Theory & Strategies for Cooperative Communications

Tutorial at IEEE ISIT 2007

Gerhard Kramer[1] J. Nicholas Laneman[2]

[1]Communications and Statistical Sciences Department, Bell Labs, Alcatel-Lucent

[2]Department of Electrical Engineering, University of Notre Dame

June 24, 2007



Tutorial Goals

- Organize topics on cooperative communications
- Provide intuition for basic concepts
- Connect with recent results
- Encourage research activity
- Contribute to networking practice



Recurring Themes and Take-Aways

1. Many approaches, benefits, and challenges to utilizing relaying and cooperative communications.
2. For a network information theorist, relaying includes, or should include, multihop (store-and-forward routing), network coding, and cooperative communications.
3. The capacity of relay systems is difficult to analyze, but the information theory is surprisingly flexible and diverse.
4. Relays add power/energy which can boost network capacity.
5. Wireless relaying lets one distribute power spatially, reducing path-loss, combating fading, and increasing coverage.
6. Designing wireless relay systems is subtle due to issues such as half-duplex constraints, synchronization, etc.



Caveats

1. We assume familiarity with: entropy, mutual information, capacity of discrete memoryless and additive white Gaussian noise channels, rate-distortion theory
2. References are not provided in the main slides; selected references are given at the end. These, along with the references therein, point to many other recent papers on relaying and cooperative communications.
3. Time constraints limit the scope to less than originally planned.



Tutorial Outline

- Part 1: Introduction and Models (GK)
- Part 2: Protocols (JNL)
- Part 3: Information Rates (GK)
- Part 4: Diversity and Summary (JNL)



Part I

Introduction and Models



Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



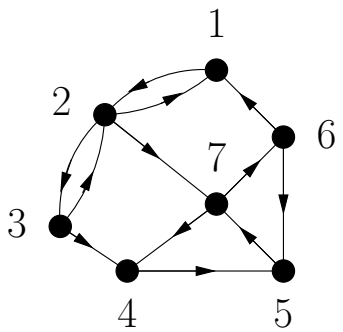
Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



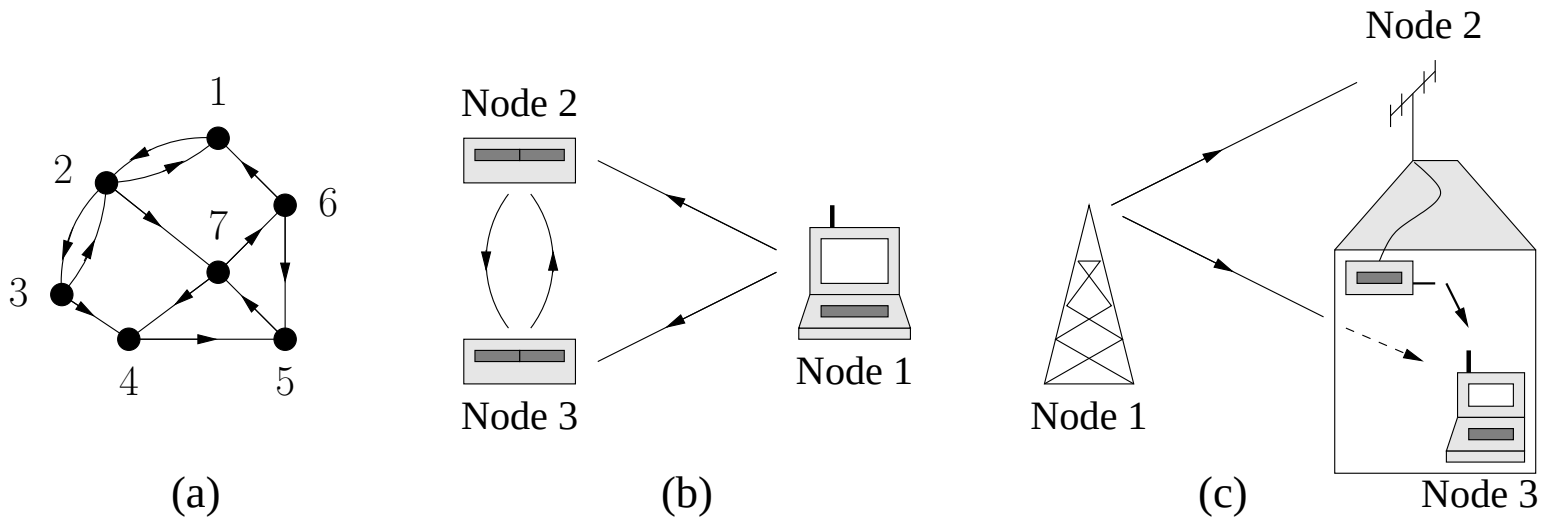
Background

- The purpose of this section is to review networking models and information theory for point-to-point channels.
- A communication network has devices and channels
- One commonly represents a network as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ is a set of nodes and $\mathcal{E}$ is a set of edges (an edge is an ordered pair of nodes).
  - Nodes represent communication devices
  - Edges represent communication channels



Network Examples

- (a) Wireline network; (b) Cellular network with cooperating base stations (mixed wireline/wireless); (c) cellular network with remote antennas; (a) ad hoc wireless network



Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



Device & Channel Properties

Network	Devices (Nodes)	Channels (Edges)
Wireline	Constraints: - processing speed/energy - input/output (ports) - delay Limited Network Knowledge	Independent Channels No Interference Limited Bandwidth Slow Changes Packet Erasures
Wireless	Constraints: - transmit energy - processing speed/energy - half-duplex - delay Limited Network Knowledge Limited Channel Knowledge	Broadcasting Interference Noise Limited Bandwidth Slow or Rapid Changes



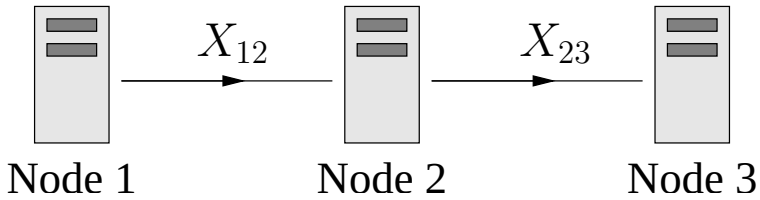
Node Coding/Broadcasting

- Network purpose: enable message exchange between nodes
- Traditional approach:
  - Nodes use packet transmission with store-and-forward
  - Channels treated as point-to-point links
  - Data packets traverse paths (sequences of nodes)
- Other possibilities:
  - Node coding: nodes process
    - Reliable message or packet bits (network coding)
    - Reliable or unreliable symbols (relaying/cooperation)
  - Broadcasting: nodes overhear wireless transmissions
- These concepts already appear in 3-node networks

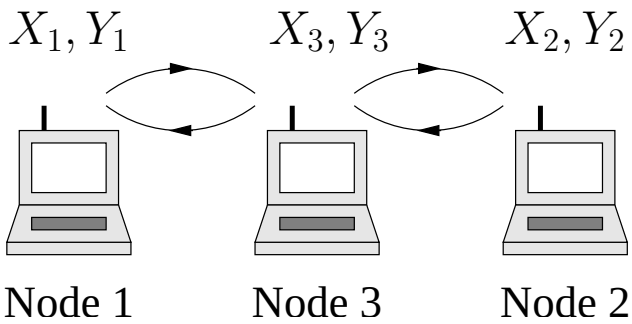


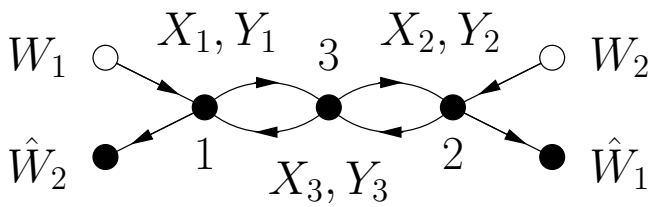
Examples

- Example 1: node 2 sends $X_{23} = a \cdot X_{12}$ (amplify-and-forward)



- Example 2: node 3 decodes binary symbols X_1 and X_2 and broadcasts $X_1 \oplus X_2$



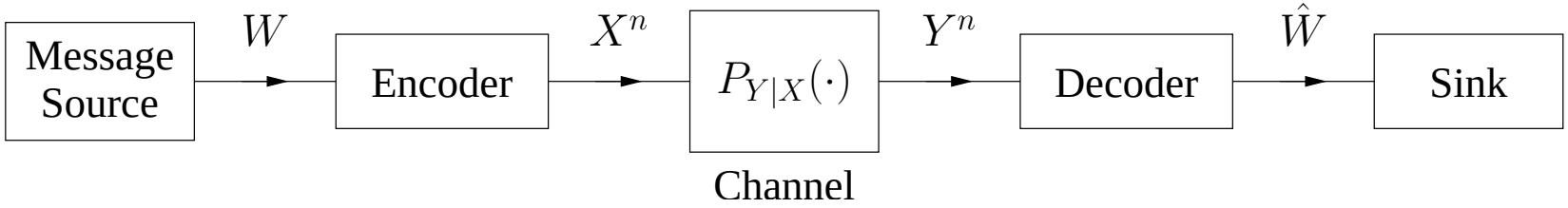


Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



Information & Communications Theory



- Point-to-point channel model:
  - Discrete-time: inputs X_i and outputs Y_i, $i = 1, 2, \ldots, n$
  - Vector notation: $X^n = \underline{X} = X_1, X_2, \ldots, X_n$
  - Memoryless & time-invariant: Y_i generated by X_i via $P_{Y|X}(\cdot)$
- Source message W has B bits; rate is $R = B/n$ bits/use
- Capacity: maximum R at which one can transmit W reliably: $\Pr[\hat{W} \neq W]$ can be made close to zero for large n
- Capacity: $C = \max_{P_X(\cdot)} I(X;Y)$



Coding

- Generate a Random Codebook
  - Choose a probability distribution $P_X(\cdot)$
  - Generate 2^{nR} codewords $\underline{x}(w)$, $w = 1, 2, \ldots, 2^{nR}$, by choosing $x_i(w)$, $i = 1, 2, \ldots, n$, independently using $P_X(\cdot)$
- Encoder and Decoder
  - Encoder transmits $\underline{x}(w)$
  - Decoder sees $\underline{y}$ and puts out the $\hat{w}$ for which $\underline{x}(\hat{w})$ is "closest" to $\underline{y}$, where "closest" is measured in terms of, e.g., maximum-likelihood, least squares, entropy-typicality, etc.
- Error Probability
  - Over the ensemble of all codebooks constructed as above, one can make $\Pr[\hat{W} \neq W] \to 0$ as $n \to \infty$ if $R < I(X;Y)$
  - One can thus find a sequence of codebooks of increasing length n for which $R < I(X;Y)$ and $\Pr[\hat{W} \neq W] \to 0$
- Finally, optimize over $P_X(\cdot)$



AWGN Channel

- Additive white Gaussian noise (AWGN) channel:

$$Y = X + Z$$

  - where X, Y, and Z are complex random variables
  - Z is Gaussian, indep. real/imag. parts, each with variance $N/2$
  - Transmit power constraint: $\sum_{i=1}^{n} \mathsf{E}[|X_i|^2]/n \leq P$

- Capacity: $C = \log_2(1+\gamma)$
  where $\gamma = P/N$ is the signal-to-noise ratio (SNR)
- Capacity achieved with Gaussian $P_X(\cdot)$ where X has independent real and imaginary parts, each with variance $P/2$
- Modulation: for practical reasons, use discrete X
  Example: M-ary phase-shift keying (PSK) with

$$X = \sqrt{P}e^{j2\pi m/M},\ m = 0, 1, \ldots, M-1$$

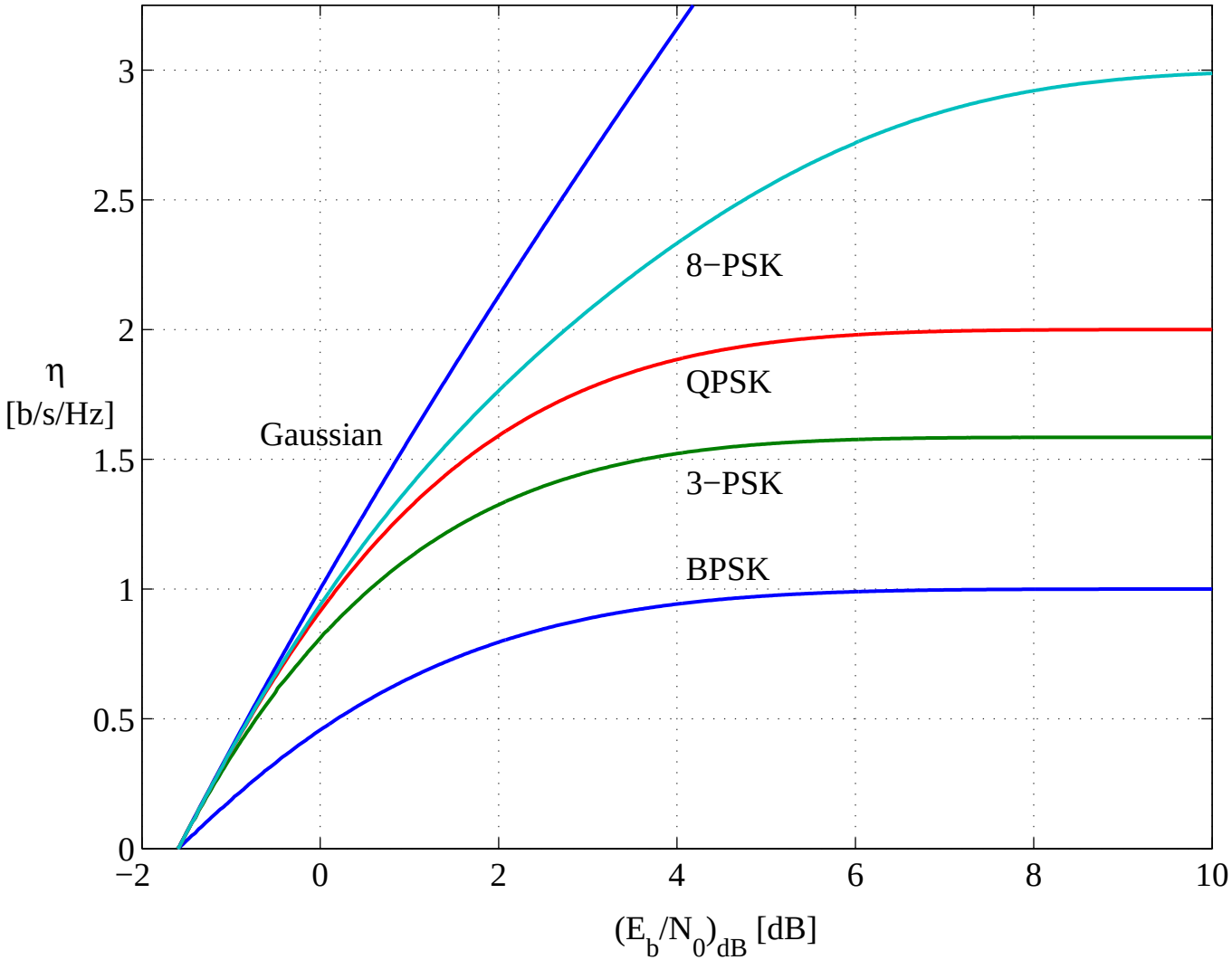


Spectral Efficiency

- Spectral efficiency: include effects of bandwidth
  - Channel passes the frequencies
    $$f_0 - W/2 < f < f_0 + W/2 \text{ Hertz}$$
    where the center frequency f_0 is much larger than W
  - Use sampling theory to represent filtered signals by n samples spaced $T_s = 1/W$ seconds apart.
  - Noise power is proportional to bandwidth: $N = N_0 W$ where N_0 is the noise power per Hz
- Transmit energy constraint: $\sum_{i=1}^{n} \mathsf{E}[|X_i|^2]/n \leq PT_s$
- Energy per information bit: $E_b = PT_s \cdot (1/R)$
- Noise sample energy or variance: NT_s
- Spectral efficiency: the maximum reliable rate R in bits/sec/Hz at a specified $E_b/N_0 = (P/N)/R$



PSK Spectral Efficiency



Outline: Part 1

1 Introduction
- Background
- Device & Channel Properties
- Information & Communications Theory

2 Models
- Wireline Models
- Wireless Models
- Discrete Memoryless Network Model



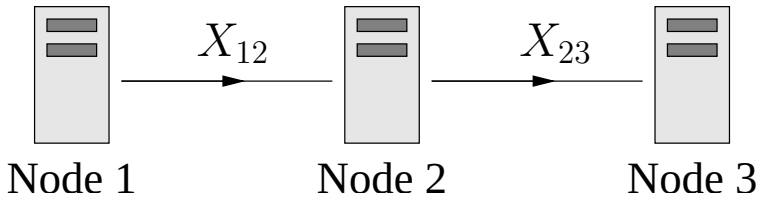
Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model

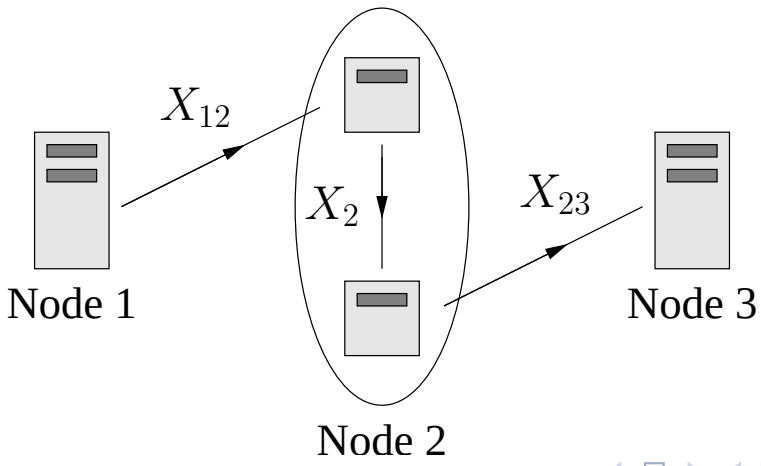


Wireline Models

- The purpose of this section is to develop channel and device models that apply to both wireline and wireless networks.
- Wireline networks: edges (u, v) have capacity C_{uv}
- Model edge variables as having alphabet sizes $2^{C_{uv}}$

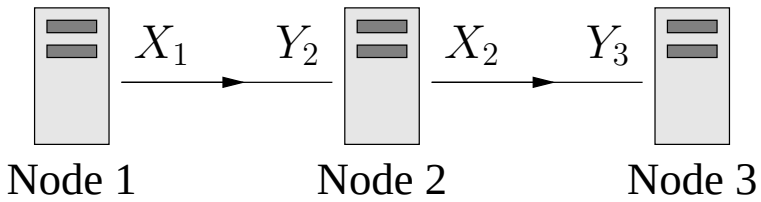


- A node constraint: node 2 has capacity C_2



Node Port Constraints

- Suppose node 2 cannot receive and transmit simultaneously
- A simple model: assign each device one input variable X_u and one output variable Y_u. Then define appropriate "channels".
- Example:



$$\text{if } X_2 = 0 \text{ then } Y_2 = X_1$$
$$\text{if } X_2 \neq 0 \text{ then } Y_2 = 0$$
$$Y_3 = X_2$$

- Wireline port constraints are the counterpart of wireless half-duplex constraints



Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



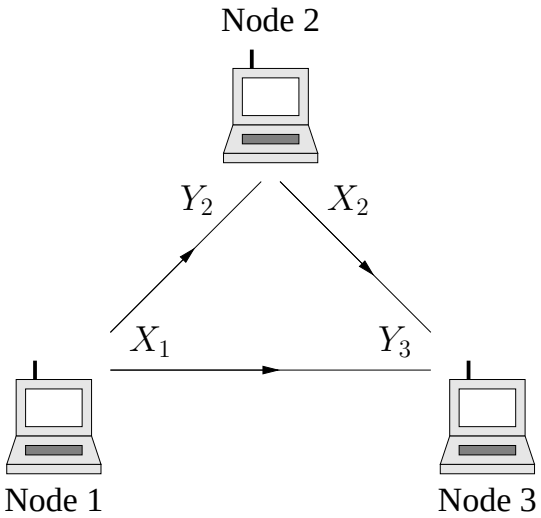
Wireless Models

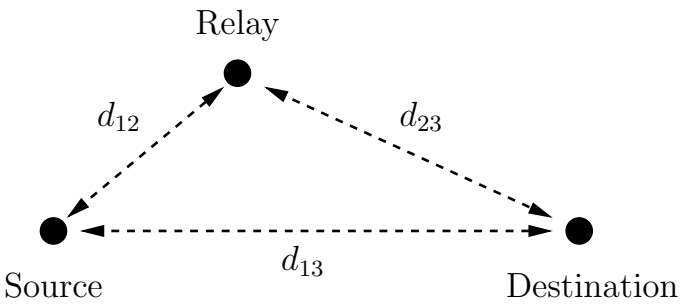
- A half-duplex relay channel:

$$Y_2 = \begin{cases} \frac{h_{12}}{d_{12}^{\alpha/2}} X_1 + Z_2 & \text{if } X_2 = 0 \\ 0 & \text{if } X_2 \neq 0 \end{cases}$$

$$Y_3 = \frac{h_{13}}{d_{13}^{\alpha/2}} X_1 + \frac{h_{23}}{d_{23}^{\alpha/2}} X_2 + Z_3$$

  - h_{uv} are fading gains
  - d_{uv} are distances
  - α is an attenuation exponent





Wireless Channel Models

- Fading gains h_{uv} vary due to
  - Spatial variations: electromagnetic scattering and absorption
  - Temporal variations: node mobility, temperature, humidity
- Space-time models for the h_{uv}:
  - Deterministic: electromagnetic wave propagation equations
  - Random: $\{h_{uv,i}\}_{i=1}^{n}$ is a realization of an integer-time stochastic process $\{H_{uv,i}\}_{i=1}^{n}$
- The random model admits uncertainty and is simpler. Issues:
  - Marginal distributions of the $H_{uv,i}$
  - Temporal correlation: joint distribution of $\{H_{uv,i}\}_{i=1}^{n}$



Fast and Slow Fading

- Marginal distributions:
  - Assume the $H_{uv,i}$, $i = 1, 2, \ldots, n$, have the same marginal distribution H_{uv} during a given communication session
  - No fading: H_{uv} is a known constant
  - Rayleigh fading: H_{uv} is complex, Gaussian, 0-mean, unit var.
- Temporal correlation: we study two extremes
  - Fast fading: $h_{uv,i}$ are independent realizations of H_{uv}
  - Slow fading: $H_{uv,i} = H_{uv}$, all i, where H_{uv} is drawn once
- A channel (edge) is fast fading if each packet encounters many channel realizations. A channel is slow fading if a packet encounters one channel realization.
- Channel type depends on: geography, device velocity, center frequency, bandwidth, delay constraints, application, etc.



Wireless Device Models

- Power and energy constraints:
  - Block constraints: $\sum_{i=1}^{n} \mathsf{E}[|X_i|^2]/n \leq P$
  - Per-symbol constraints: $\mathsf{E}[|X_i|^2] \leq P$, all i
- Half-duplex constraint: include in channel model (see below)
- Channel state information (CSI) at the transmittter (CSIT) and/or receiver (CSIT):
  - No CSIT, No CSIR: optimizing $P_X(\cdot)$ is tricky, capacity results are sensitive to model parameters
  - CSIR, No CSIT: augment Y_v with $\{H_{uv} : \text{all } u\}$
  - CSIT, No CSIR: transmitter has extra information, e.g., when broadcasting multiple message streams simultaneously
  - CSIT, CSIR: capacity-achieving $P_X(\cdot)$ includes (more) power allocation, e.g., waterfilling



Outline: Part 1

1. Introduction
   - Background
   - Device & Channel Properties
   - Information & Communications Theory
2. Models
   - Wireline Models
   - Wireless Models
   - Discrete Memoryless Network Model



Discrete Memoryless Network Model

- For wireless and other networks, reinterpret the network graph
- Device model:
  - Node u has one input variable X_u and one output variable Y_u
  - Causality: $X_{u,i} = f_{u,i}(\text{local messages}, Y_u^{i-1})$, $i = 1, 2, \ldots, n$
  - Cost constraint example: $\mathsf{E}[f(X_{1,i}, X_{2,i}, Y_{2,i}, Y_{3,i})] \leq S$, all i
- Channel model:
  - A network clock governs operations: node u transmits $X_{u,i}$ between clock tick $(i-1)$ and tick i, and receives $Y_{u,i}$ at tick i
  - Memoryless: $Y_{u,i}$ generated by the $X_{u,i}$, all u, via the channel

$$P_{Y_1 Y_2 Y_3 \ldots | X_1 X_2 X_3 \ldots}(\cdot)$$

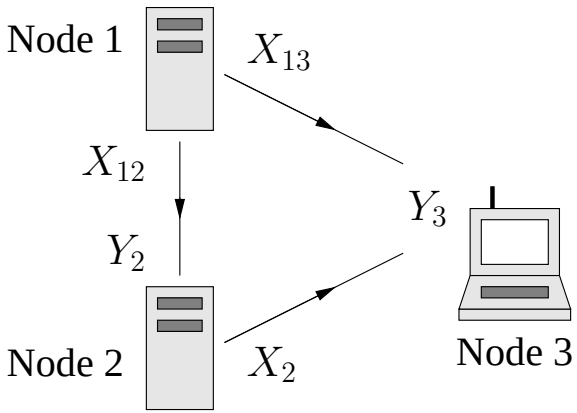
  - Delay modeled as due to device processing, not the channel
  - Device port/half-duplex constraints are part of the channel

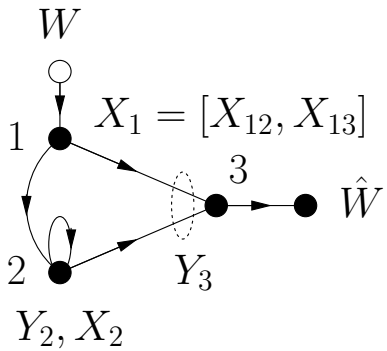


Network Graph

- Graph: draw a directed edge from node u to node v if X_u contributes to Y_v
- Example: a mixed wireline/wireless network where node 2 is half-duplex. The network channel distribution factors as

$$P_{Y_2|X_{12},X_2}(a|b,c) \cdot P_{Y_3|X_{13},X_2}(d|e,f)$$





Source and Destination Models

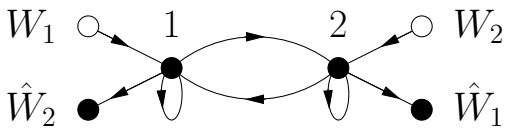
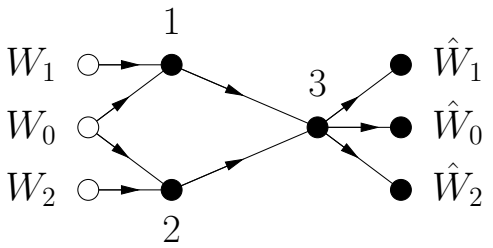
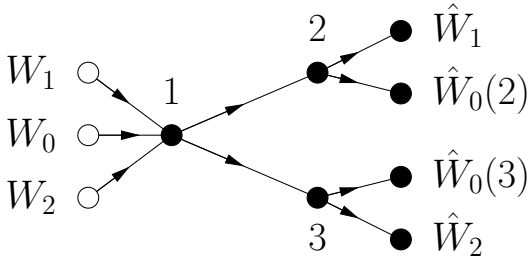
- There are M source messages W_m, $m = 1, 2, \ldots, M$
- Every node has multiple messages
- Message W_m estimated as $\hat{W}_m(u)$ at certain nodes u
- Source model:
  - Messages are statistically independent
  - Messages are fully-compressed bit strings
  - Sources are not bursty
  - Delay is not critical
- The last two assumptions are unrealistic, but they let us focus on specific issues.
- Capacity region: closure of set of $(R_1, R_2, \ldots, R_M)$ for which

$$\Pr\left[\bigcup_{m,u}\left\{\hat{W}_m(u) \neq W_m\right\}\right]$$

can be made close to zero for large n



Basic Networks

Network	Device Nodes	Channel Edges	Sources	Sinks	Graph
Point−to−Point Channel (DMC & AWGN Channel)	2	1	1	1	
Two−way Channel (2WC)	2	4	2	2	
Multiaccess Channel (MAC)	3	2	3	3	
Broadcast Channel (BC)	3	2	3	4	

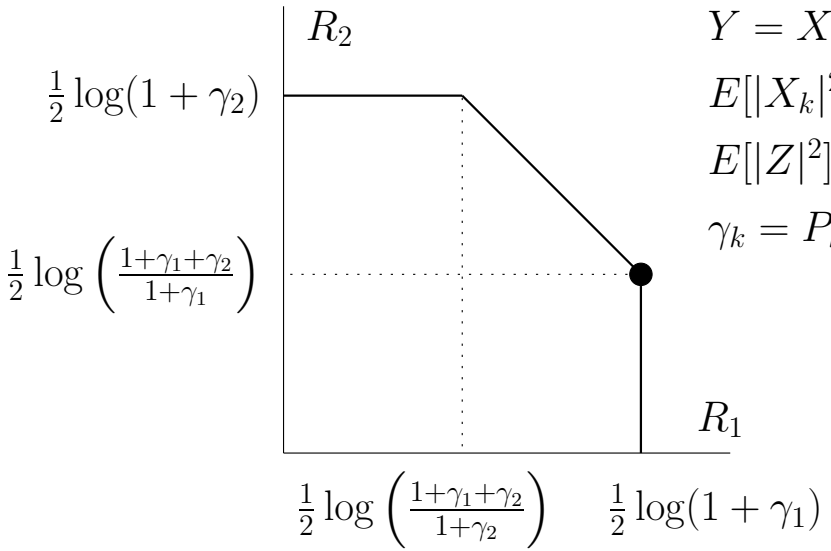


- Consider the MAC with $R_0 = 0$. The capacity region $\mathcal{C}$ is

$$\bigcup \left\{ (R_1, R_2): \begin{array}{l} 0 \le R_1 \le I(X_1; Y|X_2U) \\ 0 \le R_1 \le I(X_2; Y|X_1U) \\ R_1 + R_2 \le I(X_1X_2; Y|U) \end{array} \right\}$$

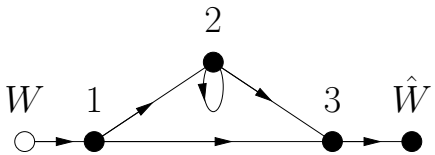
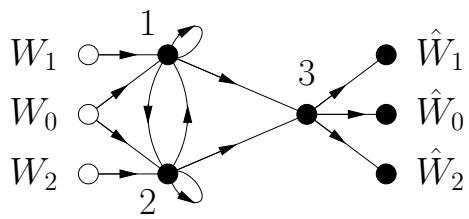
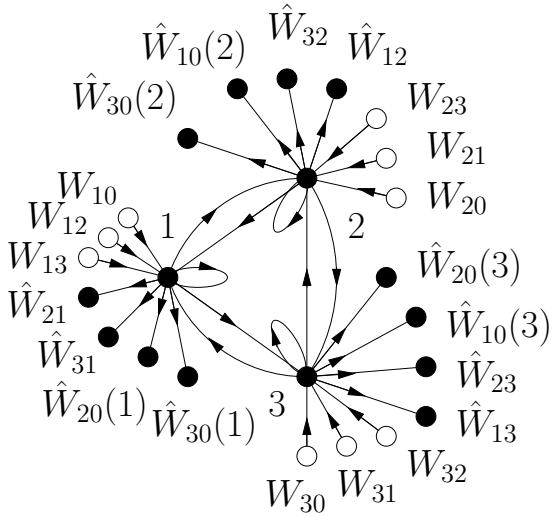
  where the union is over all $P_U(\cdot)$, $P_{X_1|U}(\cdot)$, $P_{X_2|U}(\cdot)$

- $\mathcal{C}$ for the AWGN channel is shown below.
  Example: "decode" $\underline{x}_2$ while treating $\underline{x}_1$ as interference, then decode $\underline{x}_1$ (single-user decoding with interference cancellation)

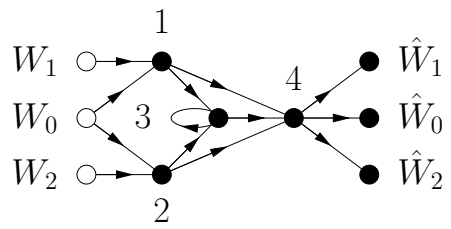
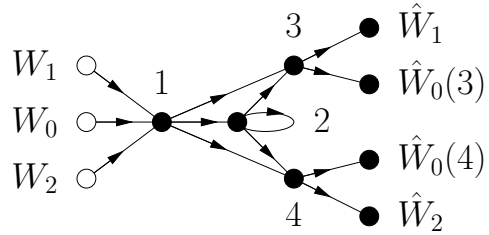
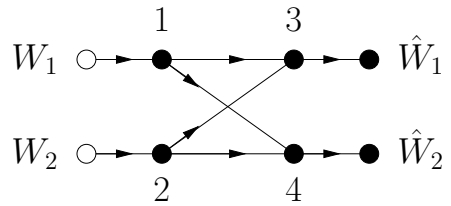
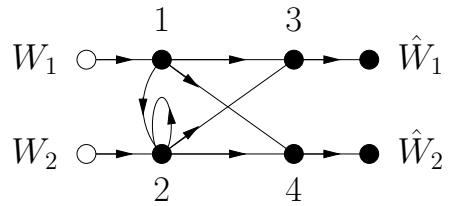


$$Y = X_1 + X_2 + Z$$
$$E[|X_k|^2] \le P_k, \ k = 1, 2$$
$$E[|Z|^2] = N$$
$$\gamma_k = P_k/N, \ k = 1, 2$$

Networks with Three Device Nodes

Network	Device Nodes	Channel Edges	Sources	Sinks	Graph
Relay Channel (RC)	3	4	1	1	
MAC with Generalized Feedback (MAC−GF)	3	6	3	3	
Three−way Channel (3WC)	3	9	9	12	

Networks with Four Device Nodes

Network	Device Nodes	Channel Edges	Sources	Sinks	Graph
MAC with a Dedicated Relay (MAC−DR)	4	6	3	3	
BC with a Dedicated Relay (BC−DR)	4	6	3	4	
Interference Channel (IC)	4	4	2	2	
Cognitive Radio Channel	4	6	2	2	

Summary

What we covered in Part 1:

1. Reviewed basic information theory
2. Introduced simple device and channel models
3. Introduced a network model with a memoryless channel
4. Reviewed network information theory channels



Part II

Protocols



Outline: Part 2

3. Protocols
   - Foundations
   - Amplify-and-Forward (AF)
   - Decode-and-Forward (DF)
   - Compress-and-Forward (CF)



Protocols

- The purpose of this section is to give a high-level overview of protocols for relaying and cooperative communications; the next section will provide more details
- A *protocol* is the sequence of interactions involved in communicating messages among nodes in a network
- Focus mainly on the simplest setting of a relay channel, with immediate generalizations
- Protocols also called "strategies", "schemes", and so forth

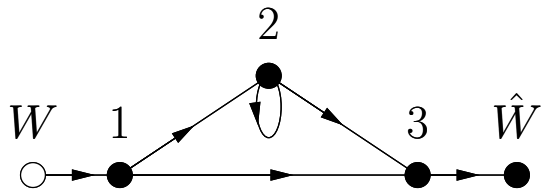


Outline: Part 2

3 Protocols
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)



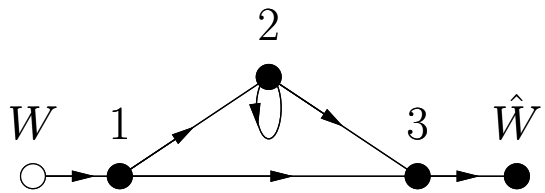
Basic Ideas



- Analogies
  - Professor lectures; TA recitations; Students learn
  - Baby cries; A parent hears; Either cries or elbows the other parent to take a turn
  - Tutorial presentation; Attendees whisper to reconcile understanding
- Key Ingredients
  - Multiple transmitters (Nodes 1 and 2)
  - Multiple receivers (Nodes 2 and 3)
  - Reconciliation, or combining, at Node 3



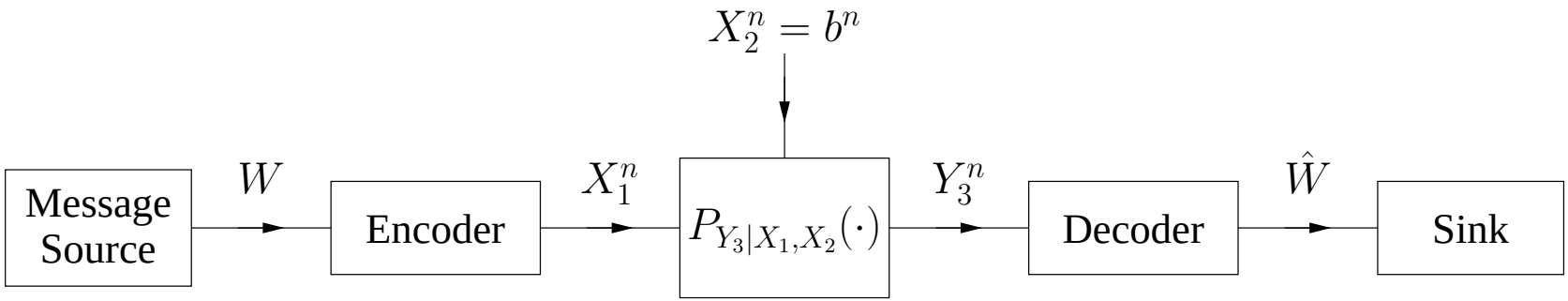
Building Blocks



- Direct transmission from Node 1 to Node 3
- Multihop transmission through Node 2
  - Amplify-and-Forward (AF)
  - Decode-and-Forward (DF)
  - Compress-and-Forward (CF)
- Broadcast transmission from Node 1 to Nodes 2 and 3
- Multiaccess transmission from Nodes 1 and 2 to Node 3
- Combinations, superpositions, mode coding, feedback, etc.



Direct Transmission



- Relay does not participate, i.e., $X_{2,i} = b$ (often 0) for all i
- A standard random coding argument shows that rates

$$R < R_{\text{dir}} := \max_{P_{X_1|X_2}(\cdot|b),b} I(X_1; Y_3|X_2 = b)$$

  are achievable
- R_{dir} is in fact capacity if the relay channel is *reversely degraded*, i.e.,

$$P_{Y_2Y_3|X_1X_2}(\cdot) = P_{Y_3|X_1X_2}(\cdot)P_{Y_2|Y_3}(\cdot)$$

  (Think: Y_2 is a noisy version of Y_3)



Multihop Transmission

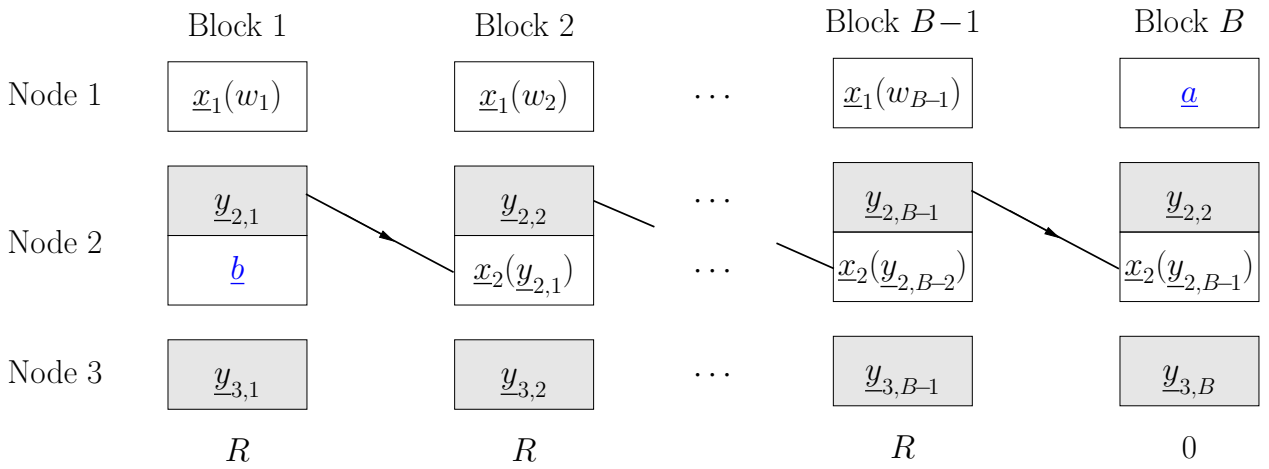
- Node 1 transmits to Node 2, and Node 2 transmits to Node 3
- Motivated by a cascade of two channels, i.e.,

$$P_{Y_2^n Y_3^n|X_1^n X_2^n}(\cdot) = P_{Y_2^n|X_1^n}(\cdot) P_{Y_3^n|X_2^n}(\cdot)$$

- Question: What should Node 2 transmit?
  - Non-Regenerative / Amplify-and-Forward (AF)
    Node 2 sets $X_{2,i} = \beta Y_{2,i-k}$, for some $k \geq 1$
  - Compress-and-Forward (CF) / Estimate-Forward (EF)
    Node 2 conveys a quantized version of Y_2^n to Node 3
  - Regenerative / Decode-and-Forward (DF)
    Node 2 decodes W and re-encodes into a codebook X_2^n
- Transmissions usually occur in a pipeline of two (or more) blocks (potentially of varying sizes)



Block Pipeline



- B blocks, each of length n channel uses
  - Block 1 fills the pipeline
  - Block B empties the pipeline
  - Blocks 2 through B can blend broadcast and multiaccess
- Extremes
  - $n=1$, B large: Like an intersymbol interference (ISI) channel
  - n large, $B=2$: No interblock interference, half-duplex
  - n, B large: Effective rate $\frac{(B-1)}{B}R \to R$
- Could consider memory within and among input blocks,

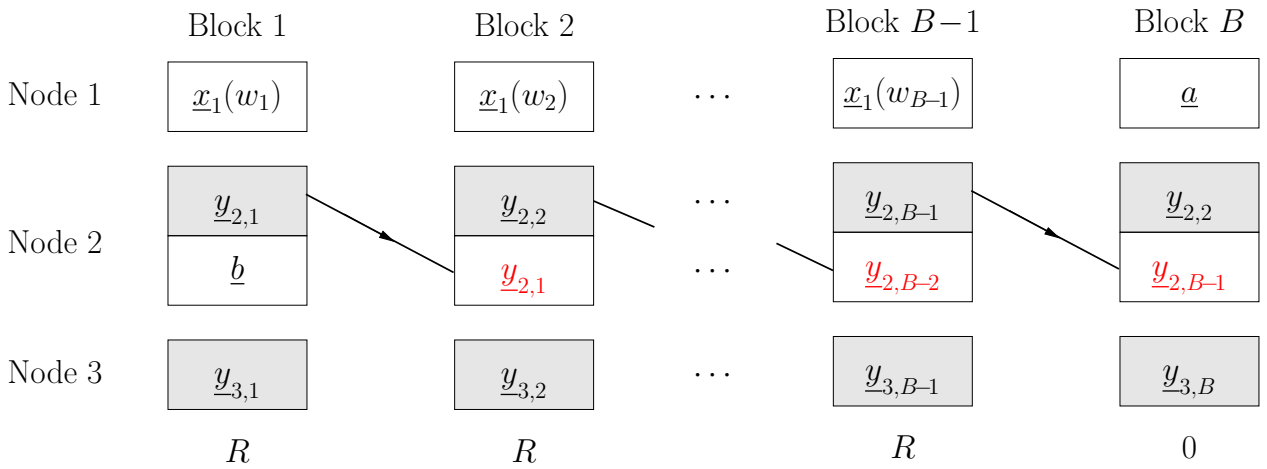


Outline: Part 2

3 Protocols
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)



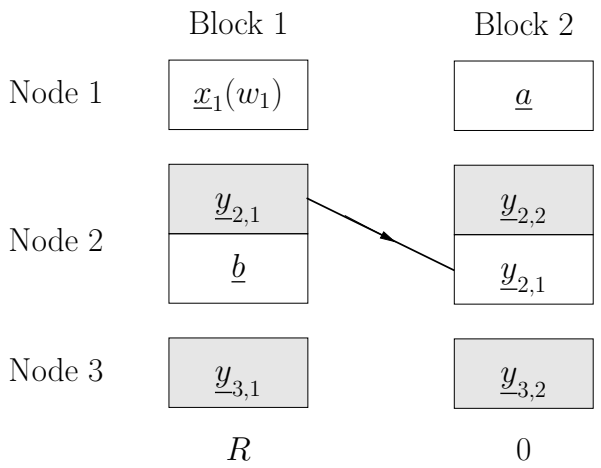
Amplify-and-Forward (AF)



- Discrete Channels
  - Often $\underline{x}_{2,i} = \underline{y}_{2,i-1}$
  - Requires $\mathcal{Y}_2 \subseteq \mathcal{X}_2$
- Continuous Channels
  - Often $\underline{x}_{2,i} = \beta \underline{y}_{2,i-1}$
  - β chosen subject to a power constraint



Multihop AF (MAF), $B=2$



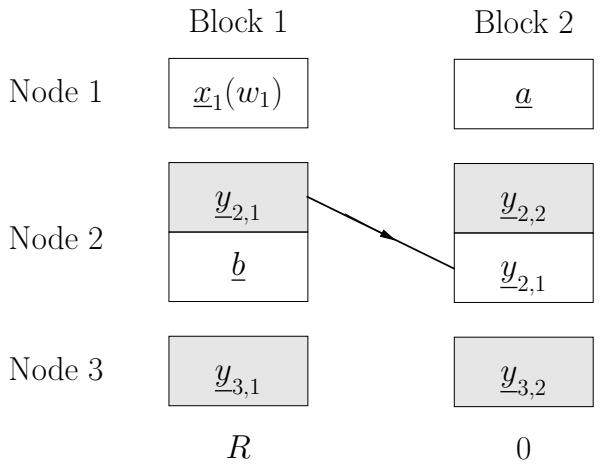
- If Node 3 decodes using only Block 2, R achievable if

$$R < R_{\text{maf}} := \max_{P_{X_{1,1}|X_{2,1}}(\cdot|b),a,b} \frac{1}{2} I(X_{1,1}; Y_{3,2}|X_{2,1}=b, X_{1,2}=a)$$

- $I(X_{1,1}; Y_{3,2}|X_{2,1}=b, X_{1,2}=a)$ computed for the effective channel $P_{Y_{3,2}|X_{1,1}X_{1,2}X_{2,1}}(\cdot|\cdot, a, b)$ in which $X_{2,2} = Y_{2,1}$
- Also known as "non-regenerative repeating"



Diversity AF (DAF), $B=2$



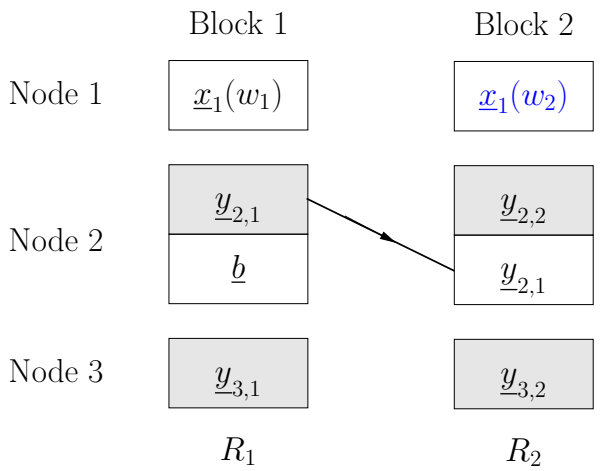
- If Node 3 decodes using both Blocks 1 and 2, then R achievable for

$$R < R_{\text{daf}} := \max_{P_{X_{1,1}|X_{2,1}}(\cdot|b),a,b} \frac{1}{2} I(X_{1,1}; Y_{3,1}Y_{3,2}|X_{2,1}=b, X_{1,2}=a)$$

- Similar to repetition coding, except $X_{2,2} = Y_{2,1}$ is a corrupted version of $X_{1,1}$



Non-Orthogonal DAF (NDAF), $B=2$



- If Node 1 sends new information in Block 2, and Node 3 decodes using both Blocks 1 and 2, then R achievable for

$$R < R_{\text{ndaf}} := \max \frac{1}{2} I(X_{1,1}X_{1,2}; Y_{3,1}Y_{3,2}|X_{2,1}=b)$$

  with max over $P_{X_{1,1}|X_{2,1}}(\cdot|b)P_{X_{1,2}|X_{2,1}}(\cdot|b)$ and b
- A combination of DAF and Direct from Node 1 to Node 3



Intersymbol Interference AF (IAF)

- Pipeline with $n = 1$, $B \to \infty$ creates an effective intersymbol interference (ISI) channel
- Input memory important in this case ("waterfilling")
- Additional improvements with "bursty AF", mainly at low SNR



AF Remarks

- Generally

$$R_{\text{maf}} < R_{\text{daf}} < R_{\text{ndaf}} < R_{\text{iaf}}$$

  but coding and decoding grows increasingly complex
- MAF, DAF, and NDAF with $B = 2$ are useful for half-duplex systems
- IAF with $B \to \infty$ is useful for full-duplex systems

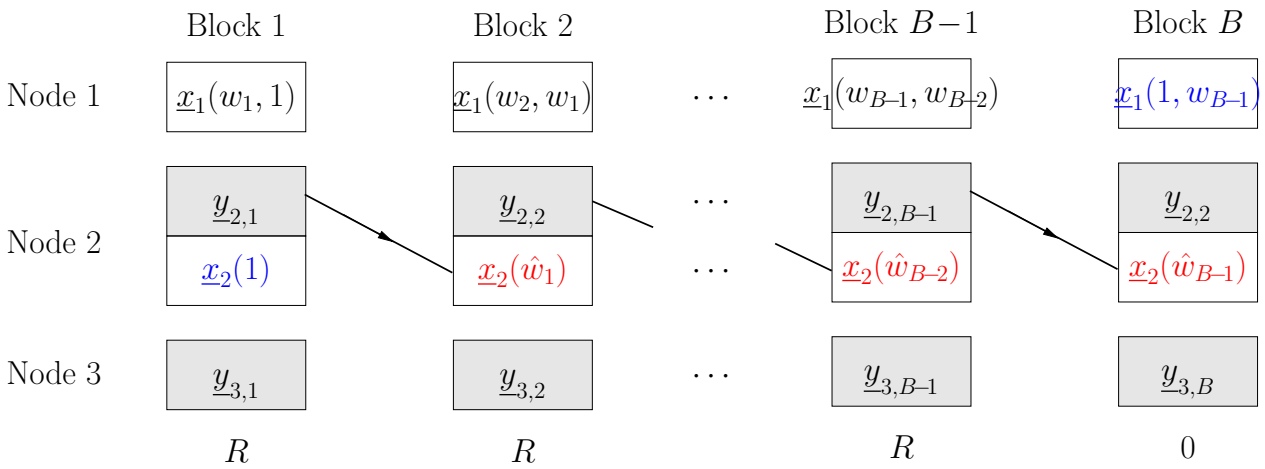


Outline: Part 2

3 Protocols
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)



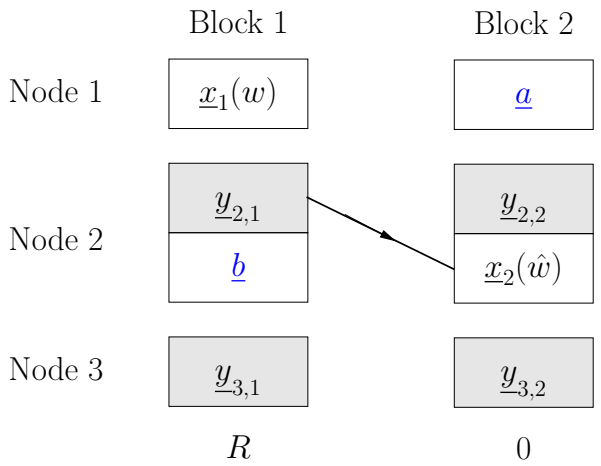
Decode-and-Forward (DF)



- Multi-user coding schemes
  - $X_{1,i}^n$ and $X_{2,i}^n$ both depend upon W_{i-1}
  - Joint distributions $P_{X_2}(\cdot)P_{X_1|X_2}(\cdot)$



Multihop DF (MDF), $B=2$



- Let q ($\overline{q}=1-q$) be the fractional length of Block 1 (Block 2)
- If Node 2 decodes using only Block 1, R achievable if

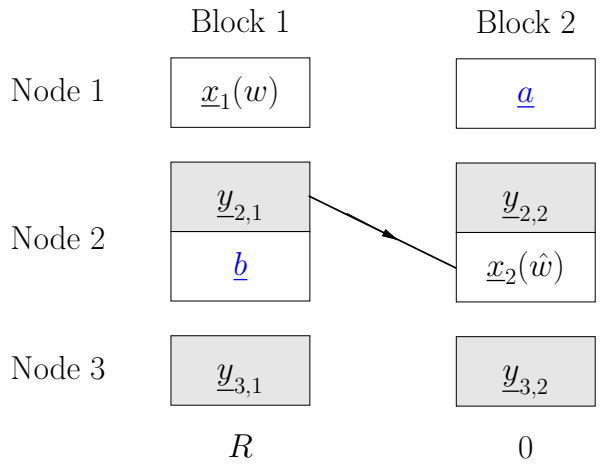
$$R < q \times I(X_1; Y_2 | X_2 = b)$$

  and $\hat{W} = W$ with high probability
- If Node 3 decodes using only Block 2, R achievable if

$$R < \overline{q} \times I(X_2; Y_3 | X_1 = a)$$



Multihop DF (MDF), $B=2$



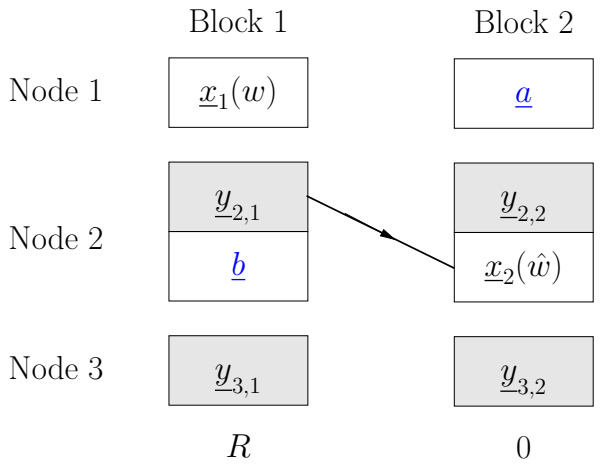
- Thus, a rate R is achievable if

$$R < R_{\text{mdf}} := \max_{0\le q\le 1} \min\{q \times \max_{P_{X_1|X_2}(\cdot|b),b} I(X_1; Y_2|X_2=b),$$
$$\overline{q} \times \max_{P_{X_2|X_1}(\cdot|a),a} I(X_2; Y_3|X_1=a)\}$$

- Routing: Time-sharing between Direct from Node 1 to Node 2 and Direct from Node 2 to Node 3
- Also known as "regenerative repeating"



Diversity DF (DDF), $B=2$



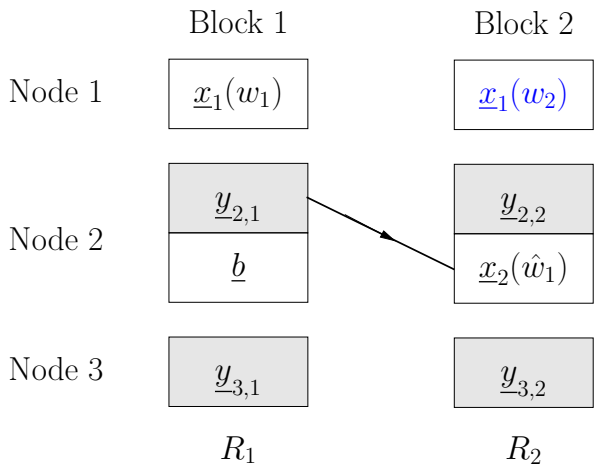
- If Node 3 decodes using both Blocks 1 and 2, then R achievable for

$$R < R_{\text{ddf}} := \max \min\{q \times I(X_{1,1}; Y_{2,1}|X_{2,1}=b), \\ q \times I(X_{1,1}; Y_{3,1}|X_{2,1}=b) \\ + \overline{q} \times I(X_{2,2}; Y_{3,2}|X_{1,2}=a)\}$$

with max over $P_{X_{1,1}|X_{2,1}}(\cdot|b)P_{X_{2,2}|X_{1,2}}(\cdot|a)$, a, b, and $0 \leq q \leq 1$



Non-Orthogonal DDF (NDDF), $B = 2$



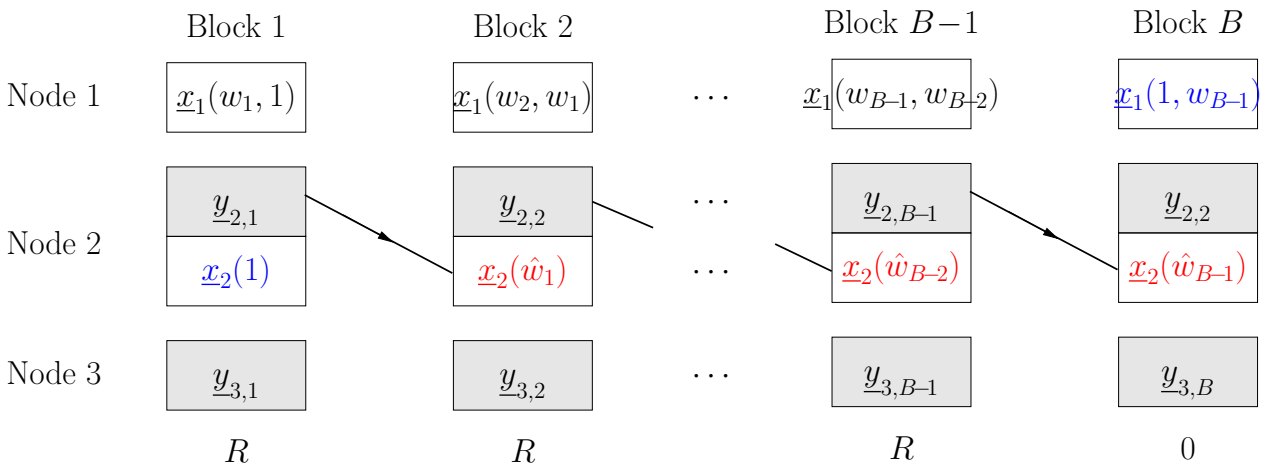
- If Node 1 sends new information in Block 2, and Node 3 decodes using both Blocks 1 and 2, then R achievable for

$$R < R_{\text{nddf},2} := \max \min\{q \times I(X_{1,1}; Y_{2,1}|X_{2,1}=b), \\ q \times I(X_{1,1}; Y_{3,1}|X_{2,1}=b) \\ + \overline{q} \times I(X_{1,2}X_{2,2}; Y_{3,2})\}$$

with max over $P_{X_{1,1}|X_{2,1}}(\cdot|b)P_{X_{1,2}X_{2,2}}(\cdot)$, a, b, and $0 \le q \le 1$



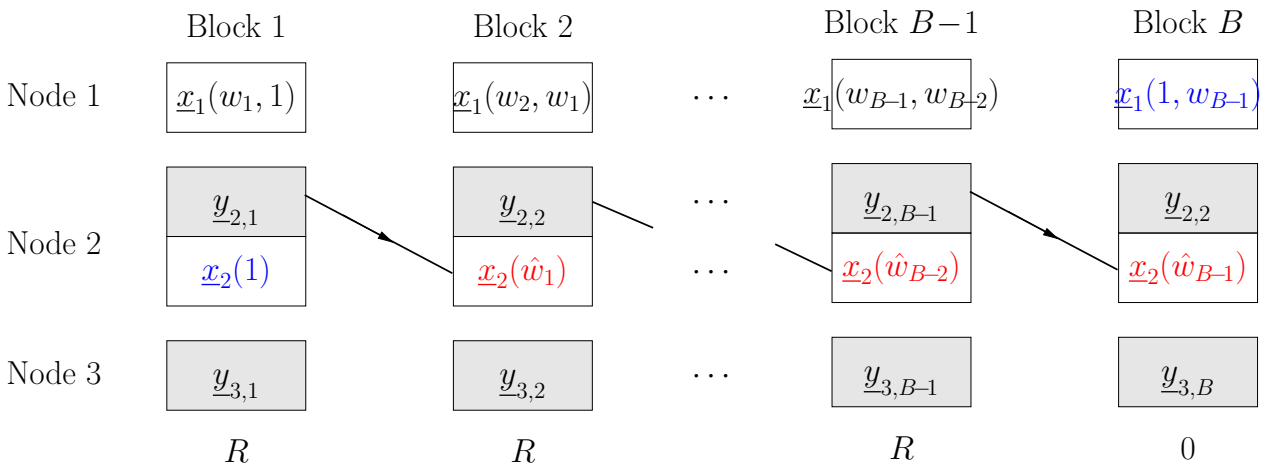
Non-Orthogonal DDF (NDDF), $B \to \infty$



- All blocks of length n
- Coding and decoding algorithms
  - Irregular encoding, successive decoding
  - Regular encoding, sliding window decoding
  - Regular encoding, backward decoding



Non-Orthogonal DDF (NDDF), $B \to \infty$



- Rate R achievable if

$$R < R_{\text{nddf},\infty} := \max \min\{I(X_1; Y_2|X_2), I(X_1X_2; Y_3)\}$$

with max over $P_{X_1X_2}(\cdot)$

- $R_{\text{nddf},\infty}$ is in fact capacity if relay channel is *physically degraded*, i.e,

$$P_{Y_2Y_3|X_1X_2}(\cdot) = P_{Y_2|X_1X_2}(\cdot)P_{Y_3|Y_2X_2}(\cdot)$$



DF Remarks

- Generally

$$R_{\text{mdf}} < R_{\text{ddf}} < R_{\text{nddf},2} < R_{\text{nddf},\infty}$$

  but coding and decoding grows increasingly complex
- MDF, DDF, and NDDF with $B = 2$ are useful for half-duplex systems
- NDDF with $B \to \infty$ is useful for full-duplex systems



Outline: Part 2

3 Protocols
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)



Multihop CF

- Basic Idea: Node 2 *quantizes* (scalar or vector) Y_2^n and communicates it to Node 3
- Details:
  - Fix distributions $P_{X_1|X_2}(\cdot|b)$ and $P_{\hat{Y}_2|Y_2X_2}(\cdot|\cdot,b)$, to be optimized later
  - Generate $\lceil 2^{nR_0} \rceil$ quantization codewords $\hat{Y}_2^n$ independently and i.i.d. according to the induced marginal distribution

$$P_{\hat{Y}_2|X_2}(\hat{y}_2|b) = \sum_{x_1,y_2} P_{X_1|X_2}(x_1|b) P_{Y_2|X_1X_2}(y_2|x_1,b) P_{\hat{Y}_2|Y_2X_2}(\hat{y}_2|y_2,b)$$

  - Upon receiving Y_2^n, Node 2 quantizes it by finding a jointly typical $\hat{Y}_2^n$ in the quantization codebook. This is likely for n large if $R_0 > I(Y_2;\hat{Y}_2|X_2=b)$



Multihop CF

- Details (cont.)
  - Use standard random coding argument for transmission from Node 2 and Node 3 to argue that

$$R_0 < \max_{P_{X_2|X_1}(\cdot|a),a} I(X_2; Y_3|X_1 = a)$$

    is achievable on the second channel
- Then $\hat{Y}_2^n$ is reproduced at Node 3, and rate R is achievable if

$$R < R_{\text{mcf}} := \max I(X_1; \hat{Y}_2|X_2 = b)$$

  with $\max$ over $P_{X_1|X_2}(\cdot|b), P_{\hat{Y}_2|Y_2,X_2}(\cdot|\cdot, b)$, a, and b such that $I(Y_2; \hat{Y}_2|X_2 = b) < \max_{P_{X_2|X_1}(\cdot|a)} I(X_2; Y_3|X_1 = a)$
- Can improve upon these rates by using non-orthogonal transmission, multi-block reception, and source coding with side information



Summary

What we covered in Part 2:

1. Summarized basic elements of relay channels, including direct, multihop, broadcast, and multiaccess transmission
2. Introduced mechanics of various kinds of relay processing, including amplify-and-forward, decode-and-forward, and compress-and-forward



Part III

Information Rates



Outline: Part 3

4. Information Rates
   - Cut-Set Bound
   - Cooperative Protocols
   - AF Rates
   - Classic Multihop Rates
   - CF Rates
   - DF Rates
   - MDF Rates
   - DF Rates with Network Coding



Outline: Part 3

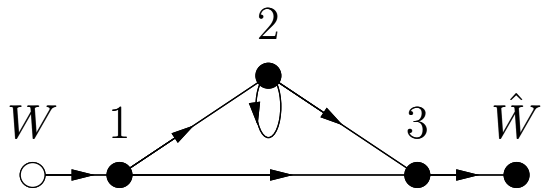
4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



Information Rates

- The purpose of this section is to refine the above analysis, study numerical examples, and develop insight based on rate.
- Consider our network model with M messages.
  The capacity region $\mathcal{C}$ is usually difficult to compute.
  A useful outer bound on $\mathcal{C}$ is the cut-set bound.
- Let $\mathcal{S} \subseteq \mathcal{N}$ and let $\mathcal{S}^c$ be the complement of $\mathcal{S}$ in $\mathcal{N}$.
  A cut separating W_m from one of its estimates $\hat{W}_m(u)$ is a pair $(\mathcal{S}, \mathcal{S}^c)$ where W_m is connected (immediately) to a node in $\mathcal{S}$ but not in $\mathcal{S}^c$, and where $u \in \mathcal{S}^c$.

- Example: relay channel has cuts $(\{1\}, \{2,3\})$ and $(\{1,2\}, \{3\})$



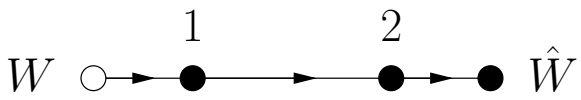
Cut-Set Bound

- Let $X_{\mathcal{S}} = \{X_u : u \in \mathcal{S}\}$
- Consider any choice of encoders, and compute
$$P_{X_{\mathcal{N}}Y_{\mathcal{N}}}(a,b) = \left[\frac{1}{n}\sum_{i=1}^{n} P_{X_{\mathcal{N},i}}(a)\right] \; P_{Y_{\mathcal{N}}|X_{\mathcal{N}}}(b|a)$$
where $P_{X_{\mathcal{N},i}}(\cdot)$ is the marginal input distribution at time i.
- Let $\mathcal{M}(\mathcal{S})$ be the set of messages separated from one of their sinks by the cut $(\mathcal{S}, \mathcal{S}^c)$.
- Cut bound: any $(R_1, R_2, \ldots, R_M) \in \mathcal{C}$ satisfies
$$\sum_{m \in \mathcal{M}(\mathcal{S})} R_m \leq I(X_{\mathcal{S}}; Y_{\mathcal{S}^c}|X_{\mathcal{S}^c})$$
- Cut-set bound for fixed $P_{X_{\mathcal{N}}}(\cdot)$: intersection over all $\mathcal{S}$ of $(R_1, R_2, \ldots, R_M)$ satisfying the above bounds
- Cut-set bound: union over $P_{X_{\mathcal{N}}}(\cdot)$ of all such regions



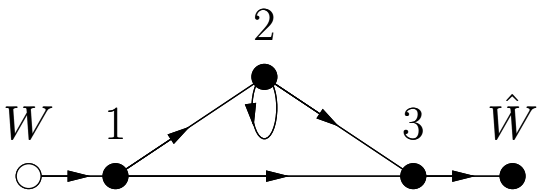
Cut-Set Bound Examples

- Point-to-point channel:



$$\mathcal{C} \subseteq \bigcup_{P_X(\cdot)} I(X;Y)$$

- Relay channel:



$$\mathcal{C} \subseteq \bigcup_{P_{X_1X_2}(\cdot)} \min\left\{I(X_1; Y_2Y_3|X_2),\ I(X_1X_2; Y_3)\right\}$$

- The cut-set bound is usually loose, e.g., for two-way channels, MACs, broadcast channels, relay channels, etc.



Outline: Part 3

4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



Cooperative Protocols

- We focus on the relay channel
- We consider rates for the following protocols:
  - Amplify-and-forward (AF)
  - Classic multihop
  - Compress-and-forward (CF)
  - Decode-and-forward (DF)
  - Multipath decode-and-forward (MDF)
  - DF with network coding
- These strategies require progressively more coordination:
  - AF and classic multihop can be used with existing devices
  - CF requires changes at the destination node and extra CSI
  - DF requires changes at both the source and destination nodes
  - Network coding requires changes at higher layers



Wireless Models for Comparisons

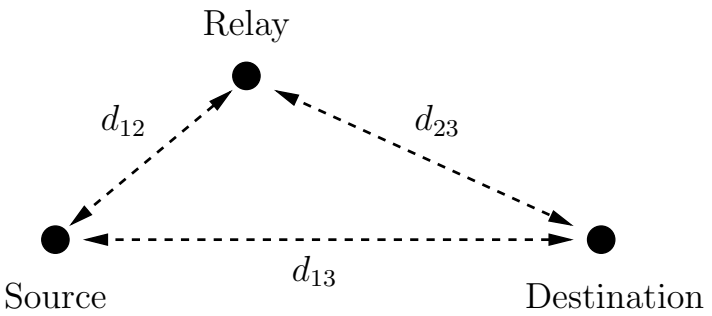
- To compare rates, we consider an idealized wireless model:
  - Relay is a full-duplex device
  - Fast fading with CSIR, No CSIT
  - In addition to no fading, we consider "uniform-phase fading" where $H_{uv} = e^{j2\pi\Phi_{uv}}$ where Φ_{uv} is uniform in $[0, 2\pi)$
- Motivation:
  - Full duplex models give basic insight for comparisons, and relays are often full-duplex, esp. for mixed wireline/wireless
  - No CSIT lets us avoid power allocation issues
  - Uniform phase "fading" give simple rate expressions and avoids (sometimes unrealistic) coherent combining gains
- We later consider half-duplex relays and slow fading

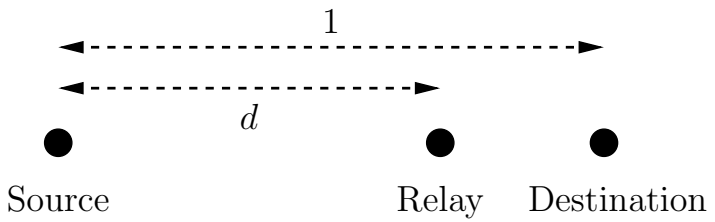


Wireless Geometry

- A general geometry is shown on the left below
- We focus on the linear geometry on the right and set all noise variances to N. The model is:

$$Y_2 = \frac{H_{12}}{|d|^{\alpha/2}} X_1 + Z_2$$
$$Y_3 = H_{13} X_1 + \frac{H_{23}}{|1-d|^{\alpha/2}} X_2 + Z_3$$





Outline: Part 3

4. Information Rates
   - Cut-Set Bound
   - Cooperative Protocols
   - AF Rates
   - Classic Multihop Rates
   - CF Rates
   - DF Rates
   - MDF Rates
   - DF Rates with Network Coding



AF Rates

- AF processing: set

$$X_{2,i} = a\, Y_{2,i-1} = a \left(\frac{H_{12,i-1}}{d_{12}^{\alpha/2}} X_{1,i-1} + Z_{2,i-1} \right)$$

  where a is chosen to satisfy the relay power constraint.

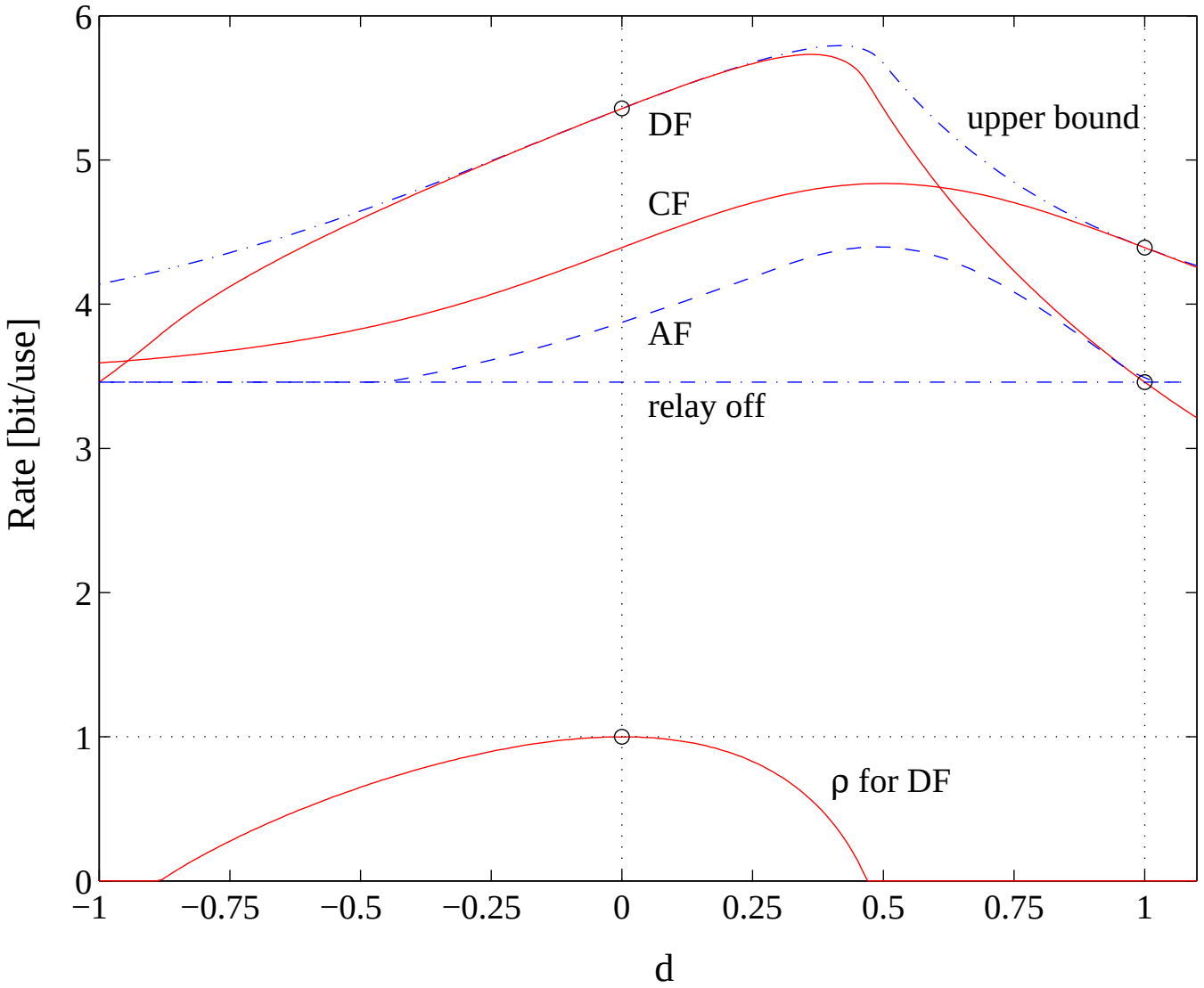
- Destination output:

$$Y_{3,i} = \frac{H_{13,i}}{d_{13}^{\alpha/2}} X_{1,i} + \frac{H_{23,i}}{d_{23}^{\alpha/2}} \left(\frac{H_{12,i-1}}{d_{12}^{\alpha/2}} X_{1,i-1} + Z_{2,i-1} \right) + Z_{3,i}$$

- AF effectively converts the channel into a unit-memory inter-symbol interference channel. The transmitter should thus perform a water-filling optimization of the spectrum of X_1^n.
- It turns out the relay should not always transmit with maximum power (see the graph below).



Rates for a Linear Geometry, No Fading

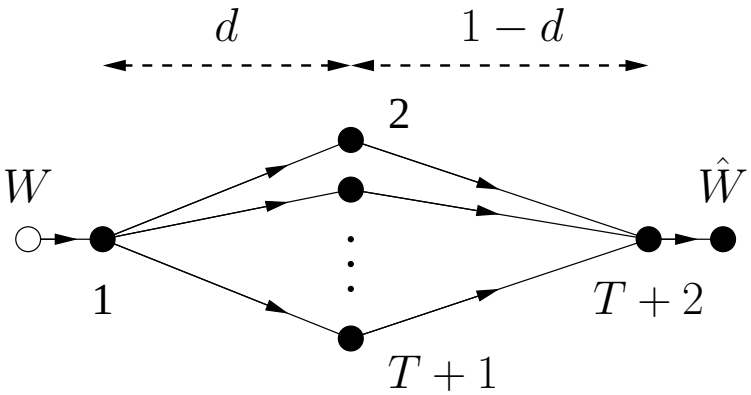


- $\frac{P_1}{N} = \frac{P_2}{N} = 10$
- $H_{uv} = 1$
- $\alpha = 2$
- AF relay off for $d < -0.5$
- $\log_2(11) \approx 3.5$
- $\log_2(21) \approx 4.4$
- $\log_2(41) \approx 5.4$



More on AF Rates

- More generally, choose row vector $\underline{a}$ and
$$X_{2,i} = \underline{a} \cdot [Y_{2,i-1}, Y_{2,i-2}, \ldots Y_{2,i-D}]^T$$
Could call this "filter-and-forward".
- AF sometimes performs well for large/small SNR or for a large number of users.
- Example: consider the 2-hop relay network shown below with T relays. AF achieves capacity for $T \to \infty$ with either per-relay or across-relay power constraints.



Outline: Part 3

4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



Classic Multihop Rates

- The source transmits W to the relay in one time slot, and the relay transmits W to the destination in the next slot
- Assign time fractions q and $\bar{q} = 1 - q$ to the first and second hops, respectively.
- For constant H_{uv}, the rate is

$$R = \min \left[q \log_2 \left(1 + \frac{P_1 |H_{12}|^2}{q d_{12}^{\alpha} N} \right), \bar{q} \log_2 \left(1 + \frac{P_2 |H_{23}|^2}{\bar{q} d_{23}^{\alpha} N} \right) \right].$$

- For $\alpha = 2$, even with best q one performs worse than using no relay for all d. But multihop helps for larger α (see below).



Outline: Part 3

4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



CF Rates

- Recall: a CF relay quantizes $\underline{y}_2$ to $\underline{\hat{y}}_2$ and transmits the quantized bits s_b reliably to the destination as $\underline{x}_2(s_b)$
- Vector quantization rate with distortion D:

$$R_Q(D) > I(Y_2; \hat{Y}_2 | X_2)$$

  where the relay cancels the effect of X_2
- Reliable transmission rate:

$$R_Q(D) < I(X_2; Y_3)$$

- The relay encoding and transmission thus requires

$$I(Y_2; \hat{Y}_2 | X_2) < I(X_2; Y_3)$$

- The destination processing requires a block encoding structure.



- The block strategy is shown below. Observe that:
  - The source transmits a new message w_b ever block b
  - $\underline{\hat{y}}_2(s_{b-1}, s_b)$ has two indices for correlation with $\underline{x}_2(s_{b-1})$
- After block $b+1$, the destination decodes the quantization bits s_b and decompresses them to $\underline{\hat{y}}_2^n(s_{b-1}, s_b)$.
- Destination decodes w_b using $\underline{\hat{y}}_2^n(s_{b-1}, s_b)$ and $\underline{y}_{3,b}^n$ at rate

$$R < I(X_1; \hat{Y}_2 Y_3 | X_2).$$

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(w_1)$	$\underline{x}_1(w_2)$	$\underline{x}_1(w_3)$	$\underline{x}_1(1)$
Relay	$\underline{y}_{2,1}$	$\underline{y}_{2,2}$	$\underline{y}_{2,3}$	$\underline{y}_{2,4}$
	$\underline{x}_2(1)$	$\underline{x}_2(s_1)$	$\underline{x}_2(s_2)$	$\underline{x}_2(s_3)$
	$\underline{\hat{y}}_2(1, s_1)$	$\underline{\hat{y}}_2(s_1, s_2)$	$\underline{\hat{y}}_2(s_2, s_3)$	$\underline{\hat{y}}_2(1, 1)$
Destination	$\underline{y}_{3,1}$	$\underline{y}_{3,2}$	$\underline{y}_{3,3}$	$\underline{y}_{3,4}$



Improved CF Rates

- Improvement: the relay transmits a hash $h(s_{b-1})$ via
$$\underline{x}_2(h(s_{b-1})).$$
The quantization codebook is now $\underline{\hat{y}}_2(h(s_{b-1}), s_b)$.
- After block $b+1$, the destination decodes $h(s_b)$ and forms $\underline{\hat{y}}_2(h(s_{b-1}), s_b)$ using the correlation of $\underline{y}_{2,b}$ and $\underline{y}_{3,b}$. The new distortion rate is:
$$R_Q(D) > I(Y_2; \hat{Y}_2 | X_2 Y_3)$$
- The overall rate is now
$$\begin{aligned} R &= I(X_1; \hat{Y}_2 Y_3 | X_2) \\ \text{s.t. } & I(Y_2; \hat{Y}_2 | X_2 Y_3) \leq I(X_2; Y_3) \end{aligned}$$
and the joint probability distribution factors as
$$P(x_1)P(x_2)P(y_2, y_3|x_1, x_2)P(\hat{y}_2|x_2, y_2)$$



CF Rates for AWGN Channels

- For AWGN channels, a natural choice is

$$\hat{Y}_2 = Y_2 + \hat{Z}_2$$

  where $\hat{Z}_2$ is Gaussian with variance $\hat{N}_2$.
- X_1 and X_2 are chosen as independent, Gaussian, and with variances P_1 and P_2, respectively.
- The smallest possible $\hat{N}_2$ is when

$$I(Y_2; \hat{Y}_2 | X_2 Y_3) = I(X_2; Y_3)$$

  which gives

$$\hat{N}_2 = N \cdot \frac{P_1 \left(|H_{12}|^2 / d_{12}^{\alpha} + |H_{13}|^2 / d_{13}^{\alpha} \right) + N}{P_2 \, |H_{23}|^2 / d_{23}^{\alpha}}$$

  and (for full-duplex relays)

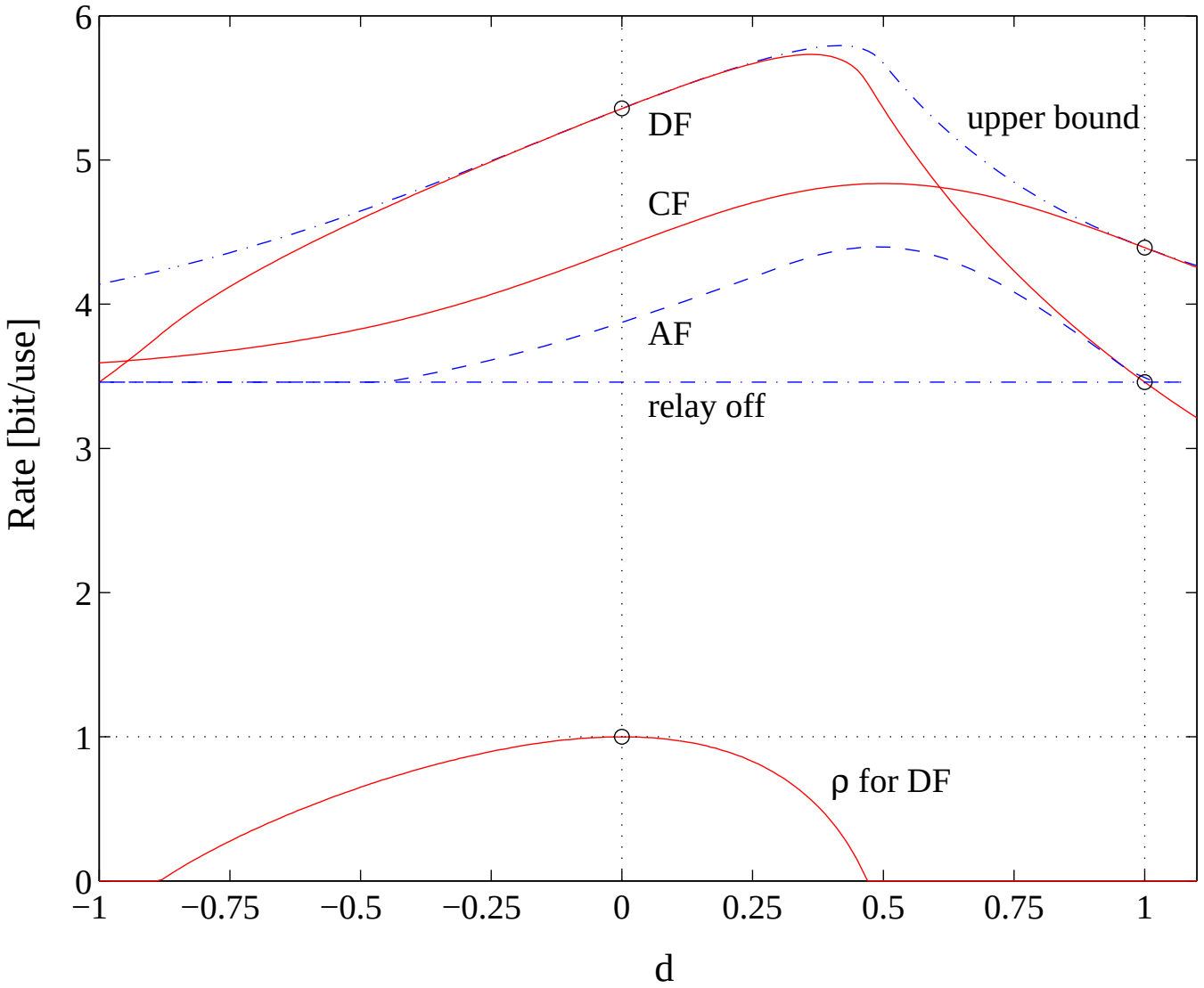
$$R = \log_2 \left(1 + \frac{P_1 \, |H_{12}|^2}{d_{12}^{\alpha} (N + \hat{N}_2)} + \frac{P_1 \, |H_{13}|^2}{d_{13}^{\alpha} N} \right) \text{ bits/use}$$



- Comments:
  - As $d_{23} \to 0$ we have $\hat{N}_2 \to 0$ and $\hat{Y}_2 \to Y_2$ and R becomes
    $$R = \max_{P_{X_1 X_2}(\cdot)} I(X_1; Y_2 Y_3 | X_2)$$
    This is a cut rate so CF is optimal as $d_{23} \to 0$.
  - Important insight: use CF when the relay is near the destination, and not AF or DF (see the rate figure).
- Improvement: time share between modes t
  - Use $P_{X_1|T}(\cdot|t)$ and $P_{X_2|T}(\cdot|t)$ a fraction $P(t)$ of the time
  - Cycle through the modes many times; the relay collects quantization bits over all modes before transmitting. Result:
    $$R = I(X_1; \hat{Y}_2 Y_3 | X_2 T)$$
    $$\text{s.t. } I(Y_2; \hat{Y}_2 | X_2 Y_3 T)| \leq I(X_2; Y_3 | T)$$
    and the joint probability distribution factors as
    $$P(t)P(x_1|t)P(x_2|t)P(y_2, y_3|x_1, x_2)P(\hat{y}_2|x_2, y_2, t)$$
- Time-sharing is useful for any strategy, e.g., AF, DF, half-duplex relaying.



Rates for a Linear Geometry, No Fading



- $\frac{P_1}{N} = \frac{P_2}{N} = 10$
- $H_{uv} = 1$
- $\alpha = 2$
- Relay adds power
- $\log_2(11) \approx 3.5$
- $\log_2(21) \approx 4.4$
- $\log_2(41) \approx 5.4$

Outline: Part 3

4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



DF Rates

- Recall the DF block structure where $\underline{x}_2(w_{b-1})$ is generated by $P_{X_2}(\cdot)$, and $\underline{x}_1(w_{b-1}, w_b)$ by $P_{X_1|X_2}(\cdot|x_{2,i}(w_{b-1}))$ for all i.
- After block b, the DF relay decodes w_b at rate

$$R < I(X_1; Y_2|X_2)$$

- The destination can use one of two decoding methods:
  - Backward decoding: decode w_b after block $b+1$ by using $\underline{y}_{3,b+1}$, $b = B, B, \ldots 1$
  - Sliding-window decoding: decode w_b after block $b+1$ by using $\underline{y}_{3,b}$ and $\underline{y}_{3,b+1}$, $b = 1, 2, \ldots B$

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(1, w_1)$	$\underline{x}_1(w_1, w_2)$	$\underline{x}_1(w_2, w_3)$	$\underline{x}_1(w_3, 1)$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w_1)$	$\underline{x}_2(w_2)$	$\underline{x}_2(w_3)$



- Backward decoder rate:

$$R < I(X_1 X_2; Y_3)$$

- Sliding-window decoder rate:

$$R < I(X_1; Y_3|X_2) + I(X_2; Y_3) = I(X_1 X_2; Y_3)$$

where the first information term is due to $\underline{y}_{3,b}$, and $I(X_2; Y_3)$ is due to $\underline{y}_{3,b+1}$ (treat $\underline{x}_1(w_b, w_{b+1})$ as interference)

- DF rate:

$$R = \max_{P_{X_1 X_2}(\cdot)} \min\{I(X_1; Y_2|X_2), I(X_1 X_2; Y_3)\}$$

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(1, w_1)$	$\underline{x}_1(w_1, w_2)$	$\underline{x}_1(w_2, w_3)$	$\underline{x}_1(w_3, 1)$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w_1)$	$\underline{x}_2(w_2)$	$\underline{x}_2(w_3)$



DF Rates for AWGN Channels

- For AWGN channels, choose Gaussian $P_{X_1X_2}(\cdot)$ with $\mathrm{E}[|X_1|^2] = P_1$, $\mathrm{E}[|X_2|^2] = P_2$, and $\rho = \mathrm{E}[X_1X_2^*]/\sqrt{P_1P_2}$.
- The DF rate is

$$R = \max_{\rho} \min \left\{ \log_2 \left(1 + \frac{P_1 \, |H_{12}|^2 \, (1-|\rho|^2)}{d_{12}^{\alpha} N} \right), \right.$$
$$\left. \log_2 \left(1 + \frac{P_1 \, |H_{13}|^2}{d_{13}^{\alpha} N} + \frac{P_2 \, |H_{23}|^2}{d_{23}^{\alpha} N} + \frac{\sqrt{P_1 P_2} \, \mathrm{Re}\{\rho \, H_{13} H_{23}^*\}}{d_{13}^{\alpha/2} d_{23}^{\alpha/2} N} \right) \right\}$$

- As $d_{12} \to 0$ we can choose $\rho \to 1$ and R becomes

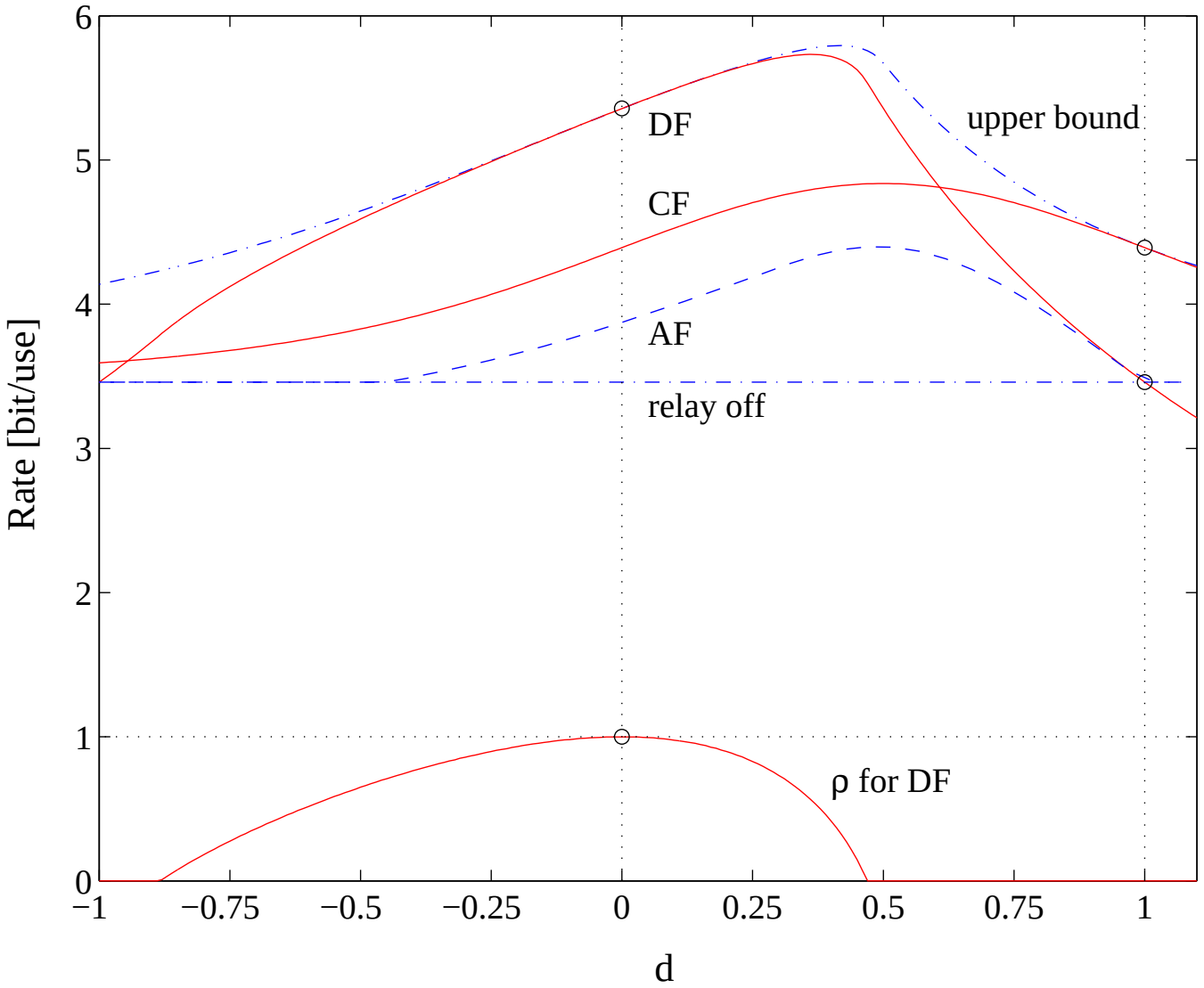
$$R = \max_{P_{X_1X_2}(\cdot)} I(X_1X_2; Y_3)$$

  This is a cut rate so DF is optimal as $d_{12} \to 0$.

- Important insight: use DF when the relay is near the source (see the rate figure).



Rates for a Linear Geometry, No Fading



- $\frac{P_1}{N} = \frac{P_2}{N} = 10$
- $H_{uv} = 1$
- $\alpha = 2$
- Relay adds power
- $\log_2(11) \approx 3.5$
- $\log_2(21) \approx 4.4$
- $\log_2(41) \approx 5.4$

- For sliding-window decoding, use the block structure below.
  - Choose the codebooks $\underline{x}_1'(\cdot)$ and $\underline{x}_2(\cdot)$ independently
  - Choose the power P_1' of $\underline{x}_1'(\cdot)$ and β so that $P_1' + \beta^2 P_2 = P_1$. This gives $\rho = \sqrt{1 - P_1'/P_1}$.
- Differences to classic multi-hop:
  1. The source and relay transmit at the same time
  2. The source & relay signals coherently combine at the dest.
  3. The destination waits and decodes using all its observations
- These differences remain for half-duplex relays.

	Block 1	Block 2	Block 3
Source	$\underline{x}_1'(w_1) + \beta \underline{x}_2(1)$	$\underline{x}_1'(w_2) + \beta \underline{x}_2(w_1)$	$\underline{x}_1'(w_3) + \beta \underline{x}_2(w_2)$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w_1)$	$\underline{x}_2(w_2)$

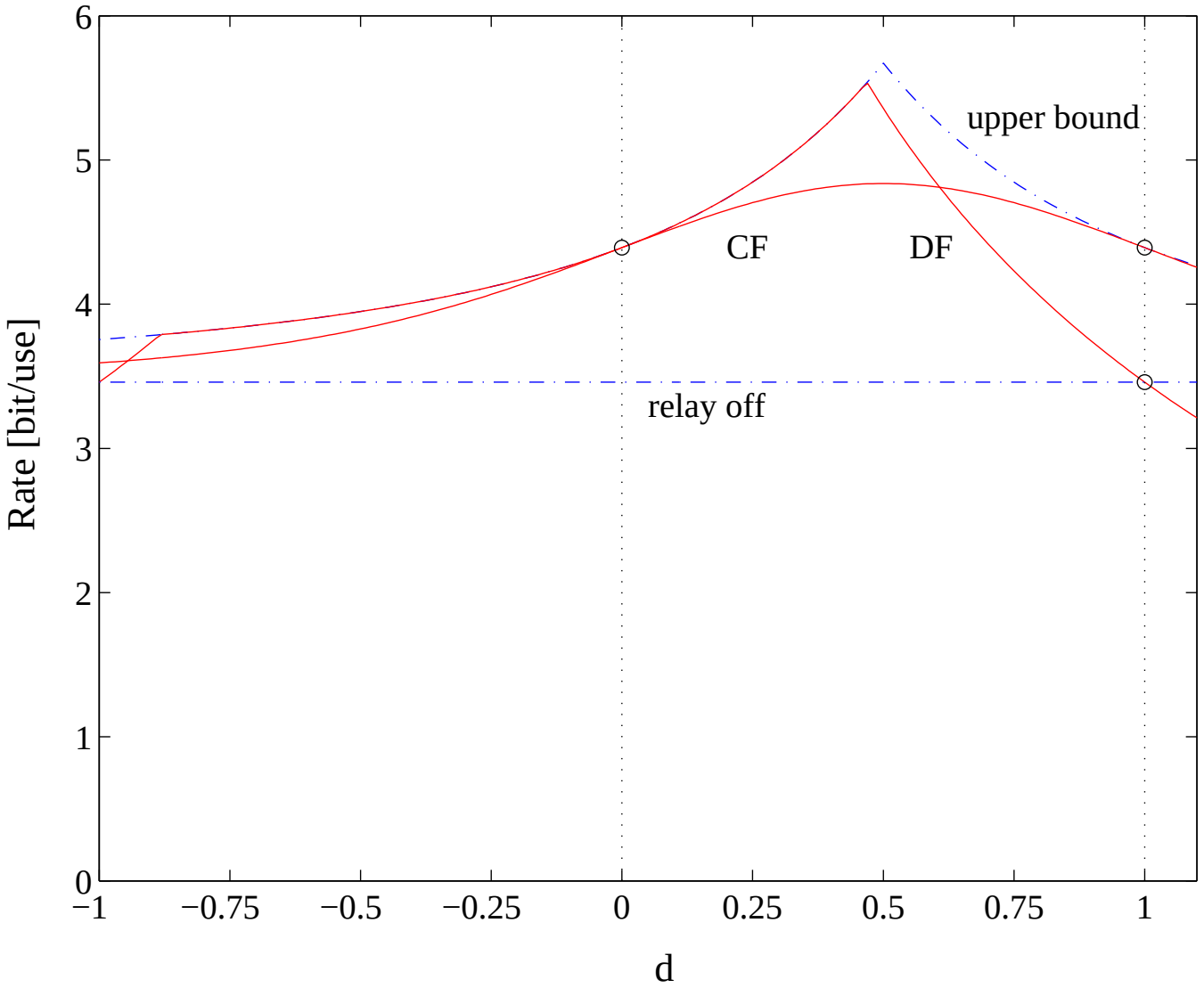


- Note: coherent combining requires CSIT at either the source or relay node and seems unrealistic in mobile environments
- For fading with uniform-phase (e.g. Rayleigh fading) and no CSIT, set $\rho = 0$ and the strategy simplifies to the form below.
- Important insight: DF achieves capacity for fading with uniform phase, no CSIT, and if the relay is near the source.
- Important insight: $\rho = 0$ does not require synchronizing transmitters, but the dest. must determine its path delays
  - Reason: for bandlimited signals and ignoring boundary effects, the interference caused by the $\underline{x}_1(\cdot)$ during sliding-window decoding does not depend on the time shift relative to $\underline{x}_2(\cdot)$

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(w_1)$	$\underline{x}_1(w_2)$	$\underline{x}_1(w_3)$	$\underline{x}_1(w_4)$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w_1)$	$\underline{x}_2(w_2)$	$\underline{x}_2(w_3)$

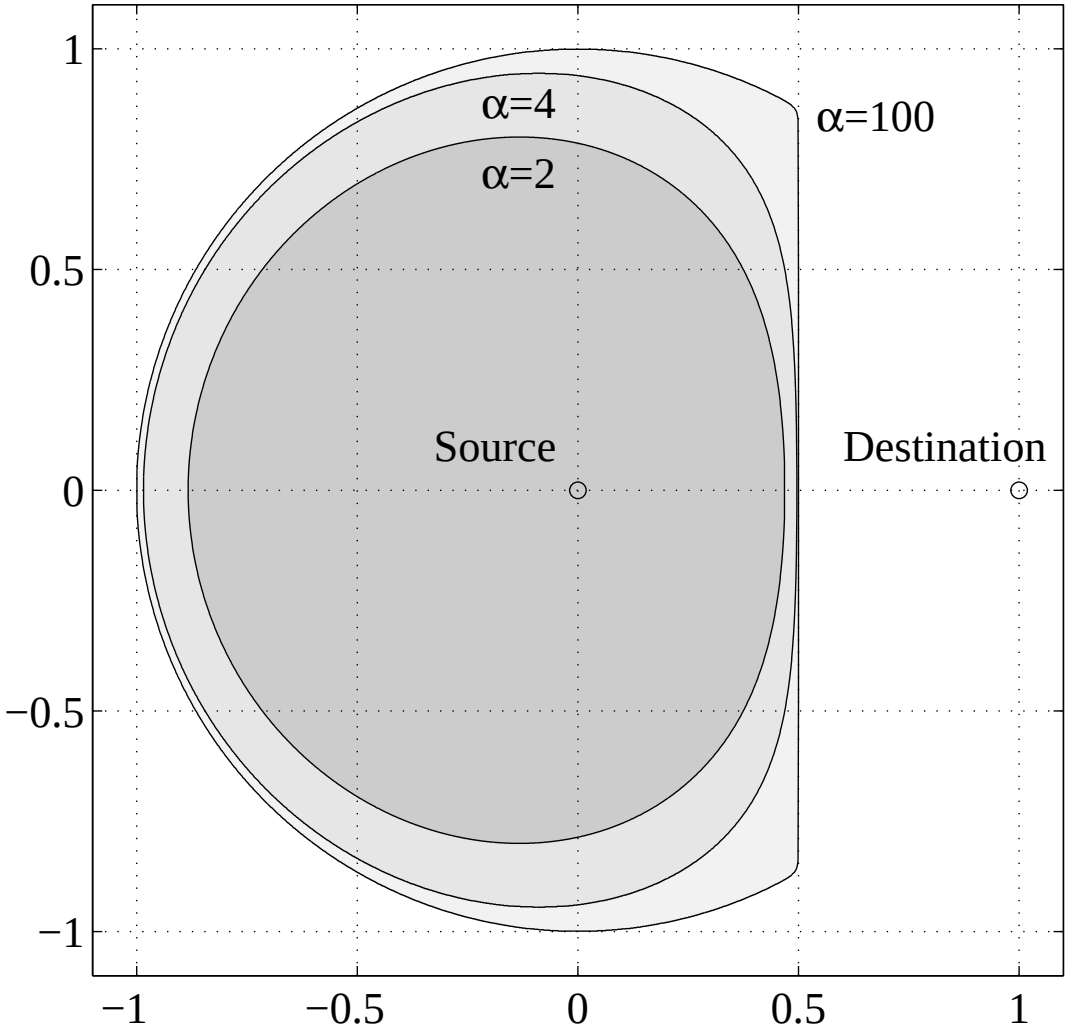


Rates for a Linear Geometry, Uniform-Phase Fading



- $\frac{P_1}{N} = \frac{P_2}{N} = 10$
- $H_{uv} = e^{j2\pi\Phi_{uv}}$
- $\alpha = 2$
- Relay adds power
- No CSIT
- Same rates as no fading and $\rho = 0$
- $\log_2(11) \approx 3.5$
- $\log_2(21) \approx 4.4$
- $\log_2(41) \approx 5.4$

Relay Positions for DF Optimality



- $\frac{P_1}{N} = \frac{P_2}{N}$
- $H_{uv} = e^{j2\pi\Phi_{uv}}$
- $\alpha = 2, 4, 100$
- Same as no fading and $\rho = 0$
- Similar plot for Rayleigh fading

Outline: Part 3

4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



MDF Rates

- The MDF block structure is shown below. The source splits W into W' and W'' with rates R' and R'', respectively.
- The source node transmits on two paths to the destination:
  - DF path: $\underline{u}(\cdot)$ plays the role of $\underline{x}_1(\cdot)$ to transmit the w'_b. The destination uses backward or sliding-window decoding:

$$R' < \min\{I(U;Y_2|X_2), I(UX_2;Y_3)\}$$

  - Direct path: superpose $\underline{x}_1(\cdot)$ on $\underline{u}(\cdot)$ to transmit the w''_b

$$R'' < I(X_1;Y_3|UX_2)$$

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{u}(1, w'_1)$	$\underline{u}(w'_1, w'_2)$	$\underline{u}(w'_2, w'_3)$	$\underline{u}(w'_3, 1)$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w'_1)$	$\underline{x}_2(w'_2)$	$\underline{x}_2(w'_3)$
Source	$\underline{x}_1(1, w'_1, w''_1)$	$\underline{x}_1(w'_1, w'_2, w''_2)$	$\underline{x}_1(w'_2, w'_3, w''_3)$	$\underline{x}_1(w'_3, 1, w''_4)$



MDF Rates for AWGN Channels

- Overall rate: $R = R' + R''$ is

$$\max \min \{I(U; Y_2|X_2) + I(X_1; Y_2|UX_2), I(X_1X_2; Y_3)\}$$

where the max. is over all $P_{UX_1X_2}(\cdot)$

- For AWGN channels, a natural choice is Gaussian $P_{UX_1X_2}(\cdot)$. For sliding-window decoding, use the block structure below:
  - Choose the codebooks $\underline{u}'(\cdot)$, $\underline{x}_2(\cdot)$, and $\underline{u}''(\cdot)$ independently
  - Choose the powers P_1' and P_1'' of the resp. $\underline{u}'(\cdot)$ and $\underline{u}''(\cdot)$ so that $P_1' + P_1'' + \beta^2 P_2 = P_1$.

	Block 1	Block 2
Source	$\underline{u}'(w_1') + \beta\underline{x}_2(1) + \underline{u}''(w_1'')$	$\underline{u}'(w_2') + \beta\underline{x}_2(w_1') + \underline{u}''(w_2'')$
Relay	$\underline{x}_2(1)$	$\underline{x}_2(w_1')$



Half-Duplex DF and MDF

- Half-duplex DF: the relay decodes all source messages. Block structure with independent X_1 and X_2:

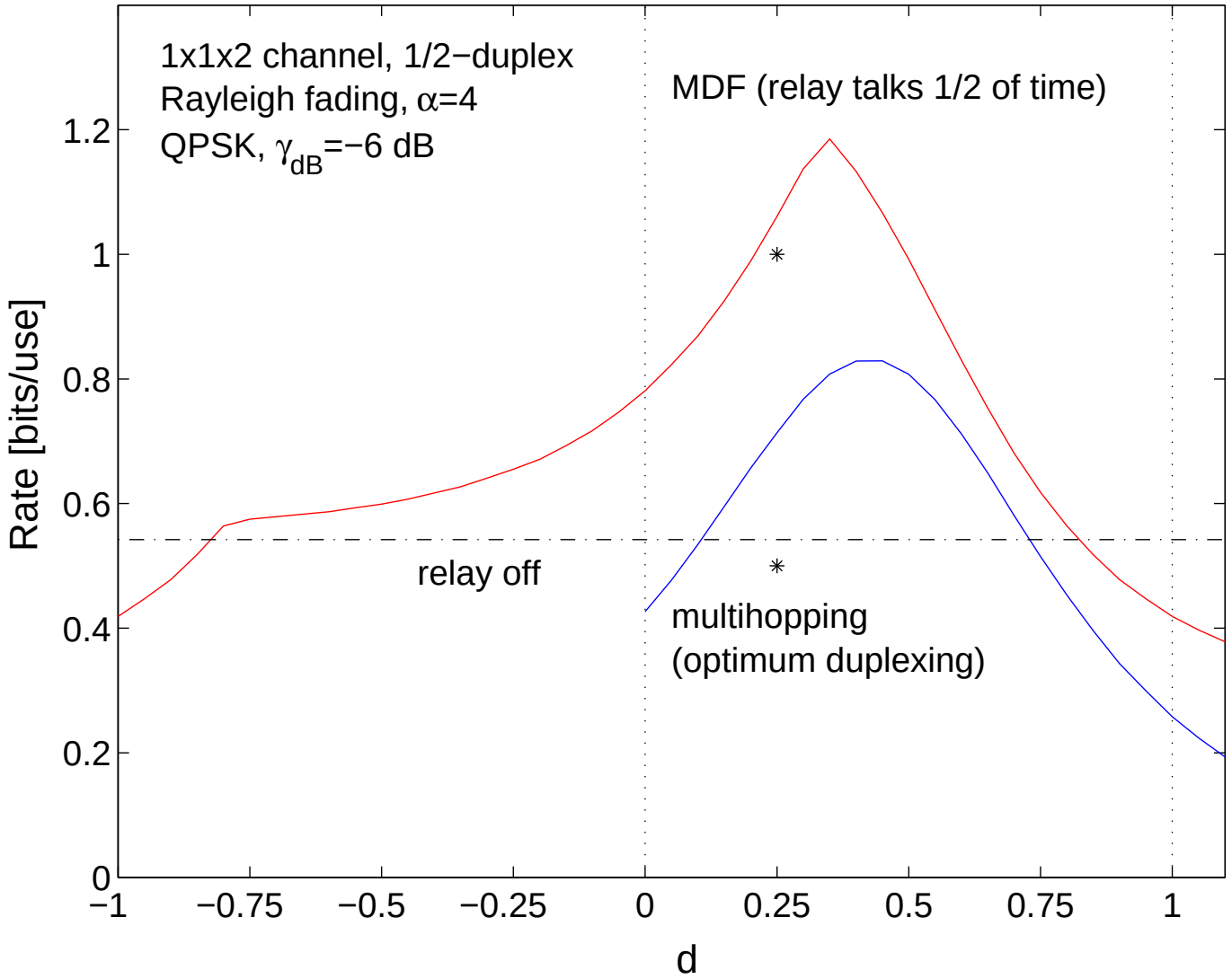
	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(w_1)$	$\underline{x}_1(w_1)$	$\underline{x}_1(w_2)$	$\underline{x}_1(w_2)$
Relay	$\underline{0}$	$\underline{x}_2(w_1)$	$\underline{0}$	$\underline{x}_2(w_2)$

- Half-duplex MDF: the relay decodes some source messages. Block structure with independent X_1 and X_2:

	Block 1	Block 2	Block 3	Block 4
Source	$\underline{x}_1(w_1)$	$\underline{x}_1(w_2)$	$\underline{x}_1(w_3)$	$\underline{x}_1(w_4)$
Relay	$\underline{0}$	$\underline{x}_2(w_1)$	$\underline{0}$	$\underline{x}_2(w_3)$



Rates for a Linear Geometry, Rayleigh Fading



- Source and relay have 1 antenna
- Destination has 2 antennas
- Low SNR: $\frac{P_1}{N} = \frac{P_2}{N} = 0.25$
- Relay adds power
- QPSK mod.
- No CSIT
- No-relay rate: $R \approx 0.54$
- At $d = 0.25$ MDF doubles the no-relay rate

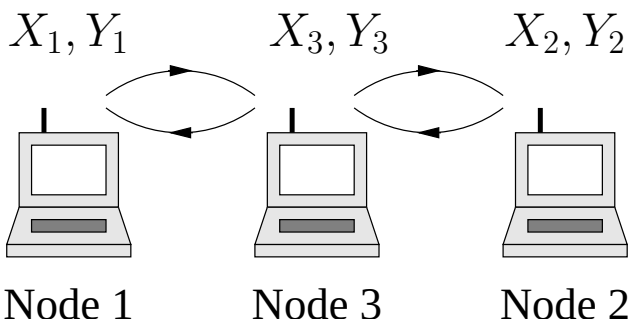
Outline: Part 3

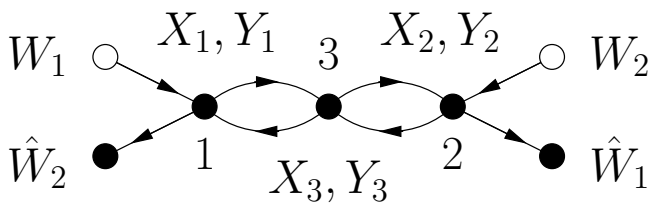
4 Information Rates
- Cut-Set Bound
- Cooperative Protocols
- AF Rates
- Classic Multihop Rates
- CF Rates
- DF Rates
- MDF Rates
- DF Rates with Network Coding



DF Rates with Network Coding

- Consider the two-way relay network below. Suppose relay node 3 decodes w_1 and w_2 reliably during a "MAC slot".
- There are many ways to "combine" DF and network coding:
  - For $R_1 = R_2$, broadcast $\underline{x}_3(w_1 \oplus w_2)$
  - For $R_1 \geq R_2$, split w_1 into w_1' and w_1'' with resp. rates R_1' and R_2 and broadcast $\underline{x}_3(w_1', w_1'' \oplus w_2)$
  - Simply broadcast $\underline{x}_3(w_1, w_2)$
  - Broadcast a modulo-lattice sum of $\underline{x}_1(w_1)$ and $\underline{x}_2(w_2)$





Summary

What we covered in Part 3:

1. Reviewed a cut-set bound
2. Refined information theory for cooperative protocols, including AF, CF, DF, MDF, and network coding
3. Examples of rate gains for a wireless relay channel
4. Gave insight on protocol choice based on geometry, duplexing constraints, and coherence



Part IV

Diversity



Outline: Part 4

5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Diversity

- The purpose of this section is to summarize how relaying and cooperation can be employed to exploit spatial diversity in wireless systems
- This form of diversity has been called *cooperative diversity*
- Start by focusing on relay channel model, but extend to cooperative multiaccess and multiaccess relay channels



Outline: Part 4

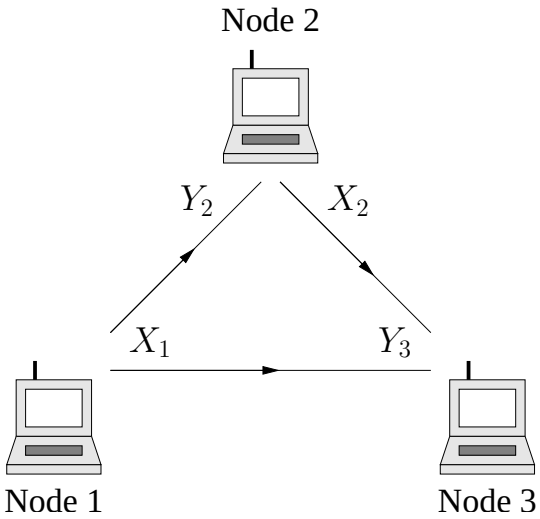
5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Basic Ideas

- Quasi-static fading / shadowing channels, e.g., $H_{uv,i} = H_{uv}$ throughout coding interval
- Outage probability and outage capacity are relevant performance metrics
- Outage requires more than one path to be severely attenuated $\Longrightarrow$ Cooperative Diversity
- Partially mimics multi-antenna array systems



Common Assumptions

- Quasi-static, Rayleigh fading (H_{uv} zero-mean, complex Gaussian)
- Channel state information (CSI) at least at the receivers
- Synchronous and coherent communications
- Half-duplex devices
- Emphasis on high SNR regime



Outage Probability and Capacity

- Treat mutual information as a random variable

$$I(X;Y) := \log \frac{P_{X,Y}(X,Y)}{P_X(X)P_Y(Y)}$$

- *Outage* occurs if mutual information falls below the target transmission rate, i.e., $I(X;Y) < R$
- Key performance metrics
  - *Outage probability*

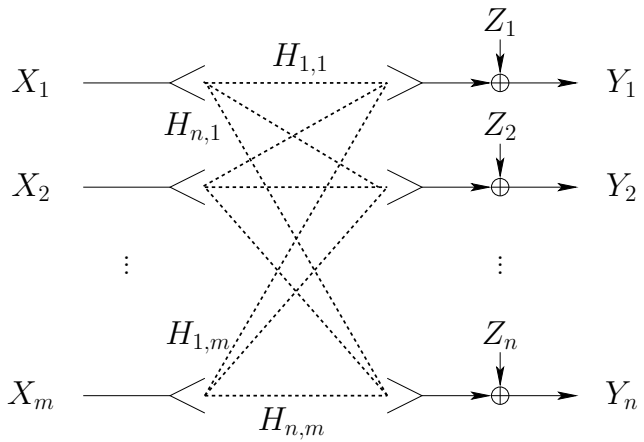
$$p_{\text{out}}(R) := \Pr[I(X;Y) < R]$$

  - *Outage capacity*

$$C_\epsilon := \sup\{R : p_{\text{out}}(R) \leq \epsilon\}$$



Multi-Antenna Array Systems



$$\underline{Y} = \underline{\underline{H}} \times \underline{X} + \underline{Z}$$

- Space-time coding
- For many i.i.d. fading distributions
  - Shannon capacity grows as

$$C \sim \min\{m, n\} \log(\gamma), \quad \gamma \gg 1$$

  - Outage probability behaves as

$$p_{\text{out}}(R, \gamma) \doteq \gamma^{-m \times n}, \quad \gamma \gg 1$$



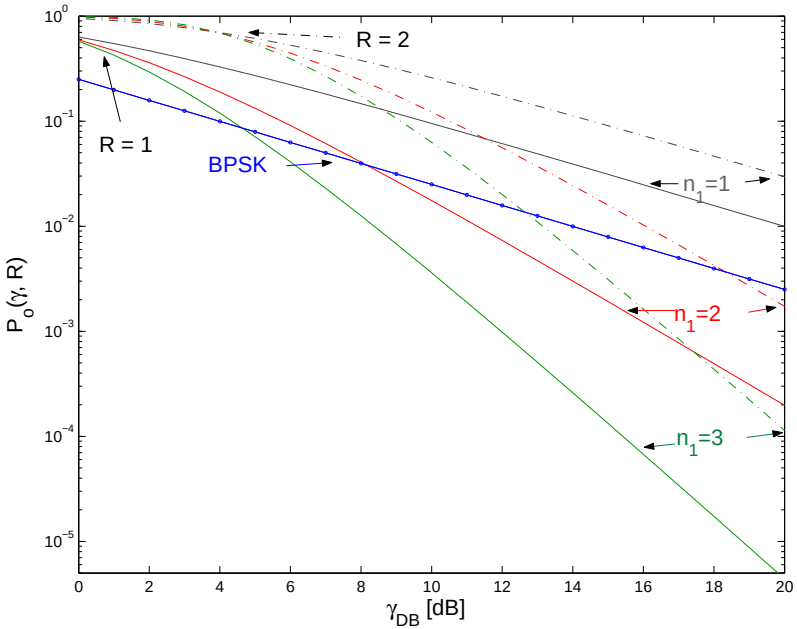
Transmit-Antenna Array Systems

- Multi-antenna system model with single receive antenna
- Mutual information for i.i.d., complex-Gaussian inputs, each with power P/m

$$I(\underline{X};Y) = \log\left(1+\frac{\gamma}{m}\sum_{k=1}^{m}|H_{1,k}|^2\right)$$

- Outage probability

$$p_{\text{out}}(R,\gamma) \sim \frac{1}{m!}\left[\frac{2^R-1}{\sigma_H^2\gamma}\right]^m$$



Key System Parameters

- Rate-normalized SNR

$$\gamma_{\text{norm}} := \frac{\gamma}{2^R - 1}$$

- Diversity and Multiplexing Gains
  - Idea: If we increase γ, we should also increase R
  - *Multiplexing gain*

$$r := \lim_{\gamma \to \infty} \frac{R(\gamma)}{\log \gamma} \quad (R(\gamma) \sim r \log \gamma)$$

  - *Diversity gain*

$$d(r) := \lim_{\gamma \to \infty} -\frac{\log p_{\text{out}}(r \log \gamma, \gamma)}{\log \gamma} \quad (p_{\text{out}}(R(\gamma), \gamma) \doteq \gamma^{-d(r)})$$

  - *Diversity-multiplexing tradeoff (DMT)*
    Tradeoff between $d(r)$ and r



Direct Transmission

- Mutual information with i.i.d., complex-Gaussian inputs

$$I_{\text{dir}} = \log(1 + |H_{1,3}|^2 \gamma)$$

- Outage probability

$$p_{\text{out,dir}}(R, \gamma) \sim \sigma_{H_{1,3}}^{-2} \left[\frac{2^R - 1}{\gamma} \right]^1, \quad \gamma \gg 1$$

- Diversity-multiplexing tradeoff

$$d_{\text{daf}}(r) = [1 - r]^+$$



Outline: Part 4

5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Diversity AF (DAF)

- DAF protocol with $B = 2$
- Equal transmit power P at Nodes 1 and 2, but each for half the coding interval
- Mutual information with Gaussian inputs

$$I_{\text{daf}} = \frac{1}{2}\log[1 + |H_{1,3}|^2\gamma + \underbrace{f(|H_{1,2}|^2\gamma, |H_{2,3}|^2\gamma)}_{\approx \min\{|H_{1,2}|^2, |H_{2,3}|^2\}\gamma}]$$

- Outage probability

$$p_{\text{out,daf}}(R,\gamma) \sim \frac{\sigma_{H_{1,3}}^{-2}(\sigma_{H_{1,2}}^{-2}+\sigma_{H_{2,3}}^{-2})}{2}\left[\frac{2^{2R}-1}{\gamma}\right]^2, \quad \gamma \gg 1$$

- Diversity-multiplexing tradeoff

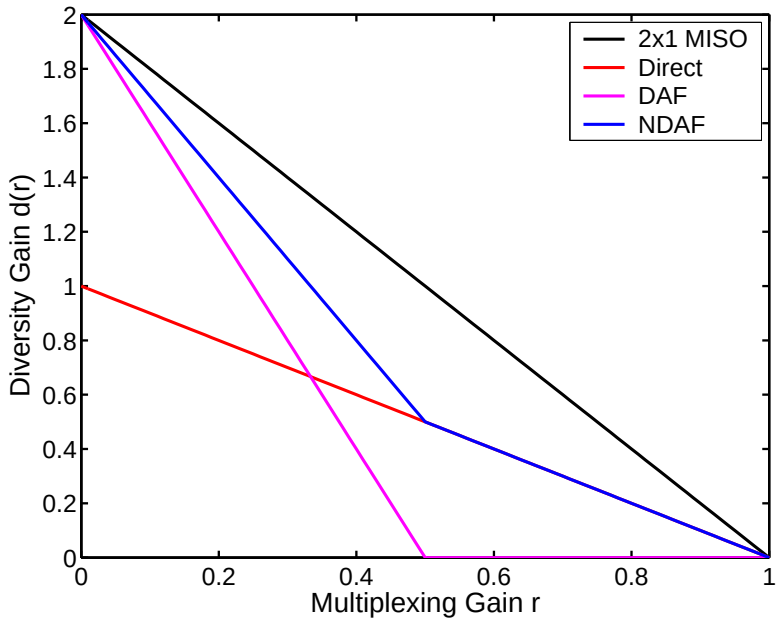
$$d_{\text{daf}}(r) = [2(1-2r)]^+$$



Non-Orthogonal DAF (NDAF)

- NDAF protocol with $B = 2$
- Improves upon Direct, DAF, and time-sharing between them
- Diversity-multiplexing tradeoff

$$d_{\text{ndaf}}(r) = [1 - r]^+ + [1 - 2r]^+$$



Outline: Part 4

5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Decode-and-Forward (DF)

- Requiring Node 2 to decode in a fixed amount of time limits diversity benefits on quasi-static fading channels
  - $d(r) \rightarrow 1$ as $r \rightarrow 0$
  - e.g., for DDF with $B = 2$ and NDDF with $B = 2, \infty$
- Two solutions (at least)
  - *Selective* protocols that react to whether or not the relay can decode
  - *Dynamic* protocols that allow the relay to receive for as long as it needs to decode



Selection DDF (SelDDF)

- Start with the DDF protocol for $B = 2$
- Condition on whether or not Node 2 can decode, and use 1 bit of overhead to convey the situation to Nodes 1 and 3
  - If Node 2 can decode after Block 1, proceed with the DDF protocol in Block 2
  - If Node 2 cannot decode after Block 1, proceed with Direct transmission in Block 2
- Variations depending upon whether Node 2 uses repetition or more powerful coding schemes
- SelDDF with repetition coding has the *same* high-SNR outage behavior and diversity-multiplexing tradeoff as DAF!

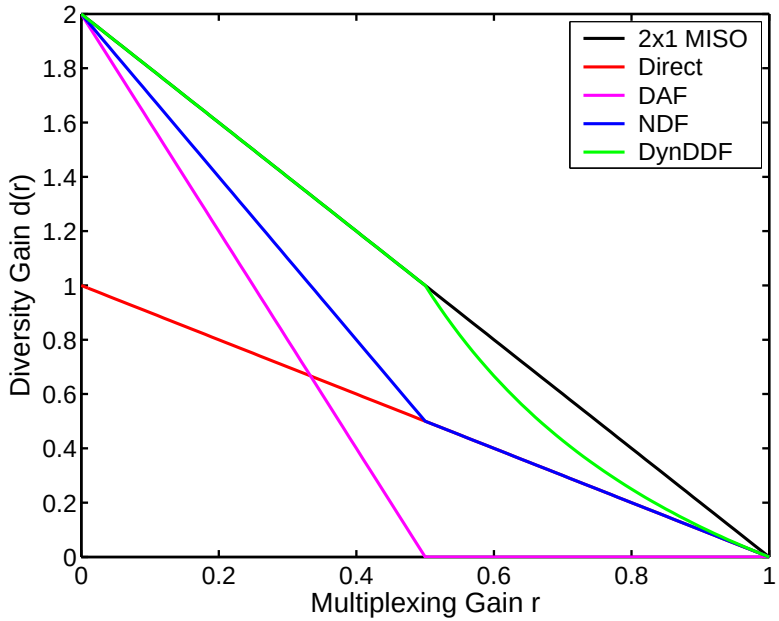
$$d_{\text{sel}-\text{ddf}}(r) = [2(1-2r)]^+$$



Dynamic DDF (DynDDF)

- Similar to the NDDF protocol for $B = 2$
- Node 2 remains in receive mode until it can decode, so that the relative lengths of the two phases of NDDF is *random*
- Diversity-multiplexing tradeoff

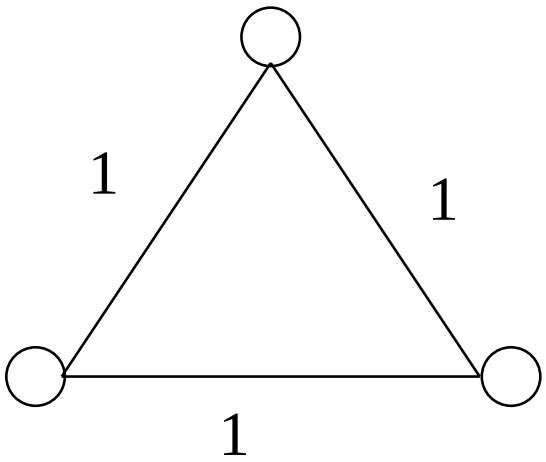
$$d_{\mathrm{dyn-ddf}}(r) = \begin{cases} 2(1-r), & 0 \le r < \frac{1}{2} \\ (1-r)/r, & \frac{1}{2} \le r \le 1 \end{cases}$$

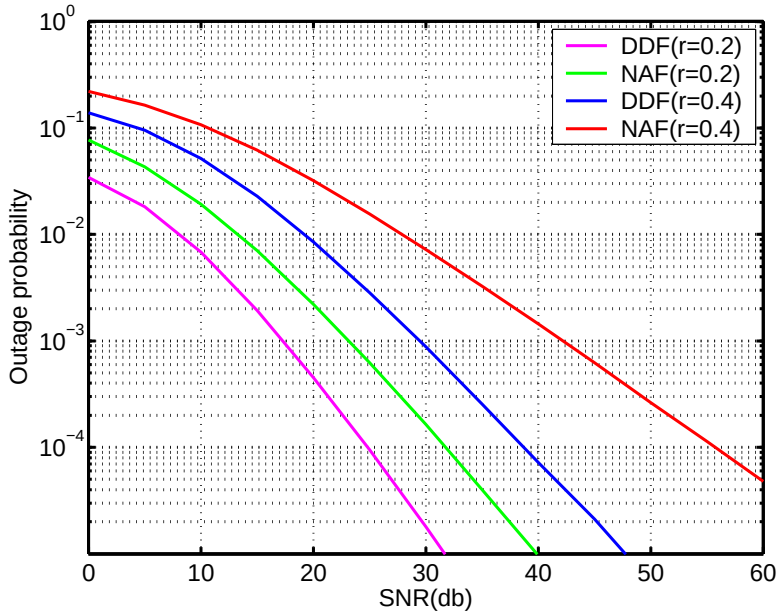


Relay Channel Outage Probabilities

System Parameters

- $R = r \log(1 + \mathrm{SNR})$
- One user, one relay





Outline: Part 4

5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Compress-and-Forward (CF)

- CF protocol with $B \to \infty$
- Relay needs CSIT via some feedback mechanism
- Gaussian test channel for quantization

$$\hat{Y}_2 = Y_2 + \hat{Z}_2, \quad \hat{Z}_2 \sim \text{Gaussian}$$

- Diversity-multiplexing tradeoff

$$d_{\text{dyn}-\text{ddf}}(r) = [2(1-r)]^+$$

- In fact, CF is DMT-optimal for half-duplex relay channel models with multiple antennas at each node, for any number of antennas!



Outline: Part 4

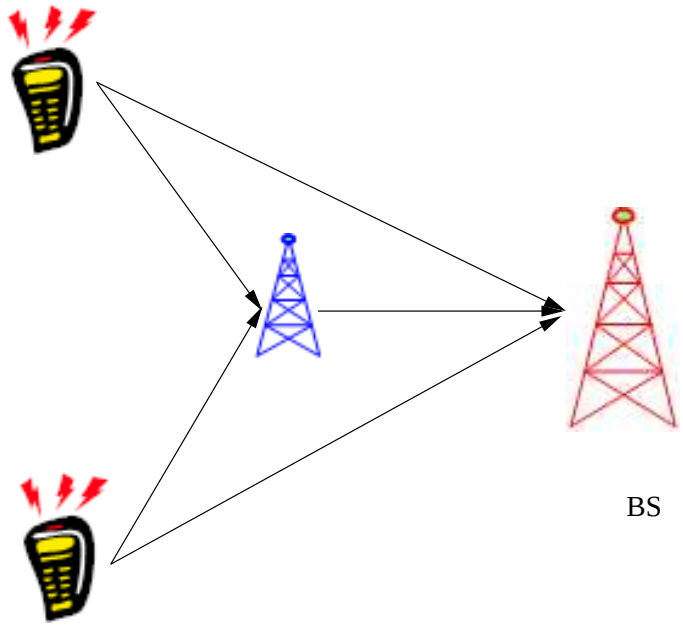
5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Multi-Access Relay Channel (MARC)

- Idea: Share a relay station among multiple mobile stations



- Questions
  - Interference/contention at the relay?
  - Backward compatibility?

MARC with NDAF and DynDDF

NDAF

- Relay amplifies the aggregate signal

$$X_{3,i} = \beta Y_{3,i-1}$$

- Simple relaying with surprisingly good performance

DynDDF

- Idea: Relay waits until *all* mobile stations can be decoded

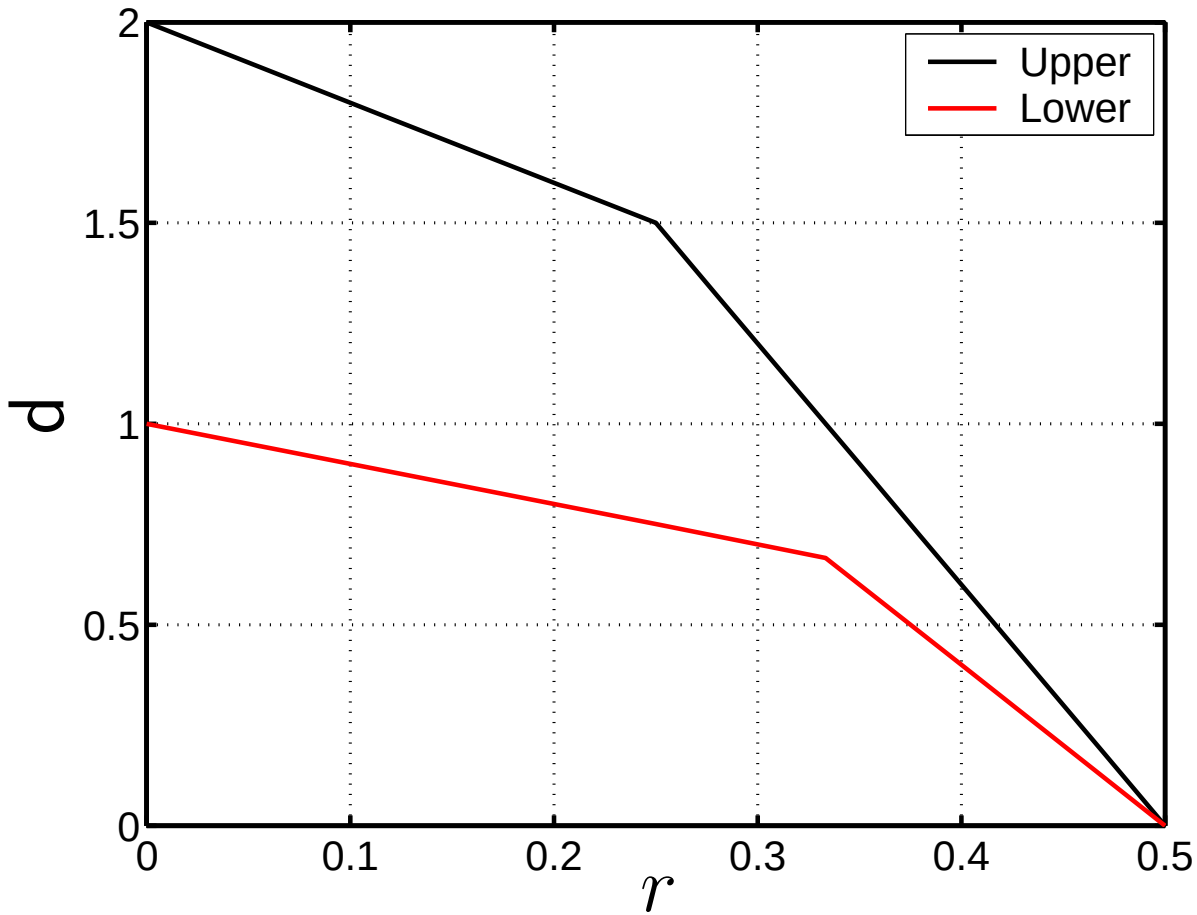
$$X_3^n = f(\hat{X}_1^n, \hat{X}_2^n)$$

- Relay complexity increases with the number of mobiles

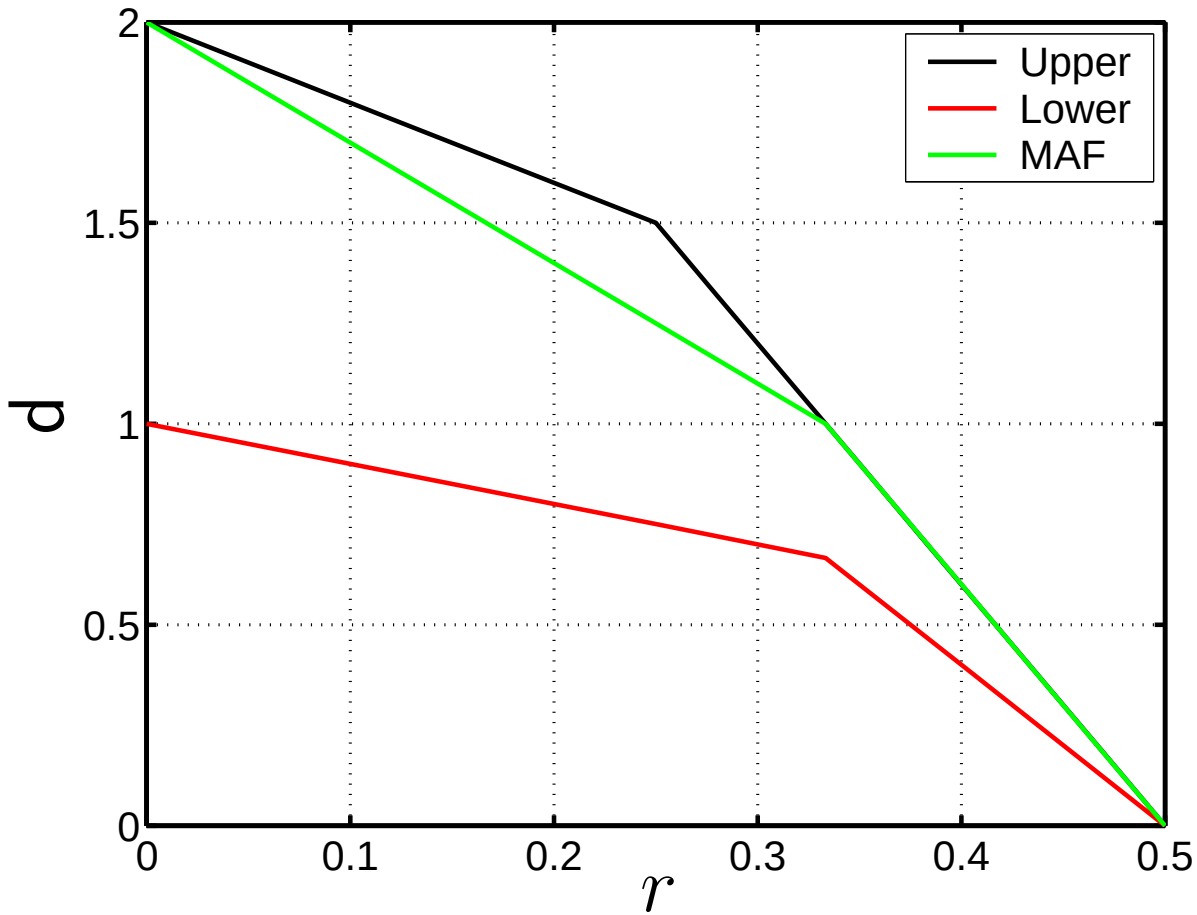
Each of these protocols is DMT-optimal for different ranges of the multiplexing gain r; CF also optimal for $r \geq \frac{4}{5}$, and strictly suboptimal to either NDAF or DynDDF for $r < \frac{4}{5}$.



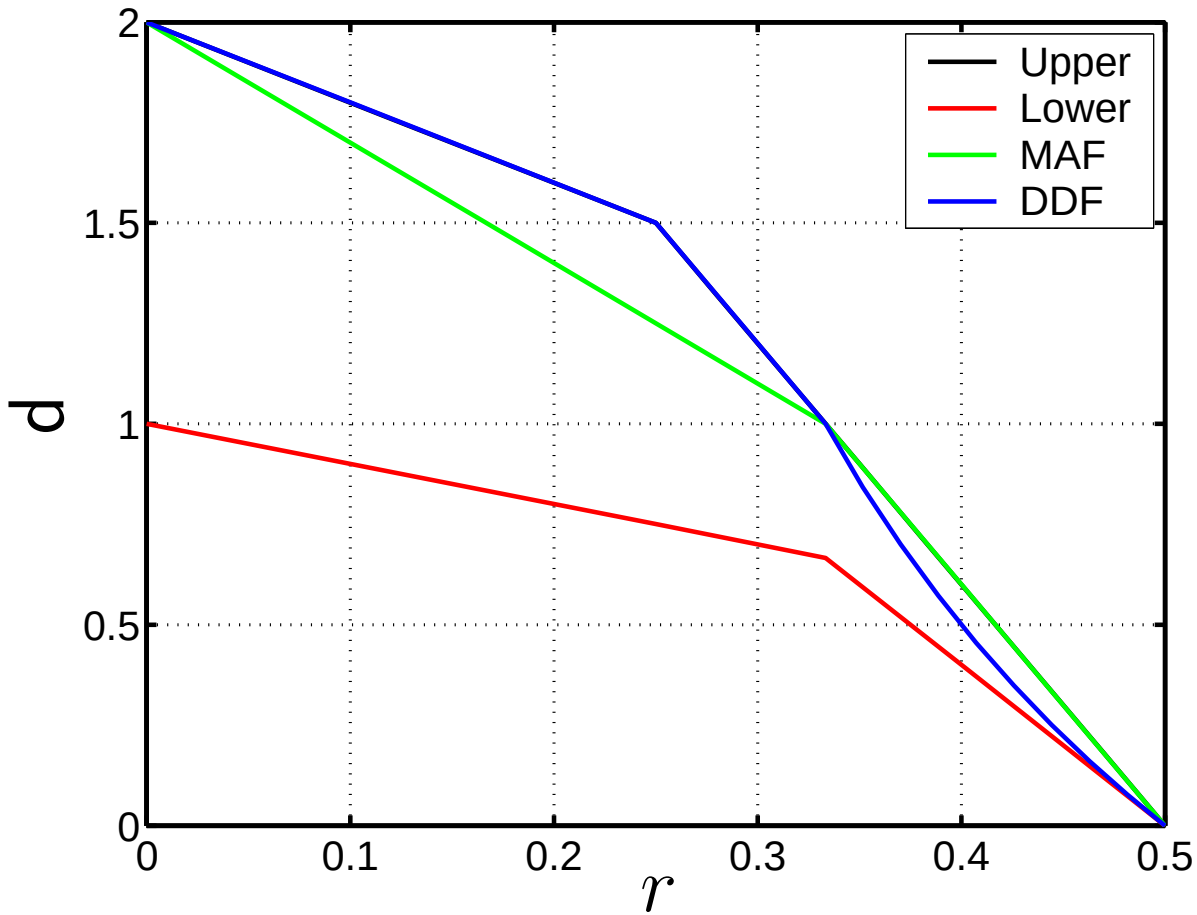
MARC Diversity-Multiplexing Tradeoffs



MARC Diversity-Multiplexing Tradeoffs



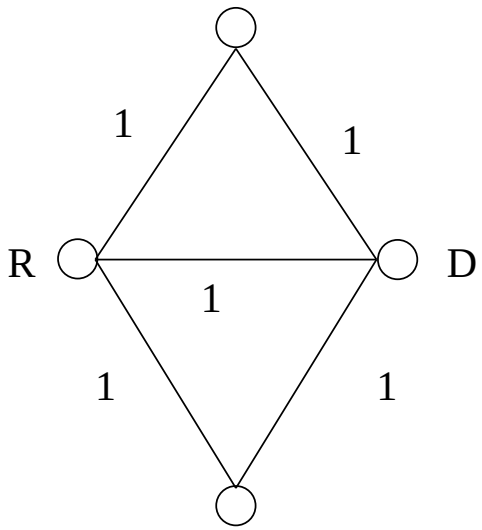
MARC Diversity-Multiplexing Tradeoffs

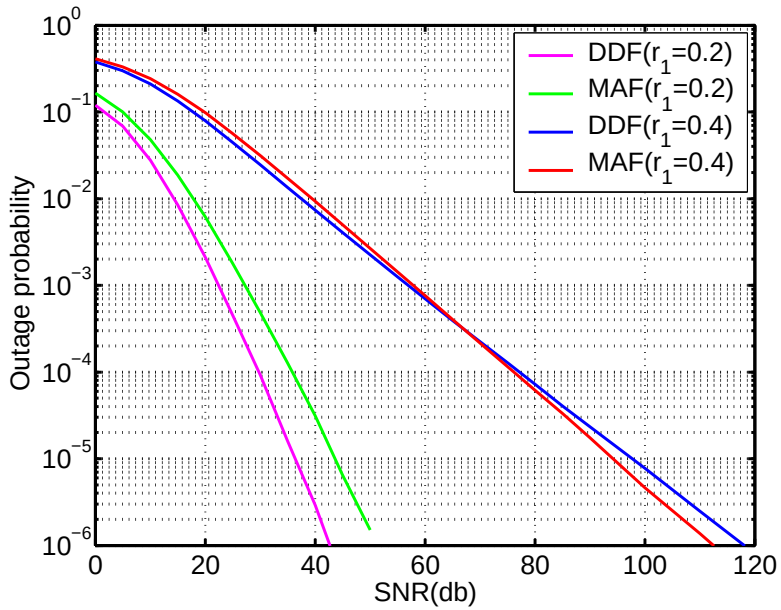


MARC Outage Probabilities

System Parameters

- $R = r \log(1 + \mathrm{SNR})$
- Two users, one relay





Summary

What we covered in Part 4:

1. Reviewed models and performance metrics for spatial diversity
2. Motivated cooperative diversity for scenarios in which other forms of diversity are not available
3. Examined how well cooperative protocols mimic multi-antenna systems in terms of *diversity* and *multiplexing*
4. Highlighted system architectures in which cooperative diversity can be exploited



Outline: Part 4

5 Diversity
- Foundations
- Amplify-and-Forward (AF)
- Decode-and-Forward (DF)
- Compress-and-Forward (CF)
- Multiaccess Relay Channel (MARC)

6 Wrap-Up



Summary

1. Many approaches, benefits, and challenges to utilizing relaying and cooperative communications
2. For a network information theorist, relaying includes, or should include, multihop (store-and-forward routing), network coding, and cooperative communications.
3. The capacity of relay systems is difficult to analyze, but the information theory is surprisingly flexible and diverse.
4. Relays add power/energy which can boost network capacity.
5. Wireless relaying lets one distribute power spatially, reducing path-loss, combating fading, and increasing coverage.
6. Designing wireless relay systems is subtle due to issues such as half-duplex constraints, synchronization, etc.



Bibliography

Classics

1. E. C. van der Meulen, "Three-Terminal Communication Channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.
2. T. M. Cover and A. A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.
3. A. B. Carleial, "Multiple-Access Channels with Different Generalized Feedback Signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 841–850, Nov. 1982.
4. F. M. J. Willems, "Informationtheoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, Oct. 1982.



Bibliography

Surveys

1. A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative Communications in Wireless Networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
2. A. Scaglione, D. Goeckel, and J. N. Laneman, "Cooperative Communications in Mobile Ad-Hoc Networks: Rethinking the Link Abstraction," *IEEE Signal Processing Mag.*, vol. 23, no. 5, pp. 18–29, Sept. 2006.
3. G. Kramer, I. Marić, and R. D. Yates, *Cooperative Communications*, ser. Foundations and Trends in Networking. Hanover, MA: now Publishers, Inc., 2007, vol. 1, no. 3-4.



Bibliography

Special Issues

1. *IEEE J. Sel. Areas Comm.*, Special Issue on Cooperative Communications and Networking, February 2007
2. *IEEE Trans. Inform. Theory*, Special Issue on Models, Theory, and Codes for Relaying and Cooperation in Communication Networks, to appear Oct. 2007

