

# Numerical Method for Nonlinear Optical Spectroscopies: Ultrafast Ultrafast Spectroscopy

Peter A. Rose<sup>1</sup> and Jacob J. Krich<sup>1,2</sup>

<sup>1</sup>*Department of Physics, University of Ottawa, Ottawa, ON, Canada*

<sup>2</sup>*School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada*

We outline a novel numerical method, called Ultrafast Ultrafast (UF<sup>2</sup>), for calculating the  $n^{\text{th}}$ -order wavepackets required for calculating  $n$ -wave mixing signals. The method is simple to implement, and we demonstrate that it is computationally more efficient than other methods in a wide range of use cases. Resulting spectra are identical to those calculated using the standard response function formalism but with increased efficiency. The computational speed-ups of UF<sup>2</sup> come from (a) non-perturbative and costless propagation of the system time-evolution (b) numerical propagation only at times when perturbative optical pulses are non-zero and (c) use of the fast Fourier transform convolution algorithm for efficient numerical propagation. The simplicity of this formalism allows us to write a simple software package that is as easy to use and understand as the Feynman diagrams that form the basis of our understanding of  $n$ -wave mixing.

## I. INTRODUCTION

Ultrafast nonlinear optical spectroscopies, in the perturbative light-matter limit, are powerful tools for elucidating details about the electronic structure and ultrafast dynamics of optically-active systems. Interpreting such spectra often requires understanding what signals would be produced by a range of parametrized system Hamiltonians; one then seeks the best agreement with experimental results by varying system parameters such as energy levels, couplings, dephasing rates, and more. This kind of fitting, as in Ref. 1, involves repeatedly rederiving the spectroscopic signals as system parameters change, which can be computationally expensive.

Simulations of nonlinear spectroscopies can be particularly challenging when the optical pulse durations are similar to relevant timescales in the system dynamics, especially when one must consider the pulses overlapping in time. For spectroscopies that rely on varying pulse durations, simulating those variations can be computationally expensive, limiting the range of systems and pulses one can study.<sup>2</sup>

In interpreting and simulating nonlinear spectroscopies, frequently the pulse shapes are ignored, and calculations are performed in the impulsive limit, which is computationally simpler but can be inaccurate. In the last several years there has been an increased interest in understanding the role of finite-pulse effects in nonlinear spectroscopies, and there are several symbolic solutions or approximations for calculating 4-wave mixing signals using finite pulses.<sup>2-6</sup> These symbolic solutions each assume a particular, idealized pulse shape, such as a Gaussian or Lorentzian profile. The motivation driving these symbolic solutions is to provide methods for speeding up the calculation of nonlinear optical spectra, sometimes while including pulse overlap effects.

We present a complementary numerical technique we call Ultrafast Ultrafast (UF<sup>2</sup>) spectroscopy, which can efficiently simulate nonlinear optical spectra using arbitrary pulse shapes and for arbitrary perturbative order

$n$ , including pulse-overlap effects. It is a fast implementation of the standard results in perturbative nonlinear spectroscopy in the electric dipole approximation. The version we present here is for closed systems.<sup>i</sup> It assumes that the system Hamiltonian is time-independent and has a finite relevant eigenbasis.<sup>ii</sup> When the pulses are finite, we perform the time propagation using the convolution theorem and the fast Fourier transform (FFT) and therefore benefit from the speed of the FFT. Working in the energy eigenbasis allows non-perturbative and costless time evolution of the system at times when the optical pulses are negligible, giving a dramatic speed-up over algorithms that involve numerical integration of the wavefunctions<sup>7</sup> or density matrices<sup>8-11</sup> for each time-step; this manuscript only discusses closed systems and therefore does not consider the system-bath dynamics included in density matrix formalisms.

In Section II we derive the mathematical framework of UF<sup>2</sup>, show how it is used, and describe some of the techniques used for increasing its efficiency. Like any perturbative technique, it is most efficient where the rotating wave approximation and phase-matching can be assumed (see Section IID), but it does not require these assumptions. In Section III we discuss the computational scaling of UF<sup>2</sup>, and in Section IIIB we demonstrate that UF<sup>2</sup> can, and does, surpass the speed of known symbolic solutions for large system sizes. We also compare it to the time-domain wavepackets method. UF<sup>2</sup> enables computationally efficient studies of the effects of varying system parameters and optical pulse shapes. We give specific examples of the speed of UF<sup>2</sup> in Section IV, where we show transient absorption (TA) spectra for a system with a Hamiltonian of size  $N \approx 10$ , and another with a Hamiltonian of size  $N \approx 20000$ . For the former case, we

<sup>i</sup> UF<sup>2</sup> can be extended to work with open systems, but we do not do so here.

<sup>ii</sup> UF<sup>2</sup> is best suited for studying energy transfer. It is not designed to study reactions.

show how UF<sup>2</sup> can be used to perform rapid parameter sweeps. The latter example gives a sense of the upper limit of what our implementation of UF<sup>2</sup> can be used for on a laptop.

UF<sup>2</sup> is not only efficient but also easy and intuitive to use. The method is built around single-sided Feynman diagrams and their double-sided counterparts, which describe the perturbative pathways contributing to a particular spectroscopic signal. Using the method is a simple process of translating a desired Feynman diagram into a set of iterative function calls. A python implementation of UF<sup>2</sup> is available for download at <https://github.com/peterarose/ultrafastultrafast>. This code includes examples of how to translate Feynman diagrams into the language of UF<sup>2</sup>, and an example implementation that calculates TA spectra.

## II. ALGORITHM

### A. Overview

We begin with a Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \hat{H}'(t), \quad (1)$$

where the light-matter interaction in the electric-dipole approximation is

$$\hat{H}'(t) = -\hat{\boldsymbol{\mu}} \cdot \mathbf{E}(t), \quad (2)$$

and is treated as a perturbation, where  $\hat{\boldsymbol{\mu}}$  and  $\mathbf{E}(t)$  are the dipole moment of the system and the external electric field, respectively. Bold face symbols indicate Cartesian vectors.  $\hat{H}_0$  describes the material system (e.g., a molecule or quantum dot) and has eigenstates  $|\phi\rangle$  and eigenvalues  $\hbar\omega_\phi$ , which are assumed to be known (either analytically or numerically). Therefore the unitary time-evolution operator  $\hat{U}_0(t - t_0) = e^{-i\hat{H}_0(t-t_0)/\hbar}$  is known. We assume that the set of eigenstates  $\{|\phi\rangle\}$  is finite, and that all relevant wavefunctions in this problem can be expressed using  $N$  eigenfunctions as

$$|\psi(t)\rangle = \sum_{\phi=1}^N e^{-i\omega_\phi t} c_\phi(t) |\phi\rangle. \quad (3)$$

The electric dipole operator  $\hat{\boldsymbol{\mu}}$  must be known in the eigenbasis of  $\hat{H}_0$ , where we define matrix elements as

$$\boldsymbol{\mu}_{\phi\phi'} = \langle\phi|\hat{\boldsymbol{\mu}}|\phi'\rangle.$$

Note that only  $\omega_\phi$  and  $\boldsymbol{\mu}_{\phi\phi'}$  are needed. The eigenfunctions themselves are unnecessary if  $\boldsymbol{\mu}_{\phi\phi'}$  can be calculated in some other way.

We describe the electric field classically as a sum of pulses, where each pulse is denoted by a lowercase letter starting from  $a$ . A typical 4-wave mixing signal would be

calculated by using up to 4 pulses. We write the electric field as a sum over  $L$  pulses,

$$\mathbf{E}(t) = \sum_{i=a,b,\dots,L} \mathbf{e}_i \varepsilon_i(t) + \mathbf{e}_i^* \varepsilon_i^*(t),$$

where  $\mathbf{e}_i$  is the possibly complex polarization vector, and the amplitude  $\varepsilon_i$  of each pulse is defined with envelope  $A_i$ , central frequency  $\omega_i$ , wavevector  $\mathbf{k}_i$ , and phase  $\phi_i$  as

$$\varepsilon_i(t) = A_i(t - t_i) e^{-i[\omega_i(t-t_i) - \mathbf{k}_i \cdot \mathbf{r} + \phi_i]}, \quad (4)$$

where  $t_i$  is the arrival time of each pulse, and we define the Fourier transform of the pulse as

$$\tilde{\varepsilon}_i(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt \varepsilon_i(t) e^{i\omega t}.$$

Then the light-matter interaction is a sum over rotating ( $\varepsilon_i$ ) and counter-rotating ( $\varepsilon_i^*$ ) terms. We express these terms individually as

$$\hat{H}'_{j(*)}(t) = -\hat{\boldsymbol{\mu}} \cdot \mathbf{e}_j^{(*)} \varepsilon_j^{(*)}(t) \quad (5)$$

so that

$$\hat{H}'(t) = \sum_{i=a,b,\dots} \hat{H}'_i(t) + \hat{H}'_{i*}(t). \quad (6)$$

In general, 4-wave mixing signals are calculated from the third-order perturbed density matrix.<sup>12</sup> For closed systems, as described here, we can decrease the computational complexity of the problem by calculating the necessary perturbed wavefunctions up to third-order. We expand the true time-dependent wavefunction using the usual perturbative expansion

$$|\psi(t)\rangle = |\psi^{(0)}(t)\rangle + |\psi^{(1)}(t)\rangle + |\psi^{(2)}(t)\rangle + \dots \quad (7)$$

where  $|\psi^{(0)}(t)\rangle$  is the initial time-independent (ignoring a trivial phase) state, and  $|\psi^{(n)}(t)\rangle$  is proportional to  $(\hat{H}'(t))^n$ . UF<sup>2</sup> is a fast method for calculating  $|\psi^{(n)}(t)\rangle$ . The total polarization field is

$$\begin{aligned} \mathbf{P}(t) &= \langle\psi(t)|\hat{\boldsymbol{\mu}}|\psi(t)\rangle \\ &= \mathbf{P}^{(1)}(t) + \mathbf{P}^{(2)}(t) + \mathbf{P}^{(3)}(t) + \dots \end{aligned}$$

To calculate  $\mathbf{P}^{(n)}(t)$ , we expand  $|\psi(t)\rangle$  using Eq. 7 and keep only terms proportional to  $(\hat{H}'(t))^n$ .<sup>13</sup> As an example, the 3<sup>rd</sup>-order polarization field is

$$\mathbf{P}^{(3)}(t) = \langle\psi^{(3)}(t)|\hat{\boldsymbol{\mu}}|\psi^{(0)}\rangle + \langle\psi^{(2)}(t)|\hat{\boldsymbol{\mu}}|\psi^{(1)}\rangle + c.c.$$

From  $\mathbf{P}^{(3)}(t)$  any 3<sup>rd</sup>-order spectroscopic signal can be determined.  $\mathbf{P}^{(3)}(t)$  is the material response emitted after three optical interactions. In 2D-photon echo (2DPE) spectroscopy, the response is due to a single interaction with 3 separate pulses, with arrival times  $t_a, t_b, t_c$ . The

delay times between pulses are  $t_b - t_a = \tau$  and  $t_c - t_b = T$ . The polarization field is therefore a function of  $\tau$  and  $T$ :  $\mathbf{P}^{(3)}(\tau, T; t)$ . Typically the quantity of interest is actually the Fourier transform partner

$$\mathbf{P}^{(3)}(\tau, T; \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt \mathbf{P}^{(3)}(\tau, T; t) e^{i\omega t}.$$

The polarization field produces an electric field, which can be heterodyne detected using a fourth local oscillator (LO) pulse. For the case of the 2DPE rephasing signal, the desired signal is calculated as<sup>12</sup>

$$S_{2D}^{(3)}(\tau, T, \omega) = \text{Im} \left[ \tilde{\varepsilon}_{LO}^*(\omega) \mathbf{e}_{LO}^* \cdot \mathbf{P}_{2D}^{(3)}(\tau, T; \omega) \right]. \quad (8)$$

We derive UF<sup>2</sup> for the classic example given in Eq. 8. In Section IV, we show example TA spectra, which arise due to the interaction of two pulses, a pump and a probe, separated by the delay time  $T$ . In the TA case, the first two interactions happen with the pump pulse, and therefore  $\tau = 0$ . The third interaction comes from the probe, which also acts as the local oscillator. TA calculations are thus a function of only two variables:

$$S_{TA}^{(3)}(T, \omega) = \text{Im} \left[ \tilde{\varepsilon}_{\text{probe}}^*(\omega) \mathbf{e}_{\text{probe}}^* \cdot \mathbf{P}_{TA}^{(3)}(T, \omega) \right]. \quad (9)$$

Returning to the more general case with  $\tau$  varying, the canonical formula for calculating the third-order time-dependent polarization field is

$$P^{(3)}(\tau, T, t) = \iiint dt_3 dt_2 dt_1 E(t - t_3) E(t - t_3 - t_2 + T) \times E(t - t_3 - t_2 - t_1 + T + \tau) R(t_1, t_2, t_3) \quad (10)$$

where we follow Mukamel and suppress the polarization of the fields,<sup>12</sup> and  $R(t_1, t_2, t_3)$  is a material response function containing all of the factors arising from  $\hat{H}_0$  and  $\hat{\mu}$ . This formulation separates the material properties ( $\hat{H}_0$  and  $\hat{\mu}$ ) from the shapes of the electric fields. If the electric fields are impulsive (i.e., short compared with any system timescale), then the convolutions in Eq. 10 are trivial and spectral signals may be calculated directly from the response function.

In this paper we focus on the case where the shape of the electric field cannot be ignored, and therefore studying the response function alone is insufficient to predict, interpret or understand spectroscopic observables. The triple-nested convolutions of Eq. 10 are so costly that they are rarely carried out. We present a fast method for calculating the  $n^{\text{th}}$ -order wavefunctions, and thence the desired polarization  $\mathbf{P}^{(n)}(t)$ , with arbitrary pulse shapes. The wavepackets can be studied in conjunction with the signal, giving intuition for the underlying physics.<sup>14</sup>

## B. Derivation

We follow standard time-dependent perturbation theory by considering that at time  $t_0$  the system begins in a time-independent state  $|\psi^{(0)}(t_0)\rangle$ , i.e., an eigenstate of  $\hat{H}_0$ . The perturbation  $\hat{H}'(t)$  produces  $|\psi(t)\rangle$  as in Eq. 7, where  $|\psi^{(n)}\rangle$  has contributions proportional to  $(\hat{H}')^n$ . Then the standard result has<sup>13</sup>

$$|\psi^{(n)}(t)\rangle = -\frac{i}{\hbar} \int_{t_0}^t ds \hat{U}_0^{-1}(s - t) \hat{H}'(s) |\psi^{(n-1)}(s)\rangle.$$

Since  $|\psi^{(0)}\rangle$  is time-independent (up to a trivial phase), we are free to send  $t_0 \rightarrow -\infty$ . We also substitute  $t' = t - s$  to arrive at

$$|\psi^{(n)}(t)\rangle = -\frac{i}{\hbar} \int_0^\infty dt' \hat{U}_0(t') \hat{H}'(t - t') |\psi^{(n-1)}(t - t')\rangle.$$

Using the decomposition of  $\hat{H}'(t)$  in Eq. 6, we define  $|\psi^{(n)}(t)\rangle$  as a sum over  $2L$  terms

$$|\psi^{(n)}(t)\rangle = \sum_{j=a,b,\dots,L} \left( \hat{K}_j + \hat{K}_{j^*} \right) |\psi^{(n-1)}(t)\rangle,$$

where

$$\hat{K}_{j^*} = -\frac{i}{\hbar} \int_0^\infty dt' \hat{U}_0(t') \hat{H}'_{j^*}(t - t').$$

A small rearrangement of terms allows us to perform this integral numerically much more quickly:

$$\hat{K}_{j^*} = -\frac{i}{\hbar} \hat{U}_0(t) \int_0^\infty dt' \hat{U}_0^{-1}(t - t') \hat{H}'_{j^*}(t - t'). \quad (11)$$

Most spectroscopic signals are not calculated using the full perturbative wavefunction  $|\psi^{(n)}(t)\rangle$ , because only specific pathways give nonzero contributions, which are visualized using Feynman diagrams, as in Fig. 1.<sup>12</sup> Each operator  $\hat{K}_{j^*}$  represents a single interaction arrow in a Feynman diagram, as in Fig. 2. UF<sup>2</sup> calculates the wavefunctions contributing to each diagram separately. For example, as shown in Fig. 1, the stimulated emission (SE) contribution requires two wavefunctions that we label  $|\psi_{ac^*}(t)\rangle \equiv \hat{K}_{c^*} \hat{K}_a |\psi^{(0)}\rangle$  and  $|\psi_b(t)\rangle \equiv \hat{K}_b |\psi^{(0)}\rangle$ .

Following Eq. 3, we write

$$|\psi_p(t)\rangle = \sum_{\phi'} e^{-i\omega_{\phi'} t} c_{\phi',p}(t) |\phi'\rangle,$$

where  $p$  can be a multi-index, such as  $ac^*$ . Then we use Eqs. 5 and 11 to write

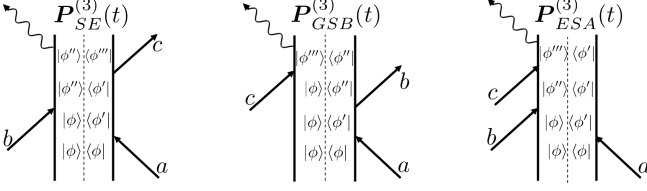


Figure 1. From left to right, stimulated emission (SE), ground-state bleach (GSB), and excited-state absorption (ESA) double-sided Feynman diagrams for the rephasing 2D photon echo signal. The dashed lines down the centers of each diagram emphasize that the ket and bra evolve independently for closed systems. We calculate the left and right sides of the diagram separately. These are the only three diagrams that contribute to the rephasing signal when the pulses are well-separated in time, though additional diagrams must be considered when pulses overlap.

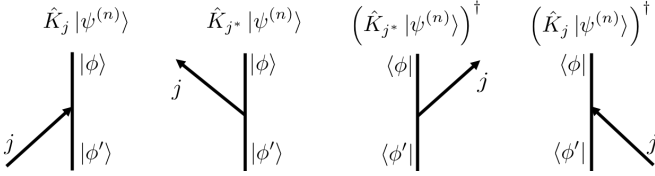


Figure 2. Building blocks of Feynman diagrams for the interaction with optical pulse  $j$ . The left two diagrams are for kets and the right two are for bras. Time moves from bottom to top, as the system begins in a linear combination of states  $|\phi'\rangle$ , and evolves to a linear combination of states  $|\phi\rangle$ . The arrow sloping up to the right is a contribution of the operator  $K_j$ , and the arrow sloping up to the left is a contribution of the operator  $K_{j^*}$ .

$$\begin{aligned} |\psi_{pj^{(*)}}(t)\rangle &\equiv \hat{K}_{j^{(*)}} |\psi_p(t)\rangle \\ &= \frac{i}{\hbar} \hat{U}_0(t) \int_0^\infty dt' \hat{U}_0^{-1}(t-t') \sum_\phi |\phi\rangle \langle\phi| \left( \hat{\mu} \cdot \mathbf{e}_j^{(*)} \varepsilon_j^{(*)}(t-t') \right) \sum_{\phi'} e^{-i\omega_{\phi'}(t-t')} c_{\phi',p}(t-t') |\phi'\rangle \end{aligned} \quad (12)$$

$$= \sum_\phi e^{-i\omega_\phi t} |\phi\rangle \frac{i}{\hbar} \int_{-\infty}^\infty dt' \theta(t') e^{i\omega_\phi(t-t')} \underbrace{\sum_{\phi'} \left( \mu_{\phi\phi'} \cdot \mathbf{e}_j^{(*)} \varepsilon_j^{(*)}(t-t') \right) e^{-i\omega_{\phi'}(t-t')} c_{\phi',p}(t-t')}_{y_\phi(t-t')}, \quad (13)$$

where  $\theta(t)$  is the unit step function. We rewrite Eq. 13 as

$$|\psi_{pj^{(*)}}(t)\rangle = \sum_\phi e^{-i\omega_\phi t} |\phi\rangle \underbrace{\frac{i}{\hbar} [\theta * y_\phi](t)}_{c_{\phi,pj^{(*)}}(t)}, \quad (14)$$

where

$$[x * y](t) = \int_{-\infty}^\infty dt' x(t') y(t-t')$$

is a convolution. Thus we arrive at a compact description of the coefficients  $c_{\phi,pj^{(*)}}(t)$ . If we make the physical

assumption that the incident pulse is localized in time, i.e.,  $\varepsilon_j(t)$  is negligible for  $t < t_{j,\min}$  or  $t > t_{j,\max}$ , then it is also true that  $y_{\phi,\alpha j^{(*)}}(t)$  is negligible when  $t < t_{j,\min}$  or  $t > t_{j,\max}$ . This assumption implies that

$$c_{\phi,pj^{(*)}}(t) = \begin{cases} 0 & t < t_{j,\min} \\ r_\phi(t) & t_{j,\min} < t < t_{j,\max} \\ C_\phi & t > t_{j,\max} \end{cases} \quad (15)$$

Therefore, we need only calculate this convolution for  $t_{j,\min} < t < t_{j,\max}$  (see Fig. 3). Note that without the rearrangement of Eq. 11, the convolution in Eq. 14

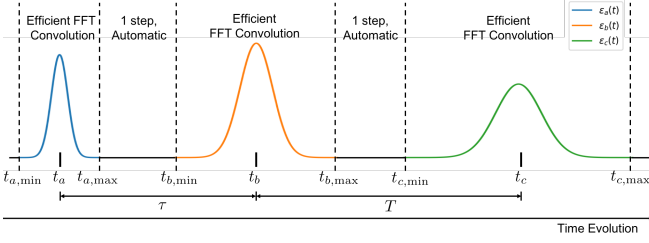


Figure 3. Series of three pulses, as would be used in a rephasing 2DPE experiment, showing the time intervals  $(t_{j,\min}, t_{j,\max})$  during which the pulses are non-negligible. UF<sup>2</sup> only performs time propagations during these time periods, using efficient FFT convolutions. At other times, the system time evolution operator  $\hat{U}_0$  gives exact time evolution.

would be more difficult to evaluate numerically because the functions  $c_{\phi,pj(*)}(t)$  would include the phase factor  $e^{-i\omega_{\phi}t}$  even after  $t_{j,\max}$  and thus not be constant. This trivial definitional difference when working symbolically makes a large difference when working numerically.

Physically, we only need to solve for the time dependence due to the interaction with the pulse while the pulse is non-zero. The rest of the time-dependence is contained entirely in  $\hat{H}_0$ , and is therefore known exactly. This realization drastically reduces the computational cost of UF<sup>2</sup> compared to techniques that must use time-stepping for both the system dynamics and the perturbation.<sup>7</sup> We solve for the function  $r_{\phi}(t)$  numerically using the convolution theorem and FFT. Reducing the numerical integration to the time-interval  $(t_{j,\min}, t_{j,\max})$  reduces the number of discrete time points required to  $\approx 10^2 - 10^3$  for reasonably well-behaved pulses (i.e., pulses that have a finite duration and frequency bandwidth).

Translating the operators  $K_{j(*)}$  into computational algorithms is straightforward. We consider the time interval  $(t_{j,\min}, t_{j,\max})$  which begins just before the  $j^{\text{th}}$  pulse arrives and ends just after the  $j^{\text{th}}$  pulse leaves. We discretize this time interval to have  $M$  equally spaced points with spacing  $dt$ , which must be small enough to resolve the pulses. Therefore, the wavefunctions that give rise to the SE signal are calculated in three steps:

1. Calculate  $|\psi_b(t)\rangle = \hat{K}_b |\psi^{(0)}\rangle$  for fixed  $\tau$ ,
2. Calculate  $|\psi_a(t)\rangle = \hat{K}_a |\psi^{(0)}\rangle$ ,
3. Calculate  $|\psi_{ac^*}(t)\rangle = \hat{K}_{c^*} |\psi_a(t)\rangle$  for fixed  $T$ .

The polarization field is then

$$\mathbf{P}_{SE}^{(3)}(t)|_{\tau,T} = \langle \psi_{ac^*}(t) | \hat{\mu} | \psi_b(t) \rangle |_{\tau,T}. \quad (16)$$

This process is repeated for all  $\tau$  and  $T$  of interest.

We demonstrate how this conceptual procedure is implemented using the syntax of UF<sup>2</sup>.

```
psi_b = uf2.up(uf2.psi0, pulse_number = 1)
psi_a = uf2.up(uf2.psi0, pulse_number = 0)
```

```
psi_ac = uf2.down(psi_a, pulse_number = 2)
P_SE = uf2.dipole_expectation(psi2_ac, psi1_b)
S_SE = uf2.polarization_to_signal(P_SE)
```

In the syntax of UF<sup>2</sup>, pulse  $a$  is pulse 0, pulse  $b$  is pulse 1, etc., and the `up` and `down` methods implement  $K_j$  and  $K_{j^*}$ , respectively. This snippet is adapted from the Jupyter notebook Example.ipynb included with the source code.

One of the strengths of UF<sup>2</sup> is that the user interacts with it by translating Feynman diagrams into nested calls to the  $\hat{K}_{i(*)}$  functions. The goal of this algorithm is to make calculating spectra as easy as writing down Feynman diagrams. We have used 3<sup>rd</sup>-order diagrams for the derivation because they are the most familiar, but calculating arbitrary order diagrams is just as straightforward. UF<sup>2</sup> is therefore easy to use and avoids the necessity of first writing out lengthy symbolic expressions for each diagram and translating each expression into code. UF<sup>2</sup> can also be used in the impulsive limit to calculate response functions, allowing this ease of use to be applied to the study of response functions as well.

### C. Optical dephasing

For closed systems,  $\mathbf{P}^{(n)}(t)$  oscillates for all time after the  $n$  optical interactions. This undamped oscillation is unphysical. In order to include optical dephasing in general, we would need to solve for the third-order density matrix,  $\rho^{(3)}(t)$ , instead of the perturbative wavepackets. From Fig. 1, three different perturbative density matrices are needed to calculate the SE, GSB and ESA terms:  $\rho_{SE}^{(3)}(t)$ ,  $\rho_{GSB}^{(3)}(t)$ , and  $\rho_{ESA}^{(3)}(t)$ . Note that for a closed system  $\rho_{SE}^{(3)}(t) = |\psi_b(t)\rangle \langle \psi_{ac^*}(t)|$ .

We can include a simple form of optical dephasing with wavefunctions by using a modified form of  $\hat{K}_j$  to describe the final,  $n^{\text{th}}$ , interaction. We include dephasing by sending  $\omega_{\phi} \rightarrow \omega_{\phi} - i\gamma$  in Eq. 13, while leaving  $\omega_{\phi'}$  unchanged. Having a constant decay rate for  $\langle \psi_{ac^*}(t) |$ , while unphysical by itself, imposes a constant decay rate on  $|\psi_b(t)\rangle \langle \psi_{ac^*}(t)| = \rho_{SE}^{(3)}(t)$ , which is physical, because  $\rho_{SE}^{(3)}(t)$  is only composed of off-diagonal elements representing optical coherences between the ground-state and the singly excited state.

### D. Useful standard approximations

The derivations in Section II B do not require the rotating wave approximation (RWA) or phase matching. However, these approximations significantly reduce the cost of the calculations, as in all perturbative spectroscopies.<sup>12,13,15</sup> Since  $|\psi^{(n)}(t)\rangle \propto (\hat{H}'(t))^n$ , the  $n^{\text{th}}$ -order wavefunction is composed of  $(2L)^n$  terms. The standard RWA and phase-matching conditions reduce the number of terms relevant to spectroscopic signals. The

RWA is valid in the limit that the pulse durations are long compared to the optical carrier frequencies, which are roughly degenerate with the optical energy gap of  $\hat{H}_0$ . In the RWA, only one of either the rotating or counter-rotating terms of Eq. 6 contributes to an interaction.<sup>13</sup> <sup>iii</sup> If the material system is dispersed over a volume much larger than the wavelength of the light then the phase-matching condition ensures that signal fields will be produced only in directions corresponding to sums and differences of wavevectors of the optical pulses.<sup>12</sup> For example Fig. 1 shows the signals produced in the rephasing direction  $(-\mathbf{k}_a + \mathbf{k}_b + \mathbf{k}_c)$  of 2D photon echo (2DPE) spectroscopy.

In addition to reducing the number of relevant diagrams, the RWA speeds up calculations because we do not need to keep track of the optical carrier frequency. This advantage is significant because UF<sup>2</sup> performs calculations using  $M$  time points with spacing  $dt$  from  $t_{j,\min}$  to  $t_{j,\max}$ . In the RWA, we set the optical carrier frequency to 0, and therefore  $dt$  only needs to resolve the pulse envelope and not the carrier frequency.

### E. Efficiency improvements

We increase the efficiency of UF<sup>2</sup> by decreasing the required dimensionality of the Hilbert space from the full size,  $N_{\text{total}}$ , to a smaller subspace of dimension  $N$ , over which the sums in Eq. 13 run. We perform this reduction in advance by determining which elements of the dipole operator  $\hat{\mu}$  will not contribute appreciably to the calculation and ignoring them. We can prune the required set of states in two ways. First, many systems have elements  $\mu_{\phi'\phi} \approx 0$ . Second, we determine which energetic transitions will not be allowed by the electric field shape.

We use Bessel's inequality to define which states  $|\phi'\rangle$  are important in order to resolve  $\mu_{\phi'\phi}$  accurately. If we begin in the eigenstate  $|\phi\rangle$ , one interaction with the dipole operator (as occurs in each application of  $K_{j(*)}$ ) yields the state

$$|\Phi\rangle = \hat{\mu}|\phi\rangle = \sum_{\phi'=1}^{N_{\text{total}}} \mu_{\phi'\phi} |\phi'\rangle,$$

where  $N_{\text{total}}$  might be infinite. We seek to restrict the sum to the smallest number of terms without significantly altering the norm of this state, which implies that we have captured all of the physically relevant states. We write the norm of  $|\Phi\rangle$  as

$$\langle\Phi|\Phi\rangle = \sum_{\phi'=1}^{N_{\text{total}}} |\mu_{\phi'\phi}|^2 = \langle\phi|\hat{\mu}^2|\phi\rangle = (\mu^2)_{\phi\phi}.$$

<sup>iii</sup> The rotating term excites a ket and de-excites a bra. The counter-rotating term de-excites a ket and excites a bra.

We find the smallest set of states such that

$$\frac{(\mu^2)_{\phi\phi} - \sum' |\mu_{\phi\phi'}|^2}{(\mu^2)_{\phi\phi}} > 1 - \epsilon, \quad (17)$$

for small  $\epsilon$ , where  $\sum'$  represents the restricted sum over the required  $N$  states. We perform this analysis for all required states  $|\phi\rangle$ .

We make this concept concrete by giving an example using the non-rephasing ground state bleach signal, which is not included in Fig. 1. The polarization field produced by that diagram is  $\mathbf{P}_{GSB,NR}^{(3)}(t) = \langle\psi^{(0)}|\hat{\mu}|\psi_{ab^*c}(t)\rangle$ . If we begin, say, in the state  $|\psi^{(0)}\rangle = |1\rangle$ , the lowest energy eigenstate of  $\hat{H}_0$ , then we must determine the smallest number of states,  $N_1$ , that satisfy Eq. 17 for  $\phi = 1$ . Then we know before calculating Eq. 13 that  $|\psi_a(t)\rangle = \sum c_{\phi,a}(t) |\phi\rangle$  will be composed of  $N_1$  terms. To find the states required to describe  $|\psi_{ab^*c}\rangle$ , we use Eq. 17 for each of the  $N_1$  states needed for  $|\psi_a(t)\rangle$ , which gives a new set of  $N_2$  states, where the subscript indicates the number of times  $\hat{\mu}$  has been applied to the initial state.

The next obvious step is to use Eq. 17 for each of those  $N_2$  states, but that step is actually unnecessary. Such an analysis would allow us to determine the  $N_3$  states required to resolve  $|\psi_{ab^*c}(t)\rangle$ . However, the final signal depends upon  $\langle\psi^{(0)}|\hat{\mu}|\psi_{ab^*c}(t)\rangle$ , and therefore the only components of  $|\psi_{ab^*c}(t)\rangle$  that matter spectroscopically are those that overlap with  $\hat{\mu}|\psi^{(0)}\rangle$ . These are precisely the  $N_1$  states we determined in the first step of this process. Therefore we only need the same  $N_1$  states when calculating  $|\psi_{ab^*c}(t)\rangle$ .

For many systems relevant to optical spectroscopy, there are well-separated manifolds of states with 0, 1, 2, etc. electronic excitations. When there are well-separated manifolds, the RWA allows a further reduction to the relevant size  $N$ . In the same non-rephasing GSB example,  $|\psi_a(t)\rangle$  is in the singly excited manifold (SEM), and  $\hat{\mu}|\psi_a(t)\rangle$  has components in both the ground state manifold (GSM) and the doubly excited manifold (DEM). In the RWA, however,  $|\psi_{ab^*c}(t)\rangle$  is only in the GSM, so  $N_2$  can be divided into its GSM and DEM portions, with only the GSM portion required for this diagram. The DEM portion of  $N_2$  is required for the ESA diagram.

In Section IV A we calculate TA spectra for a vibronic system which formally has  $N_{\text{total}} = \infty$ . With  $\epsilon = 0.001$  the spectra converge to within 1%. Using the procedure outlined here, we determine that  $N_1 = 8$  and  $N_{2,GSM} = 10$ .  $N_3$  would be 35 in order to correctly resolve  $|\psi_{ab^*c}(t)\rangle$ , but we only need to use  $N_1$  of those states to accurately reproduce the TA spectra.

In addition, we can use basic knowledge of the shape of the electric field to further restrict  $N$ . UF<sup>2</sup> performs discrete convolutions for times  $t$  with spacing  $dt$  (see end of Section II D). The spacing  $dt$  is chosen in order to resolve the shape of the electric field and implies a frequency range in which  $\hat{E}(\omega)$  is non-zero. The maximum frequency resolved is  $\frac{\pi}{dt}$ . Each element  $\mu_{\phi'\phi}$  has an associated frequency difference  $\omega_{\phi'} - \omega_{\phi}$ . If  $|\omega_{\phi'} - \omega_{\phi}| \leq \frac{\pi}{dt}$ ,

then the transition is energetically allowed by the pulse. If  $|\omega_{\phi'} - \omega_{\phi}| > \frac{\pi}{dt}$ , the transition is not allowed, and we set  $\mu_{\phi'\phi} = 0$  for all those energetically inaccessible transitions, further decreasing the relevant size  $N$ .<sup>iv</sup>

The reduction from dimension  $N_{total}$  to  $N$  not only makes  $UF^2$  more efficient but also dramatically expands the range of systems for which  $UF^2$  is tractable, as demonstrated in Section IV B.

### III. COMPUTATIONAL ADVANTAGE

We derive the computational cost of  $UF^2$ , highlighting which parameters are required for convergence and which are at the user's discretion, focusing on the case of a 2DPE rephasing spectrum, as in Eq. 8. The most important convergence parameter is the numbers of time points required to describe each optical pulse,  $M$ . If a spectrum with  $m_{\tau}$  coherence times and  $m_T$  population times is desired, then we show in Section III A that the cost of using  $UF^2$  to calculate the 2DPE spectrum scales as  $\sim m_{\tau} m_T N^2 M$  in the limit that  $N$  is large. By comparison, Section III B considers the cost of recently published symbolic solutions that, for particular pulse shapes, avoid the need for the time integrations, as in Eq. 13.<sup>2-4</sup> Though their time propagation is evaluated analytically, their cost still scales as  $\sim m_{\tau} m_T N^3$ . For sufficiently large systems,  $UF^2$  is significantly faster even than these symbolic methods, which we show can occur even at quite modest system sizes near  $N \approx 100$ .  $UF^2$ , of course, allows treatment of arbitrary pulse shapes and pulse overlap effects.

#### A. Computational cost analysis

We derive the computational cost of calculating the 2DPE rephasing signal. We briefly outline what is required for this calculation by working backwards from Eq. 8. Symbols used in this section are summarized in Table I. Calculation of the signal requires that we determine  $\mathbf{P}_{2D}^{(3)}(\tau, T; \omega)$  at the desired values of  $\tau$  and  $T$ .  $UF^2$  directly calculates  $\mathbf{P}_{2D}^{(3)}(t)$  for one pair of  $(\tau, T)$  at a time. We calculate  $\mathbf{P}_{2D}^{(3)}(t)$  at sufficient time points,  $M_t$ , in order to obtain the desired frequency resolution. The cost of the FFT to obtain  $\mathbf{P}_{2D}^{(3)}(\omega)$  from  $\mathbf{P}_{2D}^{(3)}(t)$  is negligible compared with the other costs of  $UF^2$ . Assuming that we require the polarization field at  $m_{\tau}$  values of  $\tau$  and  $m_T$  values of  $T$ , the cost of the algorithm is  $m_{\tau} m_T \text{Cost}(\mathbf{P}_{2D}^{(3)}(t))$ .  $\mathbf{P}_{2D}^{(3)}(t)$  is a sum of three Feynman

Table I. Summary of symbols used to describe computational cost of a 2DPE rephasing signal.  $M$  and  $M_t$  are convergence parameters.

Symbol	Definition
$N$	Dimension of relevant Hilbert space
$M$	Number of time points to resolve each optical pulse
$M_t$	Number of time points to resolve polarization field
$m_{\tau}$	Number of desired values of $\tau = t_b - t_a$
$m_T$	Number of desired values of $T = t_c - t_b$

diagrams (see Fig. 1), each of which has equal cost and is calculated separately. Therefore we focus on the cost of  $\mathbf{P}_{SE}^{(3)}(t)$ , and multiply by three at the end.

Using Eq. 16, the cost of  $\mathbf{P}_{SE}^{(3)}(t)$  can be broken into three parts: calculating the two necessary perturbed wavepackets, and the cost of the dipole matrix element of those wavepackets. Since  $|\psi_b(t)\rangle = K_b |\psi^{(0)}(t)\rangle$  and  $|\psi_{ac^*}(t)\rangle = K_{c^*} K_a |\psi^{(0)}(t)\rangle$ , the cost of calculating these wavepackets is the cost of three calls to the  $K_{j(*)}$  operator, which is the heart of  $UF^2$ . To determine the cost of  $K_{j(*)}$ , we first explain how wavefunctions are stored in  $UF^2$ .

Each wavefunction is represented in the same way. For example, the second-order wavefunction is

$$|\psi_{ac^*}(t)\rangle = \sum_{\phi=1}^N e^{-i\omega_{\phi} t} c_{\phi, ac^*}(t) |\phi\rangle, \quad (18)$$

where each of the  $c_{\phi, ac^*}(t)$  are calculated at  $M$  time points that are evenly spaced by  $dt$ . The spacing  $dt$  is determined by the shape of the pulse amplitudes  $A_i(t)$  (see Eq. 4). Recall from Eq. 15 that  $c_{\phi, pj(*)}(t)$  only varies in the interval  $(t_{j, \min}, t_{j, \max})$ . Outside of this interval,  $c_{\phi, pj(*)}$  is constant and can thus be trivially extended to any time points needed. For simplicity in this discussion we assume that  $t_{j, \max} - t_{j, \min}$  is the same for each pulse. Since we are discretizing a continuous convolution integral,  $M = \frac{t_{j, \max} - t_{j, \min}}{dt}$  is a convergence parameter. For any electric field that does not go strictly to zero, the choice of  $t_{j, \min}$  and  $t_{j, \max}$  must also be checked for convergence.

Inspecting Eq. 13, the cost of  $\hat{K}_j$  is  $N$  times the cost of evaluating  $\theta * y_{\phi}$  to obtain  $c_{\phi, pj}$ . Calculating  $y_{\phi}$  is dominated by the cost of taking the dipole-weighted sum over  $\phi'$  (we are neglecting the small cost of multiplying by the pulse amplitude  $\varepsilon_j^{(*)}$  and by the time evolution factors  $e^{i\omega_{\phi}(t-t')}$  and  $e^{-i\omega_{\phi'}(t-t')}$ ). The sum has a cost of  $\alpha NM$ , where  $\alpha$  is the cost of adding two complex numbers. We calculate  $\theta * y_{\phi}$  using the convolution theorem and the FFT. Since we are interested in the linear convolution, we evaluate  $\theta$  at  $2M - 1$  points and zero-pad  $y_{\phi}$  to be size  $2M - 1$ . We only calculate the FFT of  $\theta$  once, making

<sup>iv</sup> This procedure is not only helpful for speeding up calculations, but actually necessary for accurate calculations. If included, any frequency differences  $|\omega_{\phi'} - \omega_{\phi}| > \frac{\pi}{dt}$  will cause spurious signals to appear in the signal due to aliasing.

that cost negligible. We calculate the FFT of  $y_\phi$ , which has a cost of  $\beta'(2M-1)\log_2(2M-1) \approx 2\beta'M\log_2 M + O(M)$ , where  $\beta'$  depends upon the implementation of the FFT.<sup>v</sup> We multiply the two ( $\tilde{y}_\phi$  and  $\tilde{\theta}$ ), which has cost  $O(M)$ , and then take the inverse FFT of the product (cost  $2\beta'M\log_2 M + O(M)$ ). Thus the total cost of the convolution is  $4\beta'M\log_2 M + O(M)$ . The cost to obtain one coefficient  $c_{\phi,pj(*)}(t)$  to highest order in  $M$  and  $N$  is then

$$\text{Cost}(c_{\phi,pj(*)}(t)) = \alpha NM + \beta M \log_2 M + O(M), \quad (19)$$

where  $\beta = 4\beta'$ . The cost of calling  $K_{j(*)}$  is then  $N\text{Cost}(c_{\phi,pj(*)})$  or

$$\text{Cost}(K_{j(*)}) = NM(\alpha N + \beta \log_2 M) + O(MN), \quad (20)$$

and so  $\text{Cost}(K_{j(*)}) \sim MN^2$  for large  $N$ .

The cost of calculating the two perturbative wavepackets,  $|\psi_b(t)\rangle$  and  $|\psi_{ac*}(t)\rangle$ , needed for the SE diagram is  $3\text{Cost}(K_{j(*)})$ . Once we know each of the necessary wavefunctions at its  $M$  time points, we evolve the wavefunctions using  $\hat{U}_0$  to include  $M_t$  times points, in order to obtain the desired frequency resolution of the final spectrum. The  $M_t$  points are spaced by the same  $dt$ , and span from just before the last pulse arrives,  $t_{c,\min}$ , until the signal has decayed to an appropriate cut-off due to the optical dephasing described in Section II C. This evolution is of negligible cost since we know the exact form of  $\hat{U}_0(t)$ .

We then calculate the expectation value  $\mathbf{P}_{SE}^{(3)}(t) = \langle \psi_{ac*}(t) | \hat{\mu} | \psi_b(t) \rangle$ , which has a cost of  $\alpha N^2 M_t$ , so in total  $\text{Cost}(\mathbf{P}_{SE}^{(3)}(t)) = \alpha N^2 M_t + 3\text{Cost}(K_j)$ . The signal must be calculated at  $m_\tau \tau$  points and  $m_T T$  points, so we arrive at the full cost of the SE signal

$$\begin{aligned} \text{Cost}(S_{SE}(\tau, T, \omega)) = m_\tau m_T \left( \alpha N^2 M_t \right. \\ \left. + 3NM(\alpha N + \beta \log_2 M) \right) \end{aligned} \quad (21)$$

to highest order in  $N$  and  $M$ . Note that each of the diagrams in Fig. 1 has the same cost, so the cost of the full rephasing signal is simply 3 times this cost. In Section IV, we calculate the TA signal, which is composed of four Feynman diagrams, each of which has the same cost as in Eq. 21, with  $m_\tau = 1$ , since the TA signal has  $\tau = 0$ . In

Section IV we use  $M = 138$  and  $M_t \approx 1000$ , in order to achieve convergence of better than 1%.<sup>vi</sup>

Assuming that  $M \approx M_t$ , and that  $\alpha N \gg \beta \log_2 M$ ,<sup>vii</sup> the total cost scales as

$$\text{Cost}(S_{SE}(\tau, T, \omega)) \sim m_\tau m_T N^2 M.$$

Note that  $\text{UF}^2$  assumes that the eigenvalues and eigenstates of  $\hat{H}_0$  have already been attained. If the dimension of  $\hat{H}_0$  is  $N_{\text{total}}$ , then the computational cost of that diagonalization can scale as  $N_{\text{total}}^3$ , which can far exceed the cost of  $\text{UF}^2$  itself, especially since  $\text{UF}^2$  reduces the relevant Hilbert space dimension  $N$  so aggressively. For many systems, however,  $\hat{H}_0$  is sparse, allowing efficient iterative methods to be used to find its eigenvalues and eigenstates. The scaling of those algorithms is beyond the scope of this manuscript, but we simply note that there are many important cases where solving  $\hat{H}_0$  is not computationally limiting. For example, vibronic systems of coupled chromophores, each with local harmonic oscillators, are extremely sparse, and an efficient algorithm to determine their eigenstates will be detailed separately.

## B. Comparison to symbolic solutions

Perlik et al. have demonstrated that for Lorentzian pulse shapes, the triple-nested convolution integrals in Eq. 10 can be solved exactly to give symbolic expressions for  $P^{(3)}(t)$  (and therefore any signal of the form shown in Eqs. 8 and 9).<sup>3</sup> Smallwood et al. report symbolic solutions for the Fourier transforms of Eq. 8 with respect to  $\tau$  or  $T$ :  $S^{(3)}(\omega_\tau, T, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\tau e^{i\omega_\tau \tau} S^{(3)}(\tau, T, \omega)$  and  $S^{(3)}(\tau, \omega_T, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dT e^{i\omega_T T} S^{(3)}(\tau, T, \omega)$ , assuming Gaussian or chirped Gaussian pulse shapes.<sup>4</sup> Johnson et al. report symbolic expressions for the frequency-integrated TA signal

$$S_{FI,TA}(T) = \int d\omega S_{TA}(T, \omega), \quad (22)$$

which are valid for Gaussian pulses in the approximation that the pulses do not overlap in time.<sup>2</sup> Two of these papers include a simple model for optical dephasing and population decay,<sup>3,4</sup> whereas the third paper considers only closed systems.<sup>2</sup> All three can be used to study closed systems, and all comparisons here use the symbolic formulae to study closed systems.

In all of these cases, the symbolic expressions remove the need for time- or frequency-integration, but they still must sum over the intermediate states in the relevant

<sup>v</sup> We find that when calculating the convolution between a step-function  $\theta(t)$  and another function  $y(t)$ , we achieve convergence much more quickly when we use an odd number of time points, with half positive, half negative, and the point  $t = 0$  with associated value  $\theta(0) = 0.5$ . The FFT is fastest for arrays of length  $2^k$ , which is impossible in this case. To keep  $\beta'$  small, we use the FFTW library and pick a size  $2M - 1$  that has only small prime-factors.

<sup>vi</sup> The value of  $M_t$  depends upon the pulse shape and the optical dephasing  $\gamma$ . We resolve the polarization field from  $t_{c,\min}$  until  $t_c + 5/\gamma$ .

<sup>vii</sup> In our implementation of this algorithm,  $\beta \log_2 M / \alpha \approx 150$ .



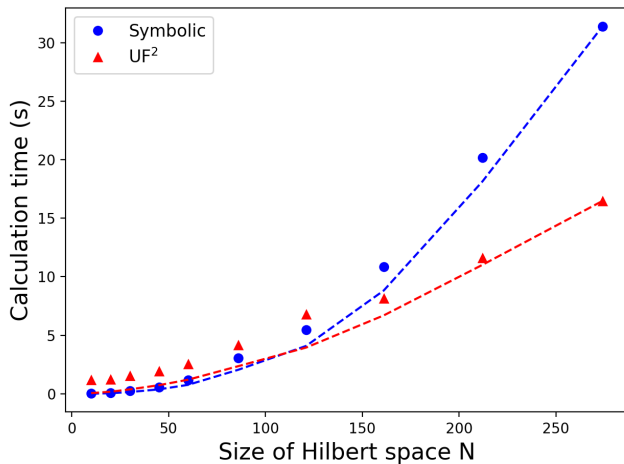


Figure 4. Time to compute the frequency-integrated pump-probe signal at 100 delay times using the symbolic solution from Ref. 2 (blue circles) and UF<sup>2</sup> (red triangles) using a laptop computer. Dashed curves  $t \propto N^3$  (blue) and  $t \propto N^2$  (red) show the predicted scaling. UF<sup>2</sup> begins to out-perform the symbolic solution between  $N = 100$  and  $N = 200$ . It is also worth noticing that UF<sup>2</sup> can calculate spectra for  $N < 1000$  in about a minute or less.

Feynman diagrams. Symbolic expressions have been obtained for the desired signal due to a single pathway of a particular diagram. The total signal from a diagram is then obtained by summing over all pathways. For example,

$$S_{SE}(\tau, T, \omega) = \sum_{\alpha} S_{SE, \alpha},$$

where  $\alpha = \{\phi, \phi', \phi'', \phi'''\}$  (see Fig. 1). Assuming that the system begins in the state  $|\phi\rangle$ , the sum over  $\alpha$  is a triple sum over all possible values of  $\phi', \phi'', \phi'''$ . Therefore the computational cost of this sum scales as  $N^3$ . We demonstrated in Section III A that UF<sup>2</sup> scales as  $N^2$  for large  $N$ . This improved scaling occurs because UF<sup>2</sup> does not handle each pathway  $\alpha$  individually. For sufficiently large system sizes, UF<sup>2</sup> calculates spectra faster than would be possible using these types of symbolic solutions. We have implemented the symbolic solution from Ref. 2 for a test problem and found that our implementation of UF<sup>2</sup> performs better for  $N \gtrsim 100$ , as shown in Figure 4. In Section IV A we show the TA spectra for a system with  $N \approx 20000$ , where UF<sup>2</sup> calculated the spectra in about 9 hours on a laptop. The symbolic solution would have required about 2 days for the same calculation.

Aside from scaling better, UF<sup>2</sup> offers the advantage that it can predict spectra for any pulse shape, whether or not the analytical form of that pulse shape is known. UF<sup>2</sup> simply requires that the pulse be known on a discrete time interval, and that the pulse go to zero outside of that domain.

### C. Comparison to time-domain wavepackets

UF<sup>2</sup> is significantly faster at simulating bound systems than the time-domain wavepackets (TDW) method as detailed in Ref. 7, which uses the split operator method to evolve the perturbative wavefunctions forward in time according to  $\hat{H}_0$ . Briefly, the split operator method performs part of the time evolution in position space, and part of the time evolution in momentum space. The perturbative interaction with the pulses is also handled numerically. The time step  $dt$  must be short enough to resolve both the system and perturbative dynamics.

Since UF<sup>2</sup> handles the evolution due to  $\hat{H}_0$  exactly, using the known form of  $\hat{U}_0(t)$ , UF<sup>2</sup> can evolve the wavefunctions forward in the absence of the pulses for near zero cost, greatly reducing the number of numerical integration steps. In addition, the time step  $dt$  used by UF<sup>2</sup> to evaluate the convolutions is determined solely by the pulse shapes, and need not be small enough to accurately resolve the dynamics in  $\hat{H}_0$ . These advantages allow UF<sup>2</sup> to dramatically outperform TDW in certain cases. For the cases shown in Section IV, which involve harmonic vibrational modes, we have found that UF<sup>2</sup> calculates spectra at least  $10^5$  times faster than TDW.

On the other hand, UF<sup>2</sup> is limited in scope. It requires that the eigenbasis of  $\hat{H}_0$  be discrete, and therefore cannot simulate bond-breaking, isomerization, or chemical reactions, where a continuous eigenbasis may be required. TDW represents wavefunctions in position space and merely requires that the potential energy surface of  $\hat{H}_0$  be known, with no diagonalization of  $\hat{H}_0$ . In these cases where UF<sup>2</sup> would break down, TDW would shine.

## IV. EXAMPLE

In this section we demonstrate the power of UF<sup>2</sup> in two different regimes. In Section IV A, we study a system that has a small Hilbert space, where wavefunctions may be represented using  $N = 10$  terms. This small size allows us to generate TA spectra in a matter of seconds, and therefore it is easy to run fast parameter sweeps and map out how spectroscopic observables change. In the second subsection, we show the TA spectra for a system that has a large Hilbert space, where wavefunctions are represented using  $N \approx 20000$  terms. Even in such a large Hilbert space, UF<sup>2</sup> produces isotropically averaged TA spectra in about 9 hours on a laptop. This second example pushes the memory constraints of a personal computer and gives a realistic sense of the upper limits of the current implementation of UF<sup>2</sup>. However, the algorithm could certainly be expanded to work with larger systems using distributed memory computing.

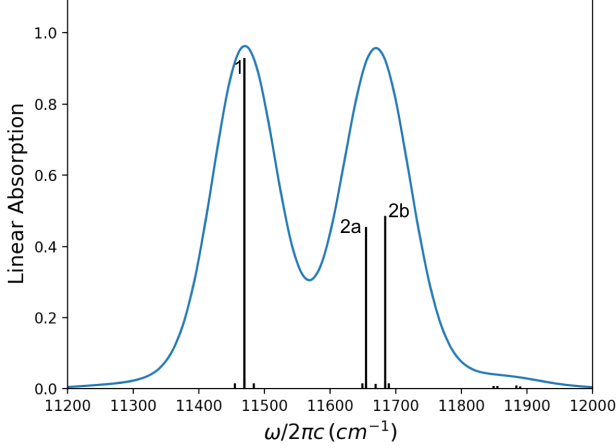


Figure 5. Linear absorption stick spectrum calculated using the same eigenstates as in Ref. 16. The frequencies of the stick spectra are the eigenenergies of the Hamiltonian Eq. 23, with  $\omega_i = 200 \text{ cm}^{-1}$  and  $\lambda_i = 5 \text{ cm}^{-1}$  for each monomer subunit. The two subunits have dipoles  $|\mu_{ag}| = |\mu_{bg}|$  and  $\mu_{ag} \cdot \mu_{bg} = 0$ . The heights of the sticks are  $|\hat{\mu}_{\phi'\phi}|^2$ . The spectrum is computed at a temperature of 80 K, using the 3 lowest energy ground vibrational states,  $|00\rangle, |01\rangle, |10\rangle$ . The blue curve represents the absorption cross-section, which is obtained by broadening the stick spectrum with homogeneous and inhomogeneous linewidths of  $\gamma = 10 \text{ cm}^{-1}$  and  $\sigma = 45 \text{ cm}^{-1}$ .

### A. Rapid parameter sweep of small system

Here we explore nonlinear optical spectra for a model system presented by Tiwari and Jonas.<sup>16</sup> Their model system is inspired by the Fenna-Matthews-Olson complex<sup>17</sup> and consists of a molecular dimer formed from two electronic two-level systems, each locally coupled to a single harmonic vibrational mode. The electronic Hamiltonian has a ground state  $|g\rangle$ , excitations on sites  $|a\rangle$  or  $|b\rangle$ , and a doubly-excited state  $|f\rangle$ , with Hamiltonian

$$H_e = E_g|g\rangle\langle g| + E_a|a\rangle\langle a| + E_b|b\rangle\langle b| + J(|a\rangle\langle b| + |b\rangle\langle a|) + (E_a + E_b)|f\rangle\langle f|,$$

where  $E_{a,b}$  are site energies and  $J$  is the electronic coupling. The vibrations have Hamiltonian

$$H_{ph} = \frac{p_a^2}{2} + \frac{p_b^2}{2} + \frac{1}{2}\omega_a^2 q_a^2 + \frac{1}{2}\omega_b^2 q_b^2,$$

where  $p_{a,b}$  and  $q_{a,b}$  are the generalized momentum and position for each vibration, with frequencies  $\omega_{a,b}$ . Standard linear coupling of the position of each oscillator to its site excitation gives

$$H_{e-ph} = \omega_a^2 d_a q_a |a\rangle\langle a| + \omega_b^2 d_b q_b |b\rangle\langle b|,$$

where  $d_{a,b}$  are the coupling strengths, corresponding to reorganization energy

$$\lambda_i = \frac{1}{2}\omega_i d_i^2,$$

and the total system Hamiltonian is

$$H_0 = H_e + H_{ph} + H_{e-ph}. \quad (23)$$

The system couples to optical pulses in the electric-dipole approximation, with transition dipole matrix elements  $\mu_{ga}$  and  $\mu_{gb}$ . Following Ref. 16, we consider the homodimer to have nearly identical subunits, with identical vibrational frequencies  $\omega_a = \omega_b$  and couplings  $d_a = d_b$ , and with  $E_b - E_a = 150 \text{ cm}^{-1}$ . We consider that  $\mu_{ga}$ ,  $\mu_{gb}$  have the same magnitude but extend the model by varying the angle  $\theta$  between these transition dipoles, corresponding to varying the angle between the two subunits. Due to the small value of  $d$  considered in Ref 16 (near 0.22), only a small number of vibrational states contribute to third-order spectra, allowing truncation of the vibrational Hilbert space, as described in Sec. II E.

Reference 16 demonstrates how the vibrational coupling  $d$  splits one exciton peak into two vibronic peaks. We plot the linear absorption spectrum for this system in Fig. 5, choosing lineshape parameters to visually reproduce results from Ref. 16. We label the peaks as indicated in Fig. 5 following Tiwari and Jonas. There are two broad peaks centered around  $f_1 = 11469.5 \text{ cm}^{-1}$  and  $f_2 = 11669.6 \text{ cm}^{-1}$ , with the latter composed of two peaks, 2a and 2b, separated by  $\Omega_1 = 29 \text{ cm}^{-1}$ . However, this splitting is invisible to linear absorption because peaks 2a and 2b smear together.

$\Omega_1$  is, however, the dominant beat frequency in a TA experiment, shown in Fig. 6, corresponding to oscillations repeating about every 1.15 ps. The TA signal is detected as two broad features centered at  $f_1$  and  $f_2$ . Notably, the oscillations centered around  $f_1$  are roughly  $90^\circ$  out of phase with those at  $f_2$ , and there is a node halfway between them. This type of nodal feature is widely observed and has been studied in many vibrational systems.<sup>18–20</sup>

Each TA spectrum is isotropically averaged, with the same homogeneous and inhomogeneous linewidths in the linear absorption spectrum. Each TA spectrum, which is of the form  $S_{TA}(T, \omega)$ , consists of  $m_T = 90$  delay times, and  $M_t = 1000$  detection frequency points, and takes about 5 seconds to generate on a 2.3GHz Macbook Pro. The equivalent calculation using the TDW code (see Section III C) would take almost 6 days to converge to the same accuracy if run on the same machine.

The efficiency of UF<sup>2</sup> enables rapid study of a wide range of parameters. The frequency-integrated TA signal from Eq. 22 is particularly sensitive to electric field shape.<sup>2,21</sup> Figure 7(a,b) shows frequency-integrated TA spectra with dipole angles  $\theta = 20^\circ, 58^\circ$  and three optical pulse durations. The responses with different pulse shapes are quite different in both cases. In order to study the differences over a wide range of  $\theta$  we study the Fourier transform of  $S_{FI,TA}(T)$  with respect to the delay time:

$$\tilde{S}_{FI,TA}(\omega_T) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dT S_{FI,TA}(T) e^{i\omega_T T},$$

and track the response at  $\omega_T = \Omega_1$  and  $\omega_T = \Omega_2 = f_{2a} - f_1 = 185.5 \text{ cm}^{-1}$  in Fig. 7(c,d). A nearly impulsive pulse

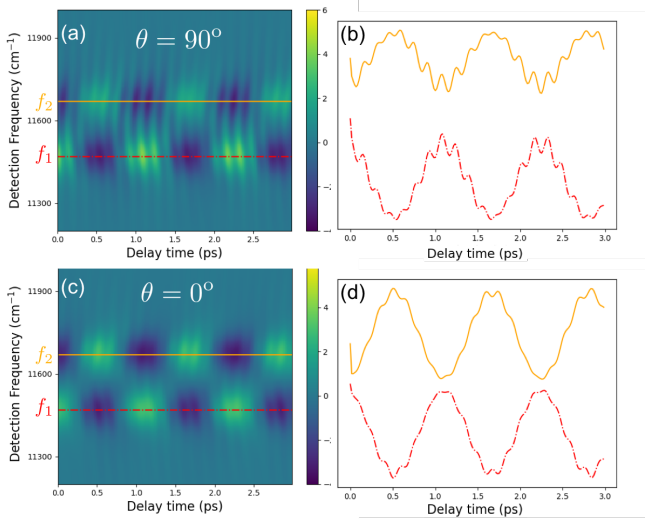


Figure 6. Transient absorption with Gaussian pump and probe pulses centered on  $f_2$  and with FWHM of 12 fs. (a) Dipole angle  $\theta = 90^\circ$ , (b) linecuts of (a) at  $f_2$  (orange) and  $f_1$  (red, dash-dotted), (c) Dipole angle of  $\theta = 0^\circ$ , (d) linecuts of (c) at  $f_2$  (orange) and  $f_1$  (red, dash-dotted). In both cases the low-frequency beats are out of phase between  $f_1$  and  $f_2$ . The high-frequency beats are in phase between  $f_1$  and  $f_2$  for  $\theta = 90^\circ$  and are reduced in amplitude and out of phase for  $\theta = 0^\circ$ .

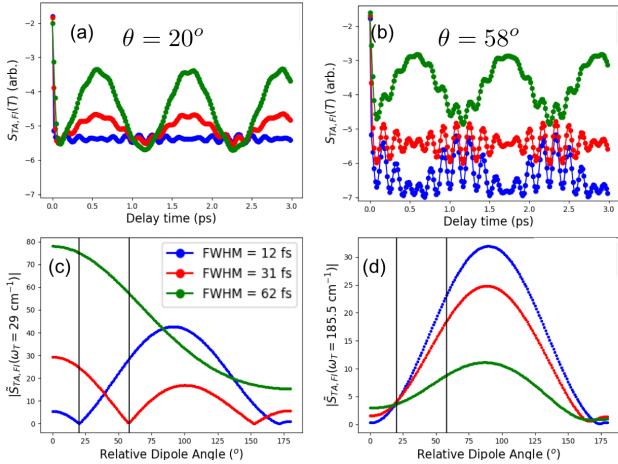


Figure 7. Demonstration of parameter variations for the same system as in Fig. 6. Panels (a) and (b) show frequency-integrated transient absorption (FITA) spectra for (a)  $\theta = 20^\circ$  and (b)  $\theta = 58^\circ$  with three different Gaussian pulse durations. There are two prominent time scales present: the  $\Omega_1$  beat (corresponding to the 1.15 ps oscillation) and the  $\Omega_2$  and  $\Omega_3$  beats (corresponding to the 180 fs and 156 fs oscillations). Panels (c) and (d) show the magnitude of these beats of the FITA spectra at (c)  $\omega_T = \Omega_1$  and (d)  $\omega_T = \Omega_2$  as a function of  $\theta$ . Black vertical lines indicate  $\theta$  from panels (a) and (b). Note that at  $\omega_T = \Omega_1$ , the FITA magnitude shows a dramatic dependence on both dipole angle and electric field shape, demonstrating the importance of taking into account electric field shapes. The signals at  $\omega_T = \Omega_2$  are less sensitive.

(12 fs FWHM - blue curves) gives a dramatically different response as compared to somewhat longer pulses (35 fs - red curves, 70 fs - green curves). Notice that the longest pulse considered here is still short compared to the fastest time scale of the system,  $\Omega_3$ , corresponding to 156 fs, so the large changes of the signals with pulse duration are a warning sign that accurate study of finite-pulse effects is important to correctly model nonlinear spectra. This study is easily tractable using  $\text{UF}^2$  but would have taken more than a year using our TDW implementation on the same computer.

## B. TA spectrum of a large system

Here we show a TA spectrum for a system that pushes the limits of our current implementation of  $\text{UF}^2$ , with  $N \approx 20000$ . The system presented here is a 6-mode harmonic oscillator, with each mode linearly coupled to a single electronic excitation. The Hamiltonian is

$$H = E_g|g\rangle\langle g| + E_e|e\rangle\langle e| + \frac{1}{2} \sum_{i=1}^6 (p_i^2 + \omega_i^2 q_i^2 + 2\omega_i^2 d_i q_i |e\rangle\langle e|),$$

with  $\omega_i = \{0.1976, 0.3032, 0.5123, 0.9124, 0.7432, 1\}$  in units of the largest vibrational frequency  $\omega_0$ , and  $d_i = \{1.4, 1, 0.8, 0.75, 0.7, 0.6\}$ . The optical pulse is a Gaussian with duration  $0.2\omega_0^{-1}$  and centered on  $E_e - E_g - \lambda$ , with

$$\lambda = \frac{1}{2} \sum_{i=1}^6 \omega_i d_i^2 = 1.297\omega_0.$$

The spectrum took about 9 hours to compute, demonstrating that  $\text{UF}^2$  can be used to study quite large systems. Note that if we had kept all of the states with  $\hbar\omega_\phi < 20\omega_0$ , with the cutoff based upon the bandwidth of the pulse, we would have had  $N_{\text{total}} > 7 \times 10^6$  states; the dipole-based pruning described in Sec. II E reduces this dimensionality to  $N \approx 20000$ , making the calculation tractable. This Hamiltonian is block diagonal, since the vibrational modes do not mix, and can therefore be solved with relative ease. In this case, a significant portion of the pruning can be performed by hand since the harmonic modes are uncoupled, but more complicated systems benefit from the automated dimension reduction performed by  $\text{UF}^2$ .

## V. CONCLUSION

$\text{UF}^2$  is a computationally efficient method for calculating perturbative nonlinear spectroscopies of systems with finite dimensional relevant Hilbert spaces. It includes all finite-duration and pulse-overlap effects, enabling accurate modeling of and fitting to nonlinear spectroscopic data. The current introduction has focused on closed

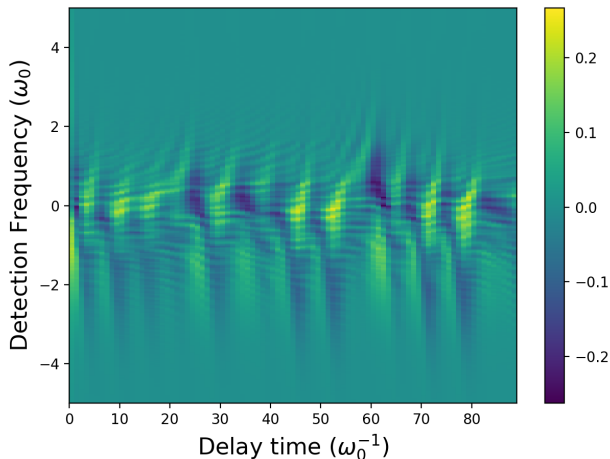


Figure 8. TA spectrum for a 6-mode harmonic oscillator with  $N \approx 20000$ .

systems, but  $UF^2$  can be readily extended to open systems with time-independent Liouvillians, in which case its cost scales as  $N^4$  rather than  $N^2$ , which is still excellent scaling. We have shown the methods that can be used to aggressively reduce the required dimension  $N$ , which enables  $UF^2$  to be used in surprisingly large problems.

The intent of the publicly available  $UF^2$  code is to present a package that makes it easy to translate Feynman diagrams into code. Since  $UF^2$  readily enables consideration of arbitrary pulse shapes, the effects of pulse durations, chirps, etc. on nonlinear spectroscopies can now be included as a matter of course in analyses of experiments.

## ACKNOWLEDGMENTS

We acknowledge helpful conversations with Joel Yuen-Zhou, Ivan Kassal, Luc Robichaud, and Eduard Dumitrescu. We also acknowledge funding from the Natu-

ral Sciences and Engineering Research Council of Canada and the Ontario Trillium Scholarship.

## REFERENCES

- <sup>1</sup>A. Perdomo-Ortiz, J. R. Widom, G. A. Lott, A. Aspuru-Guzik, and A. H. Marcus, *J. Phys. Chem. B* **116**, 10757 (2012).
- <sup>2</sup>A. S. Johnson, J. Yuen-Zhou, A. Aspuru-Guzik, and J. J. Krich, *The Journal of Chemical Physics* **141**, 244109 (2014).
- <sup>3</sup>V. Perlík, J. Hauer, and F. Šanda, *J. Opt. Soc. Am. B* **34**, 430 (2017).
- <sup>4</sup>C. L. Smallwood, T. M. Autry, and S. T. Cundiff, *J. Opt. Soc. Am. B* **34**, 419 (2017).
- <sup>5</sup>J. D. Bell, R. Conrad, and M. E. Siemens, *Opt. Lett.* **40**, 1157 (2015).
- <sup>6</sup>T. N. Do, M. F. Gelin, and H.-S. Tan, *The Journal of Chemical Physics* **147**, 144103 (2017).
- <sup>7</sup>J. Yuen-Zhou, J. J. Krich, I. Kassal, A. S. Johnson, and A. Aspuru-Guzik, *Ultrafast Spectroscopy*, 2053-2563 (IOP Publishing, 2014).
- <sup>8</sup>M. F. Gelin, D. Egorova, and W. Domcke, *The Journal of Chemical Physics* **123**, 164112 (2005).
- <sup>9</sup>M. Gelin, D. Egorova, and W. Domcke, *Chemical Physics* **312**, 135 (2005).
- <sup>10</sup>M. F. Gelin, D. Egorova, and W. Domcke, *Accounts of Chemical Research* **42**, 1290 (2009).
- <sup>11</sup>M. F. Gelin, D. Egorova, and W. Domcke, *The Journal of Chemical Physics* **131**, 194103 (2009).
- <sup>12</sup>S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (OXFORD UNIV PR, 1999).
- <sup>13</sup>M. Z. Peter Hamm, *Concepts and Methods of 2D Infrared Spectroscopy* (CAMBRIDGE UNIV PR, 2011).
- <sup>14</sup>E. J. Heller, *Accounts of Chemical Research* **14**, 368 (1981).
- <sup>15</sup>M. Cho, *Two-Dimensional Optical Spectroscopy* (CRC Press, 2009).
- <sup>16</sup>V. Tiwari and D. M. Jonas, *The Journal of Chemical Physics* **148**, 084308 (2018).
- <sup>17</sup>R. E. Fenna and B. W. Matthews, *Nature* **258**, 573 (1975).
- <sup>18</sup>M. Liebel, C. Schnedermann, T. Wende, and P. Kukura, *The Journal of Physical Chemistry A* **119**, 9506 (2015).
- <sup>19</sup>S. D. McClure, D. B. Turner, P. C. Arpin, T. Mirkovic, and G. D. Scholes, *The Journal of Physical Chemistry B* **118**, 1296 (2014).
- <sup>20</sup>J. A. Cina, P. A. Kovac, C. C. Jumper, J. C. Dean, and G. D. Scholes, *The Journal of Chemical Physics* **144**, 175102 (2016).
- <sup>21</sup>J. Yuen-Zhou, J. J. Krich, and A. Aspuru-Guzik, *The Journal of Chemical Physics* **136**, 234501 (2012).