Modern Electronic Structure Theory using Tensor Product States

Vibin Abraham

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Chemistry

Nicholas Mayhall, Chair
Daniel Crawford
Eduard Valeyev
Kyungwha Park

October 27, 2021
Blacksburg, Virginia

Keywords: Tensor Product State, Wavefunction, Strong Correlation, Singlet Fission,
Excited States

Modern Electronic Structure Theory using Tensor Product States

Vibin Abraham

(ABSTRACT)

Strongly correlated systems have been a major challenge for a long time in the field of theoretical chemistry. For such systems, the relevant portion of the Hilbert space scales exponentially, preventing efficient simulation on large systems. However, in many cases, the Hilbert space can be partitioned into clusters on the basis of strong and weak interactions. In this work, we mainly focus on an approach where we partition the system into smaller orbital clusters in which we can define many-particle cluster states and use traditional many-body methods to capture the rest of the inter-cluster correlations.

This dissertation can be mainly divided into two parts. In the first part of this dissertation, the clustered ansatz, termed as *tensor product states* (TPS), is used to study large strongly correlated systems. In the second part, we study a particular type of strongly correlated system, correlated triplet pair states that arise in singlet fission.

The many-body expansion (MBE) is an efficient tool that has a long history of use for calculating interaction energies, binding energies, lattice energies, and so on. We extend the incremental full configuration interaction originally proposed for a Slater determinant to a tensor product state (TPS) based wavefunction. By partitioning the active space into smaller orbital clusters, our approach starts from a cluster mean-field reference TPS configuration and includes the correlation contribution of the excited TPSs using a many-body expansion. This method, named *cluster many-body expansion* (cMBE), improves the convergence of MBE at lower orders compared to directly doing a block-based MBE from an RHF reference. The performance of the cMBE method is also tested on a graphene nano-sheet with a very

large active space of 114 electrons in 114 orbitals, which would require 10^{66} determinants for the exact FCI solution.

Selected CI (SCI) using determinants becomes intractable for large systems with strong correlation. We introduce a method for SCI algorithms using tensor product states which exploits local molecular structure to significantly reduce the number of SCI variables. We demonstrate the potential of this method, called tensor product selected configuration interaction (TPSCI), using a few model Hamiltonians and molecular examples. These numerical results show that TPSCI can be used to significantly reduce the number of SCI variables in the variational space, and thus paving a path for extending these deterministic and variational SCI approaches to a wider range of physical systems.

The extension of the TPSCI algorithm for excited states is also investigated. TPSCI with perturbative corrections provides accurate excitation energies for low-lying triplet states with respect to extrapolated results. In the case of traditional SCI methods, accurate excitation energies are obtained only after extrapolating calculations with large variational dimensions compared to TPSCI. We provide an intuitive connection between lower triplet energy manifolds with Hückel molecular orbital theory, providing a many-body version of Hückel theory for excited triplet states.

The n-body Tucker ansatz (which is a truncated TPS wavefunction) developed in our group provides a good approximation to the low-lying states of a clusterable spin system. In this approach, a Tucker decomposition is used to obtain local cluster states which can be truncated to prune the full Hilbert space of the system. As a truncated variational approach, it has been observed that the self-consistently optimized n-body Tucker method is not size-extensive, a property important for many-body methods. We explore the use of perturbation theory and linearized coupled-cluster methods to obtain a robust yet efficient approximation. Perturbative corrections to the n-body Tucker method have been implemented for the

Heisenberg Hamiltonian and numerical data for various lattices and molecular systems has been presented to show the applicability of the method.

In the second part of this dissertation, we focus on studying a particular type of strongly correlated states that occurs in singlet fission material. The correlated triplet pair state ^{1}(TT) is a key intermediate in the singlet fission process, and understanding the mechanism by which it separates into two independent triplet states is critical for leveraging singlet fission for improving solar cell efficiency. This separation mechanism is dominated by two key interactions: (i) the exchange interaction (K) between the triplets which leads to the spin splitting of the biexciton state into ^{1}(TT),^{3}(TT) and ^{5}(TT) states, and (ii) the triplet-triplet energy transfer integral (t) which enables the formation of the spatially separated (but still spin entangled) state ^{1}(T...T). We develop a simple ab initio technique to compute both the triplet-triplet exchange (K) and triplet-triplet energy transfer coupling (t). Our key findings reveal new conditions for successful correlated triplet pair state dissociation. The biexciton exchange interaction needs to be ferromagnetic or negligible compared to the triplet energy transfer for favorable dissociation. We also explore the effect of chromophore packing to reveal geometries where these conditions are achieved for tetracene.

We also provide a simple connectivity rule to predict whether the through-bond coupling will be stabilizing or destabilizing for the (TT) state in covalently linked singlet fission chromophores. By drawing an analogy between the chemical system and a simple spin-lattice, one is able to determine the ordering of the multi-exciton spin state via a generalized usage of Ovchinnikov's rule. In the case of meta connectivity, we predict ^{5}(TT) to be formed and this is later confirmed by experimental techniques like time-resolved electron spin resonance (TR-ESR).

This work was supported by the National Science Foundation (Award No. 1752612).

Modern Electronic Structure Theory using Tensor Product States

Vibin Abraham

(GENERAL AUDIENCE ABSTRACT)

The study of the correlated motion of electrons in molecules and materials allows scientists to gain useful insights into many physical processes like photosynthesis, enzyme catalysis, superconductivity, chemical reactions and so on. Theoretical quantum chemistry tries to study the electronic properties of chemical species. The exact solution of the electron correlation problem is exponentially complex and can only be computed for small systems. Therefore, approximations are introduced for practical calculations that provide good results for ground state properties like energy, dipole moment, etc. Sometimes, more accurate calculations are required to study the properties of a system, because the system may not adhere to the assumptions that are made in the methods used. One such case arises in the study of strongly correlated molecules.

In this dissertation, we present methods which can handle strongly correlated cases. We partition the system into smaller parts, then solve the problem in the basis of these smaller parts. We refer to this block based wavefunction as tensor product states and they provide accurate results while avoiding the exponential scaling of the full solution. We present accurate energies for a wide variety of challenging cases, including bond breaking, excited states and π conjugated molecules. Additionally, we also investigate molecular systems that can be used to increase the efficiency of solar cells. We predict improved solar efficiency for a chromophore dimer, a result which is later experimentally verified.

Dedication

To my parents, Mary and Abraham.

Acknowledgments

Firstly I would like to thank my family back in India. Namely I thank my parents for their constant support. If not for them, I would not be here. I would also like to thank my sister and brother-in-law for their support and patience and constantly sending me pictures and videos of the kids. I feel blessed to have a supportive family, including my uncles, aunts and cousins, who have been a support system for my parents in my absence. They have been very close to me and always made me at home anytime I visited.

I have made many friends during my stay in the theoretical chemistry department. First of all, I would like to thank Ruhee for being a great friend for the past 10 years and always planning the best trips. Shannon for being a great friend, proof reading my scientific documents and helping with debugging codes. Ben for countless hours of scientific discussions. Ruhee, Shannon and Ben have been my closest friends during the past few years. I've always enjoyed our game nights. I would also like to thank Varun Rishi and Ashutosh Kumar for being supportive mentors and teaching me electronic structure theory from the ground up.

I have had many close friends from Kerala in Blacksburg who helped me have a good work-life balance in the last few years. I would like to thank Jopaul for being a good friend and making amazing authentic home food (writing this while sipping the chai he made), Manu for being a great friend and planning the next adventure,Ajit for the countless hours of chess and music discussions, Shweta for being always the life of the room, and Vaishakhi for being a very caring friend and for her positive outlook. I would also like to thank Shravan, Rutuja and Sasha for being good friends for being patient with extensive use of Malayalam around them. I would also like to thank Bijo, Nandita, Sachin, Satya, Ajai, Shobak and Shobal for being good friends during the past few years.

I would also like to thank Soumi for being my best friend, for staying by me for the past few years, and for being extremely patient with me. I also thank my undergraduate friends Ruhee, Santanu and Zeba, who are also in the USA, for going on exciting trips every year and helping me to step away from my work. I also thank Siddharth, cousin Tony and Cavya for being good friends from different parts of the globe.

I am indebted to my undergraduate advisor Prof. R. B. Sunoj and the supportive members of his lab, because of whom I decided to go to graduate school. I would also like to thank Prof. Daniel Crawford for hosting me as an undergraduate summer student and making VT feel so much like home that I decided to do my graduate studies here.

I would also like to thank my supportive lab-mates during my time in the Mayhall lab: Nicole for being a volleyball and gossip partner, Harper for the constant source of wit and being the dungeon master, Luke for his amazing cold brewDaniel for dragging me to the gym and a successful coup attempt, Robert for his animal stories, Ayush for being a kind roommate and positive attitude, and Oinam for helpful scientific feedback.

Finally, I would like to thank my advisor, Prof. Nicholas Mayhall, for his guidance, support and patience during these past 5 years. He has inspired me as a scientist, investing countless hours in discussion and providing constructive criticism throughout my graduate work. He has been very encouraging during my highs and extremely patient with my lows during my PhD journey. His scientific rigour and supportive nature are things I will miss in the future.

I also thank Shannon, Ayush, Soumi and Ruhee for proof reading parts of the abstract, introduction and conclusion.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With improvements in computing technology and theoretical methods, the use of quantum chemistry methods to study larger and complicated chemical systems is becoming more feasible. For the ground state properties of a molecule in closed shell, the calculations are easier, and one can use methods which assume the system is well-described by a single determinant. One the other hand, several challenges remain for computation of molecules that require a multi-determinant description of the state for example in bond breaking. This specific type of correlation is termed as 'static correlation' or 'strong correlation', and a multideterminantal wavefunction is necessary to describe this type of system.

Strong correlation is present in a wide variety of systems. As mentioned before, bond breaking involves strong correlation. Another class of system that has strong correlation is transition metal complexes with partially filled 3d orbitals. Transition metal complexes play a very important role as single molecule magnets for quantum computing,[7] optoelectronics[8] etc. Electronic excited states play an important role in applications in solar cells[9], aritificial photosynthesis[10], etc. These excited states in molecules are generally multideterminantal and hence considered strongly correlated. Finally, polyaromatic hydrocarbons (PAHs), which have been extensively used in the past decade to synthesize carbon nanosheets and nanobelts,[11, 12, 13] have delocalized π orbitals and exhibit strong correlation.[14] PAHs also have importance in several other fields like medicinal chemistry, single molecule wires, and so on. The study of these systems require computational methods that can accurately

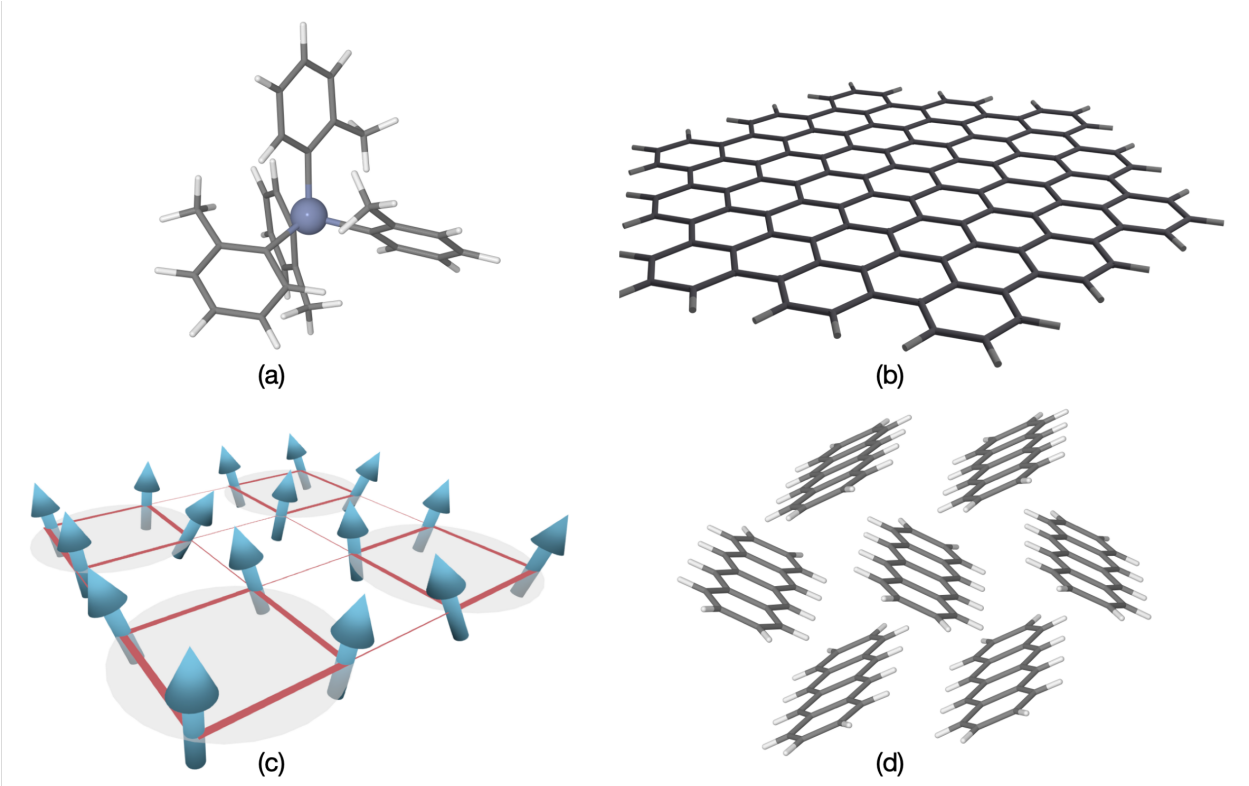


Figure 1.1: Some examples of strongly correlated systems. (a) A prototypical single molecular magnet sythesized by Bayliss et al. [1] (b) A carbon nanosheet with delocalized π orbitals (c) A two dimensional spin lattice. (d) A portion of a tetracene molecular system that undergoes singlet fission.

describe strong correlation. In this work, we present theoretical methods for treating strongly correlated systems.

1.1 Electronic Structure Theory

Quantum chemistry is the field of chemistry to study the motion of electrons and nuclei which most commonly involves solving the Schrödinger equation.[15] Modern electronic structure theory is a sub-field in quantum chemistry to study electronic degrees of freedom by solving the Schrödinger equation which gives a description of a system evolving with time.[16, 17]

From postulates of quantum mechanics, the time-dependent Schrödinger equation that describes a system of quantum particles can be written as:

$$i\hbar\frac{d}{dt}|\Psi(\mathbf{x},t)\rangle = \hat{H}|\Psi(\mathbf{x},t)\rangle \tag{1.1}$$

where $|\Psi(t)\rangle$ is the wavefunction of the system. The $\hat{H}$ is the quantum mechanical Hamiltonian operator associated with the energy of the system (E). When the Hamiltonian is not explicitly dependent on time (t), the wave function is separable and the above equation can be simplified to provide a time independent Schrödinger equation:

$$\hat{H}|\Psi(\mathbf{x})\rangle = E|\Psi(\mathbf{x})\rangle. \tag{1.2}$$

The Hamiltonian operator has a kinetic energy component as well as a potential energy component. The kinetic energy operator is similar to that in classical mechanics and is the derivative of the momentum operator. The potential energy in a molecular system arises due to Coulombic interactions between electron and nuclei. The Hamiltonian for a system with electrons and nuclei can be written as:

$$\begin{aligned}\hat{H} = &-\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2m_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_i-\mathbf{R}_A|} \\ &+\sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A-\mathbf{R}_B|}\end{aligned} \tag{1.3}$$

where m_A (Z_A) is the mass (charge) of the nuclei A, $\mathbf{R}_A$ ($\mathbf{R}_B$) is the coordinate of nuclei A (B), $\mathbf{r}_i$ ($\mathbf{r}_j$) is the coordinates for electron i (j). The first term in Equation 1.3 is the kinetic energy of electrons, the second one is the kinetic energy of nuclei, the third term is the potential energy between an electron and the nucleus. The fourth term is the electron-

electron repulsion and the fifth term is the inter nuclear repulsion.

Since the nuclei are much heavier than the electrons, the Born-Oppenheimer (BO) approximation is invoked where the nuclei are fixed with respect to the motion of electrons. We can neglect the kinetic energy of the nucleus and the potential energy of the nucleus becomes a constant. The electronic Hamiltonian becomes:

$$\hat{H}_{\text{el}} = \underbrace{-\frac{1}{2}\sum_i \nabla_i^2 - \sum_{iA} \frac{Z_{iA}}{|\mathbf{r}_i - \mathbf{R}_A|}}_{1-body} + \underbrace{\frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{2-body} + E_{\text{nuc}} \,. \tag{1.4}$$

The electronic wavefunction hence depends only parametrically on the nuclear degrees and can be hence written as

$$|\Psi_{\text{el}}\rangle = |\Psi_{\text{el}}\left(\{\mathbf{r}_i\}\,;\{\mathbf{R}_A\}\right)\rangle \,. \tag{1.5}$$

Electrons are fermions and hence they follow the anti-symmetry principle which means that the electronic many-body wavefunction must be anti-symmetric with respect to interchange of coordinates of any two electrons. The approximate electronic wavefunction for an N electron system can be written as a Slater determinant:[18]

$$|\Psi(\mathbf{x}_1, ...\mathbf{x}_N)\rangle = (N!)^{-1/2} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} . \tag{1.6}$$

A Slater determinant formed by a set of orthonormal spin orbitals are normalized and different N-electron determinants are orthogonal to each other.

The Hartree-Fock theory[19, 20] is the simplest approximate method to get the solution of

Schrödinger equation. In HF theory, for N electrons, we choose N orbitals from the pool of orbitals and minimise the energy variationally. The wavefunction for a Hartree-Fock procedure is a single Slater determinant which accounts for the anti-symmetric nature of electrons. The nonlinear differential equation from Hartee-Fock is usually solved using the Linear Combination of Atomic Orbitals (LCAO) approach introduced by Roothaan and Hall independently in 1951. [21, 22] Within the basis set limit, the Slater determinant formed from the optimized Hartree-Fock orbitals is the best variational approximation to the ground state of the system with a single determinant.

The energy not captured by Hartree-Fock method is termed as the correlation energy.

$$E^{corr} = E^{exact} - E^{HF}. \tag{1.7}$$

The correlation energy is much smaller compared to the HF energy, but is important for chemical accuracy (1 kcal/mol).

Second Quantization or occupation number representation uses field operators for the formulation of many-body quantum systems.[17] The wavefunction in the many-body Hilbert space can be manipulated using creation ($\hat{p}^\dagger$) and annihilation ($\hat{p}$) operators.

We start from orthonormal orbital basis which defines the Fock space and generate the occupation vector for a determinant using the creation operator and applying it to the fully vacuum space. The vacuum state can be represented as:

$$|\psi\rangle_0 = |0, 0, \ldots, \underbrace{0}_{p}, 0, \ldots, 0\rangle \tag{1.8}$$

Action of the creation operator on a vacuum state generates

$$|\psi\rangle_p = \hat{p}^\dagger|0\rangle = |0, 0\ldots, \underbrace{1}_{p}, 0, \ldots, 0\rangle \tag{1.9}$$

Similarly action of the annihilation operator on a state vector removes the electron from that orbital

$$\begin{aligned} |\psi\rangle_{pmn} &= |0, \ldots, \underbrace{1}_{p}, \ldots \underbrace{1}_{m} \ldots, \underbrace{1}_{n}, 0\rangle \\ |\psi\rangle_{mn} = \hat{p}|\psi\rangle_{pmn} &= |0, \ldots, \underbrace{0}_{p}, \ldots \underbrace{1}_{m} \ldots, \underbrace{1}_{n}, 0\rangle \end{aligned} \tag{1.10}$$

The anti-commutator relations for these operators can be summarized as:

$$\left[\hat{p}^\dagger, \hat{q}^\dagger\right]_+ = 0; \quad \left[\hat{p}, \hat{q}\right]_+ = 0; \quad \left[\hat{p}, \hat{q}^\dagger\right]_+ = \delta_{pq}. \tag{1.11}$$

The electronic Hamiltonian in the second quantized framework is:

$$\hat{H} = \sum_{pq} h_{pq}\hat{p}^\dagger\hat{q} + \tfrac{1}{2}\sum_{pqrs} \langle pq||rs\rangle\, \hat{p}^\dagger\hat{q}^\dagger\hat{s}\hat{r} \tag{1.12}$$

The one-electron part (h_{pq}) is:

$$h_{pq} = \langle p|h|q\rangle = \int d\mathbf{x}_1 \chi_p^*(\mathbf{x}_1)\, h(\mathbf{r}_1)\, \chi_q(\mathbf{x}_1) \tag{1.13}$$

The anti-symmetrized two-electron integral ($\langle pq||rs\rangle$)is:

$$\langle pq||rs\rangle = \langle pq|rs\rangle - \langle pq|sr\rangle = \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_p^*(\mathbf{x}_1)\, \chi_q^*(\mathbf{x}_2)\, r_{12}^{-1}\, (1 - \mathcal{P}_{12})\, \chi_r(\mathbf{x}_1)\, \chi_s(\mathbf{x}_2) \tag{1.14}$$

where $\langle pq|rs\rangle$ is the Coulombic interaction term and the $\langle pq|sr\rangle$ is the exchange term. The exchange term arise due to the anti-symmetric nature of the fermionic wavefunction.

1.2 Correlated Wavefunctions

Post Hartree-Fock methods are employed to capture the correlation energy. A full configuration interaction calculation is required to capture the complete correlation energy.[23, 24] The full CI wavefunction in occupation number representation is expressed as

$$|\Psi\rangle = \sum_{i_1,i_2,i_3\ldots i_n} \mathcal{C}_{i_1 i_2 i_3 \ldots i_n} |i_1 i_2 i_3 .. i_n\rangle \tag{1.15}$$

where $\mathcal{C}$ is the n dimensional coefficient tensor.

Unfortunately full CI scales combinatorially with system size and is not computationally feasible to be computed for molecules with more than a few atoms. Hence approximate methods are invoked to obtain the correlation energy. The correlation energy can be roughly classified into dynamic and non-dynamic (static) correlation.

Static correlation arises from degenerate or near-degenerate configurations that mix strongly. Meanwhile dynamical correlation accounts for the short-range electron-electron interaction and weak long-range interactions. Traditional single reference methods provides very accurate results for dynamic correlation even though large basis sets need to be used. The Hartree-Fock wavefunction is still a good approximate wavefunction and can be used as a reference for correlated methods. In case of static correlation, which is usually seen in cases of bond breaking, degenerate orbitals in transition metal complexes, excited states etc, the wavefunction shows multideterminantal characteristics, where more than a single determinant becomes important. In such cases, one resorts to active space based approaches where

the orbital basis is split into doubly occupied, virtual (empty) and active orbitals.

1.3 Concept of an active space

Active space based approaches are key for describing strongly correlated systems. Active spaces are generally formed from a set of important occupied and virtual orbitals from a single reference calculation. When we perform a FCI calculation within the active space, this is referred to as complete active space configuration interaction (CASCI). In the complete active-space self-consistent field (CASSCF[25]) method, the orbitals and CI coefficients are optimized. For an active space based method, the correlation outside the active space can be captured using CASPT2[26, 27], MRCI.[28, 29] Even after forming an active space, the computational complexity of the problem is still combinatorially since we do FCI within the active space. Hence application of this approach to larger active spaces is not feasible. Along this route there are multiple approaches that can be used to solve a large active space reaching the exact CASCI limit.

1.4 Solving large strongly correlated active space

Carrying out an exact CASCI calculation is an exponentially scaling problem with system size, hence we need approximate methods keeping in mind that we cannot truncate the excitation rank similar to traditional truncated coupled cluster methods since higher excited determinants can also be important. We mainly focus on selected CI, incremental CI and spin-flip based methods in this dissertation. There are other methods like density matrix renormalization group (DMRG)[30, 31], externally corrected coupled-cluster methods[32, 33, 34] and methods based on quantum monte carlo-like full configuration interaction quantum

monte carlo (FCIQMC)[35, 36] and auxillary field quntum monte carlo (AFQMC)[37] that provide promising results for strongly correlated systems. Most of these methods still scale exponentially with a smaller prefactor than FCI. U sually for half filling, accurate results can be performed for up to 60 electrons in 60 orbitals with efficient computer implementations.

1.4.1 Selected CI

In SCI approaches the wavefunction is iteratively generated by selecting configurations or determinants important for the wavefunction.[38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48] There are different criteria based on which this selection can be carried out. In the earliest selected CI approach (CIPSI), the selection is based on the magnitude of the first order wavefunction coefficients.[38] For the heat bath criteria, the selection is based on the Hamiltonian element between the configurations.[41, 49] There are other selection criteria, but irrespective of the selection criteria, the current limitation of the strongly correlated system with selected CI approach is the fast growing wavefunction dimension. It has been seen that using a natural orbital basis improves the sparsity of the wavefunction.[41, 47]

1.4.2 Incremental FCI

Another approach is to use an incremental approach to capture the correlation energy. This approach was proposed by Nesbet in the 1960s by using a many-body expansion (MBE) to capture the correlation energy using the n-th order Bethe-Goldstone equation.[50, 51, 52] This MBE approach has recently seen improvements from the Zimmerman group[53, 54, 55] as well as Eriksen and Gauss independently.[56, 57, 58, 59] Efficient pruning techniques and improved reference has helped the application of the incremental method for complicated strongly correlated systems and have been reaching FCI limit.

1.4.3 Spin-flip Approaches

Spin-flip based approaches are single-reference electronic structure methods designed for describing strongly correlated systems.[60] The spin-flip wavefunction is generated by the action of a spin-flipping excitation operator on a single high-spin reference determinant. This method provides a single well defined reference orbitals and also leads to automatic selection of active space based on the number of spin-flips. The correlation on top of spin-flip approach can be achieved using restricted active space (RAS) approach. [61, 62, 63] In this work, the last two chapters use spin-flip to study a complicated doubly excited state in singlet fission process. One of the main issues with the spin-flip approach is capturing dynamic correlation after the strong correlation part. Since we start from a very high energy excited spin configuration, the orbitals also are sometimes not good for these calculations. The equation of motion coupled-cluster (EOM-CC) based spin-flip approaches have shown improved results.[64] Recently, a single spin-flip method was proposed that can study large active spaces by extracting spin Hamiltonian parameters from the 1SF wavefunction.[65, 66]

1.5 Qualitative understanding of Tensor Product States

In the determinant basis, the FCI vector even though sparse[67] can be huge for large strongly correlated systems. Using selected CI, the largest variational dimension possible are in the order of 10^{10} determinants. For the graphene nanosheet system shown in Figure 1.1(b), the FCI vector has 10^{66} configurations. An SCI vector with 10^{10}, would often be a very small part of the full wavefunction. In this work, we discuss a new wavefunction representation by partitioning the orbital system into separate clusters. This partitioning is done such a way that the orbitals in a cluster are strongly interacting with each other. We form many-body basis in each clusters using the exact solution in each cluster and the final wavefunction is

formed in the basis of the tensor product of these cluster states.

Analyzing the systems in Figure 1.1, we can form a clustering for each of these systems. For the transition metal complex in Figure 1.1(a), we can partition the system such that the ligands are separate clusters and the metal d orbitals are another cluster. The bonding between the ligand and the metal can be another cluster. The graphene nanosheet presented in Figure 1.1(b) can be partitioned using Clar's rule and we investigate this molecule in Chapter 3. We investigate the spin lattice (Figure 1.1(c)) in Chapter 6. Finally we investigate the singlet fission material in Chapter 5 and Chapter 7.

The TPS approach is most applicable if the system can be partitioned into clusters with strong interaction within them and weak interaction between them. One of the main disadvantage of the TPS method is that we can get different results with different clustering and an optimal clustering is still an open questions. Also the TPS based approaches are not a black box approach where the user needs to have more information about the system to partition efficiently. The clustering can be automated to some extend using spectral clustering or using a Cuthill–McKee algorithm[68] as we do for diatomic systems in Chapter 4.

1.6 Summary

In the first part of this dissertation, we investigate the application of tensor product states for strongly correlated systems using various approximations from Chapter 3 to Chapter 6. In the second part, we investigate a particular type of strongly correlated states, correlated triplet pair states in singlet fission in Chapter 7 and Chapter 8. In Chapter 2, we give a brief overview of tensor product states, how to form matrix elements in this basis and discuss the mean field solution for the TPS method. In Chapter 3, we investigate the applicability of an incremental approach for correlation energy using TPS and apply to

large systems like graphene nanosheets (Figure 1.1(b)) In Chapter 4, we use a selected configuration method (SCI) along with the TPS wavefunction termed tensor product selected CI (TPSCI). In Chapter 5, we investigate the application of TPSCI method for excited states. In Chapter 6, we formulate a size-extensive version of the n-body Tucker approach using perturbation theory and linearized coupled-cluster approach and apply it to strongly correlated spin lattices (Figure 1.1(c)) In Chapter 7, we investigate the singlet fission process using simple parameters derived from a single 1 spin-flip ab initio calculation In Chapter 8, we provide simple rule using the Ising spin model to predict the multiexciton spin state ordering based on bridging molecule. Finally we present the conclusion of the dissertation in Chapter 9.

Chapter 2

Theory

Part of this Chapter is reproduced from Abraham, V.; Mayhall, N. J. Cluster many-body expansion: A many-body expansion of the electron correlation energy about a cluster mean field reference. *The Journal of Chemical Physics* **2021**,*155*, 054101, with the permission of AIP Publishing.

Part of this Chapter is adapted with permission from Abraham, V.; Mayhall, N. J. Selected Configuration Interaction in a Basis of Clus- ter State Tensor Products. *Journal of Chemical Theory and Computation* **2020**, *16*, 6098–6113. Copyright 2020 American Chemical Society.

2.1 Tensor Product State

We start by partitioning the spatial orbitals into disjoint sets, which we refer to as "clusters". Within each cluster, N, we define a set of cluster states, $|\alpha\rangle_N$, each of which is a linear combination of all possible Slater determinants involving a cluster's orbitals. In order to simplify the notation, we use lower case Roman characters to enumerate orbitals $(p, q, ...)$, upper case Roman to enumerate clusters $(I, J, ...)$, and Greek letters to enumerate local many-body cluster states $(\alpha, \beta, ...)$. A global tensor product state (TPS) over the full system can be represented using these cluster states as:

$$|\phi\rangle = |\alpha\beta\ldots\gamma\rangle = |\alpha\rangle_1 |\beta\rangle_2 ... |\gamma\rangle_N \tag{2.1}$$

By taking all possible tensor products of local cluster states (involving all sectors of a cluster's Fock space), we exactly span the original Hilbert space. Thus, the exact full CI wavefunction can be represented in this TPS basis as

$$|\Psi\rangle = \sum_{\alpha} \sum_{\beta} ... \sum_{\gamma} c_{\alpha,\beta,...,\gamma} |\alpha\beta \ldots \gamma\rangle \tag{2.2}$$

Here $c_{\alpha,\beta,...,\gamma}$ is the expansion coefficient in front of the corresponding TPS configuration.

There is freedom in how the orbitals are organized into clusters. The orbital clustering can be chosen based on orbital locality, symmetry of the system, or any other criterion. While the choice of this clustering is up to the user, the guiding principle is that the Hamiltonian should act more strongly within clusters and more weakly between them. If a reasonable clustering can be found, an accurate approximation to the FCI state can be represented using few tensor product states as compared to the full space.

For the ground state of a given system, the tensor product state formed as the direct product of the lowest energy state of each cluster can be used as an initial approximation. In other words, the solution for the global ground state can be approximated as

$$|\Psi_0\rangle = |0\rangle_1 |0\rangle_2 ... |0\rangle_n \tag{2.3}$$

However, this approximation can be quite drastic since the cluster states are not influenced by neighboring clusters, i.e., each state is the gas-phase ground state of the cluster. Alternatively, one could choose cluster states which rigorously minimize the energy of the TPS approximation, $|\Psi_0\rangle$. This approach is known as the cluster-based mean field (cMF) approximation.[69] In that work, Jimenez-Hoyos and Scuseria further minimize the energy subject to inter-cluster orbital rotations. The resulting cMF wavefunction, then has an optimal set of cluster states and orbitals (for representing a single TPS). This technique has also been

recently applied to ab initio molecular systems by Hermes and Gagliardi.[70, 71] While this significantly improves the reference wavefunction, only the ground state of each cluster is defined by the variational condition. The cMF energy is invariant to rotations of the excited cluster states. Analogous to the definition of canonical molecular orbitals in HF theory, the local cluster's excited state could be defined as simply the higher energy eigenvectors of the "one-cluster reduced Hamiltonian", which is the cMF analogue of the Fock matrix. Both being eigenfunctions of an effective mean-field operator, the cluster state energies in cMF are analogous to the orbital energies in HF theory. As such, one can form perturbative expansions about each mean-field operator, as is done in traditional MP2 theory and in the PT2 correction developed for the cMF work.[69]

2.2 Matrix Elements

In order to optimize the expansion coefficients of the TPS basis vectors described above, one needs to evaluate the Hamiltonian matrix elements between arbitrary TPS configurations. Although the matrix elements in a traditional determinant-based CI code are straightforward to evaluate, the TPS matrix elements are significantly more involved.

To start, we first partition the second quantized Hamiltonian,

$$\hat{H} = \sum_{pq} h_{pq} \hat{p}^\dagger \hat{q} + \tfrac{1}{2} \sum_{pqrs} \langle pq||rs\rangle \, \hat{p}^\dagger \hat{q}^\dagger \hat{s} \hat{r} \tag{2.4}$$

into distinct operators which are labelled by the clusters upon which they act.

$$\hat{H} = \sum_{I} \hat{H}_I + \sum_{I<J} \hat{H}_{IJ} + \sum_{I<J<K} \hat{H}_{IJK} + \sum_{I<J<K<L} \hat{H}_{IJKL} \tag{2.5}$$

For instance, $\hat{H}_I$ corresponds to the operators that are local to cluster, I. The two-body term $\hat{H}_{IJ}$ involves all Hamiltonian operators such that the operator indices occur in clusters I and J. Interactions such as 2-body charge-transfer, exchange, and dispersion fall within this set. The details of implementing the one-body and two-body terms have been worked out by Shiozaki and coworkers as part of the ASD method.[72, 73] However, unlike in ASD which is defined for two clusters, we assume an arbitrary number of clusters. As a result, we must also handle the 3-body and 4-body Hamiltonian terms explicitly.

Due to the antisymmetric nature of fermions, many of the above "local" terms require a non-local treatment. For instance, to act a creation operator on cluster 3, it must first anticommute through the first two clusters. While a general algorithm can be defined easily when using the Jordan-Wigner spin mapping (and this was the approach we took for an initial proof of principle code), this approach incurs significant overhead, and prevents efficient vectorization, due to the need to account for operator commutation with Pauli Z matrices. To avoid this, and to speed up the matrix element construction, we make the restriction that each cluster state has well-defined particle number and $\hat{S}_z$. This keeps the Pauli Z strings diagonal, allowing us to simply precompute the anticommiutation sign before doing any floating point operations. In other words, we don't allow mixing between the local cluster states in different sectors of Fock space. This also has the added benefit of extra Hamiltonian sparsity and trivial quantum number preservation of the global state (i.e., ensuring that the final state has the correct number of electrons).

To provide a concrete example, let us consider one contribution to the two-body matrix

element:

$$\hat{H}_{IJ} \Leftarrow \sum_{pqr\in I}\sum_{s\in J} \langle pq||rs\rangle\, \hat{p}^\dagger\hat{q}^\dagger\hat{s}\hat{r} \tag{2.6}$$

$$= -\sum_{pqr\in I}\sum_{s\in J} \langle pq||rs\rangle \left\{\hat{p}^\dagger\hat{q}^\dagger\hat{r}\right\}\left\{\hat{s}\right\} \tag{2.7}$$

between two arbitrary TPS configurations:

$$|\psi\rangle = |\alpha\rangle_1 \ldots |\beta\rangle_I \ldots |\gamma\rangle_J \ldots |\delta\rangle_N \tag{2.8}$$

$$|\psi'\rangle = |\alpha'\rangle_1 \ldots |\beta'\rangle_I \ldots |\gamma'\rangle_J \ldots |\delta'\rangle_N \tag{2.9}$$

$$\tag{2.10}$$

Here, we first move each group of operators to the associated cluster it acts on, keeping track of any signs.

$$\sum_{pqr}^{I}\sum_{s}^{J} \langle pq||rs\rangle \left\{\hat{p}^\dagger\hat{q}^\dagger\hat{r}\right\}\left\{\hat{s}\right\} |\alpha\rangle_1 \ldots |\beta\rangle_I \ldots |\gamma\rangle_J \ldots$$

$$= (-1)^\chi \sum_{pqr}^{I}\sum_{s}^{J} \langle pq||rs\rangle\, |\alpha\rangle_1 \ldots \hat{p}^\dagger\hat{q}^\dagger\hat{r}\, |\beta\rangle_I \ldots \hat{s}\, |\gamma\rangle_J \ldots \tag{2.11}$$

where $\chi = \sum_{K=I}^{J-1} N_K$, meaning we just sum the number of electrons in each state on clusters between the two active clusters. Next applying the bra to get the matrix element, we are left with:

$$\langle\psi'|\hat{H}_{IJ}|\psi\rangle \Leftarrow -(-1)^\chi \sum_{pqr\in I}\sum_{s\in J} \langle pq||rs\rangle\, \Gamma^{\beta'\beta}_{pqr}\Gamma^{\gamma'\gamma}_{s}$$

$$\times \prod_{K\neq I,J} \langle\omega'|\omega\rangle_K \tag{2.12}$$

where ω (ω') is the local state occupied in $|\psi\rangle$ ($|\psi'\rangle$), and the operator tensor,

$$\Gamma^{\beta'\beta}_{pqr} = \langle\beta'|\hat{p}^\dagger\hat{q}^\dagger\hat{r}|\beta\rangle_I \tag{2.13}$$

is a quantity local to cluster I, and

$$\Gamma^{\gamma'\gamma}_{s} = \langle\gamma'|\hat{s}|\gamma\rangle_J \tag{2.14}$$

is a quantity local to cluster J. In the above, all terms are zero unless the associated cluster states are from compatible cluster Hilbert spaces. For example, $|\beta'\rangle_I$ must have one more electron than $|\beta\rangle_I$. These are similar to transition density matrices, and are precomputed and accessed from memory when needed in the above expressions. Due to the orthonormality of the cluster states, for each cluster $K \neq \{I, J\}$, the overlap evaluates to a Kronecker delta indicating that the matrix element between any two TPS's is zero unless all of the non-active cluster states are in the same state. This is the TPS analogy to Slater-Condon rules, and significantly reduces the number of terms we must compute. Similar manipulations are required for all the various 1, 2, 3, and 4 body terms.

2.3 Cluster Mean Field

The cMF method, originally proposed by Jiménez-Hoyoz and Scusceria, is an ideal reference for any TPS based method, and has shown promising results for the 1D and 2D Hubbard systems.[74] In this approach the active space is partitioned into separate clusters or blocks and many-electron states are formed within these clusters. The full system wave function

can then be represented using a tensor product of the local cluster states,

$$|\psi\rangle = \sum_{\alpha,\beta,\gamma..\omega} c_{\alpha_I,\beta_J,\gamma_K..\omega_N} |\alpha_I, \beta_J, \gamma_K, ..\omega_N\rangle \tag{2.15}$$

where $c_{\alpha_I,\beta_J,\gamma_K..\omega_N}$ corresponds to the coefficient for a given state in the tensor product basis. Here we use upper case letters for representing blocks or clusters of orbitals, and Greek letters to represent the many electron cluster-states.

The ground state of the full system can be approximated by taking the lowest many-body cluster-state as

$$|\psi_0\rangle = |0_I, 0_J, 0_K, ..0_N\rangle \tag{2.16}$$

where $|0_L\rangle$ is the lowest energy cluster state for cluster L. We can write the cluster state $|0\rangle_L$ as linear combination of determinants in the cluster L.

$$|0_L\rangle = \sum_l x^L_{l,0} \left|D^l_L\right\rangle \tag{2.17}$$

where l is the determinant index in the cluster state basis.

The Hamiltonian in the clustered form can be represented as,

$$\hat{H} = \sum_I \hat{H}_I + \sum_{I<J} \hat{H}_{IJ} + \sum_{I<J<K} \hat{H}_{IJK} + \sum_{I<J<K<L} \hat{H}_{IJKL} \tag{2.18}$$

where $\hat{H}_I$, $\hat{H}_{IJ}$, $\hat{H}_{IJK}$ and $\hat{H}_{IJKL}$ correspond to Hamiltonian terms with one, two, three and four cluster interactions, respectively.

Analogous to Hartree-Fock,[16] we seek a self-consistent optimization of the cluster states $|0_L\rangle$ such that the reference TPS is variationally minimized. The Lagrangian under the

constraint that the reference state is normalized can be written as

$$\mathcal{L} = \langle\psi_0|\hat{H}|\psi_0\rangle + \epsilon\Big(\langle\psi_0|\psi_0\rangle - 1\Big) \tag{2.19}$$

Differentiating this Lagrangian with respect to the cluster basis coefficients for a given cluster state $|0_L\rangle$,

$$\frac{\partial\mathcal{L}}{\partial\langle 0_L|} = 0, \tag{2.20}$$

and substituting $|\psi_0\rangle$ and the Hamiltonian in clustered form in Equation 2.20 yields:

$$\frac{\partial}{\partial\langle 0_L|}\Bigg(\sum_{M\neq L}\langle 0_L, 0_M|\hat{H}_L + \hat{H}_M + \hat{H}_{LM}|0_L, 0_M\rangle - \epsilon(\langle 0_L|0_L\rangle - 1)\Bigg) = 0. \tag{2.21}$$

In Equation 2.21, since the differentiation is with respect to cluster L, only Hamiltonian terms that have contributions from cluster L are needed. Because Hamiltonian terms with three or four body terms (H_{IJK}, H_{IJKL}) will necessarily have an odd number of creation/annihilation operators on at least one cluster, they do not contribute to the cMF energy. This can be easily demonstrated using an example:

$$\hat{H}_{IJK} \Leftarrow \sum_{pr\in I}\sum_{q\in J}\sum_{s\in K}\langle pq||rs\rangle\,\hat{p}^\dagger\hat{q}^\dagger\hat{s}\hat{r} \tag{2.22}$$

$$= \sum_{pr\in I}\sum_{q\in J}\sum_{s\in K}\langle pq||rs\rangle\left\{\hat{p}^\dagger\hat{r}\right\}\left\{\hat{q}^\dagger\right\}\left\{\hat{s}\right\} \tag{2.23}$$

Action of this term on the reference would produce a new TPS with a new electron config-

uration where cluster I will have same electrons as before, but cluster J will have an extra electron because of the $\hat{q}^\dagger$ term and cluster K will have one less electron because of the $\hat{s}$ term. Hence, three and four body terms do not contribute to the energy evaluation at the cMF step.

After the differentiation and collecting terms, we have:

$$\left(\hat{H}_L + \sum_M \hat{V}_{L[M]} + E_M\right)|0_L\rangle - \epsilon_L|0_L\rangle = 0, \tag{2.24}$$

where $\hat{V}_{L[M]} = \langle 0_M|\hat{H}_{LM}|0_M\rangle$ is the potential from the cluster M and $E_M = \langle 0_M|\hat{H}_M|0_M\rangle$. Equation 2.24 is an eigenvalue problem where ϵ_L corresponds to the cluster state energy similar to the orbital energies in HF. For the fermionic Hamiltonian, $\hat{V}_{L[M]}$ can be represented as

$$\hat{V}_{L[M]} = \sum_{pr\in L} \hat{p}^\dagger \hat{r} \sum_{qs\in M} \langle pq||rs\rangle\, \rho^M_{qs} \tag{2.25}$$

where $\rho^M_{qs} = \langle 0_M|\hat{q}^\dagger \hat{s}|0_M\rangle$ is a one-particle density matrix for cluster M.

Because the effective potential in cluster L has contributions from each cluster M through its one-particle density matrix (ρ^M), we must solve for the cMF state self consistently by updating the effective potential iteratively until convergence.

The Hamiltonian from Equation 2.24 can be understood as the many-electron Fock-like operator for the cMF procedure,

$$\hat{H}^0 = \sum_I \hat{H}_I + \sum_{I,J} \hat{V}_{I[J]}. \tag{2.26}$$

Because these equations arise from a variational minimization of a well-defined energy functional, we can easily improve the ansatz by minimizing with respect to the orbitals as well as

the cluster state CI coefficients. As previously demonstrated,[74, 75] the orbital optimization is a key step that can improve the energy significantly. Another advantage to optimizing orbitals is that it removes ambiguities surrounding orbital localization procedures, and it also makes extensions to properties and gradients much simpler.

For a given wavefunction $|\Psi_0\rangle$, we can define the unitary transformation in the single particle basis which minimizes the energy using an anti-Hermitian matrix $\hat{\kappa}$.

$$\hat{\kappa} = \sum_{p<q} \kappa_{pq}(p^\dagger q - q^\dagger p) \tag{2.27}$$

where κ_{pq} are the orbital rotation parameters.

The single particle basis gets transformed into a new basis.

$$\tilde{\hat{p}} = e^{\hat{\kappa}} \hat{p} e^{-\hat{\kappa}}, \tag{2.28}$$

such that the energy now carries an orbital dependence.

$$E[\kappa] = \left\langle \Psi_0 | e^{-\hat{\kappa}} \hat{H} e^{\hat{\kappa}} | \Psi_0 \right\rangle \tag{2.29}$$

The orbitals are optimized when the orbital rotation gradient goes to zero. Hence the orbital gradient at each step is formed and a conjugate gradient or BFGS algorithm can be used to optimize the orbitals. The gradient can be expressed in terms of a generalized Fock matrix similar to traditional quantum chemistry methods.[17, 76]

$$G_{pq} = 2(F_{pq} - F_{qp}) \tag{2.30}$$

where F_{pq} is the generalized Fock matrix which can be formed using the one- and two-particle

density matrices of the cMF reference,

$$F_{pq} = D_{pr} h_{qr} + \Gamma_{prst} g_{qrst}. \tag{2.31}$$

Therefore, cMF with orbital optimization is identical to a CASSCF with multiple active spaces. The orbital optimization can be accelerated by forming the orbital Hessian as well,[17, 74] but we do not take this approach in the current study.

Chapter 3

Cluster Many-body Expansion

Reproduced from Abraham, V.; Mayhall, N. J. Cluster many-body expansion: A many-body expansion of the electron correlation energy about a cluster mean field reference. *The Journal of Chemical Physics* **2021**,*155*, 054101 [**Editor's Pick**], with the permission of AIP Publishing.

3.1 Introduction

Modern electronic structure methods are usually based on the Hartree-Fock reference.[16, 17] Although most of the energy is already accounted for by this reference, the missing energy, or correlation energy, is necessary in order to obtain accurate and meaningful results. A full configuration interaction (FCI)[23, 24] calculation is required for the exact correlation energy but is unfeasible when the system size is large due to its exponential scaling. Less expensive single reference electronic structure methods like density functional theory (DFT)[77] or truncated coupled cluster (CC)[78] can be used to capture part of this correlation and for most ground state properties. However, if the system has an ill-defined reference determinant, these methods tend to fail since they are dependent on the reference. This type of correlation is broadly referred to as strong or static correlation and usually arises in transition metal complexes, excited states, and bond breaking. In these cases there is usually orbital near-degeneracy and contributions from more than one determinant becomes im-

portant. One usually resorts to active space based methods for such cases, but even these approaches are plagued by the exponential scaling of the wavefunction with respect to system size. Improved computational resources and approximations have allowed application of accurate wavefunction-based quantum mechanical methods to many challenging strongly correlated systems in recent years.[79, 80, 81]

The FCI wavefunction is extremely sparse, and there are different approximate methods that can be used to exploit this sparsity.[67] Selected configuration interaction (SCI) exploits this idea and approximates the wavefunction by selecting important configurations. The first selected CI algorithm was proposed by Malrieu and coworkers in the 1970's.[38] Other selected CI methods include recent improvements to the CIPSI algorithm, [39, 40] semi-stochastic heat bath CI,[41, 42] adaptive CI, [43] coordinate descent FCI, [82] iterative CI, [45, 46] adaptive sampling CI [47] and MCCI [48]. A stochastic variant of the selected CI method, the full configuration interaction quantum Monte Carlo (FCIQMC) method, samples the determinant basis by assigning signed walkers.[35, 36] FCIQMC has also been used to develop the semi-stochastic CAD-FCIQMC where the higher excitations of FCIQMC are used with a CC formalism similarly to externally corrected coupled cluster methods.[32, 34] A selected coupled cluster method, full coupled cluster reduction (FCCR), has also been proposed and has shown very accurate results with a PT correction.[83]

Another set of approaches used for solving strongly correlated large active spaces are tensor network based methods. Density matrix renormalization group (DMRG), [30, 84] initially designed for the exact solution of 1D spin lattices, has shown impressive results for chemical systems.[31, 85, 86] DMRG is mostly applicable for pseudo-one-dimensional systems. There are also a few higher-dimensional tensor network based methods such as tree tensor network states (TTNS), [87] complete graph tensor network states (CGTNS)[88] and so on.

Similar to the approach that we will discuss below, there are also approaches in which the

active space is partitioned into orbital groups and then the system is solved by restricting the excitations between those groups. Occupation restricted multiple active space (ORMAS), [89, 90] restricted active space (RAS)[91] and generalized active space (GAS) [92, 93] can all be conceptualized in this way.

Nesbet in the 1960's proposed to use a many-body expansion (MBE) to capture the correlation energy using the n-th order Bethe-Goldstone equation.[50, 51, 52] The MBE and its variants[94, 95, 96, 97, 98, 99, 100, 101, 102] are versatile tools used in traditional chemistry applications like predicting binding energies, [103, 104] crystal lattice energies and structures,[105, 106, 107, 108, 109, 110] dipole moment and polarizability,[111, 112] vibrational frequencies,[113, 114, 115] forces[116, 117, 118, 119] and excited state energies.[100, 120, 121] Even though MBE has been used in these contexts, its ability to solve for the correlation energy of large systems was not widely exploited until recently. In recent years, with increased computing power and smart pruning, use of the MBE method to approximate FCI energy has seen new interest. One of the earliest methods where the MBE was used is the method of increment (MoI) approach by Stoll where orbital blocks were used as n-body entities.[122, 123, 124, 125] Paulus and coworkers have further used the MoI with localized orbitals to study a variety of systems and have also proposed a multi-reference version for bond breaking problems.[126, 127, 128] Ruedenberg and co-workers proposed the correlation energy extrapolated many-body expansion where they combined the correlation energy extrapolation by intrinsic scaling method with the many-body expansion using local orbitals.[129, 130] The incremental FCI (iFCI) method by Zimmerman *et al.*[53, 55] used SHCI [41] as a solver for higher order calculations and has also been extended to do orbital optimization.[131, 132] Eriksen and Gauss proposed the many-body expansion full configuration interaction (MBE-FCI) method by expanding over virtual orbitals[56, 57, 59]. A generalized MBE-FCI was later proposed [58] and has been extended to excited states.[133] Recently

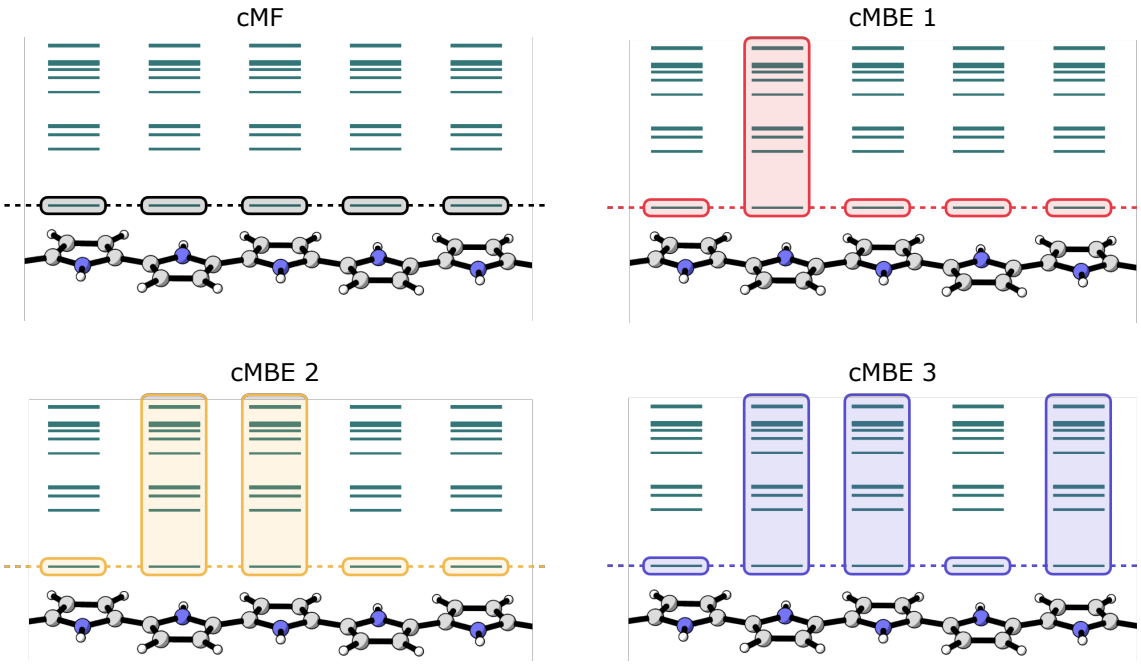


Figure 3.1: Pictorial depiction of the reference cMF state and example terms from a given subsequent cMBE expansion for the polypyrrole molecule. The green lines correspond to the cluster state energies of each cluster (here each pyrrole unit is considered as a cluster). The cMF reference as shown is a single TPS formed by the direct product of the lowest energy cluster states. The subsequent many-body expansion can be understood as including the degrees of freedom for the active clusters. We show the example terms for cMBE1, cMBE2 and cMBE3.

the incremental approach has also been used with frozen natural orbitals for reducing the dimension of the virtual space dimension at each order.[134]

As the degree of strong correlation increases, higher body corrections need to be incorporated to get exact results since the HF reference is not a reliable guess.[55, 58] The perfect pairing reference,[135] instead of a Hartree-Fock reference has shown improved results for strongly correlated systems[55] but generalizing it to cases where there are more than two orbitals in a block is challenging. Another drawback of the traditional MBE approaches expanded over an orbital basis is that as the size of the system increases, the calculation of each order gets more expensive with increased number of virtual and occupied orbitals.

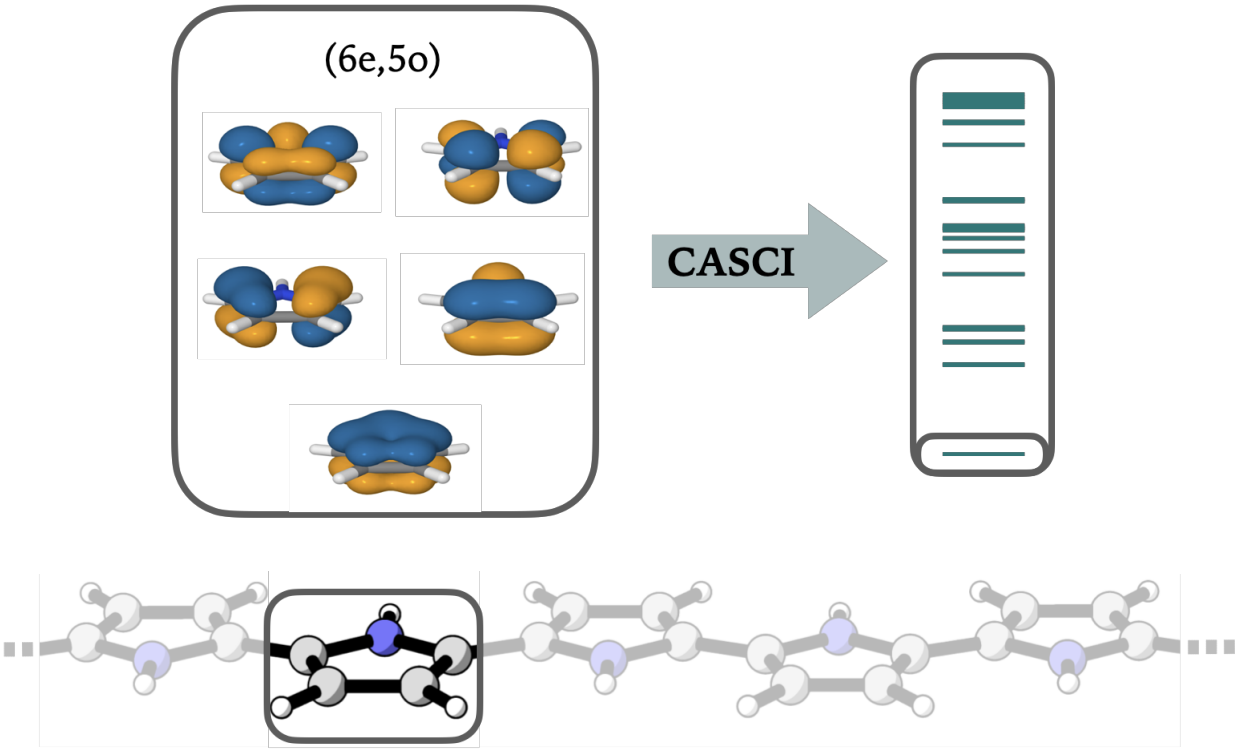


Figure 3.2: The orbitals within a given cluster are plotted on the right. Solving the exact CASCI problem inside the cluster gives the cluster states in the left panel. The lowest energy cluster state is optimized during cMF.

In this paper, we propose the use of the tensor product state (TPS) basis as an alternative to the traditional Slater determinant basis in the many-body expansion. The basic idea is to partition the system into separate clusters, solve the smaller many electron problem within each cluster, and then represent the wavefunction as a tensor product of these cluster states.[75] Using this alternate wavefunction expansion has multiple advantages when the system is *"clusterable"*. A system is said to be *"clusterable"* when the Hamiltonian has a structure to it which can be exploited to partition the orbital space into different orthonormal orbital clusters. Even though orbital locality is the most sensible clustering criteria, sometimes other factors like symmetry and even bonding and anti-bonding orbital pairings can also lead to efficient clustering criteria.[75] Most of the large molecular systems of interest like crystal lattices or polymetallic complexes have an intrinsic structure which can be used to partition the orbitals into separate groups. Recently Jiménez-Hoyos and Scuseria proposed the cluster mean field (cMF) method for fermions, where a single TPS configuration is variationally minimized.[74] cMF defines a reference TPS configuration, like Hartree-Fock is the reference determinant for Slater determinant based methods. There are different methods that use the basic structure of the TPS basis and add correlation to it using perturbation theory (PT) [74], configuration interaction [75, 136], coupled cluster [137, 138, 139], tensor networks [72, 73] and effective Hamiltonians.[140, 141, 142, 143]

In the approach described in this paper, the cluster many-body expansion (cMBE), the increments or the building blocks are the clusters themselves and not the virtual or occupied orbitals. The definitions of the reference state and expansion orders change when using many-electron cluster states as the basis. This tensor product state and corresponding orbitals are optimized using cMF.[74] Hence in this new representation, the correlation component within each cluster is exactly captured. Since the exact solution of the cluster basis is used, the size of the clusters is restricted, but with selected CI methods this could also be alleviated. Using

the cMBE method, we propose to both surpass the computational challenge and improve the convergence of the MBE by exploiting the structure of the system.

Some of the advantages of using a mean field TPS reference are:

1. The cMF reference with orbital optimization provides a better reference than Hartree-Fock since it incorporates more correlation.

2. Faster convergence means fewer terms to compute, which avoids numerical precision issues.

3. The tensor product reference allows one to exploit localized structure in the molecule for increased efficiency.

4. Although less of a black-box approach, a cluster based method can provide a more intuitive framework for chemists to analyse and interpret *ab intio* results.

Even with these advantages, the tensor product based methods do not entirely remove the issue of dimensionality. In the tensor product basis, calculating the matrix elements for the Hamiltonian is more expensive than just using the Slater-Condon rules. One of the attractive features of cMBE is that the higher body corrections can be formed in either the TPS basis or the orbital basis using an effective embedding approach. We briefly describe the cMF method along with the orbital optimization and the cMBE method in Section 3.2. We then show the performance of the cMBE method by applying it to the strongly correlated 1D and 2D Hubbard models in Section 3.3.1. In Section 3.3.2, we study the commonly benchmarked strongly correlated dichromium system. In Section 3.3.3, we apply the cMBE method to the delocalized polycyclic aromatic hydrocarbons (PAH). We also study the benzene molecule using the cc-pVDZ basis in Section 3.3.4. Finally in Section 3.4, we summarize the results and discuss future directions.

3.2 Theory

Traditional wavefunction-based methods start from a mean field determinant and expand the wavefunction as excitations from this reference. Usually the Hartree-Fock determinant is used as the reference wavefunction and contributions from the excited determinants need to be included for chemical accuracy. The Hartree-Fock orbitals are extremely delocalized and may not provide the best reference orbitals for larger systems. There are also studies in which CCSD natural orbitals or localized orbitals are used for the MBE.[58, 59] Even with these modified orbitals, the reference determinant is not changed significantly and hence can be non ideal for strongly correlated systems. We propose to use a cMF reference instead of a HF reference and expand the MBE in a TPS basis. Using this modified reference can be helpful because part of the strong correlation is already included inside the cluster, and the interaction outside can be captured using MBE or any other approach.

To understand the clustering and the TPS wavefunction, we look at the polypyyrole (**PPy**) polymer unit shown in Figure 3.1. This molecule has strong local interactions and weak inter-pyrrole interactions in its ground state and has applications in molecular switches.[144, 145] In the neutral form, it is a good test system to show the applicability of cMBE. Each pyyrole unit is a single cluster with 5 orbitals corresponding to the π space in each unit.(Figure 3.2) The exact solution of each unit can be solved since it is just a (6e,5o) active space and corresponding cluster states can be generated as shown in Figure 3.2. In Figure 3.1, for each unit, the energies of the ground and excited cluster states within each cluster have been plotted. For the reference, we form a tensor product of each of the ground states in each cluster. The many-body expansion can then be formed on top of this reference configuration. In the next section we give a brief description of the cMF method and then introduce the cMBE method in detail.

3.2.1 Cluster many-body Expansion

Although cMF provides an exact description of local correlations, as a direct product of single cluster states, it lacks entanglement between clusters. To reintroduce inter-cluster entanglement, higher energy TPS configurations need to be included to improve the wavefunction. In this framework, we define a singly excited TPS as when a single cluster is allowed to have multiple cluster states rather than just the ground cluster state.

$$|\psi_{\lambda_L}\rangle = |0_I, 0_J, ..\lambda_L, ..0_N\rangle \tag{3.1}$$

For a given single excitation, the matrix element between the reference TPS and singly excited TPS can be written as,

$$\begin{aligned}\langle 0_I, 0_J, ..\lambda_L, ..0_N|\hat{H}|0_I, 0_J, ..0_L, ..0_N\rangle = \\ \langle \lambda_L|\hat{H}|0_L\rangle = \langle \lambda_L|\hat{H}^0_L|0_L\rangle = 0.\end{aligned} \tag{3.2}$$

This matrix element is zero for a self-consistently optimized TPS reference due to the cMF stationary conditions. Hence we can define a generalized Brillouin's [16] condition for TPS's:

$$\langle \psi_S|\hat{H}|\psi_0\rangle = 0 \qquad\qquad \forall S \tag{3.3}$$

A doubly excited TPS would be when two clusters are allowed to have full degrees of freedom (Figure 3.1).

$$|\psi_{\lambda_L,\mu_M}\rangle = |0_I, 0_J, ..\lambda_L, .., \mu_M..0_N\rangle \tag{3.4}$$

where $|\lambda_L\rangle$ and $|\mu_M\rangle$ are two excited configurations in the clusters L and M respectively. In this work, we introduce an incremental approach, *cluster many-body expansion*, on top of the cMF reference to capture the rest of the correlation energy. Since cMF captures part of the correlation energy missing from Hartree-Fock, we refer to the correlation energy not captured by the cMF as *inter-cluster correlation energy*. The general many-body expansion method can be written as,

$$E_c = \sum_I \epsilon_I + \sum_{J<I} \epsilon_{IJ} + \sum_{I<J<K} \epsilon_{IJK} + ... \tag{3.5}$$

where E_c is the inter-cluster correlation energy instead of the traditional correlation energy. The two body term can be expanded as

$$\epsilon_{IJ} = E_{IJ} - E_I - E_J \tag{3.6}$$

where E_{IJ} is computed using the excited TPS configurations where cluster I and cluster J are allowed to have all excited degrees of freedom. (Figure 3.1) A three body correction for clusters I, J and K can be written as

$$\epsilon_{IJK} = E_{IJK} - E_{IJ} - E_{JK} - E_{IK} + E_I + E_J + E_K \tag{3.7}$$

As can be seen, the computational cost of higher body terms would increase drastically. If all the clusters have n states each, the 3-body term will have a variational space of n^3 in the initial Fock space configuration. For the system in Figure 3.1, this would be around 10^6 TPS configurations. This will become intractable at higher orders very quickly.

One way to tackle this problem is by using a truncated basis in each cluster. However truncating the cluster basis can affect the final energy quite a lot, especially for systems that

have non negligible interactions between clusters. As a significant improvement over energy-based trunctation, we can instead choose states which are highly entangled via the embedded Schmidt truncation introduced in our previous work.[75] However, even though this can significantly reduce the number of necessary states, the full dimension formed using the tensor product of the states of each cluster will still grow exponentially. Recently we proposed the tensor product selected configuration interaction (TPSCI) method which approximates the exact solution as a variational linear combination of tensor product states which are chosen by a selected CI procedure. TPSCI is ideal for our current purposes since it will adaptively form the wavefunction depending on the interaction between clusters.

One issue with the TPS based approaches is the expensive matrix element evaluation compared to Slater determinant based approaches. This can be considered as one of the advantages of the cMBE approach compared to other TPS based approaches since the expansion can be computed by avoiding the TPS basis altogether. For example, for a dimer term E_{IJ}, we can compute the effective integrals inside a combined cluster IJ by combining the two clusters. Hence if cluster I has n_I electrons and cluster J has n_J electrons, the combination of the two clusters forms a new cluster (IJ) with $n_I + n_J$ electrons. This is similar to forming the effective Hamiltonian inside an active space in a CASCI calculation. The effective Hamiltonian inside the new cluster can be formed and the CASCI problem can be solved in the determinant basis. Even for the orbital basis approach where we avoid the TPSs, the new combined active space can be large for higher order terms and approximate approaches ultimately have to be used. This can be solved using any approximate FCI method, such as selected CI or DMRG or even CCSD(T) if the CAS space becomes large. Although both methods can be used, it is difficult to tell a priori which one will be ideal for computing higher MBE terms for a general system. If a system can be clustered efficiently, the TPSCI approach offers unique advantages arising from the natural representation which mirrors

the physical system. Therefore, we use the orbital basis approach for smaller sized cluster systems since CASCI is cheaper and use the TPSCI based approach for large clusters.

Similar to other many-body expansion methods, the approximate dimer and trimer terms can be computed using any many-electron method. If the system is not fully strongly correlated, traditional methods like CCSD(T) can be used to solve for cMBE terms. For example, if we have a molecular crystal, coupled cluster is a good ansatz for the ground state of each monomer as well as the full system. Hence a CCSD(T) result can be achieved using a many-body expansion with CCSD(T) results for dimer and trimer systems. We can even start from a cMF reference with an approximate CCSD density in equation 2.25 and use it as the reference for the MBE. Even though these are interesting possible directions where cMBE can give very convergent results, in this study we only focus on the correlation energy. The cMBE method is also exactly size extensive even though it is not variational.

3.3 Results

In this section we present data for the cMBE method for a variety of systems. First we study the half-filled one- and two-dimensional Hubbard model. We then apply the cMBE method on the strongly correlated dichromium system and some polycyclic aromatic hydrocarbon (PAH) systems. The PAH systems are extensively delocalized and hence can be considered challenging for the cMBE approach. Finally, we apply the cMBE method on the recently bench-marked benzene molecule with cc-pVDZ basis.[2] The integrals for the molecular systems were computed using PySCF package.[146] All DMRG calculations provided for the Hubbard model in this work have been carried out using the `ITensor` package.[147]

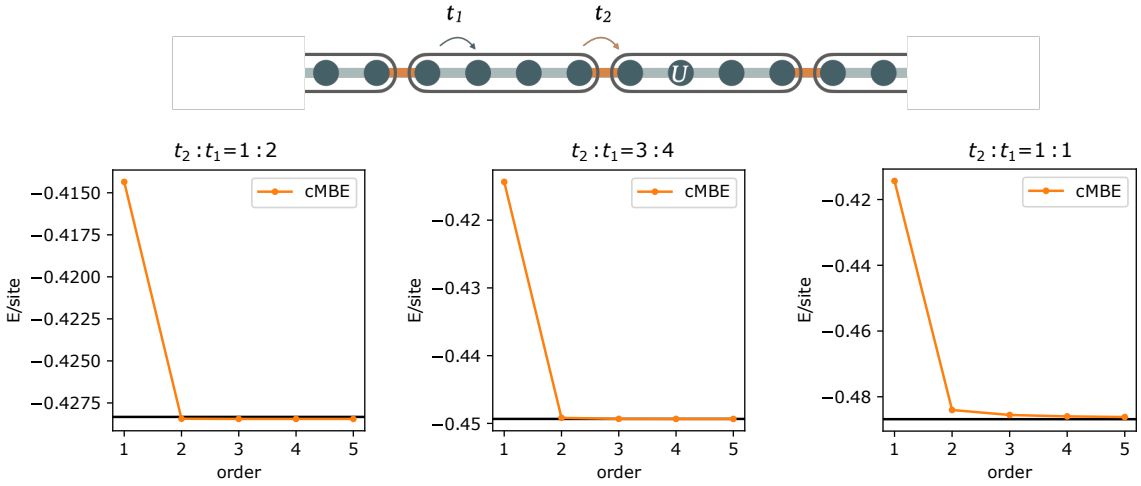


Figure 3.3: The cMBE energy per site for the 40 site Hubbard model. The black line corresponds to DMRG results.

3.3.1 Hubbard Model

In this section, we study the one and two dimensional Hubbard model[148] using the cMBE method. The Hubbard Hamiltonian used in this study can be represented using two different hopping values:

$$\hat{H} = \sum_{\langle i,j\in A\rangle\sigma} -t_1 a^\dagger_{i\sigma} a_{j\sigma} + \sum_{\langle i\in A, j\in B\rangle\sigma} -t_2 a^\dagger_{i\sigma} a_{j\sigma} + U \sum_j n_{i\uparrow} n_{j\downarrow} \tag{3.8}$$

where t_1 (t_2) are hopping terms within (between) clusters, and U is the same-site Coulomb repulsion. The Hubbard model becomes strongly correlated when the two-electron Coulomb repulsion (U) is much larger than the hopping term. For all calculations, we use $U = 5$ and $t_1 = 1$ which is strongly correlated regime. We study the effect of *clusterabilty* of the system by scanning the t_2 hopping term with respect to the t_1 parameter. For example, cMF would be exact for $t_2 = 0$ since the clusters are non-interacting.

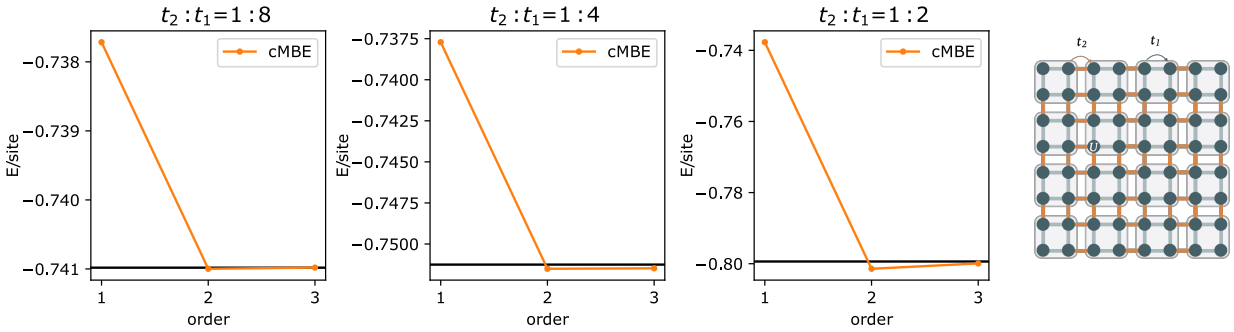


Figure 3.4: The many-body expansion for the 64 site 2D Hubbard model. The black line corresponds to reference DMRG result with M=1600 for $t_2 : t_1$ of 1 : 8 and 1 : 4 and M=3000 for ratio of 1 : 2.

1D chain

For the 1D system considered, we present results for values of $t_2 : t_1 = 1 : 1$, 3 : 4 and 1 : 2 for a 40 site periodic Hubbard system. We divide the system into 10 four site clusters. Since this is a 1D system, we used DMRG for computation of all of the higher order terms in the cMBE method. We used the original local orbitals and did not perform any orbital optimization since orbital optimization can remove some of the sparseness leading to more terms.[75]

Even though the 1D system is exactly solvable using DMRG, the cMBE gives us a good indication of *clusterability* of these systems. From Figure 3.3 it can be seen that for all cases except for $t_2 : t_1 = 1 : 1$, the cMBE expansion converges quickly, almost at second order, even at $t_2 : t_1 = 3 : 4$.

2D lattice

Here we attempt to study the two dimensional Hubbard model with 64 sites as an example. We have 16 clusters with 4 sites each. We consider $t_2 : t_1 = 1 : 8$, 1 : 4 and 1 : 2 for these systems. For the cMBE results, we use TPSCI to solve for higher body corrections. We

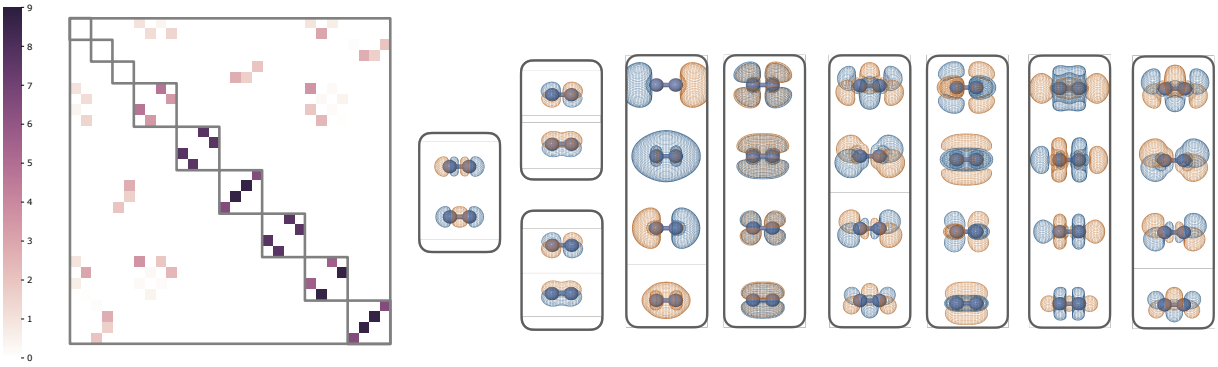


Figure 3.5: The clustering of the dichromium system. The absolute value of the exchange matrix is plotted and the clusters are selected as blocks. The ordering of the orbital is similar to the ordering of orbitals in the right. Each 4 orbital cluster corresponds to a bond in the hextuple bonded Cr dimer.

use the `FermiCluster` package developed by our group for the cMBE results.[149] The two-dimensional Hubbard model has been studied previously using an increment based approach and has shown promising results just by using third order corrections.[150]

From the results shown in Figure 3.4, we can see that for a ratio like 1 : 8, the cMBE approach converges quickly. For these two dimensional systems, the DMRG results also get complicated as we go to higher ratios. We use the variational TPSCI results for the terms at each order. The DMRG values are computed using M=1600 except for the case where $t_2 : t_1 = 1 : 2$ where we use M = 3000. From Figure 3.4 it can be seen that the cMBE and DMRG values match well for all ratios considered. We can conclude from the results that the cMBE method can be used for strongly correlated systems with reasonable inter cluster interactions.

3.3.2 Chromium Dimer

The chromium dimer at 1.5 using an Ahlrichs-SV basis set[151] is a common test system used to study methods developed for strong correlation.[93, 152, 153, 154] All orbitals up to 3s orbitals are frozen, leading to an active space with 24 electrons in 30 orbitals. This system is studied to benchmark new methods developed for static and dynamic correlation, since it has a hextuple bond with one σ, two π and three δ bonds involving the 4s orbitals and the 3d valence orbitals. There are benchmark results computed using DMRG [152, 153], FCIQMC [154] and SHCI [155] among others. In a recent study, Lehtola and coworkers showed that there are excitations as high as octuples using their cluster decomposition method.[156] The Cr_2 has also been studied using the MBE-FCI approach,[58] where it was observed that the many-body expansion converges at approximately the 10th excitation rank.

The cMBE approach requires us to partition the orbital space of the system into clusters. In the case of the chromium dimer, this may initially seem difficult. A simple clustering approach is to use each bond as a separate cluster. As shown by the exchange matrix in Figure 3.5, the $3d$ orbitals and $4d$ orbitals with the similar shape have large off-diagonal elements, implying significant interaction. Hence, we put these $3d$ and $4d$ orbitals in the same cluster. The bonding and antibonding orbital pairs formed from each atom's $3p_x$, $3p_y$, and $3p_z$ atomic orbitals, which are fully occupied, are each treated as their own cluster. The orbital clustering (shown in Figure 3.5) is:

($3p_z$), ($3p_x$), ($3p_y$), (3s, 4s), ($3d_{xy}$, $4d_{xy}$), ($3d_{yz}$, $4d_{yz}$), ($3d_{xz}$, $4d_{xz}$),($3d_{x^2-y^2}$, $4d_{x^2-y^2}$), ($3d_{z^2}$, $4d_{z^2}$)

We present the data for the Cr dimer in Table 3.1. Coupled cluster with up to quadruple exictations is unable to get a good estimate for this system.[152] The cMBE result for the 4 body correction is within chemical accuracy compared to other FCI quality results.[153, 154]

Table 3.1: Correction at each order up to four-body correction for the Cr_2 system (24e,30o). The orbital basis used is RHF. We present reference values for other methods using the same HF core.

Order	cMBE
1	-2085.9921
2	-2086.4482
3	-2086.3278
4	-2086.4211
Method	**Energy**
CCSDTQ[152]	-2086.4067
DMRG[153]	-2086.4211
FCIQMC[154]	-2086.4212
SHCI[155]	-2086.4211

We use the tightly converged TPSCI+PT results for the individual terms for the cMBE. We use the selection threshold 1e-10 and search threshold 1e-3 for the variational space.

cMBE or any other TPS based methods are ideally designed for studying spatially extended molecules where localized orbitals can be used and the drastic decay of electron correlation can be taken advantage of. The presented Cr dimer data demonstrates that the cMBE approach, even though not meant or designed for small systems such as diatomics, is capable of providing very accurate results. To be definitive, more accurate calculations incorporating 5-body terms is needed to confirm that the cMBE calculation has actually converged on the accurate point (especially since the 4-body terms provided a significant effect). However, we were not able to converge the individual 5-body terms accurately enough with the current implementation. In future work we plan to further exploit the cluster basis representation to simplify such higher-body terms.

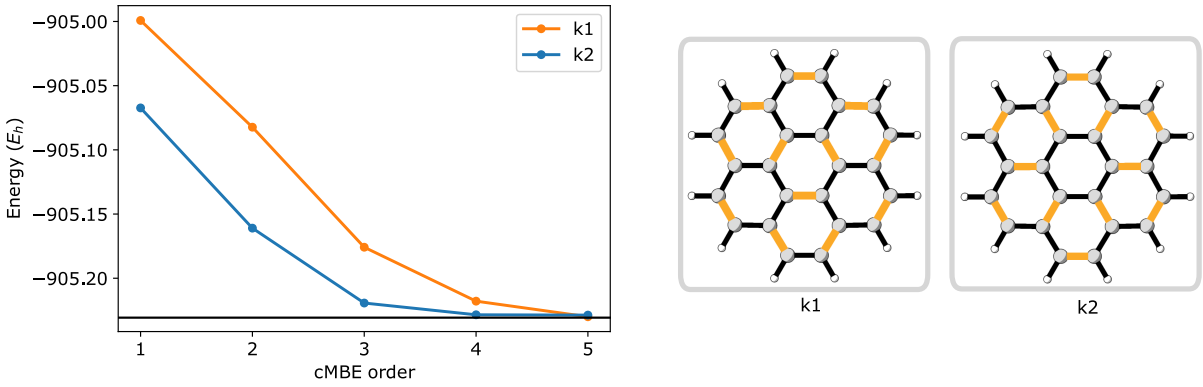


Figure 3.6: Comparison of the two different clustering options considered for the coronene molecule. The highlighted yellow region corresponds to the atom pairs in a cluster.

3.3.3 PAH system

Next, we study several polycyclic aromatic hydrocarbon (PAH) systems using the cMBE method. Due to the extended π conjugation in these systems, they have possible applications in light-emitting diodes, solar cells and so on. Large graphene nanodots [157] have applications in bioimaging and photovoltaics and can be considered as strongly correlated systems. The extended π-conjugation in these systems makes them a relatively hard problem for a fragmentation based approach like cMBE. Hence the PAH systems should be challenging for the cMBE method. Geometries for all the PAH systems in this study were optimized using B3LYP with the cc-pVDZ basis except for the hexacene system whose geometry was obtained from Ref. [158]. All structures are provided in the Supporting Information.

There exist multiple classical rules which give a good qualitative idea of the correlation and structure of PAH systems.[159, 160, 161] One such model is the Kekulé structure. The PAH molecules are mainly formed by conjugated double bonds, hence it makes sense to cluster PAH systems using one of its Kekulé structures as the guiding clustering option. Because this leads to a two orbital cluster, we can include higher order corrections to see the convergence properties of the cMBE approach.

Different Kekulé structures

For a simple PAH system, there can be multiple Kekulé clusterings possible. In this section, we study the coronene molecule using two different clustering approaches. It has been previously debated whether the coronene molecule has weak outer double bonds or two concentric π-conjugations, leading to clustering in **k1** or **k2** respectively in Figure 3.6. cMBE can be used to determine which clustering provides a more physically correct picture. Based on just the Kekulé structure, it is difficult to say which of the two clustering is ideal. We present data for both clustering approaches in Figure 3.6. Both Kekulé clusters provide the same Clar's structure in the end, but it can be seen that **k2** clustering is more convergent than **k1**. We also note that we found the cMF reference to have a lower energy for **k2** compared to **k1** which can be used as a good indicator for relatively better clustering. Using this as a metric, we can avoid doing the expensive cMBE for all possible Kekulé structures. This is usually the case with systems which do not have too many empty virtual orbital clusters. It has been previously demonstrated experimentally and theoretically that the coronene molecule has weak outer double bonds unlike benzene which clearly suggests that the **k2** clustering would be better.[162, 163] Even with **k2**, we need to go to higher orders for the expansion to converge since the system is delocalized.

Comparison between cMF and split localized RHF basis

Now, we move on to evaluate the effect of starting from a cMF reference by comparing it with the RHF reference. As mentioned, the MBE can be formulated without the cMF reference by directly using RHF orbitals. We take a few PAH molecules and localize the occupied and virtual π-orbitals separately. We then cluster the orbitals based on bonding/antibonding pairs. This leads to an automatic stable Kekulé type clustering for some PAH systems. In

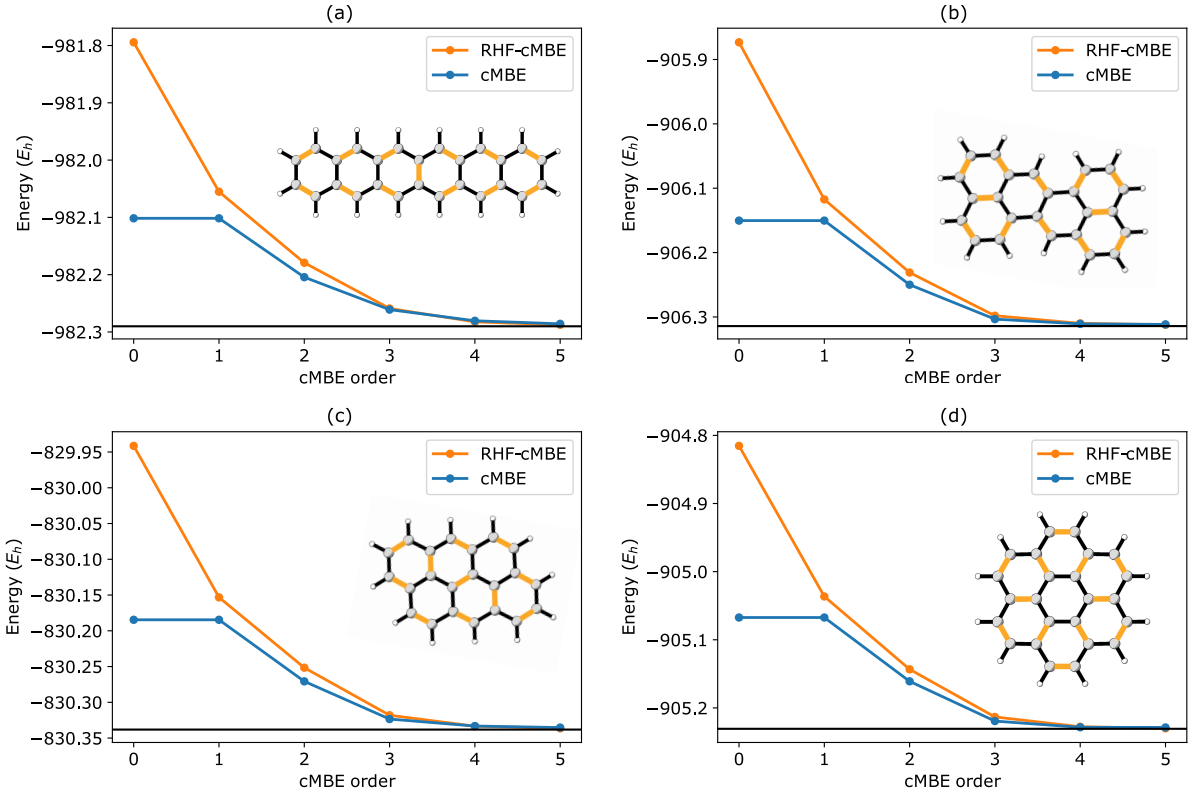


Figure 3.7: Comparison of cMF reference with using split localized RHF basis for four PAH systems. The yellow highlighted region of the molecules corresponds to the double bonds considered as clusters. The reference value corresponds to the DMRG result for (a) and extrapolated SHCI results for (b),(c) and (d).

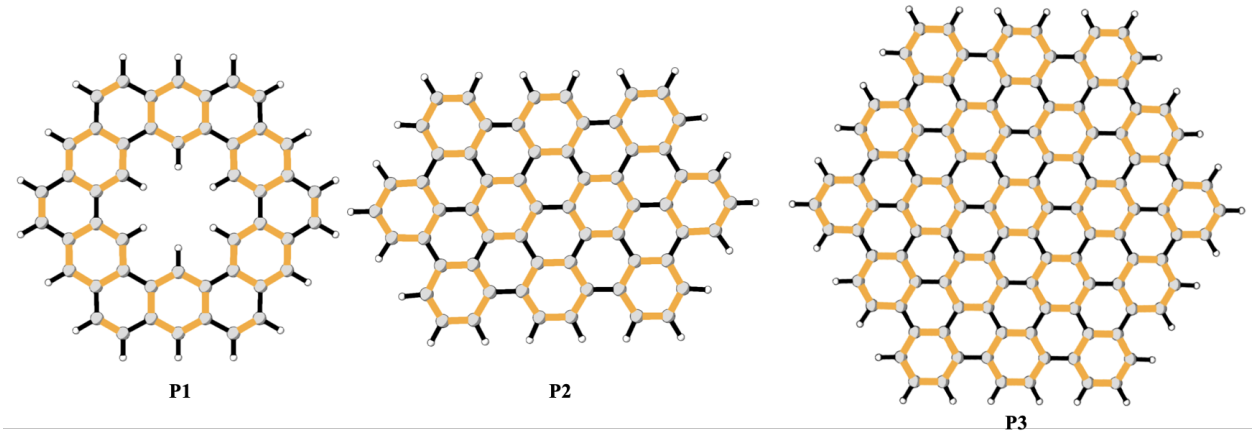


Figure 3.8: Large PAH systems considered in this work using the Clar's rule clustering. The active space comprises of the π-conjugated electrons. The yellow highlighted regions corresponds to a single cluster.

contrast to the cMF reference, where part of the correlation is already captured, the MBE using RHF reference starts from a less stable RHF solution. For the RHF based MBE, the reference is the RHF determinant with all the occupied orbitals doubly occupied. The first order correction for this RHF-cMBE amounts to a CAS calculation for a single cluster while constraining the occupied orbital in all other clusters to be doubly occupied. The two-body correction is then a CAS calculation by combining the orbitals in the two clusters while others are doubly occupied and so on.

We present data for the comparison of the cMBE with the split localized RHF-cMBE method in Figure 3.7. It can be seen that the cMF reference is much better than the RHF reference. The cMBE and the RHF-cMBE approach converge to the exact result at higher orders with cMBE having better convergence for all the systems. The reference energies are extrapolated SHCI values computed using the `Arrow`[164] package except for the hexacene molecule, for which we use the DMRG value from Ref. [158]. Pruning can be used for truncating the number of terms for larger systems. In the Supporting Information, we present two different pruning techniques, `scheme-1` and `scheme-2` and apply it to the Kekulene molecule.

Table 3.2: Data for the large PAH systems studied using the Clar's rule clustering.

	CCSD(T)	Extrp SHCI	cMF	cMBE2	cMBE3
P1 (48e,48o)	-1810.3857	-1810.3842	-1810.2330	-1810.3578	-1810.3838
P2 (60e,60o)	-2258.5624	-2258.5360	-2258.3885	-2258.5262	-2258.5559
P3 (114e,114o)	-4283.9319	-	-4284.2751	-4284.5790	-4284.5999

Clar's clustering

For the PAH systems, even though the Kekulé clustering gives very accurate results, we need to go to larger orders of cMBE. In this section, we consider a larger cluster size based on Clar's rule such that most correlation is captured within the clusters and the cMBE can be truncated at lower orders. The Clar's aromatic sextet rule has also been used with a fragmentation based DFT approach and provide promising results.[165] For larger systems, the Clar's rule based clustering would be more ideal.

First we consider the Kekulene molecule. Kekulene, similar to coronene also has two possible structures: the two superaromatic inner and outer ring or the Clar's sextet based structure. The Kekulene molecule was recently synthesized and visualized using ultra-high-resolution atomic force microscopy (AFM) and the superaromatic behaviour was not observed.[166] There are computational studies also suggesting Kekulene to have the Clar's rule based structure,[167] hence we expect more convergent behaviour using the Clar's clustering.

The Kekulene molecule (**P1**) in Clar's type clustering leads to a 12 cluster system with six two orbital clusters and six sextet clusters. The **P2** molecule, which is a part of a graphene nanosheet has an active space of 60 orbitals in 60 electrons. This system requires about 10^{34} determinants for the exact results. We also study an even larger nonographene system **P3** which has an active space of 114 orbitals in 114 electrons. The FCI space for this molecule in the π space would have 10^{66} determinants for the ground state. This system has 19 clusters and we provide data up to third order correction.

For all of these systems, we use the energy obtained using the TPSCI+PT for the tuples. The TPSCI method, being a selected CI approach, forms a smaller variational space for clusters that interact less. For example, the interaction between the π sextets at the two corners of the **P3** molecule are nearly negligible. Hence considering the two body term between these two clusters have a variational space of 94 TPS configurations. The variational space for one of the nearest neighbour interactions had approximately 4000 configurations. Both these values are much less than the full dimension for a 12 orbitals in 12 electrons active space which has 853,776 determinants. Using a RHF reference and performing CAS calculations for the many-body expansion would be intractable at 3rd order. Hence using a TPSCI procedure, we avoid forming the full space dimension for the two and three body terms for the cMBE approach. Using the Clar's clustering, the cMBE approach should converge faster and higher order corrections would not be required. Obtaining higher body corrections would require further coarse-graining of the variational space. There are multiple possible ways this can be achieved which will be discussed in future work.

In Table 3.2, we present data for the large PAH systems using the Clar's rule based clustering and compare them to extrapolated SHCI and CCSD(T) results. For the **P1** molecule, the errors are within $0.2mE_h$. For the next largest PAH system, **P2**, the storage of the PT space gets very large, hence the SHCI values could not be computed at very accurate threshold. Hence the extrapolated SHCI number for this system, as seen from Table 3.2 is not a good estimate. The **P3** molecule is a larger graphene type system and has extended electron delocalization. It is interesting to observe that the variational cMF energy is already lower in energy that CCSD(T), even before adding any many-body interactions.

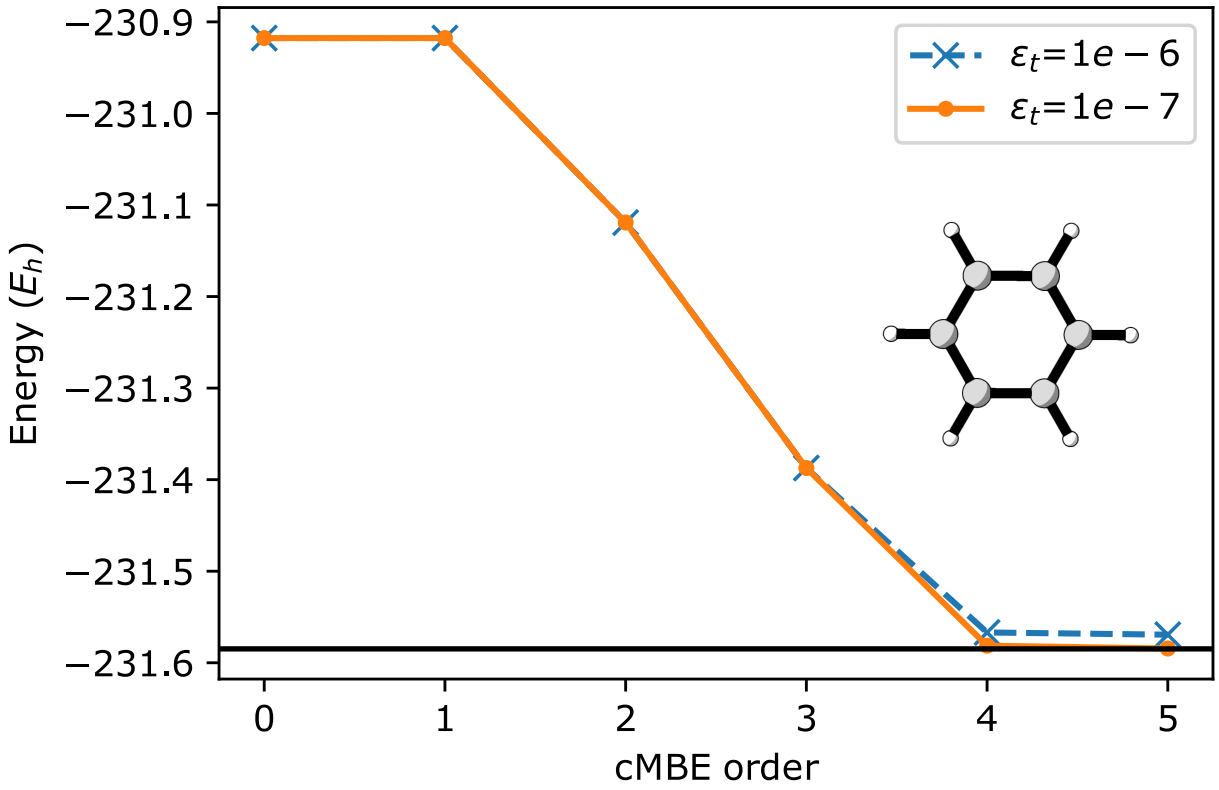


Figure 3.9: The cMBE values for the Benzene system with cc-pVDZ basis using `scheme-1` pruning. The threhsold used for two different cMBE calculation is also given in the legend. The black line corresponds to the estimated energy value from Ref. [2].

3.3.4 Benzene cc-pVDZ

In a recent review, most of the methods mentioned in the introduction section were tested in a blind challenge to approach the full configuration interaction energy of the benzene molecule in a cc-pVDZ basis.[2] Even though the benzene molecule is not strongly correlated, the numerically exact result for this large active space with 30 electrons in 108 orbitals is not trivial. The correlation energy for this system was estimated to be $-863mE_h$.[2]

There is no simple strategy to cluster the benzene molecule and partitioning the benzene system is a non-trivial problem. However since the focus of this work is not the optimal clustering but rather the cMBE method itself, we do not yet explore different clustering possibilities. We localize the π space by taking both occupied and virtual orbitals from the π bond formed by the $2p_z$ orbitals. The rest of the occupied and virtual orbitals are localized separately. We use a clustering where all the σ bonds and the corresponding antibonding orbitals are in a single cluster. The π space is then clustered using a Kekulé structure. Hence, the valence and occupied space are relatively easier to cluster compared to the large virtual space. For simplicity, we also clustered the full valence space using a Kekulé criteria by pairing the orbitals with similar shape in adjacent C atoms into a cluster. This basically means that the system is partitioned into three main parts, orbitals corresponding to C1-C2, C3-C4 and C5-C6. Within these groups the orbitals are clustered such that orbitals with same principle quantum numbers, angular momentum and shape are in the same cluster. For example, each of the $3p_x$ orbitals in C1 and C2 are in a cluster and similarly for other orbitals. This makes a total of 54 clusters with 6 C-H clusters, 6 C-C clusters and 14$\times$ 3 clusters for the valence π space and virtual orbtials. The clustering for the benzene system is pictorially depicted in Supporting Information.

For the benzene molecule, we present the data in Figure 3.9. We use the `scheme-1` pruning

for the benzene system. All the tuples are considered up to third order. We present data for two different pruning thresholds ϵ_t, `1e-6` E_h and `1e-7` E_h . The ϵ_t is the pruning cutoff as discussed in the Supproting Information. For the loose threshold, the correlation energy is not fully captured even at 5 body. This is expected since we have a large virtual orbital space and we need to include the effect of even smaller contributions. For the tighter threshold, we get correlation energy of $-862.6\ mE_h$ at 5 body which is just $.4\ mE_h$ away from the estimated FCI result. Only a very small fraction of the terms are considered using pruning. For the tight threshold calculation, only 16 % and 2 % of the total possible terms are considered after pruning at 4 body and 5 body, respectively. The most expensive calculation for the cMBE is the 5 body terms, which still is only a 10 orbital active space and is solvable without much computational resources.

3.4 Conclusion

We present a new increment based method on top of a cMF wavefunction formed by clustering the strongly interacting orbitals into clusters. We show how the cMBE method can be used to obtain the correlation energy for different types of systems including model Hamiltonians and molecular systems. The cMBE approach gives good results for systems with large active spaces as presented in this work. If only expectation values are needed, it can be used as an alternative to other methods like DMRG and selected CI when the wavefunction is approximately clusterable. We test this method on large π-conjugated systems and even the strongly correlated dichromium system.

Future work will focus on further decreasing the computational cost. One simple approach could be to use an approximate treatment (e.g., CCSD(T)) for the cluster calculations. Another future direction could be to include a PT correction from other clusters on top

of the cMBE similar to using RPA or MP correction terms for higher orders in traditional methods.[168] This would lead to higher body corrections at every correction level, which can improve the convergence of the expansion. We can also introduce a better pruning criteria than the energy based criteria used in this work. For example, a distance based criteria can also be used for the PAH systems.[169, 170]

One of the main future directions would be to extend the cMBE method to do larger systems and apply it to other properties like excited states and interaction energies. Recently the many-body expansion with diabatic states method was proposed and applied to charge transfer reactions.[121] A similar framework can be used with a multi-reference cMF to study excited states using the cMBE approach.

Another interesting direction we are interested in is to use the TPS framework and coarse-grain the degrees of freedom at each order. This would lead to a compact wavefunction and allow application of the TPS framework for higher body corrections for larger clusters.

3.5 Acknowledgements

This research was supported by the National Science Foundation (Award No. 1752612).

3.6 Appendix

3.6.1 Pruning for large active spaces

We study a larger system, Kekulene, which has a large active space with 48 electrons in 48 orbitals using the conjugated double bond type clustering. With such a large system, the

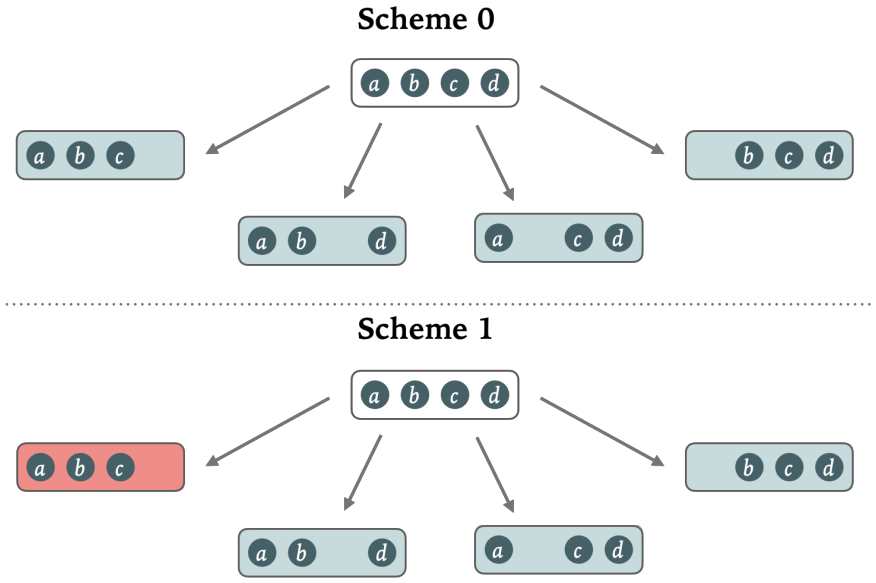


Figure 3.10: The pruning `scheme-0` and `scheme-1` used in this work. For `scheme-0` all the child tuples energy contributions need to be above the threshold (ϵ_t) for the parent tuple calculation to be carried out. For `scheme-1`, all except one child tuple's energy contribution needs to be above the threshold.

number of calculations at higher orders becomes very large. It is prudent to use pruning to cut down the number of calculations. Pruning of terms is common in MBE based approaches.[56, 59] We define two types of pruning techniques to cut down the factorial cost of the number of terms in the cMBE approach. For example, let us consider a four body term with clusters `[a,b,c,d]`. The decision to compute a tuple `[a,b,c,d]` is made by looking at the tuples `[a,b,c]` `[a,b,d]` `[b,c,d]` and `[a,c,d]`. We define the `scheme-0` pruning, where the tuple `[a,b,c,d]` is computed if the energy contribution of all the child tuples are above a given threshold ϵ_t. This is similar to the pruning introduced by Eriksen *et al.*.[59] This leads to the use of very tight ϵ_t values for good accuracy. We can further relax the criteria and define a `scheme-1` pruning, where the tuple is computed if N - 1 tuples are important out of the N possible tuples.

We can keep going further to even more relaxed criteria, but the pruning might not be very

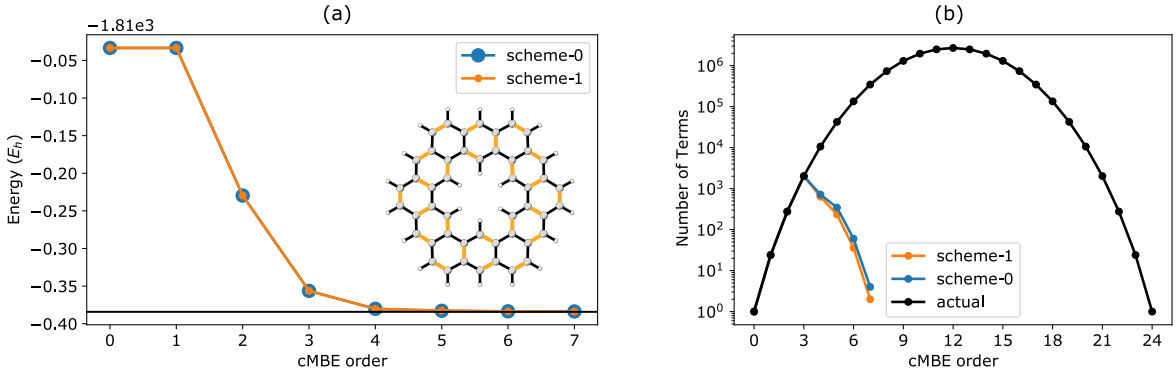


Figure 3.11: Comparison of the two pruning algorithms for the kekulene molecule using the Kekulé clustering. a) the energy per order is plotted with respect to the cMBE order. b) Number of terms computed for `scheme-1` and `scheme-0` compared with the `actual` number of terms needed at each order.

efficient. The reason for using a relaxed criteria can be understood by looking at a 1D lattice with three adjacent clusters `a,b,c`. With the `scheme-0` threshold, the tuple `[a,c]` which is a next nearest neighbour tuple may be below a given ϵ_t and will get ignored in `scheme-0`. However, `[a,b,c]` is a linear three body term which is important for the final energy. Using a `scheme-1` criteria allows one of the three tuples to be negative and hence we can have a sensible truncation. Both the pruning schemes are pictorially depicted in Figure 3.10.

We study the Kekulene molecule using both the pruning criteria. All terms up to three body were considered. For the `scheme-0`, we used a pruning threshold ϵ_t of `1e-6` mE_h for the 4-body terms. For the `scheme-1` method, a more relaxed pruning threshold ϵ_t of `5e-6` mE_h was used. The threshold was decreased by a factor of 2 for the subsequent terms. It would be interesting to study the effect of different pruning algorithms on the final energy of the system using the cMBE approach but that is beyond the scope of current study.

Results for the Kekulene system are presented in Figure 3.11. The `scheme-1` method with a relaxed threshold is a bit better than the `scheme-0` but the errors are within chemical accuracy for both pruning algorithms and hence non-distinguishable at convergence. It can

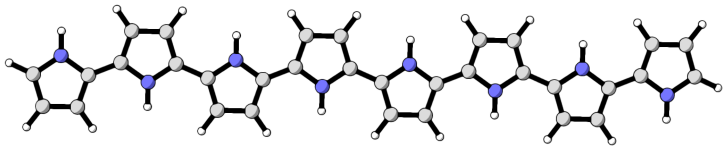


Figure 3.12: The **PPy** molecule considered in this study.

be seen from Figure 3.11 that the number of calculations is much smaller compared to the full expansion and can be computed easily.

3.6.2 Polypyyrole

We also study a polymer unit with a heteroatom. **PPy** can be used in biomaterial applications like artificial muscles because of their biocompatible material. [171, 172] Due to its high conductivity and simple synthetic procedure it also has applications to biosensors,[144] renewable cathodes, [145] efficient photocatalysts,[173] and supercapacitors.[174] Here we present the **PPy** molecule in its reduced neutral form with 8 pyyrole units as shown in Figure 3.12 as an example. The active space for a single pyyrole unit is (5o,6e) using π space orbitals. Therefore the full active space for the system becomes (40o,48e).

We present the cMBE results for **PPy** molecule in Table 3.3. We provide CCSD(T) energy for comparison. The error between the CCSD(T) result and the cMBE value is within 1mH. In the oxidized bipolaron form [175] the molecule becomes more delocalized and studying that with a single cMF reference would not be sufficient. A multi-reference version of cMF can be used in that case. The cMBE can be defined on top of this MR cMF similar to the

Table 3.3: Correction at each order up to 3 body correction for the **PPy** polymer with 8 pyrrole units.

Order	cMBE
1	-1642.4376
2	-1642.4878
3	-1642.4908
Method	Energy
CCSD(T)	-1642.4896

work developed by Fertitta *et al.* for Slater determinants[128] and will be investigated in future studies.

3.6.3 Clustering of Benzene molecule

Clustering used for the benzene benchmark.

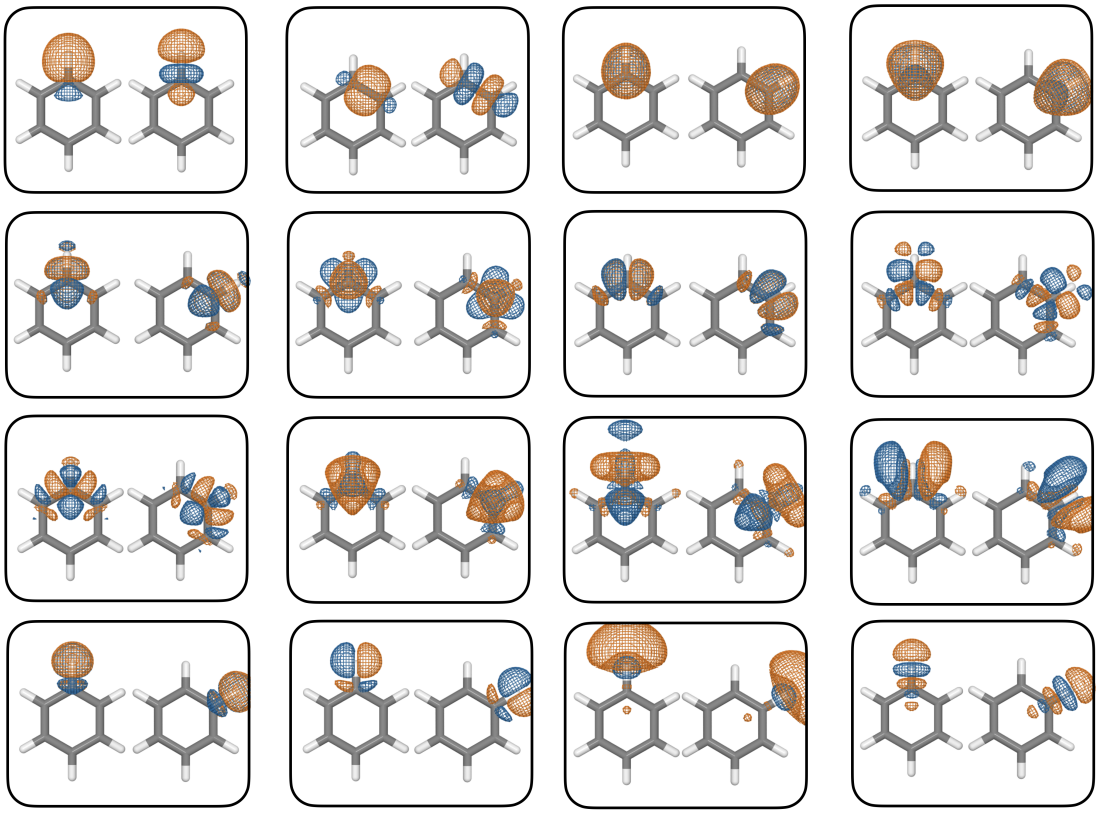


Figure 3.13: The localized orbitals of benzene using the cc-pVDZ basis for C1-C2 The C-C and C-H sigma bonding /antibonding orbitals were taken as a cluster as shown in the first two panels. The rest of the valence π orbitals and the virtual orbitals were partitioned using the Kekulé structure, where similar orbitals in C1 and C2 are paired.

Chapter 4

Tensor Product Selected Configuration Interaction

4.1 Introduction

The efficient simulation of strongly correlated electrons remains a key challenge toward better understanding several critical areas of chemical and molecular sciences including catalysis, organometallic chemistry, excited state processes, and many more. Although the term "strongly correlated" is rather ambiguously defined, we will take this to mean a system which cannot be efficiently and accurately modeled using perturbative or diagrammatic techniques starting from a single Slater determinant wavefunction. For systems dominated by one-electron interactions, Hartree-Fock (HF) represents a useful approach, such that the resulting single Slater determinant wavefunctions are qualitatively accurate. Consequently, the full configuration interaction (FCI) wavefunction written as a sum of all possible determinants bears a near unit coefficient weighting the HF ground state determinant. Methods such as perturbation theory, or coupled-cluster work extremely well in this regime. How-

ever, as the relative strength of the two-electron component increases, the HF state becomes less useful as an approximation, with many different determinants contributing significantly. This leads to a breakdown of most common approximations, such as perturbation theory, coupled-cluster theory, density functional theory, etc.

Although any algorithm which solves an *arbitrary* strongly correlated system is likely to exhibit exponential scaling, it is often the case that a molecule's Hamiltonian has some structure that can be exploited to make the problem easier. For example, if a molecule's strong correlation arises due to an orbital near-degeneracy, then active-space methods are effective in obtaining accurate results from a relatively simple computation. For systems which have a near one-dimensional structure, matrix product states provide highly efficient representations which can be solved for using density matrix renormalization group (DMRG).[30, 176, 177] Likewise, approximately two-dimensional structure can be efficiently parameterized using projected entangled pair states (PEPS),[178, 179] and recent improvements in contraction algorithms have made these algorithms more promising for molecular applications.[180] More general tensor networks have also been explored.[181, 182, 183, 184, 185, 186] For systems whose strong correlation occurs among a relatively small subset of Slater determinants (as opposed to a single particle subset defining an active-space) one might choose to perform a configuration interaction (CI) calculation using only the important Slater determinants. This is the physical motivation for so-called "selected CI" methods which have a long history in the field,[38, 187, 188] but which have seen a resurgence during the past few years.[41, 43, 47, 189, 190, 191, 192, 193]

Selected CI methods typically involve an iterative procedure in which a CI calculation is performed within a small subspace of Slater determinants, and this subspace is iteratively refined using some search algorithm to find the most important Slater determinants. The basic algorithmic steps of a selected CI program involve:

1. Determine an initial variational space (typically either the HF determinant or the set of single and double excitations).

2. Find the ground state of the Hamiltonian in the current variational space.

3. Perform some search algorithm which identifies which determinants outside of the space are most important. This importance criterion is usually based on perturbative or energy minimization estimates.

4. Construct a new variational space based on the search results, and continue until the spaces stop changing.

Although all selected CI approaches follow these general steps, they differ in various details. As one of the first approaches of this sort, the configuration interaction perturbatively selected iteratively (CIPSI)[38, 194, 195] algorithm builds the CI space by adding determinants which have first order coefficients larger than some threshold, ϵ. CIPSI has also been improved in the last 40 years by various groups. An overview of many different CIPSI approaches across the years, including a collection of useful references, can be found in the work of Yann et al. [196] CIPSI has also been recently used with Diffusion Monte Carlo where the determinantal part is constructed using SCI. [39, 197, 198] The FCI-QMC[35, 36] method proposed by Booth and co-workers can be considered as a stochastic variant of CIPSI method.[39, 47] More recently, the Adaptive Sampling CI (ASCI)[47, 199, 200] was developed where only a few determinants with large coefficients were considered for generating new determinants. With the goal of achieving better accuracy guaranteed, the Λ-CI method adds determinants based on a variational energy criterion Λ.[201] Following their earlier work, Evangelista and co-workers then proposed the adaptive CI (ACI) method which produces compact wavefunctions with tunable accuracy.[43] As a deterministic generalization of heat-bath sampling,[202] heat-bath CI (HCI)[41, 42, 49, 203] adds determinants based on the

magnitude of the Hamiltonian matrix element. This selection criteria is very cost efficient since it avoids sampling the determinants directly by using the magnitude of the integrals themselves, skipping the denominator computation for the selection step. The Monte Carlo CI (MCCI) method, proposed by Greer, repeatedly adds interacting configurations randomly to the reference space and generates a variational space. [48, 204]

All of the SCI methods mentioned above succeed when the number of significant coefficients in the FCI wavefunction is small, and they fail when this number becomes large. This becomes problematic when the degree of strong correlation increases. Luckily, the distribution of FCI wavefunction coefficients directly depends on the choice of basis. For example, it has been observed that rotating the one-particle basis to diagonalize an approximate one-particle density matrix (natural orbitals) increases the compactness of a selected CI wavefunction.[47, 199, 205] However, single particle rotations are only the simplest type of change of basis one could consider. This presents a natural question: *Can one find a basis (not necessarily comprised of Slater determinants) which yields a more compact FCI wavefunction, thus decreasing the number of variational parameters in a SCI procedure?*

In this paper, we explore a new basis designed to provide a more compact representation of the wavefunction leading to larger scale SCI calculations. This basis is defined by performing many-body rotations on disjoint sets of orbitals (or "clusters"). The resulting tensor product state (TPS) basis can incorporate a large amount of electron correlation into the basis vectors themselves. As a result, the exact FCI wavefunction written in terms of tensor product states can be more compact, requiring significantly fewer basis vectors than in the analogous expansion in terms of Slater determinants. In the following sections we describe the construction of the tensor product state basis by way of a Tucker decomposition of a sparse global state vector,[206] and a method to exploit the resulting compactness by developing a framework for performing CIPSI calculations directly in terms of tensor product states. We

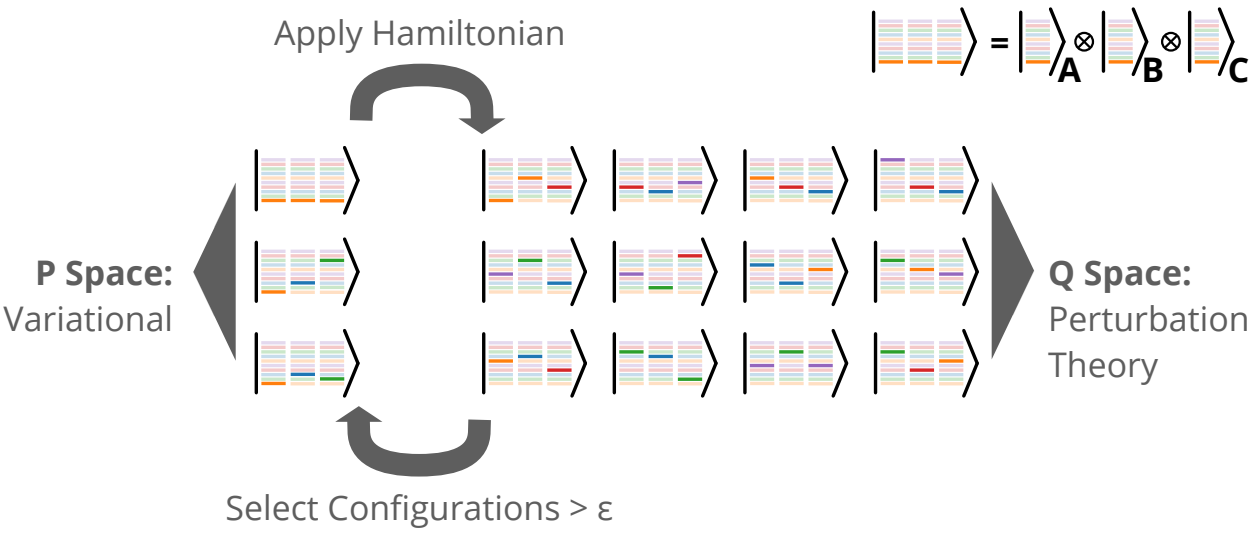


Figure 4.1: Schematic representation of the TPSCI algorithm for a three cluster problem. Each stack of lines indicates the different local states for each cluster, with different colors indicating different particle number states. Bold colors indicate that the state is activated in that basis vector. The threshold, ϵ, can be used to move states from $\mathcal{Q}$ to $\mathcal{P}$, based on the magnitude of the first order wavefunction (though other criteria could be used as well).

refer to this method as tensor product selected CI (TPSCI), and we investigate the numerical performance for a number of systems, including the Hubbard model (Sec. 4.3.1), diatomic molecular dissociation (Sec. 4.3.2), and large aromatic systems with active spaces up to 42 electrons in 42 orbitals (Sec. 4.3.3). A cartoon schematic of the TPSCI method is shown in Fig. 4.1.

4.2 Theory

In this section, we start by providing a description of our notation used to represent arbitrary tensor product states, which is similar to the work by Scuseria and co-workers in their cluster-based mean field study.[69] We then explain how the Hamiltonian matrix elements can be obtained between arbitrary tensor product states and give a layout of the TPSCI algorithm used in this paper. Finally, we discuss the Tucker decomposition technique used to further

compress the TPSCI wavefunction.

4.2.1 TPSCI Algorithm

Once the matrix elements have been implemented, then the exact state can *in principle* be obtained in the TPS basis. However, as described in the introduction, this is intractable, and we need some method for identifying important TPS configurations. To achieve this, we use the SCI strategies developed for Slater determinants and adapt them for generating a compact basis of TPS configurations. While all of the different SCI strategies described above in the introduction could be leveraged in this TPS basis, for this initial report we simply based our work on the earlier CIPSI method, often using the simplified search defined in ASCI work.[47] A schematic overview of the TPSCI method is shown in Fig. 4.1.

The algorithmic steps for TPSCI are quite similar to the Slater determinant CIPSI:

1. **Precompute cluster states and operator tensors.** Choose a technique for defining the cluster states. We find that the cMF method works well, and for many systems the orbital optimization in cMF provides significantly improved results. Using these defined cluster states, compute all 28 unique operator tensors, including those in Eq. 2.13 and 2.14, between states in compatible sectors of the local Fock space. This is a very memory intensive step, limiting the approach to clusters of about 6 spatial orbitals. However, if *unimportant* cluster states can be identified and discarded, then significant reductions in memory can be made. Many approaches to this could be imagined, and we describe one technique based on an approximate Schmidt decomposition in the Appendix. Here we also store the coefficients mapping the local Slater determinant basis to the cluster state basis.

2. **Initialize the variational $\mathcal{P}$ space.** With Slater determinants, this might be done

by choosing either the HF state or CISD space. In this work, we initialize by deciding on an initial Fock space configuration which defines how many electrons are in each cluster to begin with. We then choose the lowest energy TPS with that Fock space configuration. Alternatively, one can choose multiple Fock space configurations, and this is sometimes useful when describing delocalized states.

3. **Build the Hamiltonian matrix in the current $\mathcal{P}$ space and diagonalize.** As mentioned earlier, the matrix element evaluation (described in Sec. 2.2) is more expensive than usual determinant based codes. Although our current code builds the full Hamiltonian matrix (limiting our current work to variational spaces containing up to around 100,000 TPS), this can trivially be adapted to perform a matrix-vector product to avoid constructing the full matrix. This creates the current variational state, $|\mathcal{P}\rangle = \sum_i c_i |\mathcal{P}_i\rangle$, and variational energy, E_0.

4. **Calculate the action of the Hamiltonian on each configuration in the $\mathcal{P}$ space.** For each TPS, $|\mathcal{P}_i\rangle$, that has a variational coefficient larger than a user-defined threshold,

$$|c_i| > \epsilon_c, \tag{4.1}$$

apply each of the terms in the Hamiltonian, and collect the resulting configurations that lie in the $\mathcal{Q}$ space. This threshold, ϵ_c, is the search simplification introduced in ASCI.

For larger systems and larger clusters, the action of the Hamiltonian on the $\mathcal{P}$ space can become too large to work with efficiently. Because the Hamiltonian is relatively dense in the TPS basis, the action of the Hamiltonian on each configuration generates a large number of possible new configurations. To increase efficiency, we introduce a screening

threshold (ϵ_s) to discard negligible configurations coupled by each Hamiltonian term. Here non-negligible is defined to be terms such that:

$$|\langle \mathcal{Q}_j| \hat{H}_x |\mathcal{P}_i\rangle c_i| > \epsilon_s \tag{4.2}$$

where $\hat{H}_x$ is any of the terms in Eq. 2.5. We then add all terms $|\mathcal{Q}_j\rangle$ where Eq. 4.2 is `true`, creating a vector of coefficients in the $\mathcal{Q}$ space, $\sigma_j = \langle \mathcal{Q}_j| \hat{H} |\mathcal{P}_i\rangle c_i$.

5. **Compute first order coefficients in the $\mathcal{Q}$ space.** Using a chosen perturbative expansion (we consider either Epstein-Nesbet (EN) or the Möller-Plesset-like approach (MP) defined in Ref. [69]), compute the first-order coefficient for each configuration: $c_j^{(1)} = \langle \mathcal{Q}_j| \hat{H} |\mathcal{P}_i\rangle c_i / D_j$. If EN partitioning is chosen,[207, 208]

$$D_j = E_0 - \langle \mathcal{Q}_j| \hat{H} |\mathcal{Q}_j\rangle , \tag{4.3}$$

whereas, if MP partitioning is chosen a barycentric denominator is defined based on the effective cluster mean field operator, $\hat{F}_I$:

$$D_j = \sum_I \langle \mathcal{P}| \hat{F}_I |\mathcal{P}\rangle - \sum_I \langle \mathcal{Q}_j| \hat{F}_I |\mathcal{Q}_j\rangle , \tag{4.4}$$

Notice that this reduces to the MP-based approach defined for cMF if the $\mathcal{P}$ space contains only a single TPS.[69] If cMF is used to define the cluster states, then each mean field operator, $\hat{F}_I$, is diagonal, making it computationally efficient to compute these denominators. If, on the other hand, a different cluster basis is chosen (*vide infra*) then we simply take the diagonal elements of $\hat{F}_I$, pushing the off-diagonal elements to the perturbation. We have found that typically the MP partitioning is comparable to the EN partitioning, but at a much reduced cost.

Add any $\mathcal{Q}$ space configurations with first order coefficients larger than a threshold:

$$|c_j^{(1)}|^2 > \epsilon \tag{4.5}$$

to the $\mathcal{P}$ space, and return to step 3, exiting if no new configurations are to be added.

Finally, we can compute the full PT correction for the final variational ($\mathcal{P}$) space, by setting $\epsilon_c = 0$, computing the full matrix-vector action (σ_j), and performing a dot product:

$$\Delta E_{PT2} = \sum_j c_j^{(1)} \sigma_j \tag{4.6}$$

This is usually the most expensive step in the entire calculation.

4.2.2 Tucker Decomposition

Up to this point, we have mainly considered the use of cMF for generating the cluster state basis. This is a natural choice, as it variationally minimizes the reference TPS. However, this basis often does not lead to sufficiently compact representations after introducing entanglement. If only two clusters are present, then an SVD of the global state could be performed which would provide a maximally compact representation. When considering systems with multiple clusters, no such unique and optimal decomposition exists. However, we can still use the same physical intuition and define the cluster state basis to be the eigenvectors of the cluster's reduced density matrix. For two clusters, this is of course equivalent to an SVD of the global state, or a MPS. When the number of clusters is greater than two, this generalizes to a higher-order SVD, or Tucker decomposition.[136, 206]

$$\mathcal{T}_{i,j,\ldots,d} = \mathcal{C}_{\alpha,\beta,\ldots,\gamma} U_{i,\alpha} U_{j,\beta} \cdots U_{d,\gamma} \tag{4.7}$$

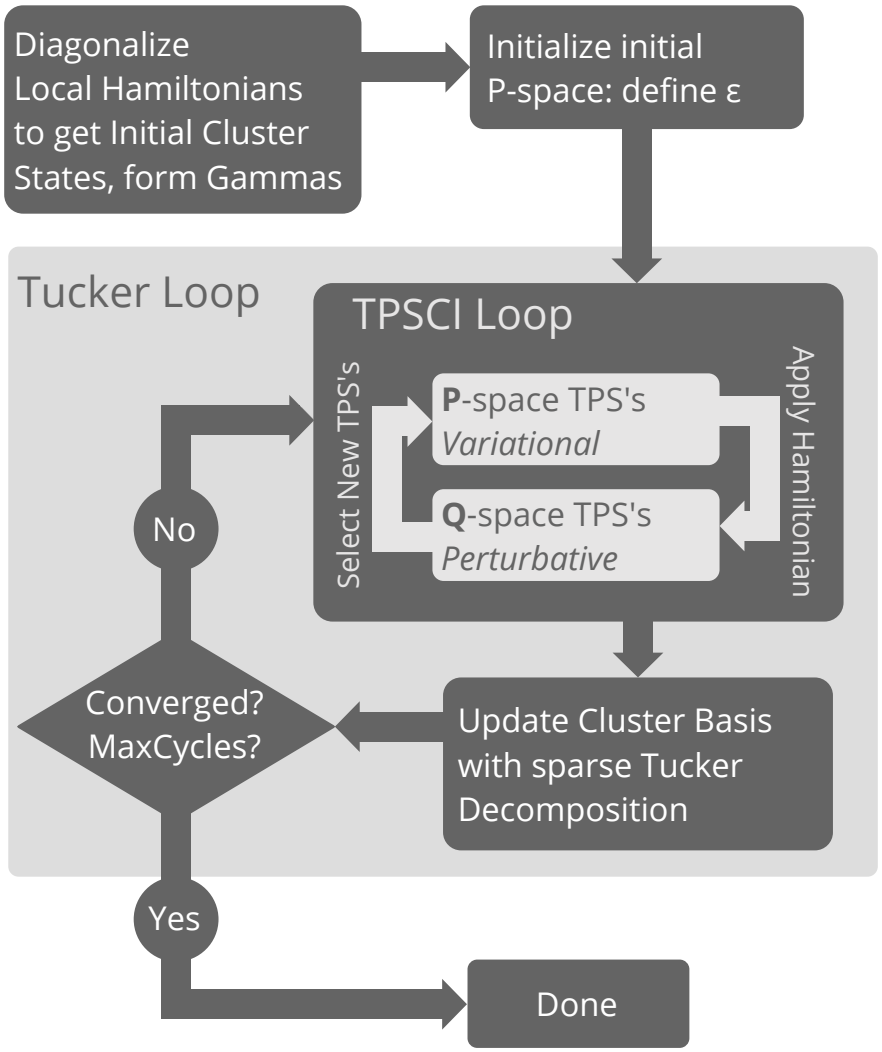


Figure 4.2: Self consistent Tucker decomposition loop. After optimizing an approximate global state via TPSCI, the sparse tensor contraction can be easily used to perform a Tucker decomposition of the state. This involves diagonalizing the single cluster reduced density matrices. To retain local quantum numbers, we only block diagonalize the RDMs within a given Fock space. Because the Tucker decomposition often significantly decreases the number of variational parameters, one can optionally start with a loose threshold ϵ_0 to get a better set of cluster states, then tighten the threshold until it reaches the desired value, ϵ, a procedure we refer to as "bootstrapping" explained in the Supplementary Information.

In the absence of truncation, Eq. 4.7 is essentially a change of basis, from i to α, from j to β, etc. Each matrix U can be obtained by unfolding the tensor along the associated axis and performing an SVD, e.g.,

$$\mathcal{T}_{i,j\ldots d} = U_{i,\alpha}\Sigma_{\alpha}V_{\alpha,j\ldots d} \tag{4.8}$$

Zero values in Σ can be dropped without approximation, revealing a subspace for the first index. After one has formed the U matrix for each index, "core tensor", $\mathcal{C}_{\alpha,\beta,\ldots,\gamma}$, can be formed by a simple change of basis. Equivalently, the U matrices can be considered as eigenvectors of the clusters reduced density matrix evaluated via,

$$\rho_{i,i'} = \mathcal{T}_{i,j,\ldots,d}\mathcal{T}_{i',j,\ldots,d} \tag{4.9}$$

In order to compute the Tucker decomposition of a TPSCI wavefunction, (here represented as a sparse tensor of TPS expansion coefficients) we also first construct the reduced density matrix, $\rho_{\alpha'\alpha}$, sequentially for each cluster. Computing the reduced density matrix for cluster, I, involves a contraction over all clusters $J \neq I$ needs to be performed.

$$\rho_{\alpha'\alpha} = c(\alpha', \beta, \ldots, \gamma)c(\alpha, \beta, \ldots, \gamma) \tag{4.10}$$

Fortunately, only a few of these coefficients are non-zero due to the sparsity of the TPSCI approach, making the computation efficient. Further, because we impose local quantum numbers $\hat{N}$ and $\hat{S}_z$, only the symmetry sublocks of the reduced density matrices are actually needed (with the rest being zero),

$$\rho_{\alpha'\alpha}^{N_{\uparrow}^1,N_{\downarrow}^1} = \mathsf{c}(\alpha'^{\mathtt{N}_{\uparrow}^1,\mathtt{N}_{\downarrow}^1}, \beta^{\mathtt{N}_{\uparrow}^2,\mathtt{N}_{\downarrow}^2}, \ldots)\mathsf{c}(\alpha^{\mathtt{N}_{\uparrow}^1,\mathtt{N}_{\downarrow}^1}, \beta^{\mathtt{N}_{\uparrow}^2,\mathtt{N}_{\downarrow}^2}, \ldots) \tag{4.11}$$

such that α and α' have the same number of up and down electrons. Once the cluster reduced density matrices are diagonalized producing new cluster states, we repeat the TPSCI calculation to get a new global vector. This process can be iterated until the density matrix stops changing, establishing a self-consistency condition. Although each iteration increases the computational cost, we find that the first iteration usually provides the most significant compression, with subsequent iterations only making smaller changes. Because of this, we have found it effective to perform just a single Tucker iteration, or to use a Tucker basis from a loose TPSCI calculation (i.e., large ϵ) as the basis for tighter TPSCI calculations in a "bootstrapping" fashion (this is explained further in the Supplementary Information). We have found that this works quite well for obtaining a compact basis at reasonable cost. This procedure is illustrated in Fig. 4.2.

If the full self-consistent Tucker solution is desired, we find that the convergence is often quite quick such that the energy stops changing significantly after only a few iterations. However, it is conceivable that this highly non-linear optimization could become a problem. In our previous work,[136] we found that a DIIS accelerated procedure which simultaneously extrapolates each cluster RDM greatly improved convergence in challenging cases. The same strategy could be applied here if needed.

4.2.3 Implementation Details

Our current code is written in Python, using NumPy for the computation of the matrix elements. Our code uses PySCF[209] for performing RHF calculations and computing the one- and two-electron integrals. As a pilot implementation, the performance of our code is far from optimal, with many opportunities existing for optimization. However, despite this, we are still able to perform rather non-trivial calculations on large active spaces, sometimes

with higher accuracy than we were able to achieve with HCI or ASCI. We expect that an implementation in a lower-level language like `C++` will increase the method's performance considerably.

One aspect of this work which is likely to impact performance is the technique chosen for storing the wavefunction. Since our states are sparse without any predictable structure in the indices, we simply choose to use a hash table to store the configurations. This is done using Python's `OrderedDict` implementation. However, because we enforce local symmetries ($\hat{N}$ and $\hat{S}_z$), we use a nested hash table. By specifying a Fock space configuration over N clusters as an immutable tuple-of-tuples $\texttt{((N}^1_\uparrow\texttt{,N}^1_\downarrow\texttt{), (N}^2_\uparrow\texttt{,N}^2_\downarrow\texttt{), ...,(N}^N_\uparrow\texttt{,N}^N_\downarrow\texttt{))}$ all TPS states with the same distribution of particle numbers can be described with a tuple of state indices: $(\alpha, \beta, \ldots, \gamma)$. Consequently, an arbitrary TPS expansion coefficient can be accessed by two sequential hash table lookups:

$$\texttt{TBL2} = \texttt{TBL1}[((\texttt{N}^1_\uparrow, \texttt{N}^1_\downarrow), (\texttt{N}^2_\uparrow, \texttt{N}^2_\downarrow), \ldots, (\texttt{N}^\texttt{n}_\uparrow, \texttt{N}^\texttt{n}_\downarrow))] \tag{4.12}$$

$$\texttt{c}(\alpha^{\texttt{N}^1_\uparrow,\texttt{N}^1_\downarrow}, \beta^{\texttt{N}^2_\uparrow,\texttt{N}^2_\downarrow}, \ldots, \gamma^{\texttt{N}^\texttt{n}_\uparrow,\texttt{N}^\texttt{n}_\downarrow}) = \texttt{TBL2}[\alpha, \beta, \ldots, \gamma] \tag{4.13}$$

This approach seems quite appropriate for dealing with variational degrees of freedom, of which there are typically only a modest number (i.e., fewer than around 100k). However, when computing the first-order perturbative correction to the wavefunction, this number of configurations can create hash tables too large to store in memory. We have addressed this problem in the short term, by simply (but at significant CPU work) pruning small values before adding them to our hash table, using ϵ_s from Eq. 4.2. However, a better solution would be to devise a low-memory PT2 correction for a TPS basis analogous to either the deterministic approach by Tubman and coworkers [210] or the semistochastic approaches by Sharma *et al.* [49] or Yann *et al.* [40] as implemented in the `Quantum Package` code [196].

In addition to the storage and manipulation of the state vector information, storing the operator tensors from Eq. 2.12 also requires some care. Our code is organized in a class-based structure, such that each cluster is an instance of a `Cluster` class, which owns all local operator tensor data, stored as dense arrays. However, because we have preserved local symmetries, we can reduce the storage by keeping only the operator transitions which are not symmetry forbidden. As such, in order to access an operator tensor, we again use a hash table to map specific Fock space transitions to operator tensors. For example, consider two states, $|\gamma\rangle_I$ and $|\delta\rangle_I$, which live in Fock spaces $(\mathtt{N}^{\mathtt{I}}_{\uparrow}, \mathtt{N}^{\mathtt{I}}_{\downarrow})$ and $(\mathtt{N}'^{\mathtt{I}}_{\uparrow}, \mathtt{N}'^{\mathtt{I}}_{\downarrow})$, respectively. The operator tensors associated with the $\hat{p}^\dagger \hat{q}^\dagger \hat{r}$ operators only have data available if symmetry allowed, i.e., if $\mathtt{N}^{\mathtt{I}}_{\uparrow} == \mathtt{N}'^{\mathtt{I}}_{\uparrow} + \mathtt{1}$ and $\mathtt{N}^{\mathtt{I}}_{\downarrow} == \mathtt{N}'^{\mathtt{I}}_{\downarrow}$. If that's satisfied, then the dense tensor can be retrieved by a hash table lookup taking in the local Fock space transition:

$$\Gamma^{\gamma\delta}_{pqr} = \mathtt{DATA}[\mathtt{string}(\mathtt{AAa})][(\mathtt{N}^{\mathtt{I}}_{\uparrow}, \mathtt{N}^{\mathtt{I}}_{\downarrow}, \mathtt{N}'^{\mathtt{I}}_{\uparrow}, \mathtt{N}'^{\mathtt{I}}_{\downarrow})] \tag{4.14}$$

where `string(AAa)` indicates a request for an operator string with three α operators (`A`=α spin, `B`=β spin), and the first two are creation (upper case) and the last is annihilation (lower case). This allows us to have a fine grained control over the tensor contractions used to form matrix elements, preventing the computation of any hard zeros. Because the Hamiltonian contains up to four-index quantities, a naive strategy would also store the full set of two-particle transition densities, $\Gamma^{\alpha,\alpha'}_{pqrs}$. However, because these terms can only contribute to local Hamiltonian terms, we can precontract these terms into a local Hamiltonian matrix, avoiding the need to store the 6 index quantities, leaving the memory bottleneck to be the 5 index terms: $\Gamma^{\alpha\alpha'}_{pqr}$. This memory bottleneck prevents us from considering exact clusters larger than 6 orbitals. The size of this tensor, $\Gamma^{\alpha\alpha'}_{pqr}$, is $\mathcal{O}(N^3 M M')$ where M is the number of states in the largest Fock space, M' is the number of states in the next largest Fock space with one electron different, and N is the number of orbitals in the cluster. Because

the number of states, M, increases factorially, it is difficult to store data for more than a few clusters having 6 orbitals. One way to reduce the memory demands is to truncate the number of cluster states. This can be done either by energy or by entanglement measures, as is outlined in the Appendix. Additional improvements can be made by manually handling the various tensor contractions. Currently these are handled in a rather abstract way which prevents much customization.

By the nature of the algorithm each non-trivial step is relatively easily parallelized. We have implemented the most expensive steps using shared memory parallelization, and have seen good scaling on the machines we've tested this on, systems with 24 or 32 cores. However, it would be relatively straightforward to parallelize over many nodes, and we plan on doing this in the near future.

4.2.4 Related works

There are several approaches in the literature which share the orbital clustering feature and TPS representation used in TPSCI. Perhaps the work most closely related to TPSCI is the Block Correlated Coupled Cluster (BCCC) approach of Li and coworkers.[137, 138, 211, 212, 213, 214] In BCCC, the orbitals are grouped into clusters and the wavefunction is represented in a TPS basis. Then, inter-cluster correlations are treated with an exponential parameterization, whose amplitudes are solved for non-linearly. The excitonically renormalized CC (XR-CC) also works in a TPS basis where the state-to-state interaction is solved for in a CC fashion.[100] Our method shares the TPS basis, but differs in both the treatment of intercluster correlation, and in the definition of the block states.

Another related cluster-based approach is the Renormalized Exciton Method (REM)[1] of

[1]One should be careful to not confuse this approach with the musical group from Athens, Georgia.

Malrieu and coworkers.[140, 141, 215, 216] In contrast to both TPSCI and BCCC, REM includes the intercluster interactions via a Bloch effective Hamiltonian. In terms of the implementation, our approach is most closely related to the Active Space Decomposition (ASD) of Shiozaki and coworkers.[217] The ASD method was extended to more than two clusters using a DMRG type wavefunction.[218, 219] One can view TPSCI as a generalization of ASD to arbitrary numbers of clusters, with the global state optimization being approximated with CIPSI rather than the exact subspace diagonalization used in their work. Again, the Tucker decomposition basis is another distinguishing aspect of our current work. Another approach which helped inspire our current work is the Cluster Mean Field (cMF) method of Scuseria and coworkers.[69] In fact, in most of the numerical calculations below, we use the fully optimized (both orbital and cluster state rotations) cMF as the reference TPS for TPSCI. The cMF method has also been extended to ab initio systems with the name variational localized active space self consistent field (vLASSCF).[220] cMF and TPSCI also share the orbital clustering, but TPSCI goes beyond the variational description of a single TPS and also defines the Cluster states through a Tucker decomposition. Other TPS methods include the rank-one basis proposed for molecular aggregates,[219] and the ab initio Frenkel Davydov Exciton Model (AIFDEM) of Herbert and coworkers for modeling the low-lying singly excited states of aggregates using monomer direct product basis.[142]

The TPSCI method also shares some features with other approaches which do not necessarily work with a TPS basis, but still involves some degree of orbital clustering. The ORMAS (occupation restricted multiple active space) method restricts the number of electrons in different orbital blocks and truncates the configuration expansion while still in the determinant basis.[89] The Restricted Active-Space (RAS) method [91] and Generalized Active-Space method are also approaches which involve orbital clustering, but the similarities essentially end there.[92] The multi-level DMRG (ML-DMRG) method proposed by Ma and coworkers

use chemical intuition based orbital ordering and parition the active space into high and low level subspaces with variable bond dimensions.[221]

4.3 Results and Discussion

The main goal of TPSCI is to make the calculation of large molecules possible when the number of determinants get intractable for determinant based selected CI. Hence one of the main focuses in this paper will be a comparison between the TPSCI method and Slater determinant based SCI to understand how clustering impacts the compactness of the representation. We compare mainly two aspects, the accuracy vs. final dimension of the variational space.

How to compare compactness of different methods? There is indeed some ambiguity in deciding how to compare TPSCI with determinant based SCI methods. Ultimately, it's not immediately clear what should be treated as a variable when counting degrees of freedom. On the one hand, since TPSCI involves diagonalizing local Hamiltonians to define the cluster basis, one might consider these local wavefunction coefficients as variables. Thus adding to the parameter count for TPSCI. Similarly, we could add orbital coefficients into the list of variables for both TPSCI and SCI. On the other hand, one might prefer to define variables to be parameters optimized by a global objective function (full system energy). In this paper we have chosen the latter definition, as this seems to be more consistent with the literature (e.g., basis set contraction coefficients aren't usually considered degrees of freedom in post-SCF calculations), and because it is more closely related to the computational cost (the initial cMF calculation is much faster than the resulting TPSCI). For TPSCI, the cluster state basis is a precomputed quantity, which is generally a trivial component of the calculation. While this can be updated via Tucker rotations, we generally don't optimize the cluster

states self-consistently. However, the cost profile would change considerably in the large cluster limit. Because we are using FCI solutions inside of a cluster (though one is not necessarily required to do this), the computational cost increases factorially with increasing cluster size. However, for the small clusters we've considered in this work (6 orbitals or less), the CMF calculation is a trivial component of the overall cost. As such, throughout the results section, we will make comparisons between different methods based on the number of degrees of freedom which are optimized to minimize the full molecule's energy. Thus the term "Dimension" will refer to the number of Slater determinants or TPS's.

In many of the results below, we use a simple convergence technique we refer to as "bootstrapping" which avoids going through larger dimensioned intermediate steps during the Tucker optimization. This approach is explained in the Supplementary Information.

In section 4.3.1 we study a simple modified Hubbard model which allows us to manually tune the impact of "clusterability" on the performance of TPSCI. Because SCI methods were unable to accurately model the Hubbard model, we use DMRG as a benchmark to compare the TPSCI results. The DMRG results were obtained using the `ITensor` library.[147] In section 4.3.2 we study smaller ab initio systems. We present data for N_2 bond dissociation curve with 6-31G basis, active space = (10e, 16o). We also present data for cc-pVDZ basis set results for N_2 and F_2 molecules at their equilibrium and stretched bond lengths, having active spaces of (10e, 26o) and (14e, 26o), respectively.

Finally in section 4.3.3 we study the ground state energies for a few π conjugated systems. The largest molecule in our test set is hexabenzocoronene, which has an active space of (42e, 42o). The geometries were optimized using B3LYP/cc-pVDZ level of theory, and the `xyz` files can be found in the Supplementary Information. The HCI data quoted for the Hubbard model, N_2 molecule and the π conjugated systems were obtained using the Arrow package.[41, 42, 49] The PT correction used for HCI is computed semistochastically (SHCI).[42] The ASCI

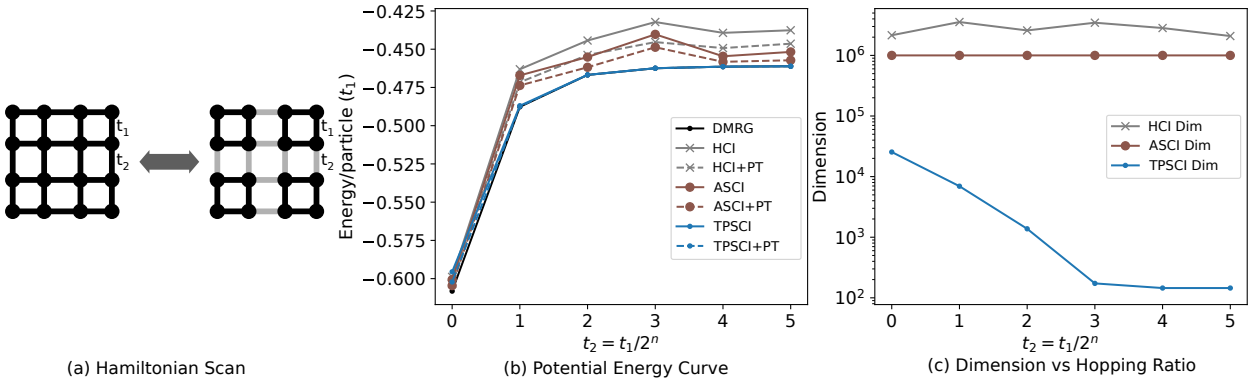


Figure 4.3: Clusterability of the Hubbard model. (a) Schematic representation of the Hubbard model used for the data, dark lines correspond to t_1 and lighter lines correspond to t_2. (b) Energy/site of the system as the $t_1 : t_2$ ratio is changed. TPSCI lines are nearly indistinguishable from the DMRG results. (c) Comparison of the dimension of the variational space as the $t_1 : t_2$ ratio is changed. DMRG result uses M=1600. The TPSCI calculations reported use (`ϵ=5e-8` ϵ_c`=1e-2` ϵ_s`=1e-7`) with MP PT correction.

data for the Hubbard model were generated using the `Q-Chem` package.[222] The integrals for all the molecular examples were computed using `PySCF`.[209]

4.3.1 Hubbard Model

Model Hamiltonians provide a useful tool for exploring the behavior of different approximate simulation techniques. In this section, we use the Hubbard model to explore how the inherent "clusterability" of a system of fermions affects the performance of TPSCI. All the Hubbard calculations are carried out at half-filling with no periodic boundary conditions.

Hubbard: Effect of Clusterability

Being motivated by the assumption that one can find some local structure in the Hamiltonian, the uniform Hubbard model is taken as our worst case scenario, and we expect our method to be inefficient in this domain. Since we are working in a TPS basis, the best case scenario

would be a Hamiltonian with no interactions between clusters. In such a case, the exact ground state becomes a single TPS. We expect a wide variety of physical systems to occur between these artificial limits. As such, we start in Fig. 4.3(a) by exploring the transition between uniform lattice to a highly clusterable lattice, by scanning the relative strength of the Hamiltonian coupling between clusters. The Hamiltonian used has two distinct hopping terms, and one electron-electron repulsion term:

$$\begin{aligned}\hat{H} = &\sum_{I} \sum_{\langle i,j\in I\rangle\sigma} -t_1 c^\dagger_{i\sigma} c_{j\sigma} + \sum_{IJ} \sum_{\langle i\in I, j\in J\rangle\sigma} -t_2 c^\dagger_{i\sigma} c_{j\sigma} \\ &+ U \sum_j n_{i\uparrow} n_{j\downarrow} \end{aligned} \tag{4.15}$$

where t_1 (t_2) denote hopping within (between) clusters, and U is the Coulomb repulsion. In order to make this a strongly correlated system, we set $U = 5t_1$ and start with a uniform lattice, $t_1 = t_2 = 1$. We then change the magnitude of of the inter-cluster hopping, t_2, scaled as $\frac{t_1}{2^n}$ where n varies from 1 to 5.

In Fig. 4.3, we observe confirmation that the accuracy and compactness of TPSCI should increase with increasing clusterabilty. We compare TPSCI with two different determinant based SCI methods, the HCI and ASCI methods. For this strongly correlated system, Slater determinant-based SCI methods were not able to find accurate results for any point on this scan, with reasonable numbers of variational parameters except for the uniform lattice. Even with the uniform lattice, it can be seen that TPSCI has 2 orders of magnitude less configurations as compared to ASCI or HCI. For the ASCI method, 1 million determinants were included in the variational space while the HCI results were computed using $\epsilon_1 =$ `5e-4` and $\epsilon_2 =$ `1e-9`. For both ASCI and HCI, we present data with the HF basis since it gave better results compared to the local basis. From these calculations it can be clearly seen that the traditional SCI methods would not scale well for larger systems with clusterable

property.

At higher ratios of $t_1 : t_2$, the TPSCI results are almost exact with a variational space of less than a few thousand configurations. This is a result of the fact that the exact ground state is moving increasingly close to a single TPS. Hence in the single particle basis, the representation is not really sparse and far more determinants than computationally feasible might need to be included for such an example. For large U/t ratio, the single particle basis would be even worse. TPSCI on the other hand does not depend on this and hence can be used as a good alternative.

Hubbard: Effect of Lattice Size

In this second example, we explore how accuracy and dimension of the variational space changes when increasing the size of the system. To do this, we fix the t_1/t_2 ratio to $\frac{1}{8}$. Consistent with the previous section, we set $U = 5t_1$. We start with the 16-site problem from above, but now increase the system size from 16, to 36, to 64 site lattices, all at half-filling and antiferromagnetic. Based on the performance of SCI on the 16 site problem, we did not attempt to compute the SCI energies for these larger lattices. Also consistent with the above section, the TPSCI calculations uses a clustering in which all t_1-coupled sites form a cluster. As such, the three different systems have 4, 9, and 16 clusters, respectively. This is shown in Figure ??.

We use both orbital optimized and frozen orbital cMF references for the TPSCI calculation. We denote the frozen orbital (orbital optimized) version as TPSCI (TPSCI') and the corresponding reference as cMF (cMF'). We compare both TPSCI and TPSCI' results with DMRG values with a fixed M value of 1600. We also plot the reference TPS (cMF and cMF') energy for these systems for comparison. Despite these being 2D systems, DMRG

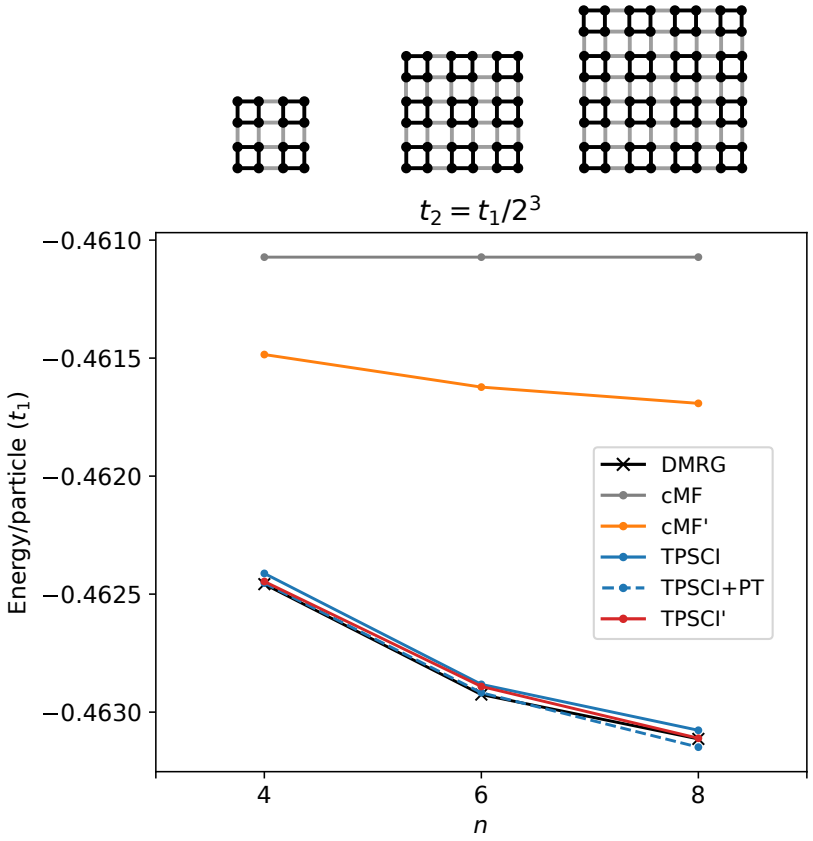


Figure 4.4: Size dependence of the Hubbard model is studied where we plot Energy/particle vs. lattice size. The intra-cluster:inter-cluster hopping ratio is fixed at $1/2^3$. For CMF (CMF') orbitals are frozen (optimized). The TPSCI (TPSCI') dimensions for each data points are: 16-site: 173 (429), 36-site: 1073 (1205) , 64-site: 1978 (2735) The DMRG calculations were carried out with a 'snake-like' path to keep stronger interactions more local.[3]

works quite well, although for the larger lattices (especially the 64 site lattice), the accuracy of TPSCI approaches that of DMRG. We note though, that it's difficult to achieve a fair comparison of TPSCI and DMRG, as both can, in principle, be systematically improved to get arbitrarily accurate results. Nonetheless, for the 64-site example, the variational energy for TPSCI' (using 2735 variables) is comparable to DMRG with M=1600. The Hubbard model in the frozen basis is extremely sparse and PT correction for this 64-site could be easily computed, since the repulsion is diagonal. For the orbital optimized version this is not the case and hence we have only included the variational energy correction.

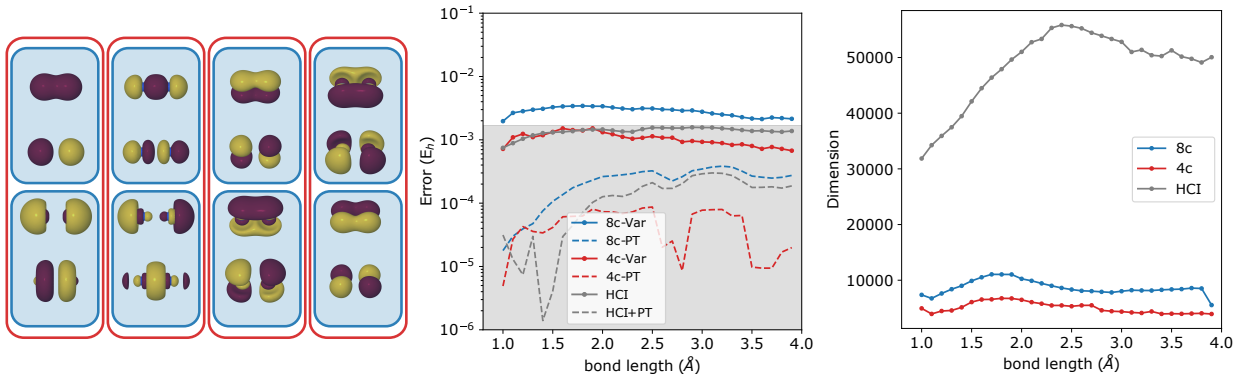
One challenge arises when studying the Hubbard model with different Hamiltonian parameters. Because the Hamiltonian enters into the selection criterion for the TPSCI method (via the first order amplitudes), we find that the TPSCI threshold value ϵ does not yield consistent convergence behavior. This means that the accuracy of the method cannot be directly linked to the selection criteria when modeling different Hamiltonians. While we only notice this problem with the Hubbard Hamiltonian, it is something we plan to investigate more in the future. One strategy would be to develop a TPSCI version of the Λ-CI method of Evangelista for growing the $\mathcal{P}$ space,[201] which is designed to have better accuracy guaranteed.

4.3.2 Molecular diatomics

Nitrogen dissociation with 6-31G

While model Hamiltonians are useful for artificially exploring the behavior of an approximation, the ultimate goal of our work is to produce an efficient method for ab initio molecular modeling. To understand the convergence behavior for molecular electronic structure, we start with a canonical example of a small strongly correlated system: N_2 dissociation.

For the Hubbard model it was straightforward to form the clusters based on sparsity of the



(a) Orbital Clustering (b) Error vs Bond Length (c) Dimension vs Bond Length

Figure 4.5: Nitrogen molecule with clustering based on bonding patterns. (a) The molecular orbitals for N_2 and the clustering choices (4c) four clusters and (8c) eight clusters (b) Error with CAS-CI results for TPSCI method with the two different clustering options and HCI method. Grey area denotes regions with chemical accuracy, 1 kcal/mol. (c) dimension of the variational space along the PES scan The TPSCI calculations reported use (`$\epsilon$=5e-8 $\epsilon_c$=1e-3 $\epsilon_s$=1e-7`) with EN perturbative correction. The HCI calculations reported use (`$\epsilon_1$=2e-4 $\epsilon_2$=1e-9`) with a semistochastic EN perturbative correction.

hopping term. In contrast, for small molecular systems the clustering is less straightforward. For a diatomic system like N_2 molecule, traditional single reference methods like CCSD provide good results at shorter bond lengths, but fail when the bond is stretched. Six orbitals in the N_2 molecule become degenerate when bond is stretched and therefore RHF reference is not good enough to represent the molecule.

In this section, we study the N_2 example with frozen 1s orbitals. Even though N_2 is a triple bonded system, the interaction between these bonds are not as strong as the bond itself. Hence, we can partition the orbitals in the triple bond (6e, 6o) into three (2e, 2o) clusters. This would mean putting the σ and σ^* bonds in a cluster, and the two π and π^* bonds in separate clusters. We are left with the lone pairs and we leave them in a separate cluster. This is similar to a perfect pairing type cluster since each cluster is a pair of bonding/antibonding orbitals.[223, 224] Hence the clustering would look like this in a minimal basis: (σ_s, σ_s^*),$(\sigma_{pz}, \sigma_{pz}^*)$, (π_{px}, π_{px}^*), (π_{py}, π_{py}^*)

In the 6-31G basis, there are extra 3s and 3p basis functions on top of the minimal basis. Hence we have a total of 16 orbitals. We can put the extra orbitals in bonding/antibonding pair clusters, similar to a perfect pairing type clustering. This makes a total of 8 clusters. We can also combine the orbitals of same angular momentum but different principle quantum numbers as one cluster. This would lead to a clustering with 4 orbitals per cluster. We refer to these two types of clusterings as 8c and 4c:

- 8c:(2s), (3s), ($2p_z$), ($3p_z$), ($2p_x$), ($3p_x$), ($2p_y$), ($3p_y$)
- 4c:(2s, 3s), ($2p_z$, $3p_z$), ($2p_x$, $3p_x$), ($2p_y$, $3p_y$)

By defining the clusters in this way in 4c, dynamic correlation for each bond is included within a cluster.

We compare these two clustering choices against HCI with ϵ_1=`2e-4` (variational part) and ϵ_2=`1e-9`. For TPSCI, we use cMF with frozen orbitals as the reference, with a selection threshold of ϵ=`5e-8`,[2] and with EN perturbative correction. The search space for each iteration was defined with ϵ_c=`1e-3` and a screening of the $\mathcal{Q}$ space couplings was set to ϵ_s=`1e-7`. We have found that setting the search threshold $\epsilon_s \leq$`1e-7` consistently provides sub-mH accuracy.

The molecular orbital clusterings, 4c and 8c, are pictorially depicted in Figure 4.5(a). In Figure 4.5(b), we show the error with respect to FCI results. The region of chemical accuracy is marked by 1 kcal/mol. As seen from the figure 4.5(b), the 8c results are not as accurate as the HCI or 4c results. These calculations can be made more accurate by using a lower threshold, but for clarity we chose to show data using same threshold for both 4c and 8c data. The PT correction of the determinant based HCI is better at lower energies, and

[2]since TPSCI selects on the probability and HCI selects on the magnitude, these thresholds are related by a square root

Table 4.1: Diatomic systems: N_2 (10e, 26o) and F_2 (14e, 26o) with cc-pVDZ basis. We compare both HCI and ASCI results with TPSCI. We also provide results with natural orbitals for ASCI calculations. The TPSCI calculations reported use (ϵ=5e-8 ϵ_c=5e-3 ϵ_s=1e-7). The HCI calculation use (ϵ_1=5e-4 ϵ_2=1e-9) and all the ASCI calculations used a variational space of 50,000 determinants, except for the 2r calculation with natural orbitals where the 50k calculation converged to an excited state so we report results from a 100k calculation. *ASCI calculation converged to the nearly degenerate quintet state.

N_2	Variational	PT2	Dim	Variational	PT2	Dim	Variational	PT2	Dim
r=1.0977		r			2r			3r	
HCI	-109.2692	-109.2769	37,577	-108.9571	-108.9668	53,028	-108.9477	-108.9564	42,782
TPSCI	-109.2694	-109.2769	8,274	-108.9607	-108.9674	15,659	-108.9522	-108.9572	12,744
ASCI	-109.2723	-109.2770	50,000	-108.9603	-108.9673	50,000	-108.9515	-108.9570	50,000
ASCI-no	-109.2738	-109.2771	50,000	-108.9641	-108.9677	100,000	-108.9524*	-108.9568	50,000
F_2	Variational	PT2	Dim	Variational	PT2	Dim	Variational	PT2	Dim
r=1.4119		r			2r			3r	
HCI	-199.0913	-199.0992	68,994	-199.0489	-199.0554	67,434	-199.0489	-199.0549	65,476
TPSCI	-199.0911	-199.0991	6,225	-199.0501	-199.0556	3,694	-199.0498	-199.0551	3,956
ASCI	-199.0923	-199.0993	50,000	-199.0498	-199.0556	50,000	-199.0499	-199.0550	50,000
ASCI-no	-199.0937	-199.0994	50,000	-199.0500	-199.0556	50,000	-199.0496	-199.0551	50,000

gets worse for the stretched geometries where the system is more strongly correlated. The variational dimension for the HCI method increases as the bond is dissociated (more strong correlation). This is in contrast to relatively constant dimension for TPSCI along the PES.

Diatomics with cc-pVDZ basis

In order to observe the impact of dynamical correlation, we have also performed TPSCI calculations for the larger basis-set, cc-pVDZ. As in the previous subsection, we cluster with bonding/antibonding pairs. We consider three different bond lengths for both N_2 and F_2.

While we follow a similar clustering pattern as the 6-31G results above, in the cc-pVDZ basis, extra d-shell orbitals are present leading to more clusters. To keep the size of the clusters small, for the present paper, we partition the d-orbitals such that the atomic pairs coming from d_{z^2}, d_{xz} and d_{xy} each form cluster. We leave the two d_{xy} and $d_{x^2-y^2}$ as a 4-orbital cluster. Taken together, this creates a total of 8 clusters for both N_2 and F_2. By adding higher principle quantum number orbitals into the cluster, we are effectively allowing the clusters to become dynamically correlated. The clustering pattern is described below,

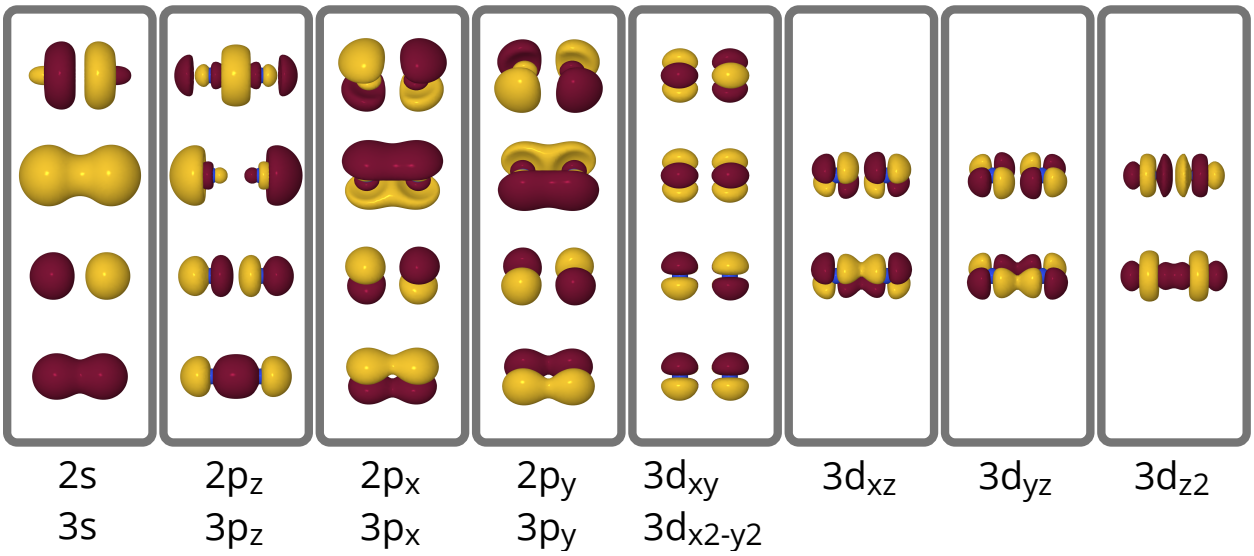


Figure 4.6: Choice of active-space orbital clustering for cc-pVDZ basis sets for N_2 and F_2

keeping in mind that each orbital listed represents two orbitals, one from each atomic center.

- (2s, 3s), ($2p_z$, $3p_z$), ($2p_x$, $3p_x$), ($2p_y$, $3p_y$), ($3d_{xy}$, $3d_{x^2-y^2}$), ($3d_{xz}$), ($3d_{yz}$), ($3d_{z^2}$)

This clustering is shown for N_2 in Fig. 4.6 In Table 4.1, we present data for three bond lengths r, 2r, 3r where r=1.0977 (1.4119) for N_2 (F_2). The thresholds for all the TPSCI calculations were (ϵ=`5e-8`, ϵ_c=`5e-3`, ϵ_s=`1e-7`), and the frozen cMF reference state was used. The HCI calculation were computed using ϵ_1 = `5e-4`, while the ASCI calculations were performed using 50K determinants. The most immediate conclusion, is that TPSCI converges to chemical accuracy with a smaller dimension than the determinant based methods. It has been previously shown that the use of natural orbitals can improve determinant based selected CI algorithms.[199, 205] By using natural orbitals with ASCI, we observe similar behavior, although even with natural orbitals, the wavefunction in determinant based SCI is not as compact as the TPSCI wavefunction. The use of natural orbitals seems to have the largest effect near equilibrium bond distances, with a smaller dependence on orbitals occuring at stretched distances.

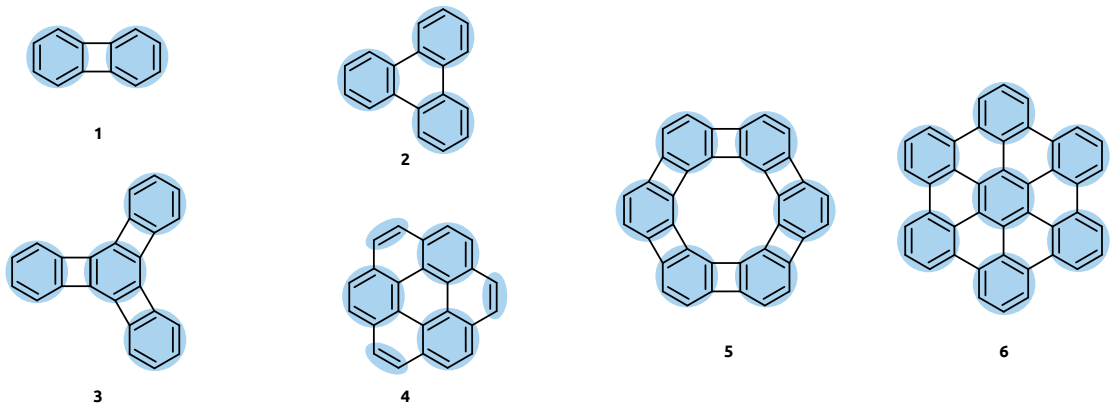


Figure 4.7: π-conjugated systems for TPSCI calculation with blue circles representing the clusters for the system.

4.3.3 π-conjugated systems

While the small diatomic molecules of the previous section provide a rigorous test of the TPSCI method for systems which are well described by existing SCI methods like HCI and ASCI, our goal in designing TPSCI is to model active spaces which are larger and more strongly correlated than what can be afforded with the traditional Slater determinant-based approaches. For this reason, we now turn our attention to systems which are expected to be good applications of TPSCI, systems which are non-linear (so a MPS is non-ideal), strongly correlated, and somewhat clusterable. Taking poly-aromatic hydrocarbons (PAH's) as test cases, might perhaps be surprising because the characteristic delocalization of the π system might seem to suggest exactly the worst case scenario for observing "clusterability". However, the success of Clar's rule in relating the number of disjoint benzene units with stability, seems to suggest that a clustering pattern chosen from Clar's rule might be effective. A suggestion that is consistent with recent results using fragment-based DFT methods.[165]

We have chosen a few example PAH's (shown in Fig. 4.7) ranging in size and clusterability, with the clustering pattern consistent with Clar's rule shown in blue. In addition to the well-

Table 4.2: Comparison between TPSCI and HCI for the π-conjugated systems used in the study. The molecules are labelled according to Figure 4.7. The TPSCI calculations reported use (`$\epsilon$=1e-7 $\epsilon_c$=1e-2 $\epsilon_s$=1e-6`). The HCI calculations use (`$\epsilon_1$=1e-5 $\epsilon_2$=1e-9`) for molecules **1**-**4**. For **5** (`$\epsilon_1$=3e-5 $\epsilon_2$=1e-9`) and for **6** (`$\epsilon_1$=4e-5 $\epsilon_2$=1e-9`).

molecule	Variational		PT2		Extrapolated		Dim	
	TPSCI	HCI	TPSCI	HCI	TPSCI	HCI	TPSCI	HCI
1 (12e, 12o)	-453.6310	-453.6310	-453.6310	-453.6310	-453.6310	-453.6310	201	174,757
2 (18e, 18o)	-680.5951	-680.5906	-680.5958	-680.5944	-680.5964	-680.5970	2,440	7,397,514
3 (24e, 24o)	-904.9121	-904.8865	-904.9136	-904.9012	-904.9142	-904.9144	3,885	21,179,338
4 (24e, 24o)	-905.2157	-905.2100	-905.2238	-905.2212	-905.2307	-905.2307	16,272	19,510,272
5 (36e, 36o)	-1353.8138	-1353.6927	-1353.8199	-1353.7509	-1353.8244	-1353.8259	10,376	20,232,920
6 (42e, 42o)	-1582.4291	-1582.2378	-1582.4402	-1582.3255	-1582.4482	-1582.4396	20,325	11,194,996

known coronene-type bonding pattern (**2, 4, 6**), we have also considered a few systems which contain rather strained 4-center ring bridging units (**1, 3, 5**). Though less stable, materials based on this bonding pattern have recently been synthesized using polymerization reactions of 1,3,5-trihydroxybenzene molecules.[12] These material can have interesting applications because of its planar structure like graphene and porous nature. For the sake of simplicity, we use the minimal STO-3G basis for these systems, since we are only studying the π-conjugated electrons.[3] The active space for molecules **1**-**6** consist of 12, 18, 24, 24, 36, 42 orbitals, respectively.

To initialize the clusters, we started with localized orbitals and then performed a cMF calculation to obtain a cluster basis, and optimal orbitals. We found that the orbital optimization (only mixing orbitals within the active space and between clusters) lowered the energy a substantial amount compared to the frozen cMF (-0.38 au in the case of **6**). To minimize the memory requirements, we used the embedded Schmidt trunction (EST) approach described in the appendix, for each of these systems discarding Schmidt vectors with singular values smaller than `1e-4`. This value was found to provide large memory (and cpu) savings, without significantly impacting the accuracy, with our tests indicating that the error was below 1mH. For each of these systems, we then performed three increasingly tight TPSCI calculations defined by the following settings:

[3] We have also used the cc-pvdz basis set with the same size active space and found similar qualitative results, but with consistently smaller amount of active-space correlation.

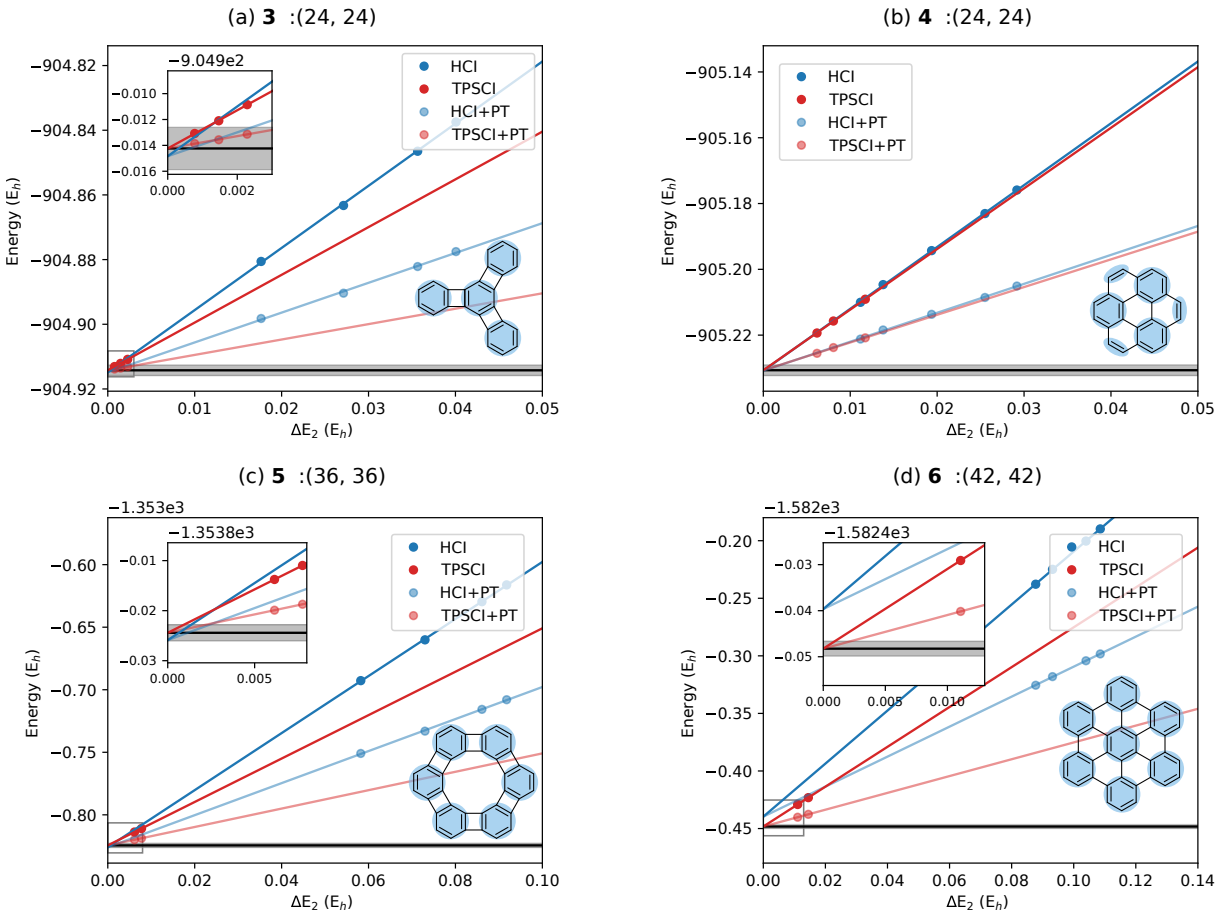


Figure 4.8: Extrapolation scheme for the four large molecules considered. The grey area corresponds to ± 1 kcal/mol about the extrapolated TPSCI energy (solid black line)

1. (ϵ=1e-6, ϵ_c=1e-2, ϵ_s=1e-6)

2. (ϵ=1e-7, ϵ_c=1e-2, ϵ_s=1e-6)

3. (ϵ=1e-8, ϵ_c=1e-2, ϵ_s=1e-6)

each time performing an MP perturbative correction setting (ϵ_c=0, ϵ_s=1e-7). Only a single Tucker iteration was performed for the ϵ=1e-6 calculation, and then that basis was used for the more accurate calculations. For molecules **5** and **6**, we were unable to obtain the PT2 correction for the ϵ=1e-8 calculations.

For HCI, we used canonical HF orbitals, as we found that natural orbitals didn't have a significant affect for these systems. The HCI was computed for molecule **1**-**4** using (ϵ_1=1e-5, ϵ_2=1e-9). For molecule **5** and **6**, the variational space became too large, and thus the tightest data we could obtain for **5** and **6** was ϵ_1=3e-5 and ϵ_1=4e-5, respectively.

In Table 4.2, we present the most accurate HCI data, alongside TPSCI data using the intermediate threshold level: (ϵ=1e-7 ϵ_c=1e-2 ϵ_s=1e-6). For biphenylene, **1**, we can see that both HCI and TPSCI give essentially the exact result for this small active space of (12e, 12o). Nonetheless, even for this small system, the TPS basis is much more compact than a determinantal basis 201 vs. 174,757. However, this shouldn't be too surprising, because **1** only has two clusters, meaning the "Tucker" decomposition performed at ϵ=1e-6 is actually an SVD, which is formally the most compact representation. For larger systems though, this is not the case, and the Tucker decomposition no longer presents a diagonal representation.

Considering the larger molecules we continue to see a similar increase in compactness by around 3 orders of magnitude. Molecules **2** and **3** collect only a few thousand configurations while still being more accurate than the HCI results with millions of determinants (Table 4.2).

Clusterability Considering molecules **3** and **4** side-by-side, provides insight into the impact of clusterability. Both systems have the same active space size, (24e, 24o), but they differ in the connectivity of the clusters. Molecule **3** is able to be clustered into four 6-site clusters, a complete Clar's tiling. Molecule **4** on the other hand can only be grouped into three 6-site clusters, and three remaining 2-site clusters. This has a significant impact on the compactness, with **4** requiring over four times the number of TPS's and being further from the extrapolated result. Nonetheless, comparison with determinant based SCI is still impressive. With approximately 16 thousand TPS configurations, the TPSCI energy for **4** is significantly lower than the determinant based SCI result with 19 million determinants.

Extrapolation Because of the slow convergence of correlation energy with variational dimension, it is often useful to use a few SCI calculations and extrapolate to the exact result. We can also do this for TPSCI, provided each of the variational and PT2 energy pairs are computed using the same cluster basis. In Fig. 4.7, we use the available energies to extrapolate to the exact result, giving us both an estimate of the FCI energy, and also an estimate of how accurate our TPSCI calculations are. TPSCI extrapolations are performed using all 3 accuracy levels, except for molecules **5** and **6**, for which only the first two calculations were used to extrapolate. Comparing these extrapolations to the HCI extrapolations we observe that even though the TPSCI energies are much closer to converged, the HCI extrapolations (with the exception of the largest system **6**) are very effective and provide comparable FCI estimates. Overall, these systems indicate that the TPSCI method provides a unique representation able to more efficiently captures correlation in clusterable systems, extending the applicablity of SCI algorithms to larger systems.

4.4 Conclusion and Future Work

In this work, we have introduced a new selected CI method using tensor products of cluster states as the basis. By folding the most important correlations into the basis vectors themselves, much more compact wavefunctions can be obtained using the basic selected CI procedures, a feature which can significantly improve the performance for strongly correlated systems. In choosing the nature of the cluster states, we found that the Tucker decomposition provided a simple and efficient way to significantly improve the compactness of the final TPSCI wavefunction. Although our current code is far from optimized, we have demonstrated advantages over determinant based methods for large active spaces in PAH's. However, even if our implementation was sufficiently optimized, it's not obvious that TPSCI will provide a faster "time to solution" compared to methods like HCI or ASCI for small systems in large basis sets due to the lack of clusterability. The TPS representation involves quite a bit of computational overhead, which is only likely to pay off for spatially extended systems (like PAH's) which provide better opportunities for effective clustering.

This initial paper presents the algorithm, an implementation, and proof of concept results. However, much work remains to be done. A few of the ongoing and future works will involve:

1. Develop efficient and low-memory PT2 correction algorithm, similar to the deterministic[210] or semistochastic approach.[40, 49]

2. Investigating the behavior of different orbital clusterings. Good performance of TPSCI requires the sensible partitioning of the orbitals into clusters. In this work, we've largely done this by hand. However, this does not necessarily provide the best clustering. We plan on developing automated procedures for orbital clustering. Preliminary results suggest that partitioning clusters based on the exchange matrix seems to work quite well. However, more rigorous approaches based on information theory[225, 226, 227]

might provide improved results.

3. Exploring the performance of cluster basis truncation. We have explored two different approaches (energy based, and entanglement based), but it's not yet obvious what the best approach is, or how the technique depends on the chemical system.

4. Improved implementation. While our current code is efficient enough to obtain all the results in this manuscript, including the (42e, 42o) results and the 64 site Hubbard lattice example, many steps are far from optimal, with obvious "hot spots" occuring in pure Python functions. Using C++ to reimplement these steps should provide significant improvements.

5. Exploring how the different components of the various SCI methods such as HCI, ACI, ASCI, etc. behave in the TPS basis.

6. Extending the method to study excited states. One possible route would be to use similar approach taken in the original CIPSI work. Another interesting direction is to form a reference state using single excitations in clusters as done in AIFDEM[228] and include important configurations avoiding collapse to the ground state. TPSCI would be especially well suited for modeling excitations in molecular aggregates, due to the lack of covalent bonds between systems making the TPS representation converge extremely quickly.

4.5 Acknowledgements

The authors would like to thank Garnet Chan for helpful suggestions regarding DMRG calculations. The authors would also like to thank Anthony Scemama for helpful commments

about the manuscript. This research was supported by the National Science Foundation (Award No. 1752612).

4.6 Supporting Information

The "bootstrapping" approach is discussed in the supporting information. The geometries used for the PAH system is provided in `coordinates.txt` file. The raw data for all the systems are also included with the supporting information in file `raw_data.xlsx`.

4.7 Appendix

4.7.1 Truncation of cluster states using approximate Schmidt vectors

The TPSCI method described in the main text presents an algorithm for systematically approximating the exact solution within a basis of tensor products of many-body cluster states. This algorithm is quite general and can be used with cluster states obtained in a number of ways. The most straightforward approach would be to simply diagonalize the Hamiltonian operator which only acts locally on the cluster (i.e., all indices correspond to orbitals within the cluster). This has the convenience that local operators become diagonal, but lacks all interactions with other clusters. A better approach, would be to use the cluster mean-field (cMF) method developed by Jiménez-Hoyos and Scuseria.[69] Ignoring the orbital optimization component of cMF first, cMF is the variational minimization of a single tensor product state wavefunction. This amounts to defining each cluster's ground state as the lowest energy eigenstate of an effective local Hamiltonian, which includes the mean-field

interactions with all clusters. Because the energy of a single TPS only depends on each cluster's ground state, the higher energy cluster states aren't determined rigorously by a variational principle as the cMF energy has a rotational invariance among the non-ground cluster state (this is analogous to problems in trying to tie physical meaning to virtual orbitals in Hartree-Fock). However, the cMF local Hamiltonian is uniquely defined, and its associated eigenbasis (i.e., the "canonical" cluster state basis) provides an improved set of vectors for defining the cluster basis used in TPSCI.

For small clusters (e.g., less than about 5 spatial orbitals) the full cluster state basis can be used without any complications. However, for larger clusters, the memory requirements needed to store the local operator tensors, $\Gamma^{I_\alpha, I_{\alpha'}}_{p^\dagger qr}$, become significant. For example, the largest sector of Fock space for a six site cluster (3 α and 3 β electrons) has a dimension of 400. Storing the associated tensors requires about 2.5Gb per cluster: This doesn't include the numerous smaller particle number spaces. However, for a system that is clusterable one should be able to discard many of these states without significantly impacting the global ground state.

Using the eigenvalue of the cMF Hamiltonian is one way to determine which cluster states to discard. However, because local energies aren't good predictors of entanglement, rather large numbers of states are needed to maintain accurate results for global quantities. To address this issue, we have developed a relatively simple approach for defining more compact cluster states, which we refer to as "embedded Schmidt truncation" (EST). For cluster I, the goal is to obtain a compact set of vectors which captures as much of its entanglement with the rest of the system as possible. The ideal set of vectors would then be obtained by diagonalizing the cluster I's reduced density matrix obtained from the exact global system's ground state. To develop a practical approximation to this, we instead decide to find the ground state of a smaller system comprised of cluster I and a small number of bath orbitals

which directly interact with I, using a mean-field description of the remaining system. This approach is based on the "Concentric Localization" concept recently used in projection based embedding,[229] and also density matrix embedding theory.[230]

In order to identify the bath orbitals for cluster I, we simply SVD the off-diagonal block of the exchange matrix between orbitals in cluster I and orbitals in clusters $J \neq I$.

$$K^I_{pq} = \sum_{LM} \sum_{r\in L} \sum_{s\in M} (pr|sq) D_{rs} \qquad \forall p \in I, q \notin I \tag{4.16}$$

$$= \sum_{s} U^I_{p,s} \lambda_s V^I_{q,s} \tag{4.17}$$

where D_{rs} is the one-particle reduced density matrix resulting from the cMF calculation. The bath orbitals for cluster I are defined to be those which directly interact with cluster I via the exchange operator, and are trivially obtained by rotating the MO coefficients, $C_{\mu,p}$, by the right singular vectors:

$$C^{\text{bath},I}_{\mu,s} = \sum_{J\neq I} \sum_{q\in J} C_{\mu,q} V^I_{q,s} \tag{4.18}$$

This creates a natural compression of exchange interactions, and the number of non-zero singular values, λ_s, is bounded by the number of orbitals in cluster I. The molecular orbitals are now organized into three subspaces,

$$\mathbf{C}^I = \mathbf{C}^{\text{cluster},I} | \mathbf{C}^{\text{bath},I} | \mathbf{C}^{\text{env},I}, \tag{4.19}$$

where $C^{\text{env},I}$ are the orbitals associated with the null space of the exchange coupling. This procedure can be continued recursively to define another bath, which completely captures the interaction between the previous bath and the environment, creating a recursive approach to organizing all the orbitals by their "nearness" to cluster I. In this work, we only consider the

first layer bath. The goal is now to obtain the exact solution to the combined cluster+bath system and then SVD the resulting ground state to define a basis of many body states to use for cluster I. Because the bath orbitals are defined via the SVD, the number of bath orbitals will always be less than or equal to the number of orbitals in that cluster. For instance, if a cluster has 6 orbitals, then to obtain the basis, one would need to compute only the ground state for a 12 orbital problem. For larger clusters, this will quickly become expensive, but since only the ground state is needed, conventional selected CI algorithms like ASCI or SCI could be used alternatively for this step.

If the environment orbitals were unentangled with the cluster and bath orbitals, then the 1RDM in the environment space would be idempotent and we could simply perform a CASCI calculation to obtain this ground state. Because this is not the case generally, we simply purify the density in the environment space, and use this background density for the CASCI core. Once the ground state of the embedded cluster|bath system is obtained, We organize the resulting CI vector into contributions to local particle number spaces and then perform an SVD. This provides us with a set of vectors with well defined particle numbers for cluster I, ordered according to their weight in the embedded ground state. This process is repeated separately for each cluster.

As described so far, this is simply a change of basis for the cluster states on I. No approximation has been made, and the final TPSCI wavefunction can still converge to the exact ground state. However, because our vectors are now weighted according to an approximate entanglement metric instead of energy, we can now perform a more aggressive truncation on the number of cluster states used to determine the dimension of the Hilbert space accessible to the TPSCI algorithm. In practice, this works extremely well when the system is well localized (e.g., polyaromatic hydrocarbons), and deleting cluster states with singular values smaller than .0001 seems to consistently have sub milliHartee impact on the

results. Of course, for some systems this might not be ideal, and alternatives might need to be considered. This will be a focus of future work.

Chapter 5

Excited States using TPSCI

5.1 Introduction

Electronic excited states play an important role in a lot of biological processes and mimicking them for applications in solar cells, artificial photosynthesis is of great interest. Theoretical methods play a central role in this and can help experiments. Traditional theoretical methods that depend on a single excitation like TDDFT[231, 232, 233] fail miserably when studying doubly excited states[234, 235] and charge transfer states.[236, 237, 238] Even single reference CCSD methods like EOM-CCSD [239, 240] fail for doubly excited states with around 1 eV error. CASSCF[25] related methods like CASPT2[26, 27], MRCI[28, 29] are more accurate alternatives. However these methods cannot be used for active spaces larger than 18 orbitals in 18 electrons. It is also very difficult to select active orbitals for state averaging when the ground and excited states differ too much in dipole moment, seen usually in cases with charge transfer excitation.

Selected configuration interaction (SCI) based approaches have been recently used to calculate accurate estimates for vertical excitation[195, 205, 241], doubly excitation[242], charge transfer[243] states. Even though these are for small to medium molecules, the variational space for these wavefunctions are in tens of millions. Hence the wavefunction will scale combinatorially for larger and more strongly correlated systems and hence will become intractable for SCI based approaches. The tensor product selected configuration interaction

discussed in Chapter 4 provides accurate results for ground states with very compact variational dimension. In this work, we investigate the TPSCI approach for excited states focusing mainly on π conjugated systems. We also present an intuitive basis for multi-exciton states.[244, 245, 246, 247]

5.2 Theory

The ground state can be approximated as the single tensor product state where all the clusters are in their ground state. This configuration can be optimized using the cMF procedure where the cluster states and orbitals are optimized simultaneously. If the system is fully decoupled, then the vertical excited state wavefunction can be represented using a diabatic basis where higher excited states in each cluster is used and the wavefunction can be formed as a linear combination of the two.

The singly excited configuration for a given cluster can be written as:

$$|\psi_{\lambda_L}\rangle = |0_I, 0_J, ..\lambda_L, ..0_N\rangle \tag{5.1}$$

In the tensor product framework, we can introduce a TPS-CIS type wavefunction where all the singly excited configurations can be included. These energy values will not generally be accurate since there is no relaxation achieved in the TPS-CIS space by higher excited configurations. Similarly doubly excited manifold also should be similarly possible in the TPS framework. In case of active space decomposition (ASD)[218] and the rank-1 matrix product state approach[219], very few excited cluster states are included to form the excited states. When studying disjoint chromophore assemblies and focussing on few low energy states, this is not a drastic approximation. This approximation becomes an issue when we

study systems which can be more entangled. We investigate this in Section 5.3.1. For the multi-state TPSCI algorithm, we can form a set of guess vectors (**P** space) either from a TPS-CIS calculation or by initializing important few configurations.

Algorithm:

1. Start from a reference **P** space.
2. Diagonalize the Hamiltonian in **P** space and update the configurations.
3. Apply the Hamiltonian on each state in the reference space **P** to form the first order interacting space (FOIS) to form **Q**.
4. Prune the **Q** based on the selection criteria ϵ. If the first order coefficient of any TPS configuration is greater than ϵ, we include the configuration to the **P** space ($|c_i^{(1)}| > \epsilon$) .
5. Check for convergence. If not converged, go back to Item 2.
6. Finally compute a state specific PT2 correction for each state.

5.3 Results and Discussions

In the following section, we investigate the efficacy of TPSCI by using it to study a several π conjugated systems or polycyclic aromatic hydrocarbons (PAHS) and compute the low lying triplet states for these systems. PAH molecules have been of very high interest recently in terms of novel material in synthesizing chiral nanographenes,[248] twisted carbon nanobelts,[249], carbon-based electronic devices [11] etc. We study how the TPSCI and traditional TPS-based methods like ASD compare to each other in a dimer case in Section

5.3.1. We also investigate the effect of charged clusters by studying the acepleiadylene systems in Section 5.3.2 and comparing it to a similar sized neutral clustering in Pyrene. We also investigate larger π conjugated systems in Section 5.3.3. We use geometries optimized at the B3LYP/cc-pVDZ level of theory for the π conjugated systems. The active space is generated using the localized $2p_z$ orbitals of an RHF calculation using cc-pVDZ basis.

5.3.1 Benzene Dimer vs Biphenylene

In this Section, we use two simple examples, a benzene dimer (A) and a biphenylene (B) molecule to show how a cluster based many body methods like TPSCI and ASD compare with each other. Biphenylene molecule has been recently used for synthesizing 2D carbon material.[11, 12, 13, 250] and application in single molecule junctions [251]. Both these systems have similar active space with 12 electrons in 12 π orbitals. Since this active space only contains 853776 ($^{12}C_6 \times {}^{12}C_6$) configurations, the exact CASCI calculation can be performed, allowing us to make a direct comparison for evaluating the performance of each approximate method. Being a dimeric system, both of these systems can also be partitioned into two clusters with 6 electrons in 6 orbitals each. For the biphenylene molecule, the 12 orbitals can also be partitioned using the Clar's structure into two clusters. Both these systems have similar clustering and orbitals but very different interactions between them. The ASD method has been successfully used in studying excitation energies of disjoint molecular systems.[72, 73] In this approach the number of many electron states in a cluster is truncated to M states and the full tensor product basis is formed. The ground and excited states are then formed in this basis. In TPSCI, we employ a selected CI approach and avoid forming the full tensor product basis. This helps in cutting down the configurations quite drastically and also apply this to more than two cluster systems.

We compute the excitation energy for the lowest three triplet states using ASD as well as TPSCI and compare with the exact CASCI results for these system. For the ASD calculation, the M value was fixed at 20 states per cluster for each possible set of Fock space configurations. For the TPSCI calculation, the M value was fixed at 100 states per cluster. Doing a CASCI in the tensor product of 100 states per cluster would be much more expensive. For the benzene dimer, as presented in Figure 5.1(a), the errors in excitation energies are negligible for ASD as well as variational TPSCI and TPSCI with a PT2 correction. The variational dimension of the ASD is around 13,000 configurations while the TPSCI wavefunction only has 540 configurations. The biphenylene system (Figure 5.1(b)) is far more entangled and the errors for the ASD approach compared to the CASCI results are more than one kcal/mol. Using TPSCI, the error is below chemical accuracy for both variational and PT2-corrected versions. The wavefunction dimension for ASD is the same for this molecule as the benzene dimer since the M value is same for both. The TPSCI variational dimension is 5142 configurations for the biphenylene but is still lower than the ASD approach and gives much more accurate results. In Figure 5.1(c) we plot the variational dimension for the TPSCI calculation of the two systems considered alongside the variational dimensions of the ASD and CASCI approaches. Since the two clusters in the biphenylene system interact closely, the ASD approach will need larger M values to get to accurate results. In TPSCI we already have a large M value and only selected TPS configurations are added to the wavefunction, making the final wavefunction compact. Therefore, TPSCI provides a framework where the important configurations for the wavefunction are automatically selected from the large tensor product basis.

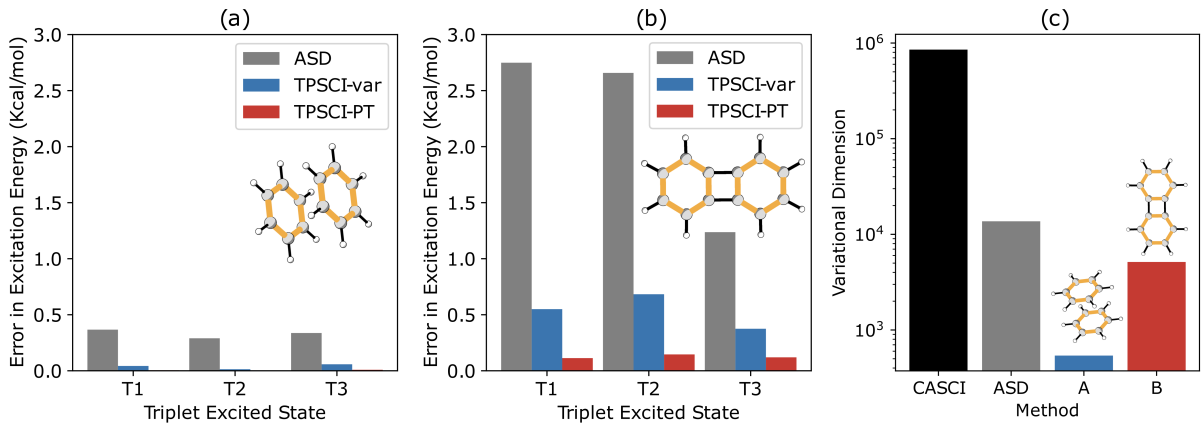


Figure 5.1: Excitation energies for (a) the benzene dimer and (b) the biphenylene molecule and (c) the variational wavefunction dimension for CASCI, ASD with both systems using TPSCI.

5.3.2 Effect of cluster type: Pyrene vs Acepleiadylene

Acepleiadylene (APD) is a promising precursor for generating nonbenzoid nanographene sheets.[252] The APD molecule is a structural isomer of the pyrene molecule but has a non zero dipole moment due to the resonance structure in its ground state. APD is similar to azulene in the resonance structure where the seven member ring has positive and the five member ring has negative charge density.Here we intend to do a comparative study of the excited states of Pyrene and the APD molecule (Figure **??**).

For the pyrene molecule, we can use the Clar's structure to cluster the π orbitals into 4 clusters with two 6 orbital sextet clusters and two outer double bond clusters. For the APD molecule, there is a 7 orbital cluster, a 5 orbital cluster and two 2-orbital clusters. The clustering is presented in Figure 5.2. The two lowest energy triplets should have major contributions from the lowest the energy triplet states in the clusters. For example, given $\mathbf{T}_A$ and $\mathbf{T}_B$ are the two triplet states in the two 6 electron clusters. The final two triplet states will have major contributions like:

Table 5.1: The occupation of each cluster for each state for the relevant Fock configurations for the pyrene molecule. The cluster ordering is: two 6 orbital clusters and then the two 2 orbital clusters.

Fock Space	S0	T1	T2
(3,3)(3,3)(1,1)(1,1)	0.941	0.598	0.806
(3,2)(3,4)(1,1)(1,1)	0.003	0.022	0.003
(2,3)(4,3)(1,1)(1,1)	0.003	0.022	0.003
(4,3)(2,3)(1,1)(1,1)	0.003	0.022	0.003
(3,4)(3,2)(1,1)(1,1)	0.003	0.022	0.003
Excitation Energy (kcal/mol)	0	60.4	96.3

$$\mathbf{T}_1 = \frac{1}{\sqrt{2}}(\mathbf{T}_A + \mathbf{T}_B)$$

$$\mathbf{T}_2 = \frac{1}{\sqrt{2}}(\mathbf{T}_A - \mathbf{T}_B)$$

For the pyrene molecule, we compute the two lowest triplet states. We use the second lowest energy cluster state to form the excited state guess. We use the extrapolated TPSCI as the reference value. We also computed the excitation energies using ASCI[47] since it is only a 16 orbital in 16 electrons system. The excitation energies using ASCI are also within chemical accuracy with the extrapolated TPSCI results for this system.

The wavefunction contribution to few important Fock configurations are presented for both pyrene and APD molecule is presented in Table 5.1 and Table 5.2 respectively. The two excited states have major wavefunction contribution from the initial neutral configuration. The lower lying triplet state has large contributions from the charge transfer configurations from the two benzene type units, and these contributions stabilize the triplet state for the pyrene system. The APD system has an additional triplet state with large charge transfer contribution compared to the other two triplet states. The charge transfer triplet state is lower in energy than the second triplet state.

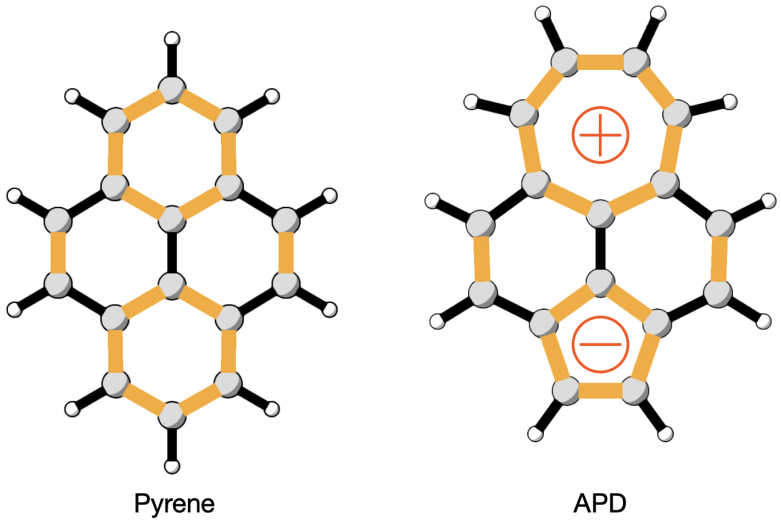


Figure 5.2: Clustering of the pyrene and APD molecule. Highlighted bonds correspond to a cluster.

Table 5.2: The occupation of each cluster for each state for the relevant Fock configurations for the APD molecule. The cluster ordering is: 7 orbital cluster, 5 orbital cluster and the two 2 orbital clusters.

Fock Space	S0	T1	T2	T3
(3,3)(3,3)(1,1)(1,1)	0.882	0.564	0.291	0.453
(4,3)(2,3)(1,1)(1,1)	0.018	0.146	0.252	0.174
(3,4)(3,2)(1,1)(1,1)	0.018	0.146	0.251	0.174
Excitation Energy (kcal/mol)	0	51	55.1	64.2

5.3.3 Medium sized PAH systems

As presented in the previous section, the triplet states of the full molecular system can be formed using TPSCI and the major contribution of the triplet states. These excited states can be formed using a simple CIS calculation since they are only single excitations in the TPS basis. However the energies are inaccurate since these states do not have any relaxation from the other states. Here we study few medium PAH systems as presented in Figure 5.3 and present how TPSCI performs for these molecules.

For a system with 4 similar clusters, the lowest 4 excited triplet states can be formed using linear combination of lowest energy triplet in each cluster. The lowest energy triplet state would be in phase linear combination of the local triplet states.

$$\mathbf{T}_1 = \frac{1}{\sqrt{4}}(\mathbf{T}_A + \mathbf{T}_B + \mathbf{T}_C + \mathbf{T}_D)$$

In Figure 5.4, we present the data for the systems considered. The active space is selected by choosing the $2p_z$ orbitals from the RHF calculation using the ccpVDZ basis and localizing them. The active space is then partitioned into clusters based on the location of each $2p_z$ orbitals. The TPSCI calculation use these three thresholds: ϵ values ($1. \times 10^{-3}$,$7. \times 10^{-4}$, $5. \times 10^{-4}$). It has to be noted that the ϵ used in the ground state TPSCI work[75] was pruning the probability and hence was square of the ϵ in this work.

For all the systems, P1-P4, the ground state variational estimate converges much faster than the excited states. For P1 and P2, the entanglement between the clusters are much smaller compared to P3 and P4 and hence P1 and P2 has smaller PT2 correction since the variational estimate for the states are good. P3 has maximum entanglement between clusters and hence has points farther from the extrapolated results. It is interesting to note that the highest

Figure 5.3: Large PAH systems used to study excited states.

energy triplet state converges faster than other three low lying states as shown in Figure 5.3 for P1,P2 and P4. For P3, the second excited triplet state converges faster since the entanglement between the top two clusters and bottom two clusters is small.

5.3.4 Singlet Fission System

Singlet fission is a multichromophore process where a bright singlet excited state is converted into two lower energy triplets. Tensor product based methods would be ideal for these systems since the chromophores can be partitioned into different clusters. We study a pentacene dimer previously studied by [253] system with eight orbitals per chromophore. We perform an ROHF calculation with $m_s = 8$ (16 unpaired orbitals) and use the localized

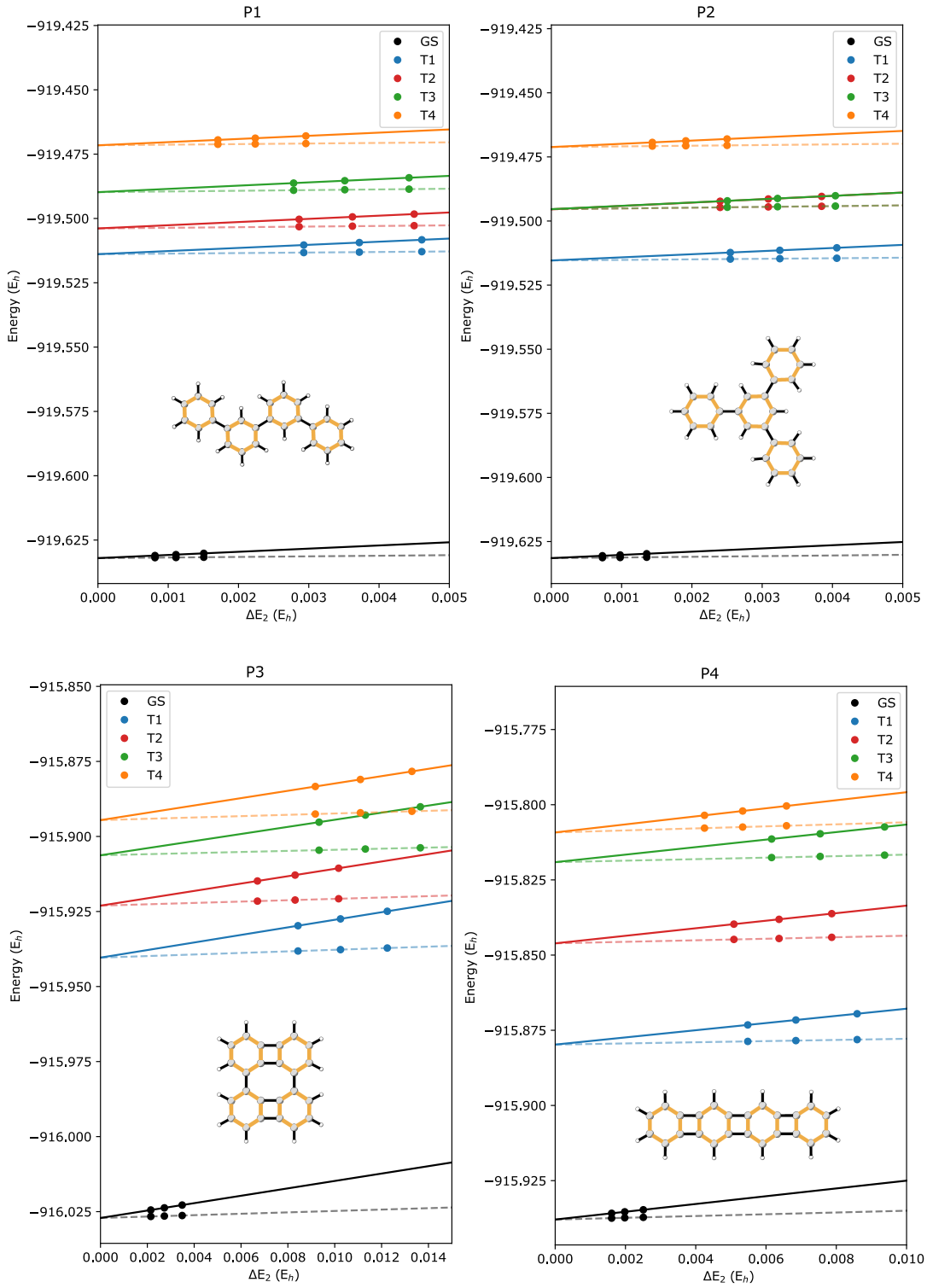


Figure 5.4: Extrapolation of the ground state and the few low lying triplet states for the medium sized PAH systems studied using TPSCI.

Table 5.3: Excitation energies (kcal/mol) for the 11 low lying excited states of the pentacene dimer along with the expectation value of the S^2 operator. We also include the type of excitation.

State	Excitation Energy	<S2>
T_1	45.97	2.000
T_1^*	46.09	2.000
T_2	58.94	2.000
T_2^*	59.13	2.000
T_3	73.06	2.000
T_3^*	73.17	2.000
^{1}(TT)	91.89	0.084
^{3}(TT)	91.94	2.000
^{5}(TT)	92.13	5.916
S_1	103.12	0.000
S_1^*	103.26	0.000

singly occupied orbitals as our active space. We partition these orbitals into two clusters based on which chromophore the orbitals belong to.

In Table 5.3, we present the lowest 11 excited states we obtained using the TPSCI algorithm with $\epsilon = 5e - 5$. The excitation energies are overestimated since we use a high spin ROHF based active space. The bright state is higher in energy than the biexciton for pentacene molecules and in the given orbital basis it is larger.[254, 255, 256] However the qualitative ordering of the energies is still present for the high spin states. As expected, the biexciton states are very close in energy. These states can be identified by analyzing the wavefunction as shown in Table 5.4.

Table 5.4: The wavefunction distribution for the three multiexciton states generated using the TPSCI algorithm for the pentacene dimer. The weight of each configuration and the number of configurations for important Fock space is provided.

Weight	# Configs	Fock space(α,β)
	^{1}(TT)	
0.225	3471	(4,4)(4,4)
0.003	2542	(4,5)(4,3)
0.003	2538	(5,4)(3,4)
0.386	1404	(5,3)(3,5)
0.382	1405	(3,5)(5,3)
	^{3}(TT)	
0.002	2542	(4,5)(4,3)
0.002	2538	(5,4)(3,4)
0.496	1404	(5,3)(3,5)
0.500	1405	(3,5)(5,3)
	^{5}(TT)	
0.772	3471	(4,4)(4,4)
0.114	1404	(5,3)(3,5)
0.114	1405	(3,5)(5,3)

5.4 Conclusion

We present an extension to the TPSCI algorithm to study excited states. We study the low lying triplet states of few π conjugated systems. The excitation energies are within chemical accuracy (1 kcal/mol) once the state specific PT2 correction is added. These low lying triplet excited states can have major contributions from the singly excited TPS states and can be captured easily. We also study a singlet fission pentacene dimer using the excited state TPSCI approach and provide equal footing for all the important states in the singlet fission process. Our model derived has all three spin states of the dark multiexciton state.

Even though we compute the excitation energies to the exact limit within the active space, these energies are way off from the experimental results. This is mainly because of the dynamic correlation that needs to be included for these states by including all the virtual orbitals. These effects need to be captured for using the TPSCI algorithm for important applications. Nevertheless addition of dynamic correlation does not change the clustering or the TPSCI algorithm as such. Recently embedding methods[257, 258] like double unitary coupled cluster (DUCC)[259, 260], self energy embedding theory [261, 262] and can be used to obtain an active space Hamiltonian which already has effect of virtual orbitals. This is one of the future works where we will use the embedding method to obtain accurate excitation energies.

Chapter 6

Size Extensive TPS

6.1 Introduction

Our group has recently developed a new method (n-body Tucker) using a Tucker decomposition and partitioning method to represent strongly correlated systems, and implemented it successfully on a Heisenberg Hamiltonian.[136] In n-body Tucker method we define clusters which are strongly correlated and each of these clusters are entangled to each other. In this approach, the cluster states are formed by diagonalizing the block reduced density matrix. Being a truncated CI type approach, the n-body Tucker is not size-extensive and hence even though it provides accurate results for medium sized systems, it cannot be applied to larger systems since the error scales with system size.

Size extensivity is defined as the correct scaling of energy with system size.[78, 263, 264] For example, the NB2 approach introduced in our previous work is not size extensive and hence can be inapplicable for larger systems. We can demonstrate this using a simple example as presented in Figure 6.1. We present data for a one dimensional Heisenberg lattice using DMRG (state of the art method for 1D systems) and our NB2 approach. In Figure 6.1, we plot the energy/site for each 1D Heisenberg model. As the length of the system is increased (reaching thermodynamic limit), any size extensive method should have an almost constant Energy/site. From Figure 6.1 it can be seen that DMRG converges quite fast around 20-30 site calculations. The NB2 being not size extensive does not asymptotically decay. For the

1D system, the DMRG method should be exact. Hence as the system size becomes larger, the error increases as NB2 does not converge. We also plot a perturbative version (NB0(2) and NB0(3)) of n-body Tucker approach and it can be seen that they converge much faster. We discuss the NB0(2) and NB0(3) methods in detail in the theory section, but it is interesting to see that the NB2 method starts off better than the NB0(2) method for around 20 sites but gets worse as the system size increases. Therefore it is necessary for a many body method to be size extensive for application to larger systems.

In this chapter, we present an alternate approach by taking similar approach from traditional quantum chemistry methods. In traditional quantum chemistry, configuration interaction, perturbation theory and coupled cluster are three major approaches to obtain the correlation energy with the latter two being more reliable since they are size extensive. In this work, we use perturbation theory instead of truncated CI to obtain the intra-cluster correlation. The excited cluster states are formed by the diagonalization of the local Fock Matrix presented in 2.

6.2 Theory

6.2.1 n-body Tucker

We start by partitioning the system into separate clusters and define a reference space by forming the tensor product of lowest energy states in each cluster. We can define this as the P space. We can simultaneously define the Q space which is the orthogonal complement to P, which consists of all excited states in each cluster. Generally the P space is much smaller than the Q space.

The final FCI coefficients in the TPS basis can be written using the projection operators P

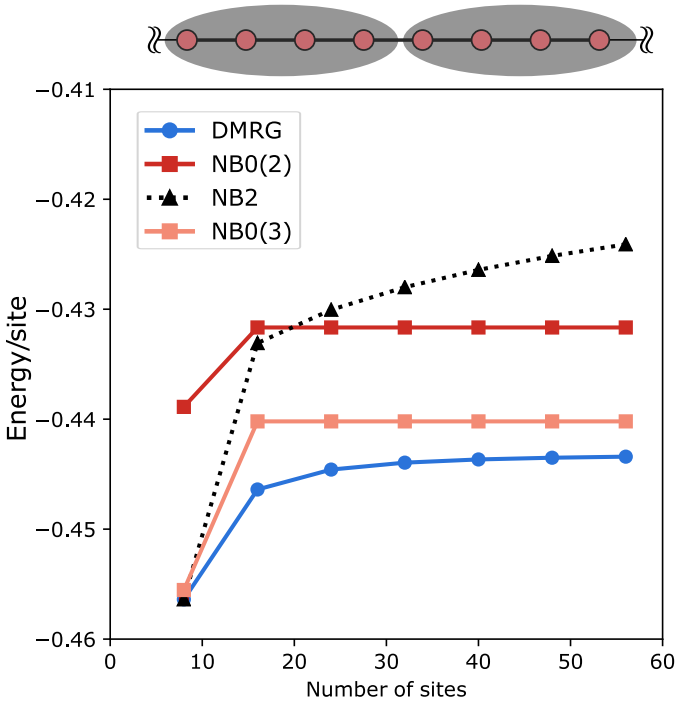


Figure 6.1: Energy/site for the 1D Heisenberg model using the different variants of n-body Tucker approaches and DMRG.

and Q in each cluster.

For a trimer example, this can be written as:

$$\mathcal{C} = \mathcal{C} * (P_A + Q_A)(P_B + Q_B)(P_C + Q_C) \tag{6.1}$$

This can be pictorially depicted, as shown in Figure 6.2. Expanding the above equation leads to:

$$\begin{aligned} \mathcal{C} \quad &= \mathcal{C}P_AP_BP_C && NB0 \\ &+\mathcal{C}Q_AP_BP_C + \mathcal{C}P_AQ_BP_C + \mathcal{C}P_AP_BQ_C && NB1 \\ &+\mathcal{C}Q_AQ_BP_C + \mathcal{C}Q_AP_BQ_C + \mathcal{C}P_AQ_BQ_C && NB2 \\ &+\mathcal{C}Q_AQ_BQ_C && NB3 \end{aligned} \tag{6.2}$$

We can define the higher order n-body approximations based on the number of Q spaces. Performing a truncated CI calculation with all the TPS configurations with up to two Q spaces will lead to a $NB2$ calculation. We can also use a perturbative correction to include intra-cluster correlation. We define the model space as the reference NB0 space and then include the higher excitations using a perturbative expansion. Since the Q space is much larger the higher body corrections like $NB3$ and above would be very expensive. However the goal of this hierarchy is to capture most important configurations within the $NB0/NB2$ space, and hence there will only be small contributions from $NB3$ and other higher-body corrections.

By expanding the exact core tensor in the basis of the projectors gives us a sum of compressed tensors. This expansion can be truncated to obtain $NB0$, $NB1$, $NB2$ approximations and so on. We have implemented the n-body Tucker method for the Heisenberg spin Hamiltonian, which is a good example of strongly correlated method. Since the method is a subspace

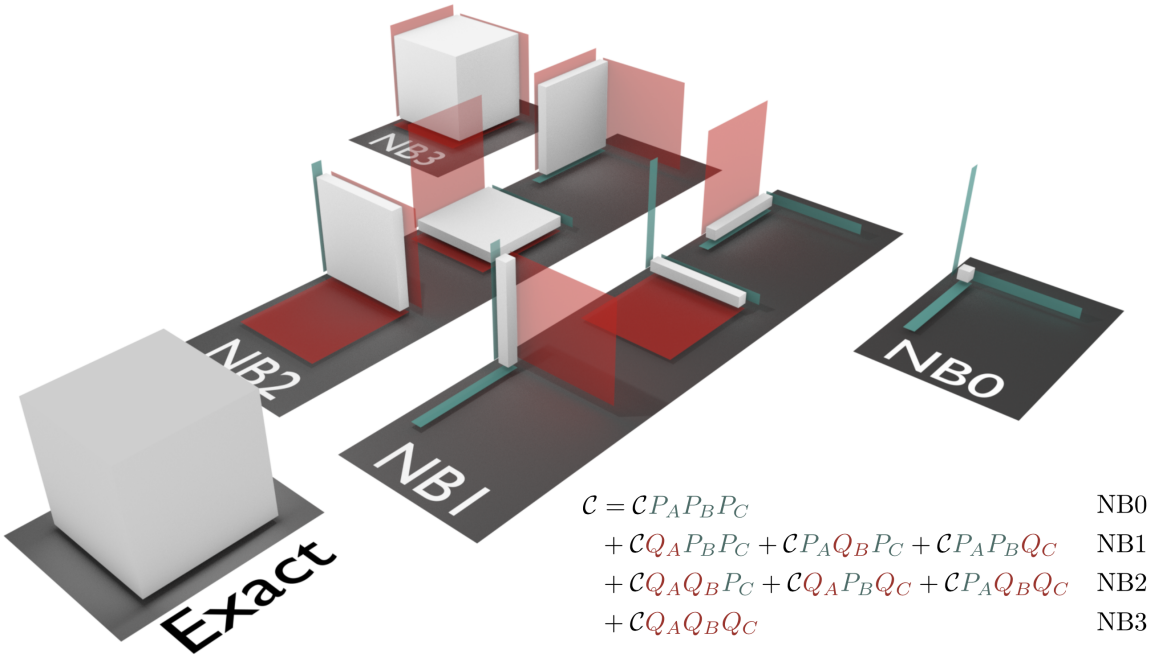


Figure 6.2: Schematic representation of the n-body Tucker approximation.

truncated variational method, n-body Tucker is not size-extensive. In this work, we are using perturbation theory to obtain an exactly size-extensive method.

6.2.2 Perturbation Theory

When the entanglement between each of the clusters is low, a perturbative method should be an effective size consistent approach to capture this entanglement. One of our research goal is to add a perturbative correction to the n-body Tucker approximation using different partitioning theories and check for higher order corrections. There are multiple ways to partition the Hamiltonian into a zero order term and the perturbation. Higher orders of the perturbative methods mostly converge to the Full CI energy but are expensive to compute. We have partitioned the Hamiltonian such that all the intra-cluster terms are contained in the H_0 term. And the H_1 terms capture all the entanglement. In the prescribed partitioning

scheme, the 1st order terms would be zero by construction when P space for each vector is 1. The Rayleigh–Schrödinger Perturbation Theory[265, 266] is size extensive and has been reliably used for many body methods.[267, 268, 269, 270]

For an Heisenberg Hamiltonian, we can define the Hamiltonian partition such that single cluster Hamiltonian terms (intra cluster) are in H_0 and all the inter cluster interactions are in H_1.

$$\mathbf{H}_0 = \sum_I H_I \tag{6.3}$$

For the spin Hamiltonian, this is equivalent to a mean-field Hamiltonian.

$$\mathbf{H}_1 = \sum_{IJ} H_{IJ} \tag{6.4}$$

Using this partition, the zeroth order energy becomes the $NB0$ energy. The energy of the system at n-th order PT correction can be written as:

$$E_0^{(n)} = \left\langle \Phi_0 \left| \hat{H}_1 \right| \psi_0^{(n-1)} \right\rangle \tag{6.5}$$

where $|\Phi_0\rangle$ is the zero order NB0 wavefunction and $|\psi_0^n\rangle$ is the $n-th$ order PT correction. The wavefunction at n-th order can be represented as:

$$\left|\psi^{(n)}\right\rangle = \left(\frac{1}{E^{(0)} - \hat{H}_0}\right) \left(\underbrace{\hat{H}_1 \left|\psi^{(n-1)}\right\rangle}_{\substack{\text{Principle} \\ \text{term}}} - \underbrace{\sum_{k=1}^{n-1} E^{(k)} \left|\psi^{(n-k)}\right\rangle}_{\substack{\text{Renormalized} \\ \text{term}}} \right) \tag{6.6}$$

We can define the resolvent for the RSPT as:

$$R_0^{(\mathrm{RS})} = \sum_{j(\neq 0)} \frac{|\Phi_j\rangle \langle \Phi_j|}{E_0 - E_j} \tag{6.7}$$

Here, the E_j is the diagonal of excited configuration Φ_j. There are two types of terms in the perturbative expansion, the principle term that is formed by n $\hat{H}_1$ terms with alternate reslovent terms and extra renormalized terms that can be formed by lower order energies. This will become more clear in the following discussions. The first order correction is zero when a mean field reference is used as the reference space. For the spin Hamiltonian, the TPS of the lowest energy local eigenstate is the reference.

$$E_0^{(1)} = \left\langle \Phi^{(0)} \right| \hat{H}_1 \left| \Phi^{(0)} \right\rangle = 0 \tag{6.8}$$

The second order correction can be written as:

$$E_0^{(2)} = \sum_{j \neq 0} \frac{\left\langle \Phi^{(0)} \left| \hat{H}_1 \right| \Phi_j \right\rangle \left\langle \Phi_j \left| \hat{H}_1 \right| \Phi^{(0)} \right\rangle}{E_0 - E_j} \tag{6.9}$$

$$E_0^{(2)} = \left\langle \Phi^{(0)} \right| \hat{H}_1 R_0 \hat{H}_1 \left| \Phi^{(0)} \right\rangle \tag{6.10}$$

In short notation, this energy can be represented as:

$$\left| \psi^{(2)} \right\rangle = R_0 \hat{H}_1 \left| \Phi^{(0)} \right\rangle \tag{6.11}$$

The renormalized terms start arising in the third order correction is:

$$E_0^{(3)} = \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle - \langle \hat{H}_1 \rangle \langle H_1 R_0 \hat{H}_1 \rangle \tag{6.12}$$

The first term is the principle term which is alternate resolvent and $\hat{H}_1$ terms and the second term is the renormalized terms. Since $\langle \hat{H}_1 \rangle = E^{(1)} = 0$ the third order correction also only has principle term.

$$E_0^{(3)} = \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle \tag{6.13}$$

The subsequent higher order corrections can be written as:

$$E_0^{(4)} = \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle - \langle \hat{H}_1 R_0 R_0 \hat{H}_1 \rangle \langle \hat{H}_1 R_0 \hat{H}_1 \rangle \tag{6.14}$$

$$E_0^{(4)} = \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle - E^{(2)} \left\langle \psi^{(1)} \middle| \psi^{(1)} \right\rangle \tag{6.15}$$

Hence at fourth order we have a renormalization term that does not vanish. If we analyze the scaling of the terms, the renormalization term contains two expectation values. Hence this term scales as N^2 with system size and is not size extensive. At the right truncation, the principle term on its own is not size extensive and have N^2 scaling terms.[78] The renormalization term exactly cancels with the non size extensive terms from the principle term at each order and leads to a final size extensive expression.[78, 271, 272] Hence it can be said that each of the expressions in the perturbative expansion is not size extensive on its own but the total sum is always size extensive.

In the context of n-body Tucker, the second order PT correction will include excitations up to NB2. The third order correction also will be within the NB2 space. The PT2 correction only uses the diagonal of the doubly excited TPS space and is cheaper than the variational NB2 calculation. For the PT3 correction, we need to compute the full doubly excited TPS

basis, and would be similar computational scaling as NB2. In the case of the PT4 correction, there will be contributions up to quadruple excited configurations. We refer to the n-th order PT correction as $NB0(n)$ in this text. Hence for any PT correction to be size extensive, the excitation order needs to be the same as the PT order. This is summarized as:

$$\text{n-body Truncation} = \text{PT order (Size Extensive)}$$
$$\text{n-body Truncation} > \text{PT order (Size Extensive)}$$
$$\text{n-body Truncation} < \text{PT order (not Size Extensive)}$$

We implemented higher order perturbative corrections up to n-th order for an isotropic Heisenberg model using Equation 6.5 and Equation 6.6. We also investigate the infinite order correction for the second order PT2 correction. This can be formed using only the principle terms in each expression while truncating the n-body expansion at the 2nd order. The energy for this method termed NB0(2)∞ can be written as:

$$E_0^{(2)\infty} = \langle \hat{H}_1 R_0 \hat{H}_1 \rangle + \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle + \langle \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 R_0 \hat{H}_1 \rangle + ... \tag{6.16}$$

This is similar to the DMBPT∞ method by Bartlett.[272] It has been shown previously that DMBPT∞ is numerically similar to CEPA[273, 274, 275] and linearized coupled cluster doubles (LCCD) method.[263, 276] For the spin Hamiltonian in the TPS basis we confirm numerically if the TPS-PT2 ∞ is numerically same as solving a TPS-CEPA. These sets of methods are important since they can be extended to multi-reference cases easily as compared to coupled cluster method.[277, 278]

6.3 Results

We investigate the effect of higher order PT corrections on a Heisenberg Hamiltonian in 1D and 2D in Section 6.3.1 and Section 6.3.2 respectively. We also investigate the TPS-LCC (also named NB2(2)∞ for the spin lattice) for a few molecular systems in Section 6.3.4. All DMRG results for the spin lattice are computed using iTensor.[147]

6.3.1 1D Spin Lattice

We used the one-dimensional spin lattice example to check the size-extensivity of the perturbative method we implemented. DMRG values are provided as reference values.

Effect of cluster size

We investigate the effect of two size clustering: a) 2 sites per cluster, and b) 4 sites per cluster. The $NB2$ method gives reasonably good results for smaller systems but as the system size increases, the NB2 method diverges from the DMRG. This is because the method is not size extensive and hence the error gets larger as the system size is increased. The $NB0(2)$ method converges quite quickly as the system size increases. As expected the NB2 method is even worse than the cheaper $NB0(2)$ calculation for systems with more than 20 sites. Hence the $NB0(2)$ method is cheaper and better than the $NB2$ method for systems with more than 20 sites for both clustering sizes. Even though it is better than the NB2 method, its still farther than the reference DMRG result. We therefore investigate higher order PT corrections.

As seen in Figure 6.3, all the perturbative methods converge to a constant as we increase the number of sites in the 1D system. It can be seen that with a larger cluster size, all the PT

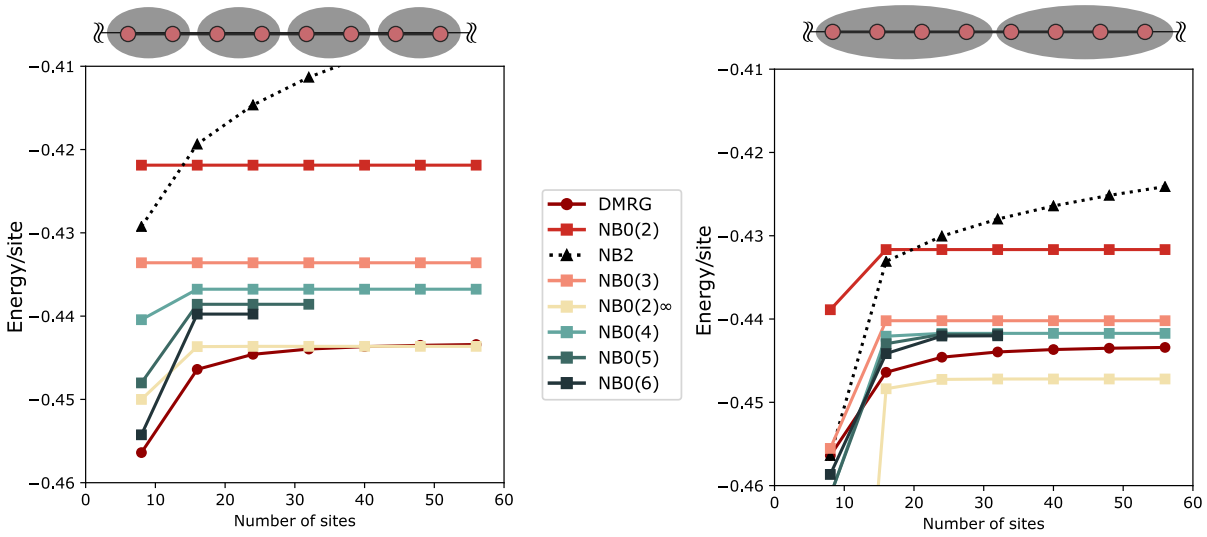


Figure 6.3: The error of the different PT orders with the exact value for different J_1/J_2 ratios. We have included the DMRG values as reference values.

corrections ($NB0(n)$) are more accurate compared to the smaller cluster. The $NB0(2)\infty$ method gives accurate results for the 2 site per cluster partitioning. This being less expensive than NB0(4) and higher PT corrections makes it an attractive approximation. Even with higher order PT corrections, it can be seen that the PT correction does not reach the exact limit for the uniform lattice. Even though at infinite order PT it would be exact, it would also need to include the same number of configurations at infinite order as FCI, making it not practically possible. The accuracy of all the approximate n-body Tucker methods increase as we increase the cluster size at a higher computational cost. The $NB0(2)\infty$ method provides very accurate results for 2 site clusters and is a promising method when the inter fragment clusters are strong. As seen from the 4site cluster example, this approach overestimates the correlation energy when the cluster size is increased similar to traditional LCC approaches.

Effect of inter cluster interaction

The applicability of a fragment based method depends a lot on the inter cluster interactions. In order to study how well our methods handle inter cluster interactions, we define a modified Heisenberg spin lattice with alternate J_1, J_2 interactions such that J_1 is within the cluster and J_2 is outside the cluster. Here we study how the perturbative method works for a given ratio of J_2/J_1. In the last section, it was seen that the PT order does not converge well for the 1:1 case for 2-site clusters. We study small J_2/J_1 ratios to up to 1:1 which is the uniform lattice studied in the previous section. We plot the error in energy/site at the thermodynamic limit in Figure 6.4. We obtain this values by increasing the system size until the contribution from each PT correction does not change per site as we further increase the system size (Figure 6.3). For lower J_2/J_1 ratios, representing systems with smaller inter-cluster interactions, even the NB0(2) correction provide accurate results. It can be seen from Figure 6.4that for lower J_2/J_1 ratios, the PT corrections converge much faster than the 1:1 case. Hence methods like $NB0(2)$ and $NB0(3)$ can be useful for systems where J_2/J_1 ratios is small with reasonable computational cost.

6.3.2 2D Spin Lattice

For the two dimensional spin lattice, we take a large spin lattice of 24 sites with a varying J_2/J_1 ratio. We study two types of 2 dimensional lattices, a cluster of 2 and a cluster of 4 as seen in the Figure 6.5. We can see that the size-extensive variants of the n-body Tucker method are an improvement to the variational n-body Tucker method at similar cost of calculation. The $NB0(3)$ and the the $NB2$ method are of similar cost and it can be seen that the $NB0(3)$ method is more accurate than the $NB2$ method for both systems. Similar conclusions can be drawn about $NB0(5)$ and $NB4$ method.

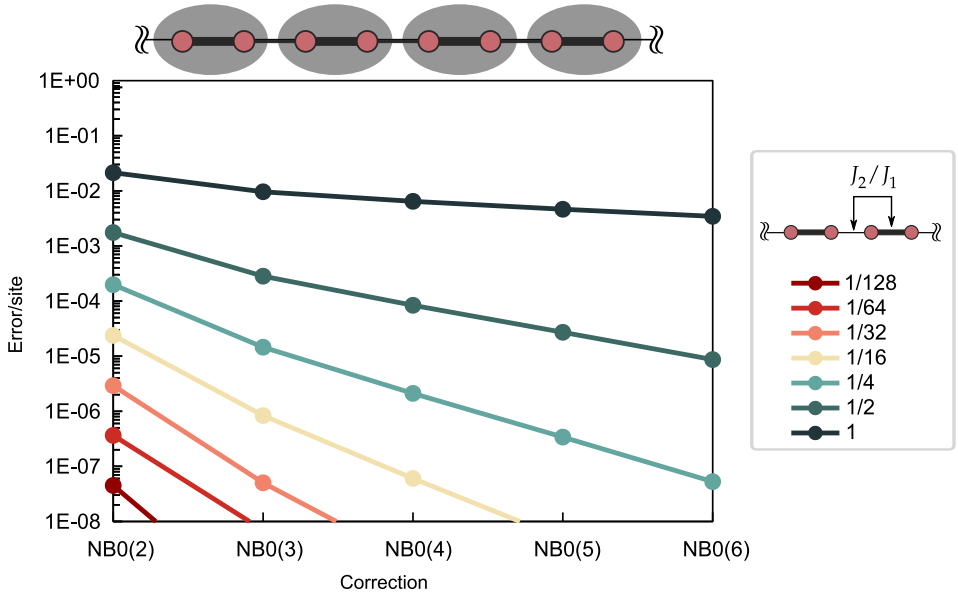


Figure 6.4: The error of the different PT orders with the exact value for different J_2/J_1 ratios at the thermodynamic limit.

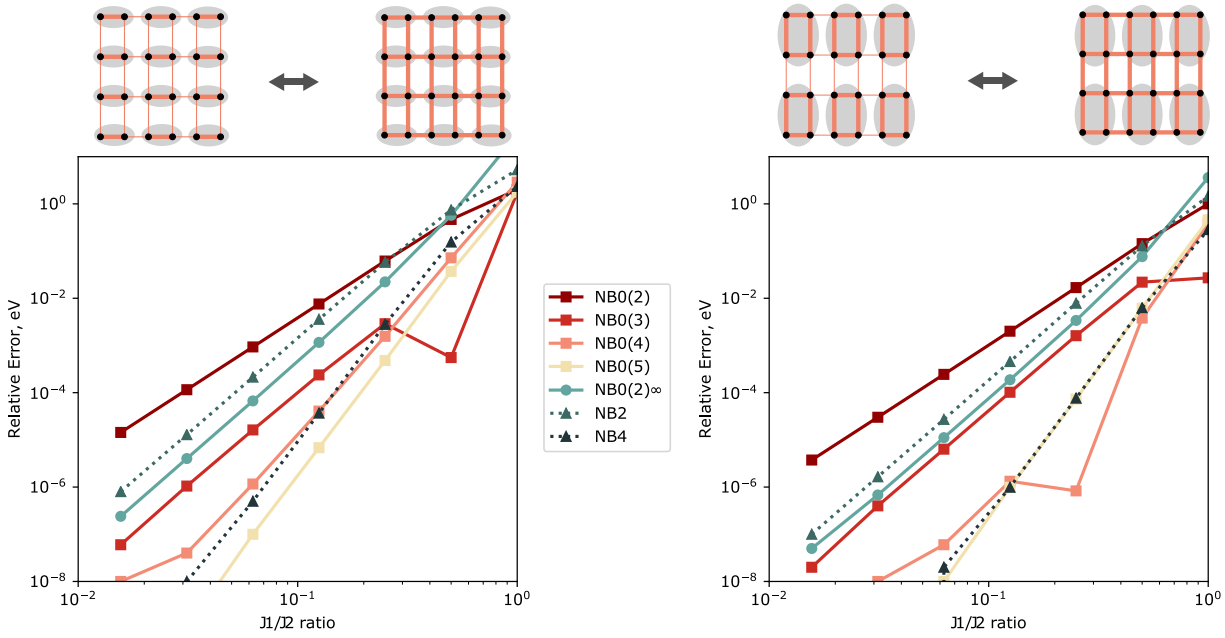


Figure 6.5: Error from the exact result for the two dimensional spin lattice. We present data for both size extensive methods (solid line) and non size extensive methods(dotted lines).

6.3.3 Multi Reference Case

In case of poly-metallic complexes, the couplings within each metal center are ferromagnetic and the couplings in between the metals are generally small and anti-ferromagnetic. To study the applicability of our methods to poly-metallic complexes, let us test it on a model system with similar characteristics. In this section, we study a small 16 site 2D spin lattice as shown in Figure 6.6. For this case, there is strong ferromagentic interaction within the cluster and anti-ferromagnetic interaction outside the cluster. For a ferromagnetic cluster, the cluster will have a high-spin ground state leading to a multi dimensional P space vector which includes all the degenerate m_s states.

For the present model, there are 5 degenerate states with $m_s = -2, -1, 0, 1, 2$. Hence for our model with four clusters, the dimension of the $NB0$ calculation which has only the P space will be $5^4 = 625$ configurations. For this, we employ the quasi degenerate perturbation theory with *diagonalize then perturb* approach for the $NB0(2)$ correction and $NB0(2)\infty$ correction.[279]

In Figure 6.6, we present the data for $NB2$, $NB0(2)$ and the $NB0(2)\infty$ method. It can be seen that the $NB0(2)\infty$ and $NB2$ have similar accuracy for all the J_2/J_1 ratios. Hence NB0(2)∞ is a promising method for both single reference and multi reference TPS wave-functions.

6.3.4 Molecular Systems using TPS-LCC

In the final section, we investigate a few molecular systems. The initial work used the spin lattice and we use similar nomenclature as to that used in our initial work.[136] In contrast to the spin-hamiltonian, the molecular hamiltonian will have up to 4-body excitations in the first order interacting space (FOIS). The TPS-LCC method used in this section is same as

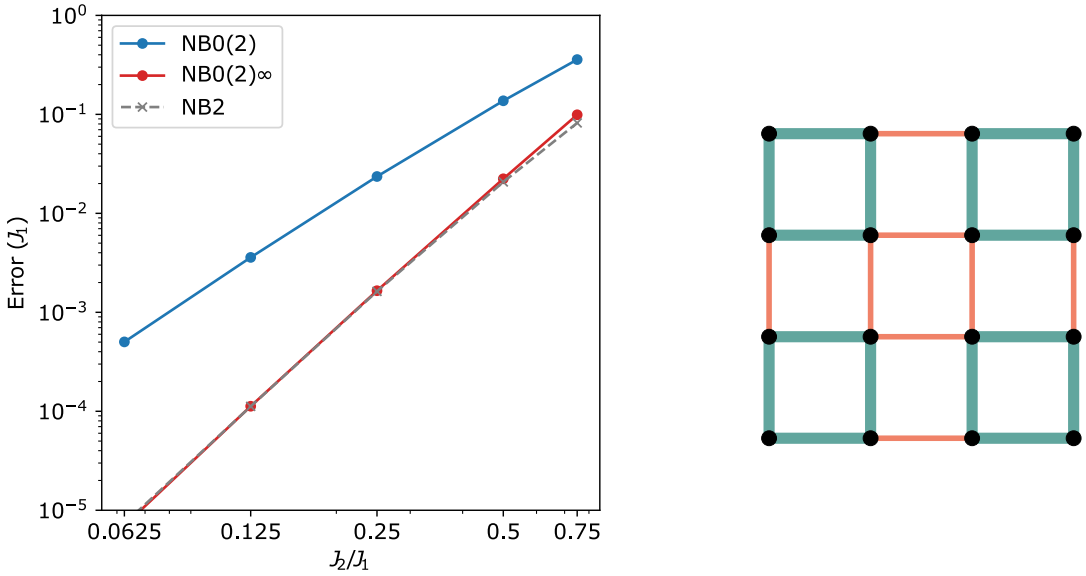


Figure 6.6: Data for the 16 site spin lattice with ferromagnetic intra-cluster interactions.

the NB0(2)∞ but since the molecular systems have has large first order wavefunction, we use a threshold to select only few important configurations. While this would not be rigorously size extensive, since we have truncated the FOIS that has upto 4 body excitations, for the present study we investigate how the TPS-LCC performs any try to converge the values numerically by increasing the threshold on the first order wavefunction.

Table 6.1: Energy of the PAH systems using the TPS-LCC and TPS-PT2 method in Hartree. We also present Extrapolated TPSCI values for comparison.

	active space	FOIS Thresh	Dim	TPS-PT2	TPS-LCC	Extrp TPSCI
1	(12,12) 2c	t=1e-5	11416	-453.6275	-453.6311	-453.6310
2	(18,18) 3c	t=1e-5	51444	-680.5892	-680.5967	-680.5964
3	(24,24) 4c	t=2e-5	66950	-904.9029	-904.9149	-904.9142
4	(24,24) 6c	t=7e-5	48222	-905.1940	-905.2296	-905.2307
5	(36,36) 6c	t=5e-5	76257	-1353.8004	-1353.8219	-1353.8244
6	(42,42)7c	t=1e-4	65757	-1582.4137	-1582.4400	-1582.4482

We also study the PAH systems presented in Figure 4.7 in Chapter 4. We present the data for the PAH systems in Table 6.1 along with the threshold used to prune the first order wavefunction. We also provide the dimension of the final wavefunction formed using this threshold. We can see that the TPS-LCC gives very accurate results for these systems compared to TPSCI results for small and medium size systems (1-4). For larger active spaces (5,6), the threshold is too drastic since the FOIS was getting very large and the LCC values are not very accurate.

We also investigate the polypyrrole molecule using the TPS based methods TPS-LCC and TPS-PT2 and compare to the Slater determinant based method. We study two polymer usits with n=4 and n=6 pyrrole units. From Figure 6.7, we can see that the error for the TPS based methods is much smaller compared to traditional Slater Determinant based methods. Specifically, we can see that the TPS-LCC method is more accurate than the CCSD(T) method and can be used as a promising alternative. Also, extension of the TPS-LCC method into multi-reference version should be easier and would be one of the major future directions.

6.4 Conclusion

We have demonstrated a new ansatz based on the tensor product of cluster state method for strongly correlated clusterable systems. Our work in this chapter emphasises the importance of size-extensivity in the many-body problem and has successfully addressed it using Raleyigh Schrödinger Perturbation Theory. We present how the error for a non size extensive method like $NB2$ increases as we increase the system size. We derive higher order PT corrections in the TPS framework and investigate the convergence of the PT corrections for different spin systems. We also derive a size extensive linearized coupled cluster type wavefunction in

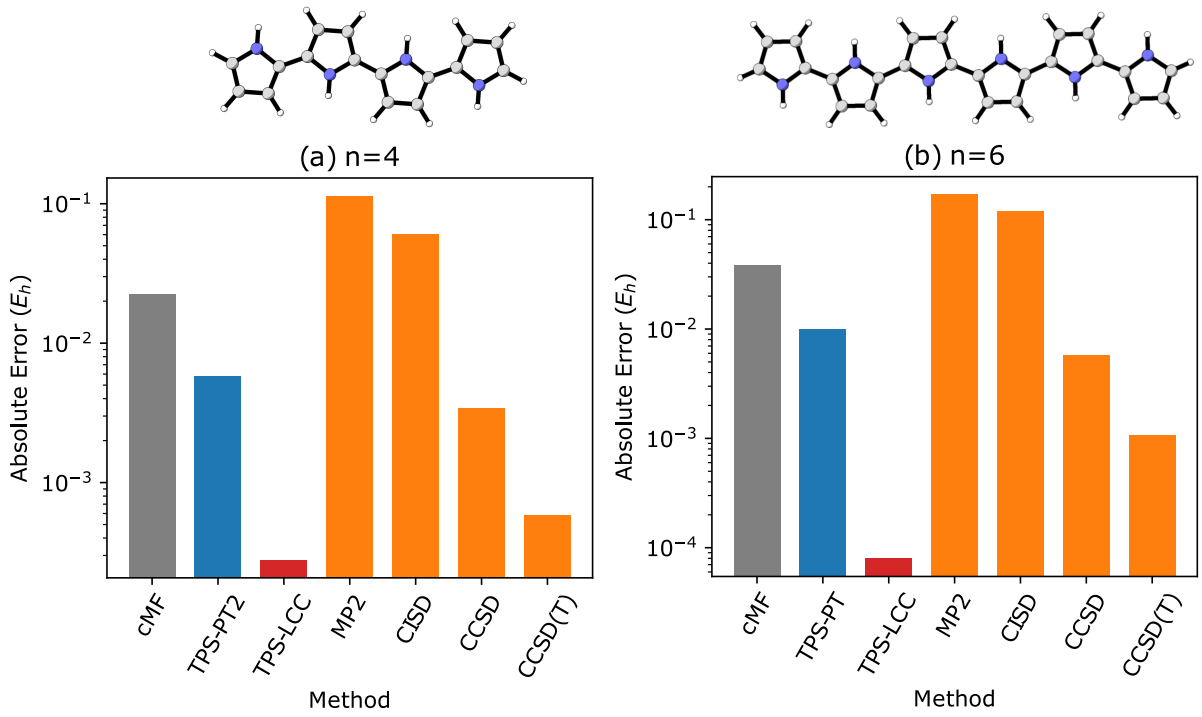


Figure 6.7: Comparison of the Slater Determinant based method with the tensor product based approach for the polypyrrole polymer system with (a)n=4 and (b)n-6 pyrrole units.

the TPS framework (TPS-LCC for molecular systems and $NB0(2)\infty$ for spin systems). We demonstrate how TPS-LCC can be a promising method for studying large strongly correlated systems. In the future, we will be investigating poly-metallic systems using the TPS-LCC approach presented in this work.

Chapter 7

Biexciton Hamiltonian for Singlet Fission

7.1 Introduction

Singlet fission[280, 281, 282, 283] is a spin allowed process where an excited singlet state (S_1) is converted to two low energy triplet states ($T_1 + T_1$). Although discovered around 50 years ago, it was recently shown that the process can be used to overcome the Shockley-Queisser limit for efficiency of single junction organic solar cells.[284] The singlet fission mechanism involves a singlet excited state (S_1) forming a spin coupled triplet excited state ^{1}(TT) before splitting into two independent triplet states.[280, 282]

The correlated triplet pair state ^{1}(TT) [285, 286, 287, 288] serves as the key spin conserving intermediate between the final triplet states and the initial singlet state for the singlet fission process.[280, 289] While the separation of the ^{1}(TT) to individual triplets was initially thought to be a fast process in *fs* region, recent studies have shown it to stay correlated

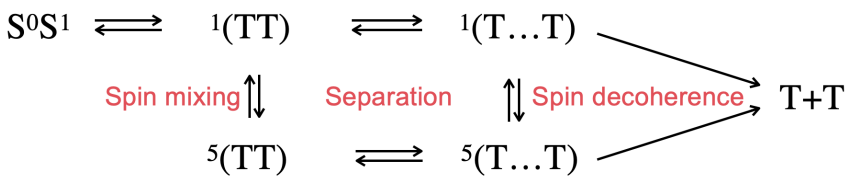


Figure 7.1: Schematic representation of the mechanism of singlet fission dissociation process.

until ns timescales.[290, 291, 292, 293, 294] Being a long lived state, multiple competing decay pathways exist for the ^{1}(TT), beyond simply separating into individual triplets. The formation of the ^{5}(TT) state, which is the quintet coupled triplet pair state is also reported recently.[291, 295, 296, 297, 298, 299] One of the main advantages of forming the ^{5}(TT) state is that it doesn't have any decay channel other than to separate into individual triplets unlike ^{1}(TT) or ^{3}(TT).[300] These long lived ^{1}(TT) states [292] might also have important applications in quantum information.[301] Being a multichromophoric process, the interaction between molecules is very important for the singlet fission process.[297, 302, 303] To realize the ideal of fast singlet-fission, the electronic interaction between chromophores should be strong enough for the triplet states to be entangled, yet not too strong such that the ^{1}(TT) becomes a trap state.

The triplet-triplet exchange interaction (K) provides insight into the energy manifold of the different spin states of biexcitonic (TT) character. This interaction is usually antiferromagnetic,[66, 304] leading to the following state energy ordering: ^{1}(TT) < ^{3}(TT) < ^{5}(TT). Closely packed (or covalently bonded) chromophores often exhibit large exchange interaction.[291] Usually, this close packing increases the electronic coupling between the initial singlet state and the ^{1}(TT) and can lead to decoherence via triplet-triplet annihilation depending on the energy level of the initial bright state and the final ^{1}(TT) state.[290, 305] For energy levels that favor TTA, it has been shown that lowering this coupling can lead to higher yields of the separated triplets from the 1TT in case of intramolecular singlet

fission.[296, 304, 306, 307]

Even after establishing the spin-state orderings of (TT), the dissociation of the correlated triplet pair state is still not well understood. Spatially separated yet spin entangled ^{1}(T...T) states have recently been proposed as a possible intermediate[308] and have subsequently been characterized experimentally.[309, 310] The triplet-triplet energy transfer (or Dexter)[311] mechanism has been proposed for the formation of these (T...T) states.[309, 310, 312, 313] The triplet-triplet energy transfer rate largely depends on the square of the triplet-triplet energy transfer coupling (t),[314, 315, 316] which is an exchange type interaction and decays exponentially with distance between donor and acceptor (similar to K).[317, 318] Formation of the ^{1}(T...T) has recently been shown to improve the final triplet yield in covalently linked tetracene trimers and tetramers.[319]

The purpose of this manuscript is two fold: First, we present an efficient ab initio-based technique to characterize the multi-excitonic states in terms of a simple model Hamiltonian consisting of physically meaningful quantities, t and K. This model Hamiltonian, in turn, allows us to identify values of t and K which lead to favorable energetics for dissociation of the (TT) state into independent triplet excitons. Second, we perform numerical calculations to resolve the connection between chromophore packing and the t and K parameters to identify which geometries might be preferable for efficient triplet separation.

Theoretical calculations[218, 228, 253, 301, 305, 320, 321, 322, 323, 324, 325, 326] play an important role in characterizing the ^{1}(TT) state, with its first theoretical report by Zimmerman and coworkers[289] much before its first experimental detection.[327] Being a doubly excited state, conventional single excitation methods (e.g., CIS, TDDFT, and RPA) cannot represent this state.[328] Hence, calculations require very expensive multi-reference methods like CASSCF[25] or MRPT[329]. Spin flip [330] methods are an alternate approach where the multi-configurational systems can be studied by starting from a well-defined high

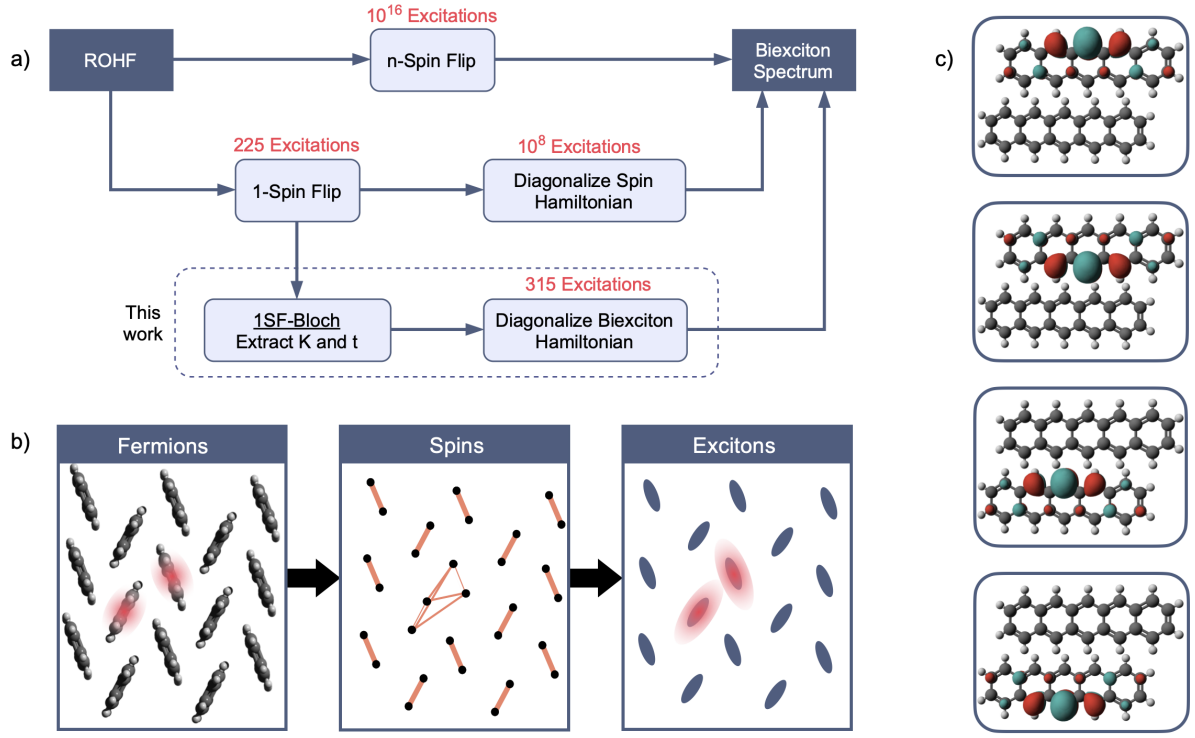


Figure 7.2: a) The final energy levels/biexciton spectrum for the bound triplet state energy manifold can be obtained using CAS-nSF and also by using CAS-1SF+ spin Hamiltonian. In this work we introduce the 1SF-Bloch method to obtain the biexciton spectrum which scales polinomially instead of exponentially. The numbers in red corresponds to the number of variables in the wavefunction for a system with 15 chromophores. b) Illustration showing the transformation from the molecular lattice into spin lattice and the subsequent biexciton model. c) Singly occupied localized orbitals in the singly occupied active space for the pentacene dimer.

spin single reference. In this work, we use a single spin flip method to obtain the two key electronic couplings (K and t) needed to describe the multiexciton space. We present numerical evidence that the multiexciton space is well represented using these two couplings.

7.2 Methods

Complete active space spin-flip (CAS-SF) and restricted active space spin-flip (RAS-SF) methods[61, 62, 63] have been successfully used to study the singlet fission process.[312, 331, 332, 333, 334, 335] Even though the RAS-SF calculation can produce reliable results for medium sized systems, the calculations get intractable as the number of chromophores increases. The reason, is that the number of spin-flip excitations increases combinatorially with number of chromophores.

Recently, our group proposed a hybrid numerical/analytical method where only a simple 1-spin flip calculation is needed to obtain the biexciton manifold instead of factorially scaling n spin flips.[66] This is carried out by extracting a spin Hamiltonian using the single spin flip method and subsequently using the spin Hamiltonian to form the excited states, hence having the singly, doubly, and higher excited triplet states.[65] However, having extracted a spin Hamiltonian does not fully alleviate the scaling problem. Even though the spin Hamiltonian is smaller compared to the ab initio Hamiltonian, the scaling is still factorial with system size. In this work, we further simplify this Hamiltonian to form a biexciton model, which only involves the doubly excited space of the spin Hamiltonian. We further demonstrate that this Hamiltonian can be formed solely by using the biexciton exchange and the biexciton hopping term.

7.2.1 Setup

While spin models have commonly been used to represent the coupled (TT) state,[280, 286, 303, 336] these models mapped each chromophore to a spin 1 site. While these approaches were able to describe the dimer systems, the 1-to-1 mapping prevented the descriptions of biexcitons on model systems containing more than two chromophores. In our recent work,[66] a 1-to-2 mapping was introduced, wherein a chromophore system with n chromophores can be cast as a $2n$ site spin lattice, allowing application to systems larger than chromophore dimers. This approach, represents each chromophore as having two $S = \frac{1}{2}$ spins, rather than one $S = 1$ spin. In brief, this process involves computing the high spin ROHF (or DFT) calculation with multiplicity $2n+1$, and then carrying out a single spin flip calculation (CAS-SF, RAS-SF, SF-TDDFT) to obtain the lowest $2n$ eigenvectors in the $2n-1$ multiplicity space (Figure 7.2a). The singly occupied orbitals are localized (Figure 7.2c) which organizes the determinant basis into either neutral or ionic configurations. We finally project the 1-SF eigenstates into the neutral determinant basis and reorthogonalize to obtain a new basis. The ab initio eigenvalues can be represented as a Bloch effective Hamiltonian matrix[65, 337, 338] and the spin half exchange interaction can be obtained from the off diagonal elements of this effective Hamiltonian.[66]

For a dimer system as shown in Figure 7.2a, starting from a quintet guess and using a single spin flip gives the two triplet states, ^{3}(TT) and ^{5}(TT) in the $m_s = +1$ space. Using these four states, we can form the spin=$\frac{1}{2}$ Hamiltonian using the procedure as mentioned in Ref. [66].

$$\hat{H}^{\text{spin}} = -2 \sum_{\text{ij}} J_{\text{ij}} \hat{S}_{\text{i}} \hat{S}_{\text{j}} \tag{7.1}$$

The ability to construct an accurate spin model offers multiple advantages such as deeper

conceptual insight, numerical efficiency, and ease of implementation of periodic boundary conditions. In the spin Hamiltonian, the interactions between sites which are on the same chromophore are much largeer than the interchromophore interactions. Hence we can write the Hamiltonian as

$$\hat{H} = \sum_{I} \hat{H}_I + \sum_{IJ} \hat{H}_{IJ} \tag{7.2}$$

where I and J are chromophore indices. The intra chromophore exchange interactions are much larger than inter terms, and hence we can define the zeroth order as the intra terms. These intra chromophore interactions give an estimate of the excitation energy for a completely localized triplet state on each chromophore. We can diagonalize $\sum_I \hat{H}_I$ to form exciton subspaces corresponding to single, double and higher excited triplet states. The energy manifold of each of these excited subspaces is well separated from the manifolds of other subspaces.[66]

Since we are interested in the biexciton space, we can therefore ignore the effect of other states including the reference lowest energy state and other excited manifolds. We can diagonalize the Hamiltonian in the biexciton space to obtain the energies of the (TT) states. We test this model by analyzing the biexciton energy by diagonalizing the spin Hamiltonian and using our model for a tetracene heptamer in Section 7.3.1. By removing the ground state from the diagonalization, this approximation provides size consistent energies[339] and can be extended to larger systems. This diabatic biexciton space has mainly two types of interactions, the exchange interaction between the triplet states and the energy transfer coupling. We will refer to our method to obtain the exchange interaction and the transfer coupling as 1SF-Bloch approach.

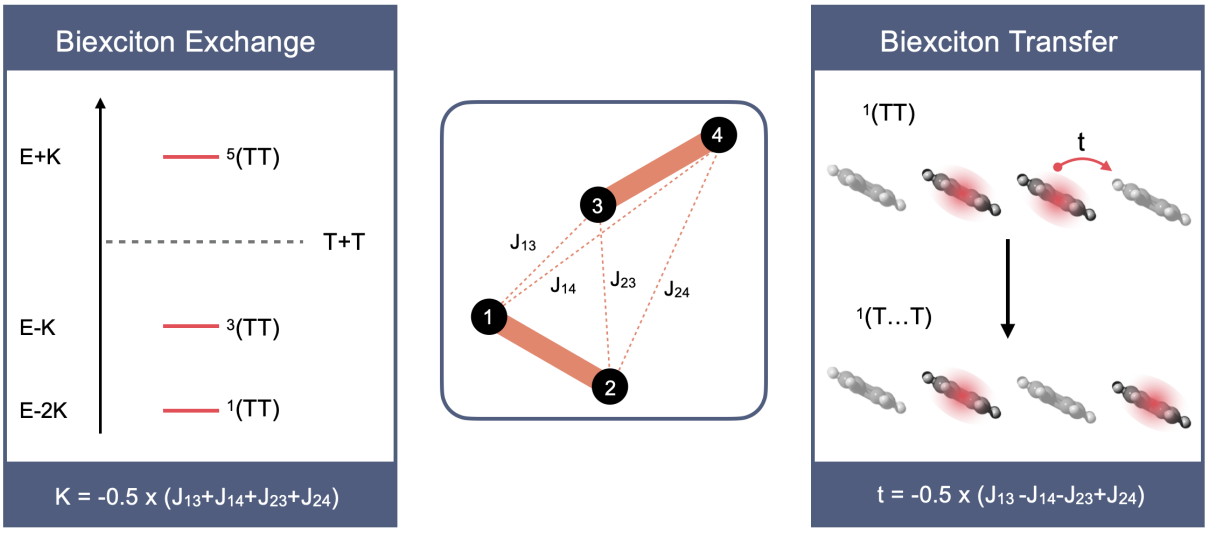


Figure 7.3: The 4 site spin model for a two chromophore system. The biexciton exchange integral splits the three spin components energetically. The biexciton transfer integral plays an important role in formation of ^{1}(T...T) from ^{1}(TT).

7.2.2 Biexciton Exchange Integral

The decoherence of the ^{1}(TT) state for molecular dimers depends largely on the inter triplet exchange interaction (K).[66, 340, 341] For a chromophore dimer system, we can compute the K value from the spin Hamiltonian derived using the 1SF calculation. The biexciton basis for a dimer system in the $m_s = 0$ space can be formed using the zero order space $|\mathrm{T}^+\mathrm{T}^-\rangle$, $|\mathrm{T}^0\mathrm{T}^0\rangle$ and $|\mathrm{T}^-\mathrm{T}^+\rangle$ where the first index corresponds to chromophore A and the second index corresponds to chromophore B.

The $|\mathrm{T}^-\rangle$, $|\mathrm{T}^0\rangle$ and $|\mathrm{T}^+\rangle$ can be written as

$$\begin{aligned} |\mathrm{T}^-\rangle &= |\beta\beta\rangle \\ |\mathrm{T}^0\rangle &= \tfrac{1}{\sqrt{2}}|\alpha\beta + \beta\alpha\rangle \\ |\mathrm{T}^+\rangle &= |\alpha\alpha\rangle \end{aligned} \tag{7.3}$$

The matrix element between any two of these states can be evaluated quite easily for the spin lattice. For example, the matrix element between the $|\mathrm{T}^+\mathrm{T}^-\rangle$ and $\left|\mathrm{T}^0\mathrm{T}^0\right\rangle$ in terms of the spin half exchange matrix element comes as,

$$\left\langle \mathrm{T}^+\mathrm{T}^- \right| H \left|\mathrm{T}^0\mathrm{T}^0\right\rangle = -\frac{1}{2}(J_{13} + J_{14} + J_{23} + J_{24}) \tag{7.4}$$

Here site 1 and 2 are in chromophore A and site 3 and 4 are in chromophore B (Figure 7.3).

This integral is the triple-triplet exchange interaction (K) between the two triplet states in the biexciton space. Hence we can write K in terms of the inter chromophore exchange interactions.

$$K_{\mathrm{AB}} = -\frac{1}{2}(J_{13} + J_{14} + J_{23} + J_{24}) \tag{7.5}$$

The 3×3 Hamiltonian for the dimer using basis $|\mathrm{T}^+\mathrm{T}^-\rangle$, $\left|\mathrm{T}^0\mathrm{T}^0\right\rangle$ and $|\mathrm{T}^-\mathrm{T}^+\rangle$ can be written as ,

$$\hat{H} = \begin{bmatrix} E - K_{\mathrm{AB}} & K_{\mathrm{AB}} & 0 \\ K_{\mathrm{AB}} & E & K_{\mathrm{AB}} \\ 0 & K_{\mathrm{AB}} & E - K_{\mathrm{AB}} \end{bmatrix} \tag{7.6}$$

where E is the reference energy for the biexciton space, and K_{AB} is the exchange interaction between the two triplets in chromophores A and B. Diagonalization of this Hamiltonian leads to the three spin components of the multiexciton, ^{1}(TT), ^{3}(TT) and ^{5}(TT) with energies $E - 2K_{\mathrm{AB}}$, $E - K_{\mathrm{AB}}$ and $E + K_{\mathrm{AB}}$ respectively and eigenvectors:

$$\left|^1(\mathrm{TT})\right\rangle = \frac{1}{\sqrt{3}}\left(\left|\mathrm{T}^+\mathrm{T}^-\right\rangle - \left|\mathrm{T}^0\mathrm{T}^0\right\rangle + \left|\mathrm{T}^-\mathrm{T}^+\right\rangle\right) \tag{7.7}$$

$$\left|^3(\mathrm{TT})\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\mathrm{T}^+\mathrm{T}^-\right\rangle - \left|\mathrm{T}^-\mathrm{T}^+\right\rangle\right) \tag{7.8}$$

$$\left|^5(\mathrm{TT})\right\rangle = \frac{1}{\sqrt{6}}\left(\left|\mathrm{T}^+\mathrm{T}^-\right\rangle + 2\left|\mathrm{T}^0\mathrm{T}^0\right\rangle + \left|\mathrm{T}^-\mathrm{T}^+\right\rangle\right) \tag{7.9}$$

The K parameter also plays an important role that it contributes to the exciton binding energy (E_b) defined as the energy difference between ^{1}(TT) and separated triplets. There are studies reporting the biexciton binding energy as low as neV[342, 343] to as high as 100's of meV.[344, 345] Connecting singlet fission chromophores through a covalent linker[346] increases the interaction between the chromophores and usually results in large E_b compared to crystals.[304] One of the major limitations with covalently connected chromophores is that the rate of recombination to form the singlet state increases, diminishing the final triplet population.[290, 305, 347] Recently, Basel *et. al* studied singlet fission using a non-conjugated linker between the pentacene monomers thereby decreasing the inter-chromophore exchange interaction.[295] Another approach is by connecting the singlet fission chromophores through meta posotion of a bridging benzene molecule to change the sign of the exchange interaction.[296, 304] The final triplet yield is increased in such cases. Hence it can be said that low/ ferromagnetic K value can lead to improved triplet yield.

The exchange interaction (Equation. 7.5) derived here is similar to the two orbital representations using the diabatic pseudo canonical orbitals with HOMO and LUMO orbitals in both the chromophore. The exchange interaction comes down to the sum of exchange integral between the orbitals in that representation as well.[285, 308, 348]

7.2.3 Biexciton Transfer Integral

Recent studies have proposed that the decoherence of the ^{1}(TT) state occurs by evolving into the ^{1}(T...T) state by triplet-triplet energy transfer.[309, 310, 312, 313, 349] The magnitude of TT coupling, which is exchange type in nature depends exponentially on the distance between triplet donor and acceptor.[311, 350, 351] Theoretical methods using singly excited methods have been previously used to estimate the triplet-triplet energy transfer integral using the fragment spin difference method.[315, 318] Charge transfer states are also important for the triplet-triplet energy transfer process,[314, 317, 352] and are indirectly (via the effective interactions) included in our model.[66]

Given the spin Hamiltonian, the triplet-triplet hopping can be extracted easily by looking at the biexciton or single exciton space. For example, consider a three chromophore system, where the biexciton can shift from one pair to another. For the trimer, the size of the biexciton space is $^3C_2 \times 3 = 9$ which includes three of each ^{1}(TT), ^{3}(TT) and ^{5}(TT). By closely analyzing the Hamiltonian, it can be seen that the biexciton subspace contains both the exchange interaction (K) between the chromophores and a new integral corresponding to triplet transfer, t. This becomes evident when we work with the spin coupled biexciton space. This makes the Hamiltonian block diagonal with dense singlet, triplet and quintet blocks. The resulting Hamiltonian after spin coupling has the K terms in the diagonal.

The triplet-triplet hopping between two diabatic biexcitons $^1(\mathrm{TT})_{AC}$ and $^1(\mathrm{TT})_{BC}$ can be defined as the matrix element between the two states using the spin Hamiltonian.

$$t_{\mathrm{AB}} = \left\langle {}^1(\mathrm{TT})_{BC} \right| \hat{H} \left| {}^1(\mathrm{TT})_{AC} \right\rangle \tag{7.10}$$

Similar to K, the t parameter between chromophore A and B can be written as a linear

combination of the inter chromophore exchange terms.

$$t_{\mathrm{AB}} = \frac{1}{2}(-J_{13} + J_{14} + J_{23} - J_{24}) \tag{7.11}$$

For the trimer, the "singlet-biexciton" block of the Hamiltonian can be represented as using basis of $^1(\mathrm{TT})_{AB}$,$^1(\mathrm{TT})_{AC}$ and $^1(\mathrm{TT})_{BC}$ can be written as,

$$\hat{H}\left[^1(\mathrm{TT})\right] = \begin{bmatrix} E - 2\mathrm{K}_{AB} & \mathrm{t}_{BC} & \mathrm{t}_{AC} \\ \mathrm{t}_{BC} & E - 2\mathrm{K}_{AC} & \mathrm{t}_{AB} \\ \mathrm{t}_{AC} & \mathrm{t}_{AB} & E - 2\mathrm{K}_{BC} \end{bmatrix} \tag{7.12}$$

The transfer integral for the quintet and triplet multiexcitons are the same as the singlet state. The off diagonal elements for the model Hamiltonians for those states look exactly same as Equation 7.12. We compare the triplet-triplet transfer coupling obtain using the 1SF-Bloch with the traditional CIS based approach[315, 353] in supporting information. The energy cleft method has also been used recently to help triplet-triplet separation in case of intramolecular singlet fission.[354]

7.2.4 Biexciton Hamiltonian

By using the 1SF-Bloch method, we have derived parameters that model the biexciton subspace. The (TT) energy manifold for all three spin states can be obtained using the biexciton space. The final biexciton Hamiltonian for the $^1(\mathrm{TT})$ state using the two parameters can be written as:

$$\hat{\mathbf{H}}_{1(\mathrm{TT})} = \sum_{A,B} -2K_{\mathrm{AB}} \left|A,B\right\rangle \left\langle A,B\right| + \sum_{ABC} t_{AB} \left|A,C\right\rangle \left\langle C,B\right| . \tag{7.13}$$

The basis $|A,B\rangle$ is the biexciton state where chromophores A and B are in their triplet excited states. The full biexciton Hamiltionian with all three spin states for an n chromophore system scales as $3 \times {}^{n}C_{2}$. Hence, the dimension of this Hilbert space scales quadratically with the number of chromophores compared to the factorial scaling for the parent spin Hamiltonian. In this work we diagonalize this Hamiltonian to look at the biexciton energy spectrum and understand the effect of K and t on the biexciton spectrum. Numerically these couplings are small and usually in meV scale which is much smaller than the excitation energies.

With the aide of the 1SF-Bloch model, we can now use a simple 1-spin flip calculation from a high spin reference to simultaneously obtain both the biexciton exchange and transfer integrals. In terms of the impact on singlet-fission triplet yield, the two parameters play opposing roles. Large antiferromagnetic exchange interaction (positive K) can lead to ^{1}(TT) being a strongly bound state and reduces final triplet yield. Meanwhile having a larger hopping integral compared to the exchange should facilitate the formation of spatially separate (T...T) state. *Hence, we expect systems with large t and small/negative K to facilitate in the dissociation of the biexciton state.*

Even though both of these interactions are on the same scale, their relative magnitudes can change based on the interaction between the chromophores. Since, the biexciton exchange interaction is a sum of the spin Hamiltonian terms, it does not really become zero for molecular assemblies when all sites have antiferromagnetic interaction. The biexciton transfer has both positive and negative contributions from the spin half exchange interactions meaning that

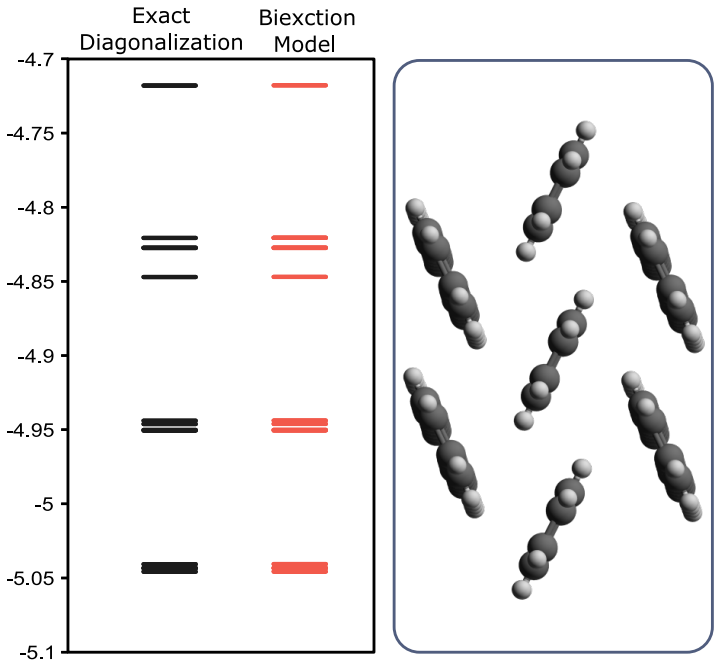


Figure 7.4: Comparison of the biexciton model with the exact result from the spin Hamiltonian for a system with seven tetracene units.

there is an orientational dependence. For example, having two chromophores perpendicular to each other cancels the exciton transfer integral while the biexciton exchange interaction stays non zero.

While our proposed model is accurate in that it can reproduce more expensive n-spin flip calculations, and provides consistent results with the FSD approach, by it's very nature it is incomplete. Most importantly, our model does not include the initial bright state. This prevents our model from describing both the initial formation of the ^{1}(TT) state, as well as the important triplet-triplet annihilation decay mechanism. Additionally, we have not included spin dipole interactions into our model. However, this is not an essential limitation, and we plan to include this in future work to study the ^{1}(TT)/^{5}(TT) mixing.

7.3 Results and Discussion

We focus on three different aspects of the biexciton model proposed in this work. In Section 7.3.1, we compare numerical data for the biexciton model to the data from the actual diagonalization of the spin Hamiltonian for a large chromophore assembly. Secondly, in Section 7.7.5, we analyze the effect of different ratios of exchange to transfer parameter and look at the biexciton spectrum for a model 1D system. Finally, in Section 7.3.3 we investigate how the exchange and transfer parameters vary as we change the chromophore orientations and suggest geometric orientations where improved triplet dissociation is possible using our model.

7.3.1 Tetracene 7-mer

As described above, the mapping of the ab initio Hamiltonian to a quantum spin lattice does not completely remove the combinatorial complexity of the problem. While the dimension of the spin-Hamiltonian Hilbert space is indeed smaller (it is the square root of the fermionic problem), it is still intractable to diagonalize for large systems. However, because we are only concerned with the biexciton states (in contrast to tri-excitons and so forth), we propose to build and diagonalize the spin-Hamiltonian only the the biexcitonic manifold. This not only reduces the computational cost to polynomial scaling, but also restores size consistency by removing the ground state from the diagonalization space.

In this section, we evaluate the magnitude of the errors obtained by diagonalizing in only the biexcitonic subspace instead of the full space. In our previous work,[66] we computed the biexciton spectrum using exact diagonalization (ED) of the full spin Hamiltonian for a large cluster with seven tetracene monomers. In Figure 7.4 we show the comparison of the ^{1}(TT) spectrum using ED and the biexciton model Hamiltonian proposed in this paper. There are

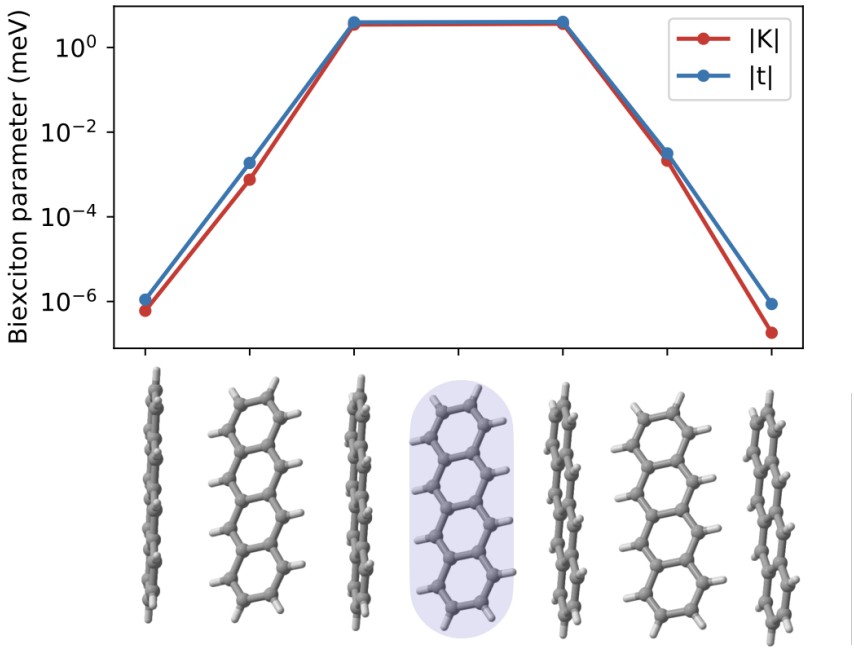


Figure 7.5: The exchange and transfer integral for the herringbone structure from a tetracene crystal structure. We present values of K and t for the central chromophore (highlighted) with the adjacent chromophores.

a total of 21 ^{1}(TT) states for this system. As seen in Figure 7.4, the spectrum obtained from diagonalizing only the biexcitonic manifold is essentially indistinguishable from the exact diagonalization. We have also derived a rigorous downfolding using quasi-degenerate perturbation theory, but since the zeroth-order results presented here are accurate enough, we only provide these slighly more accurate results in the supplementary information.

The broadening of the biexciton spectrum is largely due to static disorder of the tetracene monomers (which would vanish if periodic boundary conditions were used in the ab initio calculations, which will be a focus of future work) and other acenes having two different geometries in the herringbone crystal structure.[355] The energy gap between these two different structures is as large as 50 meV which is larger than the exchange interaction in crystal tetracene molecule.[66]

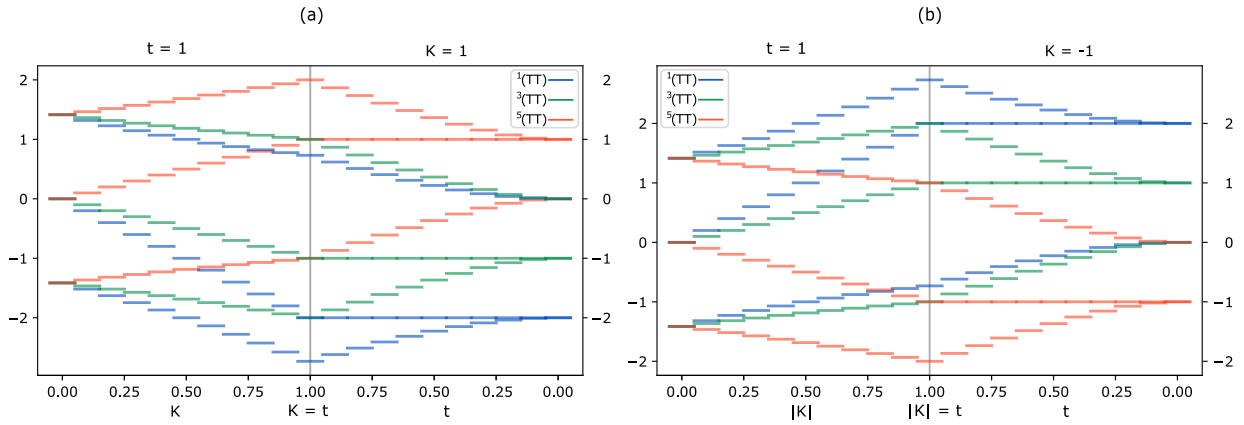


Figure 7.6: Comparison of the biexciton energy manifold for varying values of K and t for a) antiferromagnetic (positive K) and b) ferromagnetic (negative K) exchange interactions. We fix $t = 1$ meV for the left of the vertical gray line and vary K uniformly from 0 to 1. On the right side of gray line, we vary the t parameter from 1 to 0.

7.3.2 Biexciton energy manifold for 1D system

In this section, we analyse the distance dependence of the exchange and transfer parameters for a herringbone structure of a tetracene lattice. In Figure 7.5, we present the exchange and hopping parameters that couple the central chromophore with the other chromophores. It can be seen that the exchange and transfer integrals decay exponentially as the distance between the chromophores increases. For the next nearest neighbour dimer, the exchange interaction is on the scale of μeV. Because of this, we do not include the next nearest neighbour interactions in our current 1D Hamiltonian model. In the future, however, when considering spin mixing interactions, these need to be included since spin dipolar interactions are also in the μeV range.[301, 319] Each of these couplings are computed using a single spin flip CAS calculation.

We generate the biexciton energy manifold for a model 1D system with three chromophores. With three chromophores, there are total $3 \times {}^3C_2 = 3 \times 3$ biexciton states. We analyze the effect of varying K and t on the biexciton spectrum.

We present the biexciton energy values for both antiferromagnetic and ferromagnetic exchange interactions (K) in Figure 7.6. The sign of t does not change the biexciton spectrum. We fix the K parameter and vary t on the right side of the grey line and on the left we fix t and vary K for both antiferromagnetic (Figure 7.6a) and ferromagnetic (Figure 7.6b) K. For the antiferromagentic case with very low t, it can be seen that the ^{1}(TT) state is much lower in energy than the ^{3}(TT) and ^{5}(TT) states since K value dominates here.

As we increase the t parameter, we can see that one of each ^{1}(TT), ^{3}(TT) and ^{5}(TT) states becomes more stable until $K = t$. When $K = 0$, all three spin states are degenerate and would facilitate spin mixing. As we increase the K parameter, the ^{1}(TT) state becomes lowest energy for the antiferromagnetic case and the ^{5}(TT) becomes lowest energy for the ferromagnetic case. The spectrum is flipped vertically for the ferromagnetic K values. Having the ^{5}(TT) lower in energy would improve the final triplet yield since the ^{5}(TT) state does not have other decay pathways. Hence for ideal singlet fission dissociation, low/negative K values would be favourable such that ^{1}(TT) and ^{5}(TT) are very close in energy.

In the future, we will include the spin dipole interactions into the model to investigate the effect of spin mixing. A kinetic model with the diabatic states for (TT) and (T...T) including both biexciton transfer and biexciton exchange can shed more light into the biexciton dynamics and will be investigated in future work. Geometric relaxation is also needed for a complete description of the mechanism.[356]

7.3.3 Biexciton coupling for tetracene dimer

The decoherence of the ^{1}(TT) largely depends on the K and t parameters for the model presented. While large K leads to longer coherence time, large t helps in the formation of the (T...T) state. In this section, we investigate how these parameters vary as the molecular

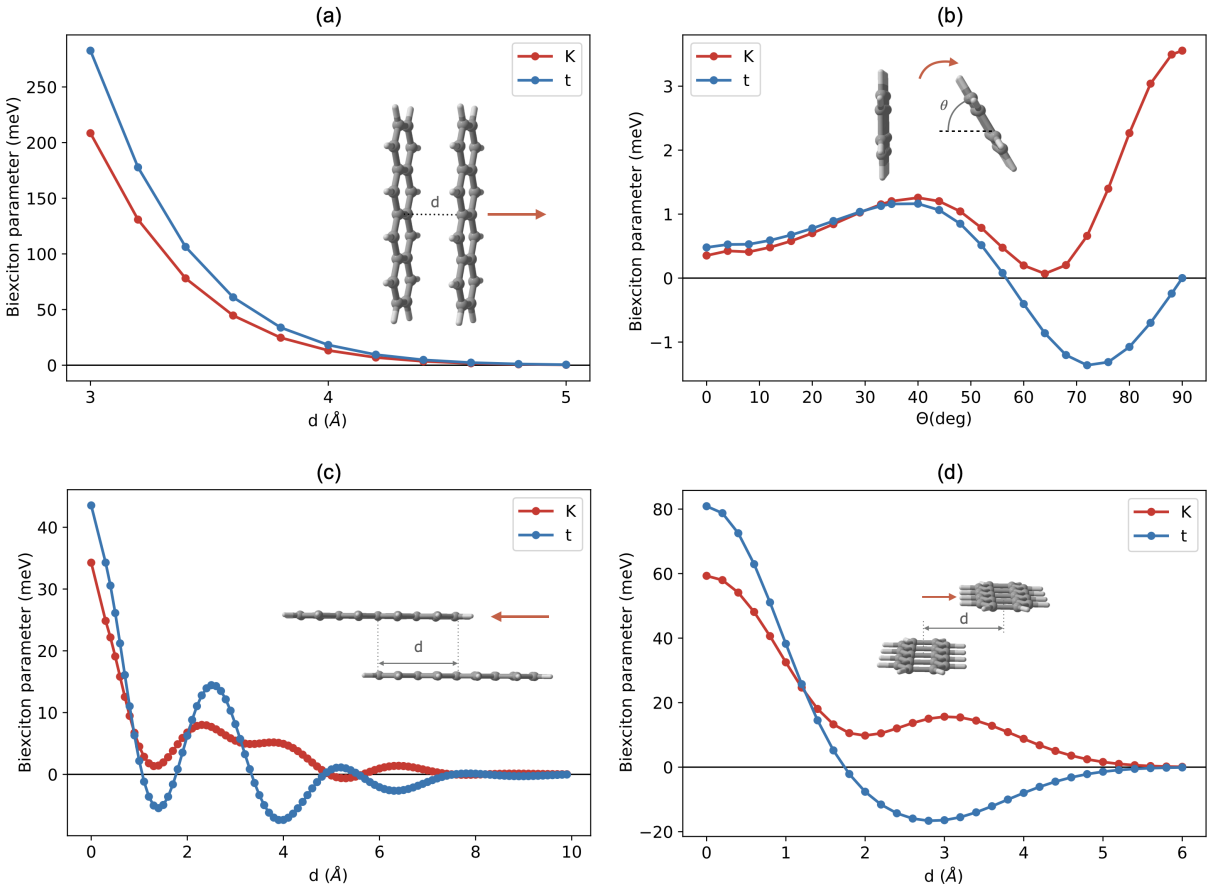


Figure 7.7: The exchange and transfer coupling for the tetracene dimer.

packing is changed. The chromophore dimer units are generated using optimized geometries of tetracene molecules at B3LYP/6-31G(d). We present t and K parameters as a function of chromophore orientation.

Center of mass-coordinate In this first panel, Figure 7.7(a), the location of one chromophore is fixed and the other chromophore is translated along the z axis. The K and t parameters decay exponentially as the distance between the chromophores (d) increases with exponential factors of $e^{-2.76}d$ and $e^{-2.77}d$, respectively. Beyond 5 the couplings fall below 1 meV. Hence having a large distance between the chromophores can increase spin

mixing as the only non-exponentially decaying term remaining are spin dipolar interactions. Since both t and K decay similarly as a function of center of mass separation (the K/t ratio stays fairly constant at around 0.73), therefore this is not a useful coordinate for optimizing triplet-triplet separation.

θ-coordinate The energetically favorable herringbone structure of simple chromophores stabilize dimers which are rotated with respect to each other. Here we analyze how this rotation angle coordinate between the two chromophores influences the biexciton exchange and transfer integrals. We fix the distance between the centre of mass of the two molecules to be 5.0 and rotate one of the chromophores. Since the center of mass is fixed, by rotating we are bringing the two chromophores closer, hence at larger angles the parameters increase in magnitude, although they are still less than 10 meV. From Figure 7.7(b), it can be seen that the exchange interaction between the two chromophores becomes very low at about 64 degrees. This is the geometry were we have realized the smallest K/t ratio, which (based on the considerations discussed above), should provide the fastest ^{1}(TT) $\rightarrow$ (T...T) separation and decoherence.

Consistent with models based on orbital overlap[357], the perpendicular geometry drives the triplet hopping term becomes zero, essentially shutting down any formation of (T...T). The spin lattice in the case of the exact perpendicular geometry is similar to a orthogonal dimer arrangement which is basically a 1D Shastry-Sutherland spin lattice.[358]

x-coordinate Next, we study the effect on biexciton exchange and transfer integrals when one chromophore is moved in the x direction (Figure 7.7(c)). We fix the distance between the planes of the two chromophores to be 3.75 . We can see that there is an interesting variation of K and t in this case and the periodicity is almost similar. While the sign of t oscillates

along this coordinate, this does not affect the hopping rate since it depends on t^2.[310, 318] In terms of minimizing K/t, one can see a favorable region occurring at around $d = 2.75$ where t reaches a local maximum. The exchange interaction between the two triplet states is ferromagnetic for the x shift from 5 to 5.6 and 7.6 to 8.2. For these geometries, the ^{5}(TT) state is lower in energy compared to the ^{1}(TT) state.

y-coordinate We also study the parameters as we move the chromophore in the y direction (Figure 7.7(d)). We fix the distance between the two chromophores to be 3.5 (similar results were obtained at 3.75). As seen in the previous case, the values of the t and K parameters have similar periodicity. In terms of K/t ratio, this coordinate minimizes this ratio only at small values of d, where the two chromophores are directly on top of one another. Hence we do not expect fast dissociation in any of the geometries along this coordinate.

7.4 Conclusion

In this work we present a simple biexciton model for the correlated triplet pair state intermediates in singlet fission materials. The biexciton energy manifold is mainly decided by two key parameters, the exciton transfer term (t) and the triplet-triplet exchange interaction (K). We demonstrate that these parameters can be derived from a simple ab initio calculation by using a single spin flip calculation from a high spin reference using the 1SF-Bloch approach as presented in this work. K and t are competing interactions, for the antiferromagnetic exchange, $K >> t$ can lead to bound biexcitons whereas $t >> K$ leads to small energy gap between ^{1}(TT) and ^{5}(TT), and favourable formation of ^{1}(T...T) states. For a ferromagnetic exchange, the quintet biexciton states is more stable than the singlet state even for larger chromophore assemblies and hence can lead to higher triplet yield.[296, 304]

These different ratios of K and t depend on chromophore interactions, hence we investigate the effect of molecular packing. For most of the geometries, the K value is antiferromagnetic. We find ferromagnetic interaction for a small region when the two chromophores are shifted along the x axis around $\sim$ 5.3 and $\sim$ 7.9. Very few chromophore arrangements lead to conditions where t is significantly larger than K. We find large exciton transfer compared to exchange along the θ rotation coordinate around 60 degrees. For these molecular packings, the biexciton dissociation should be favoured either by the fast formation of ^{1}(T...T) state or the ^{5}(TT) state.

In the future, we will be developing a Marcus theory based model which includes the entanglement of the triplet pair states.[310] The 1SF-Bloch model can be further improved by introducing hole and particle excitations either perturbatively[359, 360] or variationally,[61, 62, 63] or using an effective Hamiltonian derived using equation of motion coupled cluster (EOM-CC).[64] An intramolecular vibrational model has also been proposed to be important for the formation of the correlated triplet pair states.[361] Vibrational effects can be introduced for the model either by extracting the exchange parameters at different geometries of the vibrational mode or by introducing the vibrational coupling explicitly using phonon Hamiltonian. The 1SF-Bloch model can also be used for obtaining the coupling in triplet-triplet energy transfer reactions for artificial photosynthesis.[362, 363]

7.5 Supporting Information

We provide data for the effect of QDPT, FSD scheme for triplet energy transfer, effect of dihedral, effect of system size in the supporting information.

7.6 Acknowledgements

This research was supported by the National Science Foundation (Award No. 1752612).

7.7 Appendix

7.7.1 Biexciton transfer using CIS-FSD

The fragment spin difference (FSD) [315, 353] is a successful tool used for computing the triplet energy transfer coupling. We compare the coupling using the 1SF-Bloch model with the triplet-triplet energy transfer from the FSD scheme. In Figure 7.8, we present the comparison between the t parameter obtained using FSD scheme with our 1SF-Bloch method.

For this comparison, we consider the tetracene dimer whose center of mass is 5.0 away and at 90 degrees with each other. The relative orientation is such that one of the chromophores is perpendicular to the other. Using the 1SF-Bloch model, we can see that the J_{13} and J_{23} are equal and so is J_{14} and J_{24} for this geometry. Hence using a spin model, this coupling should be zero, and indeed we find this value to be zero.

Next, we calculate the triplet-triplet energy transfer coupling using the FSD method with the triplet states obtained using CIS method. We can see that this coupling becomes zero at 90 degrees from Figure 7.8 which has been previously reported for the charge transfer couplings as well.[357] We compute the energy transfer integral for more points near by rotating by a small angle and it can be seen that the FSD coupling and 1SF-Bloch coupling are very close for all of the points despite being formally very different approaches and using different references. This consistency between the two methods for computing t, helps validate that although our method simultaneously provides t and K, we retain the accuracy of the FSD

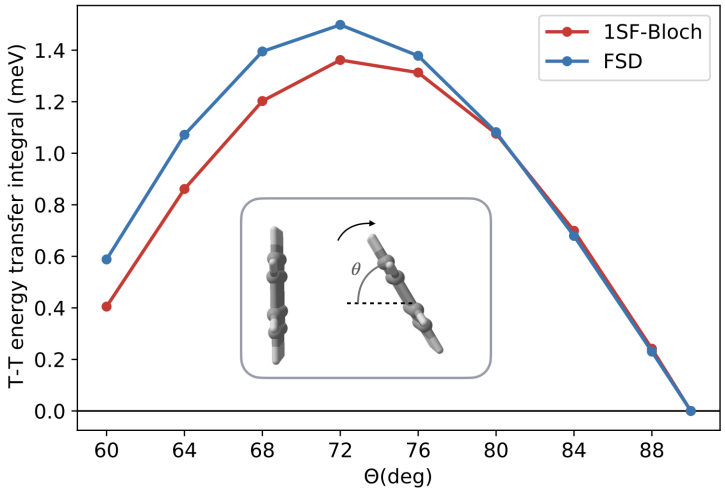


Figure 7.8: Comparison of the triplet energy transfer integral obtained using FSD on top of a CIS wavefunction with our 1SF-Bloch approach.

method which specializes on computing t.

7.7.2 Quasidegenerate perturbation theory

The biexciton model presented in the main text only uses the doubly excited space as the basis and still provides accurate results. When the chromophores are too close/intramolecular, then more accurate description is required. We can improve the energies of the biexcitons by including the effect of other exciton blocks using quasi degenerate perturbation theory (QDPT).[364, 365]

We can define a projection onto the biexciton basis and define an orthogonal projection for all the other exciton spaces and the ground state(Figure 7.9).

$$\hat{P} = \left|\Psi^0_{BE}\right\rangle \left\langle\Psi^0_{BE}\right| \tag{7.14}$$

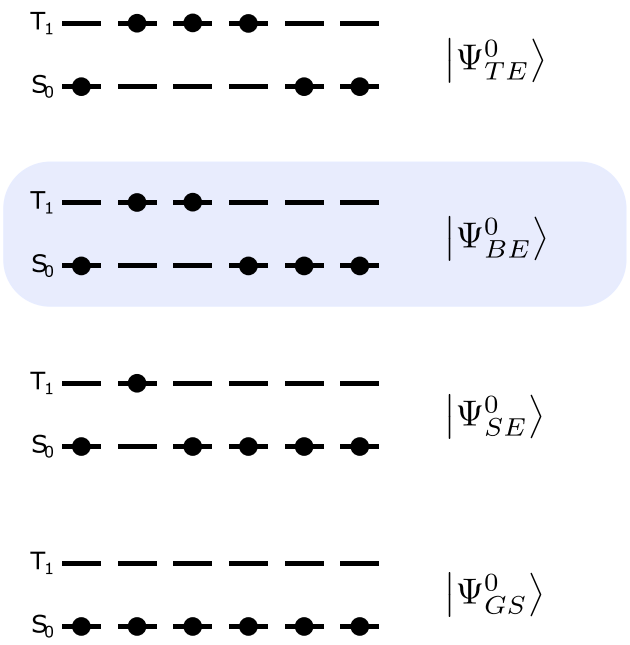


Figure 7.9: Schematic representation of the exciton basis. Singly excited (SE), biexciton (BE) and triply excited (TE) space are shown. The biexciton model only diagonalizes the Hamiltonian in the biexciton space which is highlighted.

$$\hat{Q} = 1 - \left|\Psi^0_{BE}\right\rangle \left\langle\Psi^0_{BE}\right| \tag{7.15}$$

Hence the final Hamiltonian can be written as,

$$\begin{bmatrix} \mathbf{H}^{PP} & \mathbf{H}^{PQ} \\ \mathbf{H}^{QP} & \mathbf{H}^{QQ} \end{bmatrix} \begin{bmatrix} \mathbf{C}^P \\ \mathbf{C}^Q \end{bmatrix} = E \begin{bmatrix} \mathbf{C}^P \\ \mathbf{C}^Q \end{bmatrix} \tag{7.16}$$

The effective Hamiltonian in the P space can be represented as,

$$\mathbf{H}^{eff}(E) = \mathbf{H}^{PP} - \mathbf{H}^{PQ} \left(\mathbf{H_0}^{QQ} - \mathbf{1}E\right)^{-1} \mathbf{H}^{QP} \tag{7.17}$$

The H^{QQ} is approximated as H_0^{QQ}. E is the unentangled triplet biexciton energy which is $4\times$

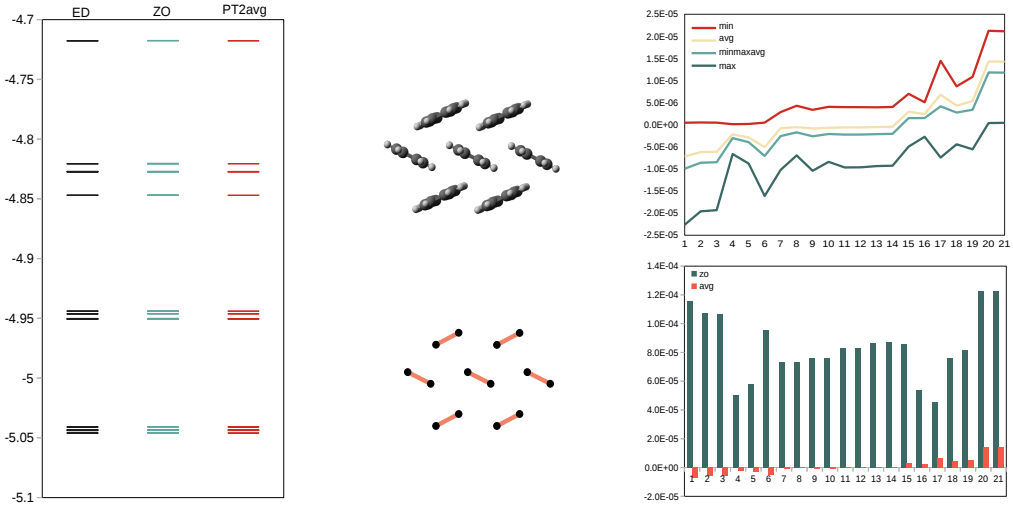


Figure 7.10: Comparison of including the PT correction with the exact result for a system with seven chromophores.

J_{intra}.For a system where the chromophores are not identical to each other, an approximate E can be used. The final energies will be dependent on the choice of approximate E provided for the expansion. Diagonalizing this final effective Hamiltonian will provide us with the energies for the (TT) states.

In Figure 7.10, we present the effect of the perturbative correction on top of the zero order (ZO) biexciton model. Since the chromophores do not have identical triplet excitations, the final energy depends on the choice of E in Equation 7.17. We plot the error for four different choice of E.

Even though adding the perturbative correction improves the results, it is not considered in our model. This is mainly because the error scale for the zeroth order model is much smaller than the energy gap we are interested in and the PT correction does not change the qualitative ordering of the states.

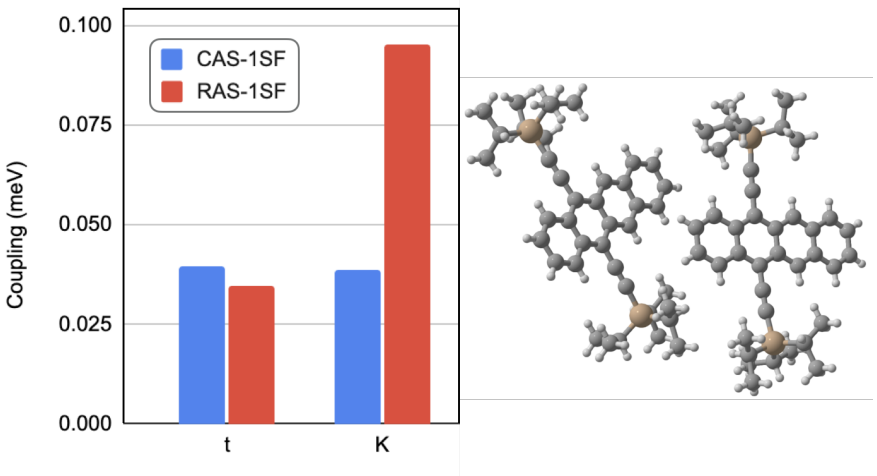


Figure 7.11: The exchange coupling constants from CAS-1SF and RAS-1SF for the TIPS-tetracene dimer.

7.7.3 TIPS-Tetracene

The TIPS tetracene crystal lattice has been an important molecule for the study of endothermic singlet fission.[327, 366] The biexciton exchange parameter has been reported [367] and DMRG-SCF calculation[340] has been used to compute the exchange interaction for the TIPS-Tetracene crystals. We report the biexciton exchange interaction using the CAS-1SF and also using the RAS-1SF where additional hole and particle excitations are considered. It can be seen from Figure 7.11 that the exciton interactions are comparable using the CAS calculation. Adding additional hole and particle excitations imcrease the exchange splitting (K) and is comparable to the experimental result.[367]

7.7.4 Dihedral Angle

Recent studies have suggested that changing the dihedral angle can help in the triplet dissociation process.[303, 347] We study the effect of dihedral angle on the K and t parameters using a simple tetracene dimer directly connected in the corner. We generate the molecular

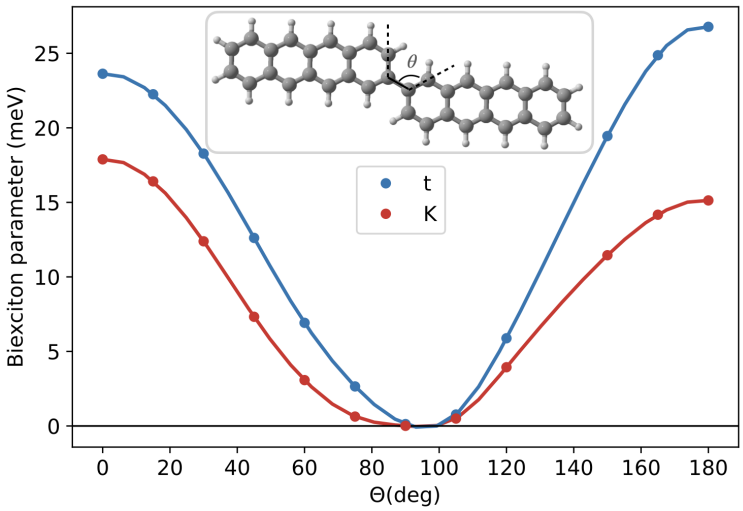


Figure 7.12: The exchange coupling constants for the directly connected tetracene dimer system as we change the dihedral angle between the two chromophores.

geometry for different dihedrals by optimizing the geometry using B3LYP/6-31G(d) at each dihedral angle while freezing the dihedral angle. We finally compute the K and t using the CAS-1SF method at each geometry. It can be seen from Figure 7.12 both exchange interaction and transfer interaction goes to zero at 90 degrees dihedral angle. The transfer integral is larger than the exchange for rest of the geometries we considered. Inclusion of a bridging linker in between the two chromophores can change the nature of the K and t and will be studied in future work.

7.7.5 Biexciton spectrum for large 1D system:Effect of system size

To understand the effect that the parameters have on the spectrum of larger chromophore assemblies, we generate the biexciton energy spectrum for a model 1D system. The exchange interaction (K) is taken to be 10 meV and the spectrum is generated for varying values of t. We first analyse the energy difference between lowest energy ^{1}(TT) and ^{5}(TT) states for a system with 40 chromophores.

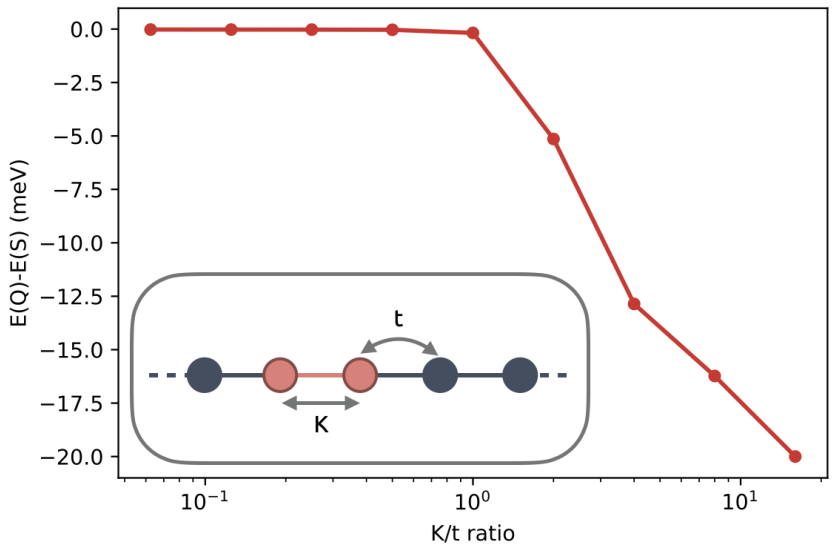


Figure 7.13: The spin gap between the ^{1}(TT) and ^{5}(TT) states for a model 1D system with 40 chromophores for different K vs. t values.

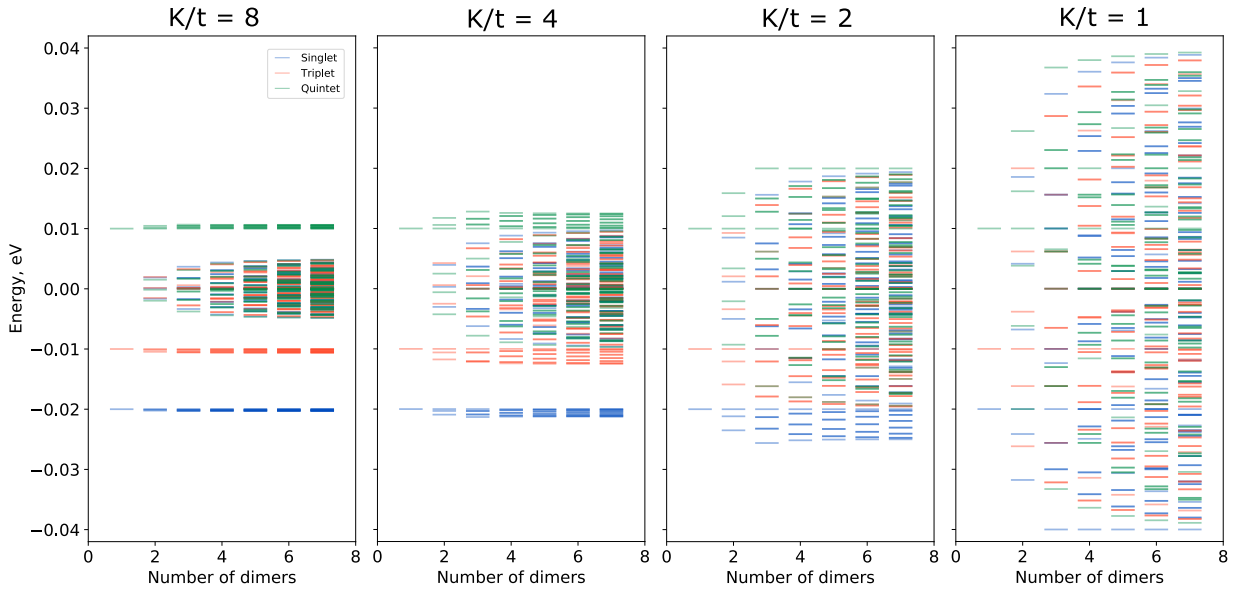


Figure 7.14: The biexciton spectrum for all three spin states is generated in the model space using a constant $K = 10meV$ value and varying the t parameter. Blue lines: ^{1}(TT). Red lines: ^{3}(TT). Green lines: ^{5}(TT).

In Figure 7.13, we can see that the spin gap is much lower when the triplet transfer term is larger than or comparable to the biexciton exchange interaction. Large K/t values imply that the exchange interaction between the biexcitons is larger than the tendency for the (TT) states to delocalize. For most of the systems, the values of K and t are comparable and hence the delocalization is favoured. In cases where K is much larger, the biexcitons tend to stay more localized. The exact ratio where the exciton transfer integral would dominate over exchange is debatable since there are other factors like vibrational motion and reorganization which can affect the biexciton dynamics.[323, 325] A kinetic model with the diabatic states for (TT) and (T...T) including both biexciton transfer and biexciton exchange can shed more light into the biexciton dynamics and will be investigated in future work.

As the hopping term tends to zero (right side of Figure 7.14), the biexcitons form band like structure around the ^{1}(TT),^{3}(TT), ^{5}(TT) states. The large band around 0 corresponds to the (T...T)-type states (Figure 7.14). This band is clearly the most dense, as there are a quadratic number of non-nearest neighbor biexcitons, whereas the more separated spin-state bands only involve nearest neighbor biexcitons, which are linear in the number of chromophores. As the hopping parameter, t, increases, biexciton delocalization begins to occur, broadening the spectrum and removing the gaps between the different spin states.

Chapter 8

Predicting Boundness of Biexcitons

8.1 Introduction

Singlet fission, a process of forming two low-energy triplet excitons from a single higher energy singlet exciton, carries the potential for significantly improving organic photovoltaic solar cells.[280, 281, 368, 369, 370, 371] The basic mechanism involves a spin allowed transition from a bright (one-electron transition) singlet excited state to a singlet-coupled pair of triplet states which are localized on neighboring chromophore monomers. Because this triplet-coupled state is predominantly a two-electron transition, it is referred to as a multi-exciton (ME) state, and consequentially, is difficult to observe experimentally[254, 372, 373, 374] or model theoretically.[375] This ME then decouples to form independent triplet states.

$$S_0S_0 \xrightarrow{h\nu} S_0S_1 \xrightarrow{k_1} (\mathrm{TT}) \xrightarrow{k_2} T + T \tag{8.1}$$

Being an interchromophoric process, the efficiency of singlet fission is related to the electronic coupling between chromophores.[280] One of the basic thermodynamic requirements for a

molecule to be singlet fission active is that the energy of the singlet state must be double or more of that of the triplet state, $E(S_1) \geq 2E(T_1)$. Linear polyacenes of four or more rings satisfy these conditions and can hence be used as singlet fission chromophores. Hence, singlet fission has been extensively studied in polyacene crystals and acene films.[254, 372, 373, 376, 377]

Recently, there have been increasing attention given to intramolecular singlet-fission dimers created by covalently linking two chromophores.[290, 378, 379, 380, 381, 382, 383, 384, 385] Because a covalent linkage can be modified by synthetic organic chemistry, it provides a much more controllable design parameter for optimizing chromophore couplings, compared to the more difficult task of modulating packing structure in molecular crystals. Contrary to the case of intermolecular singlet fission, intramolecular singlet fission systems have two mechanisms through which their triplet states can couple: through-space and through-bond.

For through-space couplings to be significant, there must exist overlap between the densities of the two chromophore units. This is because the triplets are only coupled through exchange. As the size of the organic linkage becomes larger, the chromophores are necessarily further apart and the through-space coupling decays exponentially. Thus, for many systems where the chromophore densities do not overlap, the through-space coupling will be negligible for intramolecular dimers.

In contrast, through-bond coupling can remain significant at larger distances if the organic linkage has delocalized frontier orbitals which can establish overlap between the chromophores. Conjugation (among carbon p orbitals) provides one route for establishing this delocalization, and we will focus exclusively on conjugated bridging linkers in this work.

The singlet fission process is much slower in dimers compared to crystals[386, 387] and both steps (k_1 and k_2 in Eq. 8.1) have considerable significance.[312, 384, 388] Recently,

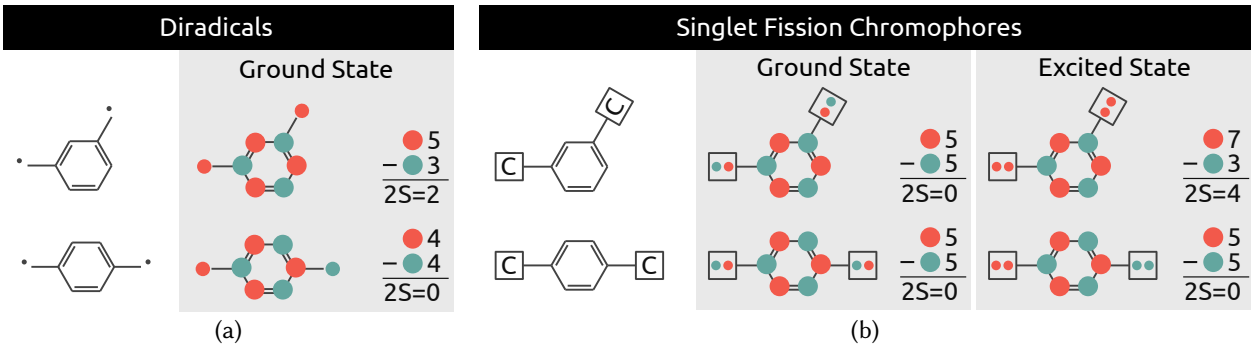


Figure 8.1: Illustration of the carbon labeling scheme. S is ground state spin calculated using Eq. 8.3. $S = |n_{\text{red}} - n_{\text{blue}}|/2$. (a) The ground state spin configuration of meta and para Xylene using Ising model. (b)The ground and excited state spin configuration of a singlet fission dimer.

Krylov *et al.* proposed a three-state kinetic model for singlet fission in which the importance of both steps in the overall rate is emphasized.[389] According to this model, the rate of the second step depends on the multi-exciton stabilizing energy as shown in Equation 8.2, $E_b = E[^5(\text{TT})] - E[(\text{TT})]$. If $E_b > 0$, the (TT) state is bound, and if $E_b < 0$, it is unbound.

$$k_2 = e^{\frac{-E_b}{2k_BT}} \tag{8.2}$$

The overall rate of the SF process cannot be discussed on the basis of E_b alone because according to this model, (TT) destabilization makes the second step faster but at the same time makes the first step slower. But a relationship between the triplet yield and boundness of the (TT) state can be drawn and is discussed later.

Having predominant ME character, the (TT) state is dark, making it experimentally difficult to observe. The nature of this short-lived correlated exciton pair (TT) is not fully understood. There has been significant effort spent to understand the mechanism of formation of the coupled triplet pair using computational chemistry.

Many of the computational approaches have applied multireference methods (ie., MRMP2,

CASPT2 or NEVPT2) to study the excited states.[253, 289, 331, 332, 390, 391] These provide straightforward description of the target states, although it has been mentioned that intruder states must be carefully avoided.[253] Using a quite different formalism, Yost et al. applied constrained DFT with configuration interaction (CDFT-CI) calculations to propose a model based on Marcus theory, which emphasized the importance of state to state energy alignment.[392] In addition to the energy requirements, several researchers have also focused on understanding the role non-adiabatic coupling plays in the creation of the ME.[228, 333, 384] Model Hamiltonian approaches have also been very effective in providing clearer insight into the singlet fission process.[308, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402] Parker and coworkers used the Active Space Decomposition strategy to derive an ab initio diabatic wavefunction basis, providing a rigorously obtained model Hamiltonian.[218] In our own group, we have demonstrated that an ab initio-based Heisenberg spin Hamiltonian can reproduce the ME spin state gaps.[66]

In the present work, we extend our previous study so as to map the singlet fission chromophores onto the simpler Ising Hamiltonian, $\hat{H} = -\sum J_{ij}\hat{S}_i^z\hat{S}_j^z$, for which simple analytical expressions allow us to make qualitative predictions without any computation. For the systems studied here, we assume only nearest-neighbor coupling such that $J_{ij} < 0$ (antiferromagnetic) when two sites are adjacent to each other and $J_{ij} = 0$ otherwise. We note that although a model space of spin configurations excludes explicit charge transfer, it can be implicitly folded into the effective interaction, J_{ij}, provided the overall charge transfer contribution is relatively small. Based on the ability to map the problem onto a spin lattice, we propose a simple qualitative rule (an application of Ovchinnikov's rule, *vide infra*) which relates the energetic ordering of the multiexciton spin states to the connectivity pattern of the bridge. *This allows one to predict without computation whether the through-bond coupling will lead to a bound or unbound* (TT) *state for intra-molecular singlet-fission chromophores*

with conjugated bridges.

In our recent work,[66] we used a Heisenberg Hamiltonian to characterize the (TT) state by mapping onto a spin lattice the localized[4, 5, 6] singly occupied ROHF orbitals of the chromophore's lowest energy triplet state. The characterization of the nominally closed shell singlet fission chromophores as diradicals was also found to be useful in earlier studies.[403, 404, 405] These localized singly occupied orbitals of each chromophore have the highest amplitude near the center of the chromophore's edges and are plotted in Fig. 8.2. Because these orbitals are spatially proximal with significant overlap, the coupling within a chromophore is always large and antiferromagnetic. However, because there is significantly less overlap between the orbitals of distinct chromophores, the coupling between chromophores is much smaller and often (but not always) antiferromagnetic. In absence of any interaction between the two chromophores, the singlet ^{1}(TT), triplet ^{3}(TT) and quintet ^{5}(TT) states will be degenerate. The introduction of coupling then, splits this spin-state degeneracy, leading to either a ferromagnetic or antiferromagnetic order. If we consider coupling between the two triplet states via direct exchange (ferromagnetic) only, the order would necessarily be ^{5}(TT) $<$ ^{3}(TT) $<$ ^{1}(TT). However, due to the presence of implicit super exchange type coupling with low lying charge transfer states (which is generally antiferromagnetic in nature) this order can vary.

8.2 Ovchinnikov's Rule

The interest in predicting spin state orderings based on topological considerations (rather than explicit calculation) has a long history. Ovchinnikov's rule[406] relates the spin, S, of the ground state of a bipartite, antiferromagnetic spin lattice to the total number of sites

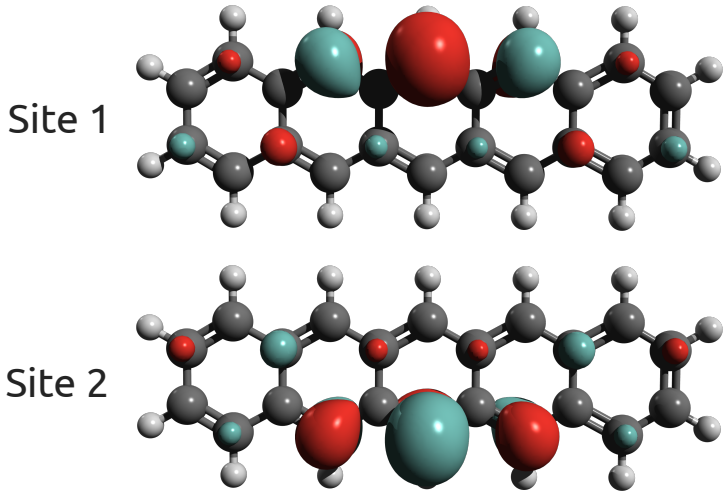


Figure 8.2: Two singly occupied molecular orbitals for triplet pentacene choromophore molecules. Boys localization[4, 5, 6] mixes the original ROHF semicanonical singly occupied orbitals to yield the above orbitals, which are localized toward the center of the outside edges of the acene.

with α (n_α) and β (n_β) spins.

$$2S = |n_\alpha - n_\beta| \tag{8.3}$$

This comes from the fact that the lowest energy spin configuration is two-fold degenerate (by m_s symmetry) with α and β spins arranged alternately on the lattice. Because the requirement to arrange spins alternately can only be met for bipartite carbon networks, this rule only holds for alternate hydrocarbons (no odd-numbered rings). Ovchinnikov follows a procedure very similar to Leib and Mattis for proving that the Ising model has a ground state spin as given by Equation 8.3.[407] An illustration of this rule for an example organic diradical is shown in Figure 8.1(a). In the top panel of Fig. 8.1(a), the radicals are attached such that the ground state is a triplet state, whereas the bottom panel has a singlet ground state. In this paper, we extend this rule, which is commonly applied in organic diradicals, to predict the qualitative energy ordering of the ^{5}(TT), ^{3}(TT), and ^{1}(TT) ME states in covalently linked singlet fission chromophores. This order is important as it determines

whether dissociation to separate triplets will be uphill or downhill.

In order to extend Ovchinnikov's rule to excited states, we consider the energies of higher-energy spin-configurations (those with one or two parallel spin neighbors) and use the same rule to predict the ordering of the spin states of the ME state. Because all nearest-neighbor J_{ij} values are antiferromagnetic, all singlet fission dimers will correctly be predicted to have singlet ground states, regardless of the parity of the bridge (linker) connection. This is demonstrated in the middle panel of Figure 8.1(b).

If we were to consider the small subset of higher energy configurations in which there is a single unfavorable interaction between two neighboring carbons with the same spin, one would expect to see two energy levels: i) one lower energy level in which the parallel interaction occurs on a chromophore, and ii) a higher energy level in which the parallel interaction occurs on the bridge. These two distinct types of spin configurations relate to the physical picture of either having a triplet state on the chromophore or the bridge. Because singlet-fission chromophores must have low-lying triplet states to satisfy the basic energetic requirements discussed above, a triplet state is much more energetically favorable to exist on the chromophore than the bridge. In other words, the exchange coupling constant between chromophore orbitals will be much smaller than that between bridge orbitals, $J_{ij}^{C} << J_{ij}^{B}$. The experimental J_{ij} values for benzene and naphthalene (examples of bridges used in this study) are roughly around 84 and 60 kcal/mol, whereas tetracene and pentacene (chromophores used in this study) are around 29 and 20 kcal/mol, respectively.[158, 408, 409, 410]

In contrast to the "single exciton" triplet states just described, the ME state must be modelled by considering spin configurations in which there are *two* instances of parallel interactions. Due to the relatively small value of J_{ij}^{C}, the lowest energy ME states will be those which have parallel spins on the chromophores. This situation is illustrated in the third panel of Figure 8.1(b). Here we see that unlike the ground state for which Ovchinnikov's

rule predicts all singlet fission dimers to be singlets, Ovchinnikov's rule predicts the lowest energy ME states to be either a singlet state or a quintet state, depending on the parity of the bridging connection. *This is the central result of this work, which implies that the boundedness of the* (TT) *state can be predicted based on the parity of the bridge connection.*

To test our hypothesis, we created a data set by combining different types of bridging ligands and chromophores (Figure 8.3). We studied three types of bridge molecules, namely benzene, naphthalene and conjugated linear polyacenes. The bridges are categorized (based on our proposed Ovchinnikov's rule) as either ferromagnetic or antiferromagnetic bridges. For example, for benzene, the bridge is antiferromagnetic if ortho or para connected, `B(o)` or `B(p)`, and ferromagnetic if meta connected, `B(m)`. For the linear polyacene bridges, an odd number of carbon atoms between the bridge, `O(n)`, leads to ferromagnetic coupling, while an even number, `E(n)`, leads to antiferromagnetic coupling. In both cases n stands for the number of double bonds present in the bridge. Because the chromophores themselves could potentially be linked in a number of ways, we have included connections for both tetracene and pentacene at different positions, also shown in Figure 8.3.

8.3 Systems and Method

Since the (TT) is a doubly excited state with respect to the ground state, only quantum chemistry methods capable of describing multiple electronic excitations can be used. For our calculations, we have used the RAS-SF method,[411] which has been used extensively for studying singlet fission systems.[333, 388, 390] Using the ROHF high-spin reference, RAS-2SF calculations were performed to compute the relative orderings of the multi-exciton spin states, shown in Table 8.1. Despite not including dynamical correlation, RAS-SF has been shown to be reliable for predicting spin states of organic polyradicals.[61] Nonetheless,

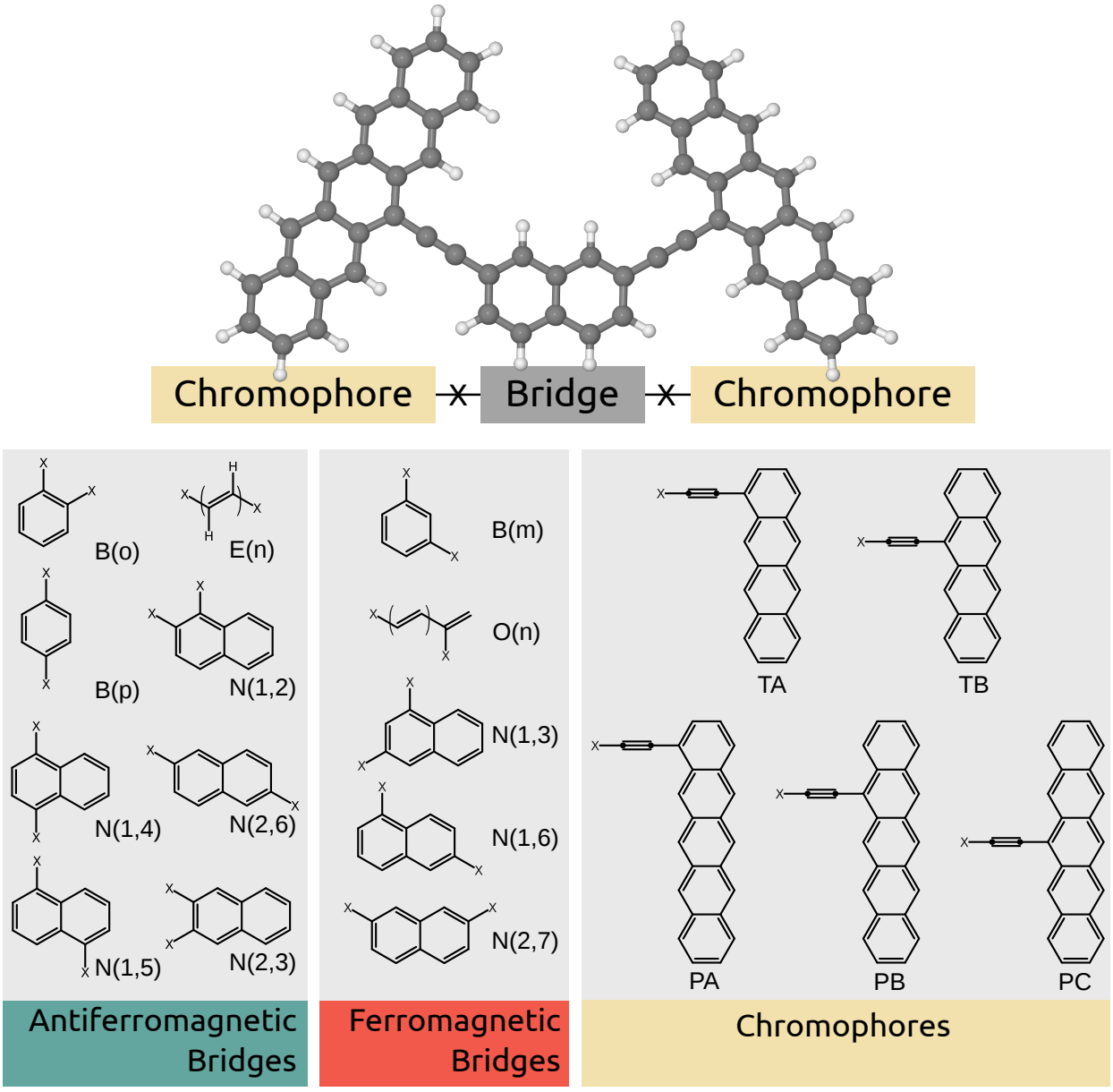


Figure 8.3: Chromophores and bridge monomers used to construct the series of covalently bound singlet-fission chromophore dimers. Molecular models are obtained by taking a pair of each chromophore, and attaching them to each of the bridge monomers. The PC-N(2,7) molecule is shown as an example.

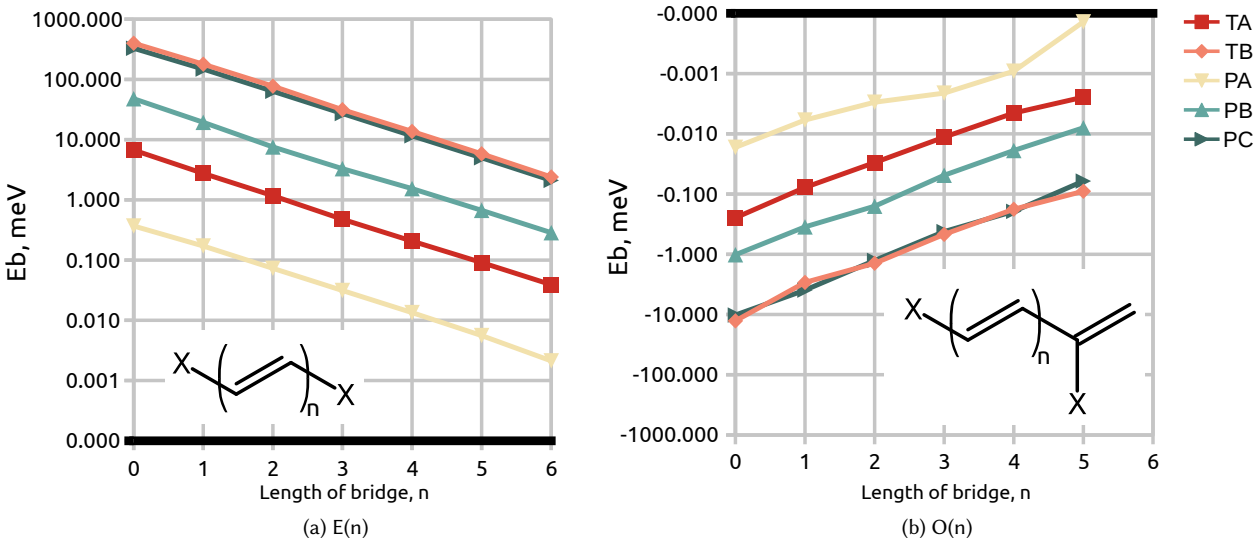


Figure 8.4: Logarithm scale plot of the binding energies, E_b, as a function of number of polyene bridge units for (a) the `E(n)` bridge , and (b) the `O(n)` bridge. Exciton binding energies obtained from the RAS-2SF/6-31g* energies of the $^5(TT)$ and $^1(TT)$ states.

we have also performed a series of calculations which include dynamical correlation perturbatively, such as SF-CAS(S), SF-CAS(2x), and SF-CAS(3x).[360, 412, 413] Because the inclusion of dynamical correlation did not change any of the results qualitatively, we discuss only the RAS-SF results in the main text, providing the perturbative results in the Supplementary Information. All RAS-SF single point energy calculations were performed using the ground state geometries optimised at the B3LYP/6-31G* level of theory. In instances where the system might have preferred a π-stacked geometry (when the bridge is sufficiently long to allow the chromophores to fold back on each other) care was taken to converge to the relatively flat, un-folded structure. This is to be sure that we are only characterizing the through-bond interactions, and not the through-space interactions. All calculations were performed using QChem 4.[222]

Table 8.1: Exciton binding energies of all combination of the covalently bound singlet fission dimers formed from the set of chromophores and bridges shown in Figure 8.3. If the exciton binding energy is positive, the (TT) state is bound, and unbound otherwise.

meV	TA	TB	PA	PB	PC
	Ferromagnetic (odd) Bridges				
O(0)	-0.248	-12.768	-0.017	-1.016	-10.211
O(1)	-0.078	-2.937	-0.006	-0.351	-3.944
O(2)	-0.030	-1.408	-0.003	-0.159	-1.263
O(3)	-0.011	-0.467	-0.002	-0.049	-0.142
O(4)	-0.005	-0.179	-0.001	-0.019	-0.194
O(5)	-0.002	-0.090	-0.003	-0.008	-0.061
B(m)	-0.090	-5.784	-0.006	-0.367	-4.597
N(1,3)	-0.071	-3.896	-0.006	-0.362	-4.541
N(1,6)	-0.043	-2.145	-0.004	-0.172	-1.379
N(2,7)	-0.038	-2.530	-0.003	-0.174	-2.212
	Antiferromagnetic (even) Bridges				
E(0)	6.700	399.164	0.369	47.762	328.825
E(1)	2.755	178.912	0.171	19.366	148.263
E(2)	1.172	77.129	0.073	7.558	63.429
E(3)	0.477	31.452	0.031	3.293	27.107
E(4)	0.207	13.693	0.013	1.539	11.582
E(5)	0.091	5.815	0.006	0.671	4.996
E(6)	0.039	2.397	0.002	0.286	2.105
B(o)	0.526	33.755	0.031	3.038	37.918
B(p)	0.869	51.799	0.053	2.596	42.247
N(1,2)	0.810	53.525	0.051	1.506	61.618
N(1,4)	1.008	39.231	0.062	2.962	25.767
N(1,5)	0.032	1.263	0.002	0.212	1.002
N(1,7)	0.133	8.299	0.007	0.838	6.596
N(2,3)	0.175	14.390	0.010	0.421	21.808
N(2,6)	0.301	18.846	0.012	1.443	15.102

8.4 Discussion

As can be seen from the RAS-SF data in Table 8.1, *our hypothesis is supported in all calculations performed*. For each calculation using a Ferromagnetic bridge, the ^{5}(TT) is lower in energy than the (TT) state, indicating an unbound ME, consistent with the hypothesis. Similarly, all bridge linkers predicted to be antiferromagnetic exhibit a bound (TT) state, with the ^{5}(TT) higher in energy than the (TT) state. The nature of the (TT) state did not change by changing the connectivity of the chromophore to the bridge, but the magnitude of the exciton binding energy was larger when the connectivity was towards the center of the chromophore. Because both singly occupied chromophore orbitals are somewhat localized toward the middle of the acene (see Figure 8.2), as the attachment point moves from the middle of the acene toward the end, the coupling (and thus exciton binding energy) decreases. This can be observed by inspecting the values for `PA`, `PB`, and `PC`, and also `TA` and `TB`.

The systematic set of polyene bridges `O(n)` and `E(n)`, allow us to inspect the distance dependence of the through-bond interactions, illustrated in Fig. 8.4. Here, we observe an exponential decay of exciton binding energy with respect to the number of C atoms between the chromophores. Although attachment position on the chromophore affects the magnitude of the coupling (i.e., `PA` is much less coupled than `PC`), it does not seem to noticably affect the rate of decay, which is ultimately determined by the bridging ligand.

Our studies agree with previous experimental and theoretical studies. Bradforth *et al* studied a covalently linked SF dimer using ortho-connected benzene as a linker (BET-B).[312] It was seen that in the case of a dimer with minimal overlap, BET-B (`B(o)` linker in our data set), the (TT) state does not separate into triplet states unless there is intermolecular triplet energy transfer. Using our model it can be clearly seen that the (TT) is bound in this case, which makes separating the (TT) states to two triplet states uphill and hence unfavorable.

Tykwinski and coworkers showed that singlet fission dimers with meta connectivity lead to good triplet quantum yield.[386, 387] This can be attributed in part to the (TT) state in these examples being unbound and the conversion from (TT) to separate triplet states being favourable. Our results are complimentary to the recent theoretical study by Nakano and coworkers,[405, 414] despite our approaches being rather different. In their approach they used Fermi's Golden Rule and a Green's function formalism to study the direct (through-space) and through-bond electronic couplings separately. They discuss the rate of the first step using these electronic couplings, and find that the electronic coupling for `B(m)` is very low and hence the first step is slower. With the first step controlling the kinetics in this system, their results correlated well with experimental rate constants. Our calculations shed light on a different aspect of these results, by focusing solely on the energetics of the second step. Although our results predict the `B(m)` to support a relatively fast exciton dissociation compared to a `B(p)` bridge, this is already so fast that the overall rate is determined by the first step, making Nakano's approach more predictive for this system.

Due to the steric limitation, some of the dimers (`B(o), N(1,2), N(2,3)`) do have through-space coupling comparable to the through-bond couplings. Any direct overlap between the densities of the two chromophores will have an additional antiferromagnetic (binding) contribution. Since each of these dimers are already antiferromagnetic, the through-bond coupling does not change the sign, and the rule is unchanged. Since the calculations in this study do not employ high level dynamic correlation, the quantitative comparison on binding energies cannot be concluded without more accurate calculations.

In addition to the systems in our test set, we have also investigated two situations where our rule is not expected to by effective, or simply cannot be applied. In this first problem case, we studied the anthracene molecule as a chromophore. Because anthracene is shorter, its triplet state is higher in energy and can easily become degenerate with the triplet state on a

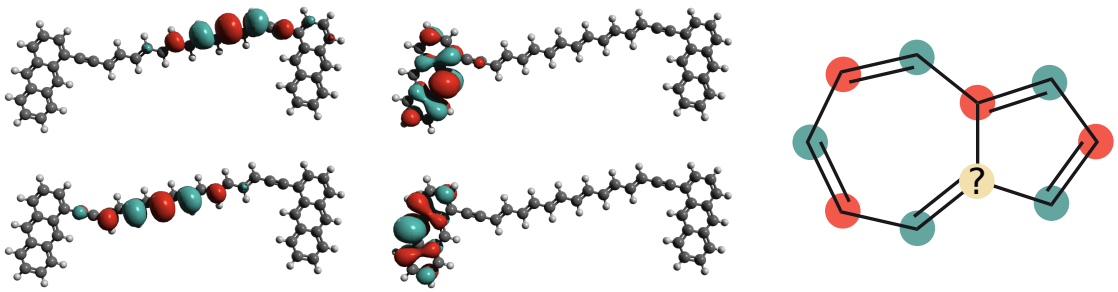


Figure 8.5: Exceptions to the rule: (a) The four singly occupied orbitals of the ROHF quintet state for anthracene system with a long bridge. Orbitals are not localized on chromophores. (b) Azulene molecule as an example of a non alternate molecule. The most stable configuration is obtained when opposite color are next to each other, which is not possible in case of azulene

longer bridge. In such a situation, the bridge and chromophore have comparable couplings ($J_{ij}^{C} \approx J_{ij}^{B}$), and the triplet states delocalize across the whole molecule, which destroys the local spin lattice picture. In fact, even the base assumption of a 4-site spin lattice is no longer applicable. As shown in Fig. 8.5(a), the four singly occupied orbitals of the quintet ROHF solution are no longer localized on just the chromophores. The ROHF orbital optimization procedure determines that it is more energetically favorable to unpair the electrons on the bridge first, then the chromophore. In this scenario, a 4-site model is no longer appropriate, and more electronic structure calculations may be necessary to determine the ordering of states.

As mentioned above, Ovchinnikov's rule is only valid for bipartite carbon networks, i.e., alternate hydrocarbons. Violating this requirement, azulene is a non-alternate system and is not able to be characterized with Ovchinnikov's rule (Figure 8.4(b). Because it has odd-numbered rings, it is not possible to identify configurations which have only favorable interactions. As expected, the excited state Ovchinnikov's rule fails for non-alternate bridges such as Azulene, and numerical computations are needed to predict the spin state ordering.

8.5 Conclusion

To conclude, we have proposed a rule based on a spin lattice model to predict the relative ordering of the various spin states (^{1}(TT), ^{3}(TT), and ^{5}(TT)) of the key multiexciton intermediate in singlet fission, which is an excited state extension of Ovchinnikov's rule. Because the energy of a pair of uncoupled triplet states (the final product of singlet fission) is bounded by the ^{1}(TT) and ^{5}(TT) energies, predicting the energetic order of ^{1}(TT) and ^{5}(TT) indicates whether $^1(\mathrm{TT}) \rightarrow T + T$ proceeds uphill or downhill. In other words, this rule allows one to predict (without computation) whether the multiexciton state is bound or unbound with respect to separated triplets. While this rule was found to be 100% predictive in our numerical test set, there are a few caveats limiting the general application of this rule: (a) This rule only applies when through-space interactions are negligible compared to through-bond. (b) only bipartite bridges (or bridges without odd numbered rings) can be characterized with this rule (c) the energetics must be appropriate, meaning that the singlet-triplet gap of the chromophores must be smaller than that of the bridge. If these conditions are met, then the spin state ordering can be predicted based solely on the parity (even or odd-ness) of the bridge, providing insight into the design of new singlet fission dimers.

Chapter 9

Conclusion

In this dissertation, we explore the applicability of tensor product states for strongly correlated systems. We partition the active space orbitals into separate clusters and form many body states in each clusters. The final wavefunction is formed in the basis of tensor product of many body states. In Chapter 2, we introduce the TPS basis and present examples for computing matrix elements and derive equations for the cluster mean field method.

In Chapters 3, we introduce a method named cluster many-body expansion (cMBE) where we form a many body expansion based approach using tensor product states for capturing correlation energy. The cMBE method provides improved results compared to traditional RHF based MBE at lower orders. We provide efficient pruning technique for larger active spaces. cMBE provides accurate energies for strongly correlated systems like chromium dimer and large carbon nanosheet.

In Chapter 4 and 5, we have introduced a new selected CI method using tensor products of cluster states as the basis. We perform the SCI algorithm in a basis of tensor products of cluster states instead of Slater determinants. This allows us to include part of the correlation within the clusters already using the cluster basis. We present numerical data to show how TPSCI significantly outperforms traditional SCI methods in case of strongly correlated systems. For the ground state energy, using a embedded Schmidt decomposition helps to truncate the cluster states and hence storage of the transition density matrices is feasible. For the excited states, we use the excited configurations formed from the cMF Fock matrix.

The TPSCI+PT2 provide good estimate of excitation energies with respect to extrapolated results. In the future, the TPSCI method will be applied on polymetallic systems. Orbital optimization with the TPSCI method is also an interesting future direction to provide CASSCF values for large active spaces.

In Chapter 6, we develop a size-extensive formulation of a truncated TPS method. We demonstrate why size-extensivity is an important property for a many body method by numerically testing a 1D spin lattice. We use RSPT to capture correlations on the reference NB0 wavefunction. We demonstrate how higher order PT corrections improve the accuracy of the method. We also form a linearized coupled cluster method in the TPS basis (TPS-LCC or $NB0(2)\infty$) and demonstrate that this can be a good approximate method for single and multi-reference TPS cases. In the future, the TPS-LCC approach would be applied to the molecular Hamiltonian using an improved dense code.

In Chapter 7, we drive a simple biexciton model for correlated triplet pair states in singlet fission. We demonstrate that the biexciton model provides accurate results for a system with seven tetracene molecules taken from a crystal structure. We use the TPS matrix elements between different biexciton states to derive biexciton transfer parameter(t) and biexciton exchange parameter (K). Depending on these two parameters, the biexciton can stay as a trap state (large antiferromagnetic K) *vs.* the biexciton can dissociate into separate triplets (large t). We investigate how these parameters vary with molecular packing in tetracene chromophores and provide optimal packing for improved biexciton separation which is key for improved solar cell efficiencies.

In Chapter 8, we provide a simple model based on Ovchinnikov's rule for the spin state ordering of the biexciton states for chromophores that are connected by a π conjugated bridging molecule. We demonstrate this by taking a wide variety of bridging molecule with odd and even C atoms and show that for the cases with odd C atoms, the ^{5}(TT) state

is higher in energy as compared to ^{1}(TT) and improve biexciton dissociation and thereby increase final triplet yield. Our hypothesis was later experimentally proved by Sakai et al. [296] independently where they saw improved triplet yield for meta linked chromophores.

The TPS methods we introduce in this work can be applied to study a variety of scientific problems. It can be also be used along with an energy decomposition analysis for investigating the nature of bonding in sophisticated chemical species.[415] Analysis of polymettaic systems like the FeMoCo system using TPS wavefunction is another attractive future direction as far as systems are considered. A multi-reference TPS-LCC would be an ideal candidate for this system. The spin exchange interactions can also be extracted for these polymetallic systems.

This work is one of the primary exploratory works on TPS based approaches and sets the basis for many possible future developments. One interesting future direction is to extend the TPSCI or cMBE to include orbital optimization which can provide accurate excitation energies. This will involve calculation of 1-rdm and 2-rdm of the wavefunction and the orbitals can be optimized similar to recently improved CASSCF solver. Another important direction is to introduce spin orbit effects into the TPS based method which would allow accurate study of compounds containing heavy elements where strong correlation is often a major challenge. With this work I hope I was able to convince the reader the importance of the TPS based methods in quantum chemistry and their potential future use to solve major scientific problems.

Bibliography

[1] Bayliss, S. L.; Laorenza, D. W.; Mintun, P. J.; Kovos, B. D.; Freedman, D. E.; Awschalom, D. D. Optically addressable molecular spins for quantum information processing. *Science* **2020**, *370*, 1309–1312.

[2] Eriksen, J. J. et al. The Ground State Electronic Energy of Benzene. *The Journal of Physical Chemistry Letters* **2020**, *11*, 8922–8929, PMID: 33022176.

[3] Stoudenmire, E.; White, S. R. Studying Two-Dimensional Systems with the Density Matrix Renormalization Group. *Annu. Rev. Condens. Matter Phys.* **2012**, *3*, 111–128.

[4] Boys, S. Construction of Some Molecular Orbitals to Be Approximately Invariant for Changes from One Molecule to Another. *Rev. Mod. Phys.* **1960**, *32*, 296–299.

[5] Subotnik, J. E.; Shao, Y.; Liang, W.; Head-Gordon, M. An efficient method for calculating maxima of homogeneous functions of orthogonal matrices: applications to localized occupied orbitals. *J. Chem. Phys.* **2004**, *121*, 9220–9.

[6] Subotnik, J. E.; Sodt, A.; Head-Gordon, M. Localized orbital theory and ammonia triborane. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5522–30.

[7] Gaita-Ariño, A.; Luis, F.; Hill, S.; Coronado, E. Molecular spins for quantum computation. *Nature Chemistry* **2019**, *11*, 301–309.

[8] Thakar, K.; Lodha, S. Optoelectronic and photonic devices based on transition metal dichalcogenides. *Materials Research Express* **2020**, *7*, 014002.

[9] Coropceanu, V.; Chen, X.-K.; Wang, T.; Zheng, Z.; Brédas, J.-L. Charge-transfer electronic states in organic solar cells. *Nature Reviews Materials* **2019**, *4*, 689–707.

[10] Pillai, S.; Ravensbergen, J.; Antoniuk-Pablant, A.; Sherman, B. D.; van Grondelle, R.; Frese, R. N.; Moore, T. A.; Gust, D.; Moore, A. L.; Kennis, J. T. M. Carotenoids as electron or excited-state energy donors in artificial photosynthesis: an ultrafast investigation of a carotenoporphyrin and a carotenofullerene dyad. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4775–4784.

[11] Fan, Q.; Yan, L.; Tripp, M. W.; Krejčí, O.; Dimosthenous, S.; Kachel, S. R.; Chen, M.; Foster, A. S.; Koert, U.; Liljeroth, P.; Gottfried, J. M. Biphenylene network: A non-benzenoid carbon allotrope. *Science* **2021**, *372*, 852–856.

[12] Du, Q.-S.; Tang, P.-D.; Huang, H.-L.; Du, F.-L.; Huang, K.; Xie, N.-Z.; Long, S.-Y.; Li, Y.-M.; Qiu, J.-S.; Huang, R.-B. A new type of two-dimensional carbon crystal prepared from 1,3,5-trihydroxybenzene. *Sci. Rep* **2017**, *7*, 40796.

[13] Liu, D.; Kim, E.; Weck, P. F.; Tománek, D. Strain-controlled magnetic ordering in 2D carbon metamaterials. *Carbon* **2020**, *161*, 219–223.

[14] Pelzer, K.; Greenman, L.; Gidofalvi, G.; Mazziotti, D. A. Strong Correlation in Acene Sheets from the Active-Space Variational Two-Electron Reduced Density Matrix Method: Effects of Symmetry and Size. *The Journal of Physical Chemistry A* **2011**, *115*, 5632–5640, PMID: 21563790.

[15] Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* **1926**, *28*, 1049–1070.

[16] Szabo, A.; Ostlund, N. S. *Modern quantum chemistry: introduction to advanced electronic structure theory*; Courier Corporation, 2012.

[17] Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; John Wiley & Sons, 2014.

[18] Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34*, 1293–1322.

[19] Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*, 89–110.

[20] Fock, V. N{ä}herungsmethode zur L{ö}sung des quantenmechanischen Mehrk{ö}rperproblems. *Zeitschrift f{ü}r Phys.* **1930**, *61*, 126–148.

[21] Hall, G. G. The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials. *Proc. R. Soc. London A Math. Phys. Eng. Sci.* **1951**, *205*, 541–552.

[22] Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[23] Knowles, P.; Handy, N. A new determinant-based full configuration interaction method. *Chemical Physics Letters* **1984**, *111*, 315 – 321.

[24] Knowles, P. J.; Handy, N. C. Unlimited full configuration interaction calculations. *The Journal of Chemical Physics* **1989**, *91*, 2396–2398.

[25] Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics* **1980**, *48*, 157–173.

[26] Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *The Journal of Physical Chemistry* **1990**, *94*, 5483–5488.

[27] Finley, J.; Åke Malmqvist, P.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chemical Physics Letters* **1998**, *288*, 299–306.

[28] Siegbahn, P. E. M. The externally contracted CI method applied to N2. *Int. J. Quantum Chem.* **1983**, *23*, 1869–1889.

[29] Werner, H.; Reinsch, E. The self-consistent electron pairs method for multiconfiguration reference state functions. *The Journal of Chemical Physics* **1982**, *76*, 3144–3156.

[30] White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **1992**, *69*, 2863–2866.

[31] Chan, G. K.-L.; Head-Gordon, M. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *The Journal of Chemical Physics* **2002**, *116*, 4462–4476.

[32] Deustua, J. E.; Shen, J.; Piecuch, P. Converging High-Level Coupled-Cluster Energetics by Monte Carlo Sampling and Moment Expansions. *Phys. Rev. Lett.* **2017**, *119*, 223003.

[33] Kinoshita, T.; Hino, O.; Bartlett, R. J. Coupled-cluster method tailored by configuration interaction. *The Journal of Chemical Physics* **2005**, *123*, 074106.

[34] Deustua, J. E.; Magoulas, I.; Shen, J.; Piecuch, P. Communication: Approaching exact quantum chemistry by cluster analysis of full configuration interaction quantum Monte Carlo wave functions. *The Journal of Chemical Physics* **2018**, *149*, 151101.

[35] Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 54106.

[36] Cleland, D.; Booth, G. H.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, 41103.

[37] Purwanto, W.; Zhang, S.; Krakauer, H. An auxiliary-field quantum Monte Carlo study of the chromium dimer. *The Journal of Chemical Physics* **2015**, *142*, 064302.

[38] Huron, B.; Malrieu, J. P.; Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **1973**, *58*, 5745.

[39] Caffarel, M.; Applencourt, T.; Giner, E.; Scemama, A. *Recent Progress in Quantum Monte Carlo*; Chapter 2, pp 15–46.

[40] Garniron, Y.; Scemama, A.; Loos, P.-F.; Caffarel, M. Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory. *J. Chem. Phys.* **2017**, *147*, 034101.

[41] Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.

[42] Li, J.; Otten, M.; Holmes, A. A.; Sharma, S.; Umrigar, C. J. Fast Semistochastic Heat-Bath Configuration Interaction. *J. Chem. Phys.* **2018**, *148*, 214110.

[43] Schriber, J. B.; Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106.

[44] Wang, Z.; Li, Y.; Lu, J. Coordinate Descent Full Configuration Interaction. *Journal of Chemical Theory and Computation* **2019**, *15*, 3558–3569, PMID: 31042383.

[45] Liu, W.; Hoffmann, M. R. iCI: Iterative CI toward full CI. *Journal of Chemical Theory and Computation* **2016**, *12*, 1169–1178, PMID: 26765279.

[46] Zhang, N.; Liu, W.; Hoffmann, M. R. Iterative Configuration Interaction with Selection. *Journal of Chemical Theory and Computation* **2020**, *16*, 2296–2316, PMID: 32069046.

[47] Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112.

[48] Greer, J. C. Estimating full configuration interaction limits from a Monte Carlo selection of the expansion space. *J. Chem. Phys.* **1995**, *103*, 1821–1828.

[49] Sharma, S.; Holmes, A. A.; Jeanmairet, G.; Alavi, A.; Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **2017**, *13*, 1595–1604.

[50] Nesbet, R. K. Atomic Bethe-Goldstone Equations. I. The Be Atom. *Phys. Rev.* **1967**, *155*, 51–55.

[51] Nesbet, R. K. Atomic Bethe-Goldstone Equations. II. The Ne Atom. *Phys. Rev.* **1967**, *155*, 56–58.

[52] Nesbet, R. K. Atomic Bethe-Goldstone Equations. III. Correlation Energies of Ground States of Be, B, C, N, O, F, and Ne. *Phys. Rev.* **1968**, *175*, 2–9.

[53] Zimmerman, P. M. Incremental full configuration interaction. *The Journal of Chemical Physics* **2017**, *146*, 104102, Publisher: American Institute of Physics.

[54] Zimmerman, P. M. Singlet–Triplet Gaps through Incremental Full Configuration Interaction. *The Journal of Physical Chemistry A* **2017**, *121*, 4712–4720, Publisher: American Chemical Society.

[55] Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *The Journal of Chemical Physics* **2017**, *146*, 224104, Publisher: American Institute of Physics.

[56] Eriksen, J. J.; Gauss, J. Many-Body Expanded Full Configuration Interaction. I. Weakly Correlated Regime. *Journal of Chemical Theory and Computation* **2018**, *14*, 5180–5191, Publisher: American Chemical Society.

[57] Eriksen, J. J.; Gauss, J. Many-Body Expanded Full Configuration Interaction. II. Strongly Correlated Regime. *Journal of Chemical Theory and Computation* **2019**, *15*, 4873–4884, Publisher: American Chemical Society.

[58] Eriksen, J. J.; Gauss, J. Generalized Many-Body Expanded Full Configuration Interaction Theory. *The Journal of Physical Chemistry Letters* **2019**, *10*, 7910–7915, Publisher: American Chemical Society.

[59] Eriksen, J. J.; Lipparini, F.; Gauss, J. Virtual Orbital Many-Body Expansions: A Possible Route towards the Full Configuration Interaction Limit. *The Journal of Physical Chemistry Letters* **2017**, *8*, 4633–4639, Publisher: American Chemical Society.

[60] Krylov, A. Spin-flip configuration interaction: an electronic structure model that is both variational and size-consistent. *Chem. Phys. Lett.* **2001**, *350*, 522–530.

[61] Casanova, D.; Head-Gordon, M. Restricted active space spin-flip configuration interaction approach: theory, implementation and examples. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9779–9790.

[62] Zimmerman, P. M.; Bell, F.; Goldey, M.; Bell, A. T.; Head-Gordon, M. Restricted active space spin-flip configuration interaction: Theory and examples for multiple spin

flips with odd numbers of electrons. *The Journal of Chemical Physics* **2012**, *137*, 164110.

[63] Bell, F.; Zimmerman, P. M.; Casanova, D.; Goldey, M.; Head-Gordon, M. Restricted active space spin-flip (RAS-SF) with arbitrary number of spin-flips. *Phys. Chem. Chem. Phys.* **2013**, *15*, 358–366.

[64] Pokhilko, P.; Krylov, A. I. Effective Hamiltonians derived from equation-of-motion coupled-cluster wave functions: Theory and application to the Hubbard and Heisenberg Hamiltonians. *The Journal of Chemical Physics* **2020**, *152*, 094108.

[65] Mayhall, N. J.; Head-Gordon, M. Computational Quantum Chemistry for Multiple-Site Heisenberg Spin Couplings Made Simple: Still Only One Spin-Flip Required. *J. Phys. Chem. Lett* **2015**, *6*, 1982–1988.

[66] Mayhall, N. J. From Model Hamiltonians to ab Initio Hamiltonians and Back Again: Using Single Excitation Quantum Chemistry Methods To Find Multiexciton States in Singlet Fission Materials. *Journal of Chemical Theory and Computation* **2016**, *12*, 4263–4273, PMID: 27472260.

[67] Ivanic, J.; Ruedenberg, K. Identification of deadwood in configuration spaces through general direct configuration interaction. *Theoretical Chemistry Accounts* **2001**, *106*, 339–351.

[68] Cuthill, E.; McKee, J. Reducing the Bandwidth of Sparse Symmetric Matrices. Proceedings of the 1969 24th National Conference. New York, NY, USA, 1969; p 157–172.

[69] Jiménez-Hoyos, C. A.; Scuseria, G. E. Cluster-based mean-field and perturbative description of strongly correlated fermion systems: Application to the one- and two-dimensional Hubbard model. *Phys. Rev. B* **2015**, *92*, 085101.

[70] Hermes, M. R.; Gagliardi, L. Multiconfigurational Self-Consistent Field Theory with Density Matrix Embedding: The Localized Active Space Self-Consistent Field Method. *J. Chem. Theory Comput.* **2019**, *15*, 972–986.

[71] Hermes, M. R.; Gagliardi, L. The Variational Localized Active Space Self-Consistent Field Method. *arXiv:2003.02995 [physics]* **2020**,

[72] Parker, S. M.; Shiozaki, T. Communication: Active Space Decomposition with Multiple Sites: Density Matrix Renormalization Group Algorithm. *J. Chem. Phys.* **2014**, *141*, 211102–211102.

[73] Parker, S. M.; Seideman, T.; Ratner, M. A.; Shiozaki, T. Communication: Active-Space Decomposition for Molecular Dimers. *J. Chem. Phys.* **2013**, *139*, 021108.

[74] Jiménez-Hoyos, C. A.; Scuseria, G. E. Cluster-based mean-field and perturbative description of strongly correlated fermion systems: Application to the one- and two-dimensional Hubbard model. *Phys. Rev. B* **2015**, *92*, 085101.

[75] Abraham, V.; Mayhall, N. J. Selected Configuration Interaction in a Basis of Cluster State Tensor Products. *Journal of Chemical Theory and Computation* **2020**, *16*, 6098–6113, PMID: 32846094.

[76] Bozkaya, U.; Sherrill, C. D. Orbital-optimized coupled-electron pair theory and its analytic gradients: Accurate equilibrium geometries, harmonic vibrational frequencies, and hydrogen transfer reactions. *The Journal of Chemical Physics* **2013**, *139*, 054104.

[77] Korth, M. Density Functional Theory: Not Quite the Right Answer for the Right Reason Yet. *Angewandte Chemie International Edition* **2017**, *56*, 5396–5398.

[78] Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and*

Coupled-Cluster Theory; Cambridge Molecular Science; Cambridge University Press, 2009.

[79] Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013**, *493*, 365–370.

[80] Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics. *Nature Chemistry* **2014**, *6*, 927–933.

[81] Li, Z.; Guo, S.; Sun, Q.; Chan, G. K.-L. Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations. *Nature Chemistry* **2019**, *11*, 1026–1033.

[82] Wang, Z.; Zhang, C.; Wang, R.; Wang, G.; Wang, X.; Xiao, M. Weakly coupled triplet pair states probed by quantum beating in delayed fluorescence in tetracene crystals. *The Journal of Chemical Physics* **2019**, *151*, 134309.

[83] Xu, E.; Uejima, M.; Ten-no, S. L. Full Coupled-Cluster Reduction for Accurate Description of Strong Electron Correlation. *Phys. Rev. Lett.* **2018**, *121*, 113001.

[84] White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B* **1993**, *48*, 10345–10356.

[85] White, S. R.; Martin, R. L. Ab initio quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics* **1999**, *110*, 4127–4130.

[86] Mitrushenkov, A. O.; Fano, G.; Ortolani, F.; Linguerri, R.; Palmieri, P. Quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics* **2001**, *115*, 6815–6821.

[87] Nakatani, N.; Chan, G. K.-L. Efficient tree tensor network states (TTNS) for quantum chemistry: Generalizations of the density matrix renormalization group algorithm. *The Journal of Chemical Physics* **2013**, *138*, 134113.

[88] Marti, K. H.; Bauer, B.; Reiher, M.; Troyer, M.; Verstraete, F. Complete-graph tensor network states: a new fermionic wave function ansatz for molecules. *New Journal of Physics* **2010**, *12*, 103008.

[89] Ivanic, J. Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. I. Method. *The Journal of Chemical Physics* **2003**, *119*, 9364–9376.

[90] Ivanic, J. Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. II. Application to oxoMn(salen) and N2O4. *The Journal of Chemical Physics* **2003**, *119*, 9377–9385.

[91] Olsen, J.; Roos, B. O.; Jo/rgensen, P.; Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics* **1988**, *89*, 2185–2192.

[92] Ma, D.; Li Manni, G.; Gagliardi, L. The generalized active space concept in multiconfigurational self-consistent field methods. *The Journal of Chemical Physics* **2011**, *135*, 044128.

[93] Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS Method for Strong Correlation and the Challenging Case of Cr2. *Journal of Chemical Theory and Computation* **2013**, *9*, 3375–3384, PMID: 26584093.

[94] Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for

Large Systems, with Applications to Water Clusters. *Journal of Chemical Theory and Computation* **2007**, *3*, 46–53, PMID: 26627150.

[95] Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. Evaluation of the Electrostatically Embedded Many-Body Expansion and the Electrostatically Embedded Many-Body Expansion of the Correlation Energy by Application to Low-Lying Water Hexamers. *Journal of Chemical Theory and Computation* **2008**, *4*, 33–41, PMID: 26619977.

[96] Schmitt-Monreal, D.; Jacob, C. R. Frozen-density embedding-based many-body expansions. *International Journal of Quantum Chemistry* **2020**, *120*, e26228.

[97] Liu, J.; Qi, L.-W.; Zhang, J. Z. H.; He, X. Fragment Quantum Mechanical Method for Large-Sized Ion–Water Clusters. *Journal of Chemical Theory and Computation* **2017**, *13*, 2021–2034, PMID: 28379695.

[98] Richard, R. M.; Herbert, J. M. A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory. *The Journal of Chemical Physics* **2012**, *137*, 064113.

[99] Veccham, S. P.; Lee, J.; Head-Gordon, M. Making many-body interactions nearly pairwise additive: The polarized many-body expansion approach. *The Journal of Chemical Physics* **2019**, *151*, 194101.

[100] Liu, J.; Sun, H.; Glover, W. J.; He, X. Prediction of Excited-State Properties of Oligoacene Crystals Using Fragment-Based Quantum Mechanical Method. *The Journal of Physical Chemistry A* **2019**, *123*, 5407–5417, PMID: 31187994.

[101] Bygrave, P. J.; Allan, N. L.; Manby, F. R. The embedded many-body expansion for energetics of molecular crystals. *The Journal of Chemical Physics* **2012**, *137*, 164102.

[102] Zhu, T.; de Silva, P.; van Aggelen, H.; Van Voorhis, T. Many-electron expansion: A density functional hierarchy for strongly correlated systems. *Phys. Rev. B* **2016**, *93*, 201108.

[103] Richard, R. M.; Lao, K. U.; Herbert, J. M. Understanding the many-body expansion for large systems. I. Precision considerations. *The Journal of Chemical Physics* **2014**, *141*, 014108.

[104] Friedrich, J.; Walczak, K. Incremental CCSD(T)(F12)|MP2-F12—A Method to Obtain Highly Accurate CCSD(T) Energies for Large Molecules. *Journal of Chemical Theory and Computation* **2013**, *9*, 408–417, PMID: 26589043.

[105] Beran, G. J. O. A New Era for ab initio Molecular Crystal Lattice Energy Prediction. *Angewandte Chemie International Edition* **2015**, *54*, 396–398, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201409823.

[106] Liu, K.-Y.; Herbert, J. M. Energy-Screened Many-Body Expansion: A Practical Yet Accurate Fragmentation Method for Quantum Chemistry. *Journal of Chemical Theory and Computation* **2020**, *16*, 475–487, PMID: 31765559.

[107] Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schütz, M.; Chan, G. K.-L. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science* **2014**, *345*, 640–643.

[108] Wen, S.; Nanda, K.; Huang, Y.; Beran, G. J. O. Practical quantum mechanics-based fragment methods for predicting molecular crystal properties. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578–7590.

[109] Nanda, K. D.; Beran, G. J. O. Prediction of organic molecular crystal geometries

from MP2-level fragment quantum mechanical/molecular mechanical calculations. *The Journal of Chemical Physics* **2012**, *137*, 174106.

[110] Müller, C.; Usvyat, D. Incrementally Corrected Periodic Local MP2 Calculations: I. The Cohesive Energy of Molecular Crystals. *Journal of Chemical Theory and Computation* **2013**, *9*, 5590–5598, PMID: 26592293.

[111] Medders, G. R.; Paesani, F. Many-Body Convergence of the Electrostatic Properties of Water. *Journal of Chemical Theory and Computation* **2013**, *9*, 4844–4852, PMID: 26583403.

[112] Peyton, B. G.; Crawford, T. D. Basis Set Superposition Errors in the Many-Body Expansion of Molecular Properties. *The Journal of Physical Chemistry A* **2019**, *123*, 4500–4511, PMID: 31058506.

[113] Howard, J. C.; Tschumper, G. S. N-body:Many-body QM:QM vibrational frequencies: Application to small hydrogen-bonded clusters. *The Journal of Chemical Physics* **2013**, *139*, 184113.

[114] Heindel, J. P.; Knodel, E. S.; Schofield, D. P. Origin of Many-Body Vibrational Frequency Shifts in Water Clusters. *The Journal of Physical Chemistry A* **2018**, *122*, 6724–6735, PMID: 30028609.

[115] Varandas, A. J. C.; Bowman, J. M.; Gazdy, B. Adjusted double many-body expansion potential energy surface for H02 based on rigorous vibrational calculations. *Chemical Physics Letters* **1995**, *233*, 405–410.

[116] Wang, H.; Yang, W. Force Field for Water Based on Neural Network. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3232–3240, PMID: 29775313.

[117] Bates, D. M.; Smith, J. R.; Tschumper, G. S. Efficient and Accurate Methods for the Geometry Optimization of Water Clusters: Application of Analytic Gradients for the Two-Body:Many-Body QM:QM Fragmentation Method to (H2O)n, n = 3–10. *Journal of Chemical Theory and Computation* **2011**, *7*, 2753–2760, PMID: 26605466.

[118] Demerdash, O.; Head-Gordon, T. Convergence of the Many-Body Expansion for Energy and Forces for Classical Polarizable Models in the Condensed Phase. *Journal of Chemical Theory and Computation* **2016**, *12*, 3884–3893, PMID: 27405002.

[119] Demerdash, O.; Mao, Y.; Liu, T.; Head-Gordon, M.; Head-Gordon, T. Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *The Journal of Chemical Physics* **2017**, *147*, 161721.

[120] Jin, X.; Glover, W. J.; He, X. Fragment Quantum Mechanical Method for Excited States of Proteins: Development and Application to the Green Fluorescent Protein. *Journal of Chemical Theory and Computation* **2020**, *16*, 5174–5188, PMID: 32551640.

[121] Paz, A. S. P.; Glover, W. J. Diabatic Many-Body Expansion: Development and Application to Charge-Transfer Reactions. *Journal of Chemical Theory and Computation* **0**, *0*, null, PMID: 33538588.

[122] Stoll, H. Correlation energy of diamond. *Phys. Rev. B* **1992**, *46*, 6700–6704.

[123] Stoll, H.; Paulus, B.; Fulde, P. On the accuracy of correlation-energy expansions in terms of local increments. *The Journal of Chemical Physics* **2005**, *123*, 144108.

[124] Stoll, H.; Paulus, B.; Fulde, P. An incremental coupled-cluster approach to metallic lithium. *Chemical Physics Letters* **2009**, *469*, 90 – 93.

[125] Stoll, H. Can incremental expansions cope with high-order coupled-cluster contributions? *Molecular Physics* **2010**, *108*, 243–248.

[126] Paulus, B. Wave-function-based ab initio correlation treatment for the buckminsterfullerene C60. *International Journal of Quantum Chemistry* **2004**, *100*, 1026–1032.

[127] Paulus, B. The method of increments—a wavefunction-based ab initio correlation method for solids. *Physics Reports* **2006**, *428*, 1 – 52.

[128] Fertitta, E.; Koch, D.; Paulus, B.; Barcza, G.; Legeza, O. Towards a multiconfigurational method of increments. *Molecular Physics* **2018**, *116*, 1471–1482.

[129] Bytautas, L.; Ruedenberg, K. The Range of Electron Correlation between Localized Molecular Orbitals. A Full Configuration Interaction Analysis for the NCCN Molecule. *The Journal of Physical Chemistry A* **2010**, *114*, 8601–8612, PMID: 20387786.

[130] Boschen, J. S.; Theis, D.; Ruedenberg, K.; Windus, T. L. Correlation Energy Extrapolation by Many-Body Expansion. *The Journal of Physical Chemistry A* **2017**, *121*, 836–844, Publisher: American Chemical Society.

[131] Zimmerman, P. M.; Rask, A. E. Evaluation of full valence correlation energies and gradients. *The Journal of Chemical Physics* **2019**, *150*, 244117, Publisher: American Institute of Physics.

[132] Dang, D.-K.; Zimmerman, P. M. Fully variational incremental CASSCF. *The Journal of Chemical Physics* **2021**, *154*, 014105.

[133] Eriksen, J. J.; Gauss, J. Ground and excited state first-order properties in many-body expanded full configuration interaction theory. *The Journal of Chemical Physics* **2020**, *153*, 154107.

[134] Verma, P.; Huntington, L.; Coons, M.; Kawashima, Y.; Yamazaki, T.; Zaribafiyan, A. Scaling Up Electronic Structure Calculations on Quantum Computers: The Frozen Natural Orbital Based Method of Increments. 2020.

[135] Beran, G. J. O.; Head-Gordon, M.; Gwaltney, S. R. Second-order correction to perfect pairing: an inexpensive electronic structure method for the treatment of strong electron-electron correlations. *J. Chem. Phys.* **2006**, *124*, 114107.

[136] Mayhall, N. J. Using Higher-Order Singular Value Decomposition To Define Weakly Coupled and Strongly Correlated Clusters: The n-Body Tucker Approximation. *Journal of Chemical Theory and Computation* **2017**, *13*, 4818–4828, PMID: 28829590.

[137] Li, S. Block-Correlated Coupled Cluster Theory: The General Formulation and Its Application to the Antiferromagnetic Heisenberg Model. *J. Chem. Phys.* **2004**, *120*, 5017–5017.

[138] Fang, T.; Li, S. Block Correlated Coupled Cluster Theory with a Complete Active-Space Self-Consistent-Field Reference Function: The Formulation and Test Applications for Single Bond Breaking. *J. Chem. Phys.* **2007**, *127*, 204108–204108.

[139] Liu, Y.; Dutoi, A. D. Excitonically renormalised coupled-cluster theory. *Mol. Phys.* **2019**, *117*, 446–461.

[140] Al Hajj, M.; Malrieu, J.-P.; Guihéry, N. Renormalized excitonic method in terms of block excitations: Application to spin lattices. *Phys. Rev. B* **2005**, *72*, 224412.

[141] Ma, Y.; Liu, Y.; Ma, H. A new fragment-based approach for calculating electronic excitation energies of large systems. *J. Chem. Phys.* **2012**, *136*, 024113.

[142] Morrison, A. F.; You, Z.-Q.; Herbert, J. M. Ab Initio Implementation of the Frenkel-

Davydov Exciton Model: A Naturally Parallelizable Approach to Computing Collective Excitations in Crystals and Aggregates. *J. Chem. Theory Comput.* **2014**, *10*, 5366–76.

[143] Morrison, A. F.; Herbert, J. M. Low-Scaling Quantum Chemistry Approach to Excited-State Properties via an ab Initio Exciton Model: Application to Excitation Energy Transfer in a Self-Assembled Nanotube. *J. Phys. Chem. Lett.* **2015**, *6*, 4390–4396.

[144] Pernaut, J.-M.; Reynolds, J. R. Use of Conducting Electroactive Polymers for Drug Delivery and Sensing of Bioactive Molecules. A Redox Chemistry Approach. *The Journal of Physical Chemistry B* **2000**, *104*, 4080–4090.

[145] Milczarek, G.; Inganäs, O. Renewable Cathode Materials from Biopolymer/Conjugated Polymer Interpenetrating Networks. *Science* **2012**, *335*, 1468–1471, Publisher: American Association for the Advancement of Science _eprint: https://science.sciencemag.org/content/335/6075/1468.full.pdf.

[146] Sun, Q. et al. Recent developments in the PySCF program package. *The Journal of Chemical Physics* **2020**, *153*, 024109.

[147] Fishman, M.; White, S. R.; Stoudenmire, E. M. The ITensor Software Library for Tensor Network Calculations. 2020.

[148] Hubbard, J.; Flowers, B. H. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1963**, *276*, 238–257.

[149] FermiCluster, A python library to run fermionic many body problem by partitioning the system into strongly interacting clusters. https://github.com/mayhallgroup/FermiCluster, 2020.

[150] Malek, J.; Flach, S.; Kladko, K. Incremental expansions for the ground-state energy of the two-dimensional Hubbard model. *Phys. Rev. B* **1999**, *59*, R5273–R5276.

[151] Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *The Journal of Chemical Physics* **1992**, *97*, 2571–2577.

[152] Olivares-Amaya, R.; Hu, W.; Nakatani, N.; Sharma, S.; Yang, J.; Chan, G. K.-L. The ab-initio density matrix renormalization group in practice. *The Journal of Chemical Physics* **2015**, *142*, 034102.

[153] Kurashige, Y.; Yanai, T. High-performance ab initio density matrix renormalization group method: Applicability to large-scale multireference problems for metal compounds. *The Journal of Chemical Physics* **2009**, *130*, 234114.

[154] Booth, G. H.; Smart, S. D.; Alavi, A. Linear-scaling and parallelisable algorithms for stochastic quantum chemistry. *Molecular Physics* **2014**, *112*, 1855–1869.

[155] Li, J.; Yao, Y.; Holmes, A. A.; Otten, M.; Sun, Q.; Sharma, S.; Umrigar, C. J. Accurate many-body electronic structure near the basis set limit: Application to the chromium dimer. *Phys. Rev. Research* **2020**, *2*, 012015.

[156] Lehtola, S.; Tubman, N. M.; Whaley, K. B.; Head-Gordon, M. Cluster decomposition of full configuration interaction wave functions: A tool for chemical interpretation of systems with strong correlation. *The Journal of Chemical Physics* **2017**, *147*, 154105.

[157] Chen, S.; Ullah, N.; Wang, T.; Zhang, R. Tuning the optical properties of graphene quantum dots by selective oxidation: a theoretical perspective. *J. Mater. Chem. C* **2018**, *6*, 6875–6883.

[158] Hachmann, J.; Dorando, J. J.; Avilés, M.; Chan, G. K.-L. The radical character of

the acenes: A density matrix renormalization group study. *The Journal of Chemical Physics* **2007**, *127*, 134309.

[159] Maksić, Z. B.; Barić, D.; Müller, T. Clar's Sextet Rule Is a Consequence of the σ-Electron Framework. *The Journal of Physical Chemistry A* **2006**, *110*, 10135–10147, PMID: 16913689.

[160] Alonso, M.; Herradón, B. A universal scale of aromaticity for π-organic compounds. *Journal of Computational Chemistry* **2010**, *31*, 917–928.

[161] Yeh, C.-N.; Chai, J.-D. Role of Kekulé and Non-Kekulé Structures in the Radical Character of Alternant Polycyclic Aromatic Hydrocarbons: A TAO-DFT Study. *Scientific Reports* **2016**, *6*, 30562.

[162] Fedik, N.; Boldyrev, A. I. Insight into The Nature of Rim Bonds in Coronene. *The Journal of Physical Chemistry A* **2018**, *122*, 8585–8590, PMID: 30296096.

[163] Dong, R.; Pfeffermann, M.; Skidin, D.; Wang, F.; Fu, Y.; Narita, A.; Tommasini, M.; Moresco, F.; Cuniberti, G.; Berger, R.; Müllen, K.; Feng, X. Persulfurated Coronene: A New Generation of "Sulflower". *Journal of the American Chemical Society* **2017**, *139*, 2168–2171, PMID: 28128953.

[164] Arrow, Fast Semistochastic Heat Bath Configuration Interaction Solver (SHCI). https://github.com/QMC-Cornell/shci, 2020.

[165] Noffke, B. W.; Beckett, D.; Li, L.-s.; Raghavachari, K. Aromatic Fragmentation Based on a Ring Overlap Scheme: An Algorithm for Large Polycyclic Aromatic Hydrocarbons Using the Molecules-in-Molecules Fragmentation-Based Method. *J. Chem. Theory Comput.* **2020**, *16*, 2160–2171.

[166] Pozo, I.; Majzik, Z.; Pavliček, N.; Melle-Franco, M.; Guitián, E.; Peña, D.; Gross, L.; Pérez, D. Revisiting Kekulene: Synthesis and Single-Molecule Imaging. *Journal of the American Chemical Society* **2019**, *141*, 15488–15493, PMID: 31525873.

[167] Jiao, H.; Schleyer, P. v. R. Is Kekulene Really Superaromatic? *Angewandte Chemie International Edition in English* **1996**, *35*, 2383–2386.

[168] Bygrave, P. J.; Allan, N. L.; Manby, F. R. The embedded many-body expansion for energetics of molecular crystals. *The Journal of Chemical Physics* **2012**, *137*, 164102, _eprint: https://doi.org/10.1063/1.4759079.

[169] Liu, K.-Y.; Herbert, J. M. Understanding the many-body expansion for large systems. III. Critical role of four-body terms, counterpoise corrections, and cutoffs. *The Journal of Chemical Physics* **2017**, *147*, 161729.

[170] Ouyang, J. F.; Bettens, R. P. A. When are Many-Body Effects Significant? *Journal of Chemical Theory and Computation* **2016**, *12*, 5860–5867, PMID: 27779845.

[171] Smela, E. Conjugated Polymer Actuators for Biomedical Applications. *Advanced Materials* **2003**, *15*, 481–494.

[172] Park, K.-H.; Jo, E. A.; Na, K. Heparin/polypyrrole (PPy) composite on gold-coated matrix for the neurite outgrowth of PC12 cells by electrical stimulation. *Biotechnology and Bioprocess Engineering* **2007**, *12*, 463.

[173] Yuan, X.; Floresyona, D.; Aubert, P.-H.; Bui, T.-T.; Remita, S.; Ghosh, S.; Brisset, F.; Goubard, F.; Remita, H. Photocatalytic degradation of organic pollutant with polypyrrole nanostructures under UV and visible light. *Applied Catalysis B: Environmental* **2019**, *242*, 284 – 292.

[174] Huang, Y.; Li, H.; Wang, Z.; Zhu, M.; Pei, Z.; Xue, Q.; Huang, Y.; Zhi, C. Nanostructured Polypyrrole as a flexible electrode material of supercapacitor. *Nano Energy* **2016**, *22*, 422 – 438.

[175] Chen, X. L.; Jenekhe, S. A. Bipolar Conducting Polymers: Blends of p-Type Polypyrrole and an n-Type Ladder Polymer. *Macromolecules* **1997**, *30*, 1728–1733.

[176] Schollwöck, U. The density-matrix renormalization group. *Rev. Mod. Phys.* **2005**, *77*, 259–315.

[177] Chan, G. K.-L.; Sharma, S. The density matrix renormalization group in quantum chemistry. *Ann. Rev. Phys. Chem.* **2011**, *62*, 465–81.

[178] Jordan, J.; Orús, R.; Vidal, G.; Verstraete, F.; Cirac, J. I. Classical Simulation of Infinite-Size Quantum Lattice Systems in Two Spatial Dimensions. *Phys. Rev. Lett.* **2008**, *101*, 250602.

[179] Verstraete, F.; Cirac, J. I. Renormalization Algorithms for Quantum-Many Body Systems in Two and Higher Dimensions. *arXiv:cond-mat/0407066* **2004**,

[180] Hyatt, K.; Stoudenmire, E. M. DMRG Approach to Optimizing Two-Dimensional Tensor Networks. *arXiv:1908.08833 [cond-mat, physics:quant-ph]* **2019**,

[181] Kovyrshin, A.; Reiher, M. Self-adaptive tensor network states with multi-site correlators. *J. Chem. Phys.* **2017**, *147*, 214111.

[182] Marti, K. H.; Bauer, B.; Reiher, M.; Troyer, M.; Verstraete, F. Complete-Graph Tensor Network States: A New Fermionic Wave Function Ansatz for Molecules. *New J. Phys.* **2010**, *12*, 103008–103008.

[183] Murg, V.; Verstraete, F.; Schneider, R.; Nagy, P. R.; Legeza, Ö. Tree Tensor Network State with Variable Tensor Order: An Efficient Multireference Method for Strongly Correlated Systems. *J. Chem. Theory Comput.* **2015**, *11*, 1027–1036.

[184] Szalay, S.; Pfeffer, M.; Murg, V.; Barcza, G.; Verstraete, F.; Schneider, R.; Legeza, r. Tensor Product Methods and Entanglement Optimization for Ab Initio Quantum Chemistry. *Int. J. Quantum Chem.* **2015**, *115*, 1342–1391.

[185] Choo, K.; Mezzacapo, A.; Carleo, G. Fermionic neural-network states for ab-initio electronic structure. *Nat. Commun.* **2020**, *11*, 2368.

[186] Nakatani, N.; Chan, G. K.-L. Efficient tree tensor network states (TTNS) for quantum chemistry: Generalizations of the density matrix renormalization group algorithm. *J. Chem. Phys.* **2013**, *138*, 134113.

[187] Bender, C. F.; Davidson, E. R. Studies in Configuration Interaction: The First-Row Diatomic Hydrides. *Phys. Rev.* **1969**, *183*, 23–30.

[188] Whitten, J. L.; Hackmeyer, M. Configuration Interaction Studies of Ground and Excited States of Polyatomic Molecules. I. The CI Formulation and Studies of Formaldehyde. *J. Chem. Phys.* **1969**, *51*, 5584–5596.

[189] Liu, W.; Hoffmann, M. R. iCI: iterative CI toward full CI. *J. Chem. Theory Comput.* **2016**, *12*, 1169–1178.

[190] Chakraborty, R.; Ghosh, P.; Ghosh, D. Evolutionary algorithm based configuration interaction approach. *Int. J. Quantum Chem.* **2018**, *118*, e25509.

[191] Ohtsuka, Y.; Hasegawa, J.-y. Selected configuration interaction method using sampled first-order corrections to wave functions. *J. Chem. Phys.* **2017**, *147*, 34102.

[192] Coe, J. P. Machine Learning Configuration Interaction. *J. Chem. Theory Comput.* **2018**, *14*, 5739–5749.

[193] Levine, D. S.; Hait, D.; Tubman, N. M.; Lehtola, S.; Whaley, K. B.; Head-Gordon, M. CASSCF with Extremely Large Active Spaces Using the Adaptive Sampling Configuration Interaction Method. *J. Chem. Theory Comput.* **2020**, *16*, 2340–2354.

[194] Evangelisti, S.; Daudey, J.-P.; Malrieu, J.-P. Convergence of an improved CIPSI algorithm. *Chem. Phys.* **1983**, *75*, 91 – 102.

[195] Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks. *J. Chem. Theory Comput.* **2018**, *14*, 4360–4379.

[196] Garniron, Y. et al. Quantum Package 2.0: An Open-Source Determinant-Driven Suite of Programs. *J. Chem. Theory Comput.* **2019**, *15*, 3591–3609.

[197] Giner, E.; Scemama, A.; Caffarel, M. Using perturbatively selected configuration interaction in quantum Monte Carlo calculations. *Can. J. Chem.* **2013**, *91*, 879–885.

[198] Scemama, A.; Garniron, Y.; Caffarel, M.; Loos, P.-F. Deterministic Construction of Nodal Surfaces within Quantum Monte Carlo: The Case of FeS. *J. Chem. Theory Comput.* **2018**, *14*, 1395–1402.

[199] Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern Approaches to Exact Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method. *J. Chem. Theory Comput.* **2020**, *16*, 2139–2159.

[200] Levine, D. S.; Hait, D.; Tubman, N. M.; Lehtola, S.; Whaley, K. B.; Head-Gordon, M.

CASSCF with Extremely Large Active Spaces Using the Adaptive Sampling Configuration Interaction Method. *J. Chem. Theory Comput.* **2020**, *16*, 2340–2354.

[201] Evangelista, F. A. Adaptive multiconfigurational wave functions. *J. Chem. Phys.* **2014**, *140*, 124114.

[202] Holmes, A. A.; Changlani, H. J.; Umrigar, C. J. Efficient Heat-Bath Sampling in Fock Space. *J. Chem. Theory Comput.* **2016**, *12*, 1561–1571.

[203] Smith, J. E. T.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and Near Exact CASSCF with Large Active Spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468–5478.

[204] Coe, J. P.; Paterson, M. J. Development of Monte Carlo configuration interaction: Natural orbitals and second-order perturbation theory. *J. Chem. Phys.* **2012**, *137*, 204108.

[205] Holmes, A. A.; Umrigar, C. J.; Sharma, S. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **2017**, *147*, 164111.

[206] Tucker, L. R. Some Mathematical Notes on Three-Mode Factor Analysis. *Psychometrika* **1966**, *31*, 279–311.

[207] Epstein, P. S. The Stark Effect from the Point of View of Schroedinger's Quantum Theory. *Phys. Rev.* **1926**, *28*, 695–710.

[208] Nesbet, R. K.; Hartree, D. R. Configuration interaction in orbital theories. *Proc. R. Soc. London, Ser. A* **1955**, *230*, 312–321.

[209] Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. PySCF: the Python-based simulations of chemistry framework. *WIRES Comput. Mol. Sci.* **2018**, *8*, e1340.

[210] Tubman, N. M.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. An Efficient Deterministic Perturbation Theory for Selected Configuration Interaction Methods. *arXiv:1808.02049 [cond-mat, physics:physics, physics:quant-ph]* **2018**,

[211] Fang, T.; Shen, J.; Li, S. Block Correlated Coupled Cluster Method with a Complete-Active-Space Self-Consistent-Field Reference Function: The Implementation for Low-Lying Excited States. *J. Chem. Phys.* **2008**, *129*, 234106–234106.

[212] Fang, T.; Shen, J.; Li, S. Block Correlated Coupled Cluster Method with a Complete-Active-Space Self-Consistent-Field Reference Function: The Formula for General Active Spaces and Its Applications for Multibond Breaking Systems. *J. Chem. Phys.* **2008**, *128*, 224107–224107.

[213] Shen, J.; Li, S. Block Correlated Coupled Cluster Method with the Complete Active-Space Self-Consistent-Field Reference Function: Applications for Low-Lying Electronic Excited States. *J. Chem. Phys.* **2009**, *131*.

[214] Xu, E.; Li, S. Block Correlated Second Order Perturbation Theory with a Generalized Valence Bond Reference Function. *J. Chem. Phys.* **2013**, *139*, 174111–174111.

[215] Zhang, H.; Malrieu, J.-P.; Ma, H.; Ma, J. Implementation of renormalized excitonic method at ab initio level. *J. Comput. Chem.* **2011**, *33*, 34–43.

[216] Ma, Y.; Ma, H. Calculating excited states of molecular aggregates by the renormalized excitonic method. *J. Phys. Chem. A* **2013**, *117*, 3655–3665.

[217] Parker, S. M.; Seideman, T.; Ratner, M. A.; Shiozaki, T. Communication: Active-space decomposition for molecular dimers. *J. Chem. Phys.* **2013**, *139*, 021108.

[218] Parker, S. M.; Seideman, T.; Ratner, M. A.; Shiozaki, T. Model Hamiltonian Analysis

of Singlet Fission from First Principles. *The Journal of Physical Chemistry C* **2014**, *118*, 12700–12705.

[219] Nishio, S.; Kurashige, Y. Rank-one basis made from matrix-product states for a low-rank approximation of molecular aggregates. *J. Chem. Phys.* **2019**, *151*, 084110.

[220] Hermes, M. R.; Pandharkar, R.; Gagliardi, L. Variational Localized Active Space Self-Consistent Field Method. *J. Chem. Theory Comput.* **2020**, *16*, 4923–4937.

[221] Ma, Y.; Wen, J.; Ma, H. Density-matrix renormalization group algorithm with multi-level active space. *J. Chem. Phys.* **2015**, *143*, 034105.

[222] Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2014**, *113*, 184–215.

[223] Hunt, W.; Hay, P.; Goddard Iii, W. Self-Consistent Procedures for Generalized Valence Bond Wavefunctions. Applications H3, BH, H2O, C2H6, and O2. *J. Chem. Phys.* **1972**, *57*, 738–748.

[224] Beran, G. J. O.; Austin, B.; Sodt, A.; Head-Gordon, M. Unrestricted Perfect Pairing: The Simplest Wave-Function-Based Model Chemistry beyond Mean Field. *J. Phys. Chem. A* **2005**, *109*, 9183–9192.

[225] Sayfutyarova, E. R.; Sun, Q.; Chan, G. K.-L.; Knizia, G. Automated Construction of Molecular Active Spaces from Atomic Valence Orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 4063–4078.

[226] Luzanov, A. V.; Prezhdo, O. High-order entropy measures and spin-free quantum entanglement for molecular problems. *Mol. Phys.* **2010**, *105*, 2879–2891.

[227] Stein, C. J.; Reiher, M. Automated identification of relevant frontier orbitals for chemical compounds and processes. *Chimia* **2017**, *71*, 170–176.

[228] Morrison, A. F.; Herbert, J. M. Evidence for Singlet Fission Driven by Vibronic Coherence in Crystalline Tetracene. *The Journal of Physical Chemistry Letters* **2017**, *8*, 1442–1448, PMID: 28277682.

[229] Claudino, D.; Mayhall, N. J. Simple and Efficient Truncation of Virtual Spaces in Embedded Wave Functions via Concentric Localization. *J. Chem. Theory Comput.* **2019**, *15*, 6085–6096.

[230] Knizia, G.; Chan, G. K.-L. Density Matrix Embedding: A Simple Alternative to Dynamical Mean-Field Theory. *Phys. Rev. Lett.* **2012**, *109*, 186404–186404.

[231] Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

[232] Marques, M.; Gross, E. TIME-DEPENDENT DENSITY FUNCTIONAL THEORY. *Annual Review of Physical Chemistry* **2004**, *55*, 427–455, PMID: 15117259.

[233] Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *The Journal of Chemical Physics* **2002**, *117*, 7433–7447.

[234] Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double excitations within time-dependent density functional theory linear response. *The Journal of Chemical Physics* **2004**, *120*, 5932–5937.

[235] Elliott, P.; Goldson, S.; Canahui, C.; Maitra, N. T. Perspectives on double-excitations in TDDFT. *Chemical Physics* **2011**, *391*, 110–119, Open problems and new solutions in time dependent density functional theory.

[236] Dreuw, A.; Head-Gordon, M. Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin−Bacteri-

ochlorin and Bacteriochlorophyll−Spheroidene Complexes. *Journal of the American Chemical Society* **2004**, *126*, 4007–4016, PMID: 15038755.

[237] Levine, B. G.; Ko, C.; Quenneville, J.; MartÍnez, T. J. Conical intersections and double excitations in time-dependent density functional theory. *Molecular Physics* **2006**, *104*, 1039–1051.

[238] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation energies in density functional theory: An evaluation and a diagnostic test. *The Journal of Chemical Physics* **2008**, *128*, 044118.

[239] Geertsen, J.; Rittby, M.; Bartlett, R. J. The equation-of-motion coupled-cluster method: Excitation energies of Be and CO. *Chemical Physics Letters* **1989**, *164*, 57–62.

[240] Comeau, D. C.; Bartlett, R. J. The equation-of-motion coupled-cluster method. Applications to open-and closed-shell reference states. *Chemical Physics Letters* **1993**, *207*, 414–423.

[241] Chien, A. D.; Holmes, A. A.; Otten, M.; Umrigar, C. J.; Sharma, S.; Zimmerman, P. M. Excited States of Methylene, Polyenes, and Ozone from Heat-Bath Configuration Interaction. *The Journal of Physical Chemistry A* **2018**, *122*, 2714–2722, PMID: 29473750.

[242] Loos, P.-F.; Boggio-Pasqua, M.; Scemama, A.; Caffarel, M.; Jacquemin, D. Reference Energies for Double Excitations. *Journal of Chemical Theory and Computation* **2019**, *15*, 1939–1956.

[243] Loos, P.-F.; Comin, M.; Blase, X.; Jacquemin, D. Reference Energies for Intramolecular Charge-Transfer Excitations. *Journal of Chemical Theory and Computation* **2021**, *17*, 3666–3686, PMID: 33955742.

[244] Bäppler, S. A.; Plasser, F.; Wormit, M.; Dreuw, A. Exciton analysis of many-body wave functions: Bridging the gap between the quasiparticle and molecular orbital pictures. *Phys. Rev. A* **2014**, *90*, 052521.

[245] Luzanov, A. V.; Casanova, D.; Feng, X.; Krylov, A. I. Quantifying charge resonance and multiexciton character in coupled chromophores by charge and spin cumulant analysis. *J. Chem. Phys.* **2015**, *142*, 224104.

[246] Anstöter, C. S.; Dean, C. R.; Verlet, J. R. R. Chromophores of chromophores: a bottom-up Hückel picture of the excited states of photoactive proteins. *Phys. Chem. Chem. Phys.* **2017**, *19*, 29772–29779.

[247] Diaz-Andres, A.; Casanova, D. Benzene Excimer and Excited Multimers: Electronic Character, Interaction Nature, and Aromaticity. *The Journal of Physical Chemistry Letters* **2021**, *12*, 7400–7408, PMID: 34328333.

[248] Reger, D.; Haines, P.; Amsharov, K. Y.; Schmidt, J. A.; Ullrich, T.; Bönisch, S.; Hampel, F.; Görling, A.; Nelson, J.; Jelfs, K. E.; Guldi, D. M.; Jux, N. A Family of Superhelicenes: Easily Tunable, Chiral Nanographenes by Merging Helicity with Planar Systems. *Angewandte Chemie International Edition* **2021**, *60*, 18073–18081.

[249] Fan, W.; Matsuno, T.; Han, Y.; Wang, X.; Zhou, Q.; Isobe, H.; Wu, J. Synthesis and Chiral Resolution of Twisted Carbon Nanobelts. *Journal of the American Chemical Society* **0**, *0*, null, PMID: 34550688.

[250] Liu, M.; Liu, M.; She, L.; Zha, Z.; Pan, J.; Li, S.; Li, T.; He, Y.; Cai, Z.; Wang, J.; Zheng, Y.; Qiu, X.; Zhong, D. Graphene-like nanoribbons periodically embedded with four- and eight-membered rings. *Nature Communications* **2017**, *8*, 14924.

[251] Gantenbein, M.; Li, X.; Sangtarash, S.; Bai, J.; Olsen, G.; Alqorashi, A.; Hong, W.;

Lambert, C. J.; Bryce, M. R. Exploring antiaromaticity in single-molecule junctions formed from biphenylene derivatives. *Nanoscale* **2019**, *11*, 20659–20666.

[252] Liu, P.; Chen, X.-Y.; Cao, J.; Ruppenthal, L.; Gottfried, J. M.; Müllen, K.; Wang, X.-Y. Revisiting Acepleiadylene: Two-Step Synthesis and -Extension toward Non-benzenoid Nanographene. *Journal of the American Chemical Society* **2021**, *143*, 5314–5318, PMID: 33784083.

[253] Zeng, T.; Hoffmann, R.; Ananth, N. The Low-Lying Electronic States of Pentacene and Their Roles in Singlet Fission. *Journal of the American Chemical Society* **2014**, *136*, 5755–5764, PMID: 24697685.

[254] Wilson, M. W. B.; Rao, A.; Clark, J.; Kumar, R. S. S.; Brida, D.; Cerullo, G.; Friend, R. H. Ultrafast dynamics of exciton fission in polycrystalline pentacene. *J. Am. Chem. Soc.* **2011**, *133*, 11830–11833.

[255] Rao, A.; Wilson, M. W. B.; Albert-Seifried, S.; Di Pietro, R.; Friend, R. H. Photophysics of pentacene thin films: The role of exciton fission and heating effects. *Phys. Rev. B* **2011**, *84*, 195411.

[256] Walker, B. J.; Musser, A. J.; Beljonne, D.; Friend, R. H. Singlet exciton fission in solution. *Nature chemistry* **2013**, *5*, 1019–1024.

[257] Libisch, F.; Huang, C.; Carter, E. A. Embedded Correlated Wavefunction Schemes: Theory and Applications. *Accounts of Chemical Research* **2014**, *47*, 2768–2775, PMID: 24873211.

[258] Ma, H.; Sheng, N.; Govoni, M.; Galli, G. Quantum Embedding Theory for Strongly Correlated States in Materials. *Journal of Chemical Theory and Computation* **2021**, *17*, 2116–2125, PMID: 33739106.

[259] Bauman, N. P.; Bylaska, E. J.; Krishnamoorthy, S.; Low, G. H.; Wiebe, N.; Granade, C. E.; Roetteler, M.; Troyer, M.; Kowalski, K. Downfolding of many-body Hamiltonians using active-space models: Extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms. *The Journal of Chemical Physics* **2019**, *151*, 014107.

[260] Kowalski, K.; Bauman, N. P. Sub-system quantum dynamics using coupled cluster downfolding techniques. *The Journal of Chemical Physics* **2020**, *152*, 244127.

[261] Zgid, D.; Gull, E. Finite temperature quantum embedding theories for correlated systems. *New Journal of Physics* **2017**, *19*, 023047.

[262] Lan, T. N.; Zgid, D. Generalized Self-Energy Embedding Theory. *The Journal of Physical Chemistry Letters* **2017**, *8*, 2200–2205, PMID: 28453934.

[263] Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annual Review of Physical Chemistry* **1981**, *32*, 359–401.

[264] *, M. N.; Shamasundar, K. R.; Mukherjee, D. Reflections on size-extensivity, size-consistency and generalized extensivity in many-body theory. *Molecular Physics* **2005**, *103*, 2277–2298.

[265] Rayleigh, J. *The Theory of Sound*; The Theory of Sound v. 1; Macmillan, 1894.

[266] Schrödinger, E. *Collected Papers on Wave Mechanics*; Collected papers of scientists; Chelsea Publishing Company, 1928.

[267] M\oller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

[268] Raghavachari, K.; Pople, J. A.; Replogle, E. S.; Head-Gordon, M. Fifth order Moeller-Plesset perturbation theory: comparison of existing correlation methods and implementation of new methods correct to fifth order. *The Journal of Physical Chemistry* **1990**, *94*, 5579–5586.

[269] Hubač, I.; Čársky, P. Correlation energy of open-shell systems. Application of the many-body Rayleigh-Schrödinger perturbation theory in the restricted Roothaan-Hartree-Fock formalism. *Phys. Rev. A* **1980**, *22*, 2392–2399.

[270] Cremer, D. Møller-Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.

[271] Frantz, L. M.; Mills, R. L. Many-body basis for the optical model. *Nuclear Physics* **1960**, *15*, 16–32.

[272] Bartlett, R. J.; Shavitt, I. Comparison of high-order many-body perturbation theory and configuration interaction for H2O. *Chemical Physics Letters* **1977**, *50*, 190–198.

[273] Ahlrichs, R. Many body perturbation calculations and coupled electron pair models. *Computer Physics Communications* **1979**, *17*, 31–45.

[274] Koch, S.; Kutzelnigg, W. Comparison of CEPA and CP-MET methods. *Theoretica chimica acta* **1981**, *59*, 387–411.

[275] Kollmar, C.; Neese, F. The coupled electron pair approximation: variational formulation and spin adaptation. *Molecular Physics* **2010**, *108*, 2449–2458.

[276] Taube, A. G.; Bartlett, R. J. Rethinking linearized coupled-cluster theory. *The Journal of Chemical Physics* **2009**, *130*, 144112.

[277] Fink, R.; Staemmler, V. A multi-configuration reference CEPA method based on pair natural orbitals. *Theoretica chimica acta* **1993**, *87*, 129–145.

[278] Szalay, P. G. Multireference averaged quadratic coupled-cluster (MR-AQCC) method based on the functional of the total energy. *Chemical Physics* **2008**, *349*, 121–125, Electron Correlation and Molecular Dynamics for Excited States and Photochemistry.

[279] Chaudhuri, R. K.; Freed, K. F. Comparison of high order perturbative convergence of multireference perturbation methods: Application to singlet states of CH2. *The Journal of Chemical Physics* **1997**, *107*, 6699–6711.

[280] Smith, M. B.; Michl, J. Singlet fission. *Chem. Rev.* **2010**, *110*, 6891–6936.

[281] Smith, M. B.; Michl, J. Recent Advances in Singlet Fission. *Annual Review of Physical Chemistry* **2013**, *64*, 361–386, PMID: 23298243.

[282] Casanova, D. Theoretical Modeling of Singlet Fission. *Chemical Reviews* **2018**, *118*, 7164–7207, PMID: 29648797.

[283] Ullrich, T.; Munz, D.; Guldi, D. M. Unconventional singlet fission materials. *Chem. Soc. Rev.* **2021**, *50*, 3485–3518.

[284] Hanna, M. C.; Nozik, A. J. Solar conversion efficiency of photovoltaic and photoelectrolysis cells with carrier multiplication absorbers. *J. Appl. Phys.* **2006**, *100*, 074510.

[285] Miyata, K.; Conrad-Burton, F. S.; Geyer, F. L.; Zhu, X.-Y. Triplet Pair States in Singlet Fission. *Chemical Reviews* **2019**, *119*, 4261–4292.

[286] Sanders, S. N.; Pun, A. B.; Parenti, K. R.; Kumarasamy, E.; Yablon, L. M.; Sfeir, M. Y.; Campos, L. M. Understanding the Bound Triplet-Pair State in Singlet Fission. *Chem* **2019**, *5*, 1988–2005.

[287] Musser, A. J.; Clark, J. Triplet-Pair States in Organic Semiconductors. *Annual Review of Physical Chemistry* **2019**, *70*, 323–351, PMID: 31174458.

[288] Kim, H.; Zimmerman, P. M. Coupled double triplet state in singlet fission. *Phys. Chem. Chem. Phys.* **2018**, *20*, 30083–30094.

[289] Zimmerman, P. M.; Zhang, Z.; Musgrave, C. B. Singlet fission in pentacene through multi-exciton quantum states. *Nature chemistry* **2010**, *2*, 648–652.

[290] Sanders, S. N.; Kumarasamy, E.; Pun, A. B.; Trinh, M. T.; Choi, B.; Xia, J.; Taffet, E. J.; Low, J. Z.; Miller, J. R.; Roy, X.; Zhu, X.-Y.; Steigerwald, M. L.; Sfeir, M. Y.; Campos, L. M. Quantitative Intramolecular Singlet Fission in Bipentacenes. *Journal of the American Chemical Society* **2015**, *137*, 8965–8972, PMID: 26102432.

[291] Tayebjee, M. J. Y.; Sanders, S. N.; Kumarasamy, E.; Campos, L. M.; Sfeir, M. Y.; McCamey, D. R. Quintet multiexciton dynamics in singlet fission. *Nat. Phys.* **2017**, *13*, 182–188.

[292] Folie, B. D.; Haber, J. B.; Refaely-Abramson, S.; Neaton, J. B.; Ginsberg, N. S. Long-Lived Correlated Triplet Pairs in a π-Stacked Crystalline Pentacene Derivative. *J. Am. Chem. Soc.* **2018**, *140*, 2326–2335.

[293] Kundu, A.; Dasgupta, J. Photogeneration of Long-Lived Triplet States through Singlet Fission in Lycopene H-Aggregates. *The Journal of Physical Chemistry Letters* **2021**, *12*, 1468–1474, PMID: 33528257.

[294] Zhou, Z.; Ma, L.; Guo, D.; Zhao, X.; Wang, C.; Lin, D.; Zhang, F.; Zhang, J.; Nie, Z. Ultrafast Dynamics of Long-Lived Bound Triplet Pair Generated by Singlet Fission

in 6,13-Bis(triisopropylsilylethynyl) Pentacene. *The Journal of Physical Chemistry C* **2020**, *124*, 14503–14509.

[295] Basel, B. S. et al. Unified model for singlet fission within a non-conjugated covalent pentacene dimer. *Nat. Commun.* **2017**, *8*, 1–8.

[296] Sakai, H.; Inaya, R.; Nagashima, H.; Nakamura, S.; Kobori, Y.; Tkachenko, N. V.; Hasobe, T. Multiexciton Dynamics Depending on Intramolecular Orientations in Pentacene Dimers: Recombination and Dissociation of Correlated Triplet Pairs. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3354–3360, PMID: 29847939.

[297] Lubert-Perquel, D.; Salvadori, E.; Dyson, M.; Stavrinou, P. N.; Montis, R.; Nagashima, H.; Kobori, Y.; Heutz, S.; Kay, C. W. M. Identifying triplet pathways in dilute pentacene films. *Nat. Commun.* **2018**, *9*, 4222.

[298] Nagashima, H.; Kawaoka, S.; Akimoto, S.; Tachikawa, T.; Matsui, Y.; Ikeda, H.; Kobori, Y. Singlet-Fission-Born Quintet State: Sublevel Selections and Trapping by Multiexciton Thermodynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 5855–5861.

[299] Weiss, L. R.; Bayliss, S. L.; Kraffert, F.; Thorley, K. J.; Anthony, J. E.; Bittl, R.; Friend, R. H.; Rao, A.; Greenham, N. C.; Behrends, J. Strongly exchange-coupled triplet pairs in an organic semiconductor. *Nat. Phys.* **2017**, *13*, 176–181.

[300] Chen, M.; Krzyaniak, M. D.; Nelson, J. N.; Bae, Y. J.; Harvey, S. M.; Schaller, R. D.; Young, R. M.; Wasielewski, M. R. Quintet-triplet mixing determines the fate of the multiexciton state produced by singlet fission in a terrylenediimide dimer at room temperature. *Proc. Natl. Acad. Sci.* **2019**, *116*, 8178–8183.

[301] Smyser, K. E.; Eaves, J. D. Singlet fission for quantum information and quantum computing: the parallel JDE model. *Scientific reports* **2020**, *10*, 1–10.

[302] Bhattacharyya, K.; Datta, A. Polymorphism Controlled Singlet Fission in TIPS-Anthracene: Role of Stacking Orientation. *The Journal of Physical Chemistry C* **2017**, *121*, 1412–1420.

[303] Collins, M. I.; McCamey, D. R.; Tayebjee, M. J. Y. Fluctuating exchange interactions enable quintet multiexciton formation in singlet fission. *The Journal of Chemical Physics* **2019**, *151*, 164104.

[304] Abraham, V.; Mayhall, N. J. Simple Rule To Predict Boundedness of Multiexciton States in Covalently Linked Singlet-Fission Dimers. *J. Phys. Chem. Lett.* **2017**, *8*, 5472–5478.

[305] Lin, H.-H.; Kue, K. Y.; Claudio, G. C.; Hsu, C.-P. First Principle Prediction of Intramolecular Singlet Fission and Triplet Triplet Annihilation Rates. *Journal of Chemical Theory and Computation* **2019**, *15*, 2246–2253, PMID: 30860838.

[306] Wang, X.; Tom, R.; Liu, X.; Congreve, D. N.; Marom, N. An energetics perspective on why there are so few triplet–triplet annihilation emitters. *J. Mater. Chem. C* **2020**, *8*, 10816–10824, Publisher: The Royal Society of Chemistry.

[307] Khan, S.; Mazumdar, S. Free Triplets Versus Bound Triplet–Triplet Biexciton in Intramolecular Singlet Fission Materials: Structure–Property Correlations. *The Journal of Physical Chemistry C* **2020**, *124*, 1171–1177.

[308] Scholes, G. D. Correlated Pair States Formed by Singlet Fission and Exciton–Exciton Annihilation. *The Journal of Physical Chemistry A* **2015**, *119*, 12699–12705, PMID: 26595530.

[309] Pensack, R. D.; Ostroumov, E. E.; Tilley, A. J.; Mazza, S.; Grieco, C.; Thorley, K. J.; Asbury, J. B.; Seferos, D. S.; Anthony, J. E.; Scholes, G. D. Observation of Two

Triplet-Pair Intermediates in Singlet Exciton Fission. *The Journal of Physical Chemistry Letters* **2016**, *7*, 2370–2375, PMID: 27281713.

[310] Lee, T. S.; Lin, Y. L.; Kim, H.; Pensack, R. D.; Rand, B. P.; Scholes, G. D. Triplet Energy Transfer Governs the Dissociation of the Correlated Triplet Pair in Exothermic Singlet Fission. *J. Phys. Chem. Lett.* **2018**, *9*, 4087–4095.

[311] Dexter, D. L. A theory of sensitized luminescence in solids. *The journal of chemical physics* **1953**, *21*, 836–850.

[312] Korovina, N. V.; Das, S.; Nett, Z.; Feng, X.; Joy, J.; Haiges, R.; Krylov, A. I.; Bradforth, S. E.; Thompson, M. E. Singlet Fission in a Covalently Linked Cofacial Alkynyltetracene Dimer. *Journal of the American Chemical Society* **2016**, *138*, 617–627, PMID: 26693957.

[313] Srimath Kandada, A. R.; Petrozza, A.; Lanzani, G. Ultrafast dissociation of triplets in pentacene induced by an electric field. *Phys. Rev. B* **2014**, *90*, 075310.

[314] Skourtis, S. S.; Liu, C.; Antoniou, P.; Virshup, A. M.; Beratan, D. N. Dexter energy transfer pathways. *Proceedings of the National Academy of Sciences* **2016**, *113*, 8115–8120.

[315] You, Z.-Q.; Hsu, C.-P. The fragment spin difference scheme for triplet-triplet energy transfer coupling. *The Journal of Chemical Physics* **2010**, *133*, 074105.

[316] Devi, L. S.; Al-Suti, M. K.; Dosche, C.; Khan, M. S.; Friend, R. H.; Köhler, A. Triplet energy transfer in conjugated polymers. I. Experimental investigation of a weakly disordered compound. *Physical Review B* **2008**, *78*, 045210.

[317] Scholes, G. D. Long-Range Resonance Energy Transfer in Molecular Systems. *Annual Review of Physical Chemistry* **2003**, *54*, 57–87, PMID: 12471171.

[318] You, Z.-Q.; Hsu, C.-P.; Fleming, G. R. Triplet-triplet energy-transfer coupling: Theory and calculation. *The Journal of Chemical Physics* **2006**, *124*, 044506.

[319] Wang, Z.; Liu, H.; Xie, X.; Zhang, C.; Wang, R.; Chen, L.; Xu, Y.; Ma, H.; Fang, W.; Yao, Y.; Sang, H.; Wang, X.; Li, X.; Xiao, M. Free-triplet generation with improved efficiency in tetracene oligomers through spatially separated triplet pair states. *Nature Chemistry* **2021**, *13*, 559–567.

[320] Li, X.; Parrish, R. M.; Martínez, T. J. An ab initio exciton model for singlet fission. *The Journal of Chemical Physics* **2020**, *153*, 184116.

[321] Ambrosio, F.; Troisi, A. Singlet fission in linear chains of molecules. *The Journal of Chemical Physics* **2014**, *141*, 204703.

[322] Yang, C.-H.; Hsu, C.-P. First-Principle Characterization for Singlet Fission Couplings. *The journal of physical chemistry letters* **2015**, *6*, 1925–1929.

[323] Tempelaar, R.; Reichman, D. R. Vibronic exciton theory of singlet fission. I. Linear absorption and the anatomy of the correlated triplet pair state. *The Journal of Chemical Physics* **2017**, *146*, 174703.

[324] Margulies, E. A.; Logsdon, J. L.; Miller, C. E.; Ma, L.; Simonoff, E.; Young, R. M.; Schatz, G. C.; Wasielewski, M. R. Direct Observation of a Charge-Transfer State Preceding High-Yield Singlet Fission in Terrylenediimide Thin Films. *Journal of the American Chemical Society* **2017**, *139*, 663–671, PMID: 27977196.

[325] Berkelbach, T. C.; Hybertsen, M. S.; Reichman, D. R. Microscopic theory of singlet exciton fission. II. Application to pentacene dimers and the role of superexchange. *The Journal of Chemical Physics* **2013**, *138*, 114103.

[326] Berkelbach, T. C. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 2017; Chapter 1, pp 1–38.

[327] Stern, H. L.; Musser, A. J.; Gelinas, S.; Parkinson, P.; Herz, L. M.; Bruzek, M. J.; Anthony, J.; Friend, R. H.; Walker, B. J. Identification of a triplet pair intermediate in singlet exciton fission in solution. *Proceedings of the National Academy of Sciences* **2015**, *112*, 7656–7661.

[328] Dreuw, A.; Head-Gordon, M. Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chemical Reviews* **2005**, *105*, 4009–4037, PMID: 16277369.

[329] Hirao, K. Multireference Møller–Plesset perturbation treatment of potential energy curve of N2. *International Journal of Quantum Chemistry* **1992**, *44*, 517–526.

[330] Casanova, D.; Krylov, A. I. Spin-flip methods in quantum chemistry. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4326–4342.

[331] Zimmerman, P. M.; Bell, F.; Casanova, D.; Head-Gordon, M. Mechanism for Singlet Fission in Pentacene and Tetracene: From Single Exciton to Two Triplets. *Journal of the American Chemical Society* **2011**, *133*, 19944–19952, PMID: 22084927.

[332] Casanova, D. Electronic Structure Study of Singlet Fission in Tetracene Derivatives. *Journal of Chemical Theory and Computation* **2014**, *10*, 324–334, PMID: 26579913.

[333] Feng, X.; Luzanov, A. V.; Krylov, A. I. Fission of Entangled Spins: An Electronic Structure Perspective. *The Journal of Physical Chemistry Letters* **2013**, *4*, 3845–3852.

[334] Kolomeisky, A. B.; Feng, X.; Krylov, A. I. A Simple Kinetic Model for Singlet Fission: A Role of Electronic and Entropic Contributions to Macroscopic Rates. *The Journal of Physical Chemistry C* **2014**, *118*, 5188–5195.

[335] Jiang, H.; Zimmerman, P. M. Charge transfer via spin flip configuration interaction: Benchmarks and application to singlet fission. *The Journal of Chemical Physics* **2020**, *153*, 064109.

[336] Benk, H.; Sixl, H. Theory of two coupled triplet states. *Molecular Physics* **1981**, *42*, 779–801.

[337] Bloch, C. Sur La Théorie Des Perturbations Des États Liés. *Nuclear Physics* **1958**, *6*, 329–347.

[338] Des Cloizeaux, J. Extension D'une Formule De Lagrange À Des Problèmes. *Nuclear Physics* **1960**, *20*, 321–346.

[339] Liu, X.; Subotnik, J. E. The Variationally Orbital-Adapted Configuration Interaction Singles (VOA-CIS) Approach to Electronically Excited States. *Journal of Chemical Theory and Computation* **2014**, *10*, 1004–1020, PMID: 26580179.

[340] Taffet, E. J.; Beljonne, D.; Scholes, G. D. Overlap-Driven Splitting of Triplet Pairs in Singlet Fission. *Journal of the American Chemical Society* **2020**, *142*, 20040–20047, PMID: 33190497.

[341] Wakasa, M.; Kaise, M.; Yago, T.; Katoh, R.; Wakikawa, Y.; Ikoma, T. What Can Be Learned from Magnetic Field Effects on Singlet Fission: Role of Exchange Interaction in Excited Triplet Pairs. *J. Phys. Chem. C* **2015**, *119*, 25840–25844.

[342] Bayliss, S. L.; Chepelianskii, A. D.; Sepe, A.; Walker, B. J.; Ehrler, B.; Bruzek, M. J.; Anthony, J. E.; Greenham, N. C. Geminate and Nongeminate Recombination of Triplet Excitons Formed by Singlet Fission. *Phys. Rev. Lett.* **2014**, *112*, 238701.

[343] Wan, Y.; Guo, Z.; Zhu, T.; Yan, S.; Johnson, J.; Huang, L. Cooperative singlet and

triplet exciton transport in tetracene crystals visualized by ultrafast microscopy. *Nat. Chem.* **2015**, *7*, 785–792.

[344] Srimath Kandada, A. R.; Petrozza, A.; Lanzani, G. Ultrafast dissociation of triplets in pentacene induced by an electric field. *Phys. Rev. B* **2014**, *90*, 075310.

[345] Trinh, M. T.; Pinkard, A.; Pun, A. B.; Sanders, S. N.; Kumarasamy, E.; Sfeir, M. Y.; Campos, L. M.; Roy, X.; Zhu, X.-Y. Distinct properties of the triplet pair state from singlet fission. *Science Advances* **2017**, *3*, e1700241.

[346] Korovina, N. V.; Pompetti, N. F.; Johnson, J. C. Lessons from intramolecular singlet fission with covalently bound chromophores. *The Journal of Chemical Physics* **2020**, *152*, 040904.

[347] Korovina, N. V.; Chang, C. H.; Johnson, J. C. Spatial separation of triplet excitons drives endothermic singlet fission. *Nature Chemistry* **2020**, *12*, 391–398.

[348] Tao, G.; Tan, Y. Modular Tensor Diagram Approach for the Construction of Spin Eigenfunctions: The Case Study of Exciton Pair States. *The Journal of Physical Chemistry A* **2020**, *124*, 5435–5443, PMID: 32551608.

[349] Grieco, C.; Doucette, G. S.; Munro, J. M.; Kennehan, E. R.; Lee, Y.; Rimshaw, A.; Payne, M. M.; Wonderling, N.; Anthony, J. E.; Dabo, I.; Gomez, E. D.; Asbury, J. B. Triplet Transfer Mediates Triplet Pair Separation during Singlet Fission in 6,13-Bis(triisopropylsilylethynyl)-Pentacene. *Advanced Functional Materials* **2017**, *27*, 1703929.

[350] Hartzler, D. A.; Slipchenko, L. V.; Savikhin, S. Triplet–Triplet Coupling in Chromophore Dimers: Theory and Experiment. *The Journal of Physical Chemistry A* **2018**, *122*, 6713–6723, PMID: 30040412.

[351] Yost, S. R.; Hontz, E.; Yeganeh, S.; Van Voorhis, T. Triplet vs Singlet Energy Transfer in Organic Semiconductors: The Tortoise and the Hare. *The Journal of Physical Chemistry C* **2012**, *116*, 17369–17377.

[352] Fujimoto, K. J. Transition-density-fragment interaction combined with transfer integral approach for excitation-energy transfer via charge-transfer states. *The Journal of Chemical Physics* **2012**, *137*, 034101.

[353] qiang You, Z.; Hsu, C.-P. Theory and calculation for the electronic coupling in excitation energy transfer. *International Journal of Quantum Chemistry* **2014**, *114*, 102–115.

[354] Pun, A. B.; Asadpoordarvish, A.; Kumarasamy, E.; Tayebjee, M. J.; Niesner, D.; McCamey, D. R.; Sanders, S. N.; Campos, L. M.; Sfeir, M. Y. Ultra-fast intramolecular singlet fission to persistent multiexcitons by molecular design. *Nature chemistry* **2019**, *11*, 821–828.

[355] Della Valle, R. G.; Venuti, E.; Brillante, A.; Girlando, A. Inherent Structures of Crystalline Tetracene. *The Journal of Physical Chemistry A* **2006**, *110*, 10858–10862, PMID: 16970382.

[356] Korovina, N. V.; Joy, J.; Feng, X.; Feltenberger, C.; Krylov, A. I.; Bradforth, S. E.; Thompson, M. E. Linker-Dependent Singlet Fission in Tetracene Dimers. *Journal of the American Chemical Society* **2018**, *140*, 10179–10190, PMID: 30016102.

[357] Valeev, E. F.; Coropceanu, V.; da Silva Filho, D. A.; Salman, S.; Brédas, J.-L. Effect of Electronic Polarization on Charge-Transport Parameters in Molecular Organic Semiconductors. *Journal of the American Chemical Society* **2006**, *128*, 9882–9886, PMID: 16866546.

[358] Sriram Shastry, B.; Sutherland, B. Exact ground state of a quantum mechanical antiferromagnet. *Physica B+C* **1981**, *108*, 1069–1070.

[359] Mayhall, N. J.; Goldey, M.; Head-Gordon, M. A Quasidegenerate Second-Order Perturbation Theory Approximation to RAS-nSF for Excited States and Strong Correlations. *Journal of Chemical Theory and Computation* **2014**, *10*, 589–599, PMID: 26580035.

[360] Mayhall, N. J.; Head-Gordon, M. Increasing spin-flips and decreasing cost: Perturbative corrections for external singles to the complete active space spin flip model for low-lying excited states and strong correlation. *The Journal of Chemical Physics* **2014**, *141*, 044112.

[361] Tempelaar, R.; Reichman, D. R. Vibronic exciton theory of singlet fission. III. How vibronic coupling and thermodynamics promote rapid triplet generation in pentacene crystals. *The Journal of Chemical Physics* **2018**, *148*, 244701.

[362] Ho, J.; Kish, E.; Méndez-Hernández, D. D.; WongCarter, K.; Pillai, S.; Kodis, G.; Niklas, J.; Poluektov, O. G.; Gust, D.; Moore, T. A.; Moore, A. L.; Batista, V. S.; Robert, B. Triplet–triplet energy transfer in artificial and natural photosynthetic antennas. *Proceedings of the National Academy of Sciences* **2017**, *114*, E5513–E5521.

[363] Tejeda-Ferrari, M. E.; Brown, C. L.; Coutinho, G. C. C. C.; Gomes de Sá, G. A.; Palma, J. L.; Llansola-Portoles, M. J.; Kodis, G.; Mujica, V.; Ho, J.; Gust, D.; Moore, T. A.; Moore, A. L. Electronic Structure and Triplet–Triplet Energy Transfer in Artificial Photosynthetic Antennas. *Photochemistry and Photobiology* **2019**, *95*, 211–219.

[364] Lindgren, I. The Rayleigh-Schrodinger perturbation and the linked-diagram theorem for a multi-configurational model space. *Journal of Physics B: Atomic and Molecular Physics* **1974**, *7*, 2441–2470.

[365] Kuramoto, Y. *Quantum Many-Body Physics*; Springer, 2020; pp 1–11.

[366] Dover, C. B.; Gallaher, J. K.; Frazer, L.; Tapping, P. C.; Petty II, A. J.; Crossley, M. J.; Anthony, J. E.; Kee, T. W.; Schmidt, T. W. Endothermic singlet fission is hindered by excimer formation. *Nature chemistry* **2018**, *10*, 305.

[367] Bayliss, S. L. et al. Site-selective measurement of coupled spin pairs in an organic semiconductor. *Proc. Natl. Acad. Sci.* **2018**, *115*, 5077–5082.

[368] Geacintov, N.; Pope, M.; Vogel, F. Effect of Magnetic Field on the Fluorescence of Tetracene Crystals: Exciton Fission. *Phys. Rev. Lett.* **1969**, *22*, 593–596.

[369] Swenberg, C.; Stacy, W. Bimolecular radiationless transitions in crystalline tetracene. *Chem. Phys. Lett.* **1968**, *2*, 327–328.

[370] Singh, S.; Jones, W. J.; Siebrand, W.; Stoicheff, B. P.; Schneider, W. G. Laser Generation of Excitons and Fluorescence in Anthracene Crystals. *J. Chem. Phys.* **1965**, *42*, 330.

[371] Merrifield, R.; Avakian, P.; Groff, R. Fission of singlet excitons into pairs of triplet excitons in tetracene crystals. *Chem. Phys. Lett.* **1969**, *3*, 386–388.

[372] Thorsmølle, V.; Averitt, R.; Demsar, J.; Smith, D.; Tretiak, S.; Martin, R.; Chi, X.; Crone, B.; Ramirez, A.; Taylor, A. Photoexcited carrier relaxation dynamics in pentacene probed by ultrafast optical spectroscopy: Influence of morphology on relaxation processes. *Physica B: Condensed Matter* **2009**, *404*, 3127 – 3130.

[373] Burdett, J. J.; Müller, A. M.; Gosztola, D.; Bardeen, C. J. Excited state dynamics in solid and monomeric tetracene: The roles of superradiance and exciton fission. *J. Chem. Phys.* **2010**, *133*, 144506.

[374] Burdett, J. J.; Bardeen, C. J. Quantum beats in crystalline tetracene delayed fluorescence due to triplet pair coherences produced by direct singlet fission. *J. Am. Chem. Soc.* **2012**, *134*, 8597–607.

[375] Kuhlman, T. S.; Kongsted, J.; Mikkelsen, K. V.; Møller, K. B.; Sølling, T. I. Interpretation of the Ultrafast Photoinduced Processes in Pentacene Thin Films. *J. Am. Chem. Soc.* **2010**, *132*, 3431–3439, PMID: 20175536.

[376] Roberts, S. T.; McAnally, R. E.; Mastron, J. N.; Webber, D. H.; Whited, M. T.; Brutchey, R. L.; Thompson, M. E.; Bradforth, S. E. Efficient singlet fission discovered in a disordered acene film. *J. Am. Chem. Soc.* **2012**, *134*, 6388–6400.

[377] Jundt, C.; Klein, G.; Sipp, B.; Moigne, J. L.; Joucla, M.; Villaeys, A. Exciton dynamics in pentacene thin films studied by pump-probe spectroscopy. *Chem. Phys. Lett.* **1995**, *241*, 84 – 88.

[378] Müller, A. M.; Avlasevich, Y. S.; Müllen, K.; Bardeen, C. J. Evidence for exciton fission and fusion in a covalently linked tetracene dimer. *Chem. Phys. Lett.* **2006**, *421*, 518–522.

[379] Müller, A. M.; Avlasevich, Y. S.; Schoeller, W. W.; Müllen, K.; Bardeen, C. J. Exciton fission and fusion in bis(tetracene) molecules with different covalent linker structures. *J. Am. Chem. Soc.* **2007**, *129*, 14240–14250.

[380] Sakuma, T.; Sakai, H.; Araki, Y.; Mori, T.; Wada, T.; Tkachenko, N. V.; Hasobe, T. Long-Lived Triplet Excited States of Bent-Shaped Pentacene Dimers by Intramolecular Singlet Fission. *J. Phys. Chem. A* **2016**, *120*, 1867–1875.

[381] Lukman, S.; Musser, A. J.; Chen, K.; Athanasopoulos, S.; Yong, C. K.; Zeng, Z.; Ye, Q.; Chi, C.; Hodgkiss, J. M.; Wu, J.; Friend, R. H.; Greenham, N. C. Tuneable

Singlet Exciton Fission and Triplet-Triplet Annihilation in an Orthogonal Pentacene Dimer. *Adv. Funct. Mater.* **2015**, *25*, 5452–5461.

[382] Lukman, S. et al. Tuning the role of charge-transfer states in intramolecular singlet exciton fission through side-group engineering. *Nat. Commun.* **2016**, *7*, 13622.

[383] Chien, A. D.; Molina, A. R.; Abeyasinghe, N.; Varnavski, O. P.; Goodson, T.; Zimmerman, P. M. Structure and Dynamics of the 1 (TT) State in a Quinoidal Bithiophene: Characterizing a Promising Intramolecular Singlet Fission Candidate. *J. Phys. Chem. C* **2015**, *119*, 28258–28268.

[384] Feng, X. et al. On couplings and excimers: lessons from studies of singlet fission in covalently linked tetracene dimers. *Phys. Chem. Chem. Phys.* **2016**, *18*, 7751–7761.

[385] Zeng, T. Through-Linker Intramolecular Singlet Fission: General Mechanism and Designing Small Chromophores. *J. Phys. Chem. Lett.* **2016**, *7*, 4405–4412.

[386] Zirzlmeier, J.; Lehnherr, D.; Coto, P. B.; Chernick, E. T.; Casillas, R.; Basel, B. S.; Thoss, M.; Tykwinski, R. R.; Guldi, D. M. Singlet fission in pentacene dimers. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 5325–30.

[387] Zirzlmeier, J.; Casillas, R.; Reddy, S.; Coto, P.; Lehnherr, D.; Chernick, E.; Papadopoulos, I.; Thoss, M.; Tykwinski, R.; Guldi, D. Solution-based intramolecular singlet fission in cross-conjugated pentacene dimers. *Nanoscale* **2016**, *8*, 10113–10123.

[388] Feng, X.; Krylov, A. I. On couplings and excimers: lessons from studies of singlet fission in covalently linked tetracene dimers. *Phys. Chem. Chem. Phys.* **2016**, *18*, 7751–61.

[389] Kolomeisky, A. B.; Feng, X.; Krylov, A. I. A Simple Kinetic Model for Singlet Fission: A Role of Electronic and Entropic Contributions to Macroscopic Rates. *J. Phys. Chem. C* **2014**, *118*, 5188–5195.

[390] Zimmerman, P. M.; Musgrave, C. B.; Head-Gordon, M. A Correlated Electron View of Singlet Fission. *Acc. Chem. Res.* **2013**, *46*, 1339–1347.

[391] Havenith, R. W. A.; de Gier, H. D.; Broer, R. Explorative computational study of the singlet fission process. *Mol. Phys.* **2012**, *110*, 2445–2454.

[392] Yost, S. R. et al. A transferable model for singlet-fission kinetics. *Nature Chemistry* **2014**, *6*, 492–497.

[393] Greyson, E. C.; Vura-Weis, J.; Michl, J.; Ratner, M. A. Maximizing singlet fission in organic dimers: theoretical investigation of triplet yield in the regime of localized excitation and fast coherent electron transfer. *J. Phys. Chem. B* **2010**, *114*, 14168–77.

[394] Teichen, P. E.; Eaves, J. D. A microscopic model of singlet fission. *J. Phys. Chem. B* **2012**, *116*, 11473–81.

[395] Piland, G. B.; Burdett, J. J.; Dillon, R. J.; Bardeen, C. J. Singlet Fission: From Coherences to Kinetics. *J. Phys. Chem. Lett.* **2014**, *5*, 2312–2319.

[396] Aryanpour, K.; Shukla, A.; Mazumdar, S. Theory of Singlet Fission in Polyenes, Acene Crystals, and Covalently Linked Acene Dimers. *J. Phys. Chem. C* **2015**, *119*, 6966–6979.

[397] Chan, W.-L.; Berkelbach, T. C.; Provorse, M. R.; Monahan, N. R.; Tritsch, J. R.; Hybertsen, M. S.; Reichman, D. R.; Gao, J.; Zhu, X.-Y. The quantum coherent mechanism for singlet fission: experiment and theory. *Acc. Chem. Res.* **2013**, *46*, 1321–9.

[398] Renaud, N.; Grozema, F. C. Intermolecular vibrational modes speed up singlet fission in perylenediimide crystals. *J. Phys. Chem. Lett.* **2015**, *6*, 360–365.

[399] Yang, C. H.; Hsu, C. P. First-Principle Characterization for Singlet Fission Couplings. *J. Phys. Chem. Lett.* **2015**, *6*, 1925–1929.

[400] Berkelbach, T. C.; Hybertsen, M. S.; Reichman, D. R. Microscopic theory of singlet exciton fission. I. General formulation. *J. Chem. Phys.* **2013**, *138*, 114102.

[401] Berkelbach, T. C.; Hybertsen, M. S.; Reichman, D. R. Microscopic theory of singlet exciton fission. II. Application to pentacene dimers and the role of superexchange. *J. Chem. Phys.* **2013**, *138*, 114103.

[402] Berkelbach, T. C.; Hybertsen, M. S.; Reichman, D. R. Microscopic theory of singlet exciton fission. III. Crystalline pentacene. *J. Chem. Phys.* **2014**, *141*, 074705.

[403] Minami, T.; Nakano, M. Diradical Character View of Singlet Fission. *J. Phys. Chem. Lett.* **2012**, *3*, 145–150.

[404] Minami, T.; Ito, S.; Nakano, M. Theoretical Study of Singlet Fission in Oligorylenes. *J. Phys. Chem. Lett.* **2012**, *3*, 2719–23.

[405] Ito, S.; Nagami, T.; Nakano, M. Design Principles of Electronic Couplings for Intramolecular Singlet Fission in Covalently-Linked Systems. *J. Phys. Chem. A* **2016**, *120*, 6236–6241.

[406] Ovchinnikov, A. A. Multiplicity of the ground state of large alternant organic molecules with conjugated bonds - (Do Organic Ferromagnetics Exist?). *Theor. Chim. Acta.* **1978**, *47*, 297–304.

[407] Lieb, E.; Mattis, D. Ordering Energy Levels of Interacting Spin Systems. *J. Mat. Phys.* **1962**, *3*, 749–751.

[408] Siebrand, W. Radiationless Transitions in Polyatomic Molecules. II. Triplet-Ground-State Transitions in Aromatic Hydrocarbons. *J. Chem. Phys.* **1967**, *47*, 2411–2422.

[409] Hajgató, B.; Szieberth, D.; Geerlings, P.; De Proft, F.; Deleuze, M. S. A benchmark theoretical study of the electronic ground state and of the singlet-triplet split of benzene and linear acenes. *J. Chem. Phys.* **2009**, *131*, 224321.

[410] Fosso-Tande, J.; Nguyen, T.-S.; Gidofalvi, G.; DePrince, A. E. Large-Scale Variational Two-Electron Reduced-Density-Matrix-Driven Complete Active Space Self-Consistent Field Methods. *J. Chem. Theory Comput.* **2016**, *12*, 2260–2271.

[411] Casanova, D.; Slipchenko, L. V.; Krylov, A. I.; Head-Gordon, M. Double spin-flip approach within equation-of-motion coupled cluster and configuration interaction formalisms: Theory, implementation, and examples. *J. Chem. Phys.* **2009**, *130*, 044103.

[412] Mayhall, N. J.; Goldey, M.; Head-Gordon, M. A quasidegenerate second-order perturbation theory approximation to RAS- n SF for excited states and strong correlations. *J. Chem. Theory Comput.* **2014**, *10*, 589–599.

[413] Mayhall, N. J.; Head-Gordon, M. Adding dynamical correlation to spin-complete spin-flip wavefunctions via second-order perturbation theory. *to be published elsewhere*

[414] Ito, S.; Nagami, T.; Nakano, M. Rational design of doubly-bridged chromophores for singlet fission and triplet-triplet annihilation. *RSC Adv.* **2017**, *7*, 34830–34845.

[415] Radenković, S.; Shaik, S. S.; Braïda, B. Na B Bond in NaBH3—: Solving the Conundrum. *Angewandte Chemie International Edition* **2021**, *60*, 12723–12726.