1 On the average of rotation matrices sharing a common (deterministic) axis

Let \( Q = \{Q_k\}_{k=1}^N \subset \text{SO}(3) \) an ensemble of rotation matrices sharing a common unitary axis of rotation denoted by \( u \). The norm of the mean value is computed as the supremum of the vectorial norm \( \| \frac{1}{N} \sum_{k=1}^N Q_k x \| = \| \frac{1}{N} \sum_{k=1}^N y_k \| \) where \( y_k := Q_k x \) and \( \|x\| = 1 \). As \( x \) is a unit vector \( y_k \) is also unitary, by using the triangular inequality we obtain that

\[
\| \frac{1}{N} \sum_{k=1}^N y_k \| \leq \frac{1}{N} \sum_{k=1}^N \| y_k \| = 1. \tag{1}
\]

Let now \( x = u \), this implies that \( y_k := Q_k u = u \) and thus the triangular inequality in the previous equation is actually an equality and the axis of rotation is the vector that satisfy the supremum, i.e., \( \|\langle Q \rangle\| = \sup_{\|x\|=1} \|\langle Q \rangle x\| = \|\langle Q \rangle u\| = 1 \).

Assume now that at least one of the rotation matrix in \( Q \) is not a rotation matrix around the axis \( u \). This implies that for any choice of \( x \in \mathbb{R}^3 \) the vectors \( y_k = Q_k x \) will not all be co-linear and thus \( \|\langle Q \rangle\| < 1 \). In other words, when at least one rotation matrix has a different rotation axis, for any choice of \( x \) the euclidean average of the unit vectors \( y_k : Q_k x \) will be a point in the interior of the unit sphere.

2 On the parametrization of junction displacement using quaternions

1. Let define \( u = \text{Cay}(Q) \) the Cayley vector of \( Q \in \text{SO}(3) \) with \( \|u\| = \tan \frac{\theta}{2} \), \( 0 \leq \theta < \pi \). We need to find \( q = q(u) \in \mathbb{R}^4 \) such that \( \|q\| = 1 \) and \( q \) satisfy the relations (4) in the statement of this exercise. Thus, we have that

\[
\| (q_0, q) \| = \sqrt{q_0^2 + \|q\|^2} = \sqrt{\cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2}} = \cos \frac{\theta}{2} \sqrt{1 + \tan^2 \frac{\theta}{2}}
= \cos \frac{\theta}{2} \sqrt{1 + \|u\|^2}
= \cos \frac{\theta}{2} \parallel (1, u) \|.
\]

Finally we found that \( q = \cos \frac{\theta}{2} (1, u) \in \mathbb{R}^4 \) is a quaternion that satisfy \( \|q\| = 1 \) and the relations (4).

2. We recall that the Euler-Rodrigues formula reads:

\[
Q(u) = \frac{1 - \|u\|^2}{1 + \|u\|^2} 1 + \frac{2}{1 + \|u\|^2} [u \times] + \frac{2}{1 + \|u\|^2} u \otimes u, \tag{2}
\]

see exercise 1.2 session 3 for details. We will use now the previous part and we will define \( q = (q_0, q) = (q_0, q_1, q_2, q_3) = (\cos \frac{\theta}{2}, \cos \frac{\theta}{2} u) \) where \( \frac{q}{\cos \frac{\theta}{2}} = u \). Then the following identities
3 Comparing different truncation fit to observed MD covariance

2. We consider now the parameter sets 1 and 2. Both parameter set where computed by nu-

1. In figure (1) we show the difference between the observed covariance and the two possible

truncations for the sequence 3 of the data set. We can observe that the relative entropy

approach computes a banded stiffness matrix which inverse does not equals the observed

covariance matrix inside the stencil. On the other hand the maximum entropy (maximum

likelihood) method computes a banded stiffness matrix which inverse is equal to the

observed covariance inside the stencil. Moreover in figure (1) we show the entry–by–entry

difference (in absolute value) between the two banded stiffness matrices. We remark that the

most of the differences between the two matrices are located around the stencil.

2. We consider now the parameter sets 1 and 2. Both parameter set where computed by nu-

merically solve the same problem involving the sum of Kullback–Leibler divergences with

model in first argument of the Kullback–Leibler divergence. The only difference between

the two parameter sets is the method of truncation used to estimate the banded observed

stiffness matrices. We recall that the observed mean is not affected by the choice of trun-
cation method. We consider two sequences: poly25(AT) and poly25(AA). In figure (2) we

show the differences in the reconstructed ground–state components between reconstruction

done with parameter set 1 (blue) and reconstruction done with parameter set 2 (red). We
can observe a general agreement between the two different shape reconstructions. In figure

(3) we show the values of the (sorted) eigenvalues of the reconstructed stiffness matrices. We
can observed that the value so of the largest eigenvalues are smaller for the stiffness matrices
reconstructed using parameter set 2. This can suggest that the cgDNA parameter set 2 could

holds

\[
\frac{1 - \|u\|^2}{1 + \|u\|^2} = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = q_0^2 - q_1^2 - q_2^2 - q_3^2,
\]

\[
\frac{2}{1 + \|u\|^2} \langle u \rangle = 2q_0 \langle q \rangle = \begin{bmatrix}
0 & -2q_0 q_3 & 2q_0 q_2 \\
2q_0 q_3 & 0 & -2q_0 q_1 \\
-2q_0 q_2 & 2q_0 q_1 & 0
\end{bmatrix},
\]

\[
\frac{2}{1 + \|u\|^2} u \otimes u = 2q \otimes q = \begin{bmatrix}
q_1^2 & q_1 q_2 & q_1 q_3 \\
q_1 q_2 & q_2^2 & q_2 q_3 \\
q_1 q_3 & q_2 q_3 & q_3^2
\end{bmatrix}.
\]

The result is obtained just by using the previous identities.

3. We recall that in the cgDNA model we have a specific scaling for the rotations, thus one has

to be careful when using, for example, the Euler-Rodrigues formula or the Cayley transform.

In fact in the cgDNA model we use a scaling in such a way that if \( u = 10 \tan \frac{\theta}{7} \). It will be explained in class later on why we are using such a scaling. For the

purpose of this exercise, you have to set

\[
\rho_2,3 = \frac{u_2}{10},
\]

where \( u_2 \) is the Cayley vector related to the second inter which is related to the rotation \( Q_2 \). In order to get the quaternions for the

matrices \( R_{2,3} \) you have to use the inverse Cayley transform of part 3), exercise 1 session 3

with \( M = R_{2,3} \) to get the Cayley vector \( \rho_{2,3} \). Once you have the Cayley vectors \( \rho_{2,3} \) and \( U \)
you can use the part 1) of this exercise to compute the quaternions \( q^{R_{2,3}} \), and \( q^{Q_2} \). Compute

the quaternion multiplication to verify that \( q^{R_2} \circ q^{Q_2} = q^{R_3} \) and then use the part 2) to verify

that \( R_3 = R(q^{R_3}) \).
be "less rigid" than the cgDNA parameter set 1. In order to show that the cgDNA parameter set 2 predicts softer DNA fragments we refer to the study of the persistence length done in the article "Absolute versus relative entropy parameter estimation in a coarse–grain model of DNA", O. Gonzalez et. al., 2017, Multiscale Model. Simul. (can be found here [https://lcvmwww.epfl.ch/publications/data/articles/141/M108609.pdf]), where the authors generated randomly an ensemble of 1000 sequences of length 220 base pairs and computed their persistence length using cgDNAmc and by fitting the tangent–tangent correlation function. In figure (4) we reported the histograms of the persistence length computed over the sequence ensemble for the two cgDNA parameter sets. We observe that the histogram computed using cgDNA parameter set 2 have a lower sequence average persistence length. This suggesting that the parameter set 2 predict a less rigid DNA molecule with respect to the parameter set 1.
Figure 2: Difference in the ground–states predicted by the cgDNA parameter set 1 (blue) and parameter set 2 (red) for the sequences poly$_{25}$(AA) (top) and poly$_{25}$(AT) (bottom).
Figure 3: Difference in the (sorted) eigenvalues of the stiffness matrix reconstructed using cgDNA parameter set 1 (blue) and parameter set 2 (red) for the sequences poly$_{25}$(AA) (top) and poly$_{25}$(AT) (bottom).
Figure 10. Normalized histograms of persistence length values $\ell_p(\mathcal{P}, s_k)$ for $k = 1, \ldots, 1000$ in units of base pair (bp) obtained with the two parameter sets $\mathcal{P}_1$ and $\mathcal{P}_2$; see text. Left: Histogram of $\ell_p(\mathcal{P}_1, s_k)$ with average at 188bp. Right: Histogram of $\ell_p(\mathcal{P}_2, s_k)$ with average at 161bp.

Figure 4: Image taken from: "Absolute versus relative entropy parameter estimation in a coarse-grain model of DNA", O. Gonzalez et. al., 2017, Multiscale Model. Simul.