List decodability of randomly punctured codes

Mary Wootters

January 17, 2014
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
   - Argument from [Rudra, W. 2013]
   - Argument from [W. 2013]

4 Conclusion
1 Background
- Coding Theory and List Decoding
- Reed-Solomon Codes

2 Results

3 Proof ideas
- Outline
- Argument from [Rudra, W. 2013]
- Argument from [W. 2013]

4 Conclusion
1 Background
   • Coding Theory and List Decoding
     ○ Reed-Solomon Codes

2 Results

3 Proof ideas
   • Outline
   • Argument from [Rudra, W. 2013]
   • Argument from [W. 2013]

4 Conclusion
Error correcting codes

Alice

Noisy channel

Bob
Error correcting codes

Let $w \in \mathbb{F}_q^n$ be the received word with an error rate $\rho$. For a given $x \in \mathbb{F}_q^k$, the decoder maps it to a codeword $c \in \mathbb{F}_q^n$. The noisy channel connects Alice to Bob.
Error correcting codes

\[ x \in \mathbb{F}_q^k \iff c \in \mathbb{F}_q^n \]

message

codeword

received word

Noisy channel

\[ w \in \mathbb{F}_q^n = c + \text{noise} \]

error rate is \( \rho \)
Error correcting codes

$w \in \mathbb{F}_q^n = c + \text{noise}$

message

$x \in \mathbb{F}_q^k \mapsto c \in \mathbb{F}_q^n$

received word

codeword

Noisy channel

Alice

Bob

Mary Wootters (University of Michigan)  List decodability of RS codes  January 17, 2014 5 / 50
Error correcting codes

\[ w \in \mathbb{F}_q^n = c + \text{noise} \]

message

\[ x \in \mathbb{F}_q^k \mapsto c \in \mathbb{F}_q^n \]

codeword

received word

error rate is \( \rho \)

Alice

Noisy channel

Bob

Mary Wootters (University of Michigan)
Error correcting codes

\[ x \in \mathbb{F}_q^k \mapsto c \in \mathbb{F}_q^n \]

message
codeword
received word

\[ w \in \mathbb{F}_q^n = c + \text{noise} \]

error rate is \( \rho \)

Noisy channel

x? (or c?)

Alice

Bob
Combinatorially

**distance**: $d(x, y)$ is the fraction of symbols on which $x, y$ disagree.
Combinatorially

**distance**: $d(x, y)$ is the fraction of symbols on which $x, y$ disagree.

“Attack at noon.”

“Run away!”

“Attack at dawn.”

“It’s cool, they’re on our side.”
Combinatorially, the distance $d(x, y)$ is the fraction of symbols on which $x, y$ disagree.
Combinatorially

**distance:** \( d(x, y) \) is the fraction of symbols on which \( x \), \( y \) disagree.
Combinatorially

\[ d(x, y) \] is the fraction of symbols on which \( x, y \) disagree.

“Attack at noon.”

“Attack at dawn.”

“Run away!”

“It’s cool, they’re on our side.”
The trade-off

Want:

- To be able to handle large error rates $\rho$. (big distance).
- To be able to send lots of information. (big rate).

$$\text{rate} = \frac{k}{n} = \frac{\log_q(|C|)}{n}$$

How many symbols Alice wanted to send

How many symbols Alice actually sent
Limitations of unique decoding

Cannot decode when $\rho$ is bigger than half the distance.
List decoding

Bob does know something, even if $\rho$ is much larger.
List decodable codes

Alice

Bob

message

$\mathbf{x} \in \mathbb{F}^k_q \leftrightarrow \mathbf{c} \in \mathbb{F}^n_q$

codeword

received word

$\mathbf{w} \in \mathbb{F}^n_q = \mathbf{c} + \text{noise}$

error rate is $\rho$

Small list $S$
so that $\mathbf{c} \in S$?

Noisy channel

Mary Wootters (University of Michigan)
List-Decoding Capacity Theorem

For large $q$:

- Can list-decode with error rate $\rho = 1 - \varepsilon$ with rate $\varepsilon$.
  - To send $\varepsilon$ information, you can corrupt at most $1 - \varepsilon$.
  - This is attainable.
List-Decoding Capacity Theorem

For large $q$:

- Can list-decode with error rate $\rho = 1 - \varepsilon$ with rate $\varepsilon$.
  - To send $\varepsilon$ information, you can corrupt at most $1 - \varepsilon$.
  - This is attainable.

- To uniquely decode, the error rate needs to be $\rho \leq \frac{1-\varepsilon}{2}$.
1 Background
- Coding Theory and List Decoding
- Reed-Solomon Codes

2 Results

3 Proof ideas
- Outline
- Argument from [Rudra, W. 2013]
- Argument from [W. 2013]

4 Conclusion
Reed-Solomon Codes

**Message:** \((m_0, \ldots, m_{k-1})\)

\[ f(X) = m_0 + m_1 X + m_2 X^2 + \cdots + m_{k-1} X^{k-1} \]

**Codeword:**

\[ (f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)) \]

- \(\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q\) are evaluation points.
- Rate is \(\frac{k}{n}\).
- Distance is \(1 - \frac{k-1}{n}\).
List Decodability and Reed-Solomon Codes

- **Johnson Bound:** Good distance $\Rightarrow$ Good list-decodability.

  RS codes of rate $\varepsilon^2$ can handle error rate $\rho = 1 - \varepsilon$.

- **Guruswami-Sudan:** Can list-decode RS codes *efficiently* up to the Johnson bound.
List Decodability and Reed-Solomon Codes

- **Johnson Bound**: Good distance $\Rightarrow$ Good list-decodability.

  RS codes of rate $\varepsilon^2$ can handle error rate $\rho = 1 - \varepsilon$.

- **Guruswami-Sudan**: Can list-decode RS codes *efficiently* up to the Johnson bound.
Since then

Best rate to which we can efficiently list-decode with $\rho = 1 - \varepsilon$?

1999  Guruswami-Sudan: RS codes get $\varepsilon^2$.
2005  Parvaresh-Vardy: PV codes beat the Johnson bound.
2008  Guruswami-Rudra: Folded RS codes get $\varepsilon$.
2012  Kopparty: Multiplicity codes get $\varepsilon$.
2012  Guruswami-Xing: Folded AG codes get $\varepsilon$.
2013  Guruswami-Wang: Derivative codes get $\varepsilon$.
2013  Guruswami-Xing: Nice subcodes of RS, AG codes get $\varepsilon$. 
Motivating Question

Are there Reed-Solomon codes which can be list-decoded beyond the Johnson bound?
Motivating Question

Are there Reed-Solomon codes which can be list-decoded beyond the Johnson bound?

- First asked by [Guruswami’99]
- Formulations also in [Guruswami’04, Rudra’07, Vadhan’12]
Motivating Question

Are there Reed-Solomon codes which can be list-decoded beyond the Johnson bound?

Why should we care?

- Reed-Solomon codes are important.
- Complexity theory applications.
- What structure (or lack of structure) is necessary to get beyond the Johnson bound?
Motivating Question

Are there Reed-Solomon codes which can be list-decoded beyond the Johnson bound?

Reasons to think \textbf{yes}:

- Reed-Solomon codes are really nice.
- Other constructions are variants on RS codes.

Reasons to think \textbf{no}:

- When \textit{all} evaluation points are used, the answer is \textbf{no}, $R = \Omega(\varepsilon^{2-\gamma})$ [Ben-Sasson, Kopparty, Radhakrishnan 2010].
- The Johnson bound is the right answer for \textit{list recovery} [Guruswami, Rudra 2006].
Motivating Question

Are there Reed-Solomon codes which can be list-decoded beyond the Johnson bound?

This talk:

- The answer is yes.
- In fact, “most” RS codes are list-decodable beyond the Johnson bound.
- Techniques are more general:
  - Random linear codes.
  - Randomly punctured codes.
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
   - Argument from [Rudra, W. 2013]
   - Argument from [W. 2013]

4 Conclusion
Theorem

Suppose \( q \geq \frac{k}{\varepsilon^2} \), and let \( C \) be the Reed-Solomon code over \( \mathbb{F}_q \) with:

- evaluation points \( \alpha_1, \ldots, \alpha_n \) chosen uniformly at random,
- rate

\[
R = \frac{k}{n} = \frac{C\varepsilon}{\log(q) \log^5(1/\varepsilon)}.
\]

Then w.h.p, \( C \) is list-decodable up to error rate

\[
\rho = 1 - \varepsilon
\]

with list sizes

\[
L = O \left( \frac{1}{\varepsilon} \right).
\]
More Generally

\[ N = \left| C_0 \right| \]

- Start with a “decently” list decodable code \( C_0 \) with bad rate.
More Generally

\[ N = |C_0| \]

- Start with a “decently” list decodable code \( C_0 \) with bad rate.
- Choose \( n \) random positions.
More Generally

- Start with a “decently” list decodable code $C_0$ with bad rate.
- Choose $n$ random positions.
- Get a code $C$ with better rate and “same” list-decodability properties.
More Generally: all $q$

- List-decoding capacity is $R = 1 - H_q(\rho)$.
- Can handle error rate $\rho = 1 - 1/q - \varepsilon$.

\[
1 - H_q(1 - 1/q - \varepsilon) \approx \min \left\{ \varepsilon, \frac{q \varepsilon^2}{2 \log(q)} \right\}
\]
More Generally

**Theorem**

Suppose that $C_0 \subset \mathbb{F}_q$, with distance $1 - 1/q - \varepsilon^2$. Choose $n$ so that

$$R \approx \frac{\min \{ \varepsilon, q\varepsilon^2 \}}{\log^5(1/\varepsilon) \log(q)}.$$ 

Then w.h.p., the resulting code $C$ is list-decodable up to error rate

$$\rho = 1 - 1/q - \varepsilon$$

with list sizes $\text{poly}(1/\varepsilon)$. 

Mary Wootters (University of Michigan) 
List decodability of RS codes 
January 17, 2014 21 / 50
More Generally

**Theorem**

Suppose that $C_0 \subset \mathbb{F}_q$, with distance $1 - \frac{1}{q} - \varepsilon^2$. Choose $n$ so that

$$R \approx \min \left\{ \varepsilon, q\varepsilon^2 \right\} \frac{\log^5(1/\varepsilon) \log(q)}{\log^5(1/\varepsilon) \log(q)}.$$

Then w.h.p., the resulting code $C$ is list-decodable up to error rate

$$\rho = 1 - \frac{1}{q} - \varepsilon$$

with list sizes $\text{poly}(1/\varepsilon)$.

Compare to

$$1 - H_q(1 - \frac{1}{q} - \varepsilon) \approx \min \left\{ \varepsilon, \frac{q\varepsilon^2}{\log(q)} \right\}.$$
Another application: random linear codes

- A **linear code** forms a subspace of $\mathbb{F}_q^n$.
- A **random linear code** is a random subspace of $\mathbb{F}_q^n$.

$$x \in \mathbb{F}_q^k \implies G = \{ c \in C \mid xG \}$$

Random **generator** matrix

$C = \{ xG \mid x \in \mathbb{F}_q^k \}$
<table>
<thead>
<tr>
<th>Source</th>
<th>Rate</th>
<th>List size</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Correct”</td>
<td>$\frac{q\varepsilon^2}{\log(q)}$</td>
<td>$\varepsilon^{-2}$</td>
</tr>
<tr>
<td>[Zyablov, Pinsker’81]</td>
<td>$\frac{q\varepsilon^2}{\log(q)}$</td>
<td>$q^{1/\varepsilon^2}$</td>
</tr>
<tr>
<td>Johnson bound</td>
<td>$\frac{q\varepsilon^4}{\log(q)}$</td>
<td>$\varepsilon^{-2}$</td>
</tr>
</tbody>
</table>

for medium-sized $q$
<table>
<thead>
<tr>
<th>Source</th>
<th>Rate</th>
<th>List size</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Correct”</td>
<td>$\frac{q\epsilon^2}{\log(q)}$</td>
<td>$\epsilon^{-2}$</td>
</tr>
<tr>
<td>[Zyablov, Pinsker’81]</td>
<td>$\frac{q\epsilon^2}{\log(q)}$</td>
<td>$q^{1/\epsilon^2}$</td>
</tr>
<tr>
<td>Johnson bound</td>
<td>$\frac{q\epsilon^2}{\log(q)}$</td>
<td>$\epsilon^{-2}$</td>
</tr>
<tr>
<td>[Cheraghchi, Guruswami, Velingker’13]</td>
<td>$\frac{\epsilon^2}{\log(q) \log^3(1/\epsilon)}$</td>
<td>$\epsilon^{-2}$</td>
</tr>
<tr>
<td>[W.’13]</td>
<td>$\frac{\epsilon^2}{\log(q)}$</td>
<td>$\epsilon^{-2}$</td>
</tr>
<tr>
<td>[Rudra,W.’13]</td>
<td>$\frac{q\epsilon^2}{\log(q) \log^5(1/\epsilon)}$</td>
<td>$\epsilon^{-2}$</td>
</tr>
</tbody>
</table>

for medium-sized $q$
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
   - Argument from [Rudra, W. 2013]
   - Argument from [W. 2013]

4 Conclusion
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
     - Argument from [Rudra, W. 2013]
     - Argument from [W. 2013]

4 Conclusion
All lists of size $> L$ have at least one codeword further than $\rho$

All lists of size $> L$ have average distance larger than $\rho$
Average-Radius List Decodability

List Decodable

All lists of size $> L$ have at least one codeword further than $\rho$

Average-Radius List Decodable

All lists of size $> L$ have average distance larger than $\rho$
What do we need for average-radius list-decodability?

\[ N = |C| \]

The average distance \( \Lambda \) from \( w \):

\[ 1 \left| \Lambda \right| = \sum_{j=1}^{n} p_{l}(\Lambda) \]

"Plurality" of symbol \( j \)

Mary Wootters (University of Michigan)
List decodability of RS codes
January 17, 2014 27 / 50
What do we need for average-radius list-decodability?

\[ N = |C| \]

Average distance of \( \Lambda \) from \( w \):

\[ |\Lambda| \sum_{j=1}^{n} p_{l_j}(\Lambda) \]

"Plurality" of symbol \( j \)

Mary Wootters (University of Michigan)
List decodability of RS codes
January 17, 2014 27 / 50
What do we need for average-radius list-decodability?

\[ N = |C| \]

Average distance of \( \Lambda \) from \( w \):

\[ \frac{1}{n} \sum_{j=1}^{pl(j)(\Lambda)} \]

"Plurality" of symbol \( j \)

Mary Wootters (University of Michigan)
What do we need for average-radius list-decodability?

\[ N = |C| \]

Average distance of \( \Lambda \) from \( w \):

\[ \frac{1}{n} \sum_{j=1}^{\text{pl}(\Lambda)} \]

"Plurality" of symbol \( j \)

Mary Wootters (University of Michigan)
List decodability of RS codes
January 17, 2014 27 / 50
What do we need for average-radius list-decodability?

\[ N = |C| \]

Average distance of \( \Lambda \) from \( w \):

\[ \frac{1}{n} \sum_{j=1}^{p} \text{pl}_j(\Lambda) \]

"Plurality" of symbol \( j \)

Mary Wootters (University of Michigan)
What do we need for average-radius list-decodability?

\[ N = |C| \]

Average distance of \( \Lambda \) from \( w \):

\[
\frac{1}{|\Lambda|} \sum_{j=1}^{n} pl_j(\Lambda).
\]
What do we need for average-radius list-decodability?

Average distance of $\Lambda$ from $w$:

$$\frac{1}{|\Lambda|} \sum_{j=1}^{n} \text{pl}_j(\Lambda).$$

“Plurality” of symbol $j$
Proof outline

To show that $C$ is list-decodable, suffices to show that

$$\mathbb{E} \max_{\Lambda} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}.$$
Proof outline

To show that $\mathcal{C}$ is list-decodable, suffices to show that

$$\mathbb{E} \max_{\Lambda} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small.}$$

The plan:

- Show

  $$\max_{\Lambda} \mathbb{E} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}$$

- Show

  $$\mathbb{E} \max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small.}$$
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
   - Argument from [Rudra, W. 2013]
   - Argument from [W. 2013]

4 Conclusion
Concentration

- Want to bound

$$\max_{\Lambda} \left| \sum_{j=1}^{n} p_{\lambda j}(\Lambda) - \mathbb{E} p_{\lambda j}(\Lambda) \right| \leq \text{small}.$$ 

- $n$ large $\Rightarrow$ more concentration.
- $n$ small $\Rightarrow$ better rate.
Approach 0: Union bound

\[
\max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small}.
\]

▶ We’ve got a sum of independent, mean zero things!
Approach 0: Union bound

\[ \max_{\Lambda} \left| \sum_{j=1}^{n} p_l_j(\Lambda) - \mathbb{E} p_l_j(\Lambda) \right| \leq \text{small.} \]

- We’ve got a sum of independent, mean zero things!
- But each \( p_l_j(\Lambda) \) is not very concentrated.
Approach 0: Union bound

\[ \max_{\Lambda} \left| \sum_{j=1}^{n} p_l j(\Lambda) - \mathbb{E}p_l j(\Lambda) \right| \leq \text{small}. \]

- We’ve got a sum of independent, mean zero things!
- But each \( p_l j(\Lambda) \) is not very concentrated.
- The summands are not concentrated enough to allow for a union bound over \( {N \choose L} \) things.
Basic idea from [Rudra, W.’13]

\[
\max_{\Lambda} \left| \sum_{j=1}^{n} \text{pl}_j(\Lambda) - \mathbb{E}\text{pl}_j(\Lambda) \right| \leq \text{small.}
\]

If \( \Lambda \) and \( \Lambda' \) are close, a union bound is wasteful.
Chaining argument

(relative) pluralities are:

Similar with **very** high probability.

Similar with high probability.

Similar with decent probability.
Chaining argument
Chaining argument
Chaining argument

\[
\max_{\Lambda} \left| \sum_{j=1}^{n} p_l_j(\Lambda) - \mathbb{E} p_l_j(\Lambda) \right| \leq \text{small.}
\]

Can afford:

- Union bound over \textbf{lots} of \textcolor{green}{green} edges.
- Union bound over \textbf{many} \textcolor{orange}{orange} edges.
- Union bound over \textbf{a few} \textcolor{red}{red} edges.

In the favorable case,

\[
\max_{|\Lambda|=L, |\Lambda|} \frac{1}{|\Lambda|} \left| \sum_{j=1}^{n} p_l_j(\Lambda) - \mathbb{E} p_l_j(\Lambda) \right| \approx \max_{|\Lambda|=1, |\Lambda|} \frac{1}{|\Lambda|} \left| \sum_{j=1}^{n} p_l_j(\Lambda) - \mathbb{E} p_l_j(\Lambda) \right|.
\]

This is well-behaved
That’s not quite right

Eventually,

$$\max_{|\Lambda|=L/2^j} \frac{1}{|\Lambda|} \sum_j p_{l_j}(\Lambda)$$

does get larger.
But it works with some tweaks

To show that $C$ is list-decodable, suffices to show that

$$\mathbb{E} \max_{\Lambda} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}.$$ 

The plan:

- Show
  $$\max_{\Lambda} \mathbb{E} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}$$

- Show
  $$\mathbb{E} \max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small}.$$
Start with a “decently” list decodable code $C_0$ with bad rate.
General result (restatement)

\[ N = |C_0| \]

- Start with a “decently” list decodable code \( C_0 \) with bad rate.
- Choose \( n \) random positions.
General result (restatement)

\[ N = |C_0| = |C| \]

- Start with a “decently” list decodable code \( C_0 \) with bad rate.
- Choose \( n \) random positions.
- Get a code \( C \) with better rate and “same” list-decodability properties.
General result (restatement)

**Theorem**

Suppose that $C_0$ is a linear code over $\mathbb{F}_q$, with distance $1 - 1/q - \varepsilon^2$. Choose $n$ so that

$$R \approx \min \left\{ \varepsilon, q\varepsilon^2 \right\} / \log^5(L) \log(q).$$

Then w.h.p., the resulting code $C$ is list-decodable up to error rate

$$\rho = 1 - 1/q - \varepsilon$$

with list sizes $\text{poly}(1/\varepsilon)$.

Compare to

$$1 - H_q(1 - 1/q - \varepsilon) \approx \min \left\{ \varepsilon, \frac{q\varepsilon^2}{\log(q)} \right\}.$$
1 Background
- Coding Theory and List Decoding
- Reed-Solomon Codes

2 Results

3 Proof ideas
- Outline
- Argument from [Rudra, W. 2013]
- Argument from [W. 2013]

4 Conclusion
A simpler argument for $q = 2$

To show that $\mathcal{C}$ is list-decodable, suffices to show that

$$\mathbb{E} \max_{\Lambda} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}.$$ 

The plan:

- Show

  $$\max_{\Lambda} \mathbb{E} \sum_{j=1}^{n} pl_j(\Lambda) \leq \text{small}$$

- Show

  $$\mathbb{E} \max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small}.$$
Two tools

- Symmetrization
- Mean width
Symmetrization

For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\| \cdot \|$, 

$$
\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| \lesssim \mathbb{E} \left\| \sum_i \xi_i Y_i \right\|
$$
Symmetrization

For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\| \cdot \|$, 

$$
\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| \lesssim \mathbb{E} \left\| \sum_i \xi_i Y_i \right\|
$$

Proof.

Let $Y'_i$ be independent copies of $Y_i$. 

$$
\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| = \mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y'_i \right\|
$$
Symmetrization

For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\|\cdot\|$, 

$$
\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| \preceq \mathbb{E} \left\| \sum_i \xi_i Y_i \right\|
$$

Proof.

Let $Y_i'$ be independent copies of $Y_i$. 

$$
\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| = \mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i' \right\| \leq \mathbb{E} \left\| \sum_i (Y_i - Y_i') \right\|
$$
Symmetrization

For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\|\cdot\|$, 

$$
E\left\| \sum_i Y_i - EY_i \right\| \lesssim E\left\| \sum_i \xi_i Y_i \right\| 
$$

Proof.

Let $Y_i'$ be independent copies of $Y_i$.

$$
E\left\| \sum_i Y_i - EY_i \right\| = E\left\| \sum_i Y_i - EY_i' \right\| \leq E\left\| \sum_i (Y_i - Y_i') \right\| = E\left\| \sum_i \xi_i (Y_i - Y_i') \right\| 
$$
Symmetrization
For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\| \cdot \|$, 
\[
\mathbb{E} \left| \sum_i Y_i - \mathbb{E} Y_i \right| \leq \mathbb{E} \left| \sum_i \xi_i Y_i \right|
\]

Proof.
Let $Y_i'$ be independent copies of $Y_i$.
\[
\mathbb{E} \left| \sum_i Y_i - \mathbb{E} Y_i \right| = \mathbb{E} \left| \sum_i Y_i - \mathbb{E} Y_i' \right| \leq \mathbb{E} \left| \sum_i (Y_i - Y_i') \right| \\
= \mathbb{E} \left| \sum_i \xi_i (Y_i - Y_i') \right| \leq 2\mathbb{E} \left| \sum_i \xi_i Y_i \right|
\]
Symmetrization

For $\xi_i = \pm 1$ i.i.d., independent random variables $Y_i$, and any norm $\| \cdot \|$, 

$$\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| \lesssim \mathbb{E} \left\| \sum_i g_i Y_i \right\|$$

Proof.

Let $Y_i'$ be independent copies of $Y_i$.

$$\mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i \right\| = \mathbb{E} \left\| \sum_i Y_i - \mathbb{E} Y_i' \right\| \leq \mathbb{E} \left\| \sum_i (Y_i - Y_i') \right\|$$

$$= \mathbb{E} \left\| \sum_i \xi_i (Y_i - Y_i') \right\| \leq C \cdot 2 \mathbb{E} \left\| \sum_i g_i Y_i \right\|$$
(Gaussian) mean width of $T = \mathbb{E} \max_{t \in T} \langle g, t \rangle$
A problem about matrices

Stack the codewords as the columns of a matrix $A$.

$$\Lambda_n$$

$$n$$

$$\Lambda$$

$$\max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small}.$$
A problem about matrices

Stack the codewords as the columns of a matrix $A$.

$$\max_{\Lambda} \left| \sum_{j=1}^{n} p_{l_j}(\Lambda) - \mathbb{E} p_{l_j}(\Lambda) \right| \leq \text{small}.$$
A problem about matrices

Stack the codewords as the columns of a matrix $A$.

$$\Lambda$$

$$n$$

$$\Lambda$$

$$\max_{\Lambda} \left| \sum_{j=1}^{n} pl_j(\Lambda) - \mathbb{E} pl_j(\Lambda) \right| \leq \text{small}.$$
To control concentration of $\|A1_\Lambda\|_1$

\[
\max_{\Lambda} ||A1_\Lambda||_1 - \mathbb{E}||A1_\Lambda||_1 \leq \text{small}.
\]

\[
\mathbb{E} \max_{\Lambda} ||A1_\Lambda||_1 - \mathbb{E}||A1_\Lambda||_1
\]
To control concentration of $\|A_{1\Lambda}\|_1$

\[
\max_{\Lambda} \|A_{1\Lambda}\|_1 - \mathbb{E}\|A_{1\Lambda}\|_1 \leq \text{small}.
\]

\[
\mathbb{E} \max_{\Lambda} \|A_{1\Lambda}\|_1 - \mathbb{E}\|A_{1\Lambda}\|_1 \leq \mathbb{E} \max_{\|x\|_1 = L} \left| \sum_{j=1}^{n} \langle A_j, x \rangle - \mathbb{E} \langle A_j, x \rangle \right|.
\]
To control concentration of $\|A_1\|_1$

$$\max_{\Lambda} \|A_1\|_1 - \mathbb{E}\|A_1\|_1 \leq \text{small}.$$
To control concentration of $\|A_1^\Lambda\|_1$

$$\max_{\Lambda} \|A_1^\Lambda\|_1 - \mathbb{E}\|A_1^\Lambda\|_1 \leq \text{small}.$$
To control concentration of $\|A_1\|_1$

\[
\max_{\Lambda} \|A_1\|_1 - \mathbb{E}\|A_1\|_1 \leq \text{small}.
\]

\[
\mathbb{E} \max_{\Lambda} \|A_1\|_1 - \mathbb{E}\|A_1\|_1 \leq \mathbb{E} \max_{\|x\|_1=L} \left| \sum_{j=1}^{n} \langle A_j, x \rangle - \mathbb{E} \langle A_j, x \rangle \right|
\]

\[
\lesssim \mathbb{E} \max_{\|x\|_1=L} \sum_{j=1}^{n} g_j \langle A_j, x \rangle
\]

\[
= \mathbb{E} \max_{\|x\|_1=L} \langle g, Ax \rangle
\]

Symmetrization

Independent centered random variables
To control concentration of $\|A^{1\Lambda}\|_1$

$$\max_{\Lambda} |\mathbb{E}\|A^{1\Lambda}\|_1 - \mathbb{E}\|A^{1\Lambda}\|_1| \leq \text{small}.$$
To control concentration of $∥A1_Λ∥_1$

\[
\max_Λ ∥A1_Λ∥_1 - \mathbb{E}∥A1_Λ∥_1 \leq \text{small}.
\]

\[
\mathbb{E} \max_Λ ∥A1_Λ∥_1 - \mathbb{E}∥A1_Λ∥_1 \leq \mathbb{E} \max_{∥x∥_1=\mathcal{L}} \left| \sum_{j=1}^{n} \langle A_j, x \rangle - \mathbb{E} \langle A_j, x \rangle \right|
\]

\[
\leq \mathbb{E} \max_{∥x∥_1=\mathcal{L}} \sum_{j=1}^{n} g_j \langle A_j, x \rangle
\]

\[
= \mathbb{E} \max_{∥x∥_1=\mathcal{L}} \langle g, Ax \rangle
\]

\[
= L \mathbb{E} \max_{y \in AB_1} \langle g, y \rangle
\]

**Symmetrization**

**Mean width of $AB_1$**

**Independent centered random variables**
Mean width of $A B_1$

- $A B_1$ is the convex hull of the columns of $A$ (and their opposites)
Mean width of $A B_1$

- $A B_1$ is the convex hull of the columns of $A$ (and their opposites)

![Diagram]

- $E \max_{y \in A B_1} \langle g, y \rangle = E \max_{\text{cols } A_j \text{ of } A} \langle g, A_j \rangle$
Mean width of $A B_1$

$A B_1$ is the convex hull of the columns of $A$ (and their opposites)

$\mathbb{E} \max_{y \in A B_1} \langle g, y \rangle = \mathbb{E} \max_{\text{cols } A_j \text{ of } A} \langle g, A_j \rangle = N(0, \|A_j\|_2^2) = N(0, n)$
Mean width of $A B_1$

- $A B_1$ is the convex hull of the columns of $A$ (and their opposites)

\[
E \max_{y \in A B_1} \langle g, y \rangle = E \max_{\text{cols } A_j \text{ of } A} \langle g, A_j \rangle = E \max \{ 2N \text{ gaussians with variance } n \}
\]

\[
N(0, \|A_j\|_2^2) = N(0, n)
\]
Mean width of $AB_1$

- $AB_1$ is the convex hull of the columns of $A$ (and their opposites)

\[
\mathbb{E} \max \langle g, y \rangle = \mathbb{E} \max_{\text{cols } A_j \text{ of } A} \langle g, A_j \rangle
\]

\[
= \mathbb{E} \max \{ \text{2N gaussians with variance } n \}
\]

\[
\lesssim \sqrt{n \log(N)}
\]
This is enough to imply

**Theorem**

Let $\mathcal{C}$ be a random linear code over $\mathbb{F}_q$ of rate

$$R = \frac{C \varepsilon^2}{\log(q)}.$$

Then $\mathcal{C}$ is list-decodable up to error rate

$$\rho = 1 - \frac{1}{q} - \varepsilon$$

with list sizes

$$L = O\left(\frac{1}{\varepsilon^2}\right).$$
1 Background
   - Coding Theory and List Decoding
   - Reed-Solomon Codes

2 Results

3 Proof ideas
   - Outline
   - Argument from [Rudra, W. 2013]
   - Argument from [W. 2013]

4 Conclusion
Recap

Results:

- **Question:** Are there Reed-Solomon codes list-decodable beyond the Johnson bound?
  - Yes.

The result is more general, and also gives improved bounds for random linear codes.

Questions:

- Remove log factors?
- Find a RS code which is optimally list-decodable? (Or a sufficient condition?)
- Applications to other pseudorandom objects?
Recap

Results:

- **Question:** Are there Reed-Solomon codes list-decodable beyond the Johnson bound?
  - Yes.
  
- The result is more general, and also gives improved bounds for random linear codes.
Recap

Results:

- **Question:** Are there Reed-Solomon codes list-decodable beyond the Johnson bound?
  - Yes.
- The result is more general, and also gives improved bounds for random linear codes.

Questions:

- Remove log factors?
- Find a RS code which is optimally list-decodable? (Or a sufficient condition?)
- Applications to other pseudorandom objects?
The end

Thanks!
Theorem

Suppose that

\[ \rho \leq (1 - 1/q) \left( 1 - \sqrt{1 - \delta \left( \frac{q}{q-1} \right)} \right). \]

Then \( C \) is \((\rho, L)\) list-decodable with

\[ L \leq qn^2 \delta. \]

Thus, for

\[ \delta = (1 - 1/q) \left( 1 - \varepsilon^2 \right), \]

we have

\[ \rho = (1 - 1/q) \left( 1 - \varepsilon \right). \]
The rate/distance trade-off of a random linear code is given by the Gilbert-Varshamov bound,

\[ R \approx 1 - H_q(\delta), \]

and so for

\[ \delta = 1 - 1/q - \varepsilon^2, \]

this is approximately

\[ R = \frac{q\varepsilon^4}{\log(q)} \]

for small \( q \) or

\[ R \approx \varepsilon^2 \]

for large \( q \).
List-Decoding Capacity Theorem

For all $L, \rho$:

Rate

\[ R \leq 1 - H_q(\rho) - \frac{1}{L} \]

⇒

There are codes with
- Error rate $\rho$
- List size $L$

Rate

\[ R \geq 1 - H_q(\rho) + \varepsilon \]

⇒

Any code needs

\[ L \geq \exp(n). \]
List-Decoding Capacity Theorem

For big $q$, the picture is:

\[ \text{Error rate } \rho \]

\[ \text{Rate } R \]

\[ 1 - H_q(\rho) \]

Upper bound for unique decoding (Singleton bound)

We are interested in the "big error, small rate" regime.
List-Decoding Capacity Theorem

For big $q$, the picture is:

\[ R = \frac{1}{2} - \frac{1}{q} \]

Upper bound for unique decoding (Singleton bound)

We are interested in the “big error, small rate” regime.

\[ 1 - H_q(\rho) \]
Applications of list-decoding

- Communication: e.g., [Micali, Peikert, Sudan, Wilson 2005] use list-decodable codes to make good uniquely-decodable codes under crypto assumptions.

- Complexity theory:
  - Good list decodable codes can be used to get:
    - Hardness amplification
    - Hardcore predicates from one-way functions
    - Extractors
    - Expanders
    - Pseudorandom generators
  - (See [Sudan’00] or [Vadhan’11] for good surveys).
  - [Cai, Pavan, Sivakumar 1999] use list-decodability of RS codes in particular for reductions about hardness of computing permanents.