Slepian-Wolf and Related Problems

BASiCS Group, Smartdust, TinyOS, Blackouts

http://basics.eecs.berkeley.edu/sensorwebs

Julius Kusuma

Laboratory for Information and Decision Systems

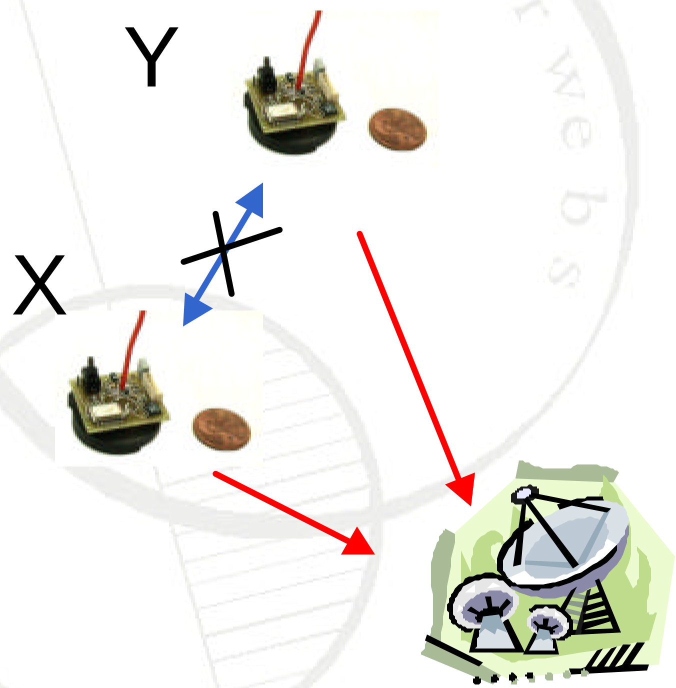
kusuma@mit.edu

Massachusetts Institute of Technology

Outline of presentation

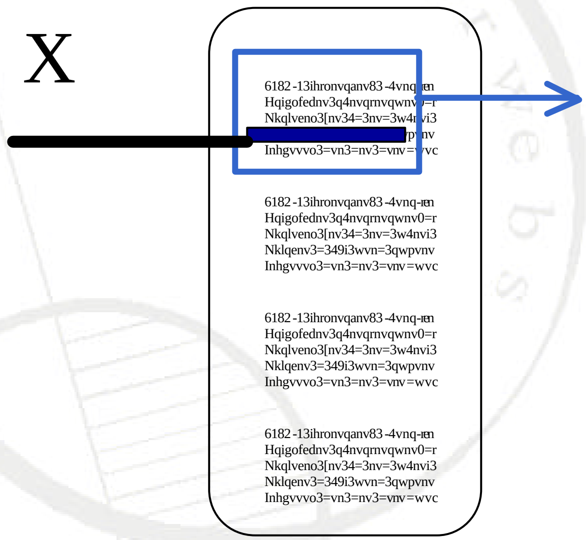
- Information-theoretic motivation: achievable performance
- Algorithmic component for distributed compression
- Code constructions
- Rate-distortion performance
- Optimization of parameters
- Deployment in sensor networks

Distributed compression: basic ideas



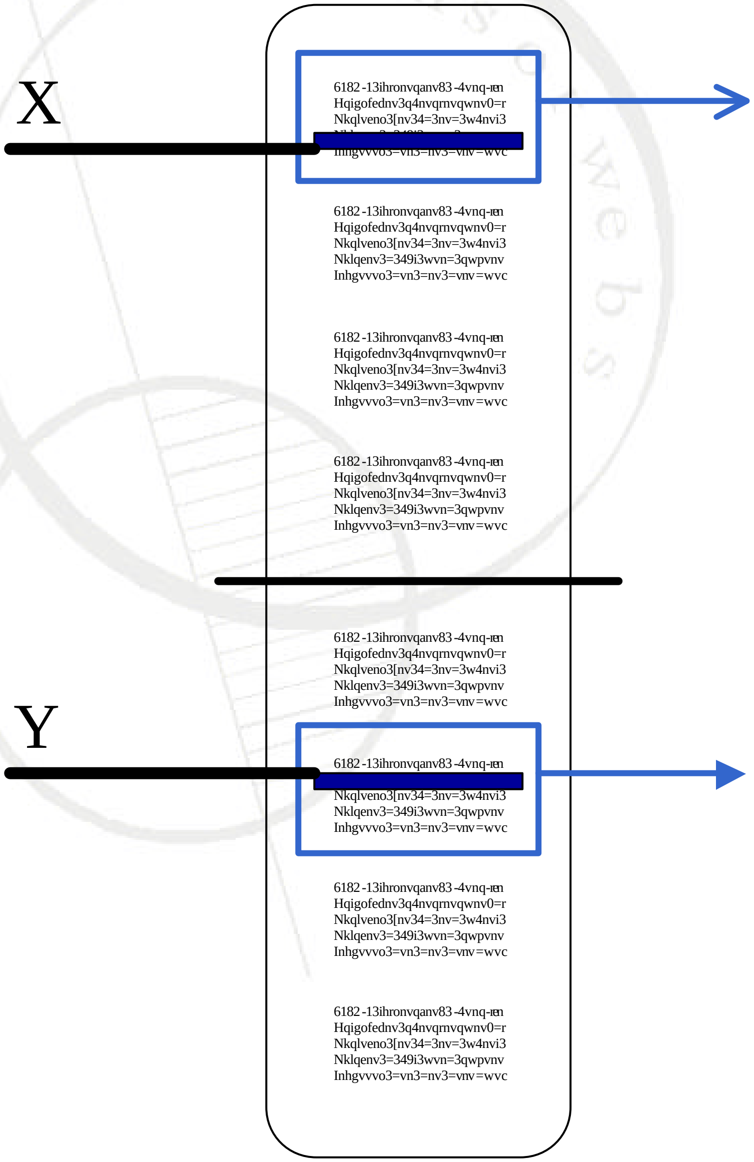
- Suppose X, Y correlated as X=Y+N
- Y available at decoder but not at encoder
- How to compress X close to H(X|Y)?
- Key idea: discount I(X;Y). H(X|Y) = H(X) – I(X;Y)
- For now X and Y iid.

Binning argument

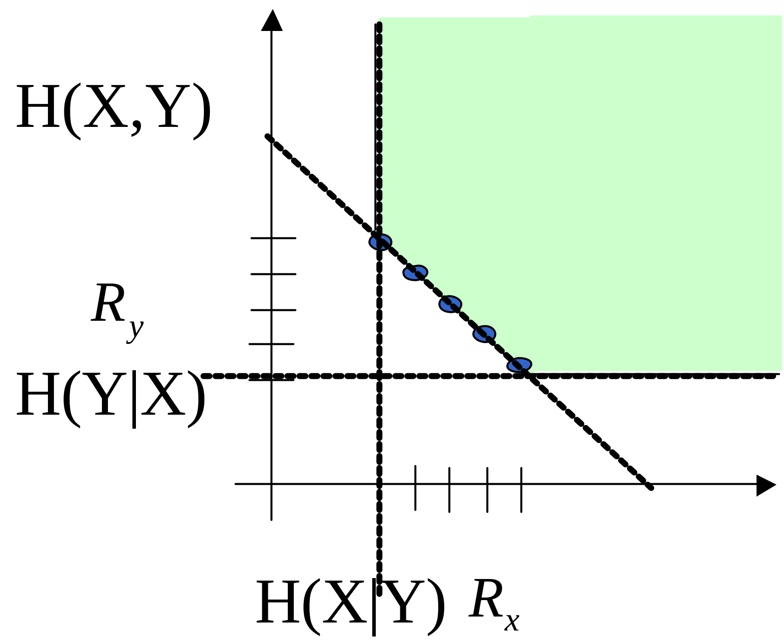


- Make a main codebook of all typical sequences. $2^{nH(X)}$ and $2^{nH(Y)}$ elements.
- Partition into $2^{nH(X|Y)}$.
- When observe X^n, transmit index of bin it belongs to
- Decoder finds member of bin that is jointly typical with Y^n.
- Can extend to “symmetric cases”

Symmetric case: joint binning

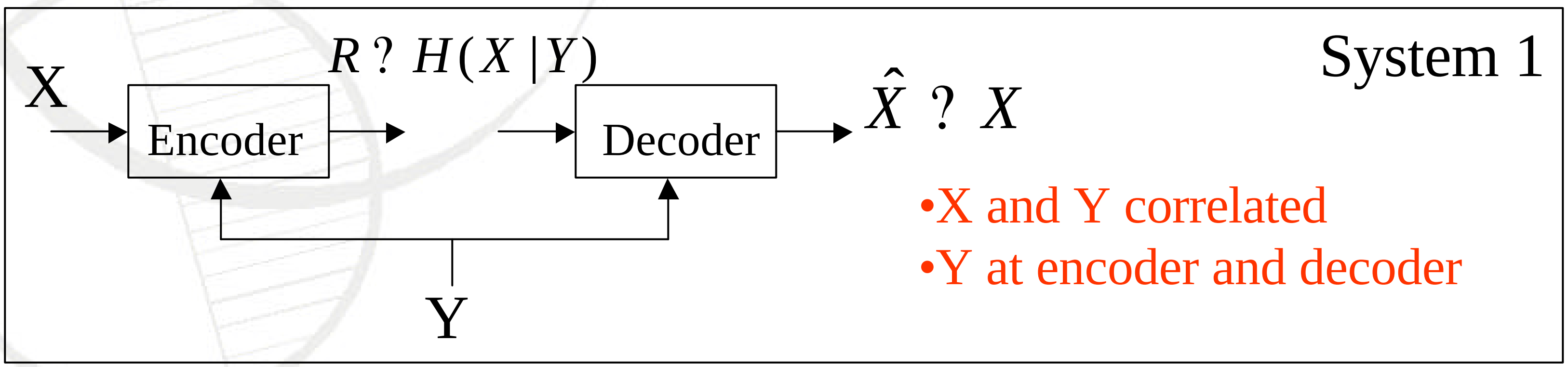


- Rate limited by:
  - R_X ? $H(X|Y)$
  - R_Y ? $H(Y|X)$
  - $R_X + R_Y$? $H(X,Y)$



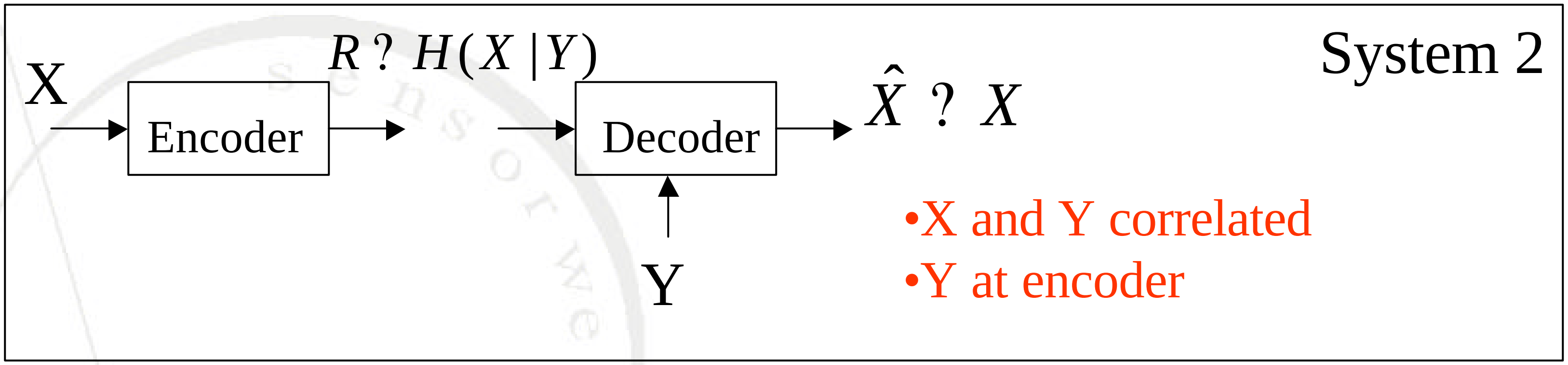
Simple binary example

- X and Y => length-3 binary data (equally likely),
- Correlation: Hamming distance between X and Y is at most 1.
  Example: When X=[0 1 0],
  Y => [0 1 0], [0 1 1], [0 0 0], [1 1 0].



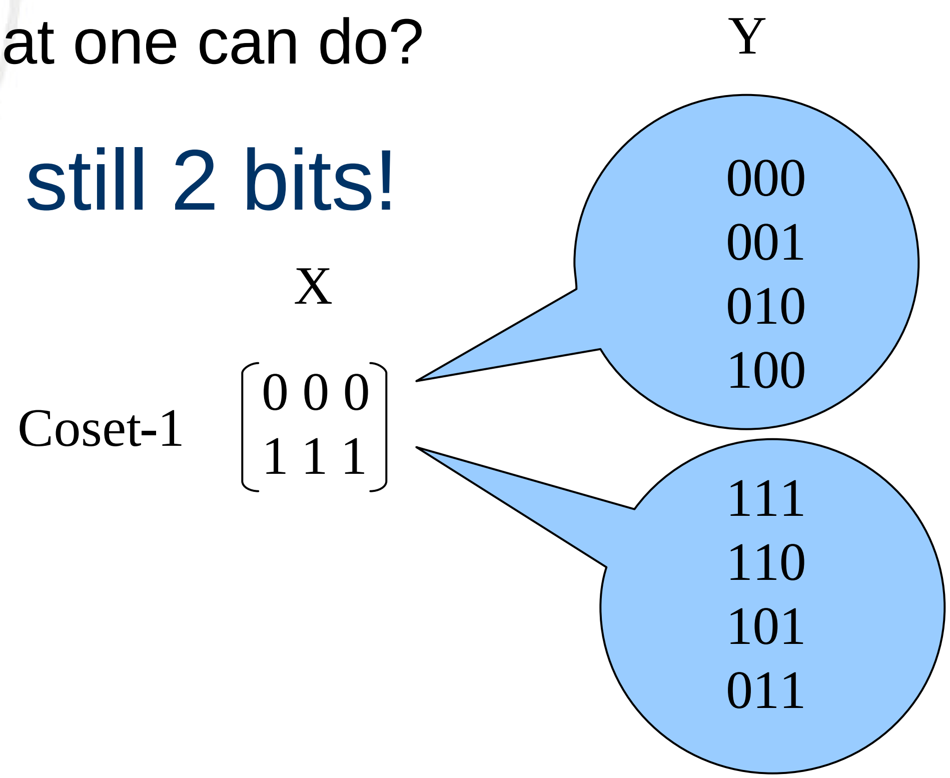
$$X+Y = \begin{cases} 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 1\ 0\ 0 \end{cases}$$

Need 2 bits to index this.



- What is the best that one can do?
- The answer is still 2 bits!

How?



Coset-1 $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ Coset-2 $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Coset-3 $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ Coset-4 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

- Encoder -> index of the coset containing X.
- Decoder reconstructs X in given coset.

Note:

- Coset-1 -> repetition code.
- Each coset -> unique "syndrome"
- DIstributed Source Coding Using Syndromes

Group interpretation of "binning"

- Rules of thumb:
  1. Want high density of elements in codebook
  2. Want members of each bin as far apart

- Consider **error-correcting codes**!
  Codes select a (normal) subgroup of all possible elements.
  Members of a subgroup is as far apart as possible.

- Error occurs when distance between side info to main info > d_{min}

- Example: (3,1) repetition codes: can compress if $d_H(X,Y)<2$

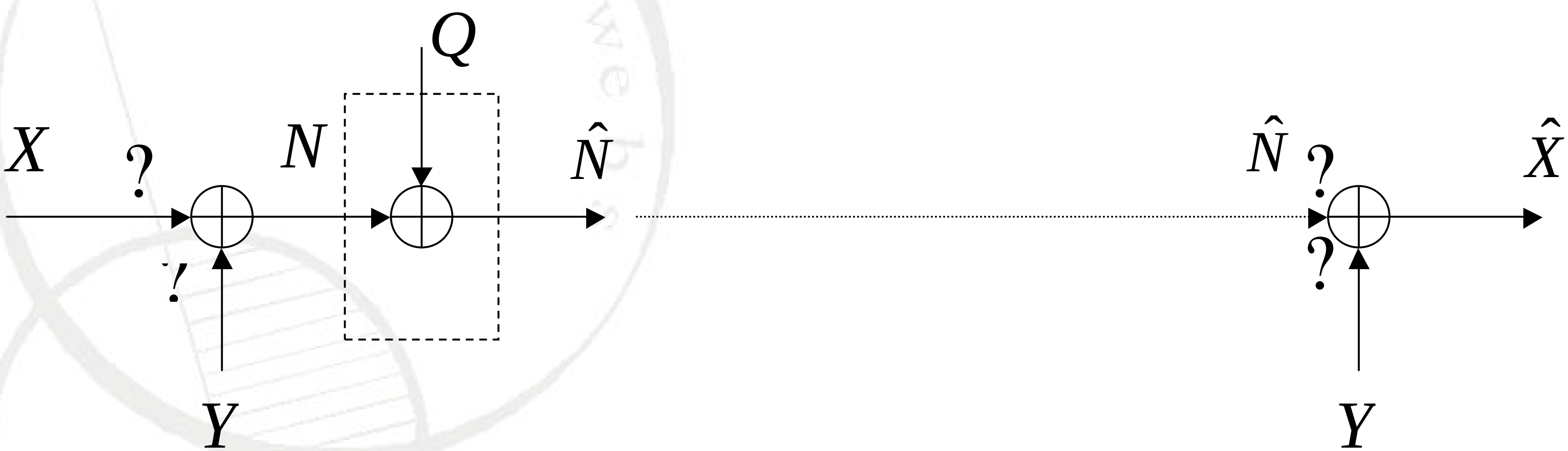
coset-00	000 111	coset-10	001 110
coset-01	010 101	coset-11	100 011

Intuition behind source coding with side info.

Why does it not matter if encoder doesn't have Y?

<u>Case I: Y present at both ends</u>

X=Y+N



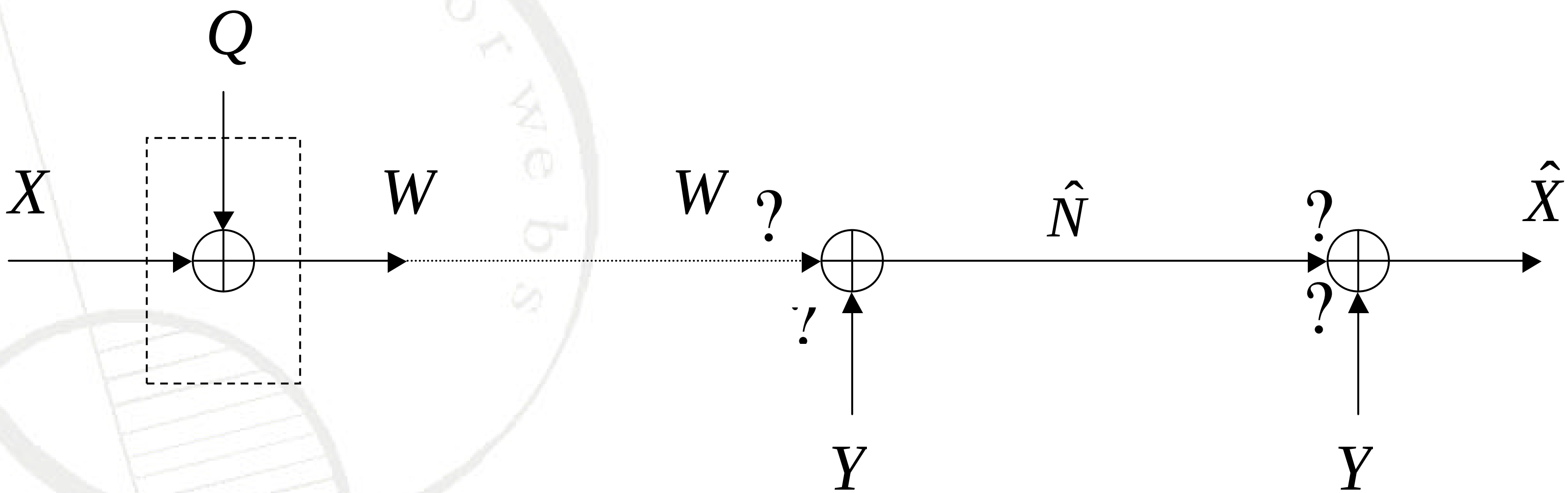
- X = Y+N, where N is Gaussian (note X and Y need not be Gaussian)
- Subtract Y and quantize only N, add Y back at the decoder.
- Transmission rate: $I(N;\hat{N}) = h(\hat{N}) - h(\hat{N} | N)$

$$= \tfrac{1}{2}\log\left(\frac{\sigma_n^2 + \sigma_q^2}{\sigma_q^2}\right)$$

Intuition (contd.)…

Case II: Y present at decoder only:

X=Y+N



- Quantize to same rate, subtract and add Y back at the decoder.
- Transmission rate:

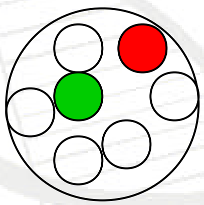
$$= I(W;X) - I(W;Y)$$

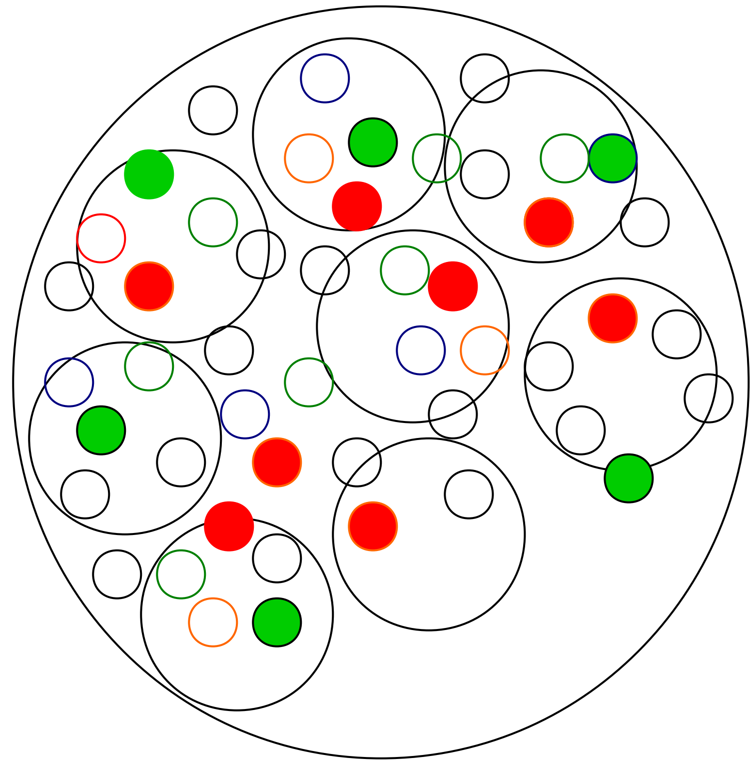
$$= h(W|Y) - h(W|X)$$

$$= h(N+Q) - h(Q)$$

$$= \tfrac{1}{2}\log\left(\frac{\sigma_n^2+\sigma_q^2}{\sigma_q^2}\right)$$

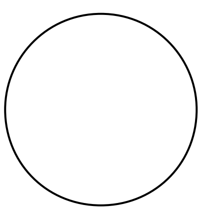
Geometric Interpretation



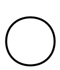


Case I: Y at both sides

Case II: Y at decoder only



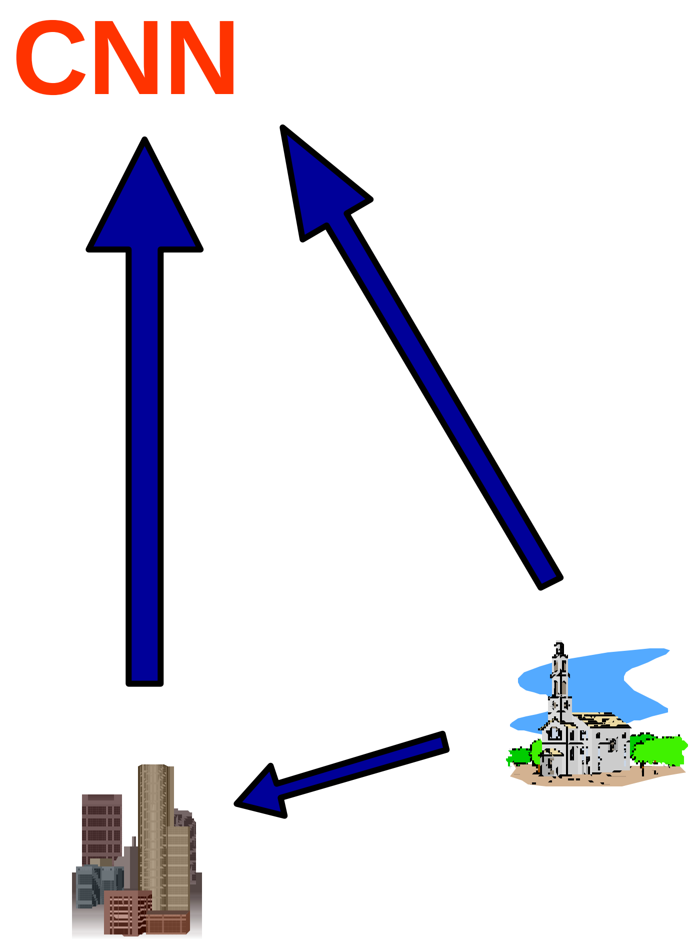
radius $?_{n?q}$



radius $?_{q}$

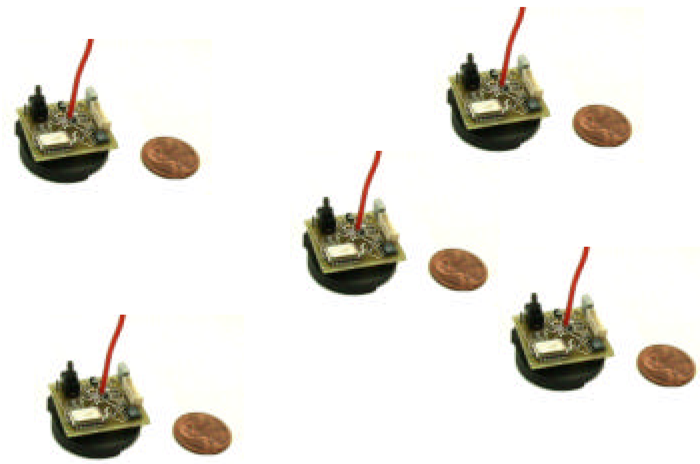
Sending the difference telepathically

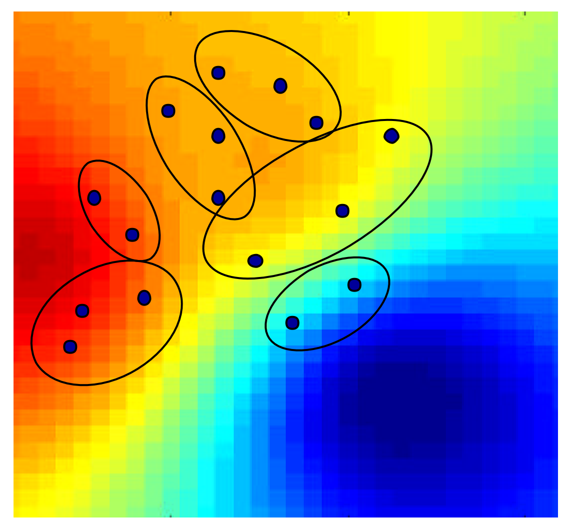
- Jump ahead to a real-world example:
  X - Temperature in Boston
  Y - Temperature in Providence
- Suppose we can bound difference, most of the time < 8 degrees
- If Boston knows the reading of Providence, can just send difference.
- But this means that the information Y must be available at both Boston and Providence!
- Establishing communication network expensive in a sensor network!



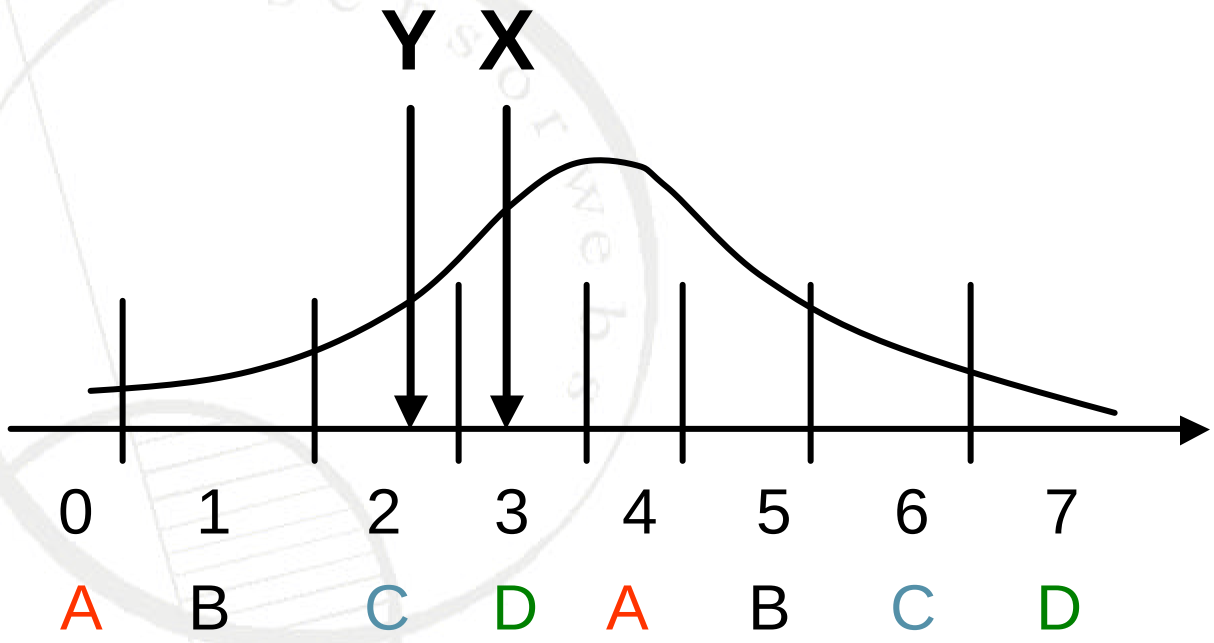
Motivations for sensor networks

- Dense sensor network ✍ high spatial redundancy
- Need to remove redundancy without communication
- Assume statistical correlation properties of neighboring nodes are known/learnt





Consider the following idea



- Difference at most 1 cell.
- Send only index of “coset”: A,B,C,D
- Decoder decide which member of coset is the correct answer

We have compressed from 3 bits to 2 bits

Coding operation of “binning”

- Performance determined by selection of main and subgroups.

- Tradeoff:
  Quantization error: determined by the main group
  Coset error: determined by the subgroup

- **Quantization error**: want main group to be dense
- **Coset error**: want intra-coset distance of subgroup to be as large as possible

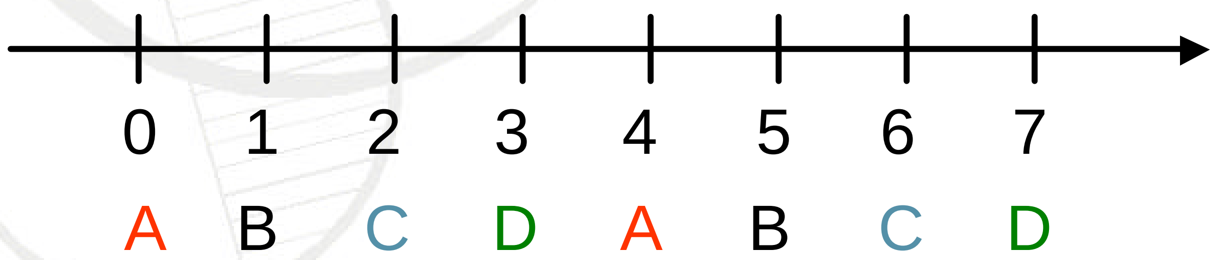
Gentle intro to groups and codes

- Key idea: algebraic codes is a subgroup of (discrete) signal set.
- For example: (7,4) Hamming code is subgroup of $\{0,1\}^7$.
- Therefore **codes induce a (geometrically uniform) partition!**
- We develop several examples in the following

Partitioning a scalar quantizer

- Start with a scalar quantizer

- Partition into PAM signals

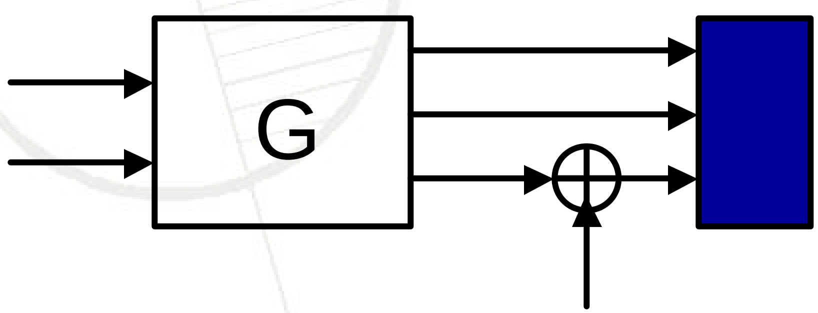


- Call this SQ-PAM

Better idea: using TCM codes

- Objective of algebraic codes: sphere packing – densest packing for distance and rate.

- Use a TCM code to partition ?^{L}.



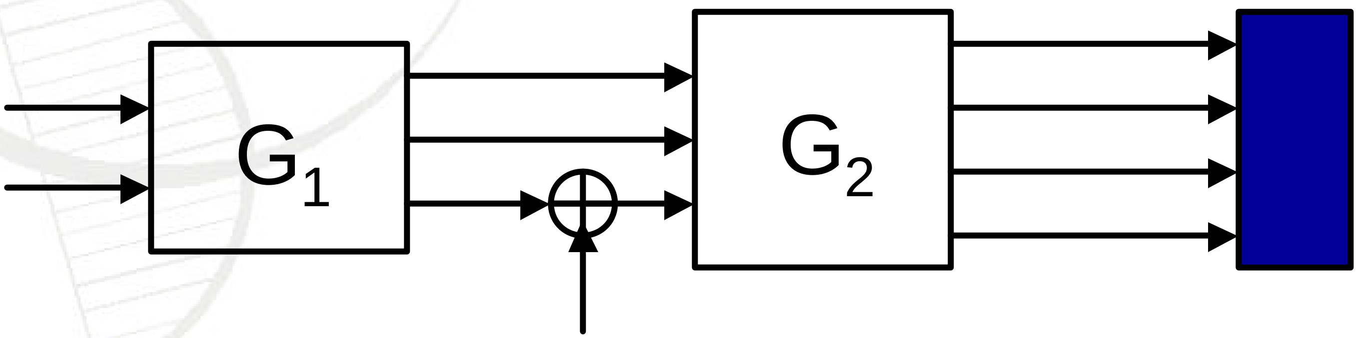
Coset selector

$$G=\begin{bmatrix} D & 1 & 0 \\ 1 & D^2 & D \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

And yet better … !

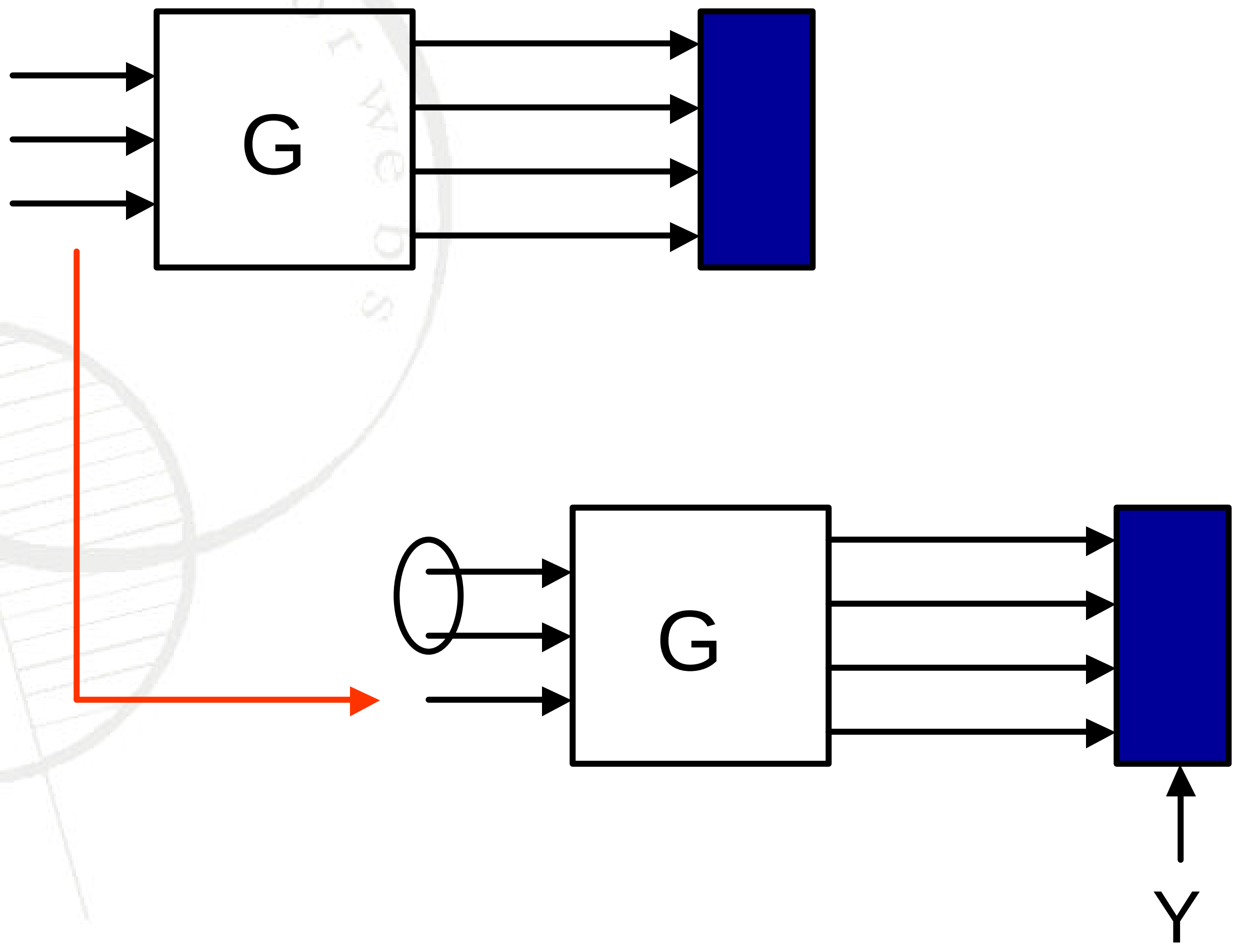
- Can also induce partition on codes themselves by choosing subcodes.



- Called: TCQ-TCM

Alternative representation

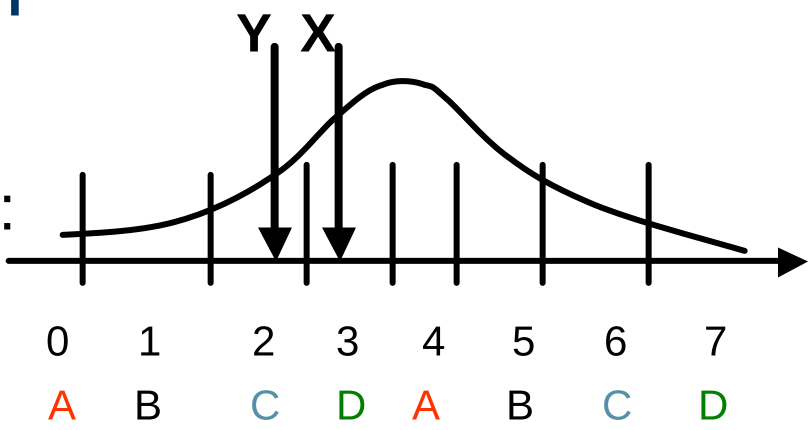
- Subspace of a code is a subgroup of the code



Note: send LSB of codewords!

SQ-PAM: Scalar quantization, pulse-amplitude modulation

- Back to previous example:

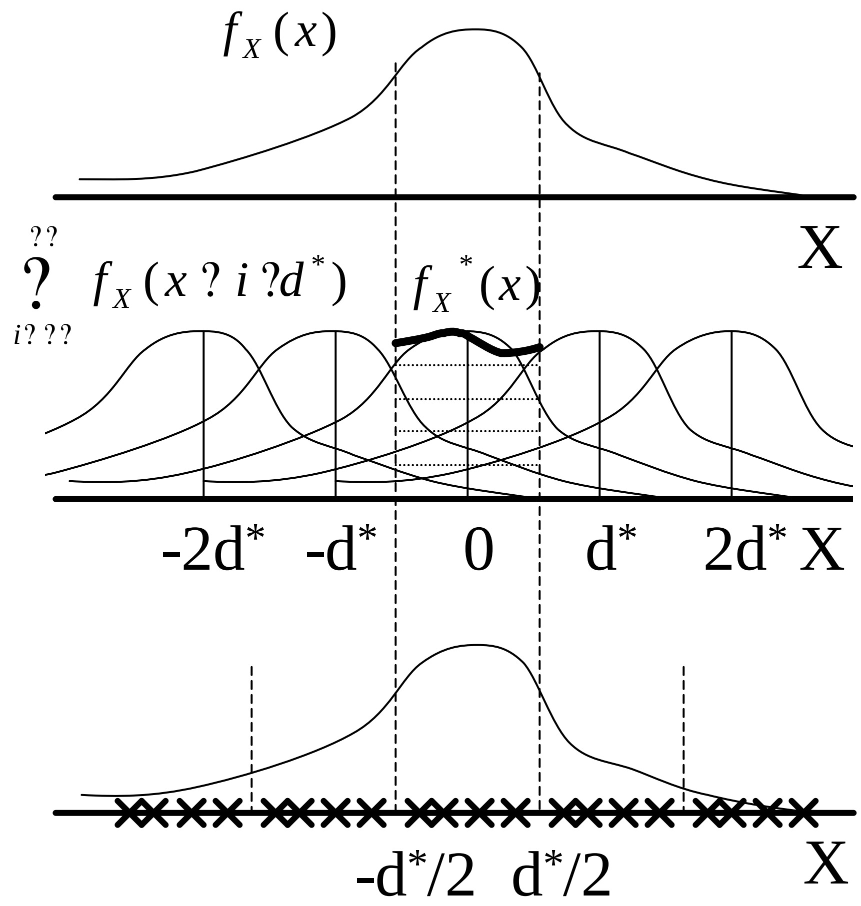


- The letters index different cosets of a PAM code.
- Start with scalar quantization.
  Encoder calculates the index of the bin.
  Transmit index of the bin.

  Decoder receives index of the bin.
  Use correlated reading to determine which member of the bin is correct.

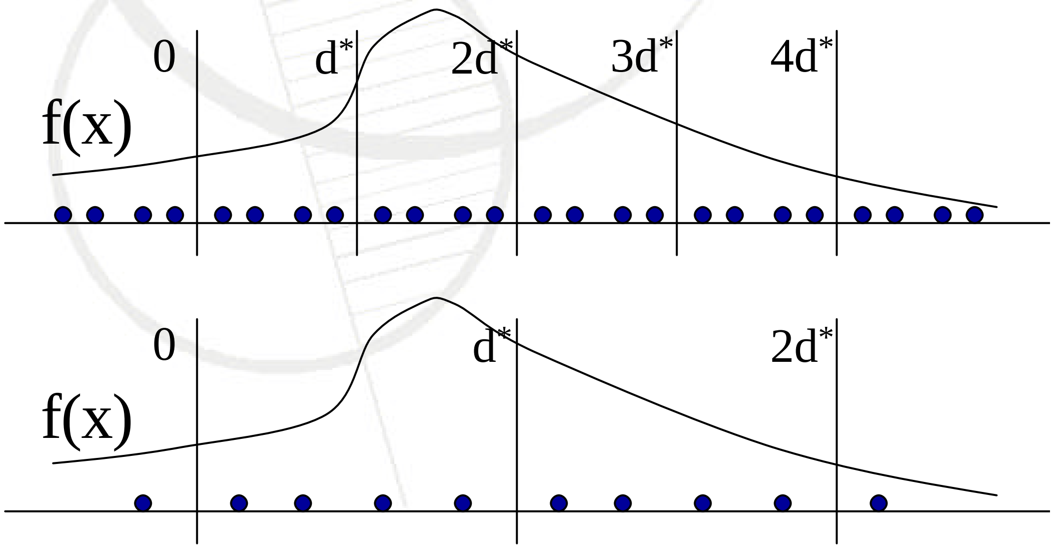
Observation: quantization indifference

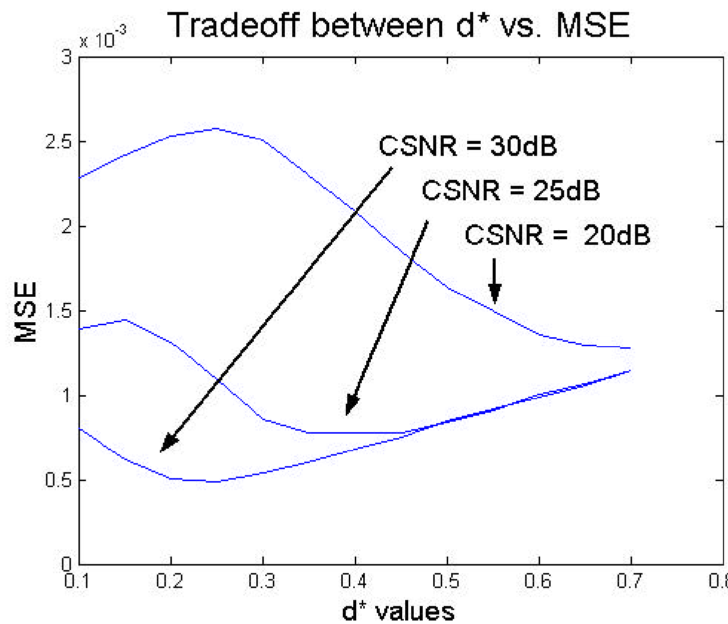
- Important note: Quantizer can't differentiate A from A
- Therefore: Must combine statistics of members of bins
- Use **PDF periodization**: repeat PDFs using parameter d*.
- Design using **f'$_x$(x)**



Caveats: choice of d*

- If too small: high coset error
- If too large: high quantization error

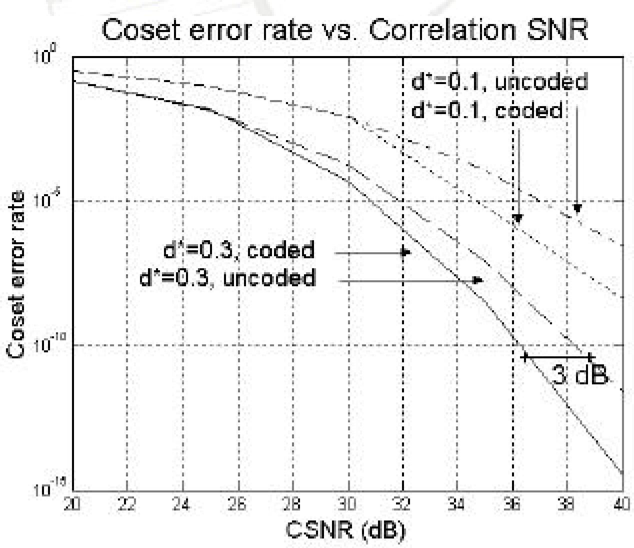




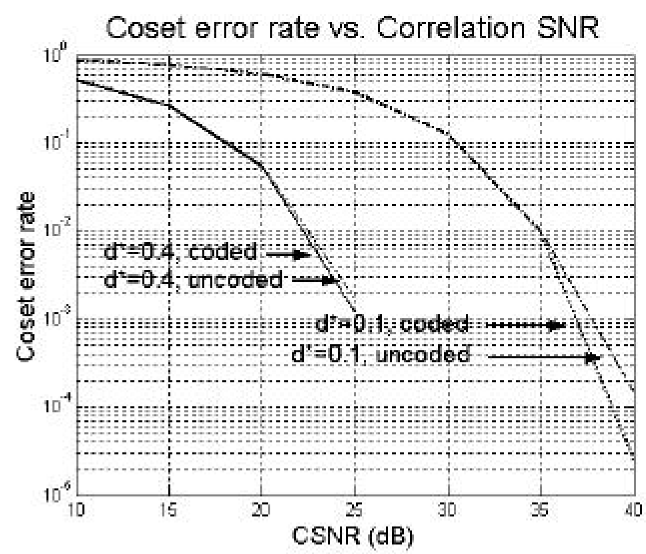
Coding performance

- Can use **SQ-TCM** (Trellis Coded Modulation), **TCQ-TCM**.

- More details in:
  Pradhan, Ramchandran, “DISCUS: Distributed Coding Using Syndromes”, DCC 1999 and 2000
  http://basics.eecs.berkeley.edu



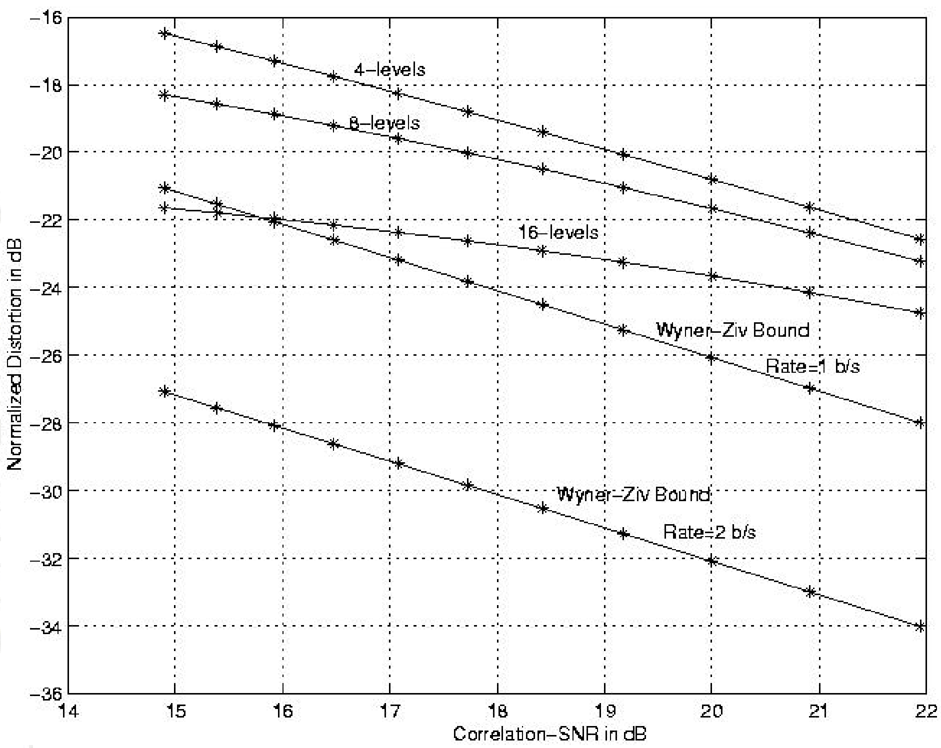
Gaussian



Uniform

Theoretical bound

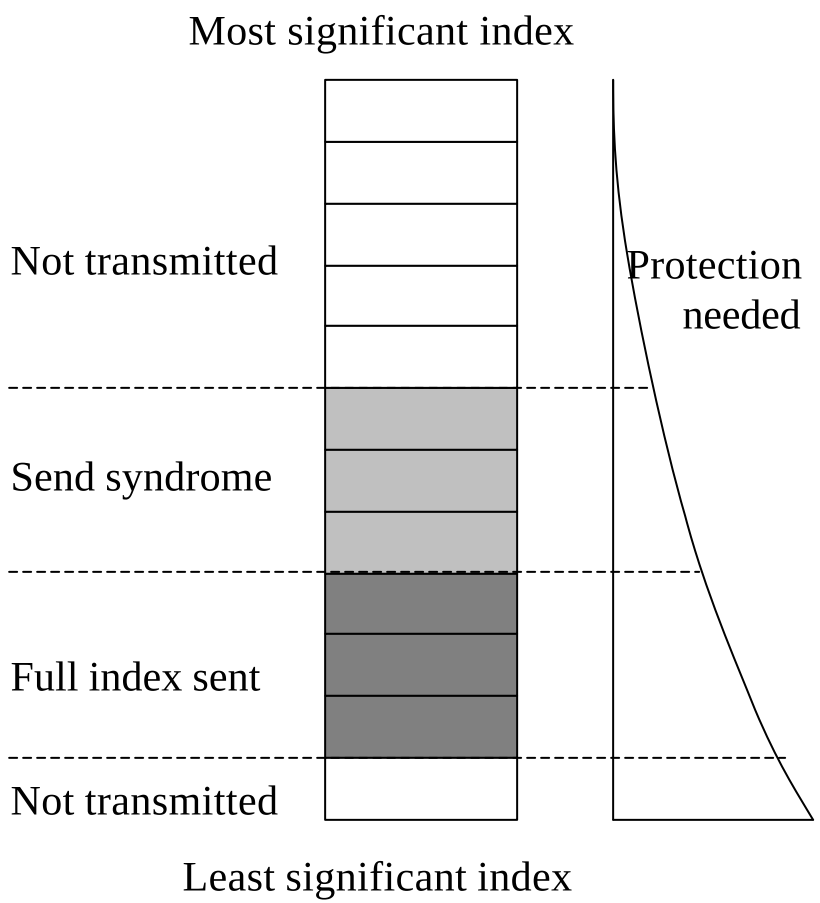
- Can get within 2-3 dB of Wyner-Ziv's bound using Trellis codes.



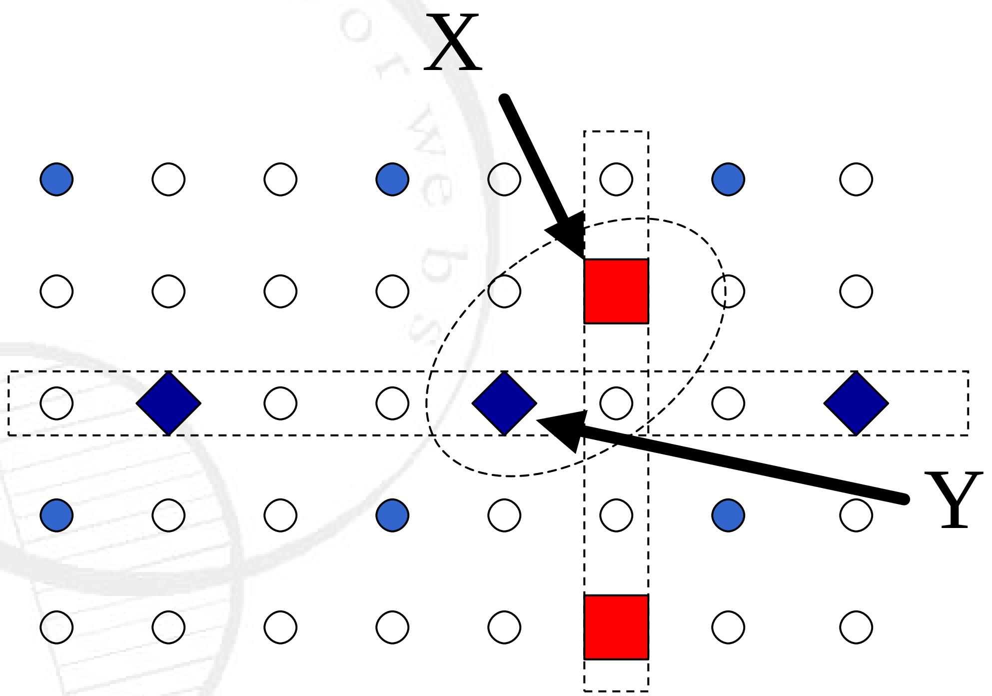
Transmission rate 1 bit per symbol

Dynamic bit allocation

- Consider iterative method: assign one bit at a time
- Can either:
  improve quantization
  improve code performance
- Iteratively assign using rules of thumb.
- Multiple levels of protection



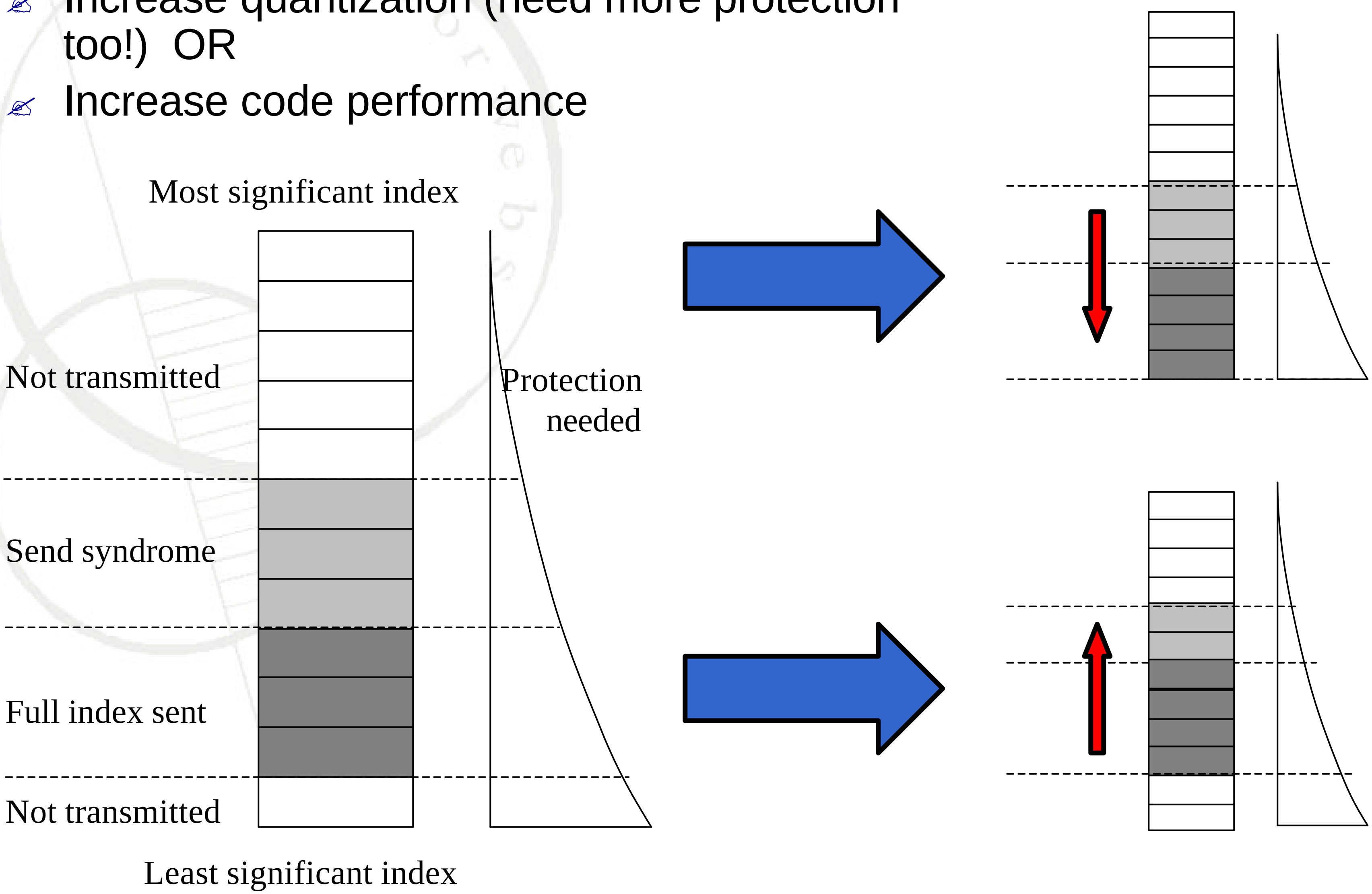
Lattice illustration of symmetric decoding



- Finding nearest neighbors of selected cosets

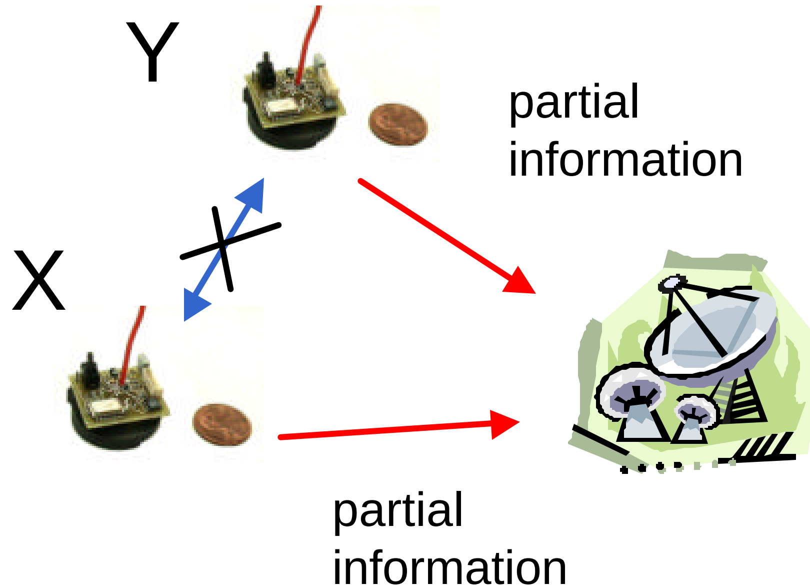
For example

- Increase quantization (need more protection too!) OR
- Increase code performance

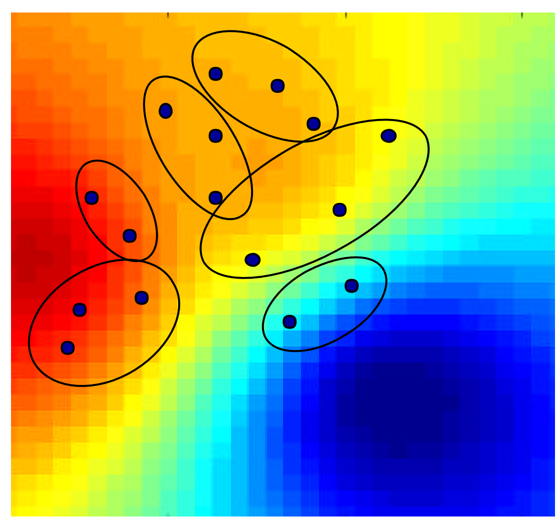


Symmetric rates and network applications

- Can also have motes all send "partial information".



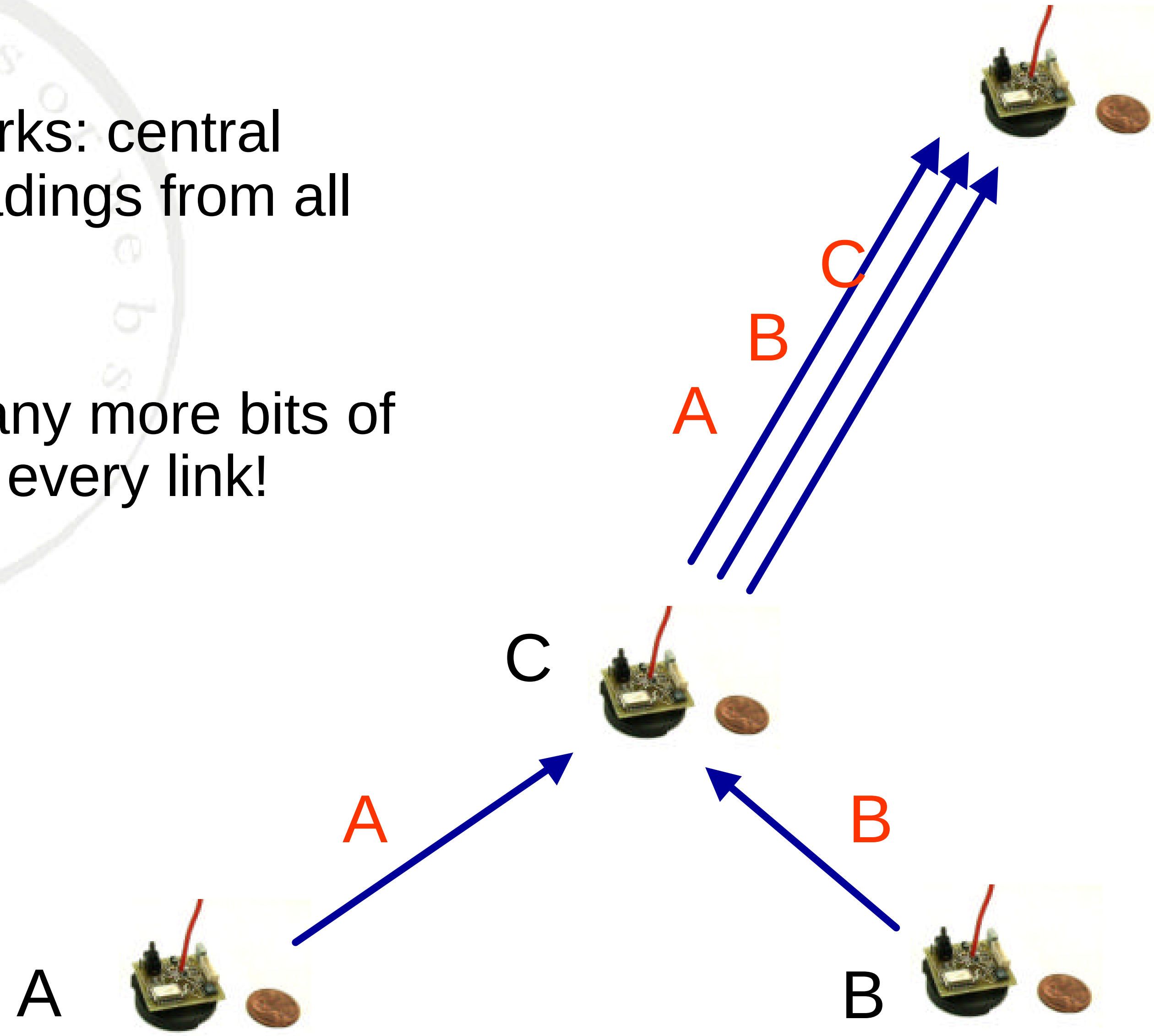
- Use clustering to enable network deployment



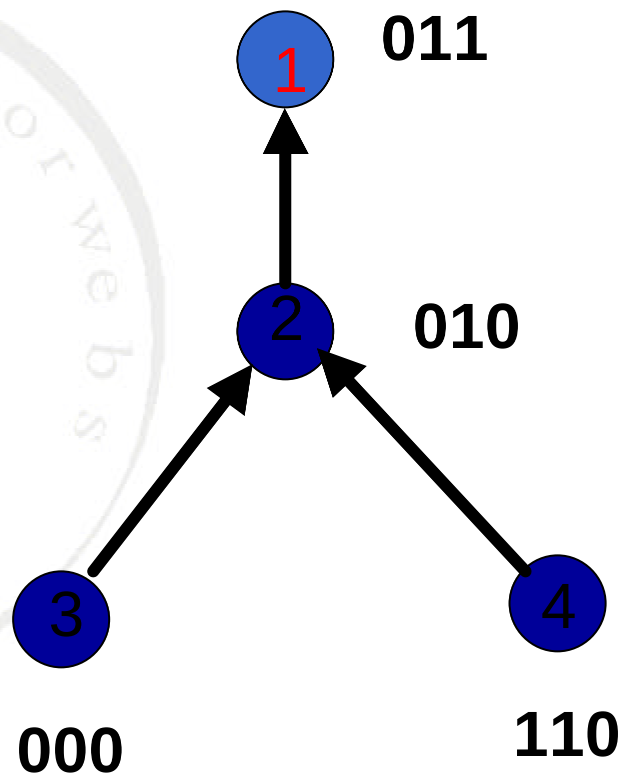
Multiplicity effect in multihop networks

- Multihop networks: central node wants readings from all other nodes.

- Get multiply many more bits of data to send at every link!



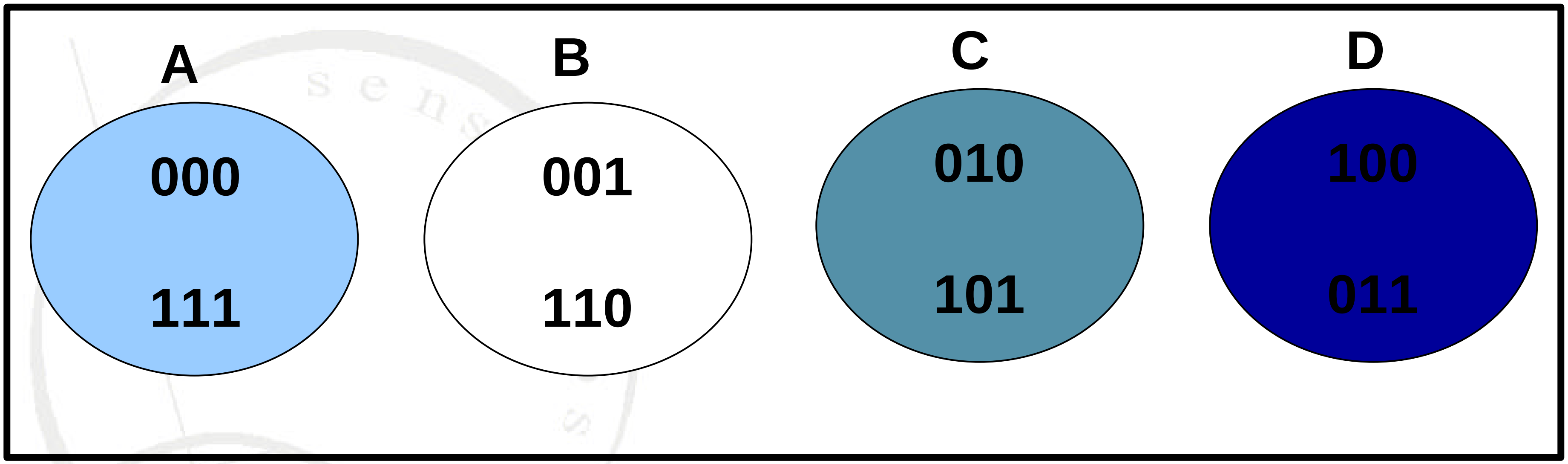
Example of simple sensor network with one gateway



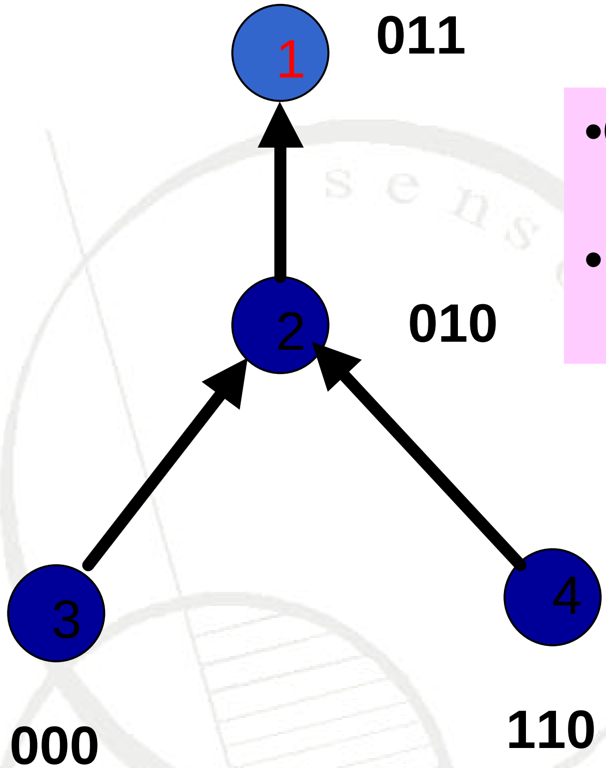
Objective: Gateway node **1** has to get sensor readings from network

Assumptions:

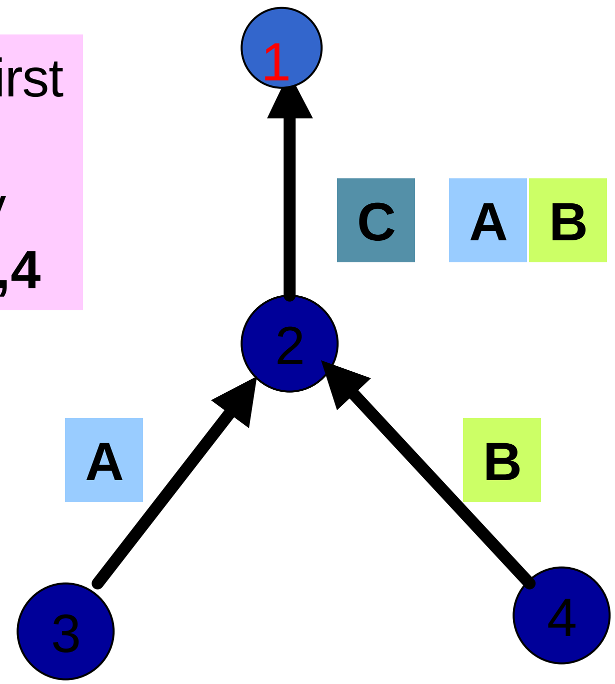
- Sensor readings are 3-bit quantized representations
- Children nodes differ in ***at most one bit*** from their parent

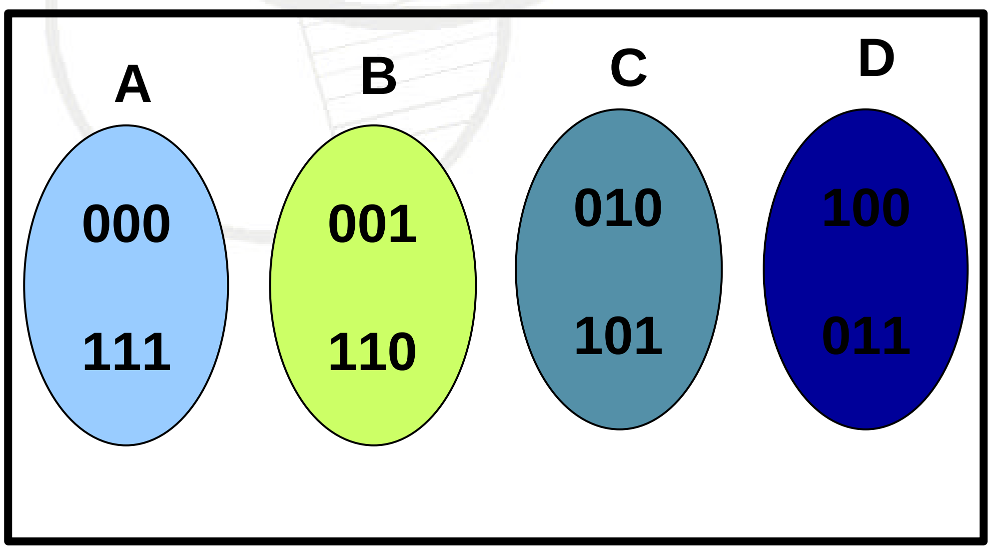


DISCUS: Set of all 3-tuples partitioned into **4** cosets **A, B, C, D.** Distributed encoder assigns each 3-tuple with a 2-bit index: **A=00; B=01; C=10; D=11**



- •Gateway node **1** first decodes node **2**
- It then recursively decodes nodes **3,4**





If each link ~ 1 m, network does **15 bit-meter** work **w/o DISCUS**

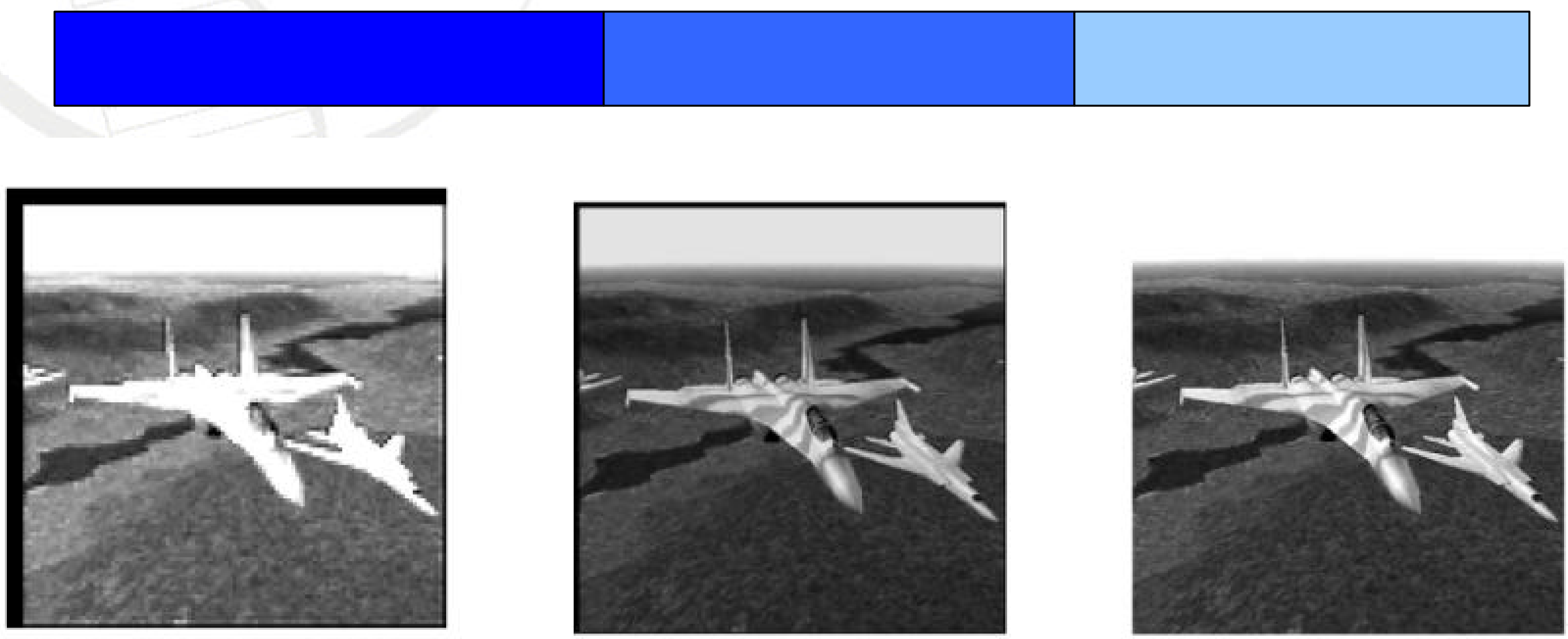
With DISCUS, network does only **10 bit-meter** work.

Real-life application: Blackouts project, etc.

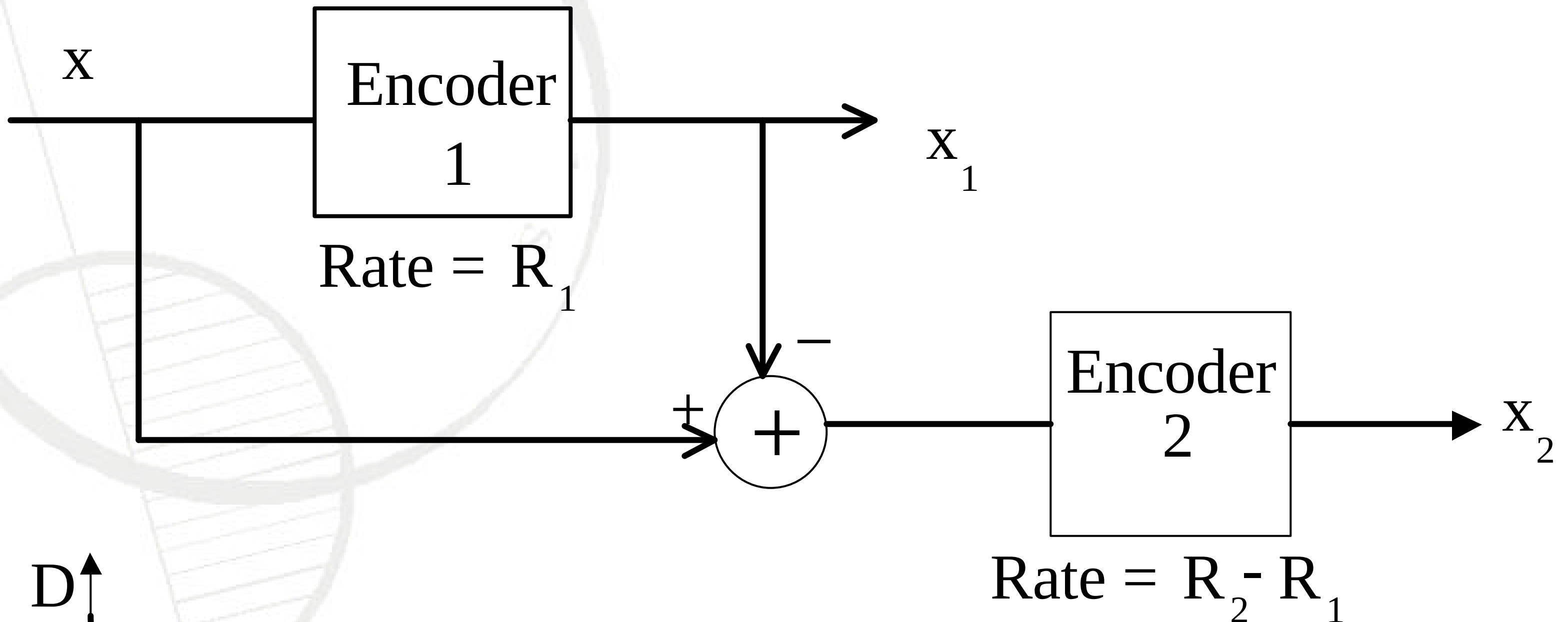
- Near real-time monitoring of room conditions
  http://blackouts.eecs.berkeley.edu
- Ad-hoc networking, data goes online (WWW)
- Earthquake, engine data (with LBNL)

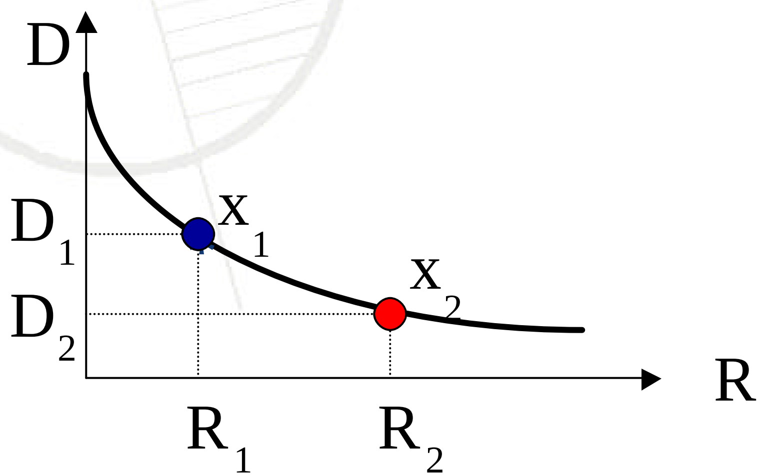
Emerging Image/Video Coding Standards (progressive format)

- Multi-resolution coding: e.g, JPEG-2000, MPEG-4.
- Bit stream arranged in importance layers (progressive)



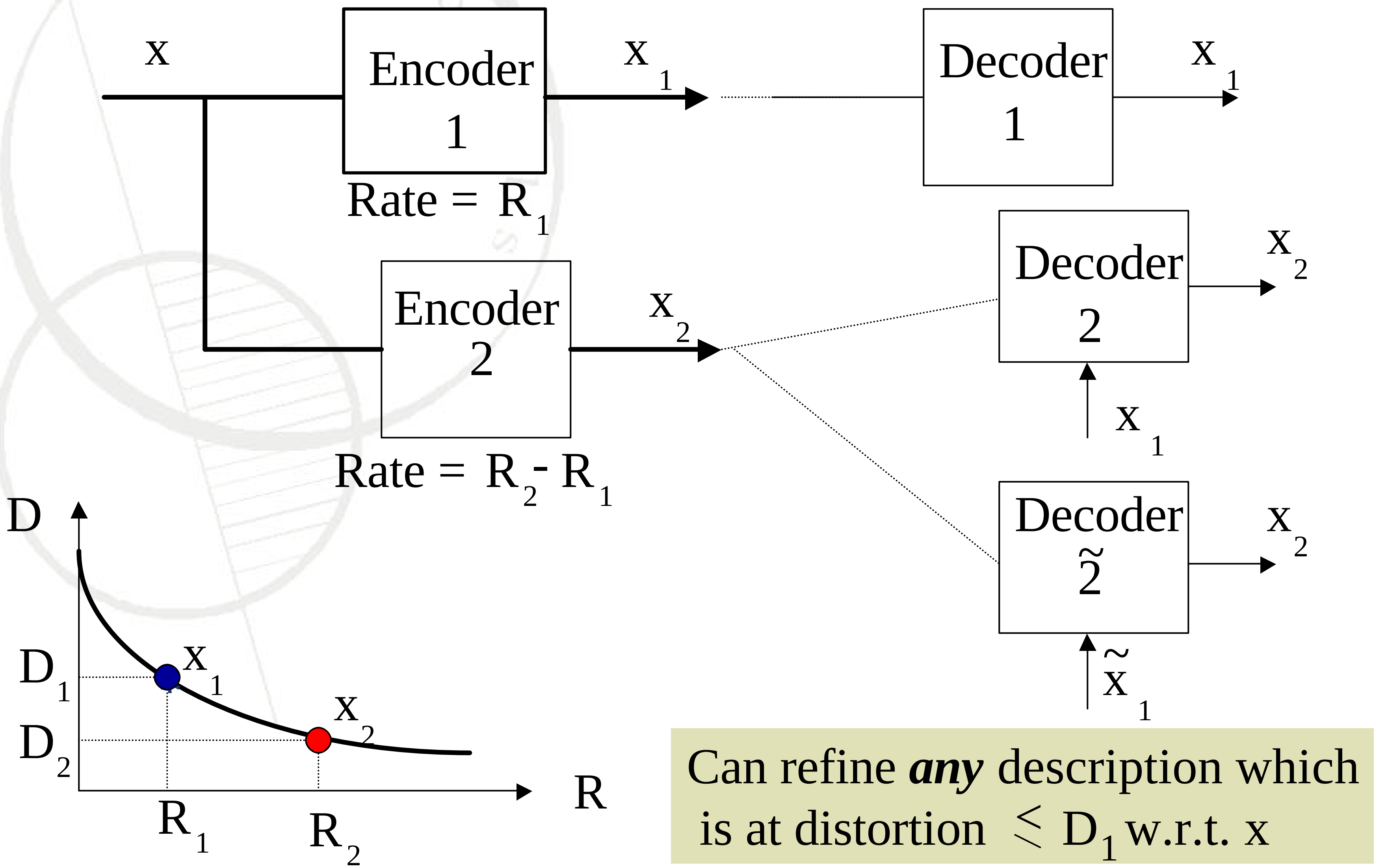
Conventional Successive Refinement





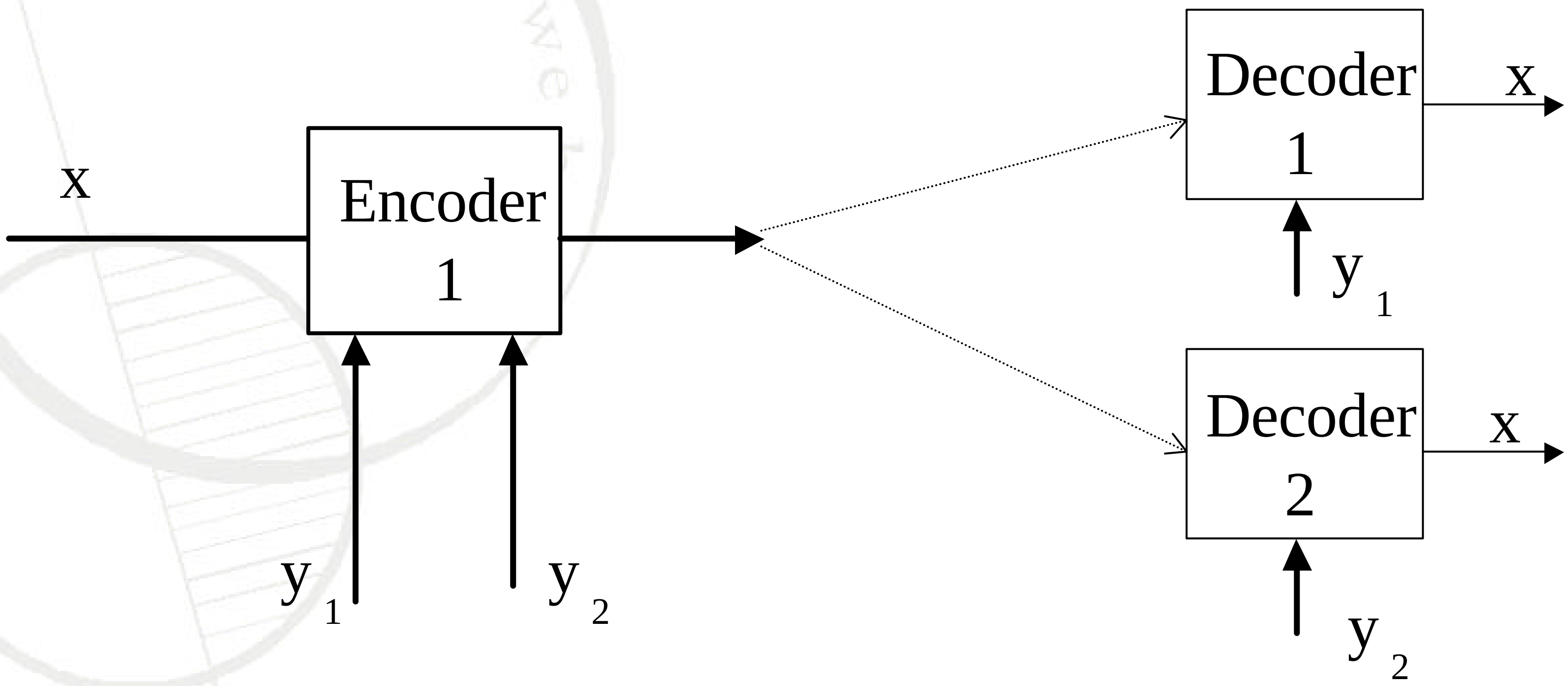
The other SR perspective ...

- Wyner-Ziv successive refinement: “**universal**”



Can refine ***any*** description which is at distortion $\leq D_1$ w.r.t. x

Multiuser successive refinement…



Even when no “explicit” need for invoking coding with side information, we believe that this is the most efficient strategy for ***broadcast source coding***.

Conclusions

- Can effectively take advantage of correlation. Improved performance compared to same bit budget without side information.
- Can use efficient encoding/decoding algorithms.
- Simple design using well-known tools.