PolyHalt DNA Encoding Analysis Tables

Fig. 1: No. of codewords for $X = 4$ conflict-free and $Y = 3$ reverse-complement-free DNA codes with varying N and E

Codeword Length (N)	Edit Distance (E)								
	1	2	3	4	5	6	7	8	9
2	3	1							
4	23	15	3	1					
6	75	66	14	6	2	1			
8	80	74	50	20	7	3	1	1	
10	84	77	70	53	24	9	3	2	1

Fig. 2: Information Density for $X = 4$ conflict-free and $Y = 3$ reverse-complement-free DNA codes with varying N and E

Codeword Length (N)	Edit Distance (E)								
	1	2	3	4	5	6	7	8	9
2	0.79								
4	1.13	0.97	0.39						
6	1.04	1.00	0.63	0.43	0.17				
8	0.79	0.77	0.71	0.54	0.35	0.20			
10	0.64	0.62	0.61	0.57	0.46	0.31	0.16	0.10	

Fig. 3: No. of codewords for $X = 2$ conflict-free and $Y = 5$ reverse-complement-free DNA codes with varying N and E

Codeword Length (N)	Edit Distance (E)								
	1	2	3	4	5	6	7	8	9
2	3	1							
4	23	15	3	1					
6	80	77	17	7	2	1			
8	82	72	56	24	9	3	1	1	
10	85	82	78	57	28	9	4	2	1

Fig. 4: Information Density for $X = 2$ conflict-free and $Y = 5$ reverse-complement-free DNA codes with varying N and E

Codeword Length (N)	Edit Distance (E)								
	1	2	3	4	5	6	7	8	9
2	0.79								
4	1.13	0.97	0.39						
6	1.05	1.04	0.68	0.47	0.17				
8	0.79	0.77	0.72	0.57	0.40	0.20			
10	0.64	0.63	0.62	0.58	0.48	0.32	0.20	0.10	

A. Variable Parameters

- **Codeword Length (N):** Systematically varying the length of the codewords enables us to observe how the encoding method adapts to different information sequencelengths. Various N values represent distinct scenarios of encoding information into DNA.
- **Edit Distance (E):** The edit distance between codewords is a critical factor affecting error correction. By adjust- ing E, we examine how the encoding method handles variations and errors in the information sequences.

B. Objective

Our goal is to identify optimal combinations of N and E that strike a balance between maximizing information density and maintaining robust error-correction capabilities.