Toward single-letter feedback capacity via structured auxiliary r.v.

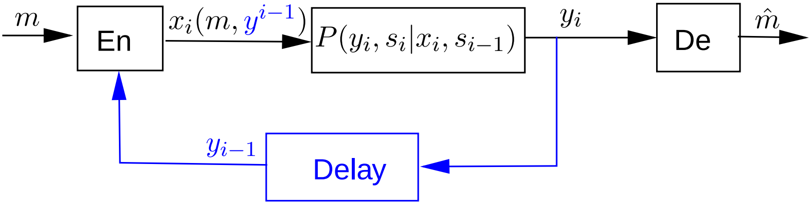
Haim Permuter

Ben-Gurion University

The ISL Colloquium
Feb. 2017

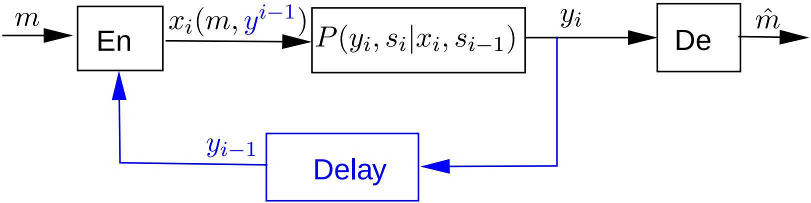
In my Ph.D. ...

considered channel with memory and feedback



In my Ph.D. ...

considered channel with memory and feedback



- Finite State Channel (FSC)

$$P(y_i, s_i|x_i, s_{i-1}, x^{i-1}, y^{i-1}, s^{i-2}) = P(y_i, s_i|x_i, s_{i-1})$$

In my Ph.D. ...

Theorem

For any FSC ***with*** *feedback* *[P.&Weissman&Goldsmith09]*

$$C_{FB} \geq \frac{1}{n} \max_{P(x^n||y^{n-1})} \min_{s_0} I(X^n \to Y^n | s_0) - \frac{\log |\mathcal{S}|}{n}$$

$$C_{FB} \leq \frac{1}{n} \max_{P(x^n||y^{n-1})} \max_{s_0} I(X^n \to Y^n | s_0) + \frac{\log |\mathcal{S}|}{n}$$

In my Ph.D. ...

Theorem

For any FSC with feedback *[P.&Weissman&Goldsmith09]*

$$C_{FB} \geq \frac{1}{n} \max_{P(x^n||y^{n-1})} \min_{s_0} I(X^n \to Y^n|s_0) - \frac{\log|\mathcal{S}|}{n}$$

$$C_{FB} \leq \frac{1}{n} \max_{P(x^n||y^{n-1})} \max_{s_0} I(X^n \to Y^n|s_0) + \frac{\log|\mathcal{S}|}{n}$$

- *directed information:*

$$I(X^n \to Y^n) = \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1})$$

In my Ph.D. ...

Theorem

For any FSC with feedback *[P.&Weissman&Goldsmith09]*

$$C_{FB} \geq \frac{1}{n} \max_{P(x^n||y^{n-1})} \min_{s_0} I(X^n \to Y^n|s_0) - \frac{\log|\mathcal{S}|}{n}$$

$$C_{FB} \leq \frac{1}{n} \max_{P(x^n||y^{n-1})} \max_{s_0} I(X^n \to Y^n|s_0) + \frac{\log|\mathcal{S}|}{n}$$

- *directed information:*

$$I(X^n \to Y^n) = \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1})$$

- *causally conditioned distribution:*

$$P(x^n||y^{n-1}) = \prod_{i=1}^{n} P(x_i|x^{i-1}, y^{i-1})$$

In my Ph.D. ...

Theorem

For any FSC with feedback *[P.&Weissman&Goldsmith09]*

$$C_{FB} \geq \frac{1}{n} \max_{P(x^n||y^{n-1})} \min_{s_0} I(X^n \to Y^n | s_0) - \frac{\log |\mathcal{S}|}{n}$$

$$C_{FB} \leq \frac{1}{n} \max_{P(x^n||y^{n-1})} \max_{s_0} I(X^n \to Y^n | s_0) + \frac{\log |\mathcal{S}|}{n}$$

- *directed information:*

$$I(X^n \to Y^n) = \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1})$$

- *causally conditioned distribution:*

$$P(x^n || y^{n-1}) = \prod_{i=1}^{n} P(x_i | x^{i-1}, y^{i-1})$$

- under mild conditions

$$C_{FB} = \lim_{n \to \infty} \frac{1}{n} \max_{P(x^n||y^{n-1})} I(X^n \to Y^n)$$

My adviser questions

$$C_{FB} = \lim_{n\to\infty} \frac{1}{n} \max_{P(x^n||y^{n-1})} I(X^n \to Y^n)$$

My adviser questions

$$\boxed{C_{FB} = \lim_{n\to\infty} \frac{1}{n} \max_{P(x^n||y^{n-1})} I(X^n \to Y^n)}$$

1. Is it computable?

My adviser questions

$$C_{FB} = \lim_{n\to\infty} \frac{1}{n} \max_{P(x^n||y^{n-1})} I(X^n \to Y^n)$$

1. Is it computable?
2. Is there single-letter expression?

My adviser questions

$$C_{FB} = \lim_{n\to\infty} \frac{1}{n} \max_{P(x^n||y^{n-1})} I(X^n \to Y^n)$$

1. Is it computable?
2. **Is there single-letter expression?**

Auxiliary random variable (r.v.)

- Auxiliary r.v. plays an important role in multi-users problems

Auxiliary random variable (r.v.)

- Auxiliary r.v. plays an important role in multi-users problems

Gelfand-Pinsker: $$C = \max_{P(u|s)P(x|u,s)} I(U;Y) - I(U;S)$$

Wyner-Ziv: $$R = \min_{P(u|x)} I(X;U|Y)$$

Auxiliary random variable (r.v.)

- Auxiliary r.v. plays an important role in multi-users problems

Gelfand-Pinsker: $C = \max_{P(u|s)P(x|u,s)} I(U;Y) - I(U;S)$

Wyner-Ziv: $R = \min_{P(u|x)} I(X;U|Y)$

- Auxiliary r.v. converts multi-letter into single-letter

Auxiliary random variable (r.v.)

- Auxiliary r.v. plays an important role in multi-users problems

Gelfand-Pinsker: $C = \max_{P(u|s)P(x|u,s)} I(U;Y) - I(U;S)$

Wyner-Ziv: $R = \min_{P(u|x)} I(X;U|Y)$

- Auxiliary r.v. converts multi-letter into single-letter
- Auxiliary r.v. are i.i.d.

Structured auxiliary r.v.

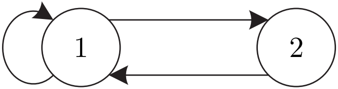
- Non i.i.d auxiliary r.v.

Structured auxiliary r.v.

- Non i.i.d auxiliary r.v.
- The auxiliary is represented by a graph

Structured auxiliary r.v.

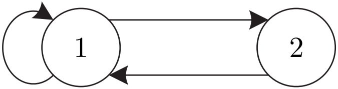
- Non i.i.d auxiliary r.v.
- The auxiliary is represented by a graph



- The graph induces a Markov process

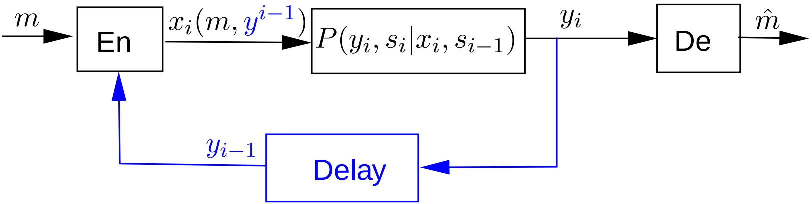
Structured auxiliary r.v.

- Non i.i.d auxiliary r.v.
- The auxiliary is represented by a graph



- The graph induces a Markov process
- The single-letter expression is evaluated with the stationary distribution

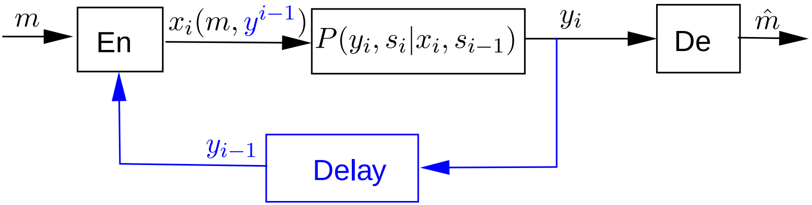
Unifilar FSC with feedback



- Finite State Channel (FSC)

$$P(y_i, s_i | x_i, s_{i-1}, x^{i-1}, y^{i-1}, s^{i-2}) = P(y_i, s_i | x_i, s_{i-1})$$

Unifilar FSC with feedback



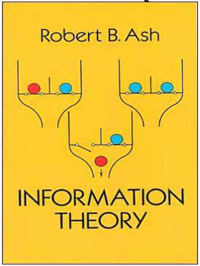
- Finite State Channel (FSC)

$$P(y_i, s_i|x_i, s_{i-1}, x^{i-1}, y^{i-1}, s^{i-2}) = P(y_i, s_i|x_i, s_{i-1})$$

- Unifilar FSC

$$S_i = f(S_{i-1}, X_i, Y_i)$$

Trapdoor Channel [Blackwell61]



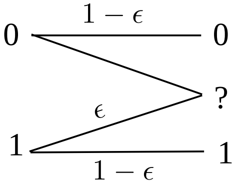
Ising Channel [Berger90]

$$y_i = \begin{cases} x_i, & \text{with prob. } \frac{1}{2} \\ x_{i-1}, & \text{with prob. } \frac{1}{2} \end{cases}$$

Dicode Erasure Channel [Pfister08]

$$y_i = \begin{cases} x_i - x_{i-1}, & \text{with prob. } 1-\epsilon \\ ?, & \text{with prob. } \epsilon \end{cases}$$

Erasure Channel with no repeated 1's



Feedback capacity

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

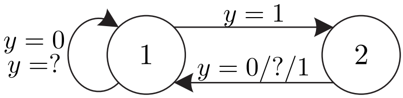
Feedback capacity

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

Q-graph defines a mapping: $(\mathcal{Q}_{i-1}, \mathcal{Y}_i) \to \mathcal{Q}_i$



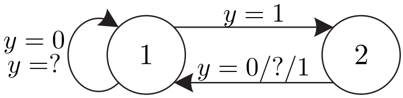
Feedback capacity

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

Q-graph defines a mapping: $(\mathcal{Q}_{i-1},\mathcal{Y}_i) \to \mathcal{Q}_i$



The Q-graph and $P(x|s,q)$ induces

$$p(s,q,x,y) = \pi(s,q)p(x|s,q)p(y|s,x)$$

Feedback capacity

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

- For all known cases the upper bound is tight $|\mathcal{Q}| \leq 4$,

Feedback capacity

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

- For all known cases the upper bound is tight $|\mathcal{Q}| \leq 4$,
- If $|\mathcal{Q}|$ unbounded then its also achievable, i.e.,

$$C_{fb} = \sup_{Q} \sup_{p(x|s,q)} I(X,S;Y|Q)$$

Sketch Proof

$$C_{fb} = \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n)$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1}) \\
&= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i, S_{i-1}(X^{i-1}, Y^{i-1}); Y_i|Y^{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1}) \\
&= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i, S_{i-1}; Y_i|Y^{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1}) \\
&= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i|Y^{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i|Y^{i-1})
\end{aligned}
$$

Sketch Proof

$$\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i|Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i|Y^{i-1}) \\
&\leq \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i|Q_{i-1}(Y^{i-1}))
\end{aligned}$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Y^{i-1}) \\
&\leq \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Q_{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Y^{i-1}) \\
&\leq \max_{P(x_i|s_{i-1},q_{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Q_{i-1})
\end{aligned}
$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i;Y_i|Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i,S_{i-1};Y_i|Y^{i-1}) \\
&\leq \max_{P(x_i|s_{i-1},q_{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i,S_{i-1};Y_i|Q_{i-1}) \\
&= \max_{P(x|s,q)} I(X,S;Y|Q)
\end{aligned}
$$

Examples

Theorem [Sabag/P./Pfiser16]

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

Ex1: Memoryless channel, $|\mathcal{S}| = 1$. Choose Q constant.

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q) = \sup_{p(x)} I(X;Y)$$

Examples

Theorem [Sabag/P./Pfiser16]

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

Ex1: Memoryless channel, $|\mathcal{S}| = 1$. Choose Q constant.

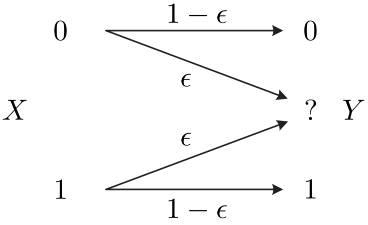
$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q) = \sup_{p(x)} I(X;Y)$$

Ex2: State known at the decoder and encoder. Choose $Q = S$

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y,S|Q) = \sup_{p(x|s)} I(X;Y|S)$$

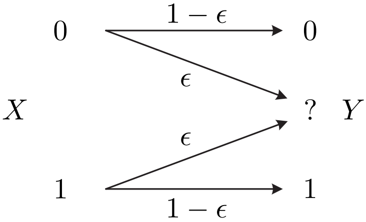
The input-constrained BEC [Sabag,Permuter,Kashyap 16]

- Binary erasure channel (BEC):



The input-constrained BEC [Sabag,Permuter,Kashyap 16]

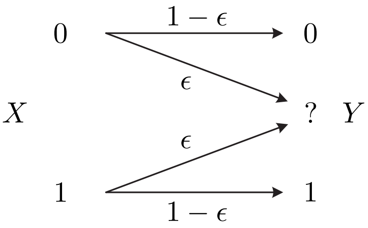
- Binary erasure channel (BEC):



- The codewords contain no-consecutive ones

The input-constrained BEC [Sabag,Permuter,Kashyap 16]

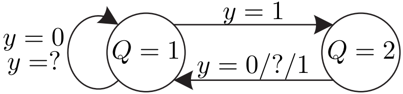
- Binary erasure channel (BEC):



- The codewords contain no-consecutive ones
- The channel state is $S_{i-1} = X_{i-1}$.

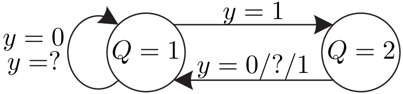
Solving BEC with no consecutive '1'

Q-graph

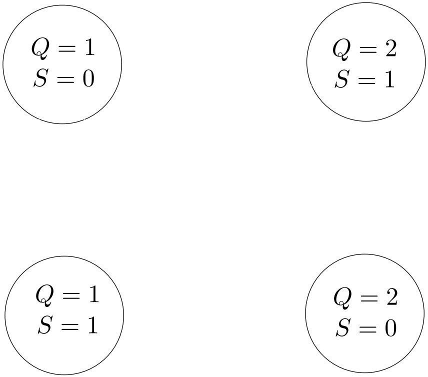


Solving BEC with no consecutive '1'

Q-graph

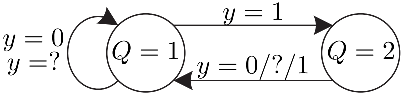


$(S,Q)-$graph

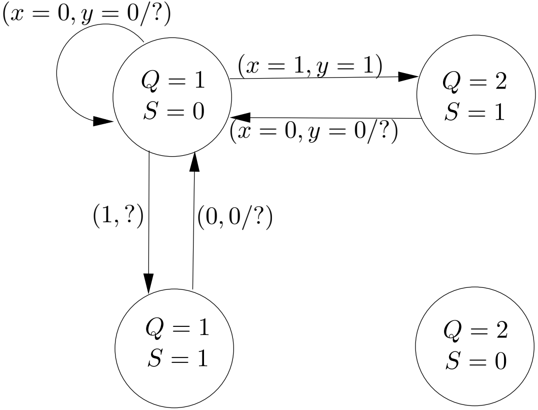


Solving BEC with no consecutive '1'

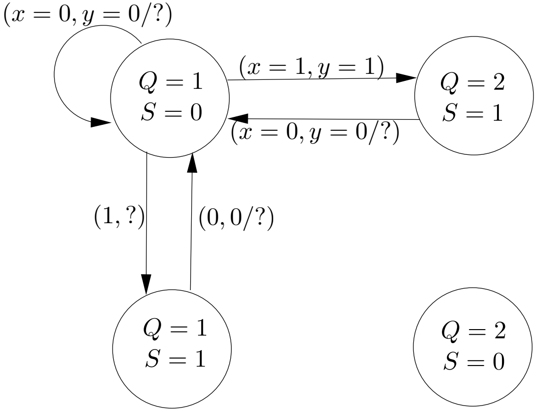
Q-graph



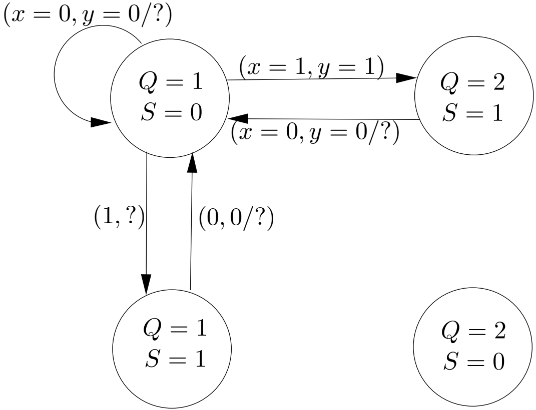
$(S,Q)-$graph



Solving BEC with no consecutive '1'

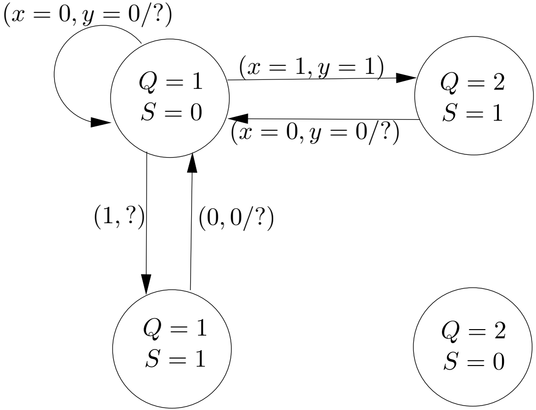


Solving BEC with no consecutive '1'



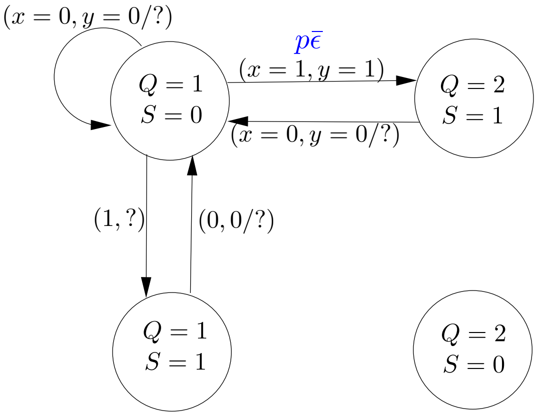
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$

Solving BEC with no consecutive '1'



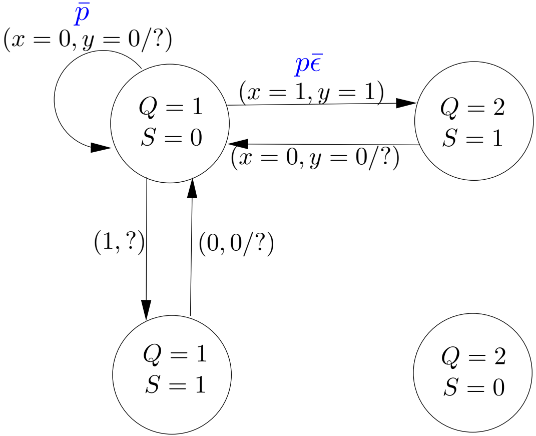
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$
$$P(x=1|s=0,q=1) = p$$

Solving BEC with no consecutive '1'



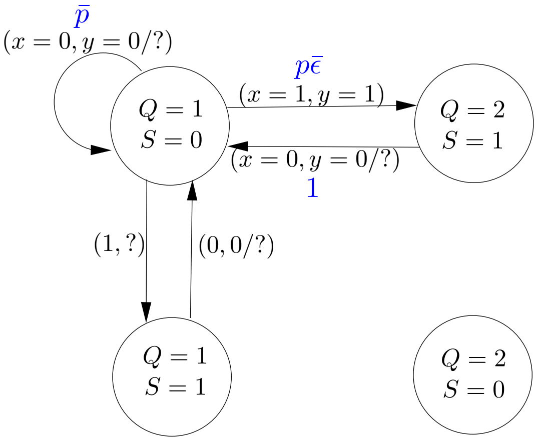
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$
$$P(x=1|s=0,q=1) = p$$

Solving BEC with no consecutive '1'



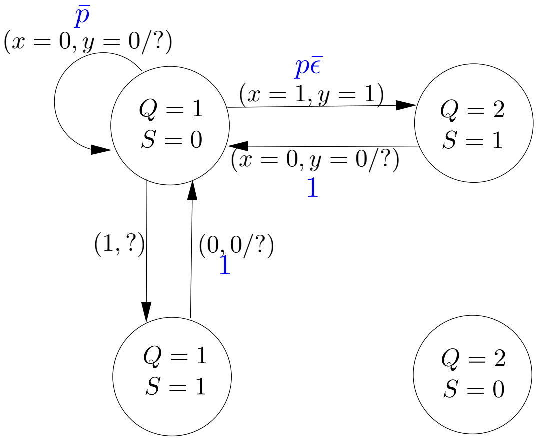
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$
$$P(x=1|s=0,q=1) = p$$

Solving BEC with no consecutive '1'



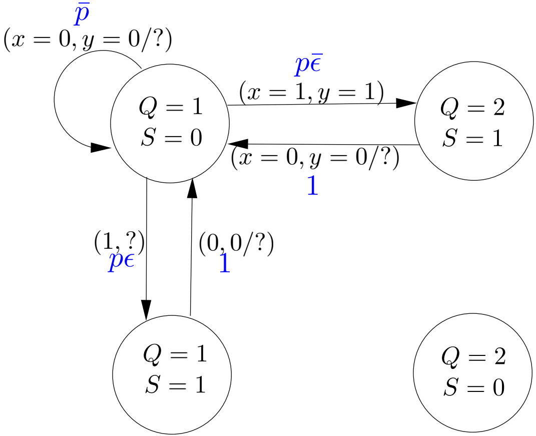
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$
$$P(x=1|s=0,q=1) = p$$

Solving BEC with no consecutive ’1’



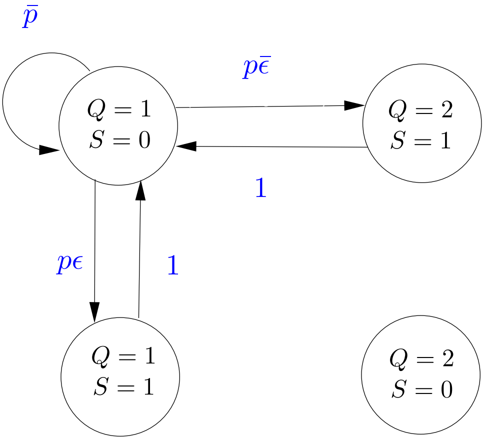
$$P(x=1|s=1,q=1)=P(x=1|s=1,q=2)=0$$
$$P(x=1|s=0,q=1)=p$$

Solving BEC with no consecutive '1'



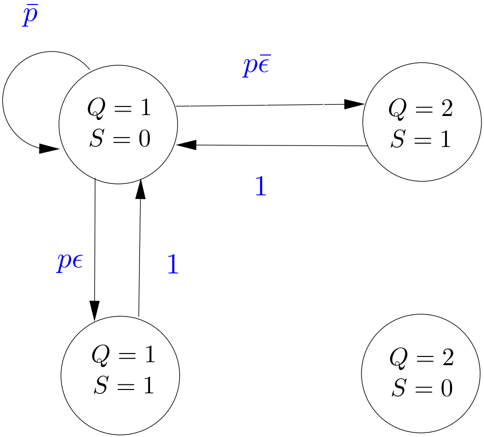
$$P(x=1|s=1,q=1) = P(x=1|s=1,q=2) = 0$$
$$P(x=1|s=0,q=1) = p$$

Solving BEC with no consecutive ’1’



$$P(x=1|s=1,q=1)=P(x=1|s=1,q=2)=0$$
$$P(x=1|s=0,q=1)=p$$

Solving BEC with no consecutive '1'



$$[\pi_{0,1}, \pi_{1,1}, \pi_{1,2}] = \left[\frac{1}{1+p}, \frac{p\epsilon}{1+p}, \frac{p\bar{\epsilon}}{1+p}\right]$$

Final step in solving BEC with no consecutive '1'

Evaluate

$$I(X, S; Y|Q)$$

at $\pi(s,q)p(x|s,q)p(y|x,s)$:

$$C_{fb} \leq \max_p \frac{H_2(p)}{p + \frac{1}{1-\epsilon}}.$$

Sketch Proof

$$
\begin{aligned}
C_{fb} &= \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} I(X^n \to Y^n) \\
&\triangleq \max_{P(x_i|x^{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1}) \\
&= \max_{P(x_i|s_{i-1},y^{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Y^{i-1}) \\
&\leq \max_{P(x_i|s_{i-1},q_{i-1})} \frac{1}{n} \sum_{i=1}^{n} I(X_i, S_{i-1}; Y_i | Q_{i-1}) \\
&= \max_{P(x|s,q)} I(X, S; Y | Q)
\end{aligned}
$$

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$

Sufficient condition

- The channel state estimation:

$$
\begin{aligned}
p(s_t|y^t) &= \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})} \\
&= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.
\end{aligned}
$$

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0,1]$.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0, 1]$.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0, 1]$.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})} = \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0, 1]$.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})} = \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

  This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0,1]$.

- An input, $p(x|s,q)$, is *BCJR-invariant* if

$$\pi(S = s|Q = \Phi(q, y)) \qquad \forall (y, q) \in \mathcal{Y} \times \mathcal{Q},$$

  where $\Phi(q, y)$ is the Q-graph mapping.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

  This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0,1]$.

- An input, $p(x|s,q)$, is *BCJR-invariant* if

$$\pi(S = s|Q = \Phi(q,y)) = B_s(\pi_{S|Q=q}, y) \;\; \forall (y,q) \in \mathcal{Y} \times \mathcal{Q},$$

  where $\Phi(q,y)$ is the Q-graph mapping.

Sufficient condition

- The channel state estimation:

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

  This mapping is denoted by $\mathrm{B}_s : \mathcal{P}(\mathcal{S}) \times \mathcal{Y} \to [0,1]$.

- An input, $p(x|s,q)$, is *BCJR-invariant* if

$$\pi(S = s|Q = \Phi(q,y)) = B_s(\pi_{S|Q=q}, y) \;\; \forall (y,q) \in \mathcal{Y} \times \mathcal{Q},$$

  where $\Phi(q,y)$ is the Q-graph mapping.

- $p(x|s,q)$ *is BCJR-invariant* $\Longrightarrow \; Y_i - Q_{i-1} - Y^{i-1} \;\; \forall i$

Final step in solving BEC with no consecutive '1'

Evaluate

$$I(X,S;Y|Q)$$

at $\pi(s,q)p(x|s,q)p(y|x,s)$:

$$C_{fb} \leq \max_{p} \frac{H_2(p)}{p+\frac{1}{1-\epsilon}}.$$

Final step in solving BEC with no consecutive '1'

Evaluate

$$I(X,S;Y|Q)$$

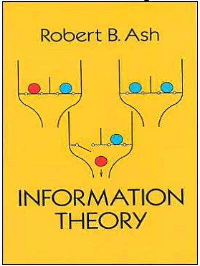
at $\pi(s,q)p(x|s,q)p(y|x,s)$:

$$C_{fb} \leq \max_p \frac{H_2(p)}{p+\frac{1}{1-\epsilon}}.$$

We show that maximizing input is BCJR, hence

$$C_{fb} = \max_p \frac{H_2(p)}{p+\frac{1}{1-\epsilon}}.$$

Trapdoor Channel [Blackwell61]



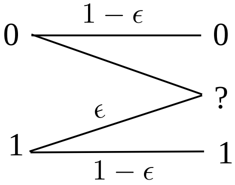
Ising Channel [Berger90]

$$y_i = \begin{cases} x_i, & \text{with prob. } \frac{1}{2} \\ x_{i-1}, & \text{with prob. } \frac{1}{2} \end{cases}$$

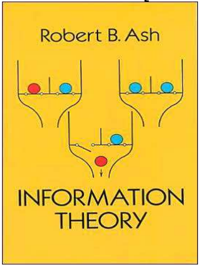
Dicode Erasure Channel [Pfister08]

$$y_i = \begin{cases} x_i - x_{i-1}, & \text{with prob. } 1-\epsilon \\ ?, & \text{with prob. } \epsilon \end{cases}$$

Erasure Channel with no repeated 1's



Trapdoor Channel [Blackwell61]



$$C_{fb} = \log \phi, \ \ \phi = \frac{\sqrt{5}+1}{2}$$

Ising Channel [Berger90]

$$y_i = \begin{cases} x_i, & \text{with prob. } \frac{1}{2} \\ x_{i-1}, & \text{with prob. } \frac{1}{2} \end{cases}$$

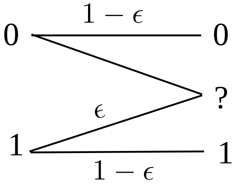
$$C_{fb} = \max_p \frac{2H_2(p)}{3+p} \approx 0.575$$

Dicode Erasure Channel [Pfister08]

$$y_i = \begin{cases} x_i - x_{i-1}, & \text{with prob. } 1-\epsilon \\ ?, & \text{with prob. } \epsilon \end{cases}$$

$$C_{fb} = \max_p (1-\epsilon)\frac{p+\epsilon H_2(p)}{\epsilon+(1-\epsilon)p}$$

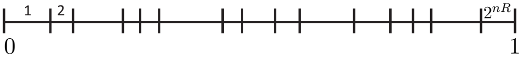
Erasure Channel with no repeated 1's



$$C_{fb} = \max_p \frac{H_2(p)}{p+\frac{1}{1-\epsilon}}$$

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:



Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:

2. The encoder matches inputs according to $p^*(x)$

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:

2. The encoder matches inputs according to $p^*(x)$

3. The decoder declares $\hat{m} = \max_m p(m|y^n)$

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:

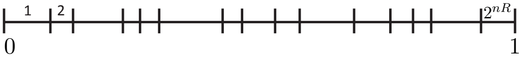
2. The encoder matches inputs according to $p^*(x)$

3. The decoder declares $\hat{m} = \max_m p(m|y^n)$

- For memoryless channels:
  - BSC (Horstein 63), AWGN channel (Schalkwijk, Kailath 66).

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:



2. The encoder matches inputs according to $p^*(x)$

3. The decoder declares $\hat{m} = \max_m p(m|y^n)$

- For memoryless channels:
  - BSC (Horstein 63), AWGN channel (Schalkwijk, Kailath 66).
  - For all memoryless channels (Shayevitz, Feder 11, 16).

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:

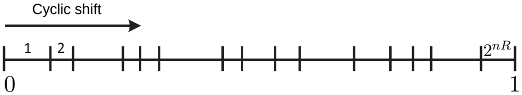
2. The encoder matches inputs according to $p^*(x)$

3. The decoder declares $\hat{m} = \max_m p(m|y^n)$

- For memoryless channels:
  - BSC (Horstein 63), AWGN channel (Schalkwijk, Kailath 66).
  - For all memoryless channels (Shayevitz, Feder 11, 16).
  - New and simple proof, (Li, El Gamal 15).

Matching schemes for memoryless channels

1. The posterior intervals $p(m|y^t)$:



2. The encoder matches inputs according to $p^*(x)$

3. The decoder declares $\hat{m} = \max_m p(m|y^n)$

- For memoryless channels:
  - BSC (Horstein 63), AWGN channel (Schalkwijk, Kailath 66).
  - For all memoryless channels (Shayevitz, Feder 11, 16).
  - New and simple proof, (Li, El Gamal 15).
    - random cyclic shift

Matching schemes for unifilar channels

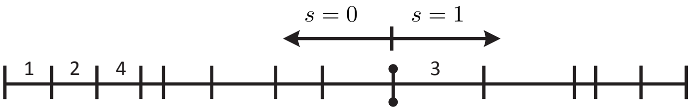
- y^{t-1} is public

Matching schemes for unifilar channels

- y^{t-1} is public
- for each message compute $s_{t-1}(m, y^{t-1})$

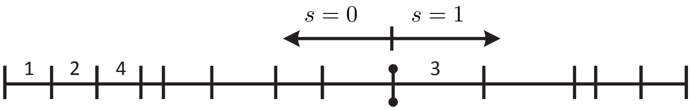
Matching schemes for unifilar channels

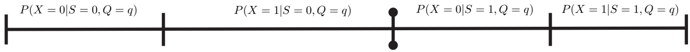
- y^{t-1} is public
- for each message compute $s_{t-1}(m, y^{t-1})$



Matching schemes for unifilar channels

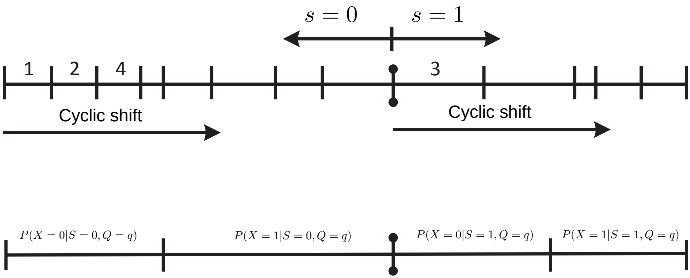
- y^{t-1} is public
- for each message compute $s_{t-1}(m, y^{t-1})$





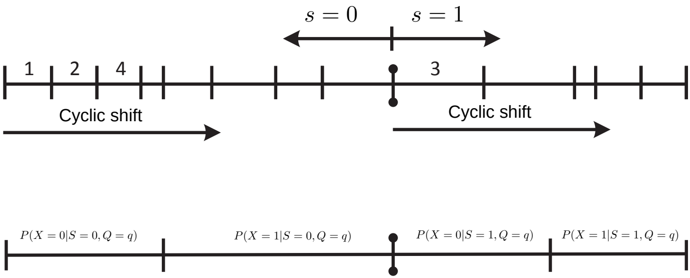
Matching schemes for unifilar channels

- y^{t-1} is public
- for each message compute $s_{t-1}(m, y^{t-1})$



Matching schemes for unifilar channels

- y^{t-1} is public
- for each message compute $s_{t-1}(m, y^{t-1})$



- new message-splitting idea

Intuition for posterior matching for unifilar channels

- The posterior principle,

$$p(m|y^t) = p(m|y^{t-1})\frac{p(y_t|m, y^{t-1})}{p(y_t|y^{t-1})}$$

Intuition for posterior matching for unifilar channels

- The posterior principle,

$$
\begin{aligned}
p(m|y^t) &= p(m|y^{t-1})\frac{p(y_t|m, y^{t-1})}{p(y_t|y^{t-1})} \\
&= p(m|y^{t-1})\frac{p(y_t|x_t(m, y^{t-1}), s_{t-1}(m, y^{t-1}))}{p(y_t|q_{t-1})}
\end{aligned}
$$

Intuition for posterior matching for unifilar channels

- The posterior principle,

$$p(m|y^t) = p(m|y^{t-1})\frac{p(y_t|m,y^{t-1})}{p(y_t|y^{t-1})}$$
$$= p(m|y^{t-1})\frac{p(y_t|x_t(m,y^{t-1}),s_{t-1}(m,y^{t-1}))}{p(y_t|q_{t-1})}$$

- The gain is

$$E\left[\log\frac{p(Y_t|X_t,S_{t-1})}{p(Y_t|Q_{t-1})}\right] = E\left[\log\frac{p(Y_t|X_t,S_{t-1},Q_{t-1})}{p(Y_t|Q_{t-1})}\right]$$
$$= I(X_t,S_{t-1};Y_t|Q_{t-1})$$

Intuition for posterior matching for unifilar channels

- The posterior principle,

$$p(m|y^t) = p(m|y^{t-1})\frac{p(y_t|m, y^{t-1})}{p(y_t|y^{t-1})}$$
$$= p(m|y^{t-1})\frac{p(y_t|x_t(m, y^{t-1}), s_{t-1}(m, y^{t-1}))}{p(y_t|q_{t-1})}$$

- The gain is

$$E\left[\log \frac{p(Y_t|X_t, S_{t-1})}{p(Y_t|Q_{t-1})}\right] = E\left[\log \frac{p(Y_t|X_t, S_{t-1}, Q_{t-1})}{p(Y_t|Q_{t-1})}\right]$$
$$= I(X_t, S_{t-1}; Y_t|Q_{t-1})$$

[Feedback capacity and coding for the BIBO channel with a no-repeated-ones input constraint, Sabag/P/Kashyap, on Arxiv]

Summary

- Introduced the idea of structured auxiliary r.v.

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \le \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel
- Tight for all channels if cardinality is unbounded

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel
- Tight for all channels if cardinality is unbounded
- Sufficient condition on finite Q

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel
- Tight for all channels if cardinality is unbounded
- Sufficient condition on finite Q
- Coding based on posterior matching scheme

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel
- Tight for all channels if cardinality is unbounded
- Sufficient condition on finite Q
- Coding based on posterior matching scheme

Ongoing work

- Algorithm for computing the bound for large Q.
- Derive cardinality bound on Q.

Summary

- Introduced the idea of structured auxiliary r.v.
- Plays important role in feedback capacity of unifilar channel

$$C_{fb} \leq \sup_{p(x|s,q)} I(X,S;Y|Q). \quad \forall Q\text{-graph}$$

- Tight for all known channel
- Tight for all channels if cardinality is unbounded
- Sufficient condition on finite Q
- Coding based on posterior matching scheme

Ongoing work

- Algorithm for computing the bound for large Q.
- Derive cardinality bound on Q.

Thank you very much!

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})} = \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})} = \frac{\sum_{x_t,s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t,s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.
- If $|\mathcal{Z}| \leq \infty$ and $P_{X|S,Z}$ satisfies the BCJR mapping, we call it *BCJR-invariant*.

Sufficient condition

- The channel state estimation (DP state):

$$p(s_t|y^t) = \frac{p(s_t, y_t|y^{t-1})}{p(y_t|y^{t-1})}$$
$$= \frac{\sum_{x_t, s_{t-1}} p(s_t, y_t, x_t, s_{t-1}|y^{t-1})}{\sum_{x_t, s_{t-1}} p(y_t, x_t, s_{t-1}|y^{t-1})}.$$

- This mapping is denoted by BCJR : $\mathcal{Z} \times \mathcal{Y} \to \mathcal{Z}$.
- If $|\mathcal{Z}| \leq \infty$ and $P_{X|S,Z}$ satisfies the BCJR mapping, we call it *BCJR-invariant*.

Theorem (Lower bound)

The feedback capacity satisfies

$$C_{fb} \geq I(X, S; Y|Q),$$

for all BCJR-invariant inputs.

Upper bound with sufficient condition

Theorem [Sabag/P./Pfiser16]

The feedback capacity of a unifilar FSC is bounded by

$$C_{fb} \leq \max_{p(x|s,q)} I(X,S;Y|Q), \quad \forall Q\text{-graph}$$

and if $p^*(x|s,q)$ is BCJR-invariant input, equality holds.