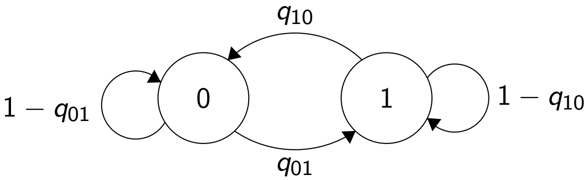
Novel Lower Bounds on the Entropy Rate of Binary Hidden Markov Processes

Or Ordentlich
MIT

ISIT,
Barcelona,
July 11, 2016

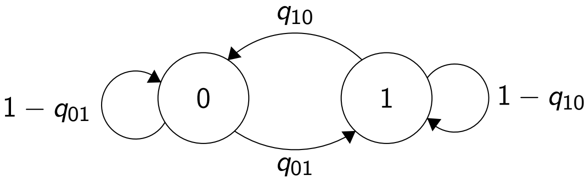
Binary Markov Processes



$$\mathbf{P} = \begin{bmatrix} 1 - q_{01} & q_{01} \\ q_{10} & 1 - q_{10} \end{bmatrix}, \qquad \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi} = [\pi_0 \ \pi_1]$$

$$X_1 \sim \text{Bernoulli}(\pi_1), \qquad \Pr(X_n = j | X_{n-1} = i, X_{n-2}, \ldots, X_1) = \mathbf{P}_{ij}$$

Binary Markov Processes



$$\mathbf{P} = \begin{bmatrix} 1 - q_{01} & q_{01} \\ q_{10} & 1 - q_{10} \end{bmatrix}, \qquad \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi} = [\pi_0 \ \pi_1]$$

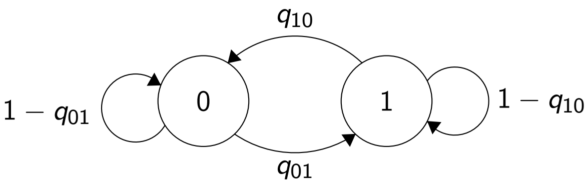
$$X_1 \sim \mathrm{Bernoulli}(\pi_1), \qquad \Pr(X_n = j | X_{n-1} = i, X_{n-2}, \ldots, X_1) = \mathbf{P}_{ij}$$

Entropy Rate

For a stationary process $\{X_n\}$ the entropy rate is defined as

$$\bar{H}(X) \triangleq \lim_{n\to\infty} \frac{H(X_1, \ldots, X_n)}{n} = \lim_{n\to\infty} H(X_n | X_n - 1, \ldots, X_1)$$

Binary Markov Processes



$$\mathbf{P} = \begin{bmatrix} 1 - q_{01} & q_{01} \\ q_{10} & 1 - q_{10} \end{bmatrix}, \qquad \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi} = [\pi_0 \ \pi_1]$$

$$X_1 \sim \text{Bernoulli}(\pi_1), \qquad \Pr(X_n = j | X_{n-1} = i, X_{n-2}, \ldots, X_1) = \mathbf{P}_{ij}$$

Entropy Rate

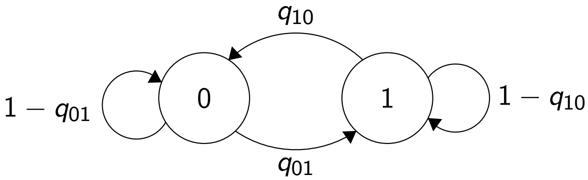
For the Markov process above

$$\bar{H}(X) = H(X_n|X_{n-1}) = \pi_0 h(q_{01}) + \pi_1 h(q_{10})$$

$$h(\alpha) \triangleq -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha)$$

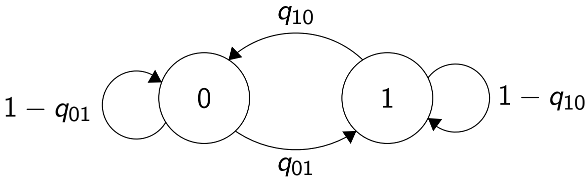
Binary Hidden Markov Processes

$\{X_n\}$:

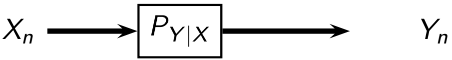


Binary Hidden Markov Processes

$\{X_n\}$:

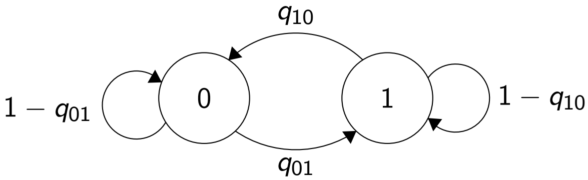


$\{Y_n\}$:

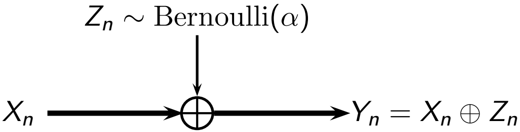


Binary Hidden Markov Processes

$\{X_n\}$:

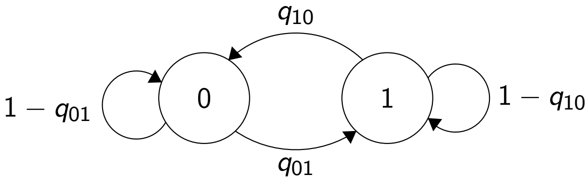


$\{Y_n\}$:

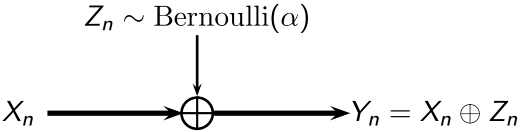


Binary Hidden Markov Processes

$\{X_n\}$:



$\{Y_n\}$:

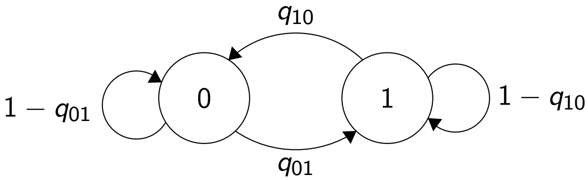


Entropy Rate Unknown

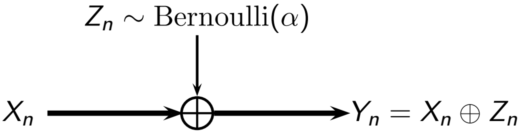
$$\bar{H}(Y) = f(\alpha, q_{10}, q_{01}) = ???$$

Binary Hidden Markov Processes

$\{X_n\}$:



$\{Y_n\}$:



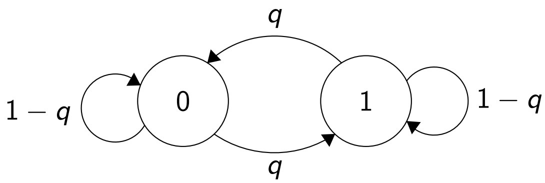
Entropy Rate Unknown

$$\bar{H}(Y) = f(\alpha, q_{10}, q_{01}) = ???$$

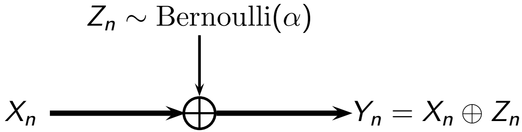
Our contribution: new lower bounds on $\bar{H}(Y)$

Binary Symmetric Hidden Markov Processes

$\{X_n\}$:



$\{Y_n\}$:



Entropy Rate Unknown

$$\bar{H}(Y) = f(\alpha, q) = ???$$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\ldots,Y_1,X_0) \leq \bar{H}(Y) \leq H(Y_n|Y_{n-1}\ldots,Y_0)$$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\ldots,Y_1,X_0) \leq \bar{H}(Y) \leq H(Y_n|Y_{n-1}\ldots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\ldots,Y_1,X_0) \le \bar{H}(Y) \le H(Y_n|Y_{n-1}\ldots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$H(Y_1,\ldots,Y_n) \ge nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\ldots,X_n)}{n}\right)\right)$$

$$a*b = a(1-b)+b(1-a), \quad h^{-1}:[0,1]\to[0,1/2]$$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\dots,Y_1,X_0) \le \bar{H}(Y) \le H(Y_n|Y_{n-1}\dots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$H(Y_1,\dots,Y_n) \ge nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$

$$\Rightarrow \bar{H}(Y) \ge h\left(\alpha * h^{-1}\left(\lim_{n\to\infty}\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$

Continuity of MGL function $\varphi(u) = h\left(\alpha * h^{-1}(u)\right)$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\ldots,Y_1,X_0) \leq \bar{H}(Y) \leq H(Y_n|Y_{n-1}\ldots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$H(Y_1,\ldots,Y_n) \geq nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\ldots,X_n)}{n}\right)\right)$$

$$\Rightarrow \bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty}\frac{H(X_1,\ldots,X_n)}{n}\right)\right)$$

$$\Rightarrow \bar{H}(Y) \geq h\left(\alpha * q\right)$$

$$\bar{H}(X) = h(q)$$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\dots,Y_1,X_0) \le \bar{H}(Y) \le H(Y_n|Y_{n-1}\dots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$\begin{aligned}
&H(Y_1,\dots,Y_n) \ge nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\dots,X_n)}{n}\right)\right)\\
\Rightarrow&\bar{H}(Y) \ge h\left(\alpha * h^{-1}\left(\lim_{n\to\infty}\frac{H(X_1,\dots,X_n)}{n}\right)\right)\\
\Rightarrow&\bar{H}(Y) \ge h\left(\alpha * q\right)
\end{aligned}$$

The same as Cover-Thomas bound of order $n = 1$

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\dots,Y_1,X_0) \le \bar{H}(Y) \le H(Y_n|Y_{n-1}\dots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$H(Y_1,\dots,Y_n) \ge nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$
$$\Rightarrow \bar{H}(Y) \ge h\left(\alpha * h^{-1}\left(\lim_{n\to\infty}\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$
$$\Rightarrow \bar{H}(Y) \ge h\left(\alpha * q\right)$$

Standard MGL gives a weak estimate

Binary Symmetric HMP - Simple Bounds

"Cover-Thomas bounds":

$$H(Y_n|Y_{n-1}\dots,Y_1,X_0) \leq \bar{H}(Y) \leq H(Y_n|Y_{n-1}\dots,Y_0)$$

Accuracy improves exponentially with n [Birch'62]

Simple lower bound by Mrs. Gerber's Lemma:

$$H(Y_1,\dots,Y_n) \geq nh\left(\alpha * h^{-1}\left(\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$

$$\Rightarrow \bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty}\frac{H(X_1,\dots,X_n)}{n}\right)\right)$$

$$\Rightarrow \bar{H}(Y) \geq h(\alpha * q)$$

Standard MGL gives a weak estimate

We will use an improved version of MGL

Samorodnitsky's MGL

- $\mathbf{X}, \mathbf{Y} \in \{0,1\}^n$ are the input and output of a $\mathrm{BSC}(\alpha)$
- $\lambda \triangleq (1-2\alpha)^2$
- The projection of $\mathbf{X}$ onto a subset of coordinates $S \subseteq [n]$ is

$$\mathbf{X}_S \triangleq \{X_i \ : i \in S\}$$

- Let V be a random subset of $[n]$ generated by independently sampling each element i with probability λ

Theorem [Samorodnitsky'15]

$$H(\mathbf{Y}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{X}_V|V)}{\lambda n}\right)\right)$$

Samorodnitsky's MGL

- $\mathbf{X}, \mathbf{Y} \in \{0,1\}^n$ are the input and output of a $\mathrm{BSC}(\alpha)$
- $\lambda \triangleq (1-2\alpha)^2$
- The projection of $\mathbf{X}$ onto a subset of coordinates $S \subseteq [n]$ is

$$\mathbf{X}_S \triangleq \{X_i \ : i \in S\}$$

- Let V be a random subset of $[n]$ generated by independently sampling each element i with probability λ

Theorem [Samorodnitsky'15]

$$H(\mathbf{Y}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{X}_V|V)}{\lambda n}\right)\right)$$

By Han's inequality $\frac{H(\mathbf{X}_V|V)}{\lambda n}$ is nonincreasing* in λ

Samorodnitsky's MGL

- $\mathbf{X}, \mathbf{Y} \in \{0,1\}^n$ are the input and output of a $\mathrm{BSC}(\alpha)$
- $\lambda \triangleq (1-2\alpha)^2$
- The projection of $\mathbf{X}$ onto a subset of coordinates $S \subseteq [n]$ is

$$\mathbf{X}_S \triangleq \{X_i \ : i \in S\}$$

- Let V be a random subset of $[n]$ generated by independently sampling each element i with probability λ

Theorem [Samorodnitsky'15]

$$H(\mathbf{Y}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{X}_V|V)}{\lambda n}\right)\right)$$

$\Rightarrow$ The new bound is stronger than MGL

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

$\varphi(x) \triangleq h\left(\alpha * h^{-1}(x)\right)$

Samorodnitsky's MGL - Proof Outline

$$
\begin{aligned}
H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) \\
&\geq \sum_{i=1}^{n} \varphi(H(X_i|Y_1^{i-1})) \\
&= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right)
\end{aligned}
$$

Need to upper bound

$I(X_i; Y_1^{i-1})$

Samorodnitsky's MGL - Proof Outline

$$
\begin{aligned}
H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\
&\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\
&= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right)
\end{aligned}
$$

Need to upper bound

$$
I(X_i; Y_1^{i-1}) = I(X_i; Y_1^{i-2}) + I(X_i; Y_{i-1} | Y_1^{i-2})
$$

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

Need to upper bound

$$\begin{aligned} I(X_i; Y_1^{i-1}) &= I(X_i; Y_1^{i-2}) + I(X_i; Y_{i-1} | Y_1^{i-2}) \\ \text{(SDPI)} &\leq I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1} | Y_1^{i-2}) \end{aligned}$$

Samorodnitsky's MGL - Proof Outline

$$
\begin{aligned}
H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\
&\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\
&= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right)
\end{aligned}
$$

Need to upper bound

$$
\begin{aligned}
I(X_i; Y_1^{i-1}) &= I(X_i; Y_1^{i-2}) + I(X_i; Y_{i-1} | Y_1^{i-2}) \\
\text{(SDPI)} &\leq I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1} | Y_1^{i-2}) \\
&= (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda \left(I(X_i; Y_1^{i-2}) + I(X_i; X_{i-1} | Y_1^{i-2}) \right)
\end{aligned}
$$

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

Need to upper bound

$$\begin{aligned} I(X_i; Y_1^{i-1}) &= I(X_i; Y_1^{i-2}) + I(X_i; Y_{i-1} | Y_1^{i-2}) \\ \text{(SDPI)} &\leq I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1} | Y_1^{i-2}) \\ &= (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda \left(I(X_i; Y_1^{i-2}) + I(X_i; X_{i-1} | Y_1^{i-2}) \right) \\ \text{(Chain Rule)} &= (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1}, Y_1^{i-2}) \end{aligned}$$

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

We have $I(X_i; Y_1^{i-1}) \leq (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1}, Y_1^{i-2})$

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

We have $I(X_i; Y_1^{i-1}) \leq (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1}, Y_1^{i-2})$

Using this to form and solve a suitable linear program [Samorodnitsky'15], or using induction [Polyanskiy-Wu'16], gives

Samorodnitsky's MGL - Proof Outline

$$\begin{aligned} H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\ &\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\ &= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right) \end{aligned}$$

We have $I(X_i; Y_1^{i-1}) \leq (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1}, Y_1^{i-2})$

Using this to form and solve a suitable linear program [Samorodnitsky'15], or using induction [Polyanskiy-Wu'16], gives

$$I(X_i; Y_1^{i-1}) \leq I(X_i; {Y_{\mathrm{BEC},(1-\lambda)}}_1^{i-1})$$

Samorodnitsky's MGL - Proof Outline

$$
\begin{aligned}
H(\mathbf{Y}) &= \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \\
&\geq \sum_{i=1}^{n} \varphi(H(X_i | Y_1^{i-1})) \\
&= \sum_{i=1}^{n} \varphi\left(H(X_i) - I(X_i; Y_1^{i-1})\right)
\end{aligned}
$$

We have $I(X_i; Y_1^{i-1}) \leq (1-\lambda) I(X_i; Y_1^{i-2}) + \lambda I(X_i; X_{i-1}, Y_1^{i-2})$

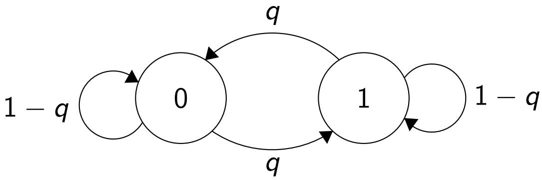
Using this to form and solve a suitable linear program [Samorodnitsky'15], or using induction [Polyanskiy-Wu'16], gives

$$
I(X_i; Y_1^{i-1}) \leq I(X_i; {Y_{\mathrm{BEC},(1-\lambda)}}_1^{i-1})
$$

From here, standard arguments give the theorem

Back to HMPs

$\{X_n\}$:



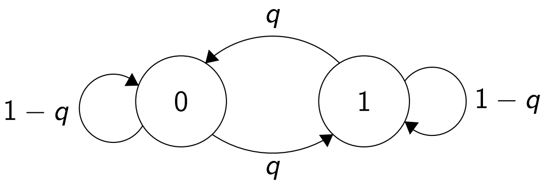
$Y_n = X_n \oplus Z_n$

Theorem [Samorodnitsky'15]

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}\right)\right)$$

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

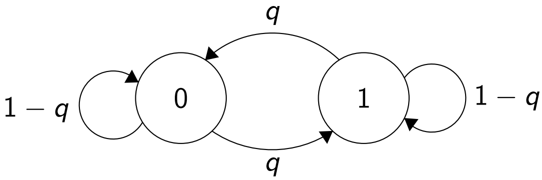
Theorem [Samorodnitsky'15]

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}\right)\right)$$

Need to find $\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}$

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

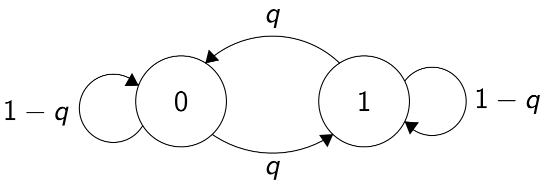
Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \mathbb{E}H(X_{G+1}|X_1)$$

where $G \sim \mathrm{Geometric}(\lambda)$.

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n \to \infty} \frac{H(X_V | V)}{\lambda n} = \mathbb{E} H(X_{G+1} | X_1)$$

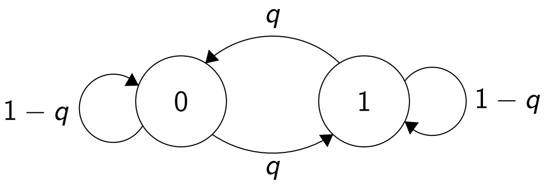
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$q^{*k} \triangleq \underbrace{q * q * \cdots * q}_{k \text{ times}}$$

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \mathbb{E}H(X_{G+1}|X_1)$$

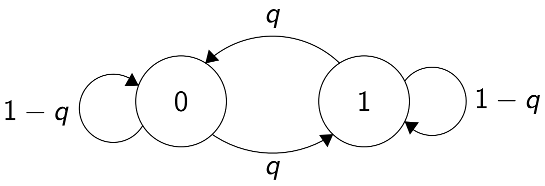
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$q^{*k} \triangleq \underbrace{q * q * \cdots * q}_{k \text{ times}} = \frac{1-(1-2q)^k}{2}$$

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \mathbb{E}h\left(q^{*G}\right)$$

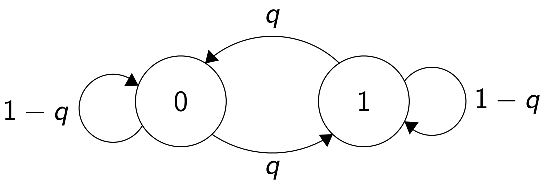
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$q^{*k} \triangleq \underbrace{q * q * \cdots * q}_{k \text{ times}} = \frac{1-(1-2q)^k}{2}$$

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

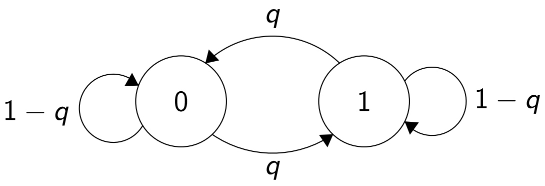
Theorem

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\mathbb{E}h\left(q^{*G}\right)\right)\right),$$

where $G \sim \text{Geometric}((1-2\alpha)^2)$.

Back to HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

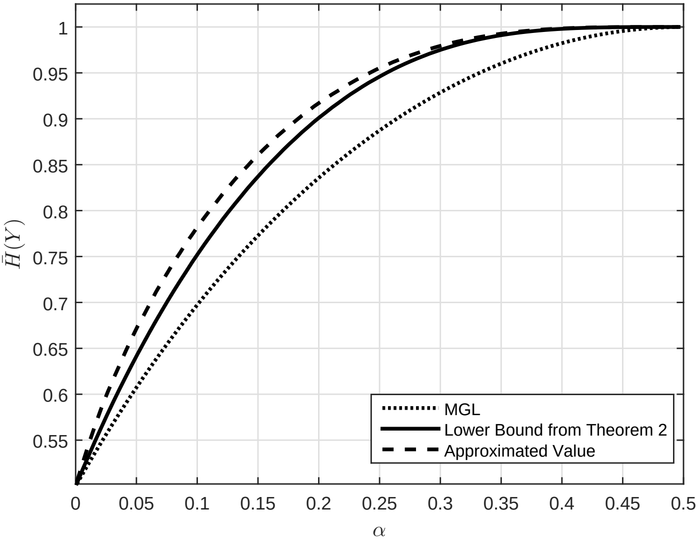
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\mathbb{E}h\left(q^{*G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

- For small α (high-SNR) bound approaches MGL
- For large α (low-SNR) much better than MGL

New Bound - Behavior with α

$q = 0.11$



New Bound - Behavior with α

Theorem (low-SNR)

Let q be fixed and $\alpha = \frac{1}{2} - \epsilon$. Then

$$\bar{H}(Y) \geq 1 - 16\epsilon^4 \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} \frac{(1-2q)^{2k}}{1-(1-2q)^{2k}} + o(\epsilon^4)$$

New Bound - Behavior with α

Theorem (low-SNR)

Let q be fixed and $\alpha = \frac{1}{2} - \epsilon$. Then

$$\bar{H}(Y) \geq 1 - 16\epsilon^4 \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} \frac{(1-2q)^{2k}}{1-(1-2q)^{2k}} + o(\epsilon^4)$$

Our result shows that

$$\limsup_{\epsilon \to 0} \frac{1 - \bar{H}(Y)}{\epsilon^4} \leq 16 \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} \frac{(1-2q)^{2k}}{1-(1-2q)^{2k}}$$

New Bound - Behavior with α

Theorem (low-SNR)

Let q be fixed and $\alpha = \frac{1}{2} - \epsilon$. Then

$$\bar{H}(Y) \geq 1 - 16\epsilon^4 \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} \frac{(1-2q)^{2k}}{1-(1-2q)^{2k}} + o(\epsilon^4)$$

Our result shows that

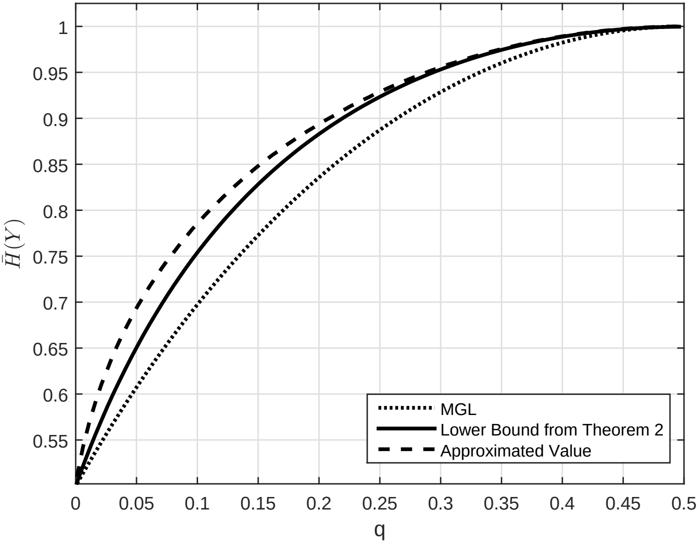
$$\limsup_{\epsilon \to 0} \frac{1-\bar{H}(Y)}{\epsilon^4} \leq 16 \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} \frac{(1-2q)^{2k}}{1-(1-2q)^{2k}}$$

Best previously known bound [E. Ordentlich-Weissman'11]:

$$\limsup_{\epsilon \to 0} \frac{1-\bar{H}(Y)}{\epsilon^4} \leq \frac{2\log(e)(1-2q)^2(1-4q+16q^2-32q^3+32q^4)}{q^2}$$

New Bound - Behavior with q

$\alpha = 0.11$



New Bound - Behavior with q

Theorem (fast transitions)

Let α be fixed and $q = \frac{1}{2} - \epsilon$. Then

$$1 - \bar{H}(Y) \leq 2 \log(e)(1 - 2\alpha)^4 \epsilon^2 + \mathcal{O}(\epsilon^4)$$

New Bound - Behavior with q

Theorem (fast transitions)

Let α be fixed and $q = \frac{1}{2} - \epsilon$. Then

$$1 - \bar{H}(Y) \leq 2\log(e)(1-2\alpha)^4\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Recovers the expression found in [E. Ordentlich-Weissman'11]

New Bound - Behavior with q

Theorem (fast transitions)

Let α be fixed and $q = \frac{1}{2} - \epsilon$. Then

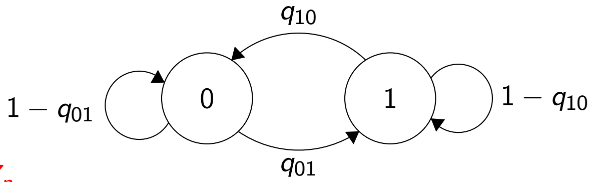
$$1 - \bar{H}(Y) \leq 2\log(e)(1 - 2\alpha)^4 \epsilon^2 + \mathcal{O}(\epsilon^4)$$

Recovers the expression found in [E. Ordentlich-Weissman'11]

Bound is tight [E. Ordentlich-Weissman'11]

Nonsymmetric HMPs

$\{X_n\}$:



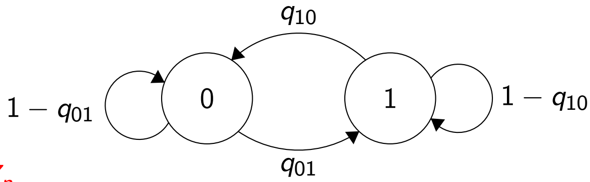
$Y_n = X_n \oplus Z_n$

Theorem [Samorodnitsky'15]

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}\right)\right)$$

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

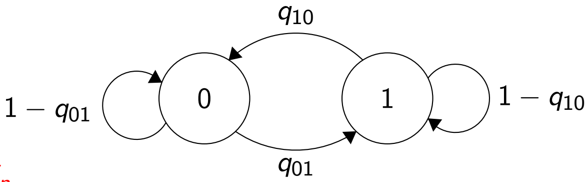
Theorem [Samorodnitsky'15]

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}\right)\right)$$

Need to find $\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n}$

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

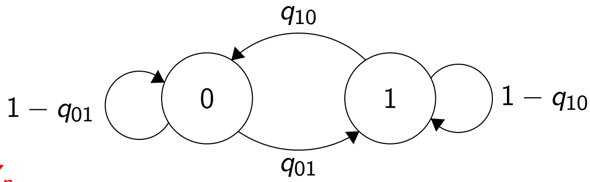
Proposition

$$\lim_{n \to \infty} \frac{H(X_V | V)}{\lambda n} = \mathbb{E} H(X_{G+1} | X_1)$$

where $G \sim \mathrm{Geometric}(\lambda)$.

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \mathbb{E}H(X_{G+1}|X_1)$$

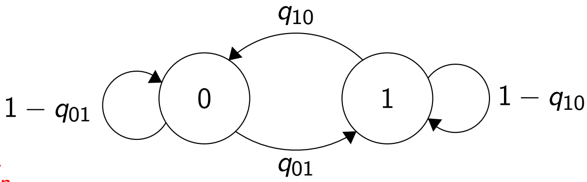
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$\mathbf{P} = \begin{bmatrix} 1-q_{01} & q_{01} \\ q_{10} & 1-q_{10} \end{bmatrix},$$

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \mathbb{E}H(X_{G+1}|X_1)$$

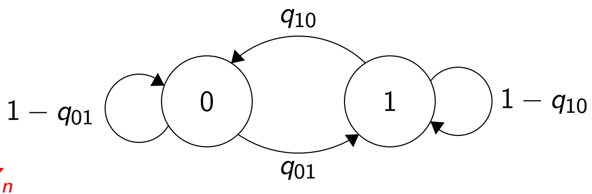
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$\mathbf{P} = \begin{bmatrix} 1-q_{01} & q_{01} \\ q_{10} & 1-q_{10} \end{bmatrix}, \qquad q_{ij}^{\#k} \triangleq \left(\mathbf{P}^k\right)_{ij} = \Pr\left(X_n = j | X_{n-k} = i\right)$$

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Proposition

$$\lim_{n\to\infty} \frac{H(X_V|V)}{\lambda n} = \pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)$$

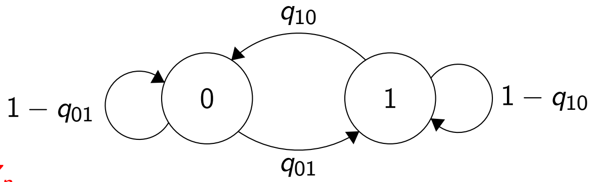
where $G \sim \mathrm{Geometric}(\lambda)$.

Define:

$$\mathbf{P} = \begin{bmatrix} 1-q_{01} & q_{01} \\ q_{10} & 1-q_{10} \end{bmatrix}, \qquad q_{ij}^{\#k} \triangleq \left(\mathbf{P}^k\right)_{ij} = \Pr\left(X_n = j | X_{n-k} = i\right)$$

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

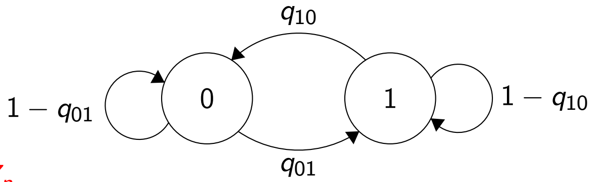
Theorem

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

Nonsymmetric HMPs

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

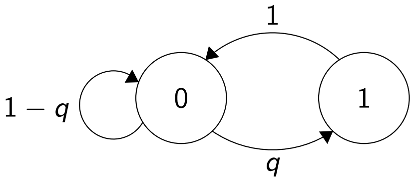
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

- For small α (high-SNR) bound approaches MGL
- For large α (low-SNR) much better than MGL

Special Case - $(1, \infty)$-RLL Constraint

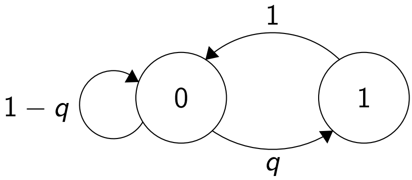
$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

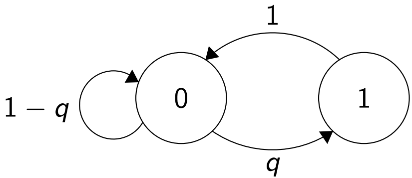
Theorem

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\# G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\# G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

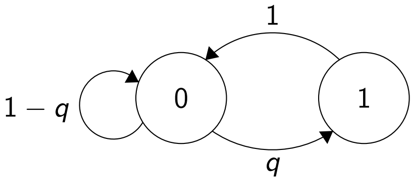
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

$$\mathbf{P} = \begin{bmatrix} 1-q & q \\ 1 & 0 \end{bmatrix}, \qquad \pi_0 = \frac{1}{1+q},\ \pi_1 = \frac{q}{1+q}$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

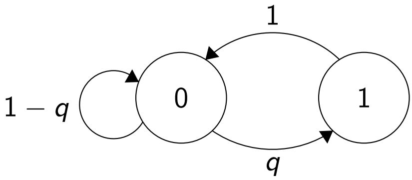
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

$$\mathbf{P}^k = \begin{bmatrix} \frac{1-(-q)^{k+1}}{1+q} & \frac{q+(-q)^{k+1}}{1+q} \\ \frac{1-(-q)^k}{1+q} & \frac{q+(-q)^k}{1+q} \end{bmatrix}, \qquad \pi_0 = \frac{1}{1+q}, \ \pi_1 = \frac{q}{1+q}$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

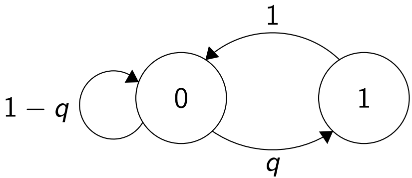
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}\left(\pi_0 \mathbb{E} h\left(q_{01}^{\#G}\right) + \pi_1 \mathbb{E} h\left(q_{10}^{\#G}\right)\right)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

$$\beta \triangleq \frac{1}{1+q}\mathbb{E} h\left(\frac{1-(-q)^{G+1}}{1+q}\right) + \frac{q}{1+q}\mathbb{E} h\left(\frac{1-(-q)^{G}}{1+q}\right)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

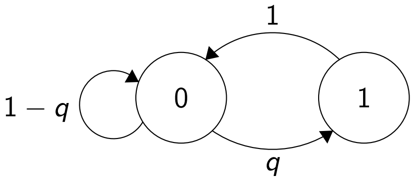
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}(\beta)\right),$$

where $G \sim \text{Geometric}((1-2\alpha)^2)$.

$$\beta \triangleq \frac{1}{1+q}\mathbb{E}h\left(\frac{1-(-q)^{G+1}}{1+q}\right) + \frac{q}{1+q}\mathbb{E}h\left(\frac{1-(-q)^{G}}{1+q}\right)$$

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

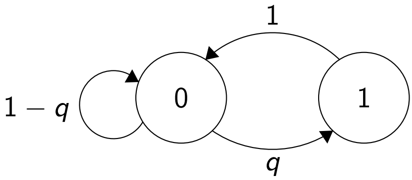
$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}(\beta)\right),$$

where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

$$\beta = \mathbb{E}H(X_{G+1}|X_1)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}(\beta)\right),$$

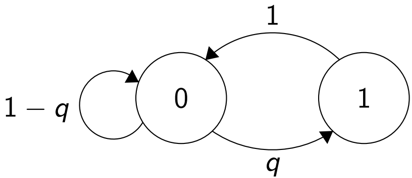
where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

Proposition

$$\beta \geq h(\pi_1) - \frac{(1-2\alpha)^2 q}{(1+q)(1-4\alpha(1-\alpha)q)}\left(2h\left(\pi_1\right) - h\left(\frac{1-q}{1+q}\right)\right)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem

$$\bar{H}(Y) \geq h\left(\alpha * h^{-1}(\beta)\right),$$

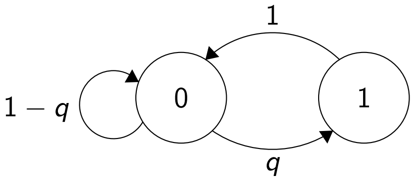
where $G \sim \mathrm{Geometric}((1-2\alpha)^2)$.

Proposition

$$\beta \geq h(\pi_1) - c\epsilon^2, \qquad \text{for } \alpha = \frac{1}{2} - \epsilon, c > 0$$

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

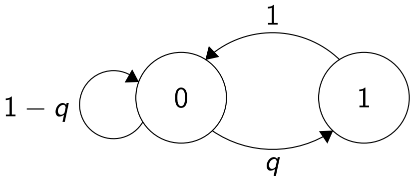
$$\bar{H}(Y) \geq h\left(\left(\frac{1}{2} - \epsilon\right) * h^{-1}\left(h(\pi_1) - c\epsilon^2\right)\right)$$

Proposition

$$\beta \geq h(\pi_1) - c\epsilon^2, \qquad \text{for } \alpha = \frac{1}{2} - \epsilon, c > 0$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



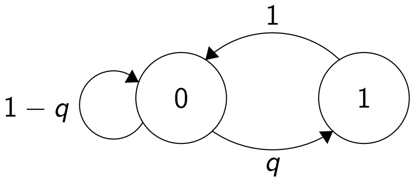
$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

$$\bar{H}(Y) \geq h\left(\left(\frac{1}{2} - \epsilon\right) * \left(\pi_1 - \frac{c}{h'(\pi_1)}\epsilon^2\right)\right)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



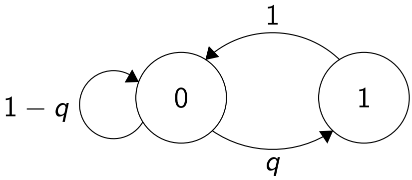
$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

$$\bar{H}(Y) \geq 1 - 2\log(e)(1 - 2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

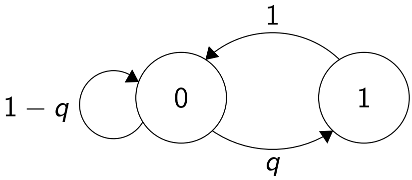
$$\bar{H}(Y) \geq 1 - 2\log(e)(1-2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Proposition: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Upper Bound

$$\bar{H}(Y) \leq H(Y_n)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

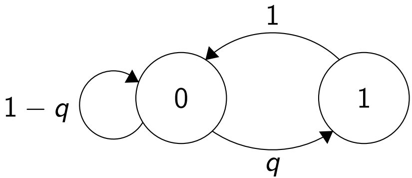
$$\bar{H}(Y) \geq 1 - 2\log(e)(1 - 2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Proposition: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Upper Bound

$$\bar{H}(Y) \leq h(\alpha * \pi_1)$$

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

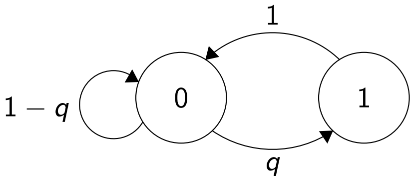
$$\bar{H}(Y) \geq 1 - 2\log(e)(1-2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Proposition: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Upper Bound

$$\bar{H}(Y) \leq h\left(\left(\frac{1}{2} - \epsilon\right) * \pi_1\right)$$

Special Case - $(1,\infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Lower Bound

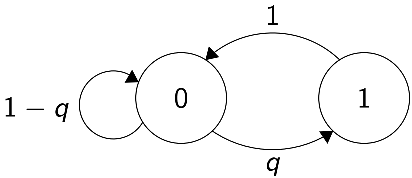
$$\bar{H}(Y) \geq 1 - 2\log(e)(1-2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Proposition: Low-SNR ($\alpha = \frac{1}{2} - \epsilon, 0 \leq q < 1$) Upper Bound

$$\bar{H}(Y) \leq 1 - 2\log(e)(1-2\pi_1)^2\epsilon^2 + \mathcal{O}(\epsilon^4)$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

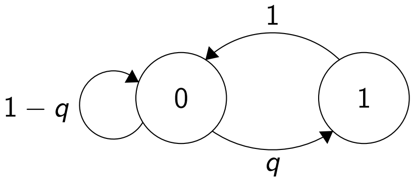
Theorem: Low-SNR

For $\alpha = \frac{1}{2} - \epsilon$ and $0 \leq q < 1$, we have

$$\lim_{\epsilon \to 0} \frac{1 - \bar{H}(Y)}{\epsilon^2} = 2 \log(e) \left(\frac{1-q}{1+q} \right)^2$$

Special Case - $(1, \infty)$-RLL Constraint

$\{X_n\}$:



$Y_n = X_n \oplus Z_n$

Theorem: Low-SNR

For $\alpha = \frac{1}{2} - \epsilon$ and $0 \leq q < 1$, we have

$$\lim_{\epsilon \to 0} \frac{1 - \bar{H}(Y)}{\epsilon^2} = 2 \log(e) \left(\frac{1-q}{1+q} \right)^2$$

Recovers result from Han-Marcus'07 and Pfister'11

Summary and Conclusions

We derived a new lower bound for the entropy rate of binary hidden Markov processes

Summary and Conclusions

We derived a new lower bound for the entropy rate of binary hidden Markov processes

The bound relies on a strengthened version of MGL, due to Samorodnitsky

Summary and Conclusions

We derived a new lower bound for the entropy rate of binary hidden Markov processes

The bound relies on a strengthened version of MGL, due to Samorodnitsky

We improved the best known bound for symmetric processes in low-SNR, and recovered the best known results in some other regimes

Summary and Conclusions

We derived a new lower bound for the entropy rate of binary hidden Markov processes

The bound relies on a strengthened version of MGL, due to Samorodnitsky

We improved the best known bound for symmetric processes in low-SNR, and recovered the best known results in some other regimes

Our technique can be generalized to hidden Markov processes over larger alphabets