Bounds on constant weight codes

Punarbasu Purkayastha
Joint work with Alexander Barg

Hamming space

- Hamming space: $\mathbb{F}_2^n = \{0,1\}^n$
- $d(\boldsymbol{u},\boldsymbol{v}) = \#\{u_i \neq v_i : i = 1,\ldots,n\}$
- $\mathcal{S}_w = \{\boldsymbol{x} \in \mathbb{F}_2^n : d(\boldsymbol{x},\mathbf{0}) = w\}$
- Code $\mathcal{C} \subset \mathcal{S}_w$. Parameters: (n, M, d, w)

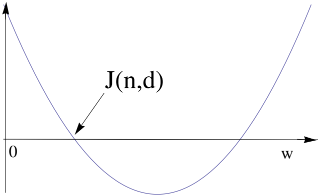
- $A(n,d)$: maximum size of $\mathcal{C}(n,M,d)$ in $\mathbb{F}_2^n$
- $A(n,d,w)$: maximum size of $\mathcal{C}(n,M,d,w)$ in $\mathcal{S}_w$

Bounds on constant weight codes

- Johnson bound '62: $\mathcal{C}(n, M, d, w)$ in $\mathcal{S}_w$

$$M \leq \frac{dn}{dn - 2wn + 2w^2}$$

- Johnson Radius:



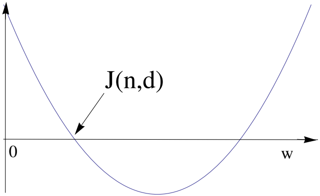
$$J(n,d) = \left\lfloor \frac{n}{2}\left(1 - \sqrt{1 - 2\frac{d}{n}}\right) \right\rfloor$$

Bounds on constant weight codes

- Johnson bound '62: $\mathcal{C}(n, M, d, w)$ in $\mathcal{S}_w$

$$M \leq \frac{dn}{dn - 2wn + 2w^2}$$

- Johnson Radius:



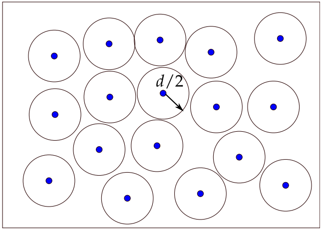
$$J(n,d) = \left\lfloor \frac{n}{2}\left(1 - \sqrt{1 - 2\frac{d}{n}}\right) \right\rfloor$$

$$A(n,d) \leq \frac{2^n A(n,d,w)}{\binom{n}{w}}$$

- Use Bassalygo-Elias inequality to estimate bounds on codes
- Johnson radius: the radius till which we have polynomial sized list for any code under list decoding

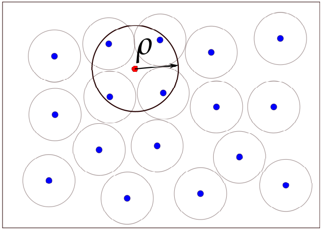
List decoding

- Unique decoding: decode up to $d/2$, output ≤ 1 codeword



List decoding

- Unique decoding: decode up to $d/2$, output ≤ 1 codeword
- List decoding:
  - Output at most L codewords
  - Desired size of L is at most polynomial in n



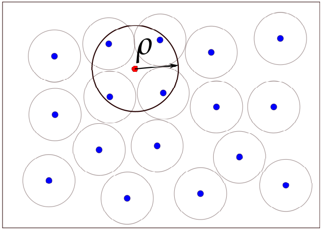
List decoding

- Unique decoding: decode up to $d/2$, output ≤ 1 codeword
- List decoding:
  - Output at most L codewords
  - Desired size of L is at most polynomial in n
  - Let $\delta = \frac{d}{n}$, $J(\delta) = \frac{J(n,d)}{n}$.

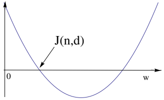
$$J(\delta) = \frac{1}{2}\left(1 - \sqrt{1 - 2\delta}\right)$$

$$\rho^{\mathsf{poly}}(\delta) \geq J(\delta)$$

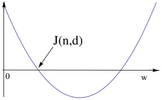
  - $\rho^{\mathsf{poly}}(\delta) \leq J(\delta) + 10^{-50}$ (Guruswami, Spharlinski '03)



- Johnson bound not valid beyond $J(n, d)$
- Many improvements known. (Agrell, et.al. '00)



- Johnson bound not valid beyond $J(n, d)$
- Many improvements known. (Agrell, et.al. '00)
- We provide two new bounds
    - Valid for some values beyond Johnson radius
    - Better than Johnson bound at points close to Johnson radius



Johnson bound

- Johnson bound '62: $\mathcal{C}(n, M, d)$ in $\mathcal{S}_w$

$$M \leq \frac{dn}{dn - 2wn + 2w^2}$$

- Can be proved by averaging argument.
- Map $0 \mapsto 1, \quad 1 \mapsto -1$. Then

$$d = n - {}^1\!/_2 \max_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l|$$

Johnson bound

- Johnson bound '62: $\mathcal{C}(n, M, d)$ in $\mathcal{S}_w$

$$M \leq \frac{dn}{dn - 2wn + 2w^2}$$

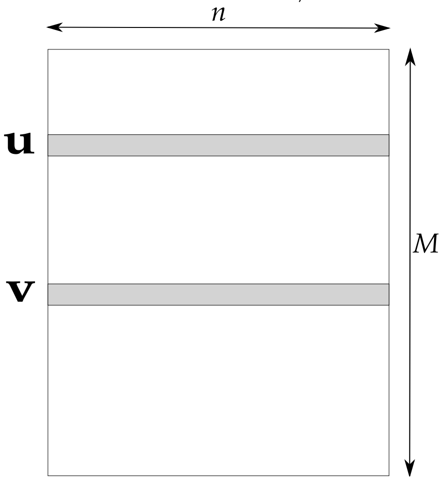
- Can be proved by averaging argument.
- Map $0 \mapsto 1, \quad 1 \mapsto -1$. Then

$$d = n - {}^1\!/_2 \max_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l|$$

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \frac{1}{\binom{M}{2}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v} \in \mathcal{C} \\ \boldsymbol{u} \neq \boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l|$$

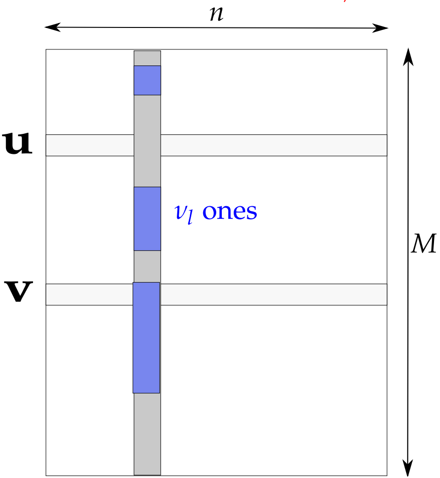
Johnson bound

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \frac{1}{\binom{M}{2}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l|$$



Johnson bound

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \frac{1}{\binom{M}{2}} \sum_{l=1}^{n} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} |u_l + v_l|$$



Johnson bound

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \frac{1}{\binom{M}{2}} \sum_{l=1}^{n} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} |u_l + v_l|$$

$$\frac{1}{\binom{M}{2}} \sum_{l=1}^{n} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} |u_l + v_l| = \frac{1}{\binom{M}{2}} \sum_{l=1}^{n} 2\left(\binom{\nu_l}{2} + \binom{M-\nu_l}{2}\right)$$

$$\geq \min_{\substack{0\leq\nu_l\leq M\\ \sum_l \nu_l = Mw}} \frac{1}{\binom{M}{2}} \sum_{l=1}^{n} 2\left(\binom{\nu_l}{2} + \binom{M-\nu_l}{2}\right)$$

Bounds on constant weight codes

New bounds — Ideas:

- Max is greater than averaged L^2 norm:

$$\max_i c_i \geq \left(\frac{1}{N} \sum_{i=1}^{N} (c_i)^2 \right)^{1/2}$$

- Max is greater than weighted norm:

$$\max_i c_i \geq \sum_i g(i) c_i$$

Bound from L^2 norm

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \left(\frac{1}{\binom{M}{2}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \left(\sum_{l=1}^{n} |u_l + v_l| \right)^2 \right)^{1/2}$$

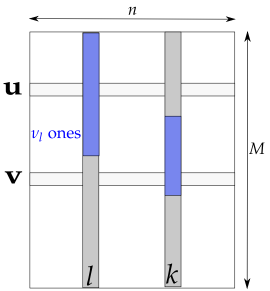
Bound from L^2 norm

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \left(\frac{1}{\binom{M}{2}} \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \left(\sum_{l=1}^{n} |u_l + v_l| \right)^2 \right)^{1/2}$$

Bound from L^2 norm

$$\sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \left(\sum_{l=1}^{n} |u_l + v_l|\right)^2 = M^2((n-2w)^2 - 2w^2) - n^2 M$$

$$+ 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$



$\zeta_{l,k} = \#\{(1,0)\} + \#\{(0,1)\}$

Bound from L^2 norm

$$\mathbf{X} \triangleq \big((\nu_l)_l, (\zeta_{l,k})_{l,k}\big)$$

$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2M + 2n\sum_{l=1}^{n}\nu_l^2 + 2\sum_{l=1}^{n-1}\sum_{k=l+1}^{n}\zeta_{l,k}^2$$

Bound from L^2 norm

$$\mathbf{X} \triangleq \big((\nu_l)_l, (\zeta_{l,k})_{l,k}\big)$$

$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2 M + 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$\geq \text{constant} + \min_{\nu_l, \zeta_{l,k}} 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$\sum_{l=1}^{n} \nu_l = Mw, \quad \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k} = Mw(n-w), \quad \zeta_{l,k} \leq \nu_l + \nu_k$$

Bound from L^2 norm

$$\mathbf{X} \triangleq \big((\nu_l)_l, (\zeta_{l,k})_{l,k}\big)$$

$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2 M + 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$\geq \text{constant} + \min_{\nu_l, \zeta_{l,k}} 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$\geq \text{constant} + \min_{\nu_l} \sum_{l=1}^{n} \nu_l^2 + \min_{\zeta_{l,k}} 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$\sum_{l=1}^{n} \nu_l = Mw, \quad \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k} = Mw(n-w),$$

Bound from L^2 norm

$$\mathbf{X} \triangleq \big((\nu_l)_l, (\zeta_{l,k})_{l,k}\big)$$

$$\begin{aligned} f(\mathbf{X}) &= M^2((n-2w)^2 - 2w^2) - n^2 M + 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2 \\ &\geq \text{constant} + \min_{\nu_l, \zeta_{l,k}} 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2 \\ &= \text{constant} + \min_{\nu_l} \sum_{l=1}^{n} \nu_l^2 + \min_{\zeta_{l,k}} 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2 \end{aligned}$$

$$\sum_{l=1}^{n} \nu_l = Mw, \quad \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k} = Mw(n-w),$$

Bound from L^2 norm

- Above minimization gives one bound, with real values

$$\nu = \nu_l = \frac{Mw}{n}, \qquad \zeta = \zeta_{l,k} = 2\frac{Mw(n-w)}{n(n-1)}$$

Bound from L^2 norm

- Above minimization gives one bound, with real values

$$\nu = \nu_l = \frac{Mw}{n}, \qquad \zeta = \zeta_{l,k} = 2\frac{Mw(n-w)}{n(n-1)}$$

- For integer $\nu_l,\ \zeta_{l,k}$, take the integer vector closest in euclidean distance to $\mathbf{X}^* = (\nu, ..., \nu, \zeta, ..., \zeta)$.
  (Note: The second inequality is no longer an equality)

Theorem

$$E = (n-2w)^2 + 4\frac{w^2(n-w)^2}{n(n-1)} - (n-d)^2 + \frac{1}{M^2}\big(2n^2\{\nu\}(1-\{\nu\}) + n(n-1)\{\zeta\}(1-\{\zeta\})\big),$$

For $E > 0$,

$$M \leq \left\lfloor \frac{d(2n-d)}{E} \right\rfloor.$$

Bound from weighted norm

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} g(\boldsymbol{u},\boldsymbol{v}) \sum_{l=1}^{n} |u_l + v_l|$$

- Weights:

$$g(\boldsymbol{u},\boldsymbol{v}) = \frac{\sum_{l=1}^{n} 2^{s|u_l+v_l|}}{\sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} 2^{s|u_l+v_l|}}$$

  Largest weights given to $\{\boldsymbol{u},\boldsymbol{v}\}$ with largest $\sum_l |u_l + v_l|$
- Let $s \to \infty$

Bound from weighted norm

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} g(\boldsymbol{u},\boldsymbol{v}) \sum_{l=1}^{n} |u_l + v_l| = \frac{f(\mathbf{X})}{g(\mathbf{X})},$$

$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2 M + 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$g(\mathbf{X}) = {}^1\!/_2\Big(M^2(n-2w) - nM + 2\sum_{l=1}^{n} \nu_l^2\Big)$$

Bound from weighted norm

$$\max_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} \sum_{l=1}^{n} |u_l + v_l| \geq \sum_{\substack{\boldsymbol{u},\boldsymbol{v}\in\mathcal{C}\\ \boldsymbol{u}\neq\boldsymbol{v}}} g(\boldsymbol{u},\boldsymbol{v}) \sum_{l=1}^{n} |u_l + v_l|$$

$$= \frac{f(\mathbf{X})}{g(\mathbf{X})},$$

$$\geq \min_{\nu_l,\ \zeta_{l,k}} \frac{f(\mathbf{X})}{g(\mathbf{X})}$$

$$\sum_{l=1}^{n} \nu_l = Mw, \quad \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k} = Mw(n-w), \quad \zeta_{l,k} \leq \nu_l + \nu_k$$

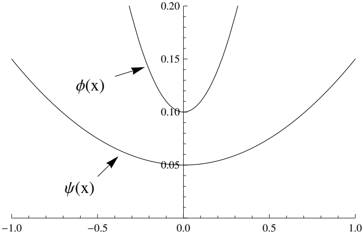
$$f(\mathbf{X}) = M^2((n-2w)^2 - 2w^2) - n^2 M + 2n \sum_{l=1}^{n} \nu_l^2 + 2 \sum_{l=1}^{n-1} \sum_{k=l+1}^{n} \zeta_{l,k}^2$$

$$g(\mathbf{X}) = {}^1\!/_2 \Big(M^2(n-2w) - nM + 2 \sum_{l=1}^{n} \nu_l^2 \Big)$$

Minimization of $f(\mathbf{X})/g(\mathbf{X})$

Lemma

- *Given* $\phi(\boldsymbol{x})$,$\psi(\boldsymbol{x})$ *and* $\phi(\boldsymbol{x}) - \psi(\boldsymbol{x})$ *are quadratic, positive, strongly convex, and are all minimized at* $\boldsymbol{x}_0$.
- *Then minimum of* $\frac{\phi(\boldsymbol{x})}{\psi(\boldsymbol{x})}$ *is obtained at either the boundary or at* $\boldsymbol{x}_0$.

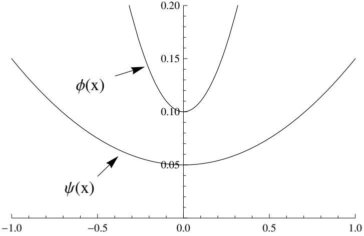


Minimization of $f(\mathbf{X})/g(\mathbf{X})$

Lemma

- *Given $\phi(\boldsymbol{x})$,$\psi(\boldsymbol{x})$ and $\phi(\boldsymbol{x}) - \psi(\boldsymbol{x})$ are quadratic, positive, strongly convex, and are all minimized at $\boldsymbol{x}_0$.*
- *Then minimum of $\frac{\phi(\boldsymbol{x})}{\psi(\boldsymbol{x})}$ is obtained at either the boundary or at $\boldsymbol{x}_0$.*

- $\phi(x) = x^2 + 0.1$,
  $\psi(x) = 0.1x^2 + 0.05$,
  $S = [-1, 1]$.
- $\phi(x), \psi(x), \phi(x) - \psi(x)$, and $\phi(x)/\psi(x)$ attain their minimum at $x = 0$.



Bound from weighted norm

Theorem

$$D = (d-2w)(1-2\frac{w}{n}) + (\frac{w}{n})^2\Big(4\frac{(n-w)^2}{n-1} - 2(n-d)\Big).$$

For $D > 0$*,*

$$M \leq \left\lfloor \frac{d}{D} \right\rfloor$$

Attains some points in table from Agrell, et.al. '00

$d=8$	5	6	7	8
12		4		
14		7	8	
15	6	10	15	
16		16		30
19	12			
20	16			
21	21			

$d=10$	6	9	10
19		19	
20			38
21	7		
26	13		

Relation to Sidelnikov's bound '75

- Improved Bassalygo-Elias bound asymptotically

$$A(n,d) \leq \frac{2^n A(n,d,w)}{\binom{n}{w}}$$

- Used "Inequality in the mean". Let $U_w \subset S^{n-1}$, $\mathcal{S}_w \to U_w$ via $0 \mapsto \frac{1}{\sqrt{n}},\ 1 \mapsto -\frac{1}{\sqrt{n}}$

Lemma

Let $C \subset U_w$ be a code of size M. Then

$$\frac{1}{M^2} \sum_{\boldsymbol{x},\boldsymbol{y} \in C} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \geq \frac{1}{|U_w|^2} \sum_{\boldsymbol{x},\boldsymbol{y} \in U_w} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \qquad (t \in \mathbb{N}).$$

Relation to Sidelnikov's bound '75

- Improved Bassalygo-Elias bound asymptotically

$$A(n,d) \leq \frac{2^n A(n,d,w)}{\binom{n}{w}}$$

- Used "Inequality in the mean". Let $U_w \subset S^{n-1}$, $\mathcal{S}_w \to U_w$ via $0 \mapsto \frac{1}{\sqrt{n}},\ 1 \mapsto -\frac{1}{\sqrt{n}}$

Lemma

Let $C \subset U_w$ be a code of size M. Then

$$\frac{1}{M^2} \sum_{\boldsymbol{x},\boldsymbol{y} \in C} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \geq \frac{1}{|U_w|^2} \sum_{\boldsymbol{x},\boldsymbol{y} \in U_w} \langle \boldsymbol{x}, \boldsymbol{y} \rangle^t \qquad (t \in \mathbb{N}).$$

- In our bounds, optimum is attained by random code
- Same bound (as L^2 norm) is obtained from above for $t = 2$.

Relation to List decoding bound of Blinovskii '86

- Problem: determine the maximum size of code $\mathcal{C}$ in $\mathbb{F}_2^n$ such that any ball of radius r has at most L codewords.
  Note: d is not used here.

Relation to List decoding bound of Blinovskii '86

- Problem: determine the maximum size of code $\mathcal{C}$ in $\mathbb{F}_2^n$ such that any ball of radius r has at most L codewords.
  Note: d is not used here.
- Blinovskii's (asymptotic) result:
  - determines upper bound on constant weight code in $\mathcal{S}_w$ with above property
  - use Bassalygo-Elias inequality:

$$|\mathcal{C}| \leq \frac{2^n M}{\binom{n}{w}}$$

  - Based on average frequency of $(L+1)$-tuples of alphabets in single column of codematrix

Relation to List decoding bound of Blinovskii '86

- Problem: determine the maximum size of code $\mathcal{C}$ in $\mathbb{F}_2^n$ such that any ball of radius r has at most L codewords.
  Note: d is not used here.
- Blinovskii's (asymptotic) result:
  - determines upper bound on constant weight code in $\mathcal{S}_w$ with above property
  - use Bassalygo-Elias inequality:

$$|\mathcal{C}| \leq \frac{2^n M}{\binom{n}{w}}$$

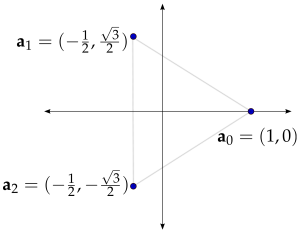
  - Based on average frequency of $(L+1)$-tuples of alphabets in single column of codematrix
- We can recover Blinovskii's result for $L = 2$ by analyzing the average frequency of $3-$tuple letters over pairs of columns.
- Improvement?

$q-$ary case

Extend bounds to $q-$ary by using mapping from Dunkl '76

Map $\mathbb{F}_q^n \to \mathbb{R}^{(q-1)n}$ via $i \mapsto \boldsymbol{a}_i$, vertices of a simplex

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \begin{cases} 1, & i = j, \\ -\frac{1}{q-1}, & i \neq j, \end{cases}$$

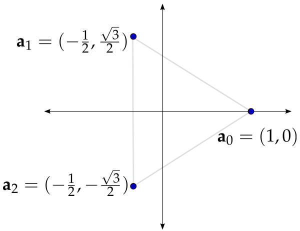


$q-$ary case

Extend bounds to $q-$ary by using mapping from Dunkl '76

Map $\mathbb{F}_q^n \to \mathbb{R}^{(q-1)n}$ via $i \mapsto \boldsymbol{a}_i$, vertices of a simplex

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \begin{cases} 1, & i = j, \\ -\frac{1}{q-1}, & i \neq j, \end{cases}$$



$$d(\boldsymbol{u}, \boldsymbol{v}) = \sum_{l=1}^{n} d(u_l, v_l) = \frac{q-1}{2q}\Big(4n - \sum_{l=1}^{n} \|\boldsymbol{u}_l + \boldsymbol{v}_l\|^2\Big),$$

Thank You!