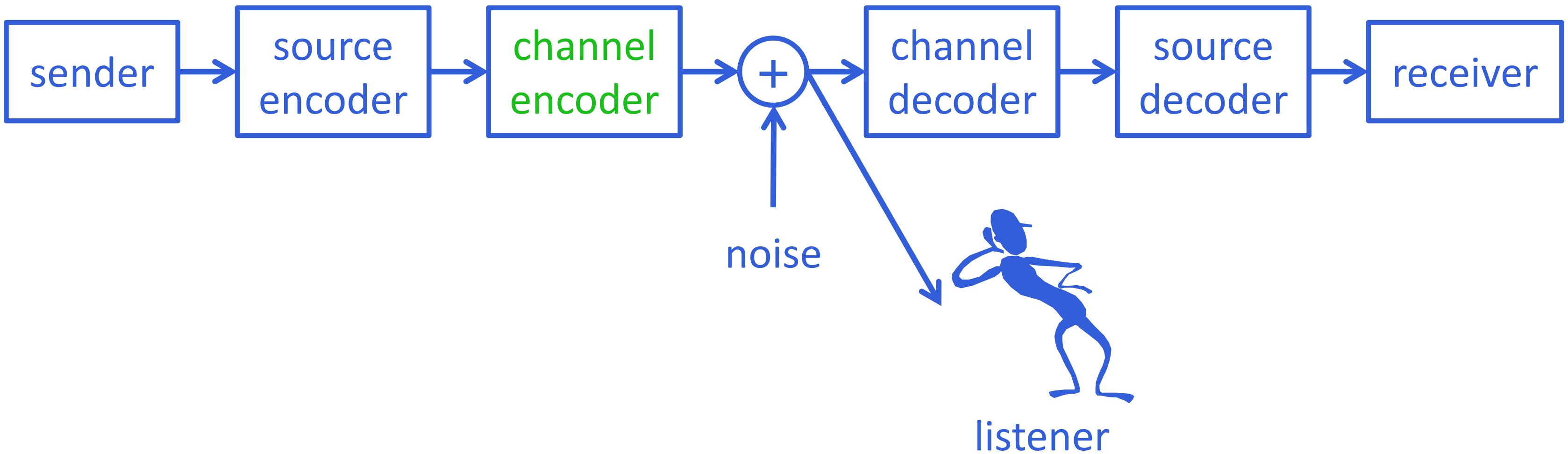
Channel-Code Detection by a Third-Party Receiver via the Likelihood Ratio Test

Arti Yardi, **Animesh Kumar**, and Saravanan Vijayakumaran
Electrical Engineering
Indian Institute of Technology Bombay
Mumbai 400076

ISIT 2014, Honolulu HI

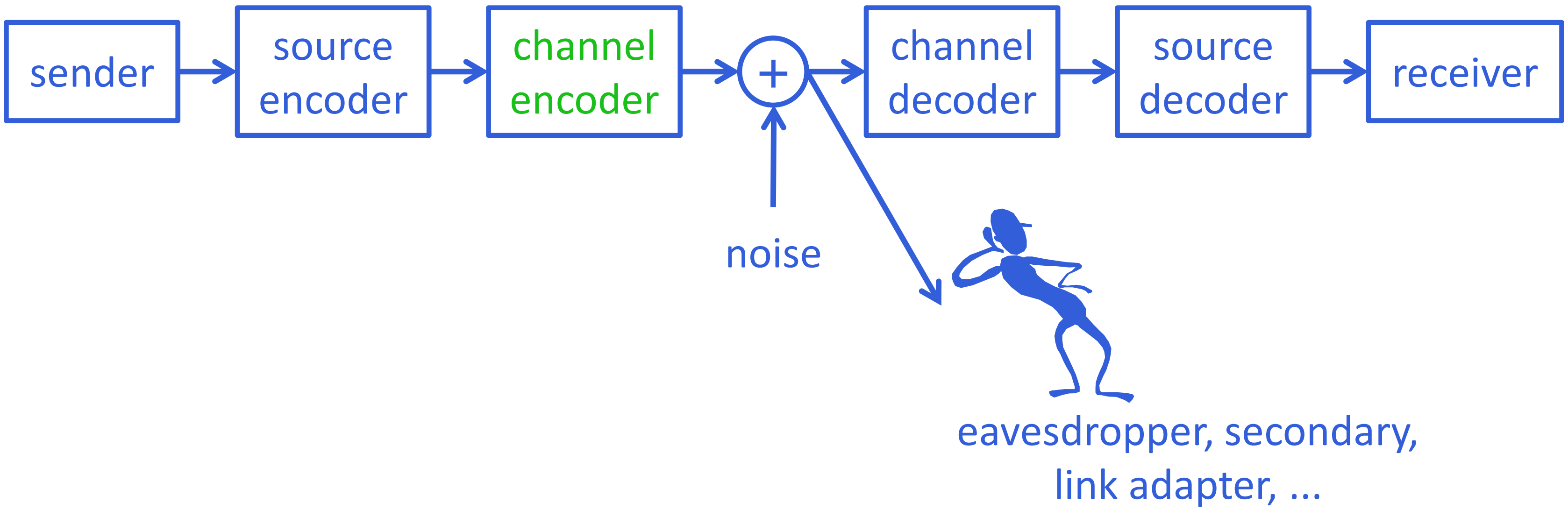
Listener on a channel



◊ The knowledge of channel encoding scheme seems essential to recover the source or message

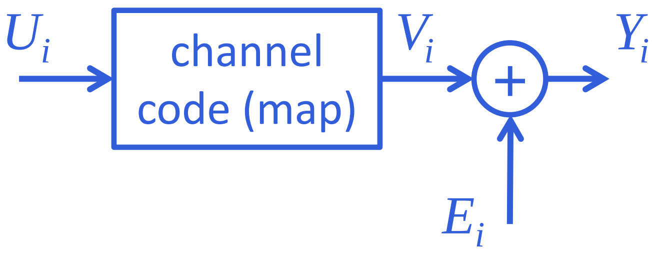
◊ Consider a **listener**, with access to “noisy” bits or symbols, who wants to ascertain the channel code used

Applications of this model



This model has applications in security, or cognitive radios (where a secondary may want to know primary's message), or in link adaptation in some wireless technologies

Zooming in to the "right" problem



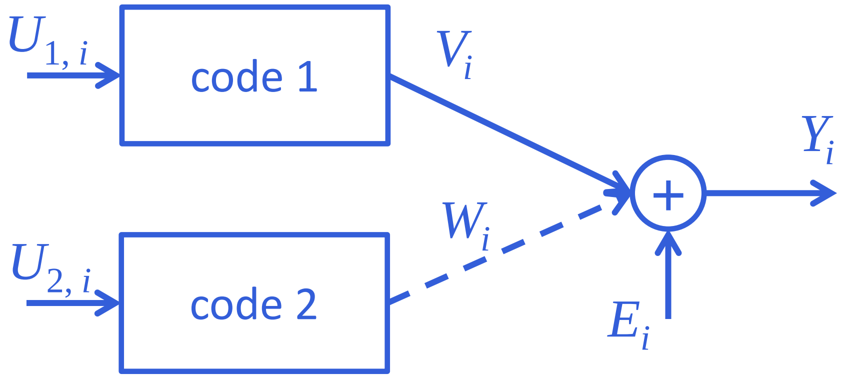
◊ Observe $Y_1, Y_2, \ldots, Y_N$ and find out the channel code (map)

◊ This problem has been explained to be NP-hard **[Valembois'01]**

◊ With some extra information on the channel code, this problem will be addressed by us

◊ We will address the problem in a hypothesis testing setup

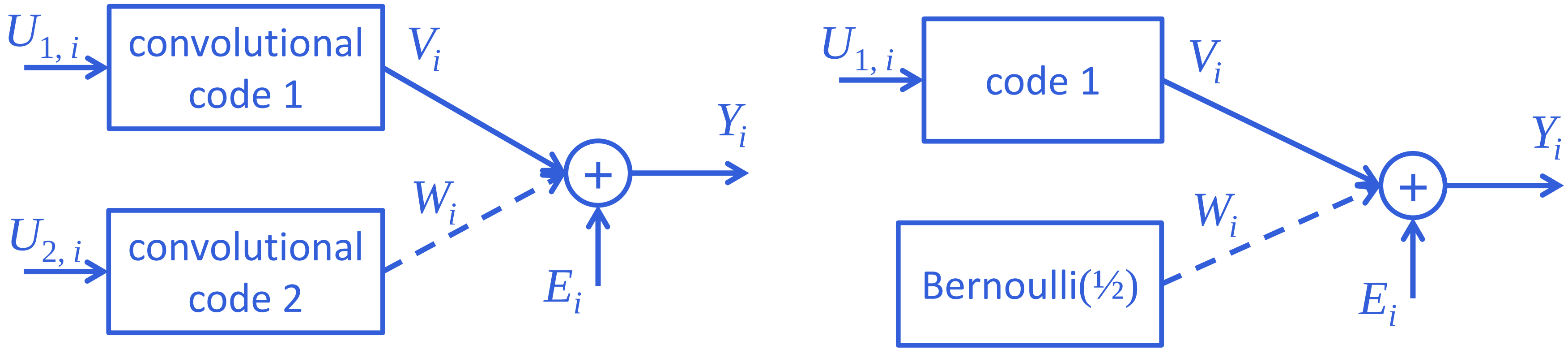
The code-detection problem: assumptions



Message words U_i are mapped to codewords V_i (or W_i) by two different binary linear block codes with parameters $[n, k_1, d_1]$ and $[n, k_2, d_2]$

- ◊ Message words are equally likely, that is, codewords are equally likely
- ◊ Block length n is the same for the two codes
- ◊ In a large deviation setting, vectors $(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_N)$ of (binary, synchronous) observations are available to detect the channel code
- ◊ Noise is IID Bernoulli(p), and indep of the hypothesis and messages

Related work



◊ Single "low-weight" parity check equations have been used for: (i) convolutional code detection **[Moosavi-Larsson'11]** and (ii) distinguishing noise from codewords **[Chabot'07]**

◊ Estimation of channel code from noise-affected bits has been studied for various settings **[Valembois'01] [Cluzeau'06] [Dingel-Hagenauer'07]**

Our key contributions

◊ We use the **likelihood ratio test** for this problem and show that the Chernoff information, that is the optimal error-probability exponent, for the code-detection problem is (strictly) positive if the two hypothesis are different

◊ Likelihood computation, though it leads to min. error probability test, can be difficult. Banking upon the (presence of) efficient BCJR or GDL based decoding, methods to compute the likelihood ratio for code-detection problem are detailed

Outline

◊ Introduction

◊ Chernoff information bound for the code-detection problem

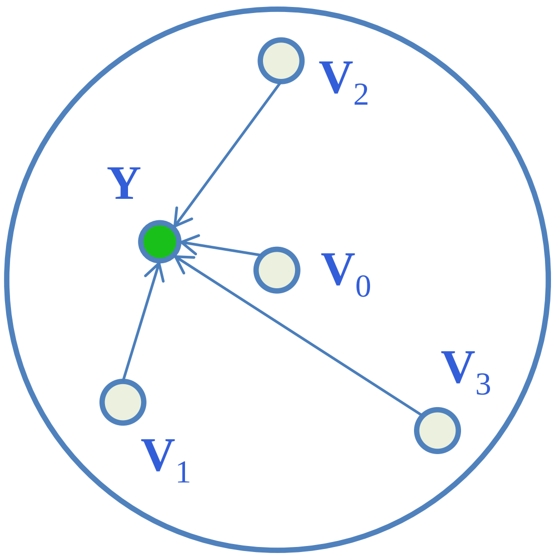
◊ Algorithms for computing likelihood ratio efficiently for code-detection

◊ Concluding remarks and future work

Likelihood computation

◊ The likelihood ratio test will involve the comparison of $f(\mathbf{Y}, H_1)$ against $f(\mathbf{Y}, H_2)$ where H_1 and H_2 are the two hypotheses

◊ The **main difference** between classical decoding and code-detection is that the likelihood depends on the entire codeword **constellation**



$$\begin{aligned} f(\mathbf{Y}, H_1) &= \sum_{\mathbf{v}_i \in \mathcal{C}_1} \mathbb{P}(\mathbf{V} = \mathbf{v}_i) f(\mathbf{Y} | \mathbf{V} = \mathbf{v}_i; H_1) \\ &= \frac{1}{2^{k_1}} \sum_{\mathbf{v}_i \in \mathcal{C}_1} p^{\mathrm{wt}(\mathbf{Y}+\mathbf{v}_i)} (1-p)^{n-\mathrm{wt}(\mathbf{Y}+\mathbf{v}_i)} \end{aligned}$$

◊ This likelihood $f(\mathbf{Y}, H_1)$ is quite challenging to compute and is the key stumbling block in further analysis

Chernoff information

◊ We have a hypotheses testing problem where two distributions, P and Q, corresponding to code 1 and code 2 have to be distinguished where

$$P = (\underbrace{p_{\mathbf{y}_0}, \ldots, p_{\mathbf{y}_{2^n-1}}}_{\text{code 1}}) \quad \text{and} \quad Q = (\underbrace{q_{\mathbf{y}_0}, \ldots, q_{\mathbf{y}_{2^n-1}}}_{\text{code 2}})$$

◊ Then the optimal exponent of detection error-probability is given by the Chernoff information **[Cover-Thomas]**. That is,

$$C(P,Q) = - \min_{0 \le \lambda \le 1} \log \left(\sum_i p_{\mathbf{y}_i}^{\lambda} q_{\mathbf{y}_i}^{1-\lambda} \right)$$

Lower bound on Chernoff information

Chernoff information is difficult to compute since **individual** terms in P and Q are NP-hard to compute. A lower bound on $C(P,Q)$ can be used for analysis **[Sason'13]**

$$C(P,Q) \geq -\frac{1}{2}\log\left(1-(d_{TV}(P,Q))^2\right)$$

where $d_{\mathrm{TV}}(P,Q)$ is (half of) L_1 distance between P and Q

$$d_{TV}(P,Q) \; = \; \tfrac{1}{2}\sum_i |p_{\mathbf{y}_i} \; - \; q_{\mathbf{y}_i}|$$

Likelihood and cosets of the block code

◊ For binary linear block codes, the likelihood only depends on which coset the vector $\mathbf{Y}$ belongs to. This is because

$\{\mathrm{wt}(\mathbf{Y}+\mathbf{v_i}),\ \mathbf{v_i}$ in Code 1$\} = \{\mathrm{wt}(\mathbf{Y}+\mathbf{v_i}+\mathbf{c}),\ \mathbf{c}$ fixed in code 1, $\mathbf{v_i}$ in code 1$\}$

$$f(\mathbf{Y}, H_1) = \frac{1}{2^{k_1}} \sum_{\mathbf{v}_i \in \mathcal{C}_1} p^{\mathrm{wt}(\mathbf{Y}+\mathbf{v}_i)} (1-p)^{n-\mathrm{wt}(\mathbf{Y}+\mathbf{v}_i)}$$

◊ That is, the coset-leaders in standard-array used for decoding can be used to ascertain likelihood for the entire row

$\mathbf{v}_0 = \vec{0}$	$\mathbf{v}_1$	$\mathbf{v}_2$	$\ldots$	$\mathbf{v}_{2^{k_1}}$
$\mathbf{g}$	$\mathbf{g}+\mathbf{v}_1$	$\mathbf{g}+\mathbf{v}_2$	$\ldots$	$\mathbf{g}+\mathbf{v}_{2^{k_1}}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$

Bounds on $(p_{\mathbf{y}} - q_{\mathbf{y}})$

If $\mathbf{y}$ is a codeword in code 1 **and** code 2, then $p_{\mathbf{y}}$ can be computed and is equal to $p_{\mathbf{0}}$. Similarly, if the same $\mathbf{y}$ is a codeword in code 2, then $q_{\mathbf{y}}$ is $q_{\mathbf{0}}$ And $|p_{\mathbf{y}} - q_{\mathbf{y}}|$ is given by $|p_{\mathbf{0}} - q_{\mathbf{0}}|$

If $\mathbf{y}$ is a codeword in code 1 **and not** in code 2, then $p_{\mathbf{y}}$ can be computed and is equal to $p_{\mathbf{0}}$. The same $\mathbf{y}$ is not a codeword in code 2, then $q_{\mathbf{y}}$ is bounded using $q_{\mathbf{0}}$ as follows

[Ancheta'81] $$\underbrace{q_{\mathbf{0}} \left(\frac{p}{1-p} \right)^{n-k_2}}_{q_{\mathbf{y}_L}} \leq q_{\mathbf{y}_i} \leq \underbrace{q_{\mathbf{0}} \frac{1-(1-2p)^{k_2+1}}{1+(1-2p)^{k_2+1}}}_{q_{\mathbf{y}_H}}$$ **[Sullivan'67]**

Main result

Bounds on $|p_{\mathbf{y}} - q_{\mathbf{y}}|$ for cases where $\mathbf{y}$ belongs in code 1 or code 2 or both

$p_0 - q_0$	$p_0 - q_{\mathbf{y}H}$
$\alpha = \max\{q_{\mathbf{0}} - p_{\mathbf{y}H}, p_{\mathbf{y}L} - q_{\mathbf{0}}, 0\}$	$\beta = \max\{p_{\mathbf{y}L} - q_{\mathbf{y}H}, q_{\mathbf{y}L} - p_{\mathbf{y}H}, 0\}$

Theorem: Assume $p_0 - q_0 \geq 0$. The $d_{\mathrm{TV}}(P, Q)$ and consequently Chernoff information has a strictly positive lower-bound for code-detection

$$d_{TV}(P,Q) \geq 2^m(p_{\mathbf{0}} - q_{\mathbf{0}}) + (2^{k_1} - 2^m)(p_{\mathbf{0}} - q_{\mathbf{y}_H}) + (2^{k_2} - 2^m)\alpha + (2^n - 2^{k_1} - 2^{k_2} - 2^m)\beta$$

where m is the dimension of code 1 intersection with code 2

Outline

◊ Introduction

◊ Chernoff information bound for the code-detection problem

◊ Algorithms for computing likelihood ratio efficiently for code-detection

◊ Concluding remarks and future work

Fast algorithms for likelihood calculation

When the two channel codes "code 1" and "code 2" can be (efficiently) decoded using (i) the GDL **[Aji-McEliece'00]** or the (ii) BCJR algorithm **[Bahl-Cocke-Jelinek-Raviv'74]**, then the likelihoods $f(\mathbf{Y}, H_1)$ against $f(\mathbf{Y}, H_2)$ can be found efficiently using some intermediate steps in the two algorithms

Algorithm based on the GDL

Using Baye's rule, it can be shown that

$$\begin{aligned}\mathbb{P}(V_i = 1|\mathbf{Y}; H_1) &= \frac{1}{2^{k_1}} \frac{1}{f(\mathbf{Y}; H_1)} \sum_{v_i = 1, \mathbf{v} \in \text{code } 1} f(\mathbf{Y}|\mathbf{V} = \mathbf{v}; H_1) \\ &= \frac{1}{2^{k_1}} \frac{1}{f(\mathbf{Y}; H_1)} L_{H_1}(1)\end{aligned}$$

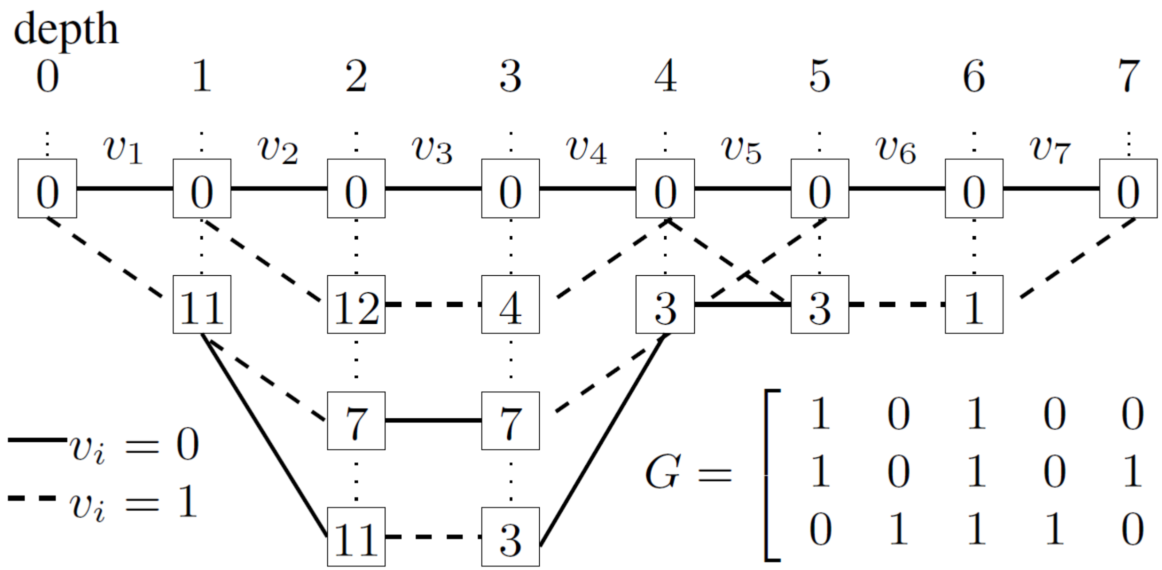
If code 1 has a junction tree, this can

be computed efficiently using GDL

The desired likelihood can be obtained using

$$\frac{1}{2^{k_1}} \frac{1}{f(\mathbf{Y}; H_1)} \left[L_{H_1}(0) + L_{H_1}(1) \right] = 1$$

Algorithm based on the BCJR algorithm



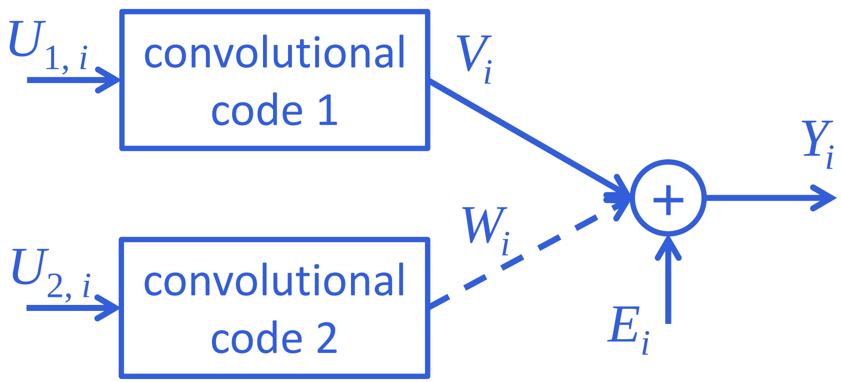
$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

◊ Let S_i be the state random variable at depth i

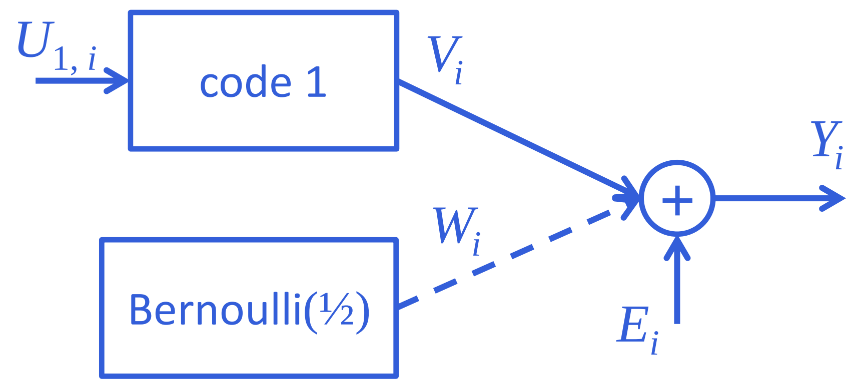
◊ The BCJR algorithm calculates Prob($S_i = m$, $\mathbf{Y}$) in an intermediate step during decoding

◊ By adding Prob($S_i = m$, $\mathbf{Y}$) over states m, $f\,(\mathbf{Y}, H_1)$ can be obtained

Recap



Parity-check method

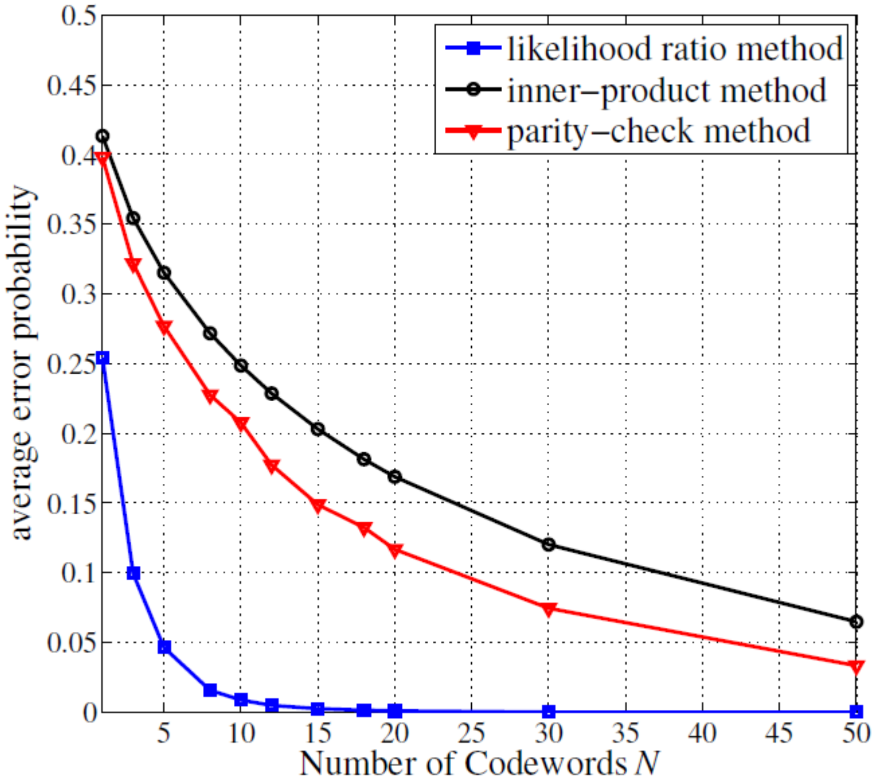


Inner-product method

◊ Single "low-weight" parity check equations have been used for: (i) convolutional code detection **[Moosavi-Larsson'11]** and (ii) distinguishing noise from codewords **[Chabot'07]**

Simulations for the average error-probability

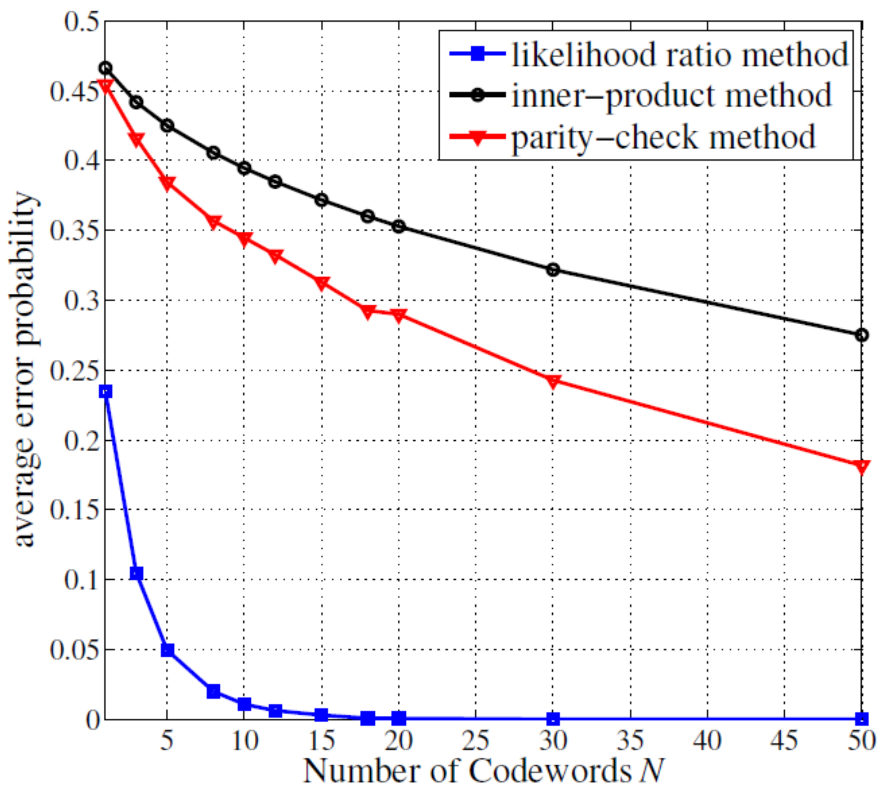
Plot of average error probability versus N for inner-product method **[Chabot'07]**, parity-check method **[Moosavi-Larsson'11]** and our method for H_1: Hamming(15,11) and H_2: BCH(15,7) hypotheses



$p = 0.1$

Simulations for the average error-probability

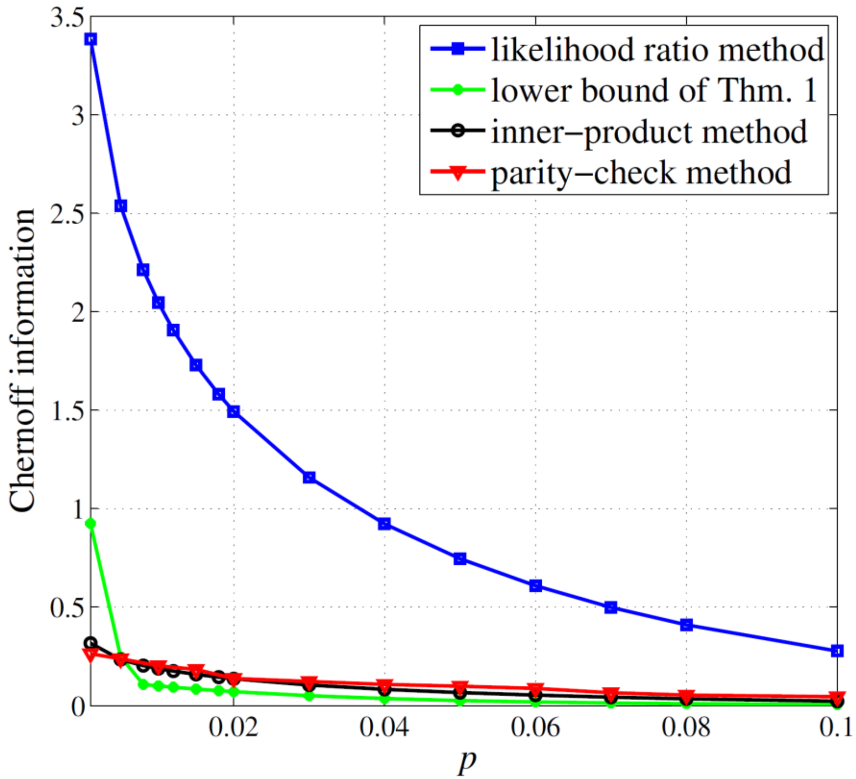
More simulations where two hypotheses are H_1: Hamming(31,26) and H_2: BCH(31,16)



$p = 0.1$

Simulations for the Chernoff information

Plot of Chernoff information for the inner-product method **[Chabot'07]**, parity-check method **[Moosavi-Larsson'11]** , our lower bound, and likelihood ratio method for H_1: Hamming(15,11) and H_2: BCH(15,7)



Outline

◊ Introduction

◊ Chernoff information bound for the code-detection problem

◊ Algorithms for computing likelihood ratio efficiently for code-detection

◊ Concluding remarks and future work

Conclusions

◊ **The likelihood test's error-exponent:** we showed that the Chernoff information for the code-detection problem is strictly positive for two hypotheses consisting of binary linear block codes

◊ **Likelihood calculation:** banking upon the existence of efficient GDL or BCJR decoding algorithms, efficient methods to compute the likelihood ratio test was shown

Future work

Code-detection problem

◊ where two hypotheses consist of linear block codes with unequal block lengths

◊ more than two hypotheses

◊ where codes which are not linear or do not have a block structure

◊ when the two hypotheses consist of LDPC codes (where decoding is efficient)

◊ ...