

Information-theoretic perspective on massive multiple-access

Yury Polyanskiy

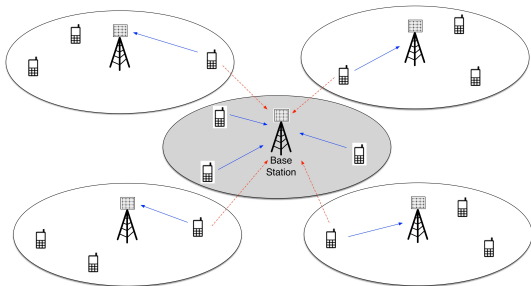
Department of EECS
MIT

`yp@mit.edu`

North American School of Info Th., Jun. 2018

Legal notice: Some images in this presentation are borrowed from publicly available sources. The copyright on these images belongs to their original creators. For full copyright information please contact the author.

How does your cell phone work?



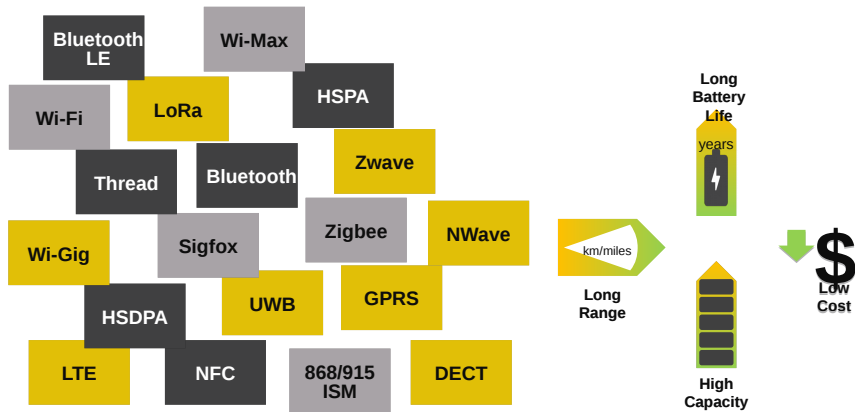
- Cell phone is powered on.
- Announces its presence on PRACH.
- Base station (periodically) gives permission to send.
- Summary:
 - ▶ Random-Access is very low duty cycle.
 - ▶ BS makes access **ORTHOGONAL** across users
 - ▶ bulk of communication is over an interference-free single-user AWGN.
- What's new in 5G?



- Smart Agriculture
- Advanced Metering systems
- Fire alarms
- Home security and automation
- Oilfield and pipeline monitoring
- M-health
- Smart parking, intelligent traffic
- Waste and recycling
- Asset tracking and geo-location
- Animal tracking and livestock

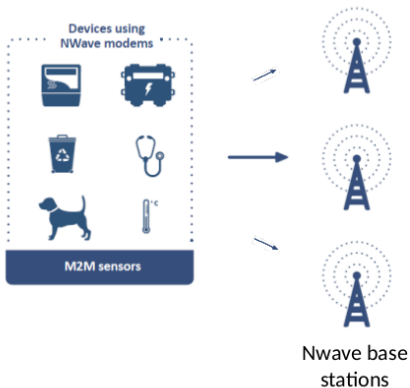
Expected density: 100-500 devices per household/office

Soup of solutions

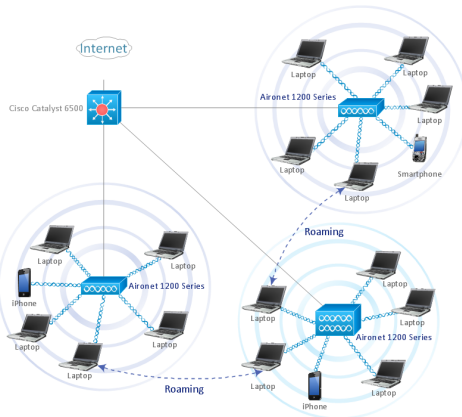


Two breeds of IoT

LPWAN



One basestation covers 10 km





Q: What drains the battery? Examples (@ 3.3V):

	Arduino (w/o reg.)	XBee (Zigbee)	LP-WAN sensor
Sleep	5 μ A	1 μ A	1-2 μ A
CPU Running	50 μ A	40 μ A	60 μ A

IoT is about battery life



Q: What drains the battery? Examples (@ 3.3V):

	Arduino (w/o reg.)	XBee (Zigbee)	LP-WAN sensor
Sleep	5 μ A	1 μ A	1-2 μ A
CPU Running	50 μ A	40 μ A	60 μ A
Radio Xmit		40 mA	20 mA

IoT is about battery life



Q: What drains the battery? Examples (@ 3.3V):

	Arduino (w/o reg.)	XBee (Zigbee)	LP-WAN sensor
Sleep	5 μ A	1 μ A	1-2 μ A
CPU Running	50 μ A	40 μ A	60 μ A
Radio Xmit		40 mA	20 mA

- Duty-cycle of 1 sec / 20 min radio lasts 6-10 yr / AA bat.
- **Caveat:** Calculation assumes **single-user**
- **Key problem:** Energy usage will grow with # of sensors deployed.
How much?
- **Sad:** depends on technology? **Happy:** IT comes to rescue!

Envisioned solution:

- To save battery: sensors sleep all the time, except transmissions.
- ... uncoordinated transmissions.
- ... they wake up, blast the packet, go back to sleep.
- Focus on low-energy (low E_b/N_0)
- Focus on fundamental limits
- ... but with low-complexity solutions (single-user-only decoding).

Envisioned solution:

- To save battery: sensors sleep all the time, except transmissions.
- ... uncoordinated transmissions.
- ... they wake up, blast the packet, go back to sleep.
- Focus on low-energy (low E_b/N_0)
- Focus on fundamental limits
- ... but with low-complexity solutions (single-user-only decoding).

Issues we need to understand:

- ① packets are short: finite-blocklength (FBL) info theory
- ② multiple-access channel: Classical MAC
- ③ low-complexity MAC: modulation, CDMA, multi-user detection
- ④ massive random-access: many users, same-codebook codes (NEW)

Envisioned solution:

- To save battery: sensors sleep all the time, except transmissions.
- ... uncoordinated transmissions.
- ... they wake up, blast the packet, go back to sleep.
- Focus on low-energy (low E_b/N_0)
- Focus on fundamental limits
- ... but with low-complexity solutions (single-user-only decoding).

Issues we need to understand:

- ① packets are short: finite-blocklength (FBL) info theory
- ② multiple-access channel: Classical MAC
- ③ low-complexity MAC: modulation, CDMA, multi-user detection
- ④ massive random-access: many users, same-codebook codes (NEW)

Supporting 10 users at 1Mbps is much easier than 1M users at 10bps.

FBL Info Theory: short intro

Case study: 1000-bit BSC

- Consider channel $BSC(n = 1000, \delta = 0.11)$
- How many data bits can we transmit with (block) $P_e \leq 10^{-3}$?
- Attempt 1: Repetition

$$k = 47 \text{ bits via } [21,1,21]\text{-code}$$

- Attempt 2: Reed-Muller

$$k = 112 \text{ bits via } [64,7,32]\text{-code}$$

- Shannon's prediction: $C = 0.5$ bit so

$$k \approx 500 \text{ bit}$$

Case study: 1000-bit BSC

- Consider channel $BSC(n = 1000, \delta = 0.11)$
- How many data bits can we transmit with (block) $P_e \leq 10^{-3}$?
- Attempt 1: Repetition

$$k = 47 \text{ bits via } [21,1,21]\text{-code}$$

- Attempt 2: Reed-Muller

$$k = 112 \text{ bits via } [64,7,32]\text{-code}$$

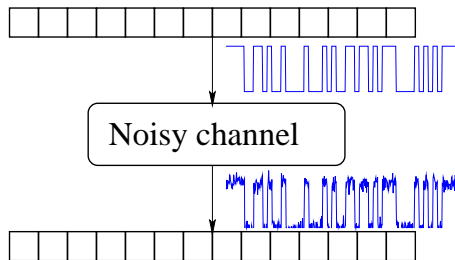
- Shannon's prediction: $C = 0.5$ bit so

$$k \approx 500 \text{ bit}$$

- Finite blocklength IT:

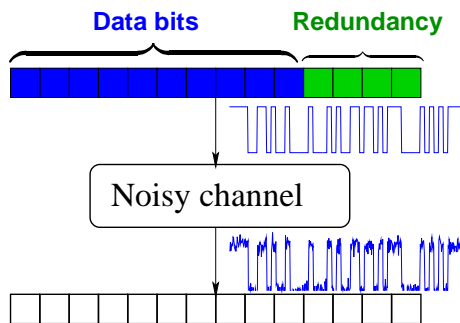
$$414 \leq k \leq 416$$

Abstract communication problem



Goal: Decrease corruption of data caused by noise

Channel coding: principles

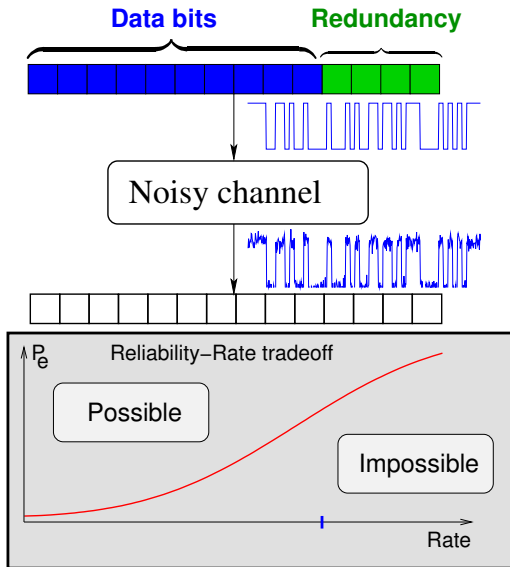


Goal: Decrease corruption of data caused by noise

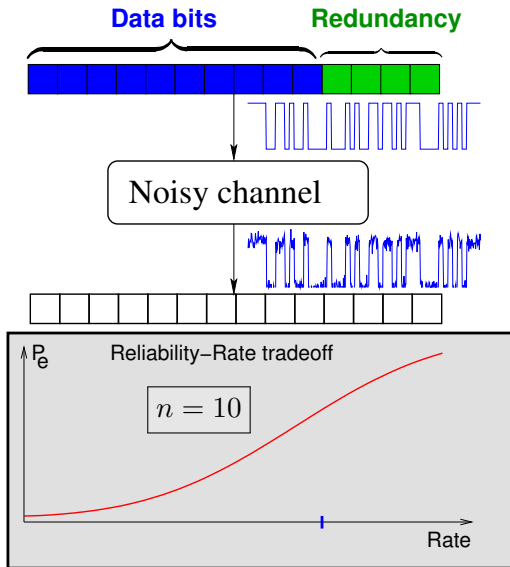
Solution: Code to diminish probability of error P_e .

Key metrics: Rate and P_e

Channel coding: principles



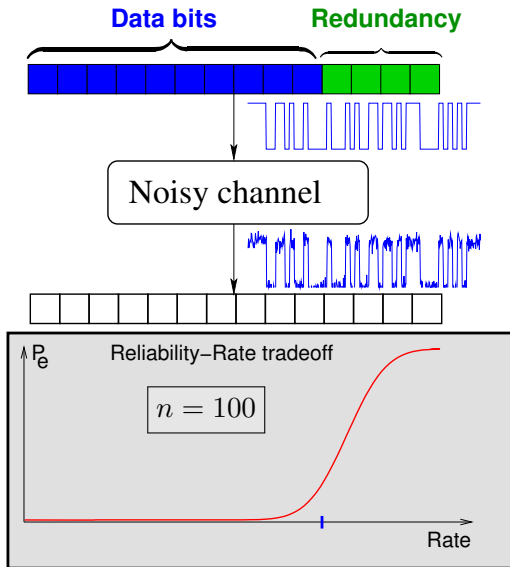
Channel coding: principles



Decreasing P_e further:

1. More redundancy
Bad: loses rate
2. Increase blocklength!

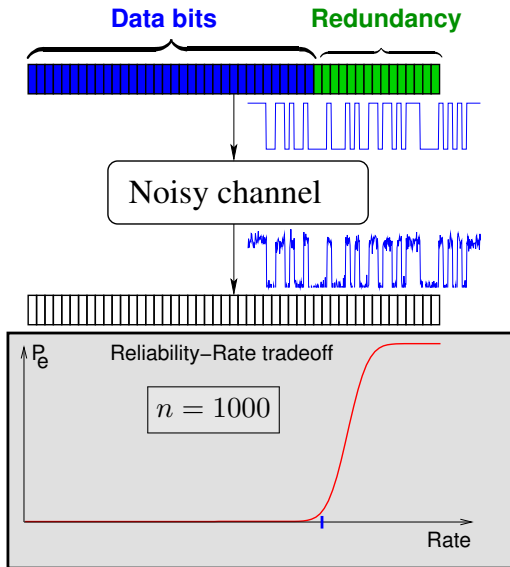
Channel coding: principles



Decreasing P_e further:

1. More redundancy
Bad: loses rate
2. Increase blocklength!

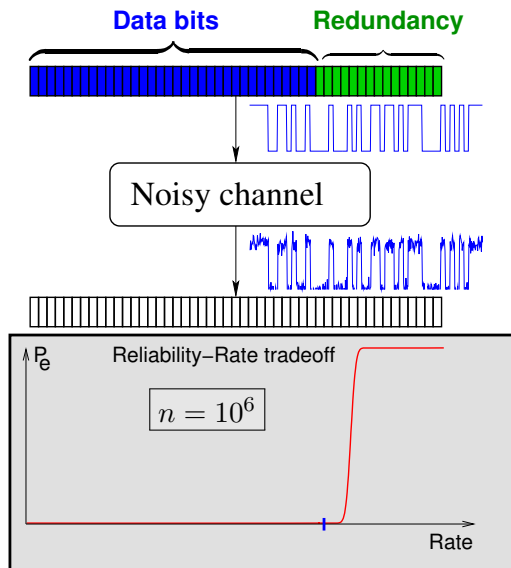
Channel coding: principles



Decreasing P_e further:

1. More redundancy
Bad: loses rate
2. Increase blocklength!

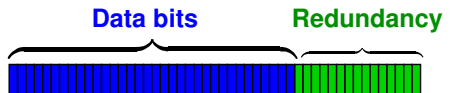
Channel coding: principles



Decreasing P_e further:

1. More redundancy
Bad: loses rate
2. Increase blocklength!

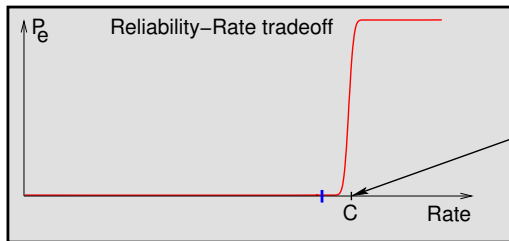
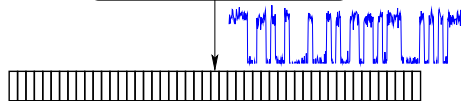
Channel coding: Shannon capacity



Shannon: Fix $R < C$

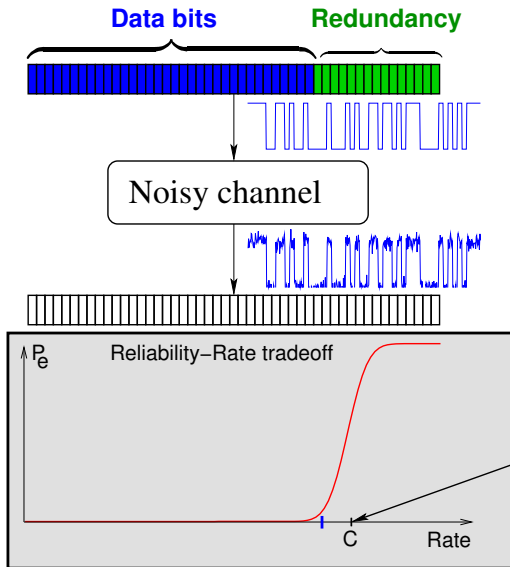
$$P_e \searrow 0 \text{ as } n \rightarrow \infty$$

Noisy channel



Channel capacity

Channel coding: Shannon capacity



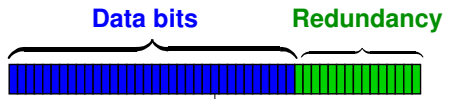
Shannon: Fix $R < C$

$$P_e \searrow 0 \text{ as } n \rightarrow \infty$$

Question:

For what n will $P_e < 10^{-3}$?

Channel coding: Gaussian approximation



Noisy channel

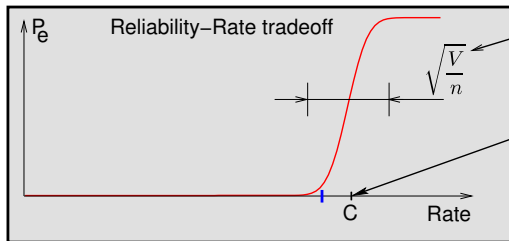


Shannon: Fix $R < C$

$$P_e \searrow 0 \text{ as } n \rightarrow \infty$$

Question:

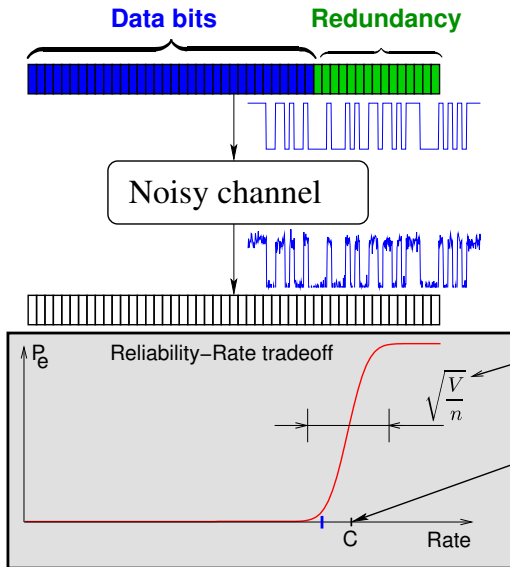
For what n will $P_e < 10^{-3}$?



Channel dispersion

Channel capacity

Channel coding: Gaussian approximation



Shannon: Fix $R < C$

$$P_e \searrow 0 \text{ as } n \rightarrow \infty$$

Question:

For what n will $P_e < 10^{-3}$?

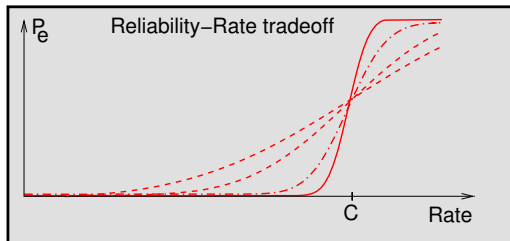
Answer:

$$n \gtrsim \text{const} \cdot \frac{V}{C^2}$$

Channel dispersion

Channel capacity

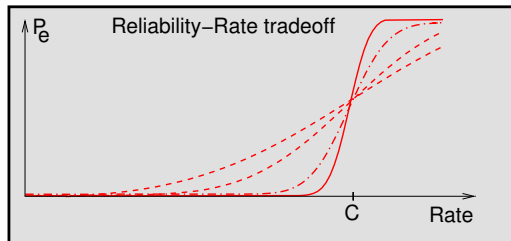
How to describe evolution of the boundary?



Classical results:

- **Vertical asymptotics:** fixed rate, reliability function
Elias, Dobrushin, Fano, Shannon-Gallager-Berlekamp
- **Horizontal asymptotics:** fixed ϵ , strong converse, \sqrt{n} terms
Wolfowitz, Weiss, Dobrushin, Strassen, Kemperman

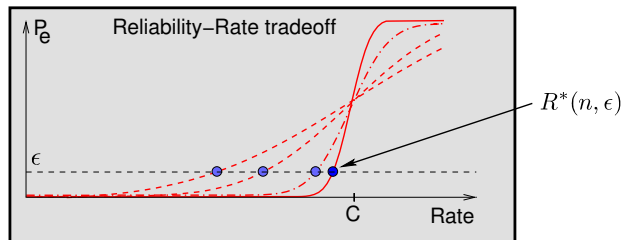
How to describe evolution of the boundary?



XXI century:

- Tight non-asymptotic bounds
- Remarkable precision of normal approximation
- Extended results on *horizontal* asymptotics
AWGN, $O(\log n)$, cost constraints, feedback, etc.

Finite blocklength fundamental limit



Definition

$$R^*(n, \epsilon) = \max \left\{ \frac{1}{n} \log M : \exists (n, M, \epsilon)\text{-code} \right\}$$

(max. achievable rate for blocklength n and prob. of error ϵ)

Rough summary: For ergodic channels

$$R^*(n, \epsilon) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon).$$

- Let $P_{Y^n|X^n} = P_{Y|X}^n$ be memoryless.
- Converse bounds (roughly):

$$R^*(n, \epsilon) \lesssim \epsilon\text{-th quantile of } \frac{1}{n} \log \frac{dP_{Y^n|X^n}}{dQ_{Y^n}}$$

- Achievability bounds (roughly):

$$R^*(n, \epsilon) \gtrsim \epsilon\text{-th quantile of } \frac{1}{n} \log \frac{dP_{Y^n|X^n}}{dQ_{Y^n}}$$

- Let $P_{Y^n|X^n} = P_{Y|X}^n$ be memoryless.
- Converse bounds (roughly):

$$R^*(n, \epsilon) \lesssim \epsilon\text{-th quantile of } \frac{1}{n} \log \frac{dP_{Y^n|X^n}}{dQ_{Y^n}}$$

- Achievability bounds (roughly):

$$R^*(n, \epsilon) \gtrsim \epsilon\text{-th quantile of } \frac{1}{n} \log \frac{dP_{Y^n|X^n}}{dQ_{Y^n}}$$

- Info-density $i(X^n; Y^n) = \log \frac{dP_{Y^n|X^n}}{dQ_{Y^n}}$ is a sum of iid.
- Choice of Q_{Y^n} is an art. Often c.a.o.d. works. Then, $\mathbb{E}[i(X^n; Y^n)] = nC$.
- So by CLT

$$R^*(n, \epsilon) \approx \epsilon\text{-quantile of } \mathcal{N}(C, V/n)$$

FBL achievability bounds

- A random transformation $A \xrightarrow{P_{Y|X}} B$
- (M, ϵ) codes:

$$W \rightarrow A \rightarrow B \rightarrow \hat{W} \quad W \sim \text{Unif}\{1, \dots, M\}$$

$$\mathbb{P}[W \neq \hat{W}] \leq \epsilon$$

- For every $P_{XY} = P_X P_{Y|X}$ define **information density**:

$$i(x; y) \triangleq \log \frac{dP_{Y|X=x}}{dP_Y}(y)$$

- ▶ $\mathbb{E}[i(X; Y)] = I(X; Y)$
- ▶ $\text{Var}[i(X; Y)|X] = V$
- ▶ Memoryless channels: $i(A^n; B^n) = \text{sum of iid.}$

$$i(A^n; B^n) \stackrel{d}{\approx} nI(A; B) + \sqrt{nV}Z, \quad Z \sim \mathcal{N}(0, 1)$$

- Goal: Prove FBL bounds.

As by-product: $R^*(n, \epsilon) \gtrsim C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$

Theorem (Dependence Testing Bound)

For any P_X there exists a code with M codewords and

$$\epsilon \leq \mathbb{E} \left[\exp \left\{ - \left| \iota_{X;Y}(X;Y) - \log \frac{M-1}{2} \right|^+ \right\} \right].$$

Highlights:

- Strictly stronger than Feinstein-Shannon
- ... and no optimization over γ !
- Easier to compute than RCU
- Easier asymptotics: $\epsilon \leq \mathbb{E} \left[e^{-n \left| \frac{1}{n} \iota(X^n; Y^n) - R \right|^+} \right]$
 $\approx Q \left(\sqrt{\frac{n}{V}} \{ I(X; Y) - R \} \right)$
- Has a form of f -divergence: $1 - \epsilon \geq D_f(P_{XY} \| P_X P_Y)$

- Codebook: random $C_1, \dots, C_M \sim P_X$ iid
- Feinstein decoder:

$$\hat{W} = \text{smallest } j \text{ s.t. } \iota_{X;Y}(C_j; Y) > \gamma$$

- j -th codeword's probability of error:

$$\mathbb{P}[\text{error} \mid W = j] \leq \underbrace{\mathbb{P}[\iota_{X;Y}(X; Y) \leq \gamma]}_{\textcircled{a}} + (j - 1) \underbrace{\mathbb{P}[\iota_{X;Y}(\bar{X}; Y) > \gamma]}_{\textcircled{b}}$$

In \textcircled{a} : C_j too far from Y

In \textcircled{b} : C_k with $k < j$ is too close to Y

- Average over W :

$$\mathbb{P}[\text{error}] \leq \mathbb{P}[\iota_{X;Y}(X; Y) \leq \gamma] + \frac{M-1}{2} \mathbb{P}[\iota_{X;Y}(\bar{X}; Y) > \gamma]$$

- Recap: for every γ there exists a code with

$$\epsilon \leq \mathbb{P} [I_{X;Y}(X;Y) \leq \gamma] + \frac{M-1}{2} \mathbb{P} [I_{X;Y}(\bar{X};Y) > \gamma] .$$

- Key step:** closed-form optimization of γ .
- Introduce $\bar{X} \perp\!\!\!\perp Y: I_{X;Y} = \log \frac{dP_{XY}}{dP_{\bar{X}Y}}$
- We have

$$P_{XY} \left[\frac{dP_{XY}}{dP_{\bar{X}Y}} \leq e^\gamma \right] + \frac{M-1}{2} P_{\bar{X}Y} \left[\frac{dP_{XY}}{dP_{\bar{X}Y}} > e^\gamma \right]$$

Bayesian dependence testing!

Optimum threshold: Ratio of priors $\Rightarrow \boxed{\gamma^* = \log \frac{M-1}{2}}$

- Change of measure argument:

$$P \left[\frac{dP}{dQ} \leq \tau \right] + \tau Q \left[\frac{dP}{dQ} > \tau \right] = \mathbb{E}_P \left[\exp \left\{ - \left| \log \frac{dP}{dQ} - \log \tau \right|^+ \right\} \right] .$$

- Take a random transformation $A \xrightarrow{P_{Y|X}} B$
(think $A = \mathcal{A}^n$, $B = \mathcal{B}^n$, $P_{Y|X} = P_{Y^n|X^n}$)
- Input distribution P_X induces $P_Y = P_{Y|X} \circ P_X$
 $P_{XY} = P_X P_{Y|X}$
- Fix code:

$$W \xrightarrow{\text{encoder}} X \rightarrow Y \xrightarrow{\text{decoder}} \hat{W}$$

$W \sim \text{Unif}[M]$ and $M = \#$ of codewords

Input distribution P_X associated to a code:

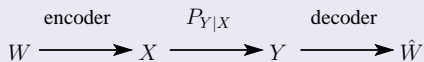
$$P_X[\cdot] \triangleq \frac{\# \text{ of codewords } \in (\cdot)}{M}.$$

- **Goal:** Upper bounds on $\log M$ in terms of $\epsilon \triangleq \mathbb{P}[\text{error}]$

As by-product: $R^*(n, \epsilon) \lesssim C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$

Theorem (Fano)

For any code



with $W \sim \text{Unif}\{1, \dots, M\}$:

$$\log M \leq \frac{\sup_{P_X} I(X; Y) + h(\epsilon)}{1 - \epsilon}, \quad \epsilon = \mathbb{P}[W \neq \hat{W}]$$

Implies *weak converse*:

$$R^*(n, \epsilon) \leq \frac{C}{1 - \epsilon} + o(1).$$

Proof: ϵ -small $\implies H(W|\hat{W})$ -small $\implies I(X; Y) \approx H(W) = \log M$

A (very long) proof of Fano via *channel substitution*

Consider two distributions on (W, X, Y, \hat{W}) :

$$\mathbb{P}: P_{WXY\hat{W}} = P_W \times P_{X|W} \times P_{Y|X} \times P_{\hat{W}|Y}$$

DAG: $W \rightarrow X \rightarrow Y \rightarrow \hat{W}$

$$\mathbb{Q}: Q_{WXY\hat{W}} = P_W \times P_{X|W} \times Q_Y \times P_{\hat{W}|Y}$$

DAG: $W \rightarrow X \text{---} Y \rightarrow \hat{W}$

Under \mathbb{Q} the channel is useless:

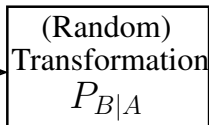
$$\mathbb{Q}[W = \hat{W}] = \sum_{m=1}^M P_W(m) Q_{\hat{W}}(m) = \frac{1}{M} \sum_{m=1}^M Q_{\hat{W}}(m) = \frac{1}{M}$$

Next step: data-processing for relative entropy $D(\cdot||\cdot)$

Input distribution

P_A

Q_A



Output distribution

P_B

Q_B

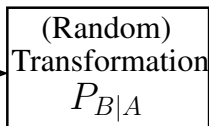


$$D(P_A\|Q_A) \geq D(P_B\|Q_B)$$

Input distribution

P_A

Q_A



Output distribution

P_B

Q_B



$$D(P_A||Q_A) \geq D(P_B||Q_B)$$

Apply to transform: $(W, X, Y, \hat{W}) \mapsto 1\{W \neq \hat{W}\}$:

$$\begin{aligned} D(P_{WXY\hat{W}}||Q_{WXY\hat{W}}) &\geq d(\mathbb{P}[W = \hat{W}]||\mathbb{Q}[W = \hat{W}]) \\ &= d(1 - \epsilon||\frac{1}{M}) \end{aligned}$$

where $d(x||y) = x \log \frac{x}{y} + (1 - x) \log \frac{1-x}{1-y}$.

So far:

$$D(P_{WXY\hat{W}} \| Q_{WXY\hat{W}}) \geq d(1 - \epsilon \| \frac{1}{M})$$

Lower-bound RHS:

$$d(1 - \epsilon \| \frac{1}{M}) \geq (1 - \epsilon) \log M - h(\epsilon)$$

Analyze LHS:

$$\begin{aligned} D(P_{WXY\hat{W}} \| Q_{WXY\hat{W}}) &= D(P_{XY} \| Q_{XY}) \\ &= D(P_X P_{Y|X} \| P_X Q_Y) \\ &= D(P_{Y|X} \| Q_Y | P_X) \end{aligned}$$

(Recall: $D(P_{Y|X} \| Q_Y | P_X) = \mathbb{E}_{x \sim P_X} [D(P_{Y|X=x} \| Q_Y)]$)

A proof of Fano via *channel substitution*: last step

Putting it all together:

$$(1 - \epsilon) \log M \leq D(P_{Y|X} \| Q_Y | P_X) + h(\epsilon) \quad \forall Q_Y \quad \forall \text{code}$$

Two methods:

- 1 Compute $\sup_{P_X} \inf_{Q_Y}$ and recall

$$\inf_{Q_Y} D(P_{Y|X} \| Q_Y | P_X) = I(X; Y)$$

- 2 Take $Q_Y = P_Y^* =$ **the caod** (capacity achieving output dist.) and recall

$$D(P_{Y|X} \| P_Y^* | P_X) \leq \sup_X I(X; Y) \quad \forall P_X$$

Conclude:

$$(1 - \epsilon) \log M \leq \sup_{P_X} I(X; Y) + h(\epsilon)$$



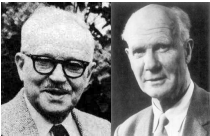
Important: Second method is particularly useful for FBL!

Tightening: from $D(\cdot||\cdot)$ to $\beta_\alpha(\cdot, \cdot)$

Question: How about replacing $D(\cdot||\cdot)$ with other divergences?



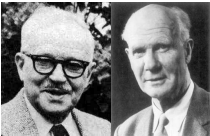
Tightening: from $D(\cdot||\cdot)$ to $\beta_\alpha(\cdot, \cdot)$

Question: How about replacing $D(\cdot||\cdot)$ with other divergences?

	$D(\cdot \cdot)$	relative entropy (KL divergence)	weak converse
	$D_\lambda(\cdot \cdot)$	Rényi divergence	strong converse
	$\beta_\alpha(\cdot, \cdot)$	Neyman-Pearson ROC curve	FBL bounds

Tightening: from $D(\cdot||\cdot)$ to $\beta_\alpha(\cdot, \cdot)$

Question: How about replacing $D(\cdot||\cdot)$ with other divergences?

	$D(\cdot \cdot)$	relative entropy (KL divergence)	weak converse
	$D_\lambda(\cdot \cdot)$	Rényi divergence	strong converse
	$\beta_\alpha(\cdot, \cdot)$	Neyman-Pearson ROC curve	FBL bounds

Note: Using β_α is aka *meta-converse*.

$$\dots \text{ and leads to } R^*(n, \epsilon) \leq C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

General meta-converse principle

Steps:

- Select auxiliary channel $Q_{Y|X}$ (art)
E.g.: $Q_{Y|X=x}$ = center of a cluster of x
- Prove converse bound for channel $Q_{Y|X}$
E.g.: $\mathbb{Q}[W = \hat{W}] \lesssim \frac{\# \text{ of clusters}}{M}$
- Compute distance $D(\mathbb{P}||\mathbb{Q})$ between two spaces

$$\mathbb{P} : P_{WXY\hat{W}} = P_W \times P_{X|W} \times P_{Y|X} \times P_{\hat{W}|Y}$$

vs.

$$\mathbb{Q} : P_{WXY\hat{W}} = P_W \times P_{X|W} \times Q_{Y|X} \times P_{\hat{W}|Y}$$

- Apply data processing: $D(P_{W,\hat{W}}||Q_{W,\hat{W}}) \leq D(P_{X,Y}||Q_{X,Y})$
- **Key observation:** This inequality connects $\mathbb{P}[\text{error}]$, $\mathbb{Q}[\text{error}]$ and distance $D(\mathbb{P}||\mathbb{Q})$.

- All in all, these methods allow us to conclude:

$$R^*(n, \epsilon) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

for a wide range of channels.

- Typically, $V = \text{Var}[i(X; Y)|X]$ for cap.ach. distribution X .

- All in all, these methods allow us to conclude:

$$R^*(n, \epsilon) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

for a wide range of channels.

- Typically, $V = \text{Var}[i(X; Y)|X]$ for cap.ach. distribution X .
- Example: **The AWGN Channel**

$$\begin{array}{c} Z \sim \mathcal{N}(0, \sigma^2) \\ \downarrow \\ X \longrightarrow \oplus \longrightarrow Y \end{array}$$

Codewords $x^n \in \mathbb{R}^n$ satisfy power-constraint: $\sum_{j=1}^n |x_j|^2 \leq nP$

$$C(P) = \frac{1}{2} \log(1 + P), \quad V(P) = \frac{\log^2 e}{2} \left(1 - \frac{1}{(1 + P)^2} \right)$$

- Curious property of Gaussian noise: $V(P) \leq \frac{\log^2 e}{2}$

- All in all, these methods allow us to conclude:

$$R^*(n, \epsilon) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon)$$

for a wide range of channels.

- Typically, $V = \text{Var}[i(X; Y)|X]$ for cap.ach. distribution X .
- Example: **The AWGN Channel**

Below for **Gaussian MAC** we focus on m.i./capacity. By FBL there \exists codes within $O(\frac{1}{\sqrt{n}})$ uniformly in P .

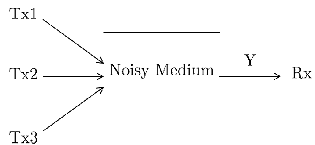
Codewords $x^n \in \mathbb{R}^n$ satisfy power-constraint: $\sum_{j=1}^n |x_j|^2 \leq nP$

$$C(P) = \frac{1}{2} \log(1 + P), \quad V(P) = \frac{\log^2 e}{2} \left(1 - \frac{1}{(1 + P)^2} \right)$$

- Curious property of Gaussian noise: $V(P) \leq \frac{\log^2 e}{2}$

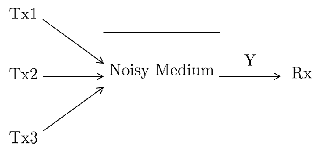
Classical multiple-access IT

IT vs networks view on MAC

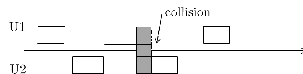


- Core problem: many users, one channel

IT vs networks view on MAC



- Core problem: many users, one channel
- Networking folks:

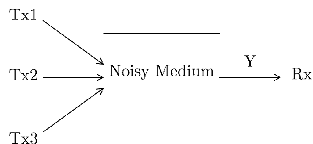


- **ALOHA** protocol (slotted) achieves:

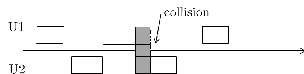
$$\sum_i R_i \approx 0.37C$$

- **Open problem:** what max fraction η^* achievable?
State of the art [Tsybakov-Lihanov'87]: $0.476 \leq \eta^* \leq 0.568$
(collision resolution codes)

IT vs networks view on MAC



- Core problem: many users, one channel
- Networking folks:

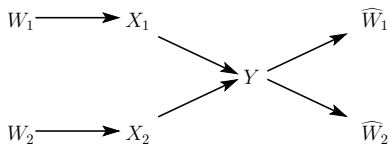


- **ALOHA** protocol (slotted) achieves:

$$\sum_i R_i \approx 0.37C$$

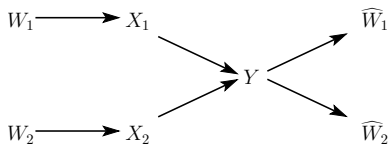
- **Open problem:** what max fraction η^* achievable?
State of the art [Tsybakov-Lihanov'87]: $0.476 \leq \eta^* \leq 0.568$
(collision resolution codes)
- **IT:** We want $\sum_i R_i \gg C$!
- **How?** By exploiting **physics** of collision.

2-user MAC: IT formalism



- 2-input channel: $P_{Y|X_1, X_2}$ (memoryless)
- Random messages $W_1 \in [2^{nR_1}]$, $W_2 \in [2^{nR_2}]$
- Encoders: $X_1^n = f_1(W_1)$, $X_2^n = f_2(W_2)$
- **Joint decoder:** $(\hat{W}_1, \hat{W}_2) = g(Y)$
- **Joint probability of error:**

$$\mathbb{P}[W_1 \neq \hat{W}_1, W_2 \neq \hat{W}_2] \geq 1 - \epsilon.$$



- 2-input channel: $P_{Y|X_1, X_2}$ (memoryless)
- Random messages $W_1 \in [2^{nR_1}]$, $W_2 \in [2^{nR_2}]$
- Encoders: $X_1^n = f_1(W_1)$, $X_2^n = f_2(W_2)$
- **Joint decoder:** $(\hat{W}_1, \hat{W}_2) = g(Y)$
- **Joint probability of error:**

$$\mathbb{P}[W_1 = \hat{W}_1, W_2 = \hat{W}_2] \geq 1 - \epsilon.$$

- FBL fundamental limit (region):

$$R^*(n, \epsilon) = \{(R_1, R_2) : \exists(2^{nR_1}, 2^{nR_2}, \epsilon)\text{-code}\}$$

- Asymptotics: $[\cdot] = \text{closure}$

$$C_\epsilon = \left[\liminf_{n \rightarrow \infty} R^*(n, \epsilon) \right], \quad C = \bigcap_{\epsilon > 0} C_\epsilon$$

Theorem (Ahlsvede-Liao (capacity) + Dueck (Strong converse))

$$C = C_\epsilon = \left[\text{co} \left\{ \bigcup_{P_{X_1}, P_{X_2}} \text{Penta}(P_{X_1}, P_{X_2}) \right\} \right]$$

$$\text{Penta}(P_{X_1}, P_{X_2}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \leq I(X_1, X_2; Y) \\ R_1 \leq I(X_1; Y | X_2) \\ R_2 \leq I(X_2; Y | X_1) \end{array} \right\}$$

- $\text{co}\{\cdot\}$ – convex hull
- **Fun fact:** w/o synchronization $C = [\bigcup \text{Penta}]$ but w/o $\text{co}\{\cdot\}$!

Theorem (Ahlsvede-Liao (capacity) + Dueck (Strong converse))

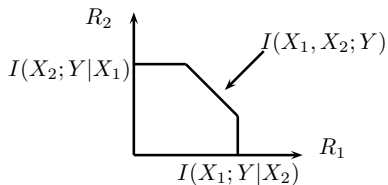
$$C = C_\epsilon = \left[\text{co} \left\{ \bigcup_{P_{X_1}, P_{X_2}} \text{Penta}(P_{X_1}, P_{X_2}) \right\} \right]$$

$$\text{Penta}(P_{X_1}, P_{X_2}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \leq I(X_1, X_2; Y) \\ R_1 \leq I(X_1; Y|X_2) \\ R_2 \leq I(X_2; Y|X_1) \end{array} \right\}$$

- $\text{co}\{\cdot\}$ – convex hull
- **Fun fact:** w/o synchronization $C = [\bigcup \text{Penta}]$ but w/o $\text{co}\{\cdot\}$!
- Not true with cost constraints. In that case need **time-sharing**:

$$C = \bigcup_{X_1, X_2, U} \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \leq I(X_1, X_2; Y|U) \\ R_1 \leq I(X_1; Y|X_2, U) \\ R_2 \leq I(X_2; Y|X_1, U) \end{array} \right\}.$$

$$\text{Penta}(P_{X_1}, P_{X_2}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \leq I(X_1, X_2; Y) \\ R_1 \leq I(X_1; Y|X_2) \\ R_2 \leq I(X_2; Y|X_1) \end{array} \right\}$$



Note: After taking $\bigcup_{P_{X_1}, P_{X_2}}$ and convex-hull, resulting region may be curvilinear!

Theorem

$$C = C_\epsilon = \left[\text{co} \left\{ \bigcup_{P_{X_1}, P_{X_2}} \text{Penta}(P_{X_1}, P_{X_2}) \right\} \right]$$

Here is a standard proof

- Weak-converse:
 - ▶ **sum-rate**

$$R_1 + R_2 \lesssim \frac{1}{n} I(X_1^n, X_2^n; Y^n) \leq \frac{1}{n} \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_i).$$

- ▶ genie gives X_1^n to decoder

$$R_2 \lesssim \frac{1}{n} I(X_2^n; Y^n | X_1^n) \leq \frac{1}{n} \sum_{i=1}^n I(X_{2i}; Y_i | X_{1i})$$

- ▶ Hence $(R_1, R_2) \in \frac{1}{n} \sum_i \text{Penta}(P_{X_{1i}}, P_{X_{2i}})$

Theorem

$$C = C_\epsilon = \left[\text{co} \left\{ \bigcup_{P_{X_1}, P_{X_2}} \text{Penta}(P_{X_1}, P_{X_2}) \right\} \right]$$

Here is a standard proof

- Achievability:
 - ▶ Fix P_{X_1}, P_{X_2} .
 - ▶ Generate codewords for user i from $(P_{X_1})^{\otimes n}$ iid
 - ▶ Decode via joint-typicality
 - ▶ Have $(M_1 - 1)(M_2 - 1)$ possibilities with both \hat{W}_1, \hat{W}_2 wrong (each w.p. $\leq 2^{-nI(X_1, X_2; Y)}$)
 - ▶ Have $M_i - 1$ possibilities with \hat{W}_i wrong (each w.p. $\leq 2^{-nI(X_i; Y|X_i)}$)
 - ▶ Hence, if $(R_1, R_2) \in \text{Penta}(P_{X_1}, P_{X_2})$ all **three** types of errors are small.
 - ▶ Let us understand this more carefully...

MAC achievability: details I

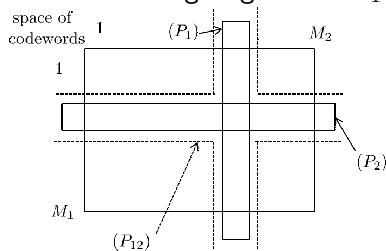
- Gen. $M_1 = 2^{nR_1}$ codewords $C_i \stackrel{iid}{\sim} (P_{X_1})^{\otimes n}$
- Gen. $M_2 = 2^{nR_2}$ codewords $D_i \stackrel{iid}{\sim} (P_{X_2})^{\otimes n}$
- True message $W_1 = i_0, W_2 = j_0$.
- Decoder sees y^n . **How to decode?**

MAC achievability: details I

- Gen. $M_1 = 2^{nR_1}$ codewords $C_i \stackrel{iid}{\sim} (P_{X_1})^{\otimes n}$
- Gen. $M_2 = 2^{nR_2}$ codewords $D_i \stackrel{iid}{\sim} (P_{X_2})^{\otimes n}$
- True message $W_1 = i_0, W_2 = j_0$.
- Decoder sees y^n . **How to decode?**
- Why is this not the same as decoding single-user $M_1 \times M_2$ -size code?

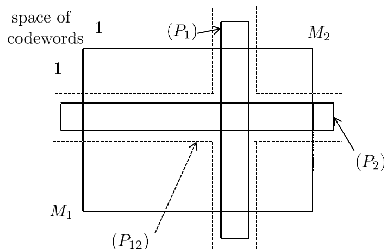
MAC achievability: details I

- Gen. $M_1 = 2^{nR_1}$ codewords $C_i \stackrel{iid}{\sim} (P_{X_1})^{\otimes n}$
- Gen. $M_2 = 2^{nR_2}$ codewords $D_j \stackrel{iid}{\sim} (P_{X_2})^{\otimes n}$
- True message $W_1 = i_0, W_2 = j_0$.
- Decoder sees y^n . **How to decode?**
- Why is this not the same as decoding single-user $M_1 \times M_2$ -size code?



- Extra structure: $(C_{i_0}, D_j) \not\perp (C_{i_0}, D_{j_0})$

MAC achievability: details II



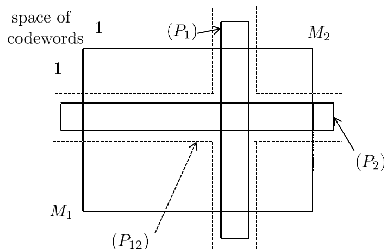
- Decoder sees y^n . **How to decode?**
- A good **test** for rejecting $(M_1 - 1)(M_2 - 1)$ codewords in (P_{12}) :
 $(T_{12}) \quad i(c_i, d_j; y^n) \leq \gamma_{12} \Rightarrow \text{remove } (i, j) \text{ from consideration}$

- $i(c, d; y^n) \triangleq \log \frac{P_{Y^n | X_1^n, X_2^n}(y^n | c, d)}{P_{Y^n}(y^n)}$
- Standard bound: $\forall i \neq i_0, j \neq j_0$:

$$\mathbb{P}[i(C_i, D_j; Y^n) > \gamma_{12}] \leq e^{-\gamma_{12}}$$

- Set $\gamma_{12} = \log(M_1 M_2) + \tau$ then test (T_{12}) filters all $(i, j) \in (P_{12})$

MAC achievability: details III



- Decoder sees y^n . **How to decode?**
- A good **test** for rejecting $(M_2 - 1)$ codewords in (P_2) :

$$(T_2) \quad i(d_j; y^n | c_i) \leq \gamma_2 \Rightarrow \text{remove } (i, j) \text{ from consideration}$$

- $i(d; y^n | c) \triangleq \log \frac{P_{Y^n | X_1^n, X_2^n}(y^n | c, d)}{P_{Y^n | X_1^n}(y^n | c)}$
- Standard bound: $\forall j \neq j_0$:

$$\mathbb{P}[i(D_j; Y^n | C_{i_0}) > \gamma_2] \leq e^{-\gamma_2}$$

- Set $\gamma_2 = \log(M_2) + \tau$ then test (T_2) filters all $(i_0, j) \in (P_2)$

- Decoder sees y^n . **How to decode?**

$$(T_{12}) \quad i(c_i, d_j; y^n) \leq n(R_1 + R_2) + \tau \quad \Rightarrow \text{remove } (i, j)$$

$$(T_1) \quad i(c_i; y^n | d_j) \leq nR_1 + \tau \quad \Rightarrow \text{remove } (i, j)$$

$$(T_2) \quad i(d_j; y^n | c_i) \leq nR_2 + \tau \quad \Rightarrow \text{remove } (i, j)$$

- This achieves:

$$\epsilon \leq 3e^{-\tau} + \mathbb{P} \left[\left\{ i(X_1^n, X_2^n; Y^n) \leq n(R_1 + R_2) + \tau \right\} \cup \left\{ i(X_1^n; Y^n | X_2^n) \leq nR_1 + \tau \right\} \cup \left\{ i(X_2^n; Y^n | X_1^n) \leq nR_2 + \tau \right\} \right].$$

- By **CLT** a (R_1, R_2) within $\frac{1}{\sqrt{n}}$ of the boundary of Penta is achievable.

- Decoder sees y^n . **How to decode?**

$$(T_{12}) \quad i(c_i, d_j; y^n) \leq n(R_1 + R_2) + \tau \quad \Rightarrow \text{remove } (i, j)$$

$$(T_1) \quad i(c_i; y^n | d_j) \leq nR_1 + \tau \quad \Rightarrow \text{remove } (i, j)$$

$$(T_2) \quad i(d_j; y^n | c_i) \leq nR_2 + \tau \quad \Rightarrow \text{remove } (i, j)$$

- This achieves:

$$\epsilon \leq 3e^{-\tau} + \mathbb{P} \left[\left\{ i(X_1^n, X_2^n; Y^n) \leq n(R_1 + R_2) + \tau \right\} \cup \left\{ i(X_1^n; Y^n | X_2^n) \leq nR_1 + \tau \right\} \cup \left\{ i(X_2^n; Y^n | X_1^n) \leq nR_2 + \tau \right\} \right].$$

- By CLT a (R_1, R_2) within $\frac{1}{\sqrt{n}}$ of the boundary of Penta is achievable.
- Typical decoding
 - ▶ Use (T_{12}) rule – this is like decoding single-user $M_1 \times M_2$ -code (LDPC+LDGM structure!)
 - ▶ After applying it, most often get only one (true) message left (!)

- Decoder sees y^n . **How to decode?**

$$(T_{12}) \quad i(c_i, d_j; y^n) \leq n(R_1 + R_2) + \tau \quad \Rightarrow \text{remove } (i, j)$$

$$(T_1) \quad i(c_i; y^n | d_j) \leq nR_1 + \tau \quad \Rightarrow \text{remove } (i, j)$$

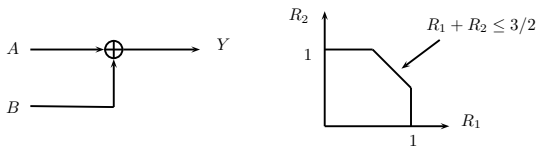
$$(T_2) \quad i(d_j; y^n | c_i) \leq nR_2 + \tau \quad \Rightarrow \text{remove } (i, j)$$

- This achieves:

$$\epsilon \leq 3e^{-\tau} + \mathbb{P} \left[\left\{ i(X_1^n, X_2^n; Y^n) \leq n(R_1 + R_2) + \tau \right\} \cup \left\{ i(X_1^n; Y^n | X_2^n) \leq nR_1 + \tau \right\} \cup \left\{ i(X_2^n; Y^n | X_1^n) \leq nR_2 + \tau \right\} \right].$$

- By CLT a (R_1, R_2) within $\frac{1}{\sqrt{n}}$ of the boundary of Penta is achievable.
- Typical decoding
 - ▶ Use (T_{12}) rule – this is like decoding single-user $M_1 \times M_2$ -code (LDPC+LDGM structure!)
 - ▶ After applying it, most often get only one (true) message left (!)
 - ▶ Unless $R_1 = I(X_1; Y | X_2) + O(\frac{1}{\sqrt{n}})$.
 - ▶ In this case, many (i, j) 's remain. **But they are all in one column!**
 - ▶ Hence decode W_2 . Conditioned on X_2 – decode M_1 -code.

Example: Binary Adder Channel (BAC)



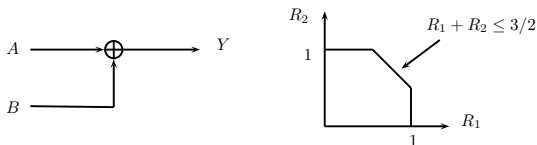
$$Y = X_1 + X_2 \quad X_i \in \{0, 1\}, Y \in \{0, 1, 2\}$$

- Maximal sum-rate:

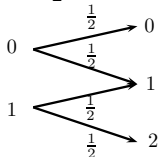
$$C_{sum} = \max_{A,B} I(A, B; Y) = \max H(A + B) = \frac{3}{2} \log 2$$

- Each user can send 1 bit/ch.use. But together $\frac{3}{2}$ bit/ch.use. **How?**

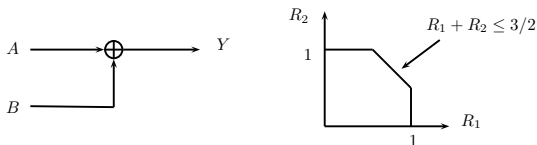
Example: Binary Adder Channel (BAC)



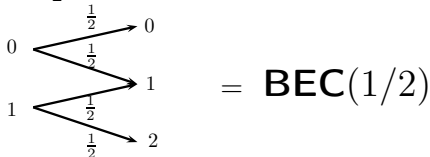
- Take $R_1 = 1$. Then $X_2 \rightarrow Y$ sees channel:



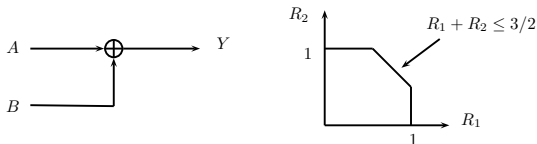
Example: Binary Adder Channel (BAC)



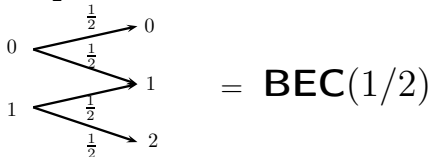
- Take $R_1 = 1$. Then $X_2 \rightarrow Y$ sees channel:



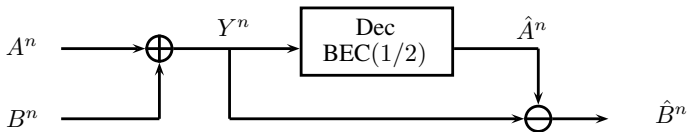
Example: Binary Adder Channel (BAC)



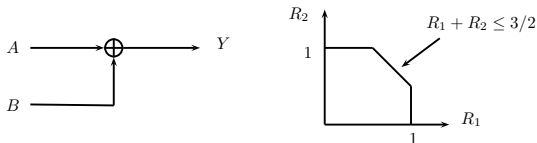
- Take $R_1 = 1$. Then $X_2 \rightarrow Y$ sees channel:



- successive interference cancellation (SIC):



Example: Binary Adder Channel (BAC)



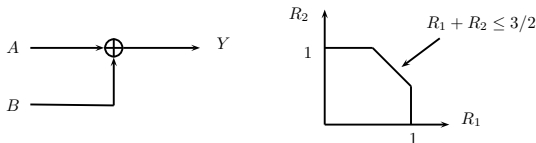
$$Y = X_1 + X_2 \quad X_i \in \{0, 1\}, Y \in \{0, 1, 2\}$$

- Analyzing FBL achievability we can show: (maximal sumrate)

$$R_{sum}^*(n, \epsilon) \geq \frac{3}{2} - \sqrt{\frac{1}{4n}} Q^{-1}(\epsilon) + O(\log n).$$

- Open problem:** Prove $R_{sum}^*(n, \epsilon) \leq \frac{3}{2} + \sqrt{\frac{1}{n}} K_\epsilon$

Example: Binary Adder Channel (BAC)



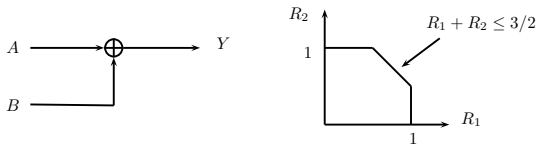
$$Y = X_1 + X_2 \quad X_i \in \{0, 1\}, Y \in \{0, 1, 2\}$$

- Analyzing FBL achievability we can show: (maximal sumrate)

$$R_{sum}^*(n, \epsilon) \geq \frac{3}{2} - \sqrt{\frac{1}{4n}} Q^{-1}(\epsilon) + O(\log n).$$

- Open problem:** Prove $R_{sum}^*(n, \epsilon) \leq \frac{3}{2} + \sqrt{\frac{1}{n}} K_\epsilon$
- ... not even asking for $K_\epsilon < 0$
- ... So far best-known result (Ahslwede): $R_{sum}^* \leq \frac{3}{2} + c\sqrt{\frac{1}{n}} \log n$

Example: Binary Adder Channel (BAC)



$$Y = X_1 + X_2 \quad X_i \in \{0, 1\}, Y \in \{0, 1, 2\}$$

- Analyzing FBL achievability we can show: (maximal sumrate)

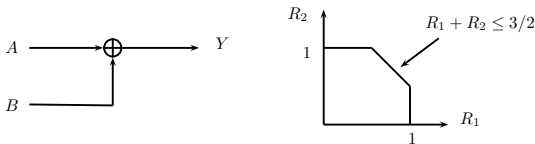
$$R_{sum}^*(n, \epsilon) \geq \frac{3}{2} - \sqrt{\frac{1}{4n}} Q^{-1}(\epsilon) + O(\log n).$$

- Open problem:** Prove $R_{sum}^*(n, \epsilon) \leq \frac{3}{2} + \sqrt{\frac{1}{n}} K_\epsilon$
- ... not even asking for $K_\epsilon < 0$
- ... So far best-known result (Ahslwede): $R_{sum}^* \leq \frac{3}{2} + c\sqrt{\frac{1}{n}} \log n$
- The state is so bad that even for $\epsilon = 0$ we only know (Fano):

$$R_{sum}^*(n, \epsilon = 0) \leq \frac{3}{2}$$

- Open problem:** Prove $\lim_{n \rightarrow \infty} R_{sum}^*(n, \epsilon = 0) < \frac{3}{2}$.

Example: Binary Adder Channel (BAC)



$$Y = X_1 + X_2 \quad X_i \in \{0, 1\}, Y \in \{0, 1, 2\}$$

- Analyzing FBL achievability we can show: (maximal sumrate)

$$R_{sum}^*(n, \epsilon) \geq \frac{3}{2} - \sqrt{\frac{1}{4n}} Q^{-1}(\epsilon) + O(\log n).$$

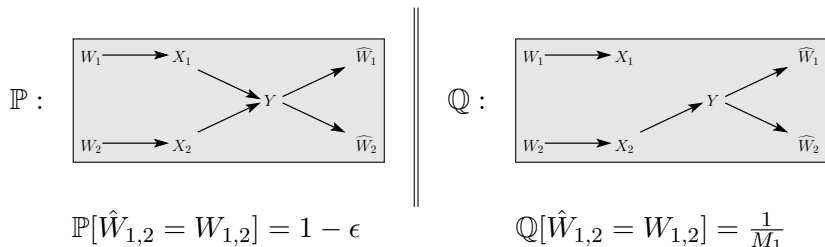
- Open problem:** Prove $R_{sum}^*(n, \epsilon) \leq \frac{3}{2} + \sqrt{\frac{1}{n}} K_\epsilon$
- Conjecture:** [Ajjanagadde-P.'15] for all $0 < \alpha < 1$

$$\max_{A^n \perp\!\!\!\perp B^n} H_\alpha(A^n + B^n) = nH_\alpha\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$

where $H_\alpha(\cdot)$ is Renyi entropy.

- If true implies **Open problem**. **How?**

MAC: revisit weak-converse (genie)



... apply data processing of $D(\cdot||\cdot)$...

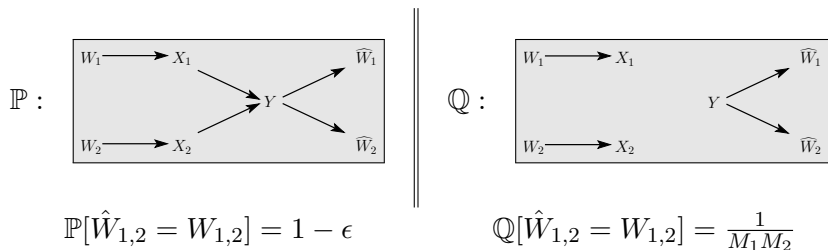


$$d(1 - \epsilon || \frac{1}{M_1}) \leq D(P_{Y|X_1 X_2} || \mathcal{Q}_{Y|X_1} | P_{X_1} P_{X_2})$$

Optimizing $\mathcal{Q}_{Y|X_1}$:

$$\log M_1 \leq \frac{I(X_1; Y | X_2) + h(\epsilon)}{1 - \epsilon}$$

MAC: revisit weak-converse (genie)



... apply data processing of $D(\cdot||\cdot)$...



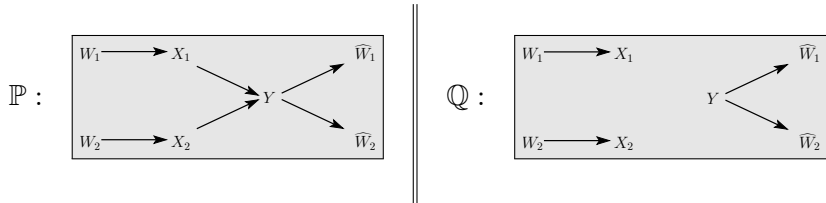
$$d(1 - \epsilon || \frac{1}{M_1}) \leq D(P_{Y|X_1 X_2} || Q_Y | P_{X_1} P_{X_2})$$

Optimizing Q_Y :

$$\log M_1 M_2 \leq \frac{I(X_1, X_2; Y) + h(\epsilon)}{1 - \epsilon}$$

Together with previous: full (pentagon) weak converse

MAC: towards strong-converse



$$\mathbb{P}[\hat{W}_{1,2} = W_{1,2}] = 1 - \epsilon$$

$$\mathbb{Q}[\hat{W}_{1,2} = W_{1,2}] = \frac{1}{M_1 M_2}$$

... use Renyi $D_\lambda(\cdot\|\cdot)$...

$$\Downarrow$$
$$D_\lambda(P_{X_1 X_2 Y} \| P_{X_1} P_{X_2} Q_Y) \geq d_\lambda(1 - \epsilon) \frac{1}{M_1 M_2}$$

Selecting $\lambda = 1 + \frac{1}{\sqrt{n}}$ yields (for BAC)

$$\log M_1, M_2 \leq \sup_{A^n \perp\!\!\!\perp B^n} H_{\alpha_n}(A^n + B^n) + K\sqrt{n}$$

with $\alpha_n = 1 - \frac{1}{\sqrt{n}}$.

- Trivially generalizes to K -user MAC:

$$\text{Penta} = \{(R_1, \dots, R_K) : \sum_{i \in S} R_i \leq I(X_S; Y | X_{S^c}) \forall S \subset [K]\}$$

- Classic IT: Fix K let $n \rightarrow \infty$.
- Use **joint probability of error**:

$$\mathbb{P}[W_1 = \hat{W}_1, \dots, W_K = \hat{W}_k] \geq 1 - \epsilon.$$

- New **FBL** issue: for $K = 100$ need 2^{100} tests in achievability.

- Trivially generalizes to K -user MAC:

$$\text{Penta} = \{(R_1, \dots, R_K) : \sum_{i \in S} R_i \leq I(X_S; Y | X_{S^c}) \forall S \subset [K]\}$$

- Classic IT: Fix K let $n \rightarrow \infty$.
- Use **joint probability of error**:

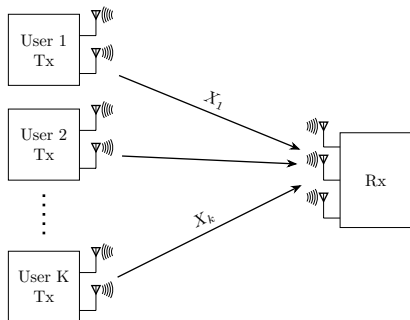
$$\mathbb{P}[W_1 = \hat{W}_1, \dots, W_K = \hat{W}_k] \geq 1 - \epsilon.$$

- New **FBL** issue: for $K = 100$ need 2^{100} tests in achievability.
- What is new today?
 - ▶ Many-user scaling [D. Guo et al]: $K = \mu n, n \rightarrow \infty$
 - ▶ New probability of error [P.'17]: $\frac{1}{K} \sum_i \mathbb{P}[W_i \neq \hat{W}_i] \leq \epsilon$
 - ▶ Same-codebook coding [P.'17]: $X_i \in \mathcal{C}$ for all i .

Gaussian MAC. Modulation

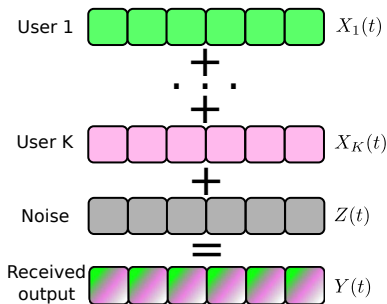
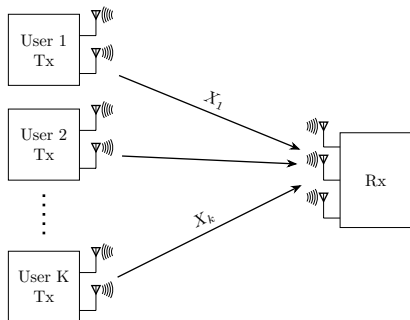
Let's put on our engineering shoes.

The classical model: K-user multiple-access channel



$$Y(t) = X_1(t) + \cdots + X_K(t) + Z(t)$$

The classical model: K-user multiple-access channel

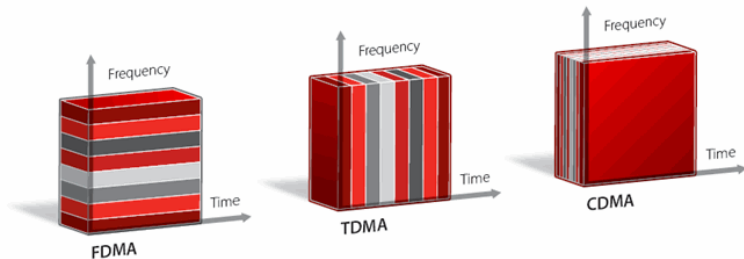
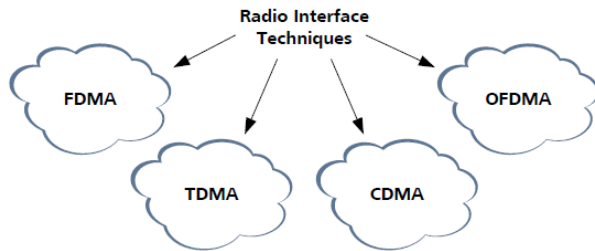


$$Y(t) = X_1(t) + \cdots + X_K(t) + Z(t)$$

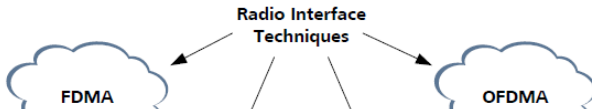
- Users send coded waveforms $X_j(t)$
- Additive Gaussian noise $Z(t)$
- Base station's job: **estimate** X_j from the knowledge of $Y(t)$

Tech note: synchronized block coding

How to avoid inter-user interference?

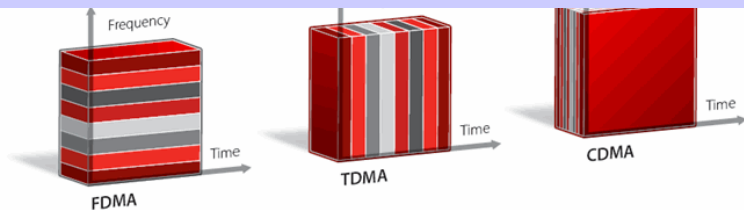


How to avoid inter-user interference?



These are called **orthogonal schemes**

Key problem: resources divided among active and inactive (!) users
(or need costly resource ack/grant protocol)
in IoT most are inactive \Rightarrow huge waste of bandwidth



This “pie-slicing” philosophy comes from:

- Given: W Hz bandwidth and duration T sec:
- By XYZ Theorem: d.o.f. $n = 2WT$

$$XYZ \in \{ \text{Kotelnikov, Nyquist, Shannon, Slepian, ...} \}$$

- TDMA, FDMA, CDMA: just different bases in \mathbb{R}^{2WT} .
(Fine print: CDMA = Orthogonal CDMA here).

This “pie-slicing” philosophy comes from:

- Given: W Hz bandwidth and duration T sec:
- By XYZ Theorem: d.o.f. $n = 2WT$

$XYZ \in \{ \text{Kotelnikov, Nyquist, Shannon, Slepian, ...} \}$

- TDMA, FDMA, CDMA: just different bases in \mathbb{R}^{2WT} .
(Fine print: CDMA = Orthogonal CDMA here).
- Is there value in having $K > n$? (non-orthogonal signalling)
- Is it even possible to have $K > n$ or even $K \gg n$?

This “pie-slicing” philosophy comes from:

- Given: W Hz bandwidth and duration T sec:
- By XYZ Theorem: d.o.f. $n = 2WT$

$XYZ \in \{ \text{Kotelnikov, Nyquist, Shannon, Slepian, ...} \}$

- TDMA, FDMA, CDMA: just different bases in \mathbb{R}^{2WT} .
(Fine print: CDMA = Orthogonal CDMA here).
- Is there value in having $K > n$? (non-orthogonal signalling)
- Is it even possible to have $K > n$ or even $K \gg n$?
- Silly: Take $n = 1$ and let user j send a bit via $\{0, 2^j\}$.

This “pie-slicing” philosophy comes from:

- Given: W Hz bandwidth and duration T sec:
- By XYZ Theorem: d.o.f. $n = 2WT$

$XYZ \in \{ \text{Kotelnikov, Nyquist, Shannon, Slepian, ...} \}$

- TDMA, FDMA, CDMA: just different bases in \mathbb{R}^{2WT} .
(Fine print: CDMA = Orthogonal CDMA here).
- Is there value in having $K > n$? (non-orthogonal signalling)
- **Is it even possible to have $K > n$ or even $K \gg n$?**
- **Silly:** Take $n = 1$ and let user j send a bit via $\{0, 2^j\}$.
- ... cheating: user K 's power is 2^{2K} larger than user 1's.

This “pie-slicing” philosophy comes from:

- Given: W Hz bandwidth and duration T sec:
- By XYZ Theorem: d.o.f. $n = 2WT$

$XYZ \in \{ \text{Kotelnikov, Nyquist, Shannon, Slepian, ...} \}$

- TDMA, FDMA, CDMA: just different bases in \mathbb{R}^{2WT} .
(Fine print: CDMA = Orthogonal CDMA here).
- Is there value in having $K > n$? (non-orthogonal signalling)
- **Is it even possible to have $K > n$ or even $K \gg n$?**
- **Silly:** Take $n = 1$ and let user j send a bit via $\{0, 2^j\}$.
- ... cheating: user K 's power is 2^{2K} larger than user 1's.
- **Challenge:** users only allowed to send ± 1 , can we have $K \gg n$?

Achieving capacity of K -user BAC with zero-error

$$Y = \sum_{j=1}^K X_j \quad X_i \in \{\pm 1\}$$

- Known: $C_{sum}(K) = H(\text{Bin}(K, 1/2)) \approx \frac{1}{2} \log K$.
- IOW, for sending 1-bit (each) the frame-length $n \approx \frac{2K}{\log_2 K} \ll K$.

How can $K > n$ users signal in n dimensions **simultaneously**?

Achieving capacity of K -user BAC with zero-error

$$Y = \sum_{j=1}^K X_j \quad X_i \in \{\pm 1\}$$

- Known: $C_{sum}(K) = H(\text{Bin}(K, 1/2)) \approx \frac{1}{2} \log K$.
- IOW, for sending 1-bit (each) the frame-length $n \approx \frac{2K}{\log_2 K} \ll K$.

How can $K > n$ users signal in n dimensions **simultaneously**?

- **Khachatrian-Martirosian: even with zero-error!**

First, recall a particularly nice orthogonal basis:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_2 = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \end{bmatrix} \quad H_{m+1} = \begin{bmatrix} H_m & H_m \\ H_m & -H_m \end{bmatrix}$$

(each user is modulating his row)

- K.-M. noticed you can add more rows!

How can $K > n$ users signal in n dimensions **simultaneously**?

- Walsh-Hadamard basis:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_2 = \begin{bmatrix} \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} \\ \boxed{1} & \boxed{-1} & \boxed{1} & \boxed{-1} \\ \boxed{1} & \boxed{1} & \boxed{-1} & \boxed{-1} \\ \boxed{1} & \boxed{-1} & \boxed{-1} & \boxed{1} \end{bmatrix} \quad H_{m+1} = \begin{bmatrix} H_m & H_m \\ H_m & -H_m \end{bmatrix}$$

- K.-M. signals:

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \boxed{1} & \boxed{1} & \boxed{1} & \boxed{-1} \end{bmatrix}$$

- Key property: $x \mapsto xA_m$ is injective on $\{\pm 1\}^{K_m}$, $K_m = \frac{m}{2}2^m + 1$

How can $K > n$ users signal in n dimensions **simultaneously**?

- Walsh-Hadamard basis:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_2 = \begin{bmatrix} \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} \\ \boxed{1} & \boxed{-1} & \boxed{1} & \boxed{-1} \\ \boxed{1} & \boxed{1} & \boxed{-1} & \boxed{-1} \\ \boxed{1} & \boxed{-1} & \boxed{-1} & \boxed{1} \end{bmatrix} \quad H_{m+1} = \begin{bmatrix} H_m & H_m \\ H_m & -H_m \end{bmatrix}$$

- K.-M. signals:

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} \\ \boxed{1} & \boxed{-1} & \boxed{1} & \boxed{-1} \\ \boxed{1} & \boxed{1} & \boxed{-1} & \boxed{-1} \\ \boxed{1} & \boxed{-1} & \boxed{-1} & \boxed{1} \\ \boxed{1} & \boxed{1} & \boxed{1} & \boxed{-1} \end{bmatrix} \quad \tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}$$

$\begin{matrix} \uparrow \\ \vdots \\ \uparrow \end{matrix} 2^m$
 $\leftarrow 2^m \quad \leftarrow 2^m \quad \rightarrow$

- Key property: $x \mapsto xA_m$ is injective on $\{\pm 1\}^{K_m}$, $K_m = \frac{m}{2}2^m + 1$
- Number of users at dimension n : $K \approx \frac{1}{2}n \log_2 n$ (optimal!)

How can $K > n$ users signal in n dimensions **simultaneously**?

- Walsh-Hadamard basis:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_2 = \begin{bmatrix} \boxed{1} & \boxed{1} & \boxed{1} & \boxed{1} \\ \boxed{1} & \boxed{-1} & \boxed{1} & \boxed{-1} \\ \boxed{1} & \boxed{1} & \boxed{-1} & \boxed{-1} \\ \boxed{1} & \boxed{-1} & \boxed{-1} & \boxed{1} \end{bmatrix} \quad H_{m+1} = \begin{bmatrix} H_m & H_m \\ H_m & -H_m \end{bmatrix}$$

- K.-M. signals:

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ \boxed{1} & \boxed{1} & \boxed{1} & \boxed{-1} \end{bmatrix} \quad A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \dots & \dots \\ 1 \dots 1 & 1 \dots -1 \dots 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$\left. \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\} 2^m$
 $\left. \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\} 2^m$
 $\left. \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right\} 2^m$

- Key property: $x \mapsto xA_m$ is injective on $\{\pm 1\}^{K_m}$, $K_m = \frac{m}{2}2^m + 1$
- Number of users at dimension n : $K \approx \frac{1}{2}n \log_2 n$ (optimal!)

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \ 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ **and** $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \ 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ **and** $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x \ y \ z]$ and

$$[x \ y \ z]\tilde{A}_m = [g \ h]$$

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \ 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

$\begin{array}{c} \uparrow \\ 2^m \\ \downarrow \end{array}$
 $\begin{array}{c} \leftarrow \\ 2^m \\ \rightarrow \end{array}$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ **and** $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x \ y \ z]$ and

$$[x \ y \ z]\tilde{A}_m = [g \ h] \Rightarrow \quad g - h = [x \ y \ z] \begin{pmatrix} 0 \\ 2A_{m-1} \\ 2I_{2^{m-1}} \end{pmatrix}$$

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \dots -1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ and $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x y z]$ and

$$[x y z]\tilde{A}_m = [g h] \Rightarrow g - h = [x y z] \begin{pmatrix} 0 \\ 2A_{m-1} \\ 2I_{2^{m-1}} \end{pmatrix}$$

- $z_1 = 0$, so by adding $(g - h)_1$ to $(g - h)_\ell$ we get:

$$(*) \quad 2z_\ell = (g - h)_1 + (g - h)_\ell - 2y \cdot v_\ell \quad \ell = 2, \dots, 2^{m-1}$$

where v_ℓ is sum of 1-st and ℓ -th column of A_{m-1}

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \dots -1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots -1 \end{bmatrix}$$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ and $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x \ y \ z]$ and

$$[x \ y \ z]\tilde{A}_m = [g \ h] \Rightarrow g - h = [x \ y \ z] \begin{pmatrix} 0 \\ 2A_{m-1} \\ 2I_{2^{m-1}} \end{pmatrix}$$

- $z_1 = 0$, so by adding $(g - h)_1$ to $(g - h)_\ell$ we get:

$$(*) \quad 2z_\ell = (g - h)_1 + (g - h)_\ell - 2y \cdot v_\ell \quad \ell = 2, \dots, 2^{m-1}$$

where v_ℓ is sum of 1-st and ℓ -th column of A_{m-1}

- **Key:** v_ℓ 's entries are $\{0, 2\}$. Take mod 4 of (*) and decode z_ℓ 's !
- Subtracting z_ℓ 's we get system:

$$[x \ y] \begin{pmatrix} A_{m-1} & A_{m-1} \\ A_{m-1} & -A_{m-1} \end{pmatrix} = [g' \ h']$$

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \ 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ and $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x \ y \ z]$ and

$$[x \ y \ z]\tilde{A}_m = [g \ h] \Rightarrow g - h = [x \ y \ z] \begin{pmatrix} 0 \\ 2A_{m-1} \\ 2I_{2^{m-1}} \end{pmatrix}$$

- $z_1 = 0$, so by adding $(g - h)_1$ to $(g - h)_\ell$ we get:

$$(*) \quad 2z_\ell = (g - h)_1 + (g - h)_\ell - 2y \cdot v_\ell \quad \ell = 2, \dots, 2^{m-1}$$

where v_ℓ is sum of 1-st and ℓ -th column of A_{m-1}

- Key:** v_ℓ 's entries are $\{0, 2\}$. Take mod 4 of (*) and decode z_ℓ 's !
- Subtracting z_ℓ 's we get system:

$$[x \ y] \begin{pmatrix} A_{m-1} & A_{m-1} \\ A_{m-1} & -A_{m-1} \end{pmatrix} = [g' \ h'] \Rightarrow xA_{m-1} = \frac{g' + h'}{2}$$

$$\tilde{A}_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & -1 \ 1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

$$A_{m+1} = \begin{bmatrix} A_m & A_m \\ A_m & -A_m \\ \vdots & \vdots \\ 1 \dots 1 & 1 \ -1 \dots 1 \\ \vdots & \vdots \\ 1 \dots 1 & 1 \dots \dots -1 \end{bmatrix}$$

- Want to show: v is decodable from $v\tilde{A}_m$ for any $v \in \{\pm 1\}^{\otimes K_m}$ and $v_{2K_{m-1}+1} = 0$.
- Equivalently: $v \in \{0, 1\}^{\otimes K_m}$ (just use $v \mapsto \frac{1+v}{2}$)
- Let $v = [x \ y \ z]$ and

$$[x \ y \ z]\tilde{A}_m = [g \ h] \Rightarrow g - h = [x \ y \ z] \begin{pmatrix} 0 \\ 2A_{m-1} \\ 2I_{2^{m-1}} \end{pmatrix}$$

- $z_1 = 0$, so by adding $(g - h)_1$ to $(g - h)_\ell$ we get:

$$(*) \quad 2z_\ell = (g - h)_1 + (g - h)_\ell - 2y \cdot v_\ell \quad \ell = 2, \dots, 2^{m-1}$$

where v_ℓ is sum of 1-st and ℓ -th column of A_{m-1}

- Key:** v_ℓ 's entries are $\{0, 2\}$. Take mod 4 of (*) and decode z_ℓ 's !
- Subtracting z_ℓ 's we get system:

$$[x \ y] \begin{pmatrix} A_{m-1} & A_{m-1} \\ A_{m-1} & -A_{m-1} \end{pmatrix} = [g' \ h'] \Rightarrow xA_{m-1} = \frac{g' + h'}{2} \Rightarrow \text{induct}$$

- When user inputs are constrained (to ± 1), can have $K \gg n$
- Total information grows with K : $H(X_1 + \dots + X_K) \sim \frac{1}{2} \log K$.
(This is similar to $\frac{1}{2} \log(1 + KP)$ in GMAC.)
- Lots of smart ideas in MAC codes.

- When user inputs are constrained (to ± 1), can have $K \gg n$
- Total information grows with K : $H(X_1 + \dots + X_K) \sim \frac{1}{2} \log K$.
(This is similar to $\frac{1}{2} \log(1 + KP)$ in GMAC.)
- Lots of smart ideas in MAC codes.
- Information theory structures it all into:

$$C = \bigcup_{X_1, \dots, X_K, U} \{(R_1, \dots, R_K) : R_S \leq I(X_S; Y | X_{S^c}, U)\}$$

- When user inputs are constrained (to ± 1), can have $K \gg n$
- Total information grows with K : $H(X_1 + \dots + X_K) \sim \frac{1}{2} \log K$. (This is similar to $\frac{1}{2} \log(1 + KP)$ in GMAC.)
- Lots of smart ideas in MAC codes.
- Information theory structures it all into:

$$C = \bigcup_{X_1, \dots, X_K, U} \{(R_1, \dots, R_K) : R_S \leq I(X_S; Y | X_{S^c}, U)\}$$

- Similar to how all the smarts (Reed-Muller, BCH, LDPC, Polar, ...) are hidden behind

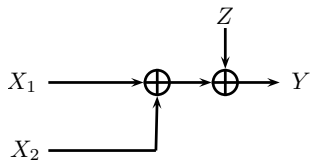
$$C = \max_X I(X; Y)$$

2-user Gaussian MAC

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$

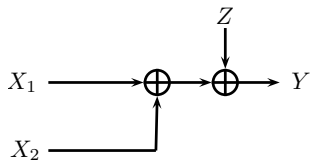


2-user Gaussian MAC

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Evaluating capacity region:

$$R_1 + R_2 \leq I(X_1, X_2; Y) \leq \frac{1}{2} \log(1 + P_1 + P_2)$$

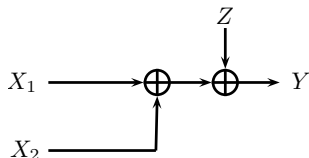
$$R_i \leq I(X_i; Y | X_{\hat{i}}) = I(X_i; X_i + Z) \leq \frac{1}{2} \log(1 + P_i)$$

2-user Gaussian MAC

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

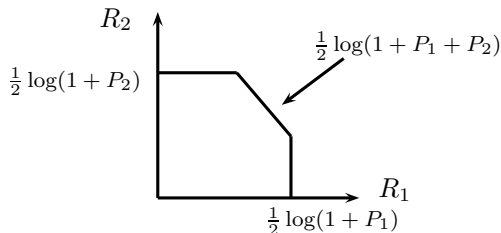
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Evaluating capacity region:

$$R_1 + R_2 \leq I(X_1, X_2; Y) \leq \frac{1}{2} \log(1 + P_1 + P_2)$$

$$R_i \leq I(X_i; Y | X_{\hat{i}}) = I(X_i; X_i + Z) \leq \frac{1}{2} \log(1 + P_i)$$



2-GMAC rates for TDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

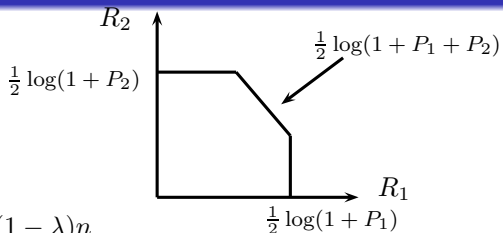
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$

- Here is a TDMA:

- ▶ Partition block: $n = \lambda n + (1 - \lambda)n$

- ▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + P_1)$

- ▶ User 2 sends in $\bar{\lambda}n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + P_2)$



2-GMAC rates for TDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

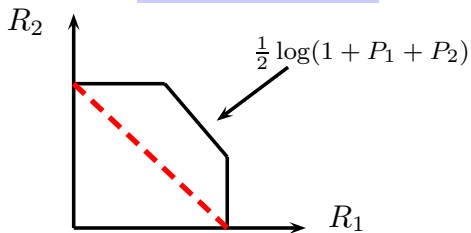
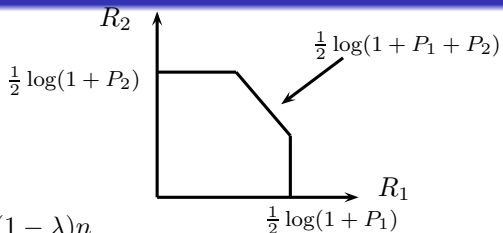
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$

• Here is a TDMA:

▶ Partition block: $n = \lambda n + (1 - \lambda)n$

▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + P_1)$

▶ User 2 sends in $\bar{\lambda}n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + P_2)$



2-GMAC rates for TDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

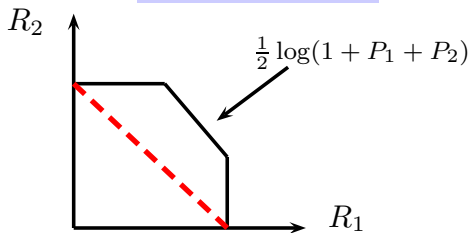
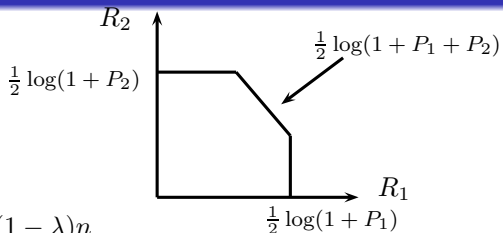
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$

- Here is a TDMA:

- ▶ Partition block: $n = \lambda n + (1 - \lambda)n$

- ▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + P_1)$

- ▶ User 2 sends in $\bar{\lambda}n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + P_2)$



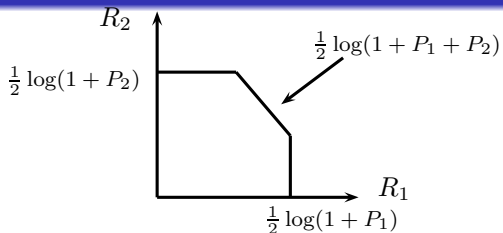
- Note: **low-complexity** decoder – two users are decoded separately.

2-GMAC rates for FDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Here is a FDMA:

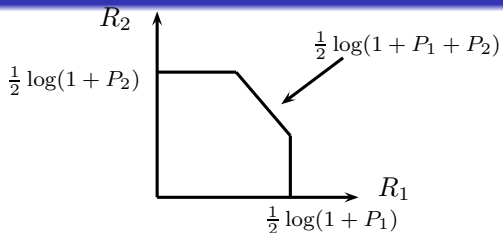
- ▶ Use **Fourier** transform to change n =time to n =frequency.
- ▶ Partition block: $n = \lambda n + (1 - \lambda)n$
- ▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + \frac{P_1}{\lambda})$
- ▶ User 2 sends in $\bar{\lambda} n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + \frac{P_2}{\bar{\lambda}})$

2-GMAC rates for FDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



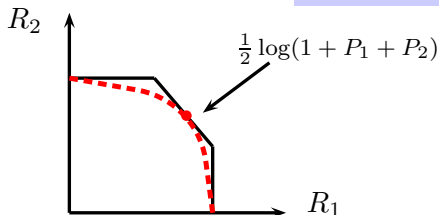
- Here is a FDMA:

- ▶ Use **Fourier** transform to change n =time to n =frequency.

- ▶ Partition block: $n = \lambda n + (1 - \lambda)n$

- ▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + \frac{P_1}{\lambda})$

- ▶ User 2 sends in $\bar{\lambda} n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + \frac{P_2}{\bar{\lambda}})$

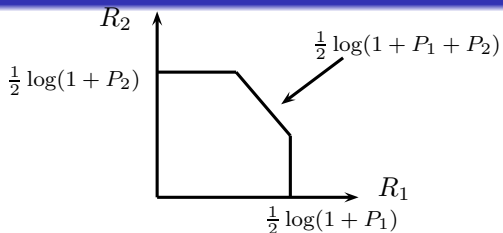


2-GMAC rates for FDMA

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



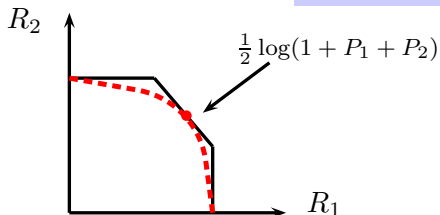
- Here is a FDMA:

- ▶ Use **Fourier** transform to change n =time to n =frequency.

- ▶ Partition block: $n = \lambda n + (1 - \lambda)n$

- ▶ User 1 sends in λn : $R_1 = \frac{\lambda}{2} \log(1 + \frac{P_1}{\lambda})$

- ▶ User 2 sends in $\bar{\lambda} n$: $R_2 = \frac{\bar{\lambda}}{2} \log(1 + \frac{P_2}{\bar{\lambda}})$



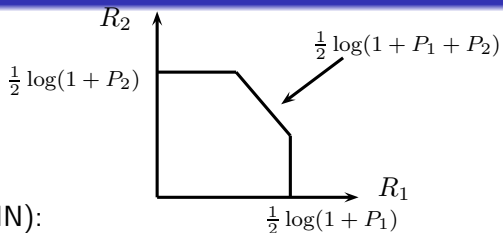
λ^* = $\frac{P_1}{P_1 + P_2}$
achieves **optimal**
sumrate

2-GMAC rates for TIN

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



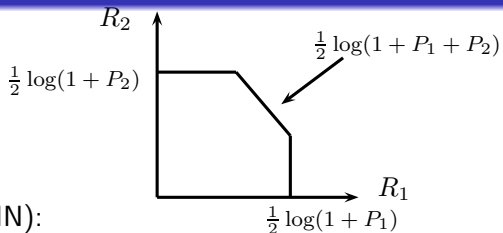
- Treat-interference-as-noise (TIN):
 - ▶ Each user treats the other as noise (**single-user decoders**)
 - ▶ Random coding ensures noise is Gaussian.
 - ▶ Rates: $R_1 = \frac{1}{2} \log(1 + \frac{P_1}{1+P_2})$, $R_2 = \frac{1}{2} \log(1 + \frac{P_2}{1+P_1})$

2-GMAC rates for TIN

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

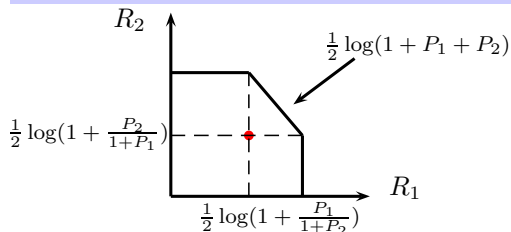
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Treat-interference-as-noise (TIN):

- ▶ Each user treats the other as noise (**single-user decoders**)
- ▶ Random coding ensures noise is Gaussian.

- ▶ Rates: $R_1 = \frac{1}{2} \log(1 + \frac{P_1}{1+P_2}), R_2 = \frac{1}{2} \log(1 + \frac{P_2}{1+P_1})$



- TIN point can be inside/outside TDMA.

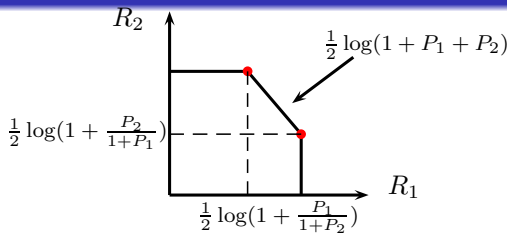
$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$

- Consider a corner point:

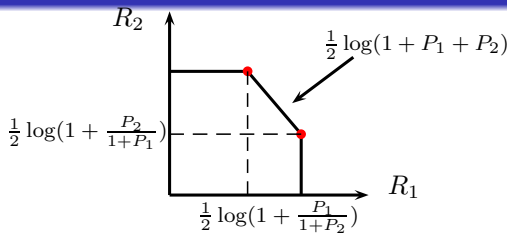
$$R_1 = \frac{1}{2} \log\left(1 + \frac{P_1}{1 + P_2}\right), \quad R_2 = \frac{1}{2} \log(1 + P_2).$$



$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Consider a corner point:

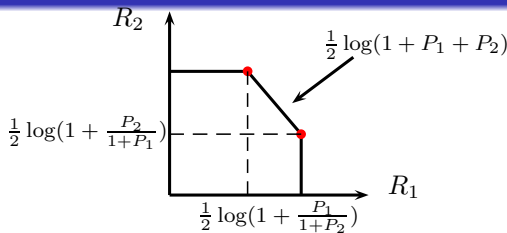
$$R_1 = \frac{1}{2} \log\left(1 + \frac{P_1}{1 + P_2}\right), \quad R_2 = \frac{1}{2} \log(1 + P_2).$$

- User 1 can be decoded by TIN. **But then can subtract it out!**

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

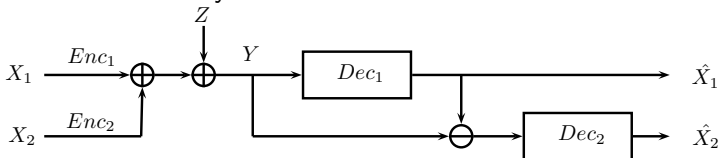
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Consider a corner point:

$$R_1 = \frac{1}{2} \log\left(1 + \frac{P_1}{1 + P_2}\right), \quad R_2 = \frac{1}{2} \log(1 + P_2).$$

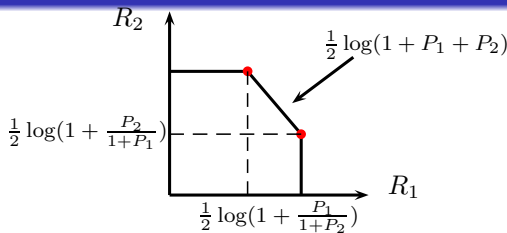
- User 1 can be decoded by TIN. **But then can subtract it out!**



$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

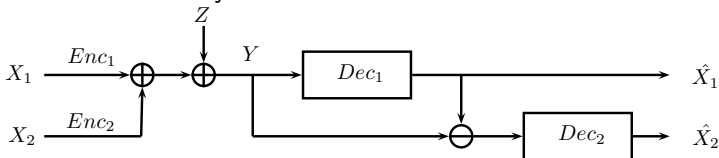
$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Consider a corner point:

$$R_1 = \frac{1}{2} \log\left(1 + \frac{P_1}{1 + P_2}\right), \quad R_2 = \frac{1}{2} \log(1 + P_2).$$

- User 1 can be decoded by TIN. **But then can subtract it out!**



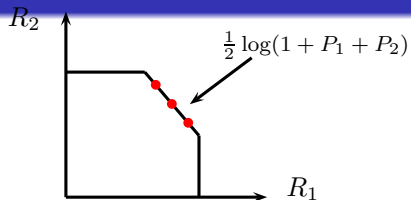
- So far: achieved three optimal points via SU-decoding. **Any more?**

Rate-splitting

$$Y = X_1 + X_2 + Z$$

$$Z \stackrel{iid}{\sim} \mathcal{N}(0, 1)$$

$$\mathbb{E}[(X_1)^2] \leq P_1, \mathbb{E}[(X_2)^2] \leq P_2$$



- Split user 1 into two virtual users 1A and 1B:

$$R_1 = R_{1A} + R_{1B}, \quad P_1 = P_{1A} + P_{1B}$$

- A funny order of decoding:

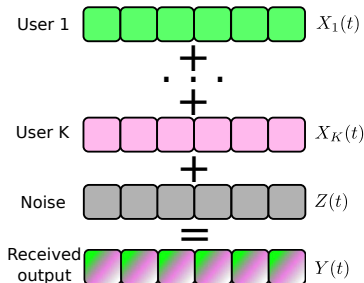
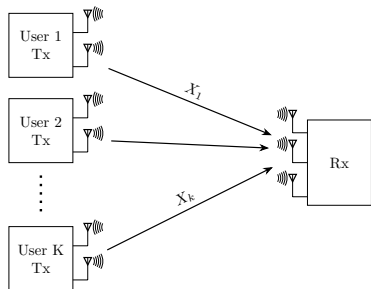
- ▶ Decode X_{1A} via TIN: $R_{1A} = \frac{1}{2} \log\left(1 + \frac{P_{1A}}{1 + P_{1B} + P_2}\right)$
- ▶ Subtract X_{1A} , decode X_2 : $R_2 = \frac{1}{2} \log\left(1 + \frac{P_2}{1 + P_{1B}}\right)$
- ▶ Subtract X_2 , decode X_{1B} : $R_{1B} = \frac{1}{2} \log(1 + P_{1B})$

- Simple check:

$$R_{1A} + R_{1B} + R_2 = \frac{1}{2} \log(1 + P_1 + P_2) \quad \text{sumrate optimal}$$

by varying $P_{1A} + P_{1B} = P_1$ can achieve any point!!

[t]

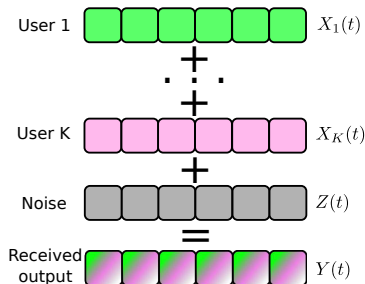
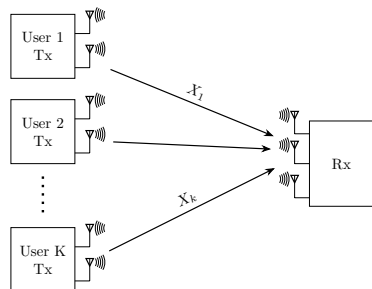


$$Y(t) = X_1(t) + \dots + X_K(t) + Z(t)$$

- Assume equal-power setting $P_i = P$. Capacity region (**sumrate**):

$$\sum_{i=1}^K R_i \leq \frac{1}{2} \log(1 + KP)$$

[t]



$$Y(t) = X_1(t) + \dots + X_K(t) + Z(t)$$

- **single-user** decoders achieve:
 - ▶ FDMA optimal at symmetric point: $R_i = \frac{1}{2K} \log(1 + KP)$
 - ▶ TIN+SIC achieves all vertices.
 - ▶ Rate-Splitting all points of optimal sumrate.
- Is that it? **Let us see...**

- So total capacity:

$$C_{sum} = \frac{1}{2} \log_2(1 + KP) \quad \textit{bit/rdof}$$

growing to ∞ as $K \rightarrow \infty$.

- So total capacity:

$$C_{sum} = \frac{1}{2} \log_2(1 + KP) \quad \text{bit/r dof}$$

growing to ∞ as $K \rightarrow \infty$.

- But at the same time, per-user rate:

$$C_{sym} = \frac{1}{2K} \log_2(1 + KP) \rightarrow 0.$$

- The crucial performance metric: HRH **energy-per-bit**

$$\frac{E_b}{N_0} \triangleq \frac{\text{total energy spent}}{2 \times \text{total \# bits}} = \frac{nKP}{2nC_{sum}}$$

- So total capacity:

$$C_{sum} = \frac{1}{2} \log_2(1 + KP) \quad \text{bit/rdof}$$

growing to ∞ as $K \rightarrow \infty$.

- But at the same time, per-user rate:

$$C_{sym} = \frac{1}{2K} \log_2(1 + KP) \rightarrow 0.$$

- The crucial performance metric: HRH **energy-per-bit**

$$\frac{E_b}{N_0} \triangleq \frac{\text{total energy spent}}{2 \times \text{total \# bits}} = \frac{nKP}{2nC_{sum}}$$

- As $K \rightarrow \infty$:

$$\frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)} \rightarrow \infty \quad !!!$$

- Capacity increases, but each user **works harder** and **moves fewer bits!**

- So total capacity:

$$C_{sum} = \frac{1}{2} \log_2(1 + KP) \quad \text{bit/r dof}$$

growing to ∞ as $K \rightarrow \infty$.

- But at the same time, per-user rate:

$$C_{sym} = \frac{1}{2K} \log_2(1 + KP) \rightarrow 0.$$

- The crucial performance metric: HRH **energy-per-bit**

$$\frac{E_b}{N_0} \triangleq \frac{\text{total energy spent}}{2 \times \text{total \# bits}} = \frac{nKP}{2nC_{sum}}$$

- As $K \rightarrow \infty$:

$$\frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)} \rightarrow \infty \quad !!!$$

- Capacity increases, but each user **works harder** and **moves fewer bits!**
- Correct scaling: $P_{tot} = KP$ should be fixed!

Spectral efficiency vs. $\frac{E_b}{N_0}$

- Studying this tradeoff is the favorite pastime of ComSoc
- Sp. eff. $\rho \triangleq \frac{\text{total \# of data bits}}{\text{total real d.o.f.}}$
- We have:

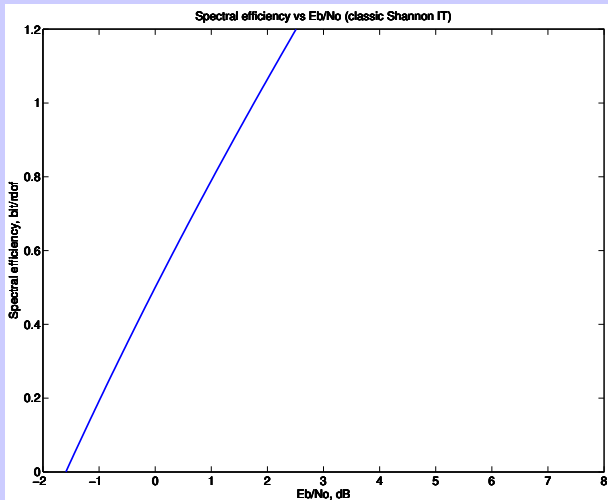
$$\rho = \frac{1}{2} \log(1 + KP), \quad \frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)}$$

- regardless of K : (and any sumrate-optimal arch)

$$\frac{E_b}{N_0} = \frac{2^{2\rho} - 1}{2\rho} \geq -1.59 \text{ dB}$$

- Stud
- Sp.e
- We l

- rega



Spectral efficiency vs. $\frac{E_b}{N_0}$

- Studying this tradeoff is the favorite pastime of ComSoc
- Sp. eff. $\rho \triangleq \frac{\text{total \# of data bits}}{\text{total real d.o.f.}}$
- We have:

$$\rho = \frac{1}{2} \log(1 + KP), \quad \frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)}$$

- regardless of K : (and any sumrate-optimal arch)

$$\frac{E_b}{N_0} = \frac{2^{2\rho} - 1}{2\rho} \geq -1.59 \text{ dB}$$

- Compare to TIN: $\rho = \frac{K}{2} \log_2\left(1 + \frac{P}{1+(K-1)P}\right) \xrightarrow{K \rightarrow \infty} \frac{1}{2 \ln 2} \frac{P_{tot}}{1+P_{tot}}$

Spectral efficiency vs. $\frac{E_b}{N_0}$

- Studying this tradeoff is the favorite pastime of ComSoc
- Sp. eff. $\rho \triangleq \frac{\text{total \# of data bits}}{\text{total real d.o.f.}}$
- We have:

$$\rho = \frac{1}{2} \log(1 + KP), \quad \frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)}$$

- regardless of K : (and any sumrate-optimal arch)

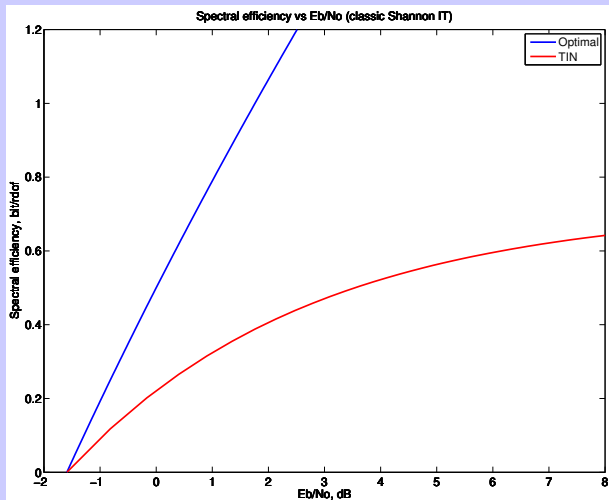
$$\frac{E_b}{N_0} = \frac{2^{2\rho} - 1}{2\rho} \geq -1.59 \text{ dB}$$

- Compare to TIN: $\rho = \frac{K}{2} \log_2\left(1 + \frac{P}{1+(K-1)P}\right) \xrightarrow{K \rightarrow \infty} \frac{1}{2 \ln 2} \frac{P_{tot}}{1+P_{tot}}$

$$\rho = \frac{1}{2 \ln 2} \frac{P_{tot}}{1 + P_{tot}}, \quad \frac{E_b}{N_0} = (1 + P_{tot}) \ln 2$$

Spectral efficiency vs. $\frac{E_b}{N_0}$

- Stud
- Sp.e
- We l
- rega
- Com



Spectral efficiency vs. $\frac{E_b}{N_0}$

- Studying this tradeoff is the favorite pastime of ComSoc
- Sp. eff. $\rho \triangleq \frac{\text{total \# of data bits}}{\text{total real d.o.f.}}$
- We have:

$$\rho = \frac{1}{2} \log(1 + KP), \quad \frac{E_b}{N_0} = \frac{KP}{\log(1 + KP)}$$

- regardless of K : (and any sumrate-optimal arch)

$$\frac{E_b}{N_0} = \frac{2^{2\rho} - 1}{2\rho} \geq -1.59 \text{ dB}$$

- Compare to TIN: $\rho = \frac{K}{2} \log_2\left(1 + \frac{P}{1+(K-1)P}\right) \xrightarrow{K \rightarrow \infty} \frac{1}{2 \ln 2} \frac{P_{tot}}{1+P_{tot}}$

$$\rho = \frac{1}{2 \ln 2} \frac{P_{tot}}{1 + P_{tot}}, \quad \frac{E_b}{N_0} = (1 + P_{tot}) \ln 2$$

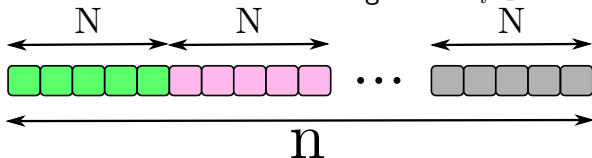
- **IMPORTANT:** $\rho \leq \frac{1}{2 \ln 2} = 0.72$ bit/rdof
- **IMPORTANT:** Essentially optimal for **low sp. eff.**

- Given that TIN is not bad for low sp.eff., let us try to achieve it.
- **Problem:** Per-user rate is $\frac{\rho}{K}$ is very small for a large K .

- Given that TIN is not bad for low sp. eff., let us try to achieve it.
- **Problem:** Per-user rate is $\frac{\rho}{K}$ is very small for a large K . Aside:
 - ▶ For IT Soc: Channel with $C = 0.5$ and channel with $C = 0.001$ are not fundamentally different.
 - ▶ For ComSoc: First channel is OK (turbo/LDPC/polar), second is a nightmare.
 - ▶ **Why?** SNR needs to be brought first to reasonable level.
 - ▶ This is the idea of **modulation**.

- Given that TIN is not bad for low sp. eff., let us try to achieve it.
- **Problem:** Per-user rate is $\frac{\rho}{K}$ is very small for a large K . Aside:
 - ▶ For IT Soc: Channel with $C = 0.5$ and channel with $C = 0.001$ are not fundamentally different.
 - ▶ For ComSoc: First channel is OK (turbo/LDPC/polar), second is a nightmare.
 - ▶ **Why?** SNR needs to be brought first to reasonable level.
 - ▶ This is the idea of **modulation**.
 - ▶ Another issue: how do you do TIN practically? A code with ± 1 entries will create a very non-Gaussian interference!

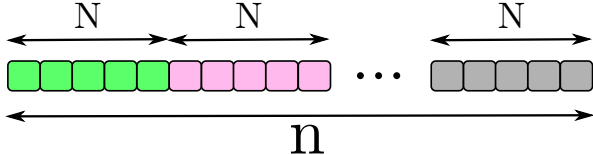
- Given that TIN is not bad for low sp. eff., let us try to achieve it.
- Problem:** Per-user rate is $\frac{\rho}{K}$ is very small for a large K .
- Solution:** each user modulates some N -signature $s_i \in \mathbb{R}^N$



- Think of N -blocks as new super-symbols. Effective channel:

$$Y^N = s_1 B_1 + s_2 B_2 + \dots + s_K B_K + Z^N, \quad \|s_i\| = 1$$

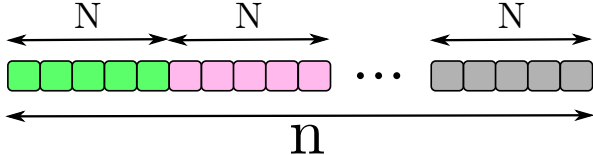
- ▶ Set $\beta = \frac{K}{N}$
- ▶ new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- ▶ new rate: $\frac{\rho N}{K} = \frac{\rho}{\beta}$ in bits / one B -symbol.
- ▶ with proper choice should have $\frac{\rho}{\beta} \sim 1$ as ComSoc likes.



- N -blocks are new super-symbols. Effective channel:

$$Y^N = s_1 B_1 + s_2 B_2 + \dots + s_K B_K + Z^N, \quad \|s_i\| = 1$$

- ▶ Set $\beta = \frac{K}{N}$
- ▶ new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- **Side observation:**
 - ▶ If s_i 's are chosen orthogonally and $K = N$, this is FDMA (hence optimal).
 - ▶ But incurs FBL loss – important when $K \sim n$. But ignore for now.
 - ▶ So why not do so?
 - ▶ Many reasons: E.g. K may vary, but N should be constant.
 - ▶ Requires central distribution of signatures among ACTIVE users.
 - ▶ Also random-like s_i 's would help TIN decoders.



- N -blocks are new super-symbols. Effective channel:

$$Y^N = s_1 B_1 + s_2 B_2 + \dots + s_K B_K + Z^N, \quad \|s_i\| = 1$$

- ▶ Set $\beta = \frac{K}{N}$
- ▶ new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.

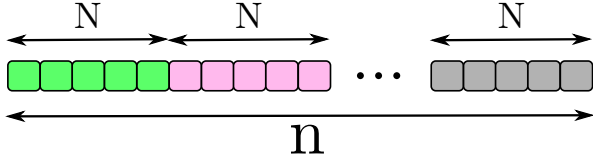
- **Idea 1:** Decode via **matched-filter** + SU decoders:

$$\hat{B}_i = \langle s_i, Y^N \rangle = B_i + \tilde{Z}_i$$

- **Idea 2:** Select s_i randomly. (attractive sys. arch.)
- When s_i 's are random and N large:

$$|\langle s_i, s_j \rangle| \approx \frac{1}{\sqrt{N}} \quad \text{w.h.p.}$$

- So SU-decoder sees effective $\text{SNR} = \frac{NP}{1+(K-1)P} = \frac{P_{tot}}{1+P_{tot}} \frac{1}{\beta}$



- N -blocks are new super-symbols. Effective channel:

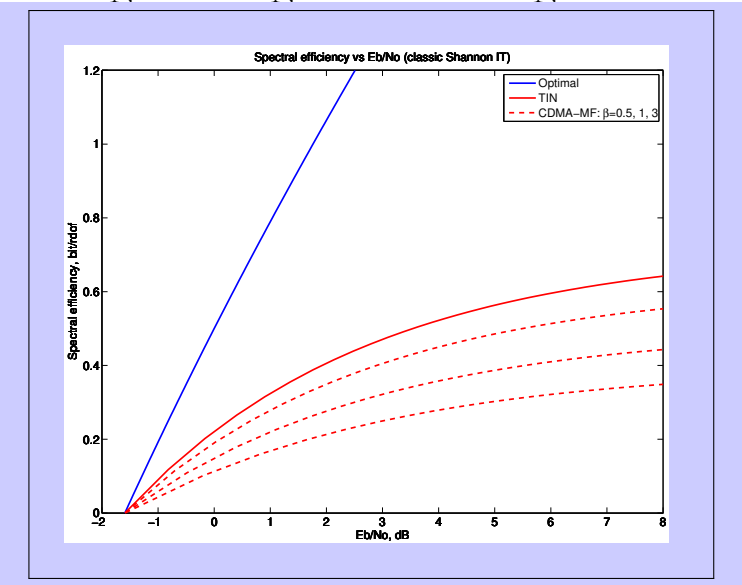
$$Y^N = s_1 B_1 + s_2 B_2 + \dots + s_K B_k + Z^N, \quad \|s_i\| = 1$$

- ▶ Set $\beta = \frac{K}{N}$
- ▶ new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- ▶ random (non-orthogonal) signatures
- ▶ matched-filter + SU-decoder
- End result:

$$\rho_{CDMA} = \frac{\beta}{2} \log_2 \left(1 + \frac{P_{tot}}{1 + P_{tot}} \frac{1}{\beta} \right) \quad \frac{E_b}{N_0} = \frac{P_{tot}}{2\rho_{CDMA}}$$

- ▶ As $\beta \rightarrow \infty$ we approach TIN.
- ▶ So classical CDMA folks (Viterbi...) were only trying to achieve TIN.

- N-b
- ▶
- ▶
- ▶
- ▶
- End



- ▶ As $\beta \rightarrow \infty$ we approach TIN.
- ▶ So classical CDMA folks (Viterbi...) were only trying to achieve TIN.

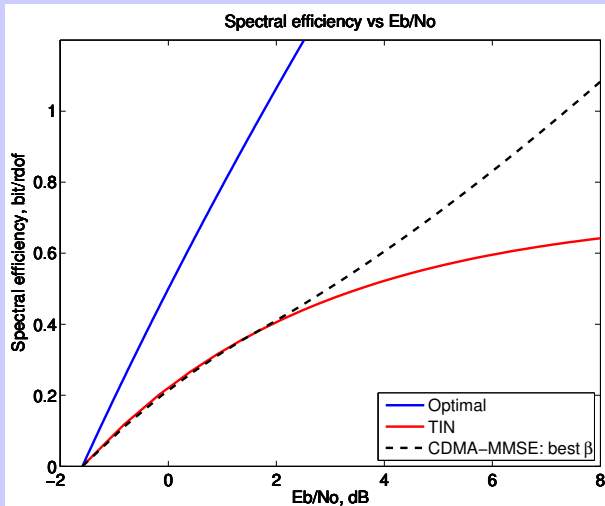
- Set $\beta = \frac{K}{N}$
- new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- random (non-orthogonal) signatures
- matched-filter + SU-decoder

$$\rho_{CDMA} = \frac{\beta}{2} \log_2 \left(1 + \frac{P_{tot}}{1 + P_{tot}} \frac{1}{\beta} \right) \quad \frac{E_b}{N_0} = \frac{P_{tot}}{2\rho_{CDMA}}$$

- multi-user detectors (MUD) improve performance of random-CDMA.
- E.g. MMSE detector yields (Tse-Hanly/Verdú-Shamai formula)

$$\rho_{MMSE} = \frac{\beta}{2} \log_2 \left(1 + P_1 - \frac{1}{4} \mathcal{F} \right), \quad P_1 = \frac{P_{tot}}{\beta}$$

where $\mathcal{F} = (\sqrt{P_1(1 + \sqrt{\beta})^2 + 1} - \sqrt{P_1(1 - \sqrt{\beta})^2 + 1})^2$



CDMA.
a)

- Set $\beta = \frac{K}{N}$
- new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- random (non-orthogonal) signatures
- matched-filter + SU-decoder

$$\rho_{CDMA} = \frac{\beta}{2} \log_2 \left(1 + \frac{P_{tot}}{1 + P_{tot}} \frac{1}{\beta} \right) \quad \frac{E_b}{N_0} = \frac{P_{tot}}{2\rho_{CDMA}}$$

- multi-user detectors (MUD) improve performance of random-CDMA.
- E.g. MMSE detector yields (Tse-Hanly/Verdú-Shamai formula)

$$\rho_{MMSE} = \frac{\beta}{2} \log_2 \left(1 + P_1 - \frac{1}{4} \mathcal{F} \right), \quad P_1 = \frac{P_{tot}}{\beta}$$

where $\mathcal{F} = (\sqrt{P_1(1 + \sqrt{\beta})^2 + 1} - \sqrt{P_1(1 - \sqrt{\beta})^2 + 1})^2$

- Allows to beat TIN's $\rho \leq 0.72$ bit/rdof bottleneck.
- Still, industry converged to **OFDM**: spectrum is too precious.

- Set $\beta = \frac{K}{N}$
- new power-constraint: $\mathbb{E}[B_i^2] \leq NP = \frac{P_{tot}}{\beta}$.
- random (non-orthogonal) signatures
- matched-filter + SU-decoder

$$\rho_{CDMA} = \frac{\beta}{2} \log_2 \left(1 + \frac{P_{tot}}{1 + P_{tot} \beta} \right) \quad \frac{E_b}{N_0} = \frac{P_{tot}}{2\rho_{CDMA}}$$

- multi-user detectors (MUD) improve performance of random-CDMA.
- E.g. MMSE detector yields (Tse-Hanly/Verdú-Shamai formula)

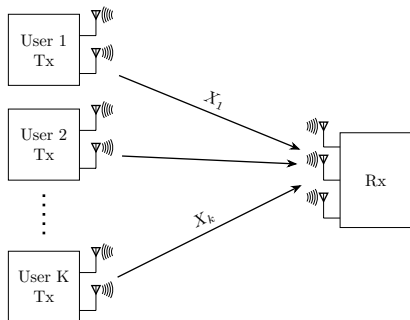
$$\rho_{MMSE} = \frac{\beta}{2} \log_2 \left(1 + P_1 - \frac{1}{4} \mathcal{F} \right), \quad P_1 = \frac{P_{tot}}{\beta}$$

where $\mathcal{F} = (\sqrt{P_1(1 + \sqrt{\beta})^2 + 1} - \sqrt{P_1(1 - \sqrt{\beta})^2 + 1})^2$

- Allows to beat TIN's $\rho \leq 0.72$ bit/r dof bottleneck.
- Still, industry converged to **OFDM**: spectrum is too precious.
- IoT: centralized orthogonalization impossible! **Comeback of MUD?**

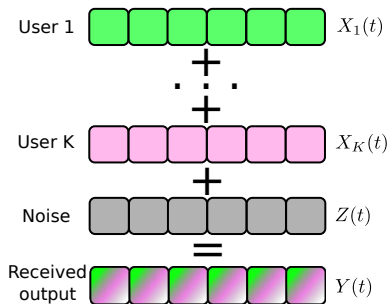
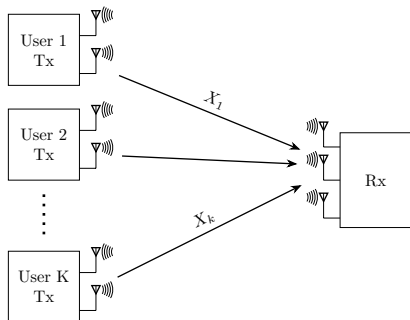
New problems: many users and random-access

The classical model: K-user multiple-access channel



$$Y(t) = X_1(t) + \cdots + X_K(t) + Z(t)$$

The classical model: K-user multiple-access channel



$$Y(t) = X_1(t) + \cdots + X_K(t) + Z(t)$$

- Before: Fix K , let $n \rightarrow \infty$. Few users. Large payloads.
- Now: Huge K . Small payload.
- Random-access: User activity – random, uncoordinated

On number of sensors (user density)

- Key metric: μ in users/rdof

$$\mu = \frac{\# \text{ of active users per frame}}{\text{size of frame}}$$

- K_{tot} sensors sending with period T_{per} (sec) in band B (Hz)

$$\mu = \frac{K_{tot}}{2BT_{per}}$$

- Futuristic example:

- ▶ City of 10^6 .
- ▶ Each house has 10^2 devices.
- ▶ Each dev sends every 10 min, $T_{per} = 600$ s.
- ▶ sub-GHz bandwidth is scarce: ISM $B = 20$ MHz.
- ▶ $\mu \approx 4 \cdot 10^{-3}$.

On number of sensors (user density)

- Key metric: μ in users/rdof

$$\mu = \frac{\# \text{ of active users per frame}}{\text{size of frame}}$$

- K_{tot} sensors sending with period T_{per} (sec) in band B (Hz)

$$\mu = \frac{K_{tot}}{2BT_{per}}$$

- Futuristic example:

- ▶ City of 10^6 .
- ▶ Each house has 10^2 devices.
- ▶ Each dev sends every 10 min, $T_{per} = 600$ s.
- ▶ sub-GHz bandwidth is scarce: ISM $B = 20$ MHz.
- ▶ $\mu \approx 4 \cdot 10^{-3}$.

- Another point of view:

- ▶ Traditional comm: focus on sp. eff. ρ vs $\frac{E_b}{N_0}$. Why?
- ▶ $\frac{\rho B}{K}$ = per-user speed?
- ▶ or is it $\frac{\rho B}{\text{speed}}$ = number of happy users?

Problem 1 “massive”: K and n are comparable.

Relevant asymptotics: $K, n \rightarrow \infty$ with $\frac{K}{n} = \mu$.

Problem 2 “user-centric” probability of error

$$P_e \triangleq \frac{1}{K} \sum_j \mathbb{P}[\hat{X}_j \neq X_j]$$

Problem 3 “random-access”

indistinguishable users (same-codebook), non-asymptotics.

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M \approx 10^2$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M \approx 10^2$ bits.
- Figures of merit: energy-per-bit and user density

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

Problem 1: “massive” number of users

- Number of users $K = \mu n$ scales linearly with blocklength!
- Q: Why scale linearly? A: # of devices waking up \asymp time.
- Q: Ok, but what μ should we look at?

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M \approx 10^2$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

Problem 1: “massive” number of users

- Number of users $K = \mu n$ **scales linearly with blocklength!**
- **Q:** Why scale linearly? **A:** # of devices waking up \propto time.
- **Q:** Ok, but what μ should we look at?
A: $\mu \sim 10^{-3}$. Here is why:
 - ▶ City of 10^6 .
 - ▶ Each house has 10^2 devices.
 - ▶ Each dev sends 1-10 times/hour.
 - ▶ sub-GHz bandwidth is scarce, unlikely to ever get > 20 MHz.
 - ▶ $\Rightarrow \frac{K}{n} \approx 10^{-3} \dots 10^{-2}$. This relation is unlikely to change soon.

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

Problem 1: “massive” number of users

- Number of users $K = \mu n$ **scales linearly with blocklength!**
- [Chen-Chen-Guo'17]: Fix per-user power to P (i.e. codeword $\|c\|_2^2 \leq nP$), then

$$\log M_{user}^*(K = \mu n, n, P) \approx \frac{1}{2\mu} \log(1 + \mu n P)$$

- Note: this corresponds to $\frac{E_b}{N_0} \rightarrow \infty$.
- **Our work:** What about finite $\frac{E_b}{N_0}$?

Problem 1 “massive”: K and n are comparable



Relevant asymptotics: $K, n \rightarrow \infty$ with $\frac{K}{n} = \mu$.

Problem 2 “user-centric” probability of error

$$P_e \triangleq \frac{1}{K} \sum_j \mathbb{P}[\hat{X}_j \neq X_j]$$

Problem 3 “random-access”

indistinguishable users (same-codebook), non-asymptotics.

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

- Regime: $K = \mu n$, $n \rightarrow \infty$.

Problem 2: “user-centric” prob. of error

- For finite $\frac{E_b}{N_0}$ we have (**Why?** See next...)

$$\mathbb{P}[W_1 = \hat{W}_1, \dots, W_K = \hat{W}_K] \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

- \Rightarrow **NEED** to switch to per-user P_e , **PUPE**:

$$P_e = \frac{1}{K} \sum_{i=1}^K \mathbb{P}[W_i \neq \hat{W}_i]$$

Theorem

Suppose K users send *one bit each* with finite energy \mathcal{E} over the GMAC (with *arbitrary* n): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

And, thus, *classical probability of error* $\rightarrow 1$ as $K \rightarrow \infty$.

Theorem

Suppose K users send **one bit each** with finite energy \mathcal{E} over the GMAC (with **arbitrary n**): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

And, thus, **classical probability of error** $\rightarrow 1$ as $K \rightarrow \infty$.

Proof:

- WLOG can assume: $Y = \sum c_i W_i + Z$, where $c_i \in \mathbb{R}^n$ and $W_i \sim \text{Ber}(1/2)$.
- **Genie**: Reveal vector of W_i 's to within Hamming-distance 1.
- New problem: **See** $Y = c_U + Z$, $U \sim [K]$. **Goal**: find U .

Theorem

Suppose K users send **one bit each** with finite energy \mathcal{E} over the GMAC (with **arbitrary** n): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

And, thus, **classical probability of error** $\rightarrow 1$ as $K \rightarrow \infty$.

Proof:

- WLOG can assume: $Y = \sum c_i W_i + Z$, where $c_i \in \mathbb{R}^n$ and $W_i \sim \text{Ber}(1/2)$.
- **Genie**: Reveal vector of W_i 's to within Hamming-distance 1.
- New problem: **See** $Y = c_U + Z$, $U \sim [K]$. **Goal**: find U .
- Fano + Capacity calculation:

$$\mathbb{P}[U = \hat{U}] \log K - \log 2 \leq I(c_U; Y)$$

Theorem

Suppose K users send **one bit each** with finite energy \mathcal{E} over the GMAC (with **arbitrary** n): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

And, thus, **classical probability of error** $\rightarrow 1$ as $K \rightarrow \infty$.

Proof:

- WLOG can assume: $Y = \sum c_i W_i + Z$, where $c_i \in \mathbb{R}^n$ and $W_i \sim \text{Ber}(1/2)$.
- **Genie**: Reveal vector of W_i 's to within Hamming-distance 1.
- New problem: **See** $Y = c_U + Z$, $U \sim [K]$. **Goal**: find U .
- Fano + Capacity calculation:

$$\mathbb{P}[U = \hat{U}] \log K - \log 2 \leq I(c_U; Y) \leq \frac{n}{2} \log \left(1 + \frac{\mathcal{E}}{n} \right) \leq \frac{\log e}{2} \mathcal{E}$$

Theorem (AWGN)

Suppose K users send one bit each with finite energy \mathcal{E} over the GMAC (with arbitrary n): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

Same proof:

Theorem (BSC)

Let G be a $K \times n$ generating matrix with $\leq \mathcal{E}$ ones per row. Then over $BSC(\delta)$ and all n :

$$1 - \mathbb{P}[\text{block error}] \leq \frac{d(\delta \| \bar{\delta}) \mathcal{E} + \log 2}{\log K}$$

Theorem (AWGN)

Suppose K users send one bit each with finite energy \mathcal{E} over the GMAC (with arbitrary n): $Y^n = \sum_{i=1}^K X_i + Z^n$. Then we have

$$\mathbb{P}[X_1 = \hat{X}_1, \dots, X_K = \hat{X}_K] \leq \frac{\mathcal{E} \frac{\log e}{2} + \log 2}{\log K}.$$

Same proof:

Theorem (BSC)

Let G be a $K \times n$ generating matrix with $\leq \mathcal{E}$ ones per row. Then over $BSC(\delta)$ and all n :

$$1 - \mathbb{P}[\text{block error}] \leq \frac{d(\delta \| \bar{\delta}) \mathcal{E} + \log 2}{\log K}$$

Puzzle: Genie + Fano method fails for BEC! (Proof by induction works.)

- Per-user probability of error as

$$P_e = \frac{1}{K} \sum_{i=1}^K \mathbb{P}[W_i \neq \hat{W}_i].$$

- Let's forget about $K = \mu n$ and consider ...
- Classical regime:** K -fixed, power P fixed, $n \rightarrow \infty$. Symmetric capacity

$$C_{sym}(K) = \frac{1}{2K} \log(1 + KP).$$

- But no strong converse (!)

$$C_{sym,\epsilon}(K) > C_{sym}(K - 1) \quad \forall \epsilon \gtrsim \frac{1 + \log_e K}{K}$$

- Lesson:** When PUPE above $\frac{\log K}{K}$, far from usual GMAC+JPE.

- Let $C_{sym,\epsilon}(K)$ be the max achievable symmetric rate (K -fixed, $n \rightarrow \infty$) under PUPE

$$\frac{1}{K} \sum_{i=1}^K \mathbb{P}[W_i \neq \hat{W}_i] \leq \epsilon.$$

K -user GMAC under PUPE: no strong converse

- Let $C_{sym,\epsilon}(K)$ be the max achievable symmetric rate (K -fixed, $n \rightarrow \infty$) under PUPE

$$\frac{1}{K} \sum_{i=1}^K \mathbb{P}[W_i \neq \hat{W}_i] \leq \epsilon.$$

Theorem (P.-Telatar'16)

We have: $C_{sym,\epsilon}(K, \epsilon) = \begin{cases} \frac{1}{2K} \log(1 + KP), & \epsilon < 1/K \\ \geq \frac{1}{2(K-1)} \log(1 + (K-1)P), & \epsilon \gtrsim \frac{1+\log_e K}{K} \end{cases}$

- Note that sequence: $\frac{1}{2K} \log(1 + KP)$ is monotonically decreasing.
- First part: by union bound PUPE $\leq \epsilon$ implies JPE $\leq K\epsilon +$ strong-converse for GMAC.

K -user GMAC under PUPE: no strong converse

- Let $C_{sym,\epsilon}(K)$ be the max achievable symmetric rate (K -fixed, $n \rightarrow \infty$) under PUPE

$$\frac{1}{K} \sum_{i=1}^K \mathbb{P}[W_i \neq \hat{W}_i] \leq \epsilon.$$

Theorem (P.-Telatar'16)

We have: $C_{sym,\epsilon}(K, \epsilon) = \begin{cases} \frac{1}{2K} \log(1 + KP), & \epsilon < 1/K \\ \geq \frac{1}{2(K-1)} \log(1 + (K-1)P), & \epsilon \gtrsim \frac{1 + \log_e K}{K} \end{cases}$

- Note that sequence: $\frac{1}{2K} \log(1 + KP)$ is monotonically decreasing.
- First part: by union bound PUPE $\leq \epsilon$ implies JPE $\leq K\epsilon +$ strong-converse for GMAC.
- Second part: Choose codebooks for symmetric-rate point of $(K-1)$ -GMAC
- Each user sends 0 w.p. ϵ . Then w.p. $1 - (1 - \epsilon)^K$ only $(K-1)$ are active.

Problem 1 “massive”: K and n are comparable



Relevant asymptotics: $K, n \rightarrow \infty$ with $\frac{K}{n} = \mu$.

Problem 2 “user-centric” probability of error



$$P_e \triangleq \frac{1}{K} \sum_j \mathbb{P}[\hat{X}_j \neq X_j]$$

Problem 3 “random-access”

indistinguishable users (same-codebook), non-asymptotics.

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

- Regime: $K = \mu n$, $n \rightarrow \infty$.
- **PUPE** definition: $P_e \triangleq \frac{1}{K} \sum_{j=1}^K \mathbb{P}[X_j \neq \hat{X}_j]$.

So what are the results?

Recap: MAC setting and performance metrics

- Perfectly synchronized K -user Gaussian MAC with blocklength n
- Each user transmits $\log_2 M$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

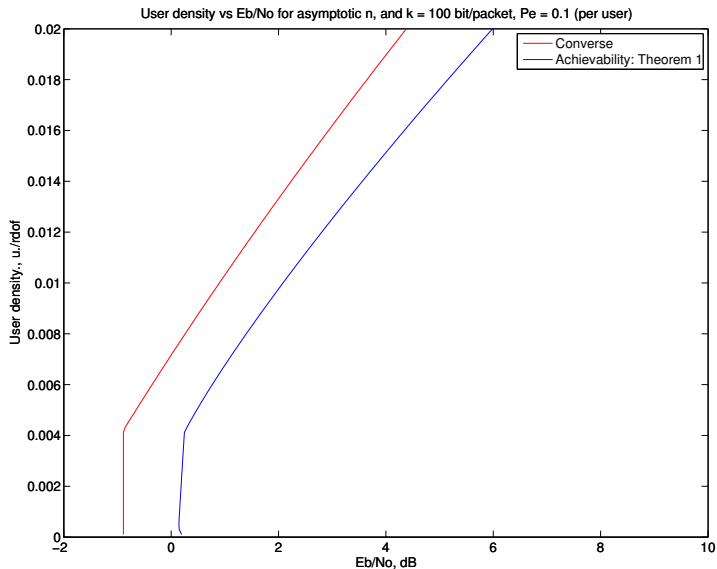
- Regime: $K = \mu n$, $n \rightarrow \infty$.
- **PUPE** definition: $P_e \triangleq \frac{1}{K} \sum_{j=1}^K \mathbb{P}[X_j \neq \hat{X}_j]$.

So what are the results?

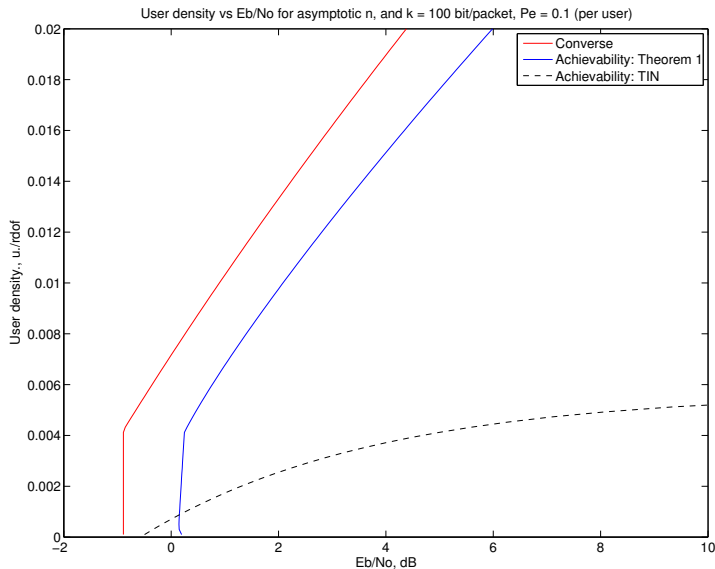
- **IoT regime:** $K = \mu n$, $nP = \text{const}$, $n \rightarrow \infty$.
- **Main result 1:** Asymptotic tradeoff $\mu^*(\epsilon, M, \frac{E_b}{N_0})$ satisfies

$$\mu_{\text{conv}} \leq \mu^* \leq \mu_{\text{ach}}$$

User density vs. Energy-per-bit: best bounds



User density vs. Energy-per-bit: CDMA (w/o MUD)



- Each user transmits $\log_2 M$ bits.
- Figures of merit: **energy-per-bit** and **user density**

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\|X^n\|^2]}{2 \log_2 M}$$

$$\mu \triangleq \frac{K}{n}$$

- $K = \mu n$, $E \triangleq nP = \text{const}$, $n \rightarrow \infty$.
- **Main result 1:** Asymptotic tradeoff $\mu^*(\epsilon, M, \frac{E_b}{N_0})$ satisfies

$$\mu_{conv} \leq \mu^* \leq \mu_{ach}$$

- Converse 1 (**Fano**): $(1 - \epsilon)\mu \log M \leq \frac{1}{2} \log(1 + K_a P) + \mu h(\epsilon)$
- Converse 2 (**PPV'11**): low $\frac{E_b}{N_0}$ **requires (!)** $\log M \gg 1$

$$\log_e M \lesssim \frac{E}{2} - \sqrt{EQ}^{-1}(\epsilon)$$

K -user GMAC with a per-user P_e

- Each user transmits $\log_2 M$ bits.
- Figures of merit

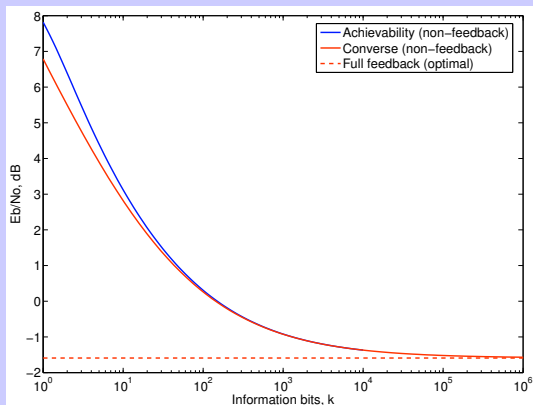
- $K = \mu n$

- Main results

- Converse

- Converse

[P.-Poor-Verdú'11]



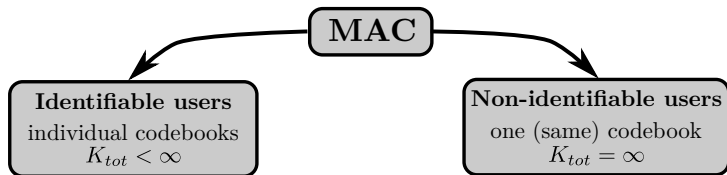
Problem 3: Information theory of random-access

It's a mess...

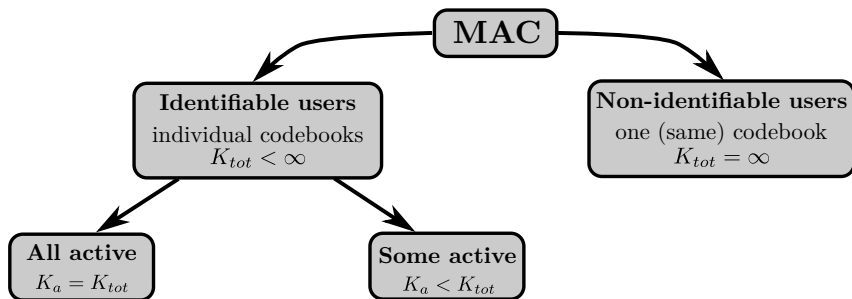
It's a mess...

- Channel model: collision vs. additive
- Noise model: noiseless, stochastic or worst-case
- Coding with or without feedback (as in CSMA)
- Probability of error: zero, vanishing or fixed > 0 .
- Probability of error: per-user vs all-users
- User activity: always-on vs sporadic
- finite blocklength vs $n \rightarrow \infty$
- Various asymptotics: $K = \text{const}, n \rightarrow \infty$ vs both $K, n \rightarrow \infty$

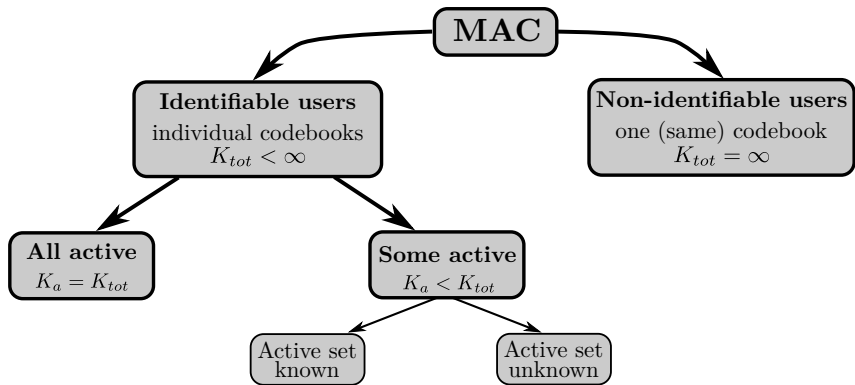
Classification by user activity

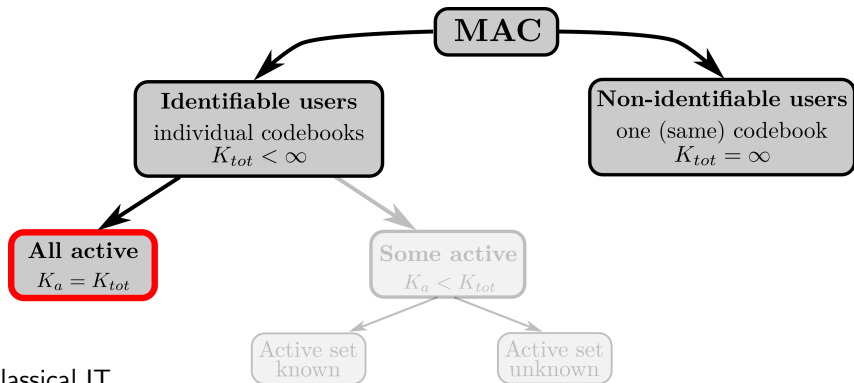


Classification by user activity



Classification by user activity





- Classical IT

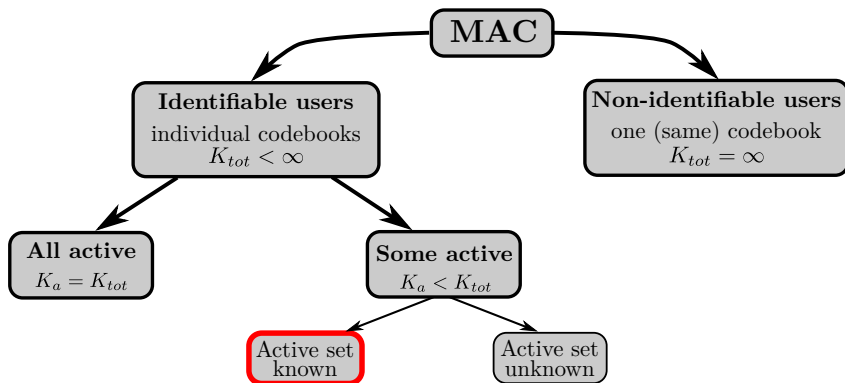
[Liao'72],[Ahlsvede'73]

- Orthogonal schemes TDMA/FDMA

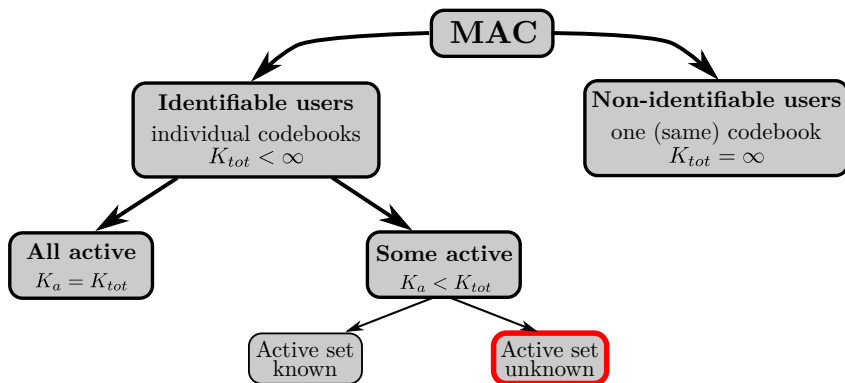
- Rate splitting [Rimoldi-Urbanke'99]

- Finite blocklength [MolavianJazi-Laneman'14-16]

- Many-user [Chen-Guo'14]

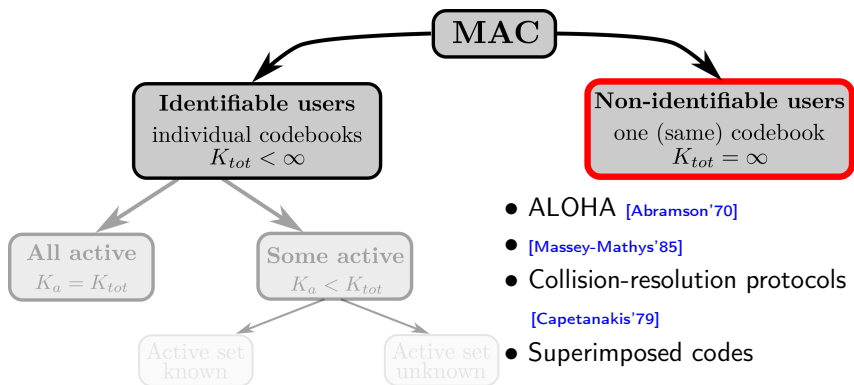


- Non-orthogonal CDMA, MUD
- Randomly-spread CDMA
 - [Tse-Hanly'99], [Verdú-Shamai'99]
- [Mathys'90]
- LDS, SCMA



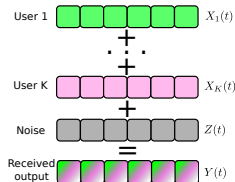
- Non-orthogonal CDMA, MUD
- Randomly-spread CDMA
[Tse-Hanly'99], [Verdú-Shamai'99]
- [Mathys'90]
- LDS, SCMA

- Many-access [Chen-Chen-Guo'17]
- Blind-detection for CDMA
- [BarDavid-Plotnik-Rom'93]
- conflict-avoiding codes
[Bassalygo-Pinsker'83], B.Tsybakov



- ALOHA [Abramson'70]
- [Massey-Mathys'85]
- Collision-resolution protocols [Capetanakis'79]
- Superimposed codes [Ericson-Gyorfi'88] [Furedi-Ruszinkó'99]
- B_r -codes [Dyachkov-Rykov'81]
- Coded Slotted ALOHA [Casini et al'07],[Liva'11]
- Compressed sensing [Jin-Kim-Rao'11]

Key definition: random-access code



Definition (P.'17)

$f : [M] \rightarrow \mathbb{R}^n$ is a **random-access code** for K_a users if \exists **list- K_a decoder** g s.t.

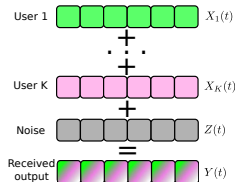
$$\mathbb{P}[W_j \notin g(f(W_1) + \dots + f(W_{K_a}) + Z)] \leq \epsilon \quad \forall j \in [K_a]$$

where $W_i \stackrel{iid}{\sim} \text{Unif}[M]$.

For $\epsilon = 0$ this was studied:

- Noiseless channels: B_r -codes [Dyachkov-Rykov'81]
- Worst-case noise: superimposed codes [Ericson-Gyorfi'88, Furedi-Ruszinkó'99]

Key definition: random-access code



Definition (P.'17)

$f : [M] \rightarrow \mathbb{R}^n$ is a **random-access code** for K_a users if \exists **list- K_a decoder** g s.t.

$$\mathbb{P}[W_j \notin g(f(W_1) + \dots + f(W_{K_a}) + Z)] \leq \epsilon \quad \forall j \in [K_a]$$

where $W_i \stackrel{iid}{\sim} \text{Unif}[M]$.

For $\epsilon > 0$ this is:

- Just compressed sensing: $Y = X\beta + Z$, X is K_a -out-of- M sparse.
- \Rightarrow studied by many, but not w.r.t. $\frac{E_b}{N_0}$ and not with $M = 2^{\Theta(n)}$.

Same-codebook codes = compressed sensing

- random-access = all users share same codebook
- ... obviously decoding is upto permutation of users
- **New problems:** capacity/error-exponent/zero-error capacity
- Equivalent to compressed-sensing [\[Jin-Kim-Rao'11\]](#)

Same-codebook codes = compressed sensing

- random-access = all users share same codebook
- ... obviously decoding is upto permutation of users
- **New problems:** capacity/error-exponent/zero-error capacity
- Equivalent to compressed-sensing [Jin-Kim-Rao'11]
- Let same-codebook (column) vectors be c_1, \dots, c_M .

$$X = (c_1 \mid \cdots \mid c_M)$$

- Let $\beta \in \{0, 1\}^M$ with $\beta_j = 1$ if codeword j was transmitted
- Then the problem is:

$$Y = X\beta + Z, \quad \text{Goal: } \mathbb{E}[\|\beta - \hat{\beta}(Y)\|] \rightarrow \min$$

(linear regression with sparsity $\|\beta\|_0 = K_a$ aka comp.sensing).

Same-codebook codes = compressed sensing

- random-access = all users share same codebook
- ... obviously decoding is upto permutation of users
- **New problems:** capacity/error-exponent/zero-error capacity
- Equivalent to compressed-sensing [Jin-Kim-Rao'11]
- Let same-codebook (column) vectors be c_1, \dots, c_M .

$$X = (c_1 \mid \cdots \mid c_M)$$

- Let $\beta \in \{0, 1\}^M$ with $\beta_j = 1$ if codeword j was transmitted
- Then the problem is:

$$Y = X\beta + Z, \quad \text{Goal: } \mathbb{E}[\|\beta - \hat{\beta}(Y)\|] \rightarrow \min$$

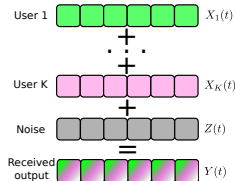
(linear regression with sparsity $\|\beta\|_0 = K_a$ aka comp.sensing).

- The famous $n \sim 2K_a \log_e M$ is just **TIN**:

$$\log_e M \approx \frac{n}{2} \log_e \left(1 + \frac{P}{1 + (K_a - 1)P} \right) \approx \frac{n}{2K_a}$$

So all the L_1 (LASSO) frenzy is just to achieve TIN (hehe...)

Key definition: random-access code



Definition (P.'17)

$f : [M] \rightarrow \mathbb{R}^n$ is a **random-access code** for K_a users if \exists **list- K_a decoder** g s.t.

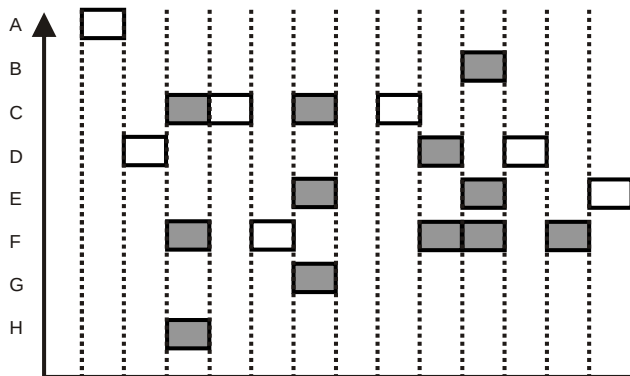
$$\mathbb{P}[W_j \notin g(f(W_1) + \dots + f(W_{K_a}) + Z)] \leq \epsilon \quad \forall j \in [K_a]$$

where $W_i \stackrel{iid}{\sim} \text{Unif}[M]$.

This definition is **answer to many prayers**, but ...

Bad news: Asymptotics of $K_a = \mu n$, $n \rightarrow \infty$ is **nonsense**.

Prototypical random-access code: ALOHA



Slotted ALOHA protocol (shaded slots indicate collision)

- n -frame is partitioned into $L = \frac{n}{n_1}$ subframes of length n_1
- Each of K_a users places his n_1 -codeword into a random subframe.
- Per-user error: $P_e \approx \mathbb{P}[\text{Bino}(K_a - 1, \frac{1}{L}) > 0] \approx \frac{K_a}{L} e^{-\frac{K_a}{L}}$

Main result 2: random-coding bound

II. RANDOM CODING BOUND

Theorem 1. Fix $P' < P$. There exists an (M, n, ϵ) random-access code for K_a -user GMAC satisfying power-constraint P and

$$\epsilon \leq \sum_{t=1}^{K_a} \frac{t}{K_a} \min(p_t, q_t) + p_0, \quad (3)$$

where

$$p_0 = \frac{\binom{K_a}{2}}{M} + K_a \mathbb{P} \left[\frac{1}{n} \sum_{j=1}^n Z_j^2 > \frac{P}{P'} \right], \quad (4)$$

$$p_t = e^{-nE(t)}, \quad (5)$$

$$E(t) = \max_{0 \leq \rho, \rho_1 \leq 1} -\rho \rho_1 t R_1 - \rho_1 R_2 + E_0(\rho, \rho_1)$$

$$E_0 = \rho_1 a + \frac{1}{2} \log(1 - 2b\rho_1)$$

$$a = \frac{\rho}{2} \log(1 + 2P't\lambda) + \frac{1}{2} \log(1 + 2P't\mu) \quad (6)$$

$$b = \rho\lambda - \frac{\mu}{1 + 2P't\mu}, \quad \mu = \frac{\rho\lambda}{1 + 2P't\lambda} \quad (7)$$

$$\lambda = \frac{P't - 1 + \sqrt{D}}{4(1 + \rho_1\rho)P't}, \quad (8)$$

$$D = (P't - 1)^2 + 4P't \frac{1 + \rho\rho_1}{1 + \rho} \quad (9)$$

$$R_1 = \frac{1}{n} \log M - \frac{1}{n} \log(t!) \quad (9)$$

$$R_2 = \frac{1}{n} \log \binom{K_a}{t} \quad (10)$$

$$q_t = \inf_{\gamma} \mathbb{P}[I_t \leq \gamma] + \exp\{n(R_1 + R_2) - \gamma\}$$

Remark: For classical regime K_a -fixed, $n \rightarrow \infty$ and $\epsilon \rightarrow 0$

$$C_{\text{random-access}}(K_a) = \frac{1}{2K_a} \log(1 + K_a P).$$

Random-coding achievability bound

- Generate M codewords: $c_i \sim \mathcal{N}(0, P)^{\otimes n}$.
- WLOG, users send c_1, c_2, \dots, c_{K_a} .
- Decoder sees

$$Y = c_1 + \dots + c_{K_a} + Z$$

- Define sum-codewords $c(S) \triangleq \sum_{i \in S} c_i$
- ML-decoder (not optimal!)

$$\hat{S} = \arg \min_S \|c(S) - Y\|.$$

- Error-analysis:

$$P_e \leq \sum_{t=1}^{K_a} \frac{t}{K_a} \mathbb{P}[t\text{-misguessed}]$$

$$\mathbb{P}[t\text{-misguessed}] \leq \mathbb{P} \left[\bigcup_{S \in \binom{K_a}{t}} \bigcup_{S' \in \binom{M-K_a}{t}} \|c(S) - c(S') + Z\| \leq \|Z\| \right]$$

Random-coding achievability bound

- Generate M codewords: $c_1, \dots, c_M \in \mathcal{C} \subset \mathcal{D}^{\otimes n}$

Analysis I:

- Condition on Z, c_1, \dots, c_{K_a}
 - Use Chernoff + Gallager ρ -trick for $\mathbb{P}[\cup_{S'} \dots | c_1^{K_a}, Z]$
 - Use another Gallager ρ -trick for $\mathbb{P}[\cup_S \dots | Z]$
 - Finally take expectation over Z
- ML-decoder (not optimal!)

$$\hat{S} = \arg \min_S \|c(S) - Y\|.$$

- Error-analysis:

$$P_e \leq \sum_{t=1}^{K_a} \frac{t}{K_a} \mathbb{P}[t\text{-misguessed}]$$

$$\mathbb{P}[t\text{-misguessed}] \leq \mathbb{P} \left[\bigcup_{S \in \binom{K_a}{t}} \bigcup_{S' \in \binom{M-K_a}{t}} \|c(S) - c(S') + Z\| \leq \|Z\| \right]$$

Random-coding achievability bound

- Generate M codewords: $c_1, \dots, c_M \in \mathcal{D}^{\otimes n}$

Analysis I:

- Condition on Z, c_1, \dots, c_{K_a}
- Use Chernoff + Gallager ρ -trick for $\mathbb{P}[\cup_{S'} \dots | c_1^{K_a}, Z]$
- Use another Gallager ρ -trick for $\mathbb{P}[\cup_S \dots | Z]$
- Finally take expectation over Z
- ML-decoder (not optimal!)

Analysis II:

- Define information density appropriately
- Use Feinstein's trick to bound

$$\mathbb{P}[\cup_S \cup_{S'} \dots] \leq \mathbb{P}[i_{\min}(X_1^{K_a}; Y) < \gamma] + \binom{K_a}{t} \binom{M}{t} e^{-\gamma}$$

$$i_{\min} = \min_S i_t(c(S); Y | c(S^c))$$

- $i_{\min} \approx$ max of Gaussian process indexed by t -subsets of $[K_a]$

$$\mathbb{P}[t\text{-misguessed}] \leq \mathbb{P} \left[\bigcup_{S \in \binom{[K_a]}{t}} \bigcup_{S' \in \binom{[M-K_a]}{t}} \|c(S) - c(S') + Z\| \leq \|Z\| \right]$$

Random-coding achievability bound

- Generate M codewords: $c_1, \dots, c_M \in \mathcal{C} \otimes n$

Analysis I:

- Condition on Z, c_1, \dots, c_{K_a}
- Use Chernoff + Gallager ρ -trick for $\mathbb{P}[\cup_{S'} \dots | c_1^{K_a}, Z]$
- Use another Gallager ρ -trick for $\mathbb{P}[\cup_S \dots | Z]$
- Finally take expectation over Z
- ML-decoder (not optimal!)

Analysis II:

- Define information density appropriately

- Use Feinstein's trick to bound

$$\mathbb{P}[\cup_S \cup_{S'} \dots] \leq \mathbb{P}[i_{\min}(X_1^{K_a}; Y) < \gamma] + \binom{K_a}{t} \binom{M}{t} e^{-\gamma}$$

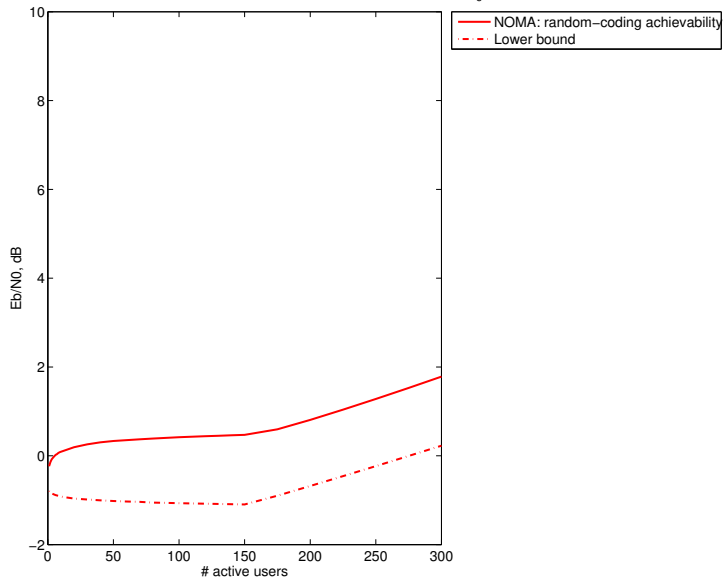
$$i_{\min} = \min_S i_t(c(S); Y | c(S^c))$$

- $i_{\min} \approx$ max of Gaussian process indexed by t -subsets of $[K_a]$

Classical IT: term $\mathbb{P}[t \text{ misguessed}] \leq \mathbb{P}[\dots \|Z\| \leq \|Z\|]$ goes $\rightarrow 0$ if $I(X_S; Y | X_{S^c}) > \sum_{i \in S} R_i$

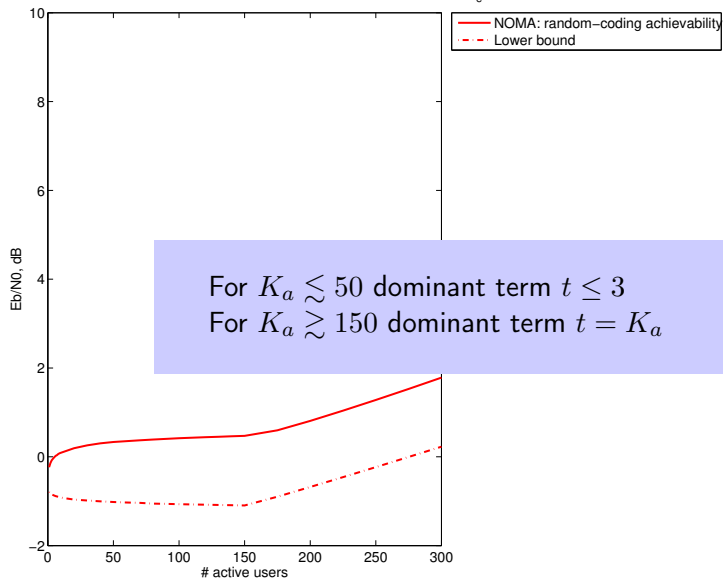
Numerical evaluation

Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$



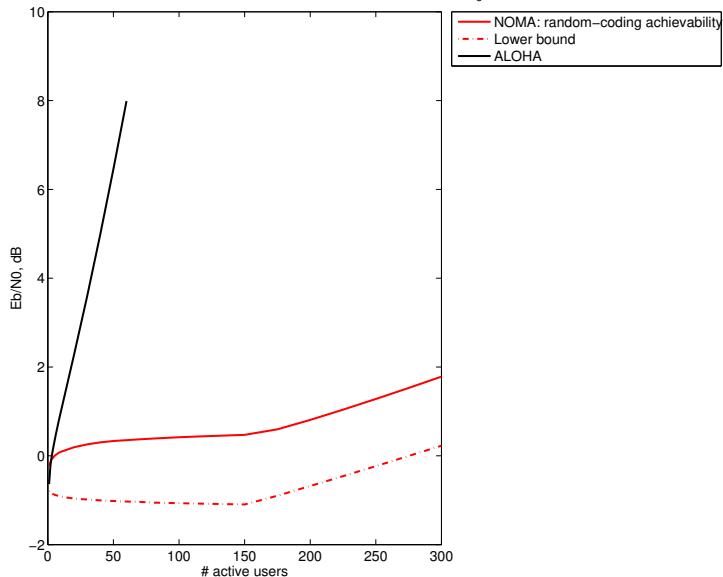
Numerical evaluation

Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$

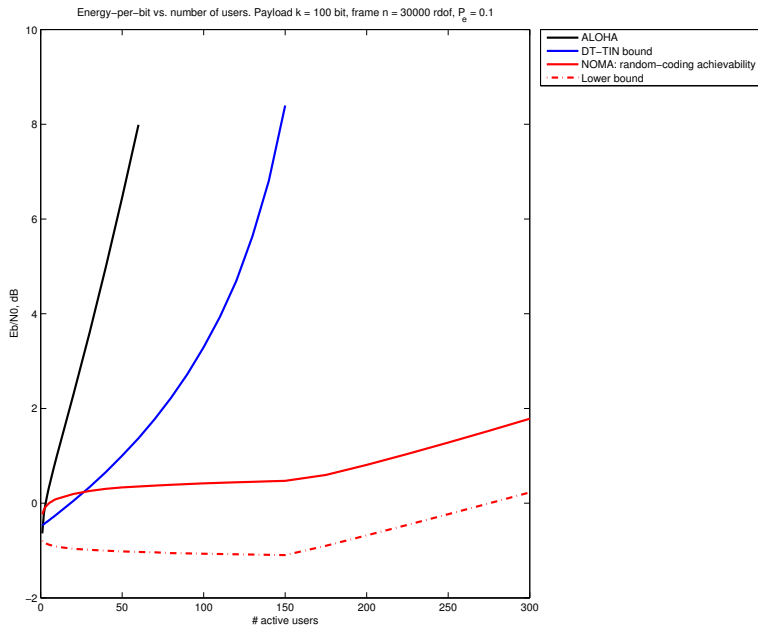


Fundamental limits vs. ALOHA

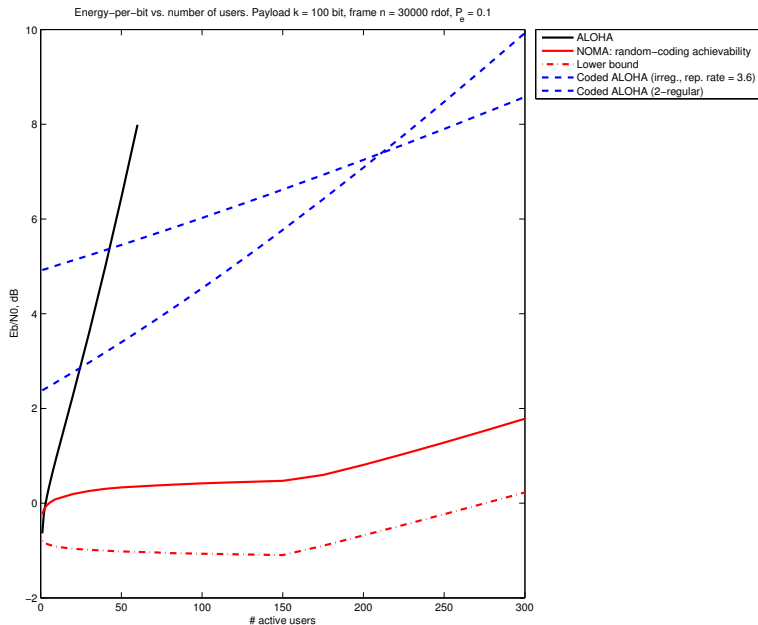
Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$



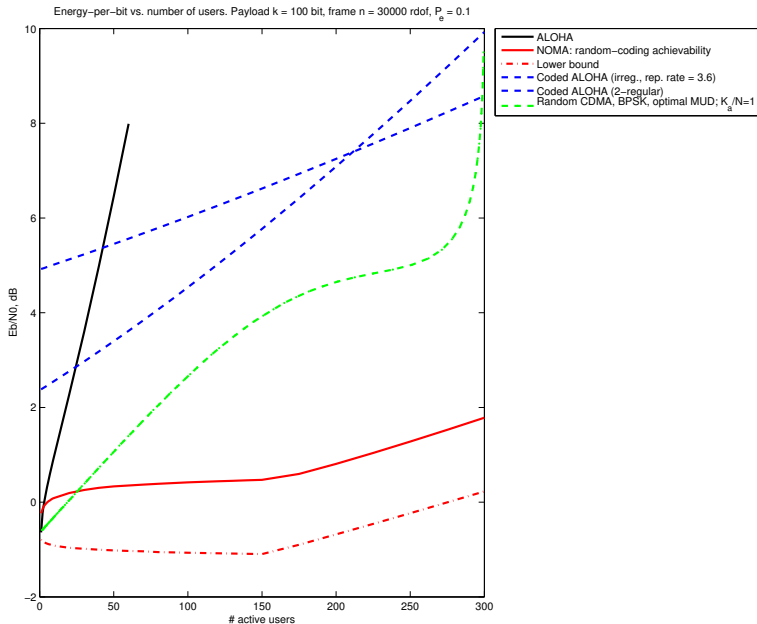
Fundamental limits vs. TIN (aka CDMA w/o MUD)



Fundamental limits vs. Coded Slotted ALOHA



... and randomly-spread CDMA w/ optimal MUD



Problem 1 “massive”: K and n are comparable



Relevant asymptotics: $K, n \rightarrow \infty$ with $\frac{K}{n} = \mu$.

Problem 2 “user-centric” probability of error



$$P_e \triangleq \frac{1}{K} \sum_j \mathbb{P}[\hat{X}_j \neq X_j]$$

Problem 3 “random-access”



indistinguishable users (same-codebook), non-asymptotics.

Low-complexity random-access over GMAC

Key challenge:

Providing multiple-access to massive number of
UNCOORDINATED
and infrequently communicating devices

Key challenge:

Providing multiple-access to massive number of
UNCOORDINATED
and infrequently communicating devices

Typical scenario:

- Huge # of users $K_{\text{tot}} \approx 10^6 - 10^7$
- Still large # of active users $K_a \approx 1 - 500$
- Small data payload, e.g. $k = 100$ bits
- Blocklength $n \sim 10^4$
- $\frac{k}{n} \ll 1$, but system spectral efficiency $\rho = \frac{K_a \cdot k}{n} \sim 1$

Key challenge:

Providing multiple-access to massive number of
UNCOORDINATED
and infrequently communicating devices

Typical scenario:

- Huge # of users $K_{\text{tot}} \approx 10^6 - 10^7$
- Still large # of active users $K_a \approx 1 - 500$
- Small data payload, e.g. $k = 100$ bits
- Blocklength $n \sim 10^4$
- $\frac{k}{n} \ll 1$, but system spectral efficiency $\rho = \frac{K_a \cdot k}{n} \sim 1$

The goal is to communicate with the smallest possible energy-per-bit

Simple scheme I: Treat interference as noise (TIN)

Theorem (DT-TIN bound)

There exists $\mathcal{C} \subset B(0, \sqrt{nP})$ of size M such that

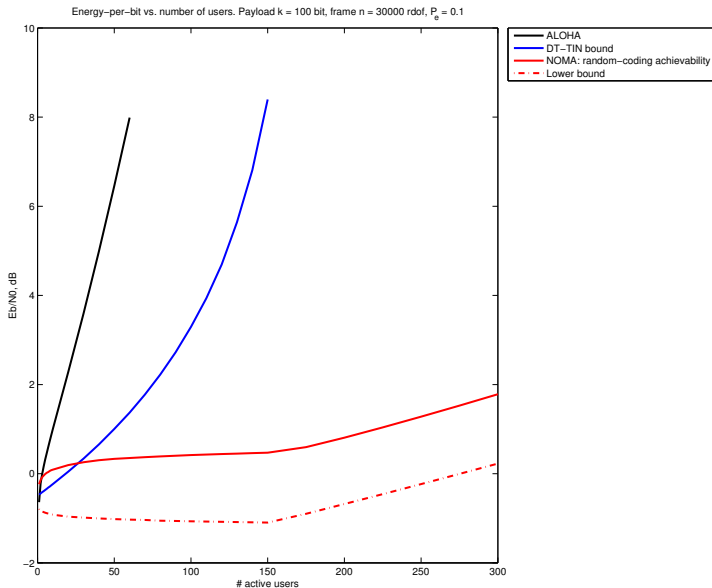
$$\mathbb{P}[X_1 \notin \{\text{top-}K_a \text{ closest } c/w \text{ to } Y\}] \lesssim \mathbb{E} \left[e^{-|i(X; X+Z) - \log M|^+} \right]$$

where $Y = X_1 + \dots + X_{K_a} + Z$, X_i – uniform on \mathcal{C} , $X \sim \mathcal{N}(0, P)^{\otimes n}$ and $Z \sim \mathcal{N}(0, 1)^{\otimes n}$.

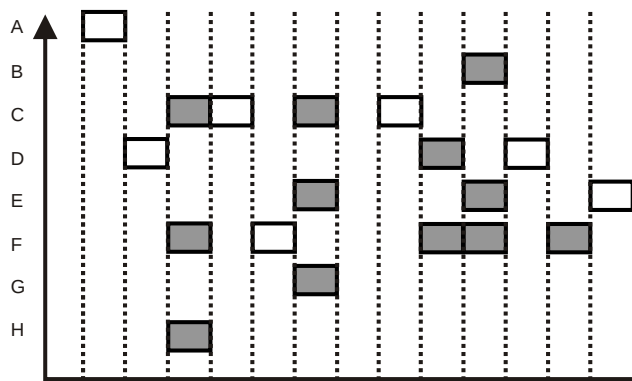
Remarks:

- Decoder searches for top- K_a closest codewords
- Achieves about $\log M \approx nC_{TIN}(P) - \sqrt{nV_{TIN}(P)}Q^{-1}(\epsilon)$
 $C_{TIN}(P) = \frac{1}{2} \log \left(1 + \frac{P}{1+(K_a-1)P} \right)$, $V_{TIN}(P) = \frac{P \log^2 e}{1+K_a P}$.
- Spectral efficiency as $K_a \rightarrow \infty$ is bounded by $\frac{\log_2 e}{2} \approx 0.72$ bit.

Simple scheme I: Treat interference as noise (TIN)



Simple scheme II: T -fold ALOHA

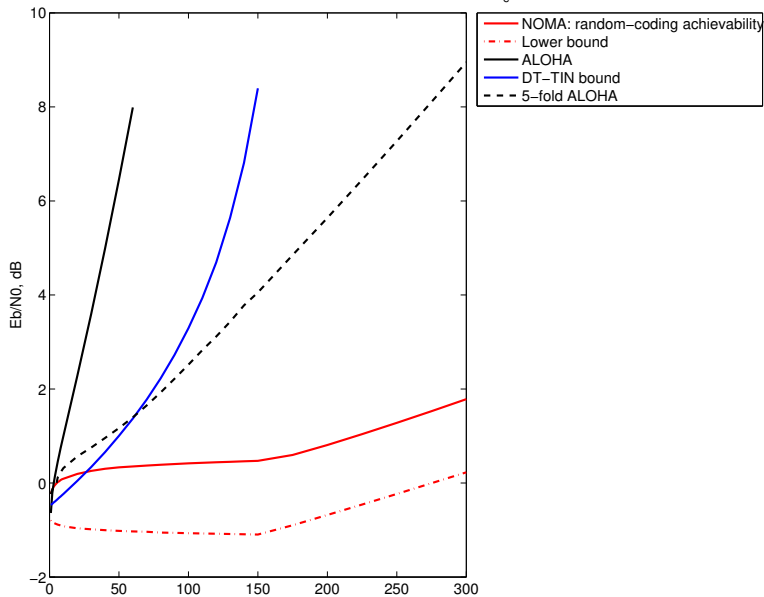


Slotted ALOHA protocol (shaded slots indicate collision)

- Each user places his n_1 -codeword into one of L subframes.
- **Assume any T -fold collision is resolvable**
- Per-user error: $P_e \approx \mathbb{P}[\text{Bino}(K_a - 1, \frac{1}{L}) > T] \approx \left(\frac{K_a}{L}\right)^T e^{-\frac{K_a}{L}}$

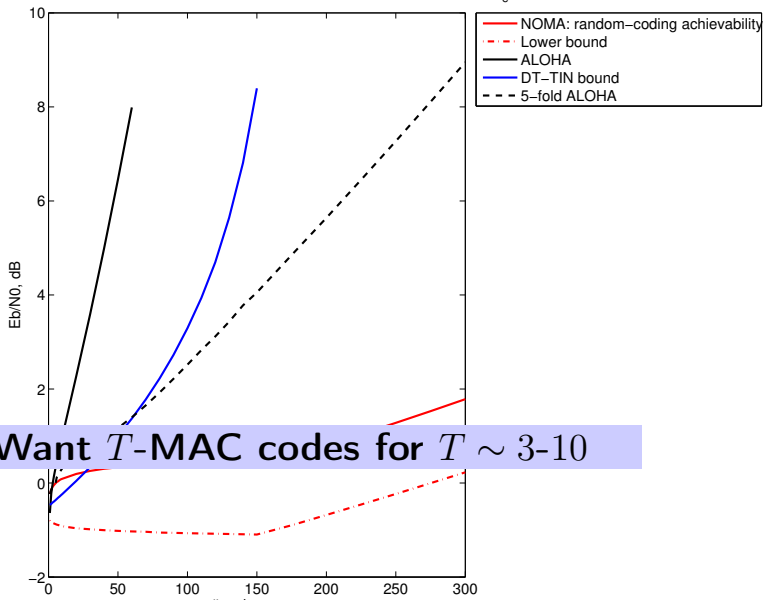
Simple scheme II: T -fold ALOHA

Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$



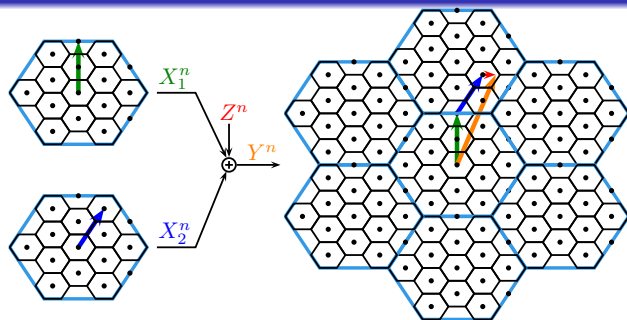
Simple scheme II: T -fold ALOHA

Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$



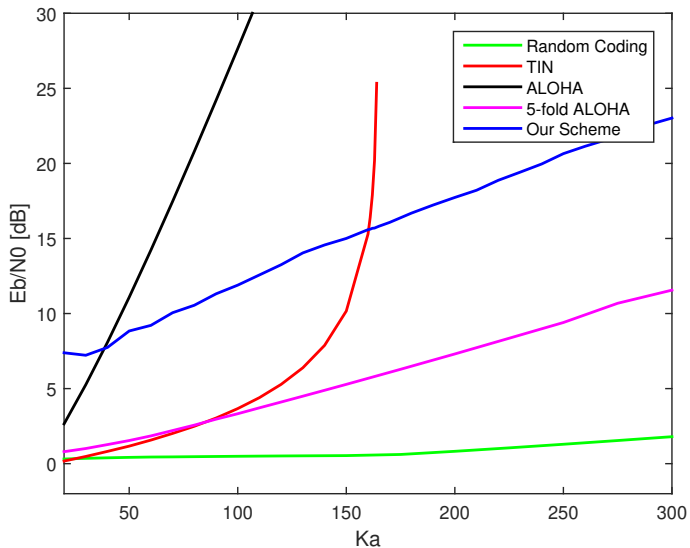
Want T -MAC codes for $T \sim 3-10$

Our scheme: high-level idea



- Send lattice points
- **Decode sum of codewords** via single-user decoder [Nazer-Gastpar'11]
- **Use a subset of points** forming a Sidon set
(all sums $c_1 + c_2$ distinct)
- Single-lattice (no MMSE scaling): $R \approx \frac{1}{2K} \log^+ P$
- Nested-lattice (with MMSE scaling): $R \approx \frac{1}{2K} \log^+ \left(\frac{1}{K} + P \right)$
Warning: issues with same-dither
- **Lose power-factor** compared to $\frac{1}{2K} \log(1 + KP)$

Sample performance of new scheme



Many ideas appeared separately:

- Compute-and-forward [Nazer-Gastpar'11]
- Explicit codes for the modulo-2 binary adder channel [Lindström'69, Bar-David et al.'93]
- 2-user codes for \mathbb{F}_q -adder MAC [Dumer-Zinoviev'78, Dumer'95]
- Concatenation of codes with good minimum distance and codes for the BAC [Ericson-Levenshtein'94]
- Concatenation of CoF inner codes with syndrome decoding for compressed sensing [Lee-Hong'16]

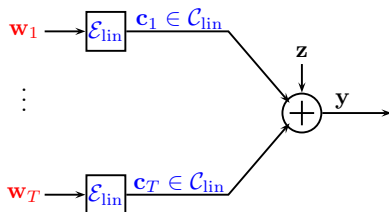
Three phases:

- Sidon set: $\{0, 1\}^k \rightarrow \mathbb{F}_p^n$
- Compute-and-forward: $\mathbb{F}_p^n \rightarrow \mathbb{R}^{n_1}$
- T -fold ALOHA: Place n_1 -codeword in a random subframe

Concatenation scheme

Inner code (CoF):

Convert T -user GMAC into a mod- p (noiseless) adder MAC.



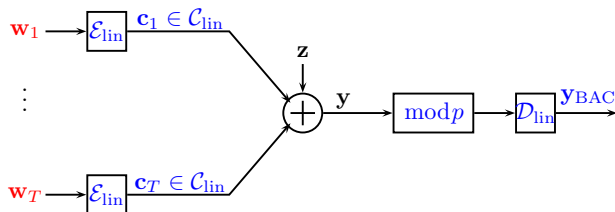
w_1, \dots, w_T are vectors in \mathbb{Z}_p

\mathcal{C}_{lin} is linear code over \mathbb{Z}_p

Concatenation scheme

Inner code (CoF):

Convert T -user GMAC into a mod- p (noiseless) adder MAC.



w_1, \dots, w_T are vectors in \mathbb{Z}_p

\mathcal{C}_{lin} is linear code over \mathbb{Z}_p

$$y_{\text{BAC}} = \left[\sum_{i=1}^T w_i \right] \text{mod } p$$

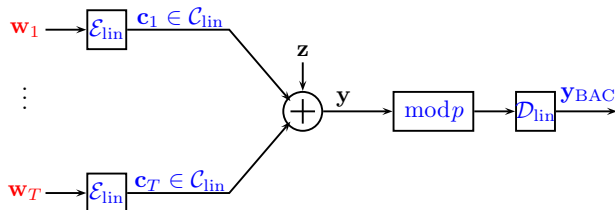
Concatenation scheme

Inner code (CoF):

Convert T -user GMAC into a mod- p (noiseless) adder MAC.

Outer code (BAC):

\mathcal{C}_{BAC} code for mod- p adder T -MAC Here: only $p = 2$



w_1, \dots, w_T are vectors in \mathbb{Z}_p

\mathcal{C}_{lin} is linear code over \mathbb{Z}_p

$$y_{\text{BAC}} = \left[\sum_{i=1}^T w_i \right] \text{mod } p$$

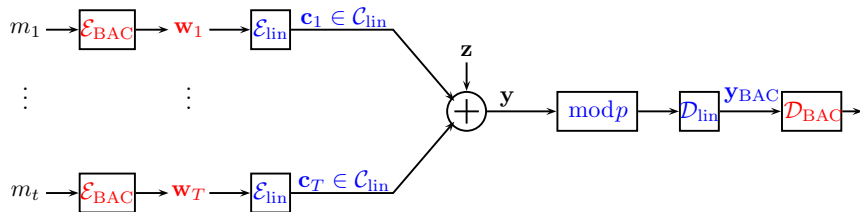
Concatenation scheme

Inner code (CoF):

Convert T -user GMAC into a mod- p (noiseless) adder MAC.

Outer code (BAC):

\mathcal{C}_{BAC} code for mod- p adder T -MAC Here: only $p = 2$



$\mathbf{w}_1, \dots, \mathbf{w}_T$ are vectors in \mathbb{Z}_p

\mathcal{C}_{lin} is linear code over \mathbb{Z}_p

$$y_{\text{BAC}} = \left[\sum_{i=1}^T \mathbf{w}_i \right] \text{mod } p$$

- $\mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$ is a binary linear code (shifted to $\pm\sqrt{P}$)
- Receive $\mathbf{y} = \sum_{i=1}^T \mathbf{x}_i + \mathbf{z}$, shift, rescale, take mod-2, get

$$\mathbf{y}_{\text{CoF}} = [\mathbf{x} + \mathbf{z}] \bmod 2$$

where $\mathbf{x} = [\sum_i \mathbf{x}_i] \bmod 2 \in \mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$

- The channel from \mathbf{x} to \mathbf{y}_{CoF} is a **BMS with folded Gsn noise**
 \implies Designing \mathcal{C}_{lin} is a standard coding task

Normal approximation: $\log |\mathcal{C}_{\text{lin}}| \approx nC - \sqrt{nV}Q^{-1}(\epsilon_{\text{code}})$

- $\mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$ is a binary linear code (shifted to $\pm\sqrt{P}$)
- Receive $\mathbf{y} = \sum_{i=1}^T \mathbf{x}_i + \mathbf{z}$, shift, rescale, take mod-2, get

$$\mathbf{y}_{\text{CoF}} = [\mathbf{x} + \mathbf{z}] \bmod 2$$

where $\mathbf{x} = [\sum_i \mathbf{x}_i] \bmod 2 \in \mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$

- The channel from \mathbf{x} to \mathbf{y}_{CoF} is a **BMS with folded Gsn noise**
 \implies Designing \mathcal{C}_{lin} is a standard coding task

Normal approximation: $\log |\mathcal{C}_{\text{lin}}| \approx nC - \sqrt{nV}Q^{-1}(\epsilon_{\text{code}})$

What is lost in the conversion $\mathbf{y} \mapsto \mathbf{y}_{\text{CoF}}$?

Sum-capacity of \mathbf{y} grows like $\log(T \cdot P)$

Capacity of \mathbf{y}_{CoF} only grows like $\log(P)$

- $\mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$ is a binary linear code (shifted to $\pm\sqrt{P}$)
- Receive $\mathbf{y} = \sum_{i=1}^T \mathbf{x}_i + \mathbf{z}$, shift, rescale, take mod-2, get

$$\mathbf{y}_{\text{CoF}} = [\mathbf{x} + \mathbf{z}] \bmod 2$$

where $\mathbf{x} = [\sum_i \mathbf{x}_i] \bmod 2 \in \mathcal{C}_{\text{lin}} \subset \{0, 1\}^n$

- The channel from \mathbf{x} to \mathbf{y}_{CoF} is a **BMS with folded Gsn noise**
 \implies Designing \mathcal{C}_{lin} is a standard coding task

Normal approximation: $\log |\mathcal{C}_{\text{lin}}| \approx nC - \sqrt{nV}Q^{-1}(\epsilon_{\text{code}})$

What is lost in the conversion $\mathbf{y} \mapsto \mathbf{y}_{\text{CoF}}$?

Sum-capacity of \mathbf{y} grows like $\log(T \cdot P)$

Capacity of \mathbf{y}_{CoF} only grows like $\log(P)$

T -fold ALOHA reduces “power-loss” to $1/T$ instead of $1/K_a$

$$\mathbf{y}_{\text{BAC}} = \left[\sum_{i=1}^T \mathbf{w}_i \right] \bmod 2, \quad \mathbf{w}_1, \dots, \mathbf{w}_T \in \mathcal{C}_{\text{BAC}}$$

Need to decode a list $\{\mathbf{w}_1, \dots, \mathbf{w}_T\}$

Symmetric-capacity: $C_{\text{sym}} = \frac{1}{T}$

$$\mathbf{y}_{\text{BAC}} = \left[\sum_{i=1}^T \mathbf{w}_i \right] \bmod 2, \quad \mathbf{w}_1, \dots, \mathbf{w}_T \in \mathcal{C}_{\text{BAC}}$$

Need to decode a list $\{\mathbf{w}_1, \dots, \mathbf{w}_T\}$

Symmetric-capacity: $C_{\text{sym}} = \frac{1}{T}$

How to construct explicit codes?

- Let $H = [\mathbf{h}_1 | \dots | \mathbf{h}_N]$ be the **parity-check matrix** of a T -error correcting code
- \Rightarrow all T -sums of columns are distinct
- Set $\mathcal{C}_{\text{BAC}} = \{\mathbf{h}_1, \dots, \mathbf{h}_N\}$
- BCH parity check matrix: $R_{\text{BAC}} = \frac{1}{T}$ (optimal!)
- Encoding: easy (just compute $\alpha, \alpha^3, \dots, \alpha^{2T-1}$)

$$\mathbf{y}_{\text{BAC}} = \left[\sum_{i=1}^T \mathbf{w}_i \right] \bmod 2, \quad \mathbf{w}_1, \dots, \mathbf{w}_T \in \mathcal{C}_{\text{BAC}}$$

Need to decode a list $\{\mathbf{w}_1, \dots, \mathbf{w}_T\}$

Symmetric-capacity: $C_{\text{sym}} = \frac{1}{T}$

How to construct explicit codes?

- Let $H = [\mathbf{h}_1 | \dots | \mathbf{h}_N]$ be the **parity-check matrix** of a T -error correcting code
- \Rightarrow all T -sums of columns are distinct
- Set $\mathcal{C}_{\text{BAC}} = \{\mathbf{h}_1, \dots, \mathbf{h}_N\}$
- BCH parity check matrix: $R_{\text{BAC}} = \frac{1}{T}$ (optimal!)
- Encoding: easy (just compute $\alpha, \alpha^3, \dots, \alpha^{2^T-1}$)

Problem: decoding complexity of BCH linear in $n = 2^k - 1$

Decoding:

- $\alpha_1, \dots, \alpha_T \in \mathbb{F}_{2^k}$ are messages
- $y_{\text{BAC}} = He' - \text{syndrome (!)} \Rightarrow$ we know $\sum_i (\alpha_i)^s, s \leq 2T$
- *Error locator*: Berlekamp-Massey yields coeffs of

$$\sigma(z) = \prod_{i=1}^T (1 + \alpha_i z)$$

- *Find roots of $\sigma(\cdot)$* e.g. via [\[Rabin'80\]](#)
- *Invert roots*: using the identity $\alpha^{-1} = \alpha^{2^k} - 1$

Total complexity: $\mathcal{O}(kT^2 \log^2(T) \log \log(T))$ operations in \mathbb{F}_{2^k}

The spectral efficiency $\rho = \frac{K_a \cdot k}{n}$ of our scheme is at most R_{lin}
What if $\rho > 1$?

Solution: - work with $p > 2$

- CoF phase requires good linear codes over \mathbb{F}_p
- BAC phase can be implemented using $H = [\mathbf{h}_1 | \dots | \mathbf{h}_n]$ of a $[n = p^s - 1, n - k = 2T]$ Reed-Solomon code over \mathbb{F}_{p^s} with

$$\mathcal{C}_{\text{BAC}} = \{\alpha \mathbf{h}_i : \alpha \in \mathbb{F}_{p^s} \setminus \{0\}, i = 1, \dots, p^s - 1\}$$

- Can use nested lattice to achieve the 1.53dB shaping gain
- **Drawback:** hard to analyze **finite blocklength**

Asymptotic optimum: $\left(\frac{E_b}{N_0}\right)^* = \frac{2^{2\rho}-1}{2\rho}$, with $\rho = \frac{K_a \cdot k}{n}$.

Let $L = \frac{K_a}{\alpha T}$ for $\alpha \in (0, 1]$ be number of subframes

$P_e \approx \mathbb{P}[T\text{-collision}] = \Pr\left(\text{Binomial}\left(K_a - 1, \frac{\alpha T}{K_a}\right) \geq T\right)$

Linear code rate $R_{\text{lin}} = \frac{\rho}{\alpha}$

$$\begin{aligned}\Delta &= \left(\frac{E_b}{N_0}\right) \text{dB} - \left(\frac{E_b}{N_0}\right)^* \text{dB} \\ &\approx 6\rho \frac{1-\alpha}{\alpha} + 10 \log_{10}(\alpha)\end{aligned}$$

T-Collision avoidance loss due to a $1/\alpha$ increase in spectral efficiency

Asymptotic optimum: $\left(\frac{E_b}{N_0}\right)^* = \frac{2^{2\rho}-1}{2\rho}$, with $\rho = \frac{K_a \cdot k}{n}$.

Let $L = \frac{K_a}{\alpha T}$ for $\alpha \in (0, 1]$ be number of subframes

$P_e \approx \mathbb{P}[T\text{-collision}] = \Pr\left(\text{Binomial}\left(K_a - 1, \frac{\alpha T}{K_a}\right) \geq T\right)$

Linear code rate $R_{\text{lin}} = \frac{\rho}{\alpha}$

$$\begin{aligned}\Delta &= \left(\frac{E_b}{N_0}\right) \text{dB} - \left(\frac{E_b}{N_0}\right)^* \text{dB} \\ &\approx 6\rho \frac{1-\alpha}{\alpha} + 10 \log_{10}(\alpha) + 10 \log_{10}(T)\end{aligned}$$

CoF loss from the reduction $\mathbf{y} \mapsto \mathbf{y}_{\text{CoF}}$

Asymptotic optimum: $\left(\frac{E_b}{N_0}\right)^* = \frac{2^{2\rho}-1}{2\rho}$, with $\rho = \frac{K_a \cdot k}{n}$.

Let $L = \frac{K_a}{\alpha T}$ for $\alpha \in (0, 1]$ be number of subframes

$$P_e \approx \mathbb{P}[T\text{-collision}] = \Pr\left(\text{Binomial}\left(K_a - 1, \frac{\alpha T}{K_a}\right) \geq T\right)$$

Linear code rate $R_{\text{lin}} = \frac{\rho}{\alpha}$

$$\begin{aligned}\Delta &= \left(\frac{E_b}{N_0}\right) \text{dB} - \left(\frac{E_b}{N_0}\right)^* \text{dB} \\ &\approx 6\rho \frac{1-\alpha}{\alpha} + 10 \log_{10}(\alpha) + 10 \log_{10}(T) - 10 \log_{10}(1 - 2^{-2\rho})\end{aligned}$$

Loss of +1 in computation rate

Asymptotic optimum: $\left(\frac{E_b}{N_0}\right)^* = \frac{2^{2\rho}-1}{2\rho}$, with $\rho = \frac{K_a \cdot k}{n}$.

Let $L = \frac{K_a}{\alpha T}$ for $\alpha \in (0, 1]$ be number of subframes

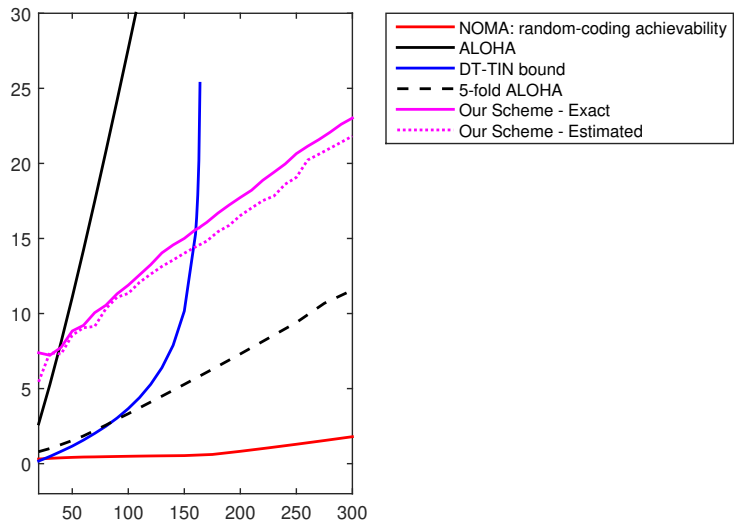
$P_e \approx \mathbb{P}[T\text{-collision}] = \Pr\left(\text{Binomial}\left(K_a - 1, \frac{\alpha T}{K_a}\right) \geq T\right)$

Linear code rate $R_{\text{lin}} = \frac{\rho}{\alpha}$

$$\begin{aligned}\Delta &= \left(\frac{E_b}{N_0}\right) \text{dB} - \left(\frac{E_b}{N_0}\right)^* \text{dB} \\ &\approx 6\rho \frac{1-\alpha}{\alpha} + 10 \log_{10}(\alpha) + 10 \log_{10}(T) - 10 \log_{10}(1 - 2^{-2\rho}) + 1.53\end{aligned}$$

Shaping loss

Low-complexity schemes: summary



- Work in progress by several groups
 - ▶ Narayanan-Chamberland
 - ▶ P.-Frolov
 - ▶ Durisi-Dalai
 - ▶ Popovski-Liva
 - ▶ ... (sorry to those I forgot)
- Methods we did not cover:
 - ▶ Coded Slotted ALOHA
 - ▶ ... including with MPR capability
 - ▶ iterative decoding same-codebook LDPCs
 - ▶ super-imposed codes
- Problem is even more interesting with fading
 - ▶ Random channel gains H_j help distinguish users.
 - ▶ With many users, order statistics of H_j 's becomes deterministic.

Envisioned solution:

- To save battery: sensors sleep all the time, except transmissions.
- ... uncoordinated transmissions.
- ... they wake up, blast the packet, go back to sleep.
- Focus on low-energy (low E_b/N_0)
- Focus on fundamental limits
- ... but with low-complexity solutions (single-user-only decoding).

Envisioned solution:

- To save battery: sensors sleep all the time, except transmissions.
- ... uncoordinated transmissions.
- ... they wake up, blast the packet, go back to sleep.
- Focus on low-energy (low E_b/N_0)
- Focus on fundamental limits
- ... but with low-complexity solutions (single-user-only decoding).

Issues we need to understand:

- ① packets are short: finite-blocklength (FBL) info theory
- ② multiple-access channel: Classical MAC
- ③ low-complexity MAC: modulation, CDMA, multi-user detection
- ④ massive random-access: many users, same-codebook codes (NEW)

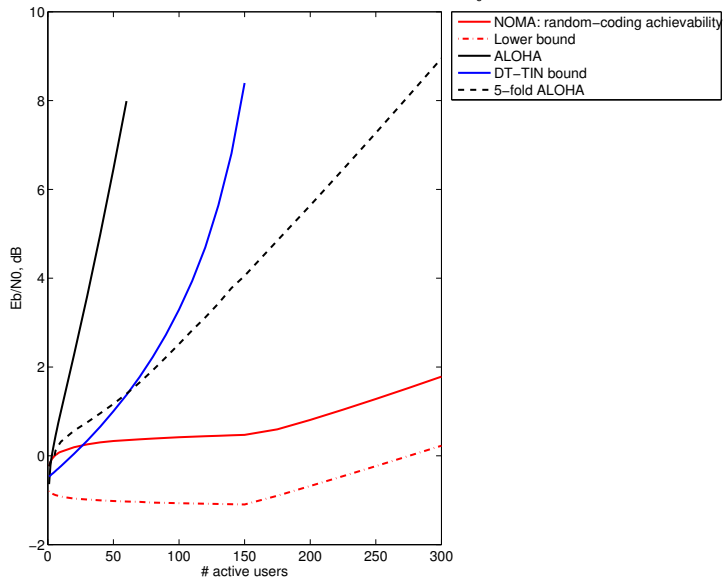
Supporting 10 users at 1Mbps is much easier than 1M users at 10bps.

Thank you!

Extra: More plots

ALOHA + codes repairing 5-fold collisions

Energy-per-bit vs. number of users. Payload $k = 100$ bit, frame $n = 30000$ rdof, $P_e = 0.1$



Other schemes...

