Short Message Noisy Network Coding

Gerhard Kramer

Institute for Communications Engineering
Technische Universität München, Germany

2013 European School of Information Theory
Ohrid, Republic of Macedonia
April 22-26, 2013

Unterstützt von / Supported by
Alexander von Humboldt
Stiftung / Foundation

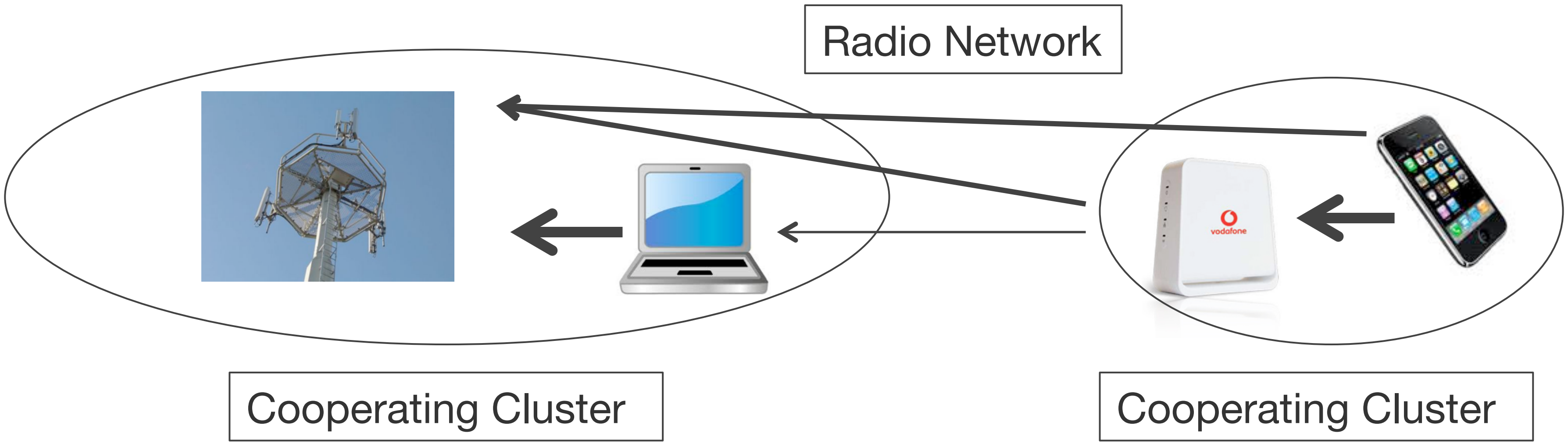
Outline

1) Cooperative Communications
2) Relaying is the core concept (1979-)
3) Relaying via Quantization (2011-)
4) Network Coding (2000-) via Relaying

1) What is Cooperative Communications?

Network communication where nodes cooperate, rather than compete, to transmit data for themselves and others

- Classic networks: TDM/FDM, admission control, routing
- Question: how should devices best operate?
  To answer this question fundamentally we need …



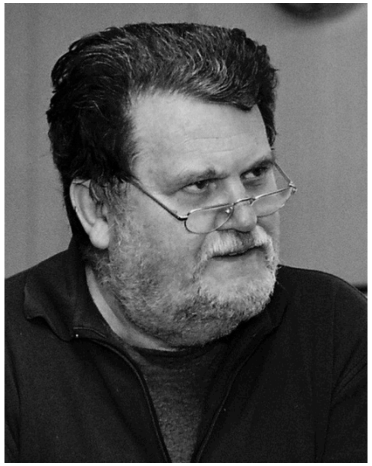
Network Information Theory

- Two Pioneers: Ahlswede and Cover
- Two of their many important contributions:
    1) Network Coding (2000) … see examples on next pages
    2) Relaying strategies (1979)

Thomas Cover (7.8.38 - 26.3.12)

1990 Shannon Lecturer

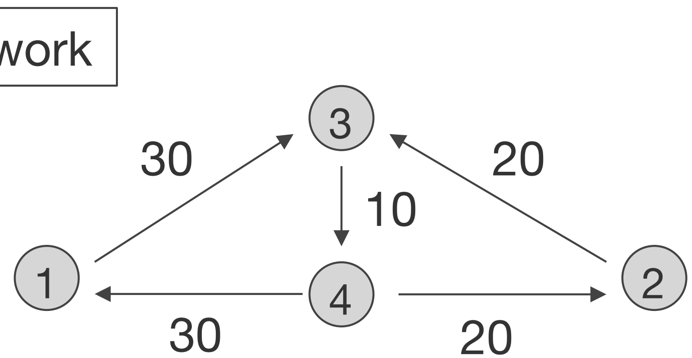
Rudolf Ahlswede (15.9.38 – 18.12.10)



2006 Shannon Lecturer

Example: Traffic Network

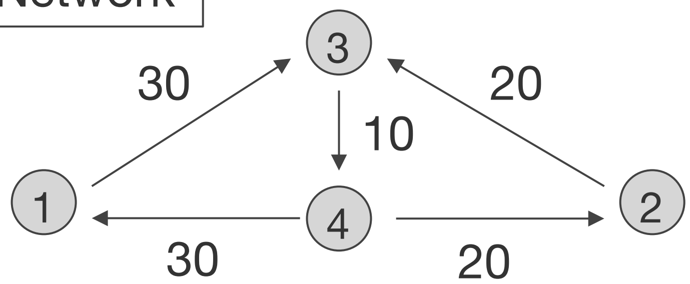
- Consider a traffic network with capacities in cars/minute.
  How many cars can flow between nodes 1 and 2 per minute?
- The bottleneck is clearly street (3,4).
  The answer is 10 cars per minute, either red or blue.
- But the answer is <u>different</u> for digital communication networks



Example: Communication Network

- Bottleneck: 10 Mbit/sec (say) but now both nodes can send 10 Mbit/sec simultaneously by using network coding
- Trick: node 3 takes bits B_1 and B_2 from nodes 1 and 2, respectively, and sends bit $C=B_1 \oplus B_2$ to node 4
- Node 1 computes $B_2=C \oplus B_1$ and Node 2 computes $B_1=C \oplus B_2$
- Many recent (2000-) results using Galois field algebra

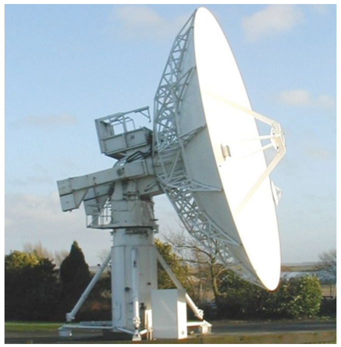
Digital Communication Network

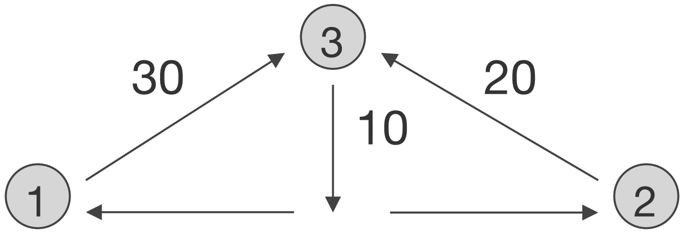


Example: Two-Way Satellite Network

- Nodes 1 & 2 send B_1 & B_2 to node 3 that broadcasts $C=B_1 \oplus B_2$
- Savings: ¾ time resources or large energy gains via coding
- Demonstrator: TUM-DLR-IQW collaboration

Satellite Network

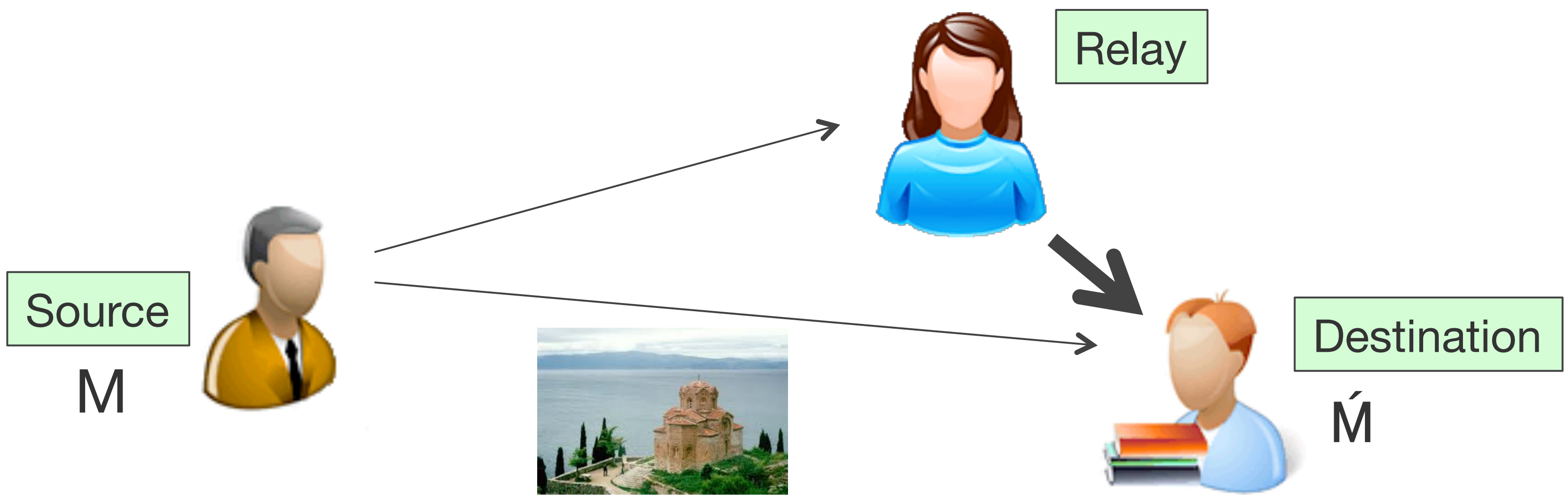




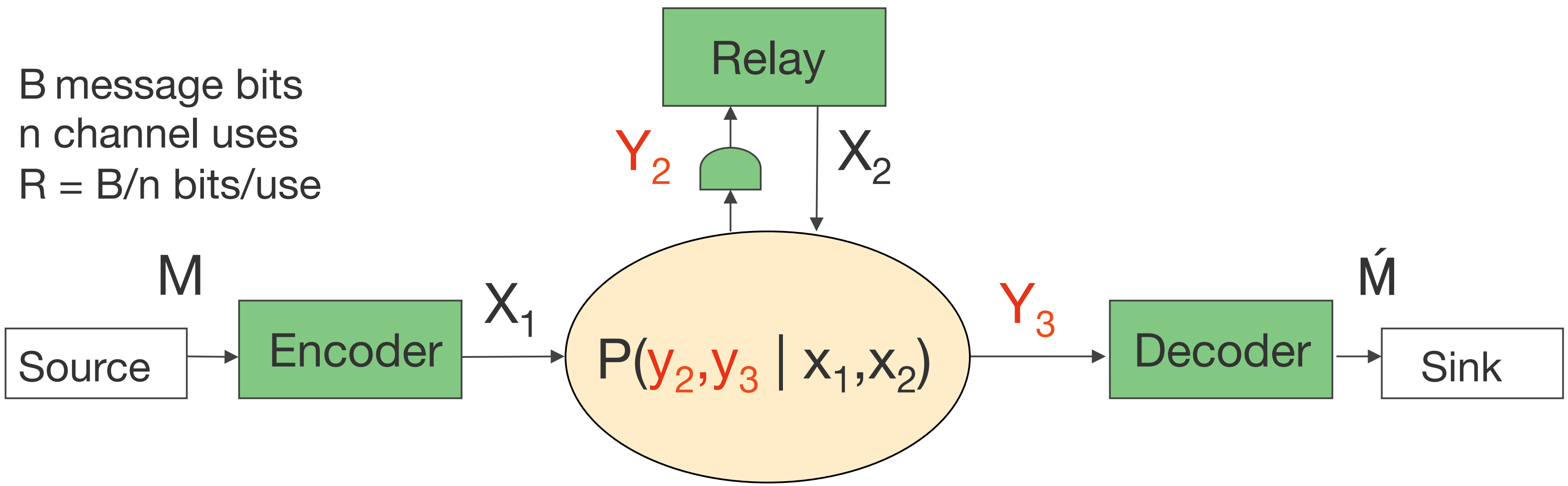
2) Relaying

The core of cooperative communications is relaying

- Examples: amplification and multi-hop
- Question: are there other good strategies?
  To answer this question fundamentally we first study a basic …

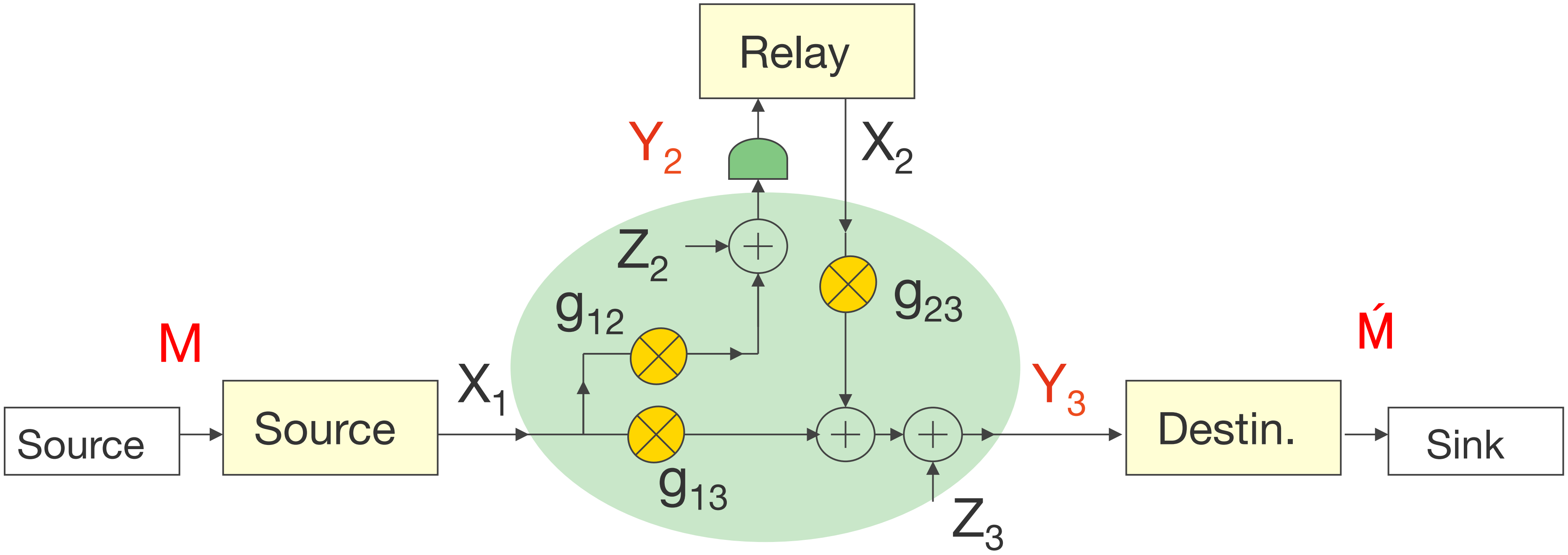


Relay Channel (Capacity an Open Problem)



- Problem: maximize R for reliability (B and n can be large)
- Network coding doesn't seem to play a role, does it?

Example: Gaussian Relay Channel



- Gaussian noise Z_t , t=2,3
- Cost: $\Sigma_i |X_{ti}|^2/n \leq P_t$, t=1,2
  (or use total power, peak power, etc.)

Known Relaying Strategies

Basic Methods and Recent Method

① Amplify-Forward (AF): amplify Y_2
Symbol Relaying: forward $f(Y_2)$ with optimized f(.)

② Decode-Forward (DF): decode message and re-encode

③ New: Compute-Forward with Lattices

Compression-Based Methods

① Classic Compress-Forward (CF), 1979

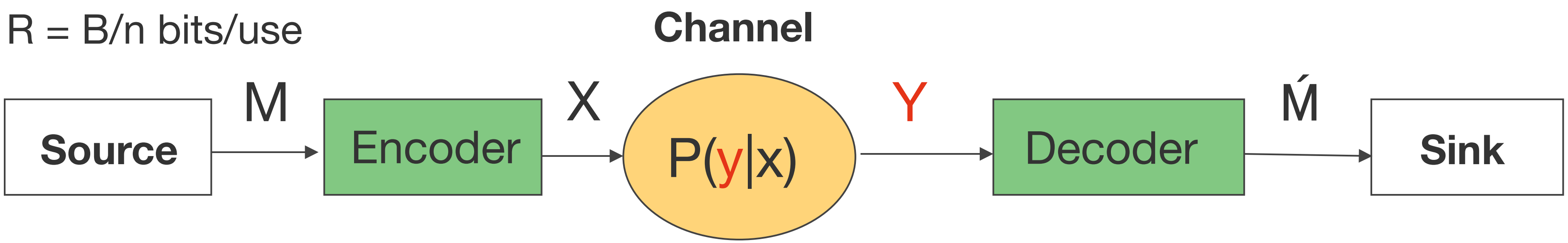
② Quantize-Map-and-Forward (QMF), 2007

③ Noisy Network Coding (NNC), 2010

④ Short-Message NNC (SNNC), 2010

Channel Coding Review **(Warning: Some IT Math!)**

B message bits
n channel uses
R = B/n bits/use

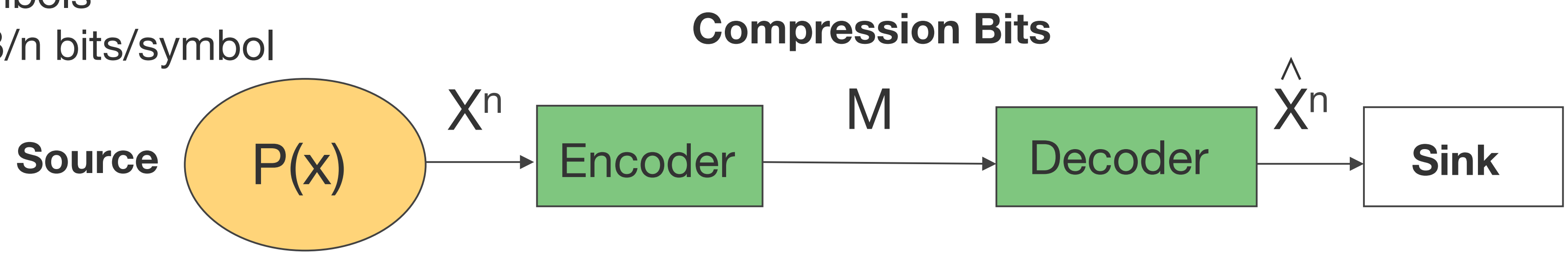


- <u>Cost</u> constraint for n symbols: $\sum_i s(X_i, Y_i) \le nS$
- Problem: find the maximum R for reliable communications (small Pr[M≠Ḿ]) under the cost constraint
- Shannon's Capacity-Cost Function:

$$C(S) = \max_{P(x)\,:\,E[s(X,Y)] \le S} I(X;Y)$$

Source Coding Review

B compression bits
n symbols
R = B/n bits/symbol

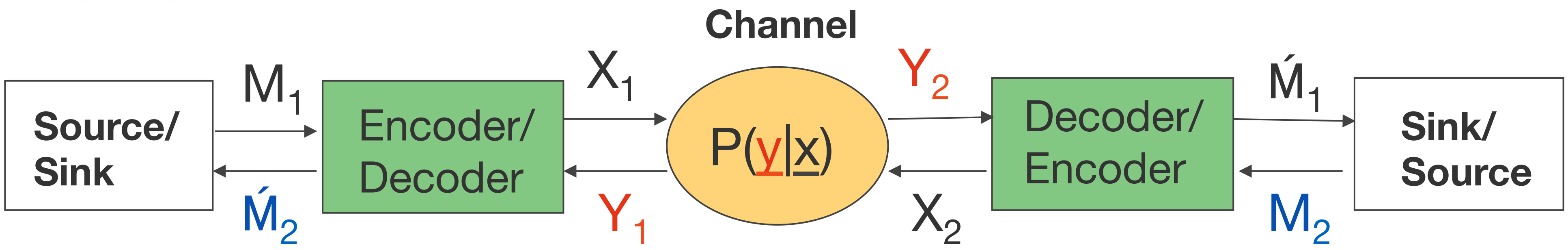


- <u>Distortion</u> constraint for n symbols: $\sum_i d(X_i , X_i) \le nD$
- Problem: find the minimum R under the distortion constraint
- Shannon's Rate-Distortion Function:

$$R(D) = \min_{P(\hat{x} \mid x) : E[d(X,\hat{X})] \le D} I(X;\hat{X})$$

Two-Way Channel Review

$R_1 = B_1/n$ bits/use
$R_2 = B_2/n$ bits/use

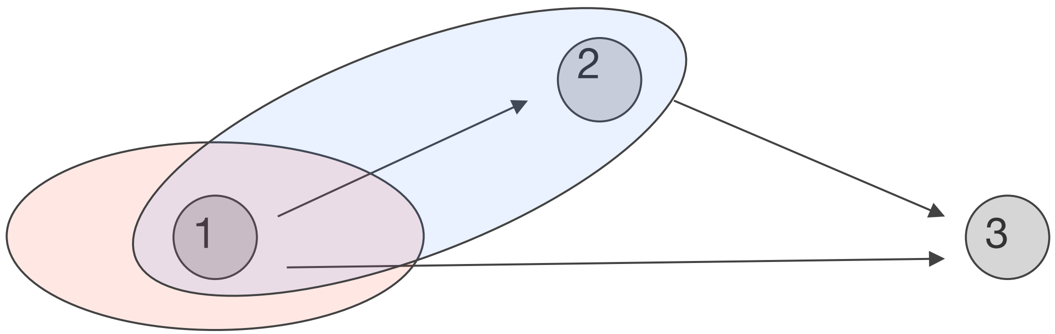


- Shannon's Capacity Bound: given $P(x_1,x_2)$ we have

$$R_1 \leq I(X_1;Y_2|X_2) \qquad R_2 \leq I(X_2;Y_1|X_1)$$

- Cut Bound: partition network nodes into 2 sets (S,S^c) and develop similar bound. Method applies to any information network (biological, physical, financial, social, etc.)

Example: Relay Channel Cut Bounds



- Two cuts: $(S,S^c)=(\{1\},\{2,3\})$ and $(S,S^c)=(\{1,2\},\{3\})$

  $R < \max \min [\, I(X_1; Y_2Y_3|X_2),\ I(X_1X_2; Y_3)\,]$

  where the max is over all $P(x_1,x_2)$

Full-Duplex vs. Half-Duplex

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{x}_{21}(1)$	$\underline{x}_{22}(q_1)$
	$\underline{y}_{21}$	$\underline{y}_{22}$

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{0}$	$\underline{x}_{22}(q_1)$
	$\underline{y}_{21}$	$\underline{0}$

- Claim: Half-duplex rates are special full-duplex rates
- The trick is to model properly: a half-duplex channel is a "Discrete Memoryless Network"
- But coding for half-duplex nodes is easier to explain

3) Relaying via Quantization

① Classic QF (here with a Half-Duplex Relay)

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{0}$	$\underline{x}_{22}(q)$
	$\underline{\hat{y}}_{21}(q)$	$\underline{0}$

- Relay quantizes $\underline{Y}_2$ to bits q representing $\underline{\hat{Y}}_2$ and transmits $\underline{X}_2(q)$
- Simple: use scalar quantization (good for high-rate quantization)
- Better: use vector quantization after canceling effect of X_2.
  Quantization: $I(Y_2; \hat{Y}_2 \,|\, X_2) < R_Q(D)$ where, e.g., $E[(Y_2-\hat{Y}_2)^2] \leq D$
- FEC Coding: $R_Q(D) < I(X_2; Y_3)$

Classic CF

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{0}$	$\underline{x}_{22}(h(q))$
	$\underline{\hat{y}}_{21}(q)$	$\underline{0}$

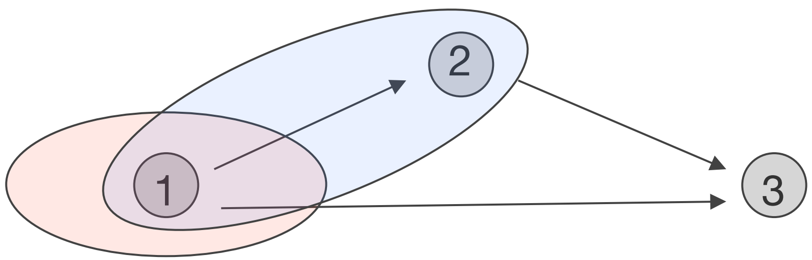
- Improvement #1: relay hashes q (aka Wyner-Ziv coding)
  Quantization bound improves to: $I(Y_2; \hat{Y}_2 | X_2 Y_3) < R_Q(D)$
- Improvement #2: bursty transmission helps at low SNR, i.e., use high power for short time intervals. Formally take into account via a "time-sharing" random variable T.

CF Rate

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{0}$	$\underline{x}_{22}(h(q))$
	$\hat{\underline{y}}_{21}(q)$	$\underline{0}$

- Final CF Rate*: with a cut-set interpretation for 2 error events

$$R < \max \min \left[\, I(X_1; \hat{Y}_2 Y_3 | X_2 T),\ I(X_1 X_2; Y_3 | T) - I(Y_2; \hat{Y}_2 | X_1 X_2 Y_3 T) \,\right]$$



*Cover-El Gamal (1979), El Gamal-et al (2006)

②③ QMF/NNC*

	Block 1	Block 2
Source	$\underline{x}_{11}(m)$	$\underline{x}_{12}(m)$
Relay	$\underline{0}$	$\underline{x}_{22}(q)$
	$\hat{\underline{y}}_{21}(q)$	$\underline{0}$

- Source repetitively encodes a long message m
  Relay quantizes only (no hashing)
  Destination decodes m and q jointly
- Advantage: theory extends nicely to many sources and relays
- Issues: long (en/de)coding delay, limited DF possibilities

* Avestimehr et al. (Allerton 2007), Lim et al. (ITW 2010)

④ SNNC*

	Block 1	Block 2
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$
Relay	$\underline{0}$	$\underline{x}_{22}(q)$
	$\hat{\underline{y}}_{21}(q)$	$\underline{0}$

- Results (since late 2010):
  (1) classic short messages achieve same rates*
  (2) can use a mixed joint/backward decoding strategy*
  (3) can use per-block processing via a multi-hop initialization**
  (4) enables DF which improves flexibility, rates, and reliability**
  (5) extension to multiple multicast***

* Wu-Xie (2010), ** K-Hou (2011), *** Hou-K (2012)

(1) Proof of Equivalence for 1 Relay

- Fix the coding distribution. NNC rate with joint decoding:

$$R < \max\{\, I(X_1; Y_3|T)^* \,,$$

$$\min\,[\, I(X_1; \hat{Y}_2 Y_3|X_2 T),\;\; I(X_1X_2; Y_3|T) - I(Y_2; \hat{Y}_2|X_1\, X_2 Y_3 T)\,]\,\} \quad (1)$$

- Additional bound for SNNC with backward decoding:

$$0 \le I(X_2; Y_3|X_1 T) - I(Y_2; \hat{Y}_2|X_1\, X_2 Y_3 T) \qquad (2)$$

- If (2) is violated, subtract (2) from 3rd expression in (1) to get:

$$R < I(X_1; Y_3|T)^*$$

- Proof method generalizes to many relays and sources **

* Destination treats X_2 as noise, ** Hou-K (2012)

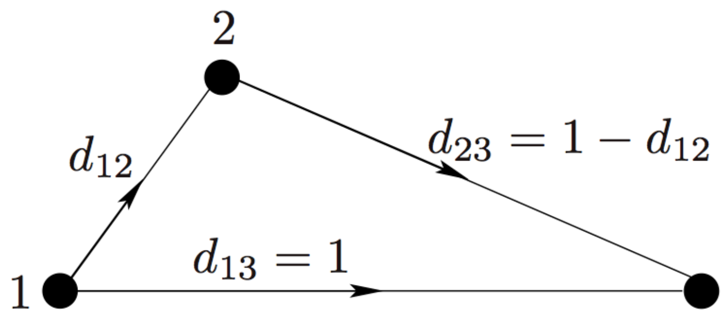
(2)(3) Full-Duplex SNNC and Backward Decoding

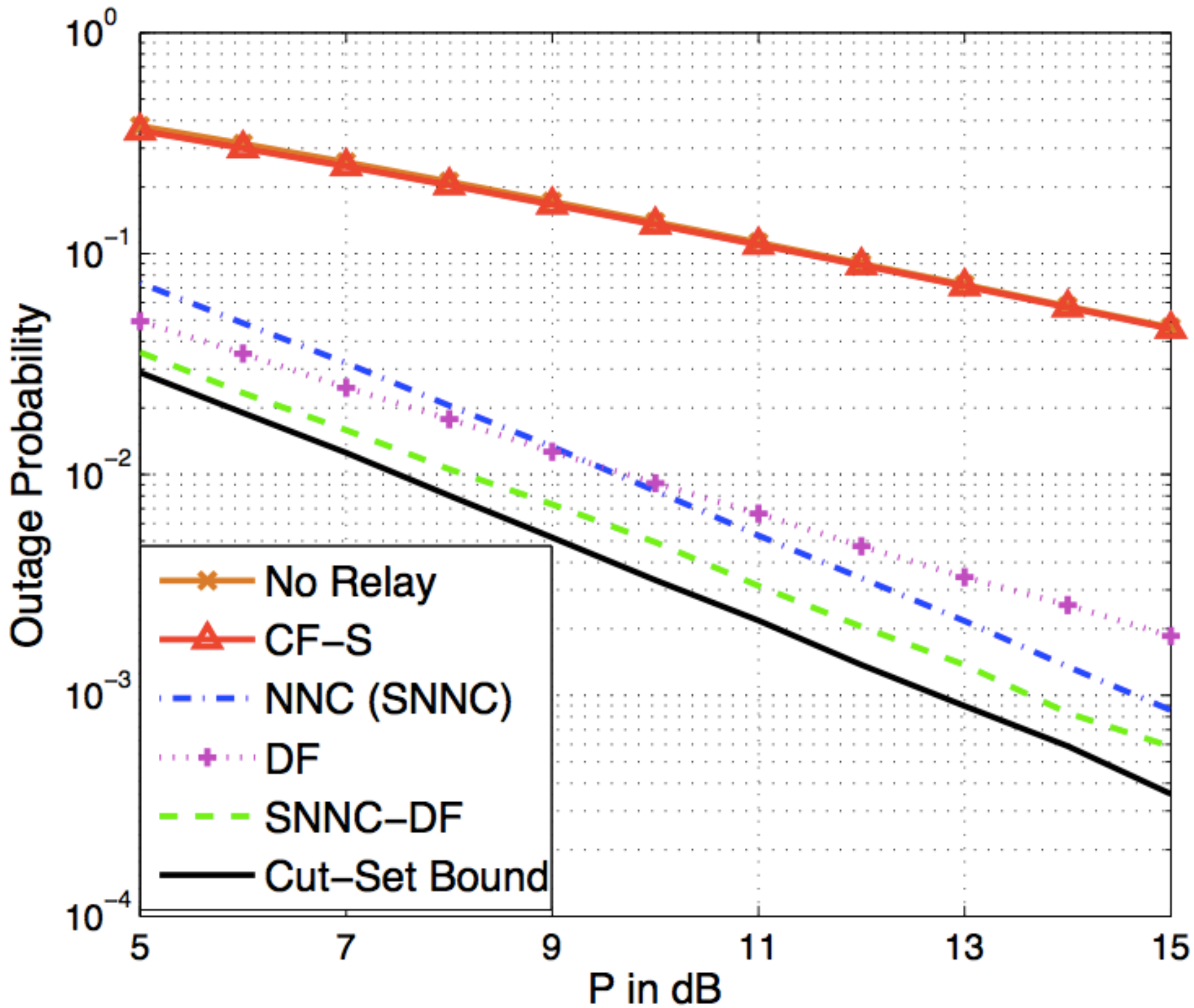
	Block 1	Block 2	Block 3
Source	$\underline{x}_{11}(m_1)$	$\underline{x}_{12}(m_2)$	$\underline{x}_{13}(m_3)$
Relay	$\underline{x}_{21}(1)$	$\underline{x}_{22}(q_1)$	$\underline{x}_{23}(q_2)$
	$\hat{\underline{y}}_{21}(q_1 \mid 1)$	$\hat{\underline{y}}_{22}(q_2 \mid q_1)$	$\hat{\underline{y}}_{23}(q_3 \mid q_2)$

- Added: superposition encode $\hat{\underline{y}}_{2b}(q_b \mid q_{b-1})$ on $\underline{x}_{2b}(q_{b-1})$
- At block 3: q_3 is known and decode m_3 and q_2 jointly
- NNC: don't care about q_b if m is recovered; get 2 bounds
- SNNC: need q_2 for the next backward step; get 3 bounds
  (To initialize: can send q_3 to destination using various methods)

(4) Enabling DF

- Single-relay, d_{12}=0.3
- Attenuation exponent 3, slow Rayleigh fading, Gaussian noise
- Per-node power: relay power P, source power 2P
- Rate target =2 bit/symbol
- SNNC gains 1 dB over NNC



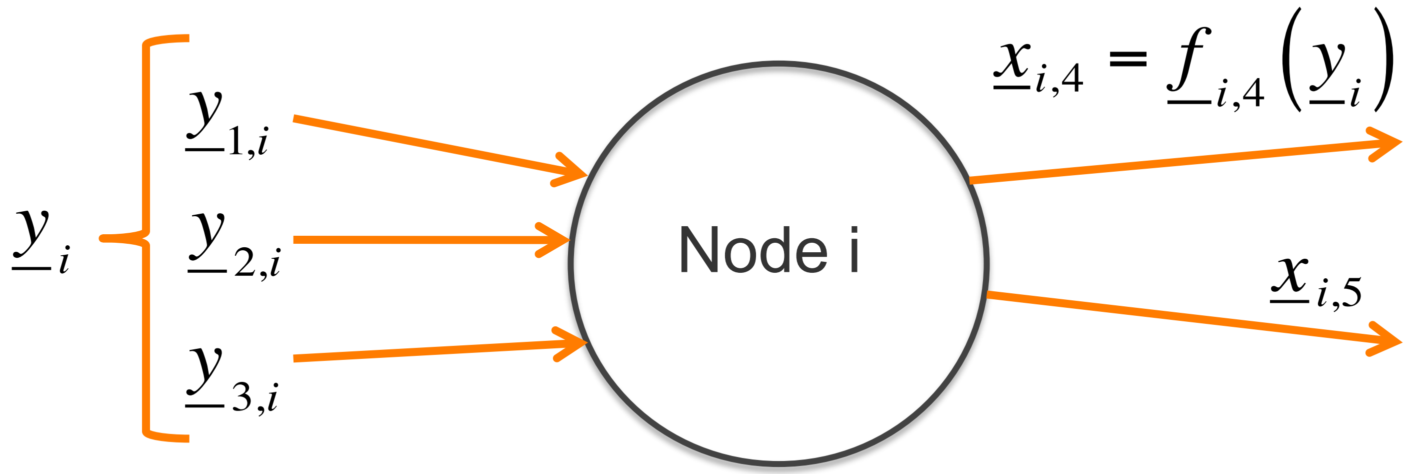


Discussion: Deterministic and Gaussian Channels

- $R < \max \min [\, I(X_1; \hat{Y}_2 Y_3 | X_2 T),\ I(X_1 X_2; Y_3 | T) - I(Y_2; \hat{Y}_2 | X_1 X_2 Y_3 T) \,]$
- Deterministic channels: $Y_2 = f(X_1, X_2)$ so choose $\hat{Y}_2 = Y_2$ and achieve cut-bound with independent inputs
(Note: capacity known and achieved by "Partial DF")
- Gaussian channel: choose $\hat{Y}_2 = Y_2 + \hat{Z}_2$ where $\hat{Z}_2 \sim N(0, N_2)$. Get
    - $I(Y_2; \hat{Y}_2 | X_1 X_2 Y_3 T=1) = I(Z_2; Z_2 + \hat{Z}_2) = \log(2N_2/N_2) = 1$ bit
    - $I(X_1; Y_2 Y_3 | X_2 T=1) - I(X_1; \hat{Y}_2 Y_3 | X_2 T=1) \leq \log(2) = 1$ bit
- R is within 1 bit of the cut-set bound with indep. X_1 and X_2
- High SNR: beamforming gains are small so virtually optimal
Low SNR: bursty signals mimic high SNR, but no beamforming

4) Network Coding via Relaying

- Classic networks: for each edge (i,j), network coding chooses $\underline{f}_{i,j}(.)$ to uniformly map $\{\underline{y}_i\}$ to $\underline{x}_{i,j}$
- Linear coding: $\underline{x}_{i,j} = A_{i,j}\,\underline{y}_i$ where $A_{i,j}$ is often taken to be random

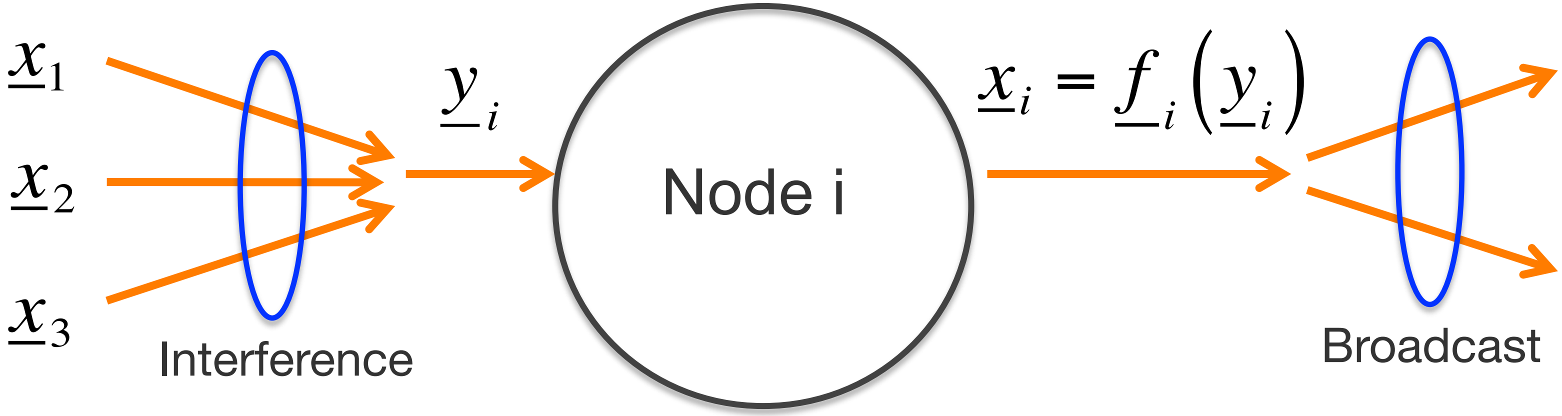


Interface: Discrete, Uniform Mapping, Independent across Nodes

Ahlswede-Cai-Li-Yeung (2000)

Network Coding for Wireless

- Nodes with interference and broadcast constraints*: For each node i, choose $\underline{f}_i(.)$ to map $\underline{y}_i$ to an $\underline{x}_i$
- Non-linear $\underline{f}_i(.)$ needed in general

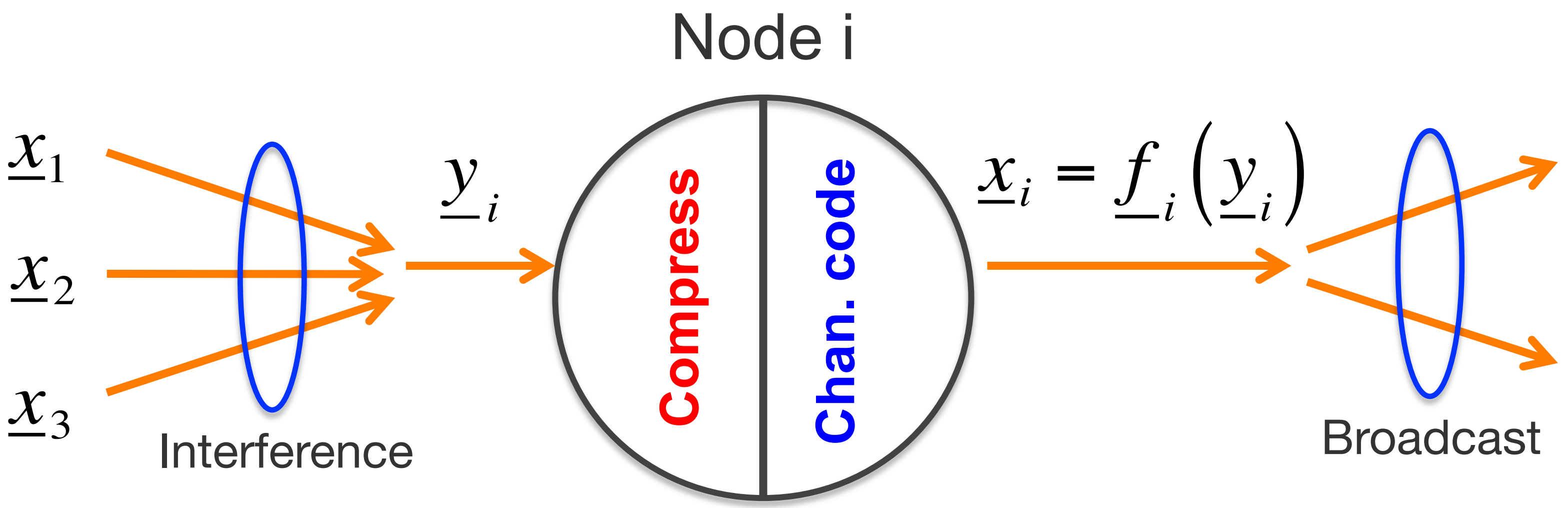


Interface: Uniform Mapping. But what if the $\underline{y}_i$ are continuous?

* Model includes classic networks as a special case

Noisy (Digital) Network Coding

- Two-step: (1) compress (quantize/hash) and (2) channel code
- Method is digital (binary interface) and non-linear in general
- Surprise(?): includes classic network coding as a special case



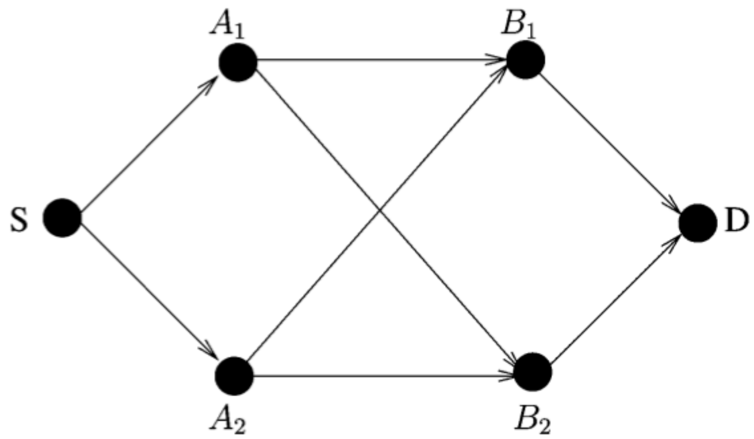
Interface: Digital, Uniform Mapping, Independent across Nodes

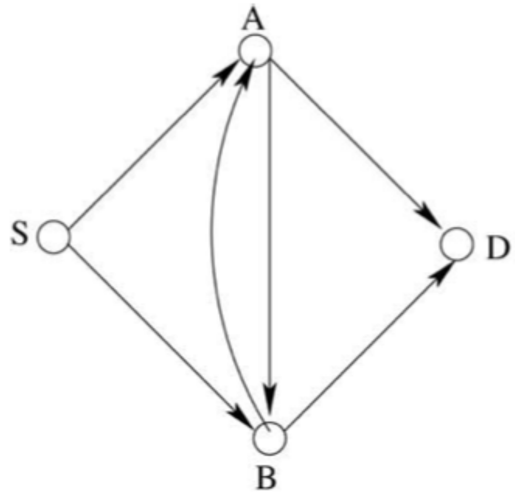
Many Nodes, either Sources or Relays

	Block b	Block b+1
Source	$\underline{x}_{1b}(m_{1b})$	$\underline{x}_{1(b+1)}(m_{1(b+1)})$
Source/Relay k	$\underline{0}$	$\underline{x}_{k(b+1)}(m_{k(b+1)},q_{kb})$
	$\hat{\underline{y}}_{kb}(q_{kb})$	$\underline{0}$

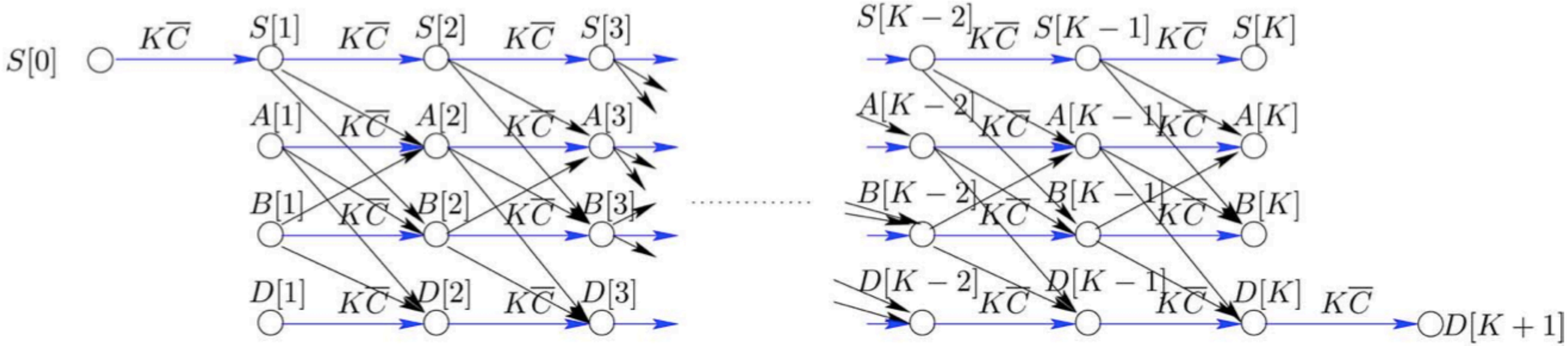
- NNC properly extends classic network coding
- SNNC achieves same rates
- Relation to Monday’s talk:
  - theory was based on layered networks so that non-layered networks require “time expansion”
  - layered analysis is useful, but is not needed

Layered Networks vs. General Networks



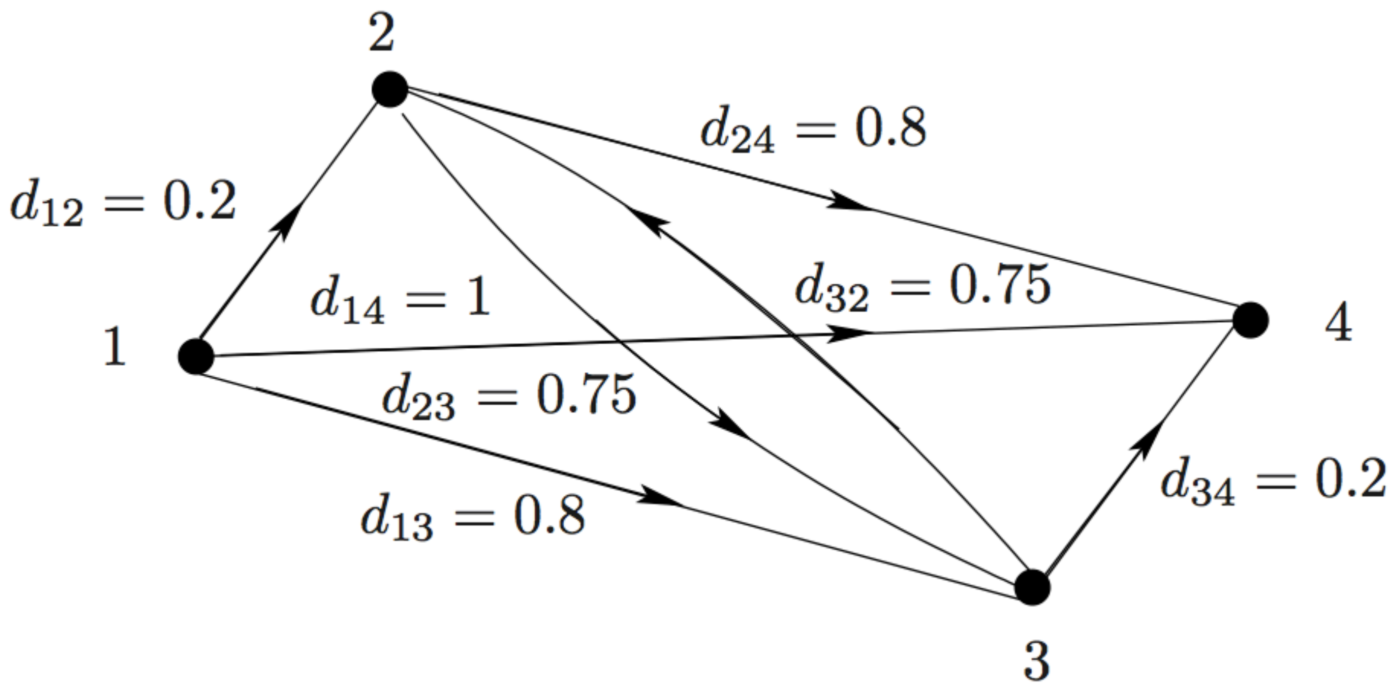


- Time-unfolded graph to get a layered network:

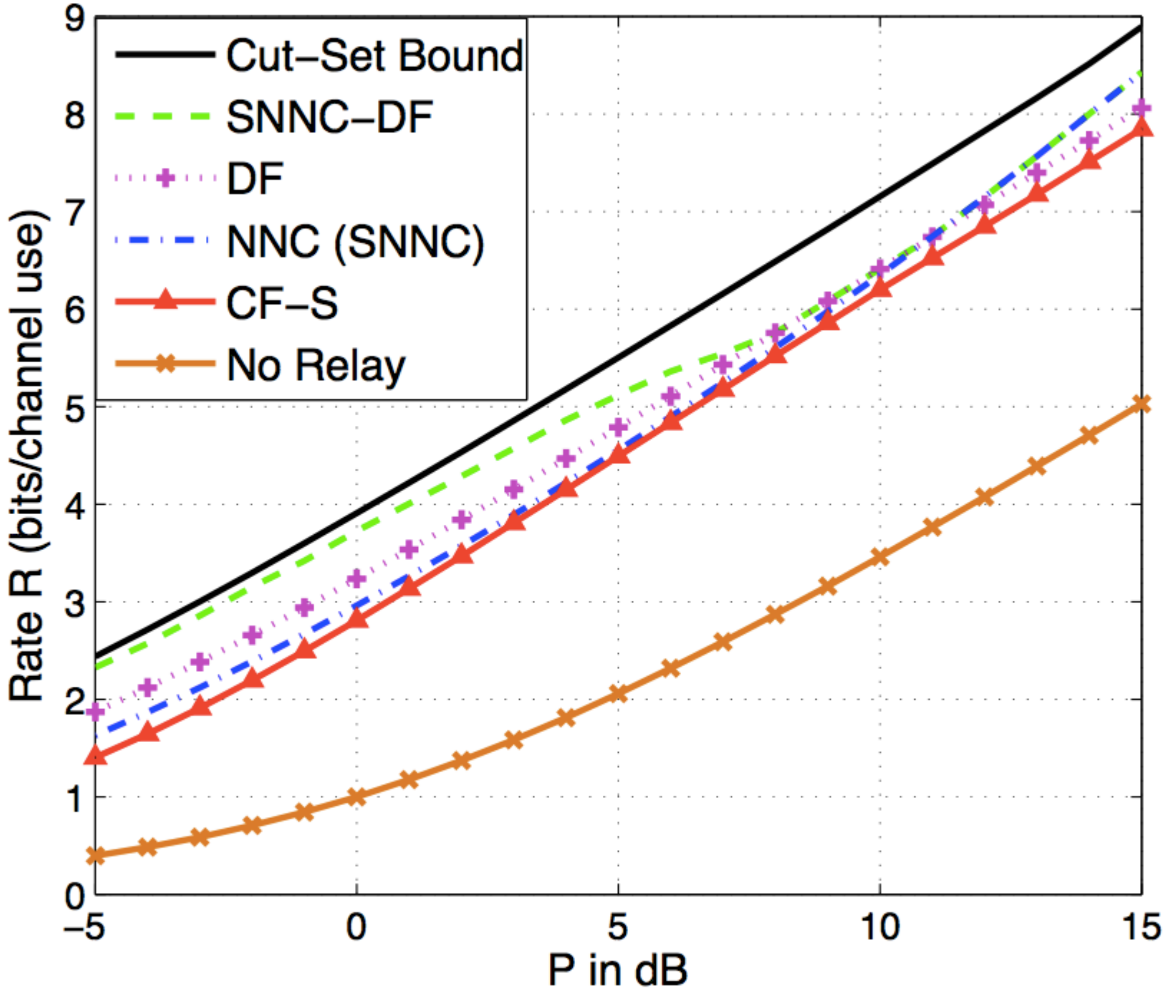


Experiment with 2 Relays (Full Duplex)

- Source (node 1), Relays (nodes 2 and 3), Destination (node 4)
- AWGN, unit-variance noise, attenuation exponent 3
- Common, per-node, per-symbol power constraint

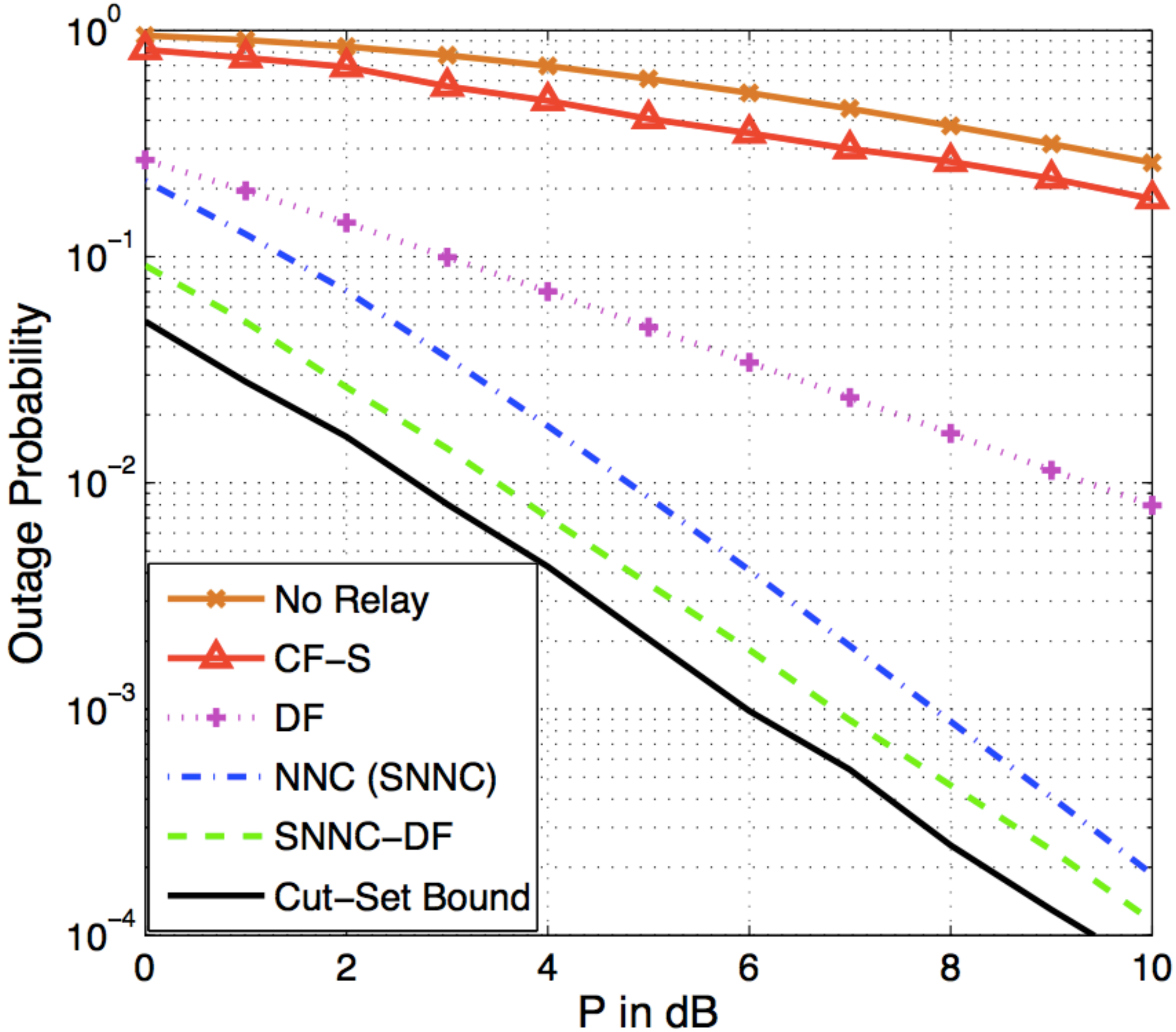


Experiment with 2 Relays (continued)



Experiment (cont'd)

- Attenuation exponent 3, slow Rayleigh fading, Gaussian noise
- Per-node power: common power constraint
- Rate target =2 bit/symbol
- SNNC gains 1 dB over NNC

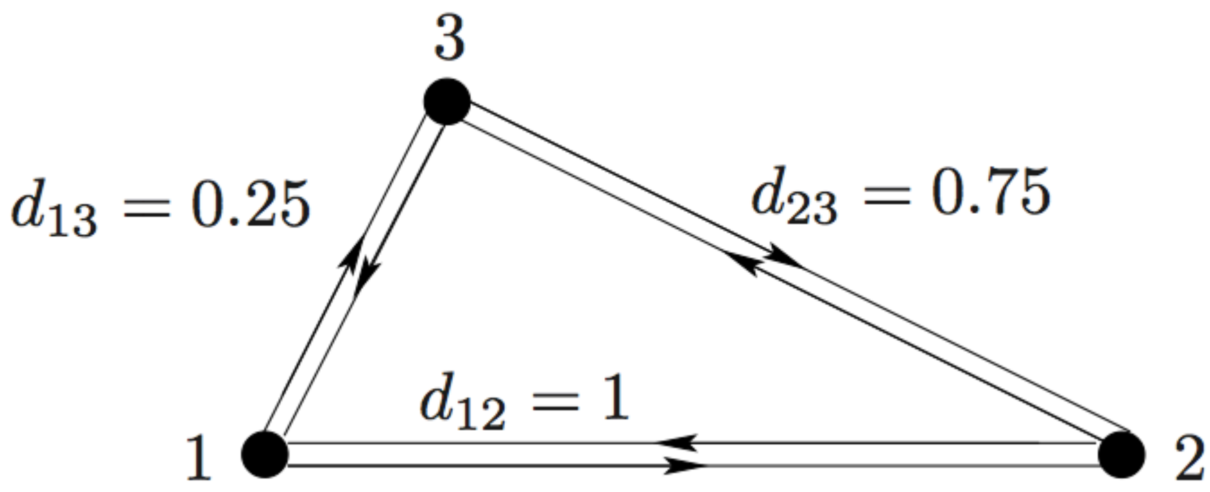


Discussion* (1 Source/Many Relays)

- $R_S < \max \min_{(S,\hat{S})} I(X_S; \hat{Y}_{\hat{S}} Y_d | X_{\hat{S}} T) - I(Y_S; \hat{Y}_S | X_S X_{\hat{S}} Y_{\hat{S}} Y_d T)$
- Deterministic (e.g. classic) networks: choose $\hat{Y}_i = Y_i$ and achieve cut-set bound with independent inputs
- Gaussian networks: choose $\hat{Y}_k = Y_k + \hat{Z}_k$, $\hat{Z}_k \sim CN(0,N)$, optimize N, to get within 0.63|V| bits of the cut-set bound (a true upper bound with dependent inputs)
- Can use short messages and multi-hop/backward decoding to enable DF and per-block processing
- Results extend to many sources & many relays

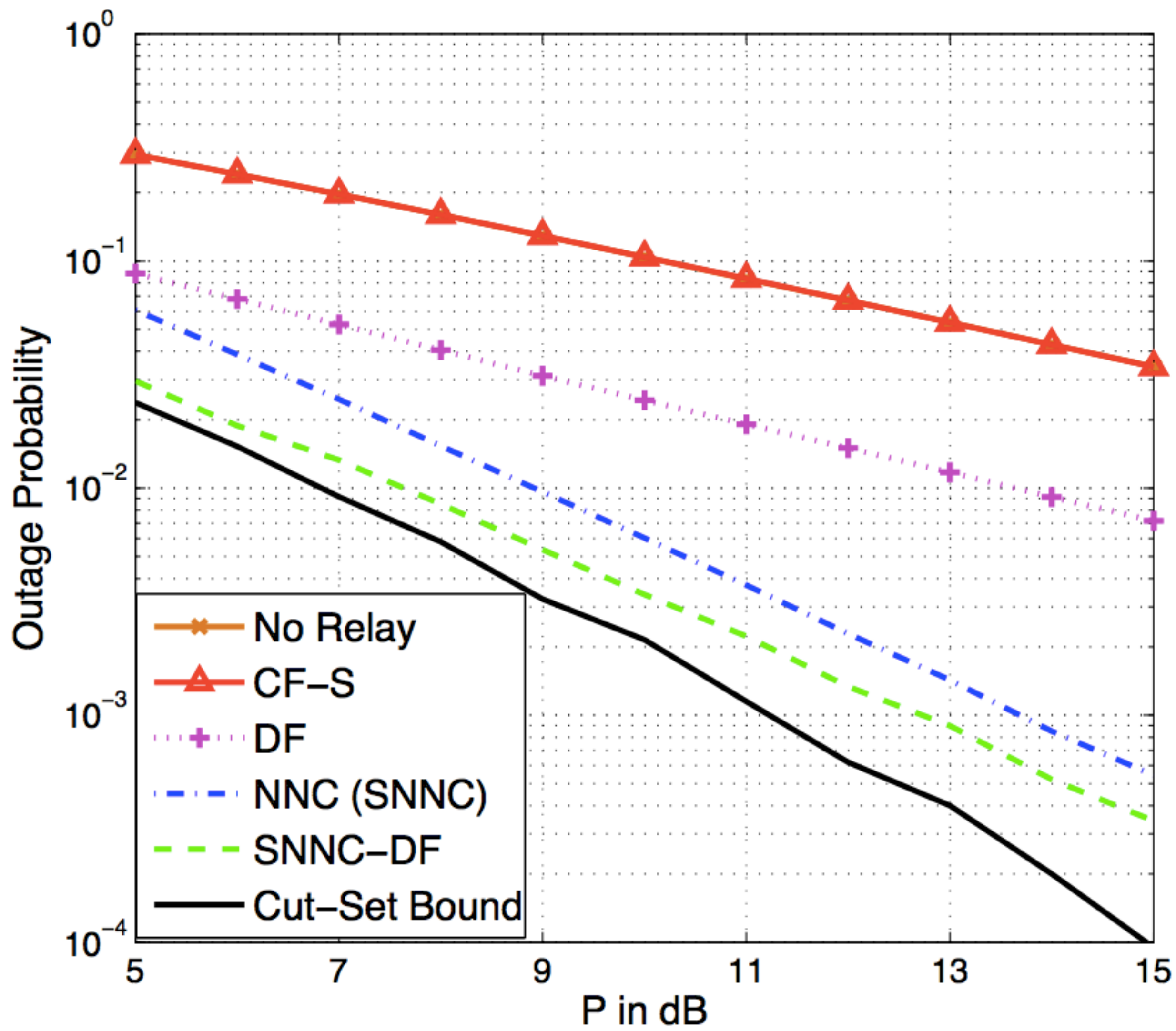
Experiment with 2 Sources, 1 Relay (Full Duplex)

- 2 Sources (nodes 1 and 2), 1 Relay (node 3)
- AWGN, unit-variance noise, attenuation exponent 3
- Per-node, per-symbol power constraint, P_1=5P, P_2=2P, P_3=P



Experiment

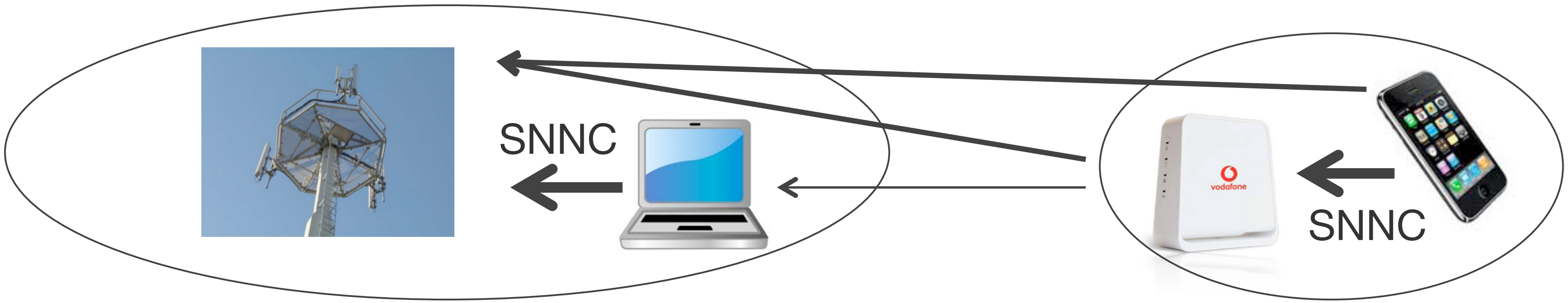
- Attenuation exponent 3, slow Rayleigh fading, Gaussian noise
- Per-node power: common power constraint
- Rate target 1=2 bit/symbol
  Rate target 2=1 bit/symbol
- SNNC gains 1-2 dB over NNC



Application Question

Does SNNC have a practical future?

- relays can operate in a distributed and autonomous fashion
- achieves the “multi-output” gains of MIMO
- SNNC with DF achieves “multi-input” gains of MIMO
- method applies to more than radio, e.g., classic & optical networks
- Difficulty and Research: how to design practical codes and decoders?



Receiver Cooperation

Transmitter Cooperation

Extra Slides

Proof* of Equivalence for 1 Source/Many Relays

- Fix a coding distribution. Let V be the set of relays.
  Let $S \subseteq T \subseteq V$ and $\hat{S}$ be the complement of S in T. Define

$$R_T(S) = I(X_1 X_S; Y_{\hat{S}} Y \mid X_{\hat{S}}) - I(Y_S; \hat{Y}_S \mid X_1 X_T Y_{\hat{S}} Y)$$
$$Q_T(S) = I(X_S; Y_{\hat{S}} Y \mid X_1 X_{\hat{S}}) - I(Y_S; \hat{Y}_S \mid X_1 X_T Y_{\hat{S}} Y)$$

- QF/NNC bounds: **$R \le \max_T \min_S R_T(S)$** **(1)**
- Backward decoding: **T must satisfy $0 \le Q_T(S)$ for all $S \subseteq T$** **(2)**
- Suppose (2) is violated for some S. Then for all B with $S \subseteq B \subseteq T$ we have $R \le R_T(B) < R_T(B) - Q_T(S) = R_{T \setminus S}(B \setminus S)$
- This means the destination can treat the X_k with $k \in S$ as noise
- Repeat argument until all bounds (2) satisfied
- Proof method generalizes to many sources (ISIT 2012)

* Kramer-Hou (ITW 2011)