

A Generalized Ray Formulation For Wave-Optics Rendering

SHLOMI STEINBERG*, RAVI RAMAMOORTHI†, and BENEDIKT BITTERLI, NVIDIA, United States
EUGENE D’EON, NVIDIA, New Zealand
LING-QI YAN, University of California, Santa Barbara, United States
MATT PHARR, NVIDIA, United States



Fig. 1. **From ray optics to wave optics.** A scene rendered with our technique, showcasing various wave effects: (a) a Bornite ore with an interfering layer of copper oxide; (b) a Brazilian Rainbow Boa, whose scales are biological diffraction-grated surfaces; and (c) a Chrysolmelidae beetle, whose colour arises due to multilayered interference reflectors in its elytron. The practical contribution of this paper is the ability to render such complex scenes, under rigorous wave-optical light transport, at a performance that surpasses the state-of-the-art by orders-of-magnitude. The objective of this paper is not the appearance reproduction of some material (a “diffractive BRDF”), but an accurate formulation of wave-optical light transport, where light is rigorously treated as waves globally throughout the entire scene. We propose the *generalized ray*: an extension of the classical ray to wave optics. The generalized ray retains the defining characteristics of the ray-optical ray: *locality* and *linearity*. These properties allow the generalized ray to serve as a “point query” of light’s behaviour, the same purpose that the classical ray fulfils in rendering. Generalized rays enable the application of backward (sensor-to-source) light transport and sophisticated sampling techniques, which are impossible with the state-of-the-art. We indicate resolution and samples-per-pixel (spp) count in all figures rendered using our method. While these figures showcase converged (high spp) results, our implementation also allows interactive rendering of all these scenes at 1 spp. Frame times (at 1 spp) for interactive rendering are indicated. See our supplemental material for the implementation, and additional renderings and videos.

Under ray-optical light transport, the classical ray serves as a linear and local “point query” of light’s behaviour. Linearity and locality are crucial to the formulation of sophisticated path tracing and sampling techniques, that enable efficient solutions to light transport problems in complex, real-world settings and environments. However, such formulations are firmly confined to the realm of ray optics, while many applications of interest—in

computer graphics and computational optics—demand a more precise understanding of light: as waves. We rigorously formulate the *generalized ray*, which enables linear and weakly-local queries of arbitrary wave-optical distributions of light. Generalized rays arise from photodetection states, and therefore allow performing backward (sensor-to-source) wave-optical light transport. Our formulations are accurate and highly general: they facilitate the application of modern path tracing techniques for wave-optical rendering, with light of any state of coherence and any spectral properties. We improve upon the state-of-the-art in terms of the generality and accuracy of the formalism, ease of application, as well as performance. As a consequence, we are able to render large, complex scenes, as in Fig. 1, and even do interactive wave-optical light transport, none of which is possible with any existing method. We numerically validate our formalism, and make connection to partially-coherent light transport.

* Also with University of Waterloo.

† Also with University of California San Diego.

Authors’ addresses: Shlomi Steinberg, p@shlomisteinberg.com; Ravi Ramamoorthi, ravir@cs.ucsd.edu; Benedikt Bitterli, benedikt.bitterli@gmail.com, NVIDIA, San Francisco, United States; Eugene d’Eon, ejdeon@gmail.com, NVIDIA, Wellington, New Zealand; Ling-Qi Yan, lingqi@cs.ucsb.edu, University of California, Santa Barbara, Santa Barbara, California, 93106, United States; Matt Pharr, matt@pharr.org, NVIDIA, San Francisco, United States.

1 INTRODUCTION

Rendering and light transport theories are often formulated within the context of *ray optics*. We may state three tenets of ray-optical light transport:

- (1) **Locality**: The light-transport primitive, i.e. the *ray*, exists at a singular position (a point) in space and propagates into a singular direction at any given time. Therefore, rays serve as perfect “point queries” of light’s behaviour.
- (2) **Linearity** in terms of radiometric quantities (like intensity or radiance): The observed intensity I of a superposition of a pair of rays, with intensities I_1, I_2 , is

$$I = I_1 + I_2 . \quad (1)$$

- (3) **Completeness**: Any observable optical phenomenon (within the ray-optical context) can be computed to arbitrarily high accuracy with a finite count of rays.

The first pair of tenets are essential for rendering and path tracing techniques. *Locality* enables the use of spatial-subdivision acceleration structures, facilitating efficient ray propagation, even in the presence of massive geometry and detail. *Linearity* gives rise to a linear rendering equation, enabling the application of powerful sampling techniques. Note that linearity must be formulated in terms of radiometric quantities, as opposed to, for example, field strength. The final tenet, *completeness*, makes these light transport theories general: any effect that can be described by the applicable physics (ray optics) may also be simulated.

The locality and linearity of ray-optical light transport allows these techniques to scale to the massive scenes that can be seen in computer-generated content in films and other media. Today, path tracing techniques are even applied in real-time applications using dedicated ray-tracing hardware [Burgess 2020]. The above comes at a price: ray optics is a simplistic understanding of light, where wave effects are ignored. Such effects include the colourful glints that appear when light is scattered by scratches in metal; the colour of the wings and scales of some species of insects, snakes and fish; stress birefringence; and, the appearance of an oil layer or metal oxide on a surface. Many important unsolved problems demand wave simulation in complex scenes, e.g., RADAR simulation in urban environments for automotive applications, and signal coverage simulation for efficient infrastructure deployment.

The appearance of diffractive materials (Fig. 1), the ability of long-wavelength RADAR radiation to diffract around an object, and other wave effects all depend on light’s wave properties, its spatial and temporal optical coherence and spectral properties. In other words, the applications and effects mentioned above cannot be simulated only at a material-level, e.g., via a “diffractive BRDF” as done in computer graphics. Instead, these phenomena are the aggregated product of the *wave-optical light transport* over the entire scene. In Fig. 12 we show how light’s wave properties induce very different appearance in the same materials. Our focus in this paper is on formulating a wave-optical light transport formalism, that allows efficient and accurate modelling of these effects.

One could be naïvely motivated to attempt to formulate a wave-optical formalism that would admit all the three important tenets above. That is, replace the wave function with an alternative analytic construct, that is perfectly local, linear in terms of its observed intensity, and wave-optically complete. Unfortunately, that is impossible: a single-point formulation of light—one that is simultaneously both *local and linear*—is incompatible with electromagnetism [Wolf 2007].

While locality and linearity cannot be simultaneously achieved, many computational electromagnetic frameworks have been designed to be perfectly local: i.e., they employ a ray-like construct to conduct their wave simulations. This is indeed possible, and such formulations may even be complete, however any such formalism inevitably abandons linearity. This means that the observed intensity of a superposition of these ray-like constructs does not follow Eq. (1), but instead is:

$$I = I_1 + I_2 + \mathfrak{I}_{12} , \quad (2)$$

where \mathfrak{I}_{12} is a *bilinear* interference term that depends on both rays. The loss of linearity greatly frustrates the application of sampling and path tracing techniques: clearly, a linear rendering equation cannot be formulated; and the angular distribution of scattered energy—the BRDF—at a point of interaction behaves non-linearly, thus practical importance sampling is a significant challenge.

Linearity is crucial for a formulation that is compatible with path tracing techniques, hence we conclude that perfect locality must be abandoned. Recognizing that fact, *physical light transport* (PLT) was introduced by Steinberg et al. [2022]; Steinberg and Yan [2021] as a framework that enables practical, *weakly-local* wave-optical path tracing. PLT forgoes locality *to a degree*, that is, in-place of rays, PLT’s light transport primitives are beams: While a ray, at any time instant, exists at a singular point, a beam occupies a small spatial region and propagates into a solid angle. The important insight here is that, under some circumstances, locality can be sacrificed to a partial degree, in order to regain perfect linearity.

PLT deals with partially-coherent light, and sets the spatial region occupied by a beam to be proportional to the spatial coherence of light. Thereby, distinct beams become mutually incoherent, the interference term in Eq. (2) vanishes, and beams always superpose linearly. Hence, PLT is able to formulate a *linear* wave-optical rendering equation [Steinberg and Yan 2021], enabling the application of some classical path tracing tools and importance sampling of light-matter interactions. PLT is able to scale to large scenes, beyond any non-linear framework, while remaining accurate within its applicability domain (weakly-coherent light).

The sampling problem. In order to recover weak locality, PLT builds upon the partial coherence of light to formulate mutually-incoherent beams. The implication is that PLT is only compatible with a *forward* model of light transport, i.e. where we trace light from light sources to a sensor. To see why, recall that the coherence properties of light depend upon the light source, mutate on propagation, and on interaction with matter. The spatial extent of a beam, and hence *the degree-of-locality of the formalism*, also depends on these properties. If we were to trace beams *backwards* (from a sensor to sources) we would not be able to localize the beam, as the

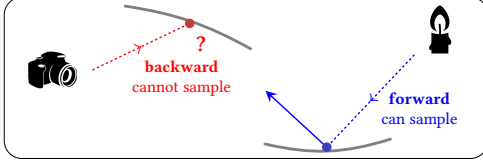


Fig. 2. **The sampling problem.** Existing light transport formalisms, like partially-coherent light transport, work by evolving light’s properties from the source, through the scene, until light is sensed by a detector. Such formalisms are inherently incompatible with backward (sensor-to-source) models of light transport: they are not able to formulate a light-matter interaction (a BSDF) without knowledge of light’s wave properties, however these properties depend on the light source and evolve throughout the scene, hence are very difficult to predict or estimate in a backward model [Steinberg et al. 2022]. Fundamental path tracing and sampling techniques, like importance sampling of interactions, cannot be applied, greatly hampering the practicality and ability of these formalisms to work with complex, real-world scenes. Solving this *sampling problem*, i.e. devising a formalism of backward wave-optical light transport, where a wide-range of sampling techniques can be applied, is the primary motivation for this paper.

coherence properties of light are unknown, therefore its spatial extent may be arbitrary, and locality is lost in its entirety (see Fig. 2).

We call this incompatibility of weakly-local formalisms with backward transport the *sampling problem*. The sampling problem is not restricted to PLT only, but applies to other formalisms as well, as we will discuss in Section 2. Moreover, in a forward model of light transport, weak locality can only be recovered by assuming or enforcing constraints on light’s properties, therefore: (i) any forward weakly-local formalism suffers from the sampling problem; (ii) and, such formalisms are never complete.

Our contributions in this paper. The backward model plays an important practical role in light transport simulations (for example, modern path tracers trace backward only or bi-directionally), and our primary motivation is to solve the sampling problem. In sharp contrast to ray optics, where a ray of light admits identical dynamics to its time-reversed counterpart, within the wave-optical context there is a fundamental change in physics between forward and backward light transport. We will discuss this further in Section 3.

The sampling problem arises due to this fundamental incompatibility between the forward and backward models. In order to properly address this sampling problem, we derive a novel formalism of backward wave-optical light transport. Our formalism is simultaneously *weakly-local*, *linear* and *complete*. There does not exist another wave-optical formalism with all these properties, and, as discussed, this would not be possible in a forward model. This is a powerful result: completeness means that our formalism can sample arbitrary wave-optical distributions of light, of any wavelength, any spectrum, any polarization, and any optical coherence.

At the heart of our proposed formalism is the *generalized ray*: a general-purpose wave-optical construct, that generalizes the classical “point query”, i.e. the ray, to wave optics. Conveniently, the weak locality of a generalized ray does not depend on light’s properties, and only depends on the sensed wavelength. In Sections 4 and 5, we derive our theory of backward wave-optical light

transport, making essentially no approximating assumptions, from well-known optical principles. This is our primary contribution. The sections marked with an asterisk (*) are mathematically dense, and readers that are less interested in the theory might wish to (initially) skip these sections.

While our proposed formalism may also serve as an alternative to the state-of-the-art (for backward-only transport), it is meant to *complement* existing weakly-local, linear forward-based formulations, by working in tandem with them to enable bi-directional wave-optical transport. In Subsection 7.1 we discuss a particular application of the formalism to high-performance wave-optical rendering, and present *sample-solve*: a simple two-stage bi-directional algorithm, that first samples the scene via backward transport with generalized rays; then, uses PLT to apply a partially-coherent forward pass that acts as a variance reduction technique. We provide a sample implementation of our rendering algorithm in our supplemental material. We will show in Subsection 7.2 that we are able to do wave-optical rendering with complex scenes (as in Fig. 1), and at interactive performance, with convergence rate that is multiple orders-of-magnitude faster compared to the state-of-the-art.

2 RELATED WORK

Wave-optical light transport. In computer graphics one of the earliest formulations of wave-optical light transport is *diffractive shaders* [Stam 1999]. Light is propagated as rays, and a ray is converted into a plane wave (implicitly an asymptotic, far-field approximation) when formulating a diffractive shader, which quantifies the interaction of that ray with a surface. A plane wave exists throughout the entire space, therefore no degree of locality could be rigorously recovered.

Wigner distribution-based light transport methods have gained some attention both in computer graphics [Cuyper et al. 2012] and optical literature [Jensen and Buot 1991; Mackay and Johnson 2021; Mout et al. 2018]. The *Wigner distribution function* (WDF) $\mathcal{W}(\vec{r}, \vec{k})$ (see formal definition in Eq. (6)) is a bilinear distribution defined w.r.t. some wave function $\psi(\vec{r})$, and is a function of spatial position \vec{r} and wavevector \vec{k} . The WDF admits some useful ray-like properties, which has motivated the definition of the WDF as a form of a “generalized radiance” [Walther 1968]. However, locality is lost when the interaction with matter is considered, as we formally show next. The scattered WDF is [Torre 2005, Chapter 8]

$$\mathcal{W}_o(\vec{r}_o, \vec{k}_o) = \int d\vec{r}_i d\vec{k}_i K(\vec{r}_o, \vec{r}_i, \vec{k}_o, \vec{k}_i) \mathcal{W}_i(\vec{r}_i, \vec{k}_i), \quad (3)$$

where $\mathcal{W}_i, \mathcal{W}_o$ are the input and output WDFs, respectively, and K is a *diffractive kernel* that is completely defined by the scattering matter [Testorf et al. 2010, Chapter 1.6]:

$$K(\vec{r}_o, \vec{r}_i, \vec{k}_o, \vec{k}_i) \triangleq \int d\vec{x}_o d\vec{x}_i h\left(\vec{r}_o + \frac{1}{2}\vec{x}_o, \vec{r}_i + \frac{1}{2}\vec{x}_i\right) \times h^*\left(\vec{r}_o - \frac{1}{2}\vec{x}_o, \vec{r}_i - \frac{1}{2}\vec{x}_i\right) e^{-i(\vec{k}_o \cdot \vec{x}_o - \vec{k}_i \cdot \vec{x}_i)}, \quad (4)$$

where \star denotes complex conjugation, and h is the *optical response function* of the system. Locality is lost because, by definition, K requires integration over the *entire scene*.

Cuypers et al. [2012] claim that the negative values taken by the WDF allow “rays to interfere later for global illumination”. This is only true when a single WDF accounts for all light in the system, and the kernel K accounts for the entire scene. The WDF is a bilinear distribution, therefore the WDF of the sum $\psi_1 + \psi_2$ takes a form identical to Eq. (2), viz.

$$\mathcal{W} = \mathcal{W}_1 + \mathcal{W}_2 + \mathcal{W}_{12}, \quad (5)$$

where $\mathcal{W}_1, \mathcal{W}_2$ are the WDFs of ψ_1, ψ_2 , respectively, and \mathcal{W}_{12} is a bilinear cross-term, that depends on both ψ_1 and ψ_2 (see Testorf et al. [2010, Chapter 1.8] for more details). If we were to limit the spatial integration in Eq. (4) to a finite region, then different kernels would account for different interactions, and to compute the total superposition WDF with contributions from rays scattered from distinct spatial regions, we must account for the bilinear term in Eq. (5). Cuypers et al. [2012] implicitly neglect that term: by assuming that the WDFs from different rays add up *linearly*. The negative values that arise in the WDF of a light distribution depend only on that distribution, while the bilinear cross-term that integrates over both distributions is the term that accounts for their mutual interference. Discarding the cross-term is equivalent to neglecting wave interference. We numerically demonstrate in Fig. 5 that limiting the integration region in Eq. (4) and not accounting for the bilinearity of the WDF produces incorrect results.

Both Stam [1999] and Cuypers et al. [2012] recognize the need for linearity, however neither admits machinery that is able to limit the spatial integration region of light-matter interactions to a well-defined finite extent, hence neither admits locality of any form. Even if such machinery was to be developed, these frameworks would then suffer (like any other weakly-local forward formalisms) from the very same sampling problem that we address in this paper.

Physical light transport (PLT) [Steinberg et al. 2022; Steinberg and Yan 2021] does derive such machinery by quantifying the coherence of light. Under the assumption of weakly-coherent light, rigorous, well-defined weak locality is recovered. That is, PLT is able to quantify the spatial extent over which a light-matter interaction (like Eq. (4)) may produce observable interference. Treating light as partially-coherent beams, PLT gives rise to a weakly-local, linear formalism of wave-optical light transport. PLT is able to accurately, and efficiently, reproduce partially-coherent effects. However, in order to do that, PLT needs to propagate the coherence properties of light forward—properties that are needed in-order to recover weak locality—hence the sampling problem inevitably arises, and practical backward light transport is not possible.

Computational electrodynamics. A variety of wave solvers, many commercially available, aim to produce a solution to Maxwell’s equations. Common solvers are variants of the finite-difference time-domain (FDTD) [Yee 1966], finite element (FEM) [Jin 2015], or boundary element (BEM) methods. These methods are usually supra-linear in scene complexity, require deterministically described fields as well as geometry at sub-wavelength resolution. Furthermore, solving for the electromagnetic field directly is neither feasible nor desirable in most applications of interest: *observable properties* of light arise when integrated over the extent of a sensor as well as over the

period of observation. These methods are practical only in exceedingly simple scenes.

Of more relevance are “physical optics” methods, most notably the “shooting-bouncing ray” (SBR) method. SBR is a perfectly-local method, by design, and uses ray tracing to track field propagation to a target aperture or geometry (i.e., a ray captures the behaviour of the field as it propagates into some solid angle). Then, the electromagnetic field is reconstructed at the target in order to evaluate the scattered electromagnetic radiation. Applications of SBR methods are extensive, and include: simulation of radar for imaging [Feng and Guo 2021]; driving-assistive technology [Castro et al. 2019]; analysis of aircraft scattering cross-section [Bilal et al. 2019]; ground-penetrating radar [Warren et al. 2016]; and, indoor positioning using WiFi [Hossain et al. 2018]. Some methods point-sample an input field and propagate a ray tube in order to approximate a diffraction integral [Andreas et al. 2015].

Another popular class of perfectly-local methods are ray tracers that handle diffractions by employing the geometric/uniform theory of diffraction (UTD) [Bilibashi et al. 2020; Son and Myung 1999; Yi et al. 2022]. Some of these have become the state-of-the-art in the simulation of light of longer wavelengths: major applications include automotive-targeted simulation of RADAR [Boban et al. 2014; Guan et al. 2020], as well as simulation of WiFi/cellular radiation (e.g., for analysis of signal coverage in a city) [Choi et al. 2023; de Adana et al. 2005]. Only a few select works are cited. Due to their importance, these applications have garnered significant attention, and better solutions would be of real interest.

SBR and UTD-based ray tracing methods are perfectly local, and some may be complete, but as discussed, such methods can never be linear. That is, rays are always *mutually interfering*, and the interference term in Eq. (2) must always be considered. As very many rays need to be traced, attempts to accelerate SBR methods have taken a research trajectory not dissimilar to early computer graphics work: Employing multi-resolution grids [Suk et al. 2001] and spatial-subdivision data structures to accelerate ray-facet intersections [Jin et al. 2006; Tao et al. 2008]; accelerating on GPUs [Gao et al. 2015; Tao et al. 2010]; and, sampling the initial ray directions using Halton sequences [Key et al. 2018]. Nevertheless, as linearity is unavoidably lost, this research greatly trails computer graphics in its ability to apply sophisticated sampling techniques, like importance sampling light-matter interactions or path guiding. Scene complexity remains a very limiting factor.

Holography. Some limited forms of wave-optical light transport have been used for computer-generated holography (CGH). These algorithms share similar problems to the methods described above. Magallón et al. [2021] propose to trace ray tubes backwards and accumulate the coherent contributions from all traced paths. Blinder et al. [2021] also consider the challenges of efficient backward wave-optical path tracing, and propose a method which discretizes and integrates a diffraction integral. Being perfectly-local but non-linear formalisms, both suffer from the issues of non-linearity. An accurate backward model of wave-optical light transport is important for CGH. Our formalism could be applied to CGH in the future, though it is beyond the scope of this paper.

Optical speckle. Tangentially related is the rendering of speckle or the integration of speckle statistics [Bar et al. 2019, 2020]. Speckle are patterns produced by coherent contributions from many scatterers. However, not all such contributions give rise to *observable* speckle. Observable optical phenomena are always *integrated* effects: the integration is over the spatial and angular extent of the observer (which, by the uncertainty principle, must be positive), and over the period of observation. Point-sampling interference effects (as perfectly-local formalisms most often do) suffer from aliasing. This is discussed in greater detail in Section 5. We show in Fig. 3 that as we integrate over the extent of an observer, scatterers that are sufficiently far apart do not contribute to observable interference effects, even in the presence of perfectly-coherent illumination.

Typical speckle integration algorithms (and forward light transport formalisms) have no means to establish whether a pair of scatterers do indeed contribute to observable speckle, and hence can be wasteful. In our formalism, generalized rays quantify exactly the spatial extent in a scene over which observable interference may indeed arise. Therefore, our formalism could be used in future work to speed up such speckle rendering, as well as support partially-coherent speckle. A sample-solve method could also be used to integrate speckle statistics. This is out of scope of this paper.

Material appearance reproduction. Also related is work that aims to reproduce the appearance of some diffractive materials. This includes the rendering of iridescent and pearlescent materials [Guilén et al. 2020]; diffractive scratches [Velinov et al. 2018; Werner et al. 2017]; statistical surface profiles [Holzschuch and Pacanowski 2017; Krywonos 2006; Steinberg and Yan 2022] or diffractive surfaces with explicit microgeometry [Falster et al. 2020; Yu et al. 2023]; and, thin-film interference at a soap bubbles [Huang et al. 2020] or due to a dielectric layer over a conductor [Belcour and Barla 2017; Kneiphof et al. 2019]. Synthesis of BSDFs that account for wave interference was discussed by Toisoul and Ghosh [2017].

The cited work does not deal with light transport in a scene, but only considers the result of an interaction with a material. Assumptions are made regarding the structure of the simulated fields: these are often restricted to plane waves, which can only be used for light transport under an asymptotic, far-field approximation; or, Gaussian beams, but a decomposition into mutually-incoherent Gaussian beams in a forward model is a crude approximation (and always suffers from the sampling problem). In this work we are not concerned with appearance reproduction, or the simulation of some particular wave-optical interaction at the material level. Instead, we focus on deriving a light transport formalism that does not suffer from the sampling problem.

3 BACKGROUND: THEORETICAL FOUNDATIONS

We briefly introduce the Wigner picture of wave optics, which is a useful setting for the discussion of the dynamics of light. See our supplemental material or Testorf et al. [2010]; Torre [2005] for additional information. We also formalize the process of photoelectric detection [Mandel and Wolf 1995; Ou and Kimble 1995]. Our interest in photoelectric detection is motivated by how our backward light transport formalism (introduced in Section 4) operates: In contrast to forward models of light transport, which propagate the

emission distributions of light forward in time, we propagate the *detection states* of a detector under time-reversed dynamics. This is key in decoupling the light transport simulation from the properties of light, thereby solving the sampling problem.

3.1 Phase-Space Optics

The basic descriptor of light under wave optics is the *wave function*, denoted $\psi(\vec{r}; t)$, where \vec{r} is position and t is time. The wave function quantifies the spatial excitations of a component of the associated electric field. Henceforth, we will fix time and, for brevity, drop t from the argument lists. In his seminal work in 1932, Wigner formulated the *Wigner distribution function* (WDF) [Wigner 1932], which describes a *position-momentum* distribution of the signal ψ :

$$\mathcal{W}(\vec{r}, \vec{k}) \triangleq \frac{1}{(2\pi)^3} \int d\vec{r}' \psi^*(\vec{r} - \frac{1}{2}\vec{r}') \psi(\vec{r} + \frac{1}{2}\vec{r}') e^{-i\vec{r}' \cdot \vec{k}}, \quad (6)$$

where \star denotes complex conjugation, \vec{r} is a spatial position, and \vec{k} is the *wavevector*, which quantifies light’s temporal frequency and direction of propagation. $|\vec{k}| = \eta \frac{2\pi}{\lambda}$ is the *wavenumber*, with η being the refractive index of the medium in which light propagates and λ is the light’s wavelength. Up to a phase term, the wave function may always be recovered from the WDF, hence the WDF provides a complete description of light.

The domain of the WDF is a space of both positions and wavevectors, referred to as *phase space*. This phase-space treatment of wave optics admits stark similarities to the ray-optical depiction of light: $\mathcal{W}(\vec{r}, \vec{k})$ is the energy density of a “light particle” at position \vec{r} with direction of propagation proportional to \vec{k} . $\mathcal{W}(\vec{r}, \vec{k})$ also propagates like the classical ray in free space and under some other interactions. Furthermore, it is well known that the WDF fulfils most of the postulates expected of a classical (ray optical) phase-space density function. The WDF does depart from the ray-optical picture in a couple of fundamental aspects:

- (1) **Bilinearity:** The WDF is a bilinear distribution, hence a linear superposition of WDFs is not the WDF of a superposition of waves (see Eq. (5)).
- (2) **Negativity:** The WDF takes somewhat counter-intuitive negative values, restricting its interpretation as an energy density. These negative values are not an exception, as a matter of fact, non-negative WDFs arise only with one specific class of wave functions [Torre 2005].

As discussed, the properties of the WDF have motivated other work to use the WDF to perform ray-like “point queries” of light behaviour for wave-optical light transport. However, either locality is entirely lost when interaction with matter considered (viz. Eq. (3)), because the kernel K must account for the entire scene; or, if we restrict K to a subset of the scene, then WDFs that arise from interactions with different such kernels must superpose bilinearly. No formalism can be simultaneously perfectly-local and linear. In addition, as a consequence of the WDF not being non-negative, it is a highly-oscillatory function, as shown in Section S4 in our supplemental material. This means that integration of the WDF with point samples heavily suffers from aliasing.

We are not interested in working with potentially-arbitrary emitted WDFs, as done by other Wigner-based formalisms. Instead, we

work in phase space to study the dynamics of *detection states*, introduced next, and show that these states admit far nicer properties than a general WDF.

Gaussian distributions. A phase-space construct central to our discussion is the (symplectic) *Gaussian WDF* [Testorf et al. 2010]:

Phase-space Gaussian

$$g_{\beta,\rho}(\vec{r}, \vec{k}; \vec{r}_0, \vec{k}_0) \triangleq \frac{1}{\pi^3} \exp\left[-\frac{1+\rho^2}{\beta^2} |\vec{r} - \vec{r}_0|^2 - \beta^2 |\vec{k} - \vec{k}_0|^2\right] \times \exp[2\rho |\vec{r} - \vec{r}_0| |\vec{k} - \vec{k}_0|]. \quad (7)$$

As will be discussed in Subsection 3.2, the (non-negative) values measured by our detectors can always be written as, and only as, a convolution of such phase-space Gaussians with an arbitrary WDF. Therefore, we will use these Gaussians as our weakly-local light transport primitives.

The parameterization of the Gaussian is chosen such that we may understand $\beta > 0$ as the initial spatial variance of the WDF g , and the correlation parameter $\rho \geq 0$ is related to propagation distance. As we will see later, propagation of light induces phase-space correlation between the spatial and wavevector variables, i.e. propagation increases ρ . In general, β, ρ can be positive-definite matrices, describing anisotropy (see our supplemental material). For simplicity, here we assume scalar parameters. g is normalized such that $\int d\vec{r} d\vec{k} g_{\beta,\rho} = 1$.

Phase-space Gaussians can also be understood as the most compact physically-realizable construct: note that the product of the spatial and wavevector variances of the Gaussian above is $\frac{1}{4}(1 + \beta^4 \rho^2)$, which fulfils the *uncertainty relation* [Mandel and Wolf 1995, Chapter 4], viz. $\sigma_r^2 \sigma_k^2 = 1/4$, if and only if $\rho = 0$.

3.2 Measurement of the WDF *

In sharp contrast to classical physics, we may not measure light without disturbing it: under wave optics, any measurement apparatus deduces information about light by interacting with it. The action of a detector on light can be quantified by the detector's WDF, and the intensity observed by the detector is [Dragoman 2005]

$$I = \int d\vec{r}' d\vec{k}' \mathcal{W}(\vec{r}', \vec{k}') \mathcal{W}_d(\vec{r}', \vec{k}'), \quad (8)$$

where $\mathcal{W}, \mathcal{W}_d$ are the WDFs of light and the detector, respectively. Even though the WDF takes negative values, the result of a measurement is always non-negative, i.e. $I \geq 0$, which formally follows directly from the Moyal formula [Torre 2005, Chapter 6.3.7].

We assume our detectors are *classical photoelectric* detectors. Photons impinging upon a photoelectric detector ionize free electrons, producing a current that is amplified and measured. All detectors of interest (e.g., a photoreceptor in the eye, a pixel in a CMOS array in a camera, or a RADAR antenna) operate in this fashion. Classicality implies that quantum effects are ignored: we assume that the photon flux is sufficiently high that electric pulses from single photons overlap and form a continuous electric current.

As photoelectric detection works by absorption of photons, it is the *photon annihilation* and *photon number* operators that are the observables [Mandel and Wolf 1995, Chapter 11]. The only eigenstates of these operators are the (squeezed) *coherent states* [Mandel

and Wolf 1995, Chapter 11, 21], whose phase-space representation takes the form of uncorrelated (minimum-uncertainty) Gaussians, i.e. the Gaussians defined in Eq. (7) with $\rho = 0$, viz.

$$\hat{g}_\beta(\vec{r}, \vec{k}; \vec{r}_0, \vec{k}_0) \triangleq \frac{1}{\pi^3} e^{-\frac{1}{\beta^2} |\vec{r} - \vec{r}_0|^2} e^{-\beta^2 |\vec{k} - \vec{k}_0|^2}. \quad (9)$$

As with the general Gaussian g , the parameter $\beta > 0$ above quantifies the trade-off between the spatial and wavevector variances.

Let a detector be described by a distribution \mathcal{D} of detection elements, each quantified by the parameters: (i) \vec{k}_0 : mean wavevector of detection (quantifies the mean wavelength and direction of propagation of detected light); (ii) β : spatial extent; and, (iii) α : detection efficiency (quantum efficiency). The WDF \mathcal{W}_d of the detector quantifies which light states are detectable [Dragoman 2005]. As the states detectable by a classical photoelectric detector are the coherent states, the detector's WDF may be written as

$$\mathcal{W}_d(\vec{r}, \vec{k}) = \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \hat{g}_\beta(\vec{r}, \vec{k}; \vec{r}_0, \vec{k}_0) \quad (10)$$

where $\mu_{\mathcal{D}}$ is the measure over \mathcal{D} . In the simplest case, \vec{k}_0, α, β are constant, and \mathcal{D} quantifies the spatial region occupied by the detector. Then, the integration is over the spatial extent, with \vec{r}_0 being the integration variable. In general, \mathcal{D} may quantify spatially-varying detection properties.

Denote the *Husimi Q distribution* of a WDF \mathcal{W} as the convolution of the WDF with a phase-space Gaussian:

$$\mathcal{Q}(\vec{r}, \vec{k}, \beta) \triangleq \int d\vec{r}' d\vec{k}' \mathcal{W}(\vec{r}', \vec{k}') \hat{g}_\beta(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0). \quad (11)$$

Substitute Eqs. (10) and (11) into Eq. (8) and formally interchange the orders of integration:

$$\begin{aligned} I &= \int d\vec{r}' d\vec{k}' \mathcal{W}(\vec{r}', \vec{k}') \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \hat{g}_\beta(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0) \\ &= \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \mathcal{Q}(\vec{r}_0, \vec{k}_0, \beta). \end{aligned} \quad (12)$$

The above implies that a classical photoelectric detector directly measures the Husimi Q distribution, and not the WDF. The “smoothing” operation of the WDF over a minimum-uncertainty Gaussian (as in Eq. (11)) is a sufficient and necessary condition to produce a bandwidth-limited, non-negative distribution [Soto and Claverie 1983], and it is the smoothed distribution that classical photoelectric detectors observe [Dragoman 2004; Leonhardt 1997].

4 THEORY OF BACKWARD WAVE-OPTICAL LIGHT TRANSPORT

Note the symmetry in the measurement formula, Eq. (8): we may understand it as the detector acting upon the incident distribution of light; or, equivalently, as light acting upon detection states. A forward model works by sourcing a light distribution \mathcal{W} from light sources, simulating its interaction with the scene, and finally integrating it over the detector's distribution. On the other hand, our backward model that we present in this Section evolves the detector distribution \mathcal{W}_d under time-reversed dynamics, and then integrates it over the sourcing distribution. In contrast to the sourced WDF, the WDF of the photoelectric detector \mathcal{W}_d (Eq. (10)) admits much nicer analytic properties.

We term the detection states' wave function as the *generalized ray*. We formally define the generalized ray and discuss their sourcing from an arbitrary photoelectric detector \mathcal{D} in Subsection 4.1. In Subsection 4.2, we will formalize the workings of our backward light transport formalism with generalized rays. In Section 5 we analyze this formalism and show that it is weakly-local, linear and complete. In Section 6 the interaction kernels will be discussed further, and we will show how to derive such kernels. In Section 7 we build upon the formalism developed here, and present a practical, interactive wave-optical renderer.

Time-reversed dynamics. Consider the WDF \mathcal{W}_s of light that is sourced from one or more light sources. Let the light quantified by \mathcal{W}_s propagate around a scene, interact with that scene, and finally impinge upon a classical photoelectric detector, with detection WDF \mathcal{W}_d , as in Eq. (10). Let the kernel for the entire process be denoted K (as defined in Eq. (4)), and we denote the action of such kernel on a WDF (as in Eq. (3)) via the *interaction operator*:

$$\mathcal{K}\{\mathcal{W}\}(\vec{r}_o, \vec{k}_o) \triangleq \int d\vec{r}_i d\vec{k}_i K(\vec{r}_o, \vec{r}_i, \vec{k}_o, \vec{k}_i) \mathcal{W}(\vec{r}_i, \vec{k}_i). \quad (13)$$

Denote the WDF of the light that impinges upon the detector, after interaction with the scene, as $\mathcal{W}_o = \mathcal{K}\{\mathcal{W}_s\}$. Apply the measurement formula, Eq. (12), to \mathcal{W}_o , yielding an expression for the result of a measurement of light after interaction:

$$\begin{aligned} I &= \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \mathcal{Q}_o(\vec{r}_o, \vec{k}_o, \beta) \\ &= \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \int d\vec{r}' d\vec{k}' \hat{g}_{\beta}(\vec{r}', \vec{k}'; \vec{r}_o, \vec{k}_o) \mathcal{K}\{\mathcal{W}_s\}(\vec{r}', \vec{k}'), \end{aligned} \quad (14)$$

where \mathcal{Q}_o is the Husimi Q distribution (Eq. (11)) of \mathcal{W}_o .

Time-reversal induces wavevector reversal, viz. $\vec{k} \rightarrow -\vec{k}$, and phase conjugation [Geru 2018, Chapter 2.3]. Hence, a system's optical response function h fulfils $h(\vec{r}_o, \vec{r}_i) = h^*(\vec{r}_i, \vec{r}_o)$. By applying this relation to Eq. (4), we define the *time-reversed diffraction kernel*:

$$K^{-1}(\vec{r}_o, \vec{r}_i, \vec{k}_o, \vec{k}_i) \triangleq K^*(\vec{r}_i, \vec{r}_o, \vec{k}_i, \vec{k}_o). \quad (15)$$

The time-reversed interaction operator \mathcal{K}^{-1} is then defined to act upon a WDF via K^{-1} , as in Eq. (13). Using Eqs. (4) and (13), we may observe that $\mathcal{K}^{-1} \equiv \mathcal{K}^{\dagger}$, for lossless kernels. Formally interchange the integration order in Eq. (14) between the integral over the primed variables and the integral over the kernel K , and apply the time-reversed interaction operator, yielding

Backward light transport formalism

$$I = \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha \int d\vec{r}' d\vec{k}' \mathcal{W}_s(\vec{r}', \vec{k}') \mathcal{K}^{-1}\{\hat{g}_{\beta}\}(\vec{r}', \vec{k}'). \quad (16)$$

Subject to the physics of photoelectric detection (Subsection 3.2), all the derivations above are exact. Eq. (14) quantifies the *forward* model of light transport: a distribution of light sourced from light sources, \mathcal{W}_s , is evolved forward in time, and then measured by a detector. On the other hand, Eq. (16) quantifies the *backward* model: each detector state of a photoelectric detector is evolved backward, under the time-reversed wave-optical dynamics (as quantified by

Eq. (15)), and then the overlap of this state with the sourcing distributions are computed. Both yield exactly the same answer. However, in contrast to the potentially arbitrary WDF \mathcal{W}_s , the artefacts of photoelectric detection—the states \hat{g} —are much “better behaved” functions: they are simple Gaussians, bandwidth limited and always non-negative.

4.1 Generalized Rays *

By applying the inverse WDF transform to g [Torre 2005, Chapter 6.3.5], we derive the wave function of the phase-space Gaussian, which we term the *generalized ray*:

Generalized ray (wave function)

$$\psi_{\beta, \rho}(\vec{r}; \vec{r}_o, \vec{k}_o) \triangleq \left(\frac{1}{\pi\beta^2}\right)^{3/4} e^{i\vec{k}_o \cdot (\vec{r} - \vec{r}_o)} e^{-\frac{1}{2\beta^2} (1-i\rho) |\vec{r} - \vec{r}_o|^2}. \quad (17)$$

A general expression with anisotropic β, ρ is available in our supplemental material. The generalized ray is defined with respect to the mean spatial position \vec{r}_o , mean wavevector \vec{k}_o , and parameters β, ρ . The reader may verify via Eq. (6) that the phase-space representation of the generalized ray (its WDF) is $g_{\beta, \rho}(\vec{r}, \vec{k}; \vec{r}_o, \vec{k}_o)$.

Initially, i.e. given $\rho = 0$, a generalized ray is one of the detection states of the detector \mathcal{D} , and its phase-space representation is a coherent state \hat{g}_{β} that minimizes the product of its space and wavevector variances. Thus, it is the *most local* construct permissible under wave optics [Torre 2005]. As its name suggests, the generalized ray can be understood as *the closest analogue to the classical ray under wave optics*. Once it is propagated away from the detector ($\rho > 0$), its WDF becomes the correlated Gaussian g , as in Eq. (7).

Sourcing. Let \mathcal{W}_s be a sourcing WDF, as before. Under the backward model of light transport (Eq. (16)), the result of a measurement at a detector element $\{\vec{r}_o, \vec{k}_o, \beta, \alpha\} \in \mathcal{D}$ can be written as

$$I(\vec{r}_o, \vec{k}_o, \beta) \triangleq \int d\vec{r}' d\vec{k}' \mathcal{W}_s(\vec{r}', \vec{k}') \mathcal{K}^{-1}\{\hat{g}_{\beta}\}(\vec{r}', \vec{k}'). \quad (18)$$

The total intensity integrated over the entire detector is then

$$I = \int_{\mathcal{D}} d\mu_{\mathcal{D}} \alpha I(\vec{r}_o, \vec{k}_o, \beta). \quad (19)$$

By Monte-Carlo integrating the integral over the detector distribution \mathcal{D} in Eq. (19), we *source* a generalized ray from the detector:

Sourcing

$$\begin{aligned} I &\approx \frac{1}{N} \sum_{n=1}^N \frac{\alpha^{(n)}}{p^{(n)}} I(\vec{r}_o^{(n)}, \vec{k}_o^{(n)}, \beta^{(n)}) \\ &= \frac{1}{N} \sum_{n=1}^N \frac{\alpha^{(n)}}{p^{(n)}} \int d\vec{r}' d\vec{k}' \mathcal{W}_s(\vec{r}', \vec{k}') \mathcal{K}^{-1}\{g^{(n)}\}(\vec{r}', \vec{k}') \end{aligned} \quad (20)$$

(normalization constants accounted for by α), with $p^{(n)}$ being the sampling probabilities. Each sampled $\{\vec{r}_o^{(n)}, \vec{k}_o^{(n)}, \beta^{(n)}, \alpha^{(n)}\} \in \mathcal{D}$ is a detector state that quantifies a sourced generalized ray, whose WDF is denoted as

$$g^{(n)}(\vec{r}, \vec{k}) \triangleq \hat{g}_{\beta^{(n)}}(\vec{r}, \vec{k}; \vec{r}_o^{(n)}, \vec{k}_o^{(n)}) \quad (21)$$

(recall that detection state \hat{g}_{β} is simply $g_{\beta, \rho}$ with $\rho = 0$). A sourced generalized ray is defined by its spatial position on the detector

$\vec{r}_0^{(n)}$, the wavevector $\vec{k}_0^{(n)}$ that quantifies wavelength and direction of propagation of the generalized ray from the detector, and the variance $\beta^{(n)}$ that quantifies its spatial extent. At sourcing, we set $\rho^{(n)} = 0$. At the detector, α is the detection efficiency, however once a generalized ray is sourced $\alpha^{(n)}$ should be understood as the intensity carried by that generalized ray.

4.2 Light Transport with Generalized Rays *

For each generalized ray sampled from a detector, viz. Eq. (20), we run a light transport simulation, as formalised by Eq. (16). That simulation is dictated by the time-reversed (total) interaction operator \mathcal{X}^\dagger . To compute the contributions of the detection element to the total intensity observed by the detector, we finally integrate the generalized ray over the light sources' sourcing distributions \mathcal{W}_s .

Often it is more practical to formulate the interaction operator \mathcal{X} , that quantifies the total action on light of the entire scene, as a sequence of interactions, viz. $\mathcal{X} \equiv \mathcal{X}_1 \circ \mathcal{X}_2 \circ \dots \circ \mathcal{X}_M$. A single interaction \mathcal{X}_j might now describe free-space propagation, or a light-matter interaction. The definition of the composed interaction operator $\mathcal{X}_1 \circ \mathcal{X}_m$ only depends on the respective interaction kernels K_j, K_m , and follows immediately from Eq. (3) (see Torre [2005, Chapter 8.3.2]). The time-reversed interaction operator trivially becomes $\mathcal{X}^\dagger \equiv \mathcal{X}_M^\dagger \circ \mathcal{X}_{M-1}^\dagger \circ \dots \circ \mathcal{X}_1^\dagger$.

We will now recursively apply each kernel \mathcal{X}_m^\dagger to a generalized ray, in order to Monte Carlo integrate the light transport through the scene. This process is similar to classical light transport: A generalized ray undergoes interactions with the scene at discretized steps. Because a generalized ray is weakly-local (as will be discussed in Section 5), each interaction is confined to a small region in space. A sequence of interactions gives rise to a *path*, enumerated via the indices of the generalized rays chosen at each interaction, starting with n , the index of the sourced generalized ray from Eq. (20). Let a path of depth $m - 1$ be denoted by a tuple $\boldsymbol{\pi} = (n_1, \dots, n_m)$, where the indices n_i denote sampled generalized rays at each interaction. We define the addition operation of a path with an index as: $\boldsymbol{\pi} + n_{m+1} = (n_1, \dots, n_m, n_{m+1})$, for some index n_{m+1} .

We denote a generalized ray's phase-space representation as the Gaussian WDF $g^\boldsymbol{\pi}$, for an arbitrary path $\boldsymbol{\pi}$, as

$$g^\boldsymbol{\pi}(\vec{r}, \vec{k}) \triangleq g_{\beta, \rho, \boldsymbol{\pi}}(\vec{r}, \vec{k}; \vec{r}_0^\boldsymbol{\pi}, \vec{k}_0^\boldsymbol{\pi}), \quad (22)$$

each defined w.r.t. the values $\beta^\boldsymbol{\pi}, \rho^\boldsymbol{\pi}, \vec{r}_0^\boldsymbol{\pi}, \vec{k}_0^\boldsymbol{\pi}$. Note that Eq. (22) generalizes Eq. (21).

Recursive light transport. Let $\boldsymbol{\pi}$ be some path of depth $m - 1 \geq 0$, i.e. $|\boldsymbol{\pi}| = m$, and $g^\boldsymbol{\pi}$ the generalized ray sampled using that path. For the initial case $m = 1$, $g^\boldsymbol{\pi} = g^{(n)}$ is a sampled sourced generalized ray from the detector, as in Eqs. (20) and (21). Denote the WDF that arises after interaction with \mathcal{X}_m^\dagger :

$$\mathcal{Q}_m^\boldsymbol{\pi} \triangleq \mathcal{X}_m^\dagger \{g^\boldsymbol{\pi}\} = \int d\vec{r}_i d\vec{k}_i K_m^\dagger(\vec{r}, \vec{r}_i, \vec{k}, \vec{k}_i) g^\boldsymbol{\pi}(\vec{r}_i, \vec{k}_i). \quad (23)$$

As interaction kernels fulfil all the properties of a WDF [Testorf et al. 2010, Chapter 1.6], the integration above is a phase-space convolution of a WDF with a Gaussian, hence $\mathcal{Q}_m^\boldsymbol{\pi}$ is indeed a Husimi Q distribution (as defined in Eq. (11)). Note, the above holds only because our generalized ray is a phase-space Gaussian (a WDF is a

non-negative husimi Q distribution if, and only if, it can be written as a convolution of an arbitrary WDF with a Gaussian).

Gaussian functions form an overcomplete functional basis [Rudin 1990]. Therefore, a Husimi Q distribution, being a non-negative, bandwidth-limited distribution (a consequence of the convolution), may always be written as a finite sum of Gaussians to arbitrarily high accuracy, viz.

$$\mathcal{Q}_m^\boldsymbol{\pi}(\vec{r}, \vec{k}) \approx \frac{1}{N_m} \sum_{n_m=1}^{N_m} \alpha^{\boldsymbol{\pi}+(n_m)} g^{\boldsymbol{\pi}+(n_m)}(\vec{r}, \vec{k}), \quad (24)$$

where $\alpha^{\boldsymbol{\pi}+(n_m)} > 0$ are the intensities of each Gaussian above (positiveness is a consequence of the non-negativity of a Husimi Q distribution). Thereby, we have sampled N_m generalized rays with paths $\boldsymbol{\pi} + (n_m)$ of depth m , arising from the action of the operator \mathcal{X}_m on the generalized ray $g^\boldsymbol{\pi}$.

Continue this process recursively, giving rise to sampled generalized rays $g^{\boldsymbol{\pi}_j}$, with $\boldsymbol{\pi}_j$ being the sampled paths, and each generalized ray is defined by the tuple $\{\vec{r}_0^{\boldsymbol{\pi}_j}, \vec{k}_0^{\boldsymbol{\pi}_j}, \beta^{\boldsymbol{\pi}_j}, \rho^{\boldsymbol{\pi}_j}\}$ and $\alpha^{\boldsymbol{\pi}_j}$. The sampled path depth is $|\boldsymbol{\pi}_j| - 1$: For $|\boldsymbol{\pi}_j| = 1$ these generalized rays are sourced from the detector \mathcal{D} , while for $|\boldsymbol{\pi}_j| > 1$ these generalized rays are sourced from a distribution that arises after interaction with $\mathcal{X}_{|\boldsymbol{\pi}_j|-1}^\dagger \circ \dots \circ \mathcal{X}_1^\dagger$.

The rendering equation. The recursive *rendering equation for backward wave-optical light transport*, that drives the process above, takes the familiar form of a Fredholm integral equation:

Rendering equation

$$\mathcal{Q}_o \Big|_{\vec{r}_0} = \int d\vec{k}_0 d\beta d\rho \mathcal{X}_r^\dagger \left\{ g_{\beta, \rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0) \right\}, \quad (25)$$

where \mathcal{X}_r^\dagger is some interaction operator (acts w.r.t. the primed variables), g is an incident generalized ray and the integration is over all possible generalized rays' wavevectors \vec{k}_0 and parameters β, ρ . In contrast to the classical rendering equation that acts upon scalars, Eq. (25) acts upon functions, all generalized rays that are incident to mean position \vec{r}_0 , and its result is a function as well, the distribution \mathcal{Q}_o . Outgoing generalized rays are then sampled from the resulting light distribution \mathcal{Q}_o , in an identical manner to Eq. (24). There is no emission term: interaction with light sources is handled by the measurement operator, introduced next. The interaction operator will be discussed further in Section 6.

Measurement. The recursive light transport process above solves for the total light transport, viz. the term $\mathcal{X}^\dagger \{g^{(n)}\}$ in Eq. (20). The observed intensity is computed by integrating each generalized ray over the sourcing distribution \mathcal{W}_s , quantified by the following *measurement operator*:

Measurement

$$\begin{aligned} \mathfrak{I}\{g\} &\triangleq \int d\vec{r}' d\vec{k}' \mathcal{W}_s(\vec{r}', \vec{k}') g_{\beta, \rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0) \\ &= \frac{1}{(2\pi)^3} \left| \int d\vec{r}' \psi_s(\vec{r}') \psi_{\beta, \rho}^*(\vec{r}'; \vec{r}_0, \vec{k}_0) \right|^2, \end{aligned} \quad (26)$$

where we apply the Moyal formula [Torre 2005, Chapter 6.3.7] in order to reformulate the phase space integration using Wigner distributions into integration in position space using wave functions, and with ψ_s being the sourced wave function. In practice, the integration above happens once a generalized ray is incident upon a light source, and under the support of that light source's WDF.

In Section 7 and Algorithm 1 we will develop a practical light transport algorithm, building upon the derivations in this Section.

5 ANALYSIS

In Section 4 we have derived our primary contributions: the theory of backward wave-optical light transport with generalized rays. We now analyse and validate these results.

Linearity. All formulations above are linear: (i) Sourcing equations, Eqs. (18) and (19), describe a linear (incoherent) superposition of the measured intensities of each detector element; (ii) the rendering equation, Eq. (25), is linear; and, (iii) measurement of a generalized ray, i.e. integration over the sourcing distribution (Eq. (26)), as well as accumulation of these intensities are also linear operations. Generalized rays always carry positive intensities. We discuss the linearity of generalized ray further, from the perspective of Shannon sampling, in Subsection 5.1.

Locality. We need to show that all integrals in our formulae can be restricted to a well-defined finite integration region. A generalized ray is a sharply-peaked Gaussian, thus in practice we ignore its tails and assume that $\psi_{\beta,\rho}(\vec{r}) = 0$ when $|\vec{r} - \vec{r}_0| > \rho$, for some $\rho > 0$. Clearly, this can be done to arbitrary precision: the tail mass of $\psi_{\beta,\rho}$ decays rapidly as a function of ρ . Therefore, all integrations are over the support of a generalized ray: in the sourcing (Eqs. (18) and (20)), measurement (Eq. (26)), and interaction operator acting upon a generalized ray (Eq. (13)), viz. $\mathcal{X}^{-t}\{g\}$, formulae are confined to the (finite) spatial extent of $\psi_{\beta,\rho}$.

We will now show that we may always restrict the integration that defines an arbitrary interaction kernel K^{-t} to a finite region. Consider the action of that kernel (Eq. (13)) on a generalized ray g :

$$\begin{aligned} \mathcal{W} &\triangleq \mathcal{X}^{-t}\{g_{\beta,\rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0)\} \\ &= \int d\vec{r}' d\vec{k}' K^{-t}(\vec{r}, \vec{r}', \vec{k}, \vec{k}') g_{\beta,\rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0). \end{aligned} \quad (27)$$

Substitute the definition of the kernel (Eq. (4)) and rewrite the generalized ray g using the definition of the WDF (Eq. (6)) via its wave function $\psi_{\beta,\rho}$ (Eq. (17)). With some basic algebra and proper variable changes the above becomes

$$\begin{aligned} \mathcal{W} &= \int d\vec{y}' d\vec{y}'' d\vec{x}_o \psi_{\beta,\rho}^*(\vec{y}') h(\vec{y}', \vec{r} + \frac{1}{2}\vec{x}_o) \\ &\quad \times \psi_{\beta,\rho}(\vec{y}'') h^*(\vec{y}'', \vec{r} - \frac{1}{2}\vec{x}_o) e^{-i\vec{k} \cdot \vec{x}_o}. \end{aligned} \quad (28)$$

That is, the system's optical response function $h(\vec{r}_o, \vec{r}_i)$ only contributes to the result \mathcal{W} when $\psi_{\beta,\rho}(\vec{r}_o)$ is non-negligible. Apply the relation $h(\vec{r}_o, \vec{r}_i) = h^*(\vec{r}_i, \vec{r}_o)$ to Eq. (28), and we also deduce that $h(\vec{r}_o, \vec{r}_i)$ only contributes when $\psi_{\beta,\rho}(\vec{r}_i)$ is non-negligible. Hence,

we may always replace the optical response function h with

$$\bar{h}(\vec{r}_o, \vec{r}_i) \triangleq \begin{cases} h(\vec{r}_o, \vec{r}_i) & \text{if } |\vec{r}_o| < \rho \text{ and } |\vec{r}_i| < \rho \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

in the definition of a diffraction kernel (Eq. (4)), thereby limiting the integration to the spatial extent ρ of the incident generalized ray, while remaining accurate to arbitrarily good precision. We have shown that our formalism achieves weak-locality.

Compare the above with related work: If we were to replace the generalized rays with plane waves, viz. $\psi_{\beta,\rho} \propto \exp(-i\vec{k} \cdot \vec{r})$, locality would no longer be recoverable as the integration in Eq. (28) must happen over the entire domain. Similarly, consider a Wigner-based perfectly-local formalism (for example Cuypers et al. [2012]) where we set $g = \delta(\vec{r} - \vec{r}_0) \delta(\vec{k} - \vec{k}_0)$. Then, Eq. (27) becomes $\mathcal{W} = K^{-t}$, and the integration region in the definition of K (Eq. (4)) can no longer be constrained (without sacrificing linearity).

Completeness. Given a photoelectric detector, whose detectable states (its WDF) can be written as Eq. (10), our derivations in Section 4 are exact. We stress that essentially all detectors of interest work via the process of photoelectric detection [Leonhardt 1997, Chapters 3 and 4]. Arbitrary detector geometry and detection properties, quantified by \mathcal{D} , are supported. Interactions of the WDF with the scene via a diffraction kernel K , as in Eqs. (3) and (4), is a general formalism [Testorf et al. 2010], and no restrictions are placed on the total interaction operator \mathcal{X}^{-t} . Likewise, the sourcing distribution \mathcal{W}_s can be arbitrary.

Because (multivariate) Gaussians serve as an overcomplete functional basis, an arbitrary Husimi Q distribution can be written as a finite superposition of Gaussians to arbitrarily good precision. We discuss this further in our supplemental material. Therefore, the recursive light transport process, formalised by Eq. (25), is well defined. The restriction of a generalized ray, as well as the definition of the interaction kernel (Eq. (4)), to a finite spatial region can also be done to an arbitrarily good precision. Therefore, our formalism is complete: able to reproduce any wave-optical effect observable by a photoelectric detector.

5.1 Linearity of Generalized Rays

Consider Young's iconic double-slit interferometer (illustrated in Fig. 3a), and assume we use a coherent laser source. This experiment will be used to (i) lend insight into how generalized rays are always able to maintain linearity, even when the incident illumination is perfectly coherent; and, (ii) numerically validate our formalism (see Subsection 5.2). In addition, while our practical interactive rendering algorithm that we introduce later (Subsection 7.2) neglects free-space diffractions, this experiment demonstrates that our formalism may reproduce such effects.

As the coherent light passes through the slits, it diffracts around the slits, resulting in a (coherent) superposition phasor φ at the screen. However, we do not observe φ at singular points, but only over regions, "pixels", on the screen (blue line on the screen in Fig. 3a). The spatial extent of these pixels must be positive, due to the uncertainty relation [Mandel and Wolf 1995]: we are never able to resolve light at infinite resolution. Because interference is averaged out over the spatial extent of a pixel, oscillations of φ that are

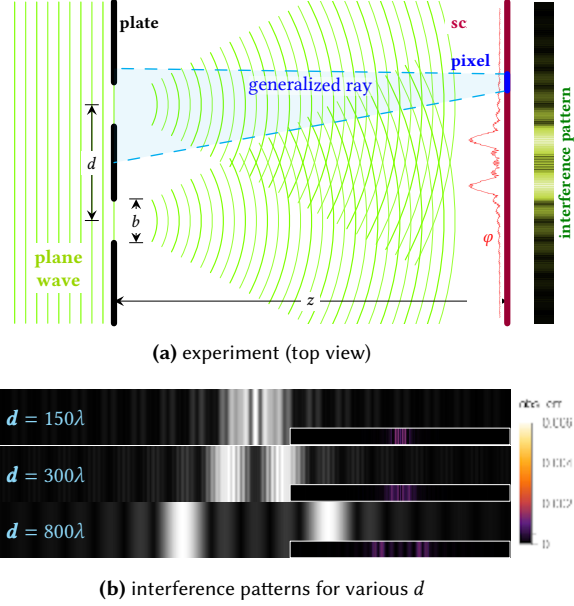


Fig. 3. **Diffraction through double slits.** (a) Schematic of Young’s double slit experiment. A pair of slits, of width b and spaced a distance d apart, are cut in a thin, conductive plate. A coherent plane wave (illustrated in green) diffracts through the slits, and is observed upon a screen, placed at a distance z from the plate. The superposition of coherent light from both slits results in a rapidly-oscillating phasor φ (illustrated in red), producing an interference pattern. (b) The experiment is performed with increasing slit distances d , and we compare our method (sampling incident light with generalized rays) with a ground truth (explicitly diffracting the plane wave through the slits). Differences are plotted in the insets at the bottom right of each pattern (intensity of peak fringe is 1). The experiment was performed with wavelength $\lambda = 1$ (arbitrary units), $z = 10000\lambda$ and $b = 40\lambda$.

more rapid than the pixel’s extent do not contribute to observable interference effects.

Let points $\vec{s}_{1,2}$ located on the slits act as point sources (i.e., the Huygens–Fresnel principle). Given identical peak amplitudes, their superposition phasor at a point \vec{r} on the screen can be written as

$$\varphi \approx \frac{e^{-ikz}}{z} \left[\exp\left(-i\frac{k}{z}\vec{s}_1^\perp \cdot \vec{r}^\perp\right) + \exp\left(-i\frac{k}{z}\vec{s}_2^\perp \cdot \vec{r}^\perp\right) \right], \quad (30)$$

where k is the wavenumber, as before, the superscript \perp denotes projection upon the plate plane ($z = 0$), and we made the typical Fraunhofer (far-field) approximation [Born and Wolf 1999] (note, we only make this approximation for the illustrative analysis here, the results in Figs. 3b and 4 were obtained using exact formulae). The observed intensity of this phasor is then proportional to

$$|\varphi|^2 \propto \frac{2}{z^2} \left(1 + \cos \left[\frac{k}{z} (\vec{s}_1^\perp - \vec{s}_2^\perp) \cdot \vec{r}^\perp \right] \right). \quad (31)$$

The interference term above oscillates with an angular frequency of $k\frac{l_s}{z}$ (as a function of screen position \vec{r}^\perp), where $l_s = |\vec{s}_1^\perp - \vec{s}_2^\perp|$ is the distance between the points on the slits.

We now make contact with sampling theory: By the Shannon sampling theorem, we may resolve that interference term without aliasing, only if we observe the phasor over a spatial extent (i.e.,

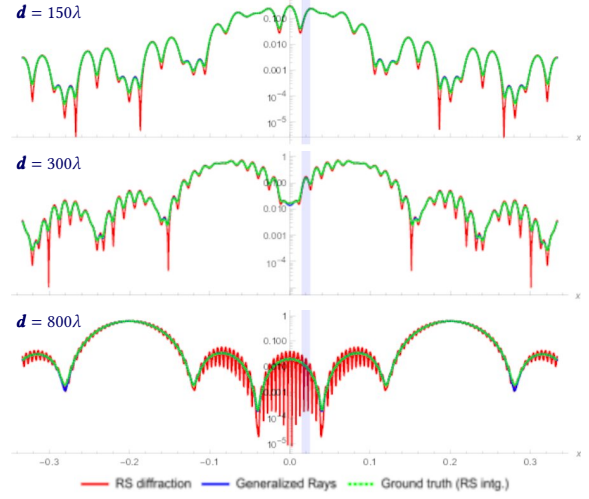


Fig. 4. Plots of light intensity as a function of x (position on screen) of the experiment in Fig. 3. In red we plot the exact Rayleigh-Sommerfeld (RS) diffraction, i.e. the *unobservable* interference that arises in singular points. The spatial extent of a detector on the screen (a pixel in each pattern in Fig. 3b) is illustrated as a cyan bar. Integrating the RS diffraction over that spatial extent of a pixel computes the numeric ground truth, plotted in dashed green. Results obtained with generalized rays are plotted in blue.

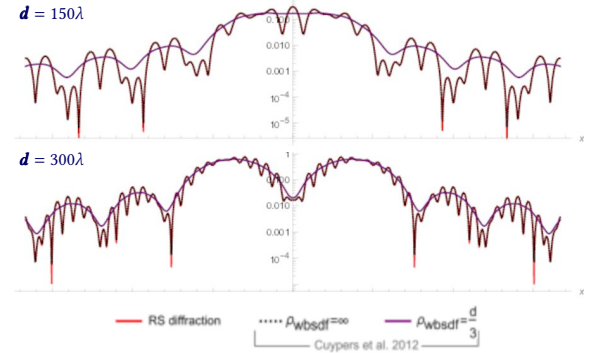


Fig. 5. **Loss of locality with Cuypers et al. [2012].** Red plot (exact RS diffraction) is as in Fig. 4. Let ρ_{wbsdf} be the radius of integration of a diffraction kernel (Eq. (4)) in Cuypers et al. [2012]. Limiting that integration radius produces erroneous results, hence *locality is entirely lost* with their method: correctness requires integration over the entire scene ($\rho_{\text{wbsdf}} = \infty$).

a pixel) no greater than $\frac{1}{2k(l_s/z)}$. Under the setting of our small-angle approximation, l_s/z approximates the angle between \vec{s}_1, \vec{s}_2 subtended from \vec{r} , denoted θ_s . Denote the spatial extent of a pixel as l_r , and we establish the *sampling relation*: $l_r(k\theta_s) \leq 1/2$. This relation resembles the uncertainty relation [Mandel and Wolf 1995, Chapter 4] but with the inequality reversed (l_r is the spatial variance and $k\theta_s$ is the angular variance scaled by the wavenumber).

When the angular extent θ_s between the interference sources $\vec{s}_{1,2}$ is large enough to violate the sampling relation above, i.e. $l_r(k\theta_s) > 1/2$, the observer is no longer able to resolve the interference term. Generalized rays at the detector ($\rho = 0$) are minimum-uncertainty

constructs, i.e. fulfil the equality in the sampling relation. Therefore, generalized rays quantify exactly the spatial and angular extent over which interference effects may be resolved. The arguments above are not limited to double-slit diffraction, and apply, in general, to any superposition from two or more point sources. To conclude: it is the *integration over the observer's spatial extent that induces decoherence*, allowing generalized rays to regain linearity under any illumination conditions.

We illustrate that decoherence in Fig. 3b, where we perform the experiment with increasing slit distances. The standard deviation of the spatial extent of a generalized ray (illustrated with dashed cyan lines in Fig. 3a) at the slits is about 100λ , i.e. a full width at half maximum of $\sim 235\lambda$. While the characteristic double-slit interference pattern is visible at first (when d is smaller than the spatial extent of a generalized ray), it slowly dissolves into a pair of (mutually-incoherent) single-slit diffraction patterns. Observe the increasing frequency of the secondary fringes in Fig. 3b with increasing d .

5.2 Numeric Validation

We perform the described double-slit experiment using a pair of methods: (i) forward transport, where we diffract the incident plane wave directly via the exact Rayleigh-Sommerfeld (RS) explicit diffraction integral of the first kind [Mandel and Wolf 1995, Chapter 3.2], and integrate over the detector (each pixel); and (ii) backward transport using our formalism, i.e. sampling with generalized rays and diffracting them through the slits over their spatial extent. The absolute difference between the two methods is shown in the inset at the bottom right of each pattern in Fig. 3b (intensity is normalized such that the peak fringe of each pattern has intensity of 1). The differences (due to ignoring generalized rays' tail) are negligible.

Once the distance d becomes larger than the spatial extent of a generalized ray that reaches the slits, generalized rays no longer solve a double-slit diffraction problem. Indeed, they don't need to: interference between the two slits does arise at singular points, but is not observable over a pixel's extent. This can be seen in Fig. 4, where for large d we can see that the rapidly-oscillating double-slit interference pattern arises everywhere, but is integrated out over the spatial extent of a pixel. In contrast to generalized rays, explicit diffraction integrals (first method above) have no means of quantifying the extent over which observable interference may arise. This means that the explicit method needs to (wastefully) integrate each sample over each pixel and over the entire plate. Therefore, our method (second method) is an order-of-magnitude faster than the explicit method (8.9 s compared to 246 s). This experiment demonstrates that generalized rays achieve well-defined weak locality.

Other methods suffer from similar deficiencies: they fail to derive machinery that is able to quantify the extent over which weak-locality can be maintained. For example, Cuypers et al. [2012] describe a WDF-based formalism with perfectly local primitives. However, as discussed in Section 2, such a formalism forgoes either locality or linearity: formulating a light-matter interaction requires integration over the entire scene, as defined by the interaction kernel (Eq. (4)). In Fig. 5 we show that if we were to limit the spatial extent of integration of an interaction kernel, while insisting on

linearity (ignore the bilinear superposition term in Eq. (5)), incorrect results are produced. Indeed, no wave-optical formalism may be simultaneously perfectly local and linear.

6 INTERACTION KERNELS

Developing a light-matter interaction operator entails deriving an analytic method to compute the diffracted distribution \mathcal{Q}_o which arises from the interaction operator acting on an arbitrary generalized ray, and express it as a finite sum of generalized rays, viz. Eq. (24). Formally, this can always be done with arbitrarily high accuracy (as discussed in Section 3). However, doing so in practice can be analytically-involved, and at times application-specific assumptions are made in order to simplify the analysis. Often, an importance sampling strategy, that is used to sample a generalized ray out of the sum in Eq. (24), is desired. In this Section we present a few important example interactions.

Because time-reversal is equivalent to wavevector reversal and phase conjugation [Geru 2018, Chapter 2.3], we may rewrite the action of an arbitrary time-reversed interaction operator on a generalized ray (that enters the rendering equation Eq. (25)) as

$$\mathcal{X}^{-t}\{g_{\beta,\rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0)\} = \mathcal{X}\{g_{\beta,\rho}(\vec{r}', \vec{k}'; \vec{r}_0, -\vec{k}_0)\} \quad (32)$$

(recall that the WDF is real). Instead of deriving time-reversed interactions, we only need to reverse the generalized ray's wavevector.

We classify light-matter interactions into two types:

- (1) SIMPLE — These arise with *linear optical systems* (also known as “ABCD systems”): propagation through a homogeneous medium with a slowly-varying refractive index (including free space), and reflection or refraction at a smooth interface. Under simple interactions the dynamics are identical to ray-optical dynamics (see our supplemental material).
- (2) DIFFRACTIVE — All other interactions, e.g., scattering by a rough surface: there are many ways to formalise a diffractive interaction, and we will discuss a few examples.

Generalized rays are unique in simultaneously (i) behaving like classical rays under simple interactions; (ii) and being *Wigner-representable* [Torre 2005], i.e. they admit well-defined wave functions. The first point only holds for constructs that are fully defined by their 2nd-order moments matrix (like a generalized ray, which is a Gaussian beam), while the second point only applies to phase-space constructs that fulfil the uncertainty relation. For example, the second point does not hold for perfectly-local Wigner-based formalisms (the WDF $\mathcal{W} \equiv \delta(\vec{r}) \delta(\vec{k})$ admits no wave function).

Free-space propagation (simple interaction). Let a generalized ray be parameterized by its mean spatial position \vec{r}_0 , mean wavevector \vec{k}_0 , and β, ρ , as before. Using its phase-space representation (Eq. (7)), observe that we may express its spatial and wavevector variances as $\sigma_r^2 = \beta^2/2$, and $\sigma_k^2 = (1 + \rho^2)/(2\beta^2)$, respectively. σ_r quantifies the spatial extent occupied by the generalized ray, and σ_k/k_0 the solid angle into which the generalized ray propagates.

Let $z > 0$ be the distance of propagation, and $\bar{z} = z/|\bar{\mathbf{k}}_0|$ the distance normalized by the mean wavevector. Then, the corresponding kernel (Eq. (4)) is

$$K_{\text{SPACE}}^{\text{FREE}}(\bar{\mathbf{r}}, \bar{\mathbf{r}}', \bar{\mathbf{k}}, \bar{\mathbf{k}}') \triangleq \delta(\bar{\mathbf{k}}' - \bar{\mathbf{k}}) \delta(\bar{\mathbf{r}}' + \bar{z}\bar{\mathbf{k}} - \bar{\mathbf{r}}), \quad (33)$$

i.e. propagation in direction $\bar{\mathbf{k}}$. Applying Eq. (13), the diffracted (propagated) distribution after interaction becomes

$$\mathcal{Q}_o(\bar{\mathbf{r}}, \bar{\mathbf{k}}) = \mathcal{X}_{\text{SPACE}}^{\text{FREE}} \left\{ g_{\beta, \rho}(\bar{\mathbf{r}}', \bar{\mathbf{k}}'; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0) \right\} = g_{\beta_o, \rho_o}(\bar{\mathbf{r}}, \bar{\mathbf{k}}; \bar{\mathbf{r}}_0 + \bar{z}\bar{\mathbf{k}}_0, \bar{\mathbf{k}}_0) \quad (34)$$

with $\beta_o^2 = \beta^2 + \bar{z}(2\rho + 2\bar{z}\frac{1+\rho^2}{2\beta^2})$ and $\rho_o^2 = (\rho + 2\bar{z}\frac{1+\rho^2}{2\beta^2})^2$.

The reader may verify that after propagation $\bar{\mathbf{k}}_0, \sigma_k^2$ remain invariant, i.e. the direction and solid angle into which the generalized ray propagates do not change; and, the spatial variance transforms as $\sigma_r^2 \rightarrow \sigma_r^2 + O(\bar{z}^2\sigma_k^2)$, i.e. the space occupied by the generalized ray increases proportionally to the propagation distance times the solid angle. Finally, the mean position $\bar{\mathbf{r}}_0$ is shifted by z . That is, all the generalized ray's parameters transform under classical ray-optical dynamics, as would be expected with free-space propagation.

Reflection/refraction at an interface (simple interaction). Assume a smooth, flat interface. Let \mathcal{X} be the interaction operator, quantifying the action of reflection or refraction at that interface. Let $\bar{\mathbf{k}}$ be the mean wavevector of the incident generalized ray, and we denote $\bar{\mathbf{k}}_0^{(r)}$ as that wavevector after the reflection (by the law of reflection) or refraction (by Snell's law) at the interface. Let a be the Fresnel coefficient for the interaction. Then,

$$\mathcal{Q}_o(\bar{\mathbf{r}}, \bar{\mathbf{k}}) = \mathcal{X}_{\text{REFLECT}}^{\text{REFLECT}} \left\{ g_{\beta, \rho}(\bar{\mathbf{r}}', \bar{\mathbf{k}}'; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0) \right\} = ag_{\beta, \rho}(\bar{\mathbf{r}}, \bar{\mathbf{k}}; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0^{(r)}). \quad (35)$$

The parameters $\bar{\mathbf{r}}_0, \beta, \rho$ are unchanged as no propagation takes place. That is, with generalized rays this interaction (and other simple interactions) mimics the classical dynamics. Reflection and refraction are polarization-dependent phenomena. Vectorization can be done trivially (generalized ray per each polarization component). We handle polarization differently, see Subsection 7.2.

Diffraction grating (diffractive interaction). A benefit of generalized rays is that they admit a well-defined wave function, $\psi_{\beta, \rho}$. Hence, to formulate a diffractive interaction one may work in phase space with WDFs, or in position space with wave functions (via a diffraction integral, electromagnetic theory, etc.). Furthermore, the generalized ray's wave function is a simple (coherent) Gaussian beam, which has been extensively studied in optical literature.

Apply the Fraunhofer diffraction integral [Born and Wolf 1999] to a one-dimensional sinusoidal grating of period Λ and height h , yielding

$$\mathcal{Q}_o(\bar{\mathbf{r}}, \bar{\mathbf{k}}) = \mathcal{X}_{\text{GRN}}^{\text{GRN}} \left\{ g_{\beta, \rho}(\bar{\mathbf{r}}', \bar{\mathbf{k}}'; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0) \right\} \approx \sum_j a_j g_{\beta, \rho}(\bar{\mathbf{r}}, \bar{\mathbf{k}}; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0^{(j)}), \quad (36)$$

where $a_j = J_j(hk/2)^2$ is the intensity of the j -order diffracted lobe, J_j is the Bessel function of 1st kind and $\bar{\mathbf{k}}_0^{(j)}$ is the diffracted wavevector in direction $\sin \theta_o = \sin \theta_i - j\frac{\lambda}{\Lambda}$, where $\theta_{i,o}$ are the incident and diffraction directions w.r.t. the grating direction and λ is the wavelength. The diffraction grating also very slightly enlarges the correlation constant ρ , but we ignore that effect for simplicity.

Scatter by moderately-rough surface (diffractive interaction). We make a simplifying assumption: ignore the beam curvature of a generalized ray, viz. $\psi_{\beta, \rho} \equiv e^{i\bar{\mathbf{k}}_0 \cdot (\bar{\mathbf{r}} - \bar{\mathbf{r}}_0)} e^{-\frac{1}{2\beta^2} |\bar{\mathbf{r}} - \bar{\mathbf{r}}_0|^2}$. This serves to transform the Gaussian beam into a spatially-modulated plane wave, easing the analysis. We then may apply the Harvey-Shack surface scatter theory [Krywonos 2006]:

$$\mathcal{Q}_o(\bar{\mathbf{r}}, \bar{\mathbf{k}}) = \mathcal{X}_{\text{SURFACE}} \left\{ g_{\beta, \rho}(\bar{\mathbf{r}}', \bar{\mathbf{k}}'; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0) \right\} \approx \int d\bar{\mathbf{k}}_0^{(r)} f_{\text{HS}}(\bar{\mathbf{k}}_0 \rightarrow \bar{\mathbf{k}}_0^{(r)}) g_{\beta, \rho}(\bar{\mathbf{r}}, \bar{\mathbf{k}}; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_0^{(r)}), \quad (37)$$

where f_{HS} is the Harvey-Shack BRDF, quantifying the scattering for incident and diffracted wavevectors $\bar{\mathbf{k}}_0$ and $\bar{\mathbf{k}}_0^{(r)}$, respectively. In practice, sampling generalized rays from the distribution \mathcal{Q}_o is done via Monte-Carlo integrating the expression above, thereby rewriting it in the form of Eq. (24):

$$\mathcal{Q}_o(\bar{\mathbf{r}}, \bar{\mathbf{k}}) \approx \frac{1}{J} \sum_j \frac{f_{\text{HS}}(\bar{\mathbf{k}}_0 \rightarrow \bar{\mathbf{k}}_{0,j}^{(r)})}{p_j} g_{\beta, \rho}(\bar{\mathbf{r}}, \bar{\mathbf{k}}; \bar{\mathbf{r}}_0, \bar{\mathbf{k}}_{0,j}^{(r)}) \quad (38)$$

(up to a normalization constant), where $\bar{\mathbf{k}}_{0,j}^{(r)}$ are the set of J sampled mean scattering directions, and p_j are the sampling probabilities. In our implementation, the Harvey-Shack BRDF f_{HS} is important sampled using the technique described by Holzschuch and Pacanowski [2017].

In Eqs. (36) and (37) we make optical approximations that are reasonable for our target application: rendering at optical frequencies. One consequence of using the (approximative) Harvey-Shack model for surfaces is that we only model the averaged scatter, where every generalized ray interacts with the entire distribution of surface frequencies. Therefore, surface imperfections, such as glints, or optical speckle, do not arise. We stress that the assumptions we made here are not a limitation of our light transport formalism. Also note that, in general, \mathcal{X} may describe cross-wavelength scattering, e.g., due to fluorescence or phosphorescence, however we ignore such effects in our implementation.

Other materials that we use in our rendered scenes are also derived under the assumption that we ignore the curvature of a generalized ray. Their interaction formula then echoes Eq. (37), but with the BRDF f_{HS} replaced with the BRDF of the relevant interaction that acts on plane waves.

7 WAVE-OPTICAL RENDERING

In this Section we introduce our rendering algorithm. We start with the accurate algorithm that applies our light transport formalism, as derived in Section 4, exactly. Later, in Subsection 7.2 we show that with a few reasonable assumptions, our formalism gives rise to an interactive wave-optical rendering algorithm that is easy to implement in a modern path tracer.

Our wave-optical light transport formalism works by propagating backwards, from the detector, generalized rays and simulating their interactions with the scene under time-reversed dynamics, and finally measuring their contributions by integrating over the sourcing distributions of light sources. A generalized ray is a time-reversed Gaussian beam sourced from the detector: Eqs. (7) and (17) define its wave function and WDF, respectively.

Algorithm 1: Our rendering algorithm.

```

1 foreach detector (pixel)  $\mathcal{D}$  do
2    $I \leftarrow 0$ ; /* will hold the computed detected intensity */
3   for  $n \leftarrow 1$  to  $N$  do /* sample the backward light transport */
4      $\{\vec{r}_0, \vec{k}_0, \beta, \rho\} \leftarrow$  source generalized ray; /* Eq. (20) */
5      $\alpha \leftarrow \frac{\alpha^{(n)}}{Np^{(n)}}$ ; /* path weight; p is sampling probability */
6     while path depth < max depth do
7       propagate the generalized ray  $\{\vec{r}_0, \vec{k}_0, \beta, \rho\}$ ;
8       /* measurement, or solve pass (Subsection 7.1) */
9        $I \leftarrow I + \alpha \cdot \mathfrak{S}\{g_{\beta, \rho}(\vec{r}', \vec{k}'; \vec{r}_0, \vec{k}_0)\}$ ; /* Eq. (26) */
10      /* light-matter interaction */
11       $\mathcal{X} \leftarrow$  interaction operator around  $\vec{r}_0$ ;
12       $\mathcal{Q}_o \leftarrow \mathcal{X}\{g_{\beta, \rho}(\vec{r}', \vec{k}'; \vec{r}_0, -\vec{k}_0)\}$ ; /* rendering equation */
13       $\{\vec{r}_0, \vec{k}_0, \beta, \rho\} \leftarrow$  sample  $\mathcal{Q}_o$ ; /* Eq. (24) */
14       $\alpha \leftarrow \alpha \frac{\alpha^{\pi+(nm)}}{N_m}$ ; /* update weight ( $\alpha^{\pi+(nm)}$ ,  $N_m$  as in
15      Eq. (24) */

```

Crucially, generalized rays are always **linear** (generalized rays never interfere), and **weakly local**: we limit the spatial extent of the generalized ray to some finite radius around its mean position \vec{r}_0 . Being Gaussian beams, with rapidly-decaying tails, this can be easily done to arbitrarily high accuracy. For example, a radius of $\varrho = 2.5\beta$, where $\beta^2/2$ is the generalized ray’s spatial variance, captures $> 99.8\%$ of its weight (such ϱ is used in Fig. 3). Finally, this formalism is **complete** (further discussed in Section 5).

In Algorithm 1 we summarise our general-purpose, accurate light transport algorithm:

- **Sourcing.** A generalized ray is fully described by its mean position \vec{r}_0 , mean wavevector \vec{k}_0 (quantifying direction of propagation and the wavelength), spatial variance $\beta^2/2$ and correlation parameter ρ . For each sample, these parameters are sourced via Eq. (20), thereby sourcing a generalized ray initiating a path.
- **Propagation.** We propagate the generalized ray through free space, until it (or a part of it) encounters matter. As we limit the spatial extent of a generalized ray to a region (ball of radius ϱ) around its mean \vec{r}_0 , tracing it reduces to beam (or cone) tracing through the scene. We formalize the transformation of the generalized ray’s parameters, $\{\vec{r}_0, \vec{k}_0, \beta, \rho\}$, under free-space propagation in Eq. (34).
- **Measurement.** If the generalized ray (or part of it) encounters a light source, we integrate the generalized ray over the sourcing distribution, as formalised by the measurement operator (Eq. (26)). When no light source is present, $\mathfrak{S} \equiv 0$.
- **Light-matter interaction.** We compute the interaction operator \mathcal{X} (Eq. (13)), that quantifies the action of matter that falls within the spatial extent of the beam (ball of radius ϱ centred at \vec{r}_0). This operator acts upon the traced generalized ray (with wavevector reversed, as in Eq. (32)), giving rise to the scattered distribution \mathcal{Q}_o (as in Eq. (23)). Then, a

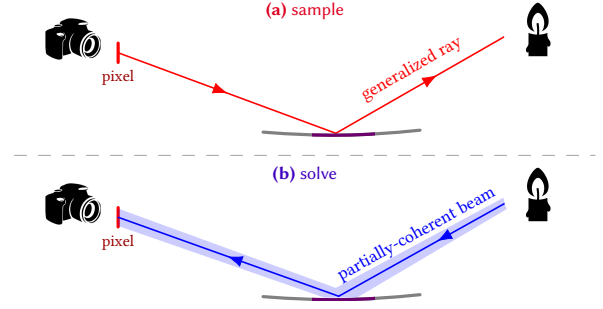


Fig. 6. **Sample-solve.** Our path tracing algorithm (a) uses generalized rays (dotted lines) to *sample* paths through the scene. Generalized rays are always linear, therefore classical sampling techniques apply essentially unchanged. Once a path is sampled (solid red path), we (b) *solve* for the partially-coherent light transport, by applying PLT [Steinberg et al. 2022] from the light source to the sensor across all intermediate interactions.

diffracted generalized ray is sampled out of this distribution and the light transport simulation continues.

The tuple π is a bookkeeping variable that keeps track of the indices of the generalized rays that were sampled at each step.

The algorithm presented above is a novel wave-optical rendering algorithm. As formulated, it is able to simulate rigorous wave-optical light transport accurately (to arbitrarily good precision) using weakly-local, linear generalized rays. This algorithm is employed for the simulation in Fig. 3. A pair of practical difficulties may arise:

- (1) In contrast to classical light transport, propagation cannot be done via ray tracing and requires significantly-slower beam tracing, as generalized rays are not perfectly local (indeed, no wave-optical formalism may be simultaneously perfectly local and linear).
- (2) Because generalized rays occupy a positive extent, the operator \mathcal{X} in the **light-matter interaction** stage may need to account for multiple different materials simultaneously (for example, when a generalized ray is incident upon an edge or a slit, or upon a surface with spatially-varying scattering characteristics). This can be understood as the interaction being a composition of two or more operators, viz. $\mathcal{X} \equiv \mathcal{X}_1 \circ \mathcal{X}_2$. Formally, this can always be done, but, in practice, formalising the composed interaction and sampling generalized rays from it (thereby accounting for the mutual interference of the scattered generalized rays from all materials that are involved in the interaction) can be difficult.

An efficient implementation of beam tracing, designed for modern GPUs, as well as deriving some composite interactions of interest (e.g., free-space diffraction of a generalized ray around geometry) would be of great interest, and are left for future work. Such work would be especially important for wave-optical simulation at longer wavelengths (e.g., RADAR). In the rest of this paper our focus is on rendering at optical frequencies. We present next our interactive wave-optical rendering implementation, where we make a simplifying application-specific assumption.

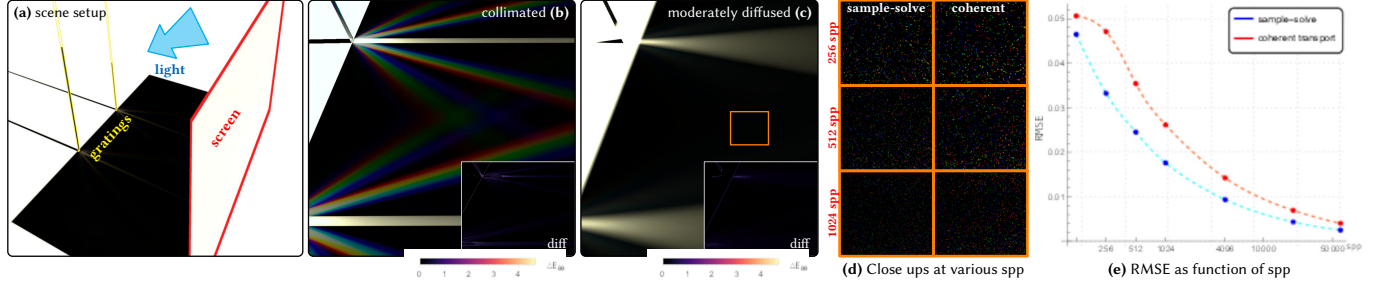


Fig. 7. **Sample-solve: PLT as a variance reduction technique.** During the solve step, we apply PLT to solve for the partially-coherent light transport over a sampled light path. (a) Light arrives from the right, illuminating a simple scene. A screen (on the right, red outline) shadows a rectangular area. On the left a pair of thin diffraction gratings (yellow outline) reflect and disperse light. (b) When the incident light is highly collimated (subtends a very small solid angle from the scene), hence is moderately coherent, the diffraction lobes are clearly visible. (c) As we increase the light’s diffusivity (increase its solid angle), the reflection from the direct contribution lobe spreads out, while diffraction lobes mostly disappear, as expected. As the only light that arrives to the shadowed region is from the diffraction gratings, this is a good scene to study the benefits of the solve stage, and to do so we render the scene with only fully-coherent transport (no PLT applied). (b-c, insets) Difference images between the partially-coherent and fully-coherent sample-solve show that indeed both converge to an identical result. (d) However, close ups on the region outlined in orange in (c) show that, while the diffraction lobes are no longer visible, they still induce error, which the partially-coherent solve stage serves to reduce. (e) Plot of error in that area as function of sample count suggests that fully-coherent transport requires about 2-5 times the sample count to achieve similar-quality renderings.

7.1 Sample-Solve

In this Subsection we draw a formal connection between our backward (sensor-to-source) formalism to *optical coherence*. The purpose of this is to connect our formalism to other forward-based models and computational optics tools, thereby enabling the application of bi-directional techniques. For our application of interest, we will show that we may use a forward PLT pass as a cheap variance reduction technique.

Relation to optical coherence. The central quantity in the study of optical coherence [Wolf 2007] is the *cross-spectral density* (CSD) of a statistical ensemble of light waves, viz.

$$\mathcal{C}(\vec{r}_1, \vec{r}_2) \triangleq \langle \psi(\vec{r}_1) \psi^*(\vec{r}_2) \rangle, \quad (39)$$

where $\langle \cdot \rangle$ denotes ensemble-averaging over the wave ensemble, and the wave function ψ is now understood as a realization from that ensemble. From the definitions of the WDF and the CSD, Eqs. (4) and (39), it is easy to observe that the CSD is the Fourier transform of the ensemble-averaged WDF, therefore we may write

$$\mathcal{C}\left(\vec{r} - \frac{1}{2}\vec{d}, \vec{r} + \frac{1}{2}\vec{d}\right) \propto \left\langle \mathcal{F}\left\{\mathcal{W}(\vec{r}, \vec{k}')\right\}(\vec{d})\right\rangle \quad (40)$$

(up to an irrelevant constant), where the Fourier transform \mathcal{F} is w.r.t. the primed variable, \vec{k}' . We may immediately conclude that optical coherence at a fixed point \vec{r} only depends on the (ensemble-averaged) light’s wavevector distribution in phase space.

Under partial coherence, the observed values are often ensemble-averaged values [Wolf 2007] (and can be understood as time averaging over the period of detection), i.e. $\langle I \rangle$. Then, partial coherence is trivially accounted for by replacing the sourcing WDF \mathcal{W}_s with its ensemble-averaged counterpart, $\langle \mathcal{W}_s \rangle$, in all our formulae. $\langle \mathcal{W}_s \rangle$ is the Fourier transform of the sourced CSD of light (Eq. (40)).

We term the *diffusivity* Ω of a WDF to be the angular variance of propagation from the mean direction of propagation. The diffusivity Ω is a 2×2 positive-definite matrix Ω , allowing for anisotropy. For example, in the isotropic case, the diffusivity can be written

as σ_k^2/k_0^2 , where σ_k^2 is the distribution’s wavevector variance. The variance in the solid angle into which the bundle propagates is then $\Omega = |\Omega|$. Then, from Eq. (40) we may formally derive the following result (see Subsection S4.2 in our supplemental material):

$$\Theta = \lambda^2 \Omega^{-1}, \quad (41)$$

where Θ is the *shape matrix* from PLT theory [Steinberg et al. 2022], i.e. the (inverse) 2nd-order moments of spatial variance. Deriving equality relations between higher-order moments is also possible, however we do not require higher-order moments.

A consequence of the above relation between optical coherence and the diffusivity of light is the well-known connection between the coherence area of light, i.e. $|\Theta|$, and the solid angle subtended by a thermal source [Mandel and Wolf 1995]: $|\Theta| = \frac{\lambda^2}{\Omega}$. Though note that the relation Eq. (41) is more general, and establishes a direct connection between optical coherence and light’s distribution in the wave-optical phase space, with no assumptions on the light sourcing process or its state-of-coherence at other regions in space.

Sample-solve. We present a simple two pass algorithm: first, we *sample* the wave-optical distribution of light using generalized rays, using the rendering algorithm we presented so far, Algorithm 1. The process continues until a generalized ray encounters a light source, thereby a path π that connects the detector to a light source is found. The effective diffusivity over that path is well defined. In place of the *measurement* stage (line 9 in Algorithm 1), we now apply a forward pass: we use PLT machinery [Steinberg et al. 2022] to *solve* for the partially-coherent light transport over the path π . More formally, instead of integrating over the ensemble-averaged sourcing WDF, $\langle \mathcal{W}_s \rangle$, we use PLT to source a partially-coherent beam that corresponds to $\langle \mathcal{W}_s \rangle$ from that light source, and retrace the steps forward (source-to-sensor) taken by the generalized ray over the path. As the PLT beam captures more information (it is “wider” than a generalized ray), this forward solve pass serves as a

variance-reduction technique, by computing the partially-coherent optical response for the sampled path. See Fig. 7.

In our domain of interest—wave-optical rendering—applying PLT makes sense: The optical coherence of light is a primary factor in limiting our ability to resolve wave-interference effects. Other applications might find a different *solve* pass to be more appropriate: For example, integrating optical speckle statistics can be done via a sample-solve approach, where first we *sample* paths connecting the detector to a light source, and then integrate statistics in a *solve* pass. Such applications are beyond the scope of this paper, however they serve to highlight the generality of this simple sample-solve approach: it bridges a gap between classical path tracing tools and wave optics, via the generalized ray construct, and enables the application of powerful sampling techniques in a wider context.

7.2 Interactive Rendering

In similar manner to optical rendering with PLT [Steinberg et al. 2022], we make the assumption that the spatial extent ϱ of a generalized ray is smaller or comparable to the scene’s smallest geometric details (we refer to the explicit geometry against which we trace rays, and not analytically defined details like surface properties). Therefore, we may neglect the generalized ray’s spatial extent and use ray tracing to propagate it. Then, the **propagation** stage in our rendering algorithm, Algorithm 1, reduces to simple ray tracing, and in the **light-matter interaction** stage we do not consider composite interactions with multiple materials.

The simplifying assumption above is reasonable for our target application: high-performance rendering at optical frequencies. In terms of optical effects, this assumption means that we neglect free-space diffractions (light “bending” around geometry) and interference across distinct materials. We stress that this assumption is not a limitation of our wave-optical light transport formalism, as presented in Section 3.

Under the assumption above, surprisingly few changes are required to convert a classical unidirectional, backward path tracer into a path tracer compatible with wave optics: generalized rays replace the classical ray as the “point queries” of light’s behaviour. In practice, doing so requires essentially only a few changes: (i) transitioning to a spectral, vectorized (polarization-aware) path tracer; (ii) transforming generalized rays on free-space propagation, as in Eq. (34); and (iii) reformulating all the BSDFs in terms of generalized rays, as in Eq. (37).

Implementation details. To **source** a generalized ray from the detector distribution \mathcal{D} of a pixel: the variance $\beta^2/2$ is taken to be a fraction of the pixel spatial size; $\rho = 0$ at sourcing; for simplicity, we set the detection efficiency $\alpha = 1$; the initial mean position \vec{r}_0 is sampled uniformly at random on the pixel; and, $-\vec{k}_0$ is the (mean) wavevector. A wavelength λ_0 (such that $|\vec{k}_0| = \frac{2\pi}{\lambda_0}$) is drawn randomly at uniform from the visible spectrum. This wavelength drives the path sampling process, hence we refer to it as the “Hero wavelength” [Wilkie et al. 2014].

The **propagation** of a generalized ray remains unchanged, as described in Section 7 and formalized by Eq. (34). **Light-matter interactions** are done as described in Section 6, and we continue to ignore

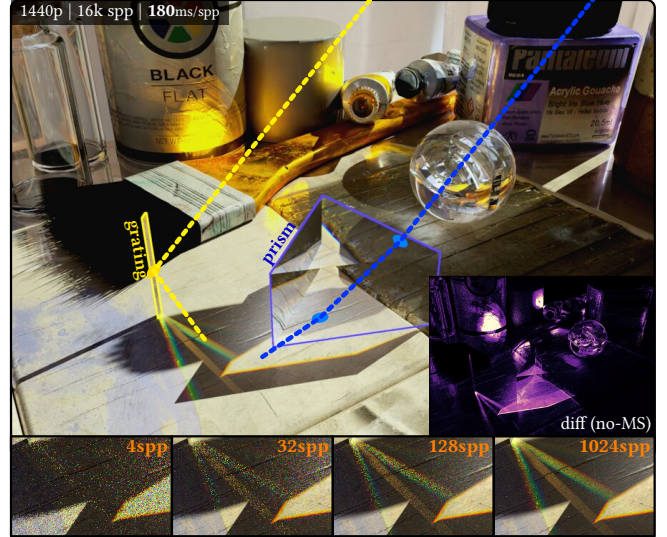


Fig. 8. **Manifold sampling.** Example application of an advanced sampling technique. When performing next event estimation (NEE), manifold sampling (MS) [Hanika et al. 2015; Zeltner et al. 2020] enables finding a light path between a surface and a light source across one or more dielectric interfaces. In the rendered scene, such a sampled path, that refracts through the dispersive prism (outlined in blue), is visualized via the dotted blue line. We also employ MS for NEE on specular reflections: note the reflections off the paint brush’s metal handle and off the paint tubes, as well as the thin diffraction grating (outlined in yellow) dispersing light into multiple diffraction lobes. See full high-resolution rendering in our supplemental material. (inset) Colour-coded difference compared to a without-MS rendering. (bottom) The diffraction grating lobes at increasing samples-per-pixel (spp).

the generalized ray’s curvature for diffractive interactions. Diffractive interaction formulae then take the form of Eq. (37), and we are able to importance sample these interactions (i.e., importance sampling the function f in Eq. (37)).

We continue to transform the generalized ray’s parameters, viz. $\{\vec{r}_0, \vec{k}_0, \beta, \rho\}$, on propagation and on interaction with matter. Therefore, we are able to quantify exactly the spatial extent over which a generalized ray interacts with matter. Materials that describe deterministic scattering features, like a surface’s explicit microgeometry, or a multilayered stack (as seen in the beetle’s wings in Fig. 1), can be accurately rendered. This is not possible with any existing linear wave-optical formalisms: none of them are able to recover weak locality, and would require integration over the entire material.

In addition, when a generalized ray encounters matter, we attempt to connect the path to a light source via next-event-estimation (NEE). We also perform “Russian roulette” early path termination. We note that NEE, “Russian roulette”, and importance sampling are all standard sampling techniques in path tracing [Pharr et al. 2016]. We are also able to apply other advanced sampling techniques in our wave-optical rendering algorithm: To demonstrate that we have implemented manifold sampling [Hanika et al. 2015; Zeltner et al. 2020], see Fig. 8.

As discussed in Subsection 7.1, the **measurement** stage in Algorithm 1 is replaced with a PLT solve pass. Every time a path has

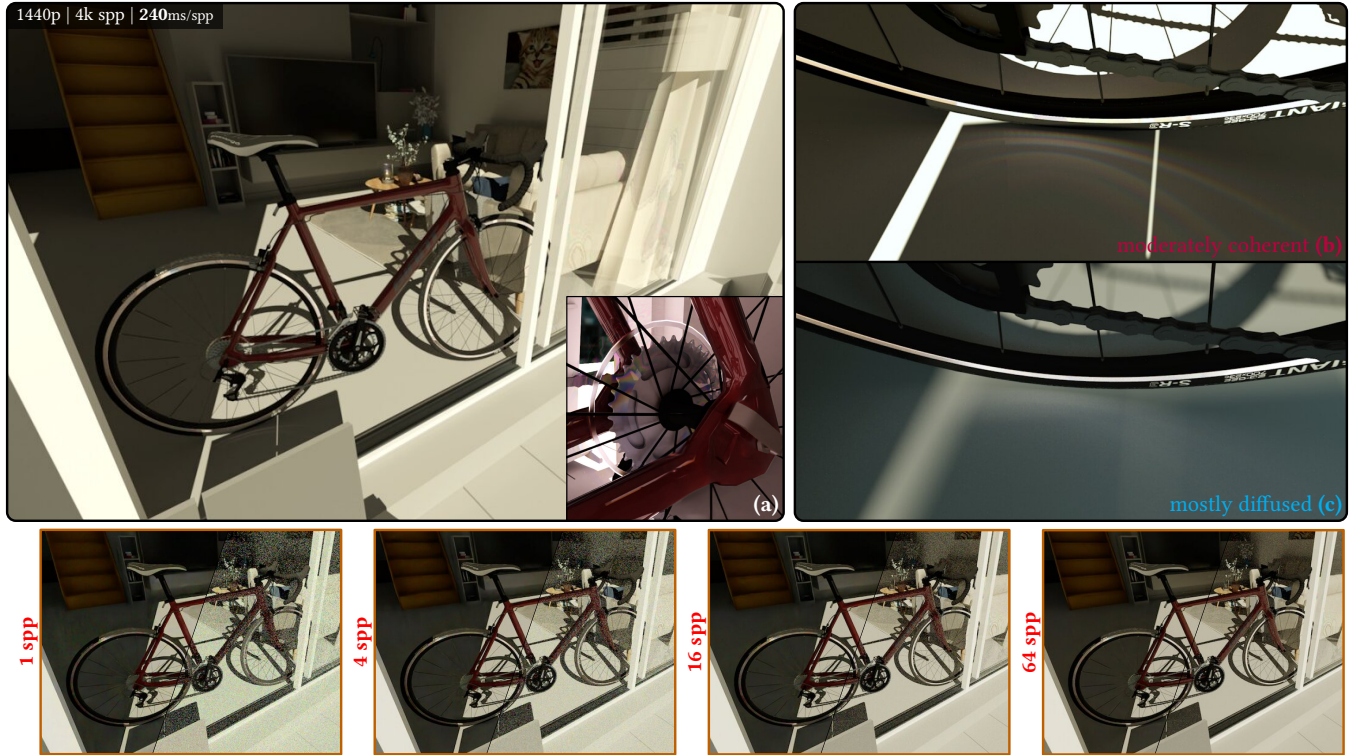


Fig. 9. **Bike scene.** The bike scene from Steinberg et al. [2022] rendered using generalized rays. The scene contains a few materials with visible wave-interference effects, most notably the birefringent dielectrics, (a) like the plastic wheel spoke guard, (b-c) as well as the diffraction grating wheel brake surface. (b) when illuminated by direct sunlight, this grating disperses light into visible diffraction lobes; however, (c) when sunlight passes through a diffuser (like clouds), only the direct lobe is visible. (bottom) Low spp renderings, with (left side) and without (right side) a denoiser, showcasing the interactive rendering performance of our wave-optical renderer.

been connected to a light source, either organically or via NEE, the solve stage kicks in to compute the partially-coherent transport over the sampled path. If the sampled path does not contain dispersive delta segments (e.g., refraction through a smooth dielectric interface), in addition to the Hero wavelength, we also importance sample 3 additional wavelengths from the light’s emission spectrum. We therefore transport one to four spectral samples per sampled path. The solve stage applies PLT in an essentially unchanged manner, see Steinberg et al. [2022] for more details.

Our implementation supports two types of light sources:

- (1) *Distant light sources* – These include analytic distant sources, with a predefined solid angle that they subtend from the scene, as well as environment maps, in which case light is sourced from a small cluster of a few pixels, defining the solid angle.
- (2) *Emissive geometry light sources* – where light is sourced from a small area on the emitting triangle.

As a design choice, our implementation ignores polarization during the sample pass (i.e., the backward transport with generalized rays). This is done for simplicity and performance: decomposing and inverting Mueller matrices, and other polarimetric interactions, can be cumbersome and expensive. No errors are introduced: the solve pass is fully vectorized.

7.3 Results

Our results are comprised of three main scenes:

- (1) **Snake enclosure** (Figs. 1 and 12). This scene is illuminated by multiple light sources: the sun, the sky (diffused sunlight), as well as a pair of industrial 4100 K fluorescent lamps with a decent colour rendering index of 82 located at the back of the enclosure. Sunlight and skylight arrive from the opening at the top. This is a difficult scene to render: most of the light that arrives at the different diffractive materials is indirect.
- (2) **Manifold sampling** (Fig. 8). A highly-detailed scene that we use as our manifold-sampling playground.
- (3) **Bike** (Fig. 9). Adapted from Steinberg et al. [2022] (appearance is not expected to match, as our materials are different).

In addition to the above, the CD scene is used for analysis of partially-coherent sampling, Fig. 10, and comparison to the state-of-the-art, Fig. 11. Fig. 12 demonstrates that wave-optics is a global process: the appearance of a material, and the observable diffractive phenomena, depend on the wave properties of light. Accurate reproduction of such effects cannot be done at the material level, but instead requires a wave-optical simulation throughout the entire scene, as we do. Our supplemental material contains additional renderings, as well as animated videos of all these scenes, showcasing the performance of our method.

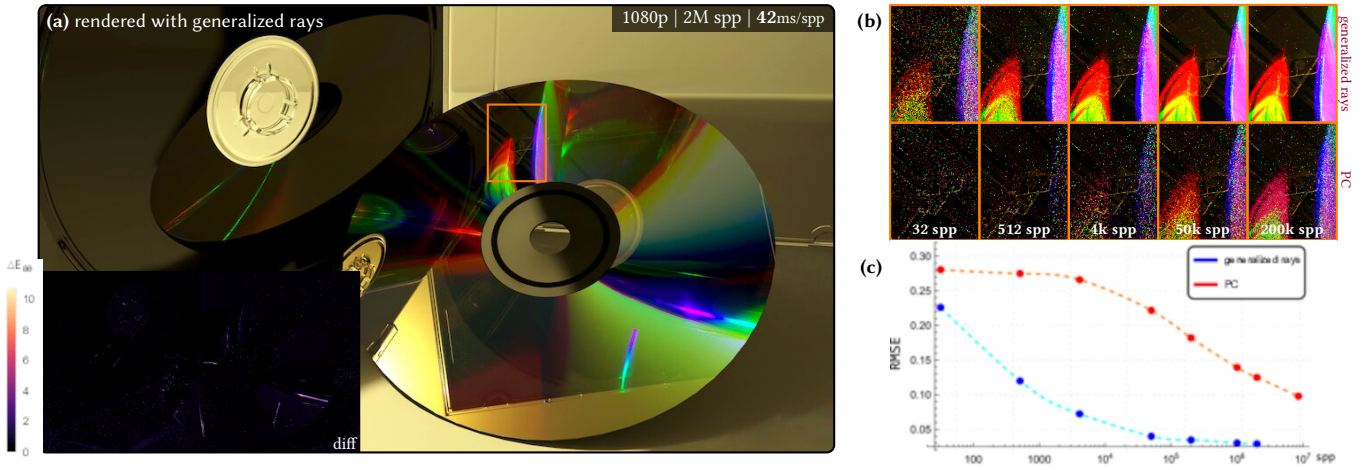


Fig. 10. **Partially-coherent sampling.** (a) A scene contains a compact disk (CD) that rests next an open CD case, upon which another closed CD case is placed. A ceiling-mounted light source illuminates the scene. While simple, the scene admits interesting light transport. (b) We emulate partially-coherent (PC) sampling of BSDFs, in an identical manner to the state-of-the-art, which manifests the sampling problem: the diffraction lobes are very sharp lobes that are difficult to sample when doing backward path tracing (and the coherence properties of light are unknown) using existing tools. The close ups show the area marked in orange rendered at various sample counts. Note the sharp difference in noise between PC sampling and our proposed light transport formalism with generalized rays. (c) Plot of noise as function of sample count, quantifying the drastic improvement in sampling performance: sampling using generalized rays reduces sample count by a factor of about 4000 for similar-quality rendering. (a, inset) A difference image between the two sampling strategies. With the exception of the very high-frequency diffraction effects, the differences are minor, and are due to the PC rendering never converging.

Performance metrics, i.e. rendering resolution and samples-per-pixel (spp) count are given in each figure, and summarised in Table 1. For the paper figures we used high-quality, converged (i.e., very high samples-per-pixel) renders. Nevertheless, our method and implementation enable interactive wave-optical rendering at 1 spp, and the frame times for interactive rendering are also given in the figures and in Table 1. Similar to other modern GPU-accelerated path tracers, we use a denoiser for interactive rendering. This enables generating acceptable images at 1 spp, and allows the user to interact with and edit the scene in real-time. The videos in our supplemental material were also rendered using a denoiser.

Low-spp images of the bike scene, with and without the denoiser, are shown in Fig. 9 (bottom). Low sample count is sufficient for the vast majority of the materials, including the diffractive birefringent dielectrics. An exception is the dispersive diffraction lobes, which are visible on the floor, and arise due to the diffraction-grated wheel brake surface. This is due to a couple of reasons: (i) A diffraction

Table 1. **Rendering performance.** Listed are the resolution and sample count used to generate the figures, as well as the interactive (1 spp) rendering frame times (at the indicated resolution). All rendering was done on a NVIDIA® GeForce RTX™ 3090 GPU.

Scene	Frame time (1 spp)	Figure	
Snake enclosure	116 ms	Figs. 1 and 12	1440p, 64k spp
MS playground	180 ms	Fig. 8	1440p, 16k spp
Bike	240 ms	Fig. 9	1440p, 4k spp
CD	42 ms	Fig. 10	1080p, 2M spp

grating scatters into many wavelength-dependent lobes, hence requires a moderate amount of spectral samples. (ii) These are rendered via manifold sampling (MS), as discussed in Fig. 8. However, MS is only initiated during the sample stage when a path is scattered from the floor into the grating. The probability of finding such a connection organically is rather low (roughly about 1 in a few dozen), as the grated surface is quite small. The difficulty of rendering the diffracted lobes of a diffraction grating is also analysed in Fig. 8 (bottom). This problem is similar to the classical problem of rendering dispersive caustics, e.g., on the bottom of a pool with a disturbed water surface; a problem where classical unidirectional path tracers struggle as well.

Comparison with the state-of-the-art. We compare our interactive wave-optical renderer to PLT [Steinberg et al. 2022]. PLT employs a bi-directional path tracer, and samples partially-coherent BSDFs. As discussed, because the coherence of light cannot be known a priori when tracing paths backwards (i.e., the sampling problem), they set a global lower limit on the coherence of light. Partially-coherent BSDFs are then sampled with respect to that limit. We emulate such partially-coherent sampling in our renderer, and compare the performance in the simple CD scene, see Fig. 10. The difference is dramatic: we observe over a thousand-fold increase in convergence performance when sampling with generalized rays compared to partially-coherent sampling.

We also compare by rendering the same CD scene directly with PLT, see Fig. 11. An increase in performance is expected, as our implementation is GPU accelerated, nevertheless, we record up to a several thousand speed gain for a similar-quality rendering (for

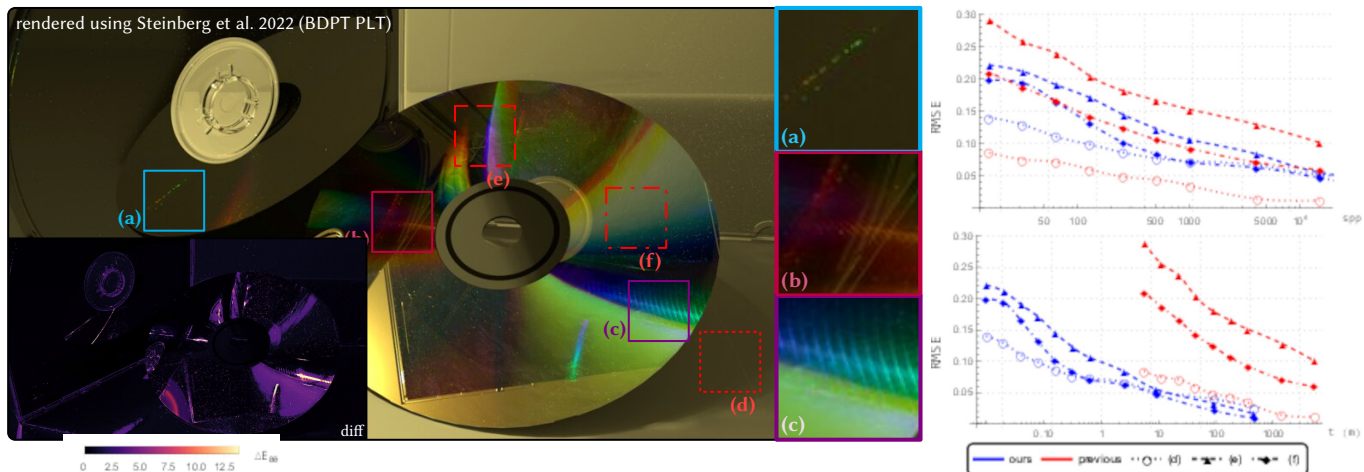


Fig. 11. **Comparison to the state-of-the-art.** We render the CD scene using the PLT bidirectional path tracer (BDPT) [Steinberg et al. 2022]. Because the illumination of the diffractive CD surface is indirect, convergence is poor, as to be expected given the analysis in Fig. 10. The image was rendered with 56 000 samples over about 315 hours, on an Intel[®] Core™ i9-10980XE 18-core CPU. (a-c) Their renderer struggles with capturing the high-frequency details of the diffraction grating. Furthermore, because they propagate a fixed number (64) of spectral samples, sampled uniformly, clear banding artefacts are reproduced, suggesting that these high-frequency details will fail to converge due to spectral aliasing. (right) To analyze and compare the convergence performance to our approach, we plot error per sample and error per rendering-time graphs for the three different regions outlined in red: (d) The light transport that arrives to this area (red dotted line) is dominated by light that is not diffracted by the CD. Therefore, the per-sample convergence performance of their BDPT renderer (which performs much more work per sample) is significantly superior, as expected. (e-f) On the other hand, these regions do sample the CD and the convergence performance of their renderer is considerably inferior. A performance improvement is expected, as our renderer is GPU accelerated. The performance at region (d), plotted with open circle markers, should then be considered as baseline: up to 100 times faster, depending on sample count. The plots pertaining to the diffractive regions (filled markers) show a vastly greater improvement of about 1000 to 10 000 times faster convergence.

image areas that admit diffractive materials). Our equal-sample convergence performance is increased by a factor of 1 to 8. This is despite the fact that their path tracer is bi-directional (ours is unidirectional) and each sample propagates 64 uniformly sampled spectral samples (compared to the up to 4 per sample with our renderer), therefore they do a considerably greater amount of work per sample. Figs. 10 and 11 highlight the sampling problem and emphasize the need for better solutions: the state-of-the-art is able to achieve decent results when diffractive materials are directly illuminated, making good use of a bi-directional path-tracer. However, it struggles greatly when it isn't trivial to connect paths from a light source to these materials.

8 CONCLUSION

The primary contribution in this work is our backward (sensor-to-source) accurate wave-optical light transport formalism. The derivations in Section 3 are exact. Our formalism is *complete*, subject to photoelectric detectors: meaning any wave-optical phenomena that may be observed by a photoelectric detector can be simulated to arbitrarily good precision, with light of optical and non-optical frequencies, and of any state of coherence or polarization. At the core of our formalism is the generalized ray: *weakly-local* and *linear* constructs that sample the distribution of light from the detector. In Fig. 3 we numerically validate this formalism, as well as delineate the highly-general process by which generalized rays are able to regain linearity in all conditions: only interfering phasors that can be

resolved over the spatial extent of the detection element, without aliasing, contribute to observable interference effects.

Some practical considerations are left for future work: (i) efficient beam tracing; and, (ii) sampling the interaction of a generalized ray with a composite interaction operator, like free-space diffraction through a slit. These considerations are especially important for rendering and light transport with longer, non-optical wavelengths, where the spatial extent of a generalized ray is significantly larger than at optical frequencies.

As a sample application, in this paper we focus on wave-optical rendering in the visible spectrum. A reasonable, application-specific simplification is made: like PLT, we assume that the spatial extent of the generalized ray is small compared to the scene's geometric details. By doing so, we have made it possible for the first time to efficiently render complex scenes under a highly-accurate wave optical model. Our novel sample-solve approach makes it possible to apply decades of work in path tracing sampling algorithms to wave optics rendering, now allowing the construction of paths from the camera and the use of advanced sampling techniques like manifold next-event estimation in the wave-optical setting. Our algorithms are highly efficient and are amenable to GPU implementation; we show interactive one sample per pixel rendering of complex scenes, on a modern GPU, at a performance that is orders-of-magnitude faster than the state-of-the-art. Our implementation is available with our supplemental material.

We stress: no wave-optical formalism may be perfectly local and linear simultaneously. We then sacrifice perfect locality in favour



Fig. 12. **Diffraction materials under different illumination conditions.** The appearance of diffractive materials depends on light's spectral, polarimetric and coherence properties. (a) During daytime, the illumination reaching the snake is dominated by indirect sunlight. At some angles, where incident light has a narrower angular spread (and thus is more coherent), the diffractive scales show clear interference patterns. (b) At night, direct light from the fluorescent lamps is the primary source of illumination, however that light is too incoherent to produce visible diffraction patterns.

of weak locality, while retaining linearity. Even in our practical, interactive rendering algorithm, where we make the simplifying assumption discussed above, we are able to accurately quantify the spatial extent over which a generalized ray interacts with matter. This is crucial for an efficient light transport simulation: without weak locality, light-matter interactions need to be integrated over the entire scene. As discussed, any forward (source-to-sensor) weakly-local, linear formalism inevitably suffers from the *sampling problem*, where weak locality cannot be established when tracing backward (sensor-to-source). Our formalism is unique in providing such machinery that is always able to regain weak locality and linearity under backward light transport.

A generalized ray can be understood as the wave-optical analogue of the classical ray: Generalized rays traverse free space, trace Eikonals through media with a slowly-varying refractive index, and reflect or refract at a smooth interface between media in an identical fashion to classical rays. A generalized ray is the most compact construct permissible under wave optics.

REFERENCES

- Birk Andreas, Giovanni Mana, and Carlo Palmisano. 2015. Vectorial ray-based diffraction integral. *Journal of the Optical Society of America A* 32, 8 (July 2015), 1403. <https://doi.org/10.1364/josaa.32.001403>
- Chen Bar, Marina Alterman, Ioannis Gkioulekas, and Anat Levin. 2019. A Monte Carlo framework for rendering speckle statistics in scattering media. *ACM Transactions on Graphics* 38, 4 (Jul 2019), 1–22. <https://doi.org/10.1145/3306346.3322950>
- Chen Bar, Ioannis Gkioulekas, and Anat Levin. 2020. Rendering near-field speckle statistics in scattering media. *ACM Transactions on Graphics* 39, 6 (Nov 2020), 1–18. <https://doi.org/10.1145/3414685.3417813>
- Laurent Belcour and Pascal Barla. 2017. A Practical Extension to Microfacet Theory for the Modeling of Varying Iridescence. *ACM Trans. Graph.* 36, 4, Article 65 (July 2017), 14 pages. <https://doi.org/10.1145/3072959.3073620>
- Ahmad Bilal, Syed Muhammad Hamza, Ziauddin Taj, and Shuaib Salamati. 2019. Comparison of SBR and MLFMM techniques for the computation of RCS of a fighter aircraft. *IET Radar, Sonar & Navigation* 13, 10 (2019), 1805–1810.
- D Bilibashi, EM Vitucci, and V Degli-Esposti. 2020. Dynamic ray tracing: Introduction and concept. In *2020 14th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 1–5.
- David Blinder, Maksymilian Chlipala, Tomasz Kozacki, and Peter Schelkens. 2021. Photorealistic computer generated holography with global illumination and path tracing. *Opt. Lett.* 46, 9 (May 2021), 2188–2191. <https://doi.org/10.1364/OL.422159>
- Mate Boban, Joao Barros, and Ozan K Tonguz. 2014. Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation. *IEEE Transactions on Vehicular Technology* 63, 9 (2014), 4146–4164.
- Max Born and Emil Wolf. 1999. *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, Cambridge New York.
- John Burgess. 2020. RTX on—The NVIDIA Turing GPU. *IEEE Micro* 40, 2 (March 2020), 36–44.
- Juan D. Castro, Sahitya Singh, Akshaj Arora, Sara Louie, and Damir Senic. 2019. Enabling Safe Autonomous Vehicles by Advanced mm-Wave Radar Simulations. In *2019 IEEE MTT-S International Microwave Symposium (IMS)*. IEEE. <https://doi.org/10.1109/mwsym.2019.8700843>
- Hyuckjin Choi, Jaeky Oh, Jaehoon Chung, George C Alexandropoulos, and Junil Choi. 2023. WiThRay: A Versatile Ray-Tracing Simulator for Smart Wireless Environments. *IEEE Access* (2023).
- Tom Cuypers, Tom Haber, Philippe Bekaert, Se Baek Oh, and Ramesh Raskar. 2012. Reflectance model for diffraction. *ACM Transactions on Graphics* 31, 5 (Aug 2012), 1–11. <https://doi.org/10.1145/2231816.2231820>
- FS de Adana, O Gutierrez, I Gonzalez, and MF Catedra. 2005. FASPRO: Fast computer tool for the analysis of propagation in mobile communications. In *INDIN'05. 2005 3rd IEEE International Conference on Industrial Informatics, 2005*. IEEE, 257–261.
- Daniela Dragoman. 2004. Phase space correspondence between classical optics and quantum mechanics. In *Progress In Optics*. Vol. 42. 424–486.
- Daniela Dragoman. 2005. Phase space formulation of quantum mechanics. Insight into the measurement problem. *Physica Scripta* 72, 4 (2005), 290.
- V. Falster, A. Jarabo, and J. R. Frisvad. 2020. Computing the Bidirectional Scattering of a Microstructure Using Scalar Diffraction Theory and Path Tracing. *Computer Graphics Forum* 39, 7 (Oct 2020), 231–242. <https://doi.org/10.1111/cgf.14140>
- Tiantian Feng and Lixin Guo. 2021. Multiview ISAR Imaging for Complex Targets Based on Improved SBR Scattering Model. (2021).
- Pengcheng Gao, Xiaobing Wang, Zichang Liang, and Wei Gao. 2015. Efficient GPU implementation of SBR for fast computation of composite scattering from electrically large target over a randomly rough surface. In *2015 IEEE International Symposium on Antennas and Propagation ; USNC/URSI National Radio Science Meeting*. IEEE. <https://doi.org/10.1109/aps.2015.7305222>
- Ion I. Geru. 2018. *Time-Reversal Symmetry*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-01210-6>
- Junfeng Guan, Sohrab Madani, Suraj Jog, Saurabh Gupta, and Haitham Hassanieh. 2020. Through fog high-resolution imaging using millimeter wave radar. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11464–11473.
- Ibón Guillén, Julio Marco, Diego Gutierrez, Wenzel Jakob, and Adrian Jarabo. 2020. A General Framework for Pearlescent Materials. *ACM Transactions on Graphics* 39, 6 (2020). <https://doi.org/10.1145/3414685.3417782>

- Johannes Hanika, Marc Droske, and Luca Fascione. 2015. Manifold next event estimation. *Comput. Graph. Forum* 34, 4 (July 2015), 87–97.
- Nicolas Holzschuch and Romain Pacanowski. 2017. A Two-scale Microfacet Reflectance Model Combining Reflection and Diffraction. *ACM Trans. Graph.* 36, 4, Article 66 (July 2017), 12 pages. <https://doi.org/10.1145/3072959.3073621>
- Ferdous Hossain, Tan Kim Geok, Tharek Abd Rahman, Mhd Nour Hindia, Kaharudin Dimiyati, and Azlan Abdaziz. 2018. Indoor millimeter-wave propagation prediction by measurement and ray tracing simulation at 38 GHz. *Symmetry* 10, 10 (2018), 464.
- Weizhen Huang, Julian Iseringhausen, Tom Kneiphof, Ziyin Qu, Chenfanfu Jiang, and Matthias B. Hullin. 2020. Chemomechanical simulation of soap film flow on spherical bubbles. *ACM Transactions on Graphics* 39, 4 (Jul 2020). <https://doi.org/10.1145/3386569.3392094>
- KL Jensen and FA Buot. 1991. The methodology of simulating particle trajectories through tunneling structures using a Wigner distribution approach. *IEEE Transactions on Electron Devices* 38, 10 (1991), 2337–2347.
- Jian-Ming Jin. 2015. *The finite element method in electromagnetics*. John Wiley & Sons.
- K.-S. Jin, T.-I. Suh, S.-H. Suk, B.-C. Kim, and H.-T. Kim. 2006. Fast Ray Tracing Using A Space-Division Algorithm for RCS Prediction. *Journal of Electromagnetic Waves and Applications* 20, 1 (Jan. 2006), 119–126. <https://doi.org/10.1163/15693930677577341>
- Cam Key, Blake Troksa, Forest Kunkel, Slobodan V. Savic, Milan M. Ilic, and Branislav M. Notaros. 2018. Comparison of Three Sampling Methods for Shooting-Bouncing Ray Tracing Using a simple Waveguide Model. In *2018 IEEE International Symposium on Antennas and Propagation ; USNC/URSI National Radio Science Meeting*. IEEE. <https://doi.org/10.1109/apusncursinrm.2018.8609320>
- Tom Kneiphof, Tim Golla, and Reinhard Klein. 2019. Real-time Image-based Lighting of Microfacet BRDFs with Varying Iridescence. *Computer Graphics Forum* 38, 4 (2019), 77–85. <https://doi.org/10.1111/cgf.13772>
- Andrey Krywonos. 2006. *Predicting surface scatter using a linear systems formulation of non-paraxial scalar diffraction*. Ph.D. Dissertation. University of Central Florida.
- Ulf Leonhardt. 1997. *Measuring the quantum state of light*. Cambridge University Press, Cambridge, England.
- Jacob Mackay and David Johnson. 2021. Millimetre wave ray tracing simulator with phase and beam effects using the Wigner distribution function. In *Passive and Active Millimeter-Wave Imaging XXIV*, Vol. 11745. SPIE, 43–57.
- Juan Magallón, Alfonso Blesa, and Francisco Serón. 2021. SLM simulation and MonteCarlo path tracing for computer-generated holograms. *SN Computer Science* 2 (2021), 1–16.
- Leonard Mandel and Emil Wolf. 1995. *Optical coherence and quantum optics*. Cambridge University Press, Cambridge.
- Marco Mout, Michael Wick, Florian Bociort, Joerg Petschulat, and Paul Urbach. 2018. Ray tracing the Wigner distribution function for optical simulations. *Optical Engineering* 57, 1 (2018), 014106–014106.
- ZY Ou and HJ Kimble. 1995. Probability distribution of photoelectric currents in photodetection processes and its connection to the measurement of a quantum state. *Physical Review A* 52, 4 (1995), 3126.
- Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2016. *Physically based rendering: From theory to implementation*. Morgan Kaufmann.
- Walter Rudin. 1990. *Functional Analysis* (2 ed.). McGraw Hill Higher Education, Maidenhead, England.
- Hae-Won Son and Noh-Hoon Myung. 1999. A deterministic ray tube method for microcellular wave propagation prediction model. *IEEE Transactions on Antennas and Propagation* 47, 8 (1999), 1344–1350.
- Francisco Soto and Pierre Claverie. 1983. When is the Wigner function of multidimensional systems nonnegative? *J. Math. Phys.* 24, 1 (1983), 97–100.
- Jos Stam. 1999. Diffraction shaders. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99*. ACM Press. <https://doi.org/10.1145/311535.311546>
- Shlomi Steinberg, Pradeep Sen, and Ling-Qi Yan. 2022. Towards Practical Physical-Optics Rendering. *ACM Transactions on Graphics* 41, 4 (Jul 2022), 1–13. <https://doi.org/10.1145/3528223.3530119>
- Shlomi Steinberg and Ling-Qi Yan. 2021. A Generic Framework for Physical Light Transport. *ACM Transactions on Graphics* 40, 4 (Aug 2021), 1–20. <https://doi.org/10.1145/3450626.3459791>
- Shlomi Steinberg and Ling-Qi Yan. 2022. Rendering of Subjective Speckle Formed by Rough Statistical Surfaces. *ACM Trans. Graph.* 41, 1, Article 2 (feb 2022), 23 pages. <https://doi.org/10.1145/3472293>
- Sungha Suk, Tae-Il Seo, Hae-Sung Park, and Hyo-Tae Kim. 2001. Multiresolution grid algorithm in the SBR and its application to the RCS calculation. *Microwave and Optical Technology Letters* 29, 6 (2001), 394–397. <https://doi.org/10.1002/mop.1188>
- Yubo Tao, Hai Lin, and Hujun Bao. 2010. GPU-Based Shooting and Bouncing Ray Method for Fast RCS Prediction. *IEEE Transactions on Antennas and Propagation* 58, 2 (Feb. 2010), 494–502. <https://doi.org/10.1109/tap.2009.2037694>
- Yu-Bo Tao, Hai Lin, and Hu Jun Bao. 2008. Kd-tree based fast ray tracing for rcs prediction. *Electromagn. Waves (Camb.)* 81 (2008), 329–341. <https://doi.org/10.2528/pier08011305>
- Markus Testorf, Bryan Hennelly, and Jorge Ojeda-Castañeda. 2010. *Phase-space optics: fundamentals and applications*. McGraw-Hill Education.
- Antoine Toisoul and Abhijeet Ghosh. 2017. Practical Acquisition and Rendering of Diffraction Effects in Surface Reflectance. *ACM Transactions on Graphics* 36, 5 (Jul 2017), 1–16. <https://doi.org/10.1145/3012001>
- Amalia Torre. 2005. *Linear ray and wave optics in phase space: bridging ray and wave optics via the Wigner phase-space picture*. Elsevier.
- Z. Velinov, S. Werner, and M. B. Hullin. 2018. Real-Time Rendering of Wave-Optical Effects on Scratched Surfaces. *Computer Graphics Forum* 37, 2 (2018), 123–134. <https://doi.org/10.1111/cgf.13347>
- A Walther. 1968. Radiometry and coherence. *JOSA* 58, 9 (1968), 1256–1259. <https://doi.org/10.1364/JOSA.58.001256>
- Craig Warren, Antonios Giannopoulos, and Iraklis Giannakis. 2016. gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. 209 (Dec 2016).
- Sebastian Werner, Zdravko Velinov, Wenzel Jakob, and Matthias Hullin. 2017. Scratch Iridescence: Wave-Optical Rendering of Diffractive Surface Structure. *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 36, 6 (Nov. 2017). <https://doi.org/10.1145/3130800.3130840>
- E. Wigner. 1932. On the Quantum Correction For Thermodynamic Equilibrium. *Phys. Rev.* 40 (Jun 1932), 749–759. Issue 5. <https://doi.org/10.1103/PhysRev.40.749>
- A Wilkie, S Nawaz, M Droske, A Weidlich, and J Hanika. 2014. Hero wavelength spectral sampling. *Comput. Graph. Forum* 33, 4 (July 2014), 123–131.
- Emil Wolf. 2007. *Introduction to the theory of coherence and polarization of light*. Cambridge University Press, Cambridge.
- Kane Yee. 1966. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on antennas and propagation* 14, 3 (1966), 302–307.
- Haofan Yi, Danping He, P Takis Mathiopoulos, Bo Ai, Juan Moreno Garcia-Loygorri, Jianwu Dou, and Zhangdui Zhong. 2022. Ray tracing meets terahertz: Challenges and opportunities. *IEEE Communications Magazine* (2022).
- Yunchen Yu, Mengqi Xia, Bruce Walter, Eric Michielssen, and Steve Marschner. 2023. A Full-Wave Reference Simulator for Computing Surface Reflectance. , 17 pages. <https://doi.org/10.1145/3592414>
- Tizian Zeltner, Iliyan Georgiev, and Wenzel Jakob. 2020. Specular manifold sampling for rendering high-frequency caustics and glints. *ACM Trans. Graph.* 39, 4 (Aug. 2020).