STANFORD

Lecture 1

Introduction & Dimensionality

April 2, 2024

JOHN M. CIOFFI

Hitachi Professor Emeritus (recalled) of Engineering

Instructor EE379B – Spring 2024

Announcements & Agenda

- Announcements
  - People Introductions
  - Web site https://cioffi-group.stanford.edu/ee379b/
  - Chapters 1-8 are used, on-line at class web site (Course Reader)
  - Read Chapter 4
  - EE379A website is also available for review
    - https://cioffi-group.stanford.edu/ee379a/

- Problem Set 1 = PS1 due Wednesday April 12 at 17:00
  1. 2.15 capacity refresher
  2. 4.3 gap-based 1-dimensional channel analysis
  3. 4.18 DMT water-fill loading
  4. 4.7 Simple Water-fill Loading
  5. 4.25 Matrix AWGN & vector coding with water-fill

  Send to kfardi@Stanford.edu with cc to cioffi@Stanford.edu

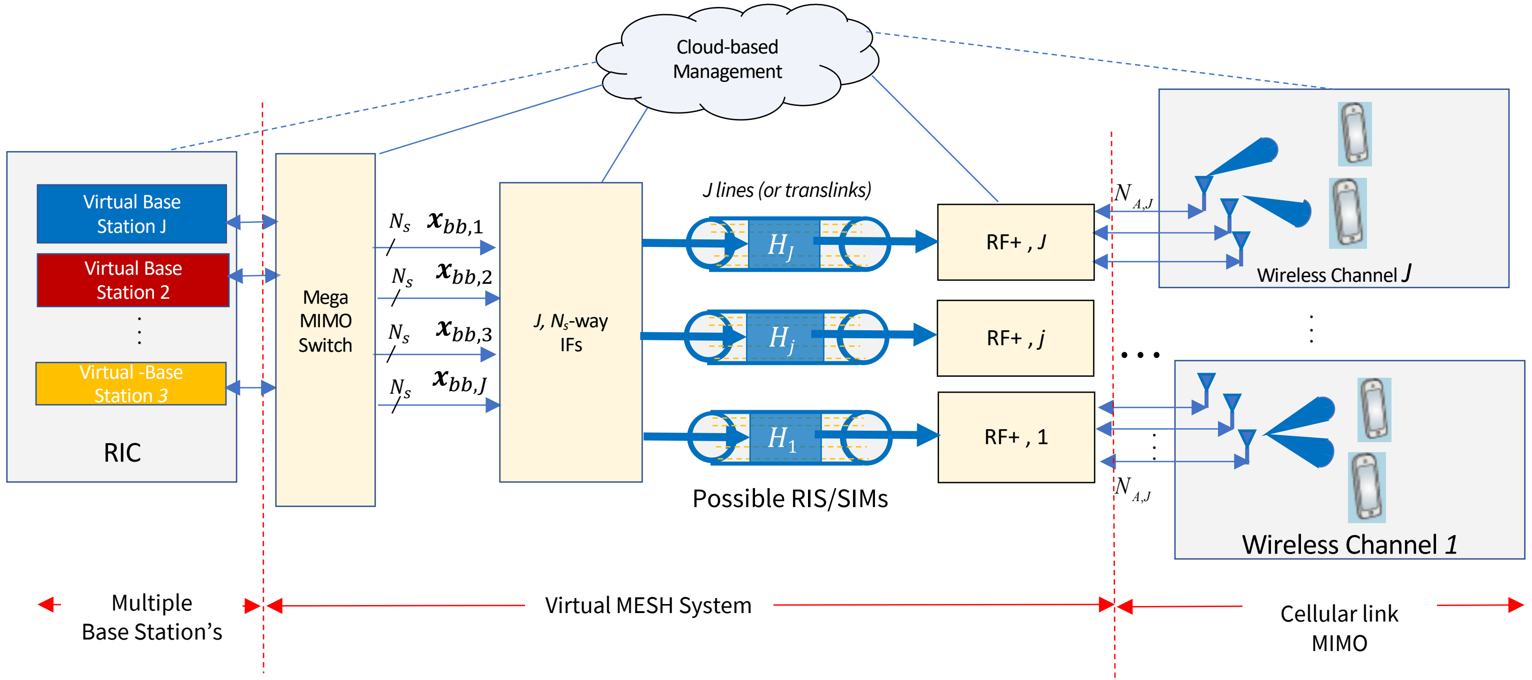
- Today
  - Course introduction
  - The scalar AWGN channel (a foundation)
  - The matrix AWGN channel
  - Water-filling energy distribution
  - Projecting forward

Multiuser Communications

Broadband and Cellular

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

Mega MIMO – Translink Convergence (“Xhaul”)



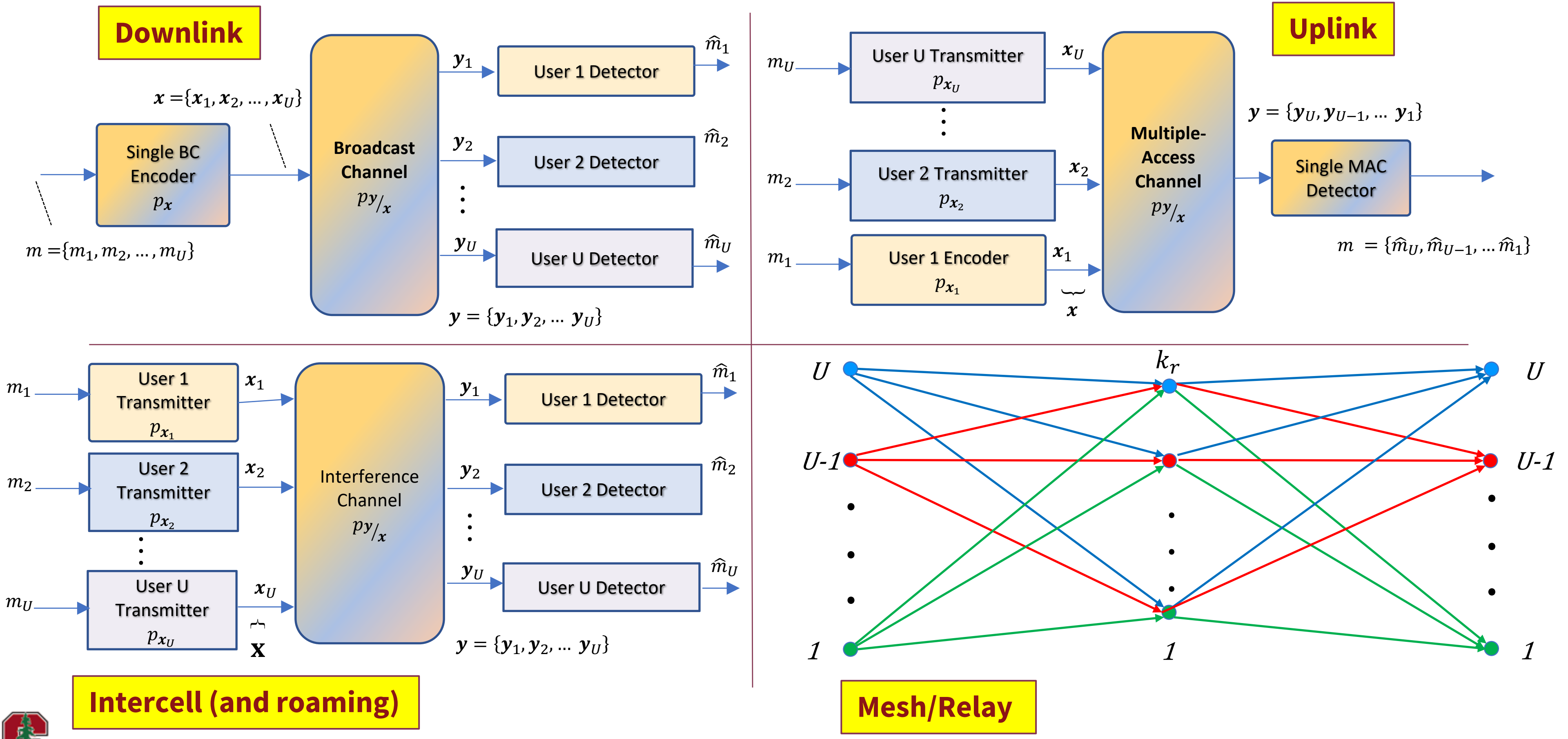
- This supports ”cell-free.
- “Virtualization” (software modulation, coding too!) moves to data-center/edge.

Metaverse Distributed Rendering

- Remote rendering: Devices (smart phones, glasses/goggles) used to augment current environment

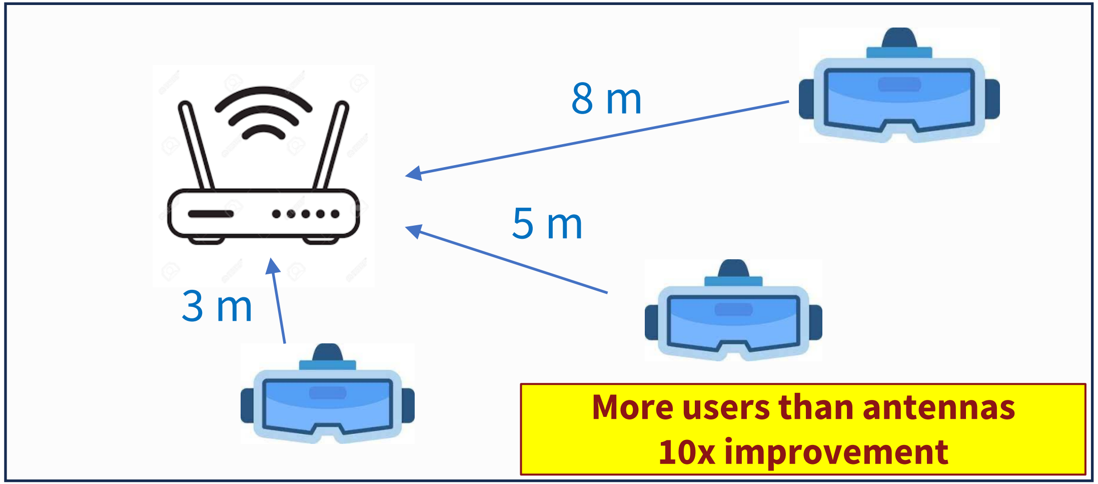
- Games
- Education
  - Instructions
- Health
- **Multiple contributors**

Multiuser Channel Basics (all others are combos)

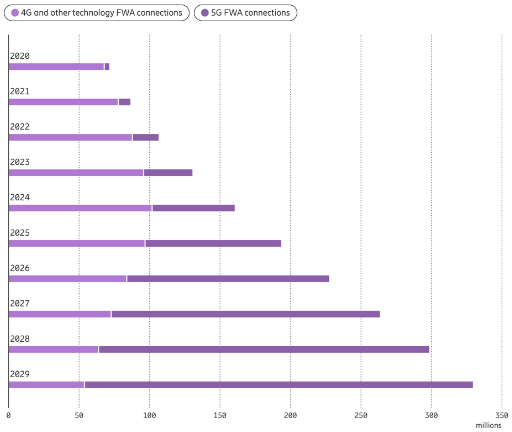


An Example with Wi-Fi

- Basic example that uses:
  - IEEE 802.11 Room B model,
  - 3 users,
  - 2 AP antennas, 1 glasses antenna each.
- Minimum Distributed AR rate
  - 500 Mbps/user in 80 MHz channel
- Similar in Fixed Wireless Access



More users than antennas 10x improvement

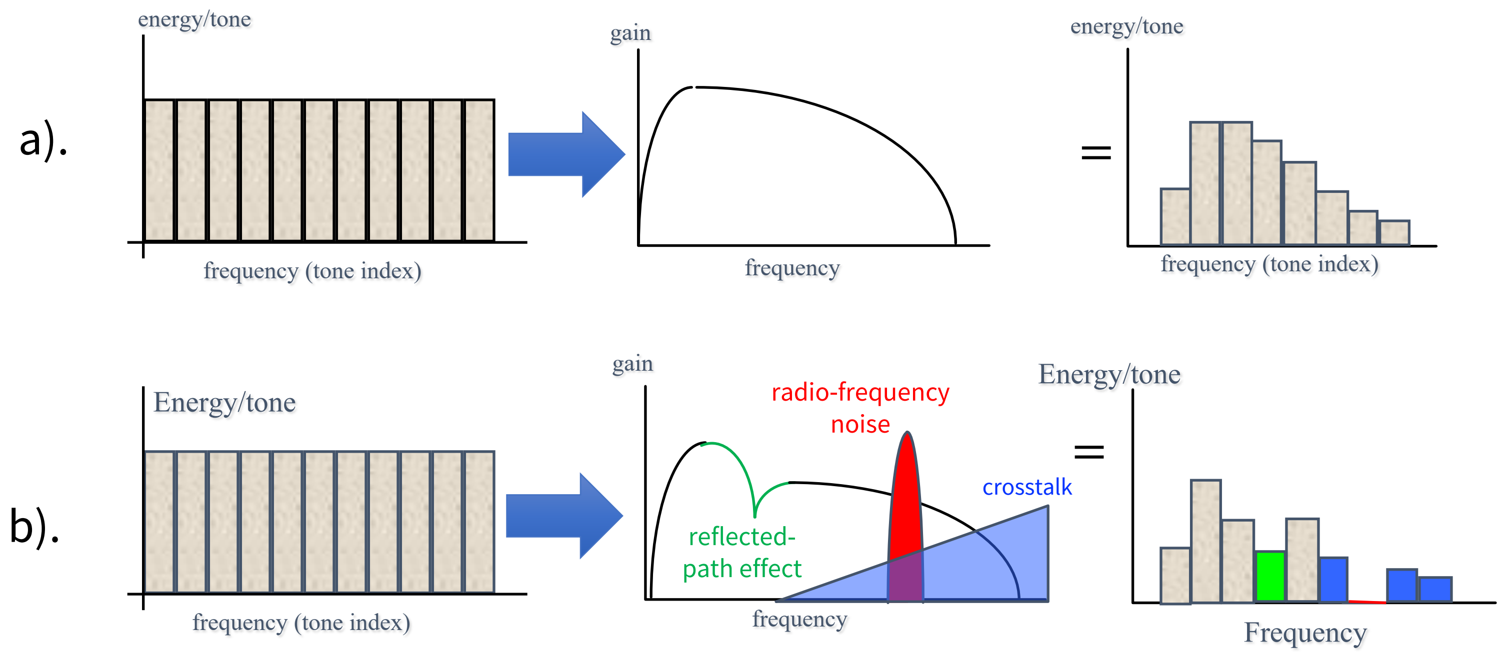


EE379B has custom Matlab that optimizes for best wireless performance.

https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/fwa-outlook#:~:text=Over%20330%20million%20FWA%20connections%20by%202029,expected%20to%20be%20over%205G.

Multicarrier Adaptive Transmitters

- Dynamic adaptation of transmit resource use to the channel situation, “**loading**” or **resource allocation**.

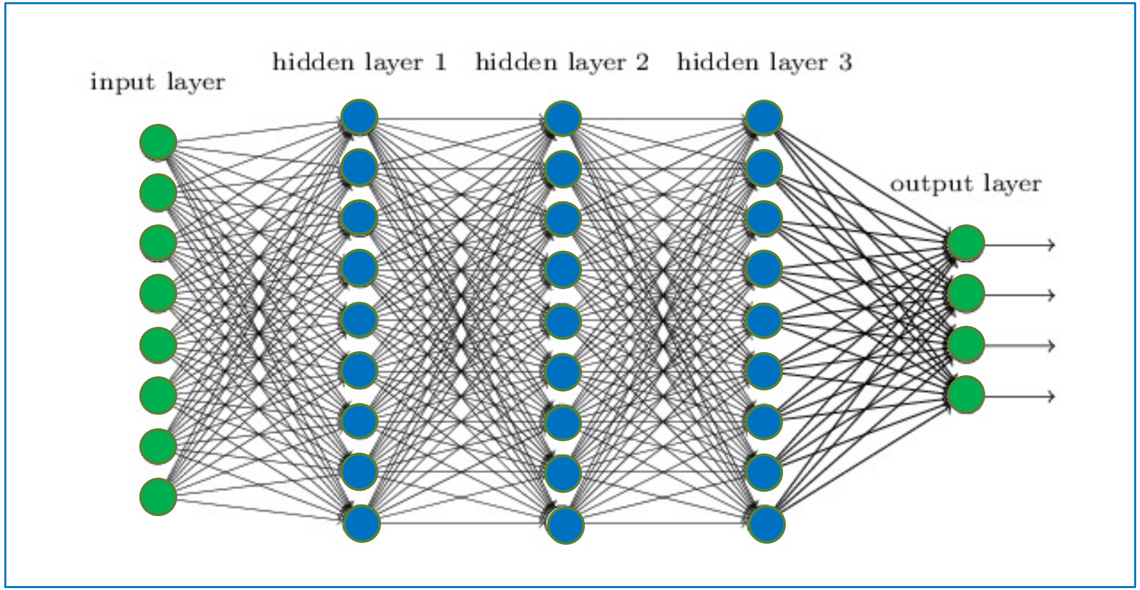


- Frequency expands to include spatial dimensions.

Use of Machine/Deep Learning?

- **Transceivers** will use two **basic operations** (unitary Q and triangular-inverse G^{-1}) in real time.
  1. **Filtering** – beamforming, spectrum adjustment (Q)
  2. **Recursive feedback** – (G)

- **Controller** that assigns resources (energy/information/bits), guides Q and G also.



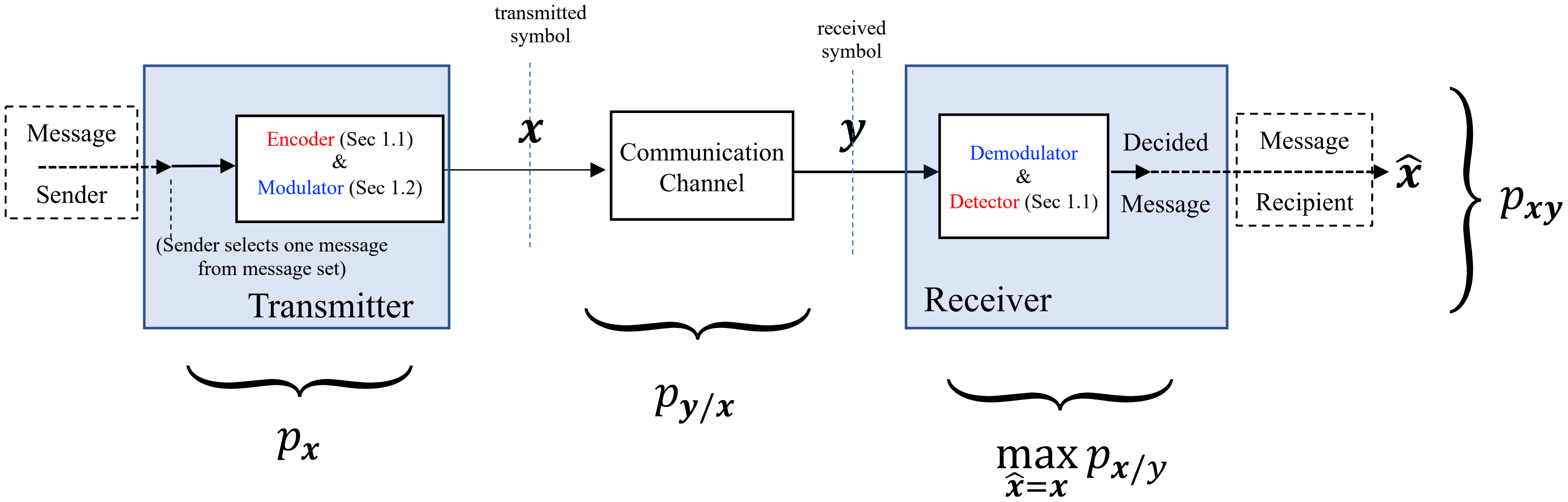
Controller also evolves this way – definitely adaptive optimization is very important – at edge.

The scalar AWGN channel

(a foundation: Section 1.3, Section 2.1-3
direct: 2.4.1, 2.4.3)

See PS1.1 (Prob 2.15 - capacity) and PS1.2 (Prob 4.3 gap)

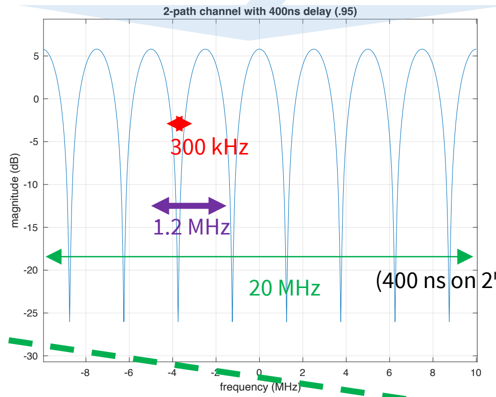
Basic Communication (digital)



- The symbol x and messages are in some 1-to-1 relationship.
- Finding the best $\hat{x}$ and designing x well → this class (good 1-to-1 assumed).
- Most general channel is represented by the conditional probability $p_{y/x}$.
- Most general source description is p_x - together, p_{xy} .
- Optimum detector (minimizes ave error probability) is Maximum a Posteriori (MAP), max $p_{x/y}$.
  - When input distribution is uniform → ML (maximum likelihood), max $p_{y/x}$.

Communication *Dimensionality*

- **Temporal** is Time-Frequency (at any fixed location).
  - 2 x bandwidth = # of dimensions/sec (wireless or wired, including "optical" – all are EM waves)



2G (2 dim/ 3.7 μs)

3G (128 dim/ 533μs)

4G/Wi-Fi (160 dim/4μs)

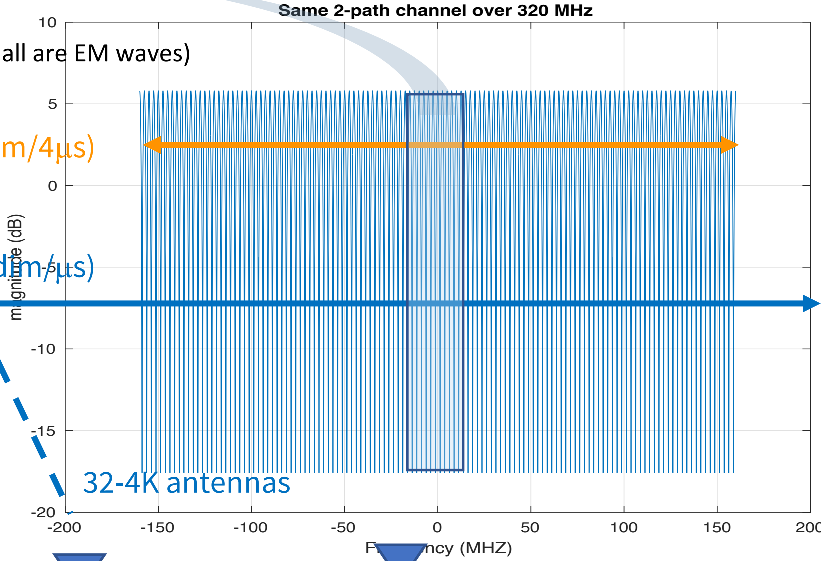
5G / Wi-Fi6 (1280 dim/4μs)

6G / Wi-Fi7 (~1000 dim/μs)

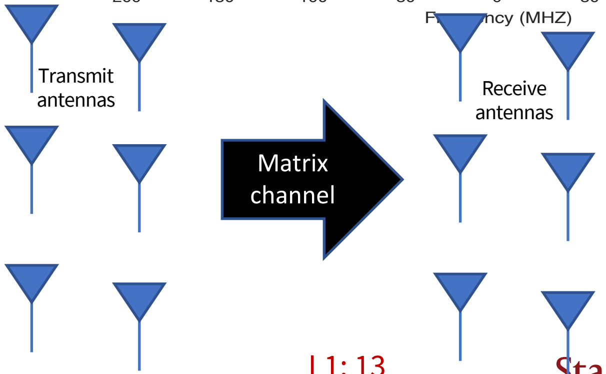
8-32 antennas

32-4K antennas

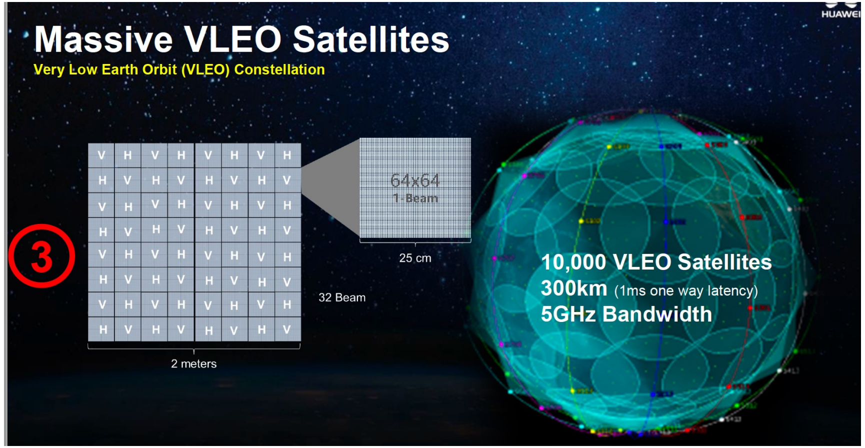
2-4 antennas



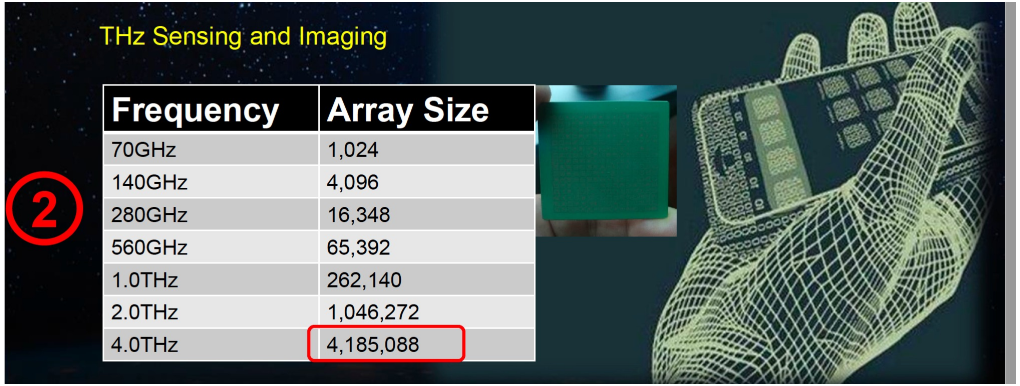
- **Spatial** is space-time (at any frequency).
  - 2D - 3D (at least .....)
  - Antenna spacing is half wavelength or more
  - 10k – 1M dimensions exist over a few microseconds.
  - Number of channels is up to # of antennas "**spatial streams**."
  - There can be **crosstalk** between wires (e.g., ethernet, telco lines).
    - Optical xtalk is between "modes" at same frequency (a spatial effect).



Even More Dimensions (smaller wavelengths)



Courtesy Huawei

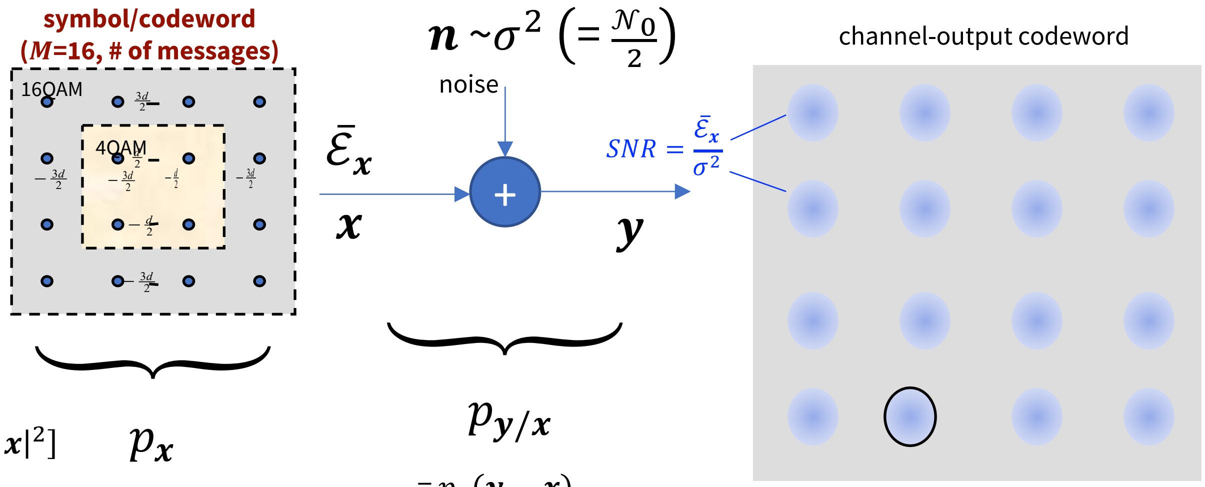


Frequency	Array Size
70GHz	1,024
140GHz	4,096
280GHz	16,348
560GHz	65,392
1.0THz	262,140
2.0THz	1,046,272
4.0THz	4,185,088

- How do we design these systems for best rates (per energy) use?
- How adaptive do they need to be?

Simple Additive White Gaussian Noise Channel

Detection Problem First, every T seconds (symbol period)



$\mathcal{E}_x = \mathbb{E}[|x|^2]$ $\quad p_x$

$p_{y/x}$

$= p_n(y - x)$

$\max_{\hat{x}=x} p_{y/x}$

- Min error probability P_e
- Pick closest point for AWGN
- Maximum Likelihood
  - Uniform input dist'n

"SISO"
(scalar)
$N = 1 \quad \text{or} \quad 2$

- QAM → 2 dimensional
- Uniform input (usually) $p_x = \frac{1}{M}$
- $b = \log_2 M$ **bits/symbol**
- $R = \frac{b}{T}$ bits/second (**data rate**)

- Add noise
- Zero mean
- Variance σ^2 (= 2-sided PSD)

$$P_e = 4 \cdot \left(1 - \frac{1}{\sqrt{M}}\right) \cdot Q\left(\sqrt{\frac{3 \cdot SNR}{M-1}}\right)$$

Subsymbol if coded

$$x \to \tilde{x} \in \mathbb{R}\,;\; x \to \tilde{x} \in \mathbb{C}$$

x has N real dimensions in general, and has $\bar{N}$ subsymbols, of dim $\tilde{N}$.

SNR, QAM, PAM reminders

$$SNR \triangleq \frac{\bar{\mathcal{E}}_x}{\sigma^2} = \frac{\text{single} - \text{sided psd}}{\text{single} - \text{sided psd}} = \frac{\text{two} - \text{sided psd}}{\text{two} - \text{sided psd}}$$

- SNR must have the same number of dimensions in numerator (signal) and denominator (noise).
- Thus, also $SNR \triangleq \frac{\bar{\mathcal{E}}_x}{\sigma^2} = \frac{2 \cdot \bar{\mathcal{E}}_x}{\mathcal{N}_0} = \frac{\mathcal{E}_x}{N \cdot \sigma^2}$ where $\bar{\mathcal{E}}_x$ is energy/real-dimension.
- Energy/dimension essentially generalizes the term power/Hz (= energy) so that is why these quantities are related to power-spectral densities (psd's)
  - 1-sided → power is integral over positive frequencies of psd.
  - 2-sided → power is integral over all frequencies of psd.
  - These two powers are the same.
  - So -40 dBm/Hz (one-sided) psd over 1 MHz is 20 dBm, or 100 mWatts of power , *practice PS1.1 (Prob 2.15) and Homework Helper 1's first part.*
- PAM is always real baseband. QAM is always complex baseband (2 real dimensions).
  - **When QAM** has only 1 bit (2 points) in constellation, it is called BPSK (not binary PAM).
  - PAM's positive-frequency bandwidth is [0, 1/2T) .... x $(1+\alpha)$ when there is $(100 \cdot \alpha)$ percent excess bandwidth.
  - QAM's positive-frequency bandwidth is [-1/2T + f_c , 1/2T + f_c) .... "
  - The PAM system looks like it uses only 1/2 the bandwidth, but the QAM system is really transmitting two dimensions per symbol (so really like 2 PAM systems in parallel with symbol rate 1/T each), so no wonder it takes twice the bandwidth of a single PAM to do so.

Codes and Gaps

Shannon's maximum reliable data rate "capacity" is

$$\mathcal{C} = log_2(1 + SNR) \text{ bits/complex-subsymbol.}$$

AWGN Max bits/sub-sym for $P_e \to 0$ (reliably decodable)

$\tilde{\boldsymbol{x}}_1$	$\tilde{\boldsymbol{x}}_2$	$\tilde{\boldsymbol{x}}_3$	$\cdots$	$\tilde{\boldsymbol{x}}_{\bar{N}}$

$codeword\ (symbol)\ \boldsymbol{x}$

Good Code $\tilde{b} \to \tilde{\mathcal{C}}$ as $\bar{N} \to \infty$.

$subsymbols$ $N = \bar{N} \cdot \tilde{N} =$ # subsymbols x (dim/subsymbol)

Section 2.1.1; also PS1.1 (Prob 2.15)

bits/dim = $\bar{b} = {}^{b}/_{N}$; **bits/subsym =** $\tilde{b} = {}^{b}/_{\bar{N}} = \tilde{N} \cdot \bar{b}$

Code construction

- QAM/PAM operates with given low P_e (10^{-6}) and at a "SNR gap" ($\Gamma = 8.8\ dB$ @10^{-6}) below capacity.
  - See basics in Section 1.3.4 – *for practice, see Section 2.4; also PS1.2 (Prob 4.3).*

$$\tilde{b} = log_2\left(1 + \frac{SNR}{\Gamma}\right) \text{ bits/complex-subsymbol} \leq \tilde{\mathcal{C}}\ .$$

$\frac{3}{2^{\tilde{b}}-1} \cdot SNR = 13.5\ dB$ (from $P_e = 10^{-6}$ formula)

- For all $\tilde{b} > 1$, simple square QAM constellations have constant gap (= 8.8 dB at P_e =10^{-6}).

It's like noise increased or power decreased for P_e
(where Γ approaches 0 dB for best codes)
Gap is function of code and of P_e, not $\tilde{b}$.

Margin

$$\tilde{b} = log_2\left(1 + \frac{SNR}{\Gamma \cdot \gamma_m}\right) \text{ bits/complex-subsymbol } \leq \tilde{C}.$$

See also PS1.2 (Prob 4.3)

- The designer wants "margin" protection against possible noise-power increase.
- **MARGIN** γ_m is this protection (usually in dB), $\gamma_m = \frac{(SNR/\Gamma)}{2^{\tilde{b}}-1}$.

Positive margin – means performing well; **Negative margin** – means not meeting design goals.

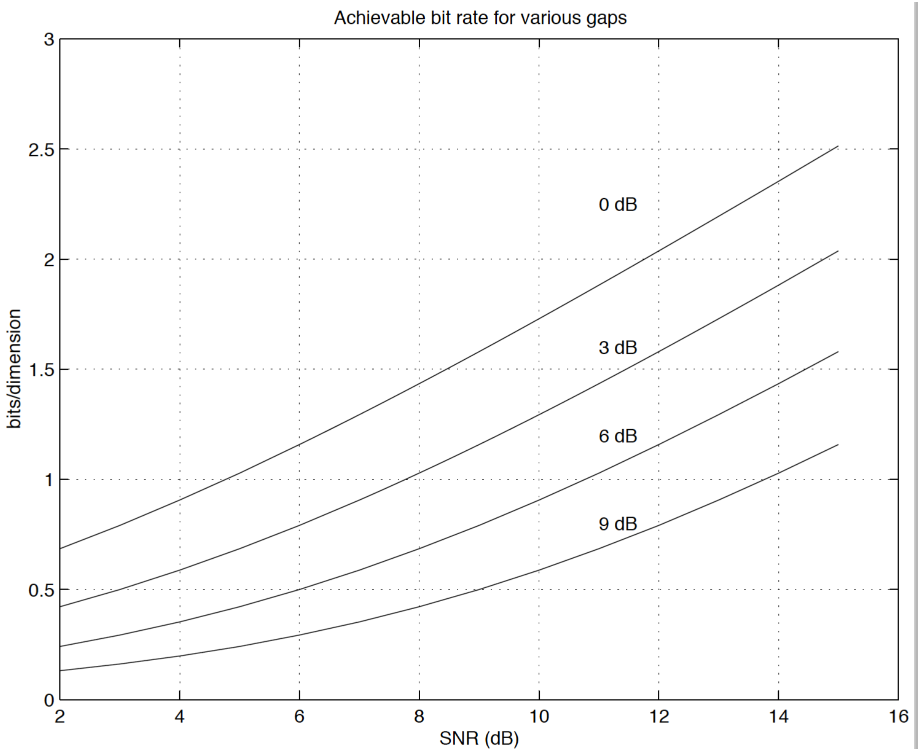
- AWGN with SNR = 20.5 dB, then $\tilde{C} = log_2\,(1 + 10^{2.05})$ = 7 bits/subsymbol.
- Suppose that 16-QAM ($\tilde{b} = 4$) is transmitted @ $P_e = 10^{-6}$ ($\Gamma = 8.8$ dB), then $\gamma_m = \frac{10^{2.05-.88}}{2^4-1} = 0$ dB.
- Suppose instead QAM with $\tilde{b}$ =5 bits/complex-subsymbol with code of 7 dB gain ($\Gamma \rightarrow$ 8.8-7=1.8 dB).
  - $\gamma_m = \frac{10^{2.05-.18}}{2^5-1} = 3.8$ dB.
- 6 bits/subsymbol with same code? → 0.7 dB margin – just barely below the desired P_e ; $\bar{P}_e = {}^{P_e}/_{N}$.

EE 379B

- The simple single-dimension AWGN is fundamental to most all designs.
- All subsequent designs will depend on good codes (small or 0 dB gap) re-use on those single dimension AWGNs.
- Designs can be optimized to get highest possible data rates for Gaussian noise:
  - single user (of course),
  - **all multiuser,**
  - channels with interference between dimensions, which includes
    - intersymbol interference (temporal),
    - crosstalk (spatial), &
    - modal (electromagnetic information theory – near field).
  - Designs are for many users with many antennas, high/low data rates, crosstalking wires, and different locations.
- The gains can be enormous (particularly with respect to EE379A coding gains).

Gap Plot & Example

- The gap is constant, independent of the bits/dimension – greatly simplifies "loading" (adapting transmission codes to the channel).

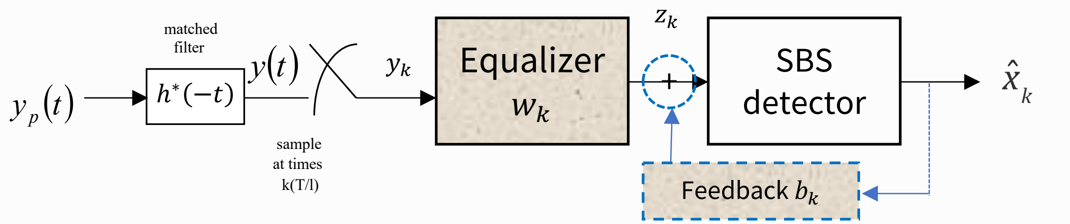


The Matrix AWGN Channel

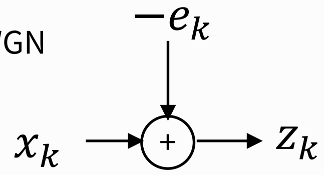
Section 2.3.5

Generating Parallel AWGNs

- Methods from EE379A?
- An “equalizer” is one choice and
  - creates parallel channels in time.
  - $z_k = x_k - e_k$.

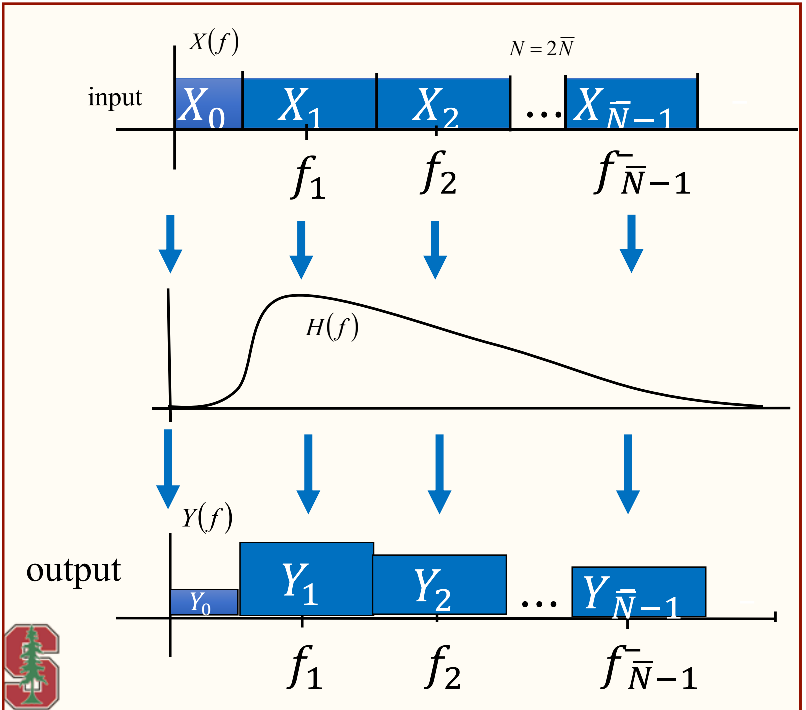


Equivalent AWGN at time k



$$SNR = \frac{\mathcal{E}_x}{\mathbb{E}[|z_k - x_k|^2]}$$

Section 3.6

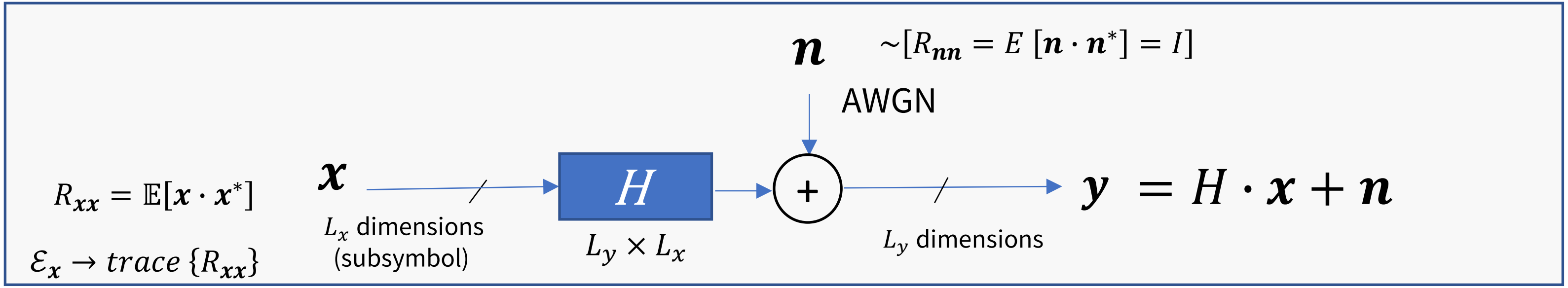


- Another?
- Multicarrier is another choice and
  - creates parallel channels in frequency.

$$Y_n \cong H_n \cdot X_n \quad (+N_n)$$

Sections 1.3.8 and 4.2.1

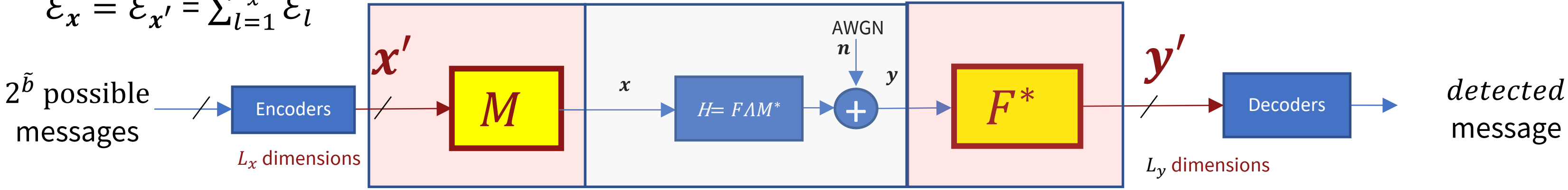
In general, a matrix AWGN channel

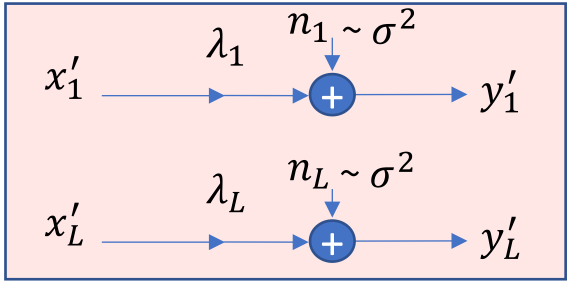


$H = F\cdot\Lambda\cdot M^*$

singular value decomposition (svd in matlab)
$F\cdot F^* = F^*\cdot F = I_{L_y}\,;\, M\cdot M^* = M^*\cdot M = I_{L_x}$
$\Lambda.\ (L_y\times L_x)$ is "diagonal" (real)

$\mathcal{E}_{\boldsymbol{x}} = \mathcal{E}_{\boldsymbol{x}'} = \sum_{l=1}^{L_x}\mathcal{E}_l$





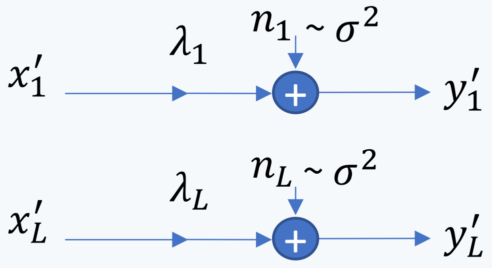
Vector Coding (MIMO)
Kasturia 1989

$L \le min(L_x, L_y)$ independent dimensions

$R_{\boldsymbol{nn}} \ne I \to (H \to R_{\boldsymbol{nn}}^{-1/2}\cdot H)$

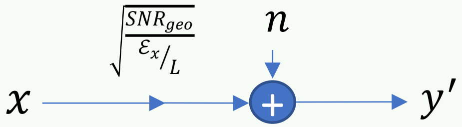
Geometric Equivalent Channel

Parallel Independent



$$\tilde{b}_l = \mathcal{C}_l = log_2(1 + SNR_l) \qquad SNR_l = \mathcal{E}_l \cdot \lambda_l^2 / \sigma^2 = \mathcal{E}_l \cdot g_l$$

$$\tilde{b} = \sum_{l=1}^{L} \tilde{b}_l = \sum_{l=1}^{L} log_2(1 + SNR_l) = L \cdot \log_2(1 + SNR_{geo})$$



$$SNR_{geo} = \Bigg[\underbrace{\prod_{l=1}^{L}(1 + SNR_l)}_{\text{geometric average}}\Bigg]^{1/L} - 1$$

Use it L times like single constant AWGN

- Vector Coding – uses SVD to translate matrix AWGN to set of equivalent parallel AWGN's.
  - Each can be individually encoded like AWGN (they are independent).
- Geometric-equivalent channel is used L times,
  - any H and R_{nn}, &
  - any set of input energies (that sum to allowed energy).

The Detection/Communication Issue

- **Generally,** MAP/ML receiver/detector implementation can be very complex.

- ***Decomposing into multiple channels can simplify design!***
  - Multiple dimensions are the key to this simplification.
  - And today, used throughout digital communication (wires, wireless, soon fiber).
  - And, with proper design, there is no loss in so doing.

The Water-Filling Energy Distribution

Sections 2.3.5, 4.1-4.3

also supplementary lecture S1A

See PS1.3 (Prob 4.18), PS1.4 (Prob 4.7), and PS1.5 (Prob 4.25)

Rate Maximization and Dual

- Transmitter chooses energy and bit allocation to maximize sum data rate over the dimensions $g_l = {[H_l]^2}/{\sigma^2}$.

$$\max_{\bar{\mathcal{E}}_l} \sum_{l=1}^{L} \frac{1}{2} \cdot \log_2\left(1 + \frac{\bar{\mathcal{E}}_l \cdot g_l}{\Gamma}\right) = \sum_{l=1}^{L} \bar{b}_l$$

$$ST: \mathcal{E}_{\boldsymbol{x}} = \sum_{l=1}^{L_x} \bar{\mathcal{E}}_l$$

Rate Adaptive (RA)

DUAL

$$\min_{\tilde{b}_l} \sum_{l=1}^{L} \bar{\mathcal{E}}_l$$

$$ST: b = \sum_{l=1}^{L_x} \tilde{b}_l$$

Margin Adaptive (MA)

- Solution (basic calculus – see Section 4.2) ; see also matlab “waterfill.m” at web site to save hand calcs.

$$\bar{\mathcal{E}}_l + \frac{\Gamma}{g_l} = constant.$$

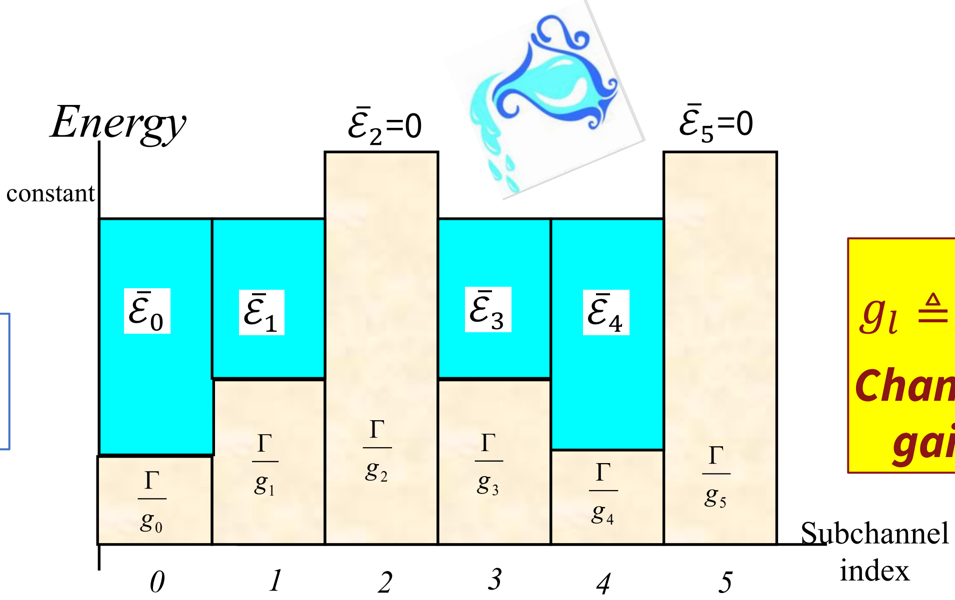
WATER-FILLING

(Shannon 1948)

Neither energies allocated nor bits allocated can be negative.

Water-filling Illustrated

- Energy is available in a pitcher:
  - note re-indexed 0 (DC) to 5.



$g_l \triangleq \frac{\lambda_l^2}{\sigma^2}$ **Channel gain**

RA: until all energy used.
MA: until target bit rate attained.

$$\tilde{b}_l = log_2\left(1 + \frac{SNR_l}{\Gamma}\right)$$ where $SNR_l = \bar{\mathcal{E}}_l \cdot g_l$.

Rate Adaptive Solution

$$g_1 \geq g_2 \geq \ldots \geq g_L$$

- Write and sum energy constraints:

$$\begin{aligned} \mathcal{E}_1 + \Gamma/g_1 &= K \\ \mathcal{E}_2 + \Gamma/g_2 &= K \\ &\vdots \\ \mathcal{E}_L + \Gamma/g_L &= K \end{aligned}$$

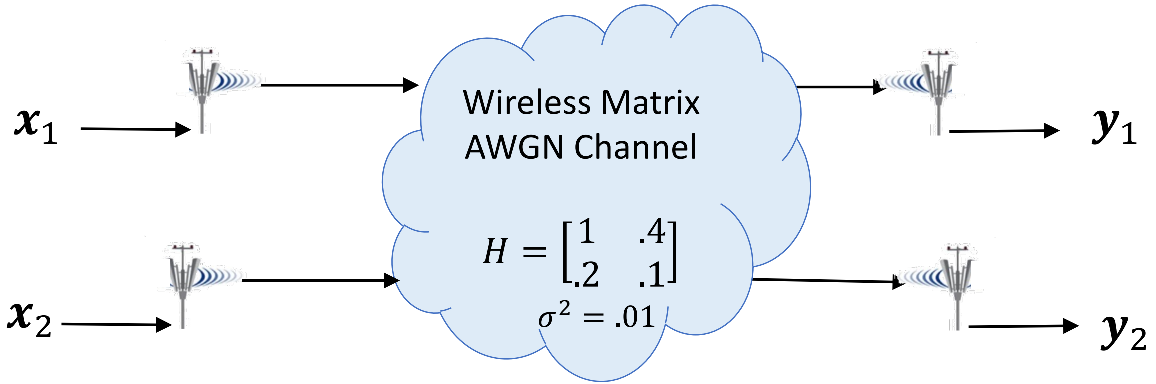
$$\sum_{l=1}^{L} \mathcal{E}_l + \Gamma \cdot \sum_{l=1}^{L} 1/g_l = L \cdot K$$

- Solve for Water-Fill Constant:

$$K = \frac{\mathcal{E}_x}{L^*} + \frac{\Gamma}{L^*} \cdot \sum_{l=1}^{L^*} 1/g_l$$

L^* is largest L such that $\mathcal{E}_l > 0$ for all $l = 1, \ldots, L^*$.

2 x 2 Antenna System with 0 dB gap



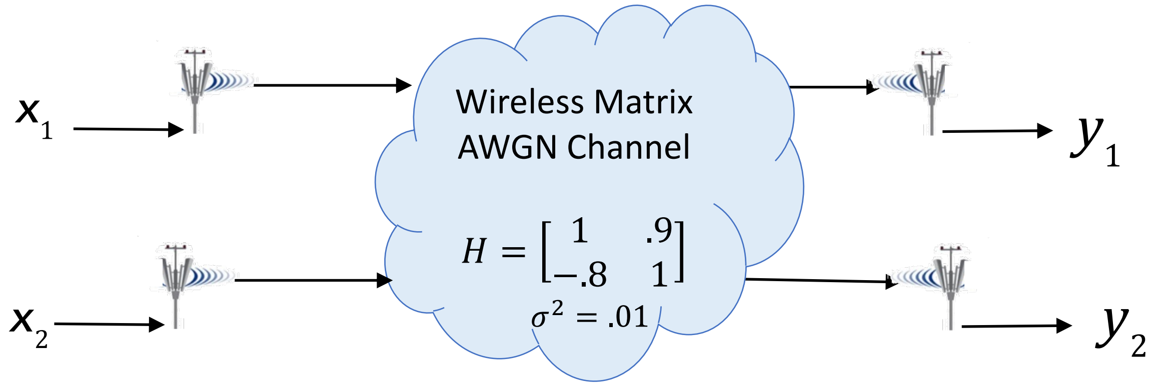
- There is crosstalk between dimensions and $\mathcal{E}_x$=2 .
  - Kind of sounds like a problem then, right?

```
>> H=[10 4
2 1];
>> [F , Lambda , Mstar]=svd(H);
>> Lambda =
  10.9985       0
        0   0.1818
>> g2=Lambda(1,1)^2  =  120.9669
>> g1=Lambda(2,2)^2  =    0.0331
>> K=1+0.5*(1/g1+1/g2) = 16.1250
>> E2=K-1/g2  =   16.1167
>> E1=K-1/g1 =  -14.1167   < 0 (whoops)
```

Just use dimension 2 → $\tilde{b} = \log_2(1 + 2 * g_2) = 6.93$ bits/subsymbol .

In this case water-fill simply puts all energy on the best dimension (returns to scalar/SISO if that is best).

2 x 2 Antenna System



- There is stronger crosstalk between dimensions.
  - Maybe worse, right? ???

```
>> H=[10 9
-8 10];
>> [F , Lambda , Mstar]=svd(H);
>> Lambda =
  13.6244       0
       0  12.6244
>> g2=Lambda(1,1)^2  =  185.6244
>> g1=Lambda(2,2)^2  =  159.3756
>> K=1+0.5*(1/g1+1/g2) =    1.0058
>> E2=K-1/g2  =    1.0004
>> E1=K-1/g1  =    0.9996
>> btilde = log2(1+E2*g2)+log2(1+E1*g1)  =   14.8693
```

Actually this is close to 2x the data rate for the previous case. Clearly, the use of both dimensions, and somewhat stronger crosstalk and signal **improves the best rate.**

In general, the increase is roughly a factor of L in data rate if H has rank L.

Energy-minimizing Margin-Adaptive Solution

$$g_1 \geq g_2 \geq \ldots \geq g_L$$

$$\bar{\mathcal{E}}_l = K - {}^{\Gamma}\!/_{g_l}$$

- Energy and sum-bit constraints

$$\tilde{b} = \sum_{l=1}^{L} \tilde{b}_l = \sum_{l=1}^{L} \log_2\left(1 + \frac{\bar{\mathcal{E}}_l \cdot g_l}{\Gamma}\right)$$

$$= \sum_{l=1}^{L} \log_2\left(\frac{K \cdot g_l}{\Gamma}\right)$$

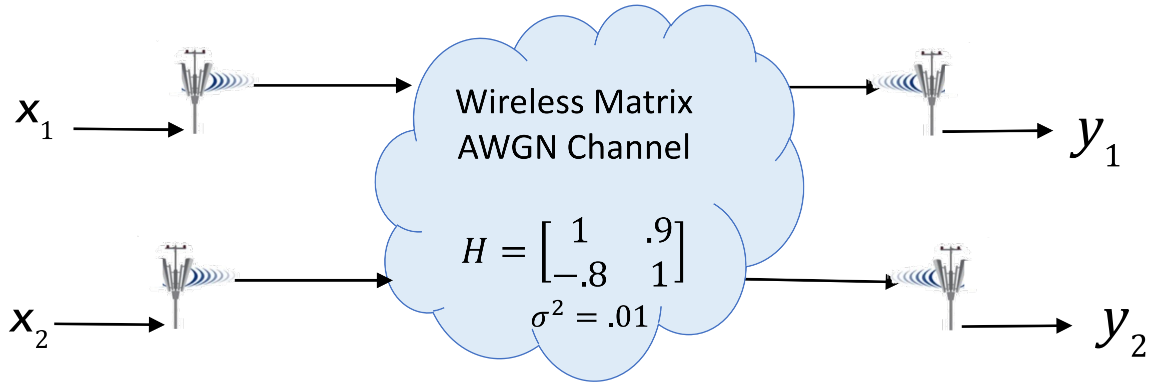
$$= \log_2\left(\prod_{l=1}^{L} \frac{K \cdot g_l}{\Gamma}\right)$$

- Solve for Water-Fill Constant

$$K = \Gamma \cdot \left(\frac{2^{\tilde{b}}}{\prod_{l=1}^{L^*} g_l}\right)^{1/L^*}$$

L^* is largest L such that $\bar{\mathcal{E}}_l > 0$ for all $l = 1, \ldots, L^*$.

2 x 2 Antenna System with MA



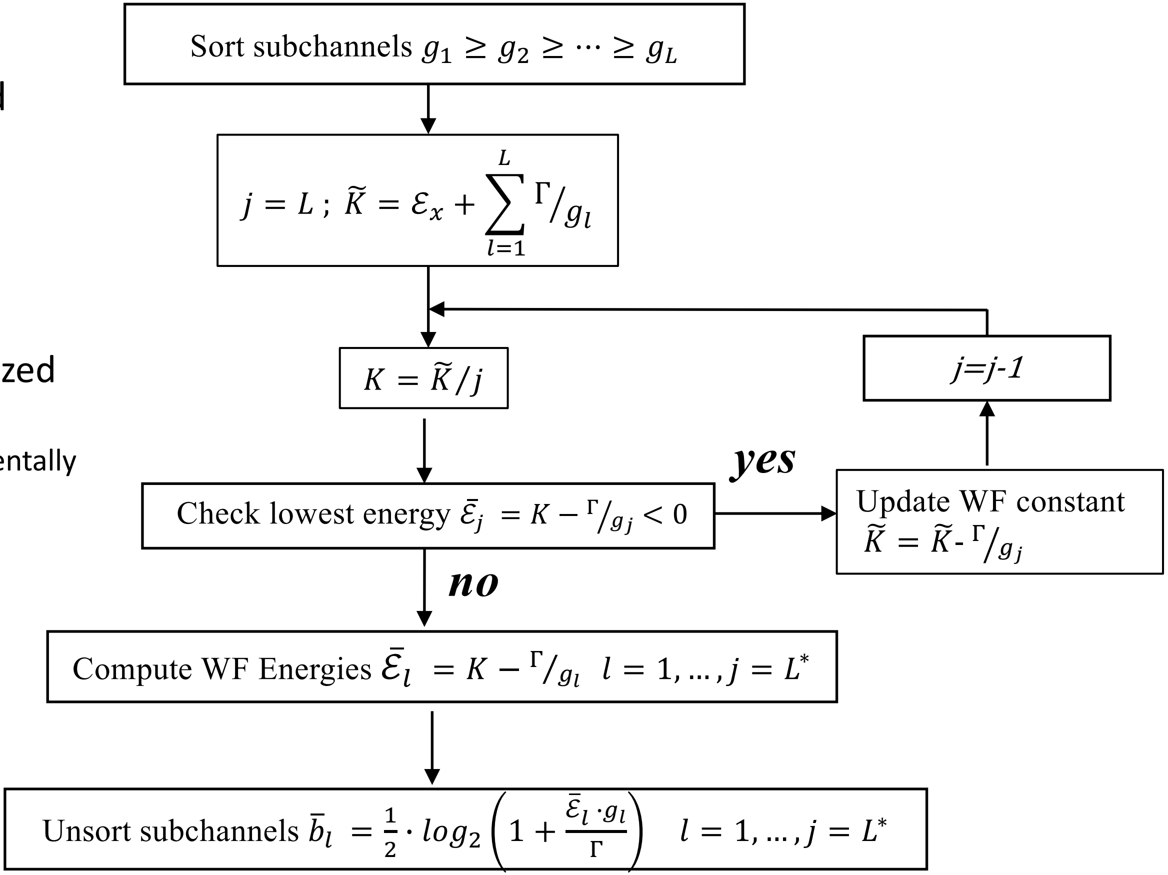
- Attempt $\tilde{b} = 14 \frac{\text{bits}}{\text{Hz}}$; The use of 2 antennas exploited channel's crosstalk,
  - Without the crosstalk, this channel supports only 7 bits/Hz (either channel has then SNR = 10).

```
>> H=[10 9
-8 10];
>> K=sqrt((2^14)/(g1*g2)) =   0.7442
>> E2=K-1/g2  =   0.7388
>> E1=K-1/g1  =   0.7379
>> margin = 10*log10(2/(E1+E2))  =   1.3 dB
```

This effect magnifies as long as most of the singular values are "decent."

RA Water-Fill Flow Chart

- Can start with all channels energized
  - Compute K, test lowest energy
  - Reduce number of dimensions incrementally
- Can also start with 1 channel energized
  - Compute K, test lowest energy
  - Increase number of dimensions incrementally
- The sort is most complex part
  - Can use pivots and bi-section
  - Avoids sort



Margin Adaptive Flowchart

Sort subchannels $g_1 \geq g_2 \geq \cdots \geq g_L$

$i = L$; $\log_2(\tilde{K}) = 2 \cdot b - \sum_{l=1}^{L} \log_2(g_l)$

Compute WF constant: $\log_2(K) = \log_2(\Gamma) + \frac{\log_2(\tilde{K})}{i}$

$\log_2(\tilde{K}) = \log_2(\tilde{K}) + \log_2(g_i)$

Check lowest energy $\bar{\mathcal{E}}_i = K - {}^{\Gamma}/_{g_i} < 0$

yes

$i = i - 1$

no

Compute WF Energies $\bar{\mathcal{E}}_l = K - {}^{\Gamma}/_{g_l}$ $l = 1, \ldots, j = L^*$

Unsort subchannels $\bar{b}_l = \frac{1}{2} \cdot log_2\left(1 + \frac{\mathcal{E}_l \cdot g_l}{\Gamma}\right)$ $l = 1, \ldots, j = L^*$

$$\gamma_{max} = \frac{\mathcal{E}_x}{\sum_{l=1}^{L^*} \bar{\mathcal{E}}_l}$$

Section 4.3.2

STANFORD

End Lecture 1