

## 1 Positive Definiteness of the cgDNA stiffness matrix and Invariance of the parameter set

1. For proving the positiveness of  $K = K(S, \mathcal{P})$  we will prove that  $z^T K z > 0$ , for any  $z \in \mathbb{R}^{12n-6}$ . Let decompose the matrix  $K$  into  $18 \times 18$  blocks:

$$z^T K z = \sum_{i=2}^{n-2} z_i^T K_i z_i + z_1^T K_1 z_1 + z_{n-1}^T K_{n-1} z_{n-1} \quad (1)$$

where

$$K_i = K^{X_i X_{i+1}} + \frac{1}{2} \begin{bmatrix} K^{X_i} & 0 & \\ 0 & 0 & 0 \\ 0 & 0 & K^{X_{i+1}} \end{bmatrix}, \forall i = 2, \dots, n-2,$$

$$K_1 = K^{X_1 X_2} + \begin{bmatrix} K^{X_1} & 0 & \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} K^{X_2} \end{bmatrix},$$

$$K_{n-1} = K^{X_{n-1} X_n} + \begin{bmatrix} \frac{1}{2} K^{X_{n-1}} & 0 & \\ 0 & 0 & 0 \\ 0 & 0 & K^{X_n} \end{bmatrix},$$

and  $z_i \in \mathbb{R}^{18}$ . Using the assumptions we can conclude that  $z^T K z > 0$ , for any  $z \in \mathbb{R}^{12n-6}$ .

2. We first remark that

$$\begin{bmatrix} EK^\alpha E & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & EK^\beta E \end{bmatrix} = E_3 \begin{bmatrix} K^\beta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & K^\alpha \end{bmatrix} E_3. \quad (2)$$

We denote by  $\bar{S} = \bar{X}_n \bar{X}_{n-1} \dots \bar{X}_2 \bar{X}_1$  the complementary of the sequence  $S$ , and  $\bar{K} := K(\bar{S}, \mathcal{P})$ . The position of the base  $j$  of  $\bar{S}$  in the original sequence is  $\sigma(j) = n - j + 1$ , for example, the first base in  $\bar{S}$  is the last one in  $S$ . Now, with the decomposition used in part 1) we have,  $\forall z \in \mathbb{R}^{12n-6}$ :

$$z^T \bar{K} z = \sum_{j=2}^{n-2} z_j^T \bar{K}_j z_j + z_1^T \bar{K}_1 z_1 + z_{n-1}^T \bar{K}_{n-1} z_{n-1}, \quad (3)$$

where the definition of  $\bar{K}_j, j = 2, \dots, n-2, \bar{K}_1, \bar{K}_{n-1}$  are defined in the same way as in the previous point. The only thing one have to be careful is to associate the right base pair to the matrices  $\bar{K}_j$  using the linear change of indices  $\sigma$ . Using the symmetry assumption and the

latter remark, we obtain :

$$\bar{K}_j = K^{\bar{X}_{\sigma(j)}} \bar{X}_{\sigma(j+1)} + \frac{1}{2} \begin{bmatrix} K^{\bar{X}_{\sigma(j)}} & 0 & & \\ 0 & 0 & 0 & \\ 0 & 0 & K^{\bar{X}_{\sigma(j+1)}} & \end{bmatrix} \quad (4)$$

$$= E_3 K^{X_{\sigma(j+1)} X_{\sigma(j)}} E_3 + \frac{1}{2} \begin{bmatrix} EK^{X_{\sigma(j)}} E & 0 & & \\ 0 & 0 & 0 & \\ 0 & 0 & EK^{X_{\sigma(j+1)}} E & \end{bmatrix} \quad (5)$$

$$= E_3 \left[ K^{X_{\sigma(j+1)} X_{\sigma(j)}} + \frac{1}{2} \begin{bmatrix} K^{X_{\sigma(j+1)}} & 0 & & \\ 0 & 0 & 0 & \\ 0 & 0 & K^{X_{\sigma(j)}} & \end{bmatrix} \right] E_3 \quad (6)$$

$$= E_3 K_{\sigma(j+1)} E_3. \quad (7)$$

A similar computation leads to

$$\bar{K}_1 = E_3 K_{n-1} E_3 \quad (8)$$

$$\bar{K}_{n-1} = E_3 K_1 E_3. \quad (9)$$

Finally we obtain that

$$\begin{aligned} z^T \bar{K} z &= \sum_{j=2}^{n-2} z_j^T E_3 K_{\sigma(j)} E z_j + z_1^T E_3 K_{n-1} E_3 z_1 + z_{n-1}^T E_3 K_1 E_3 z_{n-1} \\ &= \sum_{j=2}^{n-2} (E_3 z_j)^T K_{\sigma(j)} (E_3 z_j) + (E_3 z_1)^T K_{n-1} (E_3 z_1) + (E_3 z_{n-1})^T K_1 (E_3 z_{n-1}) \\ &= \sum_{i=1}^{n-2} (E_3 z_{n-i+1})^T K_i (E_3 z_{n-i+1}) + (E_3 z_1)^T K_{n-1} (E_3 z_1) + (E_3 z_{n-1})^T K_1 (E_3 z_{n-1}) \\ &= (E_n z)^T K (E_n z) = z^T E_n K E_n z \end{aligned}$$

Thus, we obtained that  $\bar{K} = E_n K E_n$ . We can prove in a similar manner that  $\bar{\sigma} = E_n \sigma$ . Finally, by using the fact that  $E_n = E_n^T = E_n^{-1}$  we obtain that

$$\mu(\bar{S}, \bar{\mathcal{P}}) = \bar{K}^{-1} \bar{\sigma} = E_n K^{-1} E_n E_n \sigma = E_n K^{-1} \sigma = E_n \mu(S, \mathcal{P}). \quad (10)$$

3. The number of independent Dimers is 10, thus using the symmetry condition given in the part 2 of this exercise we can consider only 10 conditions for the interior blocks. For the end blocks we gave two sets of 16 conditions differing only on the position of the multiplication by  $\frac{1}{2}$ . Using again the symmetry conditions we can consider only one set of conditions for the end blocks. Thus, the total number of independent conditions is 10 for the interior blocks and 16 for the end blocks.

## 2 Zero entries in the cgDNA parameter set

- i) We consider first  $\sigma^{\alpha\bar{\alpha}}$  with  $\alpha \in \{A, T, G, C\}$ . We use the following notation

$$\sigma^{\alpha\bar{\alpha}} = (\eta^\alpha, w^\alpha, u^{\alpha\bar{\alpha}}, v^{\alpha\bar{\alpha}}, \eta^{\bar{\alpha}}, w^{\bar{\alpha}}).$$

The symmetry condition leads to

$$\sigma^{\alpha\bar{\alpha}} = E_3 \sigma^{\alpha\bar{\alpha}}, \quad (11)$$

where  $E_n$  is a block trailing-diagonal matrix with  $2n - 1$  copies of  $E = \text{diag}(-1, 1, 1, -1, 1, 1)$ . Finally by computing the right hand side of (11) we get the following restrictions :

$$\begin{aligned} u_1^{\alpha\bar{\alpha}} &= 0 \\ v_1^{\alpha\bar{\alpha}} &= 0 \\ \begin{bmatrix} \eta^\alpha \\ w^\alpha \end{bmatrix} &= E \begin{bmatrix} \eta^{\bar{\alpha}} \\ w^{\bar{\alpha}} \end{bmatrix} \end{aligned}$$

Thus for  $\sigma^{\alpha\bar{\alpha}}$  you obtain eight restrictions. For  $\widehat{x}^{\alpha\bar{\alpha}}$  the restrictions are the same as for the  $\sigma^{\bar{\alpha}\alpha}$ .

We consider now the stiffness matrix  $K^{\alpha\bar{\alpha}}$  and the symmetry condition  $K^{\alpha\bar{\alpha}} = E_3 K^{\alpha\bar{\alpha}} E_3$ . The first thing to understand is how the blocks of  $K^{\alpha\bar{\alpha}}$  permute with the double multiplication with  $E_3$ . Schematically, we can divide  $K^{\alpha\bar{\alpha}}$  into nine  $6 \times 6$  blocks and, for sake of simplicity, we can number each block by taking into account the symmetry of the matrix  $K^{\alpha\bar{\alpha}}$ .

$$\begin{aligned} \begin{bmatrix} K_1 & K_2 & K_3 \\ K_2^T & K_4 & K_5 \\ K_3^T & K_5^T & K_6 \end{bmatrix} &= E_3 \begin{bmatrix} K_1 & K_2 & K_3 \\ K_2^T & K_4 & K_5 \\ K_3^T & K_5^T & K_6 \end{bmatrix} E_3 \\ &= \begin{bmatrix} EK_6E & EK_5^T E & EK_3^T E \\ EK_5E & EK_4E & EK_2^T E \\ EK_3E & EK_2E & EK_1E \end{bmatrix} \end{aligned}$$

The second step is to understand how the signs change in the multiplication  $EAE$ , where  $A \in \mathbb{R}^{6 \times 6}$ . We will denote by "+" the entries of  $EAE$  that do not change the sign while we will denote by "-" the entries that do change sign:

$$\begin{aligned} EAE &= \begin{bmatrix} -1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} A \begin{bmatrix} -1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} \\ &= \begin{bmatrix} + & - & - & + & - & - \\ - & + & + & - & + & + \\ - & + & + & - & + & + \\ + & - & - & + & - & - \\ - & + & + & - & + & + \\ - & + & + & - & + & + \end{bmatrix} \quad (12) \end{aligned}$$

Finally we remark that the inter block of  $K^{\alpha\bar{\alpha}}$ , the block  $K_4$ , do not change position when multiplied twice by  $E_3$ . Thus, we have the following condition

$$K_4 = EK_4E,$$

thanks to the +/- scheme, (12), we obtain the following restrictions on the entries of  $K^{\alpha\bar{\alpha}}$ :

$$\begin{aligned} K_{([7,10],[8,9,11,12])}^{\alpha\bar{\alpha}} &= 0 \\ K_{([8,9,11,12],[7,10])}^{\alpha\bar{\alpha}} &= 0 \end{aligned}$$

The other conditions are given by the two following relations:

$$\begin{aligned} K_1 &= EK_6E, \\ K_2 &= EK_5^T E, \end{aligned}$$

which means that by knowing  $K_6$  one can retrieve  $K_1$  and by knowing  $K_5$  one can compute  $K_2$ . The last restrictions are on the block  $K_3$  due to the relation

$$K_3 = EK_3^T E.$$

This relation implies that, in absolute values, the upper triangular part of  $K_3$  is equal to the upper triangular part of  $K_3^T$ , thus the total number of independent entries of  $K_3$  are the same as for a symmetric matrix. We stress that  $K_3$  is not symmetric, the entries in its upper triangular part are equal to the entries in the lower triangular part except for the entries corresponding to the +/- scheme, (12).

- ii) For any palindromic sequence  $S$  of arbitrary length  $n$  we can say that  $n$  is even and that the middle dimer (step between the  $\frac{n}{2}$  and  $\frac{n}{2} + 1$  base-pair) of  $S$  must be a palindromic dimer. We recall also that the reconstructions of  $K(S)$  and  $\sigma(S)$  is local but  $\mu(S)$  has a non local reconstruction because is computed using the inverse of  $K$  that is dense. Before finding all the restrictions on  $K, \sigma$ , and  $\mu$ , we recall that

$$K(\bar{S}) = E_n K(S) E_n, \quad (13)$$

$$\sigma(\bar{S}) = E_n \sigma(S), \quad (14)$$

$$\mu(\bar{S}) = E_n \mu(S). \quad (15)$$

and that the action on  $E_n$  on  $K, \sigma$  and  $\mu$  just swap all the intra/inter blocks in order to have that the  $i$ -th base-pair of  $S$  is the  $(n-i+1)$ -th base-pair of  $\bar{S}$ . We remark that the mid dimer  $X_{\frac{n}{2}} X_{\frac{n}{2}+1}$  of  $S$  correspond to the mid dimer of  $\bar{S}$  and that  $\bar{X}_{\frac{n}{2}} \bar{X}_{\frac{n}{2}+1} = X_{\frac{n}{2}} X_{\frac{n}{2}+1}$  because is a palindromic dimer. Denote now  $K_{mid} \in \mathbb{R}^{18 \times 18}$ ,  $\sigma_{mid} \in \mathbb{R}^{18}$ , and  $\mu_{mid} \in \mathbb{R}^{18}$  the components of respectively  $K, \sigma$ , and  $\mu$ , associated to the mid dimer step. Using the relations (13,14,15) we obtain that :

$$K_{mid}(\bar{S}) = E_3 K_{mid}(S) E_3,$$

$$\sigma_{mid}(\bar{S}) = E_3 \sigma_{mid}(S),$$

$$\mu_{mid}(\bar{S}) = \mu_{mid}(S).$$

and we can reuse the restrictions founded in the previous parts. Thus, the components of  $K, \sigma$ , and  $\mu$  related to the inter block are zero for a palindrome. Next step is to study the different between a local reconstruction and a non local reconstruction of a palindrome  $S$  of length  $n$  when  $S$  has additional palindromic dimers different from the middle one. In fact one can check that if  $X_i X_{i+1}$  is a palindromic dimer with  $i \neq \frac{n}{2}$ , we have that  $\bar{X}_{n-i} \bar{X}_{n-i+1} = X_i X_{i+1}$ , because is a palindrome. Thus, we will obtain the same restrictions we obtain for the middle components only for  $K$  and  $\sigma$  because for  $\mu$  we will not have the same components for the related base-pairs because  $\mu$  has a non local reconstruction. Please use the cgDNA MATLAB package to verify the findings. You can find the reconstruction of  $\sigma$  vector, in the function `constructSeqParams`.

### 3 Total number of unknowns in a cgDNA parameter set

Let  $\mathcal{P} = \{K^{\alpha\beta}, \sigma^{\alpha\beta}, K^\alpha, \sigma^\alpha | \alpha\beta \in D \text{ and } \alpha \in M\}$ . Using the Crick and Watson symmetries we decrease the numbers of dimers to 10 independent (four palindromic plus six others non palindromic dimers) and the number of bases to two independent ( for example the two purine A and G). In this way we decrease the total number of element in the parameter set to 1944. This number comes out by computing:

- 1) the entries of the ten symmetric matrices  $K^{\alpha\beta} \in \mathbb{R}^{18 \times 18}$ ,
- 2) the entries of the two symmetric matrices  $K^\alpha \in \mathbb{R}^{6 \times 6}$ ,
- 3) the entries of the ten vectors  $\sigma^{\alpha\beta} \in \mathbb{R}^{18}$ ,
- 4) the entries of the two vectors  $\sigma^\alpha \in \mathbb{R}^6$ .

Next we will count how many independent entries the cgDNAparamset has. For this we use the restrictions for the palindromic dimers founded in Ex. 2 of this session. One can easily found that for the fours  $\sigma^{\alpha\bar{\alpha}}$  we have 8 restrictions. For the stiffness matrices  $K^{\alpha\bar{\alpha}}$  is slightly more difficult. Using the same notation as for the solution of the Ex. 2 of this exercise sheet, we have that the following numbers of entries are not independent for  $K^{\alpha\bar{\alpha}}$  :

- 1) 8 entries of the block  $K_4$ ,
- 2) 21 entries of the block  $K_6$ ,
- 3) 36 entries of the block  $K_5$ ,
- 4) 15 entries of the block  $K_3$ ,

Finally we obtained that for each  $\alpha\bar{\alpha}$  we have 88 not independent entries (  $K^{\alpha\bar{\alpha}}$  plus  $\sigma^{\alpha\bar{\alpha}}$  ) which lead to a total of 1592 independent entries in the parameter set  $\mathcal{P} = \{K^{\alpha\beta}, \sigma^{\alpha\beta}, K^\alpha, \sigma^\alpha | \alpha\beta \in \tilde{D} \text{ and } \alpha \in \tilde{M}\}$ , where  $\tilde{D}$  is a set of ten independent dimers and  $\tilde{M}$  is a set of two independent monomers (bases).

### 4 Square roots in SE(3)

Using a direct computation one can easily find that

$$B = \begin{bmatrix} Q^{\frac{1}{2}} & (I + Q^{\frac{1}{2}})^{-1}q \\ 0_3^T & 1 \end{bmatrix} \quad (16)$$

**Remark:** Using the square root of a SE(3) element we can define the mid frames between two frames  $(R^n, r^n)$  and  $(R^{n+1}, r^{n+1})$ , or more precisely the mid rigid body motion between rigid body  $n$  and rigid body  $n + 1$ . Let first define the rigid body configuration for the rigid body  $n$  as

$$g_n = \begin{bmatrix} R^n & r^n \\ 0_3^T & 1 \end{bmatrix}, \quad (17)$$

equivalently we can define the rigid body configuration for the rigid body  $n + 1$ .

The mid rigid body motion from  $n$  to  $n + 1$  is then defined as:

$$g_{n+\frac{1}{2}} = g_n (g_n^{-1} g_{n+1})^{\frac{1}{2}}. \quad (18)$$

The rotation part of  $g_{n+\frac{1}{2}}$  is  $R_{n+\frac{1}{2}} = R_n(R_n^T R_{n+1})^{\frac{1}{2}}$  which correspond to the definition we used for the orientation of the mid frames, but the translational part  $r_{n+\frac{1}{2}}$  is more complicated than the definition we used ( mean value of the position  $r_n$  and  $r_{n+1}$ ). One of the disadvantage of using (18) is that the mid rigid body motion from  $n$  to  $n+1$  is different (in the translation part) than the mid rigid body motion from  $n+1$  to  $n$ . This is not the case for the definition we adopted in the cgDNA model.