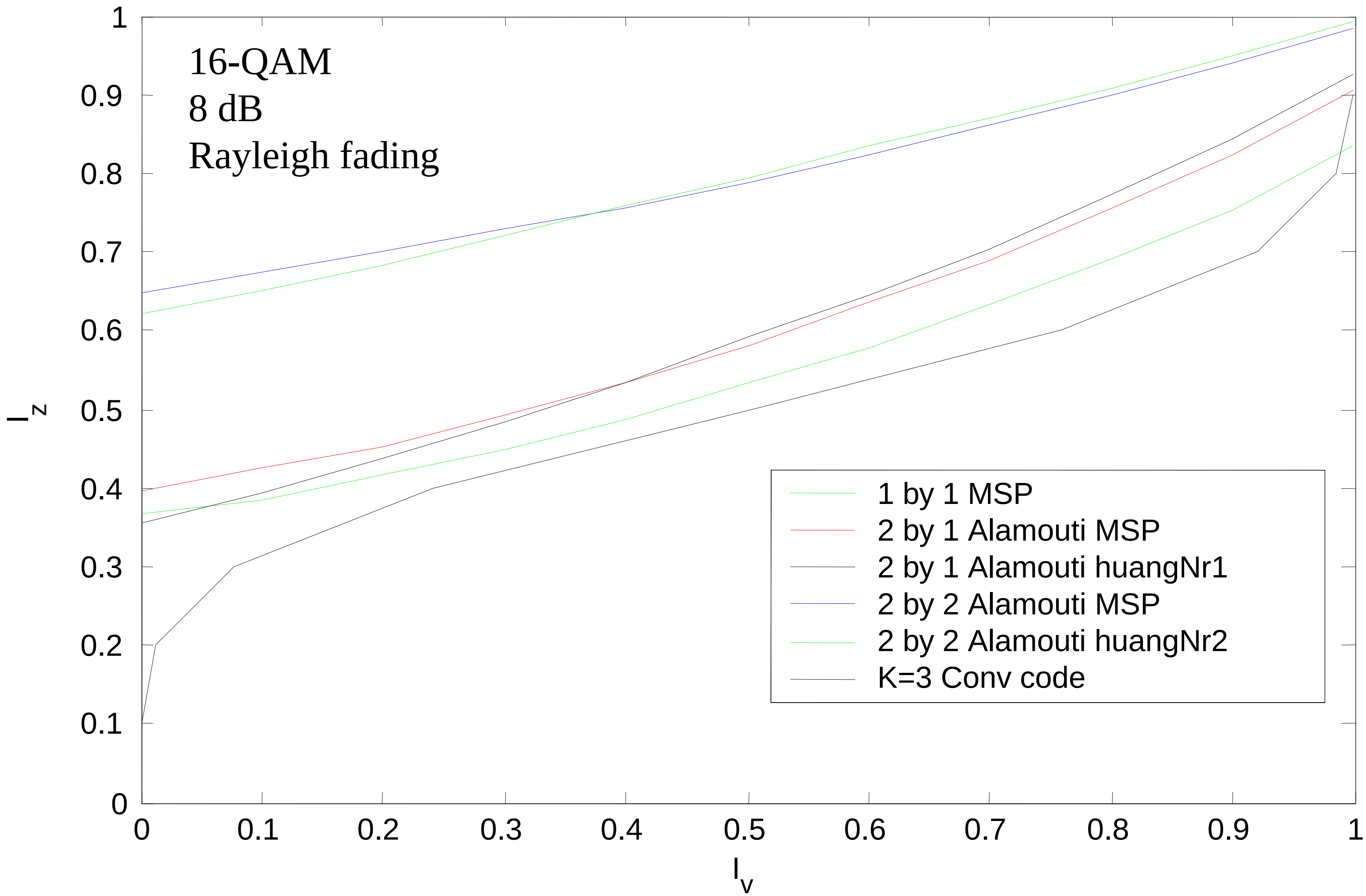
Mutual Information Characteristic



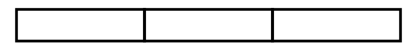
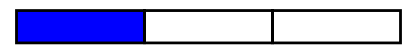
EXIT Chart


16-QAM
AWGN
6.8 dB
adding curve for a FEC code
makes this an extrinsic information
transfer (EXIT) chart
gray
SP
MSP
MSEW
Antigray
K=3 Conv code
I_z
I_v
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

EXIT Chart for Space Time Block Code



Extensions to the MAC Layer

- Hybrid-ARQ
  - Encode data into a low-rate R_M code
    - Implemented using rate-compatible puncturing.
  - Break the codeword into M distinct blocks 
    - Each block has rate $R = R_M/M$
  - Source begins by sending the first block. 
  - If destination does not signal with an ACK, the next block is sent.
    - After mth transmission, effective rate is $R_m = R/m$ 
  - This continues until either the destination decodes the message or all blocks have been transmitted.

Info Theory of Hybrid-ARQ

- Throughput of hybrid-ARQ has been studied by Caire and Tuninetti (IT 2001).
  - Let γ_m denote the received SNR during the m^{th} transmission
    - γ_m is a random.
  - Let $C(\gamma_m)$ be the capacity of the channel with SNR γ_m
    - $C(\gamma_m)$ is also random.
  - The capacity after m blocks have been transmitted is:

$$C_m = \sum_m C(\gamma_m)$$

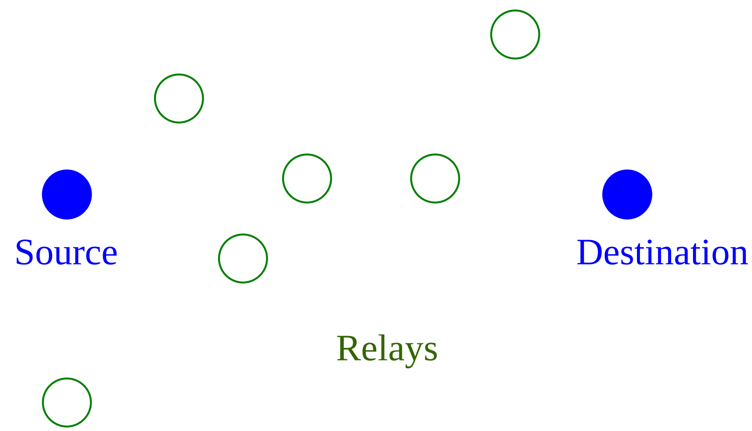
    - This is because the capacity of parallel Gaussian channels adds.
  - An ***outage*** occurs after the m^{th} block if

$$C_m < R$$

  - ***Throughput*** and ***delay*** depend on the average number of blocks required to get out of an outage.

Extensions to the Network Layer

- Now consider the following ad hoc network:

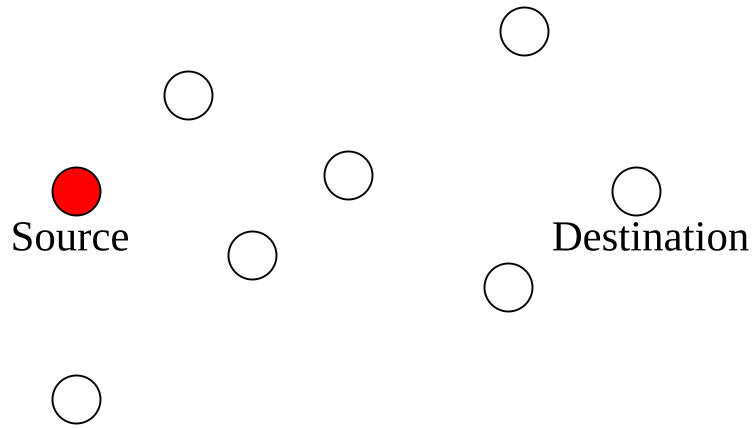


- We can generalize the concept of hybrid-ARQ
  - The retransmission could be from any relay that decoded the message.
  - In large network, relays form a subset of the network called a ***cluster***.

Generalized Hybrid-ARQ Protocol

- Source broadcasts first packet, m=1.
- Relays that can decode are added to the ***decoding set*** *D.*
  - The source is also in *D*
- The next packet is sent by a node in *D*.
  - The choice of which node depends on the protocol.
  - Geographic-Relaying: Pick the node in *D* closest to destination.
- The process continues until the destination can decode.
- We term this protocol “HARBINGER”
  - Hybrid ARq-Based INtercluster GEographic Relaying.
- Energy-latency tradeoff can be analyzed by generalizing Caire and Tuninetti’s analysis.

HARBINGER: Initialization

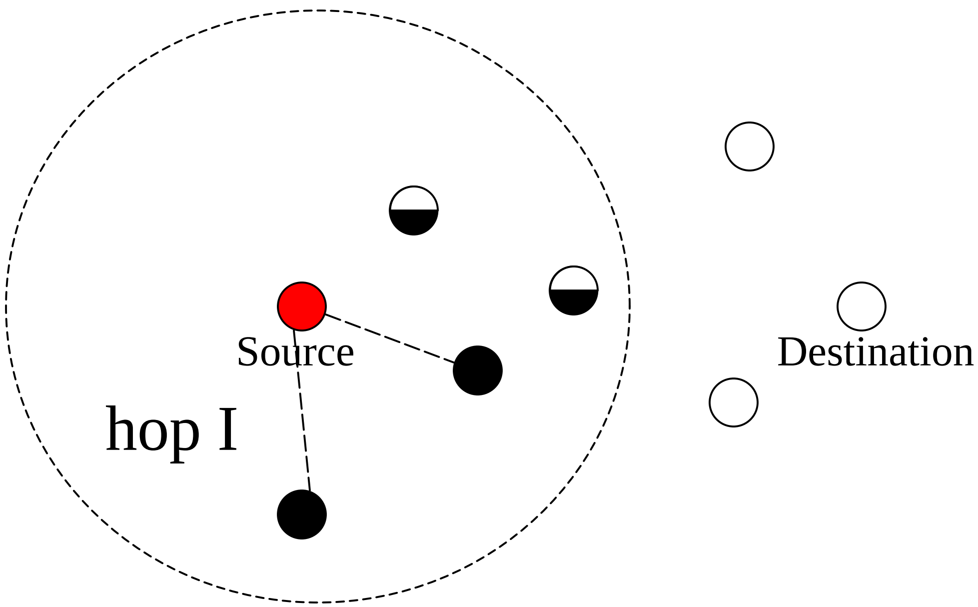


Solid circles are in the decoding set D.

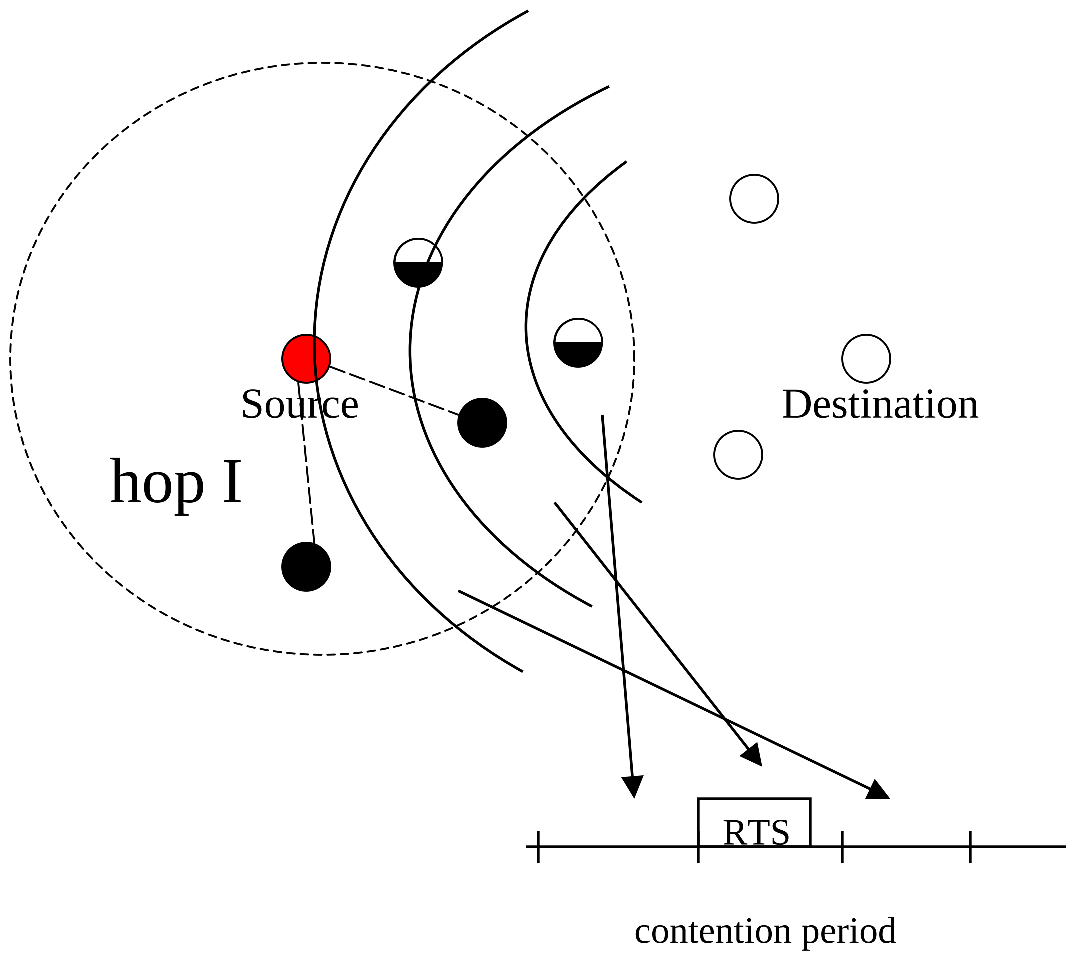
Amount of fill is proportional to the accumulated entropy.

Keep transmitting until Destination is in D.

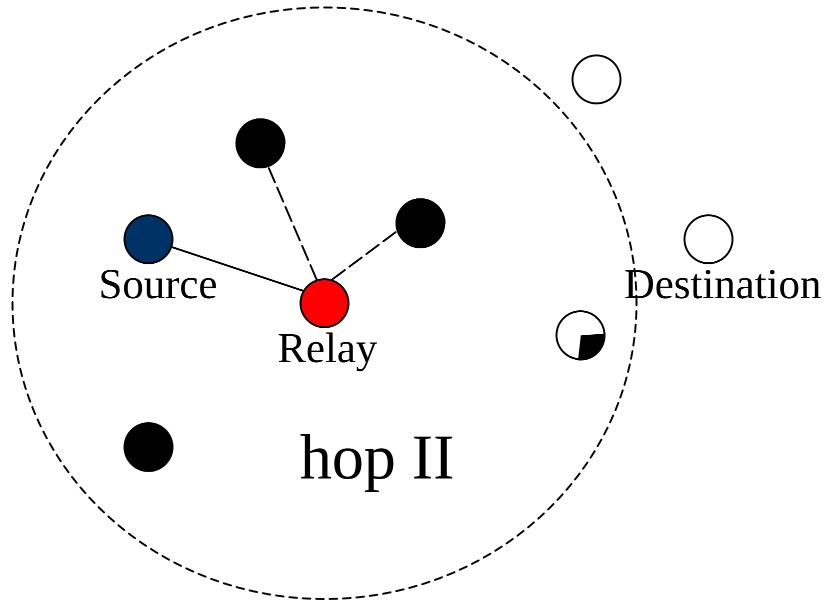
HARBINGER: First Hop



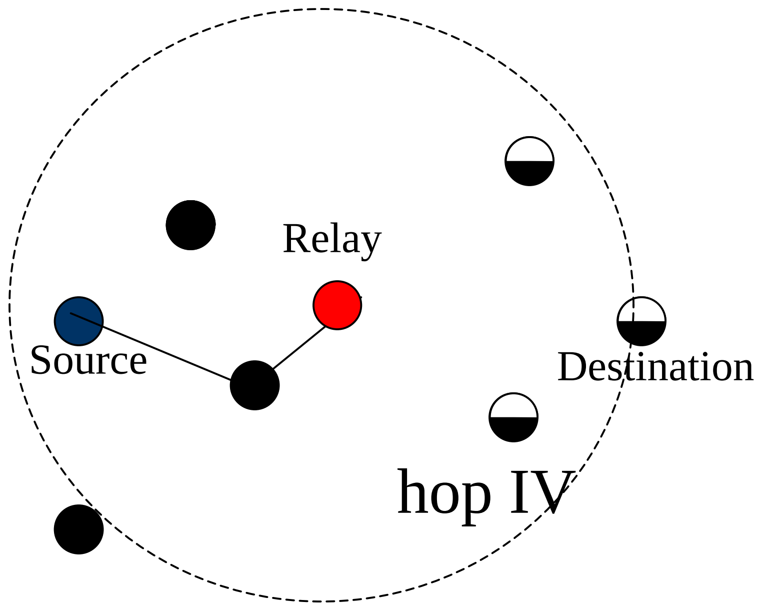
HARBINGER: Selecting the Relay for the Second Hop



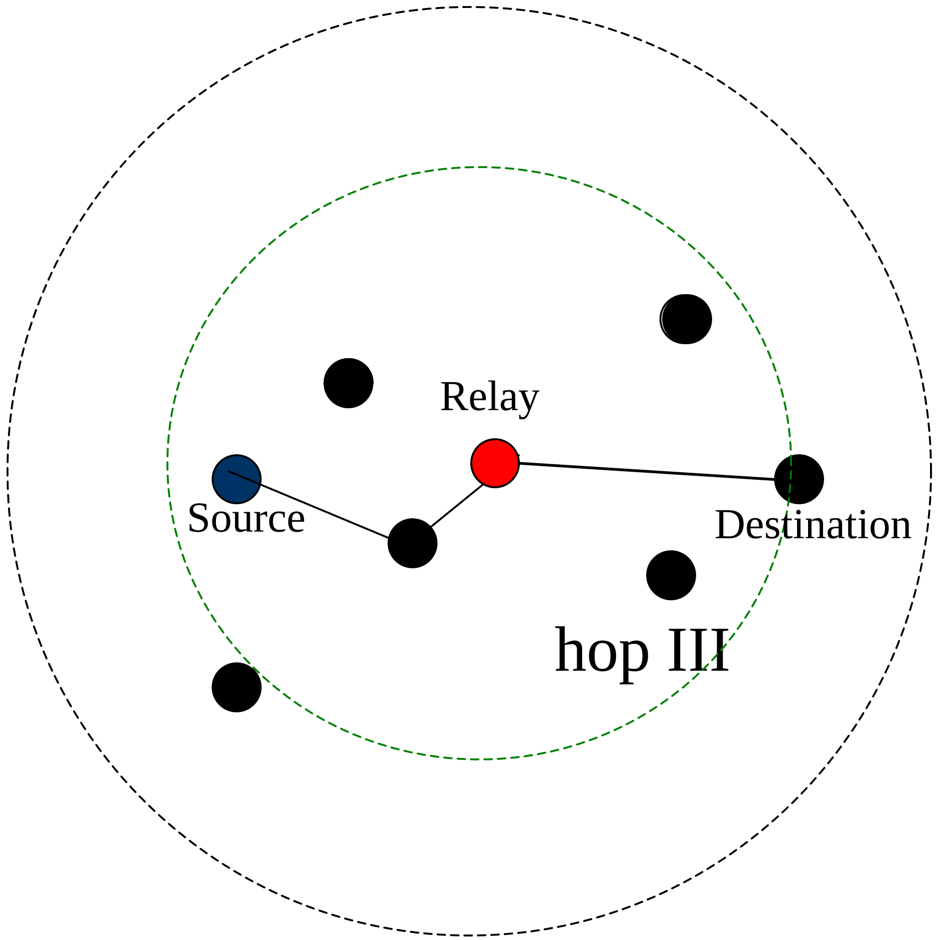
HARBINGER: Second Hop



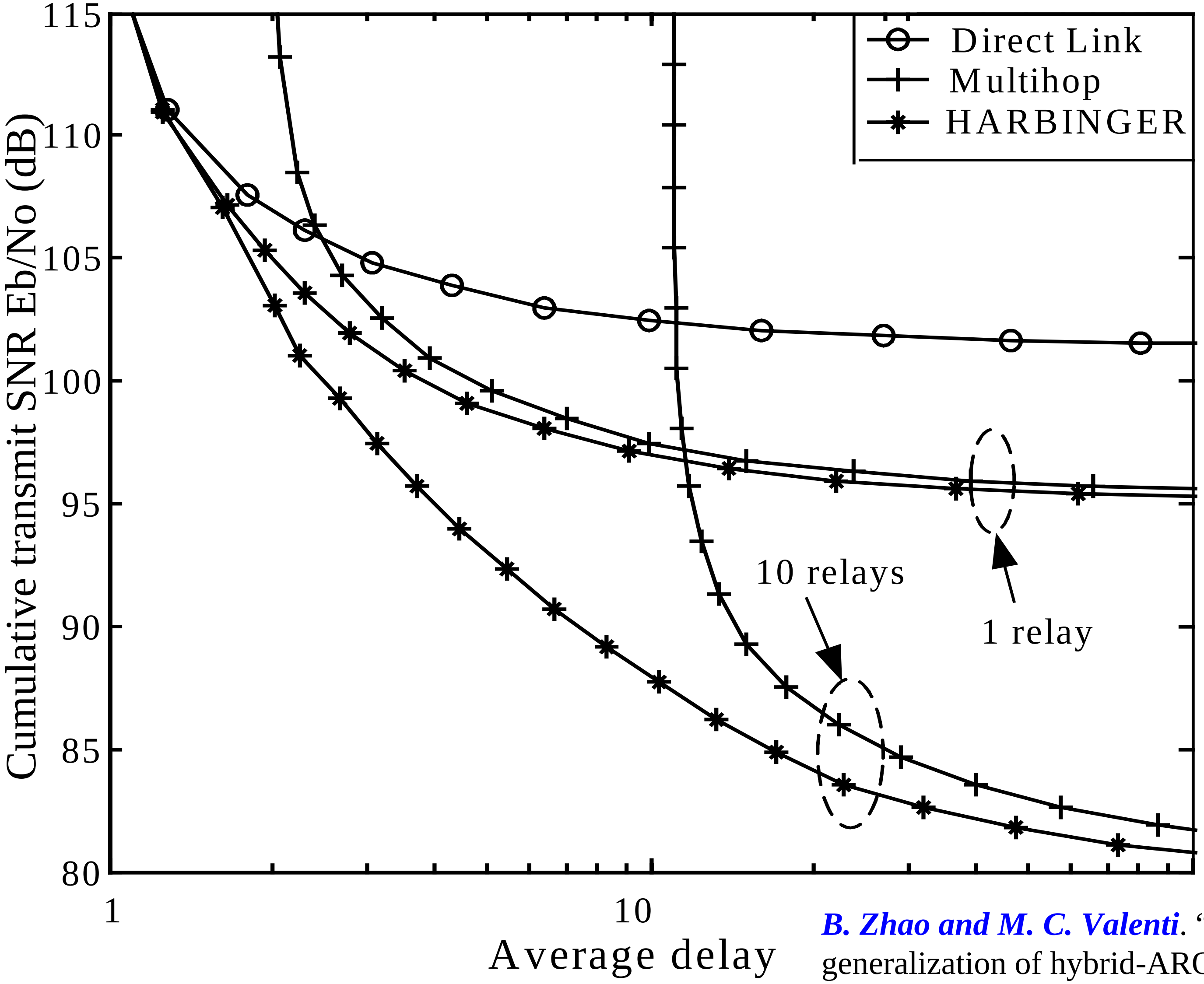
HARBINGER: Third Hop



HARBINGER: Fourth Hop



HARBINGER: Results



Topology:

Relays on straight line

S-D separated by 10 m

Coding parameters:

Per-block rate R=1

No limit on M

Code Combining

Channel parameters:

n = 3 path loss exponent

2.4 GHz

d_0 = 1 m reference dist

Unconstrained modulation

Monte Carlo Integration

B. Zhao and M. C. Valenti. "Practical relay networks: A generalization of hybrid-ARQ," *IEEE JSAC*, Jan. 2005.

Discussion

- Advantages.
  - Better energy-latency tradeoff than multihop.
    - Nodes can transmit with significantly lower energy.
    - System exploits momentarily good links to reduce delay.
  - No need to maintain routing tables (reactive).
- Disadvantages.
  - More receivers must listen to each broadcast.
    - Reception consumes energy.
  - Nodes within a cluster must remain quiet.
  - Longer contention period in the MAC protocol.
  - Results are intractable, must resort to simulation.
  - Requires position estimates.
- These tradeoffs can be balanced by properly selecting the number of relays in a cluster.

Conclusions

- Capacity analysis is a quick way to assess the impact of the modulation choice and channel model.
  - The capacity of complicated systems can be found through Monte Carlo simulation.
- Once a modulation choice is selected for the channel of interest, any off-the-shelf capacity approaching binary code can be used.
  - The interface between demodulator and decoder can be characterized by its EXIT chart.
- Capacity analysis can also be used to characterize:
  - Delay and throughput of retransmission protocols.
  - Performance of multihop routing protocols.