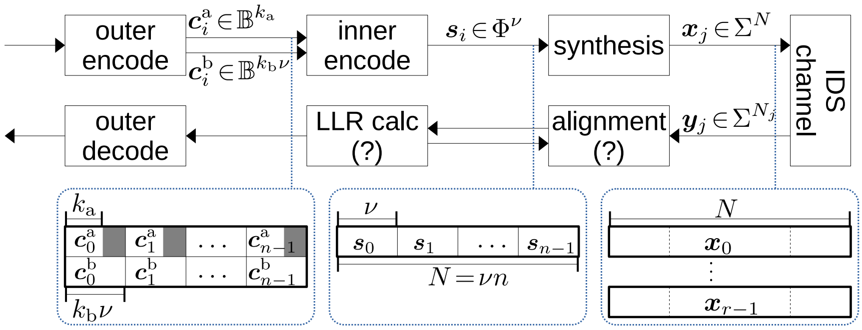
8 10
(A,T,G,C)
(0,1,2,3)



$i \in [n]$
$j \in [r]$
$N = \nu n$

Code length: ν (even)
Binary constraint vector set: $\mathcal{B} \subset \mathbb{B}^{\nu}$ $\quad (|\mathcal{B}| \geq 2^{k_{\mathrm{a}}})$
Composite symbol set:

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

$|\Phi_0| = |\Phi_1| = 2^{k_{\mathrm{b}}} \leq k+1 \qquad \Phi = \Phi_0 \cup \Phi_1$

Inner codebook:

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

(example) $\nu = 6, |\mathcal{B}| = 16, k = 7, k_{\mathrm{a}} = 4, k_{\mathrm{b}} = 3,$

$\mathcal{B}$:	001011	110100	Φ_0 :	0700	Φ_1 :	0007
	001101	110010		1600		0016
	001110	110001		2500		0025
	010011	101100		3400		0034
	010101	101010		4300		0043
	010110	101001		5200		0052
	011001	100110		6100		0061
	011010	100101		7000		0070
	011100	100011				

Lee dist?

Encode:

$\left.\begin{array}{l} f_{\mathrm{e}}^{\mathrm{a}} : \mathbb{B}^{k_{\mathrm{a}}} \to \mathcal{B} \\ f_{\mathrm{e}}^{i} : \mathbb{B}^{k_{\mathrm{b}}} \to \Phi_i \ (i \in \mathbb{B}) \end{array}\right]$ bijection

$f_{\mathrm{e}}(\boldsymbol{c}_i^{\mathrm{a}}, \boldsymbol{c}_i^{\mathrm{b}}) = \boldsymbol{s}_i = (s_{i,0}, \ldots, s_{i,\nu-1}) : \mathbb{B}^{k_{\mathrm{a}}} \times \mathbb{B}^{k_{\mathrm{b}}\nu} \to \Phi^{\nu}$

$s_{i,j} = f_{\mathrm{e}}^{b_i}(\boldsymbol{c}_{i,j}^{\mathrm{b}})$

$(b_0, \ldots, b_{\nu-1}) = f_{\mathrm{e}}^{\mathrm{a}}(\boldsymbol{c}_i^{\mathrm{a}})$

$\boldsymbol{c}_{i,j}^{\mathrm{b}} = [\boldsymbol{c}_i^{\mathrm{b}}]_{j\nu}^{j\nu+\nu-1}$

flip (RL, LB)

Decode:

```
          001011 001101 001110 ...110100
base      AAGAGG AAGGAG AAGGGA    GGAGAA
          AAGAGC AAGGAC AAGGGT    GGAGAT
          AAGACG AAGGTG AAGGCA    GGAGTA   64
          ...    ...    ...       ...
          TTCTCC TTCCTC TTCCCT    CCTCTT
composite
                                           ??
```

multi-base composite (Manchester coding)

$\nu = 1:$ (A,C,G,T)
$\nu = 2:$ (AA,AC,AG,AT, CA,CC,CG,CT, GA,GC,GG,GT, TA,TC,TG,TT)
$\nu = 3:$ (AAA, AAC, ..., TTT)

```
          001011 001101 001110 ...110100
base      AAGAGG AAGGAG AAGGGA    GGAGAA
          AAGAGC AAGGAC AAGGGT    GGAGAC
          AAGACG AAGGTG AAGGCA    GGAGCA   64
          ...    ...    ...       ...
          CCTCTT CCTTCT CCTTTC
composite
```