Open source, heterogeneous, nonlinear optics simulation

NICHOLAS KARPOWICZ*

Max Planck Institute of Quantum Optics, Hans-Kopfermann-Str. 1, 85748 Garching, Germany
nicholas.karpowicz@mpq.mpg.de

Abstract: The spatio-temporal evolution of a laser field taking part in a nonlinear optical interaction can be challenging to simulate, yet forms the basis for many experiments in ultrafast optics. To allow better insight into these phenomena, a program for nonlinear optics simulations is described, which can run on multiple hardware platforms, and is performant and open source. It was designed to deal with a number of complex problems in light-matter interaction accurately and reproducibly. The open source code allows for extensive cross-checking of its results by other researchers and growth of its capabilities over time, as well as serving to make the simulations associated with ultrafast experiments more broadly reproducible.

1. Introduction

One of the great strengths of performing experiments in optics using ultrafast lasers is the ability to modify the central wavelength and bandwidth of the light being used through nonlinear processes [1]. In the case of simple, non-cascaded nonlinear processes, analytical solutions to the evolution of the amplitude of the nonlinearily-generated wave are possible, allowing for straightforward predictions and optimizations of a given optical layout. However, in the presence of competing nonlinearities, cascaded processes, and other effects where an analytical solution is often not possible, it is common to perform a numerical solution of the nonlinear wave equation. There are multiple possible approaches to this, depending on the problem being solved, with varying degrees of approximations, numbers of spatial dimensions, assumptions regarding the underlying light-matter interaction. As a result, it may become difficult to compare to or reproduce the results of publications where the source code that produced the data forming the basis of the conclusions is not publicly available.

Effects which we may want to reproduce inside a simulation include amplification [2,3], generation of optical harmonics [4], the generation of a continuum through filamentation [5], pulse metrology [6,7], and others. Having a consistent framework capable of including all of these effects benefits extrapolation of nonlinear optical properties across experimental contexts and improve the predictability of nonlinear optical systems.

A recent example of where precise simulations are necessary to interpret measurements is recent work in attosecond physics, retrieving information carried by the directly-resolved electric field of high-frequency light in the near-infrared and visible [8,9]. Some field metrology techniques, such as electro-optic sampling (EOS) [10,11] and generalized heterodyne optical sampling techniques (GHOST) [12] use nonlinear optical effects to provide direct access to the electric field. The response function of these measurement approaches depends on the underlying light-matter interaction inside the nonlinear crystal, which is often non-resonant and low-order, making it approachable through straightforward calculation in nonlinear optics, but a fully-detailed reconstruction of the measurement apparatus requires a high-fidelity implementation of the relevant effects. Thus, a well-tested nonlinear optics platform can provide a useful link between the measured quantity and the real electric field.

The simulation presented here is implemented such that it can take advantage of diverse hardware platforms simultaneously, is optimized for performance, and is integrated with a
graphical user interface in order to provide direct visual feedback regarding the results of a simulation.

2. Physical model and underlying assumptions

Two physical models are implemented: finite-difference time domain (FDTD) where the electromagnetic field is represented on a spatial grid and propagated forward in time, and the unidirectional pulse propagation equation (UPPE), where the field is represented on a grid with two spatial dimensions and a time dimension, and this structure evolves in space as it propagates through the medium. FDTD is primarily used in thin crystals where reflections of the wave at interfaces have a significant effect on the dynamics, while the UPPE, being generally faster to solve and less resource-intensive is used for propagation through thicker media.

The physical model of the FDTD mode is simply the solution of Maxwell’s equations, with the interaction with matter mediated by a space and time dependent current induced by a set of classical oscillators whose equations of motion are solved, coupled to the electromagnetic field. The nonlinear tensors describing second and third order nonlinear optical effects are integrated into this system of oscillators such that the nonlinear optical effects have a dispersion consistent with the linear dispersion of the materials.

The boundaries of the grid are periodic in the transverse direction. In the case of the UPPE, the boundary is also periodic in the temporal direction (such that if a pulse falls behind the frame of the grid, which is locked to the phase velocity of the pulse, it will come back on the front side). In FDTD mode, the boundaries along the z-axis are handled by an absorbing layer implemented as a modification of the spatial-derivative operators such that the light propagates unidirectionally out of the grid with minimal reflections.

2.1. Slowly evolving wave approximation

The UPPE is based on the assumption that the light primarily propagates along the z-axis [1,13–15]. In this case, simplification of the Laplacian operator $\nabla^2$ can be made, into the sum of an unmodified transverse Laplacian $\nabla^2_\perp = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, and a modified z-derivative, obtained via the assumption that a given plane-wave component of the field can be described by the product of a slowly-evolving amplitude with a plane wave.

The resulting equation, depending on the level of approximation, can be written as

$$\frac{\partial}{\partial z} \tilde{\mathbf{E}}(\mathbf{x}, \omega) = i(k + \frac{1}{2k} \nabla^2_\perp) \tilde{\mathbf{E}}(\mathbf{x}, \omega) + \frac{i\omega}{2\epsilon_0 \hbar(\omega) c} \mathbf{P}^{NL}(\mathbf{x}, \omega), \quad (1)$$

or

$$\frac{\partial}{\partial z} \tilde{\mathbf{E}}(\mathbf{x}, \omega) = i\sqrt{k^2 - k^2_\perp} \tilde{\mathbf{E}}(\mathbf{x}, \omega) + \frac{i\omega^2}{2\epsilon_0 c^2 \sqrt{k^2 - k^2_\perp}} \mathbf{P}^{NL}(\mathbf{x}, \omega). \quad (2)$$

These wave equations are first-order in z. Applying them, a given electric field in a plane perpendicular to the z-axis may be propagated to the solution at a later plane along the same axis. The forward Maxwell equation [15], (1), applies the paraxial approximation. There are multiple forms of the nonlinear wave equation which can be derived with assumptions more specific to a given system under study [15]; (2) is used in this work as it allows a real-valued electric field to be the main propagated quantity, which is suitable for very short pulses where carrier-envelope phase effects are relevant.
The propagation is further simplified by representing the field in a basis evolving via Fourier optics, which changes as the field propagates through the medium:

$$\tilde{E}(\mathbf{x} + z_0, \omega) = e^{ik_0z} \tilde{E}(\mathbf{x}, \omega).$$ (3)

This operation, diagonal in the frequency domain, does not need to be solved via finite differences, and the first term on the R.H.S. of (2) is eliminated. This improves the convergence of the numerics by approximately a factor of four.

### 2.2. Dispersion of the refractive index

The dispersion of the refractive index is handled in the frequency domain in the UPPE, and in the time domain in FDTD. There are three dispersion equations available to the UPPE mode. The first of these is a general fitting equation designed to accommodate most expressions found in the literature. It has the up-to-22-parameter form:


where $k = \frac{e^2}{\epsilon_0 m}$ and the coefficients are stored in the array $a$. The equation contains four resonances producing a real-valued refractive index, with two amplitudes (multiplied by the square of wavelength and not) to accommodate most available expressions. There are three polynomial terms for using expressions where they are used as infrared corrections, and finally two Lorentzian lines which provide a complex-valued contribution for adding absorption lines.

The second form is a series of up to seven Lorentzian oscillators:

$$n^2 = a[0] + \sum_{m=0}^{6} \frac{ka[1 + 3m]}{a[2 + 3m] - \omega^2 + ia[3 + 3m] \omega} \quad (5)$$

The refractive index defined in this way may be used in both the UPPE and FDTD mode, because unlike the previous expression, it obeys the Kramers-Kronig relations, and is compatible with a time-domain calculation.

In the time domain, each oscillator has the following equations of motion for the current $J$ and polarization $P$, which are solved concurrently with the electromagnetic field in FDTD mode, for each oscillator of index $m$:

$$\frac{1}{\epsilon_0} \frac{\partial J}{\partial t} = ka[1 + 3m] E - a[3 + 3m] J - a[2 + 3m] P \quad (6)$$

$$\frac{\partial P}{\partial t} = J \quad (7)$$

$a[1 + 3m]$ provides the dipole strength, $a[2 + 3m]$ is the square of the angular frequency, and $a[3 + 3m]$ is the relaxation rate. $a[0]$ is high-frequency value of the dielectric constant $\epsilon_\infty$.

The third form of equation expresses the complex index as a series of resonances with Gaussian line-shapes, with the real part provided by the Dawson function. This also obeys the Kramers-Kronig relations and could in principle be used in FDTD mode, but has not yet been included there as it is significantly more numerically demanding than Lorentzians. In the
frequency domain, for oscillator of index $m$, it is expressed as

$$n^2 = a[0] + \sum_{m=0}^{6} \frac{1}{\sqrt{3}} a[3m] \left( F(\omega, \omega(m)) - i e^{\omega^2(m)} \right)$$

$$\omega_j(m) = \frac{\omega - a[1 + 3m]}{\sqrt{3}a[2 + 3m]}$$

where $F(\omega)$ is the Dawson function. This equation is more appropriate for solids whose absorption features are dominated by inhomogeneous broadening.

### 2.3. Dispersion of nonlinear coefficients

The dispersion of the nonlinear coefficients is taken into account through Miller’s rule [16], or an extrapolation of it to higher orders. This is the idea that the nonlinear susceptibilities will scale with the linear ones. In the case of $\chi^{(2)}$:

$$\chi^{(2)}_{jkl}(\omega_1 + \delta_1, \omega_2 + \delta_2; \omega_3 + \delta_3) \approx \frac{\chi^{(1)}_j(\omega_1 + \delta_1)\chi^{(1)}_k(\omega_2 + \delta_2)\chi^{(1)}_l(\omega_3 + \delta_3)}{\chi^{(1)}_j(\omega_1)\chi^{(1)}_k(\omega_2)\chi^{(1)}_l(\omega_3)} \chi^{(2)}_{jkl}(\omega_1, \omega_2; \omega_3)$$

This approximates the dispersive behavior of the nonlinear coefficients to a good degree of accuracy, especially if the resonances are Lorentzian or far away from all frequencies. To extend this to higher order nonlinearities, the pattern is identical, but the numbered indices extend up to higher numbers. The resulting nonlinearities are not instantaneous: the response to a $\delta$-function excitation will not be another $\delta$-function.

The dispersion of the nonlinear effects is treated the same way in the FDTD calculations, with the time-domain equivalent being the use of the properly-normalized nonlinear polarization as an additional driving term of the oscillators that produce the linear response of the material. In the case of third-order nonlinearities, a possible refinement of the accuracy of the dispersion would be to use a different set of oscillators for nonlinear effects, for example those responsible for Raman-active modes.

### 2.4. Nonlinear absorption and plasma formation

In many crystals, the damage mechanism will be multiphoton absorption followed by free-carrier absorption. In some cases, these effects will limit the intensity inside of the medium, through intensity clamping, or plasma formation may participate. This is treated classically using the Drude model in the time domain:

$$J^\text{Drude}_m(t) = \frac{e^2}{\mu_e} \exp(-\gamma t) \int_{-\infty}^{t} dt' \exp(\gamma t') N(t') E_m(t')$$

where $e$ is the electron charge, $\mu_e$ is the reduced effective mass, $\gamma$ is the momentum relaxation rate, and $N(t)$ is the density of free carriers. $E_m$ indicates the field in the $m$-direction, where $E$ with no index indicates the total field magnitude, $E = \sqrt{E_x^2 + E_y^2 + E_z^2}$.

This requires the time-dependent carrier density $N$ and must conserve energy. Exciting the system and introducing a pair of carriers costs energy, in the form of $n$ photons, where $n>1$. The field must then also undergo an absorptive nonlinearity, whose associated current in direction $m$ is, in the multiphoton limit:

$$J^\text{abs}_m(t) = \left( \beta E^2(t) \right)^{n-1} E_m(t)$$

where $\beta$ is a nonlinear absorption parameter, depending on the system and frequency of the excitation. From this rate of nonlinear absorption, the power transfer from the field to the
system will be \( \sum J_{ab}^{m}(t)E_{m}(t) \). This simple expression has been found to reproduce the excitation dynamics in wide-gap materials [8,17]. The nonlinear absorption rate depends on the wavelength of the laser, can be strongly influenced by the defect density in the crystal, and can be difficult to find in literature, although sufficient information can be obtained through a Z-scan measurement [18,19]. It will often be necessary to fit this value to measurements from the system under study if plasma effects are involved.

The number of carriers generated, assuming that they just go to the \( \Gamma \) point of the crystal is

\[
\frac{dN(t)}{dt} = \frac{2}{\Delta_{\varepsilon}} \sum_{m} J_{ab}^{m}(t)E_{m}(t)
\]  

(13)

where \( \Delta_{\varepsilon} \) is the band gap. The factor of 2 is due to the creation of an electron-hole pair.

3. Numerical implementation

In both the FDTD and UPPE modes, the numerical propagation is performed through the explicit, fourth-order Runge-Kutta method (RK4). In each case, the propagated quantity is the real-valued electromagnetic field, and no envelope approximations are applied; the full carrier wave is treated.

In the UPPE, the field grid is discretized over one or two spatial dimensions, and time. The electric field has two polarization states, \( E_{x} \) and \( E_{y} \). When it is represented in 3D, with two spatial dimensions and one time dimension, it provides a complete description of the dynamics within the approximations of the UPPE. Two reduced-dimension modes are provided. One uses a single Cartesian coordinate, \( x \), neglecting diffraction effects in the \( y \) direction, and is suitable for qualitative calculations, such as finding phasematching and estimating the effects of beam walk-off in birefringent crystals. The other 2D mode assumes cylindrical symmetry, which is suitable for the calculation of collinear beams in non-birefringent media, with significantly reduced resource requirements relative to the full 3D mode, and comparable results.

In FDTD the field is a discretized function of two or three spatial dimensions, with electric and magnetic fields in the Yee lattice arrangement [20], propagated in time. The fields have \( x \), \( y \), and \( z \) components. Additionally, a grid describing the state of the materials at the positions of the electric field grid is propagated in time along with the fields.

RK4 is not unconditionally stable when solving the systems under study. Accordingly, if the grid discretization is poorly chosen or the step size too large, the simulation will generate invalid floating point numbers in the result. However, the grid parameters at which this happens are already in the range where the accuracy is poor: such simulations would have been invalid even if the propagation was unconditionally stable, just more subtly so. For a simulation being used by the public, this is arguably an advantage, since if an invalid number appears the interface will immediately present an error that the simulation failed rather than providing a poorly-converged result. However, it is still necessary to check for adequate convergence with respect to the observable of interest for any results that will be published.

3.1. Considerations in cylindrical symmetry

Cylindrical symmetry has additional considerations when discretizing the spatial grid. The Fourier basis is not as useful in this mode as it is in the ones based on Cartesian coordinates. In order to fully apply Fourier optics, the proper basis would be entered via the Hankel transform instead. However, the Hankel transform was found to not perform well numerically as the fast Fourier transform (FFT) with an additional term in the propagation related to the transverse Laplacian. In cylindrical coordinates with radial symmetry, the transverse Laplacian becomes

\[
\nabla_{\perp}^{2} = \frac{\partial^{2}}{\partial \rho^{2}} + \frac{1}{\rho} \frac{\partial}{\partial \rho} .
\]  

(14)
The first term has the same form as one finds in Cartesian coordinates, and it may be eliminated in the same way, but the second cannot, because it is not a diagonal operator in the Fourier domain. Calculating this term explicitly in space requires an IFFT. This will be performed in any case for the calculation of the nonlinear polarization, so the additional computation is minimal.

The grid used in this case extends to the negative side of the origin, so that there is no discontinuity in the domain of the FFT at $\rho = 0$. This would seem to be a waste of memory since the positive and negative sides of the origin are identical by symmetry. If we use a grid that is truly symmetric, that is indeed the case: we would be solving everything at $\rho = \cdots - 3, -2, -1, 0, +1, +2, +3 \cdots$ and so on.

Shifting the grid points by $\frac{1}{4}$ we now have the field at $\rho = \cdots - 2\frac{3}{4}, -1\frac{3}{4}, -\frac{3}{4}, +\frac{1}{4}, +1\frac{1}{4}, +2\frac{1}{4}, +3\frac{1}{4} \cdots$. This way, the positive and negative sides of the grid contain different information: the previously-redundant negative-$\rho$ values contain the midpoints between the points on the positive $\rho$ grid. The spacing is still 1, so the grid doesn’t include high values of the transverse angular momentum, which often bring numerical instability, but we now have doubled resolution in space.

We make use of this in two ways: to reduce aliasing in the calculation of the nonlinear polarization, and to improve the accuracy of the finite-difference calculation of the radial derivative included in the $\frac{1}{\rho} \frac{\partial}{\partial \rho}$ term of the Laplacian. This results in overall easier convergence of calculations in cylindrical symmetry mode.

### 3.2. Program implementation

The core simulation is written in C++ and uses functors as a zero-cost abstraction allowing the same code to be compiled by a standard C++ compiler, the NVIDIA CUDA compiler, and in the SYCL language using Intel’s oneAPI DPC++ compiler. The language-specific implementations are encapsulated in device classes defined for each language. This allows the same source code and identical algorithm containing the parallel kernels and the surrounding logic to run on different classes of device, graphical processing units (GPUs) and central processing units (CPUs). The results from different devices only exhibit small differences resulting from their respective floating point implementations. Thus, performing a batch calculation (parameter scan), the work can be divided between CPU and GPU, running simultaneously.

The full source code is available in a public git repository [21]. This allows straightforward reproduction of simulations whenever the code is applied. Additionally, by making the project open to inspection and contribution from the wider community, values of parameters can be cross-checked in more detail than would be available in a private code-base, and external contributors may add new materials to the database of crystals available to all users.

### 4. User interface and interactions

The program is primarily run via a graphical user interface (GUI), shown in Fig. 1, implemented in GTK 4 for cross-platform compatibility. The interface contains text fields for the relevant parameters of the input electric fields (including options to load the results of a Field Resolved Optical Gating reconstruction or waveform measurement), the parameters of the crystal (primarily the selection of the active medium from a database contained in a user-editable text file), and the parameters of the space-time grid.

The simulation can also be compiled to a command-line version, primarily used to run it on multi-GPU clusters. The GUI has the option to create a SLURM script to run the simulation currently entered on the interface on systems using that job management system. A typical use-case would be to run a 2D calculation locally to verify the simulation parameters, then submit the 3D version to a cluster.
Fig. 1. An image of the GUI: the parameters of two input pulses are entered in the two left-most columns. The properties of the material and space-time grid, as well as control over the propagation mode are done using the two right-most columns and associated pull-down menus. The large box at the bottom of the window marked sequence can be used to enter a more complicated series of optical elements. Plots of the result in the time and frequency domains are provided on the right.

4.1. File formats

The final simulation grid (electric field, in $x-$ and $y-$ polarizations, as a function of space and time) is saved as a binary file, along with the integrated total spectrum, also in a binary file. A text file containing all simulation parameters is also saved. A Python module is provided that allows these results to be loaded into a data class containing the grids, spectra, and parameters for further post-processing and visualization.

4.2. Sequence programming

Sequences of optical elements can be written in compact scripts within the interface, allowing the result of one calculation to be passed to the next. These can comprise not only multiple nonlinear crystals, but also free-space propagation, spherical and parabolic mirrors, spectral filters, apertures, and so on. They are written in a simple, descriptive language. An example, for the propagation of light in a periodically-poled lithium niobate (PPLN) crystal, is:

```
set(1,14.6)
for(500/v01,0)
   nonlinear(d,d,d,v01,d)
   rotate(180)
```

This sequence sets the variable v01 to 14.6 (the half-period of the PPLN grating, in microns), and then runs a for-loop over a 500 µm interaction length, running nonlinear propagation through segments of the crystal of length v01 (the rest of the arguments of nonlinear() use the keyword d, which tells the interpreter to use the values entered in the user interface). After each segment, the system of coordinates is rotated by 180 degrees around the polarization axis, as required for quasi-phasesmatching. The sequence mode has been used to simulate both small arrangements of crystals, as well as entire laser cavities describing nonlinear effects inside of regenerative amplifiers.
4.3. Fitting and numerical optimization

The parameters of the simulation may also be numerically optimized with a global fitting routine provided by the DLIB library [22]. This allows, for example, the phasematching angles and the focusing conditions of the beam to be optimized for power simultaneously. A reference spectrum may also be loaded and the same routine used to reduce the difference between the measured and calculated spectra, which is a useful approach to determine crystal parameters which may be difficult to access through other measurements.

5. Example calculations

The example calculations below allow the output of the code to be compared against experiments in the literature, and allow different propagation modes of the program to be compared. Additional simulations paired with experiments have been published: looking at the spectral broadening and nonlinear pulse propagation dynamics of 2.3 µm pulses in TiO$_2$ [23], electro-optic sampling response functions [24], and broadband mid-infrared difference frequency generation in a multi-crystal arrangement [25].

5.1. Cross-polarized wave generation

The generation of cross-polarized waves (XPW) [26] in materials with asymmetric elements of the $\chi^{(3)}$ tensor has been intensely studied and utilized in ultrafast optics. Since the present code is able to automatically use the full $\chi^{(3)}$ tensor in calculations of nonlinear propagation, it is a useful test-case for comparing the resulting angle-dependent XPW generation process against literature. For example, the simulations in Fig. 2 can be compared against literature regarding the intensity dependence [27] and angle-dependence [28].

![Fig. 2. Simulation of XPW in BaF$_2$.](image)

(a) Variation of the efficiency with polarization angle of the incident light for two orientations of the crystal. (b) Scaling and saturation of the XPW signal, with and without the inclusion of multiphoton absorption. (c) Comparison of the temporal form of the electric field along the original polarization axis, and the cross-polarized axis, with moderate conversion (3%). (d) Comparison of the spectra on the axes.
This shows that the program can be applied to real-world calculations with little additional work on the part of the user, once the crystal properties are entered in the database. The input on the user interface of the program is simply the crystal orientation and the polarization of the light, while the required rotations and resolution of the nonlinear tensor is handled programmatically.

5.2. Comparison of UPPE and FDTD modes

The different propagation modes provide directly comparable results. In certain cases, there may be significant differences, understanding of which can be important for ensuring that the results are physical. For example, if the assumption of cylindrical symmetry is made in a strongly birefringent crystal, the effects of spatial walk-off will be lost. Similarly, calculations using the UPPE will neglect the effects of reflections at material interfaces. In the case of an uncoated crystal there can be significant changes in the results due to reflection, which become more prominent at higher intensities as the nonlinear modification of the reflectivity more strongly affects the light, and the constructive interference between the outgoing and reflected wave at the back surface of the crystal leads to a locally-enhanced intensity that can strongly increase multiphoton absorption at that surface.

A simple simulation, shown in Fig. 3 shows the differences in the transmitted few-cycle pulses obtained from a 2 µm fused silica plate in FDTD and UPPE modes. In the time domain, the most prominent effect is the post pulse resulting from multiple reflections, while in the frequency domain, one sees also a slight reshaping on the nonlinear signal.

![Fig. 3. Comparison of the results of transmission through a thin fused silica plate, when using the UPPE and FDTD models of field propagation. The UPPE model employed was 3D with the assumption of radial symmetry, while the FDTD model was 3D with no assumptions regarding symmetry. (a) The transmitted pulses from each model. The difference between the results is due to reflections on the surfaces of the sample, which are only present in FDTD. (b) Comparison of the spectra from the two propagators, showing that they are qualitatively similar, but with a small etalon effect present on the FDTD result, and small variations of intensity due to standing wave effects.](image)

The time required to perform a calculation using the UPPE and FDTD as implemented scales differently with the thickness of the crystal: UPPE propagation scales linearly with length, whereas FDTD will scale with the square of length. There, with a longer crystal, the grid must accommodate its entirety in space if all etalon effects are to be preserved, slowing down each time step linearly, and the field simply takes longer to propagate from one side to the other, requiring additional time steps and providing the second linear factor. Thus, FDTD mode is better suited to studies of thin materials (≲ 100 µm on a typical desktop PC). The UPPE scales more favorably with thicker crystals, where the surface reflections also play a relatively small role.
6. Conclusions and outlook

The open source software presented makes use of the well-understood nonlinear wave equations and nonlinear implementations of Maxwell’s equations, but makes them both easily accessible to students and researchers in ultrafast optics. It provides efficient use of the hardware available, across multiple platforms and device manufacturers. The open nature of the project will allow it to naturally grow and improve as others build on and adapt it to their own unique experiments.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. The data generated for this manuscript are available in the Edmond database of the Max Planck Society [29]. The source code of the simulation described is also publicly available [21].

References
