For questions 1-3 there are no solutions provided.

Remark for question 2: In order to be able to visualize the 3D reconstruction of the ground-state with *cgDNA 3D viewer* you have to replace, in the cgDNA `main.m` file, the following lines

```
%% Write standard Curves+ coords to file
coordinateOutputFile = 'shapes.txt';
fprintf('Saving coordinates to file <%s>... \n', coordinateOutputFile);
printShapeParms(curshapes, sequence, coordinateOutputFile);
```

with

```
%% Write standard non dimensional coords to file
coordinateOutputFile = 'shapes_nondim.txt';
fprintf('Saving non dimensional coordinates to file <%s>... \n', ...
    coordinateOutputFile);
printShapeParms(nondimshapes, sequence, coordinateOutputFile);
```

Now you can open in the *cgDNA 3D viewer* app the text file `shapes_nondim.txt`.

4 On the main outputs of the cgDNA package

4.1 Stiffness matrices

1. In Figure 1 we show the sparsity pattern of the stiffness matrix K for the sequence $S = CGCGAATTCGCG$. The sparsity pattern is 18x18 block diagonal with 6 x 6 block of overlap. Due to the overlap blocks the inverse of K is dense.

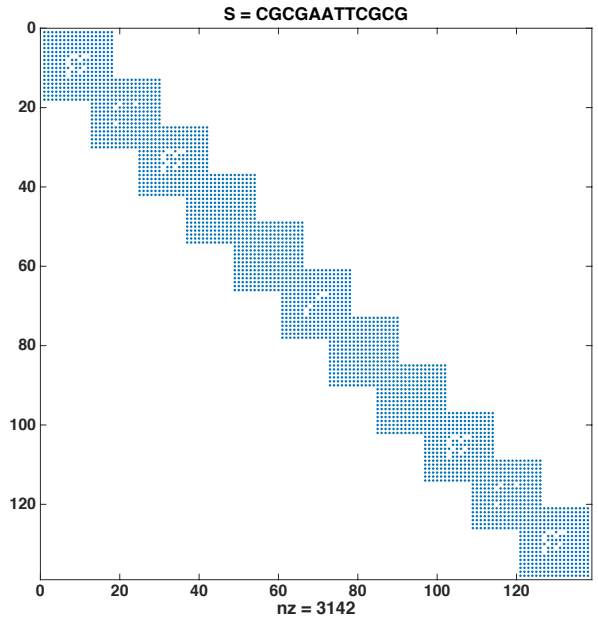

Figure 1: `spy` of the stiffness matrix K for the sequence $S = CGCGAATTCGCG$.

2. Given a sequence S, its stiffness matrix is reconstructed locally by adding 18 x 18 block based on the dimer step (two bases in a row), and by adding 6 times 6 blocks based on the monomer

(one base). We performed a single point mutation on the sequence S which leads to the following sequence $\tilde{S} = CGCGCATTCGCG$. We basically changed the fifth base A with a C. In Figure 2 we can see that the latter modification leads to the modification of two 18 x 18 blocks and one 6 x 6 overlap. In particular the single point mutation we perform changed the trimer GAA of sequence S into GCA and left the rest unchanged. Finally in Figure 2 we see the difference between the trimer GAA and the trimer GCA in the same context (i.e with the same flanking sequence).

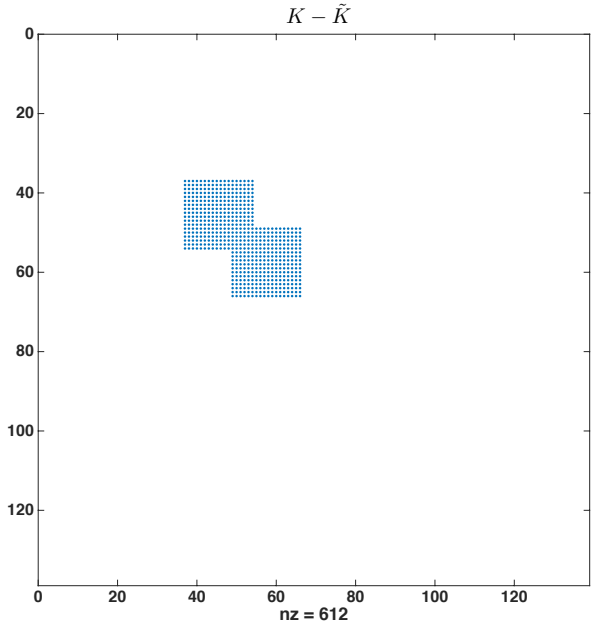

Figure 2: `spy` of $K - \tilde{K}$ where $\tilde{S} = CGCGCATTCGCG$.

3. In Figure 3 we compare the eigenvalues of K and $\tilde{K}$. We can observe that the single point mutation did not influence the eigenvalues. As said in the statement of this exercise, comparing two symmetric matrix is not a simple task. Another way to compare two symmetric positive definite matrices $K_1, K_2 \in \mathbb{R}^{n\times n}$ in a non-dimensional way that captures both differences in eigenvalues and in eigenvectors is to consider the generalised eigenvalue problem

$$K_1\mathbf{x} = \lambda K_2\mathbf{x},$$

This problem always has n positive eigenvalues, and the two matrices K_1 and K_2 are identical if and only if all the eigenvalues are 1. We will discuss this idea extensively later in the course. Here we merely show in Figure 4 the eigenvalues for the generalised eigenvalue problem between the original and point mutation cgDNA stiffness matrices.

4. In order to generate a random sequence we first set that each of the four bases have same probability to get picked. Then the sequence is constructed by picking randomly a number p from 0 to 1 and by assigning it a base, for example if $p < 0.25$ next base is A, if $p \geq 0.25$ and $p < 0.5$ next base is C, and so on. Let define K_{200} to be the stiffness matrix for the computed random sequence. In Figure 5 we plot its sorted eigenvalues. An interesting observation is that the range of the values of the eigenvalues do not increase by increasing the number of basepair. This property is a property of the cgDNA stiffness matrix.

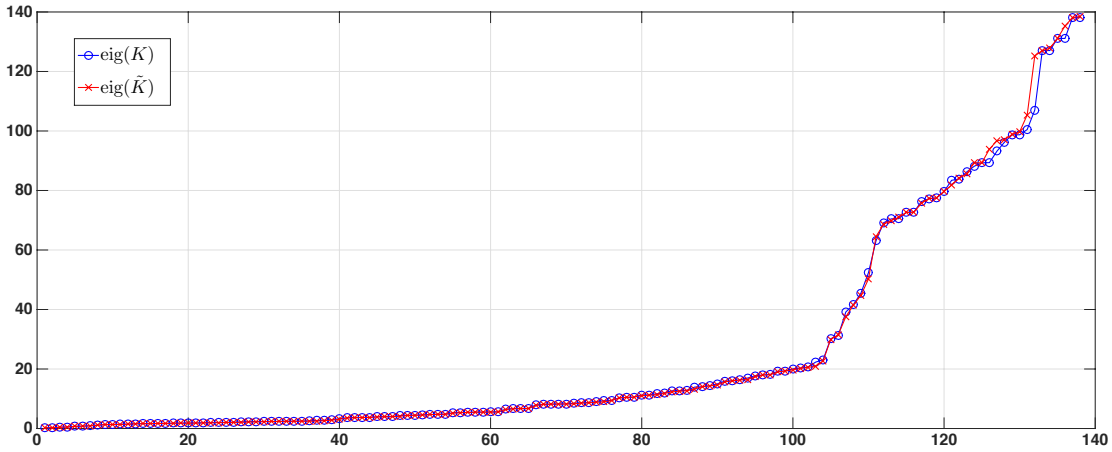

Figure 3: Sorted eigenvalues of K (circle,blue) and of $\tilde{K}$ (cross,red) as a function of degrees of freedom (ndof=138).

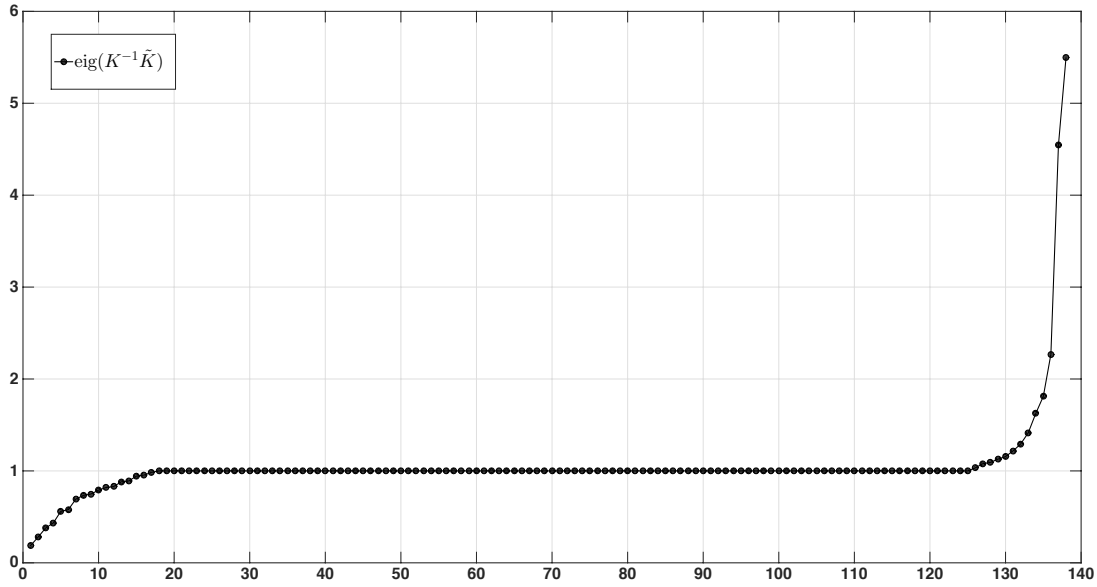

Figure 4: Sorted eigenvalues of $K^{-1}\tilde{K}$ as a function of degrees of freedom (ndof=138).

4.2 Visualization of the ground state

The first observation is that the sequence $S = CGCGAATTCGCG$ is a palindrome and that can be observed easily by inspecting the helical parameters of the ground-state shown in Figure 6. One can observe that the first column (Buckle-Shear-Tilt-Shift) are odd functions with respect to the middle junction while the other two columns are even with respect to the middle junction. The latter property is a property of the palindromes and is due to the change of readying strand properties of the cgDNA coordinates. This property will be explored in details in the coming weeks with more exercises.

1. The visualization of the 3D shape can be a useful tools for the study and analysis of the ground-state. In Figure 7 we show the bases as rigid bodies which color scheme depend upon the sequence. By observing closely we can see some features of the palindromic sequence already observed in Figure 5. For example we can observe the change in sign of the Buckle

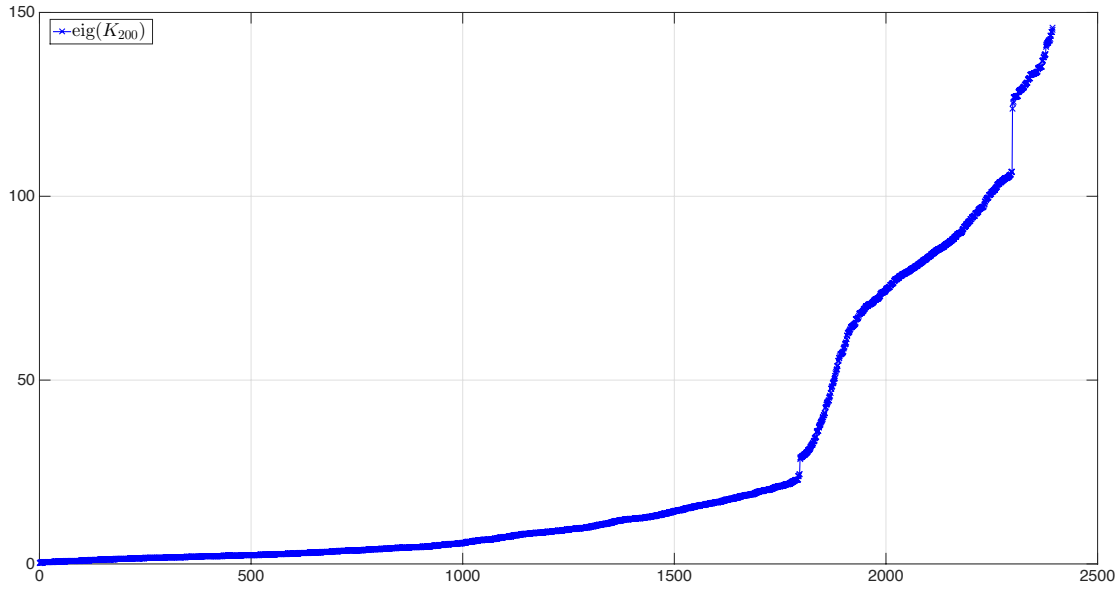

Figure 5: Sorted eigenvalues of K_{200} for a a random 200 basepair long sequence, as a function of degrees of freedom (ndof=2394).

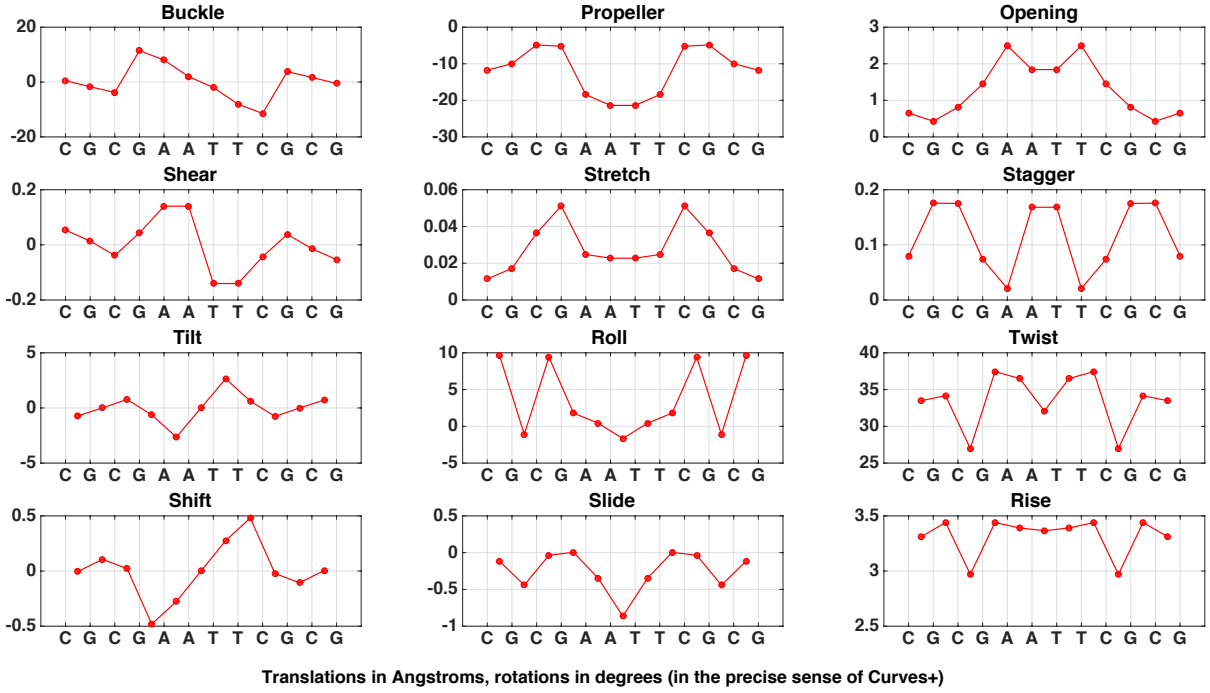

Figure 6: Helical parameters of the ground-state for the sequence S in Curves+ sense (translations in Angstroms and rotations in degrees). The rows 1 and 2 are the components of the intras while rows 3 and 4 are components of the inters. Rows 1 and 3 are the rotation parts while rows 2 and 4 are the translation parts. See Figure 1 of this exercise sheet for the cartoon of the helical parameters.

for the 4th base pair.

2. Another way of visualizing the ground-state is just to plot all the entries of the vector at once as shown in Figure 8. This way could be useful when one have to compare two or more

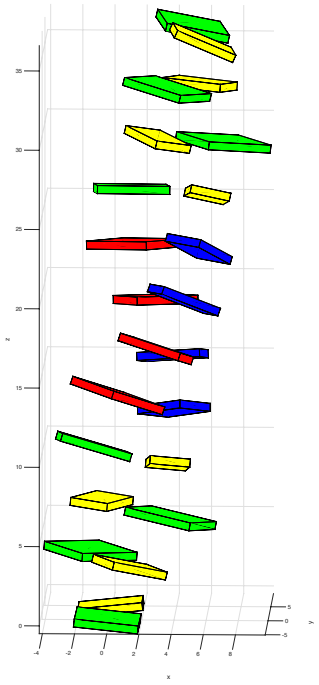

Figure 7: 3D view of the sequence S with the color scheme : yellow - C , green - G , blue - T , red - A

ground-states in a quick way, but, in the case of the sequence S, this way of visualizing the ground-state is not really useful in order to observe the palindromic property of S.

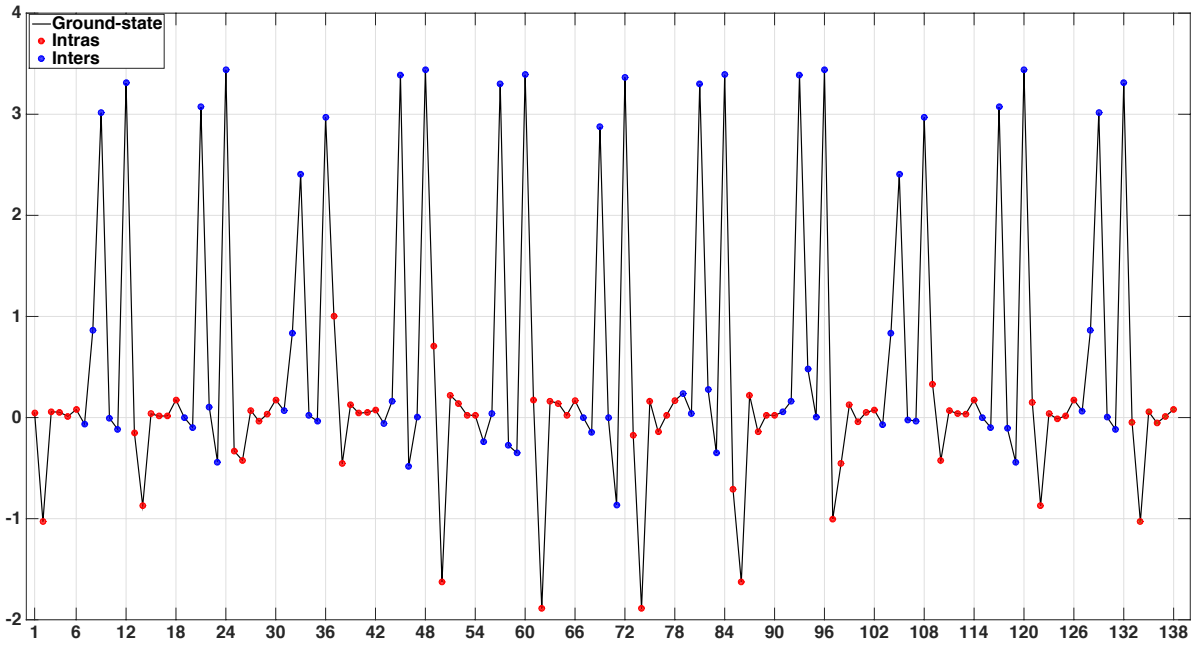

Figure 8: Non-dimensional ground-state as function of the degrees of freedom. Here we just plot all the entries of the ground-state vector at once. The red dots show the degrees of freedom corresponding to the intras while the blue dots show the one corresponding to the inters.

4.3 Understanding the function `frames`

1. Let $z = (y_1, x_1, y_2, x_2, \ldots, x_{n-1}, y_n) \in \mathbb{R}^{12n-6}$ be the ground-state for the sequence S_n. We recall that: $y_i = (\eta_i, w_i) \in \mathbb{R}^6$ are the intras, and η_i are Cayley vectors encoding the relative intra-basepair rotations while w_i are intra-basepair translations; $x_i = (u_i, v_i) \in \mathbb{R}^6$ are the inters, and u_i are Cayley vectors encoding the relative inter-basepair rotations while v_i are inter-basepair translations. The reconstruction rules for getting the absolute coordinates $\{(R_i^-, r_i^-), (R_i, r_i), (R_i^+, r_i^+)\}_{i=1}^n$ are the following:

Inter's part:

$$\begin{bmatrix} R_{i+1} & r_{i+1} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} R_i & r_i \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} P(u_i) & P(u_i)^{\frac{1}{2}} v_i \\ \mathbf{0}^T & 1 \end{bmatrix} = \prod_{j=1}^{i} \begin{bmatrix} P(u_j) & P(u_j)^{\frac{1}{2}} v_j \\ \mathbf{0}^T & 1 \end{bmatrix}, \qquad (1)$$

where we have chose as first basepair frames $(R_1, r_1) = (\mathbf{I}_3, \mathbf{0})$, and $P(u_i) = (I + \frac{1}{10}[u_i\times])(I - \frac{1}{10}[u_i\times])^{-1}$ is the Cayley transform of $[u_i\times]$

Intra's part:

$$R_i^- = R_i Q(\eta_i)^{-\frac{1}{2}},\ r_i^- = r_i - \frac{1}{2} R_i w_i \qquad (2)$$

$$R_i^+ = R_i^- Q(\eta_i),\ r_i^+ = r_i^- + R_i w_i, \qquad (3)$$

where again $Q(\eta_i) = (I + \frac{1}{10}[\eta_i\times])(I - \frac{1}{10}[\eta_i\times])^{-1}$ is the Cayley transform of $[\eta_i\times]$.

In this exercise we used two different representations of a same SE(3) element, i.e , for us an element $g \in \mathrm{SE}(3)$ can be represented as a couple $g = (R, r)$ where $R \in \mathrm{SO}(3)$ and $r \in \mathbb{R}^3$ or as a matrix $g = \begin{bmatrix} R & r \\ \mathbf{0}^T & 1 \end{bmatrix}$, where again $R \in \mathrm{SO}(3)$ and $r \in \mathbb{R}^3$. The rotation part R and the translation part r are the same in both representation. One can interpret the first representation as a frame which orientation is given by R and position is given by r while the second representation can be interpreted as a rigid body motion of the lab frame, which, in matrix representation, is just the identity matrix $\mathbf{I}_4 \in \mathbb{R}^{4\times4}$ (i.e. rotation part: $\mathbf{I}_3$, translation part $\mathbf{0} \in \mathbb{R}^3$). In the lecture we will use the latter two representations for noting elements in SE(3) that correspond to absolute coordinates of bases or basepairs.

2. The added part are at lines: 8-9 and 36-37.

```
1  nbp = (numel(shapes)+6)/12;
2
3  % absolute coordinates of the first basepair
4  G = eye(3);
5  q =[0, 0, 0]';
6
7  % store the absolute coordinates of the first basepair
8  basepair(1).G = G ;
9  basepair(1).q = q ;
10
11 % relative coordinates of the oligomer
12 [eta, w, u, v] = vector2shapes(shapes);
13
14 for i=1:nbp
15
16   % base pair:
17   r = cay(eta(i,:));
18   Gw = G * w(i,:)';
19
```

```
20    % complimentary strand
21    basepair(i).Dc = G * (sqrtm(r))';
22    basepair(i).rc = q - 0.5 * Gw;
23
24    % main strand
25    basepair(i).D = basepair(i).Dc * r;
26    basepair(i).r = basepair(i).rc + Gw;
27
28    if i<nbp
29      ru = cay(u(i,:));
30      H = G * sqrtm(ru);
31      % next base pair:
32      G = G * ru;
33      q = q + H * v(i,:)';
34
35      % store the absolute coordinates of the i-th basepair
36      basepair(i+1).G = G ;
37      basepair(i+1).q = q ;
38
39    end
40
41  end
42
43  end
```

5 A MATLAB cgDNA viewer

Here (**http://lcvmwww.epfl.ch/teaching/modelling_dna/public_files/cgDNAviewer.zip**) you can download the complete code of the cgDNA viewer and also the code for `cgDNAviewer_test.m`. Place the viewer inside your cgDNA folder.

```
1  % code to be added at the line 93 of getRigidBodyConfig.m
2  RB.conf = (R*verteces_RB')' + repmat(r',[8,1]);
```