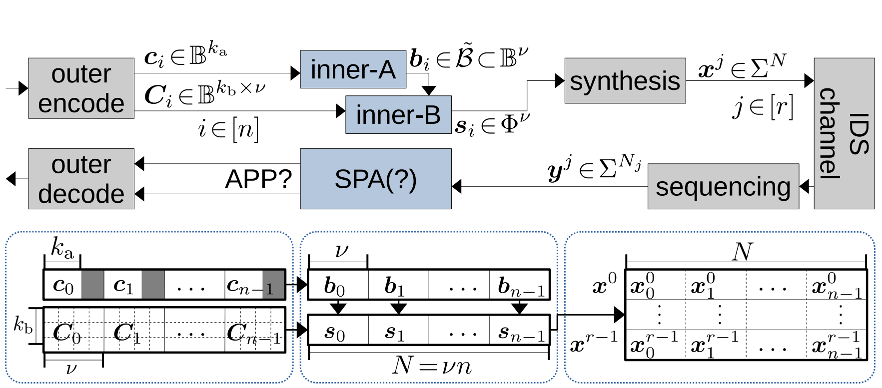


$\mathbb{B} = \{0,1\}, [n] = \{0,1,\dots,n-1\}$

Alphabet: $\Sigma = \{0,1,2,3\} \leftrightarrow \{\mathrm{A}, \mathrm{T}, \mathrm{G}, \mathrm{C}\}$
Code length: ν symbol (even)
Block length: $N = \nu n$ symbol
(block = n inner codewords)

Constraints (binary):
run length: RL-ρ
local balance: LB-(ℓ, δ) $\left|w(\boldsymbol{b}_i^{i+\ell-1}) - \frac{\ell}{2}\right| \le \delta$
indicator func: $\mathbb{1}_{\rho,\ell,\delta}(\boldsymbol{b})$
IDS channel:
error prob: $p_{\mathrm{i}}, p_{\mathrm{d}}, p_{\mathrm{s}}$
drift vector:
$\boldsymbol{d}^j = (d_0^j, \dots, d_N^j)$ (symbol level)
$\tilde{\boldsymbol{d}}^j = (\tilde{d}_0^j, \dots, \tilde{d}_n^j)$ (word level) $\tilde{d}_i^j = d_{i\nu}^j$

Binary constraint vector set [SITA2024]

$\mathcal{B} \subset \mathbb{B}^\nu \quad (|\mathcal{B}| = 2^{k_a})$

$\forall \boldsymbol{b} \in \mathcal{B}, w(\boldsymbol{b}) = \nu/2$	(LB: inside CW)
$\forall \boldsymbol{b} \in \mathcal{B}, \forall i \in [\nu]_{\mathrm{even}}, \left|\tilde{w}(\boldsymbol{b}_i^{\nu-1})\right| \le 2\delta$	(LB: CW boundary)
$\forall \boldsymbol{b} \in \mathcal{B}, \forall i \in [\nu]_{\mathrm{odd}}, \left|\tilde{w}(\boldsymbol{b}_i^{\nu-1})\right| \le 2\delta - 1$	(LB: CW boundary)
$\forall \boldsymbol{b} \in \mathcal{B}, \lambda(\boldsymbol{b}) \le \rho$	(RL: inside CW)
$\forall \boldsymbol{b} \in \mathcal{B}, \lambda(\boldsymbol{b}, \nu-1) \le \rho - 1$	(RL: CW boundary)
$\boldsymbol{b} \in \mathcal{B} \to \overline{\boldsymbol{b}} \in \mathcal{B}$	(re-balance)
$(01)^{\frac{\nu}{2}}, (10)^{\frac{\nu}{2}} \notin \mathcal{B}$	(erasure symbol)

$\tilde{\mathcal{B}} = \mathcal{B} \cup \{(01)^{\frac{\nu}{2}}, (10)^{\frac{\nu}{2}}\}$ (re-balance)

Composite symbol sets

$\Phi_0 \subseteq \left\{(\sigma_0, \sigma_1, \sigma_2, \sigma_3) \in \mathbb{Z}_{k+1}^4 \,\middle|\, \sigma_0 + \sigma_1 = k, \sigma_2 = \sigma_3 = 0\right\}$ AT

$\Phi_1 \subseteq \left\{(\sigma_0, \sigma_1, \sigma_2, \sigma_3) \in \mathbb{Z}_{k+1}^4 \,\middle|\, \sigma_0 = \sigma_1 = 0, \sigma_2 + \sigma_3 = k\right\}$ GC

$|\Phi_0| = |\Phi_1| = 2^{k_b} \le k+1 \qquad \Phi = \Phi_0 \cup \Phi_1$

Inner codebook

$\mathcal{C} = \left\{(c_0, \dots, c_{\nu-1}) \,\middle|\, c_j \in \Phi_{b_j}, (b_0, \dots, b_{\nu-1}) \in \mathcal{B}\right\}$

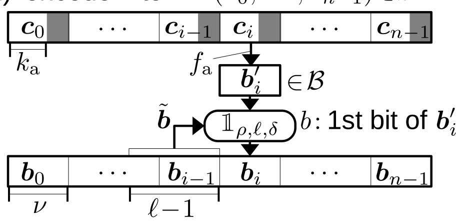
Mapping

$f_a : \mathbb{B}^{k_a} \to \mathcal{B}$ (bijection)
$f_i : \mathbb{B}^{k_b} \to \Phi_i \ (i \in \mathbb{B})$ (bijection)
$f_d : \Sigma \to \mathbb{B} \qquad f_d(x) = \mathbb{1}[x \in \{2,3\}]$

Encoding

input: $\boldsymbol{c} = (\boldsymbol{c}_0, \dots, \boldsymbol{c}_{n-1}) \quad \boldsymbol{c}_i \in \mathbb{B}^{k_a}$
$\boldsymbol{C} = (\boldsymbol{C}_0, \dots, \boldsymbol{C}_{n-1}) \quad \boldsymbol{C}_i \in \mathbb{B}^{k_b \times \nu}$
output: $\boldsymbol{s} = (\boldsymbol{s}_0, \dots, \boldsymbol{s}_{n-1}) \quad \boldsymbol{s}_i \in \Phi^\nu$

(1) encode $\boldsymbol{c}$ to $\boldsymbol{b} = (\boldsymbol{b}_0, \dots, \boldsymbol{b}_{n-1}) \in \tilde{\mathcal{B}}^n$



$$\boldsymbol{b}_i = \begin{cases} \boldsymbol{b}'_i & (\mathbb{1}_{\rho,\ell,\delta}(\tilde{\boldsymbol{b}}, \boldsymbol{b}'_i) = 1) & \text{(unchanged)} \\ \overline{\boldsymbol{b}}'_i & (\mathbb{1}_{\rho,\ell,\delta}(\tilde{\boldsymbol{b}}, \overline{\boldsymbol{b}}'_i) = 1) & \text{(invert)} \\ (\overline{b}b)^{\frac{\nu}{2}} & (\mathbb{1}_{\rho,\ell,\delta}(\tilde{\boldsymbol{b}}, (\overline{b}b)^{\frac{\nu}{2}}) = 1) & \text{(rebalance)} \\ (b\overline{b})^{\frac{\nu}{2}} & (\text{otherwise}) & \text{(rebalance)} \end{cases}$$

(2) encode $(\boldsymbol{C}_i, \boldsymbol{b}_i)$ to $\boldsymbol{s}_i$ $(i \in [n])$

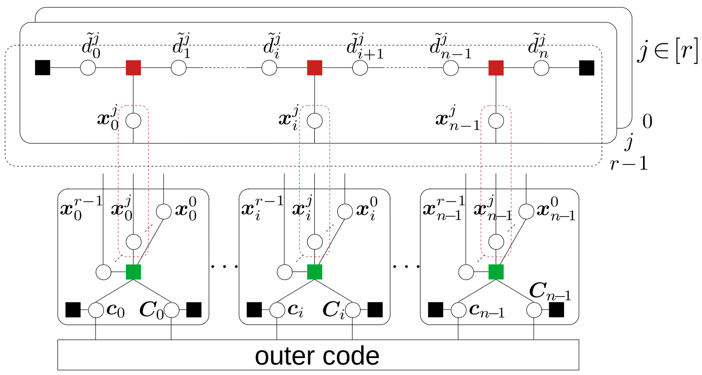
$s_{i,t} = f_{b_{i,t}}(\boldsymbol{C}_{i,t}^{\mathsf{T}}) \quad (t \in [\nu])$
$\boldsymbol{s}_i = (s_{i,0}, \dots, s_{i,\nu-1}) \in \Phi^\nu$
$\boldsymbol{b}_i = (b_{i,0}, \dots, b_{i,\nu-1}) \in \tilde{\mathcal{B}}$
$\boldsymbol{C}_{i,t}$: t th column of $\boldsymbol{C}_i$

(ex) $\nu = 6, |\mathcal{B}| = 16, k = 7, k_a = 4, k_b = 3$

```
B:0:001011  8:110100  Φ0:0:0700  Φ1: 0:0007
  1:001101  9:110010     1:1600      1:0016
  2:001110  A:110001     2:2500      2:0025
  3:010011  B:101100     3:3400      3:0034
  -:010101  -:101010     4:4300      4:0043
  4:010110  C:101001     5:5200      5:0052
  5:011001  D:100110     6:6100      6:0061
  6:011010  E:100101     7:7000      7:0070
  7:011100  F:100011       (Lee distance?)
```

```
ci :  0110(6)=011010
CiT : 011(3)               si : (0,3)=3400
      100(4)                    (1,4)=0043
      010(2)                    (1,2)=0025
      011(3)                    (0,3)=3400
      110(6)                    (1,6)=0061
      100(4)                    (0,4)=4300
```

Decoding



$$p\left(\boldsymbol{y}^0, \dots, \boldsymbol{y}^{r-1}, \boldsymbol{x}^0, \dots, \boldsymbol{x}^{r-1}, \tilde{\boldsymbol{d}}^0, \dots, \tilde{\boldsymbol{d}}^{r-1}, \boldsymbol{c}, \boldsymbol{C}\right)$$
$$= p\left(\boldsymbol{y}^0, \dots, \boldsymbol{y}^{r-1} \,\middle|\, \boldsymbol{x}^0, \dots, \boldsymbol{x}^{r-1}, \tilde{\boldsymbol{d}}^0, \dots, \tilde{\boldsymbol{d}}^{r-1}, \boldsymbol{c}, \boldsymbol{C}\right) \times$$
$$p\left(\tilde{\boldsymbol{d}}^0, \dots, \tilde{\boldsymbol{d}}^{r-1}\right) p\left(\boldsymbol{x}^0, \dots, \boldsymbol{x}^{r-1} \middle| \boldsymbol{c}, \boldsymbol{C}\right) p(\boldsymbol{c}, \boldsymbol{C})$$
$$= p(\boldsymbol{c})\, p(\boldsymbol{C}) \prod_{j=0}^{r-1} p\left(\boldsymbol{y}^j \middle| \boldsymbol{x}^j, \tilde{\boldsymbol{d}}^j\right) p\left(\tilde{\boldsymbol{d}}^j\right) p\left(\boldsymbol{x}^j \middle| \boldsymbol{c}, \boldsymbol{C}\right)$$
$$\simeq \left(\prod_{i=0}^{n-1} p(\boldsymbol{c}_i)\, p(\boldsymbol{C}_i)\right) \times$$
$$\prod_{j=0}^{r-1}\left(p(\tilde{d}_0^j) \prod_{i=0}^{n-1} \underline{p\left([\boldsymbol{y}^j]_{i\nu+\tilde{d}_i^j}^{i\nu+\nu-1+\tilde{d}_{i+1}^j} \,\middle|\, \boldsymbol{x}_i^j, \tilde{d}_i^j, \tilde{d}_{i+1}^j,\right) p\left(\tilde{d}_{i+1}^j \middle| \tilde{d}_i^j\right)}\ \underline{p\left(\boldsymbol{x}_i^j \middle| \boldsymbol{c}_i, \boldsymbol{C}_i\right)}\right)$$

$$\underline{\prod_{j=0}^{r-1} p\left(\boldsymbol{x}_i^j \middle| \boldsymbol{c}_i, \boldsymbol{C}_i\right)} = \prod_{j=0}^{r-1} \underbrace{p\left(\boldsymbol{b}_i^j \middle| \boldsymbol{c}_i\right)}_{\text{(sim)}} p\left(\boldsymbol{x}_i^j \middle| \boldsymbol{b}_i^j, \boldsymbol{C}_i\right) = \prod_{j=0}^{r-1} p\left(\boldsymbol{b}_i^j \middle| \boldsymbol{c}_i\right) \prod_{t=0}^{\nu-1} p\left(x_{i,t}^j \middle| b_{i,t}^j, \boldsymbol{C}_{i,t}\right)$$

$$\boldsymbol{b}_i^j = \left(f_d(x_{i,0}^j), \dots, f_d(x_{i,\nu-1}^j)\right)$$
$$p\left(x_{i,t}^j \,\middle|\, b_{i,t}^j, \boldsymbol{C}_{i,t}\right) = \sigma_{x_{i,t}^j}/k \qquad \left(x_{i,t}^j \in \Sigma, b_{i,t}^j \in \mathbb{B}, \boldsymbol{C}_{i,t}^{\mathsf{T}} \in \mathbb{B}^{k_b}\right)$$
$$(\sigma_0, \sigma_1, \sigma_2, \sigma_3) = f_{b_{i,t}^j}(\boldsymbol{C}_{i,t}^{\mathsf{T}})$$

non-uniform k