of $\mathcal{G}(X)$ are exactly the graphs

$$\mathcal{G}(X)_j = (V(X)_j, E(X)_j) \text{ for } 1 \leq j \leq n-1$$

with

$$V(X)_j = \{N_1...N_j, N_{j+1}...N_n, N_1...N_{n-j}, N_{n-j+1}...N_n : N_1N_2N_3...N_n \in X\}$$

and

$$E(X)_j = \{[N_1...N_j, N_{j+1}...N_n], [N_1...N_{n-j}, N_{n-j+1}...N_n] : N_1N_2N_3...N_n \in X\}.$$

These components do not have to be connected as we can see in Figure 3. However, quite often they are. In fact, $\mathcal{G}(X)_j$ consists exactly of the nodes (and their corresponding edges) that interpret the elements of X in two ways: as a pair of a j-nucleotide and a $(n-j)$-nucleotide and as a pair of a $(n-j)$-nucleotide and a j-nucleotide. Note that by symmetry we have $\mathcal{G}(X)_j = \mathcal{G}(X)_{n-j}$ for all $j < n-1$. For instance, in Figure 3 above the two components of the graph associated to the tetranucleotide code are $\mathcal{G}(X)_1$ $(= \mathcal{G}(X)_3)$ and $\mathcal{G}(X)_2$. The next observation is obvious.

Lemma 1. *Let X be an n-nucleotide code for some $n \in \mathbb{N}$. Then the following statements hold:*

(1) *If n is odd, then $\mathcal{G}(X)$ is a bipartite graph. In particular, all its components $\mathcal{G}(X)_j$ are bipartite.*

(2) *If n is even, then all components of $\mathcal{G}(X)$ are bipartite except for maybe $\mathcal{G}(X)_{\frac{n}{2}}$.*

We now start to investigate our desired objects, namely n-nucleotide circular codes.

Si l'ensemble des code est infore, on jète,

Definition 2. *Let $X \subseteq B^n$ be a code. We say that X is a circular code if for any concatenation $c_1 \cdots c_m$ of n-nucleotide words from X there is only one partition into n-nucleotide words from X when read on a circle.*

The first observation shows that for circular codes the associated graph is already simple. Recall from graph theory (Clark and Holton, 1991) that an oriented graph is *simple* if it does not contain loops, i.e. edges between a node and itself, and does not have multiple edges with the same orientation between two nodes. Note that the orientation for the multiple edges do play a role, i.e. for a simple oriented graph $\mathcal{G}$, we can still have $[x, y] \in E(\mathcal{G})$ and $[y, x] \in E(\mathcal{G})$ which means that there is a cycle (circle) of length 2. However, for circular codes this structure is also excluded. Recall that a *cycle* in $\mathcal{G}$ is an oriented closed path in

cycle (circle) length 2 excluded $X = \{AAAA, AAAC, \dots, TTTT\}$

éliminé ⟹ AT ⇆ AT <=> 208 ATAT, AAAA,