[Synchronization + Constrained coding (inner code)]

$\Sigma = \{0,1,2,3\}$ (A,T,G,C)

$\ell :$ constraint length

$\mathcal{Q} \subset \Sigma^\ell :$ valid state set

  (ex) $\mathcal{Q}_{\ell,k}^{\mathrm{R}} :$ maximum run-length k

      $\mathcal{Q}_{\ell,\varepsilon}^{\mathrm{B}} :$ local ε-balanced

      $\mathcal{Q}_\ell^{\mathrm{M}} :$ motif (?)

      $\mathcal{Q}_\ell^{\mathrm{S}} :$ synchronization (?)

      $\mathcal{Q} = \mathcal{Q}_{\ell,k}^{\mathrm{R}} \cap \mathcal{Q}_{\ell,\varepsilon}^{\mathrm{B}} \cap \mathcal{Q}_\ell^{\mathrm{M}} \cap \mathcal{Q}_\ell^{\mathrm{S}} \cap \cdots$

state: $\boldsymbol{q} = q_0 q_1 \ldots q_{\ell-2} q_{\ell-1} \in \mathcal{Q}$

left-shift: $\overleftarrow{\boldsymbol{q}}\, a = q_1 q_2 \ldots q_{\ell-1} a \in \Sigma^\ell$

state transition: $t(\boldsymbol{q}, a) = \{\boldsymbol{q}' = \overleftarrow{\boldsymbol{q}}\, a \,|\, \boldsymbol{q}' \in \mathcal{Q}\}$

              $|t(\boldsymbol{q}, a)| \in \{0,1\}$

edge label set: $\mathcal{L}(\boldsymbol{q}) = \{a \,|\, t(\boldsymbol{q}, a) \neq \phi\}$

encoding: $f_{\mathrm{e}} : \mathcal{Q} \times \mathbb{B}^2 \to \Sigma$

$\mathcal{L}(\boldsymbol{q})$	00	01	10	11
$\{0,1,2,3\}$	0	1	2	3
$\{0,1,2\}$	0	1	2	2
$\{0,1,3\}$	0	1	3	3
$\{0,2,3\}$	0	0	2	3
$\{1,2,3\}$	1	1	2	3
$\{q_0, q_1\}$	q_0	q_0	q_1	q_1
$\{q\}$	q	q	q	q
ϕ	(invalidate $\boldsymbol{q}$)			

FBA: $s_i = (\boldsymbol{q}_i, d_i)$

      $x_i = (x_i^0, x_i^1) \in \mathbb{B}^2$

outer code: $\mathcal{C}_0 :$ left-bit

           $\mathcal{C}_1 :$ right-bit
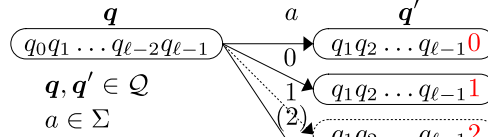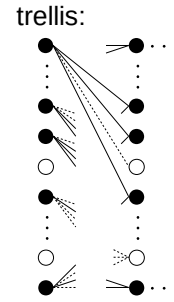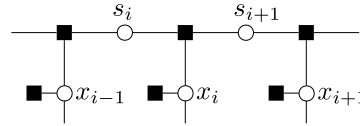
           (rate) $R(\mathcal{C}_0) > R(\mathcal{C}_1)$

evaluation: ?

  constrained channel capacity

  IDS channel SIR? (<=> finite length?)

  FBA soft output: mutual info?

[Synchronization + Constrained coding (inner code)]

$\Sigma = \{0, 1, 2, 3\}$ (=A,T,G,C)

k :   maximum run-length

δ :   balance margin $\left| w_{01}(\boldsymbol{x}_0^{i-1}) - w_{23}(\boldsymbol{x}_0^{i-1}) \right| \le \delta$ $\rightarrow$ local balance: $b = (2\delta + \ell')/\ell$

$\left| 2w_{01}(\boldsymbol{x}_0^{i-1}) - i \right| \le \delta$ $(\forall i \in [n])$

$$\ell' = \begin{cases} (\ell - 2\delta)/2 & (\ell : \text{even}) \\ (\ell - 2\delta - 1)/2 & (\ell : \text{odd}) \end{cases}$$

$w_{01}(\boldsymbol{x}_0^{i-1}) - w_{23}(\boldsymbol{x}_0^{i-1})$

$$b = \begin{cases} \frac{1}{2} + \frac{\delta}{\ell} & (\ell : \text{even}) \\ \frac{1}{2} + \frac{\delta}{\ell} - \frac{1}{2\ell} & (\ell : \text{odd}) \end{cases}$$

$$\varepsilon = \begin{cases} \frac{\delta}{\ell} & (\ell : \text{even}) \\ \frac{\delta}{\ell} - \frac{1}{2\ell} & (\ell : \text{odd}) \end{cases}$$
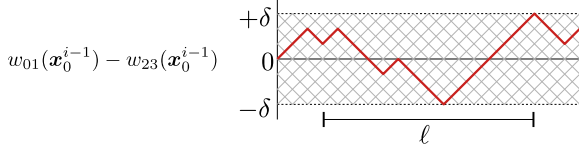
[IDS channel]

p_{i} :   insertion
p_{d} :   deletion
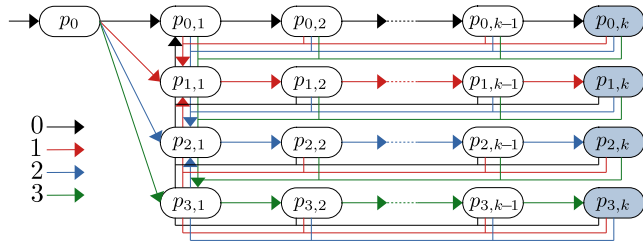$p_{\mathrm{s}}(y|x)$:   asymmetric error
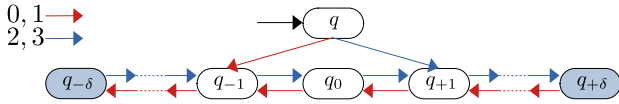$d_{\min} < 0$: drift min
$d_{\max} > 0$: drift max
$\mathcal{D} = \{d \in \mathbb{Z} | d_{\min} \le d \le d_{\max}\}$

[state transition]

run-length: $t_{\mathrm{R}}(p, a) : \mathcal{P} \times \Sigma \to \mathcal{P} \cup \{\bot\}$

balance: $t_{\mathrm{B}}(q, a) : \mathcal{Q} \times \Sigma \to \mathcal{Q} \cup \{\bot\}$

run-length + balance:
$t(p, q, a) : \mathcal{P} \times \mathcal{Q} \times \Sigma \to (\mathcal{P} \times \mathcal{Q}) \cup \{\bot\}$
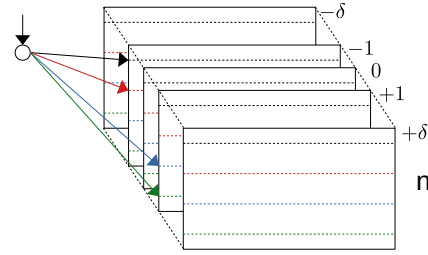
number of states:
$4k(2\delta + 1) + 1$

$$t(p, q, a) = \begin{cases} (p', q') & (t_{\mathrm{R}}(p, a) = p', t_{\mathrm{B}}(q, a) = q') \\ \bot & (\text{otherwise}) \end{cases}$$

[information mapping]

$\mathcal{L}(p, q) = \{a \mid t(p, q, a) \ne \bot\}$
$f_{\mathrm{e}} : \mathcal{P} \times \mathcal{Q} \times \mathbb{B}^2 \to \Sigma$

$\mathcal{L}(p,q)$	00	01	10	11
$\{0,1,2,3\}$	0	1	2	3
$\{0,1,2\}$	0	1	2	2
$\{0,1,3\}$	0	1	3	3
$\{0,2,3\}$	0	0	2	3
$\{1,2,3\}$	1	1	2	3
$\{q_0, q_1\}$	q_0	q_0	q_1	q_1
$\{q\}$	q	q	q	q
ϕ	(delete (p,q))			

erasures:
corrected by outer code

[FBA]

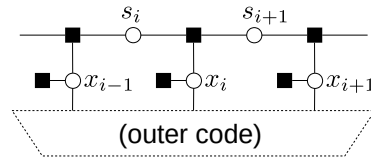$s_i = (p_i, q_i, d_i) \in \mathcal{P} \times \mathcal{Q} \times \mathcal{D}$
$x_i = (x_i^0, x_i^1) \in \mathbb{B}^2$

(outer code)

$p(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{d})$

$= p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{d}) p(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{d})$

$= p(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{d}) \prod_{i=0}^{n-1} p(\boldsymbol{y}_{i+d_i}^{i+d_{i+1}} | x_i, p_i, p_{i+1}, q_i, q_{i+1}, d_i, d_{i+1})$

$= p(p_0) p(q_0) p(d_0) \prod_{i=0}^{n-1} p(\boldsymbol{y}_{i+d_i}^{i+d_{i+1}} | x_i, p_i, p_{i+1}, q_i, q_{i+1}, d_i, d_{i+1}) p(p_{i+1}|p_i) p(q_{i+1}|q_i) p(d_{i+1}|d_i)$
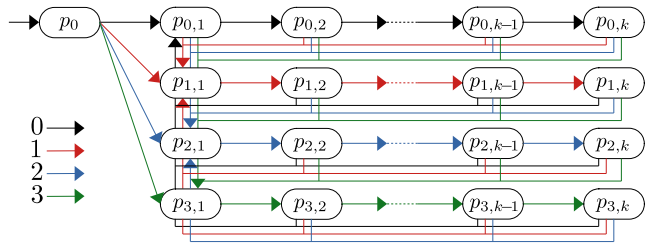
[evaluation]

* code rate
  - capacity of constrained channel
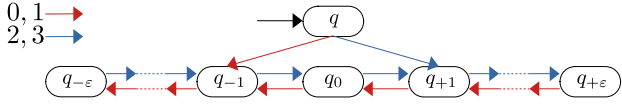 - encoding rate (erasure prob.)
* mutual information: I(channel input, FBA output)
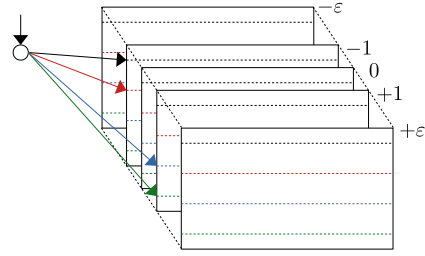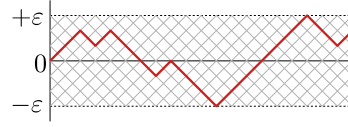 - comparison to marker, watermark ?

[run-length]

p_0

$p_{0,1}$ $p_{0,2}$ $p_{0,k-1}$ $p_{0,k}$

$p_{1,1}$ $p_{1,2}$ $p_{1,k-1}$ $p_{1,k}$

$p_{2,1}$ $p_{2,2}$ $p_{2,k-1}$ $p_{2,k}$

$p_{3,1}$ $p_{3,2}$ $p_{3,k-1}$ $p_{3,k}$

0
1
2
3

[run-length + balance]

$-\varepsilon$
-1
0
$+1$
$+\varepsilon$

number of states:
$4k(2\varepsilon + 1) + 1$

[balance]

$0, 1$
$2, 3$

q

$q_{-\varepsilon}$ q_{-1} q_0 q_{+1} $q_{+\varepsilon}$

$+\varepsilon$
0
$-\varepsilon$

info mapping: rate-1

00
01
10
11

00
01
1*

0*
1*

**

(invalidate)
(*: erasure)