Reliable Uncoded Communication in the SIMO MAC with Efficient Decoding

Mainak Chowdhury, Andrea Goldsmith, Tsachy Weissman

Department of Electrical Engineering, Stanford University

July 1, 2013, Broadcom Corp., Sunnyvale
# Table of contents

1. Introduction
2. Our work
3. Pictures and proofs
4. Extensions
5. Conclusions
Motivation: Uplink communication in large networks

Often limited by transmitter side constraints
- Power/energy requirements
- Delay constraints
- Processing capability

Examples
- Cellular uplink
- Sensor networks
- Low power radio
System model

- $H_{ij} \sim \mathcal{N}(0, 1), \nu_i \sim \mathcal{N}(0, \sigma^2)$
- $y = Hx + \nu$
- $N_U$ users, each with 1 antenna
- $N_R$ receive antennas at base station
No CSI at transmitters, perfect CSI at receiver
Warmup question

What is the number of independent streams that can be simultaneously supported by the channel?
Warmup question

What is the number of independent streams that can be simultaneously supported by the channel?

Asymptotic regime of high SNR

- \( \min(N_R, N_U) \) streams
- Theoretically achieved using coding across time
- Transmitter channel knowledge usually important
Warmup question

What is the number of independent streams that can be simultaneously supported by the channel?

**Asymptotic regime of high SNR**
- \( \min(N_R, N_U) \) streams
- Theoretically achieved using coding across time
- Transmitter channel knowledge usually important

**A different (practical) operating regime**
- Low SNR, one-shot communication, symbols from constellation set
- Does the maximum number of streams still stay the same?
Introduction

Warmup question

What is the number of independent streams that can be simultaneously supported by the channel?

Asymptotic regime of high SNR

- \( \min(N_R, N_U) \) streams
- Theoretically achieved using coding across time
- Transmitter channel knowledge usually important

A different (practical) operating regime

- Low SNR, one-shot communication, symbols from constellation set
- Does the maximum number of streams still stay the same?

Not necessarily

Different operating regimes can yield totally different intuition/insights!
We consider the regime of large $N_U$, fixed rate $R$ and QoS requirement $P_e$

**We ask**
- How many receive antennas are required?
- What kind of decoders can give reliable decoding?
- How useful is coding in large systems?

**We show**
As $N_U \to \infty$, $P_e \to 0$ for any fixed ratio $\frac{NR}{NU}$, using either of the following:
- Maximum likelihood decoder
- Interval Search and Quantize (ISQ) decoder (deterministic and randomized)
Our setting

We consider the regime of large $N_U$, fixed rate $R$ and QoS requirement $P_e$

We ask

- How many receive antennas are required?
- What kind of decoders can give reliable decoding?
- How useful is coding in large systems?

We show

As $N_U \to \infty$, $P_e \to 0$ for any fixed ratio $\frac{N_R}{N_U}$, using either of the following:

- Maximum likelihood decoder
- Interval Search and Quantize (ISQ) decoder (deterministic and randomized)

Number of data streams supported can be much larger than $\min(N_R, N_U)$!
Table of contents

1. Introduction
2. Our work
3. Pictures and proofs
4. Extensions
5. Conclusions
### Decoders

#### Decoder 1: Maximum likelihood (ML) decoder
- Compute \( \hat{x} = \arg\min_{x \in \{-1, +1\}^{N_U}} \|y - Hx\| \)
- Complexity exponential in \( N_U \)

#### Decoder 2: Interval Search and Quantize (ISQ) decoder
- Compute \( \hat{x} = \text{sign}(\arg\min_{x \in [-1, +1]^{N_U}} \|y - Hx\|) \)
- \( \text{sign}(x) \) is elementwise application of \( \text{sign}(\cdot) \)
- If solution is non unique (for \( N_R < N_U \)), return randomized projection onto solution space
- Complexity of decoding and random projection is polynomial in \( N_U \)
Let $x_0 = -1$ be transmitted.

**Theorem**

For $N_R = \alpha N_U$ for any $\alpha > 0$ and a large enough $n_0$, there exists a $c > 0$ such that the probability of block error goes to zero exponentially fast with $N_U$, i.e.

$$P(\hat{x} \neq x_0) \leq 2^{-cN_U} \text{ for all } N_U > n_0$$

**In other words**

- Equal rate transmission
- Fixed $\frac{N_R}{N_U}$ ratio
- Error probability $P_e \to 0$ with $N_U \to \infty$
Reliability in communication: (ML decoder)

Let $x_0 = -1$ be transmitted

**Theorem**

For $N_R = \alpha N_U$ for any $\alpha > 0$ and a large enough $n_0$, there exists a $c > 0$ such that the probability of block error goes to zero exponentially fast with $N_U$, i.e.

$$P(\hat{x} \neq x_0) \leq 2^{-cN_U} \text{ for all } N_U > n_0$$

**In other words**

- Equal rate transmission
- Fixed $\frac{N_R}{N_U}$ ratio
- Error probability $P_e \to 0$ with $N_U \to \infty$

*This is a combination of receiver diversity and spatial multiplexing*
Reliability in communication: (ISQ decoder)

**Theorem**

For $N_R = \alpha N_U$ for any $\alpha, k > 0$ and a large enough $n_0$, there exists a $c > 0$ such that the probability of $kN_U$ symbol errors goes to zero super-exponentially fast with $N_U$, i.e.

$$P(|\hat{x} - x_0|_0 > kN_U) \leq 2^{-cN_U \log N_U} \text{ for all } N_U > n_0$$

**In other words**

- Per user symbol error goes to zero
- Net rate of decay of error probability is at best polynomial
- Error even without noise (self interference)!
Minimum number of required Rx antennas per user

Let us consider coding across time:

- Every user needs a rate of 1 bit per channel use on average.
- Every user has average power 1.
- Sum rate required from the system is $N_U$.
- Equal rate capacity of the system is $\frac{1}{2} \log |(I + \frac{HH^T}{\sigma^2})|$.

Thus, the lowest $N_R$ to support reliable transmissions is:

$$N_R \geq 2 N_U^2 + \log N_U$$

Smallest $N_R$ ratio achievable with coding is:

$$N_R \geq \frac{1}{2} \log N_U + 2$$

Can we achieve similar ratios for reliable uncoded systems also?
Minimum number of required Rx antennas per user

Let us consider coding across time
- Every user needs a rate of 1 bit per channel use on average
- Every user has average power 1
- Sum rate required from the system is $N_U$
- Equal rate capacity of the system is $\frac{1}{2}N_R \log N_U + cN_R$

Thus, the lowest $N_R$ to support reliable transmissions is

\[
N_R \geq \frac{2N_U}{2c + \log N_U}
\]
Minimum number of required Rx antennas per user

Let us consider coding across time

- Every user needs a rate of 1 bit per channel use on average
- Every user has average power 1
- Sum rate required from the system is \( N_U \)
- Equal rate capacity of the system is \( \frac{1}{2} N_R \log N_U + c N_R \)

Thus, the lowest \( N_R \) to support reliable transmissions is

\[
N_R \geq \frac{2N_U}{2c + \log N_U}
\]

Smallest \( \frac{N_R}{N_U} \) ratio achievable with coding is

\[
\frac{N_R}{N_U} \geq \frac{2}{\log N_U + 2c}
\]
Minimum number of required Rx antennas per user

Let us consider coding across time

- Every user needs a rate of 1 bit per channel use on average
- Every user has average power 1
- Sum rate required from the system is $N_U$
- Equal rate capacity of the system is $\frac{1}{2} N_R \log N_U + c N_R$

Thus, the lowest $N_R$ to support reliable transmissions is

$$N_R \geq \frac{2N_U}{2c + \log N_U}$$

Smallest $\frac{N_R}{N_U}$ ratio achievable with coding is

$$\frac{N_R}{N_U} \geq \frac{2}{\log N_U + 2c}$$

Can we achieve similar ratios for reliable uncoded systems also?
Minimum number of required Rx antennas per user: Results for ML decoder

Theorem

For \( \frac{N_R}{N_U} = \frac{2+\epsilon}{\log N_U} \) for any \( \epsilon > 0 \) and a large enough \( n_0 \), there exists a \( c > 0 \) such that the probability of block error with the maximum likelihood decoder goes to zero exponentially fast with \( N_U \), i.e.

\[
P(\hat{x} \neq x) \leq 2^{-cN_U \log N_U} \text{ for all } N_U > n_0
\]

In other words

- Coding does not reduce the number of required antennas per user for 1 bit per channel use
- Scaling behaviour of minimum \( \frac{N_R}{N_U} \) is \( \Theta\left(\frac{1}{\log N_U}\right) \)
Results

Minimum number of required Rx antennas per user: Results for ISQ decoder

**Theorem**

For \( \frac{N_R}{N_U} = \alpha \) for any \( \epsilon > 0, k > 0 \) and a large enough \( n_0 \), there exists a \( c > 0 \) such that the probability of \( kN_U \) symbol errors with the randomized ISQ decoder goes to zero exponentially fast with \( N_U \), i.e.

\[
P_{k \neq x} \leq 2^{-cN_U \log N_U} \text{ for all } N_U > n_0
\]

**In other words**

- Per-user reliability still holds with \( N_U \to \infty \) (😊)
- Decay of per-user error probability is at most polynomial (not exponential 😞)
Table of contents

1 Introduction
2 Our work
3 Pictures and proofs
4 Extensions
5 Conclusions
Proofs

Decision region (transmitter space): ML decoder,

\[ N_R = 2, \; N_U = 3 \]

Possible transmitted codewords

Uncoded transmission in MAC channels
Decision region (receiver space): ML decoder, 
\[ N_R = 2, \ N_U = 3 \]

Possible transmitted codewords (as seen by the receiver)
Decision region (transmitter space): ISQ decoder,

\[ N_R = 2, N_U = 3 \]

Regions for transmitted codewords
Decision region (receiver space): ISQ decoder,
\[ N_R = 2, N_U = 3 \]

Regions for some selected transmitted codewords (as seen by the receiver)
Let's say $x_0$ is transmitted.

$$P_e \leq \sum_{x \neq x_0} P(\hat{x} = x)$$
Proof outline (ML): Union bound

Let’s say $x_0$ is transmitted.

$$P_e \leq \sum_{x \neq x_0} P(\hat{x} = x)$$

**Idea**

Group wrong codewords by the number of positions in which they differ from $x_0$

Thus

$$P_e \leq \sum_{i=1}^{NU} \sum_{x : d(x, x_0) = i} P(\hat{x} = x),$$

where $d(\cdot, \cdot)$ computes the hamming distance between its two arguments.
Proof outline (ML): Computing pairwise error probabilities

Let \( x_i \) be a codeword with \( i \) positions different from \( x_0 \). Then

\[
P(\hat{x} = x_i) = Q\left( \frac{\|H(x_i - x_0)\|}{2\sigma} \right)
\]

\[
= Q\left( \frac{\|\sum_{j=1}^{i} 2h_{b(j)}\|}{2\sigma} \right)
\]

\( b(j) \) is the \( j^{th} \) position where \( x_i \) and \( x_0 \) differ,

\( h_b \) is the \( b^{th} \) column of \( H \)

\[
\leq \frac{1}{2} \exp \left( -\frac{\|\sum_{j=1}^{i} 2h_{b(j)}\|^2}{8\sigma^2} \right)
\]
Proof outline (ML): Average pairwise error probability

Idea

Average over channel realizations

\[ \mathbb{E}_H(P(\hat{x} = x_i)) \leq \mathbb{E}_H \left( \frac{1}{2} \exp \left( - \frac{\| \sum_{j=1}^{i} 2h_{b(j)} \|^2}{8\sigma^2} \right) \right) \]

This is just the moment generating function of an appropriately scaled chi squared distribution.

Closed form expression

\[ \mathbb{E}_H(P(\hat{x} = x_i)) \leq \frac{1}{2} \left( 1 + \frac{i}{\sigma^2} \right)^{-\frac{NR}{2}} \]
Proof outline (ML): Average pairwise error probability

Idea
Average over channel realizations

\[ \mathbb{E}_H(P(\hat{x} = x_i)) \leq \mathbb{E}_H\left( \frac{1}{2} \exp \left( - \frac{\| \sum_{j=1}^{i} 2h_{b(j)} \|^2}{8\sigma^2} \right) \right) \]

This is just the moment generating function of an appropriately scaled chi squared distribution.

Closed form expression

\[ \mathbb{E}_H(P(\hat{x} = x_i)) \leq \frac{1}{2} \left( 1 + \frac{i}{\sigma^2} \right)^{-\frac{NR}{2}} \]

Depends only on \( i \), i.e. number of columns, not on particular columns considered
Proof outline (ML): Upper bound on total error probability

Let

$$E_H(P(\hat{x} = x_i)) = P_{e,i},$$

since it is independent of $x_i$. Then the total error probability is

$$P_e \leq \sum_{i=1}^{N_U} \sum_{x: d(x, x_0) = i} P(\hat{x} = x) = \sum_{i=1}^{N_U} \frac{1}{2} \binom{N_U}{i} P_{e,i}$$

$$\leq \sum_{i=1}^{N_U} \frac{1}{2} \left( \frac{N_U}{i} \right) \left( 1 + \frac{i}{\sigma^2} \right)^{-\frac{N_R}{2}}$$

$$\leq N_U \max_{i \in \{1, \ldots, N_U\}} \frac{1}{2} \left( \frac{N_U}{i} \right) \left( 1 + \frac{i}{\sigma^2} \right)^{-\frac{N_R}{2}}$$

$$\leq \frac{N_U}{2} \max_{i \in \{1, \ldots, N_U\}} 2^{N_U H_2 \left( \frac{i}{N_U} \right)} \left( 1 + \frac{i}{\sigma^2} \right)^{-\frac{N_R}{2}}$$
Proof outline (ML): Asymptotic limits

Define

\[ \alpha = \frac{N_R}{N_U} \]

Decoding error probability

\[ P_e \leq 2^{N_U g(N_U)} \]

where

\[ g(n) \triangleq \max_{1 \leq i \leq n} H_2 \left( \frac{i}{n} \right) - \frac{\alpha}{2} \log \left( 1 + \frac{i}{\sigma^2} \right) + \frac{\log n - \log 2}{n} \]

\[ \triangleq \max_{1 \leq i \leq n} g(i, n) \]
Proof outline (ML): Behaviour of $g(n)$, constant $\alpha$

$\alpha = 0.3, \sigma^2 = 0.5, n = 50, g(n) = \max_{1 \leq i \leq n} g(i, n) = 0.26$
Proof outline (ML): Behaviour of $g(n)$, constant $\alpha$

$\alpha = 0.3$, $\sigma^2 = 0.5$, $n = 100$, $g(n) = \max_{1 \leq i \leq n} g(i, n) = 0.08$
Proof outline (ML): Behaviour of $g(n)$, constant $\alpha$

\[ \alpha = 0.3, \sigma^2 = 0.5, n = 200, g(n) = \max_{1 \leq i \leq n} g(i, n) = -0.1 \]
Proof outline (ML): Behaviour of $g(n)$, constant $\alpha$

$\alpha = 0.3, \sigma^2 = 0.5, n = 700, g(n) = \max_{1 \leq i \leq n} g(i, n) = -0.37$
Proof outline (ML): Behaviour of $g(n)$, constant $\alpha$

Message

For fixed $\alpha > 0$ and large enough $n$,

$$g(n) \leq c(\alpha) < 0$$
Proof outline (ML): Behaviour of $g(n)$, $\alpha = \frac{2+\epsilon}{\log n}$

$\epsilon = 0.1, \sigma^2 = 0.5, n = 100, g(n) = \max_{1 \leq i \leq n} g(i, n) = 0.03$
Proof outline (ML): Behaviour of $g(n)$, $\alpha = \frac{2+\epsilon}{\log n}$

$\epsilon = 0.1, \sigma^2 = 0.5, n = 300, g(n) = \max_{1 \leq i \leq n} g(i, n) = -0.01$
Proof outline (ML): Behaviour of $g(n)$, $\alpha = \frac{2+\epsilon}{\log n}$

Message

For fixed $\alpha > 0$ and large enough $n$,

$$g(n) \leq c(\epsilon) < 0,$$

$c(\epsilon)$ goes down as $\Theta\left( -\frac{1}{\log n} \right)$

Thus

$$P_e \leq 2^{\frac{-cN_U}{\log N_U}}$$
Proof outline (ISQ): $((\epsilon, \delta))$ grid

We look at $P_{e,k}$, which is the probability of error in $kN_U$ symbols.

**Definition**

For any $\epsilon$, the $(\epsilon, \delta)$ grid $G_{\epsilon,\delta}$ is the following set

$$\{x : x_i \mod \epsilon = \delta_i, |x_i| < 1\}.$$ 

**Idea**

- Look at decoder $\hat{x} = \text{sign} (\text{argmin}_{x \in G_{\epsilon,\delta}} \|y - Hx\|)$
- Using techniques as in the previous proof,

$$P_{e,k,\epsilon,\delta} \leq \left(\frac{1}{\epsilon}\right)^{N_U} 2^{-\tilde{c}N_U \log N_U}$$

- Relate the grid error probability to the ISQ decoder error probability
Proof outline (ISQ): Grid versus interval performances

Lemma

The minimizer over $G_{\epsilon,\delta}$ will w.p. 1 be “close” to at least one solution of the ISQ decoder, in the following sense

$$||\hat{x}_{\epsilon,\delta} - \hat{x}_{ISQ}||_{\infty} \leq \epsilon.$$ 

Plausibility arguments and implications

- Easier to visualize for $N_R \geq N_U$
- For $N_R < N_U$, the same results continue to hold on average for any distribution $f(\delta)$ on the offset $\delta$
- Computationally feasible to sample a vector “close” to $\hat{x}_{\epsilon,\delta} \sim f(\delta)$
- $\hat{x}_{\epsilon,\delta}$ has “close” to zero entries in at most sublinear positions
Proof outline (ISQ) (contd.)

\[ P_{e,k} \leq 2^{-\tilde{C} N_U \log N_U}, \]

i.e. ISQ estimate differs from \( x_0 \) in at most a sublinear number of symbols.
Table of contents

1 Introduction
2 Our work
3 Pictures and proofs
4 Extensions
5 Conclusions
So far discussion focused on

- BPSK constellation: Each user transmits from \{-1, +1\}
- Gaussian channel statistics (Rayleigh fading)
Pairwise error probability can be bounded by

\[ P(\hat{x} = x_i) \leq \frac{1}{2} \exp \left( - \frac{\| \sum_{j=1}^{i} d h_{b(j)} \|^2}{8\sigma^2} \right), \]

where \( d \) is the minimum distance of the constellation i.e.

\[ d = \min_{x \in C, y \in C, x \neq y} \| x - y \| \]

In general loose

Same scaling \( \Theta \left( \frac{1}{\log N_U} \right) \) for the number of receiver antennas per user
Arbitrary fading statistics

Idea

Central limit theorem and rate of convergence of distributions!

Specifically

- Error due to $i$ mismatches depends on the statistics of $\sum_{j=1}^{i} h_b(j)$
- Berry Esseen bound guarantees $\Theta\left(\frac{1}{\sqrt{n}}\right)$ convergence, i.e.

$$\sup_{x} |F_n(x) - F_g(x)| \leq \frac{\tilde{C}}{\sqrt{n}}$$

- Can be shown that

$$P_{e,i} \leq C i^{-\frac{NR}{2}} ,$$

for all $i \geq i_0$

Thus

Asymptotic behaviour of $P_{e,i}$ in $i$ same as that for Gaussian fading
Extensions

Arbitrary fading statistics

Idea

Central limit theorem and rate of convergence of distributions!

Specifically

- Error due to $i$ mismatches depends on the statistics of $\sum_{j=1}^{i} h_b(j)$
- Berry Esseen bound guarantees $\Theta\left(\frac{1}{\sqrt{n}}\right)$ convergence, i.e.

$$\sup_x |F_n(x) - F_g(x)| \leq \frac{\tilde{C}}{\sqrt{n}}$$

- Can be shown that

$$P_{e,i} \leq C i^{-\frac{N_R}{2}}$$

for all $i \geq i_0$

But $P_e = \sum_{i=1}^{i_0} P_{e,i} + \sum_{i=i_0+1}^{N_U} P_{e,i}$

Terms with less than $i_0$ mismatches $\to 0$ with large $N_R$
Table of contents

1 Introduction

2 Our work

3 Pictures and proofs

4 Extensions

5 Conclusions
Key insights

- Sufficient degrees of freedom already present in large systems
- Receiver diversity allows reliable communication with efficient decoders
- Positive rate possible for every user without coding
- **Ongoing work:** Imperfect CSI, correlation patterns (in channel or symbols)
Thank you for your attention
Questions/Comments?