

Review

Tight-Binding Modeling of Nucleic Acid Sequences: Interplay between Various Types of Order or Disorder and Charge Transport

Konstantinos Lambropoulos *  and Constantinos Simserides * 

Department of Physics, National and Kapodistrian University of Athens, Panepistimiopolis, Zografos, 15784 Athens, Greece

* Correspondence: klambro@phys.uoa.gr (K.L.); csimseri@phys.uoa.gr (C.S.)

Received: 30 June 2019; Accepted: 11 July 2019; Published: 1 August 2019



Abstract: This review is devoted to tight-binding (TB) modeling of nucleic acid sequences like DNA and RNA. It addresses how various types of order (periodic, quasiperiodic, fractal) or disorder (diagonal, non-diagonal, random, methylation et cetera) affect charge transport. We include an introduction to TB and a discussion of its various submodels [wire, ladder, extended ladder, fishbone (wire), fishbone ladder] and of the process of renormalization. We proceed to a discussion of aperiodicity, quasicrystals and the mathematics of aperiodic substitutional sequences: primitive substitutions, Perron–Frobenius eigenvalue, induced substitutions, and Pisot property. We discuss the energy structure of nucleic acid wires, the coupling to the leads, the transmission coefficients and the current–voltage curves. We also summarize efforts aiming to examine the potentiality to utilize the charge transport characteristics of nucleic acids as a tool to probe several diseases or disorders.

Keywords: nucleic acids; aperiodic; quasiperiodic; fractal; order; disorder; energy structure; charge transport

1. Introduction

Nucleic acids are polymeric macromolecules consisting of units that are called nucleotides. The term nucleic acids is the overall name of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA's nucleotide sequence carries the genetic instructions for the development, function, growth and reproduction of living organisms and several viruses. Although RNA's primary role is to carry out the instructions encoded in DNA for protein synthesis, it also acts like a catalyst of biochemical reactions, while it is the genetic material of many viruses.

For more than sixty years now, the double-stranded structure of DNA has been known [1]. The nucleotides of each strand are composed of one of four planar, aromatic, nitrogenous bases, i.e., guanine (G), cytosine (C), adenine (A) or thymine (T), a pentose sugar (deoxyribose), and a phosphate group. Covalent, phosphodiester bonds between pentoses and phosphate groups of adjacent nucleotides form an alternating sugar-phosphate backbone. The purines (G or A) of a nucleotide belonging to a strand are joined together with the pyrimidines of the other strand (C or T, respectively) via (three or two, respectively) hydrogen bonds, forming the double helix structure. This specificity in the way bases match ensures that G is always bonded with C, and A is always bonded with T. Pairing between non-complementary bases results in mutations that can be detrimental to the development of an organism. In RNA, deoxyribose (whose 2-carbon is bonded with a hydrogen) is replaced by ribose (whose 2-carbon is bonded with a hydroxyl group), and T is replaced by uracil (U). Furthermore, RNA molecules are single-stranded; however, some viruses possess double-stranded RNA (other viruses can contain even single-stranded DNA).

Although the study of nucleic acids is mainly associated with molecular biology and genetics, today, a broad interdisciplinary community is interested in biological systems, such as nucleic acids and analogues. The base-pair stack of nucleic acids creates a nearly one-dimensional π -stack that allows charge carrier movement, i.e., charge transfer and transport. Let us distinguish between these two terms: transfer means that a carrier, created or injected at a specific nucleotide, moves to a more favorable location, while transport implies the use of electrodes and the application of external voltage between them. Charge transfer is the basis of many biological processes, e.g., in various proteins [2] including metalloproteins [3], and enzymes [4], with medical and bioengineering applications [5,6], while it plays a role in DNA damage and repair [7–9]. Charge transport might be an indicator to distinguish pathogenic from non-pathogenic mutations at an early stage [10].

From a physicist's point of view, the charge transfer and transport properties of nucleic acids are studied in order to obtain a deeper understanding of their biological functions as well as for potential applications, such as nanosensors, nanocircuits or molecular wires, due to their high yield synthesis, near-unity purification, and nanoscale self-organization [11–13]. There are many external (aqueousness, presence of counterions, extraction process, electrodes, contacts, purity, substrate), and internal (such as the base-pair sequence and geometry) factors that affect carrier motion along nucleic acids. Both *ab initio* calculations [14–22] and model Hamiltonians [23–34] have been used to theoretically explore the variety of experimental results that predict electrical behavior ranging from metallic to insulating, as well as the underlying mechanisms.

It has become evident that the influence of various types of order or disorder plays a central role in the energy structure and the charge transport properties of nucleic acids. This interplay between various types of order or disorder and charge transport is addressed in this brief review. This is done in the context of one of the most widely applied theoretical methods, i.e., with Tight-Binding (TB), because of its simplicity and low computational cost.

The rest of this review is organized as follows. In Section 2, we present the TB formulation and explain some of its most common variations applied in the literature for the study of nucleic acids. In Section 3, we overview several aperiodic substitutional sequences that highlight the influence of disorder in the properties of nucleic acids. In Section 4, we discuss the energy spectra of ordered and disordered nucleic acid sequences. In Section 5, we focus on electron transmission and on the influence of coupling the examined systems with leads. Section 6 is dedicated to the influence of various types of order or disorder on the current–voltage ($I - V$) curves of nucleic acids. Finally, in Section 7, we make some concluding remarks.

2. Tight-Binding and Its Application in Nucleic Acids

TB is an approximate method widely used in condensed matter physics to determine the electronic structure of a solid through the expansion of its wavefunction as a superposition of the wavefunctions corresponding to the isolated moieties located at each lattice site [35]. As the name of the model suggests, the main hypothesis in TB is that the system's orbitals are tightly bound at the sites at which they belong, so that the overlap with neighboring orbitals is small. Hence, the electronic wavefunction of the moiety that occupies a lattice site is rather similar to the orbital of the free moiety. As a result, the corresponding energy of the electron will be rather close to the (negative) ionization energy of the free moiety due to the weak interaction with its neighbors. This picture is applicable at the bands formed by the core electrons of metals, the valence and conduction bands of insulators and semiconductors, as well as the valence and conduction bands arising from localized d or f states (e.g., in transition metals and rare earths).

Today, several decades after its introduction [36], TB has evolved into a fast and efficient approach, employable to numerous problems regarding the electronic structure and properties of matter, requiring various degrees of accuracy [37,38]. Its main advantages include its intuitive simplicity, the ability it gives to obtain analytic results in several cases, and its low computational cost [39]. The latter makes TB applicable to large systems, currently unreachable by the more sophisticated

ab initio methods, such as Density Functional Theory (DFT). In contrast to those methods, TB is semi-empirical, in the sense that an external set of parameters is needed in order to perform calculations. These parameters are (a) the on-site energies that correspond to the energy of the electrons that belong to each lattice site, and (b) the hopping (or transfer) integrals that correspond to the coupling of orbitals which belong to neighboring sites.

Over the last few decades, TB has been widely used to describe, among others, polymers and organic systems. One-dimensional TB models are commonly applied to study the energy structure and thermal, magnetic as well as charge transfer and transport properties of π -conjugated organic systems that are candidates for molecular wires, such as nucleic acids and analogues. Those models have varying degrees of complexity, and each one of them requires a different number of parameters. As far as nucleic acids are concerned, the models employed include, inter alia, the Wire Model (WM), the Ladder Model (LM), the Extended Ladder Model (ELM), the Fishbone Model (FM) and the Fishbone Ladder Model (FLM). Generally, the studied systems consist of N monomers extended at L chains ($L \ll N$, since nucleic acids are approximately one-dimensional). The problem is reduced to the solution of the so-called system of TB equations, which is a system of coupled stationary, algebraic equations or differential equations of first order, equivalent to a discretized form of the time-independent or time-dependent Schrödinger equation. As far as nucleic acids are concerned, the stationary TB system of equations can be compactly written in the matrix form

$$E\vec{\Psi}_n = \epsilon_n \vec{\Psi}_n + \tau_{n-1}^T \vec{\Psi}_{n-1} + \tau_n \vec{\Psi}_{n+1}, \quad (1)$$

for $n = 1, 2, \dots, N$. $\vec{\Psi}_n$ is a vector matrix containing the elements of the wavefunction that correspond to monomer n , i.e., $\vec{\Psi}_n = (\psi_n^1 \ \psi_n^2 \ \dots \ \psi_n^L)^T$, ϵ_n is a symmetric $L \times L$ matrix containing the on-site energies of each site, ϵ_n^l and the hopping integrals $t_{nn'}^{ll'}$ between the sites of the monomer that belong to different chains, and τ_n is a generally non-symmetric $L \times L$ matrix containing the hopping integrals $t_{nn'}^{ll'}$ between each site of a monomer and the neighboring sites of the next monomer. Finally, E is the energy. The situation is schematically presented in Figure 1. From Bloch's theorem, it holds that $\vec{\Psi}_{N+n} = z\vec{\Psi}_n$, where z generally lies in the unit circle ($z = z^* = 1$, for cyclic boundaries, or $z = z^* = 0$ for fixed boundaries). Hence, the solution of the system of Equation (1) can be reduced to the diagonalization of the Hamiltonian matrix, written in block form as

$$\mathbf{H} = \begin{pmatrix} \epsilon_1 & \tau_1 & & & z^* \tau_0^T \\ \tau_1^T & \epsilon_2 & \tau_2 & & \\ & \tau_2^T & \epsilon_3 & \tau_3 & \\ & & \ddots & \ddots & \ddots \\ z\tau_N & & & \tau_{N-1}^T & \epsilon_N \end{pmatrix}. \quad (2)$$

Equivalently, Equation (1) can be written in the form

$$\begin{pmatrix} \vec{\Psi}_{n+1} \\ \vec{\Psi}_n \end{pmatrix} = \begin{pmatrix} \tau_n^{-1}(E - \epsilon_n) & -\tau_n^{-1}\tau_{n-1}^T \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \vec{\Psi}_n \\ \vec{\Psi}_{n-1} \end{pmatrix} = \mathbf{Q}_n(E) \begin{pmatrix} \vec{\Psi}_n \\ \vec{\Psi}_{n-1} \end{pmatrix}, \quad (3)$$

where $\mathbf{Q}_n(E)$ is called the transfer matrix of monomer n , and $\mathbf{1}, \mathbf{0}$ are the unit and zero matrix of order L . The product

$$\mathbf{M}_N(E) = \prod_{n=N}^1 \mathbf{Q}_n(E) \quad (4)$$

defines the global transfer matrix of the system, which satisfies the relation,

$$\mathbf{M}_N(E) \begin{pmatrix} \vec{\Psi}_1 \\ \vec{\Psi}_0 \end{pmatrix} = \begin{pmatrix} \vec{\Psi}_{N+1} \\ \vec{\Psi}_N \end{pmatrix} = z \begin{pmatrix} \vec{\Psi}_1 \\ \vec{\Psi}_0 \end{pmatrix}, \quad (5)$$

and contains all the information about its energetics. In fact, since z is an eigenvalue of the global transfer matrix, with eigenvector $(\bar{\Psi}_1 \ \bar{\Psi}_0)^T$, the whole eigenvector of the Hamiltonian matrix of Equation (2) can be reconstructed via a successive application of Equation (3) [40,41]. Hence, when z is an eigenvalue of $\mathbf{M}_N(E)$, E is an eigenvalue of the system's Hamiltonian. Thus, both methods can be used to determine the energy structure of the system. The form of the matrices in Equation (1) for various TB models is presented in Table 1. Some details on each of these TB models are discussed below.

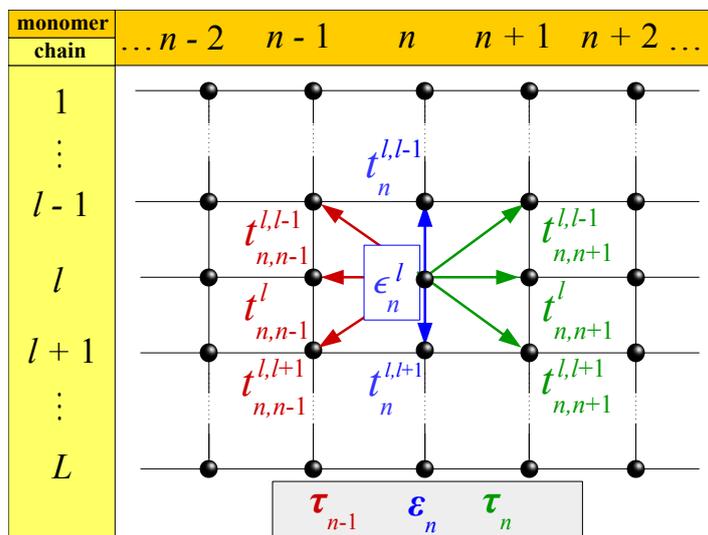


Figure 1. Schematic representation of a TB model consisting of N monomers, extended at L chains. Within the model, we take into account (a) the on-site energies of each site, ϵ_n^l , and the inter-chain hopping integrals, $t_n^{l,l'}$, i.e., between the sites of the monomer (blue), as well as (b) the inter-monomer hopping integrals, $t_{n,n'}^{l,l'}$, i.e., between each site of a monomer and the neighboring sites of the previous (red) and the next (green) monomers. The former are contained in the matrix ϵ_n , while the latter in the matrices τ_{n-1} and τ_n , respectively.

Table 1. Form of the matrices $\bar{\Psi}_n$, ϵ_n , τ_n in the TB system of equations (Equation (1)) for several models used to describe nucleic acids and analogues: the Wire Model (WM), the Ladder Model (LM), the Extended Ladder Model (ELM), the Fishbone Model (FM) and the Fishbone Ladder Model (FLM).

Model	L	$\bar{\Psi}_n$	ϵ_n	τ_n
WM	1	ψ_n	ϵ_n	$t_{n,n+1}$
LM	2	$\begin{pmatrix} \psi_n^1 \\ \psi_n^2 \end{pmatrix}$	$\begin{pmatrix} \epsilon_n^1 & t_n^{1,2} \\ t_n^{2,1} & \epsilon_n^2 \end{pmatrix}$	$\begin{pmatrix} t_{n,n+1}^{1,1} & 0 \\ 0 & t_{n,n+1}^{2,2} \end{pmatrix}$
ELM	2	$\begin{pmatrix} \psi_n^1 \\ \psi_n^2 \end{pmatrix}$	$\begin{pmatrix} \epsilon_n^1 & t_n^{1,2} \\ t_n^{2,1} & \epsilon_n^2 \end{pmatrix}$	$\begin{pmatrix} t_{n,n+1}^{1,1} & t_{n,n+1}^{1,2} \\ t_{n,n+1}^{2,1} & t_{n,n+1}^{2,2} \end{pmatrix}$
FM	3	$\begin{pmatrix} \psi_n^1 \\ \psi_n^2 \\ \psi_n^3 \end{pmatrix}$	$\begin{pmatrix} \epsilon_n^1 & t_n^{1,2} & 0 \\ t_n^{2,1} & \epsilon_n^2 & t_n^{2,3} \\ 0 & t_n^{3,2} & \epsilon_n^3 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & t_{n,n+1}^{2,2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$
FLM	4	$\begin{pmatrix} \psi_n^1 \\ \psi_n^2 \\ \psi_n^3 \\ \psi_n^4 \end{pmatrix}$	$\begin{pmatrix} \epsilon_n^1 & t_n^{1,2} & 0 & 0 \\ t_n^{2,1} & \epsilon_n^2 & t_n^{2,3} & 0 \\ 0 & t_n^{3,2} & \epsilon_n^3 & t_n^{3,4} \\ 0 & 0 & t_n^{4,3} & \epsilon_n^4 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & t_{n,n+1}^{2,2} & 0 & 0 \\ 0 & 0 & t_{n,n+1}^{3,3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

2.1. Wire Model

WM is the simplest TB model to describe nucleic acids and analogues [42,43]. It can be applied to mimic either single-stranded nucleic acids and hairpins at the single-base level [44] or double-stranded ones [45] at the base-pair level. In other words, if the WM refers to a single-stranded nucleic acid,

then the on-site energies are related to the energy levels of the four possible bases and the hopping integrals to the interaction between bases, while, if it refers to a double-stranded nucleic acid, then the on-site energies are related to the energy levels of the two possible base-pairs (incorporating the hydrogen bonding) and the hopping integrals to the interaction between base-pairs. It consists of just one chain ($L = 1$) and the parameters needed for its employment are the on-site energies of the bases or base pairs, ϵ_n , and the hopping integrals between successive bases or base pairs, t_n . A schematic representation of the WM is shown in Figure 2a.

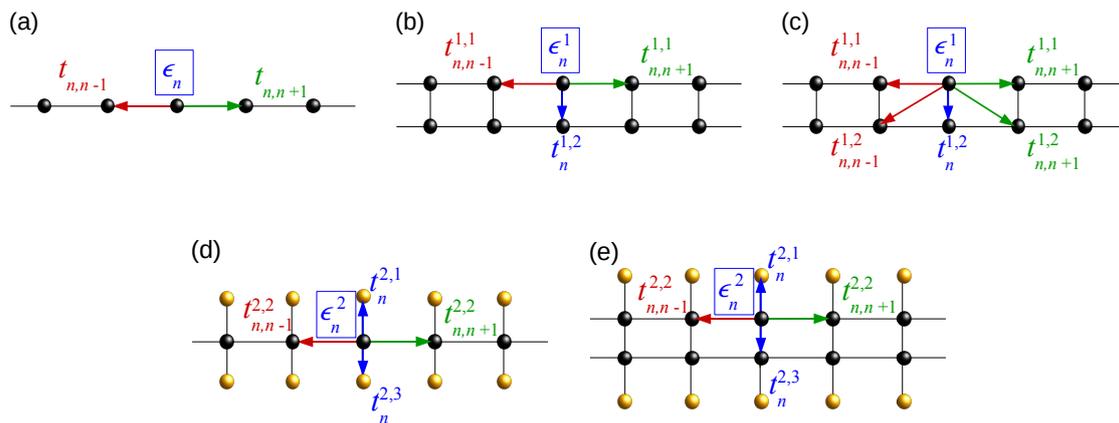


Figure 2. Schematic representation of the TB models listed in Table 1. (a) Wire Model (WM); (b) Ladder Model (LM); (c) Extended Ladder Model (ELM); (d) Fishbone Model (FM); (e) Fishbone Ladder Model (FLM).

2.2. Ladder Model

LM is the simplest model that can address the influence of base-pairing in the energetics of nucleic acids [42,46]. It consists of two chains ($L = 2$) and the parameters needed for its employment are the on-site energies of the bases, ϵ_n^l , the inter-strand hopping integrals between successive bases, $t_{n,n\pm 1}^{ll}$, and the intra-base-pair hopping integrals, $t_n^{ll'}$, due to the hydrogen bonds formed by the complementary bases in a pair. A schematic representation of the LM is shown in Figure 2b.

2.3. Extended Ladder Model

ELM is a more detailed version of the LM, including the inter-strand hopping integrals, $t_{n,n\pm 1}^{ll'}$, between the bases of successive base pairs [46,47]. A schematic representation of the ELM is shown in Figure 2c.

2.4. Fishbone Model

FM is the simplest model that can take into account the effect of the sugar-phosphate backbone [29,42]. It consists of three chains ($L = 3$). The central one corresponds to the base pairs, with each one being interconnected with the top and bottom chains, which represent the backbone sites. The latter are not connected with each other, since the insulating sugars are separating phosphate groups from one another [11,48]. Hence, the parameters needed for its employment are the on-site energies, ϵ_n^l , of the base pairs ($l = 2$) and the backbone sites ($l = 1, 3$), the intra-strand hopping integrals between successive base pairs, $t_{n,n\pm 1}^{2,2}$, and the inter-strand hopping integrals, $t_n^{ll'}$, between the base pairs and the backbone. A schematic representation of the FM is shown in Figure 2d.

2.5. Fishbone Ladder Model

FLM is a combination of the LM and the FM [29,42]. It thus includes both the effect of base-pairing and the presence of the sugar-phosphate backbone. It consists of four chains ($L = 4$). The two central

ones ($l = 2, 3$) correspond to the nitrogenous bases and the edge ones ($l = 1, 4$) to the backbone sites. Hence, the parameters needed for its employment are the on-site energies, ϵ_n^l , of the base pairs ($l = 2, 3$) and the backbone ($l = 1, 4$), the intra-strand hopping integrals between base pairs, $t_{n,n\pm 1}^{ll}$ ($l = 2, 3$) and the inter-strand hopping integrals between the bases of a base pair as well as between each base and the backbone, t_n^{ll} . A schematic representation of the FLM is shown in Figure 2e.

2.6. Additional Remarks

Apart from the models described above, one can introduce several other variants to describe nucleic acids. For example, an obvious extension would be a fishbone extended ladder model. Additionally, several other models have been proposed, including intra-backbone interactions [27,46,49], single-stranded nucleic acids with a backbone [49] and explicit inclusion of helicity [50] strain [51], and spin-orbit coupling [52] effects. We also mention that more complex models can be reduced to simpler ones via a renormalization scheme, which reduces the degrees of freedom of the system. Then, the on-site energies of the renormalized Hamiltonian are energy-dependent. This procedure is important when environmentally induced effects are considered [29]. For example, the FLM can be reduced into an LM via a one-step renormalization procedure [53], or to an even simpler WM via a two-step renormalization procedure [54,55].

Several techniques can be applied to solve the models, depending on what is studied, such as the numerical diagonalization of the Hamiltonian in Equation (2) [47,56,57], the transfer matrix method [58–60] outlined above, and the Non-Equilibrium Green's Function technique [61]. As it is apparent from Equation (3), the transfer matrix method is not applicable if the matrices τ_n are singular. Generally, this is the case, e.g., for the FM and the FLM (cf. Table 1). Then, a renormalization scheme is needed to apply the transfer matrix method.

Relevant parametrizations for nucleic acids have been proposed in many works and used within various TB models. For example, for on-site energies and hopping integrals, cf. Refs. [16,17,20,62,63], for on-site energies, cf. Refs. [64–70], and for hopping integrals, cf. Refs. [71–73]. Such parametrizations allow researchers to go beyond the chemically unrealistic treatments, such as the assumptions that all hopping integrals or on-site energies are equal, i.e., disorder in the Hamiltonian is either purely diagonal or off-diagonal, respectively, and address in more detail the complexity of nucleic acid energy structure.

3. Aperiodic One-Dimensional Wires

The dichotomy between the notions of order and disorder has expanded beyond a simple distinction between periodicity and aperiodicity, since the first observation of icosahedral diffraction patterns in the spectrum of an $\text{Al}_{0.86}\text{Mn}_{0.14}$ alloys [74] (2011 Nobel Prize in Chemistry for Prof. Dan Shechtman). The discussion that opened in the scientific community following this and other relevant discoveries led to a change in the very definition of the term crystal by the International Union of Crystallography in 1992, expanding it from referring solely to periodically arranged structures to “any solid having an essentially discrete diffraction diagram” [75]. This extended notion of crystals encompasses a whole family of structures, called quasi-periodic crystals or quasicrystals. Quasicrystals do not possess the translation symmetry that is inherent to classical (periodic) crystals; however, they possess inflation/deflation symmetry which leads to long-range order as well.

The discovery of quasicrystals has turned scientific interest into the study of specific one-dimensional aperiodic lattices, modeled with TB [76], i.e., described by Equation (1). The lattices are typically created using substitutional sequences. Apart from the interest the study of such systems has in itself, it is applicable, among other systems of physical relevance, in nucleic acids, as seen in Section 2. The ability to produce synthetic, de novo, nucleic acid sequences of interest [77], using mainly the phosphoramidite method [78] (although other promising methods have recently been proposed [79]), provides a chance not only to examine theoretical predictions regarding aperiodic structures, but also to create molecular wires with tailored properties. Below, we present some details

about substitutional sequences as well as some of the most commonly used ones in the literature of one-dimensional wires generally, and nucleic acids specifically.

3.1. Aperiodic Substitutional Sequences

Aperiodic substitutional sequences are based on an alphabet, e.g., $\mathcal{A} = \{A, B, C, D, \dots\}$ equipped with substitution rules that apply to each of its letters, $s(j), \forall j \in \mathcal{A}$. In the case of nucleic acids, the alphabet letters correspond to nitrogenous bases, i.e., G, C, A, T, U (for double-stranded chains the complementary strand is implied). The sequences start with a seed, i.e., a letter belonging to the alphabet (0th generation of the sequence). The substitution rules replace each alphabet letter by finite words consisting of alphabet letters, i.e., $s(j) = j'_1 j'_2 \dots j'_k, \forall j \in \mathcal{A}$. Iterating this procedure g times constructs the g th generation of the sequence.

Substitutional sequences can, in most cases, be described by introducing the substitution matrix, \mathbf{S} . It is a square, non-negative matrix of order $\text{card}(\mathcal{A})$ (the cardinality of a set is the number of elements of the set), and its elements are $S_{ij} = n_i[s(j)]$, where $n_i[s(j)]$ is the number of times the letter i is present in the substitution rule $s(j)$. Notice that, by definition, \mathbf{S} does not contain information about the ordering of letters in the sequence, hence more than one substitutions can have the same substitution matrix. However, the substitution matrix reveals much information about the underlying order and other properties of the corresponding sequence at the thermodynamic limit.

3.2. Primitive Substitutions and the Perron–Frobenius Eigenvalue

The matrix \mathbf{S} (and, hence, the substitution) is called primitive if there exists a natural number k such that S^k is a positive matrix. For primitive substitutions, the Perron–Frobenius theorem [80,81] guarantees that \mathbf{S} has a largest, unique, real, positive eigenvalue, λ_{PF} , and its corresponding (left and right) eigenvectors can be chosen to have strictly positive entries. The components of the right eigenvector associated with λ_{PF} , normalized such as their sum is unity, give the asymptotic relative frequencies of the letters in \mathcal{A} . Hence, using \mathbf{S} , one can determine the occurrence percentage of each nucleotide in a substitutional nucleic acid sequence.

3.3. Induced Substitutions

In addition to the previous discussion, it is also possible to determine the letter frequencies of the legal words of length k in a substitutional sequence with primitive \mathbf{S} (corresponding to nucleotide k -plets). This can be done as follows [82]; let $W = \{w = j_1 j_2 \dots j_k, \forall j \in \mathcal{A}\}$ be the set of the legal k -letter words in the sequence and $s(w) = s(j_1) s(j_2) \dots s(j_k) = j'_1 j'_2 \dots j'_n$ the word constructed from a letter-by-letter substitution of the word w . Then, the induced substitution of a k -letter word, $s_k(w) = (j'_1 j'_2 \dots j'_k) (j'_2 j'_3 \dots j'_{k+1}) \dots (j'_l j'_{l+1} \dots j'_{l+k-1})$, where l is the number of letters in $s(j_1)$, is also primitive. Hence, an induced primitive substitution matrix \mathbf{S}_k can be defined, from which the asymptotic letter frequencies of the legal k -letter words of the sequence can be determined using the Perron–Frobenius theorem. For sequences in which \mathbf{S} is defined via a helping alphabet [83], k -letter word frequencies can be deduced in the same fashion from the legal $2k$ -letter words of the helping alphabet.

3.4. The Pisot Property

A real algebraic integer (i.e., a real solution of a monic integer polynomial) is said to be a Pisot–Vijayaraghavan number if its modulus is larger than unity, and all its algebraic conjugates (i.e., the other solutions of the polynomial) have modulus strictly less than unity [84]. A substitution has the Pisot property if the matrix \mathbf{S} has a largest, unique, real, positive eigenvalue which is a Pisot–Vijayaraghavan number, and for all the other eigenvalues, λ , it holds that $|\lambda| < 1$. If the characteristic polynomial of \mathbf{S} is irreducible over the rationals, the Pisot substitution is called irreducible. Irreducible Pisot substitutions are a subclass of primitive substitutions [85].

Let us remember some definitions. Given n linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n \in \mathbb{R}^m$, the lattice generated by them is defined as $\mathcal{L}(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n) = \sum_i x_i \mathbf{b}_i, x_i \in \mathbb{Z}$. We call the set $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$

a *basis* of the lattice. We say that the *rank* of the lattice is n and its *dimension* is m . The Fourier transform of the (direct) lattice is a lattice that is called the reciprocal lattice.

Furthermore, according to the Lebesgue's decomposition theorem, any measure on \mathbb{R} can be decomposed into three parts: a pure point (or discrete) part, an absolutely continuous part, and singularly continuous part. This theorem helps to categorize the energy or Fourier spectra of aperiodic substitutional sequences.

The first connections between the irreducible Pisot property and the Fourier spectrum of a substitutional sequence were reported in Refs. [86,87], where it was conjectured that if the Perron–Frobenius eigenvalue of a substitutional system is a Pisot–Vijayaraghavan number, then the system is quasiperiodic. Later studies have revealed more details, providing a more sophisticated classification of substitutional systems with respect to the nature of their diffraction spectrum. In the one-dimensional case, sequences produced from irreducible Pisot substitutions have pure point Fourier spectra [88]. (I) The Pisot property, together with (II) the extra condition $\lambda \neq 0$, provide the means to distinguish between:

- (1) *strictly quasiperiodic* sequences, in which the rank of the reciprocal lattice n_r is finite and larger than the dimension of the physical space of the sequence m , and
- (2) *limit-quasiperiodic* sequences, in which the rank of reciprocal lattice n_r is countably infinite (in a 1–1 correspondence with the natural numbers or integers).

The distinction criterion between categories (1) and (2) is the value of the determinant of \mathbf{S} : unimodular \mathbf{S} implies strict quasiperiodicity, otherwise the structure is limit-quasiperiodic [89–91]. Limit-quasiperiodic structures can be interpreted as a superposition of an infinite number of strictly quasiperiodic structures. Examples of *strictly quasiperiodic* structures are the classical Fibonacci sequence [92] as well as all the precious means sequences [93] and the Fibonacci-class sequences [94] (cf. Table 2, where several substitutional sequences studied in the literature are listed, together with their substitution rules and matrices). *Limit-quasiperiodic* structure representatives are the mixed means sequences with $n \geq m$ [95].

For substitutions not satisfying the above-mentioned conditions (I) and (II), the situation is more complex. In such cases, the Fourier spectrum can be:

- (3) *limit-periodic*, i.e., a superposition of countably infinite periodic structures. Some examples are the period doubling sequence and metallic means sequences with $n = l(l + 1)$ [96],
- (4) *singular continuous*, i.e., non-constant, non-decreasing, continuous and has zero derivative, everywhere that the derivative exists. Examples are the Thue–Morse sequence [97–99] and metallic means sequences with $n \neq l(l + 1)$ [96], or even
- (5) *absolutely continuous*, such as the Rudin–Shapiro sequence [100,101].

Apart from the above-mentioned sequences, there are others for which the substitution is not primitive or the matrix \mathbf{S} cannot even be defined at all. Examples of non-primitive substitutions include the sequences inspired by the Cantor set [102], maybe the most well-known deterministic fractal. A sequence for which a substitution matrix cannot be defined is the classical Kolakoski(1,2) sequence [103,104], and generally Kolakoski(p,q) sequences where p is odd and q even or vice versa [105]. The situation is different when p and q are both even or odd; then, a primitive \mathbf{S} can be defined. In the former case, the sequences have been classified as limit-periodic [106]. In the latter case, the irreducible Pisot property holds when $2(p + q) \geq (p - q)^2$, and \mathbf{S} is also unimodular when $p = q \pm 2$ [105].

Table 2. Substitutional sequences studied in the literature, together with the alphabets through which they are defined, the corresponding substitution rules, and the substitution matrices. In the last row, the subscripts *o* and *e* in the substitution rules denote substitutions that are applied on odd and even positions in the sequence, respectively.

Sequence	\mathcal{A}	Substitution Rule	S
Fibonacci	{A, B}	$s(A) = AB \quad s(B) = A$	$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$
Precious means	{A, B}	$s(A) = A^n B \quad s(B) = A$	$\begin{pmatrix} n & 1 \\ 1 & 0 \end{pmatrix}$
Fibonacci-class	{A, B}	$s(A) = B^{n-1} A B \quad s(B) = B^{n-1} A$	$\begin{pmatrix} 1 & 1 \\ n & n-1 \end{pmatrix}$
Mixed means	{A, B}	$s(A) = A^n B^m \quad s(B) = A$	$\begin{pmatrix} n & 1 \\ m & 0 \end{pmatrix}$
Metallic means	{A, B}	$s(A) = AB^n \quad s(B) = A$	$\begin{pmatrix} 1 & 1 \\ n & 0 \end{pmatrix}$
Period doubling	{A, B}	$s(A) = AB \quad s(B) = AA$	$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$
Thue–Morse	{A, B}	$s(A) = AB \quad s(B) = BA$	$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$
Rudin–Shapiro	{A, B, C, D}	$s(A) = AB \quad s(B) = AC$ $s(C) = DB \quad s(D) = DC$	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
Triadic Cantor set	{A, B}	$s(A) = ABA \quad s(B) = BBB$	$\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$
Asymmetric Cantor set	{A, B}	$s(A) = ABAA \quad s(B) = BBBB$	$\begin{pmatrix} 3 & 0 \\ 1 & 4 \end{pmatrix}$
Generalized Cantor set (<i>t, d</i>)	{A, B}	$s(A) = A^{\frac{t-d}{2}} B^d A^{\frac{t-d}{2}} \quad s(B) = B^t$	$\begin{pmatrix} t-d & 0 \\ d & t \end{pmatrix}$
Kolakoski ($p = 2m, q = 2n$)	{A = pp , B = qq }	$s(A) = A^m B^m \quad s(B) = A^n B^n$	$\begin{pmatrix} m & n \\ m & n \end{pmatrix}$
Kolakoski ($p = 2m + 1, q = 2n + 1$)	{A = pp , B = pq , C = qq }	$s(A) = A^m B C^m \quad s(B) = A^m B C^n$ $s(C) = A^n B C^n$	$\begin{pmatrix} m & m & n \\ 1 & 1 & 1 \\ m & n & n \end{pmatrix}$
Kolakoski ($p = 2m, q = 2m + 1$) or ($p = 2m + 1, q = 2m$)	{ p, q }	$s_o(q) = p^q \quad s_o(p) = p^p$ $s_e(q) = q^q \quad s_e(p) = q^p$	undefinable

4. Energy Structure of Nucleic Acid Wires

The energy structure of a physical system is closely connected to many of its properties (electrical, magnetic, thermal, optical, et cetera). A useful –and closely related to experimental data–quantity that describes the energy structure of a given system is the density of states (DOS), which shows the number of states that can be occupied by electrons at each energy. It can be formally defined as

$$g(E) = \sum_k \delta(E - E_k), \tag{6}$$

where no spin degeneracies are included. The sum runs over all allowed states, each of which has an eigenenergy E_k . A closely related quantity is the integrated density of states (IDOS), defined as

$$IDOS(E) = \int_{-\infty}^E g(E') dE', \tag{7}$$

i.e., it is the number of states that have energy smaller than E . Discontinuities in the IDOS indicate the presence of energy gaps, and the height of an IDOS step gives information about the level population. For periodic systems, the regions of allowed energies lead to smooth parts in DOS or IDOS curves, separated by well defined gaps at specific energies, thus reflecting the continuous electronic band

structure of a periodic crystal. On the contrary, the DOS and IDOS of random systems are rough, indicative of the presence of a multitude of gaps between the allowed energy levels. As it has to do with deterministic aperiodic sequences with a substitution rule, which reflects their self-similarity, it has been conjectured (and proven, in several specific cases) that their energy spectrum is singular continuous, i.e., in the thermodynamic limit, it exhibits an infinity of gaps and vanishing bandwidths [107].

Furthermore, for primitive substitutions described by a Hamiltonian corresponding to the WM, the following gap-labeling theorem has been introduced by Bellissard et al. [108]:

Theorem 5.13 of Ref [108]. *Let \hat{H} be a Hamiltonian corresponding to the WM, where the coefficients (i.e., parameters) are determined by a primitive substitution on a finite alphabet. Then, the values of the IDOS of \hat{H} on the spectral gaps in $[0, 1]$ belong to the $\mathbb{Z}(\lambda_{PF}^{-1})$ module generated by the components of the eigenvectors \vec{v}_{PF} and $\vec{v}_{PF,2}$ of the substitution matrices \mathbf{S} and \mathbf{S}_2 , respectively.*

From the above theorem, it follows that, in order to obtain the position of the gaps in the (normalized) IDOS of a primitive substitutional sequence within the WM, it is sufficient to know the substitution matrices of its legal 1- and 2-letter words (c.f. Section 3.3). Specifically, the gaps can be labeled by the negative powers of λ_{PF} times integral linear combinations of the components of \vec{v}_{PF} and $\vec{v}_{PF,2}$ that lie within the interval $[0, 1]$ [108,109]. For example, in the case of Fibonacci sequences, from the diagonalization of \mathbf{S} (cf. Table 2), we get $\lambda_{PF} = \phi$ and $\vec{v}_{PF} = [\phi^{-1} \ \phi^{-2}]^T$ (where ϕ is the golden ratio). Hence, the sequence consists of $\approx 61.8\%$ A letters and $\approx 38.2\%$ B letters. The legal 2-letter words in the Fibonacci sequence are BA, AB, and AA (i.e., BB is forbidden), thus the induced 2-substitution reads (cf. Section 3.3) $s_2(AA) = (AB)(BA)$, $s_2(AB) = (AB)(BA)$, $s_2(BA) = (AA)$, leading to the induced substitution matrix

$$\mathbf{S}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}. \tag{8}$$

The Perron–Frobenius eigenvector (cf. Section 3.2) of \mathbf{S}_2 is $\vec{v}_{PF,2} = [\phi^{-3} \ \phi^{-2} \ \phi^{-2}]^T$. Hence, the gaps can be labeled by integer linear combinations of negative powers of ϕ . Since every positive power of ϕ can be reduced to a linear expression of the form $\phi^g = N_g\phi + N_{g-1}$, where N_g is the Fibonacci number of generation g , and it also holds that $\phi^g + \phi^{-g} \in \mathbb{N}^*$, the situation can be reduced to an integral linear combination of 1 and ϕ . Thus, the positions of the gaps in the IDOS of a Fibonacci sequence within the WM can be given by

$$\{\mathcal{G}_n\} = \{n\phi \pmod 1, \forall n \in \mathbb{Z}\}. \tag{9}$$

Another interesting remark, arising from the DOS values of a single-stranded Fibonacci DNA sequence consisting of G and C, is that the ratio among the distances between DOS of consecutive generations tends to ϕ [110]. The IDOS of periodic, several aperiodic, and random binary DNA sequences with G and A on the same strand, calculated within the WM, taking into account both diagonal and off-diagonal disorder is presented in Figure 3 [83]. Periodic sequences display two well defined bands, separated by a single energy gap (the largest among all cases). Thue–Morse, Fibonacci, Rudin–Shapiro, and Kolakoski sequences possess a staircase-like IDOS, while the shape of random sequence IDOS resembles, albeit it is more disrupted, to that of Rudin–Shapiro, and its main energy gap is the smallest among all cases. The fractal, Cantor set based, sequences have a very rough spectrum. For all sequences, the value of the IDOS at the largest energy gap is equal to the occurrence percentage of A. Furthermore, it has been observed that there are steps in the IDOS, the relative value of which is equal to the occurrence percentages of the possible base-pair triplets [83]. This remark holds for all categories of deterministic aperiodic sequences, either generated by a primitive substitution matrix or not, such as Kolakoski (1,2), further connecting the specific base-pair sequence of a DNA segment with its energy structure. The above-mentioned IDOS steps and the corresponding values are marked (where possible) in the left vertical axes of Figure 3.

Apart from the sequential disorder, mentioned above, other disorder types are present or can be induced in nucleic acid sequences. In Ref. [111], the authors study single poly(CG) or poly(CT) DNA strands with diluted base-pairing, i.e., for example, the G sites are randomly attached to their complementary C sites, with a probability p . The C-G base pairs are renormalized onto the first strand, leading to two inter-penetrating lattices: a periodic one containing the G or T sites and a random one containing bare and renormalized C sites. The DOS for three indicative cases, i.e., for $p = 0, 0.5$, and 1 , is presented at the top panels of Figure 4. When $p = 0$, there are two well defined bands arising from the periodicity of the segment. The band character is maintained for $p = 1$, with a smaller gap for poly(CG), while for poly(CT) the number of bands changes to three, reflecting the total number of different sites (since the renormalization procedure takes into account the original structure). When $p = 0.5$, for poly(CG), fluctuations of the same magnitude in both allowed energy regions arise and the singularities are rounded off due to the induced disorder. For poly(CT), the bands collapse at a single energy region, stronger fluctuations are present at smaller energies than at larger ones, and there is a persisting van Hove singularity exactly at the on-site energy of T. Hence, in this case, diluted base-pairing produces a gapless structure and keeps a number of states extended (around the on-site energy of T), which is an ideal scenario for charge transport.

Several human diseases are associated with aberrant DNA methylation, which is heritable during cell division but does not alter the DNA sequence. In Ref. [112], a poly(CG) single-stranded segment is considered, with methyl groups randomly connected with the 5-carbon of C bases (forming the so-called 5-methylcytosine), again, with probability p . For completely unmethylated or methylated segments, the DOS consists of two smooth bands, derived by the on-site energy of G and C, for $p = 0$, and of G and 5-methylcytosine for $p = 1$, respectively; the only difference is in the energy intervals of the allowed states. For $0 < p < 1$, the smooth profile of the DOS is degraded, since the presence of randomly distributed methyl groups along the chain introduces a small disorder, which in turn leads to an enhancement in the effective resistance that can reach one order of magnitude.

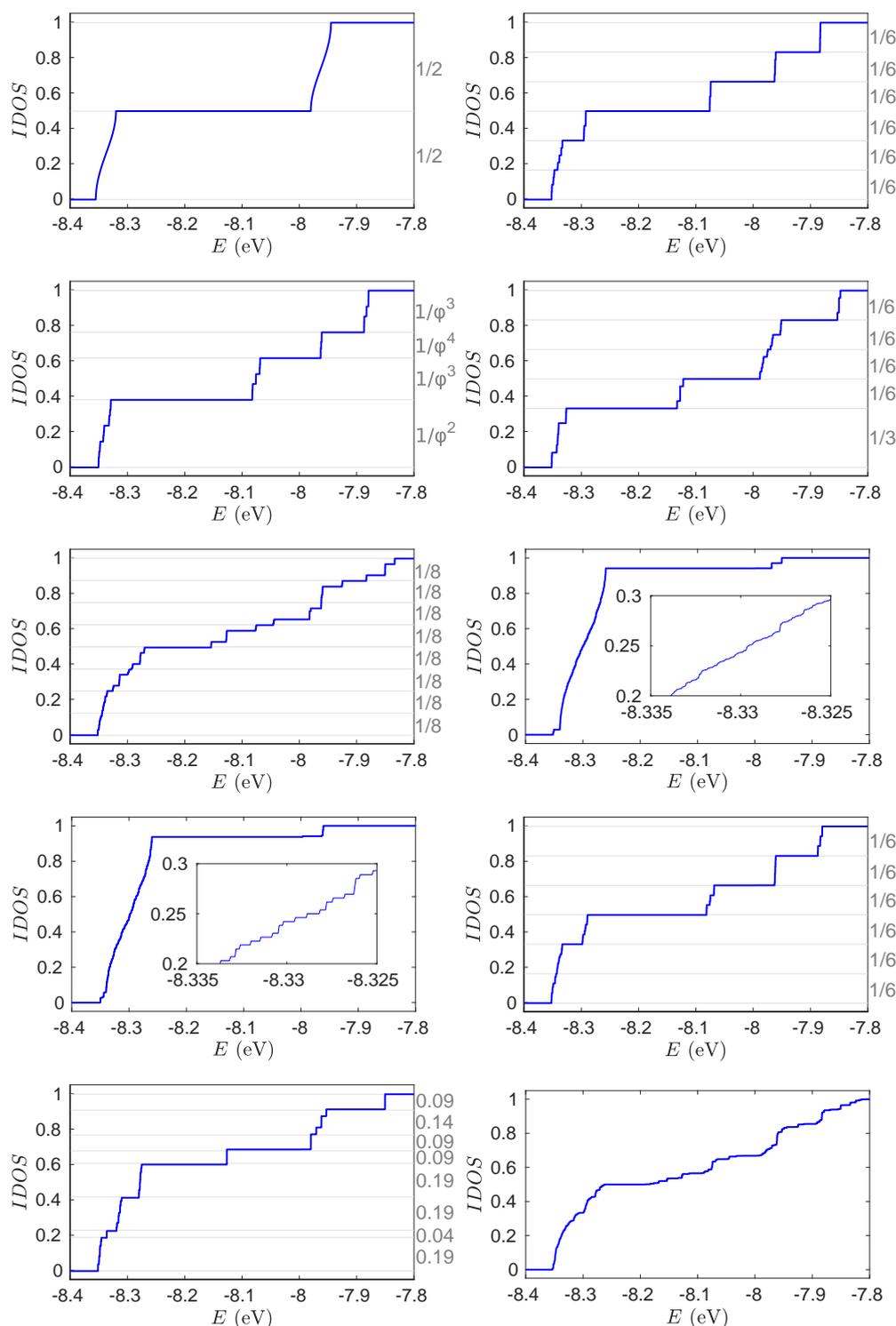


Figure 3. Normalized IDOS of various categories of binary DNA segments with purines on the same strand, within the WM. (a) Poly(GA); (b) Thue–Morse; (c) Fibonacci (d) Period-doubling; (e) Rudin–Shapiro; (f) Cantor Set; (g) Generalized Cantor Set (4,2); (h) Kolakoski(1,2); (i) Kolakoski(1,3); (j) Random (50% G, 50% A). Reprinted figure from K. Lambropoulos and C. Simserides, Periodic, quasiperiodic, fractal, Kolakoski, and random binary polymers: Energy structure and carrier transport, *Phys. Rev. E* **2019**, *99*, 032415 [83] <http://dx.doi.org/10.1103/PhysRevE.99.032415>, © 2019 by the American Physical Society.

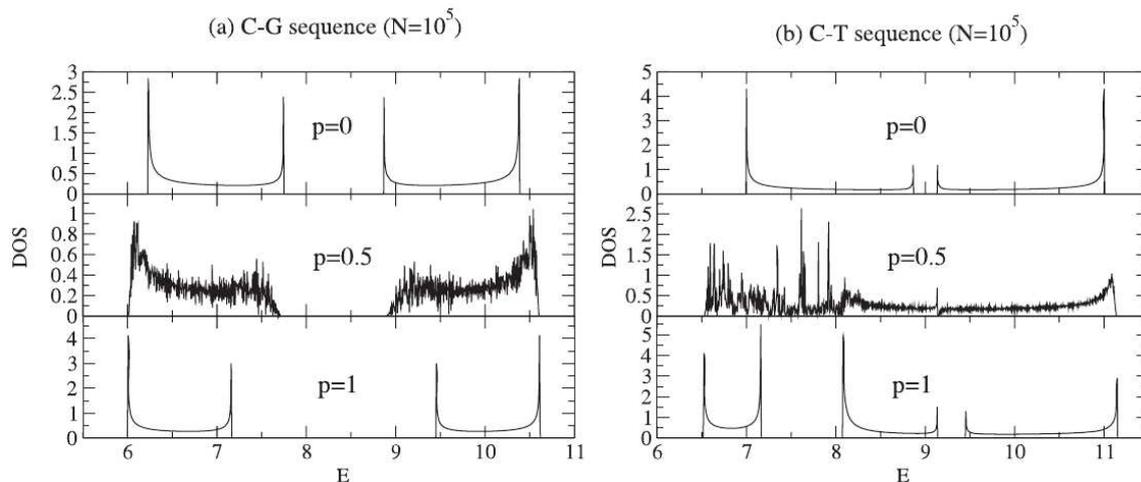


Figure 4. DOS of (a) poly(CG), and (b) poly(CT) DNA strands with diluted base-pairing at random cytosine sites with probability p . Figure reproduced with permission from F. A. B. F. de Moura, M. L. Lyra and E. L. Albuquerque, Electronic transport in poly(CG) and poly(CT) DNA segments with diluted base pairing, *J. Phys. Condens. Matter* **2008**, *20*, 075109 [111] <http://dx.doi.org/10.1088/0953-8984/20/7/075109>, © 2008 IOP Publishing. All rights reserved.

5. Coupling Nucleic Acids with Leads: Transmission Coefficients

In order to study the charge transport properties of nucleic acid nanowires within a TB framework, the system under examination is attached to two semi-infinite homogeneous metallic leads, which play the role of a carrier bath. The leads are characterized by a single on-site energy, ϵ_M , and a single hopping integral, t_M , so that the allowed energy states of the incident and outgoing waves lie in the interval $[\epsilon_M - 2|t_M|, \epsilon_M + 2|t_M|]$. Since detailed information on the nucleic acid's chemical bonding at the contacts is not known, one introduces effective parameters dealing with the tunneling probability between the frontier orbitals, roughly encompassing the bonding effects at the interface [113]. These parameters are $t_{R(L)}$ and couple the left (right) lead with the nucleic acid wire.

A first useful physical quantity to evaluate the charge transport properties of a quantum system is the transmission coefficient, $T(E)$. It is an energy-dependent quantity that describes the probability that a carrier, incident to a quantum wire, transmits through its eigenstates. Charge transport will experience a sequence-dependent contribution of backscattering, according to the distribution of potential barriers, corresponding to bases or base pairs, over the length scale of the sequence [24].

The coupling between the nucleic acid and the leads plays an important role on the transmission profiles. It has been shown that *stronger coupling does not necessarily mean higher transmission*. In Ref. [114], the authors studied the transmission profiles of a single-stranded poly(GACT) DNA chain within the WM, with purely diagonal disorder, assuming equal coupling parameters with both leads ($t_R = t_{cL} = \tau$), and arrived at the resonance condition $\tau = \sqrt{t_M t}$, where t is the hopping integral between the wire sites. From Figure 5, it is evident that, when the value of the coupling parameter is either smaller or larger than the one fulfilling the resonance condition, quite smaller transmission peaks are obtained. This result properly illustrates the influence of contacts on electrical transport. This extreme sensitivity is due to interference effects between the DNA molecular bands and the electronic structure of the leads at the lead-DNA interface.

The above-mentioned results were generalized in an analytical manner for any periodic WM, through the conditions $\omega = \frac{t_M t_u}{t_R t_L} = \pm 1$ (ideal coupling condition), where t_u couples the moieties at the end of a unit cell and at the start of the next, and $\chi = \frac{t_L}{t_R} = \pm 1$ (symmetric coupling condition) [60]. The ideal coupling condition, $\omega = \pm 1$, implies that the system and the leads are interconnected as if they were connected to themselves. When this condition is reached, the existence of fully resonant states is guaranteed at specific energies determined by the zeros of Chebyshev polynomials of the

second kind [115]. Hence, any periodic sequence can display full transmission, if appropriate couplings are utilized. Deviations from the symmetric coupling condition give rise to secondary peaks. The effect of the coupling strength and the asymmetry factors, together with the internal hoppings, is exemplified in Figure 6, for a generic periodic WM with two sites per unit cell (hence, two hopping integrals t_1 and t_2 connect the wire sites) and $N = 10$. It is evident that the ideal and symmetric coupling conditions lead to the most efficient transmission. For ideal and asymmetric coupling, except for the peaks of magnitude 1, there is one additional peak, which is of significant magnitude only when $|t_1| \approx |t_2|$. In the strong (weak) and symmetric coupling regimes, the peaks that are closer to the band gap vanish (emerge) as $\left|\frac{t_1}{t_2}\right|$ increases. When the coupling is asymmetric, transmission is enhanced only in one of the two bands.

Analogous conclusions can be obtained for more complex TB models. In Ref. [31], a poly(G)-poly(C) oligomer ($N = 5$) was studied within the FM. The authors report that, for small values of coupling, the transmission shows sharp and narrow unit resonances due to the localization of states, while, as the coupling increases, the well-arranged resonant peaks overlap. An inspection of Figure 7 of Ref. [31] indicates that there are intermediate values of $t_L (= t_R)$ in which the overall transmission is more enhanced compared to smaller and larger values.

In Ref. [116], the authors study a poly(G)-poly(C) chain within an extension of the FLM, which allows hopping between backbone sites as well as all possible diagonal hoppings (between the nitrogenous bases as well as between the bases and the backbone). Each of the two strands containing the DNA bases is connected with each lead with equal coupling parameters. For diagonal hoppings being switched either on or off, it can again be concluded that stronger coupling with the leads does not necessarily lead to enhanced transmission (cf. the panels in the first two rows in Figure 7). This is also evident by comparing the averaged transmission coefficient, which is defined as

$$T_a(E) = \frac{\int_{E_{\min}}^E T(e)de}{E - E_{\min}}, \quad (10)$$

cf. the panels in third row of Figure 7. Although $T(E)$ and $T_a(E)$ are indeed much smaller for $t_L = t_R = 0.1$ eV, an increase from 0.5 eV to 0.9 eV does not lead to transmission enhancement. In fact, for diagonal hoppings switched both on and off, T_a reaches larger values for the intermediate coupling $t_L = t_R = 0.5$ eV.

The above discussion demonstrates that, apart from the internal degree of disorder of a given sequence, other factors can significantly affect their charge transport properties.

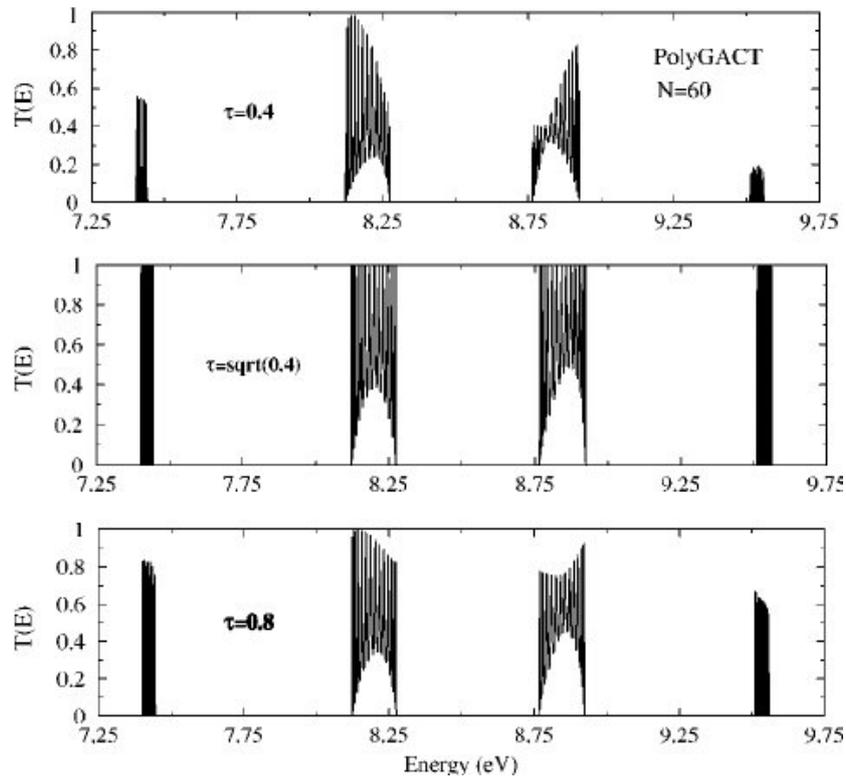


Figure 5. Transmission coefficient for a poly(GACT) chain within the WM, with $N = 60$, $t_M = 1.0$ eV, $t = 0.4$ eV, and $\tau = 0.4$ eV; i.e., $\tau = \sqrt{t_M t}$ (top), $\tau = \sqrt{0.4}$ eV (middle), $\tau = \sqrt{0.8}$ eV (bottom). Reprinted figure with permission from E. Maciá, F. Triozon, and S. Roche, Contact-dependent effects and tunneling currents in DNA molecules, *Phys. Rev. B* **2005**, *71*, 113106 [114] <http://dx.doi.org/10.1103/PhysRevB.71.113106>, © 2005 by the American Physical Society.

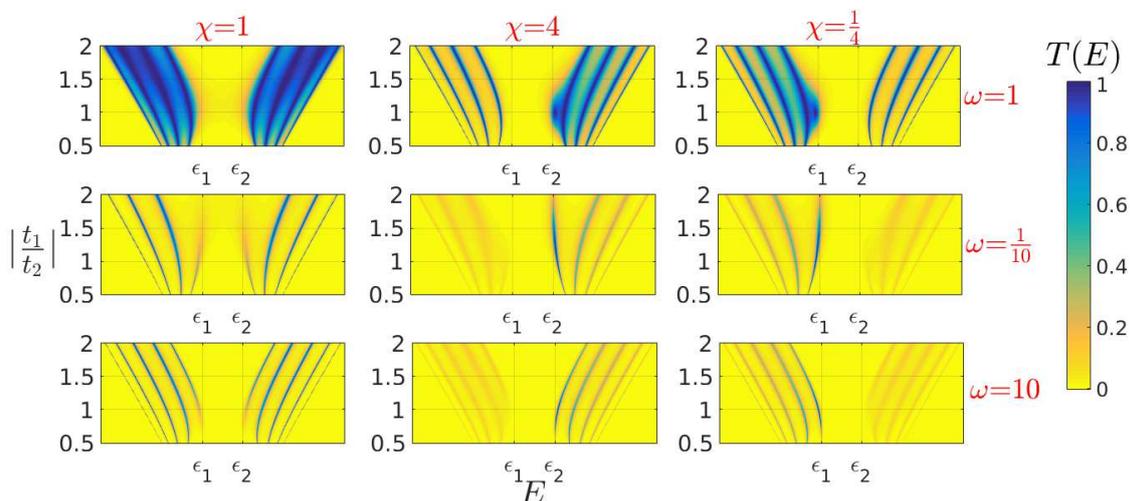


Figure 6. Transmission coefficient of a periodic WM with two sites per unit cell and $N = 10$ for ideal (top), strong (middle), and weak (bottom) coupling with the leads. (Left column) Symmetric coupling. (Middle column) Asymmetric coupling with $|\chi| > 1$. (Right column) Asymmetric coupling with $|\chi| < 1$. The leads parameters are such that all the eigenstates of the system are contained. Reprinted from Ref. [60], K. Lambropoulos and C Simserides, Spectral and transmission properties of periodic 1D tight-binding lattices with a generic unit cell: an analysis within the transfer matrix approach, *J. Phys. Commun.* **2018**, *2*, 035013 [60] <http://dx.doi.org/10.1088/2399-6528/aab065>, CC BY 3.0.

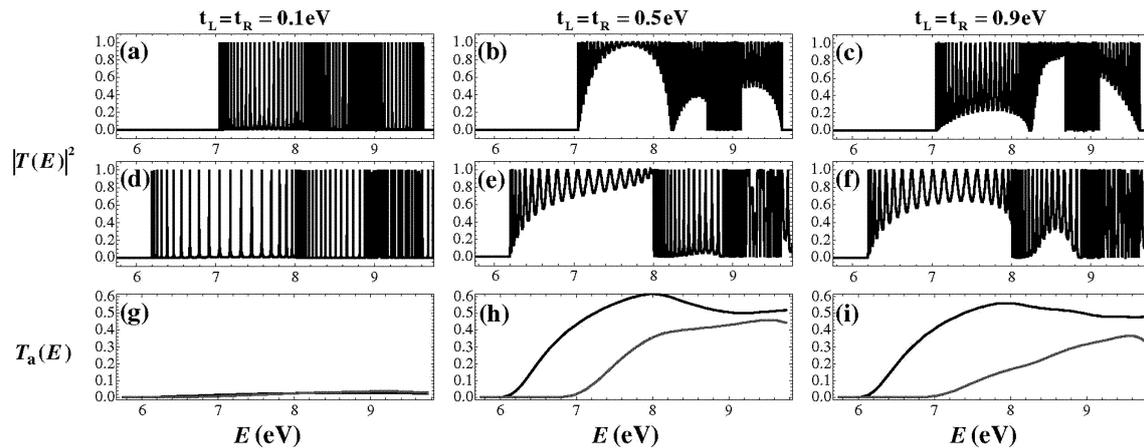


Figure 7. Transmission spectra as a function of energy without (a–c) and with (d–f) the diagonal hoppings; (g–i) average transmission spectra as a function of energy: gray line (diagonal hoppings switched off) and black line (diagonal hoppings switched on). (Left column) $t_L = t_R = 0.1$ eV. (Middle column) $t_L = t_R = 0.5$ eV. (Right column) $t_L = t_R = 0.9$ eV. Reprinted from S. Malakooti, E. R. Hedin, Y. D. Kim, and Y. S. Joe, Enhancement of charge transport in DNA molecules induced by the next nearest-neighbor effects, *J. Appl. Phys.* **2012**, *112*, 094703 [116], <http://dx.doi.org/10.1063/1.4764310>, with the permission of AIP Publishing.

6. Current–Voltage Curves

The situation is more complex as far as the calculation of $I - V$ characteristic curves is concerned. The $I - V$ curve of a given nucleic acid segment can be given, using the Landauer–Büttiker formalism [61,117,118], by the relation

$$I(V) = \frac{2e}{h} \int_{-\infty}^{\infty} T(E, V) [f_L(E - \mu_L) - f_R(E - \mu_R)] dE, \quad (11)$$

under the assumption that charge propagates from left to right. $\mu_{L(R)}$ and $f_{L(R)}(E)$ are the chemical potential and the Fermi–Dirac distribution at the left (right) lead, respectively. From Equation (11), we deduce that there are several factors, apart from the structure of the sequence under examination that have an effect on the magnitude of currents, the bias regime and the shape of the $I - V$ curves. These factors include:

- The choice of the Fermi level of the leads E_F , which coincides with ϵ_M if one electron per site is assumed. If E_F is not aligned with an allowed energy region of the segment, then no currents occur in the vicinity of $V = 0$, while a metallic behavior is expected otherwise.
- The way the external bias is applied. For example, only one of the leads' energy bands can be shifted, so that $\mu_L = E_F + eV$, and $\mu_R = E_F$, or, alternatively, both leads' bands can be symmetrically shifted so that $\mu_{L,R} = E_F \pm \frac{eV}{2}$. This choice affects both the way the voltage drop is induced in the nucleic acid sequence and the energy limits of the conductance channel. At zero temperature, the Fermi–Dirac distributions become Heaviside step functions and determine the limits of integration. Hence, Equation (11) can be simplified to

$$I(V) = \frac{2e}{h} \int_{\mu_R}^{\mu_L} T(E, V) dE, \quad (12)$$

while, at finite temperatures, it can be written in the form

$$I(V) = \frac{2e}{h} \sinh\left(\frac{eV}{2k_B T}\right) \int_{-\infty}^{\infty} \frac{T(E, V) dE}{\cosh\left(\frac{E - E_F}{k_B T}\right) + \cosh\left(\frac{eV}{2k_B T}\right)}, \quad (13)$$

i.e., the $I - V$ curve occurs from the modulation of the hyperbolic function $\sinh\left(\frac{eV}{2k_B T}\right)$ by the integral factor expression [119].

- (c) Whether or not the transmission coefficient is considered as bias-dependent. Although assuming bias-independent transmission coefficient could be a justified choice in the small bias regime, and it is indeed less computationally costly, this assumption cannot lead, under any circumstances, to the occurrence of negative differential resistance, since an increasingly larger part (as V increases) of a nonnegative function is integrated.

There are several works discussing the $I - V$ curves of nucleic acid sequences, considering different types of order or disorder. Regarding sequential order or disorder, in Ref. [83], the $I - V$ curves of periodic, deterministic aperiodic, and random binary DNA segments have been studied within the WM. The curves have been shown to have clearly distinct shapes for different sequence categories. It has also been demonstrated that periodic sequences lead to the most enhanced currents. Additionally, there are several categories deterministic aperiodic sequences (specifically, Fibonacci, Period-doubling, Cantor and generalized Cantor) that can also display significant currents, depending on the Fermi level of the leads. Random sequences represent the least efficient category, since they were found to always display smaller currents than all their deterministic aperiodic counterparts with similar base-pair content.

In Ref. [120], the authors study dry and hydrated DNA sequences with correlated and uncorrelated disorder within a WM, for $N = 50$ and at a temperature of 300 K. For different concentrations of G and A sites, the resulting currents are larger for correlated disorder, both for dry and backbone-hydrated sequences. Generally, the authors report a conductor to semiconductor to insulator transition as a function of three effects, i.e., sequence size, disorder, and hydration, suggesting that an appropriate choice of chain size and relative concentration of base pairs can be used to tailor the electrical behavior of DNA strands.

A similar transition has been reported by introducing conformal variation at the helical symmetry as well as backbone disorder into a FLM [121]. Helical symmetry is taken into account via the inclusion of hopping integrals between bases in adjacent pitches (i.e., turns of the helix). The number of base-pairs within a given pitch is denoted by n . Backbone disorder is introduced by a random distribution of backbone on-site energies, characterized by a disorder strength w . The results for poly(G)-poly(C) and poly(A)-poly(T) chains with $N = 50$, for different values of n and w are shown in Figure 8. At low disorder, the effect of n is smaller, since, in that case, any path of charge conduction is equivalent, as an electron feels almost no potential variation. As the disorder increases, the effect of n becomes more distinctive, since there is substantial variation of the effective potential at different sites and an increase of n gives an electron more shortcut pathways to move along the DNA chain. The current is enhanced with increasing n , and the effect is more vivid for strong disorder. Furthermore, for weak disorder, a cut-off voltage is observed in the $I - V$ curves, which reduces with increasing n . At strong disorder, the current is enhanced and almost linear response is observed at larger values of n , which indicates a transition from the insulating to the metallic phase.

Thermal structural disorder has been studied in Ref. [122], by introducing a random variation in the hopping integrals of a poly(G)-poly(C) chain with $N = 5$, within an FLM allowing inter-backbone hoppings. Comparing the $I - V$ curves of such systems for $T = 0$ K and $T = 300$ K, the authors report that the voltage threshold for current onset is about the same, indicating that the thermal structural disorder does not affect the voltage gaps. Above that threshold, as the temperature increases, the linear behavior of the current changes to a step-like behavior, and the current is reduced, since the static distortion increases elastic scattering of electrons through the DNA molecule.

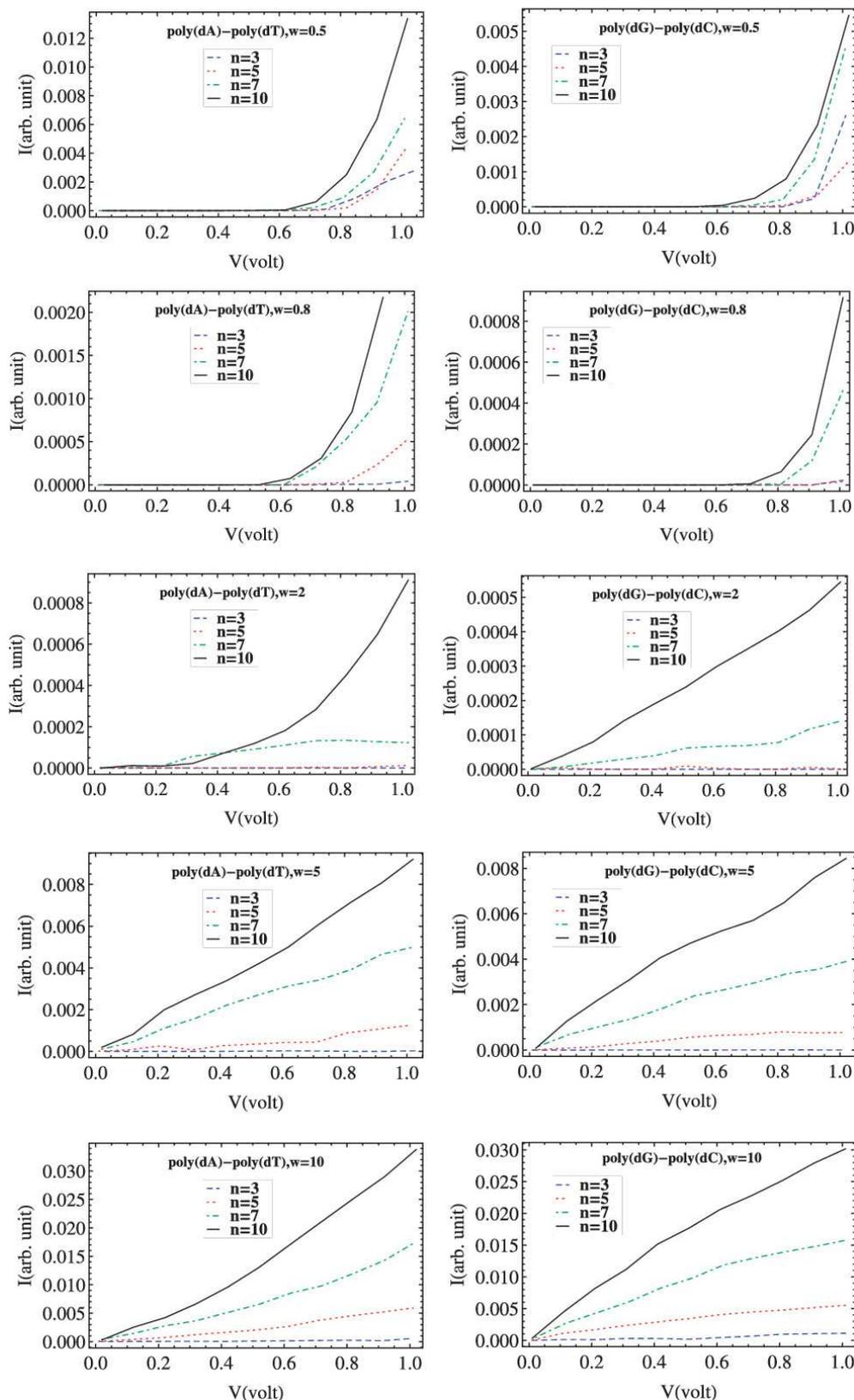


Figure 8. $I - V$ curves for poly(dA)-poly(dT) and poly(dG)-poly(dC) various disorder strengths w and pitch-size values n . For weak disorder, the cut-off voltage reduces with n , showing semiconducting behaviour. For strong disorder, the current is considerably enhanced with increasing n , giving a insulator to metal transition. Reproduced from Ref. [121], S. Kundu and S. N. Karmakar, Conformation dependent electronic transport in a DNA double-helix, AIP Adv. 2015, 5, 107122 [121] <http://dx.doi.org/10.1063/1.4934507>, CC BY 3.0.

The effect of cytosine methylation disorder on the $I - V$ curves of a single stranded GAGCTGACGTTACGG segment retrieved from the first sequenced human chromosome (chromosome 22) has been studied within the WM in Ref. [123]. The effect of all possible single, double, and triple methylation defects (out of the totally four C sites) is addressed. It is demonstrated that even a single methylated site reduces the currents by one order of magnitude. This reduction is directly associated with the fact that such sites act as additional impurity centers. The observed sensitivity of the saturation current on the position of the methylated cytosine is related to the impact of methylation on the hopping integrals to the neighboring bases. Thus, for a single methylation defect, the saturation current is strongly suppressed when cytosine is connected with guanine; for two defects, this suppression is smaller; for three methylations, the non-methylated base is the one that acts as a defect, hence the suppression of the saturation current will be larger when the cytosine has both hopping amplitudes to the neighboring bases enhanced by methylation. These results suggest the feasibility of using $I - V$ curves to develop biosensors for the purpose of diagnosis.

There also exist efforts aiming to examine the potentiality to utilize the charge transport characteristics of nucleic acids as a tool to probe several diseases or disorders. In Ref. [124], the $I - V$ characteristics of twenty seven single-stranded microRNA chains (with 21 to 23 nucleotides) related to the autism spectrum disorder have been studied. The authors classified the chains into five groups according to their conductivity (from high to negligible), suggesting that a kind of electronic biosensor can be developed to distinguish different profiles of autism disorders.

In Ref. [125], a similar treatment was employed to study DNA sequences related to the Huntington's Disease. A segment of the human chromosome 4p16.3 was modified by the addition of a variant number of CAG repeats, the number of which determines whether a person does or does not have Huntington's Disease; repeats smaller than 27 are normal; repeats between 27 and 35 are rarely associated with the disease, but it may expand in paternal transmission; repeats between 36 and 39 are associated with reduced penetration, so individuals may or may not develop the disease; 40 and above are associated with the disease [126]. The increasing presence of periodicity leads to enhanced transmission and thus to more efficient electronic transport. $I - V$ calculations revealed that the above-mentioned groups based on the number of repeats can be characterized by different value ranges for the saturation currents, indicating a promising method for identifying Huntington's disease.

7. Conclusions

This review was devoted to tight-binding (TB) modeling of nucleic acid sequences like DNA and RNA. We briefly presented the TB approach and discussed its various submodels: wire, ladder, extended ladder, fishbone (wire), and fishbone ladder. We addressed various types of orders (periodic, quasiperiodic, fractal) or disorder (diagonal, non-diagonal, random, methylation) and explained how these various types of order or disorder affect charge transport. We proceeded to a discussion of aperiodicity, quasicrystals and the mathematics of aperiodic substitutional sequences. Specifically, we discussed the notions of primitive substitutions, Perron–Frobenius eigenvalue, induced substitutions, and Pisot property. We explained how the energy structure of nucleic acid wires is affected by order or disorder. We also discussed the corresponding transmission coefficients, focusing on the effects of coupling the nucleic acids to external leads, and demonstrating that, apart from the internal degree of order or disorder of a given sequence, there are several other factors that can significantly affect their charge transport properties. We also discussed the effects that various types of order or disorder induce on the current–voltage curves and presented some efforts aiming to examine the potentiality to utilize the charge transport characteristics of nucleic acids as a tool to probe several diseases or disorders. The sensitivity that the results demonstrate regarding the choice of the nucleic acids sequence, the recruited models and parametrizations, the way the systems are coupled to external leads, the nature of the leads, the environmental conditions, etc, indicate that much work is needed in order to reach a thorough description of the effect the combination such a multitude of factors has on charge transport. Furthermore, other factors, such as the sequence geometry or the use of modified

nitrogenous bases could be potentially used to tailor the above-mentioned properties of nucleic acids and analogues.

Author Contributions: Conceptualization, K.L. and C.S.; methodology, K.L. and C.S.; software, K.L. and C.S.; validation, K.L. and C.S.; formal analysis, K.L. and C.S.; investigation, K.L. and C.S.; resources, K.L. and C.S.; data curation, K.L. and C.S.; writing—original draft preparation, K.L.; writing—review and editing, C.S.; visualization, K.L. and C.S.; supervision, C.S.; project administration, C.S.; funding acquisition, K.L. and C.S.

Funding: This research was supported by the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under the HFRI PhD Fellowship grant (GA no 260).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DNA	Deoxyribonucleic Acid
RNA	Ribonucleic acid
G	Guanine
A	Adenine
C	Cytosine
T	Thymine
U	Uracil
TB	Tight-Binding
WM	Wire Model
LM	Ladder Model
ELM	Extended Ladder Model
FM	Fishbone Model
FLM	Fishbone Ladder Model
DOS	Density of States
IDOS	Integrated Density of States

References

1. Watson, J.D.; Crick, F.H.C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.10.1038/171737a0. [[CrossRef](#)] [[PubMed](#)]
2. Page, C.C.; Moser, C.C.; Dutton, P.L. Mechanism for electron transfer within and between proteins. *Curr. Opin. Chem. Biol.* **2003**, *7*, 551–556.10.1016/j.cbpa.2003.08.005. [[CrossRef](#)] [[PubMed](#)]
3. Gray, H.B.; Winkler, J.R. Electron flow through metalloproteins. *Biochim. Biophys. Acta* **2010**, *1797*, 1563–1572.10.1016/j.bbabi.2010.05.001. [[CrossRef](#)] [[PubMed](#)]
4. Moser, C.C.; Ross Anderson, J.L.; Dutton, P.L. Guidelines for tunneling in enzymes. *Biochim. Biophys. Acta* **2010**, *1797*, 1573–1586.10.1016/j.bbabi.2010.04.441. [[CrossRef](#)] [[PubMed](#)]
5. Artés, J.M.; López-Martínez, M.; Díez-Pérez, I.; Sanz, F.; Gorostiza, P. Nanoscale charge transfer in redox proteins and DNA: Towards biomolecular electronics. *Electrochim. Acta* **2014**, *140*, 83–95.10.1016/j.electacta.2014.05.089. [[CrossRef](#)]
6. Kannan, A.M.; Renugopalakrishnan, V.; Filipek, S.; Li, P.; Audette, G.F.; Munukutla, L. Bio-Batteries and Bio-Fuel Cells: Leveraging on Electronic Charge Transfer Proteins. *J. Nanosci. Nanotechnol.* **2009**, *9*, 1665–1678.10.1166/jnn.2009.si03. [[CrossRef](#)] [[PubMed](#)]
7. Dandliker, P.J.; Holmlin, R.E.; Barton, J.K. Oxidative Thymine Dimer Repair in the DNA Helix. *Science* **1997**, *275*, 1465–1468.10.1126/science.275.5305.1465. [[CrossRef](#)]
8. Rajski, S.R.; Jackson, B.A.; Barton, J.K. DNA repair: models for damage and mismatch recognition. *Mutat. Res.* **2000**, *447*, 49–72.10.1016/s0027-5107(99)00195-5. [[CrossRef](#)]
9. Giese, B. Electron transfer through DNA and peptides. *Bioorg. Med. Chem.* **2006**, *14*, 6139–6143.10.1016/j.bmc.2006.05.067. [[CrossRef](#)]
10. Shih, C.T.; Cheng, Y.Y.; Wells, S.A.; Hsu, C.L.; Römer, R.A. Charge transport in cancer-related genes and early carcinogenesis. *Comput. Phys. Commun.* **2011**, *182*, 36–38.10.1016/j.cpc.2010.06.029. [[CrossRef](#)]

11. Endres, R.G.; Cox, D.L.; Singh, R.R.P. Colloquium: The quest for high-conductance DNA. *Rev. Mod. Phys.* **2004**, *76*, 195–214. [10.1103/revmodphys.76.195](#). [[CrossRef](#)]
12. Wohlgamuth, C.H.; McWilliams, M.A.; Slinker, J.D. DNA as a Molecular Wire: Distance and Sequence Dependence. *Anal. Chem.* **2013**, *85*, 8634–8640. [10.1021/ac401229q](#). [[CrossRef](#)] [[PubMed](#)]
13. Abouzar, M.H.; Poghossian, A.; Cherstvy, A.G.; Pedraza, A.M.; Ingebrandt, S.; Schöning, M.J. Label-free electrical detection of DNA by means of field-effect nanoplate capacitors: Experiments and modeling. *Phys. Status Solidi A* **2012**, *209*, 925–934. [10.1002/pssa.201100710](#). [[CrossRef](#)]
14. Ye, Y.J.; Shen, L.L. DFT approach to calculate electronic transfer through a segment of DNA double helix. *J. Comput. Chem.* **2000**, *21*, 1109–1117. [10.1002/1096-987X\(200009\)21:12<1109::AID-JCC7>3.0.CO;2-4](#). [[CrossRef](#)]
15. Ye, Y.J.; Jiang, Y. Electronic structures and long-range electron transfer through DNA molecules. *Int. J. Quantum Chem.* **2000**, *78*, 112–130. [10.1002/\(SICI\)1097-461X\(2000\)78:2<112::AID-QUA5>3.0.CO;2-5](#). [[CrossRef](#)]
16. Voityuk, A.A. Electronic couplings and on-site energies for hole transfer in DNA: Systematic quantum mechanical/molecular dynamic study. *J. Chem. Phys.* **2008**, *128*, 115101, doi:10.1063/1.2841421. [[CrossRef](#)]
17. Kubař, T.; Woiczikowski, P.B.; Cuniberti, G.; Elstner, M. Efficient Calculation of Charge-Transfer Matrix Elements for Hole Transfer in DNA. *J. Phys. Chem. B* **2008**, *112*, 7937–7947. [10.1021/jp801486d](#). [[CrossRef](#)]
18. Tassi, M.; Morphis, A.; Lambropoulos, K.; Simserides, C. RT-TDDFT study of hole oscillations in B-DNA monomers and dimers. *Cogent Phys.* **2017**, *4*, 1361077, doi:10.1080/23311940.2017.1361077. [[CrossRef](#)]
19. Artacho, E.; Machado, M.; Sánchez-Portal, D.; Ordejón, P.; Soler, J.M. Electrons in dry DNA from density functional calculations. *Mol. Phys.* **2003**, *101*, 1587–1594. [10.1080/0026897031000068587](#). [[CrossRef](#)]
20. Mehrez, H.; Anantram, M.P. Interbase electronic coupling for transport through DNA. *Phys. Rev. B* **2005**, *71*, 115405, doi:10.1103/PhysRevB.71.115405. [[CrossRef](#)]
21. Adessi, C.; Walch, S.; Anantram, M.P. Environment and structure influence on DNA conduction. *Phys. Rev. B* **2003**, *67*, 081405, doi:10.1103/PhysRevB.67.081405. [[CrossRef](#)]
22. Barnett, R.N.; Cleveland, C.L.; Landman, U.; Boone, E.; Kanvah, S.; Schuster, G.B. Effect of Base Sequence and Hydration on the Electronic and Hole Transport Properties of Duplex DNA: Theory and Experiment. *J. Phys. Chem. A* **2003**, *107*, 3525–3537. [10.1021/jp022211r](#). [[CrossRef](#)]
23. Cuniberti, G.; Craco, L.; Porath, D.; Dekker, C. Backbone-induced semiconducting behavior in short DNA wires. *Phys. Rev. B* **2002**, *65*, 241314, doi:10.1103/PhysRevB.65.241314. [[CrossRef](#)]
24. Roche, S.; Bicout, D.; Maciá, E.; Kats, E. Long Range Correlations in DNA: Scaling Properties and Charge Transfer Efficiency. *Phys. Rev. Lett.* **2003**, *91*, 228101, doi:10.1103/PhysRevLett.91.228101. [[CrossRef](#)]
25. Roche, S. Sequence Dependent DNA-Mediated Conduction. *Phys. Rev. Lett.* **2003**, *91*, 108101, doi:10.1103/PhysRevLett.91.108101. [[CrossRef](#)]
26. Palmero, F.; Archilla, J.F.R.; Hennig, D.; Romero, F.R. Effect of base-pair inhomogeneities on charge transport along the DNA molecule, mediated by twist and radial polarons. *New J. Phys.* **2004**, *6*, 13 doi:10.1088/1367-2630/6/1/013. [[CrossRef](#)]
27. Yamada, H. Localization of electronic states in chain models based on real DNA sequence. *Phys. Lett. A* **2004**, *332*, 65–73. [10.1016/j.physleta.2004.09.041](#). [[CrossRef](#)]
28. Apalkov, V.M.; Chakraborty, T. Electron dynamics in a DNA molecule. *Phys. Rev. B* **2005**, *71*, 033102, doi:10.1103/PhysRevB.71.033102. [[CrossRef](#)]
29. Klotsa, D.; Römer, R.A.; Turner, M.S. Electronic Transport in DNA. *Biophys. J.* **2005**, *89*, 2187–2198. [10.1529/biophysj.105.064014](#). [[CrossRef](#)]
30. Shih, C.T.; Roche, S.; Römer, R.A. Point-Mutation Effects on Charge-Transport Properties of the Tumor-Suppressor Gene *p53*. *Phys. Rev. Lett.* **2008**, *100*, 018105, doi:10.1103/PhysRevLett.100.018105. [[CrossRef](#)]
31. Joe, Y.S.; Lee, S.H.; Hedin, E.R. Electron transport through asymmetric DNA molecules. *Phys. Lett. A* **2010**, *374*, 2367–2373. [10.1016/j.physleta.2010.03.050](#). [[CrossRef](#)]
32. Yi, J. Conduction of DNA molecules: A charge-ladder model. *Phys. Rev. B* **2003**, *68*, 193103, doi:10.1103/PhysRevB.68.193103. [[CrossRef](#)]
33. Caetano, R.A.; Schulz, P.A. Sequencing-Independent Delocalization in a DNA-Like Double Chain with Base Pairing. *Phys. Rev. Lett.* **2005**, *95*, 126601, doi:10.1103/PhysRevLett.95.126601. [[CrossRef](#)]

34. Wang, X.F.; Chakraborty, T. Charge Transfer via a Two-Strand Superexchange Bridge in DNA. *Phys. Rev. Lett.* **2006**, *97*, 106602, doi:10.1103/PhysRevLett.97.106602. [[CrossRef](#)]
35. Ashcroft, N.W.; Mermin, N.D. *Solid State Physics*; Saunders College: Philadelphia, PA, USA, 1976.
36. Slater, J.C.; Koster, G.F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **1954**, *94*, 1498–1524. doi:10.1103/PhysRev.94.1498. [[CrossRef](#)]
37. Harrison, W.A. *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond (Dover Books on Physics)*, 2nd ed.; Dover Publications: New York, NY, USA, 1989.
38. Papaconstantopoulos, D.A.; Mehl, M.J. The Slater-Koster tight-binding method: A computationally efficient and accurate approach. *J. Phys. Condens. Matter* **2003**, *15*, 413–440, doi:10.1088/0953-8984/15/10/201. [[CrossRef](#)]
39. Foulkes, W.M.C. Tight-Binding Models and Coulomb Interactions for *s*, *p*, and *d* Electrons. In *Quantum Materials: Experiments and Theory*; Pavarini, E., Koch, E., van den Brink, J., Sawatzky, G., Eds.; Modeling and Simulation, Forschungszentrum Jülich: Jülich, Germany, 2016; Volume 6.
40. Molinari, L. Transfer matrices and tridiagonal-block Hamiltonians with periodic and scattering boundary conditions. *J. Phys. A Math. Gen.* **1997**, *30*, 983–997, doi:10.1088/0305-4470/30/3/021. [[CrossRef](#)]
41. Molinari, L. Spectral duality and distribution of exponents for transfer matrices of block-tridiagonal Hamiltonians. *J. Phys. A Math. Gen.* **2003**, *36*, 4081–4090. doi:10.1088/0305-4470/36/14/311. [[CrossRef](#)]
42. Cuniberti, G.; Maciá, E.; Rodríguez, A.; Römer, R.A. Tight-Binding Modeling of Charge Migration in DNA Devices. In *Charge Migration in DNA: Perspectives from Physics, Chemistry, and Biology*; Chakraborty, T., Ed.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 1–20, doi:10.1007/978-3-540-72494-0_1.
43. Albuquerque, E.L.; Fulco, U.L.; Freire, V.N.; Caetano, E.W.S.; Lyra, M.L.; de Moura, F.A.B.F. DNA-based nanobiostructured devices: The role of quasiperiodicity and correlation effects. *Phys. Rep.* **2014**, *535*, 139–209. doi:10.1016/j.physrep.2013.10.004. [[CrossRef](#)]
44. Zarea, M.; Berlin, Y.; Ratner, M.A. Effect of the reflectional symmetry on the coherent hole transport across DNA hairpins. *J. Chem. Phys.* **2017**, *146*, 114105. doi:10.1063/1.4978571. [[CrossRef](#)]
45. Simserides, C. A systematic study of electron or hole transfer along DNA dimers, trimers and polymers. *Chem. Phys.* **2014**, *440*, 31–41, doi:10.1016/j.chemphys.2014.05.024. [[CrossRef](#)]
46. Albuquerque, E.L.; Vasconcelos, M.S.; Lyra, M.L.; de Moura, F.A.B.F. Nucleotide correlations and electronic transport of DNA sequences. *Phys. Rev. E* **2005**, *71*, 021910. doi:10.1103/PhysRevE.71.021910. [[CrossRef](#)]
47. Lambropoulos, K.; Chatzieftheriou, M.; Morphis, A.; Kaklamanis, K.; Lopp, R.; Theodorakou, M.; Tassi, M.; Simserides, C. Electronic structure and carrier transfer in B-DNA monomer polymers and dimer polymers: Stationary and time-dependent aspects of a wire model versus an extended ladder model. *Phys. Rev. E* **2016**, *94*, 062403. doi:10.1103/PhysRevE.94.062403. [[CrossRef](#)]
48. Xu, M.; Endres, R.; Tsukamoto, S.; Kitamura, M.; Ishida, S.; Arakawa, Y. Conformation and Local Environment Dependent Conductance of DNA Molecules. *Small* **2005**, *1*, 1168–1172. doi:10.1002/sml.200500216. [[CrossRef](#)]
49. Iguchi, K. π -electrons in a single strand DNA: A phenomenological approach. *Int. J. Mod. Phys. B* **2004**, *18*, 1845–1910. doi:10.1142/S0217979204025051. [[CrossRef](#)]
50. Kundu, S.; Karmakar, S.N. Localization phenomena in a DNA double-helix structure: A twisted ladder model. *Phys. Rev. E* **2014**, *89*, 032719. doi:10.1103/PhysRevE.89.032719. [[CrossRef](#)]
51. Malakooti, S.; Hedin, E.; Joe, Y. Tight-binding approach to strain-dependent DNA electronics. *J. Appl. Phys.* **2013**, *114*, 014701. doi:10.1063/1.4812394. [[CrossRef](#)]
52. Varela, S.; Mujica, V.; Medina, E. Effective spin-orbit couplings in an analytical tight-binding model of DNA: Spin filtering and chiral spin transport. *Phys. Rev. B* **2016**, *93*, 155436. doi:10.1103/PhysRevB.93.155436. [[CrossRef](#)]
53. Díaz, E. Analysis of the interband optical transitions: Characterization of synthetic DNA band structure. *J. Chem. Phys.* **2008**, *128*, 175101. doi:10.1063/1.2901046. [[CrossRef](#)]
54. Maciá, E. Electronic structure and transport properties of double-stranded Fibonacci DNA. *Phys. Rev. B* **2006**, *74*, 245105. doi:10.1103/PhysRevB.74.245105. [[CrossRef](#)]
55. Maciá, E.; Roche, S. Backbone-induced effects in the charge transport efficiency of synthetic DNA molecules. *Nanotechnology* **2006**, *17*, 3002–3007. doi:10.1088/0957-4484/17/12/031. [[CrossRef](#)]
56. Lambropoulos, K.; Chatzieftheriou, M.; Morphis, A.; Kaklamanis, K.; Theodorakou, M.; Simserides, C. Unbiased charge oscillations in B-DNA: Monomer polymers and dimer polymers. *Phys. Rev. E* **2015**, *92*, 032725. doi:10.1103/PhysRevE.92.032725. [[CrossRef](#)]

57. Lambropoulos, K.; Vantaraki, C.; Bilia, P.; Mantela, M.; Simserides, C. Periodic polymers with increasing repetition unit: Energy structure and carrier transfer. *Phys. Rev. E* **2018**, *98*, 032412.10.1103/PhysRevE.98.032412. [[CrossRef](#)]
58. MacKinnon, A.; Kramer, B. The scaling theory of electrons in disordered solids: Additional numerical results. *Z. Phys. B* **1983**, *53*, 1–13.10.1007/BF01578242. [[CrossRef](#)]
59. Dwivedi, V.; Chua, V. Of bulk and boundaries: Generalized transfer matrices for tight-binding models. *Phys. Rev. B* **2016**, *93*, 134304.10.1103/PhysRevB.93.134304. [[CrossRef](#)]
60. Lambropoulos, K.; Simserides, C. Spectral and transmission properties of periodic 1D tight-binding lattices with a generic unit cell: An analysis within the transfer matrix approach. *J. Phys. Commun.* **2018**, *2*, 035013.10.1088/2399-6528/aab065. [[CrossRef](#)]
61. Datta, S. *Electronic Transport in Mesoscopic Systems*; Cambridge University Press: Cambridge, UK, 1995.10.1017/CBO9780511805776. [[CrossRef](#)]
62. Hawke, L.G.D.; Kalosakas, G.; Simserides, C. Electronic parameters for charge transfer along DNA. *Eur. Phys. J. E* **2010**, *32*, 291–305.10.1140/epje/i2010-10650-y. [[CrossRef](#)]
63. Capobianco, A.; Landi, A.; Peluso, A. Modeling DNA oxidation in water. *Phys. Chem. Chem. Phys.* **2017**, *19*, 13571–13578.10.1039/C7CP02029E. [[CrossRef](#)]
64. Sugiyama, H.; Saito, I. Theoretical Studies of GG-Specific Photocleavage of DNA via Electron Transfer: Significant Lowering of Ionization Potential and 5'-Localization of HOMO of Stacked GG Bases in B-Form DNA. *J. Am. Chem. Soc.* **1996**, *118*, 7063–7068.10.1021/ja9609821. [[CrossRef](#)]
65. Hutter, M.; Clark, T. On the Enhanced Stability of the Guanine-Cytosine Base-Pair Radical Cation. *J. Am. Chem. Soc.* **1996**, *118*, 7574–7577.10.1021/ja953370+. [[CrossRef](#)]
66. Zhang, H.; Li, X.Q.; Han, P.; Yu, X.Y.; Yan, Y. A partially incoherent rate theory of long-range charge transfer in deoxyribose nucleic acid. *J. Chem. Phys.* **2002**, *117*, 4578–4584.10.1063/1.1497162. [[CrossRef](#)]
67. Li, X.; Cai, Z.; Sevilla, M.D. Investigation of Proton Transfer within DNA Base Pair Anion and Cation Radicals by Density Functional Theory (DFT). *J. Phys. Chem. B* **2001**, *105*, 10115–10123.10.1021/jp012364z. [[CrossRef](#)]
68. Li, X.; Cai, Z.; Sevilla, M.D. Energetics of the Radical Ions of the AT and AU Base Pairs: A Density Functional Theory (DFT) Study. *J. Phys. Chem. A* **2002**, *106*, 9345–9351.10.1021/jp021322n. [[CrossRef](#)]
69. Shukla, M.K.; Leszczynski, J. A Theoretical Study of Excited State Properties of Adenine-Thymine and Guanine-Cytosine Base Pairs. *J. Phys. Chem. A* **2002**, *106*, 4709–4717.10.1021/jp014516w. [[CrossRef](#)]
70. Mantela, M.; Morphis, A.; Tassi, M.; Simserides, C. Lowest ionisation and excitation energies of biologically important heterocyclic planar molecules. *Mol. Phys.* **2016**, *114*, 709–718.10.1080/00268976.2015.1113313. [[CrossRef](#)]
71. Voityuk, A.A.; Jortner, J.; Bixon, M.; Rösch, N. Electronic coupling between Watson–Crick pairs for hole transfer and transport in desoxyribonucleic acid. *J. Chem. Phys.* **2001**, *114*, 5614–5620.10.1063/1.1352035. [[CrossRef](#)]
72. Ivanova, A.; Shushkov, P.; Rösch, N. Systematic Study of the Influence of Base-Step Parameters on the Electronic Coupling between Base-Pair Dimers: Comparison of A-DNA and B-DNA Forms. *J. Phys. Chem. A* **2008**, *112*, 7106–7114.10.1021/jp8031513. [[CrossRef](#)]
73. Migliore, A.; Corni, S.; Varsano, D.; Klein, M.L.; Di Felice, R. First, Principles Effective Electronic Couplings for Hole Transfer in Natural and Size-Expanded DNA. *J. Phys. Chem. B* **2009**, *113*, 9402–9415.10.1021/jp904295q. [[CrossRef](#)]
74. Shechtman, D.; Blech, I.; Gratias, D.; Cahn, J.W. Metallic Phase with Long-Range Orientational Order and No Translational Symmetry. *Phys. Rev. Lett.* **1984**, *53*, 1951–1953.10.1103/PhysRevLett.53.1951. [[CrossRef](#)]
75. International Union of Crystallography. Report of the Executive Committee for 1991. *Acta Crystallogr. A* **1992**, *48*, 922–946.10.1107/s0108767392008328. [[CrossRef](#)]
76. Maciá, E. On the Nature of Electronic Wave Functions in One-Dimensional Self-Similar and Quasiperiodic Systems. *ISRN Condens. Matter Phys.* **2014**, *2014*, 165943.10.1155/2014/165943. [[CrossRef](#)]
77. Kosuri, S.; Church, G.M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **2014**, *11*, 499–507.10.1038/nmeth.2918. [[CrossRef](#)]
78. Beaucage, S.; Caruthers, M. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981**, *22*, 1859–1862.10.1016/S0040-4039(01)90461-7. [[CrossRef](#)]

79. Palluk, S.; Arlow, D.H.; de Rond, T.; Barthel, S.; Kang, J.S.; Bector, R.; Baghdassarian, H.M.; Truong, A.N.; Kim, P.W.; Singh, A.K.; et al. De novo DNA synthesis using polymerase-nucleotide conjugates. *Nat. Biotechnol.* **2018**, *36*, 645–650.10.1038/nbt.4173. [[CrossRef](#)]
80. Perron, O. Zur Theorie der Matrices. *Math. Ann.* **1907**, *64*, 248–263, doi:10.1007/bf01449896. (In German) [[CrossRef](#)]
81. Frobenius, G. *Über Matrizen aus nicht Negativen Elementen*; Sitzungsberichte Königlich Preussischen Akademie Wissenschaften Berlin: Berlin, Germany, 1912; pp. 456–477. (In German)
82. Baake, M.; Grimm, U. *Aperiodic Order: Volume 1, A Mathematical Invitation*; Cambridge University Press: Cambridge, UK, 2013.
83. Lambropoulos, K.; Simserides, C. Periodic, quasiperiodic, fractal, Kolakoski, and random binary polymers: Energy structure and carrier transport. *Phys. Rev. E* **2019**, *99*, 032415.10.1103/PhysRevE.99.032415. [[CrossRef](#)]
84. Berthé, V.; Siegel, A. Basic notions on substitutions. In *Substitutions in Dynamics, Arithmetics and Combinatorics*; Fogg, N.P., Berthé, V., Ferenczi, S., Mauduit, C., Siegel, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 1–32, doi:10.1007/3-540-45714-3_1.
85. Canterini, V.; Siegel, A. Geometric Representation of Substitutions of Pisot Type. *Trans. Am. Math. Soc.* **2001**, *353*, 5121–5144. [[CrossRef](#)]
86. Bombieri, E.; Taylor, J.E. Which distributions of matter diffract? An initial investigation. *J. Phys. Colloq.* **1986**, *47*, 19–28.10.1051/jphyscol:1986303. [[CrossRef](#)]
87. Bombieri, E.; Taylor, J.E. Quasicrystals, Tilings, and Algebraic Number Theory: Some Preliminary Connections. In *The Legacy of Sonya Kovalevskaya*; Keen, L., Ed.; Contemporary Mathematics, American Mathematical Society: Providence, RI, USA, 1987; Volume 64, pp. 241–264.
88. Solomyak, B. Dynamics of Self-Similar Tilings. *Ergod. Theory Dyn. Syst.* **1997**, *17*, 695–738.10.1017/S0143385797084988. [[CrossRef](#)]
89. Luck, J.M.; Godrèche, C.; Janner, A.; Janssen, T. The nature of the atomic surfaces of quasiperiodic self-similar structures. *J. Phys. A Math. Gen.* **1993**, *26*, 1951–1999.10.1088/0305-4470/26/8/020. [[CrossRef](#)]
90. Kolář, M. New class of one-dimensional quasicrystals. *Phys. Rev. B* **1993**, *47*, 5489–5492.10.1103/PhysRevB.47.5489. [[CrossRef](#)]
91. Kolář, M.; Iochum, B.; Raymond, L. Structure factor of 1D systems (superlattices) based on two-letter substitution rules. I. delta (Bragg) peaks. *J. Phys. A Math. Gen.* **1993**, *26*, 7343–7366.10.1088/0305-4470/26/24/011. [[CrossRef](#)]
92. Sigler, L. *Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation*; Springer: New York, NY, USA, 2003.
93. Birch, J.; Severin, M.; Wahlström, U.; Yamamoto, Y.; Radnoczi, G.; Riklund, R.; Sundgren, J.E.; Wallenberg, L.R. Structural characterization of precious-mean quasiperiodic Mo/V single-crystal superlattices grown by dual-target magnetron sputtering. *Phys. Rev. B* **1990**, *41*, 10398–10407.10.1103/PhysRevB.41.10398. [[CrossRef](#)]
94. Fu, X.; Liu, Y.; Zhou, P.; Srirakool, W. Perfect self-similarity of energy spectra and gap-labeling properties in one-dimensional Fibonacci-class quasilattices. *Phys. Rev. B* **1997**, *55*, 2882–2889.10.1103/PhysRevB.55.2882. [[CrossRef](#)]
95. Maciá, E. Spectral Classification of One-Dimensional Binary Aperiodic Crystals: An Algebraic Approach. *Ann. Phys.* **2017**, *529*, 1700079, doi:10.1002/andp.201700079. [[CrossRef](#)]
96. Baake, M.; Grimm, W.; Mañibo, N. Spectral analysis of a family of binary inflation rules. *Lett. Math. Phys.* **2018**, *108*, 1783–1805.10.1007/s11005-018-1045-4. [[CrossRef](#)]
97. Prouhet, E. Mémoire sur quelques relations entre les puissances des nombres. *CR Acad. Sci. Paris* **1851**, *33*, 225. (In French)
98. Nagell, T.; Selberg, A.; Selberg, S.; Thalberg, K. (Eds.) *Selected Mathematical Papers of Axel Thue*; Universitetsforlaget: Oslo, Norway, 1977.
99. Morse, H.M. Recurrent Geodesics on a Surface of Negative Curvature. *Trans. Am. Math. Soc.* **1906**, *22*, 84–100.10.2307/1988844. [[CrossRef](#)]
100. Rudin, W. Some theorems on Fourier coefficients. *Proc. Am. Math. Soc.* **1959**, *10*, 855–895.10.1090/s0002-9939-1959-0116184-5. [[CrossRef](#)]
101. Shapiro, H.S. Extremal Problems for Polynomials and Power Series. Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1951.
102. Cantor, G. Über unendliche, lineare Punktmannigfaltigkeiten. *Math. Ann.* **1883**, *21*, 545–586. (In German) [[CrossRef](#)]

103. Vaidya, A.M.; Simon, H.; Garrett, P.; Shapiro, H.S.; Kolakoski, W.; Dapkus, F.; Gross, F.; Cohen, M.J.; Comtet, L.; Feller, E.H. Advanced Problems: 5300–5309. *Am. Math. Mon.* **1965**, *72*, 673–675, doi:10.2307/2313883. [[CrossRef](#)]
104. Oldenburger, R. Exponent trajectories in symbolic dynamics. *Trans. Am. Math. Soc.* **1939**, *46*, 453–466. doi:10.2307/1989933. [[CrossRef](#)]
105. Sing, B. Kolakoski sequences—An example of aperiodic order. *J. Non-Cryst. Solids* **2004**, *334*, 100–104. doi:10.1016/j.jnoncrysol.2003.11.021. [[CrossRef](#)]
106. Sing, B. Kolakoski-(2m,2n) are limit-periodic model sets. *J. Math. Phys.* **2003**, *44*, 899–912. doi:10.1063/1.1521239. [[CrossRef](#)]
107. Bovier, A.; Ghez, J.M. Remarks on the spectral properties of tight-binding and Kronig-Penney models with substitution sequences. *J. Phys. A Math. Gen.* **1995**, *28*, 2313–2324. doi:10.1088/0305-4470/28/8/022. [[CrossRef](#)]
108. Bellissard, J.; Bovier, A.; Ghez, J.M. Gap Labelling Theorems for One Dimensional Discrete Schrödinger operators. *Rev. Math. Phys.* **1992**, *4*, 1–37. doi:10.1142/S0129055X92000029. [[CrossRef](#)]
109. Bellissard, J.; Bovier, A.; Ghez, J.M. Discrete Schrödinger Operators with Potentials Generated by Substitutions. In *Differential Equations with Applications to Mathematical Physics*; Ames, W.F., Harrell, E.M., Herod, J.V., Eds.; Elsevier: Amsterdam, The Netherlands, 1993; Volume 192, pp. 13–23, doi:10.1016/S0076-5392(08)62368-1.
110. De Oliveira, B.P.W.; Albuquerque, E.L.; Vasconcelos, M.S. Electronic density of states in sequence dependent DNA molecules. *Surf. Sci.* **2006**, *600*, 3770–3774. doi:10.1016/j.susc.2006.01.081. [[CrossRef](#)]
111. De Moura, F.A.B.F.; Lyra, M.L.; Albuquerque, E.L. Electronic transport in poly(CG) and poly(CT) DNA segments with diluted base pairing. *J. Phys. Condens. Matter* **2008**, *20*, 075109. doi:10.1088/0953-8984/20/7/075109. [[CrossRef](#)]
112. De Moura, F.A.B.F.; Lyra, M.L.; de Almeida, M.L.; Ourique, G.S.; Fulco, U.L.; Albuquerque, E.L. Methylation effect on the ohmic resistance of a poly-GC DNA-like chain. *Phys. Lett. A* **2016**, *380*, 3559–3563. doi:10.1016/j.physleta.2016.07.069. [[CrossRef](#)]
113. Maciá Barber, E. *Aperiodic Structures in Condensed Matter: Fundamentals and Applications*; CRC Press: Boca Raton, FL, USA, 2008. doi:10.1201/9781420068283. [[CrossRef](#)]
114. Maciá, E.; Triozon, F.; Roche, S. Contact-dependent effects and tunneling currents in DNA molecules. *Phys. Rev. B* **2005**, *71*, 113106. doi:10.1103/PhysRevB.71.113106. [[CrossRef](#)]
115. Mason, J.; Handscomb, D.C. *Chebyshev Polynomials*; Chapman and Hall/CRC: London, UK, 2002.
116. Malakooti, S.; Hedin, E.R.; Kim, Y.D.; Joe, Y.S. Enhancement of charge transport in DNA molecules induced by the next nearest-neighbor effects. *J. Appl. Phys.* **2012**, *112*, 094703. doi:10.1063/1.4764310. [[CrossRef](#)]
117. Landauer, R. Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction. *IBM J. Res. Dev.* **1957**, *1*, 223–231. doi:10.1147/rd.13.0223. [[CrossRef](#)]
118. Buttiker, M. Symmetry of electrical conduction. *IBM J. Res. Dev.* **1988**, *32*, 317–334. doi:10.1147/rd.323.0317. [[CrossRef](#)]
119. Roche, S.; Maciá, E. Electronic Transport and Thermopower in Aperiodic DNA Sequences. *Mod. Phys. Lett. B* **2004**, *18*, 847–871. doi:10.1142/S021798490400744X. [[CrossRef](#)]
120. Plazas, C.A.; Fonseca-Romero, K.M.; Rey-González, R.R. Insulator-to-Semiconductor-to-Conductor Phase-Like Transition of DNA Chains. *J. Nanosci. Nanotechnol.* **2018**, *18*, 5042–5048. doi:10.1166/jnn.2018.15339. [[CrossRef](#)]
121. Kundu, S.; Karmakar, S.N. Conformation dependent electronic transport in a DNA double-helix. *AIP Adv.* **2015**, *5*, 107122, doi:10.1063/1.4934507. [[CrossRef](#)]
122. Joe, Y.S.; Lee, S.H.; Hedin, E.R.; Kim, Y.D. Temperature and Magnetic Field Effects on Electron Transport Through DNA Molecules in a Two-Dimensional Four-Channel System. *J. Nanosci. Nanotechnol.* **2013**, *13*, 3889–3896. doi:10.1166/jnn.2013.7206. [[CrossRef](#)]
123. De Almeida, M.L.; Oliveira, J.I.N.; Lima Neto, J.X.; Gomes, C.E.M.; Fulco, U.L.; Albuquerque, E.L.; Freire, V.N.; Caetano, E.W.S.; de Moura, F.A.B.F.; Lyra, M.L. Electronic transport in methylated fragments of DNA. *Appl. Phys. Lett.* **2015**, *107*, 203701, doi:10.1063/1.4936099. [[CrossRef](#)]
124. Oliveira, J.I.N.; Albuquerque, E.L.; Fulco, U.L.; Mauriz, P.W.; Sarmiento, R.G.; Caetano, E.W.S.; Freire, V.N. Conductance of single microRNAs chains related to the autism spectrum disorder. *EPL* **2014**, *107*, 68006, doi:10.1209/0295-5075/107/68006. [[CrossRef](#)]

125. Sarmento, R.G.; Silva, R.N.O.; Madeira, M.P.; Frazão, N.F.; Sousa, J.O.; Macedo-Filho, A. Electronic Transport in Single-Stranded DNA Molecule Related to Huntington's Disease. *Braz. J. Phys.* **2018**, *48*, 155–159.10.1007/s13538-018-0554-z. [[CrossRef](#)]
126. Walker, F.O. Huntington's disease. *Lancet* **2007**, *369*, 218–228.10.1016/S0140-6736(07)60111-1. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).