New Results in Coding and Communication Theory

Communication Systems Laboratory
http://www.ee.ucla.edu/~csl/

Rick Wesel

At the Aerospace Corporation
August 11, 2008

Miguel Griot

Bike Xie

Yuan-Mao Chang

Tom Courtade

Jiadong Wang

Recent Results

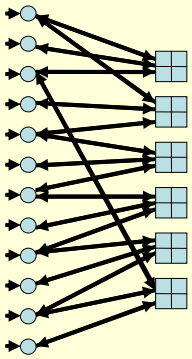
- A more efficient LDPC decoding algorithm
- Dynamically scheduled LDPC decoder that converge to a better frame error rate than previous state-of-the-art.
- A powerful family of low-rate turbo codes for “rate-less” applications at the physical layer.
- Turbo Code with closest-to-Shannon performance through novel nonlinear labeling.
- New information theoretic results and nonlinear turbo codes for multi-rate broadcast.

Recent Results

- A more efficient LDPC decoding algorithm
- Dynamically scheduled LDPC decoder that converge to a better frame error rate than previous state-of-the-art.
- A powerful family of low-rate turbo codes for “rate-less” applications at the physical layer.
- Turbo Code with closest-to-Shannon performance through novel nonlinear labeling.
- New information theoretic results and nonlinear turbo codes for multi-rate broadcast.

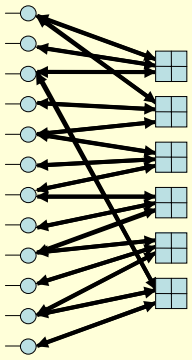
Simultaneous (Flooding) Schedule

- On every iteration
  - All variable nodes are simultaneously updated
  - All check nodes are simultaneously updated

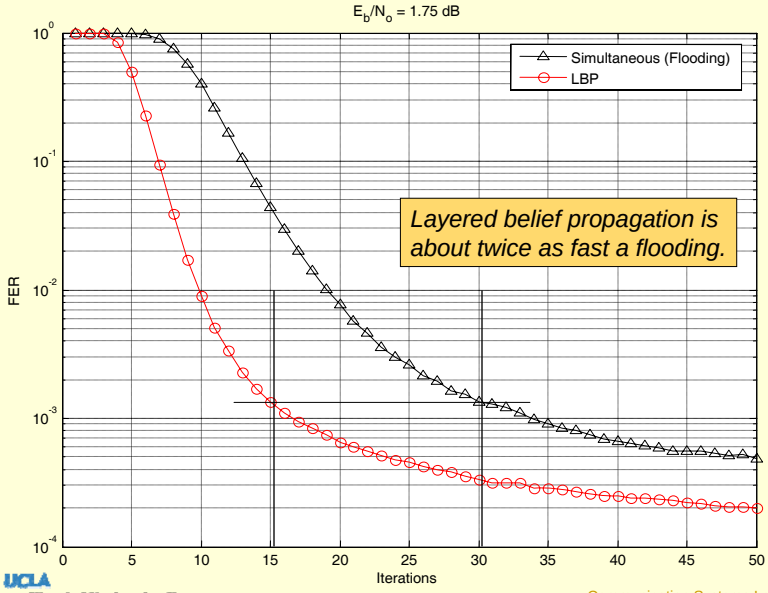


Standard Sequential Schedule (SSS) a.k.a Layered Belief Propagation (LBP)

- Update messages sequentially:
  - [Mansour 03] (CN sequence)
  - [Kfir 03] (VN sequence)
  - [Hocevar 04] (LBP)
  - [Sharon 04] (Serial Schedule)
  - [Zhang 05] (Shuffled BP)
  - [Radosavljevic 05]



Results N=1944 Rate=1/2



Layered belief propagation is about twice as fast a flooding.

Informed Dynamic Scheduling

- Sequential scheduling is a good idea
- What is the best sequence of updates?
- Use the current information in the graph to choose the next message to be sent
- This is called Informed Dynamic Scheduling (IDS)

Residual Belief Propagation (RBP) [Elidan 06]

- Define residual as:

$$r = \left\| m_{new} - m_{old} \right\|$$

- As BP converges, the residuals go to 0
- RBP is a greedy algorithm that propagates the message with the largest residual

RBP for LDPC decoding

- Messages are Log-Likelihood Ratios (LLRs)

$$\log\left(\frac{p(v_j=0)}{p(v_j=1)}\right)$$

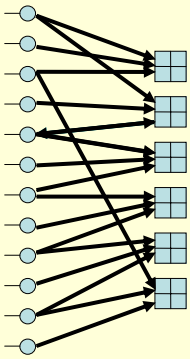
- Message-generation equations:

$$m_{v_j\to c_i} = \sum_{a\in N(v_j)\setminus c_i} m_{c_a\to v_j} + C_{v_j}$$

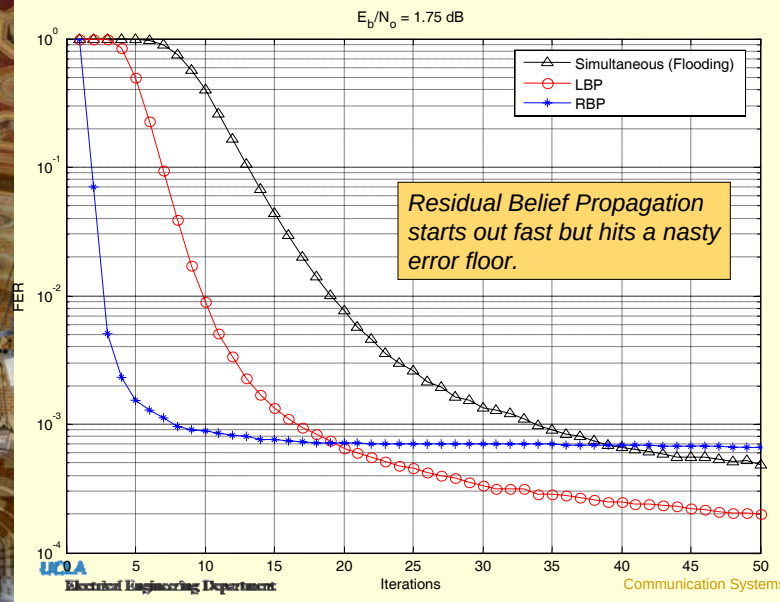
$$m_{c_i\to v_j} = 2\times \operatorname{atanh}\left(\prod_{v_b\in N(c_i)\setminus v_j} \tanh\left(\frac{m_{v_b\to c_i}}{2}\right)\right)$$

RBP for LDPC decoding

- Initially propagate the channel information
- Propagate message with biggest residual
- The variable-to-check messages that change will have the same residual so they are the biggest
- Therefore, RBP can be simplified
  - Propagate check-to-variable message with biggest residual
  - Propagate variable-to-check messages that change



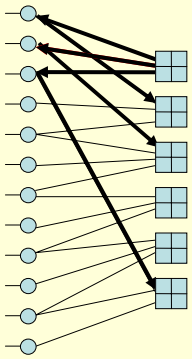
Results N=1944 Rate=1/2



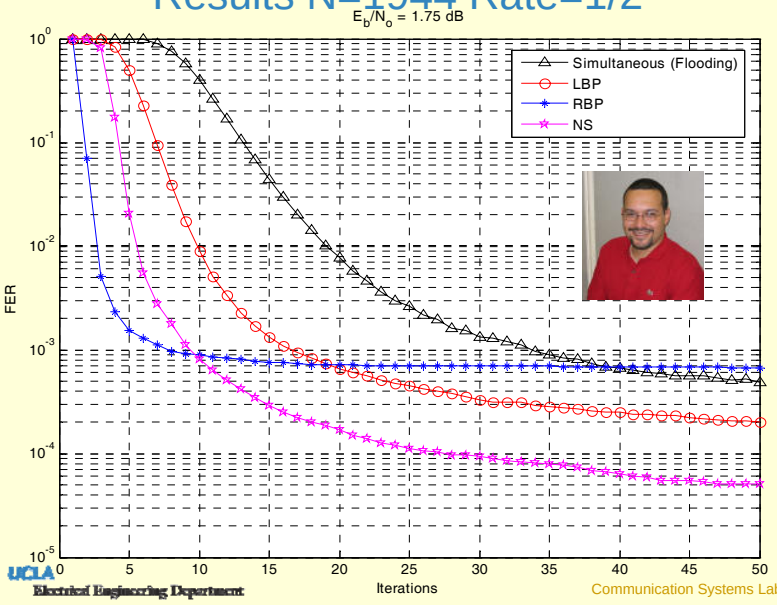
Residual Belief Propagation starts out fast but hits a nasty error floor.

Dynamic Node-wise Scheduling (NS)

- Find the check-to-variable message with the biggest residual.
- Update the check-node.
- Update the variable-to-check messages that change.



A better answer in fewer iterations: Results N=1944 Rate=1/2



UCLA Electrical Engineering Department

Communication Systems Laboratory

We understand faster, but why better?

- Performance plots show that informed dynamic scheduling strategies perform not only faster but better than LBP.
- There are several noise realizations that aren't corrected after 200 LBP iterations but are corrected after few ANS iterations.
- This difference can't be explained with the argument of the "wasted" updates.
- Informed Dynamic Scheduling solves "trapping-set" errors that neither LBP nor flooding scheduling can solve.

UCLA Electrical Engineering Department

Communication Systems Laboratory

Node-wise Scheduling (NS)

- RBP greediness makes it converge faster but less often
- Node-wise Scheduling (NS) simultaneously propagates all the outgoing messages from a check node
- The check-node updated is the one that has the message with the biggest residual
- This means that correcting messages are simultaneously sent to all variable nodes that could be in error
- NS converges slower than RBP but more often

IDS "Iterations"

- There aren't iterations in IDS strategies
- In order to compare the scheduling strategies we count an iteration after the number of check-to-variable messages propagated is the same as in a flooding or LBP iteration
- The message generation complexity of all the strategies is the same
- Residual computation makes the IDS strategies more complex than flooding and LBP

Complexity

- Residual computation requires the values of the messages to be propagated.
- Many of those computations are then "wasted" since many of those messages aren't propagated.
- Using the min-sum check-node update to compute residuals significantly reduces the complexity.
- Approximate RBP (ARBP) and Approximate NS (ANS) are the min-sum versions of RBP and NS respectively.

Min-sum

– Check-to-variable message equation:

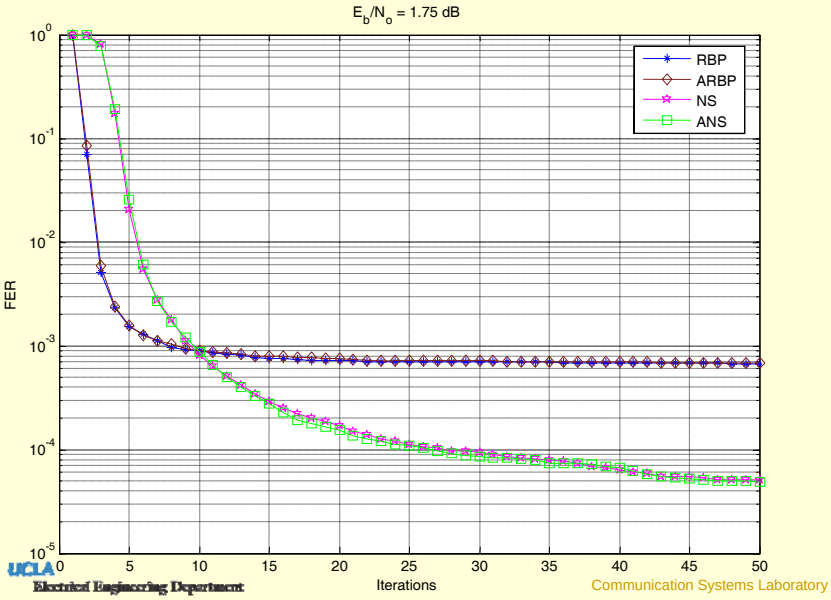
$$m_{c_i \to v_j} = \prod_{v_b \in N(c_i) \setminus v_j} \operatorname{sgn}\left(m_{v_b \to c_i}\right) \cdot \min_{v_b \in N(c_i) \setminus v_j} \left(\left|m_{v_b \to c_i}\right|\right)$$

–

– Thus a check-node update involves:

- Find the two variable-to-check messages with the smallest reliabilities
- Compute the signs of the check-to-variable messages
- Assign the correct reliability

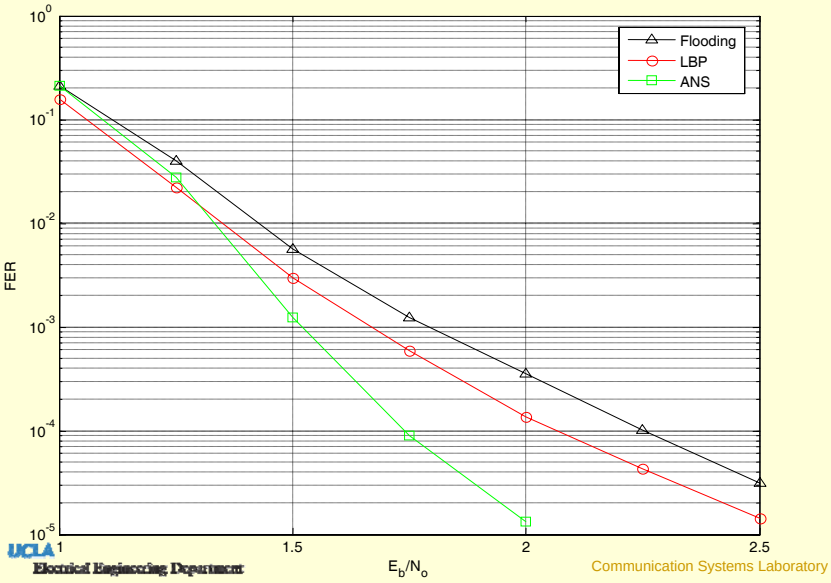
Results N=1944 Rate=1/2



UCLA Electrical Engineering Department

Communication Systems Laboratory

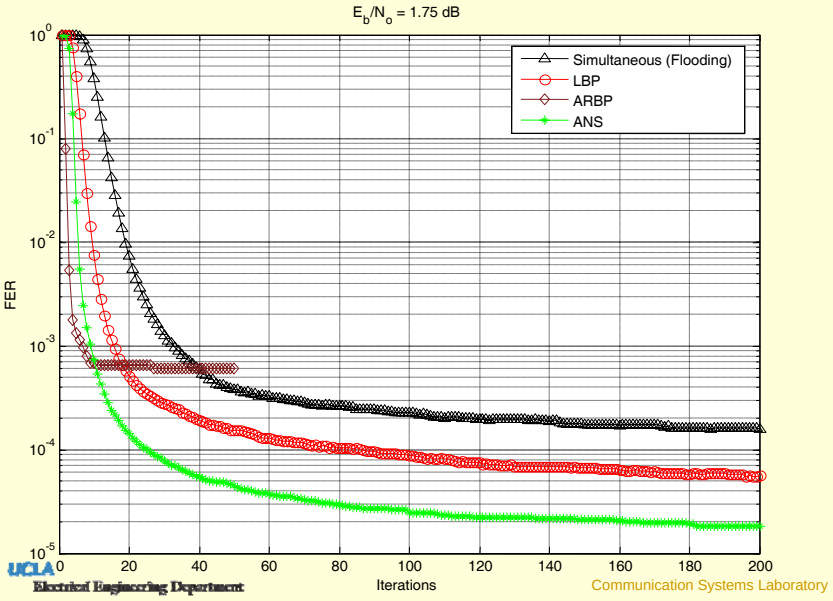
Results N=1944 Rate=1/2



UCLA Electrical Engineering Department

Communication Systems Laboratory

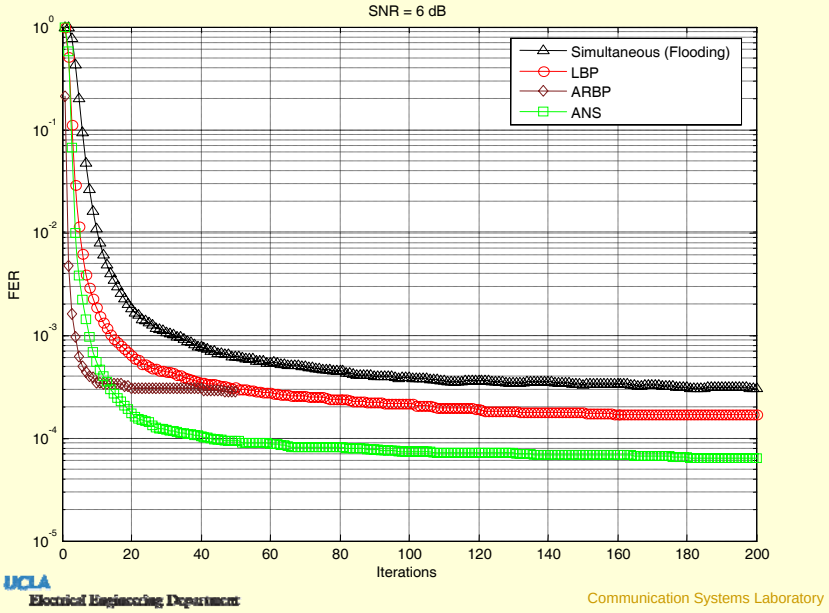
Results 802.11n code rate 1/2



UCLA Electrical Engineering Department

Communication Systems Laboratory

Results 802.11n code rate 5/6



UCLA Electrical Engineering Department

Communication Systems Laboratory

Recent Results

- A more efficient LDPC decoding algorithm
- Dynamically scheduled LDPC decoder that converge to a better frame error rate than previous state-of-the-art.
- A powerful family of low-rate turbo codes for “rate-less” applications at the physical layer.
- Turbo Code with closest-to-Shannon performance through novel nonlinear labeling.
- New information theoretic results and nonlinear turbo codes for multi-rate broadcast.

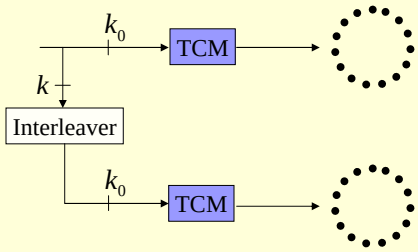
Recent Results

- A more efficient LDPC decoding algorithm
- Dynamically scheduled LDPC decoder that converge to a better frame error rate than previous state-of-the-art.
- A powerful family of low-rate turbo codes for “rate-less” applications at the physical layer.
- Turbo Code with closest-to-Shannon performance through novel nonlinear labeling.
- New information theoretic results and nonlinear turbo codes for multi-rate broadcast.

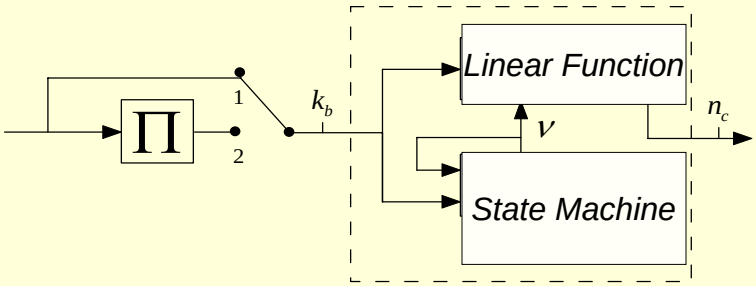
UCLA Electrical Engineering Department
Communication Systems Laboratory

Nonlinear Turbo Codes

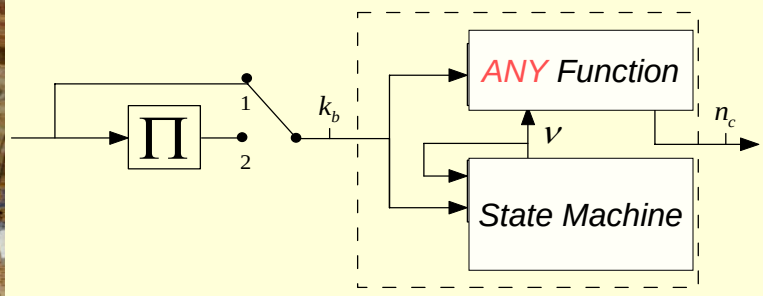
- All existing turbo codes use linear convolutional component codes.
- Miguel Griot developed new tools that allow us to design and analyze nonlinear component trellis codes.



Traditional Linear Turbo Codes

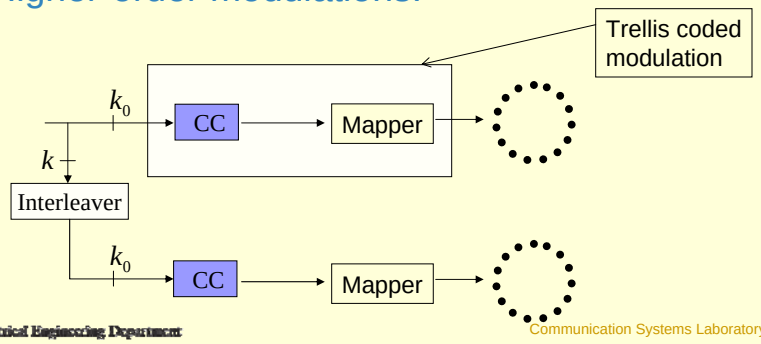


Nonlinear Turbo Codes



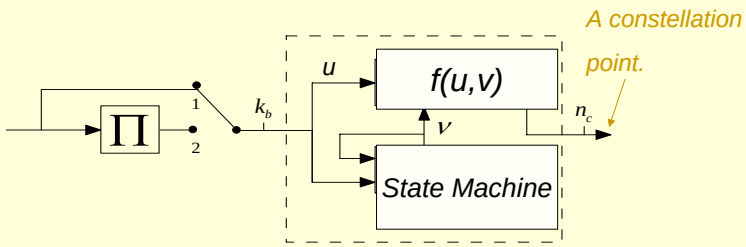
Motivations

- Applications where a non-uniform distribution of ones and zeros are required for maximum transmission rate. (Linear codes provide equally likely ones and zeros).
- Higher-order modulations:



Higher-order modulations

- The serial concatenation of a convolutional code with a mapping could be too restrictive.
- We can use any function that assigns constellation points to each of the branches of the trellis.



Analytical Bit Error Rate bound

- We provide a method to predict the BER of parallel concatenated non-linear codes over asymmetric channels, in particular the Z-Channel, under Maximum Likelihood decoding.
- We extend the uniform interleaver analysis proposed in [Benedetto '96].
- Uniform interleaver: given the two constituent codes, average over all possible interleavers of a certain length *K.*
- Key difference: **nonlinearity** of the constituent codes. We cannot assume that the all-zero codeword is transmitted. We need to average over all possible codewords.

A new uniform interleaver

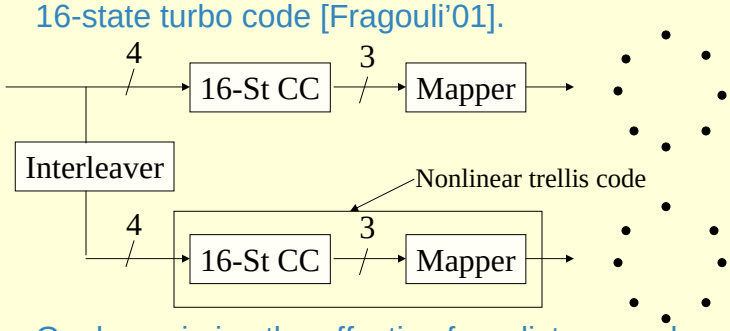
- A new probabilistic device:
- Uniform interleaver for linear codes (Benedetto): "A uniform interleaver of length *k* is a probabilistic device that maps a given input word of Hamming weight *w* into all distinct permutations of it with equal probability".
  - This definition is useful only when assuming the all-zero codeword is transmitted.
- New definition for ANY code: "A uniform interleaver of length *k* is a probabilistic device the chooses between all possible position permutations with equal probability. Then, for each position, the value of the symbol can be changed to any other value with equal probability".
  - This new definition includes the linear case and leads to similar conclusions for both linear and nonlinear constituent codes.

Trellis structure and effective free distance

- The new uniform interleaver analysis allows to show that a recursive encoder is required for nonlinear codes as well.
- Benedetto showed that an important metric to maximize in the constituent linear codes is the effective free distance. We extend this notion for nonlinear codes:
- **Effective free distance:** minimum distance in the output produced by any two possible input words with Hamming distance 2.
- This definition includes the linear case.

Example: 2 bits/s/Hz 8PSK

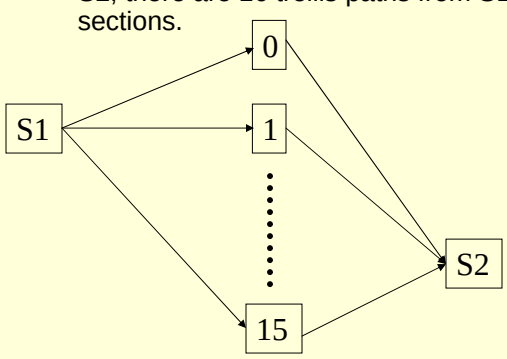
- Constrained capacity: 2.8 dB.
- We compare against the best previously published 16-state turbo code [Fragouli'01].

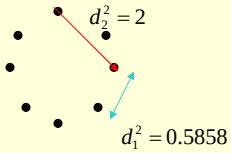


- Goal: maximize the effective free distance, where the distance in this case is the squared Euclidean distance.

Example: 2 bits/s/Hz 8PSK

- Achieving greater effective free distance than linear codes:
  - This is a fully connected trellis: given any two states S1 and S2, there are 16 trellis paths from S1 to S2 after two trellis sections.

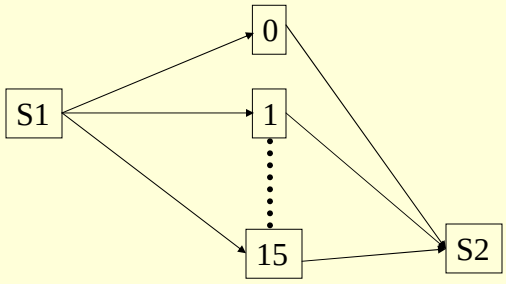


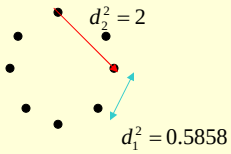


  - Using a linear convolutional code and a mapper, all the searches we and previous works have done, have at least on pair of those paths that have a distance of $2d_1^2 = 1.171573$

Example: 2 bits/s/Hz 8PSK

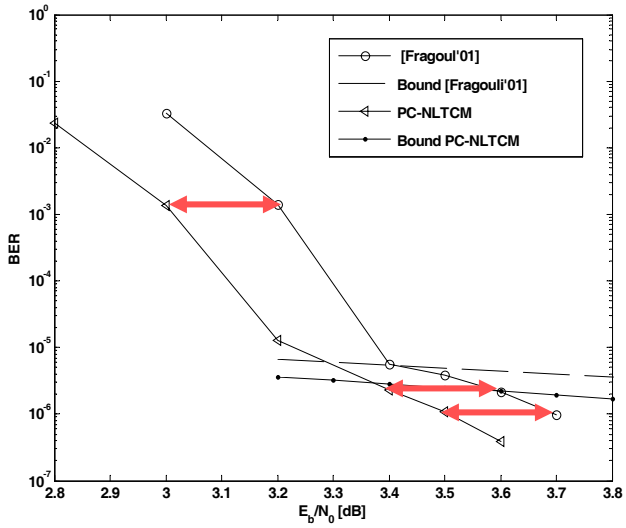
- Achieving greater effective free distance than linear codes:





  - We found a nonlinear labeling that guarantees that for each pair of those paths, there is at least one trellis sections where the constellation points have squared Euclidean distance 2.
  - There could still be a path in more than 2 trellis sections with distance less than 2. By a search in the trellis structure we avoid that.
  - The resulting nonlinear constituent code has an effective free distance equal to 2.

0.2 dB gain for rate-2/3, 16-states over the best published 10,000 bit code.



Observations

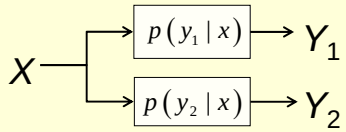
- Interleaver design plays an important role in the code's performance as shown in [Fragouli'01]. The uniform interleaver analysis averages over all possible interleaver, which explains the difference in the error flour between simulations and the bound.
- However, the BER bounding analysis helps in the design process.

Recent Results

- A more efficient LDPC decoding algorithm
- Dynamically scheduled LDPC decoder that converge to a better frame error rate than previous state-of-the-art.
- A powerful family of low-rate turbo codes for "rate-less" applications at the physical layer.
- Turbo Code with closest-to-Shannon performance through novel nonlinear labeling.
- New information theoretic results and nonlinear turbo codes for multi-rate broadcast.

Broadcast Channels

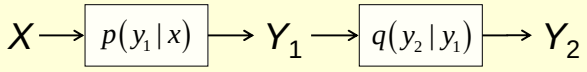
- One transmitter broadcasts information messages to several receivers.



- Receivers decode messages without collaboration.
- Still an open problem. (outer bounds [Sato78], [Marton79], inner bounds [Van75], [Cover75], [Hajek79], [Marton79])

Degraded Broadcast Channel

- Physically Degraded Broadcast Channels
  - The worse channel is a further distortion of the better channel.



- Stochastically Degraded Broadcast Channels
  - The marginal distribution of the worse channel could have been produced by a further distortion of the better channel.

$$p(y_2 \mid x) = \sum_{y_1 \in \mathcal{Y}_1} p(y_1 \mid x) q(y_2 \mid y_1)$$

Degraded Broadcast Capacity Region

[Cover72][Bergmans73][Gallager74]

$$X_2 \longrightarrow \boxed{p(x \mid x_2)} \longrightarrow X \longrightarrow \boxed{p(y_1 \mid x)} \longrightarrow Y_1 \longrightarrow \boxed{q(y_2 \mid y_1)} \longrightarrow Y_2$$

- The capacity region is the convex hull of the closure of all rate pairs (R_1, R_2) satisfying

$$R_1 \le I(X; Y_1 \mid X_2),$$
$$R_2 \le I(X_2; Y_2),$$

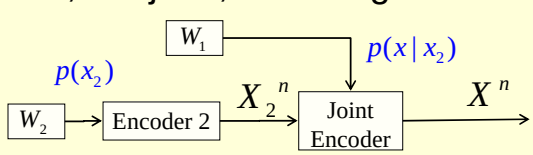
- for some joint distribution

$$p(x, x_2, y_1, y_2) = p(x_2)\, p(x \mid x_2)\, p(y_1, y_2 \mid x).$$

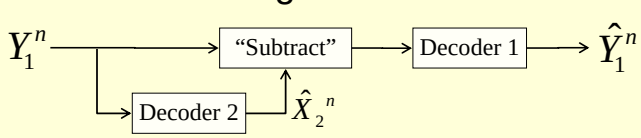
Successive Encoding and Decoding

$$X_2 \longrightarrow \boxed{p(x \mid x_2)} \longrightarrow X \longrightarrow \boxed{p(y_1 \mid x)} \longrightarrow Y_1 \longrightarrow \boxed{q(y_2 \mid y_1)} \longrightarrow Y_2$$

- Successive, but joint, encoding

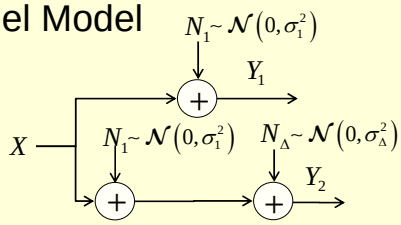


- User 2 is decoded treating user 1 as noise.
- Successive decoding for user 1

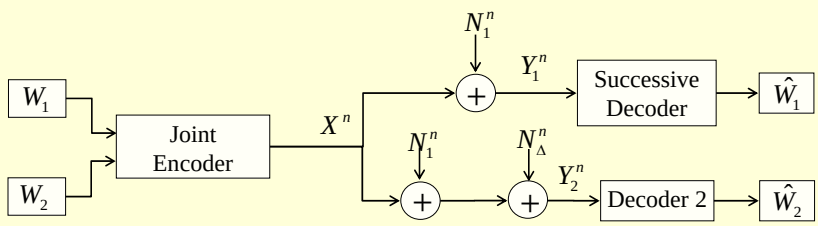


Degraded Gaussian Broadcast Channel

– Channel Model



– General Transmission Strategy



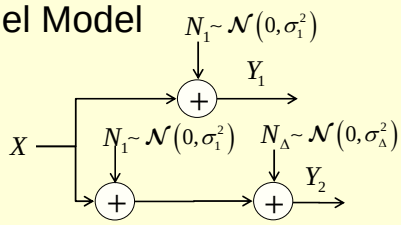
UCLA Electrical Engineering Department

Communication Systems Laboratory

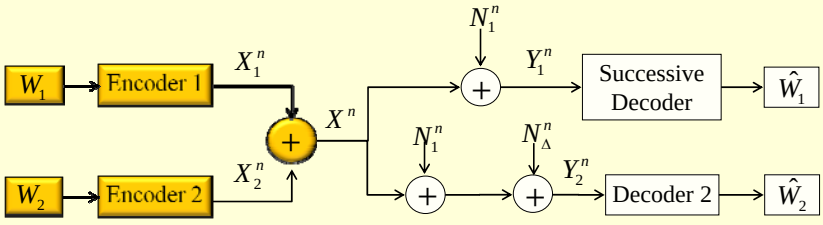
43

Degraded Gaussian Broadcast Channel

– Channel Model



– Optimal Transmission Strategy [Bergmans74]



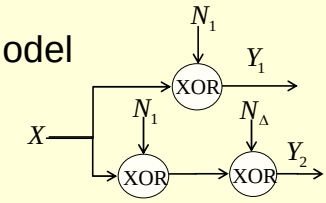
UCLA Electrical Engineering Department

Communication Systems Laboratory

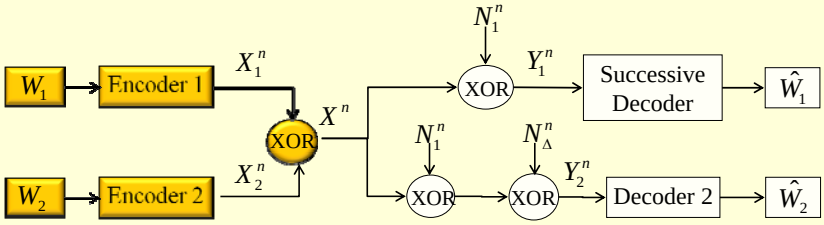
44

Degraded Binary Symmetric Broadcast Channel

– Channel Model

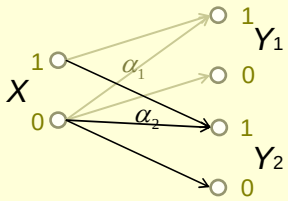


– Optimal Transmission Strategy Optimal Transmission Strategy [Cover72][Bergmans73]



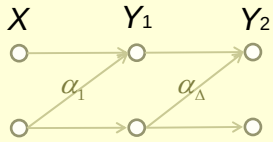
Broadcast Z Channels

- Broadcast Z Channels



$$0 < \alpha_1 < \alpha_2 < 1$$

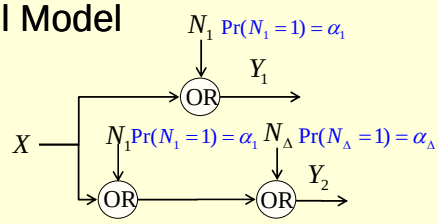
- Broadcast Z channels are stochastically degraded broadcast channels.



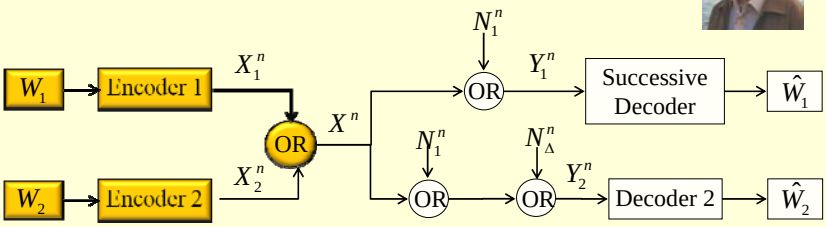
$$\alpha_\Delta = \frac{\alpha_2 - \alpha_1}{1 - \alpha_1}$$

Degraded Z Broadcast Channel

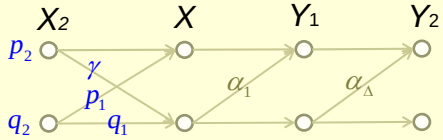
– Channel Model



– Optimal Transmission Strategy[Xie07]



Proving the Optimal Transmission Strategy



- All rate pairs on the boundary of the capacity region can be achieved with the following strategy:

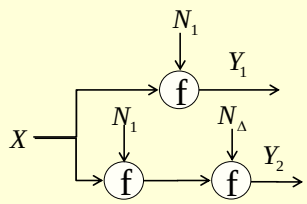
$$\gamma = 0,$$

$$\frac{1}{(1-\alpha_1)(e^{H(1-\alpha_1)/(1-\alpha_1)}+1)} \le q_1 \le 1,$$

$$H(q_1(1-\alpha_2)) - q_1(1-\alpha_2)\log\frac{1-q_2q_1(1-\alpha_2)}{q_2q_1(1-\alpha_2)} = \frac{\log(1-q_1(1-\alpha_2))}{\log(1-q_1(1-\alpha_1))}(H(q_1(1-\alpha_1)) - q_1H(1-\alpha_1)).$$

A class of degraded broadcast channels

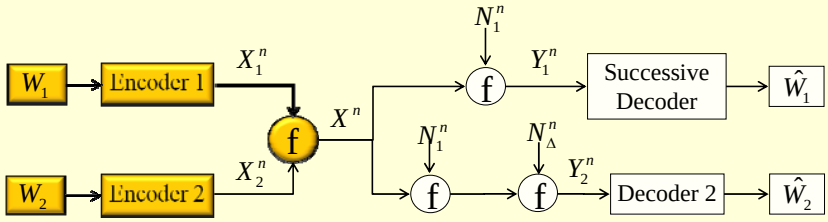
- function-like component channels



- $X, N_1, N_\Delta, Y_1, Y_2$ have the same alphabet.
- The channel function is commutative.
- The channel function is associative.

A Conjecture

- We believe that the optimal transmission strategies for this class of degraded broadcast channels are independent encoding schemes.



Explicit Broadcast Z Channel Capacity Region

– The boundary of the capacity region is

$$\begin{cases} R_1 = q_2 H(q_1(1-\alpha_1)) - q_2 q_1 H((1-\alpha_1)) \\ R_2 = H(q_2 q_1(1-\alpha_2)) - q_2 H(q_1(1-\alpha_2)) \end{cases}$$

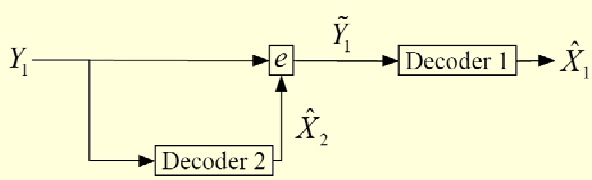
where parameters q_1, q_2 satisfy

$$\frac{1}{(1-\alpha_1)(e^{H(1-\alpha_1)/(1-\alpha_1)}+1)} \le q_1 \le 1,$$

$$H(q_1(1-\alpha_2)) - q_1(1-\alpha_2)\log\frac{1-q_2 q_1(1-\alpha_2)}{q_2 q_1(1-\alpha_2)} = \frac{\log(1-q_1(1-\alpha_2))}{\log(1-q_1(1-\alpha_1))}(H(q_1(1-\alpha_1)) - q_1 H(1-\alpha_1)).$$

Z-channel Successive Decoder

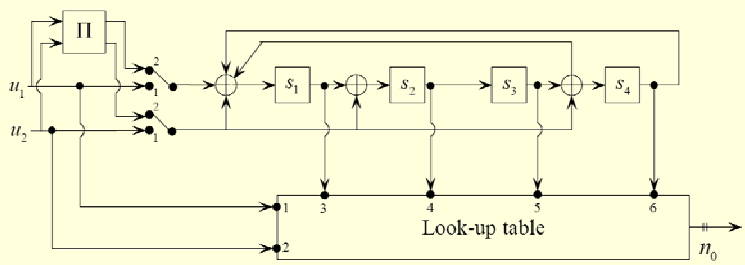
- Decoder structure of the successive decoder for user 1



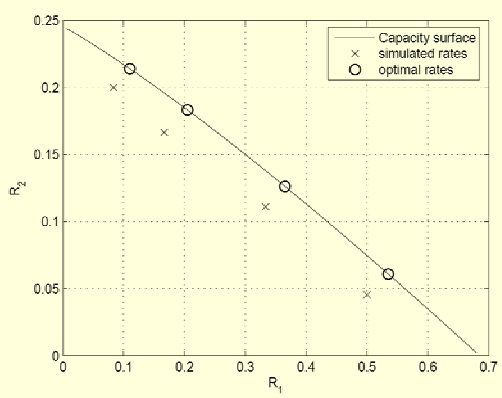
$$\tilde{y}_1 = e(y_1, \hat{x}_2) = \begin{cases} y_1 & \text{if} \quad \hat{x}_2 = 0 \\ erasure & \text{if} \quad \hat{x}_2 = 1 \end{cases}$$

Nonlinear Turbo Codes

- Nonlinear turbo codes can provide a controlled distribution of ones and zeros.
- Nonlinear turbo codes designed for Z channels are used. [Griot06]
- Encoding structure of nonlinear turbo codes:



Simulation Results



- The cross probabilities of the broadcast Z channel are

$$\alpha_1 = 0.15, \alpha_2 = 0.6.$$

- The simulated rates are very close to the capacity region.
- Only 0.04 bits or less away from optimal rates in R_1.
- Only 0.02 bits or less away from optimal rates in R_2.

R_1	R_2	p_1	p_2	K_1	K_2	BER_1	BER_2
1/12	1/5	0.106	0.56	4800	1700	2.54×10^{-5}	1.24×10^{-5}
1/6	1/6	0.196	0.5	2048	2048	7.01×10^{-6}	5.33×10^{-6}
1/3	1/9	0.336	0.3739	4608	1536	7.13×10^{-6}	6.70×10^{-6}
1/2	1/22	0.463	0.1979	5632	1024	9.27×10^{-7}	3.27×10^{-6}

References

- [Cover98] T. M. Cover, “Comments on broadcast channels,” *IEEE Trans. Inform. Theory*, vol. IT-44, pp. 2524-2530, 1998
- [Cover75] T. M. Cover, “An achievable rate region for the broadcast channel,” *IEEE Trans. Inform. Theory*, vol. IT-21 ,pp. 399-404, 1975
- [Hajek79] B. E. Hajek & M. B. Pursley, “Evaluation of an achievable rate region for the broadcast channel,” *IEEE Trans. Inform. Theory*, vol. IT-25 ,pp. 36-46, 1979
- [Gallager74] R. G. Gallager, “Capacity and coding for degraded broadcast channels,” *Probl. Pered. Inform.,* vol. 10, no. 3, pp. 3–14, July–Sept. 1974; *translated in Probl. Inform. Transm*., pp. 185–193, July–Sept. 1974.
- [Benzel79] R. Benzel, “The capacity region of a class of discrete additive degraded interference channels,” *IEEE Trans. Inform. Theory,* vol. IT-25, pp. 228–231, Mar. 1979.

References

- [Sato78] H. Sato, “An outer bound to the capacity region of broadcast channels*,” IEEE Trans. Inform. Theory*, vol. IT-24 ,pp. 374-377, 1978.
- [VAN75] E. van der Meulen, “Random coding theorems for the general discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 180-190, 1975.
- [Marton79] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 306–311,1979.
- [Cover72] T. M. Cover, “Broadcast channels,” *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.
- [Bergmans73] P. P. Bergmans, “Random coding theorem for broadcast channels with degraded components,” *IEEE Trans. Inform. Theory,* vol. IT-19, pp. 197–207, Mar. 1973.

References

- [Bergmans74] P. P. Bergmans, “A simple converse for broadcast channels with additive white Gaussian noise,” *IEEE Trans. Inform. Theory,* vol. IT-20, pp. 279–280, Mar. 1974.
- [Griot06] M.Griot, A. I. Vila Casado and R. D. Wesel, "Non-linear Turbo Codes for Interleaver-Division Multiple Access on the OR Channel". Globecom 2006, 27 Nov. - 1 Dec., San Francisco, USA.
- [Xie07] B. Xie, M. Griot, A. I. Vila Casado and R. D. Wesel, "Optimal Transmission Strategy and Capacity Region for Broadcast Z Channels", *IEEE Information Theory Workshop 2007.*