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 $\mathbf{C}^N$
  - ◇ Parameter k.
- Given vector ("signal") $s \in \mathbf{C}^N$.
  - ◇ Accessed by *sampling*: query y, learn $s(y)$.
- **Goal:**

$$\boxed{\text{Quickly find list of } \lambda \text{ such that } |\langle \varphi_\lambda, s \rangle|^2 \geq (1/k)\,\|s\|^2.}$$

- For some codebooks, leads to sparse approximation:
  - ◇ Small Λ with $\tilde{s} = \sum_{\lambda \in \Lambda} c_\lambda \varphi_\lambda \approx s$.

Definitions of Reed-Muller (-like) Codes

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

- Second-order Reed-Muller, RM(2):

$$\varphi_{P,\ell}(y) = i^{y^T P y + 2\ell^T y}.$$

- Hankel, Kerdock codes: limited allowable P's.
  - ◇ Hankel: P 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 Λ of Hankel λ such that $|\langle \varphi_\lambda, s\rangle|^2 \geq (1/k)\|s\|^2$
  - ◇ ...in time $\mathrm{poly}(k \log(N))$.
- 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 $\mathbf{Z}_2$. (Restrict multi-variate polynomial to random line, getting univariate polynomial. But low-degree univariate polys over $\mathbf{Z}_2$ are not interesting.)
- *General* sparse recovery [Gilbert-Muthukrishnan-S-Tropp, ...] requires time $\text{poly}(2^n) \gg \text{poly}(k, n)$ and/or space $\text{poly}(2^n)$

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 $\varphi_{P,0} \cdot \mathrm{RM}(1)$ for Kerdock P. ("Mutually-Unbiased Bases.")

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

- **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 φ_λ from her set.
  - ◇ Message is encoded in choice of codeword and/or coefficient c_λ.
- Receiver gets $\sum_\lambda c_\lambda \varphi_\lambda$.
- Decoder recovers all (λ, c_λ)'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 RM(2)$= \{\varphi_{P,\ell}\}$ in which P is Hankel. [Calderbank-Gilbert-Levchenko-Muthukrinshnan-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 ℓ: $\varphi_{P,\ell}$.

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 $\mathbf{Z}_2[t]$, *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 paramter 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 $h(t) = t^3 + t + 1 = t^3 + 0t^2 + 1t + 1$.

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

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 $\mathrm{Tr} : \mathbf{F}(2^n) \to \mathbf{F}(2)$ is given by
  $\mathrm{Tr}(x) = x + x^2 + x^4 + x^8 + \cdots + x^{2^{n-1}}$.
- Recall h is primitive polynomial; $h(\xi) = 0$.
- $(K_\alpha)_{j,k} := \mathrm{Tr}(\alpha\xi^{j+k})$ ("trace-Kerdock" matrix, for $\alpha \in \mathbf{F}(2^n)$)

Three lemmas, one-line proofs:

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

Facts about Trace

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

- $\mathrm{Tr}(x + y) = \mathrm{Tr}(x) + \mathrm{Tr}(y)$.
- $\mathrm{Tr}(x)^2 = \mathrm{Tr}(x^2) = \mathrm{Tr}(x)$.
  - ◇ $\mathrm{Tr}(x)$ satisfies $y^2 + y = 0$.
  - ◇ $\mathrm{Tr}(x) \in \{0, 1\}$.

Trace-Kerdocks are non-Singular

Lemma: Trace-Kerdocks are non-Singular

$K_\alpha = V^T D_\alpha V$ over $\mathbf{F}(2^n)$, where
$D_\alpha = \mathrm{diag}(\alpha, \alpha^2, \alpha^4, \alpha^8, \ldots, \alpha^{2^{n-1}})$ 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_α is over $\mathbf{F}(2)$, so $\det(K_\alpha) \in \mathbf{F}(2)$ over big field.

Trace-Kerdocks are lf-Kerdock

Lemma: Trace-Kerdocks are lf-Kerdock.

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

Feedback rule:

$$\begin{aligned}
\mathrm{Tr}(\alpha\xi^{j+k+n}) &= \mathrm{Tr}\left(\alpha\xi^{j+k}\sum_{\ell<n} h_\ell \xi^\ell\right) \\
&= \sum_{\ell<n} h_\ell \mathrm{Tr}\left(\alpha\xi^{j+k}\xi^\ell\right),
\end{aligned}$$

so feedback rule is satisfied.

lf-Kerdocks are Trace-Kerdock

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)

$\phi_{1011} = \varphi_{1000} \cdot \varphi_{0010} \cdot \varphi_{0001}$ is signal of length $2^4 = 16$.

Start with $\phi_{0000} \cong \mathbf{1}$ and flip bits, in dyadic blocks.

φ_{0000}	++++	++++	++++	++++
Flip	v v	v v	v v	v v
φ_{1000}	+−+−	+−+−	+−+−	+−+−
Flip		vvvv		vvvv
φ_{1010}	+−+−	−+−+	+−+−	−+−+
Flip			vvvv	vvvv
φ_{1011}	+−+−	−+−+	−+−+	+−+−

RM(1) Recovery

E.g., [Kushilevitz-Mansour]

Want ℓ such that $|\langle s, \varphi_\ell\rangle|^2 \geq (1/k)\|s\|^2$.

For $j \leq n$, maintain candidate list L_j for first j bits of ℓ.

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.

RM(1) Recovery, No False Negatives

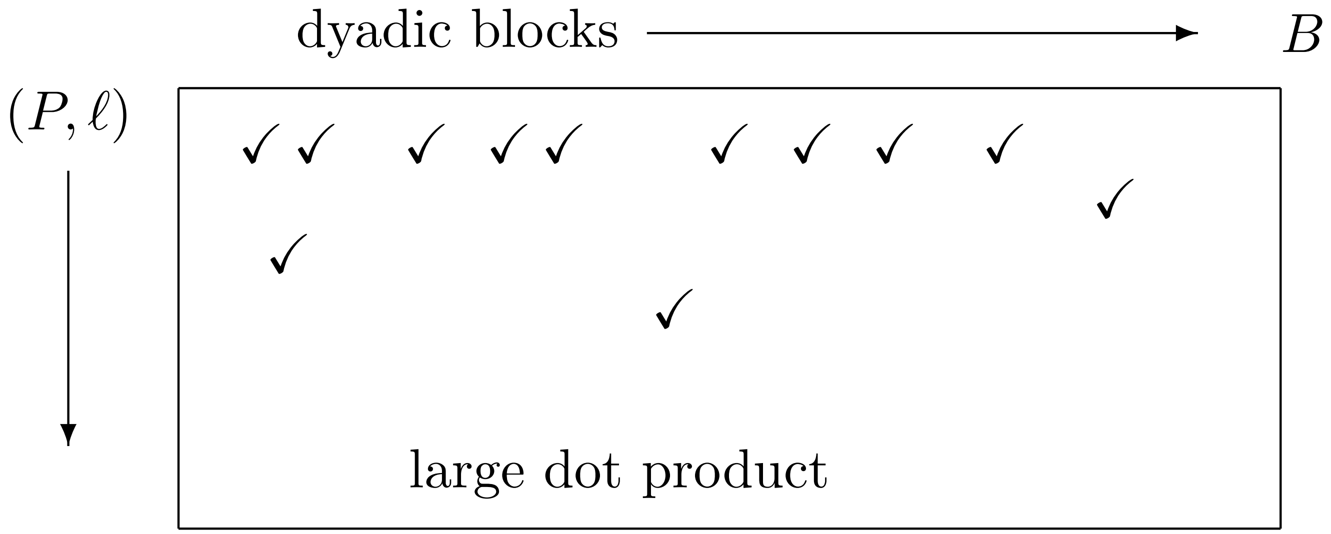
Signal $s \in \mathbf{C}^{16}$; candidate φ_ℓ with $\ell = 01**$.

Signal s	++--	++--	++-3	i-++
φ_{0100}	++--	++--	++--	++--
φ_{01**}	++--	$\pm 1\cdot$++--	$\pm 1\cdot$++--	$\pm 1\cdot$++--

- $|\langle \varphi_{0100}, s\rangle|^2$ 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 φ_{01**} since $\geq 1/O(k)$ blocks have square dot product $\geq 1/O(k)$, as estimated by sampling.
- Alternative view of dot product: $|\langle s, \varphi\rangle|^2 = |\langle s\varphi^*, \pm\mathbf{1}\rangle|^2$.

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 $\leq k^2$.



Hankel Recovery

Want P, ℓ such that $|\langle s, \varphi_\ell \rangle|^2 \geq (1/k)\,\|s\|^2$.

Find P, then use KM to find ℓ with large $|\langle s, \varphi_{P,\ell} \rangle|^2 = |\left\langle s\varphi^*_{P,0}, \varphi_{0,\ell} \right\rangle|^2$.

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.

$$P = \left(\begin{array}{c|c} P' & a \\ \hline a & b \\ & \\ & \end{array}\right)$$

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 $|\left\langle s'\varphi^*_{P',0}, \varphi_{0,\ell'}\right\rangle|^2$ 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: $\mathrm{rank}(P + P')$ 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 $\mathrm{rank}(P + P')$ low.
- Put it all together:
  - ◇ Theorem: Given signal s and parameter k, there are at most $\mathrm{poly}(k)$ Hankel vectors $\varphi_{P,\ell}$ with $|\langle \varphi_{P,\ell}, s \rangle|^2 \geq (1/k)\, \|s\|^2$.

Dickson's Theorem

If $(P,\ell) \neq (P',\ell')$, then

$$|\langle \varphi_{P,\ell}, \varphi_{P',\ell'} \rangle| \leq 2^{-\mathrm{rank}(P+P')/2}.$$

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

Bigger rank $\Rightarrow$ vectors are closer to orthogonal.

Dickson for Kerdock, proof

$$
\begin{aligned}
N^2 \langle \varphi_{P,\ell}, \varphi_{0,0} \rangle^2 &= \sum_{y,z} i^{y^T P y + 2\ell^t y + z^T P z + 2\ell^T z} \\
&= \sum_{y,w} i^{w^T P w + 2\ell^t w + 2y^T P(w+y)}, \quad w = y + z \\
&= \sum_{y,w} i^{w^T P w + 2\ell^t w + 2y^T P w + 2d^T y}, \quad d = \mathrm{diag}(P) \\
&= \sum_{w} i^{w^T P w + 2\ell^t w} \sum_{y} i^{2(w^T P + d^T) y} \\
&= N \sum_{w} i^{w^T P w + 2\ell^t w} \delta(w^T P, d^T), \quad P w = d \\
&= N i^{d^T P^{-1} d + 2\ell^t P^{-1} d}.
\end{aligned}
$$

Thus $| \langle \varphi_{P,\ell}, \varphi_{0,0} \rangle |^2 = 1/\sqrt{N}$.

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:

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

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:

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

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

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:

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

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:

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

Learn f by linear combination.

Few Low-Rank Hankels

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

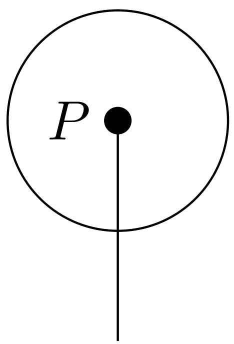
Proof: Suppose column 3 is a linear combination C of columns 0,1,2. Then C and positions 0, 1, 2 in top row determine the top half of the matrix:

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

Fill in by Hanklicity.

Hankel Vectors

Space of Hankel Vectors:

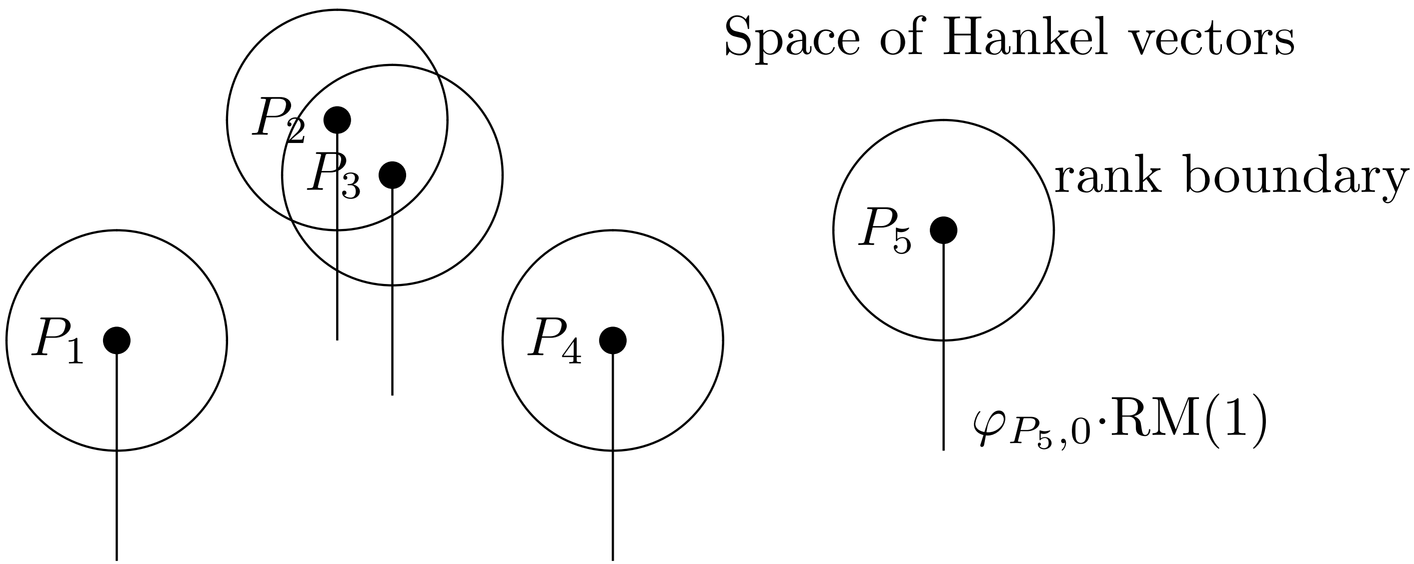


Dot: vector $\varphi_{P,0}$.

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

Stick: vectors $\varphi_{P,\ell}$, as ℓ varies.

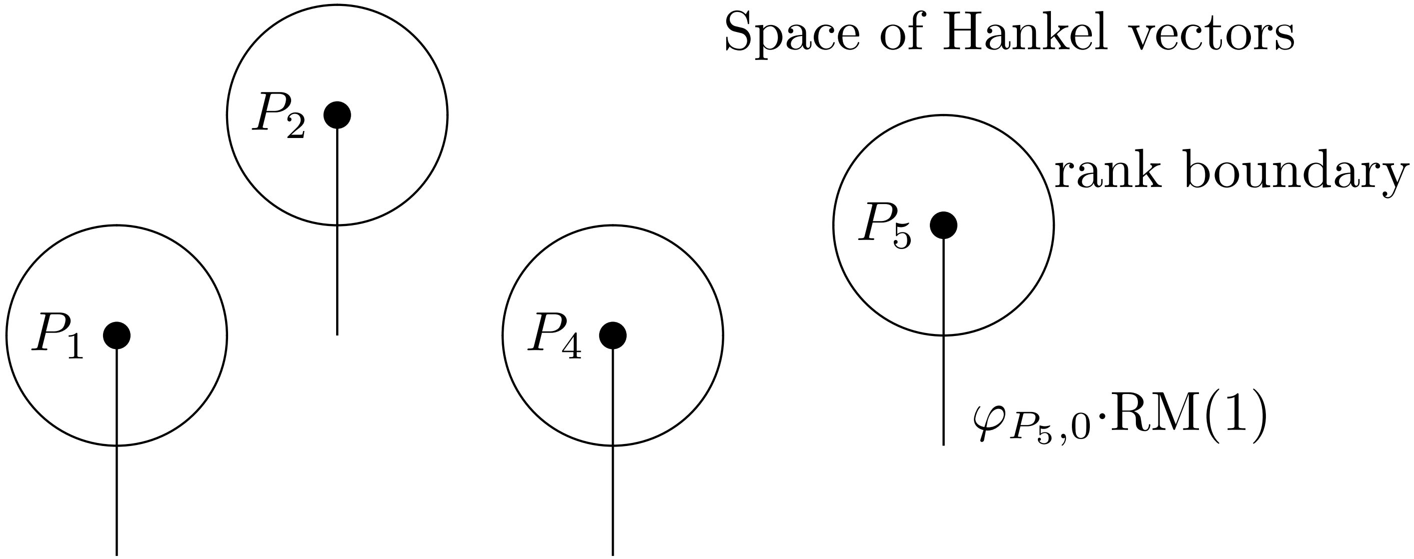
Few Large Hankel Coefficients



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.

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 reprentation to length-N vector in time 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

Alternative Algorithms

A faster alternative to KM first permutes the RM(1) labels $\ell^T \to \ell^T R$:

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

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_α to $K_{\alpha r^2}$—preserves Kerdock structure.
- For each ℓ, $\ell^T R$ is uniform over $\mathbf{Z}_2^n$ 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 sub*field* 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 $f|n$.
  - ◇ $n \to fn$, so $N \to N^f$.
  - ◇ Extend signal via trace function.
  - ◇ Cost factor $\log(N) \to \log(N^f) \leq \log^2(N)$.

Smaller Subfields

At most $O(k)$ coefs 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_λ random with $E[c_\lambda] = 0$
  - ◇ E.g., $c_\lambda = \pm 1$ for message 0 and $\pm i$ for message 1.
- ν 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

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

- 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 μ-incoherent for small μ, we get a sparse recovery algorithm for Kerdock.