UC Berkeley
UC Berkeley Electronic Theses and Dissertations

**Title**
Excitonic Structure and Energy Transfer in Photosynthetic Pigment Protein Complexes and Their Assemblies

**Permalink**
https://escholarship.org/uc/item/34s6r7wj

**Author**
Bennett, Doran I G

**Publication Date**
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

# Excitonic Structure and Energy Transfer in Photosynthetic Pigment Protein Complexes and Their Assemblies

by

Doran Bennett

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Graham R. Fleming, Chair

Professor K. Birgitta Whaley

Professor Joe Orenstein

Spring 2013

Excitonic Structure and Energy Transfer in Photosynthetic Pigment Protein Complexes and Their Assemblies

Copyright 2013

by

Doran Bennett

Abstract

Excitonic Structure and Energy Transfer in Photosynthetic Pigment Protein Complexes and Their Assemblies

by

Doran Bennett

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Graham R. Fleming, Chair

The research presented in this dissertation focuses on understanding how the spatial and energetic organization of pigment molecules within photosynthetic assemblies create efficient, natural light harvesting machines. In the first chapter we briefly describe photosynthetic light harvesting and explain the basic features of third-order spectroscopy and time-resolved fluorescence measurements used in this dissertation to characterize the relationship between the protein-bound pigments and their light harvesting processes. In chapter two, we start studying photosynthesis at the length scale of a single pigment-protein complex (PPC) and introduce a novel third-order frequency generation technique that is very sensitive to the excitonic structure of a PPCs. This technique is applied to a model pigment binding protein, the Fenna-Matthews-Olson (FMO) complex. In the third chapter, we propose a new approach to extracting information from optical, third-order measurements by combining chirplet transforms and shaped excitation pulses to glean information usually buried in the wings of the 2D electronic spectrum. Finally, in the fourth chapter, we consider longer length scale assemblies and treat photosystem II (PSII) supercomplexes which are composed of many individual PPCs. Using information gathered on isolated PPCs, we construct the first structure-based model of energy transfer in PSII supercomplexes to study the principles of light harvesting in functional photosynthetic machinery.

To Lee Marek, who planted the thought that I might want dabble in Chemistry at college,
And to Laurie Butler, who convinced me that Chemistry was fun enough to turn into a career.

# Contents

**List of Figures** **iv**

**List of Tables** **x**

**1 Introduction** **1**
1.1 Light Harvesting in Photosynthesis . . . . . 1
1.2 Experimental Probes of Excitonic Structure and Energy Transfer in Pigment Protein Complexes . . . . . 3
1.2.1 $3^{rd}$-order Measurements . . . . . 3
1.2.2 Time-Correlated Single Photon Counting (TCSPC) . . . . . 8

**2 Mapping the spatial overlap of excitons in a photosynthetic complex via coherent nonlinear frequency generation** **10**
2.1 Introduction . . . . . 10
2.2 Two-exciton states of a photosynthetic complex . . . . . 12
2.3 Nonlinear Frequency Generation Mediated by Two-Exciton States . . . . . 17
2.3.1 Simulation of the Signal for Model Systems . . . . . 19
2.4 Experimental Implementation . . . . . 22
2.5 Calculation of Frequency Shifted Nonlinear Signal for FMO . . . . . 26
2.5.1 Discussion . . . . . 26
2.6 Conclusion . . . . . 31

**3 Shaped pulse third-order spectroscopy: Chirp as a measurement axis** **32**
3.1 Introduction . . . . . 32
3.2 Chirplet Transform of Third-Order Polarization . . . . . 33
3.2.1 Third-order polarization in the semi-impulsive limit . . . . . 34
3.2.2 Chirplet Spectra . . . . . 35
3.3 Application of Chirplet Transform . . . . . 36
3.3.1 Simple Third-Order Response Function . . . . . 36
3.3.2 Frequency Dependent Dephasing . . . . . 38
3.3.3 Simulated Spectra . . . . . 41
3.4 Conclusion . . . . . 43

**4 A structure-based model of energy transfer reveals the principles of excitation migration and trapping in Photosystem II supercomplexes 45**
4.1 Introduction . . . 45
4.2 Methods . . . 47
4.2.1 Excitation Energy Transfer Theory . . . 48
4.2.2 Coarse-grain Treatments of Energy Transfer . . . 53
4.2.3 Photosystem II: Parameters . . . 55
4.2.4 Electron Transfer Scheme . . . 56
4.2.5 Simulating Fluorescence Decays . . . 57
4.2.6 Calculating Linearized Kinetics . . . 58
4.3 Model of Photosystem II . . . 59
4.4 Results & Discussion . . . 61
4.4.1 Timescales of Light Harvesting in $C_2S_2M_2$ . . . 61
4.4.2 Spatial Aspects of Light Harvesting in $C_2S_2M_2$ . . . 64
4.4.3 Quality of Fit to Fluorescence Decays is not a Proxy for Accuracy of Energy Transfer Models . . . 66
4.5 Conclusions . . . 67

**Bibliography 69**

**A Calculation of the nonlinear response functions 80**

**B Properties of Chirplet Transform 84**
B.1 Definition . . . 84
B.2 1D Subspaces . . . 84
B.2.1 $t$ Space . . . 85
B.2.2 $\omega$ Space . . . 85
B.3 2D Subspaces . . . 85
B.3.1 $(t, \omega)$ Space . . . 86
B.3.2 $(t, \beta)$ Space . . . 87

**C Photosystem II Parameters and Error Analysis 91**
C.1 Justification of protein structures used . . . 91
C.2 Phototsystem II Parameters . . . 92
C.2.1 Hamiltonians . . . 92
C.2.2 Spectral Density . . . 96
C.3 Testing Infinitely Fast Intra-Domain Equilibration . . . 98
C.4 Coarse Grained Models of Energy Transfer . . . 98
C.5 Sample Preparation and Modeling . . . 99

# List of Figures

1.1 Material Response Time Diagram. . . . . 4
1.2 The box car measurement geometry. Figure modified from Ref. [34]. . . . . 6
1.3 Absolute value 2D spectra of the FMO trimer. Image taken from Ref. [31] . 7
1.4 The basic experimental setup for TCSPC measurements. Figure courtesy of Emily-Jane Sylak Glassman. . . . . 8
1.5 The data from TCSPC measurements bins the measured photons by their time delay from the initial excitation pulse. Figure courtesy of Emily-Jane Sylak Glassman. . . . . 9

2.1 (a) A three-pigment model system with one weakly coupled pigment. All energies are in $cm^{-1}$ and the pigment energies are relative to 12550 $cm^{-1}$. (b) The one-exciton eigenstates represented in the one-pigment excitation basis (the coefficients $U^{-1}_{jm}$ in Eq. 2.6). $|e_1\rangle$ is primarily localized on pigment 3, while $|e_2\rangle$ and $|e_3\rangle$ represent delocalized out-of-phase and in-phase excitations of pigments 1 and 2. (c) The two-exciton eigenstates represented in the two-pigment excitation basis (the coefficients $v^{(k)}_{mn}$ in Eq. 2.7). The character of two-exciton states in terms of single excitons can be seen in this representation. For example $|f_1\rangle$ shares character with $|e_1\rangle$ and $|e_2\rangle$ (d) The transition dipole moment between the two- and one-exciton states represented as circles. The character of the two-exciton states in terms of single-excitons is much easier to discern (e.g. $|f_1\rangle \approx |e_1e_2\rangle$). Each two-exciton state is primarily composed of two single excitons. . . . . 15
2.2 (a) A three-pigment model system similar to Fig. 2.1, but with strongly bound pigment 3. (b) The one-exciton (b) and two-exciton wavefunctions (c) are more delocalized over all three pigments. (d) The transition dipole moment between the two- and one-exciton states reveals that each one of the two-exciton exciton states shares character with all three single excitons. Such shared character of the two-exciton states is manifested in nonlinear frequency generation as explained in the text. . . . . 16

2.3 (a) A representative FWM pathway that leads to frequency generation in a three-pigment system. Two interactions with optical frequencies $-\omega_1$ and $-\omega_2$ generates a coherence between the two-exciton state $|f_3\rangle$ and the ground state $|g\rangle$. Interaction with frequency $\omega_3$ de-excites the system to a single exciton state $|e_1\rangle$. The FWM signal at frequency $\omega_s = -\omega_1 - \omega_2 + \omega_3$ is red-shifted with respect to all excitation frequencies. The third interaction relies upon the transition dipole matrix element $\langle f_3|\hat{\mu}|e_1\rangle$. (b) In the case of strongly overlapping excitons, the mixed character of $|f_3\rangle$ allows a non-zero transition dipole to state $|e_1\rangle$, making signal generation possible. (c) When exciton $|e_1\rangle$ resides on an isolated pigment such a downward transition is not allowed, as $|f_3\rangle$ does not share any character with $|e_1\rangle$. . . . . 18

2.4 Feynman diagrams for the third order signal in the $-k_1 + k_2 + k_3$ direction (Eq. 2.16). The $S_R, S_{NR}$ and $S_{DQC}$ terms differ in the temporal order of the action of the field contributing with negative sign $(-k_1)$. Under pulse overlap where all permutations of time ordering is possible, all of the above pathways contribute to the signal. However, only the diagrams that involve the two-exciton states (labeled with subscript $f$) can lead to frequency generation outside the excitation bandwidth. . . . . 20

2.5 Calculation of frequency shifted signal for two models systems. (a) A trimer, with a weakly bound lowest energy pigment. All energies are in $cm^{-1}$, with pigment energies relative to 12550 $cm^{-1}$. (b) The linear absorption spectrum (blue) and the superposed laser excitation pulse (red). (c) The calculated $S_{2exc}$ signal shows small response outside the laser excitation band. (d-e) The same scenario for a trimer with strongly bound pigments. (f) Frequency generation outside the excitation energy band (peak at the frequency of the lowest energy pigment) is seen in the calculated $S_{2exc}$ signal. . . . . 22

2.6 Coherent multidimensional spectroscopy apparatus employed in this experiment. Four beams are sampled from an expanded laser beam and guided to a 2D reflective phase modulator based 4-f optical pulse shaper. The diffracted beams (dotted lines) from the 4f pulse shaper is retrieved with a pick-off mirror (not shown), re-collimated and focused on the sample. The pulse shaper is able to modify the spectral phase and amplitude of the pulses, leading to temporal pulse shaping. Pulse delays are also executed by the phase modulator. The nonlinear signal is retrieved in the phase-matched direction and sent to a spectrometer and spectrally dispersed on a CCD camera. . . . . 23

2.7 (a) The spectrum of excitation pulses (blue) overlaid on the FMO absorption spectrum (black). (b) The signal emerging in the $-k_1 + k_2 + k_3$ direction. A long pass filter with cutoff at 815 nm is used to only collect the frequency shifted signal. The nonlinear nature of the peak at 817 nm is shown by comparing with cases when only one of the beams is exciting the sample. In comparison, incoherent fluorescence from the lowest energy exciton (at 825 nm) shows linear dependence on the presence of the three beams. . . . . 25

2.8 The dependence of the intensity of the peak at 817 nm on the relative delay of the pulse with wavevector $+k_3$. . . . . 25
2.9 Calculation of the frequency shifted nonlinear signal for FMO. (a) Simulated linear absorption spectrum of FMO (black) with the laser excitation spectrum (red) used in the calculations. (b) The nonlinear signal based on the $S_{2exc}$ pathways (Eq. 2.17) shows a frequency shifted peak near 817 nm. (c) The frequency shifted peak persists when the signal based on the sum of all pathways is calculated. (d) The nonlinear signal when a filter similar to the one used in the experiment (Fig. 2.7) is applied. . . . . 27
2.10 Representative four-wave-mixing diagram corresponding to pathway $R_{4f}$ (Fig. 2.4). The lowest and highest frequency components of laser pulses are indicated by $\omega_L$ and $\omega_H$ respectively. The two-exciton states can mediate generation of a signal at a frequency outside the laser bandwidth. The frequency shifted signal depends on transition matrix elements between the one and two-exciton states and carries information about the mixed character of the two-exciton states. . . . . 28
2.11 The magnitude of the transition dipole matrix elements $\langle f_k|\hat{\mu}|e_j\rangle$ (represented by circles) between the 21 two-exciton states and 7 one-exciton states of FMO. The sizes follow a log scale for better visualization. The mixed nature of the two-exciton states allows them to have non-zero transitions matrix elements with more than two single excitons. One-exciton states $|e_5\rangle$-$|e_7\rangle$ (and $|e_4\rangle$ due to inhomogeneous width) can be excited after a single optical interaction. After two optical interactions, states $|f_{15}\rangle$ -$|f_{20}\rangle$ become accessible. The mixed nature of these states allows for de-excitation into single excitons in the detection band. The transition dipole matrix elements enclosed within the red square are responsible for generation of frequency shifted signal near 817 nm (Fig. 2.7). . . . . 29
2.12 The spatial extent of the wavefunctions of the single excitons of FMO. The radius of circles represents the magnitudes of the contribution of each pigment to the exciton (coefficients less that 7% are not shown). Exciton $|e_2\rangle$ in the detection range has overlap with excitons $|e_4\rangle$-$|e_6\rangle$ in the excitation band (sharing pigment 4, 5 and 7), which results into their coupling in the two-exciton states and giving rise to the frequency shifted signal. Exciton $|e_1\rangle$ does not have a significant overlap with the excitons in the excitation band, consequently it can not give rise to frequency shifted signal. . . . . 30

3.1 The measurement geometry for a chirplet spectra. Figure modified from Ref. [34]. . . . . 36

3.2 A schematic response function (whose envelope is shown in black) is the average of two contributions: a slowly dephasing high-frequency function and a quickly dephasing low-frequency response. The result is an overall chirp within the response function as the average response goes from a central frequency that is the average of the two central frequencies of the red and blue terms in the short time to a central frequency that is the blue terms central frequency in long time. . . . . 39
3.3 $\Gamma = \Gamma_0 + \tilde{\alpha} \cdot \mathrm{erf}(\frac{\omega_{eg}-\omega_0}{\gamma\sigma_0})$ is plotted for $\gamma = 2$ in red (left axis) to show that the value is now bounded so that values of the system parameters can be chosen to ensure no portion of the inhomogeneous ensemble experiences negative dephasing. The amplitude of the inhomogeneous distribution is plotted in blue (right axis). . . . . 40
3.4 The absolute value 2D rephasing spectra (a-c,e,f) of the response functions characterizing a material with frequency dependent dephasing. The figures in the top-row were calculated with $\gamma = 2$, the figures in the bottom row were calculated with $\gamma = 1$. The chirplet spectra shown in the left, middle, and right columns were calculated with $\alpha = 0$, $0.6\Gamma_0$, and $\Gamma_0$, respectively. All 2D spectra are nearly identical by eye. Figure (d) shows anti-diagonal slices taken from the left and right edges of a 2D spectra with $\alpha = 0$ (blue) and $\alpha = \Gamma_0$ (red) which quantify the small change in line shape that results from frequency dependent dephasing. . . . . 42
3.5 The chirplet spectra (a-c,e,f) of the response functions characterizing a material with frequency dependent dephasing. The figures in the top-row were calculated with $\gamma = 2$, the figures in the bottom row were calculated with $\gamma = 1$. The chirplet spectra shown in the left, middle, and right columns were calculated with $\alpha = 0$, $0.6\Gamma_0$, and $\Gamma_0$, respectively. The chirplet spectra are matched to the 2D spectra shown in Figure 3.4. Figure (d) shows a slice along the $\beta$-axis at the peak value along $t_1$ of two chirplet spectra, $\alpha = 0$ (blue) and $\alpha = \Gamma_0$ (red). . . . . 43
3.6 The overlap between a chirped pulse and a chirped response function (either of (a) the same sign or (b) different sign) is displayed for a long time delay. (a) When the two functions share the same sign of chirp, then at long delay time the pulses blue edge overlaps with the response functions red edge. This provides minimal off-resonance excitation. (b) When the two functions have the opposite sign of chirp, then at long delay time the pulses red edge overlaps with the response functions red edge which provides favorable on-resonance excitation. . . . . 44

4.1 The $C_2S_2M_2$ protein structure and pigment arrangement are shown based on the structure determined by Caffarri and coworkers [18]. (a) The protein and pigment arrangement are shown. Chlorophylls are represented as blue spheres outlining the chlorin ring. Thick (thin) black outlines surround proteins associated with the right (left) monomer unit of the $C_2S_2M_2$. (b) The protein scaffold has been removed to expose the pigments bound by each protein unit. Pigments belonging to domains contained within a single protein unit are shown in grey. Pigments assigned to a domain delocalized between more than one protein are colored according to their domain assignment. . . . . 48
4.2 The molecular structure of the three pigments (chlorophyll a, chlorophyll b, and pheophytin) that we treat in our energy transfer model for photosystem II supercomplexes. . . . . 49
4.3 Schematic of different coarse-grain treatments of energy transfer for PSII supercomplexes. The left panel shows the collection of pigments which is described in the full model by a rate matrix that containing energy transfer rates between each exitonic state. The middle panel shows the domain assignments of these pigments. Arrows between the domains show the rates of energy transfer included in the domain rate matrix. The right panel shows a compartment assignment, where two domains are assigned to one compartment and the third domain is assigned to another. . . . . 54
4.4 Charge separation model. RC represents the reaction center domain, which contains four chlorophyll (two special pair and two associated chlorophyll) and two pheophytin. Each reaction center in a PSII supercomplex performs electron transfer using the same rate constants. . . . . 57
4.5 (a) A cartoon structure of PSII supercomplex proteins is shown. The contributions of different supercomplexes to each band are represented by similar cartoons in the top-right corner of each box. A comparison between experimental (black line, taken from Caffari and coworkers[17]) and simulated (orange and green dashed-lines for B7 and B8-B11, respectively, calculated with the domain model) fluorescence decay curves for (b) B7, (c) B8, (d) B9, (e) B10, and (f) B11. The simulated decay curves are calculated using the best fit electron transfer timescales given in Table 4.1. The experimental curves discard the longest lifetime component following the work by Caffarri and coworkers (discussed in Appendix C). . . . . 61
4.6 A diagram of the linear excitation energy cascade is shown where $\sigma_0$ represents RP2 in $C_2S_2M_2$. The effective timescales of transfer between compartments ($\tau^{\text{eff}}_{i\leftarrow i+1}$) is the average time of transfer from compartment $i$ to $i-1$ averaging over all pathways starting at compartment $i$. . . . . 63

4.7 The timescale of light harvesting ($\tau_{\mathrm{LH}}$) for excitation initiated in each domain is plotted against the average distance from that domain to the nearest reaction center. The black dots denote the results from one inhomogeneous realization of the site energies. The colored contours (red indicating more points, blue indicating less) show the underlying distribution extracted from compartments calculated from 500 inhomogeneous realizations of the site energies. The underlying distribution was calculated with 2D Gaussian smoothing. 66

B.1 A sinusoidal F(t) function (black line), which is convolved with a pulse described by a gaussian envelope (red line), linear phase (blue line), and quadratic phase (magenta line). . . . . 85
B.2 (A) the TF space representation of a pure tone, (B) the T space signal, (C) the TF space representation of a chirped tone (the extra structure is the result of finite window), and (D) the T space representation of the chirp tone. . . . 86
B.3 (A) the $(t, \beta)$-space representation of a pure tone, (B) the $t$-space signal, (C) the $(t, \beta)$-space representation of a chirped tone (the extra structure is the result of finite window), and (D) the t-space chirped signal. . . . . 88
B.4 Proceeding from left to right in a row, the response function changes from unwindowed at the left to tightly windowed at the right. Proceeding from top to bottom in a column, the amplitude envelope of a pulse changes from tightly windowed at the top to almost unwindowed at the bottom. This figure explores how the chirplet transform behaves as the two window widths are varied. 89
B.5 Chirped function overlap is shown to explain the asymmetry between the positive and negative chirp directions in a $(t, \beta)$-space chirplet transform. . . 90

C.1 A comparison between the domain population curves with the largest value $\Delta^{(d)}_{\mathrm{max}}$ and $\Delta^{(d)}_{\mathrm{Integral}}$. Population trace calculated with the generalized Förster/modified Redfield (Domain) model is shown with a dashed black (sold grey) line. . . 99
C.2 A comparison between different coarse grain models of $C_2S_2M_2$ is shown both as kinetic schemes and simulated RP1 population trace. (A)-(C) The timescales of transfer ($k^{-1}$) for the various coarse-grain models are shown in units of ps. Numbers in black (red) represent transfers that move excitation closer to (farther away from) the charge separated states. (A)-(C) represent the protein model, transfer-to-trap limited model, and the trap limited model, respectively. (D) The simulated curves for the RP1 population calculated with ($\tau_{\mathrm{cs}} = 0.61$ ps, $\tau_{\mathrm{rc}} = 160$ ps, and $\tau_{\mathrm{irr}} = 520$ ps) are shown in black (Domain), blue (Protein), purple (transfer-to-trap limited), and red (trap limited). The initial excitation was distributed evenly among Chl-a molecules. . . . . 100

# List of Tables

4.1 Best fit electron transfer timescales for the Domain Model. . . . . 60
4.2 The timescales of transfer for the effective linearization scheme of $C_2S_2M_2$ supercomplex. . . . . 64
4.3 The best fit electron transfer timescales for the Domain, Protein, Transfer-to-Trap Limited, and ERPE models. . . . . 67

C.1 Transition dipole moment (TDM) magnitudes used to determine Coulombic coupling between pigments associated with different proteins[79, 93]. . . . . 92
C.2 Inhomogeneous width of site energies for different molecules in PSII supercomplex [79, 93]. . . . . 92
C.3 The Hamiltonian for Reaction Center pigments labelled by their common names[94, 93, 92]. While there is evidence of a shift in site energies for $Chl_{D1}$ and $P_{D1}$ when the RC is bound in a core complex, we have not included these changes in our Hamiltonian. . . . . 93
C.4 The Hamiltonian for CP43 pigments labelled by their numbering in the 3arc crystal structure [68]. . . . . 94
C.5 The Hamiltonian for CP47 pigments labelled by their numbering in the 3arc crystal structure[93]. . . . . 95
C.6 The spectral density number associated with each pigment-protein complex in PSII. . . . . 96
C.7 The spectral density parameters that describe core pigments in combination with eq. (C.1). . . . . 96
C.8 The spectral density number associated with each pigment-protein complex in PSII. . . . . 97
C.9 The frequencies and Huang-Rhys factors for the under-damped Brownian oscillators describing vibrational modes of Chl A molecules in PSII. . . . . 97

# Acknowledgments

I have been lucky to work with a large group of talented people during my time in the Fleming Group.

Tessa Calhoun was a much needed mentor and friend as I settled into the group and prepared for my qualifying exam.

Jahan Dawlaty was a fabulous colleague and mentor. We shared many long conversations dissecting any scientific problem we wandered across, from Fourier optics to the difference between dephasing and decoherence. He taught me everything I know about aligning optics: from my first hilarious attempt to build a three mirror cavity on our laser table to the long nights diagnosing problems with the 4-f line of the pulse shaper. He also offered the most coherent explanation of mode-locking that I have ever heard. (Please forgive the pun, they made many of our long nights building slightly less painful). I am glad to have had the opportunity to learn from Jahan during my time in the Fleming group. I wish Jahan all the best at his faculty position at the University of Southern California, I am sure he will inspire many undergraduate and graduate students with his passion for science.

Kapil Amarnath gave me the jump start I needed to start thinking about the broader picture of photosynthesis. He was a fun and engaging collaborator as we found our way through the intricate world of energy transfer in photosynthetic light harvesting assemblies. I enjoyed our long talks about the big questions.

My friends David Adamson, Alexis Smith, EmJ and John Sylak-Glassman, Jeff Mugridge, and Alison Nemirow made my time in the East Bay more fun than I would have ever expected. Most importantly of all, my wife, Nina Margaret Meigs Bennett, supported me and helped to preserve my sanity during the long road through graduate school, inspiring me to keep pushing until it was finally done.

**Chapter 2:** Jahan Dawlaty and Vanessa Huxter were a pleasure to work with both while building the pulse-shaper based non-linear spectrometer and during the long, long nights of data collection that consumed our lives. The FMO sample used was donated by Robert Blankenship and his group. Essential expertise in pulse shaping was shared by Duffy Turner and the Keith Nelson's group at MIT.

**Chapter 3:** Jahan Dawlaty was an inspiring sounding board for the crazier idea associated shaped-pulse measurements.

**Chapter 4:** Kapil was a great sounding board and collaborator. I would also like to thank Vladimir Novoderzhkin for helpful discussions about modified Redfield and LHCII calculations; Mino Yang for discussion of linearized kinetic models; Roberta Croce for providing her structure of $C_2S_2M_2$ supercomplex; Rienk van Grondelle for helpful discussions; and Julia Zaks, Eleonora De Re, and Emily Jane Sylak-Glassman for extensive comments on this chapter.

These projects would not have come together into what they are today if not for the suggestions and comments of Prof. Graham Fleming and Prof. K. Birgitta Whaley.

# Chapter 1

# Introduction

## 1.1 Light Harvesting in Photosynthesis

Sunlight provides the energy to drive the production of 105 billion tons of biological carbon annually [32]. Photosynthesis occurs in many diverse light environments ranging from ambient sunlight to extremely low light conditions in microbial mats. One striking example is the GSB1 family of green sulfur bacteria that were discovered 2500 meters beneath the surface of the Pacific, surviving off blackbody radiation of a thermal vent [6]. The ambient light intensity near the thermal vent is 5 orders of magnitude smaller than that experienced by a tree leaf on an overcast day. The basic challenge of photosynthetic light harvesting in diverse light conditions is marrying an appropriately large absorption cross-section with efficient energy transport to achieve high photochemical yields. The design principles that underpin the efficiency of the photosynthetic apparatus in natural organisms, the property that allows them to survive in both high- and low-light environments, can only be understood by describing the structural organization across many length scales.

The basic structure of a molecular light harvesting apparatus is composed of two elements: an antenna to absorb light and a reaction center (RC) where charge separation converts excitation energy into chemical energy. There are multiple different light harvesting apparatuses associated with different photosynthetic organisms each of which captures approximately 90% of absorbed excitations. The light harvesting antenna in photosynthetic organisms vary dramatically depending on the ambient light conditions of their natural environments. For plants, the antenna surrounding the reaction center of photosystem II (PSII) is composed of a collection of different pigment protein complexes (PPC) binding approximately 250 pigments/RC with the majority of pigments bound by light harvesting complex II (LHCII). For green sulphur bacteria, which survive in light conditions much dimmer than plants, a single extremely large antenna protein unit, called the chlorosome, is used to capture light. An individual chlorosome contains more than $2 \times 10^5$ chlorophylls intended to feed excitation energy to a single reaction center [56]. Unlike PSII situation where each chlorophyll is bound in a protein pocket, however, pigments within a chlorosome are self-assembled into long rod-like aggregates that are then wrapped with a protein-dense, lipid membrane. In both cases,

the chlorophyll are bound with well defined locations and orientations, allowing the pigment concentration to far exceed the 0.1 mol·chl/L at which solutions of chlorophyll would undergo rapid concentration quenching [105]. For PSII the concentration of chlorophyll is slightly above 1 mol·chl/L and the tightly packed chlorosome approaches 2 mol·chl/L. Both the high density of pigments in the antenna system and the large number of chlorophyll per reaction center give the antenna a very large absorption cross section and allow each reaction center to harvest energy efficiently even in low light conditions.

At these extremely high concentrations, chlorophylls no longer behave as isolated molecules but experience sufficiently strong coupling that delocalized excitation states (excitons) arise. In the presence of strong coupling, when one chlorophyll transition dipole is excited then the excitation will drive coherent energy transfer to the nearby strongly coupled chlorophyll. In the simplest picture, the steady state solutions are the eigenvectors of the electronic Hamiltonian that describes the absorption frequency of each chlorophyll and the couplings between chlorophylls. In the absence of fluctuations the initial excitation can be described in terms of excitons and no net energy transfer between excitons will occur. As a result of the fluctuating environment, however, the absorption frequencies of each pigment will shift as a function of time which means that the eigenvectors at the initial time will not remain eigenvectors of the system. Thus, vibrational fluctuations of the pigment-protein complexes drive energy transfer between excitonic states of the pigments. In order to understand energy transfer on the length scale of an intact photosynthetic apparatus (on the order of hundreds of nanometers), we must describe the couplings and vibrational fluctuations on the length scale of the individual pigments that compose the total system.

In this thesis we develop experimental and computational tools for revealing the relationship between the high light harvesting efficiency in photosynthetic organisms and the structural organization of pigments in the associated photosynthetic apparatus. These organizations start on the smallest scale of an isolated pigment protein complex, where the precise orientation and distance between chlorophyll determine the energy transfer characteristics that excitons experiences within the bath of vibrational modes. For green sulphur bacteria, the chlorosome packs the largest number of pigments but excitation energy must travel through the Fenna-Matthews-Olson (FMO) complex to reach the RC. The origin of efficient energy transport through FMO is related to the spatial organization of the excitonic structure of FMO. As will be discussed in Chapter 2, "Mapping the spatial overlap of excitons in a photosynthetic complex via coherent nonlinear frequency generation", we will describe an experimental technique for accessing detailed information about the spatial organization of excitons and apply this technique to FMO.

The spectroscopic technique described in Chapter 2 is one of many new non-linear spectroscopies developed in the last decade. Continuing developments in ultra-fast laser sources and pulse compression in parallel with theoretical explorations of material response have driven a rapid expansion of $3^{\text{rd}}$-order non-linear spectroscopic tools that have been applied to understanding photosynthetic PPCs. Recently, a number of groups have explored the possibility of accessing novel information by making higher-order measurements. While higher-order measurements contain additional information, they are both time intensive and

extremely difficult to analyze. One possible approach for extending the insight of 3rd-order techniques is to expand the signal in an over-complete basis that could reveal additional spectroscopic structures that are currently hidden in the line shape of the measured spectra. In Chapter 3, we describe a chirplet transform of 3rd-order spectroscopic signals and show an example of its application to a model system.

On longer length scales, assemblies of pigment-protein complexes allow for efficient energy transfer from distant antenna proteins to the reaction center. Non-linear spectroscopic techniques have revealed a great deal about the dynamics and structure of isolated pigment protein complexes, and yet the structure-function relationship of the biologically active photosynthetic assemblies are much more difficult to probe experimentally. One of the major challenges in addressing non-linear techniques to larger assemblies, e.g. photosystem II supercomplexes, is that they contain a large number of nearly iso-energetic pigments. In the case of PSII supercomplexes, the largest contains 324 pigments all within a nearly structureless two-peak linear absorption spectra. As a result, very little information is available about the time dynamics of populations through non-linear measurements. In general, a cruder time-resolved fluorescence measurement which simply reports on the lifetime of excitation within the supercomplex has been used to characterize the dynamics of light harvesting. In order to extract information about the underlying energy transport of the supercomplex, a structure based model is needed to relate the excitation lifetime to the structural and energetic organization of chlorophyll. In Chapter 4, we discuss the first structure based model of photosystem II supercomplexes and describe the insights we have obtained into the design principles that allow for efficient light harvesting in the large photosynthetic assembly.

## 1.2 Experimental Probes of Excitonic Structure and Energy Transfer in Pigment Protein Complexes

The work in this thesis focuses on measuring and interpreting experimental observables as markers of excitonic structure and energy transfer in pigment protein complexes. The experimental results we will discuss in the following chapters are (a) 3rd-order measurements of isolated PPCs and (b) time-resolved fluorescence decay curves of photosystem II supercomplexes. We will briefly describe the experimental techniques used for both these measurements here, though we leave the detailed discussion of how to simulate such data to the appropriate chapters below.

### 1.2.1 3rd-order Measurements

There are a very large number of experimental measurements that fall into the category 3rd-order, including: pump-probe [99], photon echo [22, 46], transient grating [45, 15] , 2D electronic [41, 102], 2D IR [47, 48], and double-quantum coherence [76, 50, 2]. The shared feature of all 3rd-order techniques is that the emitted signal depends on three interactions with an impinging electric field (usually laser light). This could be, for example, three

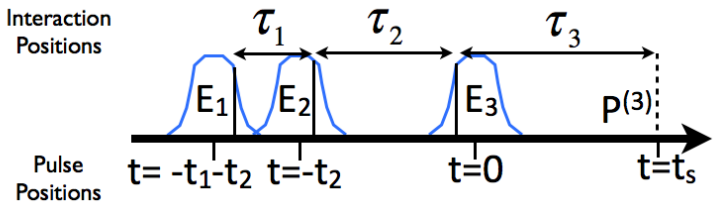



Figure 1.1: Material Response Time Diagram.

interactions with a single laser pulse or it could be one interaction each with three different pulses. For our purposes here we will assume that the electric field can be described as the sum of three independent electric fields, each with their own momentum ($\vec{k}_j$) and central frequency ($\omega_j$). For notational simplicity we define the central time of the third pulse as the zero of time. Further, the central time of $E_2(t)$ is $-t_2$ and $E_1(t)$ is $-(t_1 + t_2)$. The three electric fields ($E_1(t), E_2(t), E_3(t)$) drive a material polarization ($P^{(3)}(t)$) that itself emits an electric field ($E_s(t)$) that we call the signal field. Since, in total, four electric fields are involved in a third-order measurement, these techniques also known as four-wave mixing and the total interaction is described by three time delays, as shown in Figure 1.1. The first two times ($\tau_1$ and $\tau_2$) measure the time between the material experiencing an electric field interaction with the driving fields. The third time delay ($\tau_3$) is the time between the last interaction with the driving field and the measurement of the material polarization.

### Response Function ($S^{(3)}$)

Mathematically, the third order polarization can be written in terms of a response function ($S^{(3)}(\tau_3, \tau_2, \tau_1)$) that describes the complete microscopic details of the molecular response to the incident fields. In general the overall response function of the material which describes all possible third-order measurements, is decomposed into a sum of four different types of response functions ($R_1, R_2, R_3, R_4$) that describe different orders of interaction between the impinging electric fields.

$$S^{(3)}(\tau_3, \tau_2, \tau_1) = \sum_{n=1}^{4} (R_i(\tau_3, \tau_2, \tau_1) - R_i^*(\tau_3, \tau_2, \tau_1)) \tag{1.1}$$

Appendix A describes a method for calculating the third-order response function in terms of parameters that describe the molecular response, such as the coupling between the electronic state of a molecule and the phonon bath surrounding it. For the moment, however, we proceed simply by noting that the material response is a function of three time variables that describe the relative timing of the three interactions with an electric field. As a result of causality, the first interaction must occur no later than the second interaction and the response function is restricted to be non-zero only when $\tau_3, \tau_2, \tau_1 \geq 0$. When the electric fields

have a finite duration the polarization is given by a convolution of the material response function and the impinging electric fields, as shown in eq. (1.2).

$$P^{(3)}(t_s) = \left(\frac{i}{\hbar}\right)^3 \int_0^\infty \mathrm{d}\tau_3 \int_0^\infty \mathrm{d}\tau_2 \int_0^\infty \mathrm{d}\tau_1 S^{(3)}(\tau_3, \tau_2, \tau_1) \times (E(t_s - \tau_3) + c.c.)(E(t_s - \tau_3 - \tau_2) + c.c.)(E(t_s - \tau_3 - \tau_2 - \tau_1) + c.c.) \tag{1.2}$$

where $E(t) = E_1(t) + E_2(t) + E_3(t)$, and $E_j(t) = e^{i\vec{k_j}\cdot\vec{r} - i\omega_j t} A_j(t)$. The often used impulsive approximation assumes that each electric field is sufficiently narrow in time that it can be approximated by a $\delta$-function. The resulting polarization, then, is exactly the material response as a function of the three time delays between impinging pulses: $t_1$, $t_2$, and $t_s$.

**Phase Matching**

Each electric field has its own independently controllable momentum vector that describes the spatial phase evolution ($e^{i\vec{k}\cdot\vec{r}}$) of the light. This information is imprinted on the material as a macroscopic phase grating upon one interaction between the electric field and material transition dipoles. If a second interaction occurred with the same electric field to drive the sample to a population state, their would be no remaining grating. However, a second interaction with another electric field with a different momentum vector will oscillate between being in-phase and out-of-phase with the coherent state initiated by the previous interaction. As a result, a some populations will undergo constructive interference creating a high population and others will undergo destructive interference leaving all sample in the ground state. This resulting population grating can be described by a vector that is the difference between the momentum vectors of the two interactions ($\vec{k}_{\mathrm{pop}} = \vec{k}_1 - \vec{k}_2$). The interaction of the third pulse with the sample can be thought of as the pulse scattering off this populating grating, imparting a momentum to the resulting signal so that the final signal field emerges in the direction $\vec{k}_s = \vec{k}_3 \pm (\vec{k}_1 - \vec{k}_2)$. This is only one kind of third-order signal, more generally a third-order signal is measured anywhere that satisfies $\vec{k}_s = \pm\vec{k}_1 \pm \vec{k}_2 \pm \vec{k}_3$). The standard experimental orientation, and the one used for all measurements and calculations described in this thesis, is called the box-car geometry (shown in Figure 1.2). In the box car geometry, the signal is measured in the $\vec{k}_s = -\vec{k}_1 + \vec{k}_2 + \vec{k}_3$ direction.

Once a signal direction has been selected, this imposes a restriction on the non-zero contributions to the convolution integral shown in eq. (1.2). Using our selected signal direction the overall polarization must develop as $\vec{k}_s = -\vec{k}_1 + \vec{k}_2 + \vec{k}_3$. There are six electric field terms (and their complex conjugates) that contribute to the polarization in this signal direction; they are listed in eq. (1.3). The third-order polarization is the sum of 12 contributions given by the convolution of each term below (and its complex conjugate) with the third-order response function ($S^{(3)}$).

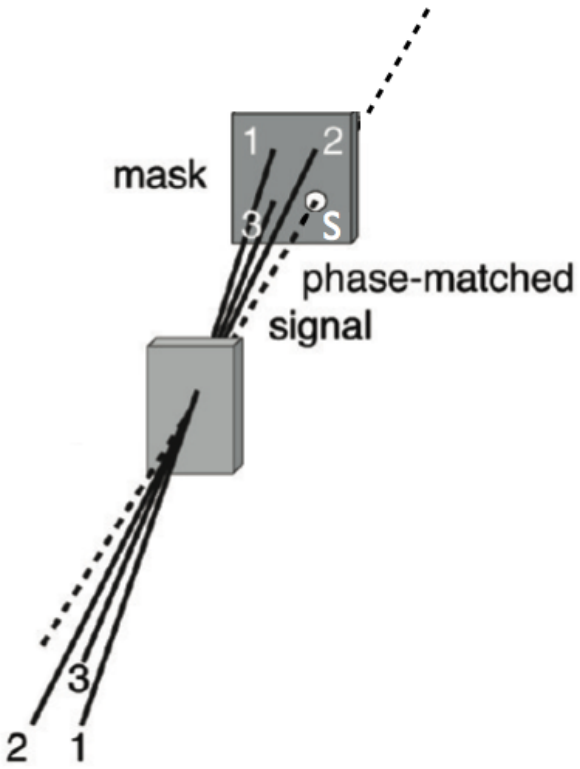



Figure 1.2: The box car measurement geometry. Figure modified from Ref. [34].

$$
\begin{aligned}
&E_3(t_s-\tau_3)E_2(t_s-\tau_3-\tau_2)E_1^*(t_s-\tau_3-\tau_2-\tau_1)\\
&+E_2(t_s-\tau_3)E_3(t_s-\tau_3-\tau_2)E_1^*(t_s-\tau_3-\tau_2-\tau_1)\\
&+E_3(t_s-\tau_3)E_1^*(t_s-\tau_3-\tau_2)E_2(t_s-\tau_3-\tau_2-\tau_1)\\
&+E_2(t_s-\tau_3)E_1^*(t_s-\tau_3-\tau_2)E_3(t_s-\tau_3-\tau_2-\tau_1)\\
&+E_1^*(t_s-\tau_3)E_2(t_s-\tau_3-\tau_2)E_3(t_s-\tau_3-\tau_2-\tau_1)\\
&+E_1^*(t_s-\tau_3)E_3(t_s-\tau_3-\tau_2)E_2(t_s-\tau_3-\tau_2-\tau_1)
\end{aligned}
\tag{1.3}
$$

**Examples of Nonlinear Spectra: Two-Dimensional Spectroscopy**

Using the box-car geometry described above, a number of different 3$^{\text{rd}}$ order signals can be observed. One possible measurement would be a 2D electronic measurement, such as those performed previously on FMO and repeated during the work leading up to the results shown in Chapter 2. 2D spectroscopy is performed by measuring the 3$^{\text{rd}}$-order response of a material as a function of the time delay between the first two pulses and the time delay

between the last interaction and signal emission by the material polarization $(t_1, t_s)$. Taking the Fourier transform of the real value of the emitted electric field results in correlation diagram between the initial excitation frequency $(\omega_1)$ and the final emission frequency $(\omega_3)$. Such spectra can be measured for different waiting times ($t_2$, the time between the second and third interaction) to observe the timescale of fluctuation driving energy transfer from the initial state to the various final states.

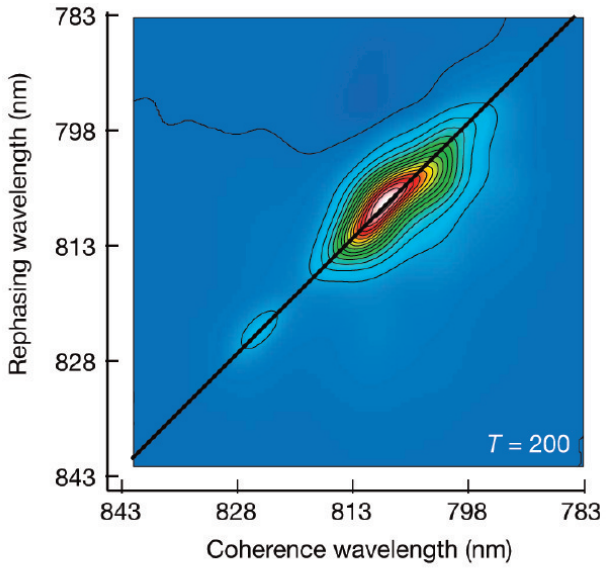



Figure 1.3: Absolute value 2D spectra of the FMO trimer. Image taken from Ref. [31]

.

Figure 1.3 shows a 2D spectra for FMO measured at $t_2 = 200$ fs. The diagonal elongation observed in the 2D spectra reports on the relative magnitudes of the inhomogeneous line shape (corresponding to fluctuations on timescales much longer than $t_2$) versus homogenous line shape (corresponding to fluctuations on timescales shorter than $t_2$). The spectral width along the diagonal corresponds to the combination of the inhomogeneous and homogeneous line shapes, while the anti-diagonal corresponds only to the homogeneous line shape. As such, the diagonal elongation of the 2D spectra of FMO shows that the chlorophylls bound in FMO experience a substantially inhomogeneous environment - related to the very different binding pockets that surrounds the various chlorophyll within the protein. At longer time, off-diagonal cross peaks will appear below the diagonal showing energy transfer from higher energy excited states to lower energy excited states. The clear separation of the homogeneous line shape and the presence of cross-peaks to display energy transfer have made 2D a powerful technique for extracting information in spectrally congested samples.

### 1.2.2 Time-Correlated Single Photon Counting (TCSPC)

Time-correlated single photon counting (TCSPC) measures the fluorescence lifetime distribution of a sample after excitation by a weak laser pulse. Under weak illumination, a small set of pigments in a sample are driven to their excited state. In a photosynthetic assembly with an intact reaction center, excitations are absorbed primarily in the antenna system (composed of light harvesting proteins) and then undergo excitation energy transfer through the surrounding proteins to the reaction center where the excitation energy is converted to chemical energy via charge separation. Alternatively, excitation can also be lost as the result of fluorescence or inter-system crossing. The lifetime distribution of fluorescent photons after excitation is a reporter on the timescales of excitation quenching in the sample either by productive photochemistry at the reaction center or by loss pathways.

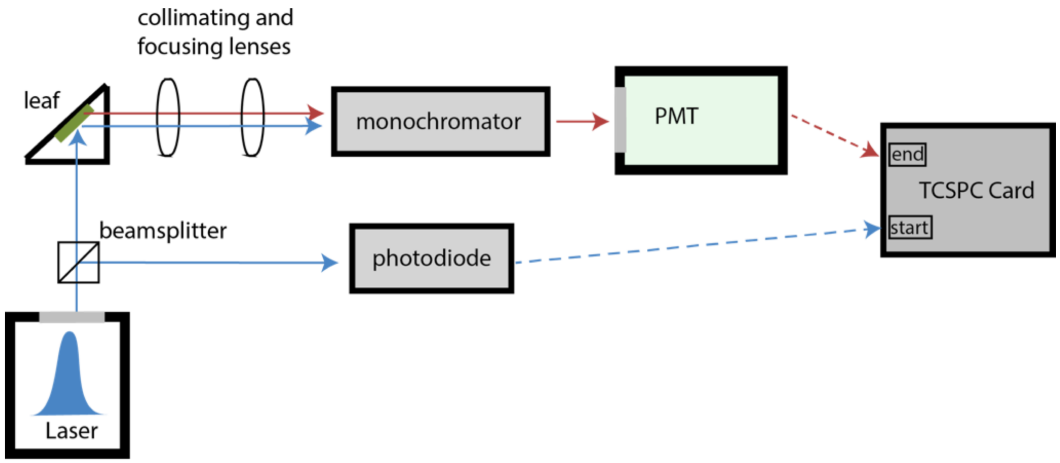



Figure 1.4: The basic experimental setup for TCSPC measurements. Figure courtesy of Emily-Jane Sylak Glassman.

.

Figure 1.4 shows the basic experimental arrangement for a TCSPC measurement. Prior to the sample a small portion of the laser pulse is split off to a photodiode that is used by the TCSPC card to determine the time delay between signal photons and the initial laser pulse. The remainder of the pulse impinges at the sample and then emitted fluorescent photons are measured at a longer wavelength by passing light off the front face of the sample through a monochromator and to a photomultiplier tube. The measured photons are binned as a function of their time delay from the initial pulse giving raw data that we have drawn schematically in Figure 1.5. The raw signal is the convolution of the fluorescence lifetimes of the sample with the instrument response function (IRF). In order to extract the fluorescence lifetime distribution from the experimental signal, we assume a functional form for the fluorescence signal (commonly a sum of three exponentials is used, $F(t) = \sum_{j=1}^{3} A_j e^{-\frac{t}{\tau_j}}$) and perform an iterative forward convolution to achieve a minimum root-mean square error

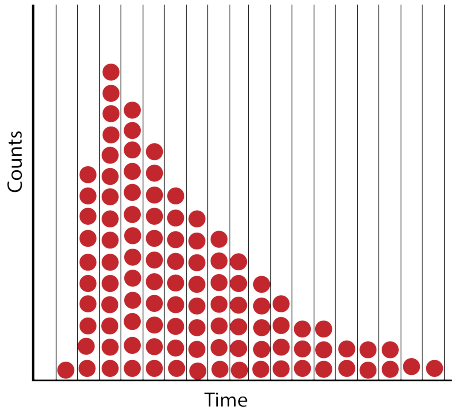



Figure 1.5: The data from TCSPC measurements bins the measured photons by their time delay from the initial excitation pulse. Figure courtesy of Emily-Jane Sylak Glassman.

.

(RMSE).

$$S(t) = \int_{-\infty}^{\infty} IRF(t - t_1)F(t_1)\mathrm{d}t_1 \tag{1.4}$$

The amplitudes ($A_j$) and lifetimes ($\tau_j$) extracted from this process are a useful measure of photochemical behavior in photosynthetic assemblies and have been used to characterize many different properties related to modulating the efficiency of light harvesting in the presence of excess light in order to avoid damage [3].

# Chapter 2

# Mapping the spatial overlap of excitons in a photosynthetic complex via coherent nonlinear frequency generation

Reproduced with permission from AIP. Previously published in *J. Chem. Phys.*, 135: 044201 (2011).

## 2.1 Introduction

Most photosynthetic organisms absorb sunlight in their antenna complexes, which are usually comprised of light absorbing pigment molecules supported inside a protein matrix. Understanding the dynamics of light absorption and energy transfer within these pigment-protein complexes (PPCs) is of practical importance, especially in the context of the need for renewable energy resources and artificial light harvesting. Due to the proximity of the pigments to each other ($\sim$10 Å) within the protein scaffold, optical transitions of nearby pigments become coupled, giving rise to delocalized excitations or excitons [113]. Recent years have witnessed rapid developments in applications of nonlinear optical spectroscopic techniques aimed at elucidating the dynamics of exciton generation and migration within antenna complexes. Two-dimensional (2D) electronic spectroscopy [20, 44, 11] and related nonlinear spectroscopy techniques have been used to demonstrate electronic couplings between pigments[12], to reveal coherent evolution of the exciton states in PPCs [31, 23, 19, 53, 87], and to elucidate the coupling of excitons to vibrations[120, 119].

When more than one photon of light is absorbed in a single PPC, more than one exciton is generated. The proximity of these excitons in the PPC and the anharmonic nature of the pigment electronic transitions[72] give rise to nonlinear interactions between them. Consequently, the two-exciton eigenstates of a complex carry information about the strength

of these interactions. Exciton-exciton interaction has been extensively studied in semiconductors [107, 51] and molecular aggregates [62]. Mukamel and co-workers have developed a series of theoretical methods towards understanding interactions between Frenkel excitons [74, 72] and their spectroscopic signatures[2]. Inclusion of the two-exciton states and exciton-exciton interactions is indispensable for correct interpretation of most third-order nonlinear spectroscopy experiments on PPCs [12, 21, 113], including 2D electronic spectroscopy. Furthermore, two-exciton states may also be relevant for light harvesting function under intense illumination. Nonlinear interactions between excitons are relevant for artificial light harvesting and is discussed in the context of carrier multiplication and the efficiency of solar cells[100].

The Fenna-Matthews-Olson (FMO) complex has been studied as a prototypical photosynthetic PPC [10, 64]. FMO is found in green sulfur bacteria and functions as an energy transfer wire between the bacteria's large antenna system (chlorosome) and the bacterial reaction center. The FMO protein forms a barrel-like structure that encompasses seven bacteriochlorophyl pigments. Recent studies have shown that an eighth pigment is relatively loosely bound to the protein[118, 110]. Here we use the previously established FMO model with seven pigments as we were not sure whether the eighth pigment has been retained in our sample. Ref ([64]) provides a review of the current state of knowledge on the optical properties of FMO.

A recent theoretical study [2] proposed double quantum coherence (DQC) 2D spectroscopy for directly studying exciton-exciton interactions in FMO. Detailed description of DQC spectroscopy and its application to molecular systems can be found in Refs. [73, 50, 49, 76]. In another recent theoretical work [54], Cho and co-workers have proposed a three-wave mixing sum frequency generation (SFG) 2D spectroscopy scheme for accessing the two-exciton states of FMO. The proposed spectroscopy is second order in nature and is normally expected to vanish for isotropic samples. It is argued that due to the chiroptical nature of exciton dipole moments a residual second order signal remains, which carries information about the doubly excited states.

While elegant and insightful, both of the above proposed methods share some restrictions in experimental implementation, such as maintaining strict time-ordering between pulses and accurate phase-retrieval. In particular, with experimentally achievable pulse-widths, a major portion of the signal is collected under pulse overlap, where many unwanted nonlinear optical processes can contaminate the signal.

Here we present a new spectroscopic technique that is suitable for studying the two-exciton states and exciton-exciton interaction in FMO with some experimental conveniences and relative conceptual simplicity. In this scheme the bandwidth of the excitation pulses is chosen to be smaller than the linear absorption band of the photosynthetic complex. We show that when three such excitation pulses interact with the sample in the well-known box geometry, a coherent nonlinear signal with a frequency outside the excitation bandwidth of the laser pulses appears in the phase matched direction. Because this frequency shifted signal appears outside the excitation bandwidth of the laser, its detection is relatively straightforward and unimpeded by lower order scattered light and coherent signals that do

not involve the two-exciton states directly. The signal is directly detected in the frequency domain, hence scanning pulse delays and recovery of the phase of the signal is not necessary. Theoretical studies have suggested the use of pulse shaping at amplitude and phase levels in spectroscopy of molecular aggregates[1]. Our method demonstrates the benefit of simple spectral amplitude shaping in elucidating the nature of excitonic wavefunctions.

In Section 2.2 we review the two-exciton eigenstates of a Frenkel exciton system. We show that when excitons interact, the two-exciton eigenstates acquire the character of many single exciton pairs. In Section 2.3 we describe the concept of nonlinear frequency generation mediated by the two-exciton states and show that it is made possible by the mixed character of the two-exciton states. Our experimental implementation of nonlinear frequency generation in the photosynthetic complex FMO is described in Section 2.4, followed by numerical simulations of the signal in Section 2.5. In section 2.6 we conclude the paper by suggesting potential applications of this method in investigating excited states of complicated Frenkel exciton systems.

## 2.2 Two-exciton states of a photosynthetic complex

In this section we briefly review the formalism for describing two-exciton states in a multi-pigment system. Then we show that the transition dipole matrix elements between the one- and two-exciton states is an indictor of the character of the two-exciton wavefunctions, arising from exciton-exciton interactions.

Details on the formalism to describe multi-pigment Frenkel exciton systems may be found in many references [20, 113]. Since the effect that we are describing in this work is electronic in nature, we disregard exciton relaxation through interactions with its vibrational bath. We consider a coupled, multi-pigment system with $N$ pigments. It is convenient to use the creation (annihilation) operator $a_m^\dagger$ ($a_m$) to indicate excitation (de-excitation) of the $m^{th}$ pigment,

$$a_m^\dagger|0\rangle = |m\rangle \tag{2.1}$$
$$a_m|n\rangle = \delta_{mn}|0\rangle \tag{2.2}$$

where $|0\rangle$ is the ground state of the complex. Similarly, $a_m^\dagger$ ($a_m$) may be used to excite (de-excite) two-pigment excitations,

$$a_m^\dagger|n\rangle = (1-\delta_{mn})|m,n\rangle \tag{2.3}$$
$$a_m|n',n\rangle = \delta_{mn'}|n\rangle + \delta_{mn}|n'\rangle \tag{2.4}$$

where $|m,n\rangle$ represents a state with the $m^{th}$ and $n^{th}$ pigments in the excited state. Equations 2.1-2.4 assume that each pigment is a two-level electronic system, which is often concisely expressed through the commutation relation $[a_m, a_n^\dagger] = \delta_{mn}(1-2a_n^\dagger a_n)$.

The coupled system Hamiltonian can be written as:

$$\hat{H}_S = \sum_m \epsilon_m a_m^\dagger a_m + \sum_{m,n}^{m\neq n} J_{mn} a_m^\dagger a_n \tag{2.5}$$

where $\epsilon_m$ is the energy of the $m^{th}$ pigment and $J_{mn}$ is the coupling energy between the $m^{th}$ and $n^{th}$ pigments. A single exciton state $|e_j\rangle$ of a photosynthetic antenna may be represented as a superposition of single pigment excitations:

$$|e_j\rangle = \sum_{m=1}^{N} U_{jm}^{-1}|m\rangle \tag{2.6}$$

where the coefficients $U_{jm}^{-1}$ are found from diagonalizing the coupled pigment Hamiltonian. Larger $J_{mn}$ and smaller energy mismatch $|\epsilon_m - \epsilon_n|$ between pigments leads to more delocalized exciton states $|e_j\rangle$. Similarly, a two-exciton eigenstate $|f_k\rangle$ can be expressed as a superposition of two-pigment excitations:

$$|f_k\rangle = \sum_{m=1}^{N-1} \sum_{n=m+1}^{N} v_{mn}^{(k)}|m,n\rangle \tag{2.7}$$

where coefficients $v_{mn}^{(k)}$ are found from diagonalizing the coupled pigment Hamiltonian for states $|m,n\rangle$. Analogous to single excitons, Eq. 2.7 shows that a two-exciton eigenstate is delocalized over many two-pigment excitations.

It is important to point out that the two-exciton states $|f_k\rangle$ in general do not correspond to two independent single excitons $|e_p,e_q\rangle$. While $|e_p\rangle$ and $|e_q\rangle$ are eigenstates of the system when individually present in the complex, their simultaneous presence does not, in general, lead to an eigenstate. This arises from the fact that in compact systems such as photosynthetic PPCs, single excitons often overlap in space (i.e. share pigments). The phase and amplitude of excitation on a shared pigment required for exciton $|e_p\rangle$ is often not the same as that required for exciton $|e_q\rangle$. Consequently, the state $|e_p,e_q\rangle$ will break into other exciton pairs and can not be an eigenstate. Such scattering of excitons from each other is due to the couplings between pigments that communicate the presence of one exciton to the other. Mukamel and co-workers have shown that the strength of such interactions depend, at least in part, on the two-level (anharmonic) nature of the the electronic energy levels of each pigment, which can not accommodate independent overlapping excitations[72, 74]. The model can be extended to three- and many-level pigments[72, 74]. However, in most theoretical works on photosynthetic PPCs the pigments are modeled as two-level systems.

Although the states $|e_p,e_q\rangle$ are not eigenstates, they may be used as a basis set for expanding the two-exciton eigenstates $|f_k\rangle$ [2].

$$|f_k\rangle = \sum_{pq}^{q\geq p} w_{pq;k}|e_p,e_q\rangle \tag{2.8}$$

where, in general, many terms in the summation are required to describe $|f_k\rangle$. Equations 2.7 and 2.8 are the pigment basis and one-exciton basis representation of the two-exciton states respectively. In the following we will show that this mixed nature of the two-exciton states is also manifested in the transition dipole moments between the one- and two-excition manifolds.

A spectroscopically relevant quantity is the transition dipole matrix elements between the ground, one- and two-exciton states. The transition dipole operator is a sum over excitations and de-excitations of all pigments weighted by the transition dipole moment of a single pigment $\mathbf{d}_m$

$$\hat{\mu} = \sum_m \mathbf{d}_m (a_m^\dagger + a_m) \tag{2.9}$$

The transition dipole matrix element between the ground and one-exciton states is give by:

$$\langle e_j | \hat{\mu} | 0 \rangle = \sum_{m=1}^{N} U_{jm}^{-1} \mathbf{d}_m \tag{2.10}$$

which is simply a vector sum over all pigment transition dipoles $\mathbf{d}_m$ with the coefficients of the exciton $|e_j\rangle$. The transition dipole matrix elements between the one and two exciton states is given by:

$$\langle f_k | \hat{\mu} | e_j \rangle = \sum_{m=1}^{N-1} \sum_{n=m+1}^{N} v_{mn}^{(k)} (U_{jn}^{-1} \mathbf{d}_m + U_{jm}^{-1} \mathbf{d}_n) \tag{2.11}$$

It is easier to understand this quantity if we consider the action of the creation operator part of $\hat{\mu}$ on the single exciton state $|e_j\rangle$.

$$\sum_m \mathbf{d}_m a_m^\dagger |e_j\rangle = \sum_{m,n}^{m \neq n} \mathbf{d}_m U_{jn}^{-1} |m, n\rangle \tag{2.12}$$

This state is a sum over all pigment-pair excitations where at least one of the pigments in each pair is part of exciton $|e_j\rangle$ weighted by the coefficients $U_{jn}^{-1}$. For example, when $|e_j\rangle$ is mostly localized in one region of a complex, pigment-pair excitations where both pigments are outside that region can not be generated via Eq. 2.12. Otherwise said, the action of the dipole operator on state $|e_j\rangle$ generates all two-exciton states that retain the character of $|e_j\rangle$. Equation 2.11 shows the overlap of a particular two-exciton state $|f_k\rangle$ with this summation. Consequently, the transition dipole matrix element $\langle f_k | \hat{\mu} | e_j \rangle$ is only nonzero if the state $|f_k\rangle$ shares character with exciton $|e_j\rangle$.

To clarify this further, we consider two model systems each composed of three interacting pigments. The two models demonstrate the effects of exciton-exciton overlap on the character of the two-exciton eigenstates. In the first model system, pigments 1 and 2 are strongly bound to each other, but weakly bound to pigment 3. The coefficients $U_{jm}^{-1}$ and $v_{mn}^{(k)}$ for the three one-exciton and three two-exciton eigenstates are shown in Fig.2.1. As expected, the weakly coupled pigment gives rise to a single exciton state $|e_1\rangle$ that is almost completely localized on pigment 3. The other two single excitons $|e_2\rangle$ and $|e_3\rangle$ are mostly delocalized as out-of-phase and in-phase excitations of pigments 1 and 2. The two-exciton eigenstates are shown in Fig. 2.1.c. The highest energy two-exciton state is almost completely localized as an excitation of pigment pair $|1, 2\rangle$, with no character of pigment 3. The expansion of such a state in the exciton basis (Eq. 2.8) will correspond to $|f_1\rangle \approx |e_2 e_3\rangle$. Similarly, the figure

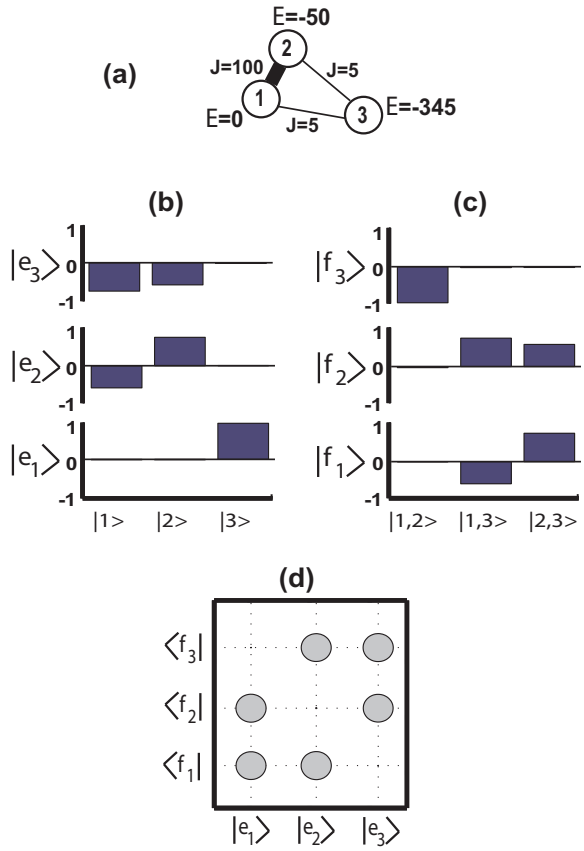



Figure 2.1: (a) A three-pigment model system with one weakly coupled pigment. All energies are in $\text{cm}^{-1}$ and the pigment energies are relative to 12550 $\text{cm}^{-1}$. (b) The one-exciton eigenstates represented in the one-pigment excitation basis (the coefficients $U^{-1}_{jm}$ in Eq. 2.6). $|e_1\rangle$ is primarily localized on pigment 3, while $|e_2\rangle$ and $|e_3\rangle$ represent delocalized out-of-phase and in-phase excitations of pigments 1 and 2. (c) The two-exciton eigenstates represented in the two-pigment excitation basis (the coefficients $v^{(k)}_{mn}$ in Eq. 2.7). The character of two-exciton states in terms of single excitons can be seen in this representation. For example $|f_1\rangle$ shares character with $|e_1\rangle$ and $|e_2\rangle$ (d) The transition dipole moment between the two- and one-exciton states represented as circles. The character of the two-exciton states in terms of single-excitons is much easier to discern (e.g. $|f_1\rangle \approx |e_1 e_2\rangle$). Each two-exciton state is primarily composed of two single excitons.

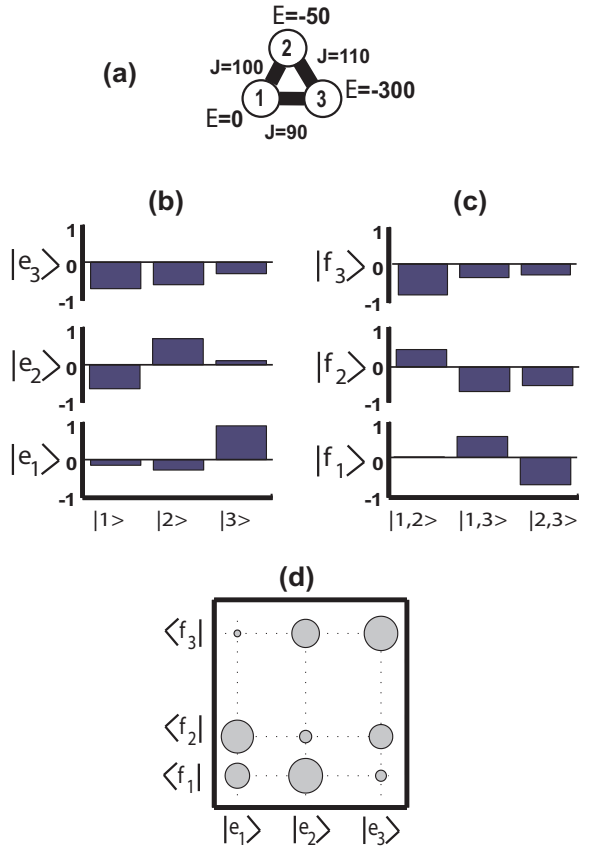



Figure 2.2: (a) A three-pigment model system similar to Fig. 2.1, but with strongly bound pigment 3. (b) The one-exciton (b) and two-exciton wavefunctions (c) are more delocalized over all three pigments. (d) The transition dipole moment between the two- and one-exciton states reveals that each one of the two-exciton exciton states shares character with all three single excitons. Such shared character of the two-exciton states is manifested in nonlinear frequency generation as explained in the text.

suggests that the other two-exciton states are primarily composed of two single excitons, i.e. $|f_1\rangle \approx |e_1e_2\rangle$ and $|f_2\rangle \approx |e_1e_3\rangle$. The character of two-exciton states is more evidently seen in the transition dipole matrix elements $\langle f_k|\hat{\mu}|e_j\rangle$ shown in Fig. 2.1.d. The lowest energy two-exciton eigenstate $|f_1\rangle$ has a transition dipole moment to only exciton $|e_1\rangle$ and $|e_2\rangle$. Consequently, the state $|f_1\rangle$ may be accessed by exciting $|e_1\rangle$ and $|e_2\rangle$, which for the most part, do not interact with each other. Similarly, the state $|f_1\rangle$ may only be de-excited via removal of either exciton $|e_1\rangle$ or $|e_2\rangle$.

In contrast, in model system 2, all three pigments are strongly bound to each other (Fig. 2.2). The one- and two-exciton states are more delocalized over the pigments as seen in Fig. 2.2.b-c. The contributions of single excitons to the two-exciton states is easily evident in the transition dipole matrix elements between them (Fig. 2.2.d). Each one of the two-exciton eigenstates has non-zero transition dipole moment to all of the three single excitons. For example, once eigenstate $|f_1\rangle$ is prepared, de-excitation to any of the three single exciton states is possible. In the next section, we shall explain the relevance of such mixed nature of two-exciton states in optical frequency generation.

We also note that Inverse Participation Ratio (IPR) may be used as a descriptor of the delocalization of the two-exciton states in terms of one-exciton states[2]. An analogous quantity may be constructed by summing the squares of the transition dipole moments between a two-exciton state and all one-exciton states.

We summarize this section by stating that exciton-exciton interactions arise due to the overlap of excitons in the complex. In the presence of overlap, the two-exciton eigenstates can no longer be described as two independent single excitons, and acquire the character of many single exciton pairs. Consequently, the two-exciton eigenstates acquire non-zero transition dipole moments to all single excitons with which they share character.

## 2.3 Nonlinear Frequency Generation Mediated by Two-Exciton States

An experimental signature of the mixed character of the two-exciton states is nonlinear frequency generation in a four-wave-mixing (FWM) experiment. In this section, we first explain the concept for frequency generation in a three-pigment system, using a representative FWM pathway. Then we discuss all FWM pathways that involve the two-exciton states, and show simulation results for frequency generation using realistic laser bandwidths for the two model systems described in the last section.

Figure 2.3 shows the energy level diagrams for a three-pigment system similar to the ones discussed in the earlier section. We consider two optical excitations at frequencies $-\omega_1$ and $-\omega_2$, which will promote the system to the two-exciton state $|f_3\rangle$, resulting into a coherence between states $|f_3\rangle$ and $|g\rangle$. Such coherence between a doubly excited state and the ground state is also known as double quantum coherence. The transition dipole moments of the two-exciton states with the single excitons were discussed in the earlier section. For the case of strongly overlapping excitons, the mixed character of $|f_3\rangle$ allows transitions to all single

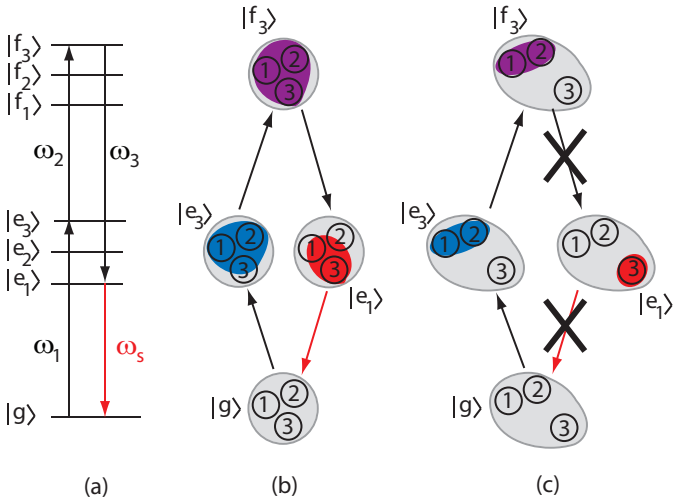



Figure 2.3: (a) A representative FWM pathway that leads to frequency generation in a three-pigment system. Two interactions with optical frequencies $-\omega_1$ and $-\omega_2$ generates a coherence between the two-exciton state $|f_3\rangle$ and the ground state $|g\rangle$. Interaction with frequency $\omega_3$ de-excites the system to a single exciton state $|e_1\rangle$. The FWM signal at frequency $\omega_s = -\omega_1 - \omega_2 + \omega_3$ is red-shifted with respect to all excitation frequencies. The third interaction relies upon the transition dipole matrix element $\langle f_3|\hat{\mu}|e_1\rangle$. (b) In the case of strongly overlapping excitons, the mixed character of $|f_3\rangle$ allows a non-zero transition dipole to state $|e_1\rangle$, making signal generation possible. (c) When exciton $|e_1\rangle$ resides on an isolated pigment such a downward transition is not allowed, as $|f_3\rangle$ does not share any character with $|e_1\rangle$.

exciton sates. Consequently, an optical interaction with frequency $+\omega_3$ can de-excite the complex to the single exciton state $|e_1\rangle$ as shown in Fig. 2.3. When exciton $|e_1\rangle$ is weakly interacting with $|e_2\rangle$ and $|e_3\rangle$, the transition dipole moment $\langle f_3|\hat{\mu}|e_1\rangle$ becomes vanishingly small, and this process becomes impossible (Fig. 2.3.c).

If all three interactions $-\omega_1$ ,$-\omega_2$ and $+\omega_3$ are derived from beams with spatially distinct directions, the frequency shifted signal can be isolated in a phase matched direction, as is commonly done in many other FWM techniques. The frequency shifted signal can be easily isolated from other FWM phase-matched signals using simple optical elements such as an interference filter.

The FWM signal discussed so far corresponds to one possible time ordering of optical interactions. Many orderings of optical interactions are possible when the excitation pulses are overlapping in time. In that case, it becomes necessary to sum over all possible FWM pathways to evaluate the frequency shifted signal. In the next section we discuss evaluation of frequency shifted third order signal from an excitonic system.

### 2.3.1 Simulation of the Signal for Model Systems

We consider an experiment where three pulses with wavevectors $k_1$, $k_2$, and $k_3$ excite the sample. All three pulses are coincident and have identical temporal profiles. The spectral content of the pulses is prepared such that it is resonant only with a portion of the linear absorption (single exciton) spectrum of the sample (Fig. 2.5). For example, the excitation pulses can be spectrally filtered such that frequencies resonant with the lowest energy exciton are removed. The frequency shifted signal is measured in the phase matched direction $-k_s = -k_1+k_2+k_3$ direction and is detected in the cleared region of the spectrum, where no excitation frequencies are present. We employed the nonlinear response function formalism [70, 20] to calculate this frequency shifted signal. Since the pulses are overlapped in time, we consider all nonlinear response terms that contribute to the signal in the phase matched direction:

$$S_R(t_3, t_2, t_1) = R_2 + R_3 - R_{1f}^* \tag{2.13}$$

$$S_{NR}(t_3, t_2, t_1) = R_1 + R_4 - R_{2f}^* \tag{2.14}$$

$$S_{DQC}(t_3, t_2, t_1) = R_{4f} - R_{3f}^* \tag{2.15}$$

where $S_R$, $S_{NR}$ and $S_{DQC}$ refer to the rephasing, non-rephasing, and double quantum coherence components [20, 73], and the subscript $f$ refers to the pathways that involve the two-exciton states of the complex. They are represented by the commonly used Feynman diagrams in Fig. (2.4). These components differ in the time ordering of the interaction of the negative and positive conjugates of the three optical fields with the sample. Since we have considered three coincident pulses, all orderings of interactions of the excitations fields are possible. The nonlinear response in the phase matched direction is a sum over all terms:

$$S(t_1, t_2, t_3) = S_R + S_{NR} + S_{DQC} \tag{2.16}$$

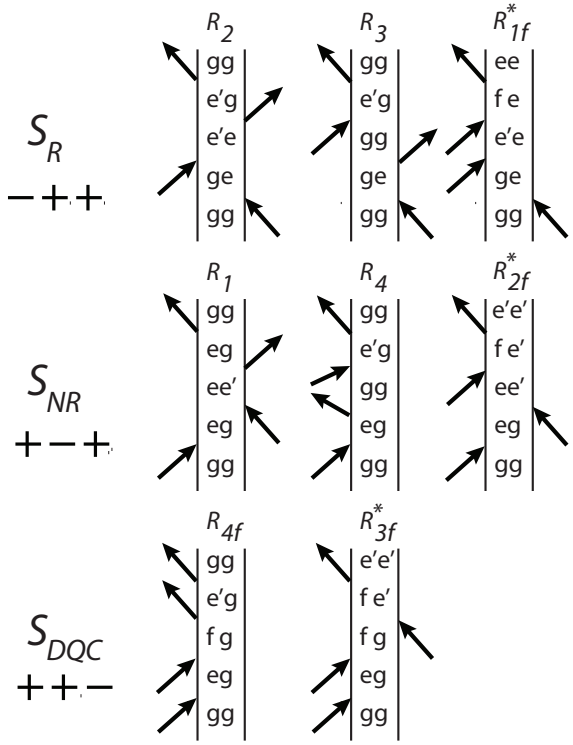



Figure 2.4: Feynman diagrams for the third order signal in the $-k_1 + k_2 + k_3$ direction (Eq. 2.16). The $S_R, S_{NR}$ and $S_{DQC}$ terms differ in the temporal order of the action of the field contributing with negative sign ($-k_1$). Under pulse overlap where all permutations of time ordering is possible, all of the above pathways contribute to the signal. However, only the diagrams that involve the two-exciton states (labeled with subscript $f$) can lead to frequency generation outside the excitation bandwidth.

In Fig. 2.4 the pathways $R_1 - R_4$ have final coherence that results into signal ($eg$ or $e'g$) has the same frequency as one of the excitation fields. Thus a signal frequency shift is not expected to occur due to these pathways. In contrast, in pathways that involve the two-exciton states $R^*_{1f}, R^*_{2f}, R^*_{3f}$ and $R_{4f}$ the emitting coherence frequency need not be the same as the excitation frequencies. For example the process discussed in Fig. 2.3 corresponds to the $R_{4f}$ pathway, where the final coherence $e'g$ has a different frequency from any of the three excitation fields (that are resonant with coherences $ge$, $ef$ and $fe'$).

To isolate the effects of the two-exciton states in frequency generation outside laser excitation spectrum, it is convenient to add all pathways that go through the two-exciton states and define:

$$S_{2exc}(t_3, t_2, t_1) = R_{4f} - R^*_{3f} - R^*_{2f} - R^*_{1f} \tag{2.17}$$

To calculate the nonlinear response functions, we have used the exciton line-shape functions as outlined in Appendix A.

The nonlinear polarization that gives rise to the signal is the result of triple convolution of the appropriate conjugates of the excitation fields with the response functions[70].

$$\begin{aligned} P^{(3)}(t) = \int_0^\infty dt_3 \int_0^\infty dt_2 \int_0^\infty dt_1 \\ S_R(t_3, t_2, t_1) E^*(t - t_3) E^*(t - t_3 - t_2) E(t - t_3 - t_2 - t_1) \\ + S_{NR}(t_3, t_2, t_1) E^*(t - t_3) E(t - t_3 - t_2) E^*(t - t_3 - t_2 - t_1) \\ + S_{DQC}(t_3, t_2, t_1) E(t - t_3) E^*(t - t_3 - t_2) E^*(t - t_3 - t_2 - t_1) \end{aligned} \tag{2.18}$$

where we have considered all three excitation fields to be temporally identical $E(t) = A(t)e^{i\omega_0 t}$, with envelope $A(t)$ and carrier frequency $\omega_0$. The fields are prepared in the frequency domain where the frequencies resonant with the lowest energy exciton are filtered out prior to converting them to the time domain format used in Eq. 2.18. Thus the time-domain shape of the spectrally filtered fields and their corresponding overlaps are taken into account. For calculation of the signal due to $S_{2exc}$, the same expression as Eq. 2.18 was used, where each term of $S_{2exc}$ was convoluted with the appropriate field conjugates. Finally we transform $P^{(3)}(t)$ to the frequency domain and look for components that appear in the cleared region of the spectrum, outside the excitation frequencies.

To demonstrate the role of two-exciton states in creating frequency shifted signal, we consider two model systems similar to the ones discussed in the earlier section. The two model systems differ in the delocalization of the lowest energy exciton (Fig. 2.5). The calculated linear absorption spectra along with the laser excitation spectra used for simulating the third order signal is shown in Fig. 2.5.b. Spectral profile of the laser excitation is Gaussian corresponding to a transform limited 40 fs pulse with a center wavelength of 780 nm. The tail of the spectrum that overlaps with the lowest energy exciton is filtered out, making the pulse resonant only with the two higher energy excitons. Figures 2.5.c , 2.5.f show the nonlinear signal calculated based on pathways that involve the two-states $S_{2exc}$ (Eq. 2.17). The signal appearing outside the laser excitation bandwidth at the frequency of the lowest energy exciton is distinctly larger in the strongly bound pigment. This demonstrates

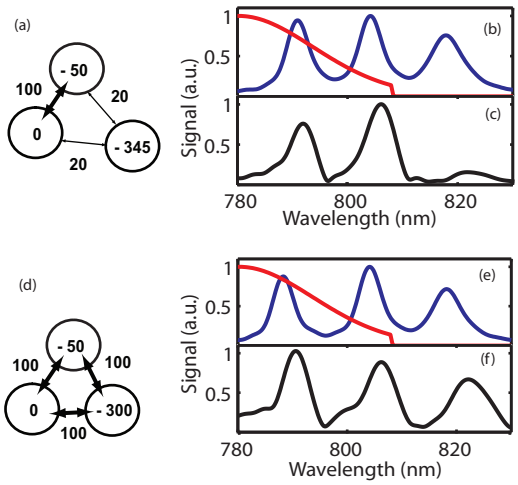



Figure 2.5: Calculation of frequency shifted signal for two models systems. (a) A trimer, with a weakly bound lowest energy pigment. All energies are in $cm^{-1}$, with pigment energies relative to 12550 $cm^{-1}$. (b) The linear absorption spectrum (blue) and the superposed laser excitation pulse (red). (c) The calculated $S_{2exc}$ signal shows small response outside the laser excitation band. (d-e) The same scenario for a trimer with strongly bound pigments. (f) Frequency generation outside the excitation energy band (peak at the frequency of the lowest energy pigment) is seen in the calculated $S_{2exc}$ signal.

that the lowest energy exciton has strong overlap with the other two-excitons, leading to their mutual coupling in the two-exciton state and the emergence of the frequency shifted signal (as discussed earlier). Experimental implementation of a scheme similar to Fig. 2.5 is relatively straight forward. The excitation pulses can be derived from a broadband short pulse laser source, and can be easily filtered out to generate a blank spectral region for detection of frequency shifted signal. Prior to detection, any unwanted signal can be filtered out with a common edge-pass filter. In the next section we demonstrate an experimental implementation of this scheme in a photosynthetic complex.

## 2.4 Experimental Implementation

In this section we describe an experimental implementation of two-exciton mediated frequency shift in the photosynthetic complex FMO. The apparatus used for the experiment is originally designed for inherently phase stabilized multidimensional spectroscopy with pulse shaping capability and is based on the work of Nelson and co-workers in Refs. [37, 117]. Fig. 2.6 shows a diagram of the apparatus. An expanded laser beam was derived from a 250 KHz Ti:sapphire regenerative amplifier (Coherent) and sampled with a mask to generate 4 parallel beams in a square geometry. The beams were directed to a long focal length (75 cm) 1:1 Keplerian telescope. The beams emerging from the telescope were directed to a

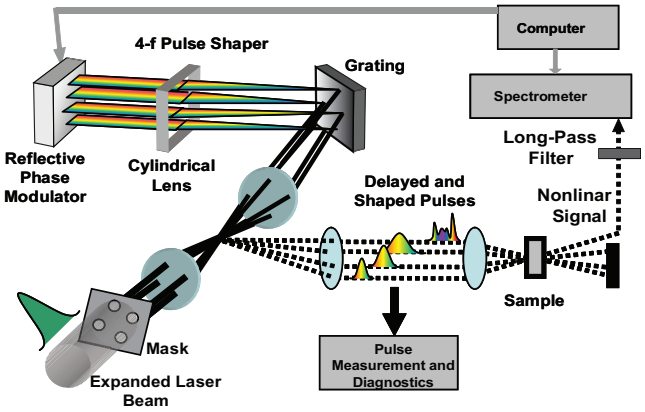



Figure 2.6: Coherent multidimensional spectroscopy apparatus employed in this experiment. Four beams are sampled from an expanded laser beam and guided to a 2D reflective phase modulator based 4-f optical pulse shaper. The diffracted beams (dotted lines) from the 4f pulse shaper is retrieved with a pick-off mirror (not shown), re-collimated and focused on the sample. The pulse shaper is able to modify the spectral phase and amplitude of the pulses, leading to temporal pulse shaping. Pulse delays are also executed by the phase modulator. The nonlinear signal is retrieved in the phase-matched direction and sent to a spectrometer and spectrally dispersed on a CCD camera.

4-f pulse shaper with a 1200 g/mm grating, a 13 cm focal length cylindrical lens, and a two-dimensional liquid crystal on silicon spatial light modulator (SLM) with 792×600 pixels (Hamamatsu). Four horizontal stripes of the SLM are dedicated for phase modulation of the four beams needed for four-wave mixing spectroscopies. In this work the signal was not heterodyne detected, consequently only three of those beams were used. The horizontal axis of the SLM is used for spectral phase application (chirp correction and time delay). To remove the higher order satellite pulses that result from the imperfections of the spectral phase profile, a constant phase grating is simultaneously applied on the vertical axis as explained in Ref. [117]. The first order of diffraction from the vertical grating of the SLM was retrieved by a pick-off mirror, re-collimated, and focused on the sample with a 10 cm lens. In a more recent implementation, the beams are sampled after passing through the 4-f pulse shaper and the telescope for enhanced spectral resolution on the SLM and better decoupling of temporal and spatial phase application. The sample was kept at 77 K in a liquid nitrogen cryostat. The signal in the $-k_1 + k_2 + k_3$ direction was collected and directed to a CCD spectrometer. The apparatus was calibrated for accurate phase application and its ability to perform a number of nonlinear optical experiments such as 2D spectroscopy on gas phase Rb, laser dyes, 2D spectroscopy on FMO, and 2D-DQC spectroscopy has been established.

FMO from *Chlorobium tepidum* was dissolved in a pH 8.0 buffer solution of 20 mM tris-HCl with 0.1% lauryldimethylamine (LDAO) to prevent aggregation. A glass forming

mixture was prepared by mixing the FMO solution with glycerol (30:70 by volume). The sample was frozen to 77K in a 200-micometer optical path-length quartz cell in a liquid nitrogen cryostat. The optical density of sample at 77K was 0.4 at 805 nm.

For frequency selective excitations, a band-pass filter was applied to all beams on the SLM. Such amplitude modulation is possible by blanking the vertical grating on the SLM for the undesired frequencies (as we only collected the first order diffraction of the vertical grating). The pulse energies delivered to the sample were less than 0.5 nJ/pulse. At the detection end, we dispersed the emitted signal in the phase-matched direction $-k_1 + k_2 + k_3$ onto the CCD of a spectrometer and collected the signal outside the bandwidth of the excitation pulses. We used a Raman long-pass filter (Semrock Optics) immediately after the mask in the phase-matched direction $-k_1 + k_2 + k_3$ to remove lower order scatter and to ensure that we only collect the signal outside the excitation pulse bandwidth.

Figure 2.7 shows the spectrum of the three excitation pulses overlaid on the low temperature (77 K) absorption spectrum of FMO. Previous work [21] has shown that the absorption spectrum of FMO is composed of seven inhomogeneously broadened excitonic transitions. The prepared pulse spectrum is expected to have overlap with at least the three highest energy excitons. The detection window covers at least the two lowest energy excitons of the complex and is determined by the cutoff wavelength of our long-pass filter (815 nm). The excitation and detection spectral regions are well separated to avoid spurious light leakage in to the detection range. Upon excitation with three pulses with zero relative delay, the retrieved signal shows two distinct peaks. The relatively narrow peak at 817 nm shows very strong three-beam dependence, revealing its coherent third order nonlinear nature. We ascribe the peak at 826 nm to fluorescence from the lowest energy exciton. It follows the shape of the lowest energy exciton with a slight red-shift of about 1 nm. Our sample only contains the FMO complex and not the reaction center where the excitation is expected to transfer. Consequently, the deposited energy relaxes to the lowest energy exciton and is eventually radiated as incoherent fluorescence in all directions. Part of the fluorescence that coincides with the phase matched direction is detected on the camera and gives rise to the peak at 826 nm. Since fluorescence can be induced by a single beam, the peak at 826 nm is mostly the addition of fluorescence from individual beams, as can be seen in Fig. 2.7. Since the powers of the three beams were not identical the fluorescence induced by them were not the same. Although, important in its own right, we are not concerned with the fluorescence signal in this work. We further studied the dependence of the peak at 817 nm on the pulse delay. Figure 2.8 shows that the signal vanishes within $\pm$100 fs of $k_3$ delay. We note that this time-scale is on the order of the temporal FWHM width of each pulse and that the signal vanishes outside pulse overlap. This suggests that the origins of this signal is due to coherent four-wave mixing and is likely not due to a delayed relaxation induced effect such as bath-induced coherence transfer [84]. In the next section we show calculation of the frequency shifted signal for FMO for our experimental conditions.

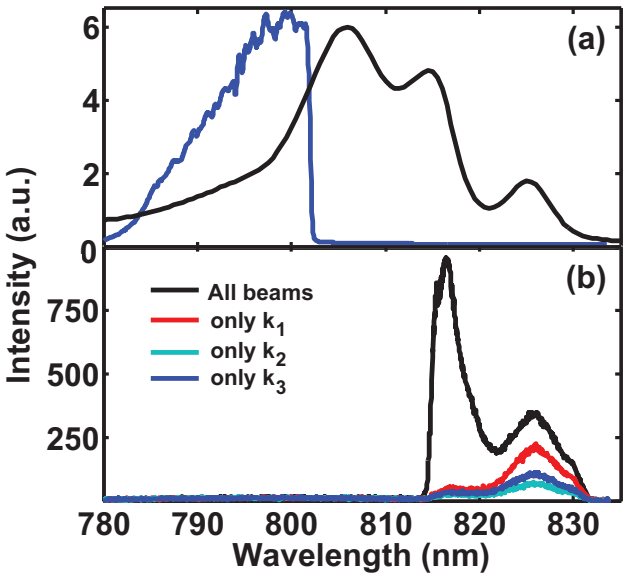



Figure 2.7: (a) The spectrum of excitation pulses (blue) overlaid on the FMO absorption spectrum (black). (b) The signal emerging in the $-k_1 + k_2 + k_3$ direction. A long pass filter with cutoff at 815 nm is used to only collect the frequency shifted signal. The nonlinear nature of the peak at 817 nm is shown by comparing with cases when only one of the beams is exciting the sample. In comparison, incoherent fluorescence from the lowest energy exciton (at 825 nm) shows linear dependence on the presence of the three beams.

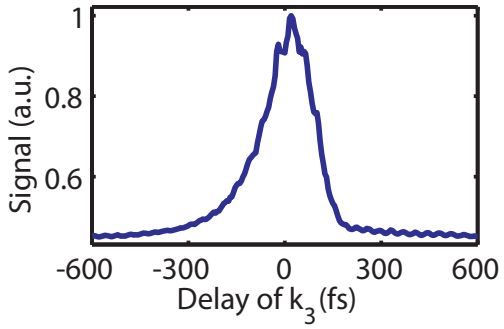



Figure 2.8: The dependence of the intensity of the peak at 817 nm on the relative delay of the pulse with wavevector $+k_3$.

## 2.5 Calculation of Frequency Shifted Nonlinear Signal for FMO

We used the procedure outlined in Section 2.3.1 to calculate the frequency shifted signal. The FMO Hamiltonian [21] and other parameters of the model are discussed in Appendix A. The results of the calculations are presented in Fig. 2.9. The top panel shows the simulated linear absorption spectrum of FMO along with a filtered laser excitation spectrum that was used in the simulations. The laser spectrum used in the simulation is very similar to the one used in the experiment (Fig. 2.7.a) The nonlinear signal from the pathways in $S_{2exc}$ (Eq. 2.17) was calculated as explained in Section 2.3.1 and is shown in Fig. 2.9.b. The signal has a peak outside the laser excitation spectrum near 817 nm, just as observed in the experiment. The frequency shifted peak persists in the total signal calculated based on all components of the response function (Eq. 2.16) as shown in Fig. 2.9.c. Applying a longpass filter to this spectrum (similar to the one used in the experiment) produces the signal shown in Fig. 2.9.d, which has the same main feature as in Fig. 2.7.b. It should be noted again that this model does not include calculation of fluorescence signal. Consequently, unlike the experimental signal, the calculated spectrum does not have a distinct peak at the lowest energy exciton (peak at 826 nm in Fig. 2.7.b).

### 2.5.1 Discussion

In this section we present the transition dipole moments between the one- and two-exiton states of FMO as a descriptor of the character of the two-exciton wavefunctions. Similar to the discussion in Section 2.3.1, we analyze the experimental and simulated signals in light of the mixed character of the two-exciton states.

For simplicity, we consider a FWM pathway similar to the one shown in Fig. 2.3.a. The one-exciton absorption spectrum and the laser excitation spectrum, along with a diagrammatic representation of the one- and two-exciton energy levels are shown in Fig. 2.10. To illustrate frequency generation via this pathway, we indicate the lowest and highest frequency components of the three laser pulses by $\omega_L$ and $\omega_H$ respectively. The figure demonstrates that two upward transitions with frequency $\omega_L$ and one downward transition with frequency $\omega_H$ results in a coherence with frequency $\omega_L+\omega_L-\omega_H < \omega_L$, which will emit a signal outside the excitation bandwidth of the lasers. Such frequency generation is only possible if all of the transitions shown in Fig. 2.10 have non-zero transition matrix elements.

The transition dipole moments between the one- and two-exciton states are shown in Fig. 2.11. Two features of this diagram are noteworthy. Firstly, many two-exciton states have significant transition dipole moments to more than two single exciton states (similar to the case in Fig. 2.2). Based on the discussion in Section 2.2, it follows that the two-exciton eigenstates in FMO can not be simply represented as two independent single excitons. The proximity of pigments in FMO causes the two-exciton eigenstates to acquire the character of many single exciton pairs. Secondly, the overall diagonal spread of the transition dipole moments in Fig. 2.11 shows a trend in the character of the two-exciton states. Overall, the

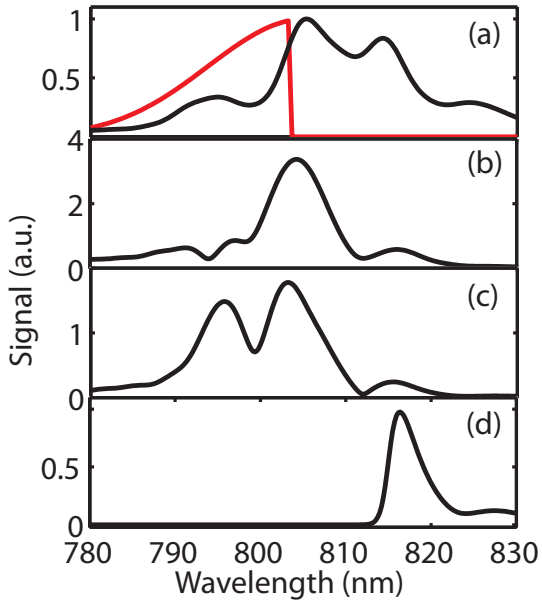



Figure 2.9: Calculation of the frequency shifted nonlinear signal for FMO. (a) Simulated linear absorption spectrum of FMO (black) with the laser excitation spectrum (red) used in the calculations. (b) The nonlinear signal based on the $S_{2exc}$ pathways (Eq. 2.17) shows a frequency shifted peak near 817 nm. (c) The frequency shifted peak persists when the signal based on the sum of all pathways is calculated. (d) The nonlinear signal when a filter similar to the one used in the experiment (Fig. 2.7) is applied.

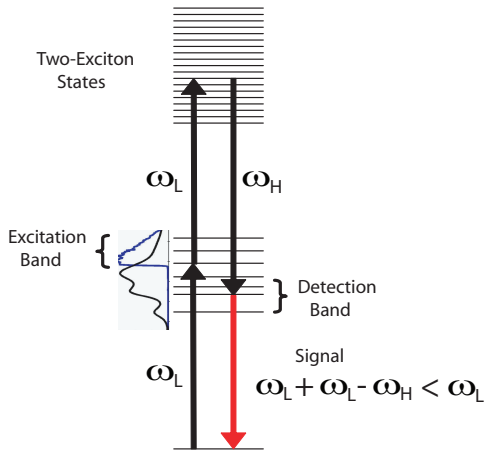



Figure 2.10: Representative four-wave-mixing diagram corresponding to pathway $R_{4f}$ (Fig. 2.4). The lowest and highest frequency components of laser pulses are indicated by $\omega_L$ and $\omega_H$ respectively. The two-exciton states can mediate generation of a signal at a frequency outside the laser bandwidth. The frequency shifted signal depends on transition matrix elements between the one and two-exciton states and carries information about the mixed character of the two-exciton states.

higher (lower) energy two-exciton states have mostly the character of higher (lower) energy single excitons.

The experimental bandwidth of our laser pulses allows excitation of excitons $|e_5\rangle$ - $|e_7\rangle$ after one optical interaction (bracket along the horizontal axis of Fig. 2.11). Due to the inhomogeneous distribution of frequencies, excitation of $|e_4\rangle$ is also possible. After two-optical interactions, a wider range of frequencies within the two-exciton manifold become available. If the lowest (highest) optical frequencies in the laser pulses are $\omega_L$ ($\omega_H$), the range of available two-exciton states is bounded by $2\omega_L$ and $2\omega_H$. The two-exciton states that are accessible after two optical interactions with our experimental laser pulses are indicated along the vertical axis in Fig. 2.11. It is seen from the figure that the two-exciton states $|f_{15}\rangle$ -$|f_{20}\rangle$ that are accessed after two-interactions, also have non-zero transition dipole moments with the single exciton state $|e_2\rangle$, indicated by the red box in Fig. 2.11. This allows for their de-excitation into state $|e_2\rangle$, resulting into emission of signal in the detection band near 817 nm (Fig. 2.7). Emergence of the frequency shifted signal provides a direct read-out of the decomposition of the two-exciton states in terms of single excitons. It should be noted that due to the cut-off frequency of the long-pass filter used in our experiment, the projection of the accessed two-exciton states $|f_{15}\rangle$ -$|f_{20}\rangle$ on to the single excitons $|e_3\rangle$ and $|e_4\rangle$ were not observed. Similarly, the (relatively small) projection of the accessed two-exciton states onto the $|e_1\rangle$ is not resonantly accessible in our experiment, as the frequency $\omega_3$ needed for de-exciting the accessed two-exciton states to $|e_1\rangle$ is not contained in the excitation pulses.

This read-out of the character of the two-exciton states can be used to map the spatial

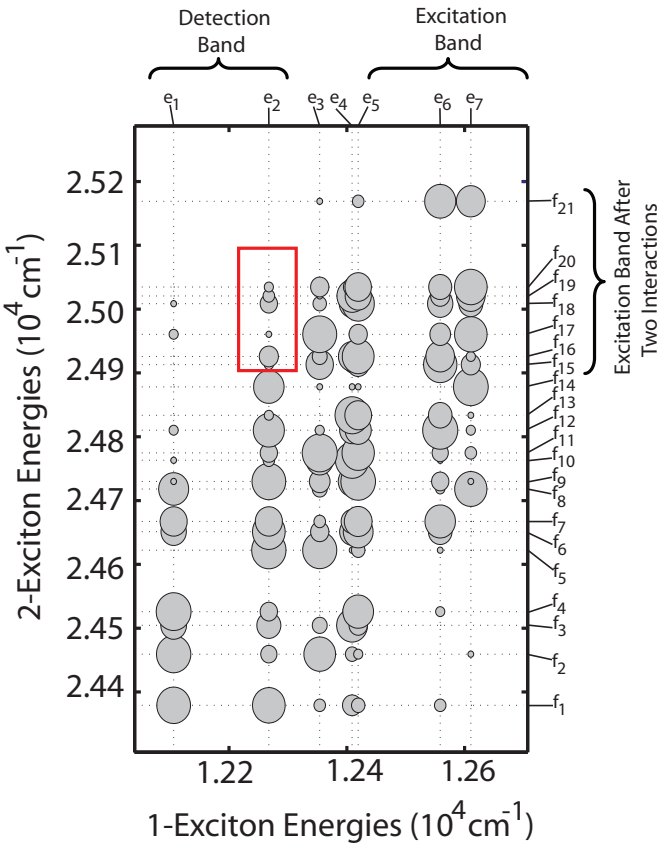



Figure 2.11: The magnitude of the transition dipole matrix elements $\langle f_k|\hat{\mu}|e_j\rangle$ (represented by circles) between the 21 two-exciton states and 7 one-exciton states of FMO. The sizes follow a log scale for better visualization. The mixed nature of the two-exciton states allows them to have non-zero transitions matrix elements with more than two single excitons. One-exciton states $|e_5\rangle$-$|e_7\rangle$ (and $|e_4\rangle$ due to inhomogeneous width) can be excited after a single optical interaction. After two optical interactions, states $|f_{15}\rangle$ -$|f_{20}\rangle$ become accessible. The mixed nature of these states allows for de-excitation into single excitons in the detection band. The transition dipole matrix elements enclosed within the red square are responsible for generation of frequency shifted signal near 817 nm (Fig. 2.7).

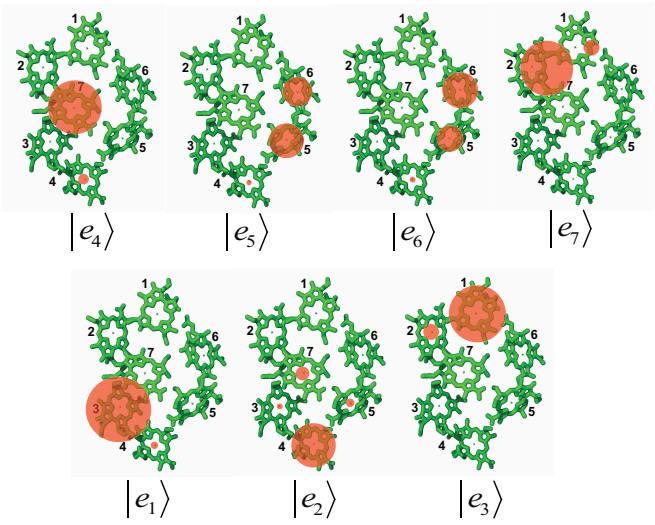



Figure 2.12: The spatial extent of the wavefunctions of the single excitons of FMO. The radius of circles represents the magnitudes of the contribution of each pigment to the exciton (coefficients less that 7% are not shown). Exciton $|e_2\rangle$ in the detection range has overlap with excitons $|e_4\rangle$-$|e_6\rangle$ in the excitation band (sharing pigment 4, 5 and 7), which results into their coupling in the two-exciton states and giving rise to the frequency shifted signal. Exciton $|e_1\rangle$ does not have a significant overlap with the excitons in the excitation band, consequently it can not give rise to frequency shifted signal.

overlap of excitons. The emergence of a strong signal at the frequency of exciton $|e_2\rangle$ indicates a strong overlap between excitons $|e_5\rangle$-$|e_7\rangle$ and $|e_2\rangle$. Similarly, the lack of a strong nonlinear signal at the frequency of exciton $|e_1\rangle$ indicates poor overlap of exciton $|e_1\rangle$ with excitons $|e_4\rangle$-$|e_7\rangle$. Figure 2.12 shows the one-exciton wavefunctions of FMO based on its relatively well-known Hamiltonian (Appendix A). In accordance with this conclusion, the spatial overlap of excitons $|e_2\rangle$ and the much smaller overlap of exciton $|e_1\rangle$ with the excitons in the excitation range ($|e_5\rangle$-$|e_7\rangle$) can be verified in Fig. 2.12.

A particular utility of the experimental ability to measure overlap of excited state wavefunctions is to predict relaxation rates between them. Relaxation of higher energy excitons to lower energy ones is mediated by the overlap of their wavefunctions and accompanied with the physiologically important process of spatial energy transport. The three dimensional arrangement of pigments in PPC's gives rise to spatially delocalized and overlapping excited states to ensure that a sufficient and often redundant number of energy transfer pathways are available for efficient function of the antenna complex. Ref. [21] studied the relation between the relaxation rates and exciton overlaps in FMO. The predicted relaxation rates from excitons $|e_4\rangle$, $|e_5\rangle$, and $|e_6\rangle$ to the relatively strongly overlapping exciton $|e_2\rangle$ are $5.38\times10^{12}\text{s}^{-1}$, $1.60\times10^{12}\text{s}^{-1}$ and $0.22\times10^{12}\text{s}^{-1}$ respectively. In contrast, the predicted relaxation rates from $|e_4\rangle$, $|e_5\rangle$, and $|e_6\rangle$ to the weakly overlapping exciton $|e_1\rangle$ are $0.08\times10^{12}\text{s}^{-1}$,

$0.03\times10^{12}\mathrm{s}^{-1}$ and $0.0\times10^{12}\mathrm{s}^{-1}$ respectively [21]. The large overlap between excitons $|e_4\rangle$-$|e_6\rangle$ in the excitation range and exciton $|e_2\rangle$ in the detection range giving rise to the signal in our work is consistent with this scheme. We anticipate that measuring exciton overlaps using the method described in this work will prove helpful for better understanding energy relaxation pathways and time-scales in excitonic systems.

In a generalized version of this experiment, we propose scanning of narrower bands of the excitation frequencies for a higher resolution investigation of the excited state character. For example, the frequencies of the $+k_2$ and $+k_3$ beams can be held constant and narrow such that they access only one two-exciton state. Then the frequency of the $-k_1$ beam can be scanned over a range of values to determine the decomposition of that state into its one-exciton components.

## 2.6 Conclusion

We have demonstrated a new spectroscopic tool for elucidating the nature of two-exciton states in Frenkel exciton systems such as photosynthetic complexes. The wavefunction of the two-exciton states contain information about the spatial overlap of single excitons and can be helpful in identifying coupling strengths and the structural connectivity of pigments. The method has certain experimental conveniences as it does not rely on time delay and phase retrieval. Furthermore, the signal is frequency shifted with respect to the excitation pulses, which makes it background free and easy to isolate.

We emphasize that the problem of exciton-exciton interaction, even in simple systems is not fully understood and requires further research. Effects such as exciton annihilation, higher order correlations, and contributions of higher excited states of pigments can complicate the exciton-exciton interaction picture. Our method is a new tool for studying such effects.

The mixed nature of the two-exciton states has been also described as entanglement of excitons[71]. Within that point of view, any experimental measurement of the mixed character of the two-exciton states (such as the one described in this work) points to the entanglement of excitons in the complex. We also point out that an intuitive description of exciton interactions can be provided in the time domain using exciton scattering theory and the nonlinear exciton equations[72]. The advantages of such description is more evident in larger systems. Use of exciton scattering theory may prove useful if this method is extended to larger systems such as whole photosystems, chlorosomes, and artificial molecular aggregates.

# Chapter 3

# Shaped pulse third-order spectroscopy: Chirp as a measurement axis

## 3.1 Introduction

As we have seen in the previous chapter, third-order spectroscopy can be a sensitive probe to the underlying electronic structure of a pigment-protein complex, such as FMO. The last decade has seen a rapid expansion in the number and variety of third-order spectroscopic probes of molecular systems [2, 35, 90]. One important reason for this development is that four-wave mixing techniques are very sensitive to many different aspects of the material response - the short time dynamics associated with solvation, timescale of interconversion between different environments, the presence of electron correlation in molecular excited states, and many others. There remain, however, a number of important features of vibrational and electronic dynamics in both homogeneous and heterogeneous environments that have not yet been determined experimentally: the possibility of non-gaussian behavior (or solvation outside of the linear response regime), the role of intra-molecular vibrations versus protein fluctuations, the possibility of correlated fluctuations in protein environments, and the role of an individual chromophore's doubly excited states in a pigment-protein complex (which was ignored in our treatment of FMO in Chapter 2).

Third-order spectroscopic techniques, such as 2D electronic spectroscopy, are limited by the same sensitivity that has made them so powerful. When a molecular system undergoes three interactions with impinging electric fields, a large number of possible pathways (such as those described by the Feynman diagrams in the top row of Figure 2.4) give rise to the measured signal. Each of these pathways, in turn, are sensitive to a large set of molecular parameters: e.g. the excited state absorption frequency of the 1-exciton and 2-exciton states, the spectral density associated with the ground, 1-exciton, and 2-exciton states, the model of the 2-exciton state overlaps, etc. Some of these parameters have been established for many systems of interest by combined experimental (both linear and non-linear spectroscopies) and

modeling work. Many of these parameters, however, remain poorly described. When trying to glean insight into the details of solvation - e.g. the possibility of correlated fluctuations in the environments of adjacent pigments - these undetermined parameters become confounding factors in our analysis. In general, two approaches have been used to bypass this challenge: first, some groups have worked quantify the analysis of 2D line shapes by connecting detailed microscopic simulations with experimental results. Other groups have worked on developing higher order spectra that have been shown to correspond to higher-order correlation functions. In this chapter, we address the possibility of extracting additional information from third-order spectroscopies using shaped excitation pulses to transfer information out of the line shape of the spectra and into more specific features such as peak position.

The difficulty with non-linear line shapes can be thought about as a problem with the choice of basis set of the representation. In the current basis - a collection of time-domain delta-functions - information is overlapped in a fashion that is difficult to de-convolute. The selection of a new basis could selectively extract and display particular contributions to the line shape function. The chirplet transform, shown in eq. (3.1), was originally developed for radar signals [60] and was then applied to analyzing audio signals (in one interesting case, counting wolf populations by their howls [30]).

$$C(t,\beta,\sigma,\omega) = \int_{-\infty}^{\infty} e^{-\frac{(t-\tau)^2}{2\sigma^2}} e^{-i\omega(t-\tau)-i\beta(t-\tau)^2} F(\tau)\mathrm{d}\tau \tag{3.1}$$

In this basis, an oscillating field is decomposed into a collection of gaussian packets (called chirplets) described by four parameters: the central frequency ($\omega$), the central time ($t$), the temporal width of the packet ($\sigma$), the time-domain chirp of the oscillation ($\beta$). In this way, the one-dimensional signal described by $F(\tau)$ is transformed into a four-dimensional object $C(t,\beta,\sigma,\omega)$. The basis-set composed of these chirplets is non-orthogonal, so this is never a minimal representation of data, but the addition of extra parameters allows for information to spread into higher dimensions which can reveal structures that are hidden when the information is compressed.

In this chapter we will describe a chirplet transform and its relationship to third-order non-linear spectroscopic measurements. We will then apply this treatment to a physical system where the influence of chirp is expected and describe the resulting chirplet spectroscopic measures. Appendix B discusses the general properties of a chirplet transform of a one-dimensional signal and foreshadows the properties of a chirplet spectrum as described in this chapter.

## 3.2 Chirplet Transform of Third-Order Polarization

If measurements are performed in the impulsive limit (the pulse width in time is much narrower than the timescale of dynamics being probed) then the measured third-order polarization is exactly the response function, as described in 1.2.1. As such, no additional information about the third-order response function is available by performing the measurement with shaped excitation pulses. However, a chirplet transform can still be used as a

signal analysis technique for the measured response function. On the other hand, if the electric fields impinging on the sample are not delta-functions then shaped excitation pulses can directly perform the chirplet transform "in sample." There is no mathematical distinction between these two cases, so in the remainder of this paper we will not distinguish between them.

### 3.2.1 Third-order polarization in the semi-impulsive limit

In the impulsive limit, third-order polarization is a function of three time variables $(t_s, t_2, t_1)$ that describe the relative timings of the incident electric fields and the emitted signal field. If we loosen this assumption to have only two of the three pulses approximated by $\delta$-functions, then the remaining pulse can be described by a chirplet. We name this case the semi-impulsive limit. At present, we select the second and third pulses ($E_2$ and $E_3$) to be impulsive, and describe the first pulse as $E_1(t) = e^{-\frac{t^2}{2\sigma_1^2}} e^{-i\omega_1 t - i\beta_1 t^2}$. Since the second and third pulse are $\delta$-functions, we assume that $E_3$ alway interacts after $E_2$, allowing us to discard three of the six electric field polarizations described in eq. (1.3). The resulting non-linear polarization is written out in eq. (3.2), and then simplified to eq. (3.3). The three remaining terms correspond to $E_1$ interacting first, $E_1$ interacting second, and $E_1$ interacting third. In all three cases the second pulse interacts before the third pulse. Using phase matching considerations we can label the component of the response functions that will give non-zero signal in each of these three cases as the rephasing $(-k_1 + k_2 + k_3)$, non-rephasing $(+k_2 - k_1 + k_3)$, and the double quantum coherence $(+k_2 + k_3 - k_1)$.

$$\begin{aligned} P^{(3)}(t_s) = \int_0^\infty \mathrm{d}\tau_3 \int_0^\infty \mathrm{d}\tau_2 \int_0^\infty \mathrm{d}\tau_1 S^{(3)}(\tau_3, \tau_2, \tau_1) \\ [\delta(t_s - \tau_3)\delta(t_s - \tau_3 + t_2 - \tau_2) E_1^*(t_s - \tau_3 + t_2 - \tau_2 + t_1 - \tau_1) \\ \delta(t_s - \tau_3) E_1^*(t_s - \tau_3 + t_2 - \tau_2 + t_1)\delta(t_s - \tau_3 + t_2 - \tau_2 - \tau_1) \\ E_1^*(t_s - \tau_3 + t_2 + t_1)\delta(t_s - \tau_3 - \tau_2)\delta(t_s - \tau_3 + t_2 - \tau_2 - \tau_1)] \end{aligned} \tag{3.2}$$

$$\begin{aligned} P^{(3)}(t_s) = &\int_0^\infty \mathrm{d}\tau_1 S_{PE}^{(3)}(t_s, t_2, \tau_1) E_1^*(t_1 - \tau_1) \\ &+ \int_0^\infty \mathrm{d}\tau_1 S_{TG}^{(3)}(t_s, t_2 - \tau_1, \tau_1) E_1^*(t_1 + \tau_1) \\ &+ \int_0^\infty \mathrm{d}\tau_2 S_{DQC}^{(3)}(t_s - \tau_2, \tau_2, t_2) E_1^*(t_2 + t_1 + \tau_2) \end{aligned} \tag{3.3}$$

The three contributions to eq. (3.3) are related to three different groups of response functions. These are the same grouping we have defined previously in eqs. (2.13)-(2.15) with a slight relabeling. The first term is $S_{PE}^{(3)} = S_R^{(3)}$ and corresponds to the portion of the

response function measured in a photon echo (PE) experiment where rephasing results in peak signal for $t_1 > 0$. The second term is $S^{(3)}_{TG} = S^{(3)}_{NR}$ is the non-rephasing component of the response function measured in transient grating (TG) experiments. The final term $S^{(3)}_{DQC}$ is given the same label here and in eq. (2.15) and is related to Feynman pathways, as shown in Figure 2.4, where the first two excitations produce a coherence between the ground and second-excited states of the sample.

### 3.2.2 Chirplet Spectra

Each term in eq. (3.3) corresponds to performing a chirplet transform on the corresponding portion of the response function and the total signal arises from the sum of each of the three contributions. In the remainder of this section, we will ignore the third term which corresponds to $E_1$ interacting last by setting out time delay between $E_2$ and $E_3$ to be greater than the pulse width of $E_1$. As a result, our chirplet spectra are described by a sum of two terms, as shown in eq. (3.4).

$$\begin{aligned} C(t_s, t_1, \beta_1) = &\int_0^\infty \mathrm{d}\tau_1 S^{(3)}_{PE}(t_s, t_2, \tau_1) e^{-\frac{(t_1-\tau_1)^2}{2\sigma_1^2}} e^{-i\omega_1(t_1-\tau_1)-i\beta_1(t_1-\tau_1)^2} \\ &+ \int_0^\infty \mathrm{d}\tau_1 S^{(3)}_{TG}(t_s, t_2-\tau_1, \tau_1) e^{-\frac{(t_1+\tau_1)^2}{2\sigma_1^2}} e^{-i\omega_1(t_1+\tau_1)-i\beta_1(t_1+\tau_1)^2} \end{aligned} \tag{3.4}$$

If the response function has no variation as a function of $t_2$ (at least over timespans comparable to the dephasing time associated with $\tau_1$) then this equation can be simplified to a single term by a substitution that treats the photon echo and transient grating signals as the positive and negative tau portions, respectively, of the same overall response. This changes the single-sided integral to a single double-sided integral ($\int_0^\infty \rightarrow \int_{-\infty}^\infty$), as shown in eq. (3.5), over an effective response function ($\tilde{S}^{(3)}(t_s, \tau_1)$) which has a photon echo response for $\tau_1 > 0$ and a transient grating response for $\tau_1 < 0$.

$$C(t_s, t_1, \beta_1) = \int_{-\infty}^\infty \mathrm{d}\tau_1 \tilde{S}^{(3)}(t_s, t_2, \tau_1) e^{-\frac{(t_1-\tau_1)^2}{2\sigma_1^2}} e^{-i\omega_1(t_1-\tau_1)-i\beta_1(t_1-\tau_1)^2} \tag{3.5}$$

Finally, to simplify both the analysis and the measurement, we calculate only the integrated intensity of the emitted signal field, allowing the chirplet spectra to depend upon only two variables, the time delay between the first pulse and the second pulse ($t_1$) and the chirp of the first pulse ($\beta_1$). Experimentally such a spectra would be measured using the same apparatus as a photon echo experiment, shown in Figure 3.1, except that the first pulse would be variably chirped in the time domain. The $\beta_1 = 0$ axis of the chirplet spectra, described in eq. 3.6, corresponds to the standard photon echo measurement which is then extended along a second axis as a function of the chirp in $E_1$.

$$C_s(t_1,\beta_1)=\int_0^\infty \mathrm{d}t_s\Big|\int_{-\infty}^{\infty}\mathrm{d}\tau_1\tilde{S}^{(3)}(t_s,\tau_1)e^{-\frac{(t_1-\tau_1)^2}{2\sigma_1^2}}e^{-i\omega_1(t_1-\tau_1)-i\beta_1(t_1-\tau_1)^2}\Big|^2 \tag{3.6}$$

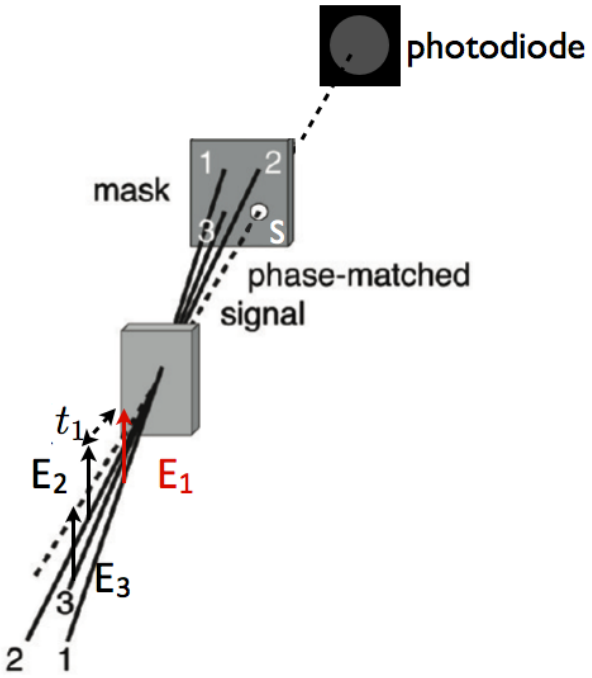



Figure 3.1: The measurement geometry for a chirplet spectra. Figure modified from Ref. [34].

## 3.3 Application of Chirplet Transform

### 3.3.1 Simple Third-Order Response Function

In the following two sections, in which we derive a simple expression for a third-order response function with frequency dependent dephasing, we draw heavily on the work of Andrei Tokmakoff[109].

In Appendix A, we describe the terms that contribute to the third-order response function appropriate for electronic spectroscopy of an N-pigment system. The resulting terms contain a complex combination of line-shape functions that depend on all three time-delays between electric field interactions. In this section we will consider the response functions appropriate for a single pigment which we treat as a two-level system with a lineshape function that is linear with respect to time, as shown in eq. (3.7). The lineshape function, $g(t)$, describes the frequency-domain broadening of an electronic transition as a result of vibrational motion

and is described mathematically in eqs. (A.9)-(A.12).

$$g(t) = \Gamma t \tag{3.7}$$

The linear form of the lineshape function correspond to the traditional Bloch picture. We note here that this approximation is unphysical since it corresponds to having $\delta$-correlated fluctuations of the energy gap between the ground and excited state of the pigment. To see this, we use eq. (A.9) and substitute the correlation function of the ground and excited state energy gap with a $\delta$-function ($C_{eg}(t) = \Gamma\delta(t)$). This substitution is shown and solved in eq. (3.8).

$$g(t) = \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 C_{ba}(\tau_2) = \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \Gamma \cdot \delta(\tau_2) = \Gamma t \tag{3.8}$$

Despite the unphysical origin, this approximation is mathematically convenient, so we maintain it for the moment in order to establish a simple form for the response function that we can use for calculating a chirplet spectrum. Using this approximation, the four non-linear line shape functions ($F_j(\tau_3, \tau_2, \tau_1)$) can be simplified by restricting the indices to $a, c = e$ and $b = g$ and then realizing that $g_{e,e}(t) = g(t)$ and $g_{e,g}(t) = g_{g,g}(t) = 0$. The ground-ground lineshape function is zero by definition since there is no energy gap between the ground-state and itself, while the cross-correlation term is assumed to be zero in the usual uncorrelated frequency fluctuation approximation [21, 20]. The resulting response functions are expressed in simplified form in eqs. (3.9) - (3.12), where we have removed any terms that modify the overall amplitude but do not influence the time-dynamics.

$$R_1(\tau_3, \tau_2, \tau_1) = e^{-i\omega_{eg}\tau_3 - i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.9}$$

$$R_2(\tau_3, \tau_2, \tau_1) = e^{-i\omega_{eg}\tau_3 + i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.10}$$

$$R_3(\tau_3, \tau_2, \tau_1) = e^{-i\omega_{eg}\tau_3 + i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.11}$$

$$R_4(\tau_3, \tau_2, \tau_1) = e^{-i\omega_{eg}\tau_3 - i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.12}$$

We construct two response functions ($R_{\mathrm{PE}}, R_{\mathrm{TG}}$) that differ in the assumed order of interaction for the first pulse ($E_1$), but contribute to the same spatial direction $-k_1+k_2+k_3$. The photon echo (PE) response function, shown in eq. (3.13), assumes that pulse one interacts first and the resulting macroscopic coherence is capable of rephasing as a result of the oppositely signed evolution of the coherence during the first ($\tau_1$) and third ($\tau_3$) response times.

$$R_{\mathrm{PE}} = R_{-k_1+k_2+k_3} = R_2 + R_3 = e^{-i\omega_{eg}\tau_3 + i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.13}$$

The transient grating (TG) response function, shown in eq. (3.14), assumes that pulse one interacts second (after pulse two) which results in the coherence being unable to rephase

during the third time period.

$$R_{\mathrm{TG}} = R_{+k2-k_1+k_3} = R_1 + R_4 = e^{-i\omega_{eg}\tau_3 - i\omega_{eg}\tau_1} e^{-\Gamma(\tau_1+\tau_3)} \tag{3.14}$$

**Inhomogeneous Broadening**

Long timescale fluctuations of the environment that shift the excitation energy of the pigment result in inhomogeneous contributions to the non-linear spectra. These long-timescale contributions can be incorporated by convoluting the response functions with a gaussian distribution of transition frequencies ($\omega_{eg}$), as shown in eqs. (3.15) and (3.16).

$$R^{(3)}_{\mathrm{PE}}(\tau_3,\tau_2,\tau_1) = \int_{-\infty}^{\infty} \mathrm{d}\omega_{eg} \sqrt{\frac{1}{2\pi\sigma_0^2}} e^{-\frac{(\omega_{eg}-\omega_0)^2}{2\sigma_0^2}} R_{\mathrm{PE}}(\tau_3,\tau_2,\tau_1) \tag{3.15}$$

$$R^{(3)}_{\mathrm{TG}}(\tau_3,\tau_2,\tau_1) = \int_{-\infty}^{\infty} \mathrm{d}\omega_{eg} \sqrt{\frac{1}{2\pi\sigma_0^2}} e^{-\frac{(\omega_{eg}-\omega_0)^2}{2\sigma_0^2}} R_{\mathrm{TG}}(\tau_3,\tau_2,\tau_1) \tag{3.16}$$

The above integration can be performed analytically to find that the total response functions are now given by eqs. (3.17) and (3.18).

$$R^{(3)}_{\mathrm{PE}}(\tau_3,\tau_2,\tau_1) = e^{-\Gamma(\tau_1+\tau_3)} e^{-(\tau_3-\tau_1)^2\frac{\sigma_0^2}{2}} e^{-i\omega_0(\tau_3-\tau_1)} \tag{3.17}$$

$$R^{(3)}_{\mathrm{TG}}(\tau_3,\tau_2,\tau_1) = e^{-\Gamma(\tau_1+\tau_3)} e^{-(\tau_3+\tau_1)^2\frac{\sigma_0^2}{2}} e^{-i\omega_0(\tau_3+\tau_1)} \tag{3.18}$$

We have now calculated the simplest response functions for an ensemble of isolated pigment in an inhomogeneous environment. Using the same approach we will now describe a model system with dynamics that cause chirp in the response function.

### 3.3.2 Frequency Dependent Dephasing

The goal of this section is to demonstrate a simple physical mechanism that results in the response functions calculated in eqs. (3.17) and (3.18) having a frequency-evolution as a function of time (chirp). In this case, we use the presence of frequency dependent dephasing which results in different members of the inhomogeneous ensemble having different rates of dephasing. Figure 3.2 shows a schematic response function with two contributions: a red contribution that dephases rapidly and a blue contribution that dephases slowly. As a function of time, the central frequency of the response function develops from an average between the red and blue contribution to being just the central frequency of the blue contribution. A more realistic situation has a continuous distribution of central frequencies each with their own dephasing rate and a resulting frequency chirp within the ensemble averaged response function. One simple example of this behavior was measured by Stiopkin and coworkers [106] in J-aggregates where the exciton delocalization varied across the band, giving a linear change in dephasing rates as a function of frequency.

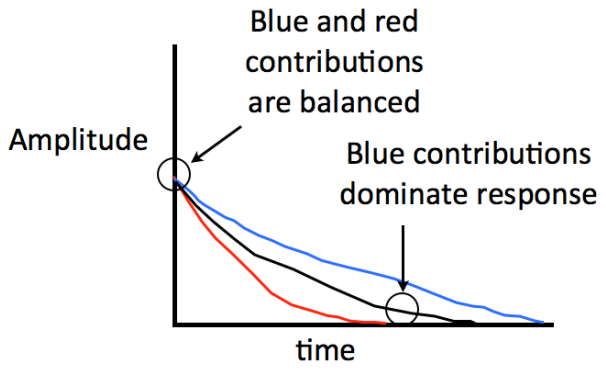



Figure 3.2: A schematic response function (whose envelope is shown in black) is the average of two contributions: a slowly dephasing high-frequency function and a quickly dephasing low-frequency response. The result is an overall chirp within the response function as the average response goes from a central frequency that is the average of the two central frequencies of the red and blue terms in the short time to a central frequency that is the blue terms central frequency in long time.

### Linear Dependence

We will start by assuming a linear form for the frequency dependent dephasing, $\Gamma = \Gamma_0 - \alpha(\omega_{eg} - \omega_0)$, as has been proposed previously in Ref.[109]. The photon echo and transient grating response functions can, using this approximation, be described by performing the integration shown in eqs. (3.15) and (3.16) again after substituting in the linear frequency dependent $\Gamma$. The resulting equations, (3.19) and (3.20), do show the presence of chirp, but the photon echo component now contains an unphysical growth term ($\exp[\frac{\sigma_0^2\alpha^2}{2}(\tau_3+\tau_1)^2]$) that will result in a non-convergent response function along the $\tau_1 = \tau_3$ line. We should note that while the transient grating component has the same growth term in the response function, it is exactly balanced by a decay term that, for $\alpha < 1$, ensures no divergent behavior.

$$R^{(3)}_{\mathrm{PE}}(\tau_3,\tau_2,\tau_1) = e^{-\Gamma_0(\tau_1+\tau_3)}e^{-\frac{\sigma_0^2}{2}(\tau_3-\tau_1)^2}e^{\frac{\sigma_0^2\alpha^2}{2}(\tau_3+\tau_1)^2}e^{-i\alpha\sigma_0^2(\tau_3^2-\tau_1^2)}e^{-i\omega_0(\tau_3-\tau_1)} \tag{3.19}$$

$$R^{(3)}_{\mathrm{TG}}(\tau_3,\tau_2,\tau_1) = e^{-\Gamma_0(\tau_1+\tau_3)}e^{-\frac{\sigma_0^2}{2}(\tau_3+\tau_1)^2}e^{\frac{\sigma_0^2\alpha^2}{2}(\tau_3+\tau_1)^2}e^{-i\alpha\sigma_0^2(\tau_3+\tau_1)^2}e^{-i\omega_0(\tau_3+\tau_1)} \tag{3.20}$$

While the exact origin of the divergent behavior in the photon echo response function is unclear, the linear approximation to dephasing causes an un-physical behavior in the far-wings of the inhomogeneous distribution. Since we describe the inhomogeneous distribution with a pure gaussian, there is no region of the frequency domain that has truly zero contribution to the response function. We have, simultaneously assumed that dephasing is linear in

frequency and therefore continues to change even in the limit $\omega_{eg} \to \pm\infty$. As a result, a very small component of signal originates from the very far wings of the response function where the dephasing is actually negative ($\omega_{eg} > \omega_0 + \frac{\Gamma_0}{\alpha}$). This component is damped by its small contribution to the inhomogeneous distribution, but since its signal undergoes exponential growth in the long-time limit it will contribute an infinite quantity of signal.

### Error Function Dependence

In order to both construct a more physical response function and test our hypothesis for the origin of the unbounded signal we have constructed a new description of the frequency dependent dephasing, $\Gamma = \Gamma_0 + \tilde{\alpha} \cdot \text{erf}(\frac{\omega_{eg}-\omega_0}{\gamma\sigma_0})$, which can be bounded to exclude the possibility of negative dephasing. Both the inhomogeneous distribution and the dephasing as a function of the central frequency $\omega_{eg}$ are plotted in Figure 3.3, for $\gamma = 2$. In this treatment, the dephasing rate is bounded above and below by $\Gamma_0 \pm \alpha$, which can be selected to exclude negative dephasing by requiring $\alpha \leq \Gamma_0$. While we did not find an analytic form for this integration, we have shown numerically that both the photon echo and transient grating response functions are convergent for any value of $\tilde{\alpha} \leq \Gamma_0$. Interestingly, the photon echo response function become divergent again for all values of $\tilde{\alpha} > \Gamma_0$, when some fraction of the ensemble experiences negative dephasing. This supports our hypothesis above that the presence of negative dephasing results in a divergent signal. In the following treatment of chirplet spectra we will use our bounded form for $\Gamma$ to simulate the expected 2D and chirplet spectra.

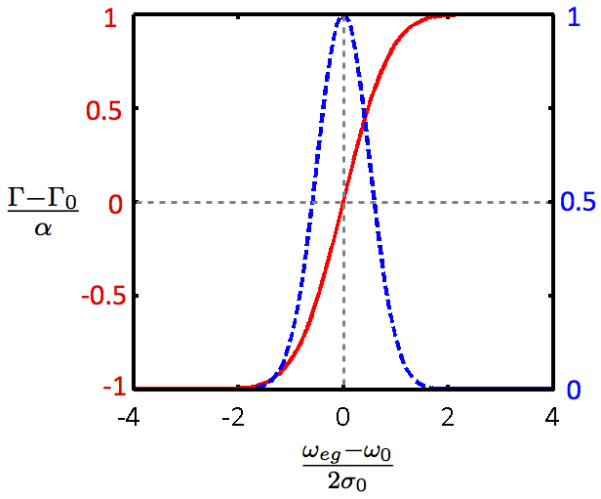



Figure 3.3: $\Gamma = \Gamma_0 + \tilde{\alpha} \cdot \text{erf}(\frac{\omega_{eg}-\omega_0}{\gamma\sigma_0})$ is plotted for $\gamma = 2$ in red (left axis) to show that the value is now bounded so that values of the system parameters can be chosen to ensure no portion of the inhomogeneous ensemble experiences negative dephasing. The amplitude of the inhomogeneous distribution is plotted in blue (right axis).

### 3.3.3 Simulated Spectra

A traditional method for characterizing the response function of the system is the rephasing portion of a 2D spectra that, in the impulsive limit, is the Fourier transform of the photon-echo component of the response function. This can be interpreted as the correlation function between the initial excitation frequency ($\omega_{\tau_1}$) and the emission frequency ($\omega_{\tau_3}$), as discussed in Chapter 1. In the presence of frequency dependent dephasing the homogeneous line width, described by the anti-diagonal width of the 2D Spectra, changes as a function of central frequency. In a 2D spectrum, this results in twisted 2D line shape where the anti-diagonal line width is narrower on one side of the spectra and wider on the other side. We have simulated 2D spectra in the presence of frequency dependent dephasing with two values of $\gamma$) (top row: $\gamma = 2$, bottom row: $\gamma = 1$) and three values of $\alpha$ (left: $\alpha = 0$, middle: $\alpha = 0.6\Gamma_0$, right: $\alpha = \Gamma_0$ column). The largest chirp in the response function occurs for the largest value of $\alpha = \Gamma_0$ and the smallest value of $\gamma = 1$. The resulting 2D spectra are plotted in Figure 3.4(a-c,e,f), where we should note that when $\alpha = 0$ the same response function is achieved for any value of $\gamma$. To the eye, all five 2D spectra look nearly identical. We can still extract the line shape distortion by considering anti diagonal slices, such as shown by the arrows in Figure 3.4(f) both with and without frequency dependent dephasing. Figure 3.4(d) shows anti-diagonal cuts through the 2D spectra from the opposite sides of the spectra, normalized to a maximum value of 1. The blue, dashed lines plot the cuts taken from a spectra with $\alpha = 0$ and, as expected, have the same width. The red, solid lines represent the cuts taken from the 2D spectra with $\alpha = \Gamma_0$ and $\gamma = 1$, which should has the largest distortion of any of the 2D spectra. While there is some variation of the homogeneous line width, as Figure 3.4 shows, it would be difficult to quantify in the presence of even a small amount of experimental noise, since it depends entirely on a careful measurement of the wings of the 2D spectra.

We can compare the slight distortion in the 2D spectra to the influence of the frequency dependence dephasing on a chirplet spectra. In a chirplet spectra, the information is not accessed via the homogeneous line width. Instead, the presence of chirp in the response function can be assessed directly. When the response function is chirped, a driving electric field (with a given temporal width) that is transform-limited will not be in resonance with the response function for the entirety of the pulse duration. This is because the frequency of the response function will walk off in time. A chirped pulse, however, can match the frequency shift and be resonant for a longer period - allowing the pulse to drive more excitation and, thereby, increase the final emitted signal intensity. The presence of chirp in a response function can, therefore, be measured as a shift in the peak signal position of a chirplet spectra.

We have simulated chirplet spectra in the presence of the same frequency dependent dephasing described in the previous paragraph. The simulated chirplet spectra are shown in Figure 3.5 (a-c,e,f), ordered to match the 2D spectra shown in Figure 3.4. Along the $\beta_1 = 0$ line, the signal corresponds to a photon-echo measurement with a peak shift to non-zero $t_1$ as the result of the large inhomogeneous distribution of excitation frequencies. For a unchirped response function (Figure 3.5(a), $\alpha = 0$), the maximum signal occurs for $\beta_1 = 0$.

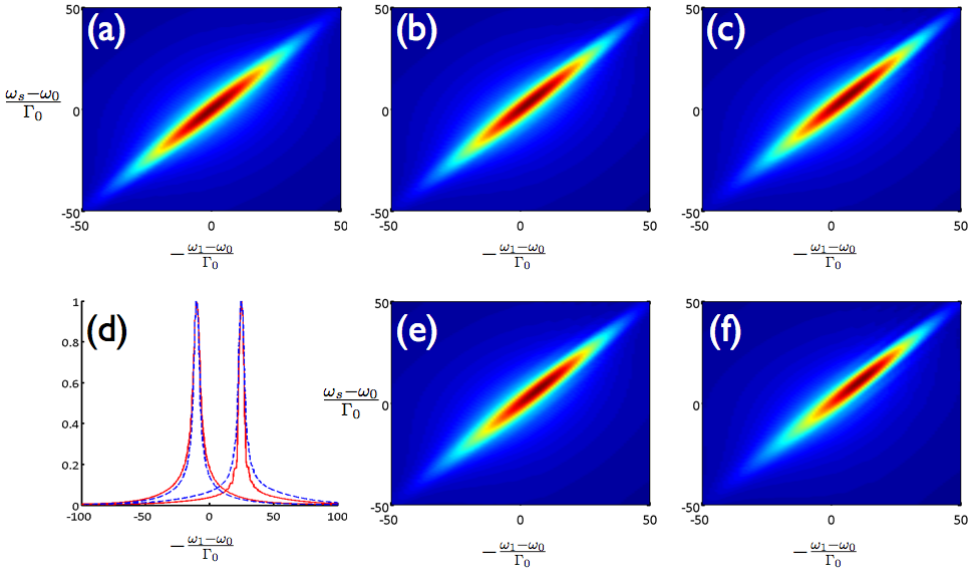



Figure 3.4: The absolute value 2D rephasing spectra (a-c,e,f) of the response functions characterizing a material with frequency dependent dephasing. The figures in the top-row were calculated with $\gamma = 2$, the figures in the bottom row were calculated with $\gamma = 1$. The chirplet spectra shown in the left, middle, and right columns were calculated with $\alpha = 0$, $0.6\Gamma_0$, and $\Gamma_0$, respectively. All 2D spectra are nearly identical by eye. Figure (d) shows anti-diagonal slices taken from the left and right edges of a 2D spectra with $\alpha = 0$ (blue) and $\alpha = \Gamma_0$ (red) which quantify the small change in line shape that results from frequency dependent dephasing.

The decay of signal is symmetric with respect to $\pm\beta_1$ and in both directions corresponds to a decreased peak-shift with increasing $\beta_1$. In the presence of a chirped response function (Figure 3.5(c-b,e-f), $\alpha > 0$) the chirplet spectra peak shifts away from $\beta = 0$, as expected given the chirplet transform properties described in Appendix B. Unlike the case of un-chirped response function, when $\alpha > 0$ the chirplet spectra is not symmetric with respect to $\pm\beta_1$. This is in agreement with the pattern described for 1D chirplet transforms in Figure B.4. One interesting asymmetry is that the signal peak appears for $\beta_1 > 0$ while signal appears for longer delay times ($t_1$) for $\beta_1 < 0$. As has been described previously, the chirplet spectra will show a peak when the chirp of the electric field ($\beta_1$) is equal to the chirp of the response function. When the electric field has the opposite sign chirp compared to the response function, however, Figure 3.6 shows that the pulse will perform more favorable excitation at long time-delay. When the response function is un-chirped then the maximum signal at long time delay and the peak signal both occur for $\beta_1 = 0$. For a chirped response function, however, these two features split and the maximum signal at long time-delay occurs for the opposite sign chirp compared to the peak signal.

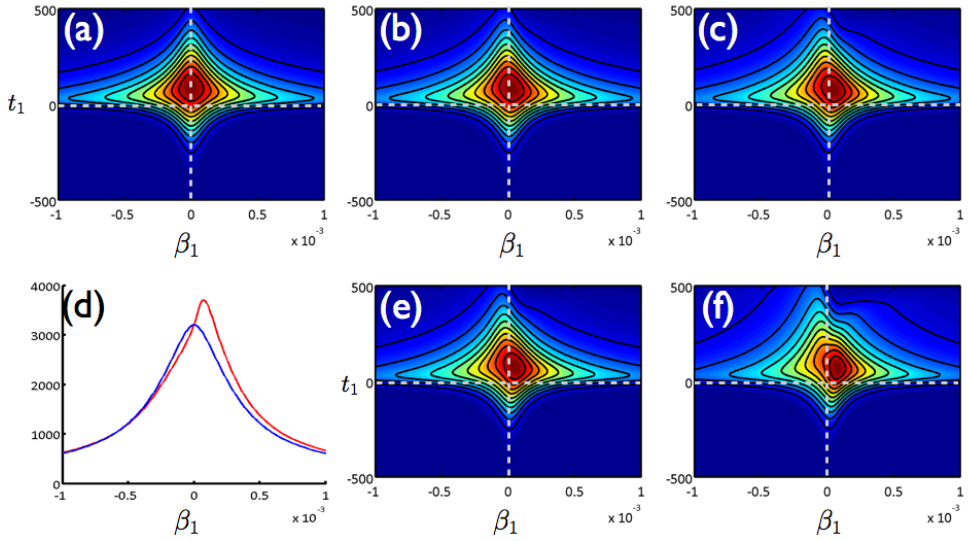



Figure 3.5: The chirplet spectra (a-c,e,f) of the response functions characterizing a material with frequency dependent dephasing. The figures in the top-row were calculated with $\gamma = 2$, the figures in the bottom row were calculated with $\gamma = 1$. The chirplet spectra shown in the left, middle, and right columns were calculated with $\alpha = 0$, $0.6\Gamma_0$, and $\Gamma_0$, respectively. The chirplet spectra are matched to the 2D spectra shown in Figure 3.4. Figure (d) shows a slice along the $\beta$-axis at the peak value along $t_1$ of two chirplet spectra, $\alpha = 0$ (blue) and $\alpha = \Gamma_0$ (red).

The shift in the peak structure of the chirplet spectra is much more obvious than the subtle line shape distortion in the wings of the 2D spectra. To further clarify this point, Figure 3.5(d) shows a slice of the chirplet spectra along the chirp axis, for the value of $t_1$ at the peak for $\alpha = 0$ (blue) and $\alpha = \Gamma_0$ (red). The peak shift noted previously becomes readily apparent in this plot where the red curve clearly peaks to the right of the blue (unchirped response function) curve. As noted previously, the line shape is symmetric with respect to $\pm\beta_1$ when $\alpha = 0$, but is not for $\alpha > 0$. Comparing Figure 3.4(d) and 3.5(d), it is clear that the shift of the peak away from zero chirp is much more easily extracted from the chirplet spectrum than the subtle charge in line shape associated with the 2D spectrum. The advantage of the chirplet transform is displayed here by removing the desired information about frequency dependent dephasing from the wings of the line shape in a 2D spectra and displaying that information in a peak shift of a chirplet spectra.

## 3.4 Conclusion

We have demonstrated one example of a change in basis allowing for clear visualization of information that is otherwise contained in the line shape of 2D spectroscopy. The chirp is

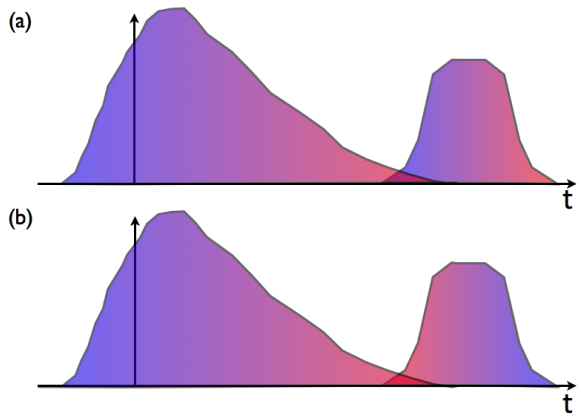



Figure 3.6: The overlap between a chirped pulse and a chirped response function (either of (a) the same sign or (b) different sign) is displayed for a long time delay. (a) When the two functions share the same sign of chirp, then at long delay time the pulses blue edge overlaps with the response functions red edge. This provides minimal off-resonance excitation. (b) When the two functions have the opposite sign of chirp, then at long delay time the pulses red edge overlaps with the response functions red edge which provides favorable on-resonance excitation.

small in all of the response functions considered here, hence the small shift of peak positions along the chirp axis of the chirplet spectra. This single observation is the most significant draw back to the current technique: unless a substantial chirp is introduced into the response function by a physical process of interest, the implementation of such a novel experimental (or signal analysis) strategy will not yield new information. Nevertheless this approach of transforming the basis set of the measurement by analogy to the experimental electric fields is promising since it allows for the simultaneous development of novel signal analysis and measurement techniques that can be directly related to different physical signals. With the continued development and increased adoption of pulse shaping technologies, the ability to implement "in system" signal analysis that increase sensitivity to particular aspects of the response function is a promising direction for future work.

# Chapter 4

# A structure-based model of energy transfer reveals the principles of excitation migration and trapping in Photosystem II supercomplexes

Reproduced with permission from JACS, in press. Unpublished work copyright 2013 American Chemical Society.

## 4.1 Introduction

Photosynthesis in plants initiates with the absorption of light by pigments bound to the antenna, or light harvesting, proteins of photosystem II (PSII). The resulting excitation energy is transferred over 10-100 nanometer distances to the reaction center (RC), where it is converted to chemical energy via charge separation. In plants, the proteins associated with PSII are located in the grana membrane [28], which is densely packed with photosystem II and the major light harvesting complexes (LHCII). PSII reversibly binds with LHCII to form PSII supercomplexes (Figure 4.1a). PSII supercomplexes and LHCIIs form a large, variably fluid array of pigment-protein complexes [52]. This array forms an energy transfer network that harvests light with 90% efficiency [17] and can respond to changes in incident sunlight intensity and wavelength [98, 43, 66]. The largest, structurally well-characterized part of the grana membrane that can harvest light is the PSII supercomplex. A detailed model of this complex would provide access to the blueprint by which PSII efficiently and robustly collects light for photosynthesis.

A complete model of light harvesting in PSII supercomplexes requires both characterizing the kinetics of light harvesting and understanding how the spatial organization of chlorophylls generates these kinetics. When considering complicated kinetics, isolating the rate limiting step or steps is a convenient simplification for understanding the overall process. Two rate limiting steps have been hypothesized for light harvesting in the smallest supercomplex,

called the core complex, which has a dimer structure and contains the core antenna proteins (CP43 and CP47) and the reaction center. The excitation radical-pair equilibrium (ERPE) model suggests that excitation will equilibrate throughout the collection of antenna pigments on a timescale much faster than excitation is converted into chemical energy[65]. The ERPE model is, therefore, also known as a trap-limited model since trapping at the RC (via charge separation) is the rate limiting step. The second hypothesis, supported by a computational model of energy transfer in the core complex[93], assumes a slow rate of transfer between the core antenna proteins and the RC. This hypothesis is known as the transfer-to-trap limited model, where transfer into the RC is the rate limiting step. Fluorescence decay curves measured on the core complex, however, can be fit with either a trap limited or transfer-to-trap limited kinetic model [114]. One experimental study has measured a series of fluorescence decay curves of various PSII supercomplexes[17]. These measurements, when fit to a slightly more detailed kinetic model that incorporated spatial structure, suggested that light harvesting is trap-limited. Taken together, the experimental measurements and models of PSII form a conflicted picture of energy transfer and trapping.

Even without knowledge of the detailed kinetics, there have been suggestions about the organizational principle that governs the structure of PSII supercomplexes and results in efficient light harvesting. One possibility is the so-called "energy funnel," in which higher energy chlorophylls are further from the RC[10]. Energy transfer down the gradient would result in unidirectional transport of excitation towards the RC. Previous computational work on PSII core complexes and LHCII trimers, however, suggests that no energy gradient exists in a supercomplex [79, 93]. In the absence of an energy gradient, efficient light harvesting in PSII supercomplexes has been attributed to a small collection of fast energy transfer steps forming efficient pathways of energy transfer that extend from the periphery of the supercomplex to the reaction center[14, 25, 102]. We call this assumption the pathways model of energy transfer since it corresponds to assuming that light harvesting in a PSII supercomplex is reducible to a series of energy transfer steps that make-up a minimal time pathway from an excited chlorophyll to the RC [89]. Fine-grain population dynamics are not readily accessible to experimental probes, so addressing the question of organizational principles in PSII supercomplexes requires a structure based model of energy transfer.

With the recent experimental work that resolved the orientations and approximate distances between proteins in the largest PSII supercomplex[18] ($C_2S_2M_2$, shown in Figure 4.1) it has become feasible to build a structure based model of energy transport for PSII supercomplexes. The number of chlorophyll in a supercomplex (>300 in $C_2S_2M_2$) still makes constructing such a model a daunting computational challenge. We can decrease the computational cost of calculating the rate matrices for a supercomplex and simplify the kinetics of energy transfer by coarse graining over groups of chlorophyll. Such a coarse graining is complicated by the large variations in electronic couplings between pigments. In the presence of small electronic couplings, excitations are localized to single pigments and energy transfer between pigments is caused by electronic coupling. In the presence of strong electronic coupling, excitations are delocalized over multiple pigments and energy transfer between delocalized states is driven by the phonon modes of the bath. While modified Redfield the-

ory interpolates energy transfer rates between both limits of electronic coupling, it does not account for the dynamic localization caused by phonon modes that are strongly coupled to the chlorophyll excited states. Previous work on two components of PSII supercomplexes, LHCII trimers[79, 81, 67, 95, 69] and the core complex [93], have used a combined generalized Förster/modified Redfield approach to incorporate an ad hoc correction for the influence of dynamic localization. In the combined approach, chlorophylls are grouped into clusters (called domains) based on the strength of their electronic coupling. In the work of Raszewski and Renger[93] on the core complex, chlorophyll pairs with electronic coupling greater than a threshold value of 34 $cm^{-1}$ were included in the same domain. The treatment of LHCII by Novoderezhkin and coworkers [79] used a threshold value of 15 $cm^{-1}$. No single threshold value will correctly balance electronic delocalization against dynamic localization for all the intermediate couplings in PSII supercomplexes. We, therefore, introduce a method for domain assignment where in cases of moderate-to-strong coupling the inclusion of a chlorophyll into a domain (or not) is selected on the basis of whether it increases (decreases) the separation of timescales between intra- and inter-domain transfer rates. The resulting domains can be used to coarse-grain the energy transfer dynamics by assuming infinitely fast intra-domain thermalization.

Using the combined generalized Förster and modified Redfield approach with our separation of timescales metric for domain assignments, we constructed a rate matrix that describes population transport through the chlorophylls bound by PSII supercomplexes. We determined empirical rates of charge separation by fitting fluorescence lifetime measurements of PSII supercomplexes. In order to answer questions about the overall kinetics of PSII supercomplexes, we used an effective linearization technique, first introduced by Yang and Fleming [125], that decomposed the average timescale of light harvesting into contributions from possible rate limiting steps. Next, we used calculations of photochemical yields in the presence or absence of different domains to assess the appropriateness of the pathways model for explaining the origin of efficient light harvesting by PSII supercomplexes. In both of these discussions we extrapolate our results to comment on energy transfer and trapping in the membrane. Finally, we provide an outlook on developing future models of energy transfer in PSII supercomplexes.

## 4.2 Methods

This section is divided into three portions. The first explains our approach to treating excitation energy transfer within the pigments of the PSII supercomplex. In this work, we exclude carotenoid molecules from our treatment, and our usage of the term pigment applies only to the chlorophyll *a* (Chl-a), chlorophyll *b* (Chl-b), and pheophytin molecules, shown in Figure 4.2. Our approach follows previous treatments [93, 79, 95] in which energy transfer within strongly bound pigment clusters is treated by modified Redfield theory and transfer between clusters is described by generalized Förster theory. The second portion describes the physical parameters we use to model the system including the spatial structure, the inter-pigment coupling, and site energies. The final portion outlines how we use the

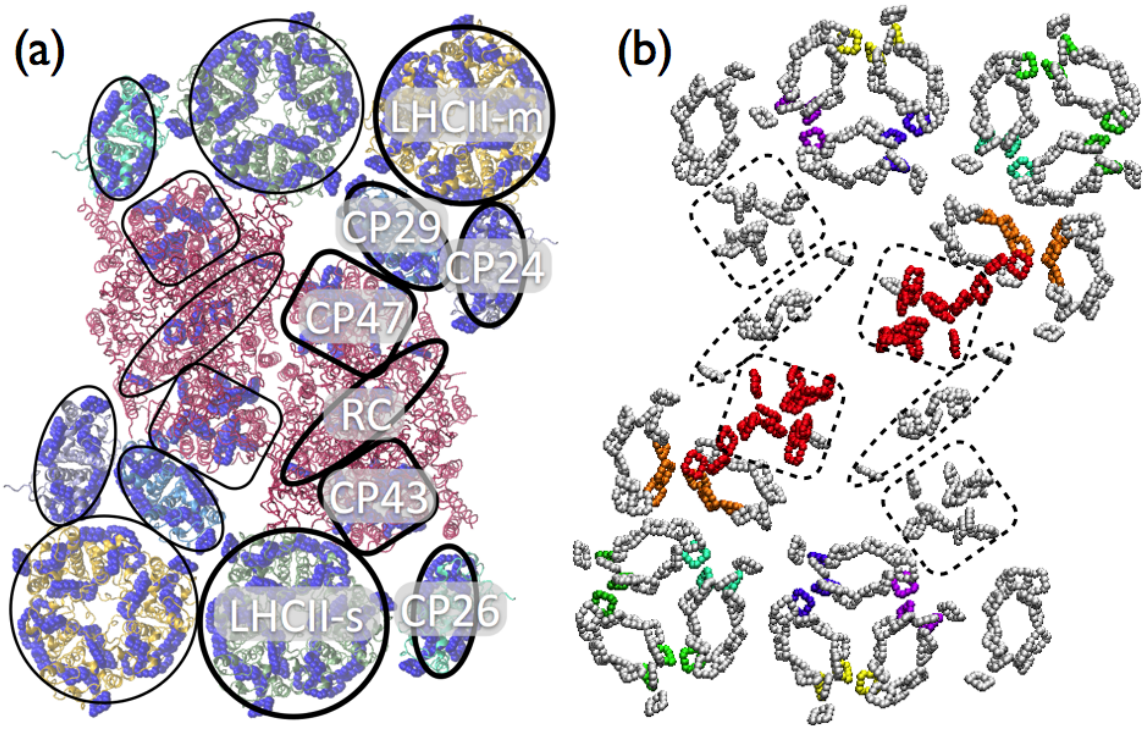



Figure 4.1: The $C_2S_2M_2$ protein structure and pigment arrangement are shown based on the structure determined by Caffarri and coworkers [18]. (a) The protein and pigment arrangement are shown. Chlorophylls are represented as blue spheres outlining the chlorin ring. Thick (thin) black outlines surround proteins associated with the right (left) monomer unit of the $C_2S_2M_2$. (b) The protein scaffold has been removed to expose the pigments bound by each protein unit. Pigments belonging to domains contained within a single protein unit are shown in grey. Pigments assigned to a domain delocalized between more than one protein are colored according to their domain assignment.

calculated rates to construct fluorescence decay spectra and a linearized kinetics model for PSII supercomplexes.

### 4.2.1 Excitation Energy Transfer Theory

#### System and Hamiltonian

To describe pigment-protein complexes (PPC) containing $N$ pigments, we account for only two electronic states of each pigment the singlet ground-state ($S_0 = |\phi_g\rangle$) and the lowest-lying singlet-excited state ($S_1 = |\phi_e\rangle$). We describe the $N$-pigment complex with product states shown in eq. (4.1), (4.2), (4.3) which span the zero-, single-, and double-excitation space, respectively. The higher-order excitation spaces can be constructed analogously. These states are collectively called the site-basis, since they correspond to excitations

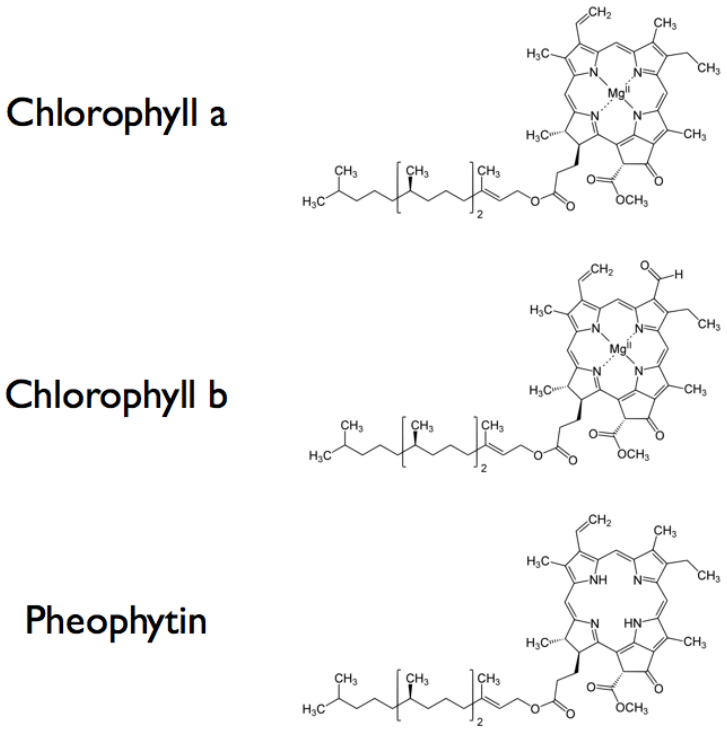



Figure 4.2: The molecular structure of the three pigments (chlorophyll a, chlorophyll b, and pheophytin) that we treat in our energy transfer model for photosystem II supercomplexes.

localized on individual pigments. In the following discussion, the indices $\mu$ and $\gamma$ index the site basis.

$$|0\rangle = \prod_{i=1}^{N} |\phi_{i,g}\rangle \tag{4.1}$$

$$|\mu\rangle = |\phi_{\mu,e}\rangle \prod_{\substack{i=1 \\ i\neq\mu}}^{N} |\phi_{i,g}\rangle \tag{4.2}$$

$$|\mu,\gamma\rangle = |\phi_{\mu,e}\rangle|\phi_{\gamma,e}\rangle \prod_{\substack{i=1 \\ i\neq\mu,\gamma}}^{N} |\phi_{i,g}\rangle \tag{4.3}$$

In light intensities appropriate for the physiological conditions of photosynthetic PPCs, the excitation energy dynamics sample only the single-excitation space [42]. The single-excitation dynamics are driven by the Hamiltonian (shown here in the site-basis) given in eq. (4.4), where $H^{\text{el}}$ contains the electronic vertical-excitation of individual pigments ($\hbar\Omega_\mu$, called the site energy) and the coupling between transitions of pigment pairs ($\hbar J_{\mu,\gamma}$, called

the excitonic coupling), $H^{\text{el-ph}}$ describes the coupling of the electronic and phonon degrees of freedom, and $H^{\text{ph}}$ is the phonon Hamiltonian (within a Harmonic-oscillator approximation) indexed by $\xi$.

$$H = H^{\text{el}} + H^{\text{el-ph}} + H^{\text{ph}} \tag{4.4}$$

$$H^{\text{el}} = \sum_{\mu=1}^{N} |\mu\rangle(\hbar\Omega_\mu + \Delta_\mu)\langle\mu| + \sum_{\mu,\gamma} |\mu\rangle\hbar J_{\mu,\gamma}\langle\gamma| \tag{4.5}$$

$$H^{\text{el-ph}} = \sum_{i=1}^{N} |\mu\rangle\hat{u}_\mu\langle\mu| \tag{4.6}$$

$$H^{\text{ph}} = \sum_{\xi} \hbar\omega_\xi(p_\xi^2 + q_\xi^2)/2 \tag{4.7}$$

The site energy of each pigment is sensitive to changes in the protein configuration that occur on timescales much longer than the timescales of excitation and fluorescence. These long timescale fluctuations are incorporated into the Hamiltonian by adding a random variable $\Delta_\mu$ that samples a Gaussian distribution with zero mean and a standard deviation of $\sigma_\mu$. Different values of $\Delta_\mu$ represent inhomogeneous realizations of the pigments. The electron-phonon coupling is written in terms of the energy gap fluctuation operator ($\hat{u}_\mu$), defined in eq. 4.8, that describes the time-evolution of the energy gap between the ground state and excited state harmonic oscillators [42]. We assume that the excited state harmonic oscillator has the same frequency but a displaced equilibrium position ($R^{(0)}_{\mu,e/g}$) with respect to the ground-state.

$$\hat{u}_\mu = -\sum_{\xi} \hbar\omega_\xi(R^{(0)}_{\mu,e} - R^{(0)}_{\mu,g})q_\xi \tag{4.8}$$

**Weak Inter-Pigment Coupling**

In the limit that the coupling between pigments is much weaker than the electron-phonon coupling, the absorption line shape is a sum of the absorptions of individual pigments in a protein bath. The absorption and fluorescence line shapes of an individual pigment, given by eqs. (4.9) and (4.10), are written in terms of the vertical transition frequency ($\Omega_\mu$), line-broadening function ($g_\mu(t)$), and reorganization energy ($\lambda_\mu$) [42, 83]. The real-value components of the Fourier transform of eq. (4.9) and (4.10) are the frequency domain line shapes of the absorption and fluorescence spectrum, respectively.

$$A_\mu(t) = e^{-i\Omega_\mu t - g_\mu(t)} \tag{4.9}$$

$$F_\mu(t) = e^{-i(\Omega_\mu - 2\lambda_\mu)t - g^*_\mu(t)} \tag{4.10}$$

$$g_\mu(t) = \int_0^\infty \frac{\mathrm{d}\omega}{\pi\omega^2}\chi''_\mu(\omega)\left[(1-\cos(\omega t))\coth\left(\frac{\omega}{2k_bT}\right) + i(\sin(\omega t) - \omega t)\right] \tag{4.11}$$

$$\lambda_\mu = \int_0^\infty \frac{\mathrm{d}\omega}{\pi\omega}\chi''_\mu(\omega) \tag{4.12}$$

The absorption and fluorescence spectra of each pigment are broadened by fluctuations of the phonon environment coupling to the electronic state of the pigments ( $H^{\text{el-ph}}$). This influence is incorporated into the line-broadening function ($g_\mu(t)$) through the spectral density ($\chi''_\mu(\omega)$) that describes the distribution of phonon modes as a function of frequency, weighted by their coupling to the electronic structure of the pigment.

Energy transfer, in this limit, results from coupling between the transition dipole moments of pigments (the off-diagonal terms in $H^{\text{el}}$) and can be described by Förster theory [42, 83]. The rate of transfer is proportional to the square of the coupling between the pigments multiplied by the overlap of the fluorescence spectrum of the donor with the absorption spectrum of the acceptor, as given in eq. (4.13).

$$k_{\gamma\leftarrow\mu} = \frac{|H^{\text{el}}_{\mu,\gamma}|^2}{\hbar^2\pi}\int_0^\infty \mathrm{d}t A_\gamma(t)F^*_\mu(t) \tag{4.13}$$

**Strong Inter-Pigment Coupling**

When the Coulombic coupling between pigments ($\hbar J_{\mu,\gamma}$) is much stronger than the electron-phonon coupling, the excited states of strongly coupled pigments will occur in coherent superpositions called excitons. Excitonic states $|m\rangle$ are the eigenvectors of the electronic Hamiltonian ($H^{\text{el}}$).

$$|m\rangle = \sum_{\mu=1}^{N} U_{\mu,m}|\mu\rangle \tag{4.14}$$

In this limit, the absorption and fluorescence line shapes are determined by both the electron-phonon coupling and the excitonic structure. The absorption and fluorescence lineshapes, described in the time-domain, are shown in eqs. (4.15) and (4.16). Excitonic lineshapes and reorganization energies are linear combinations of pigment functions, see eqs. (4.17) and (4.18) [83].

$$A_m(t) = e^{-i\Omega_m t - G_{mmmm}(t)} \tag{4.15}$$

$$F_m(t) = e^{-i(\Omega_m - 2\Lambda_{mmmm})t - G^*_{mmmm}(t)} \tag{4.16}$$

$$G_{mnpq}(t) = \sum_\mu U_{\mu,m}U_{\mu,n}U_{\mu,p}U_{\mu,q}g_\mu(t) \tag{4.17}$$

$$\Lambda_{mnpq} = \sum_\mu U_{\mu,m}U_{\mu,n}U_{\mu,p}U_{\mu,q}\lambda_\mu(t) \tag{4.18}$$

Energy transfer between excitons occurs because the electron-phonon coupling drives fluctuations in the site energies of pigments (the diagonal term of $H^{\text{el}}$ in the site basis),

which results in an off-diagonal coupling between excitonic states. Phonon driven energy transfer can be described by modified Redfield theory [124, 83], given in eqs. (4.19)-(4.20), where the electron-phonon coupling is transformed into the exciton basis and the off-diagonal terms are treated as a perturbation to second-order.

$$k_{n\leftarrow m} = 2\Re \int_0^\infty \mathrm{d}t A_n(t) F_m^*(t) V_{n,m}(t) \tag{4.19}$$

$$V_{n,m}(t) = e^{2G_{mmnn}(t)+2i\Lambda_{mmnn}(t)} \times \left[\ddot{G}_{mmnn}(t) - \{\dot{G}_{mmmn}(t) - \dot{G}_{mnnn}(t) + 2i\Lambda_{mmmn}\}^2\right] \tag{4.20}$$

The derivation of eqs. (4.19) and (4.20) have been presented previously.[124] [77] A concern with using modified Redfield theory is the appearance of negative rates of transfer as the result of numerical instabilities for some inhomogeneous realizations [124]. In the following, if either the up-hill transfer rate or down-hill transfer rate between any pair of excitons was negative, it was replaced with a rate reconstructed from the other rate of the pair using the detailed-balance criterion. If both rates were negative, the calculation was dropped and a new inhomogeneous realization was selected.

### Constructing Domains

Photosystem II contains 326 pigments with a large range of inter-pigment couplings, so neither the strong- nor weak-interpigment coupling limit is applicable to all the pigment-pigment transfers within the supercomplex. One approach is to divide PSII into groups of pigments that are strongly coupled (called domains) and approximate energy transfer within these groups by the strong coupling limit and between these groups by the weak coupling limit. Photosystem II, however, has many intermediate couplings, thus the separation between strong and weak is ambiguous and the precise grouping of chlorophylls is subjective [95, 79]. We chose to group pigments to improve the separation of timescales between inter- and intra-domain transfer.

In this treatment, pigments must have coupling above a threshold value ($V_{\mathrm{cutoff}} = 15\ \mathrm{cm}^{-1}$) and give rise to excitons that rapidly exchange energy to be included in the same domain. We have selected this value of $V_{\mathrm{cutoff}}$ following the suggestion of Novoderezhkin and coworkers[79], who pointed out that experimental evidence for excitons was found for a photosynthetic PPC with nearly isoenergetic pigments and electronic couplings of $\sim 15\ \mathrm{cm}^{-1}$ [80, 127]. Photosystem II, however, has a wide distribution of chlorophyll site energies, so pigments with electronic coupling close to $15\ \mathrm{cm}^{-1}$ often should not contained in the same domain because the large energy gaps cause slow energy transfer. To incorporate this second component into our domain defintion, we construct a proxy for energy transfer rates that depends only on the electronic Hamiltonian. Since energy transfer between excitons is driven by site fluctuations, two excitons that have substantial overlap in the site-basis will be rapidly mixed by fluctuations in site energy. This allows excitonic overlap to act as a proxy for energy transfer rates. To calculate this overlap, we removed all coupling in $H^{el}$ less

than 15 $cm^{-1}$ and then calculated the transformation matrix ($\tilde{\mathbf{U}}$) associated with this new Hamiltonian. Using the transformation matrix we calculated the excitonic overlap that two sites experience using eq. (4.21). If the overlap is larger than the threshold value ($S_{\mu,\gamma} > 0.1$, selected to reproduce separation of timescales calculated for LHCII by Novoderezhkin and coworkers [79]), the two pigments contribute to excitons that exchange energy rapidly. This overlap, however, is also sensitive to the inhomogenous realization, so the two pigments are only in the same domain if $S_{\mu,\gamma} > 0.1$ in at least 50% of all inhomogeneous realizations. As a result, our domain definitions are defined over an inhomogeneous average, rather than for each individual realization.

$$S_{\mu,\gamma} = \sum_m \frac{\tilde{U}^2_{\mu,m}\tilde{U}^2_{\gamma,m}}{\Upsilon_m} \tag{4.21}$$

$$\Upsilon_m = \sum_\mu \tilde{U}^4_{\mu,m} \tag{4.22}$$

The new hybrid basis ($|M\rangle$) was constructed by forming blocks within $H^{el}$ of the pigments belonging to the same domain and then solving for the eigenvectors of each block ($U_{\mu,M}$). These block-eigenvectors form an orthonormal basis with the property that $\langle M|H^{\mathrm{el}}|N\rangle = 0$ if $M$ and $N$ belong to the same domain and $M$ is not equal to $N$. These vectors do not form an eigenbasis of $H^{\mathrm{el}}$, but they are an eigenbasis within each domain. We then calculated energy transfer within a domain using modified Redfield theory and energy transfer between domains using generalized Förster theory. Generalized Förster theory (eqs. (4.23) and (4.24)) uses coupling matrix elements calculated between the new basis vectors ($|M\rangle$) and overlap integrals between excitonic line shape functions analogous to those given in eq. (4.15) and (4.16), since the basis is composed of states delocalized within each domain [42, 83].

$$k_{M\leftarrow N} = \frac{|V_{M,N}|^2}{\hbar^2\pi}\int_0^\infty \mathrm{d}t A_M(t)F^*_N(t) \tag{4.23}$$

$$|V_{M,N}|^2 = |\sum_{\mu,\gamma} U_{\mu,M}H^{\mathrm{el}}_{\mu,\gamma}U_{\gamma,N}|^2 \tag{4.24}$$

### 4.2.2 Coarse-grain Treatments of Energy Transfer

In this section we define two kinds of coarse-grain treatments of energy transfer for PSII supercomplexes. Figure 4.3 shows three panels that show schematically the difference between the full pigment vs the two coarse-grain treatments. In the full treatment, energy transfer rates are calculated between each exciton, so the overall rate matrix is $N_{\mathrm{pigment}} \times N_{\mathrm{pigment}}$. The coarse-grain treatments reduce the size of the rate matrix by only including effective transfer rates between groups of states.

**Domain Model**

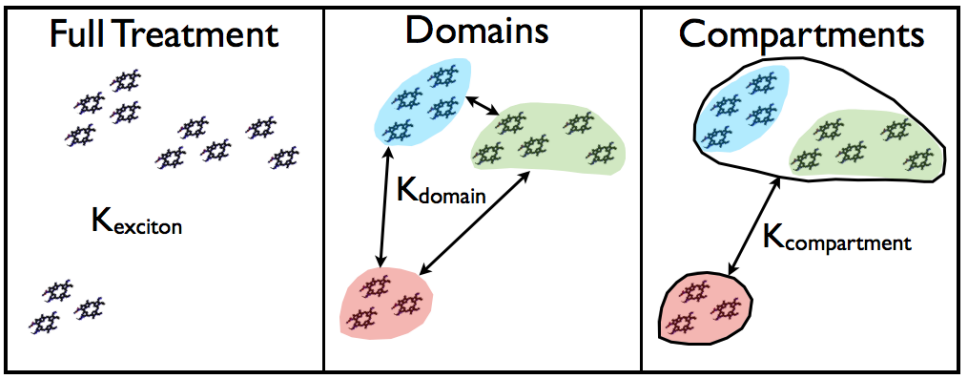



Figure 4.3: Schematic of different coarse-grain treatments of energy transfer for PSII supercomplexes. The left panel shows the collection of pigments which is described in the full model by a rate matrix that containing energy transfer rates between each exitonic state. The middle panel shows the domain assignments of these pigments. Arrows between the domains show the rates of energy transfer included in the domain rate matrix. The right panel shows a compartment assignment, where two domains are assigned to one compartment and the third domain is assigned to another.

The middle panel of Figure 4.3 shows the domain model of the rate matrix, where effective energy transfer rates are calculated between each domain, shown as a colored blob surrounding a group of pigments. The excitation energy transfer rates between two excitons in the same domain are, by construction, much faster than the rate of transfer between domains. This separation of timescales allows for excitation within a domain to thermalize rapidly with respect to transfer out of the domain. The rate of transfer from domain $d$ to domain $a$ is described by eq. 4.25, which performs a sum over the rate of transfer from all the excitons $M$ in $d$ to all the excitons $N$ in $a$ weighted by the Boltzmann population of exciton $M$ ($P_M^{(d)}$) [93].

$$k_{a\leftarrow d}^{\text{dom}} = \sum_{\substack{|M\rangle\in d \\ |N\rangle\in a}} k_{N\leftarrow M} P_M^{(d)} \tag{4.25}$$

$$P_M^{(d)} = \frac{e^{-E_M/k_BT}}{\sum_{|M\rangle\in d} e^{-E_M/k_BT}} \tag{4.26}$$

Excitation within the supercomplex can then be simplified into a collection of domain populations connected by Boltzmann-averaged rates of transfer. The domain model was found to closely reproduce the full model, as shown in the Results and Discussion Section, and will be used extensively in this paper because of the computational simplification that it allows.

**Compartment Models**

The right panel of Figure 4.3 shows an example of the compartment model where one or more domains are assigned to each compartment, represented by a black line encircling a set of pigments. The effective rates of transfer between compartments were calculated assuming rapid thermalization within each compartment. Unlike the domain model, where the coarse-grain states are defined by the original model parameters, the number and composition of compartments can be selected arbitrarily. This flexibility, however, does not ensure that the compartment model will reasonably reproduce the population transfer dynamics of the complete model. We consider compartment assignments that range in size from several domains that incorporate all the chlorophyll in a single protein to 116 domains that incorporate all the chlorophyll in the largest supercomplex ($C_2S_2M_2$). The rates between compartments were calculated using eqs. (4.25) and (4.26), except the acceptor and donor sites were compartments consisting of several domains. In the case where we defined compartments by protein units, it is possible for a domain to be shared between two compartments (if that domain contains pigments from both proteins). When this occurred, the domain was put into the compartment containing more chlorophylls assigned to that domain. The low energy domain shared by CP29 and CP24 consists of three chlorophylls from each protein, so it was decided (arbitrarily) that this domain was put into the CP29 protein compartment. In the model in which all chlorophylls in PSII form a single compartment, the effective rate of initial charge separation $k_{CS}$ (see next Section) was weighted by the Boltzmann population of excitons in the RC domain.

### 4.2.3 Photosystem II: Parameters

#### Structures

The $C_2S_2M_2$ PSII supercomplex is a two-fold symmetric dimer (Figure 4.1a). The supercomplex contains four LHCII trimers and two copies of each of the minor complexes, CP26, CP24, and CP29 as well as two copies of the reaction center core. The structure of the $C_2S_2M_2$ PSII supercomplex was recently obtained at 12 Å resolution [18]. This resolution determined the relative orientations and approximate distances between proteins in the supercomplex. The structure of the reaction center core dimer from cyanobacteria [111] and the structure of the LHCII trimer from spinach [57] have both been solved by x-ray crystallography. We have used a monomer from the LHCII structure for each of the minor complexes. The appropriateness of this substitution is discussed in the Supporting Information (SI). The CP29 crystal structure was recently published [86], but because there is as yet no Hamiltonian for this protein, we did not use it in our calculations. Still, as discussed in Appendix C, the CP29 structure showed high homology with an LHCII monomer, except that Chl 605 in the LHCII monomer has no equivalent in CP29. For this reason, that chlorophyll is deleted in our $C_2S_2M_2$ structure in the LHCII monomer representing CP29. The vertical position of the proteins in the membrane was determined using the Orientations of Proteins in Membranes Database [58]. The lateral orientation was determined using the results of Caffarri and coworkers [18]. The protein structures presented in the paper

were generated using the VMD software package[40]. The different supercomplex structures needed to model the fluorescence lifetime data from Caffarri and coworkers [17] were constructed by deleting proteins from the $C_2S_2M_2$ supercomplex. The LHCII trimers within a supercomplex (up to 2 LHCII-s and 2 LHCII-m) are all identical structures with different nomenclature only to identify their positions within the complex.

**Hamiltonian**

$H^{el}$ is divided into three types of terms: (1) site energies, (2) intra-protein couplings (the interactions of pigments contained within the same protein scaffold), and (3) inter-protein couplings (the interactions of pigments contained in different proteins scaffolds). The site energies, inhomogeneous distributions, and intra-protein couplings were set to literature values [79, 93, 68, 92, 94] and are reproduced in Appendix C. The inter-protein couplings were constructed using the dipole-dipole approximation (eq. (4.27)), where $\vec{R}_{\mu,\gamma}$ is a vector describing the center-to-center difference in position of the pigments $\mu$ and $\gamma$. The transition dipole moments are taken from the literature [79, 93, 68, 92, 94] and their values are reproduced in Appendix C. The spectral density, describing the electron-phonon coupling, for each chlorophyll is the same as that used in the literature when originally extracting the site energy for the chlorophyll. All spectral densities are reproduced in Appendix C.

$$J_{\mu,\gamma} = \frac{\vec{v}_\mu \cdot \vec{v}_\gamma}{|R_{\mu,\gamma}|^3} - 3\frac{(\vec{v}_\mu \cdot \vec{R}_{\mu,\gamma})(\vec{v}_\gamma \cdot \vec{R}_{\mu,\gamma})}{|R_{\mu,\gamma}|^5} \tag{4.27}$$

### 4.2.4 Electron Transfer Scheme

A detailed model of excitation and electron transport within the reaction center presented by Novoderezhkin and coworkers [82] proposed that charge separation can be initiated through either the excited states of special pair ($P^*_{D2}/P^*_{D1} \rightarrow P^+_{D2}P^-_{D1}$) or the associated D1 chlorophyll ($Chl^*_{D1} \rightarrow Chl^+_{D1}Phe^-_{D1}$). The electron transport scheme contains five different charge separated populations with eleven rate constants. Novoderzhkin and coworkers used pump-probe data to parameterize their electron transfer model. In this work, we will be simulating fluorescence decay curves which are only sensitive to the very coarse features of energy transfer. As such, we will decrease the number of fit parameters associated with our model. We chose to describe the electron transport scheme phenomenologically by including only two radical pair states. The two states (RP1 and RP2) and three kinetic rates that define the electron transfer process in our model are shown in Figure 4.4. The radical pair states RP1 and RP2 do not have direct physical interpretation in terms of molecular charge separated states. Instead, the kinetic rates and populations describe the overall behavior of charge separation without assuming any particular molecular identity for the electron transport through the reaction center. The three kinetic rates ($k_{CS}$, $k_{RC}$, and $k_{irr}$) describing electron transfer were determined by fitting the experimental fluorescence decay spectra with simulated curves. These are purely empirical parameters and the values we determine for them depend upon the energy transfer model used.

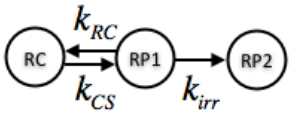



Figure 4.4: Charge separation model. RC represents the reaction center domain, which contains four chlorophyll (two special pair and two associated chlorophyll) and two pheophytin. Each reaction center in a PSII supercomplex performs electron transfer using the same rate constants.

### 4.2.5 Simulating Fluorescence Decays

The master-equation formalism (eq. (4.28)) defines energy transfer in terms of the rate matrix ($K$) which contains the transfer rates between excitons or domains and the transfer rates from the RC to the charge separated states and the population vector ($P(t)$) describing the distribution of excitation at a given time. Additional loss pathways in the form of radiative ($k_{fl}$) and non-radiative ($k_{nr}$) are also present in the system and are treated explicitly as additional compartments within the kinetic model. The fluorescence rate of each exciton was scaled by its transition dipole moment squared with the average fluorescence timescale across all states of the system set to 16 ns [7, 85]. The non-radiative rate was assumed to be equal for all excitons with a timescale of 2 ns [7, 126].

$$\dot{P}(t) = KP(t) \tag{4.28}$$

The fluorescence decay spectra were calculated as a function of the initial excitation by projecting into the basis of the eigenvectors of the rate matrix ($|u\rangle$), as shown in eqs. 4.29-4.31 [123]. The decay timescales ($\tau_u$) are the inverse of the $u^{th}$ eigenvalue of the rate matrix, the fluorescence weight factor ($0 \leq \tilde{k}_m^{fl} \leq 1$) is the relative fluorescence rate of each exciton, and the similarity transform matrix ($\Theta^{-1}$) projects from excitons ($|M\rangle$) to the eigenvectors of the rate matrix ($|u\rangle$).

$$F(t) = \sum_M \tilde{k}_M^{fl} \langle M|e^{Kt}|P(0)\rangle \tag{4.29}$$

$$F(t) = \sum_u a_u e^{-t/\tau_u} \tag{4.30}$$

$$a_u = \sum_M \tilde{k}_M^{fl} \langle M|\Theta|u\rangle\langle u|\Theta^{-1}|P(0)\rangle \tag{4.31}$$

Since experimental measurements of fluorescence lifetimes using time-correlated single photon counting have finite resolution we have implemented resolution loss by setting a lower limit to lifetime components ($\tau_u$) of 8 ps. Any amplitude $a_u$ associated with a lifetime less than the lower limit is shifted up to 8 ps.

### 4.2.6 Calculating Linearized Kinetics

To analyze the rate limiting steps for trapping within PSII, we have chosen to calculate a decomposition of the average trapping time in terms of a hierarchical linearization of the overall kinetic rate matrix. Following the work of Yang and Fleming [125], the average trapping time can be described using a sequence of rate matrices ($\mathbf{K}_i$) and kinetic compartments ($\sigma_i$) constructed to satisfy the condition that the $i^{th}$ compartment is a perfect trap within the $\mathbf{K}_{i+1}$ rate matrix. RP2 is a trap for the entire rate matrix, so we begin by defining $\sigma_0 = \text{RP2}$ and $\mathbf{K}_1$ is the domain rate matrix, composed of thermal transfer rates between domains. Subsequent kinetic compartments and rate matrices are defined in terms of eq. (4.32) and (4.33).

$$\mathbf{K}_i = \mathbf{K}_{i-1} - \mathbf{K}_{i-1}|\sigma_{i-1}\rangle\langle\sigma_{i-1}| \quad \text{for } i \geq 2 \tag{4.32}$$

$$\langle\sigma_i| = \langle\sigma_{i-1}|\frac{\mathbf{K}_i}{\sqrt{\langle\sigma_{i-1}|\mathbf{K}_i\mathbf{K}_i^T|\sigma_{i-1}\rangle}} \tag{4.33}$$

The effective forward transfer time ($\tau^{\text{eff}}_{i\leftarrow i+1}$) is the average time for transfer from compartment $i+1$ to compartment $i$ in the presence of the back-transfer away from the final charge separation state (RP2). The effective forward transfer time described in eq. (4.34) contains two contributions. The first ($k^{-1}_{i\leftarrow i+1}$) can be considered the inverse of the rate of transfer between the $i+1$ and $i$ compartments. The second term, $(1 + r_{i+2})$ corrects the forward transfer rate for trajectories that first back-transfer to compartments greater than $i+1$.

$$\tau^{\text{eff}}_{i\leftarrow i+1} = k^{-1}_{i\leftarrow i+1}(1 + r_{i+2}) \tag{4.34}$$

$$k^{-1}_{i\leftarrow i+1} = \sqrt{\langle\sigma_{i-1}|\mathbf{K}_i\mathbf{K}_i^T|\sigma_{i-1}\rangle} \times \frac{\sum_{n\in\text{Domain}}\langle n|\sigma_{i-1}\rangle}{\sum_{n\in\text{Domain}}\langle n|\sigma_i\rangle} \tag{4.35}$$

$$r_i = -\frac{\sum_{n\in\text{Domain}}^{N_{max}}\langle n|\mathbf{Q}_{i+1}(\mathbf{Q}_{i+1}\mathbf{K}_1\mathbf{Q}_{i+1})^{-1}\mathbf{Q}_{i+1}\mathbf{K}_i|\sigma_i\rangle}{\sum_{n=1}^{N_{max}}\langle n|\sigma_i\rangle} \tag{4.36}$$

$$\mathbf{Q}_i = \mathbf{I} - \sum_{n=1}^{i-1}|\sigma_n\rangle\langle\sigma_n| \tag{4.37}$$

The average trapping time is determined by both the effective rates of transfer along the chain of compartments and the initial population of excitation within the system. The contribution of $\tau^{\text{eff}}_{(i-1)\leftarrow(i)}$ to the average trapping time is weighted by $\rho_i$, the fraction of population initiated in compartments $j \geq i$.

$$\tau_{\text{trap}} = \sum_{i=0}^{N_{max}} t_{i-1\leftarrow i} \tag{4.38}$$

$$t_{i-1\leftarrow i} = \tau^{\text{eff}}_{i\leftarrow i+1}\rho_{i+1} \tag{4.39}$$

$$\rho_i = \sum_{j=i}^{N_{max}} P_j(t=0) \tag{4.40}$$

## 4.3 Model of Photosystem II

### Pigment Domain Assignments Reveal Inter-Protein Domains

Photosystem II pigments (Chl-a, Chl-b, and pheophytin) were grouped into strongly interacting clusters called domains. While it is often assumed that pigments will interact most strongly with other pigments bound within the same protein [13, 112], we have used the supercomplex structure to assign domains with an unbiased approach that depends only on the couplings and energy gaps of the relevant chlorophylls. This method has revealed a small set of pigments that form domains delocalized across multiple proteins. Figure 4.1b shows the $C_2S_2M_2$ supercomplex with the protein scaffold removed to display only the pigments. Chlorophylls belonging to a domain delocalized over more than one protein unit are shown in colors according to their domain assignment. Within LHCII trimers we see delocalization between high energy pigments belonging to different monomers (e.g. the yellow, blue, and purple chlorophylls bound by LHCII-s). CP24 and CP29 share a domain that contains the lowest energy exciton of both proteins (the orange chlorophyll). Additionally, CP29 has two pigments that belong in the same domain as 15 of the 16 CP47 chlorophylls (the red chlorophylls). The remainder of the chlorophyll (shown in grey) have the same domain assignments within a $C_2S_2M_2$ supercomplex as they would in an isolated protein and their domain assignments can be found in the supplementary information (SI).

### Testing Infinitely Fast Intra-Domain Equilibration

The domain assignments used in this work were selected to increase the separation of timescales between inter- and intra-domain transfer rates. In the limit that a complete separation of timescales is achieved, the much faster intra-domain rates will result in excitation energy thermalizing prior to transfer out of the domain. We constructed a new rate matrix (domain model) that coarse-grains the system at the domain level by assuming infinitely fast intra-domain thermalization. To assess the error associated with this domain model, we calculated the domain populations as a function of time both with and without assuming infinitely fast equilibration within domains. For these calculations, electron transfer in the RC is treated as infinitely fast and irreversible. To assess the similarity of the dynamics in both cases, we calculate two error metrics $\Delta^{(d)}_{\text{max}}$ and $\Delta^{(d)}_{\text{Integral}}$. $\Delta^{(d)}_{\text{max}}$ measures the maximum absolute deviation of the two population traces as a percentage of the maximum population of the domain calculated with the full rate matrix. The second metric, $\Delta^{(d)}_{\text{Integral}}$, reports the integral of the absolute deviation over all time as a fraction of the integral of the population calculated using the full rate matrix over all time. The mathematical expressions for these terms are described in Appendix C.

| | B8-B11 | B7 (Core Complex) |
|---|---|---|
| $\tau_{\mathrm{CS}}$ (ps) | 0.64 | 4.4 |
| $\tau_{\mathrm{RC}}$ (ps) | 160 | 130 |
| $\tau_{\mathrm{irr}}$ (ps) | 520 | 250 |

Table 4.1: Best fit electron transfer timescales for the Domain Model.

No matter what domain is initially excited, we found that $\Delta^{(d)}_{\mathrm{max}} < 2\%$ and $\Delta^{(d)}_{\mathrm{integral}} < 7\%$ for all domains of $C_2S_2M_2$. The population curve with the largest error in both of these measures is shown in Figure C.1. The difference between the dynamics calculated with the full rate matrix versus the domain rate matrix in this plot is barely differentiable by eye. The excellent agreement between the domain model and the full calculation supports our method for assigning chlorophylls to domains using a separation of timescales metric. We have used the domain model as the reference calculation for the remainder of the paper.

**Extracting Electron Transfer Rates**

The fluorescence lifetime of a PSII supercomplex is determined by both the rates of energy transfer in the antenna and electron transfer kinetics in the reaction center. We incorporated the electron transfer rates (shown in Figure 4.4) into the domain model and extracted their values by fitting fluorescence decay spectra. In previous work, Caffarri and coworkers[17] measured fluorescence decay spectra collected from 5 bands separated using sucrose density gradient centrifugation on thylakoid membranes. These bands were labeled, in order of increasing size, as B7 to B11. The protein composition of each band is shown by the cartoon in the top-right corners of Figure 4.5b-f. No single set of electron transfer rates reasonably fit both the core complex (B7) and any of the larger supercomplexes simultaneously. As a result, two different sets of electron transfer timescales ($\tau = k^{-1}$) are shown in Table 4.1, one for the fit to the core complex data (B7) and the other for the fit to the data from the larger supercomplexes (B8-B11). The core complex data is described by a much slower initial charge separation and faster irreversible electron/hole separation. The difference between the core and supercomplex electron transfer kinetics has been observed previously [17]. The origin, however, remains unclear: it could be an artifact of the modeling or the result of a physical difference in the samples not accounted for in the current structure/parameter data. In the following we use the B8-B11 electron transfer rates to explore the dynamics of light harvesting in the largest supercomplex, $C_2S_2M_2$ (B11).

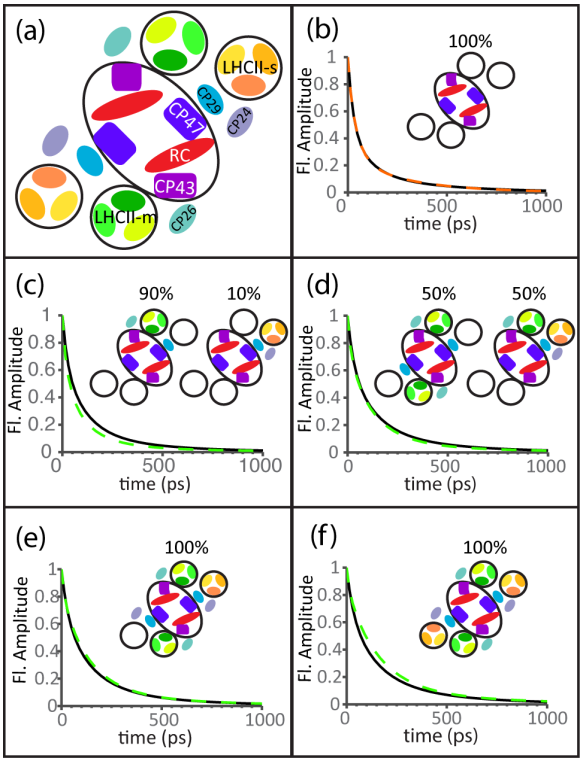



Figure 4.5: (a) A cartoon structure of PSII supercomplex proteins is shown. The contributions of different supercomplexes to each band are represented by similar cartoons in the top-right corner of each box. A comparison between experimental (black line, taken from Caffari and coworkers[17]) and simulated (orange and green dashed-lines for B7 and B8-B11, respectively, calculated with the domain model) fluorescence decay curves for (b) B7, (c) B8, (d) B9, (e) B10, and (f) B11. The simulated decay curves are calculated using the best fit electron transfer timescales given in Table 4.1. The experimental curves discard the longest lifetime component following the work by Caffarri and coworkers (discussed in Appendix C).

## 4.4 Results & Discussion

### 4.4.1 Timescales of Light Harvesting in $C_2S_2M_2$

The rate matrix provides a complete description of the kinetics of excitation energy and electron transfer, but a simplified picture would decompose the average timescale for trapping into contributions from a small set of effective forward transfer rates. In the literature [17, 115], the overall process of excitation energy initialized in an antenna protein being converted to chemical energy is decomposed into four steps:

1. the excitation energy initial migrates among the antenna proteins until it reaches the pigments that are kinetically connected to the reaction center,

2. the excitation energy transfers into the reaction center (RC),

3. excitation at the RC initiates charge separation, and

4. the electron is transferred downstream resulting in irreversible charge separation.

These steps are usually thought of as occurring sequentially, though back-transfer can occur at any point prior to the irreversible charge separation. To extract intuitive insight into the physical mechanisms associated with the rate matrix that we have calculated, we want to breakdown the timescale of the overall process of energy capture into the individual timescales associated with these four steps: 1. the migration ($\tau_{\text{mig}}$), 2. transfer-to-trap ($\tau_{\text{TtT}}$), 3. initial charge separation ($\tau_{\text{iCS}}$), and 4. the irreversible charge separation ($\tau_{\text{irrCS}}$) timescales, respectively. The breakdown of the overall timescale of capture into the component timescales is shown in eq. (4.41).

$$\tau_{\text{capture}} = \tau_{\text{mig}} + \tau_{\text{TtT}} + \tau_{\text{iCS}} + \tau_{\text{irrCS}} \tag{4.41}$$

It is generally agreed that energy capture in PSII is limited by the rate of irreversible electron hole separation ($\tau_{\text{irrCS}}$), which is one to two orders of magnitude slower that the other timescales involved [114, 65, 13, 36]. We define light harvesting as the conversion of excitation energy to chemical energy, which does not require irreversible charge separation. As such, light harvesting is described by only the first three timescales associated with energy capture ($\tau_{\text{LH}} = \tau_{\text{mig}} + \tau_{\text{TtT}} + \tau_{\text{iCS}}$). Previous work[114, 65, 115] has suggested that the kinetics of light harvesting in PSII are either trap limited ($\tau_{\text{iCS}} \gg \tau_{\text{TtT}} + \tau_{\text{mig}}$) or transfer-to-trap limited ($\tau_{\text{TtT}} \gg \tau_{\text{iCS}} + \tau_{\text{mig}}$). Knowledge of the rate-limiting step would inform our interpretation of fluorescence lifetime data from both supercomplexes and intact membranes [25, 96].

One challenge in determining the relative contributions of $\tau_{\text{mig}}$, $\tau_{\text{TtT}}$, $\tau_{\text{iCS}}$, and $\tau_{\text{irrCS}}$ is that they are not directly related to the rates of transfer in the rate matrix. These timescales describe the overall behavior of energy transport, which involves forward and reverse rates of transfer mixed with entropic effects (the connectivity of the different domains within the rate matrix). Yang and Fleming[125] have demonstrated a method that decomposes the average trapping time into contributions from different physical processes. In this treatment, a linear chain of compartments ($\sigma_i$), as shown in Figure 4.6, is constructed to reproduce the mean first-passage time to RP2 (irreversible charge separation). The compartments are linear combinations of domains that form an orthonormal set. Excitation energy flows down the chain from the periphery of the supercomplex ($\sigma_{N_{max}}$) to the final charge separation state RP2 ($\sigma_0$). The effective forward transition time ($\tau^{\text{eff}}_{i-1\leftarrow i}$, given in eq. (4.34)) describes the timescale of transfer from $\sigma_i$ to $\sigma_{i-1}$ corrected for trajectories where excitation initially performs back transfer to higher compartments in the chain prior to transferring from $\sigma_i$ to $\sigma_{i-1}$. These timescales, however, do not account for the initial distribution of excitation. A population-weighted effective transfer time ($t_{i-1\leftarrow i}$) weights $\tau^{\text{eff}}_{i-1\leftarrow i}$ for the population of excitation that will flow through compartment $i$, which takes into account the initial excitation distribution (eqs. (4.38) - (4.40)).

The most important advantage of this linear chain of compartment is that the timescales of transfer calculated along the chain can be directly related to the four timescales of energy capture. Irreversible charge separation occurs when RP1 ($\sigma_1$) transfers to RP2 ($\sigma_0$), so

$\tau_{\mathrm{irrCS}} = t_{0\leftarrow 1}$. Initial charge separation occurs when excitation energy at the RC ($\sigma_2$) is converted into a charge separated state RP1 ($\sigma_1$), so $\tau_{\mathrm{iCS}} = t_{1\leftarrow 2}$. Excitation undergoes transfer-to-trap when moving from pigments immediately adjacent to the RC to pigments within the RC, which gives $\tau_{\mathrm{TtT}} = t_{2\leftarrow 3}$. Finally, the general migration of excitation from the initial antenna protein to the pigments immediately adjacent to the RC is described by a series of energy transfer steps down the chain from the initial excitation compartment to $\sigma_3$, $\tau_{\mathrm{mig}} = \sum_{n=3}^{N_{max}} t_{n\leftarrow n+1}$. Using these assignments and the linear chain, the timescales associated with energy capture can be calculated for different initial excitations. In the following we will consider two different initial conditions: a) an even distribution of initial excitation among all the chlA molecules in $C_2S_2M_2$ and b) initial excitation on the periphery of $C_2S_2M_2$.

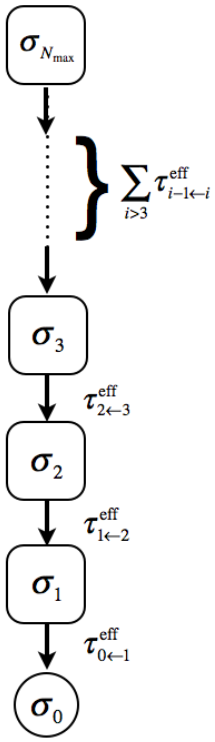



Figure 4.6: A diagram of the linear excitation energy cascade is shown where $\sigma_0$ represents RP2 in $C_2S_2M_2$. The effective timescales of transfer between compartments ($\tau^{\mathrm{eff}}_{i\leftarrow i+1}$) is the average time of transfer from compartment $i$ to $i-1$ averaging over all pathways starting at compartment $i$.

The experimental fluorescence decay curves plotted in Figure 4.5 (black lines) have initial excitation evenly distributed among Chl-a molecules in the PSII supercomplexes. This is an important initial condition to consider since most spectroscopic measurements of supercomplexes involve the excitation of a wide spatial distribution of chlorophylls. We have calculated the population-weighted effective transfer times for evenly distributed Chl-a exci-

| | Assignments | $t^{\text{chlA}}$ (ps) | $t^{\text{periphery}}$ (ps) [78] | Description |
|---|---|---|---|---|
| $\tau_{\text{irrCS}}$ | $t_{0\leftarrow 1}$ | 560 | 560 | RP2 $\leftarrow$ RP1 |
| $\tau_{\text{iCS}}$ | $t_{1\leftarrow 2}$ | 11 | 11 | RP1 $\leftarrow$ RC |
| $\tau_{\text{TtT}}$ | $t_{2\leftarrow 3}$ | 94 | 100 | RC $\leftarrow$ Antenna |
| $\tau_{\text{mig}}$ | $\sum_{n=3}^{Nmax} t_{n\leftarrow n+1}$ | 50 | 110 | Migration in Antenna |

Table 4.2: The timescales of transfer for the effective linearization scheme of $C_2S_2M_2$ supercomplex.

tation ($t^{\text{chlA}}_{i-1\leftarrow i}$), shown in Table 4.2. The contributions of different components to $\tau_{\text{LH}}$ breaks down as follows: 30% from $\tau_{\text{mig}}$, 60% from $\tau_{\text{TtT}}$, and 10% from $\tau_{\text{iCS}}$. Given the similarity of the diffusion and transfer-to-trap contributions, when Chl-a molecules are evenly excited, light harvesting in $C_2S_2M_2$ should not be considered to occur in any of the previously suggested limits, though it is closer to the transfer-to-trap limit than any other.

When a supercomplex resides within a membrane, a larger fraction of excitations will be initialized in the periphery of the supercomplex as a result of energy transfer from adjacent LHCII trimers or supercomplexes. This initial condition is quite different from those expected for measurements on isolated supercomplexes. Using our linearized kinetic model of PSII supercomplexes, however, we can explore the timescales of trapping associated with excitations entering from the edge of the supercomplex ($t^{\text{periphery}}_{i-1\leftarrow i}$). For peripheral excitations, $\tau_{\text{mig}}$ and $\tau_{\text{TtT}}$ each contribute approximately 50% of $\tau_{\text{LH}}$. Incorporating additional energy transfer into and out of the domains contained in $C_2S_2M_2$, as a result of the additional pigments bound in the membrane, will increase both $\tau_{\text{mig}}$ and $\tau_{\text{TtT}}$. It is likely, however, that $\tau_{\text{mig}}$ will increase more as the effective antenna size increases, and the average timescale for light harvesting will shift towards the diffusion limit.

### 4.4.2 Spatial Aspects of Light Harvesting in $C_2S_2M_2$

The overall efficiency of photoconversion depends on the energy transfer rates that arise from the combined energetic and spatial organization of chlorophyll in PSII supercomplexes. Our model shows no evidence of an energy funnel, in agreement with previous work on PSII antenna and core complex [93, 79]. Studies of LHCII trimers have suggested that a low energy cluster of Chl-a act as a favored exit site [102]. Using the structure of PSII supercomplexes and considering the distances between chlorophyll, Caffarri and coworkers [18] have extended this thought by proposing that the exit site of LHCII-s should experience fast transport to a nearby pigment of CP43. A series of such rapid energy transfer steps could form a favored pathway from the peripheral LHCII trimers to the RC. Such a pathway would increase the photochemical yield by allowing excitation energy to reach a reaction center prior to loss by fluorescence or intersystem crossing. Using the domain model we explore the appropriate description of the spatial distribution of energy flow through PSII supercomplexes.

We tested for the presence of favored energy transfer pathways in each quadrant of PSII.

If a small number of such pathways exist, the photochemical yield should substantially decrease when one of these energy pathways is disrupted by removing an essential domain. We selected the two unique quadrants of $C_2S_2M_2$ as the subsystems of interest: the bottom right quadrant (RC, CP43, CP26, LHCII-s) and the top right quadrant (RC, CP47, CP29, CP24, LHCII-m), as shown in Figure 4.1a. We first calculated the photochemical yield of both subsystems. Subsequently, we performed a series of calculations of the photochemical yield for each subsystem in which one of the domains of the subsystem was removed. In all calculations Chl-a's were evenly excited on the LHCII monomer farthest from the RC. There was only one case where removing a domain reduced the photochemical yield to less than 85% of the photochemical yield of the intact subsystem. The one case corresponded to removing the 17 chlorophyll domain shared between CP47 and CP29, which removes all but one of the chlorophylls bound by CP47. This results in greatly suppressed photochemical yield since, as can be seen in Figure 4.1b, the deletion of the large CP47 domain removes all chlorophyll in the vicinity of the reaction center. The general robustness to domain removal demonstrates that no single domain (excluding the 17 chlorophyll CP47 domain) is essential for efficient energy transfer to the reaction center. Instead, there are many energy transfer rates substantially faster than the timescale of loss (fluorescence and intersystem crossing) and this results in a collection of pathways that are fast enough to maintain high photochemical yield.

Next, we characterized how the timescale of light harvesting depends on a single spatial dimension, the average distance of the chlorophyll composing a domain to the nearest RC ($d_{\mathrm{RC}}$). Figure 4.7 plots the average timescale for light harvesting ($\tau_{\mathrm{LH}}$) for excitation initiated in each domain against $d_{\mathrm{RC}}$. The calculation of $\tau_{\mathrm{LH}}$ for excitation initiated in a particular domain is equivalent to calculating the mean first passage time from that domain to the first radical pair state (RP1) [89]. The dots represent the results of one calculation. The colored contours (with increasing density from blue to red) show the underlying distribution extracted from calculations with 500 realizations of the site energies. The overall process of light harvesting in PSII supercomplexes is insensitive to inhomogeneous realizations, as demonstrated by the similar width of the distributions for a single realization and for 500 realizations. Further, $\tau_{\mathrm{LH}}$ for an excitation is linearly dependent on its distance to the nearest RC, with a slope of approximately ten picoseconds per nanometer.

When a supercomplex is embedded in the grana membrane, it is surrounded by a dense array of pigment-protein complexes. A reaction center embedded in this array might be capable of efficiently capturing excitation energy initially absorbed far outside of the supercomplex it belongs to. Light harvesting is efficient anywhere that $\tau_{\mathrm{LH}}$ is substantially smaller than the $\sim$2 ns timescale of excitation loss to unproductive pathways (fluorescence and intersystem crossing). The linear dependence shown in Figure 4.7 cannot be directly extended to a membrane since the slope of the line will vary with the increased number of pigments and the unassociated LHCII may experience slower transport into the supercomplex than proteins bound by the supercomplex experience within it. However, the large disparity between the timescale of excitation loss ($\sim$2 ns) and the $\sim$200 ps timescale of light harvesting found at the periphery of $C_2S_2M_2$ suggests that photosystem II reaction centers

can efficiently harvest light from antenna assemblies much larger than that of $C_2S_2M_2$.

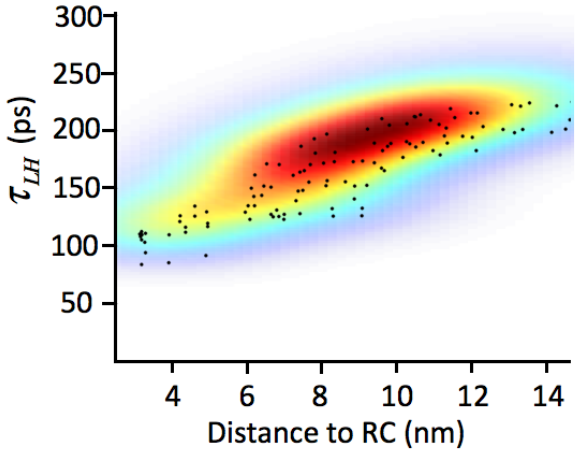



Figure 4.7: The timescale of light harvesting ($\tau_{\mathrm{LH}}$) for excitation initiated in each domain is plotted against the average distance from that domain to the nearest reaction center. The black dots denote the results from one inhomogeneous realization of the site energies. The colored contours (red indicating more points, blue indicating less) show the underlying distribution extracted from compartments calculated from 500 inhomogeneous realizations of the site energies. The underlying distribution was calculated with 2D Gaussian smoothing.

### 4.4.3 Quality of Fit to Fluorescence Decays is not a Proxy for Accuracy of Energy Transfer Models

Energy transfer models of PSII have typically been constructed by fitting kinetic models to fluorescence decay curves. The quality of fit has been used as an indicator of how well these kinetic models describe the underlying dynamics [38, 116]. However, it may be that the fluorescence lifetime data is too coarse-grained to be able to distinguish between models with many fitting parameters [114, 108]. We explored this possibility by simulating fluorescence lifetime curves using three different coarse-grain treatments of the domain model (SI, Fig. 2) that correspond to kinetic limits that have been considered previously in the literature: 1) the ERPE model [65] assumes fast equilibration within all pigments of PSII, 2) the transfer-to-trap limited (TTL) model [114] assumes fast equilibration within the antenna pigments but slow transfer from the antenna to the RC (and vice-versa), and 3) the protein model [13, 17] assumes excitation equilibrates rapidly within a protein followed by slower transfer between proteins. Each of these models was constructed by coarse graining the domain model into compartments defined by regions of fast equilibration. The rates between compartments were calculated using eqs. (4.25) and (4.26). None of these coarse-grained treatments quantitatively reproduced the population dynamics of the initial charge separated state (RP1) for $C_2S_2M_2$ calculated with the domain model. Despite any errors

| B8-B11 | Domain | Protein | TTL | ERPE |
|---|---|---|---|---|
| $\tau_{CS}$ (ps) | 0.64 | 2.1 | 2.2 | 0.28 |
| $\tau_{RC}$ (ps) | 160 | 140 | 91 | 1.0 |
| $\tau_{irr}$ (ps) | 520 | 260 | 190 | 33 |
| RMSE ($\times 10^{-2}$) | 1.7 | 1.6 | 1.4 | 1.9 |

Table 4.3: The best fit electron transfer timescales for the Domain, Protein, Transfer-to-Trap Limited, and ERPE models.

in reproducing the population dynamics of $C_2S_2M_2$, each of the three models is capable of simultaneously reproducing all four experimental fluorescence decay curves measured for the B8-B11 bands of PSII supercomplexes if we allow the three electron transfer parameters to vary. The time-constants that describe electron transfer for the best simultaneous fit to bands B8-B11 for each model are shown in Table 4.3, and the root-mean square error (RMSE) is reproduced in the last row. The electron transfer rates extracted for each model vary significantly from the domain model results, simply because the energy transfer rates describe substantially different dynamics. The quality of the fits (as reported by the RMSE), however, is very similar to that produced by the domain model. Therefore, a model with a good fit to fluorescence lifetime data does not necessarily describe the energy transfer dynamics correctly. The electron transfer rates extracted from such a fit will not accurately describe the kinetics of electron hole separation in reaction centers. More generally, caution should be taken when making mechanistic claims based on fitting kinetic models to fluorescence decay data.

## 4.5 Conclusions

An accurate model of energy transfer in PSII and grana membranes is necessary for the proper interpretation of chlorophyll fluorescence measurements, which are widely used to study photosynthesis[88]. Detailed models of energy transfer have been developed for several components of PSII [93, 79, 68, 69, 82]. We have constructed the first structure-based model of light harvesting in an assemblies of antenna proteins and RCs larger than the core complex, the PSII supercomplexes. The model describes how the specific arrangement of light harvesting proteins and pigments gives rise to efficient energy capture and the fluorescence lifetimes previously measured by Caffarri and coworkers [17]. This model can be used to put lower bounds on energy transfer timescales in the membrane. We also anticipate extending this model to explicitly treat energy transfer through the grana membrane by combining the current structural data of PSII and LHCII with images [52] and simulations [103] of protein organization in the grana.

In developing a structure-based model of energy transfer and trapping in PSII supercomplexes, we used Hamiltonians and spectral densities previously determined for isolated PPCs to calculate modified Redfield/generalized Förster energy transfer rates. In order to

simplify the calculations, we used a new metric for defining domains which improves the separation of timescales between intra- and inter-domain transfer rates. We were able to extract the electron transfer rates assuming a simple two-state charge separation scheme that incorporated three empirical rate constants to describe charge separation ($k_{\mathrm{CS}}$), charge recombination ($k_{\mathrm{RC}}$), and irreversible charge separation ($k_{\mathrm{irr}}$). While a structure for the largest PSII supercomplex has been measured, the resolution remains low so the details of coupling between proteins may change as better structures become available. Further, while the current treatment of energy transfer has the advantage of being computationally tractable for large systems, there remains an open question about how appropriate this description of energy transfer is for actual PSII supercomplexes. This can be tested in the future by comparing population dynamics calculated with more sophisticated energy transfer theories on subsets of the PSII supercomplex.

The results from analyzing our model of energy transfer in photosystem II supercomplexes are as follows:

1. The domain model is the most coarse-grained model that quantitatively reproduces the energy transfer dynamics of $C_2S_2M_2$ calculated with the full modified Redfield/generalized Förster rate matrix.

2. A linearized kinetic model of trapping showed that the timescale for light harvesting in $C_2S_2M_2$ is not well described by any limiting case but is primarily due to a combination of the diffusion and transfer-to-trap rates in the light harvesting antenna. The longer timescales of diffusion experienced by excitations initiated on the periphery of $C_2S_2M_2$ suggests that the timescale for light harvesting in grana membranes shifts towards the diffusion limit.

3. Analyzing the sensitivity of the photochemical yield of $C_2S_2M_2$ to the removal of domains has demonstrated that there is not a small set of favored energy transfer pathways that are required for efficient energy transport and capture in PSII supercomplexes. Therefore, the efficiency of energy capture is created by many rates of energy transfer between domains of chlorophyll that are much faster than the rates of excitation quenching by loss pathways such as inter-system crossing and fluorescence.

4. The timescale for light harvesting of an excitation was shown to depend linearly on its distance to the nearest reaction center. Based on this trend, the capture radius of PSII RC in the grana membrane is significantly larger than the 15 nm maximum distance to a RC achieved in $C_2S_2M_2$.

5. Energy transfer models that do not correctly treat population dynamics within PSII supercomplexes can still fit the measured fluorescence decay curves. The resulting electron transfer rates, however, do not accurately describe electron/hole separation in the reaction center.

# Bibliography

[1] D Abramavicius, B Palmieri, DV Voronine, F Sanda, and S Mukamel. Coherent multidimensional optical spectroscopy of excitons in molecular aggregates; quasiparticle versus supermolecule perspectives. *Chem. Rev.*, 109:2350–2408, 2009.

[2] D Abramavicius, DV Voronine, and S Mukamel. Double-quantum resonances and exciton-scattering in coherent 2D spectroscopy of photosynthetic complexes. *PNAS*, 105:8525, 2008.

[3] K Amarnath, J Zaks, SD Park, KK Niyogi, and GR Fleming. Fluorescence lifetime snapshots reveal two rapidly reversible mechanisms of photoprotection in live cells of Chlamydomonas reinhardtII. *PNAS*, 109:8405–8410, 2012.

[4] M Ballottari, J Girardon, L Dall'osto, and R Bassi. Evolution and functional properties of photosystem ii light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta*, 1817:143–57, 2012.

[5] LM Barter, M Bianchietti, C Jeans, MJ Schilstra, B Hankamer, BA Diner, J Barber, JR Durrant, and DR Klug. Relationship between excitation energy transfer, trapping, and antenna size in photosystem ii. *Biochemistry*, 40:4026–34, 2001.

[6] J Beatty, J Overmann, and M Lince. An obligately photosynthetic bacterial anaerobe from a deep-sea hydrothermal vent, journal = PNAS, year = 2005.

[7] NE Belyaeva, F-J Schmitt, R Steffen, VZ Paschenko, G Yu Riznichenko, YK Chemeris, G Renger, and AB Rubin. Ps ii model-based simulations of single turnover flash-induced transients of fluorescence yield monitored within the time domain of 100 ns-10 s on dark-adapted chlorella pyrenoidosa cells. *Photosynth. Res.*, 98(1-3):105–119, 2008.

[8] K Bernhardt and H Trissl. Theories for kinetics and yields of fluorescence and photochemistry: how, if at all, can different models of antenna organization be distinguished experimentally? *Biochim. Biophys. Acta., Bioenerg.*, 1409:125–42, 1999.

[9] N Betterle, M Ballottari, S Zorzan, S de Bianchi, S Cazzaniga, L Dall'Osto, T Morosinotto, and R Bassi. Light-induced dissociation of an antenna hetero-oligomer is needed for non-photochemical quenching induction. *J. Biol. Chem.*, 284:15255–15266, 2009.

[10] Robert E Blankenship. *Molecular Mechanisms of Photosynthesis*. Blackwell Publishing, 2002.

[11] T Brixner, T Mancal, IV Stiopkin, and GR Fleming. Phase-stabilized two-dimensional electronic spectroscopy. *J. Chem. Phys.*, 121:4221, 2004.

[12] T Brixner, J Stenger, HM Vaswani, M Cho, RE Blankenship, and GR Fleming. Two-dimensional spectroscopy of electronic couplings in photosynthesis. *Nature*, 434:625–628, 2005.

[13] K Broess, G Trinkunas, CD van der Weij-de Wit, JP Dekker, A van Hoek, and H van Amerongen. Excitation energy transfer and charge separation in photosystem ii membranes revisited. *Biophys. J.*, 91:3776–3786, 2006.

[14] K Broess, G Trinkunas, A van Hoek, R Croce, and H van Amerongen. Determination of the excitation migration time in photosystem ii - consequences for the membrane organization and charge separation parameters. *Biochim. Biophys. Acta, Bioenerg.*, 1777:404–409, 2008.

[15] EJ Brown, Q Zhang, and M Dantus. Femtosecond transient-grating techniques: Population and coherence dynamics involving ground and excited states. *J. Chem. Phys.*, 110:5772, 1999.

[16] C Buchel and W Kuhlbrandt. Structural differences in the inner part of photosystem ii between higher plants and cyanobacteria. *Photosynth. Res.*, 85:3–13, 2005.

[17] S Caffarri, K Broess, R Croce, and H van Amerongen. Excitation energy transfer and trapping in higher plant photosystem ii complexes with different antenna sizes. *Biophys. J.*, 100:2094–2103, 2011.

[18] S Caffarri, R Kouril, S Kereiche, EJ Boekema, and R Croce. Functional architecture of higher plant photosystem ii supercomplexes. *EMBO J.*, 28:3052–3063, 2009.

[19] TR Calhoun, NS Ginsberg, GS Schlau-Cohen, Y-C Cheng, M Ballottari, R Bassi, and GR Fleming. Quantum coherence enabled determination of the energy landscape in light-harvesting complex II. *J. Phys. Chem. B*, 113:16291–16295, 2009.

[20] M Cho. *Two-Dimensional Optical Spectroscopy*. CRC Press, 2009.

[21] M Cho, HM Vaswani, T Brixner, J Stenger, and GR Fleming. Exciton analysis in 2D electronic spectroscopy. *J. Phys. Chem. B*, 109:10542–10556, 2005.

[22] M Cho, JY Yu, T Joo, Y Nagasawa, SA Passino, and GR Fleming. The integrated photon echo and solvation dynamics. *J. Phys. Chem.*, 100:11944–11953, 1996.

[23] E Collini, CY Wong, KE Wilk, PMG Curmi, P Brumer, and GD Scholes. Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature*, 463:644–647, 2010.

[24] R Croce, MGMüller, S Caffarri, R Bassi, and AR Holzwarth. Energy transfer pathways in the minor antenna complex cp29 of photosystem ii: A femtosecond study of carotenoid to chlorophyll transfer on mutant and wt complexes. *Biophys. J.*, 84:2517–2532, 2003.

[25] R Croce and H van Amerongen. Light-harvesting and structural organization of photosystem ii: From individual complexes to thylakoid membrane. *J. Photochem. Photobiol. B*, 104(1-2):142–153, 2011.

[26] S de Bianchi, N Betterle, R Kouril, S Cazzaniga, E Boekema, R Bassi, and L Dall'Osto. Arabidopsis mutants deleted in the light-harvesting protein lhcb4 have a disrupted photosystem ii macrostructure and are defective in photoprotection. *Plant Cell*, 23:2659–79, 2011.

[27] S de Bianchi, L Dall'Osto, G Tognon, T Morosinotto, and R Bassi. Minor antenna proteins cp24 and cp26 affect the interactions between photosystem ii subunits and the electron transport rate in grana membranes of arabidopsis. *Plant Cell*, 20:1012–28, 2008.

[28] JP Dekker and EJ Boekema. Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta, Bioenerg.*, 1706(1-2):12–39, 2005.

[29] Y Dilworth, H Lee, and GR Fleming. Measuring electronic coupling in the reaction center of purple photosynthetic bacteria by two-color, three-pulse photon echo peak shift spectroscopy. *J. Phys. Chem. B*, 111:7449–7456, 2007.

[30] B Dugnol, C Fernández, G Galiano, and J Velasco. On a chirplet transform-based method applied to separating and counting wolf howls. *Signal Process.*, 88:1817–1826, 2008.

[31] GS Engel, TR Calhoun, EL Read, TK Ahn, T Mancal, Y-C Cheng, RE Blankenship, and GR Fleming. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature*, 446:782–786, 2007.

[32] CB Field. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*, 281:237–240, 1998.

[33] GR Fleming, GS Schlau-Cohen, K Amarnath, and J Zaks. Design principles of photosynthetic light-harvesting. *Faraday Discuss.*, 155:27–41, 2012.

[34] NS Ginsberg, Y-C Cheng, and GR Fleming. Two-Dimensional Electronic Spectroscopy of Molecular Aggregates. *Acc. Chem. Res.*, 42:1352–1363, 2009.

[35] NS Ginsberg, JA Davis, M Ballottari, Y-C Cheng, R Bassi, and GR Fleming. Solving structure in the CP29 light harvesting complex with polarization-phased 2D electronic spectroscopy. *PNAS*, 108:3848–3853, 2011.

[36] ML Groot, NP Pawlowicz, LJGW van Wilderen, J Breton, IHM van Stokkum, and R Van Grondelle. Initial electron donor and acceptor in isolated photosystem ii reaction centers identified with femtosecond mid-ir spectroscopy. *PNAS*, 102:13087–92, 2005.

[37] K Gundogdu, KW Stone, DB Turner, and KA Nelson. Multidimensional coherent spectroscopy made easy. *Chem. Phys.*, 341(1-3):89–94, 2007.

[38] AR Holzwarth. Data analysis of time-resolved measurements. In J. Amesz and AJ Hoff, editors, *Biophysical Techniques in Photosynthesis*, pages 75–92. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[39] AR Holzwarth, Y Miloslavina, M Nilkens, and P Jahns. Identification of two quenching sites active in the regulation of photosynthetic light-harvesting studied by time-resolved fluorescence. *Chem. Phys. Lett.*, 483(4-6):262–267, 2009.

[40] W Humphrey, A Dalke, and K Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[41] JD Hybl, AA Ferro, and DM Jonas. Two-dimensional Fourier transform electronic spectroscopy. *J. Chem. Phys.*, 115:6606, 2001.

[42] A Ishizaki, TR Calhoun, GS Schlau-Cohen, and GR Fleming. Quantum coherence and its interplay with protein environments in photosynthetic electronic energy transfer. *Phys. Chem. Chem. Phys.*, 12:7319, 2010.

[43] P Jahns and AR Holzwarth. The role of the xanthophyll cycle and of lutein in photoprotection of photosystem ii. *Biochim. Biophys. Acta, Bioenerg.*, 1817(1):182–193, 2012.

[44] D M Jonas. Two-Dimensional Femtosecond Spectroscopy. *Ann. Rev. Phys. Chem.*, 54:425–463, 2003.

[45] T Joo, Y Jia, JY Yu, MJ Lang, and GR Fleming. Third-order nonlinear time domain probes of solvation dynamics. *J. Chem. Phys.*, 104:6089–6108, 1996.

[46] XJ Jordanides, MJ Lang, XY Song, and GR Fleming. Solvation dynamics in protein environments studied by photon echo spectroscopy. *J. Phys. Chem. B*, 103:7995–8005, 1999.

[47] M Khalil, N Demirdoven, and A Tokmakoff. Coherent 2D IR spectroscopy: Molecular structure and dynamics in solution. *J. Phys. Chem. A*, 2003.

[48] M Khalil and A Tokmakoff. Signatures of vibrational interactions in coherent two-dimensional infrared spectroscopy. *Chem. Phys.*, 2001.

[49] J Kim, V M Huxter, C Curutchet, and GD Scholes. Measurement of Electron- Electron Interactions and Correlations Using Two-Dimensional Electronic Double-Quantum Coherence Spectroscopy. *J. Phys. Chem. A*, 113:12122–12133, 2009.

[50] J Kim, S Mukamel, and GD Scholes. Two-Dimensional Electronic Double-Quantum Coherence Spectroscopy. *Acc. Chem. Res.*, 42:1375–1384, 2009.

[51] CF Klingshirn. *Semiconductor Optics*. Springer, 2006.

[52] R Kouřil, J P Dekker, and EJ Boekema. Supramolecular organization of photosystem ii in green plants. *Biochim. Biophys. Acta, Bioenerg.*, 1817(1):2–12, 2012.

[53] H Lee, Y-C Cheng, and GR Fleming. Coherence dynamics in photosynthesis: protein protection of excitonic coherence. *Science*, 316:1462, 2007.

[54] H Lee, S Cheon, and M Cho. Chiroptical nature of two-exciton states of light-harvesting complex: Doubly resonant three-wave-mixing spectroscopy. *J. Chem. Phys.*, 132:225102, 2010.

[55] S Lemeille and J-D Rochaix. State transitions at the crossroad of thylakoid signalling pathways. *Photosynth. Res.*, 106(1):33–46, 2010.

[56] JM Linnanto and JEI Korppi-Tommola. Investigation on chlorosomal antenna geometries: tube, lamella and spiral-type self-aggregates. *Photosynth. Res.*, 96:227–245, 2008.

[57] ZF Liu, HC Yan, KB Wang, TY Kuang, JP Zhang, LL Gui, XM An, and WR Chang. Crystal structure of spinach major light-harvesting complex at 2.72 angstrom resolution. *Nature*, 428:287–292, 2004.

[58] MA Lomize, AL Lomize, ID Pogozheva, and HI Mosberg. Opm: orientations of proteins in membranes database. *Bioinformatics*, 22(5):623–625, 2006.

[59] S Mann and S Haykin. The chirplet transform: A generalization of Gabor's logon transform. *Vision Interface*, 91:205–212, 1991.

[60] S Mann and S Haykin. The chirplet transform: Physical considerations. *IEEE Trans. Signal Process.*, 43:2745–2761, 1995.

[61] A Marin, F Passarini, R Croce, and R Van Grondelle. Energy transfer pathways in the cp24 and cp26 antenna complexes of higher plant photosystem ii: A comparative study. *Biophys. J.*, 99:4056–4065, 2010.

[62] V May and O Kühn. *Charge and Energy Transfer Dynamics in Molecular Systems*. Wiley, 2011.

[63] A Melis. Photosystem-ii damage and repair cycle in chloroplasts: what modulates the rate of photodamage in vivo? *Trends Plant Sci.*, 4:130–135, 1999.

[64] MTW Milder, B Brüggemann, R van Grondelle, and JL Herek. Revisiting the optical properties of the FMO protein. *Photosynth. Res.*, pages 1–18, 2010.

[65] Y Miloslavina, M Szczepaniak, M G Müller, J Sander, M Nowaczyk, M Rögner, and AR Holzwarth. Charge separation kinetics in intact photosystem ii core particles is trap-limited. a picosecond fluorescence study. *Biochemistry*, 45:2436–42, 2006.

[66] J Minagawa. State transitionsthe molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochim. Biophys. Acta, Bioenerg.*, 1807(8):897–905, 2011.

[67] F Müh, M Madjet, and T Renger. Structure-based identification of energy sinks in plant light-harvesting complex ii. *J. Phys. Chem. B*, 114(42):13517–13535, 2010.

[68] F Müh, M Madjet, and T Renger. Structure-based simulation of linear optical spectra of the cp43 core antenna of photosystem ii. *Photosynth. Res.*, 111(1-2):87–101, 2012.

[69] F Müh and T Renger. Refined structure-based simulation of plant light-harvesting complex ii: Linear optical spectra of trimers and aggregates. *Biochim. Biophys. Acta, Bioenerg.*, 1817(8):1446–1460, 2012.

[70] S Mukamel. *Principles of Nonlinear Optical Spectroscopy*. Oxford University Press, USA, 1999.

[71] S Mukamel. Communications: Signatures of quasiparticle entanglement in multidimensional nonlinear optical spectroscopy of aggregates. *J. Chem. Phys.*, 132:241105, 2010.

[72] S Mukamel and D Abramavicius. Many-body approaches for simulating coherent nonlinear spectroscopies of electronic and vibrational excitons. *Chem. Rev.*, 104:2073–2098, 2004.

[73] S Mukamel, R Oszwałdowski, and L Yang. A coherent nonlinear optical signal induced by electron correlations. *J. Chem. Phys.*, 127:221105, 2007.

[74] S Mukamel, S Tretiak, T Wagersreiter, and V Chernyak. Electronic coherence and collective optical excitations of conjugated molecules. *Science*, 277:781, 1997.

[75] P Müller, X-P Li, and KK Niyogi. Non-photochemical quenching. a response to excess light energy. *Plant Phys.*, 125(4):1558–1566, 2001.

[76] A Nemeth, F Milota, T Mancal, T Pullerits, J Sperling, J Hauer, HF Kauffmann, and N Christensson. Double-quantum two-dimensional electronic spectroscopy of a three-level system: Experiments and simulations. *J. Chem. Phys.*, 133:094505, 2010.

[77] The equation for $V_{n,m}(t)$ shown in this paper has a different form from that given by Yang and Fleming because we have made use of the additional simplification provided by real valued eigenvectors.

[78] Since excitations on the periphery of $C_2S_2M_2$ are contained by the highest compartment of the chain, the population weighting of all effective rates of transfer ($\rho_i$, given in eq. (4.40)) will equal 1 and $t_{i-1\leftarrow i}^{periphery} = \tau_{i-1\leftarrow i}^{\text{eff}}$.

[79] V Novoderezhkin, A Marin, and R van Grondelle. Intra- and inter-monomeric transfers in the light harvesting lhcii complex: the redfield–förster picture. *Phys. Chem. Chem. Phys.*, 13:17093, 2011.

[80] V Novoderezhkin, M Wendling, and R Van Grondelle. Intra- and Interband Transfers in the B800B850 Antenna of Rhodospirillummolischianum: Redfield Theory Modeling of Polarized PumpProbe Kinetics. *J. Phys. Chem. B*, 107:11534–11548, 2003.

[81] VI Novoderezhkin, MA Palacios, H van Amerongen, and R van Grondelle. Excitation dynamics in the lhcii complex of higher plants: Modeling based on the 2.72 å crystal structure. *J. Phys. Chem. B*, 109(20):10493–10504, 2005.

[82] VI Novoderezhkin, E Romero, JP Dekker, and R van Grondelle. Multiple charge-separation pathways in photosystem ii: Modeling of transient absorption kinetics. *ChemPhysChem*, 12:681–688, 2011.

[83] VI Novoderezhkin and R van Grondelle. Physical origins and models of energy transfer in photosynthetic light-harvesting. *Phys. Chem. Chem. Phys.*, 12:7352–7365, 2010.

[84] AV Pakoulev, MA Rickard, NA Mathew, KM Kornau, and JC Wright. Frequency-Domain Time-Resolved Four Wave Mixing Spectroscopy of Vibrational Coherence Transfer with Single-Color Excitation. *J. Phys. Chem. A*, 112:6320–6329, 2008.

[85] MA Palacios, FL de Weerd, JA Ihalainen, R van Grondelle, and H van Amerongen. Superradiance and exciton (de)localization in light-harvesting complex ii from green plants? *J. Phys. Chem. B.*, 106:5782–5787, 2002.

[86] X Pan, M Li, T Wan, L Wang, C Jia, Z Hou, X Zhao, J Zhang, and W Chang. Structural insights into energy regulation of light-harvesting complex cp29 from spinach. *Nat. Struct. Mol. Biol.*, 18:309–315, 2011.

[87] G Panitchayangkoon, D Hayes, KA Fransted, JR Caram, E Harel, J Wen, RE Blankenship, and GS Engel. Long-lived quantum coherence in photosynthetic complexes at physiological temperature. *PNAS*, 107:12766–12770, 2010.

[88] G. C. Papageorgiu and Govindjee, editors. *Chlorophyll a Fluorescence: A Signature of Photosynthesis*, volume 19 of *Advances in Photosynthesis and Respiration*. Springer, Dordrecht, The Netherlands, 2004.

[89] S Park, MK Sener, DY Lu, and K Schulten. Reaction paths based on mean first-passage times. *J. Chem. Phys.*, 119:1313–1319, 2003.

[90] D Parkinson, H Lee, and G Fleming. Measuring electronic coupling in the reaction center of purple photosynthetic bacteria by two-color, three-pulse photon echo peak shift spectroscopy. *J. Phys. Chem. B*, 2007.

[91] F Passarini, E Wientjes, R Hienerwadel, and R Croce. Molecular basis of light harvesting and photoprotection in cp24: unique features of the most recent antenna complex. *J. Biol. Chem.*, 284:29536–46, 2009.

[92] G Raszewski, BA Diner, E Schlodder, and T Renger. Spectroscopic properties of reaction center pigments in photosystem ii core complexes: revision of the multimer model. *Biophys. J.*, 95:105–19, 2008.

[93] G Raszewski and T Renger. Light harvesting in photosystem ii core complexes is limited by the transfer to the trap: Can the core complex turn into a photoprotective mode? *JACS*, 130:4431–4446, 2008.

[94] G Raszewski, W Saenger, and T Renger. Theory of optical spectra of photosystem ii reaction centers: location of the triplet state and the identity of the primary electron donor. *Biophys. J.*, 88:986–98, 2005.

[95] T Renger, M E Madjet, A Knorr, and F Müh. How the molecular structure determines the flow of excitation energy in plant light-harvesting complex ii. *J. Plant Physiol.*, 168:1497–1509, 2011.

[96] T Renger and E Schlodder. Primary photophysical processes in photosystem ii: Bridging the gap between crystal structure and optical spectra. *ChemPhysChem*, 11:1141–1153, 2010.

[97] KH Rhee, EP Morris, J Barber, and W Kühlbrandt. Three-dimensional structure of the plant photosystem ii reaction centre at 8 å resolution. *Nature*, 396:283–286, 1998.

[98] AV Ruban, MP Johnson, and CDP Duffy. The photoprotective molecular switch in the photosystem ii antenna. *Biochim. Biophys. Acta, Bioenerg.*, 1817:167–181, 2012.

[99] S Savikhin, Y Zhu, S Lin, RE Blankenship, and WS Struve. Femtosecond spectroscopy of chlorosome antennas from the green photosynthetic bacterium Chloroflexus aurantiacus. *J. Phys. Chem.*, 98:10322–10334, 1994.

[100] RD Schaller, VM Agranovich, and VI Klimov. High-efficiency carrier multiplication through direct photogeneration of multi-excitons via virtual single-exciton states. *Nat. Phys.*, 1:189–194, 2005.

[101] GH Schatz, H Brock, and AR Holzwarth. Kinetic and energetic model for the primary processes in photosystem ii. *Biophys. J.*, 54:397–405, 1988.

[102] GS Schlau-Cohen, TR Calhoun, NS Ginsberg, EL Read, M Ballottari, R Bassi, R van Grondelle, and GR Fleming. Pathways of energy flow in lhcii from two-dimensional electronic spectroscopy. *J. Phys. Chem. B*, 113:15352–15363, 2009.

[103] AR Schneider and PL Geissler. Coexistence between fluid and crystalline phases of proteins in photosynthetic membranes. *arXiv preprint*, page 1302.6323, 2013.

[104] MK Sener, S Park, DY Lu, A Damjanovic, T Ritz, P Fromme, and K Schulten. Excitation migration in trimeric cyanobacterial photosystem i. *J. Chem. Phys.*, 120:11183–11195, 2004.

[105] W-J Shi, J Barber, and Y Zhao. Role of Formation of Statistical Aggregates in Chlorophyll Fluorescence Concentration Quenching. *J. Phys. Chem. B*, 117:3976–3982, 2013.

[106] I Stiopkin, T Brixner, M Yang, and GR Fleming. Heterogeneous Exciton Dynamics Revealed by Two-Dimensional Optical Spectroscopy. *J. Phys. Chem. B*, 110:20032–20037, 2006.

[107] KW Stone, K Gundogdu, DB Turner, X Li, ST Cundiff, and KA Nelson. Two-quantum 2D FT electronic spectroscopy of biexcitons in GaAs quantum wells. *Science*, 324:1169, 2009.

[108] L Tian, S Farooq, and H van Amerongen. Probing the picosecond kinetics of photosystem ii core complex *in vivo*. *Phys. Chem. Chem. Phys.*, 15:3146–3154, 2013.

[109] A Tokmakoff. Two-dimensional line shapes derived from coherent third-order nonlinear spectroscopy. *J. Phys. Chem. A*, 2000.

[110] DE Tronrud, J Wen, L Gay, and RE Blankenship. The structural basis for the difference in absorbance spectra for the FMO antenna protein from various green sulfur bacteria. *Photosynth. Res.*, 100:79–87, 2009.

[111] Y Umena, K Kawakami, J-R Shen, and N Kamiya. Crystal structure of oxygen-evolving photosystem ii at a resolution of 1.9 å. *Nature*, 473:55–60, 2011.

[112] L Valkunas, J Chmeliov, G Trinkunas, CDP Duffy, R van Grondelle, and AV Ruban. Excitation migration, quenching, and regulation of photosynthetic light harvesting in photosystem ii. *J. Phys. Chem. B*, 115:9252–9260, 2011.

[113] H van Amerongen, L Valkunas, and R van Grondelle. *Photosynthetic excitons*. World Scientific Pub Co Inc, 2000.

[114] CD van der Weij-de Wit, JP Dekker, R van Grondelle, and IHM van Stokkum. Charge separation is virtually irreversible in photosystem ii core complexes with oxidized primary quinone acceptor. *J. Phys. Chem. A*, 115:3947–3956, 2011.

[115] H van Oort, M Alberts, S de Bianchi, L Dall'Osto, R Bassi, G Trinkunas, R Croce, and H van Amerongen. Effect of antenna-depletion in photosystern ii on excitation energy transfer in arabidopsis thaliana. *Biophys. J.*, 98:922–931, 2010.

[116] IHM van Stokkum, DS Larsen, and R van Grondelle. Global and target analysis of time-resolved spectra. *Biochim. Biophys. Acta, Bioenerg.*, 1657:82–104, 2004.

[117] JC Vaughan, T Hornung, T Feurer, and KA Nelson. Diffraction-based femtosecond pulse shaping with a two-dimensional spatial light modulator. *Opt. Lett.*, 30:323–325, 2005.

[118] J Wen, H Zhang, ML Gross, and RE Blankenship. Native Electrospray Mass Spectrometry Reveals the Nature and Stoichiometry of Pigments in the FMO Photosynthetic Antenna Protein. *Biochemistry*, 50:3502–3511, 2011.

[119] JM Womick, SA Miller, and AM Moran. Toward the origin of exciton electronic structure in phycobiliproteins. *J. Chem. Phys.*, 133:024507, 2010.

[120] JM Womick and AM Moran. Vibronic Enhancement of Exciton Sizes and Energy Transport in Photosynthetic Complexes. *J. Phys. Chem. B*, page 4843, 2011.

[121] A E Yakushevska, W Keegstra, EJ Boekema, J P Dekker, J Andersson, S Jansson, A V Ruban, and P Horton. The structure of photosystem ii in arabidopsis: localization of the cp26 and cp29 antenna complexes. *Biochemistry*, 42:608–13, 2003.

[122] AE Yakushevska, PE Jensen, W Keegstra, H van Roon, HV Scheller, EJ Boekema, and J P Dekker. Supermolecular organization of photosystem ii and its associated light-harvesting antenna in arabidopsis thaliana. *Eur. J. Biochem.*, 268:6020–8, 2001.

[123] M Yang, A Damjanovic, HM Vaswani, and GR Fleming. Energy transfer in photosystem i of cyanobacteria *Synechococcus elongatus*: Model study with structure-based semi-empirical hamiltonian and experimental spectral density. *Biophys. J.*, 85:140–158, 2003.

[124] M Yang and GR Fleming. Influence of phonons on exciton transfer dynamics: comparison of the redfield, förster, and modified redfield equations. *Chem. Phys.*, 282(1):163–180, 2002.

[125] M Yang and GR Fleming. Construction of kinetic domains in energy trapping processes and application to a photosynthetic light harvesting complex. *J. Chem. Phys.*, 119:5614–5622, 2003.

[126] XG Zhu, Govindjee, NR Baker, E deSturler, DR Ort, and SP Long. Chlorophyll a fluorescence induction kinetics in leaves predicted from a model describing each discrete step of excitation energy and electron transfer associated with photosystem ii. *Planta*, 223:114–133, 2005.

[127] DD Zigmantas, EL Read, T Mancal, T Brixner, AT Gardiner, RJ Cogdell, and GR Fleming. Two-dimensional electronic spectroscopy of the B800-B820 light-harvesting complex. *PNAS*, 103:12672–12677, 2006.

# Appendix A

# Calculation of the nonlinear response functions

The third-order nonlinear time-domain response functions for the excitonic systems used in this work were calculated according to Ref. [20].

$$R_1(t_3,t_2,t_1) = \sum_{abc} \mu_{gc}\mu_{cb}\mu_{ba}\mu_{ag} \exp(-i\omega_{ab}t_3 - i\omega_{ac}t_2 - i\omega_{ag}t_1) F_1(t_3,t_2,t_1) \tag{A.1}$$

$$R_2(t_3,t_2,t_1) = \sum_{abc} \mu_{gc}\mu_{cb}\mu_{ba}\mu_{ag} \exp(-i\omega_{ab}t_3 - i\omega_{ac}t_2 - i\omega_{gc}t_1) F_2(t_3,t_2,t_1) \tag{A.2}$$

$$R_3(t_3,t_2,t_1) = \sum_{abc} \mu_{gc}\mu_{cb}\mu_{ba}\mu_{ag} \exp(-i\omega_{ab}t_3 - i\omega_{gb}t_2 - i\omega_{gc}t_1) F_3(t_3,t_2,t_1) \tag{A.3}$$

$$R_4(t_3,t_2,t_1) = \sum_{abc} \mu_{gc}\mu_{cb}\mu_{ba}\mu_{ag} \exp(-i\omega_{cg}t_3 - i\omega_{bg}t_2 - i\omega_{ag}t_1) F_4(t_3,t_2,t_1) \tag{A.4}$$

The subscript $g$ indicates the ground state, and the subscripts $a$ and $c$ refer to states in one-exciton manifold. For pathways that only go through the one-exciton manifold the subscript $b$ refers to the ground state. For pathways that involve the two-exciton manifold the subscript $b$ refers to the two-exciton manifold. For any two states $j$ and $k$ the transition dipole matrix element is $\mu_{jk}$ and the coherence frequency is $\omega_{jk} = \hbar^{-1}(E_j - E_k)$.

The functions $F_j(t_3,t_2,t_3)$ are the nonlinear lineshape functions and are determined by fluctuations of the electronic states due to their surrounding dissipative baths. Under the assumption of Gaussian noise for the fluctuations (second cumulant approximation) these functions are:

$$\begin{aligned} F_1(t_3,t_2,t_1) = &\exp\{-g^*_{cc}(t_2) - g^*_{bb}(t_3) - g_{aa}(t_1+t_2+t_3) - g^*_{cb}(t_2+t_3) \\ &+ g^*_{cb}(t_2) + g^*_{cb}(t_3)) + g_{ca}(t_1+t_2) - g_{ca}(t_1) + g^*_{ca}(t_2+t_3) \\ &- g^*_{ca}(t_3) + g_{ba}(t_1+t_2+t_3) - g_{ba}(t_1+t_2) + g^*_{ba}(t_3)\} \end{aligned} \tag{A.5}$$

$$\begin{aligned} F_2(t_3,t_2,t_1) = &\exp\{-g^*_{cc}(t_1+t_2) - g^*_{bb}(t_3) - g_{aa}(t_2+t_3) - g^*_{cb}(t_1+t_2+t_3) \\ &+ g^*_{cb}(t_1+t_2) + g^*_{cb}(t_3)) + g_{ca}(t_2) + g^*_{ca}(t_1+t_2+t_3) \\ &- g^*_{ca}(t_1) - g^*_{ca}(t_3) + g_{ba}(t_2+t_3) - g_{ba}(t_2) + g^*_{ba}(t_3)\} \end{aligned} \tag{A.6}$$

$$\begin{aligned} F_3(t_3,t_2,t_1) = &\exp\{-g^*_{cc}(t_1) - g^*_{bb}(t_2+t_3) - g_{aa}(t_3) - g^*_{cb}(t_1+t_2+t_3) \\ &+ g^*_{cb}(t_1) + g^*_{cb}(t_2+t_3)) + g^*_{ca}(t_1+t_2+t_3) - g^*_{ca}(t_1+t_2) \\ &+ g^*_{ca}(t_2+t_3) + g^*_{ca}(t_2) + g_{ba}(t_3) + g^*_{ba}(t_2+t_3) + g^*_{ba}(t_2)\} \end{aligned} \tag{A.7}$$

$$\begin{aligned} F_4(t_3,t_2,t_1) = &\exp\{-g_{cc}(t_3) - g_{bb}(t_2) - g_{aa}(t_1) - g_{cb}(t_2+t_3) \\ &+ g_{cb}(t_2) + g_{cb}(t_3)) - g_{ca}(t_1+t_2+t_3) + g_{ca}(t_1+t_2) \\ &+ g_{ca}(t_2+t_3) - g_{ca}(t_2) - g_{ba}(t_1+t_2) + g_{ba}(t_1) + g_{ba}(t_2)\} \end{aligned} \tag{A.8}$$

where the lineshape function $g_{ba}(t)$ is a property of a transition between states $a$ and $b$.

$$g_{ba}(t) = \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 C_{ba}(\tau_2) \tag{A.9}$$

Here $C_{ba}(t)$ is the time-correlation function (TCF) of the bath-induced fluctuations of energies of states $a$ and $b$.

$$C_{ba}(t) = \frac{1}{\hbar^2}\langle U_{bg}(t) U_{ag}(0)\rangle_B \tag{A.10}$$

where $U_{jg}(t)$ is the fluctuating part of the energy of state $j$ relative to the ground state $g$, and $\langle ...\rangle_B$ refers to thermal averaging over all bath coordinates.

First, consider a single pigment coupled to a bath of quantum harmonic oscillators where the strength of coupling to a bath mode with frequency $\omega$ is given by spectral density $\rho(\omega)$. The single pigment energy fluctuation TCF is:

$$C(t) = \int_0^\infty d\omega \rho(\omega) \coth(\hbar\omega\beta/2)\omega^2 \cos\omega t - i\int_0^\infty d\omega\rho(\omega)\omega^2 \sin\omega t \tag{A.11}$$

The corresponding line-shape function is:

$$g(t) = \int_0^\infty d\omega \rho(\omega) \coth(\hbar\omega\beta/2)(1-\cos\omega t) + i\int_0^\infty d\omega\rho(\omega)(\sin\omega t - \omega t) \tag{A.12}$$

where the bath temperature appears as $\beta = 1/k_B T$. For excitonic transitions, we assume that the energy fluctuation of the pigments contributing to the excitons are all identical and independent of each other (independent and identical bath approximations). Then the TCFs for the energy fluctuations of one-exciton states can be related to the TCF of a single pigment through:

$$\langle \delta\Omega_j(t)\delta\Omega_k(0)\rangle_B = C(t)\sum_m U_{mj}^2 U_{mk}^2 \tag{A.13}$$

where $\hbar\Omega_j$ is the energy of the $j^{th}$ single exciton states and $U_{mj}$ is the coefficient of contribution of the $m^{th}$ pigment in the $j^{th}$ exciton. Similarly, we write the energy fluctuation TCF for the two-exciton states as:

$$\langle \delta W_j(t)\delta W_k(0)\rangle_B = C(t)\sum_{m=1}^{N-1}\sum_{n=m+1}^{N}(v_{mn}^{(j)})^2\{P_m^{(k)} + P_n^{(k)}\} \tag{A.14}$$

where the auxiliary quantity $P_m^{(k)}$ is:

$$P_m^{(k)} = \sum_{j=1}^{m-1}(v_{jm}^{(k)})^2 + \sum_{j=m+1}^{N}(v_{mj}^{(k)})^2 \tag{A.15}$$

The cross-correlation functions for the one- and two-exciton states is given by:

$$\langle \delta\Omega_j(t)\delta W_k(0)\rangle_B = C(t)\sum_{m=1}^{N} U_{mj}^2 P_m^{(k)} \tag{A.16}$$

Equations (A.5) - (A.16) relate the energy fluctuations of a single pigment to the lineshape of nonlinear response functions.

For energy fluctuations of a single pigment, we assume a bath of quantum harmonic oscillators with an ohmic spectral density:

$$\rho(\omega) = (\frac{\lambda}{\hbar})\frac{1}{\omega\omega_c}\exp(-\frac{\omega}{\omega_c}) \tag{A.17}$$

where the re-organization energy $\lambda$ is the integrated strength of coupling to the bath, and the cutoff frequency $\omega_c$ determines the time-scale of fluctuations. The very slow (inhomogeneous) fluctuations of the energy levels can not be conveniently captured by the above spectral density. To account for the inhomogeneous component of the line-shape, we have added a Gaussian decay component to the lineshape functions of a single pigment (Eq. A.12).

For FMO, the directions of transition dipole moments of pigments were found from the available structure of FMO in the Protein Data Bank and published in Ref. [110]. This study shows that in some sample preparations of FMO, a loosely bound eighth chlorophyll may be incorporated with the FMO Protein. We have used the previously established FMO model with seven pigments as we were not sure whether the eighth pigment has been retained in

our sample. The magnitudes of all pigment transition dipole moments were assumed to be the same. The coupled pigment Hamiltonian used for FMO is from Ref. [21].

$$\hat{H} = \begin{pmatrix} 280 & -106 & 8 & -5 & 6 & -8 & -4 \\ -106 & 420 & 28 & 6 & 2 & 13 & 1 \\ 8 & 28 & 0 & -62 & -1 & -9 & 17 \\ -5 & 6 & -62 & 175 & -70 & -19 & -57 \\ 6 & 2 & -1 & -70 & 320 & 65 & -2 \\ -8 & 13 & -9 & -19 & 65 & 360 & 32 \\ -4 & 1 & 17 & -57 & -2 & 32 & 260 \end{pmatrix} \tag{A.18}$$

Here all diagonal energies are relative to 12130 $\text{cm}^{-1}$ and the coupling strength of pigments 5 and 6 has been adjusted from 40 $\text{cm}^{-1}$ to 65 $\text{cm}^{-1}$. We used the spectral density parameters $\omega_c = 50$ $\text{cm}^{-1}$ and $\lambda = 35$ $\text{cm}^{-1}$ . The inhomogeneous distribution of energies of pigments 1-7 were assumed to be 200, 200, 5, 150, 120, 150, 150 50 $\text{cm}^{-1}$, respectively. A calculated linear absorption spectrum based on these parameters showed reasonable match to the experimental absorption spectrum. For the model systems in Section 2.3.1, the pigment dipole directions were chosen to be in the same plane, with $45^o$ angle with respect to each other. In all simulations, the highly oscillatory components of the response functions were slowed down by subtracting a a constant reference frequency $\omega_{ref}$ ($2\omega_{ref}$) from all one-exciton (two-exciton) state energies for better numerical sampling.

# Appendix B

# Properties of Chirplet Transform

## B.1 Definition

Having established the relationship between the chirplet transform and a third-order spectroscopic signal, we will now explore the general properties of a chirplet transform using a general 1D oscillatory signal that is either a pure tone (i.e. develops in phase at a single central frequency) or a chirped tone (i.e. the central frequency of the signal changes as a function of time). Pictorially, the Chirplet Transform is the convolution of an oscillatory function F($\tau$) with a complex-valued envelope in t. Shown below is a figure plotting a sinusoidal F($\tau$) function (black line) that will be convolved with a pulse described by a gaussian envelope (red line), linear phase (blue line), and quadratic phase (magenta line).

This is mathematically defined as:

$$C(t,\omega,\sigma,\beta) = \int_{-\infty}^{\infty} e^{-\frac{(t-\tau)^2}{2\sigma^2}} e^{i\omega(t-\tau)} e^{-i\beta(t-\tau)^2} F(\tau)\mathrm{d}\tau \tag{B.1}$$

where $t$ is time delay of the envelope, $\sigma$ is the width of the envelope, $\omega$ is the central frequency of the envelope, and $\beta$ is the time-domain chirp of the envelope. Note, that all of the "chirping" is time-domain and not frequency domain - so this is not a frequency conserving action. The extension of a 1D signal into a 4D space generates an over-complete representation that is unwieldy to operate in. In this section we will explore 1D and 2D sub-spaces that can extract information from the oscillating signal.

## B.2 1D Subspaces

Two 1D subspaces of the eq. (1) above have very nice (and familiar) properties: the $t$-Space and $\omega$-Space. These spaces are defined by taking the impulsive and frequency domain limit of the convolution envelope respectively.

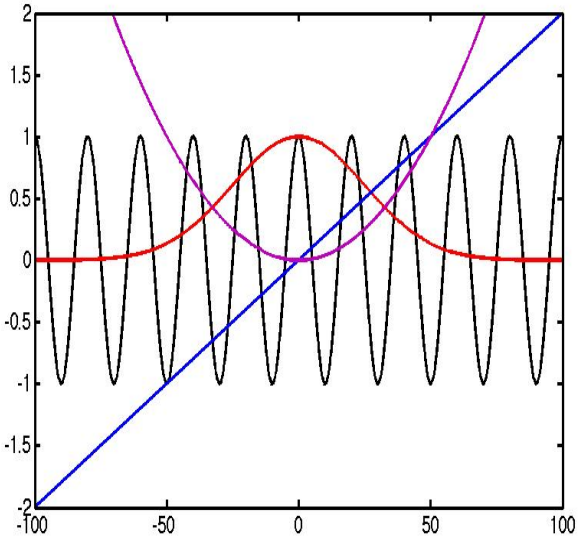



Figure B.1: A sinusoidal F(t) function (black line), which is convolved with a pulse described by a gaussian envelope (red line), linear phase (blue line), and quadratic phase (magenta line).

### B.2.1 $t$ Space

In the impulsive limit the envelope is characterized by zero width, no chirp, and a constant central frequency. This returns exactly the original function, as shown in eq. (B.2).

$$C(t, \omega = \omega_0, \sigma = 0, \beta = 0) = \int \delta(t - \tau) e^{-i\omega_0(t-\tau)} F(t) d\tau = F(t) \tag{B.2}$$

### B.2.2 $\omega$ Space

The frequency domain limit is characterized by an infinitely wide envelope (or at least much larger than the characteristic width of F(t)), no offset in $t$, and no chirp. This 1D space is exactly the Fourier Transform of the original function F(t).

$$C(t = 0, \omega, \sigma \to \infty, \beta = 0) = \int e^{-i\omega\tau} F(\tau) = \tilde{F}(\omega) dt \tag{B.3}$$

## B.3 2D Subspaces

Passing out of these lines and into a plane, we can consider interesting 2D subspaces of the Chirplet Transform. Two stand out particularly - the $(t, \omega)$ and $(t, \beta)$ spaces.

### B.3.1 $(t, \omega)$ Space

This plane will reproduce the Short-Time Fourier Transform (STFT) of the function $F(t\tau)$, also known as the time-frequency representation. Again, chirp will be set to zero, but sigma will be any constant value.

$$C(t, \omega, \sigma = \sigma_0, \beta = 0) = \int e^{-\frac{(t-\tau)^2}{2\sigma_0^2}} e^{-i\omega(t-\tau)} F(\tau) \mathrm{d}\tau \tag{B.4}$$

The advantage of this space is the ability to resolve how a non-stationary signal is developing. For example, Figure B.2 shows the absolute value results of a pure tone signal compared to a chirped signal in the $(\tau, \omega)$ plane. This has a relatively simple physical interpretation - the presence of amplitude corresponds to the presence of a frequency $\omega$ at a time $t$. Hence, the chirped signal shows a frequency development as a function of time.

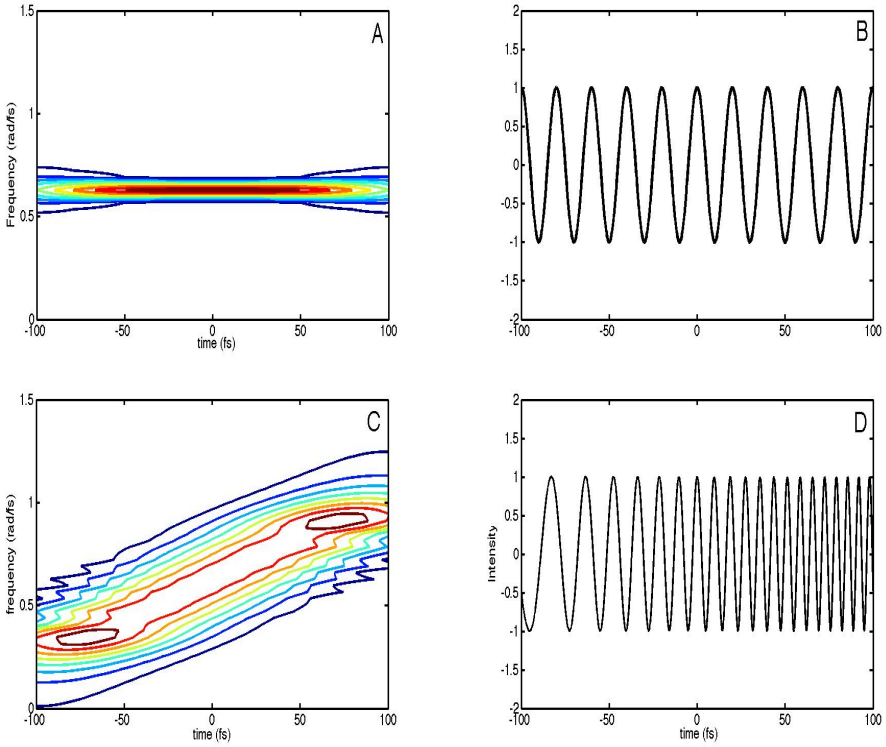



Figure B.2: (A) the TF space representation of a pure tone, (B) the T space signal, (C) the TF space representation of a chirped tone (the extra structure is the result of finite window), and (D) the T space representation of the chirp tone.

### B.3.2 $(t, \beta)$ Space

This plane corresponds to selecting a central frequency ($\omega_0$) for the envelope (presumably one that puts the majority of the envelope in resonance with the oscillatory signal) and selecting an envelope width ($\sigma_0$).

$$C(t, \omega_0, \sigma_0, \beta) = \int e^{-\frac{(t-\tau)^2}{2\sigma_0^2}} e^{-i\omega_0(t-\tau)} e^{-i\beta(t-\tau)^2} F(\tau) \mathrm{d}\tau \tag{B.5}$$

The $(t, \beta)$ plane, shows a peak on the time axis everywhere the central frequency of the pulse ($\omega_0$ matches the frequency of the transformed signal. For a pure tone (such as that shown in Figure B.3B) and envelope with a matching central frequency, the transform will have a peak everywhere along the $\tau$-axis. The decay shown in Fig. 3 (A) is an artifact of the finite integration window. If the transformed signal, however, is a chirped tone then even if it is un-windowed the transformed signal will decay along the time axis as the underlying frequency of the signal shifts away from the central frequency of the envelope. The amplitude of the transformed signal ($C(t, \omega_0, \sigma_0, \beta)$) will also increase as the chirp of the envelope ($\beta$) matches the chirp of the signal. So we should expect to see a strong peak centered at $(t_0, \beta_0)$ - the point in time where the central frequencies agree ($t_0$) and the chirp agrees ($\beta_0$). This intuition is numerically verified in Figure B.3C, which shows the amplitude of the $(t, \beta)$-plane for a chirped tone signal (shown in Figure B.3D).

It is interesting to note that there are two different kinds of envelopes present in calculating this result - an envelope present in the signal function $F(\tau)$ and the amplitude envelope associated with the chirplet. To understand the role of these two envelopes, Fig. 4 represents the chirplet transform of a chirped signal $F(\tau)$ with various pulse windows and signal windows. To understand the change in the signal shape present in figure 4 - we should notice that both the D & G spectra are approaching the limit a straight line. The straight line represents a pulse sufficiently narrow, that no matter what chirp is applied the local frequency does not substantially vary over the within the window of the pulse.

On the other hand, the E & H spectra both show substantial structure and asymmetry - though spectrum H has a more asymmetric than spectrum E. The presence of this structure in both features suggests that the pulse chirp does vary substantially over the time period of the pulse. The asymmetry is the result an interesting effect that comes out of overlap of spectral features between two chirped gaussians. To see this effect, look at Figure 5, image (a) & (b) which show two pulses of equal temporal width separated by a constant time. In image (a), the two pulses are oppositely chirped and as such still have overlapping frequencies. On the other hand, when the two pulses share the same chirp (image (b)), they do not have shared, overlapping frequencies. This gives rise to broader time range for non-zero signal when the response function and the electric field have oppositely signed chirp than in the case of shared sign of chirp.

The magnitude of the asymmetry is determined by the relative widths of the pulse and the response function. In the non-chirped case, the width is set predominately by the widest of the two envelopes. In the case where the pulse chirp being scanned is associated with

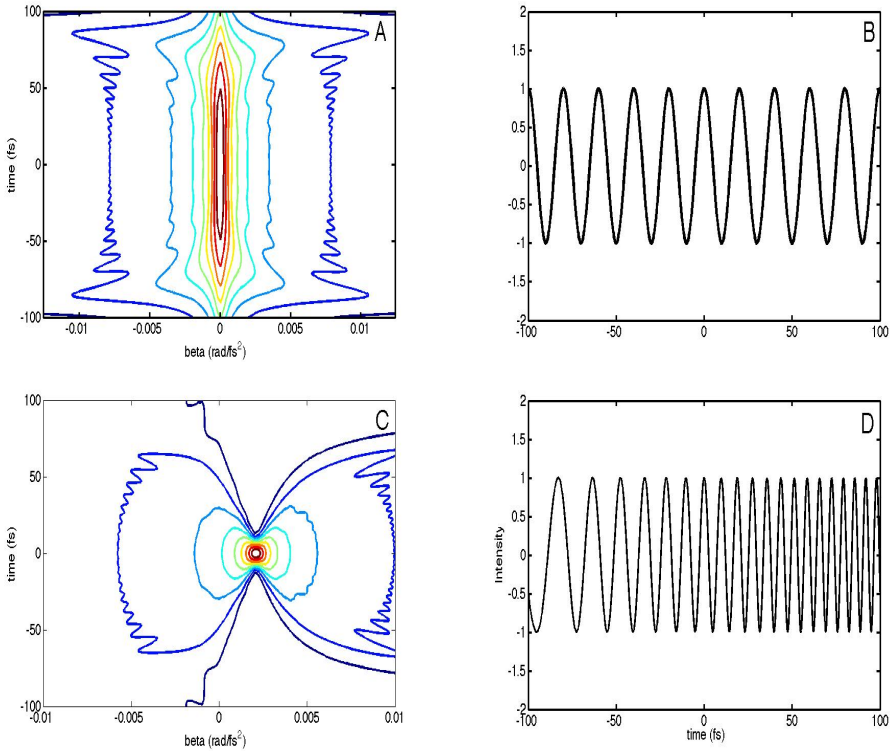



Figure B.3: (A) the $(t, \beta)$-space representation of a pure tone, (B) the $t$-space signal, (C) the $(t, \beta)$-space representation of a chirped tone (the extra structure is the result of finite window), and (D) the t-space chirped signal.

a field narrower than the envelope of the response function, then the asymmetry will be a smaller effect. On the other hand, if the pulse envelope is as wide or wider then the response function, the asymmetry will be more noticeable.

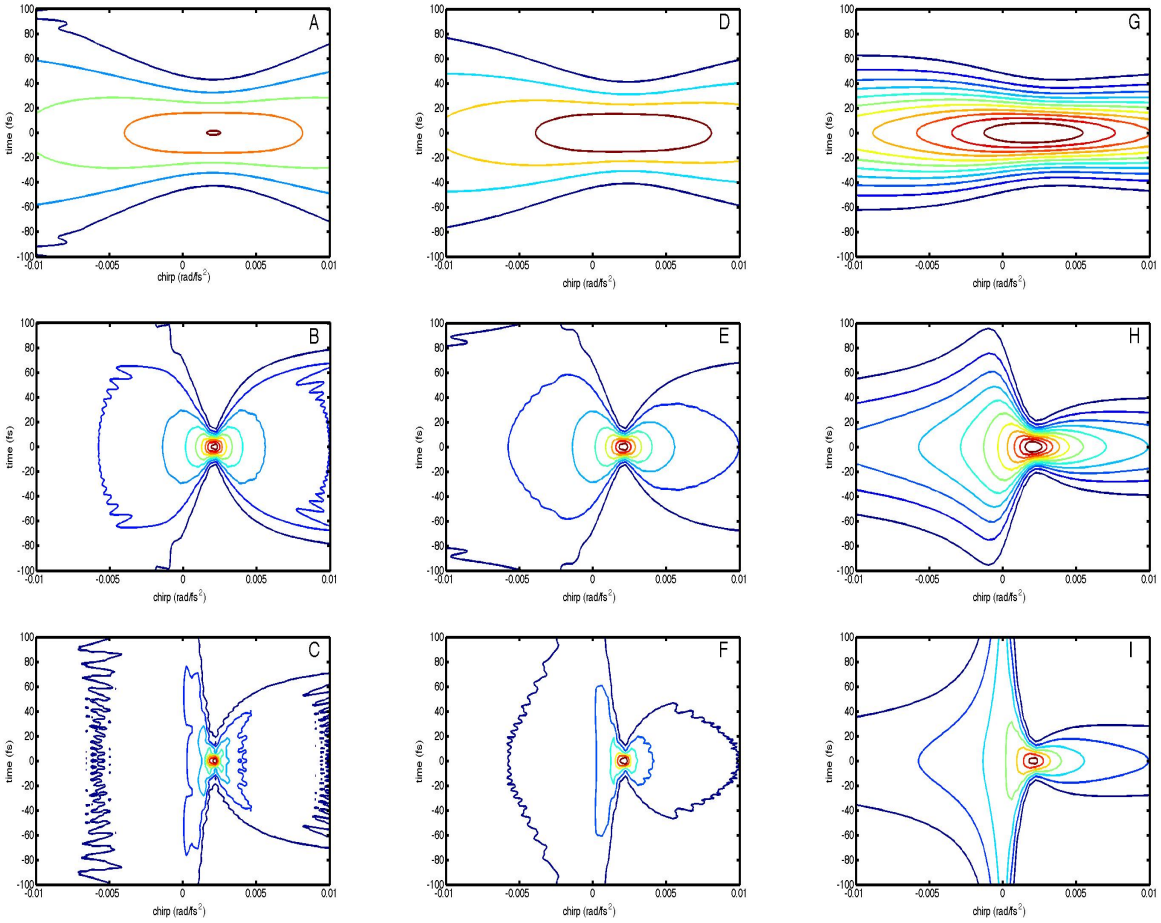



Figure B.4: Proceeding from left to right in a row, the response function changes from unwindowed at the left to tightly windowed at the right. Proceeding from top to bottom in a column, the amplitude envelope of a pulse changes from tightly windowed at the top to almost unwindowed at the bottom. This figure explores how the chirplet transform behaves as the two window widths are varied.

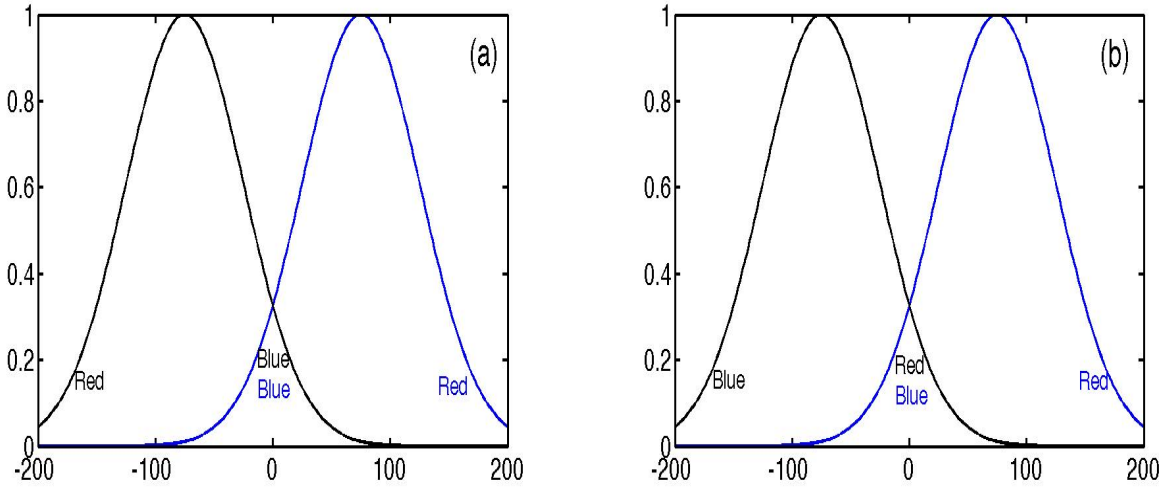



Figure B.5: Chirped function overlap is shown to explain the asymmetry between the positive and negative chirp directions in a $(t, \beta)$-space chirplet transform.

# Appendix C

# Photosystem II Parameters and Error Analysis

## C.1 Justification of protein structures used

In plants, the photosystem II supercomplex is composed of the reaction center (RC) core surrounded by the Lhcb proteins[4, 25]. As shown in Figure 1 of the main text, the core is composed of the two copies each of the proteins RC (composed of D1 and D2), CP43, and CP47, which are all attached to the oxygen evolving machinery. It is difficult to isolate pure, oxygen evolving preparations of the RC core from plants because the oxygen evolving subunits are less stable in plants, and, as a result, there are no high resolution structures of the RC core from plants[25]. The PSII core from cyanobacteria has generally been used a proxy for the PSII core from plants due to the high similarity in the two protein structures[16, 97]. We use a 1.9 Å resolution crystal structure from *Thermosynechococcus vulcanus*, a thermophilic cyanobacterium[111]. The parameters calculated in this paper may be slightly altered if the plant RC core structure is resolved and used instead of the current cyanobacteria structure.

The periphery of the PSII supercomplex is composed of light harvesting complex II (LHCII) trimers. The LHCII trimer structure from plants has been resolved to 2.72 $\mathring{A}$ resolution[57] and is used for the calculations in this paper. Between the LHCIIs and the RC core reside the minor LHCs, CP24, CP26, and CP29. No structure has been solved for CP24 and CP26, and no Hamiltonian has been calculated using the structure of CP29. We have substituted an LHCII monomer in place of the three minor LHCs. This substitution is justified for the following reasons. The sequences of the minor LHCs are homologous to those of LHCII and contain nearly all of the binding sites for Chls that LHCII has[91]. Energy transfer within the minor complexes has been shown to have roughly the same dynamics as that in LHCII[61, 24, 79]. The CP29 structure shows a high structural homology with the LHCII monomer, though it is missing Chl 605[86]. Trimers of CP26 has been shown to be able to replace LHCII in the protein complex composed of a LHCII trimer, CP24, and CP29 proteins[9]. It has also been suggested that the Lhcb1 protein (one of the proteins in the

| Protein | Molecule | TDM (Debye) |
|---|---|---|
| Reaction Center, CP47, CP43 | Chl A | 4.4 |
| Reaction Center | Pheophytin | 3.5 |
| LHCII, Minor Complexes | Chl A | 4.0 |
| LHCII, Minor Complexes | Chl B | 3.4 |

Table C.1: Transition dipole moment (TDM) magnitudes used to determine Coulombic coupling between pigments associated with different proteins[79, 93].

| Protein | Molecule | $\sigma$ ($cm^{-1}$) |
|---|---|---|
| Reaction Center | Chl A & Pheophytin | 200 |
| Reaction Center | $Chlz_{D1}$ | 120 |
| CP47 | Chl A | 180 |
| CP43 | Chl A | 180 |
| LHCII, Minor Complexes | Chl A | 80 |
| LHCII, Minor Complexes | Chl B | 96 |

Table C.2: Inhomogeneous width of site energies for different molecules in PSII supercomplex [79, 93].

LHCII trimer) can replace CP24 when it is knocked out in mutant plants, as its expression level goes up by 40% to 60% in such mutants[27, 26]. Further refinement in the organization of PSII supercomplexes is unlikely to substantially effect the overall behavior described in this work.

## C.2 Phototsystem II Parameters

In this section we present the Hamiltonian parameters used to calculate energy transfer in our model. Development of improved Hamiltonians, particularly the availability of better site energies for chlorophyll, could result in changes to our description of energy transfer.

### C.2.1 Hamiltonians

| $P_{D1}$ | $P_{D2}$ | $Chl_{D1}$ | $Chl_{D2}$ | $Pho_{D1}$ | $Pho_{D2}$ | $ChlZ_{D1}$ | $ChlZ_{D2}$ |
|---|---|---|---|---|---|---|---|
| 15,015 | 150 | −42 | −53 | −6 | 17 | 1 | 1 |
| 150 | 15,015 | −60 | −36 | 21 | −3 | 1 | 1 |
| −42 | −60 | 14,750 | 7 | 47 | −4 | 3 | 0 |
| −53 | −36 | 7 | 14,992 | −5 | 35 | 0 | 2 |
| −6 | 21 | 47 | −5 | 14,881 | 3 | −4 | 0 |
| 17 | −3 | −4 | 35 | 3 | 14,815 | 0 | −4 |
| 1 | 1 | 3 | 0 | −4 | 0 | 14,993 | 0 |
| 1 | 1 | 0 | 2 | 0 | −4 | 0 | 14,970 |

Table C.3: The Hamiltonian for Reaction Center pigments labelled by their common names[94, 93, 92]. While there is evidence of a shift in site energies for $Chl_{D1}$ and $P_{D1}$ when the RC is bound in a core complex, we have not included these changes in our Hamiltonian.

| c474 | c475 | c476 | c477 | c478 | c479 | c480 | c481 | c482 | c483 | c484 | c485 | c486 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 14,940 | −4 | −1 | −0.2 | −1.6 | 8.4 | 8.7 | 1.3 | −5 | 2.4 | −1.1 | −0.1 | 0.3 |
| −4 | 14,990 | −21.9 | 40.5 | 0.6 | −4.9 | −4.4 | 1.1 | 11.9 | −8.1 | 7 | −4 | −0.5 |
| −1 | −21.9 | 14,890 | −12.1 | −1.3 | −0.5 | −3 | −2.5 | −2.3 | −9.1 | −3.2 | 10.1 | 2.1 |
| −0.2 | 40.5 | −12.1 | 14,760 | 1.3 | 2.1 | 1.3 | 3.4 | −1.4 | 11.2 | 1.5 | −1.7 | 1.6 |
| −1.6 | 0.6 | −1.3 | 1.3 | 14,840 | −7.1 | −43.3 | −8.2 | 12 | −1.8 | −0.4 | 1.2 | −1.9 |
| 8.4 | −4.9 | −0.5 | 2.1 | −7.1 | 14,990 | 20 | 6.7 | −15.2 | 7 | 2.3 | −3.6 | 4.3 |
| 8.7 | −4.4 | −3 | 1.3 | −43.3 | 20 | 14,750 | −5.9 | 40.9 | −1.2 | −4.3 | 18.7 | −12.7 |
| 1.3 | 1.1 | −2.5 | 3.4 | −8.2 | 6.7 | −5.9 | 14,880 | 6.1 | 45.3 | 22.8 | −14.1 | 5.9 |
| −5 | 11.9 | −2.3 | −1.4 | 12 | −15.2 | 40.9 | 6.1 | 14,740 | −27.5 | 63.8 | 2.3 | −2.3 |
| 2.4 | −8.1 | −9.1 | 11.2 | −1.8 | 7 | −1.2 | 45.3 | −27.5 | 15,140 | −41.6 | −7.1 | 4.8 |
| −1.1 | 7 | −3.2 | 1.5 | −0.4 | 2.3 | −4.3 | 22.8 | 63.8 | −41.6 | 14,990 | −14.4 | 6.8 |
| −0.1 | −4 | 10.1 | −1.7 | 1.2 | −3.6 | 18.7 | −14.1 | 2.3 | −7.1 | −14.4 | 14,950 | −31.1 |
| 0.3 | −0.5 | 2.1 | 1.6 | −1.9 | 4.3 | −12.7 | 5.9 | −2.3 | 4.8 | 6.8 | −31.1 | 15,010 |

Table C.4: The Hamiltonian for CP43 pigments labelled by their numbering in the 3arc crystal structure [68].

| c511 | c512 | c513 | c514 | c515 | c516 | c517 | c518 | c519 | c520 | c521 | c522 | c523 | c524 | c525 | c526 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 14,684 | −49 | −12 | 4 | −2 | −4 | −1 | 4 | −6 | −2 | −2 | 3 | −2 | −1 | −2 | −1 |
| −49 | 15,060 | −83 | 20 | 5 | −6 | −5 | 5 | −20 | −13 | −4 | 6 | −9 | −2 | 4 | −2 |
| −12 | −83 | 14,793 | 0 | −6 | −59 | −3 | −8 | 32 | 15 | 2 | −10 | 4 | −2 | −1 | −3 |
| 4 | 20 | 0 | 14,950 | −47 | −13 | 65 | 5 | −11 | −6 | −1 | 21 | −35 | 8 | −3 | 0 |
| −2 | 5 | −6 | −47 | 14,810 | 72 | −20 | −2 | 1 | −5 | −3 | −6 | −7 | −4 | 13 | 11 |
| −4 | −6 | −59 | −13 | 72 | 14,880 | 7 | −3 | 8 | −4 | 6 | −4 | 9 | 4 | 1 | 9 |
| −1 | −5 | −3 | 65 | −20 | 7 | 15,130 | 0 | 4 | 1 | 6 | −2 | 17 | 3 | −2 | 2 |
| 4 | 5 | −8 | 5 | −2 | −3 | 0 | 14,970 | −17 | −83 | −16 | 20 | −5 | −3 | 3 | −1 |
| −6 | −20 | 32 | −11 | 1 | 8 | 4 | −17 | 14,810 | 48 | 9 | −16 | 10 | 4 | −2 | 6 |
| −2 | −13 | 15 | −6 | −5 | −4 | 1 | −83 | 48 | 15,270 | −13 | 52 | 0 | −6 | 32 | −7 |
| −2 | −4 | 2 | −1 | −3 | 6 | 6 | −16 | 9 | −13 | 14,710 | −12 | 56 | 37 | −18 | 8 |
| 3 | 6 | −10 | 21 | −6 | −4 | −2 | 20 | −16 | 52 | −12 | 15,060 | −64 | 95 | 25 | 2 |
| −2 | −9 | 4 | −35 | −7 | 9 | 17 | −5 | 10 | 0 | 56 | −64 | 14,860 | −63 | −3 | 2 |
| −1 | −2 | −2 | 8 | −4 | 4 | 3 | −3 | 4 | −6 | 37 | 95 | −63 | 14,880 | −12 | 13 |
| −2 | 4 | −1 | −3 | 13 | 1 | −2 | 3 | −2 | 32 | −18 | 25 | −3 | −12 | 14,970 | −8 |
| −1 | −2 | −3 | 0 | 11 | 9 | 2 | −1 | 6 | −7 | 8 | 2 | 2 | 13 | −8 | 14,530 |

Table C.5: The Hamiltonian for CP47 pigments labelled by their numbering in the 3arc crystal structure[93].

| Protein | Molecule | Spectral Density |
|---|---|---|
| Reaction Center | Chl A/Pheophytin | 1 |
| CP47 | Chl A | 1 |
| CP43 | Chl A | 1 |
| LHCII, Minor Complexes | Chl A | 2 |
| LHCII, Minor Complexes | Chl B | 2 |

Table C.6: The spectral density number associated with each pigment-protein complex in PSII.

| Protein | $S_0$ | $s_1$ | $s_2$ | $\omega_1$ ($\mathrm{cm}^{-1}$) | $\omega_2$ ($\mathrm{cm}^{-1}$) |
|---|---|---|---|---|---|
| RC | 0.65 | 0.8 | 0.5 | 0.532 | 1.94 |
| CP47,CP43 | 0.5 | 0.8 | 0.5 | 0.532 | 1.94 |

Table C.7: The spectral density parameters that describe core pigments in combination with eq. (C.1).

The LHCII Hamiltonian is exactly reproduced in the SI of Novoderezhkin's work [79] .

### C.2.2 Spectral Density

We have used two kinds of spectral densities to define the pigment-protein interaction within the PSII supercomplex. Table C.6 shows the spectral density associated with each pigment-protein complex within the system. The equations for both types of spectral densities are given in eqs. (C.1) and (C.2). The pigments bound by the core complex are described by eq. (C.1) using the parameters described in Table C.7[93]. The Chl-a and Chl-b pigments bound by LHCII are described by one over-damped Brownian oscillator describing the low-frequency motions of the bath and 48 under-damped Brownian oscillators describing high-frequency intra-molecular vibrations, as shown in eq. (C.2) [79] . The parameters describing the spectral density used for LHCII pigments are shown in Tables C.8 and C.9.

$$\chi_1''(\omega) = (\pi\hbar)\frac{S_0}{s_1+s_2}\sum_{i=1,2}\frac{s_i(\omega^5)}{7!2\omega_i^4}e^{-\sqrt{\frac{\omega}{\omega_i}}} \tag{C.1}$$

$$\chi_2''(\omega) = 2\lambda_0\frac{\omega\Gamma_0}{\omega^2+\Gamma_0^2} + \sum_{j=1}^{N_{\mathrm{vib}}} 2S_j\omega_j^3\frac{\omega\Gamma_{\mathrm{vib}}}{(\omega_j^2-\omega^2)^2+\omega^2\Gamma_{\mathrm{vib}}^2} \tag{C.2}$$

| Protein | Molecule | $\lambda_0$ ($cm^{-1}$) | $\Gamma_0$ ($cm^{-1}$) |
|---|---|---|---|
| LHCII, Minor Complexes | chl A | 37 | 30 |
| LHCII, Minor Complexes | chl B | 48 | 30 |

Table C.8: The spectral density number associated with each pigment-protein complex in PSII.

| $\omega_j$ ($cm^{-1}$) | $S_j$ | $\omega_j$ ($cm^{-1}$) | $S_j$ | $\omega_j$ ($cm^{-1}$) | $S_j$ |
|---|---|---|---|---|---|
| 97 | $3.95 \cdot 10^{-2}$ | 752 | $2.1 \cdot 10^{-2}$ | 1,286 | $3.7 \cdot 10^{-3}$ |
| 138 | $4.56 \cdot 10^{-2}$ | 795 | $3.9 \cdot 10^{-3}$ | 1,304 | $4.7 \cdot 10^{-3}$ |
| 213 | $5.42 \cdot 10^{-2}$ | 916 | $1.73 \cdot 10^{-2}$ | 1,322 | $2.47 \cdot 10^{-2}$ |
| 260 | $4.44 \cdot 10^{-2}$ | 986 | $8.4 \cdot 10^{-3}$ | 1,338 | $3.2 \cdot 10^{-3}$ |
| 298 | $3.58 \cdot 10^{-2}$ | 995 | $1.85 \cdot 10^{-2}$ | 1,354 | $4.7 \cdot 10^{-3}$ |
| 342 | $5.92 \cdot 10^{-2}$ | 1,052 | $9.9 \cdot 10^{-3}$ | 1,382 | $5.4 \cdot 10^{-3}$ |
| 388 | $2.71 \cdot 10^{-2}$ | 1,069 | $5.2 \cdot 10^{-3}$ | 1,439 | $5.4 \cdot 10^{-3}$ |
| 425 | $1.73 \cdot 10^{-2}$ | 1,110 | $9.1 \cdot 10^{-3}$ | 1,487 | $6.4 \cdot 10^{-3}$ |
| 518 | $3.21 \cdot 10^{-2}$ | 1,143 | $3.33 \cdot 10^{-2}$ | 1,524 | $5.2 \cdot 10^{-3}$ |
| 546 | $2.2 \cdot 10^{-3}$ | 1,181 | $1.43 \cdot 10^{-2}$ | 1,537 | $1.78 \cdot 10^{-2}$ |
| 573 | $6.9 \cdot 10^{-3}$ | 1,190 | $5.4 \cdot 10^{-3}$ | 1,553 | $7.4 \cdot 10^{-3}$ |
| 585 | $2.5 \cdot 10^{-3}$ | 1,208 | $1.5 \cdot 10^{-2}$ | 1,573 | $3.7 \cdot 10^{-3}$ |
| 604 | $1.6 \cdot 10^{-3}$ | 1,216 | $1.43 \cdot 10^{-2}$ | 1,580 | $3.7 \cdot 10^{-3}$ |
| 700 | $1.6 \cdot 10^{-3}$ | 1,235 | $5.7 \cdot 10^{-3}$ | 1,612 | $3.7 \cdot 10^{-3}$ |
| 722 | $3.2 \cdot 10^{-3}$ | 1,252 | $5.2 \cdot 10^{-3}$ | 1,645 | $3 \cdot 10^{-3}$ |
| 742 | $3.21 \cdot 10^{-2}$ | 1,260 | $5.2 \cdot 10^{-3}$ | 1,673 | $8 \cdot 10^{-4}$ |

Table C.9: The frequencies and Huang-Rhys factors for the under-damped Brownian oscillators describing vibrational modes of Chl A molecules in PSII.

## C.3 Testing Infinitely Fast Intra-Domain Equilibration

In order to assess the reliability of assuming infinitely fast equilibration within domains we calculate the population as a function of time for each domain of the system (the sum of the excitonic populations associated with the domain) with both the full generalized Förster/modified Redfield rate matrix and the domain model that assumes infinitely fast thermalization within each domain. For each domains population curve as a function of time, we calculate the maximum absolute difference in population between the two methods as a fraction of the maximum population calculated with the full rate matrix (as shown in eq. (C.3), where $d$ is the index of the domain). This results in no more than 2% error ($\Delta^{(d)}_{\text{max}} \leq 0.02$) no matter which domain is initially excited. We can also characterize the total error over the coarse of a calculation by comparing the integral of the error to the integral of the population curve (as shown in eq. (C.4)). The integrated error fraction shows less than 7% error ($\Delta^{(d)}_{\text{Integral}} < 0.07$) no matter what domain is initially excited. Figure C.1 plots the population curve that has both the largest maximum error and the largest integrated error.

$$\Delta^{(d)}_{\text{max}} = \max\left(\frac{|P^{(d)}_{\text{exc}}(t) - P^{(d)}_{\text{dom}}(t)|}{\max\ P^{(d)}_{\text{exc}}(t)}\right) \tag{C.3}$$

$$\Delta^{(d)}_{\text{Integral}} = \frac{\int_0^\infty |P^{(d)}_{\text{exc}}(t) - P^{(d)}_{\text{dom}}(t)|\mathrm{d}t}{\int_0^\infty P^{(d)}_{\text{exc}}(t)\mathrm{d}t} \tag{C.4}$$

## C.4 Coarse Grained Models of Energy Transfer

The coarse-grain descriptions of energy transfer are calculated by grouping collections of domains into compartments and calculating the Boltzmann-averaged rate constants for transfer between each compartment, as described in the Materials and Methods. Figure C.2 (A)-(C) show the timescales of transfer between compartments for the $C_2S_2M_2$ using the protein, transfer-to-trap limited, and trap limited models, respectively. Time constants for energy moving closer (farther) to the final photochemical product RP2 are shown in black (red). When only a single number is shown for a connection, the rates of forward and backward transfer are symmetric. Where no line is shown between proteins, energy transfer timescales were found to be longer than 100 ps and were excluded for clarity. Using the coarse-grained rate matrix, the dynamics of initial charge separation were calculated and the resulting RP1 population trace is shown in Fig. C.2 (D). As expected from our previous analysis of the temporal structure of energy transfer in $C_2S_2M_2$, the trap-limited model is found to very poorly reproduce the population dynamics as seen by the overestimate of the early-time charge separation. Both the protein and transfer-to-trap limited models qualitatively reproduce the energy transfer dynamics of the $C_2S_2M_2$, but both models still over-estimate the rate at which excitation energy arrives at the RC and performs charge separation.

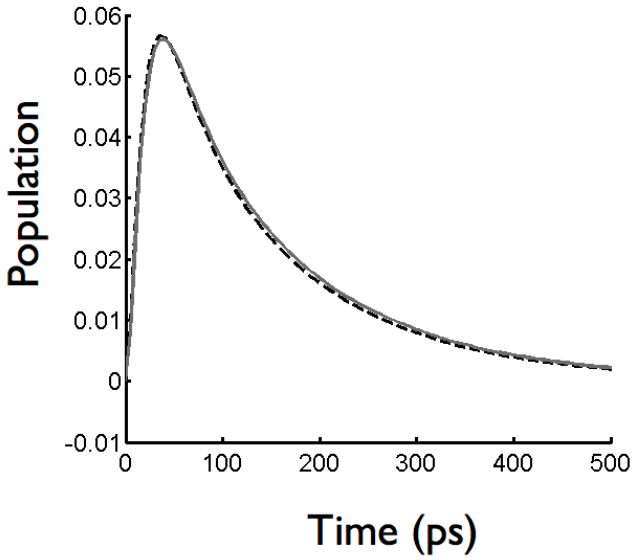



Figure C.1: A comparison between the domain population curves with the largest value $\Delta^{(d)}_{\max}$ and $\Delta^{(d)}_{\text{Integral}}$. Population trace calculated with the generalized Förster/modified Redfield (Domain) model is shown with a dashed black (sold grey) line.

## C.5 Sample Preparation and Modeling

The long lifetime components of the fluorescence decay curves between 1.7 ns and 3.3 ns that have been assigned to contamination by disconnected chlorophyll. We have followed the suggestion of Caffarri and coworkers [17] in removing these long lifetime components prior to fitting the experimental data. The long lifetime components of a fluorescence decay curve, however, report on the balance between the rate of irreversible charge separation and charge recombination at the reaction center. Therefore accurately determining the origin of the long time fluorescence decay in the preparations of PSII supercomplexes is essential for accurately modeling electron transfer kinetics.

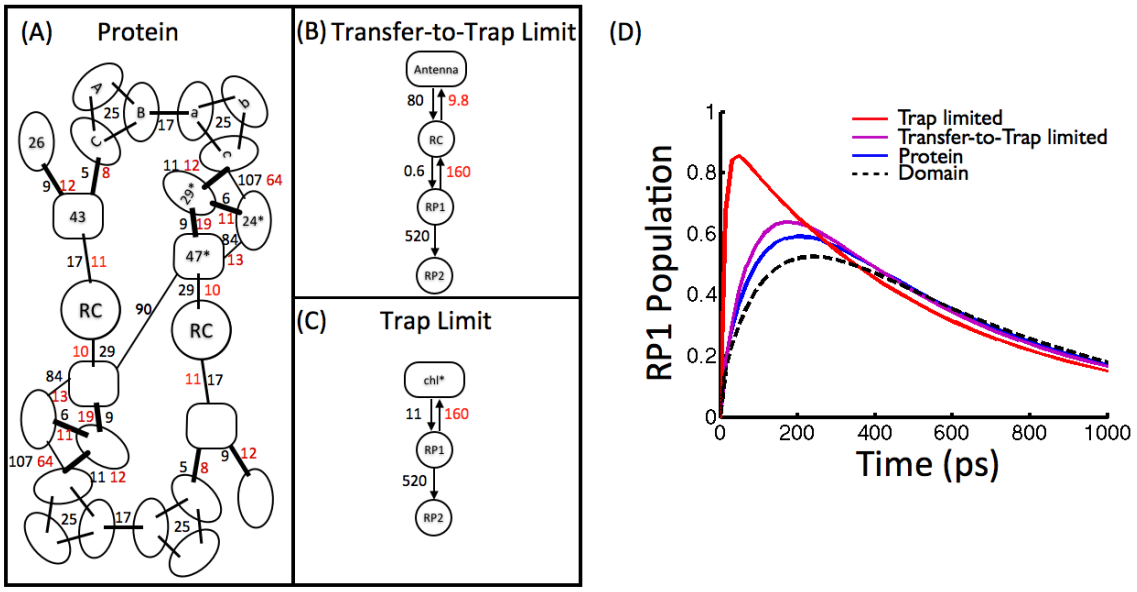



Figure C.2: A comparison between different coarse grain models of $C_2S_2M_2$ is shown both as kinetic schemes and simulated RP1 population trace. (A)-(C) The timescales of transfer ($k^{-1}$) for the various coarse-grain models are shown in units of ps. Numbers in black (red) represent transfers that move excitation closer to (farther away from) the charge separated states. (A)-(C) represent the protein model, transfer-to-trap limited model, and the trap limited model, respectively. (D) The simulated curves for the RP1 population calculated with ($\tau_{cs}$ = 0.61 ps, $\tau_{rc}$ = 160 ps, and $\tau_{irr}$ = 520 ps) are shown in black (Domain), blue (Protein), purple (transfer-to-trap limited), and red (trap limited). The initial excitation was distributed evenly among Chl-a molecules.