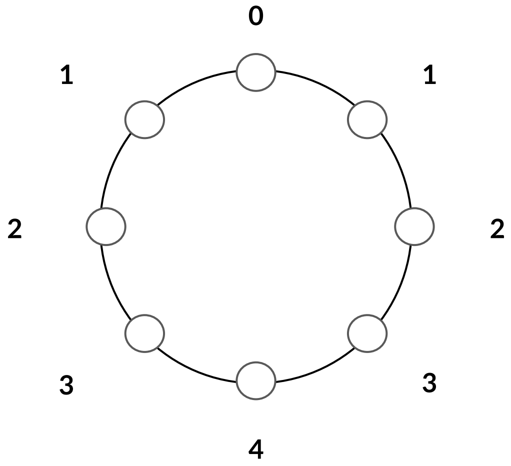
A method to find the volume of a sphere in the Lee metric, and its applications

Sagnik Bhattacharya,
Adrish Banerjee (IIT Kanpur)

ISIT 2019
July 9, 2019



Question

How do we find bounds on the size of codes for discrete metrics other than the Hamming metric? What about the Lee metric?

What can we do with the answer?

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

1. J. C.-Y. Chiang and J. K. Wolf, “On channels and codes for the Lee metric” (1971)

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

- Chiang and Wolf[1] gave the following bounds for the Lee metric

1. J. C.-Y. Chiang and J. K. Wolf, "On channels and codes for the Lee metric" (1971)

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

- Chiang and Wolf[1] gave the following bounds for the Lee metric $V_{(d-1)/2}^{(n)} \leq q^{n(1-R(d))}$

  Hamming Bound

1. J. C.-Y. Chiang and J. K. Wolf, “On channels and codes for the Lee metric” (1971)

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

- Chiang and Wolf[1] gave the following bounds for the Lee metric

$$V_{(d-1)/2}^{(n)} \leq q^{n(1-R(d))}$$

Hamming Bound

$$V_d^{(n)} > q^{n(1-R(d))}$$

Gilbert-Varshamov Bound

1. J. C.-Y. Chiang and J. K. Wolf, "On channels and codes for the Lee metric" (1971)

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

- Chiang and Wolf[1] gave the following bounds for the Lee metric $V_{(d-1)/2}^{(n)} \leq q^{n(1-R(d))}$ $V_d^{(n)} > q^{n(1-R(d))}$

  Hamming Bound — Gilbert-Varshamov Bound

- So if we know $V_t^{(n)}$ we will also know bounds on $R(d)$

1. J. C.-Y. Chiang and J. K. Wolf, "On channels and codes for the Lee metric" (1971)

What can we do with the answer?

Notation - The volume of a Lee ball of radius t and in a code with blocklength n is given by $V_t^{(n)}$

- Chiang and Wolf[1] gave the following bounds for the Lee metric $V_{(d-1)/2}^{(n)} \leq q^{n(1-R(d))}$ $V_d^{(n)} > q^{n(1-R(d))}$

  Hamming Bound Gilbert-Varshamov Bound

- So if we know $V_t^{(n)}$ we will also know bounds on $R(d)$
- Berlekamp[2] also gave the EB bound for the Lee metric

1. J. C.-Y. Chiang and J. K. Wolf, “On channels and codes for the Lee metric” (1971)
2. R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

So the question is to figure out $V_t^{(n)}$

So the question is to figure out $V_t^{(n)}$

- Chiang and Wolf[3] gave the following expression.

$$V_r^{(n)}(z) = \left(\sum_{i=0}^{r} \frac{1}{i!} \frac{d^i}{dz^i} A^{(n)}(z) \right)_{z=0}$$

3. J. C.-Y. Chiang and J. K. Wolf, "On channels and codes for the Lee metric" (1971)

So the question is to figure out $V_t^{(n)}$

- Chiang and Wolf[3] gave the following expression

$$V_r^{(n)}(z) = \left(\sum_{i=0}^{r} \frac{1}{i!} \frac{d^i}{dz^i} A^{(n)}(z) \right)_{z=0}$$

which is **mathematically intractable.**

3. J. C.-Y. Chiang and J. K. Wolf, “On channels and codes for the Lee metric” (1971)

So the question is to figure out $V_t^{(n)}$

- Chiang and Wolf[3] gave the following expression

$$V_r^{(n)}(z) = \left(\sum_{i=0}^{r} \frac{1}{i!} \frac{d^i}{dz^i} A^{(n)}(z) \right)_{z=0}$$

  which is **mathematically intractable.**

- Roth[4] gave the following expression for radius $t < q/2$

$$V_t^{(n)} = \sum_{i=0}^{n} 2^i \binom{n}{i} \binom{t}{i}$$

3. J. C.-Y. Chiang and J. K. Wolf, “On channels and codes for the Lee metric” (1971)
4. R. Roth, Introduction to Coding Theory (2006)

So the question is to figure out $V_t^{(n)}$

- Chiang and Wolf[3] gave the following expression

$$V_r^{(n)}(z) = \left(\sum_{i=0}^{r} \frac{1}{i!} \frac{d^i}{dz^i} A^{(n)}(z) \right)_{z=0}$$

which is **mathematically intractable.**

- Roth[4] gave the following expression for radius $t < q/2$

$$V_t^{(n)} = \sum_{i=0}^{n} 2^i \binom{n}{i} \binom{t}{i}$$

But the **parameter regime is too small**.

3. J. C.-Y. Chiang and J. K. Wolf, "On channels and codes for the Lee metric" (1971)
4. R. Roth, Introduction to Coding Theory (2006)

Our contribution(s)

- We use the generating function for a metric and Sanov’s theorem to find the volume of a sphere of given radius.

Our contribution(s)

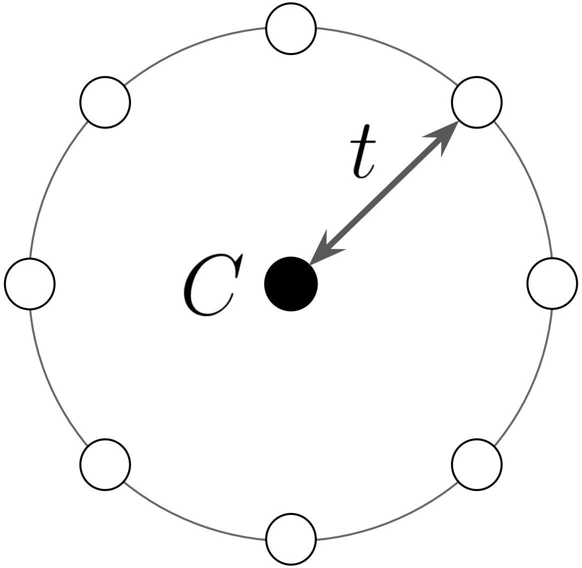
- We use the generating function for a metric and Sanov's theorem to find the volume of a sphere of given radius.
- It reduces to the known results for the Hamming metric

Our contribution(s)

- We use the generating function for a metric and Sanov's theorem to find the volume of a sphere of given radius.
- It reduces to the known results for the Hamming metric
- It allows us to find bounds on the rate for the Lee metric

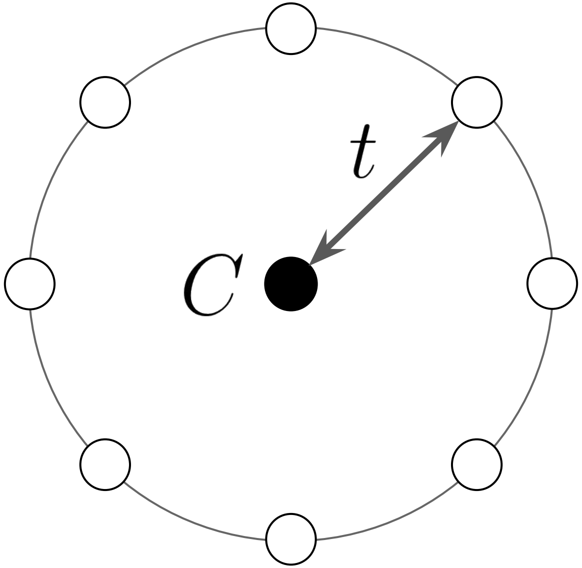
Our Methods

$$A_t^{(n)} := |\{x \in [q]^n : \mathrm{dist}(C, x) = t\}|$$



Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

$$A_t^{(n)} := |\{x \in [q]^n : \mathrm{dist}(C, x) = t\}|$$



$$V_t^{(n)} = \sum_{j=0}^{t} A_j^{(n)}$$

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

The Generating Function

- The generating function for the $A_j^{(n)}$

$$A^{(n)}(z) = \sum_j A_j^{(n)} z^j$$

The Generating Function

- Because the distance function is additive over the n coordinates…

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

The Generating Function

- Because the distance function is additive over the n coordinates…
- The generating function is multiplicative

The Generating Function

- Because the distance function is additive over the n coordinates…
- The generating function is multiplicative and $A^{(n)}(z) = [A^{(1)}(z)]^n$

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

The Generating Function

- Because the distance function is additive over the n coordinates…
- The generating function is multiplicative and $A^{(n)}(z) = [A^{(1)}(z)]^n$
- $A^{(1)}(z)$ gives the weights for a single symbol only.

The Generating Function

- Because the distance function is additive over the n coordinates…
- The generating function is multiplicative and $A^{(n)}(z) = [A^{(1)}(z)]^n$
- $A^{(1)}(z)$ gives the weights for a single symbol only.

$$A^{(1)}(z) = 1 + (q-1)z$$

q-ary Hamming Metric

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

The Generating Function

- Because the distance function is additive over the n coordinates…
- The generating function is multiplicative and $A^{(n)}(z) = [A^{(1)}(z)]^n$
- $A^{(1)}(z)$ gives the weights for a single symbol only.

$$A^{(1)}(z) = 1 + (q-1)z$$

q-ary Hamming Metric

$$A^{(1)}(z) = 1 + 2z + 2z^2 + \ldots + 2z^{\frac{q-1}{2}}$$

Lee metric for odd q

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

The Generating Function

- Because the distance function is additive over the n coordinates...
- The generating function is multiplicative and $A^{(n)}(z) = [A^{(1)}(z)]^n$
- $A^{(1)}(z)$ gives the weights for a single symbol only.

$$A^{(1)}(z) = 1 + (q-1)z$$

q-ary Hamming Metric

$$A^{(1)}(z) = 1 + 2z + 2z^2 + \ldots + 2z^{\frac{q-1}{2}}$$

Lee metric for odd q

$$A^{(1)}(z) = 1+2z+2z^2+\ldots+2z^{\frac{q-2}{2}}+z^{\frac{q}{2}}$$

Lee metric for even q

Ref - R. Berlekamp, Algebraic Coding Theory - Revised Edition (2015)

New Question

How do we find

$$\sum_j A_j^{(n)} \text{ ?}$$

Involving a probability distribution

We start by dividing both sides of $A^{(n)}(z) = \sum_j A_j^{(n)} z^j$ by q^n

Involving a probability distribution

We start by dividing both sides of $A^{(n)}(z) = \sum_j A_j^{(n)} z^j$ by q^n to get

$$\left[\frac{A^{(1)}(z)}{q}\right]^n = \sum_j \frac{A_j^{(n)}}{q^n} z^j = \sum_j B_j^{(n)} z^j$$

Involving a probability distribution

We start by dividing both sides of $A^{(n)}(z) = \sum_j A_j^{(n)} z^j$ by q^n to get

$$\left[\frac{A^{(1)}(z)}{q}\right]^n = \sum_j \frac{A_j^{(n)}}{q^n} z^j = \sum_j B_j^{(n)} z^j$$

Involving a probability distribution

Say we start from

$$A^{(1)}(z) = 1 + 2z + 2z^2 + \ldots + 2z^{\frac{q-1}{2}}$$

Involving a probability distribution

Say we start from

$$A^{(1)}(z) = 1 + 2z + 2z^2 + \ldots + 2z^{\frac{q-1}{2}}$$

Dividing by q, we get

$$\frac{A^{(1)}(z)}{q} = \frac{1}{q} + \frac{2}{q}z + \frac{2}{q}z^2 + \ldots + \frac{2}{q}z^{\frac{q-1}{2}}$$

Involving a probability distribution

Say we start from

$$A^{(1)}(z) = 1 + 2z + 2z^2 + \ldots + 2z^{\frac{q-1}{2}}$$

Dividing by q, we get

$$\frac{A^{(1)}(z)}{q} = \frac{1}{q} + \frac{2}{q}z + \frac{2}{q}z^2 + \ldots + \frac{2}{q}z^{\frac{q-1}{2}}$$

Which defines a discrete random variable X that takes value *0* w.p. *1/q*, *1* w.p. *2/q* and so on.

Involving a probability distribution

$$\left[\frac{A^{(1)}(z)}{q}\right]^n = \sum_j \frac{A_j^{(n)}}{q^n} z^j = \sum_j B_j^{(n)} z^j$$

Involving a probability distribution

$$\left[\frac{A^{(1)}(z)}{q}\right]^n = \sum_j \frac{A_j^{(n)}}{q^n} z^j = \sum_j B_j^{(n)} z^j$$

The $B_j^{(n)}$ give the probability that the sum of n *i.i.d.* samples drawn according to X add up to j

Sanov's Theorem

Theorem 1 (Sanov's Theorem)**.** *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\sim Q(x)$. Let $E \subseteq \mathcal{P}$ be a set of probability distributions and $\mathcal{P}$ be the set of all types from the n realisations $X_1, X_2, \ldots, X_n$. Then,*

$$Q^n(E) = Q^n(E \cap \mathcal{P}_n) \leq (n+1)^{|\mathcal{X}|} 2^{-nD(P^*||Q)}$$

where $|\mathcal{X}|$ is the support of each X_i, $D(\cdot||\cdot)$ is the K-L divergence, $Q^n(E)$ is the probability that the empirical distribution obtained from an n-long sample $X_1, \ldots, X_n$ each $\sim Q(x)$ belongs to the set E, and

$$P^* = \arg\min_{P \in E} D(P||Q)$$

is the distribution in E that is closest to Q in relative entropy. If we also have that the set E is the closure of its interior, then we also have the result

$$\frac{1}{n} \log Q^n(E) \to -D(P^*||Q)$$

Retaining terms upto first order in the exponent, we have

$$2^{-nD(P^*||Q)-o(n)} \leq Q^n(E) \leq 2^{-nD(P^*||Q)+o(n)}$$

Sanov, I. N. (1957) "On the probability of large deviations of random variables"
Cover, Thomas M.; Thomas, Joy A. (2006). Elements of Information Theory (2 ed.)

Using Sanov's theorem - finding E

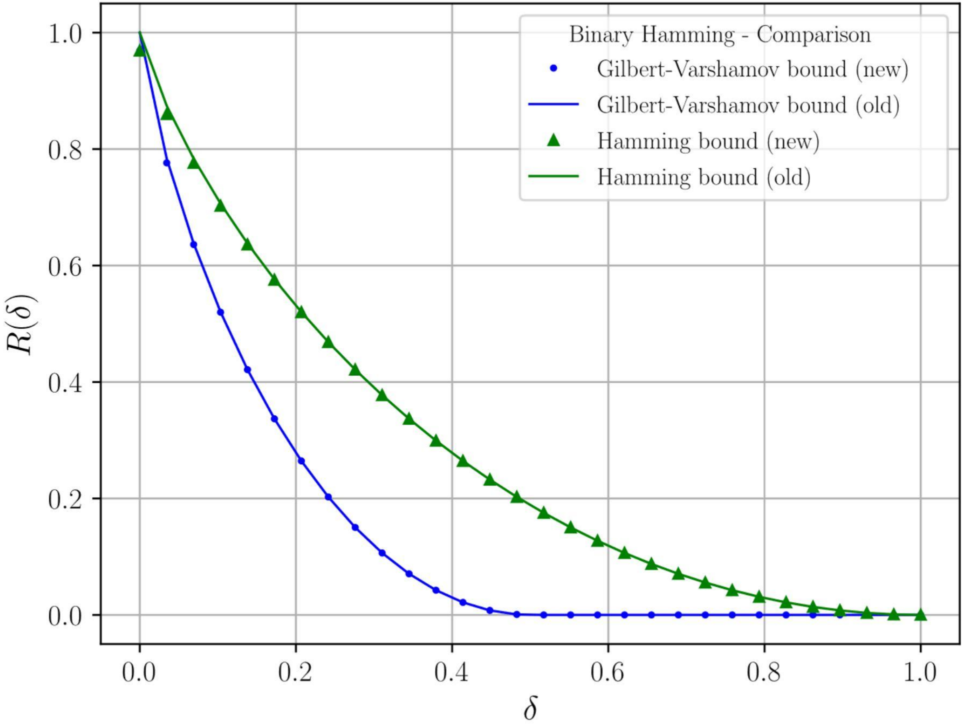
The natural choice while calculating the k^{th} coefficient is the set of all distributions with mean less than or equal to k.

Sanity Check - Hamming metric

The Hamming metric does not require convex optimisation. The random variable in the Hamming case is **Bernoulli**, and the KL divergence minimising distribution is not hard to find. The result is familiar.

$$q^{nH_q(p)-o(n)} \leq V_{pn}^{(n)} \leq q^{nH_q(p)+o(n)}$$

Sanity check for the Hamming metric



Convex Optimisation

In general, we need to use convex optimisation to find P^*

Convex Optimisation

In general, we need to use convex optimisation to find P^*

Strong duality holds for the problem, so any solution to the dual implies an upper bound for the primal.

Convex optimisation - functional form

The dual program is given by

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -p\lambda - \log\left(\sum_j \mathbb{P}_X(j)e^{-j\lambda}\right) \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

Convex optimisation - functional form

The dual program is given by

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -p\lambda - \log\left(\sum_j \mathbb{P}_X(j) e^{-j\lambda}\right) \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

Since **any** $\lambda(p)$ will give an upper bound by strong duality,

Convex optimisation - functional form

The dual program is given by

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -p\lambda - \log\left(\sum_j \mathbb{P}_X(j)e^{-j\lambda}\right) \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

Since **any** $\lambda(p)$ will give an upper bound by strong duality, we can choose the function to be $\lambda(p) = c(q)(\overline{D}^{\frac{1}{q}} - p^{\frac{1}{q}})$

Functional form - intuition

$$\lambda(p) = c(q)(\overline{D}^{\frac{1}{q}} - p^{\frac{1}{q}})$$

Functional form - intuition

$$\lambda(p) = c(q)(\overline{D}^{\frac{1}{q}} - p^{\frac{1}{q}})$$

- $\lambda(\overline{D})$ should be zero

Functional form - intuition

$$\lambda(p) = c(q)(\overline{D}^{\frac{1}{q}} - p^{\frac{1}{q}})$$

- $\lambda(\overline{D})$ should be zero
- $\lambda(p)$ should be monotonically increasing

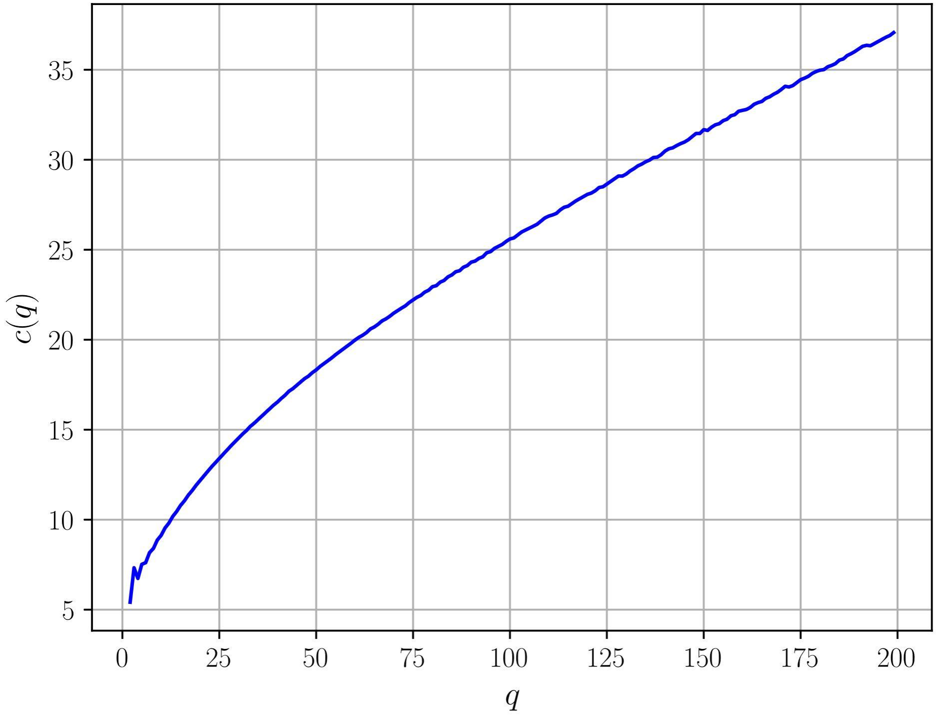
Functional form - intuition

$$\lambda(p) = c(q)(\overline{D}^{\frac{1}{q}} - p^{\frac{1}{q}})$$

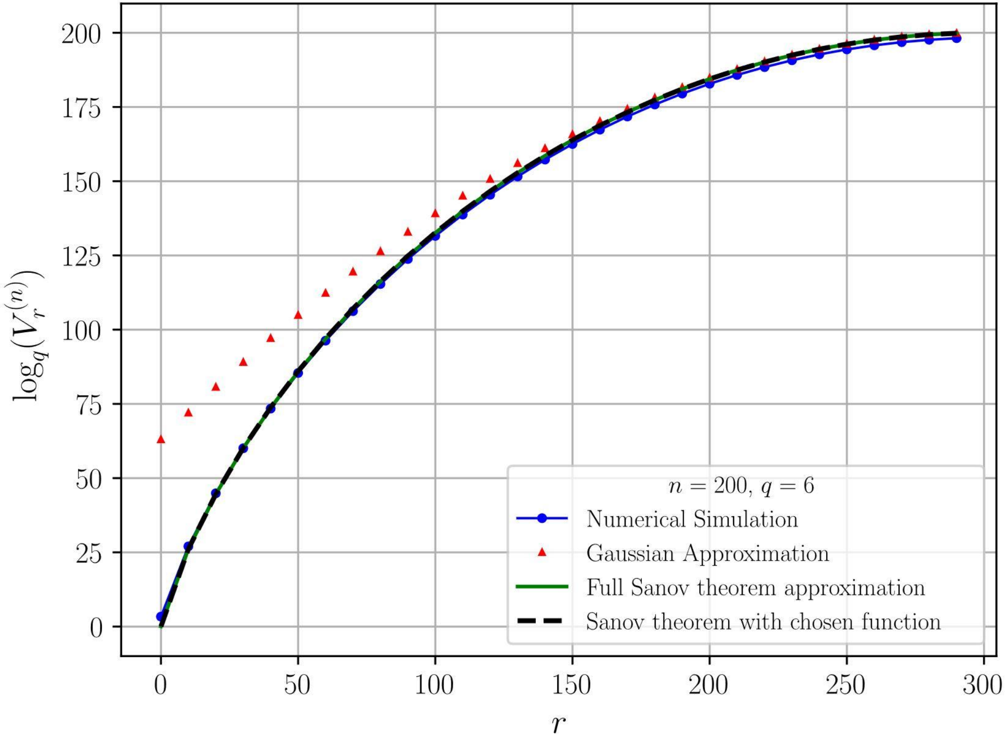
- $\lambda(\overline{D})$ should be zero
- $\lambda(p)$ should be monotonically increasing

Open - more analytical justification of why this is the form.

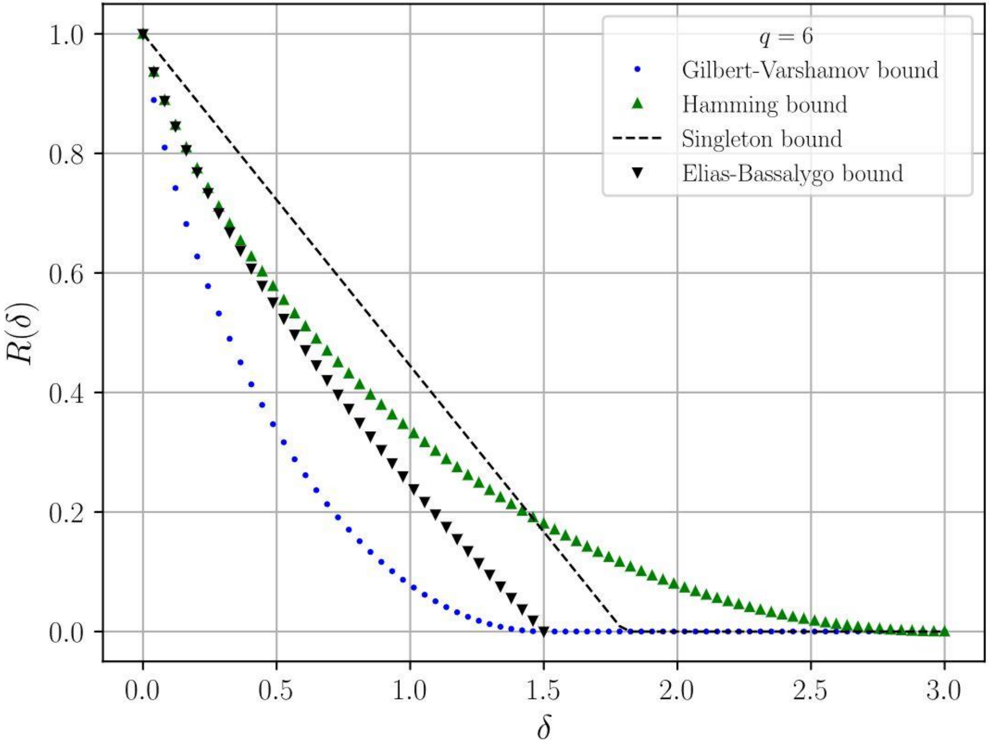
What is the value of $c(q)$?



Immediate result - asymptotic sizes of Lee balls



Also - bounds on codes in Lee metric



Key takeaways

- Solving an algebraic problem using analytic techniques

Key takeaways

- Solving an algebraic problem using analytic techniques
- Method generalises to all discrete metrics with the following property

Key takeaways

- Solving an algebraic problem using analytic techniques
- Method generalises to all discrete metrics with the following property - the set of distances of all symbols to one fixed symbol remains the same when the fixed symbol is replaced by some other symbol

Key takeaways

- Solving an algebraic problem using analytic techniques
- Method generalises to all discrete metrics with the following property - the set of distances of all symbols to one fixed symbol remains the same when the fixed symbol is replaced by some other symbol
- Should generalise to other discrete metrics too, but the expressions would be more complicated.

Thank you!