

## 1 Explicit computation of apparent persistence length for a tractable probability density function (the HWLC)

During the Lecture of 10.11.2017, we introduced a simplified model of DNA in order to compute explicitly the expectation of the tangent-tangent correlation. In this exercise we will numerically check this result. A PDF version of the Lecture notes containing all the details of the simplified model and the computation can be found on the web page of the course as a supplement to the note.

1. Implement in MATLAB the Euler-Rodrigues formula ( Eq. 2 session 3 ) for computing  $Q(u)$ . Check that your code produces a rotation matrix for any  $u$ .
2. Let  $\langle Q(u) \rangle$  be the expectation of  $Q(u)$  with respect to  $\rho(u) = \frac{1}{Z} \exp\{-\frac{1}{2}(u-\hat{u}) \cdot K(u-\hat{u})\}$ , where  $\hat{u}$  is a Cayley vector with  $0^\circ$  of tilt and roll, and  $36^\circ$  of twist and  $K = \text{diag}(100, 100, 100)$ . Compute using Monte Carlo the values of the following expectations  $\langle Q(u) \rangle_{(1,3)}$  and  $\langle Q(u) \rangle_{(2,3)}$ , for  $N = 100, 1000, 10000, 100000$  samples. What do you obtain and why?
3. Now we focus only on the entries  $\langle Q(u) \rangle_{(3,3)}$  which represent the tangent-tangent correlation. Compare now the explicit result obtained in class for  $\langle Q(u) \rangle_{(3,3)}$ , i.e,

$$\langle Q(u) \rangle_{(3,3)} = 1 - \frac{2}{1 + \hat{u}_3} \left( \frac{1}{K_1} + \frac{1}{K_2} \right), \quad (1)$$

with the result obtained via Monte Carlo simulation, for the following cases:

( use  $N = 100, 1000, 10000, 100000$  samples )

- i)  $\hat{u} : 0^\circ$  tilt,  $0^\circ$  roll,  $36^\circ$  twist ,  $K = \text{diag}(543, 543, 543)$ , what is the persistence length?
  - ii)  $\hat{u} : 0^\circ$  tilt,  $0^\circ$  roll,  $0^\circ$  twist ,  $K = \text{diag}(600, 600, 600)$ , what is the persistence length?
  - iii)  $\hat{u} : 0^\circ$  tilt,  $0^\circ$  roll,  $72^\circ$  twist ,  $K = \text{diag}(393, 393, 393)$ , what is the persistence length?
4. For each case, redo the computations with smaller values of  $K_3$ . What can you say?
  5. For each case, redo the computations with smaller values of  $K_1$ . What can you say?
  6. For each case, redo the computations with smaller values of  $K_1$  and  $K_2$  (not necessarily with same value). What can you say?
  7. (optional) How small can  $K_1$  and  $K_2$  be taken with the analytical formula remaining as a good approximation ?

## 2 On the parametrization of junction displacement using quaternions

Nowadays it is rather fast to samples from multivariate distribution, thus the main part on a Monte Carlo code is in the evaluation of the chosen deterministic function. For example in the cgDNAMc we have made two different choice of functions of the cgDNA coordinates:

1.  $\phi(\mathbf{x}) = (R_1^T R_n)_{(3,3)}$ ,

2.  $\phi(\mathbf{x}) = R_1^T (r_n - r_1)$ ,

where  $(A)_{(3,3)}$  means the (3,3) entry of a matrix  $A \in \mathbb{R}^{3 \times 3}$ ,  $(R_1, r_1)$  is a fixed base-pair frame chosen to be the first (but not necessarily the first one of the DNA fragment), and  $(R_n, r_n)$  is the  $n$ th base-pair frame after the fixed one. By choosing the above function, in cgDNAMc we have to perform many matrix multiplications in  $SO(3)$  in order to be able to evaluate the functions for each sampled configuration. For efficiency, in the cgDNAMc code, these multiplications are implemented using the quaternions multiplication. We have already seen how to parametrize a rotation matrix using three numbers or the Cayley vectors. In this exercise we will study the parametrization of a rotation matrix by four numbers called *Euler-Rodrigues parameters* or *quaternions*.

Any vector  $q = (q_0, q_1, q_2, q_3) \in \mathbb{S}^3 = \{x \in \mathbb{R}^4 | x \cdot x = 1\}$  can be interpreted as a right-handed rotation in  $\mathbb{R}^3$  through an angle  $\theta$  and around a unit axis  $\mathbf{n} \in \mathbb{R}^3$ , where  $\theta$  and  $\mathbf{n}$  solve :

$$\cos \frac{\theta}{2} = q_0, \quad \text{and} \quad \mathbf{n} \sin \frac{\theta}{2} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \mathbf{q}. \quad (2)$$

1. Let  $Q \in SO(3)$  a rotation matrix about a unit axis  $\mathbf{n}$  through an angle  $0 \leq \theta < \pi$ . Let  $u = \text{Cay}(Q) \in \mathbb{R}^3$  be the Cayley parametrisation of  $Q$ . Find the quaternion parametrisation of  $Q$  in term of the Cayley vector  $u$ . [Hint: We recall that  $\|u\| = \tan \frac{\theta}{2}$ ].
2. Using the previous part, show that the Euler-Rodrigues formula (2) of exercise 1.2 session 3 implies the following quaternion parametrisation:

$$Q(q) = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_0^2 & 2(q_1 q_2 - q_3 q_0) & 2(q_1 q_3 + q_2 q_0) \\ 2(q_1 q_2 + q_3 q_0) & -q_1^2 + q_2^2 - q_3^2 + q_0^2 & 2(-q_1 q_0 + q_2 q_3) \\ 2(q_1 q_3 - q_2 q_0) & 2(q_1 q_0 + q_2 q_3) & -q_1^2 - q_2^2 + q_3^2 + q_0^2 \end{bmatrix} \quad (3)$$

3. From a computational point of view the interest of using quaternions instead of rotation matrices lies in the following equivalence: Let  $Q_i = Q(q_i) \in SO(3)$ , for  $i = 1, 2, 3$ , we have

$$Q_3 = Q_1 Q_2 \iff q_3 = q_1 \circ q_2, \quad (4)$$

where the symbol  $\circ$  mean the multiplication operator for the quaternion that for two quaternions  $q = (q_0, \mathbf{q})$  and  $p = (p_0, \mathbf{p})$  reads

$$q \circ p = (q_0 p_0 - \mathbf{q} \cdot \mathbf{p}, q_0 \mathbf{p} + p_0 \mathbf{q} + \mathbf{q} \times \mathbf{p}). \quad (5)$$

We could derive (5) and prove the equivalence (4) but here we will just check them numerically. For that purpose use the cgDNA package to reconstruct the ground-state of a short fragment of DNA (10-12 base-pair). Then, check the following

$$R_3 = R_2 Q_2 \iff q^{R_3} = q^{R_2} \circ q^{Q_2}, \quad (6)$$

where  $R_i$  is the orientation of the  $i$ th base-pair, and  $Q_2$  is the rotational part of the second junction displacement, and  $q^{M_i}$  is the quaternion related to the rotation matrix  $M_i$ .

**Remark:** In the cgDNA model the Cayley vectors are scaled in a way that their norm equal  $10 \tan \frac{\theta}{2}$ , where  $\theta$  is the angle of the rotation. Thus, you have to rescale the cgDNA Cayley vectors such that their norm are  $\tan \frac{\theta}{2}$  if you want to use the relation between Cayley vectors and quaternions.