On the VC-Dimension of Binary Error-Correcting Codes

Sihuang Hu

Joint work with Nir Weinberger and Ofer Shayevitz

TEL AVIV UNIVERSITY אוניברסיטת תל אביב

VC-Dimension

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, ...

VC-DIMENSION

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, ...
- Shattered coordinate set

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \end{matrix}$$

VC-DIMENSION

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, …
- Shattered coordinate set

$$
\begin{array}{|ccc|cccc}
\color{blue}0 & \color{blue}0 & \color{blue}0 & 0 & 0 & 0 & 0 \\
\color{blue}1 & \color{blue}0 & \color{blue}0 & 1 & 0 & 1 & 1 \\
\color{blue}0 & \color{blue}1 & \color{blue}0 & 1 & 1 & 1 & 0 \\
\color{blue}0 & \color{blue}0 & \color{blue}1 & 0 & 1 & 1 & 1 \\
\color{blue}1 & \color{blue}1 & \color{blue}0 & 0 & 1 & 0 & 1 \\
\color{blue}1 & \color{blue}0 & \color{blue}1 & 1 & 1 & 0 & 0 \\
\color{blue}0 & \color{blue}1 & \color{blue}1 & 1 & 0 & 0 & 1 \\
\color{blue}1 & \color{blue}1 & \color{blue}1 & 0 & 0 & 1 & 0
\end{array}
$$

VC-Dimension

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, …
- Non-Shattered coordinate set

$$
\begin{array}{ccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}
$$

VC-DIMENSION

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, …
- Non-Shattered coordinate set

$$
\begin{matrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{1} & \mathbf{0} & 0 & \mathbf{1} & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 \\
\mathbf{1} & \mathbf{0} & 1 & \mathbf{1} & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 0
\end{matrix}
$$

VC-Dimension

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, …
- Shattered coordinate set

$$
\begin{array}{|ccc|cccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}
$$

- **VC-dimension**: maximum size of a shattered coordinate set

VC-Dimension

- Vapnik, Chervonenkis (1971):
  Statistical Learning, Computational Geometry, …
- Shattered coordinate set

$$\begin{array}{|ccc|cccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}$$

- **VC-dimension**: maximum size of a shattered coordinate set
- Fact: The VC-dimension of the above code is 3.

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$|\mathcal{C}| = 2^k$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\quad\quad$ $\mathrm{VCD}(\mathcal{C}) = k$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$|\mathcal{C}| = 2^k$ $\quad\quad\quad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$$

Two (Extremal) Examples

- $\mathcal{C} =$ Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C} =$ Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i} \qquad \mathrm{VCD}(\mathcal{C}) = k$$

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$\leq |\mathcal{C}| \leq$$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$2^D \leq |\mathcal{C}| \leq$$

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\quad\quad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i} \qquad \mathrm{VCD}(\mathcal{C}) = k$$

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$2^D \leq |\mathcal{C}| \leq \sum_{i=0}^{D} \binom{n}{i}.$$

TWO (EXTREMAL) EXAMPLES

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$|\mathcal{C}| = 2^k$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$ $\qquad$ $\mathrm{VCD}(\mathcal{C}) = k$

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$2^D \leq |\mathcal{C}| \leq \sum_{i=0}^{D} \binom{n}{i}.$$

Upper bound: Sauer–Shelah lemma. PF. The Shifting Method.

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\quad$ $\mathrm{VCD}(\mathcal{C}) = k$ $\quad$ Attains Lower Bound

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$ $\quad$ $\mathrm{VCD}(\mathcal{C}) = k$

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$2^D \leq |\mathcal{C}| \leq \sum_{i=0}^{D} \binom{n}{i}.$$

Upper bound: Sauer–Shelah lemma. PF. The Shifting Method.

Two (Extremal) Examples

- $\mathcal{C}$ = Linear code of dimension k and with generator matrix

$$\mathbf{G} := \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{pmatrix}$$

$|\mathcal{C}| = 2^k$ $\quad$ $\mathrm{VCD}(\mathcal{C}) = k$ $\quad$ Attains Lower Bound

- $\mathcal{C}$ = Hamming ball $B(k)$ of radius k (all vectors of weight $\leq k$)

$|\mathcal{C}| = \sum_{i=0}^{k} \binom{n}{i}$ $\quad$ $\mathrm{VCD}(\mathcal{C}) = k$ $\quad$ Attains Upper Bound

Bounds

Let $\mathcal{C}$ be a length-n binary code with VC-dimension D. Then

$$2^D \leq |\mathcal{C}| \leq \sum_{i=0}^{D} \binom{n}{i}.$$

Upper bound: Sauer–Shelah lemma. PF. The Shifting Method.

CODING-THEORETIC MOTIVATION

- Binary-input channel with errors and erasures
- e erasures (puncturing)

Coding-Theoretic Motivation

- Binary-input channel with errors and erasures
- e erasures (puncturing)
- Detection thresholds (in the remaining $n - e$ coordinates):
  - $\#$ *detectable bit errors* > 0
  - $\#$ *detectable error patterns* > 0

Coding-Theoretic Motivation

- Binary-input channel with errors and erasures
- e erasures (puncturing)

- Detection thresholds (in the remaining $n - e$ coordinates):
  - $\#$ *detectable bit errors* > 0 $\iff$ $e < \mathrm{dist}(\mathcal{C}) - 1$
  - $\#$ *detectable error patterns* > 0

Coding-Theoretic Motivation

- Binary-input channel with errors and erasures
- e erasures (puncturing)

- Detection thresholds (in the remaining $n - e$ coordinates):
  - \# *detectable bit errors* $> 0 \iff e < \mathrm{dist}(\mathcal{C}) - 1$
  - \# *detectable error patterns* $> 0 \iff e < n - \mathrm{VCD}(\mathcal{C})$

Coding-Theoretic Motivation

- Binary-input channel with errors and erasures
- e erasures (puncturing)
- Detection thresholds (in the remaining $n - e$ coordinates):
  - # *detectable bit errors* $> 0 \iff e < \mathrm{dist}(\mathcal{C}) - 1$
  - # *detectable error patterns* $> 0 \iff e < n - \mathrm{VCD}(\mathcal{C})$

WANTED

Large Minimum Distance & Small VC-Dimension

Coding-Theoretic Motivation

- Binary-input channel with errors and erasures
- e erasures (puncturing)

- Detection thresholds (in the remaining $n - e$ coordinates):
  - $\#$ *detectable bit errors* $> 0 \iff e < \mathrm{dist}(\mathcal{C}) - 1$
  - $\#$ *detectable error patterns* $> 0 \iff e < n - \mathrm{VCD}(\mathcal{C})$

WANTED

Large Minimum Distance & Small VC-Dimension

Examples

$\mathcal{C}$	Mini. Dist.	VCD
Linear codes	Large	Large
Hamming balls	Small	Small
Random codes	Large	Large

Asymptotic Problem Formulation

- Rate $R \in [0, 1]$ is (δ, d)-achievable if $\exists \{\mathcal{C}_n\}_{n=1}^{\infty}$ such that

$$
\begin{aligned}
\mathrm{dist}(\mathcal{C}_n) &\geq \delta n \\
\mathrm{VCD}(\mathcal{C}_n) &\leq dn \\
\limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{C}_n| &\geq R
\end{aligned}
$$

Asymptotic Problem Formulation

- Rate $R \in [0, 1]$ is (δ, d)-achievable if $\exists \{\mathcal{C}_n\}_{n=1}^{\infty}$ such that

$$\begin{aligned} \mathrm{dist}(\mathcal{C}_n) &\geq \delta n \\ \mathrm{VCD}(\mathcal{C}_n) &\leq dn \\ \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{C}_n| &\geq R \end{aligned}$$

Maximal rate

$$C(d, \delta) \stackrel{\text{def}}{=} \sup \{R : \ R \text{ is } (\delta, d)\text{-achievable}\}$$

Asymptotic Problem Formulation

- Rate $R \in [0, 1]$ is (δ, d)-achievable if $\exists \, \{\mathcal{C}_n\}_{n=1}^{\infty}$ such that

$$\begin{aligned} \mathrm{dist}(\mathcal{C}_n) &\geq \delta n \\ \mathrm{VCD}(\mathcal{C}_n) &\leq dn \\ \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{C}_n| &\geq R \end{aligned}$$

Maximal rate

$$C(d, \delta) \stackrel{\text{def}}{=} \sup \{R : \ R \text{ is } (\delta, d)\text{-achievable}\}$$

- In this talk: two upper bounds + two lower bounds

(Known) Upper Bounds

Second MRRW bound (ignore VCD constraint)

$$C(d, \delta) \leq R_{LP}(\delta)$$

(Known) Upper Bounds

Second MRRW bound (ignore VCD constraint)

$$C(d, \delta) \leq R_{LP}(\delta)$$

Sauer–Shelah lemma (ignore min-dist constraint)

$$C(d, \delta) \leq h(d).$$

(Known) Upper Bounds

Second MRRW bound (ignore VCD constraint)

$$C(d,\delta) \le R_{LP}(\delta)$$

Sauer–Shelah lemma (ignore min-dist constraint)

$$C(d,\delta) \le h(d).$$

Bounded VC-dimension: Dudley (1978), Haussler (1995)

(Known) Upper Bounds

Second MRRW bound (ignore VCD constraint)

$$C(d,\delta) \le R_{LP}(\delta)$$

Sauer–Shelah lemma (ignore min-dist constraint)

$$C(d,\delta) \le h(d).$$

Bounded VC-dimension: Dudley (1978), Haussler (1995)

Corollary of Haussler's packing lemma (1995)

$$C(d,\delta) \le \frac{2d}{\delta+2d} \cdot h\left(\left\langle \frac{\delta+2d}{2} \right\rangle\right)$$

Here $\langle a \rangle \stackrel{\text{def}}{=} \min\{a, \frac{1}{2}\}$.

(Known) Upper Bounds

Second MRRW bound (ignore VCD constraint)

$$C(d, \delta) \leq R_{LP}(\delta)$$

Sauer–Shelah lemma (ignore min-dist constraint)

$$C(d, \delta) \leq h(d).$$

Bounded VC-dimension: Dudley (1978), Haussler (1995)

Corollary of Haussler's packing lemma (1995)

$$C(d, \delta) \leq \frac{2d}{\delta + 2d} \cdot h\left(\left\langle \frac{\delta + 2d}{2} \right\rangle\right)$$

Here $\langle a \rangle \stackrel{\text{def}}{=} \min\{a, \frac{1}{2}\}$.

PF. The Probabilistic Method – Linearity of Expectation.

An Upper Bound via Projection

- $\mathcal{C}$ = a code with VC-dimension dn and minimum distance δn.

An Upper Bound via Projection

- $\mathcal{C}$ = a code with VC-dimension dn and minimum distance δn.
- $\mathcal{C}|_{[sn]}$: projection of $\mathcal{C}$ on the first sn coordinates

$$
\begin{array}{cccc|cc}
x_1 & x_2 & \cdots & x_{sn} & \cdots & x_n \\
y_1 & y_2 & \cdots & y_{sn} & \cdots & y_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
z_1 & z_2 & \cdots & z_{sn} & \cdots & z_n
\end{array}
$$

An Upper Bound via Projection

- $\mathcal{C}$ = a code with VC-dimension dn and minimum distance δn.
- $\mathcal{C}|_{[sn]}$: projection of $\mathcal{C}$ on the first sn coordinates

$$\begin{array}{cccc|cc} x_1 & x_2 & \cdots & x_{sn} & \cdots & x_n \\ y_1 & y_2 & \cdots & y_{sn} & \cdots & y_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1 & z_2 & \cdots & z_{sn} & \cdots & z_n \end{array}$$

By Sauer–Shelah lemma

$$|\mathcal{C}|_{[sn]}| \leq \sum_{i=0}^{dn} \binom{sn}{i}$$

An Upper Bound via Projection

- $\mathcal{C}$ = a code with VC-dimension dn and minimum distance δn.
- $\mathcal{C}|_{[sn]}$: projection of $\mathcal{C}$ on the first sn coordinates

$$\begin{array}{cccc|cc} \boxed{\begin{array}{cccc} x_1 & x_2 & \cdots & x_{sn} \\ y_1 & y_2 & \cdots & y_{sn} \\ \vdots & \vdots & \vdots & \vdots \\ z_1 & z_2 & \cdots & z_{sn} \end{array}} & \begin{array}{cc} \cdots & x_n \\ \cdots & y_n \\ \vdots & \vdots \\ \cdots & z_n \end{array} \end{array}$$

By Sauer–Shelah lemma

$$|\mathcal{C}|_{[sn]}| \leq \sum_{i=0}^{dn} \binom{sn}{i}$$

- Fix $\mathbf{u} \in \mathcal{C}|_{[sn]}$, consider its possble suffixes $\mathcal{Z}(\mathbf{u})$

$$u_1 u_2 \cdots u_{sn} \begin{cases} \boxed{\begin{array}{ccc} u_{sn+1} & \cdots & u_n \\ y_{sn+1} & \cdots & y_n \\ \vdots & \vdots & \vdots \\ z_{sn+1} & \cdots & z_n \end{array}} \end{cases}$$

An Upper Bound via Projection

- $\mathcal{C}$ = a code with VC-dimension dn and minimum distance δn.
- $\mathcal{C}|_{[sn]}$: projection of $\mathcal{C}$ on the first sn coordinates

$$\begin{array}{c} \boxed{\begin{matrix} x_1 & x_2 & \cdots & x_{sn} \\ y_1 & y_2 & \cdots & y_{sn} \\ \vdots & \vdots & \vdots & \vdots \\ z_1 & z_2 & \cdots & z_{sn} \end{matrix}} \begin{matrix} \cdots & x_n \\ \cdots & y_n \\ \vdots & \vdots \\ \cdots & z_n \end{matrix} \end{array}$$

By Sauer–Shelah lemma

$$|\mathcal{C}|_{[sn]}| \leq \sum_{i=0}^{dn} \binom{sn}{i}$$

- Fix $\mathbf{u} \in \mathcal{C}|_{[sn]}$, consider its possble suffixes $\mathcal{Z}(\mathbf{u})$

$$u_1 u_2 \cdots u_{sn} \left\{ \boxed{\begin{matrix} u_{sn+1} & \cdots & u_n \\ y_{sn+1} & \cdots & y_n \\ \vdots & \vdots & \vdots \\ z_{sn+1} & \cdots & z_n \end{matrix}} \right.$$

$|\mathcal{Z}(\mathbf{u})|$: second MRRW bound.

An Upper Bound via Projection

$$|\mathcal{C}| = \sum_{\mathbf{u} \in \mathcal{C}|_{[sn]}} |\mathcal{Z}(\mathbf{u})|$$

An Upper Bound via Projection

$$|\mathcal{C}| = \sum_{\mathbf{u} \in \mathcal{C}|_{[sn]}} |\mathcal{Z}(\mathbf{u})| \leq \sum_{i=0}^{dn} \binom{sn}{i} \cdot \max_{\mathbf{u} \in \mathcal{C}|_{[sn]}} |\mathcal{Z}(\mathbf{u})|$$

An Upper Bound via Projection

$$|\mathcal{C}| = \sum_{\mathbf{u} \in \mathcal{C}|_{[sn]}} |\mathcal{Z}(\mathbf{u})| \le \sum_{i=0}^{dn} \binom{sn}{i} \cdot \max_{\mathbf{u} \in \mathcal{C}|_{[sn]}} |\mathcal{Z}(\mathbf{u})|$$

Upper bound via projection (H., Weinberger, Shayevitz'17)

$$C(d,\delta) \le \min_{0 \le s \le 1-2\delta} \left\{ s \cdot h\left(\left\langle \frac{d}{s} \right\rangle\right) + (1-s) \cdot R_{LP}\left(\frac{\delta}{1-s}\right) \right\}$$

Lower Bound: First Thought

1. Pick a binary code $\mathcal{C}$ attaining GV bound

Lower Bound: First Thought

1. Pick a binary code $\mathcal{C}$ attaining GV bound
2. Intersect $\mathcal{C}$ with a (shifted) Hamming ball $B(dn)$

Lower Bound: First Thought

1. Pick a binary code $\mathcal{C}$ attaining GV bound

2. Intersect $\mathcal{C}$ with a (shifted) Hamming ball $B(dn)$

Elias–Bassalygo Lemma

For $\mathcal{C} \subseteq \{0,1\}^n$, there exists a Hamming ball of radius dn with at least

$$\frac{|C| \cdot |B(dn)|}{2^n} = \frac{|C| \cdot \sum_{i=0}^{dn} \binom{n}{i}}{2^n}$$

codewords in it.

Lower Bound: First Thought

1. Pick a binary code $\mathcal{C}$ attaining GV bound

2. Intersect $\mathcal{C}$ with a (shifted) Hamming ball $B(dn)$

Elias–Bassalygo Lemma

For $\mathcal{C} \subseteq \{0,1\}^n$, there exists a Hamming ball of radius dn with at least

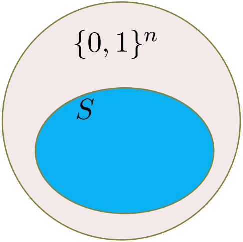
$$\frac{|C| \cdot |B(dn)|}{2^n} = \frac{|C| \cdot \sum_{i=0}^{dn} \binom{n}{i}}{2^n}$$

codewords in it.

3. We can do better via a direct approach!

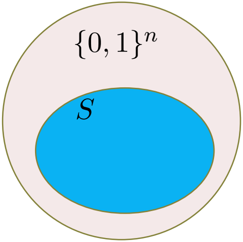
Lower Bound "Recipe"

1. Pick $S \subseteq \{0,1\}^n$ with "nice" structure, e.g. constant weight, Markov type



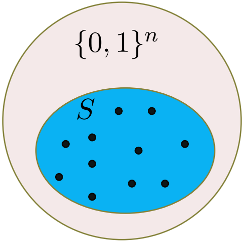
Lower Bound "Recipe"

1. Pick $S \subseteq \{0,1\}^n$ with "nice" structure, e.g. constant weight, Markov type
2. Give an upper bound for $\mathrm{VCD}(\mathcal{C})$ of a code $\mathcal{C} \subseteq S$ with $\mathrm{dist}(\mathcal{C}) \geq \delta n$.



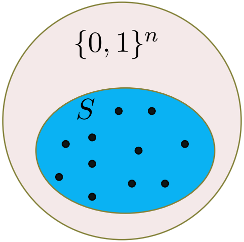
Lower Bound "Recipe"

1. Pick $S \subseteq \{0,1\}^n$ with "nice" structure, e.g. constant weight, Markov type
2. Give an upper bound for $\mathrm{VCD}(\mathcal{C})$ of a code $\mathcal{C} \subseteq S$ with $\mathrm{dist}(\mathcal{C}) \geq \delta n$.
3. Compute GV bound for codes in S



Lower Bound "Recipe"

1. Pick $S \subseteq \{0,1\}^n$ with "nice" structure, e.g. constant weight, Markov type
2. Give an upper bound for $\mathrm{VCD}(\mathcal{C})$ of a code $\mathcal{C} \subseteq S$ with $\mathrm{dist}(\mathcal{C}) \geq \delta n$.
3. Compute GV bound for codes in S
4. Combine above bounds and evaluate $n \to \infty$ asymptotics.



Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$

Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$
- $\mathcal{C}(\subseteq S)$: a code with $\mathrm{dist}(\mathcal{C}) = \delta n$

Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$
- $\mathcal{C}(\subseteq S)$: a code with $\mathrm{dist}(\mathcal{C}) = \delta n$
- Clearly, $\mathrm{VCD}(\mathcal{C}) \leq wn$.

Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$
- $\mathcal{C}(\subseteq S)$: a code with $\mathrm{dist}(\mathcal{C}) = \delta n$
- Clearly, $\mathrm{VCD}(\mathcal{C}) \leq wn$.

Claim

$$\mathrm{VCD}(\mathcal{C}) \leq \left(w - \frac{\delta}{2}\right) n + 1$$

Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$
- $\mathcal{C}(\subseteq S)$: a code with $\text{dist}(\mathcal{C}) = \delta n$
- Clearly, $\text{VCD}(\mathcal{C}) \leq wn$.

Claim

$$\text{VCD}(\mathcal{C}) \leq \left(w - \frac{\delta}{2}\right) n + 1$$

Proof: Suppose $\text{VCD}(\mathcal{C}) = dn$. Then $\exists\, \mathbf{x}, \mathbf{y} \in \mathcal{C}$ such that

$$\begin{array}{rcccccccc} \mathbf{x} = & (1 & 1 & 1 & \cdots & 1 & x_{dn+1} & \cdots & x_n) \\ \mathbf{y} = & (0 & 1 & 1 & \cdots & 1 & y_{dn+1} & \cdots & y_n) \end{array}$$

$$\delta n \leq \text{dist}(\mathbf{x}, \mathbf{y}) = 2wn - 2|\text{supp}(\mathbf{x}) \cap \text{supp}(\mathbf{y})| \leq 2wn - 2(dn - 1)$$

Lower Bound I: Constant Weight Codes

- $S = \{\mathbf{s} \in \{0,1\}^n : |\mathbf{s}| = wn\}$
- $\mathcal{C}(\subseteq S)$: a code with $\mathrm{dist}(\mathcal{C}) = \delta n$
- Clearly, $\mathrm{VCD}(\mathcal{C}) \leq wn$.

Claim

$$\mathrm{VCD}(\mathcal{C}) \leq \left(w - \frac{\delta}{2}\right) n + 1$$

GV bound for constant weight codes

$$A(n, \delta n, wn) \geq \frac{\binom{n}{wn}}{\sum_{i=0}^{\delta n/2-1} \binom{wn}{i}\binom{n-wn}{i}}$$

Constant-weight lower bound (H., Weinberger, Shayevitz'17)

Let $d, \delta \in [0, \frac{1}{2}]$, and let $w = d + \frac{\delta}{2}$. Then

$$C(d, \delta) \geq \begin{cases} h(w) - \max\limits_{0 \leq x \leq \delta/2} \left[w\, h\left(\frac{x}{w}\right) + (1-w) h\left(\frac{x}{1-w}\right)\right], & w < \frac{1}{2} \\ 1 - h(\delta), & \text{otherwise.} \end{cases}$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$\boxed{00}110$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$0\boxed{01}10$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00\boxed{11}0$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$001\boxed{10}$$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

- $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

- $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- Fact: For any $T \subseteq S$, $\mathrm{VCD}(T) \leq dn$.

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

- $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- Fact: For any $T \subseteq S$, $\mathrm{VCD}(T) \leq dn$.

- $(S, M, \delta n)$-code: a subset of S with size M and mini. dist. $\geq \delta n$.

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

- $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \le dn - 1\}$
- Fact: For any $T \subseteq S$, $\mathrm{VCD}(T) \le dn$.

- $(S, M, \delta n)$-code: a subset of S with size M and mini. dist. $\ge \delta n$.

Generalized GV [Kolesnik, Krachkovsky'91]

There exists an $(S, M, \delta n)$-code such that

$$M \ge \frac{|S|}{4\,|B|\,/|S|}$$

where $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \le \delta n - 1\}$.

Lower Bound II: Markov Type

- Number of switches of a sequence

$$00110$$

- $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- Fact: For any $T \subseteq S$, VCD$(T) \leq dn$.

- $(S, M, \delta n)$-code: a subset of S with size M and mini. dist. $\geq \delta n$.

Generalized GV [Kolesnik, Krachkovsky'91]

There exists an $(S, M, \delta n)$-code such that

$$M \geq \frac{|S|}{\cancel{4}\,|B|\,/|S|}$$

where $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.

Remark. Improved by Gu and Fuja (1993), Tolhuizen (1997, implied by Turán's theorem).

- Question: How to compute $|S|$ and $|B|$?

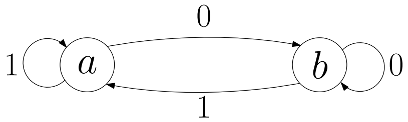
Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

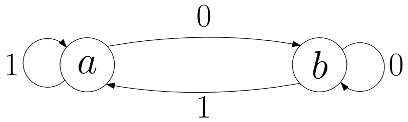
- Labeled graph $G = (V_G, E_G, L_G)$



Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$

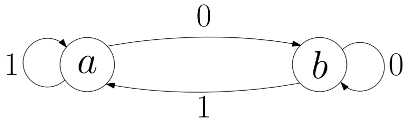


$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$



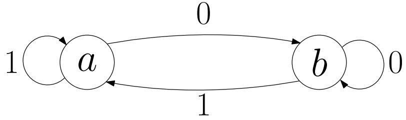
$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

$$aabba \xrightarrow{Labeling} 1001$$

Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$



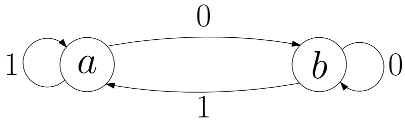
$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

$$\color{red}{aa}\color{black}{bba} \xrightarrow{Labeling} \color{red}{1}\color{black}{001}$$

Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$

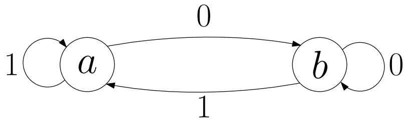


$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

$$aabba \xrightarrow{Labeling} 1001$$

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$



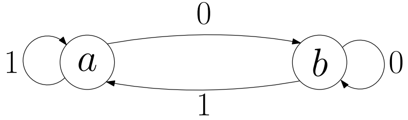
$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

$$aabba \xrightarrow{Labeling} 1001$$

Stationary Markov Chains on Graphs

- Question: How to compute $|S|$ and $|B|$?
  Follows the same line of reasoning as in (Marcus and Roth 1992), where they developed improved GV bound for constrained systems, based on stationary Markov chain analysis.

- Labeled graph $G = (V_G, E_G, L_G)$



$$\text{a walk in graph} \xrightarrow{Labeling} \text{a binary sequence}$$

$$aabba \xrightarrow{Labeling} 1001$$

Stationary Markov Chains on Graphs

- A stationary Markov chain on a finite directed graph G is a function $P : E_G \to [0, 1]$ such that
  (I) $\sum_{e \in E_G} P(e) = 1$;

- A stationary Markov chain on a finite directed graph G is a function $P : E_G \to [0, 1]$ such that
  (I) $\sum_{e \in E_G} P(e) = 1$;
  (II) $\sum_{e \in E_G^+(u)} P(e) = \sum_{e \in E_G^-(u)} P(e)$ for every $u \in V_G$.

- A stationary Markov chain on a finite directed graph G is a function $P : E_G \to [0, 1]$ such that
  (I) $\sum_{e \in E_G} P(e) = 1$;
  (II) $\sum_{e \in E^+_G(u)} P(e) = \sum_{e \in E^-_G(u)} P(e)$ for every $u \in V_G$.

- $\mathcal{M}(G)$ = the set of all stationary Markov chains on G.

Stationary Markov Chains on Graphs

- A stationary Markov chain on a finite directed graph G is a function $P : E_G \to [0, 1]$ such that
  (I) $\sum_{e \in E_G} P(e) = 1$;
  (II) $\sum_{e \in E_G^+(u)} P(e) = \sum_{e \in E_G^-(u)} P(e)$ for every $u \in V_G$.

- $\mathcal{M}(G)$ = the set of all stationary Markov chains on G.

- For a $P \in \mathcal{M}(G)$, we introduce two dummy r.v. X, Y such that

$$\Pr\{X = u, Y = v\} = \begin{cases} P((u, v)) & \text{if } (u, v) \in E_G \\ 0 & \text{otherwise.} \end{cases}$$

Stationary Markov Chains on Graphs

- A stationary Markov chain on a finite directed graph G is a function $P : E_G \to [0, 1]$ such that
  (I) $\sum_{e \in E_G} P(e) = 1$;
  (II) $\sum_{e \in E_G^+(u)} P(e) = \sum_{e \in E_G^-(u)} P(e)$ for every $u \in V_G$.

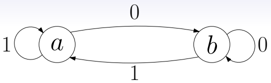
- $\mathcal{M}(G) =$ the set of all stationary Markov chains on G.

- For a $P \in \mathcal{M}(G)$, we introduce two dummy r.v. X, Y such that

$$\Pr\{X = u, Y = v\} = \begin{cases} P((u,v)) & \text{if } (u,v) \in E_G \\ 0 & \text{otherwise.} \end{cases}$$

  Then the condition (II) amounts to saying that the marginal distributions of X and Y are equal.

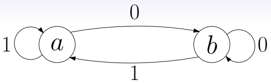
Stationary Markov Chains on Graphs



- The (switch) function $f : E_G \to \mathbb{R}$ is defined by

$$f((a,a)) = f((b,b)) = 0,$$
$$f((a,b)) = f((b,a)) = 1.$$

Stationary Markov Chains on Graphs

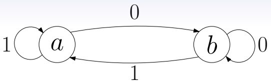


- The (switch) function $f : E_G \to \mathbb{R}$ is defined by

$$f((a,a)) = f((b,b)) = 0,$$
$$f((a,b)) = f((b,a)) = 1.$$

- $\Gamma_n(G)$ = the set of all cycles in G of length n that start and end at some fixed point u.

Stationary Markov Chains on Graphs



- The (switch) function $f : E_G \to \mathbb{R}$ is defined by

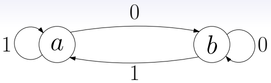
$$f((a,a)) = f((b,b)) = 0,$$
$$f((a,b)) = f((b,a)) = 1.$$

- $\Gamma_n(G)$ = the set of all cycles in G of length n that start and end at some fixed point u.
- For a cycle $\gamma = e_1 e_2 \dots e_n \in \Gamma_n(G)$, we refer to

$$P_\gamma(e) := \frac{1}{n} |\{i \in \{1, 2, \dots, n\} : e_i = e\}|$$

as the empirical distribution of the cycle γ,

STATIONARY MARKOV CHAINS ON GRAPHS



- The (switch) function $f : E_G \to \mathbb{R}$ is defined by

$$f((a,a)) = f((b,b)) = 0,$$
$$f((a,b)) = f((b,a)) = 1.$$

- $\Gamma_n(G) =$ the set of all cycles in G of length n that start and end at some fixed point u.
- For a cycle $\gamma = e_1 e_2 \ldots e_n \in \Gamma_n(G)$, we refer to

$$P_\gamma(e) := \frac{1}{n}|\{i \in \{1,2,\ldots,n\} : e_i = e\}|$$

as the empirical distribution of the cycle γ, and to

$$\mathbf{E}_{P_\gamma}(f) = \sum_{e \in E_G} P_\gamma(e) f(e)$$

as the empirical average of f on the cycle γ.

Fact

For a cycle $\gamma = e_1 e_2 \cdots e_n \in \Gamma_n(G)$, the value $nE_{P_\gamma}(f) - f(e_1)$ is equal to the number of switches of the corresponding binary sequence $L_G(e_1)L_G(e_2)\cdots L_G(e_n)$.

Second-Order Types of Markov Chains

Fact

For a cycle $\gamma = e_1 e_2 \cdots e_n \in \Gamma_n(G)$, the value $nE_{P_\gamma}(f) - f(e_1)$ is equal to the number of switches of the corresponding binary sequence $L_G(e_1)L_G(e_2)\cdots L_G(e_n)$.

- Recall that $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$

Second-Order Types of Markov Chains

Fact

For a cycle $\gamma = e_1 e_2 \cdots e_n \in \Gamma_n(G)$, the value $nE_{P_\gamma}(f) - f(e_1)$ is equal to the number of switches of the corresponding binary sequence $L_G(e_1)L_G(e_2)\cdots L_G(e_n)$.

- Recall that $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- $\mathcal{M}(G; f, [0, d))$: the set of all stationary Markov chains P on G such that $\mathbf{E}_P(f) \in [0, d)$.

Second-Order Types of Markov Chains

Fact

For a cycle $\gamma = e_1 e_2 \cdots e_n \in \Gamma_n(G)$, the value $nE_{P_\gamma}(f) - f(e_1)$ is equal to the number of switches of the corresponding binary sequence $L_G(e_1)L_G(e_2)\cdots L_G(e_n)$.

- Recall that $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- $\mathcal{M}(G; f, [0,d))$: the set of all stationary Markov chains P on G such that $\mathbf{E}_P(f) \in [0,d)$.
- $\Gamma_n(G; f, [0,d)) := \{\gamma \in \Gamma_n(G) : \mathbf{E}_{P_\gamma}(f) \in [0,d)\}$.

SECOND-ORDER TYPES OF MARKOV CHAINS

Fact

For a cycle $\gamma = e_1 e_2 \cdots e_n \in \Gamma_n(G)$, the value $nE_{P_\gamma}(f) - f(e_1)$ is equal to the number of switches of the corresponding binary sequence $L_G(e_1)L_G(e_2)\cdots L_G(e_n)$.

- Recall that $S = \{\mathbf{x} \in \{0,1\}^n : \sum x_i \oplus x_{i+1} \leq dn - 1\}$
- $\mathcal{M}(G; f, [0,d))$: the set of all stationary Markov chains P on G such that $\mathbf{E}_P(f) \in [0,d)$.
- $\Gamma_n(G; f, [0,d)) := \{\gamma \in \Gamma_n(G) : \mathbf{E}_{P_\gamma}(f) \in [0,d)\}$.
- $|S| = |\Gamma_n(G; f, [0,d))|$

Second-Order Types of Markov Chains

Lemma (Marcus, Roth'92)

Let G be a primitive graph and $f : E_G \to \mathbb{R}^k$ be a function on the edges of G. Let U be an open and nonempty subset of $\mathbb{R}^k$. Then

$$\lim_{n\to\infty} \frac{1}{n} \log |\Gamma_n(G; f, U)| = \sup_{P \in \mathcal{M}(G;f,U)} H_P(Y|X).$$

Lemma (Marcus, Roth’92)

Let G be a primitive graph and $f : E_G \to \mathbb{R}^k$ be a function on the edges of G. Let U be an open and nonempty subset of $\mathbb{R}^k$. Then

$$\lim_{n\to\infty} \frac{1}{n} \log |\Gamma_n(G; f, U)| = \sup_{P\in\mathcal{M}(G;f,U)} H_P(Y|X).$$

This lemma is a consequence of well-known results on second-order types of Markov chains, cf. Boza’71, Davisson, Longo, Sgarro’81, Natarajan’85, Csiszár, Cover, Choi’87, and Csiszár’98.

SECOND-ORDER TYPES OF MARKOV CHAINS

Lemma (Marcus, Roth'92)

Let G be a primitive graph and $f : E_G \to \mathbb{R}^k$ be a function on the edges of G. Let U be an open and nonempty subset of $\mathbb{R}^k$. Then

$$\lim_{n\to\infty} \frac{1}{n} \log |\Gamma_n(G; f, U)| = \sup_{P \in \mathcal{M}(G; f, U)} H_P(Y|X).$$

Lemma (Convex duality principle, Marcus, Roth'92)

Let G be a graph and let $f : E_G \to \mathbb{R}^k$, $g : E_G \to \mathbb{R}^l$ be functions on the edges of G. Set $\phi = [f, g] : E_G \to \mathbb{R}^{k+l}$. Then for any $\boldsymbol{r} \in \mathbb{R}^k$ and $\boldsymbol{s} \in \mathbb{R}^l$,

$$\sup_{\substack{P \in \mathcal{M}(G): \\ \mathbf{E}_P(f) = \boldsymbol{r} \\ \mathbf{E}_P(g) \le \boldsymbol{s}}} H_P(Y|X) = \inf_{\substack{\boldsymbol{x} \in \mathbb{R}^k \\ \boldsymbol{z} \in \mathbb{R}^l_{\ge 0}}} \{\boldsymbol{x} \cdot \boldsymbol{r} + \boldsymbol{z} \cdot \boldsymbol{s} + \log \lambda_{G;\phi}(\boldsymbol{x}, \boldsymbol{z})\}.$$

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \text{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.

- Graph $G \times G$:
  (I) Vertex set $V_{G\times G} = V_G \times V_G = \{\langle u, u'\rangle : u, u' \in V_G\}$
  (II) Edge set $E_{G\times G} = E_G \times E_G = \{\langle e, e'\rangle : e, e' \in E_G\}$

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.

- Graph $G \times G$:
  (I) Vertex set $V_{G\times G} = V_G \times V_G = \{\langle u, u' \rangle : u, u' \in V_G\}$
  (II) Edge set $E_{G\times G} = E_G \times E_G = \{\langle e, e' \rangle : e, e' \in E_G\}$

- We define two functions $f^{(1)}$ and $f^{(2)}$ on $E_{G\times G}$ by

$$f^{(1)}(\langle e, e' \rangle) = f(e), \quad f^{(2)}(\langle e, e' \rangle) = f(e')$$

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.
- Graph $G \times G$:
  (I) Vertex set $V_{G\times G} = V_G \times V_G = \{\langle u, u' \rangle : u, u' \in V_G\}$
  (II) Edge set $E_{G\times G} = E_G \times E_G = \{\langle e, e' \rangle : e, e' \in E_G\}$
- We define two functions $f^{(1)}$ and $f^{(2)}$ on $E_{G\times G}$ by

$$f^{(1)}(\langle e, e' \rangle) = f(e), \quad f^{(2)}(\langle e, e' \rangle) = f(e')$$

  a function $\Delta : E_{G\times G} \to \mathbb{R}$ by

$$\Delta(\langle e, e' \rangle) = \begin{cases} 1 & \text{if } L_G(e) \neq L_G(e') \\ 0 & \text{otherwise,} \end{cases}$$

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.
- Graph $G \times G$:
  (I) Vertex set $V_{G\times G} = V_G \times V_G = \{\langle u, u' \rangle : u, u' \in V_G\}$
  (II) Edge set $E_{G\times G} = E_G \times E_G = \{\langle e, e' \rangle : e, e' \in E_G\}$
- We define two functions $f^{(1)}$ and $f^{(2)}$ on $E_{G\times G}$ by

$$f^{(1)}(\langle e, e' \rangle) = f(e), \quad f^{(2)}(\langle e, e' \rangle) = f(e')$$

  a function $\Delta : E_{G\times G} \to \mathbb{R}$ by

$$\Delta(\langle e, e' \rangle) = \begin{cases} 1 & \text{if } L_G(e) \neq L_G(e') \\ 0 & \text{otherwise,} \end{cases}$$

  and a function $\varphi : E_{G\times G} \to \mathbb{R}^3$ by $\varphi = [f^{(1)}, f^{(2)}, \Delta]$.

The Computation of $|B|$

- Recall that $B = \{(\mathbf{w}, \mathbf{w}') \in S \times S : \mathrm{dist}(\mathbf{w}, \mathbf{w}') \leq \delta n - 1\}$.
- Graph $G \times G$:
  - (I) Vertex set $V_{G\times G} = V_G \times V_G = \{\langle u, u'\rangle : u, u' \in V_G\}$
  - (II) Edge set $E_{G\times G} = E_G \times E_G = \{\langle e, e'\rangle : e, e' \in E_G\}$
- We define two functions $f^{(1)}$ and $f^{(2)}$ on $E_{G\times G}$ by

$$f^{(1)}(\langle e, e'\rangle) = f(e), \quad f^{(2)}(\langle e, e'\rangle) = f(e')$$

a function $\Delta : E_{G\times G} \to \mathbb{R}$ by

$$\Delta(\langle e, e'\rangle) = \begin{cases} 1 & \text{if } L_G(e) \neq L_G(e') \end{cases}$$

Lemma

Set

$$\mathcal{G}([0,d),\delta) = \sup_{Q \in \mathcal{M}(G\times G;\varphi,[0,d)\times[0,d)\times[0,\delta))} H_Q(Y|X)$$

Then $|\mathcal{G}([0,d),\delta)| = \lim_{n\to\infty} \frac{1}{n} \log |B|$.

Spectral Radius

- For a function $f : E_G \to \mathbb{R}^k$, let $A_{G;f}(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^k$, be the matrix function indexed by the states of G with entries

$$[A_{G;f}(\boldsymbol{x})]_{u,v} = \begin{cases} 2^{-\boldsymbol{x}\cdot f((u,v))} & \text{if } (u,v) \in E_G \\ 0 & \text{otherwise.} \end{cases}$$

Spectral Radius

- For a function $f : E_G \to \mathbb{R}^k$, let $A_{G;f}(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^k$, be the matrix function indexed by the states of G with entries

$$[A_{G;f}(\boldsymbol{x})]_{u,v} = \begin{cases} 2^{-\boldsymbol{x} \cdot f((u,v))} & \text{if } (u,v) \in E_G \\ 0 & \text{otherwise.} \end{cases}$$

and let $\lambda_{G;f}(\boldsymbol{x})$ denote the spectral radius of $A_{G;f}(\boldsymbol{x})$.

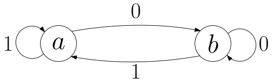
Spectral Radius

- For a function $f : E_G \to \mathbb{R}^k$, let $A_{G;f}(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^k$, be the matrix function indexed by the states of G with entries

$$[A_{G;f}(\boldsymbol{x})]_{u,v} = \begin{cases} 2^{-\boldsymbol{x}\cdot f((u,v))} & \text{if } (u,v) \in E_G \\ 0 & \text{otherwise.} \end{cases}$$

and let $\lambda_{G;f}(\boldsymbol{x})$ denote the spectral radius of $A_{G;f}(\boldsymbol{x})$.

Example



$$A_{G;f}(x) = \begin{array}{c c} & \begin{array}{cc} a & b \end{array} \\ \begin{array}{c} a \\ b \end{array} & \left[\begin{array}{cc} 1 & 2^{-x} \\ 2^{-x} & 1 \end{array} \right] \end{array}$$

$$\lambda_{G;f}(x) = 2^{-x} + 1$$

Lower Bound II: Markov Type

Markov Type Lower Bound

$$C(d,\delta) \geq \sup_{p\in[0,d]} \left\{ 2\,h(p) - \inf_{\substack{x\in\mathbb{R}\\ z\in\mathbb{R}_{\geq 0}}} \{2px + \delta z + \log \lambda_{G\times G;\varphi'}(x,z)\} \right\}$$

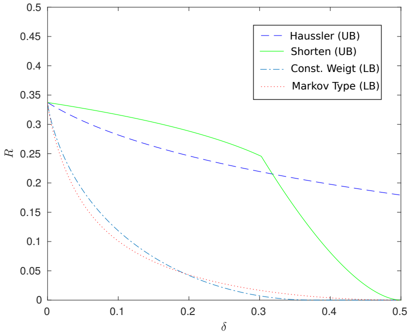
where

$$\lambda_{G\times G;\varphi'}(x,z) = \frac{1}{2}\Big((4^{-x}+1)(2^{-z}+1)+$$
$$\sqrt{(4^{-x}+1)^2 4^{-z} - 2(16^{-x} - 6\cdot 4^{-x}+1)2^{-z} + (4^{-x}+1)^2}\Big).$$

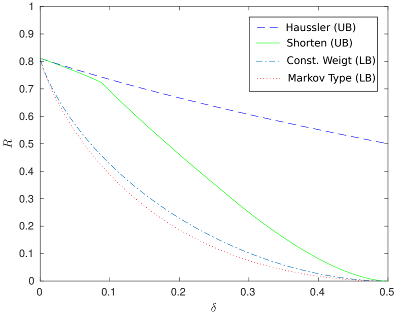
Proof. Follows the same line of reasoning as in [Marcus and Roth (1992)] where improved GV bound for constrained systems has been developed, based on stationary Markov chain analysis.

EXAMPLES: $\mathrm{VCD}(\mathcal{C}) = dn$

$d = \frac{1}{16}$

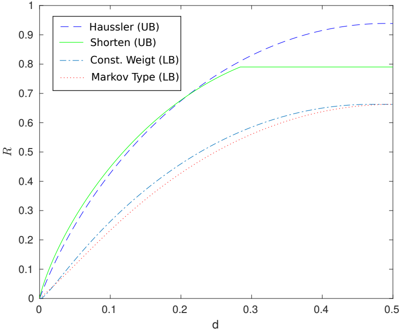


$d = \frac{1}{4}$

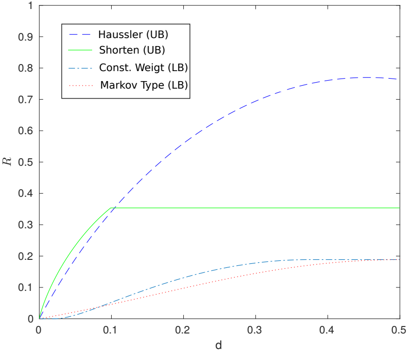


EXAMPLES: $\mathrm{dist}(\mathcal{C}) = \delta n$

$\delta = \frac{1}{16}$



$\delta = \frac{1}{4}$



FURTHER RESEARCH

- Upper bound
  - Minimum distance for prefixes in projection
  - Smart shifting?
  - VCD constraint in linear programming?
- Lower bound
  - Asymmetric Markov type
  - Other sets S
- Applications

Further Research

- Upper bound
  - Minimum distance for prefixes in projection
  - Smart shifting?
  - VCD constraint in linear programming?
- Lower bound
  - Asymmetric Markov type
  - Other sets S
- Applications

Thank You !