A Free-Space Diffraction BSDF

SHLOMI STEINBERG, University of Waterloo, Canada
RAVI RAMAMOORTHI*, NVIDIA, United States
BENEDIKT BITTERLI, NVIDIA, United States
ARSHIYA MOLLAZAINALI, University of Waterloo, Canada
EUGENE D'EON, NVIDIA, New Zealand
MATT PHARR, NVIDIA, United States

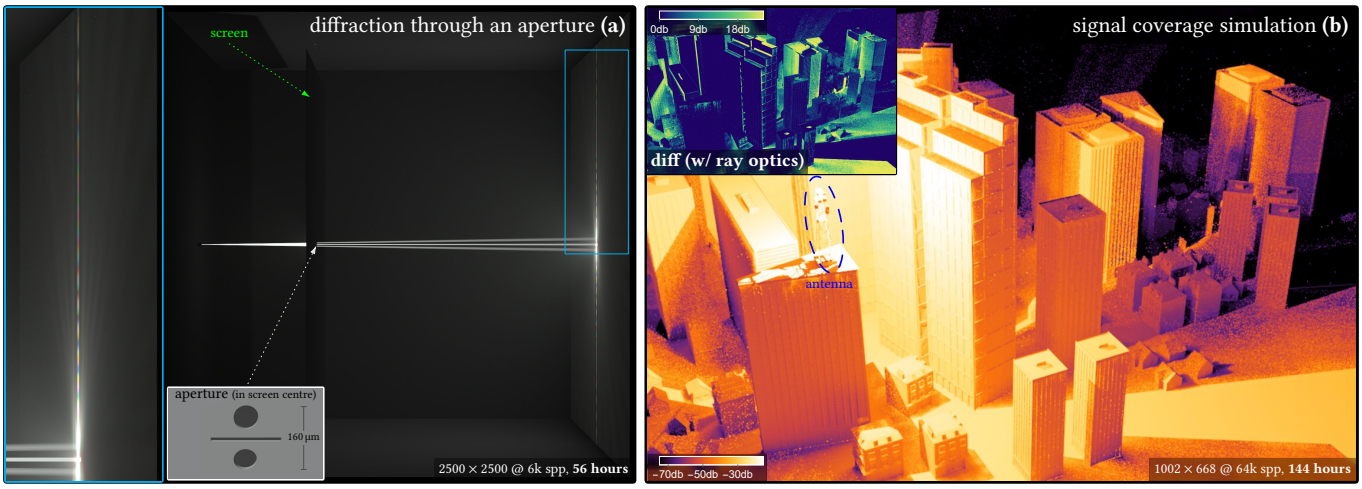


Fig. 1. **Rendering of free-space diffractions.** (a) A 1 cm^3 Cornell box is filled with a homogeneous participating medium and illuminated by a collimated beam that enters from the centre of the left wall. The beam consists of several optical wavelengths: blue ($\lambda = 475$ nm), green ($\lambda = 535$ nm) and red ($\lambda = 620$ nm). A conductive screen is placed in the beam's path with an aperture cut in the centre of screen (geometry shown in inset). Scene geometry, including the screen and its aperture, are simple triangular meshes: no other information about the aperture is fed to the renderer, and all geometry in the scene dynamically diffracts light. When light rays impinge upon the screen, our method looks at the geometry around the interaction point, and on-the-fly constructs the *free-space diffraction BSDF*. This BSDF accurately quantifies the wave-optical effect of light *diffracting* through the aperture and around geometry—i.e. light "bending" around the screen into the region shadowed by the screen—giving rise to the colourful diffraction pattern seen on the right wall. (b) The irradiance (power per surface area) of long-wavelength cellular radiation ($\lambda = 10$ cm) impinging upon the visible surfaces is colour-coded, demonstrating the ability of our method to handle complex, real-life scenes with fine geometrical details. Radiation diffracts around the buildings, simulating wave-optical signal coverage that cannot be correctly predicted by ray optics. See Fig. 7 for additional information. All rendering is done with a typical bi-directional path tracer.

Free-space diffractions are an optical phenomenon where light appears to "bend" around the geometric edges and corners of scene objects. In this paper we present an efficient method to simulate such effects. We derive an edge-based formulation of Fraunhofer diffraction, which is well suited to the common (triangular) geometric meshes used in computer graphics. Our method dynamically constructs a free-space diffraction BSDF by considering the geometry around the intersection point of a ray of light with an object, and we present an importance sampling strategy for these BSDFs. Our method is unique in requiring only ray tracing to produce free-space diffractions, works with general meshes, requires no geometry preprocessing, and is designed to work with path tracers with a linear rendering equation. We show that we are able to reproduce accurate diffraction lobes, and, in contrast to any existing method, are able to handle complex, real-world geometry. This work serves to connect free-space diffractions to the efficient path tracing tools from computer graphics.

CCS Concepts: • **Computing methodologies** → **Rendering**; **Scientific visualization**; **Computer graphics**; • **Applied computing** → *Physics*.

1 INTRODUCTION

The simulation of light transport arises in a multitude of applications and targets different regions of the electromagnetic spectrum. An important phenomenon that is challenging to capture with modern rendering frameworks is the wave-optical effect of light "bending" around geometry: when electromagnetic light waves impinge upon an obstacle, they *diffract* and reform, with some of the energy

*Also with University of California San Diego.

Authors' addresses: Shlomi Steinberg, p@shlomisteinberg.com, University of Waterloo, Waterloo, Canada; Ravi Ramamoorthi, ravir@cs.ucsd.edu, NVIDIA, San Francisco, United States; Benedikt Bitterli, benedikt.bitterli@gmail.com, NVIDIA, San Francisco, United States; Arshiya Mollazainali, arshiya.mollazainali@uwaterloo.ca, University of Waterloo, Waterloo, Canada; Eugene d'Eon, ejdeon@gmail.com, NVIDIA, Wellington, New Zealand; Matt Pharr, matt@pharr.org, NVIDIA, San Francisco, United States.

scattered into regions shadowed by the obstacle. We refer to such effects as *free-space diffractions*. See Fig. 1.

Free-space diffractions arise with light of any wavelength. However, in the visible regime, only exceptional settings produce meaningful observable effects of free-space diffraction (e.g. coherent laser illuminating one or more very thin slits). The importance of these effects rises dramatically with longer wavelengths: RADAR radiation diffracts around pedestrians and cyclists (thereby masking them from the receiver); WiFi and cellular radiation (e.g., in signal coverage simulations in urban environments) diffracts around indoor objects, and by an edge of a building. At these frequencies, free-space diffractions are the predominant phenomena that govern the distribution of energy in a scene. These effects allow light to bend into the "shadow" region (in the classical sense) of objects, in a manner that is contradictory to classical ray optics.

Practical, efficient simulation of long-wavelength light transport in complex, real-world scenes remains an open problem and a very active area of research. As free-space diffraction effects play a paramount role in these long-wavelength regimes, practical numeric methods that are able to simulate free-space diffractions have been extensively studied. The most important and by-far the most widely used ones are the *geometric theory of diffraction* (GTD) and its extension the *uniform theory of diffraction* (UTD). Due to their similarity we refer to both simply as "UTD". UTD is used to compute the propagation of electromagnetic energy over or through obstacles. It is an asymptotic method: ray-optical paths (rays) are used for propagation, and edges are assumed to be very long compared to wavelength. UTD quantifies how these rays diffract by geometric edges and corners, with different geometries requiring distinct diffraction formulae.

At a first glance, UTD is attractive because propagation is done via simple ray tracing, thereby enabling light transport simulations in scenes that are far more complex than anything a traditional wave solver would be able to handle. However, these rays carry electromagnetic phase and interfere with each other. This interference is mandatory to reproduce diffraction phenomena. However interference marks a profound departure from the common ray optical setting: *rays do not superpose linearly*, therefore a linear rendering equation (that governs the recursive light transport simulation) cannot be formulated.

As we will show (Fig. 5), even in the simplest of settings, UTD requires a very large number of samples to yield converged results: (i) diffracting edges need to be sampled at a sub-wavelength frequency, in order to correctly capture edge geometry; and (ii) sub-wavelength sampling is needed at the target (where we observe the diffraction pattern, or compute other interactions). Because interference between the rapidly-oscillating rays shapes the interference patterns, when insufficient samples are used, high-frequency, high-intensity noise arises. In addition, mutually-interfering rays induce scene-wide noise that produces no meaningful interference patterns: rays that arrive to the same spot from far-apart regions in the scene give rise to what can be understood as *optical speckle*: the superposition of effectively-random phasors. In other words, our integrand is now a rapidly-oscillating complex-valued function, making numeric integration more difficult. See Section S4 in our supplemental material for a discussion of UTD and its shortcomings.

The difficulties mentioned above make UTD, and other methods with mutually-interfering rays, **a poor fit for use with traditional light transport tools**, specifically *path tracing* and the related techniques developed by the computer graphics community. Other methods, like explicit interference integrals, are also not a good fit: only some specific aperture geometries yield a closed-form solution for such integrals. As a consequence, the state-of-the-art that simulates light transport with free-space diffractions, often using a UTD-based approach, only works on either 2D scenes or 3D scenes that would be considered exceedingly simple by the computer graphics community. These methods only consider very shallow paths, usually one or at most two interactions with the scene, and often require highly simplified geometry (e.g., a single quad for the entire face of a building).

The poor scalability of existing methods is a problem. Real-world scenes with complex geometries, where deeper paths contribute non-negligibly, are of real practical interest for the applications listed above: e.g., cellular radiation interacts with the scene many times while propagating through a city. Path tracing techniques, with a linear rendering equation, have been extensively researched and are responsible for the vast majority of the computer-generated content in multimedia and films; these techniques are able to handle the complex, real-world scenes where we would like to be able to simulate free-space diffractions.

The purpose of this paper is to introduce a novel free-space diffraction BSDF (*bi-directional scattering distribution function*; quantifying the angular scattering distribution), that is designed to operate with a typical path tracer, where propagation continues to be done by simple ray tracing. Our BSDF requires essentially no changes to the internal workings of the path tracer. Conceptually, our method can be summarised as follows:

- Propagation continues as usual, via ray tracing.
- When a ray is incident upon geometry, we look at the geometry around the impact point and identify relevant edges.
- If no edges that induce free-space diffraction are detected (e.g., the ray is incident upon a wall with no openings), then path tracing continues unaltered.
- Otherwise, from this set of edges, a BSDF is constructed that quantifies how light propagates around the geometry, as well as which portion of the incident energy undergoes free-space diffraction. That BSDF is easy to evaluate and can be importance sampled.

To allow the use of (linear) ray tracing only, a few approximations are made. Most notably, we only diffract rays that impinge upon scene geometry; rays that pass close to obstacles continue unimpeded, even though a fraction of their energy should also undergo diffraction. We compensate by having rays that *do* diffract also account for the energy of rays that should have diffracted, but did not (see Section 4). As we will show, this reproduces the diffraction lobes accurately.

The guiding optical principle for our formulation of free-space diffraction is the well-known theory of *Fraunhofer diffraction*, summarised in Subsection 3.1. This choice is motivated by our ability to derive a novel closed-form expression for Fraunhofer edge diffraction, under illumination by a spatially-modulated incident

beam. See Fig. 3 for an overview of our method. The benefits of our method over the state-of-the-art are:

(1) **Compatibility with linear path tracers**: We formulate a classical BSDF, that operates on radiometric quantities. No interference between rays is required, and our BSDF is easy to implement in a modern path tracer.
(2) **Performance**: Our method calculates the aggregated angular scatter that is induced by the *entire* aperture, and is devoid of the high-frequency noise that is prevalent with interfering rays and non-linear methods. In contrast, methods that depend on mutually-interfering, phase-carrying rays require dense sampling—multiple samples per wavelength.
(3) **Formulation as an angular scattering function (BSDF)**: Our formulations produce a classical BSDF that only depends on the incident and scattering directions. UTD, and similar methods, require knowledge of the distance of propagation of each ray from the edge to the destination, which is unknown, a priori, in a typical path tracing simulation.
(4) **Ease of use**: Our method is designed to work with (non-degenerate, non self-intersecting) triangular meshes, which are widespread in computer graphics. No special mesh preprocessing or treatment is required. We do not require extending the geometric silhouettes (in order to capture rays that do not intersect objects), as for example done by Xia et al. [2020] for fibre rendering. On the other hand, UTD diffraction formulae depend on an edge's underlying geometry, like wedge opening angle, corner topology, and so on.

2 RELATED WORK

GTD/UTD-based ray tracing. Methods based on the Geometric/uniform theory of diffraction (UTD) [Bilibashi et al. 2020; Paknys 2016; Son and Myung 1999; Yi et al. 2022] remain the state-of-the-art in the simulation of long-wavelength radiation, with applications in, for example, automotive-targeted simulation of RADAR [Boban et al. 2014; Guan et al. 2020], or simulation of WiFi/cellular radiation (e.g., for analysis of signal coverage in an urban environment) [Celaya-Echarri et al. 2020; Choi et al. 2023; de Adana et al. 2005]. Due to their importance, these applications have garnered significant attention, and better solutions would be of real interest. Only select works are cited.

Significant effort has been made in order to accelerate such ray tracing approaches: e.g., by 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]; or, acceleration on GPUs [Gao et al. 2015; Tao et al. 2010]. Unsurprisingly, this line of research is similar to early computer graphics research that aimed to accelerate ray tracing. Nevertheless, UTD-based methods, and similar methods, are unable to match common computer graphics techniques—like path tracing—in their computational efficiency. This is because UTD requires mutually-interfering rays that carry high-frequency phase, prohibiting the formulation of a linear (in terms of intensity or radiance) rendering equation. Hence, scene complexity remains a very limiting factor: all works listed above apply UTD in a two-dimensional setting, or a very simple three-dimensional scene with a single diffraction (no deep paths).

Other difficulties arise: (i) UTD cannot be formulated as an angular scattering function (e.g., a BSDF), as UTD formulae depend on propagation distance after interaction (frustrating importance sampling); (ii) edges scatter a ray into a one-dimensional set of directions—the *Keller cone*; and (iii) UTD requires dedicated formulae for diffraction by different objects and topologies: wedges, spheres, cylinders, corners, and so on. The above makes UTD-based approaches difficult to apply with path tracing techniques. In contrast, our method is formulated as a BSDF, quantifies continuous scattering over the hemisphere (enabling next-event estimation), and is designed to operate on a typical triangular mesh.

Acoustics. UTD-based and Fraunhofer-based diffraction has also been used for non-electromagnetic waves, specifically sound [Pisha et al. 2020; Schissler et al. 2021; Tsingos et al. 2001]. Our formulation of an edge-based Fraunhofer diffraction BSDF could be applied to such acoustics simulations as well with minor modifications, however acoustics are beyond the scope of this paper, and are left for future work.

Fraunhofer diffraction of polygonal apertures. Formulating Fraunhofer diffraction as a superposition of edge-based diffractions of a polygonal aperture (usually via the divergence theorem), as we do in this paper, has been done before [Huang et al. 2006; Komrska 1982]. However, existing work assumes an incident plane wave. This presents a problem: an unmodulated plane wave exists throughout the entire space, but we would like to limit the diffracting aperture to a small region around the intersection point of a ray with geometry. The naïve approach—clamping the plane wave to a finite region—would not work: it produces sharp boundaries, i.e. fake diffractive edges. The Fraunhofer diffraction (i.e., the Fourier transform) of such discontinuities gives rise to hallucinated diffraction lobes. We derive novel Fraunhofer edge-diffracted formulae where the incident field falls off smoothly away from its mean.

Additional related work. Some methods point-sample an input field and propagate a ray tube in order to approximate a diffraction integral [Andreas et al. 2015; Mout et al. 2018]. These methods rely on mutually-interfering rays, hence suffer from similar deficiencies to UTD. In computer graphics, phase-carrying mutually-interfering rays have been used in limited settings: for the rendering of rainbows [Sadeghi et al. 2012], and for appearance modelling of layered interference effects in various materials [Belcour and Barla 2017; Huang et al. 2020; Steinberg 2019; Wu and Zheng 2016].

Steinberg et al. [2023, 2022] formalise wave-optical light transport frameworks where light is decomposed into beams, and the spatial extent of these beams—the region over which they are able to coherently interfere—is quantified. These frameworks do not develop the machinery to construct free-space diffraction BSDFs. The presented paper is meant to fill that gap, and is designed to operate in tandem with these frameworks: we provide the machinery to efficiently construct and importance sample free-space diffractions BSDFs, which diffract these light beams. The search region for the diffracting geometry is the spatial extent of these beams. However, to keep our implementation simple, we do not use such frameworks, instead the region over which our method searches for diffracting geometry is set ad hoc to a constant.

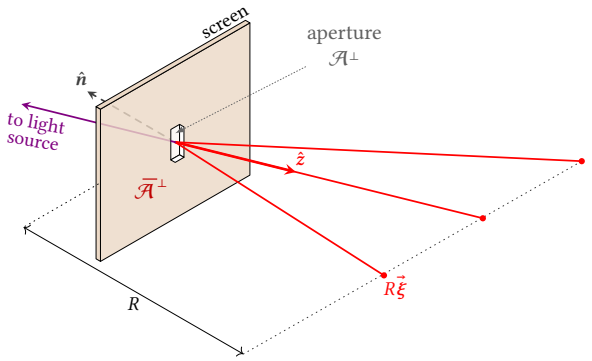


Fig. 2. **Fraunhofer diffraction.** Light is incident upon a thin, conductive screen. The screen admits one or more openings—an *aperture*—defined by the set of two-dimensional points upon the screen plane denoted $\mathcal{A}^\perp$. The complement of this set, denoted $\overline{\mathcal{A}}^\perp$, is the screen. Light passes through the aperture and *diffracts*, thereby bending into regions that are potentially shadowed by the screen, with three directions of diffraction illustrated with red lines. The normal of the screen plane is $\hat{\boldsymbol{n}}$, the direction of incidence is $-\hat{\boldsymbol{z}}$, and the direction of light that passes through the aperture without undergoing diffraction (the *direct term*) is $\hat{\boldsymbol{z}}$. R is the z distance between the screen and the observation points.

3 BACKGROUND

3.1 Fraunhofer Diffraction

Consider light that is incident upon a two-dimensional aperture (a geometric opening in a conductive screen). In the Fraunhofer region (the optical far-field region), diffraction can be formulated as a Fourier transform [Born and Wolf 1999]:

$$\psi\left(\vec{\xi}\right) \triangleq \frac{k}{2\pi R}\int_{\mathcal{A}^\perp} \mathrm{d}^2\vec{\boldsymbol{q}}^\perp\, \varphi(\vec{\boldsymbol{q}}^\perp)\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{\boldsymbol{q}}^\perp} \tag{1}$$

(we discard an irrelevant phase term), where $k = \frac{2\pi}{\lambda}$ is the *wavenumber*, with λ being the wavelength, R is the distance between the aperture and the plane of observation, $\mathcal{A}^\perp \subset \mathbb{R}^2$ is the aperture in the screen, φ is the complex amplitude field of the light incident upon the screen, ψ is the complex amplitude of the diffracted field, and $R\vec{\xi}$ are points on the observation plane. See Fig. 2.

Regarding the notation: We use the projection symbol $^\perp$ to make explicit the fact that the aperture, as well as the points $\vec{\boldsymbol{q}}^\perp \in \mathcal{A}^\perp$ upon this aperture, are two-dimensional constructs. Later, we will indeed project scene geometry upon a virtual, planar screen, giving rise to a two-dimensional projected aperture.

The Fraunhofer (far-field) approximation implies that R is large compared to the characteristic length of the diffracting aperture $\mathcal{A}^\perp \subset \mathbb{R}^2$. This is the typical setting of interest: we are interested in how diffracted energy continues to interact with the scene, and rarely observe it very close to the diffracting aperture. Furthermore, very large apertures do not produce observable diffraction patterns (as they are much larger than light's wavelength). Analytic bounds on the accuracy of Fraunhofer and Fresnel diffraction were studied by Mezouari and Harvey [2003]. We note that the distance R will drop out as we carry out our derivations, meaning that knowledge of the distance to the point of observation is not needed in practice (the reader may assume $R = 1$ for the rest of the paper).

We denote the total power contained in the incident beam of light φ that impinges upon the screen and upon the aperture as

$$\mathcal{P}_{\text{in}} \triangleq \int_{\mathbb{R}^2} \mathrm{d}^2\vec{\boldsymbol{q}}^\perp \left|\varphi(\vec{\boldsymbol{q}}^\perp)\right|^2, \tag{2}$$

$$\mathcal{P}_{\mathcal{A}^\perp} \triangleq \int_{\mathcal{A}^\perp} \mathrm{d}^2\vec{\boldsymbol{q}}^\perp \left|\varphi(\vec{\boldsymbol{q}}^\perp)\right|^2, \tag{3}$$

respectively. Clearly $\mathcal{P}_{\mathcal{A}^\perp} \leq \mathcal{P}_{\text{in}}$. By energy conservation considerations, the total diffracted power contained in the field must equal the total power incident upon the aperture: $\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2|\psi(\vec{\xi})|^2 = \mathcal{P}_{\mathcal{A}^\perp}$, which follows from Eqs. (1) and (3) and Parseval's theorem. The R^2 term is due to integrating over the observation plane, see Subsection S1.1 in our supplemental for more information.

3.2 Babinet's Principle

Let the complement of the two-dimensional aperture $\mathcal{A}^\perp$ be denoted as $\overline{\mathcal{A}}^\perp$, such that:

$$\mathcal{A}^\perp \cup \overline{\mathcal{A}}^\perp = \mathbb{R}^2 \qquad \text{and} \qquad \mathcal{A}^\perp \cap \overline{\mathcal{A}}^\perp = \varnothing, \tag{4}$$

where $\varnothing$ is the empty set. We also denote the fields diffracted by the complement aperture and the entire screen as

$$\bar{\psi}\left(\vec{\xi}\right) \triangleq \frac{k}{2\pi R}\int_{\overline{\mathcal{A}}^\perp} \mathrm{d}^2\vec{\boldsymbol{q}}^\perp\, \varphi(\vec{\boldsymbol{q}}^\perp)\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{\boldsymbol{q}}^\perp} \quad \text{and} \tag{5}$$

$$\widetilde{\varphi}\left(\vec{\xi}\right) \triangleq \frac{k}{2\pi R}\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\boldsymbol{q}}^\perp\, \varphi(\vec{\boldsymbol{q}}^\perp)\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{\boldsymbol{q}}^\perp}, \tag{6}$$

respectively. $\widetilde{\varphi}$ is the field that is "diffracted" by the entire screen $\mathbb{R}^2$, i.e. no obstacles are present, hence the field has simply propagated a distance of R (the notation reflects the fact that the field does not undergo diffraction). Energy conservation dictates that the total power contained in $\widetilde{\varphi}$ must equal the total incident power, viz. $\int_{\mathbb{R}^2} R^2|\widetilde{\varphi}(\vec{\xi})|^2 = \mathcal{P}_{\text{in}}$, and the total power contained in the field diffracted by the complement of the aperture must be

$$\mathcal{P}_{\overline{\mathcal{A}}^\perp} \triangleq \int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2\left|\bar{\psi}\left(\vec{\xi}\right)\right|^2 = \mathcal{P}_{\text{in}} - \mathcal{P}_{\mathcal{A}^\perp}. \tag{7}$$

From Eq. (4), observe that

$$\int_{\mathbb{R}^2} \equiv \int_{\mathcal{A}^\perp} + \int_{\overline{\mathcal{A}}^\perp},$$

$$\text{therefore} \qquad \widetilde{\varphi}\left(\vec{\xi}\right) = \psi\left(\vec{\xi}\right) + \bar{\psi}\left(\vec{\xi}\right). \tag{8}$$

The left-hand side in relation Eq. (8) is the non-diffracted field $\widetilde{\varphi}$, i.e. the energy that continues to propagate unrestricted. As that energy does not diffract, $\widetilde{\varphi}$ must be zero when $\vec{\xi}$ lies outside the light's beam. In such cases, the terms on the right-hand-side above, ψ, and $\bar{\psi}$, must annihilate, i.e. they quantify diffraction lobes that perfectly destructively interfere (have identical amplitude but a phase difference of π). This is known as *Babinet's principle*, which states that *the diffraction by an aperture produces an identical diffraction pattern to diffraction by its complement.*

Babinet's principle can be rigorously formulated under full electrodynamics [Zangwill 2013], however the above is sufficient for our purposes.

4 FREE-SPACE DIFFRACTIONS

We are now ready to formulate our theory. An extensive supplemental document is provided, where we carry out the derivations and analysis in detail. See Fig. 3 for an overview of our free-space diffractions rendering method.

Our primary contributions are derived in this Section:

(1) In Subsection 4.1 we show that the diffracted field ψ can be written as a sum of a *ray-optical direct term* and a *wave-optical diffracted term*. The diffracted term is induced by the aperture's complement $\overline{\mathcal{A}}^{\perp}$ (a consequence of Babinet's principle). The consequence of this contribution is that we may continue to ray trace free-space propagation: rays that propagate through an aperture quantify the (ray-optical) direct term, which require no diffraction, while diffractions around the obstacle only need to be computed when a ray is incident upon the obstacle. To make both terms physical, we need to slightly bias energy in our light transport simulation, as discussed in Subsection 4.1. The meaning of this approximation is discussed further in Section S3 in our supplemental material.

(2) In Subsection 4.2 we derive a novel formulation of edge-based diffractions, under the context of the Fraunhofer diffraction theory. The incident light is modelled as an arbitrary spatially smoothly-varying incident beam. This formulation enables us to construct a free-space diffraction BSDF, see Algorithm 1 for an overview. We will further discuss this BSDF, including an importance sampling strategy in Section 5.

4.1 Direct and Diffracted Terms

For now, let φ quantify a beam of light that is incident upon a screen with aperture $\mathcal{A}^{\perp}$, as in Subsection 3.1. We will discuss φ in greater detail later.

Consider Babinet's principle (Subsection 3.2). By reordering the terms and taking the complex magnitude squared of both sides in Eq. (8), we write

$$\left|\psi\right|^2 = \left|\widetilde{\varphi} - \overline{\psi}\right|^2 = \left|\widetilde{\varphi}\right|^2 + \left|\overline{\psi}\right|^2 - 2\,\mathrm{Re}\,\overline{\psi}\widetilde{\varphi}^{\star}\,, \tag{9}$$

where $\star$ denotes complex conjugation. Following the discussion in Subsection 3.2, we note that the total power contained in the left-hand side above is $\mathcal{P}_{\mathcal{A}^{\perp}}$, while the total power contained in the $|\widetilde{\varphi}|^2 + |\overline{\psi}|^2$ term is $\mathcal{P}_{\text{in}} + \mathcal{P}_{\overline{\mathcal{A}}^{\perp}}$. Therefore, by integrating the power contained in both sides in Eq. (9), we must conclude that the total power contained in the $\mathrm{Re}\,\overline{\psi}\widetilde{\varphi}^{\star}$ lobe is

$$\begin{aligned}\mathrm{Re}\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2 \overline{\psi}\widetilde{\varphi}^{\star} &= \frac{1}{2}\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2 \left(\left|\widetilde{\varphi}\right|^2 + \left|\overline{\psi}\right|^2 - \left|\psi\right|^2\right) \\ &= \frac{\mathcal{P}_{\text{in}} + \mathcal{P}_{\overline{\mathcal{A}}^{\perp}} - \mathcal{P}_{\mathcal{A}^{\perp}}}{2} = \mathcal{P}_{\overline{\mathcal{A}}^{\perp}}\,,\end{aligned} \tag{10}$$

where we used Eq. (7). The total power contained in the $\mathrm{Re}\,\overline{\psi}\widetilde{\varphi}^{\star}$ lobe is then the total power that is incident upon $\overline{\mathcal{A}}^{\perp}$, i.e. the complement of the aperture, which is the obstacle that diffracts light.

We denote the following key terms (illustrated in Fig. 3a):

$$\text{direct term:} \qquad d \triangleq |\widetilde{\varphi}|^2 - \mathrm{Re}\,\overline{\psi}\widetilde{\varphi}^{\star}\,, \tag{11}$$

$$\text{diffracted term:} \qquad w \triangleq \left|\overline{\psi}\right|^2 - \mathrm{Re}\,\overline{\psi}\widetilde{\varphi}^{\star}\,. \tag{12}$$

Then, we may rewrite the diffracted intensity (Eq. (9)) as

$$|\psi|^2 = d + w\,. \tag{13}$$

$|\psi|^2$ is the intensity of the wave diffracted by the aperture. This is precisely the term of interest that quantifies the intensity of the free-space diffracted field, and the term we wish to render. In Eq. (13), we have rewritten it exactly as an incoherent sum (i.e, a sum where no interference arises) of two terms:

Direct term. The power contained in the direct term is

$$\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2\, d\left(\vec{\xi}\right) = \mathcal{P}_{\mathcal{A}^{\perp}} \tag{14}$$

(directly from Eqs. (10) and (11)), i.e. exactly the power that is incident upon the opening $\mathcal{A}^{\perp}$. Also, we may immediately see from the definition of the direct term that

$$d\left(\vec{\xi}\right) \approx 0 \quad \text{when} \quad |\widetilde{\varphi}| \approx 0\,, \tag{15}$$

i.e., the direct term vanishes when $|\widetilde{\varphi}|$ vanishes. Recall that $\widetilde{\varphi}$ is the unobstructed, free-space propagated incident light, that undergoes no diffraction, and hence is non-vanishing only in a narrow range of directions around $\vec{\xi} = 0$. We define the *central lobe* to be that set of directions: i.e., the directions $\vec{\xi}$ where $|\widetilde{\varphi}(\vec{\xi})|$ is non-negligible.

For example, given an incident beam with a Gaussian profile with a spatial standard deviation of 25λ, the free-space propagated field $\widetilde{\varphi}$ (Eq. (6)) describes a beam where its full width at half maximum subtends a cone with a tiny half-angle of about $0.3°$. This cone constitutes the central lobe.

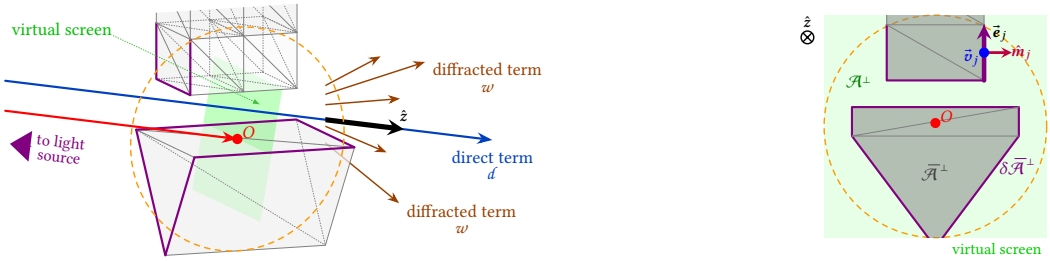
By construction, the direct term describes the unobstructed, free-space propagated lobe, which contains $\mathcal{P}_{\mathcal{A}^{\perp}}$ total power, i.e. precisely the total power that is incident upon the opening $\mathcal{A}^{\perp}$. Intuitively, the direct term is simply the light that falls upon the opening, and continues to propagate without diffracting. Then, this term is described well via ray optics: rays that pass through the opening continue to propagate unobstructed, and rendering the direct term reduces to ray tracing the propagation of light in free space.

Diffracted term. The diffracted term is the energy that diffracts through the aperture (or around the obstacle), and gives rise to energy that propagates into directions outside the central lobe. The diffracted term cannot be described ray-optically. In our rendering algorithm, we consider the diffracted term when rays impinge upon the obstacle, i.e. the aperture complement $\overline{\mathcal{A}}^{\perp}$. This is a manifestation of Babinet's principle: In directions that lie outside the central lobe, i.e. where $|\widetilde{\varphi}| \approx 0$, $|\overline{\psi}|^2$ and $|\psi|^2$ both describe an identical diffraction pattern.

However, as currently defined, the diffraction term w is not physical: clearly, the total power contained in this lobe is 0:

$$\int_{\mathbb{R}^2} \mathrm{d}^2\vec{\xi}\, R^2\, w\left(\vec{\xi}\right) = \mathcal{P}_{\overline{\mathcal{A}}^{\perp}} - \mathcal{P}_{\overline{\mathcal{A}}^{\perp}} = 0 \tag{16}$$

(using Eqs. (10) and (12)). Therefore, either $w \equiv 0$ and no diffraction happens, or in some directions within the central lobe, where



(a) **Direct and diffracted terms** (b) **Diffracting aperture**

Fig. 3. **Overview of our method.** Consider a geometric mesh (triangles in gray). We continue to path trace as usual. (a) Light rays that pass through without intersecting the mesh (rays in blue) give rise to the *direct term*: energy that does not undergo diffraction. When rays intersect the mesh (rays in red, intersection point O, propagation direction $\hat{\boldsymbol{z}}$), we project all the front-facing triangles, that we find inside a search radius (dashed orange line) around O, upon the *virtual screen* (in green, defined as the plane that passes through the intersection point O with normal $-\hat{\boldsymbol{z}}$, i.e. is orthogonal to the incident ray). (b) The projected geometry in the vicinity of O upon the virtual screen gives rise to the two-dimensional diffracting obstacle $\overline{\mathcal{A}}^\perp$. Its complement $\mathcal{A}^\perp$ is the opening in the screen (referred to as the *aperture*), and its boundary is $\delta\overline{\mathcal{A}}^\perp$. The boundary is composed of the edges (in violet) that belong to exactly one front-facing triangle. Edges are parameterized by their middle point $\vec{\boldsymbol{v}}_j$, edge vector $\vec{\boldsymbol{e}}_j$, and outward-facing (into the aperture $\mathcal{A}^\perp$) normal $\hat{\boldsymbol{m}}_j$. This set of edges gives rise to the *free-space diffraction BSDF*: quantifying the angular distribution of the *diffracted term*, i.e. rays that interact with the mesh and diffract through the aperture. Note, we do not ignore the depth of the aperture (before projection) by accounting for the z distance light travels over it.

$\widetilde{\varphi}$ is non-negligible, the diffracted term must take aphysical negative values, viz. $w(\vec{\boldsymbol{\xi}}) < 0$ for some $\vec{\boldsymbol{\xi}}$. To understand the reason, note that the energy for the diffracted lobes should be taken out of energy that impinges both upon the opening as well as the diffracting obstacle. However, our direct term is defined as the ray-optical term, where energy propagates through the aperture unobstructed.

Our approach is as follows: we clamp the diffracted term w to 0, thereby avoiding aphysical negative intensities. This clamped diffracted field clearly has positive total power. In order to conserve energy, we need to "lend" that power from elsewhere, and we take that power from the energy that is incident upon the obstacle $\overline{\mathcal{A}}^\perp$. That is, we take energy from the back-scattered or absorbed energy, and redirect it into the diffracted lobes. We are only required to do so where free-space diffractions occur: in the immediate regions of the geometric edges of objects. How we do so in practice will be discussed later in Section 6.

Implications of clamping the diffracted term. The relations Eqs. (9) and (11) to (13) are analytically exact. The only approximation that we made in this Subsection is biasing energy from back-scattering or absorption towards the direct term. The diffracted lobes remain accurate (because outside the central lobe, where $\widetilde{\varphi}$ vanishes, w remains unchanged). Because this only applies where both $\overline{\mathcal{A}}^\perp, \mathcal{A}^\perp$ are non-negligible (i.e. only around the edges of geometry, where both an obstacle and an opening are present), in effect the bias introduced into our rendering algorithm can be understood as making all geometry very slightly (several wavelengths) smaller. The powerful benefit of this decomposition into a direct and diffracted term is that free-space propagation is fully described by ray tracing.

4.2 Edge-Diffracted Waves

In this Subsection we will devise a closed-form expression for the Fraunhofer-diffracted field $\overline{\psi}$, for any polygonal aperture and under a piecewise-linear approximation of the incident beam's field, as we will discuss. This closed-form solution is expressed via a few simple analytic functions, and will be instrumental in formulating our free-space diffraction BSDF.

Assume that the scene geometry is composed of non-degenerate, non self-intersecting triangular meshes. Let a ray be incident upon geometry. Let the triangular mesh around the interaction point between the ray and geometry be composed of the set of triangles $\{T_t\}$. We denote the plane that passes through the intersection point of the ray with the geometry as the virtual screen (see Fig. 3), and project that triangular mesh in the vicinity of the intersection point upon that screen, giving rise to the *projected polyhedron*:

$$\overline{\mathcal{A}}^\perp \triangleq \sum\nolimits_t T_t^\perp \,, \tag{17}$$

where $T_t^\perp$ denotes the projection of the triangle T_t upon the screen. Thereby, we have defined the aperture complement $\overline{\mathcal{A}}^\perp$ as the projected triangular mesh from the vicinity of the intersection of a ray with geometry. "Vicinity" formally means the region over which light is able to reproduce diffraction effects. Modern wave-optical light transport frameworks are able to rigorously quantify that region (roughly multiple dozens of wavelengths), and we will discuss that in more detail later. The aperture $\mathcal{A}^\perp$ is then the complement of $\overline{\mathcal{A}}^\perp$, and is the regions on the virtual screen where no projected geometry exists. The aperture construction steps, as detailed above, are summarised in lines 5-8 in Algorithm 1.

For now, we assume that the triangles are front facing (we cull back-facing triangles) and that the projected, two-dimensional triangles do not overlap, viz. $\forall j \neq l, T_t^\perp \cap T_l^\perp = \varnothing$. For each projected triangle, let its vertices be $\vec{\boldsymbol{u}}_{t,1}, \vec{\boldsymbol{u}}_{t,2}, \vec{\boldsymbol{u}}_{t,3} \in \mathbb{R}^2$, and we denote its edge vectors (connecting the edge's vertices) and outward-facing edge normals as

$$\vec{\boldsymbol{e}}_{t,j} \triangleq \vec{\boldsymbol{u}}_{t,j-1} - \vec{\boldsymbol{u}}_{t,j+1} \qquad \text{and} \qquad \hat{\boldsymbol{m}}_{t,j} \triangleq \pm\frac{\vec{\boldsymbol{e}}_{t,j} \times \hat{\boldsymbol{z}}}{|\vec{\boldsymbol{e}}_{t,j} \times \hat{\boldsymbol{z}}|} \tag{18}$$

Algorithm 1: Free-space diffraction BSDF construction.

Function ConstructBSDF(ray, *intersection point*):
  screen ← frame(ray) ; /* Construct orthogonal frame around ray */
  edges ← {} ; /* Array will hold all diffracting edges */
  $\mathcal{P}^{(\mathrm{PL})}_{\overline{\mathcal{A}}^{\perp}}, \widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}} \leftarrow 0$; /* Total powers in diffracted field */
  /* Query all triangles around the intersection point */
  foreach T_t *within search radius* **do**
    if T_t *is back facing* **then**
      continue
    $T_t^{\perp} \leftarrow$ screen.project(T_t)
    foreach *edge* $(\vec{\boldsymbol{v}}_j, \vec{\boldsymbol{e}}_j)$ *in* $T_t^{\perp}$ **do** /* Find diffracting edges */
      if *edge belongs to another front-facing triangle* **then**
        continue
      $\{\hat{\boldsymbol{m}}_j, a_j, b_j\} \leftarrow$ Eqs. (18) and (24)
      edges ← edges ∪ $\{(\vec{\boldsymbol{v}}_j, \vec{\boldsymbol{e}}_j, \hat{\boldsymbol{m}}_j, a_j, b_j)\}$
    /* Accumulate total power in diffracted field and central lobe */
    $\mathcal{P}^{(\mathrm{PL})}_{\overline{\mathcal{A}}^{\perp}}$ += $\mathcal{P}^{(\mathrm{PL})}_{T_t^{\perp}}$; /* Eq. (31) */
    $\widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}} \leftarrow$ compute power in central lobe using Eqs. (34) and (35)
  return $\left(\text{edges}, \text{screen}, \mathcal{P}^{(\mathrm{PL})}_{\overline{\mathcal{A}}^{\perp}}, \widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}}\right)$; /* The tuple defines the BSDF */

respectively, where the vertex indices should be understood in the modulo sense, i.e., $\vec{\boldsymbol{u}}_{t,4} \equiv \vec{\boldsymbol{u}}_{t,1}$ and so on. $-\hat{\boldsymbol{z}}$ is the projected triangle's normal (i.e., $\hat{\boldsymbol{z}}$ is the ray's direction of propagation). The sign of $\hat{\boldsymbol{m}}$ depends on the triangle winding: positive for counterclockwise winding, and negative otherwise. The cross product above should be understood in the sense where $\vec{\boldsymbol{e}}_{t,j}$ are three-dimensional vectors on the virtual screen. We write the edge lengths as $e_{t,j} = |\vec{\boldsymbol{e}}_{t,j}|$.

The incident beam. While we use rays for propagation, free-space diffraction is a wave-optical phenomenon, and we still understand light as waves. Our method is designed to operate within the context of a *weakly-local* wave-optical light transport framework (most notably Steinberg et al. [2023, 2022]), which provide the machinery to understand light as a collection of electromagnetic beams. Then, ray tracing is understood as tracing the centre of a beam φ. Such light transport frameworks quantify rigorously the beam profile φ, and the region around its centre where it is non-negligible.

Despite the above, the contributions in this paper require no understanding of wave-optical light transport frameworks. The reader may understand φ as a beam with a Gaussian profile and a known spatial variance, and that is the beam profile that is used in all of our results (indeed, Steinberg et al. [2023] formally understands light as a collection of such Gaussian beams). However, for this paper the incident field φ may be an arbitrary smooth C^2 function that decays to 0 away from its centre. The region in which φ is non-negligible is the vicinity in which scene geometry may diffract the beam. When a ray is incident upon geometry, we look for all triangles T_t within that region, and construct the aperture complement $\overline{\mathcal{A}}^{\perp}$ as in Eq. (17).

Piecewise-linear approximation. Given an incident beam φ and a diffracting obstacle $\overline{\mathcal{A}}^{\perp}$, composed of the projected scene triangles,

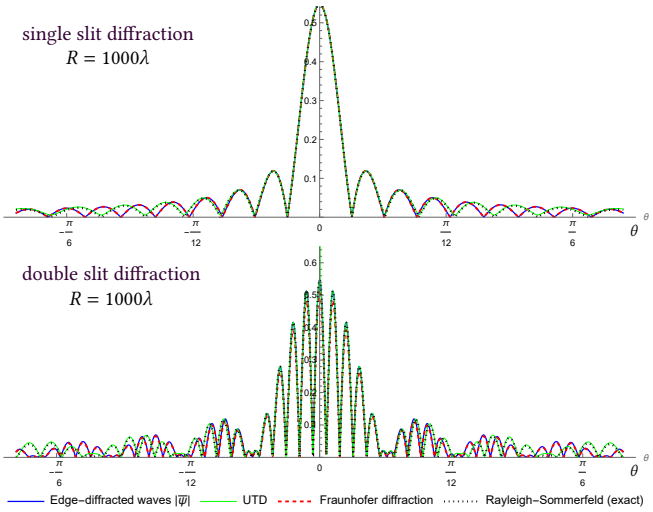


Fig. 4. **Comparison with UTD and diffraction integrals.** Square root of light intensity diffracted by a single-slit (top) and double-slit (bottom) apertures, as functions of scattering direction θ. We compare our formulation of edge-diffracted waves $\overline{\psi}$ (blue), the Fraunhofer diffraction integral (dashed red), UTD-based ray diffraction (green), and the exact Rayleigh-Sommerfeld diffraction integral [Born and Wolf 1999] (dotted black). All methods agree on the lower-order fringes. Minor deviations in the high-order fringes arise due to the short distance R from the aperture.

we piecewise-linearly approximate (PLA) φ over the projected triangular mesh:

$$\varphi_{\mathrm{PL}}(\vec{\boldsymbol{q}}^{\perp}) \triangleq \begin{cases} \sum_j \gamma_{t,j}\varphi_{t,j} & \text{given} \quad \exists t \text{ s.t. } \vec{\boldsymbol{q}}^{\perp} \in T_t^{\perp} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

$$\text{with} \qquad \varphi_{t,j} \triangleq \varphi(\vec{\boldsymbol{u}}_{t,j})\mathrm{e}^{-\mathrm{i}kz(\vec{\boldsymbol{u}}_{t,j})}\,, \tag{20}$$

where, t is the index of the (unique) triangle that contains $\vec{\boldsymbol{q}}^{\perp}$ if any, $\gamma_{t,j}$ are the *barycentric coordinates* of the point $\vec{\boldsymbol{q}}^{\perp}$ on the triangle (see Subsection S1.2 in our supplemental material), and $\varphi_{t,j}$ are shorthands denoting the values φ takes on the vertices of the triangle. For any point $\vec{\boldsymbol{u}}^{\perp} \in T_t^{\perp}$ that belongs to a triangle on the virtual screen, $z(\vec{\boldsymbol{u}}^{\perp})$ is defined to be the z coordinate of that point on T_t before projection, such that $z = 0$ on the virtual screen, $z > 0$ for points that lie behind the screen. As we assume no overlapping triangles, this function is well defined.

Our results remain accurate as long as the piecewise-linear approximation above is a good approximation, i.e. $\varphi_{\mathrm{PL}} \approx \varphi$. In Fig. 3 in our supplemental material, we show that for a Gaussian beam this approximation produces essentially exact results as long as the edge lengths of the triangles do not exceed the spatial variance of the incident beam profile. In our implementation we tessellate on-the-fly the triangles, as needed, as we build $\overline{\mathcal{A}}^{\perp}$.

Note that we correctly account for the depth light travels across the diffracting obstacle via the complex phasor in Eq. (20).

Fraunhofer edge diffraction. We now present a closed-form solution to the problem of Fraunhofer diffraction (Eq. (1)) of the PLA approximant of the incident beam over a polygonal aperture. This is a

primary contribution. Our starting point is the Fraunhofer diffraction integral of φ_{PL} over the diffracting obstacle $\overline{\mathcal{A}}^{\perp}$, viz.

$$\bar{\psi}\left(\vec{\xi}\right) \triangleq \frac{k}{2\pi R}\int_{\overline{\mathcal{A}}^{\perp}} \mathrm{d}^2\vec{q}^{\perp}\,\varphi_{\text{PL}}(\vec{q}^{\perp})\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{q}^{\perp}}\,. \tag{21}$$

In our supplemental material, Section S2, we detail the derivations that follow. First, we rewrite the above as a line integral over the boundary of the aperture complement, denoted $\delta\overline{\mathcal{A}}^{\perp}$, via an application of the divergence theorem:

$$\bar{\psi}\left(\vec{\xi}\right) \approx \frac{\mathrm{i}}{2\pi R\xi^2}\oint_{\delta\overline{\mathcal{A}}^{\perp}} \mathrm{d}s\left(\hat{\boldsymbol{m}}\cdot\vec{\xi}\right)\varphi_{\text{PL}}(\vec{q}^{\perp})\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{q}^{\perp}}\,, \tag{22}$$

where $\xi = |\vec{\xi}|$ and $\hat{\boldsymbol{m}}$ is the outwards-pointing normal of the boundary. This boundary may be self intersecting or disjoint, meaning there exists more than one obstacle or opening.

A triangle edge $\vec{\boldsymbol{e}}_{t,j}$ is part of the boundary $\delta\mathcal{A}^{\perp}$ if and only if that triangle edge does not belong to any other (front-facing) triangle. In our implementation, we assume that all the scene geometry is composed of triangle strips, meaning an edge can belong to at most a couple of adjacent triangles. As we look for triangles around the intersection point of a ray with geometry, we identify the edges that compose the boundary $\delta\overline{\mathcal{A}}^{\perp}$. We denote the edges of the boundary as $\vec{\boldsymbol{e}}_j$, their outward-facing normals as $\hat{\boldsymbol{m}}_j$, and the edge middle points as $\vec{\boldsymbol{v}}_j$. An edge's vertices are then $\vec{\boldsymbol{v}} \pm \frac{1}{2}\vec{\boldsymbol{e}}_j$, and the edge's length is $e_j = |\vec{\boldsymbol{e}}_j|$. See Fig. 3b.

Clearly, the line integral in Eq. (22) is a piecewise sum over the line integrals of the edges $\vec{\boldsymbol{e}}_j$ that compose the obstacle boundary. We term the line integral over a single edge $\vec{\boldsymbol{e}}_j$ as the *edge-diffracted wave*, and produce a closed-form solution:

$$\bar{\psi}_j\left(\vec{\xi}\right) \triangleq \frac{\mathrm{e}^{-\mathrm{i}k\vec{\xi}\cdot\vec{\boldsymbol{v}}_j}}{kR|\boldsymbol{\Xi}_j|}\left[(a_j-b_j)\alpha_1\left(\boldsymbol{\Xi}_j^{-\intercal}\vec{\xi}\right) + \mathrm{i}\tfrac{a_j+b_j}{2}\alpha_2\left(\boldsymbol{\Xi}_j^{-\intercal}\vec{\xi}\right)\right]. \tag{23}$$

Then, the total diffracted field (Eq. (22)) is simply $\bar{\psi} = \sum_j \bar{\psi}_j$. The linear transform $\boldsymbol{\Xi}$ defines the edge's geometry, and a_j, b_j are the values the incident light, φ_{PL}, takes on the edge's vertices, viz.

$$a_j, b_j \triangleq \varphi\left(\vec{\boldsymbol{v}}_j \mp \tfrac{1}{2}\vec{\boldsymbol{e}}_j\right)\mathrm{e}^{-\mathrm{i}kz\left(\vec{\boldsymbol{v}}_j \mp \frac{1}{2}\vec{\boldsymbol{e}}_j\right)}\,, \tag{24}$$

$$\text{and}\quad \boldsymbol{\Xi}_j^{-1} \triangleq k\begin{pmatrix} | & | \\ \vec{\boldsymbol{e}}_j & e_j\hat{\boldsymbol{m}}_j \\ | & | \end{pmatrix}. \tag{25}$$

The following important auxiliary functions are also defined:

$$\alpha_1\left(\vec{\zeta}\right) \triangleq \frac{\zeta_y}{2\pi\zeta^2\zeta_x}\left(\cos\tfrac{\zeta_x}{2} - \operatorname{sinc}\tfrac{\zeta_x}{2}\right) \tag{26}$$

$$\text{and}\quad \alpha_2\left(\vec{\zeta}\right) \triangleq \frac{\zeta_y}{2\pi\zeta^2}\operatorname{sinc}\tfrac{\zeta_x}{2}\,, \tag{27}$$

where $\zeta_{x,y}$ are the Cartesian components of the argument $\vec{\zeta} \in \mathbb{R}^2$, and $\zeta^2 = \vec{\zeta}\cdot\vec{\zeta}$ is the squared length, as before. When $\zeta_x = 0$, $\alpha_{1,2}$ are understood in the limit $\zeta_x \to 0$. See Section S2 in our supplemental material for more information and additional analysis of the functions $\alpha_{1,2}$.

Eq. (23), with the definitions Eqs. (24) to (27), is the primary result in this Subsection. We conclude that we may express the Fraunhofer-diffracted field $\bar{\psi}$, of any polygonal aperture and with any PLA-approximated incident field φ_{PL}, via a closed-form solution and using just a pair of simple functions, $\alpha_{1,2}$, linearly transformed and superimposed. The functions $\alpha_{1,2}$ do not depend on the aperture geometry or light's wavelength—a fact we will use to be able to importance sample the diffracted term w. A numeric comparison (Fig. 4) of our edge-diffracted waves $\bar{\psi}$ (Eq. (23)) with Fraunhofer diffraction (Eq. (1)) shows that our derivations are very accurate.

Observe that the edge-diffracted fields of a pair of opposing edges (i.e., $\vec{\boldsymbol{e}}_l = -\vec{\boldsymbol{e}}_j, \hat{\boldsymbol{m}}_l = -\hat{\boldsymbol{m}}_j, a_l = b_j, b_l = a_j$) exactly annihilate each other, viz. $\bar{\psi}_j = -\bar{\psi}_l$ (where we use the fact that α_1 is odd in both x and y, and α_2 is even in x and odd in y). This is the formal justification to us going from an integral over a collection of triangles (Eq. (21)) to a line integral over the aperture boundary (Eq. (22)), and means that edges that belong to a pair of front-facing triangles are discarded. Lines 9-13 in Algorithm 1 summarise the detection of the diffracting edges of the boundary.

The clamped diffracted term. Having derived an expression for the field $\bar{\psi}$ diffracted by the complement aperture $\overline{\mathcal{A}}^{\perp}$, we turn our attention back to the diffracted term w (Eq. (12)). As discussed, in directions of the central lobe, w will mostly take negative values. Negative intensities are not physical. Our solution is to extract the central lobe out of $\bar{\psi}$, giving rise to the clamped diffracted term.

In Subsection S2.1 in our supplemental material we detail how we extract from each edge-diffracted wave the energy contained in the tightest central lobe that can be physically resolved, yielding the *edge-diffracted wave with its central lobe removed*, denoted

$$\widehat{\psi}_j\left(\vec{\xi}\right) \triangleq \sqrt{1-\chi\left(\boldsymbol{\Xi}_j^{-\intercal}\vec{\xi}\right)}\,\bar{\psi}_j\left(\vec{\xi}\right), \tag{28}$$

where we use the transform defined in Eq. (25). χ is the modulating Gaussian function:

$$\chi\left(\vec{\zeta}\right) \triangleq \exp\left(-\tfrac{1}{2\sigma_\zeta^2}\zeta^2\right), \qquad \text{with}\quad \sigma_\zeta = \sqrt{3}\,. \tag{29}$$

We may now define the *clamped diffracted term* as the total diffracted field, with the central lobe removed:

$$\widehat{w} \triangleq \left|\sum_j \widehat{\psi}_j\right|^2. \tag{30}$$

Indeed, $\widehat{w}(\vec{\xi}) \approx 0$ when $\vec{\xi} \approx 0$, i.e. around the central lobe, as desired. In contrast to the diffracted term w, defined in Eq. (12), the clamped diffracted term $\widehat{w}$ is always non-negative and physical. Clearly $\widehat{w}(\vec{\xi}) = w(\vec{\xi})$ for $\vec{\xi}$ outside the central lobe.

Total powers contained in the fields. In order to normalize and formulate the clamped diffracted term in Eq. (30) as a BSDF, as well as to be able to importance sample said BSDF, we need to compute the total power contained in the discussed fields. In Subsection S2.2 in our supplemental material we derive the relations that follow.

The total power in the diffracted field $\bar{\psi}$, under the PLA approximation, is (by energy conservation) the total power that impinges upon the diffracting obstacle, i.e. the aperture complement $\overline{\mathcal{A}}^{\perp}$:

$$\mathcal{P}^{(\text{PL})}_{\overline{\mathcal{A}}^{\perp}} \triangleq \sum_t \mathcal{P}^{(\text{PL})}_{T_t^{\perp}}\,, \quad \text{with}\quad \mathcal{P}^{(\text{PL})}_{T_t^{\perp}} = \tfrac{1}{6}|S_t|\,\mathrm{Re}\sum_{1\le j\le l\le 3}\varphi_{t,j}\varphi^{\star}_{t,l}\,, \tag{31}$$

where S_t is the signed area (Eq. (S1.9) in our supplemental material) of the projected triangle $T_t^{\perp}$, and recall the definition of $\varphi_{t,j}$

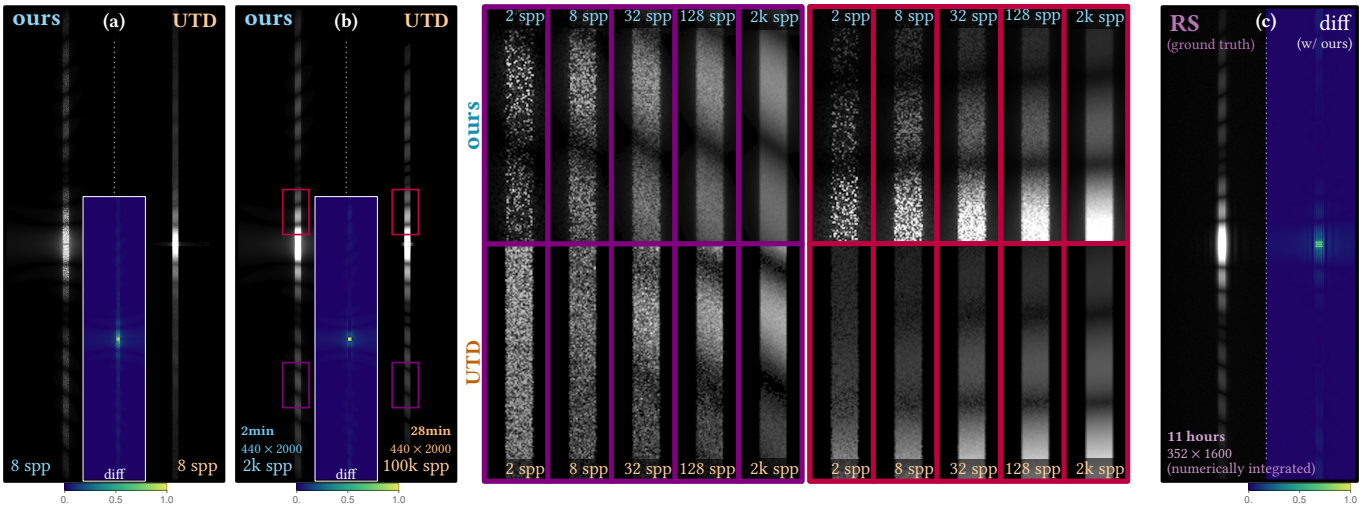


Fig. 5. **UTD-based ray tracing: interfering rays induce heavy noise.** Double-slit diffraction patterns produced by our method and UTD. UTD-based ray tracing employs rays that oscillate rapidly at light's frequency and interfere with each other. This gives rise to numeric issues and significant noise: at insufficient sample counts, UTD is unable to reconstruct the pattern and induces heavy, high-frequency and high-intensity noise. Despite the simple aperture, hundreds of samples are required for the dark fringes to emerge. (a) Observe the 8spp UTD pattern: rays scatter energy over the *entire pattern*. In contrast, our method captures the entire aperture geometry with a single sample, and the diffraction pattern with its fringes materializes even at few spp, making our method much more suitable for a path tracing simulation. (b) Once converged both methods predict similar diffraction lobes. The rendered patterns show the right wall of a Cornell box, which is illuminated by light passing through a screen with a pair of slits (similar to Fig. 8). We render the scene with a simplified form of path tracing to compare our method to UTD: for every given point on the right wall, (ours) each sample selects a random point around the aperture on the screen, and connects the given wall point to the light source through the aperture by evaluating our free-space diffraction BSDF; or, (UTD) each sample chooses a random edge, then, if there exists a valid UTD path between that edge and the given wall point, UTD makes the connection. The insets on the right show close-ups of the regions highlighted in (b), for various spp. Runtime per 1spp: (UTD) 16 ms; (ours) 62 ms. (c) We also compare our method against the numerically-integrated solution of the Rayleigh-Sommerfeld (RS) diffraction integral, which serves as a ground truth.

(Eq. (20)). Eq. (31) is a sum over the power that impinges upon each projected triangle. We compute this quantity as we identify the triangles in the vicinity of a ray and build the aperture complement.

The total power contained in a single edge-diffracted wave with its central lobe removed (Eq. (28)) is

$$\widehat{\mathcal{P}}_j^{(\mathrm{PL})} \triangleq e_j^2 \left[\left| a_j - b_j \right|^2 \mathfrak{I}_1 + \left| \tfrac{a_j + b_j}{2} \right|^2 \mathfrak{I}_2 \right], \tag{32}$$

$$\text{with} \qquad \mathfrak{I}_1 \triangleq 0.0045255085 \quad \text{and} \quad \mathfrak{I}_2 \triangleq 0.11487543\,. \tag{33}$$

The quantities $\mathfrak{I}_{1,2}$ are numerically-integrated total powers in $\alpha_{1,2}$, after removal of their central lobes. We make a slight approximation in Eq. (32) and neglect a cross-term, for simplicity. This is negligible: the relative error is less than 10^{-7}. The exact expression is given in Eq. (S2.33) in our supplemental material.

To compute the total power contained in the clamped diffracted term will need to approximate the total power contained in the central lobe, denoted $\widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}}$. First, we compute the value the diffracted field $\bar{\psi}$ takes at $\vec{\xi} = 0$ in closed-form directly from the diffraction integral Eq. (21):

$$\bar{\psi}(0) = \frac{k}{2\pi R} \int_{\bar{\mathcal{A}}^\perp} \mathrm{d}^2\vec{q}^{\,\perp}\, \varphi_{\mathrm{PL}}(\vec{q}^{\,\perp}) = \frac{k}{6\pi R} \sum_t |S_t| \sum_j \varphi_{t,j}\,. \tag{34}$$

See Eq. (B.2) in our supplemental material for derivation. We assume that around $\vec{\xi} = 0$, the Gaussian χ (Eq. (29)) dominates the falloff of the central lobe. Then,

$$\widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}} \approx 2\pi \left(\frac{3}{\sqrt{2} k e_{\max}} \right)^2 \left| \bar{\psi}(0) \right|^2, \tag{35}$$

where $e_{\max}$ is the length of the longest edge that participates in diffraction. Eq. (35) should be understood as the power in a Gaussian-like lobe with peak $|\psi(0)|^2$. We discuss this approximation further in Section 6. Then, the total power contained in the clamped diffracted term is

$$\widehat{\mathcal{P}}^{(\mathrm{PL})}_{\bar{\mathcal{A}}^\perp} \triangleq \mathcal{P}^{(\mathrm{PL})}_{\bar{\mathcal{A}}^\perp} - \widetilde{\mathcal{P}}^{(\mathrm{PL})}_{\mathrm{central}}\,. \tag{36}$$

Lines 14-15 in Algorithm 1 use the expressions above in order to compute the total powers contained in the diffracted field and central lobe.

5 THE BSDF

In this section we formulate the diffracted term $\widehat{w}$ (Eq. (30)) as a *bidirectional scattering distribution function* (BSDF), and devise a simple importance sampling strategy. A BSDF quantifies the scattered radiance to incident irradiance ratio. Using the relation between a differential surface area $\mathrm{d}A$ and the differential solid angle $\mathrm{d}\Omega$ subtended by that surface, viz. $R^2\,\mathrm{d}\Omega = \mathrm{d}A$, where R is the distance, we may write (see Steinberg and Yan [2021, Eq. 18]):

$$\text{radiance} = R^2 \frac{\partial \text{intensity}}{\partial A_s \cos\theta}\,, \tag{37}$$

Algorithm 2: Evaluating the BSDF.

Function `EvalBSDF(bsdf,`$\hat{\boldsymbol{\omega}}_o$`):`
  $\left[\sin\theta_x, \sin\theta_y\right]^\intercal \leftarrow$ bsdf.screen.project($\hat{\boldsymbol{\omega}}_o$) ; /* Project the scattering direction upon the screen, yielding the directional sines */
  $\vec{\xi} \leftarrow \left[\tan\theta_x, \tan\theta_y\right]^\intercal$
  $\widehat{\psi} \leftarrow 0$
  foreach edge *in* bsdf.edges **do** /* Loop over all BSDF edges */
    $\widehat{\psi}$ += $\widehat{\psi}_j\left(\vec{\xi}\right)\Big|_{R=1}$; /* Edge-diffracted wave. Eq. (28) with $R = 1$. */
  return $\frac{1}{\text{bsdf}.\mathcal{P}^{(\text{PL})}_{\overline{\mathcal{A}}^\perp}\cos\theta}\left|\widehat{\psi}\right|^2$

Algorithm 3: Importance sampling the BSDF.

Function `ImportanceSampleBSDF(bsdf)`:
  /* We use precomputed iCDF tables to sample from the following */
  $\vec{\zeta}_1, \vec{\zeta}_2 \leftarrow$ importance sample from $(1-\chi)|\alpha_1|^2$, $(1-\chi)|\alpha_2|^2$
  $j \leftarrow$ select edge index from bsdf.edges, with selection probabilities proportional to edge powers $\widehat{\mathcal{P}}^{(\text{PL})}_j$; /* Eq. (32). */
  $\vec{\zeta} \leftarrow$ select from $\left\{\vec{\zeta}_1, \vec{\zeta}_2\right\}$ with probabilities proportional to $\left\{|a_j - b_j|^2, \left|\frac{a_j+b_j}{2}\right|^2\right\}$.
  $\vec{\xi} \leftarrow \Xi_j^\intercal \vec{\zeta}$; /* Ξ_j as defined in Eq. (25). */
  pdf $\leftarrow$ Eq. (40) with $R = 1$
  return $\left(\vec{\xi}, \text{pdf}\right)$; /* $\vec{\xi}$ in local (screen) frame. */

with A_s being the sourcing area (i.e. aperture area) and θ the inclination angle.

Using the above, we define the *free-space diffraction BSDF*, for scattering direction $\hat{\boldsymbol{\omega}}_o$ as

$$f(\hat{\boldsymbol{\omega}}_o) \triangleq \frac{R^2}{\mathcal{P}^{(\text{PL})}_{\overline{\mathcal{A}}^\perp}\cos\theta}\widehat{w}\left(\vec{\xi}\right), \tag{38}$$

where $\vec{\xi} = \left[\tan\theta_x, \tan\theta_y\right]^\intercal$, where $\theta_{x,y}$ are the local (with respect to the virtual screen) axis-aligned scattering directions defined by $\hat{\boldsymbol{\omega}}_o$, and $\theta = \hat{\boldsymbol{\omega}}_o \cdot \hat{\boldsymbol{z}}$ is the inclination angle, with $-\hat{\boldsymbol{z}}$ being the direction of incidence, as before. The total power contained in the diffracted field, $\mathcal{P}^{(\text{PL})}_{\overline{\mathcal{A}}^\perp}$, defined in Eq. (31), acts as the normalization factor as well as the (integrated) sourcing area. Algorithm 2 summarises the evaluation of the BSDF. Simple dimensional analysis shows that f has units of sr^{-1}.

Note that the direction of incidence is not a parameter to the BSDF, instead it is implicitly captured by $\widehat{w}$ during construction of the aperture complement $\overline{\mathcal{A}}^\perp$: we project scene geometry upon the virtual screen that is orthogonal to the incident ray. The scattering direction $\hat{\boldsymbol{\omega}}_o = \hat{\boldsymbol{z}}$ (where $\theta_x = \theta_y = 0$ and $\vec{\xi} = 0$) is the direction of the central lobe and direct term. Also note that the propagation distance term R that appears above cancels out with the R term in Eq. (23), and knowledge of propagation distance after diffraction is not required. The BSDF is not reciprocal in general (due to the construction process and projection of the aperture complement). It conserves energy, by construction, and is always non-negative.

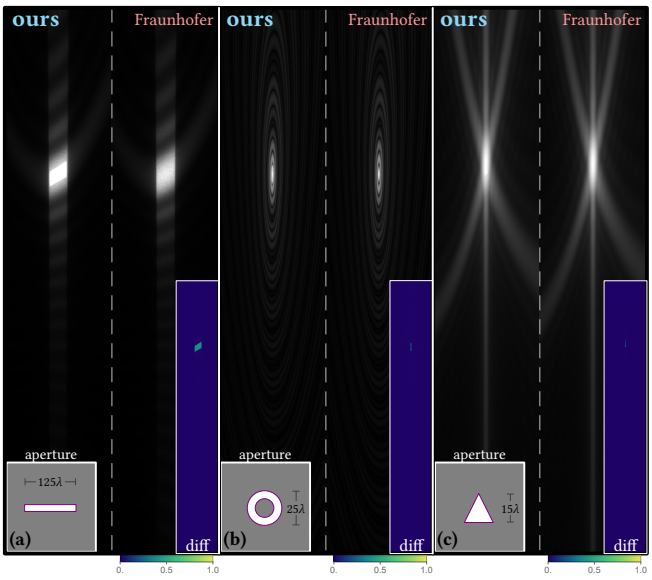


Fig. 6. **Validation against numerically-integrated Fraunhofer diffraction patterns.** Similar experiment to Fig. 5, with the upper half of the right wall displayed. Light is incident at an angle of 30° to the screen. The diffracting aperture is illustrated at the bottom right of each pattern: (a) a single slit; (b) a ring, where our method discretizes each circle into 128 edges; and, (c) a triangle. The Fraunhofer patterns are numerically integrated (of Eq. (1)). With the exception of the centres of the patterns, i.e. the direct term, both methods produce identical results. Contrast and gamma adjusted for visualization purposes.

5.1 Importance Sampling

Consider the clamped diffracted term, that appears in the BSDF (Eq. (38)) above. From Eq. (30) we note

$$\widehat{w}\left(\vec{\xi}\right) = \left|\sum_j \widehat{\psi}_j\right|^2 = \underbrace{\sum_j \left|\widehat{\psi}_j\right|^2}_{\text{edge-diffracted lobes}} + \underbrace{2\sum_{l>j} \text{Re}\,\widehat{\psi}_j\widehat{\psi}_l^\star}_{\text{interference factors}}. \tag{39}$$

The non-negative intensities of the edge-diffracted waves $|\widehat{\psi}_j|^2 \geq 0$, termed the *edge-diffracted lobes*, redistribute energy across the scattering directions, while the oscillating terms $\text{Re}\,\widehat{\psi}_j\widehat{\psi}_l^\star$, the *interference factors*, shape the diffraction pattern and its fringes.

Observe that if $|\widehat{\psi}_j| = 0$, then $\widehat{\psi}_j\widehat{\psi}_l = 0$ for all l, i.e. the edge-diffracted lobes may act as proposal distributions for the interference factors. Our importance sampling strategy is then to sample from the (incoherent) superposition of the edge-diffracted lobes, ignoring the interference factors. The PDF of the proposal distribution is then

$$p\left(\vec{\xi}\right) \triangleq \frac{R^2}{\cos\theta}\frac{1}{\sum_j \widehat{\mathcal{P}}^{(\text{PL})}_j}\sum_j \left|\widehat{\psi}_j\right|^2, \tag{40}$$

which makes the likelihood ratio

$$\frac{f}{p} = \frac{\sum_j \widehat{\mathcal{P}}^{(\text{PL})}_j}{\mathcal{P}^{(\text{PL})}_{\mathcal{A}^\perp}}\frac{\widehat{w}\left(\vec{\xi}\right)}{\sum_j \left|\widehat{\psi}_j\right|^2}. \tag{41}$$

Sampling from an edge-diffracted lobe. An edge-diffracted lobe can be written as

$$\left|\widehat{\psi_j}\left(\vec{\xi}\right)\right|^2 = \left(1-\chi\left(\vec{\xi}\right)\right)\frac{|\Xi_j|^{-2}}{k^2R^2}\left[\left|a_j-b_j\right|^2\alpha_1\left(\vec{\zeta}\right)^2 + \left|\frac{a_j+b_j}{2}\right|^2\alpha_2\left(\vec{\zeta}\right)^2 + 2\left(b_j \operatorname{Im} a_j - a_j \operatorname{Im} b_j\right)\alpha_1\left(\vec{\zeta}\right)\alpha_2\left(\vec{\zeta}\right)\right], \tag{42}$$

where $\zeta = \Xi_j^{-\intercal}\vec{\xi}$. The above is an (incoherent) superposition of $(1-\chi)\alpha_1^2$ and $(1-\chi)\alpha_2^2$ lobes, as well as a cross term. As the cross term carries negligible power, we select one of the first two terms, with probabilities proportional to $|a_j-b_j|^2$ and $\frac{1}{4}|a_j+b_j|^2$, respectively, and draw samples from the $(1-\chi)\alpha_{1,2}^2$ lobes.

Analytically devising a sampling approach for $(1-\chi)\alpha_{1,2}^2$ is difficult, however, these functions are generic—do not depend on edge geometry or on light's properties—furthermore, these functions are "well-behaved", with no high-frequency details. Therefore, we precompute lookup tables for the inverse CDFs. Noting that these functions are also either even or odd in x and y, we only need to consider the quadrant $\{x, y \geq 0\}$. Two-dimensional tables, 4 megabytes in size each, are sufficient for accurate sampling. Importance sampling of the BSDF is summarised in Algorithm 3.

6 DISCUSSION

Our method reformulates the Fraunhofer theory of diffraction over polygonal apertures as a superposition of edge-diffracted waves, under illumination by an incident beam with a smoothly-varying intensity profile. To achieve closed-form formulae, we make the piecewise-linear approximation (PLA), where the incident beam's peak amplitude is linearly interpolated over the aperture mesh. The finer the mesh, the more accurate the approximation. Assuming a Gaussian incident beam, we show in our supplemental material (Fig. 3), that if the edges of the aperture mesh are at most one standard deviation (of the incident beam's Gaussian profile), then the PLA approximant is essentially exact. In practice, we tessellate on-the-fly the triangles that compose the aperture (this optional stage is done before line 8 in Algorithm 1) such that all edge lengths are sufficiently small.

Our formulations yield a BSDF—quantifying the angular distribution of diffracted energy—which, in contrast to other methods, requires no knowledge of the propagation distance after diffraction, requires no mutually-interfering rays nor propagating the high-frequency phase of light. In other words, unlike any other method, our formulation of free-space diffractions fits neatly into a typical, *linear* path tracing framework. The search radius (within which we look for triangles and build the diffracting aperture, viz. line 5 in Algorithm 1) is expected to be provided by a wave-optical light transport framework. These frameworks quantify for every beam the region over which wave-interference effects may arise (i.e., the region over which light remains optically coherent).

To keep our implementation simple, we do not use such frameworks. We implemented our method in Mitsuba 0.6 [Jakob 2010], and rendering was done with the BDPT (bi-directional path tracer) integrator. Our free-space diffractions BSDF is implemented as a BSDF plugin. Light beams are modelled as Gaussian beams, with a fixed (spatial) standard deviation $\sigma = 25\lambda$. The search radius of the free-space diffraction BSDF is set to $3\sigma = 75\lambda$. This is a rough upper bound on the expected optical coherence of light, given the applications and scenes we discuss and render. Our implementation is a proof-of-concept implementation: not meant to produce a ground truth, but demonstrate our method's ability to practically reproduce free-space diffractions, in complex scenes.

Results. Figs. 1, 7 and 8 were rendered using the implementation detailed above. Rendering resolution, sample-per-pixel (spp) count, and rendering runtimes are listed in the figures. All rendering was done on an AMD Ryzen™ Threadripper™ PRO 5975WX 32-Cores.

In Fig. 1a we render a small Cornell box illuminated by a collimated beam of visible-spectrum frequencies. The box is filled with a participating medium, which makes the incident beam of light markedly visible. A screen with an aperture cut in its centre is placed inside the box, giving rise to free-space diffractions. Notice the shift between the fringes that arise from the different spectral components of light, giving rise to a colourful pattern. A high sample count is required to integrate the volumetric scattering effects.

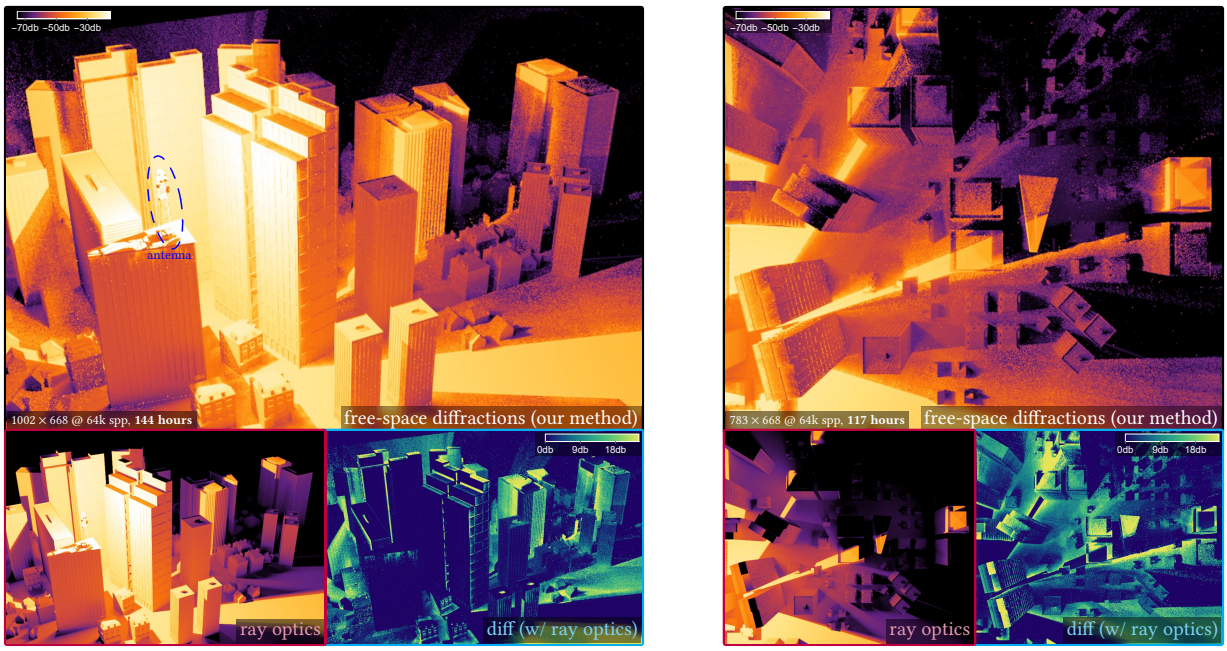
In Fig. 8 we render the Cornell box with various apertures cut in the screen. We use a spot light source, positioned at the center of the left wall. The spot light radiates into a wide cone, much wider than Fig. 1a, therefore convergence is slow as very few samples find the screen centre and the aperture.

Figs. 1b and 7 render a more complex, real-world scene: an urban environment with a variety of buildings, illuminated by radiation with 10 cm (cellular) wavelength. The scene was not optimized for long-wavelength rendering: meshes contain wavelength-scale edges and details. This scene serves to demonstrate the ability of our method to handle complex scenes, path traced at high resolution and with deep paths. We stress that this would be wholly hopeless with existing methods: for example, no existing UTD work simulates anything of such degree of complexity (due to the reasons outlined in Fig. 5). Choi et al. [2023, Fig. 6] (done with UTD) predicts a roughly similar difference between geometric optics and free-space diffractions, but in a far simpler scene.

Note that Figs. 1b, 7 and 8 are non-imaging renderings: the pixel values are the irradiance that impinges upon a visible surface, without propagation to the (virtual) observer.

Practical considerations. At every interaction point where free-space diffraction occurs (i.e., where the constructed aperture and its complement are non empty), we continue to path trace in the direction of the sampled diffraction direction ($\vec{\xi}$ in Algorithm 3). To avoid the exitant ray re-intersecting the diffracting mesh, we change the interaction point to a point inside the aperture $\mathcal{A}^\perp$ (done via rejection sampling). This is the only change required of the path tracing internals.

We presented an importance sampling strategy for our BSDF, summarised in Algorithm 3. In Section S3 in our supplemental material, we derive a formal quality metric for this importance sampling strategy (Eq. (S3.2)), and show that this strategy works best when the edge lengths e_j (of the diffracting edges that compose the boundary $\delta\mathcal{A}^\perp$) are all of similar lengths. The on-the-fly tessellation of the aperture mesh also serves to avoid very long edges. Very short edges could either be discarded (as they carry negligible energy: observe from Eq. (32) that the power contained in an edge-diffracted wave grows quadratically with edge length), or merged



(a) Simulation of signal coverage in a city

(b) Top-down aerial view of occluded area

Fig. 7. **Path tracing simulation of signal coverage.** (a) We simulate the propagation of cellular radiation ($\lambda = 10$ cm) in an urban scene, consisting of various buildings. The light source is placed on top of the highlighted antenna. (b) Also shown is a top-down view upon the region shadowed by the large buildings. Visualized is the colour-coded irradiance impinging upon the visible surfaces. The scene consists of 181 000 triangles, and the meshes were not optimized for long-wavelength rendering: they admit many small details and wavelength-scale edges, making the computations of free-space diffractions expensive. For comparison, displayed are the ray optics-only renderings. Observe the difference (compared with ray optics) insets: the long-wavelength radiation diffracts around the building edge's into the shadow regions, yielding a signal distribution that deviates sharply from the ray optics-only simulation. Also notice the multiple interactions (reflections and diffractions) of radiation with the scene—effects which are very difficult to simulate with existing methods.

into another edge, if possible. For simplicity, our implementation performs no special handling of short edges. This does not affect correctness, only performance.

Looking for triangles is expensive: rendering the city scene (Fig. 7) is roughly 50 times more expensive compared with the ray optics-only rendering. The penalty in other scenes is more modest. A significant amount of wasted cost lies in constructing our BSDF in situations where no diffractions occur (where the aperture is empty, e.g., a ray impinging upon a wall away from edges). A mechanism for giving up early on BSDF construction is left for future work.

Direct and diffracted terms. A unique aspect of our method is the formal decomposition of the diffracted field into the direct d and diffracted $\widehat{w}$ terms (Fig. 3), enabling light transport via ray tracing only. This decomposition is approximative: we effectively bias some of the energy that falls upon $\overline{\mathcal{A}}^{\perp}$, i.e. around the edges of geometry (for example, around the edges of the buildings in Fig. 7), into the diffracted lobes. In practice, once a ray has intersected with geometry and we have constructed the diffracting aperture, we select an interaction: (i) free-space diffraction; or (ii) interacting with the mesh's material (BRDF). In our implementation, we select the former with constant probability of 0.9 (chosen ad hoc): free-space diffractions admit multiple lobes and typically require many samples to integrate. To conserve energy, the mesh material's albedo must be at most the ratio $\tilde{P}_0^{(\mathrm{PL})}/\mathcal{P}_{\overline{\mathcal{A}}^{\perp}}^{(\mathrm{PL})}$: $\mathcal{P}_{\overline{\mathcal{A}}^{\perp}}^{(\mathrm{PL})}$ is the total power contained in the diffracted term (Eq. (31)), and $\tilde{P}_0^{(\mathrm{PL})}$ is the total power we remove from it (Eq. (36)). In our test scenes, the material's albedo is low enough that sufficient energy is absorbed naturally.

Time and space complexity. The time complexity of the BSDF construction (Algorithm 1) is linear in the number of triangles found within the search region around the interaction point. The time complexity of the evaluation (Algorithm 2) and importance sampling (Algorithm 3) of the BSDF are linear in the number of detected diffracting edges. As this number of edges is often a fraction of the number of triangles, sampling and evaluation are fast and the main cost is in the BSDF construction.

Only storage of the aperture's boundary edges is required, hence storage complexity is linear in the number of diffracting edges.

Accuracy. Figs. 4 and 6 serve as our primary validation: these figures show that our method produces essentially exact diffraction patterns compared with numerically-integrated Fraunhofer diffraction results. The only difference lies in the central lobe—the energy that propagates through the aperture without diffracting. We also validate against numerically-integrated Rayleigh-Sommerfeld (RS) diffraction integrals in Figs. 4 and 5c, as well as in Fig. 5 in our supplemental material. We also compare with a boundary element method (BEM) solver in Fig. 6 in our supplemental material.

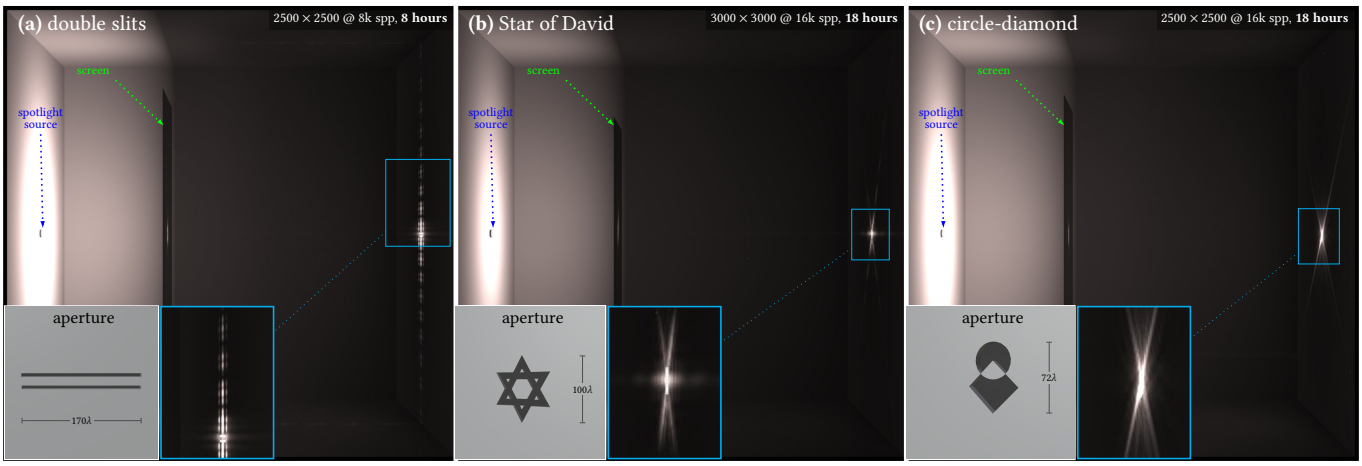


Fig. 8. **Diffraction by different apertures in a path tracing simulation.** A Cornell box (side length of $2 \times 10^4 \lambda$) with a light source placed at the centre of the left wall. A screen, with various apertures cut in its centre, is placed between the light and the right wall, producing different diffraction patterns.

In Fig. 5 we compare the diffraction pattern produced by our method with UTD-based ray tracing. The primary purpose of Fig. 5 is to demonstrate the poor behaviour of UTD at insufficient sample counts: a single ray interacting with an edge carries energy into the entire Keller cone around that edge (the UTD diffraction directions). In path tracing simulation, such rays produce high-frequency, high-intensity noise—a consequence of a non-linear rendering equation. Very many rays are required to sample the aperture, and shape the diffraction pattern via constructive/destructive-interference between these rays. Observe the 8spp patterns in Fig. 5: UTD produces noisy output over the entire pattern. Even for the simple double-slit aperture, UTD requires thousands of samples for converged results. Difficulties with UTD are further discussed in Subsection S4.1 in our supplemental material. Also see Fig. 5 in our supplemental material for comparison with UTD and an UTD failure case.

Finally, we also analyse the behaviour of a highly non-planar aperture: in Fig. 9 we plot the diffraction pattern produced by a modified double-slit aperture, where one slit is slanted and shifted out-of-plane compared with the other slit. Our method reproduces the small angular shift of the pattern, as also predicted by UTD, and matches UTD well for the first cluster of fringes.

Limitations. Applications of our methods need to be mindful of the following limitations:

(1) **Overlapping geometry**: Front-facing triangles that overlap on the virtual screen may lead to the detection of incorrect diffracting edges, producing hallucinated diffraction lobes. Detection of overlapping triangles can be expensive, and we do not do so in our implementation. For the vast majority of the geometry in our results this is not a concern, but should be addressed in future work.

(2) **Power in the central lobe**: The total powers contained in the edges $\widehat{\mathcal{P}}_j^{(\mathrm{PL})}$ (Eq. (32)), used for importance sampling, and the total power contained in the diffracted field $\mathcal{P}_{\overline{\mathcal{A}}^\perp}^{(\mathrm{PL})}$ (Eq. (31)), used to normalize the BSDF (Eq. (38)), are exact terms. On the other hand, the total power in the central lobe $\tilde{P}_0^{(\mathrm{PL})}$ (Eq. (35)) is an approximation. This term is used to compute how much energy we need to "lend" from the materials around the geometrical edges, in order to conserve energy, as discussed above. We analyze the error in this term in Fig. 10. We also present a more accurate, though more involved, approximation in Appendix B in our supplemental material.

(3) **Dielectric apertures**: The method is formulated under the assumption that the diffracting obstacle is a conductor, and fully absorbs radiation. Formulation of dielectric apertures is possible, and left for future work. It should be noted that any theory of diffraction (including UTD) requires special handling for dielectrics.

7 CONCLUSION

We presented a practical method for simulating free-space diffractions under the setting of path tracing. Our method requires no mutually-interfering rays, can be importance sampled, requires little modifications to a typical modern path tracer's internals, and is able to accurately reproduce diffraction fringes (compared with Fraunhofer diffraction). No preprocessing of the scene geometry is needed. Compared with the state-of-the-art, the presented method's ability to work with linear path tracers enables a significant increase in the efficiency of the simulation, and allows us to handle complex, real-world scenes.

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

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

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.

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.

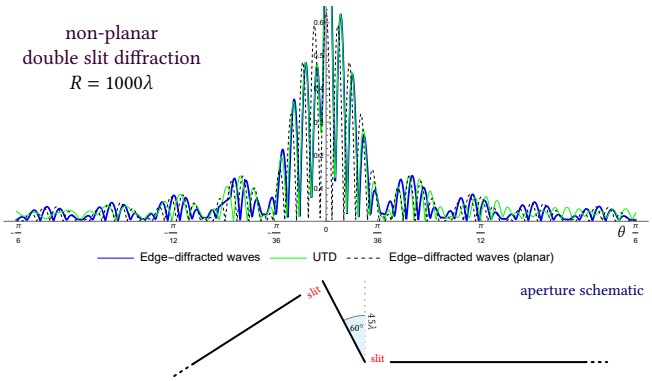


Fig. 9. **Diffraction by non-planar apertures.** The double-slit experiment, but with the left slit slanted at a $-30°$ angle, and additionally shifted about 45 wavelengths forward (in direction $\hat{z}$), such that the middle wall between the slits is slanted at a $60°$ angle, creating a highly non-planar aperture. We plot the square root of diffracted intensity, and compare our method (blue plot) with UTD (green plot). We also plot the planar double-slit experiment (as in Fig. 4) with our method (dashed black). Compared with the planar aperture, observe an overall rightward shift in the fringes.

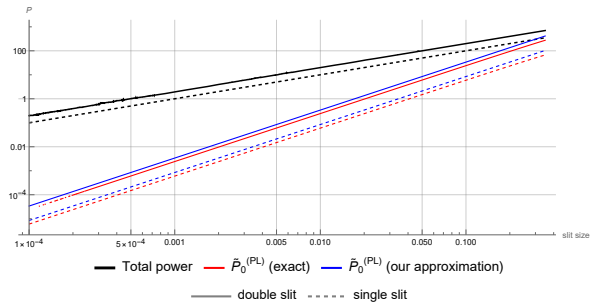


Fig. 10. **Analysis of our approximation for the central lobe total power.** We plot the total power contained in the diffracted field $|\widehat{\psi}|^2$ (black), numerically-integrated total power contained in the central lobe (red), and our approximation for the total power contained in the central lobe $\tilde{P}_0^{(PL)}$ (blue), as functions of slit width. Solid and dashed plots refer to the one-dimensional double-slit and single-slit, respectively, experiments (as in Fig. 4). As slit width increases power is increasingly concentrated in the central lobe, as expected. Sufficient agreement between the exact (red) and our approximation (blue) plots can be seen.

Max Born and Emil Wolf. 1999. *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light.* Cambridge University Press, Cambridge New York.

Mikel Celaya-Echarri, Leyre Azpilicueta, Peio Lopez-Iturri, Francisco Falcone, Manuel Garcia Sanchez, and Ana Vazquez Alejos. 2020. Validation of 3D simulation tool for radio channel modeling at 60 GHz: A meeting point for empirical and simulation-based models. *Measurement (Lond.)* 163, 108038 (Oct. 2020), 108038.

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).

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.

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

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.

J G Huang, J M Christian, and G S McDonald. 2006. Fresnel diffraction and fractal patterns from polygonal apertures. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 23, 11 (Nov. 2006), 2768–2774.

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 (Aug. 2020). https://doi.org/10.1145/3386569.3392094

Wenzel Jakob. 2010. Mitsuba renderer. http://www.mitsuba-renderer.org.

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/10.1163/156939306775777341

Jiří Komrska. 1982. Simple derivation of formulas for Fraunhofer diffraction at polygonal apertures. *J. Opt. Soc. Am.* 72, 10 (Oct. 1982), 1382.

Samir Mezouari and Andy Robert Harvey. 2003. Validity of Fresnel and Fraunhofer approximations in scalar diffraction. *J. Opt. Pure Appl. Opt.* 5, 4 (July 2003), S86–S91.

Marco Mout, Andreas Flesch, Michael Wick, Florian Bociort, Joerg Petschulat, and Paul Urbach. 2018. Ray-based method for simulating cascaded diffraction in high-numerical-aperture systems. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 35, 8 (Aug. 2018), 1356.

Robert Paknys. 2016. Uniform theory of diffraction. In *Applied Frequency-Domain Electromagnetics.* John Wiley & Sons, Ltd, Chichester, UK, Chapter 8, 268–316.

Louis Pisha, Siddharth Atre, John Burnett, and Shahrokh Yadegari. 2020. Approximate diffraction modeling for real-time sound propagation simulation. *J. Acoust. Soc. Am.* 148, 4 (Oct. 2020), 1922.

Iman Sadeghi, Adolfo Munoz, Philip Laven, Wojciech Jarosz, Francisco Seron, Diego Gutierrez, and Henrik Wann Jensen. 2012. Physically-based Simulation of Rainbows. *ACM Transactions on Graphics (Presented at SIGGRAPH)* 31, 1 (Feb. 2012), 3:1–3:12. https://doi.org/10/gfzndf

Carl Schissler, Gregor Mückl, and Paul Calamia. 2021. Fast diffraction pathfinding for dynamic sound propagation. *ACM Trans. Graph.* 40, 4 (Aug. 2021), 1–13.

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.

Shlomi Steinberg. 2019. Analytic Spectral Integration of Birefringence-Induced Iridescence. *Computer Graphics Forum* 38, 4 (Jul 2019), 97–110. https://doi.org/10.1111/cgf.13774

Shlomi Steinberg, Ravi Ramamoorthi, Benedikt Bitterli, Eugene d'Eon, Ling-Qi Yan, and Matt Pharr. 2023. A generalized ray formulation for wave-optics rendering. (2023).

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

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

Nicolas Tsingos, Thomas Funkhouser, Addy Ngan, and Ingrid Carlbom. 2001. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH01).* ACM. https://doi.org/10.1145/383259.383323

Fukun Wu and Changwen Zheng. 2016. *Interference Shader for Multilayer Films.* Springer International Publishing, 62–74. https://doi.org/10.1007/978-3-319-29971-6_4

Mengqi (Mandy) Xia, Bruce Walter, Eric Michielssen, David Bindel, and Steve Marschner. 2020. A wave optics based fiber scattering model. *ACM Transactions on Graphics* 39, 6 (Nov. 2020), 1–16. https://doi.org/10.1145/3414685.3417841

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).

Andrew Zangwill. 2013. *Modern electrodynamics.* Cambridge University Press, Cambridge.