National and Kapodistrian University of Athens
Department of Physics
Section of Condensed Matter Physics

Energy structure and charge transport-transfer in molecular wires:
carbynes and periodic, deterministic aperiodic, and random DNA

Konstantinos Lambropoulos
Ph.D. Thesis
Athens, 2020

"I've been writing for two years now
and I've composed just one idyll."
–C. P. Cavafy, *The First Step*

Contents

Preface

The research described in the present PhD thesis was conducted in the National and Kapodistrian University of Athens, between April 2016 and February 2020, under the supervision of Assistant Professor Constantinos Simserides.

It is most likely that these lines would have never been written, hadn't I met my soon-to-become BSc, Msc and, eventually, PhD supervisor, back in 2013, when I was almost finishing my undergraduate studies. His combination of scientific insight with spontaneous decency exemplifies the qualities of a great teacher and person. This is the place to deeply thank Constantinos Simserides for his guidance, support and -above all- trust throughout all these years.

I would also like to thank the other two members of my Advisory Committee, Emeritus Professor Georgios Triberis and Assistant Professor Vlasios Likodimos for their support, comments, and having an open door anytime they were needed.

Apart from Constantinos Simserides and Vlasios Likodimos, my Examination Committee was also comprised of Professor Nikolaos Stefanou, Professor Phivos Mavropoulos, Professor Enrique Maciá, Associate Professor Rosa di Felice, and Associate Professor Spiros Gardelis. I am both honored and grateful that they accepted to offer their time and effort to review and comment on this document.

Special thanks go to all current and present members and collaborators of our research group, "Physics of Nanosctructures and Biomaterials", Marilena Mantela, Andreas Morphis, Maria Tassi, Konstantinos Kaklamanis, Marina Theodorakou, Christina Zacharaki, Georgios Georgiadis, Maria Chatzieleftheriou, Richard Lopp, Christina Vantaraki, and Panagiota Bilia, whom I was pleased to work with.

Constant support, encouragement, and sacrifices by my parents, who instilled in me a high value on education, have laid the groundwork of who I am today. I deeply thank them for everything.

I also thank Elli, for being there.

I hope my computer was more useful to the person who stole it than it was to me, when the first draft of the present thesis was almost ready.

Finally, I wish to acknowledge financial support for my PhD research 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).

H.F.R.I.
Hellenic Foundation for Research & Innovation

GSRT
GENERAL SECRETARIAT FOR RESEARCH AND TECHNOLOGY

Abstract

We theoretically study the energy structure and the charge transport and transfer properties of π-conjugated molecular wires, using variants of the Tight-Binding Method. Charge transport implies the use of leads and the application of external voltage. We start with an analytical and numerical study of the spectral and transmission properties of periodic Tight-Binding wires with a generic unit cell, focusing on the effect factors such as the strength and asymmetry of coupling between the leads an the system induce on the transmission profiles. Our method is then applied to the study of atomic carbon wires (or carbynes). Carbynes represent the ultimate nanoscale structure, having a thickness of just one carbon atom, and are promising for electronic applications. We show that Ohmic, semiconducting or rectifying behavior occurs, depending on the carbyne structure and the leads, and we reproduce experimental results regarding the current-voltage curves. We move on to an examination of the energy structure, localization and charge transport in periodic, deterministic aperiodic and random binary DNA sequences. The ability to produce nucleic acid sequences of interest provides the chance to create molecular wires with tailored properties. In our study, we focus on the interplay between the sequence structure and the aforementioned properties. Charge transfer means that an extra carrier (hole or electron), created or injected at a specific location, moves to more favorable locations. We focus on aspects of this phenomenon such as the frequency content of transfer, the mean over time probabilities to find the carrier at each site of a DNA segment and the pure mean transfer rates. We start with small DNA segments (composed of one, two, and three base pairs). Our results are compared with the more complex, yet more computationally costly, Real-Time Time-Dependent Density Functional Theory. We move on to the study of several classes of periodic DNA segments (monomer-polymers, dimer-polymers, polymers with increasing repetition unit). Finally, we compare periodic DNA segments with deterministic aperiodic and random ones regarding charge transfer, and make some remarks regarding experimental charge transfer rates.

Περίληψη

Το DNA παίζει καθοριστικό ρόλο στην ανάπτυξη, λειτουργία και αναπαραγωγή των εμβίων οργανισμών, καθώς η αλληλουχία των αζωτούχων βάσεών του φέρει το γενετικό τους κώδικα. Η μελέτη του συνήθως συνδέεται με επιστημονικούς κλάδους όπως η γενετική και η μοριακή βιολογία. Ωστόσο, οι αξιοσημείωτες ιδιότητές του έχουν κεντρίσει τα τελευταία χρόνια το ενδιαφέρον μιας ευρείας διεπιστημονικής κοινότητας. Από τη σκοπιά της Φυσικής Συμπυκνωμένης Ύλης, η ηλεκτρονική δομή και οι ιδιότητες μεταφοράς και μεταβίβασης φορτίου του DNA μελετώνται με σκοπό τόσο την κατανόηση της σχέσης μεταξύ της ενεργειακής του δομής και των βιολογικών του λειτουργιών, όσο και τις δυνητικές τους εφαρμογές. Η αλληλουχία των ζευγών βάσεών του δημιουργεί έναν σχεδόν μονοδιάστατο π-δρόμο, λόγω της επικάλυψης των π μοριακών τροχιακών τους, ο οποίος επιτρέπει τη μετακίνηση φορτίου, δηλαδή τη μεταφορά και τη μεταβίβαση. Ο όρος *μεταφορά (transport)* αναφέρεται στη μετακίνηση φορτίου μέσω της χρήσης ηλεκτροδίων και της εφαρμογής εξωτερικής διαφοράς δυναμικού, ενώ ο όρος *μεταβίβαση (transfer)* αναφέρεται στη μετακίνηση ενός φορέα από μια τοποθεσία όπου δημιουργείται ή εγχέεται αρχικά σε άλλες ευνοϊκότερες τοποθεσίες.

Πέραν του ενδιαφέροντος που παρουσιάζει από βιολογικής και νανοεπιστημονικής άποψης, το DNA μπορεί να ιδωθεί κι ως ένα πρότυπο μονοδιάστατο μοριακό σύρμα. Η τεράστια μείωση του μήκους του καναλιού (διαύλου) στις ηλεκτρονικές διατάξεις κατά τις τελευταίες δεκαετίες (από περίπου 1 μm το 1984 σε περίπου 10 nm το 2014) ήγειρε το ερώτημα εάν και πως μπορούν να χρησιμοποιηθούν σε διατάξεις ως κανάλια ορισμένα πολυμερή, ολιγομερή, (μακρο)μόρια κ.ο.κ.. Μια από τις πλέον απαραίτητες προϋποθέσεις που θα πρέπει να πληρούν τέτοιες δομές προκειμένου να χρησιμοποιηθούν είναι η ακαμψία (persistence). Το DNA βρίσκεται ανάμεσα στα πλέον άκαμπτα (persistent) γνωστά πολυμερή, με μήκος ακαμψίας (persistence length) περίπου 50 nm – με άλλα λόγια, περίπου 150 ζεύγη βάσεων. Το μήκος ακαμψίας (persistence length) ποσοτικοποιεί την ακαμψία μια δομής, υπό την έννοια ότι μια δομή με μήκος μικρότερο από το μήκος ακαμψίας συμπεριφέρεται περίπου ως ελαστική δοκός, ενώ, από την άλλη, δομές με πολύ μεγαλύτερο μήκος είναι πιθανότερο να καμφθούν. Πέραν όμως του μήκους ακαμψίας του, το DNA παρέχει τη δυνατότητα κατασκευής επιλεγμένων αλληλουχιών, βασισμένων σε ένα αλφάβητο, π. χ. το {G, A, C, T} όπου G = Γουανίνη, A = Αδενίνη, C = Κυτοσίνη, T = Θυμίνη]. Συνεπώς υπάρχει η δυνατότητα ρύθμισης των ιδιοτήτων μιας διάταξης μέσω της επιλογής της κατάλληλης αλληλουχίας βάσεων για μια δεδομένη λειτουργία. Με αυτόν τον τρόπο, μπορούν να κατασκευαστούν ποικίλες περιοδικές, ντετερμινιστικές απεριοδικές (π.χ., οιονεί κρυσταλλικές και μορφοκλασματικές) ή τυχαίες ακολουθίες.

Εδώ και τουλάχιστον δύο δεκαετίες γίνονται απόπειρες κατανόησης της μετακίνησης φορτίου μέσω του DNA. Σήμερα, γνωρίζουμε ότι υπάρχουν πολλοί εξωγενείς παράγοντες (υδαρότητα, παρουσία αντισταθμιστικών ιόντων –counter ions, διαδικασία εξαγωγής από το κύτταρο, ηλεκτρόδια, επαφές, καθαρότητα, υπόστρωμα) αλλά και ενδογενείς παράγοντες (αλληλουχία ζευγών βάσεων, γεωμετρία), οι οποίοι επηρεάζουν τη μεταφορά και τη μεταβίβαση φορτίου, οδηγώντας σε ποικίλα και πολλές φορές αντικρουόμενα συμπεράσματα. Η παρούσα Διδακτορική Διατριβή εστιάζει στον σημαντικότερο ίσως ενδογενή παράγοντα, ο οποίος επιπλέον επάγει διάφορους τύπους τάξης ή αταξίας, δηλαδή στην επίδραση της αλληλουχίας βάσεων, η οποία επηρεάζει την ηλεκτρονική επικάλυψη κατά μήκος του π-δρόμου.

Μια άλλη, απλούστερη του DNA, π-συζευγμένη δομή, η οποία θα μπορούσε να είναι υποψήφια ως μοριακό σύρμα είναι τα ατομικά σύρματα άνθρακα ή καρβύνες. Οι καρβύνες αποτελούν τρόπον τινά την απόλυτη νανοδομή, έχοντας πάχος μόλις ενός ατόμου, αποτελούμενες από sp υβριδισμένα άτομα άνθρακα. Οι δυο δυνατές διαμορφώσεις των καρβυνών είναι η κουμουλενική και η πολυυνική. Στην πρώτη, τα διαδοχικά άτομα άνθρακα ισαπέχουν, ενώ στη δεύτερη όχι. Οι καρβύνες παρουσιάζουν ενδιαφέρουσες ιδιότητες όπως το ρυθμίσιμο ενεργειακό χάσμα, η ανθεκτικότητα σε θερμοκρασία δωματίου και η μεγάλη ακαμψία. Έχουν μήκος ακαμψίας περίπου 14 nm – με άλλα λόγια, περίπου 100 άτομα άνθρακα, το οποίο τις καθιστά πολλά υποσχόμενες για ηλεκτρονικές εφαρμογές.

Η παρούσα Διδακτορική Διατριβή αποτελεί μια συστηματική θεωρητική (αναλυτική και αριθμητική) μελέτη της ενεργειακής δομής και των ιδιοτήτων μεταφοράς και μεταβίβασης φορτίου στις προαναφερθείσες δομές (DNA και καρβύνες), στο πλαίσιο του προτύπου Ισχυρής Δέσμευσης (Tight Binding). Το πρότυπο αυτό αποτελεί μια γρήγορη και αποτελεσματική μέθοδο, εφαρμόσιμη σε μεγάλο αριθμό προβλημάτων, τα οποία αφορούν την ηλεκτρονική δομή και τις ιδιότητες της ύλης και απαιτούν ποικίλους βαθμούς ακρίβειας. Στα βασικά της πλεονεκτήματα περιλαμβάνονται η διαισθητική απλότητά της, η δυνατότητα εξαγωγής αναλυτικών αποτελεσμάτων σε πολλές περιπτώσεις, καθώς και το χαμηλό της υπολογιστικό κόστος. Το τελευταίο καθιστά το πρότυπο Ισχυρής Δέσμευσης εφαρμόσιμο σε μεγάλα συστήματα, τα οποία επί του παρόντος δεν είναι διαχειρίσιμα με τις –λεπτομερέστερες– μεθόδους από πρώτες αρχές. Σε αντίθεση με τις μεθόδους αυτές, το πρότυπο Ισχυρής Δέσμευσης είναι ημιεμπειρικό, υπό την έννοια ότι χρειάζεται ένα εξωτερικό σύνολο παραμέτρων προκειμένου να εκτελεστούν υπολογισμοί. Οι παράμετροι αυτές είναι (α΄) οι επιτόπιες ενέργειες (on-site energies), οι οποίες αντιστοιχούν στην ενέργεια των ηλεκτρονίων σε κάθε θέση (site) του πλέγματος, και (β΄) τα ολοκληρώματα μεταπήδησης (hopping integrals) ή εναλλακτικά ολοκληρώματα μεταβίβασης (transfer integrals), τα οποία αντιστοιχούν στη σύζευξη μεταξύ των τροχιακών γειτονικών θέσεων του πλέγματος. Τα τροχιακά που λαμβάνονται υπόψη στην παρούσα Διδακτορική Διατριβή είναι το Υψηλότερο Κατειλημμένο Μοριακό Τροχιακό (HOMO) και το Χαμηλότερο Μη-κατειλημμένο Μοριακό Τροχιακό (LUMO).

Η μετακίνηση φορτίου μπορεί να διακριθεί σε: (i) μη συνεκτική (incoherent) ή θερμική (thermal) μεταπήδηση μεταξύ πλησιεστέρων ή πιο μακρινών θέσεων, και (ii) συνεκτική (coherent) μεταπήδηση ή φαινόμενο σήραγγας (tunneling) ή υπεραντάλλαγή (superexchange). Ο συνεκτικός μηχανισμός αναμένεται να κυριαρχεί σε χαμηλές θερ-

μοκρασίες. Στο θερμικό μηχανισμό ο φορέας εντοπίζεται και ανταλλάσσει ενέργεια με το περιβάλλον κατά τη μεταβίβασή του. Η παρούσα Διδακτορική Διατριβή αφιερώνεται στη συνεκτική μεταβίβαση και μεταφορά φορτίου. Στη συνεκτική περίπτωση, υπάρχει πεπερασμένη πιθανότητα παρουσίας του φορέα σε κάθε θέση, παρ' όλο που αυτές με τις κατάλληλες επιτόπιες ενέργειες –για δεδομένη αρχική τοποθέτηση του φορέα– ευνοούνται.

Η Διδακτορική Διατριβή διαρθρώνεται σε τρία Μέρη και αποτελείται από εννέα Κεφάλαια.

Το Μέρος I (Εισαγωγή) αποτελείται από τα Κεφάλαια 1–2.

Στο Κεφάλαιο 1 αναφέρονται τα γενικά χαρακτηριστικά των υπό μελέτη δομών (DNA και καρβυνών), παρουσιάζεται το επιστημονικό πλαίσιο, η βασική ορολογία και τα ερευνητικά ερωτήματα που μελετώνται στο πλαίσιο της Διδακτορικής Διατριβής και παρουσιάζεται η διάρθρωσή της.

Στο Κεφάλαιο 2 παρουσιάζεται το θεωρητικό πλαίσιο του προτύπου Ισχυρής Δέσμευσης και τα μαθηματικά εργαλεία που χρησιμοποιούνται στη Διδακτορική Διατριβή. Εξετάζονται σε γενικές γραμμές μονοδιάστατα πρότυπα Ισχυρής Δέσμευσης, τα οποία χρησιμοποιούνται για τη μελέτη της ενεργειακής δομής και των ιδιοτήτων μεταφοράς και μεταβίβασης φορτίου σε οργανικά π-συζευγμένα συστήματα, τα οποία είναι υποψήφια ως μοριακά σύρματα. Τα μελετώμενα συστήματα αποτελούνται από N μονομερή και L αλυσίδες ($N \gg L$, καθώς εκτείνονται σε μια διάσταση) και το πρόβλημα ανάγεται στην επίλυση του λεγόμενου στάσιμου συστήματος εξισώσεων Ισχυρής Δέσμευσης, το οποίο ισοδυναμεί με μια διακριτοποιημένη μορφή της χρονοανεξάρτητης εξίσωσης Schrödinger. Παρουσιάζεται η μέθοδος των Πινάκων Μεταβίβασης (Transfer Matrix Method) και η χρήση της για τη γενική επίλυση του συστήματος των εξισώσεων Ισχυρής Δέσμευσης για κυκλική συνοριακή συνθήκη και για συνοριακή συνθήκη πακτωμένων άκρων. Στη συνέχεια, παρουσιάζονται διάφορες παραλλαγές του προτύπου Ισχυρής Δέσμευσης οι οποίες χρησιμοποιούνται για τη μελέτη του DNA: πρότυπο σύρματος (wire), κλίμακας (ladder), επεκταμένης κλίμακας (extended ladder), ψαροκόκαλου (fishbone), κλίμακας-ψαροκόκαλου (ladder-fishbone). Τέλος, παρατίθενται ορισμένα στοιχεία για τις ντετερμινιστικές απεριοδικές, αλληλουχίες, τις ιδιότητές τους και την κατηγοριοποίησή τους.

Το Μέρος II (Μεταφορά φορτίου) αποτελείται από τα Κεφάλαια 3–5.

Στο Κεφάλαιο 3 εφαρμόζεται η μέθοδος των πινάκων μεταβίβασης για τη μελέτη της ενεργειακής δομής αλλά και της διέλευσης φορτίου σε περιοδικά πρότυπα σύρματος με μια γενική μονάδα επανάληψης. Δηλαδή $N = um$, όπου u είναι ο αριθμός των μονομερών τα οποία δημιουργούν τη μονάδα επανάληψης, η οποία επαναλαμβάνεται m φορές. Μέσω των ιδιοτήτων των πινάκων μεταβίβασης, προκύπτουν σχέσεις διασποράς και αναδρομικοί τύποι για των προσδιορισμό της ενεργειακής δομής, για κυκλική συνοριακή συνθήκη και συνοριακή συνθήκη πακτωμένων άκρων, αντίστοιχα. Μελετώνται επίσης οι πυκνότητες καταστάσεων. Στη συνέχεια, μελετώνται συστηματικά οι συντελεστές διέλευσης υπό μηδενική τάση, μέσω της σύνδεσης των

συστημάτων με ημιάπειρους ακροδέκτες (ηλεκτρόδια, leads). Ερευνάται η επίδραση παραγόντων όπως η ισχύς και η ασυμμετρία της σύζευξης συστήματος-ακροδεκτών, το εύρος και το κέντρο της ζώνης των ακροδεκτών και οι μεταπηδήσεις εντός του συστήματος. Εισάγεται η συνθήκη βέλτιστης σύζευξης, για πρώτη φορά στη βιβλιογραφία σε αυτή τη γενική μορφή, μεταξύ του συστήματος των ακροδεκτών και συζητείται το φυσικό της περιεχόμενο.

Στο Κεφάλαιο 4, ο φορμαλισμός και τα συμπεράσματα του Κεφαλαίου 3 εφαρμόζονται προκειμένου να μελετηθούν η ενεργειακή δομή και η μεταφορά φορτίου σε ατομικά σύρματα άνθρακα (κουμουλενικές και πολυυνικές καρβύνες). Δείχνεται ότι το πρότυπο σύρματος προβλέπει ποικίλες βασικές πτυχές του φαινομένου και αναπαράγει τα συμπεράσματα πρόσφατων πειραματικών αποτελεσμάτων. Τα αποτελέσματα αυτά περιλαμβάνουν τη μεταλλική (ημιαγωγική) συμπεριφορά των χαρακτηριστικών καμπυλών ρεύματος–τάσης σε κουμουλενικές (πολυυνικές) καρβύνες, την επίδραση της εξασθένησης της σύζευξης μεταξύ του ενός άκρου σου συστήματος με τους ακροδέκτες, καθώς και την ανορθωτική συμπεριφορά των χαρακτηριστικών ρεύματος–τάσης όταν δεν ευθυγραμμίζονται τα επίπεδα Fermi των ακροδεκτών και του συστήματος.

Στο Κεφάλαιο 5 μελετώνται συστηματικά και συγκριτικά, στο πλαίσιο του προτύπου σύρματος, διάφορες κατηγορίες περιοδικών, ντετερμινιστικών απεριοδικών (οιονεί κρυσταλλικών, μορφοκλασματικών και άλλων) καθώς και τυχαίων δυαδικών αλληλουχιών DNA ως προς την ενεργειακή δομή, τον εντοπισμό και τη μεταφορά φορτίου. Δυαδικές σημαίνει πως για να δημιουργηθούν οι αλληλουχίες χρησιμοποιούνται εναλλαγές των λέξεων {G,A}. Εξετάζεται η σχέση μεταξύ της πολυπλοκότητας των αλληλουχιών (μέσω της συχνότητας εμφάνισης κάθε παραμέτρου, των τριάδων ζευγών βάσεων που εμπλέκονται και των συσχετίσεων) και των προαναφερθεισών ιδιοτήτων. Όσον αφορά την ενεργειακή δομή, προκύπτει ότι η συχνότητα εμφάνισης των τριάδων ζευγών βάσεων αντανακλάται στην ολοκληρωμένη πυκνότητα καταστάσεων. Η τριάδα ζευγών βάσεων είναι η φυσική μονάδα του προτύπου Ισχυρής Δέσμευσης, λαμβάνοντας υπ' όψιν τους πρώτους γείτονες. Επίσης, η τριάδα ζευγών βάσεων είναι αυτό που στη μοριακή βιολογία αποκαλούν κωδικόνιο (codon) κάθενα από τα οποία μεταφράζεται σε ένα αμινοξύ. Παρότι οι περιοδικές ακολουθίες αναμενόμενα οδηγούν σε πολύ μεγαλύτερα ρεύματα, προκύπτει ότι ανάλογα με τη δομή και τις παραμέτρους, υπάρχουν κατηγορίες ντετερμινιστικών απεριοδικών ακολουθιών οι οποίες είναι αποτελεσματικές όσον αφορά τη μεταφορά, ενώ οι τυχαίες ακολουθίες αποτελούν τη λιγότερο αποτελεσματική περίπτωση.

Το Μέρος III (Μεταβίβαση φορτίου) αποτελείται από τα Κεφάλαια 6–9.

Στο Κεφάλαιο 6 παρουσιάζεται η βασική θεωρία σχετικά με το πρόβλημα της μεταβίβασης ενός επιπλέον φορτίου (οπής ή ηλεκτρονίου, μέσω των καταστάσεων HOMO και LUMO, αντίστοιχα) σύμφωνα με το πρότυπο Ισχυρής Δέσμευσης. Επίσης, εισάγονται τα φυσικά μεγέθη που θα μελετηθούν στα επόμενα κεφάλαια.

Το Κεφάλαιο 7 εστιάζει στη μελέτη της μεταβίβασης φορτίου, μέσω των προτύπων σύρματος και επεκταμένης κλίμακας, σε μικρά τμήματα DNA και συγκεκριμένα σε όλες τις δυνατές αλληλουχίες που αποτελούνται από ένα, δύο και τρία ζεύγη βά-

σεων. Συζητούνται πτυχές του φαινομένου όπως οι μέσες πιθανότητες παρουσίας του επιπλέον φορέα σε κάθε βάση ή ζεύγος βάσεων, το συχνοτικό περιεχόμενο της μεταβίβασης φορτίου και οι καθαροί μέσοι ρυθμοί μεταβίβασης. Επιπλέον, περιλαμβάνεται και μια συγκριτική συζήτηση των προαναφερθέντων προτύπων Ισχυρής Δέσμευσης με την λεπτομερέστερη ab initio μέθοδο με την ονομασία Real-Time Time-Dependent Density Functional Theory. Τα αποτελέσματα των τριών μεθόδων βρίσκονται σε καλή συμφωνία.

Στο Κεφάλαιο 8, η μελέτη επεκτείνεται στην εξέταση της μεταβίβασης φορτίου σε μεγαλύτερες, περιοδικές αλληλουχίες DNA. Αρχικά, εξετάζονται συγκριτικά, στο πλαίσιο των προτύπων σύρματος κι επεκταμένης κλίμακας, όλες οι πιθανές περιοδικές ακολουθίες DNA με μονάδα επανάληψης ένα ή δύο ζεύγη βάσεων. Μελετώνται το φάσμα ενεργειακών ιδιοτιμών, οι πυκνότητες καταστάσεων, τα χάσματα HOMO-LUMO, οι μέσες πιθανότητες παρουσίας του φορέα σε κάθε βάση ή ζεύγος βάσεων, το συχνοτικό περιεχόμενο της μεταβίβασης και οι καθαροί μέσοι ρυθμοί μεταβίβασης. Οι δυο μέθοδοι δίνουν συμπληρωματικά αποτελέσματα, με το πρότυπο επεκταμένης κλίμακας, να παρέχει περισσότερη πληροφορία σχετικά με τη συμπεριφορά του επιπλέον φορτίου. Στη συνέχεια η μελέτη προχωρά στην εξέταση ορισμένων περιοδικών αλληλουχιών DNA με αυξανόμενη μονάδα επανάληψης, οι οποίες αποτελούνται είτε από το ίδιο είτε από διαφορετικά ζεύγη βάσεων. Στη συνεκτική μεταβίβαση υπάρχει πεπερασμένη πιθανότητα παρουσίας του φορέα στις διάφορες θέσεις, όμως οι θέσεις με τις καταλληλότερες επιτόπιες ενέργειες, για δεδομένη αρχική τοποθέτηση του φορέα, ευνοούνται. Στο φυσικό DNA είναι πιο πιθανό να δημιουργηθεί μια οπή σε μια Γουανίνη επειδή έχει το υψηλότερο HOMO από όλες τις βάσεις και ένα ηλεκτρόνιο είναι πιο πιθανό να δημιουργηθεί σε μια θυμίνη επειδή αυτή έχει το χαμηλότερο LUMO από όλες τις βάσεις. Όμως, εδώ δείχνουμε συστηματικά πως – συνεκτικά - εάν, επί παραδείγματι, μια οπή έχει δημιουργηθεί αρχικά σε μια αδενίνη, η μεταβίβαση φορτίου θα γίνει κυρίως μέσω των αδενινών και παρομοίως για άλλες αρχικές συνθήκες.

Στο Κεφάλαιο 9 παρουσιάζονται αποτελέσματα για τη μεταβίβαση φορτίου σε ντετερμινιστικές απεριοδικές (οιονεί κρυσταλλικές και μορφοκλασματικές) και τυχαίες αλληλουχίες DNA, αποτελούμενες είτε από το ίδιο είτε από διαφορετικά ζεύγη βάσεων, στο πλαίσιο του προτύπου σύρματος. Προκύπτει πως, σε ό,τι αφορά τη μεταβίβαση φορτίου από το ένα άκρο της ακολουθίας στο άλλο, η μεταβίβαση διαμέσου ντετερμινιστικών απεριοδικών αλληλουχιών είναι γενικότερα κατά τάξεις μεγέθους μικρότερη από ό,τι διαμέσου περιοδικών αλληλουχιών. Επίσης προκύπτει πως τα απλούστερα περιοδικά πολυμερή, δηλαδή τα ομοπολυμερή, αποτελούν απρόσιτο όριο για τις απεριοδικές αλληλουχίες. Επιπλέον, το τυχαίο ανακάτεμα ζευγών βάσεων των ντετερμινιστικών απεριοδικών αλληλουχιών οδηγεί γενικά σε αμελητέους ρυθμούς μεταβίβασης. Ωστόσο, για διαφορετικές αρχικές τοποθετήσεις του επιπλέον φορέα, προκύπτει ότι στις ντετερμινιστικές απεριοδικές αλληλουχίες υπάρχουν καλά ορισμένες περιοχές στις οποίες υπάρχει πιθανότητα παρουσίας του φορέα. Τέλος, στο Κεφάλαιο αυτό καταγράφονται και ορισμένες παρατηρήσεις σχετικά με τους πειραματικούς ρυθμούς μεταβίβασης φορτίου στη βιβλιογραφία.

List of publications and conferences

We include a list of the published research activity of the author as a Ph.D. candidate.

Publications in refereed international scientific journals

1. K. Lambropoulos, K. Kaklamanis, A. Morphis, M. Tassi, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides, "Wire and extended ladder model predict THz oscillations in DNA monomers, dimers and trimers", J. Phys.: Condens. Matter **28**, 495101 (2016) (19 pages).

2. K. Lambropoulos, M. Chatzieleftheriou, 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) (22 pages).

3. M. Tassi, A. Morphis, K. Lambropoulos, and C. Simserides, "RT-TDDFT Study of Charge oscillations in B-DNA Monomers and Dimers", Cogent Physics **4**, 1361077 (2017) (12 pages).

4. K. Lambropoulos, and C. Simserides, "Electronic structure and charge transport properties of atomic carbon wires", Phys. Chem. Chem. Phys. **19**, 26890–26897 (2017).

5. 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. **2**, 035013 (2018) (19 pages).

6. K. Lambropoulos, C. Vantaraki, P. Bilia, M. Mantela, and C. Simserides, "Periodic polymers with increasing repetition unit: Energy structure and carrier transfer", Phys. Rev. E **98**, 032412 (2018) (14 pages, 23 pages supplementary material).

7. K. Lambropoulos and C. Simserides, "Periodic, quasiperiodic, fractal, Kolakoski, and random binary polymers: Energy structure and carrier transport", Phys. Rev. E **99**, 032415 (2019) (17 pages).

8. M. Mantela, K. Lambropoulos, M. Theodorakou and C. Simserides, "Quasi-Periodic and Fractal Polymers: Energy Structure and Carrier Transfer", Materials **12**, 2177 (2019) (30 pages).

9. 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) (26 pages).

Conference presentations

(presenting author's name is underlined)

1. Biosensors 2016, 26th Anniversary World Congress on Biosensors, 25-27 May 2016, Gothenburg, Sweden.
   -Poster: "Three methods predict THz oscillations in DNA monomers, dimers, and trimers"
   A. Morphis, K. Lambropoulos, M. Tassi, K. Kaklamanis, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides

2. 13th International Conference on Nanosciences & Nanotechnologies (NN16), 5-8 July 2016, Thessaloniki, Greece.
   -Poster: "Charge oscillations in DNA monomers and dimers"
   M. Tassi, A. Morphis, K. Kaklamanis, K. Lambropoulos, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides

3. 32nd Panhellenic Conference on Solid-State Physics and Materials Science, 18-21 September 2016, Ioannina, Greece.
   -Talk: "Extra Carrier Transfer Oscillations in DNA Monomers, Dimers and Trimers"
   M. Tassi, A. Morphis, K. Lambropoulos, K. Kaklamanis, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides

4. 39th Progress In Electromagnetics Research Symposium in Singapore (PIERS 2017 in Singapore), 19-22 November 2017, Singapore.
   -Talk: "Two TB Approaches (a Wire Model and an Extended Ladder Model) as Well as RT-TDDFT, Predict THz Oscillations in DNA Monomers, Dimers and Trimers"
   K. Lambropoulos, K. Kaklamanis, A. Morphis, M. Tassi, G. Georgiadis, and C. Simserides
   -Talk: "Frequency Content of Carrier Oscillations along B-DNA Polymers"
   K. Lambropoulos, M. Mantela, and C. Simserides
   -Poster: "RT-TDDFT Study of Charge Oscillations in B-DNA Monomers and Dimers"
   M. Tassi, A. Morphis, K. Lambropoulos, and C. Simserides

5. Mediterranean Institute of Fundamental Physics (MIFP), March Meeting 2018, 28 February-3 March 2018, Castelli Romani, Rome, Italy.

-Invited talk: "Electronic structure and charge transport properties of atomic carbon wires (carbynes)"
K. Lambropoulos and C. Simserides

6. 33rd Panhellenic Conference on Solid-State Physics and Materials Science, 17-19 September 2018, Nicosia, Cyprus.
-Talk: "Quasi-periodic and fractal polymers: energy structure and carrier transfer"
M. Mantela, K. Lambropoulos, M. Theodorakou, and C. Simserides

7. 41st PhotonIcs & Electromagnetics Research Symposium in Rome (PIERS 2019 in Rome), 17 - 20 June 2019, Rome, Italy.
-Talk: "Frequency Content of Carrier Oscillations along B-DNA Aperiodic and Periodic Polymers"
M. Mantela, K. Lambropoulos, C. Vantaraki, P. Bilia, A. Morphis, and C. Simserides

Conference proceedings

1. K. Lambropoulos, M. Mantela, and C. Simserides, "Frequency Content of Carrier Oscillations along B-DNA Polymers", 2017 Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL), Singapore, 19–22 November.

2. M. Mantela, K. Lambropoulos, C. Vantaraki, P. Bilia, A. Morphis, and C. Simserides, "Frequency Content of Carrier Oscillations along B-DNA Aperiodic and Periodic Polymers", PhotonIcs & Electromagnetics Research Symposium in Rome (PIERS 2019 in Rome), *to be published.*

Other

1. Participation of the group "Physics of Nanostructures and Biomaterials" at Researchers Night at National Technical University of Athens (NTUA).
(a) 30 September 2016, Ch. Zacharaki, M. Mantela, K. Lambropoulos, A. Kosma, M. Tassi, A. Morphis, C. Simserides
(b) 29 September 2017, M. Tassi, A. Morphis, K. Lambropoulos, K. Kaklamanis, M. Mantela, Ch. Zacharaki, M. Theodorakou, C. Simserides

2. Participation of the group "Physics of Nanostructures and Biomaterials" in the National and Kapodistrian University of Athens Open Day.
(a) 24 November 2017, K. Kaklamanis, K. Lambropoulos (presenters)
(b) 2 November 2018, K. Lambropoulos, M. Mantela (presenters)
(c) 1 March 2019, K. Lambropoulos, M. Mantela, C. Simserides (presenters)

3. Talks in the annual PhD student seminars organized by the Section of Condensed Matter Physics of the Physics Department, National and Kapodistrian University of Athens.
   (a) K. Lambropoulos, "Electronic structure and charge transfer in periodic DNA sequences: description within the wire model and the extended ladder, model", 10 February 2017.
   (b) K. Lambropoulos, "Charge transport in one-dimensional periodic systems within the wire model: application in carbynes", 9 February 2018.
   (c) K. Lambropoulos, "Periodic, aperiodic, and random DNA sequences: Energy structure and charge transport", 8 February 2019.

List of abbreviations

We include a list of terms that are abbreviated in the present thesis, in alphabetical order.

A	Adenine
ACF	Autocorrelation Function
ACS	Asymmetric Cantor Set
BLA	Bond Length Alternation
C	Cytosine
CS	Cantor Set
DFT	Density Functional Theory
DNA	Deoxyribonucleic Acid
DOS	Density of States
ELM	Extended Ladder Model
FIR	Far-Infrared
FLM	Fishbone Ladder Model
FM	Fishbone Model
G	Guanine
GCS	Generalized Cantor Set
GTM	Global Transfer Matrix
HOMO	Highest Occupied Molecular Orbital
IDOS	Integrated Density of States
LM	Ladder Model
LUMO	Lowest Unoccupied Molecular Orbital
MIR	Mid-Infrared
NIR	Near-Infrared
PD	Period-Doubling
RNA	Ribonucleic Acid
RS	Rudin-Shapiro
RT-TDDFT	Real-Time Time-Dependent Density Functional Theory
T	Thymine
TB	Tight-Binding
TC	Transmission Coefficient
TM	Thue-Morse
TWMF	Total Weighted Mean Frequency
U	Uracil
UCTM	Unit Cell Transfer Matrix
vHS	van Hove Singularities
WM	Wire Model
WMF	Weighted Mean Frequency

Part I

Introduction

1

Charge transport and transfer in DNA and carbynes

Nucleic acids are polymeric macromolecules consisting of units that are called nucleotides. The term nucleic acids is the collective 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, 2]. The nucleotides of each strand are composed of one of four planar, heterocyclic, nitrogenous bases, i.e., guanine (G), cytosine (C), adenine (A) or thymine (T), a pentose sugar (deoxyribose), and a phosphate group. The four nitrogenous bases, together with uracil (U), which is exclusively found in RNA, are depicted in Fig. 1.1. Covalent, phosphodiester bonds between pentoses and phosphate groups of adjacent nucleotides form an alternating sugar-phosphate backbone. These phosphodiester bonds are formed between the $3'$ and $5'$ carbons of the sugars of successive nucleotides, so the respective $5'$ and $3'$ carbons at each end of the strand remain unbonded. Thus, DNA is said to have directionality. 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. Moving along the DNA growth axis, if the first strand has directionality $5' - 3'$, the second strand has directionality $3' - 5'$. The 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 re-

placed by U. Furthermore, RNA molecules are single-stranded; however, some viruses possess double-stranded RNA (other viruses can contain even single-stranded DNA).

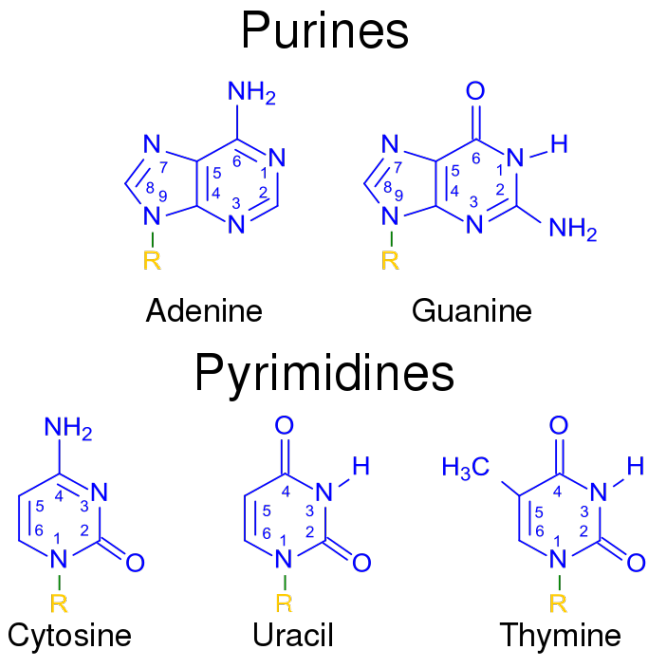

Figure 1.1: The nitrogenous bases found in nucleic acids. Adenine and Guanine (purines, composed of two rings) and Cytosine (pyrimidine, composed of one ring) are native in both DNA and RNA. Thymine / Uracil (pyrimidines) are native only in DNA / RNA. Image source: https://commons.wikimedia.org/wiki/File:Nucleotides_1.svg. Cropped from the original.

In B-DNA, which is the most common DNA conformation, successive base pairs are separated and twisted approximately by 3.4 Å and $36°$, respectively, relatively to the (right-handed) double helix growth axis (Fig. 1.2).

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, due to the overlap of their π molecular orbitals, 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 location, moves to a more favorable one, while transport implies the use of electrodes and the application of external voltage between them.

Charge transfer in biological molecules attracts considerable scientific attention, because it constitutes the basis of many biological processes e.g. in various proteins [3] including metalloproteins [4] and enzymes [5] with medical and bioengineering applications [6, 7]. Charge transfer plays a central role in DNA damage and repair [8–10]. Charge transport might be an indicator to distinguish pathogenic from non-pathogenic mutations at an early stage [11]. It could also probe DNA of different origin or organisms [12], mutations, and diseases [13, 14]. From a physicist's point of view, the charge transfer and transport properties of DNA are studied in order to obtain a deeper understanding of its biological functions as well as for potential applications, such as

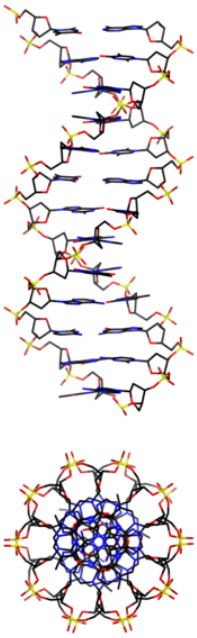

Figure 1.2: Schematic representation of B-DNA, i.e., the most common DNA conformation. (top) Front view. (bottom) top view. Since successive base pairs are separated and twisted approximately by 3.4 Å and 36°, respectively, then, approximately, the helix step is 34 Å and contains 10 base pairs. Image source: Dna-conformations by Mauroesguerroto, used under CC BY-SA 4.0. Cropped from the original.

nanosensors, nanocircuits or molecular wires, due to their high yield synthesis, near-unity purification, and nanoscale self-organization [15–17].

Apart from its biological and nanoscientific importance, DNA can be viewed as a prototype system for one-dimensional molecular wires. The remarkable reduction of channel lengths in electronics during the past decades (from $L \sim 1$ µm in the 1980s to $L \sim 10$ nm in 2014, cf. Fig. 1.3, which for a lattice constant of ~ 0.5 nm corresponds to 20 atoms) raises the question whether polymers, oligomers, (macro)molecules etc. can be recruited in devices as channels. Stiffness is one of the necessary characteristics such structures should posses for this purpose. DNA is among the stiffest known polymers with $\ell_p \approx 50$ nm or 150 base pairs [18], where ℓ_p is the persistence length. The persistence length of a polymer somehow quantifies its stiffness, in the sense that a piece shorter than ℓ_p behaves rather like a flexible elastic beam, while much longer pieces are more likely to bend. On the other hand, if we stretch and join the DNA of all chromosomes of a single cell, that would give us a length of the order of a meter and would consist of billions of base pairs. Apart from its persistence length, DNA provides the ability to construct sequences of interest based on an alphabet, i.e., {G, A, C, T}, so to potentially tailor the properties of a device by choosing the appropriate ones for a given task. This way, e.g. various periodic as well as deterministic aperiodic (e.g., quasiperiodic, fractal) or random sequences can be constructed. Hence, this gives us the opportunity to study various aspects of order or disorder, using DNA sequences, cf. Fig. 1.4.

For at least 20 years, we try to understand carrier movement through DNA [19–29]. Today, we know that 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. Favoring geometries and base-pair sequences, the use of non-natural bases [30], isomers and

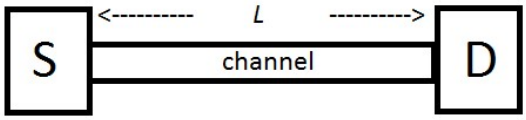

Figure 1.3: Schematic representation of a Source (S)–Channel–Drain (D) device. The urge to achieve continuously smaller channel lengths, L, in such devices, raises the question whether (macro)molecules such as DNA oligomers and polymers or carbynes can be recruited.

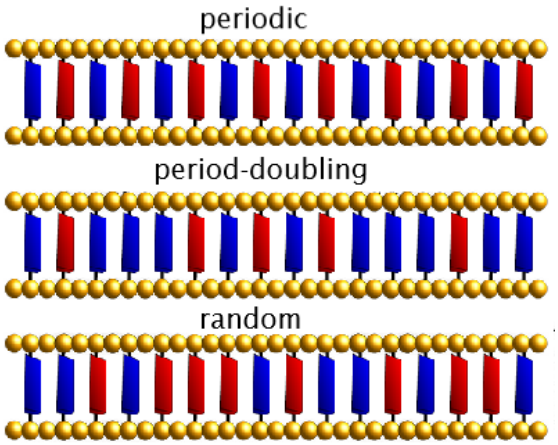

Figure 1.4: The ability to build nucleic acid sequences of choice, e.g. periodic (top), deterministic aperiodic (period doubling, middle) or random (bottom), based on the alphabet constructed by its building blocks, gives the ability to choose the appropriate segment for a given task. This picture refers, for simplicity, to binary double-stranded sequences, i.e., built by two letters only, e.g. G and A (with the complementary bases, C and T, respectively, being implied).

tautomers of the bases [31], the use of the triplet acceptor anthraquinone for hole injection [32], and so on, are being investigated. Structural fluctuations [33] is another important factor which influences carrier movement through DNA [34–37]. Charge transport in oxidatively damaged DNA under structural fluctuations has also been investigated [38]. Also, electric charge oscillations govern the serum response factor-DNA recognition [39]. G runs support delocalization over four to five G bases and resistance oscillations in such DNA segments have been observed theoretically and experimentally [40]. Finally, the carrier transfer rate through DNA can be manipulated by chemical modification [41]. Various degrees of order or disorder are present in DNA, e.g. random attachment of cations (K^+, Ca^{2+}, Na^+) due to the presence of negative charges at the backbone, structural variability, since DNA is soft and its structure changes, (generally) non periodicity in the base sequence, etc. In this thesis, we mainly focus on maybe the most important of the intrinsic factors which also induces various types of order or disorder, i.e. the effect of alternating the base sequence, which affects the electronic overlaps across the π-stack.

Charge movement is usually ascribed to two types of mechanisms [42]: (i) *incoherent* or *thermal hopping* between nearest neighboring or more distant sites and (ii) *coherent hopping* or *tunneling* or *superexchange*. The term tunneling implies quantum

mechanical tunneling, between two sites, e.g., the carrier donor and the carrier acceptor, through a bridge. The term superexchange, not to be confused with the similar term in magnetism, emanates from the distant interaction between the two sites, e.g., the donor and the acceptor, through a bridge. The coherent mechanism is expected to dominate carrier movement in the low temperature regime. Typically, in thermal hopping, charge is localized, the carrier exchanges energy with the environment during its transfer and this way it can travel far longer than via the coherent mechanism. If d_0 is a typical nearest neighbor distance, e.g., 3.4 Å, and two base pairs stand off Δr having energy difference ΔE, then, maybe one could presume an equation like $k = k_0 \exp(-\Delta E/k_B T) \exp(-\Delta r/d_0)$ –or a similar one with other mathematical form– to qualitatively describe thermal hopping [43–54]. In this PhD thesis, we will focus on the coherent case. In this case, all sites have finite occupation probabilities, but, for a given parametrization, those with adequate on-site energies and for a particular initial placement of the carrier are more favored [55–60]. In natural DNA, it is more likely that a hole will be created at guanine because its Highest Occupied Molecular Orbital (HOMO) has the largest energy among all bases and an electron will be created at thymine because its Lowest Unoccupied Molecular Orbital (LUMO) has the smallest energy among bases [61]. However, coherently, if, e.g., a hole is created at an adenine, charge transfer will mainly be through adenines; cf. e.g. Ref. [58].

Another, simpler than DNA, π-conjugated structure that could be candidate for a molecular wire is atomic carbon wires or carbynes. They represent the ultimate nanoscale structure, having a thickness of just one atom, and is made of carbon atoms that adopt sp hybridization. Each one of them is connected to its neighbors with $sp_z sp_z \sigma$ bonds and has two (one p_x and one p_y) electrons that form two, energy-degenerate, π-stacks. The two possible conformations of carbynes are cumulenic and polyynic. In the former, there is no difference between the distances of consecutive atoms [i.e., the bond length alternation (BLA) is zero]; in the latter, BLA is non-zero. Carbynes possess interesting properties, such as tunable band gap, extreme stiffness and elastic modulus, high flexibility, and room temperature persistence. They have a persistence length $\ell_p \approx 14$ nm, that is, ~ 100 carbon atoms, which makes them promising for electronic applications.

Both ab initio calculations [62–70] and model Hamiltonians [14, 71–81] 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. In this PhD thesis, we analytically and numerically study the coherent charge transport and transfer properties of DNA sequences, as well as the transport properties of carbynes, in the context of one of the most widely applied theoretical methods, i.e., with Tight-Binding (TB). Our main goal is to gain insight on the interplay between the sequential order or disorder of DNA segments and their charge transport and transfer properties, taking into account both diagonal and off-diagonal disorder in the Hamiltonian by using realistic TB parametrizations. In Chapter 2, we introduce and discuss the TB method and its variants that are used to study the properties of molecular wires more generally, and specifically of DNA. We also present some details regarding one-dimensional aperiodic substitutional sequences. The second part of this PhD thesis is dedicated to charge transport. In Chapter 3, we study the spectral and transmission

properties of periodic wires with a generic unit cell within the simplest variant of TB. The formalism developed there allows us to theoretically (analytically and numerically) study and reproduce experimental results of cumulenic and polyynic carbynes; this is done in Chapter 4. Next, in Chapter 5, we comparatively study the energy structure and charge transport properties of periodic, deterministic aperiodic and random binary DNA sequences. The third part of this PhD thesis focuses on charge transfer. In Chapter 6, we delineate the basic theory for TB modeling of the transfer of an extra hole or electron along DNA, and we also introduce the physical quantities studied. Chapter 7 is dedicated to charge transfer in small DNA segments, i.e., all possible monomers, dimers, and trimers, which are studied within two TB variants. A discussion regarding comparison between TB and the more detailed ab-initio Real-Time Time-Dependent Density Functional Theory (RT-TDDFT) is also included. In Chapter 8, we move on to examine several classes of periodic DNA polymers; monomer-polymers and dimer-polymers, comparatively within two TB variants, and polymers with increasing repetition unit. Finally, in Chapter 9, we discuss results regarding charge transfer in deterministic aperiodic and random DNA sequences, and make some remarks regarding experimental charge transfer rates.

2

Theoretical Framework[1]

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 [83]. 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 somehow 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 [84], 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 [85, 86]. Its main advantages include its intuitive simplicity, the ability it gives to obtain analytic results in several cases, and its low computational cost [87]. 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. The orbitals under consideration here are the HOMO and the LUMO.

[1] A brief version of the content of this chapter can be found published in Ref. [82], under CC BY 4.0.

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. 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 TB system of 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.

2.1 The stationary Tight-Binding system of equations

The HOMO or LUMO Hamiltonian of a system consisting of $N \times L$ sites is

$$\hat{H} = \sum_{n,l} \hat{H}_n^{l(at)} + \hat{V}, \tag{2.1}$$

where $\hat{H}_n^{l(at)}$ is the Hamiltonian of the isolated moiety that lies at monomer n ($n = 1, 2, \ldots, N$) and at chain l ($l = 1, 2, \ldots, L$), and $\hat{V}$ is the difference between the actual energy of the system and the total energy of the isolated moieties. Within TB, we consider the states $\{|n, l\rangle\}$ of the isolated moieties as localized enough so that $\langle n', l'|n, l\rangle = \delta_{n',n}\delta_{l',l}$, i.e., so that they form a complete orthonormal basis. If this was not the case, an overlap matrix should be taken into account. Due to completeness of $\{|n, l\rangle\}$, $\hat{H}$ can be written as

$$\hat{H} = \sum_{n,l} \epsilon_n^{l(at)} |n, l\rangle\langle n, l| + \sum_{n,l,n',l'} \langle n, l|\hat{V}|n', l'\rangle |n, l\rangle\langle n', l'| \tag{2.2}$$

where $\epsilon_n^{l(at)}$ is the eigenenergy of the isolated moiety. If we suppose the vacuum state $|\emptyset\rangle$, and define the operators

$$\hat{a}_{n,l}^\dagger = |n, l\rangle\langle\emptyset| \tag{2.3a}$$

$$\hat{a}_{n,l} = |\emptyset\rangle\langle n, l| \tag{2.3b}$$

as creation and annihilation operators, respectively, of state $|n, l\rangle$, then Eq. (2.2) takes the form

$$\hat{H} = \sum_{n,l} \epsilon_n^{l(at)} \hat{a}_{n,l}^\dagger \hat{a}_{n,l} + \sum_{n,l,n',l'} \langle n, l|\hat{V}|n', l'\rangle \, \hat{a}_{n,l}^\dagger \hat{a}_{n',l'}. \tag{2.4}$$

If, within TB, we consider as non-negligible only the interactions between the moieties in two successive chains in the same monomer (i.e., between n, l and $n, l \pm 1$), between the moieties in two successive monomers in the same chain (i.e, between n, l and $n \pm$

$1, l$), and between diagonally located moieties in the previous and next chain of two successive monomers (i.e., between n, l and $n+1, l \pm 1$ as well as $n-1, l \pm 1$), then the following hopping integrals can be defined:

$$\begin{aligned}
\langle n,l|\hat{V}|n,l\rangle &= t_n^l \\
\langle n,l|\hat{V}|n,l\pm 1\rangle &= t_n^{l,l\pm 1} \\
\langle n,l|\hat{V}|n\pm 1,l\rangle &= t_{n,n\pm 1}^{l} \\
\langle n,l|\hat{V}|n+1,l\pm 1\rangle &= t_{n,n+1}^{l,l\pm 1} \\
\langle n,l|\hat{V}|n-1,l\pm 1\rangle &= t_{n,n-1}^{l,l\pm 1}.
\end{aligned} \tag{2.5}$$

Hence, Eq. (2.4) becomes

$$\begin{aligned}
\hat{H} = \sum_{n,l} \; (\epsilon_n^l \hat{a}_{n,l}^\dagger \hat{a}_{n,l} \; &+ \; t_n^{l,l+1} \hat{a}_{n,l}^\dagger \hat{a}_{n,l+1} \; + \; t_n^{l,l-1} \hat{a}_{n,l}^\dagger \hat{a}_{n,l-1} \; + \\
&+ \; t_{n,n+1}^{l} \hat{a}_{n,l}^\dagger \hat{a}_{n+1,l} \; + \; t_{n,n-1}^{l} \hat{a}_{n,l}^\dagger \hat{a}_{n-1,l} \; + \\
&+ \; t_{n,n+1}^{l,l+1} \hat{a}_{n,l}^\dagger \hat{a}_{n+1,l+1} \; + \; t_{n,n-1}^{l,l+1} \hat{a}_{n,l}^\dagger \hat{a}_{n-1,l+1} \; + \\
&+ \; t_{n,n+1}^{l,l-1} \hat{a}_{n,l}^\dagger \hat{a}_{n+1,l-1} \; + \; t_{n,n-1}^{l,l-1} \hat{a}_{n,l}^\dagger \hat{a}_{n-1,l-1}),
\end{aligned} \tag{2.6}$$

where

$$\epsilon_n^l = \epsilon_n^{l(at)} + t_n^l \tag{2.7}$$

is the on-site energy of the moiety. It is common in TB models to consider only the energy of the isolated moiety as the on-site energy, without the correction that comes from its presence in the lattice, since, typically, in numeric calculations, the parameters are recruited from experimental data obtained for isolated moieties. This is a first approach, but it is not always the case. For example, in Ref. [61], in which a TB parametrization that will be used in this PhD thesis was presented, the on-site energies of Carbon, Nitrogen and Oxygen within planar organic molecules with sp^2 hybridization were obtained (and are different from the isolated atoms ionization energies) by fitting with TB experimental data of the ionization energy and the first $\pi - \pi^*$ transition of those molecules. This experimental data was obtained e.g. by absorption and photoelectron spectroscopy; see, e.g., Refs. [133-148] of Ref. [61], and Ref [88].

From Eq. (2.6) it can be seen that, if we define the matrices

$$\boldsymbol{\epsilon}_n = \begin{pmatrix}
\epsilon_n^1 & t_n^{1,2} & & & \\
t_n^{2,1} & \epsilon_n^2 & t_n^{2,3} & & \\
& \ddots & \ddots & \ddots & \\
& & t_n^{L-2,L-1} & \epsilon_n^{L-1} & t_n^{L-1,L} \\
& & & t_n^{L,L-1} & \epsilon_n^L
\end{pmatrix}, \tag{2.8}$$

which represents the "on-site energy" of the monomer,

$$\boldsymbol{\tau}_{n,n\pm 1} = \begin{pmatrix}
t_{n,n\pm 1}^1 & t_{n,n\pm 1}^{1,2} & & & \\
t_{n,n\pm 1}^{2,1} & t_{n,n\pm 1}^2 & t_{n,n\pm 1}^{2,3} & & \\
& \ddots & \ddots & \ddots & \\
& & t_{n,n\pm 1}^{L-2,L-1} & t_{n,n\pm 1}^{L-1} & t_{n,n\pm 1}^{L-1,L} \\
& & & t_{n,n\pm 1}^{L,L-1} & t_{n,n\pm 1}^{L}
\end{pmatrix}, \tag{2.9}$$

which represents the "transfer parameter" from monomer n to monomer $n \pm 1$, and

$$\hat{\boldsymbol{a}}_n = \begin{pmatrix} \hat{\alpha}_{n,1} \\ \hat{\alpha}_{n,2} \\ \vdots \\ \hat{\alpha}_{n,L} \end{pmatrix}, \tag{2.10}$$

which represents the annihilation operator of monomer n, then the Hamiltonian can by compactly written as

$$\hat{H} = \sum_{n=1}^{N} \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\epsilon}_n \hat{\boldsymbol{a}}_n + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_{n,n+1} \hat{\boldsymbol{a}}_{n+1} + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_{n,n-1} \hat{\boldsymbol{a}}_{n-1}. \tag{2.11}$$

Since $\hat{H}$ is an Hermitian operator, i.e.,

$$\hat{H} = \hat{H}^\dagger, \tag{2.12}$$

and given that the on-site energies and the hopping integrals are real numbers, it holds that

$$\boldsymbol{\epsilon_n} = \boldsymbol{\epsilon}_n^T, \tag{2.13a}$$

$$\boldsymbol{\tau}_{n,n\pm1} = \boldsymbol{\tau}_{n\pm1,n}^T. \tag{2.13b}$$

Hence, $\boldsymbol{\epsilon}_n$ is a symmetric matrix containing the on-site energies of all moieties belonging to monomer n, ϵ_n^l, in its main diagonal, and the hopping integrals, $t_n^{ll'}$, between the sites of the monomer that belong to different chains in its superdiagonal and subdiagonal. $\boldsymbol{\tau}_{n,n\pm1}$ are (generally) not symmetric matrices containing the hopping integrals between successive monomers, i.e. between the moieties of a monomer and the ones that lie in the same (main diagonal), next (superdiagonal), and previous (subdiagonal) chain of the next $(n+1)$ or previous $(n-1)$ monomer, i.e. $t_{n,n\pm1}^l$, $t_{n,n\pm1}^{l,l+1}$, and $t_{n,n\pm1}^{l-1,l}$, respectively. Both $\boldsymbol{\epsilon}_n$ and $\boldsymbol{\tau}_{n,n\pm1}$ are tridiagonal matrices of order L. We notice that if we took interactions between more chains into account, the number of non-zero off-diagonals of the matrices in Eqs. (2.8) and (2.9) would increase (by 2 for each additional chain), while, if we took interactions between more monomers into account, the number of $\boldsymbol{\tau}$ matrices would increase (by 2 for each additional monomer).

To simplify the matrix notation, we define

$$\boldsymbol{\tau}_{n,n+1} \coloneqq \boldsymbol{\tau}_n. \tag{2.14}$$

Hence, given Eq. (2.13b)

$$\boldsymbol{\tau}_{n,n-1} = \boldsymbol{\tau}_{n-1,n}^T \coloneqq \boldsymbol{\tau}_{n-1}^T. \tag{2.15}$$

Thus, the Hamiltonian of Eq. (2.11) is rewritten as

$$\hat{H} = \sum_{n=1}^{N} \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\epsilon}_n \hat{\boldsymbol{a}}_n + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_n \hat{\boldsymbol{a}}_{n+1} + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_{n-1}^T \hat{\boldsymbol{a}}_{n-1}. \tag{2.16}$$

The situation is schematically presented in Fig. 2.1.

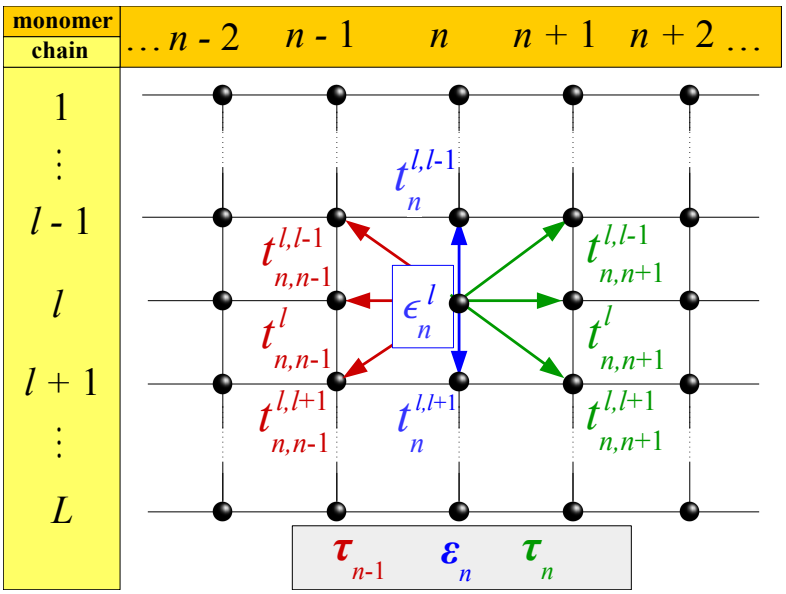

Figure 2.1: Schematic representation of a TB model consisting of N monomers, extended at L chains. Within the model, we take into account (**a**) the on-site energies of each site, ϵ_n^l, and the inter-chain hopping integrals, $t_n^{ll'}$, i.e., between the sites of the monomer (blue), as well as (**b**) the inter-monomer hopping integrals, $t_{nn'}^{ll'}$, 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 $\boldsymbol{\epsilon}_n$, while the latter in the matrices $\boldsymbol{\tau}_{n-1}$ and $\boldsymbol{\tau}_n$, respectively.

Since the states of the isolated moieties, $\{|n,l\rangle\}$, constitute a complete basis, the state of the whole system, $|\Psi\rangle$, can be written as a linear combination of them, i.e.,

$$|\Psi\rangle = \sum_{n,l} \psi_n^l \, |n,l\rangle \,, \tag{2.17}$$

where ψ_n^l is the occupation probability amplitude in the moiety that lies in strand l of monomer n. If we define the state of the isolated monomer n as

$$|n\rangle = \begin{pmatrix} |n,1\rangle & |n,2\rangle & \dots & |n,L\rangle \end{pmatrix} \tag{2.18}$$

and the probability amplitudes that correspond to monomer n as

$$\vec{\boldsymbol{\Psi}}_n = \begin{pmatrix} \psi_n^1 \\ \psi_n^2 \\ \vdots \\ \psi_n^L \end{pmatrix} , \tag{2.19}$$

then $|\Psi\rangle$ can be written as

$$|\Psi\rangle = \begin{pmatrix} |1\rangle & |2\rangle & \dots & |N\rangle \end{pmatrix} \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_2 \\ \vdots \\ \vec{\boldsymbol{\Psi}}_N \end{pmatrix} \tag{2.20}$$

or

$$|\Psi\rangle = \sum_{n=1}^{N} |n\rangle \, \vec{\boldsymbol{\Psi}}_n. \tag{2.21}$$

Since, obviously, $\langle n'|n\rangle = \boldsymbol{I}\delta_{n',n}$, where $\boldsymbol{I}$ is the identity matrix of order L, $\{|n\rangle\}$, form a complete basis, on which the operators $\hat{\boldsymbol{a}}_n^\dagger$ and $\hat{\boldsymbol{a}}_n$ act.

Substituting Eqs. (2.16) and (2.21) in the time-independent Schrödinger equation

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle \tag{2.22}$$

we get

$$\begin{aligned}\sum_{n'} |n'\rangle \, E\vec{\boldsymbol{\Psi}}_{n'} &= \sum_{n,n'} (\hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\epsilon}_n \hat{\boldsymbol{a}}_n + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_n \hat{\boldsymbol{a}}_{n+1} + \hat{\boldsymbol{a}}_n^\dagger \boldsymbol{\tau}_{n-1}^T \hat{\boldsymbol{a}}_{n-1})\,|n'\rangle \, \vec{\boldsymbol{\Psi}}_{n'} \\ &= \sum_{n} |n\rangle \, \boldsymbol{\epsilon}_n \vec{\boldsymbol{\Psi}}_n + |n\rangle \, \boldsymbol{\tau}_n \vec{\boldsymbol{\Psi}}_{n+1} + |n\rangle \, \boldsymbol{\tau}_{n-1}^T \vec{\boldsymbol{\Psi}}_{n-1}.\end{aligned} \tag{2.23}$$

Hence, multiplying by $\langle n''|$ from the left, we conclude that the derivation of the eigenenergies and eigenstates of the system is reduced to the solution of a system of N coupled L-order matrix equations of the form

$$E\vec{\boldsymbol{\Psi}}_n = \boldsymbol{\epsilon}_n \vec{\boldsymbol{\Psi}}_n + \boldsymbol{\tau}_{n-1}^T \vec{\boldsymbol{\Psi}}_{n-1} + \boldsymbol{\tau}_n \vec{\boldsymbol{\Psi}}_{n+1}, \tag{2.24}$$

for all $n = 1, 2, \ldots, N$.

2.2 Solution of the system of equations

The solution of Eq. (2.24) is equivalent with the solution of the problem

$$\begin{pmatrix} E\boldsymbol{I} - \boldsymbol{\epsilon}_1 & -\boldsymbol{\tau}_1 & & & \\ -\boldsymbol{\tau}_1^T & E\boldsymbol{I} - \boldsymbol{\epsilon}_2 & -\boldsymbol{\tau}_2 & & \\ & \ddots & \ddots & \ddots & \\ & -\boldsymbol{\tau}_{N-2}^T & E\boldsymbol{I} - \boldsymbol{\epsilon}_{N-1} & -\boldsymbol{\tau}_{N-1} & \\ & & -\boldsymbol{\tau}_{N-1}^T & E\boldsymbol{I} - \boldsymbol{\epsilon}_N & \end{pmatrix} \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_2 \\ \vdots \\ \vec{\boldsymbol{\Psi}}_{N-1} \\ \vec{\boldsymbol{\Psi}}_N \end{pmatrix} = \begin{pmatrix} \boldsymbol{\tau}_0^T \vec{\boldsymbol{\Psi}}_0 \\ \\ \\ \\ \boldsymbol{\tau}_N \vec{\boldsymbol{\Psi}}_{N+1} \end{pmatrix}. \tag{2.25}$$

The boundaries of the problem are determined by $\vec{\boldsymbol{\Psi}}_0$ and $\vec{\boldsymbol{\Psi}}_{N+1}$. From Bloch's theorem, it holds that $\vec{\boldsymbol{\Psi}}_{N+n} = \zeta \vec{\boldsymbol{\Psi}}_n$, where ζ represents a phase factor. We will use this general form for ζ and then impose either fixed of cyclic boundaries ($\zeta = 0$ or $\zeta = 1$, respectively). So, Eq. (2.25) reaches the form of the eigenvalue-eigenvector problem

$$\boldsymbol{H}\vec{\boldsymbol{\Psi}} = E\vec{\boldsymbol{\Psi}}, \tag{2.26}$$

where

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{\epsilon}_1 & \boldsymbol{\tau}_1 & & & \zeta^* \boldsymbol{\tau}_0^T \\ \boldsymbol{\tau}_1^T & \boldsymbol{\epsilon}_2 & \boldsymbol{\tau}_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \boldsymbol{\tau}_{N-2}^T & \boldsymbol{\epsilon}_{N-1} & \boldsymbol{\tau}_{N-1} \\ \zeta \boldsymbol{\tau}_N & & & \boldsymbol{\tau}_{N-1}^T & \boldsymbol{\epsilon}_N \end{pmatrix} \tag{2.27}$$

is the Hamiltonian matrix, which is a block-tridiagonal matrix of order N (with blocks of order L) with perturbed corners, and

$$\vec{\boldsymbol{\Psi}} = \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_2 \\ \vdots \\ \vec{\boldsymbol{\Psi}}_N \end{pmatrix}. \tag{2.28}$$

Since $\boldsymbol{H}$ is Hermitian, it must hold that $\boldsymbol{\tau}_N = \boldsymbol{\tau}_0$. *This matrix can be numerically diagonalized for either fixed or cyclic boundaries.*

2.2.1 The Transfer Matrix method

Another way to address the problem, is to write Eq. (2.24) in the alternate form

$$\begin{pmatrix} \vec{\boldsymbol{\Psi}}_{n+1} \\ \vec{\boldsymbol{\Psi}}_n \end{pmatrix} = \boldsymbol{P}_n(E) \begin{pmatrix} \vec{\boldsymbol{\Psi}}_n \\ \vec{\boldsymbol{\Psi}}_{n-1} \end{pmatrix}, \tag{2.29}$$

with

$$\boldsymbol{P}_n(E) = \begin{pmatrix} \boldsymbol{\tau}_n^{-1}(E\boldsymbol{I} - \boldsymbol{\epsilon}_n) & -\boldsymbol{\tau}_n^{-1}\boldsymbol{\tau}_{n-1}^T \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix}, \tag{2.30}$$

where $\boldsymbol{0}$ is the zero matrix of order L. The block matrix of order 2 (with elements of order L), $\boldsymbol{P}_n$, is called the *Transfer Matrix* of monomer n. From Eq. (2.30) it is obvious that, in order for $\boldsymbol{P}_n$ to be well-defined, $\boldsymbol{\tau}_n$ must be invertible for all n, i.e., it must hold that $\det(\boldsymbol{\tau}_n) \neq 0$[2]. Since we have N monomers, we can write

$$\begin{pmatrix} \vec{\boldsymbol{\Psi}}_{N+1} \\ \vec{\boldsymbol{\Psi}}_N \end{pmatrix} = \boldsymbol{M}_N(E) \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_0 \end{pmatrix}, \tag{2.31}$$

where

$$\boldsymbol{M}_N(E) = \prod_{n=N}^{1} \boldsymbol{P}_n(E) = \boldsymbol{P}_N \boldsymbol{P}_{N-1} \ldots \boldsymbol{P}_2 \boldsymbol{P}_1 \tag{2.32}$$

is called the *Global Transfer Matrix* (GTM) of the system. For $\vec{\boldsymbol{\Psi}}_{N+1} = \zeta \vec{\boldsymbol{\Psi}}_1$ and $\vec{\boldsymbol{\Psi}}_0 = \zeta^* \vec{\boldsymbol{\Psi}}_N$, Eq. (2.31) takes the form

$$\boldsymbol{M}_N(E) \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_0 \end{pmatrix} = \zeta \begin{pmatrix} \vec{\boldsymbol{\Psi}}_1 \\ \vec{\boldsymbol{\Psi}}_0 \end{pmatrix}. \tag{2.33}$$

Hence, ζ is an eigenvalue of the GTM with an eigenvector composed of $\vec{\boldsymbol{\Psi}}_1$ and $\vec{\boldsymbol{\Psi}}_0$. Then, given the eigenvector of the GTM, and through successive applications of Eq. (2.29), the full eigenvector $\vec{\boldsymbol{\Psi}}$ of the Hamiltonian for E can be constructed. Thus E corresponds to an eigenvalue of $\boldsymbol{H}$. In other words, Eq. (2.26), which gives the eigenvalues and eigenvectors of $\boldsymbol{H}$ is satisfied if and only if Eq. (2.33) is simultaneously satisfied [90].

[2]There are cases in which a transfer matrix can be constructed for non-invertible $\boldsymbol{\tau}_n$, but this holds when only $\boldsymbol{\tau}_n$ is nilpotent of index 2, i.e., when $\boldsymbol{\tau}_n^2 = \boldsymbol{0}$ [89]. However, this is not generally true for those models presented for DNA in Sec. 2.3 that include a non-invertible $\boldsymbol{\tau}_n$.

2.2.2 Properties of the GTM

Recursive relations

The block elements of the GTM are recursively given by

$$\boldsymbol{M}_N^{11(12)} = \boldsymbol{\tau}_N^{-1}(E\boldsymbol{I} - \boldsymbol{\epsilon}_N)\boldsymbol{M}_{N-1}^{11(12)} - \boldsymbol{\tau}_N^{-1}\boldsymbol{\tau}_{N-1}^T\boldsymbol{M}_{N-2}^{11(12)}, \tag{2.34a}$$

$$\boldsymbol{M}_N^{21(22)} = \boldsymbol{M}_{N-1}^{11(12)}, \tag{2.34b}$$

with initial conditions

$$\boldsymbol{M}_1^{11} = \boldsymbol{\tau}_1^{-1}(E\boldsymbol{I} - \boldsymbol{\epsilon}_1), \tag{2.35a}$$

$$\boldsymbol{M}_1^{12} = -\boldsymbol{\tau}_1^{-1}\boldsymbol{\tau}_N^T, \tag{2.35b}$$

$$\boldsymbol{M}_0^{11} = \boldsymbol{I}, \tag{2.35c}$$

$$\boldsymbol{M}_0^{12} = \boldsymbol{0}. \tag{2.35d}$$

From Eqs. (2.34) and (2.35) it is obvious that the elements of $\boldsymbol{M}_N$ occur, for every N, from the knowledge of the first row of $\boldsymbol{M}_{N-1}$ and $\boldsymbol{M}_{N-2}$.

Determinant

The determinant of the GTM is

$$\det(\boldsymbol{M}_N) = \prod_{n=N}^{1}\det(\boldsymbol{P}_n) = \prod_{n=N}^{1}\frac{\det\left(\boldsymbol{\tau}_{n-1}^T\right)}{\det(\boldsymbol{\tau}_n)} = \frac{\det\left(\boldsymbol{\tau}_0^T\right)}{\det(\boldsymbol{\tau}_N)} = 1. \tag{2.36}$$

In other words, the Hermitianity of the Hamiltonian imposes that the GTM is unimodular by construction.

Symplectic property

The inverse Transfer Matrix of a monomer has the form [91]

$$\boldsymbol{P}_n(E)^{-1} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -(\boldsymbol{\tau}_{n-1}^T)^{-1}\boldsymbol{\tau}_n & (\boldsymbol{\tau}_{n-1}^T)^{-1}(E\boldsymbol{I} - \boldsymbol{\epsilon}_n) \end{pmatrix}. \tag{2.37}$$

The inverse GTM is

$$\boldsymbol{M}_N^{-1}(E) = (\boldsymbol{P}_N\boldsymbol{P}_{N-1}\dots\boldsymbol{P}_2\boldsymbol{P}_1)^{-1} = \boldsymbol{P}_1^{-1}\boldsymbol{P}_2^{-1}\dots\boldsymbol{P}_{N-1}^{-1}\boldsymbol{P}_N^{-1}. \tag{2.38}$$

If we define the matrices

$$\mathfrak{s}_n = \begin{pmatrix} O & \boldsymbol{\tau}_n^T \\ -\boldsymbol{\tau}_n & O \end{pmatrix}, \quad \forall n, \tag{2.39}$$

for which it holds that [91]

$$\mathfrak{s}_n^{-1} = \begin{pmatrix} O & -\boldsymbol{\tau}_n^{-1} \\ (\boldsymbol{\tau}_n^T)^{-1} & O \end{pmatrix}, \tag{2.40}$$

we notice that

$$\boldsymbol{P}_n^{-1} = \mathfrak{s}_{n-1}^{-1}\boldsymbol{P}_n^T\mathfrak{s}_n. \tag{2.41}$$

Hence, the inverse GTM becomes

$$\boldsymbol{M}_N^{-1} = \mathfrak{s}_0^{-1}\boldsymbol{P}_1^T\boldsymbol{P}_2^T\ldots\ \boldsymbol{P}_{N-1}^T\boldsymbol{P}_N^T\mathfrak{s}_N. \tag{2.42}$$

Thus, since $\mathfrak{s}_N = \mathfrak{s}_0$, we conclude that

$$\boldsymbol{M}_N^{-1} = \mathfrak{s}_N^{-1}\boldsymbol{M}_N^T\mathfrak{s}_N. \tag{2.43}$$

From Eq. (2.43), it also follows that

$$\boldsymbol{M}_N^T\mathfrak{s}_N\boldsymbol{M}_N = \mathfrak{s}_N. \tag{2.44}$$

Eq. (2.44) defines the *symplectic property* of the GTM. The GTM itself is called symplectic. This is the property that imposes that the GTM is unimodular. Hence, the GTM belongs in the real symplectic group $\mathrm{Sp}(2L,\mathbb{R})$, which is a subgroup of the special linear group $\mathrm{SL}(2L,\mathbb{R})$ of the unimodular real square matrices of order $2L$.

Furthermore, the symplectic property [Eq. (2.44)] implies that

$$\begin{aligned}
\boldsymbol{M}_N^T\mathfrak{s}_N\boldsymbol{M}_N\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix} &= \mathfrak{s}_N\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix} \xRightarrow{\text{Eq. (2.33)}}\\
\boldsymbol{M}_N^T\mathfrak{s}_N\zeta\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix} &= \mathfrak{s}_N\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix} \Rightarrow\\
\boldsymbol{M}_N^T\mathfrak{s}_N\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix} &= \zeta^*\mathfrak{s}_N\begin{pmatrix}\vec{\boldsymbol{\Psi}}_1\\ \vec{\boldsymbol{\Psi}}_0\end{pmatrix}.
\end{aligned} \tag{2.45}$$

Thus ζ^* is an eigenvalue of M_N^T. Due to Eq. (2.43), $\boldsymbol{M}_N^T$ and $\boldsymbol{M}_N^{-1}$ are similar, hence ζ^* is also an eigenvalue of $\boldsymbol{M}_N^{-1}$. Then, given that $\boldsymbol{M}_N$ and $\boldsymbol{M}_N^T$ have the same eigenvalues (since they are transpose to one another), then the eigenvalues of GTM, its transpose and its inverse will come in the pairs ζ and ζ^*.

2.2.3 Imposing the boundaries

Fixed Boundaries

For fixed boundaries, i.e. when $\zeta = \zeta^* = 0$ ($\vec{\boldsymbol{\Psi}}_0 = \vec{\boldsymbol{\Psi}}_{N+1} = 0$), Eq. (2.33) takes the form

$$\boldsymbol{M}_{\boldsymbol{N}}^{\boldsymbol{11}}\vec{\boldsymbol{\Psi}}_1 = 0\vec{\boldsymbol{\Psi}}_1, \tag{2.46}$$

hence, the energy eigenvalues of the system can be found by the condition

$$\det\left(\boldsymbol{M}_N^{11}\right) = 0. \tag{2.47}$$

Cyclic Boundaries

For non-trivial ζ, the general form of the characteristic polynomial of $\boldsymbol{M}_N$ is

$$\det(\zeta I_{2L} - \boldsymbol{M}_N) = 0, \tag{2.48}$$
$$\det(\zeta^* I_{2L} - \boldsymbol{M}_N) = 0, \tag{2.49}$$

since $\boldsymbol{M}_N$ is symplectic. Eq. (2.48) can equivalently be written as [92]

$$\sum_{l=0}^{2L} (-1)^l A_l(E) \zeta^l = 0, \tag{2.50}$$

where

$$A_l(E) = \frac{1}{l} \sum_{p=1}^{l} (-1)^{p+1} A_{l-p}(E) \operatorname{Tr}((\boldsymbol{M}_N)^p), \quad \forall l > 0 \tag{2.51}$$

with initial condition $A_0(E) = 1$. $\operatorname{Tr}(\boldsymbol{M}_N)$ is the trace of $\boldsymbol{M}_N$. The coefficients $A_l(E)$ are $2L+1$ in number. However, if we multiply Eq. (2.50) by $(\zeta^*)^{2L}$ we can deduce that

$$A_l(E) = A_{2L-l}(E), \quad \forall l. \tag{2.52}$$

Thus, the characteristic polynomial is palindromic and their number of different $A_l(E)$ coefficients is reduced to $L+1$. Exploiting this fact, we multiply Eq. (2.50) by $(\zeta^*)^l$, to get

$$(-1)^L A_L(E) + \sum_{l=0}^{L-1} (-1)^l A_l(E) [\zeta^{L-l} + (\zeta^*)^{(L-l)}] = 0. \tag{2.53}$$

Hence, for cyclic boundaries ($\zeta = \zeta^* = 1$), the eigenvalues of the system are given by solving the polynomial

$$(-1)^L \frac{A_L(E)}{2} + \sum_{l=0}^{L-1} (-1)^l A_l(E) = 0, \tag{2.54}$$

i.e., only GTM traces up to the Lth power are needed.

If the cyclically bounded system is periodic, with a unit cell of u monomers repeated m times, then $\boldsymbol{M}_N = (\boldsymbol{M}_u)^m$, $\boldsymbol{M}_u$ is the *Unit Cell Transfer Matrix* (UCTM). In this case, Bloch theorem implies that $\zeta = e^{ikua}$, where ua is the lattice constant, so Eq. (2.53) for the UCTM takes the form of the dispersion relation (with a change in indices)

$$(-1)^L \frac{A_L(E)}{2} + \sum_{l=1}^{L} (-1)^{L-l} A_{L-l}(E) \cos(kual) = 0. \tag{2.55}$$

Furthermore, by definition,

$$\cos(kual) = T_l(x), \tag{2.56}$$

where $T_l(x)$ are the Chebyshev polynomials of the first kind, of degree l with respect to the variable $x := \cos(kua)$, which occur from the determinant of a tridiagonal matrix of order l that has the form

$$T_l(x) = \begin{vmatrix} x & 1 & & \\ 1 & 2x & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & 2x \end{vmatrix}. \tag{2.57}$$

They are recurrently given by

$$T_l(x) = 2xT_{l-1}(x) - T_{l-2}(x), \quad \forall l \geq 2, \tag{2.58}$$

with initial conditions

$$T_1(x) = x, \tag{2.59a}$$

$$T_0(x) = 1. \tag{2.59b}$$

The coefficients of the Chebyshev polynomials of the first kind up to degree 21 can be found, e.g., in Tables C.2a and C.2b of Ref. [93].

Hence, Eq. (2.53) reaches the form

$$(-1)^L \frac{A_L(E)}{2} + \sum_{l=1}^{L} (-1)^{L-l} A_{L-l}(E) T_l(x) = 0. \tag{2.60}$$

Eq. (2.60) is a polynomial of degree L with respect to $x = \cos(kua)$, the solutions of which give the energy eigenvalues of the system in the form of a dispersion relation. The only things to determine are the traces of the powers of the UCTM (cf. Eq. (2.51)) and the Chebyshev polynomials of the first kind up to degree L. For each of the $-m$ in number– values of k, the dispersion relation gives L solutions for x, each of which is a polynomial of energy of degree u, thus creating uL energy bands. Furthermore, since $\cos(kua) = \cos(-kua)$, the number of discrete eigenvalues will be $\frac{N+2u}{2}L$ for even m and $\frac{N+u}{2}L$ for odd m. In other words, there are $\frac{N-2u}{2}L$ degenerate eigenvalues for even m and $\frac{N-u}{2}L$ degenerate eigenvalues for odd m.

2.3 Models for DNA

The above analysis holds for any TB system composed of N monomers and L chains. As far as nucleic acids –and especially DNA– are concerned, the most widely applied 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), which are briefly described below. The form of the matrices involved in Eq. (2.24) for each of these models is presented in Table 2.1.

Table 2.1: Form of the matrices $\vec{\boldsymbol{\Psi}}_n$, $\boldsymbol{\epsilon}_n$, and $\boldsymbol{\tau}_n$ in the TB system of equations [Eq. (2.24)] for several models used to describe nucleic acids and especially DNA: 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	$\vec{\boldsymbol{\Psi}}_n$	$\boldsymbol{\epsilon}_n$	$\boldsymbol{\tau}_n$
WM	1	ψ_n	ϵ_n	$t_{n,n+1}$
LM	2	$\begin{pmatrix}\psi_n^1\\ \psi_n^2\end{pmatrix}$	$\begin{pmatrix}\epsilon_n^1 & t_n^{1,2}\\ t_n^{2,1} & \epsilon_n^2\end{pmatrix}$	$\begin{pmatrix}t_{n,n+1}^1 & 0\\ 0 & t_{n,n+1}^2\end{pmatrix}$
ELM	2	$\begin{pmatrix}\psi_n^1\\ \psi_n^2\end{pmatrix}$	$\begin{pmatrix}\epsilon_n^1 & t_n^{1,2}\\ t_n^{2,1} & \epsilon_n^2\end{pmatrix}$	$\begin{pmatrix}t_{n,n+1}^1 & t_{n,n+1}^{1,2}\\ t_{n,n+1}^{2,1} & t_{n,n+1}^2\end{pmatrix}$
FM	3	$\begin{pmatrix}\psi_n^1\\ \psi_n^2\\ \psi_n^3\end{pmatrix}$	$\begin{pmatrix}\epsilon_n^1 & t_n^{1,2} & 0\\ t_n^{2,1} & \epsilon_n^2 & t_n^{2,3}\\ 0 & t_n^{3,2} & \epsilon_n^3\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0\\ 0 & t_{n,n+1}^2 & 0\\ 0 & 0 & 0\end{pmatrix}$
FLM	4	$\begin{pmatrix}\psi_n^1\\ \psi_n^2\\ \psi_n^3\\ \psi_n^4\end{pmatrix}$	$\begin{pmatrix}\epsilon_n^1 & t_n^{1,2} & 0 & 0\\ t_n^{2,1} & \epsilon_n^2 & t_n^{2,3} & 0\\ 0 & t_n^{3,2} & \epsilon_n^3 & t_n^{3,4}\\ 0 & 0 & t_n^{4,3} & \epsilon_n^4\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & t_{n,n+1}^2 & 0 & 0\\ 0 & 0 & t_{n,n+1}^3 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$

2.3.1 Wire Model

WM is the simplest TB model to describe nucleic acids and analogues [94, 95]. It can be applied to mimic either single-stranded nucleic acids and hairpins at the single-base level [96] or double-stranded ones [55] 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 2.2(a).

2.3.2 Ladder Model

LM is the simplest model that can address the influence of base-pairing in the energetics of nucleic acids [94, 97]. 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 intra-strand hopping integrals between successive bases, $t_{n,n\pm1}^{ll}$, and the intra-base-pair hopping integrals, $t_n^{ll'}$, due to the hydrogen bonds between the complementary bases in a pair. A schematic representation of the LM is shown in Figure 2.2(b).

2.3.3 Extended Ladder Model

ELM is a more detailed version of the LM, including the inter-strand hopping integrals, $t^{ll'}_{n,n\pm1}$, between the bases of successive base pairs [58, 97]. A schematic representation of the ELM is shown in Figure 2.2(c).

2.3.4 Fishbone Model

FM is the simplest model that can take into account the effect of the sugar-phosphate backbone [71, 77, 94]. 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 sugars are separated by phosphate groups from one another [15, 98, 99]. Hence, the parameters needed for its employment are the on-site energies, ϵ^l_n, of the base pairs ($l = 2$) and of the backbone sites ($l = 1, 3$), the intra-strand hopping integrals between successive base pairs, $t^{2,2}_{n,n\pm1}$, and the inter-strand hopping integrals, $t^{ll'}_n$, between the base pairs and the backbone sites. A schematic representation of the FM is shown in Figure 2.2(d).

2.3.5 Fishbone Ladder Model

FLM is a combination of the LM and the FM [77, 94]. 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^l_n, of the base pairs ($l = 2, 3$) and the backbone ($l = 1, 4$), the intra-strand hopping integrals between base pairs, $t^{ll}_{n,n\pm1}$ ($l = 2, 3$) and the inter-strand hopping integrals between the bases of a base pair as well as between each base and the backbone, $t^{ll'}_n$. A schematic representation of the FLM is shown in Figure 2.2(e).

2.3.6 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 [75, 97, 100], single-stranded nucleic acids with a backbone [100] and explicit inclusion of helicity [101] strain [102], and spin–orbit coupling [103] 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 become *energy-dependent*. This procedure is important when environmentally induced effects are considered [77]. For example, the FLM can be reduced into an LM via a one-step

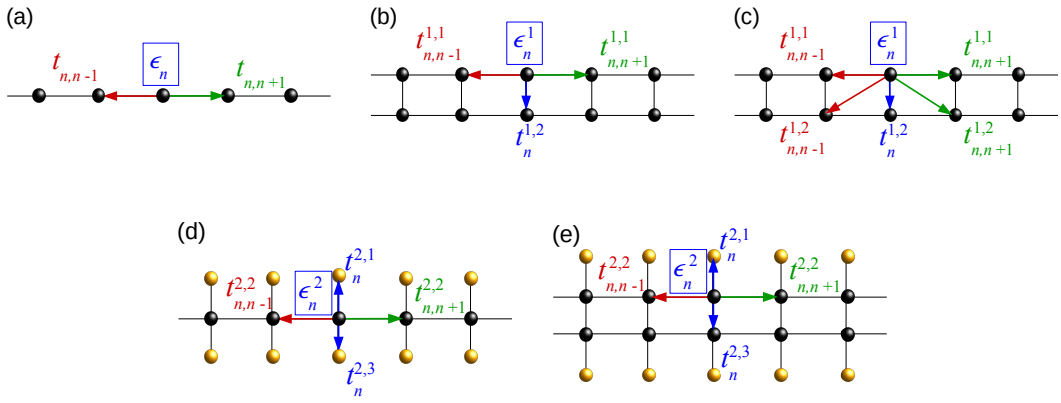

Figure 2.2: Schematic representation of the TB models listed in Table 2.1. (a) Wire Model (WM); (b) Ladder Model (LM); (c) Extended Ladder Model (ELM); (d) Fishbone Model (FM); (e) Fishbone Ladder Model (FLM).

renormalization procedure [104], or to an even simpler WM via a two-step renormalization procedure [105, 106].

Several techniques can be applied to solve the models, depending on what is studied, such as the numerical diagonalization of the Hamiltonian in Equation (2.16) [56, 58, 59], the transfer matrix method [89, 107, 108] outlined above, and the Non-Equilibrium Green's Function technique [109]. As discussed in Subsec. 2.2.1, the transfer matrix method is not applicable if the matrices $\boldsymbol{\tau}_n$ are singular. Generally, this is the case, e.g., for the FM and the FLM (cf. Table 2.1). 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. [61, 64, 65, 68, 110], for on-site energies, cf. Refs. [31, 111–116], and for hopping integrals, cf. Refs. [117–119]. Such parametrizations allow researchers to go beyond 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. In this PhD thesis, we take into account *both diagonal and off-diagonal disorder using such realistic TB parametrizations.*

2.4 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 $Al_{0.86}Mn_{0.14}$ alloys [120] (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" [121]. 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. Since their first discovery, more than 100 types of synthetic quasicrystals have been grown. Up to 2016, three naturally formed quasicrystals have been discovered in the Khatyrka meteorite; the first one, i.e., $Al_{63}Cu_{24}Fe_{14}$ has icosahedral symmetry, and is called icosahedrite [122]; the second one, $Al_{71}Ni_{24}Fe_5$, has decagonal symmetry, and is called decagonite [123]; the third one, $Al_{62.0(8)}Cu_{31.2(8)}Fe_{6.8(4)}$, has also icosahedral symmetry, and is the first example of a quasicrystal discovered in nature prior to being synthesized in the laboratory [124]. The electron and X-ray diffraction patterns of natural quasicrystals show a higher degree of structural perfection than their synthetic counterparts [125]. The electron diffraction spectra of icosahedrite and decagonite are presented in Fig. 2.3.

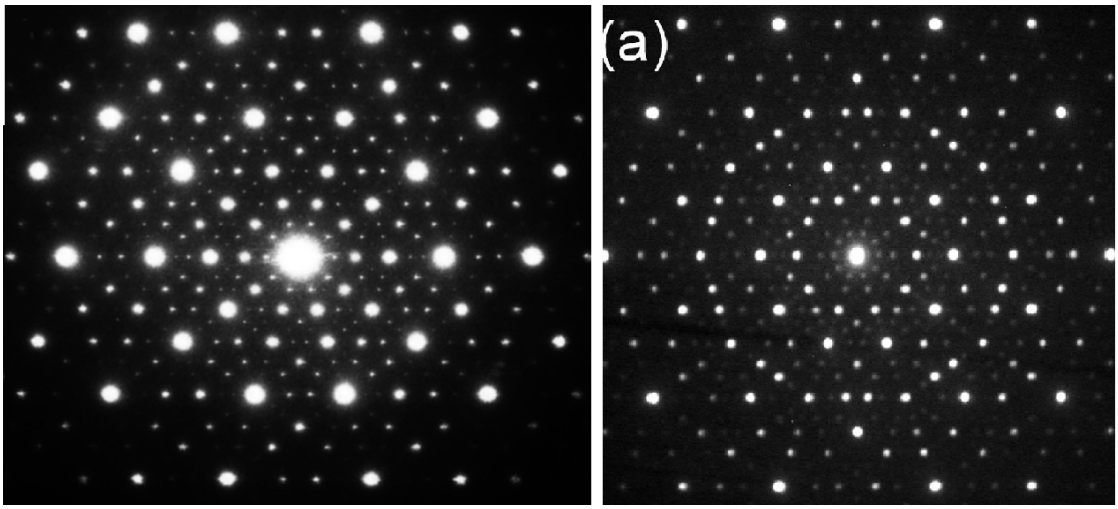

Figure 2.3: (left) The electron diffraction pattern of icosahedrite, the first discovered natural quasicrystal. Image source: https://commons.wikimedia.org/wiki/File:Icosahedrite_Diffraction_Pattern.jpg, used under CC BY-SA 4.0. (right) The electron diffraction pattern of decagonite, the second discovered natural quasicrystal. Image reproduced from Ref. [126], used under CC BY 4.0. Cropped from the original.

Another interesting category of structures are fractals. Fractals do posses inflation/deflation symmetry. However, structurally, fractals constitute a class of aperiodic systems that is distinct from quasicrystals. A quasicrystalline lattice is composed by two (or more) incommensurate lattices, while a fractal one is not [127].

Advances in the field of quasicrystals and other related structures have turned scientific interest into the study of specific one-dimensional aperiodic lattices, modeled with TB [128], that is, described by Eq. (2.24). 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. The ability to produce synthetic, de novo, nucleic acid sequences of interest [129], using mainly the phosphoramidite method [130] (although other promising methods have recently been proposed [131]), provides a chance not only to examine theoreti-

cal 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.

2.4.1 Substitutional Sequences

Aperiodic substitutional sequences are based on an alphabet, e.g., $\mathcal{A} = \{\text{A, B, C, D, }\dots\}$ 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 gth generation of the sequence.

Substitutional sequences can, in most cases, be described by introducing the substitution matrix, $\boldsymbol{S}$. It is a square, non-negative matrix of order $card(\mathcal{A})$ (the cardinality of a set is the number of elements of the set), and its elements are $\boldsymbol{S}_{ij} = n_i[s(j)]$, 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, $\boldsymbol{S}$ 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. Below we present some details about the most widely-known deterministic aperiodic sequences. Those, as well as some other sequences, together with their substitution rules and matrices, are presented in Table 2.2.

Fibonacci

The Fibonacci sequence, named after the Italian mathematician Leonardo Pisano (Fibonacci) who introduced it in his 1212 book *Liber Abaci*, in a study of the population growth of rabbits [132], is a number sequence the terms of which are generated by the addition of the two previous terms, with given initial conditions. However, this sequence appears many centuries before in Indian mathematics, in connection with Sanskrit prosody [133]. For example, the possible ways to arrange short (S) and double, long (L) syllables with given total duration measured as g S syllables is the Fibonacci number of the $g + 1$ generation. If $\mathcal{N}_g$ is the Fibonacci number of generation g, and we set $\mathcal{N}_0 = \mathcal{N}_1 = 1$, the recurrence relation $\mathcal{N}_g = \mathcal{N}_{g-1} + N_{g-2}$ produces the number sequence $1, 1, 2, 3, 5, 8, 13, 21, 34 \dots$. Using the two-letter alphabet $\{\text{A, B}\}$, we can define the Fibonacci word F_g by the substitution rule $s(\text{A}) = \text{AB}$, $s(\text{B}) = \text{A}$, starting with $F_0 = \text{B}$. $F_1 = \text{A}$, $F_2 = \text{AB}$, $F_3 = \text{ABA}$, $F_4 = \text{ABAAB}$, etc. Obviously, the length of the word F_g is $\mathcal{N}_g$.

Thue-Morse

The Thue-Morse (TM) sequence (aka Prouhet-Thue-Morse sequence) was studied by Eugene Prouhet in the field of number theory [134], defined by Alex Thue in the field of combinatorics [135], and rediscovered by Marston Morse in the context of differential geometry [136]. It is a binary sequence of 0s and 1s, starting with 0, with its g^{th} generation constructed by appending the Boolean complement of the previous generation to the sequence. With the two-letter alphabet $\{\text{A}, \text{B}\}$, we can define the TM word TM_g by the substitution rule $s(\text{A}) = \text{AB}$, $s(\text{B}) = \text{BA}$, starting with $TM_0 = \text{A}$. $TM_1 = \text{AB}$, $TM_2 = \text{ABBA}$, $TM_3 = \text{ABBABAAB}$, etc. The length of the word TM_g is 2^g.

Period-Doubling

The Period-Doubling (PD) sequence is closely connected with the TM sequence. Specifically, its elements are given by the first differences of the elements of the TM binary sequence modulo 2. Using the two-letter alphabet $\{\text{A}, \text{B}\}$, we can define the PD word PD_g by the substitution rule $s(\text{A}) = \text{AB}$, $s(\text{B}) = \text{AA}$, starting with $PD_0 = \text{A}$. $PD_1 = \text{AB}$, $PD_2 = \text{ABAA}$, $PD_3 = \text{ABAAABAB}$, etc. The length of the word PD_g is 2^g.

Rudin-Shapiro

The Rudin-Shapiro (RS, aka Golay-Rudin-Shapiro) sequence is the sequence of the appended coefficients of the RS polynomials [137, 138]. It contains only ± 1 and is generated by starting with $+1, +1$ and employing the rules

$$\begin{aligned} +1, +1 &\to +1, +1, +1, -1 \\ +1, -1 &\to +1, +1, -1, +1 \\ -1, +1 &\to -1, -1, +1, -1 \\ -1, -1 &\to -1, -1, -1, +1. \end{aligned}$$

Using the four-letter alphabet $\{\text{A}, \text{B}, \text{C}, \text{D}\}$, we can define the RS word RS_g by the substitution rule $s(\text{A}) = \text{AB}$, $s(\text{B}) = \text{AC}$, $s(\text{C}) = \text{DB}$, $s(\text{D}) = \text{DC}$, starting with $RS_0 = \text{A}$. $RS_1 = \text{AB}$, $RS_2 = \text{ABAC}$, etc. The length of the word RS_g is 2^g.

Cantor Set

The Cantor Set (CS), named after mathematician Georg Cantor who introduced it [139], is one of the most well-known deterministic fractals. It is obtained as follows: given the continuous interval $[0, 1]$, the middle third, $(\frac{1}{3}, \frac{2}{3})$ is deleted, resulting in the union $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Then, the open middle third of each remaining interval is deleted, and the process is repeated *ad infinitum*. Using the two-letter alphabet $\{\text{A}, \text{B}\}$, we can define the CS word CS_g by the substitution rule $s(\text{A}) = \text{ABA}$, $s(\text{B}) = \text{BBB}$, starting with $CS_0 = \text{A}$. $CS_1 = \text{ABA}$, $CS_2 = \text{ABABBBABA}$, etc. All generations are palindromic words. The length of the word CS_g is 3^g.

Generalized Cantor Set

In accordance with the rationale described above, one can imagine the construction of a generalized CS word, $GCS_g(t, d)$, produced by the two-letter alphabet {A, B}, where t is the total number of letters substituting each letter of the sequence in the next generation and d is the number of letters that correspond to the "deleted" middle segment ($t > d$). t and d are mutually odd or even, to preserve the palindromicity of the words. For example, the generalized word $GCS_g(4, 2)$ is given by the rule s(A) = ABBA, s(B) = BBBB, starting with $GCS_0(4, 2) =$ A. The length of the word GCS_g is t^g.

Asymmetric Cantor Set

Similarly, the Asymmetric Cantor Set (ACS) is built by splitting a straight line segment in four, removing the second quarter, then removing the second quarter of each of the three new straight line segments, and the process is repeated *ad infinitum*. Using, e.g., the two-letter alphabet {A, B}, the ACS word ACS_g is defined by the substitution rules s(A) = ABAA, s(B) = BBBB, starting with $ACS_0(4, 2) =$ A. The length of the word GCS_g is t^g.

Kolakoski

The Kolakoski $\{p, q\}$ sequences are a family of sequences of the integers $p \neq q$ that are their own run length encodings (a *run* is defined here as the maximal subsequence of identical numbers). The classic and most well-known sequence of this class, i.e., Kolakoski$(1, 2)$ [140], also referred to as Oldenburger-Kolakoski sequence, was popularized by recreational mathematician William Kolakoski [141], but it was independently introduced by Rufus Oldenburger [142]. This family of sequences possesses different properties in different cases. For example, for specific values of p and q, they may show pure-point or continuous diffraction spectra [143]. Each generation, $KOL_g(p, q)$, of the sequences can be seen as the run length encoding of the next generation, starting with $KOL_0(p, q) = q^p$ and following the substitution rule

$$\begin{aligned} s(q) &= p^q \qquad \text{if } q \text{ was at odd } n, \\ s(q) &= q^q \qquad \text{if } q \text{ was at even } n, \\ s(p) &= p^p \qquad \text{if } p \text{ was at odd } n, \\ s(p) &= q^p \qquad \text{if } p \text{ was at even } n. \end{aligned}$$

For example $KOL_0(1, 2) = 2$, $KOL_1(1, 2) = 11$, $KOL_2(1, 2) = 12$, $KOL_3(1, 2) = 122$, $KOL_4(1, 2) = 12211$, $KOL_5(1, 2) = 1221121$, etc. Accordingly, using the two-letter alphabet {A, B}, we can define the $KOL(p, q)$ word $KOL_g(p, q)$ by assigning A to p and B to q. Thus, e.g., $KOL_5(1, 2)$ = ABBAABA. The length of $KOL(1, 2)$ as the generation increases is given by the OEIS sequence A001083 [144]. Generally, the length of the word $KOL_g(p, q)$ is equal to the sum of the terms of $KOL_{g-1}(p, q)$.

Table 2.2: Some of the most well-known substitutional sequences, together with the alphabets through which they are defined and the corresponding substitution rules and substitution matrices. In the last row, the subscripts o and e denote substitutions that are applied on odd and even positions in the sequence, respectively.

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

2.4.2 Primitive Substitutions and the Perron–Frobenius Eigenvalue

The substitution matrix $\boldsymbol{S}$ (and, hence, the substitution) is called *primitive* if there exists a natural number k such that $\boldsymbol{S}^k$ is a positive matrix. For primitive substitutions, the Perron–Frobenius theorem [145, 146] guarantees that $\boldsymbol{S}$ 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 ($g \to \infty$) relative frequencies of the letters in $\mathcal{A}$.* Hence, using $\boldsymbol{S}$, one can determine the asymptotic occurrence percentage of each nucleotide in a substitutional nucleic acid sequence.

2.4.3 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 $\boldsymbol{S}$ (corresponding to nucleotide k-plets). This can be done as follows [147]; 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 $\boldsymbol{S}_k$ 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 $\boldsymbol{S}$ is defined via a helping alphabet (see, Chapter 5), k-letter word frequencies can be deduced in the same fashion from the legal $2k$-letter words of the helping alphabet.

Let us demonstrate, for illustration, the procedure to determine the induced substitution matrix $\boldsymbol{S}_3$ of the possible 3-letter words of the Period-Doubling sequence, for illustration. In this case, $k = 3$, $W = \{\text{AAA, AAB, ABA, BAA, BAB}\}$ and l is always 2. Hence,

$$
\begin{aligned}
s(\text{AAA}) = \text{ABABAB} &\to s_3(\text{AAA}) = (\text{ABA})(\text{BAB}),\\
s(\text{AAB}) = \text{ABABAA} &\to s_3(\text{AAB}) = (\text{ABA})(\text{BAB}),\\
s(\text{ABA}) = \text{ABAAAB} &\to s_3(\text{ABA}) = (\text{ABA})(\text{BAA}),\\
s(\text{BAA}) = \text{AAABAB} &\to s_3(\text{BAA}) = (\text{AAA})(\text{AAB}),\\
s(\text{BAB}) = \text{AAABAA} &\to s_3(\text{BAB}) = (\text{AAA})(\text{AAB}).
\end{aligned}
$$

So, the induced primitive substitution matrix $\boldsymbol{S}_3$ reads

$$
\boldsymbol{S}_3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}. \tag{2.61}
$$

2.4.4 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 [148]. A substitution has the Pisot property if the matrix $\boldsymbol{S}$ 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 $\boldsymbol{S}$ is irreducible over the rationals, the Pisot substitution is called irreducible. Irreducible Pisot substitutions are a subclass of primitive substitutions [149].

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

Furthermore, according to the Lebesgue's decomposition theorem [150], 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. [151, 152], 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 [153]. (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 (i.e., 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 $\boldsymbol{S}$: unimodular $\boldsymbol{S}$ [i.e. $\det(\boldsymbol{S}) = \pm 1$] implies strict quasiperiodicity, otherwise the structure is limit-quasiperiodic [154–156]. 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 as well as all the precious means sequences [157] and the Fibonacci-class sequences [158] (cf. Table 2.2). *Limit-quasiperiodic* structure representatives are the mixed means sequences with $n \geq m$ [159] (cf. Table 2.2).

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 $n = l(l+1)$ [160],

(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 [134–136] and metallic means sequences with $n \neq l(l+1)$ [160], or even

(5) *absolutely continuous*, such as the Rudin–Shapiro sequence.

Apart from the above-mentioned sequences, there are others for which the substitution is not primitive or the matrix $\boldsymbol{S}$ cannot even be defined at all. Examples of non-primitive substitutions include the sequences inspired by the Cantor set (cf. Table 2.2). A sequence for which a substitution matrix cannot be defined is the classical Kolakoski$(1,2)$ sequence, and generally Kolakoski(p,q) sequences where p is odd and q even or vice versa [143]. The situation is different when p and q are both even or odd; then, a primitive $\boldsymbol{S}$ can be defined (cf. Table 2.2). In the former case, the sequences have been classified as limit-periodic [161]. In the latter case, the irreducible Pisot property holds when $2(p+q) \geq (p-q)^2$, and $\boldsymbol{S}$ is also unimodular when $p = q \pm 2$ [143].

Part II

Charge transport

3

Periodic Wire Model with a generic unit cell[3]

As mentioned previously, the one-dimensional TB lattice with one orbital per site and nearest neighbor hoppings, i.e., the WM, can be applied to study, among others, the charge transfer and transport properties of π-conjugated organic systems which are candidates for molecular or atomic wires, such as single- and double- stranded DNA chains, as well as cumulenic and polyynic carbynes . In this chapter, we focus on the periodic WM with a generic unit cell composed of u cites, which is repeated m times. Hence, we have u different on-site energies and u different hopping integrals. Many works that study the electronic structure of periodic TB models refer only to specific cases and make a number of simplifying assumptions, such as supposing equal hoppings within the lattice and/or between the lattice and the leads. Here, we address the matter in its generality. *We explicitly take into account the difference in the on-site energies and the hopping integrals between different moieties inside the unit cell, as well as between the lattice and the leads.* The solution of the TB system of equations is equivalent to the diagonalization of a real symmetric tridiagonal u-Toeplitz matrix of order $N = mu$ with (cyclic boundaries) or without (fixed boundaries) perturbed upper right and lower left corners. The properties of such matrices have been extensively studied in the literature [162–166] due to their theoretical interest and their many applications. Apart from the interest these systems have in themselves, they could also serve as a starting point to gain more understanding about the *inclusion of chemical complexity effects* in the electronic structure of periodic lattices, as well as of quasi-periodic, fractal and amorphous atomic or molecular lattices.

Various quantum chain models, studied in the context of low-dimensional systems, have been treated with the transfer matrix method [89, 167–170]. Here, we use the

[3]Part of this chapter can be found published in Ref. [108], under CC BY 3.0.

transfer matrix method [171, 172] to solve the TB system of equations and determine expressions for its eigenvalues, for both cyclic and fixed boundary conditions, by combining the *spectral duality relations* [90, 173] with the *connection of the elements of the powers of a* 2×2 *unimodular matrix to the Chebyshev polynomials of the second kind* [127, 174]. For cyclic boundaries, we derive the dispersion relation, while, for fixed boundaries, we show that the spectrum of eigenvalues (eigenspectrum) is produced by a recurrent relation containing the Chebyshev polynomials of the second kind and the elements of the transfer matrix of the unit cell. We discuss the dispersion relations and eigenspectra through representative examples for systems with unit cells composed by $u = 1, 2, 3, 4$ sites, supposing different on-site energies and hopping integrals, that is, *the most general case*. We examine the density of states (DOS) and the occurrence of van Hove Singularities (vHS) and provide analytical relations for $u = 1, 2, 3, 4$. We determine the transmission coefficient (TC) at zero bias, by interconnecting the systems to semi-infinite leads The situation is interpreted in Fig 3.1. Many works have been dedicated to the TC determination, within various contexts cf. e.g. Refs. [175–181]. Here, we analyze in detail the transmission profiles of periodic TB systems, examining the influence of the *strength and asymmetry* of the system-leads coupling, the *leads bandwidth and band center*, and the *intrinsic hoppings* of the system. Such a systematic study reveals the significant effects each parameter induces to the transmission profile, and addresses aspects, such as the coupling asymmetry, that, although are not regularly taken into account in the literature, play a crucial role in the degree of the system's transparency to the incident waves. The number and position of the transmission peaks is determined in all cases, and representative examples are given for unit cells consisting up to $u = 6$ sites. We introduce a *generic optimal coupling condition* and demonstrate its physical meaning. This condition generalizes for any periodic one-dimensional TB chain the condition reported in Ref. [180], where a periodic single-stranded DNA chain with $u = 4$ and symmetric coupling to the leads was studied, with the hopping integrals of the system assumed identical.

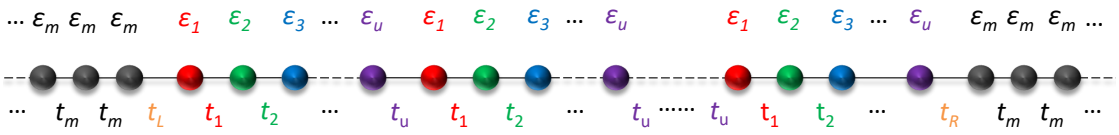

Figure 3.1: Schematic representation of a WM with a generic unit cell of u sites, sandwiched between two semi-infinite metallic leads, and connected to them from the left (right) side with hopping integral t_L (t_R). This setup addresses the situation in its generality, going beyond simplifying assumptions, such as supposing equal hoppings within the lattice and/or between the lattice and the leads.

3.1 Transfer matrix method – Wire Model

The TB system of equations of a WM with N sites is

$$E\psi_n = \epsilon_n\psi_n + t_{n,n-1}\psi_{n-1} + t_{n,n+1}\psi_{n+1}, \tag{3.1}$$

If the system is periodic, with a unit cell composed of u sites, and repeated m times, so that $N = mu$, the solution of Eq. (3.1) is reduced to the diagonalization of the Hamiltonian matrix [cf. Eq. (2.27)]

$$\boldsymbol{H}(\zeta) = \begin{pmatrix} \epsilon_1 & t_1 & & & & & & & \zeta^* t_u \\ t_1 & \epsilon_2 & t_2 & & & & & & \\ & \ddots & \ddots & \ddots & & & & & \\ & & t_{u-2} & \epsilon_{u-1} & t_{u-1} & & & & \\ & & & t_{u-1} & \epsilon_u & t_u & & & \\ & & & & t_u & \epsilon_1 & t_1 & & \\ & & & & & \ddots & \ddots & \ddots & \\ & & & & & & \ddots & \ddots & \ddots \\ \zeta t_u & & & & & & & t_{u-1} & \epsilon_u \end{pmatrix}, \tag{3.2}$$

which is a tridiagonal real symmetric u-Toeplitz matrix of order N, with (for $\zeta = 1$, i.e. for cyclic boundaries) or without (for $\zeta = 0$, i.e. for fixed boundaries) perturbed upper left and lower right corners. Here, we have simplified the notation for the hopping integrals, so that $t_{n,n+1} := t_n, t_{n,n-1} := t_{n-1}$. Hence, for the site with index $n = u+1$, it holds that $t_{u+1} = t_1$. The condition $\det(E\boldsymbol{I}_N - \boldsymbol{H}(\zeta)) = 0$, i.e. the roots of the characteristic polynomial of $\boldsymbol{H}(\zeta)$, which is a polynomial of energy of degree N, gives the eigenvalues of the system.

Within the transfer matrix method, Eq. (3.1) can equivalently be written in the form (cf. Table 2.1)

$$\begin{pmatrix} \psi_{n+1} \\ \psi_n \end{pmatrix} = \begin{pmatrix} \frac{E-\epsilon_n}{t_n} & -\frac{t_{n-1}}{t_n} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_n \\ \psi_{n-1} \end{pmatrix} = \boldsymbol{P}_n(E) \begin{pmatrix} \psi_n \\ \psi_{n-1} \end{pmatrix}, \tag{3.3}$$

hence, the GTM, $\boldsymbol{M}_N(E)$, of the periodic system is

$$\begin{pmatrix} \psi_{N+1} \\ \psi_N \end{pmatrix} = \boldsymbol{M}_N(E) \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix} := \boldsymbol{M}_u(E)^m \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}. \tag{3.4}$$

$\boldsymbol{M}_u(E)$ is the UCTM, i.e.,

$$\boldsymbol{M}_u(E) = \prod_{n=u}^{1} \boldsymbol{P}_n(E). \tag{3.5}$$

Starting from the fact that $\boldsymbol{M}_K = \boldsymbol{P}_K \boldsymbol{M}_{K-1}$, see Eq. (3.5), the elements of the 2×2 UCTM are recurrently given by (for $u \geq 2$)

$$M_u^{11(12)} = \frac{E-\epsilon_u}{t_u} M_{u-1}^{11(12)} - \frac{t_{u-1}}{t_u} M_{u-2}^{11(12)} \tag{3.6a}$$

$$M_u^{21(22)} = M_{u-1}^{11(12)} \tag{3.6b}$$

with initial conditions $M_1^{11} = (E - \epsilon_1)/t_1$, $M_1^{12} = -t_u/t_1$, $M_0^{11} = M_0^{21} = 1$, $M_0^{12} = M_0^{22} = 0$. M^{ij} is the element ij of matrix $\boldsymbol{M}$. Hence, M_u^{11}, M_u^{12}, M_u^{21}, and M_u^{22}, are polynomials of energy of degree u, $u-1$, $u-1$, and $u-2$, respectively[4].

3.1.1 Cyclic Boundaries

For the cyclically bounded periodic system, Bloch theorem implies that $\zeta = 1 = e^{ikmua}$, where ua is the lattice constant of the system. Hence, $k = \frac{2\pi\nu}{mua}$, where ν is an integer taking m values, such that k belongs to the 1st Brillouin Zone. Thus, Eq. (2.33) takes the form

$$\boldsymbol{M}_u(E) \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix} = e^{ikua} \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}. \tag{3.7}$$

From the symplectic property, we know that if e^{ikua} is an eigenvalue of $\boldsymbol{M}_u$, so is e^{-ikua} (see Subsec. 2.2.2). Hence, the eigenvalues of $\boldsymbol{M}_u$ are given by the condition

$$\det\left(e^{ikua}\boldsymbol{I}_2 - \boldsymbol{M}_u\right) = \det\left(e^{-ikua}\boldsymbol{I}_2 - \boldsymbol{M}_u\right) = 0, \tag{3.8}$$

from which we can easily arrive at the dispersion relation

$$\boxed{\cos(kua) = \frac{\mathrm{Tr}(\boldsymbol{M}_u)}{2} := z(E).} \tag{3.9}$$

Substituting each of the N energy eigenvalues obtained by the above dispersion relation to Eq. (3.3), leads, via its iterative application, to the construction of the full corresponding eigenvector of the Hamiltonian matrix $\boldsymbol{H}(\zeta = 1)$ [90]. Hence, the solutions of Eq. (3.9) produce the eigenvalues of the tridiagonal symmetric u-Toeplitz matrix of size mu with perturbed corners. Additionally, from Eq. (3.9) it follows that all eigenvalues are symmetric around the center of the 1st Brillouin Zone.

From the eigenvalues and eigenvectors of Eq. (3.7), it can also easily be shown (details, e.g., in Ref. [174]) that the m-th power of the unimodular UCTM, $\boldsymbol{M}_u(E)$, i.e. the GTM, $\boldsymbol{M}_N(E)$, is given by [127]

$$(\boldsymbol{M}_u)^m = U_{m-1}(z)\boldsymbol{M}_u - U_{m-2}(z)\boldsymbol{I}_2, \tag{3.10}$$

where $U_m(z)$ are the Chebyshev polynomials of the second kind of degree m that satisfy the recurrence relation

$$U_m(z) - 2zU_{m-1}(z) + U_{m-2}(z) = 0, \forall m \geq 1, \tag{3.11}$$

with initial conditions $U_1(z) = 2z$, $U_0(z) = 1$ [93]. Hence, the GTM of the cyclic system can be written as

$$\boldsymbol{M}_N = \begin{pmatrix} U_{m-1}M_u^{11} - U_{m-2} & U_{m-1}M_u^{12} \\ U_{m-1}M_u^{21} & U_{m-1}M_u^{22} - U_{m-2} \end{pmatrix} \tag{3.12}$$

[4]Of course, the same relationships hold for the determination of the GTM of either periodic or aperiodic systems, if u is substituted by N.

or, alternatively, using Eq. (3.11), as

$$\boldsymbol{M}_N = \begin{pmatrix} U_m - U_{m-1}M_u^{22} & U_{m-1}M_u^{12} \\ U_{m-1}M_u^{21} & U_{m-1}M_u^{22} - U_{m-2} \end{pmatrix}. \tag{3.13}$$

Eq. (3.13) and the formula connecting the Chebyshev polynomials of the first and second kind, $2T_m(z) = U_m(z) - U_{m-2}(z)$ [93], lead to an alternate form from which the eigenvalues for cyclic boundaries can be found, i.e.

$$T_m(z) = 1. \tag{3.14}$$

3.1.2 Fixed Boundaries

For the system with fixed boundaries, i.e. when $\zeta = 0$, Eq. (2.33) takes the form

$$\boldsymbol{M}_N \begin{pmatrix} \psi_1 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} \psi_1 \\ 0 \end{pmatrix} \Rightarrow M_N^{11}\psi_1 = 0\psi_1. \tag{3.15}$$

Hence, ψ_1 is an element of the eigenvector of $\boldsymbol{H}(\zeta = 0)$ [cf. Eqs. (2.26), (2.28)] if the eigenvalues E of $\boldsymbol{H}(\zeta = 0)$ satisfy the condition $M_N^{11}(E) = 0$. Since both $M_N^{11} = 0$ and $\det(E\boldsymbol{I}_N - \boldsymbol{M}(0)) = 0$ are polynomial equations of E of the same degree (N) that are mutually satisfied [173], Eq. (3.15) produces the eigenvalues of the tridiagonal symmetric u-Toeplitz matrix of size mu. Exploiting Eqs. (3.12)-(3.13), the determination of these eigenvalues is reduced to the solution of

$$U_{m-1}(z)M_u^{11} = U_{m-2}(z), \tag{3.16}$$

or, alternatively,

$$U_m(z) = U_{m-1}(z)M_u^{22}. \tag{3.17}$$

3.1.3 Phonons

Closing this Section, we should notice that the above analysis holds also for the problem of phonon dispersion relations in periodic one-dimensional Lattices with nearest neighbor hoppings. The only thing that changes is the form of the transfer matrix of the sites, i.e. $\boldsymbol{P}_n$ in Eq. (3.3) takes the form

$$\boldsymbol{P}_n(\omega_{\text{ph}}) = \begin{pmatrix} 1 + \frac{K_{n-1}}{K_n} - \frac{m_n\omega_{\text{ph}}^2}{K_n} & \frac{K_{n-1}}{K_n} \\ 1 & 0 \end{pmatrix}. \tag{3.18}$$

ω_{ph} is the phonon frequency, m_n is the mass at site n, K_n (K_{n-1}) are the spring constants connecting sites n and $n+1$ $(n-1)$, and then ψ_n correspond to the oscillation amplitudes.

3.2 Dispersion relations and Eigenspectra

As shown in Sec. 3.1, the eigenvalues of any periodic WM, for either cyclic or fixed boundary conditions, can be determined by Eqs. (3.9) or (3.16)-(3.17), respectively, via which, the problem is reduced to the determination of the elements of UCTM using Eqs. (3.6). We have verified that the eigenvalues obtained by the above mentioned equations coincide with those obtained by the numerical diagonalization of matrix $H(\zeta = 0, 1)$. Below, we discuss the eigenvalues of such systems, as the simplest examples of the generic periodic system. Similar expressions to the ones presented below can be produced for larger values of u, using Eqs. (3.6).

3.2.1 Unit cell of one site ($u = 1, m = N$)

The UCTM is

$$\boldsymbol{M}_1(E) = \begin{pmatrix} \frac{E-\epsilon_1}{t_1} & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.19}$$

For *cyclic boundaries*, the dispersion relation of Eq. (3.9) takes the well known form

$$E = \epsilon_1 + 2t_1 \cos(ka) = \epsilon_1 + 2t_1 \cos\left(\frac{2\pi\nu}{N}\right). \tag{3.20}$$

$\nu = 1, 2, \ldots, N$. The dispersion relation is presented in Fig. 3.2, for a fixed on-site energy, $\epsilon_1 = 0$, where the effect of the hopping integral, t_1, alteration is also demonstrated. Increasing $|t_1|$ linearly increases the bandwidth. Although changing the sign of t_1 does not affect the eigenenergies' values, it moves the position of their extrema from the middle to the edge of the 1st Brillouin Zone.

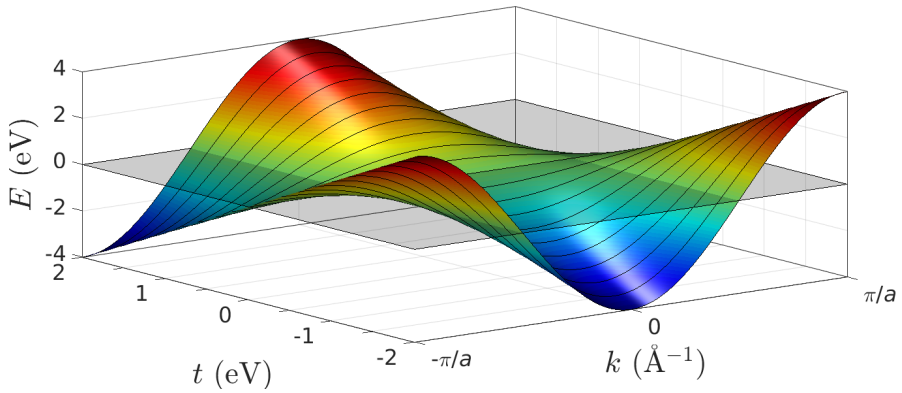

Figure 3.2: Band structure of the one-dimensional periodic lattice with *cyclic boundaries* [Eq. (3.20)], with a unit cell of one site ($u = 1$), for a fixed on-site energy $\epsilon_1 = 0$, as a function of the hopping integral t_1.

For *fixed boundaries*, it follows from Eq. (3.17) that the eigenvalues are produced by the zeros of $U_N(z)$, i.e. $z(E) = \cos\left(\frac{\mu\pi}{N+1}\right), \forall\mu = 1, 2, \ldots, N$ [93], hence

$$E = \epsilon_1 + 2t_1 \cos\left(\frac{\mu\pi}{N+1}\right). \tag{3.21}$$

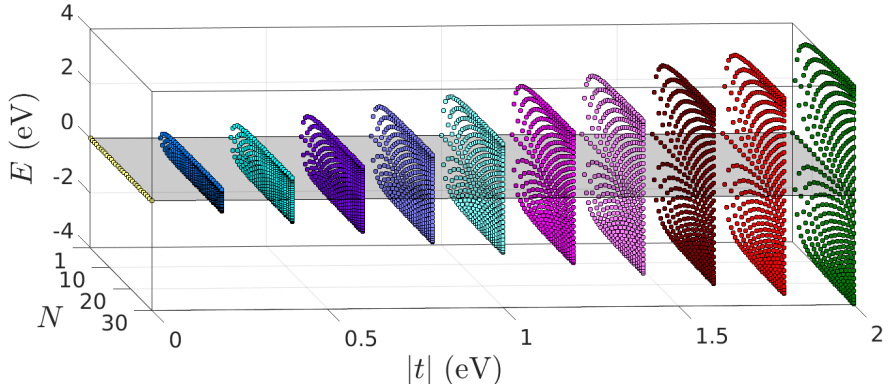

Figure 3.3: Eigenspectrum of the one-dimensional periodic lattice with *fixed boundaries* [Eq. (3.21)], with a unit cell of one site ($u = 1$), for N up to 30 sites, for a fixed on-site energy $\epsilon_1 = 0$, as a function of the absolute value of the hopping integral $|t_1|$.

The eigenspectrum for fixed boundaries as N increases, is presented in Fig. 3.3, for a fixed on-site energy, $\epsilon_1 = 0$, where the effect of the hopping integral absolute value $|t_1|$ alteration is also demonstrated. Eq. (3.21) is similar to Eq. (3.20) regarding the band center and bandwidth, and both equations tend to the same DOS as $N \to \infty$.

Eqs. (3.20) and (3.21) can be derived analytically by direct diagonalization of the Hamiltonian matrices cf. e.g. Ref. [56]. Here, we just demonstrated how they are derived with the transfer matrix method.

3.2.2 Unit cell of two sites ($u = 2, m = \frac{N}{2}$)

The UCTM is

$$\boldsymbol{M}_2(E) = \begin{pmatrix} \frac{(E-\epsilon_2)(E-\epsilon_1)}{t_2 t_1} - \frac{t_1}{t_2} & -\frac{E-\epsilon_2}{t_1} \\ \frac{E-\epsilon_1}{t_1} & -\frac{t_2}{t_1} \end{pmatrix}. \tag{3.22}$$

For *cyclic boundaries*, Eq. (3.9) takes the form

$$E = \frac{\epsilon_1 + \epsilon_2}{2} \pm \sqrt{\left(\frac{\epsilon_1 - \epsilon_2}{2}\right)^2 + t_1^2 + t_2^2 + 2t_1 t_2 \cos(k2a)}. \tag{3.23}$$

This dispersion relation is presented in Fig. 3.4, for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, where the effect of altering the ratio between the hopping integrals $\frac{t_1}{t_2}$, is also demonstrated. Again, changing the sign of $\frac{t_1}{t_2}$ moves the position of the eigenvalues' extrema from the middle to the edge of the 1st Brillouin Zone. As $\left|\frac{t_1}{t_2}\right|$ increases from 0 to 1 the widths of the two bands (which are equal to each other) increase and the bandgap decreases until it gets equal to the energy gap between the on-site energies (for $\left|\frac{t_1}{t_2}\right| = 1$). For $\left|\frac{t_1}{t_2}\right| > 1$ the bandwidths continue to increase, albeit much slower, and the bandgap starts to increase as well.

For *fixed boundaries*, it follows from Eq. (3.17) that the eigenvalues are the solutions

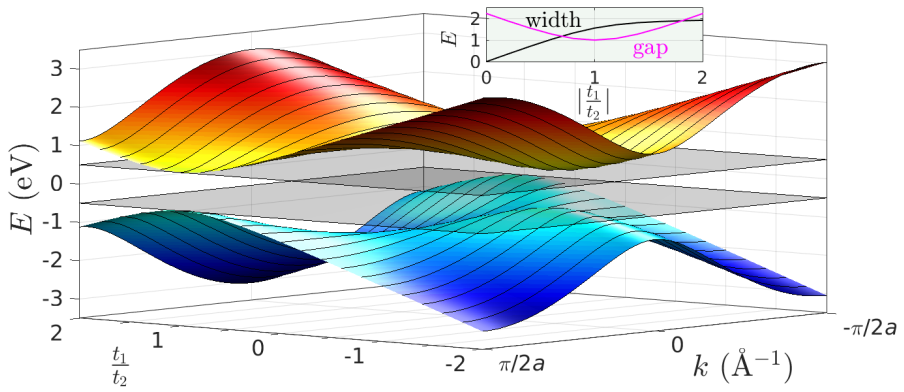

Figure 3.4: Band structure of the one-dimensional periodic lattice with a unit cell of two sites ($u = 2$) with *cyclic boundaries* [Eq. (3.23)], for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, as a function of the ratio between the hopping integrals $\frac{t_1}{t_2}$. (Inset) The bandwidths and the bandgap as a function of $|t_1/t_2|$.

of

$$U_{m-1}(z) = -\frac{t_1}{t_2} U_m(z). \tag{3.24}$$

The eigenspectrum for m up to 15 unit cells (N=30) is shown in Fig. 3.5, for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, where the effect of the absolute ratio between the hopping integrals $\left|\frac{t_1}{t_2}\right|$ alteration is also demonstrated. From Fig. 3.5 we notice that as the value of $\left|\frac{t_1}{t_2}\right|$ decreases down to 1, the bandgap decreases until it gets equal, for large N, to the energy gap between the on-site energies (for $\left|\frac{t_1}{t_2}\right| = 1$). Further decrease of $\left|\frac{t_1}{t_2}\right|$ below 1 leads to an increase in the gap between the main bands, but there is a single additional eigenvalue in the upper (lower) band that, for large N, reaches the value ϵ_1 (ϵ_2). This behavior does not occur for cyclic boundaries, hence these states inside the band gap arise solely from the boundary effect. As it will be shown below, this effect occurs generally for fixed boundaries, although it arises from more complex relations between the parameters of the system.

3.2.3 Unit cell of three sites ($u = 3$, $m = \frac{N}{3}$)

The UCTM is

$$\boldsymbol{M}_3(E) = \begin{pmatrix} \frac{(E-\epsilon_3)(E-\epsilon_2)(E-\epsilon_1)}{t_3 t_2 t_1} - \frac{(E-\epsilon_3)t_1}{t_2 t_3} - \frac{(E-\epsilon_1)t_2}{t_3 t_1} & -\frac{(E-\epsilon_3)(E-\epsilon_2)}{t_1 t_2} + \frac{t_2}{t_1} \\ \frac{(E-\epsilon_2)(E-\epsilon_1)}{t_1 t_2} - \frac{t_1}{t_2} & -\frac{(E-\epsilon_2)t_3}{t_1 t_2} \end{pmatrix}. \tag{3.25}$$

For *cyclic boundaries*, Eq. (3.9) takes the form

$$\cos(k3a) = \frac{1}{2}\prod_{i=1}^{3} \frac{(E-\epsilon_i)}{t_i} - \sum_{ijk}^{\circlearrowleft 3} \frac{(E-\epsilon_i)t_j}{2t_k t_i}, \tag{3.26}$$

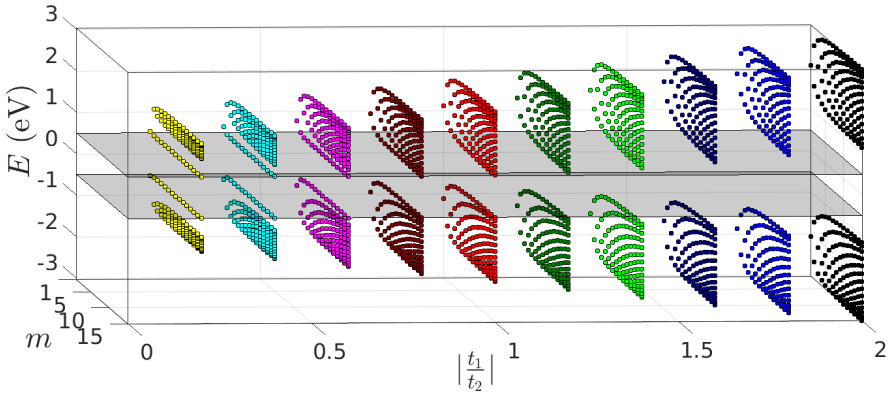

Figure 3.5: Eigenspectrum of the one-dimensional periodic lattice with a unit cell of two sites ($u = 2$) with *fixed boundaries* [Eq.(3.24)], for m up to 15 unit cells ($N = 30$), for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, as a function of the absolute ratio between the hopping integrals $|t_1/t_2|$.

where $\bigcirc^3$ denotes all cyclic permutations of the indices $i \neq j \neq k \in \{1, 2, 3\}$. This dispersion relation is presented in Fig 3.6 (left), for the example of a system with on-site energies $(\epsilon_1, \epsilon_2, \epsilon_3) = (6.3, 5.8, 6.1)$ eV and hopping integrals $(t_1, t_2, t_3) = (0.5, 0.6, 0.8)$ eV. The three bands that occur for this system have widths $\approx 0.51, 0.89, 0.38$ eV (from lower to higher energy) and the corresponding gaps are $\approx 0.22, 0.60$ eV.

For *fixed boundaries*, it follows from Eq. (3.17) that the eigenvalues are the solutions of

$$U_m(z) = -\frac{(E - \epsilon_2)t_3}{t_2 t_1} U_{m-1}(z). \tag{3.27}$$

The eigenspectrum for m up to 20 unit cells ($N = 60$) is presented in Fig. 3.6 (right), for the same system as in Fig. 3.6 (left). The bands that are shaped as N increases have widths $\approx 0.51, 0.88, 0.37$ eV, in agreement with the cyclic boundaries, but there are also three eigenvalues that exceed these bands, two of which lie well within the second gap. Generally, our calculations for $u = 3$ show that these mid-gap eigenvalues are three at most and their exact number and position depends on complex relations between the TB parameters. Such relations for 3-Toeplitz matrices of order $N = 3m+2$ can be found in Ref. [164], where the authors obtain in that case two such eigenvalues at most. In our case, where $N = 3m$, we find that, as N increases, these eigenvalues converge at most to three of the zeros of $M_3^{12} \times M_3^{21}$. Specifically, for the parameters of Fig. 3.6, the positions of the mid-gap eigenvalues, which are shown in Fig. 3.6 (right), converge quickly, as N increases, to $\frac{\epsilon_2+\epsilon_3}{2} \pm \sqrt{(\frac{\epsilon_2-\epsilon_3}{2})^2 + t_2^2}$, i.e. the zeros of M_3^{12}, and $\frac{\epsilon_1+\epsilon_2}{2} + \sqrt{(\frac{\epsilon_1-\epsilon_2}{2})^2 + t_1^2}$, i.e. the maximum of the zeros of M_3^{21}.

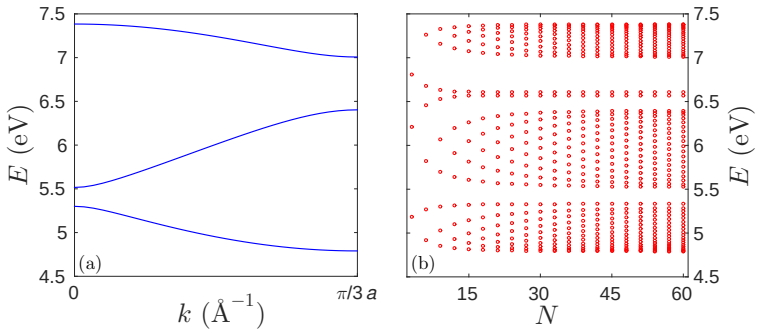

Figure 3.6: (left) Dispersion relation for *cyclic boundaries* and (right) eigenspectrum for *fixed boundaries* for m up to 20 unit cells ($N = 60$) of a one-dimensional periodic lattice with a unit cell of three sites, with on-site energies $(\epsilon_1, \epsilon_2, \epsilon_3) = (6.3, 5.8, 6.1)$ eV and hopping integrals $(t_1, t_2, t_3) = (0.5, 0.6, 0.8)$ eV, as determined by Eqs. (3.26) and (3.27), respectively.

3.2.4 Unit cell of four sites ($u = 4, m = \frac{N}{4}$)

The UCTM is

$$\boldsymbol{M}_4 = \begin{pmatrix} \frac{(E-\epsilon_4)(E-\epsilon_3)(E-\epsilon_2)(E-\epsilon_1)}{t_4t_3t_2t_1} - \frac{(E-\epsilon_4)(E-\epsilon_3)t_1}{t_2t_3t_4} & -\frac{(E-\epsilon_4)(E-\epsilon_3)(E-\epsilon_2)}{t_3t_2t_1} + \frac{(E-\epsilon_4)t_2}{t_3t_1} \\ -\frac{(E-\epsilon_4)(E-\epsilon_1)t_2}{t_1t_3t_4} - \frac{(E-\epsilon_2)(E-\epsilon_1)t_3}{t_1t_2t_4} + \frac{t_1t_3}{t_2t_4} & +\frac{(E-\epsilon_2)t_3}{t_2t_1} \\ \frac{(E-\epsilon_3)(E-\epsilon_2)(E-\epsilon_1)}{t_3t_2t_1} - \frac{(E-\epsilon_3)t_1}{t_3t_2} - \frac{(E-\epsilon_1)t_2}{t_1t_3} & -\frac{(E-\epsilon_3)(E-\epsilon_2)t_4}{t_3t_2t_1} + \frac{t_2t_4}{t_1t_3} \end{pmatrix}. \tag{3.28}$$

For *cyclic boundaries*, Eq. (3.9) takes the form

$$\cos(k4a) = \frac{1}{2}\prod_{i=1}^{4}\frac{(E-\epsilon_i)}{t_i} - \sum_{ijkl}^{\circlearrowleft 4}\frac{(E-\epsilon_i)(E-\epsilon_j)t_k}{2t_it_jt_l} + \frac{t_1t_3}{2t_2t_4} + \frac{t_2t_4}{2t_1t_3}, \tag{3.29}$$

where $\circlearrowleft^4$ denotes all cyclic permutations of the indices $i \neq j \neq k \neq l \in \{1, 2, 3, 4\}$. This dispersion relation is presented in Fig 3.7 (left), for a system with on-site energies $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (7.0, 9.0, 7.5, 8.5)$ eV and hopping integrals $(t_1, t_2, t_3, t_4) = (1.2, 0.9, 1.0, 0.8)$ eV. The four bands that occur for this system have widths $\approx 0.33, 0.61, 0.58, 0.30$ eV (from lower to higher energy) and the corresponding gaps are $\approx 0.42, 1.50, 0.53$ eV.

For *fixed boundaries*, it follows from Eq. (3.17) that the eigenvalues are solutions of

$$U_m(z) = \left(-\frac{(E-\epsilon_3)(E-\epsilon_2)t_4}{t_3t_2t_1} + \frac{t_2t_4}{t_1t_3}\right)U_{m-1}(z). \tag{3.30}$$

The eigenspectrum for m up to 20 unit cells ($N = 80$) is presented in Fig. 3.7 (right), for the same system as in Fig. 3.7 (left). The bands that are shaped as N increases

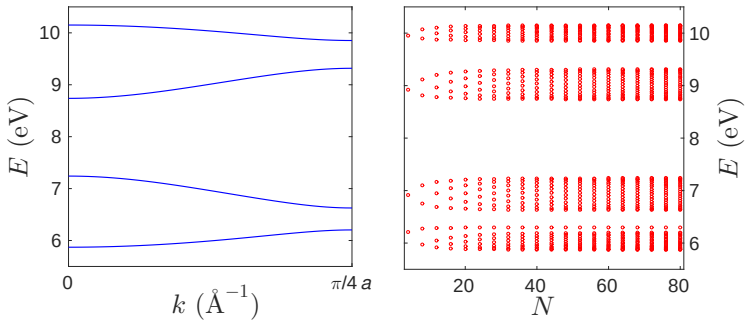

Figure 3.7: (left) Dispersion relation for *cyclic boundaries* and (right) eigenspectrum for *fixed boundaries* for m up to 20 unit cells ($N = 80$) of a one-dimensional periodic lattice with a unit cell of four sites, with on-site energies $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (7.0, 9.0, 7.5, 8.5)$ eV and hopping integrals $(t_1, t_2, t_3, t_4) = (1.2, 0.9, 1.0, 0.8)$ eV, as determined by Eqs. (3.29) and (3.30), respectively.

have widths $\approx 0.33, 0.61, 0.57, 0.30$ eV, in agreement with the cyclic boundaries, but there is also one eigenvalue that exceeds the first band. Generally, there are four mid-gap eigenvalues at most, the exact number and position of which depends on complex relations between the TB parameters. Again, we find that, as N increases, these eigenvalues converge at most to four of the zeros of $M_4^{12} \times M_4^{21}$. Specifically, for the parameters of Fig. 3.7, the position of the mid-gap eigenvalue, which is shown in Fig. 3.7 (right), converges quickly, as N increases, to the minimum zero of M_4^{21}.

3.3 Density of states

For a periodic WM with cyclic boundaries, the DOS can be obtained with the help of the UCTM as

$$g(E) = \frac{m}{\pi} \frac{\mathrm{d}}{\mathrm{d}E} |\mathrm{acos}(z)| = \frac{m}{\pi} \frac{\left|\frac{\mathrm{d}z}{\mathrm{d}E}\right|}{\sqrt{1-z^2}}, \tag{3.31}$$

where no spin degeneracies are taken into account. It is obvious that $2u$ vHS are expected at the points where $z(E) = \pm 1$, i.e. at the energies that lie at the center and edges of the 1st Brillouin Zone. Below, we provide analytical expressions of the DOS for systems with unit cell of 1, 2, 3, and 4 sites, as the simplest examples of the generic periodic system. For large N, the DOS for cyclic boundaries converges with the DOS for fixed boundaries.

3.3.1 Unit cell of one site ($u = 1, m = N$)

Eq. (3.31) takes the form [58]

$$g(E) = \frac{N}{\pi \sqrt{4t_1^2 - (E - \epsilon_1)^2}}. \tag{3.32}$$

Two vHS occur at $E = \epsilon_1 \pm 2|t_1|$. The DOS (over the total number of sites N) for $u = 1$ is presented in Fig. 3.8, for a fixed on-site energy, $\epsilon_1 = 0$, where the effect of alternating

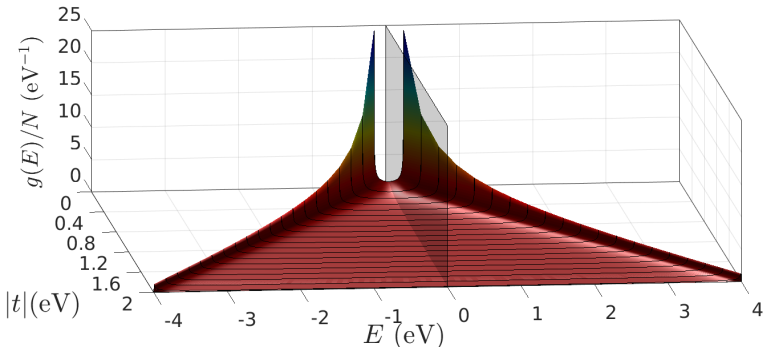

Figure 3.8: DOS (over N) of the one-dimensional periodic lattice with a unit cell of one site [Eq. (3.32)], for a fixed on-site energy $\epsilon_1 = 0$, as a function of the magnitude of the hopping integral $|t_1|$.

the magnitude of the hopping integral, $|t_1|$, is also demonstrated. Increasing the value of $|t_1|$ linearly increases the bandwidth and reduces the sharpness of the vHS that occur at the band edges.

3.3.2 Unit cell of two sites ($u = 2$, $m = \frac{N}{2}$)

Eq. (3.31) takes the form

$$g(E) = \frac{N}{2\pi} \frac{|2E - \epsilon_1 - \epsilon_2|}{\sqrt{4t_1^2 t_2^2 - [(E-\epsilon_1)(E-\epsilon_2) - t_1^2 - t_2^2]^2}}. \tag{3.33}$$

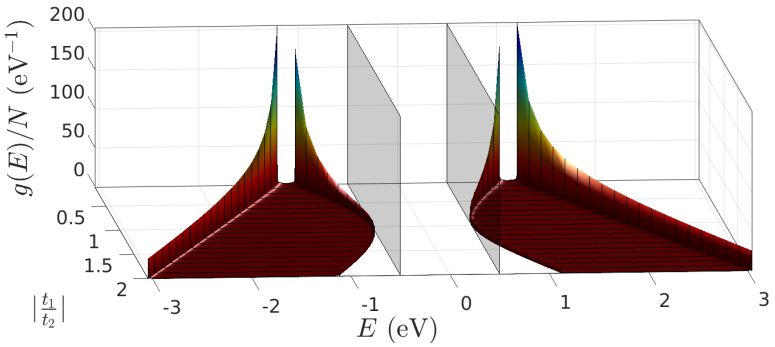

Figure 3.9: DOS (over N) of the one-dimensional periodic lattice with a unit cell of two sites [Eq. (3.33)], for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, as a function of the absolute ratio between the hopping integrals, $\left|\frac{t_1}{t_2}\right|$.

Four vHS occur at $E = \frac{\epsilon_1+\epsilon_2}{2} \pm \sqrt{\left(\frac{\epsilon_1-\epsilon_2}{2}\right)^2 + (t_1 \pm t_2)^2}$. The DOS (over the total number of sites N) for $u = 2$ is presented in Fig. 3.9, for fixed on-site energies $(\epsilon_1, \epsilon_2) = (-0.5, 0.5)$ eV, where the effect of altering the magnitude of the ratio between the hopping integrals, $\left|\frac{t_1}{t_2}\right|$, is also demonstrated. The DOS is symmetric around the mean of the on-site energies. Furthermore, the effect of altering $\left|\frac{t_1}{t_2}\right|$ on the bandwidths and

bandgap, discussed in Subsec. 3.2.2 (cf. inset of Fig. 3.4) is clearly illustrated. The vHS that are closer to the gap are less sharp. The sharpness of all vHS decreases dramatically as $\left|\frac{t_1}{t_2}\right|$ goes from 0 to 1, and slower for $\left|\frac{t_1}{t_2}\right| > 1$.

3.3.3 Unit cell of three sites ($u = 3$, $m = \frac{N}{3}$)

Eq. (3.31) takes the form

$$g(E) = \frac{N}{3\pi} \frac{\left| \sum_{ijk}^{\circlearrowleft 3} \frac{(E-\epsilon_i)(E-\epsilon_j)}{2t_i t_j t_k} - \frac{t_j}{2t_k t_i} \right|}{\sqrt{1 - \left[\frac{1}{2} \prod_{i=1}^{3} \frac{(E-\epsilon_i)}{t_i} - \sum_{ijk}^{\circlearrowleft 3} \frac{(E-\epsilon_i)t_j}{2t_k t_i} \right]^2}}. \tag{3.34}$$

In Fig. 3.10, we illustrate the DOS of a system with three sites per unit cell, for the parameters used in Fig. 3.6, as obtained by Eq. (3.34). As a careful comparison of Figs. 3.6 (right) and 3.10 suggests, the relative sharpness of the vHS as well as the position of each band's DOS minimum can be expected from the density of points in the eigenspectrum, as N increases. This is another demonstration that in the large N limit, the boundary effects play insignificant role in the electronic structure of the system, since the mid-gap eigenvalues that occur for fixed boundaries are negligible in number compared to those that lie within the bands.

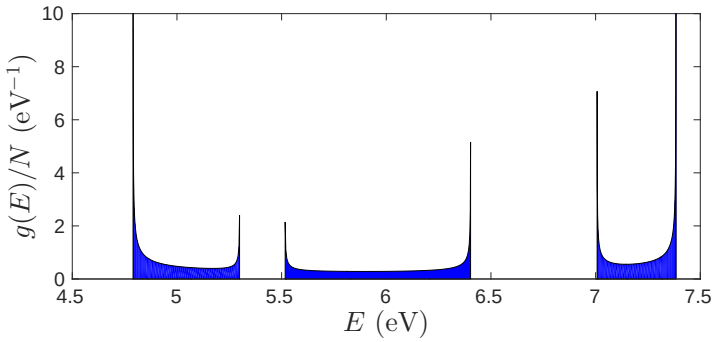

Figure 3.10: DOS (over N) of a one-dimensional periodic lattice with a unit cell of three sites [Eq. (3.34)], with on-site energies $(\epsilon_1, \epsilon_2, \epsilon_3) = (6.3, 5.8, 6.1)$ eV and hopping integrals $(t_1, t_2, t_3) = (0.5, 0.6, 0.8)$ eV.

3.3.4 Unit cell of four sites ($u = 4$, $m = \frac{N}{4}$)

Eq. (3.31) takes the form

$$g(E)=\frac{N}{4\pi}\frac{\left|\sum\limits_{ijkl}^{\circlearrowleft 4}\frac{(E-\epsilon_i)(E-\epsilon_j)(E-\epsilon_k)}{2t_it_jt_kt_l}-\frac{(2E-\epsilon_i-\epsilon_j)t_k}{2t_it_jt_l}\right|}{\sqrt{1-\left[\frac{1}{2}\prod\limits_{i=1}^{4}\frac{(E-\epsilon_i)}{t_i}-\sum\limits_{ijkl}^{\circlearrowleft 4}\frac{(E-\epsilon_i)(E-\epsilon_j)t_k}{2t_it_jt_l}+\frac{t_1t_3}{2t_2t_4}+\frac{t_2t_4}{2t_1t_3}\right]^2}}. \tag{3.35}$$

In Fig. 3.11, we illustrate the DOS of a system with four sites per unit cell, for the parameters used in Fig. 3.7, as obtained by Eq. (3.35).

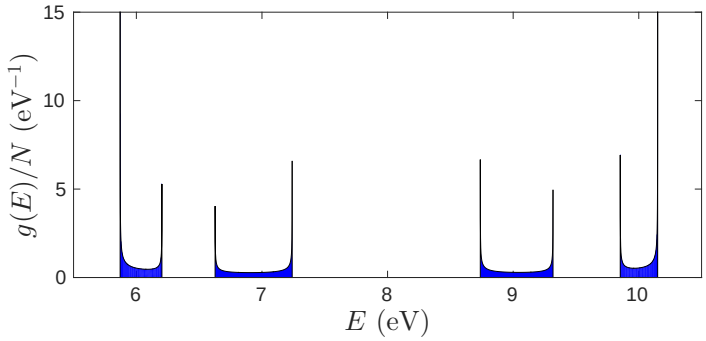

Figure 3.11: DOS (over N) of a one-dimensional periodic lattice with a unit cell of four sites [Eq. (3.35)], with on-site energies $(\epsilon_1,\epsilon_2,\epsilon_3,\epsilon_4)=(7.0,9.0,7.5,8.5)$ eV and hopping integrals $(t_1,t_2,t_3,t_4)=(1.2,0.9,1.0,0.8)$ eV.

3.4 Transmission coefficient

The TC describes the probability that a carrier, incident to a quantum system, transmits through its eigenstates. To obtain the TC of a one-dimensional TB lattice, we connect the ends of the system under examination to semi-infinite homogeneous leads, which play the role of a carrier bath. Hence, sites $[-\infty,0]\cup[N+1,\infty]$ belong to the left and right lead, respectively. Within sites $[1,N]$ lies *the periodic lattice under examination*, which is considered as a *perturbation in the homogeneous infinite lead*. The coupling between the ends of the periodic wire and the left and right lead is described by the effective parameters t_L and t_R, respectively. Here, we generally choose them to be different, as a reflection of the difference in coupling of the same material (lead) with different moieties (the end sites of the system). The effect of this asymmetry in coupling will be discussed in detail below. The leads' band lies in the energy interval $[\epsilon_m-2|t_m|,\epsilon_m+2|t_m|]$, where ϵ_m is the on-site energy of the leads and t_m the hopping integral between the leads' sites, since their dispersion relation at zero bias is identical to Eq. (3.20). Hence, the energy center and bandwidth of the leads are ϵ_m and $4|t_m|$, respectively. If we imagine the lead as a homogeneous system with one electron per

site, then the band is half-filled (hence the leads are metallic), and ϵ_m is the Fermi energy of the metal. The GTM of the whole lead-system-lead complex is $\boldsymbol{M}_R\tilde{\boldsymbol{M}}_N\boldsymbol{M}_L$, where $\boldsymbol{M}_{L/R}$ represents the sites of the ideal, *cyclically bounded left/right lead*, and $\tilde{\boldsymbol{M}}_N = \boldsymbol{P}_R\boldsymbol{M}_N\boldsymbol{P}_L$, where

$$\boldsymbol{P}_R = \begin{pmatrix} \frac{t_u}{t_R} & 0 \\ 0 & \frac{t_R}{t_m} \end{pmatrix}, \quad \boldsymbol{P}_L = \begin{pmatrix} \frac{t_m}{t_L} & 0 \\ 0 & \frac{t_L}{t_u} \end{pmatrix} \tag{3.36}$$

are the matrices that describe the coupling of the three subsystems. The waves at the left and the right lead can be expanded as

$$\psi_{\{n\}<1} = e^{iq_Lna} + re^{-iq_Lna}, \quad \psi_{\{n\}>N} = te^{iq_Rna}, \tag{3.37}$$

where $q_{L/R}$ is the wavenumber. We have assumed, without any loss of generality, that the incident waves come from the left and we have normalized their amplitude. Hence, the TC at zero bias (where $q_R = q_L := q$) is defined as $T(E) = |t|^2$. The GTM of the perturbation's region obeys to the equation

$$\begin{pmatrix} \psi_{N+1} \\ \psi_N \end{pmatrix} = \tilde{\boldsymbol{M}}_N \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}. \tag{3.38}$$

Substituting Eq. (3.37) to Eq. (3.38) yields

$$\begin{pmatrix} t \\ 0 \end{pmatrix} = \tilde{\mathfrak{M}}_N \begin{pmatrix} 1 \\ r \end{pmatrix}, \tag{3.39}$$

where

$$\tilde{\mathfrak{M}}_N = \begin{pmatrix} e^{iqNa} & 0 \\ 0 & e^{-iqNa} \end{pmatrix}^{-1} \begin{pmatrix} e^{iqa} & e^{-iqa} \\ 1 & 1 \end{pmatrix}^{-1} \tilde{\boldsymbol{M}}_N \begin{pmatrix} e^{iqa} & e^{-iqa} \\ 1 & 1 \end{pmatrix}. \tag{3.40}$$

Hence,

$$t = \frac{\det\left(\tilde{\mathfrak{M}}_N\right)}{\tilde{\mathfrak{M}}_N^{22}}, \tag{3.41}$$

After some manipulations, we find that the TC at zero bias is given by

$$T(E) = \frac{4\sin^2(qa)}{\mathrm{Tr}\left(\tilde{\boldsymbol{M}}_N\right)^2\sin^2(qa) + \left[\tilde{M}_N^{12} - \tilde{M}_N^{21} + (\tilde{M}_N^{11} - \tilde{M}_N^{22})\cos(qa)\right]^2}. \tag{3.42}$$

Substituting the elements of $\tilde{\boldsymbol{M}}_N$, and taking into account the unimodularity of $\boldsymbol{M}_N$, leads to the form

$$T(E) = \frac{1}{1 + \Lambda(E)}, \tag{3.43}$$

where

$$\Lambda(E) = \frac{\left[W_N(E) + X_N^+(E)\cos(qa)\right]^2}{4\sin^2(qa)} + \frac{X_N^-(E)^2}{4}, \tag{3.44}$$

and

$$W_N(E) = M_N^{11}\omega - M_N^{22}\omega^{-1}, \tag{3.45}$$

$$X_N^{\pm}(E) = M_N^{12}\chi \pm M_N^{21}\chi^{-1}. \tag{3.46}$$

The term

$$\omega = \frac{t_m\, t_N}{t_R\, t_L}, \tag{3.47}$$

which is included only in $W_N(E)$, expresses the coupling **strength** (specifically, the less the value of ω, the stronger the coupling), in means of the deviation of the real coupling of the system to the leads from the *ideal coupling* [in which the product of the coupling hopping integrals (t_L, t_R) becomes equal to the product of the hopping integrals of the isolated lead $(t_m = t_L)$ and the isolated system $(t_N = t_R)$ as if each was cyclically bounded to itself]. The term

$$\chi = \frac{t_L}{t_R}, \tag{3.48}$$

which is included only in $X_N^{\pm}(E)$, expresses the coupling **asymmetry**, i.e. the difference in coupling strength between the leads and the left/right end of the system. The maxima/minima of TC are the minima/maxima of $\Lambda(E)$, which is a polynomial of energy of degree N.

For a periodic system, taking into account Eq. (3.13), it holds that

$$W_N(E) = U_m(z)\omega - U_{m-1}(z)M_u^{22}(\omega + \omega^{-1}) + U_{m-2}(z)\omega^{-1}, \tag{3.49}$$

$$X_N^{\pm}(E) = U_{m-1}(z)X_u^{\pm}(E), \tag{3.50}$$

$$X_u^{\pm}(E) = (M_u^{12}\chi \pm M_u^{21}\chi^{-1}), \tag{3.51}$$

and t_N in Eq. (3.47) replaced by t_u.

In the following Subsections, we discuss in detail the effect of the coupling strength and asymmetry, as well as the lead properties, to the transmission profile of any periodic system. In all cases considered below, we will assume that the bandwidth of the leads, as determined by $|t_m|$, is such that it contains all the eigenstates of the system, in order to gain the full picture of the transmission profiles.

3.4.1 Ideal coupling

For ideal coupling between the leads and the system, i.e. when $|\omega| = |\omega^{-1}| = 1$, the recurrent formula Eq. (3.11) results in $W_N = \mathrm{sgn}(\omega)U_{m-1}(z)(M_u^{11} - M_u^{22})$. Hence,

$$\Lambda(E) = \left\{ \frac{\left[\mathrm{sgn}(\omega)(M_u^{11} - M_u^{22}) + X_u^{+}(E)\cos(qa)\right]^2}{4\sin^2(qa)} + \frac{X_u^{-}(E)^2}{4} \right\} U_{m-1}(z)^2. \tag{3.52}$$

From Eqs. (3.52) and (3.43) it is obvious that if $U_{m-1}(z) = 0$, then $T(E) = 1$. Hence, for ideal coupling, there are $(m-1)u = N - u$ energies in which transmission becomes full. These energies lie in the zeros of $U_{m-1}(z)$, i.e., they are the solutions of

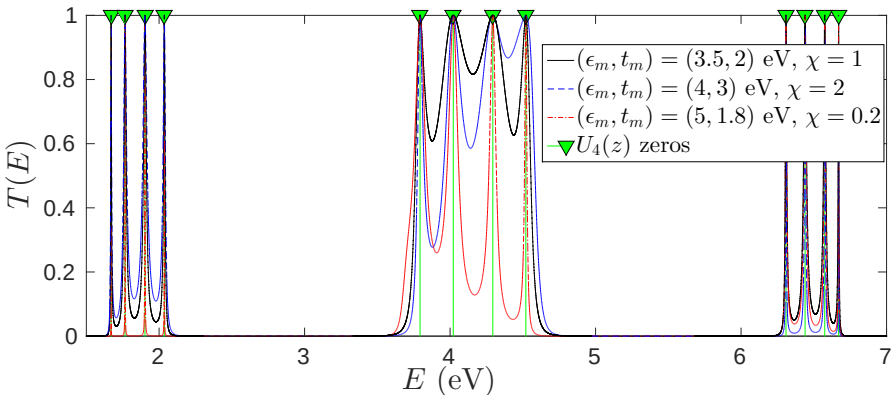

Figure 3.12: Transmission coefficient of a periodic system with $u = 3$ and $m = 5$ ($N = 15$), for the ideal coupling condition $\omega = 1$. The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3) = (3, 4, 5.5)$ eV and $(t_1, t_2, t_3) = (1, 0.8, 1.5)$ eV. The energies that correspond to full transmission do not depend on the energy center and bandwidth of the leads or the coupling asymmetry; they depend solely on the energy structure of the system, their number is $(m-1)u = 12$, and they are given by the zeros of $U_4(z)$.

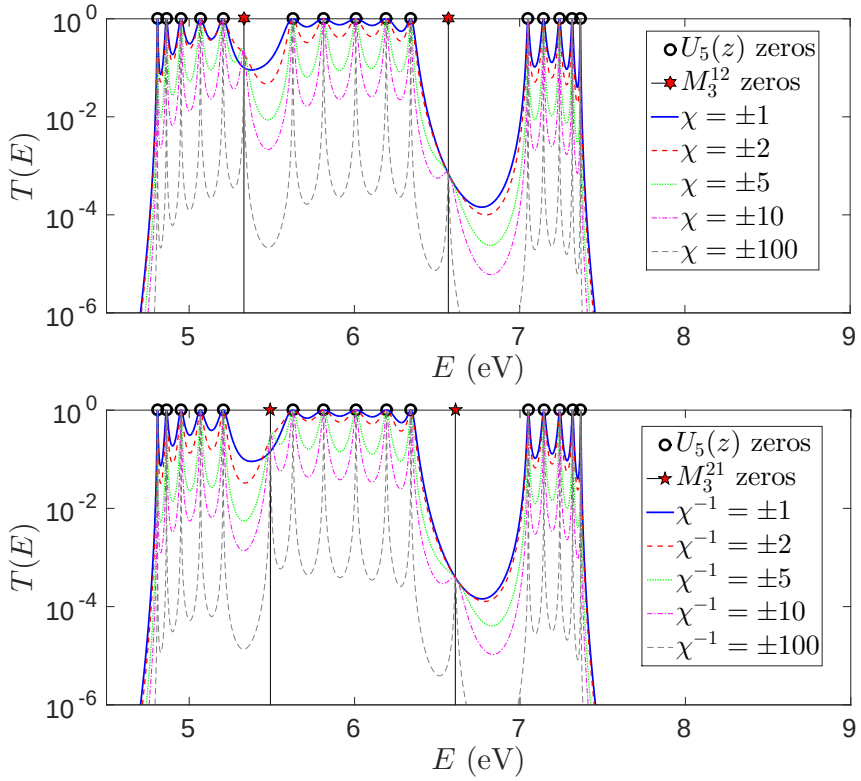

Figure 3.13: Dependence of the transmission coefficient on the coupling asymmetry χ for a periodic system with $u = 3$ sites per unit cell and $m = 6$ unit cells ($N = 18$), for the ideal coupling condition $|\omega| = 1$. The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3) = (6.3, 5.8, 6.1)$ eV and $(t_1, t_2, t_3) = (0.5, 0.6, 0.8)$ eV. The on-site energy and hopping integral of the leads are chosen $E_m = 6$ eV, $t_m = 2$ eV. [Top] $T(E)$ for $|\chi| > 1$. [Bottom] $T(E)$ for $|\chi| < 1$. The transmission becomes full in $(m-1)u = 15$ energies, which are given by the zeros of $U_5(z)$. As $|\chi|$ increases (decreases) above (below) 1, $u - 1$ secondary peaks occur, the energy of which is related to the zeros of M_u^{12} (M_u^{21}).

$z(E) = \cos\left(\frac{\mu\pi}{m}\right)$, $\mu = 1, 2, \ldots, m-1$. Thus, their position depends solely on the energy structure of the system and not on the energy center/bandwidth of the leads or the coupling asymmetry. This is depicted in Fig. 3.12, for a periodic system with $u = 3$ sites per unit cell and $m = 5$ unit cells ($N = 15$). Finally, we should mention that this full transmission condition that occurs for ideal coupling is solely a result of the periodicity of the system, i.e. of the fact that the elements of $\boldsymbol{M}_N$ are related to the Chebyshev polynomials of the second kind. The effect of the energy center and bandwidth of the leads will be separately addressed below.

The role of coupling asymmetry

The role of coupling asymmetry is depicted on Fig. 3.13, for the system of Fig. 3.6 with $m = 6$ unit cells ($N = 18$). We notice that increasing (decreasing) $|\chi|$ above (below) 1 leads to a significant decrease in the lower envelope of $T(E)$, i.e., the "curve" shaped by the local minima of $T(E)$, and a sharpening of the transmission peaks, which reflects the enhancement of backscattering effects. Furthermore, such an increase (decrease) of $|\chi|$ above (below) 1 leads to the appearance of $u-1$ secondary peaks, the positions of which are related to the zeros of M_u^{12} (M_u^{21}). These peaks occur because this increase (decrease) of $|\chi|$ leads to the domination of the terms $X^{\pm}(E)$ inside the curly brackets of Eq. (3.52).

The role of the leads

Let us take as an example the system of Fig. 3.7 with $m = 30$ unit cells ($N = 120$) and suppose symmetric coupling, $|\chi| = 1$. For a fixed bandwidth of the leads, $4|t_m|$, we observe that changing the energy center of the leads, ϵ_m, has a significant effect on the lower envelopes of $T(E)$ (Fig. 3.14 [Top]). Generally, the increase of ϵ_m leads to a shift of the maxima of the lower envelopes to smaller energies. In Ref. [181], where the TC of a periodic single-stranded DNA segment with $u = 2$ is studied, the authors state that the position of the maxima of the lower envelopes increases with ϵ_m (cf. Fig. 3 of Ref. [181]), i.e., the reverse behavior. With our treatment, we found that this holds only if $\forall t_m \gtrless 0$, $\cos(qa) > 0 \iff E - \epsilon_m \lessgtr 0$, i.e., if the dispersion relation of the leads is inconsistently incorporated into the calculations. Taking the dispersion relation in such a manner, we were able to reproduce Fig. 3 of Ref. [181], and noticed that the transmission profiles are generally alternated, giving rise to the behavior the authors mention. If we fix ϵ_m, we find that the increase of the leads' bandwidth, $4|t_m|$, has significantly smaller effect on the transmission profiles, compared to the increase of ϵ_m (Fig. 3.14 [Bottom]).

3.4.2 Strong or weak coupling

From Eq. (3.49) it follows that in the very strong coupling regime ($|\omega^{-1}| \gg 1$)

$$W_N(E) \simeq \omega^{-1}(U_{m-2}(z) - U_{m-1}(z)M_u^{22}), \tag{3.53}$$

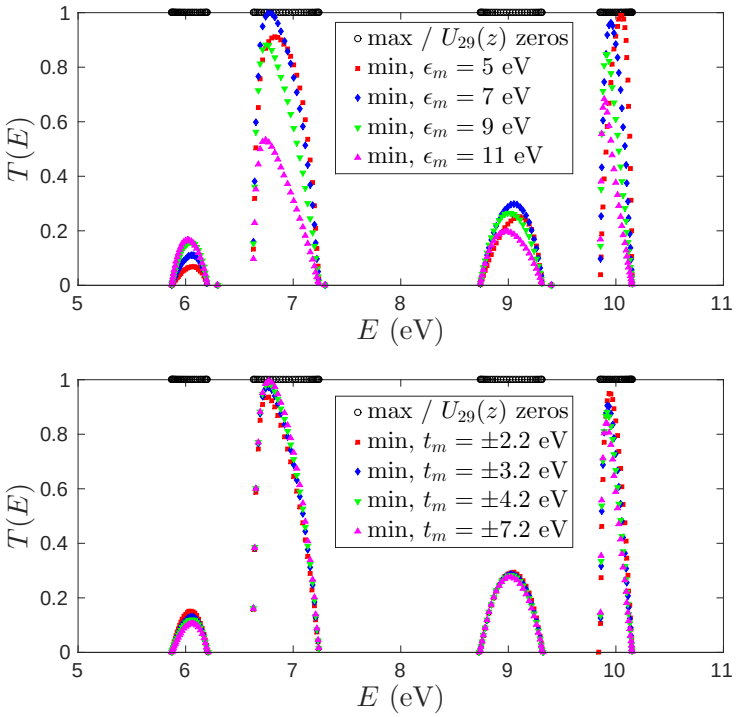

Figure 3.14: Dependence of the transmission coefficient on the leads properties for a periodic system with $u = 4$ sites per unit cell and $m = 30$ unit cells ($N = 120$), for the ideal coupling condition and symmetric coupling $|\omega| = |\chi| = 1$. The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (7, 9, 7.5, 8.5)$ eV and $(t_1, t_2, t_3, t_4) = (1.2, 0.9, 1, 0.8)$ eV. [Top] $T(E)$ for fixed bandwidth $4|t_m|$ ($t_m = 3$ eV) and varying energy center ϵ_m. [Bottom] $T(E)$ for fixed energy center ($\epsilon_m = 8$ eV) and varying t_m. The transmission becomes full in $(m-1)u = 116$ energies, which are given by the zeros of $U_{29}(z)$.

while, in the very weak coupling regime ($|\omega| \gg 1$)

$$W_N(E) \simeq \omega(U_m(z) - U_{m-1}(z)M_u^{22}). \tag{3.54}$$

If the coupling is strong or weak and symmetric, the term $W_N(E)$ becomes dominant in Eq. (3.44). Hence, for very strong or weak symmetric coupling, the transmission peaks occur in the region of the zeros of $W_N(E)$, as given by Eq. (3.53) or (3.54), hence, $N-2$ or N peaks are expected, respectively. A depiction of the effect of the decrease (increase) of coupling strength factor ω below (above) the ideal coupling condition, for symmetric coupling, is presented in Fig. 3.15 (3.16), for a system with $u = 6$ and $m = 4$ ($N = 24$). The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5, \epsilon_6) = (10, 8.3, 9.7, 8.8, 9.1, 8)$ eV and $(t_1, t_2, t_3, t_4, t_5, t_6) = (0.9, 1, 0.8, 0.6, 1.1, 0.7)$ eV. The leads' parameters are $\epsilon_m = 9$ eV, $t_m = 2$ eV.

The role of coupling asymmetry

If the coupling is strong (weak) and asymmetric, three cases can be distinguished:

(a) If the coupling asymmetry is significantly smaller than the coupling strength or weakness [$\max(|\chi|, |\chi^{-1}|) \ll |\omega^{-1}|$ for strong coupling or $\max(|\chi|, |\chi^{-1}|) \ll$

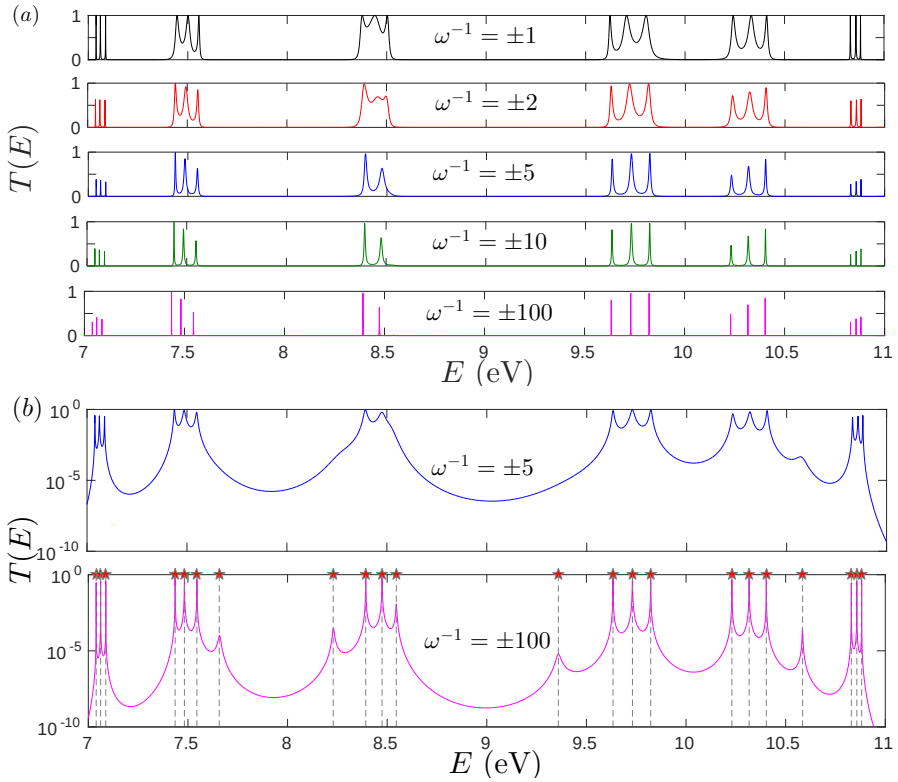

Figure 3.15: (a) Dependence of the transmission coefficient $T(E)$ on the coupling strength factor ω, in the strong coupling regime ($|\omega^{-1}| > 1$), for symmetric coupling ($|\chi| = 1$), for a periodic system with $u = 6$ and $m = 4$ ($N = 24$). (b) Transmission coefficients for strong ($|\omega| = \frac{1}{5}$) and very strong ($|\omega| = \frac{1}{100}$) coupling in logarithmic scale. As the coupling strength $|\omega^{-1}|$ increases above the ideal coupling condition, $N - 2$ peaks arise. For very strong coupling, the peaks' positions are determined by the zeros of Eq. (3.53), depicted in dashed lines, in the bottom panel of (b).

$|\omega|$ for weak coupling], $W_N(E)$ continues to be dominant in Eq. (3.44), hence, the transmission peaks occur again in the region of the zeros of $W_N(E)$ as given by Eq. (3.53) for strong coupling or by Eq. (3.54) for weak coupling.

(b) If the coupling asymmetry is of comparable magnitude with the coupling strength or weakness [$\max(|\chi|, |\chi^{-1}|) \approx |\omega^{-1}|$ for strong coupling or $\max(|\chi|, |\chi^{-1}|) \approx |\omega|$ for weak coupling], the peaks position cannot be determined without the full solution of Eq. (3.44).

(c) If the coupling asymmetry is significantly larger than the coupling strength or weakness [$\max(|\chi|, |\chi^{-1}|) \gg |\omega^{-1}|$ for strong coupling or $\max(|\chi|, |\chi^{-1}|) \gg |\omega|$ for weak coupling], the terms $X^{\pm}(E)$ become dominant in Eq. (3.44), hence the transmission peaks occur in the region of the zeros of $U_{m-1}(z)M_u^{12}$ for $|\chi| > 1$ or of $U_{m-1}(z)M_u^{21}$ for $|\chi| < 1$. The peaks number in this case is $N - 1$.

This behavior is depicted in Fig. 3.17, in the case of weak and asymmetric coupling, for the same system as in Figs. 3.15 and 3.16.

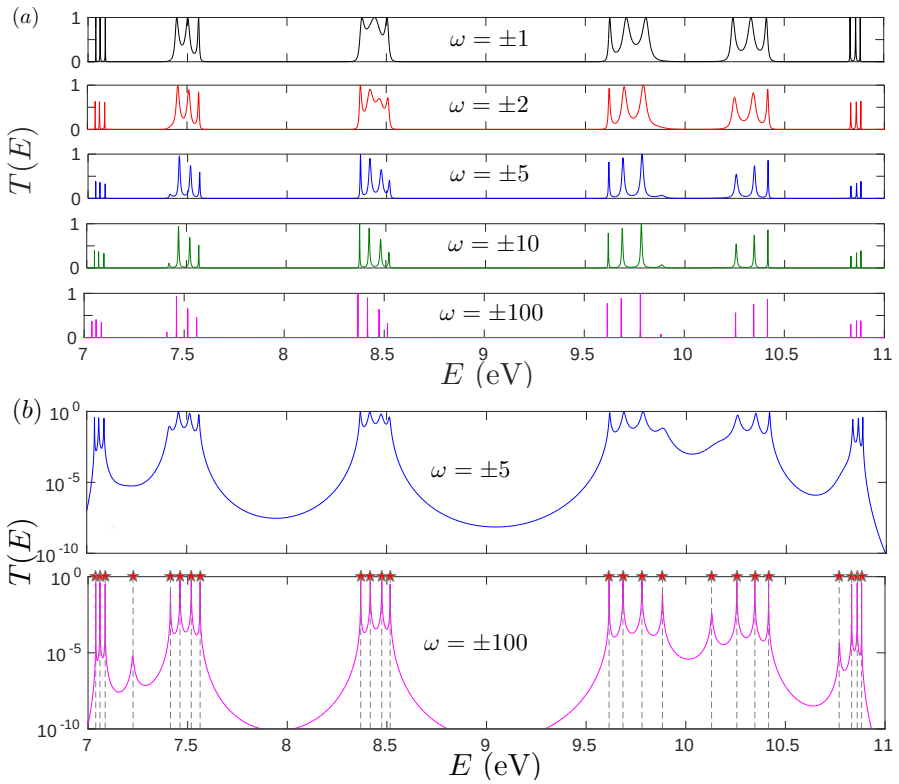

Figure 3.16: (a) Dependence of the transmission coefficient $T(E)$ on the coupling strength factor ω, in the weak coupling regime ($|\omega| > 1$), for symmetric coupling ($|\chi| = 1$), for a periodic system with $u = 6$ and $m = 4$ ($N = 24$). (b) Transmission coefficients for weak ($|\omega| = 5$) and very weak ($|\omega| = 100$) coupling in logarithmic scale. As the coupling weakness $|\omega|$ increases above the ideal coupling condition, N peaks arise. For very weak coupling, the peaks' positions are determined by the zeros of Eq. (3.54), depicted in dashed lines, in the bottom panel of (b).

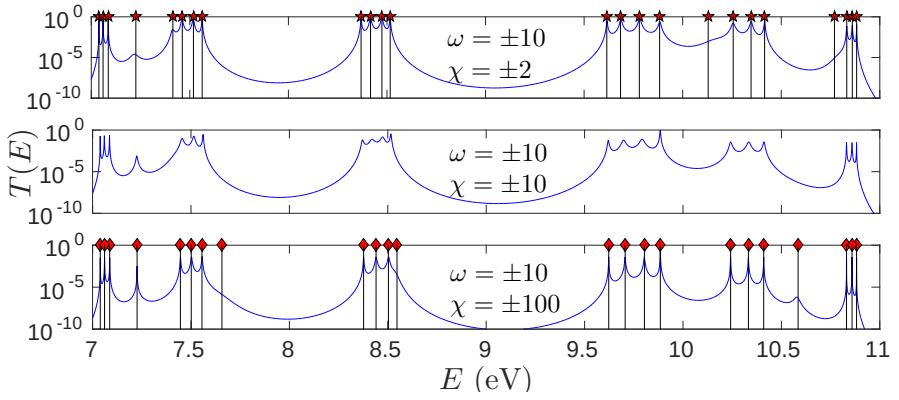

Figure 3.17: Dependence of the transmission coefficient $T(E)$ on the coupling asymmetry $|\chi|$, in the weak coupling regime ($|\omega| = 10$), for $|\chi| = 2$ (top), 10 (middle), 100 (bottom), for a periodic system with $u = 6$ and $m = 4$ ($N = 24$). For $|\chi| \ll |\omega|$, the transmission peaks occur in the region of the N zeros of Eq. (3.54), which are also depicted in the top panel. For $|\chi| \gg |\omega|$, the transmission peaks occur in the region of the $N - 1$ zeros of $U_{m-1}(z)M_u^{12}$, which are also depicted in the bottom panel. The case in which the asymmetry of coupling is comparable to is weakness, is an intermediate regime.

The role of the leads

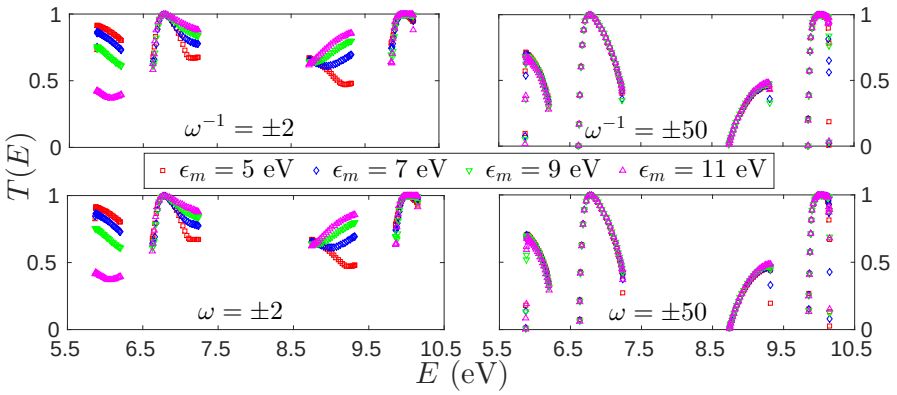

Figure 3.18: Dependence of the upper envelope of the transmission coefficient on the energy center of the leads, ϵ_m, for a system $u = 4$ sites per unit cell and $m = 30$ unit cells ($N = 120$), for symmetric coupling $|\chi| = 1$, in the strong (top left), very strong (top right), weak (bottom left) and very weak (bottom right) coupling regimes. The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (7, 9, 7.5, 8.5)$ eV and $(t_1, t_2, t_3, t_4) = (1.2, 0.9, 1, 0.8)$ eV. As the strength/weakness of coupling increases, the transmission profile becomes less dependent on the increase of ϵ_m.

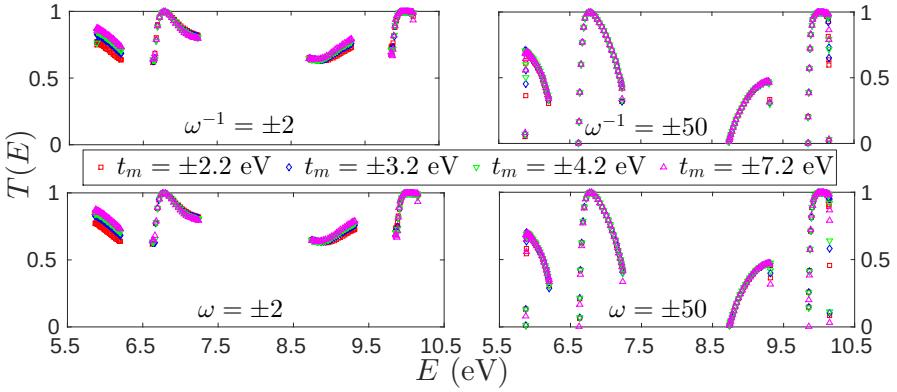

Figure 3.19: Dependence of the upper envelope of the transmission coefficient on the bandwidth of the leads, as determined by $|t_m|$, for a system $u = 4$ sites per unit cell and $m = 30$ unit cells ($N = 120$), for symmetric coupling $|\chi| = 1$, in the strong (top left), very strong (top right), weak (bottom left) and very weak (bottom right) coupling regimes. The TB parameters of the system are $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (7, 9, 7.5, 8.5)$ eV and $(t_1, t_2, t_3, t_4) = (1.2, 0.9, 1, 0.8)$ eV. As the strength/weakness of coupling increases, the transmission profile becomes less dependent on the increase of $|t_m|$. Generally, increasing $|t_m|$ has less effect on the transmission profiles than increasing ϵ_m.

As it is obvious from the previous discussion, the term that incorporates the properties of the leads, $\cos(qa)$, is significant in Eq. (3.44) only if the coupling asymmetry is of comparable magnitude with the coupling strength or weakness. In this case, the effect of the alteration of the leads' energy center and bandwidth to the transmission profile is not negligible. Both the upper and the lower envelope are affected. The lower envelope is affected in the same manner as for the ideal coupling condition. The upper envelope is also more affected by altering of ϵ_m that it is by altering $|t_m|$. If the coupling

asymmetry is much larger than its strength or weakness, the transmission profiles remain almost unchanged. The above remarks are summarized in Figs. 3.18 and 3.19, where the effect of increasing ϵ_m and $|t_m|$, respectively, to the upper envelopes of the TC in the strong and weak coupling regime, is displayed, for the system of Fig. 3.7 with $m = 30$ unit cells ($N = 120$).

3.4.3 Optimal Coupling Condition

From the above discussion, we conclude that in the strong and weak coupling regimes, the transmission peaks sharpen and the TC does not generally reach the full transmission condition $T(E) = 1$. Hence, *stronger coupling does not generally lead to larger transmission and the ideal coupling condition* $|\omega| = 1$*, together with the symmetric coupling condition, provides the optimal coupling condition.* This remark generalizes, for any periodic one-dimensional lattice, and provides with a physical meaning, the optimal coupling condition reported in Ref. [180], for a single stranded DNA segment with four sites per unit cell, in which all the hopping integrals of the DNA system were assumed to be equal. In Ref. [180], the authors analytically derive the TC for this system, impose the condition $U_{m-1}(z) = 0$ and apply extreme conditions to $T(E_p)$, where p is indexing the energies in which the above condition holds. This treatment allows for the determination of the TC value in the specific energies, E_p, and the authors find that this value is 1 for the condition $|\omega| = 1$ (in our nomenclature). As we have also shown, this is true. For $|\omega| = 1$, when $U_{m-1}(z) = 0$, then $T(E_p) = 1$. But, as we also have shown in the above discussion, the condition $U_{m-1}(z) = 0$ does not generally correspond to the transmission peaks, i.e. E_p are not always the positions of the maxima. Hence, in fact, the treatment of Ref. [180] rather examines the transmission values in the specific energies E_p and not the general occurrence of the transmission peaks. The condition $|\omega| = 1$ is optimal, not because in any other case $T(E_p) \leq 1$, since E_p are not generally the positions of peaks, but because in any other condition the transmission peaks are not generally confined to reach the value 1, as shown above. For $|\omega| = 1$ this value is reached, because it is only then that $\Lambda(E)$ is exactly proportional to the Chebyshev polynomials of the second kind, as shown in Eq. (3.52). This discussion demonstrates that our treatment of the coupling condition in terms of the transmission peaks' position, rather than the study of the specific energies for which $U_{m-1}(z) = 0$, addresses the effect properly.

Analogous conclusions regarding the coupling conditions can be obtained for more complex TB models [82]. In Ref. [78], a poly(G)-poly(C) DNA 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 Fig. 7 of Ref. [78] 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. [182], the authors study a poly(G)-poly(C) DNA chain within an extension of the FLM, which allows hopping between backbone sites as well as all possible di-

agonal 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 Fig. 2 of Ref. [182]). This is also evident by comparing the averaged TC, which is defined as

$$T_a(E) = \frac{\int_{E_{\text{min}}}^{E} T(e)de}{E - E_{\text{min}}}, \tag{3.55}$$

cf. the panels in third row in Fig. 2 of Ref. [182]. 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.

3.4.4 The intrinsic hoppings of the system

The previous discussion for the effect of the coupling strength ω and the coupling asymmetry χ on the transmission profiles is specified in this Subsection for the two simplest cases of periodic one-dimensional TB lattices ($u = 1, 2$). The effect of the intrinsic hoppings of the system is also demonstrated.

In Fig. 3.20, we present the TC of a system with $u = 1$ site per unit cell and $m = N = 10$ unit cells, for the ideal, strong and weak coupling conditions, with or without coupling asymmetry, as the magnitude of the hopping integral of the system, $|t_1|$, increases up to $|t_m|$. We have chosen $\epsilon_1 = \epsilon_m$. The number of peaks is always 9 for ideal coupling, 8 for strong coupling and 10 for weak coupling, as expected by the above discussion. It is also evident that the ideal and symmetric coupling conditions lead to the most efficient transmission, although in this simple case, the full transmission condition $T(E) = 1$ is reached in all cases of symmetric coupling. In the upper left panel of Fig. 3.20 it can be seen that when $|t_1| = |t_m|$ the structures of the leads and the system become identical, hence the system is totally transparent.

In Fig. 3.21, we present the TC of a system with $u = 2$ sites per unit cell and $m = 5$ unit cells ($N = 10$), for the ideal, strong and weak coupling conditions, with or without coupling asymmetry, as the magnitude of the ratio between the hopping integrals of the system, $\left|\frac{t_1}{t_2}\right|$, increases. We have chosen $\epsilon_m = \frac{\epsilon_1+\epsilon_2}{2}$ and the bandwidth of the leads, as determined by $|t_m|$, is chosen so as to contain all the eigenstates of the system. Again, it is evident that the ideal and symmetric coupling condition leads to the most efficient transmission. For ideal and asymmetric coupling, except for the peaks of magnitude 1 that occur in the zeros of $U_4(z)$, there is one additional peak near $E = \epsilon_1$ ($E = \epsilon_2$), when $\chi > 1$ ($\chi < 1$), i.e. in the region of the zeros of M_2^{12} (M_2^{21}). This peak is of significant magnitude only when $|t_1| \approx |t_2|$. In the strong (weak) coupling regime, 8 (10) peaks occur, as expected; the peaks that are closer to the band gap vanish (emerge) as $\left|\frac{t_1}{t2}\right|$ increases. When the coupling is asymmetric, transmission is enhanced only in one of the two bands.

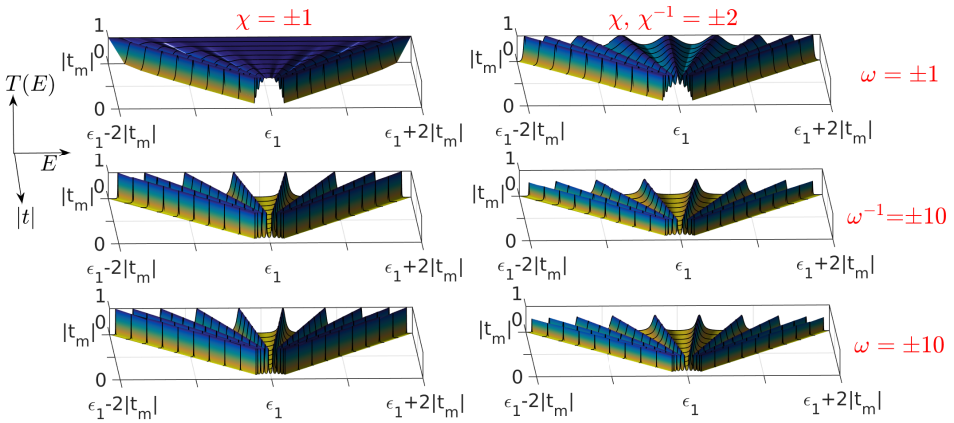

Figure 3.20: Transmission coefficient of a periodic system with $u = 1$ site per unit cell and $m = N = 10$ unit cells, for ideal (top), strong (middle) and weak (bottom) coupling with the leads. (Left column) Symmetric coupling. (Right column) Asymmetric coupling. The on-site energy of the system coincides with that of the leads ($\epsilon_1 = \epsilon_m$) and the hopping integral of the system increases until it is equal to that of the leads.

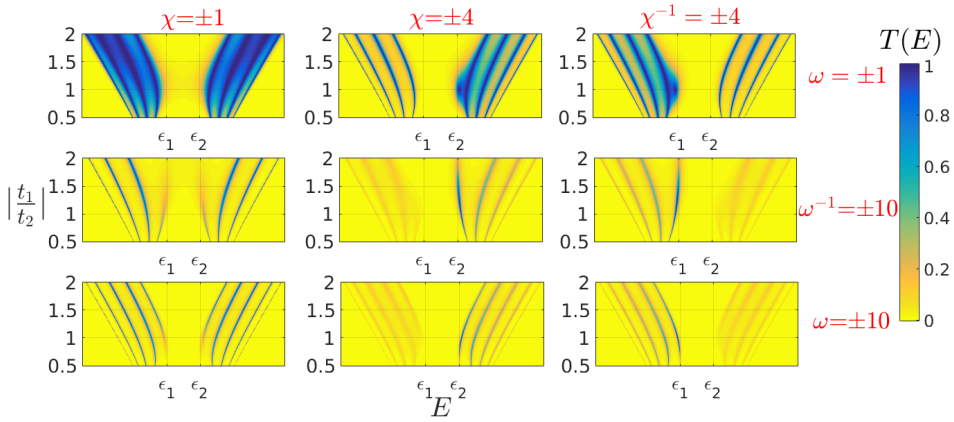

Figure 3.21: Transmission coefficient of a periodic system with $u = 2$ sites per unit cell and $m = 5$ unit cells ($N = 10$), for ideal (top), strong (middle) and weak (bottom) coupling with the leads. (Left column) Symmetric coupling. (Middle column) Asymmetric coupling with $|\chi| > 1$. (Right column) Asymmetric coupling with $|\chi| < 1$. We have chosen $\epsilon_m = \frac{\epsilon_1+\epsilon_2}{2}$ and the bandwidth of the leads is chosen so as to contain all the eigenstates of the system.

3.5 Conclusion

We employed the transfer matrix method to obtain the energy eigenvalues of a periodic WM, with a single orbital per site, u sites per unit cell, and nearest neighbor hoppings, for either cyclic or fixed boundaries. The solution of such a system is identical to the diagonalization of a real symmetric u-Toeplitz matrix of order mu, with or without perturbed upper left/lower right corners. The dispersion relation (cyclic boundaries) and the eigenspectrum (fixed boundaries) can be obtained with the help of the elements of the UCTM and the Chebyshev polynomials of the second kind. The DOS was also obtained with the help of the UCTM. The properties of the eigenvalues and the DOS were discussed through representative examples for the simplest cases, i.e. for systems with $u = 1, 2, 3, 4$, where the difference in the hopping integrals between

different sites inside the unit cell was explicitly taken into account. Furthermore, we attached the periodic systems under examination to semi-infinite homogeneous leads, and determined the TC at zero bias. We showed that in general the ideal and symmetric coupling condition is also the optimal coupling condition, and demonstrated the role of the coupling strength/weakness and asymmetry, as well as of the leads' properties, to the transmission profiles (number, position, sharpness of peaks). Finally, we demonstrated the effect of the intrinsic hoppings of the system to the TC, for systems with $u = 1, 2$, as representative examples.

4

Atomic carbon wires[5]

During the past years, the great scientific advances and achievements in the field of carbon allotropes, launched by the discovery of fullerenes in 1985 [184], the synthesis and determination of the structure of carbon nanotubes in 1991 [185] (although the first report on them can be traced back to 1952 [186]), and the rediscovery of graphene in 2004 [187], have drawn attention to other forms of carbon such as graphyne [188] and atomic carbon wires (also called carbynes). Carbynes are one-dimensional sp-hybridized chains of carbon and represent the ultimate nanoscale structure, a molecular wire with one-atom thickness [189, 190]. The very interesting properties these structures hold (tunable band gap, extreme stiffness and elastic modulus, high flexibility, room temperature persistence) have drawn a large scientific attention during this decade. The first direct electrical characterization of atomic carbon chains was reported in 2013 [191], and, although up until 2016 the longest synthesized carbon chain consisted of 44 atoms [192], more than 6,000 contiguous acetylenic carbons have been produced inside double-walled carbon nanotubes [193]. The thermal [194–197], vibrational [198–200], mechanical [196, 201], and electronic and charge transport properties [191, 202–211] of atomic carbon wires are fields of intense contemporary research.

The carbyne structure is characterized by the BLA (in the cumulenic case, BLA $= 0$, while in the polyynic case, BLA $\neq 0$). Although it has been demonstrated that Peierls distortion [212], the driving force of which is the electron phonon-coupling [213], energetically favors the polyynic configuration [214, 215], leading the state with BLA $\neq 0$ to be more stable, it has been theoretically shown [216] and experimentally observed [210] that under no strain, the cumulenic phase is also possible. This preservation of the cumulenic symmetry over the broken-symmetry polyynic carbyne in an unstrained chain has been attributed to the elimination of the Peierls distortion due to the zero-point atomic vibrations [216]. Another recent work has shown that capping a

[5]Reproduced from Ref. [183] with permission from the PCCP Owner Societies.

small finite atomic carbon wire between sp^2 conjugated end groups with a small number of aromatic units leads to a reduction of the BLA and the band gap, and hence to a more cumulene-like structure [217].

The aim of this Chapter is to apply the WM in order to study the electronic structure and charge transport properties of cumulenic and polyynic carbynes, using the formulation presented in Chapter 3. We show that TB predicts various key aspects and can efficiently catch and explain some of the most important recent experimental results. These results include the metallic (semiconducting) behavior of the current-voltage ($I-V$) characteristics for cumulenic (polyynic) carbynes, the effect of weakening the coupling between the leads and one end of the system, as well as the rectifying behavior polyynic carbynes demonstrate when there is a mismatch between their Fermi levels and the ones of the leads.

The cumulenic form of carbyne, in which BLA $= 0$, can be interpreted as

$$\cdots = \mathrm{C} = \mathrm{C} = \mathrm{C} = \mathrm{C} = \mathrm{C} = \mathrm{C} = \cdots .$$

Hence, in a TB treatment, cumulenic carbyne can be considered as a homogeneous chain with one p_x electron per atom and one atom per unit cell, with on-site energy ϵ_{C}, and a single hopping integral, t. On the contrary, its polyynic form, in which BLA $\neq 0$, the structure can be interpreted as

$$\cdots \equiv \mathrm{C} - \mathrm{C} \equiv \mathrm{C} - \mathrm{C} \equiv \mathrm{C} - \mathrm{C} \equiv \cdots .$$

Hence, in a TB treatment, polyynic carbyne can be considered as a chain with one p_x electron per atom and two atoms per unit cell, again with on-site energy ϵ_{C}, and two hopping integrals, t_s and t_ℓ, for the short (s) and long (ℓ) separations, respectively. An identical treatment holds for the p_y electrons, due to energy degeneracy.

4.1 Electronic structure of carbynes

Let us suppose a carbyne chain made up of N carbon atoms. According to the analysis presented in Chapter 3, the GTM of the chain is $\boldsymbol{P}_1^N$ for cumulenic carbynes, while, for polyynic carbynes, it is $(\boldsymbol{P}_2\boldsymbol{P}_1)^{N/2}$. Taking into account the spectral duality relations [90, 173] and the relation of the elements of the (unimodular) global transfer matrix with the Chebyshev polynomials of second kind [127, 174], we can determine the electronic structure of carbynes for either cyclic or fixed boundary conditions [183].

To obtain physically valid numerical results from a TB approach, an appropriate set of external parameters (on-site energies and hopping integrals) is necessary. In Ref. [218], the authors have calculated by ab initio methods the bond lengths of atomic carbon wire structures, i.e. $d = 1.282$ Å for the cumulenic form of carbyne, and $d_\ell = 1.301$ Å, $d_s = 1.265$ Å for the polyynic form. Taking these values, we obtain the hopping integrals we will use below to calculate the $I-V$ characteristics of cumulenic and polyynic carbynes via the Harrison formula [85] for the $pp\pi$ bond, i.e.,

$$t = -0.63\frac{\hbar^2}{md^2}, \tag{4.1}$$

where $\hbar$ is the reduced Planck constant and m is the electron mass. Hence, for cumulenic carbynes, $t = -2.92$ eV, while, for polyynic carbynes $t_\ell = -2.84$ eV and $t_s = -3.00$ eV. We should mention that the inclusion of longer-range terms (second, third, etc neighbors) is not expected to have a significant effect of the energy structure. The Harrison formula that we employed is valid for interatomic distances of the order of the covalent bond. In the case of larger distances, other, exponentially decaying formulas are more accurate [219, 220]. Using such an exponential law, in the fashion done in Ref. [221], e.g. for the bond length we have used for cumulenic carbynes, we find that the hopping integral for the second neighbors would be smaller by one order of magnitude. The on-site energy of the HOMO of the sp-hybridized carbon is set to zero. As it will be shown below, altering $\epsilon_{\rm C}$ does not affect the energy structure of the carbynes, apart from an energy shift, but, when their transport properties are considered, the choice of the metallic leads' on-site energy relative to that of carbon is of major significance.

For cyclic boundaries ($\psi_0 = \psi_N$, $\psi_{N+1} = \psi_1$), the following dispersion relations hold for carbynes

$$E = \epsilon_{\rm C} + 2t\cos(k_{\rm c}a_{\rm c}), \quad \text{cumulenic,} \tag{4.2a}$$

$$E = \epsilon_{\rm C} \pm \sqrt{t_1^2 + t_2^2 + 2t_1t_2\cos\left(k_{\rm p}a_{\rm p}\right)}, \ \text{polyynic,} \tag{4.2b}$$

where $k_{\rm c,p} \in 1^{\rm st}$ Brillouin Zone, $a_{\rm c} = d$, and $a_{\rm p} = d_s + d_\ell$. Since there is only one electron per atom, cumulenic carbynes display metallic behavior (there is one half-filled band), while polyynic carbynes display semiconducting behavior (one filled and one empty band). The Fermi level is $\epsilon_{\rm C}$. The transition from metal to semiconductor is displayed in Fig. 4.1(a), where Eq. (4.2b) is plotted as a function of the ratio t_1/t_2. We have fixed $t_2 = t_\ell = -2.84$ eV, and let t_1 gain values in the interval that corresponds to distances 1.20 - 1.53 Å, i.e. for bond lengths varying from the triple to the single bond in individual gas-phase carbon molecules [222], using the Harrison formula. We notice that for $t_1 = t_s = -3.00$ eV, i.e. for the set of parameters used below for the calculation of the $I - V$ characteristics [black curve in Fig. 4.1(a)], the bandgap is $E_g = 0.32$ eV, in perfect accordance with the gap reported in Ref. [218], from which the bond lengths were taken. Hence, our treatment is in excellent agreement with the much more time-consuming and complicated ab initio method [218]. Generally, as t_1/t_2 increases up to 1 (i.e. the cumulenic case), the width of each band linearly increases and the bandgap linearly decreases to zero. For $t_1/t_2 > 1$, the width of each band remains constant and the gap starts to increase linearly (see inset of Fig. 4.1(a)).

For fixed boundaries ($\psi_0 = \psi_{N+1} = 0$), the eigenenergies of carbynes can be obtained recursively via the relations

$$U_N(z_c) = 0, \quad \text{cumulenic,} \tag{4.3a}$$

$$\frac{t_1}{t_2}U_{\frac{N}{2}}(z_p) = -U_{\frac{N}{2}-1}(z_p), \quad \text{polyynic,} \tag{4.3b}$$

where $z_c = {\rm Tr}(\boldsymbol{P}_1)/2 = \frac{E-\epsilon_{\rm C}}{2t}$, and $z_p = {\rm Tr}(\boldsymbol{P}_2\boldsymbol{P}_1)/2 = \frac{(E-\epsilon_{\rm C})^2-t_1^2-t_2^2}{2t_1t_2}$. The eigenspectrum, i.e. the spectrum of eigenvalues, as derived by Eq. (4.3b), for polyynic carbynes

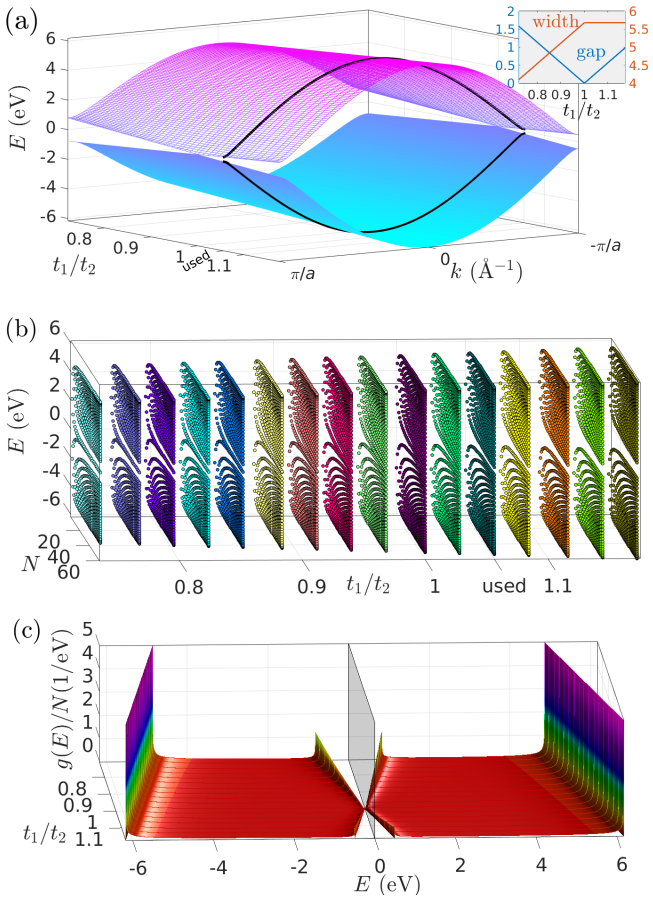

Figure 4.1: (a) Dispersion relation [(inset) width of each band and bandgap], (b) eigenspectra for up to $N = 60$, (c) density of states of polyynic carbynes as a function of the hopping integral ratio, t_1/t_2. "Used" corresponds to the ratio t_s/t_ℓ (see text).

is displayed in Fig. 4.1(b), as a function of t_1/t_2, for up to $N = 60$ atoms. It is again evident that, as t_1/t_2 increases above 1, an increasing gap occurs. This gap vanishes when $t_1/t_2 = 1$, i.e. when the transition to cumulenic carbyne occurs. For $t_1/t_2 < 1$ a bandgap starts to open again, but there is an additional eigenvalue in each band that as N increases, reaches ϵ_C. This effect occurs straight from the choice of the boundary conditions. The eigenspectra of such systems have also been discussed, from a mathematical point of view, in Refs. [162, 164, 223]. The fixed boundary conditions correspond to a finite carbon wire structure, the ends of which induce a length-dependent modulation in the BLA [190]. In our model, this would result in an alteration of the ratio t_1/t_2. Hence, for a finite structure, the modulation of the BLA could cause a shift along the horizontal axis of Fig. 4.1(b), as N increases.

The DOS of carbynes is given by the expressions

$$g(E) = \frac{N}{\pi\sqrt{4t^2 - (E - \epsilon_C)^2}}, \quad \text{cumulenic,} \tag{4.4a}$$

$$g(E) = \frac{N|E - \epsilon_C|}{\pi\sqrt{4t_1^2t_2^2 - [(E - \epsilon_C)^2 - t_1^2 - t_2^2]^2}}, \quad \text{polyynic,} \tag{4.4b}$$

where the double spin and double band degeneracies are not incorporated. Two van Hove singularities occur in cumulenic carbynes exactly at the band edges, i.e. for $E = \epsilon_{\mathrm{C}} \pm 2t$. In polyynic carbynes, the singularities increase to four, again at the band edges, i.e. for $E = \epsilon_C \pm |t_1 \pm t_2|$. The DOS of polyynic carbynes is displayed in Fig. 4.1(c) as a function of t_1/t_2. We observe that the singularities at the top (bottom) of the valence (conduction) band are much less sharp that the ones at the bottom (top). These less sharp singularities vanish as t_1/t_2 reaches 1, i.e. when the transition to cumulenic carbyne occurs.

4.2 Transmission coefficient

Suppose we interconnect the ends of a given atomic carbon wire with two semi-infinite homogeneous leads with lattice constant a and one electron per site, so that they are metallic. Let $\epsilon_{mL(R)}$ and $t_{mL(R)}$ be the on-site energy and the hopping integral of the left (right) lead, so that their energy structure is described by the dispersion relation $E = \epsilon_{mL(R)} + 2t_{mL(R)} \cos\big(q_{L(R)}a\big)$. Hence $\epsilon_{mL(R)} = E_{F,L(R)}$, i.e. the Fermi level of the leads, and the bandwidth is $4\big|t_{mL(R)}\big|$. The coupling of the leads with the end carbon atoms is described by the hopping integrals $t_{L(R)}$. If we suppose identical leads and, without any loss of generality, that the incident waves come from the left lead, then $T(E) = \frac{1}{1+\Lambda(E)}$ and the polynomials $W_N(E)$, and $X_N^{\pm}(E)$ involved in the $\Lambda(E)$ (cf. Eqs. (3.49, 3.50) take the form

$$W_N(E) = U_N(z_c)\omega + U_{N-2}(z_c)\omega^{-1}, \tag{4.5a}$$

$$X_N^{\pm}(E) = U_{N-1}(z_c)(\pm\chi^{-1} - \chi), \tag{4.5b}$$

for cumulenic carbynes, and

$$W_N(E) = U_{\frac{N}{2}}(z_p)\omega + \frac{t_2}{t_1}U_{\frac{N}{2}-1}(z_p)(\omega + \omega^{-1}) + U_{\frac{N}{2}-2}(z_p)\omega^{-1}, \tag{4.6a}$$

$$X_N^{\pm}(E) = U_{\frac{N}{2}-1}(z_p)(E - \epsilon_{\mathrm{C}})(\pm\frac{\chi^{-1}}{t_2} - \frac{\chi}{t_1}), \tag{4.6b}$$

for polyynic carbynes.

The zero-bias TC of a polyynic carbyne with $N = 10$ is plotted as a function of t_1/t_2 in Fig. 4.2, for various coupling strength and asymmetry conditions. We have chosen the Fermi level of the leads, $E_F(= \epsilon_m)$ to be equal to that of the carbyne, ϵ_{C}. The coupling integral of the leads is taken $t_m = -4$ eV, so that their bandwidth contains all the eigenstates of the polyynic chain. As it is evident from Fig. 4.2, the ideal and symmetric coupling conditions ($\omega = \chi = 1$) leads to the most enhanced transmission. When the coupling strength factor deviates from the ideal case, the transmission peaks sharpen and the transparency of the system is significantly reduced. If the not-ideal coupling is also asymmetric, then the value of the transmission peaks also reduces, leading to a further decrease of the system's transparency. The exact number, as well as the position of the transmission peaks in each coupling strength and asymmetry condition can

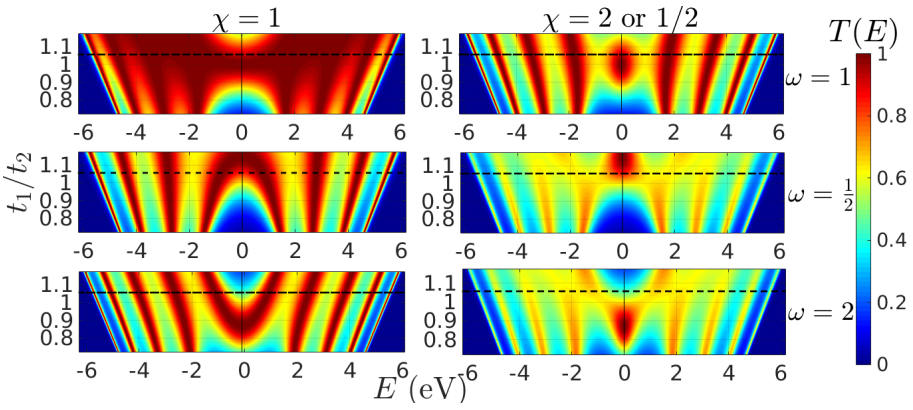

Figure 4.2: Zero-bias TC for a polyynic carbyne with $N = 10$ atoms, as a function of the ratio t_1/t_2. Identical leads have been assumed, with $\epsilon_m = \epsilon_C$ and $t_m = -4$ eV. (Top) Ideal coupling condition. (Middle) Strong coupling condition with $\omega = 1/2$. (Bottom) Weak coupling condition with $\omega = 2$. (Left column) Symmetric coupling condition. (Right column) Asymmetric coupling condition, with $\chi = 2, 1/2$. The TC that corresponds to the parameters $t_1 = t_s = -3.00$ eV, $t_2 = t_\ell = -2.84$ eV, as derived by the bond distances of Ref. [218], is noted with horizontal dashed lines in each panel.

be determined by examining the degree of the terms that become dominant in $\Lambda(E)$ (cf. Chapter 3). In other words, the number and position of the transmission peaks are strongly related to the coupling strength and asymmetry, as well the ratio t_1/t_2. This suggests that the peaks are not exclusively determined by the energy levels of the system.

4.3 Current-Voltage curves

The $I - V$ curve of a one-dimensional TB model can be given, using the Landauer-Büttiker formalism [109, 224–227], by the relation

$$I(V) = \frac{2e}{h} \int_{-\infty}^{\infty} T(E, V)[f_L(E - \mu_L) - f_R(E - \mu_R)]dE, \tag{4.7}$$

under the assumption that charges propagate 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. The factor 2 comes from spin degeneracy. From Eq. (4.7) we deduce that, apart from the energy structure of the system under examination, which is reflected in the TC, as well as the coupling strength and asymmetry, which have a significant effect on the transmission profile (c.f. Sec. 3.4), there are several factors that have an effect on the magnitude of currents, the bias regime and the shape of the $I - V$ curves [82]. These factors include:

1) Whether or not the TC is considered as bias-dependent. Although assuming bias-independent TC could be a justified choice at 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. If an external bias is applied between the (generally, not identical) leads, then the transmission coefficient takes the form

$$T(E,V)=\frac{2\sin(q_La)\sin(q_Ra)\dfrac{t_{mL}}{t_{mR}}}{\mathrm{Tr}\Big(\tilde{\boldsymbol{M}}_N^T\tilde{\boldsymbol{M}}_N\Big)/2+(\tilde{M}_N^T\tilde{M}_N)^{12}\cos(q_La)-(\tilde{M}_N\tilde{M}_N^T)^{12}\cos(q_Ra)\\ -\tilde{M}_N^{11}\tilde{M}_N^{22}\cos[(q_R+q_L)a]-\tilde{M}_N^{12}\tilde{M}_N^{21}\cos[(q_R-q_L)a]}. \tag{4.8}$$

This equation can be derived by calculating the probability current at each lead through the continuity equation and is used throughout the present thesis everywhere the $I-V$ curves are presented. For identical leads and $V=0$, Eq. (4.8) leads to Eq. (3.42).

2) The way the external bias is applied. For example, (a) only one of the leads' energy bands can be shifted, so that $\mu_L=E_F+eV$, and $\mu_R=E_F$, or, alternatively, (b) both leads' bands can be symmetrically shifted so that $\mu_{\substack{L\\R}}=E_F\pm\frac{eV}{2}$. This choice affects both the way the voltage drop is induced in the examined system and the energy limits of the conductance channel. At finite temperatures, Eq. (4.7) can be written in the form

$$I(V)=\frac{2e}{h}\sinh\left(\frac{eV}{2k_BT}\right)\int_{-\infty}^{\infty}\frac{T(E,V)dE}{\cosh\left(\frac{E-E_F}{k_BT}\right)+\cosh\left(\frac{eV}{2k_BT}\right)}, \tag{4.9}$$

i.e., the $I-V$ curve occurs from the modulation of the hyperbolic function $\sinh\left(\frac{eV}{2k_BT}\right)$ by the integral factor expression [228], while, at zero temperature, the Fermi-Dirac distributions become Heaviside step functions and determine the limits of integration. Hence, Eq. (4.7) can be simplified to

$$I(V)=\frac{2e}{h}\int_{\mu_R}^{\mu_L}T(E,V)dE. \tag{4.10}$$

3) The choice of the Fermi level of the leads. If this level is not aligned with a transmission resonance, then no currents occur in the vicinity of $V=0$, while a metallic-like behavior can be expected otherwise.

Here, we assume that a constant bias V is applied at the left lead, so that its dispersion relation is shifted by $V_b=eV$, and a voltage drop within the carbyne system, so that the TC becomes voltage-dependent, i.e. $T(E)=T(E,V)$, cf. Eq. (4.8). We apply a linear voltage drop between the end sites of the leads, separated by a distance L, so that the on-site energy ϵ_C of the nth carbon atom is shifted by $V_b(n)=V_b\frac{d_n}{L}$, where d_n is the distance of the nth atom from the end of the left lead. At zero temperature, the Fermi-Dirac distributions at the leads are $f_L(E)=\theta(E_L^F+V_b-E)$, $f_R(E)=\theta(E_R^F-E)$, where θ denotes the Heaviside step function. The conducting states lie in the intersection of the common energies between the leads and the bandwidth of carbyne chains. Then, the electrical current can be computed using the Landauer-Büttiker formalism as

$$I(V)=\frac{4e}{h}\int_{E_R^F}^{E_L^F+V_b}T(E,V)dE, \tag{4.11}$$

cf. Eq. (4.10). Here, the factor 2 changes to 4 due to the additional double degeneracy of the bands formed by the p_x and p_y electrons. As shown in the previous section, the coupling strength and asymmetry play a significant role in the transmission profile. So, when quantitative reproduction of experimental results is considered, explicit information about the size of the chain, electronic structure of the leads and the strength of coupling between the leads and the chain is necessary. In the following, we study some of the experimental results of Ref. [210], showing that our model can catch some key aspects of the experimental observations, and focus on the influence of the leads' parameters to the $I-V$ characteristics. Since we have no explicit information about the above mentioned details, we will display our results for $N=100$, as a representative case of relatively long atomic carbon wires (the recruitment of which in devices is experimentally sought) that can be treated with TB. Our calculations show that altering N has only some quantitative and not qualitative effect on the physical properties. Although the currents observed in Ref. [210] are no larger than some nA, later experimental findings report currents that lie in the μA regime [211], as our below discussed results indicate. Quantitative differences could also be attributed to the temperature effects, since the experiments were carried out at room temperatures. Without specifying the nature of the leads, we will choose $E_{L,R}^F=\epsilon_{\rm C}$ and $t_{mL,R}=3.16$ eV, a value that corresponds to the overlap of neighboring atoms in a single graphitic layer plane [229]. For the carbon chain, we will use t, for the cumulenic case, and $t_1=t_s$, $t_2=t_\ell$ for the polyynic case. As for the coupling hopping integrals, we choose $t_L=t_R=-2.67$ eV for cumulenic carbynes, and $t_L=-2.05$ eV, $t_R=-3.33$ eV for polyynic carbynes, i.e. values that correspond to double, single and triple bonds, respectively, between gas-phase carbon molecules [222], according to the Harrison formula (see Sec. 4.1). The results for the $I-V$ characteristics for cumulenic and polyynic carbynes with $N=100$ for the above mentioned parameters are shown in Fig. 4.3.

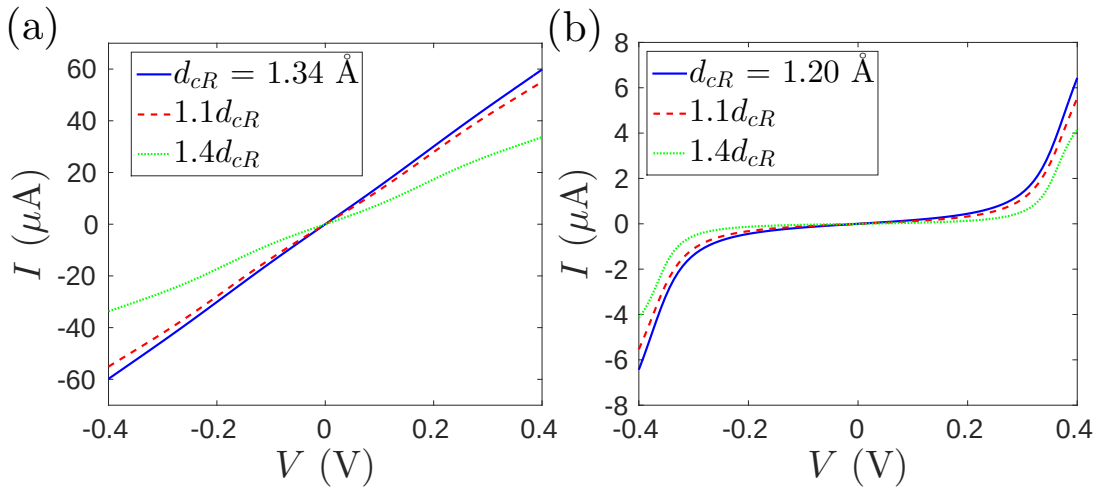

Figure 4.3: $I-V$ characteristics of (a) cumulenic (b) polyynic carbynes with $N=100$ carbon atoms. The effect of enlarging the distance between the end of the chain and the right lead, d_{cR} (continuous blue curves), is also demonstrated, for 10% (red dashed curves) and 40% (green dotted curves) increase.

From Fig. 4.3 it is evident that cumulenic carbynes display a typical Ohmic behav-

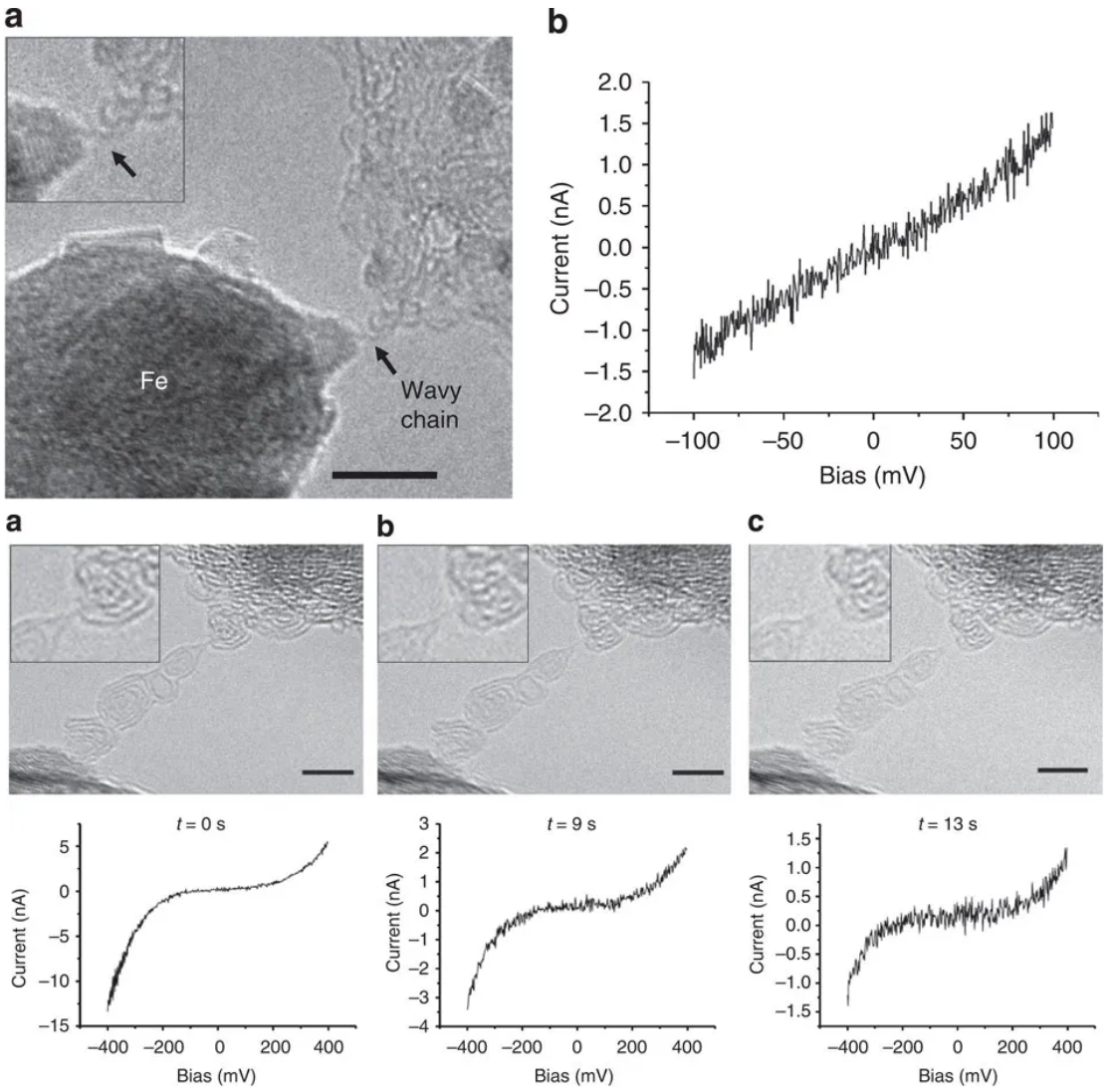

Figure 4.4: (top) TEM image and (b) the corresponding $I - V$ curve for an unstrained atomic carbon chain, corresponding to the cumulenic phase. The $I - V$ curve is linear, corresponding to Ohmic conductivity. (bottom) The temporal evolution of a strained carbon chain, corresponding to the polyynic phase (a) at an arbitrary time zero; (b) after 9 s; (c) after 13 s, i.e., as the carbyne chains coupling with one of the leads becomes increasingly poorer. The corresponding $I - V$ curves illustrate indicate semiconducting behavior, although with a decreasing conductivity from (a) to (c). Images reproduced from Ref. [210], used under CC BY 4.0.

ior (linear curves), while the S-shaped curves of polyynic carbynes are indicative of a semiconducting behavior. Furthermore, the currents for cumulenic carbynes are of larger magnitude than for polyynic carbynes, in the same bias regime, in agreement with Ref. [210]; see Fig. 4.4. In the same work, the experimental $I - V$ curves were measured at different times, as the carbyne chains coupling with one of the leads becomes increasingly poorer, resulting to reduced currents. We interpreted this effect by calculating the $I - V$ curves for increasing values of d_{cR}, i.e. the distance between the last carbon atom and the left lead. This results in a decrease of the coupling hopping

integral t_R, which is determined by the Harrison formula. In Fig. 4.3 we show that the decrease of the carbyne coupling to the right lead results in reduced currents, in agreement with the experimental observation of Ref. [210], cf. Fig. 4.4.

Let us now examine the influence of the leads' properties to the $I-V$ characteristics. Suppose that there is a mismatch between the Fermi level of a polyynic carbyne and that of the leads, which is defined as $\Delta = \epsilon_m - \epsilon_C$, and all the rest parameters are kept the same. The calculated $I-V$ curves and the differential conductance, $\frac{dI}{dV}$, for different values of Δ are presented in Fig. 4.5. From the $I-V$ curves (left panel of Fig. 4.5) it can be observed that increasing $|\Delta|$ from zero leads to a reduction of the gap region in the vicinity of $V = 0$, until $|\Delta|$ exceeds $\frac{E_g}{2} = 0.16$ eV, in which case the gap region around $V = 0$ vanishes. This happens because for $\Delta > \frac{E_g}{2} (< -\frac{E_g}{2})$, the Fermi level of the metal lies within the conduction (valence) band of the polyynic carbyne, leading to non-vanishing currents for small biases. The alternation of the gap region due to the energetic difference between the contacts and the leads has also been suggested by recent ab initio calculations, where a significant difference in the $I-V$ curves is found when the atom chain is sandwiched between graphene layers and end-capped carbon nanotubes [211]. From the right panel of Fig. 4.5 it can be seen that the differential conductance has an oscillatory behavior. Although for $\Delta = 0$, $\frac{dI}{dV}$ is symmetric around $V = 0$, the increase of $|\Delta|$ induces an increasing asymmetry, such that $\frac{dI}{dV}(\Delta, V) = \frac{dI}{dV}(-\Delta, -V)$. Furthermore, it is evident that, for $|\Delta| > \frac{E_g}{2}$, the valley of the differential conductance in the vicinity $V = 0$ vanishes, giving rise to large conductance in the small bias regime.

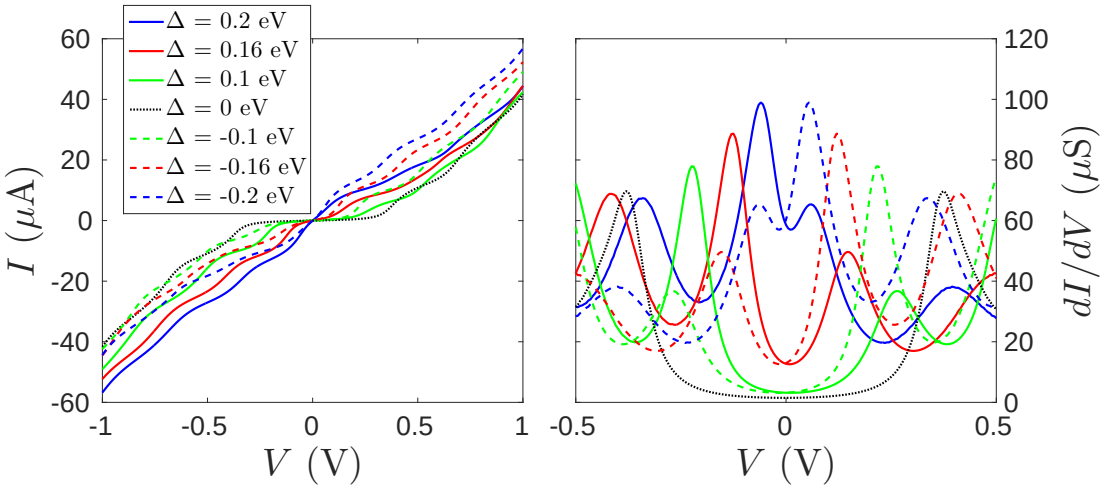

Figure 4.5: $I-V$ characteristics (left) and differential conductance (right) of a polyynic carbyne with $N = 100$ atoms for different values of the energy difference between the Fermi level of the carbyne and the metallic leads.

Next, we examine the effect of altering the leads' bandwidth on the $I-V$ curves. For a specific value of bias, the increase of leads' bandwidth, in terms of increasing $|t_m|$, for fixed coupling hopping integrals, results in increasing ω (thus weakening the coupling), so the transmission peaks become sharper. This is evident by comparing the peaks that occur for $V = 0.36$ V and for $V = 0.50$ V in Fig.4.6(a), where $T(E)$ is

plotted in the interval of the incident carrier energies in each case. Furthermore, from Fig. 4.6(a) it can be seen that for $V = 0.50$ V, the peaks become wider, their position alters, and their value is smaller. The synergetic effect of these changes to $T(E)$ can quantitatively be demonstrated by simply taking the average of the TC, $\bar{T}(E)$, in each case [inset of Fig. 4.6(a)]. By observing the inset of Fig. 4.6(a), we can expect that the increase of the leads' bandwidth results in reduced current, for the same value of bias. Finally, we can expect that the increase of the current value between 0.36 and 0.50 V will be significantly smaller for increased leads' bandwidth, i.e. the $I-V$ curve in this case will have a more step-like behavior. These observations are supported by the explicit calculation of the current-voltage curves via Eq. (4.11), which are presented in Fig. 4.6(b). The progressive increase of the leads' bandwidth results generally in reduced currents and gradually to a step-like increase of the current with applied bias. The occurrence of this step-like behavior can be explained in terms of the "molecular eigenvalue staircase" [230]. As the bias increases, more transmission peaks (resonances) are included in the broadening conductance channel, resulting in gradual jumps in the $I-V$ curves.

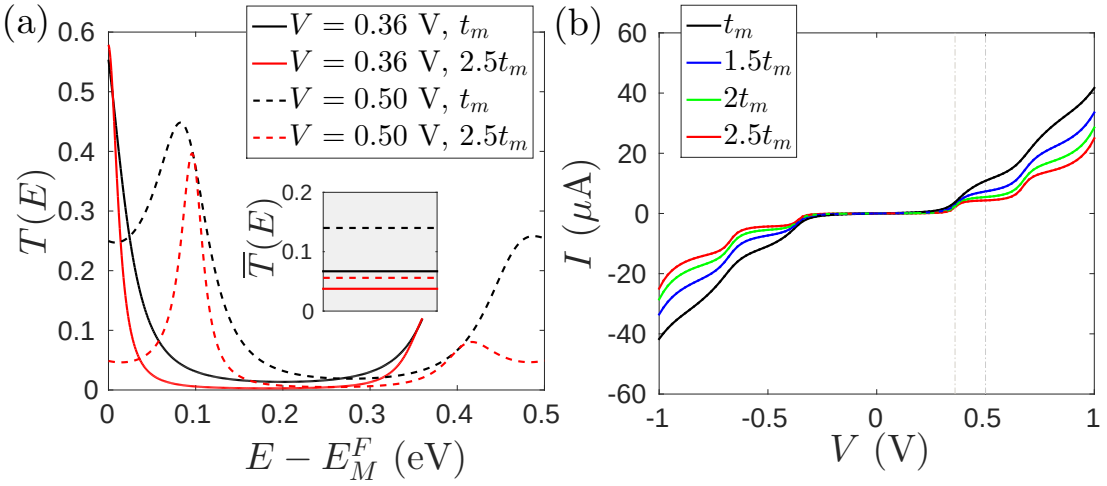

Figure 4.6: (a) Transmission coefficients of a polyynic carbyne with $N = 100$ atoms for $V = 0.36$ V and $V = 0.50$ V for different leads' bandwidths ($t_m = 3.16$ eV and $2.5t_m$). The average values of the transmission coefficient in the energy interval of the incident carriers are presented in the inset. (b) $I-V$ characteristics of the same system, for different leads' bandwidths, as determined by t_m.

Finally, we examine how considering different contacts affects the $I-V$ characteristic of the carbon chain. We assume that the Fermi level of the left (right) lead is shifted 0.2 eV above (below) the Fermi level of the polyynic carbyne and that $t_{mL} = 3.16$ eV $= 0.5t_{mR}$. The calculated $I-V$ curve under these conditions, which is presented in Fig. 4.7, displays a rectifying behavior. This is in accordance with the experimental findings of Ref. [210], also shown in Fig. 4.7, in which a rectifying behavior is observed when the carbyne chains are lying between two different contacts.

The significant influence the contacts have on the charge transport properties of carbon wires demonstrated by our results is in accordance with the conclusions of experimental works using the Scanning tunneling microscope break junction (STM-BJ) technique in order to study the conductance of end-capped oligoynes [204, 205].

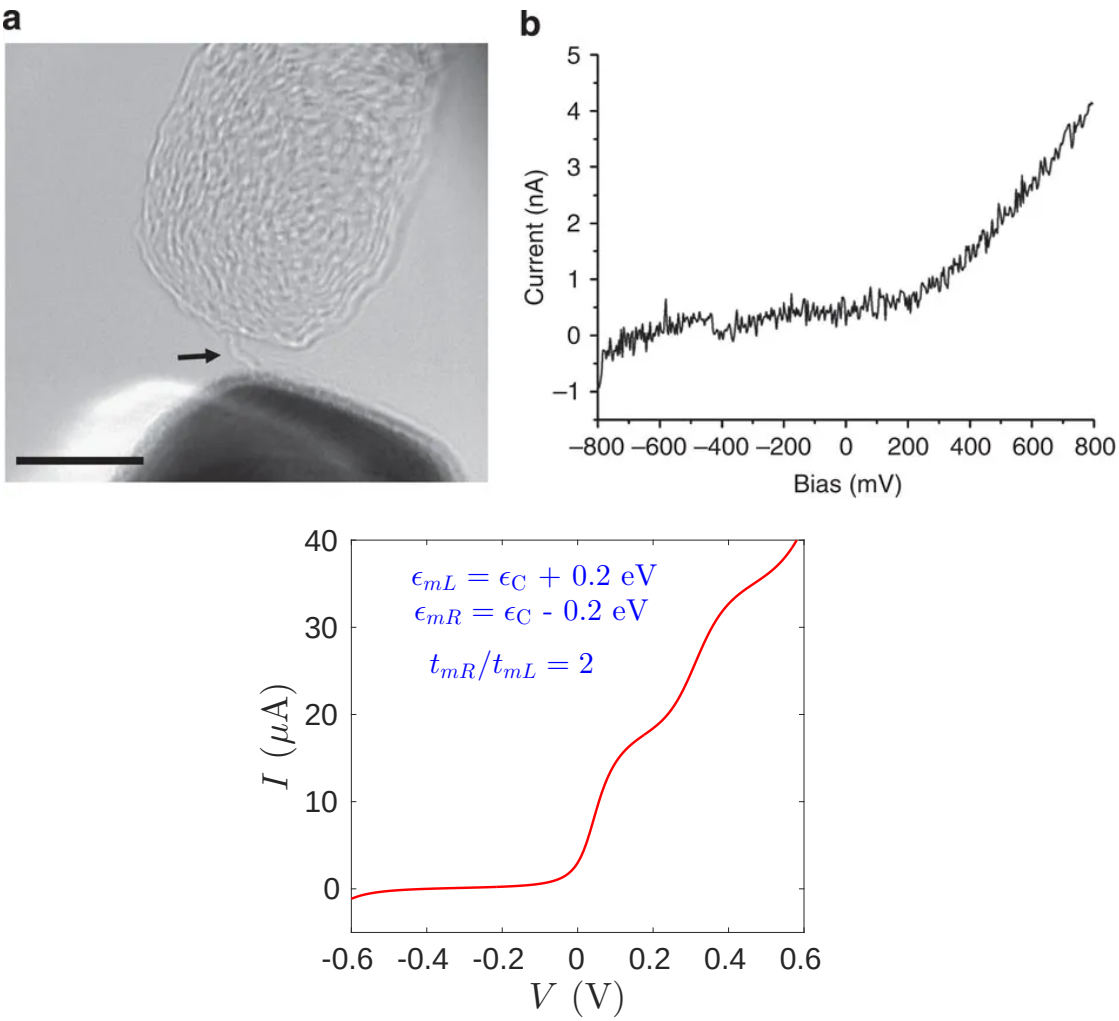

Figure 4.7: (top) Carbon chain between two dissimilar electrodes. The bottom electrode is Fe; the top electrode is C. (a) TEM image. Scale bar, 5 nm. (b) $I-V$ curve. Image reproduced from Ref. [210], used under CC BY 4.0. (bottom) Our calculated $I-V$ curve of a polyynic carbyne with $N = 100$ atoms, sandwiched between two different leads.

4.4 Conclusion

We employed a WM using the formulation of Chapter 3 in order to obtain the electronic structure of cumulenic and polyynic carbynes, in terms of their dispersion relations (for cyclic boundaries), eigenspectra (fixed boundaries), and DOS, obtaining a bandgap for polyynic carbynes at 0.32 eV, in agreement with more complex ab initio calculations. We analytically determined the TC at zero-bias by attaching the carbynes to semi-infinite metallic leads, and displayed the effect of the coupling strength and asymmetry to the transparency of the system to the incident carriers. Then, we calculated the I-V characteristics of cumulenic and polyynic carbynes, which display metallic and semiconducting behavior, respectively. Furthermore, we studied the effect on the I-V curves of factors such as the weakening of the coupling of the system to one of the leads, the relative position of the Fermi levels of the carbyne and the leads,

the leads' bandwidth and, finally, the difference in the energy structure between the leads. Weakening the coupling of the system to the leads reduced currents. Altering the relative position of the Fermi levels of the carbyne and the leads has an effect in the gap region and significantly changes the differential conductance. Increasing the bandwidth reduces the currents and leads to a more step-like increase with bias. Finally, introducing a difference in the energy structure of the leads results in rectifying I-V curves. Our results account for the main experimental features reported in the most recent experimental findings.

5

Periodic, deterministic aperiodic, and random binary DNA segments[6]

In this Chapter, we focus on the energy structure and charge transport properties of periodic, deterministic aperiodic and random binary DNA sequences, i.e., sequences based on the binary alphabet {G,A}. This means that in the $5'$-$3'$ strand of double helix B-DNA we have either G or A. We will use this strand to denote the segments. In the complementary, $3'$-$5'$, strand we have C and T, respectively. For example, the notation GGAG means that we have the GGAG bases in the $5'$-$3'$ strand and the complementary ones, CCTC, in the $3'$-$5'$ strand. All studied sequences start with G.

Several works have been devoted to the study of transfer and transport in specific DNA structures (periodic [56, 180, 232], quasiperiodic [95, 105, 233], random and natural [12, 72, 73, 234, 235]) using variants of the Tight-Binding (TB) method. Here, we employ the WM, with the sites of the chain being the base pairs (i.e., the on-site energies refer to a base pair and the hopping integrals correspond to adjacent base-pairs in the 5'-3' direction), to study the spectral, localization and charge transport properties of periodic, deterministic aperiodic [Thue-Morse (*TM*), Fibonacci (*F*), Period Doubling (*PD*), Rudin-Shapiro (*RS*), Cantor set (*CS*), generalized Cantor set ($GCS(4,2)$), Kolakoski ($KOL(1,2)$ and $KOL(1,3)$); see Sec. 2.4] and random DNA binary segments.

TB parameters may change at different levels of theory, and their values can tune the results, having both quantitative and qualitative effects. In Ref. [236] there are some nice tables showing the variance of on-site energies and coupling parameters for different triplets (or triads) of base pairs. When dealing with charge transport properties, it is usual in the literature to use only one hopping parameter and/or on-site energy, to simplify the problem. Here, we go beyond these simplifying hypotheses.

[6]The content of this chapter can be found published in Ref. [231]: K. Lambropoulos and C. Simserides, Phys. Rev. E **99**, 032415 (2019). © 2019 by the American Physical Society.

We use a TB parametrization that allows for different hopping (or transfer or coupling) parameters (or integrals). Specifically, we use the parametrization of presented in Ref. [55] for HOMO (see Table A.1 in Appendix A), which will be henceforth referred to as S parametrization. This leads to quantitative and qualitative consequences. Our treatment gives a clearer picture, as it will be discussed below. In this spirit, we calculate –among other quantities– autocorrelation functions, integrated density of states, Lyapunov exponents, TC and $I - V$ curves, taking into account the different on-site energies as well as the different hopping parameters.

We notice that, in what follows, for substitutional sequences based on a binary alphabet, the substitution rules of Table 2.2 hold as such, that is, $\mathcal{A} = \{\mathrm{G}, \mathrm{A}\}$, while for the rest, and, specifically, *RS* and $KOL(1,3)$ helping alphabets are defined. In the former, $\mathcal{A} = \{\mathrm{GG}, \mathrm{GA}, \mathrm{AG}, \mathrm{AA}\}$. In the latter, where $1 \to \mathrm{G}$, $3 \to \mathrm{A}$, $\mathcal{A} = \{\mathrm{GG}, \mathrm{GA}, \mathrm{AA}\}$ (cf. Table 2.2).

5.1 Sequence properties

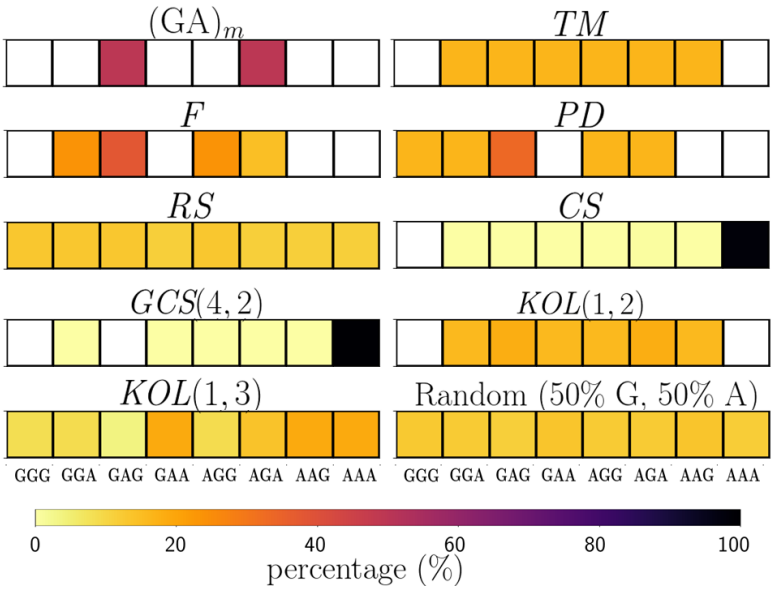

Figure 5.1: Classification of the DNA segments studied in this chapter based on the number and occurrence percentage of base-pair triplets. The boxes correspond to each of the 8 possible triplets. For each segment, white boxes correspond to forbidden triplets, and the color of the rest corresponds to their occurrence percentage (calculated for large N).

To obtain a clear picture of the interplay between sequence intricacy and energy profile of the segments, as well as its effect on localization and transport properties, we present some details on the sequence characteristics of the studied segments. Since we deal with binary DNA sequences, a useful classification of their properties can be done through the study of the different base-pair triplets that are found in each category [159]. A triplet is made of a base pair and its next and previous neighbors. Since in a realistic treatment we need to simultaneously consider the difference in the

on-site energies and the hopping integrals (as done here), the total number of possible triplets (2^3 for a binary sequence) corresponds to the total number of different transfer matrices that can be found in the GTM. Thus, a triplet is a natural unit in TB nearest neighbor interactions. Furthermore, triplets are of great importance in nucleic acids themselves. Codons are sequences of three DNA or RNA nucleotides that correspond with a specific amino acid or stop signal during protein synthesis. Each of the $3^4 = 64$ codons corresponds to a single amino acid (or stop signal), and the full set of codons is the genetic code. Finally, we notice, it has been claimed that the on-site energy of a base depends on its flanking bases, an idea beyond the scope of our present calculations [236].

The number of triplets in each studied category of DNA segments as well as the occurrence percentage of each triplet (for large N) are depicted in Fig. 5.1. From Fig. 5.1 it is obvious that the periodic $(GA)_m$ segment represents the most ordered case (2 triplets with equal occurrence percentages). *F* and *PD* segments possess 4 and 5 different triplets, respectively, and have one dominant GAG triplet. *TM* and *KOL*$(1, 2)$ segments posses 6 equidistributed triplets. *RS*, random and *KOL*$(1, 3)$ segments posses all possible triplets; in the first two cases they are equidistributed; in the latter there are some predominant triplets. Finally, the Cantor Set family segments posses many of the possible triplets [7 for *CS*, 6 for *GCS*$(4, 2)$]. However, the AAA triplets are predominant, asymptotically reaching 100% occurrence percentage as N increases. For segments with primitive substitution rules, the values at which the occurrence percentage of each possible triplet converge can analytically be found from the procedure described in Subsec. 2.4.1. For example, the occurrence percentages of the possible triplets in PD segments converge to the components of the normalized right eigenvector corresponding to $\lambda_{PF} = 2$ of matrix S_3 [Eq. (2.61)], i.e. $[\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}]^T$.

The intricacy of the sequence determines the total number of TB parameters (on-site energies and hopping integrals) and the occurrence percentage of each inside a given segment. In Fig. 5.2, we present the scaling of each TB parameter occurrence percentage for all studied segments. Among other things, we observe: The occurrence percentage of t_{GA} is always equal to that of t_{AG}. In all deterministic aperiodic cases, the occurrence percentages reach specific values as the generation, g, increases. Comparing *F* and *PD* sequences, although the former sequence is simpler (cf. Fig. 5.1), it has the same total number of TB parameters with the latter, since it has the additional triplet GGG. Again, we notice that, for sequences with primitive substitution rules, the values at which the occurrence percentage of each on-site energy and hopping integral converge coincide with the letter frequencies of the possible one- and two-letter words in the sequence, which can be found from the procedure described in Subsec. 2.4.1.

Having obtained an estimate of the intricacy of the sequences, we move to the estimation of the correlations of their energy landscape. We will do this by calculating the autocorrelation function (ACF) [97] for the quantities $\frac{E_n}{t_n}, n = 1, \ldots, N$. This ratio is used to fully capture the energy intricacy of the sequences. The degree the base pairs are correlated with their jth order neighbors, i.e., the correlation between sites with indices n and $n + j$ for fixed j, is expressed by the lag-j normalized ACF, $ACF(j)$. Using the notation $y_k = \frac{E_k}{t_k}$, with $\bar{y}$ being the mean value of $y_{\{n\}}$, it is given by

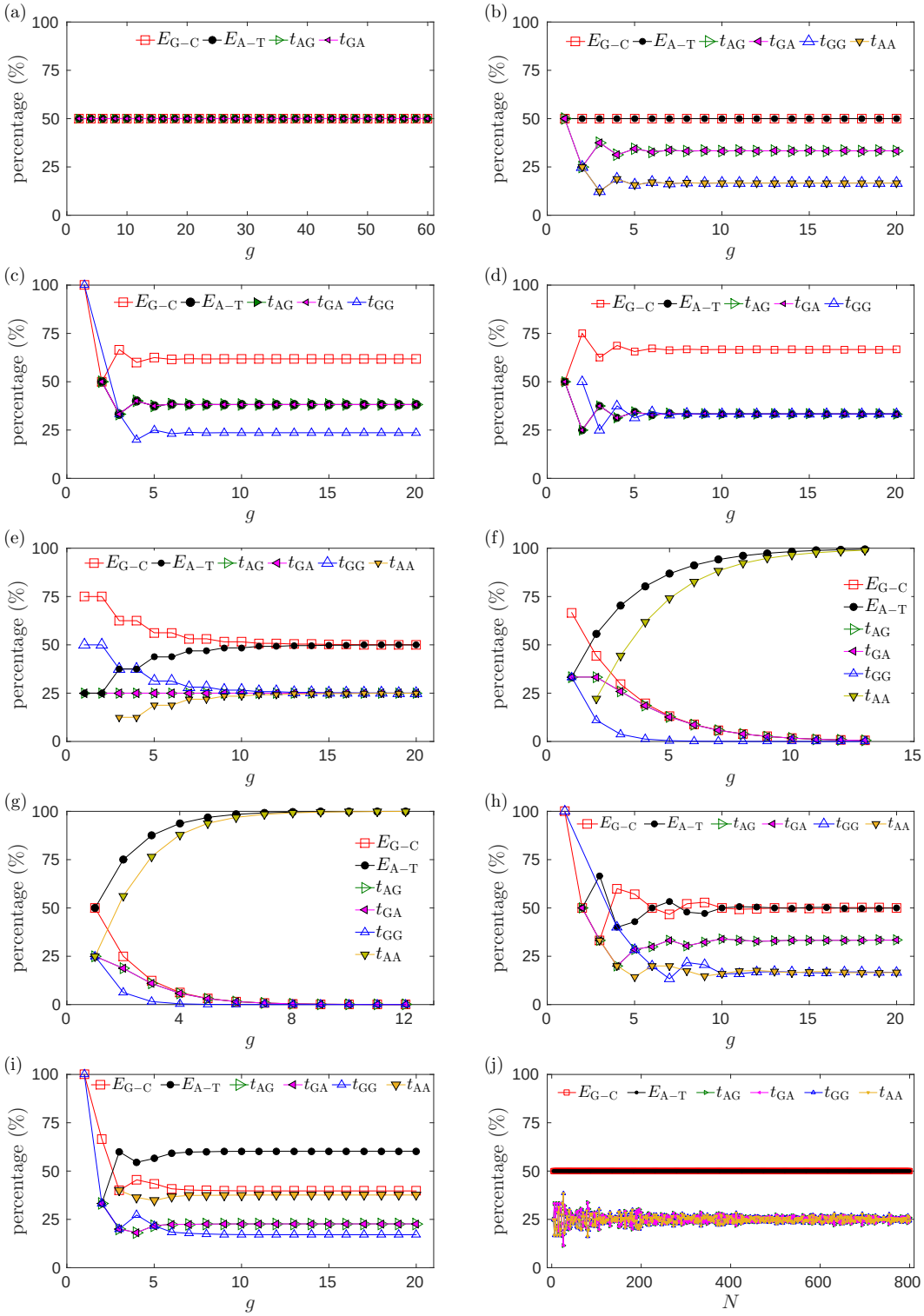

Figure 5.2: Scaling of the occurrence percentage of each TB parameter in various categories of DNA segments. (a) Periodic $(GA)_m$. (b) TM_g. (c) F_g. (d) PD_g. (e) RS_g. (f) CS_g. (g) GCS_g. (h) $KOL_g(1,2)$. (i) $KOL_g(1,3)$. (j) Random (50% G, 50% A).

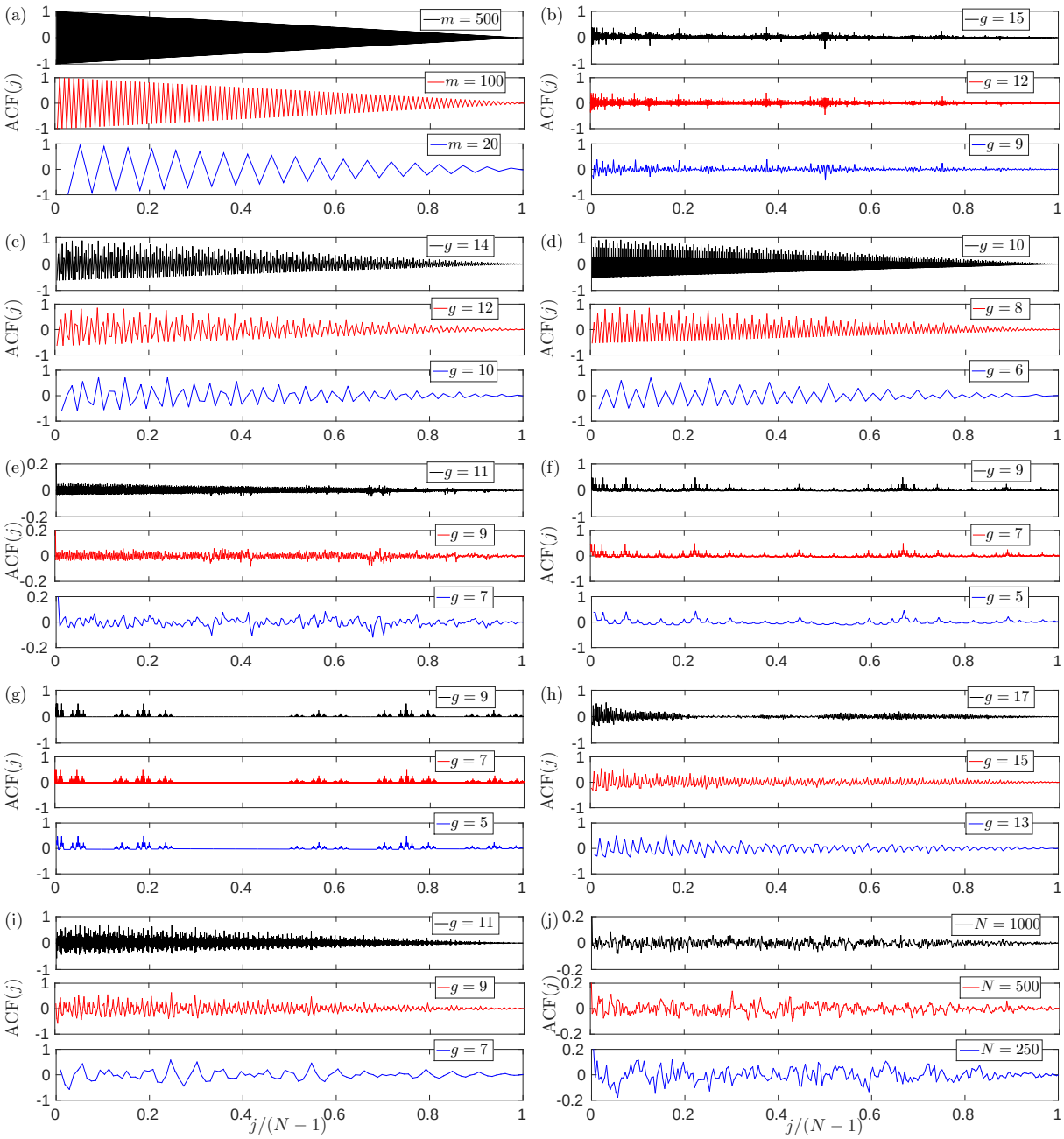

Figure 5.3: Scaling of the autocorrelation function of various categories of DNA segments. (a) Periodic $(GA)_m$. (b) TM_g. (c) F_g. (d) PD_g (e) RS_g. (f) CS_g. (j) $GCS_g(4,2)$. (h) $KOL_g(1,2)$. (i) $KOL_g(1,3)$. (j) Random (50% G content, 50% A content).

$$ACF(j) = \frac{\sum_{k=1}^{N-j}(y_k - \bar{y})(y_{j+k} - \bar{y})}{\sum_{k=1}^{N}(y_k - \bar{y})^2}. \tag{5.1}$$

In Fig. 5.3, we present the ACF all the categories of studied segments, for three

different lengths for each. The horizontal axes are normalized over the total number of neighbors $(N-1)$, thus corresponding to the *relative* neighbor distances. We notice that the ACF of each category has a characteristic shape. Furthermore, from the inspection of Fig. 5.3, we observe that there is a correspondence between the degree of intricacy of the segments and the strength of correlations. Random and *RS* sequences, which posses 8 equidistributed triplets, display weak correlations. *KOL*$(1,2)$ and *TM* sequences, which posses 6 equidistributed triplets, display somehow stronger correlations. Then follow *KOL*$(1,3)$, *CS*, and *GCS*$(4,2)$ sequences, which posses predominant triplets. The fractal sequences of the Cantor Set family possess strong correlations in the regions where G is present, interrupted by long, largely homogeneous, regions where it is not present. Deterministic aperiodic segments with the least possible triplets (*F* and *PD*, with 4 and 5 triplets, respectively) display strong correlations, and the periodic case is the dominant one.

Finally, we mention that by comparing the ACF of each category for different N, we can come to conclusions about their inflation/deflation symmetry. Sequences with this symmetry have similar autocorrelations at similar relative neighbor distances. This is the case for all studied aperiodic sequences, apart from *KOL*$(1,2)$, for which a substitution matrix cannot be defined, and the random ones [cf. Fig. 5.3(h) and (j), respectively]. As far as the *KOL*(p,q) family segments are concerned, we have checked no inflation or deflation symmetry exists when $|p-q| = 2\nu+1$, $\nu \in \mathcal{N}$, in contrast with the cases $|p-q| = 2\nu$, such as *KOL*$(1,3)$, shown in Fig. 5.3(i).

5.2 Eigenspectra and density of states

The energy structure of a physical system is closely connected to many of its properties (electrical, magnetic, thermal, optical, et cetera). In this Section, we study the eigenspectra, DOS, and integrated density of states (IDOS) of the above mentioned categories of DNA segments. The DOS shows the number of states that can be occupied by electrons at each energy, and can be formally defined as

$$g(E) = \sum_k \delta(E - E_k), \tag{5.2}$$

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

$$IDOS(E) = \int_{-\infty}^{E} g(E')dE', \tag{5.3}$$

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 [237].

Furthermore, for primitive substitutions described by a Hamiltonian corresponding to the WM, the following gap-labeling theorem has been introduced by Bellissard et al. [238]:

> **Theorem 5.13 of Ref [238].** *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 $\vec{\mathbf{v}}_{PF}$ and $\vec{\mathbf{v}}_{PF,2}$ of the substitution matrices* 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. Subsec. 2.4.3). Specifically, the gaps can be labeled by the negative powers of λ_{PF} times integral linear combinations of the components of $\vec{\mathbf{v}}_{PF}$ and $\vec{\mathbf{v}}_{PF,2}$ that lie within the interval $[0, 1]$ [238, 239]. For example, in the case of Fibonacci sequences, from the diagonalization of S (cf. Table 2.2), we get $\lambda_{PF} = \phi$ and $\vec{\mathbf{v}}_{PF} = [\phi^{-1} \quad \phi^{-2}]^T$ (where ϕ is the golden ratio). Hence, the sequence consists of ≈61.8% G letters and ≈38.2% A letters [cf. the numerically obtained occurrence percentages of on-site energies in Fig. 5.2(c)]. The legal 2-letter words in the Fibonacci sequence are AG, GA, and GG (i.e., AA is forbidden), thus the induced 2-substitution reads (cf. Subsec. 2.4.3) s_2(GG) = (GA)(AG), s_2(GA) = (GA)(AG), s_2(AG) = (GG), leading to the induced substitution matrix

$$\mathbf{S}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}. \tag{5.4}$$

The Perron–Frobenius eigenvector (cf. Subsec. 2.4.2) of $\mathbf{S}_2$ is $\vec{\mathbf{v}}_{PF,2} = [\phi^{-3} \quad \phi^{-2} \quad \phi^{-2}]^T$ [cf. the numerically obtained occurrence percentages of hopping integrals energies in Fig. 5.2(c)] . 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 by given by

$$\{\mathcal{G}_n\} = \{n\phi \mod 1, \forall n \in \mathbb{Z}\}. \tag{5.5}$$

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 [240].

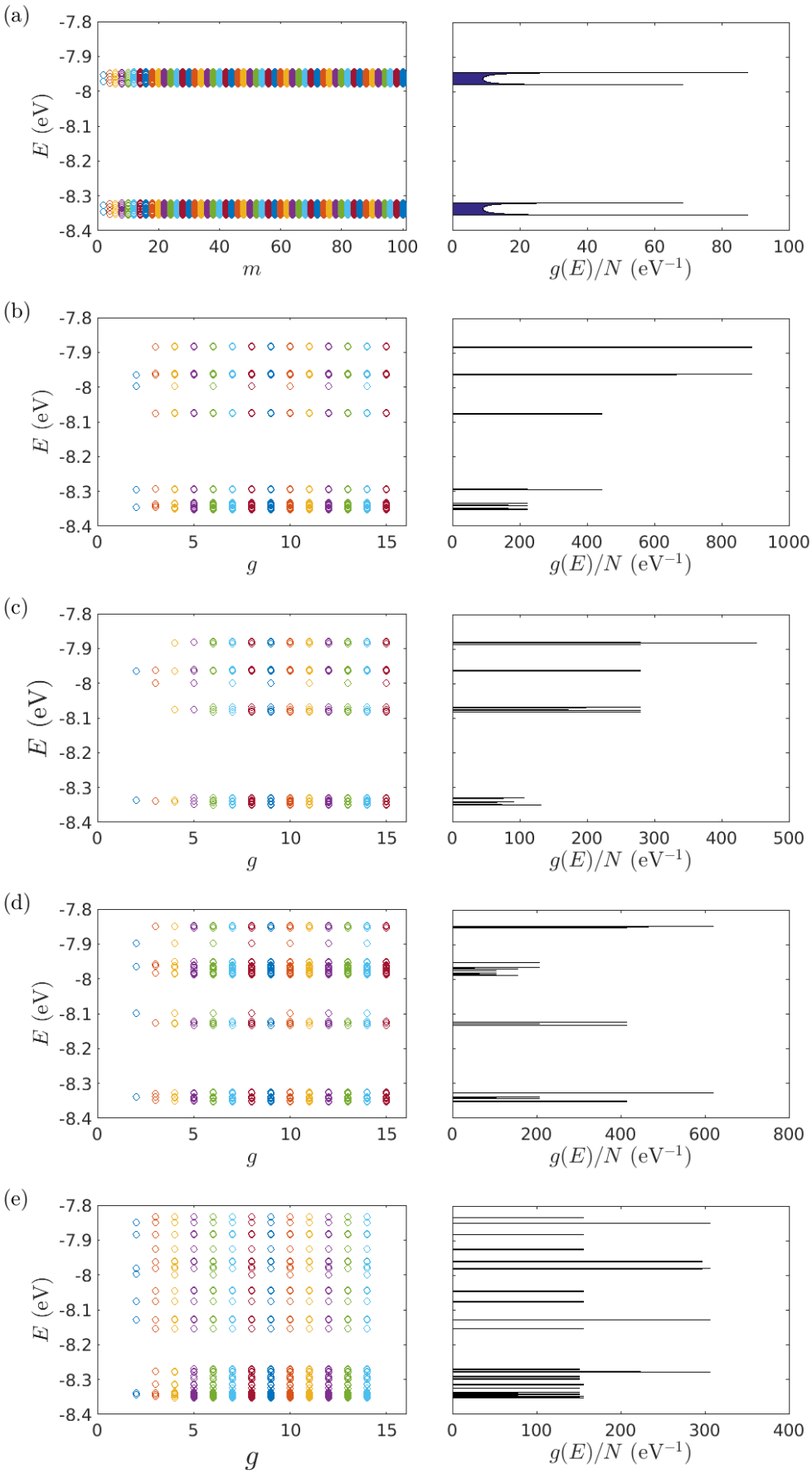

Figure 5.4: Eigenspectra and DOS of periodic and various deterministic DNA sequences. (a) Periodic $(GA)_m$. (b) TM_g. (c) F_g. (d) PD_g. (e) RS_g.

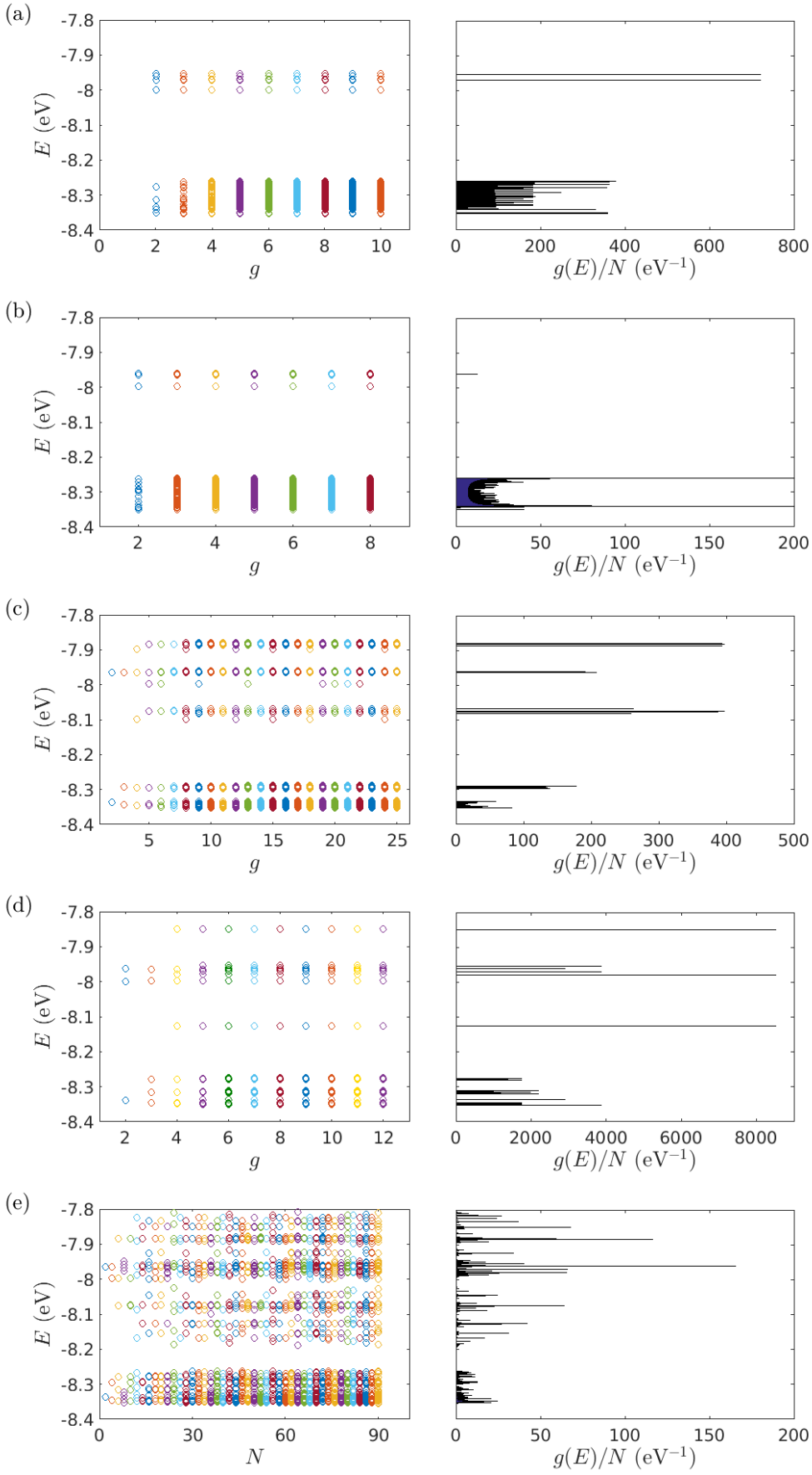

Figure 5.5: Eigenspectra and DOS of various fractal, Kolakoski and random DNA sequences. (a) CS_g. (b) $GCS_g(4,2)$. (c) $KOL_g(1,2)$. (d) $KOL_g(1,3)$. (e) Random (50% G and 50% A).

For fixed boundary conditions ($\psi_{N+1} = \psi_0 = 0$), the eigenspectrum, i.e., the eigenenergies E_j, $j = 1, 2, \ldots, N$ of a sequence, can be given by the roots of the polynomial $M_N^{11}(E)$ (see Chapter 2). Here, the eigenspectra of the sequences have been calculated by numerical diagonalization of the Hamiltonian matrix, which is real, tridiagonal and symmetric. For periodic sequences with a repetition unit composed of u sites, i.e., for $N = mu$, the matrix is u-Toeplitz, i.e., its elements have the property $h_{i,j} = h_{i+u,j+u}$. The eigenspectra of such sequences can alternatively be obtained recursively with the help of the Chebyshev polynomials of the second kind (see Chapter 3). The eigenspectra and the corresponding DOS for all the categories of DNA segments studied in this work are presented in Figs. 5.4-5.5. We notice that for all studied deterministic aperiodic sequences, the allowed energies do not exceed the energy interval defined by the eigenspectrum of the random sequence. This also holds for periodic polymers with only G and A in the $5'$-$3'$ strand, as their repetition unit increases [59] (see Chapter 8). Hence, the above mentioned interval of the random sequence represents a limit. Two subsets of the aforementioned interval gather around the on-site energies of G and A, so will be henceforth referred to as G and A energy regions. Comparing Fig. 5.4 which shows periodic and various deterministic aperiodic sequences with Fig. 5.5 which shows fractal, Kolakoski and random sequences, we observe that the former form subbands which are rather acute in the deterministic aperiodic cases, while in the latter the DOS is more fragmented and spiky. Regarding the random case, the plots correspond to a single realization of the lattice; thus, several DOS peaks are likely related to Azbel resonances, most of which would fade away by averaging over many realizations.

The normalized IDOS for all categories of DNA segments, for large N, is presented in Fig. 5.6. In each panel, the largest energy gap, i.e., the region between two consecutive discontinuities of the IDOS, corresponds to the separation between the upper limit of the allowed energies in the A region and the lower limit of the allowed energies in the G region. The value of the normalized IDOS in this gap corresponds to the relative number of A in the sequence. Periodic $(\mathrm{GA})_m$ segments possess two narrow, well defined, continuous bands, which can be recursively obtained; also, an analytical expression for the DOS exists [Eq. (3.33)]. *TM*, *F*, *PD*, *RS*, and *KOL* family sequences posses step-like IDOS, which indicates that the eigenenergies concentrate at specific energy regimes, separated by small gaps. *TM* and *KOL*$(1, 2)$ sequences resembling IDOS features, in accordance with their similar triplet number and occurrence percentages. Cantor set family sequences have allowed energies predominantly in the A region. Although at fist glance, the IDOS in this region may seem rather homogeneous, it can be seen from the insets of Fig. 5.6(f)-(g), that the spectrum is very rough. The random sequence IDOS has a shape that resembles to that of the *RS* sequence, although it is much more disrupted. We have also observed that all periodic and deterministic aperiodic segments possess IDOS steps such that their relative value is equal to the occurrence percentages of the possible base-pair triplets (cf. Fig. 5.1). These steps and the corresponding relative IDOS values are marked in the corresponding panels of Fig. 5.6 (except for the fractal segments in which the non-AAA triplets have very small occurrence percentages and cannot be depicted). For example, in the *F* segments there are four clear IDOS steps with relative heights $\phi^{-2}, \phi^{-3}, \phi^{-4}, \phi^{-3}$, respectively, where

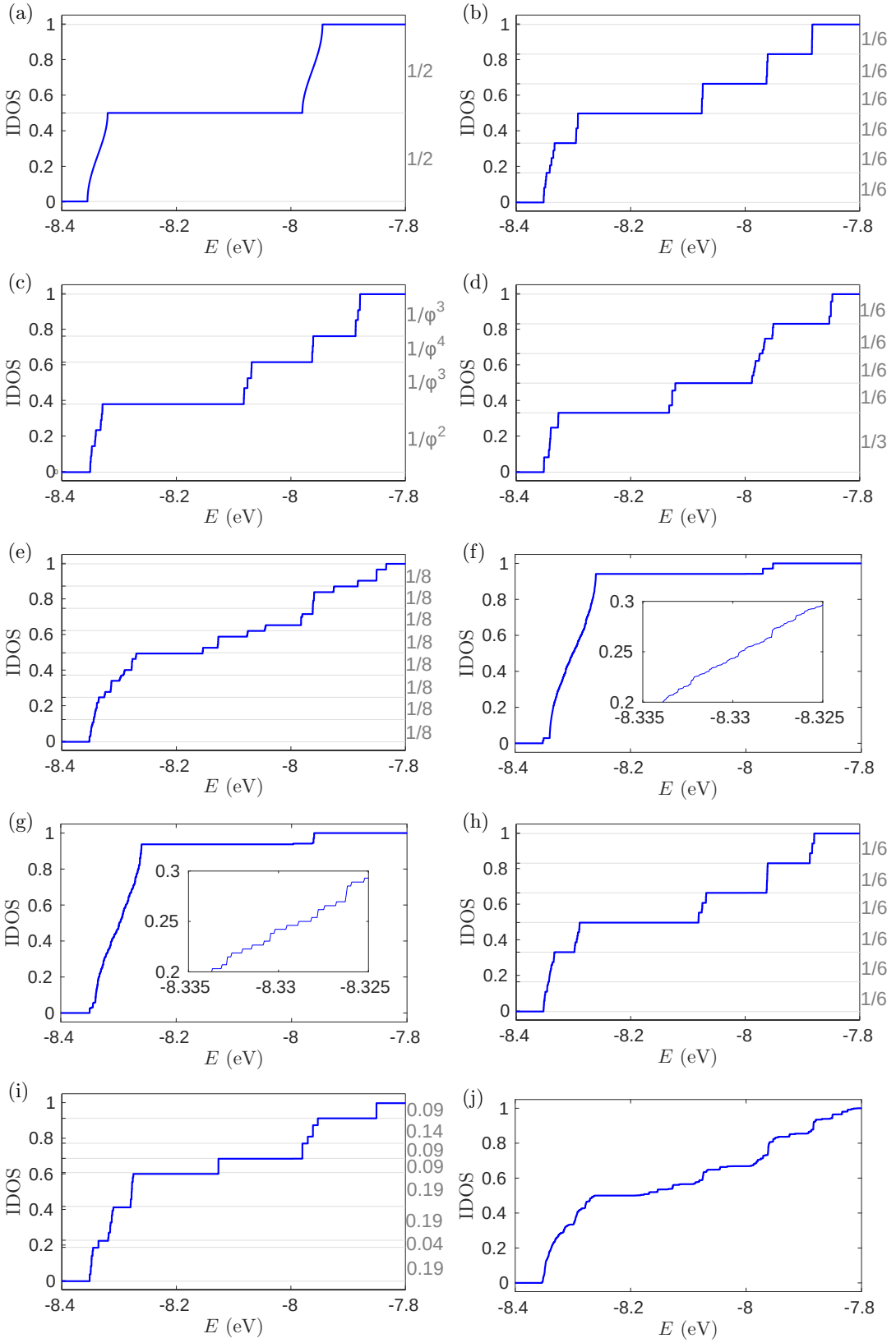

Figure 5.6: Normalized IDOS of various categories of DNA segments. (a) Periodic $(GA)_m$. (b) *TM*. (c) *F* (ϕ is the golden ratio.). (d) *PD*. (e) *RS*. (f) *CS*. (g) *GCS*. (h) $KOL(1,2)$. (i) $KOL(1,3)$. (j) Random (50% G, 50% A).

ϕ is the golden ratio; this has also been reported before [241]. Our observation holds for all categories of deterministic aperiodic sequences, either generated by a primitive substitution matrix or not, such as $KOL(1,2)$, further connecting the specific base-pair sequence of a DNA segment with its energy structure.

5.3 Localization

For the GTM of a given segment, $\boldsymbol{M}_N(E)$, there exists a limiting matrix $\boldsymbol{L}(E)$ such that

$$\boldsymbol{L}(E) = \lim_{N\to\infty} [\boldsymbol{M}_N(E)^T \boldsymbol{M}_N(E)]^{\frac{1}{2N}}. \tag{5.6}$$

The existence of $\boldsymbol{L}(E)$ is guaranteed by the Oseledec multiplicative ergodic theorem [242]. The Lyapunov Exponents of the segment are connected with the νth eigenvalue of $\boldsymbol{L}(E)$, $L_\nu(E)$, through

$$\gamma_\nu(E) = \ln[L_\nu(E)]. \tag{5.7}$$

If the GTM is a $2d \times 2d$ symplectic matrix, as in our case ($d = 1$), the Lyapunov exponents are distinct and have the property $-\gamma_1 < -\gamma_2 < \cdots < -\gamma_d < \gamma_d < \cdots < \gamma_2 < \gamma_1$, hence $\sum_{\nu=1}^{2d} \gamma_\nu = 0$[243, 244]. Since the Lyapunov exponents control the growth/decay rate of the solutions of Eq. (2.24), they are associated with the system's inverse localization length. In the case of symplectic GTMs, the localization length is given by the inverse of the smallest positive Lyapunov exponent, $\gamma_d(E)$ [244].

Since we deal with finite segments, the numerical Lyapunov exponents presented below correspond to finite values of N, hence the limit is dropped. To avoid numerical overflows when the matrix product is constructed, we use a QR decomposition scheme: We start with the initial matrix $\boldsymbol{M}_N(E)^T\boldsymbol{M}_N(E) = \boldsymbol{P}_1^T\boldsymbol{P}_2^T\ldots\boldsymbol{P}_N^T\boldsymbol{P}_N\ldots\boldsymbol{P}_2\boldsymbol{P}_1$. We perform a QR decomposition of $\boldsymbol{P}_1$, i.e. $\boldsymbol{P}_1 = \boldsymbol{Q}_1^{(1)}\boldsymbol{R}_1^{(1)}$, so that $\boldsymbol{M}_N(E)^T\boldsymbol{M}_N(E) = \boldsymbol{P}_1^T\boldsymbol{P}_2^T\ldots\boldsymbol{P}_N^T\boldsymbol{P}_N\ldots(\boldsymbol{P}_2\boldsymbol{Q}_1^{(1)})\boldsymbol{R}_1^{(1)}$. By consecutively performing QR decompositions at $\boldsymbol{P}_j\boldsymbol{Q}_{j-1}^{(1)}$, we arrive at $\boldsymbol{M}_N(E)^T\boldsymbol{M}_N(E) = \boldsymbol{Q}_{2N}^{(1)}\prod_{j=2N}^{1}\boldsymbol{R}_j^{(1)} := \boldsymbol{Q}^{(1)}\boldsymbol{R}^{(1)}$. Hence, the matrix $\boldsymbol{R}^{(1)}\boldsymbol{Q}^{(1)}$ and the initial matrix are similar, i.e., they have the same eigenvalues. By iterating this procedure, we arrive at a form $\boldsymbol{R}^{(k)}\boldsymbol{Q}^{(k)}$, where $\boldsymbol{Q}^{(k)}$ converges to a unit matrix and $\boldsymbol{R}^{(k)} = \prod_{j=2N}^{1}\boldsymbol{R}_j^{(k)}$, i.e., a product of upper triangular matrices with positive diagonal entries in descending order. Hence, the eigenvalue $L_\nu(E)$ is given by the $\frac{1}{2N}$th power of the diagonal element of $\boldsymbol{R}^{(k)}$, $R^{(k)\nu\nu}$. The Lyapunov exponents are thus

$$\gamma_\nu(E) = \frac{1}{2N}\sum_{j=1}^{2N}\ln[R_j^{(k)\nu\nu}]. \tag{5.8}$$

In our case, where $d = 1$, the only exponent to be determined is $\gamma_1(E)$. The index 1 will be dropped below.

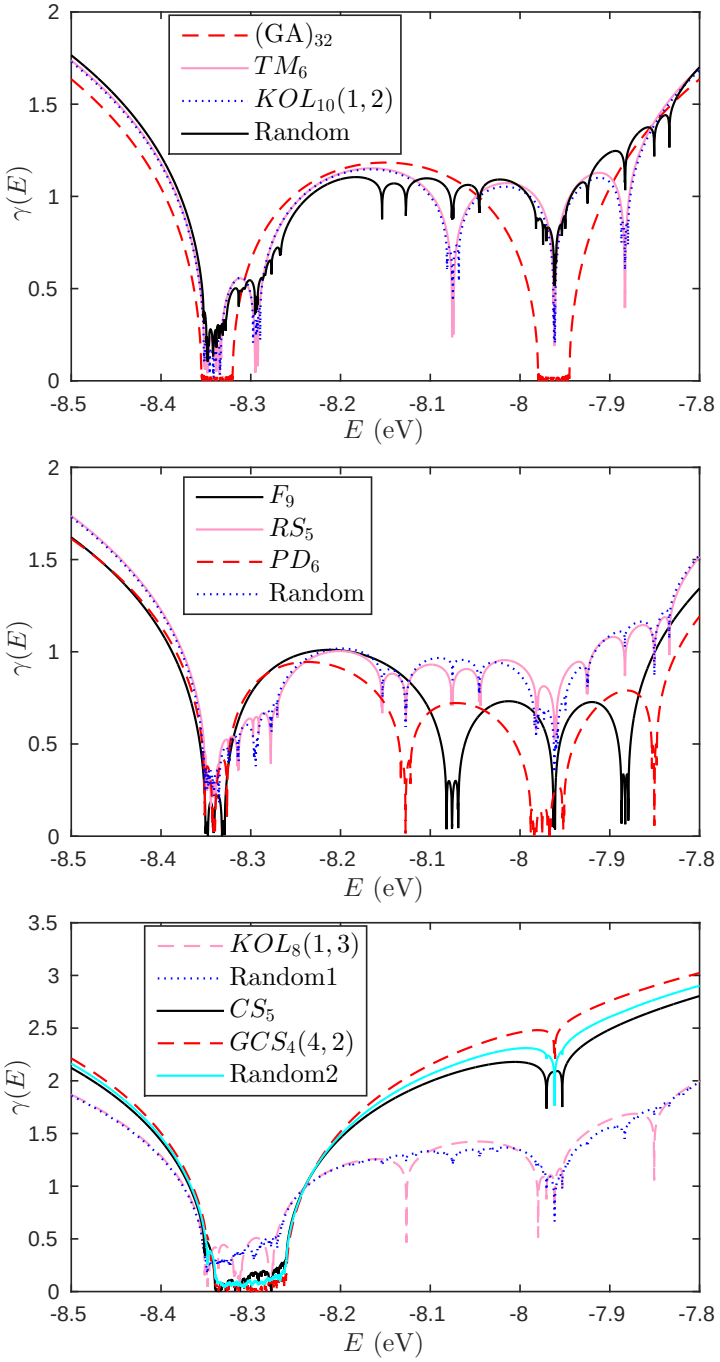

Figure 5.7: Lyapunov exponents of various categories of DNA segments. (Top) Segments with 50% G content: periodic $(GA)_m$ (red-dashed), *TM* (pink-filled), *KOL* $(1, 2)$ (blue-dotted) and random (black-filled). (Middle) Segments with more than 50% G content: *F* (61.82%, black-filled), *RS* (56.25%, pink-filled), *PD* (67.19%, red-dashed) and random (56.25%, blue-dotted). (Bottom) Segments with less than 50% G content: *KOL*$(1, 3)$ (40.00%, pink-dashed), *CS* (13.17%, black-filled), *GCS*$(4, 2)$ (6.25%, red-dashed) and two random (40.00% blue-dotted, 10.00% cyan-filled).

The Lyapunov exponents of all categories of periodic and deterministic aperiodic DNA segments, for large N, are presented in Fig. 5.7, together with some sequences with randomly rearranged base pairs. We have grouped together the segments accord-

ing to the percentages of G and A they posses. Cases with similar G and A content are depicted in Fig. 5.7(a), with dominant G content in Fig. 5.7(b) and with dominant A content in Fig. 5.7(c). Segments grouped together have similar sizes where possible.

Starting with Fig. 5.7(a), we notice that the Lyapunov exponents follow the trend of the autocorrelation functions; stronger correlations lead generally to less localized states. Periodic $(GA)_m$ segments have vanishing exponents inside their bands; this is a signature of the Bloch character of the wavefunctions. *TM* and *KOL*$(1, 2)$ sequences have non-vanishing exponents of similar magnitude. This similarity is direct consequence of the similar base-pair triplet distribution those two categories possess (cf. Fig. 5.1). The random sequence has generally much more localized states. As a general remark, we notice that the Lyapunov exponents in the A energy region are rather smaller than the ones in the G energy region.

The conclusion that segments with stronger correlations possess less localized states is also evident from Fig. 5.7(b). Furthermore, the Lyapunov exponents of *F* and *PD* segments reach very small values in both base-pair energy regions, while those of *RS* and random segments do not. *F* (*PD*) segments posses larger energy intervals of less localized states in the A (G) region than *PD* (*F*), while for *RS* and random segments the exponents follow resembling trends. The dominance of smaller exponents in *PD* segments over *F* segments in the G region can be explained by the enhanced presence of t_{GG} (which is of large magnitude) in the former, induced by the occurrence of GGG triplets (cf. Fig. 5.1).

In segments with dominant A content, which are depicted in Fig. 5.7(c), the Lyapunov exponents in the A energy region are much smaller than those in the G region. *KOL*$(1, 3)$ segments posses less localized states than random ones with similar G content in their common allowed energy intervals. The more dominant A becomes, the less (more) localized are the states in the A (G) region; this is the case for segments *CS*, *GCS*$(4, 2)$ and random sequences with similar G content. In these cases, there are large A-rich regions within the segments, interrupted by Gs, which act like a disorder. The more homogeneous regions the segments possess, the less localized their eigenstates will be in the A energy region. Comparing these segments in Fig. 5.7(c), we can see that, generally, as the percentage of G decreases, the exponents become smaller in the A region; however, there are always energies at which the fractal sequences, which possess stronger correlations, are more delocalized than the random one. The very small percentage of G leads to highly localized states in the corresponding energy interval.

5.4 Transmission coefficient

We connect the segments under examination to semi-infinite homogeneous metallic leads, which act as carrier baths. The leads' energy spectrum is given by the dispersion relation $E = E_M + 2t_M \cos(qa)$, where E_M is the on-site energy of the leads and t_M is the hopping integral between the leads' sites. The coupling between the segment and the left (right) lead is described by the effective parameters $t_{L(R)}$. We choose $\epsilon_m =$

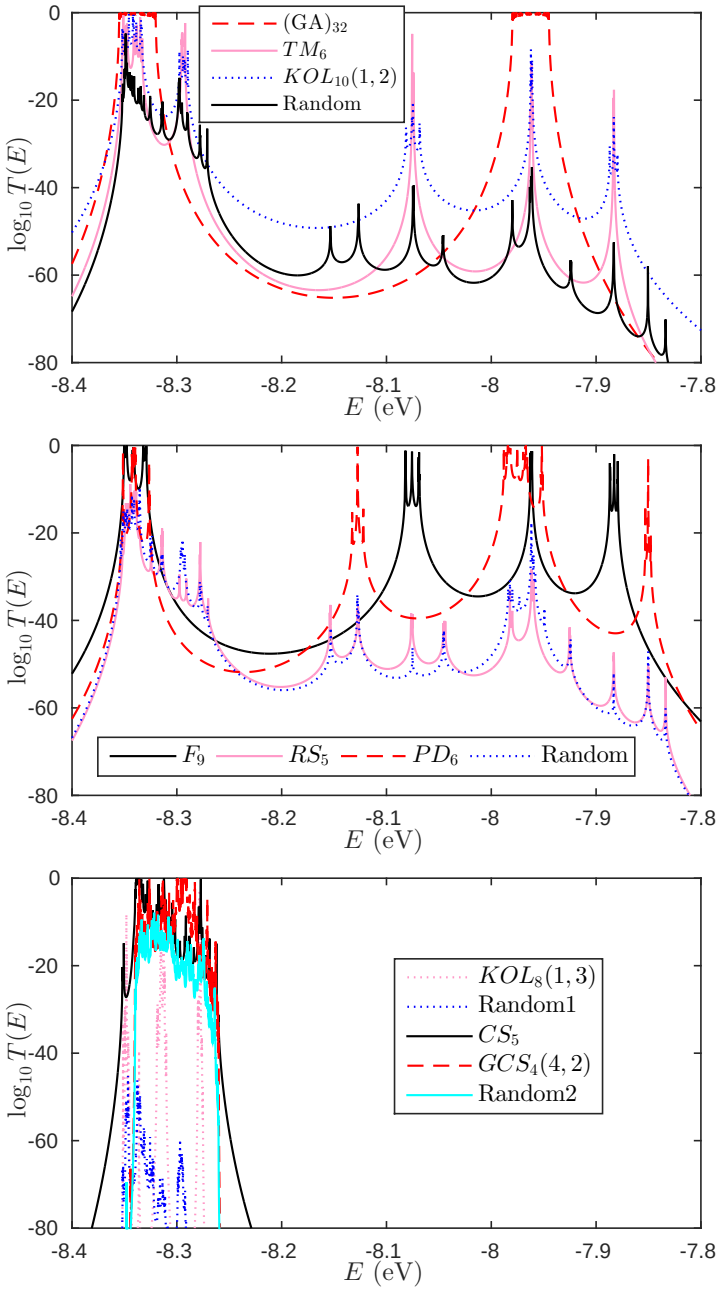

Figure 5.8: Transmission coefficients of various categories of DNA segments. (Top) Segments with 50% G content: periodic $(GA)_m$ (red/dashed), *TM* (pink-filled), *KOL* $(1, 2)$ (blue-dotted) and random (black-filled). (Middle) Segments with more than 50% G content: *F* (61.82%, black-filled), *RS* (56.25%, pink-filled), *PD* (67.19%, red-dashed) and random (56.25%, blue-dotted). (Bottom) Segments with less than 50% G content: *KOL*$(1, 3)$ (40.00%, pink-dotted), *CS* (13.17%, black-filled), *GCS*$(4, 2)$ (6.25%, red-dashed), and two random (40.00% blue- dotted, 10.00% cyan-filled).

$\frac{(\epsilon_{A-T}+\epsilon_{G-C})}{2} = -8.15$ eV and $t_M = -0.25$ eV, so that all eigenstates of the systems under examination are contained within the leads' bandwidth. We also choose the coupling parameters $t_{L(R)}$ to satisfy the ideal and symmetric coupling conditions, $|\omega| = |\chi| = 1$, which have been shown in Chapter 3 to be the optimal coupling conditions

for periodic segments.

In Fig. 5.8 we present the TC. At first glance, TC qualitatively follows the trend of the Lyapunov exponents (cf. Fig. 5.7). The less localized the eigenstates are, the more transparent the segments are to the incident waves at their energy region. Periodic $(GA)_m$ segments display the most enhanced transmission, and reach the full transmission condition at specific energies (see Subsec. 3.4.1); this does not hold in general for deterministic aperiodic and random segments. Furthermore, apart from periodic $(GA)_m$, *F*, and *PD* segments, transmission in the G energy region is from very small to negligible. These categories, together with the Cantor Set family ones, display the most enhanced transmission. *TM* and *KOL*$(1, 2)$ sequences display some energies at which transmission is rather significant. Deterministic aperiodic segments are more transparent than random ones with similar base-pair content, with the exception of *RS*, that generally follows the trend of its randomly redistributed counterpart. Finally, we notice that the sequences shown in Fig. 5.8(c) have negligible transmission in the G energy region. This is due to the small role t_{GG} plays, since it rarely occurs within the segments.

5.5 Current-Voltage Curves

We apply a symmetric constant bias voltage V_b between the leads, so that their chemical potential takes the form $\mu_{\substack{L\\R}} = E_M \pm \frac{V_b}{2}$. Then, a linear voltage drop within the DNA segment is induced and the transmission coefficient becomes bias-dependent. The energy regime between the leads' chemical potentials defines the conductance channel. The electrical current at zero temperature can be computed using the Landauer-Büttiker formalism [109, 224, 226] as

$$I(V) = \frac{2e}{h} \int_{E_M - \frac{V_b}{2}}^{E_M + \frac{V_b}{2}} T(E, V_b)\, dE, \tag{5.9}$$

since the Fermi-Dirac distributions, $f(E_M \pm \frac{V_b}{2})$, are Heaviside step-functions [cf. Eq. (4.10)]. The factor 2 in Eq. (5.9) comes from the double spin-degeneracy of each electronic level.

Again, we choose the coupling parameters to satisfy the ideal and symmetric coupling conditions, $|\omega| = |\chi| = 1$. We set the leads hopping integral $t_M = -0.5$ eV to ensure that the leads' bands are wide enough to capture the whole picture. The choice of the leads Fermi level, E_M, plays a major role in both the shape of the $I - V$ curves and the magnitude of the currents. This is demonstrated in Fig. 5.9, where the $I - V$ curve of a periodic $(GA)_{16}$ segment is determined as a function of E_M. It is evident that larger currents (~ 0.1 μA) occur at small biases when E_M lies within the bands of the segment. When this is not the case, voltage thresholds appear, and the (smaller in magnitude) turn-on currents emerge at biases that increase in a linear fashion with changing E_M. The magnitude of the currents becomes gradually smaller as E_M moves

further away from the segments' bands, and is negligible when E_M lies well outside the bands. Finally, we should mention that the $I-V$ curves are symmetric with respect to the difference between E_M and $\frac{(\epsilon_{A-T}+\epsilon_{G-C})}{2}$. The above mentioned conclusions hold also qualitatively for segments consisting of identical monomers with crosswise purines, such as $(GC)_m$, where only one on-site energy (ϵ_{G-C}) is involved, with the difference that the curves are symmetric with respect to the difference between E_M and ϵ_{G-C}.

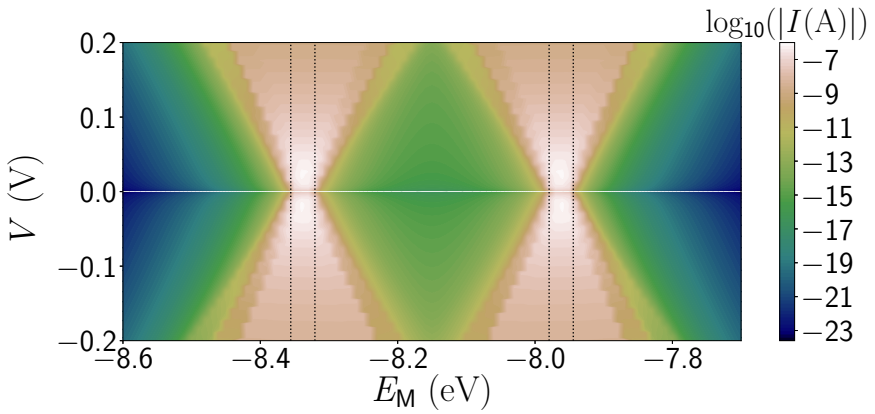

Figure 5.9: The role of the leads' Fermi level, E_M, to the $I-V$ curve of a $(GA)_{16}$ segment. The vertical dotted lines encompass the bands of the segment.

Given the previous discussion and Fig. 5.9, we chose to study the $I-V$ curves of all segments for two values of E_M, specifically -7.95 eV and -8.35 eV (i.e. at the center of the periodic segment's bands), to capture both G and A energy regions. In the following, we will only present curves the currents of which reach the pA regime. Our results are depicted in Figs. 5.10 and 5.11, for $E_M = -8.35$ eV and $E_M = -7.95$ eV, respectively.

From Fig 5.10(a), it is evident that periodic segments can carry significantly larger currents ($\sim 0.1\ \mu$A) than other categories. The deterministic aperiodic *TM* and *KOL*$(1,2)$ segments display quite smaller currents than the periodic ones, of similar magnitude ($\sim$ 1 nA), but with clearly distinct shapes. The similarity of current magnitudes between *TM* and *KOL*$(1,2)$ segments is in accordance with the similarity in the values of the Lyapunov exponents and zero-bias transmission coefficient for these cases, cf. Figs. 5.7(a) and 5.8(a), respectively. The random segment displays significantly smaller currents compared to the rest categories, reaching ~ 10 pA.

As far as segments with dominant G content are concerned, we can see in Fig. 5.10(b) that *F* and *PD* segments can carry significantly larger currents than the *RS* and random ones. This is again in accordance with the magnitude of the Lyapunov exponents and the transmission coefficients for these cases, cf. Figs. 5.7(b) and 5.8(b). In the A energy region, there is a larger energy range in which *F* segments display less localized states and higher transmission than *PD* ones. This is fact is reflected on the magnitude of the currents ($\sim$ 1 nA for *F*, ~ 0.1 nA for *PD*). *RS* and random segments display currents in the ~ 10 pA regime, but their curves have different shapes.

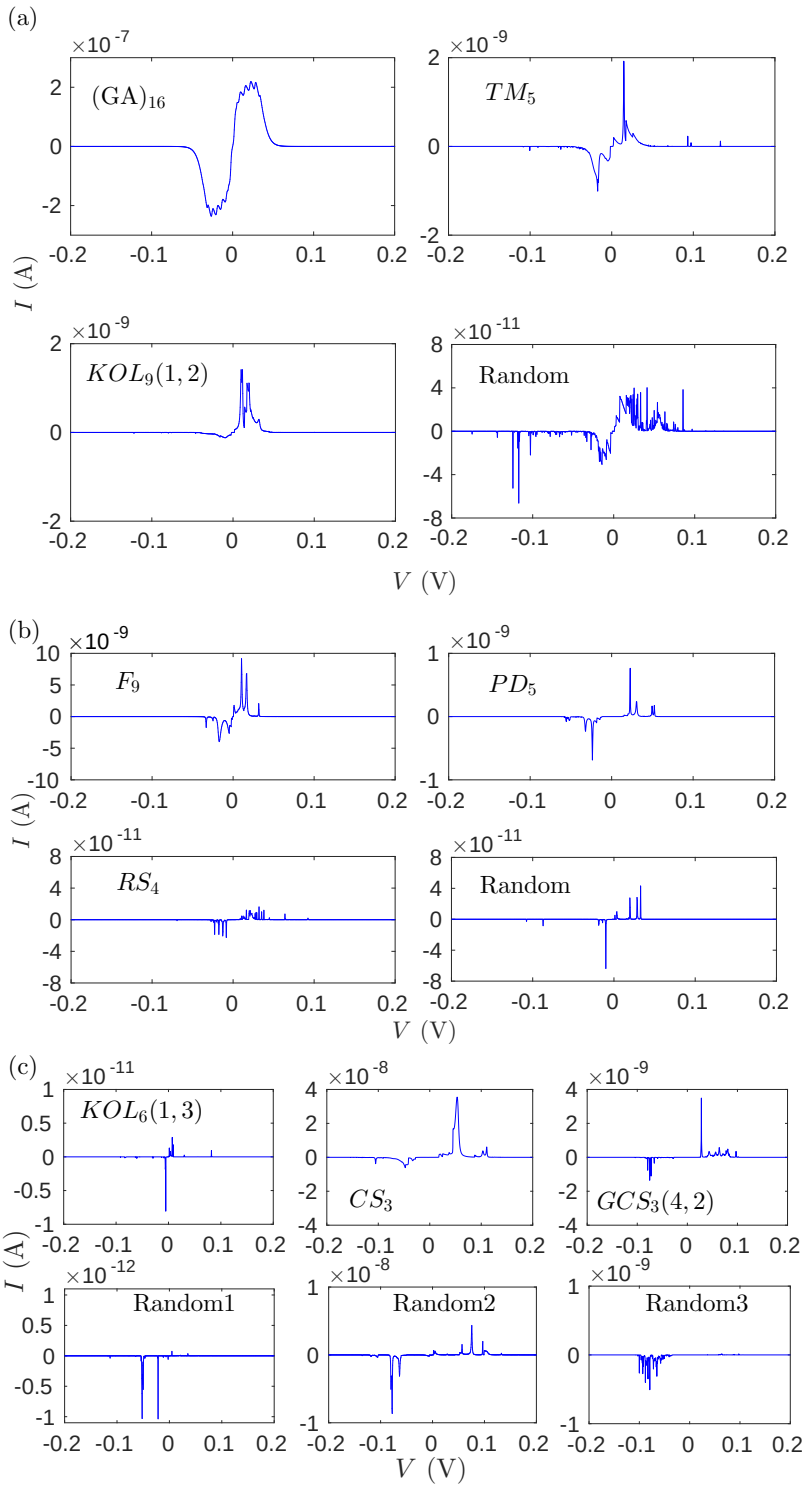

Figure 5.10: $I-V$ curves of various categories of DNA segments for $E_M = -8.35$ eV. Categories as in Figs. 5.7 and 5.8. (a) Periodic $(GA)_m$, *TM*, *KOL* $(1,2)$ segments and a random segment with similar G content. (b) *F*, *PD*, *RS* segments, and a random segment with similar G content. (c) (top) *KOL*$(1,3)$, *CS*, *GCS*$(4,2)$ segments. (Bottom) Random rearrangements of *KOL*$(1,3)$, *CS*, *GCS*$(4,2)$ segments, respectively.

Sequences with dominant A content are depicted in Fig. 5.10(c). $KOL(1,3)$ sequences display rather small currents, that hardly reach 10 pA, due to the fact that the hopping integral with the largest occurrence percentage, i.e. t_{AA}, is of rather small value. Albeit their small magnitude, the currents of $KOL(1,3)$ sequences are larger than of their random rearrangement, which hardly reach 1 pA. In Cantor set family sequences, A content is much larger than G content, leading to large parts of the segment being essentially homogeneous. Hence, although t_{AA} has a small value, rather large currents occur (~ 10 nA for *CS*, ~ 1 nA for $GCS(4,2)$). In this class of sequences, G, which, due to its small presence acts as a disorder in an otherwise homogeneous segment, is gathered in specific regions. Therefore, the currents they display are about one order of magnitude larger than their random rearrangements (~ 1 nA and ~ 10 nA, respectively).

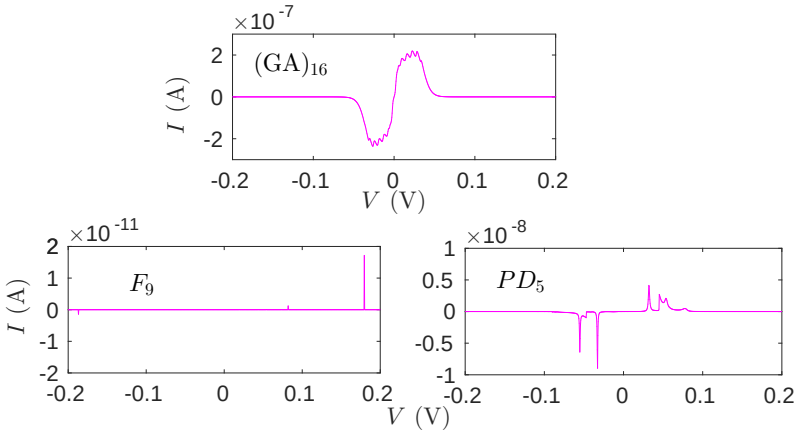

Figure 5.11: $I-V$ curves of various categories of DNA segments for $E_M = -7.95$ eV.

As discussed in previous Sections, in the G energy region the eigenstates of most segment categories are highly localized and display very small or negligible transmission. This, for $E_M = -7.95$ eV, leads to currents that lie well below the pA regime. The only cases that do not follow this trend are the periodic, *F*, and *PD* segments, the $I-V$ curves of which are depicted in Fig. 5.11. The periodic segments curve in this case is identical to the one for $E_M = -8.35$ eV, due to the symmetry of the $I-V$ curves with respect to the difference between E_M and $\frac{(\epsilon_{A-T}+\epsilon_{G-C})}{2}$, cf. Fig. 5.9. The rest two cases display energy intervals in the G region for which less localized states and enhanced transmission occur, as shown in previous Sections. Close to E_M, the interval for *F* segments is much smaller than the one for *PD* segments, leading to a great difference in the current magnitudes between the two cases: a single spike of ~ 100 pA for *F* segments, currents in the ~ 10 nA regime for *PD* segments. This is due to the presence of GGG triplets in *PD* segments, which leads to enhanced presence of t_{GG} (the magnitude of which is large), compared to *F* segments, cf. Fig 5.2(c)-(d).

5.6 Effect of parameters

It is common in the literature that all hopping parameters between different moieties are considered equal, for simplicity. Let us provide some example results occurring for identical hopping parameters, with reference to the Lyapunov exponents: In this case, *F* segments posses more delocalized states in the G region (results not presented here), in contrast with the discussion of Fig. 5.7(b). Additionally, for all studied sequences, if we take equal hopping parameters, the act of substituting G with A and vice versa leads to a mere reflection of $\gamma(E)$ relative to the mean value of the on-site energies, $\frac{(\epsilon_{A-T}+\epsilon_{G-C})}{2}$ (results not presented here). This is not the case when different hopping parameters are considered. Their relative presence and magnitude can lead to significant differences in the electronic properties. Another example is the *TM* sequence. If we equalize all hopping parameters, the Lyapunov exponent is also symmetric relative to $\frac{(\epsilon_{A-T}+\epsilon_{G-C})}{2}$ (results not presented here), a scenario that does not hold for different hopping parameters, cf. Fig. 5.7(a). Of course, the inclusion of different hopping parameters plays significant role not only in the Lyapunov exponents, but also in all properties that are determined by the electronic structure, such as the transmission coefficient and the $I-V$ curves. To conclude, besides the fact that, in terms of chemical complexity, taking identical hopping parameters is unrealistic, our treatment reveals that considering different hopping parameters leads to a better understanding of the interplay between sequence intricacy and transport properties, both quantitatively and qualitatively.

Furthermore, as far as transport properties are concerned, different results occur for different parameter values. For example, we have been able to reproduce the results reported for the transmission coefficients in Refs. [73, 180, 245], and for the $I-V$ curves in Ref. [180], using the corresponding parametrizations, which are different from the S parametrization used here (all with equal hopping integrals). Different shapes as well as current-voltage regimes can be obtained, if the parameters are modified. For example, in Ref. [13] where microRNA chains are studied, taking different hopping integrals between nucleotides but of significantly larger magnitude than the ones used here, the authors report currents in the nA regime for voltages up to 16 V. These curves have been reproduced as well. The difference in the current-voltage regimes can also be seen be comparing the $I-V$ curves of the homogeneous $(\mathrm{G})_m$ and $(\mathrm{A})_m$ segments (Fig.5.12), which, due to their sequential simplicity, represent the most efficient cases for charge transport. The curves have been calculated for $E_M = \epsilon_{G-C}$ (ϵ_{A-T}) for the former (latter) case, i.e., in the center of the bands, with $t_M = -0.5$ eV, and ideal and symmetric coupling conditions. Since the leads are aligned with the band centers, the only defining factor of the current-voltage regime is the value of the hopping parameter t_{GG} (t_{AA}). Since $t_{GG} > t_{AA}$, $(\mathrm{G})_m$ segments display greater currents than $(\mathrm{A})_m$ segments (~ 10 μA vs. ~ 1 μA) and lie in a larger bias regime. Generally, increasing the value of the hopping parameter results in increase of both the current magnitude and the voltage regime, until the states of the segment reach the bandwidth of the leads. For both $I-V$ curves, the conductance at zero bias is equal to the quantum of conductance, i.e., $\frac{\partial I}{\partial V}\big|_{V=0} = G_0 = \frac{2e^2}{h} \approx 7.748\times10^{-5}$ S.

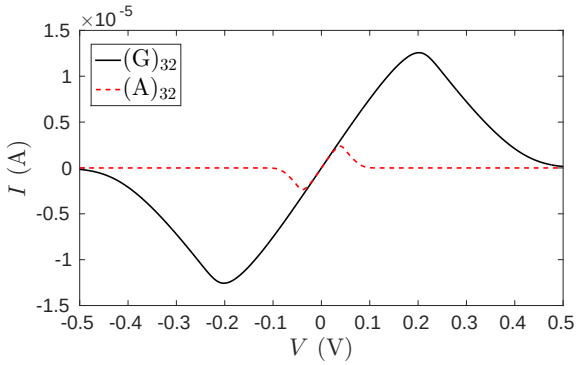

Figure 5.12: $I-V$ curves of $(G)_{32}$ and $(A)_{32}$ segments.

As discussed in Sec. 5.5 (cf. Fig. 5.9), the occurrence of voltage gaps in the $I-V$ curves depends on the relative position of the Fermi level of the leads and the eigenenergies of the segments. For example, a typical semiconducting I-V curve occurs for $(G)_{30}$ segments, if we set $E_M - \epsilon_{G-C} = 0.3$ eV (i.e. for E_M lying outside the band of the segment), with a voltage gap of ≈ 0.7 V and currents ~ 1 nA. This is in accordance with the experimental I-V curves reported for the same system in Ref. [20], where the authors also attribute the voltage gap to the offset between the Fermi level of the electrode and the energy levels of the $(G)_{30}$ segment.

5.7 Conclusion

We comparatively studied periodic, deterministic aperiodic, and random binary DNA sequences within the framework of the WM. We used B-DNA and as a prototype system and the binary alphabet {G,A}. All segments had their purines on the same strand. We gained a better understanding of the interplay between the intricacy of the segments and their spectral, localization and charge transport properties. We took differences in hopping parameters between successive monomers into account. This led to a more realistic evaluation of the role the sequence intricacy plays in the aforementioned properties.

We determined the number and occurrence percentage of all possible base-pair triplets that can be found within these segments, as well as their autocorrelation functions. Our results showed that there is a relation between the number of possible triplets, the existence of dominant triplets and the strength of correlations.

We calculated the eigenenergies, the DOS, and the IDOS. The allowed eigenenergies of all studied deterministic aperiodic segments lie within the interval defined by the eigenspectrum of random sequences. In all deterministic aperiodic segments, either if a substitution matrix can be defined or not, there exist energy steps in the relative normalized IDOS, equal to the occurrence percentages of the possible monomer triplets. This observation establishes a clear relation between the sequence intricacy and the spectral properties.

Furthermore, we calculated the Lyapunov exponents and showed that the sequence

intricacy, the relative presence of each monomer, and the values of the TB parameters play major role in the degree of eigenstates localization. Generally, sequences with strong correlations posses less localized states.

Next, we connected the segments to semi-infinite homogeneous leads and studied the zero-bias TC, reaching similar conclusions regarding their transparency to incident carriers.

We also studied the $I - V$ characteristics of the segments, using the Landauer-Büttiker formalism. We showed that the shape of the curves and the magnitude of the currents strongly depends on the leads' on-site energy (Fermi level). The current-voltage characteristics were calculated for two values of the latter, corresponding to positions that catch the energy regions of interest. For the parametrization used, we found that periodic binary segments can carry currents in the μA regime. Several deterministic aperiodic segments (specifically, Fibonacci, Period-doubling, Cantor and generalized Cantor) can also display rather large currents, namely in the nA regime, depending on the Fermi level of the leads. Random sequences hold the smallest currents, in accordance with the weak correlations they posses.

Finally, the $I - V$ curves of the homogeneous $(\mathrm{G})_m$ and $(\mathrm{A})_m$ segments, due to their sequential simplicity, represent the most efficient cases for charge transport with conductance at zero bias equal to the quantum of conductance. Typical semiconducting $I - V$ curves occur for these segments when there is a mismatch between their eigenstates and Fermi level of the leads, in accordance with experimental results.

Part III

Charge transfer

6
The time-dependent problem

In this Chapter, we briefly present the theoretical framework of unbiased coherent transfer of an extra charge carrier initially inserted (e.g. by oxidation or reduction for holes or electrons, respectively) in one-dimensional molecular wires modeled with TB, as well as the physical quantities studied in the rest of this part of the present PhD thesis, which are employed to study short DNA segments (monomers, dimers, and trimers), several classes of periodic DNA polymers, and deterministic aperiodic and random DNA polymers. We assume that an extra hole moves through HOMOs, while an extra electron moves through LUMOs.

To describe the spatiotemporal evolution of an extra carrier created at a particular site of the polymer, we consider the state of the polymer, $|\Psi_n(t)\rangle$, as a linear combination of the monomer's states with time dependent coefficients, i.e.,

$$|\Psi(t)\rangle = \sum_{n=1}^{N} |n\rangle \, \vec{\mathbf{\Psi}}_n(t), \tag{6.1}$$

where

$$\vec{\mathbf{\Psi}}_n(t) = \begin{pmatrix} \psi_n^1(t) \\ \psi_n^2(t) \\ \vdots \\ \psi_n^L(t) \end{pmatrix}. \tag{6.2}$$

$\psi_n^l(t)$ is the probability amplitude to find the carrier at the lth chain of the nth monomer at time t. Plugging Eq. (6.1) and the TB Hamiltonian of Eq. (2.16) into the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle , \tag{6.3}$$

we arrive at a system of $N \times L$ $(:= M)$ coupled differential matrix equations of the form

$$i\hbar \frac{\mathrm{d}\vec{\boldsymbol{\Psi}}_n(t)}{\mathrm{d}t} = \boldsymbol{\epsilon}_n \vec{\boldsymbol{\Psi}}_n + \boldsymbol{\tau}_{n-1}^T \vec{\boldsymbol{\Psi}}_{n-1} + \boldsymbol{\tau}_n \vec{\boldsymbol{\Psi}}_{n+1}, \tag{6.4}$$

i.e., the time-dependent TB system of equations. Eq. (6.4) is equivalent to a 1st order matrix differential equation of the form

$$\frac{\mathrm{d}\vec{\boldsymbol{\Psi}}(t)}{\mathrm{d}t} = -\frac{i}{\hbar} \boldsymbol{H} \vec{\boldsymbol{\Psi}}(t), \tag{6.5}$$

where $\vec{\boldsymbol{\Psi}}(t)$ is a vector matrix composed of the vector-coefficients $\vec{\boldsymbol{\Psi}}_n(t), n = 1, 2, \ldots, N$, or, alternatively, the coefficients $\psi_n^l(t) := \psi_\mu(t)$, $\mu = 1, 2, \ldots, M$, with the indices μ, n, l being connected via the relation

$$\mu = L(n-1) + l. \tag{6.6}$$

This indexing will be used in the following. Eq. (6.5) can be solved with the eigenvalue method, i.e., by looking for solutions of the form

$$\vec{\boldsymbol{\Psi}}(t) = \vec{\boldsymbol{\Psi}} e^{-\frac{i}{\hbar}Et} \Rightarrow \frac{\mathrm{d}\vec{\boldsymbol{\Psi}}(t)}{\mathrm{d}t} = -\frac{i}{\hbar} E \vec{\boldsymbol{\Psi}} e^{-\frac{i}{\hbar}Et}. \tag{6.7}$$

Hence, Eq. (6.5) leads to the eigenvalue problem of Eq. (2.26), that is, $\boldsymbol{H}\vec{\boldsymbol{\Psi}} = E\vec{\boldsymbol{\Psi}}$. Having determined the eigenvalues and eigenvectors of $\boldsymbol{H}$, the general solution of Eq. (6.5) is

$$\vec{\boldsymbol{\Psi}}(t) = \sum_{k=1}^{M} c_k \psi_k e^{-\frac{i}{\hbar}E_k t}. \tag{6.8}$$

In other words, the coefficients $\Psi_\mu(t), \mu = 1, 2, \ldots, M$ are given by a superposition of the time evolution of the stationary states with time-independent coefficients c_k. Hence, this is a coherent phenomenon. The coefficients c_k are determined from the initial conditions. In particular, if we define the $M \times M$ eigenvector matrix $\boldsymbol{V}$, with elements $\psi_{\mu k}$, then it can be shown that the vector matrix $\vec{\boldsymbol{c}}$, composed of the coefficients $c_k, \quad k = 1, 2, \ldots, M$, is given by the expression

$$\vec{\boldsymbol{c}} = \boldsymbol{V}^T \vec{\boldsymbol{\Psi}}(0). \tag{6.9}$$

Suppose that initially the extra carrier is placed at the λth site, i.e., $\Psi_\lambda(0) = 1$, $\Psi_\mu(0) = 0, \forall \mu \neq \lambda$. Then,

$$\vec{\boldsymbol{c}} = \begin{pmatrix} \psi_{\lambda 1} \\ \vdots \\ \psi_{\lambda k} \\ \vdots \\ \psi_{\lambda N} \end{pmatrix}. \tag{6.10}$$

In other words, the coefficients c_k are given by the row of the eigenvector matrix which corresponds to the monomer the carrier is initially placed at.

We also mention that changing the view of a polymer from one (e.g., top) to the other (e.g., bottom) side of the growth axis, reflects the Hamiltonian matrix $\boldsymbol{H}$ of the polymer on its main antidiagonal. This reflected Hamiltonian, $\boldsymbol{H}^{\text{equiv}}$, describes the *equivalent polymer*. $\boldsymbol{H}$ and $\boldsymbol{H}^{\text{equiv}}$ are connected by the similarity transformation $\boldsymbol{H}^{\text{equiv}} = \boldsymbol{J}^{-1}\boldsymbol{H}\boldsymbol{J}$, where $\boldsymbol{J}(= \boldsymbol{J}^{-1})$ is the unit antidiagonal matrix of order M. Therefore, $\boldsymbol{H}$ and $\boldsymbol{H}^{\text{equiv}}$ have identical eigenspectra (hence the equivalent polymers' DOS is identical) and their eigenvectors are connected by $\psi_{\mu k} = \psi^{\text{equiv}}_{(N-\mu+1)k}$. Generally,

$$\text{equiv(YX...Z)} = \text{Z}_{\text{compl}} \ldots \text{Y}_{\text{compl}}\text{X}_{\text{compl}}, \tag{6.11}$$

where, e.g., X_{compl} denotes the complementary base of X.

6.1 Physical Quantities

From Eq. (6.8), it follows that the probability to find the extra carrier at the μth monomer is

$$|\Psi_\mu(t)|^2 = \sum_{k=1}^{M} c_k^2 \psi_{\mu k}^2 + 2\sum_{k=1}^{M}\sum_{\substack{k'=1 \\ k'<k}}^{M} c_k c_{k'} \psi_{\mu k}\psi_{\mu k'} \cos(2\pi f_{kk'} t), \tag{6.12}$$

where

$$f_{kk'} = \frac{1}{T_{kk'}} = \frac{E_k - E_{k'}}{h}, \ \forall k > k' \tag{6.13}$$

are the frequencies ($f_{kk'}$) or periods ($T_{kk'}$) involved in charge transfer. If there are no degenerate eigenenergies (which holds for all cases studied in this part of this PhD thesis) then the number of different $f_{kk'}$ or $T_{kk'}$ involved in carrier transfer is $S = \binom{M}{2} = \frac{M!}{2!(M-2)!} = \frac{M(M-1)}{2}$. If eigenenergies are symmetric relative to some central value, then S decreases (there exist degenerate $f_{kk'}$ or $T_{kk'}$). Specifically, in that case, $S = \frac{M^2}{4}$, for even M and $S = \frac{M^2-1}{4}$ for odd M.

From Eq. (6.12), in the absence of degeneracy and for real c_k, $\psi_{\mu k}$, it follows that the mean over time probability to find the extra carrier at the μ-th monomer is

$$\left\langle |\Psi_\mu(t)|^2 \right\rangle = \sum_{k=1}^{M} c_k^2 \psi_{\mu k}^2. \tag{6.14}$$

In the following Chapters, we will deal with fixed boundaries, for which the above equation always holds.

Furthermore, from Eq. (6.12), it can be shown that the one-sided Fourier amplitude spectrum that corresponds to the probability $|\Psi_\mu(t)|^2$ is given by

$$|\mathcal{F}_\mu(f)| = \sum_{k=1}^{M} c_k^2 \psi_{\mu k}^2 \delta(f) + 2\sum_{k=1}^{M}\sum_{\substack{k'=1 \\ k'<k}}^{M} |c_k c_{k'} \psi_{\mu k}\psi_{\mu k'}| \delta(f - f_{kk'}). \tag{6.15}$$

Hence, the Fourier amplitude of frequency $f_{kk'}$ is $2|c_k v_{\mu k} c_{k'} v_{\mu k'}|$. We can further define the *weighted mean frequency* (WMF) of monomer μ as

$$f_{WM}^{\mu} = \frac{\sum_{k=1}^{M} \sum_{\substack{k'=1 \\ k'<k}}^{M} |c_k \psi_{\mu k} c_{k'} \psi_{\mu k'}| f_{kk'}}{\sum_{k=1}^{N} \sum_{\substack{k'=1 \\ k'<k}}^{M} |c_k \psi_{\mu k} c_{k'} \psi_{\mu k'}|}. \tag{6.16}$$

WMF expresses the mean frequency content of the extra carrier oscillation at monomer μ. Having determined the WMF for all monomers, we can now obtain a measure of the overall frequency content of carrier oscillations in the polymer: Since f_{WM}^{μ} is the weighted mean frequency of monomer μ and $\langle |\Psi_\mu(t)|^2 \rangle$ is the mean probability of finding the extra carrier at monomer μ, we define the *total weighted mean frequency* (TWMF) as

$$f_{TWM} = \sum_{\mu=1}^{M} f_{WM}^{\mu} \left\langle |\Psi_\mu(t)|^2 \right\rangle. \tag{6.17}$$

A quantity that evaluates simultaneously the magnitude of coherent charge transfer and the time scale of the phenomenon is the *pure* mean transfer rate [55]

$$k_{\lambda\mu} = \frac{\left\langle |\Psi_\mu(t)|^2 \right\rangle}{t_{\lambda\mu}}. \tag{6.18}$$

$t_{\lambda\mu}$ is the *mean transfer time*, i.e., having placed the carrier initially at monomer λ, the time it takes for the probability to find the extra carrier at monomer μ, $|C_\mu(t)|^2$, to become equal to its mean value, $\left\langle |\Psi_\mu(t)|^2 \right\rangle$, for the first time. For the *pure* mean transfer rates,

$$\begin{gathered} k_{\lambda\mu} = k_{\mu\lambda} = \\ k^{\text{equiv}}_{(N-\lambda+1)(N-\mu+1)} = k^{\text{equiv}}_{(N-\mu+1)(N-\lambda+1)}, \end{gathered} \tag{6.19}$$

where the superscript "equiv" refers to the equivalent polymer in the sense of Eq.(6.11).

7

Short DNA segments: monomers, dimers, and trimers

In this Chapter, we recruit the formulation presented in Chapter 6 to comparatively study extra electron or hole oscillations in DNA monomers (single base pairs), dimers (two base-pair sequences), and trimers (three base-pair sequences) within two variants of the TB model, i.e., the WM and the ELM[7]. We also compare aspects our TB results with RT-TDDFT[8]. As previously, we denote the segments using only the $5' - 3'$ strand, while the other one is implied. The mean over time probability, to find the extra carrier at a specific base (base pair) within the WM (ELM) is denoted with $\langle |A_n(t)|^2 \rangle$ ($\langle |A_n(t)|^2 \rangle$ for the 1st, $5' - 3'$, strand or $\langle |B_n(t)|^2 \rangle$ for the 2nd, $3' - 5'$, strand), with $n = 1, 2, \ldots, N$; cf. Eqs. (6.12), (6.14), (6.6).

7.1 Monomers

The WM cannot be used for charge transfer in monomers, since it considers a base pair as a single site. Hence, we use only the ELM, supposing that initially we place the carrier at one of the bases. A snapshot of electron oscillations between G and C in the base pair G-C, according to the HKS parametrization (see Table. A.2 in Appendix A), is given in Fig. 7.1.

[7]The content of Secs. 7.1, 7.2, and 7.3, as well as of Appendix B can be found published in Ref. [57]: K. Lambropoulos, K. Kaklamanis, A. Morphis, M. Tassi, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides, "Wire and extended ladder model predict THz oscillations in DNA monomers, dimers and trimers", J. Phys. Condens. Matter **28**, 595101 (2016).

[8]The content of Sec. 7.4 can be found published in Ref. [66], under CC BY 4.0, where more results and computational details are included.

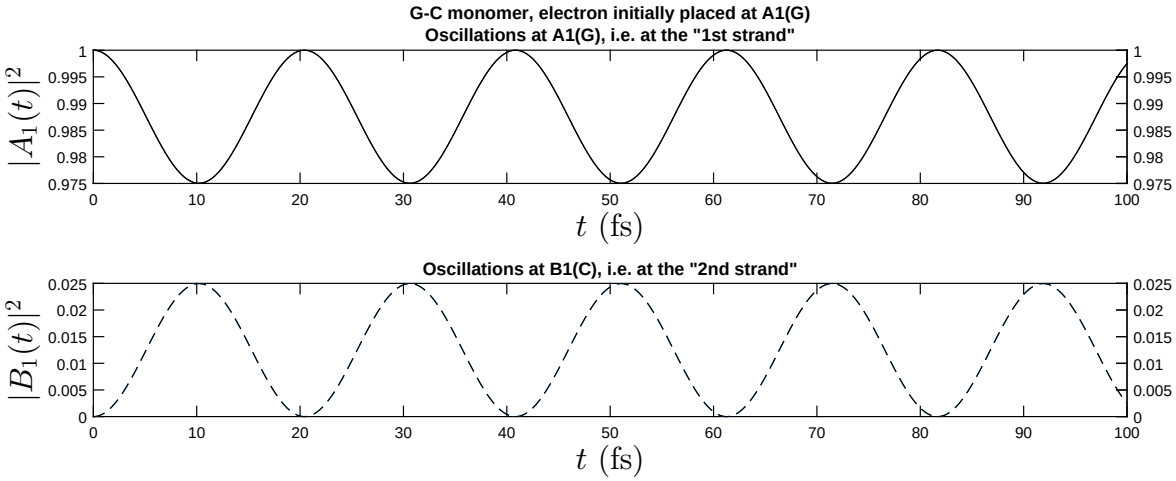

Figure 7.1: Electron oscillations within G-C, within the ELM and with the HKS parametrization.

It can be proven that an extra hole or electron oscillates between the bases of the two possible monomers (G-C and A-T) with frequency (or period)

$$f = \frac{1}{T} = \frac{\sqrt{(2t)^2 + \Delta^2}}{h}, \tag{7.1}$$

where t is the hopping integral between the complementary bases and Δ is the energy offset between the on-site energies of the complementary bases. Our results for A-T and G-C, both for holes and electrons, are shown in Fig. 7.2, with the HKS parametrization (Table A.2) as well as a parametrization taken from Ref. [68] ("MA parametrization"). For the HKS parametrization, $f \approx$ 50-200 THz ($T \approx$ 5-20 fs), while for MA parametrization, $f \approx$ 250-550 THz ($T \approx$ 2-4 fs). These ranges correspond to wavelengths $\approx$ 545 nm - 6000 nm i.e. from visible to near-infrared and mid-infrared[9]. It can also be proven that the maximum transfer percentage p [e.g., max $(|B_1(t)|^2)$ for initial conditions $A_1(0) = 1, B_1(0) = 0$, or vice versa], is given by

$$p = \frac{(2t)^2}{(2t)^2 + \Delta^2}. \tag{7.2}$$

We observe that the carrier is not very likely to be transferred between the monomer bases (p is very small in all cases). The pure maximum transfer rate defined as pf is also here very small in all cases. The pure mean transfer rate k is also shown. It can be analytically proven and numerically shown that here $k = 2pf$. T, f, p, pf and k do not depend on which base the carrier is initially placed at.

DNA base pairs within the ELM are two-level system of given stationary states (the two HOMOs or the two LUMOs) with a "perturbation" represented by the hopping integral, which impels an extra carrier to oscillate between these stationary states. Mathematically, the problem is equivalent to a two-level system (e.g. atom) under the

[9]ISO 20473 specifies: Near-Infrared (NIR) $0.78 - 3$ μm, Mid-Infrared (MIR) $3 - 50$ μm, Far-Infrared (FIR) $50 - 1000$ μm

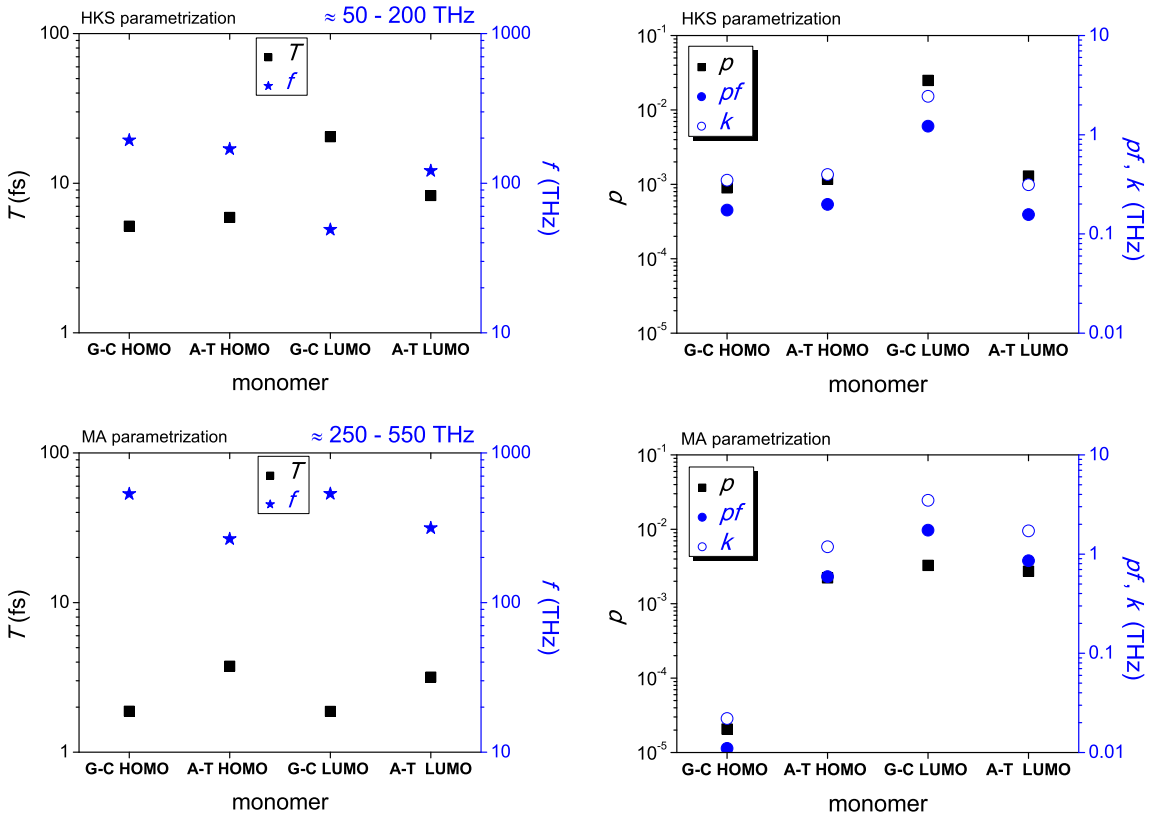

Figure 7.2: Charge oscillations in A-T and G-C base pairs within the ELM: period T, frequency f, maximum transfer percentage p, pure maximum transfer rate pf and pure mean transfer rate k. 1st row: HKS parametrization. 2nd row: MA parametrization.

influence of an electric field, which impels an electron to oscillate between the two eigenstates [semiclassical approach after Rotating Wave Approximation or the time-dependent problem with a Jaynes-Cummings Hamiltonian in a full quantum mechanical approach]. These problems are well known in the context of quantum optics [246]. The same applies to the dimer problem within the WM, which is discussed below.

7.2 Dimers

The possible dimers result from the combination by two of the four bases, since at the base located in the $5' - 3'$ strand always corresponds its complementary base. The number of possible combinations of four nitrogenous bases elements by two is $a^p = 16$. However, six of them are equivalent to other six i.e. GG≡CC, AA≡TT, AG≡CT, AC≡GT, TG≡CA, TC≡GA; cf. Eq. (6.11). Hence, the possible dimers are 10.

Within the WM, it has been shown that carrier movement in all dimers is strictly periodic [55, 247, 248]. The frequencies (or periods) are given by Eq. (6.13), where, now t is the hopping integral between the base pairs and Δ is the energy gap between the on-site energies of the base pairs. Using the the S parametrization (Table A.1), we

found [55, 247, 248] $f \approx 0.25$-100 THz, i.e. $T \approx 10$-4000 fs, i.e. wavelengths $\approx 3-1200$ μm, in other words mainly in the MIR and the FIR range. We also found that the maximum transfer percentage $p = 1$ for dimers made of identical monomers, but $p < 1$ for dimers made of different monomers and that the values of f, T, p, pf, k do not depend on which of the two monomers the carrier is initially placed at. Using the HKS parametrization (Table A.1) results in the same frequency range, although, the predicted frequencies vary slightly due to the different values of the TB parameters. A summarizing graph is given in Fig. 7.3.

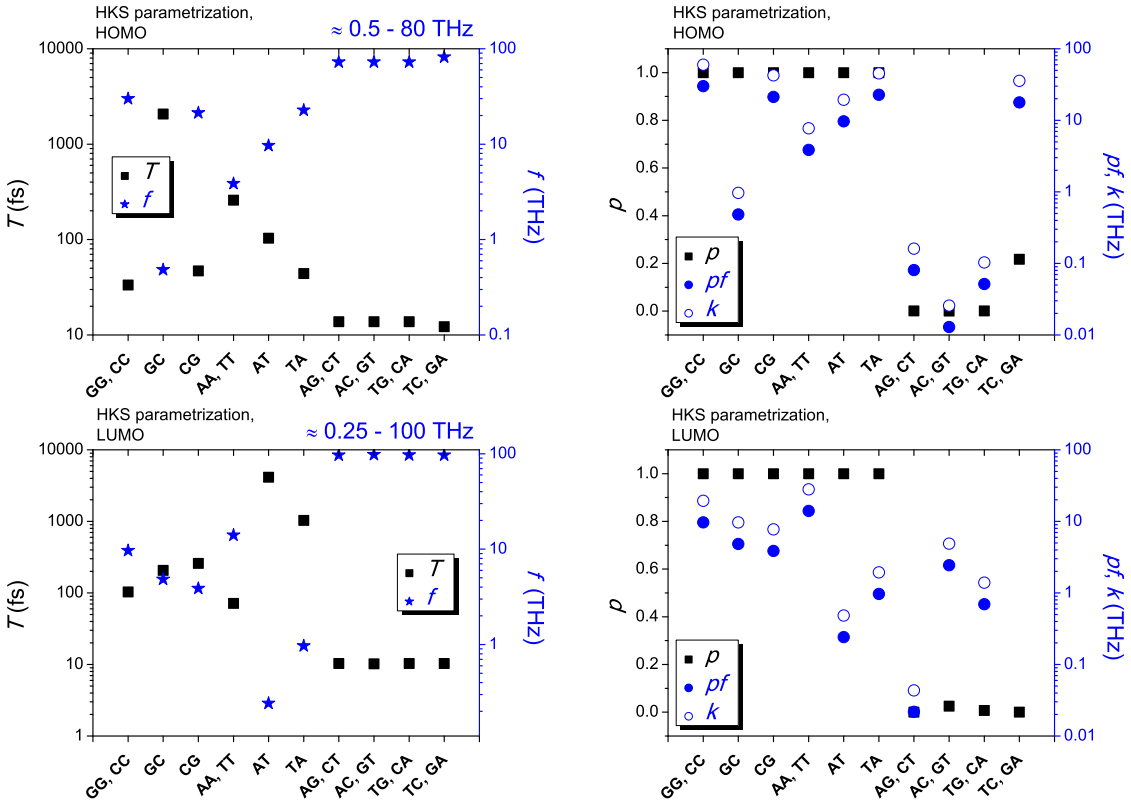

Figure 7.3: Charge oscillations in all possible DNA dimers within the WM and with the HKS parametrization: period T, frequency f, maximum transfer percentage p, pure maximum transfer rate pf and pure mean transfer rate k. It can be analytically proven and numerically shown that here $k = 2pf$.

Within the ELM, examples of snapshots of hole oscillations in GG, GC and CT dimers, using the HKS parametrization (Table A.2), are given in Fig. 7.4. Here, one cannot strictly determine periodicity in the carrier movement between the four bases, since there are more than one frequencies involved in charge transfer, hence f, T, p and pf cannot strictly be defined in the way they are within the WM.

Both TB models allow to determine the mean over time probability to find the carrier at a site (base pair for the WM or base for the ELM). A comparison between the mean probabilities obtained with the two TB models is shown in Fig. 7.5, using the HKS parametrization (Tables A.1, A.2). Comparing the two TB approaches, we reach the following conclusions: (a) Carrier transfer is large in dimers made of identical

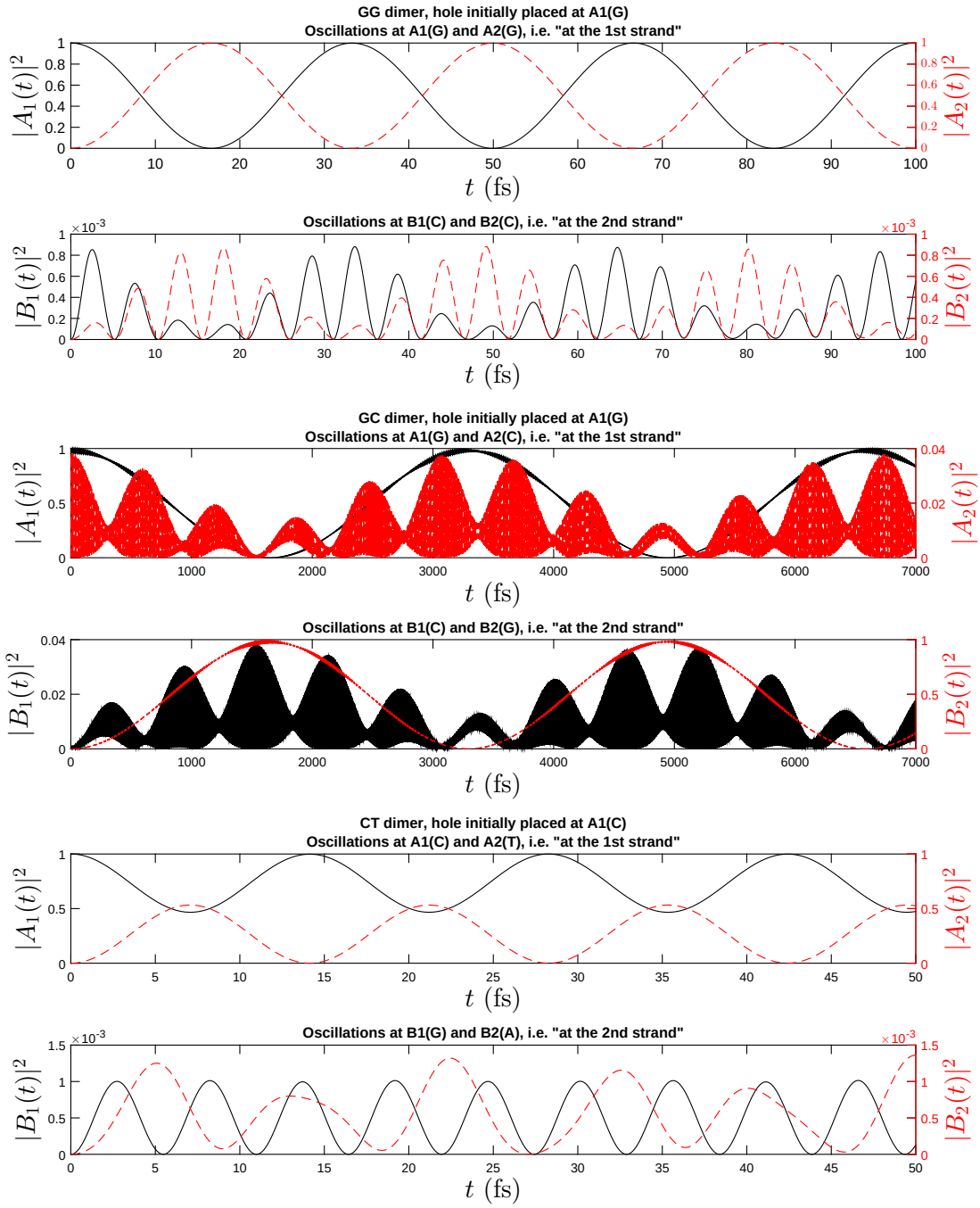

Figure 7.4: From top to bottom: Hole oscillations within the GG dimer, the GC dimer and the CT dimer, according within the ELM and with the HKS parametrization. Continuous black (dashed red) lines correspond to the 1st (2nd) base pair.

monomers: the carrier is shared between the two monomers which make up the dimer. (b) For dimers made of identical monomers, if purines are crosswise to purines, the carrier changes strand (from strand 1 to strand 2 or vice versa), while if purines are on the same strand, the carrier is transferred through the strand it was initially placed at. (c) For dimers made of different monomers, the carrier is transferred mainly through

the strand it was initially placed at. Transfer is very small; the carrier basically remains in the base it was initially placed at, while a small percentage passes to the other base of the same strand.

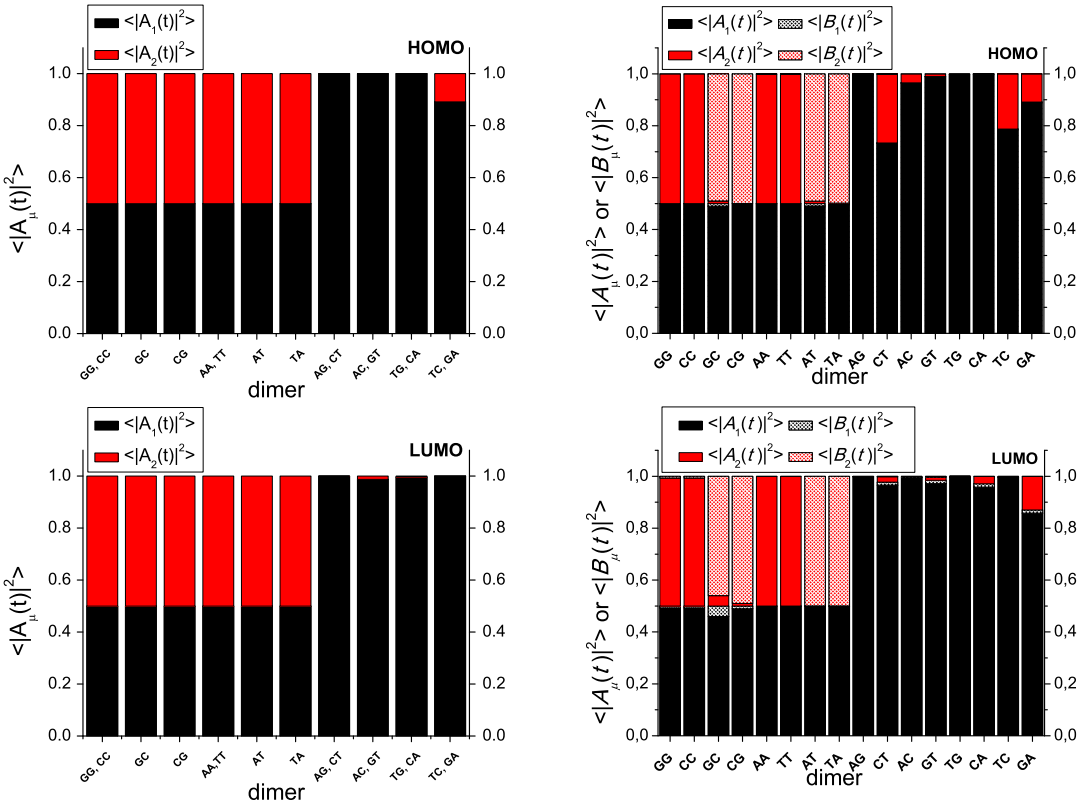

Figure 7.5: Mean over time probabilities to find an extra carrier [hole (1st row) or electron (2nd row)] at each site of a DNA dimer, as determined within the WM (left column) and and the ELM (right column), using the HKS parametrization. In the former, the carrier is initially placed at the 1st monomer, while, in the latter, it is initially placed at the base of the 1st monomer that belongs to the 1st strand.

Let us now compare the two TB models with respect to the frequency content. As mentioned above, within the ELM, but Fourier analysis using Eq. (6.15) shows similar frequency content in the THz domain. The Fourier spectra of three representative examples of DNA dimers, using the ELM and with the HKS parametrization (Table A.2) are depicted in Appendix B, Figs. B.1, B.2, B.3.

We start with GG, a dimer made of identical monomers with purines on purines (Fig. B.1). If we initially place the hole at A1(G) or A2(G), we obtain the main Fourier amplitude at $f \approx 30$ THz. If we initially place the hole at B1 (C) or B2 (C), we obtain the main Fourier amplitude at $f \approx 32$ THz. The rest of the frequencies show up with almost negligible amplitudes. These results are in accordance with the WM and HKS parametrization, where for the GG dimer we obtain $f \approx 30$ THz. Within the WM and with the S parametrization we had obtained $f \approx 48$ THz [55, 247]. The amplitudes at the main frequencies are ≈ 0.5, expressing the fact that for GG the mean probability to find the hole at a base is almost exclusively equally divided between the base the carrier was initially placed at and the other base of the same strand, cf. also Fig. 7.5.

We continue with GC, a dimer made of identical monomers with crosswise purines (Fig. B.2). If we initially place the hole at A1(G) or B2(G), we obtain the main Fourier amplitude at $f \approx 0.3$ THz. If we initially place the hole at B1 (C) or A2 (C), we obtain the main Fourier amplitude peak at $f \approx 1.6$ THz. The rest of the frequencies show up with almost negligible amplitudes. These results are in accordance with the WM and HKS parametrization, where for the GC dimer we obtain $f \approx 0.5$ THz. Within the WM and with the S parametrization we had obtained $f \approx 4.8$ THz [55, 247]. The amplitudes at the main frequencies are ≈ 0.5, expressing the fact that for GC the mean probability to find the hole at a base is, approximately, almost exclusively, equally divided between the base the carrier was initially placed at and the diagonally located base at the opposite strand, cf. also Fig. 7.5.

We finish with CT, a dimer made of different monomers (Fig. B.3). The main Fourier amplitude is at $f \approx 70.75$ THz, if the hole is initially placed at C or T, otherwise hole transfer is negligibly small. Within the WM and HKS parametrization, for the CT dimer we obtain $f \approx 72.5$ THz. Within the WM and with the S parametrization we had obtained $f \approx 74$ THz [55, 247]. The amplitudes at the main frequencies are ≈ 0.25, when the hole is initially placed at C or T and tiny when the hole is initially placed at G or A, expressing the fact that for CT the mean probability to find the hole at a base is approximately 0.75 at the base the carrier was initially placed at and approximately 0.25 at the other base of the same strand when these bases are C and T, but, hole transfer is negligibly small when the hole is initially placed at G or A.

Generally, a careful inspection in Figs. B.1, B.2 and B.3 shows that the Fourier analysis confirms conclusions (a),(b),(c).

We now turn our discussion to carrier pure mean transfer rates $k_{\lambda\mu}$, within the ELM and with the HKS parametrization (Table A.2). Our results are presented in Fig. 7.6. The specific values of $k_{\lambda\mu}$ depend, of course, on the TB hopping parameters used [249]. For dimers made of identical monomers with purine on purine (GG≡CC, AA≡TT), hole transfer is almost entirely of intrastrand character i.e. it is along the $5'$-$3'$ or $3'$-$5'$ directions. Moreover, since k_{13} and k_{24} satisfy Eq. (6.18), in Fig. 7.6 we observe a symmetry in the alternation of colors for the couples of equivalent dimers GG≡CC and AA≡TT. For dimers made of identical monomers with crosswise purines (CG, GC, TA, AT), there is significant diagonal hole transfer, and furthermore, for CG and TA the stronger hole transfer is along the $3'$-$3'$ direction. For dimers made of different monomers (AG≡CT, AC≡GT, TG≡CA, TC≡GA), hole transfer is almost exclusively of intrastrand character i.e. along the $5'$-$3'$ or $3'$-$5'$ directions; since k_{13} and k_{24} satisfy Eq. (6.18), we observe the same symmetry in the alternation of colors. For the couple TG≡CA, although k_{13} and k_{24} are the biggest among all other $k_{\lambda\mu}$, they are very small. Electron transfer in dimers made of identical monomers with purine on purine (GG≡CC, AA≡TT) is qualitatively similar to hole transfer in such dimers. For dimers made of identical monomers with crosswise purines (GC, CG, AT, TA), electron transfer is slightly different than hole transfer, in the sense that diagonal channels are important but are not, quantitatively, identically important. Electron transfer in dimers made of different monomers has a significant intrastrand character, but there is also intra-base-pair character in some cases. For the same reasons described above, we observe symmetry in color alternation for k_{ij} of equivalent dimers.

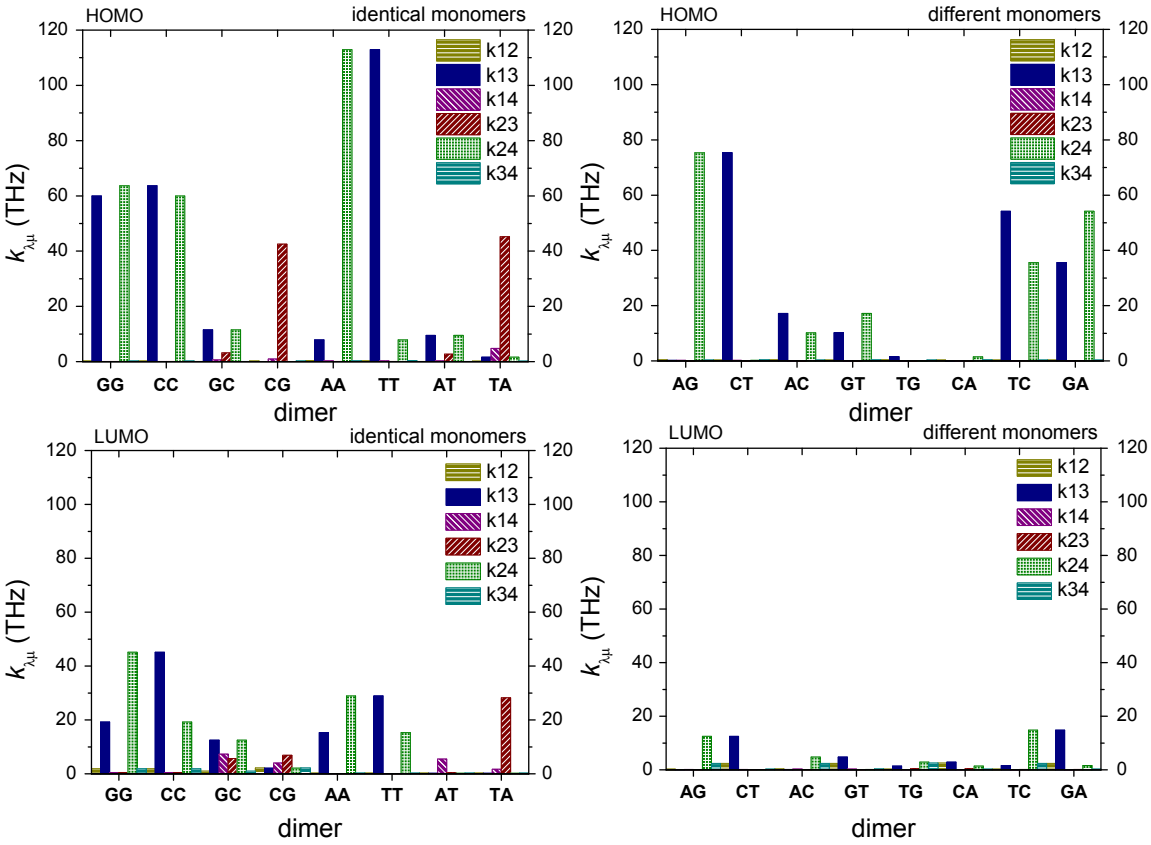

Figure 7.6: Pure mean transfer rates $k_{\lambda\mu}$ between bases λ and μ either for HOMO (hole transfer, 1st row) or for LUMO (electron transfer, 2nd row), for all dimers, within the ELM and with HKS parametrization. The 1st column corresponds to dimers made of identical monomers and the 2nd column to dimers made of different monomers.

7.3 Trimers

We compare the WM and the ELM for trimers made of identical monomers. Within the WM, we have shown [55, 247] that these trimers an extra carrier oscillates strictly periodically with

$$f = \frac{1}{T} = \frac{\sqrt{t^2 + t'^2}}{h}, \tag{7.3}$$

where t, t' are the hopping integrals between the base pairs (when all purines are on the same strand, $t = t'$). Using the S parametrization (Table A.1), we found $f \approx$ 0.5-33 THz ($T \approx$ 30-2000 fs) [247]; using the HKS parametrization (Table A.1), we find $f \approx$ 0.5-21 THz ($T \approx$ 48-2000 fs). In other words, for trimers made of identical monomers, the frequency range is narrower than for dimers. For 0 times crosswise purines, the maximum transfer percentage $p = 1$, while for 1 or 2 times crosswise purines $p < 1$ [55, 247]. For trimers made of different monomers, carrier movement is not strictly periodic within the WM. Generally, increasing the number of monomers above three, the system becomes more complex and periodicity is lost [247]; even in the simplest cases, e.g. tetramers made of identical monomers with all the purines on the same

strand, there is no periodicity [248]. Within the ELM, as in dimers, strict periodicity cannot be determined for any trimer, hence T, f, p, and pf cannot be defined.

The mean probabilities to find an extra carrier at a base and the mean transfer rates in GGG and AAA, within the ELM and with the HKS parametrization (Table A.1), are shown in Fig. 7.7. Remarkably, the probabilities to find the carrier at each base pair are either $\approx 0.375, 0.25, 0.375$ or $\approx 0.25, 0.5, 0.25$ depending on the initial placement of the carrier, in agreement with the rules we have established in Ref. [56] for the WM.

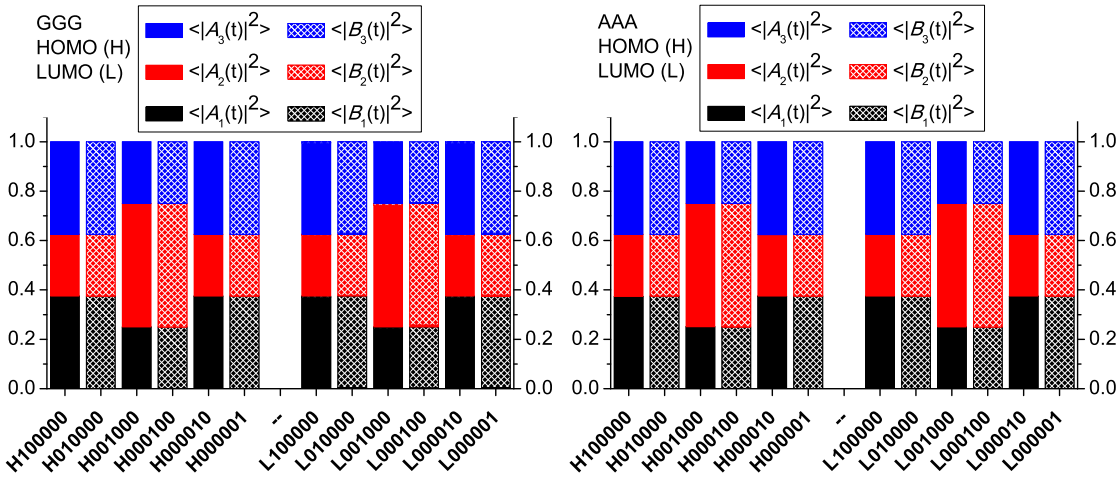

Figure 7.7: Mean over time probabilities to find the carrier at each base of GGG and AAA trimers after having placed it initially at a particular base, within the ELM and with the HKS parametrization. 100000 means that the carrier was initially placed at base A1, 010000 at base B1, etc.

The Fourier spectra show that the two models predict similar frequency content. Specific examples of Fourier spectra for oscillations in trimers within the ELM and with the HKS parametrization are presented in Appendix B, for three representative categories; made of identical monomers with no crosswise purines, i.e., GGG, Fig. B.4, made of identical monomers with crosswise purines, i.e., GCG, and made of different monomers with and without crosswise purines, i.e., CAC and CTC, Fig. B.5. For hole transfer in GGG, within the WM and with the S parametrization we found $f \approx 34.2$ THz [55, 247], with the HKS parametrization, we find $f \approx 21.2$ THz. This is in remarkable agreement with the frequencies obtained by Fourier Transform, within the ELM and with the HKS parametrization shown in Fig. B.4. Specifically, e.g. for initial placement of the hole at base A1(G), the main frequencies are around 21.2 THz (a double peak) and 42.4 THz (a single peak), while, e.g. for initial placement of the hole at base B1(C), the main frequencies are around 22.5 THz (a double peak) and 45 THz (a single peak). In other words, for hole transfer in GGG, with the HKS parametrization, the period predicted by the WM agrees with the approximate period predicted by the ELM.

The pure mean transfer rates $k_{\lambda\mu}$ (Fig. 7.8) confirm the intrastrand character of charge transfer in GGG and AAA. For hole transfer in GCG, with the HKS parametrization, the two models give similar results, indicating rather weak transfer. For example, placing the hole initially at the first base pair (within the WM) or placing the hole ini-

tially at the first base (within the ELM), the probability to find the hole at the first base pair is ≈ 0.9990 for the WM and 0.9848 for the ELM; at the last base pair, it is ≈ 0.0008 for the WM and 0.0006 for the ELM. This is also mirrored in the very small Fourier amplitudes for GCG (cf. Fig. B.5).

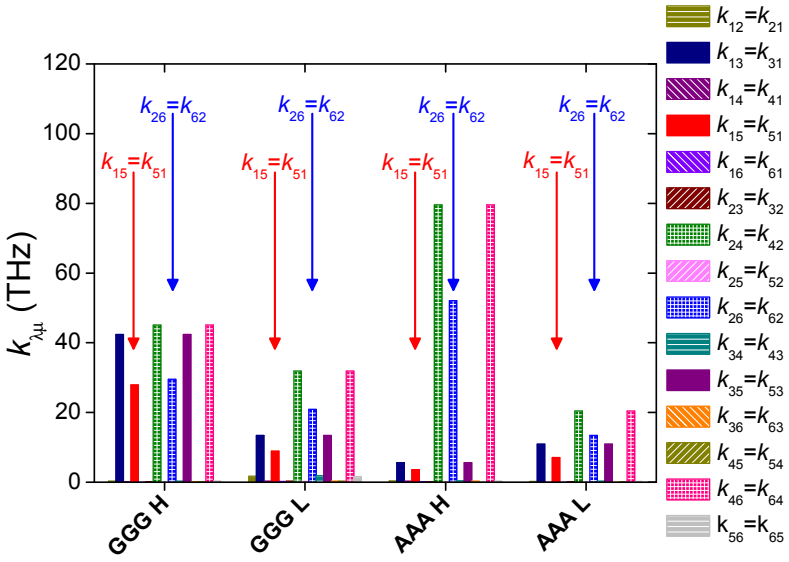

Figure 7.8: Pure mean transfer rates, $k_{\lambda\mu}$, from base λ to base μ, of GGG and AAA trimers, after having placed the charge initially at a particular base, within the ELM and with the HKS parametrization. The arrows indicate intrastrand transfer from the strand start to the strand end.

7.4 Comparison with RT-TDDFT

The results presented above for TB are in accordance with the more exact, yet more computationally costly, ab initio method, RT-TDDFT. Years after its initial development, DFT [250, 251], an efficient method for treating ground state properties of many electron systems, e.g., molecules or solids, was extended [252] to time dependent systems (TDDFT). Specifically, the Time-Dependent Kohn-Sham equations with an effective potential energy $v_{\mathrm{KS}}(\mathbf{r},t)$, uniquely described by the time-dependent charge density, $\rho(\mathbf{r},t)$, are, in atomic units,

$$i\frac{\partial}{\partial t}\Psi_j(\mathbf{r},t) = \left[-\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}(\mathbf{r},t)\right]\Psi_j(\mathbf{r},t) = \left[-\frac{1}{2}\nabla^2 + v_{\mathrm{ext}}(\mathbf{r},t) + v_{\mathrm{H}}(\mathbf{r},t) + v_{\mathrm{xc}}[\rho](\mathbf{r},t)\right]\Psi_j(\mathbf{r},t). \tag{7.4}$$

The charge density is the sum over all occupied orbitals $j = 1, 2, \ldots N_{occ}$, i.e.,

$$\rho(\mathbf{r},t) = \sum_{j=1}^{N_{occ}} |\Psi_j(\mathbf{r},t)|^2. \tag{7.5}$$

$v_{\text{ext}}(\mathbf{r}, t)$ includes external fields and nuclear potentials, $v_{\text{H}}(\mathbf{r}, t)$ is the Hartree potential energy. Exchange and correlation effects are included in $v_{\text{xc}}[\rho](\mathbf{r}, t)$. RT-TDDFT is based on a direct numerical integration of Eq. (7.4). This differs from the traditional linear-response approach, which is not actually a time-resolved method but instead solves Eq. (7.4) in the frequency domain for the excitation energies of a system subject to a small perturbation [253]. Within RT-TDDFT, the Time-Dependent Kohn-Sham equations are solved and the electron density is obtained at each time step. The electron density is then used for the calculation of the Hamiltonian in the next cycle of the self-consistent process.

Here, we restrict ourselves in a comparison of RT-TDDFT results for hole transfer in DNA monomers and dimers with TB. The calculations have been conducted using the, range-separated, CAM-B3LYP functional [254], and the the 6-31++G** [255] basis set. Larger systems have not been studied yet within RT-TDDFT due to computational restrictions. Computational results as well as additional results within RT-TDDFT can be found in Ref. [66].

Regarding monomers, RT-TDDFT predicts hole oscillations between the two bases of negligible magnitude and in the THz regime, in agreement with the TB results.

In Fig. 7.9, we present the mean probabilities to find the extra hole at each monomer of a DNA dimer, having placed it initially either at the 1st monomer (10) or at the 2nd monomer (01). We observe that for dimers made of identical monomers, there are almost equal probabilities to find the carrier at each monomer, while for dimers made of different monomers the probabilities to not find the carrier at the monomer of initial placement are very small, with the exception of the GA dimer where it is ≈ 0.3. This, again, agrees qualitatively with the TB picture, and especially with the WM (cf. the upper left panel of Fig 7.5).

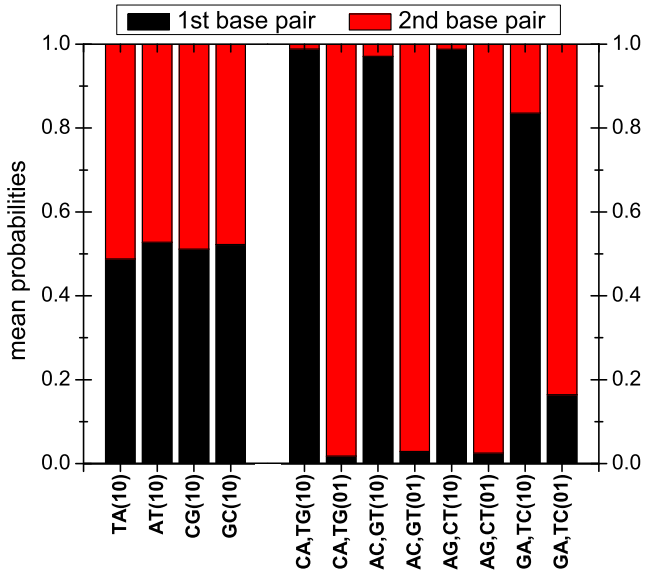

Figure 7.9: The mean probabilities to find the hole at each base pair of a dimer, having placed it initially at the 1st (10) or at the 2nd (01) base pair.

In Fig. 7.10 we depict the maximum transfer percentage, p, as well as the electronic

coupling energy (also called the electron transfer matrix element), $|V_{\mathrm{RP}}|$, between the reactant and product states R and P [51, 256–258], which reflects the magnitude of the interaction between the two monomers. In our case, the reactant state corresponds to the hole at the monomer of initial placement and the product state corresponds to the hole at the other monomer. We also show the energy difference $|\delta E|$ between reactant and product states. For dimers made of identical monomers, $|\delta E| \approx 0$, hence p is indeed close to 1. For dimers made of different monomers, we observe that AC, CA and AG have large $|\delta E| \approx$ 0.45-0.75 eV and small $|V_{\mathrm{RP}}| \approx 0 - 1$ eV, therefore their the maximum transfer percentage is insignificant. An exception is the dimer GA which not only has a large $|\delta E| \approx 0.65$ eV, but also the largest $|V_{\mathrm{RP}}|$ of all dimers $\approx$ 5 eV which makes charge transfer significant, with the maximum transfer percentage ≈ 0.3. This is in analogy with TB Eq. (7.2), where the energy difference between the two monomers is represented by Δ, and the strength of the interaction by t.

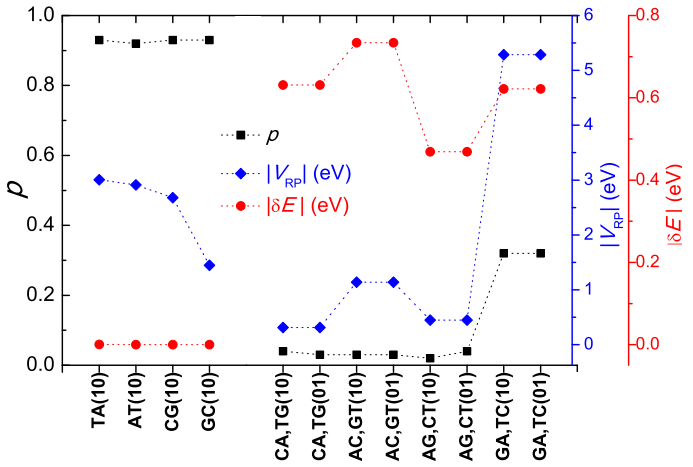

Figure 7.10: Maximum transfer percentage, i.e., oscillation amplitude, electron transfer coupling energy, $|V_{\mathrm{RP}}|$, and energy difference $|\delta E|$ between reactant and product states. Initially, the hole is placed at the 1st monomer (10) or at the 2nd monomer (01). Dotted lines serve as guides to the eye.

The frequencies of hole transfer have been obtained by Fourier analysis (Fig. 7.11). As expected by TB, for dimers made of different monomers, we generally get higher frequencies than for dimers made of identical monomers. Furthermore, from Fig. 7.11 we observe that for dimers made of identical monomers the frequencies follow $|V_{\mathrm{RP}}|$, which does not hold for dimers made of different monomers. This, again, agrees with the TB prediction [Eq. (7.1)].

Finally, in Fig. 7.12 we compare our RT-TDDFT results with those obtained within the TB WM (a) with the HKS parametrization, and (b) with the HKS parametrization (Table A.1). We observe that the maximum transfer percentages obtained by RT-TDDFT are in good agreement with those obtained by TB. For the periods, the results, both for RT-TDDFT and TB are quite close especially when we compare them with (b) the S parametrization.

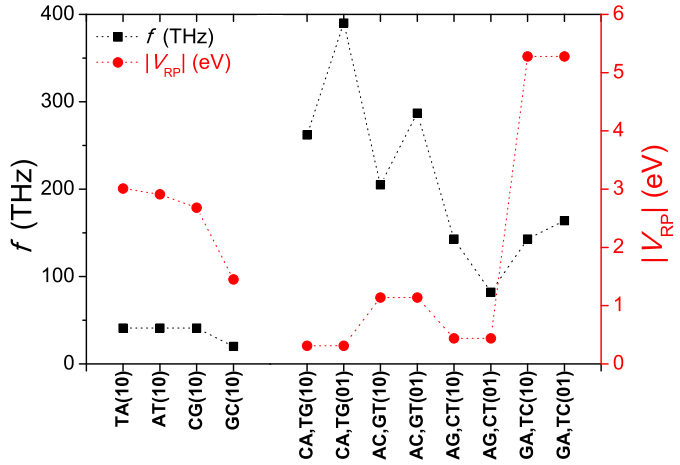

Figure 7.11: Hole oscillation frequency versus $|V_{\mathrm{RP}}|$ for DNA dimers. Dotted lines serve as guides to the eye.

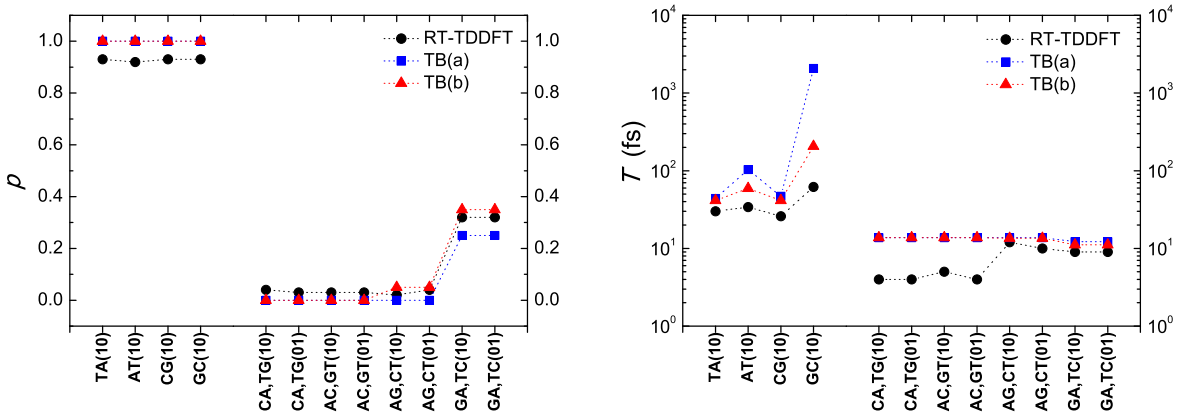

Figure 7.12: Comparison of RT-TDDFT and the WM with (a) the HKS parametrization and (b) the S parametrization: maximum transfer percentage (left panel) and oscillation period (right panel). Dotted lines serve as guides to the eye.

7.5 Conclusion

Using two TB approaches, i,e, the WM, in which a site coincides with a DNA base pair, and and the ELM, in which a site coincides with a DNA base, we studied various aspects of charge transfer in DNA monomers, dimers and trimers„ e.g. frequency content, maximum transfer percentages and transfer rates between sites, and mean probabilities to find the carrier at a particular site. We also successfully compared the two TB approaches. Naturally, the ELM allows for greater detail.

For DNA monomers, i.e. for A-T and G-C, within the ELM, we predicted electron or hole oscillations in the range $f \approx$ 50-550 THz ($T \approx$ 2-20 fs), i.e. wavelengths $\approx$ 545 nm - 6 µm, from visible to NIR and MIR. We found that the maximum transfer percentage

p and the pure maximum transfer rate pf between the bases are very small.

For DNA dimers, within the WM, we predicted electron or hole oscillations in the range $f \approx$ 0.25-100 THz ($T \approx$ 10-4000 fs) i.e. wavelengths $\approx$ 3-1200 μm, approximately in the MIR and FIR. For dimers made of identical monomers the maximum transfer percentage $p = 1$, but for dimers made of different monomers $p < 1$. Within the ELM, carrier oscillations are not strictly periodic but the frequency content is similar to that predicted within the WM. For the mean probabilities to find the carrier at a particular site, the two approaches give coherent, complementary results. The ELM shows that for dimers made of identical monomers, when purines are crosswise to purines, interstrand carrier transfer dominates, i.e. we have significant diagonal transfer, justifying the inclusion of diagonal hoppings, while if purines are on the same strand, intrastrand carrier transfer dominates. For dimers made of different monomers, we carrier transfer is mainly intrastrand character but the transfer percentage is small.

Within the WM, for trimers made of identical monomers, the carrier oscillates periodically with $f \approx$ 0.5-33 THz ($T \approx$ 30-2000 fs) if we use the S parametrization. Using the HKS parametrization, $f \approx$ 0.5-21 THz. For 0 times crosswise purines $p = 1$, for 1 or 2 times crosswise purines $p < 1$. Within the ELM, the carrier oscillations are not strictly periodic but the frequency content is similar to that predicted within the WM. For the mean probabilities to find the carrier at a particular site, the two approaches give coherent, complementary results.

Finally, we compared our TB results for monomers and dimers with RT-TDDFT, demonstrating that they are in good agreement.

8 Periodic DNA polymers

After our study of the excess charge transfer in short DNA segments, we move on to larger systems, i.e., periodic DNA polymers. Specifically, in this Chapter we will focus on some specific classes: monomer-polymers and dimer polymers[10], as well as polymers with increasing repetition unit[11]. As previously, only the $5' - 3'$ strand is used to denote the segments. Furthermore, polymers made of identical monomers are denoted by I, while polymers made of different monomers are denoted with D, followed by u, i.e., the number of monomers in the repetition unit. For example, D2 denotes a polymer made up by different monomers with a repetition unit of 2 monomers (in other words, the symbolism encompasses GA≡TC, GT≡AC, TG≡CA, AG≡CT).

8.1 Monomer-polymers and dimer polymers

In this Section, we perform a comparative study of aspects of charge transfer within all possible I1, I2, and D2 polymers, using the WM, with the S parametrization (Table A.1), and the ELM with the HKS parametrization (Table A.2).

8.1.1 Energy structure

Using the HOMO or LUMO energies of the bases that constitute a base pair, we can estimate the HOMO or LUMO energy of the base pair [61]. Specifically, supposing that $|\Psi_{bp}\rangle = \psi_{b1}|\psi_{b1}\rangle + \psi_{b2}|\psi_{b2}\rangle$, and taking the time-independent Schrödinger

[10]The content of Sec. 8.1 can be found published in Ref. [58]: K. Lambropoulos, M. Chatzieleftheriou, A. Morphis, K. Kaklamanis, R. Lopp, M. Theodorakou, M. Tassi, and C. Simserides, Phys. Rev. E **94**, 062403 (2016). © 2016 by the American Physical Society.

[11]The content of Sec. 8.2 can be found published in Ref. [59]: K. Lambropoulos, C. Vantaraki, P. Bilia, M. Mantela, and C. Simserides, Phys. Rev. E **98** 032412 (2018). © 2018 by the American Physical Society.

equation $\hat{H}\,|\Psi_{bp}\rangle = \epsilon^{bp}\,|\Psi_{bp}\rangle$ we find that the base pair eigenenergies are $\epsilon^{bp}_{1,2} = \frac{\epsilon_{b1}+\epsilon_{b2}}{2} \pm \sqrt{(\frac{\epsilon_{b1}-\epsilon_{b2}}{2})^2 + t^2}$, where ϵ_{b1} and ϵ_{b2} are the on-site energies of the bases and $t = \langle\psi_{b1}|\hat{H}|\psi_{b2}\rangle$ is the intra-base-pair hopping integral i.e. between the two bases that constitute a base pair. However, due to the weak hydrogen bonding between the bases that constitute a base pair, t is very small, of the order of 10 meV (cf. Table A.2). As a result, practically, $\epsilon^{bp}_{1,2} \approx \epsilon_{b1}, \epsilon_{b2}$ (with accuracy of 1 meV). Hence, we make the following *Observation*: Approximately, the HOMO of the base pair is the highest HOMO of the two bases and the LUMO of the base pair is the lowest LUMO of the two bases. This is expressed in Table 8.1, where we show all energies in eV with accuracy of 0.1 eV.

Table 8.1: On-site HOMO/LUMO energies of B-DNA bases and base pairs. All energies are given in eV.

base	**A**	**T**	**G**	**C**
ϵ^{b}_{H}	-8.3	-9.0	-8.0	-8.8
ϵ^{b}_{L}	-4.4	-4.9	-4.5	-4.3
gap	3.9	4.1	3.5	4.5
base pair	**A-T**		**G-C**	
ϵ^{bp}_{H}	-8.3		-8.0	
ϵ^{bp}_{L}		-4.9	-4.5	
gap	3.4		3.5	

Our numerical results for the eigenspectra of representative examples of I1, I2, and D2 polymers, which are presented in Figs. 8.1, 8.2, and 8.3, respectively, indicate that, as, increasing N, a polymer is formed, the energy eigenvalues are distributed around the on-site energies of the base pairs within WM or the bases within the ELM. Hence, the HOMO (LUMO) eigenspectrum of a given polymer within the WM corresponds to the upper (lower) part of its eigenspectrum within the ELM. We also mention that it can be proven (Appendix C, Theorem 1) within the WM, any sign alternation of the hopping integrals does not have any effect on the energy structure of the systems.

For I1 polymers, an analytical expression for the eigenvalues exists within the WM (cf. Eq. (3.21)). All eigenvalues are symmetric around the on-site energy ϵ^{bp} of the monomers and lie in the interval $\left[\epsilon^{bp} - 2\left|t^{bp}\right|, \epsilon^{bp} + 2\left|t^{bp}\right|\right]$. For odd N, the trivial eigenvalue, ϵ^{bp}, exists. The symmetry of the eigenvalues and the presence of the trivial one for odd N hold for all possible cases of I polymers within the WM, as it is shown in Appendix C, Theorem 2. Within the WM, an analytical expression can also be found for the eigenvectors [56]. The eigenvectors (hence, the occupation probabilities, too) are *eigenspectrum independent* [56], i.e., they do not depend on the TB parameters ϵ^{bp}, t^{bp}. Furthermore, the occupation probabilities display *palindromicity* [56], i.e., the occupation probability of each eigenstate of the μ-th monomer is equal to that of the $(N-\mu+1)$-th monomer ($|\psi_{\mu k}|^2 = \left|\psi_{(N-\mu+1)k}\right|^2$). Within the ELM, up to our knowledge, there are no analytical expressions for eigenvalues and eigenvectors. The eigenvalues are distributed in two subbands of different width, around the on-site energies

of the bases. Furthermore, in accordance with the above mentioned *Observation*, the upper (lower) subband of the HOMO (LUMO) eigenspectrum corresponds to the band calculated within the WM. Furthermore, our numerical results for the eigenvectors within the ELM indicate that, for μ odd ($5' - 3'$ strand), $|\psi_{\mu k}|^2 \approx |\psi_{(2N-\mu)k}|^2$, while, for μ even ($3' - 5'$ strand), $|\psi_{\mu k}|^2 \approx |\psi_{(2N-\mu+2)k}|^2$, i.e., the occupation probabilities of the eigenstates display approximate *strand-palindromicity*. For HOMO poly(dG)-poly(dC), a case where, according to the HKS parametrization, the hopping integrals between diagonally located bases of successive monomers in the $3'$-$3'$ and $5'$-$5'$ directions are equal, strand palindromicity is strict. This also holds for all types of I1 polymers, if the ELM is reduced to the simpler LM by neglecting $3'$-$3'$ and $5'$-$5'$ inter-strand interactions.

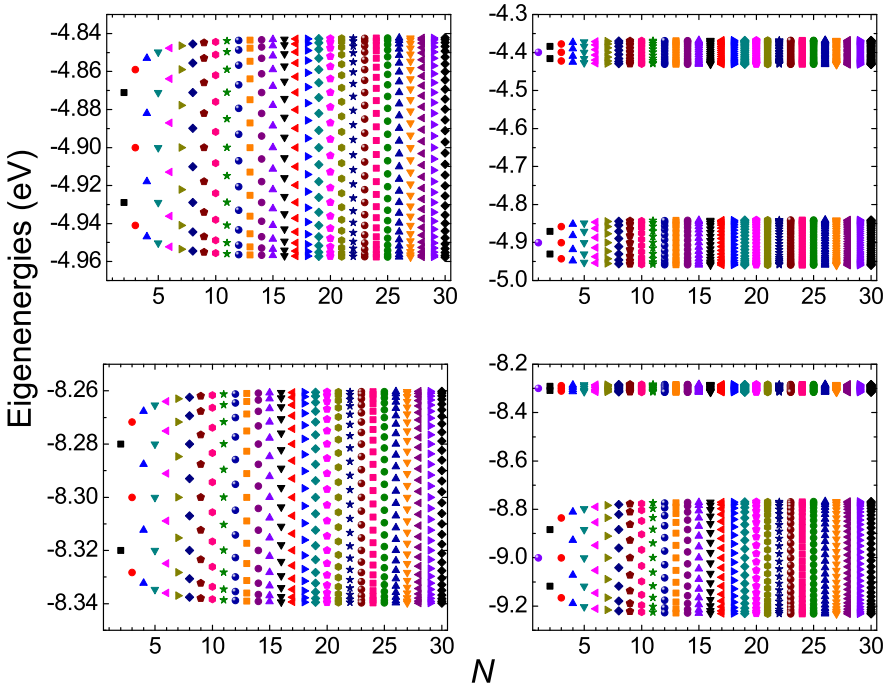

Figure 8.1: An example of I1 polymers: LUMO (first row) and HOMO (second row) eigenspectra of poly(dA)-poly(dT), within the WM (first column) and the ELM (second column).

For I2 polymers, analytical expressions for the eigenvalues when N is odd exist within the WM [56]. For N odd, the eigenvalues can be expressed explicitly in terms of Chebyshev zeros [162]. The eigenvalues lie in the interval $\left[\epsilon^{bp} - \sqrt{t_1^2 + t_2^2 + 2|t_1 t_2|}\,,\ \epsilon^{bp} + \sqrt{t_1^2 + t_2^2 + 2|t_1 t_2|}\right]$. For N even, there is no explicit formula, although the eigenvalues can be produced from Eq. (3.24). The eigenvalues lie in the same interval as for N odd. As mentioned above, all eigenvalues are symmetric around the on-site energy of the monomers, ϵ^{bp}, which is also an eigenvalue for N odd. The calculated HOMO/LUMO eigenspectrum for an example of I2 polymers, (GCGC...), displaying all the above mentioned properties, is shown in the left column of Fig. 8.2. Furthermore,

within the WM and for N odd, analytical expressions for the eigenvectors exist [163]. These eigenvectors (hence, the occupation probabilities, too) are *partially eigenspectrum dependent* [56], i.e., they depend on t_1, t_2 but not on ϵ^{bp}. Furthermore, for μ even, the occupation probability of each eigenstate of the μ-th monomer is equal to that of the $(N-\mu+1)$-th monomer ($|\psi_{\mu k}|^2 = |\psi_{(N-\mu+1)k}|^2$), i.e., for N odd, the occupation probabilities of I2 polymers display *partial palindromicity* [56]. Finally, for N odd, polymer equivalence [Eq. (6.11)] leads to the property $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = |\psi_{(N-\mu+1)k}|^2(\mathrm{XY}\ldots)$. For N even, we are aware of no analytical expressions for the eigenvectors, but our numerical results show that the occupation probabilities display *palindromicity* [56], i.e., for each eigenstate, the occupation probability of the μ-th monomer is equal to that of the $(N-\mu+1)$-th monomer ($|\psi_{\mu k}|^2 = |\psi_{(N-\mu+1)k}|^2$). Within the ELM, up to our knowledge, there are no analytical expressions for eigenvalues and eigenvectors. As an example of I2 polymers, we show in the right column of Fig. 8.2 the calculated HOMO/LUMO eigenspectra for GCGC.... The eigenvalues are distributed in two subbands of different width, around the on-site energies of the bases. Moreover, in accordance with the *Observation*, the upper (lower) subband of the HOMO (LUMO) eigenspectrum corresponds to the band calculated within the WM. For the eigenvectors within the ELM, for N odd, equivalence leads to the property $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = |\psi_{(2N-\mu+1)k}|^2(\mathrm{XY}\ldots)$, while for N even, $|\psi_{\mu k}|^2 = |\psi_{(2N-\mu+1)k}|^2$. In other words, the occupation probabilities of the eigenstates display *base-palindromicity*.

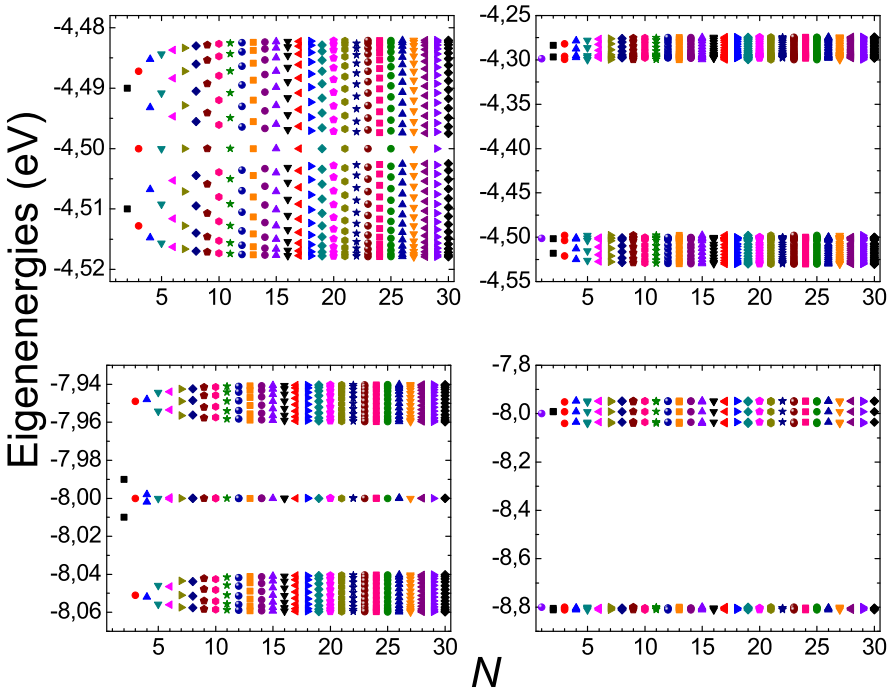

Figure 8.2: An example of I2 polymers: LUMO (first row) and HOMO (second row) eigenspectra of GCGC..., within the WM (first column) and the ELM (second column).

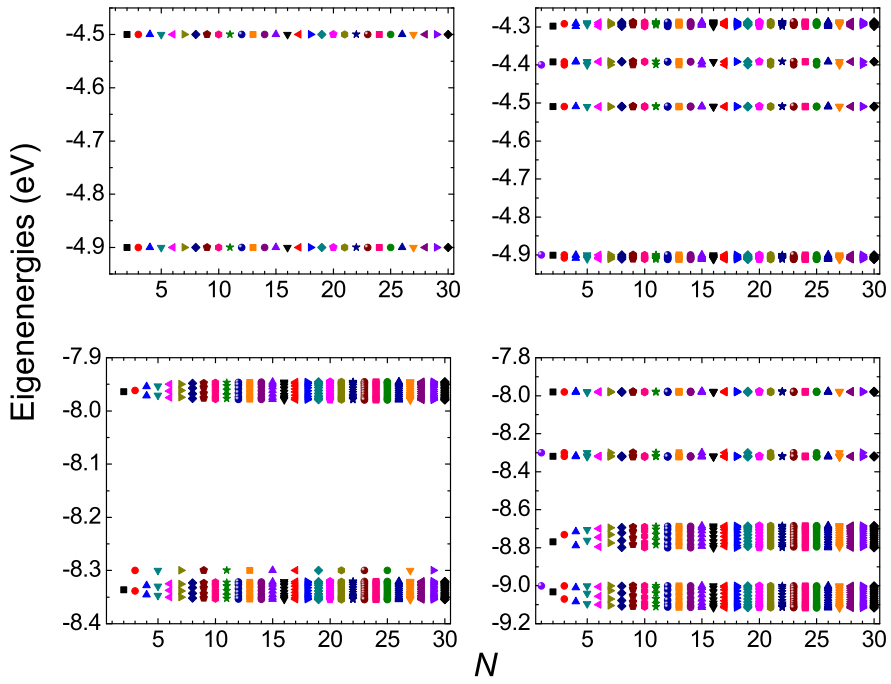

Figure 8.3: An example of D2 polymers: LUMO (first row) and HOMO (second row) eigenspectra of TCTC..., within the WM (first column) and the ELM (second column).

For D2 polymers, analytical expressions for the eigenvalues when N is odd exist within the WM [56]. Let us denote $\Sigma = (\epsilon_1^{bp} + \epsilon_2^{bp})/2$ and $\Delta = (\epsilon_1^{bp} - \epsilon_2^{bp})/2$. Then, the eigenvalues include ϵ_1^{bp}, while the rest eigenvalues lie in the interval $[\Sigma - \sqrt{\Delta^2 + t_1^2 + t_2^2 + 2|t_1 t_2|}, \Sigma + \sqrt{\Delta^2 + t_1^2 + t_2^2 + 2|t_1 t_2|}]$. For N even, there is no explicit formula, although the eigenvalues can be produced from Eq. (3.24). Our numerical results show that all eigenvalues are symmetric around Σ, and lie in the same interval as for N odd. The calculated HOMO/LUMO eigenspectrum for an example of D2 polymers (TCTC...), displaying all the above mentioned properties, is shown in the left column of Fig. 8.3. Within the WM and for N odd, analytical expressions can also be found for the eigenvectors [164]. The eigenvectors (hence, the occupation probabilities, too) are *eigenspectrum dependent* [56], i.e., they depend on all $\epsilon_1^{bp}, \epsilon_2^{bp}, t_1$, t_2. Furthermore, for μ even, the occupation probability of each eigenstate of the μ-th monomer is equal to that of the $(N-\mu+1)$-th monomer ($|\psi_{\mu k}|^2 = |\psi_{(N-\mu+1)k}|^2$), i.e., for N, odd, the occupation probabilities of D2 polymers display *partial palindromicity*. For N odd, equivalence leads to $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = |\psi_{(N-\mu+1)k}|^2(\mathrm{Y_{compl}X_{compl}}\ldots)$. For N even, up to our knowledge, no analytical expressions for the eigenvectors exist, but equivalence leads to $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = |\psi_{(N-\mu+1)k}|^2(\mathrm{X_{compl}Y_{compl}}\ldots)$. Our numerical results show that, for all μ, $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = |\psi_{\mu(N-k+1)}|^2(\mathrm{X_{compl}Y_{compl}}\ldots)$. Within the ELM, there are no analytical expressions in the literature for eigenvalues and eigenvectors, as far as we know. The calculated HOMO/LUMO eigenspectra for an example of D2 polymers (TCTC...) are demonstrated in the right column of

Fig. 8.3. The eigenvalues are distributed in four subbands of different width, around the on-site energies of the bases. Moreover, in accordance with the *Observation*, the two upper (lower) subbands of the HOMO (LUMO) eigenspectrum correspond to the bands calculated within the WM. For the eigenvectors within the ELM, equivalence leads to the properties $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = \left|\psi_{(2N-\mu+1)k}\right|^2(\mathrm{Y_{compl}X_{compl}}\ldots)$, for N odd, and $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = \left|\psi_{(2N-\beta+1)k}\right|^2(\mathrm{X_{compl}Y_{compl}}\ldots)$, for N even. Our numerical results indicate that there are no palindromic properties.

The DOS can be determined directly by the eigenspectra. It nicely represents the corresponding eigenspectral properties mentioned above. In Figs. 8.4, 8.5, and 8.6, we illustrate the numerically determined DOS for some representative examples of I1, I2 and D2 polymers, respectively, in the large N limit. We observe that, due to the fact that the eigenenergies become denser and denser as we approach the band or subband edges, vHS occur at the edges of each band or subband. We also notice that, in the large N limit, the polymer boundaries play insignificant role in the electronic structure, hence, for the same set of TB parameters, the polymers' DOS is essentially the same. For example, in the large N limit, either GCGC... or CGCG..., for either N odd or N even have practically the same DOS.

The numerically derived DOS for an example of I1 polymers is presented in Fig. 8.4. Within the WM, there is no minigap, but within the ELM there is a minigap of $\approx$ 0.545 eV; in accordance with the *Observation*, the upper subband of the HOMO band calculated within the ELM, corresponds to the HOMO band calculated within the WM. The minigap is due to the different HOMO on-site energies of the two bases (-8.0 eV for G, -8.8 eV for C).

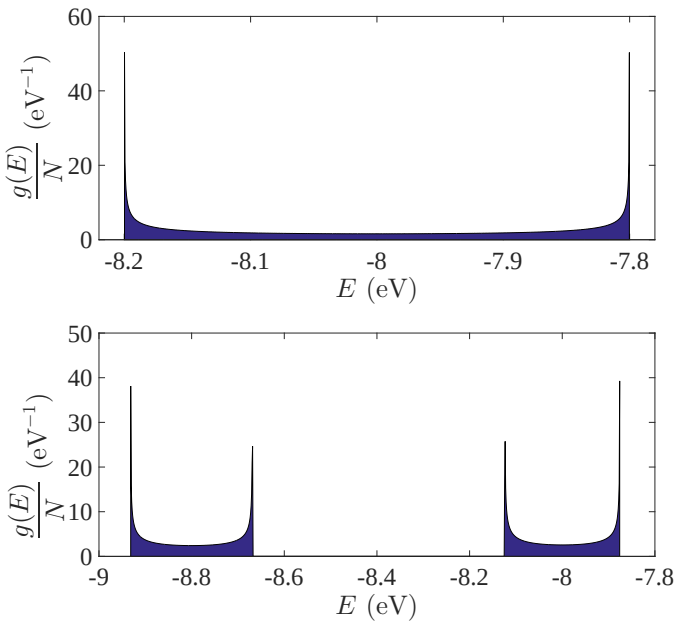

Figure 8.4: DOS for an example of I1 polymers, poly(dG)-poly(dC) ($N = 10^5$, HOMO), within the WM (top) and the ELM (bottom).

A DOS example of I2 polymers is shown in Fig. 8.5. Within the WM, there is a small

($\approx$ *0.004 eV*) minigap. Within the ELM, there is a minigap of ≈ 0.200 eV; in accordance with the *Observation*, the lower subband of the LUMO band calculated within the ELM corresponds to the LUMO band calculated within the WM. Within the ELM, there are two additional small ($\approx$ *0.003 eV*, 0.001 eV) minigaps, hardly noticeable at this scale. The *italicized* ELM minigap corresponds to the WM minigap, also *italicized*. Within the ELM, apart from the vHS at the subband edges, there is an additional singularity inside the second subband, which is hardly seen at this scale and an additional singularity inside the third subband, which is almost invisible at this scale.

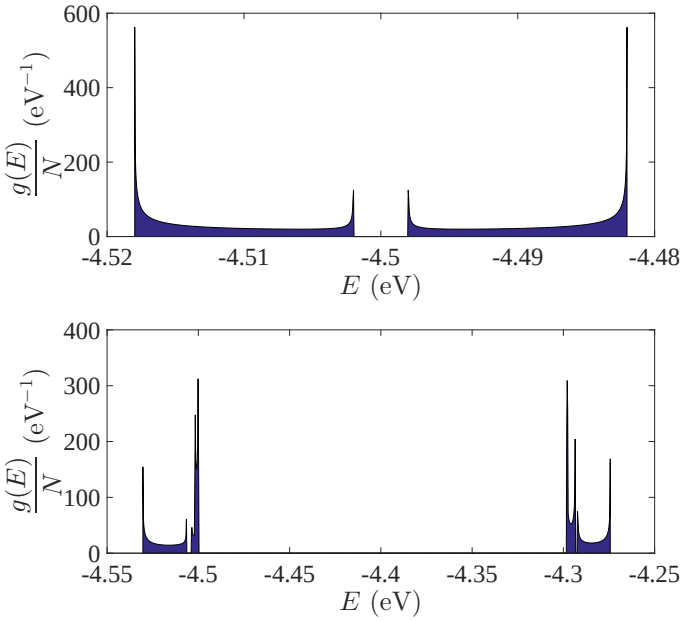

Figure 8.5: DOS for an example of I2 polymers, CGCG... ($N = 10^5$, LUMO), within the WM (top) and the ELM (bottom).

A DOS example of D2 polymers is shown in Fig. 8.6. Within the WM, there is a minigap a little greater than *0.340 eV*, due to the different HOMO on-site energies of the two base pairs (-8.0 eV for G-C, -8.3 eV for A-T). Within the ELM, four minibands are formed approximately around the HOMO on-site energies of the four bases (-9.0 eV for T, -8.8 eV for C, -8.3 eV for A and -8.0 eV for G), with three relevant minigaps (0.205 eV, 0.362 eV, *0.334 eV*). Two of these minibands are very narrow. In accordance with the *Observation*, the higher two subbands of the HOMO band calculated within the ELM correspond to the HOMO band calculated within the WM. Again, the *italicized* ELM minigap corresponds to the WM minigap.

Finally, in Fig. 8.7, we present the HOMO-LUMO energy gaps, in the large N limit, for all three types of polymers. The energy gap of a monomer is the difference between its LUMO and HOMO levels. The energy gap of a polymer is the difference between the lowest level of the LUMO regime and the highest level of the HOMO regime, because we assume that the orbitals –one per site– which contribute to the HOMO (LUMO) band are occupied (empty), since in both possible monomers there is an even number of p_z electrons contributing to the π stack [61]. Both TB approaches predict similar

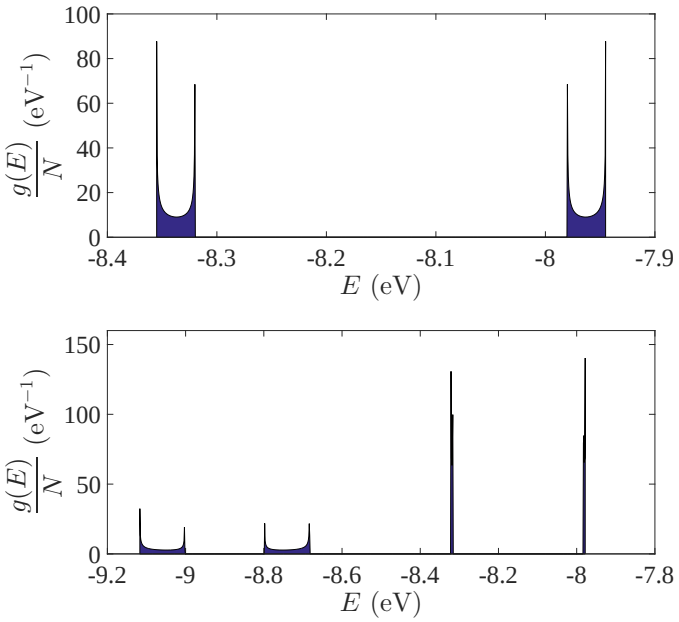

Figure 8.6: DOS for an example of D2 polymers, CTCT... ≡ AGAG... ($N = 10^5$, HOMO), within the WM (top) and the ELM (bottom).

gaps, in the range $\approx 3.04 - -3.42$ eV. We also compare the polymer gaps with the two possible monomer gaps. The decrease of the energy gap, as we move from monomer to polymer, is larger for D2 polymers.

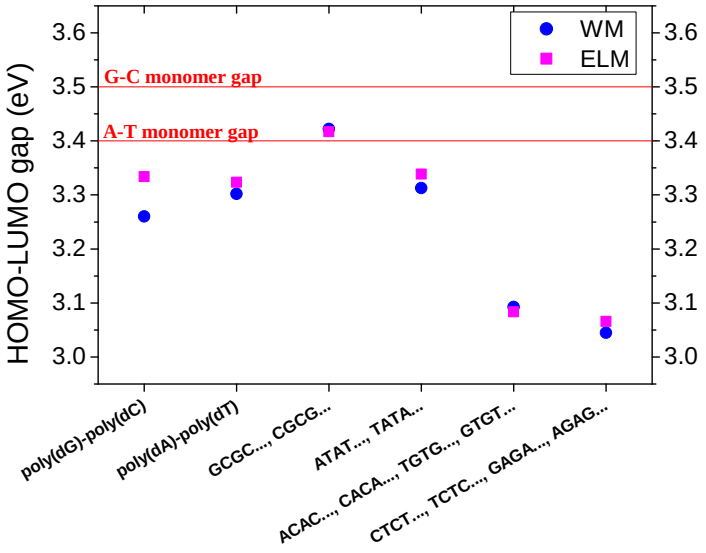

Figure 8.7: HOMO-LUMO gaps for all possible I1, I2, and D2 polymers, within the WM (blue dots) and the ELM (purple squares). The horizontal lines denote the HOMO-LUMO gaps of the two possible monomers.

8.1.2 Mean over time probabilities

We will study the mean over time probabilities to find an extra hole (HOMO) or electron (LUMO) at each site of all possible I1, I2, and D2 polymers within both TB approaches. For the WM, we initially place the carrier at the first monomer, while for the ELM we place it at each base of the first monomer (we call them *initial condition 1* for the base in the $5'-3'$ strand and *initial condition 2* for the base in the $3'-5'$ strand). Within the WM, from Eqs. (6.12), (6.10), it follows that the mean over time probability to find the extra carrier at the μ-th monomer is

$$\left\langle |\Psi_\mu(t)|^2 \right\rangle = \sum_{k=1}^{N} \psi_{1k}^2 \psi_{\mu k}^2. \tag{8.1}$$

Within the ELM, from Eq. (6.12) and the initial condition 1 or the initial condition 2 it follows that the mean probability to find the extra carrier at the μ-th base is

$$\left\langle |\Psi_\mu(t)|^2 \right\rangle = \begin{cases} \sum_{k=1}^{2N} \psi_{1k}^2 \psi_{\mu k}^2 \\ \sum_{k=1}^{2N} \psi_{2k}^2 \psi_{\mu k}^2 \end{cases}. \tag{8.2}$$

From Eqs. (8.1) and (8.2), we conclude that the palindromicity and eigenspectrum (in)dependence properties for the occupation probabilities, discussed in the previous subsection, hold also for the mean over time probabilities. Finally, for equivalent polymers it can be shown that for the WM it holds that $\langle |\Psi_N(t)|^2 \rangle_{\text{YX...}} = \langle |\Psi_N(t)|^2 \rangle_{\text{equiv(YX...)}}$, while for the ELM $\langle |\Psi_{2N}(t)|^2 \rangle_{\text{YX...}} = \langle |\Psi_{2N}(t)|^2 \rangle_{\text{equiv(YXYX...)}}$ (initial condition 1) and $\langle |\Psi_{2N-1}(t)|^2 \rangle_{\text{YXYX...}} = \langle |\Psi_{2N-1}(t)|^2 \rangle_{\text{equiv(YXYX...)}}$ (initial condition 2).

For I1 polymers, examples of our numerical results (for $N = 12$) for the mean over time probabilities to find an extra hole or electron at each base pair, within the WM, or at each base, within the ELM, are shown in Fig. 8.8. Within the WM (top panel of Fig. 8.8), the mean over time probabilities to find the carrier at a specific monomer display *palindromicity* and *eigenspectrum independence* [56]. Specifically, it can analytically be shown that

$$\left\langle |\Psi_1(t)|^2 \right\rangle = \left\langle |\Psi_N(t)|^2 \right\rangle = \frac{3}{2(N+1)}, \forall N \geq 2, \tag{8.3a}$$

$$\left\langle |\Psi_2(t)|^2 \right\rangle = \cdots = \left\langle |\Psi_{N-1}(t)|^2 \right\rangle = \frac{1}{N+1}, \forall N \geq 3. \tag{8.3b}$$

Within the ELM, the mean over time probabilities to find the carrier at a specific base display *approximate strand-palindromicity*. Moreover, adding the mean probabilities of the bases that constitute each monomer, it follows that the mean probabilities to find the carrier at a specific monomer are approximately palindromes and almost identically equal, for all cases, to the mean probabilities within TB I, which are strictly

palindromes (cf. Eq. (8.3)). This quantitative agreement suggests that the eigenspectrum independence predicted within the simpler WM approach, leads to essentially the same results as those derived by the more complicated ELM. From Fig. 8.8 we observe that, within the ELM, the carrier moves almost exclusively through the strand it was initially placed at, i.e. carrier movement is mainly of intra-strand character. Furthermore, within the ELM, our results for the two initial conditions are in complete agreement.

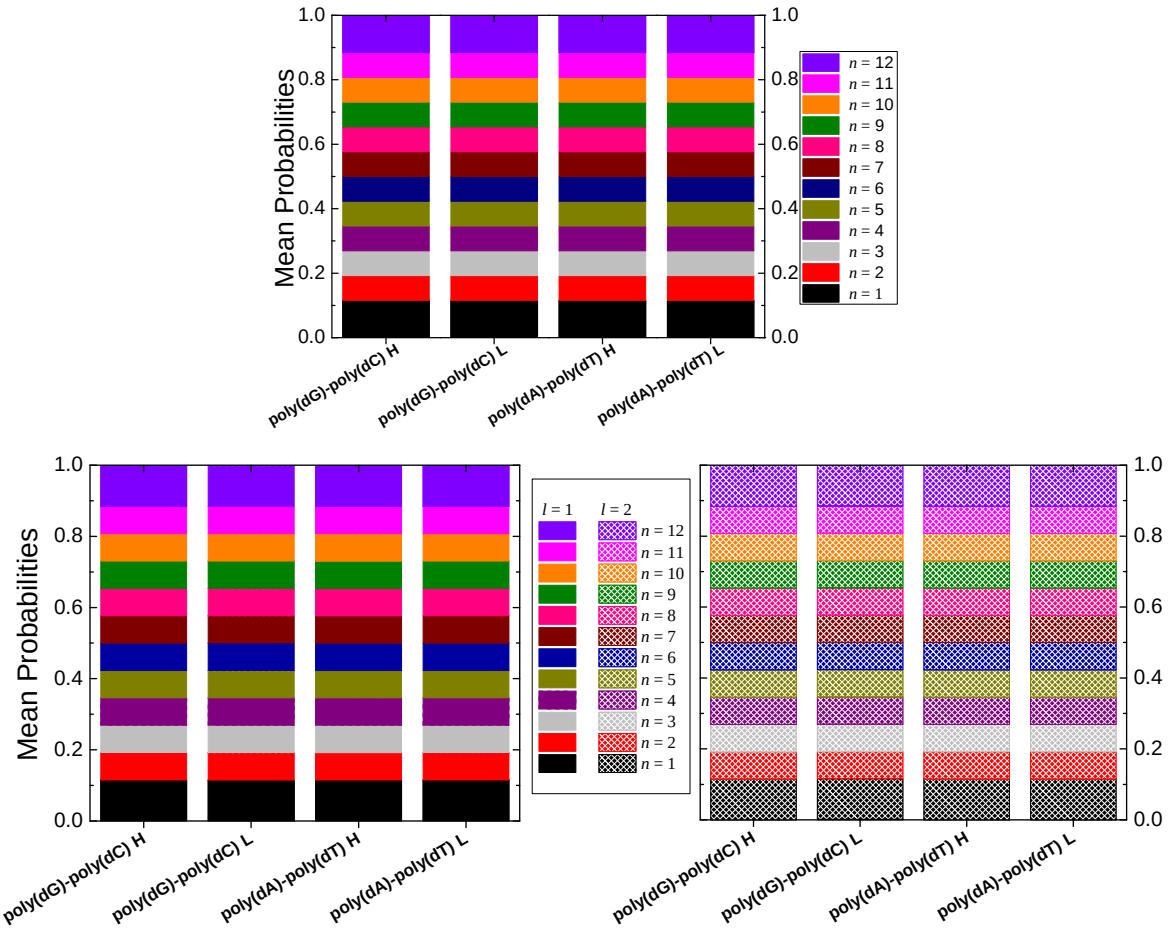

Figure 8.8: Mean over time probabilities to find an extra hole (HOMO) or electron (LUMO) at each base pair, within the WM (top), or at each base, within the ELM (bottom left: initial condition 1, bottom right: initial condition 2) for all I1 polymers, for $N = 12$.

For I2 polymers, examples of our numerical results for the mean over time probabilities to find an extra hole or electron at each base pair, within the WM, or at each base, within the ELM, are shown in Figs. 8.9 and 8.10, for odd or and even N, respectively. Within the WM, the mean probabilities to find the carrier at a specific monomer display [56] *partial eigenspectrum dependence* (i.e., dependence on the hopping parameters but not on the on-site energy), *partial palindromicity* (i.e., only for even μ) for N odd and *palindromicity* (i.e., for all μ) for N even. Within the ELM, for N even,

the mean probabilities to find the carrier at a specific base display *base-palindromicity*. Moreover, adding the mean probabilities of the bases that constitute each base pair, the mean probabilities to find the carrier at a specific base pair are palindromes, in accordance with the prediction of the WM. From the bottom panels of Figs. 8.9 and 8.10, we observe that, within the ELM, the carrier moves preferably through the bases that are identical with the one it was initially placed at, in other words it moves crosswise through identical bases, i.e., carrier movement is mainly of inter-strand character. For N odd, both TB approaches show that there are some cases, in which the carrier hardly moves from its initial site. If we add or subtract a monomer, i.e. for N even, both TB approaches show that a large percentage of the carrier is transferred at the end monomer. Furthermore, both TB approaches show that the mean probability to find the carrier at the last monomer is generally bigger for N even than for N odd.

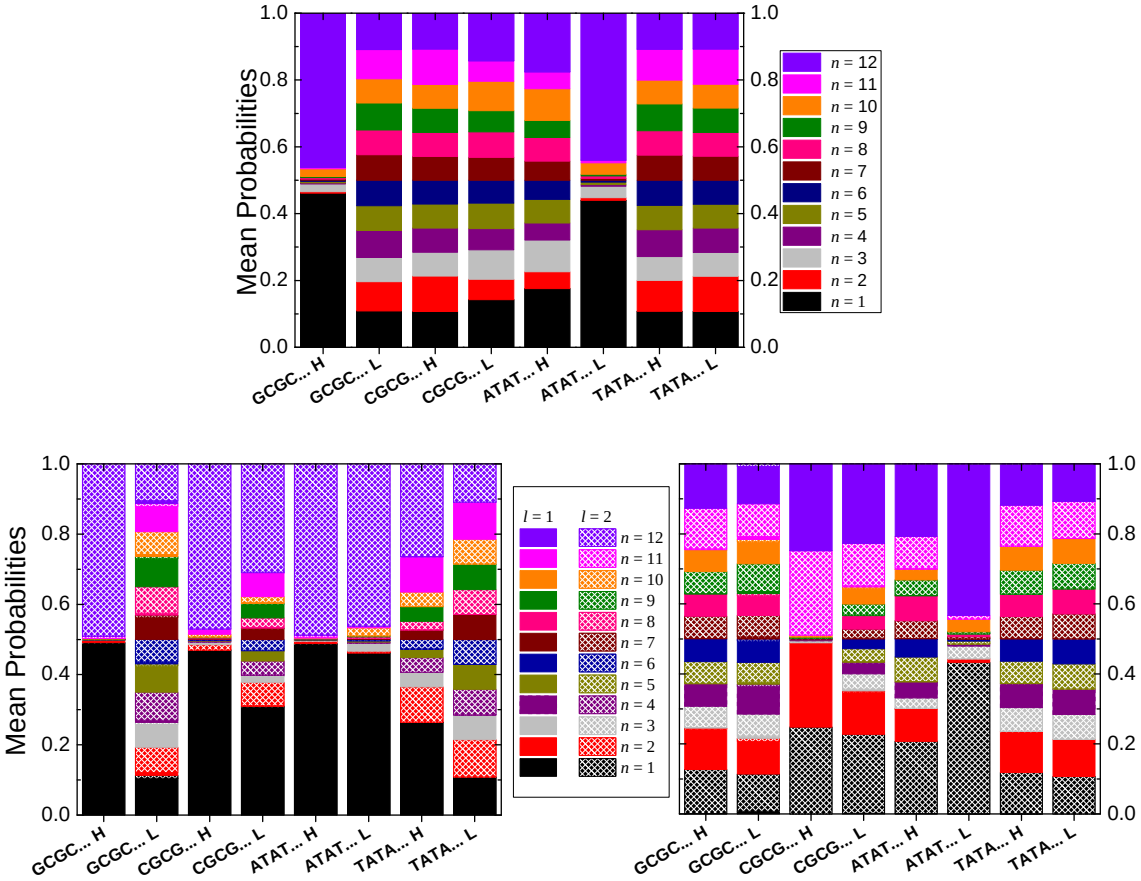

Figure 8.9: Mean over time probabilities to find an extra hole (HOMO) or electron (LUMO) at each base pair, within the WM (top), or at each base, within the ELM (bottom left: initial condition 1, bottom right: initial condition 2) for all I2 polymers, for $N = 12$.

For D2 polymers, examples of our numerical results for the mean over time probabilities to find an extra hole or electron at each base pair, within the WM, or at each base, within the ELM, are shown in Figs. 8.11 and 8.12, for odd or and even N, respectively. Within the WM, given that $|\psi_{\mu k}|^2(\mathrm{YX}\ldots) = \left|\psi_{\mu(N-k+1)}\right|^2(\mathrm{X_{compl}Y_{compl}}\ldots)$, for

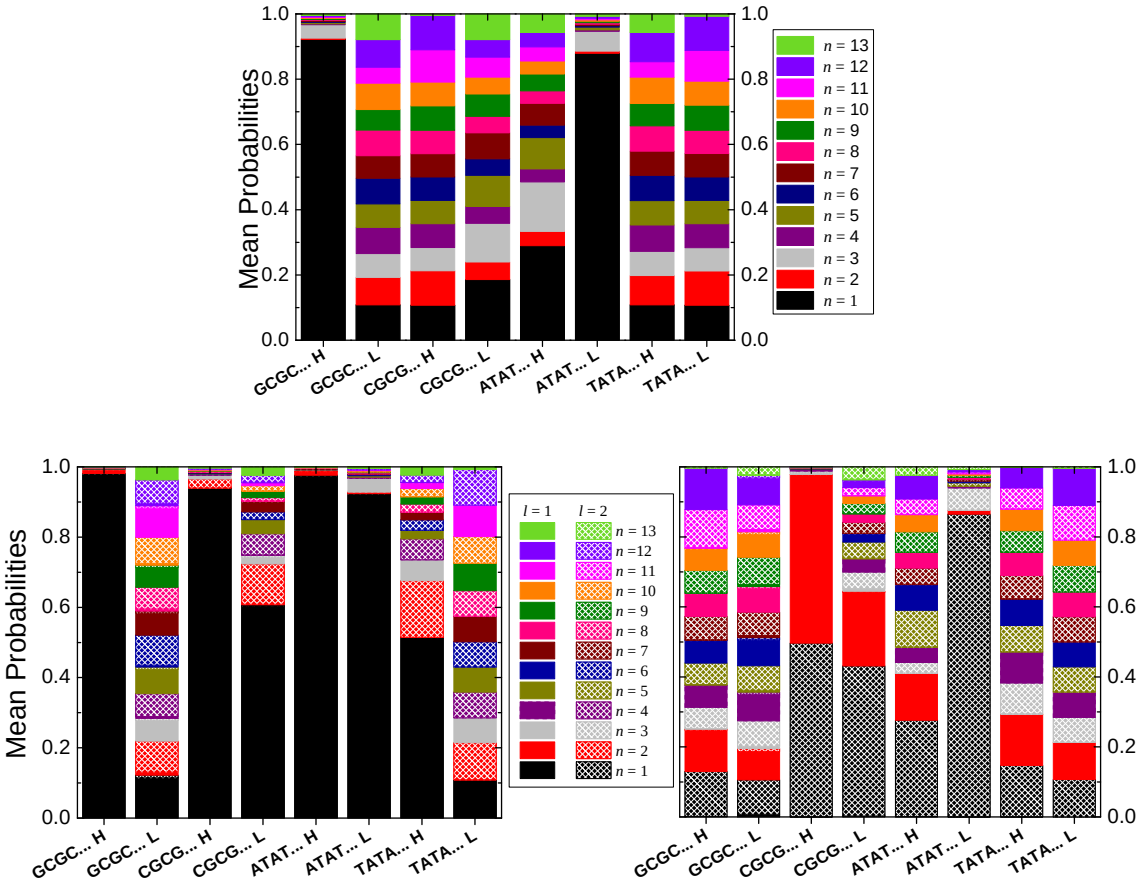

Figure 8.10: Same as in Fig. 8.9, but for $N = 13$.

all μ, Eq. (8.1) leads to identical mean probabilities for (*i*) TCTC... and GAGA..., (*ii*) CTCT... and AGAG..., (*iii*) ACAC... and GTGT..., and (*iv*) CACA... and TGTG.... Also, within the WM, we observe that the carrier moves preferably through the monomers that are identical with the one it was initially placed at, i.e., from the first monomer to the third, and so forth. Within the ELM, it carrier moves preferably through the bases that are identical with the one it was initially placed at, i.e., through the same strand from the one or the other base of the first monomer to the identical base of the third monomer, and so forth, i.e., carrier movement is mainly of inter-strand character. Both TB approaches show that the mean probability to find the carrier at the last monomer is bigger for N odd than for N even, cf. Figs. 8.11-8.12.

8.1.3 Frequency content

We mentioned above that within the WM, for I1 and I2 polymers, all eigenvalues are symmetric around the on-site energy of the base pairs (cf. Theorem 2). Hence, the total number of frequencies involved in charge transfer is $\frac{N^2-1}{4}$ for N odd and $\frac{N^2}{4}$ for N even. For D2 polymers with N even, the eigenvalues are symmetric around

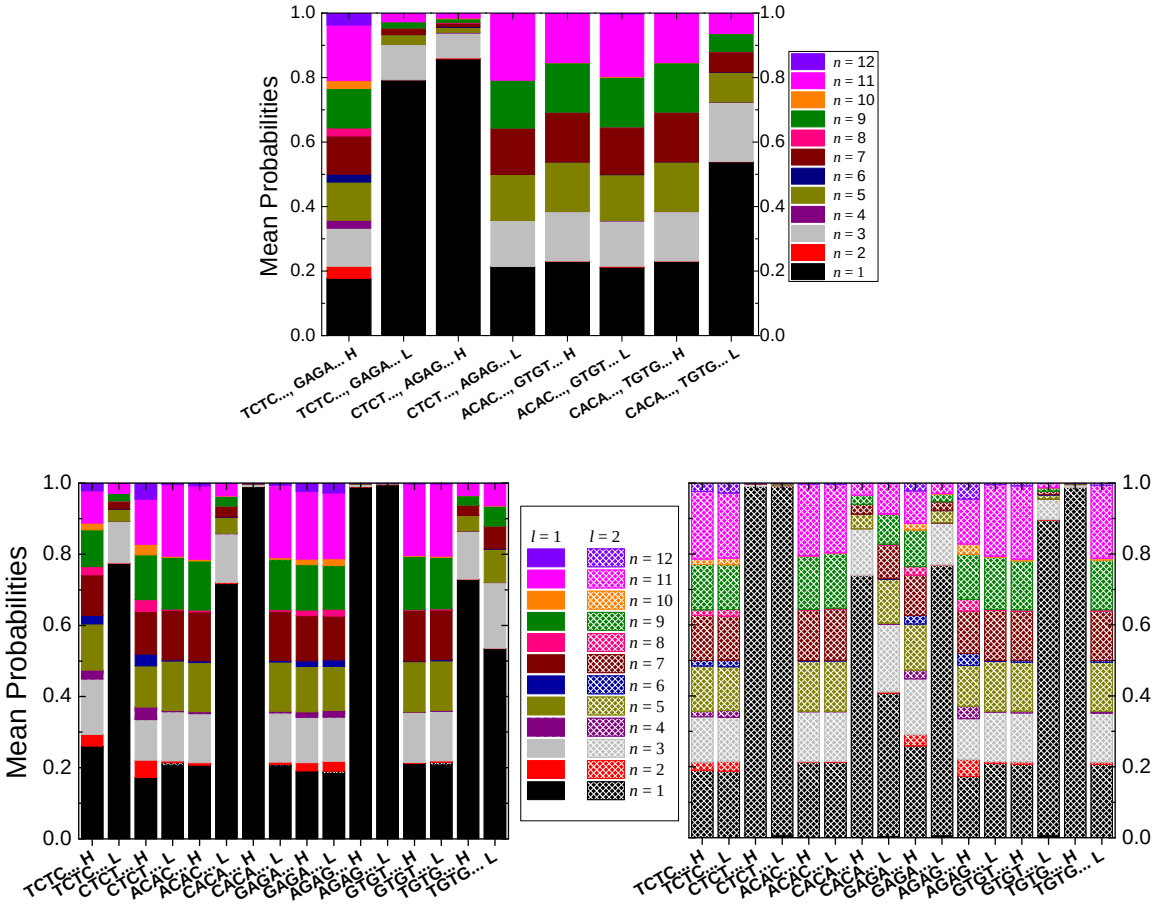

Figure 8.11: Mean over time probabilities to find an extra hole (HOMO) or electron (LUMO) at each base pair, within the WM (top), or at each base, within the ELM (bottom left: initial condition 1, bottom right: initial condition 2) for all D2 polymers, for $N = 12$.

$\frac{E_1^{bp}+E_2^{bp}}{2}$, hence the total number of frequencies is $\frac{N^2}{4}$, too. For type D2 polymers with N odd, the eigenvalues include E_1^{bp} and the total number of frequencies is $\frac{(N-1)(N+3)}{4}$. Within the ELM, there are no symmetries like those mentioned for the WM, hence the total number of frequencies for all three types of polymers is $N(2N-1)$. From Eq. (6.15) it follows that all the palindromicity and equivalence properties presented in the above subsection for the mean over time probabilities, $\langle|\Psi_j(t)|^2\rangle$, hold for the Fourier spectra, $|\mathcal{F}_j(f)|$, too. In the following, we focus on the Fourier spectra that correspond to charge transfer from the first to the last monomer i.e. on $|\mathcal{F}_1(f)|$ and $|\mathcal{F}_N(f)|$ for the WM, and on $|\mathcal{F}_1(f)|$, $|\mathcal{F}_2(f)|$, $|\mathcal{F}_{2N-1}(f)|$ and $|\mathcal{F}_{2N}(f)|$ for the ELM. Both TB approaches show that the frequency content is mainly in the THz domain, cf. Figs. 8.13, 8.14, 8.15, for examples of I1, I2, and D2 polymers.

For I1 polymers, the main frequencies within the WM are in the range $\approx$ 0.3 - 97 THz. Within the ELM, they are in the range $\approx$ 0.1 - 110 THz. The main frequency content is between the FIR the MIR. As an example, we show in Fig. 8.13 the Fourier spectra, at the first and the last monomer, of an extra hole in poly(dA)-poly(dT) with

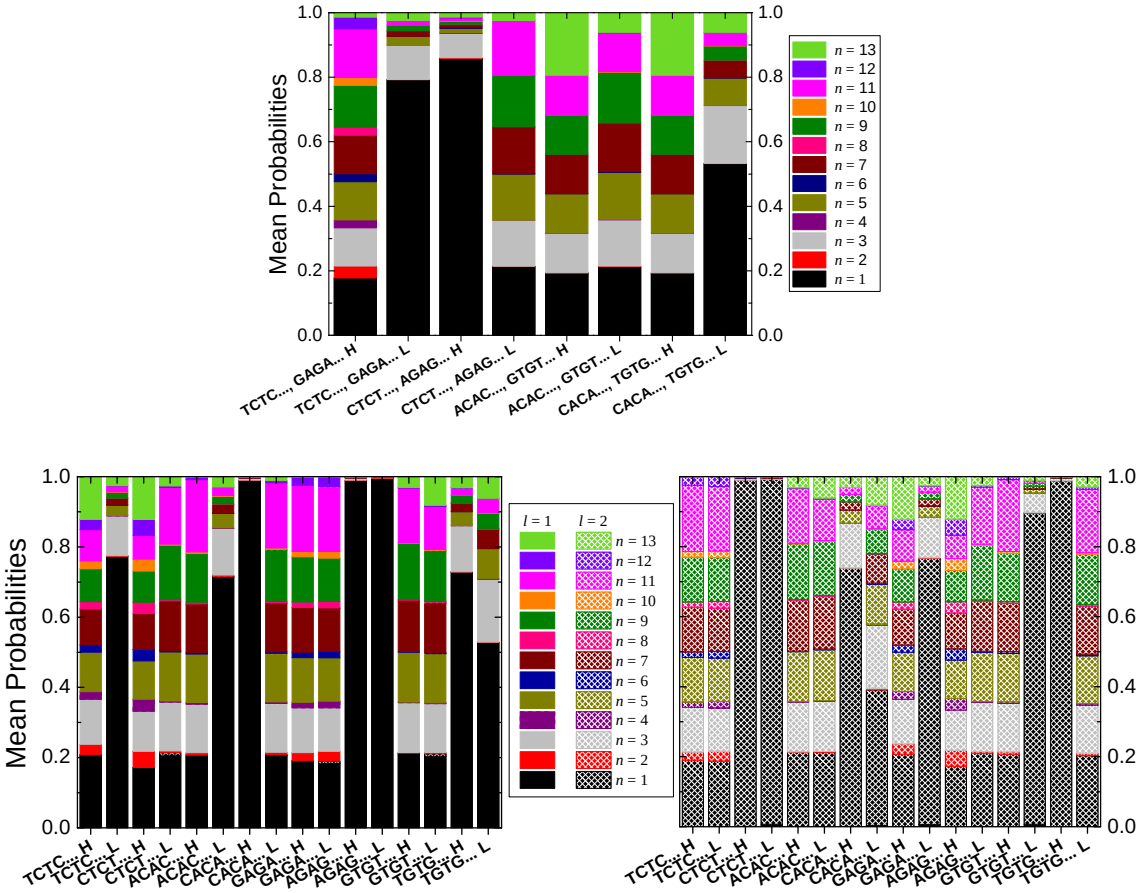

Figure 8.12: Same as in Fig. 8.11, but for $N = 13$.

$N = 20$. From the top panel of Fig. 8.13, we observe that the Fourier amplitudes for the first and the last monomer are equal, mirroring the efficient hole transfer in poly(dA)-poly(dT), cf. also Fig. 8.8. The same conclusion can also be drawn from the bottom panels of Fig. 8.13, for both initial conditions; additionally, the ELM underlines the intra-strand character of carrier transfer and shows that initial conditions 1 and 2 lead to similar form of Fourier spectra.

For I2 polymers, the main frequencies within the WM are in the range $\approx$ 0.01 - 40 THz, i.e., between microwaves and the MIR. Within the ELM, they are in the range $\approx$ 0.01 - 210 THz, i.e., between microwaves and the NIR. As an example, we show in Fig. 8.14, the Fourier spectra, at the first and the last monomer, of an extra electron in ATAT... with $N = 14$. From the top panel of Fig. 8.14 we observe that the Fourier amplitudes for the first and the last monomer are approximately equal, mirroring the finally large electron transfer in ATAT... for N even, cf. also Fig. 8.9. However, this large transfer is very slow, its main frequency is very small but with a large amplitude. The same conclusion can also be drawn from the bottom panels of Fig. 8.14, where we can additionally observe the inter-strand character of charge transfer and that initial conditions 1 and 2 lead to similar form of Fourier spectra.

For D2 polymers, the main frequencies within the WM are in the range $\approx$ 0.4

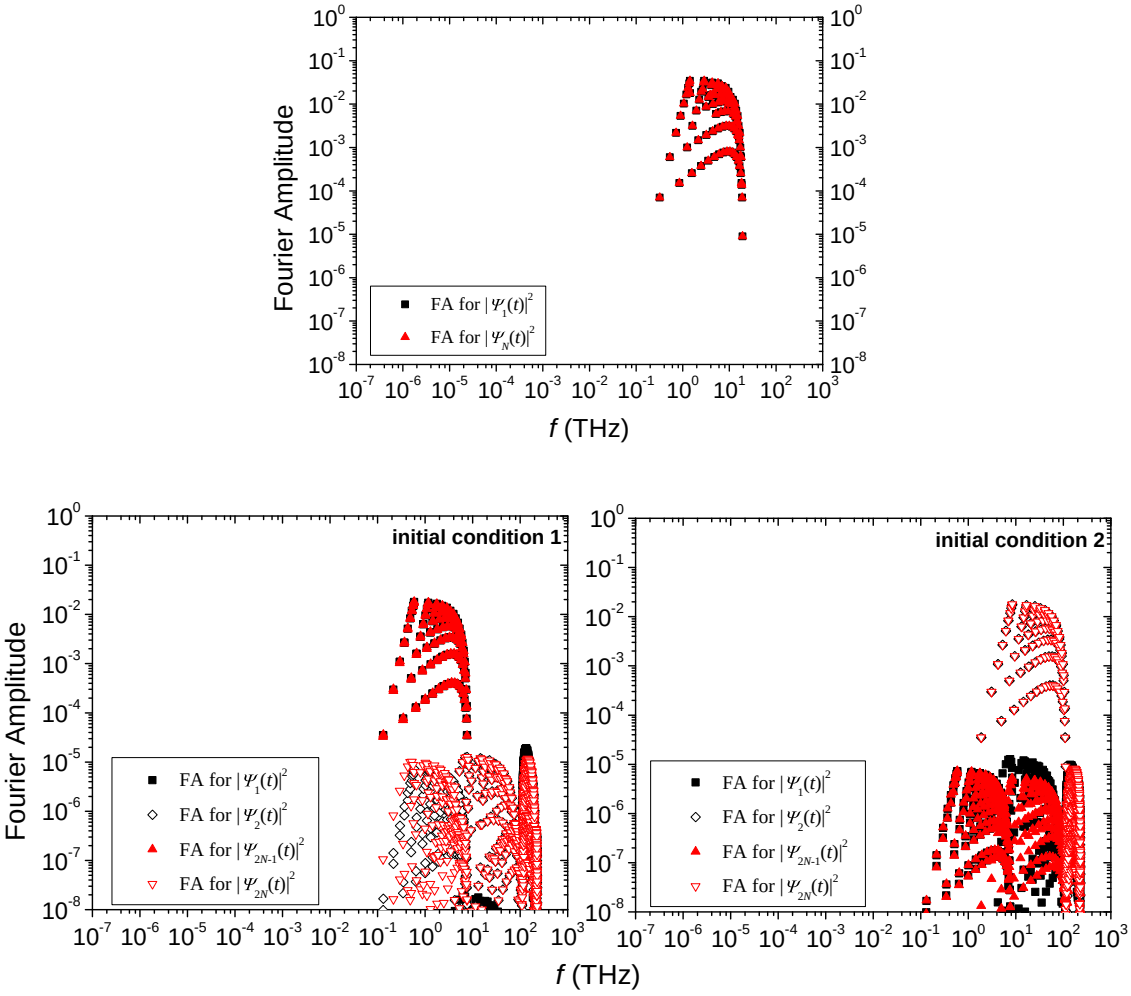

Figure 8.13: Hole transfer Fourier spectra at the first and the last monomer (WM, top) or at the bases of the first and the last monomer (ELM, bottom) for an example of I1 polymers, poly(dA)-poly(dT) with $N = 20$.

GHz - 40 THz, i.e., between radiowaves and the MIR. Within the ELM, they are in the range $\approx$ 0.02 - 190 THz, i.e., between microwaves and the FIR. As an example, we show in Fig. 8.15 the Fourier spectra, at the first and the last monomer, of an extra hole in TCTC... with $N = 21$.From the top panel of Fig. 8.15 we observe that the Fourier amplitudes for the first monomer are much larger than the ones for the last monomer, mirroring the inefficient hole transfer in TCTC... for N odd, cf. also Fig. 8.12. From the bottom panels of Fig. 8.15, we can additionally observe the intra-strand character of charge transfer and that initial conditions 1 and 2 lead to somehow different form of Fourier spectra, initial condition 1 being more efficient than initial condition 2 for hole transfer, cf. also Figs. 8.12.

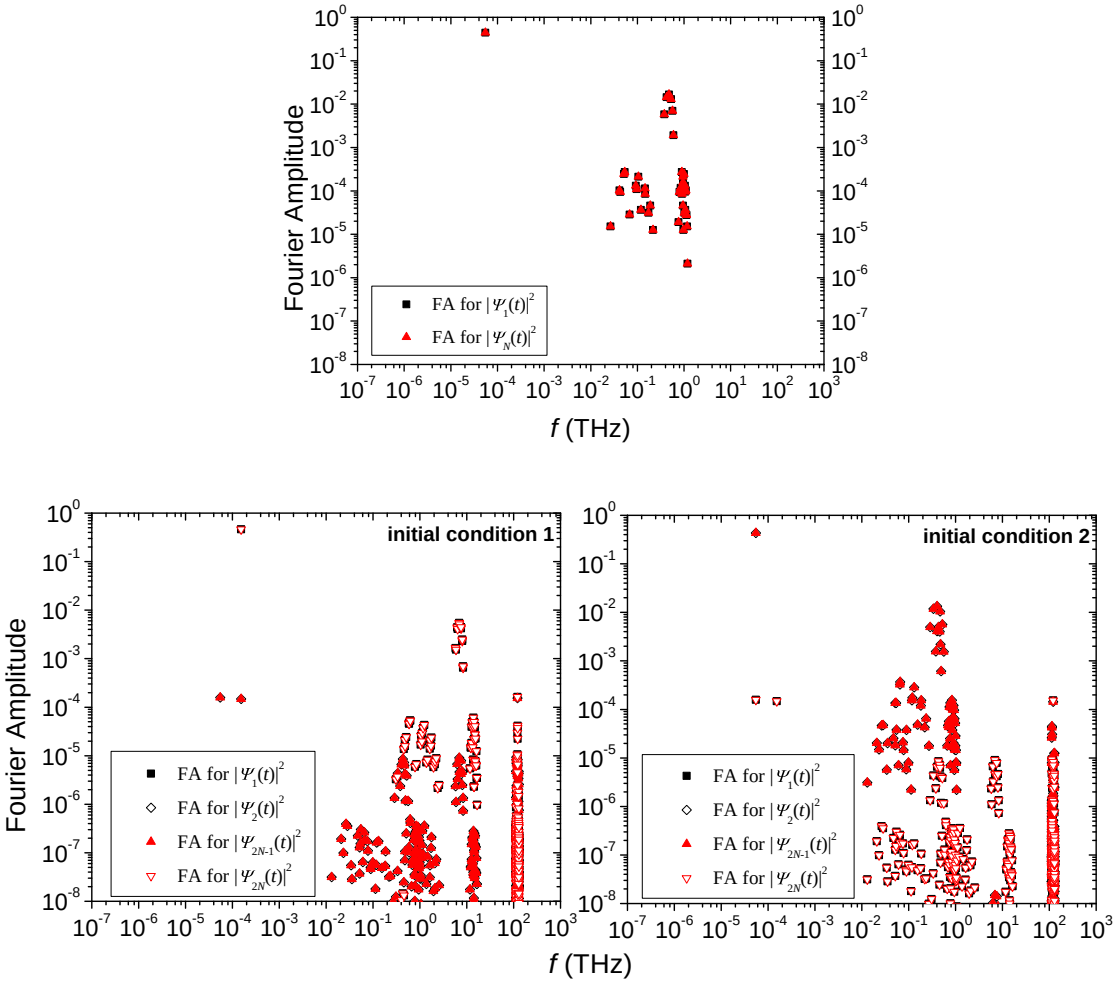

Figure 8.14: Electron transfer Fourier spectra at the first and the last monomer (WM, top) or at the bases of the first and the last monomer (ELM, bottom) for an example of I1 polymers, ATAT... with $N = 14$.

8.1.4 Pure mean transfer rates

As previously, we focus on pure mean transfer rates between the first and the last monomer, within either the WM or the ELM.

As a characteristic example of I1 polymers, we present in Fig. 8.16 the hole pure mean transfer rates for poly(dG)-poly(dC), from the first to the last monomer, within the WM and the ELM. Specifically, (I) for the WM we illustrate $k_{1,N}$ on the left panel, and (II) for the ELM we illustrate the pairs $k_{1,2} = k_{2N-1,2N}$ (cf. Eq. 6.19), $k_{1,2N-1}$ and $k_{2,2N}$, $k_{1,2N}$ and $k_{2,2N-1}$, on the right panel. We have already noticed that, within the ELM, carrier transfer is almost exclusively of intra-strand character. Hence, $k_{1,2N-1} \approx k_{2,2N}$ are the largest transfer rates. Comparing $k_{1,N}$ for the WM with $k_{1,2N-1} \approx k_{2,2N}$ for the ELM, we observe an excellent agreement, both qualitatively and quantitatively. Within the ELM, the intra-base-pair rates $k_{1,2} = k_{2N-1,2N}$ are small and the inter-strand rates $k_{1,2N} \approx k_{2,2N-1}$ insignificant. Increasing N, the intra-strand transfer

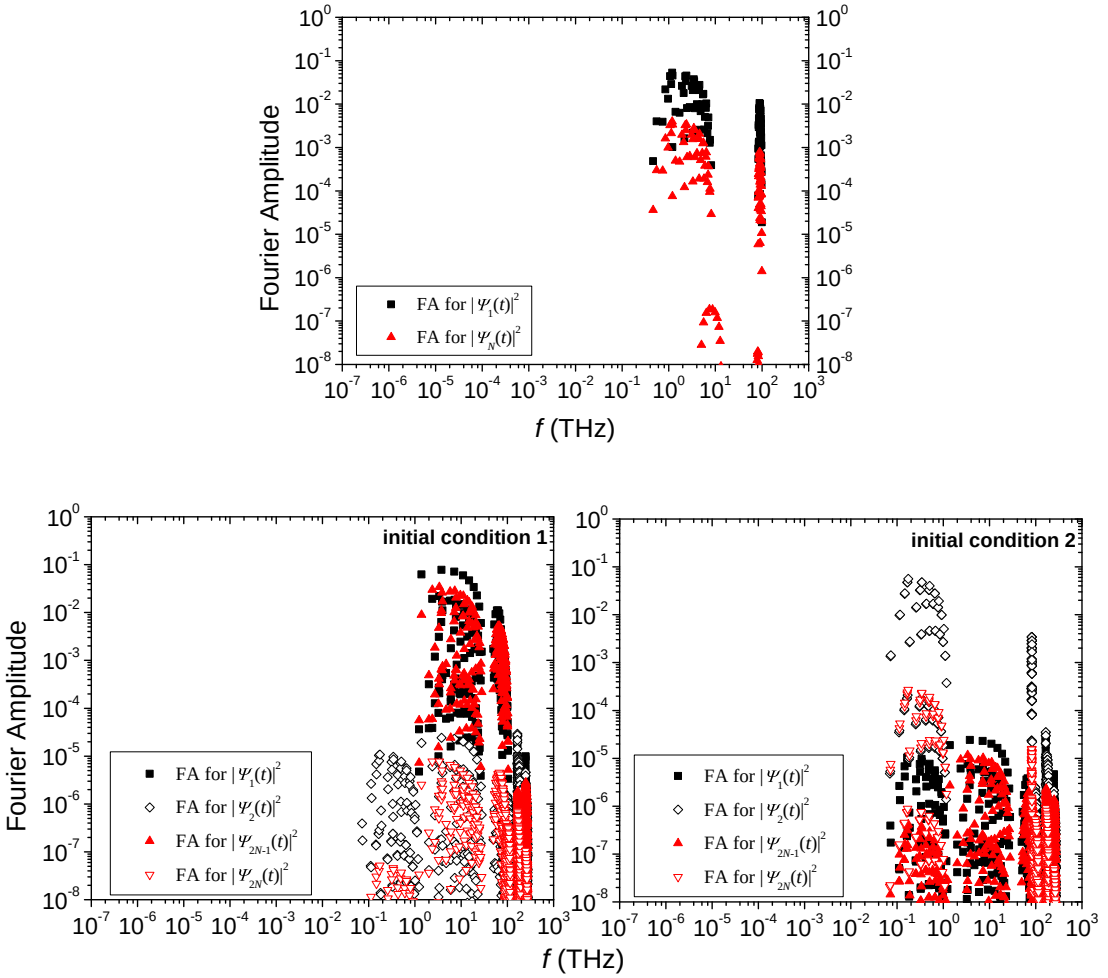

Figure 8.15: Hole transfer Fourier spectra at the first and the last monomer (WM, top) or at the bases of the first and the last monomer (ELM, bottom) for an example of I1 polymers, TCTC... with $N = 21$.

rates $k_{1,2N-1} \approx k_{2,2N}$ decrease reaching gradually the level of the the intra-base-pair rates $k_{1,2} = k_{2N-1,2N}$, at which point, finally, charge transfer along the polymer is insignificant. Increasing N, the insignificant inter-strand rates $k_{1,2N} \approx k_{2,2N-1}$ also gradually decrease further.

We have already mentioned that both TB approaches predict that for some cases of I2 polymers, for N even, the carrier is transferred at a large percentage to the last monomer but the transfer is very slow. Such a case is presented in Fig. 8.17. Specifically, we show the electron pure mean transfer rates for ATAT..., from the first to the last monomer, the WM and the ELM. Specifically, (I) for the WM, we illustrate $k_{1,N}$ on the left panel, and (II) for the ELM, we illustrate the largest transfer rates on the right panel. We have already demonstrated that, within the ELM, the extra carrier is transferred almost exclusively crosswise, through identical bases. Hence, for the ELM, the largest transfer rates are $k_{1,2N-1}$ and $k_{2,2N}$ for N odd, and $k_{1,2N}$ and $k_{2,2N-1}$ for N even. We depict these largest transfer rates in the right panel of Fig. 8.17. In other

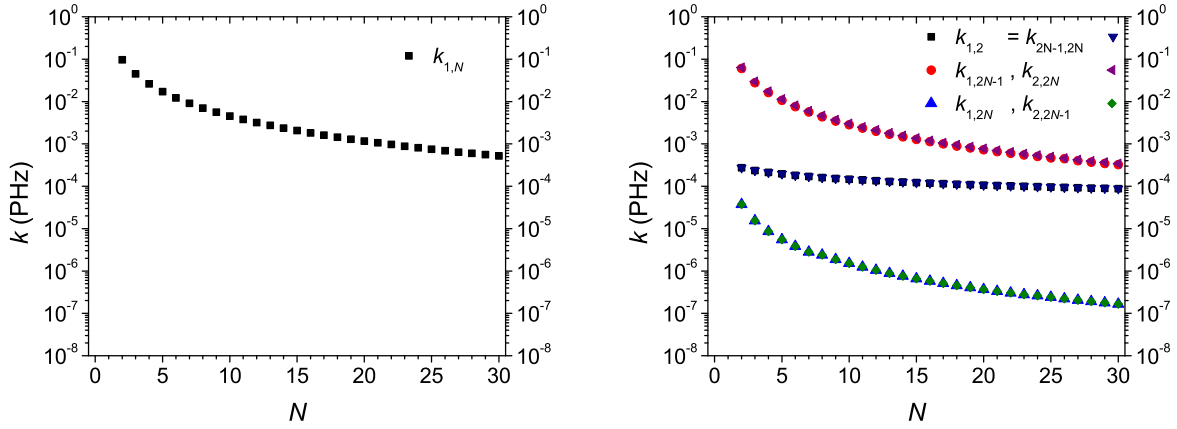

Figure 8.16: Hole pure mean transfer rates (I) $k_{1,N}$ for the WM (left), and (II) $k_{1,2}$, $k_{2N-1,2N}$, $k_{1,2N-1}$, $k_{2,2N}$, $k_{1,2N}$, and $k_{2,2N-1}$ for the ELM (right), for an example of I1 polymers, poly(dG)-poly(dC).

cases of I2 polymers the pure mean transfer rates fall over N in a different manner, which is somehow similar to D2 polymers (see below).

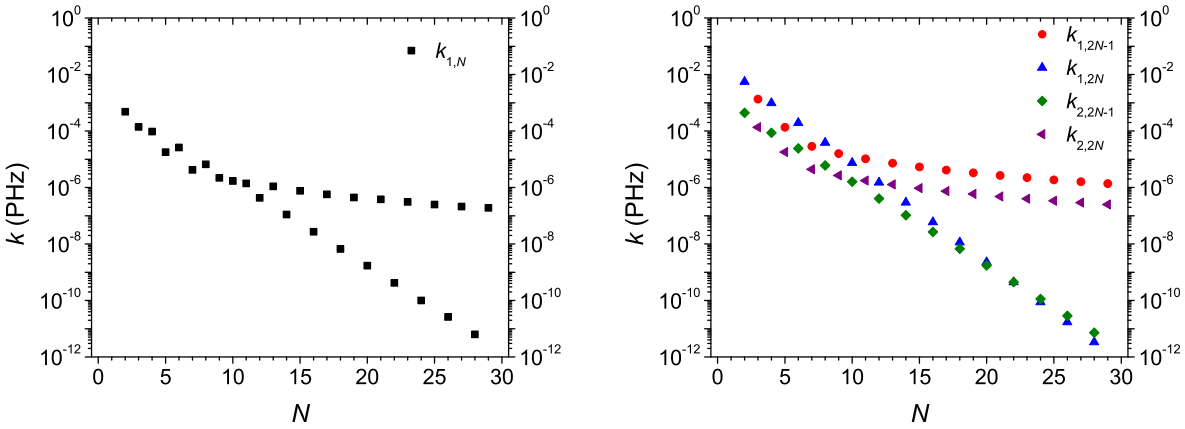

Figure 8.17: Electron pure mean transfer rates (I) $k_{1,N}$ for the WM (left), and, $k_{1,2N-1}$ for N odd, $k_{1,2N}$ for N even, $k_{2,2N}$ for N odd, $k_{2,2N-1}$ for N even, for an example of I2 polymers in which, for N even, the carrier is transferred at a large percentage to the last monomer but the transfer is very slow, ATAT....

As a characteristic example of D2 polymers, we present in Fig. 8.18 the hole pure mean transfer rates for TCTC..., from the first to the last monomer, within the WM and the ELM. Specifically, (I) for the WM we illustrate $k_{1,N}$ on the left panel, and (II) for the ELM we illustrate $k_{1,2N-1}$ and $k_{2,2N}$ on the right panel. We have already mentioned that, within the ELM, the extra carrier is transferred almost exclusively through the strand it was initially placed at, i.e., for D2 polymers the charge transfer is mainly of intra-strand character. Hence, for the ELM, we show $k_{1,2N-1}$ and $k_{2,2N}$, which are the largest transfer rates. Above, we demonstrated that, for I1 polymers, $k_{1,2N-1} = k_{2,2N}$; as shown in Fig. 8.18, this does not hold for D2 polymers.

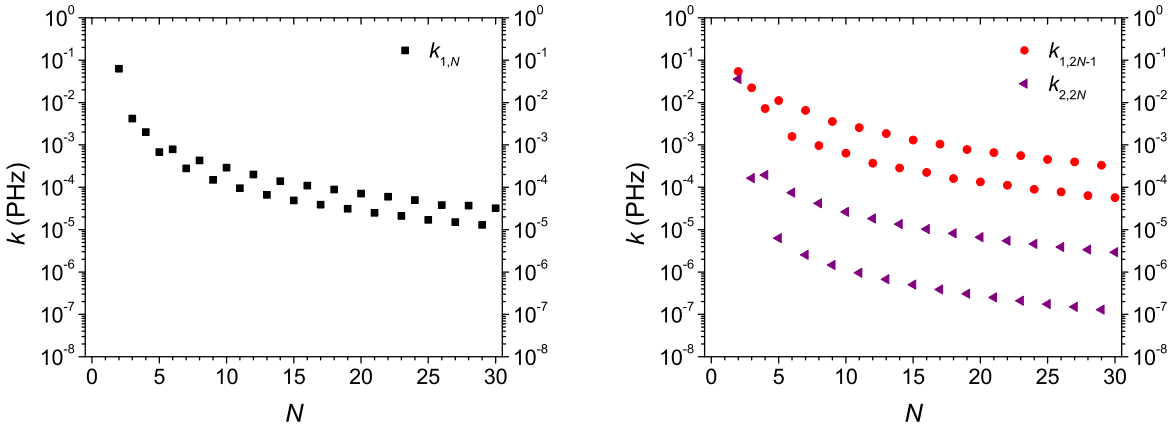

Figure 8.18: Hole pure mean transfer rates $k_{1,N}$ for the WM (left), and $k_{1,2N-1}$, $k_{2,2N}$ for the ELM (right), for an example of D2 polymers, TCTC....

8.2 Polymers with increasing repetition unit

After our comparative study of all possible I1, I2, and D2 polymers within the WM and the ELM, we move on to the study of DNA polymers with increasing repetition unit. As demonstrated in Sec. 8.1, the two TB models give coherent, complementary results, so now we will focus on the simpler WM using the S parametrization (Table A.1). The types of segments studied are presented in Table 8.2.

Table 8.2: The types of polymers mentioned in this Section. I (D) denotes polymers made of the identical (different) monomers. u is the number of monomers in the repetition unit. We only mention the $5' - 3'$ strand.

(I,D)u	sequence example
I1	G... or A...
I2	GC...
I3	GGC...
I4	GGCC...
I6	GGGCCC...
I8	GGGGCCCC...
I10	GGGGGCCCCC...
I20	GGGGGGGGGGCCCCCCCCCC...
D2	GA...
D4	GGAA...
D6	GGGAAA...
D8	GGGGAAAA...
D10	GGGGGAAAAA...
D20	GGGGGGGGGGAAAAAAAAAA...

8.2.1 Energy structure

In Figs. 8.19 and 8.20, we show the HOMO and LUMO eigenspectra of [I2, I4, I6, I8, I10, I20 and I1 polymers] and [D2, D4, D6, D8, D10, D20 and I1 (G...), I1 (A...) polymers], and in Figs. 8.21 and 8.22 we plot the corresponding DOS. The HOMO and LUMO bands of each polymer consist of u subbands, e.g., for I6 or D6 polymers, the number of subbands is 6. Some eigenenergies protrude periodically from the subbands at certain relationships between N and u. At the limits of subbands, vHS occur. The subbands are separated by small energy gaps, which, increasing u, decrease.

For I polymers (cf. Figs. 8.19, 8.21), all eigenvalues are symmetric around the monomer on-site energy and for N odd the trivial eigenvalue, equal to the monomer on-site energy, exists (Theorem 2). Increasing u, the eigenspectra tend to the eigenspectra of I1 polymers, and, of course, the DOS tends to the DOS of I1 polymers.

For D polymers (cf. Figs. 8.20, 8.22), increasing u, the eigenenergies gather around the two monomer on-site energies. Increasing u, the eigenspectra gather within the limits defined by the union of eigenspectra of I1 (G...) and I2 (A...) polymers. In Fig. 8.21, increasing u, the subbands become narrower, but they are wide enough (e.g. for I10 of a few meV) so that their DOS minima remain low enough; therefore we don't have to change the vertical presentation scale. In Fig. 8.22, the already very narrow subbands, increasing u, become even narrower (e.g for D10 two to nine orders of magnitude narrower than for I10), which drives the DOS minima in each subband much higher; therefore, to depict the DOS, we have to increase the vertical presentation scale. Finally, we remind that the eigenvalues for $N = mu$ can be obtained analytically and recursively with the help of the Chebyshev polynomials of the second kind, via the Transfer Matrix Method (see Chapter 3).

We now turn our attention to the HOMO-LUMO gaps. At the large-N limit, increasing u, the gaps of I2, I4, I6, ... polymers approach the gap of I1 polymer (cf. left panel of Fig. 8.23). Indeed, increasing the repetition unit in the mode GC, GGCC, GGGCCC, ..., finally results in a G...GC...C polymer which is almost G... with just a switch from G to C at the middle of the polymer. Hence, at the large-N limit, the energy gap of I1 polymers is the smallest of these series of polymers. For the same reason, increasing u, the eigenspectra and the DOS of I2, I4, I6, ... polymers tend to the eigenspectra and the DOS of I1 polymers (cf. Figs. 8.19, 8.21).

At the large-N limit, increasing u, the gaps of D2, D4, D6, ... polymers approach the gap of the union of I1 (G...) and I1 (A...) polymers (cf. right panel of Fig. 8.23), which is ≈ 0.5 eV lower than the gaps of the relevant homopolymers (G..., A...). Increasing the repetition unit in the mode GA..., GGAA..., GGGAAA... and so on, finally results in a G...GA...A polymer which is *energetically* almost a union of separated G... and A... polymers. This happens due to the large difference of G-C and A-T on-site energies in comparison with the t_{GA} hopping integral. Increasing u, the lowering of the energy gap in the case of D polymers [≈ 0.6 (0.7) eV relative to the A-T (G-C) monomer gap] is much bigger than in the case of I polymers [≈ 0.25 eV relative to the G-C monomer gap].

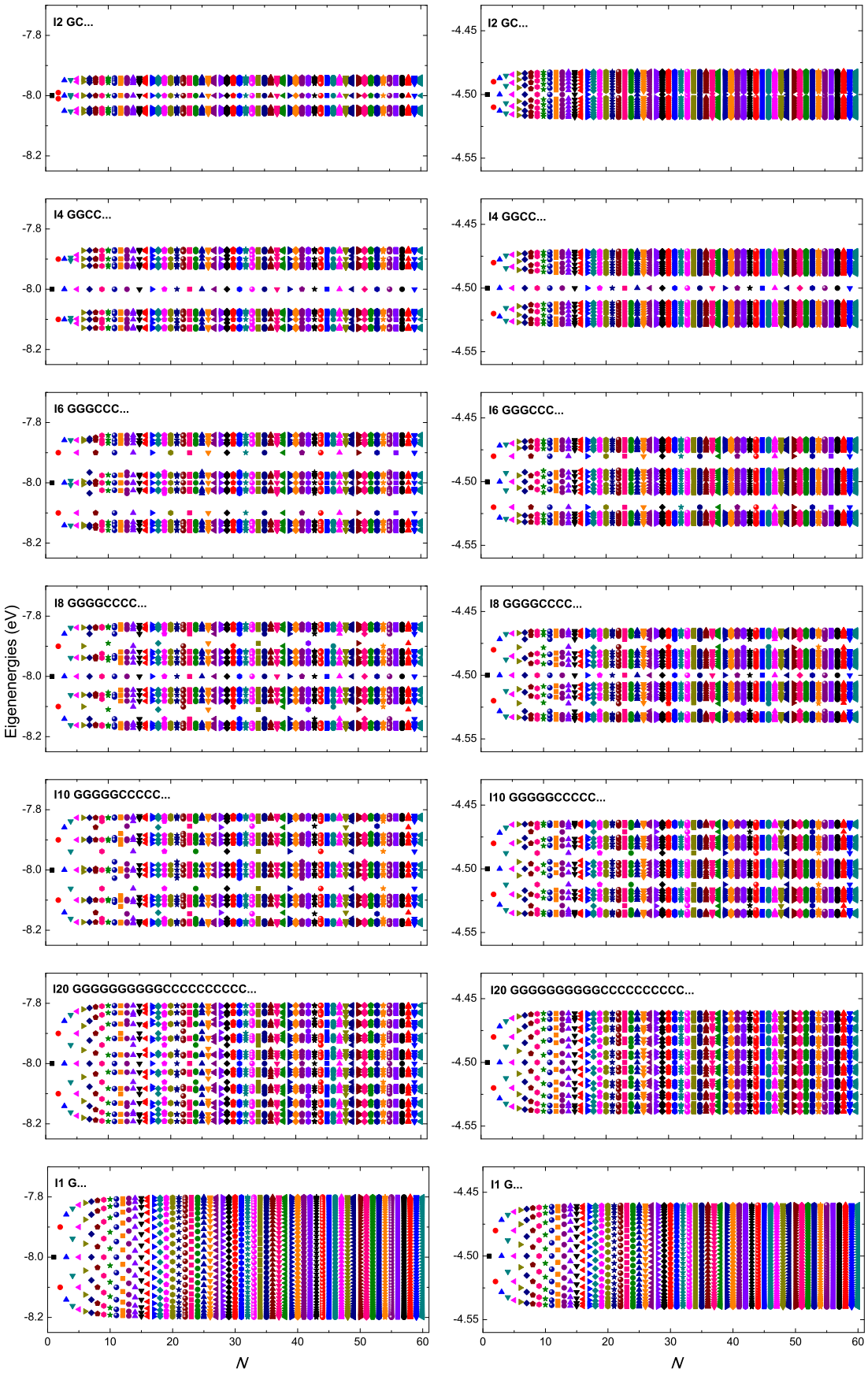

Figure 8.19: Eigenspectra of periodic I polymers. HOMO regime (left) and LUMO regime (right).

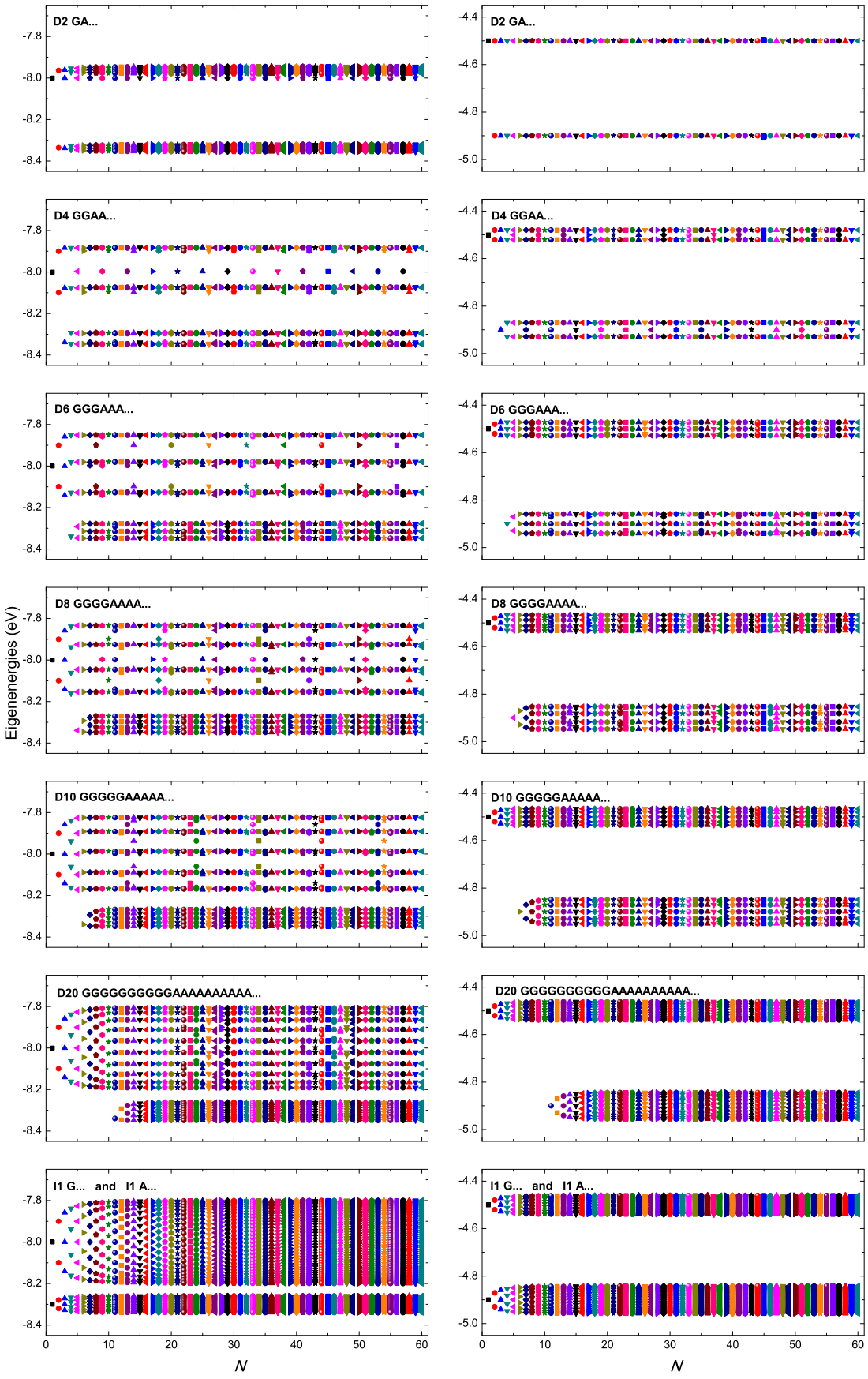

Figure 8.20: Eigenspectra of periodic D polymers as well as of I1 (G...) and I1 (A...) polymers plotted together. HOMO regime (left) and LUMO regime (right).

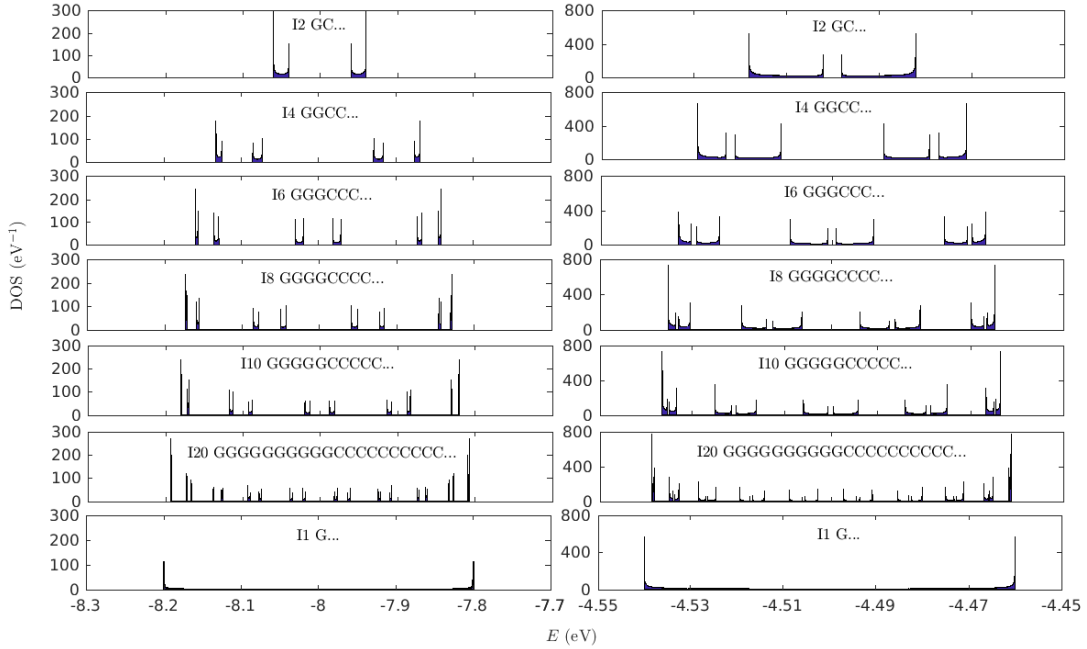

Figure 8.21: DOS of periodic I polymers for the HOMO (left) and the LUMO (right) regime.

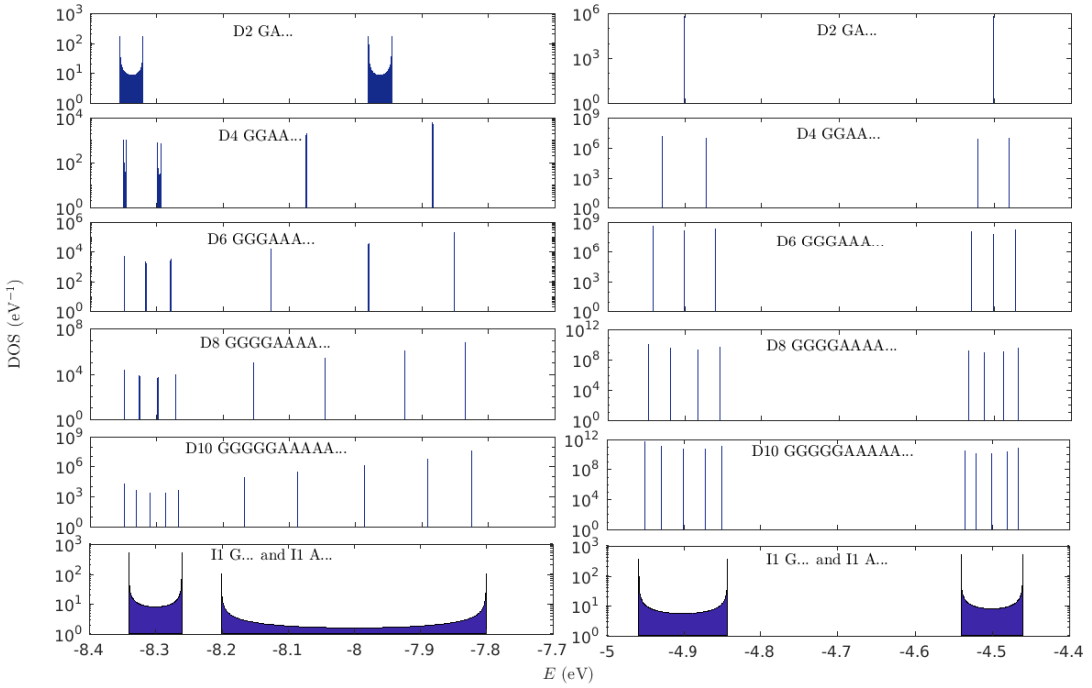

Figure 8.22: DOS of periodic D polymers as well as of I1 (G...) and I1 (A...) polymers plotted together, for the HOMO (left) and the LUMO (right) regime.

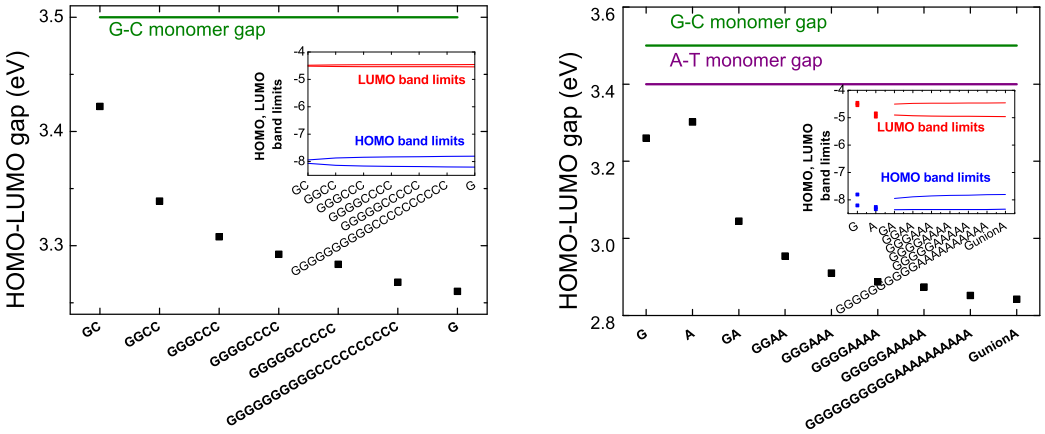

Figure 8.23: HOMO-LUMO gaps. Left: I2 (GC...), I4 (GGCC...), I6 (GGGCCC...), I8 (GGGGCCCC...), I10 (GGGGGCCCCC...), I20 (GGGGGGGGGGCCCCCCCCCC...) polymers, as well as I1 (G...) polymers. The horizontal green line at 3.5 eV shows the energy gap of G-C base pair. Inset: HOMO and LUMO band upper and lower limits. Other variants of I polymers follow the same trend, e.g. the gap of I3 (GGC...) is $\approx$ 3.36 eV. Right: D2 (GA...), D4 (GGAA...), D6 (GGGAAA...), D8 (GGGGAAAA...), D10 (GGGGGAAAAA...), D20 (GGGGGGGGGGAAAAAAAAAA...) polymers, as well as the union of I1 (G...) and I1 (A...) polymers. The horizontal green line at 3.5 eV (purple line at 3.4 eV) shows the energy gap of the G-C (A-T) base pair. Inset: HOMO and LUMO band upper and lower limits.

8.2.2 Mean over time Probabilities

The main aspects of our results for the mean over time probabilities for I2, I4, ... polymers are summarized in Figs. 8.24 and 8.25 for some example cases. For N equal to natural multiples of u ($N = mu, m \in \mathbb{N}^*$), *palindromicity* holds, i.e., the probabilities are palindromic. This is due to the fact that for $N = mu$, the Hamiltonian matrices of these polymers are palindromic, i.e. reading them from top left to bottom right and vice versa gives the same result. The palindromicity for $N = mu$ is shown in Fig. 8.24, for an example I8 (GGGGCCCC) polymer, for all possible initial placements of an extra hole. It is evident that palindromicity holds for all initial conditions. Hence, in these polymer cases, the appropriate choice of the monomer the carrier is injected to, can lead to enhanced presence at specific sites at its other end, leading to more efficient transfer. For $N \neq mu$, palindromicity is lost. This is shown in Fig. 8.25, for an example case of an I6 (GGGCCC) polymer. In the HOMO regime, all studied polymers with $N \neq mu$, show increased mean (over time) probabilities at the $\frac{u}{2}$ initial monomers. For example, for type I6 polymers, for $N \neq 6m$, we have increased probabilities at the first, second and third monomer (left panel of Fig. 8.25). This property is so evident in the HOMO regime due to the magnitude of the hopping integrals: t_{GG} is the greater of all, and t_{GC} is much smaller than t_{CG}. In the LUMO regime, this property cannot be clearly seen, because t_{GG} is the greater of all, but t_{CG} and t_{GC} have similar values. For the same reason, in the LUMO regime, for $N = mu + \frac{u}{2}$, we have an almost palindromic behavior (see e.g. the right panel of Fig. 8.25). As seen previously, for I1 polymers and initial placement of the carrier at the first monomer, the mean over time probabilities are given by Eq. (8.3). Increasing u, the relevant probabilities of I2, I4, I6, ... polymers tend to the I1 probabilities.

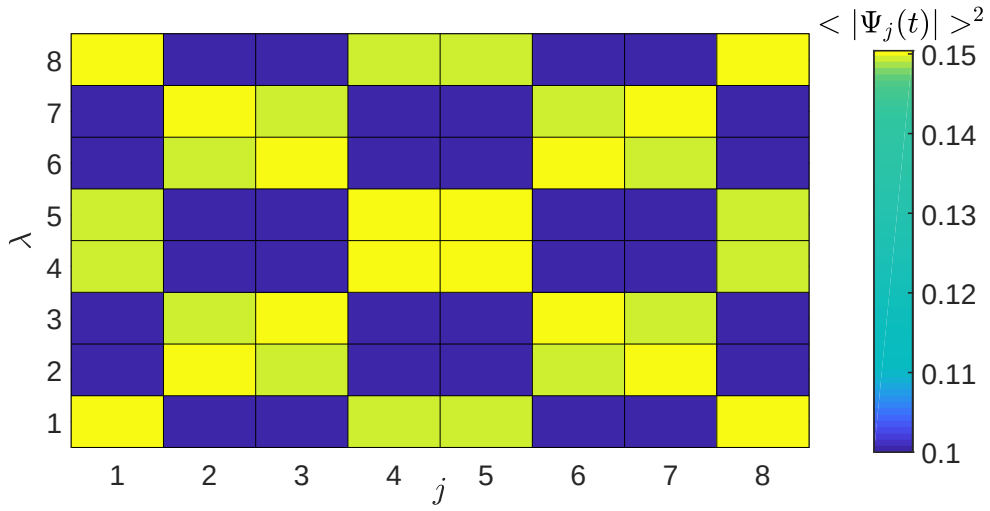

Figure 8.24: Mean (over time) probabilities to find an extra hole in a GGGGCCCC polymer for all possible initial placements.

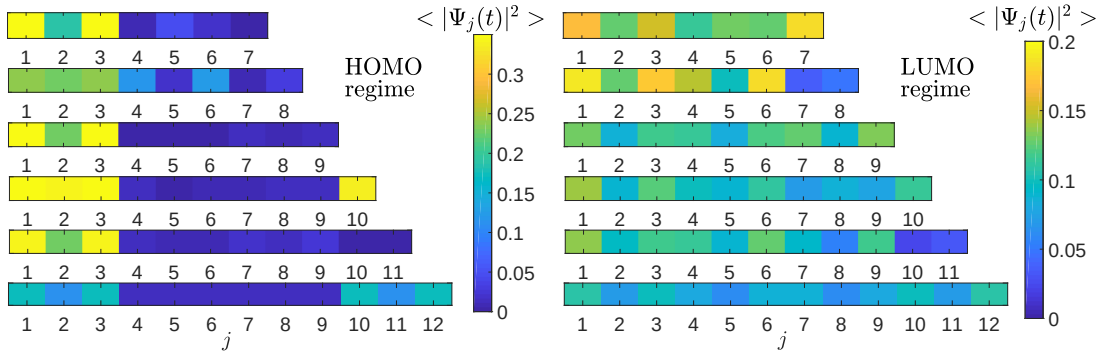

Figure 8.25: Mean over time probabilities to find an extra hole (left) and electron (right), initially placed at the first monomer, in a GGGCCC... polymer ($u = 6$) made up of $N = u + \tau, \tau = 1, \ldots, u$ monomers.

The main features of our results for the mean over time probabilities for D2, D4, ... polymers are summarized in Fig. 8.26 for the example case of an I8 (GGGGAAAA) polymer, for all possible initial placements of an extra hole. A basic observation for polymers made of different monomers is that if we initially place the carrier at a G-C monomer the probability to find it at an A-T monomer is small, and vice versa.

Detailed numerical results displaying all the above mentioned features, having placed the hole or electron initially at the first monomer, for $N = u+\tau, \tau = 0, 1, \ldots, u-1$, can be found in the Supplemental Material of Ref. [59], and specifically in Figs. A.1-A.5 for periodic I polymers, and in Figs. A.6-A.10, for periodic D polymers.

8.2.3 Frequency Content

The Fourier spectra of the time-dependent probability to find the carrier at each monomer, are, generally, in the THz regime.

For $N = mu$, $m \in \mathbb{N}^*$, for I1, I2, I4, I6, ... polymers, the Fourier spectra of the

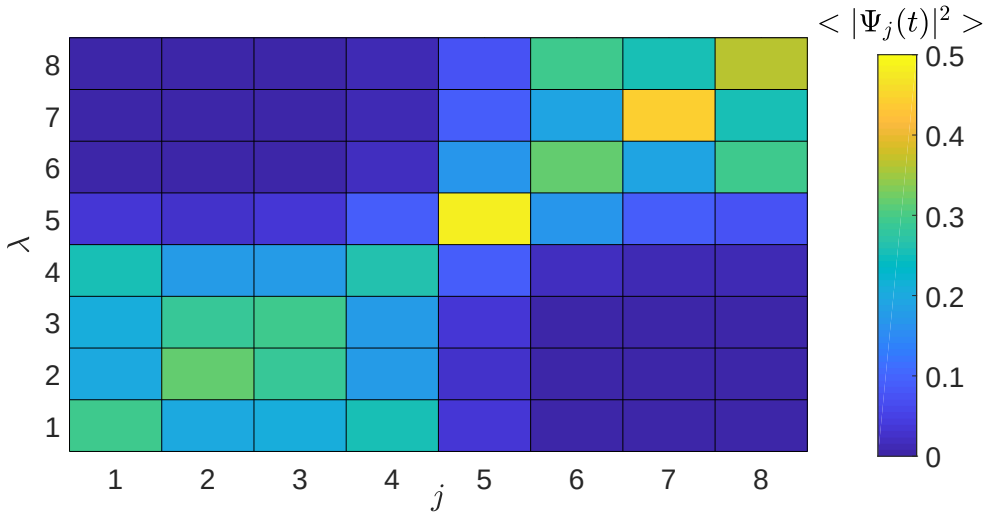

Figure 8.26: Mean (over time) probabilities to find an extra hole in a GGGGAAAA polymer for all possible initial placements.

time-dependent probability to find an extra carrier at the various monomers, either for the HOMO or the LUMO regime, are palindromic, i.e., they are identical for the μ-th and $(N-\mu+1)$-th monomer. This stems from the palindromicity characterizing the Hamiltonian matrices for $N = mu$, $n \in \mathbb{N}^*$. The Fourier spectra of the probability to find an extra carrier at the first and at the last monomer, having placed it initially at the first monomer, for I and D polymers, for the HOMO and the LUMO regime, for $N = u + \tau$, $\tau = 0, 1, \ldots, u-1$, as well as similar diagrams for greater N, can be found in Refs. [259, 260]. Since for $N \neq mu$, $n \in \mathbb{N}^*$ the Hamiltonian matrices are not palindromic, the Fourier spectra are also not palindromic. Preliminary analysis of the frequency content of I1, I2, D2, I3, I4 and I6 polymers, within the WM and the ELM, including the Fourier spectra, the WMFs and the TWMF as a function of N can be found in Ref. [261], with the HKS parametrization.

Next, we focus on the TWMF as a function of N for various types of periodic I polymers (cf. Fig. 8.27). In I2 (GC...) polymers, only two hopping integrals are involved: $t_{\mathrm{GC}}, t_{\mathrm{CG}}$. In I4 (GGCC...), I6 (GGGCCC...), ... polymers, three hopping integrals are involved: $t_{\mathrm{GG}}, t_{\mathrm{GC}}, t_{\mathrm{CG}}$. This is the reason that in the limit of large N, the TWMF for I2 polymers tends to a different frequency region than for I4, I6, ... polymers. For I4, I6, ... polymers, increasing u, the role of t_{GG} gradually increases, hence, this series of polymers has as a limit I1 (G...) polymers, where only one hopping integral is involved: t_{GG}. In particular, the TWMF of I4, I6, ... polymers, in the limit of large N, tends to the TWMF of I1 polymers.

Next, we focus on the TWMF as a function of N for various types of periodic D polymers (cf. Fig. 8.28). In D2 (GA...) polymers, only two hopping integrals are involved: t_{GA}, t_{AG}. In D4 (GGAA...), D6 (GGGAAA...), ... polymers, four hopping integrals are involved: $t_{\mathrm{GG}}, t_{\mathrm{GA}}, t_{\mathrm{AG}}, t_{\mathrm{AA}}$. This is the reason that in the limit of large N, the TWMF for D2 polymers tends to a different frequency region than for D4, D6, ... polymers. For D4, D6, ... polymers, increasing u, the role of t_{GG} gradually increases;

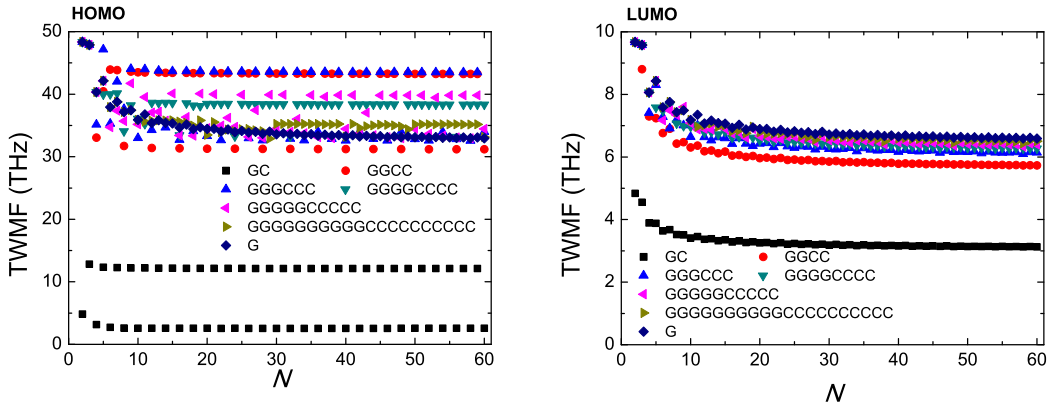

Figure 8.27: Total Weighted Mean Frequency (TWMF) as a function of the number of monomers N in the polymer, having placed the carrier initially at the first monomer, for I1 (G...), I2 (GC...), I4 (GGCC...), I6 (GGGCCC...), I8 (GGGGCCCC...), I10 (GGGGGCCCCC...), and I20 (GGGGGGGGGGCCCCCCCCCC...) polymers, for the HOMO and the LUMO regime.

the same happens with the role of t_{AA}. However, if we place the carrier initially at the first G-C monomer, the probability to find it at any A-T monomer is very small (e.g., cf. Fig. 8.26). Hence, for initial placement of the carrier at a G-C monomer (like in Fig. 8.28), the TWMF of D4, D6, ... polymers, in the limit of large N, tends to the TWMF of I1 (G...) polymers, where only one hopping integral is involved: t_{GG}.

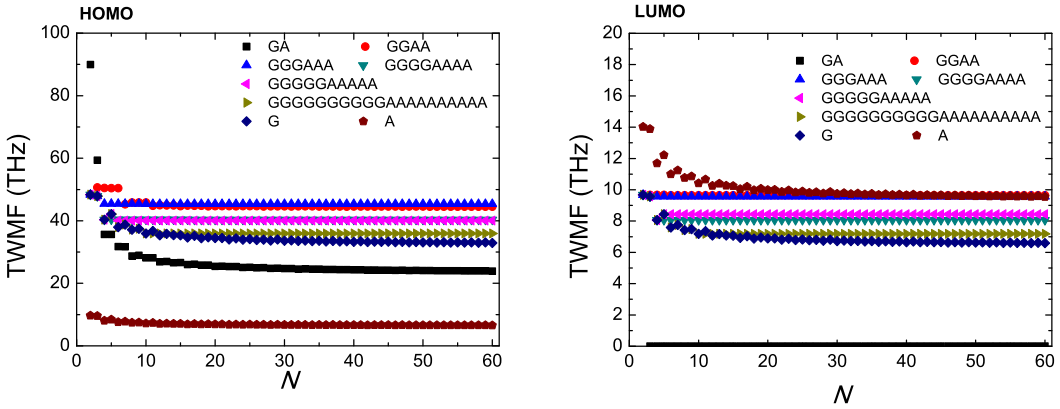

Figure 8.28: Total Weighted Mean Frequency (TWMF) as a function of the number of monomers N in the polymer, having placed the carrier initially at the first monomer, for I1 (G...), I1 (A...), D2 (GA...), D4 (GGAA...), D6 (GGGAAA...), D8 (GGGGAAAA...), D10 (GGGGGAAAAA...), and D20 (GGGGGGGGGGAAAAAAAAAA...) polymers, for the HOMO and the LUMO regime.

The frequencies involved in charge transfer are given by Eq. (6.13). Hence, the maximum frequency is determined by the maximum difference of eigenenergies, i.e., by the upper and lower limits of the HOMO or LUMO band. Since increasing u, the eigenspectra of I2, I4, I6, ... polymers tend to the eigenspectra of I1 polymers, the maximum frequencies of these polymers also tend to the maximum frequency of I1 polymers. Furthermore, since increasing u, the eigenspectra of D2, D4, D6, ... polymers

tend to the eigenspectra of the union of I1 (G...) and I1 (A...) polymers, the maximum frequencies of these polymers also tend to the maximum frequency of the union of I1 (G...) and I1 (A...) polymers. Numerical results regarding the behavior of maximum frequency can be found in Fig. A.11 in the Supplemental Material of Ref. [59].

8.2.4 Pure mean transfer rates

Next, we study the pure mean transfer rates, $k_{1,N}$, from the first to the last monomer. For simplicity, we drop the indices. An impressive case where appropriate sequence choice can increase k by many orders of magnitude is shown in Fig. 8.29, where we depict $k(N)$ for $N = mu$, either for HOMO or for LUMO, for type I1 (G...), I2 (GC...), I4 (GGCC...), I6 (GGGCCC...), I8 (GGGGCCCC...) and I10 (GGGGGCCCCC...) polymers. These polymers are palindromic, hence there is enhanced presence of the extra carrier at the last monomer. Results for any N can be found elsewhere [259].

In all cases, $k(N)$ is a decreasing function. The electron k range is many orders of magnitude narrower than the hole k range, due to the much smaller difference between the hopping integrals (t_{GG}, t_{GC}, t_{CG}) involved.

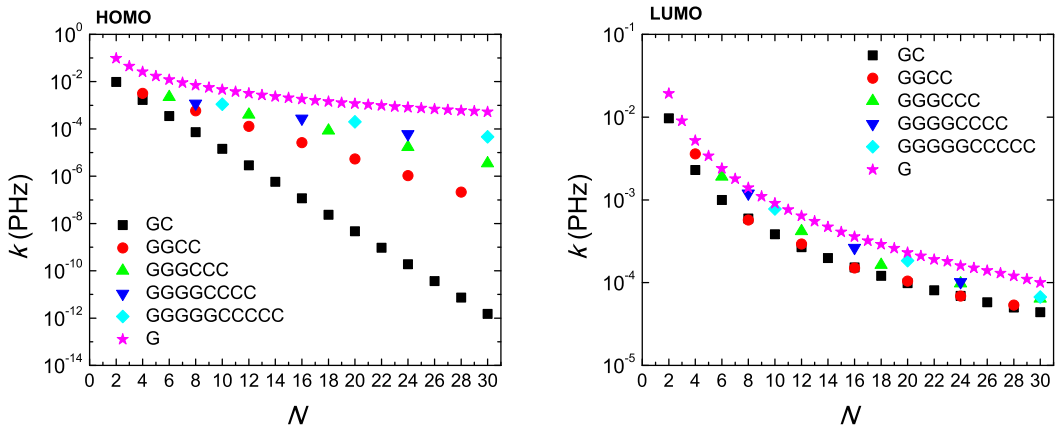

Figure 8.29: *Pure* mean transfer rate k of type I1 (G...), I2 (GC...), I4 (GGCC...), I6 (GGGCCC...), I8 (GGGGCCCC...) and I10 (GGGGGCCCCC...) polymers, as function of the number of monomers N in the polymer, for N equal to natural multiples of their u, for HOMO (upper panel) and LUMO (lower panel).

In Fig. 8.29 we observe that for $N = mu$, starting from type I2 (GC...) polymers and increasing u, i.e. for types I4 (GGCC...), I6 (GGGCCC...), ... polymers, k takes increasingly larger values. In other words, the degree of transfer difficulty is greater for type I2 (GC...) polymers and decreases gradually for types I4, I6, ... I10 polymers. And so it will be if we still increase u taking similar types of polymers. However, $k(N)$ has an upper limit which is $k(N)$ of type I1 polymers. The latter polymers are structurally simpler (more precisely, they have the simplest possible structure), a fact that favors charge transfer along them, so their transfer rates are higher than those of the other polymer types. As u increases, the influence of t_{GC} and t_{CG} becomes less significant, hence this upper limit appears.

Generally, for an extra electron (LUMO), $\ln k(\ln N)$ is approximately linear, of the form $\ln k = \ln k_0 - \eta \ln N$, a relation that generally does not hold for an extra hole (HOMO). However, for type I1 (G...) polymers the above mentioned linear relation holds both for HOMO and LUMO. For HOMO, generally $k(d)$, where $d = (N-1)\times 3.4$ Å is the charge transfer distance, is approximately of the form $\ln k = \ln k_0 - \beta d$, a relation that does not generally hold for an extra electron (LUMO).

To gain further insight, we have performed the exponential fits $k = k_0 e^{-\beta d}$ and $k = A + k_0 e^{-\beta d}$ as well as the power-law fit $k = k_0' N^{-\eta}$. In all cases we studied k from the first to the last monomer, under the condition $N = mu$, $N < 40$. We observe that for I2 (GC...), I4 (GGCC...), I6 (GGGCCC...), I8 (GGGGCCCC...), I10 (GGGGGCCCCC...) polymers the HOMO regime is better characterized by exponential fits and the LUMO regime by power-law fits. For type I1 (G...) polymers the power-law fits are better both for the HOMO and the LUMO regimes. his fact can also be easily seen in Fig. 8.29, where in the HOMO regime all I2, I4, I6, I8, I10 polymers show a linear relation between $\ln k$ and d, while, type I1 polymers do not satisfy this linear relation. In the LUMO regime, none of I2, I4, I6, I8, I10 polymers satisfies a linear relation between $\ln k$ and d. In the LUMO regime, all I2, I4, I6, I8, I10 polymers as well as I1 polymers satisfy an almost linear $\ln k$ - $\ln N$ relation.

For polymers made of different monomers $k(N)$, is depicted in Fig. 8.30. We observe that, while for type I1 polymers (G... and A...) k drops $\approx$ by only 2 to 3 orders of magnitude, increasing N from 2 to 30, as the number of A in the repetition unit increases, $k(N)$ drops dramatically by many more orders of magnitude. Again, this behavior shows that the pure mean transfer rate can be increased by many orders of magnitude by appropriate choice of the repetition unit. All in all, our results suggest that type I1 polymers are the best for electron or hole transfer.

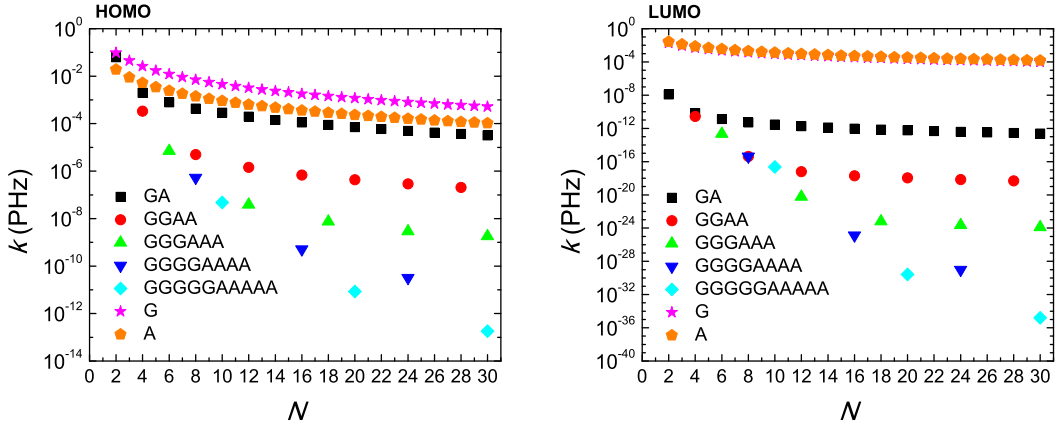

Figure 8.30: *Pure* mean transfer rate k of type I1 (G...), I1 (A...), D2 (GA...), D4 (GGAA...), D6 (GGGAAA...), D8 (GGGGAAAA...) and D10 (GGGGGAAAAA...) polymers, as function of the number of monomers N in the polymer, for N equal to natural multiples of their u, for HOMO (upper panel) and LUMO (lower panel).

8.3 Conclusion

In the first part of this Chapter, we employed two TB approaches, the WM and the ELM, to examine carrier transfer in DNA monomer polymers (I1) and dimer polymers (I2 and D2). The two TB approaches give coherent, complementary results.

For the time-independent problem, we studied the HOMO and the LUMO eigenspectra and the occupation probabilities, the DOS and the HOMO-LUMO gap. We saw that the upper (lower) subband of the HOMO (LUMO) eigenspectrum calculated within the ELM corresponds to the band calculated within the WM. The occupation probabilities within the WM and the ELM show various degrees of palindromicity and eigenspectrum (in)dependence of the probabilities to find the carrier at a site. The DOS displays nice van Hove singularities at the (sub)band edges. As expected, the polymer HOMO-LUMO gaps are smaller than the HOMO-LUMO gaps of the two possible monomers, reaching a level of 3.4 to 3.0 eV. The smallest HOMO-LUMO gaps occur for D2 polymers.

We also investigated the mean over time probabilities to find the carrier at each site (base pair for the WM and base for the ELM), the Fourier spectra and the pure mean transfer rates from a certain site to another. The mean over time probabilities illustrate clearly the basically intra-strand character of carrier transfer in I1 and D2 polymers. However, while in I1 polymers the carrier moves successively through all bases of the same strand, in D2 polymers it moves through the bases that are identical with the one it was initially placed at, i.e., it moves through the same strand from the one or the other base of the first monomer to the identical base of the third monomer, and so forth. Carrier transfer is basically of inter-strand character in I2 polymers. The Fourier spectra give us a nice representation of the frequency content of charge transfer. Both TB approaches show that this frequency content is mainly in the THz domain; the details depend on the type of polymers and the TB approach used. The pure mean transfer rates k show both how fast carrier transfer is and how much of the carrier is transferred from the initial site to the final one. Our results indicate that the fall of k as a function of N becomes generally steeper as the intricacy of the energy structure increases.

Having demonstrated the agreement between the two TB approaches, in the second part of this Chapter, we employed the –simpler– WM to study the transfer of an extra carrier, electron or hole, along N-monomer periodic polymers (made of the same monomer, i.e. I1, I2, I4, I6, I8, I10, I20, as well as made of different monomers, i.e. D2, D4, D6, D8, D10, D20).

For periodic I polymers, we show that the eigenenergies are always symmetric relative to the monomer on-site energy. Increasing the repetition unit, we have witnessed convergence of types I2, I4, I6, ... polymers, to type I1 polymers, in terms of eigenspectra, DOS, energy gaps, mean over time probabilities to find the carrier at the first and last monomers, frequency content (total weighted mean frequency), and pure mean transfer rates, i.e. for all the properties we studied. In other words, increasing the repetition unit, the physical properties of I2, I4, I6, ... polymers tend to those of the relevant homopolymer. The homopolymer has the smallest HOMO-LUMO gap. Gen-

erally, for homopolymers, the magnitude of k is larger and the fall of $k(N)$ is less steep. As the repetition unit increases, the influence of t_{GC} and t_{CG} becomes less significant, hence k of the homopolymer acts as an upper limit. Moreover, we have ascertained palindromicity of physical properties such as the mean (over time) probabilities and the Fourier spectra, when the number of monomers is a natural multiple of the repetition unit.

For periodic D polymers, the eigenenergies gather around the two monomers' on-site energies. Increasing the repetition unit, the gap decreases, converging to the gap of the union of the two relevant homopolymers, which is $\approx$ 0.5 eV lower than the gaps of the relevant homopolymers. As far as the mean probabilities are regarded, if we initially place the carrier at a G-C monomer, the probability to find it at an A-T monomer is small, and vice versa. Increasing the repetition unit, k (from the first to the last monomer) falls dramatically.

Some further general remarks: For both I and D polymers, the frequency content of carrier transfer (in terms of the TWMF) lies within the THz regime. For both I and D polymers, although $k(N)$ is a decreasing function, it can be increased, for the same N, by many orders of magnitude with appropriate sequence choice. The homopolymers (e.g. G... and A...), i.e. the simplest cases, display higher pure mean transfer rates, hence they are more efficient in terms of electron and hole transfer.

9 Deterministic aperiodic and random DNA polymers

This Chapter is devoted to charge transfer in several categories of deterministic aperiodic and in random binary DNA polymers, within the WM. The content of Secs. 9.1 and 9.3 can be found published in Ref. [60], under CC BY 4.0. Since the author of the present PhD thesis is not the first author of Ref. [60], we will restrict to a presentation of the main results of this work in Sec. 9.1, we will make some additional remarks regarding different initial conditions 9.2, and discuss experimental transfer rates in the context of our TB approach in Sec. 9.3. As previously, we will use only the $5' - 3'$ strand to denote the segments, and I polymers refer to polymers made up of identical monomers, while D polymers refer to polymers made up of different monomers. For deterministic aperiodic binary polymers, the notation X(Y) means that that the first monomer in the sequence is X, and the other monomer in the sequence is Y.

9.1 Main results

9.1.1 Energy structure

The energy structure various categories of deterministic aperiodic binary I and D DNA polymers have been calculated using the HKS parametrization (see Appendix A, Table A.1). Our numerical results are illustrated in Figs. 1-4 of Ref. [60]. For respective results for D polymers within the S parametrization, cf. Figs. 5.4, 5.5. For both I and D polymers, in fractal polymers the DOS is fragmented and spiky, while in the rest cases the DOS has rather acute subbands. Also, as proven in Appendix C, Theorem 2, for all I polymers, all eigenvalues are symmetric relative to the monomer's on-site energy

(this, obviously, also holds for the DOS). For D polymers, the eigenenergies (and the DOS) gather around the two monomer's on-site energies.

Regarding HOMO-LUMO gaps, the G-C (A-T) monomer gap is always greater than the gaps of I polymers made of G and C (A and T). D polymers have smaller HOMO–LUMO gaps than I polymers (cf. left panel of Fig. 5 of Ref. [60]). Furthermore, the lower HOMO (LUMO) band limit of D polymers is always between the lower and upper HOMO (LUMO) band limit of I polymers consisted of A and T, while the upper HOMO (LUMO) band limit of D polymers is always between the lower and upper HOMO (LUMO) band limit of I polymers consisted of G and C (cf. right panel of Fig. 5 of Ref. [60]).

9.1.2 Mean over time probabilities

The main aspects of our results for the mean over time probabilities for I and D polymers, for initial placement of an extra hole (HOMO) or electron (LUMO) at the first monomer, are summarized in Figures 6 and 7, as well as in Appendix A of Ref. [60], for some example cases. Usually, these probabilities are distributed to monomers close to the one the carrier was initially placed.

The probabilities of finding the extra carrier at each monomer of a polymer depend on the sequence on-site energies and magnitude of hopping parameters between successive monomers. This can more easily be seen in I polymers, where only the hopping integrals affect the energy structure. For the TM G(C) polymers, the probabilities are palindromic for odd generation numbers. This is due to the fact that the Hamiltonian matrices of these polymers are palindromic, i.e., reading them from top left to bottom right and vice versa gives the same result. This property stems directly from the sequence structure. For CS A(T) polymers, the mean over time probability for an extra hole is almost totally distributed at the four (or three for $g = 1$) starting monomers, regardless of N, while for an extra electron the probabilities are almost semi-palindromic, i.e., $\left\langle |\Psi_{\mu}(t)|^2 \right\rangle = \left\langle |\Psi_{N-\mu+1}(t)|^2 \right\rangle, \mu = 2, 4, ..., N-1$. In this case, even if the sequence structure is the same for HOMO and LUMO, the magnitude of hopping integrals has a stronger effect on the results. Another example is the RS A(T) sequence where the mean over time probability for an extra electron is almost totally distributed at the four starting monomers, regardless of N, while for holes it is basically distributed at monomers 1, 2, 3 and 6. Regarding the extra hole in ACS C(G) polymers, the probability is much higher for monomers 1, 2, 9, 10 of every 32-monomer period. Generally, for I polymers, the mean over time probabilities are significant only rather close to the first monomer, although in some cases we observe non-negligible probabilities at more distant monomers.

Generally, for D polymers, the mean over time probabilities are almost negligible further than the first monomer. An exception is the RS A(G) sequence where the probabilities for both HOMO and LUMO are almost totally distributed at the three starting monomers of each polymer, regardless its length. Likewise, the mean over time probability for the extra electron in CS A(G) polymers is almost totally distributed at the first and third monomer of each polymer, regardless its length. An extra electron in

PD A(G) reaches somehow more distant monomers.

A general picture is that for deterministic aperiodic polymers, the mean over time probabilities generally decline a little away from the first monomer, at which the carrier is initially placed. This situation is different from the general picture in periodic polymers studied in Chapter 8 (cf. Refs. [56, 58, 59]), where, as a rule, non-negligible probabilities exist at distant—from the first—sites.

9.1.3 Frequency content

The frequencies involved in charge transfer are given by Eq. (6.13). The Fourier spectra of the time-dependent probability to find an extra electron or hole at each monomer are generally in the THz regime, mainly in the FIR and MIR part of the electromagnetic spectrum. When the dominant frequencies, i.e., those with greater Fourier amplitudes, are smaller (larger), the carrier transfer is slower (faster). Extensive examples of the Fourier spectra of the probability to find an extra carrier at the first and at the last monomer, having placed it initially at the first monomer, for I and D deterministic aperiodic polymers, for the HOMO and the LUMO regime, can be found in Refs. [262, 263].

In Fig. 9 of Ref. [60], the TWMF as a function of N for the various types of deterministic aperiodic polymers is depicted. We notice that the TWMF generally stabilizes as the generation number increases. In all cases of deterministic aperiodic polymers studied in Ref. [60], the TWMFs are in the region $\approx 10^{-2}$–10^{2} THz. We notice, for comparison, that In various cases of *periodic* I and D polymers studied in Sec. 8.2 (cf. Ref. [59]), the TWMFs were found in the region $\approx 10^{0}$–10^{2} THz.

9.1.4 Pure mean transfer rates

We focus on the pure mean transfer rates from the first to the last monomer, $k_{1,N}$, or from now on, just k. $k(N)$ either for HOMO or for LUMO, for I and D deterministic aperiodic polymers, is depicted in Fig. 10 of Ref. [60]. In all cases, $k(N)$ is a decreasing function. Generally, the degree of coherent transfer difficulty is greater for D polymers. Overall, our results suggest that I polymers, which are simpler cases in terms of energy intricacy, are more efficient regarding coherent hole and electron transfer.

In each panel of Fig. 10 of Ref. [60], $k(N)$ of homopolymers (e.g., A...), which are the "champions" among periodic polymers in terms of efficiency of coherent carrier transfer, i.e., in terms of magnitude of k and of slower decrease of $k(N)$ (see Chapter 8), is included. It seems that $k(N)$ of homopolymers is an unreachable limit for aperiodic polymers. Comparing the periodic polymers studied in Chapter 8 (cf. Refs. [56, 58, 59]) with deterministic aperiodic polymers, in terms of $k(N)$, we realize that although generally periodic polymers are more efficient, specific aperiodic polymers can be better than specific periodic ones. However, the general picture is that charge transfer in deterministic aperiodic polymers is orders of magnitude worse than in periodic polymers.

Furthermore, in each panel of Fig. 10 of Ref. [60], a random shuffle of the sequence of best deterministic aperiodic polymers, in terms of $k(N)$, is included. In all cases, except for HOMO CS , this random shuffle deteriorates severely $k(N)$. For CS, A(T) and T(A) have identical $k(N)$ because the CS rules for A(T) and T(A) produce equivalent polymers, cf. Eq. (6.11). For equivalent polymers, $k(N)$ from the first to the last monomer are identical, cf. Eq. (6.19). For example, TAT $\equiv$ ATA, TATAAATAT $\equiv$ ATATTTATA, TATAAATATAAAAAAAAATATAAATAT $\equiv$ ATATTTATATTTTTTTTTATATTTATA, and so on. Similarly, of course, the CS rules for G(C) and C(G) produce equivalent polymers, which have identical $k(N)$. For CS HOMO, the best sequences in terms of $k(N)$ are A(T) and T(A), where the hopping integrals involved are $t_{\mathrm{AA}} = t_{\mathrm{TT}} = -8$ meV, $t_{\mathrm{AT}} = 20$ meV, $t_{\mathrm{TA}} = 47$ meV, and we have just one on-site energy, that of A-T. From these hopping integrals, t_{AA} has the smallest absolute value. Given the structure of the CS sequences, making the random shuffle, the number of t_{AA} decreases, while the numbers of the larger hopping integrals, t_{AT} and t_{TA} increase. For this reason, in HOMO CS, the random shuffle increases $k(N)$. In LUMO CS, this argument is inverted because now the best sequences in terms of $k(N)$ are G(C) and C(G), where the hopping integrals involved are $t_{\mathrm{GG}} = t_{\mathrm{CC}} = 20$ meV, $t_{\mathrm{GC}} = -10$ meV, $t_{\mathrm{CG}} = -8$ meV, and we have just one on-site energy, that of G-C. In this case, the random shuffle decreases the number of the larger hopping integrals $t_{\mathrm{GG}} = t_{\mathrm{CC}}$ and increases the number of the smaller hopping integrals t_{GC} and t_{CG}. However, apart from the exception of HOMO CS, generally speaking, the conclusion is that deterministic aperiodic polymers possess some kind of order, i.e., a well-defined construction rule, that makes them more efficient than random polymers in terms of $k(N)$; therefore, when this rule is destroyed, transfer efficiency diminishes.

9.2 Different initial conditions

In Figs. 9.1 and 9.2, we present our numerical results for periodic, various deterministic aperiodic, and random I and D, respectively, DNA polymers, within the WM and with the S parametrization (see Appendix A, Table A.1), for all possible initial placements of an extra hole.

Among the first things that can be observed from Figs. 9.1 and 9.2 is that periodic and random polymers represent the extreme cases; in the former, the carrier is essentially distributed all over the polymer, while, in the latter, the carrier is only present only very close to the monomer it was initially placed at. Furthermore, for each category of deterministic sequences with a definable substitution matrix [i.e., all cases presented in Figs. 9.1 and 9.2 apart from KOL(1,2)], there are specific regions of the polymer in which the carrier can be found; the relative length of these regions is scale invariant, i.e., it does not depend on the generation number; the corresponding relative regions just become more fragmented (results not shown here). This is another footprint of inflation/deflation symmetry. These regions seem to be somewhat more strictly defined for I polymers than for D polymers. The above remark indicates that the region of a deterministic polymer at which an extra carrier is located can be found can be predicted, given the site it was initially inserted at. Additionally, regarding de-

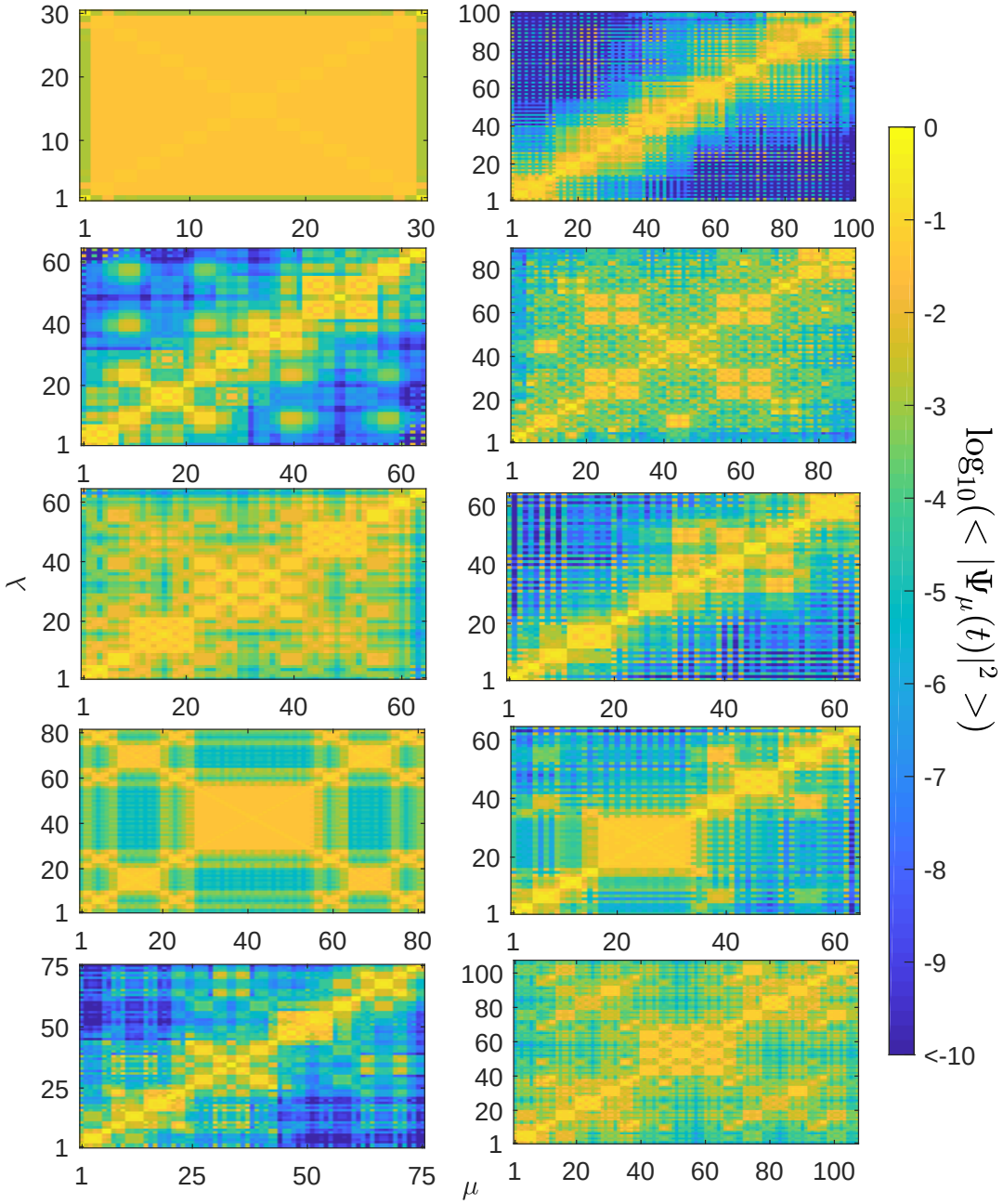

Figure 9.1: Mean over time probabilities of hole transfer for various categories of I DNA polymers and for all possible initial conditions. First row: periodic GCGC.... (left) and random (50% G, 50& C, right). Second row: TM (left) and Fibonacci (right). Third row: PD (left) and RS (right). Fourth row: CS (left) and ACS (right). Fifth row: KOL(1,2) (left) KOL(1,3) (right).

terministic aperiodic sequences, there is a correspondence between the number and occurrence percentage of possible triplets, as well as the strength of correlations, discussed in Chapter 5 (cf. Figs. 5.2, 5.3), and the extent of charge charge presence along the polymer. For example, random and RS sequences are the least efficient cases, in accordance with the 8 equidistributed triplets and the weak correlations they posses.

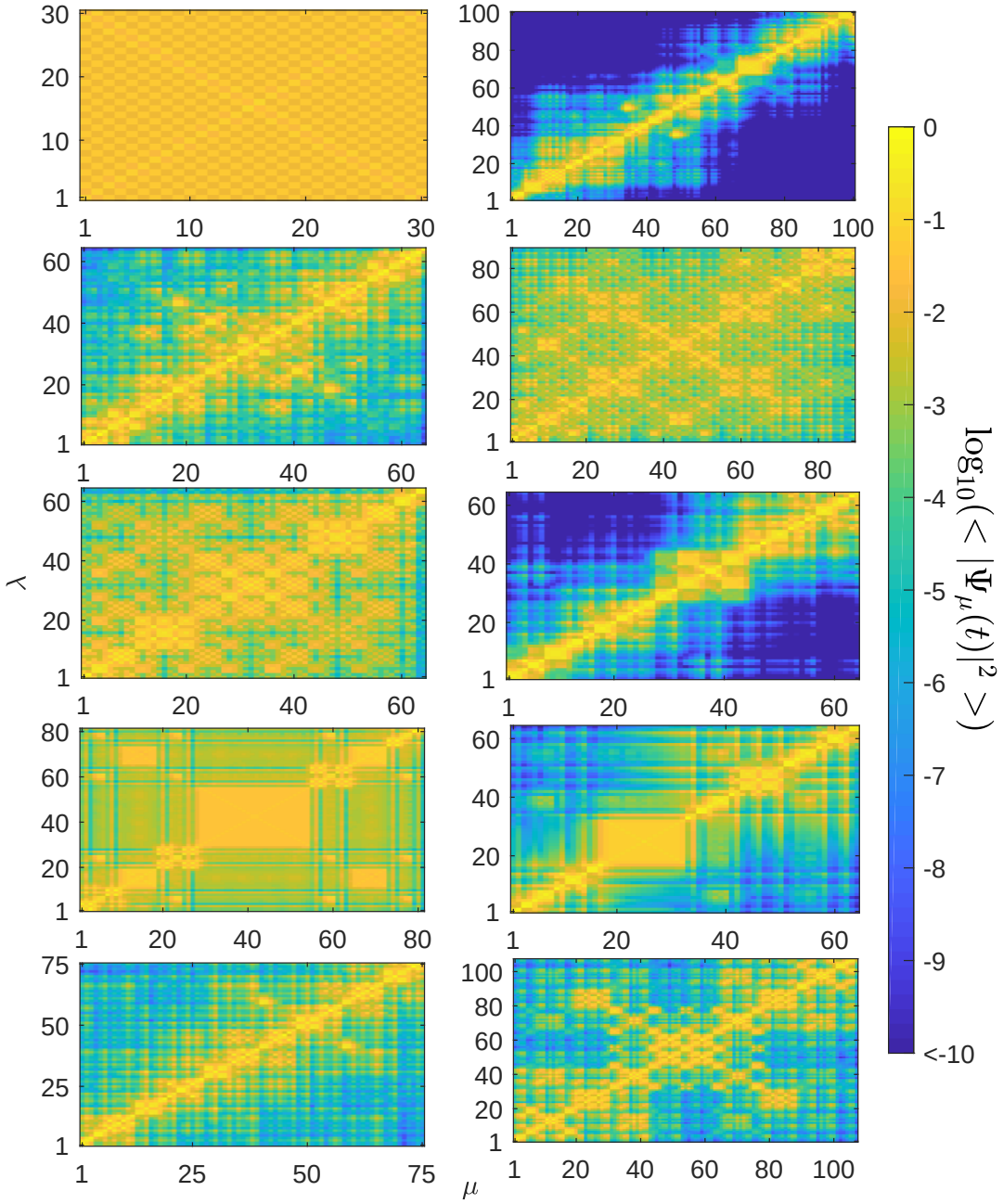

Figure 9.2: Mean over time probabilities of hole transfer for various categories of D DNA polymers and for all possible initial conditions. First row: periodic GAGA.... (left) and random (50% G, 50& C, right). Second row: TM (left) and Fibonacci (right). Third row: PD (left) and RS (right). Fourth row: CS (left) and ACS (right). Fifth row: KOL(1,2) (left) KOL(1,3) (right).

TM and KOL(1,2) sequences, with similar triplet distribution and similar correlations, have resembling mean probability distribution (except for I TM polymers with odd generation numbers, which are palindromic–a case not shown here). The most efficient deterministic aperiodic sequences are Fibonacci, PD and fractal sequences, again in accordance with the discussion of Chapter 5. All in all, it can be deduced that, al-

though end-to-end charge transfer in deterministic aperiodic polymers seems to be rather inefficient, as shown in Sec. 9.1, further investigation shows that there are specific (for each category) regions at which the the excess carrier can efficiently travel rather far along the polymer, given its chosen initial placement.

9.3 Some remarks on experimental transfer rates

Comparison of the coherent pure mean transfer rates k of DNA with experimentally obtained transfer rates is a rather complicated issue. In the past, the experimental transfer rates in donor—bridge (DNA)—acceptor systems were obtained using the concentrations of different products generated e.g., when a hole is (PY) or is not (PN) transferred. The concentrations of PY and PN were indirectly measured by methods like polyacrylamide gel electrophoresis and piperidine treatment [264, 265]. Although these methods revealed some aspects of hole transfer like the sequence dependence and the ability of transfer, they do not provide the kinetics of hole transfer in DNA [266]. Although, generally, greater concentration of PY implies greater charge transfer, there is no proof that the concentrations of PN and PY are proportional to the degree of transfer.

Quantum mechanically, only a fraction of the carrier reaches the acceptor through the bridge. For the same reason, the definition of transfer time is problematic. The transfer rate should depend both on the amount and the speed of transfer. However, the concentration of PY is not strictly proportional to the amount of carrier transfer and not strictly inversely proportional to the time of transfer. A more direct experimental approach is time-resolved spectroscopy, e.g., transient absorption, to observe the products of charge transfer [266–268].

Our point of view is different, since the quantity we use, the *pure* mean transfer rate [55], given by Eq. (6.18), uses simultaneously the magnitude of coherent charge transfer and the time scale of the phenomenon. However, our method applies to coherent transfer only and cannot cover incoherent mechanisms like thermal hopping.

It is a common assertion in the literature that when the fall of the transfer rate with respect to the length of a given DNA segment is described by an exponential fit, the mechanism of transfer is superexchange, whereas when it is described by a power law fit, the mechanism of transfer is multi-step hopping. However, we stress that the fitted parameters produced this way should be treated with care, especially when it comes to attributing them to specific mechanisms. For example, in Ref. [266], where the hole transfer kinetics of various short DNA segments were experimentally investigated with time-resolved spectroscopy, the authors present an exponential decay length $\beta = 1.6$ Å^{-1} by fitting the experimental hole transfer rates of G(A)$_n$G DNA oligomers ($n = 0, 1, 2$) to the exponential law $K = K_0 \mathrm{e}^{-\beta d}$, where d is the charge transfer distance, i.e., $d = 3.4 \times (N - 1)$ Å. Using the transfer rate values of Ref. [266], we observed that, although β, determined as the slope of the linear fit $\ln(K) = \ln(K_0) - \beta d$ is indeed $\cong 1.6$ Å^{-1}, a direct exponential fit gives $\beta \cong 1.3$ Å^{-1}, suggesting that the law of decay is not exactly exponential. On the contrary, the fits of our theoretically obtained *pure* mean transfer rates, k, for the same system, give $\beta \cong 1.84$ Å^{-1} for β

determined as the slope of the linear fit $\ln(k) = \ln(k_0) - \beta d$, and $\beta \cong 1.79$ Å^{-1} for a direct exponential fit $k = k_0 \mathrm{e}^{-\beta d}$, suggesting closer convergence to an exponential decay. Similarly, in Ref. [269], the authors experimentally study, with time-resolved spectroscopy, hole transfer through $(GA)_n$ and $(GT)_n$ sequences, where $n = 2$–12 is the number of repetition units. The authors fitted the obtained transfer rates to the power law $K = K_0' \mathcal{N}^{-\eta}$, where $\mathcal{N}$ is the number of hopping steps between guanines (in our notation, $\mathcal{N} = \frac{N}{2} - 1$), reported the same exponent for both sequences, i.e., $\eta = 2$, and suggested that this value provides evidence that the long-distance hole transfer occurs by multi-step hopping between guanines. From the rate values provided in Table I of Ref. [269], we observed that, although η as a slope of the linear fit $\ln(K) = \ln\left(K_0'\right) - \eta \ln(\mathcal{N})$ is indeed 2 for both sequences, a direct power law fit yields $\eta \cong 1.4$ for $(GA)_n$ and $\eta \cong 1.3$ for $(GT)_n$, suggesting that the rate decay does not follow exactly a power law. On the contrary, the fits of our theoretically obtained *pure* mean transfer rates, k, for $(GA)_n$, give $\eta \cong 1.40$ for η determined as the slope of the linear fit $\ln(k) = \ln\left(k_0'\right) - \eta \ln(\mathcal{N})$, and $\eta \cong 1.56$ Å^{-1} for a direct power law fit $k = k_0' \mathcal{N}^{-\eta}$. The respective values for $(GT)_n$ are $\eta \cong 2$ for both fits. Hence, our theoretical results suggest that the fall of k, as the length of the bridge increases, convergences to a power law and that the fall of the transfer rate is less steep when purines are on the same strand compared to the case when purines are crosswise.

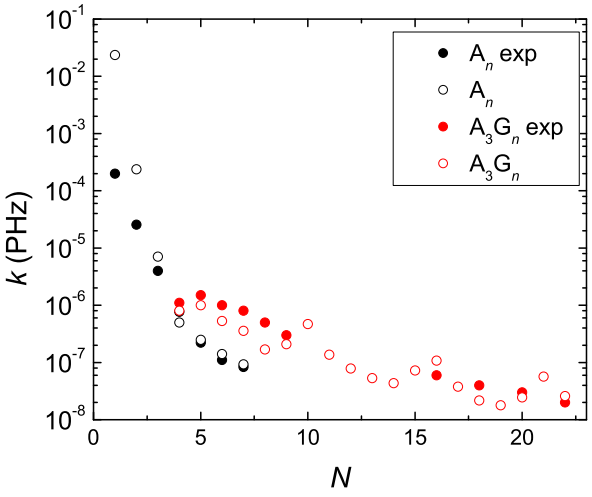

Figure 9.3: Comparison of experimental hole transfer rates K for A_n and A_3G_n segments [270] (full circles) with our theoretical coherent pure mean transfer rates k (empty circles), as a function of the number of monomers N in the polymer. The parametrization is described in the main text.

DNA is a dynamical structure, i.e., the geometry is not fixed. Large variations of the TB parameters are expected in real situations and also large variations of the TB parameters have been obtained by different theoretical methods by different authors, cf. e.g., Ref. [55] and references therein. Hence, the parameters any TB model uses have to be utilized with care. In Ref. [270], the authors report experimentally deduced (by transient absorption spectroscopy) charge separation rates, in capped A_n ($n =$

1–7) and A_3G_n ($n = 1$–19) DNA hairpins with a stilbenedicarboxamide hole donor and a stilbenediether hole acceptor. We computed our theoretical coherent pure mean transfer rates, k, for the same systems with a modified HKS parametrization: $t_{AA} \to 1.6t_{AA}$, $t_{AG} \to 2.1t_{AG}$, $t_{GG} \to 2.25t_{AG}$, cf. Table A.1). In order to mimic the donor and the acceptor, we added two sites at the ends of the TB chain, with on-site energies $\epsilon_{\text{don}} = \epsilon_{A-T} - 0.1$ eV, $\epsilon_{\text{ac}} = \epsilon_{G-C} + 0.1$ eV. We used for the hopping integral from the donor (last base pair) to the first base pair (acceptor) 100 meV (250 meV). Our results, along with the experimental ones, are depicted in Figure 9.3. Apart from the A_1 and A_2 systems, for which we find much larger rates, the pure mean transfer rates k are of the same order of magnitude, in good quantitative agreement with the experimental transfer rates K. Actually, the same sequences A_n ($n = 1$–7) and A_3G_n ($n = 1$–19) analyzed in Ref. [270] had also been analyzed by the same group in Ref. [271]. In Ref. [271], the authors mention a time resolution of ca. 180 fs. Hence, roughly, only transfer rates $K < (1/180)$ PHz $\approx (1/200)$ PHz $= 5 \times 10^{-3}$ PHz can be detected by this technique.

9.4 Conclusion

We discussed coherent charge transfer on deterministic aperiodic and random binary I and D DNA segments, within the WM. Regarding the energy structure of the polymers, the eigenenergies lie around the monomers' on-site energies. For I polymers, the eigenenergies are always symmetric relative to the (constant) monomer on-site energy. For both I and D polymers, in fractal cases the DOS is fragmented and spiky, while in the rest cases the DOS has rather acute subbands. D polymers pose smaller HOMO–LUMO gaps than I polymers. The lower HOMO (LUMO) band limit of D polymers is always between the lower and upper HOMO (LUMO) band limit of I polymers consisted of A and T, while the upper HOMO (LUMO) band limit of D polymers is always between the lower and upper HOMO (LUMO) band limit of I polymers consisted of G and C.

Next, we studied the mean over time probabilities to find an extra hole or electron at each monomer of the polymer, having it initially placed at the first monomer. For I polymers, the mean over time probabilities are significant only rather close to the first monomer, although in some cases we observe non-negligible probabilities at more distant monomers. For D polymers, the mean over time probabilities are generally negligible further than the first monomer. This is in contrast with periodic polymers, studied in Chapter 8, where, generally, non-negligible probabilities exist at distant—from the initial—sites.

Furthermore, we determined the frequency content of coherent extra carrier transfer via the total weighted mean frequency of the polymer, using the weighted mean frequencies of the Fourier spectra that correspond to the probabilities to find the carrier at each monomer. The TWMF generally stabilizes after a few generations. In all cases of deterministic aperiodic polymers studied, the TWMF lies in the regime $\approx 10^{-2} - 10^{2}$ THz. This is different from various cases of periodic I and D polymers studied Sec. 8.2, where the TWMFs were found in the region $\approx 10^{0} - 10^{2}$ THz.

The study of the pure mean transfer rates, $k(N)$, shows that I polymers, which are simpler cases in terms of energy intricacy, are more efficient than D polymers regarding coherent hole and electron transfer. Comparing the periodic polymers studied in Chapter 8 (see also Refs. [56, 58, 59]) with the aperiodic polymers studied here, in terms of $k(N)$, we realize that, although generally periodic polymers are more efficient, specific aperiodic polymers can be better than specific periodic ones. However, the general picture is that charge transfer in aperiodic polymers is orders of magnitude worse than in periodic polymers. The structurally simplest periodic polymers, i.e., homopolymers, represent an unreachable limit for all aperiodic polymers. Furthermore, a random shuffle of a deterministic aperiodic monomer sequence destroys the deterministic character of its construction rules, thus leading to vanishing transfer rates.

Further comparative investigation of the mean over time probabilities to find a carrier at each monomer of periodic, deterministic aperiodic, and random polymers for all possible initial placements reveals that, although end-to-end charge transfer in deterministic aperiodic polymers seems to be rather inefficient, there are specific (for each category) regions at which the the excess carrier can efficiently travel rather far along the polymer, given its chosen initial placement. The extent of these regions is in accordance with our discussion on the number and occurrence percentage of possible triplets, as well as the strength of correlations, discussed in Chapter 5. Less number of triplets, the existence of dominant triplets and stronger correlations lead to larger regions in which the charge can be found.

As far as comparison with experiments regarding the transfer rates is concerned, large variations of the TB parameters are expected in real situations, hence modifications are necessary. Using a modified HKS parametrization, we were able to find hole pure mean transfer rates k of similar magnitude with experimental transfer rates K obtained by time-resolved spectroscopy.

Conclusion

This PhD thesis was devoted to a comparative and systematic, both analytical and numerical, study of charge transport and transfer properties π conjugated molecular wires, and specifically of DNA and carbynes, using variants of the TB method. Regarding charge transport, the computational code used throughout Part II was solely developed by the author, while, regarding charge transfer the code used throughout Part III is a product of several amendments, additions, and optimizations made by the author to already existing programs that had been developed within the group "Physics of Nanostructures and Biomaterials", again, with his participation.

Our first goal regarding charge transport was to study the simplest variant of one-dimensional TB models, i.e., the WM, in a somewhat more generic manner. We focused on periodic systems whose chemical diversity is measured in terms of the number of different kinds of atoms and bonds in the unit cell, and found out that some properties, which are usually regarded as specific properties of the systems considered, may turn out to be quite general consequences of the adopted quantum lattice model instead. We showed a recipe to produce their energy structure for cyclic and fixed boundaries and systematically studied all factors that affect their transmission profiles. We introduced an optimal coupling condition between the considered systems and the leads they are attached to, based on two factors, i.e., the coupling strength and asymmetry factors.

Having obtained this knowledge, we moved on to apply our model to cumulenic and polyynic carbynes. Our numerical results have shown that this model is able to catch several experimentally detectable features of charge transport, such as the metallic (semiconducting) behavior of cumulenic (polyynic) carbynes and the rectifying behavior of the current-voltage curves when there is a mismatch between the Fermi levels of the leads to which polyynic carbynes are attached.

Next we focused on the interplay between the sequence structure, the energy structure and the charge transport properties of several categories of DNA segments, i.e., periodic, various deterministic aperiodic, and random sequences. We shown that there is a direct correspondence between the number and occurrence percentage of the possible base-pair triplets present in each sequence category and the aforementioned properties. Apart from a natural unit in a TB model with nearest neighbor interactions, base-pair triplets constitute the so-called codons, each one of which corresponds to a specific amino acid or stop signal during protein synthesis. Specifically, for all deterministic aperiodic sequences studied, we found that there are IDOS steps the relative height of which is equal to the occurrence percentage of the possible triplets within them. The strength of correlations within the segments, also affected by the number of triplets and the possible presence of dominant ones, is reflected on the localization and

transmission properties of the segments. By systematically studying the current voltage curves for all segment categories (which all have distinct shapes), we found that although periodic polymers represent the most efficient case, there are deterministic aperiodic sequences that can carry significant currents. Homopolymers are a limiting case in terms of transport efficiency, and random sequences are systematically the least efficient. We have also shown that the use of a realistic parametrization, going beyond the somewhat simplistic assumptions of solely diagonal or off-diagonal disorder, reveals much more detail about the studied properties.

In order to understand coherent charge transfer of an extra hole or electron along DNA, we started by focusing on small segments, within the WM and the ELM. We showed that the two models give complementary results, predicting carrier oscillations with frequencies in the THz regime for DNA monomers, dimers, and trimers. Within monomers, charge transfer is very small to negligible. In dimers made up of identical monomers, the whole extra carrier is transferred from the one monomer to the other (when purines are on the same strand, transfer is of intra-strand character, while, when purines are crosswise, transfer is of inter-strand character). In dimers made up of different monomers, transfer is of intra-strand character, albeit a only small percentage of the carrier changes monomer. Our results for the above mentioned TB models were compared with respective ones obtained by RT-TDDFT, and are in general agreement.

The two models, i.e. the WM and the ELM, were also compared in a study of charge transfer along the simplest periodic polymers, monomer- and dimer-polymers. They give coherent, complementary results regarding the mean over time probabilities to find the extra carrier at a specific site, the frequency content of transfer, and the pure mean transfer rates, with the ELM allowing for greater detail. This complementarity stems from the fact that due to the weak $pp\pi$ coupling between the complementary bases of a base pair, the HOMO (LUMO) of the base pair within the WM is very close to the highest of the HOMOs (lowest of the LUMOs) of its two constituents within the ELM. We also studied some other classes of periodic polymers, with increasing repetition unit in the fashion GC, GGCC, etc, or GA, GGAA, etc, within the WM. We found that as the repetition unit increases, in the former case, all studied quantities related to charge transfer (eigenspectra, DOS, energy gaps, mean over time probabilities to find the extra carrier at each base pair, TWMFs, pure mean transfer rates) have the homopolymers as a limit regarding their efficiency, while, in the latter case, the limit is the union of the two possible homopolymers.

The fact that the -structurally simplest- homopolymers are the most efficient case for charge transfer is demonstrated by a comparative study of periodic, deterministic aperiodic, and random DNA segments. However, there are specific aperiodic polymers that can be more efficient than periodic ones. Furthermore, although end-to-end transfer is generally orders of magnitude worse in deterministic aperiodic polymers, a further study shows that there are regions of initial carrier placement, specific for each deterministic aperiodic polymer category, for which the extra carrier can be found rather far away. The extent of transfer is, just as in charge transport, directly related to the number and occurrence percentage of the possible base-pair triplets in each segment category.

Finally, we have shown that our coherent pure mean transfer rates can quantita-

tively reproduce experimental transfer rates obtained with transient absorption spectroscopy techniques.

Qualitatively, the results presented in this PhD thesis are not polymer-specific and could be applied to study other polymers or, more generally, one-dimensional structures, as well. DNA, due to the ability it provides to construct sequences of choice, could be seen as a prototype system. However, its large persistence length, combined with the fact that, as we have shown, its charge transport and transfer properties are strongly sequence-dependent, make it a promising candidate as a channel in nanoelectronic applications. Furthermore, the knowledge we have obtained from the systematic study of periodic DNA polymers could be employed to investigate nucleic acid repetition units related with repeat expansion disorders, i.e., inherited diseases caused when the number of nucleotide repeats exceeds a critical threshold.

Although we have focused on what is the most important intrinsic factor that affects charge transfer and transport in DNA, i.e., the base-pair sequence, this is not the only intrinsic factor that matters. DNA is flexible, and its structural deformations would have a significant effect on the TB parameters. This is a subject currently studied within our group.

Finally, having studied the charge transport properties of carbynes within TB, we are currently extending our investigations to charge transfer in open and closed polyynic and cumulenic carbon atomic wires and carbon-nitrogen wires, comparatively within TB and RT-TDDFT. To this end, we have been awarded with a grant under the project ”Support of Researchers with emphasis on New Researchers - Phase B”.

A
Tight-Binding parametrizations for DNA

We list the parametrizations used in the TB models employed for DNA in the present thesis.

Wire Model

Table A.1: Hopping integrals for all possible successive DNA base pairs, in the $5' - 3'$ direction, for the two parametrizations used in this thesis to employ the WM. In each column, the left value corresponds to HOMO, while the right value to LUMO. All hopping integrals are given in meV. For both parametrizations, the two relevant on-site energies of the two base pairs are $\epsilon_{\text{G-C}} = -8.0$ eV, $\epsilon_{\text{A-T}} = -8.3$ eV for HOMO, and $\epsilon_{\text{G-C}} = -4.5$ eV, $\epsilon_{\text{A-T}} = -4.9$ eV for LUMO.

Base-pair sequence	S parametrization [55]		HKS parametrization [61]	
GG≡CC	−100	20	−62	20
AA≡TT	−20	−29	−8	−29
GC	10	−10	1	−10
CG	−50	−8	−44	−8
AT	35	0.5	20	0.5
TA	50	2	47	2
AG≡CT	−30	3	−5	3
GA≡TC	−110	−1	−79	−1
CA≡TG	−10	17	−4	17
AC≡GT	10	32	2	32

Extended Ladder Model

Table A.2: Hopping integrals for all adjacent DNA bases, in the $5'-3'$, $3'-3'$, and $5'-5'$ directions, for the parametrization used in this thesis to employ the ELM. In each column, the left value corresponds to HOMO, while the right value to LUMO. All hopping integrals are given in meV. The relevant on-site energies of the four base bases are $\epsilon_{\mathrm{G}} = -8.0$ eV, $\epsilon_{\mathrm{C}} = -8.8$ eV, $\epsilon_{\mathrm{A}} = -8.3$ eV, $\epsilon_{\mathrm{T}} = -9.0$ eV for HOMO, and $\epsilon_{\mathrm{G}} = -4.5$ eV, $\epsilon_{\mathrm{C}} = -4.3$ eV, $\epsilon_{\mathrm{A}} = -4.4$ eV, $\epsilon_{\mathrm{T}} = -4.9$ for LUMO. The intra-base-pair hopping integrals are $t_{\text{G-C}} = -12$ meV, $t_{\text{A-T}} = -12$ meV for HOMO, and $t_{\text{G-C}} = -9$ meV, $t_{\text{A-T}} = 16$ meV for LUMO.

HKS parametrization [61]						
Adjacent bases	$5'-3'$		$3'-3'$		$5'-5'$	
GG	−62	20	−44	−5	3	−2
CC	−66	−47	1	0.3	1	2
AA	−8	16	48	29	2	6
TT	−117	−30	0.5	0.2	4	2
GC	80	43	4	−4	4	−3
CG	−1	15	4	−4	4	−3
AT	68	7	−3	3	9	2
TA	26	−7	−3	3	9	2
AG	−5	1	−3	−6	4	3
CT	−107	63	0.5	−0.2	2	−2
GA	−79	30	−3	−6	4	3
TC	−86	22	0.5	−0.2	2	−2
CA	5	−12	−5	−3	5	−2
TG	28	−17	5	2	5	3
AC	68	−3	−5	−3	5	−2
GT	73	−32	5	2	5	3

B

Fourier Spectra of DNA dimers and trimers

We present representative Fourier spectra for DNA dimers and trimers within the ELM.

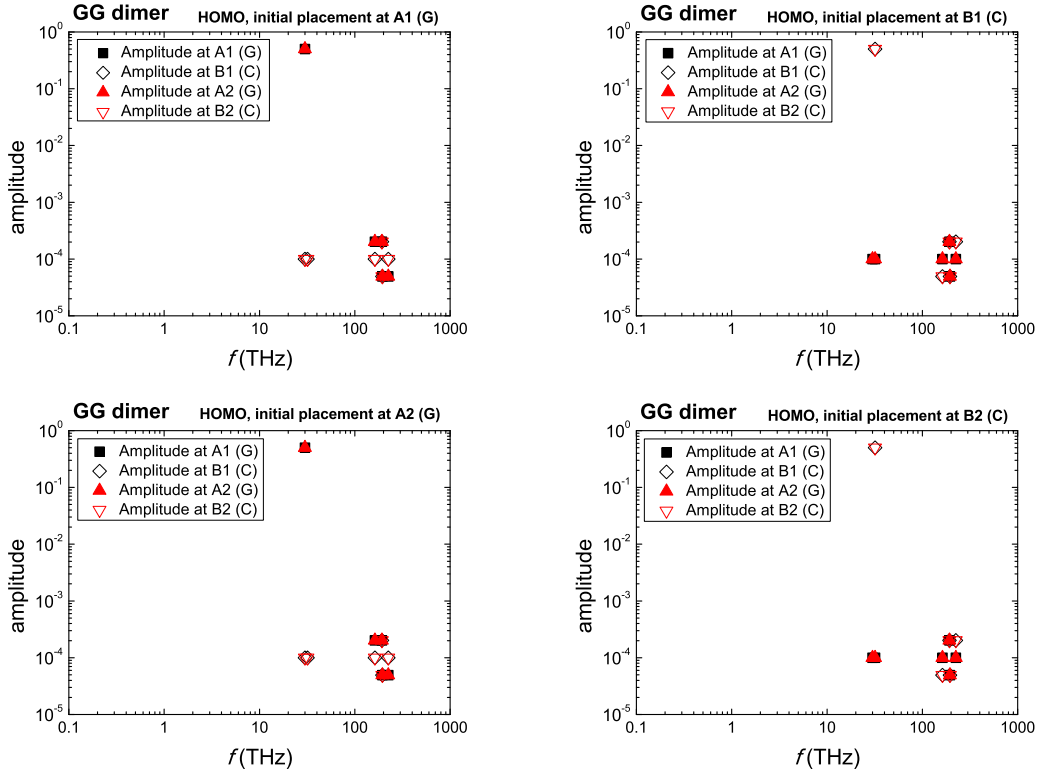

Figure B.1: Fourier spectra, within the ELM and with the HKS parametrization, of the GG dimer. A hole is placed initially at a base and we depict the frequency spectrum at all bases, A1(G), B1(C), A2(G), B2(C).

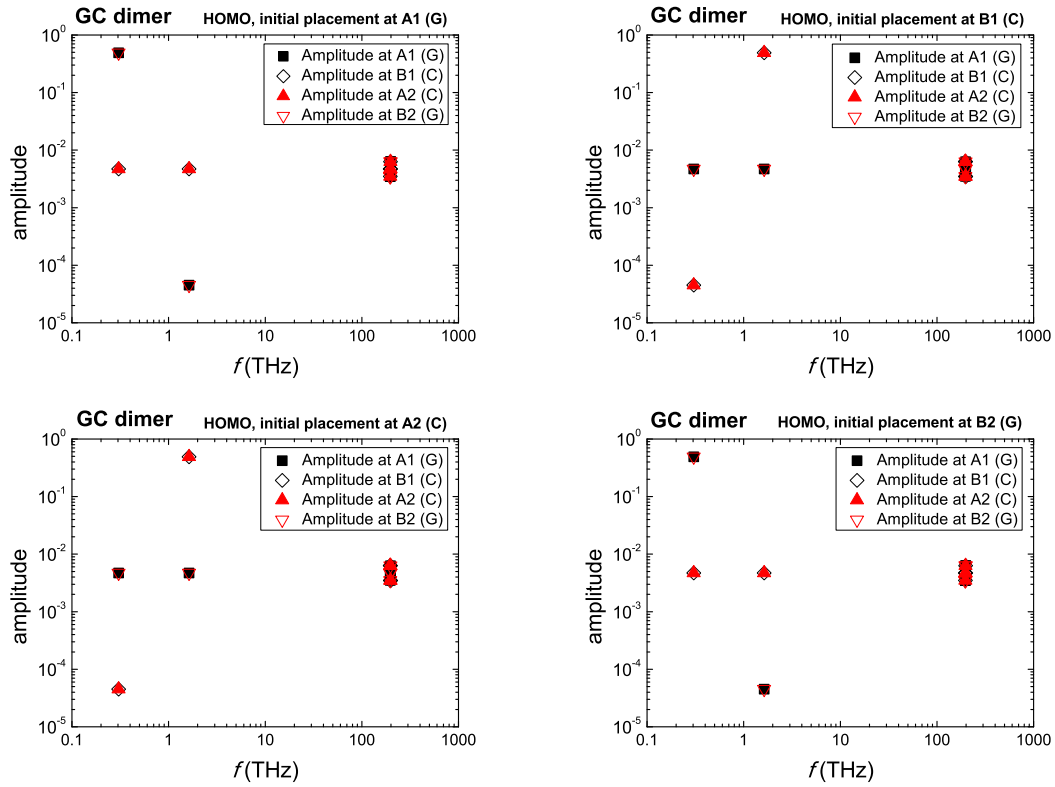

Figure B.2: Fourier spectra, within the ELM and with the HKS parametrization, of the GC dimer. A hole is placed initially at a base and we depict the frequency spectrum at all bases, A1(G), B1(C), A2(C), B2(G).

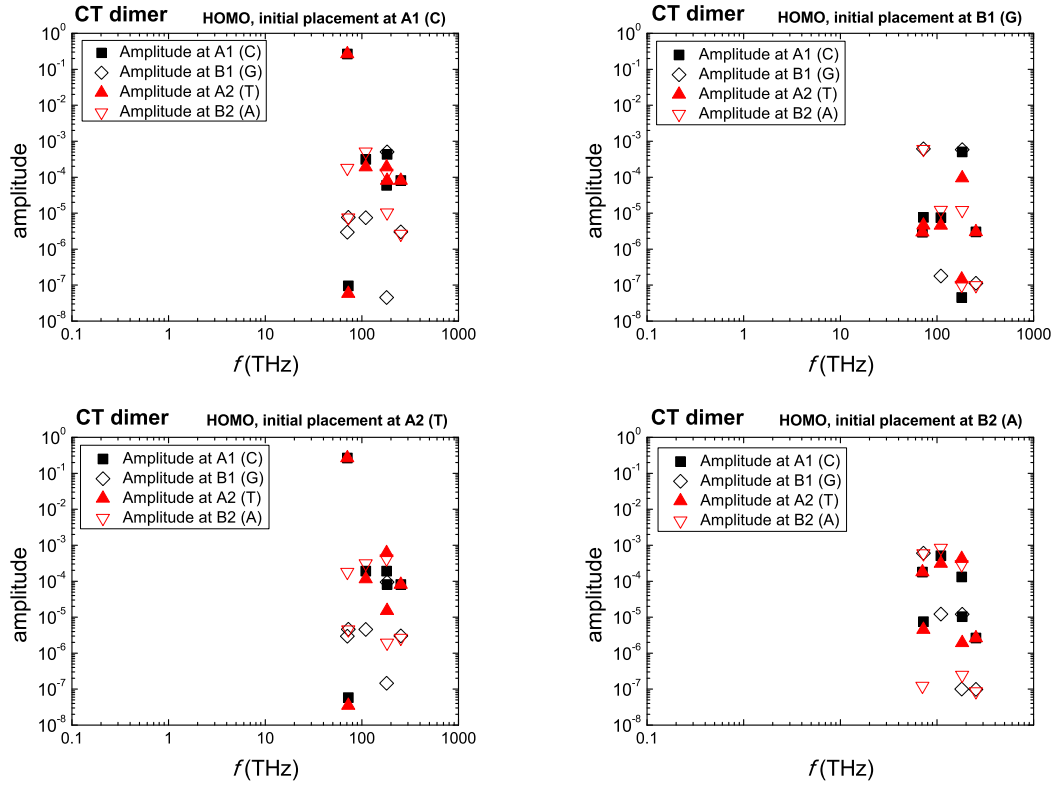

Figure B.3: Fourier spectra, within the ELM and with the HKS parametrization, of the CT dimer. A hole is placed initially at a base and we depict the frequency spectrum at all bases, A1(C), B1(G), A2(T), B2(A).

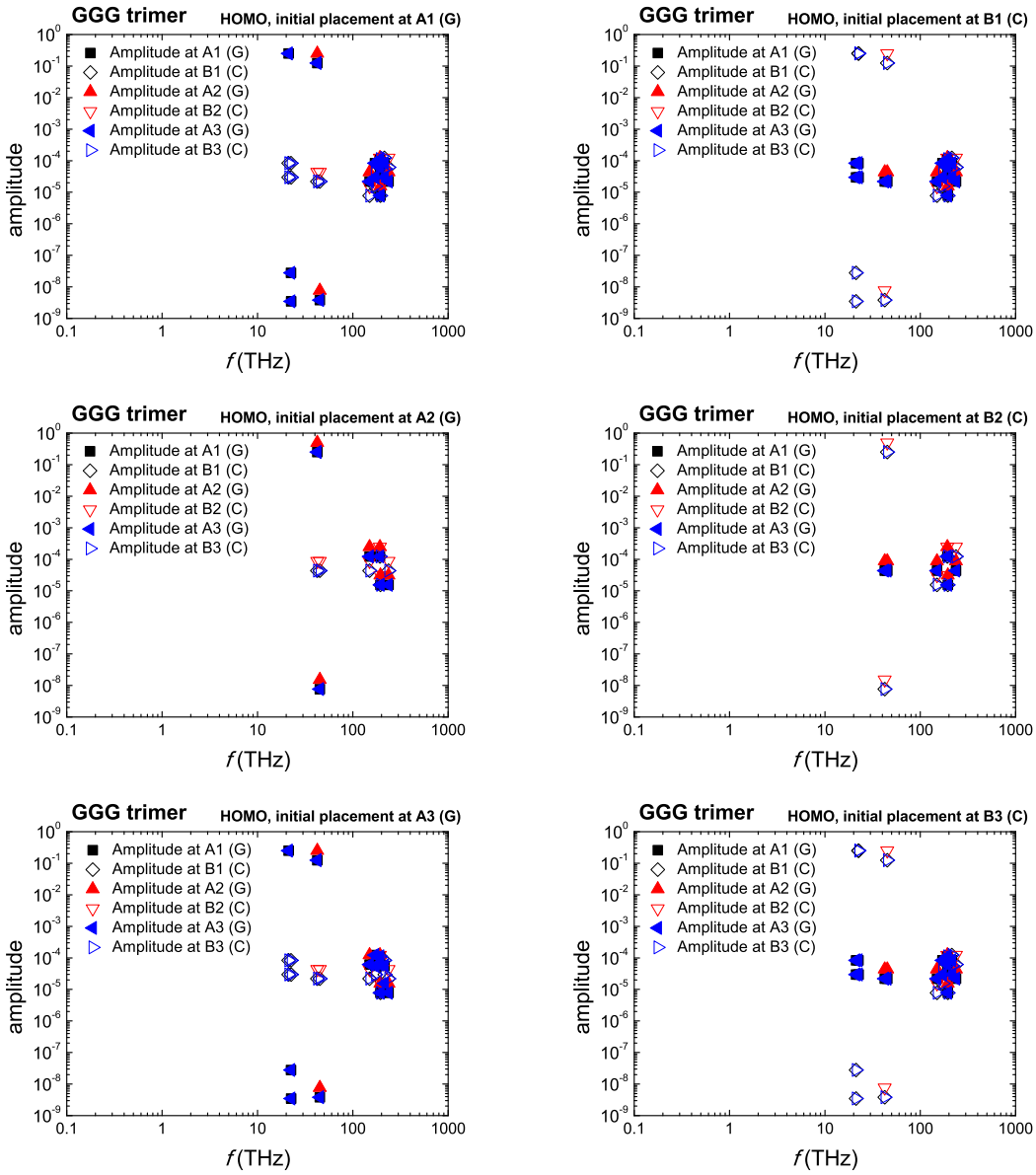

Figure B.4: Fourier spectra, within the ELM and with the HKS parametrization, of hole transfer in the GGG trimer. A hole is placed initially at a base and we depict the frequency spectrum at all bases, A1(G), B1(C), A2(G), B2(C), A3(G), B3(C).

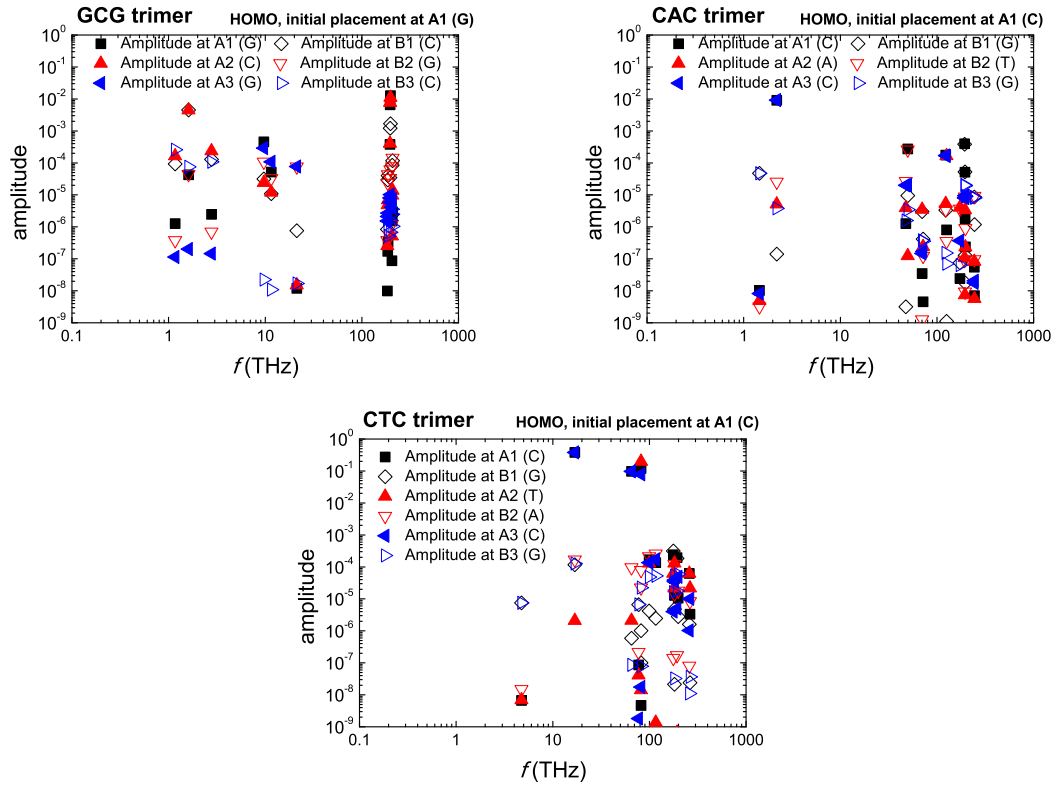

Figure B.5: Fourier spectra, within the ELM and with the HKS parametrization, of the GCG, CAC, CTC trimers for initial placement of a hole at base A1.

C

Two theorems on tridiagonal symmetric matrices

Theorem 1. *The eigenvalues of real tridiagonal symmetric matrices do not depend on the sign of their off-diagonal entries [272]*

Proof. Let $\boldsymbol{T}$ be a tridiagonal, irreducible[12], symmetric $N \times N$ matrix with diagonal elements $\boldsymbol{T}_k = a_k \; \forall k = 1, \dots N$ and off-diagonal elements $\boldsymbol{T}_{(k,k+1)} = r_{k+1} e^{-i\theta_{k+1}}$, $r_{k+1} > 0, \forall k = 1, \dots N-1$. Since $\boldsymbol{T}$ is Hermitian, $\boldsymbol{T}_{(k+1,k)} = r_{k+1} e^{i\theta_{k+1}}$. Now, suppose a diagonal $N \times N$ matrix $\boldsymbol{D}$, with elements $d_1 = 1, d_k = d_{k-1} e^{i\theta_k}, \forall k = 2, \dots, N$. Then, $\boldsymbol{D}$ is unitary, and the similarity transformation $\tilde{\boldsymbol{T}} = \boldsymbol{D}^{-1} \boldsymbol{T} \boldsymbol{D}$ leads to the matrix $\tilde{\boldsymbol{T}}$ with diagonal elements $\tilde{\boldsymbol{T}}_k = a_k$ and non-diagonal elements $\tilde{\boldsymbol{T}}_{(k,k+1)} = r_{k+1}$. Hence, the tridiagonal Hermitian matrix $\boldsymbol{T}$ has the same eigenvalues with the tridiagonal real symmetric matrix $\tilde{\boldsymbol{T}}$, which has positive non-diagonal entries [272]. Let us further suppose that $\boldsymbol{T}$ is real. Then, $\theta_k = 0$ or $\theta_k = \pi$, depending on whether $\boldsymbol{T}_{(k,k+1)} > 0$ or $\boldsymbol{T}_{(k,k+1)} < 0$. The elements of $\boldsymbol{D}$ will then be $d_k = d_{k-1}(\pm 1), \forall k = 2, \dots, .$ Hence, the matrix $\tilde{\boldsymbol{T}}$, which has positive entries, has the same eigenvalues with $\boldsymbol{T}$, which differs by $\tilde{\boldsymbol{T}}$ in that its off-diagonal elements have negative signs in arbitrary positions. We also notice, if $\vec{\boldsymbol{v}}$ is an eigenvector of $\boldsymbol{T}$, then $\boldsymbol{D}^{-1} \vec{\boldsymbol{v}}$ is an eigenvector of $\tilde{\boldsymbol{T}}$. ■

Theorem 2. *The eigenvalues of real tridiagonal symmetric matrices with the same entry in the main diagonal are symmetric around the value of this entry.*

Proof. Matrices of even order: A tridiagonal symmetric matrix, $\boldsymbol{T}$, of order $N = 2n$, with the same entry in the main diagonal, let us denote it by ϵ, can be written on the

[12] A matrix is called irreducible if it cannot be transformed through permutations into a block triagonal. In our case, i.e., tridiagonal matrices, this simply means that all the off-diagonal elements are non-zero.

form

$$\boldsymbol{T} = \epsilon \boldsymbol{I} + \boldsymbol{T}_{\text{GK}}, \tag{C.1}$$

where

$$\boldsymbol{T}_{GK} = \begin{pmatrix} 0 & t_1 & & & \\ t_1 & 0 & t_2 & & \\ & t_2 & \ddots & \ddots & \\ & & \ddots & 0 & t_{2n-1} \\ & & & t_{2n-1} & 0 \end{pmatrix} \tag{C.2}$$

is called a tridiagonal symmetric Golub-Kahan matrix. It can easily be shown that

$$\boldsymbol{T}_{GK} = \boldsymbol{P}^T \boldsymbol{B} \boldsymbol{P}, \tag{C.3}$$

where $\boldsymbol{P}$ is the perfect shuffle matrix[13] and

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{O} \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} t_1 & t_2 & & \\ & t_3 & t_4 & \\ & & \ddots & \ddots \\ & & & t_{2n-1} \end{pmatrix}. \tag{C.4}$$

Furthermore, by performing the Singular Value Decomposition of the upper bidiagonal matrix $\boldsymbol{A}$, i.e., by writing it as $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{W}^T$, and replacing in in $\boldsymbol{B}$, we obtain

$$\boldsymbol{B} = \boldsymbol{J} \begin{pmatrix} -\boldsymbol{S} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S} \end{pmatrix} \boldsymbol{J}^T, \quad \boldsymbol{J} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{U} & \boldsymbol{U} \\ -\boldsymbol{W} & \boldsymbol{W} \end{pmatrix}. \tag{C.5}$$

Thus, by combining Eqs. (C.5) and (C.3), we arrive at

$$\boldsymbol{T}_{GK} = \boldsymbol{P}^T \boldsymbol{J} \begin{pmatrix} -\boldsymbol{S} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S} \end{pmatrix} \boldsymbol{J}^T \boldsymbol{P}. \tag{C.6}$$

From Eq. (C.6), we conclude that the eigenvalues of $\boldsymbol{T}_{GK}$ are given by the positive and negative values of the diagonal matrix $\boldsymbol{S}$, i.e., they are symmetric around zero [273]. Hence, given Eq. (C.1), the eigenvalues of $\boldsymbol{T}$ are symmetric around ϵ.

Matrices of even order: If $\boldsymbol{T}$ is of order $N = 2n + 1$, we add a zero row and a zero column to $\boldsymbol{T}_{GK}$ in the odd-order version of Eq. (C.2), so that it is again of even order, and follow the aforementioned procedure. Then, two degenerate trivial eigenvalues will appear apart from the symmetric ones in Eq. (C.6) [274]. Thus, the eigenvalues of $\boldsymbol{T}$ occur by omitting the zero row and column in $\boldsymbol{T}_{GK}$, hence they are symmetric around ϵ, which is also an eigenvalue. ■

[13] $\boldsymbol{P} = (\hat{\boldsymbol{e}}_{n+1} \ \hat{\boldsymbol{e}}_1 \ \hat{\boldsymbol{e}}_{n+2} \ \hat{\boldsymbol{e}}_2 \ \ldots \ \hat{\boldsymbol{e}}_{2n} \ \hat{\boldsymbol{e}}_n)$, where $\hat{\boldsymbol{e}}_k$ denotes the column-matrix with zero elements except the kth which is unit.

Bibliography

[1] J. D. Watson and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," Nature **171**, 737–738 (1953).

[2] R. Franklin and R. , Gosling, "Molecular Configuration in Sodium Thymonucleate," Nature **171**, 740–741 (1953).

[3] C. C. Page, C. C. Moser, and P. L. Dutton, "Mechanism for electron transfer within and between proteins," Curr. Opin. Chem. Biol. **7**, 551 – 556 (2003).

[4] C. C. Moser, J. L. Ross Anderson, and P. L. Dutton, "Guidelines for tunneling in enzymes," Biochim. Biophys. Acta **1797**, 1573–1586 (2010).

[5] H. B. Gray and J. R. Winkler, "Electron flow through metalloproteins," Biochim. Biophys. Acta **1797**, 1563–1572 (2010).

[6] J. M. Artés, M. López-Martínez, I. Díez-Pérez, F. Sanz, and P. Gorostiza, "Nanoscale charge transfer in redox proteins and DNA: Towards biomolecular electronics," Electrochim. Acta **140**, 83–95 (2014).

[7] A. M. Kannan, V. Renugopalakrishnan, S. Filipek, P. Li, G. F. Audette, and L. Munukutla, "Bio-Batteries and Bio-Fuel Cells: Leveraging on Electronic Charge Transfer Proteins," J. Nanosci. Nanotechnol. **9**, 1665–1678 (2009).

[8] P. J. Dandliker, R. E. Holmlin, and J. K. Barton, "Oxidative Thymine Dimer Repair in the DNA Helix," Science **275**, 1465–1468 (1997).

[9] S. R. Rajski, B. A. Jackson, and J. K. Barton, "DNA repair: models for damage and mismatch recognition," Mutat. Res. **447**, 49–72 (2000).

[10] B. Giese, "Electron transfer through DNA and peptides," Bioorg. Med. Chem. **14**, 6139–6143 (2006).

[11] 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**, 36–38 (2011).

[12] C. J. Páez, P. A. Schulz, N. R. Wilson, and R. A. Römer, "Robust signatures in the current–voltage characteristics of DNA molecules oriented between two graphene nanoribbon electrodes," New J. Phys. **14**, 093049 (2012).

[13] J. I. N. Oliveira, E. L. Albuquerque, U. L. Fulco, P. W. Mauriz, R. G. Sarmento, E. W. S. Caetano, and V. N. Freire, "Conductance of single microRNAs chains related to the autism spectrum disorder," EPL **107**, 68006 (2014).

[14] C.-T. Shih, S. Roche, and R. A. Römer, "Point-Mutation Effects on Charge-Transport Properties of the Tumor-Suppressor Gene $p53$," Phys. Rev. Lett. **100**, 018105 (2008).

[15] R. G. Endres, D. L. Cox, and R. R. P. Singh, "Colloquium: The quest for high-conductance DNA," Rev. Mod. Phys. **76**, 195–214 (2004).

[16] C. H. Wohlgamuth, M. A. McWilliams, and J. D. Slinker, "DNA as a Molecular Wire: Distance and Sequence Dependence," Anal. Chem. **85**, 8634–8640 (2013).

[17] M. H. Abouzar, A. Poghossian, A. G. Cherstvy, A. M. Pedraza, S. Ingebrandt, and M. J. Schöning, "Label-free electrical detection of DNA by means of field-effect nanoplate capacitors: Experiments and modeling," Phys. Status Solidi A **209**, 925–934 (2012).

[18] G. S. Manning, "The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force," Biophys. J. **91**, 3607–3616 (2006).

[19] H.-W. Fink and C. Schönenberger, "Electrical conduction through DNA molecules," Nature **398**, 407–410 (1999).

[20] D. Porath, A. Bezryadin, S. de Vries, and C. Dekker, "Direct measurement of electrical transport through DNA molecules," Nature **403**, 635–638 (2000).

[21] A. J. Storm, J. van Noort, S. de Vries, and C. Dekker, "Insulating behavior for DNA molecules between nanoelectrodes at the 100 nm length scale," Appl. Phys. Lett. **79**, 3881–3883 (2001).

[22] K.-H. Yoo, D. H. Ha, J.-O. Lee, J. W. Park, J. Kim, J. J. Kim, H.-Y. Lee, T. Kawai, and H. Y. Choi, "Electrical Conduction through Poly(dA)-Poly(dT) and Poly(dG)-Poly(dC) DNA Molecules," Phys. Rev. Lett. **87**, 198102 (2001).

[23] B. Xu, P. Zhang, X. Li, and N. Tao, "Direct Conductance Measurement of Single DNA Molecules in Aqueous Solution," Nano Lett. **4**, 1105–1108 (2004).

[24] H. Cohen and C. Nogues and R. Naaman and D. Porath, "Direct measurement of electrical transport through single DNA molecules of complex sequence," Proc. Natl. Acad. Sci. USA **102**, 11589–11593 (2005).

[25] G. P. Triberis, C. Simserides, and V. C. Karavolas, "Small polaron hopping transport along DNA molecules," J. Phys. Condens. Matter **17**, 2681–2690 (2005).

[26] F. C. Grozema, Y. A. Berlin, and L. D. A. Siebbeles, "Sequence-dependent charge transfer in donor-DNA-acceptor systems: A theoretical study," Int. J. Quantum Chem. **75**, 1009–1016 (1999).

[27] F. C. Grozema, Y. A. Berlin, and L. D. A. Siebbeles, "Mechanism of Charge Migration through DNA: Molecular Wire Behavior, Single-Step Tunneling or Hopping?" J. Am. Chem. Soc. **122**, 10903–10909 (2000).

[28] Y. A. Berlin, A. L. Burin, and M. A. Ratner, "Elementary steps for charge transport in DNA: thermal activation vs. tunneling," Chem. Phys. **275**, 61 – 74 (2002).

[29] M. H. Lee, S. Avdoshenko, R. Gutierrez, and G. Cuniberti, "Charge migration through DNA molecules in the presence of mismatches," Phys. Rev. B **82**, 155455 (2010).

[30] M. Fuentes-Cabrera, X. Zhao, P. R. C. Kent, and B. G. Sumpter, "Electronic Structure of xDNA," J. Phys. Chem. B **111**, 9057–9061 (2007).

[31] M. Mantela, A. Morphis, M. Tassi, and C. Simserides, "Lowest ionisation and excitation energies of biologically important heterocyclic planar molecules," Mol. Phys. **114**, 709–718 (2016).

[32] F. D. Lewis and M. R. Wasielewski, "Dynamics and efficiency of photoinduced charge transport in DNA: Toward the elusive molecular wire," Pure Appl. Chem **85**, 1379–1387 (2013).

[33] M. Zoli, "Twist-stretch profiles of DNA chains," J. Phys.: Condens. Matter **29**, 225101 (2017).

[34] R. Gutiérrez, R. A. Caetano, B. P. Woiczikowski, T. Kubař, M. Elstner, and G. Cuniberti, "Charge Transport through Biomolecular Wires in a Solvent: Bridging Molecular Dynamics and Model Hamiltonian Approaches," Phys. Rev. Lett. **102** (2009).

[35] R. Gutiérrez and R. Caetano and P. B. Woiczikowski and T. Kubař and 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).

[36] L. D. Siebbeles and Y. A. Berlin, "Quantum motion of particles along one-dimensional pathways with static and dynamic energy disorder," Chem. Phys. **238**, 97–107 (1998).

[37] P. B. Woiczikowski, T. Kubař, R. Gutiérrez, R. A. Caetano, G. Cuniberti, and M. Elstner, "Combined density functional theory and Landauer approach for hole transfer in DNA along classical molecular dynamics trajectories," J. Chem. Phys. **130**, 215104 (2009).

[38] M. H. Lee, G. Brancolini, R. Gutiérrez, R. Di Felice, and G. Cuniberti, "Probing Charge Transport in Oxidatively Damaged DNA Sequences under the Influence of Structural Fluctuations," J. Phys. Chem. B **116**, 10977–10985 (2012).

[39] J. Štěpánek, V. Kopecký, P.-Y. Turpin, Z. Li, B. Alpert, and C. Zentz, "DNA Electric Charge Oscillations Govern Protein-DNA Recognition," PLoS One **10**, e0124444 (2015).

[40] C. Liu, L. Xiang, Y. Zhang, P. Zhang, D. N. Beratan, Y. Li, and N. Tao, "Engineering nanometre-scale coherence in soft matter," Nat. Chem. **8**, 941–945 (2016).

[41] K. Kawai and T. Majima, "Increasing the hole transfer rate through DNA by chemical modification," in *Chemical Science of π-Electron Systems*, edited by T. Akasaka, A. O. S. Fukuzumi, H. Kandori, and Y. Aso (Springer, Tokyo, 2015).

[42] D. Rawtani, B. Kuntmal, and Y. Agrawal, "Charge transfer in DNA and its diverse modelling approaches," Front. Life Sci. **9**, 214–225 (2016).

[43] R. A. Marcus, "Nonadiabatic processes involving quantumlike and classical-like coordinates with applications to nonadiabatic electron transfers," J. Chem. Phys. **81**, 4494–4500 (1984).

[44] R. A. Marcus, "Electron transfer reactions in chemistry. Theory and experiment," Rev. Mod. Phys. **65**, 599–610 (1993).

[45] A. Miller and E. Abrahams, "Impurity Conduction at Low Concentrations," Phys. Rev. **120**, 745–755 (1960).

[46] I. I. Fishchuk, A. Kadashchuk, S. T. Hoffmann, S. Athanasopoulos, J. Genoe, H. Bässler, and A. Köhler, "Unified description for hopping transport in organic semiconductors including both energetic disorder and polaronic contributions," Phys. Rev. B **88**, 125202 (2013).

[47] I. I. Fishchuk, A. Kadashchuk, S. T. Hoffmann, S. Athanasopoulos, J. Genoe, , H. Bässler, and A. Köhler, "Analytic model of hopping transport in organic semiconductors including both energetic disorder and polaronic contributions," AIP Conf. Proc. **1610**, 47–52 (2014).

[48] N. Vukmirović and L.-W. Wang, "Carrier hopping in disordered semiconducting polymers: How accurate is the Miller–Abrahams model?" Appl. Phys. Lett. **97**, 043305 (2010).

[49] J. O. Oelerich, F. Jansson, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii, "Energy position of the transport path in disordered organic semiconductors," J. Phys.: Condens. Matter **26**, 255801 (2014).

[50] N. J. van der Kaap, I. Katsouras, K. Asadi, P. W. M. Blom, L. J. A. Koster, and D. M. de Leeuw, "Charge transport in disordered semiconducting polymers driven by nuclear tunneling," Phys. Rev. B **93**, 140206(R) (2016).

[51] R. A. Marcus and N. Sutin, "Electron transfers in chemistry and biology," Biochim. Biophys. Acta **811**, 265 – 322 (1985).

[52] N. Tessler, Y. Preezant, N. Rappaport, and Y. Roichman, "Charge Transport in Disordered Organic Materials and Its Relevance to Thin-Film Devices: A Tutorial Review," Adv. Mater. **21**, 2741–2761 (2009).

[53] D. Segal, A. Nitzan, W. B. Davis, M. R. Wasielewski, and M. A. Ratner, "Electron Transfer Rates in Bridged Molecular Systems 2. A Steady-State Analysis of Coherent Tunneling and Thermal Transitions," J. Phys. Chem. B **104**, 3817–3829 (2000).

[54] T. Shimazaki, Y. Asai, and K. Yamashita, "Theoretical Rate Constants of Super-Exchange Hole Transfer and Thermally Induced Hopping in DNA," J. Phys. Chem. B **109**, 1295–1303 (2005).

[55] C. Simserides, "A systematic study of electron or hole transfer along DNA dimers, trimers and polymers," Chem. Phys. **440**, 31 – 41 (2014).

[56] K. Lambropoulos, M. Chatzieleftheriou, A. Morphis, K. Kaklamanis, M. Theodorakou, and C. Simserides, "Unbiased charge oscillations in B-DNA: Monomer polymers and dimer polymers," Phys. Rev. E **92**, 032725 (2015).

[57] K. Lambropoulos, K. Kaklamanis, A. Morphis, M. Tassi, R. Lopp, G. Georgiadis, M. Theodorakou, M. Chatzieleftheriou, and C. Simserides, "Wire and extended ladder model predict THz oscillations in DNA monomers, dimers and trimers," J. Phys.: Condens. Matter **28**, 495101 (2016).

[58] K. Lambropoulos, M. Chatzieleftheriou, 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).

[59] K. Lambropoulos, C. Vantaraki, P. Bilia, M. Mantela, and C. Simserides, "Periodic polymers with increasing repetition unit: Energy structure and carrier transfer," Phys. Rev. E **98**, 032412 (2018).

[60] M. Mantela, K. Lambropoulos, M. Theodorakou, and C. Simserides, "Quasi-Periodic and Fractal Polymers: Energy Structure and Carrier Transfer," Materials **12**, 2177 (2019).

[61] L. G. D. Hawke, G. Kalosakas, and C. Simserides, "Electronic parameters for charge transfer along DNA," Eur. Phys. J. E **32**, 291 (2010), ibid: Erratum to: Electronic parameters for charge transfer along DNA, **34**, 118 (2011).

[62] Y.-J. Ye and L.-L. Shen, "DFT approach to calculate electronic transfer through a segment of DNA double helix," J. Comput. Chem. **21**, 1109–1117 (2000).

[63] Y.-J. Ye and Y. Jiang, "Electronic structures and long-range electron transfer through DNA molecules," Int. J. Quantum Chem. **78**, 112–130 (2000).

[64] A. A. Voityuk, "Electronic couplings and on-site energies for hole transfer in DNA: Systematic quantum mechanical/molecular dynamic study," J. Chem. Phys. **128**, 115101 (2008).

[65] T. Kubař, P. B. Woiczikowski, G. Cuniberti, and M. Elstner, "Efficient Calculation of Charge-Transfer Matrix Elements for Hole Transfer in DNA," J. Phys. Chem. B **112**, 7937–7947 (2008).

[66] M. Tassi, A. Morphis, K. Lambropoulos, and C. Simserides, "RT-TDDFT study of hole oscillations in B-DNA monomers and dimers," Cogent Physics **4**, 1361077 (2017).

[67] E. Artacho, M. Machado, D. Sánchez-Portal, P. Ordejón, and J. M. Soler, "Electrons in dry DNA from density functional calculations," Mol. Phys. **101**, 1587–1594 (2003).

[68] H. Mehrez and M. P. Anantram, "Interbase electronic coupling for transport through DNA," Phys. Rev. B **71**, 115405 (2005).

[69] C. Adessi, S. Walch, and M. P. Anantram, "Environment and structure influence on DNA conduction," Phys. Rev. B **67**, 081405 (2003).

[70] R. N. Barnett, C. L. Cleveland, U. Landman, E. Boone, S. Kanvah, and G. B. Schuster, "Effect of Base Sequence and Hydration on the Electronic and Hole Transport Properties of Duplex DNA: Theory and Experiment," J. Phys. Chem. A **107**, 3525–3537 (2003).

[71] G. Cuniberti, L. Craco, D. Porath, and C. Dekker, "Backbone-induced semiconducting behavior in short DNA wires," Phys. Rev. B **65**, 241314 (2002).

[72] S. Roche, D. Bicout, E. Maciá, and E. Kats, "Long Range Correlations in DNA: Scaling Properties and Charge Transfer Efficiency," Phys. Rev. Lett. **91**, 228101 (2003).

[73] S. Roche, "Sequence Dependent DNA-Mediated Conduction," Phys. Rev. Lett. **91**, 108101 (2003).

[74] F. Palmero, J. F. R. Archilla, D. Hennig, and F. R. Romero, "Effect of base-pair inhomogeneities on charge transport along the DNA molecule, mediated by twist and radial polarons," New J. Phys. **6**, 13 (2004).

[75] H. Yamada, "Localization of electronic states in chain models based on real DNA sequence," Phys. Lett. A **332**, 65 – 73 (2004).

[76] V. M. Apalkov and T. Chakraborty, "Electron dynamics in a dna molecule," Phys. Rev. B **71**, 033102 (2005).

[77] D. Klotsa, R. A. Römer, and M. S. Turner, "Electronic Transport in DNA," Biophys. J. **89**, 2187 – 2198 (2005).

[78] Y. S. Joe, S. H. Lee, and E. R. Hedin, "Electron transport through asymmetric dna molecules," Phys. Lett. A **374**, 2367 – 2373 (2010).

[79] J. Yi, "Conduction of DNA molecules: A charge-ladder model," Phys. Rev. B **68**, 193103 (2003).

[80] R. A. Caetano and P. A. Schulz, "Sequencing-Independent Delocalization in a DNA-Like Double Chain with Base Pairing," Phys. Rev. Lett. **95**, 126601 (2005).

[81] X. F. Wang and T. Chakraborty, "Charge Transfer via a Two-Strand Superexchange Bridge in DNA," Phys. Rev. Lett. **97**, 106602 (2006).

[82] 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).

[83] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).

[84] J. C. Slater and G. F. Koster, "Simplified LCAO Method for the Periodic Potential Problem," Phys. Rev. **94**, 1498–1524 (1954).

[85] W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond (Dover Books on Physics)*, 2nd ed. (Dover Publications, Ney York, 1989).

[86] D. A. Papaconstantopoulos and M. J. Mehl, "The Slater-Koster tight-binding method: a computationally efficient and accurate approach," J. Phys.: Condens. Matter **15**, R413–R440 (2003).

[87] W. M. C. Foulkes, "Tight-Binding Models and Coulomb Interactions for s, p, and d Electrons," in *Quantum Materials: Experiments and Theory*, Modeling and Simulation, Vol. 6, edited by E. Pavarini, E. Koch, J. van den Brink, and G. Sawatzky (Forschungszentrum Jülich, 2016).

[88] V. M. Orlov, A. Smirnov, and Y. M. Varshavsky, "Ionization potentials and electron-donor ability of nucleic acid babes and their analogues," Tetrahedron Lett. **17**, 4377–4378 (1976).

[89] V. Dwivedi and V. Chua, "Of bulk and boundaries: Generalized transfer matrices for tight-binding models," Phys. Rev. B **93**, 134304 (2016).

[90] L. Molinari, "Spectral duality and distribution of exponents for transfer matrices of block-tridiagonal Hamiltonians," J. Phys. A: Math. Gen. **36**, 4081–4090 (2003).

[91] T.-T. Lu and S.-H. Shiou, "Inverses of 2 × 2 block matrices," Comput. Math. Appl. **43**, 119–129 (2002).

[92] L. L. Pennisi, "Coefficients of the Characteristic Polynomial," Math. Mag. **60**, 31–33 (1987).

[93] J. Mason and D. C. Handscomb, *Chebyshev Polynomials* (Chapman and Hall/CRC, 2002).

[94] 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 Heidelberg, Berlin, Heidelberg, 2007) pp. 1–20.

[95] 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 – 209 (2014).

[96] M. Zarea, Y. Berlin, and M. A. Ratner, "Effect of the reflectional symmetry on the coherent hole transport across DNA hairpins," J. Chem. Phys. **146**, 114105 (2017).

[97] E. L. Albuquerque, M. S. Vasconcelos, M. L. Lyra, and F. A. B. F. de Moura, "Nucleotide correlations and electronic transport of DNA sequences," Phys. Rev. E **71**, 021910 (2005).

[98] V. M. K. Bagci and A. A. Krokhin, "Metal–insulator transition in DNA molecules induced by long-range correlations in the sequence of nucleotides," Chaos Soliton. Fract. **34**, 104–111 (2007).

[99] M. Xu, R. Endres, S. Tsukamoto, M. Kitamura, S. Ishida, and Y. Arakawa, "Conformation and Local Environment Dependent Conductance of DNA Molecules," Small **1**, 1168–1172 (2005).

[100] K. Iguchi, "π-electrons in a single strand DNA: a phenomenological approach," Int. J. Mod. Phys. B **18**, 1845–1910 (2004).

[101] S. Kundu and S. N. Karmakar, "Localization phenomena in a dna double-helix structure: A twisted ladder model," Phys. Rev. E **89**, 032719 (2014).

[102] S. Malakooti, E. Hedin, and Y. Joe, "Tight-binding approach to strain-dependent DNA electronics," J. Appl. Phys. **114**, 014701 (2013).

[103] S. Varela, V. Mujica, and E. Medina, "Effective spin-orbit couplings in an analytical tight-binding model of DNA: Spin filtering and chiral spin transport," Phys. Rev. B **93**, 155436 (2016).

[104] E. Díaz, "Analysis of the interband optical transitions: Characterization of synthetic DNA band structure," J. Chem. Phys. **128**, 175101 (2008).

[105] E. Maciá, "Electronic structure and transport properties of double-stranded Fibonacci DNA," Phys. Rev. B **74**, 245105 (2006).

[106] E. Maciá and S. Roche, "Backbone-induced effects in the charge transport efficiency of synthetic DNA molecules," Nanotechnology **17**, 3002–3007 (2006).

[107] A. MacKinnon and B. Kramer, "The scaling theory of electrons in disordered solids: Additional numerical results," Z. Phys. B **53**, 1–13 (1983).

[108] 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. **2**, 035013 (2018).

[109] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).

[110] A. Capobianco, A. Landi, and A. Peluso, "Modeling DNA oxidation in water," Phys. Chem. Chem. Phys. **19**, 13571–13578 (2017).

[111] H. Sugiyama and I. Saito, "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. **118**, 7063–7068 (1996).

[112] M. Hutter and T. Clark, "On the Enhanced Stability of the Guanine–Cytosine Base-Pair Radical Cation," J. Am. Chem. Soc. **118**, 7574–7577 (1996).

[113] H. Zhang, X.-Q. Li, P. Han, X. Y. Yu, and Y. Yan, "A partially incoherent rate theory of long-range charge transfer in deoxyribose nucleic acid," J. Chem. Phys. **117**, 4578–4584 (2002).

[114] X. Li, Z. Cai, and M. D. Sevilla, "Investigation of Proton Transfer within DNA Base Pair Anion and Cation Radicals by Density Functional Theory (DFT)," J. Phys. Chem. B **105**, 10115–10123 (2001).

[115] X. Li, Z. Cai, and M. D. Sevilla, "Energetics of the Radical Ions of the AT and AU Base Pairs: A Density Functional Theory (DFT) Study," J. Phys. Chem. A **106**, 9345–9351 (2002).

[116] M. K. Shukla and J. Leszczynski, "A Theoretical Study of Excited State Properties of Adenine-Thymine and Guanine-Cytosine Base Pairs," J. Phys. Chem. A **106**, 4709–4717 (2002).

[117] A. A. Voityuk, J. Jortner, M. Bixon, and N. Rösch, "Electronic coupling between Watson–Crick pairs for hole transfer and transport in desoxyribonucleic acid," J. Chem. Phys. **114**, 5614–5620 (2001).

[118] A. Ivanova, P. Shushkov, and N. Rösch, "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 **112**, 7106–7114 (2008).

[119] A. Migliore, S. Corni, D. Varsano, M. L. Klein, and R. Di Felice, "First Principles Effective Electronic Couplings for Hole Transfer in Natural and Size-Expanded DNA," J. Phys. Chem. B **113**, 9402–9415 (2009).

[120] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, "Metallic Phase with Long-Range Orientational Order and No Translational Symmetry," Phys. Rev. Lett. **53**, 1951–1953 (1984).

[121] International Union of Crystallography (IUCr), "Report of the Executive Committee for 1991," Acta Crystallogr. A **48**, 922–946 (1992).

[122] Bindi, L. and Steinhardt, P. J. and Yao, N. and Lu, P. J., "Icosahedrite, $Al_{63}Cu_{24}Fe_{13}$, the first natural quasicrystal," Am. Mineral. **96**, 928–931 (2011).

[123] L. Bindi, N. Yao, C. Lin, L. S. Hollister, C. L. Andronicos, V. V. Distler, M. P. Eddy, A. Kostin, V. Kryachko, G. J. MacPherson, W. M. Steinhardt, M. Yudovskaya, and P. J. Steinhardt, "Decagonite, $Al_{71}Ni_{24}Fe_5$, a quasicrystal with decagonal symmetry from the Khatyrka CV3 carbonaceous chondrite," Am. Mineral. **100**, 2340–2343 (2015).

[124] L. Bindi, C. Lin, M. Chi, and P. J. Steinhardt, "Collisions in outer space produced an icosahedral phase in the Khatyrka meteorite never observed previously in the laboratory," Sci. Rep. **6**, 38117 (2016).

[125] L. Bindi and C. J. Stanley, "Natural versus synthetic quasicrystals: analogies and differences in the optical behavior of icosahedral and decagonal quasicrystals," Rend. Lincei Sci. Fis. Nat. (2019), 10.1007/s12210-019-00859-9.

[126] L. Bindi, N. Yao, C. Lin, L. S. Hollister, C. L. Andronicos, V. V. Distler, M. P. Eddy, A. Kostin, V. Kryachko, G. J. MacPherson, W. M. Steinhardt, M. Yudovskaya, and P. J. Steinhardt, "Natural quasicrystal with decagonal symmetry," Sci. Rep. **5**, 9111 (2016).

[127] E. Maciá Barber, *Aperiodic Structures in Condensed Matter: Fundamentals and Applications* (CRC Press, Boca Raton, 2008).

[128] E. Maciá, "On the Nature of Electronic Wave Functions in One-Dimensional Self-Similar and Quasiperiodic Systems," ISRN condens. matter phys. **2014**, 165943 (2014).

[129] S. Kosuri and G. M. Church, "Large-scale de novo DNA synthesis: technologies and applications," Nat. Methods **11**, 499–507 (2014).

[130] S. Beaucage and M. Caruthers, "Deoxynucleoside phosphoramidites–a new class of key intermediates for deoxypolynucleotide synthesis," Tetrahedron Letters **22**, 1859–1862 (1981).

[131] S. Palluk, D. H. Arlow, T. de Rond, S. Barthel, J. S. Kang, R. Bector, H. M. Baghdassarian, A. N. Truong, P. W. Kim, A. K. Singh, N. J. Hillson, and J. D. Jay Keasling, "De novo DNA synthesis using polymerase-nucleotide conjugates," Nat. Biotechnol. **36**, 645–650 (2018).

[132] L. Sigler, *Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation* (Springer Verlag, New York, 2003).

[133] P. Singh, "The so-called fibonacci numbers in ancient and medieval India," Hist. Math. **12**, 229 – 244 (1985).

[134] E. Prouhet, "Mémoire sur quelques relations entre les puissances des nombres," C. R. Acad. Sci. Paris **33**, 225 (1851), (in French).

[135] T. Nagell, A. Selberg, S. Selberg, and K. Thalberg, eds., *Selected mathematical papers of Axel Thue* (Universitetsforlaget, Oslo, 1977).

[136] H. M. Morse, "Recurrent Geodesics on a Surface of Negative Curvature," Trans. Amer. Math. Soc. **22**, 84–100 (1906).

[137] H. S. Shapiro, *Extremal problems for polynomials and power series*, Master's thesis, Massachusetts Institute of Technology (1951).

[138] W. Rudin, "Some theorems on Fourier coefficients," Proc. Amer. Math. Soc. **10**, 855–855 (1959).

[139] G. Cantor, "Über unendliche, lineare Punktmannigfaltigkeiten," Math. Ann. **21**, 545–586 (1883), (in German).

[140] The On-Line Encyclopedia of Integer Sequences, published electronically at https://oeis.org, (2010), sequence A000002.

[141] W. Kolakoski, "Advanced Problems: Problem 5304," Am. Math. Mon. **72**, 674 (1965).

[142] R. Oldenburger, "Exponent trajectories in symbolic dynamics," Trans. Amer. Math. Soc. **46**, 453–466 (1939).

[143] B. Sing, "Kolakoski sequences – an example of aperiodic order," J. Non. Cryst. Solids **334-335**, 100–104 (2004).

[144] The On-Line Encyclopedia of Integer Sequences, published electronically at https://oeis.org, (2010), sequence A001083.

[145] O. Perron, "Zur Theorie der Matrices," Math. Ann. **64**, 248–263 (1907), (in German).

[146] G. Frobenius, "Über Matrizen aus nicht Negativen Elementen," Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin , 456–477 (1912).

[147] M. Baake and U. Grimm, *Aperiodic Order: Volume 1, A Mathematical Invitation* (Cambridge University Press, Cambridge, 2013).

[148] V. Berthé and A. Siegel, "Basic notions on substitutions," in *Substitutions in Dynamics, Arithmetics and Combinatorics*, edited by N. P. Fogg, V. Berthé, S. Ferenczi, C. Mauduit, and A. Siegel (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002) pp. 1–32.

[149] V. Canterini and A. Siegel, "Geometric Representation of Substitutions of Pisot Type," Trans. Amer. Math. Soc. **353**, 5121–5144 (2001).

[150] E. Hewitt and K. Stromberg, *Real and Abstract Analysis: A modern treatment of the theory of functions of a real variable* (Springer Berlin Heidelberg, 1965).

[151] E. Bombieri and J. E. Taylor, "Which distributions of matter diffract? An initial investigation," J. Phys. Colloq. **47 (C3)**, 19–28 (1986).

[152] E. Bombieri and J. E. Taylor, "Quasicrystals, Tilings, and Algebraic Number Theory: Some Preliminary Connections," in *The Legacy of Sonya Kovalevskaya*, Contemporary Mathematics, Vol. 64, edited by L. Keen (American Mathematical Society, 1987) pp. 241–264.

[153] B. Solomyak, "Dynamics of Self-Similar Tilings," Ergod. Th. & Dynam. Sys. **17**, 695–738 (1997).

[154] J. M. Luck, C. Godrèche, A. Janner, and T. Janssen, "The nature of the atomic surfaces of quasiperiodic self-similar structures," J. Phys. A: Math. Gen. **26**, 1951–1999 (1993).

[155] M. Kolář, "New class of one-dimensional quasicrystals," Phys. Rev. B **47**, 5489–5492 (1993).

[156] M. Kolář, B. Iochum, and L. Raymond, "Structure factor of 1D systems (superlattices) based on two-letter substitution rules. I. delta (Bragg) peaks," J. Phys. A: Math. Gen. **26**, 7343–7366 (1993).

[157] J. Birch, M. Severin, U. Wahlström, Y. Yamamoto, G. Radnoczi, R. Riklund, J.-E. Sundgren, and L. R. Wallenberg, "Structural characterization of precious-mean quasiperiodic Mo/V single-crystal superlattices grown by dual-target magnetron sputtering," Phys. Rev. B **41**, 10398–10407 (1990).

[158] X. Fu, Y. Liu, P. Zhou, and W. Sritrakool, "Perfect self-similarity of energy spectra and gap-labeling properties in one-dimensional Fibonacci-class quasilattices," Phys. Rev. B **55**, 2882–2889 (1997).

[159] E. Maciá, "Spectral Classification of One-Dimensional Binary Aperiodic Crystals: An Algebraic Approach," Ann. Phys. (Berl.) **529**, 1700079 (2017).

[160] M. Baake, W. Grimm, and N. Mañibo, "Spectral analysis of a family of binary inflation rules," Lett. Math. Phys. **108**, 1783–1805 (2018).

[161] B. Sing, "Kolakoski-(2m,2n) are limit-periodic model sets," J. Math. Phys. **44**, 899–912 (2003).

[162] M. Gover, "The eigenproblem of a tridiagonal 2-Toeplitz matrix," Linear Algebra Appl. **197-198**, 63 – 78 (1994).

[163] S. Kouachi, "Eigenvalues and eigenvectors of tridiagonal matrices," Electron. J. Linear Algebra **15**, 115–133 (2006).

[164] R. Álvarez-Nodarse, J. Petronilho, and N. Quintero, "On some tridiagonal k-Toeplitz matrices: Algebraic and analytical aspects. Applications," J. Comput. Appl. Math. **184**, 518 – 537 (2005).

[165] W.-C. Yueh and S. S. Cheng, "Explicit eigenvalues and inverses of tridiagonal Toeplitz matrices with four perturbed corners," ANZIAM J. **49**, 361–387 (2008).

[166] C. M. da Fonseca, "The characteristic polynomial of some perturbed tridiagonal k-Toeplitz matrices," Appl. Math. Sci. **1**, 59–67 (2007).

[167] J. M. Cerveró and A. Rodríguez, "Infinite chain of different deltas: A simple model for a quantum wire," Eur. Phys. J. B **30**, 239–251 (2002).
[168] J. M. Cerveró and A. Rodríguez, "Absorption in atomic wires," Phys. Rev. A **70**, 052705 (2004).
[169] A. Rodríguez and J. M. Cerveró, "One-dimensional disordered wires with Pöschl-Teller potentials," Phys. Rev. B **74**, 104201 (2006).
[170] A. Rodríguez, "One-dimensional models of disordered quantum wires: general formalism," J. Phys. A: Math. Gen. **39**, 14303 (2006).
[171] D. H. Lee and J. D. Joannopoulos, "Simple scheme for surface-band calculations. I," Phys. Rev. B **23**, 4988–4996 (1981).
[172] A. D. Stone, J. D. Joannopoulos, and D. J. Chadi, "Scaling studies of the resistance of the one-dimensional Anderson model with general disorder," Phys. Rev. B **24**, 5583–5596 (1981).
[173] L. Molinari, "Transfer matrices and tridiagonal-block Hamiltonians with periodic and scattering boundary conditions," J. Phys. A: Math. Gen. **30**, 983–997 (1997).
[174] P. Yeh, A. Yariv, and C.-S. Hong, "Electromagnetic propagation in periodic stratified media. I. General theory," J. Opt. Soc. Am. **67**, 423–438 (1977).
[175] R. Tsu and L. Esaki, "Tunneling in a finite superlattice," Appl. Phys. Lett. **22**, 562–564 (1973).
[176] E. G. Emberly and G. Kirczenow, "Theoretical study of electrical conduction through a molecule connected to metallic nanocontacts," Phys. Rev. B **58**, 10911–10920 (1998).
[177] T. Kostyrko, "An analytic approach to the conductance and I - V characteristics of polymeric chains," J. Phys.: Condens. Matter **14**, 4393–4405 (2002).
[178] D. R. Smith, S. Schultz, P. Markoš, and C. M. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients," Phys. Rev. B **65**, 195104 (2002).
[179] J. Baringhaus, M. Ruan, F. Edler, A. Tejeda, M. Sicot, A. Taleb-Ibrahimi, A.-P. Li, Z. Jiang, E. H. Conrad, C. Berger, C. Tegenkamp, and W. A. de Heer, "Exceptional ballistic transport in epitaxial graphene nanoribbons," Nature **506**, 349–354 (2014).
[180] E. Maciá, F. Triozon, and S. Roche, "Contact-dependent effects and tunneling currents in DNA molecules," Phys. Rev. B **71**, 113106 (2005).
[181] A.-M. Guo and H. Xu, "Effects of interbase electronic coupling and electrode on charge transport through short DNA molecules: A numerical study," Phys. Lett. A **364**, 48 – 53 (2007).
[182] 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. **112**, 094703 (2012).
[183] K. Lambropoulos and C. Simserides, "Electronic structure and charge transport properties of atomic carbon wires," Phys. Chem. Chem. Phys. **19**, 26890–26897 (2017).
[184] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, "C60: Buckminsterfullerene," Nature **318**, 162–163 (1985).

[185] S. Iijima, "Helical microtubules of graphitic carbon," Nature **354**, 56–58 (1991).
[186] L. V. Radushkevich and V. M. Lukyanovich, "The structure of carbon forming in thermal decomposition of carbon monoxide on an iron catalyst," Russ. J. Phys. Chem. **26**, 88–95 (1952), (in Russian).
[187] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric Field Effect in Atomically Thin Carbon Films," Science **306**, 666–669 (2004).
[188] D. Malko, C. Neiss, F. Viñes, and A. Görling, "Competition for Graphene: Graphynes with Direction-Dependent Dirac Cones," Phys. Rev. Lett. **108**, 086804 (2012).
[189] F. Banhart, "Chains of carbon atoms: A vision or a new nanomaterial?" Beilstein J. Nanotechnol. **6**, 559–569 (2015).
[190] C. S. Casari, M. Tommasini, R. R. Tykwinski, and A. Milani, "Carbon-atom wires: 1-d systems with tunable properties," Nanoscale **8**, 4414–4435 (2016).
[191] O. Cretu, A. R. Botello-Mendez, I. Janowska, C. Pham-Huu, J.-C. Charlier, and F. Banhart, "Electrical Transport Measured in Atomic Carbon Chains," Nano Lett. **13**, 3487–3493 (2013).
[192] W. A. Chalifoux and R. R. Tykwinski, "Synthesis of polyynes to model the sp-carbon allotrope carbyne," Nat. Chem. **2**, 967–971 (2010).
[193] L. Shi, P. Rohringer, K. Suenaga, Y. Niimi, J. Kotakoski, J. C. Meyer, H. Peterlik, M. Wanko, S. Cahangirov, A. Rubio, Z. J. Lapin, L. Novotny, P. Ayala, and T. Pichler, "Confined linear carbon chains as a route to bulk carbyne," Nat. Mater. **15**, 634–639 (2016).
[194] G. Wu and J. Dong, "Anomalous heat conduction in a carbon nanowire: Molecular dynamics calculations," Phys. Rev. B **71**, 115410 (2005).
[195] M. Wang and S. Lin, "Ballistic Thermal Transport in Carbyne and Cumulene with Micron-Scale Spectral Acoustic Phonon Mean Free Path," Sci. Rep. **5**, 18122 (2015).
[196] X. Liu, G. Zhang, and Y.-W. Zhang, "Tunable Mechanical and Thermal Properties of One-Dimensional Carbyne Chain: Phase Transition and Microscopic Dynamics," J. Phys. Chem. C **119**, 24156–24164 (2015).
[197] Y. Deng and S. W. Cranford, "Thermal conductivity of 1D carbyne chains," Comp. Mater. Sci. **129**, 226 – 230 (2017).
[198] M. Tommasini, A. Milani, D. Fazzi, M. Del Zoppo, C. Castiglioni, and G. Zerbi, "Modeling phonons of carbon nanowires," Physica E **40**, 2570 – 2576 (2008).
[199] A. Milani, M. Tommasini, and G. Zerbi, "Carbynes phonons: A tight binding force field," J. Chem. Phys. **128**, 064501 (2008).
[200] N. R. Agarwal, A. Lucotti, D. Fazzi, M. Tommasini, C. Castiglioni, W. Chalifoux, and R. R. Tykwinski, "Structure and chain polarization of long polyynes investigated with infrared and Raman spectroscopy," J. Raman Spectrosc. **44**, 1398–1410 (2013).
[201] A. K. Nair, S. W. Cranford, and M. J. Buehler, "The minimal nanowire: Mechanical properties of carbyne," EPL **95**, 16002 (2011).
[202] L. Ravagnan, P. Piseri, M. Bruzzi, S. Miglio, G. Bongiorno, A. Baserga, C. S. Casari, A. Li Bassi, C. Lenardi, Y. Yamaguchi, T. Wakabayashi, C. E. Bottani,

and P. Milani, "Influence of Cumulenic Chains on the Vibrational and Electronic Properties of $sp - sp^2$ Amorphous Carbon," Phys. Rev. Lett. **98**, 216103 (2007).

[203] B. Standley, W. Bao, H. Zhang, J. Bruck, C. N. Lau, and M. Bockrath, "Graphene-Based Atomic-Scale Switches," Nano Lett. **8**, 3345–3349 (2008).

[204] C. Wang, A. S. Batsanov, M. R. Bryce, S. Martín, R. J. Nichols, S. J. Higgins, V. M. García-Suárez, and C. J. Lambert, "Oligoyne Single Molecule Wires," J. Am. Chem. Soc. **131**, 15647–15654 (2009).

[205] P. Moreno-García, M. Gulcur, D. Z. Manrique, T. Pope, W. Hong, V. Kaliginedi, C. Huang, A. S. Batsanov, M. R. Bryce, C. Lambert, and T. Wandlowski, "Single-Molecule Conductance of Functionalized Oligoynes: Length Dependence and Junction Evolution," J. Am. Chem. Soc. **135**, 12228–12240 (2013).

[206] M. Tommasini, D. Fazzi, A. Milani, M. Del Zoppo, C. Castiglioni, and G. Zerbi, "Effective hamiltonian for π electrons in linear carbon chains," Chem. Phys. Lett. **450**, 86 – 90 (2007).

[207] A. La Magna, I. Deretzis, and V. Privitera, "Insulator-metal transition in biased finite polyyne systems," Eur. Phys. J. B **70**, 311–316 (2009).

[208] P. N. D'yachkov, V. A. Zaluev, E. Y. Kocherga, and N. R. Sadykov, "Tight Binding Model of Quantum Conductance of Cumulenic and Polyynic Carbynes," J. Phys. Chem. C. **117**, 16306–16315 (2013).

[209] Y. Zhang, Y. Su, E. S.-W. Wang, L.and Kong, X. Chen, and Y. Zhang, "A one-dimensional extremely covalent material: monatomic carbon linear chain," Nanoscale Res. Lett. **6**, 577 (2011).

[210] A. La Torre, A. Botello-Mendez, W. Baaziz, J.-C. Charlier, and F. Banhart, "Strain-induced metal-semiconductor transition observed in atomic carbon chains," Nat. Commun. **6**, 6636 (2015).

[211] F. Ben Romdhane, J.-J. Adjizian, J.-C. Charlier, and F. Banhart, "Electrical transport through atomic carbon chains: The role of contacts," Carbon **122**, 92–97 (2017).

[212] R. E. Peierls, *Quantum Theory of Solids* (Oxford University Press, Oxford, 2001).

[213] A. Milani, M. Tommasini, and G. Zerbi, "Connection among Raman wavenumbers, bond length alternation and energy gap in polyynes," J. Raman Spectrosc. **40**, 1931–1934 (2009).

[214] A. Milani, M. Tommasini, M. Del Zoppo, C. Castiglioni, and G. Zerbi, "Carbon nanowires: Phonon and π-electron confinement," Phys. Rev. B **74**, 153418 (2006).

[215] M. Liu, V. I. Artyukhov, H. Lee, F. Xu, and B. I. Yakobson, "Carbyne from First Principles: Chain of C Atoms, a Nanorod or a Nanorope," ACS Nano **7**, 10075–10082 (2013).

[216] V. I. Artyukhov, M. Liu, and B. I. Yakobson, "Mechanically Induced Metal–Insulator Transition in Carbyne," Nano Lett. **14**, 4224–4229 (2014).

[217] A. Milani, M. Tommasini, V. Barbieri, A. Lucotti, V. Russo, F. Cataldo, and C. S. Casari, "Semiconductor-to-Metal Transition in Carbon-Atom Wires Driven by sp2 Conjugated End Groups," J. Phys. Chem. C **121**, 10562–10570 (2017).

[218] S. Cahangirov, M. Topsakal, and S. Ciraci, "Long-range interactions in carbon atomic chains," Phys. Rev. B **82**, 195444 (2010).

[219] M. Menon and R. E. Allen, "Simulations of atomic processes at semiconductor surfaces: General method and chemisorption on GaAs(110)," Phys. Rev. B **38**, 6196–6205 (1988).

[220] N. Lathiotakis and A. N. Andriotis, "The applicability of scaling laws in tight-binding molecular-dynamics," Solid State Commun. **87**, 871 – 875 (1993).

[221] L. G. D. Hawke, G. Kalosakas, and C. Simserides, "Empirical LCAO parameters for π molecular orbitals in planar organic molecules," Mol. Phys. **107**, 1755–1771 (2009).

[222] W. M. Haynes, ed., *CRC Handbook of Chemistry and Physics*, 95th ed. (CRC Press, 2014).

[223] F. Marcellán and J. Petronilho, "Eigenproblems for tridiagonal 2-Toeplitz matrices and quadratic polynomial mappings," Linear Algebra Appl. **260**, 169–208 (1997).

[224] R. Landauer, "Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction," IBM J. Res. Dev. **1**, 223–231 (1957).

[225] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, "Generalized many-channel conductance formula with application to small rings," Phys. Rev. B **31**, 6207–6215 (1985).

[226] M. Buttiker, "Symmetry of electrical conduction," IBM J. Res. Dev. **32**, 317–334 (1988).

[227] G. C. Liang, A. W. Ghosh, M. Paulsson, and S. Datta, "Electrostatic potential profiles of molecular conductors," Phys. Rev. B **69**, 115302 (2004).

[228] S. Roche and E. Maciá, "Electronic Transport and Thermopower in Aperiodic DNA Sequences," Mod. Phys. Lett. B **18**, 847–871 (2004).

[229] M. S. Dresselhaus and G. Dresselhaus, "Intercalation compounds of graphite," Adv. Phys. **51**, 1–186 (2002).

[230] V. Mujica, M. Kemp, A. Roitberg, and M. Ratner, "Current-voltage characteristics of molecular wires: Eigenvalue staircase, Coulomb blockade, and rectification," J. Chem. Phys. **104**, 7296–7305 (1996).

[231] K. Lambropoulos and C. Simserides, "Periodic, quasiperiodic, fractal, Kolakoski, and random binary polymers: Energy structure and carrier transport," Phys. Rev. E **99**, 032415 (2019).

[232] R. G. Sarmento, E. L. Albuquerque, P. D. Sesion, U. L. Fulco, and B. P. W. de Oliveira, "Electronic transport in double-strand poly(dG)–poly(dC) DNA segments," Phys. Lett. A **373**, 1486 – 1491 (2009).

[233] R. G. Sarmento, G. A. Mendes, E. L. Albuquerque, U. L. Fulco, M. S. Vasconcelos, O. Ujsághy, V. N. Freire, and E. W. S. Caetano, "The DNA electronic specific heat at low temperature: The role of aperiodicity," Phys. Lett. A **376**, 2413 – 2417 (2012).

[234] S. Kundu and S. N. Karmakar, "Electronic specific heat of DNA: Effects of backbones and disorder," Phys. Lett. A **379**, 1377 – 1383 (2015).

[235] S. Fathizadeh, S. Behnia, and J. Ziaei, "Engineering DNA Molecule Bridge between Metal Electrodes for High-Performance Molecular Transistor: An Environmental Dependent Approach," J. Phys. Chem. B **122**, 2487–2494 (2018).

[236] K. Senthilkumar, F. C. Grozema, C. Fonseca Guerra, F. M. Bickelhaupt, F. D. Lewis, Y. A. Berlin, M. A. Ratner, and L. D. A. Siebbeles, "Absolute Rates of Hole Transfer in DNA," J. Am. Chem. Soc. **127**, 14894–14903 (2005).

[237] A. Bovier and J.-M. Ghez, "Remarks on the spectral properties of tight-binding and Kronig-Penney models with substitution sequences," J. Phys. A: Math. Gen. **28**, 2313–2324 (1995).

[238] J. Bellisard, A. Bovier, and J.-M. Ghez, "Gap Labelling Theorems for One Dimensional Discrete Schrödinger operators," Rev. Math. Phys. **4**, 1–37 (1992).

[239] J. Bellisard, A. Bovier, and J.-M. Ghez, "Discrete Schrödinger Operators with Potentials Generated by Substitutions," in *Differential Equations with Applications to Mathematical Physics*, Mathematics in Science and Engineering, Vol. 192, edited by W. F. Ames, E. M. Harrell, and J. V. Herod (Elsevier, 1993) pp. 13–23.

[240] B. P. W. de Oliveira, E. L. Albuquerque, and M. S. Vasconcelos, "Electronic density of states in sequence dependent DNA molecules," Surf. Sci. **600**, 3770–3774 (2006).

[241] E. Maciá, F. Domínguez-Adame, and A. Sánchez, "Energy spectra of quasiperiodic systems via information entropy," Phys. Rev. E **50**, R679–R682 (1994).

[242] V. I. Oseledets, "A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems," Trans. Moscow Math. Soc. **19**, 197–231 (1968).

[243] A. Crisanti, G. Paladin, and A. Vulpiani, *Products of Random Matrices in Statistical Physics* (Springer Verlag, Berlin, Heidelberg, 1993).

[244] J. A. Scales and E. S. Van Vleck, "Lyapunov Exponents and Localization in Randomly Layered Media," J. Comput. Phys. **133**, 27 – 42 (1997).

[245] A.-M. Guo, "Long-range correlation and charge transfer efficiency in substitutional sequences of DNA molecules," Phys. Rev. E **75**, 061915 (2007).

[246] C. Simserides, *Quantum Optics and Lasers* (Hellenic Academic Libraries Link, Athens, 2016) (in Greek).

[247] K. Lambropoulos, K. Kaklamanis, G. Georgiadis, and C. Simserides, "THz and above THz electron or hole oscillations in DNA dimers and trimers," Ann. Phys. (Berl.) **526**, 249–258 (2014).

[248] K. Lambropoulos, *Charge transfer in small DNA segments: description at the base-pair level*, Diploma thesis, National and Kapodistrian University of Athens, Greece (2014), (in Greek).

[249] K. Kaklamanis, *Charge transfer in small DNA segments: description at the single-base level*, Diploma thesis, National and Kapodistrian University of Athens, Greece (2015), (in Greek).

[250] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," Phys. Rev. **136**, B864–B871 (1964).

[251] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," Phys. Rev. **140**, A1133–A1138 (1965).

[252] E. Runge and E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems," Phys. Rev. Lett. **52**, 997–1000 (1984).

[253] K. Lopata and N. Govind, "Modeling Fast Electron Dynamics with Real-Time Time-Dependent Density Functional Theory: Application to Small Molecules and Chromophores," J. Chem. Theory Comput. **7**, 1344–1355 (2011).

[254] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP)," Chem. Phys. Lett. **393**, 51 – 57 (2004).

[255] W. J. Hehre, R. Ditchfield, and J. A. Pople, "Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules," J. Chem. Phys. **56**, 2257–2261 (1972).

[256] M. D. Newton and N. Sutin, "Electron Transfer Reactions in Condensed Phases," Annu. Rev. Phys. Chem. **35**, 437–480 (1984).

[257] J. R. Bolton, N. Mataga, and G. McLendon, eds., *Electron Transfer in Inorganic, Organic, and Biological Systems* (American Chemical Society, Washington, D.C., 1991).

[258] A. Farazdel, M. Dupuis, E. Clementi, and A. Aviram, "Electric-field induced intramolecular electron transfer in spiro .pi.-electron systems and their suitability as molecular electronic devices. A theoretical study," J. Am. Chem. Soc. **112**, 4206–4214 (1990).

[259] C. Vantaraki, *Carrier transfer in periodic polymer B-DNA segments based on the G-C monomer: Base-pair-level description within the Tight-Binding approach*, Diploma thesis, National and Kapodistrian University of Athens, Greece (2017), (in Greek).

[260] P. Bilia, *Carrier transfer in periodic polymer B-DNA segments based on the G-C and A-T monomers with purine on purine: Base-pair-level description within the Tight-Binding Approach*, Diploma thesis, National and Kapodistrian University of Athens, Greece (2019), (in Greek).

[261] K. Lambropoulos, M. Mantela, and C. Simserides, "Frequency content of carrier oscillations along b-dna polymers," in *2017 Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL)* (2017) pp. 186–193.

[262] M. Mantela, *Charge transfer in aperiodic B-DNA segments: Tight-Binding description at the base-pair level]*, Master's thesis, National and Kapodistrian University of Athens, Greece (2017), (in Greek).

[263] M. Theodorakou, *Charge transfer in aperiodic B-DNA segments, made of different base pairs: Tight-Binding description at the base-pair level*, Master's thesis, National and Kapodistrian University of Athens, Greece (2018), (in Greek).

[264] E. Meggers, M. E. Michel-Beyerle, and B. Giese, "Sequence Dependent Long Range Hole Transport in DNA," J. Am. Chem. Soc. **120**, 12950–12955 (1998).

[265] B. Giese, J. Amaudrut, A.-K. Köhler, M. Spormann, and S. Wessely, "Direct observation of hole transfer through DNA by hopping between adenine bases and by tunnelling," Nature **412**, 318–320 (2001).

[266] K. Kawai and T. Majima, "Hole Transfer Kinetics of DNA," Acc. Chem. Res. **46**, 2616–2625 (2013).

[267] F. D. Lewis, T. Wu, Y. Zhang, R. L. Letsinger, S. R. Greenfield, and M. R. Wasielewski, "Distance-Dependent Electron Transfer in DNA Hairpins," Science **277**, 673–676 (1997).

[268] C. Wan, T. Fiebig, O. Schiemann, J. K. Barton, and A. H. Zewail, "Femtosecond direct observation of charge transfer between bases in DNA," Proc. Natl. Acad. Sci. USA **97**, 14052–14055 (2000).

[269] T. Takada, K. Kawai, M. Fujitsuka, and T. Majima, "Direct observation of hole transfer through double-helical DNA over 100 Å," Proc. Natl. Acad. Sci. USA **101**, 14002–14006 (2004).

[270] S. M. Conron, A. K. Thazhathveetil, M. R. Wasielewski, A. L. Burin, and F. D. Lewis, "Direct Measurement of the Dynamics of Hole Hopping in Extended DNA G-Tracts. An Unbiased Random Walk," J. Am. Chem. Soc. **132**, 14388–14390 (2010).

[271] J. Vura-Weis, M. R. Wasielewski, A. K. Thazhathveetil, and F. D. Lewis, "Efficient Charge Transport in DNA Diblock Oligomers," J. Am. Chem. Soc. **131**, 9722–9727 (2009).

[272] J. K. Cullum and R. A. Willoughby, "Tridiagonal Matrices," in *Lanczos Algorithms for Large Symmetric Eigenvalue Computations* (Society for Industrial and Applied Mathematics, 2002) pp. 76–91.

[273] O. Marques and P. B. Vasconcelos, "Computing the bidiagonal svd through an associated tridiagonal eigenproblem," in *High Performance Computing for Computational Science – VECPAR 2016*, edited by I. Dutra, R. Camacho, J. Barbosa, and O. Marques (Springer International Publishing, Cham, 2017) pp. 64–74.

[274] R. Ralha, "Perturbation Splitting for More Accurate Eigenvalues," SIAM J. Matrix Anal. Appl. **31**, 75–91 (2009).