Importance Sampling Microfacet-Based BSDFs using the Distribution of Visible Normals
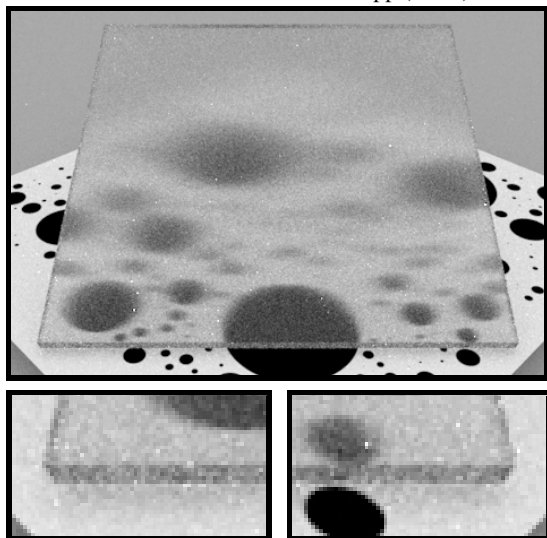
Eric Heitz, Eugene d'Eon

Importance Sampling Microfacet-Based BSDFs using the Distribution of Visible Normals

E. Heitz[1] and E. d'Eon[2]

[1]INRIA ; CNRS ; Univ. Grenoble Alpes
[2]Jig Lab, Wellington, New Zealand

Previous: BSDF Importance sampling using the distribution of normals. 64 spp (10.0s)

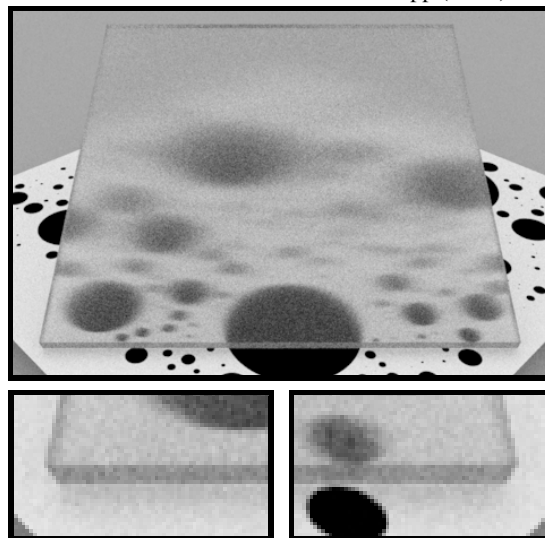Our: BSDF Importance sampling using the distribution of visible normals. 58 spp (10.4s)

Figure 1: *A dielectric glass plate (n = 1.5) with anisotropic GGX roughness ($\alpha_x = 0.05$, $\alpha_y = 0.4$) on all faces (with the Smith masking function). For a similar sample budget and the same render time, our method (right) significantly reduces the variance and converges faster than the common technique used in previous work (left).*

Abstract

We present a new approach to microfacet-based BSDF importance sampling. Previously proposed sampling schemes for popular analytic BSDFs typically begin by choosing a microfacet normal at random in a way that is independent of direction of incident light. To sample the full BSDF using these normals requires arbitrarily large sample weights leading to possible fireflies. Additionally, at grazing angles nearly half of the sampled normals face away from the incident ray and must be rejected, making the sampling scheme inefficient. Instead, we show how to use the distribution of visible normals directly to generate samples, where normals are weighted by their projection factor toward the incident direction. In this way, no backfacing normals are sampled and the sample weights contain only the shadowing factor of outgoing rays (and additionally a Fresnel term for conductors). Arbitrarily large sample weights are avoided and variance is reduced. Since the BSDF depends on the microsurface model, we describe our sampling algorithm for two models: the V-cavity and the Smith models. We demonstrate results for both isotropic and anisotropic rough conductors and dielectrics with Beckmann and GGX distributions.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Keywords: Microfacet BSDF, Importance Sampling

1. Introduction

Rendering realistic surfaces requires the use of complex material models. The most common way to define surface reflectance is the Bidirectional Scattering Distribution Functions (BSDF), which incorporates light transports at a surface's interface such as reflection, described by a Bidirectionnal Reflectance Distribution Functions (BRDF), or transmission, described by a Bidirectional Transmittance Distribution Functions (BTDF). Microfacet theory is the most common theoretical framework for deriving analytic physically-based BSDFs for rough surfaces. For Monte Carlo rendering algorithms to make images efficiently using BSDFs, an efficient importance sampling must be provided. In this paper we present a new strategy for deriving importance sampling for analytic microfacet BSDFs. In constrast to previous approaches, we exploit the structure of the microfacet theory to directly sample from the distribution of visible normals, which produces higher quality and lower variance samples in comparison to previous approaches.

A key element of our approach is to move away from normal distributions directly and use, instead, the distribution of slopes for the microsurface. In slope space, we note two key invariance relationships that produce an optimal sampling strategy for any surface roughness or anisotropy, provided the optimal strategy for an isotropic surface with unit roughness is known.

In Section 2, we review the microfacet-based BSDF model and we explain why the common associated importance sampling scheme is not optimal. In Section 3, we show that importance sampling the BSDF with the distribution of visible normals makes a better use of the sampling space and produces sample weights in $[0,1]$. This method does not create firefly artifacts and leads to a significative variance reduction (Figure 2). To make this practical, in Section 4, we show how to importance sample the distribution of visible normals analytically with the V-cavity microsurface model. In Section 5, we show how to importance sample the distribution of visible normals with the Smith microsurface model. We propose analytic sampling procedures for Beckmann and GGX distributions in the associated supplemental material. For other distributions, if an analytic sampling procedure is not available we propose to use precomputed data. Thanks to the shape invariance property of the distribution of visible slopes, we can do the precomputations for only one BSDF with roughness $\alpha = 1$ and we use the same data with arbitrary roughness ($\alpha \neq 1$) and anisotropy ($\alpha_x \neq \alpha_y$). In this way, we importance sample an entire class of anisotropic parametric BSDFs with a small amount of precomputed data.

2. Background and Previous Work

2.1. Microfacet Theory

Microfacet theory was originally developed in the early work of Torrance et al. [TS67, CT82]. Several alternative models
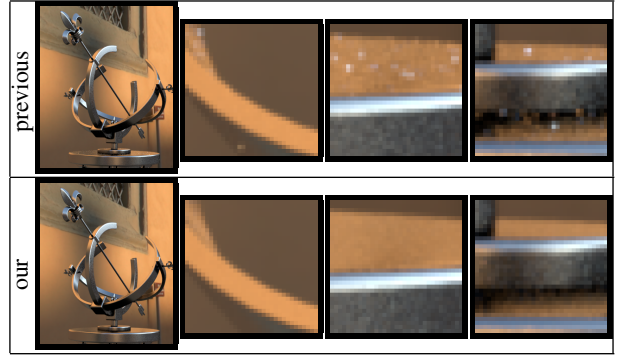
Figure 2: *An isotropic rough conductor with GGX distribution ($\alpha = 0.05$) and the Smith masking function (64 samples per pixel).*

or approximations are built on top of this theory [War92, vGSK98, AS00, KSK01]. This paper is built on the top of two major previous works:

- The extension of microfacet theory to transmittance by Stam [Sta01]. Closely related, Walter et al. [WMLT07] describe BSDF importance sample strategies for different distribution of normals.
- Heitz' investigation of the properties of the masking-shadowing function in microfacet-based BRDFs [Hei14]. He shows how the masking function is related to the microsurface model and to the correct normalization of the BRDF.

In the following, we review microfacet theory as presented in these two previous works.

Figure 3: *The physically-based BSDF model is constructed by applying the operators reflect and transmit on the distribution of visible normals. The ratio of reflection and transmission is given by the Fresnel term F.*

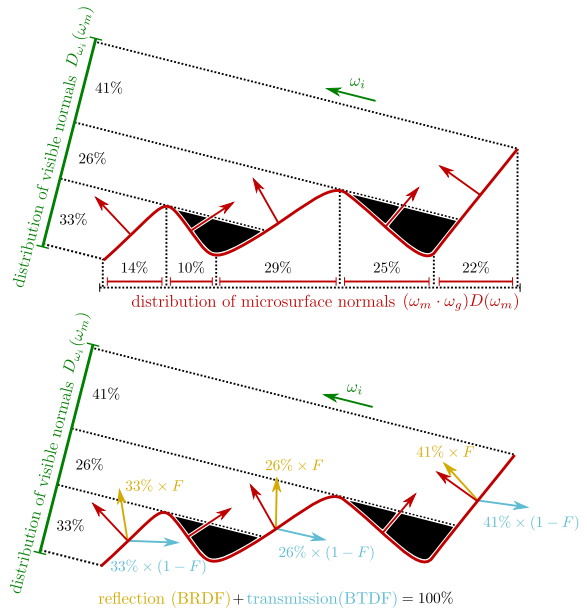|	
$\omega_g = (0,0,1)$	geometric normal
$\omega_m = (x_m, y_m, z_m)$	microfacet normal
$\omega_i = (x_i, y_i, z_i)$	incident direction† $(\omega_i \cdot \omega_g > 0)$
$\omega_o = (x_o, y_o, z_o)$	outgoing direction
$D(\omega_m)$	distribution of normals
$D_{\omega_i}(\omega_m)$	distribution of visible normals
$G_1(\omega_i, \omega_m)$	masking function
$G_2(\omega_i, \omega_o, \omega_m)$	masking-shadowing function
$\omega_1 \cdot \omega_2$	dot product
$|\omega_1 \cdot \omega_2|$	absolute value of the dot product
$\langle \omega_1, \omega_2 \rangle$	clamped dot product
$\chi^+(a)$	Heaviside function: 1 if $a > 0$ and 0 if $a \le 0$

Table 1: *Notation.*

The Distribution of Normals D describes the statistical distribution of the normals ω_m of which the microsurface is made. The area of the microsurface projected onto the geometric normal ω_g is the unit area of the geometric surface:

$$\int_\Omega (\omega_m \cdot \omega_g) D(\omega_m) \, d\omega_m = 1. \tag{1}$$

Parametric models typically specify *roughness parameters* α_x and α_y in two orthogonal tangent directions ($\alpha_x = \alpha_y$ for isotropic BSDFs). When these parameters need to be included explicitly we write $D(\omega_m, \alpha_x, \alpha_y)$ instead of $D(\omega_m)$.

The Distribution of Visible Normals D_{ω_i} describes the statistical distribution of normals ω_m visible from the incident direction† ω_i and is defined by

$$D_{\omega_i}(\omega_m) = \frac{G_1(\omega_i, \omega_m) |\omega_i \cdot \omega_m| D(\omega_m)}{|\omega_i \cdot \omega_g|}, \tag{2}$$

where the term $|\omega_m \cdot \omega_i|$ is the projected area of the normal ω_m in direction ω_i, $|\omega_i \cdot \omega_g|$ is the projected area of the geometric surface in direction ω_i, and G_1 is the masking function. The relation between D and D_{ω_i} is illustrated in Figure 3. Heitz [Hei14] shows that in a mathematically well-defined BSDF model, the masking function G_1 should be such that the distribution is normalized:

$$\int_\Omega D_{\omega_i}(\omega_m) \, d\omega_m = 1.$$

Our approach is to use the knowledge of the BSDF model to improve the sampling strategy. We can do it only for mathematically well-defined models. Two common microsurface models have masking functions G_1 that satisfy this criteria: the V-cavity and the Smith models.

---

† To ease the notation, we assume that the incident direction ω_i is located in the upper part of the hemisphere, i.e. $\omega_i \cdot \omega_g > 0$.

The V-cavity Microsurface Model was introduced by Torrance et al. [TS67, CT82]. It assumes a continuous distribution of separate microsurfaces rather than just one microsurface [Hei14]. Each microsurface is composed of two normals $\omega_m = (x_m, y_m, z_m)$ and $\omega_m' = (-x_m, -y_m, z_m)$ (see Figure 11 in Appendix B) and the contribution of each surface is weighted by PDF $(\omega_m \cdot \omega_g) D(\omega_m)$ in the BSDF. The masking function of a V-cavity microsurface is given by

$$G_1(\omega, \omega_m) = \chi^+\left(\frac{\omega \cdot \omega_m}{\omega \cdot \omega_g}\right) \min\left(1, 2\frac{|\omega_m \cdot \omega_g| \, |\omega \cdot \omega_g|}{|\omega, \omega_m|}\right), \tag{3}$$

where the heaviside function ensures that backfacing microfacets are discarded for the correct side of the microsurface. The traditional masking-shadowing function accounts for height correlation in the BRDF: ω_o is in the upper hemisphere ($\omega_o \cdot \omega_g > 0$) and masking and shadowing occur on the same side of the microfacet and are correlated. We generalize this to the BTDF: ω_o is in the lower side of the microsurface and masking and shadowing occur on opposite sides of the microfacet and are anticorrelated. The masking-shadowing function is

$$G_2(\omega_i, \omega_o, \omega_m) =$$

$$\begin{cases} \min\left(G_1(\omega_i, \omega_m), G_1(\omega_o, \omega_m)\right), & \text{if } \omega_o \cdot \omega_g > 0, \\ \max\left(G_1(\omega_i, \omega_m) + G_1(\omega_o, \omega_m) - 1, 0\right), & \text{otherwise.} \end{cases}$$

Kelemen et al. proposed a simplified version of this masking-shadowing function to aid improve efficiency and permit importance sampling [KSK01]. However, their function does not satisfy Equation (2), which makes their model physically ill-defined. In Section 4, we will show that using the V-cavity function can actually lead to a simple importance sampling scheme, which we find to be more efficient than the Kelemen model.

The Smith Microsurface Model. The Smith masking function was originally derived from a raytracing formulation [Smi67] for Beckmann distributions D and generalized to other distributions by Brown [Bro80]. Recently, Heitz showed that the Smith masking function can be derived from the equation of the conservation of the projected area [Hei14]. The Smith microsurface model is the most accurate model for compactly describing geometrical-optics reflectance and transmission from explicit random height fields (such as an ocean surface), and is better than the V-cavity model for approximating reflectance of measured material [Hei14] . However, specific masking functions G_1 must be derived for each new distribution of normals. Analytic solutions are available for Beckmann [Smi67] and GGX [WMLT07]. Heitz showed that the same formulas can be used with the anisotropic extensions of D [Hei14]. Different forms of the Smith masking-shadowing function are available, the most simple is the non-correlated form: $G_2(\omega_i, \omega_o, \omega_m) = G_1(\omega_i, \omega_m) G_1(\omega_o, \omega_m)$.

The Scattering Model is given by the local illumination equation:

$$L(\omega_i) = \int_\Omega L(\omega_o) \, |\omega_o \cdot \omega_g| \, f(\omega_i, \omega_o) \, d\omega_o, \qquad (4)$$

where f is the BSDF and $L(\omega_o)$ is the radiance in the outgoing direction. The construction of the BSDF model is detailed by Walter et al. [WMLT07] and is illustrated in Figure 3(b): the distribution of visible normals is transformed into a distribution of outgoing directions ω_o by *reflection* and *transmission*. The BSDF $f(\omega_i, \omega_o) = f_r(\omega_i, \omega_o) + f_t(\omega_i, \omega_o)$ is the sum of the BRDF

$$f_r(\omega_i, \omega_o) = \frac{F(\omega_i, \omega_{hr}) \, G_2(\omega_i, \omega_o, \omega_{hr}) \, D(\omega_h)}{4 \, |\omega_i \cdot \omega_g| \, |\omega_o \cdot \omega_g|},$$

and the BTDF

$$f_t(\omega_i, \omega_o) =$$

$$\frac{|\omega_i \cdot \omega_{ht}| \, |\omega_o \cdot \omega_{ht}|}{|\omega_i \cdot \omega_g| \, |\omega_o \cdot \omega_g|} \frac{n_0^2 \, (1 - F(\omega_i, \omega_{ht})) \, G_2(\omega_i, \omega_o, \omega_{ht}) \, D(\omega_{ht})}{(n_i(\omega_i \cdot \omega_{ht}) + n_o(\omega_o \cdot \omega_{ht}))^2},$$

where ω_{hr} and ω_{ht} are the half vectors of reflection and transmission respectively, and n_i and n_o are the indices of refraction of the incident and transmitted sides respectively (we refer the reader to [WMLT07] for more details).

2.2. Importance Sampling the BSDF

Typically, importance sampling is used to solve Equation (4). The most common importance sampling strategy for parametric BSDFs uses the distribution of normals to generate the samples [War92, AS00, APS00, KSK01, Ash07, WMLT07]. In this paper, we call it the "previous method" and it is illustrated in Algorithm 1. In order to importance sample the cosine weighted BSDF $|\omega_o \cdot \omega_g| \, f(\omega_i, \omega_o)$, previous algorithms start by generating normals ω_m with the PDF $(\omega_m \cdot \omega_g) D(\omega_m)$ rather than sampling outgoing directions ω_o directly. Then, a light transport operator is applied on the sampled normal to generate the outgoing direction. The transport operators, *reflect*

$$\omega_o = 2|\omega_i \cdot \omega_m|\omega_m - \omega_i,$$

and *transmit*

$$\omega_o = \left(nc - \text{sign}(\omega_i \cdot \omega_g)\sqrt{1 + n(c^2 - 1)} \right) \omega_m - n\omega_i,$$

with $c = (\omega_i \cdot \omega_m)$, and $n = \frac{n_i}{n_o}$,

are chosen randomly and with a probability depending on the Fresnel term: $F(\omega_i, \omega_m)$ is the reflection probability and $1 - F(\omega_i, \omega_m)$ is the transmission probability. The PDF resulting from this procedure is the PDF used to generate the normals multiplied by the Jacobian of the chosen operator:

$$\left\| \frac{\partial \omega_{hr}}{\partial \omega_i} \right\| = \frac{1}{4 \, |\omega_i \cdot \omega_{hr}|},$$

$$\left\| \frac{\partial \omega_{ht}}{\partial \omega_i} \right\| = \frac{n_o^2 \, |\omega_o \cdot \omega_{ht}|}{(n_i(\omega_i \cdot \omega_h) + n_o(\omega_o \cdot \omega_h))^2}.$$

Since the same Jacobians are present in the sampling PDF and in the BRDF/BTDF expressions, they simplify out in the sample weight expression. Also, since reflection and transmission are sampled randomly according to the Fresnel term, the value of Fresnel is present in the sampling PDF as well as in the BRDF/BTDF, and simplifies out in the sample weight expression. In both cases the final expression for the sample weight reduces to

$$\text{weight}(\omega_o) = \frac{|\omega_i \cdot \omega_m| \, G_2(\omega_i, \omega_o, \omega_m)}{|\omega_i \cdot \omega_g| \, |\omega_m \cdot \omega_g|}.$$

Discussion This sampling algorithm is almost perfect at normal incidence but we can see in Figure 4 that it has two main problems at grazing angles. The first problem is that up to half of the generated normals ω_m are backfacing ($\omega_i \cdot \omega_m < 0$). The associated weights are 0 and so half of the sample space is wasted. The second problem at grazing angles is that the sample weights can be arbitrarily high with the Smith model leading to unwanted variance, especially for paths with multiple surface interactions (though the sample weights are at most 2 with the V-cavity model). This is responsible for the firefly artifacts in Figure 2. Walter et al. [WMLT07] notice this problem and propose using a BSDF with a slightly different roughness α' for sampling the NDF at grazing angles: $\alpha' = (1.2 - 0.2\sqrt{|\omega_i \cdot \omega_m|}\alpha)$. This is denoted as "Walter's trick" in Figure 4. Walter's trick significatively reduces the maximal weight with Beckmann distributions but does not help much with GGX (in our experience, despite trying offset values greater than 1.2) and also does not alleviate the problem of wasted samples.

Algorithm 1 Sample ω_o (previous)

> $\omega_m \leftarrow$ Sample $(\omega_m \cdot \omega_g) D(\omega_m)$
> **if** $\mathcal{U} < F(\omega_i, \omega_m)$ **then**
> > $\omega_o \leftarrow$ reflect(ω_i, ω_m)
> **else**
> > $\omega_o \leftarrow$ transmit(ω_i, ω_m)
> **end if**
> weight$(\omega_o) \leftarrow \frac{|\omega_i \cdot \omega_m| \, G_2(\omega_i, \omega_o, \omega_m)}{|\omega_i \cdot \omega_g| \, |\omega_m \cdot \omega_g|}$

Algorithm 2 Sample ω_o (our)

> $\omega_m \leftarrow$ Sample $D_{\omega_i}(\omega_m)$ ▷ Algorithm 3 or 4
> **if** $\mathcal{U} < F(\omega_i, \omega_m)$ **then**
> > $\omega_o \leftarrow$ reflect(ω_i, ω_m)
> **else**
> > $\omega_o \leftarrow$ transmit(ω_i, ω_m)
> **end if**
> weight$(\omega_o) \leftarrow \frac{G_2(\omega_i, \omega_o, \omega_m)}{G_1(\omega_i, \omega_m)}$

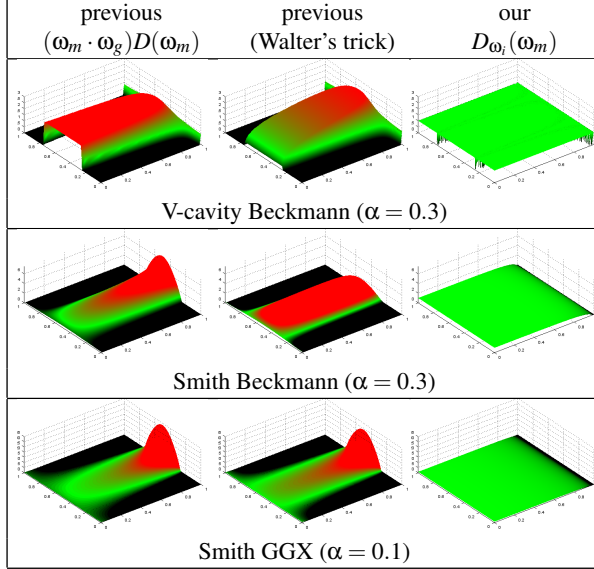|	previous	our
$\text{PDF}(\omega_m)$	$(\omega_m \cdot \omega_g)\,D(\omega_m)$	$D_{\omega_i}(\omega_m)$
$\text{weight}(\omega_o)$	$\dfrac{\lvert\omega_i\cdot\omega_m\rvert\,G_2(\omega_i,\omega_o,\omega_m)}{\lvert\omega_i\cdot\omega_g\rvert\,\lvert\omega_m\cdot\omega_g\rvert}$	$\dfrac{G_2(\omega_i,\omega_o,\omega_m)}{G_1(\omega_i,\omega_m)}$
V-cavities G	$\text{weight}(\omega_o) \in [0,2]$	$\text{weight}(\omega_o) \in [0,1]$
Smith G	$\text{weight}(\omega_o) \in [0,\infty[$	$\text{weight}(\omega_o) \in [0,1]$

Table 2: *Properties of the importance sampling techniques.*

Figure 4: *The sampling weights with different PDFs at grazing angle ($\theta_i = 1.5$). The sampling space is parametrized by the two random numbers $(\mathcal{U}_1, \mathcal{U}_2)$ used to generate ω_m. Our sampling scheme completely fills the sample space with weights very near 1.0. (The plot of our method with the V-cavity model is noisy because we use a supplemental random number $\mathcal{U}_3$, see Section 4).*

3. Our Importance Sampling Scheme

The core idea behind our new sampling scheme is to use the distribution of visible normals D_{ω_i} as the sampling PDF rather than the view-independent PDF $(\omega_m \cdot \omega_g)\,D(\omega_m)$, as shown in Algorithms 2. The differences with the previous method are highlighted in red. By doing this we are directly simulating what is illustrated in Figure 3(b): an incident ray ω_i intersects a random normal chosen within the distribution of visible normals for this ray and is reflected or transmitted. Only shadowing needs to be added, since it does not directly emerge from this model but is there to enforce single scattering on the microsurface [Hei14]. The value of the weight $\frac{G_2(\omega_m,\omega_i,\omega_g)}{G_1(\omega_m,\omega_i)}$ is the amount of visible shadowing. Note that when masking and shadowing are not correlated, i.e. $G_2(\omega_i,\omega_o,\omega_m) = G_1(\omega_i,\omega_m)\,G_1(\omega_o,\omega_m)$, then the weight comes down to shadowing alone $G_2(\omega_i,\omega_o,\omega_m)/G_1(\omega_i,\omega_m) = G_1(\omega_o,\omega_m)$. This makes sense, because if masking and shadowing are not correlated, then the visible shadowing is simply the shadowing.

Table 2 compares the previous importance sampling method with our method for the V-cavity and the Smith microsurface models. Since $G_1 \leq G_2$ the weight in our method is always less than 1: firefly artifacts do not appear (see Figure 2). This can be observed in Figure 4. We also see that our method does not waste the sampling space like previous methods. This is because we never generate backfacing normals.

Sections 4 and 5 are dedicated to implement method **Sample** $D_{\omega_i}(\omega_m)$ for the V-cavity and the Smith microsurface models, respectively.

4. Sampling $D_{\omega_i}(\omega_m)$ with the V-cavity Model

Importance sampling $D_{\omega_i}(\omega_m)$ with the V-cavity model is summarized in Algorithm 3. We start by sampling from the PDF $(\omega_m \cdot \omega_g)\,D(\omega_m)$ as in the previous method. This gives us a random microsurface with dual normals ω_m and ω_m' (see Figure 11). Then, we choose randomly between ω_m or ω_m' using selection probabilities proportional to $\langle\omega_i,\omega_m\rangle$ and $\langle\omega_i,\omega_m'\rangle$ respectively, which correspond to their projected areas toward ω_i. This random choice, weighted by the projected area, is what transforms the PDF $(\omega_m \cdot \omega_g)D(\omega_m)$ into $D_{\omega_i}(\omega_m)$. In practice, we implement this by a random swap of ω_m by ω_m' with probability $\frac{\langle\omega_i,\omega_m'\rangle}{\langle\omega_i,\omega_m\rangle + \langle\omega_i,\omega_m'\rangle}$.

Algorithm 3 Sampling $D_{\omega_i}(\omega_m)$ (V-cavity model)

$\omega_m \leftarrow$ Sample $(\omega_m \cdot \omega_g)\,D(\omega_m)$ ▷ 1. choose microsurface
if $\mathcal{U} < \frac{\langle\omega_i,\omega_m'\rangle}{\langle\omega_i,\omega_m\rangle + \langle\omega_i,\omega_m'\rangle}$ **then** ▷ 2. choose normal
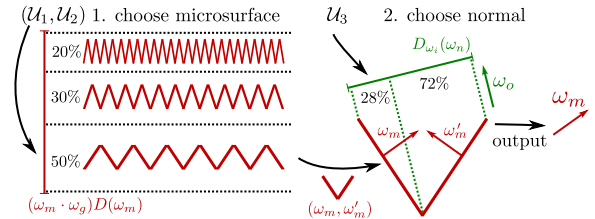$\quad\omega_m \leftarrow \omega_m'$
end if

Figure 5: *Illustration of Algorithm 3.*

Figure 5 illustrates how the algorithm works. When the incident ray approaches the geometric surface, it starts by choosing a V-cavity microsurface. Since the V-cavity model assumes separate microsurfaces, they have all the same visibility and can be chosen with the view-independent PDF $(\omega_m \cdot \omega_g)D(\omega_m)$. Then, on the chosen V-cavity microsurface, the ray must choose between ω_m and ω_m' with a probability weighted by their respective visibilities. We incorporate this second choice in the final PDF with a random swap between the two normals. Thus, Our algorithm requires an additional random number $\mathcal{U}$. The formal proof for this algorithm is given in Appendix B.

5. Sampling $D_{\omega_i}(\omega_m)$ with the Smith Model

In 5.1 we show that we can sample from the distribution of visible normals D_{ω_i} by sampling from the distribution of visible slopes $P_{\omega_i}^{22}$. Then, in 5.2, we will show that the shape of $P_{\omega_i}^{22}$ is stretch invariant. In practice, this means that a single analytic or tabulated inversion can be used with any roughness and anisotropy. These computations are detailed in 5.3. The complete algorithm we use to sample D_{ω_i} is summarized and discussed in 5.4.

5.1. Formulation in Slope Space

In Equation (2) we have already defined D_{ω_i}, but in this section we use an alternative equivalent formulation:

$$D_{\omega_i}(\omega_m) = \frac{\langle \omega_i, \omega_m \rangle D(\omega_m)}{\int_\Omega \langle \omega_i, \omega_m \rangle D(\omega_m) \, d\omega_m}$$
$$= \frac{1}{c} \langle \omega_i, \omega_m \rangle D(\omega_m), \qquad (5)$$

where $c = \int_\Omega \langle \omega_i, \omega_m \rangle D(\omega_m) \, d\omega_m$ is the normalization factor of the PDF. This normalization factor does not appear in Equation (2) but is implicitly contained in the masking function G_1 [Hei14]. In the following, we use Heitz's [Hei14] formulation of Equation (5) in slope space. We recall that $\tilde{m} = (x_{\tilde{m}}, y_{\tilde{m}}) = \left(-\frac{x_n}{z_n}, -\frac{y_n}{z_n}\right)$ are the slopes associated to the normal $\omega_m = (x_n, y_n, z_n)$ and reciprocally $\omega_m = \frac{(-x_{\tilde{m}}, -y_{\tilde{m}}, 1)}{\sqrt{x_{\tilde{m}}^2 + y_{\tilde{m}}^2 + 1}}$. If the microsurface is a heightfield, the distribution of normals is linked to the distribution of slopes $P^{22}(x_{\tilde{m}}, y_{\tilde{m}})$ by

$$(\omega_m \cdot \omega_g) D(\omega_m) = \left\| \frac{\partial \tilde{m}}{\partial \omega_m} \right\| P^{22}(\tilde{m} = (x_{\tilde{m}}, y_{\tilde{m}})),$$

where $\left\| \frac{\partial \tilde{m}}{\partial \omega_m} \right\| = \frac{1}{(\omega_m \cdot \omega_g)^3}$ is the Jacobian of the slope to normal transformation. The distributions of visible normals and of visible slopes are linked in the same way by

$$D_{\omega_i}(\omega_m) = \left\| \frac{\partial \tilde{m}}{\partial \omega_m} \right\| P_{\omega_i}^{22}(\tilde{m} = (x_{\tilde{m}}, y_{\tilde{m}})). \qquad (6)$$

In slope space, Equation (5) is written:

$$P_{\omega_i}^{22}(x_{\tilde{m}}, y_{\tilde{m}}) =$$
$$\frac{1}{c} \chi^+(x_i x_{\tilde{m}} + y_i y_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + y_i y_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}), \quad (7)$$

where the normalization factor in slope space is: $c = \int\int \chi^+(x_i x_{\tilde{m}} + y_i y_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + y_i y_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}) \, dx_{\tilde{m}} dy_{\tilde{m}}$.

→ If we sample $P_{\omega_i}^{22}$ and transform the output slope into a normal ω_m, the resulting PDF is D_{ω_i}. This is the first idea of Algorithm 4.

5.2. Stretch Invariance

Stretching a microsurface is illustrated in Figure 6. The slopes of the microsurface are divided by the stretching coefficients: $(x_{\tilde{m}}, y_{\tilde{m}})$ becomes $\left(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}\right)$. The stretching operator $\mathcal{S}^{\lambda_x, \lambda_y}(-)$ that transforms the normalized incident vector

$\omega_i = (x_i, y_i, z_i)$ into another normalized incident vector[‡] is defined as

$$\mathcal{S}^{\lambda_x, \lambda_y}(\omega_i) = \frac{(\lambda_x x_i, \lambda_y y_i, z_i)}{\sqrt{(\lambda_x x_i)^2 + (\lambda_y y_i)^2 + z_i^2}}.$$

Heitz [Hei14] calls P^{22} *shape-invariant* if it has the same shape after stretching, i.e. if for any $\lambda_x, \lambda_y > 0$

$$P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y) = \frac{1}{\lambda_x \lambda_y} P^{22}\left(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y}\right), \quad (8)$$

where $P^{22}(x_{\tilde{m}}, y_{\tilde{m}})$ is written $P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y)$ to specify the roughness parameters explicitly. This is usually the case when the distribution comes from a distribution of slopes. Heitz showed that if P^{22} is shape-invariant, then the masking function G_1 is stretch-invariant. We introduce a stronger property: *if the distribution of slopes P^{22} is shape-invariant, then the distribution of visible slopes $P_{\omega_i}^{22}$ is also shape-invariant.* This is written

$$P_{\omega_i}^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y) = \frac{1}{\lambda_x \lambda_y} P_{\mathcal{S}^{\lambda_x, \lambda_y}(\omega_i)}^{22}\left(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y}\right), \quad (9)$$

and is illustrated in Figure 7. The formal proof of this result is provided in Appendix A. A practical consequence of Equation (9) is that the distribution of visible slopes can always be expressed with an isotropic roughness α (we choose arbitrarily $\alpha = 1$):

$$P_{\omega_i}^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y) = \frac{1}{\alpha_x \alpha_y} P_{\mathcal{S}^{\alpha_x, \alpha_y}(\omega_i)}^{22}\left(\frac{x_{\tilde{m}}}{\alpha_x}, \frac{y_{\tilde{m}}}{\alpha_y}, 1, 1\right).$$
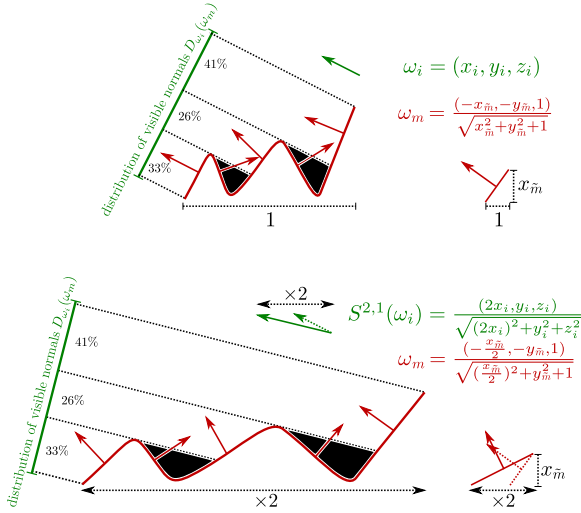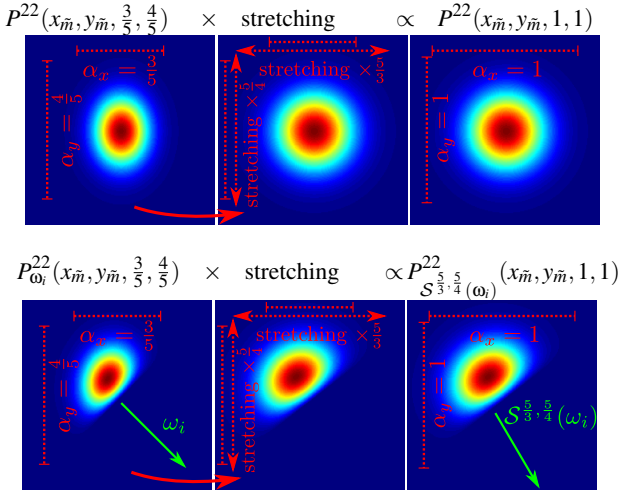
The factor $\frac{1}{\alpha_x \alpha_y}$ in the right-hand side is the Jacobian of the transformation from $\left(\frac{x_{\tilde{m}}}{\alpha_x}, \frac{y_{\tilde{m}}}{\alpha_y}\right)$ to $(x_{\tilde{m}}, y_{\tilde{m}})$.

→ If a sampling algorithm is available for $\left(\frac{x_{\tilde{m}}}{\alpha_x}, \frac{y_{\tilde{m}}}{\alpha_y}\right)$ with PDF $P_{\mathcal{S}^{\alpha_x, \alpha_y}(\omega_i)}^{22}\left(\frac{x_{\tilde{m}}}{\alpha_x}, \frac{y_{\tilde{m}}}{\alpha_y}, 1, 1\right)$, we can use it to sample $(x_{\tilde{m}}, y_{\tilde{m}})$ with PDF $P_{\omega_i}^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y)$. This is the second idea of Algorithm 4.

We note that a similar observation was made by Becker and Max [BM93]: they show that the distribution of visible normals is invariant to height-stretching, which is equivalent to scaling the slope amplitude $\|\tilde{m}\|$, where the slope coordinates $x_{\tilde{m}}$ and $y_{\tilde{m}}$ are scaled by the same factor. This allows for isotropic roughness changes only. Instead of height-stretching we use slope-stretching, where $x_{\tilde{m}}$ and $y_{\tilde{m}}$ can be scaled separately, which permits fully anisotropic roughness changes.

[‡] Note that the normals are not modified in the same way because they are not vectors but covectors. The normals after stretching are given by the new slopes: $\omega_m = \frac{\left(-\frac{x_{\tilde{m}}}{\lambda_x}, -\frac{y_{\tilde{m}}}{\lambda_y}, 1\right)}{\sqrt{\left(\frac{x_{\tilde{m}}}{\lambda_x}\right)^2 + \left(\frac{y_{\tilde{m}}}{\lambda_y}\right)^2 + 1}}$.

Figure 6: *Stretching the configuration.*

$P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \frac{3}{5}, \frac{4}{5}) \quad \times \quad \text{stretching} \quad \propto \quad P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$

$P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, \frac{3}{5}, \frac{4}{5}) \quad \times \quad \text{stretching} \quad \propto P^{22}_{\mathcal{S}^{\frac{5}{3}, \frac{5}{4}}(\omega_i)}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$

Figure 7: *Stretch invariance in slope space of P^{22} and $P^{22}_{\omega_i}$ (Beckmann distribution).*

5.3. Sampling $P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$

After stretching the configuration, the distribution of visible slopes is $P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$. In order to sample $(x_{\tilde{m}}, y_{\tilde{m}})$ with PDF $P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$, we first assume that the frame is such as $\omega_i = (x_i, 0, z_i)$. This is possible since we are now using an isotropic distribution of normals ($\alpha_x, \alpha_y = 1$). The expression of Equation (7) simplifies to

$$P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1) = \frac{1}{c} \chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}).$$

This expression emphasizes that the projection factor $\chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i)$ is not related to the slope component $y_{\tilde{m}}$. We use this property to sample $P^{22}_{\omega_i}$ in two steps a. and b.

5.3.1. a. Sampling $P^{2-}_{\omega_i}(x_{\tilde{m}}, 1, 1)$

We start by sampling the marginalized random variable $x_{\tilde{m}}$ given by the PDF

$$
\begin{aligned}
& P^{2-}_{\omega_i}(x_{\tilde{m}}, 1, 1) \\
&= \frac{1}{c} \int_{-\infty}^{+\infty} \chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1) \, dy_{\tilde{m}} \\
&= \frac{1}{c} \chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) \int_{-\infty}^{+\infty} P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1) \, dy_{\tilde{m}} \\
&= \frac{1}{c} \chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) P^{2-}(x_{\tilde{m}}, 1, 1), \qquad (10)
\end{aligned}
$$

where $P^{2-}(x_{\tilde{m}}, 1, 1)$ is the marginalized PDF of $x_{\tilde{m}}$. We invert

$$
\begin{aligned}
\mathcal{U}_1 &= \int_{-\infty}^{x_{\tilde{m}}} P^{2-}_{\omega_i}(x'_{\tilde{m}}, 1, 1) \, dx'_{\tilde{m}} \\
&\Rightarrow x_{\tilde{m}} = C^{2-^{-1}}_{\omega_i}(\mathcal{U}_1, 1, 1), \qquad (11)
\end{aligned}
$$

where $C^{2-}_{\omega_i}$ is the cumulative distribution fonction (CDF) associated to the marginalized PDF $P^{2-}_{\omega_i}$.

This equation can be solved analytically for Beckmann and GGX distributions (see our supplemental material). For other distributions, if an analytic solution is not available, we propose to precompute the solutions and store them in a 2D table $x_{\tilde{m}} = T_{x_{\tilde{m}}}[\theta_i, \mathcal{U}_1]$. Indeed, $x_{\tilde{m}}$ depends only on two variables: θ_i (because $x_i = \sin \theta_i$ and $z_i = \cos \theta_i$) and $\mathcal{U}_1$.

5.3.2. b. Sampling $P^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1)$

Then, we sample the conditional random variable $y_{\tilde{m}}|x_{\tilde{m}}$ given by the PDF

$$
\begin{aligned}
& P^{2|2}_{\omega_i}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1) \\
&= \frac{\frac{1}{c}\chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)}{\int_{-\infty}^{+\infty} \frac{1}{c}\chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y'_{\tilde{m}}, 1, 1) \, dy'_{\tilde{m}}} \\
&= \frac{P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)}{\int_{-\infty}^{+\infty} P^{22}(x_{\tilde{m}}, y'_{\tilde{m}}, 1, 1) \, dy'_{\tilde{m}}} = P^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1). \quad (12)
\end{aligned}
$$

This shows that the sampling of $y_{\tilde{m}}$ does not depend on the incident direction ω_i. Hence, we invert

$$
\begin{aligned}
\mathcal{U}_2 &= \int_{-\infty}^{y_{\tilde{m}}} P^{2|2}(y'_{\tilde{m}}|x_{\tilde{m}}, 1, 1) \, dy'_{\tilde{m}} \\
&\Rightarrow y_{\tilde{m}} = C^{2|2^{-1}}(\mathcal{U}_2|x_{\tilde{m}}, 1, 1), \qquad (13)
\end{aligned}
$$

where $C^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1)$ the CDF associated to the view-independent conditional PDF $P^{2|2}$.

This equation can be solved analytically for Beckmann and GGX distributions (see our supplemental material). For other distributions, if an analytic solution is not available, we propose to precompute the solutions and store them in a 2D table $y_{\tilde{m}} = T_{y_{\tilde{m}}}[x_{\tilde{m}}, \mathcal{U}_2]$. Indeed, $y_{\tilde{m}}|x_{\tilde{m}}$ depends only on two variables: $x_{\tilde{m}}$ and $\mathcal{U}_2$.

5.4. Sampling $D_{\omega_i}(\omega_m)$

In Algorithm [4](#), we sample $(x_{\tilde{m}}, y_{\tilde{m}})$ with PDF $P^{22}_{\mathcal{S}^{\alpha_x,\alpha_y}(\omega_i)}(x_{\tilde{m}}, y_{\tilde{m}}, 1, 1)$ by sampling $x_{\tilde{m}}$ with marginalized PDF $P^{2-}_{\omega_i}(x_{\tilde{m}}, 1, 1)$ and $y_{\tilde{m}}$ with conditional PDF $P^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1)$. Then, we transform this slope vector back to the canonical frame where $\omega_i = (\cos\phi_i \sin\theta_i, \sin\phi_i \sin\theta_i, \cos\theta_i)$ by applying a rotation of angle ϕ_i and we unstretch the slopes by multiplying by respectively α_x and α_y. The sampled PDF at this point is $P^{22}_{\omega_i}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y)$. Finally, we transform the slopes into a normal. The sampled PDF is $D_{\omega_i}(\omega_m, \alpha_x, \alpha_y)$.

Algorithm 4 Sample $D_{\omega_i}(\omega_m)$ (Smith profile)

$\omega_i' \leftarrow \mathcal{S}^{\alpha_x,\alpha_y}(\omega_i)$ ▷ 1. stretch

$x_{\tilde{m}} \leftarrow C^{2-\ -1}_{\omega_i'}(\mathcal{U}_1, 1, 1)$ ▷ 2.a. sample $P^{2-}_{\omega_i'}(x_{\tilde{m}}, 1, 1)$

$y_{\tilde{m}} \leftarrow C^{2|2\ -1}(\mathcal{U}_2|x_{\tilde{m}}, 1, 1)$ ▷ 2.b. sample $P^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1)$

$\begin{pmatrix} x_{\tilde{m}} \\ y_{\tilde{m}} \end{pmatrix} \leftarrow \begin{pmatrix} \cos\phi_i' & -\sin\phi_i' \\ \sin\phi_i' & \cos\phi_i' \end{pmatrix} \begin{pmatrix} x_{\tilde{m}} \\ y_{\tilde{m}} \end{pmatrix}$ ▷ 3. rotate

$\begin{pmatrix} x_{\tilde{m}} \\ y_{\tilde{m}} \end{pmatrix} \leftarrow \begin{pmatrix} \alpha_x x_{\tilde{m}} \\ \alpha_y y_{\tilde{m}} \end{pmatrix}$ ▷ 4. unstretch

$\omega_m \leftarrow \frac{(-x_{\tilde{m}}, -y_{\tilde{m}}, 1)}{\sqrt{x_{\tilde{m}}^2 + y_{\tilde{m}}^2 + 1}}$ ▷ 5. compute normal
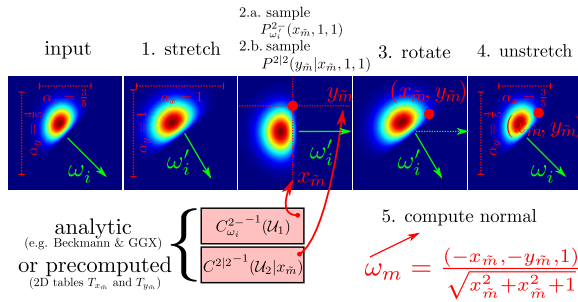

Figure 8: *Illustration of Algorithm [4](#).*

Analytic Computation of $C^{2-\ -1}_{\omega_i}$ and $C^{2|2\ -1}$ The cornerstone of Algorithm [4](#) is the inversion of the CDFs $C^{2-}_{\omega_i}$ and $C^{2|2}$. We found this to be possible analytically for Beckmann and GGX distributions, for which we provide the detailed derivations of the inversion procedures and their C++ implementations in the associated supplemental material.

Precomputation of $C^{2-\ -1}_{\omega_i}$ and $C^{2|2\ -1}$ If an analytic inversion procedure of the CDFs $C^{2-}_{\omega_i}$ and $C^{2|2}$ is not available, we propose to use precomputed data $T_{x_{\tilde{m}}}[\theta_i, \mathcal{U}_1]$ and $T_{y_{\tilde{m}}}[x_{\tilde{m}}, \mathcal{U}_2]$ parametrized in slope space. One advantage of the slope space is that the equation of $P^{22}_{\omega_i}$ becomes separable: $\chi^+(x_i x_{\tilde{m}} + z_i)(x_i x_{\tilde{m}} + z_i)$ does not depend on y_i and thus can be pulled out one of the two integrals. This property is used in Equations ([10](#)) and ([12](#)) and makes the precomputations easier than they would have been with spherical coordinates. The problem of using spherical coordinates

was discussed by Becker and Max [BM93]: they store tables parametrized in (θ, ϕ) and they either assume that θ and ϕ can be sampled independently, which is approximate and leads to biased sampling, or they need to store a 3D table to sample θ for a given ϕ (or the reciprocal). This would require an enormous 3D table and would be impractical (>1GB). However, in slope space, Equation ([12](#)) shows that $P^{2|2}_{\omega_i}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1) = P^{2|2}(y_{\tilde{m}}|x_{\tilde{m}}, 1, 1)$ and the inversion does not depend on θ_i, which allows us to use a 2D precomputed table indexed by $x_{\tilde{m}}$ and $\mathcal{U}_2$. Thus, we need to precompute only two 2D tables $T_{x_{\tilde{m}}}[\theta_i, \mathcal{U}_1]$ and $T_{y_{\tilde{m}}}[x_{\tilde{m}}, \mathcal{U}_2]$. More importantly, thanks to the invariance property in slope space, we can use the same data with different roughnesses and anisotropy. In our implementation, $T_{x_{\tilde{m}}}$ and $T_{y_{\tilde{m}}}$ are both of size 1024^2. Popular heavy-tailed distributions of slopes like GGX have very long tails but also a very sharp peak at the same time, and so require a 1024^2 table to capture them well. The total memory footprint for every parametric anisotropic BSDF with the same shape is then 32MB.

Limitation Note that this technique works only when the distribution of normals D is built from a shape-invariant distribution of slopes P^{22}. For instance, this is the case for the Beckmann [Smi67] and GGX [WMLT07], but not for the Phong distribution (which lacks a Smith masking function [WMLT07]).

6. Results and Validation

In Figure [1](#) we show a dielectric glass plate with anisotropic GGX roughness (Smith masking) on all faces. The illuminant is constant uniform distant white radiance from all directions (a white environment map). In all renders we use forward path tracing with unbounded path lengths. Multiple importance sampling (MIS) mixes BSDF and next-event-estimation (light sampling) at each path vertex. Figure [1](#) uses 64 independent samples per pixel. Our new BSDF sampling scheme runs in roughly the same time as prior work and provides a significant reduction in variance. Figure [9](#) shows a configuration where fireflies are still present with the previous method even with 4000 samples per pixel. Due to the default Gaussian filter in Mitsuba, a single bright sample may contribute significantly to several pixels (appearing as a 2x2 bright pixel block).

Figure [10](#) compares the convergence of importance sampling with the previous method to our method in a configuration shown in Figure [2](#), i.e. a rough conductor lit by an environment map. The plot shows that both methods converge to the same value, which was expected because the previous method and our method are both mathematically well-defined. However, the convergence rate is not the same.

In the two first rows we used the Smith model. We see that the convergence rate is very sensitive to the fact that the sampling weight can be arbitrary high with the previous method. This produces firefly artifacts that take a long time

to disappear. In contrary, our method does not produce these artifacts since all BSDF sample weights are bounded to 1. In the two last rows we compare the methods used with the V-cavity model. While the previous method does not suffer from very high sampling weights (they are bounded to 2), the convergence of our method is faster. This is because we do not waste the sampling space by generating backfacing normals with weight 0.

More results and more convergence comparisons are available in the associated supplemental material.
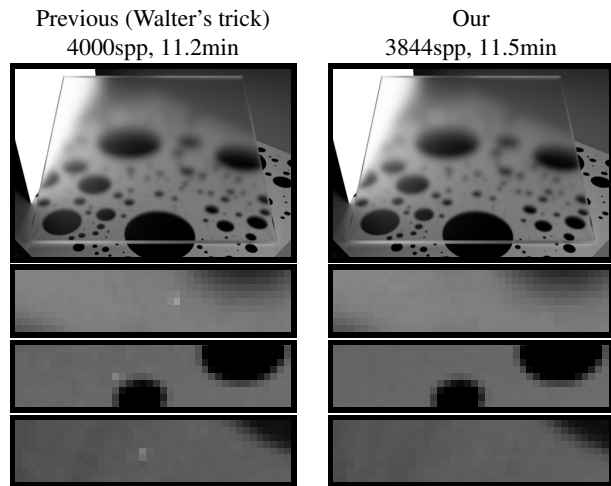

Figure 9: *A dielectric glass plate (n = 1.5) with isotropic Beckmann roughness (α = 0.1) on all faces (with the Smith masking function). Fireflies are persistent with the previous method even with a large number of samples (4000 per pixel).*

7. Conclusion

We have presented a new importance sampling scheme for microfacet-based BSDFs. The scope of our scheme is broad, encompassing many popular isotropic and anisotropic analytic BRDF and BSDF models and allows a variety of normal distribution functions. The performance of our scheme is comparable to previous methods and is suitable for both offline and GPU rendering. We have shown that importance sampling the BSDF using the distribution of visible normals rather than the distribution of normals is practical and makes convergence faster, significantly reducing artifacts at grazing incidence angles, especially for significant roughness levels. With the V-cavity microsurface model, this is implemented with a simple one conditional swap of the normal. With the much more accurate Smith microsurface model, we provide analytic sampling algorithms for Beckmann and GGX distributions in our supplemental material. Otherwise, we have shown that a small set of precomputed data can be used to sample BSDFs with any roughness and anisotropy, due to key shape invariance properties of the distribution of visible slopes.
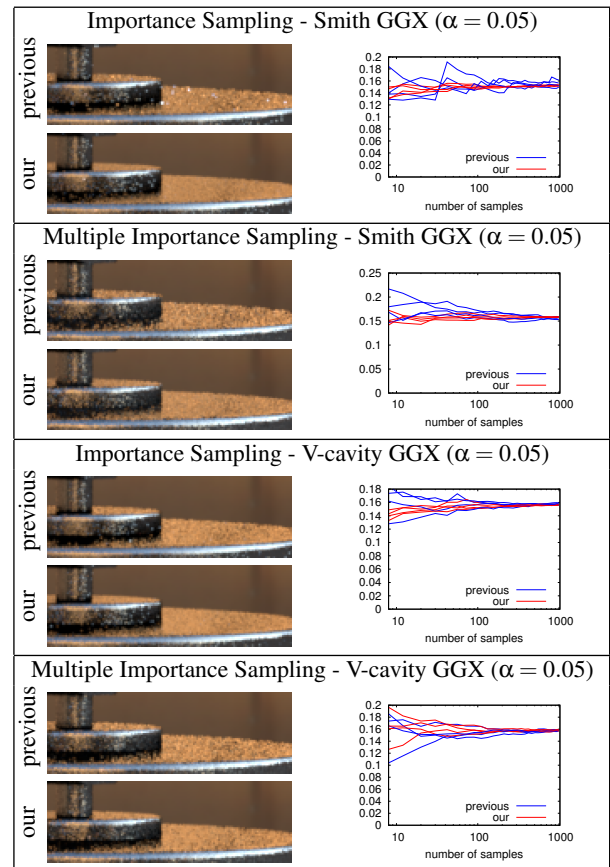

Figure 10: *Comparisons of our importance sampling technique to previous methods (with and without MIS). The images are computed with 8 samples per pixel. The plots shows the convergence of several running estimates of the same pixel of the pedestal for both techniques. The pixel variance is seen directly in the spread of the various estimates.*

References

[APS00] ASHIKMIN M., PREMOŽE S., SHIRLEY P.: A microfacet-based brdf generator. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (2000), SIGGRAPH '00, pp. 65–74. 4

[AS00] ASHIKHMIN M., SHIRLEY P.: An anisotropic phong brdf model. *Journal of Graphics Tools 5* (2000), 25–32. 2, 4

[Ash07] ASHIKHMIN M.: *Distribution-based BRDFs*. Tech. rep., 2007. 4

[BM93] BECKER B. G., MAX N. L.: Smooth transitions between bump rendering algorithms. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques* (1993), SIGGRAPH '93, pp. 183–190. 6, 8

[Bro80] BROWN G.: Shadowing by non-gaussian random surfaces. *IEEE Trans. on Antennas and Propagation 28*, 6 (Nov 1980), 788–790. 3

[CT82] COOK R. L., TORRANCE K. E.: A reflectance model for

computer graphics. *ACM Trans. Graph. 1*, 1 (Jan. 1982), 7–24. 2, 3

[Hei14] HEITZ E.: Understanding the masking-shadowing function in microfacet-based brdfs. *Journal of Computer Graphics Techniques (JCGT) 3*, 2 (2014). To appear. 2, 3, 5, 6

[KSK01] KELEMEN C., SZIRMAY-KALOS L.: A microfacet based coupled specular-matte brdf model with importance sampling. In *Eurographics Short Presentations* (2001). 2, 3, 4

[Smi67] SMITH B.: Geometrical shadowing of a random rough surface. *IEEE Trans. on Antennas and Propagation 15* (1967), 668–671. 3, 8

[Sta01] STAM J.: An illumination model for a skin layer bounded by rough surfaces. In *Rendering Techniques* (2001), pp. 39–52. 2

[TS67] TORRANCE K. E., SPARROW E. M.: Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am. 57*, 9 (Sep 1967), 1105–1112. 2, 3

[vGSK98] VAN GINNEKEN B., STAVRIDI M., KOENDERINK J. J.: Diffuse and specular reflectance from rough surfaces. *Journal of Appl. Opt. 37*, 1 (Jan 1998), 130–139. 2

[War92] WARD G. J.: Measuring and modeling anisotropic reflection. *SIGGRAPH Comput. Graph. 26*, 2 (July 1992), 265–272. 2, 4

[WMLT07] WALTER B., MARSCHNER S. R., LI H., TORRANCE K. E.: Microfacet models for refraction through rough surfaces. In *Proc. Eurographics Symposium on Rendering* (2007), EGSR'07, pp. 195–206. 2, 3, 4, 8

Appendix A: Shape Invariance of $P_{\omega_i}^{22}$

We demonstrate the result of Equation (9):

$$P_{\omega_i}^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y)$$

$$= \frac{\chi^+(-x_i x_{\tilde{m}} - y_i y_{\tilde{m}} + z_i)(-x_i x_{\tilde{m}} - y_i y_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y)}{\iint \chi^+(-x_i x_{\tilde{m}} - y_i y_{\tilde{m}} + z_i)(-x_i x_{\tilde{m}} - y_i y_{\tilde{m}} + z_i) P^{22}(x_{\tilde{m}}, y_{\tilde{m}}, \alpha_x, \alpha_y) \, dx_{\tilde{m}} dy_{\tilde{m}}}$$

$$= \frac{\frac{\chi^+(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)}{\sqrt{(\lambda_x y_i)^2 + (\alpha_y y_i)^2 + z_i^2}} \frac{1}{\lambda_x \lambda_y} P^{22}(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y})}{\iint \frac{\chi^+(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)}{\sqrt{(\lambda_x y_i)^2 + (\lambda_y y_i)^2 + z_i^2}} \frac{1}{\lambda_x \lambda_y} P^{22}(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y}) \, dx_{\tilde{m}} dy_{\tilde{m}}}$$

$$= \frac{\frac{\chi^+(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)}{\sqrt{(\lambda_x y_i)^2 + (\alpha_y y_i)^2 + z_i^2}} \frac{1}{\lambda_x \lambda_y} P^{22}(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y})}{\iint \frac{\chi^+(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)(-\lambda_x x_i \frac{x_{\tilde{m}}}{\lambda_x} - \lambda_y y_i \frac{y_{\tilde{m}}}{\lambda_y} + z_i)}{\sqrt{(\lambda_x y_i)^2 + (\lambda_y y_i)^2 + z_i^2}} P^{22}(\frac{y_{\tilde{m}}}{\lambda_y}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y}) \, d\frac{x_{\tilde{m}}}{\lambda_x} d\frac{y_{\tilde{m}}}{\lambda_y}}$$

$$= \frac{1}{\lambda_x \lambda_y} P_{S^{\lambda_x,\lambda_y}(\omega_i)}^{22}\left(\frac{x_{\tilde{m}}}{\lambda_x}, \frac{y_{\tilde{m}}}{\lambda_y}, \frac{\alpha_x}{\lambda_x}, \frac{\alpha_y}{\lambda_y}\right),$$

where we start to write Equation (7). In the second step we develop by applying Equation (8) and by multiplying the numerator and the denominator by $\frac{1}{\sqrt{(\lambda_x x_i)^2 + (\alpha_y y_i)^2 + z_i^2}}$. In the thrid step, at the denominator we make the change of variable: $\frac{1}{\lambda_x \lambda_y} dx_{\tilde{m}} dy_{\tilde{m}} = d\frac{x_{\tilde{m}}}{\lambda_x} d\frac{y_{\tilde{m}}}{\lambda_y}$.

Appendix B: Proof for Algorithm 3

To verify that Algorithm 3 samples PDF $D_{\omega_i}(\omega_m)$, we will expand G_1 from Equation (3) in the definition of $D_{\omega_i}(\omega_m)$ from Equation (2). The min in Equation (3) is used to differentiate between 2 cases illustrated in Figure 11.

Case (a) If ω_m' is backfacing, then the shadowing occuring on ω_m is $G_1(\omega_i, \omega_m) = 2\frac{(\omega_m \cdot \omega_g)(\omega_i \cdot \omega_g)}{\langle \omega_i, \omega_m \rangle}$, and by using the fact that $\langle \omega_i, \omega_m' \rangle = 0$ and $|\omega_i \cdot \omega_m| > 0$ we get

$$D_{\omega_i}(\omega_m) = 2\frac{|\omega_m \cdot \omega_g||\omega_i \cdot \omega_g|}{|\omega_i, \omega_m|} \frac{|\omega_i, \omega_m| D(\omega_m)}{|\omega_i \cdot \omega_g|}$$

$$= 2\frac{|\omega_i \cdot \omega_m|}{|\omega_i \cdot \omega_m|} |\omega_m \cdot \omega_g| D(\omega_m)$$

$$= 2\frac{\langle \omega_i, \omega_m \rangle}{\langle \omega_i, \omega_m \rangle + 0} |\omega_m \cdot \omega_g| D(\omega_m)$$

$$= 2\frac{\langle \omega_i, \omega_m \rangle}{\langle \omega_i, \omega_m \rangle + \langle \omega_i, \omega_m' \rangle} (\omega_m \cdot \omega_g) D(\omega_m).$$

Case (b) If ω_m and ω_m' are both not backfacing for ω_i, then there is no masking ($G_1 = 1$) and Equation (2) becomes

$$D_{\omega_i}(\omega_m) = \frac{\langle \omega_i, \omega_m \rangle D(\omega_m)}{|\omega_i \cdot \omega_g|}$$

$$= 2\frac{\langle \omega_i, \omega_m \rangle}{\langle \omega_i, \omega_m \rangle + \langle \omega_i, \omega_m' \rangle} (\omega_m \cdot \omega_g) D(\omega_m),$$

because $|\omega_i \cdot \omega_g| = \frac{\langle \omega_i, \omega_m \rangle + \langle \omega_i, \omega_m' \rangle}{2\langle \omega_m, \omega_g \rangle}$.

Conclusion In either case the distribution of visible normals is given by the same expression. The factor $\frac{\langle \omega_i, \omega_m \rangle}{\langle \omega_i, \omega_m \rangle + \langle \omega_i, \omega_m' \rangle}$ represents the repartition of visibility between ω_m and ω_m'. In Algorithm 3, the conditional replacement of ω_m by ω_m' multiplies the original PDF ($\omega_m \cdot \omega_g) D(\omega_m)$ by the factor $\frac{\langle \omega_i, \omega_m \rangle}{\langle \omega_i, \omega_m \rangle + \langle \omega_i, \omega_m' \rangle}$, and the factor 2 is present in the sampled PDF because ω_m can be chosen for 2 microsurfaces (the one associated to himself and the one to ω_m') and reciprocally for ω_m'.
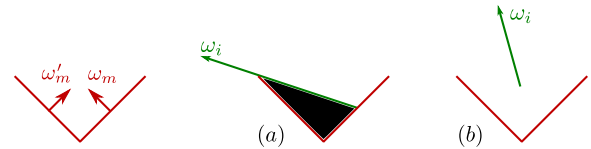

Figure 11: *Masking configurations with the V-cavity model. Either one microfacet is backfacing and completely masked (a) or the two microfacets are completely visible (b).*