Algebraic Structure in Network Information Theory

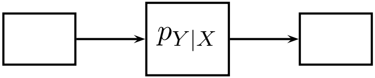
Michael Gastpar* and Bobak Nazer†

*EPFL / Berkeley †Boston University

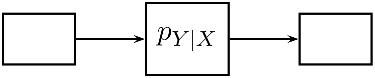
ISIT 2011

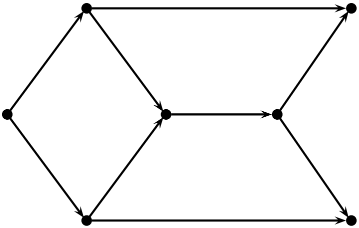
July 31, 2011

Motivation

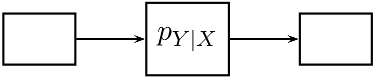


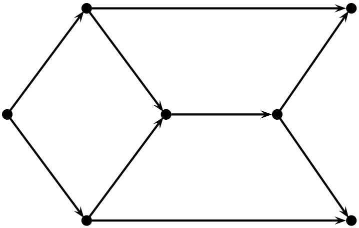
Motivation

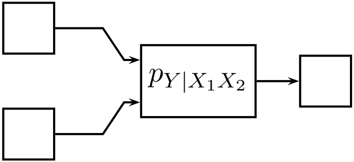


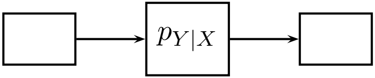


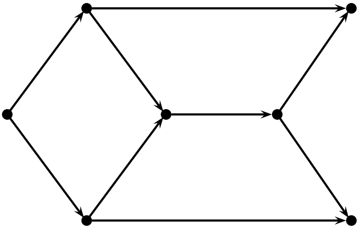
Motivation

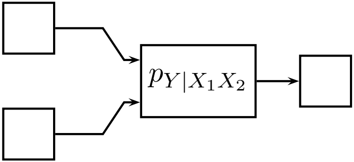


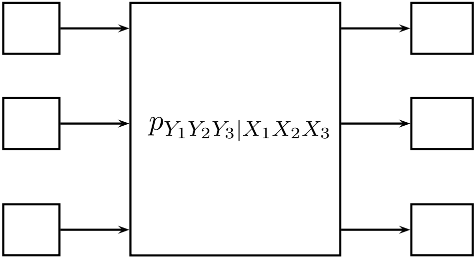





$p_{Y|X}$




$p_{Y|X_1X_2}$


$p_{Y_1Y_2Y_3|X_1X_2X_3}$

In the interest of telling a certain story,

- this tutorial does not attempt to provide an authoritative chronological account of the results;
- this tutorial does not claim to be complete (although a certain effort in this direction was made);

What This Tutorial Is Not *About*

We will *not* address the following very interesting questions (and apologize for a potentially misleading title):

- Complexity of coding schemes
- New families of algebraic codes
- Algebraic coding theory
-

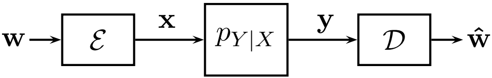
What This Tutorial Is *About*

- Achievable rates that seem out of reach for "classical" arguments.
- Novel communication strategies where algebraic arguments appear to be of key importance.
- Recipes for how to apply these strategies to networks.
- Elements missing from Information Theory books.

I. **Discrete Alphabets**

II. **AWGN Channels**

III. **Network Applications**



The Usual Suspects:

- Message $\mathbf{w} \in \{0,1\}^k$
- Encoder $\mathcal{E} : \{0,1\}^k \to \mathcal{X}^n$
- Input $\mathbf{x} \in \mathcal{X}^n$
- Estimate $\hat{\mathbf{w}} \in \{0,1\}^k$
- Decoder $\mathcal{D} : \mathcal{Y}^n \to \{0,1\}^k$
- Output $\mathbf{y} \in \mathcal{Y}^n$

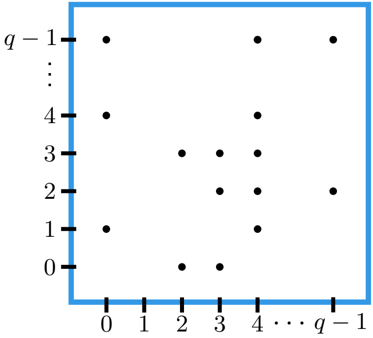
- Memoryless Channel $p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i)$
- Rate $R = \dfrac{k}{n}$.
- (Average) Probability of Error: $\mathbb{P}\{\hat{\mathbf{w}} \neq \mathbf{w}\} \to 0$ as $n \to \infty$. Assume $\mathbf{w}$ is uniform over $\{0,1\}^k$.

- Generate 2^{nR} codewords $\mathbf{x} = [X_1 \; X_2 \; \cdots \; X_n]$ independently and elementwise i.i.d. according to some distribution p_X

$$p(\mathbf{x}) = \prod_{i=1}^{n} p_X(x_i)$$

- Bound the average error probability for a random codebook.

- If the average performance over codebooks is good, there must exist at least one good fixed codebook.



(Weak) Joint Typicality

- Two sequences $\mathbf{x}$ and $\mathbf{y}$ are (weakly) jointly typical if

$$\left| -\frac{1}{n}\log p(\mathbf{x}) - H(X) \right| < \epsilon$$

$$\left| -\frac{1}{n}\log p(\mathbf{y}) - H(Y) \right| < \epsilon$$

$$\left| -\frac{1}{n}\log p(\mathbf{x},\mathbf{y}) - H(X,Y) \right| < \epsilon$$

- For our considerations, weak typicality is convenient as it can also be stated in terms of differential entropies.

- If $\mathbf{x}$ and $\mathbf{y}$ are i.i.d. sequences, the probability that they are jointly typical goes to 1 as n goes to infinity.

Decoder looks for a codeword that is jointly typical with the received sequence $\mathbf{y}$

Error Events

1. Transmitted codeword $\mathbf{x}$ is not jointly typical with $\mathbf{y}$.
   $\Longrightarrow$ Low probability by the Weak Law of Large Numbers.
2. Another codeword $\tilde{\mathbf{x}}$ is jointly typical with $\mathbf{y}$.

Cuckoo's Egg Lemma

Let $\tilde{\mathbf{x}}$ be an i.i.d. sequence that is independent from the received sequence $\mathbf{y}$.

$$\mathbb{P}\Big\{(\tilde{\mathbf{x}}, \mathbf{y}) \text{ is jointly typical}\Big\} \leq 2^{-n(I(X;Y)-3\epsilon)}$$

See **Cover and Thomas**.

- We can upper bound the probability of error via the union bound:

$$\mathbb{P}\{\hat{\mathbf{w}} \neq \mathbf{w}\} \leq \sum_{\tilde{\mathbf{w}} \neq \mathbf{w}} \mathbb{P}\Big\{(\mathbf{x}(\tilde{\mathbf{w}}), \mathbf{y}) \text{ is jointly typical.}\Big\}$$

$$\leq 2^{-n(I(X;Y)-R-3\epsilon)} \qquad \leftarrow \text{Cuckoo's Egg Lemma}$$

- If $R < I(X;Y)$, then the probability of error can be driven to zero as the blocklength increases.

Theorem (Shannon '48)

The capacity of a point-to-point channel is $C = \max_{p_X} I(X;Y)$.

- Linear Codebook: A linear map between messages and codewords (instead of a lookup table).

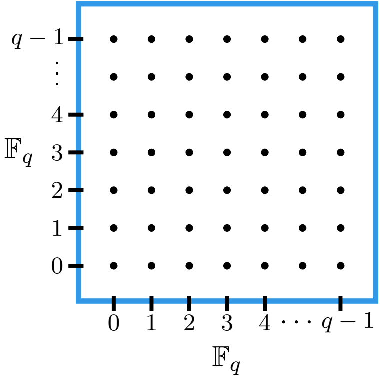
q-ary Linear Codes

- Represent message $\mathbf{w}$ as a length-k vector over $\mathbb{F}_q$.
- Codewords $\mathbf{x}$ are length-n vectors over $\mathbb{F}_q$.
- Encoding process is just a matrix multiplication, $\mathbf{x} = \mathbf{G}\mathbf{w}$.

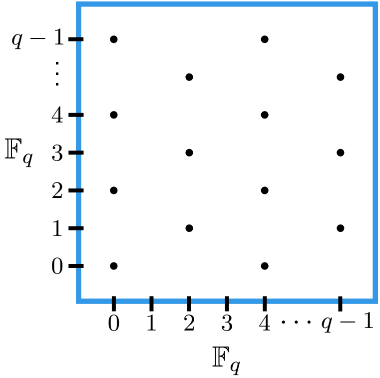
$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1k} \\ g_{21} & g_{22} & \cdots & g_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nk} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix}$$

- Recall that, for prime q, operations over $\mathbb{F}_q$ are just $\mod q$ operations over the reals.
- Rate $R = \frac{k}{n} \log q$

- Linear code looks like a regular subsampling of the elements of $\mathbb{F}_q^n$.
- **Random linear code:** Generate each element g_{ij} of the generator matrix $\mathbf{G}$ elementwise i.i.d. according to a uniform distribution over $\{0, 1, 2, \ldots, q-1\}$.
- How are the codewords distributed?



- Linear code looks like a regular subsampling of the elements of $\mathbb{F}_q^n$.
- **Random linear code:** Generate each element g_{ij} of the generator matrix $\mathbf{G}$ elementwise i.i.d. according to a uniform distribution over $\{0, 1, 2, \ldots, q-1\}$.
- How are the codewords distributed?



It is convenient to instead analyze the shifted ensemble $\bar{\mathbf{x}} = \mathbf{G}\mathbf{w} \oplus \mathbf{v}$ where $\mathbf{v}$ is an i.i.d. uniform sequence. (See **Gallager.**)

Shifted Codeword Properties

1. Marginally uniform over $\mathbb{F}_q^n$. For a given message $\mathbf{w}$, the codeword $\bar{\mathbf{x}}$ looks like an i.i.d. uniform sequence.

$$\mathbb{P}\{\bar{\mathbf{x}} = \mathsf{x}\} = \frac{1}{q^n} \quad \text{for all } \mathsf{x} \in \mathbb{F}_q^n$$

2. Pairwise independent. For $\mathbf{w}_1 \neq \mathbf{w}_2$, codewords $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2$ are independent.

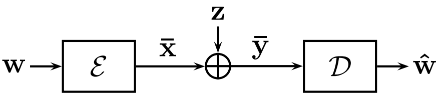
$$\mathbb{P}\{\bar{\mathbf{x}}_\mathbf{1} = \mathsf{x}_1, \bar{\mathbf{x}}_\mathbf{2} = \mathsf{x}_2\} = \frac{1}{q^{2n}} = \mathbb{P}\{\bar{\mathbf{x}}_1 = \mathsf{x}_1\}\mathbb{P}\{\bar{\mathbf{x}}_2 = \mathsf{x}_2\}$$

- Cuckoo's Egg Lemma only requires independence between the true codeword $\mathbf{x}(\mathbf{w})$ and the other codeword $\mathbf{x}(\tilde{\mathbf{w}})$. From the union bound:

$$\begin{aligned}\mathbb{P}\{\hat{\mathbf{w}} \neq \mathbf{w}\} &\leq \sum_{\tilde{\mathbf{w}} \neq \mathbf{w}} \mathbb{P}\Big\{(\mathbf{x}(\tilde{\mathbf{w}}), \mathbf{y}) \ \text{ is jointly typical.}\Big\} \\ &\leq 2^{-n(I(X;Y)-R-3\epsilon)}\end{aligned}$$

- This is exactly what we get from pairwise independence.
- Thus, there exists a good fixed generator matrix $\mathbf{G}$ and shift $\mathbf{v}$ for any rate $R < I(X;Y)$ where X is uniform.

Removing the Shift



- For a binary symmetric channel (BSC), the output can be written as the modulo sum of the input plus i.i.d. Bernoulli(p) noise,

$$\bar{\mathbf{y}} = \bar{\mathbf{x}} \oplus \mathbf{z}$$
$$\bar{\mathbf{y}} = \mathbf{G}\mathbf{w} \oplus \mathbf{v} \oplus \mathbf{z}$$

- Due to this symmetry, the probability of error depends *only* on the realization of the noise vector $\mathbf{z}$.
  $\implies$ For a BSC, $\mathbf{x} = \mathbf{G}\mathbf{w}$ is a good code as well.

- We can now assume the existence of good generator matrices for channel coding.

Random I.I.D. vs. Random Linear

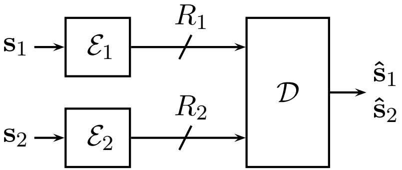
- What have we gotten for linearity (so far)?
  Simplified encoding. (Decoder is still quite complex.)

- What have we lost?
  Can only achieve $R = I(X;Y)$ for uniform X instead of $\max_{p_X} I(X;Y)$.

- In fact, this is a fundamental limitation of group codes, **Ahlswede '71**.

- Workarounds: symbol remapping **Gallager '68**, nested linear codes

- Are random linear codes strictly worse than random i.i.d. codes?



- Joint i.i.d. sources $p(\mathbf{s}_1, \mathbf{s}_2) = \prod_{i=1}^{m} p_{S_1 S_2}(s_{1i}, s_{2i})$
- **Rate Region:** Set of rates (R_1, R_2) such that the encoders can send $\mathbf{s}_1$ and $\mathbf{s}_2$ to the decoder with vanishing probability of error

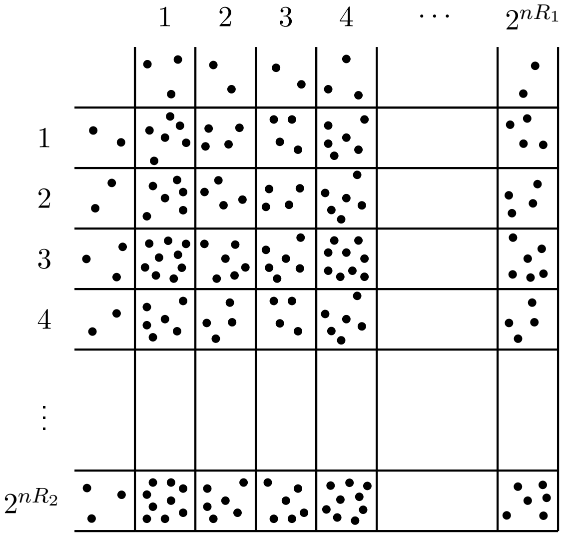
$$\mathbb{P}\{(\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2) \neq (\mathbf{s}_1, \mathbf{s}_2)\} \to 0 \text{ as } m \to \infty$$

- Codebook 1: Independently and uniformly assign each source sequence $\mathbf{s}_1$ to a label $\{1, 2, \ldots, 2^{mR_1}\}$
- Codebook 2: Independently and uniformly assign each source sequence $\mathbf{s}_2$ to a label $\{1, 2, \ldots, 2^{mR_2}\}$
- Decoder: Look for jointly typical pair $(\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2)$ within the received bin. Union bound:

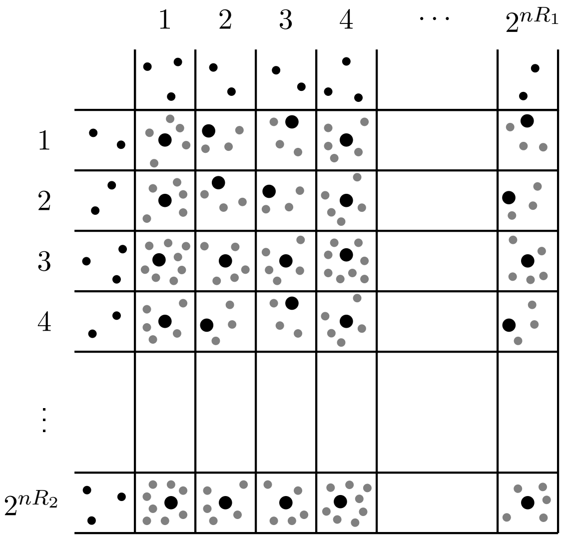
$$\mathbb{P}\Big\{\text{jointly typical } (\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2) \neq (\mathbf{s}_1, \mathbf{s}_2) \text{ in bin } (\ell_1, \ell_2)\Big\}$$
$$\leq \sum_{\text{jointly typical } (\tilde{\mathbf{s}}_1, \tilde{\mathbf{s}}_2)} 2^{-m(R_1+R_2)}$$
$$\leq 2^{m(H(S_1,S_2)+\epsilon)} 2^{-m(R_1+R_2)}$$

- Need $R_1 + R_2 > H(S_1, S_2)$.
- Similarly, $R_1 > H(S_1|S_2)$ and $R_2 > H(S_2|S_1)$

Slepian-Wolf Problem: Binning Illustration



Slepian-Wolf Problem: Binning Illustration



Random Linear Binning

- Assume source symbols take values in $\mathbb{F}_q$.
- Codebook 1: Generate matrix $\mathbf{G}_1$ with i.i.d. uniform entries drawn from $\mathbb{F}_q$. Each sequence $\mathbf{s}_1$ is binned via matrix multiplication, $\mathbf{w}_1 = \mathbf{G}_1\mathbf{s}_1$.
- Codebook 2: Generate matrix $\mathbf{G}_2$ with i.i.d. uniform entries drawn from $\mathbb{F}_q$. Each sequence $\mathbf{s}_2$ is binned via matrix multiplication, $\mathbf{w}_2 = \mathbf{G}_2\mathbf{s}_2$.
- Bin assignments are uniform and pairwise independent (except for $\mathbf{s}_\ell = \mathbf{0}$)
- Can apply the same union bound analysis as random binning.

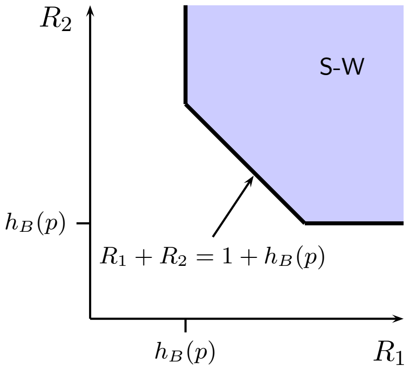
Slepian-Wolf Rate Region

Slepian-Wolf Theorem

Reliable compression possible if and only if:

$$R_1 \geq H(S_1|S_2) = h_B(p)$$
$$R_2 \geq H(S_2|S_1) = h_B(p)$$
$$R_1 + R_2 \geq H(S_1, S_2) = 1 + h_B(p)$$

Random linear binning is as good as random i.i.d. binning!

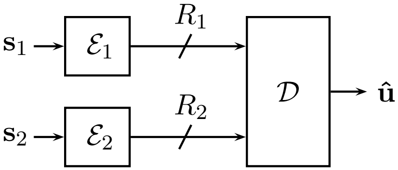


Example: Doubly Symmetric Binary Source

$S_1 \sim \text{Bern}(1/2) \qquad U \sim \text{Bern}(p) \qquad S_2 = S_1 \oplus U$

Körner-Marton Problem

- Binary sources
- $\mathbf{s}_1$ is i.i.d. Bernoulli$(1/2)$
- $\mathbf{s}_2$ is $\mathbf{s}_1$ corrupted by Bernoulli(p) noise
- Decoder wants the modulo-2 sum .



$\mathbf{u} = \mathbf{s}_1 \oplus \mathbf{s}_2$

Rate Region: Set of rates (R_1, R_2) such that there exist encoders and decoders with vanishing probability of error

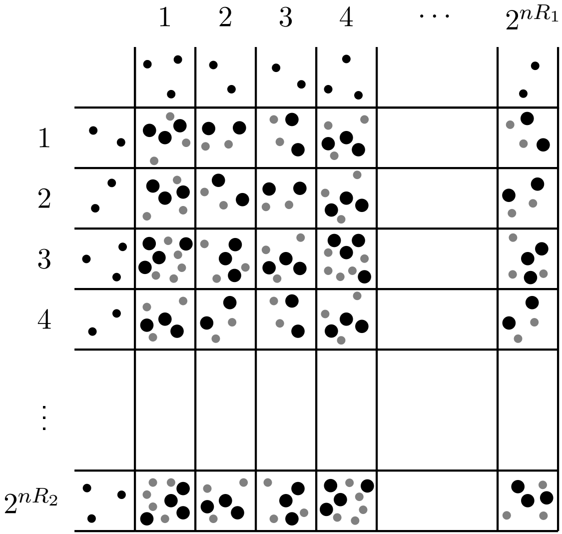
$$\mathbb{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\} \to 0 \text{ as } m \to \infty$$

Are any rate savings possible over sending $\mathbf{s}_1$ and $\mathbf{s}_2$ in their entirety?

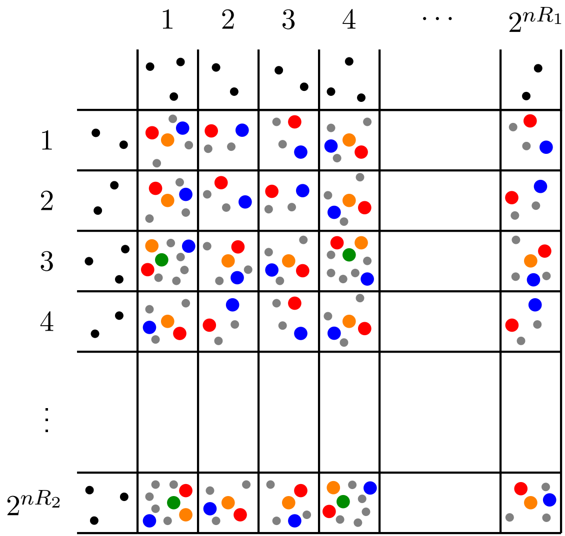
Random Binning

- Sending $\mathbf{s}_1$ and $\mathbf{s}_2$ with random binning requires $R_1 + R_2 > 1 + h_B(p)$?
- What happens if we use rates such that $R_1 + R_2 < 1 + h_B(p)$?
- There will be exponentially many pairs $(\mathbf{s}_1, \mathbf{s}_2)$ in each bin!
- This would be fine if all pairs in a bin have the same sum, $\mathbf{s}_1 + \mathbf{s}_2$. But this probability goes to zero exponentially fast!

Körner-Marton Problem: Random Binning Illustration



Körner-Marton Problem: Random Binning Illustration



Linear Binning

- Use the same random matrix $\mathbf{G}$ for linear binning at each encoder:

$$\mathbf{w}_1 = \mathbf{G}\mathbf{s}_1 \qquad \mathbf{w}_2 = \mathbf{G}\mathbf{s}_2$$

- Idea from **Körner-Marton '79:** Decoder adds up the bins.

$$\begin{aligned} \mathbf{w}_1 \oplus \mathbf{w}_2 &= \mathbf{G}\mathbf{s}_1 \oplus \mathbf{G}\mathbf{s}_2 \\ &= \mathbf{G}(\mathbf{s}_1 \oplus \mathbf{s}_2) \\ &= \mathbf{G}\mathbf{u} \end{aligned}$$

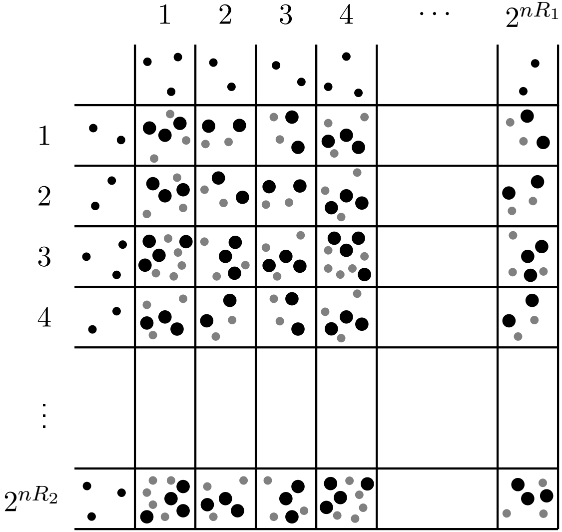
- $\mathbf{G}$ is good for compressing $\mathbf{u}$ if $R > H(U) = h_B(p)$.

Körner-Marton Theorem

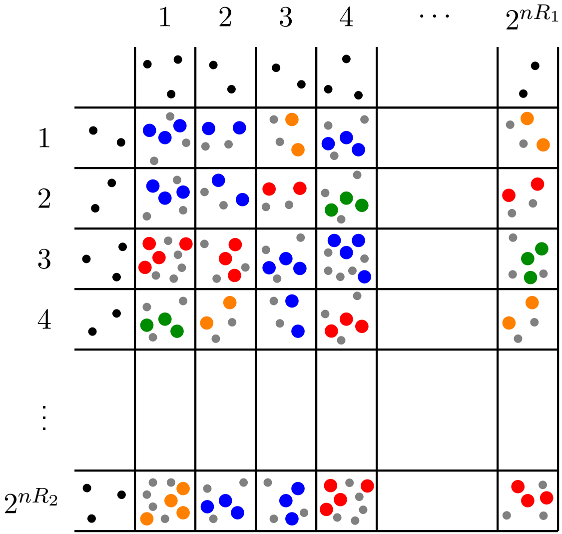
Reliable compression of the sum is possible if and only if:

$$R_1 \geq h_B(p) \qquad R_2 \geq h_B(p) \ .$$

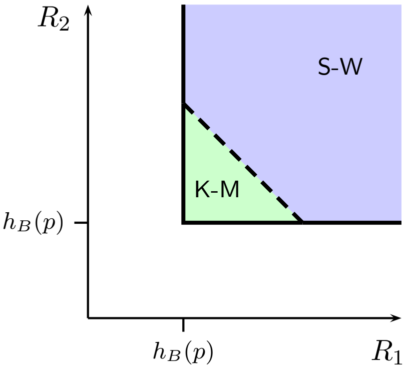
Körner-Marton Problem: Linear Binning Illustration



Körner-Marton Problem: Linear Illustration

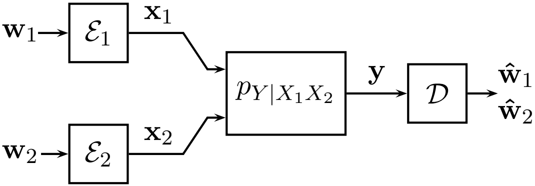


Körner-Marton Rate Region



Linear codes can improve performance!

(for distributed computation of dependent sources)



- **Rate Region:** Set of rates (R_1, R_2) such that the encoders can send $\mathbf{w}_1$ and $\mathbf{w}_2$ to the decoder with vanishing probability of error

$$\mathbb{P}\{(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2) \neq (\mathbf{w}_1, \mathbf{w}_2)\} \to 0 \text{ as } m \to \infty$$

- Cuckoo's egg lemma applies to all three error events.
- For example, event that only $\hat{\mathbf{w}}_1$ is wrong:

$$\mathbb{P}\{\hat{\mathbf{w}}_1 \neq \mathbf{w}_1, \hat{\mathbf{w}}_2 = \mathbf{w}_2\} \leq \sum_{\tilde{\mathbf{w}}_1 \neq \mathbf{w}_1} \mathbb{P}\Big\{(\mathbf{x}_1(\tilde{\mathbf{w}}_1), \mathbf{x}_2(\mathbf{w}_2), \mathbf{y}) \text{ jointly typical}\Big\}$$
$$\leq 2^{-n(I(X_1;Y|X_2)-R_1-3\epsilon)}$$

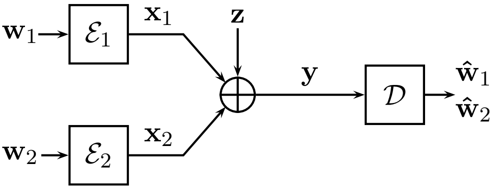
Rate Region (Ahlswede, Liao)

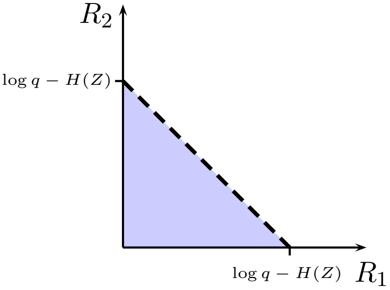
Convex closure of all (R_1, R_2) satisfying

$$R_1 < I(X_1;Y|X_2)$$
$$R_2 < I(X_2;Y|X_1)$$
$$R_1 + R_2 < I(X_1, X_2;Y)$$

for some $p(x_1)p(x_2)$.

- Linear codes can achieve any rate available for uniform $p(x_1), p(x_2)$.
- For finite field MACs, can achieve the whole capacity region.





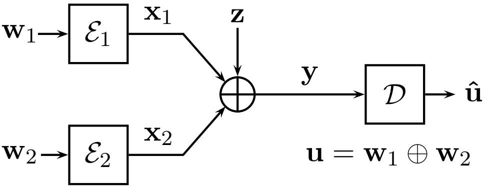
- Receiver observes noisy modulo sum of codewords $\mathbf{y} = \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \mathbf{z}$

Finite Field MAC Rate Region

All rates (R_1, R_2) satisfying

$$R_1 + R_2 \leq \log q - H(Z)$$

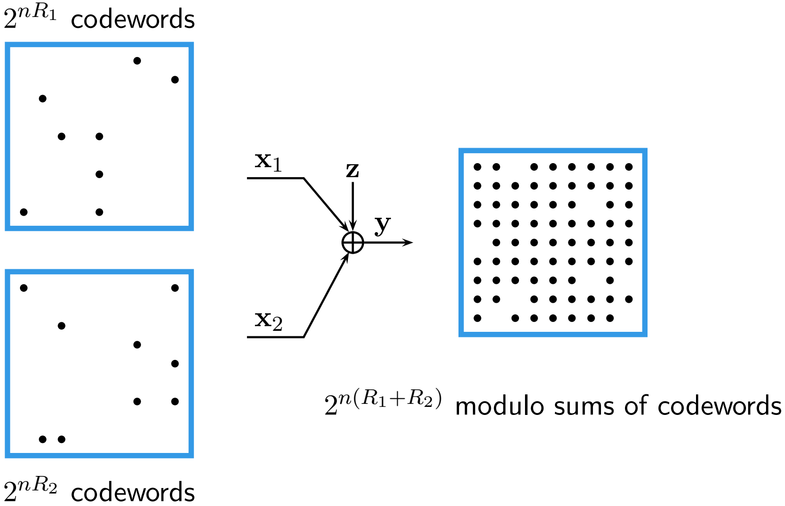
- Independent msgs $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{F}_q^k$.
- Want the sum $\mathbf{u} = \mathbf{w}_1 \oplus \mathbf{w}_2$ with vanishing prob. of error $\mathbb{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\} \to 0$



I.I.D. Random Coding

- Generate 2^{nR_1} i.i.d. uniform codewords for user 1.
- Generate 2^{nR_2} i.i.d. uniform codewords for user 2.
- With high probability, (nearly) all sums of codewords are distinct.
- This is ideal for multiple-access but not for computation.
- Need $R_1 + R_2 \leq \log q - H(Z)$

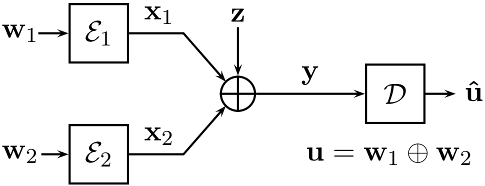
Random i.i.d. codes are not good for computation



Computation over Finite Field Multiple-Access Channels

Independent msgs $\mathbf{w}_1, \mathbf{w}_2$.

Want the sum $\mathbf{u} = \mathbf{w}_1 \oplus \mathbf{w}_2$ with vanishing prob. of error $\mathbb{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\} \to 0$



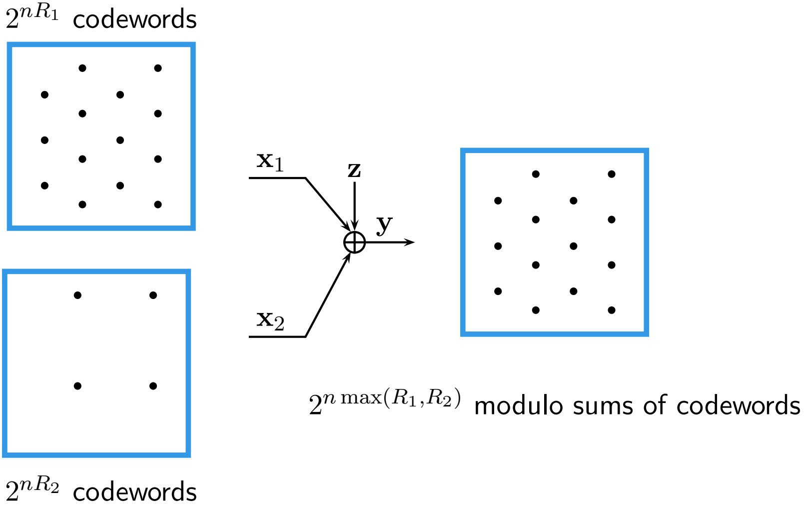
Random Linear Coding

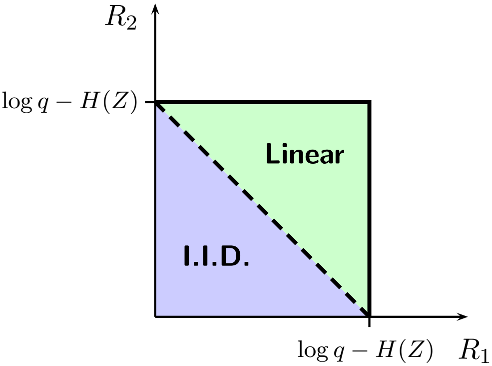
- Same linear code at both transmitters $\mathbf{x}_1 = \mathbf{G}\mathbf{w}_1$, $\mathbf{x}_2 = \mathbf{G}\mathbf{w}_2$.
- Sums of codewords are themselves codewords:

$$\begin{aligned} \mathbf{y} &= \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \mathbf{z} \\ &= \mathbf{G}\mathbf{w}_1 \oplus \mathbf{G}\mathbf{w}_2 \oplus \mathbf{z} \\ &= \mathbf{G}(\mathbf{w}_1 \oplus \mathbf{w}_2) \oplus \mathbf{z} \\ &= \mathbf{G}\mathbf{u} \oplus \mathbf{z} \end{aligned}$$

- Need $\max(R_1, R_2) \leq \log q - H(Z)$

Random linear codes are good for computation

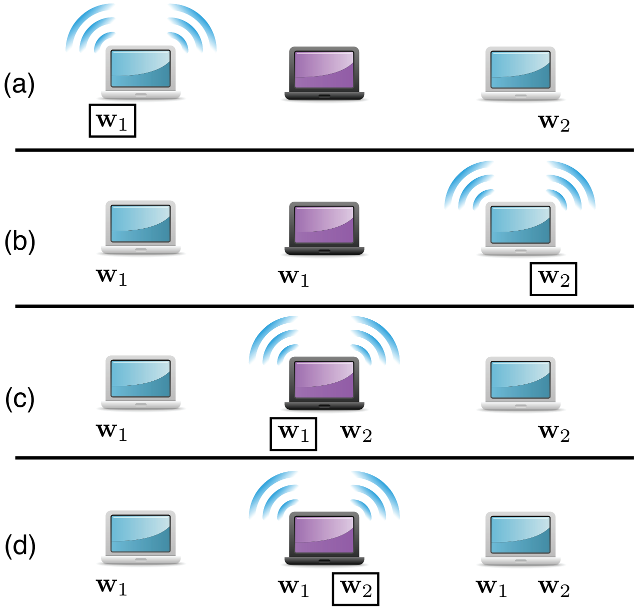




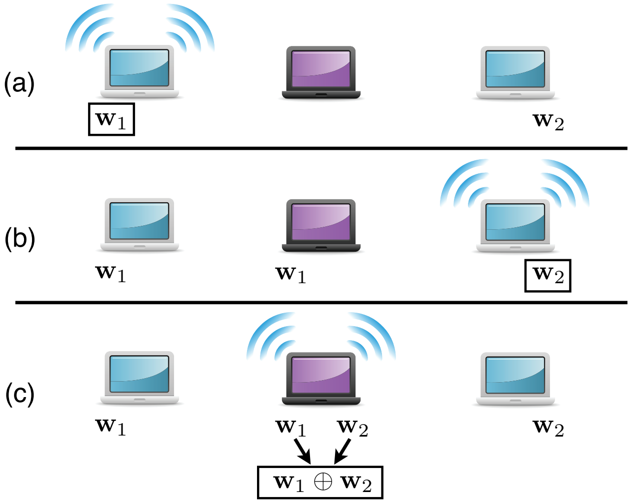
- **I.I.D. Random Coding**: $R_1 + R_2 \leq \log q - H(Z)$
- **Random Linear Coding**: $\max(R_1, R_2) \leq \log q - H(Z)$
- Linear codes double the sum rate *without any dependency*.
- Is this useful for *sending messages* (no computation)?

- Elegant example proposed by **Wu-Chou-Kung '04.**
- Closely related to butterfly network from **Ahlswede-Cai-Li-Yeung '00.**

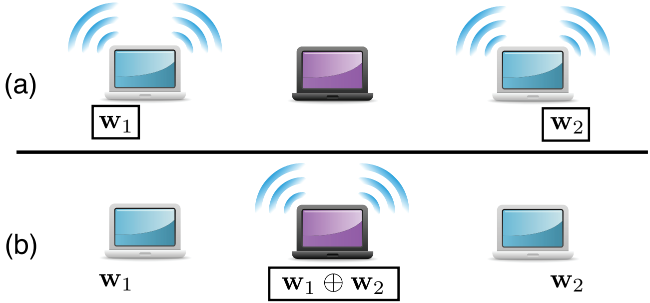
Two-Way Relay Channel – Time-Division

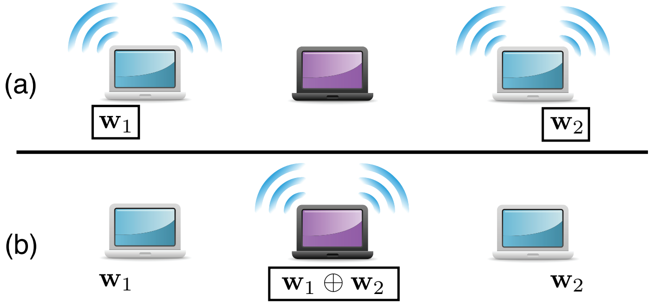


Two-Way Relay Channel – Network Coding



Two-Way Relay Channel – Physical-Layer Network Coding

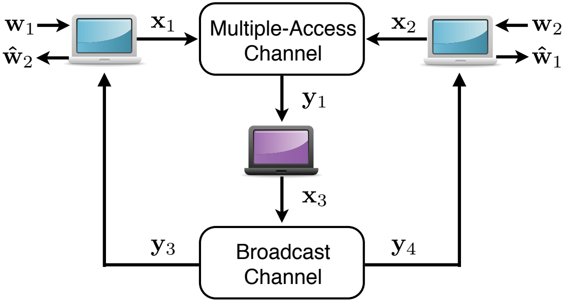




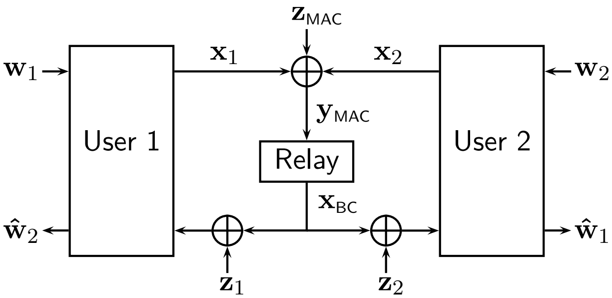
- Physical-layer network coding: exploiting the wireless medium for network coding. Independently and concurrently proposed by **Zhang-Liew-Lam '06, Popovski-Yomo '06, Nazer-Gastpar '06.**
- Sometimes referred to as Analog Network Coding **Katti-Gollakota-Katabi '08.**
- Some recent surveys **Liew-Zhang-Lu '11, Nazer-Gastpar '11.**

q-ary Two-Way Relay Channel

q-ary Two-Way Relay Channel



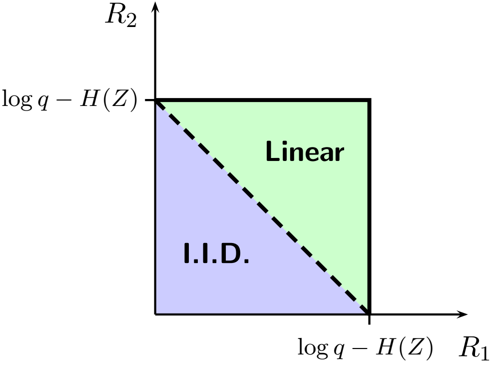
q-ary Two-Way Relay Channel



- i.i.d. noise sequences with entropy $H(Z)$.
- Rates R_1 and R_2.

- Upper Bound:
  $\max(R_1, R_2) \leq \log q - H(Z)$
- Random i.i.d.: Relay decodes $\mathbf{w}_1, \mathbf{w}_2$ and transmits $\mathbf{w}_1 \oplus \mathbf{w}_2$.
  $R_1 + R_2 \leq \log q - H(Z)$
- Random linear: Relay decodes and retransmits $\mathbf{w}_1 \oplus \mathbf{w}_2$
  $\max(R_1, R_2) \leq \log q - H(Z)$

q-ary Two-Way Relay Channel



- **I.I.D. Random Coding**: $R_1 + R_2 \leq \log q - H(Z)$
- **Random Linear Coding**: $\max(R_1, R_2) \leq \log q - H(Z)$
- Linear codes can double the sum rate *for exchanging messages.*

- *Observation:* For linear codes, the codeword statistics are uniform. This follows straightforwardly from the fact that the sum of any two codewords is again a codeword.
- *Question:* Can we retain some algebraic structure *and* have non-uniform codeword statistics?
- Idea: Nested Linear Codes (see, for instance, **Conway and Sloane '92, Forney '89, Zamir-Shamai-Erez '02** ...):

- Consider a linear code $\mathcal{C}_c$ of rate $1 - k/n$:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1,n-k} \\ g_{21} & g_{22} & \cdots & g_{2,n-k} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{n,n-k} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_{n-k} \end{bmatrix}$$

  with parity check matrix $\mathbf{H}_c$.

- For every binary sequence $\mathbf{u}$ of length k, define its coset as

$$\mathcal{C}_c(\mathbf{u}) = \{\mathbf{x} : \mathbf{H}_c\mathbf{x} = \mathbf{u}\}$$

- The *coset leader* is the one sequence in $\mathcal{C}_c(\mathbf{u})$ that has the *smallest* Hamming weight.

- For any sequence $\mathbf{x}$ we write $\mathbf{x} \bmod \mathcal{C}_c$ to denote the coset leader corresponding to $\mathbf{H}_c\mathbf{x}$.
- **Observation:** This satisfies all the usual properties of the modulo operation, such as

$$(\mathbf{x} \oplus \mathbf{y}) \bmod \mathcal{C}_c = (\mathbf{x} \bmod \mathcal{C}_c \oplus \mathbf{y} \bmod \mathcal{C}_c) \bmod \mathcal{C}_c$$

Theorem

There exists a binary linear code of rate $1 - k/n$ *such that all* 2^k *coset leaders satisfy* $w_{Hamming} \leq m$, *where*

$$k/n \geq H_b(m/n) - \epsilon$$

Note: Such a code is thus a good *covering* code.

Next step: *Decimate* coset leaders: retain only those belonging to a ("fine") code.

That way, we end up with a code of $2^{k-k'}$ codewords satisfying two properties:

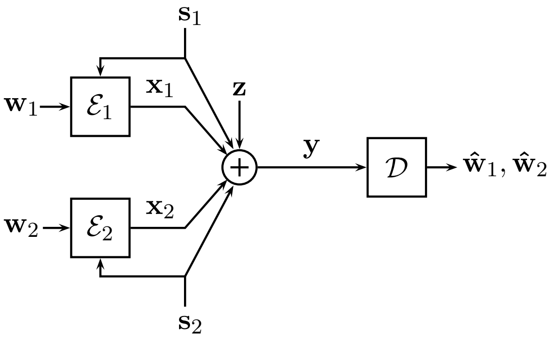
1. Noise protection just like the fine code
2. The sum of any two codewords, modulo "the coarse code," is again a codeword

On the BSC with crossover probability p, this code achieves a rate

$$R = H_b(m/n) - H_b(p).$$

Note that this is *not* the capacity of this channel.

Philosof-Zamir '09, Philosof-Zamir-Erez '09:



Without *input constraints*, the problem is trivial.

But now, consider

$$w_H(\mathbf{x}_1) \le m \quad \text{and} \quad w_H(\mathbf{x}_2) \le m.$$

Distributed Dirty Paper Coding

- Choose codewords $\mathbf{t}_1$ and $\mathbf{t}_2$. Transmit

$$\mathbf{x}_1 = (\mathbf{t}_1 \oplus \mathbf{s}_1) \text{ mod } \mathcal{C}_c \quad \text{and} \quad \mathbf{x}_2 = (\mathbf{t}_2 \oplus \mathbf{s}_2) \text{ mod } \mathcal{C}_c$$

- Choose coarse code to satisfy Hamming input constraints. Receive:

$$\mathbf{y} = [(\mathbf{x}_1 \oplus \mathbf{s}_1) \text{ mod } \mathcal{C}_c] \oplus [(\mathbf{x}_2 \oplus \mathbf{s}_2) \text{ mod } \mathcal{C}_c] \oplus \mathbf{s}_1 \oplus \mathbf{s}_2 \oplus \mathbf{z}$$

- The key step is the following pre-processing step at the decoder:

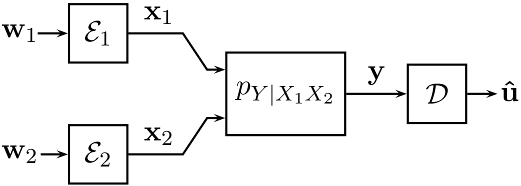
$$\begin{aligned} \mathbf{y} \text{ mod } \mathcal{C}_c &= (\mathbf{x}_1 \oplus \mathbf{s}_1 \oplus \mathbf{x}_2 \oplus \mathbf{s}_2 \oplus \mathbf{s}_1 \oplus \mathbf{s}_2 \oplus \mathbf{z}) \text{ mod } \mathcal{C}_c \\ &= (\mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \mathbf{z}) \text{ mod } \mathcal{C}_c \end{aligned}$$

- Last step: show that the noise is essentially unchanged by the modulo operation.

- Can show that this achieves the capacity (see **Philosof-Zamir-Erez '09**.)

Independent msgs $\mathbf{w}_1, \mathbf{w}_2$.

Want the sum $\mathbf{u} = \mathbf{w}_1 \oplus \mathbf{w}_2$ with vanishing prob. of error $\mathbb{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\} \to 0$



Achievable Strategy (Nazer-Gastpar '08)

Use the same linear code, $\max(R_1, R_2) \leq I(X_1 \oplus X_2; Y)$ (for binary, uniform inputs)

- General Functions: $U_i = f(W_{1i}, W_{2i})$
- Some achievable strategies, very hard in general (functional compression is a special case)
- For network communication, don't really care what functions in the middle, only care about msgs

I. **Discrete Alphabets**

II. **AWGN Channels**

III. **Network Applications**

Main References

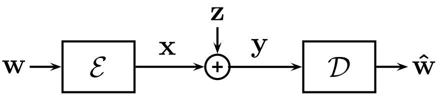
Nested lattice results in this section are almost entirely drawn from:

- U. Erez and R. Zamir, *Achieving* $\frac{1}{2}\log(1+\mathsf{SNR})$ *on the AWGN channel with lattice encoding and decoding*, IEEE Transactions on Information Theory, vol. 50, pp. 2293-2314, October 2004.
- U. Erez, S. Litsyn, and R. Zamir, *Lattices which are good for (almost) everything*, IEEE Transactions on Information Theory, vol. 51, pp. 3401-3416, October 2005.
- R. Zamir, *Lattices are everywhere*, in Proceedings of the 4th Annual Workshop on Information Theory and its Applications, La Jolla, CA, February 2009.

- Signal X is a scalar Gaussian r.v. with mean 0 and variance P.
- Noise Z is an independent scalar Gaussian r.v. with mean 0 and variance N.
- Estimate X from noisy observation $Y = X + Z$.
- Mean-squared error: $\mathbb{E}[(Y - X)^2] = \mathbb{E}[Z^2] = N$.
- Minimum mean-squared error (MMSE):

$$\begin{aligned}\mathbb{E}[(\alpha Y - X)^2] &= \mathbb{E}[(\alpha X + \alpha Z - X)^2] \\ &= \mathbb{E}[\alpha^2 Z^2 + (1-\alpha)^2 X^2] \qquad \text{Part of error due to } X \\ &= \alpha^2 N + (1-\alpha)^2 P\end{aligned}$$

- Optimal $\alpha = \dfrac{P}{N+P}$ yields $\mathbb{E}[(\alpha Y - X)^2] = \dfrac{PN}{N+P}$.



- Codewords must satisfy power constraint:

$$\|\mathbf{x}\|^2 \leq nP \ .$$

- i.i.d. Gaussian noise with variance N:

$$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, N\mathbf{I}) \ .$$

- **Shannon '48**: Channel capacity:

$$C = \frac{1}{2} \log \left(1 + \frac{P}{N}\right)$$

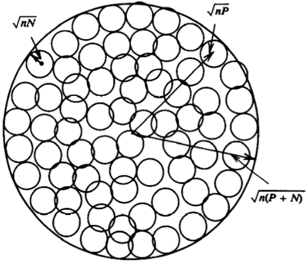


Figure 10.2. Sphere packing for the Gaussian channel.

(Cover and Thomas, *Elements of Information Theory*)

- In high dimensions, noise starts to look spherical.

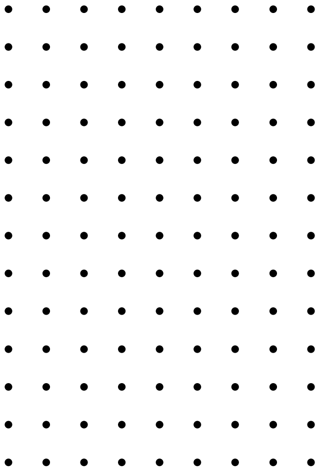
- A lattice Λ is a discrete subgroup of $\mathbb{R}^n$.
- Can write a lattice as a linear transformation of the integer vectors,

$$\Lambda = \mathbf{B}\mathbb{Z}^n \ ,$$

for some $\mathbf{B} \in \mathbb{R}^{n \times n}$.

Lattice Properties

- Closed under addition: $\lambda_1, \lambda_2 \in \Lambda \implies \lambda_1 + \lambda_2 \in \Lambda$.
- Symmetric: $\lambda \in \Lambda \implies -\lambda \in \Lambda$



$\mathbb{Z}^n$ is a simple lattice.

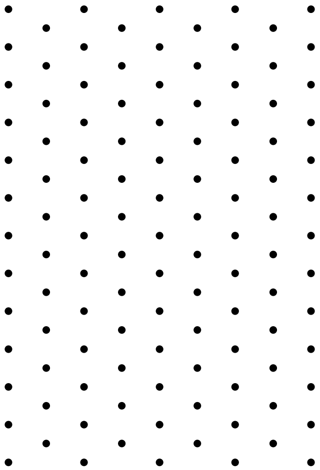
- A lattice Λ is a discrete subgroup of $\mathbb{R}^n$.
- Can write a lattice as a linear transformation of the integer vectors,

$$\Lambda = \mathbf{B}\mathbb{Z}^n ,$$

for some $\mathbf{B} \in \mathbb{R}^{n \times n}$.

Lattice Properties

- Closed under addition: $\lambda_1, \lambda_2 \in \Lambda \implies \lambda_1 + \lambda_2 \in \Lambda$.
- Symmetric: $\lambda \in \Lambda \implies -\lambda \in \Lambda$



$\mathbf{B}\mathbb{Z}^n$

Voronoi Regions

- Nearest neighbor quantizer:

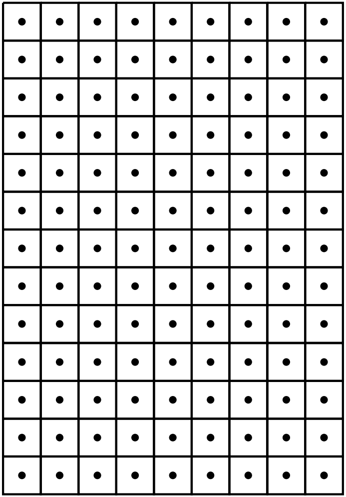
$$Q_\Lambda(\mathbf{x}) = \arg\min_{\lambda \in \Lambda} \|\mathbf{x} - \lambda\|_2$$

- The Voronoi region of a lattice point is the set of all points that quantize to that lattice point.

- Fundamental Voronoi region $\mathcal{V}$: points that quantize to the origin,

$$\mathcal{V} = \{\mathbf{x} : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}$$

- Each Voronoi region is just a shift of the fundamental Voronoi region $\mathcal{V}$



Voronoi Regions

- Nearest neighbor quantizer:

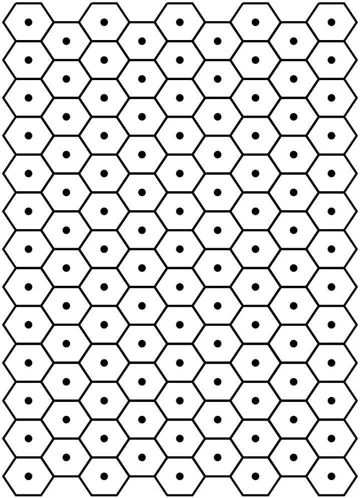
$$Q_\Lambda(\mathbf{x}) = \arg\min_{\lambda \in \Lambda} \|\mathbf{x} - \lambda\|_2$$

- The Voronoi region of a lattice point is the set of all points that quantize to that lattice point.

- Fundamental Voronoi region $\mathcal{V}$: points that quantize to the origin,

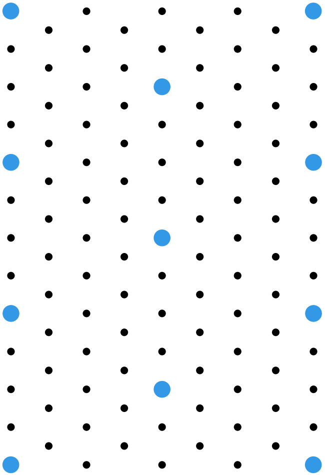
$$\mathcal{V} = \{\mathbf{x} : Q_\Lambda(\mathbf{x}) = \mathbf{0}\}$$

- Each Voronoi region is just a shift of the fundamental Voronoi region $\mathcal{V}$



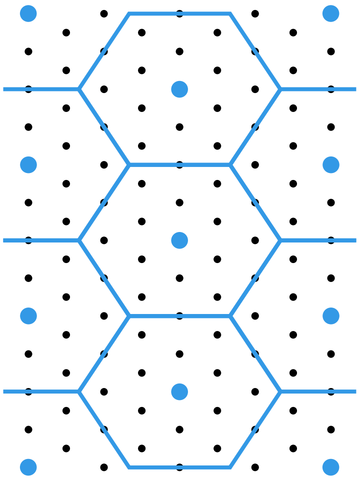
- Two lattices Λ and Λ_{FINE} are nested if $\Lambda \subset \Lambda_{\mathsf{FINE}}$
- **Nested Lattice Code:** All lattice points from Λ_{FINE} that fall in the fundamental Voronoi region $\mathcal{V}$ of Λ.
- $\mathcal{V}$ acts like a power constraint

$$\text{Rate} = \frac{1}{n} \log \left(\frac{\text{Vol}(\mathcal{V})}{\text{Vol}(\mathcal{V}_{\mathsf{FINE}})} \right)$$



- Two lattices Λ and Λ_{FINE} are nested if $\Lambda \subset \Lambda_{\mathsf{FINE}}$
- **Nested Lattice Code:** All lattice points from Λ_{FINE} that fall in the fundamental Voronoi region $\mathcal{V}$ of Λ.
- $\mathcal{V}$ acts like a power constraint

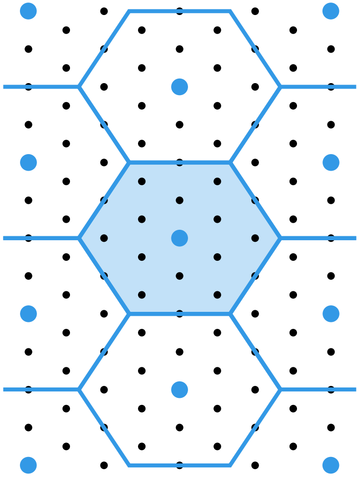
$$\text{Rate} = \frac{1}{n} \log \left(\frac{\text{Vol}(\mathcal{V})}{\text{Vol}(\mathcal{V}_{\mathsf{FINE}})} \right)$$



Nested Lattices

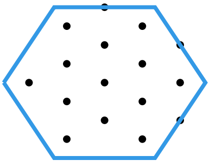
- Two lattices Λ and Λ_{FINE} are nested if $\Lambda \subset \Lambda_{\mathsf{FINE}}$
- **Nested Lattice Code:** All lattice points from Λ_{FINE} that fall in the fundamental Voronoi region $\mathcal{V}$ of Λ.
- $\mathcal{V}$ acts like a power constraint

$$\text{Rate} = \frac{1}{n} \log \left(\frac{\mathsf{Vol}(\mathcal{V})}{\mathsf{Vol}(\mathcal{V}_{\mathsf{FINE}})} \right)$$



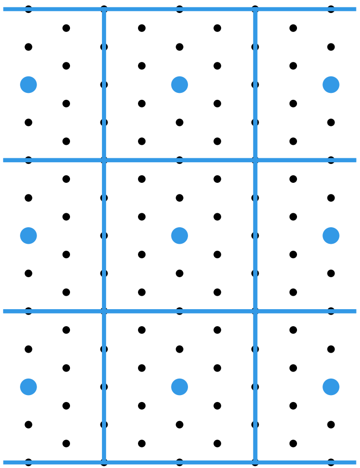
- Two lattices Λ and Λ_{FINE} are nested if $\Lambda \subset \Lambda_{\mathsf{FINE}}$
- **Nested Lattice Code:** All lattice points from Λ_{FINE} that fall in the fundamental Voronoi region $\mathcal{V}$ of Λ.
- $\mathcal{V}$ acts like a power constraint

$$\text{Rate} = \frac{1}{n} \log \left(\frac{\mathsf{Vol}(\mathcal{V})}{\mathsf{Vol}(\mathcal{V}_{\mathsf{FINE}})} \right)$$

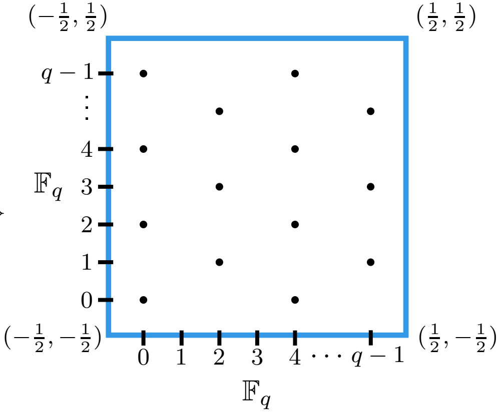


- Two lattices Λ and Λ_{FINE} are nested if $\Lambda \subset \Lambda_{\text{FINE}}$
- **Nested Lattice Code:** All lattice points from Λ_{FINE} that fall in the fundamental Voronoi region $\mathcal{V}$ of Λ.
- $\mathcal{V}$ acts like a power constraint

$$\text{Rate} = \frac{1}{n}\log\left(\frac{\text{Vol}(\mathcal{V})}{\text{Vol}(\mathcal{V}_{\text{FINE}})}\right)$$



- Choose an $n \times k$ generator matrix $\mathbf{G} \in \mathbb{F}_q^{n \times k}$ for q-ary code.
- Integers serve as coarse lattice, $\Lambda = \mathbb{Z}^n$.
- Map elements $\{0, 1, 2, \ldots, q-1\}$ to equally spaced points between $-1/2$ and $1/2$.
- Place codewords $\mathbf{x} = \mathbf{G}\mathbf{w}$ into the fundamental Voronoi region $\mathcal{V} = [-1/2, 1/2)^n$



Modulo Operation

- Modulo operation with respect to lattice Λ is just the residual quantization error,

$$[\mathbf{x}] \bmod \Lambda = \mathbf{x} - Q_\Lambda(\mathbf{x}) \ .$$

- Mimics the role of $\bmod q$ in q-ary alphabet.

- Distributive Law:

$$\Big[\mathbf{x}_1 + [\mathbf{x}_2] \bmod \Lambda\Big] \bmod \Lambda$$
$$= [\mathbf{x}_1 + \mathbf{x}_2] \bmod \Lambda$$

Modulo Operation

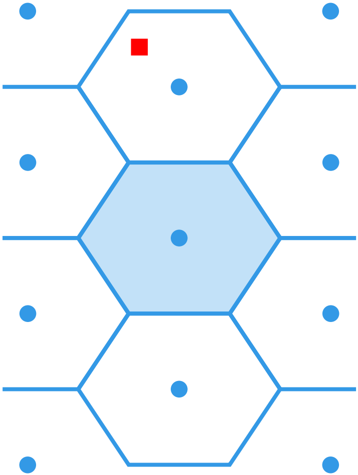
- Modulo operation with respect to lattice Λ is just the residual quantization error,

$$[\mathbf{x}] \bmod \Lambda = \mathbf{x} - Q_\Lambda(\mathbf{x}) \ .$$

- Mimics the role of $\bmod q$ in q-ary alphabet.

- Distributive Law:

$$\Big[\mathbf{x}_1 + [\mathbf{x}_2] \bmod \Lambda\Big] \bmod \Lambda = [\mathbf{x}_1 + \mathbf{x}_2] \bmod \Lambda$$



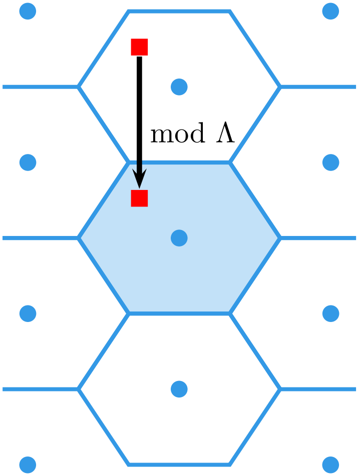
- Modulo operation with respect to lattice Λ is just the residual quantization error,

$$[\mathbf{x}] \bmod \Lambda = \mathbf{x} - Q_\Lambda(\mathbf{x}) \ .$$

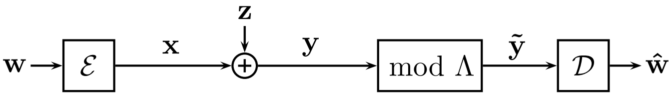
- Mimics the role of $\bmod q$ in q-ary alphabet.

- Distributive Law:

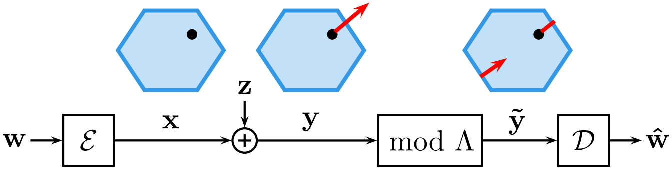
$$\Big[\mathbf{x}_1 + [\mathbf{x}_2] \bmod \Lambda\Big] \bmod \Lambda$$
$$= [\mathbf{x}_1 + \mathbf{x}_2] \bmod \Lambda$$



$\bmod \Lambda$ *AWGN Channel*

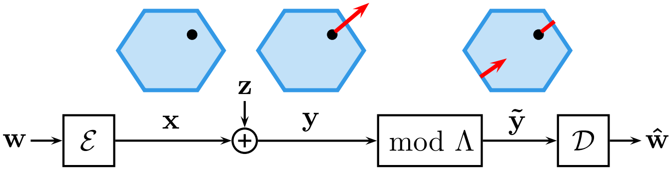


- Codebook lives on Voronoi region $\mathcal{V}$ of coarse lattice Λ.
- Take $\bmod \Lambda$ of received signal prior to decoding.
- What is the capacity of the $\bmod \Lambda$ channel?



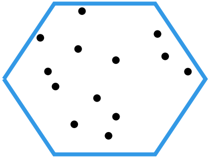
- Codebook lives on Voronoi region $\mathcal{V}$ of coarse lattice Λ.
- Take $\bmod \Lambda$ of received signal prior to decoding.
- What is the capacity of the $\bmod \Lambda$ channel?

$\bmod \Lambda$ *AWGN Channel*

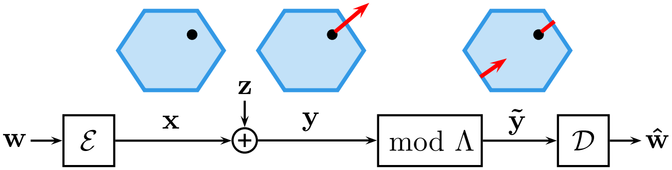


- Codebook lives on Voronoi region $\mathcal{V}$ of coarse lattice Λ.
- Take $\bmod \Lambda$ of received signal prior to decoding.
- What is the capacity of the $\bmod \Lambda$ channel?

Using random i.i.d. code drawn over $\mathcal{V}$: $C = \frac{1}{n} \max_{p(\mathbf{x})} I(\mathbf{x}; \tilde{\mathbf{y}})$



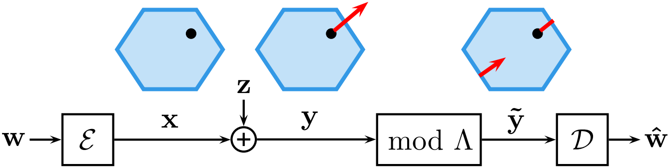
mod Λ *AWGN Channel Capacity*



$$nC = \max_{p(\mathbf{x})} I(\mathbf{x};\tilde{\mathbf{y}})$$

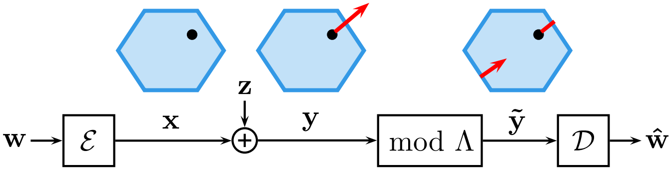
$$= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\tilde{\mathbf{y}}|\mathbf{x}) \Big)$$

mod Λ *AWGN Channel Capacity*



$$
\begin{aligned}
nC &= \max_{p(\mathbf{x})} I(\mathbf{x}; \tilde{\mathbf{y}}) \\
&= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\tilde{\mathbf{y}}|\mathbf{x}) \Big) \\
&= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h\Big([\mathbf{z}] \bmod \Lambda \Big) \Big)
\end{aligned}
$$

Distributive Law



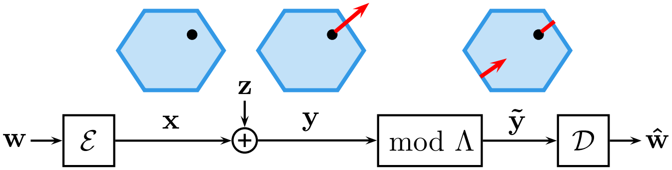
$$nC = \max_{p(\mathbf{x})} I(\mathbf{x}; \tilde{\mathbf{y}})$$

$$= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\tilde{\mathbf{y}}|\mathbf{x}) \Big)$$

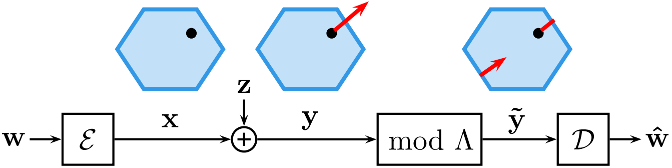
$$= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h\Big([\mathbf{z}] \bmod \Lambda \Big) \Big)$$ Distributive Law

$$\geq \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\mathbf{z}) \Big)$$ Point Symmetry of Voronoi Region

mod Λ *AWGN Channel Capacity*



$$
\begin{aligned}
nC &= \max_{p(\mathbf{x})} I(\mathbf{x}; \tilde{\mathbf{y}}) \\
&= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\tilde{\mathbf{y}}|\mathbf{x}) \Big) \\
&= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h\Big([\mathbf{z}] \bmod \Lambda \Big) \Big) \qquad \text{Distributive Law} \\
&\geq \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - h(\mathbf{z}) \Big) \qquad \text{Point Symmetry of Voronoi Region} \\
&= \max_{p(\mathbf{x})} \Big(h(\tilde{\mathbf{y}}) - \frac{n}{2} \log(2\pi e N) \Big) \qquad \text{Entropy of Gaussian Noise}
\end{aligned}
$$



- Channel output entropy is equal to the logarithm of the Voronoi region volume if it is uniform over $\mathcal{V}$:

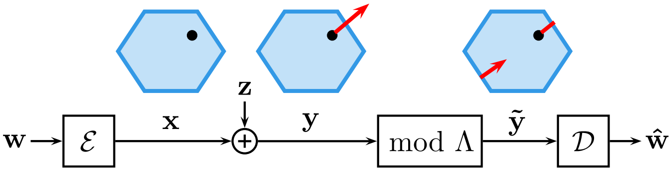
$$h(\tilde{\mathbf{y}}) = \log(\mathsf{Vol}(\mathcal{V})) \quad \text{if } \tilde{\mathbf{y}} \sim \mathsf{Unif}(\mathcal{V})$$

- $\tilde{\mathbf{y}} = [\mathbf{x} + \mathbf{z}] \bmod \Lambda$ is uniform over $\mathcal{V}$ if $\mathbf{x}$ is uniform over $\mathcal{V}$.

- Random i.i.d. coding over the Voronoi region $\mathcal{V}$ can achieve:

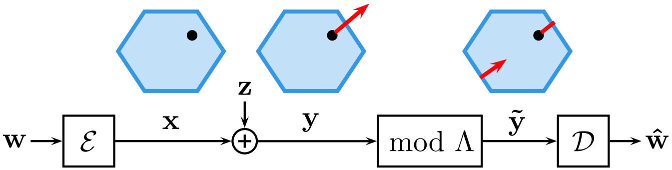
$$R = \frac{1}{n}\log(\mathsf{Vol}(\mathcal{V})) - \frac{1}{2}\log(2\pi e N)$$

Power Constraints and Second Moments



- Must scale lattice Λ so that the uniform distribution over the Voronoi region $\mathcal{V}$ meets the power constraint P.
- Set second moment $\sigma_\Lambda^2 = \dfrac{1}{n\mathsf{Vol}(\mathcal{V})}\displaystyle\int_{\mathcal{V}} \|\mathbf{x}\|^2 d\mathbf{x}$ equal to P.

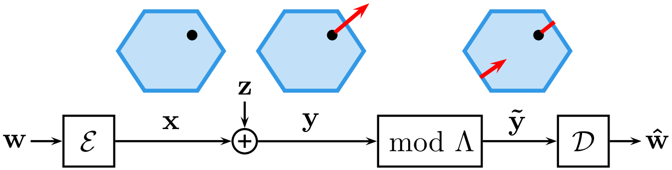
Power Constraints and Second Moments



- Must scale lattice Λ so that the uniform distribution over the Voronoi region $\mathcal{V}$ meets the power constraint P.
- Set second moment $\sigma_\Lambda^2 = \dfrac{1}{n\mathsf{Vol}(\mathcal{V})} \displaystyle\int_{\mathcal{V}} \|\mathbf{x}\|^2 d\mathbf{x}$ equal to P.

Normalized Second Moment: $$G(\Lambda) = \frac{\sigma_\Lambda^2}{(\mathsf{Vol}(\mathcal{V}))^{2/n}}$$

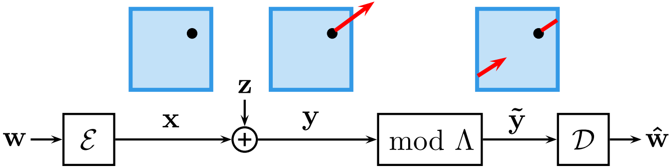
$$\implies \frac{1}{n}\log(\mathsf{Vol}(\mathcal{V})) = \frac{1}{2}\log\left(\frac{\sigma_\Lambda^2}{G(\Lambda)}\right) = \frac{1}{2}\log\left(\frac{P}{G(\Lambda)}\right)$$



- Random i.i.d. coding over the Voronoi region $\mathcal{V}$ can achieve:

$$\begin{aligned} C &\geq \frac{1}{n}\log(\mathsf{Vol}(\mathcal{V})) - \frac{1}{2}\log(2\pi e N) \\ &= \frac{1}{2}\log\left(\frac{P}{G(\Lambda)}\right) - \frac{1}{2}\log(2\pi e N) \\ &= \frac{1}{2}\log\left(\frac{P}{N}\right) - \frac{1}{2}\log(2\pi e G(\Lambda)) \end{aligned}$$

What is $G(\Lambda)$?

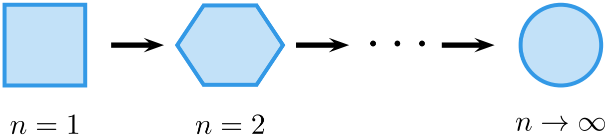


- The normalized second moment $G(\Lambda)$ is a dimensionless quantity that captures the shaping gain.
- Integer lattice is not so bad, $G(\mathbb{Z}^n) = 1/12$.
- Capacity under $\mod \mathbb{Z}^n$ is at least

$$C \geq \frac{1}{2}\log\left(\frac{P}{N}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{12}\right)$$
$$\approx \frac{1}{2}\log\left(\frac{P}{N}\right) - 0.255$$

Theorem (Zamir-Feder-Poltyrev '94)

There exists a sequence of lattices $\Lambda^{(n)}$ *such that* $\lim_{n\to\infty} G(\Lambda^{(n)}) = \frac{1}{2\pi e}$.



- Best possible normalized second moment is that of a sphere.
- Using a sequence $\Lambda^{(n)}$ with an asymptotically good $G(\Lambda^{(N)})$ allows to approach

$$R = \frac{1}{2}\log\left(\frac{P}{N}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{2\pi e}\right)$$
$$= \frac{1}{2}\log\left(\frac{P}{N}\right)$$

Asymptotically Good $G(\Lambda)$

- Can actually get this with a linear code tiled over $\mathbb{Z}^n$ (see, for instance, **Erez-Litsyn-Zamir '05**.)
- Many works looking at this from different perspectives.
- We will just assume existence.

Recall the two key properties of random linear codes $\mathbf{G}$ from earlier:

Codeword Properties

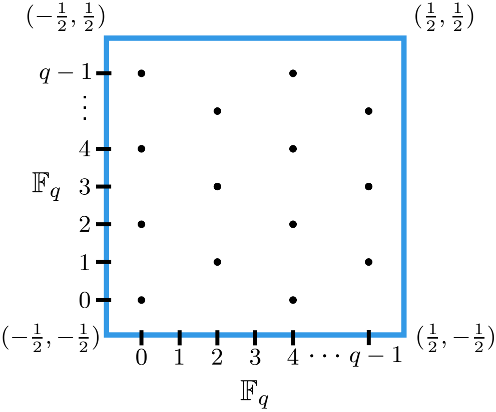
1. Marginally uniform over $\mathbb{F}_q^n$. For a given message $\mathbf{w} \neq \mathbf{0}$, the codeword $\mathbf{x} = \mathbf{G}\mathbf{w}$ looks like an i.i.d. uniform sequence.

$$\mathbb{P}\{\mathbf{x} = \mathsf{x}\} = \frac{1}{q^n} \quad \text{for all } \mathsf{x} \in \mathbb{F}_q^n$$

2. Pairwise independent. For $\mathbf{w}_1, \mathbf{w}_2 \neq \mathbf{0}$, $\mathbf{w}_1 \neq \mathbf{w}_2$, codewords $\mathbf{x}_1, \mathbf{x}_2$ are independent.

$$\mathbb{P}\{\mathbf{x_1} = \mathsf{x}_1, \mathbf{x_2} = \mathsf{x}_2\} = \frac{1}{q^{2n}} = \mathbb{P}\{\mathbf{x}_1 = \mathsf{x}_1\}\mathbb{P}\{\mathbf{x}_2 = \mathsf{x}_2\}$$

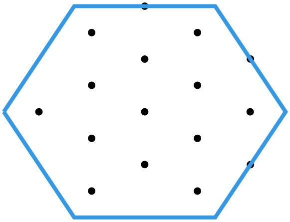
- Instead of an "inner" random codes, we can use a q-ary linear code.
- This is exactly a nested lattice.
- Each codeword has a uniform marginal distribution over the grid.
- Rate loss due to finite constellation which goes to 0 as $q \to \infty$.
- Codewords are pairwise independent so we can apply the union bound.



$$\mathbf{x} = [\gamma \mathbf{G} \mathbf{w}] \mod \mathbb{Z}^n$$

Linear Codes for $\bmod \Lambda$ *Channels*

- General coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$.
- First, apply generator matrix for linear code $\mathbf{G}\mathbf{w}$. Then scale down by γ and tile over $\mathbb{Z}^n$.
- Multiply by $\mathbf{B}$ and apply $\bmod \Lambda$ to get codebook.
- As q gets large, each codeword's marginal distribution looks uniform over $\mathcal{V}$.
- Codewords are pairwise independent so we can apply the union bound.



$$\mathbf{x} = [\mathbf{B}\gamma\mathbf{G}\mathbf{w}] \bmod \Lambda$$

MMSE Scaling

- **Erez-Zamir '04:** Prior to taking $\bmod \Lambda$, scale by α.

$$\begin{aligned}\tilde{\mathbf{y}} &= [\alpha\mathbf{y}] \bmod \Lambda \\ &= [\alpha\mathbf{x} + \alpha\mathbf{z}] \bmod \Lambda \\ &= [\mathbf{x} + \underbrace{\alpha\mathbf{z} - (1-\alpha)\mathbf{x}}_{\text{Effective Noise}}] \bmod \Lambda\end{aligned}$$

- For now, ignore that the effective noise is not independent of the codeword. Effective noise variance $N_{\text{EFFEC}} = \alpha^2 N + (1-\alpha)^2 P$.
- Optimal choice of α is the MMSE coefficient $\alpha_{\text{MMSE}} = \dfrac{P}{N+P}$.

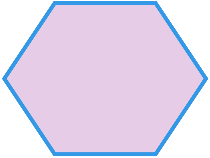
$$N_{\text{EFFEC}} = \alpha_{\text{MMSE}}^2 N + (1-\alpha_{\text{MMSE}})^2 P = \frac{PN}{N+P}$$

$$C = \frac{1}{2}\log\left(\frac{P}{N_{\text{EFFEC}}}\right) = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

- Now the noise is dependent on the codeword.
- Dithering can solve this problem (just as in the discrete case).
- Map message $\mathbf{w}$ to a lattice codeword $\mathbf{t}$.
- Generate a random dither vector $\mathbf{d}$ uniformly over $\mathcal{V}$.
- Transmitter sends a dithered codeword:

$$\mathbf{x} = [\mathbf{t} + \mathbf{d}] \bmod \Lambda$$

- $\mathbf{x}$ is now independent of the codeword $\mathbf{t}$.

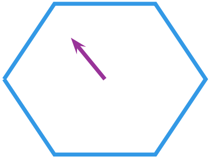


Dithering

- Now the noise is dependent on the codeword.
- Dithering can solve this problem (just as in the discrete case).
- Map message $\mathbf{w}$ to a lattice codeword $\mathbf{t}$.
- Generate a random dither vector $\mathbf{d}$ uniformly over $\mathcal{V}$.
- Transmitter sends a dithered codeword:

$$\mathbf{x} = [\mathbf{t} + \mathbf{d}] \bmod \Lambda$$

- $\mathbf{x}$ is now independent of the codeword $\mathbf{t}$.



Decoding – Remove Dither First

- Transmitter sends dithered codeword $\mathbf{x} = [\mathbf{t} + \mathbf{d}] \bmod \Lambda$.
- After scaling the channel output $\mathbf{y}$ by α, the decoder subtracts the dither $\mathbf{d}$.

$$
\begin{aligned}
\tilde{\mathbf{y}} &= [\alpha\mathbf{y} - \mathbf{d}] \bmod \Lambda \\
&= [\alpha\mathbf{x} + \alpha\mathbf{z} - \mathbf{d}] \bmod \Lambda \\
&= [\mathbf{x} - \mathbf{d} + \alpha\mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \\
&= \Big[[\mathbf{t} + \mathbf{d}] \bmod \Lambda - \mathbf{d} + \alpha\mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda \\
&= [\mathbf{t} + \alpha\mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \qquad \text{Distributive Law}
\end{aligned}
$$

- Effective noise is now independent from the codeword $\mathbf{t}$.
- By the probabilistic method, (at least) one good fixed dither exists. No common randomness necessary.

- Linear code embedded in the integer lattice:

$$R = \frac{1}{2}\log\left(\frac{P}{N}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{12}\right)$$

- Linear code embedded in the integer lattice, MMSE scaling:

$$R = \frac{1}{2}\log\left(1+\frac{P}{N}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{12}\right)$$

- Linear code embedded in a good shaping lattice, MMSE scaling:

$$R = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

Theorem (Erez-Zamir '04)

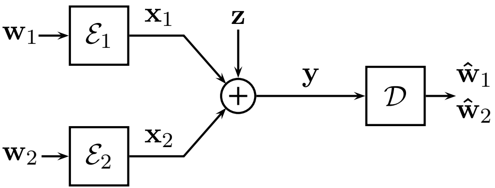
Nested lattice codes can achieve the AWGN capacity.

Rate Region

$$R_1 < \frac{1}{2}\log\left(1 + \frac{P_1}{N}\right)$$

$$R_2 < \frac{1}{2}\log\left(1 + \frac{P_2}{N}\right)$$

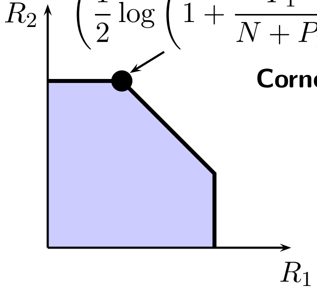
$$R_1 + R_2 < \frac{1}{2}\log\left(1 + \frac{P_1 + P_2}{N}\right)$$



Power constraints P_1, P_2. Noise variance N.

Successive Cancellation

$$\left(\frac{1}{2}\log\left(1 + \frac{P_1}{N + P_2}\right),\ \frac{1}{2}\log\left(1 + \frac{P_2}{N}\right)\right)$$



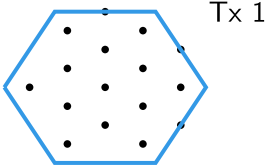
Corner Point

1. Decode $\mathbf{x}_1$, treating $\mathbf{x}_2$ as noise.
2. Subtract $\mathbf{x}_1$ from $\mathbf{y}$.
3. Decode $\mathbf{x}_2$.

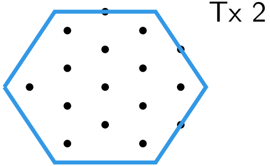
Codebook Generation

Select a nested lattice code:

- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.



Encoding



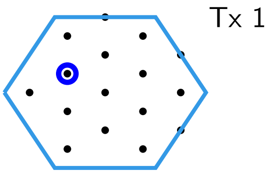
Codebook Generation

Select a nested lattice code:

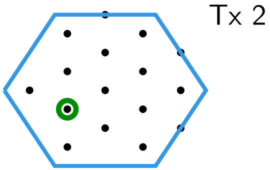
- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.

Encoding

- Map messages $\mathbf{w}_1, \mathbf{w}_2$ to lattice points $\mathbf{t}_1, \mathbf{t}_2$.



$$\mathbf{t}_1 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_1}] \bmod \Lambda$$



$$\mathbf{t}_2 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_2}] \bmod \Lambda$$

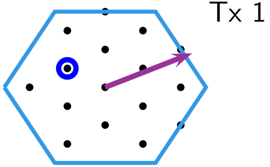
Codebook Generation

Select a nested lattice code:

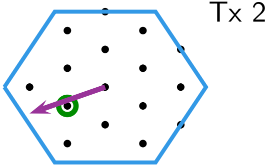
- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.

Encoding

- Map messages $\mathbf{w}_1, \mathbf{w}_2$ to lattice points $\mathbf{t}_1, \mathbf{t}_2$.
- Choose independent dithers $\mathbf{d}_1, \mathbf{d}_2$ uniformly over Voronoi region $\mathcal{V}$.



$\mathbf{t}_1 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_1}] \bmod \Lambda$



$\mathbf{t}_2 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_2}] \bmod \Lambda$

Lattice Achievability "Recipe" – Multiple-Access Corner Point

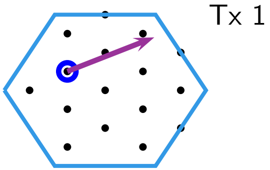
Codebook Generation

Select a nested lattice code:

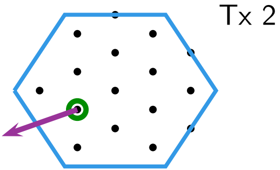
- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.

Encoding

- Map messages $\mathbf{w}_1, \mathbf{w}_2$ to lattice points $\mathbf{t}_1, \mathbf{t}_2$.
- Choose independent dithers $\mathbf{d}_1, \mathbf{d}_2$ uniformly over Voronoi region $\mathcal{V}$.
- Add dithers to lattice points and take $\bmod \Lambda$ to get transmitted signals $\mathbf{x}_1, \mathbf{x}_2$.



$$\mathbf{t}_1 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_1}] \bmod \Lambda$$
$$\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda$$



$$\mathbf{t}_2 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_2}] \bmod \Lambda$$
$$\mathbf{x}_2 = [\mathbf{t}_1 + \mathbf{d}_2] \bmod \Lambda$$

Lattice Achievability "Recipe" – Multiple-Access Corner Point

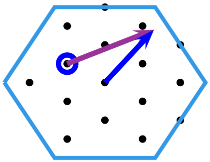
Codebook Generation

Select a nested lattice code:

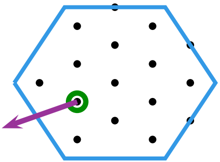
- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.

Encoding

- Map messages $\mathbf{w}_1, \mathbf{w}_2$ to lattice points $\mathbf{t}_1, \mathbf{t}_2$.
- Choose independent dithers $\mathbf{d}_1, \mathbf{d}_2$ uniformly over Voronoi region $\mathcal{V}$.
- Add dithers to lattice points and take $\bmod \Lambda$ to get transmitted signals $\mathbf{x}_1, \mathbf{x}_2$.



$$\mathbf{t}_1 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_1}] \bmod \Lambda$$
$$\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda$$



$$\mathbf{t}_2 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_2}] \bmod \Lambda$$
$$\mathbf{x}_2 = [\mathbf{t}_1 + \mathbf{d}_2] \bmod \Lambda$$

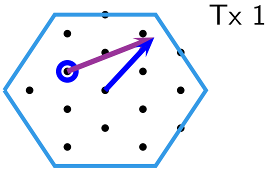
Codebook Generation

Select a nested lattice code:

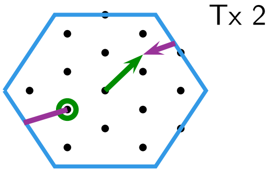
- Coarse lattice $\Lambda = \mathbf{B}\mathbb{Z}^n$ for shaping.
- Fine lattice from q-ary linear code $\mathbf{G}$ for coding.

Encoding

- Map messages $\mathbf{w}_1, \mathbf{w}_2$ to lattice points $\mathbf{t}_1, \mathbf{t}_2$.
- Choose independent dithers $\mathbf{d}_1, \mathbf{d}_2$ uniformly over Voronoi region $\mathcal{V}$.
- Add dithers to lattice points and take $\bmod \Lambda$ to get transmitted signals $\mathbf{x}_1, \mathbf{x}_2$.



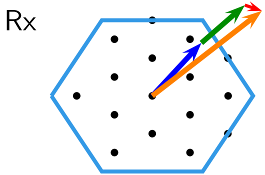
$$\mathbf{t}_1 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_1}] \bmod \Lambda$$
$$\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda$$



$$\mathbf{t}_2 = [\mathbf{B}\gamma\mathbf{G}\mathbf{w_2}] \bmod \Lambda$$
$$\mathbf{x}_2 = [\mathbf{t}_1 + \mathbf{d}_2] \bmod \Lambda$$

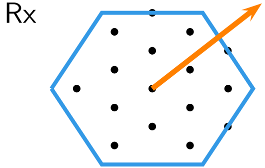
Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

Decoding



Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

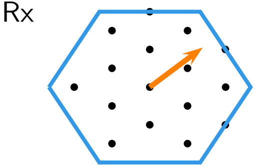
Decoding



Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

Decoding

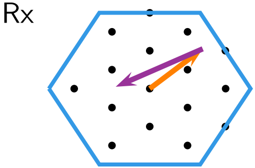
- Scale by α.



Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

Decoding

- Scale by α.
- Subtract dither $\mathbf{d}_1$.

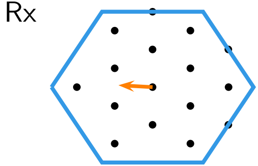


Lattice Achievability "Recipe" – Multiple-Access Corner Point

Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

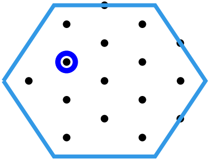
Decoding

- Scale by α.
- Subtract dither $\mathbf{d}_1$.
- Take $\bmod \Lambda$.



Receiver observes $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}$.

Decoding



- Scale by α.
- Subtract dither $\mathbf{d}_1$.
- Take $\bmod \Lambda$.
- Decode to nearest codeword.

$$
\begin{aligned}
&[\alpha \mathbf{y} - \mathbf{d}_1] \bmod \Lambda \\
&= [\alpha(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}) - \mathbf{d}_1] \bmod \Lambda \\
&= [\mathbf{x}_1 - \mathbf{d}_1 + \alpha \mathbf{z} + \alpha \mathbf{x}_2 - (1-\alpha)\mathbf{x}_1] \bmod \Lambda \\
&= \Big[[\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda - \mathbf{d}_1 + \alpha \mathbf{z} + \alpha \mathbf{x}_2 - (1-\alpha)\mathbf{x}_1 \Big] \bmod \Lambda \\
&= [\mathbf{t}_1 + \underbrace{\alpha \mathbf{z} + \alpha \mathbf{x}_2 - (1-\alpha)\mathbf{x}_1}_{\text{Effective Noise}}]
\end{aligned}
$$

Lattice Achievability "Recipe" – Multiple-Access Corner Point

- Effective noise after scaling is $N_{\mathsf{EFFEC}} = \alpha^2(N + P_2) + (1-\alpha)^2 P_1$.
- Minimized by setting α to be the MMSE coefficient:

$$\alpha_{\mathsf{MMSE}} = \frac{P_1}{N + P_1 + P_2}$$

- Plugging in, we get

$$N_{\mathsf{EFFEC}} = \frac{(N + P_2)P_1}{N + P_1 + P_2}$$

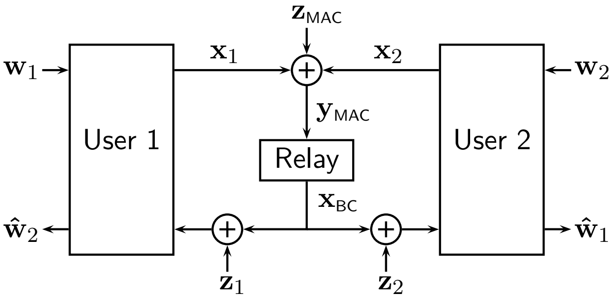
- Resulting rate is

$$R = \frac{1}{2}\log\left(\frac{P_1}{N_{\mathsf{EFFEC}}}\right) = \frac{1}{2}\log\left(1 + \frac{P_1}{N + P_2}\right)$$

- To obtain different rates for $\mathbf{x}_1$ and $\mathbf{x}_2$, use nested linear codes $\mathbf{G}_1$ and $\mathbf{G}_2$ inside Voronoi region $\mathcal{V}$.

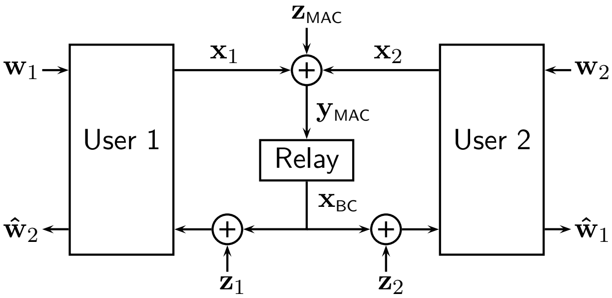
AWGN Two-Way Relay Channel – Symmetric Rates

AWGN Two-Way Relay Channel – Symmetric Rates



- Equal power constraints P.
- Equal noise variances N.
- Equal rates R.

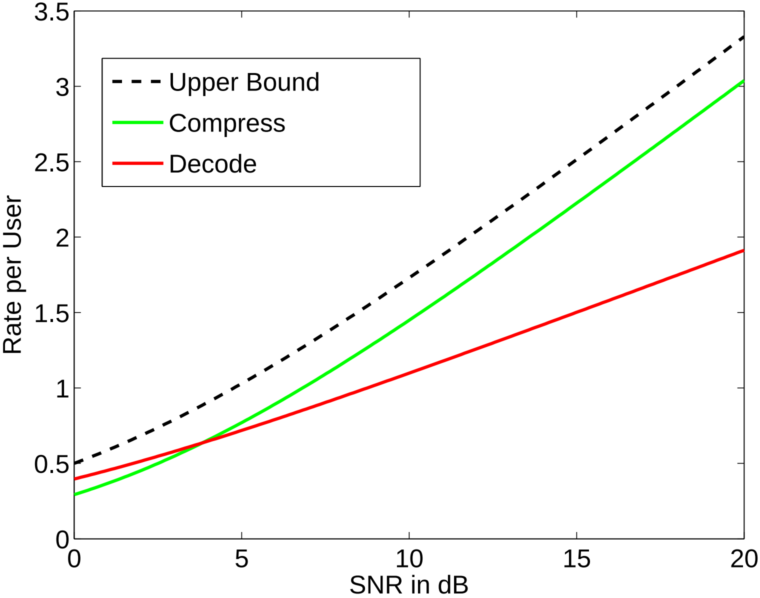
AWGN Two-Way Relay Channel – Symmetric Rates



- Equal power constraints P.
- Equal noise variances N.
- Equal rates R.

- Upper Bound:
  $R \leq \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$
- Decode-and-Forward: Relay decodes $\mathbf{w}_1, \mathbf{w}_2$ and transmits $\mathbf{w}_1 \oplus \mathbf{w}_2$.
  $R = \frac{1}{4}\log\left(1 + \frac{2P}{N}\right)$
- Compress-and-Forward: Relay transmits quantized $\mathbf{y}$.
  $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\frac{P}{3P+N}\right)$

AWGN Two-Way Relay Channel – Symmetric Rates



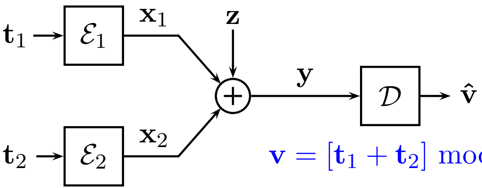
Decoding the Sum of Lattice Codewords

Encoders use the same nested lattice codebook.

Transmit lattice codewords:

$$\mathbf{x}_1 = \mathbf{t}_1$$

$$\mathbf{x}_2 = \mathbf{t}_2$$



$$\mathbf{v} = [\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda$$

Decoder recovers modulo sum.

$$\begin{aligned}
&[\mathbf{y}] \bmod \Lambda \\
&= [\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}] \bmod \Lambda \\
&= [\mathbf{t}_1 + \mathbf{t}_2 + \mathbf{z}] \bmod \Lambda \\
&= \Big[[\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda + \mathbf{z} \Big] \bmod \Lambda \qquad \text{Distributive Law} \\
&= [\mathbf{v} + \mathbf{z}] \bmod \Lambda
\end{aligned}$$

$$R = \frac{1}{2} \log \left(\frac{P}{N} \right)$$

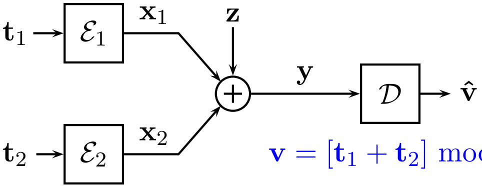
Decoding the Sum of Lattice Codewords – MMSE Scaling

Encoders use the same nested lattice codebook.

Transmit dithered codewords:

$$\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda$$

$$\mathbf{x}_2 = [\mathbf{t}_2 + \mathbf{d}_2] \bmod \Lambda$$



$\mathbf{v} = [\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda$

Decoder scales by α, removes dithers, recovers modulo sum.

$$\begin{aligned}
&[\alpha\mathbf{y} - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= [\alpha(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{z}) - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= [\mathbf{x}_1 + \mathbf{x}_2 - (1-\alpha)(\mathbf{x}_1 + \mathbf{x}_2) + \alpha\mathbf{z} - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= \Big[[\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda - (1-\alpha)(\mathbf{x}_1 + \mathbf{x}_2) + \alpha\mathbf{z} \Big] \bmod \Lambda \\
&= [\mathbf{v} - (1-\alpha)(\mathbf{x}_1 + \mathbf{x}_2) + \alpha\mathbf{z}] \bmod \Lambda
\end{aligned}$$

Effective Noise $\quad N_{\text{EFFEC}} = (1-\alpha)^2 2P + \alpha^2 N$

- Effective noise after scaling is $N_{\mathsf{EFFEC}} = (1-\alpha)^2 2P + \alpha^2 N$.
- Minimized by setting α to be the MMSE coefficient:

$$\alpha_{\mathsf{MMSE}} = \frac{2P}{N + 2P}$$

- Plugging in, we get

$$N_{\mathsf{EFFEC}} = \frac{2NP}{N + 2P}$$

- Resulting rate is

$$R = \frac{1}{2} \log \left(\frac{P}{N_{\mathsf{EFFEC}}} \right) = \frac{1}{2} \log \left(\frac{1}{2} + \frac{P}{N} \right)$$

- Getting the full “one plus” term is an open challenge. Does not seem possible with nested lattices.

- Map messages to lattice points

$$\mathbf{t}_1 = \phi(\mathbf{w}_1) = [\mathbf{B}\gamma\mathbf{G}\mathbf{w}_1] \bmod \Lambda$$
$$\mathbf{t}_2 = \phi(\mathbf{w}_2) = [\mathbf{B}\gamma\mathbf{G}\mathbf{w}_2] \bmod \Lambda$$

- Mapping between finite field messages and lattice codewords preserves linearity:

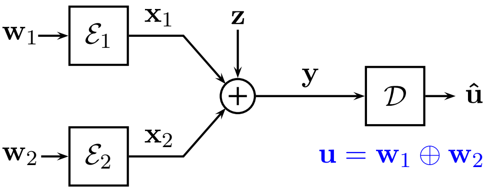
$$\phi^{-1}\Big([\mathbf{t}_1 + \mathbf{t}_2] \mod \Lambda\Big) = \mathbf{w}_1 \oplus \mathbf{w}_2$$

- This means that after decoding a $\bmod \Lambda$ equation of lattice points we can immediately recover the finite field equation of the messages. See **Nazer-Gastpar '11** for more details.

Finite Field Computation over a Gaussian MAC

Map messages to lattice points:

$$\mathbf{t}_1 = \phi(\mathbf{w}_1)$$
$$\mathbf{t}_2 = \phi(\mathbf{w}_2)$$



Transmit dithered codewords:

$$\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda$$
$$\mathbf{x}_2 = [\mathbf{t}_2 + \mathbf{d}_2] \bmod \Lambda$$

- If decoder can recover $[\mathbf{t}_1 + \mathbf{t}_2] \bmod \Lambda$, it also can get the sum of the messages

$$\mathbf{w}_1 \oplus \mathbf{w}_2 = \phi^{-1}\Big([\mathbf{t}_1 + \mathbf{t}_2] \mod \Lambda\Big) .$$

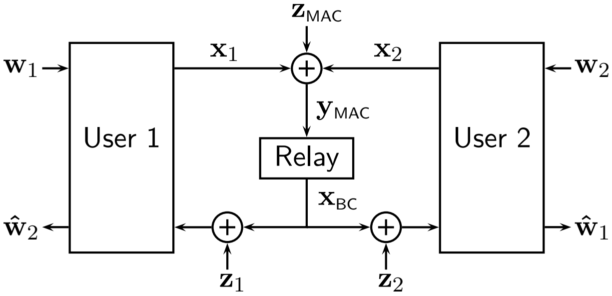
- Achievable rate $R = \frac{1}{2} \log \left(\frac{1}{2} + \frac{P}{N} \right)$.

AWGN Two-Way Relay Channel – Symmetric Rates

- Equal power constraints P.
- Equal noise variances N.
- Equal rates R.

- Upper Bound:
  $R \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right)$
- Compute-and-Forward: Relay decodes $\mathbf{w}_1 \oplus \mathbf{w}_2$ and retransmits.
  $R = \frac{1}{2} \log \left(\frac{1}{2} + \frac{P}{N} \right)$
- **Wilson-Narayanan-Pfister-Sprintson '10:** Applies nested lattice codes to the two-way relay channel.

AWGN Two-Way Relay Channel – Symmetric Rates



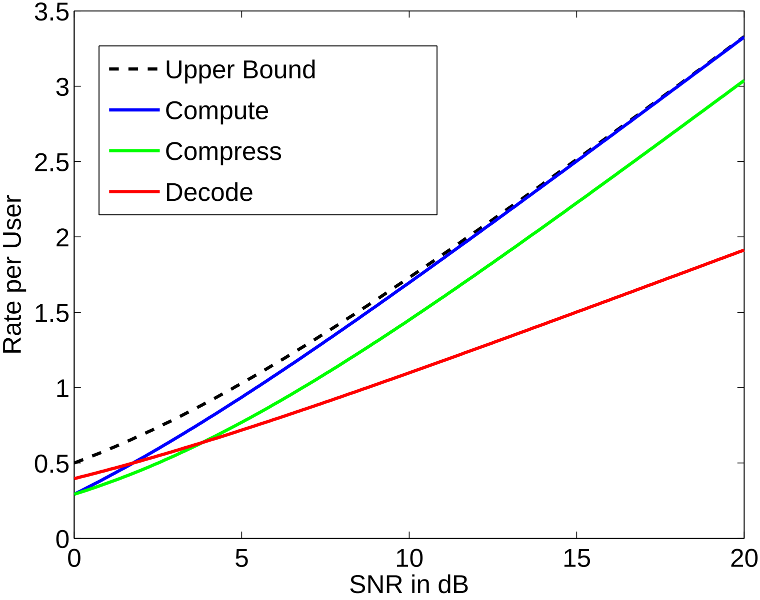
- Equal power constraints P.
- Equal noise variances N.
- Equal rates R.

- Upper Bound:
  $R \leq \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$

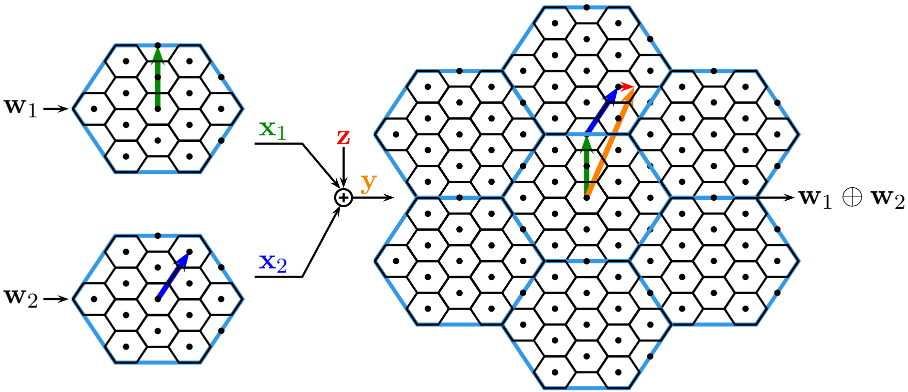
- Compute-and-Forward: Relay decodes $\mathbf{w}_1 \oplus \mathbf{w}_2$ and retransmits.
  $R = \frac{1}{2}\log\left(\frac{1}{2} + \frac{P}{N}\right)$

- **Wilson-Narayanan-Pfister-Sprintson '10:** Applies nested lattice codes to the two-way relay channel.

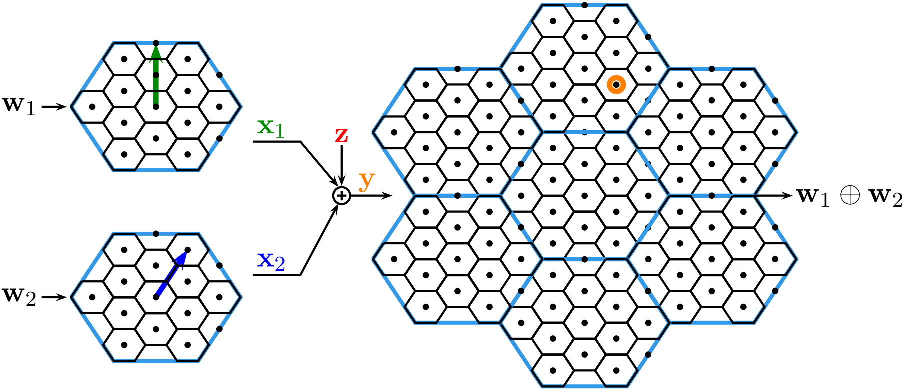
AWGN Two-Way Relay Channel – Symmetric Rates



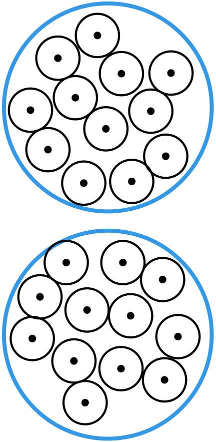
Compute-and-Forward Illustration



Compute-and-Forward Illustration



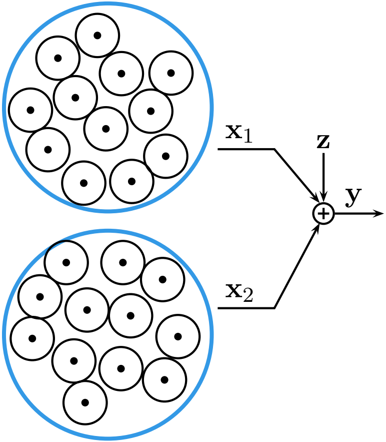
Random i.i.d. codes are not good for computation



2^{nR} codewords each.

2^{n2R} possible sums of codewords.

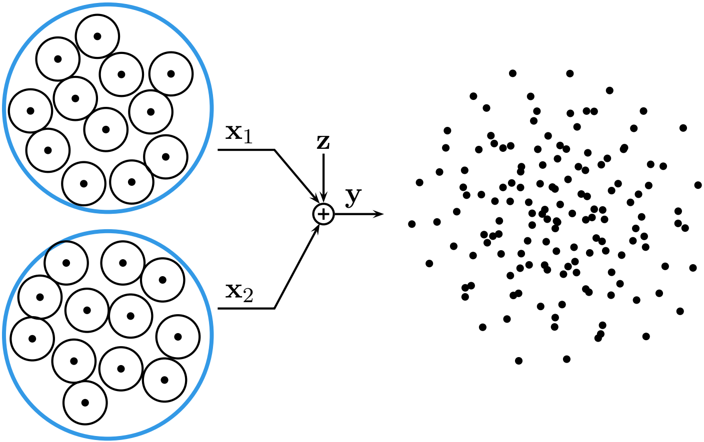
Random i.i.d. codes are not good for computation



2^{nR} codewords each.

2^{n2R} possible sums of codewords.

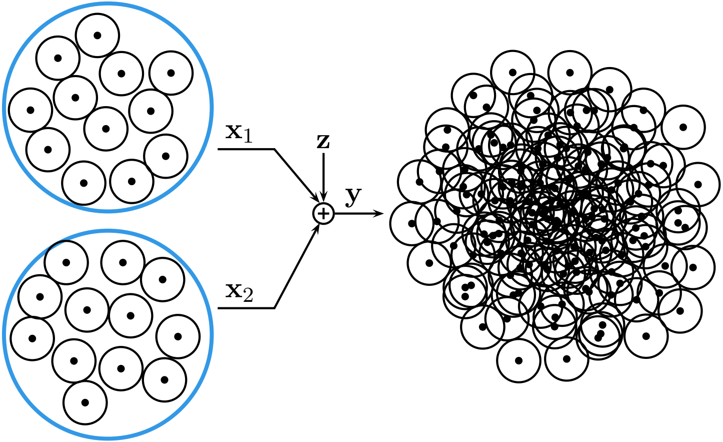
Random i.i.d. codes are not good for computation



2^{nR} codewords each.

2^{n2R} possible sums of codewords.

Random i.i.d. codes are not good for computation

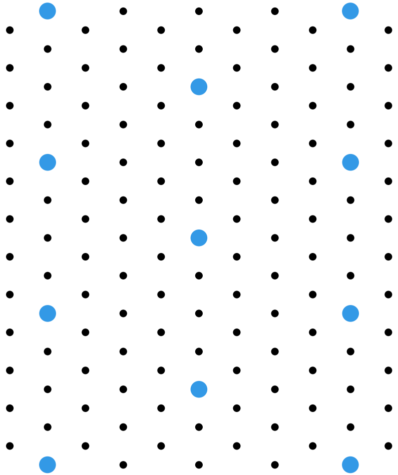


2^{nR} codewords each.

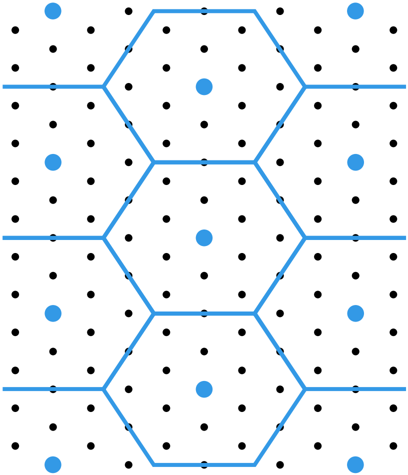
2^{n2R} possible sums of codewords.

Unequal Power Constraints – Double Nesting

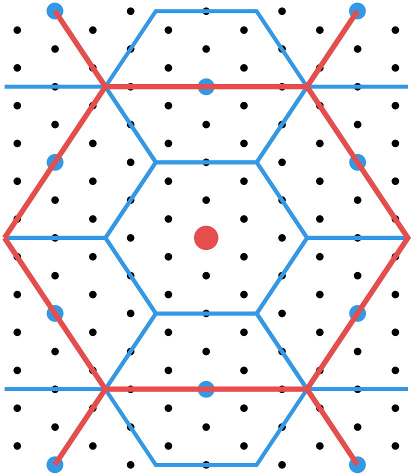
- What if the power constraints are not equal?
- Idea from **Nam-Chung-Lee '10**:
- Draw the codewords from the same fine lattice Λ_{FINE}.
- Use two nested coarse lattices Λ_1 and Λ_2 to enforce the power constraints P_1 and P_2.



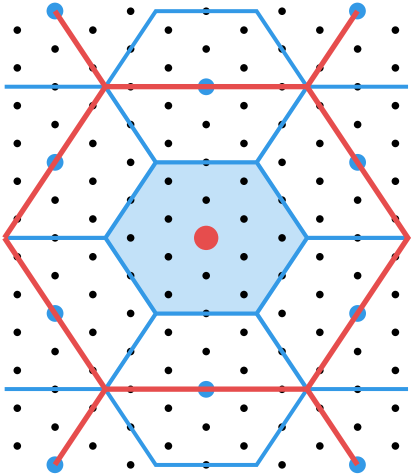
- What if the power constraints are not equal?
- Idea from **Nam-Chung-Lee '10**:
- Draw the codewords from the same fine lattice Λ_{FINE}.
- Use two nested coarse lattices Λ_1 and Λ_2 to enforce the power constraints P_1 and P_2.



- What if the power constraints are not equal?
- Idea from **Nam-Chung-Lee '10**:
- Draw the codewords from the same fine lattice Λ_{FINE}.
- Use two nested coarse lattices Λ_1 and Λ_2 to enforce the power constraints P_1 and P_2.

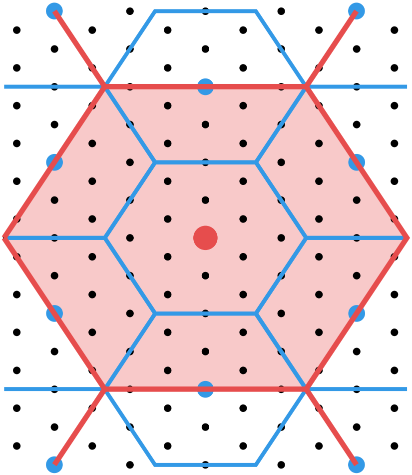


- What if the power constraints are not equal?
- Idea from **Nam-Chung-Lee '10**:
- Draw the codewords from the same fine lattice Λ_{FINE}.
- Use two nested coarse lattices Λ_1 and Λ_2 to enforce the power constraints P_1 and P_2.

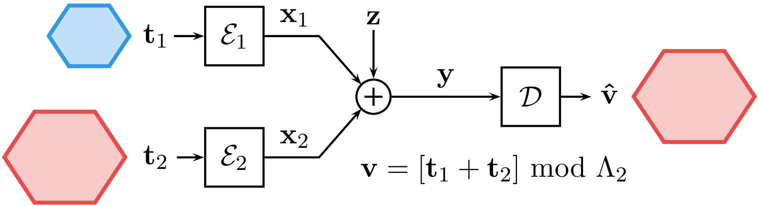


Unequal Power Constraints – Double Nesting

- What if the power constraints are not equal?
- Idea from **Nam-Chung-Lee '10**:
- Draw the codewords from the same fine lattice Λ_{FINE}.
- Use two nested coarse lattices Λ_1 and Λ_2 to enforce the power constraints P_1 and P_2.



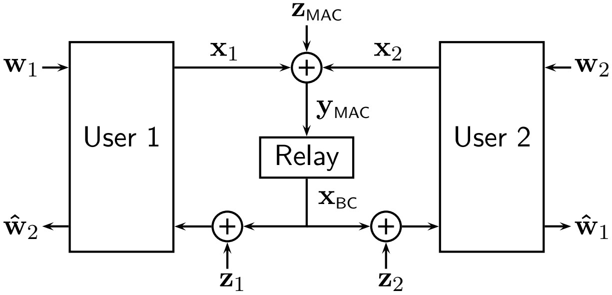
Unequal Power Constraints – Double Nesting



- Encoder 1 sends $\mathbf{x}_1 = [\mathbf{t}_1 + \mathbf{d}_1] \bmod \Lambda_1$. Coarse lattice Λ_1 has second moment P_1.
- Encoder 2 sends $\mathbf{x}_2 = [\mathbf{t}_2 + \mathbf{d}_2] \bmod \Lambda_2$. Coarse lattice Λ_2 has second moment $P_2 > P_1$.
- Decoder performs MMSE scaling, remove dithers, recovers $\bmod \Lambda_2$ sum.

$$R_1 = \frac{1}{2}\log\left(\frac{P_1}{P_1+P_2} + \frac{P_1}{N}\right) \qquad R_2 = \frac{1}{2}\log\left(\frac{P_2}{P_1+P_2} + \frac{P_2}{N}\right)$$

AWGN Two-Way Relay Channel



- User powers P_1, P_2.
- MAC noise variance N_{MAC}.
- Relay power P_{BC}.
- Broadcast noise variances N_1, N_2.

Theorem (Nam-Chung-Lee '10)

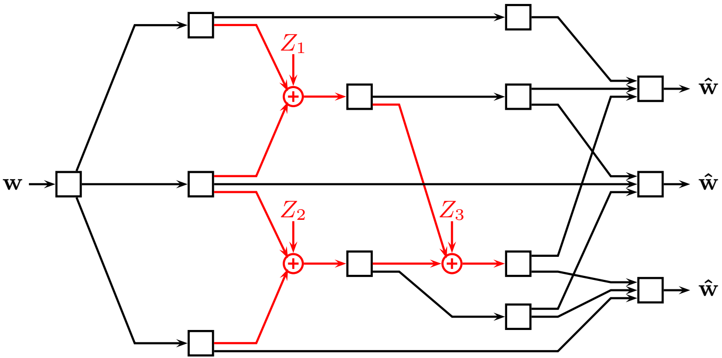
Capacity region is within $1/2$ *bit of:*

$$R_1 \leq \min\left(\frac{1}{2}\log\left(\frac{P_1}{P_1+P_2}+\frac{P_1}{N_{MAC}}\right),\ \frac{1}{2}\log\left(1+\frac{P_{BC}}{N_2}\right)\right)$$

$$R_2 \leq \min\left(\frac{1}{2}\log\left(\frac{P_2}{P_1+P_2}+\frac{P_2}{N_{MAC}}\right),\ \frac{1}{2}\log\left(1+\frac{P_{BC}}{N_1}\right)\right)$$

Moreover, "constant gap" goes to zero as powers increase.

Multiple-Access Networks



- Multicast demands
- Multi-access interference
- No broadcast constraints

- Compute-and-forward is well-suited for multicasting over multiple-access networks.
- Equal transmitter powers: **Nazer-Gastpar '07.**
  Unequal transmitter powers: **Nam-Chung-Lee '09.**

I. Discrete Alphabets

II. AWGN Channels

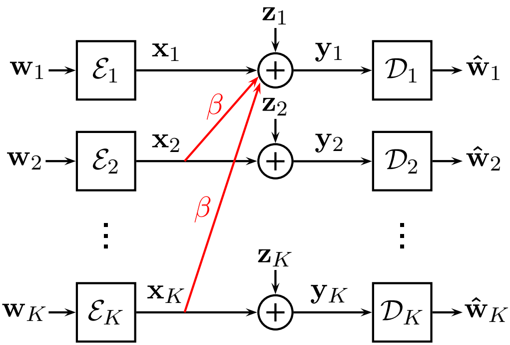
III. Network Applications

Many-to-One Interference Channel – Symmetric Very Strong Case

- Equal rates R.
- Only receiver 1 sees interference:

$$\mathbf{y}_1 = \mathbf{x}_1 + \beta \sum_{\ell=2}^{K} \mathbf{x}_\ell + \mathbf{z}_1$$

- How big does β have to be to achieve $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$? (i.e. "very strong" case)



- Scheme A: Decode $\mathbf{w}_2, \ldots, \mathbf{w}_K$ at receiver 1 and remove prior to decoding $\mathbf{w}_1$.

$$R \leq \frac{1}{2(K-1)} \log\left(1 + \frac{\beta^2 (K-1) P}{N + P}\right)$$

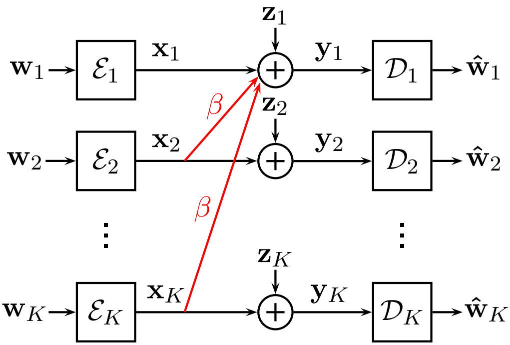
- Scheme B: Decode $\mathbf{w}_2 \oplus \cdots \oplus \mathbf{w}_K$ at receiver 1 and remove prior to decoding $\mathbf{w}_1$.

Many-to-One Interference Channel – Symmetric Very Strong Case

Encoders use the same nested lattice codebook.

Transmit dithered codewords:

$$\mathbf{x}_\ell = [\mathbf{t}_\ell + \mathbf{d}_\ell] \bmod \Lambda$$



Decoder scales by β^{-1}, removes dithers, recovers modulo sum.

$$\left[\beta^{-1}\mathbf{y}_1 - \sum_{\ell=2}^{K} \mathbf{d}_\ell\right] \bmod \Lambda = \left[\sum_{\ell=2}^{K}(\mathbf{x}_\ell - \mathbf{d}_\ell) + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)\right] \bmod \Lambda$$

(Distributive Law) $$= \left[\left[\sum_{\ell=2}^{K} \mathbf{t}_\ell\right] \bmod \Lambda + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)\right] \bmod \Lambda$$

$$\left[\beta^{-1}\mathbf{y}_1 - \sum_{\ell=2}^{K} \mathbf{d}_\ell\right] \bmod \Lambda = \left[\left[\sum_{\ell=2}^{K} \mathbf{t}_\ell\right] \bmod \Lambda + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)\right] \bmod \Lambda$$

- Effective noise variance $N_{\text{EFFEC}} = \beta^{-2}(P+N)$.
- Can decode $\bmod \Lambda$ sum of lattice points at rate $R = \frac{1}{2}\log\left(\frac{\beta^2 P}{P+N}\right)$.
- Setting equal to "very strong" condition $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$ we get

$$\beta^2 = \frac{(P+N)^2}{PN}$$

- How can we recover $\mathbf{w}_1$?
- We need to first subtract the real sum of the codewords. So far, we only have the modulo-sum.

Successive Cancellation of Sums

- First, add back in dithers to get modulo sum of codewords:

$$\left[\left[\sum_{\ell=2}^{K} \mathbf{t}_\ell\right] \bmod \Lambda + \left[\sum_{\ell=2}^{K} \mathbf{d}_\ell\right] \bmod \Lambda\right] \bmod \Lambda = \left[\sum_{\ell=2}^{K} \mathbf{x}_\ell\right] \bmod \Lambda$$

Successive Cancellation of Sums

- First, add back in dithers to get modulo sum of codewords:

$$\left[\left[\sum_{\ell=2}^{K} \mathbf{t}_\ell\right] \bmod \Lambda + \left[\sum_{\ell=2}^{K} \mathbf{d}_\ell\right] \bmod \Lambda\right] \bmod \Lambda = \left[\sum_{\ell=2}^{K} \mathbf{x}_\ell\right] \bmod \Lambda$$

- Subtract from $\mathbf{y}_1$ to expose the coarse lattice point nearest to the real sum $\sum_{\ell=2}^{K} \mathbf{x}_\ell$:

$$\beta^{-1}\mathbf{y}_1 - \left[\sum_{\ell=2}^{K} \mathbf{x}_\ell\right] \bmod \Lambda = Q_\Lambda\Big(\sum_{\ell=2}^{K} \mathbf{x}_\ell\Big) + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)$$

- Coarse lattice point easier to decode than fine lattice point:

$$Q_\Lambda\left(Q_\Lambda\Big(\sum_{\ell=2}^{K} \mathbf{x}_\ell\Big) + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)\right) = Q_\Lambda\Big(\sum_{\ell=2}^{K} \mathbf{x}_\ell\Big) \quad \text{w.h.p.}$$

Successive Cancellation of Sums

- First, add back in dithers to get modulo sum of codewords:

$$\left[\left[\sum_{\ell=2}^{K}\mathbf{t}_\ell\right] \bmod \Lambda + \left[\sum_{\ell=2}^{K}\mathbf{d}_\ell\right] \bmod \Lambda\right] \bmod \Lambda = \left[\sum_{\ell=2}^{K}\mathbf{x}_\ell\right] \bmod \Lambda$$

- Subtract from $\mathbf{y}_1$ to expose the coarse lattice point nearest to the real sum $\sum_{\ell=2}^{K}\mathbf{x}_\ell$:

$$\beta^{-1}\mathbf{y}_1 - \left[\sum_{\ell=2}^{K}\mathbf{x}_\ell\right] \bmod \Lambda = Q_\Lambda\Big(\sum_{\ell=2}^{K}\mathbf{x}_\ell\Big) + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)$$

- Coarse lattice point easier to decode than fine lattice point:

$$Q_\Lambda\left(Q_\Lambda\Big(\sum_{\ell=2}^{K}\mathbf{x}_\ell\Big) + \beta^{-1}(\mathbf{x}_1 + \mathbf{z}_1)\right) = Q_\Lambda\Big(\sum_{\ell=2}^{K}\mathbf{x}_\ell\Big) \quad \text{w.h.p.}$$

- Finally, get back the real sum

$$\left[\sum_{\ell=2}^{K}\mathbf{x}_\ell\right] \bmod \Lambda + Q_\Lambda\Big(\sum_{\ell=2}^{K}\mathbf{x}_\ell\Big) = \sum_{\ell=2}^{K}\mathbf{x}_\ell$$

Successive Cancellation of Sums

- We now have the sum of interfering codewords and can cancel them out:

$$\mathbf{y}_1 - \beta \sum_{\ell=2}^{K} \mathbf{x}_\ell = \mathbf{x}_1 + \mathbf{z}_1$$

- Can apply standard MMSE lattice decoding to recover lattice point $\mathbf{t}_1$ and then map back to $\mathbf{w}_1$.
- Overall, structured coding permits

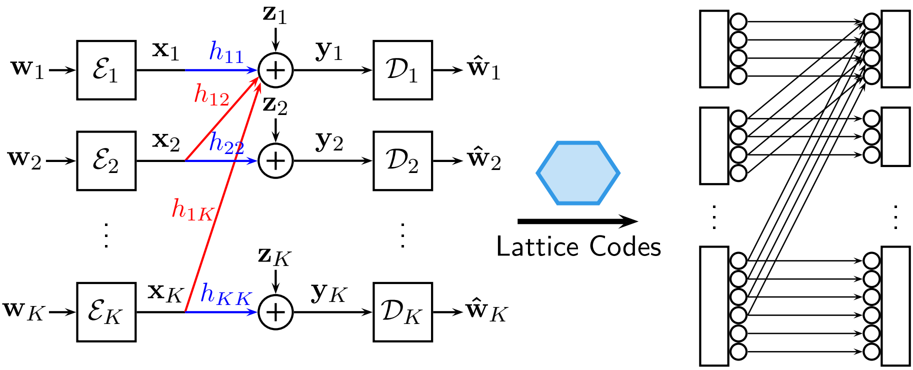
$$\beta^2 \geq \frac{(P+N)^2}{PN}$$

- Compare to decoding interfering codewords in their entirety:

$$\beta^2 \geq \frac{\left((1+\frac{P}{N})^{K-1} - 1\right)(N+P)}{(K-1)P}$$

- Originally shown in **Sridharan-Jafarian-Vishwanath-Jafar '08** using spherical shaping region. Nested lattice scheme from **Nazer '11**.

Many-to-One Interference Channel – Approximate Capacity

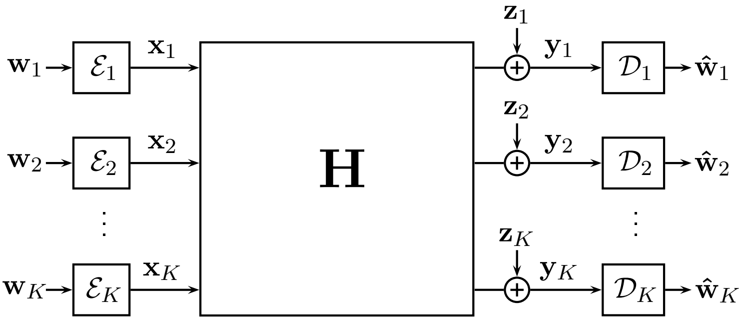


- Deterministic model by **Avestimehr-Diggavi-Tse '11** shows how to decompose by signal scale.

Theorem (Bresler-Parekh-Tse '10)

Lattices codes combined with the deterministic model can approach the capacity region to within $(3K+3)(1+\log(K+1))$ *bits per user.*

Interference Channel – Symmetric Very Strong Case

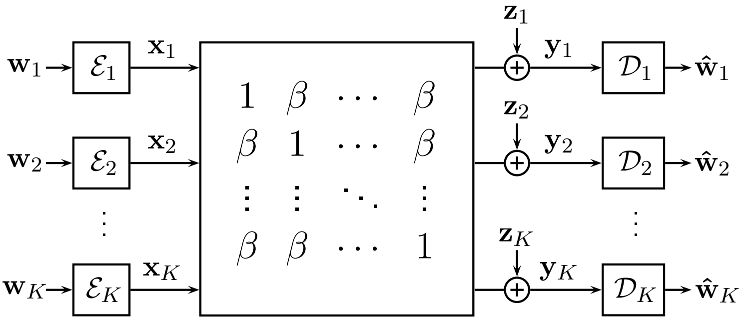


- Equal rates R. How big does β have to be to achieve $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$? (i.e. "very strong" case)
- Can use the many-to-one decoder at every receiver to get

$$\beta^2 \geq \frac{(P+N)^2}{PN}$$

- What about asymmetric interference channels?

Interference Channel – Symmetric Very Strong Case

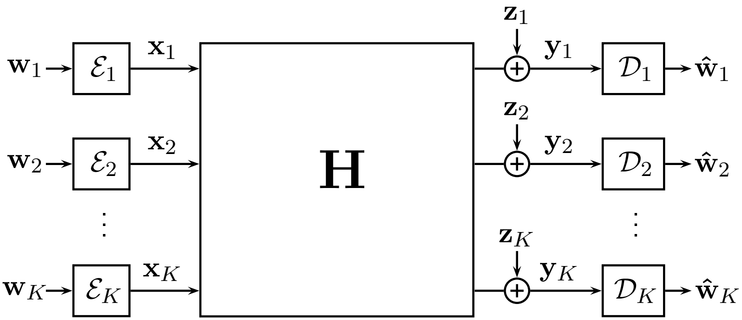


- Equal rates R. How big does β have to be to achieve $R = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$? (i.e. "very strong" case)
- Can use the many-to-one decoder at every receiver to get

$$\beta^2 \geq \frac{(P+N)^2}{PN}$$

- What about asymmetric interference channels?

Interference Channel



- Not clear how to map to a deterministic model using lattices.
- "Real" interference alignment scheme of **Motahari et al. '08** uses a lattice structure to get $K/2$ DoF (up to a set of measure one)
- Some special cases at finite SNR: **Jafarian-Viswanath '09,'10, Ordentlich-Erez '11**
- Much more known for time-varying channels: **Cadambe-Jafar '08, Nazer et al. '11, much more**

Summary

- So far we have seen that lattices are very effective for scenarios where there is a single interference bottleneck.
- Also effective for multiple bottlenecks but less is known.
- We have so far assumed that the fading coefficients are known at the transmitters.
- In general, transmitters may not have access to channel state information.

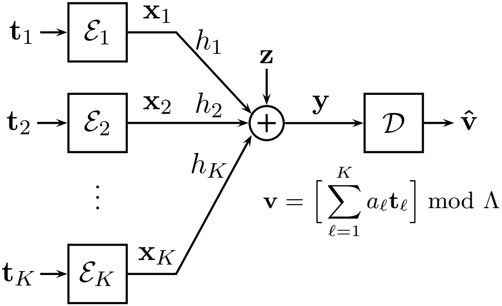
Computation over Fading Channels

Transmitters do not know channel realization.

Encoders use the same nested lattice codebook.

Transmit dithered codewords:

$$\mathbf{x}_\ell = [\mathbf{t}_\ell + \mathbf{d}_\ell] \bmod \Lambda$$



- Decoder removes dithers and recovers integer combination

$$\mathbf{v} = \left[\sum_{\ell=1}^{K} a_\ell \mathbf{t}_\ell\right] \bmod \Lambda$$

- Receiver can use its knowledge of the channel gains to match the equation coefficients a_ℓ to the channel coefficients h_ℓ.

- Distributive Law also holds for integer combinations. Let $a, b \in \mathbb{Z}$.

$$\begin{aligned}
&\Big[a[\mathbf{x}_1] \bmod \Lambda + b[\mathbf{x}_2] \bmod \Lambda\Big] \bmod \Lambda \\
&= \Big[a\Big(\mathbf{x}_1 - Q_\Lambda(\mathbf{x}_1)\Big) + b\Big(\mathbf{x}_2 - Q_\Lambda(\mathbf{x}_2)\Big)\Big] \bmod \Lambda \\
&= \Big[a\mathbf{x}_1 + b\mathbf{x}_2 - aQ_\Lambda(\mathbf{x}_1) - bQ_\Lambda(\mathbf{x}_2)\Big] \bmod \Lambda \\
&= [a\mathbf{x}_1 + b\mathbf{x}_2] \bmod \Lambda
\end{aligned}$$

- Last step follows since since $aQ_\Lambda(\mathbf{x}_1)$ and $bQ_\Lambda(\mathbf{x}_2)$ are elements of the lattice Λ.

- Transmit dithered codewords $\mathbf{x}_\ell = [\mathbf{t}_\ell + \mathbf{d}_\ell] \bmod \Lambda$
- Decoder removes dithers and recovers integer combination

$$\left[\mathbf{y} - \sum_{\ell=1}^{K} a_\ell \mathbf{d}_\ell\right] \bmod \Lambda$$

$$= \left[\sum_{\ell=1}^{K} h_\ell \mathbf{x}_\ell + \mathbf{z} - \sum_{\ell=1}^{K} a_\ell \mathbf{d}_\ell\right] \bmod \Lambda$$

$$= \left[\sum_{\ell=1}^{K} a_\ell (\mathbf{x}_\ell - \mathbf{d}_\ell) + \sum_{\ell=1}^{K} (h_\ell - a_\ell)\mathbf{x}_\ell + \mathbf{z}\right] \bmod \Lambda$$

$$= \left[\left[\sum_{\ell=1}^{K} a_\ell \mathbf{t}_\ell\right] \bmod \Lambda + \underbrace{\sum_{\ell=1}^{K} (h_\ell - a_\ell)\mathbf{x}_\ell + \mathbf{z}}_{\text{Effective Noise}}\right] \bmod \Lambda \quad \text{Distributive Law}$$

Computation over Fading Channels – Effective Noise

- Effective noise due to mismatch between channel coefficients $\mathbf{h} = [h_1 \cdots h_K]^T$ and equation coefficients $\mathbf{a} = [a_1 \cdots a_K]^T$.

$$N_{\mathsf{EFFEC}} = N + P\|\mathbf{h} - \mathbf{a}\|^2$$

$$R = \frac{1}{2}\log\left(\frac{P}{N + P\|\mathbf{h} - \mathbf{a}\|^2}\right)$$

Computation over Fading Channels – Effective Noise

- Effective noise due to mismatch between channel coefficients $\mathbf{h} = [h_1 \cdots h_K]^T$ and equation coefficients $\mathbf{a} = [a_1 \cdots a_K]^T$.

$$N_{\mathsf{EFFEC}} = N + P\|\mathbf{h} - \mathbf{a}\|^2$$
$$R = \frac{1}{2}\log\left(\frac{P}{N + P\|\mathbf{h} - \mathbf{a}\|^2}\right)$$

- Can do better with MMSE scaling.

$$N_{\mathsf{EFFEC}} = \alpha^2 N + P\|\alpha\mathbf{h} - \mathbf{a}\|^2$$
$$R = \max_{\alpha}\frac{1}{2}\log\left(\frac{P}{\alpha^2 N + P\|\alpha\mathbf{h} - \mathbf{a}\|^2}\right)$$
$$= \frac{1}{2}\log\left(\frac{N + P\|\mathbf{h}\|^2}{N\|\mathbf{a}\|^2 + P(\|\mathbf{h}\|^2\|\mathbf{a}\|^2 - (\mathbf{h}^T\mathbf{a})^2)}\right)$$

- See **Nazer-Gastpar '11** for more details.

- The rate expression simplifies in some special cases.

$$R = \frac{1}{2}\log\left(\frac{N + P\|\mathbf{h}\|^2}{N\|\mathbf{a}\|^2 + P(\|\mathbf{h}\|^2\|\mathbf{a}\|^2 - (\mathbf{h}^T\mathbf{a})^2)}\right)$$

- Integer channels: $\mathbf{h} = \mathbf{a}$.

$$R = \frac{1}{2}\log\left(\frac{1}{\|\mathbf{a}\|^2} + \frac{P}{N}\right)$$

- Recovering a single message: Set $\mathbf{a} = \delta_m$, the m^{th} unit vector.

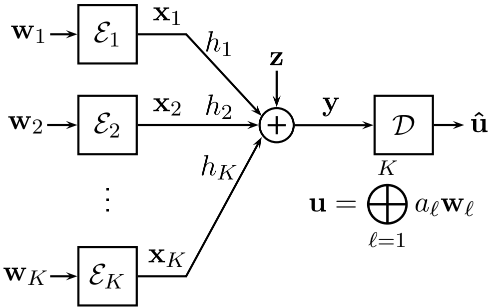
$$R = \frac{1}{2}\log\left(1 + \frac{h_m^2 P}{N + P\sum_{\ell\neq m} h_\ell^2}\right)$$

Transmitters do not know channel realization.

Encoders use the same nested lattice codebook.

Transmit dithered codewords:

$$\mathbf{x}_\ell = [\mathbf{t}_\ell + \mathbf{d}_\ell] \bmod \Lambda$$

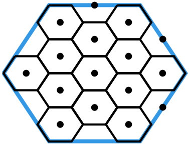


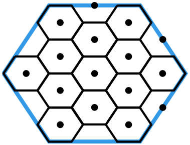
- Recall that mapping $\mathbf{t}_\ell = \phi(\mathbf{w}_\ell)$ between messages and lattice points preserves linearity.

$$\phi^{-1}\left(\left[\sum_{\ell=1}^{K} a_\ell \mathbf{t}_\ell\right] \bmod \Lambda\right) = \left[\sum_{\ell=1}^{K} a_\ell \mathbf{w}_\ell\right] \bmod q = \bigoplus_{\ell=1}^{K} a_\ell \mathbf{w}_\ell$$

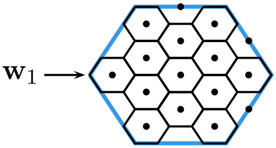
- Digital interface that fits well with **network coding**.

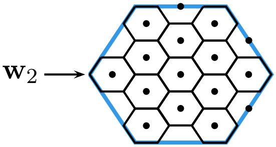
All users pick the same nested lattice code:



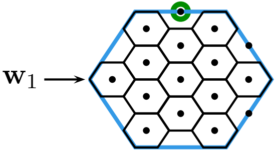


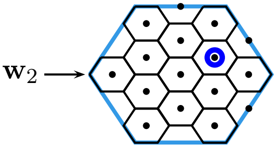
Choose messages over field $\mathbf{w}_\ell \in \mathbb{F}_q^k$:



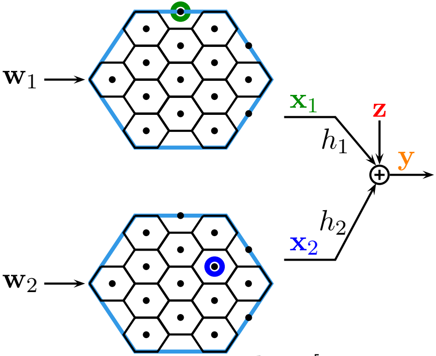


Map $\mathbf{w}_\ell$ to lattice point $\mathbf{t}_\ell = \phi(\mathbf{w}_\ell)$:





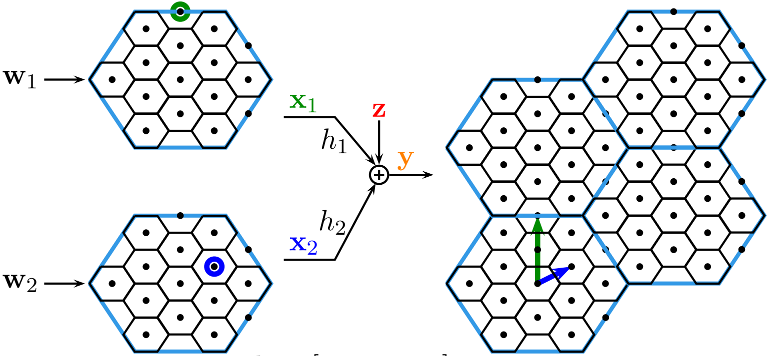
Transmit lattice points over the channel:



$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

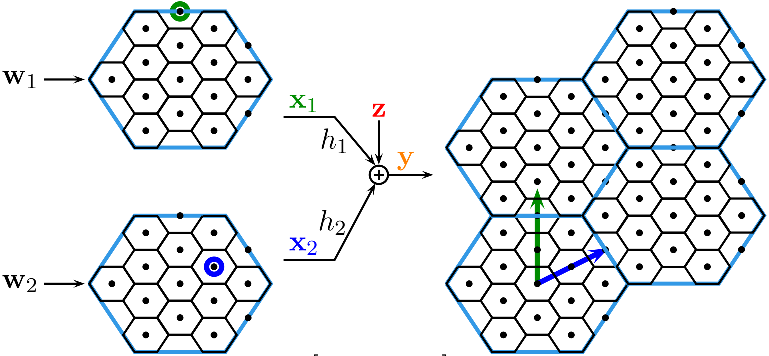
Transmit lattice points over the channel:



$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

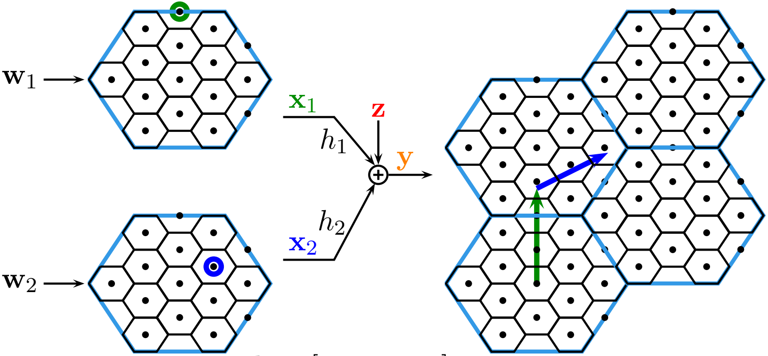
Lattice codewords are scaled by channel coefficients:



$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Scaled codewords added together plus noise:

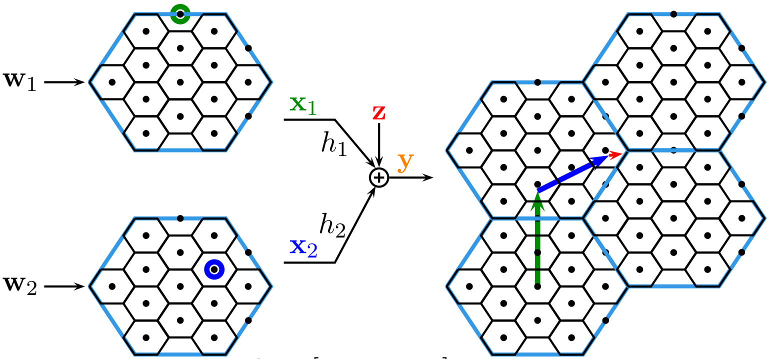


$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Computation Coding

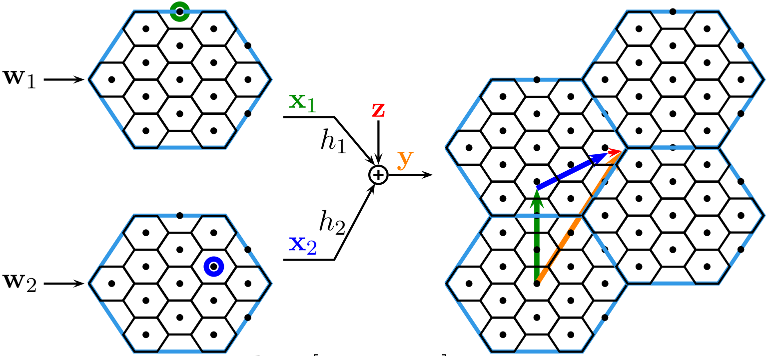
Scaled codewords added together plus noise:



$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Extra noise penalty for non-integer channel coefficients:

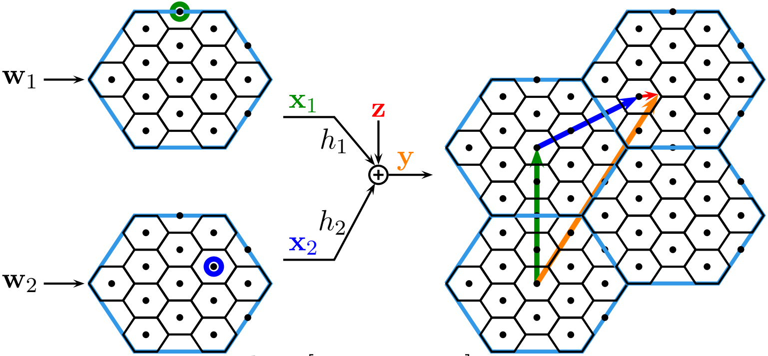


$$\mathbf{h} = [\ 1.4 \quad 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $N + P\|\mathbf{h} - \mathbf{a}\|^2$

Scale output by α to reduce non-integer noise penalty:

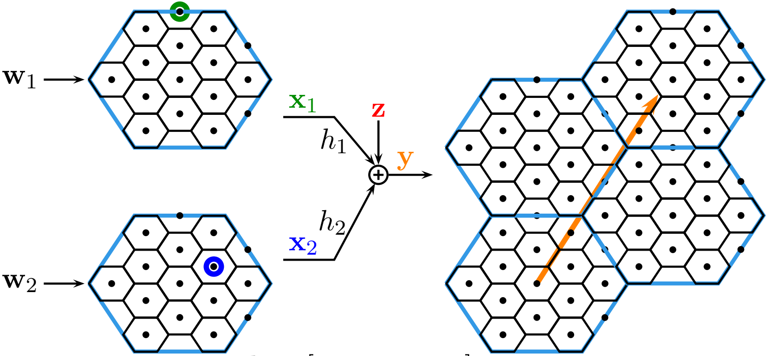


$$\alpha\mathbf{h} = [\ \alpha 1.4 \quad \alpha 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $\alpha^2 N + P\|\alpha\mathbf{h} - \mathbf{a}\|^2$

Scale output by α to reduce non-integer noise penalty:

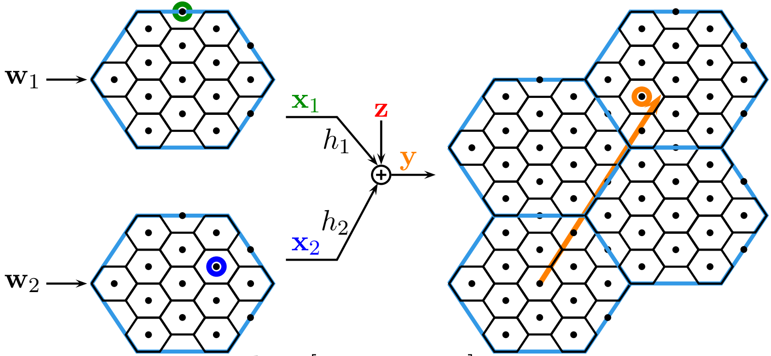


$$\alpha\mathbf{h} = [\ \alpha 1.4 \quad \alpha 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $\alpha^2 N + P\|\alpha\mathbf{h} - \mathbf{a}\|^2$

Decode to closest lattice point:

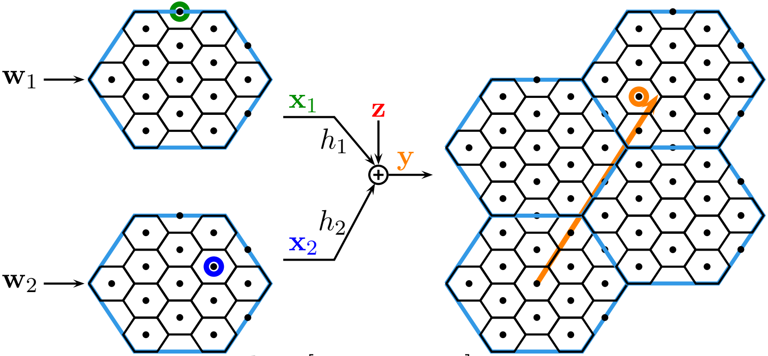


$$\alpha\mathbf{h} = [\ \alpha 1.4 \quad \alpha 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $\alpha^2 N + P\|\alpha\mathbf{h} - \mathbf{a}\|^2$

Compute sum of lattice points modulo the coarse lattice:

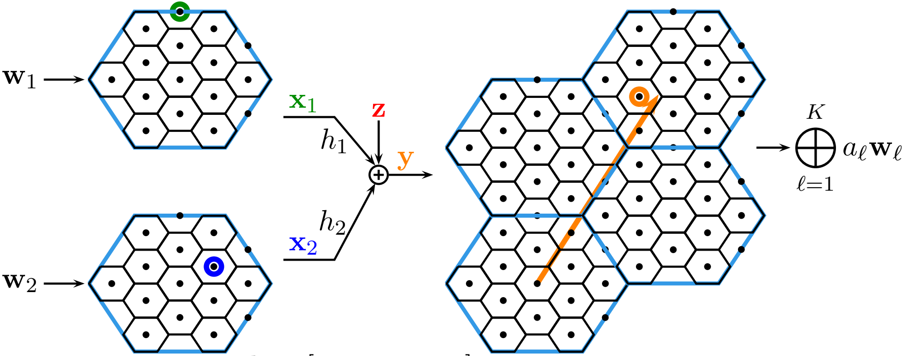


$$\alpha \mathbf{h} = [\ \alpha 1.4 \quad \alpha 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $\alpha^2 N + P\|\alpha \mathbf{h} - \mathbf{a}\|^2$

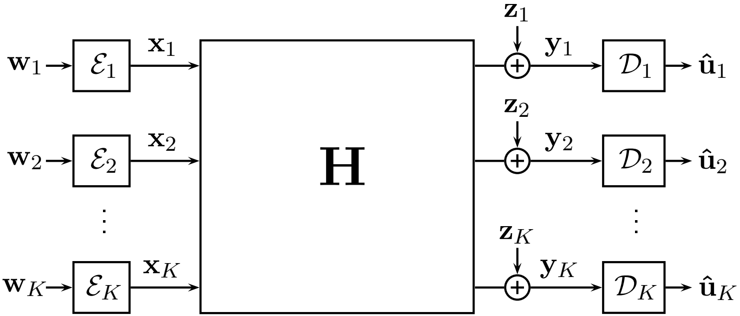
Map back to equation of message symbols over the field:



$$\alpha \mathbf{h} = [\ \alpha 1.4 \quad \alpha 2.1\ ]$$

$$\mathbf{a} = [\ 2 \quad 3\ ]$$

Effective noise: $\alpha^2 N + P \|\alpha \mathbf{h} - \mathbf{a}\|^2$

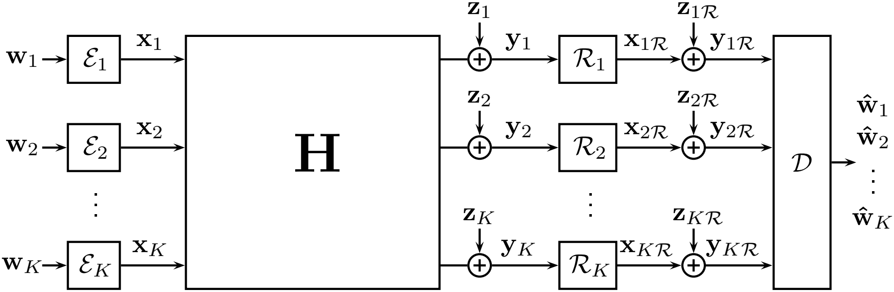


- Equal rates R. No channel state information (CSI) at transmitters.
- Receivers use their CSI to select coefficients, decode linear equation

$$\mathbf{u}_k = \bigoplus_{\ell=1}^{K} a_{k\ell} \mathbf{w}_\ell$$

- Reliable decoding possible if

$$R < \min_{k : a_{k\ell} \neq 0} \frac{1}{2} \log \left(\frac{N + P \|\mathbf{h}_k\|^2}{N \|\mathbf{a}_k\|^2 + P \left(\|\mathbf{h}_k\|^2 \|\mathbf{a}_k\|^2 - (\mathbf{h}_k^T \mathbf{a}_k)^2 \right)} \right)$$



- Equal rates R. $\mathbf{H}$ is a Hadamard matrix, $\mathbf{H}\mathbf{H}^T = K\mathbf{I}$

Upper Bound

$$\frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

Compute-and-Forward

$$\frac{1}{2}\log\left(\frac{1}{K}+\frac{P}{N}\right)$$

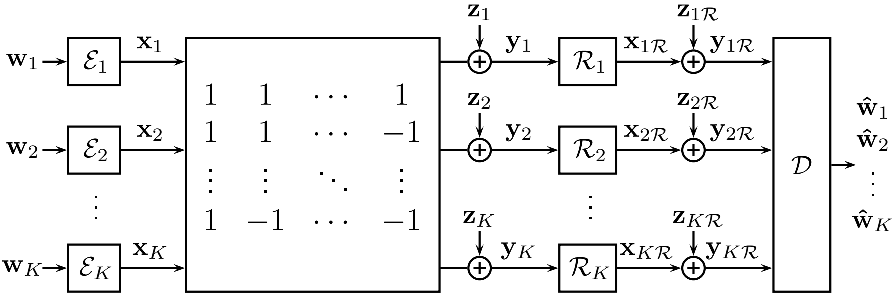
Compress-and-Forward

$$\frac{1}{2}\log\left(1+\frac{P}{N}\frac{P}{N+KP}\right)$$

Decode-and-Forward

$$\frac{1}{2K}\log\left(1+\frac{KP}{N}\right)$$

Case Study – Hadamard Relay Network



- Equal rates R. $\mathbf{H}$ is a Hadamard matrix, $\mathbf{H}\mathbf{H}^T = K\mathbf{I}$

Upper Bound

$$\frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

Compute-and-Forward

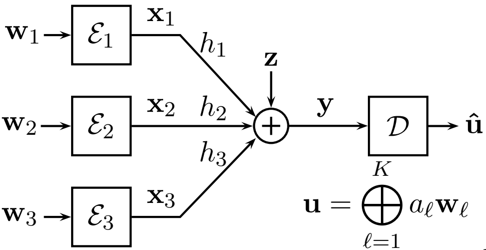
$$\frac{1}{2}\log\left(\frac{1}{K}+\frac{P}{N}\right)$$

Compress-and-Forward

$$\frac{1}{2}\log\left(1+\frac{P}{N}\frac{P}{N+KP}\right)$$

Decode-and-Forward

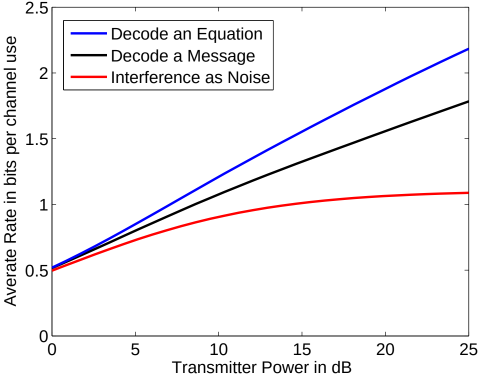
$$\frac{1}{2K}\log\left(1+\frac{KP}{N}\right)$$



$$\mathbf{u} = \bigoplus_{\ell=1}^{K} a_\ell \mathbf{w}_\ell$$

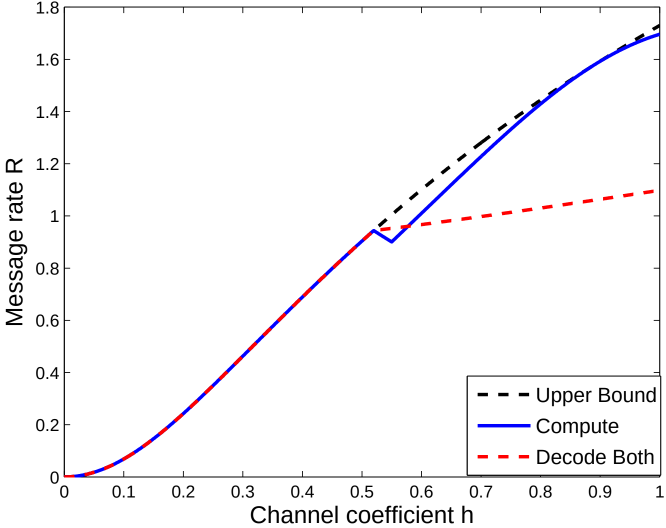
Relay either decodes some linear function of messages or an individual message.

- Three transmitters that do not know the fading coefficients.
- Average rate plotted for i.i.d. Gaussian fading.



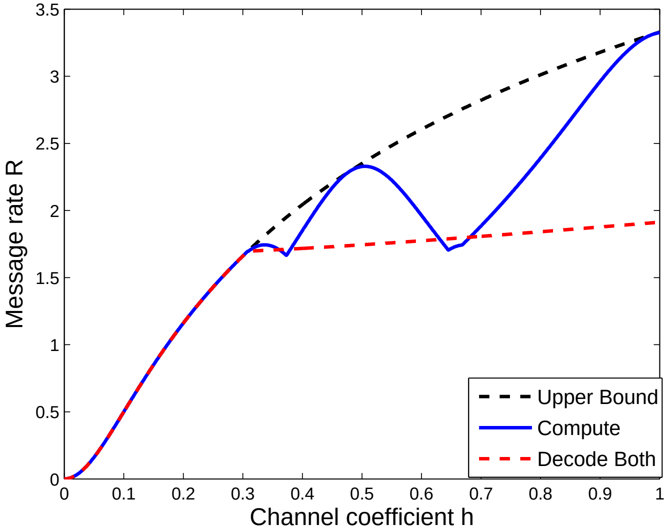
- Receiver observes $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$.
- Recovers $a\mathbf{w}_1 \oplus b\mathbf{w}_2$ for $a, b \neq 0$.

10dB



- Receiver observes $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$.
- Recovers $a\mathbf{w}_1 \oplus b\mathbf{w}_2$ for $a, b \neq 0$.

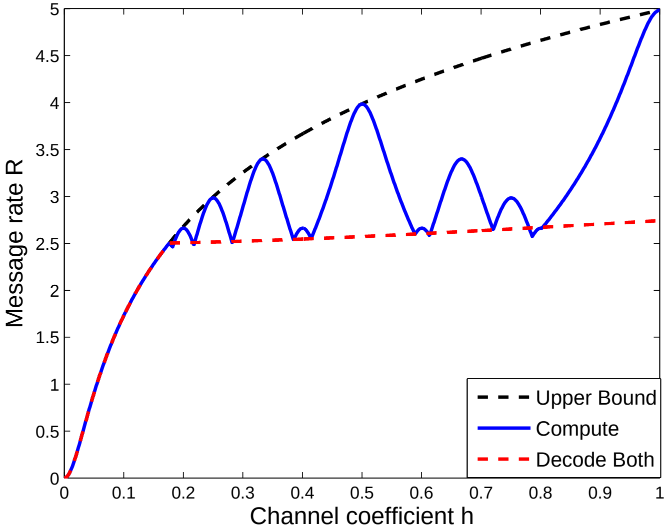
20dB



Computation over Fading Channels – No CSIT

- Receiver observes $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$.
- Recovers $a\mathbf{w}_1 \oplus b\mathbf{w}_2$ for $a, b \neq 0$.

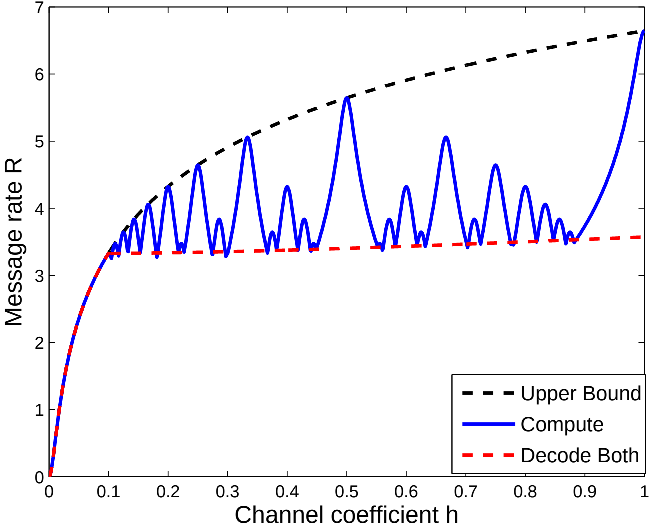
30dB



Computation over Fading Channels – No CSIT

- Receiver observes $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$.
- Recovers $a\mathbf{w}_1 \oplus b\mathbf{w}_2$ for $a, b \neq 0$.

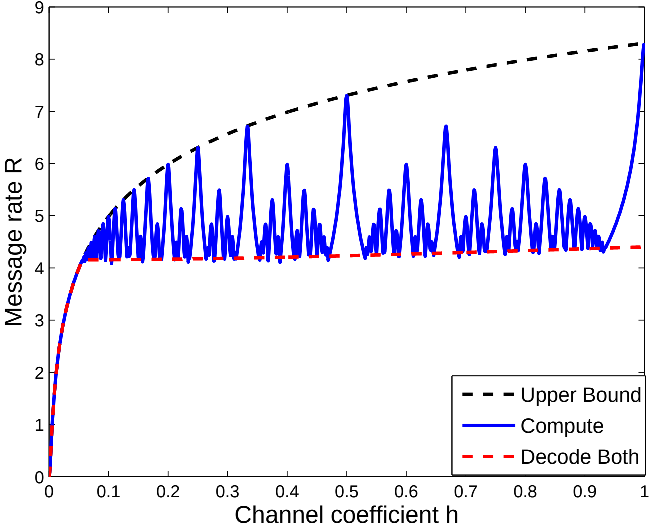
40dB

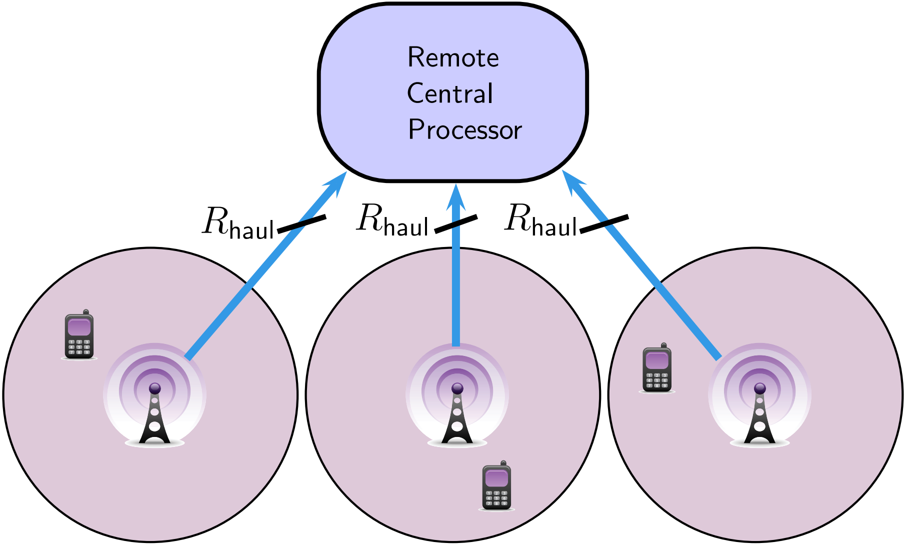


Computation over Fading Channels – No CSIT

- Receiver observes $\mathbf{y} = \mathbf{x}_1 + h\mathbf{x}_2 + \mathbf{z}$.
- Recovers $a\mathbf{w}_1 \oplus b\mathbf{w}_2$ for $a, b \neq 0$.

50dB

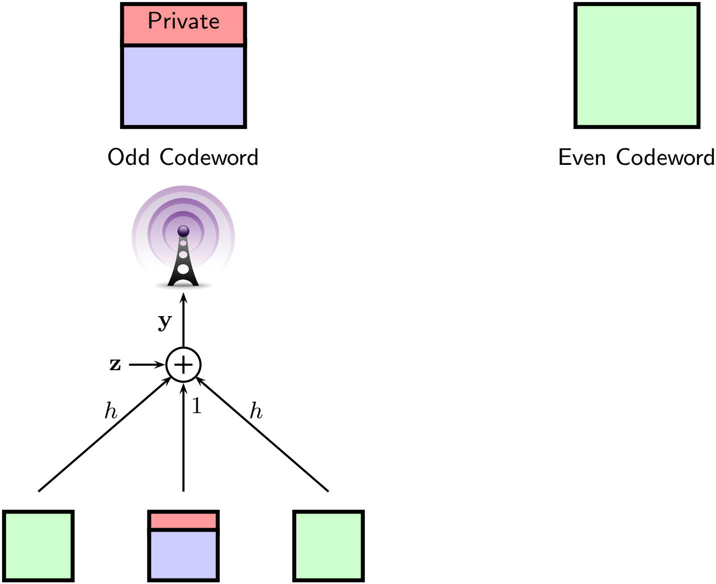




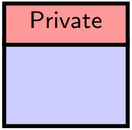
- Well-studied cellular model: **Wyner '94, Shamai-Wyner '97, Sanderovich et al. '09**

Structured Superposition

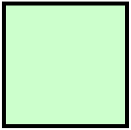
Structured Superposition



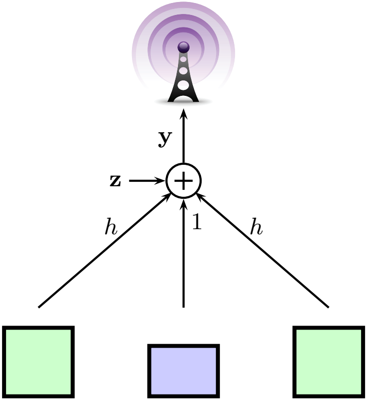
Structured Superposition



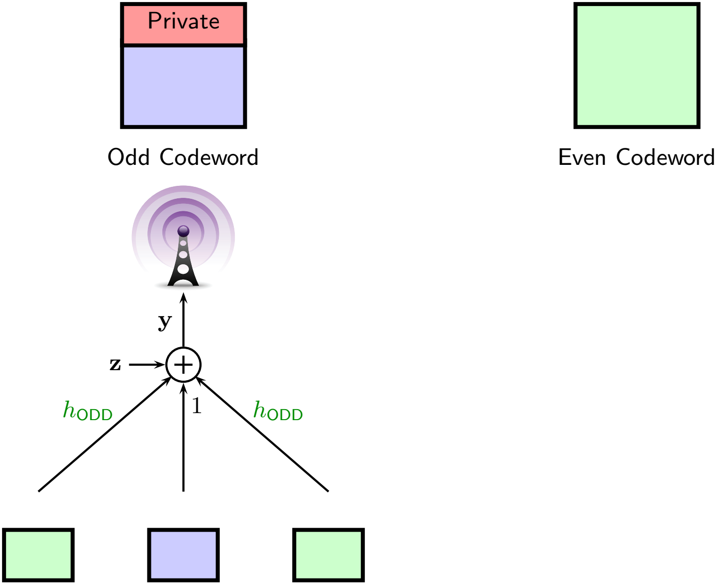
Odd Codeword



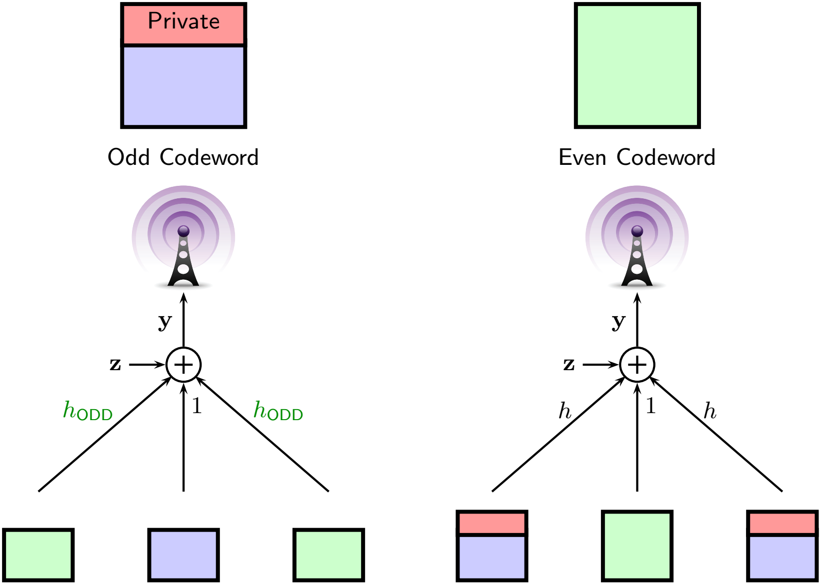
Even Codeword



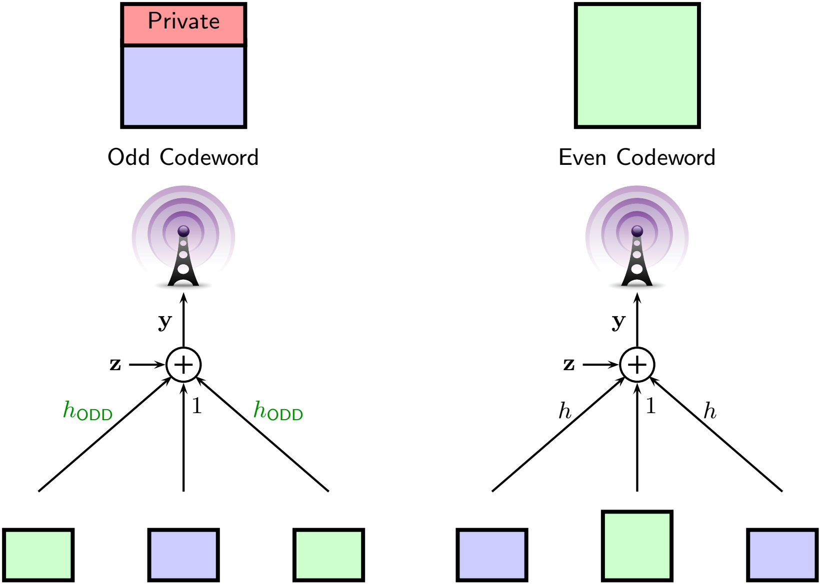
Structured Superposition



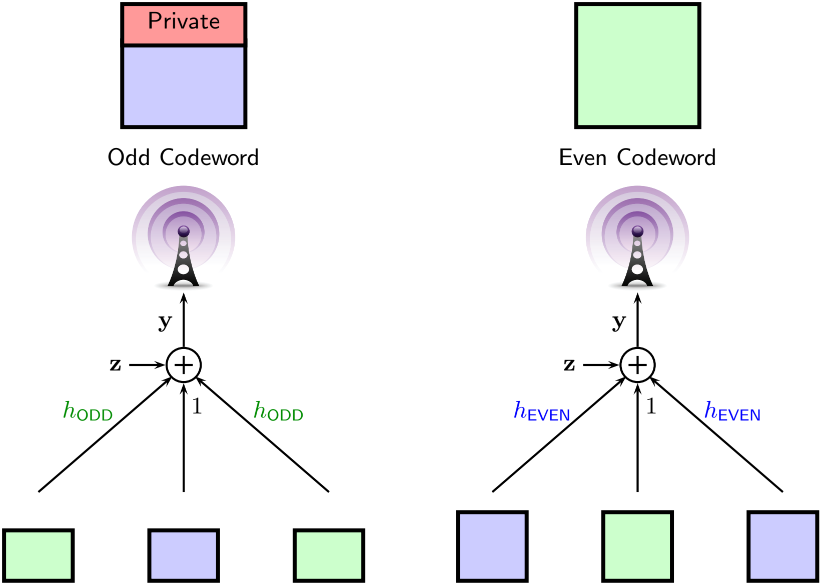
Structured Superposition



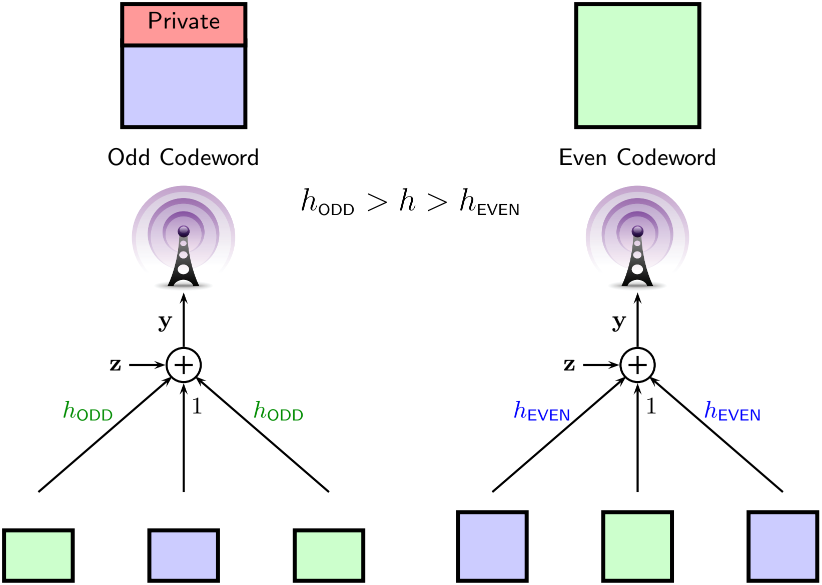
Structured Superposition

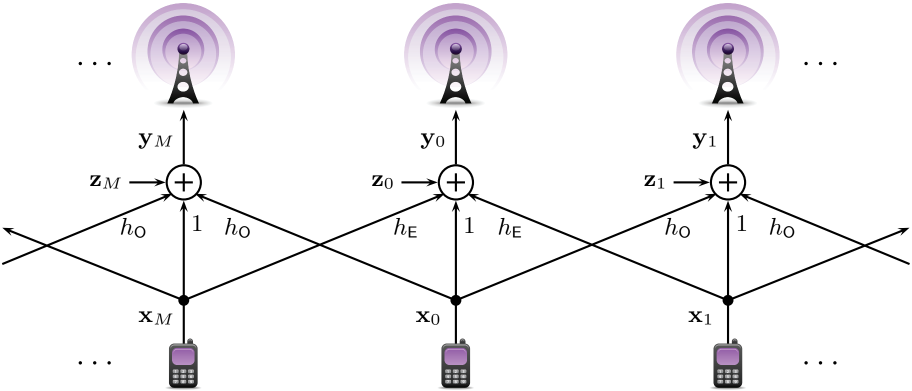


Structured Superposition

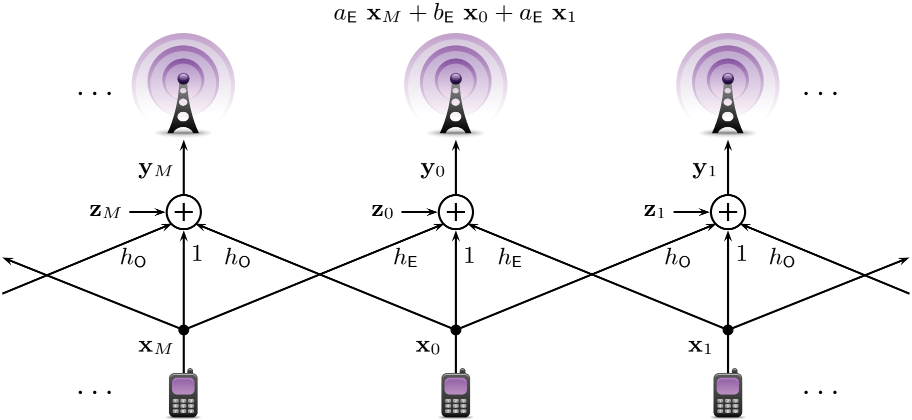


Structured Superposition



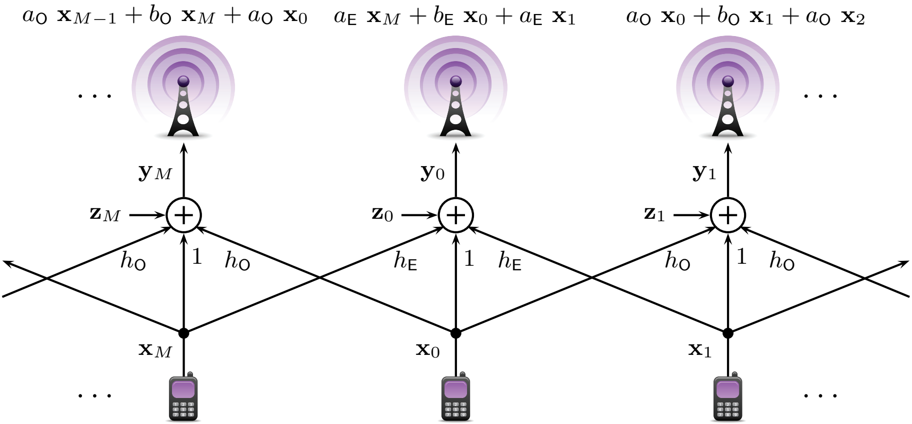


Nazer et al. '09: Each cell-site sees either h_{E} or h_{O} which is strictly better than h.



Nazer et al. '09: Each cell-site sees either h_E or h_O which is strictly better than h.

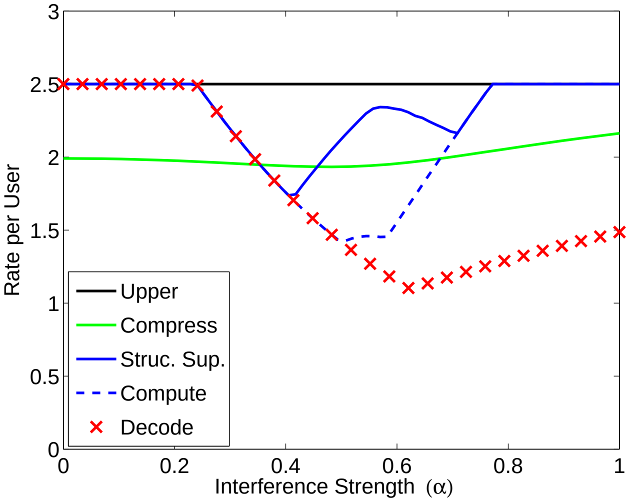
Structured Superposition



Nazer et al. '09: Each cell-site sees either h_E or h_O which is strictly better than h.

Structured Superposition: Performance

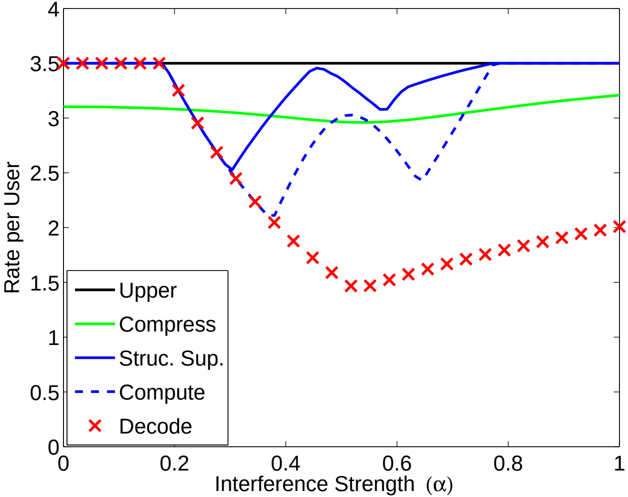
SNR $= 10$dB, Backhaul Rate $R_{\text{haul}} = 2.5$



- Compress-and-forward rate taken from **Sanderovich et al. '09**
- Layering can reduce "non-integer loss."

Structured Superposition: Performance

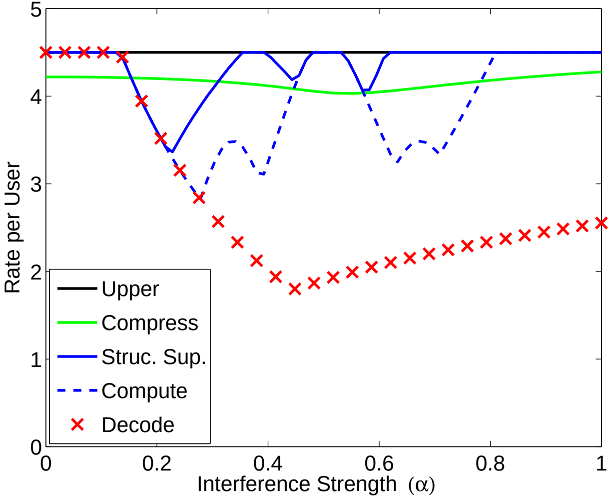
SNR $= 15$dB, Backhaul Rate $R_{\text{haul}} = 3.5$



- Compress-and-forward rate taken from **Sanderovich et al. '09**
- Layering can reduce "non-integer loss."

Structured Superposition: Performance

SNR $= 20$dB, Backhaul Rate $R_{\text{haul}} = 4.5$



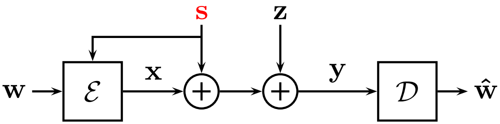
- Compress-and-forward rate taken from **Sanderovich et al. '09**
- Layering can reduce "non-integer loss."

- Choose equation coefficients to maximize rate:

$$R_{\textsf{COMP}} = \max_{\mathbf{a} \in \mathbb{Z}^K} \max_{\alpha} \frac{1}{2} \log \left(\frac{P}{\alpha^2 N + P \|\alpha \mathbf{h} - \mathbf{a}\|^2} \right)$$

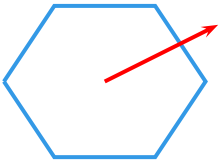
- Equivalently $\min_{\mathbf{a} \in \mathbb{Z}^K} \min_{\alpha} \alpha^2 N + P \|\alpha \mathbf{h} - \mathbf{a}\|^2$.
- Closely connected to Diophantine approximation, i.e. approximating irrationals with rationals.
- **Niesen-Whiting '11** shows that DoF $= \lim_{P \to \infty} \dfrac{R_{\textsf{COMP}}}{\frac{1}{2}\log(1+P)} \leq 2$
- Also shows that by combining compute-and-forward with interference alignment can get DoF to K.

$\mathbf{s}$ is interference known noncausally to the encoder.



Assume $\mathbf{s}$ i.i.d. Gaussian, very large variance P_S.

Erez-Shamai-Zamir '05:
Encoder subtracts $\alpha\mathbf{s}$, dithers, and takes $\mod \Lambda$.



$$\mathbf{x} = [\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \mod \Lambda$$

Decoder scales by α, removes dither, takes $\mod \Lambda$, and recovers $\mathbf{t}$.
Interference is cancelled.

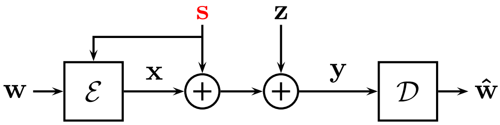
$$\begin{aligned}
[\alpha\mathbf{y} - \mathbf{d}] \bmod \Lambda &= [\mathbf{x} + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \\
&= \Big[[\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \bmod \Lambda + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda \\
&= \Big[\mathbf{t} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda
\end{aligned}$$

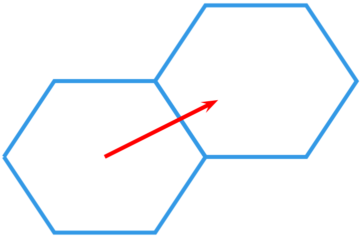
$\mathbf{s}$ is interference known noncausally to the encoder.

Assume $\mathbf{s}$ i.i.d. Gaussian, very large variance P_S.

Erez-Shamai-Zamir '05:
Encoder subtracts $\alpha\mathbf{s}$, dithers, and takes $\bmod \Lambda$.

$$\mathbf{x} = [\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \mod \Lambda$$

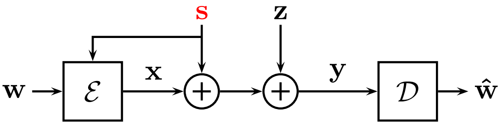




Decoder scales by α, removes dither, takes $\bmod \Lambda$, and recovers $\mathbf{t}$.
Interference is cancelled.

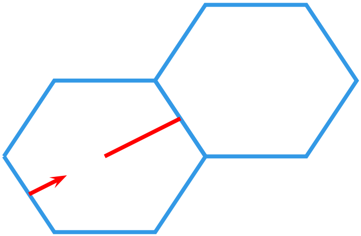
$$\begin{aligned}
[\alpha\mathbf{y} - \mathbf{d}] \bmod \Lambda &= [\mathbf{x} + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \\
&= \Big[[\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \bmod \Lambda + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda \\
&= \Big[\mathbf{t} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda
\end{aligned}$$

$\mathbf{s}$ is interference known noncausally to the encoder.



Assume $\mathbf{s}$ i.i.d. Gaussian, very large variance P_S.

Erez-Shamai-Zamir '05:
Encoder subtracts $\alpha\mathbf{s}$, dithers, and takes $\bmod \Lambda$.



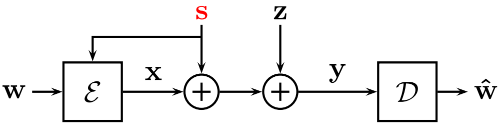
$$\mathbf{x} = [\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \bmod \Lambda$$

Decoder scales by α, removes dither, takes $\bmod \Lambda$, and recovers $\mathbf{t}$.
Interference is cancelled.

$$\begin{aligned}
[\alpha\mathbf{y} - \mathbf{d}] \bmod \Lambda &= [\mathbf{x} + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \\
&= \Big[[\mathbf{t} - \alpha\mathbf{s} + \mathbf{d}] \bmod \Lambda + \alpha\mathbf{s} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda \\
&= \Big[\mathbf{t} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda
\end{aligned}$$

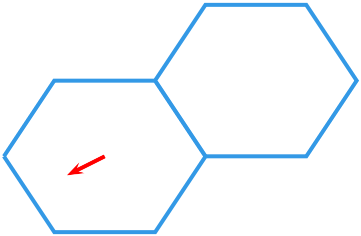
Dirty Paper Coding

$\color{red}{\mathbf{s}}$ is interference known noncausally to the encoder.



Assume $\color{red}{\mathbf{s}}$ i.i.d. Gaussian, very large variance P_S.

Erez-Shamai-Zamir '05:
Encoder subtracts $\alpha\color{red}{\mathbf{s}}$, dithers, and takes $\bmod \Lambda$.

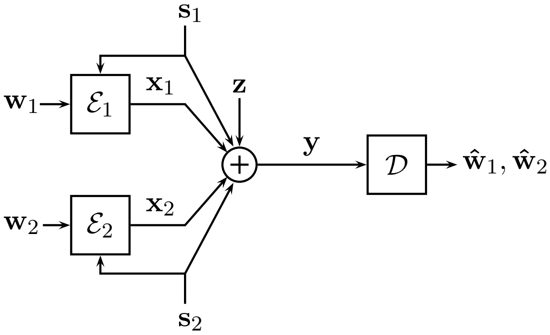


$$\mathbf{x} = [\mathbf{t} - \alpha\color{red}{\mathbf{s}} + \mathbf{d}] \mod \Lambda$$

Decoder scales by α, removes dither, takes $\bmod \Lambda$, and recovers $\mathbf{t}$.
Interference is cancelled.

$$\begin{aligned}
[\alpha\mathbf{y} - \mathbf{d}] \bmod \Lambda &= [\mathbf{x} + \alpha\color{red}{\mathbf{s}} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}] \bmod \Lambda \\
&= \Big[[\mathbf{t} - \alpha\color{red}{\mathbf{s}} + \mathbf{d}] \bmod \Lambda + \alpha\color{red}{\mathbf{s}} - \mathbf{d} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda \\
&= \Big[\mathbf{t} + \mathbf{z} - (1-\alpha)\mathbf{x}\Big] \bmod \Lambda
\end{aligned}$$

Dirty Gaussian Multiple-Access Channel



Philosof-Zamir-Erez-Khisti '11:

- Encoder 1 knows interference $\mathbf{s}_1$.
- Encoder 2 knows interference $\mathbf{s}_2$.
- Need to cancel out interference in a distributed fashion.
- Assume i.i.d. Gaussian interference with very large variance P_S. Random i.i.d. methods yield rate that goes to 0 as P_S goes to infinity.

Dirty Gaussian Multiple-Access Channel

Subtract (part of) the interference signals ahead of time:

$$\mathbf{x}_1 = [\mathbf{t}_1 - \alpha \mathbf{s}_1 + \mathbf{d}_1] \bmod \Lambda$$
$$\mathbf{x}_2 = [\mathbf{t}_2 - \alpha \mathbf{s}_2 + \mathbf{d}_2] \bmod \Lambda$$

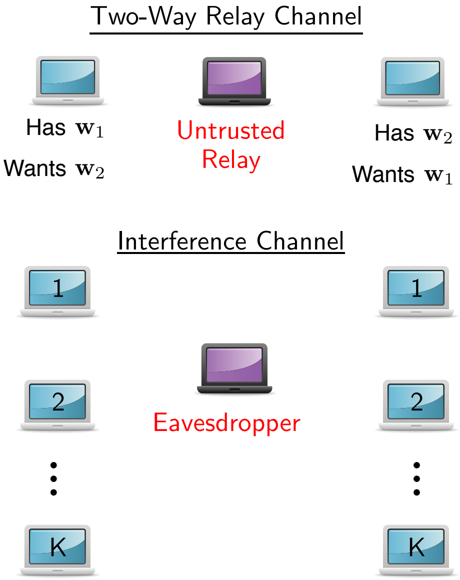
Decoder removes dithers:

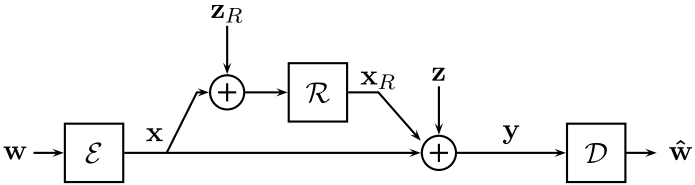
$$\begin{aligned}
&[\alpha \mathbf{y} - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= [\alpha(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{z}) - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= [\mathbf{x}_1 + \mathbf{x}_2 + \alpha(\mathbf{s}_1 + \mathbf{s}_2) - (1-\alpha)(\mathbf{x}_1 + \mathbf{x}_2) + \alpha \mathbf{z}) - \mathbf{d}_1 - \mathbf{d}_2] \bmod \Lambda \\
&= \Big[\mathbf{t}_1 + \mathbf{t}_2 + (1-\alpha)(\mathbf{x}_1 + \mathbf{x}_2) + \alpha \mathbf{z}\Big] \bmod \Lambda
\end{aligned}$$

Select $\alpha = 2P/(2P+N)$ to obtain

$$R_1 + R_2 \leq \left[\frac{1}{2}\log\left(\frac{1}{2} + \frac{P}{N}\right)\right]^+$$

- **He-Yener '09:** Lattice codes are useful for physical-layer secrecy.
- Random i.i.d. codes achieve 0 secure-degrees-of-freedom.
- Basic result: Random lattice codes achieve positive secure-degrees-of-freedom.





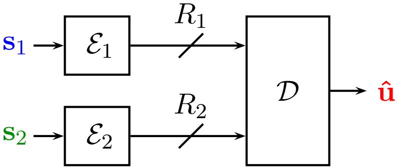
What can we prove with lattice codes for the AWGN relay channel?

- The full decode-and-forward rate can be achieved.
  See **Song-Devroye '10, Nockleby-Aazhang '11.**
- The full compress-and-forward rate can be achieved.
  See **Song-Devroye '11.**

- Correlated Gaussian sources.

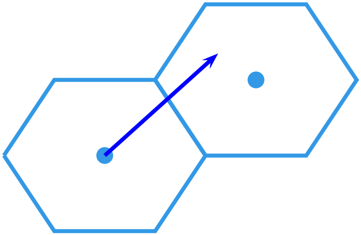
$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Decoder wants the difference.
- Nested lattices are also good for Gaussian source coding.



$$\mathbf{u} = \mathbf{s}_1 - \mathbf{s}_2$$

$$D = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{u}} - \mathbf{u}\|^2$$

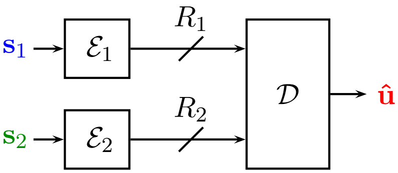


Distributed Source Coding: "Gaussian Körner-Marton Problem"

- Correlated Gaussian sources.

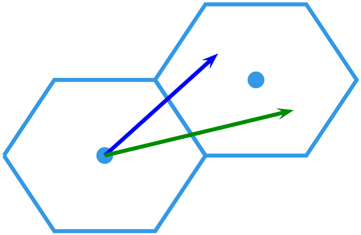
$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Decoder wants the difference.
- Nested lattices are also good for Gaussian source coding.
- **Krithivasan-Pradhan '09:** with high probability, $\mathbf{s}_1$ and $\mathbf{s}_2$ will land near the same coarse lattice point.



$$\mathbf{u} = \mathbf{s}_1 - \mathbf{s}_2$$

$$D = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{u}} - \mathbf{u}\|^2$$

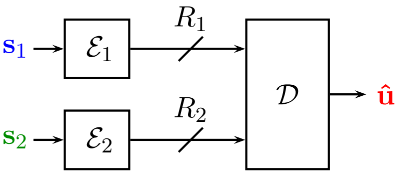


Distributed Source Coding: "Gaussian Körner-Marton Problem"

- Correlated Gaussian sources.

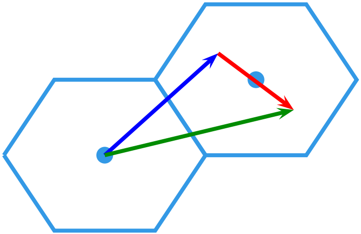
$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Decoder wants the difference.
- Nested lattices are also good for Gaussian source coding.
- **Krithivasan-Pradhan '09:** with high probability, $\mathbf{s}_1$ and $\mathbf{s}_2$ will land near the same coarse lattice point.



$$\mathbf{u} = \mathbf{s}_1 - \mathbf{s}_2$$

$$D = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{u}} - \mathbf{u}\|^2$$

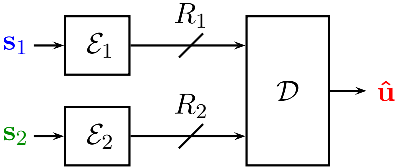


- Correlated Gaussian sources.

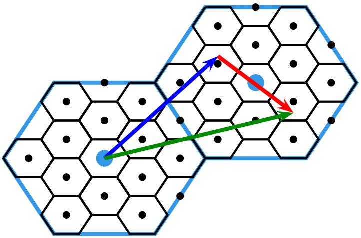
$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Decoder wants the difference.
- Nested lattices are also good for Gaussian source coding.
- **Krithivasan-Pradhan '09:** with high probability, $\mathbf{s}_1$ and $\mathbf{s}_2$ will land near the same coarse lattice point.
- Only need to send:

$$\mathbf{t}_1 = \left[Q_{\Lambda_{\mathsf{FINE}}}(\mathbf{s}_1)\right] \bmod \Lambda$$
$$\mathbf{t}_2 = \left[Q_{\Lambda_{\mathsf{FINE}}}(\mathbf{s}_2)\right] \bmod \Lambda$$



$$\mathbf{u} = \mathbf{s}_1 - \mathbf{s}_2$$
$$D = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{u}} - \mathbf{u}\|^2$$



Three-User Gaussian Distributed Source Coding

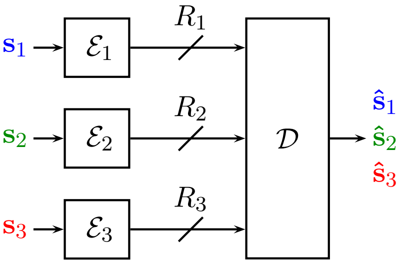
- Correlated Gaussian sources.

$$\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$$

- Third source is the difference:

$$\mathbf{s}_3 = \mathbf{s}_1 - \mathbf{s}_3$$

- Structured codes make new rate points accessible in distributed Gaussian source coding.
- Example: Set $R_1 = 0$ and $R_2 = 0$.
- See **Tavildar-Wagner-Viswanath '10, Krithivasan-Pradhan '09, Maddah-Ali–Tse '10.**



$$D_1 = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{s}}_1 - \mathbf{s}_1\|^2$$
$$D_2 = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{s}}_2 - \mathbf{s}_2\|^2$$
$$D_3 = \frac{1}{n}\mathbb{E}\|\hat{\mathbf{s}}_3 - \mathbf{s}_3\|^2$$

- **Feng-Silva-Kschischang '10** develop practical nested lattice codes that work quite well for blocklengths as small as 100.
- **Hern and Narayanan '10** develop multi-level codes to use fields of size 2^k.
- **Ordentlich and Erez '10** propose mapping by set partitioning to go from binary codewords to higher order constellations.
- Further emerging work includes **Osmane and Belfiore '11**

Concluding Remarks

- Codes with algebraic structure lead to the highest known achievable rates for some communication scenarios of great interest.
- This applies to *source coding*, *channel coding*, and also *joint source-channel coding.*
- We have discussed a set of tools to apply and analyze *random linear* and *random lattice* codes to communication network scenarios.
- However, there is currently no general unified theory of how to generally use algebraic structure in the context of network information theory.

References – Random I.I.D. Codes

- C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- R. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, Inc., 1968.
- I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1982.
- T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.
- A. El Gamal and Y.-H. Kim, *Lecture Notes on Network Information Theory*, 2010, http://arxiv.org/abs/1001.3404.
- R. L. Dobrushin, "Asymptotic optimality of group and systematic codes for some channels," Theory of Probability and its Applications, vol. 8, no. 1, pp. 47–59, 1963.
- R. Ahlswede, "Multi-way communication channels," in *Proc. IEEE Int. Symp. Inf. Theory, Prague*. Publishing House of the Hungarian Academy of Sciences, 1971, pp. 23–52.
- H. Liao, "Multiple access channels," PhD thesis, University of Hawaii, Honolulu, 1972.
- D. Slepian and J. Wolf, "Noiseless Coding of Correlated Information Sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.
- T .M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 226–228, March 1975.
- R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

References – Random Linear Codes

- P. Elias, “Coding for Noisy Channels,” in *IRE Convention Record*, pp. 37–46, 1955.
- R. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, Inc., 1968.
- I. Csiszár, “Linear Codes for Sources and Source Networks: Error Exponents, Universal Coding,” *IEEE Transactions on Information Theory*, vol. 28, no. 4, pp. 585–592, 1982.
- H.-A. Loeliger, “Averaging bounds for lattices and linear codes,” *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.
- R. Zamir, S. Shamai (Shitz), and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.
- S.-Y. R. Li, R. W. Yeung, and N. Cai, “Linear network coding,” *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- R. Koetter and M. Medard, “An algebraic approach to network coding,” *IEEE/ACM Trans. on Netw.*, vol. 11, pp. 782–795, Oct. 2003.
- M. Effros, M. Médard, T. Ho, S. Ray, D. R. Karger, R. Koetter, and B. Hassibi, “Linear network codes: A unified framework for source, channel, and network coding,” in *DIMACS Workshop on Network Information Theory, Piscataway, NJ*, 2003.
- T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, “A random linear network coding approach to multicast,” *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

References – Linear Codes Help

- J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.
- B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
- T. Philosof and R. Zamir, "The rate loss of single-letter characterization: The "dirty" multiple access channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2442–2454, June 2009.
- T. Philosof, R. Zamir, and U. Erez "The capacity region of the binary dirty MAC," *Proc. of Inf. Theory Workshop*, Volos, Greece, June 2009.
- B. Nazer and M. Gastpar, "The case for structured random codes in network capacity theorems," *European Transactions on Telecommunications*, vol. 19, pp. 455–474, June 2008.
- D. Krithivasan and S. S. Pradhan, "Distributed source coding using Abelian group codes: A new achievable rate-distortion region," *IEEE Transactions on Information Theory*, vol. 57, no.3, pp. 1495–1519, March 2011.
- D. Gündüz, O. Simeone, A. J. Goldsmith, H. V. Poor, and S. Shamai; , "Multiple Multicasts With the Help of a Relay," *IEEE Transactions on Information Theory*, vol. 56, no.12, pp. 6142–6158, Dec. 2010
- J. Goseling, M. Gastpar, and J. Weber, "Line and lattice networks under deterministic interference models," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3080–3099, May 2011.

References – Physical-Layer Network Coding

Y. Wu, P. A. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," Microsoft Research, Redmond, WA, Tech. Rep. MSR-TR-2004-78, Aug. 2004.

S. Zhang, S. Liew, and P. Lam, "Hot topic: Physical-layer network coding," in *Proc. ACM Int. Conf. Mobile Comp. Netw.*, Los Angeles, CA, Sep. 2006.

B. Nazer and M. Gastpar, "Computing over multiple-access channels with connections to wireless network coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, Jul. 2006.

P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proc. IEEE Veh. Tech. Conf.*, Melbourne, Australia, May 2006.

S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," *ACM SIGCOMM*, Kyoto, Japan, August 2007.

M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 11, no. 56, pp. 5641–5654, Nov. 2010.

W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within $1/2$ bit," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

W. Nam, S.-Y. Chung, and Y. H. Lee, "Nested Lattice Codes for Gaussian Relay Networks with Interference," in *IEEE Transactions on Information Theory*, Submitted February 2009. http://arxiv.org/abs/0902.2436

I. Maric, A. Goldsmith, and M. Médard, "Analog network coding in the high-SNR regime," in *Proceedings of the IEEE Wireless Network Coding Conference (WiNC 2010)*, (Boston, MA), June 2010.

B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 438–460, March 2011.

S.-C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," in *Physical Communication*, to appear 2011. http://arxiv.org/abs/1105.4261.

References – Relaying

- T. Cover and A. El Gamal, “Capacity Theorems for the Relay Channel,” *IEEE Transactions on Information Theory.*, vol. 25, no. 5, pp. 572–584, 1979.
- G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Transactions on Information Theory*, vol. 51, pp. 3037–3063, September 2005.
- S. Avestimehr, S. Diggavi, and D. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Transactions on Information Theory*, vol. 57, pp. 1872–1905, April 2011.
- S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, “Noisy network coding,” *IEEE Transactions on Information Theory*, vol. 57, pp. 3132–3152, May 2011.
- Y. Song and N. Devroye, “List decoding for nested lattices and applications to relay channels,” in *Proc. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Sep. 2010.
- Y. Song and N. Devroye, “A lattice compress-and-forward strategy for canceling known interference in Gaussian multi-hop channels,” *Conf. on Inf. Sci. and Sys., Baltimore,* March 2011.
- M. Nokleby, B. Aazhang, "Lattice coding over the relay channel,” IEEE Int. Conf. Comm., Kyoto, Japan, June 2011.

References – Lattices

J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups. New York: Springer, 1992.*

R. de Buda, "Some optimal codes have structure," IEEE Journal on Sel. Areas Comm., vol. 7, no. 6, pp. 893–899, Aug. 1989.

T. Linder, C. Schlegel, and K. Zeger, "Corrected proof of de Buda's theorem," IEEE Transactions on Information Theory, vol. 39, no. 5, pp. 1735–1737, Sep. 1993.

G. Poltyrev, "On coding without restrictions for the AWGN channel," IEEE Transactions on Information Theory, vol. 40, no. 2, pp. 409–417, Mar. 1994.

R. Zamir and M. Feder, "On lattice quantization noise," IEEE Transactions on Information Theory, *vol. 42, no. 4, pp. 1152–1159, July 1996.*

H.-A. Loeliger, "Averaging bounds for lattices and linear codes," IEEE Transactions on Information Theory, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," IEEE Transactions on Information Theory, vol. 44, no. 1, pp. 273–278, Jan. 1998.

G. Forney, M. Trott, and S.-Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," IEEE Transactions on Information Theory, vol. 46, no. 3, pp. 820–850, May 2000.

R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," IEEE Transactions on Information Theory, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+SNR)$ *on the AWGN channel with lattice encoding and decoding," IEEE Transactions on Information Theory, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.*

U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," IEEE Transactions on Information Theory, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

R. Zamir, "Lattices are everywhere," in Proc. Workshop Inf. Theory Applications, La Jolla, CA, Feb. 2009.

References – Lattices Help: Interference Channels

G. Bresler and A. Parekh and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Transactions on Information Theory.*, vol. 56, no. 9, pp. 4566-4592, Sep. 2010.

S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K-user Gaussian very strong interference channels," in *GLOBECOM*, Monticello, IL, Sep. 2008.

S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai, "A layered lattice coding scheme for a class of three user Gaussian interference channels," in *Proc. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Sep. 2008.

R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Transactions on Information Theory*, vol. 55, pp. 4932–4946, November 2009.

A. S. Motahari, S. O. Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Transactions on Information Theory*, Submitted November 2009. See http://arxiv.org/abs/0908.2282.

S. Vishwanath and S. A. Jafar, "Generalized degrees of freedom of the symmetric Gaussian K-User interference channel," *IEEE Transactions on Information Theory*, vol.56, no.7, pp.3297-3303, July 2010.

B. Bandemer and A. El Gamal, Interference decoding for deterministic channels, *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2966–2975, May 2011

A. Jafarian and S. Vishwanath, "Gaussian interference networks: Lattice alignment," Proc. of Inf. Theory Workshop, Cairo, Egypt, January 2010.

O. Ordentlich and U. Erez, "Interference Alignment at Finite SNR for Time-Invariant Channels," Submitted to *IEEE Transactions on Information Theory* April 2011. http://arxiv.org/abs/1104.5456

H. Huang and V. K. N. Lau, Y. Du and S. Liu, "Robust lattice alignment for K-user MIMO interference channels with imperfect channel knowledge," *IEEE Transactions on Information Theory,* vol. 59, no. 7, pp. 3315–3325, July 2011.

References – Lattices Help: Compute-and-Forward

- B. Nazer and M. Gastpar, “Compute-and-forward: Harnessing interference through structured codes,” *IEEE Transactions on Information Theory*, to appear October 2011. http://arxiv.org/abs/0908.2119
- M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, “Joint physical layer coding and network coding for bidirectional relaying,” *IEEE Transactions on Information Theory*, vol. 11, no. 56, pp. 5641–5654, Nov. 2010.
- W. Nam, S.-Y. Chung, and Y. H. Lee, “Capacity of the Gaussian two-way relay channel to within $1/2$ bit,” *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.
- U. Niesen and P. Whiting, “The Degrees of Freedom of Compute-and-Forward,” arXiv:1101.2182v1 [cs.IT], 2011.
- L. Ong, C. M. Kellett, and S. J. Johnson, “Capacity theorems for the AWGN multi-way relay channel,” in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, Jun. 2010.
- C. Feng, D. Silva, and F. Kschischang, “An algebraic approach to physical-layer network coding,” in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, June 2010.
- B. Hern and K. Narayanan, “Multilevel coding schemes for compute-and-forward,” see *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, June 2011.
- O. Ordentlich, J. Zhan, U. Erez, B. Nazer, and M. Gastpar. “Practical Code Design for Compute-and-Forward”, *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, June 2011.
- A. Osmane and J.-C. Belfiore, “The Compute-and-Forward Protocol: Implementation and Practical Aspects,” Submitted to IEEE Communications Letters 2011. http://arxiv.org/abs/1107.0300
- M. Nokleby and B. Aazhang, ”Cooperative computation in wireless networks,” *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, June 2011.
- B. Nazer, “Compute-and-Forward: Improved Successive Cancellation,” To be submitted.

References – Lattices Help: Cellular Uplink, “Dirty” MAC, Secrecy

- A. D. Wyner, Shannon-theoretic approach to a Gaussian cellular multiple-access channel, *IEEE Transactions on Information Theory.*, vol. 40, pp. 1713 1727, Nov. 1994.
- A. Sanderovich, O. Somekh, H. V. Poor, S. and Shamai (Shitz), “Uplink macro diversity of limited backhaul cellular network,” *IEEE Transactions on Information Theory*, vo. 55, no. 8, pp. 3457–3478. August 2009.
- A. Sanderovich, M. Peleg, and S. Shamai (Shitz), “Scaling laws in decentralized processing of interfered Gaussian channels,” in *Proc. Int. Zurich Seminar Comm.*, Zurich, Switzerland, March 2008.
- B. Nazer, A. Sanderovich, M. Gastpar, and S. Shamai, “Structured Superposition for Backhaul Constrained Cellular Uplink,” in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun. 2009.
- U. Erez, S. Shamai, and R. Zamir, “Capacity and lattice strategies for canceling known interference,” *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3820–3833, November 2005.
- T. Philosof and R. Zamir, “The rate loss of single-letter characterization: The “dirty” multiple access channel,” *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2442–2454, June 2009.
- T. Philosof, R. Zamir, and U. Erez “The capacity region of the binary dirty MAC,” *Proc. of Inf. Theory Workshop*, Volos, Greece, June 2009.
- T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, “Lattice strategies for the dirty multiple access channel,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5006–5035, August 2011.
- X. He and A. Yener, “Providing secrecy with structured codes: Tools and applications to two-user Gaussian channels,” *Submitted to IEEE Transactions on Information Theory*, Jul. 2009, see http://arxiv.org/abs/0907.5388
- S. Agrawal and S. Vishwanath, “On the secrecy rate of interference networks using structured codes,” in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun. 2009.
- X. He and A. Yener, “The Gaussian many-to-one interference channel with confidential messagse,” *Submitted to IEEE Transactions on Information Theory*, Apr. 2010, see http://arxiv.org/abs/1005.0624

References – Lattices Help: Distributed Source Coding

- D. Krithivasan and S. S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function," *IEEE Transactions on Information Theory*, vol. 55, pp. 5268–5651, December 2009.
- S. Tavildar and P. Viswanath and A. B. Wagner, "The Gaussian Many-Help-One Distributed Source Coding Problem," *IEEE Transactions on Information Theory.*, vol. 56, no. 1, pp. 564–571, 2010.
- A. B. Wagner, On distributed compression of linear functions, *IEEE Transactions on Information Theory.*, Vol. 57, No. 1, pp. 79-94, 2011.
- D. Krithivasan and S. S. Pradhan, "Distributed source coding using Abelian group codes: A new achievable rate-distortion region," *IEEE Transactions on Information Theory*, vol. 57, no.3, pp. 1495–1519, March 2011.
- M. A. Maddah-Ali, and D. N. C. Tse, "Interference neutralization in distributed lossy source coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, Jun. 2009.