From microfacets to participating media: A unified theory of light transport with stochastic geometry

DARIO SEYB, Dartmouth College, USA
EUGENE D’EON, NVIDIA, New Zealand
BENEDIKT BITTERLI, NVIDIA, USA
WOJCIECH JAROSZ, Dartmouth College, USA

Stochastic geometry models have enjoyed immense success in graphics for modeling interactions of light with complex phenomena such as participating media, rough surfaces, fibers, and more. Although each of these models operates on the same principle of replacing intricate geometry by a random process and deriving the average light transport across all instances thereof, they are each tailored to one specific application and are fundamentally distinct. Each type of stochastic geometry present in the scene is firmly encapsulated in its own appearance model, with its own statistics and light transport average, and no cross-talk between different models or deterministic and stochastic geometry is possible.

In this paper, we derive a theory of light transport on stochastic implicit surfaces, a geometry model capable of expressing deterministic geometry, microfacet surfaces, participating media, and an exciting new continuum in between containing aggregate appearance, non-classical media, and more. Our model naturally supports spatial correlations, missing from most existing stochastic models.

Our theory paves the way for tractable rendering of scenes in which all geometry is described by the same stochastic model, while leaving ample future work for developing efficient sampling and rendering algorithms.

CCS Concepts:
• Computing methodologies → Volumetric models.

Additional Key Words and Phrases: volumetric light transport, stochastic processes, implicit surfaces

ACM Reference Format:

1 INTRODUCTION
Computer graphics has seen significant progress over the past few decades, with advances in light transport, material fidelity, and
scene complexity greatly expanding the set of images that can be rendered. However, scene representations have evolved only slowly since their inception. Today, there is still a dichotomy that permeates through light transport theory where scene components are treated either as participating media or as hard (deterministic) surfaces. Beyond being intellectually unsatisfying, this distinction creates real problems in practice. Many rendering concepts invariably create effects that lie somewhere between surfaces and media. For example, inverse rendering and real-world scene reconstruction generate surfaces that contain some level of uncertainty due to measurement errors. Filtering and level-of-detail (LOD) of complex scenes turn hard surfaces into volumetric effects that still feature surface-like reflectance and correlations that are poorly represented by classical media. Classic scene representations can handle these effects only via specialized models or not at all, and recent efforts to unify these models [Dupuy et al. 2016] have not come to clear conclusions.

In this paper, we investigate a principled treatment of light transport in scenes containing stochastic implicit surfaces (SIS) [Gamito 2009]—a powerful modeling framework that represents surfaces as the zero crossings of a random field. In particular, we identify Gaussian processes (GPs) as a useful building block. By manipulating the mean, variance, and correlation of a Gaussian field, we can represent surfaces, media, and a new range of in-between effects including non-exponential media with arbitrary heterogeneity and microfacets with correlated microstructures.

Our key contribution is the introduction of new algorithms for computing light transport in correlated disorder, addressing a long-standing challenge in computer graphics and related fields. Stochastic environments, characterized by non-uniform distributions of geometry or particles, are found in numerous settings like tissue [Tuchin 2000], leaves [Fukshansky et al. 1993], oceans [Kirk 1975], reactors [Williams 1974], and molecular clouds [Boissé 1990]. Hence, our algorithms, primarily designed for computer graphics applications, also hold potential for fields like reactor engineering and remote sensing. Unlike traditional media, our stochastic surfaces create oriented intersections, permitting the use of conventional BSDFs for medium interactions, and can also delineate discrete phases of classical homogeneous material with stochastic boundaries.

Our rendering algorithms represent a unified and principled approach to simulating light transport in such complex scenes, but this currently comes at a cost—both in terms of computational resources, where they are more demanding than traditional graphics scene representations, and in terms of the availability of closed-form results. For example, we currently only support next-event estimation in a limited set of circumstances, preventing us from efficiently rendering scenes with sparse light sources when using mirror micro-BSDFs. Still, we believe that the method presented here will not only pave the way for novel applications in graphics but also extend its potential to scientific and industrial domains and we provide a reference implementation at https://github.com/daseyb/gpis-light-transport.

2 RELATED WORK

2.1 Stochastic Implicit Surfaces in Graphics and Beyond

Stochastic implicit surfaces have been used previously in graphics for terrain rendering [Gamito 2009] and procedural generation using e.g. Perlin noise [Ebert et al. 2003]. Here, a single realization of the surface is rendered, leading to purely solid surfaces. We consider the average light transport through all realizations, allowing us to express volumetric effects as well. Using a Gaussian process as a stochastic implicit surface has previously been used for surface reconstruction [Martens et al. 2017; Williams and Fitzgibbon 2006], where the GP is fitted to a set of point observations and the surface extracted from zero-crossings of the GP mean. (Co)variance in the GP has also been used to express reconstruction uncertainty to inform e.g. robot grasping [Dragiev et al. 2011]. Additionally, there has been a renewed interest in using GPs to quantify the uncertainty inherent in reconstructing surfaces from point clouds via Poisson surface reconstruction [Sellán and Jacobson 2022, 2023]. The posterior variance can then be used to e.g. optimize additional capture locations for maximum uncertainty reduction. Our method can use the retrieved posterior mean and covariance functions and faithfully incorporate the covariance during rendering, visualizing uncertainty as "fuzziness" in the surface.

2.2 Uncertainty Visualization

How to visualize uncertainty in volumetric datasets as "fuzziness" [Fout and Ma 2012] is a problem well explored in scientific data visualization and Brodie et al. [2012] provide a comprehensive review. Closest to our method, Pfaffelmoser et al. [2011] compute "isosurface-first-crossing probabilities" (IFCPs) in Gaussian fields, but for efficiency reasons, they limit themselves to Markov processes. We will show (Section 5.3) that while Markov processes make IFCPs tractable, they produce "non-smooth" processes which do not allow for scattering and hence cannot be used as geometry models when the aim is to compute global illumination.

2.3 Volumetric Scene Representations

Volumetric scene representations have received increased interest in recent years because they are easily differentiable and thus amenable to inverse rendering. Methods based on Neural Radiance Fields (NeRF) [Barron et al. 2021; Mildenhall et al. 2020; Müller et al. 2022] encode a classical (emissive) volume using a neural network and train it from images. More recent works [Fridovich-Keil et al. 2022] drop the neural component and represent the volume entirely with an octree. These methods have proven very effective in novel view rendering, but struggle at scene reconstruction because hard surfaces must be approximated with volume densities. Methods such as NeuS [Wang et al. 2021] or NeUDF [Liu et al. 2023] instead choose implicit surfaces as the underlying representation, where surfaces are defined by zero-crossings of a coordinate network. The hard surface prior leads to higher quality surface reconstructions for solid objects, but these methods fail to encode both surface priors and uncertainty in their representation, leading to poor results on scenes that contain both poorly- and well-resolved detail.
2.4 Transport in Discrete Stochastic Media

Similar to our method, techniques that compute transport through collections of packed objects [Moon et al. 2007] or granular media [Meng et al. 2015; Müller et al. 2016] consider the ensemble average light transport over many possible configurations of objects or grains. Unlike us, they pre-compute aggregate transport — averaging not only over different geometry configurations but also all possible light paths connecting two points. Reusing the same aggregate transport tables in different regions of their medium makes their pre-computation tractable, but introduces approximations and hence they often only opt to use it for higher-order scattering where error is less visible. We do not rely on tabulated distributions and instead focus on principled derivations of ensemble average light transport in a specific class of stochastic geometry. Discrete stochastic media have a rich history outside of graphics [Estrade et al. 2012; Lu and Torquato 1992; Torquato and Lu 1993], and Gaussian processes have previously been used to represent microstructures formed by binary mixtures [Jiao et al. 2007, 2008; Torquato 1986, 2005]. To build practical transport algorithms in such media, strong assumptions about the chord lengths along a transect are often made, such as a renewal [Roberts and Torquato 1999] or Markov [Pomraning 1991] assumption, which we avoid making.

2.5 Non-exponential Radiative Transport

Similar to methods for discrete stochastic media, these works assume that the photons experience free-path lengths given by a renewal process, where correlations “reset” after each bounce, and collisions on different segments of a path are assumed independent, which can lead to significant error [d’Eon 2023]. While algorithms like the chord-length sampling [Zimmerman and Adams 1991], conditional point sampling [Vu and Olson 2021], and 1D-point-process sampling [d’Eon 2023] have improved results by providing additional memory along a particle’s history, they are restricted to piecewise-stationary random systems and often make strong assumptions about the stochastic geometry. Our work transcend these limitations by leveraging the properties of Gaussian processes in the context of implicit surfaces. We show that our formalism can exhibit non-exponential attenuation as well but also allows for longer-reaching correlations across multiple segments or the entire path, achieving higher accuracy (but requiring higher cost). Additionally, work on non-exponential transport often only considers how correlation affects free-flights, while phase functions are assumed deterministic. In our method, appearance and transmittance are tightly coupled.

3 BACKGROUND & NOTATION

In the following, we will present the definitions and introduce any notation required for the derivations in Section 4. Readers familiar with the theory are welcome to skip individual subsections.

3.1 Gaussian Processes & Fields

A Gaussian process GP(μ, k) is a distribution over functions f : Ω → R such that for any finite set of locations x1, . . . , xN = X ⊆ Ω, the evaluations of the function follow an n-dimensional Gaussian distribution fX ∼ N(μ(X), k(X, X)). Here, we abuse notation slightly with μ(X) = [μ(x1), . . . , μ(xN)]T being an n-dimensional mean vector and k(X, X) an n × n covariance matrix, with entries k(x, x′) = k(x, x′). These are obtained by evaluating the mean function μ(x) : Ω → R and covariance kernel k(x, y) : Ω × Ω → R for each location or pair of locations in X, respectively. When the domain Ω is a subset of Rd, we call the resulting Gaussian process a Gaussian field [Adler and Taylor 2009]. In this work, we will only investigate 3D Gaussian fields, but for other applications, for example, in machine learning, d can be very large. For a more thorough introduction to Gaussian processes we recommend Williams and Rasmussen [2006], whose notation we follow in this section.

3.1.1 Restrictions to sub-domains. We can restrict the input domain of a Gaussian process (e.g. from 3D space R3 to the points xR = {x + tw|t ∈ R} along a line) with no change in the statistics. That is, if f ∼ GP(μ, k)R3 and g ∼ GP(μ, k)xR, then f(xR) ∼ g(xR), i.e. they are equal in distribution. This is a useful property because we will often look at lower dimensional “slices” of a Gaussian field, and we can do so without having to change the mean or covariance kernel.

3.1.2 Kernel Functions. Covariance kernels can take many forms and are the primary “design variable” for Gaussian processes. In Fig. 2 (left) we show realizations produced by a set of common kernels. Intuitively, the kernel prescribes how “similar” values are at nearby points in space, and small changes to the kernel can radically change the properties, such as differentiability, of the resulting sample functions. The only restriction on kernels is that they must be positive semi-definite. Kernels composed via multiplication and addition are again valid kernels, which allows the design of complex kernels out of simple building blocks.

We call a kernel stationary if k(x, y) = k(y − x), i.e. the kernel does not change shape depending on where we evaluate it. A further subset are isotropic kernels, k(y − x) = k(||y − x||), which only depend on the distance between points. More general, non-stationary kernels for which k(x, y) ≠ k(y − x) are not widely discussed in Gaussian process literature since their theoretical analysis is difficult. That said, the statistics of real-world datasets are very rarely truly stationary [Noack and Sethian 2022], and we will be using a

Fig. 2. Gaussian processes provide a very flexible way to define sets of functions (each being a realizations of the process), which can have wildly different properties based on the covariance kernel of the process. On the left we draw realizations from two prior (i.e. unconditioned) processes with different covariance kernels. On the right, we show realizations from the corresponding posterior process, conditioned on the two black points (dashed lines are the posterior means of the same hedge kernels).
non-stationary kernel to allow for spatially varying geometry and material properties such as surface roughness.

3.1.3 Conditioned Processes. Gaussian processes can be conditioned on observations, and the conditioned posterior process is again Gaussian. For example, given measurements \( m \) at locations \( C \), we can derive the conditioned GP \( \mathcal{GP}(\mu(x), k(x,y)) \) with the condition \( \xi_m \sim f(C) = m \) as

\[
\begin{align*}
    f & \sim \mathcal{GP}(\mu_{\xi_m}(x), k_{\xi_m}(x,y)) , \quad \text{with} \\
    \mu_{\xi_m}(x) & = \mu(x) + k(x,C)k(C,C)^{-1}(m - \mu(C)) \\
    k_{\xi_m}(x,y) & = k(x,x) - k(x,C)k(C,C)^{-1}k(C,x)
\end{align*}
\]

Sampling from the conditioned GP then produces realizations that pass through the values \( m \), but interpolate according to the chosen covariance kernel in-between as shown in Fig. 2 (right). Unfortunately, computing the posterior (i.e. conditioned) mean and covariance involves inverting the \( n \times n \) matrix \( k(C,C) \), where \( n = |C| \). This has \( O(n^3) \) time complexity and becomes intractable for large \( n \).

3.1.4 Sampling from Gaussian Processes. Efficiently generating large numbers of correlated samples from a Gaussian process also poses a challenge. Samples can be drawn directly from \( \mathcal{GP}(\mu, k) \) via

\[
f(X) = \mu(X) + k(X,X)^{1/2} \eta,
\]

where \( \eta \sim \mathcal{N}(0,1) \) and \( A^{1/2} \) is the matrix square root s.t. \( A = A^{1/2}A^{1/2} \). In either case, the time complexity is \( O(p^2) \) with \( p = |X| \), and drawing correlated samples from a Gaussian process quickly becomes intractable for large \( p \). This method of directly drawing values of a sample function is called the function-space view of Gaussian processes [Williams and Rasmussen 2006]. There are other methods of simulating Gaussian processes, such as the weight-space view [Rahimi and Recht 2007], but these are usually limited in the types of processes they support (e.g. only ones with stationary kernels). Still, the weight-space view will be valuable when validating our method, even though we cannot use it in more complex applications.

3.1.5 Derivatives of Gaussian Processes and ”Multi-Task” GPs. Due to the linearity of the derivative operator, the derivative of a Gaussian process is again a Gaussian process and takes the form

\[
\mathcal{GP}'(\mu(x), k(x,y)) = \mathcal{GP}(\mu'(x), k_{x,y}(x,y))
\]

where \( k_{x,y}(x,y) = \frac{\partial^2 k(x,y)}{\partial x \partial y} \). Additionally, we can reason about the joint value-derivative distribution

\[
\left[ \begin{array}{c}
    f(X) \\ f'(Y)
\end{array} \right] \sim \mathcal{N} \left( \left[ \begin{array}{c}
    \mu(X) \\ \mu'(Y)
\end{array} \right], \left[ \begin{array}{cc}
    k_{x,y}(X,Y) & k_{x,y}(X,Y) \\
    k_{x,y}(Y,X) & k_{x,y}(Y,Y)
\end{array} \right] \right).
\]

This will come in useful when we reason about distributions of normals on Gaussian process implicit surfaces in Section 4, since the normals of a GPIS are aligned with the spatial derivative of the underlying GP.

3.2 Implicit Surfaces

An implicit surface is defined as the level-set of a scalar function \( f : \Omega \to \mathbb{R}, \{ x \in \Omega \mid f(x) = l \} \). Without loss of generality, we consider the zero level-set, where \( l = 0 \). In addition to the surface, we can partition \( \Omega \) into \( \{ f(x) < 0 \} \) and \( \{ f(x) > 0 \} \) regions. Computing the intersection of a ray with an implicit surface can be done via root-finding, i.e. finding distance \( s \) along ray \( x + s \omega \) such that \( f(x_s) = 0 \), where \( x_s = x + s \omega \), i.e. a point which is a distance \( s \) along a ray \( (x, \omega) \). Many root-finding methods are available, depending on which properties of \( f \) are known [Hart 1996; Stol and De Figueiredo 1997]. The intersection point \( x_s \) with \( s = \arg \min_{t \in \mathbb{R}_+} f(x_t) = 0 \) always depends on a particular ray and a particular implicit surface \( f \). When we want to make this clear, we indicate the dependence on the surface as \( x_s^f \). In addition to intersections, we are often interested in the normal \( n_s \) of the surface at a point \( x_s \). Implicit surfaces make it easy to calculate the normal vector as \( n_s = \nabla f(x_s) \), where \( \nabla \) computes the normalized gradient \( \nabla f(x_s)/\|\nabla f(x_s)\| \). Just as before, we sometimes will highlight the dependence of the normal at the intersection on a particular implicit surface as \( n_s^f \).

3.2.1 Stochastic Implicit Surfaces. A stochastic implicit surface (SIS) is the distribution of level sets of a stochastic process. In this paper, our main focus is on Gaussian process implicit surfaces (GPIS), where the stochastic process is Gaussian. Each realization \( f \sim \mathcal{GP} \) of the process generates an implicit surface at the zero level set \( f(x) = 0 \).
Fig. 5. We show the difference between the “naive” method in Eq. (8) (top), which requires global realization sampling, and the progressive sampling method we present in Eq. (14), which only samples realizations along path segments as needed (bottom). Both compute the ground truth ensemble average light transport in a given GPIS (μ shown on the left). In the naive method, tracing is simple since each path sees a deterministic f (red: inside, blue: outside, black: zero-level set), but sampling large numbers of correlated samples from a general Gaussian process is prohibitively expensive even for modest resolutions (scaling with $O(n^9)$ where $n$ is the discretization resolution). In our method, we only ever sample 1D slices of the process, which scales with $O(m^3)$, where $m$ is the number of steps taken along each ray. While still expensive compared to classical rendering techniques, it makes it feasible, for the first time, to compute GPIS light transport in a principled manner.

In Fig. 3, we show how the deterministic intersection distance and normal in implicit surface rendering turn into distributions when considering SISes.

3.3 Light Transport

We write the surface rendering equation [Kajiy 1986] for light transport in a scene defined by the implicit surface $f$ as

$$L^f(x, \omega) = \int_{S^2} \rho(x) L^f(x', \omega_x) \, d\omega_x$$

(7)

where $x'$ and $n'$ are defined in Section 3.2 and, for brevity, we use the shorthand $\rho(x) = \rho(x_0, x_1, n_0 | n_1, \omega_1 | x_0, n_0, \omega_0)$ for the cosine-weighted BSDF of the implicit surface, which we assume to be a deterministic function of position $x_0$, normal $n_0$ and incoming and outgoing directions $\omega_0$, $\omega_1$ respectively. We illustrate this form of the rendering equation in Fig. 4.

4 ENSEMBLE-AVERAGED LIGHT TRANSPORT IN GPIS

The radiance $L^f$ introduced in the previous section describes light transport for a fixed implicit surface $f$ (Eq. (7)). We now consider the case when $f$ is a realization of a Gaussian process (a GPIS), and will derive methods for computing the average transport over all realizations. Formally, we investigate the ensemble averaged light transport over all realizations $f$ of the Gaussian process $\mathbf{GP}(\mu, k | \zeta)$.

$$\langle L^f(x, \omega) \rangle_{\zeta} = \int_{\mathbf{GP}} L^f(x, \omega) \, d\nu_{\mu,k}(f | \zeta),$$

(8)

where $\nu_{\mu,k}(f | \zeta)$ is the classical Wiener measure [Taylor 2006, Ch. 16] of $f$ with respect to the conditioned GP $(\mu, k | \zeta)$, i.e. the probability density of sampling $f \sim \mathbf{GP}(\mu, k | \zeta)$. We will use the condition $\zeta$ as an “editing” tool in Section 6, since it allows us to only consider realizations that, e.g., pass through a certain point in space. For brevity, the dependency on $\mu, k$ is made implicit going forward. We can form a straightforward Monte Carlo of Eq. (8) via

$$\langle L^f(x, \omega) \rangle_{\zeta} = \frac{1}{N} \sum_{j=1}^{N} \langle L^f_j(x, \omega) \rangle, \quad f_j \sim \mathbf{GP}(\mu, k | \zeta),$$

(9)

where we simply average the light transport across $N$ independent realizations of the Gaussian process. Within each realization, $L^f$ may in turn be estimated with Monte Carlo using e.g. path tracing.

While Eq. (9) produces correct results on average, it is wholly impractical: For each sample, an entire 3D realization $f_j$ must be constructed. If the GPIS is discretized into a volume of sidelength $O(n)$, constructing each realization takes $O(n^9)$ time, making it infeasible. In the following we will derive an alternative form of Eq. (8) that interweaves the Monte Carlo rendering process with the random sampling of realizations. As a result, we will be able to construct realizations in a “just in time” fashion, only sampling the parts of $f$ that are required to continue the light transport path.

Explicit delta free-flights in fixed $f$. As a first step we make it explicit that free-flight distributions are delta functions in fixed implicit surfaces, rewriting Eq. (7) as

$$L^f(x, \omega) = \int_0^\infty \int_{S^2} \rho(x) \delta(f(x, n)) \mathcal{I}(0, t) \, L^f(x, \omega_t) \, d\omega_t \, dn, dt,$$

(10)

where $\delta(f(x, n)) = \delta(f(x) - 0) \cdot \delta(\mathcal{I}(x, n) - n)$ and

$$\mathcal{I}(0, t) = \begin{cases} 1 & \forall x \in (0, t) : f(x) > 0 \\ 0 & \text{otherwise} \end{cases}.$$
distance, normal, and exit direction allows us to rewrite Eq. (8) as

$$
\langle L^f (x, \omega) \rangle_\xi = \int_0^\infty \int_{S^2} \rho(x_t) \left( I_0^f (x_t, n) \right) \left( \int_0^{\infty} \int_{S^2} \rho(x_t) \gamma_{x_t} (0, n | \xi) \left( \int_0^{\infty} \int_{S^2} \rho(x_t) \gamma_{x_t} (0, n | \xi) \right) \right) \text{d}x_t \text{d}n. \quad (12)
$$

So far, we have gained little — the ensemble average on the right-hand side is still over full 3D realizations $f$.

From delta functions to conditioned ensembles. Using a relation we derive in Section 1.1.2 of the supplemental, we can pull the delta function out of the ensemble average — i.e. instead of “filtering” realizations after sampling, we only average over realizations having the required normal and zero-crossing in the first place, giving us

$$
\langle L^f (x, \omega) \rangle_\xi = \int_0^\infty \int_{S^2} \rho(x_t) \gamma_{x_t} (0, n | \xi) \left( \right) \text{d}x_t \text{d}n, \quad (13)
$$

where the condition $\xi = (f(x_t) = 0 \land \nabla f(x_t) = n)$ restricts the realizations inside the ensemble average to the ones where the delta function is non-zero and $\gamma_{x_t} (0, n | \xi)$ accounts for the density of sampling such realizations. This already constricts the space of realizations we need to consider in the ensemble average on the right-hand side as we move along the path.

Splitting realizations and progressive sampling. Finally, we note that $I_0^f (x_t, n)$ only depends on the 1D restriction of $f$ to the ray segment $(x, x_t)$. Gaussian processes, being closed under conditioning, allow us to decompose sampling $f$ into two parts: Sampling the values $f_{x_t}$ along the ray segment and then sampling $f$ over the remainder of the domain. As long as we condition the second step on $(x, x_t)$, i.e. the results of the first one, the final distribution is the same as if we had sampled all parts of the domain at the same time. Again, using a relation shown in Section 1.1.3 of the supplemental we split the ensemble average on the right-hand side in two, giving us

$$
\langle L^f (x, \omega) \rangle_\xi = \int_0^\infty \int_{S^2} \rho(x_t) \gamma_{x_t} (0, n | \xi) \left( \int_0^{\infty} \int_{S^2} \rho(x_t) \gamma_{x_t} (0, n | \xi) \right) \text{d}x_t \text{d}n, \quad (14)
$$

where $\langle \rangle_{(x, x_t)}$ represents the conditioned average over realization restricted to a path segment. Eq. (14) is vastly more efficient to estimate, as at each level of recursion, realizations only need to be evaluated on path segments $(x, x_t)$ and not the entire scene, reducing the cost to $O(m^3)$, where $m$ is now the number of steps taken along a ray segment (see Fig. 5). Ensemble averaging is done for each path segment individually; however, Eq. (14) is exact, and no approximations are made. This is made possible because each subsequent segment on the path is conditioned on the observed values of the realization seen on all prior segments, ensuring a globally consistent view. However, the cost of sampling realizations on later segments increases as more and more conditions are added.

4.1 Memory Models

While sampling from a 1D process (14) is wildly more practical than sampling global 3D realizations $(8)$, conditioning subsequent path segments on all values sampled along prior segments is still computationally intensive. To enable a more practical implementation, we introduce the concept of memory models, which “forget” some of the conditioning on prior segments in Eq. (14). This trades off computational cost for approximation error. Effectively, we replace the list of conditions $\alpha = \xi \land \gamma_{x_t} (0, n | \xi)$ on the recursive ensemble average in Eq. (14) by a simpler form. Figure 6 shows a didactic visualization of the different models we present in this section. Many more are possible, but we restrict ourselves to the most interesting subset.

4.1.1 The Global-$n$ memory model. The Global-$n$ model conditions each segment on only the $n$ prior segments $\zeta^{n-1} := \zeta^{k-n} \land \ldots \land \zeta^{k}$, where each $\zeta^k$ is $\xi \land \gamma_{x_t} (0, n | \xi)$ of Eq. (14) for the $k$-th path segment. In the limit, for $n = \infty$, the ground truth model is recovered. While still expensive, this model allows setting an upper bound on the evaluation cost of Eq. (14) that does not grow arbitrarily with the length of the light path.

4.1.2 The Renewal memory model. On the other end of the spectrum, the Renewal model retains only minimal state, i.e. that a surface must be present at $x_t$ on adjacent segments, meaning $\zeta^k (f) := $
whether keeping additional vertices, similar to the Global-n model, we condition on.

ALGORITHM 1: A renewal approximation in volumes with stochastic density can introduce significant errors. Our Renewal+ model aims to reduce this error by additionally remembering the previous normal \( \mathbf{n}_k \) at \( x_k \), i.e. we condition on \( \gamma'_k(f) := \gamma'_k(f) \wedge (\nabla f(x_k) = \mathbf{n}_k) \). This enforces gradient consistency across segment realizations, which reduces approximation error compared to the Renewal model significantly (see Fig. 7) while being much more tractable than the Global model. Conditioning on higher order derivatives may be interesting to explore in future work.

Note that both the Renewal and Renewal+ memory models only retain information about the last collision. We leave the question of whether keeping additional vertices, similar to the Global-n model, would improve results for future work, since these and other memory models are trivial to experiment with.

4.2 Progressive sampling via function-space GPs

To round out this section, we present a method for rendering GPIS light transport that we will use to produce subsequent images and other results. This method supports fully non-stationary covariance models that we will use to produce subsequent images and other results. This method supports fully non-stationary covariance models (see Section 6) and any choice of memory model. For an alternative weight-space implementation that supports only stationary kernels but can serve as a “ground truth”, see Section 3 of the supplemental.

We can write a one-sample Monte Carlo estimator of Eq. (14) as

\[
\langle L_i(x^n, \omega) \rangle_\xi = \frac{\rho(t, n, \omega_t, f_{(x_k, x)})}{\rho(t, n, \omega_t, f_{(x_k, x)})} \langle L_i(x^n, \omega_t) \rangle_{\xi'},
\]

\[
t, n, \omega_t, f_{(x_k, x)} \sim \rho(t, n, \omega_t, f_{(x_k, x)}).
\]

**ALGORITHM 1:** Simple light transport in a GPIS

```python
def L(t, n, \xi) -> color:
    (t, n, \xi') -> nextHit(t, \xi)
    return L(t, n, \xi') + L(t, n, \xi)

def L(t, n, \xi, \omega) -> color:
    (\omega', p_{\omega'}) -> sampleDirection(t, \xi, \omega)
    \rho -> evalBSDF(f(t), \omega, \omega', \xi)
    return \rho \cdot (n(t, \xi, \omega'), \omega') / p_{\omega'} \cdot L(Ray(f(t), \omega', \xi), \xi)

def nextHit(t, n, \xi) -> color:
    \( f \leftarrow \text{sampleRealization}(\xi) \) //Sample a 1D realization along the ray.
    \( t \leftarrow \arg\min_s f(s) \leq 0 \) //Find first zero crossing.
    \( n \leftarrow \text{sampleNormal}(\xi, \xi \wedge f) \) //Sample normal at intersection point.
    \( \xi' \leftarrow \text{M}(\xi \wedge f \wedge n) \) //Filter state based on memory model.
    return (t, n, \xi')
```

Fig. 8. We first determine a conservative envelope around the GPIS based on the point-wise probability of being inside the surface. We then place \( n \) (here \( n = 4 \)) sample locations along the ray segment at fixed distances \( \Delta_s \) based on the covariance kernel. If the segment is longer than \( n \cdot \Delta_s \), we stop the process, effectively inserting a “null-collision” (grey outline), and continue using a new set of conditioned samples. Once an interpolation of the discrete samples crosses \( 0 \) we have found the intersection point \( x_1 \) (black outline). We then sample a normal \( n \) and a direction from the micro BSDF and continue in the scattered direction \( \omega_t \).

where, \( \rho(t, n, \omega_t, f_{(x_k, x)}) \) is a joint probability density of sampling intersection distance \( t \), normal at the intersection point \( n \), outgoing direction \( \omega_t \) and a 1D realization \( f_{(x_k, x)} \) that has positive values along the line segment \( (x, x_k) \) and a zero crossing at \( t \). The question now is whether we can choose \( \rho \sim \rho(t, n, \xi') \). This would both reduce variance, due to importance sampling part of the integrand, and also save us from having to compute \( \Gamma(t, n, \xi) \), for which we don’t have an analytic expression. To achieve this, we carefully consider the order of sampling the quantities we need as follows

1. Sample a 1D realization along the ray \( f_{(x_k, x)} \sim G\hat{P}(x_k, x)|\xi \).
2. Find the intersection distance \( t = \arg\min_{\xi} \xi \in (0, \infty) f_{(x_k, x)} \).
3. Sample a normal \( n \sim G\hat{P}_{(x)}(\xi) \).
4. Sample the BSDF \( \omega_t \sim \rho(x_k) \).

This perfectly importance samples the factor in front of the recursive term in Eq. (15). Afterward, we can recursively compute \( \langle L_i(x_k, \omega_t) \rangle_{\xi'} \) in exactly the same way, and we continue until we hit a light source or exit the scene. These steps are reflected in Alg. 1.

4.2.1 Practical considerations. Still, the list of steps does not tell the whole story. The main issue lies in sampling \( f_{(x_k, x)} \) given that \( f \) is a function of \( x \), i.e. an uncountably infinite set of values. Using function space sampling, we can only sample a finite set of \( m \) values, and the cost grows with \( \mathcal{O}(m^2) \). Hence, we have to approximate \( f_{(x_k, x)} \) somehow using a finite (and hopefully small) number of samples. Luckily, the processes we consider are smooth, after all we require continuity and second-order differentiability. This means that sample functions should be well approximated by interpolation of a relatively small number of samples. The second issue is that \( f_{(x_k, x)} \) is defined over the positive real number line, i.e. it contains all points along the ray to infinity. Hence when discretizing the number line, we would still need to sample an infinite number of values. We solve this with a two-layer strategy: First, we derive appropriate scene bounds. Then, we split the ray into segments with a fixed number of sample points, spaced appropriately for the given covariance function, and progressively sample the realization for each segment, much like we progressively sample realizations between path vertices. This process is visualized in Fig. 8 and we
provide details on each individual substep in Section 3.2 of the supplemental.

Note that despite the seemingly intricate procedure described in this section, our method is not difficult to implement, assuming one has access to a path tracer and methods to sample conditioned values from a Gaussian process—which are available in many libraries and easy to implement from scratch. Notably, we did not have to change any of the core integrator code—a GPIS is implemented as a medium with a special phase function and free-flight distance sampling routine; the data for mean and covariance is accessed through the same volume data API that is used for heterogeneous density in classical media; we expand our renderer’s existing architecture for non-exponential transport to keep track of the path history.

5 THE APPEARANCE SPACE OF GPISES

Now that we have a theory that allows for representing and rendering GPIS light transport, the remaining question is how to obtain such GPISES in the first place.

Fig. 10. At the top we show example realizations of Gaussian processes with different covariance kernels, all having the same microfacet roughness $\alpha$. On the bottom are renders of the ensemble average light transport. Our method makes it trivial to explore light transport in a range of microgeometry without having to generate, store, and process large amounts of data.

Much like traditional scene representations, a GPIS can be manually authored by an artist, either by explicitly specifying its parameters (Section 6.2.1) or implicitly by constraining the GPIS at select points (Section 6.2.2). GPISES may also be combined via CSG operations (Section 6.2.3). The novel degrees of freedom in a GPIS allow to simultaneously specify a surface as well as its “fuzziness”/“certainty”, unlocking a new spectrum of authoring workflows.

Beyond artistic control, we show how to render GPISES with parameters derived from data, for example via stochastic Poisson surface reconstruction (Section 6.3). Because GPISES can express correlations at the micro- and mesoscale, they are also a useful tool for level-of-detail, and we show how to obtain GPISES that represent prefiltered implicit surfaces at any desired resolution (Section 6.4).

Inherent to GPISES is the averaging over stochastic microgeometry, and as such they naturally subsume many classical stochastic appearance models. In the following subsections, we focus on two in particular that have seen much popularity: Microfacet surfaces, and participating media. The microgeometry assumed by either of these models can be naturally represented by a GPIS (see Fig. 9). However, because GPISES make fewer approximating assumptions, we can actually achieve more accurate renderings of the appearance described by these classical models, by incorporating correlations along paths through the microgeometry. Beyond making connections to existing appearance models, our method allows us to easily experiment with new microgeometry, and it provides a ground truth against which analytic approximations may be compared (see Fig. 10).

5.1 Joint distribution of free-flight distances and normals

The appearance of a microfacet surface is characterized by a normal distribution, a microfacet BSDF, and a shadowing-and-masking term. Similarly, a participating medium is described by its phase function and the transmittance function (or equivalently, its free-flight distribution). Though obscured by the rigorous treatment of correlations, these quantities also appear in our recursive GPIS light transport formulation, which we make explicit in the following.
112.9 From microfacets to participating media: A unified theory of light transport with stochastic geometry

Fig. 11. An appropriately chosen surface-type GPIS can match microfacet models. We show renderings of a scene lit by an environment map, containing a surface-type GPIS with mirror microfacets (left) and a Beckmann microfacet model (right) with roughness matched via Eq. (20). The insets show the vNDF for a ray hitting the surface at a grazing angle, as computed by our method and derived from microfacet theory. For the relatively smooth surface shown here, the Smith approximation is close to the ground truth computed with our method as expected [Bourlier et al. 2000].

Under the Renewal+ and Renewal models, Eq. (14) can be simplified further, as \( \langle L(x,t) \rangle_{ξ, ζ'} \) no longer depends on a specific realization \( f(x,t) \) along the incoming ray segment. This yields

\[
\langle L(x,t) \rangle_{ξ, ζ'} \approx \int_0^\infty \int_0^\infty \rho(x,t) T(t,n|ξ) \langle L(x,t,ω) \rangle_{ξ, ζ'} \, dω \, dt,
\]

(17)

with

\[
T(t,n|ξ) = \int_{\mathcal{GP}(x,t)} 1_{f(x,t)|ξ} \, dy \, (f(x,t)|ξ) \wedge ζ,
\]

(18)

being the probability of sampling any realization \( f(x,t)|ξ > 0 \) conditioned on \( ξ \wedge f(x,t) = 0 \) and \( f(x,t) = n \). We provide a step-by-step derivation in Section 1.3 of the supplemental.

The “GPIS density” \( T(t,n|ξ) \) allows us to connect our method to participating media and microfacet surfaces, and derive parameters for our model to match their appearance.

5.2 Surface-type GPISes

Microfacet surfaces, much like GPISes, describe surfaces as realizations of a stochastic process—in fact, the popular Beckmann model [Beckmann 1965] even assumes Gaussian process height fields. A crucial microsurface property is the distribution of visible normals (vNDF) \( D_0(n|ω) \), which describes the distribution of surface normals visible from direction \( ω \). We can derive a vNDF directly from the GPIS density as

\[
D_0(n|ω) = \int_0^\infty T(t,n|ω,ξ) \, dt.
\]

(19)

This gives the vNDF for any GPIS. To match microfacet theory, we can explicitly induce a heightfield surface via a linearly varying mean along the z-axis \( μ(p) = μ(p_z) \) and a strongly anisotropic kernel \( k(p,q) = k(p_x,p_y,q_x,q_y) \) that induces complete correlation along the z-axis. What remains is how to determine the parameters of the covariance kernel to match a given microfacet model.

Classical microfacet theory disregards certain correlations [Smith 1967], making the vNDF only depend on \( α \), the “roughness” of the surface, which we can derive for any kernel [Pharr et al. 2016] as

\[
α = \sqrt{-2k''(0)}.
\]

(20)

As long as the kernel is twice differentiable, we can easily find a GPIS that matches a given microfacet surface by (1) aligning the gradient of the mean with the macro surface normal, and (2) picking a kernel with perfect correlation along the macro surface normal and \( k''(0) = −α^2/2 \). In Fig. 11, we verify that our method can produce results consistent with microfacet theory by rendering a heightfield GPIS matched to a classical microfacet surface via Eq. (20). Note that without the Smith approximation to shadowing and masking, the vNDF depends on the full shape of the covariance kernel. In our method, we do not apply this approximation and we can show (Fig. 12) that, while the Smith approximation is excellent in some scenarios, for grazing angles and certain kernels, taking correlations into account does matter.

Fig. 12. We show the difference between the vNDF computed by our method and the vNDF predicted for a Beckmann surface with Smith shadowing and masking at various incident ray angles (left to right) and for different covariance kernels (top vs. bottom) for a relatively rough surface. As expected [Bourlier et al. 2000], at non-grazing angles, the Smith approximation is excellent, independent of the covariance kernel, whereas at grazing angles non-integrable covariance kernels (such as the rational quadratic kernel shown here) induce significantly different vNDFs.

Fig. 13. As we move from the heightfields produced by strongly anisotropic covariance kernels (left) towards isotropic ones (right), the NDF (inset, blue) goes from the one predicted for Beckmann surfaces (dashed orange) towards a von-Mises-Fisher distribution, as discussed by d’Eon [2021].

ACM Trans. Graph., Vol. 43, No. 4, Article 112. Publication date: July 2024.
The central quantity in volumetric transport is the free-flight distribution over the sphere of all possible normals at the intersection point \( x \):

\[
\Gamma(t \mid \zeta) = \int_{S^2} \Gamma(t, \mathbf{n} \mid \zeta) \, d\mathbf{n}.
\]  

(21)

\( \Gamma(t \mid \zeta) \) represents the probability density of finding the first zero crossing of the GPIS at distance \( t \). Assuming the Renewal model, it is sufficient to examine the case when \( \zeta_c := f(x) = 0 \) (i.e. the current ray starts on a zero crossing, e.g. after leaving a scattering event) and \( \zeta_u := f(x) > 0 \) (i.e. the ray origin is uncorrelated with the GPIS e.g. starting at the camera or leaving a non-GPIS surface). This is analogous to recent work in non-exponential transport [Bitterli et al. 2018], and the same properties apply: \( \Gamma(t \mid \zeta_c) \) can be derived from \( \Gamma(t \mid \zeta_u) \) and vice-versa (see Fig. 15).

### 5.3.1 Matching participating media

Just as in microfacet theory, disregarding correlations along light paths allows us to describe a medium by a single number, its density \( \sigma_t \), i.e. a value indicating how many particles occupy a certain region of space. As before, we want to derive the parameters, i.e. \( \sigma_t \), of this simplified classical model from the parameters of a GPIS and vice versa. Since volume-type GPISes have not been studied much in graphics literature, this requires us to pull in knowledge from a wider range of literature on stochastic processes. There, the free-flight distributions \( \Gamma(t \mid \zeta_u/c) \) are known as the first-passage-time (for \( \zeta_u \)) and the excursion-time density (for \( \zeta_c \)). The first-passage-time is the “time” (in our case distance) at which a realization of a stochastic process first crosses a particular level. The excursion time is then the time that a realization of a stochastic process takes to return to the given level for the first time. This corresponds to the distance from one intersection point to another. In the Gaussian process case, it is known that excursion times are well-defined only for smooth processes [Bray et al. 2013; Torquato and Lu 1993]. A process is smooth if the Taylor expansion of \( k(x, y) \) has the form \( k(x, y) = 1-a\|x-y\|^a+O(\|x-y\|^{a''}) \), with \( a = 2 \) [Bray et al. 2013]. For non-smooth processes—like the Ornstein-Uhlenbeck (OU) process and Brownian motion—the excursion time density is simply a delta function at 0. For us this means that any photon leaving a scatterer would scatter again immediately, no progress is made and no light transport is possible. Hence we have to limit ourselves to smooth processes. Unfortunately, it is also true that only Markov processes have known analytic solutions for first-passage times [Buonocore et al. 1987]. Since the only Gaussian Markov process is the OU process, which is not smooth, there are

### NDFs for non-heightfield surfaces

Our theory is not restricted to heightfields, and we can model non-heightfield surfaces and their corresponding (v)NDFs by relaxing the correlation constraint along the z-axis. In Fig. 13, we show empirically that these start to approach prior results for aggregate NDFs [d’Eon 2021]. Intuitively, this is explained by examining individual realizations, which start to resemble porous surfaces rather than bumpy heightfields.

**Opposition effect.** Rough surfaces lit from the viewing direction exhibit the “opposition effect”, a noticeable brightening of the surface around the observer’s shadow (Fig. 14). The main contributing factor to this effect—“shadow hiding”—is not captured in shadowing/masking terms derived from Smith theory, but can be reproduced in varying degrees of accuracy using our theory, depending on the memory model. The Renewal+ model provides a reasonable performance/accuracy trade-off.

### 5.3 Volume-type GPISes

The central quantity in volumetric transport is the free-flight distribution. We can derive it for our model by integrating the GPIS density over the sphere of all possible normals at the intersection of computing the ensemble average light transport: Classical microfacet theory with uncorrelated shadowing and masking (top left) does not produce the opposition effect at all. Our memory models produce results closer to the ground truth as we increase the types of correlation we keep track of: the Renewal model loses energy due to the lack of gradient continuity between scattering paths; and the Global model near perfectly reproduces the bright halo around the observer’s shadow.

**Fig. 15.** Left: Realizations sampled along rays starting in free-space. The resulting distribution of zero crossing distances is the first-passage time density and corresponds to \( \Gamma(t \mid \zeta_u) \). Right: Realizations sampled along rays starting at a GPIS intersection. The distribution of zero crossing distances here is the excursion time density and correspond to \( \Gamma(t \mid \zeta_c) \). Realizations are sampled from a GP with squared exponential covariance and zero mean.
We have discussed two extreme ends of the appearance spectrum that we can cover with GPIS, but for both of them, specialized methods exist which perform extremely well compared to our more general method. The value that we see in our method is that it not only does not require an a priori choice between either end of the spectrum, but that it can also represent cases that don’t fit neatly into either category. Particularly interesting is that we can represent mesoscale detail as geometry that is fine enough to require relatively large amounts of storage to represent explicitly, while being large enough to allow for efficient microscale models break down. This covers, for example, tree bark and leaves, grains, and fibers or fur, and we show some didactic examples in Fig. 10. We leave the further investigation to future work, but we believe that at mesoscale level uncertainty, our representation behaves much like certain classes

Fig. 16. We numerically compute transmittance using RIND (dashed) for a range of homogeneous GPISes, allowing us to derive the extinction coefficient \( \sigma_t \) of the closest exponential medium (solid black). Based on the differences between the dashed and solid lines, we can confirm that processes with a higher mean \( \mu \) and integrable kernels produce more exponential first-passage times.

Fig. 17. Here we show results for an index-matched homogeneous medium in a “wedge” (thicker on the left, thinner on the right), set in front of a checkerboard, and lit from above. On the left we render a volume-type GPIS with parameters based on the persistence exponent match discussed in this section and a mirror BSDF. On the right, we render the classical medium with the \( \sigma_t \) we used to match the GPIS and an isotropic phase function.

Fig. 18. Isotropic volume-type GPISes induce phase functions that approach ones predicted by prior work. For a GPIS with a mirror BSDF, we recover a close-to-isotropic phase function (left), whereas for a GPIS with a diffuse BSDF we approach the phase function for Lambertian spherical scatterers described by [d’Eon 2021] (right).
of anisotropic media [Jakob et al. 2010]. In Gaussian process terms, mesoscale detail appears when the variance of the process is on the same scale as variations in the mean. For example, in Fig. 10, the mean shape has an extent of 50 units in each direction, while the variance is 25 for all examples.

6 A NON-STATIONARY GPIS MODEL FOR ACQUISITION & CREATION

In this section, we present a concrete, flexible model that covers a wide range of practical use cases. In particular this model supports the whole range of volume- to surface-type GPIS discussed in Section 5 as well as stochastic mesoscale detail. We restrict ourselves to non-stationary Gaussian process implicit surfaces with a parameterized prior mean function and covariance kernel with parameters \( \Phi \), as well as a set of conditioning points \( C \), where each point \( c \in C \) is associated with a location \( c_\mathbf{x} \), a value \( c_\mathbf{y} \), and a normal derivative direction \( c_\mathbf{v} \). Points that condition the value of the process are encoded as \( c_\mathbf{y} = 0 \) and points that condition the derivative of the process as \( c_\mathbf{v} \neq 0 \). The resulting process, \( \text{GP}(\mu_{\Phi|C}(\mathbf{x}), k_{\Phi|C}(\mathbf{x}, \mathbf{y})) \), has mean and covariance given by

\[
\mu_{\Phi|C}(\mathbf{x}) = \mu_{\Phi}(\mathbf{x}) + k_{\Phi}(\mathbf{x}, C_\mathbf{x})k_{\Phi}(C_\mathbf{x}, C_\mathbf{x})^{-1}(C_\mathbf{y} - \mu_{\Phi}(C_\mathbf{x})) \tag{23}
\]

\[
k_{\Phi|C}(\mathbf{x}, \mathbf{y}) = k_{\mathbf{\Phi}}(\mathbf{x}, \mathbf{y}) - k_{\mathbf{\Phi}}(\mathbf{x}, C_\mathbf{x})k_{\Phi}(C_\mathbf{x}, C_\mathbf{x})^{-1}k_{\mathbf{\Phi}}(C_\mathbf{x}, \mathbf{y}), \tag{24}
\]

which are simply the standard posterior mean and covariance kernel we introduced in Section 3. We are left with having to choose a prior covariance kernel \( k_{\Phi} \).

6.1 A non-stationary kernel

Since appearance depends strongly on the parameters of the kernel we would like to be able to vary them spatially, e.g. to allow for spatially varying roughness in a surface-type GPIS. Building non-stationary covariance kernels with control over local behavior is difficult in general since it can be hard to prove that the global function is positive semi-definite. For example, we can’t simply average between different locally stationary kernels. In the following, we choose to use the non-stationary covariance kernel derived by Paciorek and Schervish [2006]:

\[
k_{\Phi}^{NS}(\mathbf{x}, \mathbf{y}) = \sigma_{\Phi}(\mathbf{x})\sigma_{\Phi}(\mathbf{y}) \left[ \frac{\Sigma_{\Phi}(\mathbf{x})}{\Sigma_{\Phi}(\mathbf{x}) + \Sigma_{\Phi}(\mathbf{y})} \right]^{-\frac{1}{2}} k_{\Phi}^S(\sqrt{\Sigma_{\Phi}(\mathbf{x}, \mathbf{y})}), \tag{25}
\]

where \( k_{\Phi}^S(t) \) is any stationary covariance kernel and

\[
Q_{\Phi}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^\top \left( \Sigma_{\Phi}(\mathbf{x}) + \Sigma_{\Phi}(\mathbf{y}) \right)^{-1} (\mathbf{x} - \mathbf{y}). \tag{26}
\]

This kernel is characterized by the parameters of the global stationary kernel \( k_{\Phi}^S \) as well as two spatially varying fields: the “local variance” \( \sigma_{\Phi} : \mathbb{R}^3 \rightarrow \mathbb{R} \) and the “local anisotropy” \( \Sigma_{\Phi} : \mathbb{R}^3 \rightarrow S_2^3 \) where \( S_2^3 \) is the set of positive semi-definite 3x3 matrices. Note that \( \Sigma_{\Phi} \) is in concept very close to how anisotropy is expressed in the SGGX [Heitz et al. 2015] microflake phase function. The kernel \( k_{\Phi}^{NS} \) is very flexible, has intuitive parameters, and has been used successfully to model complex priors [Dechheimer and Davison 2023]. We show an example of the results that we can generate by varying these parameters spatially in Fig. 19.

6.1.1 Model Storage. We store the mean, variance, and anisotropy fields, \( \mu_{\Phi}(\mathbf{x}), \sigma_{\Phi}(\mathbf{x}), \) and \( \Sigma_{\Phi}(\mathbf{x}) \), on a voxel grid and retrieve values at arbitrary evaluation locations via interpolation. For \( \Sigma_{\Phi}(\mathbf{x}) \), we use the encoding and linear interpolation method proposed by Heitz et al. [2015]. This allows for smooth interpolation while guaranteeing that the interpolated matrices are positive semi-definite.

6.2 Manual editing of GPIS

The most basic way to create GPIS scenes is to specify the underlying geometry primitive. For the anisotropy field, we can similarly provide CSG operations and warping functions, offset the isosurface, and more. Similarly, the variance field can be edited like any other classical implicit surface. That is, we can easily apply

Fig. 20. We show the effect of the different editing methodologies we discuss in Section 6.2. For example, we can start with a volume-type GP (i.e. homogeneous mean and isotropic covariance, left) and use prior shaping to insert a solid object (e.g. a plane with varying uncertainty, middle) and then deform that object using posterior conditioning (9 handles, right).
“anisotropy brushes” that paint the appropriate vector fields to parameterize the anisotropy. We created most of the simple scenes in this paper using this type of editing, as it provides very direct control over all aspects of the GPIS.

6.2.2 Explicit control via posterior conditioning. Additionally, we can allow the user to exploit the conditioning capabilities of the Gaussian process representation. To do so, we enable the user to specify points at which they can place value and derivative constraints. This is, in principle, an extension to Witkin and Heckbert’s control particle approach [Hart et al. 2002; Witkin and Heckbert 1994] for controlling implicit surfaces. The GPIS is then conditioned on taking on the specified values using classical Gaussian process regression techniques as discussed in Section 3.1.3. When rendering, we can simply add the “global” conditioning to our path conditioning variable $\xi$, since they behave exactly like any other conditioning we introduce along the path (except that they are not “forgotten” by any memory model).

6.2.3 CSG on GPISes. Ideally, one would like to perform CSG operations on GPISes directly. Unfortunately, GPISes are only closed under linear transforms (e.g. the sum of two GPISes is a GPIS, but the product is not). Many CSG operations contain non-linear operations, such as the min or the max. This means that we must support a certain set of non-Gaussian stochastic implicit surfaces to support the full range of CSG operations. Luckily, this is trivial to do in our current implementation. We can simply generate a realization for each leaf node in the CSG tree and then apply the CSG operations to the realizations to get the final sample. Our implementation hence supports the non-Gaussian SISes produced by the non-linear combination of “leaf GPs”. This flexibility does, however, mean that the cost of generating realizations depends on the number of leaves in the CSG tree, i.e. we can not “bake down” to a more compact representation. Additionally, we must keep track of path conditioning variables for each leaf GP. A lossy alternative is to fit a GPIS to the non-Gaussian SIS. We show the results of this when applied to the intersection of two sphere GPISes with different variances in Fig. 21. CSG GPISes are not only useful for artistic control but have also proven themselves as a rich model for real-world microstructure in other fields [Torquato and Lu 1993].

![CSG dist. (union) best Gaussian fit](image)

Fig. 21. We can use CSG operations to combine multiple GPISes (left, blue: point distribution). The resulting process is not Gaussian anymore, but our rendering algorithm is still able to visualize it by sampling from each GPIS individually and applying the CSG operations on each set of samples (middle). If we fit the closest Gaussian to the CSG distribution (left, orange), we can represent the result approximately with a single GPIS (right).

6.3 Stochastic Poisson surface reconstruction

Recently, Sellán and Jacobson [2022] presented a surface reconstruction method, based on the widely used Poisson surface reconstruction algorithm, that recovers both the mean implicit surface as well as a spatially varying variance field. We can visualize the output of their method using our rendering algorithm and show a simple example in Fig. 22. While Sellán and Jacobson [2022] do show a rendering-related example and even take the full correlations along single rays into account [Sellán and Jacobson 2023], they are mainly interested in the uncertainty of primary hit locations and do not consider global illumination.

6.4 Filtering implicit surfaces

Inspired by the connections between microfacet surfaces and GPISes that we discussed in Section 5.2, we present a method to use GPISes for implicit surface downsampling by representing the sub-resolution surface details via their statistics. In the following, we show how to leverage this intuition for a principled way to construct low-capacity GPISes that approximate high-capacity surface models well. This is, in spirit, similar to methods that learn the statistics of textures from examples [Galerne et al. 2012; Guehl et al. 2020] or downsample normal maps by fitting distributions over the filtered area [Dupuy et al. 2013; Olano and Baker 2010].

![downsample an implicit surface with non-stationary statistics](image)

Fig. 23. We downsample an implicit surface with non-stationary statistics (left) by first separating it into a mean field (inset) and a residual (middle-left). We then use a short-time Fourier transform based method to recover an approximation to the non-stationary autocorrelation function (ACF) of the residual (middle-right). Finally, we fit a stationary kernel to each of the ACF blocks (right). We can then interpolate the parameters of this kernel to compute the global non-stationary covariance kernel in Eq. (25).
We can go one step further and instead find the Gaussian process when the number of parameters is large. Instead, we exploit the local geometry information and any derived quantities, such as light transport, as well as possible. In general, we parameterize the downsampled approximation \( f_\theta \) by some finite set of values \( \Phi \), chosen such that the residual \( R_\Phi(x) = f_\Phi(x) - f(x) \) is minimal over the domain, i.e. \( \Phi \) is chosen to minimize \( \int_\Omega R_\Phi(x)^2 \, dx \). Traditionally, this is done using maximum likelihood estimation, but this discards any information other than the mean. This results in a drastic difference in the light transport as we show in Fig. 24. We can go one step further and instead find the Gaussian process with mean \( \mu_\Phi(x) = f_\Phi(x) \) that is most likely to sample \( f(x) \). That is, we effectively capture the original function as well as possible with the limited capacity we have and then estimate the statistics of the residual. The resulting process has the form \( \text{GP}(\mu_\Phi(x), k_\Theta(x, y)) \).

We now need to find the parameters \( \Theta \) of the kernel \( k_\Theta \) such that the prior process has statistics that are as close to possible to the residual. This is a common task in Gaussian process regression and the parameters can be found by minimizing the negative log marginal likelihood, but this process is expensive and can be unstable when the number of parameters is large. Instead, we exploit the local stationarity of our model and compute kernel parameters assuming stationarity over some finite region. We do so by numerically computing the autocorrelation function (ACF) of the residual using a Fourier transform-based method over subregions of the domain. We then find kernel parameters by fitting our model to the numerically retrieved ACF in each region. This gives us a grid of kernel parameters, which we can interpolate to approximate the full non-stationary kernel. We illustrate this process in Fig. 23. Note that here we only discuss fitting based on geometric error. We could alternatively define other loss functions, such as an appearance-based one, or one that aims to match transmittance between points in space.

Given some high-resolution representation of an implicit surface \( f(x) \), we want to downsample it, for example, to use less storage, while preserving the geometry information and any derived quantities, such as light transport, as well as possible. In general, we parameterize the downsampled approximation \( f_\theta \) by some finite set of values \( \Phi \), chosen such that the residual \( R_\Phi(x) = f_\Phi(x) - f(x) \) is minimal over the domain, i.e. \( \Phi \) is chosen to minimize \( \int_\Omega R_\Phi(x)^2 \, dx \). Traditionally, this is done using maximum likelihood estimation, but this discards any information other than the mean. This results in a drastic difference in the light transport as we show in Fig. 24. We can go one step further and instead find the Gaussian process with mean \( \mu_\Phi(x) = f_\Phi(x) \) that is most likely to sample \( f(x) \). That is, we effectively capture the original function as well as possible with the limited capacity we have and then estimate the statistics of the residual. The resulting process has the form \( \text{GP}(\mu_\Phi(x), k_\Theta(x, y)) \).

We now need to find the parameters \( \Theta \) of the kernel \( k_\Theta \) such that the prior process has statistics that are as close to possible to the residual. This is a common task in Gaussian process regression and the parameters can be found by minimizing the negative log marginal likelihood, but this process is expensive and can be unstable when the number of parameters is large. Instead, we exploit the local stationarity of our model and compute kernel parameters assuming stationarity over some finite region. We do so by numerically computing the autocorrelation function (ACF) of the residual using a Fourier transform-based method over subregions of the domain. We then find kernel parameters by fitting our model to the numerically retrieved ACF in each region. This gives us a grid of kernel parameters, which we can interpolate to approximate the full non-stationary kernel. We illustrate this process in Fig. 23. Note that here we only discuss fitting based on geometric error. We could alternatively define other loss functions, such as an appearance-based one, or one that aims to match transmittance between points in space.

7 LIMITATIONS OF GPISES

Gaussian process implicit surfaces inherently share the characteristics of implicit surfaces, enabling robust representation capabilities. However, this also implies certain challenges inherent to implicit surfaces. For instance, while editing tools are generally more developed for explicit representations like meshes, GPISES face challenges in applying techniques like texture mapping due to complexities in parameterizing the 3D surface. Further, the lack of analytic expressions for transmittance and normal distributions prevents our current formalism from being applied to techniques such as differentiable rendering. Future progress in this respect would be of great interest in light of the recent focus on neural radiance field techniques and related representations.

Performance. In terms of rendering, implicit surfaces, including those not based on Gaussian processes, are typically slower than triangle mesh rendering, which may limit their use in real-time applications like video games. Nonetheless, for problems where strong spatial correlations in stochastic geometry significantly influence the light transport, our proposed rendering framework is orders of magnitude more efficient than a brute force global-sampling approach and enables novel appearances that were not previously practical. The performance of our method is mainly determined by the amount of correlations one keeps track of—both across ray intersections and within one ray segment, see Section 3 of the supplemental. We have found that the acceleration techniques we provide make it possible to produce images of simple scenes (e.g. Figs. 1, 9 and 19) in a “reasonable” amount of time on a standard desktop PC, i.e. around 30 seconds for a 512 × 512 pixel render at 1 sample per pixel when using the Renewal+ memory model. While more than an order of magnitude slower than rendering e.g. microfacet surfaces directly, it is fast enough for us to explore the expressive power of GPISES and efficiently iterate on improvements in future work. In contrast, consider the global-sampling method of sampling many realizations and averaging many rendered images. In Fig. 1, for example, the variations in the surface of the rhinoceros are on the order of 1 mm with the statue being roughly 1.6 m tall. To resolve these details we would need to discretize to a grid of around 3200³ and to sample a realization at that resolution then requires inverting a 3200³ × 3200³ matrix, which would need 4,294,967,200,000 GB for storage alone.

Discrete sampling/ray marching. The function-space algorithm we present in Section 4.2 represents realizations along ray segments using a discrete set of samples. Finding the first zero-crossing is then done via ray marching. If the step size is too large, we encounter the usual issues that appear when finding implicit surface ray intersections with ray marching, such as missing small surface features. In practice, we choose the step size relative to the correlation length of the covariance kernel (e.g. we set \( \Delta t \) such that \( k(x, x + \Delta t \vec{d})/k(x, x) > 0.95 \)). A more advanced “stochastic root finding” algorithm and smarter sample distribution are interesting.
topics for future work that have the potential to improve the performance of GPIS rendering without large structural changes to the method we present here.

Lack of next-event estimation. As mentioned in Section 1, we currently only support next-event estimation (NEE) when using diffuse or glossy micro-BSDFs, but not when using a perfect mirror micro-BSDF. This is because we treat the normal sampled at the intersection point (see Alg. 1) as a deterministic quantity when evaluating the micro-BSDF. While this choice does reduce implementation complexity, it is not a fundamental limitation of our approach—after all, the normal is sampled from a distribution, and hence, we should be able to apply NEE. We present some thoughts on how this could be done in Section 3.3 of the supplemental.

8 FUTURE WORK AND CONCLUSION

We see the example use cases that we presented in Section 6 as validating the expressiveness of GPISes, but concede that they fall short of being “applications” immediately useful in practice. Hence we see much room for future work.

8.1 Correlated appearance

For a full rendering method, we need to be able to retrieve additional properties, such as surface albedo, at intersection points. Of course, we can simply store these in a separate volumetric data structure, which we do in this work, but Gaussian processes provide us with an elegant mechanism for a more principled solution. Multi-output or multi-task GPs predict vector-valued functions \( f_N : \mathbb{R}^D \rightarrow \mathbb{R}^N \) and allow for cross-correlations in their outputs. We should be able to extend the fitting approach presented in Section 6.4 to include BSDF parameters.

8.2 Solving elliptic partial differential equations

A central property of this volumetric geometry representation is that it can be used for a wide variety of geometry processing tasks instead of just rendering. In particular, we are optimistic that a large part of the derivations in Section 4 are applicable to other recursive integral equations, such as the ones that appear in walk-on-sphere [Sawhney and Crane 2020] and walk-on-boundary [Sugimoto et al. 2023] methods. This should allow us to solve certain classes of elliptic partial differential equations while taking uncertainty in the domain boundary and boundary conditions into account.

8.3 More general stochastic processes

So far we have limited ourselves to Gaussian processes. This is a common choice because they possess many desirable properties, such as being closed under conditioning and taking derivatives. More general stochastic processes would allow for an even more expressive function space, but at the cost of fewer analytical results. There is existing work that aims to use Gaussian processes to approximate more general stochastic processes [Gurley 1997; Hong et al. 2023; Shields et al. 2011], but any Gaussian process will always produce Beckmann normal statistics. This is limiting because prior work has shown that normal distributions in the real-world often have longer tails [Walter et al. 2007]. A logical first choice for a more complex class of stochastic processes would be the student-t process, which is essentially a weighted sum of infinitely many Gaussian processes. There is recent work on student-t microfacet surfaces [d’Eon 2023] that shows that the resulting vNDFs match real-world results very well.

8.4 Advanced weight-space methods

The global realization sampling method via the weight-space approximation is very attractive because it conceptually simplifies GPIS rendering. Unfortunately, the standard formulation of weight-space sampling via an expansion of basis functions according to the process’ spectral density is limited to stationary covariance kernels. We are not aware of any prior work that extends this to general non-stationary kernels other than methods that use the Karhunen-Loève expansion [Huang et al. 2001], which are generally not seen as practical in applications. That said, we think it is worthwhile to explore how to extend weight-space sampling to support a, maybe limited, range of non-stationary kernels. We could imagine assuming local stationarity and blending basis-function representations spatially would produce acceptable results for certain applications.

8.5 Uncertainty Quantification

A natural question is whether we can propagate the uncertainty in the scene geometry to uncertainty in the rendered results. That is, instead of just computing the ensemble average, we would like to extract higher-order moments, primarily variance. This problem is known as “uncertainty quantification” [Smith 2013] and progress here would be useful in a range of real-world applications. Doing this efficiently when we only have access to Monte Carlo estimates of the ensemble average is non-trivial and we discuss some reasons for this in Section 4 of the supplemental. Intuitively, it requires us to untangle the variance due to uncertainty in the geometry from the variance due to Monte Carlo sampling of light paths. Also known as separating parametric uncertainty from intrinsic uncertainty, methods based on polynomial chaos [Mueller et al. 2023] seem promising but assume relatively low intrinsic uncertainty. Extending these to the high-variance Monte Carlo models common in graphics would allow us to not only compute variation due to geometry but also other scene parameters.

8.6 Conclusion

In this paper, we have shown that reasoning about the light transport in GPISes allows us to make deep connections between existing stochastic representations in graphics. We have done so by relying on Monte Carlo sampling of the involved integrals without being overly concerned with performance. However, real-world applications rely on the availability of performant methods. While compute power grows with time, and methods that were previously seen as intractable become widely used, we believe that there is much room for advancing the theory of GPIS rendering. For example, there is much interest in GPs for large data sets, and we expect that advances there would directly translate to faster rendering. Analytic approximations for first passage times are important in many areas of research and would allow us to decouple rendering times from sampling efficiency. Tabulating transport functions that cannot be
We would like to thank Alex Evans for his valuable contributions.

ACKNOWLEDGMENTS

We would like to thank Alex Evans for his valuable contributions in the early stages of this project and Silvia Sellán for helpful discussions on stochastic geometry representations. We thank Aaron LeFohn for supporting this research. Additionally, this work was generously supported by NVIDIA, a Dartmouth Senior Faculty grant, and a Neukom ComP2X grant. Dario Seyb and Wojciech Jarosz were partially funded by NSF award 1844538. We have used assets in the public domain / licensed under CC0 in many figures of this paper and thank the original creators for providing them. The volume data for the explosion in Fig. 1 is courtesy of JangaFX, and the two stat-tions next to it are from threedscans.com. The "shader ball" in Figs. 7 and 9 is based on code in the SDF dataset published by Takikawa et al. [2022]. The environment map used in Fig. 11 is part of the Tungsten example assets. The dragon figure in Fig. 19 is part of the OpenVDB example assets. The bunny model used in Fig. 22 is based on a model in the Stanford 3D Scanning Repository. Finally, the tree in Fig. 24 is from polyhaven.com. ChatGPT was utilized to iterate on early versions of the abstract and introduction—all final text was written by the authors.

REFERENCES


Eugene d’Eon. 2023. Student’s-T and beyond: Practical tools for multiple-scattering BSDFs with general NDFs. In ACM SIGGRAPH 2023 Talks. ACM, New York, NY, USA. https://doi.org/10/mr42


ACM Trans. Graph., Vol. 43, No. 4, Article 112. Publication date: July 2024.