

Manifold Path Guiding for Importance Sampling Specular Chains

ZHIMIN FAN*, State Key Lab for Novel Software Technology, Nanjing University, China

PENGPEI HONG*[†], University of Utah, United States of America

JIE GUO[‡], State Key Lab for Novel Software Technology, Nanjing University, China

CHANGQING ZOU, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University, China

YANWEN GUO[‡], State Key Lab for Novel Software Technology, Nanjing University, China

LING-QI YAN, University of California, Santa Barbara, United States of America

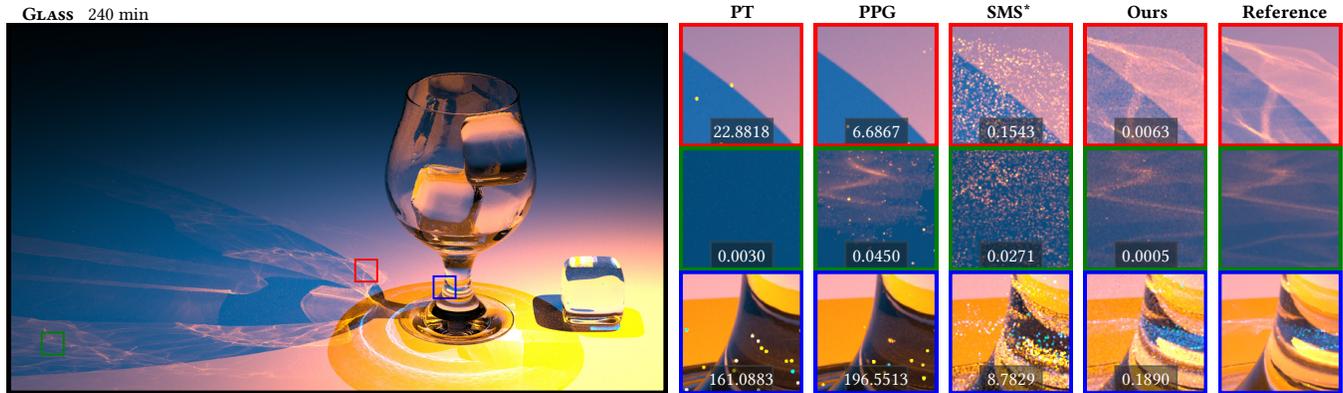


Fig. 1. Rendering a complex scene involving challenging light paths with multiple consecutive specular bounces. The long specular chains (> 5 in this scene) create great obstacles to existing path sampling algorithms, while the proposed *manifold path guiding* method addresses this issue, faithfully reproducing high-frequency caustics and noticeably reducing the variance. Here, we compare our method with Path Tracing (PT), Practical Path Guiding (PPG) [Müller 2019; Müller et al. 2017], and an extension (supporting various chain types) of Specular Manifold Sampling (SMS*) [Zeltner et al. 2020] at the same rendering time. Quantitative error in terms of MSE is reported for each closeup.

Complex visual effects such as caustics are often produced by light paths containing multiple consecutive specular vertices (dubbed *specular chains*), which pose a challenge to unbiased estimation in Monte Carlo rendering. In this work, we study the light transport behavior within a sub-path that is comprised of a specular chain and two non-specular separators. We show that the specular manifolds formed by all the sub-paths could be exploited to provide coherence among sub-paths. By reconstructing continuous energy distributions from historical and coherent sub-paths, seed chains can be

*Joint first authors.

[†]The work was done at Nanjing University.

[‡]Corresponding authors.

Authors' addresses: Zhimin Fan, zhiminfan2002@gmail.com, State Key Lab for Novel Software Technology, Nanjing University, Nanjing, China; Pengpei Hong, hpommpy@gmail.com, University of Utah, Salt Lake City, United States of America; Jie Guo, guojie@nju.edu.cn, State Key Lab for Novel Software Technology, Nanjing University, Nanjing, China; Changqing Zou, changqing.zou@zju.edu.cn, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China; Yanwen Guo, ywguo@nju.edu.cn, State Key Lab for Novel Software Technology, Nanjing University, Nanjing, China; Ling-Qi Yan, lingqi@cs.ucsb.edu, University of California, Santa Barbara, Santa Barbara, United States of America.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2023/12-ART257 \$15.00

<https://doi.org/10.1145/3618360>

generated in the context of importance sampling and converge to admissible chains through manifold walks. We verify that importance sampling the seed chain in the continuous space reaches the goal of importance sampling the discrete admissible specular chain. Based on these observations and theoretical analyses, a progressive pipeline, *manifold path guiding*, is designed and implemented to importance sample challenging paths featuring long specular chains. To our best knowledge, this is the first general framework for importance sampling discrete specular chains in regular Monte Carlo rendering. Extensive experiments demonstrate that our method outperforms state-of-the-art unbiased solutions with up to 40× variance reduction, especially in typical scenes containing long specular chains and complex visibility.

CCS Concepts: • **Computing methodologies** → **Ray tracing**.

Additional Key Words and Phrases: Specular chain, Importance sampling, Caustics

ACM Reference Format:

Zhimin Fan, Pengpei Hong, Jie Guo, Changqing Zou, Yanwen Guo, and Ling-Qi Yan. 2023. Manifold Path Guiding for Importance Sampling Specular Chains. *ACM Trans. Graph.* 42, 6, Article 257 (December 2023), 14 pages. <https://doi.org/10.1145/3618360>

1 INTRODUCTION

Monte Carlo (MC) integration using stochastic samples has long been the de facto solution to the problem of physically-based light transport simulation [Christensen and Jarosz 2016; Fascione et al. 2018; Keller et al. 2015]. Over the past decades, great efforts have

been devoted to significantly improving the convergence rate and reducing the noise for MC-based rendering algorithms.

However, several common scenes still lack effective and robust sampling strategies to handle certain classes of light paths well. In particular, long paths with multiple consecutive specular or near-specular scattering events (widely appearing in scenes with curved metals, water, and glasses as shown in Fig. 1) have always been a nightmare for existing physically-based rendering engines, causing significantly slow convergence. The key difficulty lies in finding valid path samples that are not only “important” but also satisfy all physical constraints at specular reflective/refractive vertices [Jakob and Marschner 2012].

To reduce the variance for scenes with multiple specular scattering events, many approaches have been tried. Early attempts leverage Metropolis Light Transport (MLT) [Veach and Guibas 1997] to explore high-energy regions in the path space. However, even if some specific mutation strategies have been designed to handle paths containing specular chains [Jakob and Marschner 2012; Kaplanyan et al. 2014], it is still difficult to find the so-called Specular-Diffuse-Specular (SDS) paths. Another line of work leverage fitted energy distributions to directly sample the difficult paths [Müller et al. 2017; Vorba et al. 2014]. Despite their success in production, they fail to handle pure specular cases (i.e., specular paths adjacent to point light sources).

Recently, Zeltner et al. [2020] proposed Specular Manifold Sampling (SMS), a general sampling strategy tailored for specular light paths. Unfortunately, this method favors short specular paths with one or two vertices, and its performance reduces dramatically when the paths become longer. Moreover, SMS samples specular paths according to the size of convergence basins and ignores their energy distributions which are critical for variance reduction. Until now, importance sampling arbitrarily long *specular chains* (i.e., light paths containing multiple consecutive specular vertices) is still a far less explored field in computer graphics.

Our goal in this paper is to analyze the best possible importance sampling strategy for specular chains and find a practical solution to approach it without incurring significant overhead to the existing rendering pipelines. Due to the discrete nature of the *admissible specular chain* space and its high dimensionality, conducting importance sampling directly in that space may be difficult. To address this key issue, we exploit a continuous space in which seed chains can be importance sampled and converge to corresponding admissible chains via manifold walk [Jakob and Marschner 2012]. We further reduce the dimensionality of the space by in-depth analyzing the light transport behavior of admissible specular chains, thus making the task tractable and easily supporting glossy cases.

We design and implement a practical rendering pipeline, named *manifold path guiding*, to construct continuous distributions for importance sampling seed chains. After initialization, the pipeline progressively learns specular manifolds via reconstructing continuous spatial-directional distributions from historical sub-paths, generating new seed chains following the distributions, and performing manifold walks to obtain new admissible sub-paths. The learned distributions converge after several iteration steps and are then used in the final rendering stage. We show through extensive experiments that the proposed method allows us to faithfully produce

complex caustics stemming from arbitrarily long specular chains, and it outperforms previous unbiased MC solutions (including those tailored for specular paths) with more than one order of magnitude variance reduction in equal rendering time.

In summary, our contributions are:

- (1) A formulation of the task of specular chains importance sampling that covers any number of consecutive specular vertices,
- (2) A general solution that obtains and utilizes a continuous distribution exploiting the continuity of specular manifold,
- (3) A practical pipeline that progressively gathers sub-path samples, achieving fast convergence even for paths with long specular chains and complex visibility.

2 RELATED WORK

Bidirectional Monte Carlo and MCMC methods. Importance sampling difficult paths has been a long-standing challenge in rendering. Bidirectional approaches [Lafortune and Willems 1993; Veach and Guibas 1995a] try to solve this problem by sampling from the view-point and the emitters independently and then properly connecting sub-paths. However, they fail to handle the SDS paths.

Biased methods, such as photon mapping [Hachisuka and Jensen 2009; Jensen and Christensen 1995] and regularization [Jendersie and Grosch 2019; Kaplanyan and Dachsbacher 2013; Weier et al. 2021], are also designed to find difficult paths for caustics. Due to their biased nature, these methods tend to lose details from high-frequency caustics.

Another family of Monte Carlo methods samples difficult paths by running a Markov chain of paths, known as MCMC approaches [Sik and Krivánek 2020]. Pioneered by Metropolis Light Transport (MLT) [Veach and Guibas 1997], many efforts have been devoted to significantly enhancing the performance of path mutations [Jakob and Marschner 2012]. Although MCMC-based methods reduce variance significantly, they are plagued with splotchy non-uniform noise and temporal flickering artifacts in animations.

Unlike these methods, our approach works in a regular and unbiased Monte Carlo manner. We stride over the limitation by connecting chains between separators. This belongs to the non-local sampling strategies [Veach 1998].

Path guiding. Monte Carlo rendering techniques rely heavily on importance sampling when constructing light transport paths. So far, the most promising sampling distributions are obtained based on learned scene priors [Vorba et al. 2019].

Existing works have studied various distribution representations. For instance, Lafortune and Willems [1995] stored incident radiance of previously traced rays in a 5D tree, and Jensen [1995] estimated histograms of incoming radiance from photons. Gaussian Mixture Models (GMM) [Vorba et al. 2014] are found to be flexible, while quad-trees [Müller et al. 2017] succeed in practice due to their adaptability. Some approaches further take consideration of the correlation between consecutive vertices in the full path, e.g., using product importance sampling [Herholz et al. 2016], parallax-aware warping [Ruppert et al. 2020] and spatial correlation [Dodik et al. 2022; Schüßler et al. 2022]. Recently, Li et al. [2022] utilized representative specular paths to enable effectively guided rendering of caustics,

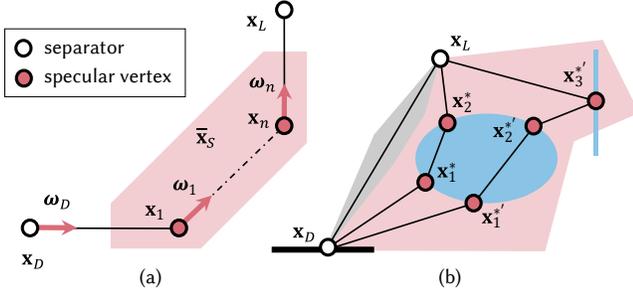


Fig. 2. **(a) Notations for a specular chain** \bar{x}_S (in the red-shaded area) which is comprised of n specular vertices x_1, x_2, \dots, x_n . By connecting two separators x_D and x_L with a specular chain \bar{x}_S , we produce a sub-path (x_D, \bar{x}_S, x_L) . **(b) Throughput between two separators**, which comes from both direct and indirect connections. In this example, the indirect part is comprised of two admissible chains $\bar{x}_S^* = x_1^*, x_2^*$ and $\bar{x}_S'^* = x_1'^*, x_2'^*, x_3'^*$. Thus, the admissible chain space $\mathcal{P}_S^* = \{\bar{x}_S^*, \bar{x}_S'^*\}$.

but their methods struggled to support specular or low-roughness materials.

Guided path sampling achieves great success when high-frequency details are absent in the radiance distribution. However, when faced with paths containing multiple specular interactions, the radiance distribution usually involves high-frequency variations that bottleneck existing guiding techniques [Loubet et al. 2020]. We address this issue by importance sampling the specular chains directly. The proposed method allows us to find admissible paths in a very efficient way and enables the creation of high-frequency caustics from pure specular surfaces with low variance.

Specular light transport. Due to the intrinsic difficulties in specular light transport, many specialized methods have been proposed to sample challenging specular paths.

A line of approaches performs exhaustive searching and root-solving to find all specular chains connecting two endpoints [Mitchell and Hanrahan 1992]. Due to the high computational complexity, these methods either only work for one specular bounce [Loubet et al. 2020; Walter et al. 2009] or incur significant performance degradation as the number of bounces increases [Wang et al. 2020].

Other methods lower the computational burden by stochastic sampling at the cost of introducing variance. Manifold Exploration Metropolis Light Transport (MEMLT) [Jakob and Marschner 2012] allows random walks on a specular manifold with Newton’s method. This, as a pioneer work in this field, is extended to half-vector space [Kaplanyan et al. 2014] and then introduced to regular Monte Carlo sampling as Manifold Next-Event Estimation (MNEE) [Hanika et al. 2015]. MNEE heuristically generates a deterministic initial chain connecting the shading point to the light source, followed by a manifold walk for a feasible solution. With fixed initialization, MNEE can only find at most one specular chain connecting a given pair of endpoints, resulting in energy loss. Specular Manifold Sampling (SMS) [Zeltner et al. 2020] addressed the above issue with random initialization. Accompanied by an unbiased reciprocal probability

Table 1. List of important notations.

Symbol	Definition
\bar{x}_S, \bar{x}_S^*	Specular chain, admissible chain
(x_D, \bar{x}_S, x_L)	Sub-path
x_1, \dots, x_n	Specular vertices of \bar{x}_S
ω_D, ω_D^*	Direction from x_D to x_1, x_1^*
τ_1, \dots, τ_n	Type of the scattering event of each vertex
$T(x_D, \bar{x}_S, x_L)$	Sub-path throughput

estimator, it is able to preserve energy for complex caustics. Unfortunately, the simple treatment of initial guesses by uniform sampling will cause a high variance.

In comparison, we take the energy contribution of each chain into consideration during sampling, forming a new and generic importance sampling paradigm suitable for a wide range of specular chains, including very long chains and chains with complex visibility.

3 BACKGROUND

We first review some background knowledge regarding our topic.

Paths with specular chains. Veach [1998] introduced *non-local sampling* that evaluates the full path integral in Monte Carlo rendering by sampling separators (i.e., diffuse¹ vertices) first and then connecting specular chains (i.e., chains of consecutive specular vertices) between adjacent separators. Hence, importance sampling a full path boils down to importance sampling separators and chains.

While there are existing approaches (e.g., irradiance caching [Ward et al. 1988]) for separator sampling, connecting chains is still a challenge and is our major concern in this paper.

Without loss of generality, we denote x_D and x_L as two separators, where x_D is a diffuse vertex and x_L is on the light source. \bar{x}_S represents a chain connecting x_D and x_L , which is comprised of specular vertices x_1, x_2, \dots, x_n . We call (x_D, x_L) as a *configuration* and (x_D, \bar{x}_S, x_L) as a *sub-path*. For brevity, we also refer to x_D as x_0, x_L as x_{n+1} . These symbols are demonstrated in Fig. 2(a) and summarized in Table 1.

Specular manifolds. To satisfy all physically-based specular constraints along a path, existing solutions mostly build on top of Jakob and Marschner [2012]’s path space manifold. In this context, an *admissible chain* \bar{x}_S^* [Hanika et al. 2015] refers to a specular chain that satisfies all the constraints, which could have a non-zero throughput once it passes the visibility test. In general, we could safely assume that there is a discrete and finite set of admissible chains connecting two separators. Although infinite cases theoretically exist, they rarely occur in natural rendering [Wang et al. 2020; Zeltner et al. 2020] as discussed in our supplemental document.

Throughput between separators. Between two separators x_D and x_L , we use the term *throughput* to refer to the irradiance received at

¹We utilize Heckbert’s notation [Heckbert 1990] for path classification: a full path is described by regular expressions in the form $E(S|D)^*L$, with each letter representing a vertex of a path. For ease of discussion, we assume that each path is comprised of pure specular and diffuse vertices. The extension to glossy will be discussed later.

\mathbf{x}_D contributed by the unit-area surface at \mathbf{x}_L . The throughput can be subdivided into two independent parts [Veach 1998]. The first part comes from the direct connection between \mathbf{x}_D and \mathbf{x}_L , which can be trivially evaluated and thus will not be considered in the following discussions. The second part corresponds to all connections passing through at least one intermediate specular vertex. We formulate it as $\sum_{\bar{\mathbf{x}}_S^* \in \mathcal{P}_S^*} T(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L)$, a summation of all admissible chains, where the admissible chain space $\mathcal{P}_S^* = \mathcal{P}_S^*(\mathbf{x}_D, \mathbf{x}_L)$ is a finite, countable set containing all admissible chains connecting \mathbf{x}_D and \mathbf{x}_L as shown in Fig. 2(b). The *sub-path throughput* $T(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L)$ is part of the throughput contributed by a specified admissible chain $\bar{\mathbf{x}}_S^*$. Its value is determined by the product of several terms [Hanika et al. 2015; Jakob 2013]:

$$T(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L) = \kappa(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L) G(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L) L_o(\mathbf{x}_n^*, \mathbf{x}_L) \quad (1)$$

with κ being a unitless specular scattering value that folds the reflection/refraction coefficients at each specular vertex (e.g., the Fresnel term for reflection), G denoting the Generalized Geometric Term (GGT) [Jakob and Marschner 2012; Wang et al. 2020] which relates the differential solid angle at \mathbf{x}_D to the differential surface area at \mathbf{x}_L (including the visibility function) and $L_o(\mathbf{x}_n^*, \mathbf{x}_L)$ representing the outgoing radiance at \mathbf{x}_L towards \mathbf{x}_n^* .

Manifold walks and seed chain sampling. Existing solutions of specular path sampling [Hanika et al. 2015; Kaplanyan et al. 2014; Zeltner et al. 2020] are mostly built on the top of *manifold walks* [Jakob and Marschner 2012]. Through combining Newton’s method and a reprojection step, it forms a general and efficient solver that can reach an admissible chain from a *seed chain* $\bar{\mathbf{x}}_S$ (i.e., the initial state of the solver).

In general, through the solver, discrete admissible chains in the discrete space \mathcal{P}_S^* can be indirectly sampled by sampling a seed chain first from a continuous probability distribution $p(\bar{\mathbf{x}}_S | \mathbf{x}_D, \mathbf{x}_L)^2$ in the continuous specular chain space \mathcal{P}_S (space formed by all chains of specular vertices) and then solve for an admissible chain via manifold walks. A conditional probability distribution $P(\bar{\mathbf{x}}_S^* | \bar{\mathbf{x}}_S, \mathbf{x}_D, \mathbf{x}_L)$ demonstrates the convergence behavior of manifold walk. This also defines the *convergence basin* [Zeltner et al. 2020]:

$$\mathcal{B}(\bar{\mathbf{x}}_S^* | \mathbf{x}_D, \mathbf{x}_L) = \{\bar{\mathbf{x}}_S | P(\bar{\mathbf{x}}_S^* | \bar{\mathbf{x}}_S, \mathbf{x}_D, \mathbf{x}_L) = 1\}. \quad (2)$$

As illustrated in Fig. 3, MNEE [Hanika et al. 2015] uses a Dirac delta distribution as $p(\bar{\mathbf{x}}_S | \mathbf{x}_D, \mathbf{x}_L)$. When multiple solutions for a given configuration exist, MNEE becomes biased (or helps little to reduce variance when combined with a path tracer) [Hanika et al. 2015]. SMS [Zeltner et al. 2020] adopts a uniform seed chain distribution to enable unbiased sampling. Consequently, the variance can be very high since the probability density ignores the impact of the throughput. The situation becomes worse when the number of specular bounces increases, where the size of each convergence basin becomes extremely small while the probability of finding each chain is still proportional to the size of its convergence basin. This inspires us to reduce the variance and support long specular chains by finding a better seed chain distribution $p(\bar{\mathbf{x}}_S | \mathbf{x}_D, \mathbf{x}_L)$ that is continuous and takes energy distributions into consideration.

²We use an upper-case P to denote discrete probability mass functions and use a lower-case p to represent continuous probability density functions.

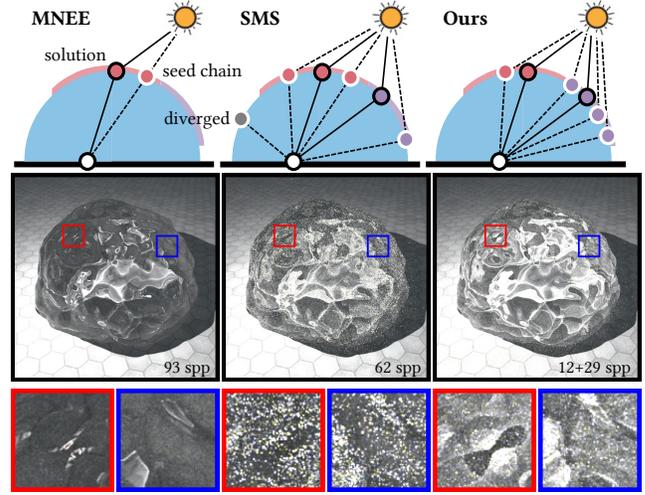


Fig. 3. MNEE [Hanika et al. 2015] easily loses energy due to incomplete solutions, while SMS [Zeltner et al. 2020] produces results with high variance. Our importance sampling method addresses these issues. The bottom row shows equal-time (3 min) rendering results of these methods.

4 PROBLEM FORMULATION

To obtain a proper seed chain distribution mentioned above, we first formulate the problem. Evaluating the throughput summation using Monte Carlo techniques requires sampling the admissible chain $\bar{\mathbf{x}}_S^*$ in \mathcal{P}_S^* for a given configuration $(\mathbf{x}_D, \mathbf{x}_L)$. To achieve this goal, we require a sampling technique that generates random samples of discrete admissible chains with probability $P(\bar{\mathbf{x}}_S^{*(i)} | \mathbf{x}_D, \mathbf{x}_L)$, where i denotes the index of samples used in Monte Carlo estimations. Then we resort to the following estimator:

$$\left\langle \sum_{\bar{\mathbf{x}}_S^* \in \mathcal{P}_S^*} T(\mathbf{x}_D, \bar{\mathbf{x}}_S^*, \mathbf{x}_L) \right\rangle = \frac{1}{N} \sum_{i=1}^N \frac{T(\mathbf{x}_D, \bar{\mathbf{x}}_S^{*(i)}, \mathbf{x}_L)}{P(\bar{\mathbf{x}}_S^{*(i)} | \mathbf{x}_D, \mathbf{x}_L)}. \quad (3)$$

The estimation is unbiased if

$$T(\mathbf{x}_D, \bar{\mathbf{x}}_S^{*(i)}, \mathbf{x}_L) > 0 \Rightarrow P(\bar{\mathbf{x}}_S^{*(i)} | \mathbf{x}_D, \mathbf{x}_L) > 0, \quad (4)$$

but the value of the sampling probability has a dramatic impact on the variance. Ideally, if the probability of sampling each admissible chain is proportional to the throughput, i.e.,

$$P(\bar{\mathbf{x}}_S^{*(i)} | \mathbf{x}_D, \mathbf{x}_L) \propto T(\mathbf{x}_D, \bar{\mathbf{x}}_S^{*(i)}, \mathbf{x}_L), \quad (5)$$

the estimation will reach zero variance and achieve the best possible importance sampling³.

Unfortunately, this task is not as simple as it seems and differs significantly from sampling non-specular vertices due to the discrete property of the \mathcal{P}_S^* space. Traditional path sampling techniques designed for continuous target distributions will fail since they hit each admissible chain with almost zero probability.

Our key insight is that, from a *well-designed continuous distribution*, a *seed chain* can converge to an admissible chain satisfying both the constraints and the requirements of importance sampling.

³The discussion here is under the assumption of diffuse separators, where the throughput equals the contribution to the full path.

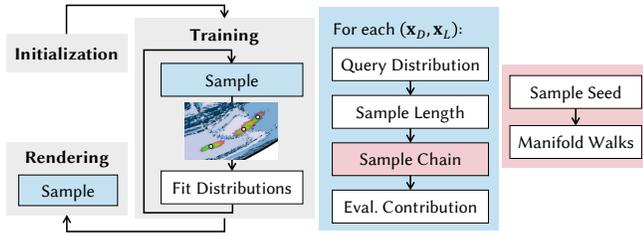


Fig. 4. **An overview of our proposed pipeline.** The pipeline employs a progressive manner to gather sub-path samples (indicated by green plus signs), which are then utilized to fit continuous distributions (marked in orange). These distributions are essential for the importance sampling of seed chains, which ultimately converge to admissible chains (marked by white dots) if they fall within the convergence basins (colored in blue with varying brightness to distinguish different basins).

Therefore, one key to the above problem is to find such a continuous distribution $p(\bar{x}_S | x_D, x_L)$ from existing information about the discrete admissible chains.

5 MANIFOLD PATH GUIDING

In this section, we present a practical rendering method, *manifold path guiding*, which allows us to address the above problem and hence achieve the goal of importance sampling specular chains.

To begin with, we will provide a brief summary of the pipeline, followed by discussions of each stage and some practical issues. Please refer to the supplemental document for pseudo-code snippets in Python style with explanations.

5.1 Overview

In general, as shown in Fig. 4, our pipeline consists of three stages:

- In the *initialization* stage, we set an initial distribution that specifies how to sample seed chains without the knowledge of historical sub-path samples.
- During the *training* stage, progressive training is achieved by reconstructing distributions from previously found sub-path samples, sampling new seed chains, and performing manifold walk.
- In the *rendering* stage, the final converged distribution is adopted to estimate the throughput between separators.

As aforementioned, we focus on specular chain sampling. Therefore, in our current implementation, we only handle the case in which x_L lies on light sources. We pick x_D by path tracing and x_L by uniformly sampling all emitters.

Our main goal is to reconstruct continuous distributions using the historical sub-path samples from previous iterations. After progressive refinement, the continuous distribution is expected to reflect the actual energy distribution and approach the optimal distribution for seed chain importance sampling.

Therefore, the foundation of our online training and rendering pipeline is a seed chain importance sampling strategy that leverages the historical sub-path samples. This is made possible by the following analyses of the continuity of the specular manifold and the light transport behaviors of specular paths.

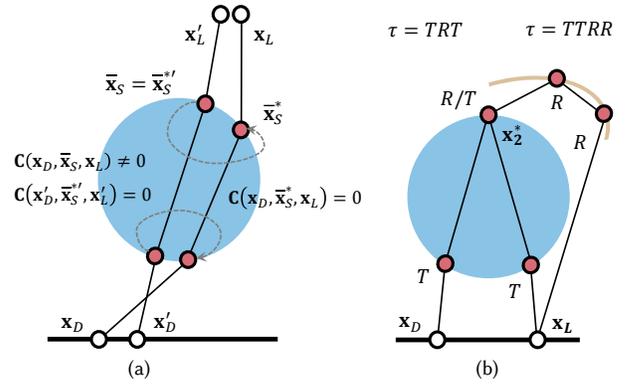


Fig. 5. (a) Two configurations (x_D, x_L) and (x'_D, x'_L) that are close to each other. They tend to generate admissible sub-paths (x_D, \bar{x}_S^*, x_L) and $(x'_D, \bar{x}_S'^*, x'_L)$ through similar specular chains \bar{x}_S^* and $\bar{x}_S'^*$. An admissible chain \bar{x}_S^* can be acquired starting from a seed chain $\bar{x}_S = \bar{x}_S'^*$ using manifold walk. (b) The scattering type is essential to ensure the uniqueness of the simplified chain description since reflection and refraction can happen at the same vertex (x_2^*) simultaneously.

5.2 Exploitation of continuity

Although the admissible chains under a specific configuration are usually discrete, all sub-paths in the scene form manifolds in the specular chain space \mathcal{P}_S , i.e., the specular manifolds [Jakob and Marschner 2012; Zeltner et al. 2020].

The continuity of the specular manifold is guaranteed by the Implicit Function Theorem [Jakob and Marschner 2012]. It tells that the function $\bar{x}_S^*(x_D, x_L)$ is continuous in an infinitely small neighborhood around an admissible sub-path, and \bar{x}_S^* lies in the convergence basin⁴ of that admissible sub-path. This allows us to use the admissible chain $\bar{x}_S'^*$ of a nearby configuration (x'_D, x'_L) to be the seed chain \bar{x}_S of the current configuration (x_D, x_L) . From this seed chain \bar{x}_S , an admissible sub-path (x_D, \bar{x}_S^*, x_L) could be obtained after convergence. Fig. 5 (a) gives a visual explanation of this process.

Moreover, the continuity of specular manifolds also results in the continuity of the sub-path throughput⁵, which means the throughput is nearly constant in the neighborhood. Consequently, we can optimally importance sample the admissible chain under the current configuration by sampling the admissible chain under all nearby configurations according to their throughput. This will be particularly efficient when the throughput has a small value of derivatives to x_D and x_L . Now the problem has been significantly simplified – the distribution of all nearby admissible chains is a continuous distribution, which can be fitted with any possible continuous distribution reconstruction technique.

⁴Manifold walks have quadratic convergence to an admissible chain when the seed is close to the solution [Jakob and Marschner 2012; Zeltner et al. 2020]. This guarantees the existence of a convergence basin in the *neighborhood* of an admissible chain, which can be used to solve for an admissible chain starting from an approximated one.

⁵In the neighborhood N of an admissible sub-path, \bar{x}_S^* has continuous partial derivatives to x_D and x_L , which reveals the continuity of GGT. Assuming that κ and L in Eq. (1) are locally continuous, the sub-path throughput T is also continuous.

5.3 Dimensionality reduction

As the length of the specular chain increases, there will be a risk of the curse of dimensionality. We avoid this by using a marginal. Let τ denote the *type* of a specular chain \bar{x}_S . It is a string consisting of T and/or R , with i -th letter τ_i describing the scattering type at i -th vertex of the chain. For a specific configuration $(\mathbf{x}_D, \mathbf{x}_L)$, once τ is given and $\mathbf{x}_1 = \mathbf{r}(\mathbf{x}_D, \omega_D)$ is determined, the rest of the vertices $\mathbf{x}_2, \dots, \mathbf{x}_n$ can be deduced using the following recursion:

$$\mathbf{x}_i = \mathbf{r}(\mathbf{x}_{i-1}, \mathbf{s}_{\tau_{i-1}}(\mathbf{x}_{i-1}, \overrightarrow{\mathbf{x}_{i-2}\mathbf{x}_{i-1}})) \quad (6)$$

where the operator $\mathbf{r}(\mathbf{x}, \omega)$ returns the intersection of the ray (\mathbf{x}, ω) with all specular surfaces \mathcal{M}_S , and the operator \mathbf{s} returns the scattered direction for given scattering type. Consequently, an admissible sub-path $(\mathbf{x}_D, \bar{x}_S, \mathbf{x}_L)$ can be determined with $(\mathbf{x}_D, \omega_D^*, \mathbf{x}_L, \tau)$. Note that the scattering type is necessary to ensure uniqueness since reflection and refraction can happen at the same vertex simultaneously, as visually explained in Fig. 5(b).

Following the above representation, we can sample (ω_D, τ) and use Eq. (6) to obtain a seed chain \bar{x}_S . It converts a high-dimensional sampling problem into a much lower one, and the probability of sampling a seed chain can be further subdivided into two factors, i.e.,

$$p(\bar{x}_S | \mathbf{x}_D, \mathbf{x}_L) = P(n | \mathbf{x}_D, \mathbf{x}_L) p(\omega_D, \tau | \mathbf{x}_D, \mathbf{x}_L, n). \quad (7)$$

With the above analysis, we divide the sampling task of a seed chain into three parts: sampling the number of bounces n , sampling the scattering type τ , and sampling the direction ω_D . This lays the foundation for efficient data-driven importance sampling. Note that the order of the last two steps may be changed due to practical considerations.

5.4 Initialization

The initial distribution $p_0(\bar{x}_S | \mathbf{x}_D, \mathbf{x}_L)$ decides how initial seed chains are sampled only according to a priori knowledge. In theory, any distribution that covers all the admissible chains can be selected. Better choices can be adopted with the help of other existing algorithms. For instance, we can reconstruct energy distributions from photons tracing from emitters. Since reconstructed distributions are not fully conservative, we could further incorporate a uniform distribution to avoid making some regions never discovered [Li et al. 2022; Müller et al. 2017; Zhu et al. 2021b].

In our implementation, we use the following initialization:

- The length of the chain is sampled by simulating a Russian Roulette (RR) defined in the underlying path tracer.
- The direction ω_D is determined by uniformly sampling positions as the first specular vertex \mathbf{x}_1 on all specular objects, yielding $p_0(\omega_D | \mathbf{x}_D, \mathbf{x}_L)$.
- The remaining specular vertices $\mathbf{x}_2, \dots, \mathbf{x}_n$ are generated by ray tracing, deciding the scattering type τ (i.e., reflection or refraction) at each vertex according to their BSDF.

This covers all the possible (ω_D, τ) pairs and is equivalent to covering all the admissible chains according to Eq. (7). Consequently, each solution will be reached with non-zero probability.

5.5 Importance sampling seed chains

The continuity discussed in Section 5.2 allows us to perform importance sampling leveraging historical samples of neighboring configurations. Now we reconstruct discrete distributions for n and τ and continuous distributions for ω_D , rewriting Eq. (7) as

$$p(\bar{x}_S | \mathbf{x}_D, \mathbf{x}_L) = P(n | \mathbf{x}_D, \mathbf{x}_L) P(\tau | \mathbf{x}_D, \mathbf{x}_L, n) p(\omega_D | \mathbf{x}_D, \mathbf{x}_L, \tau). \quad (8)$$

Gathering neighboring samples. We first perform a neighbor searching process in a set⁶ of previously obtained sub-path samples \mathcal{S} , gathering a set of sub-paths with the nearest configurations to $(\mathbf{x}_D, \mathbf{x}_L)$, denoted as $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L)$. Then, we evaluate the weight of a sub-path sample, $w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L)$, which equals to the product of its throughput and the reciprocal probability estimation:

$$w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) = T(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) \left\langle \frac{1}{P(\bar{x}'_S | \mathbf{x}'_D, \mathbf{x}'_L)} \right\rangle. \quad (9)$$

Here, the throughput and reciprocal probability are computed in earlier iterations and stored in memory.

Sampling chain length. The length of a specular chain determines the dimensionality of the specular chain space and thus needs to be sampled first. By accumulating and normalizing the weights for each length, we obtain a discrete probability distribution $P_e(n | \mathbf{x}_D, \mathbf{x}_L)$ ⁷ from $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L)$:

$$P_e(n | \mathbf{x}_D, \mathbf{x}_L) = \frac{\sum_{(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) \in \mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, n)} w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L)}{\sum_{(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) \in \mathcal{S}(\mathbf{x}_D, \mathbf{x}_L)} w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L)}, \quad (10)$$

where $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, n)$ denotes all the samples of length n in $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L)$. This will be used when sampling the length of a new seed chain.

Sampling scattering type. Once n is determined, we sample the scattering type from a discrete distribution reconstructed by summing up the total weights of each type:

$$P_e(\tau | \mathbf{x}_D, \mathbf{x}_L, n) = \frac{\sum_{(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) \in \mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, n, \tau)} w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L)}{\sum_{(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L) \in \mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, n)} w(\mathbf{x}'_D, \bar{x}'_S, \mathbf{x}'_L)}. \quad (11)$$

Let $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, \tau)$ denote the set of samples with type τ . The final step is sampling the direction from $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, \tau)$.

Sampling the direction. We blur each chain in $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, \tau)$ with a wide kernel in the directional domain. Specifically, we estimate the directional footprint [Hey and Purgathofer 2002] of i -th chain by finding a sample in the $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L, \tau)$ with the nearest ω_D^* to it and evaluating their directional distance σ'_i . Then, we place a von Mises-Fisher (vMF) lobe [Fisher 1953] along the direction ω_D^* :

$$v(\omega_D; \boldsymbol{\mu}_i, \kappa_i) = \frac{\kappa_i}{4\pi \sinh(\kappa_i)} e^{\kappa_i \boldsymbol{\mu}_i \cdot \omega_D} \quad (12)$$

where we set $\boldsymbol{\mu}_i = \omega_D^*$, $\kappa_i = \sigma_i'^{-2}$. With the weights of each lobe being proportional to the corresponding sub-path sample's weight defined in Eq. (9), we obtain a mixture of vMF lobes $p_e(\omega_D | \mathbf{x}_D, \mathbf{x}_L, \tau)$.

⁶ \mathcal{S} can include either all the samples already encountered [Reibold et al. 2018] or only the samples found in the last iteration [Müller et al. 2017]. We use the latter by default.

⁷ We use subscript e for fitted distributions and subscript 0 for initial distributions.

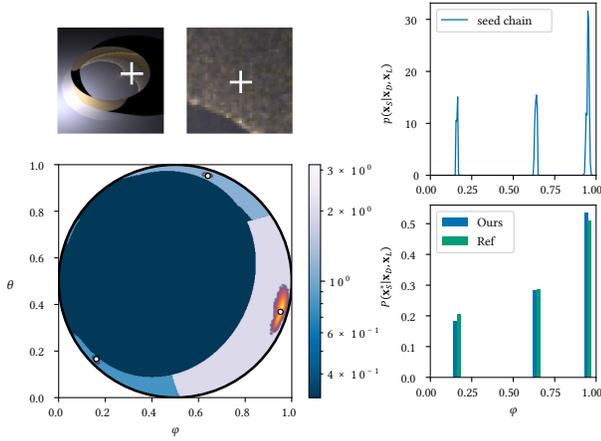


Fig. 6. **An example validating our solution of importance sampling discrete specular chains.** We choose a typical configuration in the RING scene: \mathbf{x}_L on a small emitter and \mathbf{x}_D marked by the white plus sign. The convergence basin and our sampling distribution are visualized in spherical coordinates. Two diagrams in the right panel visualize the discrete distribution of admissible chains $P(\bar{\mathbf{x}}_S^{(i)} | \mathbf{x}_D, \mathbf{x}_L)$ for this configuration (bottom-right) and the continuous distribution $p(\bar{\mathbf{x}}_S | \mathbf{x}_D, \mathbf{x}_L)$ reconstructed from historical samples (top-right).

Validation and discussion. In Fig. 6, we verify that the proposed method can achieve nearly optimal importance sampling for discrete admissible chains. Here, we marginalize the reconstructed distribution of seed chains along the horizontal flatland, shown in the top-right diagram. In the bottom-right diagram, we evaluate the corresponding sampling probability (blue) and the expected probability according to their throughput (green). As seen, the probability of finding each admissible chain closely matches their throughput. In theory, this is still possible with high-frequency details and complex visibility, despite requiring more samples.

Defensive sampling. In order to cover all admissible chains, we further perform Multiple Importance Sampling (MIS) [Veach and Guibas 1995b] with the initial distribution. We use a one-sample MIS model for sampling both n and (ω_D, τ) :

$$\begin{aligned} P(n | \mathbf{x}_D, \mathbf{x}_L) &= \alpha P_0(n | \mathbf{x}_D, \mathbf{x}_L) + (1 - \alpha) P_e(n | \mathbf{x}_D, \mathbf{x}_L) \\ p(\omega_D, \tau | \mathbf{x}_D, \mathbf{x}_L) &= \alpha p_0(\omega_D, \tau | \mathbf{x}_D, \mathbf{x}_L) + (1 - \alpha) p_e(\omega_D, \tau | \mathbf{x}_D, \mathbf{x}_L) \end{aligned} \quad (13)$$

where α is the probability of choosing P_0 . Following the convention in many path guiding methods, we always set $\alpha = 0.5$ [Müller et al. 2017; Reibold et al. 2018; Zhu et al. 2021a].

5.6 Spatial structures

In Section 5.5, we gather nearby sub-path samples and construct vMF lobes to represent the directional distribution. However, the choice of spatial hierarchy significantly affects efficiency and robustness.

Performing an accurate k-nearest neighbor search (implemented with a 6D kd-tree) would be computationally expensive. Instead, we use an approximated solution. We organize all the samples in a 6D adaptive binary tree (STree) following [Müller et al. 2017] and

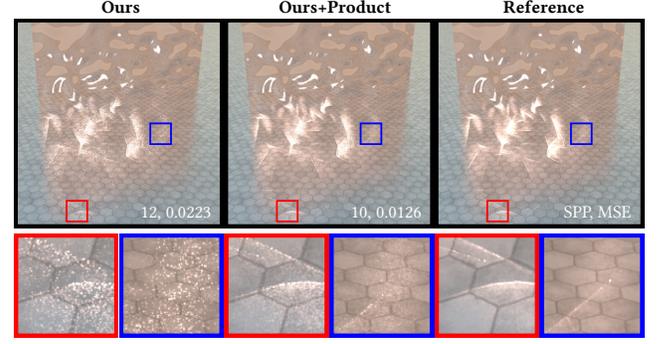


Fig. 7. **Equal-time (5 min) comparison** of our approach without and with product importance sampling with the separator’s BSDF. A normal-mapped reflective specular plane is placed on a glossy ($\alpha = 0.05$) floor.

set the spatial threshold (i.e., the maximum number of samples in a leaf node) to $\sqrt{|S|}$. This is generally similar to [Müller 2019], but we use the number of total samples rather than the samples per pixel and omit the coefficient. Our strategy is independent of image resolution and works well in all our test scenes⁸. When subdividing nodes, we perform spatial filtering by allowing a small overlapping between child nodes⁹. This prevents artifacts introduced by spatial division [Müller 2019; Ruppert et al. 2020; Zhu et al. 2021a].

5.7 Handling glossy cases

Our earlier discussions are under the assumption of diffuse separators and pure specular chains. However, our method can be generalized to support glossy separators and chains.

Glossy separators. When the separator is not diffuse, to achieve optimal importance sampling, we should perform product importance sampling [Herholz et al. 2016] with the separator’s BSDF. Since we use sub-path samples to reconstruct the distributions, we can directly modify their weights before sampling, and Eq. (9) becomes

$$w(\mathbf{x}'_D, \bar{\mathbf{x}}_S', \mathbf{x}'_L) = \rho_{\mathbf{x}'_D}(\omega'_i, \omega_D^*) T(\mathbf{x}'_D, \bar{\mathbf{x}}_S', \mathbf{x}'_L) \left\langle \frac{1}{P(\bar{\mathbf{x}}_S' | \mathbf{x}'_D, \mathbf{x}'_L)} \right\rangle. \quad (14)$$

Here, ω'_i in the BSDF term $\rho_{\mathbf{x}'_D}(\omega'_i, \omega_D^*)$ is the direction towards \mathbf{x}'_D and is already determined. All we need is to traverse through each sub-path sample in $\mathcal{S}(\mathbf{x}_D, \mathbf{x}_L)$ and apply the above weighting scheme. This can significantly reduce noise when separators have small roughness, as shown in Fig. 7.

Glossy chains. For glossy chains, once manifold offset [Jakob and Marschner 2012; Kaplanyan et al. 2014] is sampled, it boils down to pure specular cases, and the admissible chains corresponding to the offset are still finite. This is unbiased since each chain can be found with non-zero probability. Like prior work [Hanika et al. 2015; Kaplanyan et al. 2014; Zeltner et al. 2020], we sample the *offset*

⁸One can further fine-tune the performance by using $c\sqrt{|S|}$, as shown in our supplemental document. Generally, we recommend $c = 1$, which is simple and straightforward.

⁹Specifically, we copy ϵk sub-path samples from the left subtree to the right subtree and vice versa. We also discard the samples far from the bounding box of a node. Overall, we observe that $\epsilon = 10\%$ works well across all our test scenes.

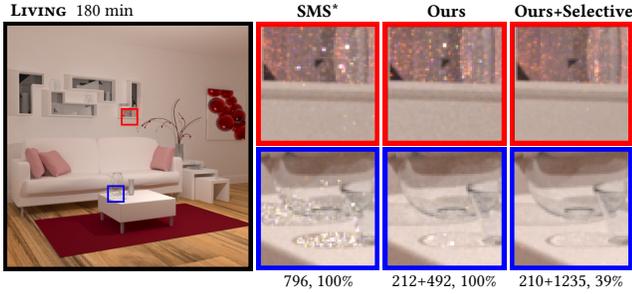


Fig. 8. **Selective activation** on a realistic scene, where caustics only cover a small part of images. We report SPP and the activation rates.

normal from the microfacet distribution before manifold sampling. Then, we sample seed chains and admissible chains constrained to this offset normal.

5.8 Practical considerations

Reciprocal probability estimation. To estimate the reciprocal probability in Eq. (9), we repeat independent trials once the sampled admissible chain was found, in the same way as in SMS [Zeltner et al. 2020]. However, some portions of the reciprocal probability can be evaluated analytically. In particular, $P(n|\mathbf{x}_D, \mathbf{x}_L)$ is determined when we sample the length of the chain. Therefore, it can be factored out of the stochastic estimation, i.e.,

$$\left\langle \frac{1}{P(\bar{\mathbf{x}}_S^*|\mathbf{x}_D, \mathbf{x}_L)} \right\rangle = \frac{1}{P(n|\mathbf{x}_D, \mathbf{x}_L)} \left\langle \frac{1}{P(\bar{\mathbf{x}}_S^*|\mathbf{x}_D, \mathbf{x}_L, n)} \right\rangle \quad (15)$$

and only $P(\bar{\mathbf{x}}_S^*|\mathbf{x}_D, \mathbf{x}_L, n)$ needs to be estimated in our implementation. In other words, we *reuse* n across trials.

Online training and rendering. We use an online training strategy that progressively gathers sub-path samples. We separate the training stage into several iterations with increasing budgets. Specifically, we double the sample count in each iteration [Müller et al. 2017]. We stop the training stage immediately when 30% of the budget is already used. Besides, when a training iteration is finished, we start a new round if less than half of the training budget is used. Otherwise, we continue the current iteration until the rendering stage starts. Considering that high variance may occur during the training phase, we never splat the samples produced in the training stage into the final image. Besides, we only use samples from the last iteration to reconstruct sampling distributions.

Selective activation. Our learned distribution can also be employed to determine when specular chain sampling should be activated [Loubet et al. 2020], which avoids the costly sampling process for non-caustic regions [Zeltner et al. 2020]. In the rendering stage, if no samples (excluding those used for filtering) are in the corresponding leaf node of the spatial hierarchy for a given configuration, we simply switch to using path tracing. In other words, we perform a combination with path tracing according to the existence of nearby sub-path samples. This significantly reduces overhead and leads to more samples in equal time, as illustrated in Fig. 8.

6 RESULTS

We have implemented our approach in the Mitsuba 2 Renderer [Nimier-David et al. 2019] as a new integrator. In this section, we will describe our experimental setup, compare our approach to related methods, and validate our building blocks.

All our test scenes have been rendered on a compute node with 32 2.50 GHz cores of Xeon E5-2682 v4 processors. Since we focus on high-frequency light transport, we only use artificial point and small area light sources to cast caustics and handle them using the proposed approach. Other light transport portions involving environmental lighting are handled by conventional path tracing. We set the max length of paths to 15. Russian roulette starts from bounce 5 with probability $\gamma = 0.95$. Unless otherwise mentioned, we do *not* enable product importance sampling or selective activation for a more fair comparison with related methods.

The reference images of LAMP, FLOWER, and DOUBLE SLABS are generated using path tracing, whereas the others are rendered with our modified SMS at high sample rates. Some of our results report the number of samples per pixel in the form of “ $n + m$ ”, meaning n spp for training and m spp for rendering.

6.1 Comparisons with previous approaches

Comparison with unbiased MC approaches. In Fig. 9, we show the equal-time comparison of our approach against PT, PPG [Müller 2019], VAPG [Rath et al. 2020], VMMPG [Ruppert et al. 2020], the original SMS [Zeltner et al. 2020], and our modified SMS on three scenes: LAMP, FLOWER and DOUBLE SLABS.

Conventional guiding approaches, such as PPG, rely on online learned distributions to perform path sampling. They generally work well but will encounter difficulties in dealing with glossy or specular interactions since the radiance distribution involves high-frequency variations. Visual comparisons clearly show that a typically guided path tracer, such as PPG, VAPG, and VMMPG, fails to handle challenging light paths involving specular vertices, leading to many outliers and energy loss.

Although our approach also adopts reconstructed energy distributions as guidance, its usage is quite different. First, our distribution is reconstructed from admissible sub-path samples, which is beneficial for exploring high-frequency regions. Second, the reconstructed distribution is used to sample seed chains instead of light paths. Therefore, even inaccurate distributions can still work well. Consequently, our method succeeds in covering all challenging lighting effects with much fewer fireflies.

SMS is the state-of-the-art approach tailored for the unbiased simulation of specular light transport. The original SMS can only support a user-specified and fixed chain length. We extend it (denoted as SMS*) using our initialization (Section 5.4) to support various specular chain lengths and scattering types. Although this enables handling all types of specular chains, it has the risk of higher variance (see the LAMP scene) than the original SMS since the sample budget is amortized for various types of chains.

Thus, to further demonstrate the effect of our importance sampling strategy, we also choose a dominant type of chain for each scene (*RT* for LAMP, *TTR* for FLOWER, and *TTTT* for DOUBLE SLABS) and present the results rendered by the original SMS in Fig. 9. Even

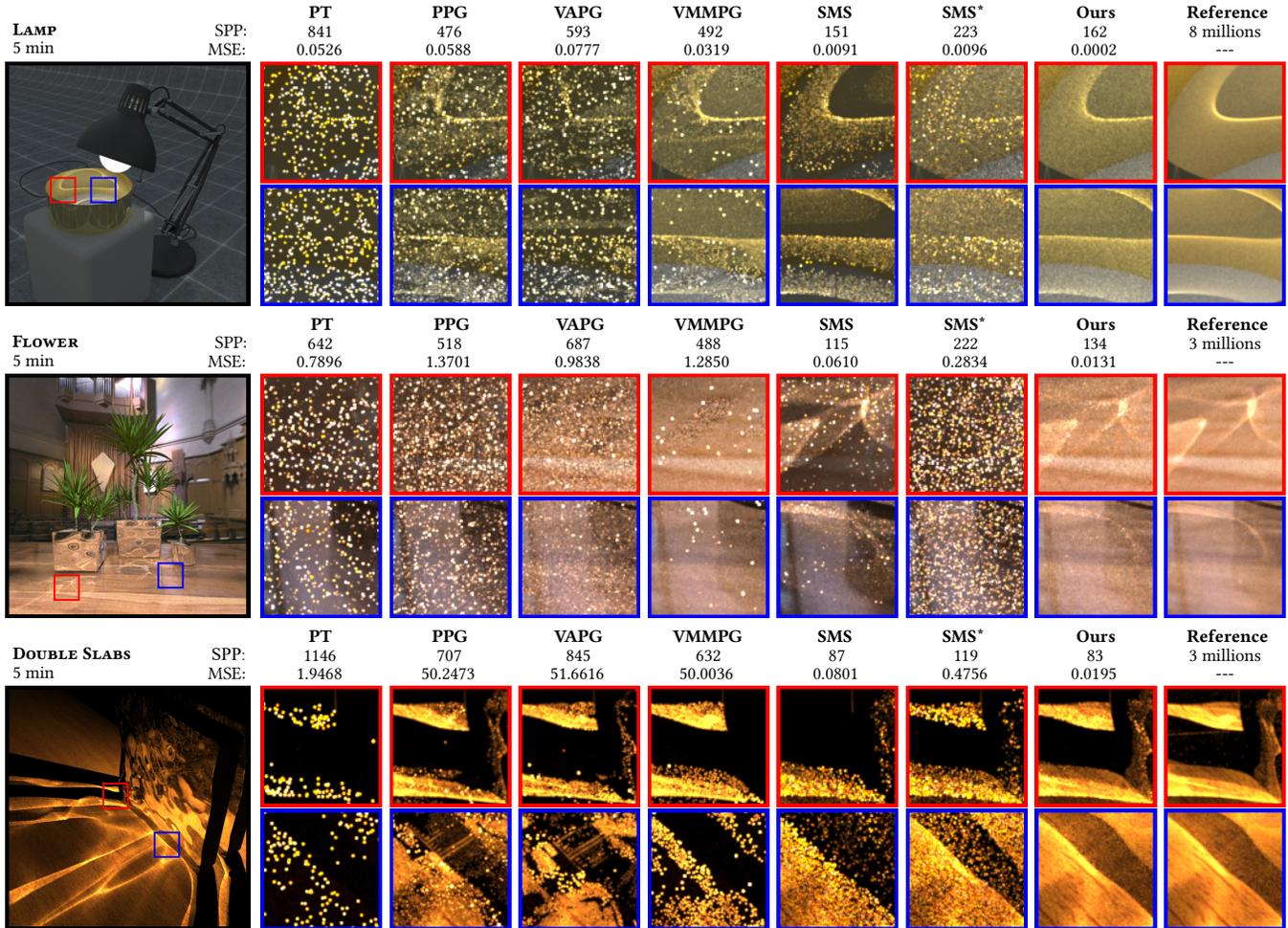


Fig. 9. **Equal-time comparisons** among Path Tracing (PT), the improved Practical Path Guiding (PPG) [Müller 2019; Müller et al. 2017], Variance-Aware Path Guiding (VAPG) [Rath et al. 2020], parallax-aware VMM Path Guiding (VMMPG) [Ruppert et al. 2020], the original Specular Manifold Sampling (SMS) [Zeltner et al. 2020], a modified version of SMS (SMS*) and our method on three scenes.

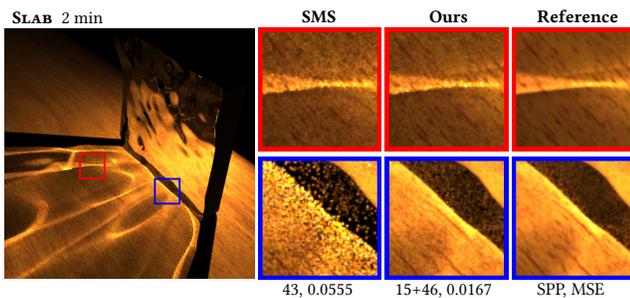


Fig. 10. **Equal-time comparison** between the original SMS [Zeltner et al. 2020] and our method on the SLAB scene. In order to validate the effect of directional sampling, we use a fixed chain length ($n = 2$) and also neglect reflection on the dielectric surface.

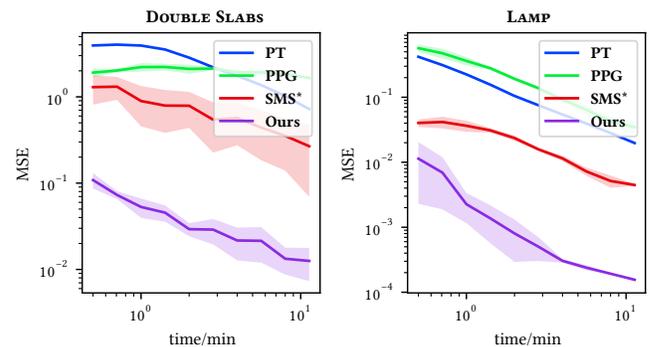


Fig. 11. **Convergence curves of two testing scenes.** We compare our method with PT, PPG, and SMS* on the evolution of MSE with respect to the total rendering (including training) time. The standard deviation of the MSE is shown as shaded regions.

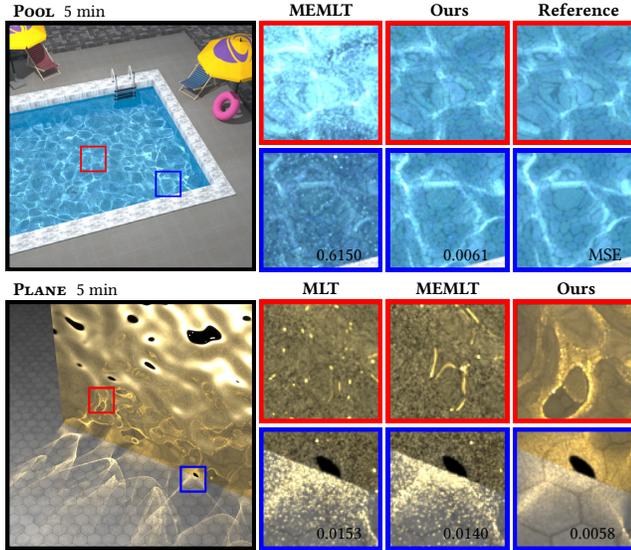


Fig. 12. **Equal-time (5 min) comparison** between MLT [Veach and Guibas 1997], MEMLT [Jakob and Marschner 2012] and our method.

in this way, the original SMS still tends to produce noticeable noise since it samples seed chains uniformly. This is particularly obvious in areas that correspond to admissible chains with extremely small convergence basins, e.g., the bottom of the slabs.

Our method exploits progressively generated sub-path samples to reconstruct and refine continuous distributions for sampling seed chains, thus providing high-quality, unbiased rendering for challenging SDS paths. Even at relatively low sample rates, our method still outperforms its competitors with much lower variance.

We further validate the benefit of our directional importance sampling in Fig. 10, where we disable the sampling of chain length and type in our method and compare with the original SMS to validate the effect of directional importance sampling. As highlighted in the closeups, due to efficient importance sampling, our approach generates more samples and produces results with far less noise. A complete ablation study is shown in the supplemental document, which validates the effectiveness of length sampling, type sampling, and directional sampling, respectively.

In Fig. 11, we evaluate the accuracy in terms of MSE with respect to the running time, showing the fast convergence provided by our method. Thanks to the energy-based sampling distribution, our method consistently outperforms previous ones as the running time increases and achieves a large margin at a high sampling rate.

Comparison with unbiased MCMC approaches. In Fig. 12, we compare our approach with MLT [Veach and Guibas 1997] and MEMLT [Jakob and Marschner 2012], two typical MCMC approaches. The POOL scene is a typical example of challenging paths, and the PLANE scene contains a normal-mapped conductor with roughness ($\alpha = 0.002$). Both MLT and MEMLT require admissible chains as their seed paths. Simply relying on PT or BDPT to find valid seed paths is inefficient for SDS paths. This results in either energy loss if no seed

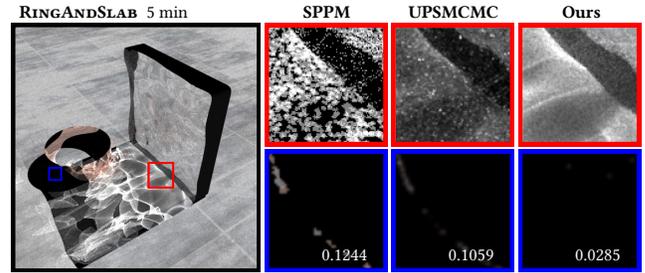


Fig. 13. **Equal-time (5 min) comparison** between SPPM [Hachisuka and Jensen 2009], UPSMCMC [Šik et al. 2016], and our method.

path is found for a specific region (blue closeups in POOL) or over-bright artifacts where Markov chains get stuck in narrow regions (red closeups in both scenes). As our method works in the context of regular Monte Carlo sampling, it is able to produce images free from blotchy artifacts.

Comparison with biased approaches. Biased approaches excel at producing noise-free results for complex caustics and hence prevail in the industry. In Fig. 13, we conduct an equal-time comparison with Stochastic Progressive Photon Mapping (SPPM) [Hachisuka and Jensen 2009] and Metropolised Bidirectional Estimator (UPSMCMC) [Šik et al. 2016]. A slab and a ring are lit by a distant point light encapsulated in a glass sphere.

For a meaningful comparison with SPPM, we decrease the photon lookup radius to the point that the bias is less perceptible. Despite this, the noise level is still higher than ours. It also suffers from visible light leaking at the bottom of the ring. A similar issue occurs for UPSMCMC. Its MCMC nature also results in an uneven convergence characterized by an overly dark region and visible fireflies.

As an unbiased method, our technique performs independent Monte Carlo estimation without spatial relaxing or reuse, thereby avoiding these problems and producing caustics of high quality with sharp details and low noise.

Fig. 3 shows a visual comparison with MNEE [Hanika et al. 2015]. As aforementioned, MNEE cannot find all valid admissible chains in this scene since it relies on deterministic initialization. Generally, this method performs well in simple cases but works inefficiently and tends to be biased (or suffers from extremely high variance if combined with path tracing) in complex scenes with many solutions. The stochastic sampling strategy adopted in our method makes it possible to cover all solutions for a configuration and hence ensures unbiasedness.

6.2 Validation of building blocks

In theory, our pipeline supports most feasible spatial neighbor searching and directional density estimation techniques. However, the decision will significantly affect the quality of the sampling. Here, we endeavor to generate such insights by comparing various alternative choices of pipeline components in Fig. 14 and Fig. 15.

Distribution reconstruction. We employ STree for approximated spatial nearest neighbor searching when reconstructing the energy distribution, and perform an accurate directional density estimation.

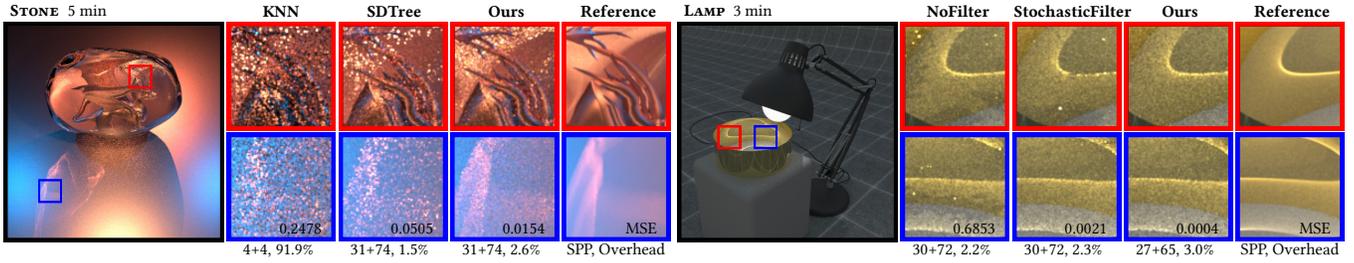


Fig. 14. **Choices of building blocks.** **Left:** we perform equal-time comparisons of various spatial and directional structures. **Right:** comparisons of different spatial filtering strategies. We report the full-image MSE, SPP, and the overhead of sampling. More scenes are shown in the supplemental document.

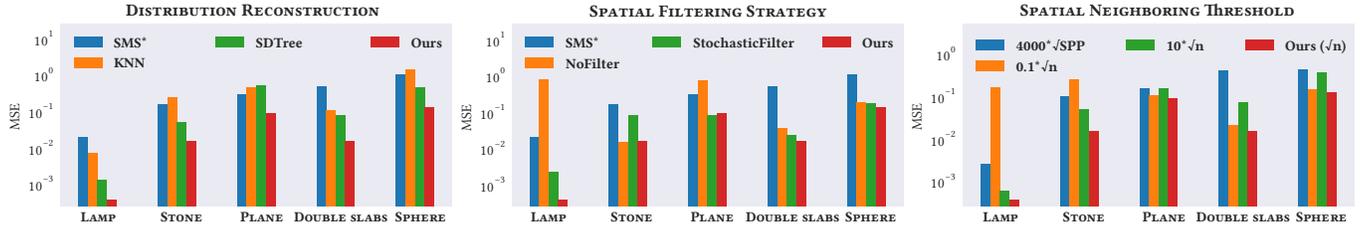


Fig. 15. **Equal-time comparison of different distribution reconstruction techniques, spatial filtering strategies, and neighboring thresholds** on five scenes featuring different light transport effects and complexity. The corresponding images are shown in the supplemental document.

To evaluate their effectiveness, we compare our method against adaptive quad-tree [Müller et al. 2017] (*SDTree*) and a variant of our approach with accurate spatial nearest neighbor searching (*KNN*).

Despite its overall feasibility, *SDTree* introduces more noise due to its inaccuracy in reconstructing high-frequency distributions. In contrast, our accurate local density estimation yields more stable convergence with fewer outliers. Accurate *KNN* introduces substantial overhead (around 90% in most scenes) since gathering thousands of nearest samples is costly. On the contrary, querying the *STree* only requires descending from the root to the leaf node. Consequently, the total overhead of querying and sampling distributions never exceeds 5% in all our experiments.

Spatial filtering. We have employed the spatial filtering (*Ours*) to improve robustness. Along with this, we include comparisons with alternative choices: no spatial filtering (*NoFilter*) and Müller [2019]’s stochastic filtering (*StochasticFilter*) that randomly jitter the sample’s position. Simply disable filtering produces inferior results with visible artifacts. Although stochastic filtering has been widely adopted in previous works [Ruppert et al. 2020; Zhu et al. 2021a], it also leads to slightly less accuracy and more fireflies. In contrast, our approach produces results with less noise and fewer outliers with nearly no performance degradation, striking an appropriate balance between overhead and performance.

Spatial neighboring threshold. We also validate our automatic spatial neighboring threshold in Fig. 15. Our automatic threshold always predicts a reasonable value and generally outperforms alternative choices, including Müller [2019]’s threshold ($4000 \cdot \sqrt{SPP}$) and variants of our threshold with different coefficients.

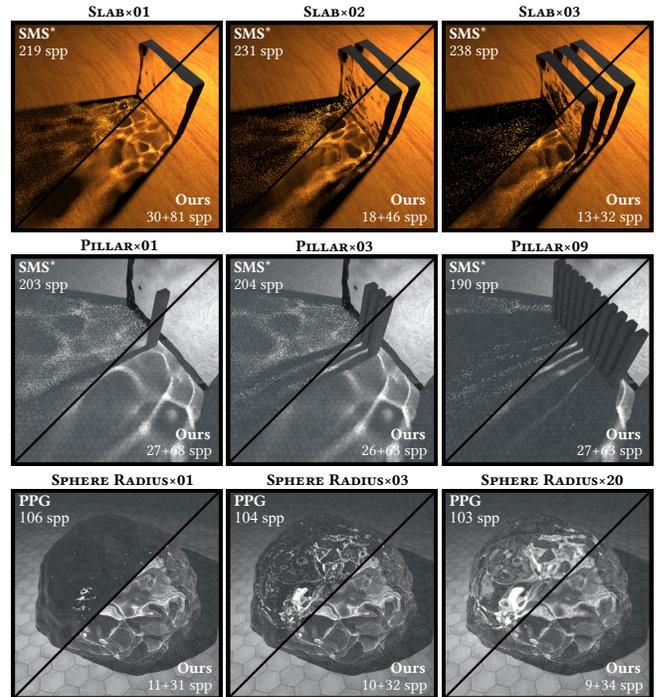


Fig. 16. **Equal-time (5 min) comparison** on scenes with varying chain lengths (top), visibility complexity (middle), and emitter radii (bottom).

6.3 Impact of scene complexity

Our method adapts well to scenes with various intricate structures.

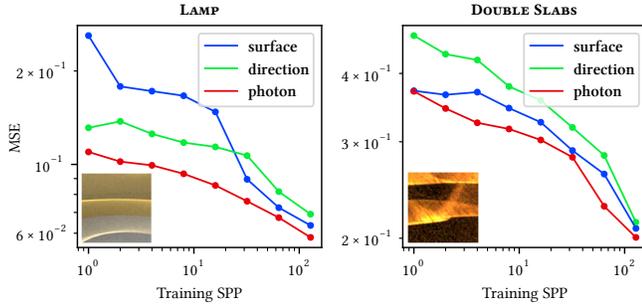


Fig. 17. **The impact of training budget and initialization strategies on the rendered image quality.** The number of training samples per pixel is given as the x-axis. Each image is finally generated with 32 spp.

Chain length. The first row of Fig. 16 conducts a visual comparison against SMS on a series of scenes with an increasing number of transparent slabs. Since SMS neglects energy distributions during specular chain sampling, the variance increases prominently as the number of specular bounces increases. In contrast, by exploiting the continuity of specular manifolds, our method keeps a low variance even in scenes with long specular chains.

Visibility. The second row of Fig. 16 shows a series of scenes with increasing degrees of complex visibility stemming from multiple pillars near a transparent slab. SMS hardly samples solutions that pass the visibility test, while we consider visibility by reconstructing distributions from unblocked admissible chains. As a result, our method avoids high variance introduced by occlusions and works well in complex scenes.

Size of light sources. In the last row of Fig. 16, we present a series of scenes with various emitter sizes. As the size of the emitter decreases, the frequency of caustics increases. While path-guiding methods work well for large area lights, producing high-frequency caustics can be challenging. Instead, our proposed method addresses this issue by exploring admissible chains through manifold walks, allowing us to handle small area lights and complex caustic patterns effectively.

6.4 Choice of strategies

Our pipeline incorporates some design choices. We conduct the following experiments to validate their effectiveness.

Training budgets. Fig. 17 shows the influence of the training budget on the rendering with the same sample rate. Generally, the variance reduces significantly as the number of training samples increases. This indicates that the training process gradually makes the sampling distribution optimal.

Initialization. We also compare different initialization strategies in Fig. 17. Here, *surface* means uniformly sampling x_1 on specular surfaces, *direction* means uniformly choosing ω_D on the hemisphere, and *photon* means reconstructing the initial distribution from 250,000 photons traced by a photon mapper. The comparison shows that initialization using photons leads to better results at the

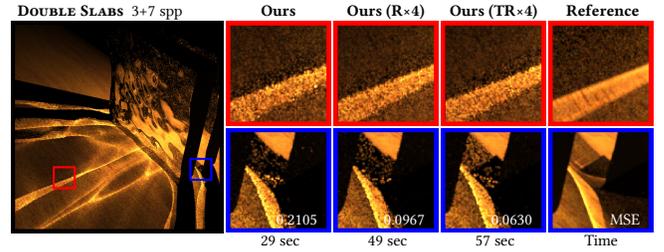


Fig. 18. The variance of reciprocal probability estimation can be reduced by repeating the process multiple times (4 in this example) and using the average, either in the rendering phase only (R×4) or both training and rendering phases (TR×4). This example also showcases the notable overhead of the estimation process.

early stage of training. However, after convergence, different initialization strategies perform similarly. This shows that our method is robust to the initial state.

6.5 Performance analysis

In Table 2, we report the rendering statistics of several test scenes used in this paper. As seen, for each scene, 50%-98% of the time is spent on specular chain sampling. In particular, a small amount of time (no more than 5% in all the scenes) is spent on guiding (i.e., querying the spatial and directional distributions). In complex scenes, the reciprocal probability estimation has a non-negligible run-time cost since many trials are required.

Our implementation needs to store the original sample, which could cause relatively heavy storage. Our consideration is twofold. On the one hand, the storage required for each raw sample is only 40 bytes. On the other hand, it is worthy of being done because the local density estimation is more accurate than existing online fitting approaches when faced with near-delta radiance distributions.

7 DISCUSSION AND FUTURE WORK

There are several challenges not yet solved in our method.

Overhead and variance of reciprocal probability estimation. Our importance sampling works as a promising variance reduction technique, but it only considers the variance of the sampling process itself. We still rely on repeating the whole chain sampling process to find the reciprocal probability of an admissible chain. Unfortunately, the overhead and variance may be high in complex scenes where there are many solutions with small convergence basins for a configuration. Figure 18 is an example showing the overhead and visual impact of the variance of reciprocal probability estimation. While the noise of the estimation visibly affects the image, how to balance the variance and overhead remains a challenge for future research.

Online training pitfalls. The training process relies on an initialization sampling strategy. This leads to a general issue that too complex or too large specular surfaces, as well as light sources, will make the training process slow. Insufficient training will lead to fireflies or temporal instability. One case is demonstrated in Fig. 19.

Table 2. Rendering statistics of our experiments.

Scene	Figure	Chain types (dominant)	Budget (min.)		Sample per pixel		Time on specular (min.)			#Sub-path samples
			Total	Training	Training	Rendering	Guide	PDF	Total	
GLASS	Fig. 1	$TT, TTTT, TTTTTT$	240.0	72.0	80	195	3.592	77.886	215.328	12 519 949
LAMP	Fig. 9	T, RT	5.0	1.5	46	116	0.233	0.388	3.377	1 347 096
FLOWER	Fig. 9	TT, TTR	5.0	1.5	16	134	0.053	1.971	3.983	1 189 876
DOUBLE SLABS	Fig. 9	$TT, TTTT$	5.0	1.5	24	59	0.096	1.433	4.127	2 066 782
WATER	Fig. 19	$TTTT, TTTTTT$	240.0	72.0	41	99	2.621	223.176	236.883	1 667 185

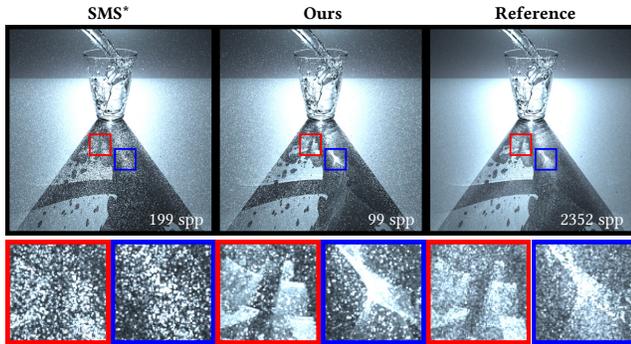


Fig. 19. Equal-time (4 hours) rendering of the WATER scene, demonstrating a failure case for our approach due to tiny caustic casters (many small droplets in this scene).

However, this is a general stubborn problem for rendering algorithms involving MC sampling. Tiny geometries make the path space complicated and hard to explore by sampling without geometry awareness [Otsu et al. 2018]. Using a deterministic strategy to acquire sampling distributions may alleviate this issue. Outlier removal [Reibold et al. 2018], or biased variants [Zeltner et al. 2020], can also be used in practice.

8 CONCLUSION

We have studied the problem of importance sampling specular chains that involve multiple consecutive specular bounces. Existing methods often overlook energy distributions in the discrete admissible chain space, resulting in high variance. To solve this problem, we have developed a comprehensive approach incorporating a specially designed continuous space for seed chain sampling. Following that, we have implemented a practical pipeline that leverages manifold path guiding to explore all the specular chains in a given scene. Additionally, our findings suggest that seed chain sampling has the potential to address the longstanding unbiased caustics rendering problem in a practical setting.

Our work is the first study focusing on importance sampling strategies for multi-bounce specular light transport. We hope this work could promote new research interests in the Monte Carlo simulation of challenging light transport.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their valuable suggestions.

The glass model in the scene GLASS has been created by the cgtrader user DeepDreamDimension, and the texture in the scene LAMP is from the cgtrader user mehran1989. The scene FLOWER has been created by the cgtrader user Sirriarikan. The scene SPHERE, SLAB, RING, DOUBLE SLABS, POOL, PLANE, RINGANDSLAB and PILLAR are modified from the test scenes of [Zeltner et al. 2020]. We also thank the following people for providing the models or scenes that appear in our figure: Wig42 (LIVING), UP3D (LAMP), Delatronic (STONE), and aXeI (WATER). These scenes are acquired from [Bitterli 2016] with modifications. Besides, our test scenes use textures from CC0 Textures and cgbookcase. We also use environment maps courtesy of HDRI Haven and Paul Debevec.

This work was supported by the National Natural Science Foundation of China (No. 61972194 and No. 62032011) and the Natural Science Foundation of Jiangsu Province (No. BK20211147).

REFERENCES

- Benedikt Bitterli. 2016. Rendering resources. <https://benedikt-bitterli.me/resources/>.
- Per H. Christensen and Wojciech Jarosz. 2016. The Path to Path-Traced Movies. *Found. Trends. Comput. Graph. Vis.* 10, 2 (oct 2016), 103–175.
- Ana Dodik, Marios Pappas, Cengiz Öztireli, and Thomas Müller. 2022. Path Guiding Using Spatio-Directional Mixture Models. *Computer Graphics Forum* 41, 1 (2022), 172–189. <https://doi.org/10.1111/cgf.14428>
- Luca Fascione, Johannes Hanika, Rob Pieké, Ryusuke Villemin, Christophe Hery, Manuel Gamito, Luke Emrose, and André Mazzone. 2018. Path Tracing in Production. In *ACM SIGGRAPH 2018 Courses* (Vancouver, British Columbia, Canada) (SIGGRAPH '18). Association for Computing Machinery, Article 15, 79 pages.
- Rory A. Fisher. 1953. Dispersion on a sphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 217 (1953), 295 – 305.
- Toshiya Hachisuka and Henrik Wann Jensen. 2009. Stochastic Progressive Photon Mapping. *ACM Trans. Graph.* 28, 4, Article 141 (2009), 8 pages. <https://doi.org/10.1145/1661412.1618487>
- Johannes Hanika, Marc Droske, and Luca Fascione. 2015. Manifold Next Event Estimation. *Computer Graphics Forum* 34, 4 (2015), 87–97. <https://doi.org/10.1111/cgf.12681>
- Paul S. Heckbert. 1990. Adaptive Radiosity Textures for Bidirectional Ray Tracing. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques* (Dallas, TX, USA) (SIGGRAPH '90). Association for Computing Machinery, New York, NY, USA, 145–154. <https://doi.org/10.1145/97879.97895>
- Sebastian Herholz, Oskar Elek, Jiří Vorba, Hendrik Lensch, and Jaroslav Krivánek. 2016. Product Importance Sampling for Light Transport Path Guiding. *Computer Graphics Forum* 35, 4 (2016), 67–77. <https://doi.org/10.1111/cgf.12950>
- Heinrich Hey and Werner Purgathofer. 2002. Importance Sampling with Hemispherical Particle Footprints. In *Proceedings of the 18th Spring Conference on Computer Graphics* (Budmerice, Slovakia) (SCCG '02). Association for Computing Machinery, New York, NY, USA, 107–114. <https://doi.org/10.1145/584458.584476>
- Wenzel Jakob and Steve Marschner. 2012. Manifold Exploration: A Markov Chain Monte Carlo Technique for Rendering Scenes with Difficult Specular Transport. *ACM Trans. Graph.* 31, 4, Article 58 (jul 2012), 13 pages. <https://doi.org/10.1145/2185520.2185554>
- Wenzel Alban Jakob. 2013. *Light transport on path-space manifolds*. Ph.D. Dissertation. Cornell University, New York.

- Johannes Jendersie and Thorsten Grosch. 2019. Microfacet Model Regularization for Robust Light Transport. *Computer Graphics Forum* 38, 4 (2019), 39–47. <https://doi.org/10.1111/cgf.13768>
- Henrik Wann Jensen. 1995. Importance Driven Path Tracing using the Photon Map. In *Rendering Techniques '95*, Patrick M. Hanrahan and Werner Purgathofer (Eds.). Springer Vienna, Vienna, 326–335.
- Henrik Wann Jensen and Niels Jørgen Christensen. 1995. Photon maps in bidirectional Monte Carlo ray tracing of complex objects. *Comput. Graph.* 19 (1995), 215–224.
- Anton S. Kaplanyan and Carsten Dachsbacher. 2013. Path Space Regularization for Holistic and Robust Light Transport. *Computer Graphics Forum* 32, 2pt1 (2013), 63–72. <https://doi.org/10.1111/cgf.12026>
- Anton S. Kaplanyan, Johannes Hanika, and Carsten Dachsbacher. 2014. The Natural-Constraint Representation of the Path Space for Efficient Light Transport Simulation. *ACM Trans. Graph.* 33, 4, Article 102 (jul 2014), 13 pages. <https://doi.org/10.1145/2601097.2601108>
- A. Keller, L. Fascione, M. Fajardo, I. Georgiev, P. Christensen, J. Hanika, C. Eisenacher, and G. Nichols. 2015. The Path Tracing Revolution in the Movie Industry. In *SIGGRAPH 2015 Courses* (Los Angeles, California) (*SIGGRAPH '15*). Association for Computing Machinery, New York, NY, USA, Article 24, 7 pages.
- Eric P. LaFortune and Yves D. Willems. 1993. Bi-directional path tracing. *Proceedings of Third International Conference on Computational Graphics and Visualization Techniques*.
- Eric P. LaFortune and Yves D. Willems. 1995. A 5D Tree to Reduce the Variance of Monte Carlo Ray Tracing. In *Rendering Techniques '95*, Patrick M. Hanrahan and Werner Purgathofer (Eds.). Springer Vienna, Vienna, 11–20.
- He Li, Beibei Wang, Changhe Tu, Kun Xu, Nicolas Holzschuch, and Ling-Qi Yan. 2022. Unbiased Caustics Rendering Guided by Representative Specular Paths. In *SIGGRAPH Asia 2022 Conference Papers* (Daegu, Republic of Korea) (*SA '22*). Association for Computing Machinery, New York, NY, USA, Article 41, 8 pages. <https://doi.org/10.1145/3550469.3555381>
- Guillaume Loubet, Tizian Zeltner, Nicolas Holzschuch, and Wenzel Jakob. 2020. Slope-Space Integrals for Specular next Event Estimation. *ACM Trans. Graph.* 39, 6, Article 239 (nov 2020), 13 pages. <https://doi.org/10.1145/3414685.3417811>
- Don P. Mitchell and Pat Hanrahan. 1992. Illumination from curved reflectors. *Proceedings of the 19th annual conference on Computer graphics and interactive techniques* (1992).
- Thomas Müller. 2019. “Practical Path Guiding” in Production. In *ACM SIGGRAPH Courses: Path Guiding in Production, Chapter 10* (Los Angeles, California). ACM, New York, NY, USA, 18:35–18:48. <https://doi.org/10.1145/3305366.3328091>
- Thomas Müller, Markus Gross, and Jan Novák. 2017. Practical Path Guiding for Efficient Light-Transport Simulation. *Computer Graphics Forum* 36 (07 2017), 91–100. <https://doi.org/10.1111/cgf.13227>
- Merlin Nimier-David, Delio Vicini, Tizian Zeltner, and Wenzel Jakob. 2019. Mitsuba 2: A Retargetable Forward and Inverse Renderer. *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 38, 6 (Dec. 2019). <https://doi.org/10.1145/3355089.3356498>
- Hisanari Otsu, Johannes Hanika, Toshiya Hachisuka, and Carsten Dachsbacher. 2018. Geometry-Aware Metropolis Light Transport. *ACM Trans. Graph.* 37, 6, Article 278 (dec 2018), 11 pages. <https://doi.org/10.1145/3272127.3275106>
- Alexander Rath, Pascal Grittmann, Sebastian Herholz, Petr Vévoda, Philipp Slusallek, and Jaroslav Krivánek. 2020. Variance-Aware Path Guiding. *ACM Trans. Graph.* 39, 4, Article 151 (aug 2020), 12 pages. <https://doi.org/10.1145/3386569.3392441>
- Florian Reibold, Johannes Hanika, Alisa Jung, and Carsten Dachsbacher. 2018. Selective Guided Sampling with Complete Light Transport Paths. *ACM Trans. Graph.* 37, 6, Article 223 (dec 2018), 14 pages. <https://doi.org/10.1145/3272127.3275030>
- Lukas Ruppert, Sebastian Herholz, and Hendrik P. A. Lensch. 2020. Robust Fitting of Parallax-Aware Mixtures for Path Guiding. *ACM Trans. Graph.* 39, 4, Article 147 (aug 2020), 15 pages. <https://doi.org/10.1145/3386569.3392421>
- Vincent Schöblier, Johannes Hanika, Alisa Jung, and Carsten Dachsbacher. 2022. Path Guiding with Vertex Triplet Distributions. *Computer Graphics Forum* 41, 4 (2022), 1–15. <https://doi.org/10.1111/cgf.14582>
- Martin Šik and Jaroslav Krivánek. 2020. Survey of Markov Chain Monte Carlo Methods in Light Transport Simulation. *IEEE Transactions on Visualization and Computer Graphics* 26 (2020), 1821–1840.
- Eric Veach. 1998. *Robust Monte Carlo Methods for Light Transport Simulation*. Ph. D. Dissertation. Stanford University, Stanford, CA, USA. Advisor(s) Guibas, Leonidas J. AAI9837162.
- Eric Veach and Leonidas Guibas. 1995a. Bidirectional Estimators for Light Transport. In *Photorealistic Rendering Techniques*, Georgios Sakas, Stefan Müller, and Peter Shirley (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 145–167.
- Eric Veach and Leonidas J. Guibas. 1995b. Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '95)*. Association for Computing Machinery, New York, NY, USA, 419–428. <https://doi.org/10.1145/218380.218498>
- Eric Veach and Leonidas J. Guibas. 1997. Metropolis Light Transport. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., USA, 65–76. <https://doi.org/10.1145/258734.258775>
- Jiří Vorba, Johannes Hanika, Sebastian Herholz, Thomas Müller, Jaroslav Krivánek, and Alexander Keller. 2019. Path Guiding in Production. , Article 18 (2019), 77 pages. <https://doi.org/10.1145/3305366.3328091>
- Jiří Vorba, Ondřej Karlík, Martin Šik, Tobias Ritschel, and Jaroslav Krivánek. 2014. On-Line Learning of Parametric Mixture Models for Light Transport Simulation. *ACM Trans. Graph.* 33, 4, Article 101 (jul 2014), 11 pages. <https://doi.org/10.1145/2601097.2601203>
- Martin Šik, Hisanari Otsu, Toshiya Hachisuka, and Jaroslav Krivánek. 2016. Robust Light Transport Simulation via Metropolised Bidirectional Estimators. *ACM Trans. Graph.* 35, 6, Article 245 (dec 2016), 12 pages. <https://doi.org/10.1145/2980179.2982411>
- Bruce Walter, Shuang Zhao, Nicolas Holzschuch, and Kavita Bala. 2009. Single Scattering in Refractive Media with Triangle Mesh Boundaries. *ACM Trans. Graph.* 28, 3, Article 92 (jul 2009), 8 pages. <https://doi.org/10.1145/1531326.1531398>
- Beibei Wang, Miloš Hašan, and Ling-Qi Yan. 2020. Path Cuts: Efficient Rendering of Pure Specular Light Transport. *ACM Trans. Graph.* 39, 6, Article 238 (nov 2020), 12 pages. <https://doi.org/10.1145/3414685.3417792>
- Gregory J. Ward, Francis M. Rubinstein, and Robert D. Clear. 1988. A ray tracing solution for diffuse interreflection. *Proceedings of the 15th annual conference on Computer graphics and interactive techniques* (1988).
- Philippe Weier, Marc Droske, Johannes Hanika, Andrea Weidlich, and Jiří Vorba. 2021. Optimised Path Space Regularisation. *Computer Graphics Forum* 40, 4 (2021), 139–151. <https://doi.org/10.1111/cgf.14347>
- Tizian Zeltner, Iliyan Georgiev, and Wenzel Jakob. 2020. Specular Manifold Sampling for Rendering High-Frequency Caustics and Glints. *ACM Trans. Graph.* 39, 4, Article 149 (jul 2020), 15 pages. <https://doi.org/10.1145/3386569.3392408>
- Shilin Zhu, Zexiang Xu, Tiancheng Sun, Alexandr Kuznetsov, Mark Meyer, Henrik Wann Jensen, Hao Su, and Ravi Ramamoorthi. 2021a. Hierarchical Neural Reconstruction for Path Guiding Using Hybrid Path and Photon Samples. *ACM Trans. Graph.* 40, 4, Article 35 (jul 2021), 16 pages. <https://doi.org/10.1145/3450626.3459810>
- Shilin Zhu, Zexiang Xu, Tiancheng Sun, Alexandr Kuznetsov, Mark Meyer, Henrik Wann Jensen, Hao Su, and Ravi Ramamoorthi. 2021b. Photon-Driven Neural Reconstruction for Path Guiding. *ACM Trans. Graph.* 41, 1, Article 7 (nov 2021), 15 pages. <https://doi.org/10.1145/3476828>