List Decoding of Noisy Reed-Muller-like Codes

Martin J. Strauss
University of Michigan

Joint work with
A. Robert Calderbank (Princeton)
Anna C. Gilbert (Michigan)
Joel Lepak (Michigan)
Euclidean List Decoding

• Fix
  ◦ structured spanning codebook $\mathcal{C} = \{\varphi_\lambda\}$ of vectors in $\mathbb{C}^N$
  ◦ Parameter $k$.

• Given vector (“signal”) $s \in \mathbb{C}^N$.
  ◦ Accessed by sampling: query $y$, learn $s(y)$.

• Goal:
  
  Quickly find list of $\lambda$ such that $|\langle \varphi_\lambda, s \rangle|^2 \geq (1/k)\|s\|^2$.

• For some codebooks, leads to sparse approximation:
  ◦ Small $\Lambda$ with $\tilde{s} = \sum_{\lambda \in \Lambda} c_\lambda \varphi_\lambda \approx s$. 
Definitions of Reed-Muller (-like) Codes

For $y, \ell \in \mathbb{Z}_2^n$; $P$ a binary symmetric matrix:

- Second-order Reed-Muller, RM(2):
  \[ \varphi_{P,\ell}(y) = iy^T P y + 2\ell^T y. \]

- Hankel, Kerdock codes: limited allowable $P$’s.
  \[ \diamond \text{ Hankel: } P \text{ is constant along reverse diagonals.} \]

- First-order Reed-Muller, RM(1):
  \[ \varphi_{0,\ell}(y) = i^{2 \ell^T y} = (-1)^{\ell^T y}. \]

Sometimes omit normalization factor $1/\sqrt{N}$; makes $\|\varphi\|_2 = 1$. 


Our Results

- Theorem: There’s a Kerdock code that is a subcode of Hankel.
- Theorem: We give a list-decoding algorithm for length-\(N\) Hankel.
  - Return list \(\Lambda\) of Hankel \(\lambda\) such that \(\left|\langle \varphi_\lambda, s \rangle \right|^2 \geq \left(1/k\right) \|s\|^2\)
  - ...in time poly\((k \log(N))\).
- Corollary: We give a fast list-decoding algorithm for Kerdock.
- Corollary: We give a fast sparse recovery algorithm for Kerdock.
Overview

• Motivation
• New construction of Kerdock
• List decoding for Hankel
• Alternatives and conclusion
Significance

- First “simple” construction of a Kerdock code, as Hankel subcode. (Isomorphic to an existing “complicated” construction [Calderbank-Cameron-Kantor-Seidel].)
- To our knowledge, first extension of RM(1) list decoding to large codebook with small alphabet.
- Sparse recovery for the important Kerdock code.
  - Wireless communication—Multi-User Detection (Joel Lepak)
  - Quantum information
- Hankel and Kerdock compromise between RM(1) and RM(2)
  - Code parameters
  - Learning
Related Work

- List decoding over a single ONB [Kushilevitz-Mansour] doesn’t (directly) give a result for the union of many ONBs (Kerdock, Hankel)
- Test for RM(2) [Alon-Kaufman-Krivelevich-Litsyn-Ron] is not a test for Kerdock and doesn’t do list decoding.
- Decoding RM(2) with low noise [AKKLR] doesn’t help with high noise.
- Work over large alphabets [Sudan, …] doesn’t help over $\mathbb{Z}_2$. (Restrict multi-variate polynomial to random line, getting univariate polynomial. But low-degree univariate polys over $\mathbb{Z}_2$ are not interesting.)
- General sparse recovery [Gilbert-Muthukrishnan-S-Tropp, …] requires time poly($2^n$) $\gg$ poly($k, n$) and/or space poly($2^n$)
Fundamental Properties of Kerdock

Used in our recovery algorithm and of independent interest. [Calderbank-Cameron-Kantor-Seidel]

• **Geometry.** Union of $N$ ONB’s, each of the form $\varphi_{P,0} \cdot \text{RM}(1)$ for Kerdock $P$. ("Mutually-Unbiased Bases.")

\[ |\langle \varphi_{P,\ell}, \varphi_{P',\ell'} \rangle| = \begin{cases} 
1, & P = P', \ell = \ell' \\
0, & P = P', \ell \neq \ell' \\
1/\sqrt{N}, & P \neq P'. 
\end{cases} \]

• **Algebra.** Add’n and invertible mult’n of Kerdock matrices $P$.
  ◊ Map one ONB to another and permute elements of one ONB.

• **Multiscale Similarity.** Some structure is preserved on some restrictions to subspaces.
Multi-User Detection

• Each subscriber gets a set of codewords.
• To speak, a user picks a codeword $\varphi_\lambda$ from her set.
  ◊ Message is encoded in choice of codeword and/or coefficient $c_\lambda$.
• Receiver gets $\sum_\lambda c_\lambda \varphi_\lambda$.
• Decoder recovers all $(\lambda, c_\lambda)$’s.

RM(2) and Hankel won’t work. Kerdock supports more users than RM(1) for fixed blocklength.
Quantum Key Distribution

Four polarization directions:

- vertical $|v\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, horizontal $|h\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and two diagonals

$|v\rangle + |h\rangle = \begin{pmatrix} +1 \\ +1 \end{pmatrix}$ and $|v\rangle - |h\rangle = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$.

arranged in two mutually unbiased bases,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, H = \begin{pmatrix} +1 & +1 \\ +1 & -1 \end{pmatrix} / \sqrt{2}.$$ 

Punchline: Diagonal particle measured in $I$ comes out $|v\rangle$ or $|h\rangle$.

- Kerdock gives optimal construction of larger MUBs.
Compromise between RM(1) and RM(2)

- Code parameters.
- Learning. RM(1) is linear functions; RM(2) is quadratics. Kerdock and Hankel are some quadratics, namely, $f(y_0, y_1, y_2, y_3, \ldots)$ has term $2y_0y_4$ iff it has $2y_1y_3$ and $y_2^2 = y_2$, etc. E.g.:

\[
\begin{pmatrix}
y_0 & y_1 & y_2 & y_3 & y_4
\end{pmatrix}
\begin{pmatrix}
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
y_0 \\
y_1 \\
y_2 \\
y_3 \\
y_4
\end{pmatrix}
\]
Overview

- Motivation ✓
- New construction of Kerdock
- List decoding for Hankel
- Alternatives and conclusion
Definition of Hankel

A matrix is *Hankel* if it is constant on reverse diagonals,

\[ P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 \\ p_1 & p_2 & p_3 & p_4 \\ p_2 & p_3 & p_4 & p_5 \\ p_3 & p_4 & p_5 & p_6 \end{pmatrix} \]

The *Hankel* code is the subcode of \( \text{RM}(2) = \{ \varphi_{P,\ell} \} \) in which \( P \) is Hankel. [Calderbank-Gilbert-Levchenko-Muthukrishnan-S]
Definition of Kerdock

A set of matrices is a *Kerdock set* if the sum of any two is non-singular or zero.

Each Kerdock set of matrices leads to *some* Kerdock code.

- Kerdock matrix $P$ and vector $\ell$: $\varphi_{P,\ell}$.

Note:

- There are at most $N = 2^n$ matrices in a Kerdock set, since each matrix in the set has a distinct top row.
- We’ll construct a maximum-sized set.
Our Construction of a Kerdock Set

Fix primitive polynomial \( h(t) = h_0 + h_1 t + \cdots + h_n t^n \) over \( \mathbb{Z}_2[t] \), e.g., \( n = 4 \). A matrix \( P \) is \( lf \)-Kerdock if

- \( P \) is Hankel,

\[
P = \begin{pmatrix}
p_0 & p_1 & p_2 & p_3 \\
p_1 & p_2 & p_3 & p_4 \\
p_2 & p_3 & p_4 & p_5 \\
p_3 & p_4 & p_5 & p_6
\end{pmatrix}
\]

- (Top row \( p_0, p_1, p_2, p_3 \) unconstrained)
- Each other parameter is a linear combination of top-row parameters, using linear-feedback rule with coefficients in \( h \).
Example

Primitive polynomial $h(t) = t^3 + t + 1 = t^3 + 0t^2 + 1t + 1$.

$$P = \begin{pmatrix} p_0 & p_1 & p_2 \end{pmatrix}$$

Top row unconstrained.
Example

Primitive polynomial $h(t) = t^3 + t + 1 = t^3 + 0t^2 + 1t + 1$.

$P = \begin{pmatrix} p_0 & p_1 & p_2 \\ p_1 & p_2 & p_3 = \\ p_2 & p_3 = & p_4 = \end{pmatrix}$

Top row unconstrained.

Extend to Hankel.
Example

Primitive polynomial \( h(t) = t^3 + t + 1 = t^3 + 0t^2 + 1t + 1. \)

\[
P = \begin{pmatrix}
p_0 & p_1 & p_2 \\
p_1 & p_2 & p_3 = p_0 + p_1 \\
p_2 & p_3 = p_0 + p_1 & p_4 = p_1 + p_2
\end{pmatrix}
\]

Top row unconstrained.

Extend to Hankel.

Use feedback rule for lower half.
Proof of Correctness

Theorem: A set of lf-Kerdock matrices is a Kerdock set.

Sufficient to show that lf-Kerdocks are non-singular. Definitions:

- Additive $\text{Tr} : \mathbb{F}(2^n) \rightarrow \mathbb{F}(2)$ is given by
  $$\text{Tr}(x) = x + x^2 + x^4 + x^8 + \cdots + x^{2^n-1}.$$ 

- Recall $h$ is primitive polynomial; $h(\xi) = 0$.

- $(K_\alpha)_{j,k} := \text{Tr}(\alpha \xi^{j+k})$ ("trace-Kerdock" matrix, for $\alpha \in \mathbb{F}(2^n)$)

Three lemmas, one-line proofs:

- Trace-Kerdocks are non-singular.
- Trace-Kerdocks are lf-Kerdock.
- lf-Kerdocks are trace-Kerdock.
Facts about Trace

Recall \( \text{Tr}(x) = x + x^2 + x^4 + x^8 + \cdots + x^{2^n-1} \). Squaring is linear in characteristic 2, so

- \( \text{Tr}(x + y) = \text{Tr}(x) + \text{Tr}(y) \).
- \( \text{Tr}(x)^2 = \text{Tr}(x^2) = \text{Tr}(x) \).
  - \( \text{Tr}(x) \) satisfies \( y^2 + y = 0 \).
  - \( \text{Tr}(x) \in \{0, 1\} \).
**Trace-Kerdocks are non-Singular**

**Lemma:** Trace-Kerdocks are non-Singular

\[ K_\alpha = V^T D_\alpha V \text{ over } \mathbb{F}(2^n), \text{ where} \]
\[ D_\alpha = \text{diag}(\alpha, \alpha^2, \alpha^4, \alpha^8, \ldots, \alpha^{2^{n-1}}) \]
and vandermonde \( V \) is given by

\[
V = \begin{pmatrix}
1 & \xi & \xi^2 & \xi^3 & \xi^4 & \cdots \\
1 & \xi^2 & \xi^4 & \xi^6 & \xi^8 & \cdots \\
1 & \xi^4 & \xi^8 & \xi^{12} & \xi^{16} & \cdots \\
1 & \xi^8 & \xi^{16} & \xi^{24} & \xi^{32} & \cdots \\
\vdots
\end{pmatrix}.
\]

\( K_\alpha \) is over \( \mathbb{F}(2) \), so \( \det(K_\alpha) \in \mathbb{F}(2) \) over big field.
Trace-Kerdocks are lf-Kerdock

**Lemma:** Trace-Kerdocks are lf-Kerdock.

A trace-Kerdock \((K_\alpha)_{j,k} := \text{Tr}(\alpha \xi^{j+k})\) is Hankel by inspection.

Feedback rule:

\[
\text{Tr}(\alpha \xi^{j+k+n}) = \text{Tr} \left( \alpha \xi^{j+k} \sum_{\ell<n} h_\ell \xi^\ell \right) \\
= \sum_{\ell<n} h_\ell \text{Tr} \left( \alpha \xi^{j+k} \xi^\ell \right),
\]

so feedback rule is satisfied.
Lemma: lf-Kerdocks are Trace-Kerdock.

There are $2^n$ distinct matrices of each type. Above we showed that all trace-Kerdocks are lf-Kerdock.
Overview

• Motivation ✓
• New construction of Kerdock ✓
• List decoding for Hankel
  ◇ Review of list decoding for RM(1).
  ◇ (Simple) extension of algorithm to Hankel.
  ◇ Hankel structure keeps intermediate and final lists small.
• Alternatives and conclusion
Tensor-Product View of RM(1)

\[ \varphi_{1011} = \varphi_{1000} \cdot \varphi_{0010} \cdot \varphi_{0001} \] is signal of length \(2^4 = 16\).

Start with \(\varphi_{0000} \simeq 1\) and flip bits, in dyadic blocks.

<table>
<thead>
<tr>
<th>(\varphi_{0000})</th>
<th>++++</th>
<th>++++</th>
<th>++++</th>
<th>++++</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flip</td>
<td>v v</td>
<td>v v</td>
<td>v v</td>
<td>v v v v</td>
</tr>
<tr>
<td>(\varphi_{1000})</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
</tr>
<tr>
<td>Flip</td>
<td>vvvv</td>
<td>vvvv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\varphi_{1010})</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
</tr>
<tr>
<td>Flip</td>
<td>vvvv</td>
<td>vvvv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\varphi_{1011})</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
<td>+---</td>
</tr>
</tbody>
</table>
RM(1) Recovery

\[ \text{E.g., [Kushilevitz-Mansour]} \]

Want \( \ell \) such that \( |\langle s, \varphi_\ell \rangle|^2 \geq \frac{1}{k} \|s\|^2 \).

For \( j \leq n \), maintain candidate list \( L_j \) for first \( j \) bits of \( \ell \).

Extend candidates one bit at a time—\( j \) to \((j + 1)\)—and test.

Need to show, with high probability:

- No false negatives
  - True candidates are found
- Few (false) positives
  - List remains small; algorithm is efficient.
  - Can remove false positives at the end.

\[ \text{25} \]
RM(1) Recovery, No False Negatives

Signal $s \in \mathbb{C}^{16}$; candidate $\varphi_\ell$ with $\ell = 01**$.

<table>
<thead>
<tr>
<th>Signal $s$</th>
<th>+++-</th>
<th>+++-</th>
<th>+++3</th>
<th>i-++</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi_{0100}$</td>
<td>+++-</td>
<td>+++-</td>
<td>+++-</td>
<td>+++-</td>
</tr>
<tr>
<td>$\varphi_{01**}$</td>
<td>+++-</td>
<td>$\pm 1$·++++</td>
<td>$\pm 1$·++++</td>
<td>$\pm 1$·++++</td>
</tr>
</tbody>
</table>

- $|\langle \varphi_{0100}, s \rangle|^2$ is high compared with $\|s\|^2$.
- $|\langle \varphi_{0100}, s \rangle|^2$ consists of contributions from dyadic blocks, many of which are high.
- Each dyadic block’s contribution is sum of small contributions.
- Keep candidate $\varphi_{01**}$ since $\geq 1/O(k)$ blocks have square dot product $\geq 1/O(k)$, as estimated by sampling.
- Alternative view of dot product: $|\langle s, \varphi \rangle|^2 = |\langle s\varphi^*, \pm 1 \rangle|^2$. 

26
RM(1) Recovery, Few (False) Positives

- Markov: Most blocks get not much more than $E[]$ share of $\|s\|^2$.
- Parseval: In each of $B$ dyadic blocks, $\leq k$ large dot products.
- So total number of $\checkmark$’s is $\leq kB$.
- Thus: number $\varphi_{P,\ell}$’s with $\geq B/k$ $\checkmark$’s is $\leq k^2$.

\[
\begin{array}{c}
\text{dyadic blocks} \\
(P, \ell) \\
\text{large dot product}
\end{array}
\rightarrow
\begin{array}{c}
B \\
\checkmark \checkmark \checkmark \checkmark \checkmark \checkmark \checkmark \checkmark \\
\checkmark \\
\checkmark
\end{array}
\]
Want $P, \ell$ such that $|\langle s, \varphi_\ell \rangle|^2 \geq (1/k) \|s\|^2$.

Find $P$, then use KM to find $\ell$ with large

$|\langle s, \varphi_{P,\ell} \rangle|^2 = |\langle s\varphi_{P,0}^*, \varphi_{0,\ell} \rangle|^2$.

For $j \leq n$, maintain candidate list for upper-left $j$-by-$j$ submatrix $P'$ of $P$.

Extend Hankel $P'$ one row/column at a time—four possibilities—and test.

\[
P = \begin{pmatrix}
P' & a \\
ap & b
\end{pmatrix}
\]
Hankel Recovery, cont’d

Keep candidate $P'$ if, on many dyadic blocks, for restricted signal $s'$, there is some RM(1) vector $\varphi_{\ell'}$ with $|\langle s' \varphi_{P',0}^*, \varphi_{0,\ell'} \rangle|^2$ large.

- Divide out RM(2) part, $\varphi_{P',0}'$.
- See if result is well-approximated by RM(1).
- Use KM to determine this.

With high probability, no false negatives:

- Algorithm works for all RM(2) just like for RM(1).

Need to show few (false) positives. Sufficient to show:

- few positives within each dyadic block.
- few large Hankel coefficients to any signal, $s$. 
There are Few Large Hankel Coefficients

Want: Approximate Parseval for the Hankel codebook.

• Dickson: \( \text{rank}(P + P') \) high \( \Rightarrow \) \( \langle \varphi_P, \ell, \varphi_{P'}, \ell' \rangle \) small.

• Incoherence: All dot products small \( \Rightarrow \) appropriate approximate Parseval.

• For each \( P \), there are few \( P' \) with \( \text{rank}(P + P') \) low.

• Put it all together:

\[ \diamond \text{Theorem: Given signal } s \text{ and parameter } k, \text{ there are at most } \text{poly}(k) \text{ Hankel vectors } \varphi_{P, \ell} \text{ with } |\langle \varphi_{P, \ell}, s \rangle|^2 \geq (1/k) \|s\|^2. \]
Dickson’s Theorem

If \((P, ℓ) \neq (P', ℓ')\), then

\[
|\langle ϕ_P,ℓ, ϕ_{P'},ℓ' \rangle| \leq 2^{-\text{rank}(P+P')/2}.
\]

Relates dot products to the rank of \(P\)-matrix sums mod 2.

Bigger rank \(⇒\) vectors are closer to orthogonal.
Dickson for Kerdock, proof

\[ N^2 \langle \varphi_P, \ell, \varphi_0, 0 \rangle^2 = \sum_{y,z} i^T P y + 2\ell^T y + z^T P z + 2\ell^T z \]

\[ = \sum_{y,w} i^T w^T P w + 2\ell^T w + 2y^T P (w+y), \quad w = y + z \]

\[ = \sum_{y,w} i^T w^T P w + 2\ell^T w + 2y^T P w + 2d^T y, \quad d = \text{diag}(P) \]

\[ = \sum w i^T w^T P w + 2\ell^T w \sum y i^T 2(w^T P + d^T)y \]

\[ = N \sum w i^T w^T P w + 2\ell^T w \delta(w^T P, d^T), \quad P w = d \]

\[ = N i^T d^T P^{-1} d + 2\ell^T P^{-1} d. \]

Thus \( |\langle \varphi_P, \ell, \varphi_0, 0 \rangle|^2 = 1/\sqrt{N} \).
Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0, 1, 2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

$$
\begin{pmatrix}
a & b & c \\
b & c \\
c
\end{pmatrix}
$$
Few Low-Rank Hankels

Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0, 1, 2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

$$
\begin{pmatrix}
    a & b & c & d \\
    b & c & d \\
    c & d \\
\end{pmatrix}
$$

Learn $d$ from linear combination.
Few Low-Rank Hankels

Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0,1,2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

\[
\begin{pmatrix}
a & b & c & d \\
b & c & d \\
c & d \\
d
\end{pmatrix}
\]

Fill in by Hanklicity.
Few Low-Rank Hankels

Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0,1,2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

\[
\begin{pmatrix}
a & b & c & d \\
b & c & d & e \\
c & d \\
d \\
\end{pmatrix}
\]

Learn $e$ from linear combination
Theorem: At most \(2^{O(r)}\) Hankel matrices have rank at most \(r\).

Proof: Suppose column 3 is a linear combination \(C\) of columns 0,1,2. Then \(C\) and positions 0, 1, 2 in top row determine the top half of the matrix:

\[
\begin{pmatrix}
a & b & c & d & e \\
b & c & d & e \\
c & d & e \\
d & e \\
e \\
\end{pmatrix}
\]

Fill in by Hanklicity.
Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0, 1, 2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

\[
\begin{pmatrix}
  a & b & c & d & e \\
  b & c & d & e \\
  c & d & e & f \\
  d & e \\
  e
\end{pmatrix}
\]

Learn $f$ by linear combination.
Theorem: At most $2^{O(r)}$ Hankel matrices have rank at most $r$.

Proof: Suppose column 3 is a linear combination $C$ of columns 0,1,2. Then $C$ and positions 0, 1, 2 in top row determine the top half of the matrix:

$$
\begin{pmatrix}
  a & b & c & d & e & f \\
  b & c & d & e & f \\
  c & d & e & f \\
  d & e & f \\
  e & f \\
  f
\end{pmatrix}
$$

Fill in by Hanklicity.
Hankel Vectors

Space of Hankel Vectors:

Dot: vector $\varphi_{P,0}$.

Ball: $\varphi_{P',0}$ with rank($P + P'$) $\leq 2\log(k)$

Stick: vectors $\varphi_{P,\ell}$, as $\ell$ varies.
Claim: Few lollipops with heavy vector \( \varphi_{P,\ell} \) \((|\langle \varphi_{P,\ell}, s \rangle|^2 \text{ large})\).

- Each dot & stick meets few lollipops. -Few low-rank Hankels.
Few Large Hankel Coefficients

Claim: Few lollipops with heavy vector $\varphi_{P,\ell}$ ($|\langle \varphi_{P,\ell}, s \rangle|^2$ large).

- Each dot & stick meets few lollipops. -Few low-rank Hankels.
- Disjoint lollipops are nearly orthogonal. -Dickson
- No large sets of heavy vectors in nearly-orthogonal subset. -Incoherence (approximate Parseval)
Sparse Recovery of Kerdock

Corollary: There is an algorithm to recover a near-best $k$-term Kerdock representation to length-$N$ vector in time $\text{poly}(k \log(N))$.

Uses incoherence of Kerdock: for Kerdock $\varphi \neq \psi$, we have $|\langle \varphi, \psi \rangle| \leq 1/\sqrt{N}$.
Overview

- Motivation ✓
- New construction of Kerdock ✓
- List decoding for Hankel ✓
- Alternatives and conclusion
A faster alternative to KM first permutes the RM(1) labels $\ell^T \rightarrow \ell^T R$:

$$(Ts)(y) = s(Ry) = i^{2\ell^T (Ry)} = i^{2(\ell^T R)y},$$

for random invertible $R$. Simulate by substituting $Ry$ for $y$.

Us: Recall $K_{\alpha} = V^T D_{\alpha} V$. Use

$R = V^{-1} D_r V = (V^T V)^{-1} (V^T D_r V) = K_1^{-1} K_r$.

- Maps $K_{\alpha}$ to $K_{\alpha r^2}$—preserves Kerdock structure.
- For each $\ell$, $\ell^T R$ is uniform over $\mathbb{Z}_2^n$ for such $R$.

Can randomize KM in our inner loop while preserving Kerdock structure.

Get faster recovery algorithm, but only for Kerdock.
Mutliscale Similarity

- Restricting Hankel to dyadic block gives Hankel
- Restricting Kerdock to subfield gives Kerdock.
  - No large dot products (v. $\leq k^8$ for Hankel)
  - More efficient algorithm
  - Bit-by-bit extension won’t work—we have new algorithm.
  - Can assume existence of subfields of the correct size.
Subfields

Need subfield of size $2^f \geq k^2$, to get $(1/k)$-incoherence.

- So need $f|n$.
  - $n \rightarrow fn$, so $N \rightarrow N^f$.
  - Extend signal via trace function.
  - Cost factor $\log(N) \rightarrow \log(N^f) \leq \log^2(N)$.
Smaller Subfields

At most $O(k)$ coef{s with $|\langle \varphi, s \rangle|^2 \geq (1/k) \|s\|^2$ in subfield of size $2^f = k^2$.

Now suppose $s = \sum_{\lambda \in \Lambda} c_{\lambda} \varphi_{\lambda} + \nu$, where

- $|\Lambda| = k$
- $|c_{\lambda}| \approx 1$
- $c_{\lambda}$ random with $E[c_{\lambda}] = 0$
  - $\Diamond$ E.g., $c_{\lambda} = \pm 1$ for message 0 and $\pm i$ for message 1.
- $\nu$ Gaussian with $\|\nu\|^2 \leq k$.

(Plausible in wireless applications.) Then...
Smaller Subfields, cont’d

(...assuming random unit coefficients and noise.)

For subfield size $k$, there are constants $c_1 > c_2$ with

$$
\begin{cases}
|\langle \varphi_\lambda, s \rangle|^2 > \left(\frac{c_1}{k}\right)\|s\|^2, & \lambda \in \Lambda; \\
|\langle \varphi_\lambda, s \rangle|^2 < \left(\frac{c_2}{k}\right)\|s\|^2, & \lambda \notin \Lambda.
\end{cases}
$$

- So list decoding works. (Ongoing work by Lepak.)
Extension: Delsarte-Goethals

- Hierarchy of codes between RM(1) and RM(2).
- Sum of two matrices has rank at least $n - g$.
  - Dickson: Get incoherent codebook
- Number of codewords between $N^2$ (Kerdock) and $N^{\Theta(\log(N))}$ (RM(2)).
Recap

- We construct a Kerdock code as a subcode of Hankel.
- We give a list-decoding algorithm for Hankel.
- (Corollary) We give a list-decoding algorithm for Kerdock.
- (Corollary) Since Kerdock is $\mu$-incoherent for small $\mu$, we get a sparse recovery algorithm for Kerdock.