On the length scale dependence of DNA conformational change under local perturbation

Soumyadip Banerjee\textsuperscript{a}, Kushal Shah\textsuperscript{b}, Shaunak Sen\textsuperscript{a}
\textsuperscript{a}Department of Electrical Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016, India
\textsuperscript{b}Department of Electrical Engineering and Computer Science
Indian Institute of Science Education and Research (IISER)
Bhopal 462066, Madhya Pradesh, India

Email: soumyadip@ee.iitd.ac.in\textsuperscript{a}, kushals@iiserb.ac.in\textsuperscript{b}, shaunak.sen@ee.iitd.ac.in\textsuperscript{a}

This paper is a preprint of a paper submitted to BioSystems. If accepted, the copy of record will be available at the Biosystems online archive

Abstract

Conformational change of a DNA molecule is frequently observed in multiple biological processes and has been modelled using a chain of strongly coupled oscillators with a nonlinear bistable potential. While the mechanism and properties of conformational change in the model have been investigated and several reduced order models developed, the dynamics as a function of the length of the oscillator chain is relatively less clear. To address this, we used a modified Lindstedt-Poincare method and numerical computations. We calculate a perturbation expansion of the frequency of the nonzero modes of the model, finding that approximating these modes with their unperturbed dynamics as in a previous reduced model may not hold when the length of the DNA model increases. We investigate the conformational change to local perturbation in models of varying lengths, finding that for chosen input and parameters, there are two regions of DNA length in the model, first where the minimum energy required to undergo the conformational change increases with DNA length and second, where it is almost independent of the length of the DNA model. We analyzed the conformational change in these models by adding randomness to the local perturbation, finding that the robustness of the conformational states of the model to random perturbation decreases with an increase in the DNA length. These results should help to understand the role that the length of a DNA molecule plays in influencing its conformational dynamics.

1 Introduction

A DNA sequence can manifest itself in various conformations and is observed to play an important role not only in various biological processes such as transcription, replication, DNA repairing \cite{1,2}, but also in building nanostructures and nanodevices \cite{3,4}. The conformational state that a DNA posses can undergo changes and
are usually induced by direct interaction of an enzyme with the DNA molecule. In fact in an experimental study by Harada et al. [7], the enzyme RNA polymerase is observed to rotate a DNA molecule at its active site. Also in a recent review [8], various possible mechanisms such as translocation and base-pair separation due to an enzyme called helicase are discussed which can lead to DNA unwinding. These studies show the importance of local interaction of enzymes in triggering such conformational changes.

To study these conformational changes, a possible approach is to use a simple coarse grained model of a DNA. In a recent work [9,10], a coarse grained model is used to quantitatively explain large localized amplitude fluctuations in a DNA molecule called ‘breathing’. Also in the work by Mezic [11], a similar coarse grained model of a bio-molecule is considered, where the individual bases are modeled as pendula coupled with adjacent bases specifically through torsional coupling and with in a base pair through a Morse potential. The study shows local structured perturbation to trigger conformational change or flipping in the model efficiently in terms of time and energy. In addition to this, the work reported two important findings— First, the conformational state is robust to random perturbation and second the flipping dynamics is robust to changes in the bio-molecular size.

To explain the mechanism leading to such properties, the standard tool of averaging theory [12,13] is normally used. Recently, the tool has been used to obtain a model with only a single degree of freedom [14–16]. Although the model predicts an activation condition for conformational change to happen, it fails to capture the transition in the flipping process. A more accurate reduced-order is proposed in Toit et al., which unlike the one-degree model, takes into account the dynamics of the non-zero modes of the system, responsible for inducing the transition in the flipping process. In this reduced-order model, the dynamics of the perturbed non-zero modes is replaced by their unperturbed dynamics. As per this model, the non-zero modes provides time-dependent aperiodic forcing to the large scale mode represented by the average angle of the pendula.

The reduced models mentioned works under a basic assumption that the non-zero modes influences only the zeroth or the reactive mode. However, a study of a similar system consisting of coupled non-linear oscillatory chains shows strong resonant interaction can exist between specific modes depending on the number of oscillators present in the system [19]. In such cases, it may restrict the use of reduced-order models in explaining the flipping mechanism of a coarse model of DNA molecule, especially when it has a large number of base pairs. Also, as per the recent works [14,17,18], the mechanism predicted by the reduced model is seen to have a direct influence on the flipping behavior of the model subjected to local perturbation. Although, the mechanism and properties of flipping are investigated, the dynamics is relatively unclear when the length of the model changes.

In this paper, we ask about the impact of an increase in the size of the DNA strand on the applicability of the reduced-order model and the flipping dynamics. We report that the assumptions required for the $\frac{1}{2}$ degree model do not hold as DNA length increases. We further performed a quantitative analysis of the flipping properties of the full model subjected to local perturbation. For a given parameter and input conditions, we report two different regimes of operation—first, where for most of the range, the energy required to flip increases with the DNA length and second where the energy becomes almost independent to it. Finally, we add randomness to the base selection criteria and perturb amplitude in the local
perturbation process, finding that a DNA molecule is more likely to flip as the DNA size increases when perturbed randomly.

2 DNA Model

We first present a previously developed model of a DNA [9,11], where the biomolecule is approximated as a chain of coupled pendula attached to a circular strand acting as a backbone (Figure 1(a)). The pendula represent bases of a DNA, where the mass of each base is assumed to be concentrated at the bob of the pendulum. In the model, there are two strands—one in which the attached pendula are free to rotate on a plane orthogonal to the length of the strand and two in which the pendula are fixed. The motion of each pendulum is governed by its interaction with two kinds of potential—one, through a harmonic potential which takes into account the torsional coupling between the adjacent bases of the same strand, and two through a non-linear Morse potential interaction that models bonds between adjacent inter-strand bases. This model is mathematically represented through the following equations,

\[
mh^2 \ddot{\theta}_k - S (\theta_{k+1} - 2\theta_k + \theta_{k-1}) + D \partial_\theta U(\theta_k) = 0, \quad k = 1, 2, ..., N
\]  

(1)

where, \(\theta_k\) represents the angular position of the \(k^{th}\) pendulum, \(N\) is the total number of pendula, \(m\) is the mass of each pendulum, and \(h\), the length of the pendulum. In this model, we adopt a periodic boundary condition, therefore \(\theta_1 = \theta_{N+1}\) and \(\theta_0 = \theta_{N}\). The second term in (1) corresponds to the torque due to the torsional coupling, where \(S\) being the torsional constant and the third term represent the torque due to Morse potential interaction, where \(D\) is the Morse potential amplitude. The Morse potential with unit amplitude is given as \(U(\theta) = (e^{-a(h(1 - \cos \theta)) - x_0} - 1)^2\), where, \(a\) is the decay coefficient controlling the range over which the molecular forces between the bases act [28], and \(x_0\) is the equilibrium distance between two nearest inter-strand bases respectively. The non-dimensional form of the above equation can be obtained by introducing a new time-scale \(\tau = \sqrt{mh^2/S}\). Under nominal parameter condition, \(S = 42\) eV , \(m = 300\) AMU, \(h = 10\), the time-scale \(\tau = 2.67\) ps. Calculating time derivatives with respect to \(\tau\) we have,

\[
\ddot{\theta}_k - (\theta_{k+1} - 2\theta_k + \theta_{k-1}) + \epsilon \partial_\theta U(\theta_k) = 0,
\]  

(2)

where, \(\epsilon = \frac{D}{S}\) denotes the relative amplitude of the Morse potential with respect to the harmonic torsional potential. As the hydrogen bonds between the base pairs of the complementary strands are weak compared to the covalent bonds that make up the strands [20], the dimensionless parameter \(\epsilon \ll 1\). For the given parametric regime, the Morse potential has two global minima located symmetrically about \(\theta = \pi\) as shown in Figure 1(b). Being a conservative system, the model has total energy as a constant of motion, given by the Hamiltonian, \(E\),

\[
E = \sum_{k=1}^{N} \frac{\dot{\theta}_k^2}{2} + \frac{1}{2}(\theta_k - \theta_{k-1})^2 + \epsilon U(\theta_k)
\]  

(3)

A structural property of the above described model is that it has two conformational states, wherein each, the pendula are equally displaced from their complementary
Figure 1: DNA molecule modeled as a chain of coupled pendula under the influence of a bistable potential. Schematic showing a coarse grained model of a DNA, where its bases are modeled as pendula pivoted to a strand acting as a backbone. On the same strand the pendula are coupled through linear torsional force and interstrand interaction are through a non-linear Morse potential. The interaction is represented by a double arrow. (b) The spatial profile of the Morse potential having two equilibrium position located symmetrically about the $\theta = \pi$ axis. Typical value of the parameters used to obtain the profile are — $a = 7 \text{Å}$, $h = 10 \text{Å}$ and $x_0 = 5 \text{Å}$ and the two equilibrium angular positions are given as $\theta_0 = \cos^{-1} \left(1 - \frac{x_0}{h}\right) = 1.047 \text{ rad}$ and $2\pi - \theta_0 = 5.236 \text{ rad}$. The inset in the figure shows a magnified view of the profile near one of the equilibrium angular position. (c) Phase portrait showing the contours of the projected Hamiltonian of the system, where the angular positions are equal. The separatrix in the projected phase space shown as a blue dashed line separates two regions — one, inside the separatrix, where all the pendula oscillates around the stable equilibrium point (indicated on the figure as dots labeled as $E_1$ and $E_2$), exhibiting a ‘breathing motion’ and two, outside the separatrix, where the pendula periodically crosses the $\theta = \pi$ position, exhibiting a ‘flipping motion’ [15].

The states correspond to two stable equilibrium points located symmetrically in the phase space of the system. Figure 1(c) shows the location of the points in a projected phase portrait of the system where the angular positions of the pendula are equal. A system in one of these conformed states can be subjected to a local perturbation [11], where a single or a group of spatially close pendula are disturbed from their resting position. Under such perturbation, the trajectory of the system in phase space can either remain close to the stable equilibrium point or can transit to a region near the other equilibrium point.
3 Flipping dynamics

We associate first, two characteristic features to the local perturbation process—one, the number of targeted pendula to be perturbed and two, the initial energy imparted to the system, termed as ‘perturbation energy’. We consider an example of such perturbation through the following set of initial conditions,

\[
\begin{align*}
\left( \theta_k(0), \dot{\theta}_k(0) \right) &= (\theta_i, 0) \quad k = [N/2, N/2 + 1], \\
&= (\theta_0, 0) \quad k \neq [N/2, N/2 + 1],
\end{align*}
\]

where, \(\theta_0\) is the equilibrium angular position. Physically, the initial conditions mean that the two pendula at the centre are pushed into the repulsive region where the new angular position is \(\theta_i\), while the other pendula are located at \(\theta_0\). Note that, as the above model has a translational invariance, we can choose the target pendula from any arbitrary location.

Using the above initial condition and the parametric condition, \(a = 0.7, h = 10, x_0 = 5\) and \(\epsilon = 1/1400\), we simulate (2) for 30 pendula (\(N = 30\)), for a duration of 2000 units with a time step of 0.01 units. We have used the fourth-order Runge-Kutta method [21], to perform the simulation here and in the rest of the paper. We find that the method conserves energy relatively well during the simulation (Figure 2(c)).

The DNA dynamics under the above local disturbance is as follows. Initially, the position of the centre two pendula is disturbed, through which energy is imparted into the system in the form of potential energy at the time, \(T = 0\). The perturbation energy provided to the targeted pendulums eventually spreads to all other pendula on both sides, thereby perturbing them from their resting positions. In the initial phase, the pendula move in a correlated manner near the stable equilibrium point. In this phase, the angular positions of the pendula remain bounded and the system is said to be in a ‘breathing state’ [32], where the base pairs of a DNA molecule undergoes a closed-open motion. Figure 2(a) shows the snapshots of the system’s configuration at two such time instants (\(T = 10\) units and \(T = 100\) units), where the position of the pendula are seen to be bounded near the stable equilibrium point. After a certain period, a second phase is seen, where the angular positions of the pendula start to increase. The fourth and the fifth snapshot (\(T = 250\) and \(T = 450\)) captures moments of this coherent rise of the angular positions of the pendula. It is in this phase, all pendula escape from near one stable equilibrium point and collectively moves towards the second equilibrium point (snapshot of it shown at time \(T = 600\)).

The transition from a bounded to an unbounded motion can be tracked using the average angular position of the pendula, which is calculated as,

\[\theta_{av} = \frac{1}{N} \sum_{k=1}^{N} \theta_k,\]

and the corresponding average angular velocity, denoted by \(v\). Figure 2(b) shows the average angular trajectory of the system following the contour of its projected Hamiltonian. The contours pass through a region near the equilibrium point (\(\theta_{av} = \theta_0\) and \(v = 0\)) in the phase space of the average co-ordinates, called the resonance zone [11], where the system has relatively strong interaction with the Morse potential. Note that, the collective transition of the pendula is detected when the average angular position crosses the \(\pi\) mark. This event is referred to as a ‘flip’ in the rest
Figure 2: Flip dynamics of the DNA model under the effect of a local perturbation. In the local perturbation process the middle two pendula (at \( N_2 \) and \( N_2 + 1 \)) are shifted to a new angular position \( \theta_i = 0.635 \) rad and the rest of pendula are at equilibrium position \( \theta_c = 1.047 \) rad. (a) Snapshots showing the time evolution of the DNA configuration from one stable equilibrium zone to the other obtained by simulating (2). (b) Phase portrait showing the trajectory of the DNA system projected on to the average co-ordinates \((\theta_{av}, v)\). (c) Figure shows the total energy of the system as a function of time. (d) Plot of the cross-correlation coefficient \( \rho \) as function of \( t_d \) and pendulum number \( k \). The color bar indicates the value of \( \rho \). The inset shows the standard deviation \( \sigma \) of the angular position of a pendulum from the average position of all the pendula during the simulation time.

of the paper and the time it takes for the event to happen is referred to as the ‘flip time’.

To understand the collective behavior of the pendula we calculate the cross-correlation coefficient \( \rho \) between angular position of the \( N_{th} \) pendulum at time \( t \) and of the \( k_{th} \) pendulum at time \( t - t_d \), given by

\[
\rho(k, t_d) = \frac{\int_{-T_0}^{T_0} \theta_{av}(t) \theta_k(t - t_d) dt}{\sqrt{\int_{-T_0}^{T_0} \theta_{av}^2(t) dt} \sqrt{\int_{-T_0}^{T_0} \theta_k^2(t - t_d) dt}}
\]

(6)

where, \( t_d \) is the displacement time and \( T_0 \) is the total simulation run time. Here, \( \rho \) is calculated using a built-in function in Matlab called \( xcorr \). We find from the plot shown in Figure 2(d), the motion of the pendula are highly correlated. Also, we find that with an increase in displacement time, \( t_d \), the correlation coefficient drops sharply.

In the above example, the collective motion of the system is such that the angular position of each pendulum remains in the vicinity of the average angular position.
of all the pendula with a standard deviation, $\sigma \approx 0.2$ (please see inset of Figure 2(d)). In fact, during the rotation or the flipping period, the motion of the system resembles a rigid body motion \cite{22}, where the angular distance between any two pendula remains almost equal. These rigid motions can be efficiently triggered by low-frequency modes of the system \cite{14}. In fact, in a recent work \cite{23}, using normal mode analysis, low-frequency modes have also been found to effectively induce large amplitude collective motion of atoms in a biomolecule.

The mechanism behind the above described flipping behavior can be explained through the dynamics of the Fourier modes \cite{14}. These modes are obtained by projecting the spatial angular co-ordinates on to the Fourier space through the following co-ordinate transformation,

$$\Theta = T\tilde{\Theta},$$

(7)

where, $\Theta = [\theta_1 \ \theta_2 \ldots \ldots \ldots \ \theta_N]'$ is the set of angular variables in the real space and $\tilde{\Theta} = [\tilde{\theta}_0 \ \tilde{\theta}_1 \ldots \ldots \ldots \ \tilde{\theta}_{N-1}]'$ is the set of modal co-ordinates in the Fourier space, and $T$ is a real symmetric orthonormal matrix which relates the two spaces \cite{18} and its columns are the eigenvectors in the configuration space. Note that, $\tilde{\theta}_0$ is proportional to the average angular position of the pendula, $\tilde{\theta}_0 = \sqrt{N}\theta_{av}$. Using (2) and (7), the equation for the modal co-ordinates can be obtained as

$$\ddot{\tilde{\theta}}_w = -\alpha^2_w \tilde{\theta}_w - \frac{\epsilon}{N} \sum_{n=1}^{N-1} T_{nw} G \left( \sum_{w'=0}^{N-1} T_{nw'} \tilde{\theta}_{w'} \right)$$

(8)

where $w$ is the mode number, $T_{nw}$ represent the $n^{th}$ component of the eigenvector corresponding to mode $w$ defined in the configuration space of the unperturbed system, $G(\theta) = \partial \theta U$ and $\alpha_w = 2\sin \left( \frac{\pi w}{N} \right)$ corresponds to the eigenvalue of the state matrix of the unperturbed system, which is obtained when $\epsilon$ is set to zero in (5). In the unperturbed model, the modes are the decoupled simple harmonic oscillators, which are free to oscillate at their characteristic frequency, $\alpha_w$. However, in the perturbed model, the modes become coupled and are allowed to exchange energy among themselves. The total energy of the system in the coupled condition can be given in terms of the modal co-ordinates as,

$$H = \sum_{w=0}^{N-1} \left( \frac{\dot{\tilde{p}}_w^2}{2} + \frac{1}{2} \alpha^2_w \tilde{\theta}_w^2 \right) + \epsilon \sum_{n=1}^{N} U \left( \sum_{w=0}^{N-1} T_{nw} \tilde{\theta}_w \right)$$

(9)

where, the first term of the Hamiltonian is a summation of the modal energies and the second term with $\epsilon$ as a factor represents the energy associated with the interacting modes. In a coupled-oscillatory system, such as the one studied here, modes can interact with each other through internal resonance \cite{24}. A condition for it is given in as \cite{14},

$$\left( \vec{k} \cdot \vec{\alpha} \right) < \frac{1}{c|\vec{k}|v}, \quad |\vec{k}| = |k_0|+|k_1|+\ldots+|k_{N-1}|,$$

(10)

where, $\vec{k} \cdot \vec{\alpha} = k_0\alpha_0 + k_1\alpha_1 + \ldots + k_{N-1}\alpha_{N-1}$, $k \in \mathbb{Z}^N - [0]$ and $c$ and $v$ are constants. The inequality is satisfied if the left-hand side is equal to zero. This can only happen if the modal frequencies are commensurable. The right-hand side takes into account the resonance zone in the phase space where the modal frequencies
are nearly commensurate [16]. An example of such resonance is the ‘nearly 0:1 resonance’ [18], where the zeroth order mode resonates with all the non-zero order modes of the system. Nonlinear interaction within the system can also perturb the frequencies of the modes having incommensurable frequencies, and make them commensurable for resonance to take place. The resonance condition in such cases may exist for a finite duration as it will depend on how long the system stays close to the resonance zone [16].

4 Validating reduced order model

To understand the influence of resonance in such systems, we further analyze the modal dynamics of the coupled pendulum system studied in the previous section. Tracking the time evolution of the average angle and the energy of selected modes (Figure 3(a) — (b)), we find that the change in the modal energies depends on how close the average mode is towards the resonance zone. A discrete change in the energy level of some modes is observed whenever the system comes relatively close to the equilibrium position, indicating resonance between modes. However, when the system is relatively far from the equilibrium point, the energy in the nonzero mode remains almost constant and the system approximately follows an integrable motion [25], which corresponds to nonzero modes, represented as,

$$\vec{\theta}_{w}^{approx} = A_w \cos \alpha_w t + B_w \sin \alpha_w t = \sqrt{\frac{2E_w}{\alpha_w}} \cos (\phi_w), \quad w = 1, 2 \cdots N - 1 \quad (11)$$

where, the modal phase, $\phi_w = \alpha_w t + \psi_w$ and $A_w$ and $B_w$ are constants which depend on the initial conditions, $E_w$ is the energy in the $w^{th}$ mode and $\psi_w = -\tan^{-1} \left( \frac{B_w}{A_w} \right)$. The above approximation is used to obtain a reduced order model in [17]. In the model, it is assumed that from the onset of perturbation, the nonzero modes follow unperturbed dynamics. In such cases, the only mechanism triggering the flip event is the driving of the zeroth order mode by the unperturbed nonzero modes, as such interaction would not perturb the dynamics of the nonzero modes. However, if the trajectory of the system in phase space passes through any resonance region, then the energy associated with the nonzero modes can get largely perturbed, in which case, reducing the order of the full model by averaging techniques may not be possible [26]. A basic assumption of the reduced-order model is that the frequency of the perturbed modes is equal to that of the unperturbed modes. Although the assumption holds outside the resonance zone [14], it may not hold inside it. A possible region to test the assumption would be close to the equilibrium point. Previously in [18], an approximate expression of the angular frequency of the perturbed mode is obtained, but it was using the partial averaging method, which works outside the resonance zone [14]. To compute the perturbed frequency close to the equilibrium point, we use the modified Lindstedt-Poincare method [27], as it takes into account the internal resonance by inherently eliminating secular terms arising due to it.

Following procedure outlined in [27], we expand the modal co-ordinate of the $w^{th}$ mode, $\bar{\theta}_w$ and the unperturbed modal frequency ($\alpha_w$) in (8) in different orders of $\epsilon$,

$$\bar{\theta}_w = \bar{\theta}_{0,w} + \epsilon \bar{\theta}_{1,w} + \cdots, \quad (12)$$
\[
\alpha_w^2 = \Omega_w^2 + \epsilon \Omega_{w,1}^2 + \cdots, \quad (13)
\]
where \(\theta_{0,w}, \theta_{1,w}, \ldots\) are the correction terms, \(\Omega_w\) is the perturbed angular frequency of the mode \('w'\), and \(\Omega_{w,1}, \Omega_{w,2} \cdots\) are terms chosen to eliminate secular terms. Collecting different orders of \(\epsilon\), in (8), we get,
\[
\epsilon^0 : \quad \ddot{\theta}_{0,w} + \Omega_w^2 \theta_{0,w} = 0,
\]
\[
\epsilon^1 : \quad \ddot{\theta}_{1,w} + \Omega_w^2 \theta_{1,w} = -\Omega_{w,1}^2 \theta_{0,w} - \sum_{n=1}^{N} T_{nw} G \left( \frac{1}{\sqrt{N}} \theta_{0,0} + \sum_{w'} T_{nw'} \theta_{0,w'} \right), \quad (14)
\]
Next, we perturb the system initially placed at the equilibrium position, by providing \(E_w\) amount of energy to the \(w^{th}\) nonzero mode. If the energy provided to the mode is such that \(a h \sqrt{\frac{E_w}{N}} < 1\), then the expression of the perturbed modal frequency can be written as (please see appendix for details)
\[
\Omega_w^2 = \alpha_w^2 + \epsilon 2a^2 x_0 (2h - x_0) + \epsilon \mathcal{O} \left( a^2 h^2 \frac{E_w}{N} \right), \quad (15)
\]
Under the perturbing condition considered, the perturbed modal frequency is different from the unperturbed modal frequency by an error approximately represented by the second term. For instance, if the energy provided to the system is arbitrarily small, the second term dominates over the other error terms. Further, it can be seen that the reduced order model is valid if the first term on right is much larger than the rest of the terms. For the given case, the reduced order model may not hold if,
\[
\alpha_1^2 << 2a^2 x_0 (2h - x_0) + \epsilon \mathcal{O} \left( a^2 h^2 \frac{E_w}{N} \right), \quad (16)
\]
Here, the frequency corresponding to the first mode, \(\alpha_1\), is chosen as it is a minimum possible characteristic frequency of the unperturbed system. On rewriting the above
Figure 4: Reduced order model may fail to hold in DNA model of large sizes. Figure shows the normalized frequency distribution of the relative error in the modal frequency for DNA system of different sizes. The frequency is calculated taking the derivative of $\phi$ using the first principle. The relative error in the x-axis is calculated as $|\dot{\phi} - \alpha_1\dot{\alpha}_1|$. The occurrence in the y-axis is used to denote the height of the bars. Here the height of each bar represents the relative number of cases in which the relative error is within the bin specified by the width of the bar, which is $\frac{1}{3}$. The inset in the figure shows the predicted phase $\phi$, as a function of time. The analysis is done under different perturbing condition (a) In each case, the first mode is excited such that the average energy per unit pendulum, $\frac{\dot{\phi}}{E_1} = 10^{-4}$. For the above condition, the factor $\frac{1}{2\alpha_1}\sqrt{\frac{E}{\dot{\phi}}}$ = 0.035. The duration of simulation is $5\alpha_1^{-1}$. (b) Local perturbation given to each of the DNA system such that the DNA just undergoes flip. Two pendula are perturbed towards the repulsive region such that the final position lies within the range 0.4 – 0.9 rads. The duration of the simulation is restricted to the flipping time.

The above inequality (18), with the right-hand side being the limiting value, may provide a restriction on the DNA size for the reduced-order model to hold.

To test this we perform a modal analysis on a DNA system with three different sizes $N = 10, 30$ and 100. First, we excite the first mode of the system in such a way that in each case the average energy per pendulum is the same. Second, we analyze how close the frequency of the first mode remains to its unperturbed value. We do this by applying the actual modal data to the reduced order model and predicting the phase, $\phi$ of the first mode, using the following relation derived from (11),

$$\phi(t) = -\frac{1}{\alpha_1} tan^{-1}\left(\frac{\dot{\theta}_1}{\bar{\theta}_1(t)}\right),$$

(19)

where, $\dot{\theta}_1$ and $\bar{\theta}_1(t)$ is obtained from (8). We then track the rate at which the phase, $\phi(t)$, changes to predict the perturbed modal frequency. The frequency distribution equation and assuming $\alpha_1 = 2sin\left(\frac{\pi}{N}\right) \approx \frac{2\pi}{N}$, we have,

$$N >> \sqrt{\frac{4\pi^2}{2\epsilon a^2 x_0 (2h - x_0) + \epsilon O\left(a^2 h^2 \frac{E}{N}\right)^2}}.$$

(17)

$$\Rightarrow N >> \sqrt{\frac{4\pi^2}{2\epsilon a^2 x_0 (2h - x_0)}} \approx 28,$$

(18)
plots in Figure 4(a) show how far the predicted modal frequency deviates from its unperturbed value corresponding to the DNA models of different lengths. The data shows that as $N$ increases, the relative error also increases. These deviations are seen to occur for a certain period where the slope of the modal phase becomes relatively steeper (please see the inset of Figure 4(a)). Further, we test the prediction made above for finite local perturbation as in Section 3, where the perturbing condition is chosen such that the system just flips (Figure 4(b)). The results are consistent with the trend in Figure 4(a), although the relative values are lower. Taken together, we infer that the reduced-order model might not hold for large $N$, as the modal frequencies are more likely to deviate from their unperturbed value. These results suggest that for the large length models, resonant interaction between modes may have to be considered to explain the flipping dynamics.

5 DNA length can influence flipping behavior

The foregoing analysis indicates that the length of a DNA molecule can influence its flipping mechanism. To understand how the length might affect the flipping behavior, we subject the DNA model of varying length, to local perturbation and test properties related to the flipping behavior. One such property is the energy threshold, which is the minimum energy required for flipping to take place. A possible implication of the DNA length on the energy threshold can be seen from the following example. Consider the case, where all pendula are perturbed such that they are equally pushed towards the repulsive region at the same time. Under such condition, the energy threshold can be calculated to be [17],

$$E_{\text{min}} = N \epsilon \left( e^{-a(2h-x_0)} - 1 \right)^2 = 7.14 \times 10^{-4} \times 10^{-4}, \quad (20)$$

We note from the above expression that the energy threshold is proportionally related to the DNA length, $N$. As the energy threshold inherently depends on the nature of the perturbation [11], the above condition may not be directly applicable for a local perturbation process. However, it may restrict the energy threshold for flipping to happen for such processes. To understand it, we select a local perturbation where a group of adjacent pendula, less than $N$ are targeted and are pushed equally towards the repulsive region. We call this process a 'uniform local perturbation'. For example, let $m$ number of adjacent targeted pendula, less than $N$, be perturbed such that their angular positions are shifted equally to $\theta = 0^\circ$. Under this condition, the maximum perturbation energy, $E_L$, that can be transferred to the system is derived from [3], as,

$$E_L = E \left( \dot{\theta}_1(0) = 0, \dot{\theta}_2(0) = 0, \cdots, \theta_1(0) = 0^\circ, \theta_2(0) = 0^\circ, \cdots, \theta_m(0) = 0^\circ, \theta_{m+1}(0) = \theta_0, \cdots \right)$$

$$- E \left( \dot{\theta}_1(0) = 0, \dot{\theta}_2(0) = 0, \cdots, \theta_1(0) = \theta_0, \theta_2(0) = \theta_0, \cdots \right) = \theta_0^2 + \epsilon m U(\theta = 0^\circ) \quad (21)$$

As $E_L$ is independent of the DNA length, $N$, flipping may not happen if, $E_L < E_{\text{min}}$. Considering this fact, we analyze the flipping dynamics of the DNA system by scaling the number of targeted pendula along with its length. In the present case, we choose 2% of the total number of pendulums, $N$. For instance, the perturbation provided to each pendulum is such that the final deviated position lie within the
Figure 5: Flipping properties changes with increase in DNA length.

(a) Scatter plot showing the time required to flip as a function of perturbation energy. For a DNA of a given length, 50 uniform local perturbations were given to $m = 2\%$ of $N$ number of pendula. Each perturbation shifts the targeted pendula to a new position with in a range $0.4 - 0.6$. For each case of perturbation the duration of simulation is fixed to 2000 units. (b) Scatter plot showing the energy threshold as a function of $N$. (c) Average phase portrait obtained for $N=100$, $300$, $600$ and $900$, when flip is obtained at the threshold point. The inset in the average phase portrait plot shows a snapshot of the DNA model configuration at the flipping point, where the average co-ordinate crosses $\theta_{av} = \pi$ mark at the first instance (indicated by the arrow). Also plotted the cross-correlation coefficient $\rho$ corresponding to each $N$.

such local perturbation can approximately model an enzyme’s interaction with a DNA molecule, where, it can actively perturb the bases while sliding along the strand at a sufficient speed \cite{17}. Figure 5 shows the flipping behavior of the DNA model of selected lengths (within a range, $N = 100 - 1000$), subjected to uniform local perturbation. We start with studying the variation of the flip time with the perturbation energy for DNA model of different lengths as shown in Figure 5(a). For a given length and within the perturbation energy window of $0.5 - 2$ units, it is observed that large perturbation energy corresponds to a smaller flip time. Adding to it, for a model of a given length, we find the existence of an energy threshold, below which flip event is not observed. Also, we notice that these observations are in line with the pattern found in \cite{11, 17}, where such flipping properties are studied of a similar model under the influence of structured
perturbation. We also find that for the uniform local perturbation case, the energy threshold changes with $N$ as shown in Figure 5(b). Interestingly, it shows two different regimes, where the DNA model may be operating. First, in the length range $N \approx 100 - 600$, where for most of the range, the energy threshold is observed to increase with $N$, whereas in the length range $N \approx 600 - 1000$, the threshold remains almost constant. Note that this is different from a result reported of a similar analysis in Mezic’s work [11], where a similar model is studied under half perturbation (number of targeted pendula being 50% of $N$), but for smaller length models. In his work, the flipping behavior is seen to be robust over the range of $N$ considered.

Further to understand the flipping behavior at the threshold point, we plot the trajectory of the DNA system projected on to the average co-ordinates, corresponding to four different DNA lengths, as shown in Figure 5(c). In case for DNA lengths, $N = 100$ and $N = 300$, the perturbed system initially remains in a ‘breathing’ state for a certain period, before undergoing a flip to the other equilibrium point. Also, we computed the cross-correlation of the motion of the centre pendulum with the other pendula, finding that their motions are well correlated and approximately follow a rigid body dynamics similar to the dynamics obtained in the Section 3 example. We also notice that during the transition period the angular velocity is lower for $N = 300$ as compared to $N = 100$. A similar trend is seen in [29], where a low-resolution model of a DNA is considered, where the motion of the bases is subjected to stochastic fluctuation. They found that the average time taken by the molecule to rotate by a fixed amount increases with the number of base pairs in the DNA molecule. In the second regime, however, we find a significant change in the flipping behavior. Unlike in case of the previous length models belonging to regime I, in regime II, the motion of system during the flipping period deviates from the rigid body type motion. A possible evidence of it can be seen in the correlation plot shown in Figure 5(c), where it is seen that the pendula which are relatively far from the centre pendulum, their motion are less correlated. Although the distant pendula such as the centre and the first pendulum are weakly correlated during the simulation time, a relatively good correlation exists between them when the motion of the centre pendula is observed after a certain lag. For instance, in case of $N = 900$, the lag, $t_d \approx 500$. This may indicate rotational deformation of the DNA model along its length. Further evidence of the deformation can be seen in the inset of Figure 5(c), which shows the DNA in a locally unwinded state [33] at the flipping instant. Note that the flipping behavior in this regime is comparatively different from the behavior shown in a similar model in earlier works [11,17], which shows a rigid behavior as observed in the lower length models of regime I.

In the above computation, we studied the flipping behavior of the DNA model to uniform local perturbation where adjacent target pendula are chosen and identically perturbed. However, from a biological perspective, such a perturbation process can be inherently stochastic. For instance, in the case of an enzyme interacting with a DNA molecule, the process may model an enzyme taking random step size during the translocation process [8] and perturbing the bases as it moves. A possible way of adding this in the DNA model can be through a random choice of the initial conditions. In [11,17], the flipping properties of a similar DNA model, with smaller lengths, have been tested by subjecting it to such random perturbations. Similar to these analyses, we next investigate how randomness in the local perturbation process may affect flipping in relatively larger length models.
Figure 6: Flipping becomes more robust to random perturbation with increase in DNA length.

(a) Figure shows a schematic of a modified model of the local perturbation. For \( m \) number of targeted pendula, the spacing between the \( k \)th and \( k+1 \)th pendulum is \( x-1 \), where \( x \) is an integer chosen randomly within a range \( 1-\sigma+1 \), following a uniform probability distribution \( f_{\sigma+1}^m(x) = \frac{1}{\sigma+1} \), where \( \sigma \) is the maximum spacing between the consecutive pendula selected. For each pendulum chosen it is given a push towards the repulsive region where its new position is a random variable \( y \) with a uniform probability distribution \( f_M^m(y) = \frac{1}{M} \), where \( M \) is the total number of samples the angular range of \( y \) is divided into. (b) Scatter plot showing for each case of DNA length, the time to flip as a function of perturbation energy for zero spacing (\( \sigma = 0 \)). In each case of \( N \), 50 perturbation were given and for each perturbation, the number of pendula targeted is \( m = 2\% \) of \( N \) and each targeted pendulum is deviated to a new angle within a range \( 0.4-0.6 \) rad divided into \( M = 50 \) segments. (c) Histogram plot showing the frequency distribution of the number of flips as a function of maximum spacing \( \sigma \) for different DNA sizes.

We start by introducing randomness in both steps of the local perturbation process. One, in the selection of the target pendula and two, in the amount of deviation given to each of the selected pendula. We describe the method using the following rules:

1. Let \( m = 2\% \) of \( N \) pendula be targeted for perturbation. If \( P_k \) represents the position of the \( k \)th target pendulum, then the position of the \( k+1 \)th target pendulum is given as

\[
P_{k+1} = P_k + x \quad 1 \leq k \leq m-1 \quad x \in [1, 2, ... \sigma + 1],
\]

where, \( P_1 = 1 \), \( x \) is a random variable following a uniform discrete probability distribution, \( f_{\sigma+1}^m(x) = \frac{1}{\sigma+1} \), and \( x-1 \) is the spacing between the two adjacent target pendula (see Fig 6(a)). A zero spacing between the pendula means that...
they are adjacent to each other. Also, we provide an approach to control the randomness in this process. In order to control it, we first measure it using the Boltzmann-Shannon entropy \[30\], \(H\). For a given sample length \(\sigma + 1\), the entropy is given as 
\[
H = -\sum_{i=1}^{\sigma+1} p_i \log_2 p_i = \log_2 (\sigma + 1),
\]
where, 
\[
p_i = f_{m+1}^M (x) = \frac{1}{M}.
\]
Note, as the randomness associated with this process is directly related to the sample size \(\sigma + 1\), we use the \(\sigma\) parameter to control it. For instance, if \(\sigma = 0\), it corresponds to zero entropy or randomness in the base selection process. Also, note that as the average spacing between the targeted pendulum is \(\frac{\sigma}{2}\), the randomness in the perturbation step can be directly related to how localised the perturbation is.

2. After selection, each targeted pendulum is pushed in to the repulsive region where its new position is a random variable \(y\) following a uniform discrete probability distribution 
\[
f_M^M (y) = \frac{1}{M},
\]
where \(M = 50\) is the total number of segments the angular range of \(y\) is divided into (see Figure 6(a)).

First, we study the case where adjacent pendula are targeted and pushed randomly within the angular range \(0.4 \rightarrow 0.6\) rad. In this case, for different length models, we compute the variation of flip time with the perturbation energy (Figure 6(b)). We find that the lower length models \((N \approx 100 \rightarrow 600)\) show similar flipping properties when compared to the case when they are subjected to uniform local perturbation (Figure 5(a)). However, when the properties of the larger length models are compared, the pattern is observed to be different, with the frequency of flip events to be lower within the energy range \(0.5 \rightarrow 2\) units. Second, we introduce randomness in the base selection process and investigate how it may influence the flipping behavior. For each length of the DNA model and a given entropy in the base selection process, we provide 50 random perturbation to the selected bases using rule 2 and count the occurrence of a flip event. The result obtained in Figure 6(c) suggests that the system with lower lengths \(\approx 100 \rightarrow 600\) are relatively immune to conformational change when subjected to perturbation with large randomness. This is different from the case where the randomness in the perturbation process is relatively low, which shows more flipping events. A reason for it could be that when randomness is reduced, the perturbation tends to be more local, which increases the chances of flipping in the model [17]. However, in the case for other lengths, the trend does not hold and the flip event becomes relatively sensitive to random perturbation. Note, although the observation for the lower length models is in line with the trend seen in [11, 14, 17], it is different for the larger length models. This indicates \(N\) may play a vital role in controlling the robustness of flipping behavior of a DNA molecule to random perturbation.

6 Discussion

Developing coarse models of flipping in a DNA molecule is important in understanding how structural features of the molecule could influence its dynamics. In this paper, we investigated whether and how the DNA size affects the applicability of a previously developed reduced order model and the flipping properties of the full model under local perturbation. Using the modified Lindstedt-Poincare method, we investigated the perturbed frequency of the modes close to the equilibrium point for single mode perturbation. The findings of our study suggests that the approximations on which the reduced order model is based may not hold for DNA models
with sufficiently large sizes as the modal frequency is observed to deviate significantly from their unperturbed value. Further, in order to understand how increase in DNA length may affect the flipping behavior, we numerically simulated the full model with comparatively large sizes by subjecting it to uniform local perturbation. For the given parameter and input properties, we find that for most of the range, N≈ 100 – 600, threshold energy required to flip increases with N, whereas after N≈600, the energy remains almost constant. Finally, we find that with increase in DNA size, flipping of the DNA model becomes more robust to random perturbation in the sense that such perturbation can readily trigger flipping.

It is interesting to note that the length of a DNA molecule can influence its effective rigidity. For lower lengths, the DNA model is essentially seen to follow a rigid body dynamics, whereas for larger lengths it behaves as a flexible body where the DNA backbone is seen to undergo rotational deformation along its length. The change in the behavior is similar to the change observed in the bending property of a semi-flexible polymer modeled as a worm like chain [36], when its length crosses a characteristic bending length scale called the persistence length [35]. For instance, for lengths shorter than the persistence length, the molecule shows high bending rigidity, whereas for longer lengths it behaves as a flexible body which can be bent easily. A possible reason for the change in the bending behavior could be the strong influence of the low frequency modes on the overall motion of the DNA system [37]. In fact the strong deformation observed in our model could be the result of considerable amount of energy being funneled into the lower modes as compared to the higher modes. According to our study, this may occur because of the large perturbation the frequency of the lower modes undergo, which makes it more likely to be commensurate with the frequency of the higher modes, resulting in resonance and energy exchange between them.

We note that apart from the length of the model, there are other physical factors which can influence the DNA dynamics such as the stiffness of the bonds present in the DNA molecule. This is evident in the perturbation term of the modal frequency expression (15) obtained for the single mode perturbation case. For instance, for the present case, two factors on which the stiffness depends are — one, ϵ which indirectly depends on the strength of the torsional interaction in the molecule’s backbone and two, directly on the decay coefficient a, which is directly related to the stiffness of the hydrogen bonds between the base pairs. Our future work would be to investigate whether a suitable values of these parameter could be obtained such that the reduced order model would hold when the length of the full model is large.

In conclusion, our work highlights the importance of length of a DNA model in controlling its flipping dynamics and also may help us to understand how an enzyme can use this property of the molecule to efficiently trigger its unwinding.

7 Acknowledgements

S.B. would like to thank the Visvesvaraya Ph.D. Scheme for Electronics & IT, Department of Electronics and Information Technology (DeitY), Government of India, for financial support (File No. IITD/IRD/MI01233).
8 Appendix

Consider in the DNA model all the pendula are initially resting at the equilibrium angular position, \( \theta_0 \). Next, we excite its \( w^{th} \) nonzero mode through the following set of initial conditions,

\[
\left( \tilde{\theta}_{0,w'}(0), \tilde{\theta}_{0,w'}'(0) \right) = \begin{cases} 
(0, \sqrt{2E_w}) & w' = w \text{ and } w' \neq 0, \\
(0, 0) & w' \neq w \text{ and } w' \neq 0,
\end{cases}
\]

(23)

Note, here \( E_w \) is the energy imparted to the \( w^{th} \) mode at time \( t = 0 \). Under the above condition, the \( \epsilon^0 \) order solutions of the modes can be written as

\[
\epsilon^0 : \quad \tilde{\theta}_{0,w'} = A_{w'} \sin (\Omega_w t), \quad w' = w \text{ and } w \neq 0
\]

\[
= 0, \quad w' \neq w \text{ and } w' \neq 0
\]

(24)

where, \( A_{w'} = \sqrt{2E_w} \). Next, using (14) and (23) we obtain the \( \epsilon \) order dynamics of the \( w^{th} \) nonzero mode given as

\[
\epsilon : \quad \ddot{\bar{\theta}}_{1,w} + \Omega_w^2 \bar{\theta}_{1,w} = -\Omega_w^2 A_w \sin (\Omega_w t) - \sum_{n=1}^{N} T_{nw} G (\theta_0 + T_{nw} A_w \sin (\Omega_w t)),
\]

(25)

On expanding function \( G (\theta) \) about \( \theta_0 \), in (25) we get,

\[
\ddot{\bar{\theta}}_{1,w} + \Omega_w^2 \bar{\theta}_{1,w} = -\Omega_w^2 A_w \sin (\Omega_w t) - G (\theta_0) \sum_{n=1}^{N} T_{nw} - \partial_\theta G (\theta_0) A_w \sin (\Omega_w t) \sum_{n=1}^{N} T_{nw}^2 - \cdots
\]

(26)

Note here, \( \partial_\theta^{k-1} G (\theta_0) \) represent \( \partial_\theta^{k-1} G (\theta) \) evaluated at \( \theta_0 \). The second term on the right hand side vanishes as \( G (\theta_0) = 0 \). Next, we evaluate \( \sum T_{nw}^k \) in (26). If \( w < N/2 \), then each element of the \( T \) matrix can be written in the form as

\[
T_{nw} = \sqrt{\frac{2N}{\pi}} \cos \left( \frac{2\pi nw}{N} \right)
\]

(27)

Let \( \frac{2\pi nw}{N} = \phi \), then for even \( k \), we have,

\[
\sum_{n=1}^{N} T_{nw}^k = \left( \frac{2}{N} \right)^{k/2} \frac{1}{2^k} \sum_{n=1}^{N} \cos^{k} \left( \frac{2\pi nw}{N} \right)
\]

(28)

Let \( \frac{2\pi nw}{N} = \phi \), then for even \( k \), we have,
Let \( 2m - k = z \), and replacing \( \phi = \frac{2\pi m}{N} \) above, we simplify the summation term over \( n \) within the summation term over \( m \)

\[
\sum_{n=1}^{N} \cos (z\phi) = \sum_{n=1}^{N} \cos (n\phi') \quad \phi' = \frac{2\pi zw}{N}
\]

\[
= \text{Re} \left[ \sum_{n=1}^{N} e^{in\phi'} \right] 
= \text{Re} \left[ e^{i\phi'} \left( e^{iN\phi'} - 1 \right) \left( e^{rac{N\phi'}{2}} - e^{-i\frac{N\phi'}{2}} \right) \right] 
= \text{Re} \left[ e^{i(N+1)\phi'} \left( \sin \frac{N\phi'}{2} \right) \sin \frac{\phi'}{2} \right] 
= 0, \quad \text{as } N\phi' = 2\pi zw
\]

Using the above relation in (28), we obtain,

\[
\sum_{n=1}^{N} T_{nw}^{k} = \frac{1}{2^k} \frac{1}{N^\frac{k}{2}-1} C_k^\frac{k}{2}
\]

(30)

For odd \( k \), following the same procedure as above, we get,

\[
\sum_{n=1}^{N} T_{nw}^{k} = \sum_{n=1}^{N} \left( \frac{2}{N} \right)^{k/2} \frac{1}{2^k} \left[ 2 \sum_{m=0}^{k-1} C_m^k \cos (2m - k) \phi \right] 
= 0
\]

(31)

From above result the terms in (26) with odd \( k \) vanishes. For even \( k \), to calculate the coefficient of the secular term, we first extract the coefficient of \( \sin (\Omega_w t) \) from \( \sin^{k-1} (\Omega_w t) \). Let \( k - 1 = k' \) and \( \Omega_w t = \Phi \), then,

\[
sin^{k-1} (\Omega_w t) = \sin^{k'} (\Phi) = \left( \frac{e^{i\Phi} - e^{-i\Phi}}{2i} \right)^{k'} 
= \frac{1}{(2i)^{k'}} \sum_{m=0}^{k'} (-1)^{k'-m} C_m^{k'} e^{im\Phi} e^{-i(k'-m)\Phi}
= \frac{1}{(2i)^{k'}} \sum_{m=0}^{k'} (-1)^{k'-m} C_m^{k'} e^{i(2m-k')m\Phi}
\]

(32)

The secular term can be obtained from the above expansion, by combining the terms corresponding to \( m = \frac{k'-1}{2} \) and \( m = \frac{k'+1}{2} \),

\[
sin^{k'} (\Phi) = \frac{1}{(2i)^{k'}} \left[ \ldots + C_{\frac{k'+1}{2}}^{k'} (-1)^{\frac{k'+1}{2}} e^{i\Phi} + C_{\frac{k'+1}{2}}^{k'} (-1)^{\frac{k'+1}{2}} e^{-i\Phi} + \ldots \right] 
= \ldots + \frac{1}{(2i)^{k'}} C_{\frac{k'+1}{2}}^{k'} (-1)^{\frac{k'+1}{2}} \left[ \frac{e^{i\Phi} - e^{-i\Phi}}{2i} \right] + \ldots
= \ldots + \frac{1}{2^{k'-1}} C_{\frac{k'+1}{2}}^{k'} \sin \Phi + \ldots
\]

(33)
Combining the results of (26) and (33) we get the coefficient of secular term which is equated to zero,

$$-\Omega_{w,1}^2 A_w - \sum_{k=even,k>1} \frac{1}{(k-1)!} \partial_{\theta}^{k-1} G(\theta_0) A_w^{k-1} \frac{1}{2^{k-2}} C_{k-2}^{\frac{1}{2}} \frac{1}{2} N \frac{1}{2} - 1 \bar{C}_{k}^{\frac{1}{2}} = 0 \quad (34)$$

On simplifying the second term,

$$\sum_{k=even,k>1} \frac{1}{(k-1)!} \partial_{\theta}^{k-1} G(\theta_0) A_w^{k-1} \frac{1}{2^{k-2}} C_{k-2}^{\frac{1}{2}} \frac{1}{2} N \frac{1}{2} - 1 \bar{C}_{k}^{\frac{1}{2}} =$$

$$\sum_{k=even,k>1} \partial_{\theta}^{k-1} G(\theta_0) A_w^{k-1} \frac{1}{2^{k-2}} \frac{1.3.5 \cdots k - 1}{(k!)^2 (k - 1)!} N \frac{1}{2} - 1 \quad (35)$$

We obtain the perturbed frequency as,

$$\Omega_{w}^2 = a_w^2 + \epsilon \partial_{\theta}^2 U (\theta_0) + \epsilon \partial_{\theta}^4 U (\theta_0) A_w^2 \frac{1}{16} N + \cdots$$

$$+ \epsilon \partial_{\theta}^6 U (\theta_0) A_w^{k-2} \frac{1}{2^{k-2}} \frac{1.3.5 \cdots k - 1}{(k!)^2 (k - 1)!} N \frac{1}{2} - 1 \cdots, \text{where} \ k = 2, 4, 6 \cdots \quad (36)$$

Next we find the order of the function $\partial_{\theta}^k U (\theta_0)$, we use the Faà di Bruno’s formula [36], through which we write the $k^{th}$ derivative of $U(\theta)$ evaluated at $\theta = \theta_0$, in terms of derivative of $U$ with respect to a new variable, $z = h (1 - \cos \theta)$. It is given as,

$$\partial_{\theta}^k U (\theta_0) = \sum_{m_1, m_2, \ldots, m_k} \frac{k!}{m_1! m_2! \cdots m_k!} \partial_z^K U (z(\theta_0)) (\partial_z z(\theta_0))^{m_1} \left( \frac{\partial_z^2 z(\theta_0)}{2!} \right)^{m_2} \cdots \left( \frac{\partial_z^k z(\theta_0)}{k!} \right)^{m_k} \quad (37)$$

where, $m_1, m_2, \ldots, m_k$ are non-negative integers and $K = m_1 + m_2 + \cdots + m_k$, where the sum obeys a partition law given as,

$$m_1 + 2 m_2 + \cdots + k m_k = k \quad (38)$$

The $K^{th}$ derivative of $U$ with respect to $z$ inside the summation of (37) is given as,

$$\partial_z^K U = 2 \left( (-1)^{K-1} e^{-(z - \bar{x}_0)} + (-1)^K 2^{K-1} e^{-(z - \bar{x}_0)} \right) \quad (39)$$

As per the properties of asymptotic notations [36], the following can be expressed as, $\partial_z z(\theta_0) = \hbar \sin \theta_0 = O(\bar{h})$, $\partial_z^2 z(\theta_0) = h \cos \theta_0 = O(\bar{h}) \cdots \partial_z^k z(\theta_0) = O(\bar{h})$, then,

$$(\partial_z z)^{m_1} \left( \frac{\partial_z^2 z}{2!} \right)^{m_2} \cdots \left( \frac{\partial_z^k z}{k!} \right)^{m_k} = O(\bar{h}^{m_1}) O(\bar{h}^{m_2}) \cdots O(\bar{h}^{m_k}) = O(\bar{h}^{m_1 + m_2 + \cdots + m_k})$$

[If $f_1 = O(g_1)$ and $f_2 = O(g_2)$, then, $f_1 f_2 = O(g_1 g_2)$] \quad (40)

As, $m_1 + m_2 + \cdots + m_k \leq m_1 + 2 m_2 + \cdots + k m_k = k$, then of all the combination of $m_1, m_2, \ldots, m_k$, the combination which gives, $m_1 + m_2 + \cdots + m_k = k$, determines the order of the above expression. Also, the value of $m_1, m_2, \ldots, m_k$, which would satisfy the condition $m_1 + m_2 + \cdots + m_k = k$ and $m_1 + 2 m_2 + \cdots + k m_k = k$ would be $m_1 = k$ and $m_l = 0 \forall l \neq 1$. This gives

$$\partial_{\theta}^k U = \partial_z^k U O(\bar{h}^k) \quad (41)$$
Also, from (39), the order of \( \partial^k_x U = O \left( 2^k \right) \), which gives,

\[
\partial^k_x U = O \left( 2^k \bar{h}^k \right) \tag{42}
\]

Lastly, as \( \frac{1.3.5 \cdots k-1}{(\frac{k}{2})!(\frac{k}{2}-1)!} < 1 \) \( \forall k \) \(^1\) then using (42), the order of the \( k^{th} \) term can be obtained as,

\[
\epsilon \partial^k_x U \left( \theta_0 \right) A_{w}^{k-2} \frac{1}{2^{k-2}} \frac{1.3.5 \cdots k - 1}{(\frac{k}{2})!} \frac{1}{N^{\frac{k}{2}-1}} = \epsilon O \left( \frac{A_{w}^{k-2} \bar{h}^k}{N^{\frac{k}{2}}} \right) \tag{45}
\]

Now, if \( \frac{A_{w} \bar{h}}{N} \) < 1, then \( 3^{rd}, 4^{th} \cdots \) terms of (39) can be written as,

\[
O \left( \frac{A_{w}^2 \bar{h}^4}{N} \right) + O \left( \frac{A_{w}^4 \bar{h}^6}{N^2} \right) + \cdots = O \left( \frac{A_{w}^3 \bar{h}^2}{N} \right) + O \left( \frac{A_{w}^4 \bar{h}^4}{N^2} \right) + \cdots = O \left( \frac{A_{w}^2 \bar{h}^2}{N} \right) \tag{46}
\]

Using the property, \( O \left( cg \right) = O \left( g \right) \), where, \( c \) is a constant and if, \( f_1 = O \left( g_1 \right) \) and \( f_2 = O \left( g_2 \right) \), then, \( f_1 + f_2 = O \left( \max \left( g_1, g_2 \right) \right) \). Finally, using the above result, (39) can be written as,

\[
\Omega^2_w = \alpha^2_w + \epsilon 2a^2x_0 \left( 2h - x_0 \right) + \epsilon O \left( \frac{A_{w}^2 \bar{h}^2}{N} \right)
\]

\[
= \alpha^2_w + \epsilon 2a^2x_0 \left( 2h - x_0 \right) + \epsilon O \left( a^2h^2 \frac{E_w}{N} \right) \tag{47}
\]

where, \( A_{w} = \sqrt{2E_w} \) and \( \bar{h} = ah \).

References

[1] Bochman, M.L., Paeschke, K. and Zakian, V.A., 2012. DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet., 13(11), p.770.

[2] Barat-Gueride, M., Dufresne, C., Rickwood D., 1989. Effect of DNA conformation on the transcription of mitochondrial DNA. Eur. J. Biochem. 183(2), p.297-302.

[3] Nouspikel, T. D. N. A., 2009. DNA repair in mammalian cells. Cell. Mol. Life Sci., 66(6), p.994-1009.

\(^1\)Let \( P \left( k \right) = \frac{1.3.5 \cdots k-1}{(\frac{k}{2})!(\frac{k}{2}-1)!} \), where \( k = 4, 6 \cdots \) then, using mathematical induction,

\[
P \left( 4 \right) = \frac{3}{4} < 1 \tag{43}
\]

Let \( P \left( k \right) < 1 \) holds true,

\[
P \left( k + 2 \right) = \frac{1.3.5 \cdots k - 1 \cdot k + 1}{(\frac{k}{2})!(\frac{k}{2}-1)!} = \frac{1.3.5 \cdots k - 1 \cdot k + 1}{(\frac{k}{2} + 1)^2 (\frac{k}{2})^2 (\frac{k}{2} - 1)!}
\]

\[
= P \left( k \right) \frac{k + 1}{(\frac{k}{2} + 1)^2 (\frac{k}{2})} = P \left( k \right) \frac{1}{k(k+1) + 2k + \frac{k}{\pi+1}} < 1 \forall k
\]

Therefore, the statement \( P \left( k \right) < 1 \) holds true.
[4] Wilner, O. I. and Willner, I., 2012. Functionalized DNA nanostructures. Chem. Rev., 112(4), p.2528-2556.
[5] Simmel, F. C., and Dittmer, W. U., 2005. DNA nanodevices. Small, 1(3), p.284-299.
[6] Niemeyer, C. M., and Adler, M., 2002. Nanomechanical devices based on DNA. Angew. Chem, 41(20), p.3779-3783.
[7] Harada, Y., Ohara, O., Takatsuki, A., Itoh, H., Shimamoto, N. and Kinoshita Jr, K., 2001. Direct observation of DNA rotation during transcription by Escherichia coli RNA polymerase. Nature, 409(6816), p.113.
[8] Patel, S.S. and Donmez, I., 2006. Mechanisms of helicases. J. Biol. Chem., 281(27), p.18265-18268.
[9] Peyrard, M., Cuesta-Lopez, S. and James, G., 2009. Nonlinear analysis of the dynamics of DNA breathing. J. Biol. Phys., 35(1), p.73.
[10] Peyrard, M., 2004. Nonlinear dynamics and statistical physics of DNA. Nonlinearity, 17(2), R1.
[11] Mezić, I., 2006. On the dynamics of molecular conformation. PNAS, 103(20), p.7542-7547.
[12] Sanders, J.A., Verhulst, F. and Murdock, J.A., 2007. Averaging methods in nonlinear dynamical systems (Vol. 59, p.21-44). New York: Springer.
[13] Lichtenberg, A. J. and Lieberman, M. A., 2013. Regular and chaotic dynamics (Vol. 38, p.103). Springer Science and Business Media.
[14] Eisenhower, B. and Mezić, I., 2010. Targeted activation in deterministic and stochastic systems. Phys. Rev. E, 81(2), p.026603.
[15] Eisenhower, B. and Mezic, I., 2007, December. A mechanism for energy transfer leading to conformation change in networked nonlinear systems. 46th IEEE Conference on Decision and Control, IEEE., p.3976-3981.
[16] Eisenhower, B. A., 2009. Targeted escape in large oscillator networks. University of California, Santa Barbara.
[17] Du Toit, P., Mezić, I. and Marsden, J., 2009. Coupled oscillator models with no scale separation. PHYSICA D: Nonlinear Phenomena, 238(5), p.490-501.
[18] Koon, W. S., Owhadi, H., Tao, M. and Yanao, T., 2013. Control of a model of DNA division via parametric resonance. CHAOS, 23(1), p.013117.
[19] Manevitch, L. I., and Smirnov, V.V., 2010. Limiting phase trajectories and the origin of energy localization in nonlinear oscillatory chains. Phys. Rev. E 82(3): p.036602.
[20] Sinden, R. R., 2012. DNA structure and function. (p.13-14). New York: Academic Press.
[21] Chapra, S. C. and Canale, R. P., 2010. (p.709) Numerical methods for engineers. Boston: McGraw-Hill Higher Education.,
[22] Goldstein, H., Poole, C., and Safko, J., 2002. (p.134-174) Classical mechanics.
[23] Delarue, M., Sanejouand, Y. H., 2002. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model., J. Mol. Biol., 320(5), p.1011-1024.
[24] Leonid, M. and Arkadiy, M., 2005. The mechanics of nonlinear systems with internal resonances.(p.190). London: Imperial College Press.

[25] Wiggins, S., 1990. Introduction to applied nonlinear dynamical systems and chaos (p.31-32, Vol. 2). New York: Springer-Verlag.

[26] Neishtadt, A., 1996. Scattering by resonances. Celest. Mech. Dyn. Astron, 65(1-2), 1-20.

[27] He, J. H., 2002. Modified Lindstedt-Poincare methods for some strongly nonlinear oscillations: Part I: expansion of a constant. International Journal of Non-Linear Mechanics, 37(2), 309-314.

[28] Fidiani, E. 2016, March. Modeling of diatomic molecule using the Morse potential and the Verlet algorithm. AIP Conference Proceedings, AIP Publishing, 1719(1), p.030001.

[29] Robinson, B. H., Forgacs, G., Dalton, L. R. and Frisch, H. L., 1980. A simple model for internal motion of DNA based upon EPR studies in the slow motion region. J. Chem. Phys. B, 73(9), p. 4688-4692.

[30] Kvlseth, T., 2016. On the measurement of randomness (uncertainty): A more informative entropy. Entropy, 18(5), p.159.

[31] Craik, A. D., 2005. Prehistory of Faà di Bruno’s formula. Am. Math. Mon., 112(2), p.119-130.

[32] Chou, K. C., 1984. Low-frequency vibrations of DNA molecules. Biochem. J., 221(1), p. 27-31.

[33] Yakushevich, L. V., 2006. Nonlinear physics of DNA (p.44), John Wiley & Sons.

[34] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C., 2009. Introduction to algorithms. MIT press.

[35] Calladine, C. R. and Drew, H., 1997. Understanding DNA: the molecule and how it works (p.107), Academic press.

[36] Marantan, A. and Mahadevan, L., 2018. Mechanics and statistics of the worm-like chain. Am. J. Phys, 86(2), 86-94.

[37] Matsumoto, A., and Olson, W. K. (2002). Sequence-dependent motions of DNA: a normal mode analysis at the base-pair level. Biophys. J., 83(1), 22-41.