Asymptotic Density and Counting Results for Lee Metric Codes

Nadja Willenborg

Institute of Computer Science, University of St. Gallen

February 22th, 2023

ALCOCRYPT 2023

Joint work with Anna-Lena Horlemann and Violetta Weger

Introduction

- The container method: upper bound for the number of independent sets, codes, sum free sets, monotone boolean functions etc.

Introduction

- The container method: upper bound for the number of independent sets, codes, sum free sets, monotone boolean functions etc.
- Idea: reduce an enumeration problem to evaluating the number of independent sets.

Introduction

- The container method: upper bound for the number of independent sets, codes, sum free sets, monotone boolean functions etc.
- Idea: reduce an enumeration problem to evaluating the number of independent sets.
- For $0 < t \leq 10\sqrt{n}$ the number of t-error correcting Hamming metric codes is at most $2^{(1+o(1))H_m(n,t)}$.

D. Dong, N. Mani and Y. Zhao "On the number of error correcting codes",arXiv preprint arXiv:2205.12363, 2022.

Introduction

- The container method: upper bound for the number of independent sets, codes, sum free sets, monotone boolean functions etc.
- Idea: reduce an enumeration problem to evaluating the number of independent sets.
- For $0 < t \leq 10\sqrt{n}$ the number of t-error correcting Hamming metric codes is at most $2^{(1+o(1))H_m(n,t)}$.

D. Dong, N. Mani and Y. Zhao "On the number of error correcting codes",arXiv preprint arXiv:2205.12363, 2022.

- **Question**: How can we count the number of t-error correcting Lee metric codes?

- The container method: upper bound for the number of independent sets, codes, sum free sets, monotone boolean functions etc.
- Idea: reduce an enumeration problem to evaluating the number of independent sets.
- For $0 < t \leq 10\sqrt{n}$ the number of t-error correcting Hamming metric codes is at most $2^{(1+o(1))H_m(n,t)}$.

D. Dong, N. Mani and Y. Zhao "On the number of error correcting codes",arXiv preprint arXiv:2205.12363, 2022.

- **Question**: How can we count the number of t-error correcting Lee metric codes?
- $\binom{X}{\leq t}$: family of subsets of X with at most t elements.

The Lee Metric

- We consider codes over $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.

The Lee Metric

- We consider codes over $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.
- We denote by $x := x + m\mathbb{Z}, 0 \leq x \leq m-1$ its residue classes.

The Lee Metric

- We consider codes over $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.
- We denote by $x := x + m\mathbb{Z}, 0 \leq x \leq m-1$ its residue classes.
- The Lee weight of $x \in \mathbb{Z}_m$ is defined by

$$\omega^{\mathsf{L}} : \{0, \ldots, m-1\} \to \{0, \ldots, \lfloor \frac{m}{2} \rfloor\}, \; x \mapsto \min\{x, m-x\}.$$

The Lee Metric

- We consider codes over $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.
- We denote by $x := x + m\mathbb{Z}, 0 \leq x \leq m-1$ its residue classes.
- The Lee weight of $x \in \mathbb{Z}_m$ is defined by

$$\omega^{\mathsf{L}} : \{0, \ldots, m-1\} \to \{0, \ldots, \lfloor \frac{m}{2} \rfloor\}, \; x \mapsto \min\{x, m-x\}.$$

- The Lee weight for $x \in \mathbb{Z}_m^n$ is $\omega^{\mathsf{L}}(x) := \sum_{i=1}^{n} \omega^{\mathsf{L}}(x_i)$.

The Lee Metric

- We consider codes over $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.
- We denote by $x := x + m\mathbb{Z}, 0 \leq x \leq m-1$ its residue classes.
- The Lee weight of $x \in \mathbb{Z}_m$ is defined by

$$\omega^{\mathsf{L}} : \{0, \ldots, m-1\} \to \{0, \ldots, \lfloor \frac{m}{2} \rfloor\}, \; x \mapsto \min\{x, m-x\}.$$

- The Lee weight for $x \in \mathbb{Z}_m^n$ is $\omega^{\mathsf{L}}(x) := \sum_{i=1}^{n} \omega^{\mathsf{L}}(x_i)$.
- The Lee distance of $x, y \in \mathbb{Z}_m^n$ is $d^{L}(x, y) := \omega^{\mathsf{L}}(x - y)$.

Error Correcting Codes

- $\mathcal{C} \subseteq \mathbb{Z}_m^n$ is a t-error correcting code if $\min\{d^L(x,y) : x, y \in \mathcal{C},\ x \neq y\} \geq 2t+1$.

Error Correcting Codes

- $\mathcal{C} \subseteq \mathbb{Z}_m^n$ is a t-error correcting code if $\min\{d^L(x,y) : x,y \in \mathcal{C},\, x \neq y\} \geq 2t+1$.
- Sphere-packing bound: The size of a t-error correcting code is at most $H_m(n,t) := \frac{m^n}{\mathbf{v}_m^L(n,t)}$, where $\mathbf{v}_m^L(n,t) = |\{y \in \mathbb{Z}_m^n : d^L(0,y) \leq t\}|$.

Error Correcting Codes

- $\mathcal{C} \subseteq \mathbb{Z}_m^n$ is a t-error correcting code if $\min\{d^L(x,y) : x, y \in \mathcal{C},\ x \neq y\} \geq 2t+1$.
- Sphere-packing bound: The size of a t-error correcting code is at most $H_m(n,t) := \frac{m^n}{\mathbf{v}_m^L(n,t)}$, where $\mathbf{v}_m^L(n,t) = |\{y \in \mathbb{Z}_m^n : d^L(0,y) \leq t\}|$.
- If there is a t-error correcting code achieving the sphere-packing bound, then the number of t-error correcting codes is at least $2^{H_m(n,t)}$.

Error Correcting Codes

- $\mathcal{C} \subseteq \mathbb{Z}_m^n$ is a t-error correcting code if $\min\{d^L(x,y) : x, y \in \mathcal{C},\ x \neq y\} \geq 2t+1$.
- Sphere-packing bound: The size of a t-error correcting code is at most $H_m(n,t) := \frac{m^n}{\mathbf{v}_m^L(n,t)}$, where $\mathbf{v}_m^L(n,t) = |\{y \in \mathbb{Z}_m^n : d^L(0,y) \leq t\}|$.
- If there is a t-error correcting code achieving the sphere-packing bound, then the number of t-error correcting codes is at least $2^{H_m(n,t)}$.
- We will prove a corresponding upper bound.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.
- So the family of independent sets of G precisely represent the set of t-error correcting codes.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.
- So the family of independent sets of G precisely represent the set of t-error correcting codes.
- $G[\mathcal{C}]$ the induced subgraph of $\mathcal{C} \subseteq V$.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.
- So the family of independent sets of G precisely represent the set of t-error correcting codes.
- $G[\mathcal{C}]$ the induced subgraph of $\mathcal{C} \subseteq V$.
- $\Delta(G)$ the maximum degree of G.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.
- So the family of independent sets of G precisely represent the set of t-error correcting codes.
- $G[\mathcal{C}]$ the induced subgraph of $\mathcal{C} \subseteq V$.
- $\Delta(G)$ the maximum degree of G.
- $\mathcal{I}(G)$ the family of independent sets in G.

Graph Theory Tools

- $G = (V, E)$ denotes a graph with vertices $V = \mathbb{Z}_m^n$ and edges

$$E = \{(x, y) \in V^2 : d^L(x, y) \leq 2t\}.$$

- Any set $I \subseteq V$ that contains no edge is called an independent set.
- So the family of independent sets of G precisely represent the set of t-error correcting codes.
- $G[\mathcal{C}]$ the induced subgraph of $\mathcal{C} \subseteq V$.
- $\Delta(G)$ the maximum degree of G.
- $\mathcal{I}(G)$ the family of independent sets in G.
- $i(G)$ the number of independent sets in G.

Lemma

Let $1 \leq t \ll n^{3/5}, \epsilon > 0$ and n be sufficiently large. Then there is a collection $\mathcal{F}$ of subsets of $\mathbb{Z}_m^n$ with the following properties

(i) $|\mathcal{F}| \leq 2^{\epsilon H_m(n,t)}$.

(ii) Every t-error correcting code $\mathcal{C} \subseteq \mathbb{Z}_m^n$ is contained in F for some $F \in \mathcal{F}$.

(iii) $i(G[F]) \leq 2^{(1+\epsilon)H_m(n,t)}$ for every $F \in \mathcal{F}$.

How can we find such a collection of Containers?

- We build a container algorithm that certifies that the following function is well-defined:

How can we find such a collection of Containers?

- We build a container algorithm that certifies that the following function is well-defined:

Lemma

Let $\epsilon > 0, C > 0$ be a sufficiently large constant, $72 \leq t \leq n^{3/5}$ and let n be sufficiently large. Then there exists a function

$$f : \binom{V(G)}{\leq \frac{CH_m(n,t)\log(n)}{n^{3/5}}} \to 2^{V(G)},$$

such that for any $I \in \mathcal{I}(G)$ there is some $P \subseteq I$ that satisfies

(i) $I \subseteq P \cup f(P)$,

(ii) $|P| \leq \frac{CH_m(n,t)\log(n)}{n^{3/5}}$,

(iii) $i(G[P \cup f(P)]) \leq 2^{(1+\epsilon)H_m(n,t)}$.

Container Algorithm

Suppose that an order $v_1, \ldots, v_{m^n}$ of V and $I \in \mathcal{I}(G)$ are given. Set $G_0 := G$ and $P := \emptyset$. In each step i do the following:

i) If $i(G_{i-1}) \leq 2^{(1+\epsilon)H_m(n,t)}$, set $f(P) = V(G_{i-1})$, terminate and output P and $f(P)$.

ii) Let u be the vertex of maximum degree in G_{i-1}.

iii) If $u \notin I$, define $G_i := G_{i-1} \setminus \{u\}$ and proceed to step $i+1$.

iv) If $u \in I$, then add u to P, define $G_i := G_{i-1} \setminus (\{u\} \cup N_G(u))$ and proceed to step $i+1$.

DJ. Kleitman and KJ Winston "On the number of graphs without 4-cycles",Discrete Mathematics, 1982.

Container Algorithm

Suppose that an order $v_1, \ldots, v_{m^n}$ of V and $I \in \mathcal{I}(G)$ are given. Set $G_0 := G$ and $P := \emptyset$. In each step i do the following:

i) If $i(G_{i-1}) \leq 2^{(1+\epsilon)H_m(n,t)}$, set $f(P) = V(G_{i-1})$, terminate and output P and $f(P)$.

ii) Let u be the vertex of maximum degree in G_{i-1}.

iii) If $u \notin I$, define $G_i := G_{i-1} \setminus \{u\}$ and proceed to step $i+1$.

iv) If $u \in I$, then add u to P, define $G_i := G_{i-1} \setminus (\{u\} \cup N_G(u))$ and proceed to step $i+1$.

DJ. Kleitman and KJ Winston "On the number of graphs without 4-cycles",Discrete Mathematics, 1982.

Remarks

- Goal: Set up an almost precise bound.
- $\mathcal{F} := \{P \cup f(P) : P \subseteq I, I \in \mathcal{I}(G)\}$.

Supersaturation 1

Let $t \leq n^{3/5}$ and $\mathcal{C} \subseteq \mathbb{Z}_m^n$. If $|\mathcal{C}| \geq n^4 H_m(n,t)$, then

$$\Delta(G[\mathcal{C}]) \gtrsim \begin{cases} \frac{n^{6/5}|\mathcal{C}|}{H_m(n,t)} & \text{if } m \text{ is even} \\ \frac{n^{12/5}|\mathcal{C}|}{H_m(n,t)} & \text{if } m \text{ is odd} \end{cases}.$$

Supersaturation 1

Let $t \leq n^{3/5}$ and $\mathcal{C} \subseteq \mathbb{Z}_m^n$. If $|\mathcal{C}| \geq n^4 H_m(n,t)$, then

$$\Delta(G[\mathcal{C}]) \gtrsim \begin{cases} \frac{n^{6/5}|\mathcal{C}|}{H_m(n,t)} & \text{if } m \text{ is even} \\ \frac{n^{12/5}|\mathcal{C}|}{H_m(n,t)} & \text{if } m \text{ is odd} \end{cases}.$$

Supersaturation 2

Let $\epsilon > 0$, $72 \leq t \leq n^{3/5}, \mathcal{C} \subseteq \mathbb{Z}_m^n$ and n be sufficiently large. If $\Delta(G[\mathcal{C}]) \leq n^5$, then

$$i(G[\mathcal{C}]) \leq 2^{(1+\epsilon)H_L(n,t)}.$$

Our Results

Theorem

(i) For $1 \leq t \leq n^{3/5}$, the number of t-error correcting Lee codes of length n in $\mathbb{Z}_m^n$ is at most $2^{(1+o(1))H_m(n,t)}$.

Our Results

Theorem

(i) For $1 \leq t \leq n^{3/5}$, the number of t-error correcting Lee codes of length n in $\mathbb{Z}_m^n$ is at most $2^{(1+o(1))H_m(n,t)}$.

(ii) For $n^{3/5} < t \leq n^{4/5}$, the number of t-error correcting Lee codes of length n in $\mathbb{Z}_m^n$ is at most $2^{o(H_m(n,t))}$.

Bipartite Graph Model

- Let $S \geq 1$ be an integer, $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ a bipartite graph with

$$\mathcal{V} = \{\{x, y\} \in \mathbb{Z}_m^n \times \mathbb{Z}_m^n : x \neq y,\ d^L(x, y) \leq 2t\}$$
$$\mathcal{W} = \{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}$$
$$\mathcal{E} = \{(\{x, y\}, \mathcal{C}) \in \mathcal{V} \times \mathcal{W} : \{x, y\} \subseteq \mathcal{C}\}.$$

A. Gruica and A. Ravagnani "Common complements of linear subspaces and the sparseness of MRD codes",SIAM Journal on Applied Algebra and Geometry, 2022.

Bipartite Graph Model

- Let $S \geq 1$ be an integer, $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ a bipartite graph with

$$\mathcal{V} = \{\{x, y\} \in \mathbb{Z}_m^n \times \mathbb{Z}_m^n : x \neq y,\ d^L(x, y) \leq 2t\}$$
$$\mathcal{W} = \{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}$$
$$\mathcal{E} = \{(\{x, y\}, \mathcal{C}) \in \mathcal{V} \times \mathcal{W} : \{x, y\} \subseteq \mathcal{C}\}.$$

A. Gruica and A. Ravagnani "Common complements of linear subspaces and the sparseness of MRD codes",SIAM Journal on Applied Algebra and Geometry, 2022.

- The set of isolated vertices in $\mathcal{B}$ is exactly the family of t-error correcting codes.

Bipartite Graph Model

- Let $S \geq 1$ be an integer, $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ a bipartite graph with

$$\mathcal{V} = \{\{x, y\} \in \mathbb{Z}_m^n \times \mathbb{Z}_m^n : x \neq y,\ d^L(x, y) \leq 2t\}$$
$$\mathcal{W} = \{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}$$
$$\mathcal{E} = \{(\{x, y\}, \mathcal{C}) \in \mathcal{V} \times \mathcal{W} : \{x, y\} \subseteq \mathcal{C}\}.$$

A. Gruica and A. Ravagnani "Common complements of linear subspaces and the sparseness of MRD codes",SIAM Journal on Applied Algebra and Geometry, 2022.

- The set of isolated vertices in $\mathcal{B}$ is exactly the family of t-error correcting codes.
- $\mathcal{B}$ is left-regular of degree $\binom{m^n-2}{S-2}$.

Bipartite Graph Model

- Let $S \geq 1$ be an integer, $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ a bipartite graph with

$$\mathcal{V} = \{\{x, y\} \in \mathbb{Z}_m^n \times \mathbb{Z}_m^n : x \neq y,\ d^L(x, y) \leq 2t\}$$
$$\mathcal{W} = \{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}$$
$$\mathcal{E} = \{(\{x, y\}, \mathcal{C}) \in \mathcal{V} \times \mathcal{W} : \{x, y\} \subseteq \mathcal{C}\}.$$

A. Gruica and A. Ravagnani "Common complements of linear subspaces and the sparseness of MRD codes",SIAM Journal on Applied Algebra and Geometry, 2022.

- The set of isolated vertices in $\mathcal{B}$ is exactly the family of t-error correcting codes.
- $\mathcal{B}$ is left-regular of degree $\binom{m^n-2}{S-2}$.
- Together with another "strong" regularity condition we can bound the number of isolated vertices in $\mathcal{B}$.

We define

$$\delta_m^L(n, S, t) := \frac{|\{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S, d^L(\mathcal{C}) \geq 2t + 1\}|}{|\{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}|}$$

as the density function of t-error correcting codes within the set of codes of size S.

We define

$$\delta_m^L(n, S, t) := \frac{|\{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S, d^L(\mathcal{C}) \geq 2t+1\}|}{|\{\mathcal{C} \subseteq \mathbb{Z}_m^n : |\mathcal{C}| = S\}|}$$

as the density function of t-error correcting codes within the set of codes of size S.

Theorem

Let $(H_m(n, 2t+1))_{n\geq 1}$ and $(S_n)_{n\geq 1}$ be sequences with $S_n \geq 1$ for all $n \geq 1$. We have

$$\lim_{n\to+\infty} \delta_m^L(n, S_n, t) = \begin{cases} 1 & \text{if } S_n \in o(\sqrt{H_m(n, 2t+1)}) \text{ as } n \to +\infty \\ 0 & \text{if } S_n \in \omega(\sqrt{H_m(n, 2t+1)}) \text{ as } n \to +\infty \end{cases}.$$

Sphere Covering Bound

GV-bound:

$$|\mathcal{C}| \geq \frac{m^n}{\mathbf{v}_m^{\mathsf{L}}(n, 2t)}$$

Sphere Covering Bound

GV-bound:

$$|\mathcal{C}| \geq \frac{m^n}{\mathbf{v}_m^{\mathsf{L}}(n, 2t)}$$

Corollary

Non-linear Lee metric codes achieving the sphere covering bound are asymptotically sparse with respect to n (or m).

Conclusion and Future Work

- Using the container method we can efficiently count all t-error correcting codes and obtain a precise upper bound.

Conclusion and Future Work

- Using the container method we can efficiently count all t-error correcting codes and obtain a precise upper bound.
- Using the bipartite graph model we can consider code families of a specific size and count t-error correcting codes within this family.

Conclusion and Future Work

- Using the container method we can efficiently count all t-error correcting codes and obtain a precise upper bound.
- Using the bipartite graph model we can consider code families of a specific size and count t-error correcting codes within this family.
- Is one model more advantageous than the other?

Conclusion and Future Work

- Using the container method we can efficiently count all t-error correcting codes and obtain a precise upper bound.
- Using the bipartite graph model we can consider code families of a specific size and count t-error correcting codes within this family.
- Is one model more advantageous than the other?
- Linear codes are more difficult to handle.

Conclusion and Future Work

- Using the container method we can efficiently count all t-error correcting codes and obtain a precise upper bound.
- Using the bipartite graph model we can consider code families of a specific size and count t-error correcting codes within this family.
- Is one model more advantageous than the other?
- Linear codes are more difficult to handle.
- Is there a suitable graph model to count linear Lee metric codes?

End

Thank you for your attention!