



STANFORD

*Lecture 6*

# **Multuser Channels and the Capacity Region**

*April 15, 2026*

**JOHN M. CIOFFI**

Hitachi Professor Emeritus of Engineering

Instructor EE379B – Spring 2026

# Announcements & Agenda

## ■ Announcements

- Mid term

## ■ Problem Set 4 = PS4 (due May 2)

1. 2.21 Multiuser Channel Types
2. 2.22 Multiuser Detector Margin
3. 2.23 Mutual-Information Vector
4. 2.24 Time-Division Multiplexing region
5. 2.25 MAC regions

## ■ Agenda

- Multi-User (MU) Introduction
  - Where used?; What is a multi-user data rate?; order & decodability
- The 3 basic MU types and the matrix AWGN
- Rate Bounds and Detection
- General MU Capacity Region and other optima

See S6 on basic  
Queuing and Scheduling



# Multiuser (MU) Introduction

*(definitions and fundamentals)*

*Section 2.6 intro*

# U>1 users

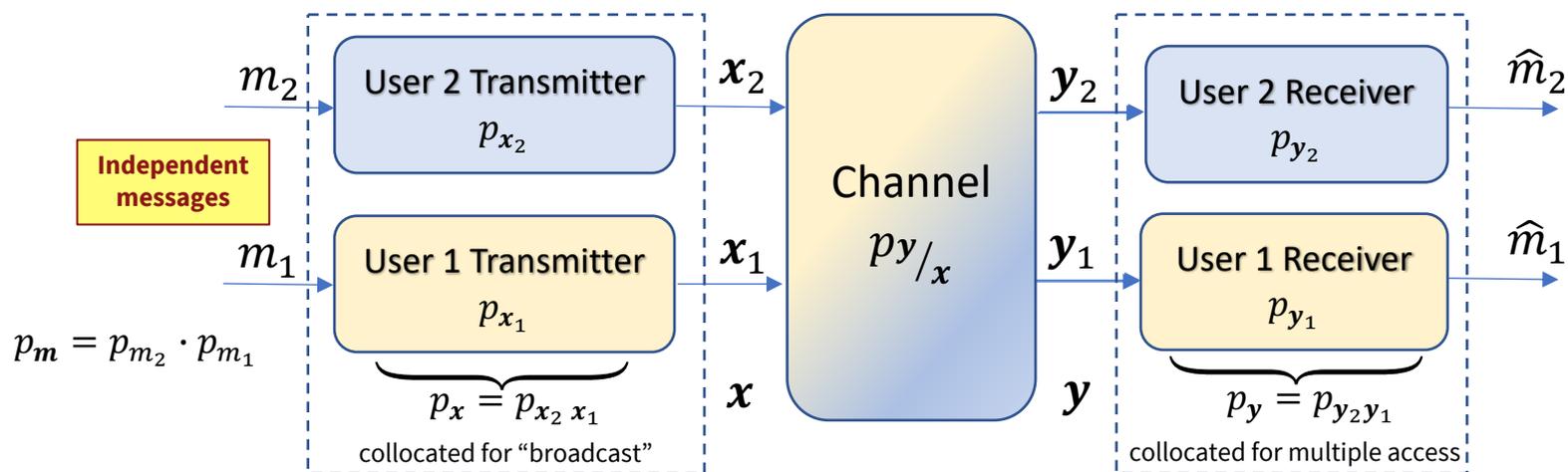


- Downlink/stream – one to many (“broadcast”)
- Uplink/stream – many to one (“multiple access”)
- Relay signals (“mesh”)
- Overlapping combinations (Wi-Fi, or cell, or really all) – “interference”



# MU Mathematical Model (Section 2.6)

- There is a joint probability distribution  $p_{xy}$  that determines all marginals (e.g., input) and conditionals (channel),  $p_{y/x} = \frac{p_{xy}}{p_x}$ .

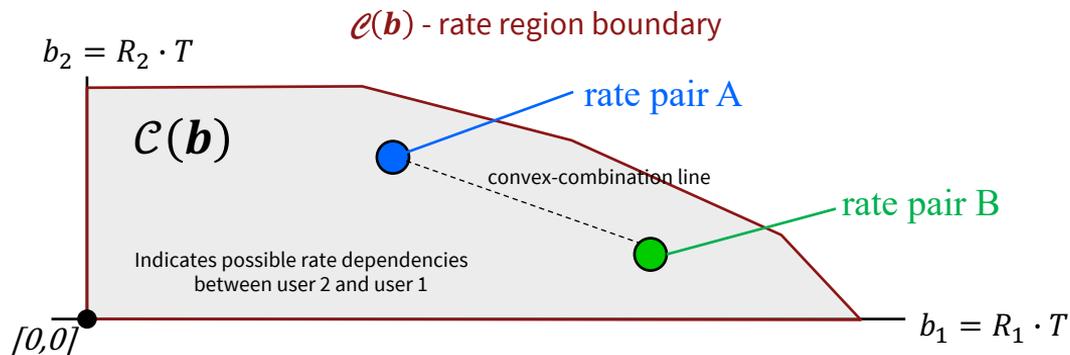


- User 1 & 2's data rates are mutually dependent (otherwise just two single-user channels).
- $b \rightarrow \mathbf{b} = \begin{bmatrix} b_2 \\ b_1 \end{bmatrix} = \mathbf{R} \cdot T = \begin{bmatrix} R_2 \cdot T \\ R_1 \cdot T \end{bmatrix}$ ; the bits/sub-symbol becomes a  $U$ -dimensional vector,  $u = 1, \dots, U$ .
- Single-user is a (degenerate) subset of multiuser.



# The Rate Region

- “Reliably decodable” set of users’ bits/subsymbol vectors that can be achieved  $P_e \rightarrow 0$  (AEP).



- All “convex combinations” (on the line connecting points) must trivially be achievable too.
- What is  $\mathcal{C}(\mathbf{b})$  if independent single-user channels?  
“crosstalk free”  
rectangle (2), prism (3), Orthotope ( $U$ )
- The region is “convex hull” (union) of achievable points over all “allowed”  $p_{xy}$ , or really over  $p_x$ ,
  - because  $p_{y/x}$  (the general MU channel description) is given (fixed).



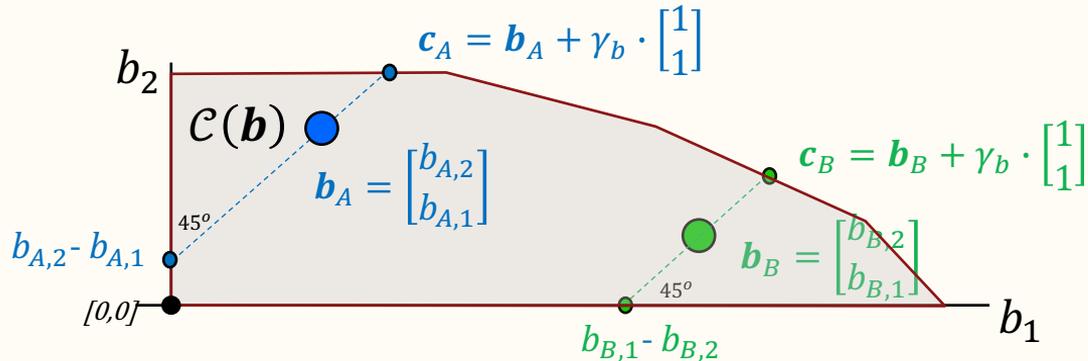
# Multuser Margin

- Single-user (energy) margin  $\bar{b} = \frac{1}{2} \cdot \log_2 \left( 1 + \frac{SNR}{\Gamma \cdot \gamma_m} \right)$  measures safety for  $\bar{b}$  if SNR changes.

$$\gamma_m = \frac{1}{\Gamma} \cdot \frac{SNR}{2^{2\bar{b}} - 1}$$

- The **bit gap** is  $\gamma_b = \bar{C} - b$  where  $\bar{C} = \frac{1}{2} \cdot \log_2(1 + SNR) = \bar{b} + \bar{\gamma}_b$  - so measures rate gap to  $\bar{C}$ .
  - $\Gamma \cdot \gamma_m \cong 6 \cdot \bar{\gamma}_b$  dB,  $\bar{\gamma}_b = 0$  if the code achieves capacity (6 dB/bit-dimension).

- Multuser bit gap** measures to  $\mathbf{c}_{b'}$ ,  $\in \mathcal{C}(\mathbf{b})$ , the rate region boundary, so  $\gamma_{b'} \cdot \mathbf{1} = \mathbf{c}_{b'} - \mathbf{b}'$



$$b_2 = b_1 + (b_{A,2} - b_{A,1})$$

$$b_1 = b_2 + (b_{B,1} - b_{B,2})$$

$$u_{max} \triangleq \arg \left( \max_u b'_u \right)$$

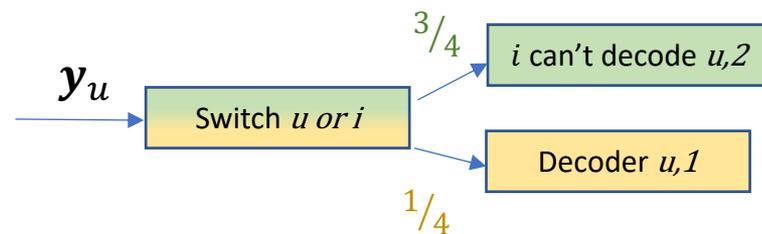
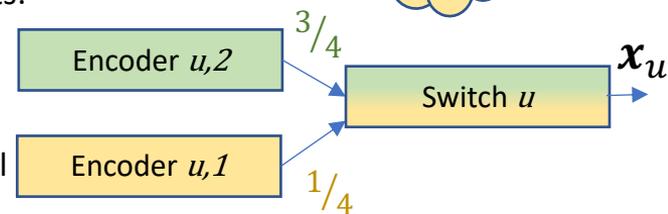
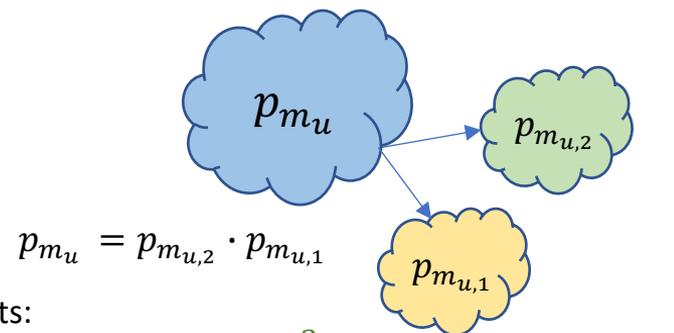
$$b_{u_{max}} = \left( \sum_{i \neq u_{max}} b_i \right) + b'_{u_{max}} - \left( \sum_{i \neq u_{max}} b'_i \right)$$

- Multuser (energy) margin still is then same as single-user margin.  $\Gamma \cdot \gamma_m \cong 6 \cdot \bar{\gamma}_b$  dB



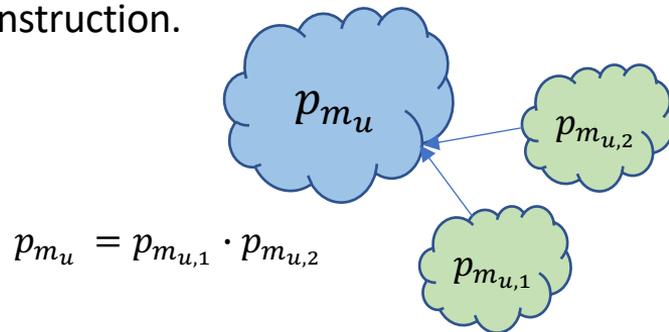
# User Components (a.k.a. “time/dimension-sharing”)

- Two independent **user components** or **subusers** have the
  - same transmitter and same receiver (“different components of same user”).
- These two subusers (codes used) can be separately encoded and decoded.
- Bits per symbol is  $b_u = b_{u,1} + b_{u,2}$ .
- Other users’ receivers  $i \neq u$  may decode all, none, or some of these components:
  - which they should do and remove if possible, or
  - otherwise they are averaged in marginal (remains as noise when Gaussian).
- The two subusers may simultaneously share dimensions, apportioning fractional information (or energy when Gaussian) to each.
- $U$  can increase to  $U + 1$ , or more generally to  $U \leq U' \leq U^2$  components.
- $\mathcal{C}(\mathbf{b})$ , and  $b$ , can also expand to  $U'$  dimensions:
  - Original  $\mathcal{C}(\mathbf{b})$  adds together the sub-users’ dimensional rates, and thus decreases its dimensionality.
- Some information theorists call this “**time-sharing**,”
  - but user components is more accurate and general, and extrapolates to all types of dimensions and combinations.



# Macro Users

- Two users (or user components) that have identical impact/influence create **a macro user**.
  - $p_{\dots x_u \dots x_i y} = p_{\dots x_i \dots x_u y}$  - interchange of the users does not change the joint probability distribution.
  - These two could be considered one macro user, where any partition of this macro user's rate to the two original users is feasible.
- Simple example is  $y = x_1 + x_2 + n$ , where both users 1 and 2 share the same energy.
  - This is  $\sim$  single-user channel with macro user  $x = x_1 + x_2$ , for which any division of  $b = b_1 + b_2$  is possible.
- Macro/Micro users can simplify some capacity-region construction.



# The real variation – the users' rates



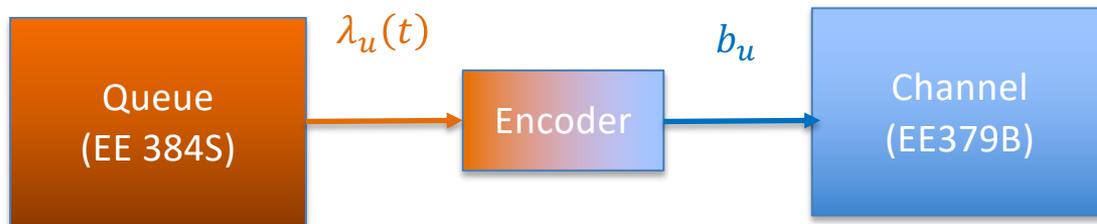
Network Designer

says the “channel has a continuous rate that apportions dynamically as needed to any user,  
 $\sum_u \lambda_u(t) = \text{constant}.$ ”



Modem Designer

says the “source has a continuous rate that is always on, each  $b_u$  is constant.”

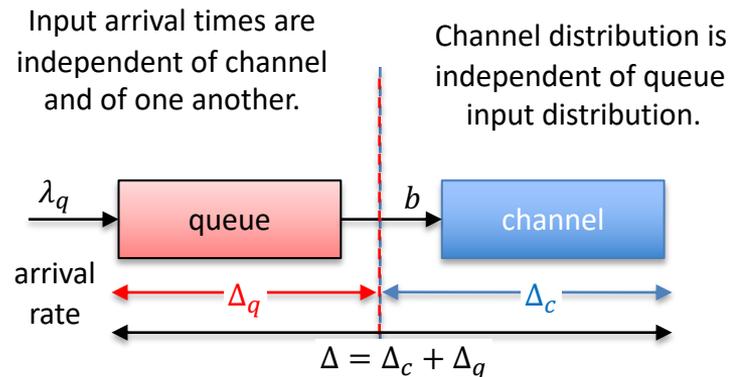


- Neither of these two design perspectives is (always) correct.
  - See also Appendix A's queuing theory basics.



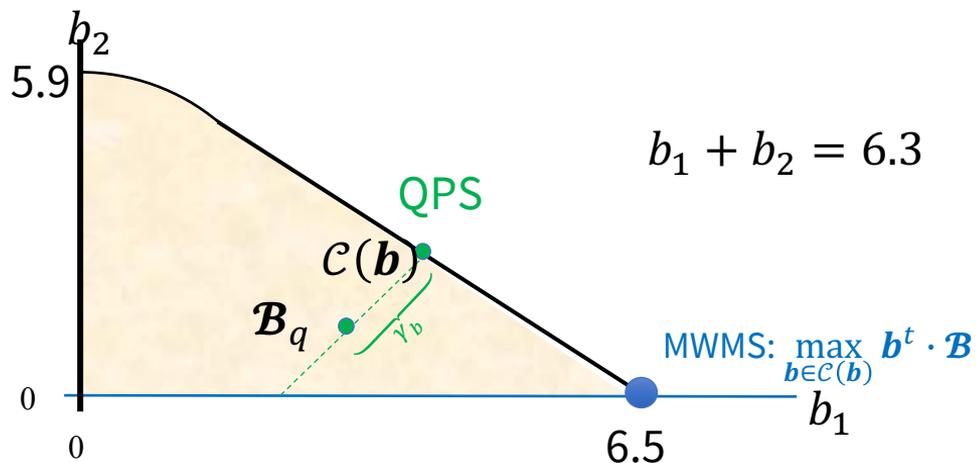
# Queuing Basics

- Arrivals are independent of channel variation.
- $\mathcal{B} = \lambda_q \cdot \Delta \rightarrow \mathbb{E}[\mathcal{B}] = \mathbb{E}[\lambda_q] \cdot \mathbb{E}[\Delta] =$   
number of bits in system (Little's Theorem).
- $\mathbb{E}[\lambda_q] \leq \mathbb{E}[b]$  for stable operation.
- Multiuser Form
  - $\mathbb{E}[\mathcal{B}] = \mathbb{E}[\lambda_q] \odot \mathbb{E}[\Delta]$ .



# Solution: Queue Proportional Scheduling

- Send data rate in capacity region that has user rate vector as scaled version of user queue depths.



We'll learn later how to find if a point is admissible (the green QPS point on the boundary).

- The design point is proportional to users relative queue depths, and has margin  $\gamma_b$ .
- QPS (Queue Proportional Scheduling) has lowest average delay of all scheduling methods.
- Less jitter than MWMS, fair among users (QPS empties the queues faster).



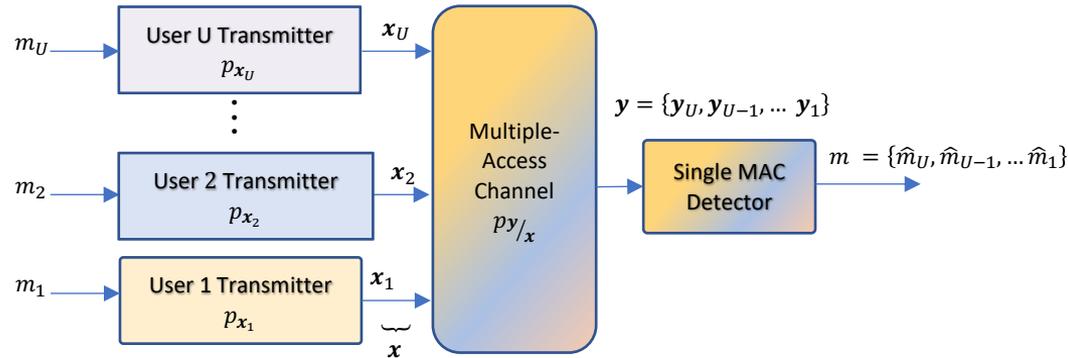
# The 3 Basic MUs & Matrix AWGN

PS4.1 - 2.21 Multiuser Channel Types

*Section 2.6.1*

# Multiple Access Channel (MAC)

Sec 2.6.1 and 2.7



- User transmitters are in different locations (cannot coordinate to encode/modulate  $x$ ).
  - All use a good single-user code (see 379A, Chapter 8).
- Single receiver detects all users and:
  - separates the users,
  - reliably decodes,  $P_e \rightarrow 0$ , by decoding and removing some (none or all) other users first, which
  - suggests “user **order**”  $\pi$  (vector “priority”) is important (decode  $\pi$ ’s 1<sup>st</sup>/bottom element first, ...  $U$  ... last at top).
  - If subusers, then up to  $U'$  subusers might be decoded, where again  $U \leq U' \leq U^2$ .
- **Order is fundamental** to best MU design – the MAC has  $U!$  possible orders (each has a  $\mathbf{b}$ )
  - and all potential convex combinations thereof.
  - There is also an input  $p_x$  choice (or code choice), and all potential convex combinations thereof.
  - The MAC simplifies so that  $U' = U$ .

**Design**

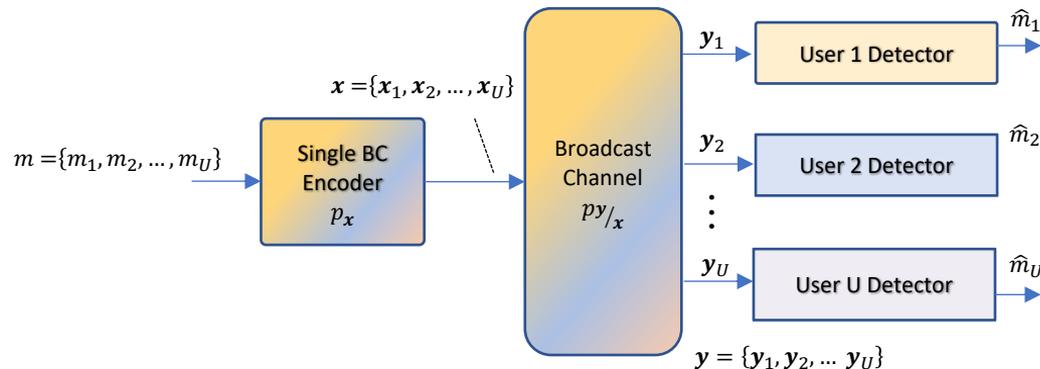
1.  $\pi$

2.  $p_x$



# Broadcast Channel (BC)

Sec 2.6.1 and 2.8



- The BC is the “**dual**” of a special type of MAC.
  - This eventually allows a common design method, at least for Matrix AWGN case.
- Receivers are in different places, so they cannot “co-process”  $y$ 's user outputs.
- Transmitter can co-encode/generate  $x$ , although input messages remain independent.
  - Who encodes first? (may be at disadvantage)
  - Who encodes last? (knowing other users' signals is an advantage)
  - What then is the **order**?
  - Can reduce to just  $U!$  Orders, although  $U' > U$ .

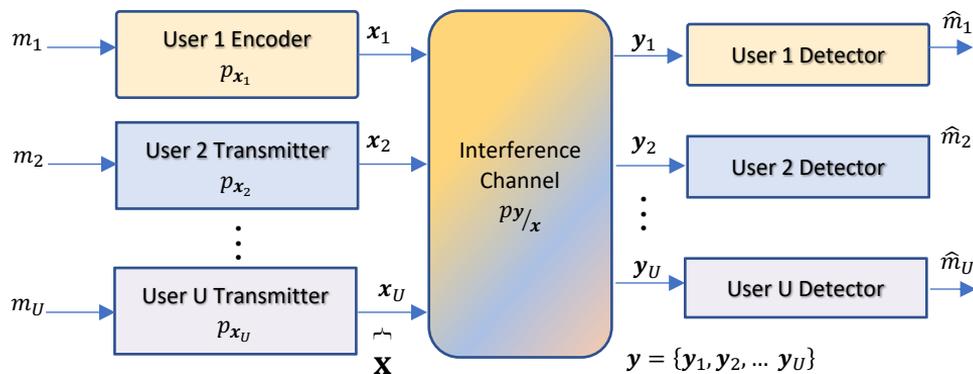
**Design**

1.  $\pi$

2.  $p_x$



# Interference Channel (IC)



Sec 2.6.1 and 2.9

- Users' transmitters, and also receivers, are in different locations.
  - No co-encoding of user messages nor coordinated reception is possible.
  - Views: A set of MACs with same inputs or a set of BCs with same outputs.
- Each receiver can use a decoding **order** to detect others first, if decoding is possible.
  - The rcvr treats other users as noise if not possible to decode/remove first.
  - Each receiver's order is column of **matrix order  $\Pi$** .
- There are  $(U'!)^U$  possible IC **orders**: ...  $U'!$  at each receiver, with  $U' \leq U^2$ .
  - Each user may have a subuser component for every user's receiver to detect, large **FINITE**.

**Design**

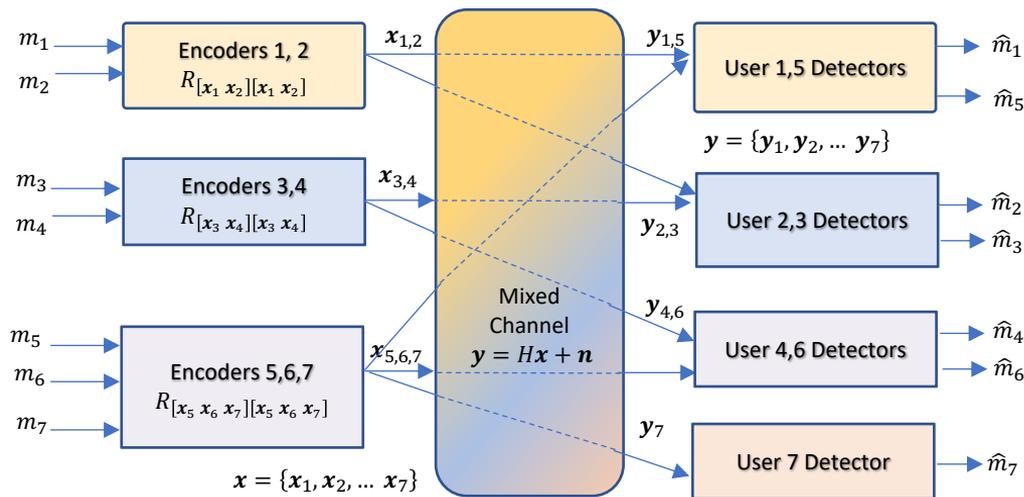
1.  $\Pi$

2.  $p_x$



# Other MU Types / Combinations

- Mixed



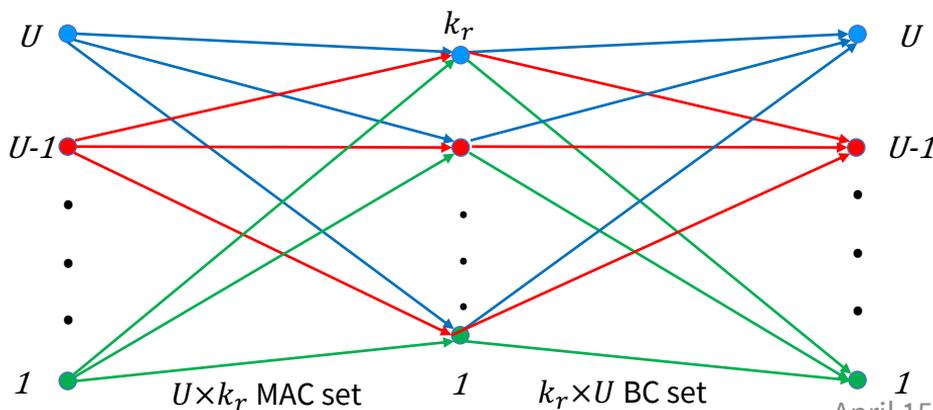
Different Macro views:

IC of 4 MAC macros:  
 $\{1,5\}, \{2,3\}, \{4,5\}$   
 $\{7\}$  is single user.

IC of 3 BC macros:  
 $\{1,2\}, \{3,4\}, \{5,6,7\}$ .

Macro design can shrink  $\mathcal{C}(\mathbf{b})$  dimensionality.

- Mesh/Relay

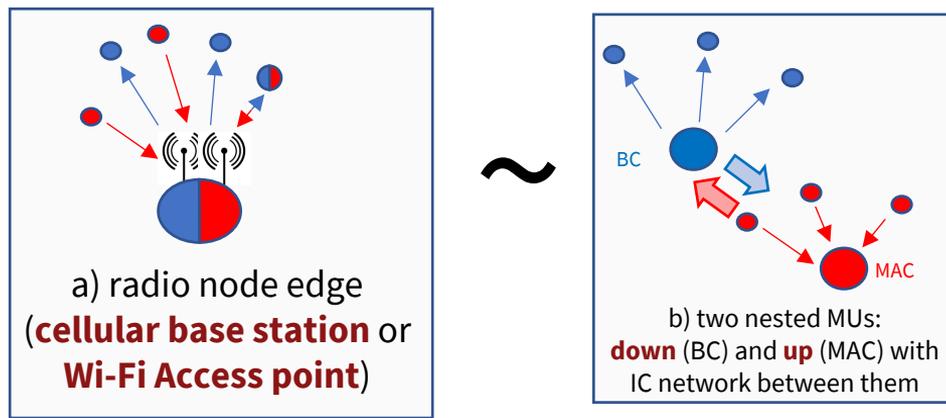


$k_r$  relays

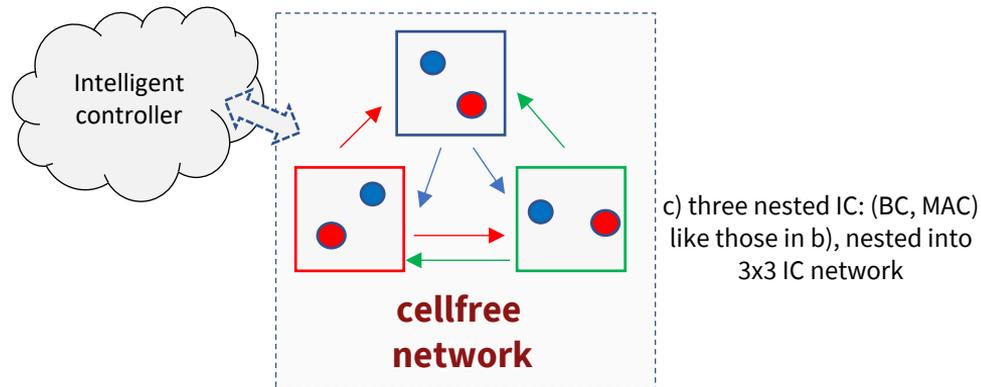
- Single User

# Nested MU Channel Examples

- Each **nested MU** channel  $\rightarrow$  1 macro user.
- These macro users crosstalk into each other.
  - Some users with macro group may decode
  - with any given order in that group.
  - Those not decoded are undecodable “noise.”



- Design treats as IC of macro users:
  - Orders users within the macro’s subusers.
  - Rcvrs decode all subusers within the local macro group
  - or use “multi-level” waterfilling (end of 379B).



# Rate Bounds & Detection

PS4.3 - 2.23 Mutual-Information Vector

# Chain-Rule Reminder/Review

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) = \sum_{n=1}^N \mathcal{I}(\mathbf{x}_n; \mathbf{y} / [\mathbf{x}_{n-1} \cdots \mathbf{x}_1])$$

Lemma 2.3.4

- Think of the input components  $\mathbf{x}_n$  as users, so  $U \rightarrow N$  and  $u \rightarrow n$  (may replace  $U$  with  $U'$  in general).
- Any receiver output (or combination of them),  $\mathbf{y}$ , has chain-rule decomposition(s); for the given  $p_{\mathbf{x}\mathbf{y}}$ , this  $\mathcal{I}(\mathbf{x}; \mathbf{y})$  represents a maximum (sum-user) data rate by AEP.
- Each sum term has similar interpretation, given the “previously decoded” (given) other users.
- The capacity region points correspond to chain-rule  $\mathcal{I}$  terms in  $\mathcal{C}(\mathbf{b})$  for each user receiver in that point’s construction.
- User **decoding order** characterizes the different “chain-rule” compositions.



# Some data rate bounds

- **Sum-Rate bound:**  $b = \sum_{u=1}^U b_u \leq \mathcal{I}(\mathbf{x}; \mathbf{y})$  - full transmit/receiver coordination is vector coding.
- **Average User  $u$  bound:**  $\mathcal{I}_u(x_u; \mathbf{y}) \leq \mathcal{I}(\mathbf{x}; \mathbf{y})$  - this does NOT bound  $b_u$  (could remove other user(s) first)
  - $\mathcal{I}_u(x_u; \mathbf{y})$  treats **all** other users as "noise."
- The **conditional mutual information** ~first decodes the conditioning users' messages correctly (reliably) and then removes them from the detection process.
  - For AWGN, this operation corresponds to remodulating, filtering by the known channel, and subtracting the result from receiver  $u$ 's signal.
  - There are ways to simplify this prior-user removal process.
- All chain-rule bounds apply also if  $U \rightarrow U' = U^2$ .



# Fundamental: User Priority $\rightarrow$ “order”

- Receiver  $u$  decodes who first?, last?

$$\boldsymbol{\pi}_u = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \text{ or } \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \text{ or } \dots \text{ (} U'! \text{ Choices)}$$

- Why is this important?

- The as-yet un-decoded users are “noise” (averaged to compute marginal dist’n  $p_{y_u/x_i}$ , upon which ML detector is based).
- Are the other users reliably decodable? (They must be treated as “noise” if not.)

**If others decoded first, then it is successive decoding or “generalized decision feedback,” for some order  $\boldsymbol{\pi}_u$ .**

- Same for all users – so there is a “global” order possibility:  $\boldsymbol{\Pi} = [\boldsymbol{\pi}_U \ \cdots \ \boldsymbol{\pi}_1]$  with  $(U')^U$  choices

- The designer might check all  $\boldsymbol{\Pi}$ , and then take convex combinations for each and every allowed  $p_x$ .
  - It simplifies in many situations (including MAC, BC, and IC).

- Order vector and inverse:**

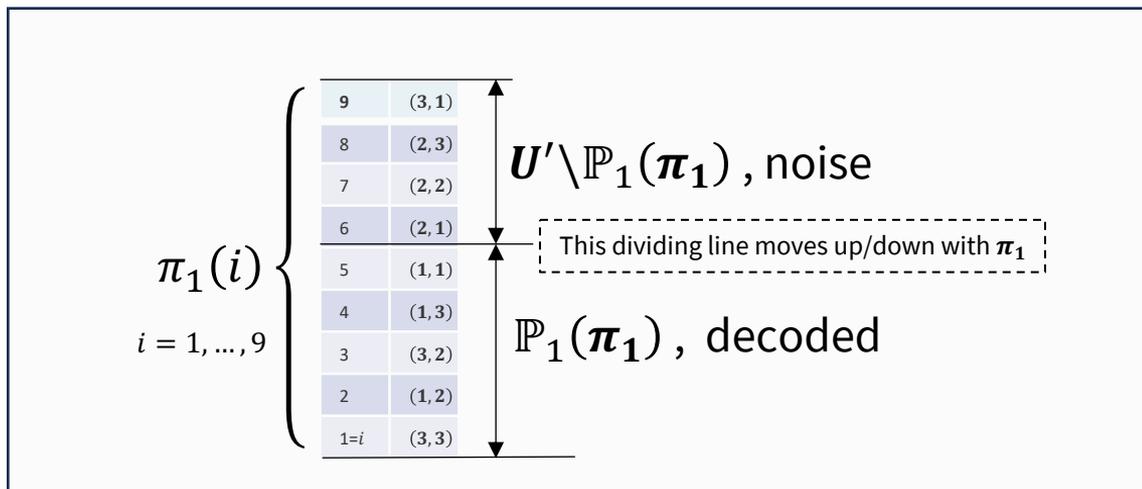
- Any permutation vector has inverse.
- Same as 379A interleave, different use.

$$\boldsymbol{\pi}_u = \begin{bmatrix} \pi(U') \\ \vdots \\ \pi(1) \end{bmatrix} \quad \boldsymbol{\pi}_u^{-1} = \begin{bmatrix} U' \\ \vdots \\ 1 \end{bmatrix} \quad j = \pi(i) \rightarrow i = \pi^{-1}(j)$$



# Prior-User Set

- $U^2$  subusers subdivide into ordered pairs  $(u, u')$  where  $u = 1, \dots, U$  and  $u' = 1, \dots, U$ .
  - Receiver  $u$  observes components from users  $u' = 1, \dots, U$ .
- In order  $\pi(u, u')$ , when receiver  $u$  has decoded all  $u' = 1, \dots, U$  (all its subuser components), it is done.
  - Any higher-in-order user components are noise (marginals).
  - Can reindex order so that  $\pi(u, u') = \pi(i), i = 1, \dots, U^2$  where  $i$  counts from bottom up.
  - The prior-user set for this order  $\mathbb{P}_u(\boldsymbol{\pi})$  occurs for smallest  $i$  that contains  $\{(u, 1), (u, 2), \dots, (u, U)\} \subseteq \mathbb{P}_u(\boldsymbol{\pi})$ .
- Example for  $U = 3$ :



# More Formal Prior-User Set

- Define the receiver-indexed sets

- $S_1 \triangleq \{(1,1), \dots, (1,U)\}$
- $S_2 \triangleq \{(2,1), \dots, (2,U)\}$
- $\vdots$
- $S_U \triangleq \{(U,1), \dots, (U,U)\}$
- $S = S_1 \cup S_2 \cup \dots \cup S_U$ .

$$i_u^*(\boldsymbol{\pi}_u) \triangleq \arg \min_j [\pi_u(S_u) \subseteq \{1:j\}]$$

- Prior-User Set is  $P_u(\boldsymbol{\pi}_{i_u^*}) \triangleq \{(u,i) | \pi(u,i) \leq i_u^*(\boldsymbol{\pi}_u)\}$ .
- For given order and input  $p_x$ , find each user's worst data rate.
- Example – let's say this particular order  $\boldsymbol{\Pi}$  just happens to have the 4 subusers/user aligned successively at each receiver,
  - which simplifies illustration from 16 to 4.

rcvr/ User $i$	$\pi_4(i)$	$\pi_3(i)$	$\pi_2(i)$	$\pi_1(i)$
$i = 4$	3	3	4	3
$i = 3$	4	2	3	2
$i = 2$	1	4	2	1
$i = 1$	2	1	1	4
$\mathbb{P}_u(\boldsymbol{\pi}_u)$	{1,2}	{2,4,1}	{1}	{4}

$$\boldsymbol{\Pi} = \begin{bmatrix} 3 & 3 & 4 & 3 \\ 4 & 2 & 3 & 2 \\ 1 & 4 & 2 & 1 \\ 2 & 1 & 1 & 4 \end{bmatrix}$$

- Designer lists data-rate entries (mutual information bounds really) for those users who are decoded, and  $\infty$  for those treated as noise.

$\mathfrak{S}$	$\mathfrak{S}_4$	$\mathfrak{S}_3$	$\mathfrak{S}_2$	$\mathfrak{S}_1$
top	$\infty$	$\mathfrak{I}_3(3/1,2,4)$ 20	$\infty$	$\infty$
	$\mathfrak{I}_4(4/1,2)$ 10	$\mathfrak{I}_3(2/1,4)$ 9	$\infty$	$\infty$
	$\mathfrak{I}_4(1/2)$ 5	$\mathfrak{I}_3(4/1)$ 4	$\mathfrak{I}_2(2/1)$ 4	$\mathfrak{I}_1(1/4)$ 2
bottom	$\mathfrak{I}_4(2)$ 1	$\mathfrak{I}_3(1)$ 2	$\mathfrak{I}_2(1)$ 2	$\mathfrak{I}_1(4)$ 5

$$\mathfrak{I}_{\min}(\boldsymbol{\Pi}, p_{xy}) = \begin{bmatrix} 4 \\ 20 \\ 1 \\ 2 \end{bmatrix}$$



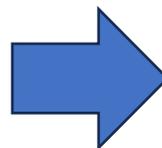
# The minimum Mutual – Info Vector $\mathcal{I}_{min}$

- **Mutual-information-like** quantity
  - follows the prior-user set.

$$\mathcal{I}_u(\mathbf{x}_{\pi_u(i)}; \mathbf{y}_u / \mathbb{P}_{\pi_u(i)}(\boldsymbol{\pi}_u)) \triangleq \begin{cases} \infty & i > \pi_u^{-1}(u) \\ \mathcal{I}_u(\mathbf{x}_{\pi_u(i)}; \mathbf{y}_u / \mathbb{P}_{\pi_u(i)}(\boldsymbol{\pi}_u)) & i \leq \pi_u^{-1}(u) \end{cases}$$

- A worst rate for each and every user  $\mathcal{I}_{min,u}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}})$  to compare to the user's implemented rate  $b_u$ .

$$\mathcal{I}_{min,u}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) = \min_{i \in \{1, \dots, U\}} \{ \mathcal{I}_i(\mathbf{x}_u; \mathbf{y}_i / \mathbb{P}_u(\boldsymbol{\pi}_i)) \}$$



The  $\mathcal{I}_{min}$  vector

$$\mathcal{I}_{min}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) = \begin{bmatrix} \mathcal{I}_{min,U}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) \\ \vdots \\ \mathcal{I}_{min,u}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) \\ \vdots \\ \mathcal{I}_{min,1}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) \end{bmatrix}$$

$$\mathbf{b} \preceq \mathcal{I}_{min}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}})$$

- This tacitly implies sum over each user's subuser components.

**Lemma 2.6.1 [Best Decodable Set]** When good codes (with  $\Gamma = 0$  dB), given  $\boldsymbol{\Pi}$  and  $p_{\mathbf{x}\mathbf{y}}$ , and with

$$\mathbf{b} \preceq \mathcal{I}_{min}(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}) \quad , \quad (2.234)$$

then

$$\mathbb{P}_u(\boldsymbol{\pi}_u) \subseteq \mathcal{D}_u(\boldsymbol{\Pi}, p_{\mathbf{x}\mathbf{y}}, \mathbf{b}) \quad (2.235)$$

and receiver  $u$  reliably achieves the data rate  $b = \mathcal{I}_u(\mathbf{x}_u; \mathbf{y}_u / \mathbb{P}_u(\boldsymbol{\pi}_u))$  with order  $\boldsymbol{\pi}_u$ .



# Example: sum of 3 users (MAC)

$$y = x_1 + x_2 + x_3 + n$$

real subsymbols

Order $\Pi$	$b_1$	$b_2$	$b_3$
$[1\ 2\ 3]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\mathcal{E}_1 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\mathcal{E}_1 + \mathcal{E}_2 + \sigma^2}\right)}{2}$
$[1\ 3\ 2]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\mathcal{E}_1 + \mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\mathcal{E}_1 + \sigma^2}\right)}{2}$
$[3\ 1\ 2]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\mathcal{E}_1 + \mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\sigma^2}\right)}{2}$
$[2\ 3\ 1]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\mathcal{E}_2 + \mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\mathcal{E}_2 + \sigma^2}\right)}{2}$
$[2\ 1\ 3]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\mathcal{E}_2 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\mathcal{E}_1 + \mathcal{E}_2 + \sigma^2}\right)}{2}$
$[3\ 2\ 1]^*$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_1}{\mathcal{E}_2 + \mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_2}{\mathcal{E}_3 + \sigma^2}\right)}{2}$	$\frac{\log_2\left(1 + \frac{\mathcal{E}_3}{\mathcal{E}_2 + \mathcal{E}_3 + \sigma^2}\right)}{2}$

Position in order determines whether other signals are noise or pre-decoded and then pre-subtracted.

- With the MAC's one receiver:
  - the  $\pi_u \equiv \pi$  vectors,
  - $U' = U$ ,
  - $U! = 6$  for this example, so there are only 6 orders to consider.
- There are many other situations that simplify also.

If energies are  $\sigma_n^2 = .001$ ,  $\mathcal{E}_1 = 3.072$ ,  $\mathcal{E}_2 = 1.008$ , and  $\mathcal{E}_3 = .015$ , then with Gaussian codes ( $p_x$  Gaussian) the order  $[123]^*$  corresponds to  $b_1 = 1$ ,  $b_2 = 3$ , and  $b_3 = 2$ .



# Optimum Detectors (2.6.3)

- Section 2.6.3 formalizes (general, including non-Gaussian, case) optimum detection.
- There are various integrals/sums and definitions.
- More simply, each user's optimum detector -- for any given order  $\pi$  -- must:
  - first (asymptotically/reliably) detect all other “earlier” users (MMSE $\rightarrow$ MAP, chain rule) reliably (“no errors”), and
  - then consider all “later” users as noise (this generalizes to integration over margin distribution on non AWGNs).
- The error-probability calculation follows single-user, simply with any “pre-users” no longer present and any “post-users” averaged (treated like noise).
- Thus, these are the same 379A decoders – just more complex notation for the multiuser case,
  - with some modulation-level preprocessing.



# Multi-User Detection (MUD) – 2.6.2

- Optimum remains  $\max_{\hat{x}_u} \{p_{x_u/y}\}$  where the  $y$  is the receiver input for detection (MAP detection).

$$P_e(u) = 1 - P_c(u) = 1 - \sum_{i=1}^{M_u} P_{c/i}(u) \cdot p_i(u)$$

$P_{c/i}$  is the probability for message  $i$   
averaged over all the possible  $y$ 's for which  $i$  is selected  
(Decision Region).

- But the receiver now might estimate another user earlier (order), so  $P_e$  becomes order dependent.
- More generally for any given  $p_{xy}$ :
  - All **decodable users**,  $\mathcal{D}_u(\Pi, p_{xy}, \mathbf{b})$ , are first detected and then “cancelled” – they contribute no “noise” (earlier in order). Abbreviate here  $\mathcal{D}_u(\Pi)$  inputs, but remains a function of all 3  $(\Pi, p_{xy}, \mathbf{b})$ .
  - Other users,  $\bar{\mathcal{D}}_u(\Pi) \setminus u$  are not first detected and are “averaged” (treated as noise).

$$p_{\mathbf{x}_u / [\mathbf{y} \mathbf{x}_{i \in \mathcal{D}_u(\Pi)}]}(\chi_u, \mathbf{x}_{i \in \mathcal{D}_u(\Pi)}, \mathbf{y}) = \int_{\chi \in \mathbf{x}_{\{\bar{\mathcal{D}}_u(\Pi) \setminus u\}}} p_{\mathbf{x} / [\mathbf{y} \mathbf{x}_{\{i \in \mathcal{D}_u(\Pi)\}}]}(\chi, \mathbf{x}_{i \in \mathcal{D}_u(\Pi)}, \mathbf{y}) \cdot d\chi$$

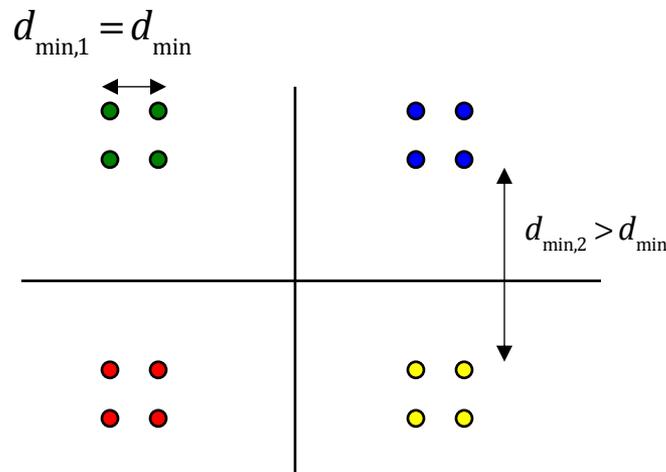
Integration/sum  
is over the noise ave

$$p_{\mathbf{x} / [\mathbf{y} \mathbf{x}_{i \in \{\mathcal{D}_u(\Pi)\}}]}(\chi_u, \chi_{i \in \mathcal{D}_u(\Pi)}, \mathbf{y}) = \frac{p_{\mathbf{y} / \mathbf{x}}(\chi_{i \in \bar{\mathcal{D}}_u(\Pi)}, \mathbf{x}_{i \in \mathcal{D}_u(\Pi)}, \mathbf{y}) \cdot p_{\mathbf{x}}(\chi_{i \in \bar{\mathcal{D}}_u(\Pi)})}{p_{\mathbf{y} / \mathbf{x}}_{\{i \in \mathcal{D}_u(\Pi)\}}(\mathbf{x}_{i \in \mathcal{D}_u(\Pi)}, \mathbf{y})}$$

Term inside integral  
Derives from  $p_{xy}$ .



# Simple Example



**Design can (optimally) reuse all the single-user good codes !!!**

- The decoder should decode first red, green, blue, yellow; this treats the variation within each color as “noise.”
- Then the decoder re-centers the constellation and decides further which of the 4 same-color points.
  - This effectively cancels the noise from the first step.
- Yes, an overall decoder performs the same if earlier decisions are correct, but the basic concept expands.
  - Again, MMSE (which is chain rule) is optimum detector if previous users (asymptotically reliable – no errors) are correct.





# End Lecture 6

# MU Matrix AWGN Channels

- $\mathcal{C}(\mathbf{b})$  for a multi-user AWGN channel  $\mathbf{y} = \mathbf{H} \cdot \mathbf{x} + \mathbf{n}$  will have all users input distributions as Gaussian at the region's (non-zero) boundary,  $\mathcal{C}(\mathbf{b})$ .
  - Each of these points is a mutual information that for each receiver/user  $b_u = \mathcal{I}$  has a chain-rule decomposition.
  - For any subset of output dimensions  $\mathbf{y}$  and any subset of inputs  $\mathbf{x}_u$ ,  $\mathcal{I}(\mathbf{x}; \mathbf{y}) = \mathcal{I}(\mathbf{x}_u; \mathbf{y} / \mathbf{x}_{U \setminus u}) + \mathcal{I}(\mathbf{x}_{U \setminus u}; \mathbf{y})$ .
    - With independent input messages, these are separable and can be separately maximized.
    - The second term is a “single-user,”  $U \setminus u$ , channel, and this channel thus has optimum Gaussian input.
    - The uncanceled users' crosstalk may contribute in MMSE sense to noise, which then is sum of Gaussians that is also Gaussian.
  - (Proof by induction: last user is single-user channel, which has Gaussian; then next to last has Gaussian xtalk and noise, so it also is Gaussian ...), the optimum  $\mathbf{u}$  is also Gaussian. This also works for any user subset  $\mathbf{u}$ . **QED.**

**In general, with user components, treat  $U \rightarrow U'$ .**

