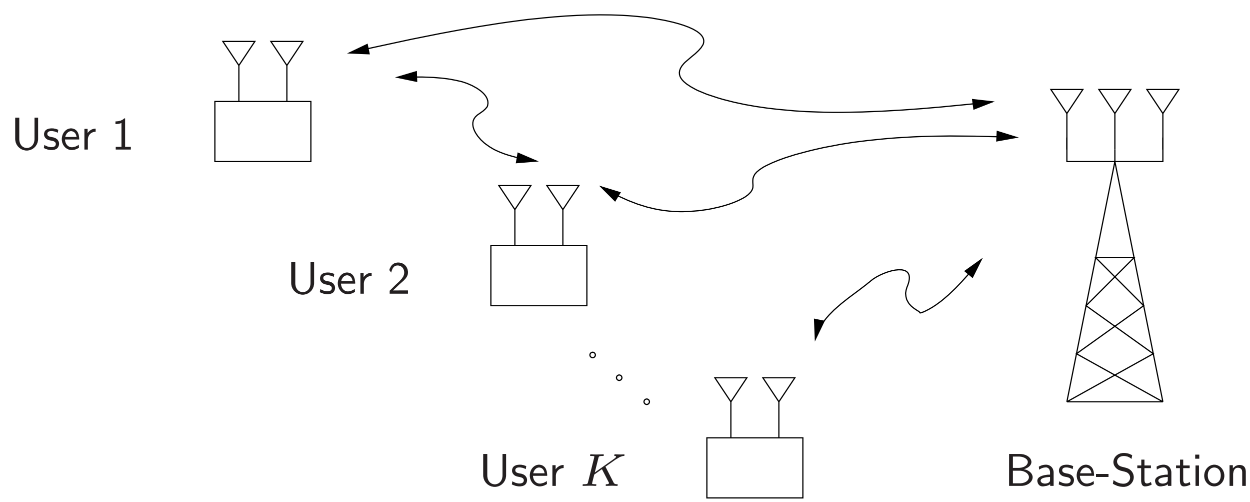
Parity Forwarding Strategies for the Relay Channel

Wei Yu

Electrical and Computer Engineering Department
University of Toronto

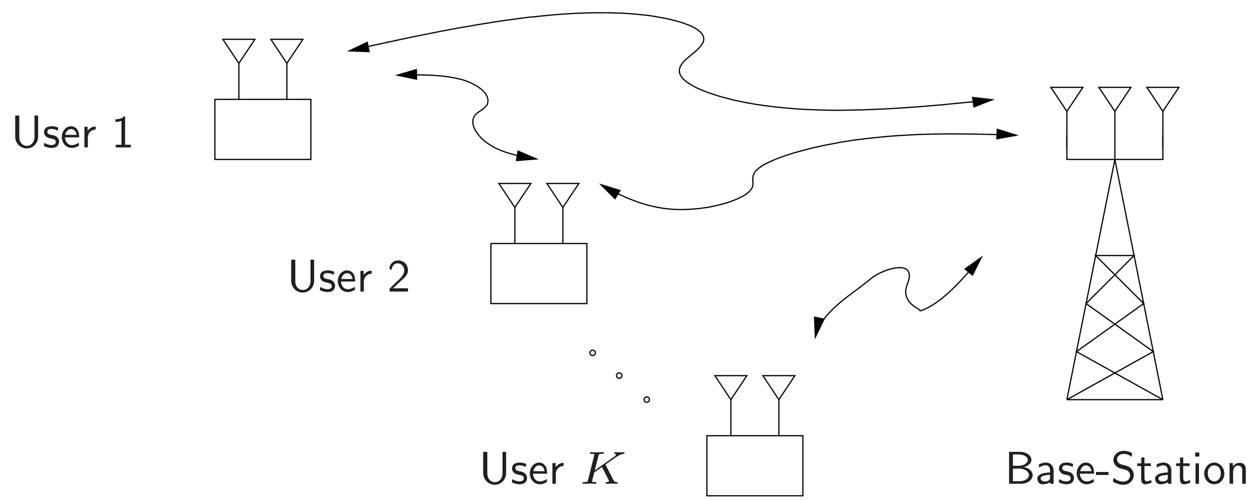
(Joint work with Peyman Razaghi, Marko Aleksic)

Relay Network



- Fundamental relay strategies (Cover and El Gamal '79)
  - "Decode-and-forward" and "Quantize-and-forward"

Relay Network

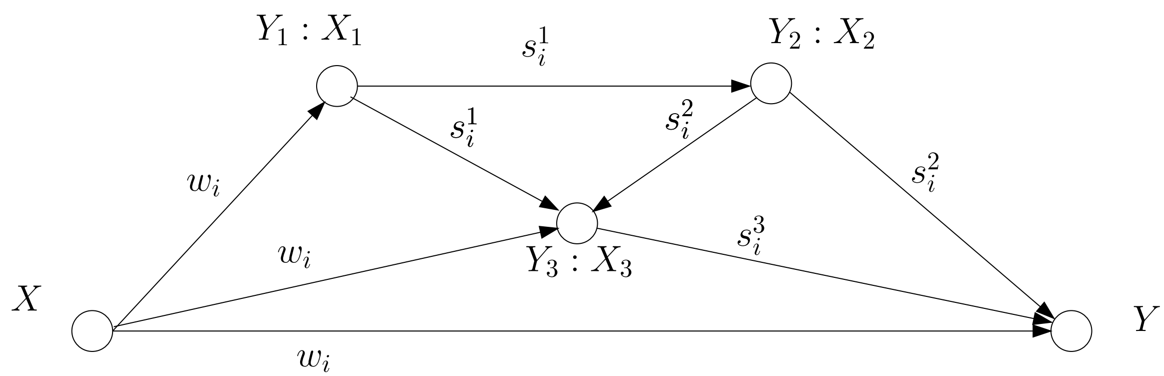


- Fundamental relay strategies (Cover and El Gamal '79)
  - "Decode-and-forward" and "Quantize-and-forward"
- Both are examples of a "parity-forwarding" strategy.

Outline of This Talk

- Decode-and-forward as a parity-forwarding strategy.
  - Part I: LDPC code design to approach the DF capacity.
- Generalization of decode-and-forward to multi-relay networks.
  - Part II: Binning and parity-forwarding for multi-relay networks.
- Is quantize-and-forward ever optimal?
  - Part III: Capacity of a class of modulo-sum relay channels.

Information Flow in a Relay Network

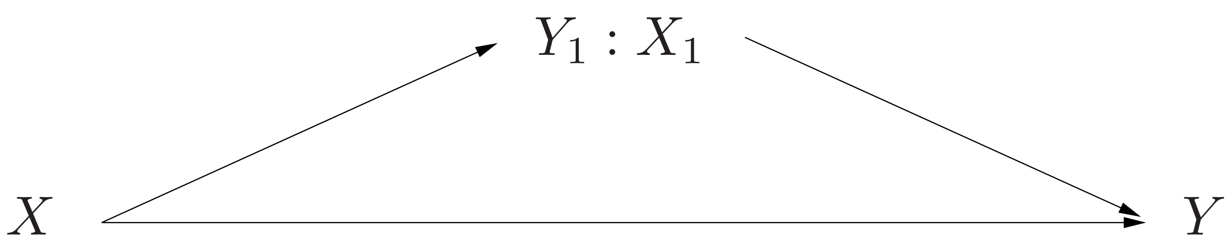


- Relay nodes summarize its own knowledge using parity bits.
- Design challenges:
  - Routing of information in a network.
  - Efficient codes to facilitate decoding at the relays/destination.

Part I: LDPC Code Design for Decode-and-Forward

Binning in Decode-and-Forward

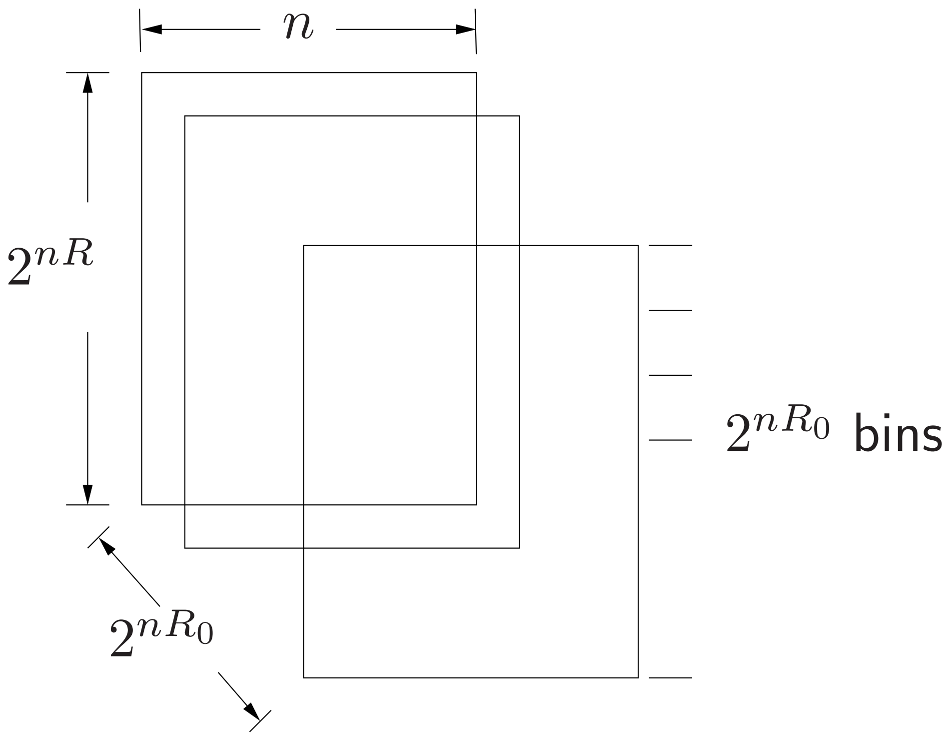
- Consider Cover and El Gamal's strategy for degraded relay channel:



- Two elements: Block-Markov coding and Binning
  - The relay provides a bin index of the transmitter codeword.

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

Code Construction



Doubly indexed codebook $X^n(w|s)$

$w \in \{1, 2, \cdots, 2^{nR}\}$

$s \in \{1, 2, \cdots, 2^{nR_0}\}$

2^{nR} is partitioned into 2^{nR_0} bins

Second Codebook $X_1^n(s)$

What is binning?

- Binning is ubiquitous in multiuser information theory
  - Writing on dirty paper (Gel'fand-Pinsker)
  - Source coding with encoder side information (Wyner-Ziv)
  - Relay communication (Cover-El-Gamal)
- Binning is a way of conveying "partial" information.

What is binning?

- Binning is ubiquitous in multiuser information theory
  - Writing on dirty paper (Gel'fand-Pinsker)
  - Source coding with encoder side information (Wyner-Ziv)
  - Relay communication (Cover-El-Gamal)
- Binning is a way of conveying "partial" information.

Bin index is equivalent to parity-checks

Bin Index as Parity-Checks

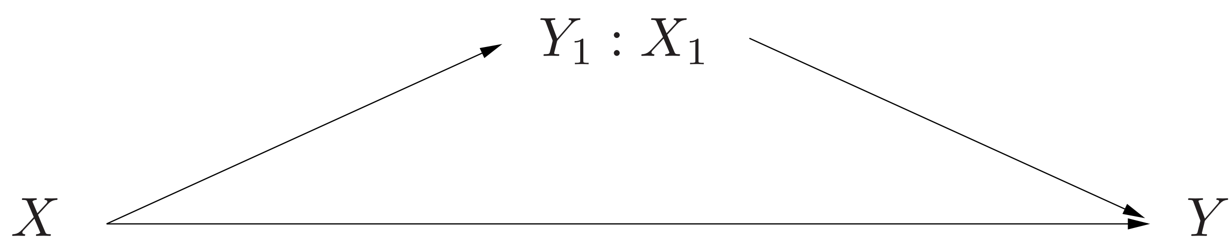
How do we partition a codebook of size 2^{nR} into 2^{nR_0} bins?

... by forming nR_0 parity check bits,
and using the parity check bits as bin indices.

Same idea as DISCUS for Slepian-Wolf coding (Pradhan-Ramchandran)
or structured binning (Zamir-Shamai-Erez)

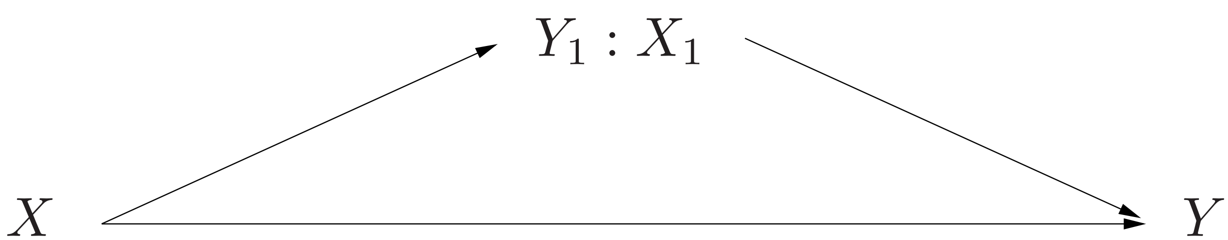
Decode and Forward

- X_1 decodes X and re-encodes parities (or a bin index) of X.



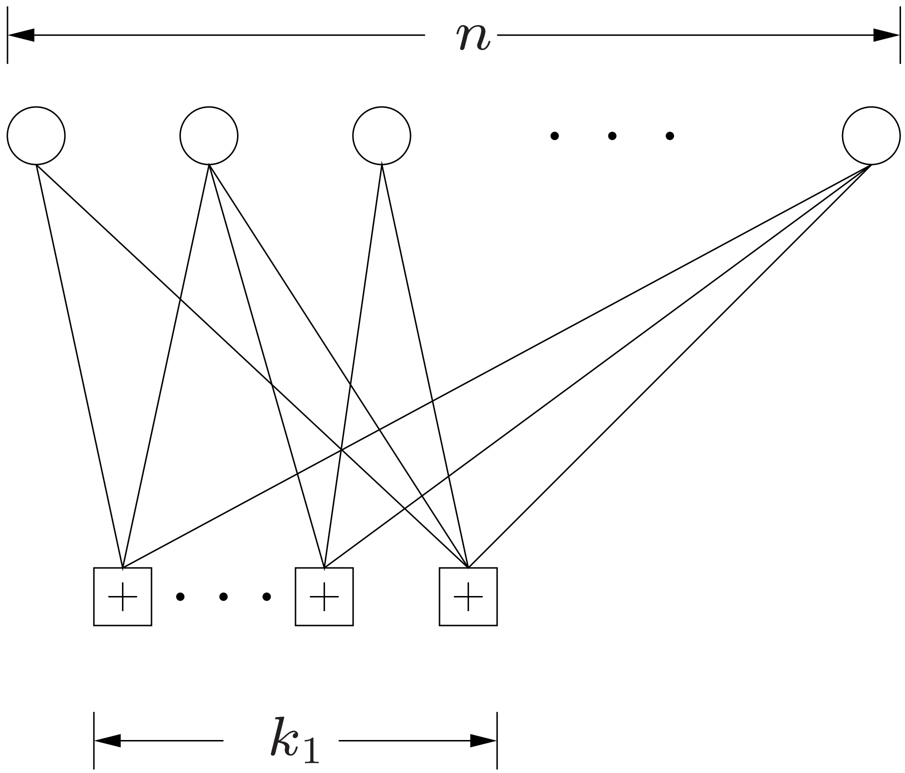
Decode and Forward

- X_1 decodes X and re-encodes parities (or a bin index) of X.



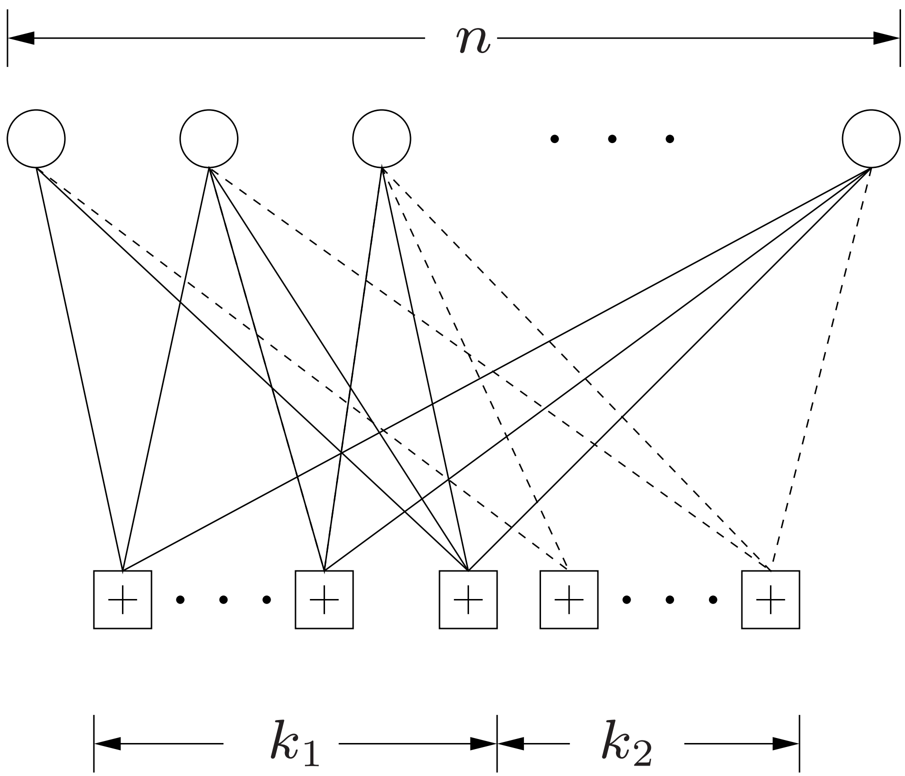
- A good code for the relay channel must be capacity-approaching
  - for the $X - Y_1$ link at R;
  - for the $X - Y$ link at $R - R_0$ with extra parities!

Bi-Layer LDPC Code for the Relay Channel



(n, k_1) must be capacity approaching for $X - Y_1$

Bi-Layer LDPC Code for the Relay Channel



(n, k_1) must be capacity approaching for $X - Y_1$

$(n, k_1 + k_2)$ is capacity approaching for $X - Y$

Bi-Layer LDPC Code

Code Design Problem for the Relay Channel

Design a single LDPC code so that:

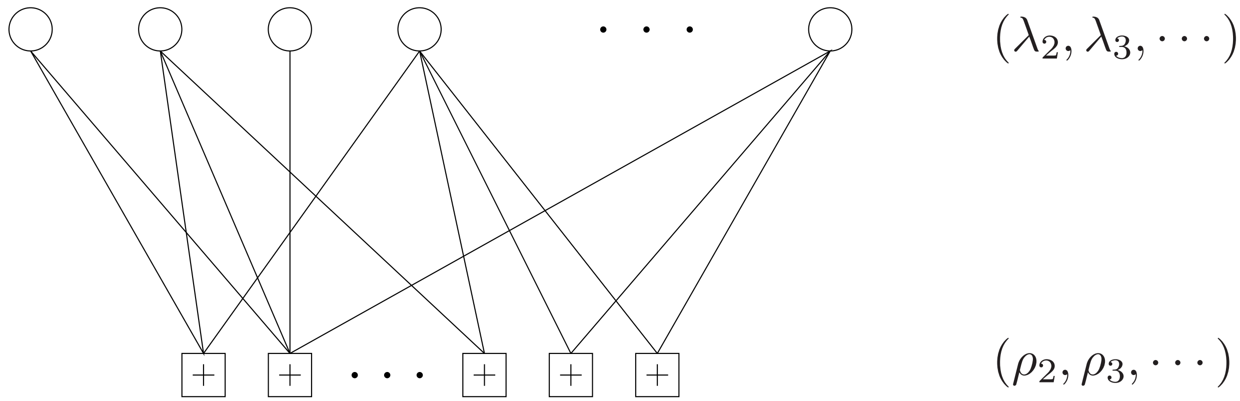
- The entire graph is capacity-achieving at $R - R_0$ with SNR_{low}.
- The sub-graph is capacity-achieving at R with SNR_{high}.

Does such code exist?

Yes. We can design LDPC degree sequence to achieve the above.

A problem of *universal codes*!

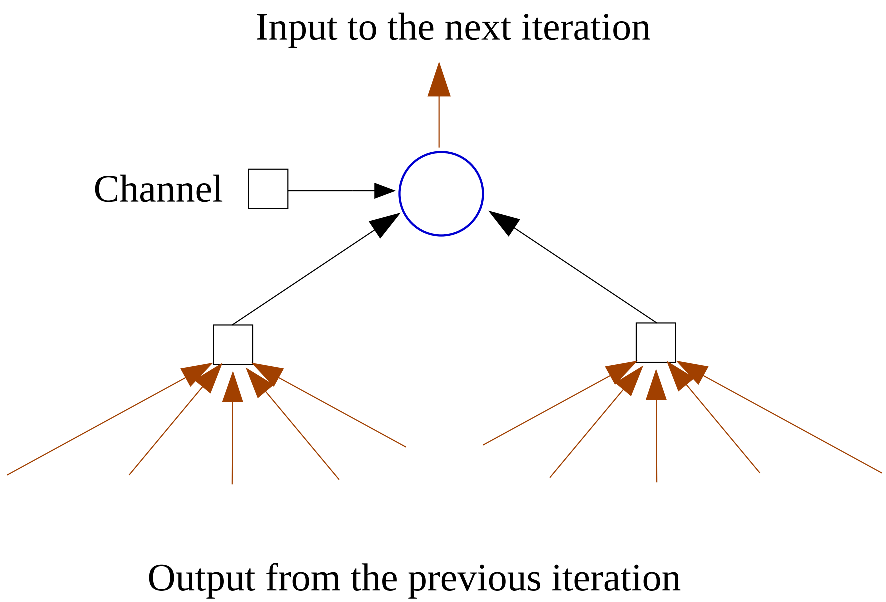
Irregular Low-Density Parity-Check Codes



- An ensemble of irregular LDPC codes is defined by its variable-degree distribution $\{\lambda_2, \lambda_3, \ldots\}$ and its check-degree distribution $\{\rho_2, \rho_3, \ldots\}$.

- Degree distribution is related to rate by: $$R = 1 - \frac{\sum_i \frac{\rho_i}{i}}{\sum_i \frac{\lambda_i}{i}}$$

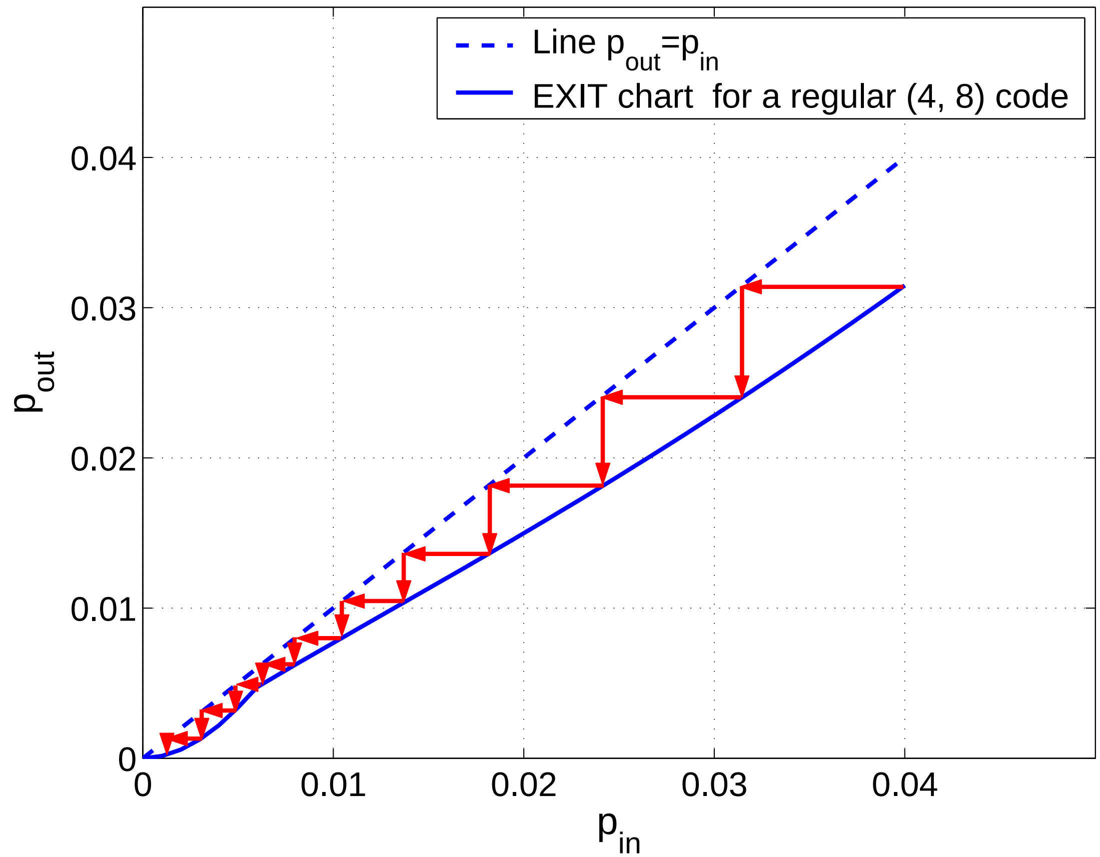
Iterative Decoding Algorithm



- A message is a belief about the incident variable node
- Decoder passes messages between check and variable nodes iteratively.

- Analysis Tool: Density Evolution (Urbanke-Richardson)
- This talk: Extrinsic Information Transfer (EXIT) charts (ten Brink)

Tracking Extrinsic Probability of Error



EXIT Chart for BSC with $\epsilon = 0.04$

- Mutual Inform. EXIT Chart (ten Brink '01)
- Prob. of Error EXIT Chart (Gallager '63, Ardakani '04)

Shaping the EXIT Chart

- For an irregular LDPC code, P_{out} at the output of variable nodes is computed using Bayes's rule.
- Assume a fixed check degree distribution, the resulting P_{out} is equivalent to a linear combination of corresponding P_{out} of regular codes
- Therefore, the EXIT chart of an irregular code is a linear combination of elementary EXIT charts of regular codes, making P_e-EXIT chart a powerful design tool.

$$\boxed{f(p) = \sum_i \lambda_i f_i(p)}$$

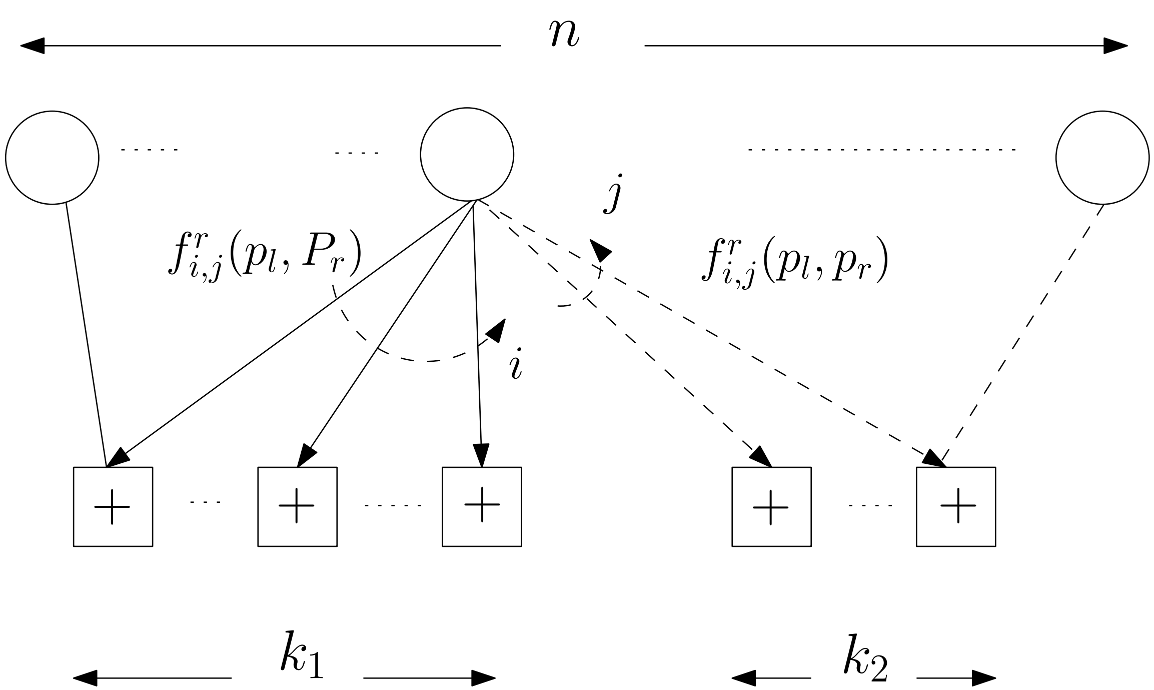
Linear Programming Approach to LDPC Code Design

- Consider a code design problem for a standard BSC or AWGN channel:
  - Fix check degree sequence ρ_i.

$$\begin{aligned} \text{maximize} \quad & 1 - \frac{\sum \rho_i / i}{\sum \lambda_i / i} \\ \text{subject to} \quad & \sum \lambda_i f_i(p) < p \end{aligned}$$

  - Choose variable degree sequence λ_i to maximize rate, subject to decodability constraints, by solving a *linear programming* problem.
- This talk: Generalizing this approach to design *bi-layer* codes.

Bi-Layer Density Evolution



Keep track of the probability of error in left and right graphs (p_l, p_r).
Define left and right elementary EXIT charts $f^l_{i,j}(p_l, p_r)$ and $f^r_{i,j}(p_l, p_r)$.

Designing Bi-Layer LDPC Codes for the Relay Channel

$$
\begin{aligned}
\text{maximize} \quad & 1 - \frac{\sum_i \rho_i / i}{\sum_i \nu_i / i} \\
\text{subject to} \quad & \nu_i = \frac{1}{\eta} \sum_j \frac{i}{i+j} \lambda_{i,j} \\
& \sum_i \nu_i f_i^s(p) < p \\
& \sum_{i,j} \lambda_{i,j} \frac{f_{i,j}^l(p_l, p_r) i + f_{i,j}^r(p_l, p_r) j}{i+j} < \eta p_l + (1-\eta) p_r
\end{aligned}
$$

Practical design: Fix one layer, optimize the second layer.

Performance

Example: Optimal $\lambda_{i,j}$ (left degree i and right degree j) for a relay channel with $R_{source-relay} = 0.7520$ and $R_{source-destination} = 0.6280$.

(i,j)	$j=0$	$j=1$	$j=2$	$j=3$
$i=2$	0.1153	0.0623	0	0
$i=3$	0.1220	0.0921	0	0
$i=5$	0	0.1897	0	0
$i=8$	0	0	0.0591	0
$i=9$	0	0	0.0166	0
$i=20$	0	0	0.3296	0.0132

Gap to capacity: **0.19dB** for source-relay, **0.34dB** for source-destination.

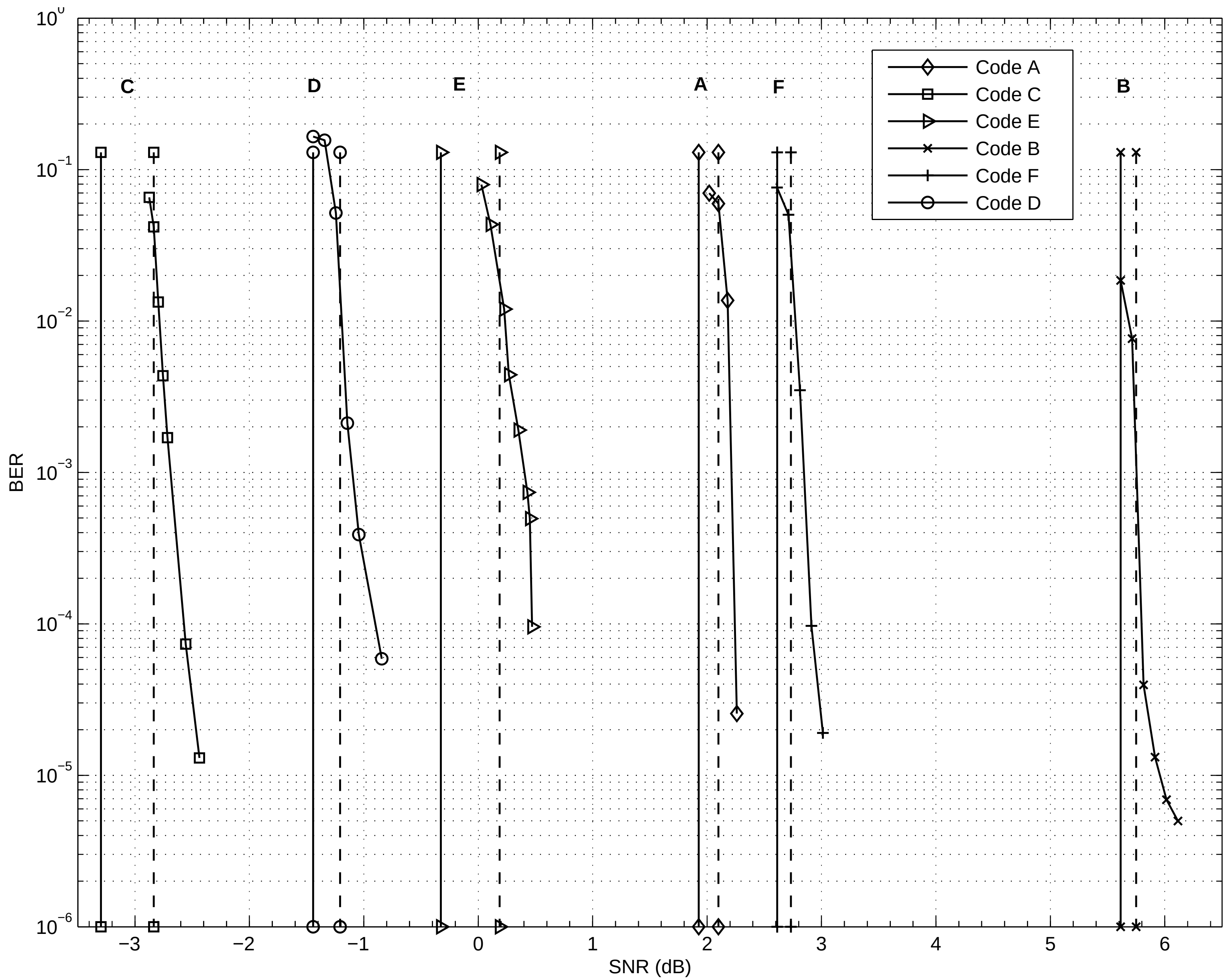


Fig. 11. Empirical bit error probability curves for the designed codes. Solid straight lines represent Shannon limits for each code, and dashed lines represent the convergence threshold computed by density evolution.

How Hard is Binning?

- Implementing binning:
  - Binning for quantization is hard. (e.g. Gel'fand-Pinsker, Wyner-Ziv)
  - Binning for error-correcting is practical! (e.g. DF in relay channel)

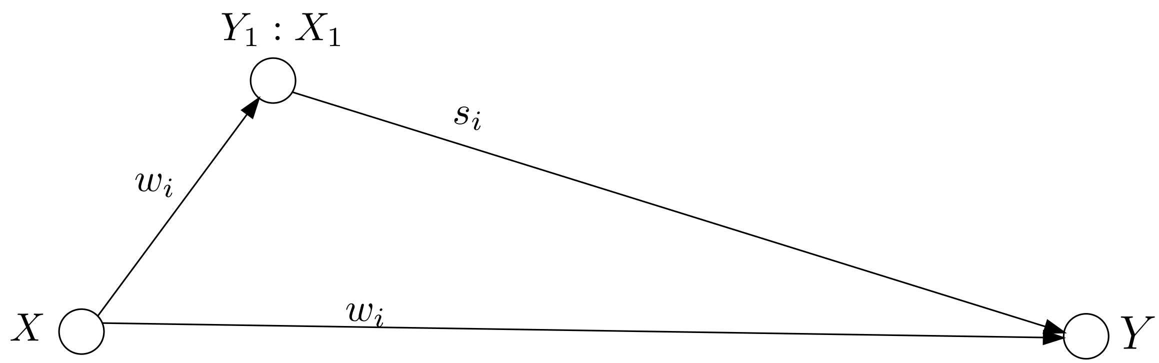
- Main message:

$$\boxed{\text{Binning for Relay Channel} = \text{Parity Forwarding}}$$

- The coding problem $\Rightarrow$ Designing a *universal* code.

Part II: Parity-Forwarding for Multi-Relay Networks

Parity Forwarding for One-Relay Network

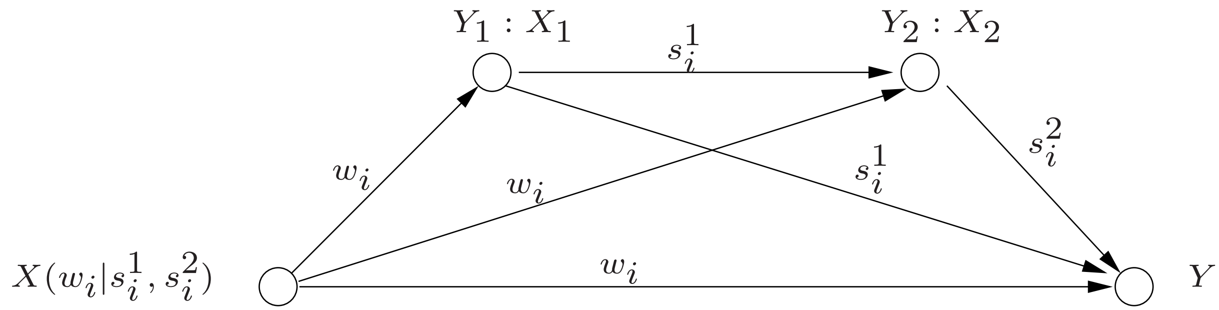


Key equations for Cover-El-Gamal strategy:

$$
\begin{aligned}
R &< I(X;Y_1|X_1) && \text{decodability at the relay} \\
R_0 &< I(X_1;Y) && \text{parity-forwarding from relay to destination} \\
R - R_0 &< I(X;Y|X_1) && \text{final decoding at the destination}
\end{aligned}
$$

"Degraded" means that relay is able to decode the source message.

Two-Relay Network

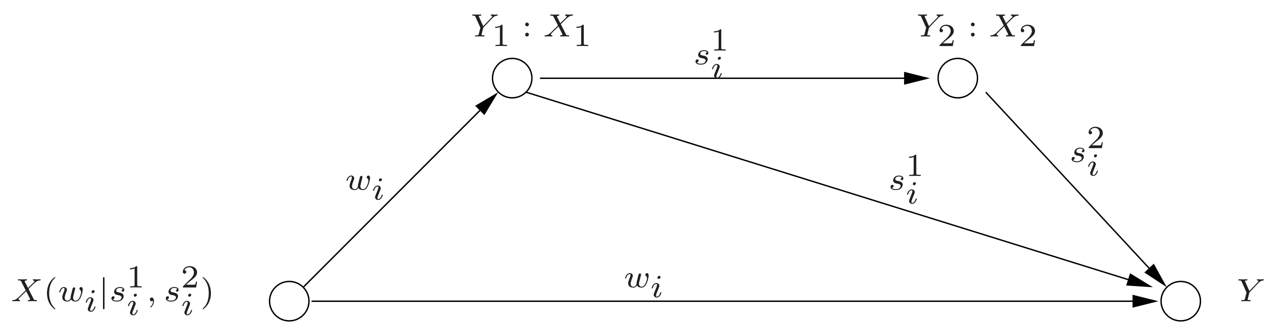


- What does degradedness mean for multi-relay networks?
  - Both relays are capable of decoding the source message. – Proof via regular encoding. (Xie-Kumar'05, Kramer-Gastpar-Gupta'05)

$$C = \max_{p(x,x_1,x_2)} \min\{I(X;Y_1|X_1,X_2), I(X,X_1;Y_2|X_2), I(X,X_1,X_2,;Y)\}$$

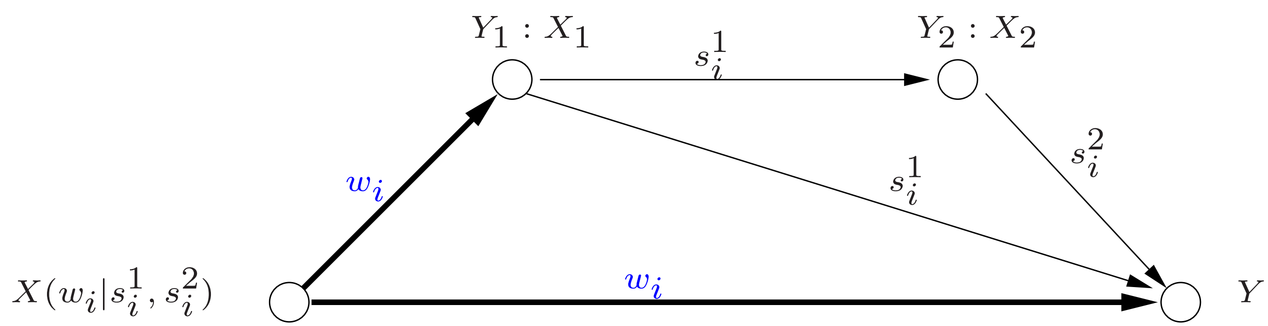
  - We call the above *serially degraded* relay channel.

Another Case: Doubly Degraded Two-Relay Network



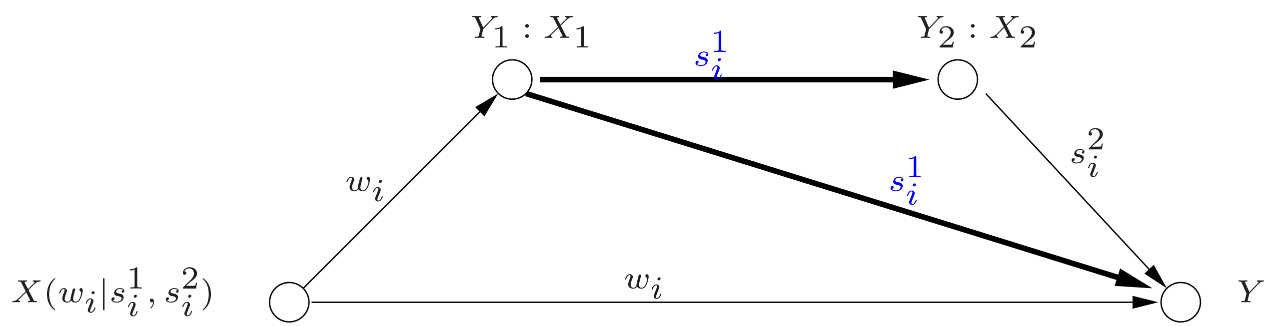
- Suppose that the link from source to the second relay is weak:
  - We do not require the second relay to decode the source message.
  - But, we use the second relay to help the first relay transmit the help-message to the destination.
- We call this a *doubly degraded* relay network.

Doubly Degraded Two-Relay Network



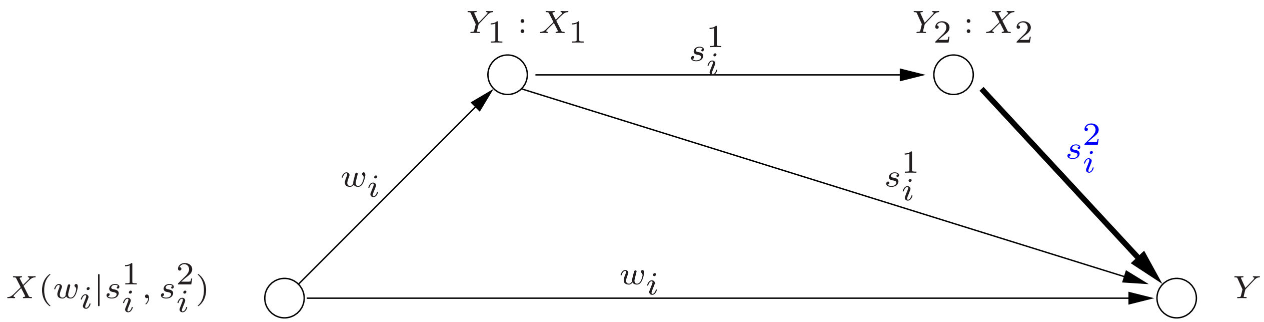
- Four-step block-Markov coding:
  - Source transmits w_i to both Y_1 and Y.
  - First relay decodes w_i and transmits s_i^1 (parities of w_{i-1}) to Y_2, Y.
  - Second relay decodes s_i^1 and transmits s_i^2 (parities of s_{i-1}^1) to Y.
  - Destination decodes s_i^2 first, then s_{i-1}^1, finally w_{i-2}.

Doubly Degraded Two-Relay Network



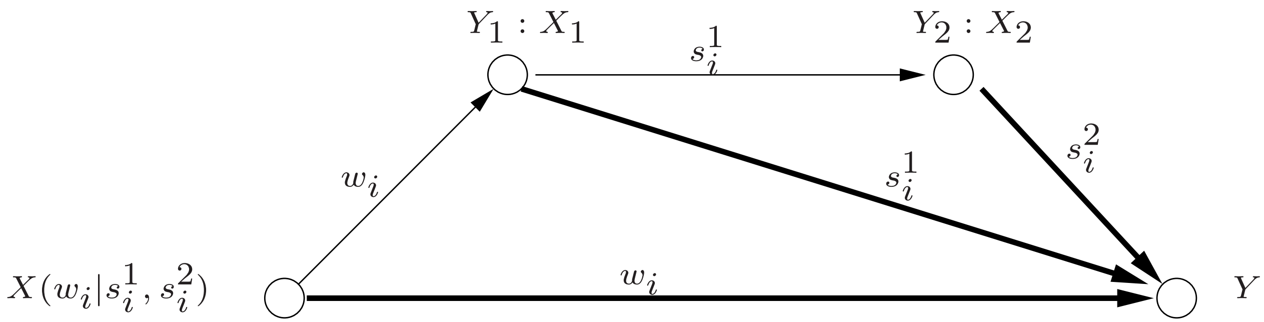
- Four-step block-Markov coding:
  - Source transmits w_i to both Y_1 and Y.
  - First relay decodes w_i and transmits s_i^1 (parities of w_{i-1}) to Y_2, Y.
  - Second relay decodes s_i^1 and transmits s_i^2 (parities of s_{i-1}^1) to Y.
  - Destination decodes s_i^2 first, then s_{i-1}^1, finally w_{i-2}.

Doubly Degraded Two-Relay Network



- Four-step block-Markov coding:
  - Source transmits w_i to both Y_1 and Y.
  - First relay decodes w_i and transmits s_i^1 (parities of w_{i-1}) to Y_2, Y.
  - Second relay decodes s_i^1 and transmits s_i^2 (parities of s_{i-1}^1) to Y.
  - Destination decodes s_i^2 first, then s_{i-1}^1, finally w_{i-2}.

Doubly Degraded Two-Relay Network



- Four-step block-Markov coding:
  - Source transmits w_i to both Y_1 and Y.
  - First relay decodes w_i and transmits s_i^1 (parities of w_{i-1}) to Y_2, Y.
  - Second relay decodes s_i^1 and transmits s_i^2 (parities of s_{i-1}^1) to Y.
  - Destination decodes s_i^2 first, then s_{i-1}^1, finally w_{i-2}.

Capacity for Doubly Degraded Two-Relay Network

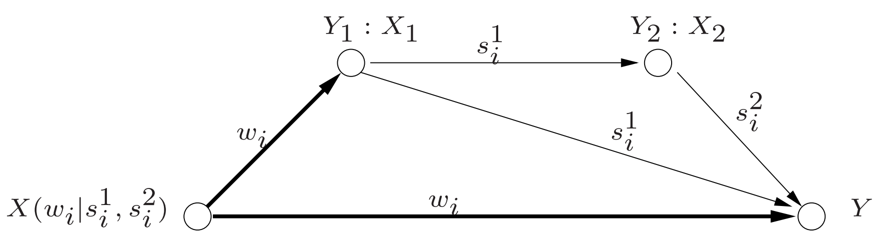
Definition 1. *A doubly degraded two-relay network is defined by* $p(y, y_1, y_2|x, x_1, x_2)$*, where* $X - (X_1, X_2, Y_1) - (Y_2, Y)$*,* $X_1 - (X_2, Y_2) - Y$ *and* $X - (X_1, X_2, Y) - Y_2$ *form Markov chains.*

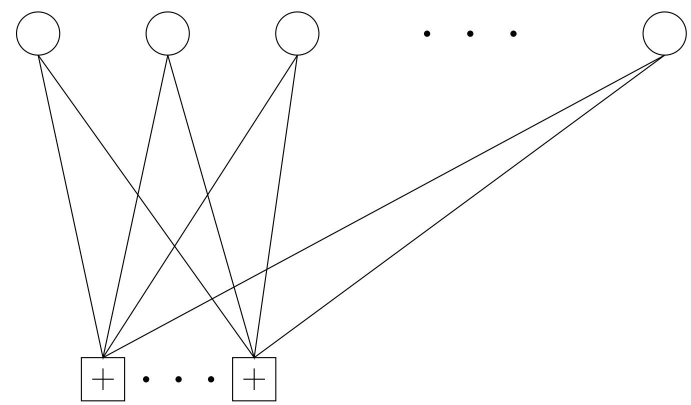
Theorem 1. *The following rate maximized over* $p(x, x_1, x_2)$ *is achievable*

$$\begin{aligned} R &< I(X; Y_1|X_1, X_2). \\ R &< I(X; Y|X_1, X_2) + I(X_1; Y_2|X_2) \\ R &< I(X; Y|X_1, X_2) + I(X_1; Y|X_2) + I(X_2; Y) \\ &= I(X, X_1, X_2; Y). \end{aligned}$$

It is also the capacity if the two-relay network is doubly degraded.

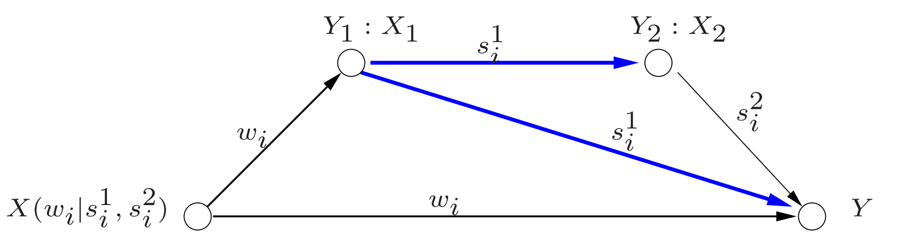
Coding for Doubly Degraded Relay Network

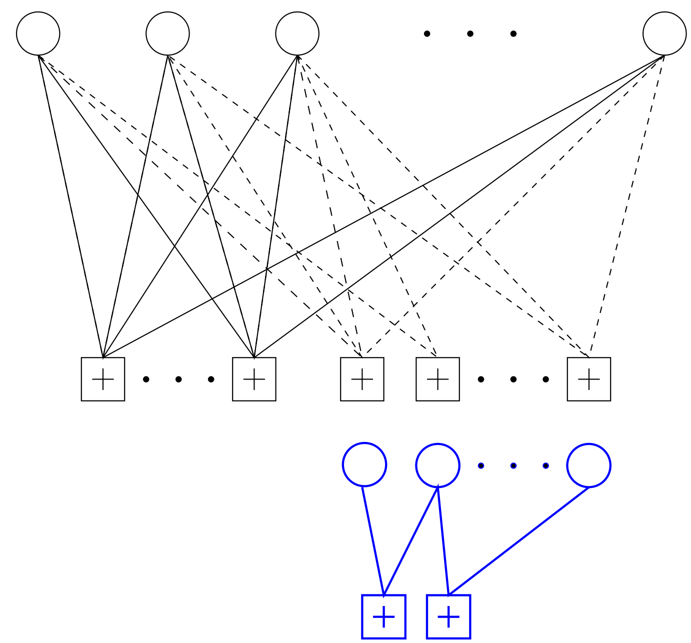




Cascade of bi-layer codes!

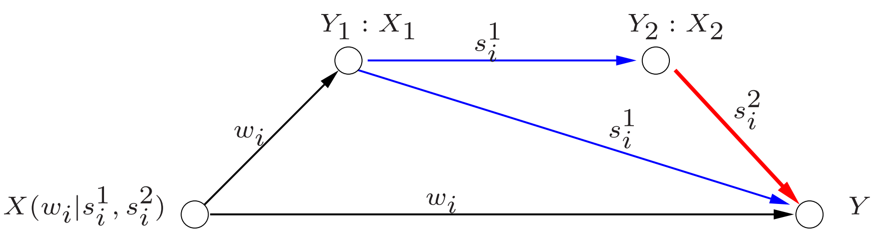
Coding for Doubly Degraded Relay Network

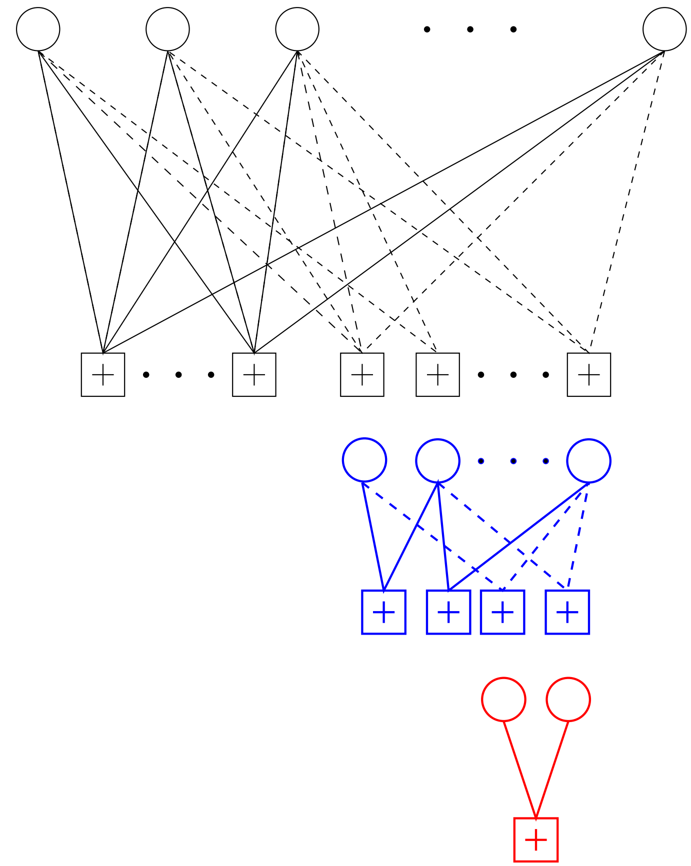




Cascade of bi-layer codes!

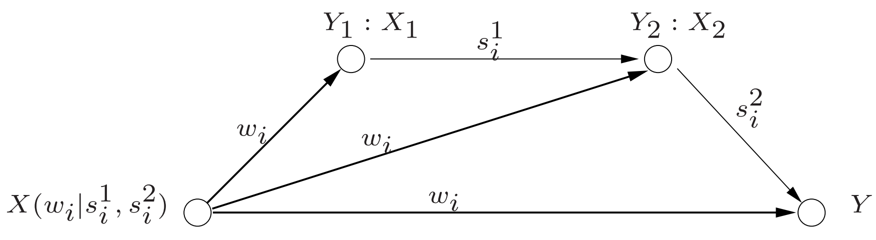
Coding for Doubly Degraded Relay Network

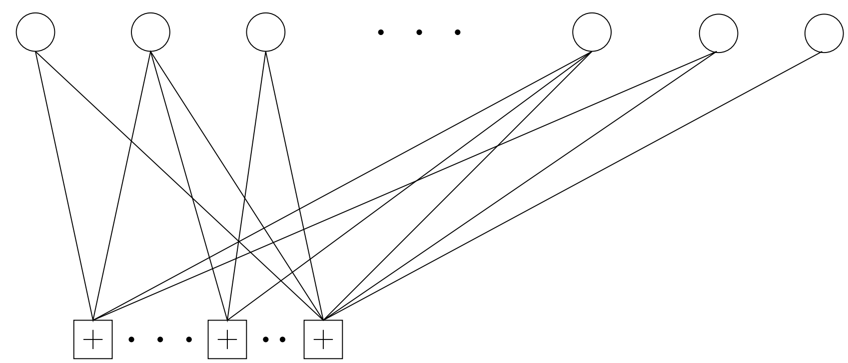




Cascade of bi-layer codes!

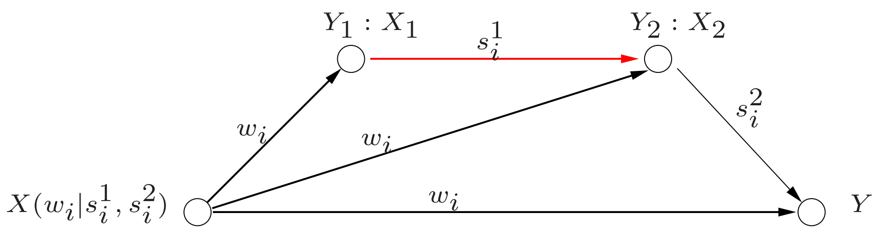
Another Case: Tri-Layer LDPC Codes

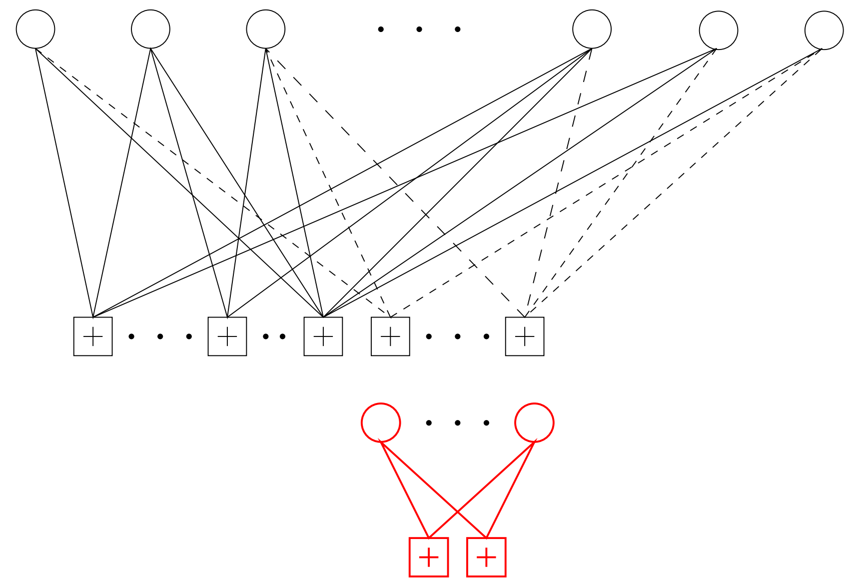




Embedded tri-layer code!

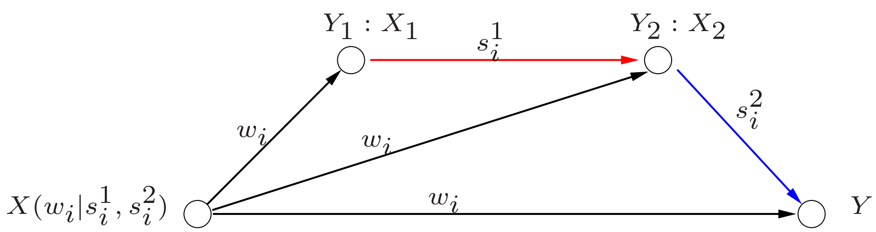
Another Case: Tri-Layer LDPC Codes

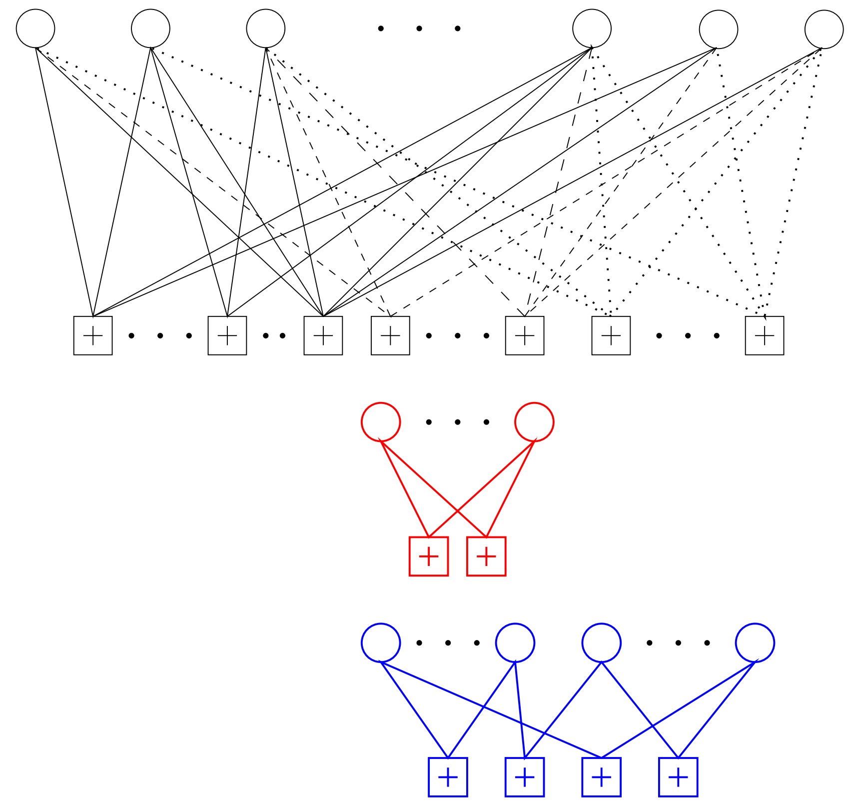




Embedded tri-layer code!

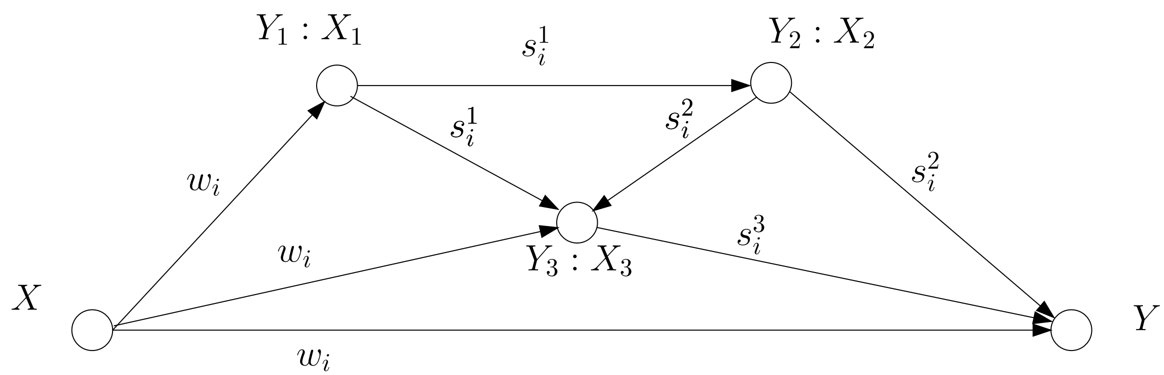
Another Case: Tri-Layer LDPC Codes





Embedded tri-layer code!

General Relay Networks



- The order at which different nodes help each other can be visualized:
  - Node X_1 helps Node Y_3 to decode w_i by sending s_i^1.
  - Node X_2 helps Node Y_3 to decode s_i^1 by sending s_i^2.
  - Node X_3 helps the destination in decoding both s_i^2 and w_i.
- This is like a routing protocol. Coding problem: *universal* codes!

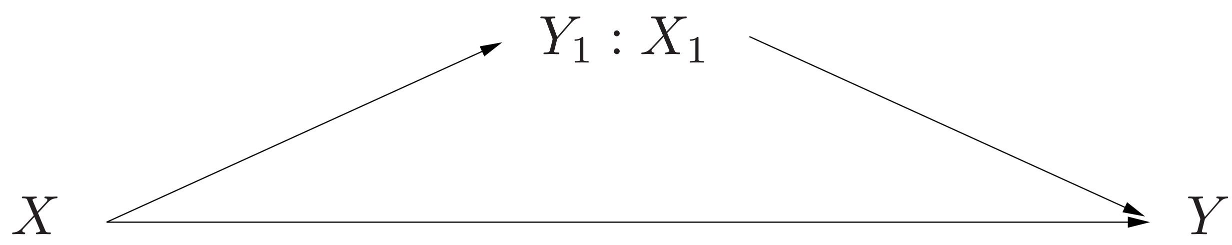
Connections with Fountain Codes and Network Coding

- Parity-generation achieves universal coding in an erasure network.
- Parity-formation achieves maximum single-source multicast throughput in network coding.
- Parity-forwarding achieves decode-and-forward rate in relay networks!

Part III: Quantize-and-Forward for a Modulo-Sum Relay Channel

Binning in Quantize-and-Forward

- In QF, neither the relay nor the destination can decode source message:

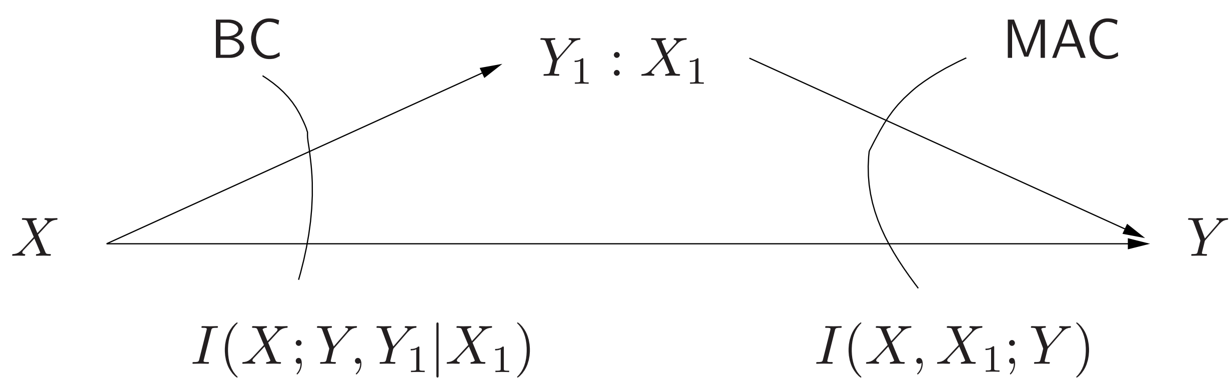


  - The relay summarizes its observation in U. Send a bin index of U.
  - The destination decodes U using Y as side information.
  - The destination then decodes X with the help of U.

$$C = \sup_{p(x)p(x_1)p(u|y_1,x_2)} I(X;YU|X_1), \quad \text{s.t. } I(X_1;Y) \geq I(Y_1;U|X_1,Y)$$

Is QF Ever Optimal?

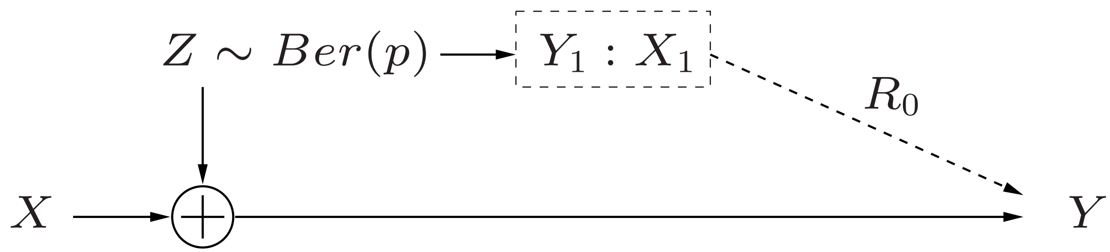
- Cut-set upper bound



$$C \le \max_{p(x,x_1)} \min\{I(X;Y,Y_1|X_2), I(X,X_1;Y)\}.$$

Is Cut-Set Bound Ever Tight for QF?

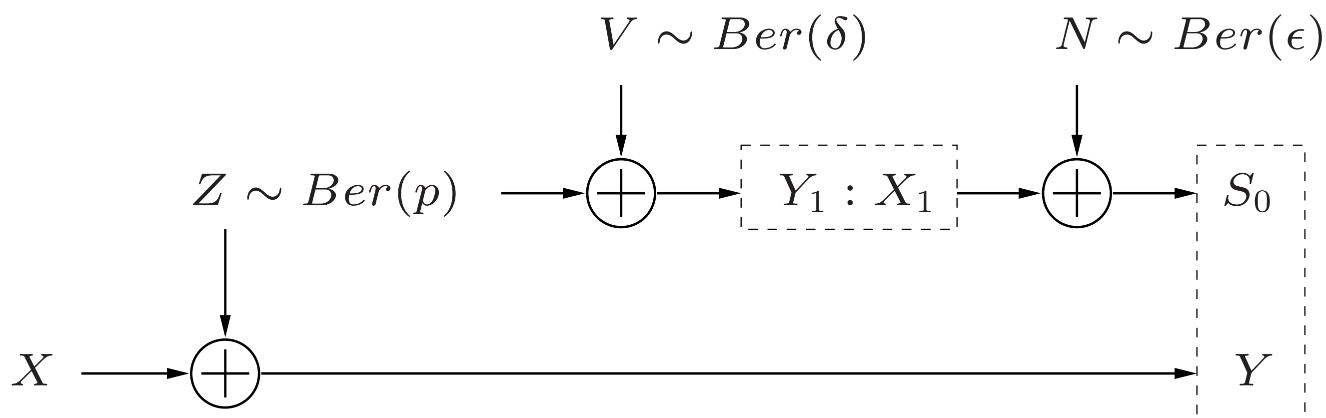
- Yes! QF achieves the cut-set bound in the deterministic channels studied by Cover and Kim '07. A simple example:



  - Quantize Z at rate R_0 minimizing Hamming distortion.
  - Destination adds quantized Z to channel.

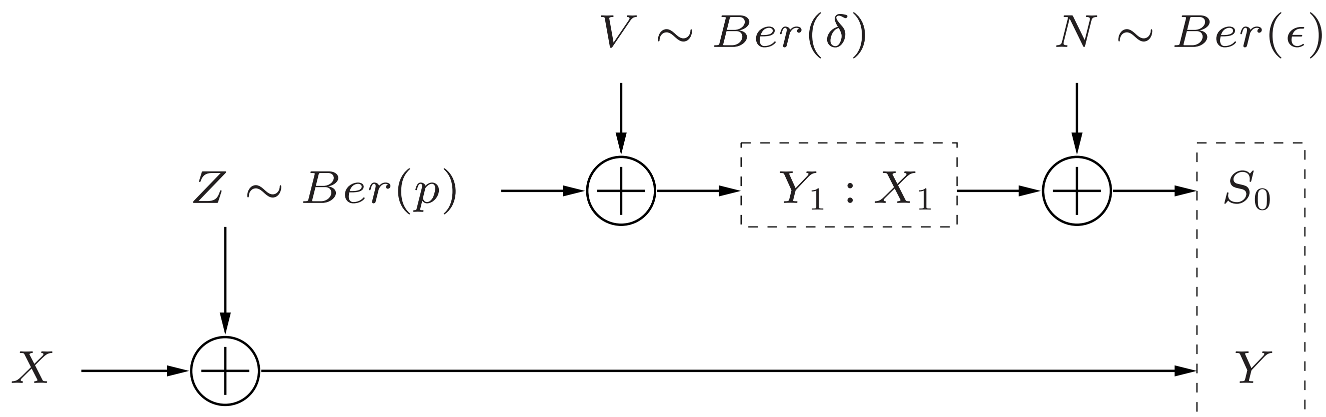
- This talk: What if the Relay observes a noisy Z?
  - QF still optimal, now at capacities *strictly below* the cut-set bound.

A Binary Relay Channel



- Relay observes a corrupted version of the noise.
  - DF is useless.
  - Forward Y_1 uncoded?
  - QF? ... but maximizing $H(X)$ limits the side information.

A Binary Relay Channel: Capacity



Theorem 2. *The capacity of the above channel is:*

$$C = \max_{p(u|y_1):I(U;Y_1)\leq R_0} 1 - H(Z|U)$$

where max is over U*'s with* $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$*, and* $R_0 = \max_{p(x_1)} I(X_1; S_0)$*.*

Capacity: Achievability

Follows from the QF strategy of Cover and El Gamal:

- Codebook Generation:
  - Set $p(x) = Ber(\frac{1}{2})$.
  - Fix $p(u|y_1)$ such that $I(U; Y_1) \leq R_0$.
  - Generate conventional rate-distortion codebook U at Y_1.
- Encoding:
  - Block-Markov coding
  - Relay quantizes Y_1^n with U^n, and sends quantization index.
- Achievable rate: $R < I(X; YU) = I(X; Y|U) = 1 - H(Z|U)$

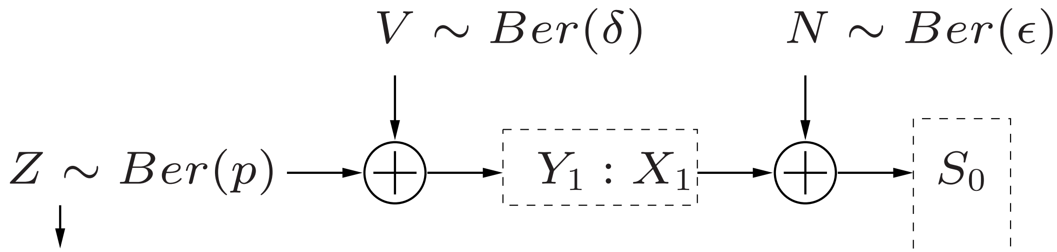
Capacity: Converse

Starting with Fano's inequality:

$$
\begin{aligned}
nR = H(W) &= I(W;Y^n,S_0^n) + H(W|Y^n,S_0^n) \\
&\leq I(W;Y^n,S_0^n) + n\epsilon_n \\
&\leq I(X^n;Y^n,S_0^n) + n\epsilon_n \\
&= I(X^n;Y^n|S_0^n) + n\epsilon_n \\
&= H(Y^n|S_0^n) - H(Y^n|S_0^n,X^n) + n\epsilon_n \\
&\leq n - H(Z^n|S_0^n,X^n) + n\epsilon_n \\
&= n - H(Z^n|S_0^n) + n\epsilon_n
\end{aligned}
$$

How to proceed? Need to modify a result from rate distortion.

Wyner's Lemma



Lemma 1. *The following holds for any encoding scheme at the relay*

$$H(Z^n|S_0^n) \geq \min_{p(u|y_1):I(U;Y_1)\leq R_0} nH(Z|U)$$

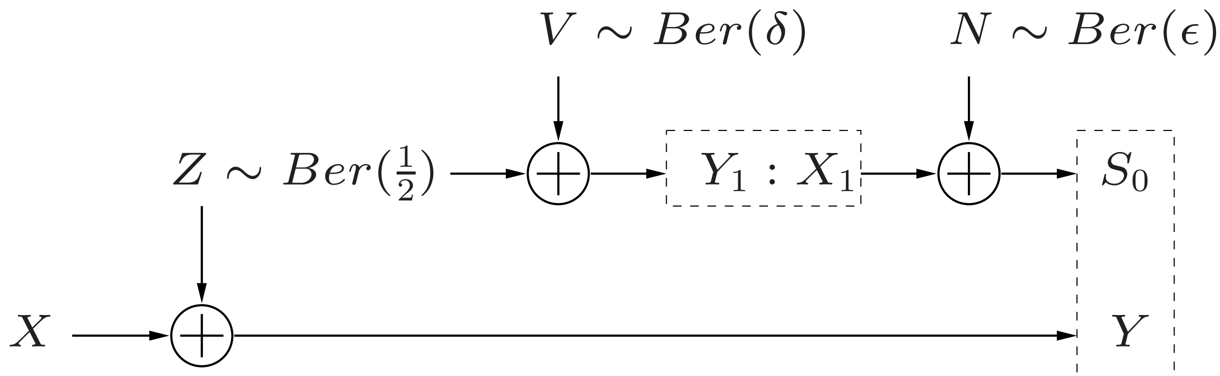
.

Proof: Expand $H(Z^n|S_0^n) \geq \sum_{i=1}^n H(Z_i|S_0^n, Y_1^{i-1})$. Let $U_i = (S_0^n, Y_1^{i-1})$.

Back to the Converse

$$
\begin{aligned}
nR = H(W) &= I(W;Y^n,S_0^n) + H(W|Y^n,S_0^n) \\
&\leq I(W;Y^n,S_0^n) + n\epsilon_n \\
&\leq I(X^n;Y^n,S_0^n) + n\epsilon_n \\
&= I(X^n;Y^n|S_0^n) + n\epsilon_n \\
&= H(Y^n|S_0^n) - H(Y^n|S_0^n,X^n) + n\epsilon_n \\
&\leq n - H(Z^n|S_0^n,X^n) + n\epsilon_n \\
&= n - H(Z^n|S_0^n) + n\epsilon_n \\
&\leq \max_{p(u|y_1):I(U;Y_1)\leq R_0} n(1 - H(Z|U)) + n\epsilon_n
\end{aligned}
$$

Computing the Capacity

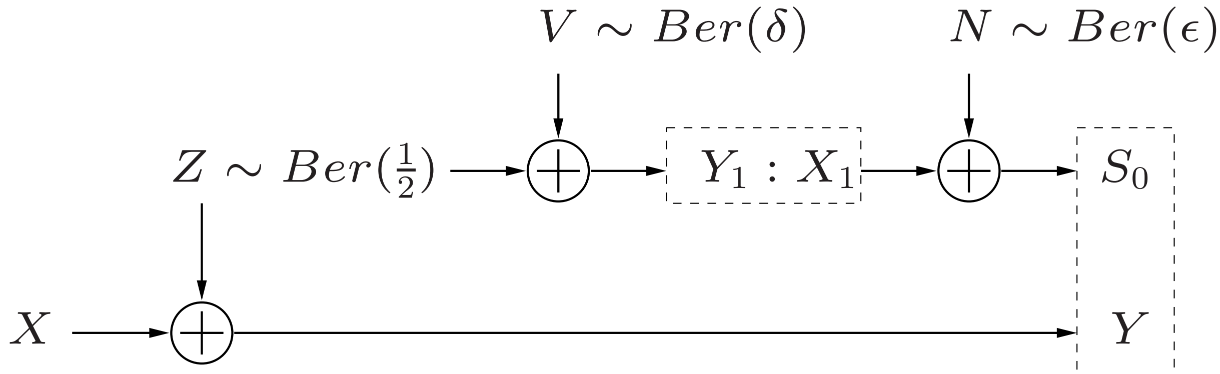


- Need to evaluate

$$C = \max_{p(u|y_1):I(U;Y_1)\le R_0} 1 - H(Z|U).$$

- In general hard to do, but possible for a special case: $Z \sim Ber(\frac{1}{2})$.

Computing the Capacity



- When $Z \sim Ber(\frac{1}{2})$, $Z = V + Y_1$ with V, Y_1 independent.
  - Rewrite constraint $I(U;Y_1) \leq R_0$ as $H(Y_1|U) \geq 1 - R_0$.
  - Now the goal is to

$$\min_{p(u|y_1):H(Y_1|U)\geq 1-R_0} H(Z|U)$$

Entropy Power Inequality for Binary Random Variable

Lemma 2. [Wyner and Ziv] *Suppose* $Z = Y_1 + V$*, and* $V \sim Ber(\delta)$*. If*

$$H(Y_1|U) \geq \alpha,$$

then

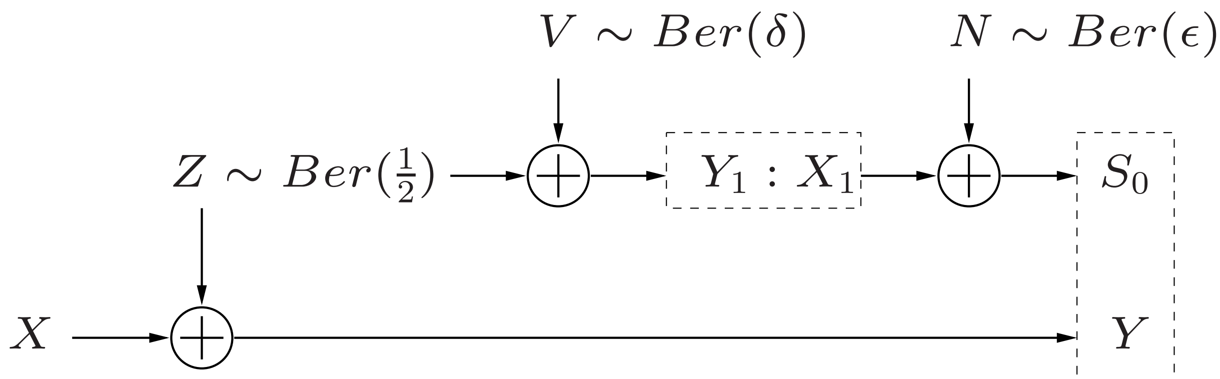
$$H(Z|U) \geq h(h^{-1}(\alpha) * \delta),$$

with equality iff Y_1 *given* U *is a* $Ber(h^{-1}(\alpha))$ *random variable.*

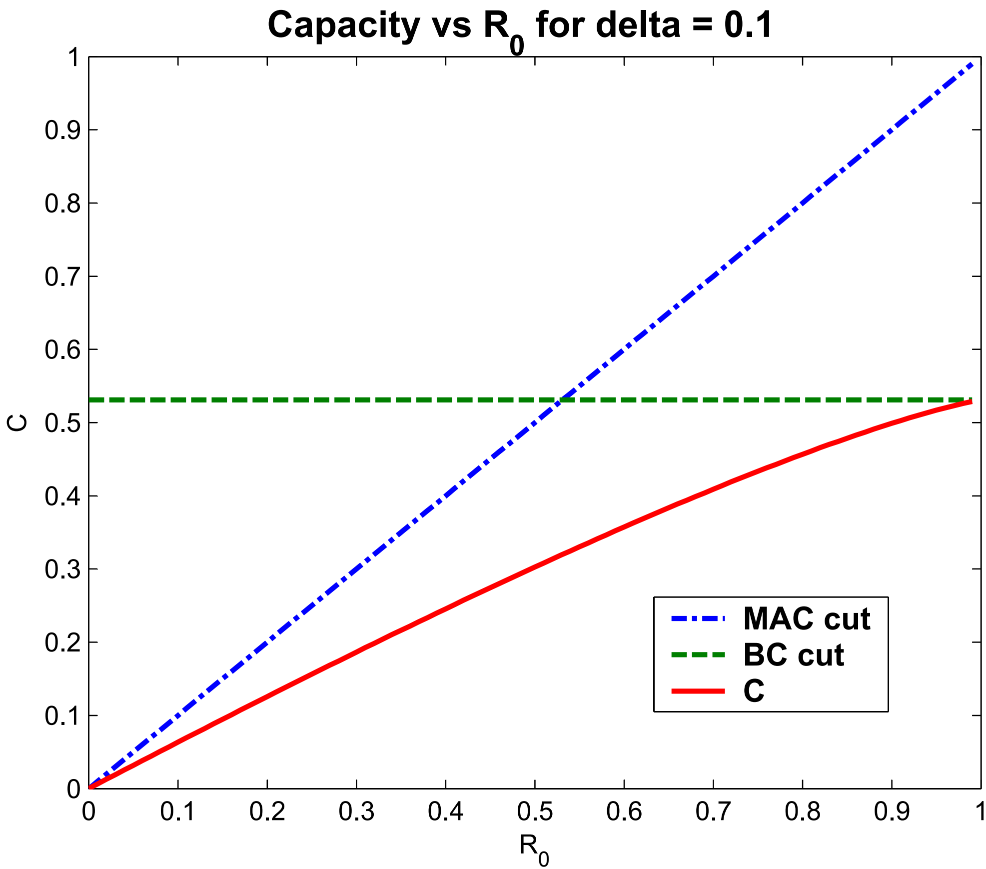
- But this is achievable with standard binary quantization:
  - Let U quantize Y_1 at rate R_0 minimizing Hamming distance D.
  - $H(Y_1|U)$ is $Ber(D)$ and $H(D) = H(Y_1) - R_0 = 1 - R_0$.

$$\boxed{C = 1 - h(h^{-1}(1 - R_0) * \delta)}$$

Capacity is Below Cut-Set Bound

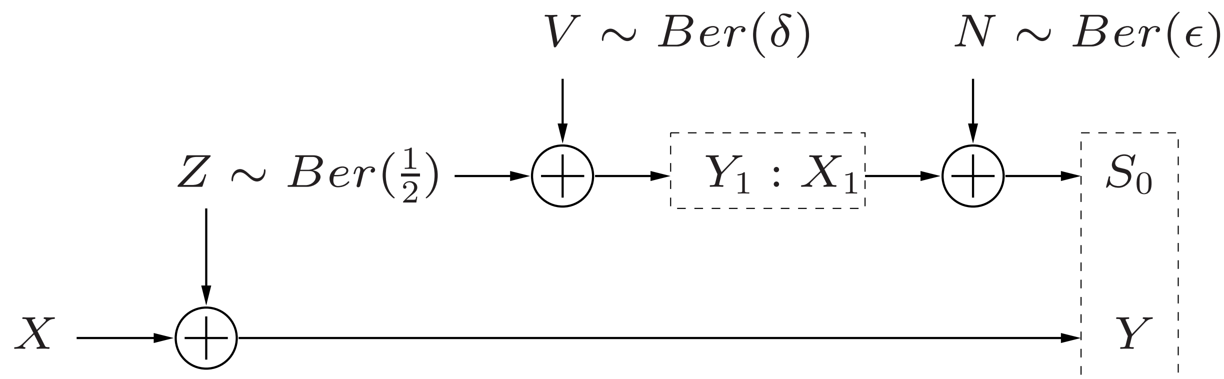


- MAC Cut: Direct channel has zero capacity.
  - $C_{MAC} = \max_{p(x_1)} I(X_1; S_0) = R_0$.
- BC Cut: Receiver adds $Y_1 = Z + V$ and $Y = X + Z$ together.
  - $C_{BC} = 1 - h(\delta)$.



Capacity vs R_0 for delta = 0.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
C
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
R_0
MAC cut
BC cut
C

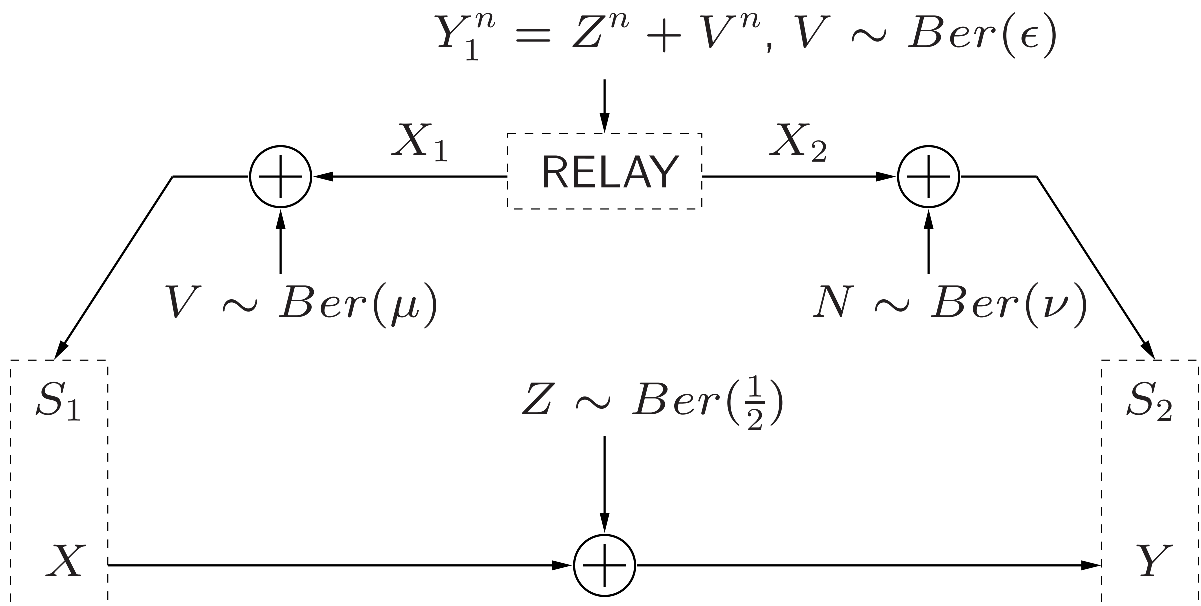
Quantize and Add



$$C = 1 - h(h^{-1}(1 - R_0) * \delta)$$

- No Wyner-Ziv and only addition at the receiver.
  - Relay quantizes Y_1 with U. Sends U to receiver.
  - Receiver does: $Y = X + Z + U = X + Y_1 + V + U = X + Z' + V$.

Connection with Channels with Side Information



Let $R_1 = \max_{p(x_1)} I(X_1; S_1)$ and $R_2 = \max_{p(x_2)} I(X_1; S_2)$.

Theorem 3. $C = 1 - h(h^{-1}(1 - R_1 - R_2)) * \delta)$

Conclusions and Summary

- Both decode-and-forward and quantize-and-forward can be interpreted as parity-forwarding strategies.
- Parity-forwarding can be efficiently implemented using LDPC codes.
- Multi-relay networks can be degraded in more than one way; parity-forwarding is capacity-achieving in degraded networks.
- Quantize-and-forward can be optimal if relay only observes noise.
- Cut-set bound is not tight in general.

For Further Details

- Peyman Razaghi and Wei Yu: "Bilayer Low-Density Parity-Check Codes for Decode-and-Forward in Relay Channels", to appear in *IEEE Transactions on Information Theory*, October 2007.

- Peyman Razaghi and Wei Yu, "Parity forwarding for multiple-relay networks", submitted to *IEEE Transactions on Information Theory*, November 2007.

- Marko Aleksic, Peyman Razaghi, and Wei Yu: "Capacity of a Class of Modulo-Sum Relay Channels", submitted to *IEEE Transactions on Information Theory*, June 2007.