Eurographics Symposium on Rendering 2019
T. Boubekeur and P. Sen
(Guest Editors)

Volume 38 (2019), Number 4
Author's Version

Microfacet Model Regularization for Robust Light Transport

Johannes Jendersie and Thorsten Grosch

TU Clausthal, Germany

Path Tracing (PT), 1000 spp | Bidirectional Path Tracing (BPT), 1000 spp | Reference

Figure 1: *Wrist watch with many glossy interactions (model courtesy of heraSK). While still noisy, both regularized variants are much more capable of rendering difficult light paths. The reference was rendered with Vertex Connection and Merging using 155k spp.*

Abstract

Today, Monte Carlo light transport algorithms are used in many applications to render realistic images. Depending on the complexity of the used methods, several light effects can or cannot be found by the sampling process. Especially, specular and smooth glossy surfaces often lead to high noise and missing light effects.
Path space regularization provides a solution, improving any sampling algorithm, by modifying the material evaluation code. Previously, Kaplanyan and Dachsbacher [KD13] introduced the concept for pure specular interactions. We extend this idea to the commonly used microfacet models by manipulating the roughness parameter prior to the evaluation. We also show that this kind of regularization requires a change in the MIS weight computation and provide the solution. Finally, we propose two heuristics to adaptively reduce the introduced bias.
Using our method, many complex light effects are reproduced and the fidelity of smooth objects is increased. Additionally, if a path was sampleable before, the variance is partially reduced.

CCS Concepts

• ***Computing methodologies*** → *Ray tracing;* • ***Mathematics of computing*** → *Sequential Monte Carlo methods;*

1. Introduction

While being around for a long time, Monte Carlo sampling methods become more and more important in computer graphics, because they are used in a growing number of application areas. On the one hand increasingly complex methods like *Vertex Connection and Merging* (VCM) [GKDS12, HPJ12] are developed which can handle increasingly difficult light paths. On the other hand production rendering and interactive applications still rely on the simpler algorithms like path tracing, because of their lower overhead.

Path space regularization [KD13] helps to find difficult light path situations in any of the rendering methods. By allowing a biased result, non-sampleable or high variance paths can be sampled with a lower variance. While the work of Kaplanyan and Dachsbacher [KD13] focused on the otherwise infeasible specular events, this paper extends the idea to general microfacet models. We distinguish specular events (deterministic reflections/refractions) and glossy events (randomly disturbed/rough specular). We summarize the difficult smooth glossy and specular vertices as *shiny* in the following. Paths including glossy events are in general sampleable, but may result in a very high variance. By smoothing the material evaluation we can reduce this variance. Further, regularization is able to produce caustics and pure *shiny* paths in simple rendering methods like *Path Tracing* (PT) – at the cost of blurriness – which would not be found otherwise.

Our approach is to invert the *Bidirectional Scattering Distribution Function* (BSDF) of common microfacet models to determine

This is the accepted version of: Microfacet Model Regularization for Robust Light Transport, Computer Graphics Forum (Proc. EGSR) 38, 4 (2019), pp. 39-47, which has been published in final form at https://onlinelibrary.wiley.com/doi/10.1111/cgf.13768. This article may be used for non-commercial purposes in accordance with the Wiley Self-Archiving Policy.

a roughness parameter such that the evaluation in any possible direction will result in at most a chosen threshold. This change of parametrization will be applied whenever a connection (in the sense of bidirectional path tracing [VG95]) is established. For an already sufficiently rough surface the regularization will not change the parameter and thus limit the bias to difficult paths only. Also, we show how the *Multiple Importance Sampling* (MIS) weight changes according to the replacement of BSDFs. Further, we introduce two heuristics to fully avoid regularization if possible and to reduce the visible bias otherwise. Our contributions are:

- Application of path space regularization to microfacet models
- An MIS weight in the presence of regularization
- Two heuristics to reduce the bias adaptively
- Evaluation of two regularization strategies:
  roughness parameter (*mollification*) $\leftrightarrow$ *virtual merge*

2. Related Work

The problem of light transport can be described by an integral equation – the rendering equation [Kaj86]. Monte Carlo methods solve this equation numerically by stochastic sampling. The first solution, *Path Tracing* (PT) by Kajiya [Kaj86], produced paths starting at the observer only. Due to its simplicity it is widely used in production rendering [FHF*17]. However, it is unable to sample any path containing a number of specular events before reaching an infinitesimal small light source. The same applies to paths with glossy events and small (non-infinitesimal) light sources which can be found, but produce an impractical amount of noise.

A possible solution is *Manifold Next Event Estimation* (MNEE) [HDF15] where a connection to a light source passing through *shiny* surfaces is iteratively perturbed on the surfaces until the correct path is found or the process is canceled. Regularization, as described here, avoids the expensive iterative process.

Other than improving PT there are also more complex rendering techniques involving paths from the observer and the light source. Bidirectional Path Tracing (BPT) [VG95] does so by creating any possible connection between a light- and a view-subpath. Unlike PT it is able to find caustics (a diffuse vertex on the view-path followed by *shiny* vertices). On the other hand, BPT still fails for *shiny-diffuse-shiny* paths. When allowing a small bias these paths can be found by merging the end vertices of the two involved caustic paths, as done by *Photon Mapping* (PM) [Jen96]. *Photon Mapping* can be seen as a regularization in the spatial domain. Using *Progressive Photon Mapping* [HOJ08, HJ09] it is possible to converge to the correct result over time (*consistent*). Knaus and Zwicker [KZ11] simplify the implementation of a *consistent* photon mapper by applying global convergence statistics. Finally, VCM [GKDS12, HPJ12] combines PM with BPT and therefore increases the robustness further. Unfortunately, even VCM is unable to find pure *shiny* paths and can thus still profit from regularization.

Another family of Monte Carlo integration methods uses Markov Chains (MCMC) which converge to an unknown density function by stochastic decisions. The first such rendering method was *Metropolis Light Transport* [VG97]. Since then a number of improvements were made, including the combination of MLT and VCM by Šik et al. [ŠOHK16]. More about MCMC in general can be found in a recent survey from Šik and Křivánek [ŠK18]. However, MLT methods often get stuck in local optima and have similar difficulties to find *shiny* paths as the other methods. Regularization helps those methods in finding the complex paths, which was the original intention of regularization [KD13].

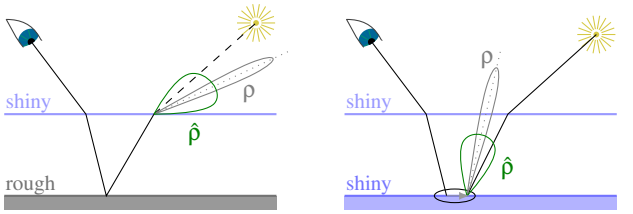


Figure 2: *Example paths with regularization of a connection (left) and a merge (right). By smoothing the BSDF ρ into $\hat{\rho}$, a larger set of paths has a non-zero contribution.*

Bouchard et al. [BIOP13] used the same regularization as Kaplanyan and Dachsbacher but introduced a custom MIS weight to select between unbiased and biased samplers. Our non-adaptive regularization generalizes this for microfacet models while our heuristics reduce the bias further.

We believe that regularization approaches and guided sampling methods [HEV*16, MGN17] benefit from each other. The idea of this methods is to learn the radiance and importance distribution in the scene to improve the local sampling quality. While regularization is strong if the found paths are close to the actual solution, guidance helps to generate these close paths more often.

3. Roughness-based BSDF Regularization

The idea of regularization is to remove difficult-to-sample light transport paths by softening *shiny* BSDFs in a mostly energy-preserving way. This can be achieved by controlling the roughness parameter α of common microfacet models with a *Bidirectional Scattering Distribution Function* (BSDF) $\rho(\alpha, \overleftarrow{\mathbf{w}}, \vec{\mathbf{w}})$, where $\overleftarrow{\mathbf{w}}$ is the incident and $\vec{\mathbf{w}}$ the exitant direction of evaluation.

Setting a maximum evaluation value τ, we can invert an upper bound $\bar{\rho}(\alpha) \geq \rho(\alpha, \overleftarrow{\mathbf{w}}, \vec{\mathbf{w}})$ on the BSDF, to obtain a threshold for the roughness parameter

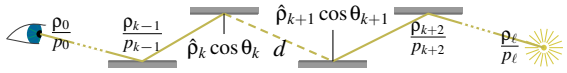
$$\hat{\alpha} = \bar{\rho}^{-1}(\tau). \tag{1}$$

Then, ρ is evaluated as $\hat{\rho} = \rho(\max(\alpha, \hat{\alpha}), \overleftarrow{\mathbf{w}}, \vec{\mathbf{w}})$, which applies regularization only if the model was too *shiny* before. Rougher surfaces keep their appearance (use the original α). For anisotropic models each of the two roughness parameters can be bounded independently. While it is be possible to maintain a constant ratio of the two parameters, this would only introduce unnecessary bias for the larger of the two parameters.

We apply the modification of the roughness parameter only if a connection or a merge is evaluated, as shown in Figure 2. During the random walk, as shown below, we use the original BSDF and *Probability Density Function* (PDF) without modification. This avoids the introduction of an additional bias where it would not be

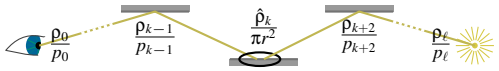
necessary. The samples of the resulting estimator for connection paths of length ℓ have the form

$$\hat{I}_C \;:\; \prod_{i=0}^{k-1} \frac{\rho_i}{p_i} \cdot \frac{\hat{\rho}_k \cos\theta_k \cos\theta_{k+1} \hat{\rho}_{k+1}}{d^2} \cdot \prod_{i=k+2}^{\ell} \frac{\rho_i}{p_i} \tag{2}$$



and for merge paths:

$$\hat{I}_M \;:\; \prod_{i=0}^{k-1} \frac{\rho_i}{p_i} \cdot \frac{\hat{\rho}_k}{\pi r^2} \cdot \prod_{i=k+1}^{\ell} \frac{\rho_i}{p_i}. \tag{3}$$



Thereby, $\rho_i/p_i = \rho(\alpha_i, \overleftarrow{\mathbf{w}}_i, \vec{\mathbf{w}}_i)/p(\alpha_i, \vec{\mathbf{w}}_i)$ are the Monte Carlo sampling events in the two sub-paths, $\hat{\rho}$ is the regularized BSDF, d^2 is distance between the two connected vertices and $1/\pi r^2$ is the kernel for the photon merge as used in VCM. For brevity we used the symbol ρ for BSDFs, camera response and light radiance, allowing to put the first and last term into the products.

All methods like PT, BPT and VCM are described by different combinations of the above two sampler types with the addition of the random hit sampler. For example, a PT combines random hits with connections towards the light sources ($k = \ell - 2$ in Equation (2)), also called next event estimation.

3.1. Variance with and without Regularization

Modifying the estimators in the proposed way (Equations (2) and (3)) limits the *relative* variance of a sample. Formally, we can examine the variance of the estimators without $V[I]$ and with regularization $V[\hat{I}]$ to understand the effect of regularization. For independent random experiments X and Y we have $V[XY] = E[X]^2V[Y] + E[Y]^2V[X] + V[X]V[Y]$ as the joint variance of the product. This allows us to factorize our estimators into two parts: one containing all Monte Carlo events ρ_i/p_i and one for the central term.

If $\rho_i \approx c_i \cdot p_i$ for some constant factor c_i, which is the case for most BSDF, camera and light samplers, the variance of all Monte Carlo events is negligible:

$$V\left[\prod_i \frac{\rho_i}{p_i}\right] \approx 0 \qquad E\left[\prod_i \frac{\rho_i}{p_i}\right] \approx \prod_i c_i.$$

An appropriate sampler for microfacet models was introduced by Heitz and d'Eon [Hd14]. Thus, in most cases the variance in the rendered images does not originate in this part of the estimator. If there are weak importance sampling methods for some complex materials it might be useful to clamp the throughput of each event via $\min(c_{\max}, \rho_i/p_i)$. However, this is not the reason for variance in difficult paths. Especially for specular events the sampler is a deterministic reflection with $V = 0$ and clamping it would not reduce the variance of the respective paths.

Hence, the variance of our estimators is dominated by the remaining central terms

$$V[\hat{I}_C] \propto V\left[\frac{\hat{\rho}_k \cos\theta_k \cos\theta_{k+1} \hat{\rho}_{k+1}}{d^2}\right] \gg 0 \tag{4}$$

$$\text{and } V[\hat{I}_M] \propto V\left[\frac{\hat{\rho}_k}{\pi r^2}\right] \gg 0. \tag{5}$$

The expected value can be computed as the double integral of the term over the *footprints* of the sampled paths. Thereby, the footprint is the distribution of possible endpoints of the respective sub-path. Then, the variance depends on the size of the footprints as well as on the variation of the terms over this area. If the two endpoints of the paths vary, there is also a variation in the angular domain which causes a high variance if the BSDF ρ has a sharp peak. Using the smoothed $\hat{\rho}$ reduces the variance in equations (4) and (5) in this case. It is not possible to limit the total variance of the path by only modifying the BSDF. All variance from sampling events, the cosine term $\cos\theta$ and the distance d cannot be controlled this way. However, by keeping those terms we have a contribution-relative limit rather than an absolute one.

3.2. Regularizing Microfacet Models

In this section, practical estimates of the upper bound $\bar{\rho}^{-1}$ (Equation (1)) are introduced for the two most common models. In general the bound must be independent of the incident and exitant direction. Otherwise, the BSDF would be a different one for each evaluated pair of directions, leading to an undesired modification of the shape and energy of the BSDF.

The most used microfacet reflectance model is the Torrance-Sparrow [TS67] model

$$\rho_r(\overleftarrow{\mathbf{w}}, \vec{\mathbf{w}}) = \frac{D(\mathbf{h})G(\overleftarrow{\mathbf{w}}, \vec{\mathbf{w}})F(\langle\overleftarrow{\mathbf{w}}, \mathbf{h}\rangle)}{4|\langle\overleftarrow{\mathbf{w}}, \mathbf{n}\rangle\langle\vec{\mathbf{w}}, \mathbf{n}\rangle|} \tag{6}$$

Here, the function D is the microfacet distribution, G the geometrical shadowing between microfacets, F the Fresnel reflectance, $\mathbf{n}$ the surface normal and $\mathbf{h} = (\overleftarrow{\mathbf{w}} + \vec{\mathbf{w}})/\|\overleftarrow{\mathbf{w}} + \vec{\mathbf{w}}\|$ the half vector between incident $\overleftarrow{\mathbf{w}}$ and exitant direction $\vec{\mathbf{w}}$.

Analogous, a microfacet transmittance model was introduced by Walter et al. [WMLT07]:

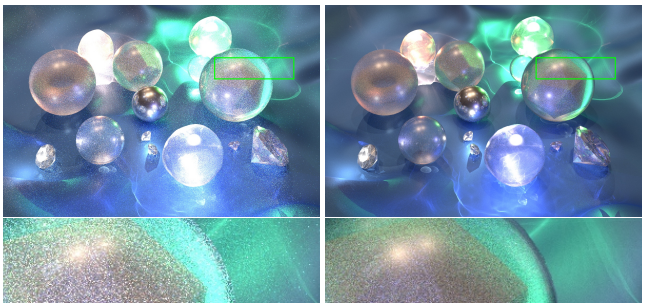
$$\rho_t(\overleftarrow{\mathbf{w}}, \vec{\mathbf{w}}) = \frac{D(\mathbf{h}_t)G(\overleftarrow{\mathbf{w}}, \vec{\mathbf{w}})(1 - F(\langle\overleftarrow{\mathbf{w}}, \mathbf{h}_t\rangle))}{(\eta_i\langle\overleftarrow{\mathbf{w}}, \mathbf{h}_t\rangle + \eta_t\langle\vec{\mathbf{w}}, \mathbf{h}_t\rangle)^2} \frac{\eta_t^2|\langle\overleftarrow{\mathbf{w}}, \mathbf{h}_t\rangle\langle\vec{\mathbf{w}}, \mathbf{h}_t\rangle|}{|\langle\overleftarrow{\mathbf{w}}, \mathbf{n}\rangle\langle\vec{\mathbf{w}}, \mathbf{n}\rangle|} \tag{7}$$

where η_i and η_t are the refraction indices on the incident and the transmitted side of the surface and $\mathbf{h}_t = -(\eta_i\overleftarrow{\mathbf{w}} + \eta_t\vec{\mathbf{w}})/\|\eta_i\overleftarrow{\mathbf{w}} + \eta_t\vec{\mathbf{w}}\|$ is the transmission half vector.

For the bound we can set the Fresnel terms F and $1 - F$ to one, respectively. This maximizes the BSDF in a conservative way.

The microfacet distribution D is usually maximized if $\overleftarrow{\mathbf{w}} = \vec{\mathbf{w}} = \mathbf{n}$. It turns out that the maximum for D is the same for the GGX, the Beckmann and the Cosine distribution, namely $1/\pi\alpha^2$. For details please refer to the supplemental.

Unfortunately, the denominator $|\langle\overleftarrow{\mathbf{w}}, \mathbf{n}\rangle\langle\vec{\mathbf{w}}, \mathbf{n}\rangle|$ in both models causes a problem. It effectively says that any model, independent of the roughness, converges to a specular model at grazing angles. Thus, the strict upper bound is infinity in general, depending on the



(a) Sample-MIS, $\sigma = 17.66$ (b) Reg-MIS, $\sigma = 1.91$

Figure 3: *MIS weight comparison for 1000spp BPT ($\tau = 30$). The inset shows a glass sphere with a diffuse icosahedron inside. Using the sampling-PDFs yields a higher variance σ^2 (a) than the corrected strategy Eq. (13) using modified PDFs (b).*

microfacet shadowing G. We discuss bounds for the two most common shadowing models, V-cavity [TS67] and Smith [Smi67] in the supplemental. Both (and some more) are described well in Heitz's survey on microfacet shadowing [Hei14].

Putting all together we get the bound functions

$$\bar{\rho}_r(\alpha) = \frac{1}{4\pi\alpha^2}\bar{G}(\alpha) \tag{8}$$

$$\text{and } \bar{\rho}_t(\alpha) = \frac{\max(\eta_i,\eta_t)^2}{\pi\alpha^2(\eta_i-\eta_t)^2}\bar{G}(\alpha) \tag{9}$$

$$\text{with } \bar{G}(\alpha) = \begin{cases} 1 & \text{V-cavity} \\ 4/\alpha^2 & \text{Smith, GGX} \\ 4\pi/\alpha^2 & \text{Smith, Beckmann or Cosine} \end{cases}$$

where details on the maximization of the remaining term in the refractive case are shown in the supplemental material.

In general, microfacet models lose energy at high roughness values due to missing multiple scattering between the facets. Our approach is energy-preserving, only if this loss is compensated. Otherwise, the bound is conservative and could be reduced according to the energy loss to minimize the bias.

3.3. MIS Weights under Regularization

Simply replacing the BSDFs in a common method like BPT leads to a high variance which is shown in Figure 3. The reason is that the usual MIS-weight computation is invalid for the new estimators.

Consider the *multi-sample* model

$$F = \sum_{i=1}^{m}\frac{1}{n_i}\sum_{j=1}^{n_i} w_i(x_j)\hat{I}_i(x_j) \tag{10}$$

where $\hat{I}$ are m different sampling strategies as described in Equation (2) and (3), n_i is the number of samples created with these strategies and x_j is the path of a specific sample.

Following the proof in Veach's thesis [Vea97, p 288] the variance

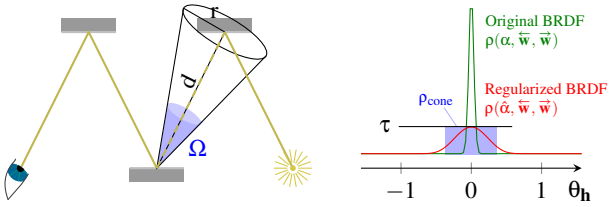


Figure 4: *Left: A cone in which a connection is accepted with a constant BSDF $\rho_{cone} = 1/\Omega = \tau$ (i.e. $\int_\Omega \tau\, d\omega = 1$). Right: Our parametrization limiting the maximum BSDF value by τ.*

of this estimator is

$$V[F] = \left(\int_\Omega \sum_{i=1}^{m}\frac{w_i^2(x)}{n_i}\hat{I}_i(x)^2 p_i^*(x)\,\mathrm{d}\mu(x)\right) - \left(\sum_{i=1}^{m}\frac{1}{n_i}\mu_i^2\right) \tag{11}$$

with $p_i^*(x)$ being the sampling probability of the associated estimator $\hat{I}_i$ which is a product of sampling probabilities $p(\mathbf{w})$ with respect to the area measure: $p_i^*(x) = \prod_j p_j(\vec{\mathbf{w}}_j)\,\mathrm{d}A_j$. On the right, μ_i is the true expected value of the estimator including the weight, i.e. $\mu_i = E[w_i\hat{I}_i]$.

At this point Veach assumed that $\hat{I}_i(x)$ has the form $f_i(x)/p_i^*(x)$ and that f_i[†] is constant per path: $f_i(x) = f_j(x)$, which allowed him to remove f_i from the weight optimization. This is not true for our regularization strategy. Due to regularizing a different ρ in each of the strategies, f_i is not constant anymore. It differs for each of the sampling strategies.

Applying the same approach of optimizing the first term in Equation (11) using Lagrange multipliers, we get a different form of the *balance heuristic*

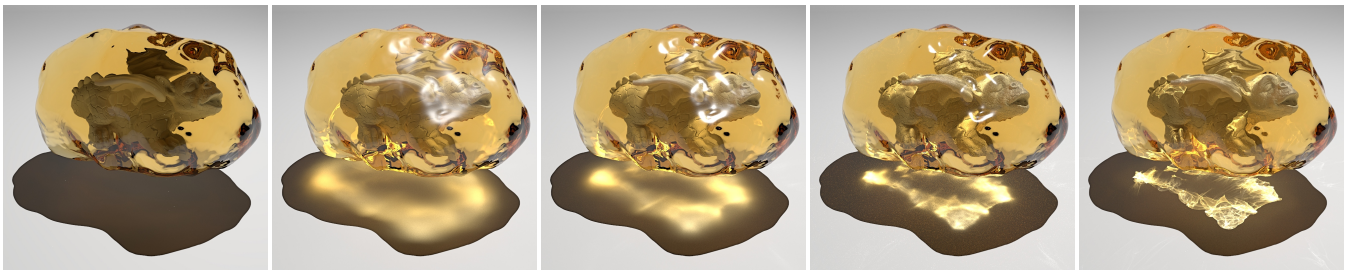
$$w_i = \frac{n_i(1/\hat{I}_i)}{\sum_k n_k(1/\hat{I}_k)} = \frac{n_i}{\sum_k n_k(\hat{I}_i/\hat{I}_k)}. \tag{12}$$

In the case that no regularization is used, Equation (12) falls back to the known form $n_i p_i^*/\sum_k n_k p_k^*$, as was shown by Jendersie [Jen18]. Similar, we can cancel out terms to express Equation (12) with respect to regularized PDFs $\hat{p}^* = \prod_j \hat{p}_j\,\mathrm{d}A$:

$$w_i = \frac{n_i\hat{p}_i^*}{\sum_k n_k\hat{p}_k^*}. \tag{13}$$

We set $\hat{p}_j = (\hat{\rho}_j/\rho_j)\cdot p_j$ in each event j on the path, such that the difference between regularization and no regularization becomes part of the PDF. Note that we still sample the non-regularized PDFs p_j. This is only an artificial change to simplify the implementation of the MIS, if integrated into a common renderer.

We show in the supplemental that $\hat{p}_j$ can be obtained directly by evaluating the PDF with the parameter $\hat{\alpha}$, instead of computing the ratio of BSDFs $(\hat{\rho}_j/\rho_j)$. For the Smith model, an additional quotient of geometry terms $G_{\hat{\alpha}}(\vec{\mathbf{w}})/G_\alpha(\vec{\mathbf{w}})$ must be added, which is not necessary for the V-cavity model.



(a) PT, 16k spp (b) PT, $\tau = 10$, 16k spp (c) PT, $\tau = 100$, 24k spp (d) PT, $\tau_0 = 100$, 32k spp (e) VCM, 32k spp

Figure 5: *Bias comparison of a non consistent, regularized PT (b, c) with a consistent one (d) and a reference rendering (e). The scene shows the PBRT-dragon (courtesy of Christian Schüller) inside an amber, illuminated by a point light source.*

3.4. A Consistent Parametrization

Commonly, a start radius r_0 is used in photon mapping to define the amount of allowed spatial bias. All photons within a disc of the specified radius around the search location are merged with the current sub-path. To produce *consistent* results, the radius r is often reduced progressively over iterations [HOJ08, HJ09, KZ11, KD13].

As shown in Kaplanyan and Dachsbacher's work [KD13], the optimal radius in iteration N for a regularization with two degrees of freedom, like photon mapping, is

$$r_N = r_0 \cdot N^{-1/6}. \tag{14}$$

They applied this parametrization to mollify specular vertices by computing a cone opening angle from r_N and the distance d (Figure 4 left). Then, if a connection is found within this cone, it is accepted and the uniform BSDF of a cone-shaped distribution

$$\rho_{\text{cone}} = \frac{1}{2\pi\left(1 - \cos\left(\arctan\left(r_N/d\right)\right)\right)} \tag{15}$$

is evaluated. We define that the threshold τ should equal ρ_{cone} (Figure 4 right), because the radius cannot be applied directly to arbitrary BSDFs. A normalized BSDF with the same maximum value will have a similar width and variance like ρ_{cone}.

However, we found that the segment length d causes problems for several reasons: First, the BSDF loses its reciprocity and second, while evaluating different directions, the smoothness of the BSDF would change, leading to arbitrary shapes. To become independent of d we inverted Equation (15) for r_N and reinserted it (for details please refer to the supplemental) which leads to

$$\tau_N = \frac{1}{2\pi\left[1 - \cos\left(\arctan\left(\frac{\sqrt{4\pi\tau_0 - 1}}{2\pi\tau_0 - 1} \cdot N^{-1/6}\right)\right)\right]} \tag{16}$$

as a *consistent* threshold for iteration N. Since we still apply the decreasing series to a radius we can rely on the proofs made previously [KD13].

† f is the path measurement contribution function

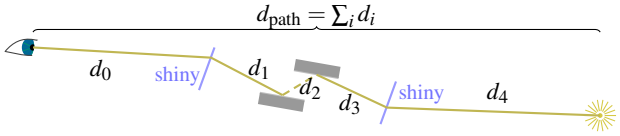


Figure 6: *Example of a sampleable, high variance path which can be avoided by regularizing other connections.*

3.5. Adaptive Regularization

In Figure 5 we can see examples of a regularized path tracer with and without consistent parametrization. While image (d) converges to the reference (e) in theory, it will not do so for any practical iteration count. However, we can restrict the bias to difficult paths only, instead of using an unconditional regularization. For example, the highlights could be rendered with a much larger τ.

We introduce two heuristics to reduce the bias: the sampler *quality* and path *diffusion*. If there is a sampler which can sample the current path well, we want to disable regularization for this path. Otherwise, we would like to hide the bias by increasing the threshold dependent on the path diffusion. If the path was scattered on a rough surface, it is likely that we are not able to see the bias introduced by the blurring of highlights. On the other hand, bias is well visible if the path is almost deterministic – in which case the variance will be small and we can reduce the amount of regularization.

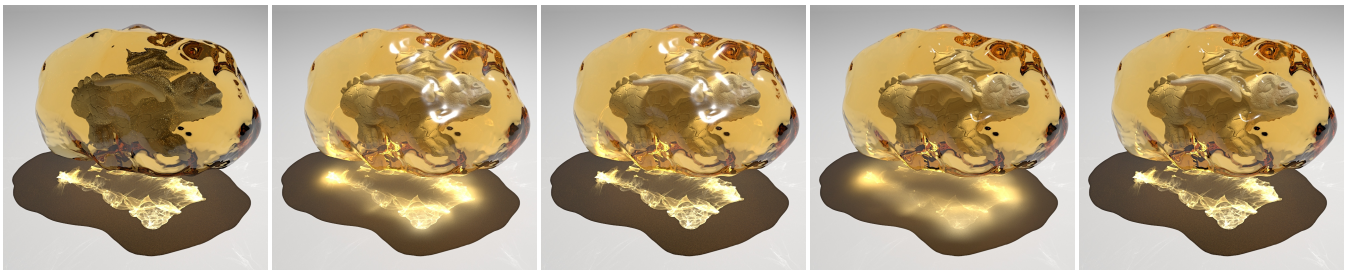
3.5.1. Sampler Quality

Our first heuristic estimates an upper bound q of the smallest possible connection term on the path:

$$q = \min_k \left(\bar{\rho}_k \cdot \bar{\rho}_{k+1} \cdot \left(d_k / d_{\text{path}}\right)^{-2}\right). \tag{17}$$

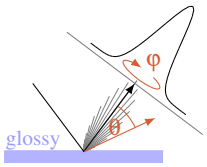
According to the assumption of low variance samplers ρ/p in Section 3.1, the connection term dominates the path's variance. Therefore, $q < \tau^2$ implies a sufficient sampler but does not guarantee one.

We added the relative connection length d_k/d_{path}, with d_{path} being the sum of all segment lengths, to judge the influence of d in Equation (4). An absolute small distance on an equally short path means a high contribution without much variance. Contrary, a



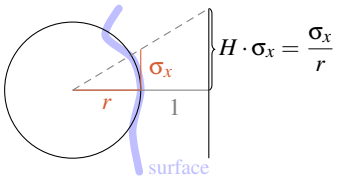
(a) No regularization (b) Non-adaptive (c) Sampler Quality (Eq. (18)) (d) Diffusion (Eq. (19)) (e) Both

Figure 7: *BPT with different strategies of adaptive regularization ($\tau_0 = 30$, 16k spp). The sampler quality heuristic (c) preserves the caustic, but keeps blurred specular paths. The diffusion heuristic (d) results in sharper highlights and sharper directly visible refractions, but blurs the caustic even more. In (e) both are combined and the overall bias is reduced. Also compare with the ground truth VCM from Figure 5.*



$$g(\theta,\varphi) = \frac{1}{2\pi\sigma_{\sphericalangle}^2}\exp\left(-\frac{\tan^2\theta}{2\sigma_{\sphericalangle}^2}\right)$$
$$g(0,0) = \frac{1}{2\pi\sigma_{\sphericalangle}^2} = \bar{\rho}$$

Figure 8: *Tangential standard deviation of the reflection direction through the BSDF.*



$$H \cdot \sigma_x = \frac{\sigma_x}{r}$$

Figure 9: *Curvature equals the inverse radius of a sphere. Multiplying with the spatial footprint yields a tangens-deviation again.*

small distance on a long path may lead to a high variance sampler. An example is shown in Figure 6 and can be observed in practice in Figure 7 (a) on the dragon's wings.

If $q < \tau^2$, at least one unbiased sampler is expected to be as good as any regularized sampler we could get, so we set

$$\tau' = \begin{cases} \tau & \text{if } q \geq \tau^2 \\ \infty & \text{otherwise} \end{cases} \tag{18}$$

to disable regularization in the presence of a good sampler. Comparing Figure 7 (b) with (c), we can see that the caustic is successfully preserved without introducing noise in other regions.

The heuristic has an inevitable discontinuity, if the best sampler's q is approximating τ^2. In the moment of crossing the threshold, regularization will be switched off which can happen due to a change of τ or a change in path length (e.g. animation). A smoother transition from ∞ to τ is possible by exchanging Equation (18).

3.5.2. Path Diffusion

The motivation for the second heuristic is that we would like to hide the bias in regions of the paths where we cannot see it directly and to reduce the blurring of shiny paths. For example, in Figure 3 (b, closeup) the shiny-diffuse-shiny paths look very cloudy. By regularizing only the second shiny surface (towards the light) we can avoid the visible blur on the first surface. To that end, we modify τ by the tangential standard deviation $\sigma_{\sphericalangle}$ on the view sub-path up to the current vertex k:

$$\tau' = \frac{\tau}{\sigma_{\sphericalangle,k}} \tag{19}$$

We determine $\sigma_{\sphericalangle,k}$ similar to the 5D covariance tracing from Belcour et al. [BSS*13], but reduce the estimate to 2D. We are neither interested in anisotropic shapes nor in a temporal domain. The computation for our simplified case can be summarized by:

$$\sigma_{\mathbf{x},k} = \sigma_{\mathbf{x},k-1} + \sigma_{\sphericalangle,k-1} \cdot d \tag{20}$$
$$\sigma_{\sphericalangle,k} = \sigma_{\sphericalangle,k-1} + (2\pi\bar{\rho}_{k-1})^{-1/2} + \sigma_{\mathbf{x},k-1} \cdot H_{k-1} \tag{21}$$

where $\sigma_{\mathbf{x}}$ describes the positional standard deviation (footprint) and H is the mean curvature of the local surface. The term $(2\pi\bar{\rho})^{-1/2}$ accounts for the BSDF-based diffusion as follows: If we assume a bivariate Gaussian distribution in the tangential plane to the reflection direction (Figure 8), we can set its maximum $g(0,0)$ to $\bar{\rho}$. This equality can be solved for the standard deviation parameter which gives us the above term.

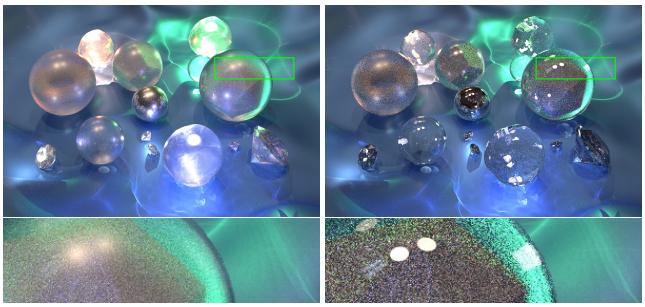
The curvature H accounts for the macroscopic scattering which needs to be converted into an absolute diffusion by using the expected footprint size $\sigma_{\mathbf{x}}$. The geometrical interpretation for the curvature diffusion is shown in Figure 9.

3.6. Alternative Regularization Strategies

For the regularization of pure specular events, the cone-BSDF can be used with either the radius or our parametrization to match the regularization of the microfacet BSDFs [KD13].

Kaplanyan and Dachsbacher also proposed to simulate a merge by accepting a connection if a random sample is within a cone. The virtual merge algorithm is as follows:

1. Sample the current BSDF $\rightarrow$ sample with $\rho_s, p_s, \mathbf{w}_s$
2. Compare angle between $\mathbf{w}_s$ and the connection direction
   a. If within allowed cone return $\tau \cdot f_s / p_s$
   b. Otherwise discard connection



(a) Non-adaptive, $\tau_N = 30$ $\sigma = 1.91$ (b) Virtual merge, $\tau_N = 30$ $\sigma = 2.66$

Figure 10: *Virtual merge strategy compared to the roughness-based approach. Both images are rendered with BPT (1000 spp).*

This approach has two drawbacks. First, the additional random sampling introduces variance on its own and also consumes precious random numbers, and second, it is difficult to adopt the MIS weight computation.

One possible MIS strategy would be to replace the PDF of the regularized vertex with τ (see supplemental). The reasoning behind this simple solution is that the difference between a random walk and a random connection is the chance $1/\tau$ to accept the connection. Both include the sampling probability p_s, such that it cancels out. Note that this is not an optimal choice, because the true chance $1/\int p$ over the cone would be required instead.

We tested this regularization strategy, as shown in Figure 10. While providing sharper versions of the highlights, it also shows a higher noise level than the mollification approach. As expected, the additional random sampling produces a high variance in the tail regions of the regularized vertices.

4. Results

Our goal was to reduce the variance of arbitrary Monte Carlo light transport algorithms by blurring paths which are hard to sample. In Figure 11 we show an equal time comparison with and without regularization, starting at a very small threshold. In all cases, both heuristics are enabled by applying Equations (18) and (19) successively. To measure the error we used the sample standard deviation σ and the root mean square error of *relative* residuals $\text{RMS}RE = \sqrt{\frac{1}{n}\sum_i (a_i - b_i)^2/(\frac{a_i+b_i}{2})^2}$. Using absolute RMSE penalizes bright noise much more than missing features and turned out to be unusable. For example in the TOY DRAGONS scene, the PT and the VCM rendering have almost the same RMSE.

In many cases regularization is able to produce more difficult light effects in the simpler rendering methods PT and BPT. There, caustics or SDS paths are missing if rendered without regularization. In some cases (e.g. TOY DRAGONS: reflected caustics) our PT does a better job than BPT, because it connects to all light sources instead of a single light path like BPT, thus lowering the overall variance. If regularization is applied to VCM, the *shiny* paths are added and additional samples for SDS paths are provided. This improves the fidelity of *shiny* objects like jewels or the glass dragon.

The noise level σ is often higher in the regularized variants, because the otherwise missing light effects add a considerable amount of new noise sources. Visually, the overall error is still smaller due to the reduced bias (also see SSIM convergence experiment in the supplemental). Comparing the three regularized variants to each other we can see that using the more expensive method produces less variance in general. The reason is that regularization only turns an impossible path into an unlikely one, whereas adding more samplers can produce the same path with a higher probability.

4.1. Parameter Choice

In general τ can be chosen independent of the scene, but fine tuning can improve the balance between noise and variance. For PT and BPT the threshold τ_0 should be in $[10, 100]$ for reasonable noise levels of caustics/SDS paths. Also see Figure 5 (b, c) for a comparison of different parameters. Higher thresholds are possible for pure *shiny* paths without problems. Therefore, setting $\tau_0 \in [1000, 5000]$ in VCM is reasonable as well. However, using smaller thresholds in VCM also helps to decrease variance (e.g. Figure 11, MARBLES).

5. Conclusions and Future Work

We have shown how the regularization framework from Kaplanyan and Dachsbacher [KD13] can be applied to more general microfacet BSDF models. It turns out that the *mollification*, i.e. the blurring of the BSDFs, yields better results with respect to variance and visual appearance than using virtual merges. In both cases it is necessary to change the MIS weight computation for which we have found an appropriate strategy.

With general regularization it is possible to render complex scenes (with respect to paths) with simpler methods. Additionally, variance can be reduced in methods like VCM. Regularization will not produce the unbiased ground truth in a practical amount of time, if the underlying algorithm cannot handle the light effects. Its main advantage is to decrease the variance at the cost of bias, while producing results which are more similar to the ground truth than without regularization.

We would prefer VCM over regularized BPT in this raw configuration, because its path reuse decreases the variance considerably. However, this might be different for renderers employing caching and guidance strategies which are likely to reduce variance in regularized SDS configurations. For example using a sampling guidance method [HEV*16, MGN17] would help in SDS paths, to proceed into the direction of the next highlight.

A different application could be the use in improved BPT methods. Popov et al. proposed to resample a connection attempt from a set of possible connections with respect to the expected contribution [PRDD15]. Similarly, Chaitanya et al.'s Matrix Bidirectional Path Tracing [CBH*18] reshuffles a set of connections to minimize the connection distances of the set, which again increases the expected contribution. Both strategies are capable of reducing the variance in BPT and could do well in cases of regularized paths. Especially, Popov's method [PRDD15], which includes the BSDF values, could select appropriate light path vertices on a smooth surface for regularized connections.

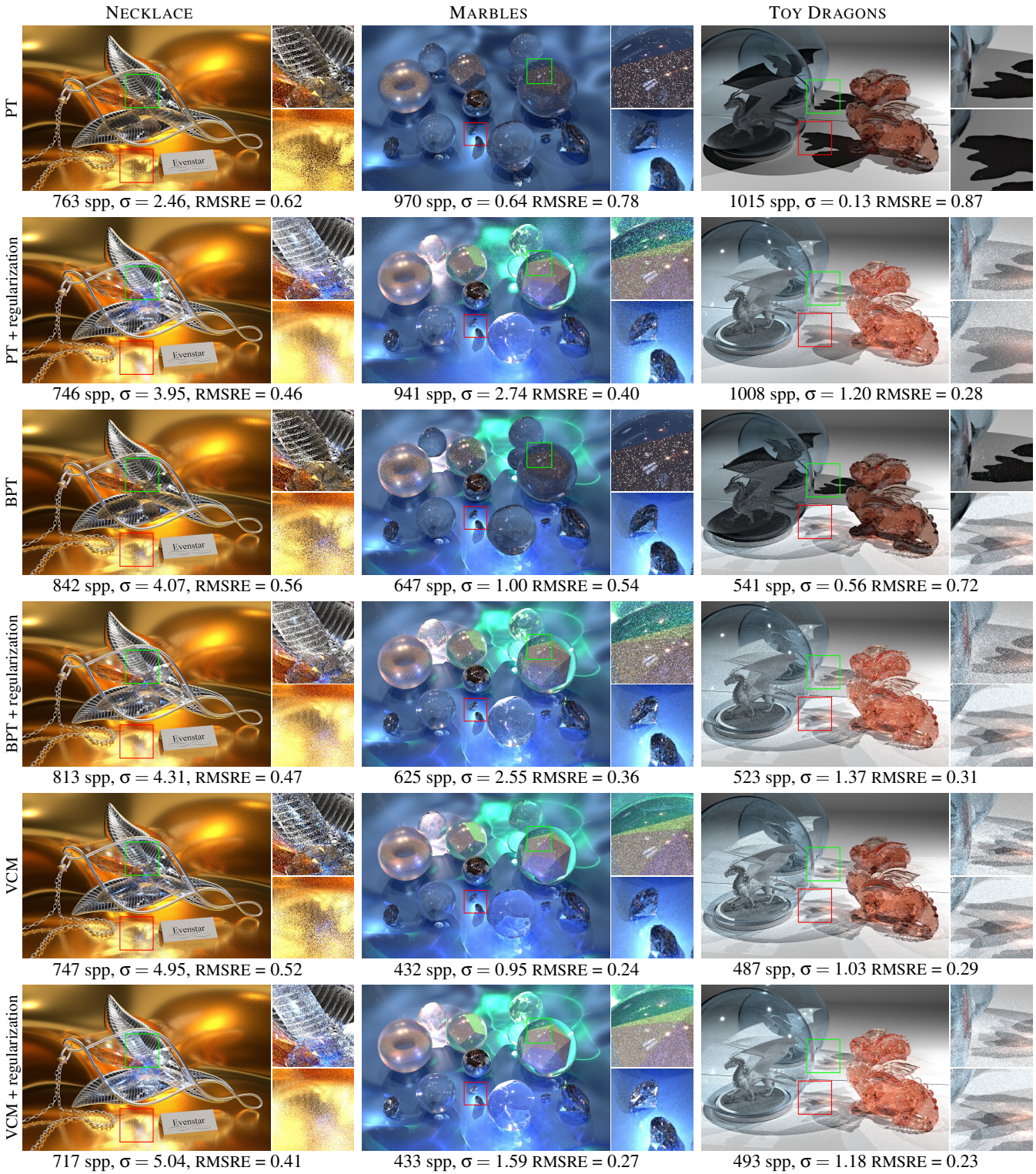


Figure 11: *Equal-time comparison over 1h process time. All regularized images are rendered with the roughness-based, adaptive, consistent approach ($\tau_0 = 10$).*
NECKLACE*: A metal necklace with several crystals on a slightly rough golden floor; 10 point lights (7 inside the crystals) and 1 area light.*
MARBLES*: Glass, metal spheres and gems on a rough blue folded cloth; 6 point lights (3 colored inside the wobbly spheres), env.-map.*
TOY DRAGONS*: Two dragons in front of a mirror, one inside a thin glass hull; 2 point lights.*

To decrease the regularization bias, it would be nice to have closed form convolutions for the analytic BSDF models. The current approach redistributes the energy of a glossy peak along all angles – including the tails. An analytic convolution would have similar results as the virtual merge approach, but without the additional random decision. An approximate solution to this problem was given by Heitz et al. [HDHN16] using cosine lobes.

Acknowledgments

Many of the scenes, including the teaser, were taken from www.blendswap.com and are protected through CC-BY-3.0. The NECKLACE scene was made by the user *pizzahouse6*. The left dragon in TOY DRAGONS was made by Kevin Ramos (*Delatronic*) while the other dragon (AMBER, TOY DRAGONS right) is a courtesy of Christian Schüller taken from the PBRT repository.

References

[BIOP13] BOUCHARD G., IEHL J.-C., OSTROMOUKHOV V., POULIN P.: Improving Robustness of Monte-Carlo Global Illumination with Directional Regularization. In *SIGGRAPH Asia 2013 Technical Briefs* (2013), SA '13, ACM, pp. 22:1–22:4. doi:10.1145/2542355.2542383. 2

[BSS*13] BELCOUR L., SOLER C., SUBR K., HOLZSCHUCH N., DURAND F.: 5d Covariance Tracing for Efficient Defocus and Motion Blur. *ACM Transaction on Graphics (TOG) 32*, 3 (July 2013), 31:1–31:18. doi:10.1145/2487228.2487239. 6

[CBH*18] CHAITANYA C. R. A., BELCOUR L., HACHISUKA T., PREMOZE S., PANTALEONI J., NOWROUZEZAHRAI D.: Matrix Bidirectional Path Tracing. In *Proc. of Eurographics Symposium on Rendering (EGSR)* (2018), pp. 023–032. doi:10.2312/sre.20181169. 7

[FHF*17] FASCIONE L., HANIKA J., FAJARDO M., CHRISTENSEN P., BURLEY B., GREEN B.: Path Tracing in Production - Part 1: Production Renderers. In *ACM SIGGRAPH Courses* (2017), pp. 13:1–13:39. doi:10.1145/3084873.3084904. 2

[GKDS12] GEORGIEV I., KŘIVÁNEK J., DAVIDOVIČ T., SLUSALLEK P.: Light Transport Simulation with Vertex Connection and Merging. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia) 31*, 6 (2012), 192:1–192:10. doi:10.1145/2366145.2366211. 1, 2

[Hd14] HEITZ E., D'EON E.: Importance Sampling Microfacet-Based BSDFs Using the Distribution of Visible Normals. *Computer Graphics Forum (CGF) 33*, 4 (2014), 103–112. doi:10.1111/cgf.12417. 3

[HDF15] HANIKA J., DROSKE M., FASCIONE L.: Manifold Next Event Estimation. *Computer Graphics Forum (Proc. EGSR) 34*, 4 (July 2015), 87–97. doi:10.1111/cgf.12681. 2

[HDHN16] HEITZ E., DUPUY J., HILL S., NEUBELT D.: Real-time Polygonal-light Shading with Linearly Transformed Cosines. *ACM Transaction on Graphics (TOG) 35*, 4 (July 2016), 41:1–41:8. doi:10.1145/2897824.2925895. 9

[Hei14] HEITZ E.: Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. *Journal of Computer Graphics Techniques (JCGT) 3*, 2 (June 2014), 48–107. URL: http://jcgt.org/published/0003/02/03/. 4

[HEV*16] HERHOLZ S., ELEK O., VORBA J., LENSCH H., KŘIVÁNEK J.: Product Importance Sampling for Light Transport Path Guiding. *Computer Graphics Forum (CGF) 35*, 4 (2016), 67–77. doi:10.1111/cgf.12950. 2, 7

[HJ09] HACHISUKA T., JENSEN H. W.: Stochastic Progressive Photon Mapping. *ACM Transactions on Graphics (TOG) 28*, 5 (Dec. 2009), 141:1–141:8. doi:10.1145/1618452.1618487. 2, 5

[HOJ08] HACHISUKA T., OGAKI S., JENSEN H. W.: Progressive Photon Mapping. *ACM Transactions on Graphics (TOG) 27 number 5* (2008), 130. doi:10.1145/1409060.1409083. 2, 5

[HPJ12] HACHISUKA T., PANTALEONI J., JENSEN H. W.: A Path Space Extension for Robust Light Transport Simulation. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia) 31*, 6 (Nov. 2012), 191:1–191:10. doi:10.1145/2366145.2366210. 1, 2

[Jen96] JENSEN H. W.: Global Illumination using Photon Maps. In *Proc. of Eurographics Workshop on Rendering (EGWR)* (1996), Springer, pp. 21–30. URL: http://dl.acm.org/citation.cfm?id=275458.275461. 2

[Jen18] JENDERSIE J.: *Path Throughput Importance Weights*. Tech. Rep. arXiv:1806.01005, June 2018. URL: https://arxiv.org/abs/1806.01005. 4

[Kaj86] KAJIYA J. T.: The Rendering Equation. *Computer Graphics (Proc. ACM SIGGRAPH) 20*, 4 (Aug. 1986), 143–150. doi:10.1145/15886.15902. 2

[KD13] KAPLANYAN A., DACHSBACHER C.: Path Space Regularization for Holistic and Robust Light Transport. *Computer Graphics Forum (CGF) 32* (2013), 63–72. doi:10.1111/cgf.12026. 1, 2, 5, 6, 7

[KZ11] KNAUS C., ZWICKER M.: Progressive Photon Mapping: A Probabilistic Approach. *ACM Transactions on Graphics (TOG) 30*, 3 (2011), 25. doi:10.1145/1966394.1966404. 2, 5

[MGN17] MÜLLER T., GROSS M., NOVÁK J.: Practical Path Guiding for Efficient Light-Transport Simulation. In *Proc. of Eurographics Symposium on Rendering (EGSR)* (June 2017). doi:10.1111/cgf.13227. 2, 7

[PRDD15] POPOV S., RAMAMOORTHI R., DURAND F., DRETTAKIS G.: Probabilistic Connections for Bidirectional Path Tracing. *Computer Graphics Forum (Proc. of EGSR) 34*, 4 (2015), 75–86. doi:10.1111/cgf.12680. 7

[ŠK18] ŠIK M., KŘIVÁNEK J.: Survey of Markov Chain Monte Carlo Methods in Light Transport Simulation. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* (Nov. 2018). doi:10.1109/TVCG.2018.2880455. 2

[Smi67] SMITH B. G.: Geometrical Shadowing of a Random Rough Surface. *IEEE Transactions on Antennas and Propagation 15*, 5 (Sept. 1967), 668–671. doi:10.1109/TAP.1967.1138991. 4

[ŠOHK16] ŠIK M., OTSU H., HACHISUKA T., KŘIVÁNEK J.: Robust Light Transport Simulation via Metropolised Bidirectional Estimators. *ACM Transaction on Graphics (TOG) 35*, 6 (2016), 245:1–245:12. doi:10.1145/2980179.2982411. 2

[TS67] TORRANCE K. E., SPARROW E. M.: Theory for Off-Specular Reflection From Roughened Surfaces. *Journal of the Optical Society of America 57*, 9 (Sept. 1967), 1105–1114. doi:10.1364/JOSA.57.001105. 3, 4

[Vea97] VEACH E.: *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, 1997. URL: http://graphics.stanford.edu/papers/veach_thesis/. 4

[VG95] VEACH E., GUIBAS L. J.: Bidirectional Estimators for Light Transport. In *Photorealistic Rendering Techniques*. Springer Berlin Heidelberg, 1995, pp. 145–167. URL: http://dx.doi.org/10.1007/978-3-642-87825-1_11. 2

[VG97] VEACH E., GUIBAS L. J.: Metropolis Light Transport. In *Proceedings of SIGGRAPH '97* (1997), pp. 65–76. doi:10.1145/258734.258775. 2

[WMLT07] WALTER B., MARSCHNER S. R., LI H., TORRANCE K. E.: Microfacet Models for Refraction Through Rough Surfaces. In *Proc. of Eurographics Symposium on Rendering (EGSR)* (2007), Eurographics Association, pp. 195–206. doi:10.2312/EGWR/EGSR07/195-206. 3