# A Fast Reconstruction Algorithm for Deterministic Compressive Sensing using Second Order Reed-Muller Codes

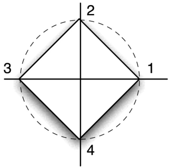
Stephen Howard[1] Robert Calderbank[2] Stephen Searle[3]

[1]Defence Science & Technology Organisation
PO Box 1500
Edinburgh 5111, Australia.

[2]Electrical Engineering
Princeton University
NJ 08544, USA.

[3]Electronic & Electrical Engineering
The University of Melbourne
Victoria 3010 Australia.

# $D_4$: The Symmetry Group of the Square




Generated by matrices $x = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$

$xz = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ anticlockwise rotation by $\dfrac{\pi}{2}$

$z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ reflection in the horizontal axis

$D_4$ is the set of eight $2 \times 2$ matrices $\varepsilon\, D(a, b)$ given by

$$\varepsilon\, D(a, b) = \varepsilon \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)^a \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)^b \text{ where } \varepsilon = \pm 1 \text{ and } a, b = 0 \text{ or } 1.$$

$x^2 = z^2 = I_2$

$$\left.\begin{aligned} zx &= \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \\ xz &= \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \end{aligned}\right] \; xz = -zx$$

# The Hadamard Transform

$H_2 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} + & + \\ + & - \end{smallmatrix}\right)$ reflects the lattice of subgroups across the central axis of symmetry

$$\begin{tikzcd}[column sep=small]
 & & D_4 \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & & \\
 & \langle \pm x \rangle \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \langle xz \rangle \arrow[d, dash] & \langle \pm z \rangle \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\langle x \rangle \arrow[drr, dash] & \langle -x \rangle \arrow[dr, dash] & \langle -I_2 \rangle \arrow[d, dash] & \langle -z \rangle \arrow[dl, dash] & \langle z \rangle \arrow[dll, dash] \\
 & & \langle I_2 \rangle & &
\end{tikzcd}$$

$$H_2^2 = I_2 \text{ and } H_2^{-1} = H_2$$

$$H_2 x H_2 = z$$

$$H_2 z H_2 = x$$

$$H_2[\varepsilon x^a z^b] H_2 = \varepsilon (H_2 x^a H_2)(H_2 z^b H_2) = \varepsilon z^a x^b = (-1)^{ab} x^b z^a$$

$$H_2[\varepsilon D(a, b)] H_2 = (-1)^{ab} \varepsilon D(b, a)$$

# The Heisenberg-Weyl Group $\mathcal{W}(\mathbb{Z}_2^m)$

$\mathcal{W}(\mathbb{Z}_2^m)$ is the $m$-fold Kronecker product of $D_4$ extended by $iI_{2^m}$.

$i^\lambda p_{m-1} \otimes \ldots \otimes p_0$ where $p_j = I_2, x, z,$ or $xz$ for $j = 0, 1, \ldots, m-1$

There are $2^{2m+2}$ elements, each represented by a pair of binary $m$-tuples

$$xz \otimes x \otimes z \otimes xz \otimes I_2 \leftrightarrow D(\overset{a}{11010}, \overset{b}{10110})$$

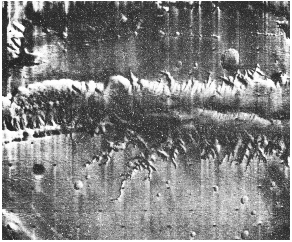
The operators $D(a, 0)$ are the *time shifts* of the binary world.

The operators $D(0, b)$ are the *frequency shifts* of the binary world.

# Walsh Functions

$$H_{2^m}^T = H_2^T \otimes \ldots \otimes H_2^T = H_{2^m}$$

Walsh functions of length $2^m$ are the rows (columns) of $H_{2^m}$ and their negatives.



Part of the Grand Canyon on Mars. This photograph was transmitted by the Mariner 9 spacecraft on January 19th, 1972 – gray levels are mapped to Walsh functions of length 32.

Walsh functions are the *sinusoids* of the binary world – eigenfunctions of the *time-shift operators* $D(a, 0)$.

# First Order Reed Muller Codes and Walsh Functions

Walsh functions are obtained by exponentiating codewords in the first order Reed Muller code.

**Example** $(m = 3)$ $RM(1,3)$

$$
\begin{array}{cc}
(\gamma, b) & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \\
 & v
\end{array}
= \begin{array}{c} (\ldots\ldots b.v + \gamma \ldots\ldots) \\ | \\ (-1)^{\gamma}(-1)^{b.v} = \varepsilon(-1)^{b.v} \end{array}
$$

**Symmetry:** Focus on orthonormal bases of eigenvectors for maximal commutative subgroups.

$$
\begin{tikzcd}
\text{Maximal Commutative Subgroup} & X = \{\varepsilon D(a,0)\} \arrow[r, leftrightarrow, "H_{2^m}"] \arrow[d, dash] & Z = \{\varepsilon D(0,b)\} \arrow[d, dash] \\
\text{Orthonormal Basis} & \text{Walsh Sequences} \arrow[r, leftrightarrow, "H_{2^m}"] & \text{Dirac Sequences}
\end{tikzcd}
$$

# Nonlinear Decoding of $RM(1, m)$

$$(\gamma, b) \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} = \underbrace{\left( \underbrace{\quad} \mid \underbrace{\quad} \mid \underbrace{\quad} \mid \underbrace{\quad} \right)}_{\substack{\text{each pair of entries} \\ \text{sums to } b_0}}$$

$e_j$: binary vector with a 1 in position $j$ and zeros elsewhere

**Input:** $[(-1)^{b.x}]_{x \in \mathbb{Z}_2^m}$

**Time Shift:** $[(-1)^{b.(x+e_j)}]_{x \in \mathbb{Z}_2^m}$

**Sample by Sample Multiply:**

$$((-1)^{b.e_j}, \ldots, (-1)^{b.e_j})$$

$\uparrow$ codeword in $RM(0, m)$.

# Generating Orthonormal Bases of $\mathbb{C}^N$

**Remark:** One basis for each coset of $RM(1, m+1)$ in $RM(2, m+1)$

Maximal Commutative Subgroup

Orthonormal Basis

$$\begin{tikzcd} X \arrow[r] \arrow[d, dash] & X_P = d_P^{-1} X d_P \arrow[d, dash] \\ H_{2^m} \arrow[r] & H_{2^m} d_P \end{tikzcd}$$

$$d_P = \mathrm{diag}[i^{vPv^T}]$$

**Example:** $m = 3$, $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

$$H_8 = \frac{1}{2\sqrt{2}} \left[ \begin{array}{cccc|cccc} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ \hline + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{array} \right] \quad d_P = \begin{array}{cl} \left[ \begin{array}{cccc|cccc} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ \hline & & & & i & & & \\ & & & & & i & & \\ & & & & & & -i & \\ & & & & & & & i \end{array} \right] & \begin{array}{l} 000 \\ 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{array} \end{array}$$

# Second Order Reed-Muller Functions

- Second Order Reed-Muller Functions: For $\mathbf{b} \in \mathbb{Z}_2^m$ and binary symmetric matrix $P$

$$\begin{aligned}\phi_{P,\mathbf{b}}(\mathbf{a}) &= i^{(2\mathbf{b}+P\mathbf{a})^T\mathbf{a}} \\ &= i^{\mathbf{a}^T P\mathbf{a}}(-1)^{\mathbf{b}^T\mathbf{a}}\end{aligned}$$

  *These are the chirps of the binary world.*

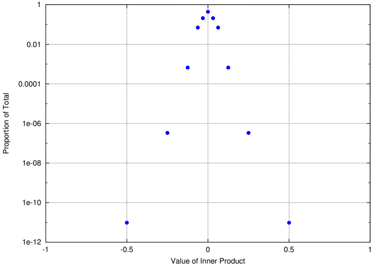
- For each $P$ we have an ON basis of $2^m$ vectors. There are $2^{m(m+1)/2}$ such matrices $P$.
- Matrices $P$ with zero diagonal give real ($\pm 1$) vectors. There are $2^{m(m-1)/2}$ such matrices.
- Fix vector $\psi$ in the ON basis corresponding to $P$ and let $\phi$ vary over the ON basis corresponding to $Q$, then if $\ell = \text{rank}(P-Q)$, (Phases $\pm 1$ and $\pm i$ occur equally often)

$$|(\psi,\phi)| = \begin{cases} \frac{1}{\sqrt{2^\ell}}, & 2^\ell \text{ times}, \\ 0, & 2^m - 2^\ell \text{ times}. \end{cases}$$

# Distribution of Inner Products

| Value | Proportion |
|---|---|
| $-1/2^h$ | $v_{2h}2^{2h-1}/2^{m(m+1)/2}$ |
| $0$ | $1-2\sum_h v_{2h}2^{2h-1}/2^{m(m+1)/2}$ |
| $1/2^h$ | $v_{2h}2^{2h-1}/2^{m(m+1)/2}$ |

for $h = 1, \cdots, [m/2]$.




$v_r$ is the number of zero diagonal $m \times m$ binary symmetric matrices with rank $r$ is given by

$$v_{2h} = 2^{h(h-1)} \frac{(2^m-1)(2^{m-1}-1)\dots(2^{m-2h+1}-1)}{(2^{2h}-1)(2^{2h-2}-1)\dots(2^2-1)}$$

and $v_{2h+1} = 0$.

# Nonlinear Decoding: Second Order Reed Muller Codes

The method of chirp detection:

$$\text{Initial codeword: } \left[ i^{xAx^T + 2bx^T} \right]_{x \in \mathbb{Z}_2^m}$$

Time-shift by $v$:

$$\text{Shift codeword: } \left[ i^{(x+v)A(x+v)^T + 2b(x+v)^T} \right]_{x \in \mathbb{Z}_2^m}$$

Sample by sample multiply:

$$i^{vAv^T + 2bv^T} \left[ (-1)^{vAx^T} \right]_{x \in \mathbb{Z}_2^m}$$

$\uparrow$ first order Reed Muller codeword.

# Compressed Sensing

- Compressive sensing: Apply an $n \times N$ matrix to a *sparse* vector $x \in \mathbb{R}^N$ to get $n$ measurements

$$y = \Phi x$$

- Restricted Isometry Property (RIP): A matrix is said to have the RIP of order $k$ if $\Phi$ acts as a near isometry on *all* $k$ sparse vectors.
- Reconstruction: Finding $x$ from the measured or observed data $y = \Phi x$ requires a search over $\mathcal{F}(y)$. The successful decoders take advantage of the geometry of $\mathcal{F}(y)$. The most prominent example, called Basis Pursuit (BP) decodes $y$ by taking the vector $x^* \in \mathcal{F}(y)$ with smallest $\ell_1$ norm.

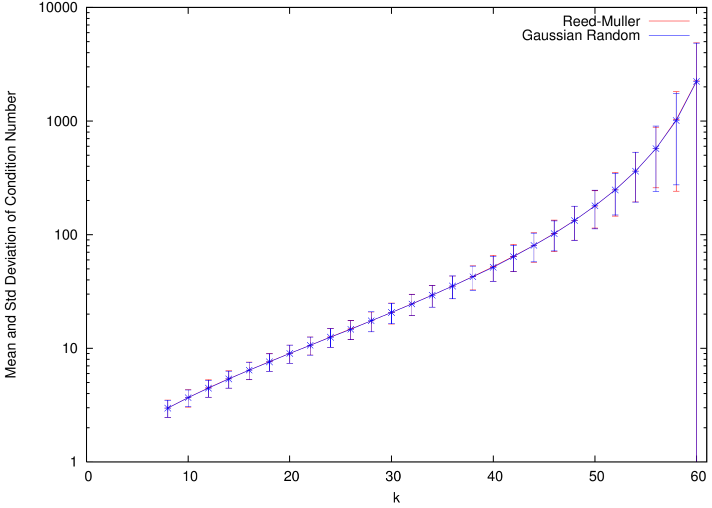
# Some Philosophy

- Lesson from coding theory: MacKay makes the point that in coding theory distance is not everything and he argues that the minimum distance of a code is not of fundamental importance to the goal of achieving reliable communication over noisy channels. Reliable communication is achieved by focusing on typical rather than worst case performance, and this involves constraining the spectrum of all possible distances rather than simply the minimum distance.
- RIP guarantees worst case performance, and in this sense it plays the same role as minimum distance in coding theory.
- Approach: Construct a compressive sensing matrix $\Phi$ that comes by design with a very fast reconstruction algorithm. This matrix may not be RIP uniformly with respect to all $k$-sparse vectors, but that acts as a near isometry on $k$-sparse vectors with very high probability.

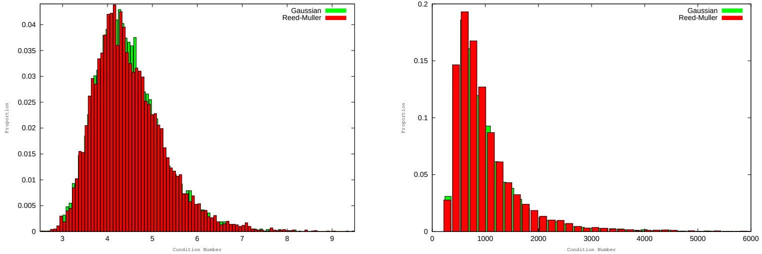
# Reed-Muller Compressed Sensing Matrix

- Matrix $\Phi$ with $2^m$ rows (measurements) and columns which are all real $2^{\text{nd}}$-order Reed-Muller functions.
- Columns are labeled by a zero diagonal $m \times m$ binary symmetric matrix $P$ and a binary vector $\mathbf{b}$.
- There are $2^{m(m+1)/2}$ columns!
- The set of columns forms a tight frame since it consists of $2^{m(m-1)/2}$ ON bases.

# Distribution of Condition Numbers $m = 6$



# Distribution of Condition Numbers (2)




Distribution of condition numbers for $k$-Gram matrices of Gaussian random ($\sim \mathcal{N}(0,1)$) and Reed-Muller compressed sensing matrices for $m = 6$ with $k = 10$ and $k = 56$

# Fast Reconstruction Algorithm

- Suppose $x$ is a $k$-sparse vector. The reconstruction problem takes the form

$$y = \Phi x = \sum_{j=1}^{k} c_k \phi_{P_j,\mathbf{b}_j}$$

- The Second-Order Reed-Muller codes have the property

$$\phi_{P,\mathbf{b}}(\mathbf{a})\phi_{P,\mathbf{b}}(\mathbf{a}+\mathbf{e}) = (-1)^{\mathbf{b}^T\mathbf{e}}(-1)^{\mathbf{e}^T P\mathbf{a}}$$

  i.e., shift and multiply gives a Walsh function. *We can identify a $2^{nd}$-order RM function with $m$ fast Hadamard transforms*

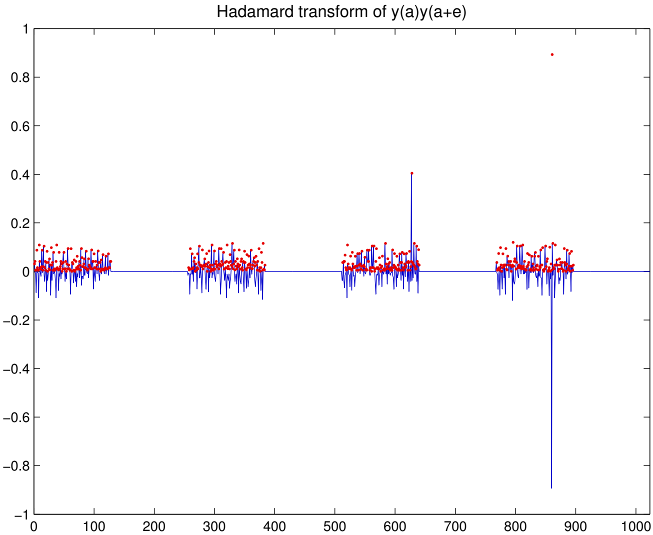
- Approach: Compute

$$y(\mathbf{a})y(\mathbf{a}+\mathbf{e}) = \sum_{j=1}^{k} c_k^2 \phi_{P_j,\mathbf{b}_j}(\mathbf{a})\phi_{P_j,\mathbf{b}_j}(\mathbf{a}+\mathbf{e}) + \text{Chirps},$$

  then take Hadamard transform and search.

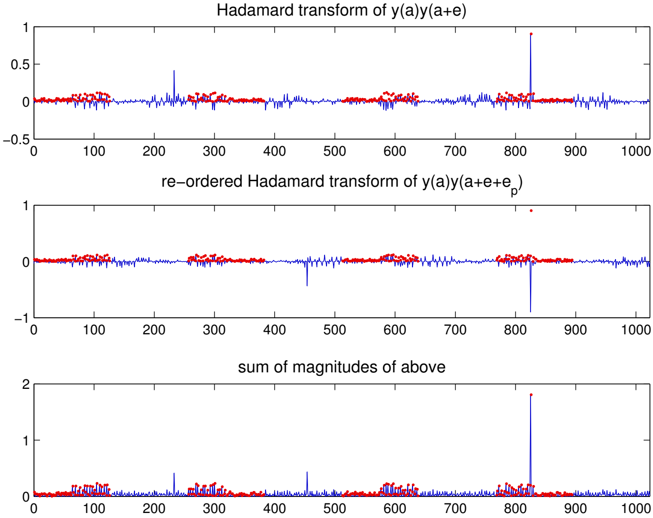
# Fast Reconstruction Algorithm (2)

- Example $n = 2^{10}$ and $N = 2^{55}$



# Fast Reconstruction Algorithm (3)

- Example $n = 2^{10}$ and $N = 2^{55}$

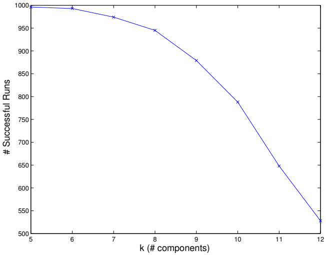


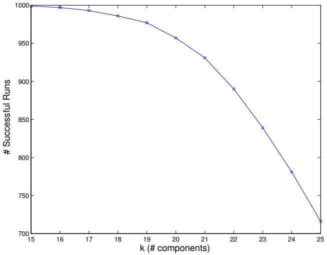
# Fast Reconstruction Algorithm (4)

- Number of measurements rule of thumb for Basis Pursuit

$$n > k \log_2 \left(1 + \frac{N}{k}\right)$$

- $k = 20$ and $N = 2^{55}$ implies $n > 1014$ (we have $n = 1024$)
  $k = 7$ and $N = 2^{36}$ implies $n > 232$ (we have $n = 256$)



# Discussion

- In terms of accuracy of reconstruction the number of common columns of the P matrices should be kept to a minimum.
- As the $\Phi_{RM}$ matrices have an abundance of columns, we can afford to place conditions of the P matrices.
- Placing restriction on the rank of the differences of the P matrices, that is, taking all $P \in DG_{2h}$, the performance of the algorithm is increases as the number of columns in $\Phi$ decreases.
- A detailed analysis of the algorithm and the proof that the matrix $\Phi_{RM}$ acts as a near isometry on $k$-sparse vectors with very high probability will be given elsewhere.