Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Estimation of Nonlinear Sequence-dependent Constitutive Law for DNA Molecules

Harish J. Palanthandalam-Madapusi and Sachin Goyal

Abstract— **Long length-scale structural deformations of DNA play a central role in many biological processes including gene expression. While all-atom based Molecular Dynamics (MD) tools fall short to simulate the long-length and time scale mechanics of DNA, continuum rod model of DNA has emerged as a viable tool. The continuum rod model predictions are however very sensitive to the constitutive law (material properties) of the molecule which, in turn, vary along the molecule's length according to its base-pair sequence. Identification of the sequence-dependent constitutive law from experimental data and feasible all-atom MD simulations remains a significant challenge. The primary goal of the paper is to formulate the general identification problem in a form that is amenable to system identification/estimation methods. To this end, we suggest several simplifications to the continuum rod model of DNA. Finally, we suggest a possible approach to solving the identification problem for a simplified case with a nonlinear constitutive law. A secondary goal of this paper is to bring this important problem to the attention of the system identification community.**

I. Introduction

DNA is a long chain biopolymer molecule that is a central substance in the working of all life [1]. Located within the nucleus of our cells, DNA contains the coded (genetic) information needed to synthesize proteins and thus sustain life. Replication and segregation of DNA enable the transfer of this genetic information from one cellular generation to the next. These biological functions are significantly influenced by the structural deformation of the molecule which is tied to its chemical structure, the base-pair sequence. How these structural deformations originate from the chemical make-up of the molecule is an open and active area of research. To elucidate this connection, we begin by summarizing the basic chemistry and structure of DNA, the multiple length-scales involved, and the major biological functions that DNA performs.

Figure 1 illustrates a DNA molecule on three different length scales as reproduced from several sources [1-3]. The smallest length scale (far left) shows a segment of the familiar 'double-helix' which has a diameter of approximately 2 nanometers (nm). One complete helical turn is depicted here and this extends over a length of approximately 3.6 nm. The double helices, which wind like the supports of a spiral staircase, are composed of two polynucleotide chains which in turn are made up of four different nucleotides. Each nucleotide is made from a five-carbon sugar to which one or more phosphate groups and a nitrogen containing base are attached. There are four types of bases that include adenine (A), guanine (G), cytosine (C) and thymine (T). The four bases bond in only two, complementary pairs, namely A with T and C with G. The sugar-phosphate groups of the nucleotides are covalently linked into long chains (highlighted in orange) that form the backbone of DNA. Pairing of the two polynucleotide strands is achieved by hydrogen bonding between the nucleotide bases (highlighted in blue) that fill the small voids between the single DNA strands. It is this linear sequence of base-pairs that constitutes the genetic code. This chemical structure and the rules for 'base-paring' follow from the seminal discoveries of Franklin and Gosling [4] and Watson and Crick [5]. There are approximately 10.5 base-pairs in one helical turn for the common "B" form of DNA which also forms a right-handed helix as depicted in figure 1.

On an intermediate spatial scale (middle of figure 1), the double helix appears as a solid "strand" of DNA that might extend over tens to hundreds of helical turns (approximately tens to hundreds of nanometers). This is the approximate length scale of a 'gene' which is a portion of a DNA strand (i.e. a specific base-pair sequence) that controls a discrete hereditary characteristic. The base-pair sequence within a gene constitutes a chemical code for the production of a specific protein elsewhere within the cell. The major biological function of DNA is to store these chemical codes and to make them available for protein production through a process known as transcription. In addition, the same chemical codes are passed from one cellular generation to the next through a process known as replication. These biological processes are strongly influenced by the structure of the molecule on even longer length scales.

The longer-length scale structures of DNA are illustrated to the far right in figure 1. Here, the long DNA strand may contain thousands to millions of base-pairs and resemble a highly curved and twisted filament with lengths ranging from micron to millimeter scales. The long-length scale curving/twisting of this strand is called supercoiling and two generic types of supercoils are illustrated to the far right of figure 1. One type, referred to as an interwound supercoil (or plectoneme), leads to an interwoven structure where the strand wraps upon itself with many sites of apparent 'self-contact'. By contrast, a solenoidal supercoil possesses no self-contact and resembles a coiled spring or telephone cable. With the aid of proteins, DNA must supercoil for several key reasons. First, supercoiling provides an organized

H. J. Palanthandalam-Madapusi is with the Department of Mechanical and Aerospace Engineering, Syracuse University, Syracuse, NY 13244, hjpalant@syr.edu

S. Goyal is with the Theoretical and Applied Mechanics Department, Cornell University, Cornell, NY 14853, sgoyal@cornell.edu

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

means to compact these very long molecules (by as much as 10^5) enabling them to fit within the small confines of the cell nucleus. An unorganized compaction would hopelessly tangle the strand and render it useless as a medium for storing the coded information. Second, supercoiling plays an important role in transcription and replication. For instance, the formation of simple loops of DNA on long-length scales is known to regulate the transcription of certain genes, for example refer to Schleif [6], Semsey et al. [7] and other citations in Goyal et al. [8].

The long-length scale looping of DNA is dominated by the bending and torsion of the molecule. Many 'coarse-grain' models of DNA have emerged as efficient computational tools to simulate these large (nonlinear) deformations as discussed in Goyal [9]. The coarse-grain models can also be combined with all-atom molecular dynamics (MD) simulations by multi-scale techniques to capture the fine-length scale descriptions of the molecule's conformation. One of the most computationally efficient coarse-grain models for simulating DNA deformations is an elastic rod. The use of rod theory is reasonably well-established in the literature on DNA modeling as reviewed by Schlick [10] and Olson [11]. Rod models approximate DNA as a continuum with prescribed elasticity properties which ultimately vary with length according to the local base-pair sequence of the molecule. In particular, the stiffness variation may arise due to the fact that G≡C pairs possess three hydrogen-bonds while A=T pairs possess only two hydrogen bonds. Thus, one may expect G≡C rich regions of DNA to be stiffer than A=T rich regions. Moreover, the base-pairs in general do not stack in a straight line. Their stacking gives rise to a sequence-dependent 'intrinsic curvature' to the molecule that also affects its structural behavior.

A detailed review of rod models of DNA is provided in Goyal et al. [9, 12] which often begin by assuming homogeneous and isotropic elasticity. The use of an isotropic (circular) rod to represent the structure of the double helix is specifically addressed by Maddocks and co-workers [13, 14] who conclude that bending anisotropy at the base-pair scale quickly averages to an effective isotropic rod on long-length scales due to the high intrinsic twist (10.5 base-pairs/per helical turn) of the double-helix. The majority of rod models also employ equilibrium formulations and hence can not capture dynamic transitions as described in Goyal et al. [9, 15] where the influences of non-homogenous stiffness and intrinsic curvature are also studied.

In this paper, we use the continuum rod model equations developed in [9, 15] to formulate the problem of estimating the constitutive law governing DNA. Since both static and dynamic structural deformations of DNA molecules play a crucial role in biological activity, the continuum rod model is a set of partial differential equations with both time and space as the independent variables. Therefore, we make suitable assumptions to simplify the continuum rod model and then cast it in state-space form, thus enabling us to use techniques from system identification and Estimation literature.

Furthermore, by considering a simplified case of homogeneous DNA, we show that the problem of estimating the nonlinear constitutive law can be treated as the problem of estimating unknown inputs to a state-space model. Thus by using input-reconstructions methods [45-48], we demonstrate the viability of this approach for a simplified case.

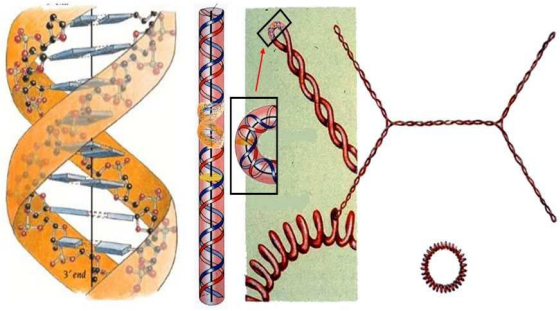

Fig. 1. Three different length scales of the DNA molecule. The smallest scale (left) shows double-helix structure (sugar-phosphate chains and base-pairs). Intermediate scale (middle) shows how multiple double-helices form a continuous molecule of double-stranded DNA (ds-DNA). Largest scale (right) shows how the molecule curves and twists in forming supercoils including the idealized plectonemic and solenoidal supercoils depicted here. (Courtesy: Branden and Tooze, 1999, From Introduction to Protein Structure by Carl Branden & John Tooze (reproduced by permission of Garland Science/Taylor & Francis, LLC) [2] and Lehninger et al. (copied with permission from W.H. Freeman) [3]).

II. Problem Statement and Broader Impact

The elastic properties of DNA follow from its interatomic interactions and must define a base-pair sequence-dependent constitutive law for the bulk (continuum) model. The basic form of this constitutive law and its sequence-dependent mapping is an open and highly active area of research. The continuum rod behavior has been shown to be highly sensitive to the variations in the constitutive law [8, 15, 38]. Thus the experimentally observable behaviors which are highly sensitive to the constitutive law can be leveraged to calibrate the the constitutive law using system identification/ inverse modeling techniques. Some useful calibration experiments are cyclization (DNA loop closure) experiments as described in Goyal et al. [38] and Manning et al. [16]. These experiments give an estimate of the elastic energy of deformation of the DNA.

In addition, the fact that the constitutive law conspires from the interatomic interactions, motivates the concept of mapping the material law from all atom details using Molecular Dynamics (MD) simulations. MD tools can capture ab initio calculations by accurately modeling interatomic potentials, but their viability is limited to only short length and time scale simulations. Nevertheless, MD simulations could also serve as data for calibrating the constitutive law for continuum (bulk) models. In particular, MD simulations solve for positions of each atom and forces on each atom under various loading conditions which can in principle be mapped into the stress distribution and geometry of the continuum model. This mapped data then could perhaps be

used to calibrate the constitutive law. The utility of this approach extends beyond DNA molecules to the whole area of material science where now-a-days researchers are trying to map constitutive laws of various materials form all-atom MD simulations and also trying to design detailed atomic structures to achieve desired material behavior.

We begin by summarizing the computational rod model of Goyal et al. [9, 15] in Section 2.

III. Computational Rod Model

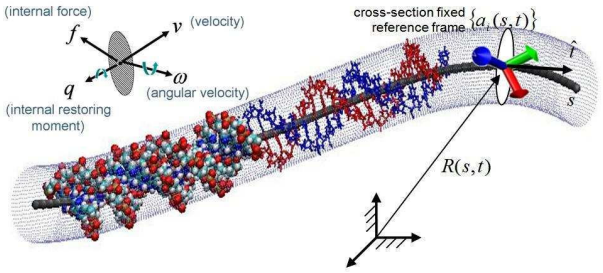

Fig. 2. Dynamical rod model of double stranded DNA on long-length scales. Helical axis of duplex defines the rod centerline which forms a three-dimensional space curve located by $R(s,t)$.

The dynamics of a flexible rod can be represented by the rigid-body dynamics of a local, cross-section fixed frame as illustrated in figure 2. The orientation of the body-fixed frame $\{\mathbf{a_i}(s,t)\}$ varies with both position s and time t. The deformation of the rod centerline, which represents the helical axis of the molecule, is determined by the curvature and twist vector $\kappa(s,t)$. This vector defines the rotation of the body-fixed frame with length s. In a stress-free state, DNA conforms to its natural geometry represented by $\kappa_{\mathbf{0}}(s)$, which is also called its intrinsic curvature.

The rigid-body dynamics of the body-fixed frame is described by its translational velocity vector $\mathbf{v}(s,t)$ and its angular velocity vector $\omega(s,t)$. The stress distribution across the cross-section results in a net internal (tensile and shear) force vector $\mathbf{f}(s,t)$ and (bending and torsional) moment vector $\mathbf{q}(s,t)$. As described in Goyal et al. [15], the governing equations for the rod dynamics can be derived from the first principle by looking at a free body diagram of an infinitesimal rod segment and they are written as

$$\frac{\partial \mathbf{v}}{\partial s} + \kappa \times \mathbf{v} = \omega \times \hat{\mathbf{t}} \quad \text{(III.1)}$$

$$\frac{\partial \omega}{\partial s} + \kappa \times \omega = \frac{\partial \kappa}{\partial t} \quad \text{(III.2)}$$

$$\frac{\partial \mathbf{q}}{\partial s} + \kappa \times \mathbf{q} = \mathbf{I}\frac{\partial \omega}{\partial t} + \omega \times \mathbf{I}\omega + \mathbf{f} \times \hat{\mathbf{t}} - \mathbf{Q} \quad \text{(III.3)}$$

$$\frac{\partial \mathbf{f}}{\partial s} + \kappa \times \mathbf{f} = m\left(\frac{\partial \mathbf{v}}{\partial t} + \omega \times \mathbf{v}\right) - \mathbf{F} \quad \text{(III.4)}$$

where $\hat{\mathbf{t}}(s,t)$ denotes the unit tangent vector, $m(s)$ denotes the mass of the rod per unit arc length and $\mathbf{I}(s)$ denotes the tensor of principal mass moments of inertia per unit arc length. The computational rod model [15] in general can also capture interactions of the rod with the external environment via the external forces $\mathbf{F}(s,t)$ and moments $\mathbf{Q}(s,t)$ distributed per unit arc length. The partial derivatives are all relative to the body-fixed frame $\{\mathbf{a_i}(s,t)\}$. Equation (III.1) represents inextensibility and unshearability constraints; Eq. (III.2) is a constraint on the curvature/twist and angular velocity vectors that ensures continuity of cross-section orientation; Eqs. (III.4,III.3) are the Newton-Euler equations for an infinitesimal rod element.

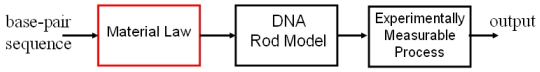

Fig. 3. Block diagrams illustrating the various components in DNA experiments and the estimation procedure. The block diagram on top illustrates an DNA experiment, while the block-diagram on the bottom illustrates the estimation of the DNA material law.

IV. Sequence-dependent Constitutive Law

The internal moment $\mathbf{q}(s,t)$ is a restoring moment that results from the deformation $(\kappa(s,t) - \kappa_{\mathbf{0}}(s))$. The relationship between the deformation $(\kappa(s,t) - \kappa_{\mathbf{0}}(s))$ and the restoring moment $\mathbf{q}(s,t)$ is governed by interatomic interactions/ material properties and is called the constitutive law. A conceivable form of the nonlinear constitutive law can be written as:

$$\kappa(s,t) - \kappa_{\mathbf{0}}(s) = \mathbf{g}(\mathbf{q}(s,t), s) \quad \text{(IV.1)}$$

Observe that the non-homogenous intrinsic curvature $\kappa_{\mathbf{0}}(s)$ and the explicit dependence of the function $\mathbf{g}$ on s captures the non-homogenous nature of the constitutive law. The non-homogenous nature (dependence on s) of the constitutive law maps from the varying sequence of base-pairs along the length of the DNA. The various components of the continuum rod model are seen in Figure 3. Many experimental and computational efforts are underway to determine the sequence-dependence of the constitutive law for DNA; see, for example [17, 18]. All of the efforts, so far assume a linear approximation of the constitutive law. But it should also be noted that the assumption of linear material law has been recently questioned and debated [19-21]; and therefore even the identification of the functional form is an open area of research. Nevertheless, even the linear approximation of the material law is not yet well characterized.

V. Problem Formulation

To estimate the constitutive law, we first note that most DNA experiments measure steady-state characteristics of the molecules [16, 38]. Therefore we first consider spatial variations alone by freezing time. Thus equations (1)-(4) can be simplified by eliminating the time-variable. Note that in the steady-state configuration $\mathbf{v} = \omega = 0$, and partial derivative with respect to t are zero. Hence (1)-(4) reduce

to the following two ordinary differential equations in terms of the spatial variable s.

$$\frac{d\mathbf{q}}{ds} + \kappa \times \mathbf{q} = \mathbf{f} \times \hat{\mathbf{t}}. \tag{V.1}$$

$$\frac{d\mathbf{f}}{ds} + \kappa \times \mathbf{f} = 0, \tag{V.2}$$

With the independent variable as s, we can now write these equations in the state-space form by defining the state vector $x(s)$ to be $x(s) \in \begin{bmatrix} \mathbf{q}(s) & \mathbf{f}(s) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^6$.

Typically measurements in DNA experiments are bulk properties of the molecules such as the elastic strain energy which can be expressed as

$$U = \int_0^L \int_{\kappa_0(s)}^{\kappa(s)} \mathbf{q}(s) \cdot dk \; ds. \tag{V.3}$$

Note that (V.3) is a function of the states at several values of the independent variable s. Therefore, this does not fit into the traditional state-space formulation where measurements are available for all values of the independent variable.

Since this is a sizeable challenge, we first consider a simpler case and demonstrate that the above problem can in-principle be solved.

A. DNA Cantilever Experiment

We first consider a simple fictitious cantilever experiment with a DNA molecule clamped at one end and a constant shear force applied at the free end. We assume that the DNA molecule is uniform and thus has intrinsic curvature $\kappa_0(s) \equiv 0$. Moreover, we assume that the material behavior of DNA is decoupled in the principal directions of bending and torsion. Therefore, by choosing the body-fixed frame along these principal directions, the vector material law (IV.1) is decoupled into the following scalar material laws

$$\kappa_1(s) = g_1(\mathrm{q}_1(s)), \tag{V.4}$$

$$\kappa_2(s) = g_2(\mathrm{q}_2(s)), \tag{V.5}$$

$$\kappa_3(s) = g_3(\mathrm{q}_3(s)). \tag{V.6}$$

Finally, we assume that measurements are available at all values s.

Let the first two axes in the body-fixed frame $a_1(s)$ and $a_2(s)$ correspond to the principal bending axes and the third axis $a_3(s)$ correspond to the principal torsion axis. Now, if the applied shear force is acting along axis a_1, then the DNA molecule bends in-plane about axis a_2. Thus the first and third components of $\kappa(s)$ and $\mathbf{q}(s)$ along with the second component of $\mathbf{f}$ are zero. Thus the elastic rod model reduces to the following three equations

$$\frac{dq_2}{ds} = -\mathrm{f}_1(s), \tag{V.7}$$

$$\frac{df_1}{ds} = -\mathrm{f}_3(s)\kappa_2(s), \tag{V.8}$$

$$\frac{df_3}{ds} = \mathrm{f}_1(s)\kappa_2(s), \tag{V.9}$$

with the material-law equation (V.5).

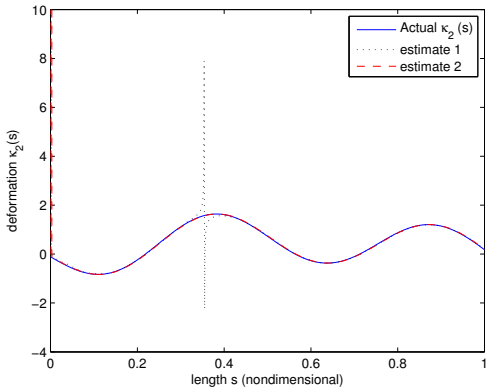

Fig. 4. Actual and estimated deformation $\kappa_2(s)$ for a cantilever DNA molecule with in-plane deformations.

By treating the free end as the origin and prescribing the boundary conditions at the free end, the above problem can be treated as an initial value problem. Moreover, as seen from (V.7) - (V.9), when the material law is unknown $\kappa_2(s)$, which is an unknown function of q_2, can be treated as an unknown input to the elastic rod model.

Note that (V.7) - (V.9) are a set nonlinear state space equations with states $x(s) = \begin{bmatrix} \mathrm{q}_2(s) & \mathrm{f}_1(s) & \mathrm{f}_3(s) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^3$. Next, since $\kappa_2(s)$ is an unknown function of $\mathrm{q}_2(s)$, we cannot use standard state estimation techniques as the dynamics itself is unknown.

Alternatively, we treat $\kappa_2(s)$ as an input to the system. Therefore, we can use unknown-input estimation such as the unscented unbiased minimum variance filter [46-48].

We assume that measurements of $\mathbf{f}(s)$ and $\mathbf{q}(s)$ are available. We then use the unscented unbiased minimum-variance (UUMV) filter [47] with s as the independent variable, to estimate $\kappa_2(s)$. As a second step we use standard least-squares fitting to estimate the functional relationship between $\kappa_2(s)$ and $\mathrm{q}_2(s)$.

VI. Results

For simulation we choose $g_2(\cdot)$ to be an arctangent function and set the initial conditions to be $x(0) = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}$, where all numbers are dimensionless.

As $\kappa_2(s)$ appears twice in the state equations (V.7) - (V.8), there are two ways to estimate $\kappa_2(s)$ using the UUMV filter. Figure 4 shows actual $\kappa_2(s)$ and two independent estimates. Although, the independent variable is the length coordinate s, it is treated as an initial value problem, and as seen from the Figure 4, the estimate converges to the actual value in the first few steps. Also, around $s = 0.35$, a spike is seen in one of the estimates because the value of $\mathrm{f}_3(s)$, which is the coefficient of $\kappa_2(s)$ in (V.8), crosses zero. Thus there is a loss of input observability and hence the error in the estimate.

Once $\kappa_2(s)$ is estimated, the function $g_2(\cdot)$ can be estimated by treating the unknown function as an expansion of sinusoidal basis functions and using a standard least-squares to fit the unknown coefficients of the basis function

expansion. Figure 5 shows the actual and estimate of the constitutive law $g_2(\cdot)$, when an arctangent function was used for simulations. Figure 6 shows the actual and estimate of the constitutive law $g_2(\cdot)$, when a saturation function is used for simulations.

Finally, we note that by running three separate experiments with excitation along one principal axis at a time, we can use the same procedure above to estimate all three material laws (V.4) - (V.6), and thus the complete nonlinear material law. The estimated material law can then be used to predict deformations for any general loading configuration.

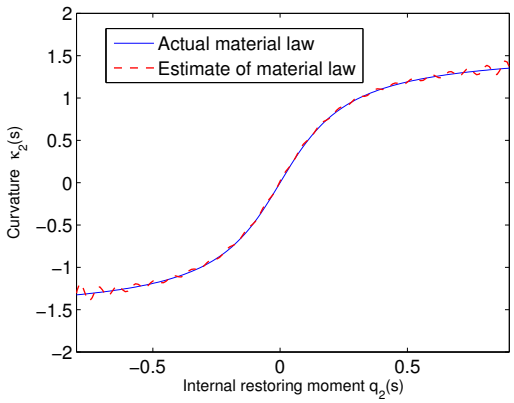

Fig. 5. Actual and estimated arctangent material law for a DNA Molecule with the in-plane cantilever deformation.

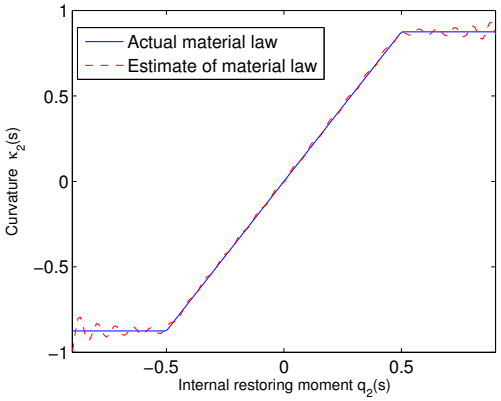

Fig. 6. Actual and estimated saturation material law for a cantilever beam with the in-plane cantilever deformation.

VII. Conclusion

In this paper, using a continuum rod model for DNA, we considered identification of the sequence-dependent constitutive law from experimental data and feasible all-atom MD simulations. We reformulated the problem in a form that is amenable to system identification/estimation methods. Furthermore, by making suitable assumptions, we derived several simplifications to the continuum rod model of DNA. Finally, we considered a possible approach to estimating the nonlinear constitutive law for a simplified case.

Acknowledgment

The authors gratefully acknowledge support from Cornell University and Syracuse University.

VIII. References

[1] Calladine, C. R., Drew, H. R., Luisi, B. F., and Travers, A. A., 2004, Understanding DNA, the Molecule and How It Works, Elsevier Academic Press, Amsterdam.

[2] Branden, C., and Tooze, J., 1999, Introduction to Protein Structure, Garland Publishing, New York.

[3] Lehninger, A. L., Nelson, D. L., and Cox, M. M., 2005, Lehninger Principles of Biochenistry, W. H. Freeman, New York.

[4] Franklin, R. E., and Gosling, R. G., 1953, "Molecular Configuration in Sodium Thymonucleate," Nature, 171(4356), pp. 740-741.

[5] Watson, J. D., and Crick, F. H. C., 1953, "Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid," Nature, 171(4356), pp. 737-738.

[6] Schleif, R., 1992, "DNA Looping," Annual Review of Biochemistry, 61, pp. 199-223.

[7] Semsey, S., Virnik, K., and Adhya, S., 2005, "A Gamut of Loops: Meandering DNA," Trends in Biochemical Sciences, 30(6), pp. 334-341.

[8] Goyal, S., Lillian, T., Blumberg, S., Meiners, J. C., Meyhofer, E., and Perkins, N. C., 2007, "Intrinsic Curvature of DNA Influences LacR-Mediated Looping," Biophysical Journal, 93(12).

[9] Goyal, S., 2006, "A Dynamic Rod Model to Simulate Mechanics of Cables and DNA," Ph.D. Dissertation in Mechanical Engineering, University of Michigan, Ann Arbor.

[10] Schlick, T., 1995, "Modeling Superhelical DNA - Recent Analytical and Dynamic Approaches," Current Opinion in Structural Biology, 5(2), pp. 245-262.

[11] Olson, W. K., 1996, "Simulating DNA at Low Resolution," Current Opinion in Structural Biology, 6(2), pp. 242-256.

[12] Goyal, S., Lillian, T., Perkins, N. C., and Meyhofer, E., 2005, "Cable Dynamics Applied to Long-Length Scale Mechanics of DNA." CD-ROM Proceedings of Sixth International Symposium of Cable Dynamics. Chaleston, South Carolina.

[13] Kehrbaum, S., and Maddocks, J. H., 2000, "Effective Properties of Elastic Rods with High Intrinsic Twist." 16th IMACS World Congress.

[14] Rey, S., and Maddocks, J. H., 2000, "Buckling of an Elastic Rod with High Intrinsic Twist." 16th IMACS World Congress.

[15] Goyal, S., Perkins, N. C., and Lee, C. L., 2005, "Nonlinear Dynamics and Loop Formation in Kirchhoff Rods with Implications to the Mechanics of DNA and Cables," Journal of Computational Physics, 209(1), pp. 371-389.

[16] Manning, R. S., Maddocks, J. H., and Kahn, J. D., 1996, "A Continuum Rod Model of Sequence-Dependent DNA Structure," Journal of Chemical Physics, 105(13), pp. 5626-5646.

[17] Beveridge, D. L., Barreiro, G., Byun, K. S., Case, D. A., Cheatham, T. E., Dixit, S. B., Giudice, E., Lankas, F., Lavery, R., Maddocks, J. H., Osman, R., Seibert, E., Sklenar, H., Stoll, G., Thayer, K. M., Varnai, P., and Young, M. A., 2004, "Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. I. Research Design and Results on D(C(P)G) Steps," Biophysical Journal, 87(6), pp. 3799-3813.

[18] Dixit, S. B., Beveridge, D. L., Case, D. A., Cheatham, T. E., Giudice, E., Lankas, F., Lavery, R., Maddocks, J. H., Osman, R., Sklenar, H., Thayer, K. M., and Varnai, P., 2005, "Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. II: Sequence Context Effects on the Dynamical Structures of the 10 Unique Dinucleotide Steps," Biophysical Journal, 89(6), pp. 3721-3740.

[19] Cloutier, T. E., and Widom, J., 2004, "Spontaneous Sharp Bending of Double-Stranded DNA," Molecular Cell, 14(3), pp. 355-362.

[20] Du, Q., Smith, C., Shiffeldrim, N., Vologodskaia, M., and Vologodskii, A., 2005, "Cyclization of Short DNA Fragments and Bending Fluctuations of the Double Helix," Proceedings of the National Academy of Sciences of the United States of America, 102(15), pp. 5397-5402.

[21] Wiggins, P. A., Phillips, R., and Nelson, P. C., 2005, "Exact Theory of Kinkable Elastic Polymers," Physical Review E, 71(2).

[22] Mehta, R. A., and Kahn, J. D., 1999, "Designed Hyperstable Lac Repressor-DNA Loop Topologies Suggest Alternative Loop Geometries," Journal of Molecular Biology, 294(1), pp. 67-77.

[23] Morgan, M. A., Okamoto, K., Kahn, J. D., and English, D. S., 2005, "Single-Molecule Spectroscopic Determination of Lac Repressor-DNA Loop Conformation," Biophysical Journal, 89(4), pp. 2588-2596.

[24] Chung, J., and Hulbert, G. M., 1993, "A Time Integration Algorithm for Structural Dynamics with Improved Numerical Dissipation - the Generalized-Alpha Method," Journal of Applied Mechanics-Transactions of The ASME, 60(2), pp. 371-375.

[25] Gobat, J. I., and Grosenbaugh, M. A., 2001, "Application of the Generalized-Alpha Method to the Time Integration of the Cable Dynamics Equations," Computer Methods in Applied Mechanics and Engineering, 190(37-38), pp. 4817-4829.

[26] Gobat, J. I., Grosenbaugh, M. A., and Triantafyllou, M. S., 2002, "Generalized-Alpha Time Integration Solutions for Hanging Chain Dynamics," Journal of Engineering Mechanics-ASCE, 128(6), pp. 677-687.

[27] Shimada, J., and Yamakawa, H., 1984, "Ring-Closure Probabilities for Twisted Wormlike Chains - Application to DNA," Macromolecules, 17(4), pp. 689-698.

[28] Zhang, Y. L., and Crothers, D. M., 2003, "Statistical Mechanics of Sequence-Dependent Circular DNA and its Application for DNA Cyclization," Biophysical Journal, 84(1), pp. 136-153.

[29] Popov, Y. O., and Tkachenko, A. V., 2005, "Effects of Kinks on DNA Elasticity," Physical Review E, 71(5).

[30] Wilson, D. P., Lillian, T. D., Goyal, S., Perkins, N. C., Tkachenko, A., and Meiners, J. C., 2007, "Modeling the Entropic Cost of DNA Looping," Biophysical Journal, pp. 540A-540A.

[31] Howard, J., 2001, Mechanics of Motor Proteins and the Cytoskeleton, Sinauer Associates, Sunderland.

[32] Hagerman, P. J., 1988, "Flexibility of DNA," Annual Review of Biophysics and Biophysical Chemistry, 17, pp. 265-286.

[33] Gabrielian, A., and Pongor, S., 1996, "Correlation of Intrinsic DNA Curvature with DNA Property Periodicity," FEBS Letters, 393(1), pp. 65-68.

[34] Weaver, R., 2007, Molecular Biology, McGraw-Hill, Boston.

[35] Baumann, C. G., Smith, S. B., Bloomfield, V. A., and Bustamante, C., 1997, "Ionic Effects on the Elasticity of Single DNA Molecules," Proceedings of the National Academy of Sciences of the United States of America, 94(12), pp. 6185-6190.

[36] Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V., 1996, "The Elasticity of a Single Supercoiled DNA Molecule," Science, 271(5257), pp. 1835-1837.

[37] Balaeff, A., Mahadevan, L., and Schulten, K., 2006, "Modeling DNA Loops Using the Theory of Elasticity," Physical Review E, 73(3).

[38] Goyal, S., Perkins, N.C., and Meiners, J.C., 2008, "Resolving the Sequence-Dependent Stiffness of DNA Using Cyclization Experiments and a Computational Rod Model," Journal of Computational and Nonlinear Dynamics, 3, 011003.

[39] Kalman, R. E., 1960, "A new approach to linear filtering and prediction problems," Transactions of the ASME Journal of Basic Engineering, 82: 3545.

[40] Daum, F. E., 2005, "Nonlinear filters: Beyond the Kalman filter," IEEE Aerospace and Electronics Systems Magazine, 20(8), 5769.

[41] Julier, S. J. and Uhlmann, J. K., 2004, "Unscented filtering and nonlinear estimation," Proceedings of the IEEE, 92: 401422.

[42] Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., 2000, "A new method for the nonlinear transformation of means and covariances in filters and estimators," IEEE Transactions on Automatic Control, 45(3): 477 482.

[43] Haykin, S. (Ed.), 2001, "Kalman Filtering and Neural Networks." New York NY, USA: Wiley Publishing.

[44] M. K. Sain and J. L. Massey. "Invertibility of linear time-invariant dynamical systems." IEEE Trans. on Automatic Contr., AC-14(2):141149, April 1969.

[45] P. K. Kitanidis. "Unbiased Minimum-variance Linear State Estimation." Automatica, 23(6):775578, 1987.

[46] H. J. Palanthandalam-Madapusi and D. S. Benrstein. "Unbiased minimum-variance filtering for input recon- struction." In Proc. of Amer. Contr. Conf., New York, NY, July 2007.

[47] H. J. Palanthandalam-Madapusi, Anouck Girard and D. S. Benrstein. "Wind-field Reconstruction Using Flight Data." Proc. of Amer. Contr. Conf., June 2008, Seattle, WA, pp. 1863 – 1868.

[48] S. Gillijns and B. De Moor. "Unbiased minimum-variance input and state estimation for linear discrete-time systems." Automatica, Volume 43, Issue 1, January 2007, Pages 111-116.