5th Euro-NGI Conference on Next Generation Internet Networks
Aveiro, Portugal, June 2009

Network Information Theory

Principles and Applications

João Barros

Instituto de Telecomunicações
Faculdade de Engenharia
Universidade do Porto

Faculdade de Engenharia da Universidade do Porto

Instituto de Telecomunicações – Porto

Shannon Communications Lab @ FEUP & IT Porto

Faculty / Post-Doctoral Researchers

Prof. João Barros	(PhD TU München in 2004, Fulbright at Cornell Univ., Sab. MIT, JB)
Dr. Fausto Vieira	(PhD UPC Barcelona in 7/2008)
Dr. Tiago Vinhoza	(PhD PUC Rio in 8/2008)
Dr. Ian Marsh	(PhD KTH in 6/2009)

PhD students

Rui Costa	(Lic. in Math., MSc in CS, PhD Student)
Sergio Crisóstomo	(MSc in ECE, FCT PhD Student in CS at Porto/Klagenfurt,)
Mari Nistor	(MSc in ECE. PhD Student, MAP-Tele,)
Luísa Lima	(Lic. in ERSI, FCT PhD Student in CS at Porto/MIT,)
Gerhard Maierbacher	(MSc in ECE at TUM, FCT PhD Student in CS at Porto)
Paulo Oliveira	(Lic. in ERSI, FCT PhD Student in CS at Porto/MIT)
João Paulo Vilela	(Lic. in ERSI., currently MSc in CS, PhD Student in 2007,)
Rui Meireles	(Lic. in CS, PhD student MAP/CMU)
Mate Boban	(Lic. in ECE, PhD student MAP/CMU)
Pedro Gomes	(Lic. in ECE, PhD student MAP/CMU)
Saurabh Shintre	(MSc IIT Bombay)

Information Theory

Multi-user Information Theory

Interplay IT and Estimation Theory

Rate-Distortion Theory

Network coding

Information Processing

Scalable Distributed Compression

Distributed Inference

Secure Quantization

Distributed Storage

Information Networks

Integration in Heterogeneous Networks

Data Gathering in Sensor Networks

Vehicular Ad-hoc Networks

Small-World Networks

Information Security

Information-Theoretic Security

Cooperatively Secure Routing

Secure Network Coding

Secret Key Agreement

Funding

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

(Portuguese NSF)

- WITS: Wireless Information-Theoretic Security
- CALLAS: Calculii and Languages for Sensor Networks
- SeNeCom: Secure Network Communications
- 8 PhD Fellowships

European Union

- DYNAMO: Foundations and Algorithms of Dynamic Networks
- DAIDALOS I and II: Integration of Heterogenous Networks
- N-CRAVE: Network Coding in Highly Volatile Networks
- EURO-NFI: Network of Excellence on Future Internet
- WiPhySec: Wireless Physical-Layer Security
- NeCo: Network Coding Opportunities
- SENECA: Secure Network Coding
- NET-PEEC: Network Coding for P2P

FUNDAÇÃO
LUSO-AMERICANA

DAAD

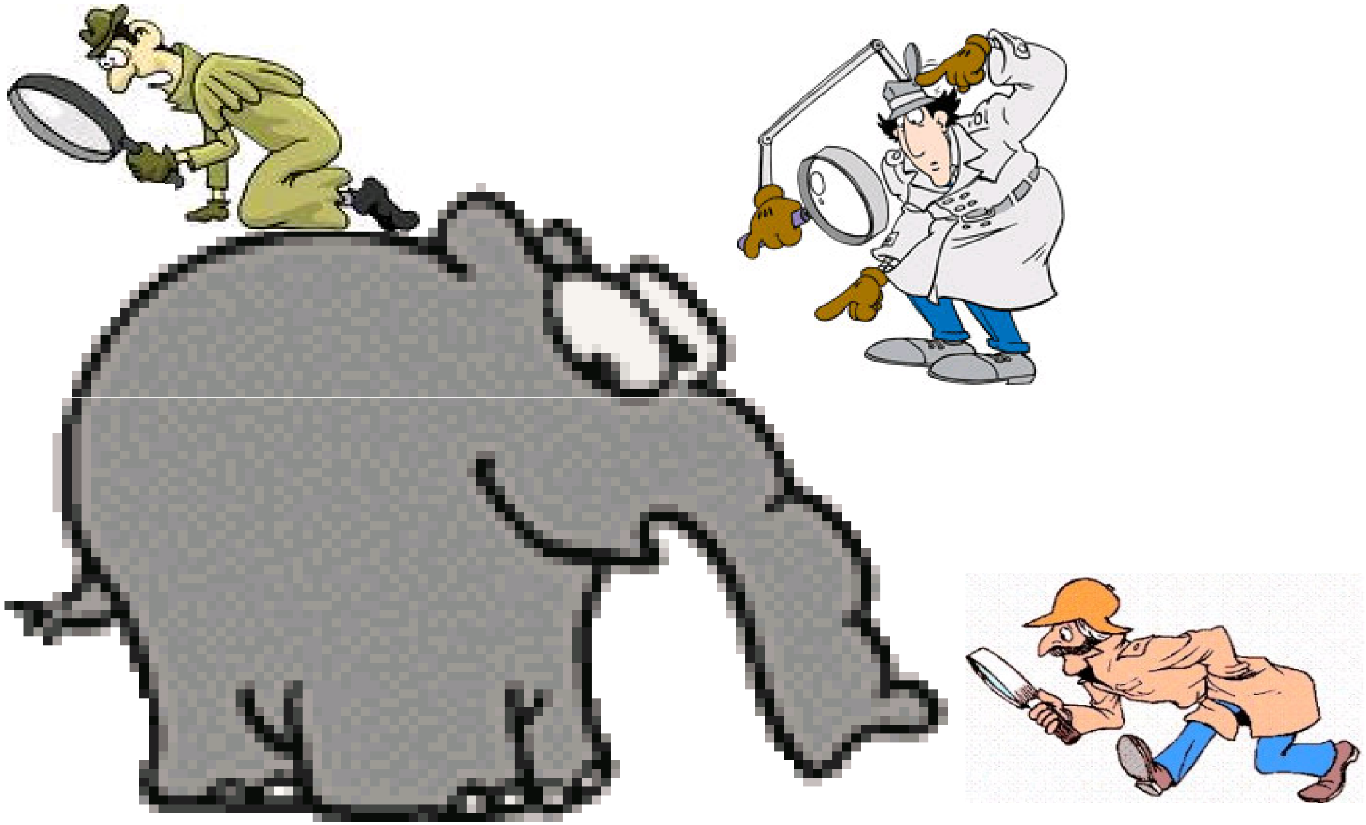
NTT
Do Co Mo
DoCoMo Euro-Labs

Information and Communication Technologies Institute
Carnegie Mellon | PORTUGAL
AN INTERNATIONAL PARTNERSHIP

MIT Portugal

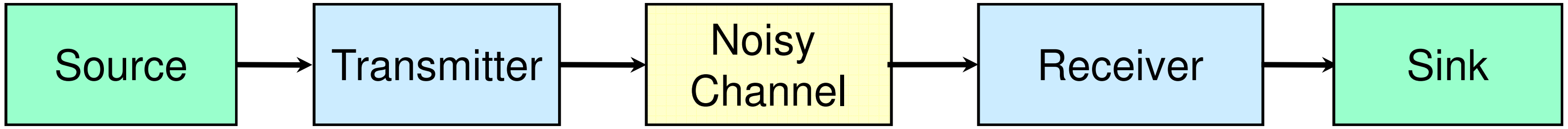
Network Information Theory

The eternal problem…



Coding Theorems

Claude Shannon:



Source

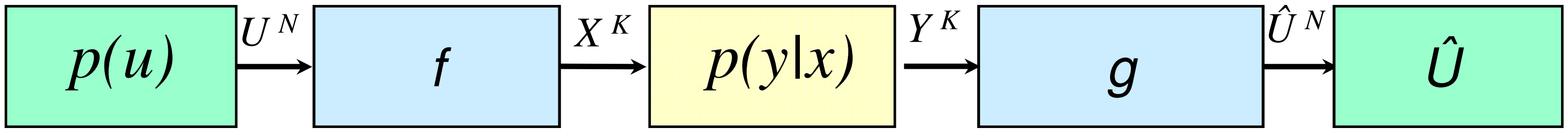
Transmitter

Noisy Channel

Receiver

Sink

Mathematical Model



$p(u)$

U^N

f

X^K

$p(y|x)$

Y^K

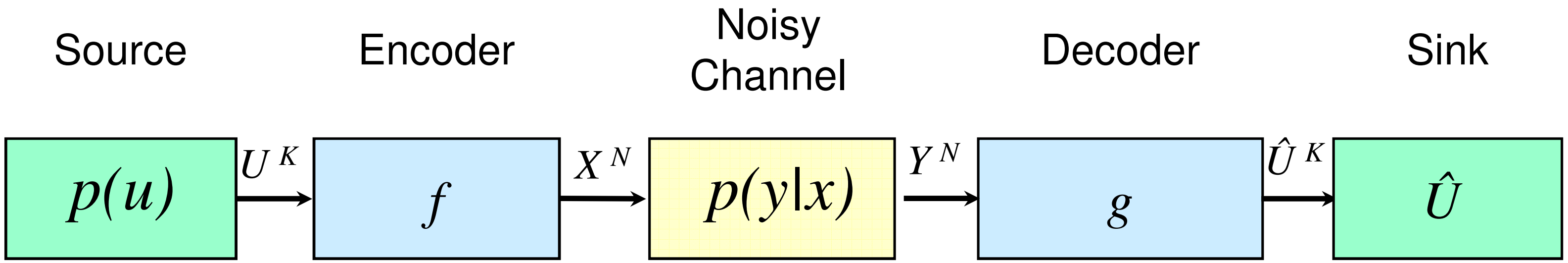
g

$\hat{U}^N$

$\hat{U}$

(almost) perfect reconstruction

Basic Definitions



Coding rate: $R = K / N$

Entropy: $$H(U) = \sum_{i=1}^{L} p(u_i) \log_2 \frac{1}{p(u_i)}$$

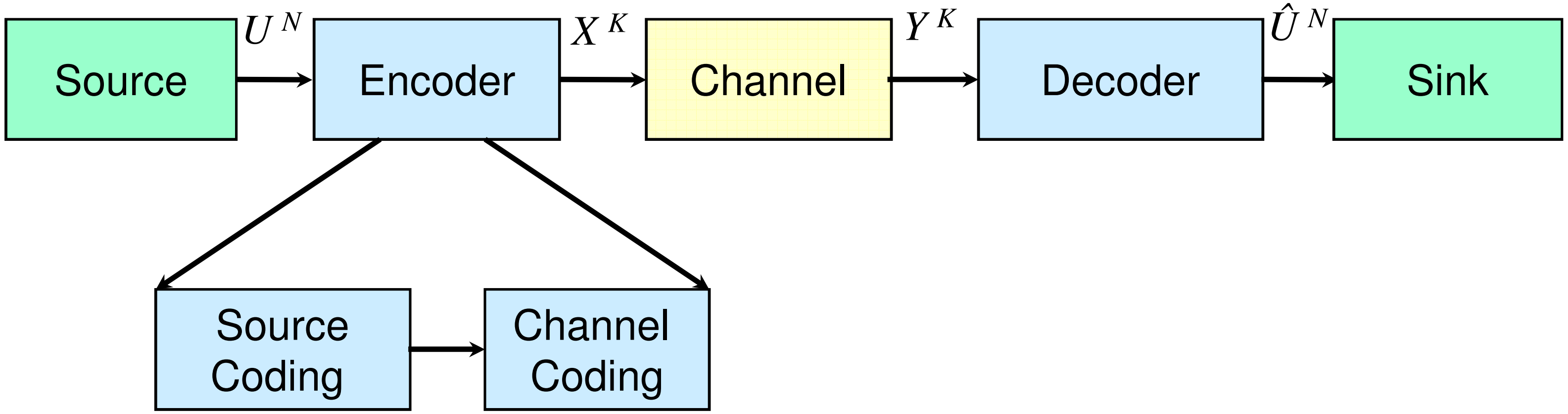
Kullback-Leibler Distance:

$$D(p \,\|\, q) = \sum_{i=1}^{L} p(u_i) \log_2 \frac{p(u_i)}{q(u_i)}$$

Mutual Information:

$$I(X;Y) = D(p(x_i, y_j) \,\|\, p(x_i)\, p(y_j))$$

Fundamental Theorems



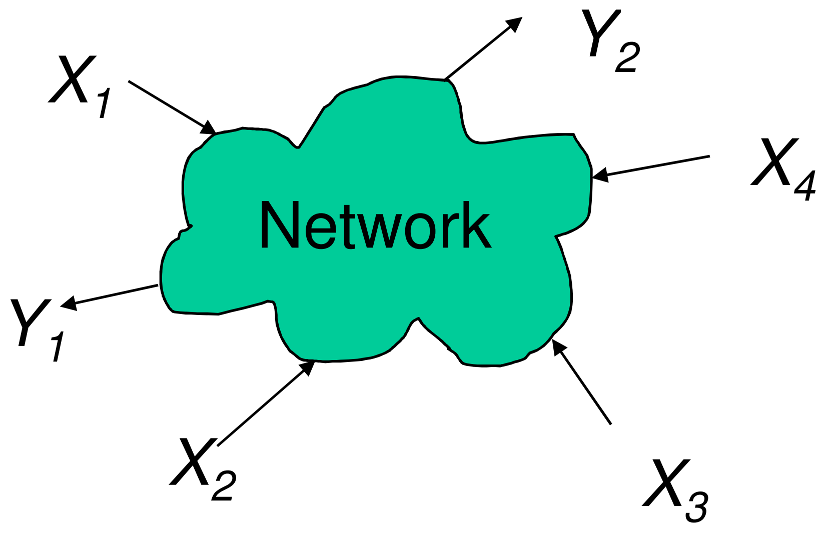
Source Coding: $R > H(U)$

Channel Coding: $R < I(X;Y) < C$

Source/Channel Separation: $H(U) < I(X;Y) < C$

Network Information Theory

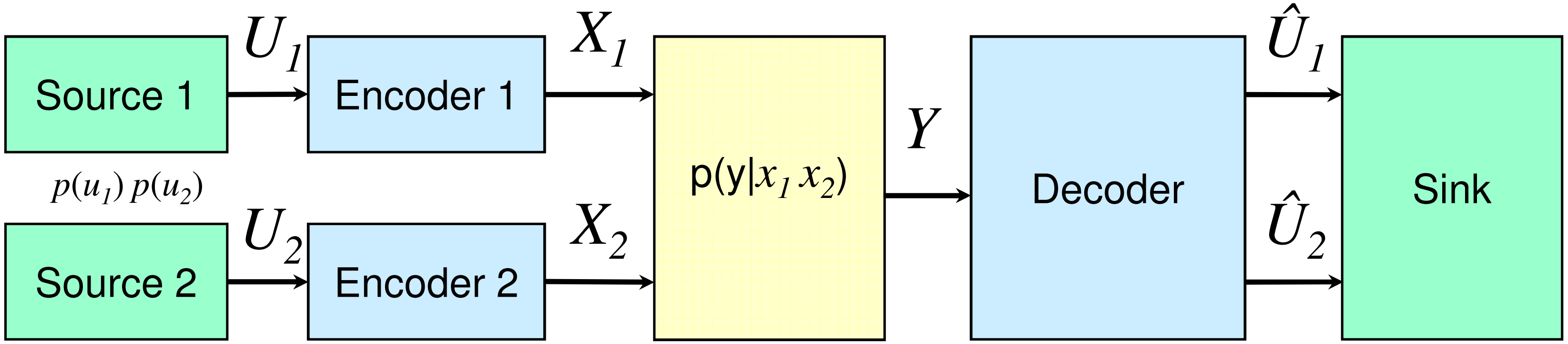
- Information theory has been very successful at characterizing the fundamental limits of point-to-point communications.
- How about ***networks***?



- Interference
- Cooperation
- Feedback

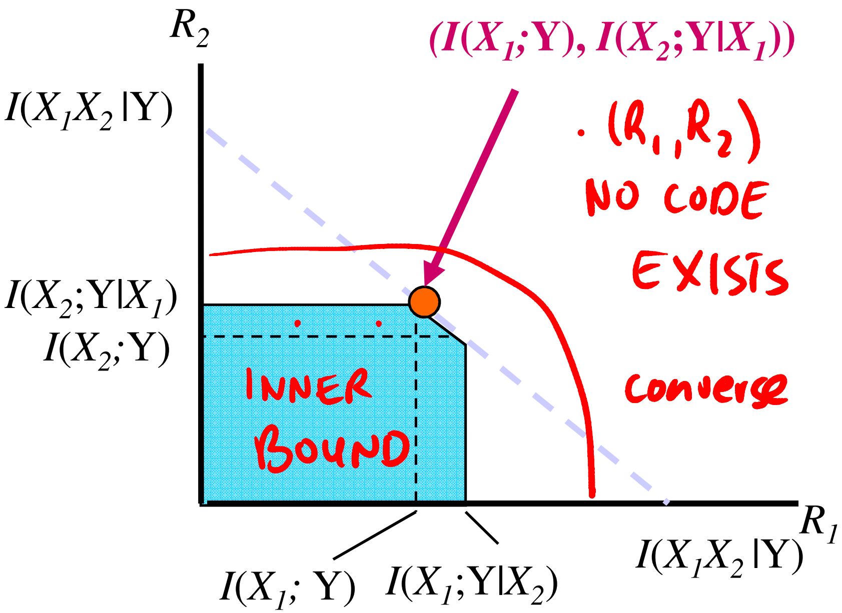
?

Multiple Access Channel

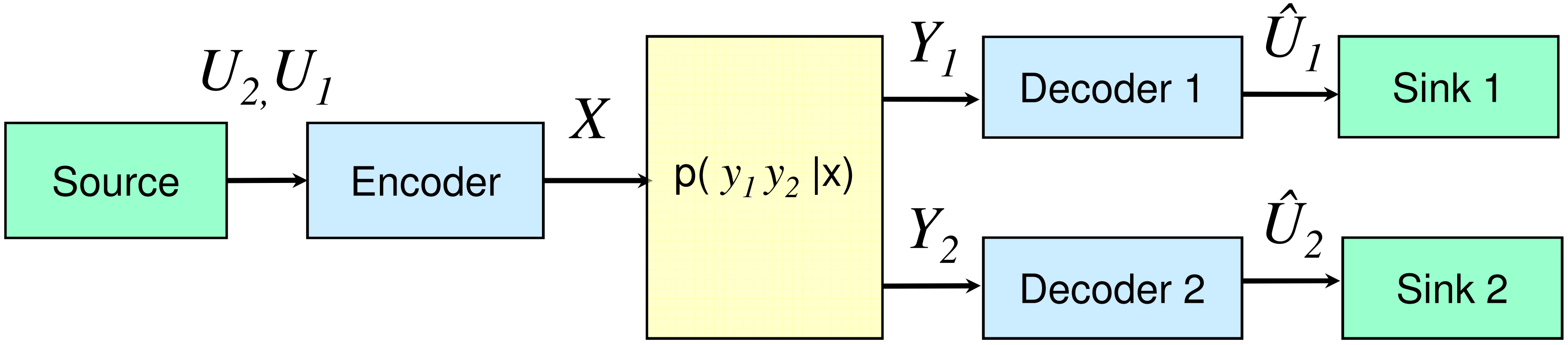


- Model for uplink in wireless networks

- Complete solution is known.



Broadcast Channel



- Model for downlink in wireless networks
- Superposition Coding
- Solution only known for the degraded case with

  Markov chain X-Y_1-Y_2
- Auxiliary r.v. W

$$R_1 \leq I(W;Y_2)$$

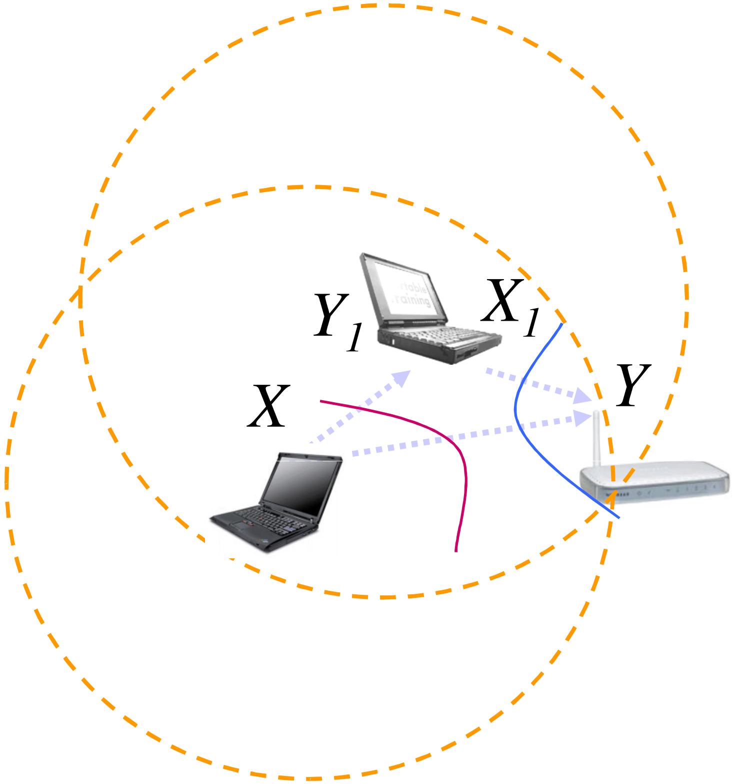
$$R_2 \leq I(X;Y_2|W)$$

Relay Channel

- User Cooperation
- Physically Degraded
- Max-flow Min-cut Bound

$$C = \sup_{p(x,x_1)} \min\{\underline{I(X,X_1;Y)}, \underline{I(X;Y,Y_1 | X_1)}\}$$

- General solution still unknown



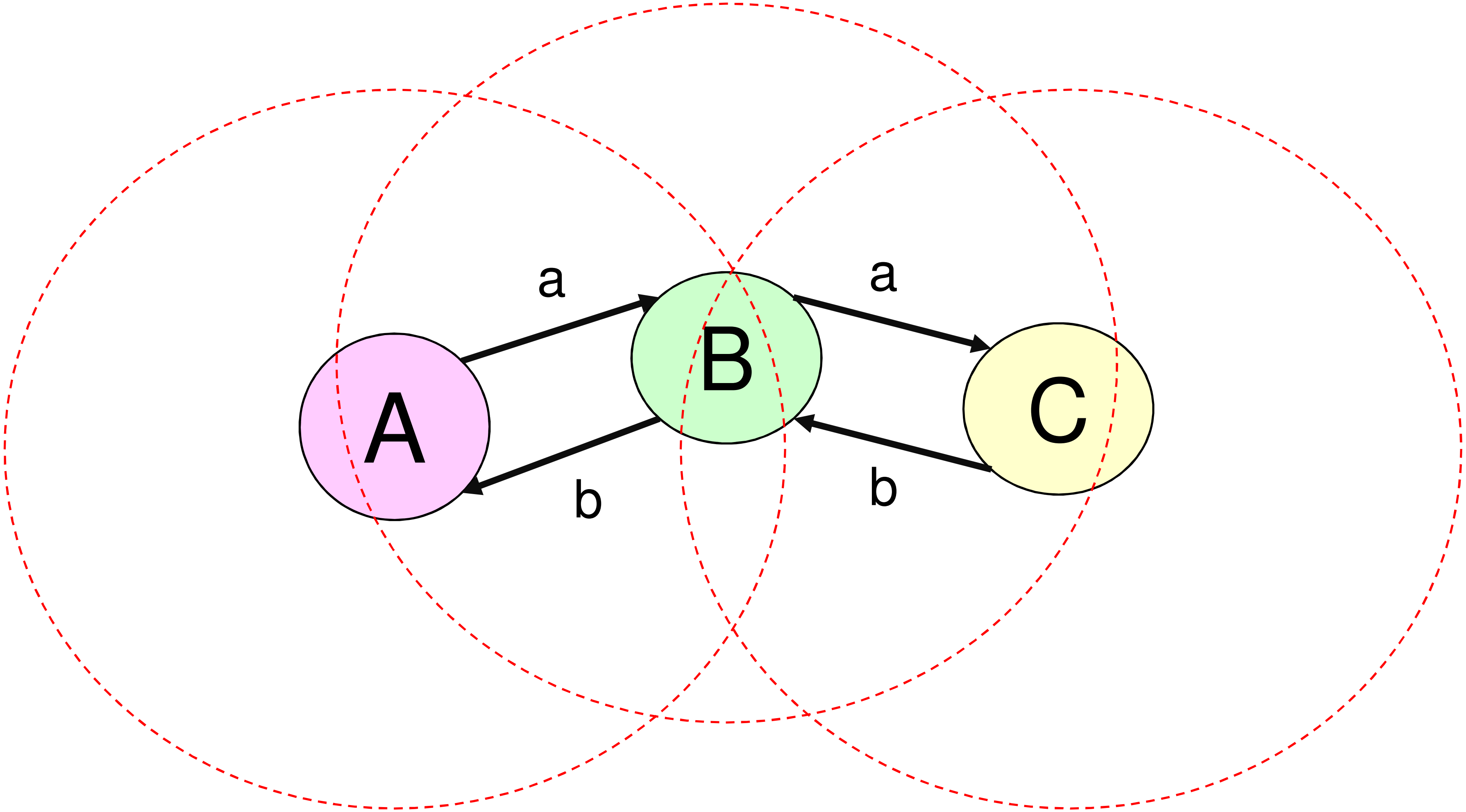
Network Coding

Store-and-Forward versus Network Coding

- In today’s networks, information is viewed as a commodity, which is transmitted in packets and forwarded from router to router pretty much as water in pipes or cars in highways.

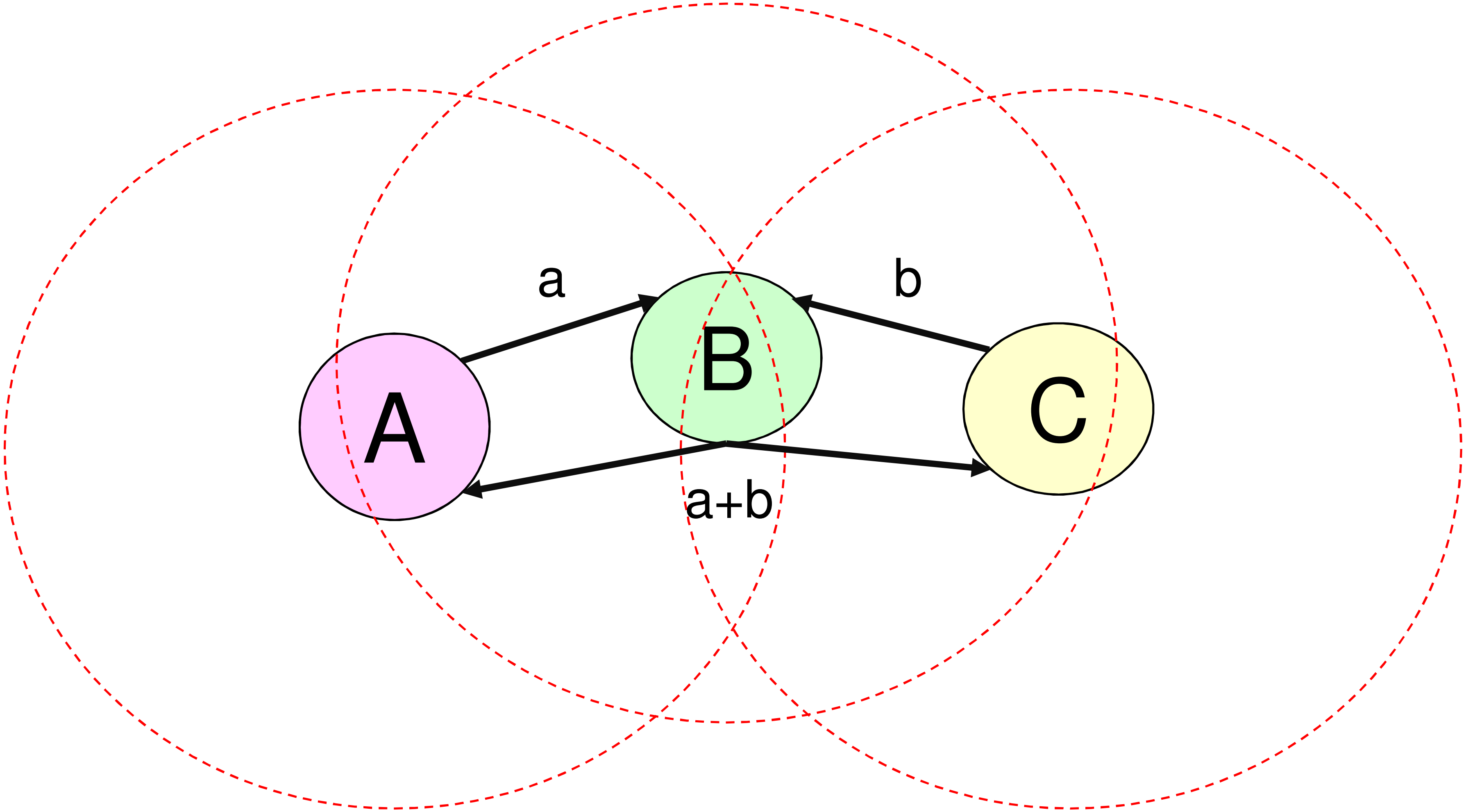
- In contrast, network coding allows intermediate nodes to mix different information flows by combining different input packets into one or more output packets.

A simple three-node example



In the current networking paradigm we require 4 transmissions.

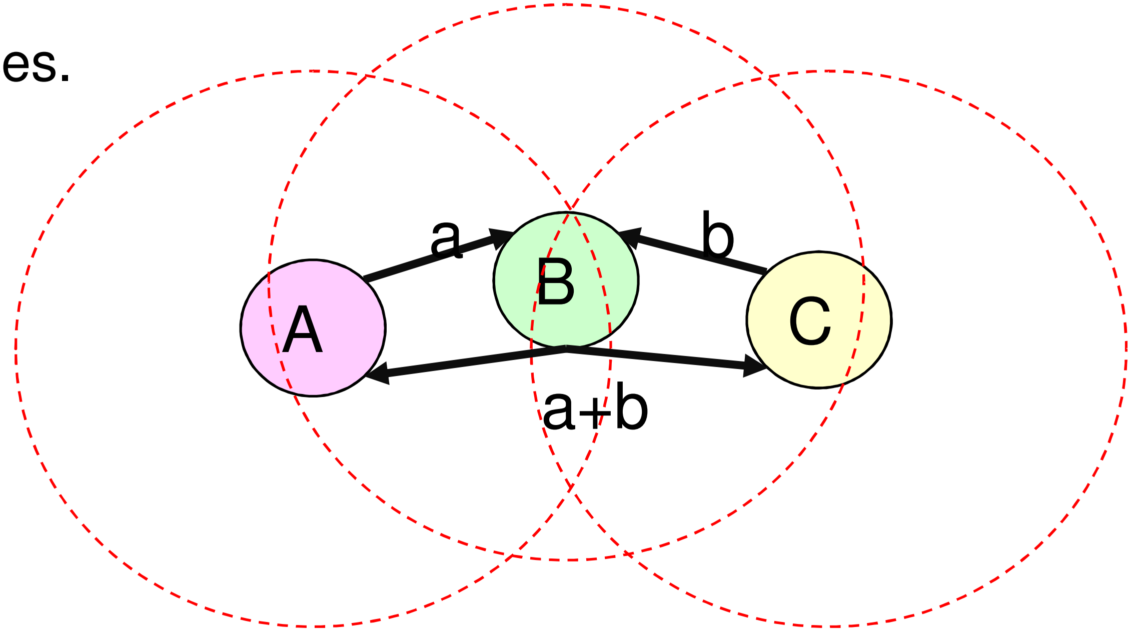
Network Coding



With network coding we require **only 3** transmissions.

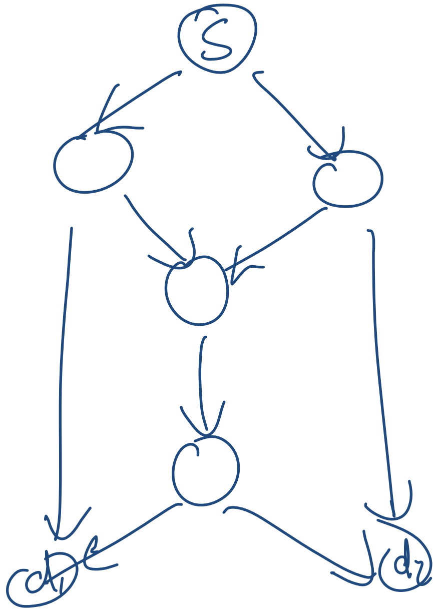
Basic Principles of Network Coding

- To receive the requested data, the destination node does not require specific packets.
- It is sufficient to receive a sufficient number of packets, from which the destination node can recover (or decode) the data.
- This allows us to trade off computation and communication resources.



Foundations of Network Coding

- **[Ahlswede, Cai, Li and Yeung, 2000]**
  - Max-flow min-cut capacity of a general multicast network can only be achieved by allowing intermediate nodes to mix different data flows
- **[Li, Yeung and Cai, 2003]**
  - Linear network coding sufficient to reach multicast capacity of a network
- **[Koetter and Médard, 2003]**
  - Algebraic framework
- **[Ho et al, 2003]**
  - Randomized network coding



Algebraic Framework for Network Coding

- Binary vector of length m: element in F_{2^m}

- Random processes at nodes

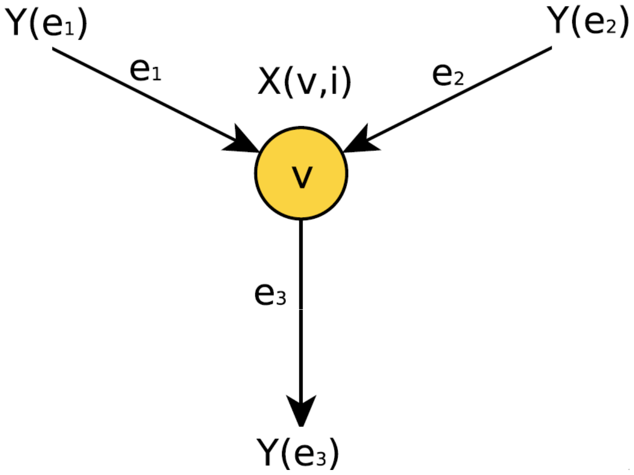
$$Y(e_3) = \sum_i \alpha_i X(v,i) + \sum_{j=1,2} \beta_j Y(e_j)$$

- Transfer matrix

$$\underline{z} = \underline{x}M \qquad M = A(I-F)^{-1}B^T$$

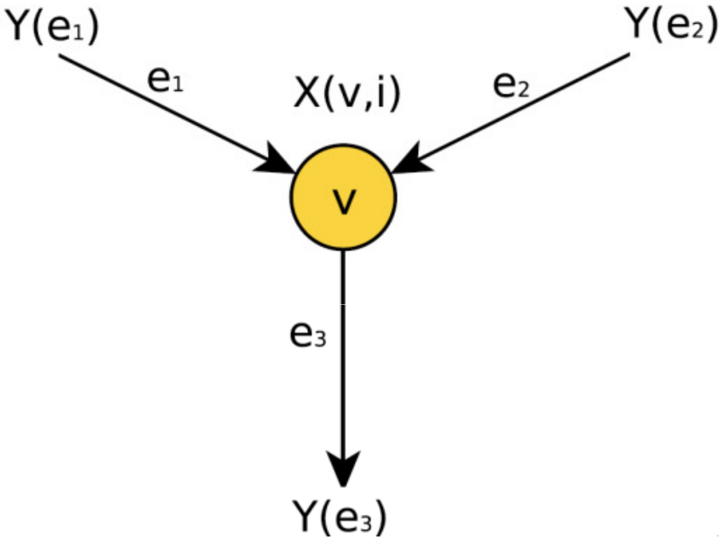
- Generalized MIN-CUT MAX-FLOW Condition

$$|M| \neq 0$$



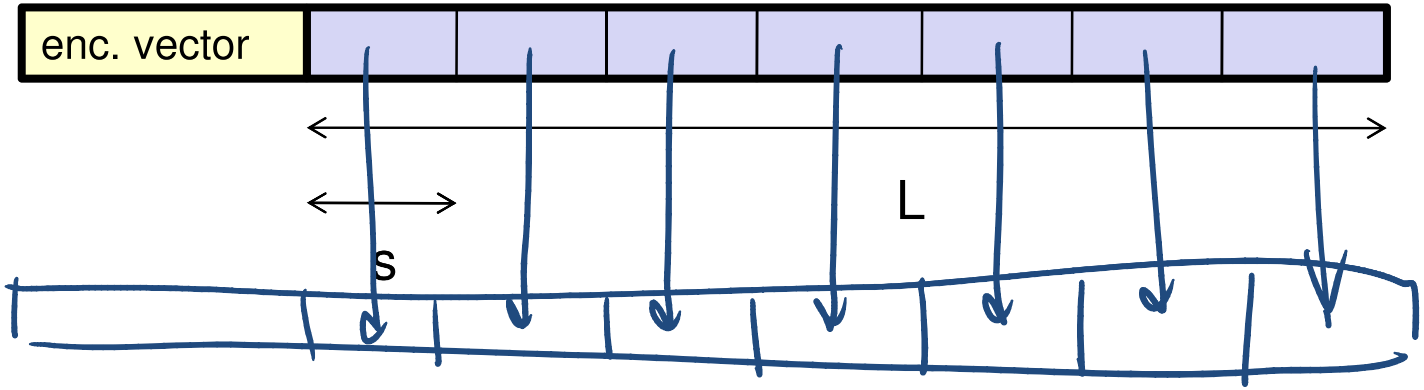
Random Linear Network Coding

- Coefficients chosen independently at random
- With high probability, transfer matrix is non-singular
- Multicast problem, *some or all* coefficients chosen independently from F_q, ***d*** receivers, η number of links with associated random coefficients
    - Probability that random code is valid is $\left(1-\frac{d}{q}\right)^{\eta}$



Packetized Network Coding

- Assume each packet carries *L* bits
- *s* consecutive bits can be viewed as a symbol in F_q

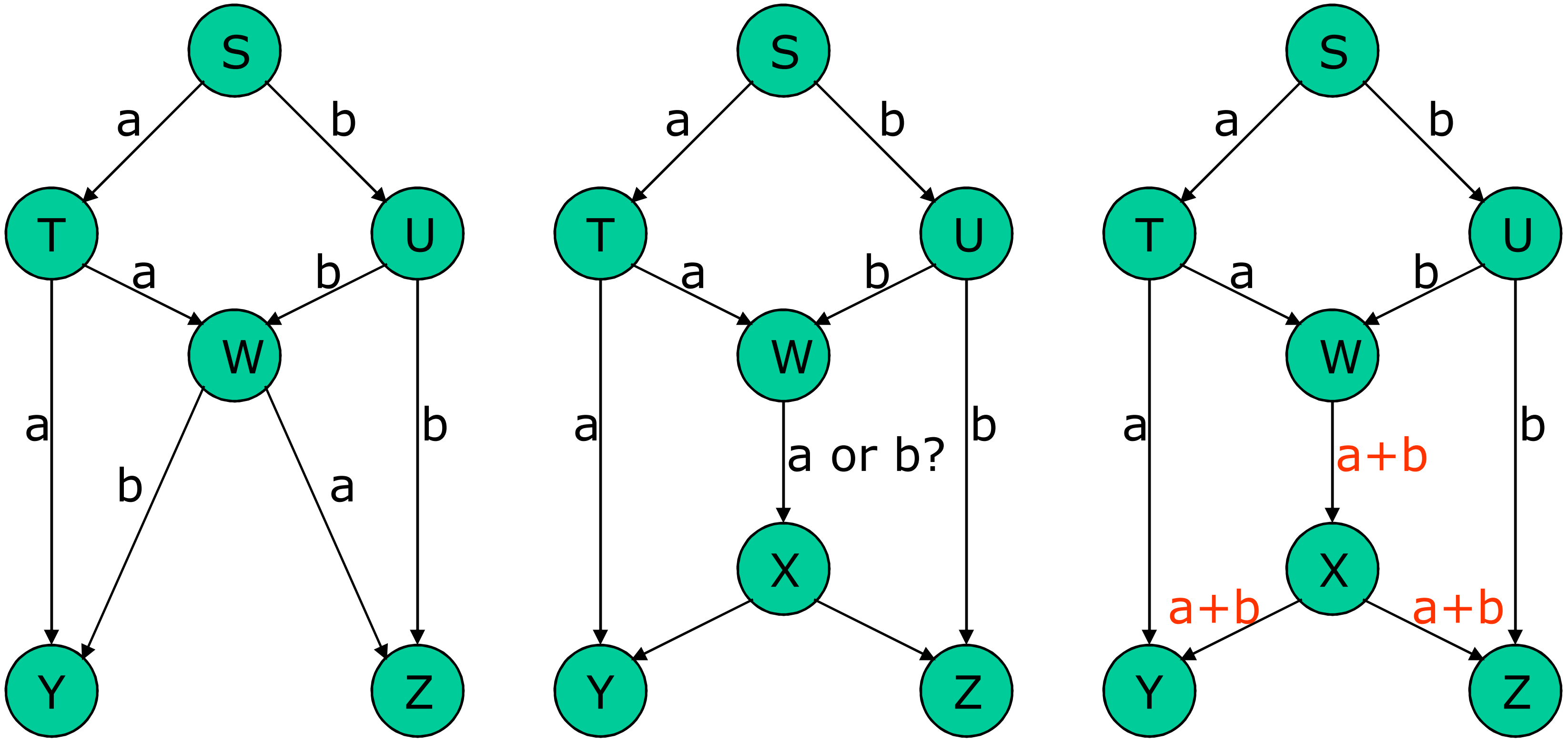


- Perform network coding on a symbol by symbol basis.
- Output packet also has length L.
- Send the coefficients (the “encoding vector”) in the header.
- Information is spread over multiple packets.

Practical Considerations

- **Encoding:** Elementary linear operations which can be implemented in a straightforward manner (with shifts and additions).
- **Decoding:** Once a receiver has enough linearly independent packets, it can decode the data using Gaussian elimination, which requires $\mathcal{O}(n^3)$ operations.
- **Generations:** To manage the complexity and memory requirements, we mix only generations with fixed number of packets and limit the field size. Each keeps a buffer sorted by generation number. Non-innovative packets are discarded.
- **Delay:** Since we must wait until we have enough packets to decode, there is some delay (not very significant, since we require less transmissions in many relevant scenarios)

Throughput benefits of network coding

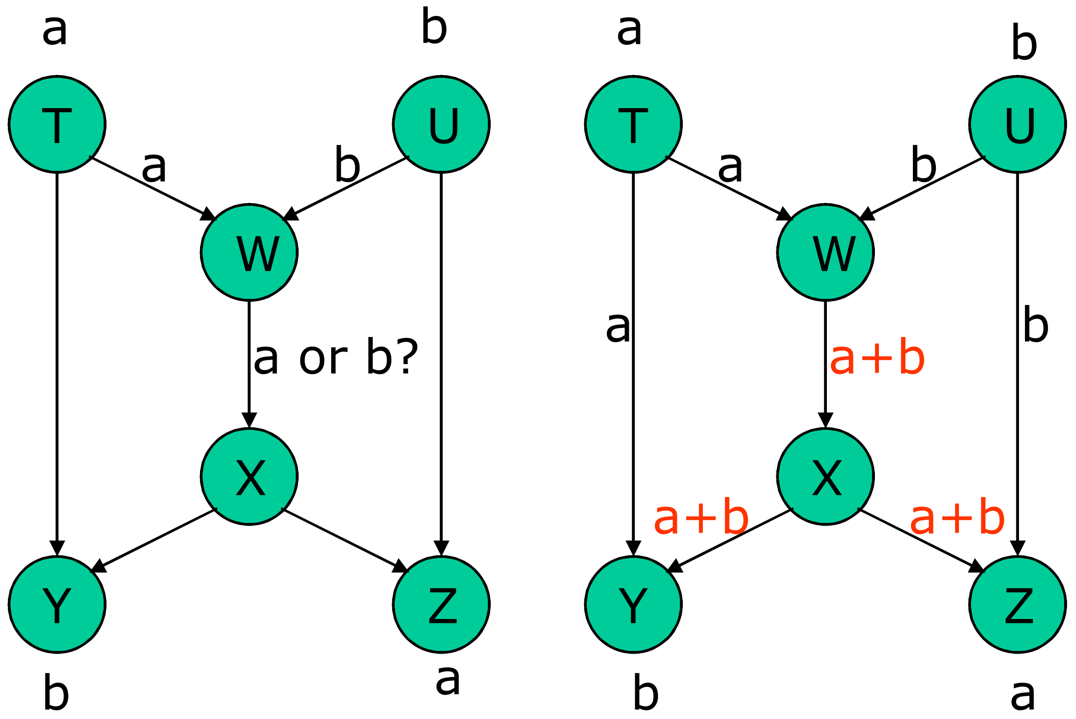


How can we send a and b to nodes Y and Z simultaneously?

(Linear)Network Coding

achieves Max-Flow bound in Multicast Networks

Works also for unicast sessions

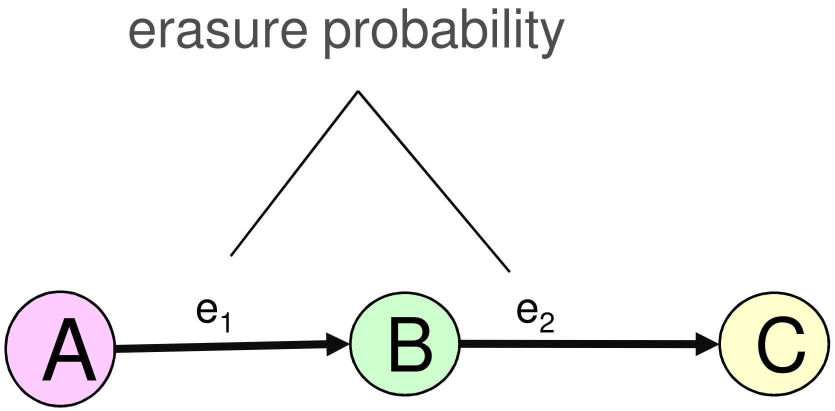


How can T send a to Z and U send b to Y simultaneously?

(Linear)Network Coding improves throughput,
and can also help with robustness to failures.

Reliability in the presence of erasures and errors

- State-of-the-art reliable communication over noisy links relies on:
  - Automatic Repeat Request (ARQ) techniques (at a price in delay)
  - Forward Error Correction (FEC) mechanisms (at a price in rate)
- Network coding can achieve optimum delay and rate.



- End-to-end FEC (no delay)

$$R \leq (1-e_1)(1-e_2)$$

- Decode-and- forward (delay)

$$R \leq \min\{(1-e_1),(1-e_2)\}$$

- Network Coding (no delay)

$$R \leq \min\{(1-e_1),(1-e_2)\}$$

Coupon Collector's Problem

- n nodes, O(n) messages
- every node should get n messages
- Centralized gossiping algorithm:

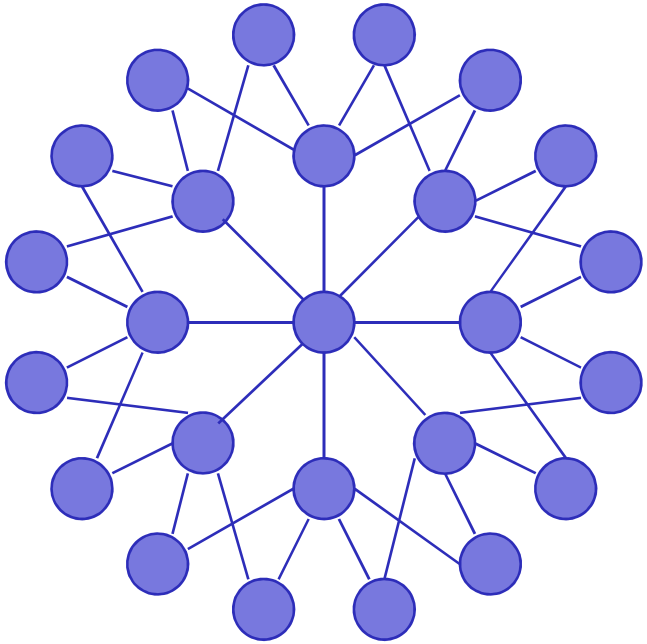
Θ(n) rounds with Θ(n) pairs of nodes exchanging one message per round.

- Decentralized gossiping algorithm:

Θ(n log n) rounds, because some messages become hard to collect.

- Random linear network coding:

all packets are "equal" and each node only needs to get n packets, therefore Θ(n) are sufficient – and still fully decentralized.



Other benefits

- **Simpler algorithms:** The multicast routing problem is NP-hard (packing Steiner trees), however with network coding there exist polynomial time algorithms.
- **Robustness:** Random network coding is completely decentralized and preserves the information in the network, even in highly volatile networking scenarios.

Applications of Network Coding

First real-life application in July 2007:

Microsoft Secure Content Downloader (a.k.a. Avalanche)

- **Distributed Storage and Peer-to-Peer:** robustness against failures in highly volatile networks;
- **Wireless Networks:** Information dissemination using opportunistic transmission;
- **Sensor Networks:** Data gathering with extremely unreliable sensing devices;
- **Network Management:** Assessing critical network parameters (e.g. topology changes and link quality)

Capacity of Random Networks

Small World Phenomena

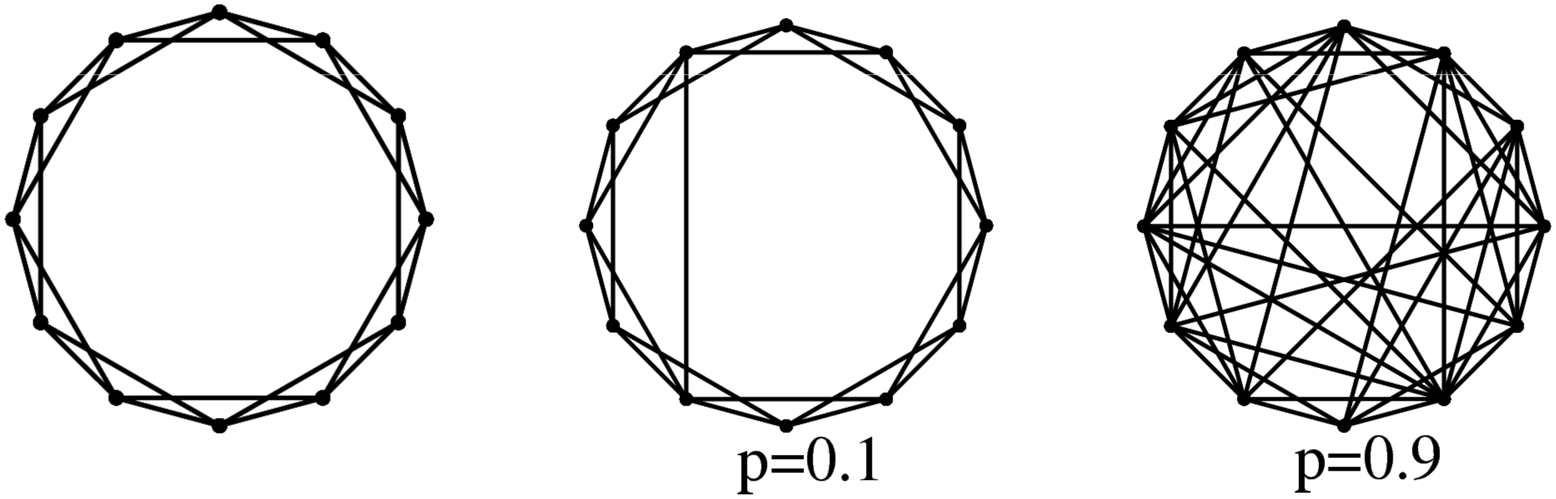
Many complex natural and man-made networks share the following common properties:

- Large networks ($n >> 1$)
- Sparse connectivity (avg degree $k << n$)
- No central node ($k_{max} << n$)
- Large clustering coefficient (larger than in random graphs of same size)
- Short average paths (~log n, close to those of random graphs of the same size)

Small World Models

Small-World Network (SWN) with Shortcuts

- Start with a k-connected ring lattice $(V_L,\ E_L)$
- Add each edge $e \notin E_L$ to the graph with probability *p*.

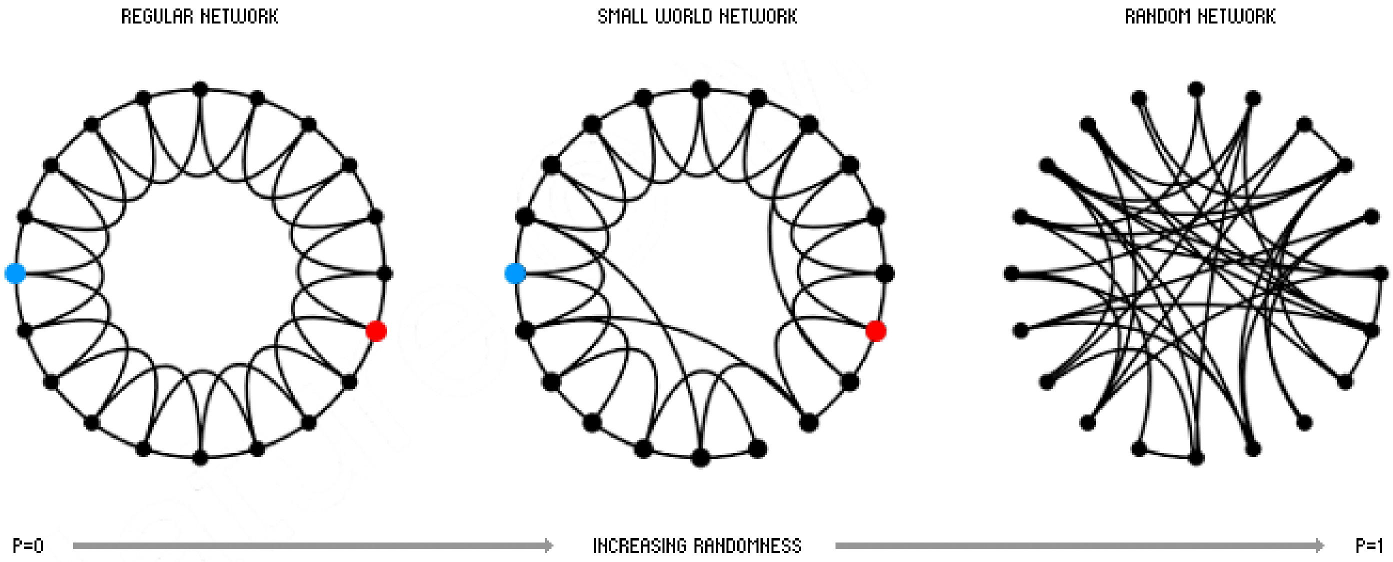


[Watts & Strogatz ’98, Newman & Watts ’99]

Rewiring

Small-World Network with Rewiring

- Start with a *k*-connected ring lattice;
- Choose an edge in the initial lattice;
- With probability *p,* rewire this edge to a random node;
- Repeat until all the edges have been considered once.



[Watts & Strogatz '98, Watts '99]

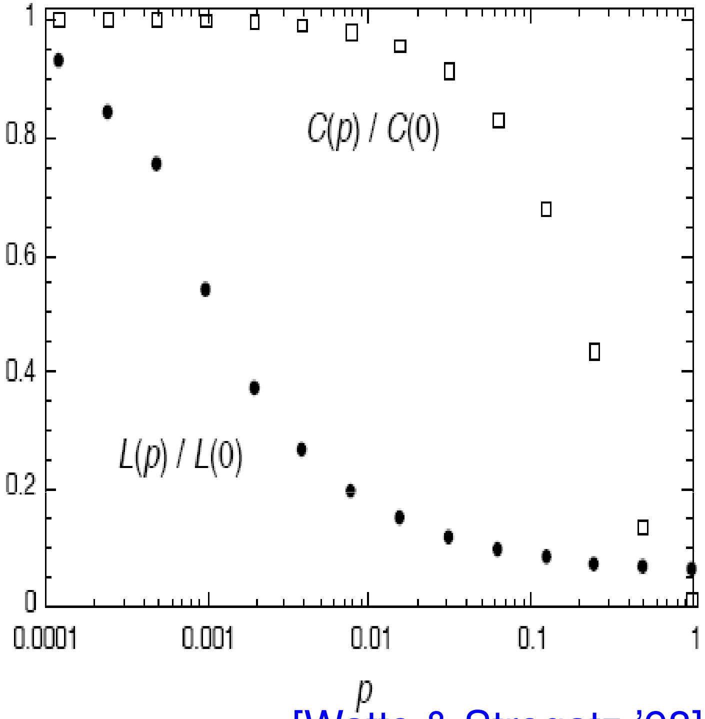
Fundamental Properties

- Characteristic Path Length

  (drops sharply)

- Clustering Coefficient

  (remains almost constant)



[Watts & Strogatz '98]

$$C = \frac{\sum_{v \in V} C(v)}{|V|}$$

with $C(v) = \dfrac{\#\, links\ between\ neighbours\ of\ \ v}{\#\, possible\ links\ between\ neighbours\ of\ \ v}$

▶ **Other parameters**: sparsity, degree distribution, *betweenness*,...

...

Why do we care about capacity of SWN?

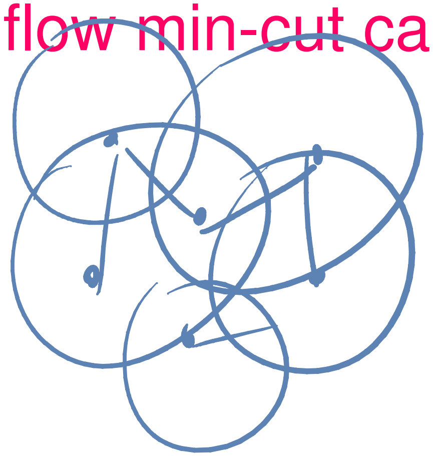
- The combination of strong local connectivity and long range shortcut links renders small-world topologies particularly well suited for
  - Resource discovery in mobile ad-hoc networks [Helmy'03]
  - Heterogenous networks [Reznik, Kulkarni, Verdu,'04]
  - Peer-to-peer communications [Manku, Naor, Wieder,'04]
  - Cellular wireless networks [Dixit, Yanmaz, Tonguz,'05]

- The notion of capacity plays a key role in many of the systems for which small world topologies are currently envisioned.
- Capacity and network information flow may help explain why small-world networks appear so frequently.

Max-Flow Min-Cut Capacity

- The max-flow min-cut bound [Ford,Fulkerson] gives the fundamental limits of information flow for:
  - one unicast session
  - multiple independent sources and one sink
  - correlated sources and one sink
  - multicast (with network coding)

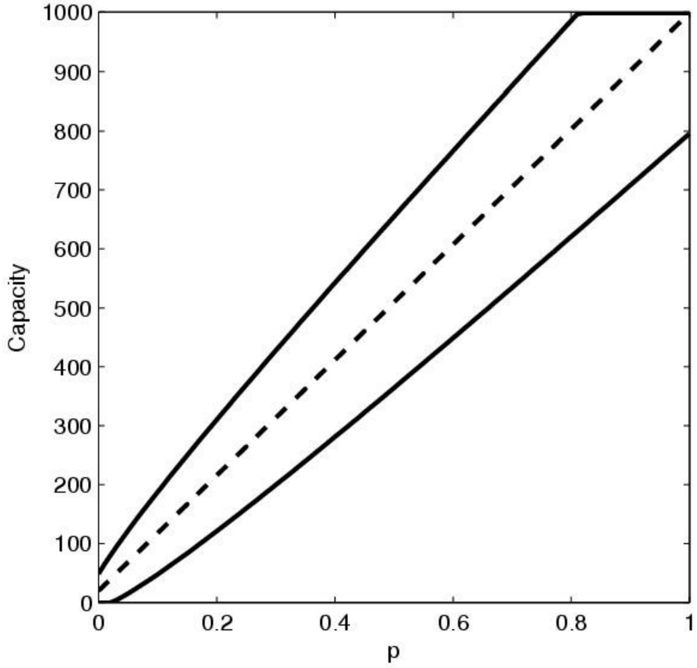
- We will be concerned with the max-flow min-cut capacity of small-world networks.



Capacity of SWN with added Shortcuts

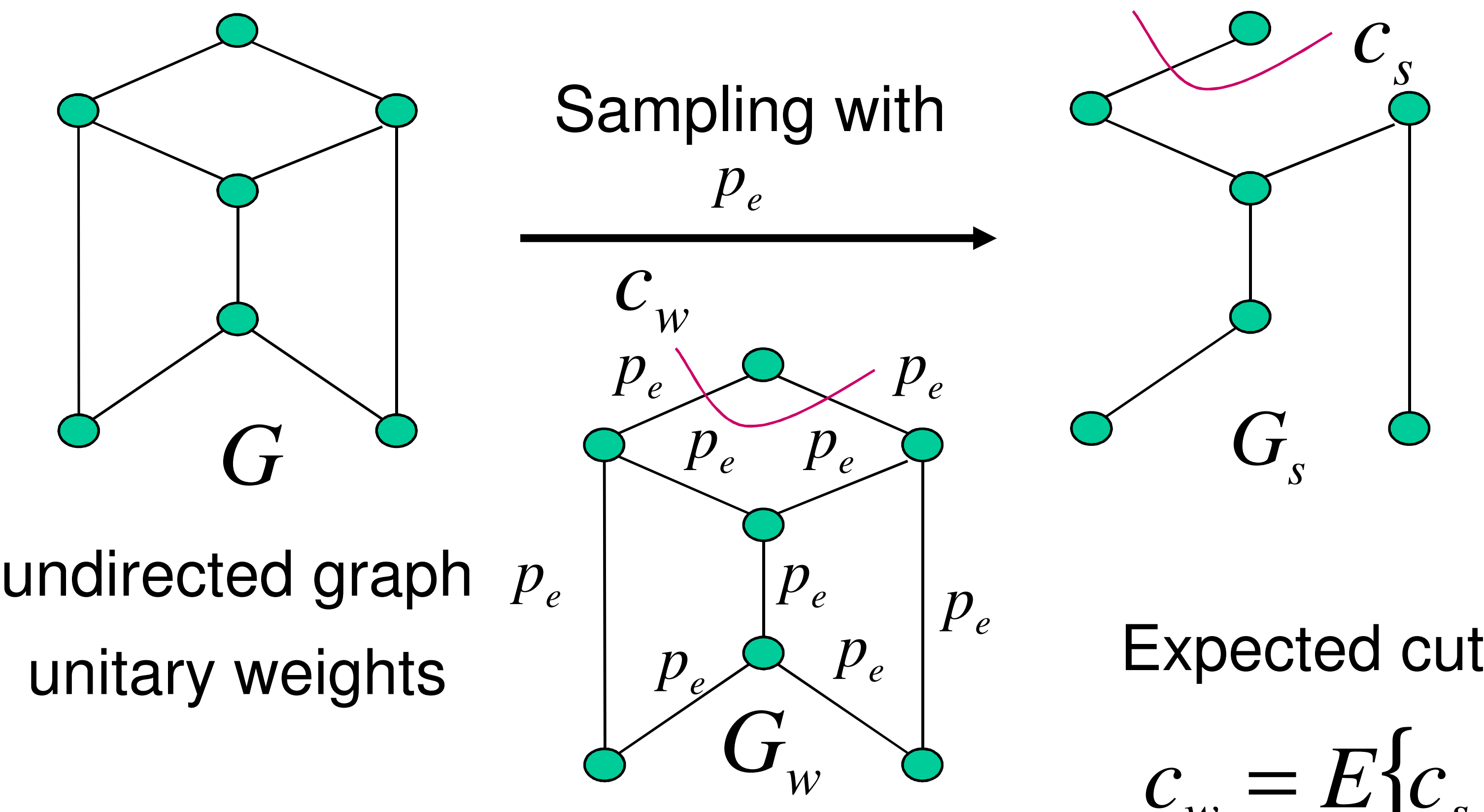
Theorem 1 With high probability, the value of the capacity of a small-world network with added shortcuts lies between $(1-\varepsilon)c_w$ and $(1+\varepsilon)c_w$, with $c_w = k+(n-1-k)p$ and $\varepsilon = \sqrt{2(d+2)\ln(n)/c_w}$.

Bounds for the capacity of a small-world network with added shortcuts, for *n*=1000, *k*=20 and *d*=1.



Key ingredient for the proofs

Random Sampling on Graphs [Karger'94]



undirected graph

unitary weights

Expected cut

$$c_w = E\{c_s\}$$

The generation of small-world networks can be viewed as a random sampling process on a graph.

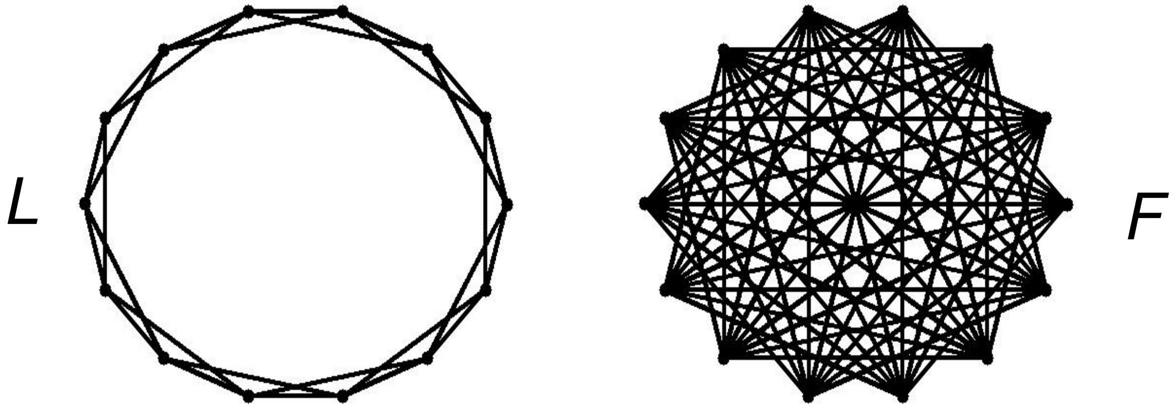
Random Sampling on Graphs

Theorem (Karger, 1994): Let $\varepsilon = \sqrt{2(d+2)\ln(n)/c_w}$ Then, with probability $1 - O(1/n^d)$, every cut in G_s has value between $(1-\varepsilon)$ and $(1+\varepsilon)$ times its expected value (i.e., the value of the same cut in G_w).

Corollary Let $\varepsilon = \sqrt{2(d+2)\ln(n)/c_w}$. Then with high probability, the value of c_s lies between $(1-\varepsilon)c_w$ and $(1+\varepsilon)c_w$.

A simple lemma

Lemma 1: Let $L=(V_L,E_L)$ be a *k*-connected ring lattice and let $G=(V_L,E)$ be a fully connected graph, in which edges $e\in E_L$ have weight $w_1 \geq 0$ and edges $f \notin E_L$ have weight $w_2 \geq 0$. Then, the global minimum cut in G is $kw_1+(n-1-k)w_2$



Each node in *L* has k edges of weight w_1; each node in *F* has $n-1-k$ edges of weight w_2 .

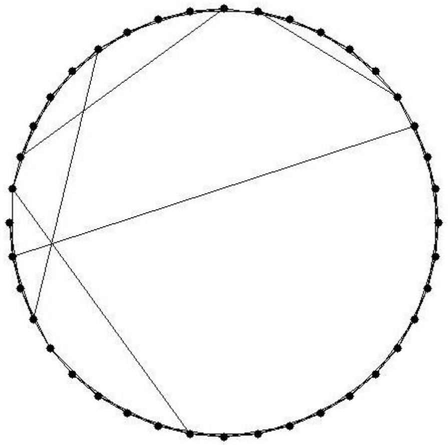
Sketch of Proof for Theorem 1

- Take a fully connected graph.
- Define the weight of the edges as follows:
  - edges on the ring lattice have unitary weight (they are not removed)
  - edges on the remaining edges (shortcuts) have weight *p*
- Apply lemma 1 with $w_1 = 1$ and $w_2 = p$
- The global minimum cut in G_w is $c_w = k + (n-1-k)p$

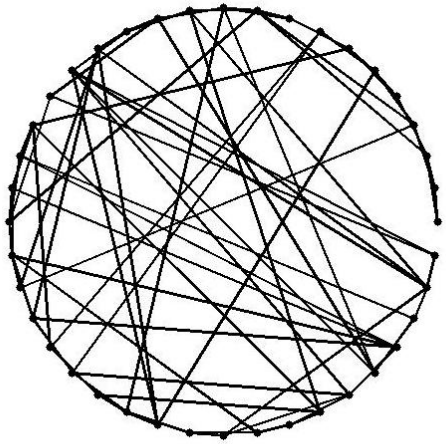
Capacity Bounds for SWNs with Rewiring

Theorem 3 ***(Rewiring does not alter capacity):***

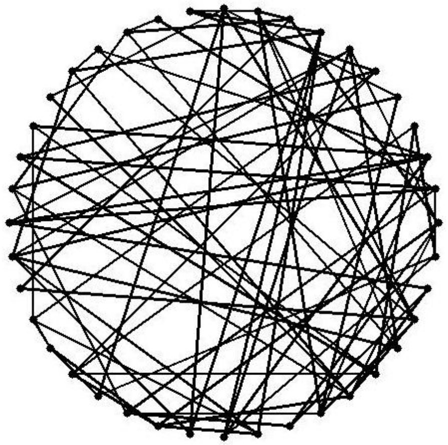
With high probability, the capacity of a small-world network with rewiring has a value in the interval $\left[(1-\varepsilon)k, k\right]$, with $\varepsilon = \sqrt{2(d+2)\ln(n)/k}$



p=0.1



p=0.5



p=0.9

Sketch of Proof

- Take a fully connected graph
- Assign weight *1-p* to edges of the lattice
- To calculate the remaining weights consider the following events:

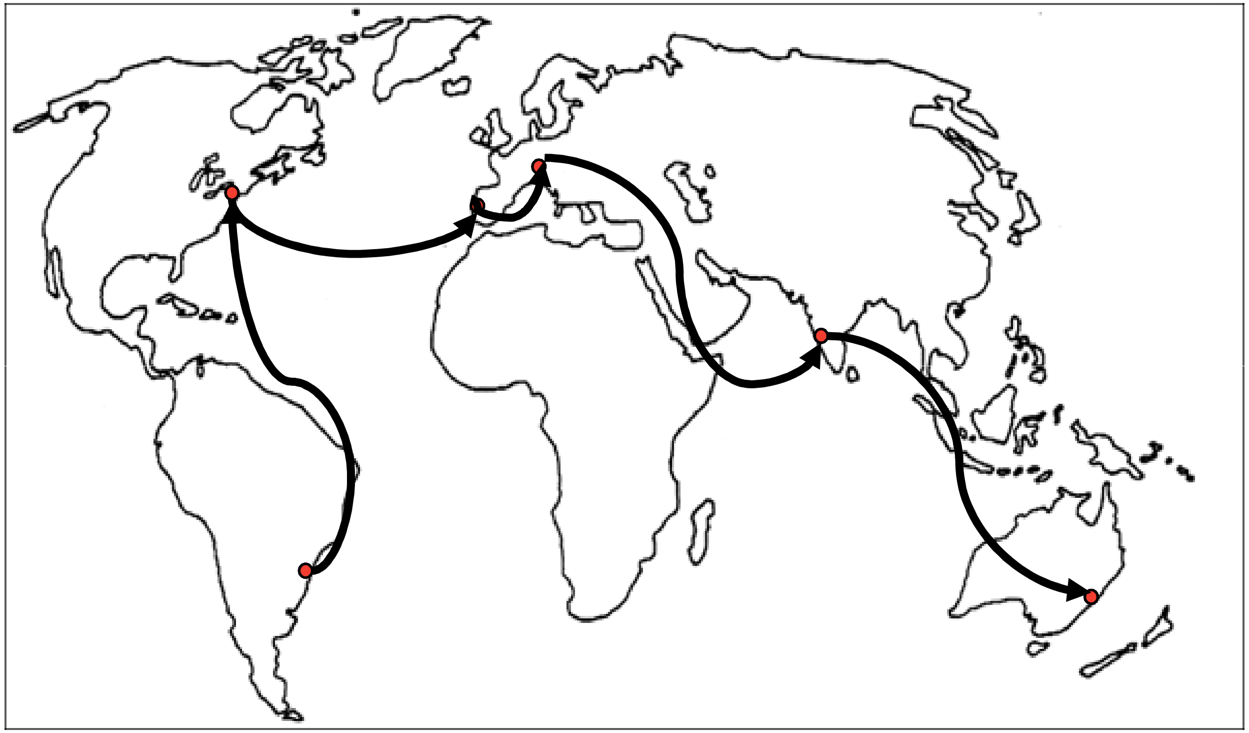
$R(i, j)$: “Rewire the edge $(i, j) \in E_L$”; $\left(P(R(i, j)) = p, \forall i, j\right)$

$C_i(j, l)$: “Rewire $(i, j) \in E_L$ to $(i, l) \notin E_L$”

- ▶ Based on these events show that $\Rightarrow P(i \leftrightarrow j) \geq \frac{pk}{n-k-1}$
- ▶ Use lemma 1 to show that $c_w > k(1-p) + (n-k-1)\frac{pk}{n-k-1} = k$
- ▶ Karger’s Corollary 1 yields lower bound ;
- ▶ Prove by contradiction that $c_s < k$

Navigability of Small World Networks

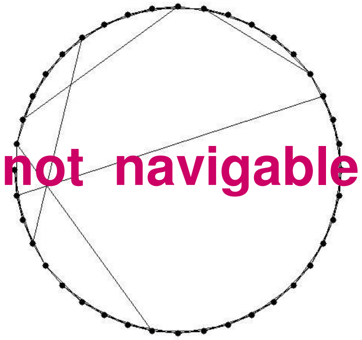
If short paths exist…

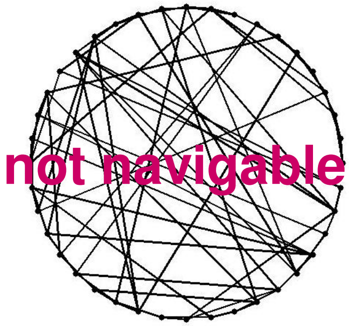


…how do people find them?

Navigability

- Given a source s and a destination t, define a greedy local search algorithm that
  1. knows the positions of the nodes on the graph
  2. knows the neighbors and shortcuts of the current node
  3. knows the neighbors and shortcuts of all nodes seen so far
  4. operates greedily, each time moving as close to t as possible



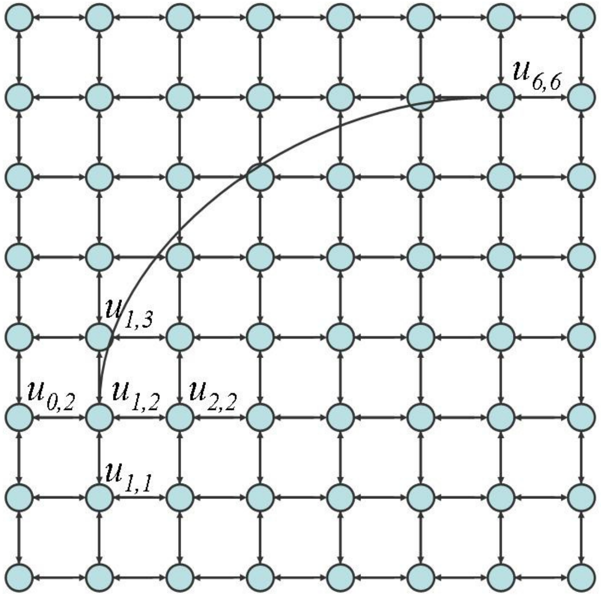


[Kleinberg'00]

▶ Such an algorithm does not work for the previous models.

Kleinberg’s model

- Consider a directed 2-dimensional lattice
- For each vertex u add q shortcuts
  - choose vertex v as the destination of the shortcut with probability proportional to $[d(u,v)]^{-r}$
  - when r = 0, we have uniform probabilities



- This model is navigable only for r=2 (otherwise efficient *distributed* routing algorithms do not exist)

Capacity Bounds for Kleinberg networks

Lemma 1: Let G_w be the weighted graph associated with a Kleinberg network, and c_w be the global minimum cut in G_w. Then, for $h<n-1$,

$$c_w=\frac{h(h+3)}{2}+\sum_{x=1}^{h+1}\sum_{y=h+2-x}^{n}f(x,y)+\sum_{x=h+2}^{n}\sum_{y=1}^{n}f(x,y)$$

where

$$f(x,y)=q.\left(g_{(x,y)}(1,1)+g_{(1,1)}(x,y)\right)$$

$$g_{(x,y)}(a,b)=\left(1-\frac{(x+y-2)^{-r}}{s(a,b)}\right)^{q-1}.\frac{(x+y-2)^{-r}}{s(a,b)}$$

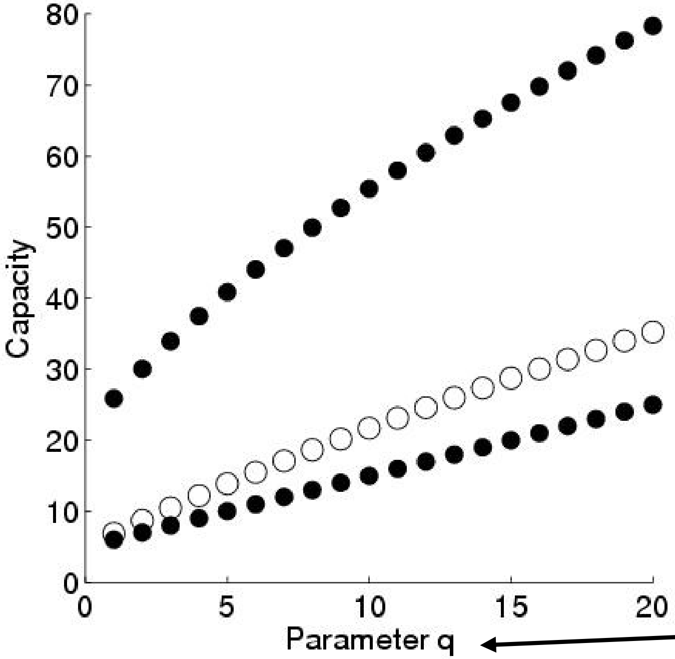
$$s(1,1)=\sum_{i=h+1}^{n-1}(i+1).i^{-r}+\sum_{i=0}^{n-2}(n-1-i).(n+i)^{-r}$$

The proof is a bit too technical for this talk…

Capacity Bounds Kleinberg Networks

Theorem 1: For $h < n-1$, the capacity of a Kleinberg SWN lies, with high probability, in the interval $[M, (1+\varepsilon).c_w]$ where c_w is given by Lemma 1,

$$M = \max\left\{\frac{h(h+3)}{2} + q,\ (1-\varepsilon).c_w\right\} \text{ and } \varepsilon = \sqrt{2(d+2).\ln(n^2)/c_w} .$$



Bounds for the capacity of a Kleinberg network for

n =80, *h* =2, *r* =2 and *d* =1.

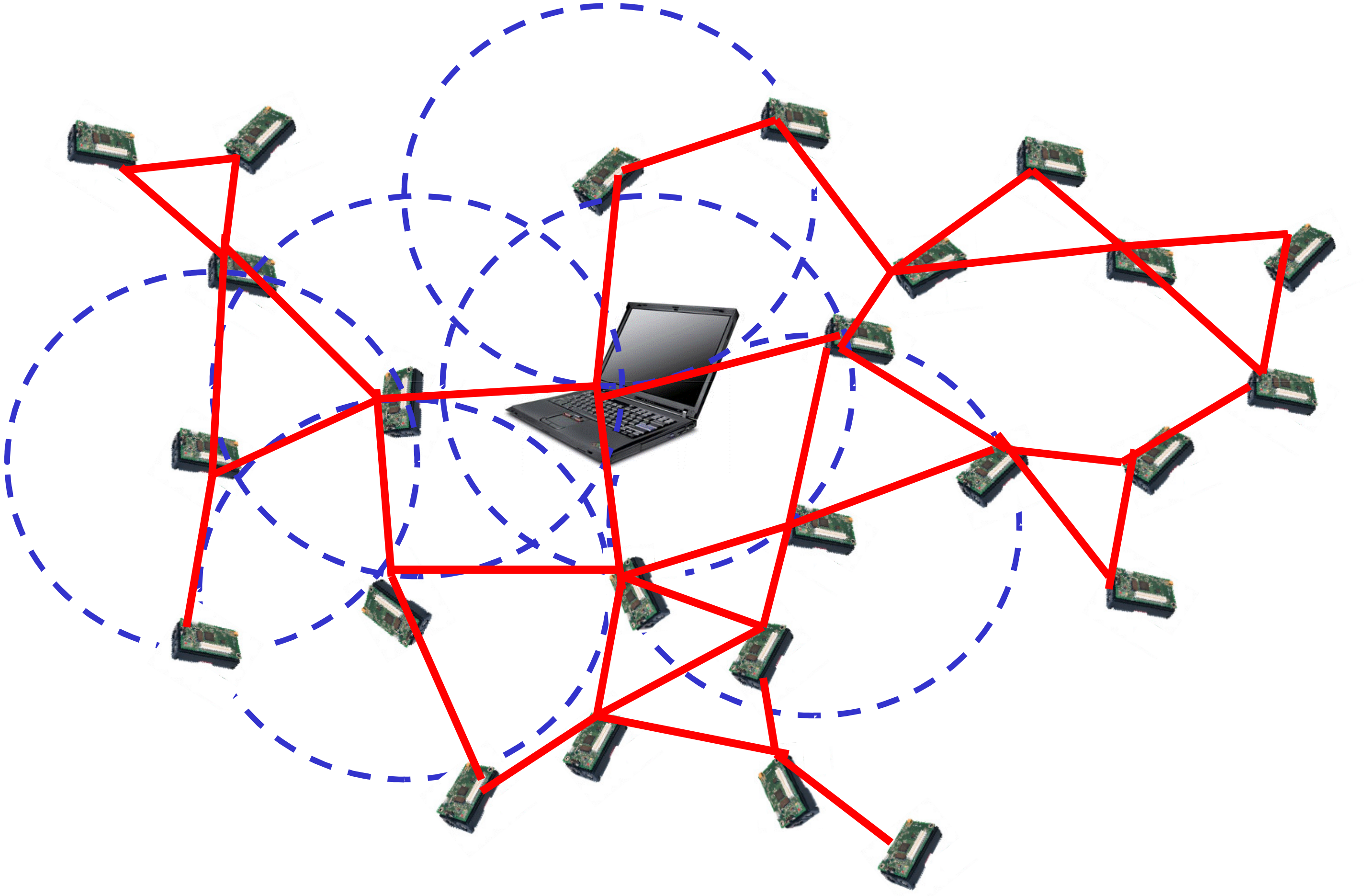
Number of shortcuts per node

Comments

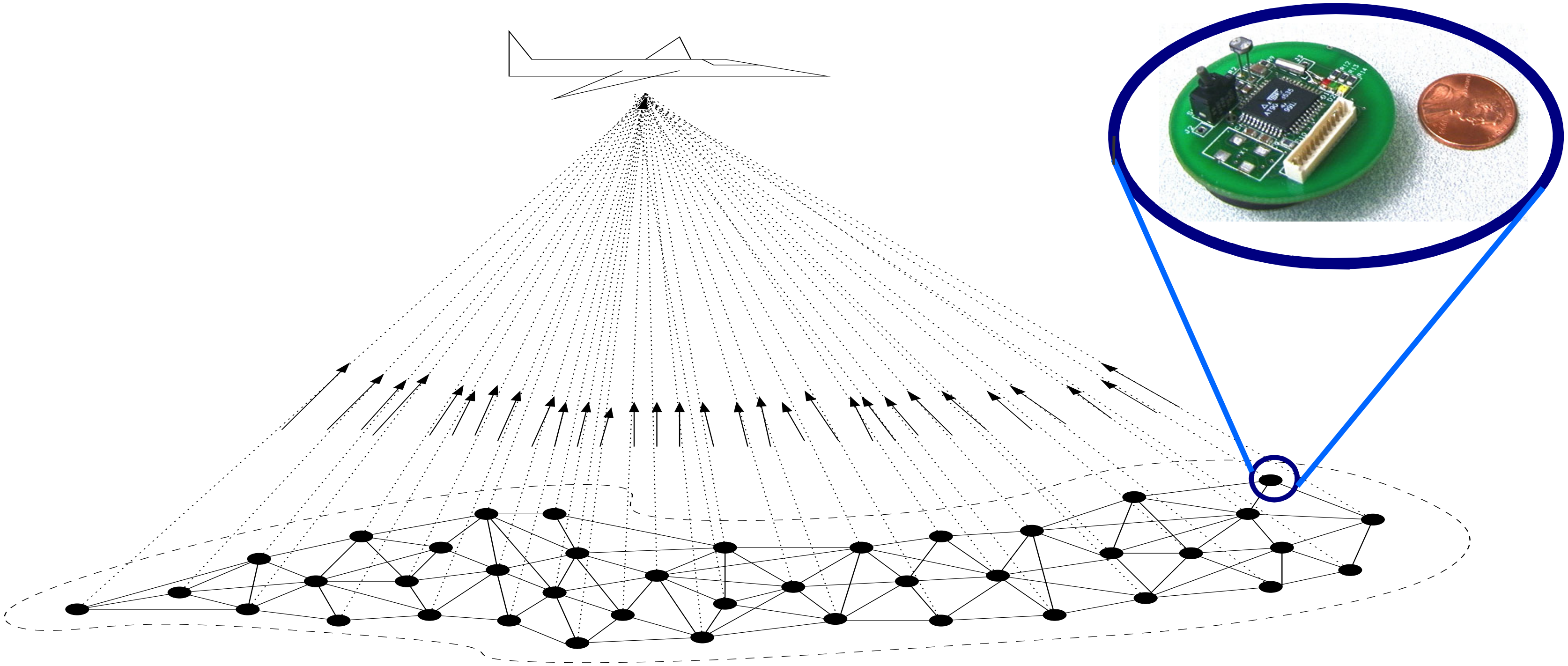
- Small-world networks appear to be “everywhere” and have appealing properties both in theory and in practice.
- Although connectivity parameters have been studied extensively before, to the best of our knowledge these are the first results on the *capacity of SWNs.*
- When it comes to capacity, rewiring is not the same as adding shortcuts – w.h.p. rewiring does not alter the capacity.
- Navigable small-world networks are particularly interesting because they allow for highly efficient distributed routing (and network coding?) algorithms.

Fundamental Limits of Sensor Networks

Sensor Networks



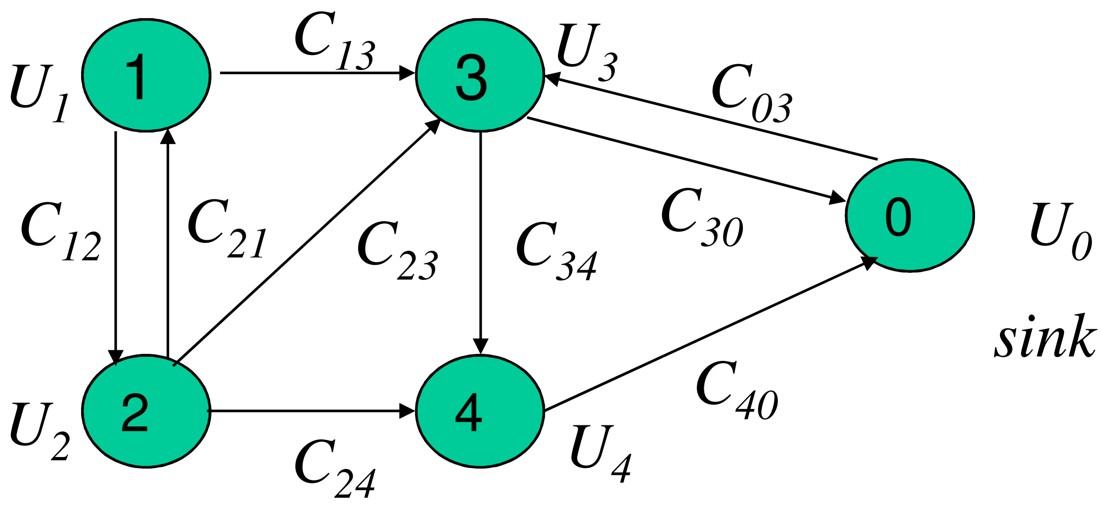
Motivation: Wireless Sensor Networks



Main Task: To collect and transmit data about some physical process

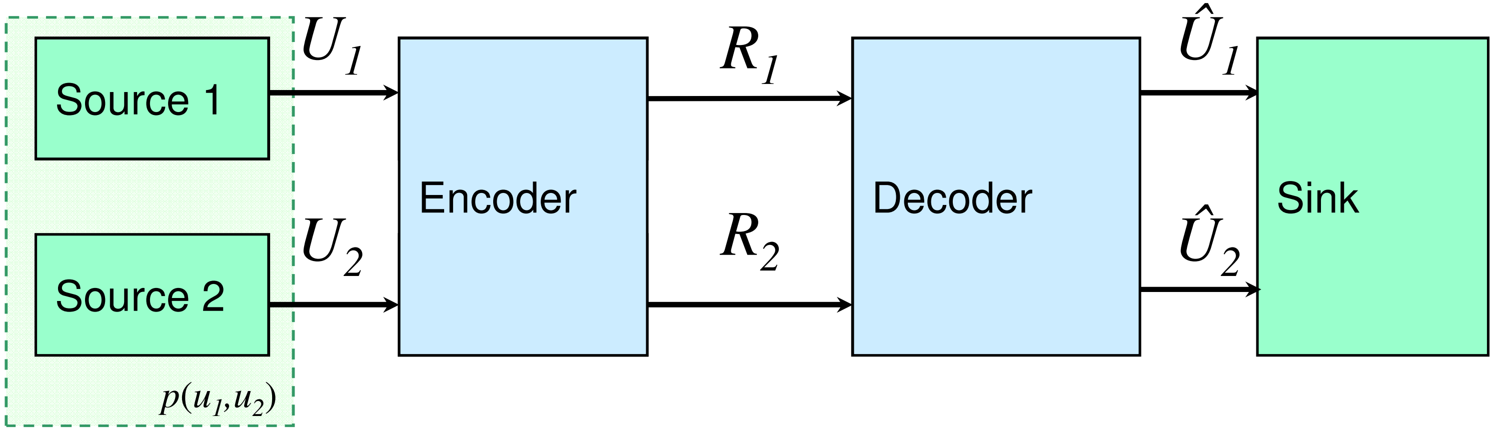
➡ Collected data is typically correlated!

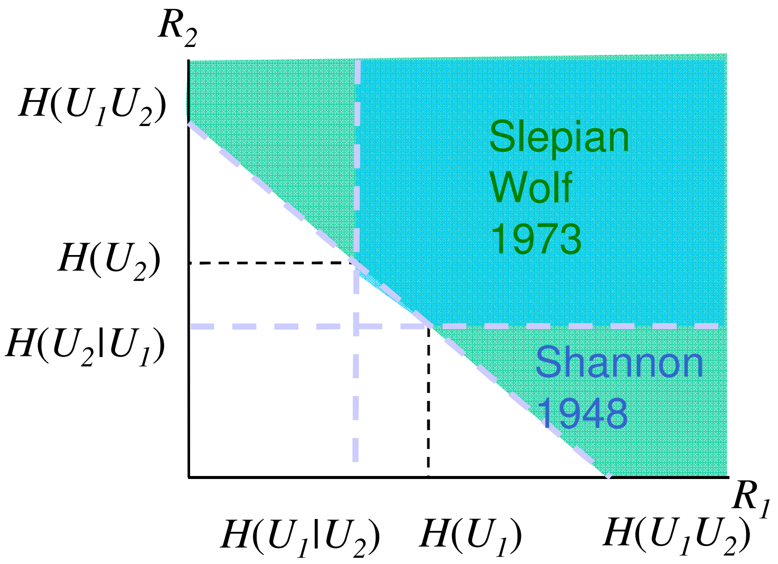
Informal Problem Statement



Under what conditions on the sources and the channels is reliable communication possible?

Encoding Correlated Sources



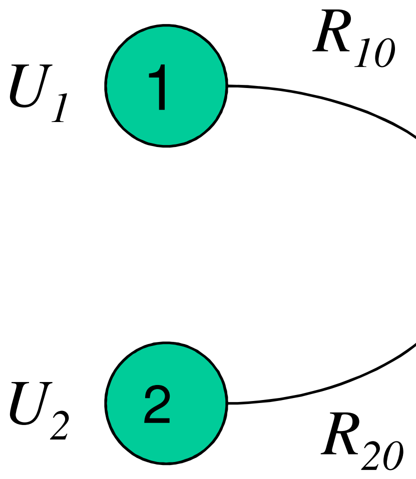


$$R_1 > H(U_1|U_2)$$

$$R_2 > H(U_2|U_1)$$

$$R_1+R_2 > H(U_1U_2)$$

A network flow perspective



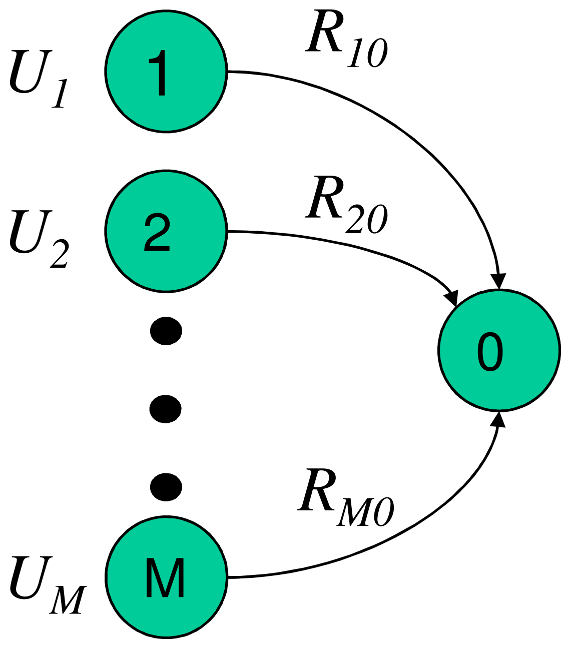
$$R_{10} > H(U_1|U_2)$$

$$R_{20} > H(U_2|U_1)$$

$$R_{10}+R_{20} > H(U_1U_2)$$

The Slepian-Wolf Theorem gives necessary and sufficient conditions for feasible flows that guarantee perfect reconstruction at node 0.

Many correlated sources



Perfect reconstruction is possible if and only if

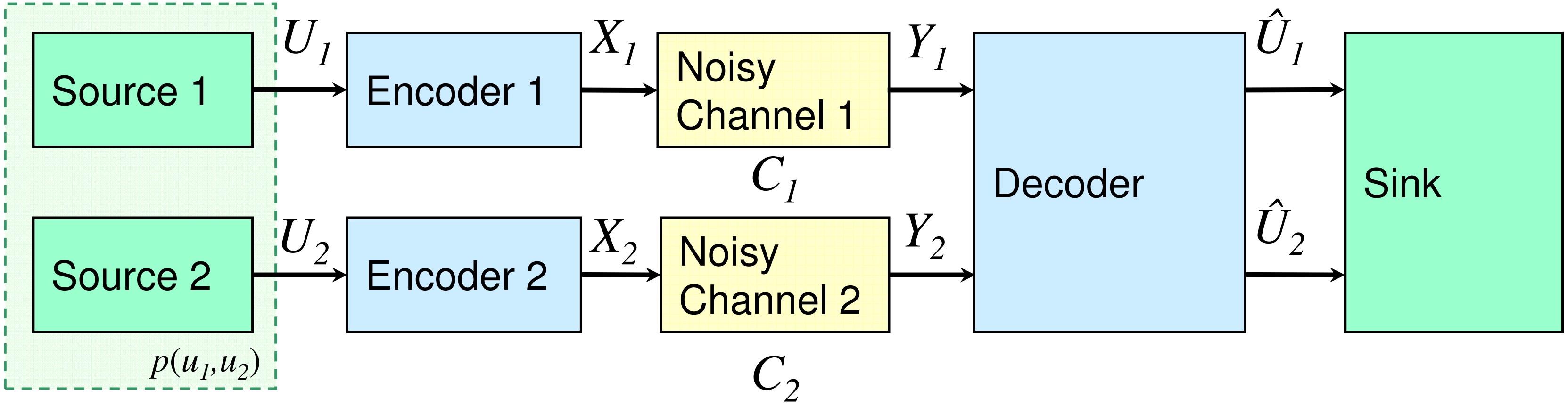
$$\sum_{i\in S} R_{i0} > H(U(S)\,|\,U(S^c))$$

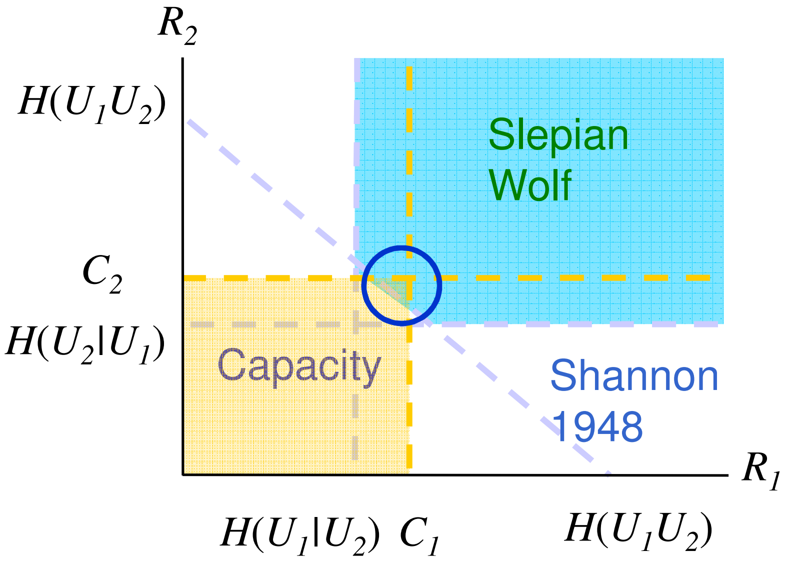
for all sets $S \subset \{1,2,....,M\}$,

$S \cap S^c = 0$,

$S \neq 0$

Noisy Channels





Barros, Servetto 2002:

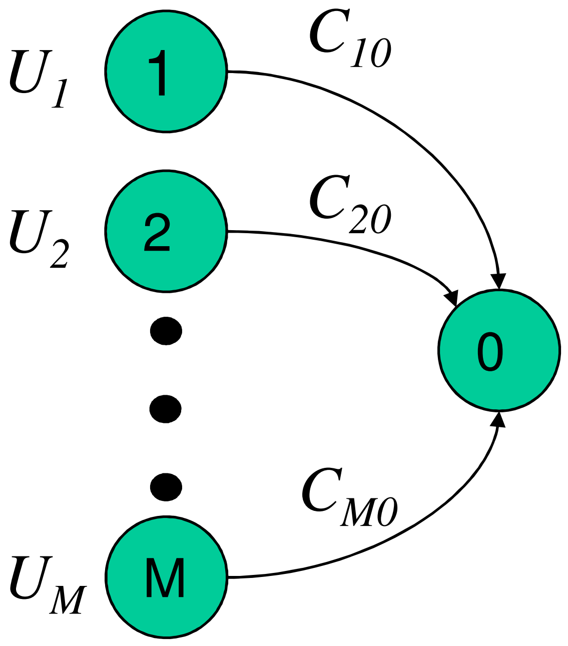
Perfect reconstruction is possible if and only if

$$H(U_1|U_2) < C_1$$

$$H(U_2|U_1) < C_2$$

$$H(U_1U_2) < C_1+C_2$$

Multiple Sources and Channels



Perfect reconstruction is possible if and only if

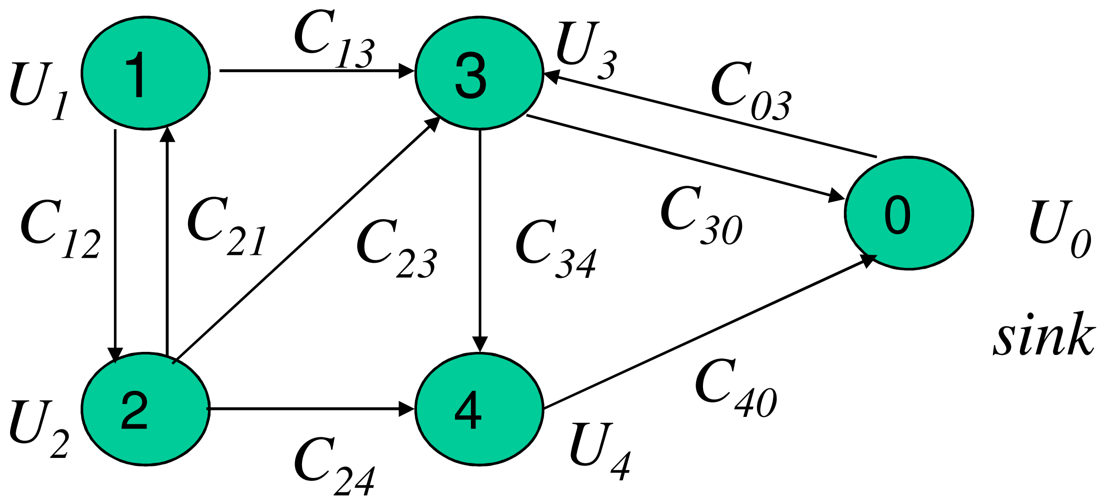
$$H(U(S)\,|\,U(S^c) < \sum_{i\in S} C_{i0}$$

for all sets $S \subset \{1,2,\ldots.,M\},$

$S \cap S^c = 0,$

$S \neq 0$

General Problem Statement



The network is described by a directed graph $G=(V,E)$.

After W rounds of communication node 0 must produce a perfect reconstruction of all sources.

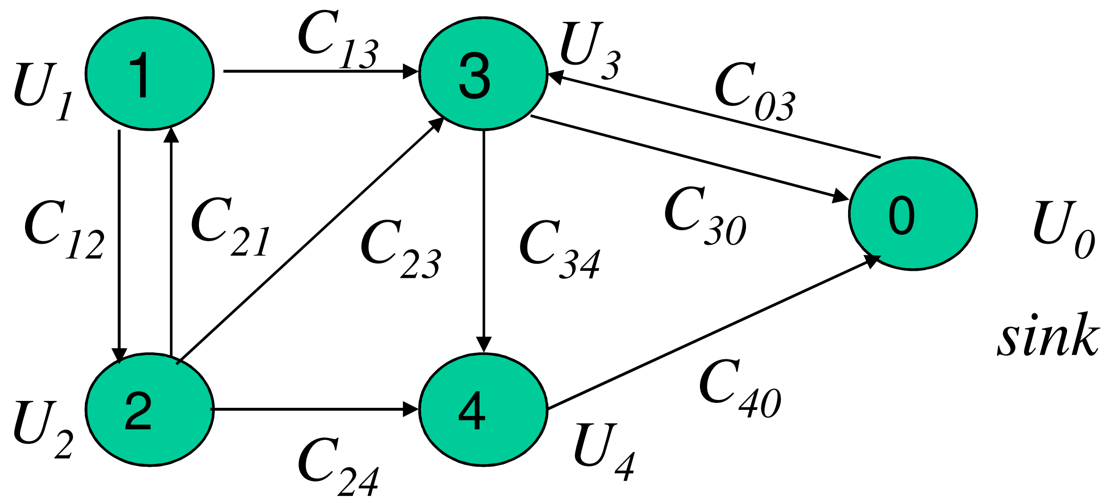
In each round the sent codewords depend on all previously received channel outputs.

Coding Strategy

Use capacity-achieving channel codes to turn the noisy network into a noiseless network.

Use network source codes for (1) distributed compression and (2) routing to the destination.

Network Source Codes



Use classical Slepian-Wolf codes at some operating point $(R_1,R_2\ldots,R_M)$.

View this as a flow network and consider a flow f with M sources and demands $(R_1,R_2\ldots,R_M)$ *at node 0.*

If *f* exists, then *f* determines the number of bits that each node must send to its neighbours.

Achievability

Slepian Wolf Theorem: $$\sum_{i \in S} R_i > H(U(S) \,|\, U(S^c))$$

Elementary flow concepts:

a flow is feasible if $$\sum_{i \in S} R_i < \sum_{\substack{i \in S \\ j \in S^c}} C_{ij}$$

i.e. the total amount of flow injected on one side of the cut has to be lower than the capacity of the links carrying that quantity of flow to the other side.

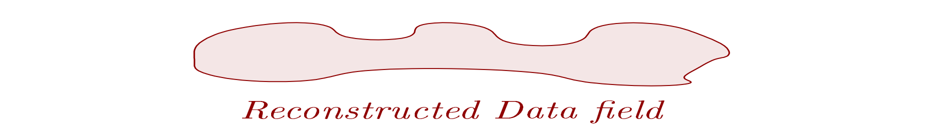
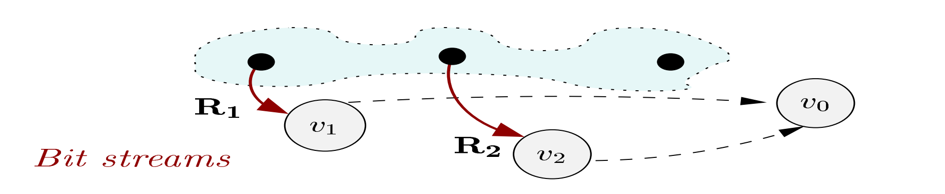
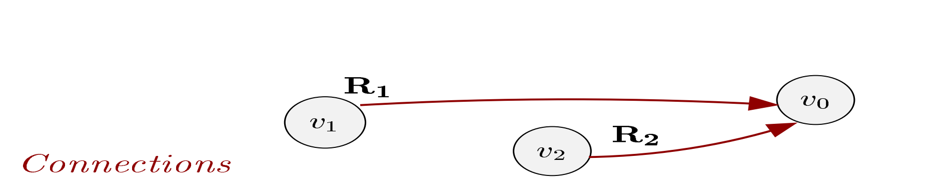
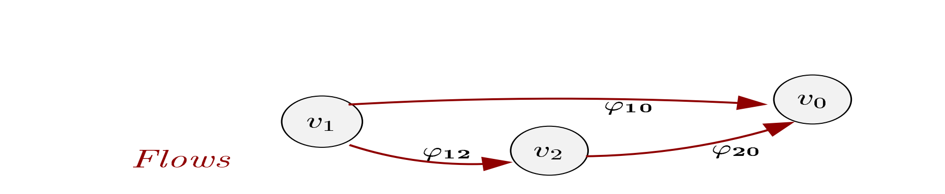
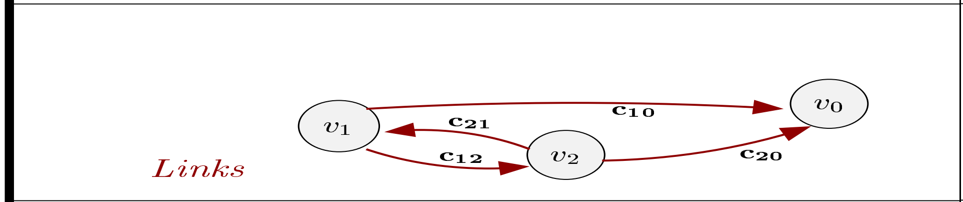
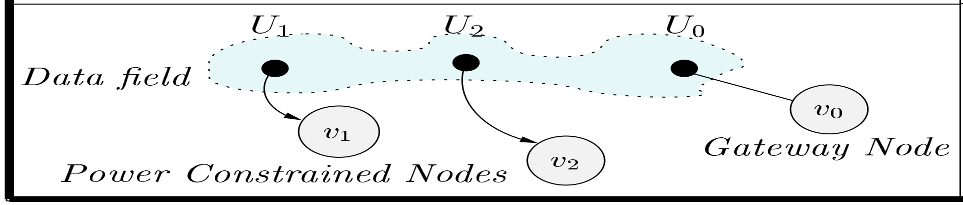
Thus, $$\boxed{H(U(S) \,|\, U(S^c)) < \sum_{\substack{i \in S \\ j \in S^c}} C_{ij}}$$

Barros, Servetto 2005

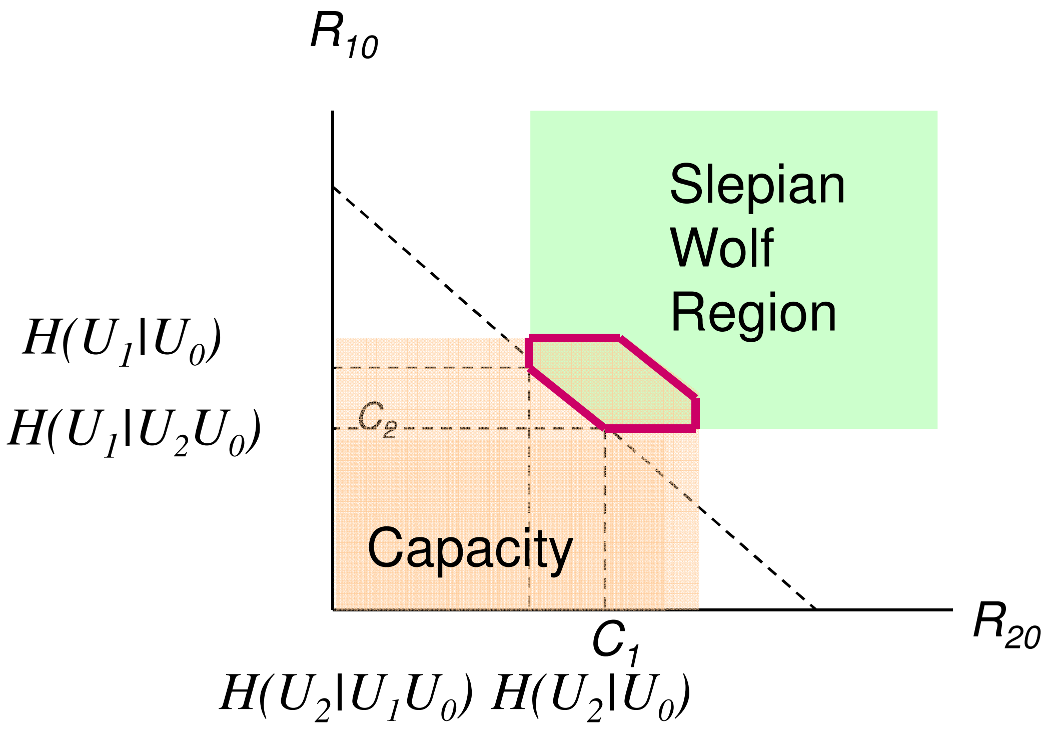
Converse Proof

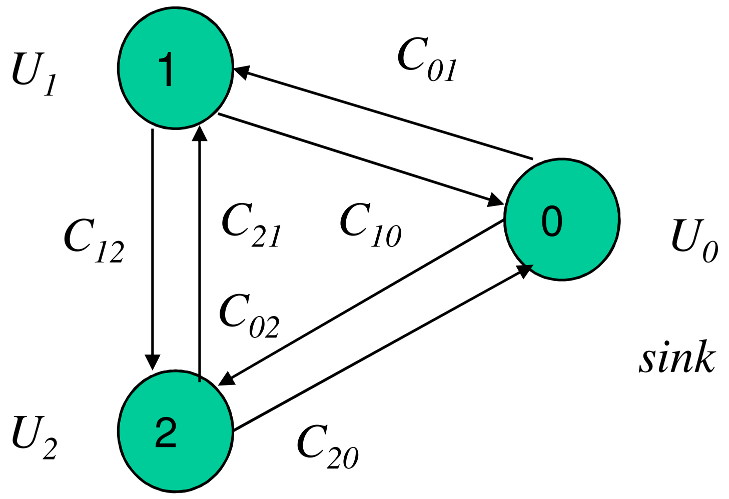
- Max-Flow Min-Cut Bounds do not apply here, because of correlated sources.
- All the painstaking steps of a classical converse proof are required.
- Key ideas:
  - Take “snapshots” of source blocks
  - After a finite time all the information about a snapshot has crossed the network and arrived at the decoder.
  - Exploit Markov relationship between super-blocks long enough to accomodate this notion of network delay.

An Optimal Protocol Stack

 *Reconstructed Data field*	Application Layer *(User of the data)*
 *Bit streams*	Presentation Layer *(Distributed Sampling/Compression)* *(Interpolation)*
 *Connections*	Transport Layer
 v_1 φ_{10} v_0 φ_{12} v_2 φ_{20} *Flows*	Network Layer *(Feasible Flow Computation)*
 v_1 c_{10} v_0 c_{21} c_{12} v_2 c_{20} *Links*	Link Layer *(MAC/Power/Error Control)* Independent channels
 *Data field* *Power Constrained Nodes* *Gateway Node*	Physical Layer

Network Optimization





Is the rate polytope non-empty?

If yes, what is an optimal flow?

Linear Programming

Linear cost model (e.g. energy per bit)

$$\min \kappa(f) = \sum_{(v_i, v_j) \in E} w(v_i, v_j) \cdot f(v_i, v_j)$$

subject to:

$$f(i,j) \leq c_{ij} \qquad 0 \leq i, j \leq M$$

flow constraints

$$f(i,j) = -f(j,i) \qquad 0 \leq i, j \leq M$$

$$\sum_{i \in V} f(i,j) = 0 \qquad 1 \leq i \leq M$$

coding constraints

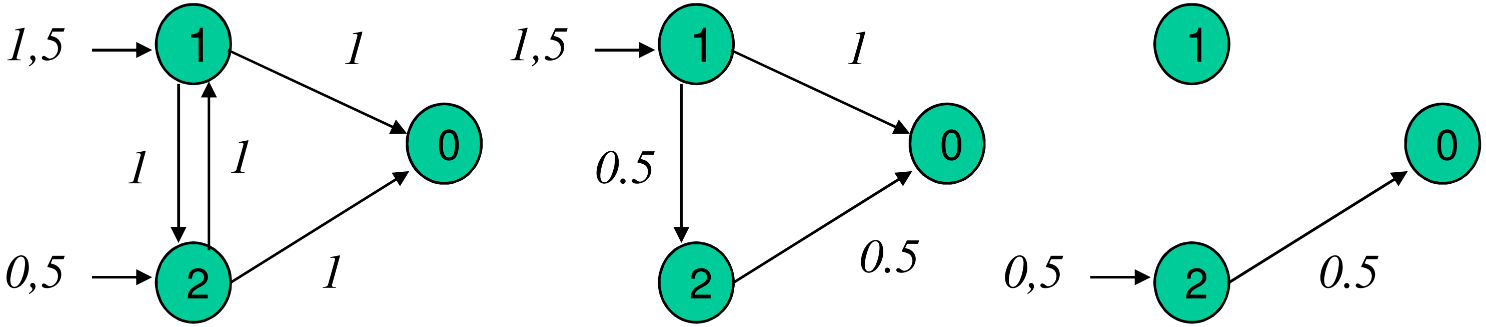
$$H(U_S \,|\, U_{S^c}) \leq \sum_{i \in S} f(s,i) \leq \sum_{i \in S, j \in S^c} c_{ij}$$

$$f(s,i) = R_i \qquad 1 \leq i \leq M$$

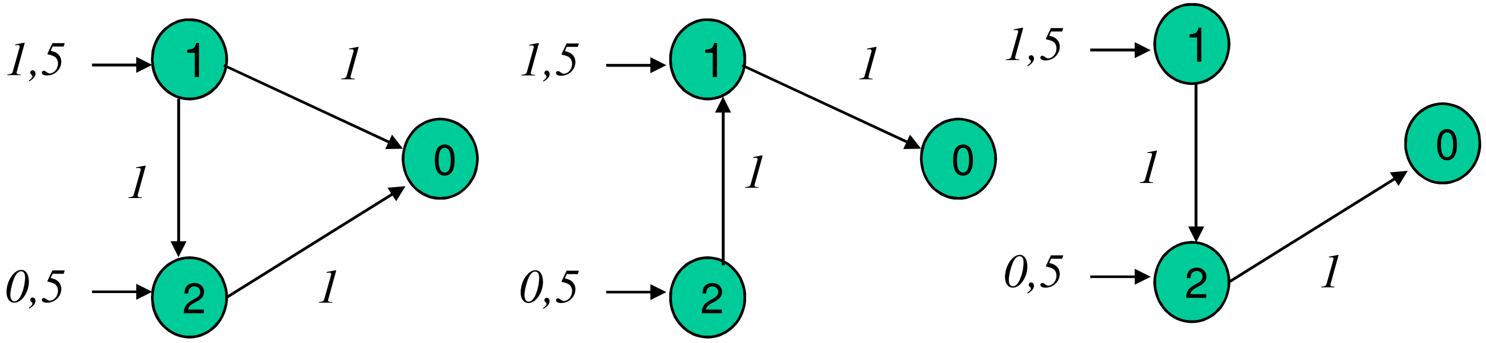
Beware of Trees

Barros, Servetto 2004

This example is solvable…

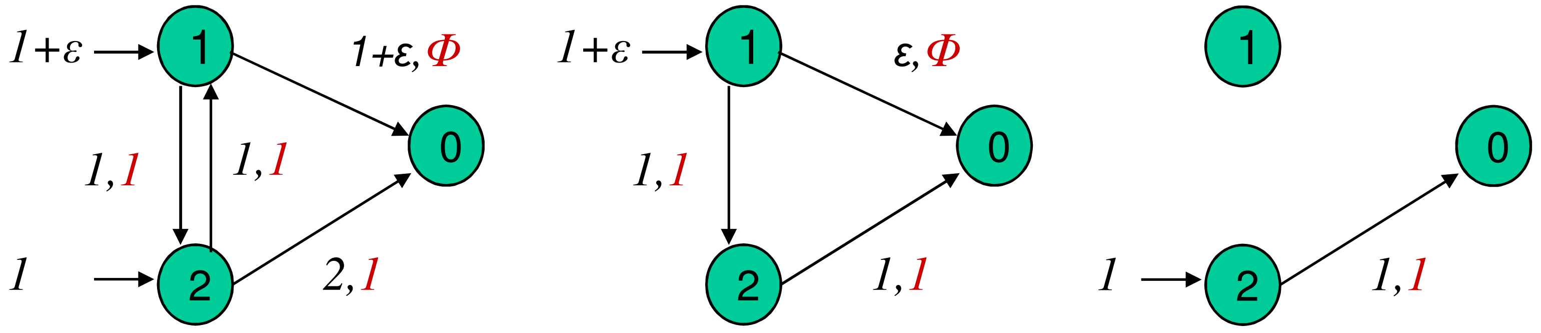


…but not with trees!

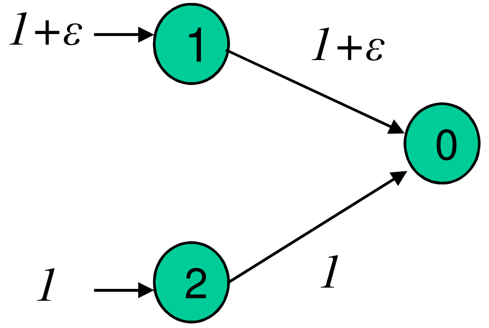


Beware of Trees... indeed.

This example is also solvable…



…but the cost of using a tree is huge!



Ratio (Φ(1+ε)+1) / (ε Φ+3)

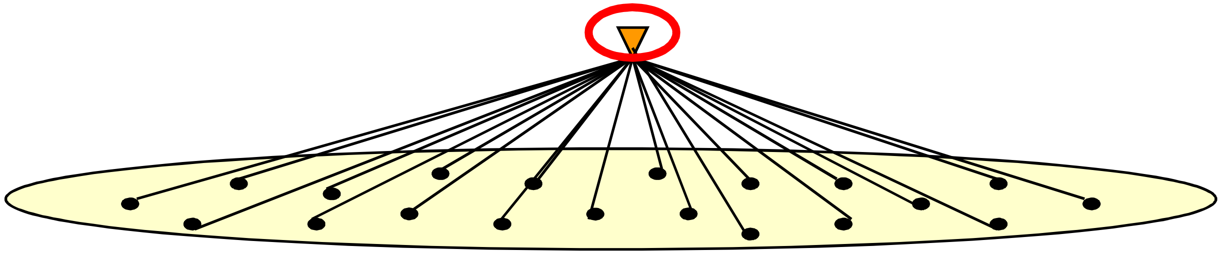
for large Φ, we get about
1+1/ε, unbounded for small ε!

Scalable Decoding on Factor Graphs

Fundamental Challenge

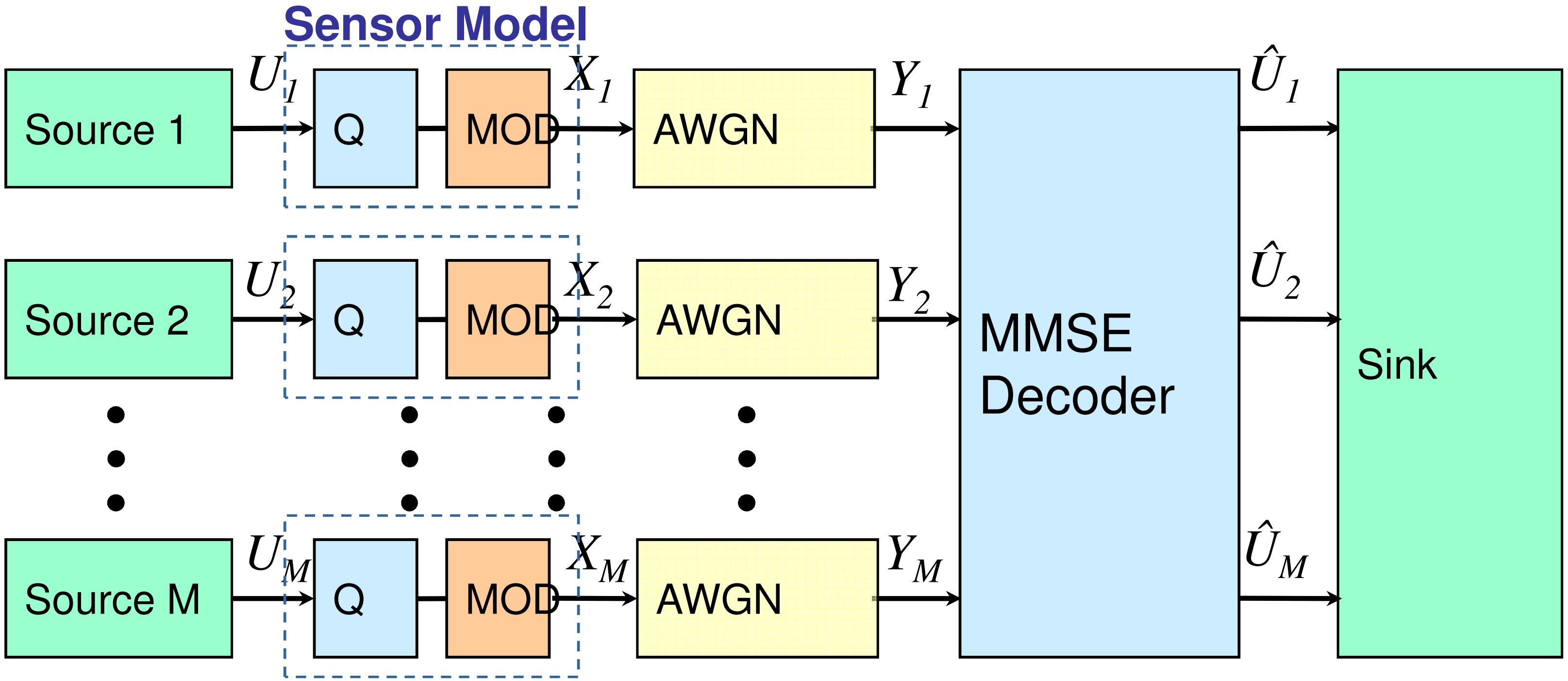
How can we design a scalable system that achieves near-to-optimal performance with manageable complexity?

Two Key Issues



- The algorithmic complexity should be moved from the sensor nodes to the fusion center.
- We should exploit the correlation in the sensor data as much as possible (distributed compression, error correction, improved estimation)

System Model



Source Model:

$$p(\overline{u}) < \frac{1}{\sqrt{2\pi \det \Sigma}} \exp\left(-\frac{1}{2}\overline{u}^T \Sigma^{-1} \overline{u}\right)$$

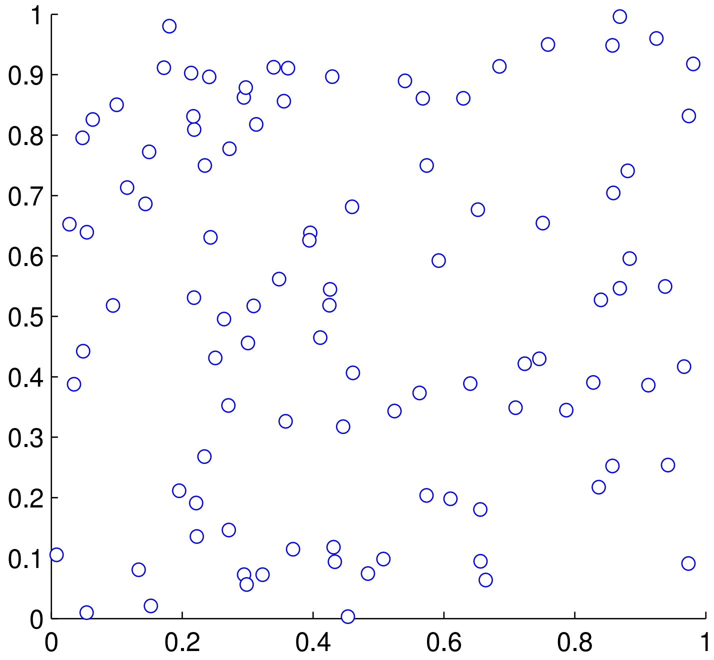
Correlated Sensor Data

$$p(\overline{u}) < \frac{1}{\sqrt{2\pi \det \sum}} \exp\left(-\frac{1}{2} \overline{u}^{T} \sum^{-1} \overline{u} \right)$$

The correlation matrix Σ determines the statistical dependence of the collected sensor data.

In general the correlation will depend on the topology of the network, in particular the distances between the sensors.

Problem Statement



hundreds of sensors

uniformly distributed on the unit square

We want to minimize the mean square error for each of the transmitted samples.

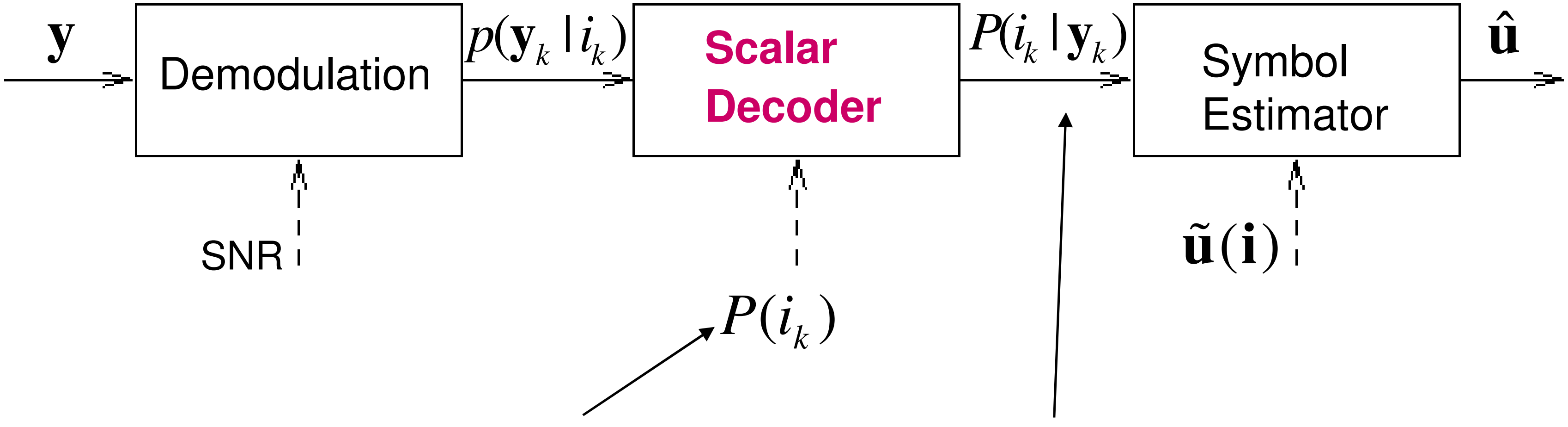
$$E\{(u_k - \hat{u}_k)^2\} \longrightarrow \min$$

What is the optimal MMSE decoder?

Optimal MMSE Decoding

Ignoring the spatial correlation between the sensor measurements, we get the *scalar* conditional mean estimator:

$$\hat{u}_k = E\{\tilde{u}(I_k) \mid \mathbf{Y}_k = \mathbf{y}_k\}$$



$\mathbf{y}$
Demodulation
SNR
$p(\mathbf{y}_k \mid i_k)$
Scalar Decoder
$P(i_k)$
$P(i_k \mid \mathbf{y}_k)$
Symbol Estimator
$\tilde{\mathbf{u}}(\mathbf{i})$
$\hat{\mathbf{u}}$

a priori probability

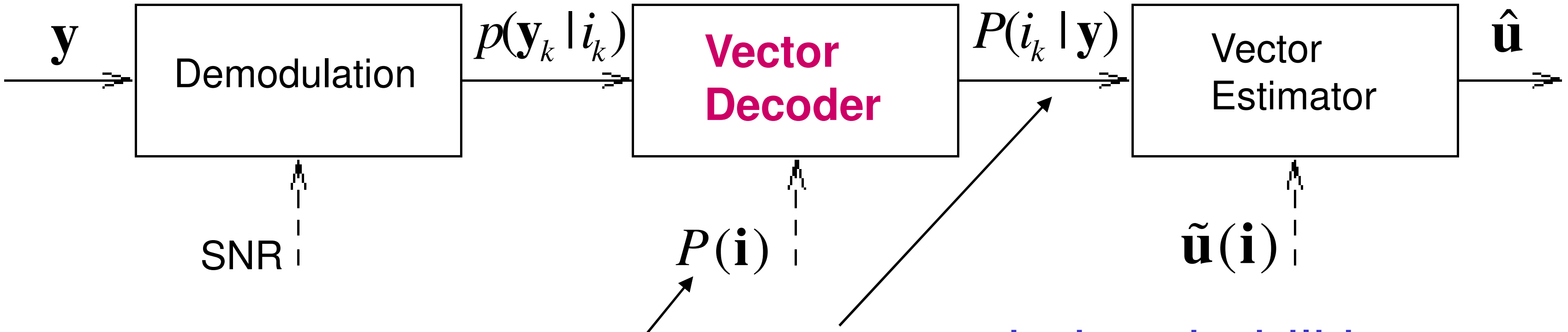
a posteriori probabilities

Optimal MMSE Decoding

However, for optimal decoding we should take the spatial correlation between sensor measurements, into account.

In this case the optimal MMSE Decoder is given by

$$\hat{u}_k = E\{\tilde{u}(I_k) \mid \mathbf{Y} = \mathbf{y}\}$$



a priori probabilities

a posteriori probabilities

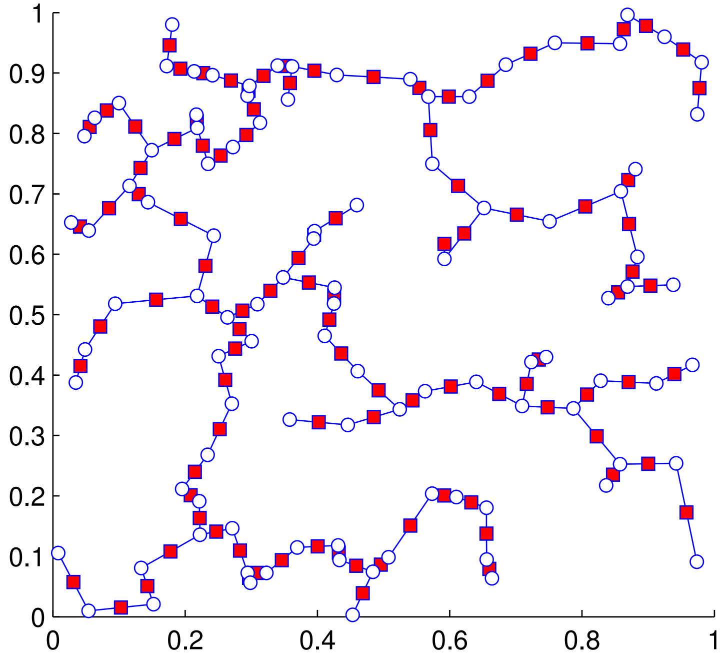
Main Problem: Complexity

Decoder	Complexity (# Multiplications)	$M=100$ sensors, $Q=1$ bit/sample
Scalar	$O(M2^{Q})$	≈ 200
Vector	$O(M2^{QM})$	$\approx 10^{32}$

The complexity of the optimal MMSE decoder grows exponentially with the number of sensors!

Solution: *Scalable Decoding*

Barros, Tuechler 2006



Use a factor tree to approximate the correlation structure of the data

Minimize the Kullback Leibler Distance

Apply the sum-product algorithm to obtain the desired estimates

Under mild assumptions on the graph the decoding complexity grows linearly with the number of sensors !

Comments

- Shannon Theory offers very powerful tools to help understand some of the fundamental aspects of sensor networks.
- Factor Graphs methods provide scalable solutions for joint source-channel coding and source-optimized clustering in large-scale sensor networks.
- When computing functions, the factorisation and the message updates need to be adapted.
- We provided algorithms for three basic functionals, but there are many other cases with practical interest.
- Many (tough) open problems will surely require combinatorics, graph theory, algorithms... and lots of applied math and computer science.

Information-Theoretic Security

Today’s Layered Architecture

Standard Protocol Stack

Application	Programs and applications
Transport	End-to-end reliability, cong. control
Network	Routing and forwarding
Link	Medium access control
Physical	Channel coding and modulation

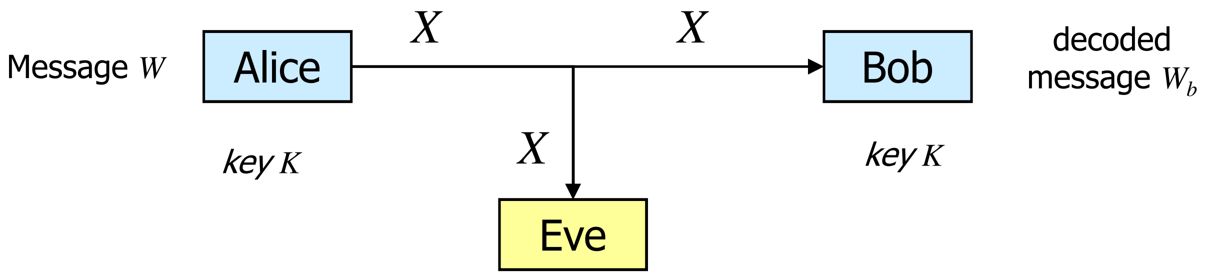
Where is security ?

Security: a patchwork of add-ons…

Application	End-to-end cryptography
Transport	Secure Sockets Layer (SSL)
Network	Virtual private networks (IPSec)
Link	Admission control (e.g.WPA)
Physical	**Physical-layer security ?**

Information-Theoretic-Security – are we biased?

A typical graduate course in cryptography and security always starts by discussing Shannon's notion of perfect secrecy (widely accepted as the strictest notion of security):



$p(w|x)=p(\underset{w}{\cancel{x}})$

Then, it emphasizes its conceptual beauty.

Then, it states that it is basically “useless” for any practical application.

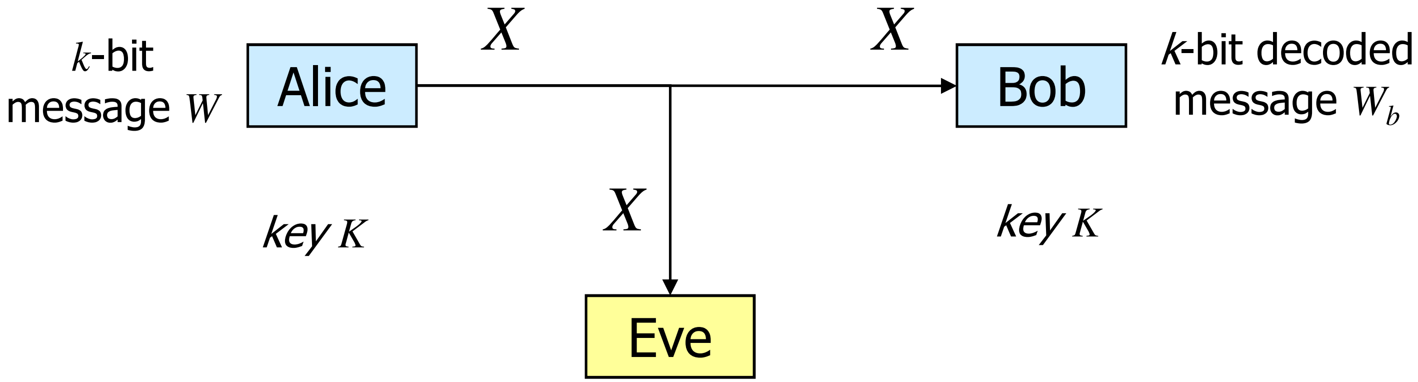
→ Computational Security

Main Questions

- What are the fundamental security limits at the physical layer?
- Which notions of security are we talking about?
- Is information-theoretic security practical?
- What kind of code constructions can we use?
- How do we build protocols based on information-theoretic security?
- Can we combine physical-layer security with classical cryptography?
- How can we secure new wireless networking paradigms?

Theoretical Foundations

Notions of Security



Computational Security

- Alice sends a k-bit message W to Bob using an encryption scheme;
- Security schemes are based on (unproven) assumptions of intractability of certain functions;
- Typically done at upper layers of the protocol stack

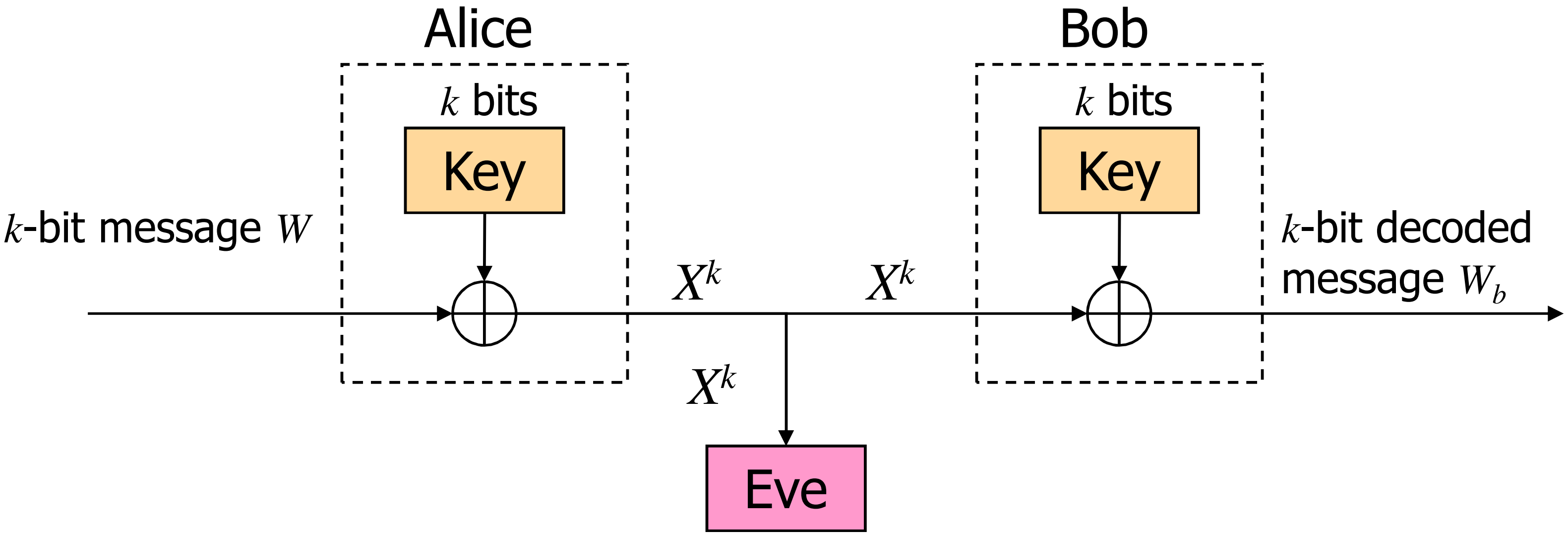
Information-Theoretic (Perfect or unconditional) Security

- strictest notion of security, no computability assumption
- Prob{W | Eve's knowledge}=Prob{W}

  $H(W|X)=H(W)$ or $I(X;W)=0$
- e.g. One-time pad

  [Shannon, 1949] **:** $H(K) \geq H(M)$

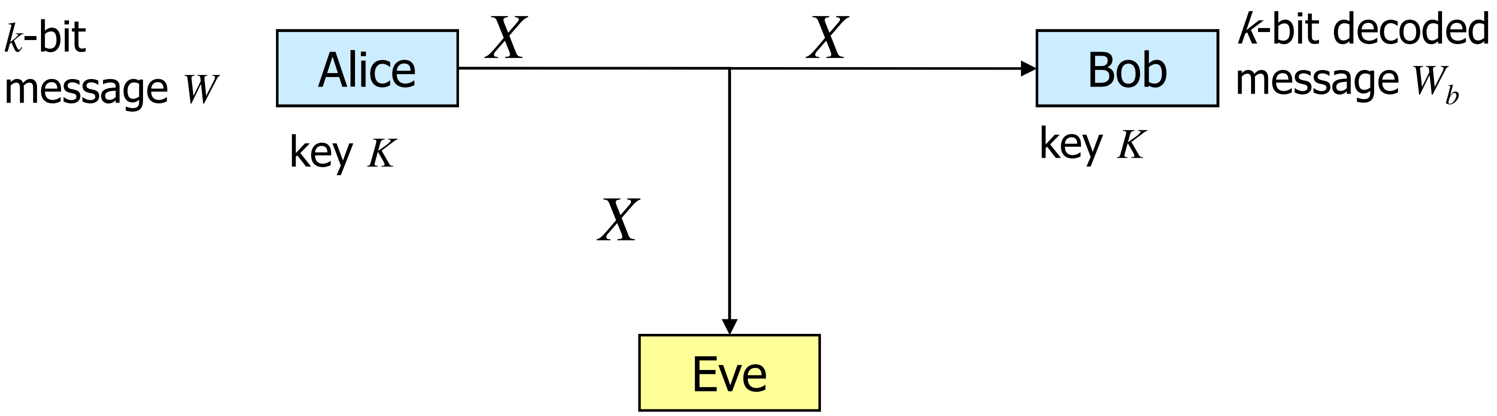
One-time Pad



If Eve does not know the key and $P(\text{Key}=k\text{-tuple})=1/2^k$

then we have $p(w|x^k) = p(w)$.

Shannon’s Model



This model is somewhat pessimistic, because most communications channels are actually noisy.

[Wyner, 1975]

Wyner's Wiretap Channel (I)

Reliability & Security

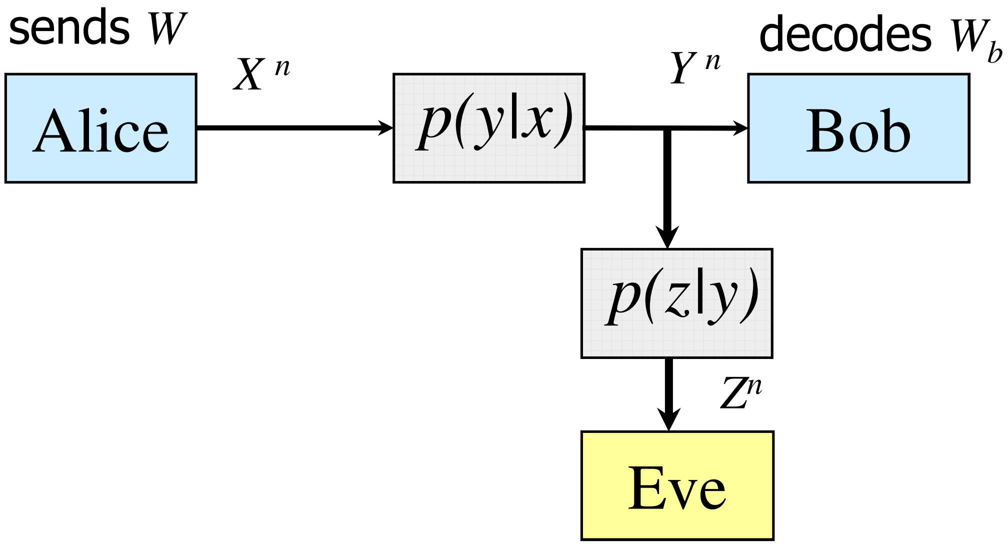
For Bob and Alice,

$\text{Prob}\{W \neq W_b | Y^n\} \rightarrow 0$

With respect to Eve,

$(1/n)\ \text{I}(W; Z^n) \rightarrow 0$

as $n \rightarrow \infty$



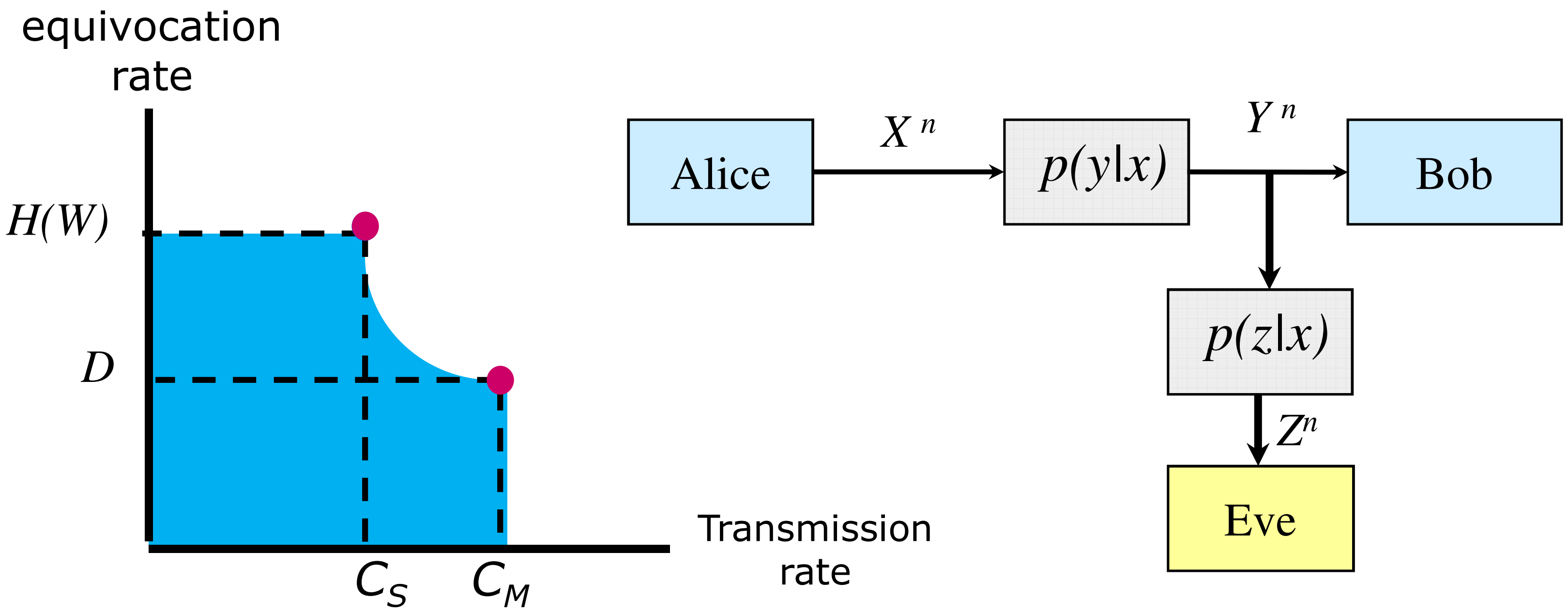
Secrecy Capacity:

Largest transmission rate at which both conditions can be satisfied.

Positive secrecy capacity only in the degraded case.

Wyner’s Wiretap Channel (II)

[Wyner, 1975]

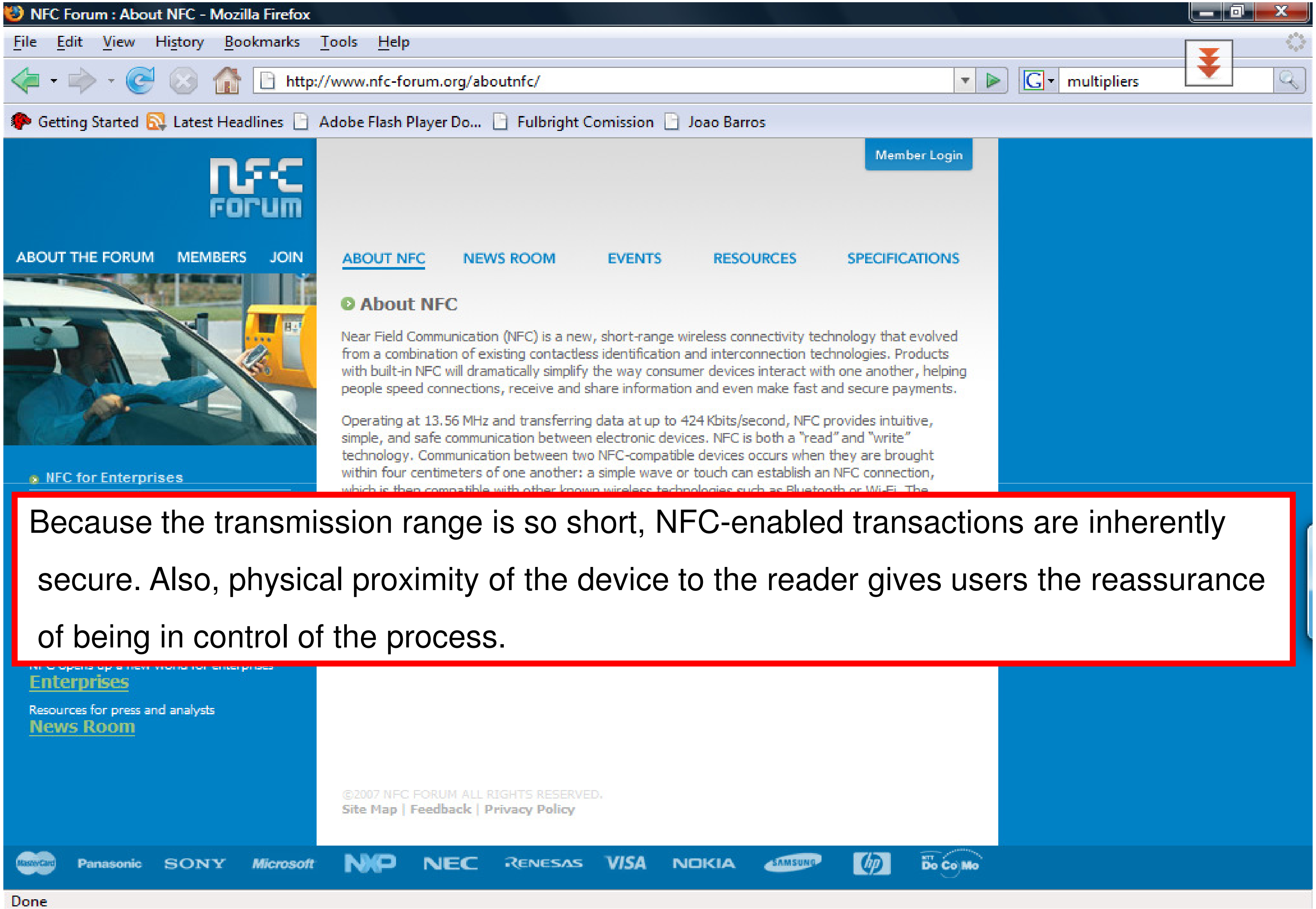


Proof Idea:

- Alice assigns multiple codewords to each message, picks one at random and thus exhausts Eve’s capacity.
- Converse uses Fano’s inequality and classical arguments.

Rate-equivocation region:

- Two critical corner points (C_M, D) and $(C_S, H(W))$
- Unusual shape (not convex)



About NFC

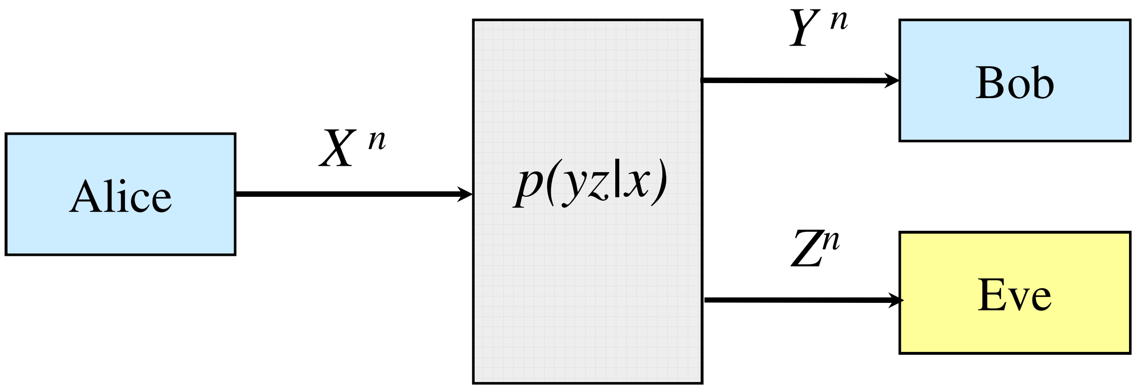
Near Field Communication (NFC) is a new, short-range wireless connectivity technology that evolved from a combination of existing contactless identification and interconnection technologies. Products with built-in NFC will dramatically simplify the way consumer devices interact with one another, helping people speed connections, receive and share information and even make fast and secure payments.

Operating at 13.56 MHz and transferring data at up to 424 Kbits/second, NFC provides intuitive, simple, and safe communication between electronic devices. NFC is both a "read" and "write" technology. Communication between two NFC-compatible devices occurs when they are brought within four centimeters of one another: a simple wave or touch can establish an NFC connection, which is then compatible with other known wireless technologies such as Bluetooth or Wi-Fi. The

NFC for Enterprises
NFC opens up a new world for enterprises
Enterprises
Resources for press and analysts
News Room

Because the transmission range is so short, NFC-enabled transactions are inherently secure. Also, physical proximity of the device to the reader gives users the reassurance of being in control of the process.

Broadcast Channel with Confidential Messages



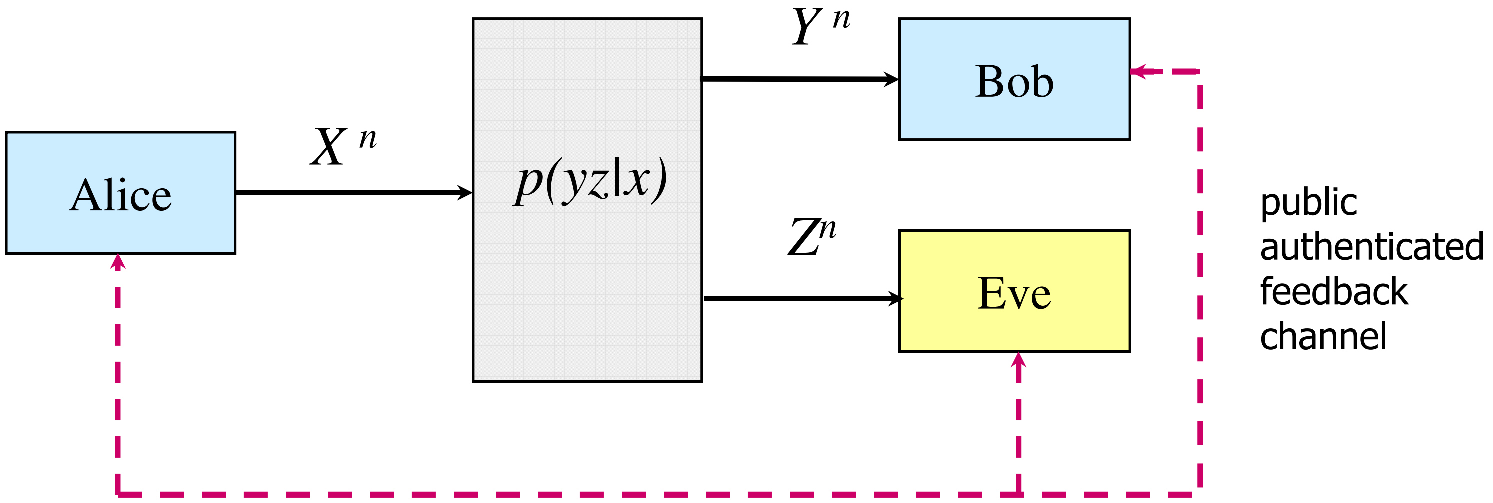
$$\boxed{C_S = \max_{\substack{p(u,x) \\ U-X-YZ}} (I(U;Y) - I(U;Z))}$$

[Csiszár & Koerner, 1978]

Secrecy capacity is strictly positive if Bob’s channel is *less noisy* than Eve’s, i.e. $I(X;Y) > I(X;Z)$

Feedback (Public Discussion)

[Maurer, 93]



Secret Key agreement scheme

- Clever protocol allows Alice and Bob to increase their secrecy capacity by exchanging information over the feedback channel

- This requires a public authenticated feedback channel!

Increasing the Secrecy Capacity via Feedback

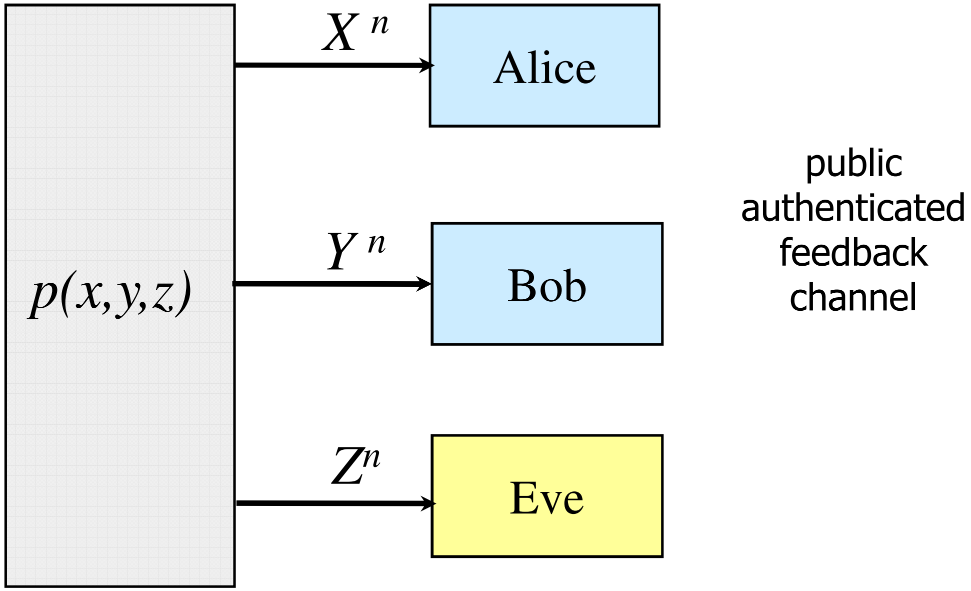
- Suppose Alice, Bob and Eve are connected via binary symmetric channels and a public authenticated feedback channel is available.

	Noisy Channel	Error-free public communication	Computation	
Alice	X	V+X+E	V+X+E+X	V+E
Bob	X+E	V+X+E	V	V
Eve	X+D	V+X+E	V+X+E+X+D	V+E+D

- Bob and Eve observe different noises (D, E).
- Bob feeds back random value V plus what he observed $(X+E)$
- Eve ends up with more noise than Bob (as in the wiretap channel)

[Ahlswede and Csiszar, 93]

Source Model



- Alice and Bob share common randomness.
- Eve gets to see a correlated random variable.
- Alice and Eve generate a secret key using the public authenticated channel.

Some recent work on (weak) secrecy capacity

- Secure space-time communications (Hero, 2003)
- Secrecy rates for the relay channel (Oohama, 2004)
- Secrecy capacity of SIMO channels (Parada and Blahut, 2005)
- Secure MIMO with artificial noise (Negi and Goel, 2005)
- Gaussian MAC and cooperative jamming (Tekin and Yener, 2005)
- Secrecy capacity of slow fading channels (Barros and Rodrigues, 2006)
- Multiple access channel with confidential messages (Liang and Poor, Liu et al., 2006)
- Secure broadcasting with multiuser diversity (Khisti, Tchamkerten, and Wornell, *2006)*
- Ergodic secrecy capacity (Gopala, Lai and El Gamal, Liang, Poor and Shamai 2007)
- Strong secrecy for wireless channels (Barros and Bloch, 2008)

… *and many more.*

Comments

- Information Theory provides you with tools to determine fundamental security limits in particular at the physical layer;
- There exist codes which can guarantee both reliability and information-theoretic security;
- Secure communication over wireless channels is possible even when the eavesdropper has a better channel (on average);
- When it comes to security, fading is a friend and not a foe.