

## Ultrafast excitonic dynamics in DNA: Bridging correlated quantum dynamics and sequence dependence

Dennis Herb,<sup>\*</sup> Mirko Rossini<sup>✉, \*</sup> and Joachim Ankerhold<sup>✉</sup>

*Institute for Complex Quantum Systems, Ulm University, 89069 Ulm, Germany  
and Center for Integrated Quantum Science and Technology (IQST), Ulm-Stuttgart, Germany*



(Received 5 March 2024; accepted 21 May 2024; published 25 June 2024)

After photoexcitation of DNA, the excited electron (in the LUMO) and the remaining hole (in the HOMO) localized on the same DNA base form a bound pair, called the Frenkel exciton, due to their mutual Coulomb interaction. In this study, we demonstrate that a tight-binding (TB) approach, using TB parameters for electrons and holes available in the literature, allows us to correlate relaxation properties, average charge separation, and dipole moments to a large ensemble of double-stranded DNA sequences (all 16384 possible sequences with 14 nucleobases). This way, we are able to identify a relatively small subensemble of sequences responsible for long-lived excited states, high average charge separation, and high dipole moment. Further analysis shows that these sequences are particularly T rich. By systematically screening the impact of electron-hole interaction (Coulomb forces), we verify that these correlations are relatively robust against finite-size variations of the interaction parameter, not directly accessible experimentally. This methodology combines simulation methods from quantum physics and physical chemistry with statistical analysis known from genetics and epigenetics, thus representing a powerful bridge to combine information from both fields.

DOI: [10.1103/PhysRevE.109.064413](https://doi.org/10.1103/PhysRevE.109.064413)

### I. INTRODUCTION

Photoexcited energy and charge transfer in biologically relevant systems is one of the fundamental processes in nature and the basis for all life on earth. Exploring its microscopic mechanisms has been a major field of research over the past decades in biology, chemistry, and physics. One prominent example is the formation of electron-hole pairs (excitons) on complex molecular structures under external radiation. It not only serves as a model system to better understand molecular energy transport, e.g., in light-harvesting complexes in photosynthesis [1–5] but has also triggered the design of nature-inspired tools such as efficient organic solar cells [6,7], energy storage devices [8], and organic photodetectors and phototransistors [6].

In particular, photoexcited charge dynamics on DNA, the central repository of genetic information in the cell, has been intensively studied because of its relevance to induce DNA lesions, alter DNA binding properties, and trigger carcinogenic mutations. As a result, a substantial body of literature has appeared proposing and reviewing charge transfer processes in DNA [9–16] and proposing applications ranging from genetics, e.g., understanding DNA damage and repair [17–19] and distinguishing pathogenic from nonpathogenic mutations [20], to nanotechnology, e.g., designing nanosensors, nanocircuits, and molecular wires [21,22].

DNA is known to be photostable with rapidly decaying excitations. Improved spectroscopic methods, such as pump-probe techniques, have made it possible to experimentally

study the dynamics of ultrafast excited states in DNA [23,24]. These states can be generated naturally by ultraviolet (UV) radiation, for example as a result of direct exposure to sunlight. In both cases, an electron is excited from the highest occupied molecular orbital (HOMO) to the lowest unoccupied one (LUMO), leaving a hole in the HOMO. These charge carriers interact as a result of attractive Coulomb forces and, when bound to a single nitrogenous DNA base, form an electron-hole pair known as a Frenkel exciton. As a localized excitation, the Frenkel exciton plays an important role in processes such as energy transfer in DNA-based systems [25–27].

A theoretical description of the propagation of charges through DNA is rather challenging. It is sensitive not only to extrinsic effects, such as the surrounding aqueous solution, but also to intrinsic influences, such as fluctuations in the dynamical structure [28–30], including the phosphate backbone. As a consequence, the complexity of the DNA molecule and its environment renders atomistic *ab initio* calculations even in equilibrium not feasible for any structures beyond 3–4 nucleotides, that is, far from capturing long-range charge transfer in double-strand aggregates. Therefore, so-called tight-binding models have been developed as powerful alternatives [30,31].

Here, we follow this strategy with the goal of combining two methodologies, namely, mesoscopic modeling that captures the essential features of excited-state charge dynamics and a statistical analysis of sizable sets of double-strand DNA sequences. With model parameters taken from calculations developed in Refs. [32,33], we achieve sufficiently accurate results with the required high numerical efficiency to screen larger ensembles.

<sup>\*</sup>These authors contributed equally to this work.

Most of the previous studies have focused on the hole transfer along the DNA strand for various reasons: First, experiments aimed at investigating the motion of charges along DNA often make use of DNA oxidation [19,34] (lack of an electron) to create an initial charge carrier; second, *ab initio* calculations (such as DFT) proved to be more accurate in retrieving low-energy structures and properties of HOMO molecular orbitals. However, more recent literature [33] reports a comprehensive collection of structural data such as binding energies and overlap integrals, which provides the basis to develop models including both electron and hole motion.

In this work, we go beyond by promoting a simplistic, yet relatively accurate, unified model to analyze the dynamics of electrons and holes along DNA double strands. In contrast to similar models discussed in the literature, we refine the model by adding two crucial ingredients, namely the Coulomb interaction among both charge carriers and the finite lifetime of excitons due to local charge recombination.

Our hypothesis for this study is that there is a direct relation between effective exciton lifetime and, induced by the mobility of electron and hole along the DNA, their mean spatial separation; the latter depends on the specificities of the DNA sequences encoded in the tight-binding parameters. If this were true, at least for a subensemble of DNA sequences of certain length, the consequences could be far reaching: A relatively stable charge distribution resulting from charge delocalization within pockets of the DNA molecule would result in the emergence of electrical dipole forces that, in turn, may affect local conformational structures of the DNA complex and its local electrical properties. Both phenomena could in principle have an impact on the likelihood of binding of regulatory proteins to the DNA [35], a crucial aspect of gene regulation and transcriptional processes.

In order to test this hypothesis, we combine efficient quantum dynamical simulations with the screening of a large set of DNA sequences. More specifically, our systematic analysis using advanced computational implementation includes up to 16 384 possible DNA double-strand sequences of length seven bases such that, based on a statistical analysis, we are able to combine information from quantum physics or physical chemistry with that from genetics or epigenetics. This way, we aim to identify DNA sequences which may be of particular interest to be explored further in experimental settings of genetics and epigenetics.

Remarkably, we first find that our predictions in terms of exciton lifetime and spatial separation for the most stable species are in good agreement with data obtained in previous experiments [23,24,36–38]. Our results show further that it is only a small subset of sequences that stands out as responsible for relatively long-lived excited states and high-average charge separation. The role of the electron-hole interaction is very decisive: Assuming that the Coulomb interaction between charge carriers is completely screened (no interaction) leads statistically to significantly different subsets compared to the case when it is finite. However, if it is taken as finite and is varied in a reasonable window of values [32,39], the interesting subset turns out to be relatively robust. This is important since the precise value of this Coulomb interaction is not known and is experimentally not directly accessible.

Although still in the embryonic stage, we hope that the presented methodology may trigger future research by, for example, developing even more elaborate tight-binding models and optimized numerical tools based on AI [40,41] for even larger sets of DNA.

The present paper is organized as follows. In Sec. II we provide a description of the model used in our study. In Sec. III A we systematically investigate exciton lifetime and average charge separation without electron-hole interaction. We introduce a Coulomb interaction between the charges in Sec. III B and present its effect on exciton lifetime and average charge separation in Sec. III C. We conclude in Sec. IV.

## II. THEORETICAL MODELING

We consider each of the four DNA bases (A, C, G, T) as a site in a (two-dimensional) tight-binding lattice (see Fig. 1) and numerate the sites by an index tuple  $(ij)$  where  $i \in \{1, 2\}$  refers to the upper or lower strand, and  $j \in \{1, \dots, N\}$  denotes the base index for each DNA strand with  $N$  being the number of bases per strand. Two charges (electron and hole) are created by an initial excitation. This is reflected in the Hilbert space of the system  $\mathcal{H}_S = \mathcal{H}_e \otimes \mathcal{H}_h$  with basis states  $|e_{ij}h_{kl}\rangle$  that describe all possible electron-hole configurations on the lattice.

The overlap of the LUMO (and HOMO) orbitals of neighboring DNA bases allows an electron (and a hole) to move and spread along the DNA molecule. The Hamiltonian for both the electron and hole dynamics can be described by two terms,  $H_{\text{self}}^p$  and  $H_{\text{trans}}^p$  (where  $p \in \{e, h\}$  describes either electron or hole), containing, respectively, the on-site and transfer energies:

$$\begin{aligned} H_{\text{self}}^p &= \sum_{j=1}^N (\epsilon_{1j}^p |p_{1j}\rangle \langle p_{1j}| + \epsilon_{2j}^p |p_{2j}\rangle \langle p_{2j}|) \quad (1) \\ H_{\text{trans}}^p &= \sum_{j=1}^{N-1} (t_{1j}^p |p_{1j}\rangle \langle p_{1,j+1}| + t_{2j}^p |p_{2j}\rangle \langle p_{2,j+1}|) + \text{H.c.} \\ &+ \sum_{j=1}^N h_j^p |p_{1j}\rangle \langle p_{2j}| + \text{H.c.} \\ &+ \sum_{j=1}^{N-1} (r_j^{+p} |p_{1j}\rangle \langle p_{2,j+1}| + r_j^{-p} |p_{1,j+1}\rangle \langle p_{2,j}|) + \text{H.c.} \quad (2) \end{aligned}$$

This model is commonly known as the extended ladder model (ELM) [30]. We can now build the Hamiltonian for the electron and hole dynamics along the DNA double strand as  $H_e = (H_{\text{self}}^e + H_{\text{trans}}^e) \otimes \mathbb{I}_h$  and  $H_h = \mathbb{I}_e \otimes (H_{\text{self}}^h + H_{\text{trans}}^h)$ , where  $\mathbb{I}_{e(h)}$  describes the identity on the Hilbert space  $\mathcal{H}_{e(h)}$  of the electron (hole).

To account for electron-hole attraction due to the Coulomb forces, we introduce two-particle interaction terms of the form:

$$H_{\text{int}} = \sum_{i,k=1}^2 \sum_{j,l=1}^N U_{ijkl} |e_{ij}h_{kl}\rangle \langle e_{ij}h_{kl}|. \quad (3)$$

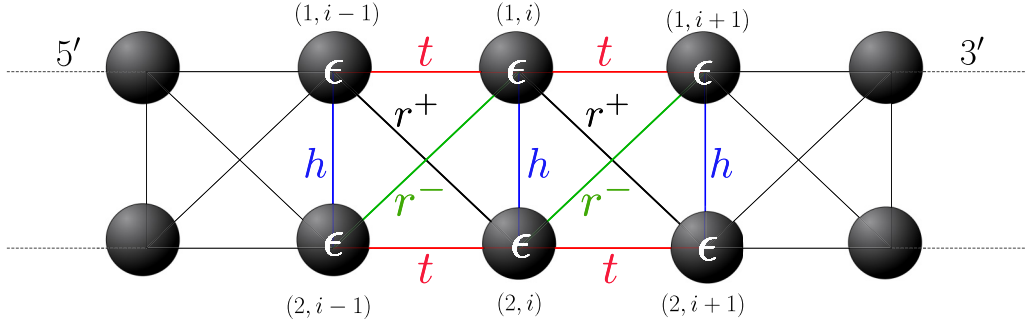


FIG. 1. Extended ladder model (ELM) with on-site energies  $\varepsilon$ , horizontal intrastrand TB parameters  $t$ , vertical interstrand TB parameters  $h$  and diagonal interstrand TB parameters  $r_{\pm}$ .

Normally, to account for these many-body interaction terms requires investigating an exponentially large Hilbert space with respect to the dimension of the system (number of DNA bases). In this case, however, we are able to enforce the constraint that there is always either one exciton or no exciton in the system, allowing us to investigate only a quadratically large Hilbert space.

More specifically, for the interaction matrix elements we assume an algebraic spatial decay according to  $U_{ijkl} = J/(1 + r/r_0)$  with  $r_0 = 1.0 \text{ \AA}$  following [26,27]. Since two neighboring DNA bases are separated (in the intrastrand and interstrand direction) by  $D \sim 3.4 \text{ \AA}$ , we simplify in the tight-binding description the distance dependence to be of the form  $r/r_0 \approx (D/r_0) \cdot (|i - k| + |j - l|)$  and only include interactions up to next-neighbor bases in agreement with a short-range approximation used in the literature, i.e.,

$$U_{ijkl} = \begin{cases} J & \text{if } (|i - k| + |j - l|) = 0 \\ J/(1 + D/r_0) & \text{if } (|i - k| + |j - l|) = 1. \\ 0 & \text{else} \end{cases} \quad (4)$$

The first case describes the two charges being located at the same base, while the second holds when electron and hole are sitting on neighboring bases. In this way, we arrive at the interacting TB Hamiltonian for the ELM:  $H = H_e + H_h + H_{\text{int}}$ . The parameters for this model are taken from Hückel TB model calculations in Ref. [33] and are given in the Supplemental Material (SM) [42].

Typical energies that are needed to create an excitation on the DNA are given by the energy difference between HOMO and LUMO and range between 3.5 eV and 4.5 eV [33]. The typical temperature  $T \approx 310 \text{ K}$  for DNA under ambient conditions instead corresponds to energies in the range of the meV, much too small to spontaneously induce excitons, thus reflecting the stability of native DNA. Hence, external stimuli give rise to excitonic phenomena. In this sense, in a minimal setting, intrinsic loss mechanisms, i.e., exciton recombination, are the only relevant impact of residual degrees of freedom to which energy from the charge sector can be dissipated. In the spirit of quantum optical modeling, these can be described by operators of the form

$$A_{ij} = \sqrt{\gamma_{ij}} |0\rangle \langle e_{ij} h_{ij}| \quad (5)$$

with  $|e_{ij} h_{ij}\rangle$  denoting an electron  $e$  (hole  $h$ ) at site  $i$  on strand  $j$  and local (monomer) transition rates  $\gamma_{ij}$ . They act as

dissipators with the tendency to destroy quantum coherences in a Lindblad-type equation for the reduced density operator of the exciton, namely,

$$\frac{d}{dt} \rho(t) = -\frac{i}{\hbar} [H, \rho] + \sum_{i,j} A_{ij} \rho A_{ij}^\dagger - \frac{1}{2} (A_{ij}^\dagger A_{ij} \rho + \rho A_{ij}^\dagger A_{ij}).$$

Here, the total Hamiltonian enters the deterministic part while the second part describes local exciton recombination to the monomer ground state, see Fig. 2 (for further details see below).

The above time-evolution equation is now represented in the above site basis. Time-dependent expectation values are obtained from the respective density operator  $\rho(t)$  at time  $t$ . For example, the population in the electron (hole) sector only follows from

$$P_{ij}^{p'}(t) = \langle p_{ij} | \rho_p(t) | p_{ij} \rangle$$

with the partial trace  $\rho_p = \text{Tr}_{p'}[\rho]$ ,  $p \neq p' \in \{e, h\}$ . The exciton population at the same site is given by

$$P_{ij}^{\text{ex}}(t) = \langle e_{ij} h_{ij} | \rho(t) | e_{ij} h_{ij} \rangle.$$

To incorporate the possibility for electron-hole recombination, the underlying Hilbert space is expanded by adding the state  $|0\rangle$  to the set of above basis states. It represents the ground state of the DNA sequence, as illustrated in Fig. 2, with a population given by  $P_0(t) = \langle 0 | \rho(t) | 0 \rangle$ .

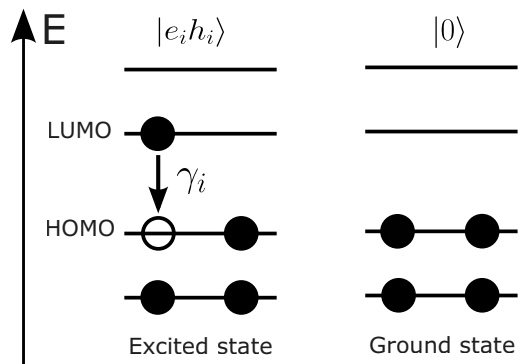


FIG. 2. HOMO and LUMO for excited state and ground state of the DNA molecule. Whenever electron and hole are located at the same base they recombine with a rate  $\gamma_i$  such that the DNA relaxes to its ground state.

Based on this modeling, the effective exciton lifetime as well as the charge separation on the lattice is determined. For the latter, we introduce the charge separation operator  $\hat{d}$

$$\langle e_{ij}h_{kl} | \hat{d} | e_{ij}h_{kl} \rangle = D \cdot (|i - k| + |j - l|)$$

whose time average is an estimate for the amount of charge separation during the dynamics, i.e.,

$$\bar{d} = \frac{1}{T} \int_0^T \text{Tr}\{\rho(t) \hat{d}\} dt.$$

When the electron and hole are temporarily separated in space, they form an electric dipole with an effective dipole moment, which is directly determined by the average charge separation.

### III. RESULTS

#### A. Excited-state lifetime and average dipole moment

In this section, we explore the exciton dynamics with particular focus on lifetime and spatial separation, which in turn allows us to reveal their correlations with respect to DNA sequences. Our analysis of exciton lifetimes as such appears to be of relevance in the context of pump-probe experiments presented in the literature [10,23,43], thus addressing the yet not fully understood topic of relatively long-lasting charge-hole excitations (in the range of ps) after photoexciting specific DNA strands. The second observable, related to exciton mobility, provides information about the charge distribution along DNA strands and sheds light on the question over which length scales (number of bases) quantum coherences may survive, a topic that has been discussed but not yet conclusively answered in previous work [36,44]. In fact, possible implications are far reaching since long-range spatial separation of electron-hole pairs may give rise to relatively strong electrical dipole forces, which in turn may open pathways for conformational changes or modify the binding affinity for CpG methylation or other enzymatic reactions.

More specifically, we conducted a systematic screening of lifetime and charge separation for all double-strand DNA sequences ranging between  $N = 3$  and  $N = 7$  bases in each strand (i.e., total number of bases between 6 and 14), thus covering up to  $S_7 \equiv 4^7 = 16,384$  DNA sequences. Each sequence has its individual footprint in form of a specific set of parameters for the ELM, see Eqs. (2), (3), obtained from atomistic calculations [32,33]. Monomer exciton recombination rates in Eq. (5) are not known theoretically and have thus been inferred from pump-probe experiments in Ref. [23]. There, values have been reported in the range of hundreds of femtoseconds for the recombination of an exciton confined within a single DNA base. We have then chosen to assume a constant rate for all modeled bases corresponding to an average lifetime of 300 fs. Variations on these time scales do not significantly affect the quantitative results for lifetime and charge separation, and thus have only a marginal impact on relative dependencies between different DNA sequences (see SM [42]).

We observed that exciton lifetimes vary substantially within the total set  $S_N$  of DNA sequences, with several sequences showing lifetimes on the order of picoseconds (e.g.,

>2.5 ps for TTTTAA) versus many other sequences showing lifetimes comparable to the monomer lifetime of approximately 300 fs (see SM [42]). The range between the longest and shortest lifetimes spanned nearly one order of magnitude. Data of the top 30 sequences (ordered accordingly in a subset  $S_{\text{time}} \subset S_7$ ) are shown in Fig. 3(a) with typical lifetimes of the order of 2 ps and a variation of approximately 25%.

We also note that even though sequences such as TTTTAA–TTAAAA or GCCCCC–GGGGGC are the same (since the opposite strand of a DNA sequence is uniquely identified), they yield different results in our setting because the initial excitation occurs on different bases (top left vs bottom right base excitation).

The charge separation analysis revealed a wide range of results among different sequences, with some exhibiting negligible separation, while others displaying charge separations exceeding 10 Å, equivalent to an average separation of approximately three base pairs. Following a similar analysis to the one above, we identified an ordered set  $S_{\text{space}} \subset S_7$  of 30 sequences with variations of approximately 20%, see Fig. 3(b). Notably, charge separations within  $D_{\text{large}}$  qualitatively match previous findings in Refs. [24,36–38], where a mixed diffusion mechanism for charge transfer along DNA strands is proposed in which quantum delocalization plays a relevant role in patches of up to 3–4 base pairs, i.e., approximately  $\sim 10$ – $13$  Å.

Now, based on this analysis we turn to consider correlations between lifetime and charge separation. Indeed, within the total set  $S_7$ , we find robust correlations with correlation factor exceeding  $\chi \gtrsim 0.95$  (see SM [42]). However, when considering only the respective top subsets  $S_{\text{time}}$  and  $S_{\text{space}}$  a remarkably weaker correlation appears: less than  $\chi \approx 0.4$ . For these specific sequences, a strong correspondence between lifetime and dipole moment is not clearly recognizable. The first finding is intuitive: Excitonic recombination is substantially inhibited when electron and hole sit on different sites (bases) as they are then efficiently screened by the surrounding background. This implies that larger charge mobility along the chain leads in general to a longer excitonic lifetimes. The second finding suggests though that details are more subtle and, among the top sequences, additional mechanisms besides the mere particle separation play a crucial role such as on-site energies.

A general trend can be found when relative distributions of A-T versus G-C base pairs in  $S_7$  are studied (see Fig. 4). Sequences containing regions rich in G-C base pairs are more likely to feature long-lived excitations than sequences rich in A-T base pairs. In all these data, increasing (decreasing) monomer relaxation rates within a window around 300 fs leads to a mere shift of absolute values towards smaller (larger) exciton relaxation times but overall preserving relative relations.

#### B. Electron-hole interaction and charge distribution

While single charge transfer along DNA has been largely investigated [32,39,45], the role of charge-charge interactions has received much less attention. In the context of the current work, this is particularly true for electron-hole correlations due to attractive Coulomb forces. In fact, to include this



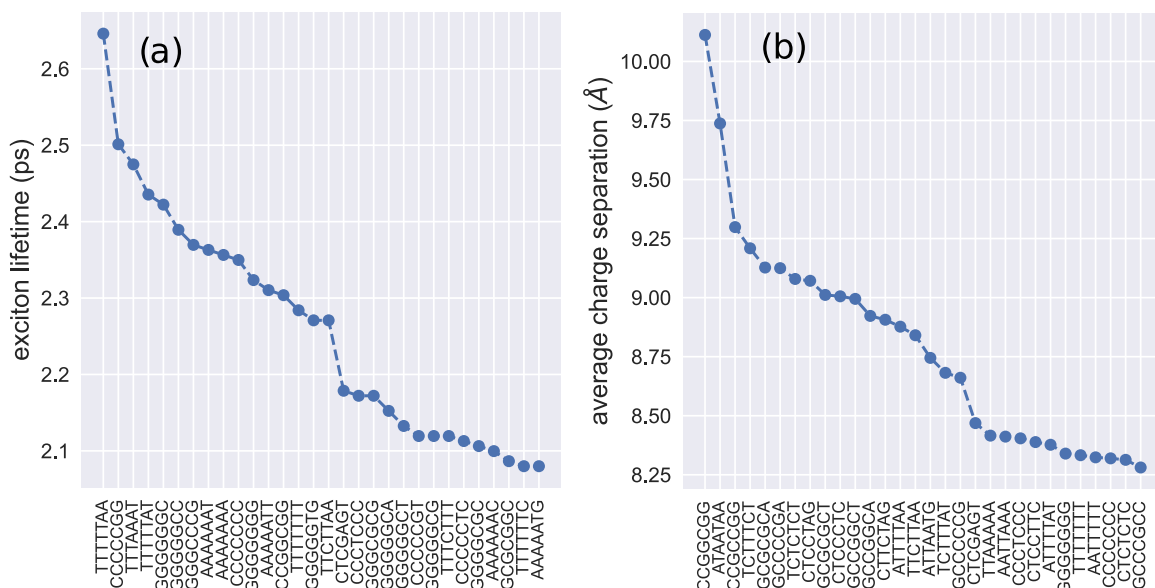


FIG. 3. (a) Top 30 DNA sequences with the longest exciton lifetime (subset  $S_{\text{time}}$ ) and (b) the highest average charge separation (subset  $S_{\text{space}}$ ) from the analysis of all possible double-strand sequences with seven DNA base pairs for each strand  $S_7$ . We assume a uniform relaxation rate  $\gamma = 2 \text{ meV}/\hbar$  and perfect charge screening (no electron-hole interaction  $J = 0$ ). The letters denote the DNA bases on the upper strand in the 5'-3' direction.

interaction may be decisive for the dynamics of excitons as it implies to account for two competing processes: electron-hole interaction with the tendency to localize charge pairs on individual sites (nucleotides) and the mobility of both charge carriers in LUMO and HOMO with the tendency for delocalization. Details thus depend on energy profiles, strength of hybridization, and screening properties (range of Coulomb interaction) of a specific DNA sequence.

Here, we start this analysis considering the impact of electron-hole correlations within the so-called wire model [31]. The latter is a simplified version of the ELM in which the bases forming a base pair are treated as one larger molecule composed of two adjacent molecules with electronic overlap such that HOMO and LUMO are localized only on one of the bases of a pair [33]. This model has been previously studied in the literature and further motivated by Mantela *et al.* [46,47], thus providing benchmark data that we used to validate our modeling. It further allows to achieve a transparent dynamical

understanding of electron-hole interaction which will serve as the basis to consider it in the full ELM in the next section.

Coulomb forces are usually considered as well-established and well-known interactions, whose intensity can be easily evaluated given specific boundary conditions. However, the situation in macromolecular aggregates such as DNA is more complex, though, because the strength of this interaction can vary due to several factors inherent to surrounding degrees of freedom. An aqueous environment acts as a polarizable medium (water has a relatively high dielectric constant of about 80 at physiological temperatures), thus influencing and screening the Coulombic forces in various ways. Additionally, factors such as local hydration and the presence of counterions can contribute to variations in the effective dielectric constant of the surroundings. Structural deformations, which encompass mechanical stress, temperature changes, and interactions with surrounding biomolecules or macromolecules, as well as conformational changes such as bending or twisting of

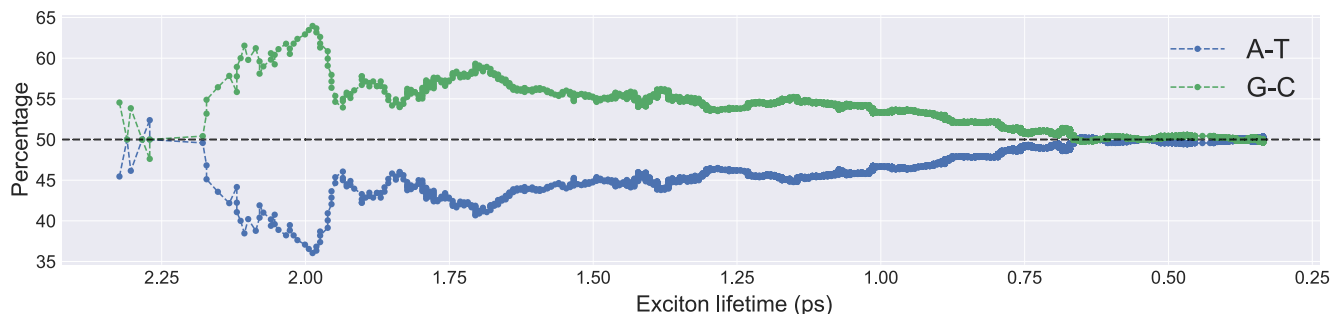


FIG. 4. Relative distribution of base-pairs (A-T pair and G-C pair) in sequences of seven base pairs over the expected exciton lifetime. We obtain the distribution at time  $t$  by counting the base pairs in all sequences with a lifetime greater than  $t$  and dividing by the number of base pairs. Since it is necessary to average over several sequences before the results become relevant, we have omitted the first ten sequences with the largest exciton lifetimes.

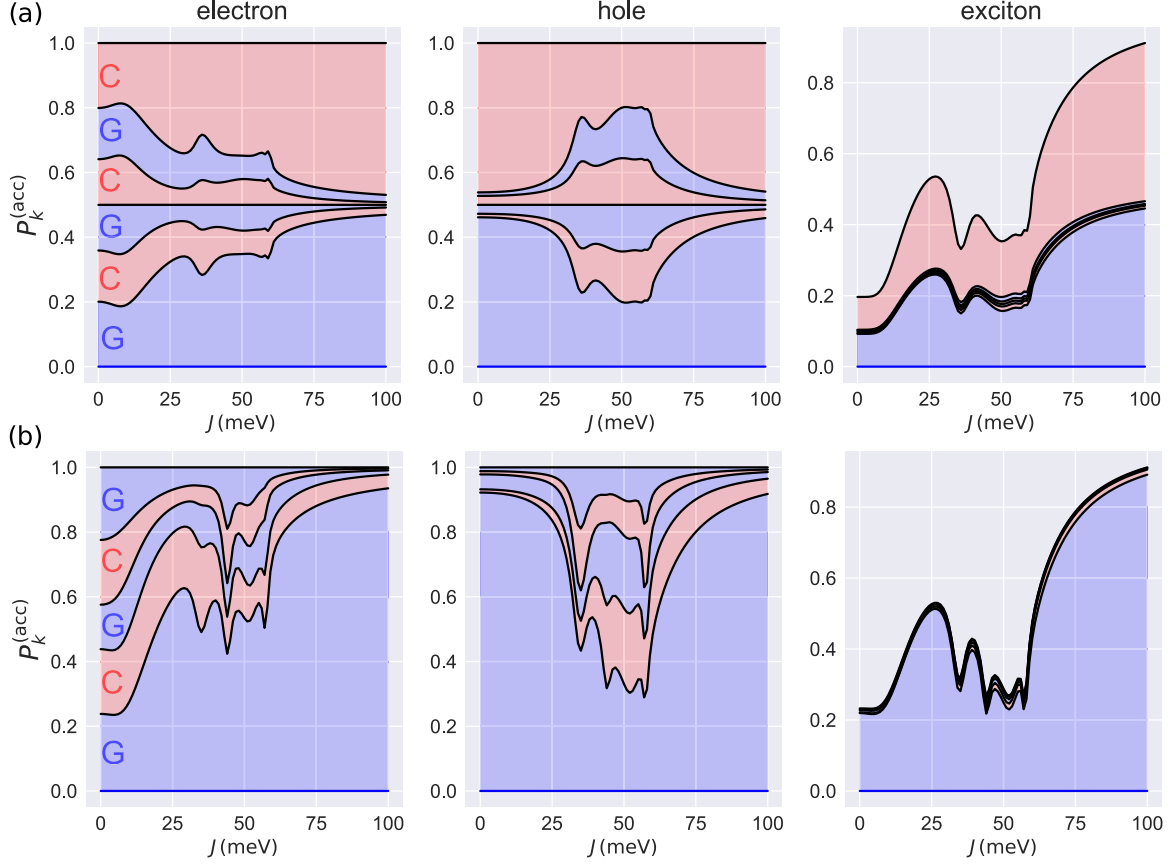


FIG. 5. Accumulated average populations  $P_k^{(acc)} = \sum_{i=0}^k (P_i)$  of (a) GCGCGC and (b) GCGCG with initial excitation on the left base pair according to the wire model description. The hopping parameters are as follows: GC  $-10$  meV, CG  $-8$  meV for the electron and GC  $-10$  meV, CG  $50$  meV for the hole.

the DNA helix, can further impact the spatial arrangement of the charge carriers and modify the strength and range of the Coulombic forces. Consequently, the strength of the Coulomb interaction within DNA can exhibit considerable variability. Due to these reasons, in our study, we investigate a comprehensive range of interaction strengths, ranging from considerably weaker to stronger with respect to the DNA hopping energies.

More specifically, we investigate the effect of electron-hole interaction on time-averaged populations for GC periodic sequences, modeled as one-dimensional chains whose base units model the G-C DNA nucleobase pair. A similar analysis is shown in the SM in Fig. S3 [42], where we report results for a uniform G sequence. Since the base units of the chain are the same (G-C nucleobase pairs), we note that the difference between the alternating GC sequence and the uniform G sequence lies in the staggered (or uniform) site-site overlap parameters along the chain. Accordingly, within this wire model, this sequence becomes degenerate with respect to on-site energies but not with respect to coupling energies as the 5'-3' DNA directionality implies that the G-C molecular overlap is different from C-G one. When we set the electron-hole interaction strength to zero, i.e.,  $J = 0$  in Eq. (4), our results match the findings of the above-mentioned works [32,39,45]: The electron spreads over all DNA bases, while the hole is either distributed between the left and right edges [consider  $J = 0$  for even-length chains, Fig. 5(a)] or localized

on the left edge [ $J = 0$  for odd-length chains, Fig. 5(b)]. This different behavior can be easily explained by considering the physics of one-dimensional topological systems, such as the Su-Schrieffer-Heeger (SSH) chain [48]. By inspecting the interaction energies of electrons and holes in Table VII in the SM [42], one notices that in the electronic sector one has  $|t_{GC}| > |t_{CG}|$ , while the opposite applies to holes. In SSH chains this in turn induces two different regimes, namely a trivial and topological one, characterized by exactly the two dynamical scenarios seen for electrons and holes here.

Now, when switching on the electron-hole interaction  $J \neq 0$ , we observe the tendency to wash out these differences, and two different regimes can be distinguished: In the intermediate regime ( $J$  is of the same order as the site-site overlap parameters), the electron population at the edges increases slightly while the hole spreads across the molecule and populates the bases in the bulk, very similar to the behavior of the electron in the uncorrelated regime. While the excitonic population (population of electron and hole sitting at the same base) thus increases, the spreading of the hole along the chain is suppressed. With increasing interaction strength, the role of electron and hole switch, with the hole pulling the electron towards the edges. In this strongly interacting regime, both charges display a similar localized distribution, leading to the observed increase in the excitonic population.

Interestingly, we can identify an even-odd effect by comparing results shown in Fig. 5(a) and Fig. 5(b). Sequences

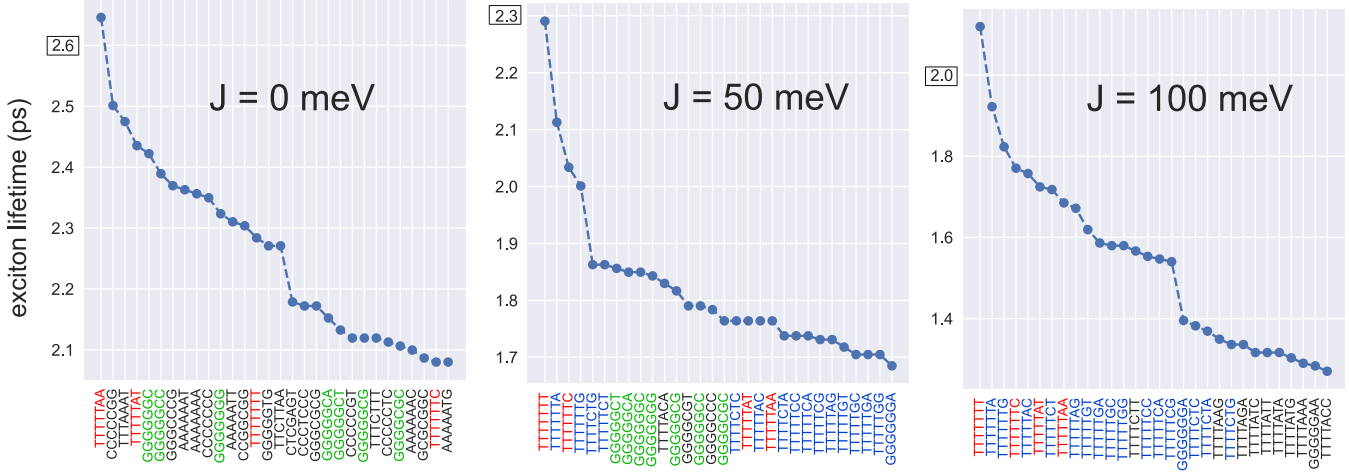


FIG. 6. Expected values of excitonic lifetimes for different values of the electron-hole interaction strength: 0 meV, 50 meV, and 100 meV. The colors of the sequences indicate the copresence of the sequence in different plots. Red: sequence present in all three plots; green: sequence present in the first two plots; blue: sequence present in the second and third plots.

with an even or odd number of sites (base pairs) show a rather different behavior in the hole sector. In even-length sequences, the hole tends to be distributed at the edges of the strand, not populating the bulk, in both cases when it is not interacting with the electron and when it does so strongly. The same is true for the excitonic distribution in the strong interaction regime. In contrast, in odd-length sequences, both charges tend to localize only at the very first site of the chain, where the exciton has been created by an external stimulus. Different studies on the delocalization of quantum particles in SSH-like topological systems confirm this even-odd duality that we observe in this context [48,49]. These findings support the fact that the mobility of charge carriers and of more elaborate structures like Frenkel excitons, is way more likely to be influenced by long-range transport when placed in a wire sequence with an even number of sites rather than an odd number, in agreement with Ref. [30]. Nonetheless, both cases show very similar values for the total excitonic population along the strand.

### C. Electron-hole interaction and exciton lifetime

We now come back to the statistical analysis presented in Sec. III A and extend it by including electron-hole interaction in the ELM. Inspired by the discussion in the previous section, as illustrative scenarios two specific values for the interaction parameter, namely,  $J = 50$  meV and  $J = 100$  meV are chosen.

Figure 6 shows the top 30 sequences with the longest excitonic lifetimes  $S_{\text{time}}^{J \neq 0}$ . One immediately sees the tendency that lifetimes are slightly reduced for all sequences compared to the situation with  $J = 0$  (see also the SM [42]). This is easily understood by the fact that an attractive electron-hole interaction promotes local exciton recombination. However, quite unexpected is the finding that the sets  $S_{\text{time}}^{J \neq 0}$  of DNA sequences for finite  $J$  are strongly correlated while correlations to the set for  $J = 0$  ( $S_{\text{time}}$ ) are considerably weaker. This in turn implies that the mere insertion of a finite charge-charge

correlation is relevant (within a reasonable window) but not the precise value of the interaction parameter. Since the latter is not directly accessible experimentally, estimated values may already provide sufficiently reliable predictions for sequence dependences.

Details of this analysis are shown in Fig. 7. The left panel shows correlations for the entire list  $S_7$  of DNA sequences discussed already in Sec. III A. The correlations between the noninteracting and either of the two interacting lists (blue and green) are weak to absent within the subsets  $S_{\text{time}}$  resp.  $S_{\text{time}}^{J \neq 0}$  of the longest lifetimes, but moderate between the subsets of shorter lifetimes. As mentioned above, most strikingly, the correlations between lists with finite  $J$  are, instead, permanently very strong, independent of the value of  $J$ . Even in the right panel of Fig. 7, where data for the 100 DNA sequences with the longest lifetimes are shown, correlations lie constantly above  $\chi \sim 0.9$ , except for the top sequences where still  $\chi > 0.7$ . Very similar results are obtained by performing the same analysis for the mean charge distribution as reported in Fig. S4 of the SM [42].

We thus conclude that, with respect to the observables under analysis, considering the effects of electron-hole interactions is crucial for a realistic description of the dynamics of charges and excitons along DNA strands. The mere introduction of electron-hole interactions into the model seems to be much more relevant than the actual tuning of interaction strength.

In Fig. 8 we report the same analysis as in Fig. 4 for the relative distribution of base pairs over the exciton lifetime, for the two scenarios with finite electron-hole interaction. We find that the effect of the electron-hole interaction translates into a much greater influence of the TA base pair to ensure a higher lifetime of an exciton within a given sequence. This reversal of tendency also appears in Fig. 6, when considering the sequences that newly appear in  $S_{\text{time}}^{J \neq 0}$  as compared to the set with  $J = 0$ . This suggests that T-rich sequences may more strongly support long-lived dipole structures and may thus be more susceptible to reveal properties that are of potential

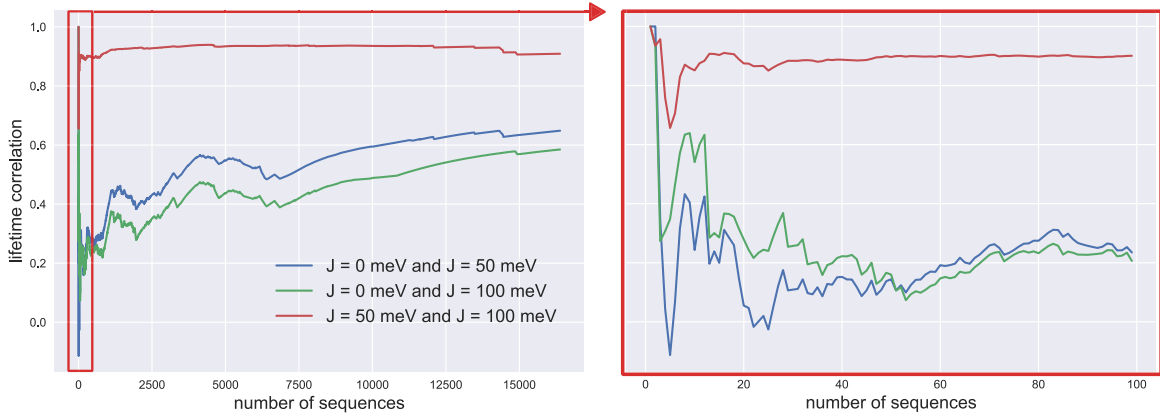


FIG. 7. Correlation plots between lists of lifetimes of DNA sequences, for different values of the electron-hole interaction strength (see legend). The right plot is a zoom into the first 100 sequences of the left plot. The correlation between different non-zero electron-hole interactions is always higher than the correlations with zero electron-hole interactions.

relevance for epigenetic processes due to photoexcitation (e.g., UV radiation in skin cells).

#### IV. CONCLUSIONS

In this work we established a methodological bridge between quantum dynamical simulations of electron-hole charge transfer along double-stranded DNA and the statistical analysis of a large ensemble of DNA sequences. Although single charge transfer has been studied quite frequently, correlated electron-hole dynamics including recombination has rarely been investigated [26,27,50,51]. By using a tight-binding model with parameters taken from the literature, we are able to obtain results for all possible permutations of DNA segments of up to seven bases in each strand, far beyond what is feasible

in atomistic calculations. The efficiency of this method allows us to correlate specific physical parameters to the entire ensemble of corresponding DNA sequences, thus to establishing relative dependences between them. This way, predictions can be made for subsets that support relatively long-lived, delocalized charge-hole pairs, which, in turn, imply relatively strong electrical dipole moments over distances of 3–4 base pairs. Notably, these sequences display a particular enrichment in thymine (T) bases. The resulting charge imbalance within the DNA molecule induces molecular polarization, which has the potential to influence various biological processes, such as sequence- and conformation-dependent protein binding and transcriptional regulation.

Our findings are quantitatively in agreement with experimental data in terms of exciton lifetimes (about 2–3 ps)

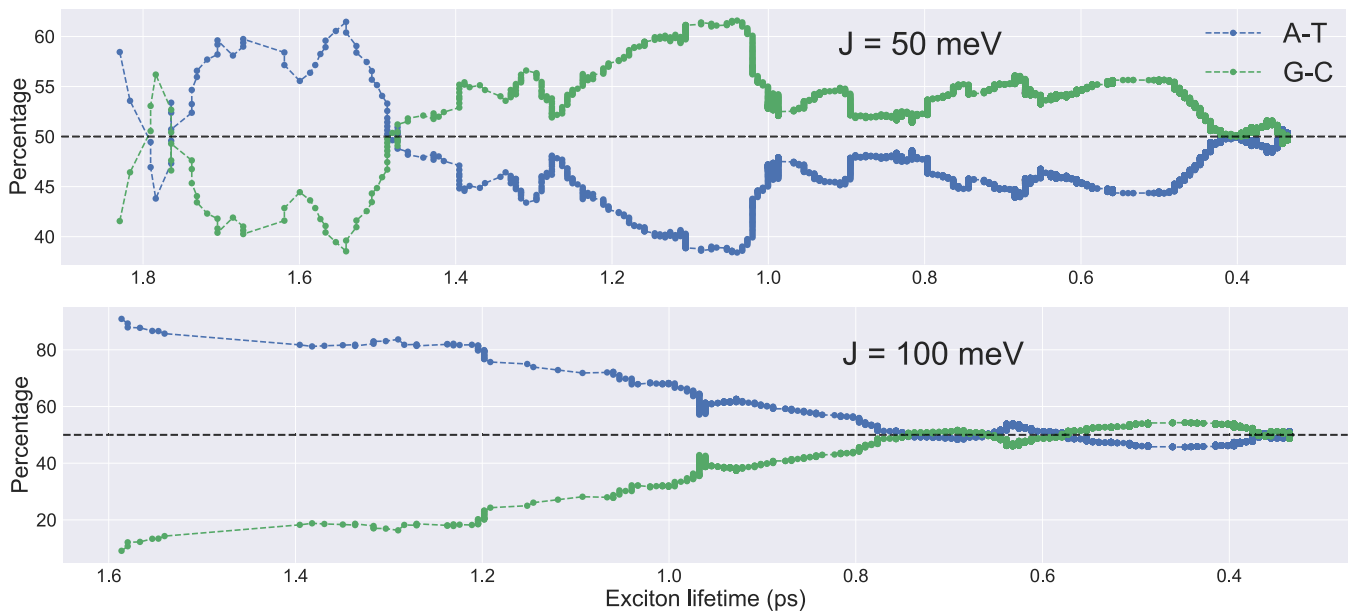


FIG. 8. Relative distribution of base pairs (A-T pair and G-C pair) in sequences of seven bases per strand during the expected exciton lifetime, for the electron-hole interaction = 50 (top) or 100 (bottom) meV. We obtain the distribution at time  $t$  by counting the base pairs in all sequences with a lifetime greater than  $t$  and dividing by the total number of base pairs. Since it is necessary to average over several sequences before the results become relevant, we have omitted the first ten sequences with the largest exciton lifetimes.



[10,23,43] and coherence lengths (about 10 Å) [36,44]. However, the tight-binding model presented is still too simplistic to allow quantitative predictions for individual DNA sequences, e.g., including a realistic description of the DNA backbone, but it does allow the model to characterize relative correlations within an ensemble of sequences. In particular, it turns out that most subsets are relatively robust against variations of the charge-hole interaction parameter, as long as it remains finite, within a broad window of values.

Finally, we want to discuss some implications that can be cast into the following questions: Might some DNA-binding proteins use DNA-mediated (correlated) charge transfer (CT) for long-range signaling or activation? CT chemistry provides an approach to detect base stacking perturbations and lesions: Might nature take advantage of this chemistry? Telomeric DNA (the region of DNA at the ends of all linear chromosomes) and regions flanking protein-coding exons are guanine rich: Could the higher degree of degeneracy (due to repeated Gs) in these areas facilitate the tunneling of excess charges to these noncoding regions, thus avoiding dangerous

complications (as speculated in Ref. [52])? We believe that the methodology presented in this work and its extension to larger system sizes and respectively larger ensembles of DNA sequences may open new avenues to address these questions from a perspective, which combines quantum dynamics with the large-scale screening of DNA base compositions. Thus, this research may complement existing techniques from physical chemistry, biochemistry, and genetics to achieve a broader understanding of properties of DNA with potential applications in fields such as nanotechnology [22], epigenetics [35], and DNA-related studies including DNA damage, repair, and mutagenesis.

#### ACKNOWLEDGMENTS

The authors gratefully acknowledge Marco Trenti for his assistance with numerical simulations. The authors appreciate the financial support of the Center for Integrated Quantum Science and Technology (IQST) and the BMBF through the QCOMP project in Cluster4Future QSens.

- 
- [1] F. Fassioli, R. Dinshaw, P. Arpin, and G. Scholes, Photosynthetic light harvesting: excitons and coherence, *J. R. Soc. Interface.* **11**, 20130901 (2014).
- [2] T. Mirkovic, E. E. Ostroumov, J. M. Anna, R. van Grondelle, Govindjee, and G. D. Scholes, Light absorption and energy transfer in the antenna complexes of photosynthetic organisms, *Chem. Rev.* **117**, 249 (2017).
- [3] A. Chenu and G. D. Scholes, Coherence in energy transfer and photosynthesis, *Annu. Rev. Phys. Chem.* **66**, 69 (2015).
- [4] J. Cao, R. J. Cogdell, D. F. Coker, H.-G. Duan, J. Hauer, U. Kleinekathöfer, T. L. C. Jansen, T. Mančal, R. J. D. Miller, J. P. Ogilvie, V. I. Prokhorenko, T. Renger, H.-S. Tan, R. Tempelaar, M. Thorwart, E. Thyryhaug, S. Westenhoff, and D. Zigmantas, Quantum biology revisited, *Sci. Adv.* **6**, eaaz4888 (2020).
- [5] A. Maffioni, F. Caycedo-Soler, Susana F. Huelga, and M. B. Plenio, Design principles for long-range energy transfer at room temperature *Phys. Rev. X* **11** 041003 (2021).
- [6] O. Ostroverkhova, Organic optoelectronic materials: Mechanisms and applications, *Chem. Rev.* **116**, 13279 (2016).
- [7] O. Inganäs, Organic photovoltaics over three decades, *Adv. Mater.* **30**, 1800388 (2018).
- [8] X. F. Lu, Y. Fang, D. Luan, and X. W. D. Lou, Metal-organic frameworks derived functional materials for electrochemical energy storage and conversion: A mini review, *Nano Lett.* **21**, 1555 (2021).
- [9] R. Improta, F. Santoro, and L. Blancafort, Quantum mechanical studies on the photophysics and the photochemistry of nucleic acids and nucleobases, *Chem. Rev.* **116**, 3540 (2016).
- [10] C. Crespo-Hernandez, B. Cohen, P. Hare, and B. Kohler, Ultrafast excited-state dynamics in nucleic acids, *Chem. Rev.* **104**, 1977 (2004).
- [11] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, DNA excited-state dynamics: From single bases to the double helix, *Annu. Rev. Phys. Chem.* **60**, 217 (2009).
- [12] K. Kleinermanns, D. Nachtigallová, and M. S. de Vries, Excited state dynamics of DNA bases, *Int. Rev. Phys. Chem.* **32**, 308 (2013).
- [13] W. J. Schreier, P. Gilch, and W. Zinth, Early events of DNA photodamage, *Annu. Rev. Phys. Chem.* **66**, 497 (2015).
- [14] G. Schuster, *Long-Range Charge Transfer in DNA* (Springer, New York, 2004).
- [15] R. G. Endres, D. L. Cox, and R. R. P. Singh, *Colloquium: The quest for high-conductance DNA*, *Rev. Mod. Phys.* **76**, 195 (2004).
- [16] L. D'Agostino, Native DNA electronics: is it a matter of nanoscale assembly? *Nanoscale* **10**, 12268 (2018).
- [17] S. R. Rajski, B. A. Jackson, and J. K. Barton, DNA repair: models for damage and mismatch recognition, *Mutat. Res.* **447**, 49 (2000).
- [18] S. R. Rajski and J. K. Barton, How different DNA-binding proteins affect long-range oxidative damage to DNA, *Biochem.* **40**, 5556 (2001).
- [19] B. Giese, Electron transfer through DNA and peptides, *Bioorg. Med. Chem.* **14**, 6139 (2006).
- [20] C.-T. Shih, Y.-Y. Cheng, S. A. Wells, C.-L. Hsu, and R. A. Römer, Charge transport in cancer-related genes and early carcinogenesis, *Comput. Phys. Commun.* **182**, 36 (2011).
- [21] T. Chakraborty, *Charge Migration in DNA: Perspectives from Physics, Chemistry, and Biology* (Springer, Berlin, 2007).
- [22] E. L. Albuquerque, U. L. Fulco, V. N. Freire, E. W. S. Caetano, M. L. Lyra, and F. A. B. F. de Moura, DNA-based nanobiostructured devices: The role of quasiperiodicity and correlation effects, *Phys. Rep.* **535**, 139 (2014).
- [23] C. Crespo-Hernandez, B. Cohen, and B. Kohler, Base stacking controls excited-state dynamics in A · T DNA, *Nature (London)* **436**, 1141 (2005).
- [24] T. Takaya, C. Su, K. de La Harpe, C. E. Crespo-Hernández, and B. Kohler, UV excitation of single DNA and RNA strands produces high yields of exciplex states between

- two stacked bases, *Proc. Natl. Acad. Sci. USA* **105**, 10285 (2008).
- [25] X. Wang, R. Sha, W. B. Knowlton, N. C. Seeman, J. W. Canary, and B. Yurke, Exciton delocalization in a DNA-templated organic semiconductor dimer assembly, *ACS Nano* **16**, 1301 (2022).
- [26] E. R. Bittner, Lattice theory of ultrafast excitonic and charge-transfer dynamics in DNA, *J. Chem. Phys.* **125**, 094909 (2006).
- [27] E. R. Bittner, Frenkel exciton model of ultrafast excited state dynamics in AT DNA double helices, *J. Photochem. Photobiol. A* **190**, 328 (2007).
- [28] R. Gutiérrez, R. A. Caetano, B. P. Woiczikowski, T. Kubar, M. Elstner, and G. Cuniberti, Charge transport through biomolecular wires in a solvent: Bridging molecular dynamics and model Hamiltonian approaches, *Phys. Rev. Lett.* **102**, 208102 (2009).
- [29] R. Gutiérrez, R. Caetano, P. B. Woiczikowski, T. Kubar, M. Elstner, and G. Cuniberti, Structural fluctuations and quantum transport through DNA molecular wires: A combined molecular dynamics and model Hamiltonian approach, *New J. Phys.* **12**, 023022 (2010).
- [30] K. Lambropoulos and C. Simserides, Tight-binding modeling of nucleic acid sequences: Interplay between various types of order or disorder and charge transport, *Symmetry* **11**, 968 (2019).
- [31] G. Cuniberti, E. Maciá, A. Rodríguez, and R. A. Römer, Tight-binding modeling of charge migration in DNA devices, in *Charge Migration in DNA: Perspectives from Physics, Chemistry, and Biology*, edited by T. Chakraborty (Springer, Berlin, 2007), pp. 1–20.
- [32] C. Simserides, A systematic study of electron or hole transfer along DNA dimers, trimers and polymers, *Chem. Phys.* **440**, 31 (2014).
- [33] L. Hawke, G. Kalosakas, and C. Simserides, Electronic parameters for charge transfer along DNA, *Eur. Phys. J. E* **32**, 291 (2010).
- [34] B. Giese, Long-distance electron transfer through DNA, *Annu. Rev. Biochem.* **71**, 51 (2002).
- [35] R. Siebert, O. Ammerpohl, M. Rossini, D. Herb, S. Rau, M. Plenio, F. Jelezko, and J. Ankerhold, A quantum physics layer of epigenetics: a hypothesis deduced from charge transfer and chirality-induced spin selectivity of DNA, *Clin. Epigenet.* **15**, 145 (2023).
- [36] J. C. Genereux and J. K. Barton, Mechanisms for DNA charge transport, *Chem. Rev.* **110**, 1642 (2010).
- [37] I. Buchvarov, Q. Wang, M. Raytchev, A. Trifonov, and T. Fiebig, Electronic energy delocalization and dissipation in single- and double-stranded DNA, *Proc. Natl. Acad. Sci. USA* **104**, 4794 (2007).
- [38] U. Kadhane, A. I. S. Holm, S. V. Hoffmann, and S. B. Nielsen, Strong coupling between adenine nucleobases in DNA single strands revealed by circular dichroism using synchrotron radiation, *Phys. Rev. E* **77**, 021901 (2008).
- [39] K. Lambropoulos, M. Chatzieftheriou, A. Morphis, K. Kaklamanis, R. Lopp, M. Theodorakou, M. Tassi, and C. Simserides, 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* **94**, 062403 (2016).
- [40] R. Korol and D. Segal, Machine learning prediction of DNA charge transport, *J. Phys. Chem. B* **123**, 2801 (2019).
- [41] A. Ullah and P. O. Dral, Predicting the future of excitation energy transfer in light-harvesting complex with artificial intelligence-based quantum dynamics, *Nature Commun.* **13**, 1930 (2022).
- [42] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.109.064413> for ultrafast excitonic dynamics in DNA: Bridging quantum dynamics and sequence dependence.
- [43] A. Trifonov, M. Raytchev, I. Buchvarov, M. Rist, J. Barbaric, H.-A. Wagenknecht, and T. Fiebig, Ultrafast energy transfer and structural dynamics in DNA, *J. Phys. Chem. B* **109**, 19490 (2005).
- [44] E. M. Boon and J. K. Barton, Charge transport in DNA, *Curr. Opin. Struct. Biol.* **12**, 320 (2002).
- [45] K. Lambropoulos, K. Kaklamanis, A. Morphis, M. Tassi, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieftheriou, and C. Simserides, Wire and extended ladder model predict THz oscillations in DNA monomers, dimers and trimers, *J. Phys.: Condens. Matter* **28**, 495101 (2016).
- [46] M. Mantela, C. Simserides, and R. D. Felice, LCAO electronic structure of nucleic acid bases and other heterocycles and transfer integrals in B-DNA, including structural variability, *Materials* **14**, 4930 (2021).
- [47] M. Mantela, K. Lambropoulos, and C. Simserides, Charge transport properties of ideal and natural DNA segments, as mutation detectors, *Phys. Chem. Chem. Phys.* **25**, 7750 (2023).
- [48] J. K. Asbóth, L. Oroszlány, and A. Pályi, The Su-Schrieffer-Heeger (SSH) model, in *A Short Course on Topological Insulators: Band Structure and Edge States in One and Two Dimensions* (Springer, Berlin, 2016), pp. 1–22.
- [49] T. Cramer, S. Krapf, and T. Koslowski, DNA charge transfer: An atomistic model, *J. Phys. Chem. B* **108**, 11812 (2004).
- [50] S. Tornow, R. Bulla, F. B. Anders, and G. Zwignagl, Multiple-charge transfer and trapping in DNA dimers, *Phys. Rev. B* **82**, 195106 (2010).
- [51] E. M. Conwell, P. M. McLaughlin, and S. M. Bloch, Charge-transfer excitons in DNA, *J. Phys. Chem. B* **112**, 2268 (2008).
- [52] S. Delaney and J. K. Barton, Long-range DNA charge transport, *J. Org. Chem.* **68**, 6475 (2003).