Institute for the Wireless Internet of Things at Northeastern University

Mathematical Picture Language Seminar

An introduction to Forward Error Correction and Guessing Random Additive Noise Decoding

Ken Duffy
Professor, Department of Electrical and Computer Engineering
Professor, Department of Mathematics
Faculty Member, Institute for the Wireless Internet of Things
Northeastern University
k.duffy@northeastern.edu
epic.sites.northeastern.edu

A trip back in time

Dublin 2005

Institute for the Wireless Internet of Things at Northeastern

DUBLIN INSTITUTE FOR ADVANCED STUDIES

J.T. LEWIS MEMORIAL CONFERENCE

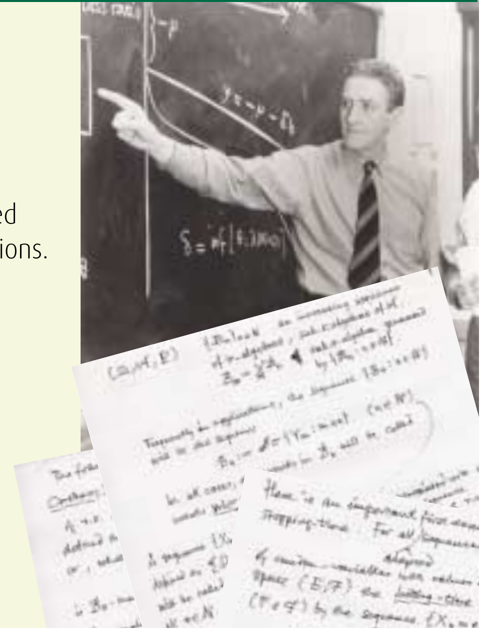
JUNE 14TH -17TH 2005

Dublin Institute of Technology, Dublin, Ireland.

The conference will focus on three broad areas of applied mathematics in which John Lewis made major contributions. These are:

(i) quantum mechanics;
(ii) statistical mechanics;
(iii) communications theory.

The conference will consist of plenary talks and parallel sessions in the above topics. The emphasis will be squarely on modern developments.



Plenary Speakers

JENNIFER CHAYES
DAVID EVANS
GEORGE W. FORD
ARTHUR JAFFE
FRANK KELLY
CHRISTOPHER KING
DEREK McAULEY
CATHLEEN MORAWETZ
NEIL O'CONNELL
RAYMOND RUSSELL
ANDRE VERBEURE
MARC YOR

ORGANISING COMMITTEE:

Tony Dorlas dorlas@stp.dias.ie
Ken Duffy ken.duffy@nuim.ie
Brendan Goldsmith brendan.goldsmith@dit.ie

CONFERENCE ADMINISTRATOR:

Marguerite Carter, CNRI, Focas Institute, DIT, Kevin Street, Dublin 8, Ireland

T: 00353 1 4027903
F: 00353 1 4027901
E: cnri@dit.ie
W: http://www.cnri.dit.ie/lewis_2005.html

Venue: DIT, Aungier Street, Dublin 2, Ireland.

Supported by NDP NATIONAL DEVELOPMENT PLAN sfi science foundation ireland fondúireacht eolaíochta éireann

Dublin 2005

Acknowledgements

Collaborators and Acknowledgements

Institute for the Wireless Internet of Things
at Northeastern

Muriel Medard
MIT

Rabia Yazicigil
Boston University

Science Foundation Ireland sfi For what's next

NSF

DARPA

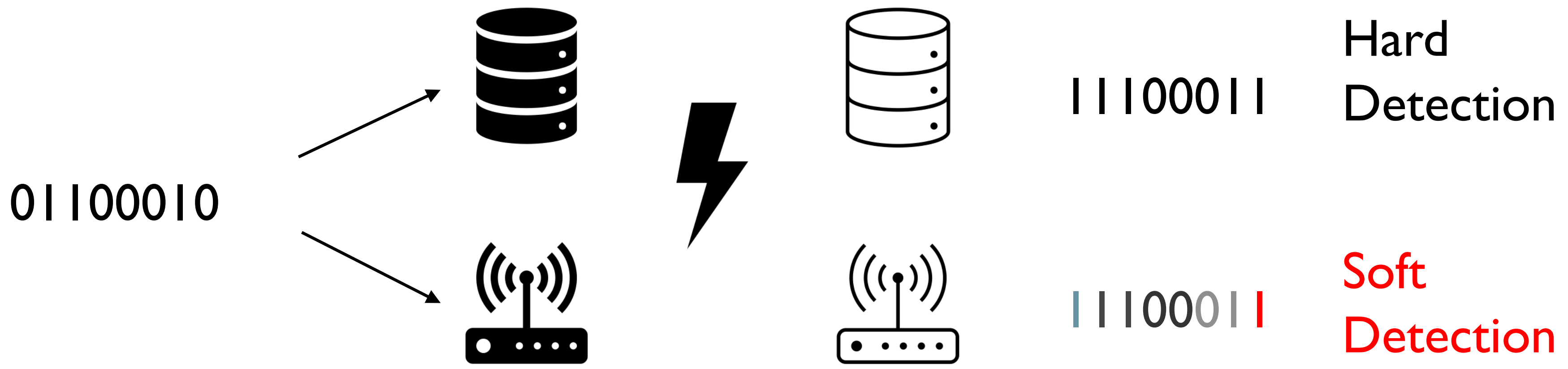
- Wei An
- Joe Griffin
- Basak Ozaydin
- Amit Solomon
- Kathleen Yang

- Kishori Konwar
- Jiange Li
- Hadi Sarieddeen
- Peihong Yuan
- Kevin Galligan
- Moritz Grundei

- Vaibhav Bansal
- Qijun Liu
- Jonathan Ngo
- Arslan Riaz
- Alperen Yasar
- Furkan Ercan

Context

Error correction coding



Error correction

Shannon (1948):

- Error detection and correction is possible only if a subset of strings are code-words.
- Out of 2^n possible strings, 2^k are code-words, giving a rate of $R=k/n$.
- The highest rate a code-book can be depends on a statistic of the corruption that
  w
- B
- H

My intentions in this talk?
Entirely dishonorable and epsilontics will be left to the listener!

Joe Doob

> …it is not always clear that the author's mathematical intentions are honorable. (MR0026286)

> Following tradition, the "detailed epsilontics" of the proof of the fundamental theorem are omitted. (MR0055621)

Shannon, *Bell. Sys. Tech. J.*, 1948. Duffy, *London Math. Soc. Newsletter*, 2021.

Error correction coding

Institute for the Wireless Internet of Things at Northeastern

Shannon (1948):

- Error detection and correction is possible only if a subset of strings are code-words.
- Out of 2^n possible strings, 2^k are code-words, giving a rate of $R=k/n$.
- The highest rate a code-book can be depends on a statistic of the corruption that we now call the Shannon Entropy.
- Best correction performance bang for buck comes at long code-lengths.
- In practice, for communication and storage of digital data, almost all error correction codes are linear in the binary field of two elements, F_2

$$a^k G = c^n .$$

- Linear codes = perfect grammar.

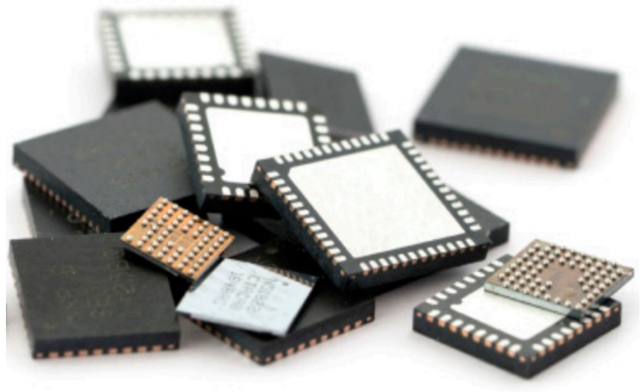
Berlekamp, McEliece & Van Tilborg (1978): optimal hard detection decoding of linear codes is NP-complete.

Shannon, *Bell. Sys. Tech. J.*, 1948. Berlekamp, McEliece & Van Tilborg, *IEEE Tran. Inf. Theory*, 1978.

Complexity Limits of Forward Error Correction

Institute for the Wireless Internet of Things at Northeastern

Practical consequence is the current paradigm: co-design of restricted code (i.e. grammar) and decoder pairs.



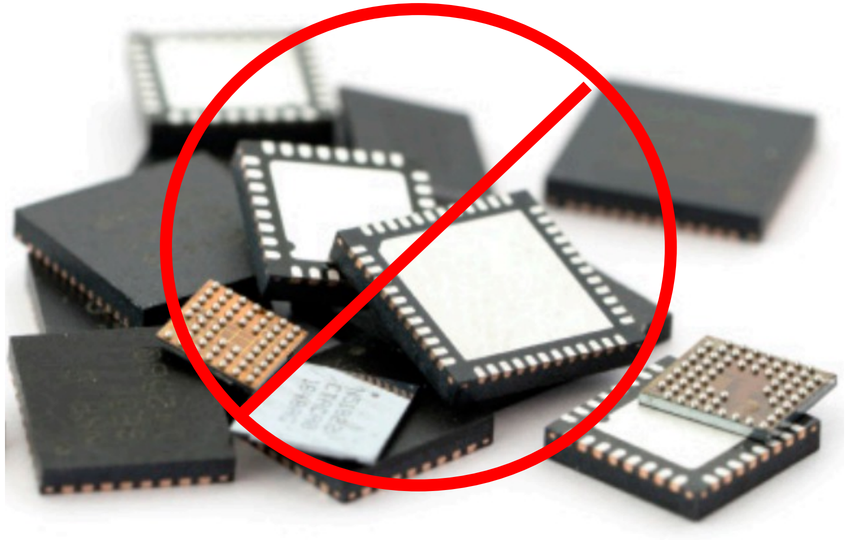
- Error Detection Only ← Cyclic Redundancy Check
- Majority Logic ← RM
- Berlekamp-Massey ← BCH
- CA-SCL ← CA-Polar
- No decoder ← Random Linear Code

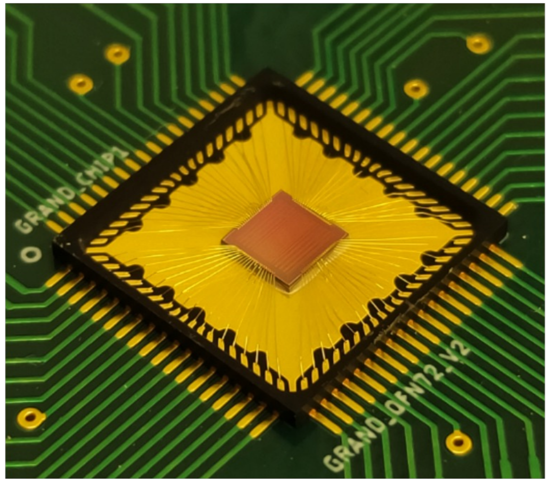
- Distinct chip required to decode each code.
- Requires standardization.

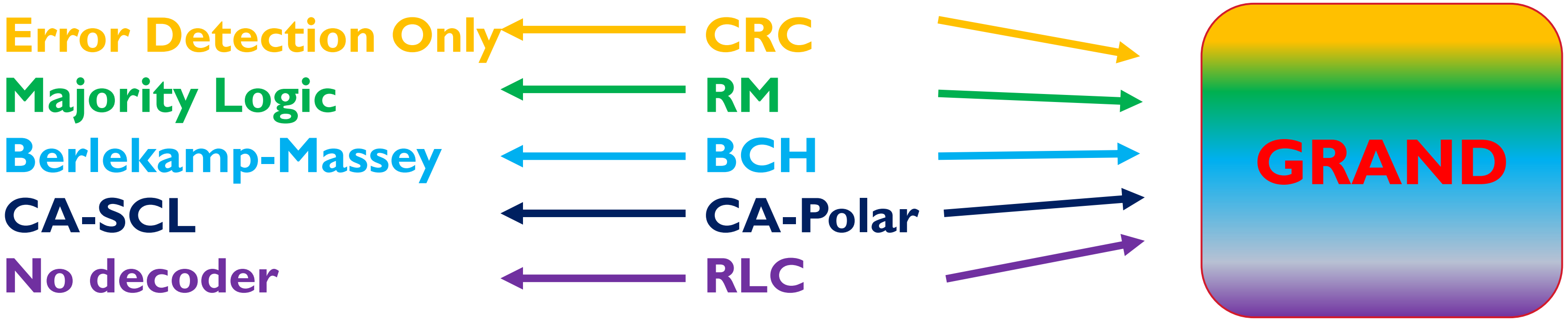
5G

LDPC – 1960s
CA-Polar – 2010s

Guessing Random Additive Noise Decoding







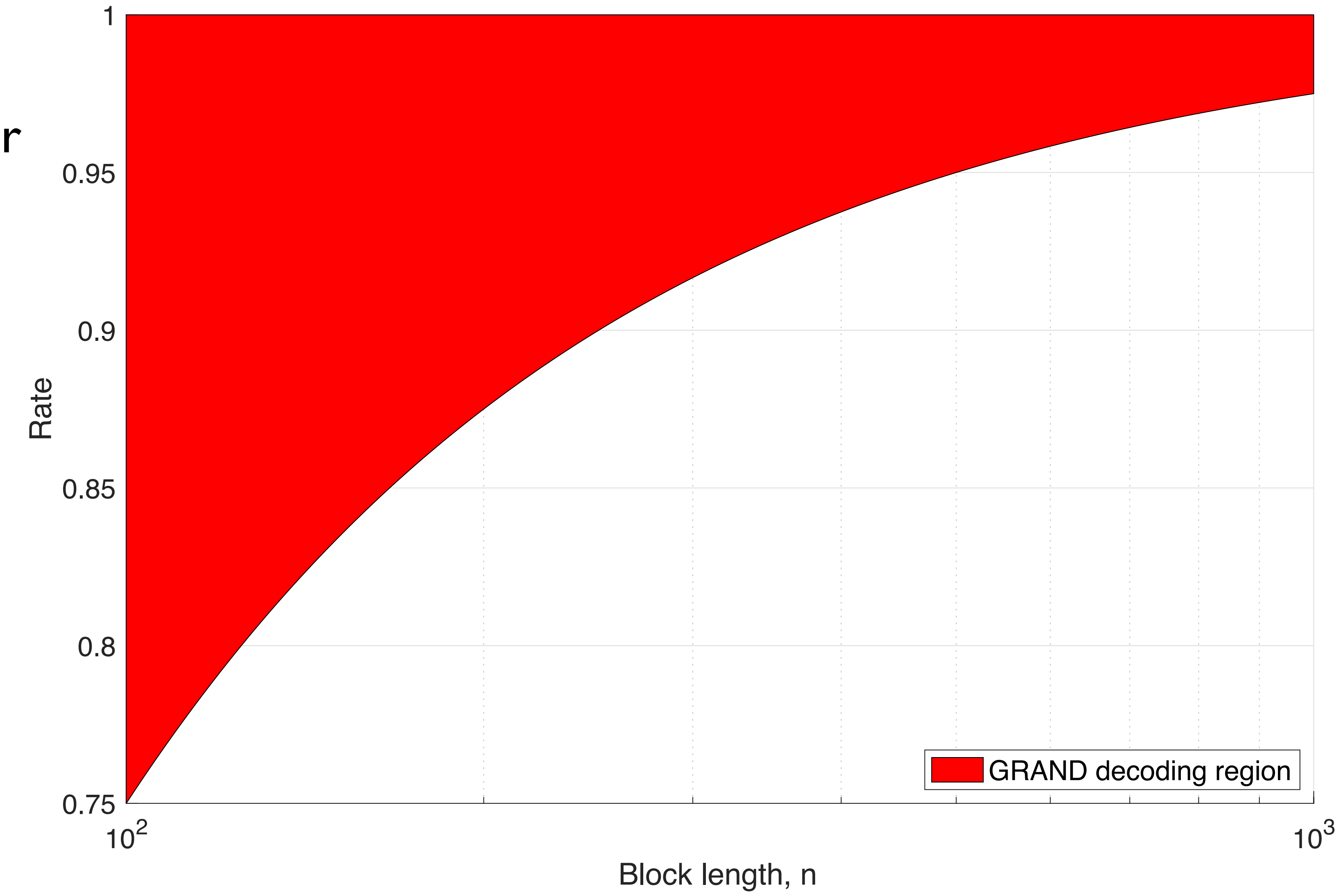
Guessing Random Additive Noise Decoding

Institute for the Wireless Internet of Things at Northeastern

Practical decoding region

Institute for the Wireless Internet of Things at Northeastern

- A function of the redundancy, n-k, rather than k/n.



Idea behind GRAND

Channel output is input plus noise effect

$$Y^n = \underbrace{X^n}_{2^{nR}} \oplus \underbrace{N^n}_{2^{nH}}$$

Standard decoder: identify X^n using structure of code-book
GRAND: identify N^n using structure of the noise

Inputs: Code-book membership test, Y^n.
Output: Decoding $c^{*,n}$.
$y^n \leftarrow \text{demod}(Y^n)$.
$d \leftarrow 0$.
while $d = 0$ **do**
  $z^n \leftarrow$ next most likely noise effect
  if $y^n \ominus z^n$ is in the code-book **then**
    $c^{*,n} \leftarrow y^n \ominus z^n$
    $d \leftarrow 1$
    return $c^{*,n}$.
  end if
end while

- **Universal** decoders suitable for moderate redundancy codes.
- **Complexity** a function of noise and redundancy, not code-rate.
- Highly **parallelizable**.

Duffy, Li, Médard, *IEEE Tran. Inf. Theory*, 19. Duffy, Li, Médard, *IEEE ISIT*, 18.

GRAND is max. likelihood if channel match

Institute for the Wireless Internet of Things at Northeastern

- Channel output is input plus independent noise: $Y^n = X^n \oplus N^n$

- Max. likelihood decoding:

$$c^{n,*} \in \arg\max \left\{ p(y^n | c^{n,i}) : c^{n,i} \in \mathcal{C}_n \right\}$$
$$= \arg\max \left\{ P(N^n = y^n \ominus c^{n,i}) : c^{n,i} \in \mathcal{C}_n \right\}$$

- Max. likelihood decoding by sequential guessing

$$P(N^n = y^n \ominus c^{n,*}) \geq P(N^n = y^n \ominus c^{n,i}) \text{ for all } c^{n,i} \in \mathcal{C}_n$$

- As maximum likelihood decoding is optimal for uniform sources, automatically get existing capacity results.

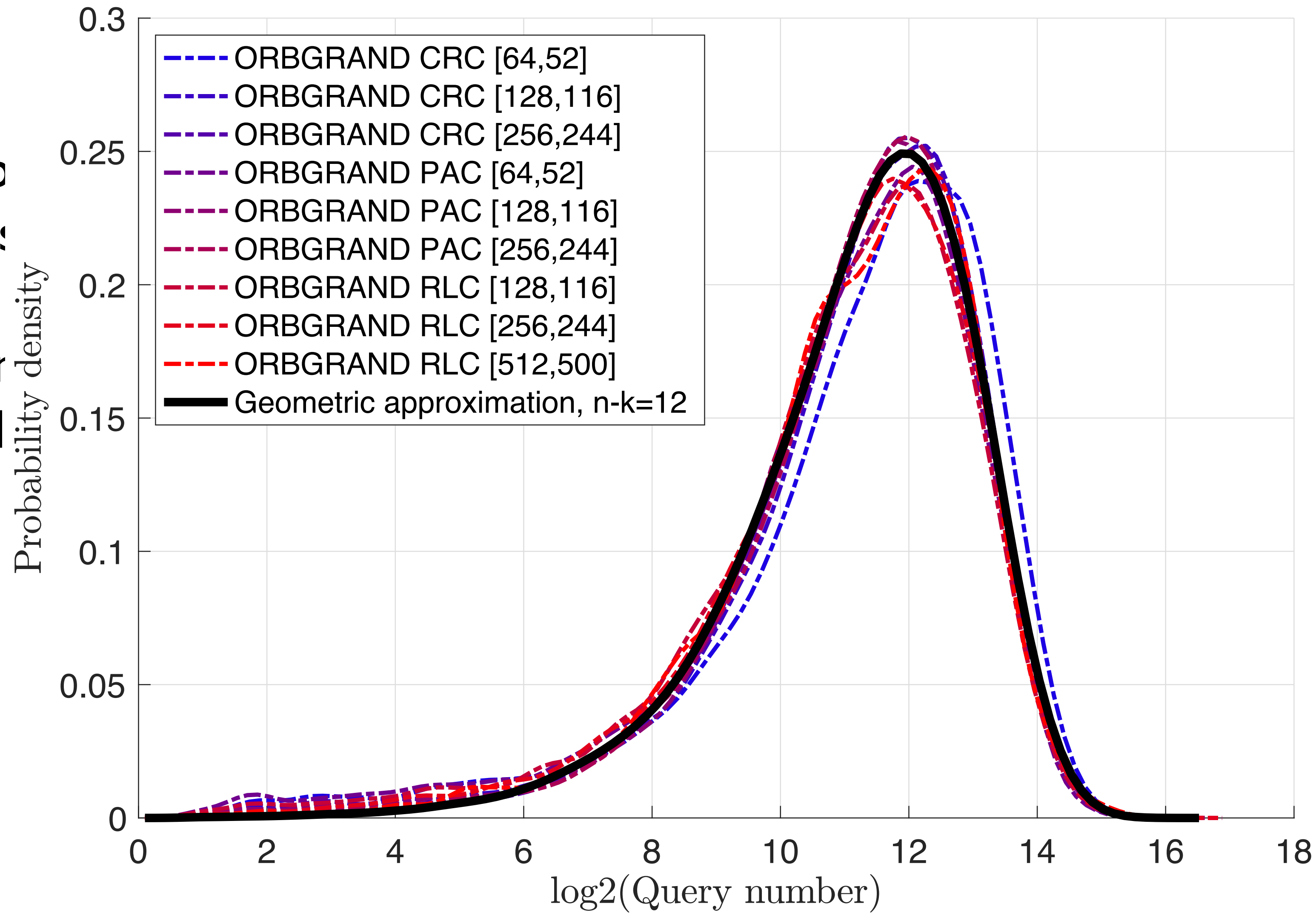
- New way of thinking enables new derivation of old results & new ones.

Number of queries to an error

Consider a ngs.

If the codeb query identifies a codeword i

Hence the r is approximately geometrical

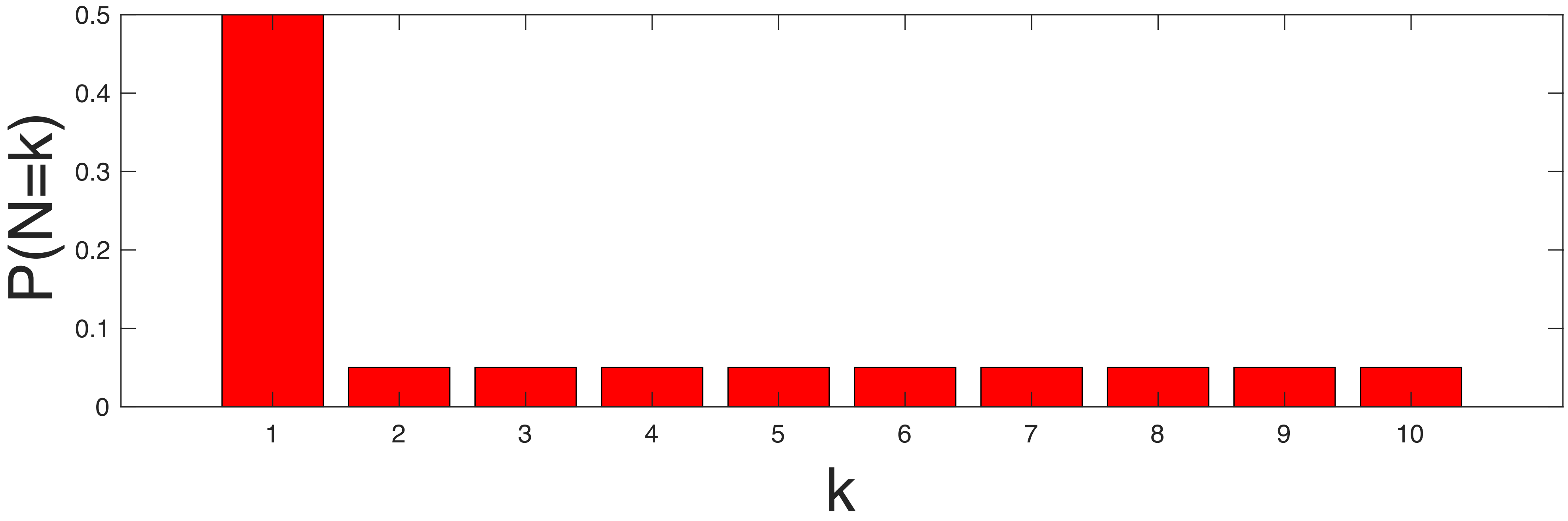


Guesswork

- Given you know the distribution from which an object is selected, Guesswork is the number of yes/no queries until a randomly selected object is identified:

$$G(z^{n,i}) \leq G(z^{n,j}) \text{ iff}$$

$$P(N^n = z^{n,i}) \geq P(N^n = z^{n,j})$$



Massey, *IEEE ISIT*, 1994

Number of queries to a correct decoding

Institute for the Wireless Internet of Things at Northeastern

Moments of # queries to correct decoding:

$$\Lambda(\alpha) = \lim_{n\to\infty} \frac{1}{n} \log E\left(G(N^n)^\alpha\right) = \begin{cases} \alpha H_{1/(1+\alpha)} & \text{if } \alpha > -1 \\ -H_\infty & \text{if } \alpha \le -1 \end{cases}$$

Probabilities of # queries to correct decoding:

$$P\left(G(N^n) \approx 2^{ng}\right) \approx \exp\left(-n \sup_\alpha \left(\alpha g - \Lambda(\alpha)\right)\right)$$

Probabilities of # queries to incorrect decoding a rate R codebook:

$$P\left(U^n \approx 2^{nu}\right) \approx \begin{cases} \exp\left(-n(1-R-u)\right) & \text{if } u \in [0, 1-R] \\ 0 & \text{otherwise} \end{cases}$$

Likelihood of error: $P(U^n \le G(N^n))$ Complexity: $\min(U^n, G(N^n))$

Arikan, *IEEE Trans. Inf. Theory*, 96. Malone & Sullivan, *IEEE Trans. Inf. Theory*, 04. Pfister & Sullivan, *IEEE Trans. Inf. Theory*, 04.

Christiansen and Duffy, *IEEE Trans. Inf. Theory*, 13.

Theorems – Channel Coding, Error Exponent

Proposition 1 *(Channel Coding Theorem With GRAND). Under Assumptions 1 and 2, with I^U defined in equation* (10) *and I^N in equation* (8)*, we have the following.*

1) If the code-book rate is less than the capacity, $R < 1-H$, then

$$\lim_{n\to\infty} \frac{1}{n} \log P(U^n \le G(N^n)) = -\inf_{a\in[H,1-R]} \{I^U(a) + I^N(a)\} < 0,$$

so that the probability that GRAND does not correctly identify the transmitted code-word decays exponentially in the block length n. If, in addition, x^* exists such that

$$\frac{d}{dx} I^N(x)|_{x=x^*} = 1, \tag{12}$$

then the error rate simplifies further to

$$\begin{aligned}\epsilon(R) &= -\lim_{n\to\infty} \frac{1}{n} \log P(U^n \le G(N^n)) \\ &= \begin{cases} 1 - R - H_{1/2} & \text{if } R \in (0, 1-x^*) \\ I^N(1-R) & \text{if } R \in [1-x^*, 1-H). \end{cases}\end{aligned} \tag{13}$$

Moreover,

$$s(R) = \lim_{n\to\infty} \frac{1}{n} \log P(U^n \ge G(N^n)) = 0$$

so that the probability that GRAND does not provide the true channel does not decay exponentially in n.

Proposition 3. *(GRANDAB Coding Theorem and Guessing Complexity). Under the assumptions of Theorems 1 and 2. If the code-book rate is less than the capacity, $R < 1-H$, then the GRANDAB error rate is*

$$\begin{aligned}&\lim_{n\to\infty} \frac{1}{n} \log P\left(\{U^n \le G(N^n)\} \cup \left\{\frac{1}{n} \log G(N^n) \ge H + \delta\right\}\right) \\ &\quad = -\min\left\{\inf_{a\in[H,1-R]} \{I^U(a) + I^N(a)\}, I^N(H+\delta)\right\} < 0,\end{aligned}$$

so that probability that the ML decoding is not the transmitted code-word decays exponentially in the block length n. If, in addition, x^* defined in equation (12) *exists then this simplifies to what we call the GRANDAB error rate*

$$\epsilon^{AB}(R) = \min\left(\epsilon(R), I^N(H+\delta)\right) \tag{21}$$

where $\epsilon(R)$ is the ML decoding error rate in equation (13)*. The expected number of guesses until GRANDAB terminates, $\{D^n_{AB}\}$, satisfies*

$$\lim_{n\to\infty} \frac{1}{n} \log E(D^n_{AB}) = \min\left(H_{1/2}, 1-R, H+\delta\right).$$

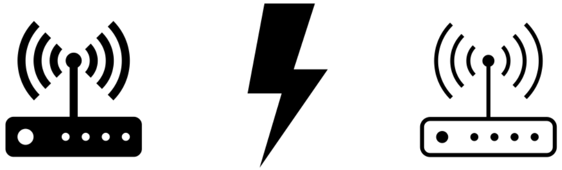
For rates above capacity, $R > 1-H$, the success probability is identical to that for ML decoding, given in equation (14).

Duffy, Li, Médard, *IEEE Tran. Inf. Theory*, 19.

Decoding with soft information

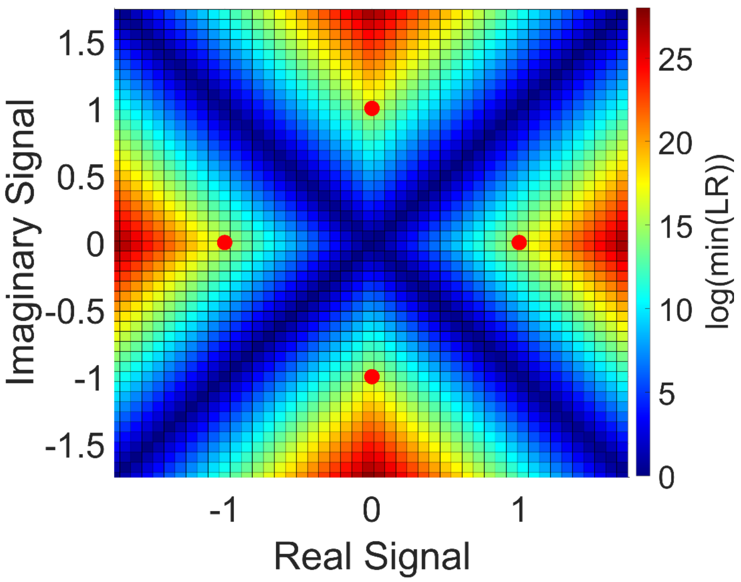
Institute for the Wireless Internet of Things at Northeastern

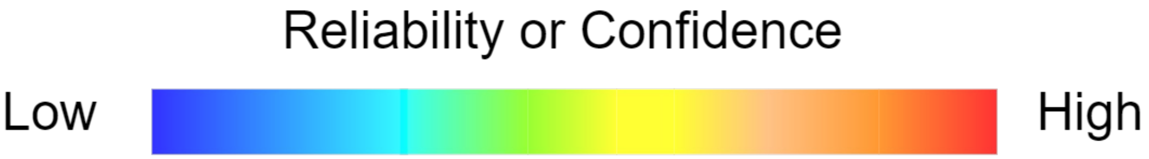
01100010 →

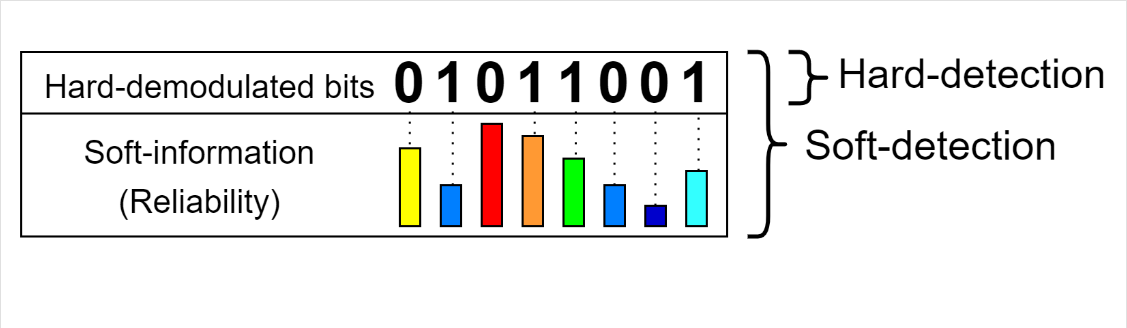


11100011

Soft Detection







Duffy, An, Médard, *IEEE Trans. Sig. Process.*, 23. Duffy, Médard, An, *IEEE Trans. Commun.*, 21. Duffy, *IEEE ICASSP*, 21. Solomon, Duffy, Médard, *IEEE ICC*, 20.

Bits are flipped independently?

Because they're engineered to be so to match decoder expectations

Collect data as rows:

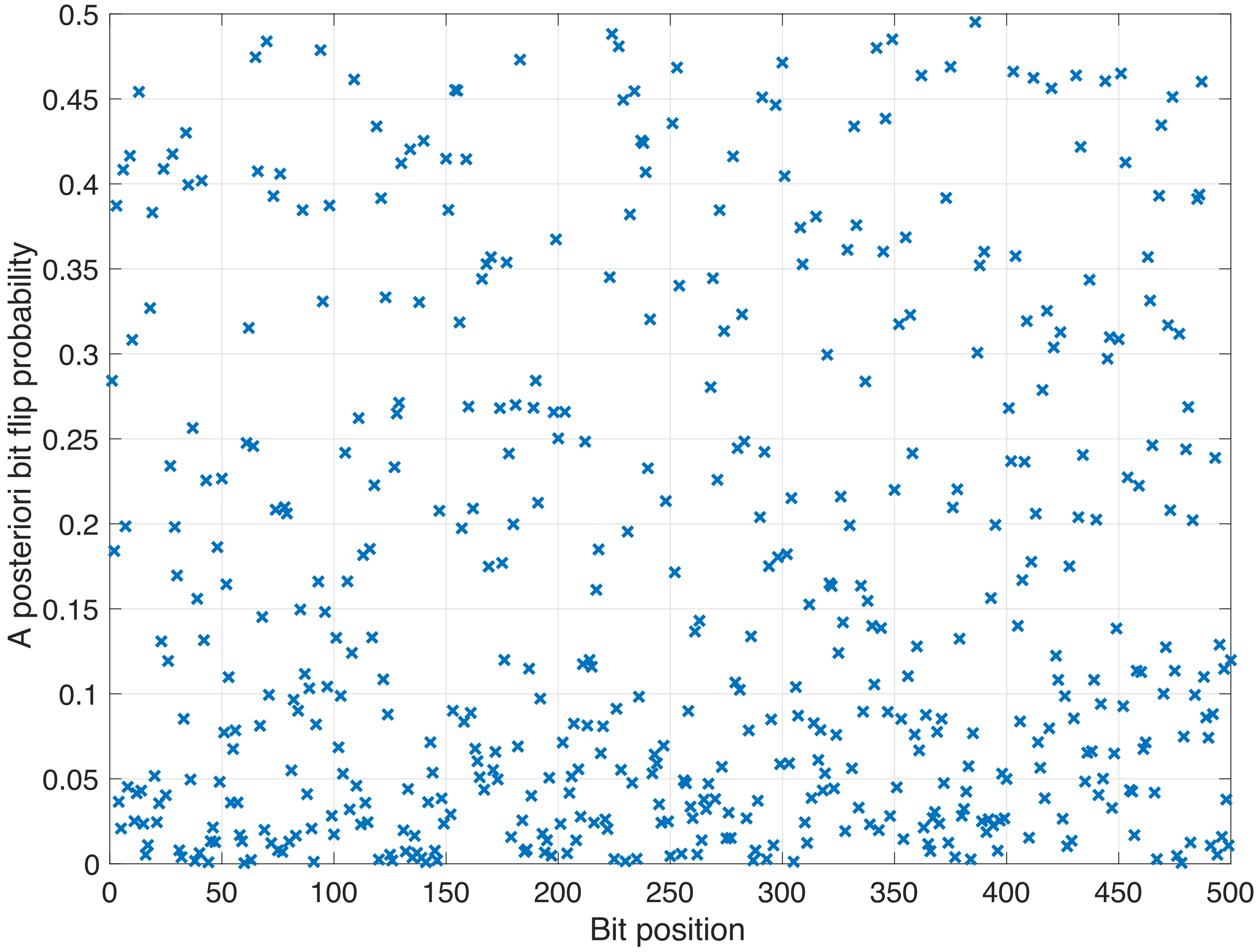
$$\left(\begin{array}{cccccc} c_{1,1} & c_{1,2} & c_{1,3} & \cdots & c_{1,n-1} & c_{1,n} \\ \hline c_{2,1} & c_{2,2} & c_{2,3} & \cdots & c_{2,n-1} & c_{2,n} \\ \hline \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \hline c_{n,1} & c_{n,2} & c_{n,3} & \cdots & c_{n,n-1} & c_{n,n} \end{array}\right)$$

Transmit as columns:

$$\left(\begin{array}{c|c|c|c|c|c} c_{1,1} & c_{1,2} & c_{1,3} & \cdots & c_{1,n-1} & c_{1,n} \\ c_{2,1} & c_{2,2} & c_{2,3} & \cdots & c_{2,n-1} & c_{2,n} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ c_{n,1} & c_{n,2} & c_{n,3} & \cdots & c_{n,n-1} & c_{n,n} \end{array}\right)$$

A posteriori bit flip probabilities

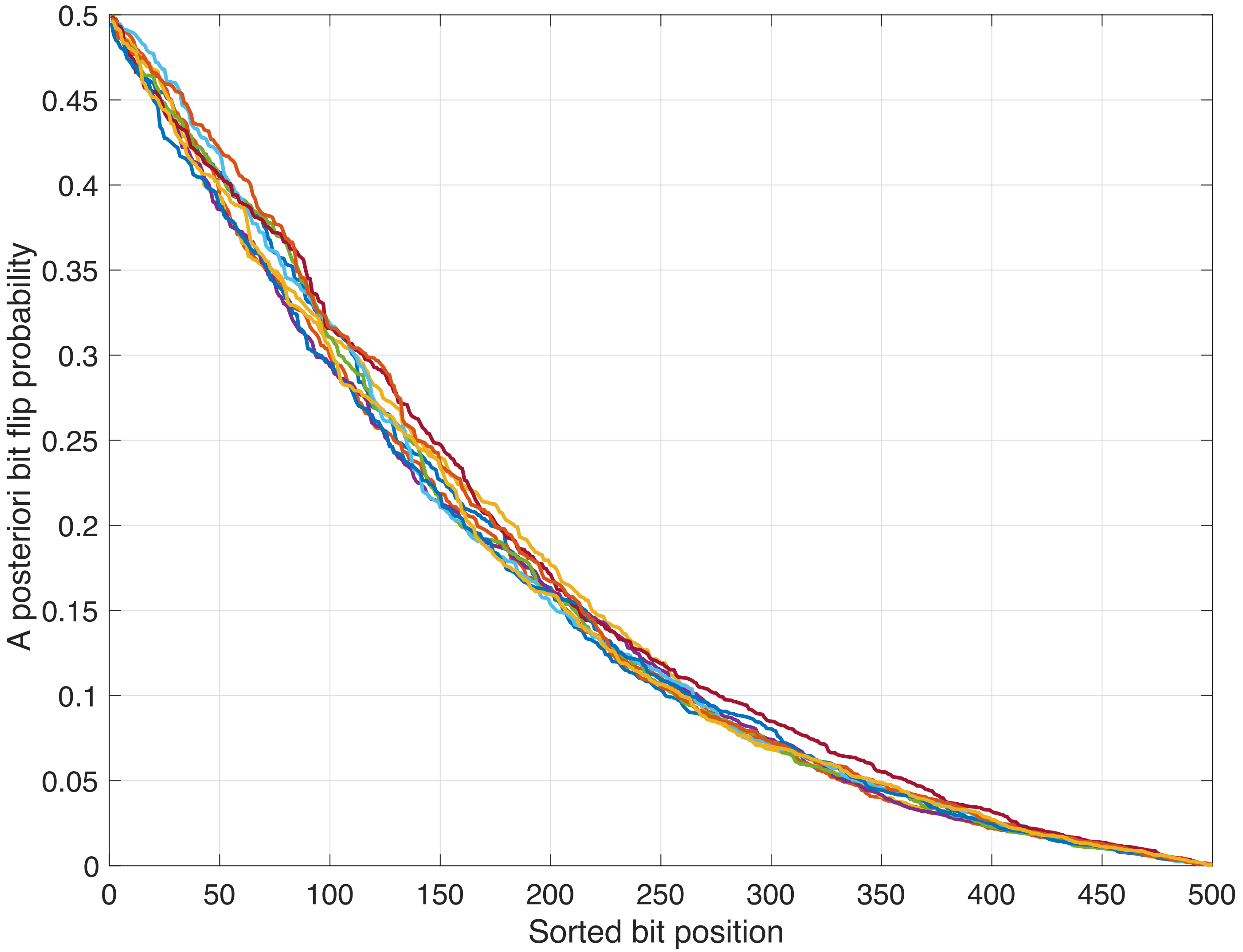
Standard interleaved channel model:
Additive White Gaussian Noise



Rank ordered reliabilities

Institute for the Wireless Internet of Things at Northeastern

- For each set of received reliabilities, rank order from least reliable to most.
- Consistency across different samples for the same reasons empirical cumulative distribution functions converge

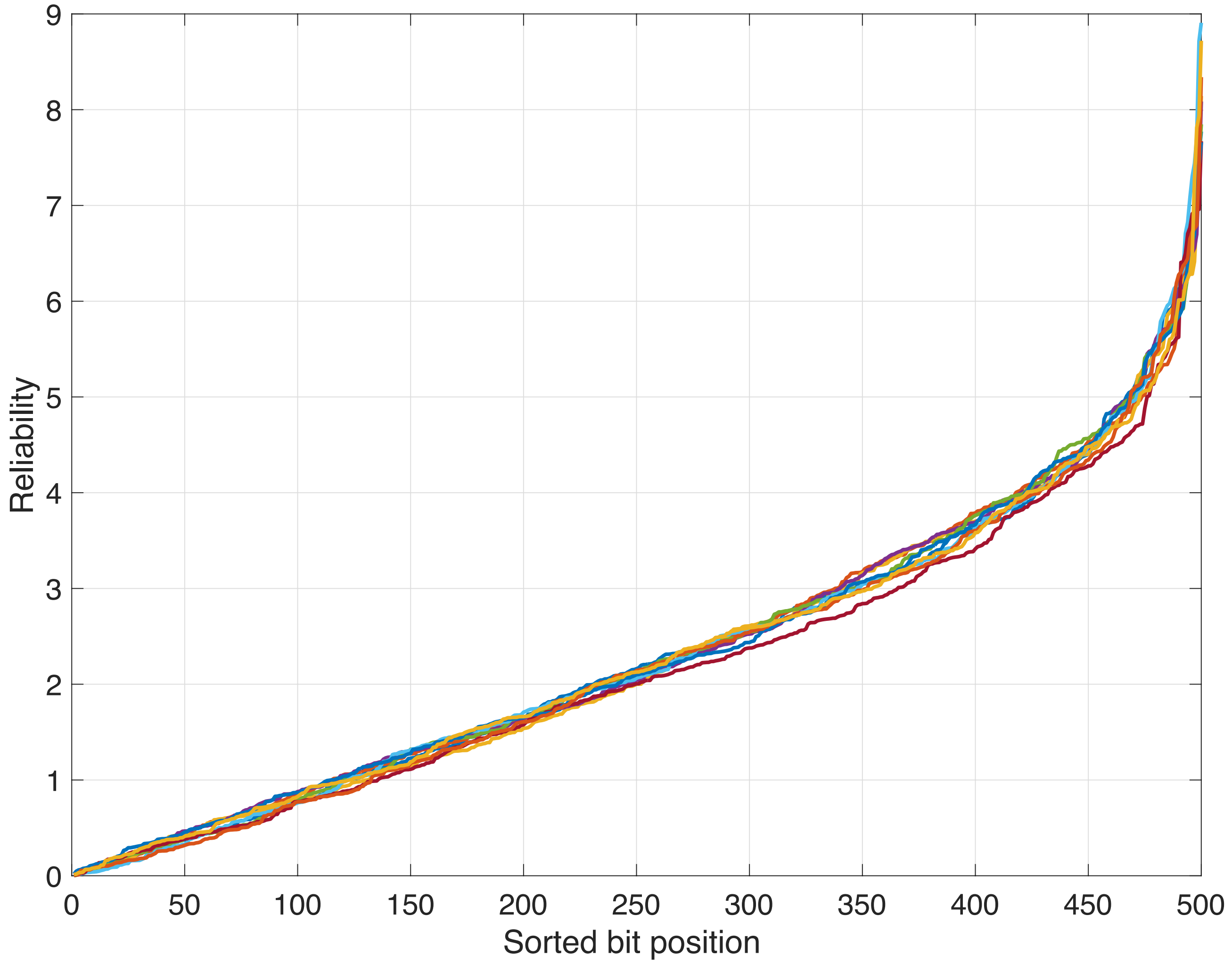


Rank ordered reliabilities

Institute for the Wireless Internet of Things at Northeastern

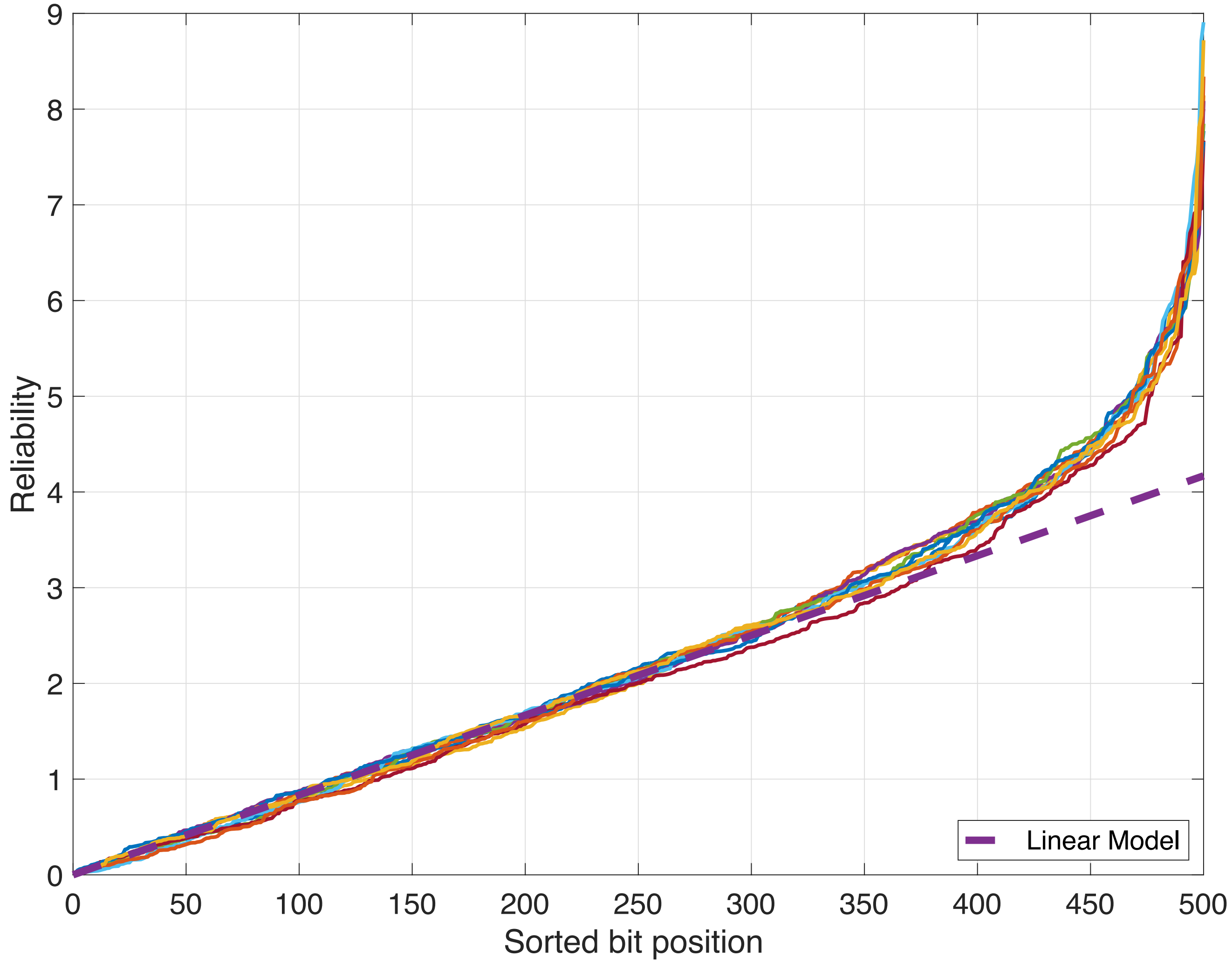
- Standard to not think of probabilities, but an invertible transformation.
- Reliability is the absolute value of the log likelihood ratio of the hypotheses that a bit is a 1 or a 0

$$|\mathrm{LLR}| = \log\left(\frac{1-p}{p}\right)$$



Rank ordered reliabilities

- Put your statistical modelling hat on
- I.e. it's a line



Rank ordered reliabilities

Institute for the Wireless Internet of Things at Northeastern

$$\mathbb{P}(N^n = z^n) = \prod_{i=1}^{n}(1-p_i) \prod_{i:z_i=1} \frac{p_i}{1-p_i}$$

$$\mathbb{P}(N^n = z^n) \propto \prod_{i:z_i=1} \frac{p_i}{1-p_i} = e^{-\sum_{i:z_i=1} |\mathrm{LLR}_i|}$$

If, rank ordered from least reliable

$$|\mathrm{LLR}_i| \approx \beta i \text{ for } i = 1, \ldots, n$$

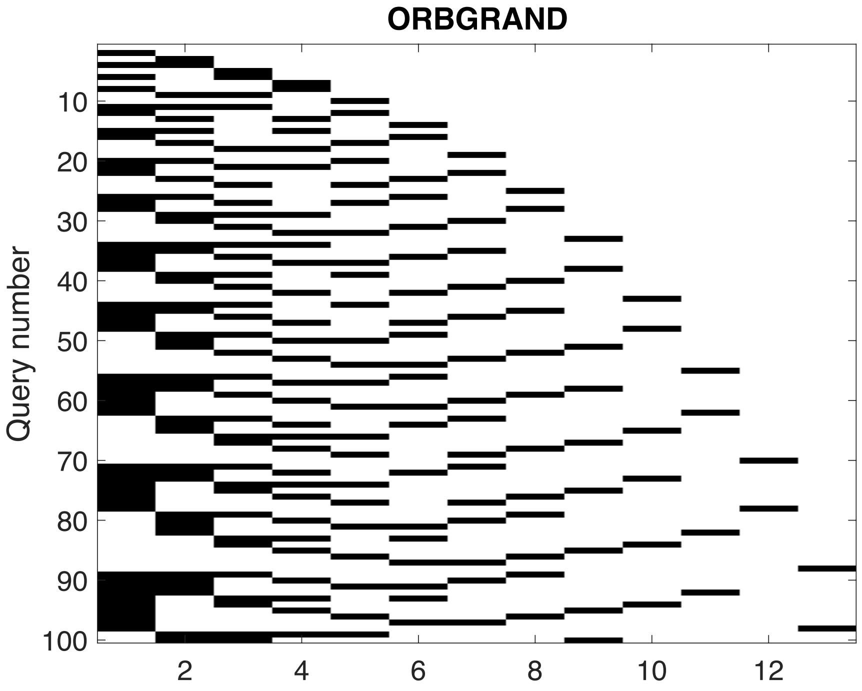
then

$$\sum_{i:z_i=1} |\mathrm{LLR}_i| = \beta \sum_{i=1}^{n} i z_i \propto \sum_{i=1}^{n} i z_i = w_{\mathrm{L}}(z^n)$$

and sequences rank ordered by logistic weight.

Ordered Reliability Bits GRAND

Institute for the Wireless Internet of Things at Northeastern



Decreasing likelihood of noise effects = increasing Logistic Weight (sum of rank-ordered position of bits flipped)

$$w_{\mathrm{L}}(z^n) = \sum_{i=1}^{n} i z_i$$

Once bits are rank ordered, ORBGRAND uses a fixed guessing order

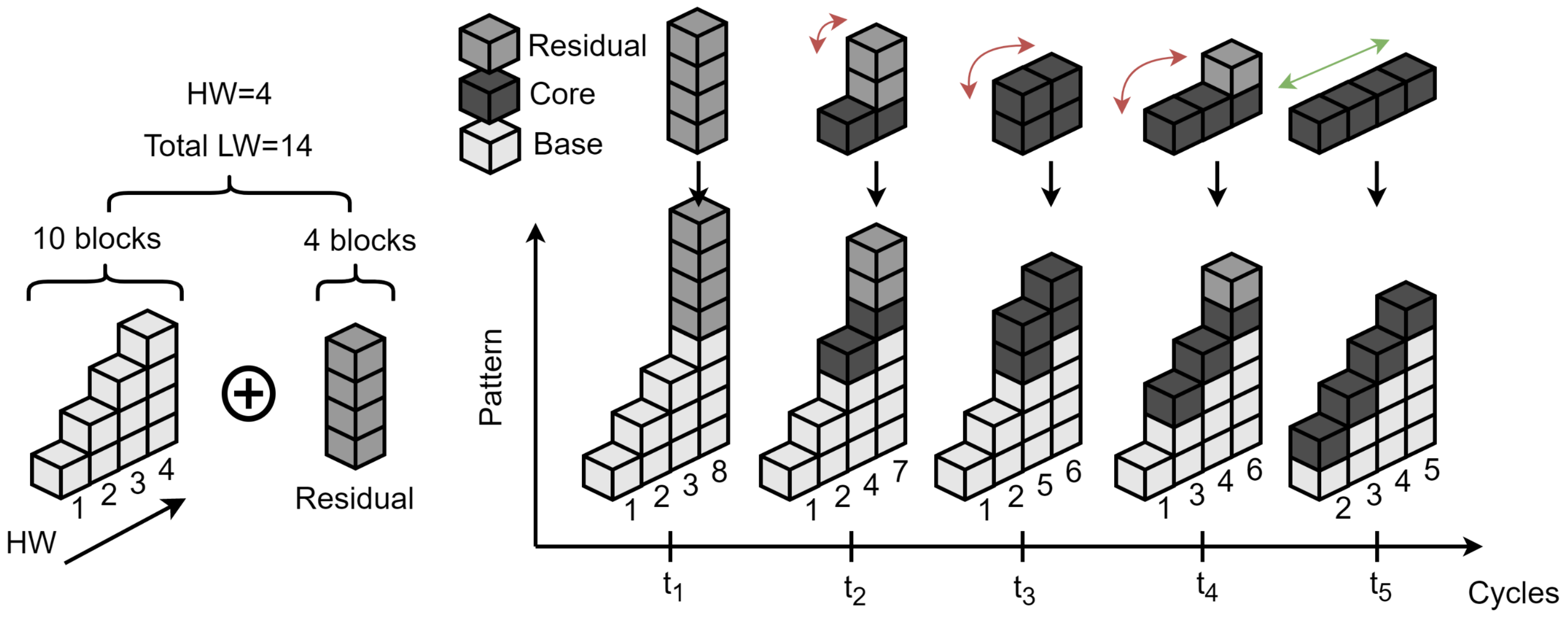
Generating patterns for a given logistic weight corresponds to solving an integer partition problem: sum of distinct integers (bit flip positions), each no greater than n, that add to given value.

Duffy, An, Medard, *IEEE Trans. Signal Proc.*, 22. Duffy, *IEEE ICASSP*, 21.
Liu, Wei, Chen, Zhan, *IEEE Trans. Info. Theory*, 23.

Rank ordered reliabilities

Institute for the Wireless Internet of Things at Northeastern

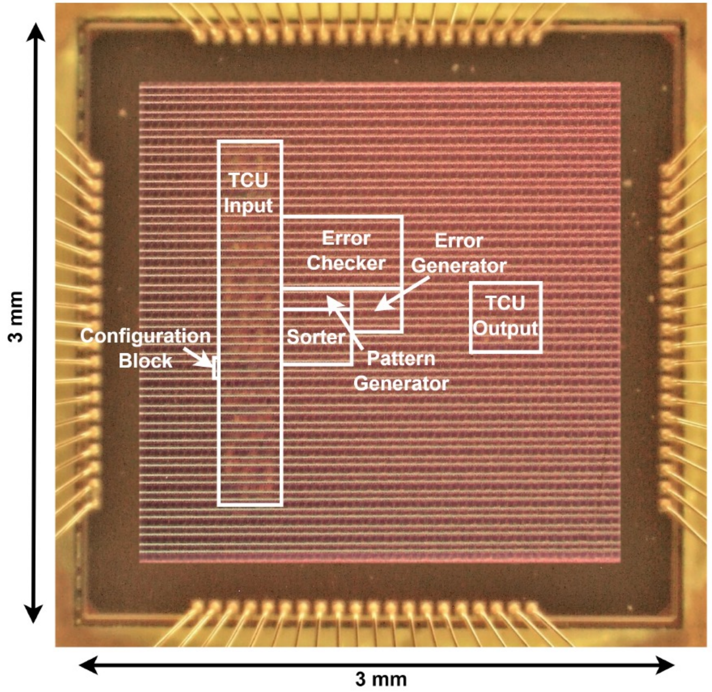
Logistic weight: $w_{\mathrm{L}}(z^n) = \sum_{i=1}^{n} i z_i$ Hamming weight: $w_{\mathrm{H}}(z^n) = \sum_{i=1}^{n} z_i$

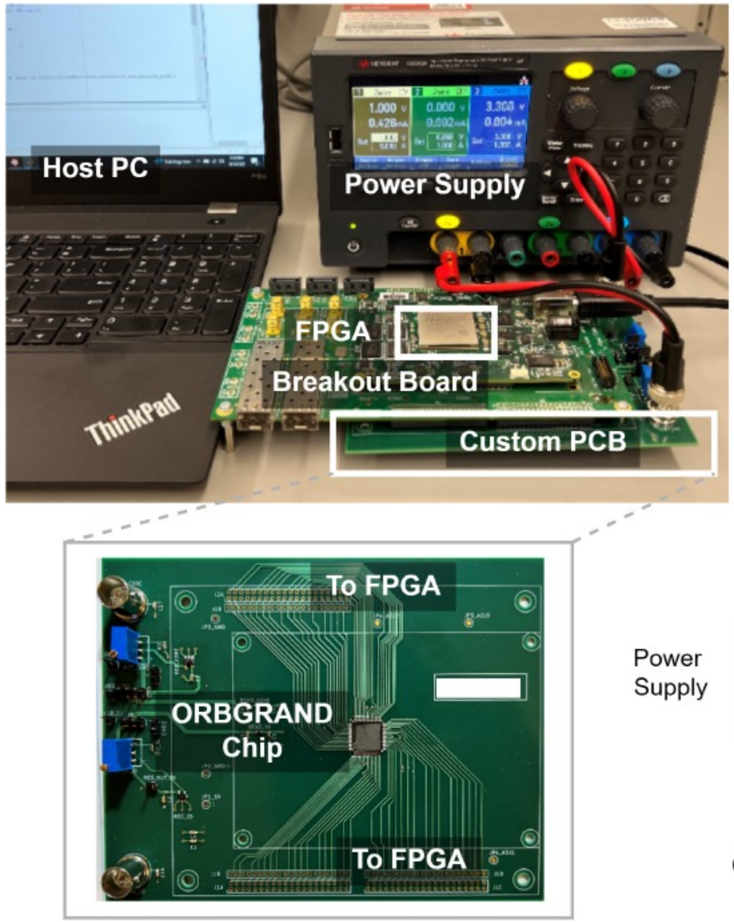


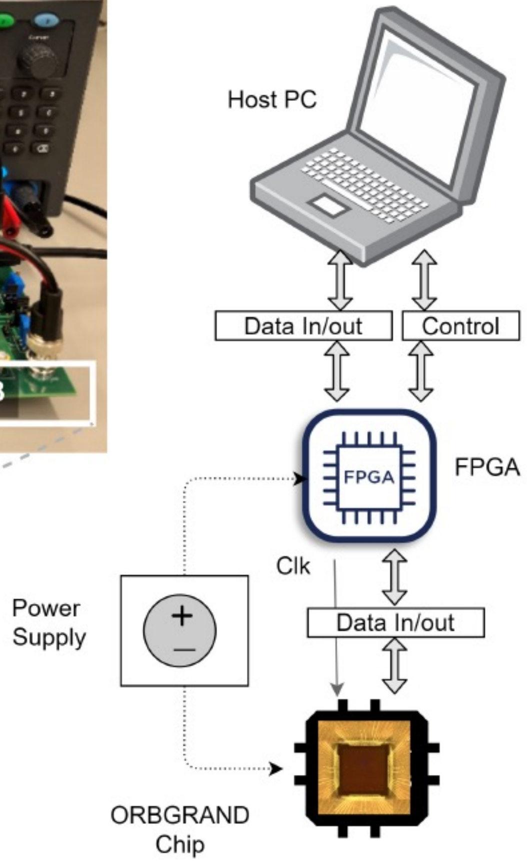
Duffy, An, Medard, *IEEE Trans. Signal Proc., 22.*

ORBGRAND in hardware

Institute for the Wireless Internet of Things at Northeastern







Riaz, Yasar, Ercan, An, Ngo, Galligan, Médard, Duffy, Yazicigil, *IEEE ISSCC,* 23.

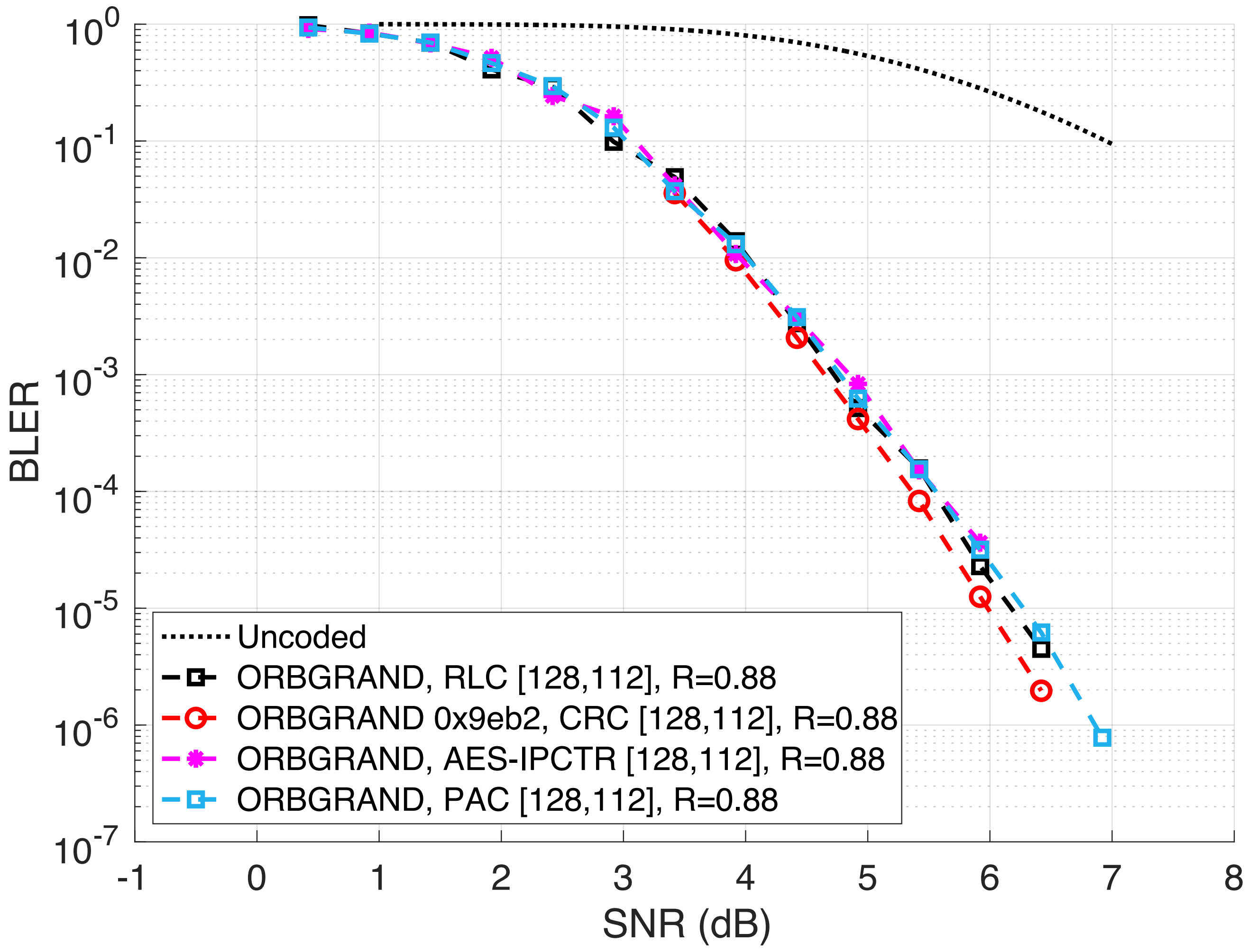
Other circuits designs, e.g. : Abbas, Tonnellier, Ercan, Jalaleddine, Gross, *IEEE Trans. VLSI*, 22. Condo, *IEEE Trans Circuits Syst*, 21. Condo, Bioglio, Land, IEEE Globecom, 21

ORBGRAND code performance

Institute for the Wireless Internet of Things at Northeastern

Block Error Rate (BLER) – fraction of blocks decoded incorrectly vs. Signal-to-Noise Ratio (SNR)

Binary phase shift keying modulation and additive white Gaussian noise



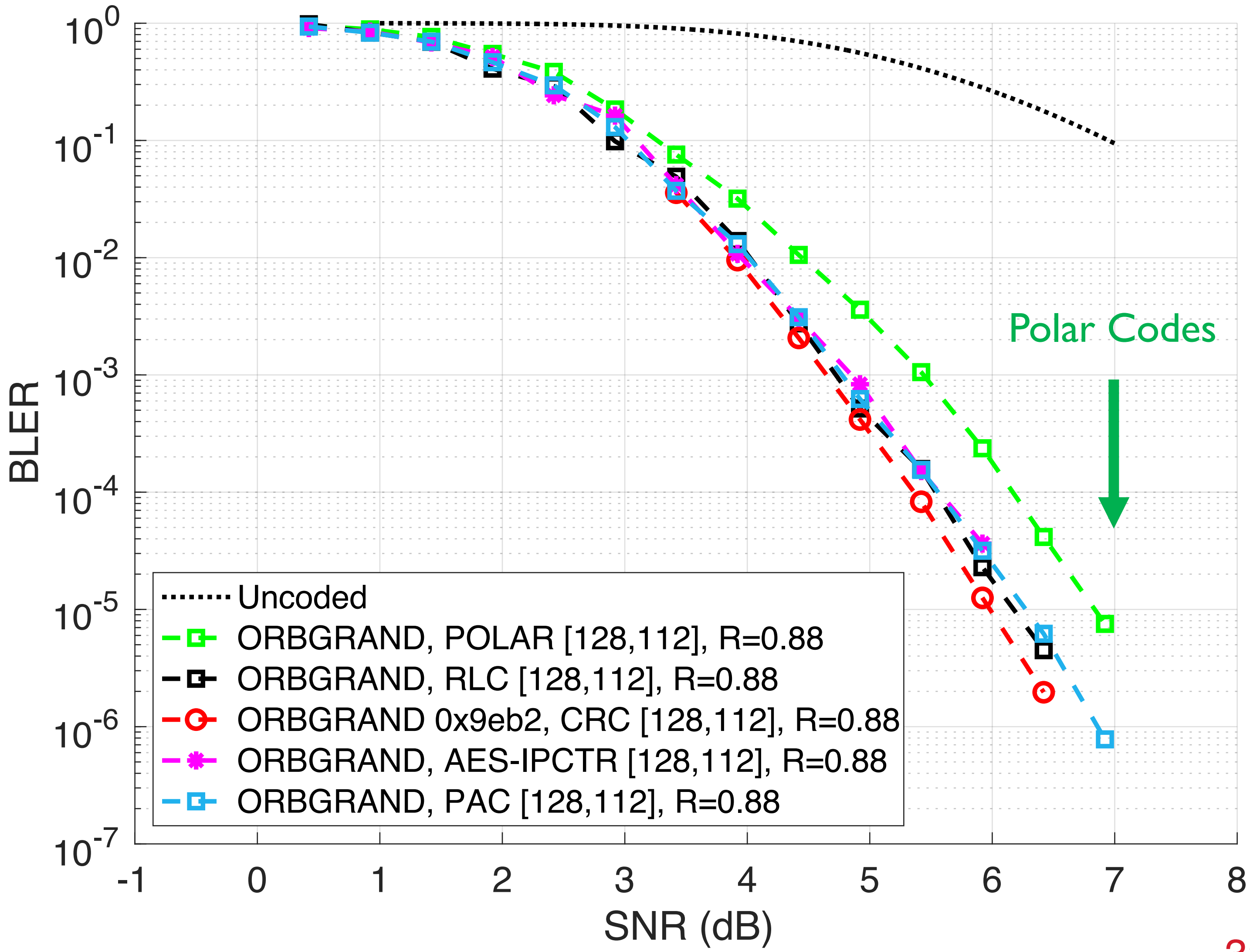
Coffey, Goodman, *IEEE Trans. Inf. Theory*, 90. Koopman, https://users.ece.cmu.edu/~koopman/crc/. Cohen, D'Oliverira, Duffy, Woo, Médard, *IEEE Comun. Lett.*, 23. Arikan, *arXiv*, 19.

ORBGRAND code performance

Institute for the Wireless Internet of Things at Northeastern

Block Error Rate (BLER) – proportion of blocks decoded incorrectly vs. Signal-to-Noise Ratio (SNR)

Most celebrated recent code construction almost uniquely underperforms



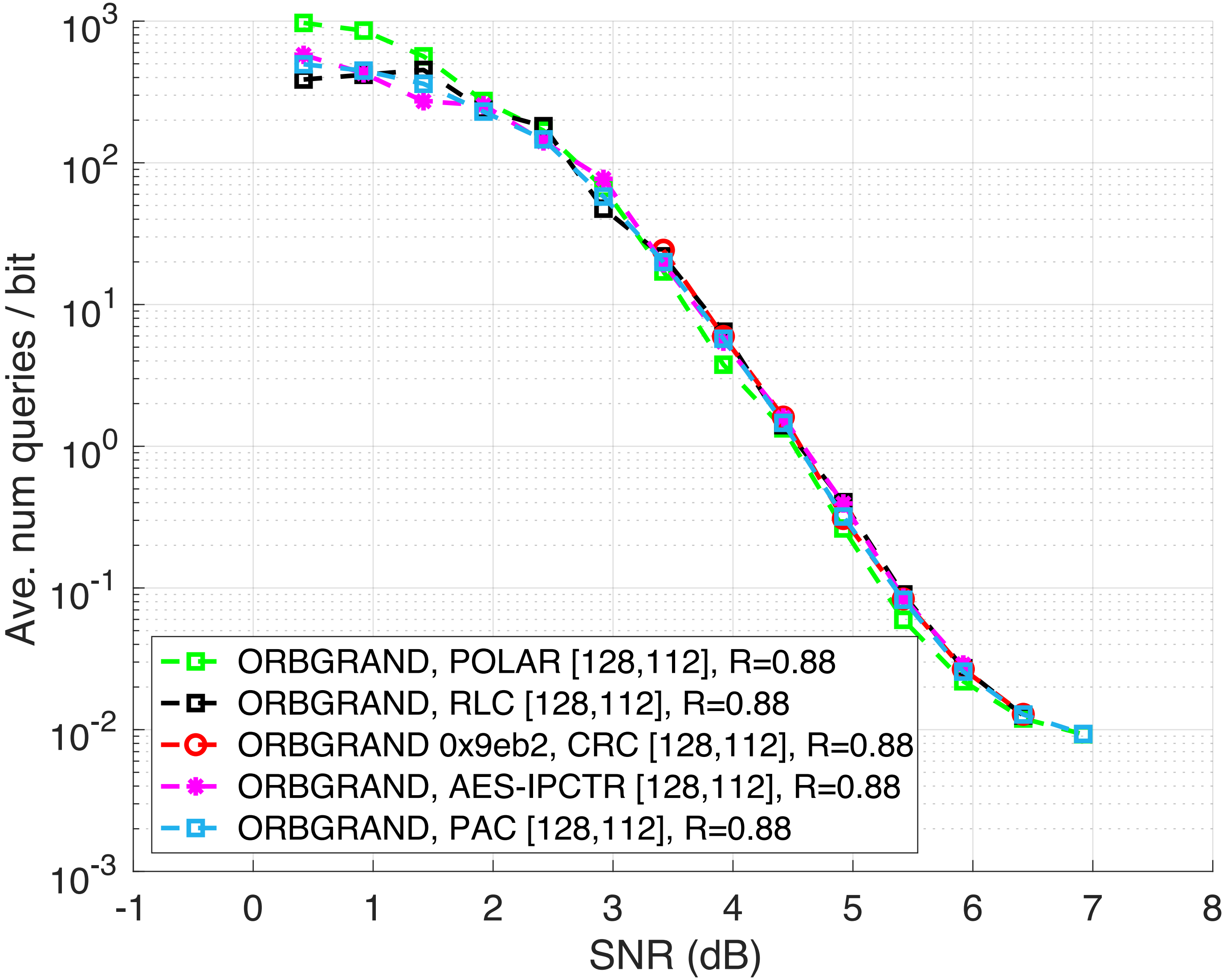
Arikan, *IEEE Trans. Inf. Theory*, 09.

ORBGRAND decoding complexity

Institute for the Wireless Internet of Things at Northeastern

Guesswork vs. Signal-to-Noise Ratio (SNR)

A measure of decoding complexity

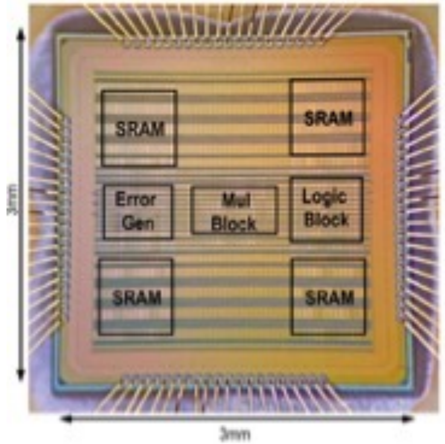


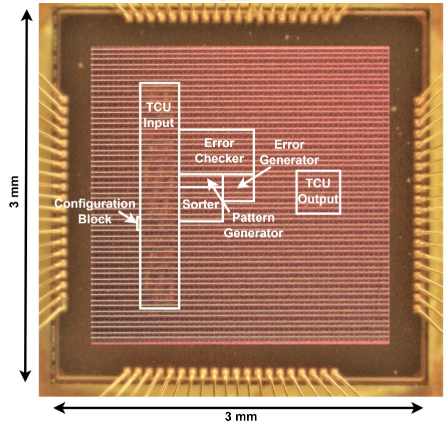
Duffy, An, Medard, *IEEE Trans. Signal Proc., 22.*

GRAND

Institute for the Wireless Internet of Things at Northeastern

- Decodes **any** moderate redundancy code, of any length, with **max accuracy**.
- **Hard** and **soft** detection variants have been developed.
- Inherently **highly parallelizable**, resulting **in low latency**.
- In silicon prototypes establish **energy efficiency**.
- **Uniquely** provides an accurate estimate of the likelihood of correct decoding.
- **Only** universal decoder that decode **in channels with correlated noise**.
- Essentially **all** long, low-rate codes are composed of smaller components and **GRAND** is being developed for use with them.
- Offers a **single**, **energy efficient, precise** decoder for a **broad swathe** of codes with a **small footprint**.
- Much more to come, **in practice and in theory** (with epsilontics)…





GRAND

granddecoder.mit.edu

Massachusetts Institute of Technology

Northeastern University

BOSTON UNIVERSITY

NU Math (will be) hiring

Quantum Information Science

Tenure Track / Tenure Open Rank Professorship

Details will be available soon at:

https://hr.northeastern.edu/careers/job-listings/