

Proxy Tracing: Unbiased Reciprocal Estimation for Optimized Sampling in BDPT

FUJIA SU*, Peking University, China

BINGXUAN LI*, Peking University, China

QINGYANG YIN, Peking University, China

YANCHEN ZHANG, Peking University, China

SHENG LI[†], School of Computer Science, National Biomedical Imaging Center, Peking University, China

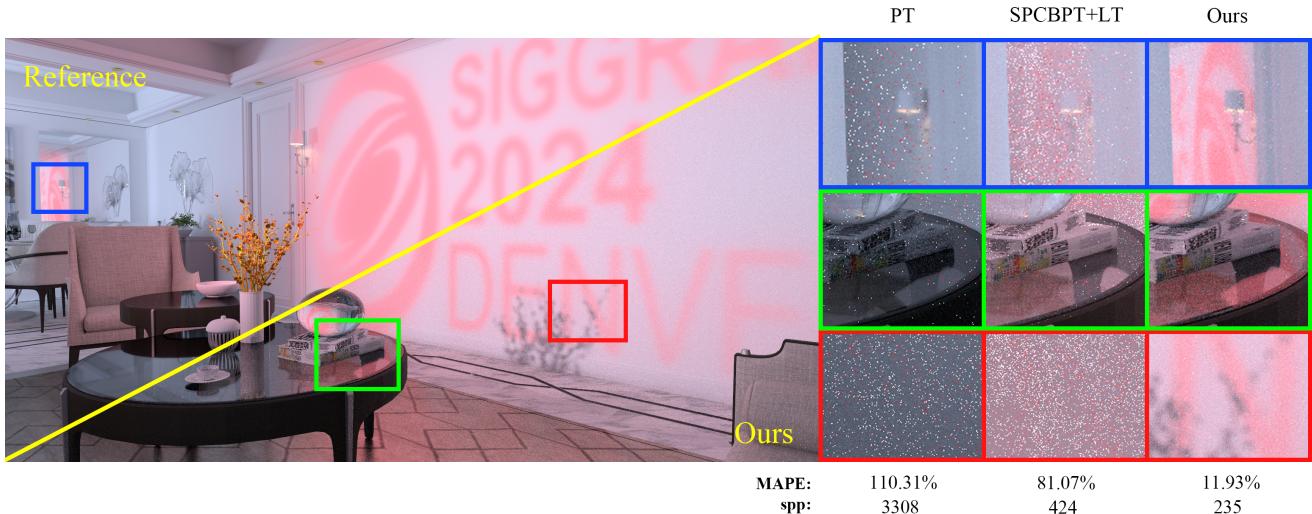


Fig. 1. Equal-time (60s) comparison between unidirectional Path Tracing (PT) combined with next event estimation (NEE), Subspace-based Probabilistic Connections Bidirectional Path Tracing (SPCBPT) [Su et al. 2022] with Light Tracing (LT) enabled, and our approach. Mean absolute percentage error (MAPE) is used as the metric for comparison. In this *Projector* scene, a convex lens refracts the textured light from the projector and then projects it on the wall. Our method can efficiently sample the specular-involved difficult paths like L_{DDSSDE} and $L_{DDSSDSE}$, leading to indistinguishable results that closely match the reference. Our approach outperforms the other approaches significantly.

Robust light transport algorithms, particularly bidirectional path tracing (BDPT), face significant challenges when dealing with specular or highly glossy involved paths. BDPT constructs the full path by connecting sub-paths traced individually from the light source and camera. However, it remains difficult to sample by connecting vertices on specular and glossy surfaces with narrow-lobed BSDF, as it poses severe constraints on sampling

*Fujia Su and Bingxuan Li contributed to this work equally.

[†]Sheng Li is the corresponding author.

Authors' addresses: Fujia Su, sufujia@pku.edu.cn, Peking University, China; Bingxuan Li, lbx@stu.pku.edu.cn, Peking University, China; QingYang Yin, 2000013161@stu.pku.edu.cn, Peking University, China; Yanchen Zhang, xy_zml@foxmail.com, Peking University, China; Sheng Li, lisheng@pku.edu.cn, School of Computer Science, National Biomedical Imaging Center, Peking University, China.

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.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2024/7-ART97 \$15.00 https://doi.org/10.1145/3658216

in the feasible direction. To address this issue, we propose a novel approach, called *proxy sampling*, that enables efficient sub-path connection of these challenging paths. When a low-contribution specular/glossy connection occurs, we drop out the problematic neighboring vertex next to this specular/glossy vertex from the original path, then retrace an alternative sub-path as a proxy to complement this incomplete path. This newly constructed complete path ensures that the connection adheres to the constraint of the narrow lobe within the BSDF of the specular/glossy surface. Unbiased reciprocal estimation is the key to our method to obtain a probability density function (PDF) reciprocal to ensure unbiased rendering. We derive the reciprocal estimation method and provide an efficiency-optimized setting for efficient sampling and connection. Our method provides a robust tool for substituting problematic paths with favorable alternatives while ensuring unbiasedness. We validate this approach in the probabilistic connections BDPT for addressing specular-involved difficult paths. Experimental results have proved the effectiveness and efficiency of our approach, showcasing high-performance rendering capabilities across diverse settings.

CCS Concepts: • Computing methodologies → Rendering; Ray tracing.

Additional Key Words and Phrases: Bidirectional path tracing, reciprocal estimation, importance sampling, proxy sampling, difficult path

ACM Reference Format:

Fujia Su, Bingxuan Li, QingYang Yin, Yanchen Zhang, and Sheng Li. 2024. Proxy Tracing: Unbiased Reciprocal Estimation for Optimized Sampling in BDPT. *ACM Trans. Graph.* 43, 4, Article 97 (July 2024), 21 pages. <https://doi.org/10.1145/3658216>

1 INTRODUCTION

Bidirectional path tracing (BDPT) is a widely used rendering technique known for its robust performance in synthesizing realistic images. BDPT employs a variety of sampling strategies to minimize variance, with the majority relying on connection. In a connection sampling strategy, an eye sub-path is traced from the camera and a light sub-path is traced from the light source, respectively. These two sub-paths are then connected to construct a full path, and this works well when the end vertices of both sub-paths lie on diffuse surfaces. However, when one of the end vertices lies on a specular or highly glossy surface, the connected full path tends to be inefficient, contributing negligibly to the final image. This is due to the strict constraints required by the Bidirectional Scattering Distribution Function (BSDF) of specular materials with narrow lobe characteristics, which are difficult to satisfy when sub-paths are sampled independently for connections. This issue is exemplified in Figure 2 (a), demonstrating the challenges in connection sampling with specular surfaces.

To improve the efficiency of connection sampling in BDPT, several algorithms based on probabilistic connections have been developed [Nabata et al. 2020; Popov et al. 2015; Su et al. 2022]. These algorithms specifically focus on the efficient resampling of light sub-paths given eye sub-path. However, the challenge of handling high-frequency specular constraints remains. The sampling precision required in such cases is beyond the capabilities of these algorithms. Another effective technique to reduce variance is Multiple importance sampling (MIS) [Veach and Guibas 1995b], which combines multiple sampling strategies to reduce variance. Yet, even with MIS, managing complex paths like $L_{DD}SDSE^1$ remains a formidable task. In this context, we use a hyphen (-) to indicate connections between the eye sub-path and the light sub-path. Strategies that involve specular connections, such as $L_{DD} - SDSE$, $L_{DD}S - DSE$, $L_{DD}SD - SE$, and $L_{DD}SDS - E$, exhibit poor performance. The only strategy left might involve tracing the whole $L_{DD}SDSE$ path unidirectionally from the camera. This is usually difficult, particularly when the specular surface S is far from the diffuse surface D . As a result, none of the existing sampling strategies in BDPT can sample this path efficiently, leading to a significant increase in variance.

Our objective is to tackle challenging paths in BDPT that involve specular or glossy materials with high-frequency BSDF. Our key idea is to substitute the less contributive vertices in a light sub-path with newly calculated vertices. This substitution is aimed at creating a more contributive full path, as depicted in Figure 2. To accomplish this, we discard the vertex adjacent to the last specular vertex (i.e.,

¹We adopt Veach's notation [Veach 1998] to represent paths and light source. Here, L_{DD} , S , D , E denote a vertex on the a diffusely emitting light source with finite area, specular surface, diffuse surface, and camera, respectively. Different from Veach's original notation, a simple E is used for the camera because the camera's type is irrelevant in our method. Also, when a surface is highly glossy (metallic/transparent material with roughness smaller than 0.2 in Disney principled BSDF), we mark the vertex on it as S .

the penultimate vertex) in light sub-path, and retrace a new vertex instead, termed the *proxy vertex*. This new vertex is traced based on the direction of the eye sub-path and the BSDF of the specular surface at the light vertex's location. By incorporating the proxy vertex into the full path, we ensure adherence to the specular lobe constraints of the BSDF. This adaptation provides a highly effective strategy for sampling these complex paths.

However, when vertices are discarded from a sub-path, it results in an incomplete sub-path which, theoretically, could be originated from an infinite number of initial sub-paths. This presents a significant challenge in unbiasedly estimating its Probability Density Function (PDF), as it requires the integration of all possible original sub-paths.

Stochastic Multiple Importance Sampling (SMIS) [West et al. 2020] or Marginal Multiple Importance Sampling (MMIS) [West et al. 2022] offer a solution by treating the various original sub-paths as different techniques within a continuous technique space and then applying multiple importance sampling. Nevertheless, SMIS/MMIS face considerable challenges when dealing with the complex sampling distributions that are central to our proxy sampling. Therefore, these approaches are ill-suited for tackling the unique requirements of our methodology.

Instead, we explore an efficient solution based on reciprocal estimation, a method already employed in photon mapping [Qin et al. 2015], volumetric rendering [Novák et al. 2018], specular manifold sampling [Zeltner et al. 2020], and differential rendering [Bangaru et al. 2020]. Existing reciprocal estimators are typically constrained by high computational costs and variance. To overcome these limitations, we introduce a novel reciprocal estimator. This new estimator is designed for optimized efficiency and is distinguished by a theoretically derived upper bound on its variance, enhancing its applicability in complex rendering scenarios.

We highlight that our unbiased reciprocal estimator enables the seamless institution of problematic paths with more favorable ones. This replacement is accomplished without compromising the unbiased nature of the Monte Carlo estimation in the path integral framework, and importantly, it is implemented without incurring excessive costs. Thus, our approach enables efficient sampling of the challenging paths using the connection strategy in BDPT.

Finally, our estimator can be integrated with the probabilistic connections BDPT algorithm [Popov et al. 2015] to further improve efficiency. We have validated our proxy tracing approach across a range of scene settings through comprehensive experimental comparisons. These comparisons were conducted on both the latest BDPT approaches and sampling approaches, demonstrating the superiority of our approach with high efficiency in handling difficult light transport scenarios.

Overall, our main contributions can be summarized in two key aspects:

- We present a novel path sampling technique using path substitution to greatly enhance BDPT's performance in handling difficult paths.
- We introduce a novel reciprocal estimator. This novel estimator is more efficient and practically applicable, making it a feasible methodology for our path sampling technique.

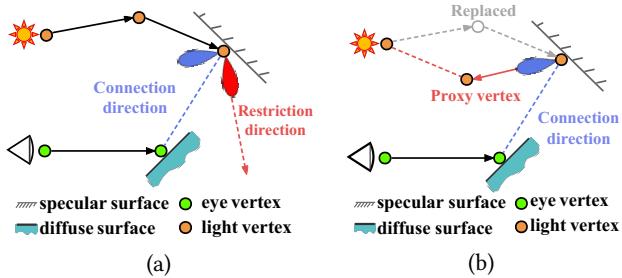


Fig. 2. Motivation of Proxy tracing. (a) When the end vertex of a light sub-path lies on a glossy surface, the probability of connecting it to the eye sub-path within the narrow BSDF lobe is very low. (b) To address this, our method modifies the light sub-path by introducing a proxy vertex instead of the original one, which satisfies the narrow lobe constraint on the glossy surface, thereby enabling us to handle difficult specular-involved paths by improving the connection probabilities.

2 RELATED WORK

Our proposed method aims to improve the efficiency of connection sampling strategies in bidirectional path tracing (BDPT), with a focus on difficult paths involving specular materials. While many existing algorithms, including probabilistic connection methods [Nabata et al. 2020; Popov et al. 2015; Su et al. 2022], have been developed for BDPT, our approach can be easily combined with these methods and is specifically designed for efficient sampling of difficult paths.

Caustics, traditionally considered a challenging aspect of specular-involved paths, are commonly rendered using the photon mapping algorithm [Jensen and Christensen 1995]. Conventional photon mapping methods are inherently biased. Progressive photon mapping iteratively evaluates the photon density to produce a biased but consistent rendering result by reducing the kernel radius for radiance estimation [Hachisuka and Jensen 2009; Hachisuka et al. 2008], or by finding a kernel radius that deems the radiance to be unbiased [Lin et al. 2023, 2020]. Georgiev et al. [2012]; Hachisuka et al. [2012] propose the VCM/UPS framework to combine BDPT with the photon mapping method. More recently, path guiding strategy has been employed to enable reliable rendering of caustics for path tracing [Li et al. 2022].

The manifold method [Hanika et al. 2015; Zeltner et al. 2020] also focuses on specular path handling. By iteratively solving the appropriate position for the specular vertex in a given initial path, thereby producing a path satisfying the specular constraint. This method incorporates reciprocal estimation, as different initial paths may transform into the same final path. Differently, our method fixes the vertex on the specular surface, while the manifold method fixes the neighboring vertices next to the specular vertex. Both the manifold method and reciprocal estimation require multiple iterations and tracing to get an unbiased result. In our method, we retrace only once per connection and significantly reduce the overhead of reciprocal estimation through light sub-path reuse. This efficiency allows our method to integrate seamlessly with existing BDPT methodologies. In addition, Tokuyoshi and Harada [2018,

2019] proposed the Hierarchical Russian roulette (HRR) and used BVH for connections within the BRDF range, focusing on the SDS path but requiring extensive light sub-path tracing. Our method, conversely, needs fewer light sub-paths and applies to a broader range of path types.

Path guiding algorithms [Jensen 1996; Lafortune and Willem 1999; Rath et al. 2020] can improve the performance of unidirectional path tracing to better handle the difficult paths during light transport. Path guiding uses the incident radiance distribution to guide path tracing and builds models such as spatio-directional trees [Müller et al. 2017] or Gaussian mixture model [Vorba et al. 2014] to represent the incident radiance field. Unidirectional path tracing is efficient when the diffuse surface is close to the glossy surface but less so when the surfaces are far apart. Our algorithm, an enhancement of the connection algorithm, performs well, especially when the diffuse surface is distant from the glossy surface. Therefore, our method can be easily incorporated into path-guiding algorithms to provide efficient and robust sampling for difficult paths.

Reciprocal estimation is essential for our method to compute the PDF of proxy tracing. The method to estimate reciprocal unbiasedly was first proposed by Booth [2007] and was introduced to graphics by Qin et al. [2015] to implement unbiased photon gathering. Zeltner et al. [2020] also utilized this method in specular manifold sampling. Blanchet et al. [2015] provided the approach to estimate $g(\int x dx)$ unbiasedly by Taylor expansion when function g is nonlinear and smooth. Reciprocal estimation is a special case of the nonlinear smooth function g . Zackary et al. [2022] applied Taylor expansion to evaluate the bias of the biased estimator and transform this estimator into an unbiased one. We adopt the reciprocal estimation framework of the previous method and propose an optimized setting for the reciprocal estimation. Russian roulette and splitting (RRS) method is involved in the reciprocal estimation. Rath et al. [2022a] propose the efficiency-optimized setting for the RRS process in path tracing, which inspires us to explore the optimized RRS function in reciprocal estimation.

Traditional MIS methods can only combine the contribution of a certain number of sampling techniques. Continuous MIS (CMIS) [West et al. 2020] extends MIS to an uncountable infinite number of sampling strategies. CMIS utilizes an optimally balanced heuristic to integrate a continuous spectrum of techniques, with its practical counterpart, Stochastic MIS (SMIS), being applicable in most scenes. Marginal MIS (MMIS) [West et al. 2022] further extends this concept to multiple technique spaces, allowing for integration with classical multi-sample MIS estimators. Both SMIS and MMIS can be used to solve our problem, serving as an alternative for reciprocal estimation.

Probabilistic connections methods [Nabata et al. 2020; Popov et al. 2015; Su et al. 2022] were developed for resampling the light sub-path based on the information of the eye sub-path. Although these resampling strategies of the light sub-path may not be sufficient to satisfy specular constraint, they can be combined with our method to cover the sampling of various difficult paths. By connecting sub-paths and generating valid specular paths, we can overcome difficulties associated with sampling specular materials, as illustrated in Figure 2(b). We also adopt the idea of light sub-path

reuse in the probabilistic connections method to reduce the overhead of reciprocal estimation; as well as the idea of selecting the appropriate light sub-path for connection to ensure the best possible efficiency and rendering quality.

3 PRELIMINARIES

3.1 Path Integral and Bidirectional Path Tracing

From the path integral formulation [Veach and Guibas 1995a], the pixel measurement I is as:

$$I = \int_{\Omega} f(\bar{x}) d\mu(\bar{x}), \quad (1)$$

where Ω is the path space, and $\bar{x} = x_0 x_1 \dots x_k$ is a path of length $k \geq 1$, x_0 and x_k are on a light source and the camera, respectively. $d\mu(\bar{x}) = dA(x_0) \dots dA(x_k)$ is the differential area product, and f is the measurement contribution function as:

$$f(\bar{x}) = L_e(x_0, x_1) T(\bar{x}) W_e(x_{k-1}, x_k),$$

$$T(\bar{x}) = GV(x_0, x_1) \left[\prod_{i=1}^{k-1} \rho(x_{i-1}, x_i, x_{i+1}) GV(x_i, x_{i+1}) \right],$$

where $L_e(x_0, x_1)$ is the radiance emitted $x_0 \rightarrow x_1$, W_e is the pixel sensitivity, ρ is the bidirectional scattering distribution function (BSDF), and GV is the geometry term including the visibility.

BDPT estimates Equation 1 by Monte Carlo integration. It traces the light sub-path \bar{y} from the light source in PDF $p(\bar{y})$ and eye sub-path \bar{z} from the camera in PDF $p(\bar{z})$, then connects these two sub-paths to sample the full path $\bar{x} = \bar{y}\bar{z}$. The PDF for sampling \bar{x} is $p(\bar{x}) = p(\bar{y})p(\bar{z})$. The contribution of the full path $f(\bar{x})$ can be divided into three parts $f_z(\bar{z})f_{yz}(\bar{y}, \bar{z})f_y(\bar{y})$ as [Popov et al. 2015]:

$$f_y(\bar{y}) = L_e(y_0, y_1) GV(y_0, y_1) \prod_{i=1}^{s-2} \rho(y_{i-1}, y_i, y_{i+1}) GV(y_i, y_{i+1}),$$

$$f_{yz}(\bar{y}, \bar{z}) = \rho(y_{s-2}, y_{s-1}, z_{t-1}) GV(y_{s-1}, z_{t-1}) \rho(y_{s-1}, z_{t-1}, z_{t-2}),$$

$$f_z(\bar{z}) = W_e(z_0, z_1) GV(z_0, z_1) \prod_{i=1}^{t-2} \rho(z_{i-1}, z_i, z_{i+1}) GV(z_i, z_{i+1}).$$

Here, s is the vertices number of the light sub-path, and t is the vertices number of the eye sub-path.

The PDF for a sub-path to be traced $p(\bar{y})$ can be written as the PDF product of its vertices as

$$p(\bar{y}) = p(y_0)p(y_1) \dots p(y_{s-1}) \quad (2)$$

and the PDF for a vertex $p(y_i)$ to be traced usually depends on its precedent vertex $p(y_{i-1})$ and the incident direction $y_{i-2} \rightarrow y_{i-1}$ to trace from BSDF at y_{i-1}

$$p(y_i) = p(y_{i-2} \rightarrow y_{i-1} \rightarrow y_i) = p(y_i | y_{i-1}, y_{i-2} \rightarrow y_{i-1}). \quad (3)$$

3.2 Probabilistic Connections

The probabilistic connections method, an enhancement of BDPT, boosts sub-path connection efficiency [Popov et al. 2015]. It begins each iteration by caching a set of light sub-paths. These are then resampled when connecting with the eye sub-paths, improving

efficiency by reusing the cached light sub-paths for all eye sub-paths and selectively providing higher sampling density to paths with greater contributions.

SPCBPT introduces the concept of subspace within path space to implement efficient probabilistic connections [Su et al. 2022]. Sub-paths are classified into subspaces based on position, normal, and incident direction, ensuring low discrepancy and shared sampling probabilities within each subspace. The method employs a two-stage process for light sub-path selection: first, selecting a light subspace based on the eye subspace, followed by resampling the light sub-path from this chosen subspace. In probabilistic algorithms like SPCBPT, $p(\bar{y})$ and $p(\bar{z})$ can easily approximate the shape of $f_y(\bar{y})$ and $f_z(\bar{z})$ by tracing from BSDF. In most cases, the probabilistic connections method can provide good sampling for f_{yz} . However, when y_{s-1} or z_{t-1} is on specular surface, the high variance of $f_{yz}(\bar{y}, \bar{z})$ will make the Monte Carlo estimator $\hat{I} = f/p = \frac{f_y(\bar{y})}{p(\bar{y})} \frac{f_z(\bar{z})}{p(\bar{z})} f_{yz}(\bar{y}, \bar{z})$ less efficient. Even the probabilistic connection method can hardly find the appropriate light sub-path that satisfies the high-frequency specular constraint.

Although multiple importance sampling (MIS) can leverage multiple sampling strategies in BDPT and select an appropriate strategy for the specular material [Kondapaneni et al. 2019; West et al. 2020], the fundamental challenges of path sampling and connection are still not fully addressed. Sampling paths involving specular materials, such as $L_{DD}(S|D)^* SDS^* E$ remains difficult. In this type of path, all the possible sampling strategies rely on the eye sub-path to trace from its first diffuse surface to the next specular surface and the PDF for the $D \rightarrow S$ bounce can be very small. Consequently, it is challenging for BDPT to sample such path effectively, which motivates the development of our proposed method.

Constructing a path with specular connections, like $L_{DD}(S|D)^* SDS^* E$, often fails to satisfy specular constraint, thus yielding negligible contributions to the final image. To address this issue, we propose retracing the neighbor of the specular vertex that can satisfy specular constraint, converting the invalid path into valuable ones with high contribution. In section 4, we will introduce our methodology of proxy sampling. This technique involves discarding the low-contribution segment of the path and retracing the complementary proxy path to enhance sampling performance.

3.3 Continuous, Stochastic and Marginal MIS

Multiple Importance Sampling (MIS) [Veach and Guibas 1995b] is a Monte Carlo integration framework for combining various sampling techniques, yet traditionally limited to countable sets. Continuous MIS (CMIS) [West et al. 2020] extends this to an uncountable technique space \mathcal{T} , broadening its applicability. The one-sample estimator in CMIS is

$$\langle I \rangle_{CMIS} = \frac{w(t, x)f(x)}{p(t)p(x|t)}, \quad (4)$$

where weighting function $w : \mathcal{T} \times \mathcal{X} \rightarrow \mathbb{R}$ with the property $\int_{\mathcal{T}} w(t, x) dt = 1$. For the estimator to be unbiased, the weighting

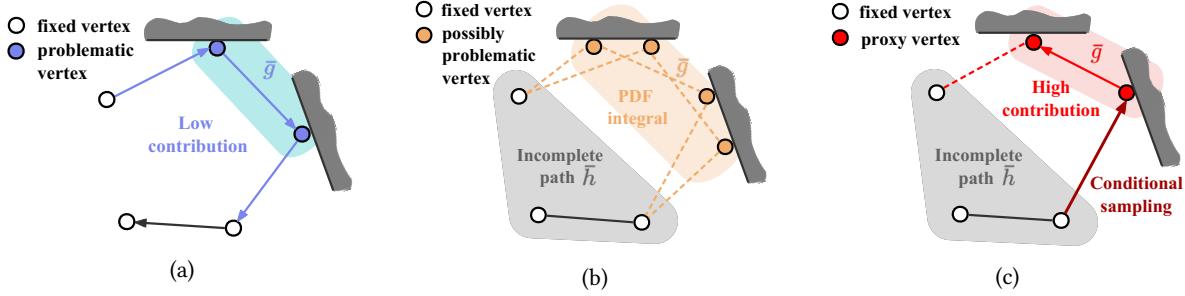


Fig. 3. Semantic illustration of proxy sampling. (a) An original path, sampled from the initial distribution, often includes poorly distributed vertices, reducing the overall contribution of the full path; (b) our approach discards the problematic vertices, resulting in an incomplete path. The PDF for an incomplete path requires integrating over all potential problematic paths that could lead to the identical incomplete path; (c) proxy vertices are then traced to merge with the incomplete sub-path and construct a new path. The data from the incomplete sub-path assists in tracing these proxy vertices, providing a more favorable conditional distribution for constructing a full path with a higher contribution.

function w must satisfy

$$\int_{\mathcal{T}} w(t, x) dt = 1 \text{ whenever } f(x) \neq 0, \quad (5a)$$

$$w(t, x) = 0 \text{ whenever } p(t, x) = 0. \quad (5b)$$

A variance-optimal weighting function, as derived in [West et al. 2020], is

$$\tilde{w}(t, x) = \frac{p(t)p(x|t)}{\int_{\mathcal{T}} p(t', x) dt'} = \frac{p(t, x)}{p(x)}, \quad (6)$$

apply this optimal weighting function in Equation 4, we get the balance-heuristic CMIS estimator

$$\langle I \rangle_{CMIS} = \frac{\tilde{w}(t, x)f(x)}{p(t, x)} = \frac{f(x)}{p(x)}. \quad (7)$$

Evaluating the balance-heuristic CMIS estimator requires evaluating the marginal PDF integral $p(x)$ in its denominator. Typically, $p(x)$ is not readily available in a closed form. Our solution is to create an unbiased estimator for the reciprocal, $1/p(x)$. [West et al. 2020] circumvented reciprocal estimation by proposing Stochastic MIS (SMIS), which can be seen as an approximation for balance-heuristic CMIS.

In SMIS, n independent sample pairs $(t_1, x_1), \dots, (t_n, x_n)$ are sampled. Then the SMIS estimator is constructed as

$$\langle I \rangle_{SMIS} = \sum_{i=1}^n \frac{\dot{w}(t_i, x_i)f(x_i)}{p(x_i|t_i)} = \sum_{i=1}^n \frac{f(x_i)}{\sum_{j=1}^n p(x_i|t_j)}, \quad (8)$$

where $\dot{w}(t_i, x) = \frac{p(x|t_i)}{\sum_{j=1}^n p(x|t_j)}$. SMIS effectively samples techniques from the space \mathcal{T} and then applying MIS. SMIS exhibits bias when there's a positive probability that the combined sampling distribution cannot cover $f(x)$'s support. Therefore, it requires nearly every technique in \mathcal{T} being unbiased, i.e. $p(x|t) > 0$ whenever $f(x) > 0$, if the techniques are chosen in a stochastic way. Further challenges arise when each technique only provide efficient sampling for small range of x . In that case, SMIS requires an large number of techniques and appropriate importance sampling in \mathcal{T} , both can be difficult to implement. The computational cost, scaling as $O(n^2)$, further limits the number of combinable techniques. We conclude these factors as the main limitations of SMIS.

Marginal MIS (MMIS) [West et al. 2022] generalized SMIS to multiple technique spaces. While MMIS enhances the robustness of SMIS by incorporating multiple technique spaces, it still inherits the limitations of SMIS. If none of the technique spaces within MMIS are capable of efficiently sampling $f(x)$, then the MMIS estimator can not achieve high efficiency, as it just combines these spaces.

These identified limitations of SMIS/MMIS underpin our decision to favor reciprocal estimation in our approach. We will provide further illustration in subsection 4.3, where we conduct a comparison between SMIS/MMIS and our reciprocal estimation.

4 OUR METHODOLOGY

When sampling a path \bar{x} from a distribution $p(\bar{x})$, certain vertices in \bar{x} may contribute to high variance due to poor distribution. To mitigate this, we employ a dropout-and-complement scheme, removing these *problematic* vertices and retracing new ones as substitutes. The retracing process is designed to be independent of the original problematic vertices while leveraging information from the fixed vertices in the original path. In this section, we formulate this dropout-and-complement scheme and show that its core principle in a reciprocal estimation of PDF for unbiasedness, and is effectively addressed by our proposed estimator eventually.

4.1 Proxy Sampling Formulation

Formally, we use \bar{g} to denote the candidate vertices that can be replaced in path \bar{x} . \bar{h} to denote the incomplete path that contains fixed vertices which will be used in the retracing. We use function $D(\bar{x}, \bar{g})$ to denote the operation of dropping out problematic vertices in \bar{x} from complete \bar{x} , obtaining an incomplete path \bar{h} :

$$\bar{h} = D(\bar{x}, \bar{g}). \quad (9)$$

We also define the repair operation $R(\bar{h}, \bar{g})$ to fill in this incomplete path \bar{h} with proxy vertices \bar{g} and get the complete path \bar{x} :

$$\bar{x} = R(\bar{h}, \bar{g}). \quad (10)$$

We illustrate the key aspects of proxy sampling in Figure 3. We first apply dropout operation $D(\bar{x}_o, \bar{g}_o)$ to obtain an incomplete path \bar{h} from original path \bar{x}_o (see Figure 3(a)). Next, we retrace proxy

vertices \bar{g} from distribution $p(\bar{g}|\bar{h})$ and generate the final path by applying the repair operation $\bar{x} = R(\bar{h}, \bar{g})$ (see Figure 3(c)). Hereby, The PDF for sampling the final path $p_d(\bar{x})$ is given by

$$p_d(\bar{x}) = p(\bar{h})p(\bar{g}|\bar{h}), \quad (11)$$

where $p(\bar{h})$ is used instead of $p(\bar{x}_0)$ to compute the PDF. As shown in Figure 3(b), an incomplete path \bar{h} can result from different original paths with different alternative vertices \bar{g} , so the PDF of incomplete path \bar{h} should be computed as

$$p(\bar{h}) = \int_{A^u} p[R(\bar{h}, \bar{g})]d\mu(\bar{g}), \quad (12)$$

where A is the surfaces of a scene and u is the vertex count (also path length) of \bar{g} .

Proxy sampling can provide better performance than the original distribution $p(\bar{x})$ when the conditional distribution $p(\bar{g}|\bar{h})$ is efficient. For instance, when a specular vertex and one of its neighbor vertices are located within \bar{h} , the conditional distribution for sampling the other neighbor vertex in \bar{g} is efficient by sampling from the specular BSDF.

However, computing the PDF $p(\bar{h})$ for sampling an incomplete path, as given in Equation 12, requires an integral that cannot be computed analytically in most cases. We can get an unbiased estimator $\tilde{p}(\bar{h})$ of $p(\bar{h})$ by applying the Monte Carlo method to Equation 12. Unfortunately, $\frac{1}{\tilde{p}(\bar{h})}$ is not an unbiased estimator for $\frac{1}{p(\bar{h})}$ [Chandler 1987]. This means the path integral estimator $f(\bar{x})/[\tilde{p}(\bar{h})p(\bar{g}|\bar{h})]$ is biased due to the presence of $\tilde{p}(\bar{h})$ in the denominator. Therefore, we need an unbiased reciprocal estimation for $\frac{1}{p(\bar{h})}$ in the proxy sampling. By deriving a reciprocal estimation formulation for the path integral, we can effectively estimate the PDF of the incomplete paths without introducing bias.

4.2 Our Reciprocal Estimation

In the reciprocal estimation, we adopt the Taylor expansion framework of Blanchet et al. [2015]. Without loss of generality, we use β to denote the integral of a function $\beta = \int_{\mathcal{R}} f(x)dx$ in an integral domain \mathcal{R} . We use a supporting distribution q to sample x from \mathcal{R} . Our objective is to get an unbiased estimation for the reciprocal of β as

$$\alpha = \frac{1}{\beta} = \frac{1}{\int_{\mathcal{R}} f(x)dx}.$$

We first compute the difference between β and a constant B as $A = B - \beta$, then we have

$$\alpha = \frac{1}{B - A}.$$

A Taylor expansion of $\frac{1}{B-A}$ at $A=0$ yields

$$\frac{1}{B-A} = \frac{1}{B} + \frac{A}{B^2} + \frac{A^2}{B^3} \dots = \frac{1}{B} \sum_{n=0} \frac{A^n}{B^n}. \quad (13)$$

The integral formulation of Equation 13 is

$$\alpha = \frac{1}{B} \sum_{n=0} I_n = \frac{1}{B} \sum_{n=0} \int_{\mathcal{R}^n} \prod_i \frac{B - f(x_i)}{B} d\mu(\bar{x}), \quad (14)$$

where $I_n = \frac{A^n}{B^n}$ and \bar{x} is constructed by N samples from domain \mathcal{R} .

The product of n random variables with $\frac{A}{B}$ -expectation is an unbiased estimation for $\frac{A^n}{B^n}$. Therefore, we can estimate I_n without bias by generating n independent samples from the supporting distribution q and building estimator \tilde{I}_n as

$$\tilde{I}_n = \prod_i^n \left[1 - \frac{f(x_i)}{Bq(x_i)} \right]. \quad (15)$$

Here, $E[1 - \frac{f(x_i)}{Bq(x_i)}] = \frac{A}{B}$, so \tilde{I}_n is an unbiased estimator for $I_n = \frac{A^n}{B^n}$.

We use $1 - \frac{f(x_i)}{Bq(x_i)}$ instead of $\frac{B-f(x_i)}{Bq(x_i)}$ to construct the $\frac{A}{B}$ -expectation random variable because the optimal supporting distribution q for the former is $q \propto f(x)$ while the latter is $q \propto 1 - f(x)$, which is typically more challenging and unconventional to construct in most scenarios.

We can make an unbiased estimation to α by combining infinite \tilde{I}_n as

$$\tilde{I} = \frac{1}{B} \sum_n \tilde{I}_n. \quad (16)$$

Blanchet et al. [2015] proposes two solutions to address the infinite summation in Equation 16. One approach constructs a single-term estimator: Sample an integer N from \mathbb{Z} with a probability mass function (PMF), p_N , generate N samples to estimate \tilde{I}_N , and use $\tilde{I}_N/[Bp_N]$ as the unbiased estimator for \tilde{I} . The other also samples N and generates N samples, but instead of estimating \tilde{I}_N , it uses the N samples to estimate any $\tilde{I}_{N'}$ for $N' \leq N$ by computing the prefix product. The resulting estimator for \tilde{I} is then given by $\sum_{i \leq N} \tilde{I}_i/[Bp(j \geq i)]$. Both solutions rely on appropriate distribution to sample integer N . Although Blanchet et al. [2015] introduced a cost-minimized distribution, applying this distribution to our proxy tracing method still presents challenges. This arises from several non-closed terms required by the distribution, such as the root of a complex non-linear equation, which are impractical to solve for our estimation. Consequently, we have chosen to deviate from their methodology and instead focus on addressing the reciprocal estimation problem with a simpler and more practical solution.

Since Equation 14 is close to the path integral formulation in Equation 1, it can be estimated in a Russian Roulette and Splitting (RRS) manner, similar to unidirectional path tracing. This involves constructing \bar{x} from the first sample x_0 and controlling the sampling using an RRS function $r(\bar{x})$. When a prefix \bar{x} is sampled and $R(\bar{x}) \leq 1$, we finish the sampling in probability $1 - r(\bar{x})$; otherwise, continue the sampling. If $r(\bar{x}) \geq 1$, we split $\lfloor r(\bar{x}) \rfloor + 1$ samplings on the prefix \bar{x} in probability $r(\bar{x}) - \lfloor r(\bar{x}) \rfloor$; otherwise, split $\lfloor r(\bar{x}) \rfloor$ samplings. The estimator for a path of k vertices \bar{x} is

$$\tilde{I}(\bar{x}) = \frac{\prod_{0 \leq i < k} \left[1 - \frac{f(x_i)}{Bq(x_i)} \right]}{\prod_{1 \leq j < k} r(\bar{x}_j)}, \quad (17)$$

where \bar{x}_j is the j -vertex prefix of \bar{x} . $\tilde{I} = [1 + \sum \tilde{I}(\bar{x})]/B$ provides unbiased estimation to α , where the extra $\frac{1}{B}$ comes from \tilde{I}_0 .

We propose an RRS function for a \bar{x} of k vertices as

$$r(\bar{x}) = |1 - \frac{f(x_{k-1})}{Bq(x_{k-1})}|. \quad (18)$$

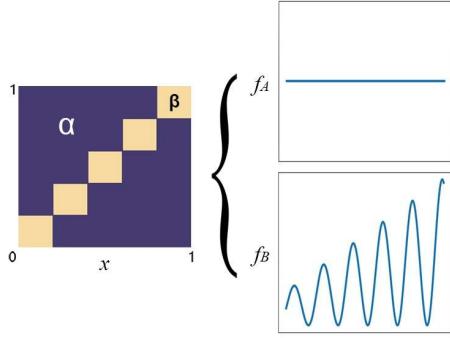


Fig. 4. Illustration of the sampling distribution $p(t, x)$ alongside the target function $f(x)$. Parameters α and β in the graph correspond to the PDF in this region. Different t associates with different sampling techniques $p(x|t)$. SMIS works by first sampling n techniques t_1, t_2, \dots, t_n , then performing MIS between $p(x|t_1), p(x|t_2), \dots, p(x|t_n)$.

We prove that the proposal RRS function is sub-optimal, i.e., optimal in some assumptions. We provide a detailed explanation of our estimator with analysis and proof in Appendix A. With our proposal RRS function, estimator $\tilde{I}(\bar{x})$ can be transformed into

$$\tilde{I}(\bar{x}) = \left[1 - \frac{f(x_{k-1})}{Bq(x_{k-1})} \right] S(\bar{x}), \quad (19)$$

where $S(\bar{x}) = \prod_{0 \leq j < k-1} \text{sgn}[1 - \frac{f(x_j)}{Bq(x_j)}]$ is used to denote the sign product of all the prefix vertices of \bar{x} .

In Appendix F, we also discuss the optimal setting of B . With our proposed RRS function, we construct a sign estimator $\tilde{I}_{\text{sign}}(\bar{x}) = S(\bar{x})$ which can also be used for reciprocal estimation but produces higher variance than $\tilde{I}(\bar{x})$. We provide the variance upper bound as well as the work-minimized B setting for the sign estimator. The recommended B setting is $B = \max(\frac{f(x)}{q(x)})$, i.e., the upper bound of $\frac{f(x)}{q(x)}$. We use the upper bound B setting for our estimator. Because $V(\tilde{I}) \leq V(\tilde{I}_{\text{sign}})$, we highlight that using this B can minimize the upper bound of reciprocal estimation.

In contrast to Blanchet et al. [2015], our RRS method does not directly sample N . Instead, it determines the number of samples by continuous application of RRS on current prefix samples. The probability of sampling \bar{x} is decided by both \bar{x} 's length and the prefix's importance (product of $r(x_i)$). Note that Booth [2007] uses a similar Russian roulette method in reciprocal estimation (splitting not involved). However, the RR function used is $r(\bar{x}) = \min(1, \frac{1}{r} \prod_{i < k} |1 - \frac{f(x_i)}{Bq(x_i)}|)$ where r is a hyperparameter that must be manually set. In contrast, our Russian roulette method uses a sub-optimal RRS function that is supported by rigorous variance analysis, without any manually-set parameters required. Additionally, both our method and Booth's method can be adapted to the geometric distribution when dealing with binary functions $f(x)$, as demonstrated by Qin et al. [2015].

Table 1. Results for $\alpha = 0.1$ and $\beta = 4.6$. We compared the variances of SMIS- n and our reciprocal estimator under conditions of equal cost. SMIS16 shows the best performance within this specific setting.

	SMIS2	SMIS4	SMIS8	SMIS16	OURS
f_A	22.0	15.2	7.5	2.1	4.3
f_B	203.5	142.4	71.6	23.8	51.7

Table 2. Results for $\alpha = 0.01$ and $\beta = 4.96$. Our reciprocal estimator exhibited superior efficiency. Note that SMIS- n 's variance increases dramatically compared to Table 1 while our estimator maintains nearly consistent variance across tests, showing our method's robustness in handling this type of distribution.

	SMIS2	SMIS4	SMIS8	SMIS16	OURS
f_A	240.2	173.6	91.0	21.6	4.5
f_B	2344.8	1658.4	874.9	197.1	54.0

4.3 Comparative Analysis of SMIS/MMIS and our Reciprocal Estimation

In this section, we will demonstrate the conditions under which our reciprocal estimation method outperforms SMIS/MMIS, and explain our preference for the former. Our analysis is structured in two parts: Initially, we will conduct unit experiments to compare the performance of each method directly. Subsequently, we will explore complex scenarios where SMIS/MMIS exhibit limitations, thereby emphasizing the strengths of our proxy sampling approach.

4.3.1 Unit Test. Our unit test is depicted in Figure 4, where we employ a sampling distribution $p(t, x)$ characterized by parameters α and β . These parameters are adjustable to estimate $f(x)$. It is important to note that only one parameter can be varied at a time, as they must satisfy the condition $0.8\alpha + 0.2\beta = 1$ to ensure the validity of the distribution.

Table 1 and Table 2 show the variance of SMIS- n and our reciprocal estimator under conditions of equal cost. In Table 1 where $\alpha = 0.1$, SMIS16 performs best. However, as α is reduced to 0.01 in Table 2, our reciprocal estimator exhibited superior efficiency. The variance of SMIS- n increases dramatically as $\alpha \rightarrow 0$. This phenomenon can be attributed to the characteristics of the distribution $p(t, x)$. As $\alpha \rightarrow 0$, $p(x|t)$ associate with different t targets different segments of $f(x)$. Notably, each individual sampling technique $p(t, x)$ is characterized by significant variance, and the combination of a limited subset fails to mitigate this issue. Only by combining all five distinct sampling methods (see Figure 4) can SMIS get comprehensive coverage of $f(x)$'s support.

When $\alpha = 0$, this situation falls outside SMIS's scope, and SMIS will produce bias. This bias arises because a combination of different sampling techniques does not consistently cover the entire support of $f(x)$. While increasing the number of techniques used in SMIS may reduce this bias, a certain level of positive bias remains inevitable as long as the number of techniques is finite.

Although MMIS was not included in the direct experimental comparison, it is crucial to recognize that MMIS shares the same

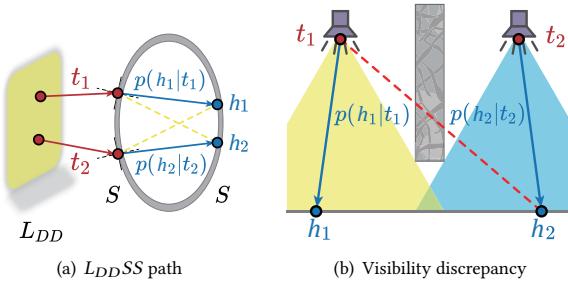


Fig. 5. Two failure cases in SMIS. (a) In handling this L_{DDSS} type light sub-path, our proxy sampling designates the $L_{DD}S$ as the dropout vertices, represented by \bar{g}_o , while the residual S forms the incomplete path \bar{h} . Within SMIS framework, \bar{g}_o is associated with technique t , and $p(\bar{h}|t)$ is proportional to the PDF that traces from \bar{g}_o 's last vertex to the first vertex (denote as h without the bar) following \bar{g}_o in \bar{h} . The probabilities $p(h_2|t_1)$ and $p(h_1|t_2)$, signified by yellow dashed lines, are nearly zero. (b) In scenario where visibility varies significantly, different techniques in SMIS primarily sample separate regions of $f(x)$. This leads to an inherent bias in SMIS because there is a positive probability that the combined sampling distribution cannot cover $f(x)$'s support.

limitations as SMIS. This similarity arises because MMIS operates by integrating multiple SMIS strategies.

4.3.2 SMIS/MMIS in Difficult Situations. Our proxy sampling is designed to handle specular-involved complex situations, while SMIS/MMIS will face challenges in such contexts. Recall Sec. 4.1, we first trace the original path \bar{x}_o from the original sampling method, and "dropout" vertices \bar{g}_o from the original path \bar{x}_o to get the incomplete path \bar{h} , and complement \bar{h} with newly traced proxy vertices \bar{g} . The "dropout" vertices \bar{g}_o is actually the technique in CMIS context, since it determines the sampling of \bar{h} . Also, A^u is the technique space. In the section, we use notations $t = \bar{g}_o$ and $\mathcal{T} = A^u$ to fit in with the CMIS context. The balance-heuristic CMIS estimator is then constructed as

$$\langle I \rangle_{CMIS} = \frac{f(\bar{x})}{p(\bar{x})} = \frac{f(\bar{x})}{p(\bar{h})p(\bar{x}|\bar{h})} = \frac{1}{\int_{\mathcal{T}} p(\bar{h}, t) dt} \cdot \frac{f(\bar{x})}{p(\bar{x}|\bar{h})}. \quad (20)$$

In SMIS, n independent pairs of techniques and corresponding samples, denoted as $(t_1, \bar{x}_1), \dots, (t_n, \bar{x}_n)$, are drawn. The SMIS estimator is constructed as

$$\langle I \rangle_{SMIS} = \sum_{i=1}^n \frac{1}{\sum_{j=1}^n p(\bar{h}_i|t_j)} \cdot \frac{f(\bar{x}_i)}{p(\bar{x}_i|\bar{h}_i)}. \quad (21)$$

Here, $p(\bar{x}_i|\bar{h}_i) = p(\bar{g}_i|\bar{h}_i)$ is the retracing PDF; $p(\bar{h}|t)$ refers to the PDF for all the vertices in \bar{h} to be traced in the original sampling method when the vertices in t given. Note that computing $p(\bar{h}_i|t_j)$ requires a visibility test between \bar{h}_i and t_j when $i \neq j$, as illustrated in Figure 5. This requirement results in a quadratic increase in the cost of visibility tests within SMIS. Consequently, utilizing smaller numbers of techniques is generally more practical in proxy tracing.

We present two scenarios illustrating the lesser effectiveness of SMIS. The first involves multiple specular vertices following a

light vertex, as depicted in Figure 5 (a). Constraints in reflective or refractive directions result in $p(h_2|t_1)$ and $p(h_1|t_2)$ being almost zero. This associates with Figure 4 when $\alpha \rightarrow 0$, where different techniques predominantly sample different regions of $f(x)$. In cases where the S vertex is highly glossy, SMIS faces the challenge of requiring numerous techniques to minimize variance. Conversely, if the S vertex is purely specular, it leads to bias in SMIS. This is observed in the *Projector* scene shown in our experiments.

Another case is when different lights do not share the same visibility, as shown in Figure 5 (b). In scenes with multiple light sources possessing complex visibilities, SMIS tends to exhibit obvious bias. This corresponds to SMIS's first limitation and will be further demonstrated in the scene *Hallway*. While MMIS can eliminate bias by combining SMIS with unbiased sampling techniques like unidirectional path tracing, the inherent high variance of SMIS makes the overall estimator inefficient.

The analysis above leads us to prefer reciprocal estimation over SMIS/MMIS for our method, with experimental findings detailed later. We also highlight that our reciprocal estimator can be integrated into the CMIS framework, serving as a potent SMIS alternative. The CMIS estimator (Equation 7) is effectively computed using reciprocal estimation, which shows superior performance in certain scenarios. This is exemplified in our application to photon plane sampling [Deng et al. 2019], a method formerly using SMIS [West et al. 2020]. Detailed results in Sec. 6.7 highlight our method's advantages.

5 PROXY TRACING FOR BDPT

With the proxy sampling method and reciprocal estimation discussed above, we therefore can improve the performance of probabilistic BDPT with specular connection by modifying the problematic vertices in the path sampled by specular connection. If the eye sub-path is on the specular surface, we can always continue the tracing of the eye sub-path and use the MIS method to get a robust sampling. Therefore, We focus on the case where the light sub-path is on the specular surface.

5.1 Proxy Construction

When a light sub-path \bar{y} terminates at a specular vertex y_{s-1} and needs to be connected with an eye sub-path \bar{z} , we keep the eye sub-path \bar{z} and the specular vertex y_{s-1} unchanged in the proxy sampling. Then, we retrace the light sub-path \bar{y} based on the location of z_{t-1} that can satisfy specular restriction.

Not all the vertices in the light sub-path \bar{y} need to be retraced. To ensure a valid BSDF at vertex y_{s-1} , we only retrace y_{s-2} by sampling the BSDF $B(z_{t-1} \rightarrow y_{s-1} \rightarrow y_{s-2})$. However, for a light sub-path that ends with multiple consecutive specular vertices, only retracing y_{s-2} is insufficient to satisfy the multiple specular constraints in the path. Therefore, if there are u consecutive specular vertices $y_{s-1}y_{s-2}\dots y_{s-u}$ at the end of a light sub-path, we will retrace u vertices $\bar{g} = y_{s-2}y_{s-3}\dots y_{s-u-1}$ to complete the proxy sampling.

The specular or diffuse attributes of each vertex in the sequence in a proxy path should be consistent with those of the original sequence in the original path that has been substituted. Otherwise, there could be multiple possible proxy paths \bar{g} for a given full path

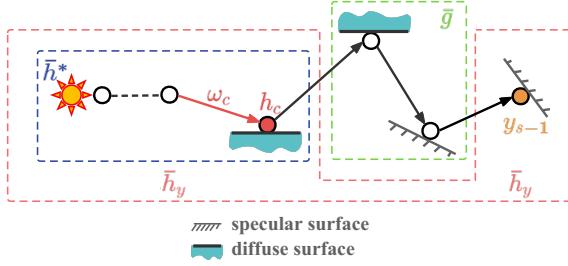


Fig. 6. A light sub-path can be segmented into several components: the path to be dropped out \bar{g} , the residual path \bar{h}^* , and the terminal specular vertex y_{s-1} located on a specular surface. The control vertex h_c represents the terminal of \bar{h}^* and the control direction ω_c is the incident direction at h_c . Reciprocal estimation is determined by these elements: the control vertex h_c , control direction ω_c , specular vertex y_{s-1} , and u , which is the vertex count of \bar{g} .

\bar{x} . For instance, an L_{DDDSDE} path $y_0y_1y_2z_1z_0$ can be sampled from either $\bar{g} = y_1$ or $\bar{g} = y_1y_2$ if the original L_{DDSSDE} path is retraced. Dealing with multiple alternative paths \bar{g} is difficult and impractical. So, we stop the retracing and discard the connection when the retraced alternate path \bar{g} has a specular/diffuse mismatch with the original. This ensures that the alternate path \bar{g} as the proxy for a connection can be uniquely determined.

5.2 PDF Evaluation

We discuss the PDF for the incomplete sub-path sampling $p(\bar{h})$. \bar{h} consists of the complete eye sub-path \bar{z} and the incomplete light sub-path \bar{h}_y . Eye sub-path tracing is irrelevant to the tracing of the light sub-path. Therefore $p(\bar{h})$ is the product of $p(\bar{h}_y)$ and $p(\bar{z})$ as:

$$p(\bar{h}) = p(\bar{h}_y)p(\bar{z}). \quad (22)$$

Since $p(\bar{z})$ can be accurately computed, we only need to make a reciprocal estimation of the incomplete light sub-path $1/p(\bar{h}_y)$. We further analyze the tracing of incomplete light sub-path \bar{h}_y . A light sub-path $\bar{g} = R(\bar{h}_y, \bar{g})$ can be divided into several parts, as shown in Figure 6. They are the alternate path \bar{g} , the precedent path before reaching the alternate path $\bar{h}^* = y_0y_1\dots y_{s-u-2}$, and the last specular vertex y_{s-1} . According to Equation 3, the probability density of a vertex y_i being traced is determined by its precedent two vertices. Therefore, the probability density of tracing the light sub-path \bar{g} can be

$$p(\bar{g}) = p[R(\bar{h}_y, \bar{g})] = p(\bar{h}^*)p(y_{s-1}|\bar{g}, h_c)p(\bar{g}|h_c, \omega_c), \quad (23)$$

where $h_c = y_{s-u-2}$ is the last vertex of \bar{h}^* , $\omega_c = y_{s-u-3} \rightarrow y_{s-u-2}$ is the incident direction of h_c . Under extreme conditions, h_c is set to null when $y_0 \in \bar{g}$, ω_c is set to null when $y_0 \in \bar{g}$ or $y_1 \in \bar{g}$. h_c affects the sampling of y_{s-1} only when the vertex count of \bar{g} equals one. There is no vertex before \bar{h}^* , so the tracing of \bar{h}^* is irrelevant to the alternate path \bar{g} . Only y_{s-u-2} and y_{s-u-3} can make sense to the tracing of \bar{g} and y_{s-1} . Therefore, we call h_c and ω_c as the control vertex and control direction.

Then, according to Equation 12, the PDF for the incomplete light sub-path \bar{h}_y is as

$$\begin{aligned} p(\bar{h}_y) &= \int_{A^u} p[R(\bar{h}_y, \bar{g})]d\mu(\bar{g}) \\ &= p(\bar{h}^*) \int_{A^u} p(\bar{g}|h_c, \omega_c)p(y_{s-1}|\bar{g}, h_c)d\mu(\bar{g}) \\ &= p(\bar{h}^*)\mathcal{P}(u, h_c, y_{s-1}, \omega_c), \end{aligned} \quad (24)$$

where $\mathcal{P}(u, h_c, y_{s-1}, \omega_c) = \int_{A^u} p(\bar{g}|h_c, \omega_c)p(y_{s-1}|\bar{g}, h_c)d\mu(\bar{g})$ is the exact probability density that involves PDF integral and requires reciprocal estimation.

Note that a reciprocal estimation is determined by the four variables included in Equation 24. They are the glossy count u , control vertex h_c , control direction ω_c , and the specular vertex y_{s-1} . These variables can be determined during the light sub-path tracing and are completely irrelevant to the eye sub-path. Therefore, the reciprocal estimation can be done in the light sub-path tracing instead of the run-time connection. This allows us to employ light sub-path reuse [Davidović et al. 2014] when tracing incomplete light sub-path \bar{h}_y . Specifically, we can trace and estimate the reciprocal for a small number of incomplete light sub-paths \bar{h}_y and connect those incomplete sub-paths with eye sub-paths from all the pixels. By doing so, we can save a significant amount of cost associated with reciprocal estimation.

5.3 Supporting Distribution Construction

Our reciprocal estimation requires a supporting distribution q to sample the proxy path \bar{g} in the estimation of $\frac{1}{\mathcal{P}(u, h_c, y_{s-1}, \omega_c)}$. We use three sampling strategies for the proxy path \bar{g} sampling and apply these strategies to the supporting distribution $q(\bar{g})$.

5.3.1 Light Source Sampling. If the only vertex that requires retracing is on the light source, i.e., when the alternate path $\bar{g} = y_0$, we can simply sample the light source surface to obtain y_0 . However, this method is disabled when $u > 1$, as in the case where $u = 2$ and alternate path $\bar{g} = y_0y_1$, with y_1 being sampled by tracing from y_0 . To satisfy the constraint on y_1 (which has to be on a specular surface), we need to trace y_2 from $y_0 \rightarrow y_1$, but y_2 is the specular vertex and fixed, making it impossible. As a result, $p(y_0 \rightarrow y_1 \rightarrow y_2)$ would be close to zero and the alternate path \bar{g} would be ineffective for rendering.

5.3.2 Tracing from the Control Vertex. If $u = 1$ and y_0 are not included in the alternate path \bar{g} , we trace the only alternate vertex y_{s-2} from the control vertex h_c . If the control vertex is on the light source, i.e., $h_c = y_0$, we sample the cosine hemisphere space at h_c to determine the tracing direction; otherwise, we sample the BSDF at h_c using ω_c as the incident direction to determine the tracing direction. This method is also disabled when $u > 1$ for the same reason as light source sampling.

5.3.3 Tracing from the Specular Vertex y_{s-1} . This method samples the alternate path \bar{g} by tracing from the specular vertex y_{s-1} . The initial direction is sampled from the cosine hemisphere space of y_{s-1} . For each intermediate vertex y_{s-i} , continue the tracing for

y_{s-i-1} based on the BSDF at y_{s-i} and direction $y_{s-i+1} \rightarrow y_{s-i}$ until all u vertices are traced.

The final supporting distribution is the uniform mixture of the available sampling strategies. That is, if $u = 1$ and $\bar{g} = y_0$, we use light source sampling and tracing from y_{s-1} to sample the alternate path \bar{g} ; otherwise, if $u = 1$ we sample the alternate path \bar{g} from the control vertex h_c and the specular vertex y_{s-1} ; for $u > 1$, the only available strategy is to sample the alternate path from the specular vertex y_{s-1} .

5.4 MIS Weighting Function for Reciprocal Estimation

While our method provides an efficient strategy for sampling specular paths, it may not be efficient in all cases. For instance, when the end of eye sub-path z_{t-1} is very close to the specular vertex y_{s-1} , a unidirectional path tracing that proceeds to trace the eye sub-path $z_{t-1} \rightarrow y_{s-1}$ may be more effective. Thus, it is necessary for our method to combine with the existing methods by multiple importance sampling.

A well-known MIS weighting function is the balance heuristic [Veach and Guibas 1995b] that assigns MIS weight proportional to the PDF of each strategy. This heuristic can minimize the upper bound of variance of the combined estimator. However, the PDF-proportional balance heuristic is derived when the PDF for each strategy can be uniquely determined. In the reciprocal estimation, our estimation of $\frac{1}{p(\bar{h})}$ is a random variable that may introduce extra variance to the combined estimator. Therefore, we need to derive the MIS weight for reciprocal estimation as follows.

Adopting the way similar to Veach and Guibas [1995b], our objective is to minimize $E(I^2)$, the upper bound of the combined estimator I is as

$$E(I^2) = \sum_t \int_{\Omega} w_t^2(\bar{x}) f^2(\bar{x}) E\left[\frac{1}{\tilde{p}_t^2(\bar{x})}\right] p_t(\bar{x}) d\mu(\bar{x}), \quad (25)$$

where t refers to the identity of sampling strategy, w_t is the corresponding MIS weighting function for strategy t , and $\frac{1}{\tilde{p}(\bar{x})}$ represents an unbiased estimation for $\frac{1}{p(\bar{x})}$.

The optimal MIS weighting function for w_t in Equation 25 is as:

$$w_t(\bar{x}) = \frac{1/(E[\frac{1}{\tilde{p}_t^2(\bar{x})}]) p_t(\bar{x})}{\sum_i 1/(E[\frac{1}{\tilde{p}_i^2(\bar{x})}]) p_i(\bar{x})}. \quad (26)$$

If $p_t(\bar{x})$ can be precisely computed without any variance, then $E[\frac{1}{\tilde{p}_t^2(\bar{x})}] = \frac{1}{p_t^2(\bar{x})}$ and Equation 26 reduces to the PDF-proportional balance heuristic. For the proxy sampling, its performance is decreased by the variance of reciprocal estimation, resulting in less MIS weight.

5.5 Subspace for Statistic

According to Sec. 4.2, the reciprocal estimation requires the upper bound B for optimization. In Sec. 5.4, $E[\frac{1}{\tilde{p}_i^2(\bar{x})}]$ and $p_i(\bar{x})$ are required to compute the MIS weight. However, those statistical data are unavailable in the practical estimation. We need an approximation for the statistical data.

In Equation 24, reciprocal estimation can be identified by four variables: the specular vertex y_{s-1} , the control vertex h_c , the control

direction ω_c , and the glossy count u . We adopt the same mechanism of subspace proposed by Su et al. Su et al. [2022] and trace a small number of $L(S|D)^*SDS^*E$ paths by PT at the beginning of rendering. The control and specular vertices from these paths are used to build control subspace and specular subspace, respectively, to classify h_c , ω_c , and y_{s-1} . A control subspace maps (h_c, ω_c) to control subspace label C ; a specular subspace maps y_{s-1} to a specular subspace label S . This allow us to approximate $\mathcal{P}(u, h_c, y_{s-1}, \omega_c)$ as

$$\mathcal{P}(u, h_c, y_{s-1}, \omega_c) \approx \mathcal{P}(u, C, S). \quad (27)$$

We do not use $\mathcal{P}(u, C, S)$ directly to estimate the reciprocal. Instead, we compute and store the statistical data for reciprocal estimation based on (u, S, C) and these statistical data are used to support rendering tasks such as selecting upper bound B and computing MIS weight.

5.6 Probabilistic Light Selection in Subspace

Our algorithm enables sampling of a specular path by connection strategy, which can further improve rendering efficiency when combined with existing probabilistic connection algorithms. However, some adaptations should be made to these connection algorithms for proper integration.

Most of the connection algorithms rely on the local contribution $f_y(\bar{y})$ of light sub-path to construct their resampling distribution. However, retracing of the proxy vertices may significantly alter the light sub-path, causing the $f_y(\bar{y})$ of the original sub-path to differ greatly from the new proxy sub-path. Therefore, connection algorithms should not use the $f_y(\bar{y})$ of the original sub-path in their contribution construction.

In our implementation, we employ the Subspace-based Probabilistic Connection for BPT (SPCBPT) [Su et al. 2022] to select the appropriate light sub-path for proxy sampling. We utilize the specular subspace discussed in section 5.5 to cache the incomplete light sub-paths and resample them based on their respective specular subspace.

A subspace sampling matrix $\Gamma[T, S]$ is required to determine the PMF for an eye subspace T to sample the specular subspace S in the first stage of subspace sampling. We construct the subspace sampling matrix based on the expected contribution of the proxy path. Specifically, $\Gamma[T, S]$ is proportional to the expected contribution for an incomplete sub-path $\bar{h}_y \in S$ and an eye sub-path $\bar{z} \in T$. We learn the Γ matrix on-the-fly during rendering. In the second sampling stage to sample the sub-path from the subspace, the sub-path is simply sampled uniformly in the specular subspace S to obtain the incomplete sub-path for the subsequent proxy sampling.

6 EXPERIMENTS AND RESULTS

We validate our approach in various scenarios and make comparisons with unidirectional path tracing (PT) with MIS combination of next event estimation (NEE) as the baseline, the state-of-the-art probabilistic connections method SPCBPT [Su et al. 2022], an SPCBPT with Light Tracing (LT) enabled, and an LVCBPT [Davidović et al. 2014] with LT enabled, which is a GPU accelerated version of BDPT that serves as a BDPT baseline. Our method is implemented based on SPCBPT and uses SPCBPT to handle the non-specular paths and focuses on the sampling of the specular involved path. Mean absolute

percentage error (MAPE) is used as the metric for comparison. We also provide the comparison and combination with path guiding of Müller et al. [2017], and the effect of subspace sampling. In Sec. 6.6, we provide additional results comparing SMIS/MMIS sampling with our reciprocal estimation technique in proxy tracing. Furthermore, in Sec. 6.7, we showcase the application of reciprocal estimation as a substitute for SMIS in photon plane volume rendering, achieving superior performance in scenarios where SMIS previously excelled.

6.1 Experimental Setting

6.1.1 Renderer Setting. All the algorithms are implemented based on the OptiX architecture [Parker et al. 2010], and run on an NVIDIA GeForce RTX 3090 GPU with an Intel Core i9-12900K CPU on a Windows system. All images are rendered in a resolution of 1920×1000 . We build 300 eye subspaces and 300 light subspaces for SPCBPT. Compared to the default setting of 1000 subspaces in SPCBPT, a smaller number of subspaces can avoid overfitting and works better for scenes with easy visibility, as discussed in Su et al. [2022]. We use PT to pretrace 1,000,000 full paths to train the sampling distribution of SPCBPT. In each iteration, SPCBPT, LVCBPT and our method all trace $M = 10,000$ light sub-paths for resampling. Our approach shares the same parameter and configuration with SPCBPT as recommended in Su et al. [2022], unless otherwise stated.

In our proxy sampling method, we impose constraints on the path processing. We drop out the problematic vertices and obtain the incomplete light sub-path only if there is no control vertex or the control vertex is on the light source. This constraint limits the light sub-path type to $L_{DD}S^+$ and L_{DDDS}^+ , as the contribution of $L_{DDDS}^+S^+$ is negligible and not visible in the rendered image. Additionally, the glossy count u is limited to $1 \leq u \leq 4$. Paths that do not satisfy these constraints are handled by SPCBPT. In each iteration, we perform reciprocal estimation on up to 400 incomplete light sub-paths and discard the others to reduce overhead as much as possible. The incomplete sub-paths are divided into 100 specular subspaces for probabilistic connections. We use only 10 subspaces for the control vertex since the control vertex is limited to the light source. To reduce the variance of reciprocal estimation, we repeat the reciprocal estimation 5 times for each incomplete light sub-path and use the average as the final estimation. We learn the subspace sampling matrix in the first 40 iterations and then freeze the matrix for stable rendering performance. We only sample the specular light sub-path and perform proxy sampling when the eye sub-path travels to its first diffuse surface. Additionally, we divide the image into grids, with each grid containing 10×10 pixels. If the contribution of our target path exceeds 5% in a grid, we sample the incomplete light sub-path and perform the proxy sampling in probability 100%; otherwise, the probability is set to 20%.

We adopt the DISNEY principled BSDF [Brent Burley 2012] for the BSDF model of highly glossy material. The roughness of most of the specular material in our experiment is set to 0.01. A material is identified as specular when the material is metallic or transparent with roughness smaller than 0.2.

6.1.2 Scene Settings. Six benchmarks, including *Bedroom*, *Breakfast*, *Projector*, *Hallway*, *Kitchen*, and *Water*, are tested in our experiment. These benchmarks cover different illumination conditions and glossy surface settings.

The *Bedroom* scene is illuminated by a light source positioned above a mirror. Some of the light directly illuminates the scene, while the other light is reflected off the mirror and into the scene as our target path. The *Mirror* on the wall shows the effect of L_{DDSDSE} path. In the *Breakfast* scene, two glossy angular spheres are positioned above two cylindrical light sources on the table to create red and blue caustic patterns on the wall. The *Projector* scene features a textured light source enclosed in a box with a convex lens, creating a projector effect. The convex lens refracts the textured light before projecting it onto the wall. *Hallway* is mainly illuminated by two light sources enclosed in a glass lamp, requiring the light to travel through the refractive glass to enter the scene. In the *Kitchen* scene, the light comes from the outside and is located behind a two-layer glass window. Finally, the *Water* scene displays a classic caustic pattern created by small light sources and complex refractive surfaces.

The preprocessing times for SPCBPT to pre-trace (using PT) the full paths never exceed 5 seconds, which depends on the efficiency of PT. On average, each pixel only needs to trace approximately 0.5 path to obtain 1,000,000 full paths for preprocessing. This time cost is negligible compared to the convergence time of PT. Therefore, the preprocessing time is not taken into consideration during our comparisons. *Projector*, *Hallway*, and *Water* are rendered by PT in 10M iterations and the rest scenes are rendered by SPCBPT in 100K iterations.

6.2 Performance Evaluation

We show the performance of different methods in Figure 1 and Figure 7. We also highlight the zoom-in regions for a detailed comparison. When rendering difficult specular-involved paths, both PT and SPCBPT can only rely on the sampling of the unidirectional path tracing. However, within the given time budget, PT makes a much higher number of iterations than that of SPCBPT due to its very low overhead. Therefore, PT can sample this specular path more efficiently than SPCBPT. Conversely, SPCBPT can select the appropriate light sub-path in the connection and produce much better results in indirect illumination. The performance difference is evident in the highlight region of *Breakfast* and *Hallway* at Figure 7. Our method focuses on sampling the specular-involved path and utilizes the strategy from SPCBPT to sample the residual path. In most cases, we can achieve better rendering performance on difficult paths like $L_{DD}(S|D)^*SDS^E$ than PT and SPCBPT using the same time budget and fewer iterations. Our method takes extra computational overhead and the rendering speed for each iteration of our method is slower than SPCBPT at around 25% ~ 50%. The extra overhead comes from reciprocal estimation, retracing for the proxy vertices, and the MIS weight computation.

Figure 8 shows the convergence over time in terms of MAPE in all scenarios. Our approach outperforms others significantly in terms of convergence performance compared to other methods. In particular, in *Breakfast*, *Projector*, and *Hallway*, neither PT nor SPCBPT



Fig. 7. Equal-time comparison: PT with NEE integration, SPCBPT [Su et al. 2022], SPCBPT with LT enabled, and our approach. Rendering time cost is 180s for Water and 60s for other scenes. The leftmost side shows the full frame synthesized by our approach. The number of Iterations (SPP, samples per pixel) and MAPE are shown in the zoom-in region. Our approach shows superior performance in terms of MAPE and produces less noisy results than other approaches.

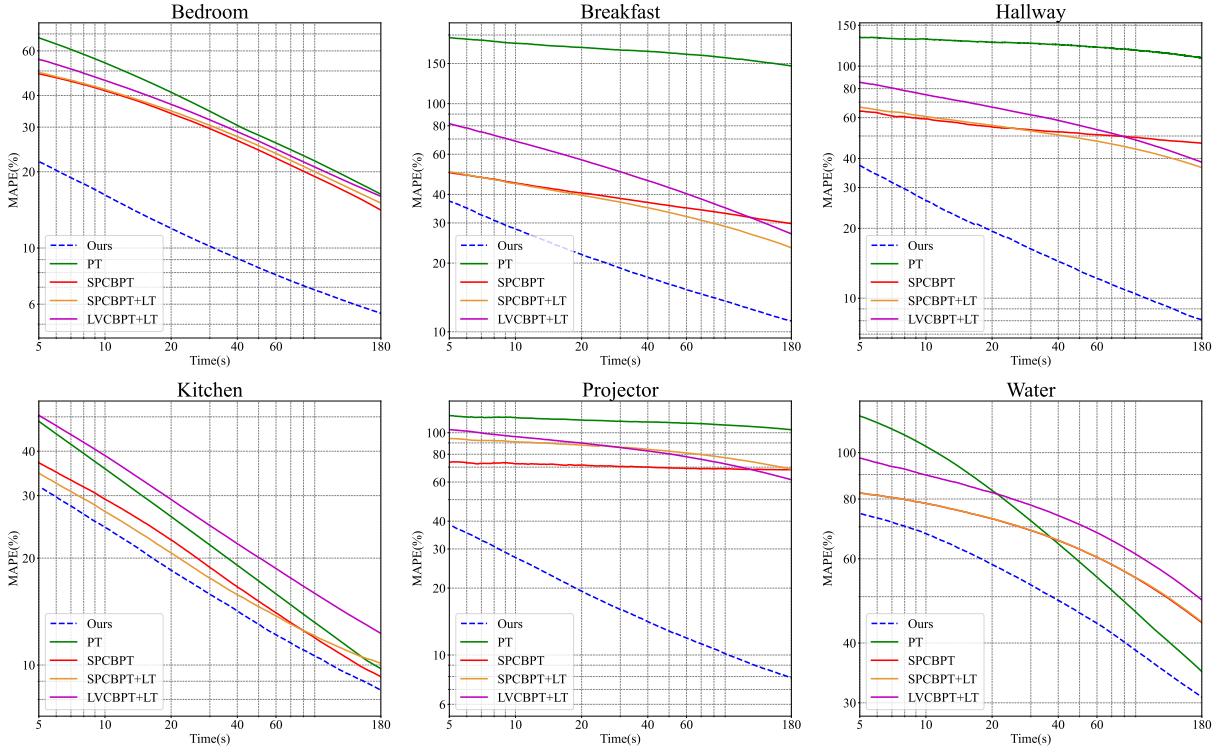


Fig. 8. Convergence over time (3 min) in terms of MAPE for PT, SPCBPT [Su et al. 2022], SPCBPT with LT, LVCBPT [Davidović et al. 2014] with LT, and ours. Our approach shows superior convergence performance in all tested scenarios. In *Water*, the SPCBPT curve overlaps with the SPCBPT+LT curve because a specular and transparent surface seals the box, making LT have no contribution.

Table 3. Time cost (measured in seconds) and iterations to reach a specified MAPE by different methods. Our approach shows the best performance consistently and outperforms others significantly in both time cost and iteration required.

Scene	Bedroom (20%)		Breakfast (30%)		Hallway (35%)		Kitchen (10%)		Projector (20%)		Water (35%)	
	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time
PT	6533	113.35	>14324	>200	>7631	>200	9248	168.42	>11071	>200	21225	178.37
SPCBPT	582	81.05	1529	171.45	>1295	>200	788	142.44	>1357	>200	>2691	>200
SPCBPT+LT	683	90.22	609	78.68	>1311	>200	1052	189.38	>1412	>200	>2666	>200
LVCBPT+LT	983	100.27	1739	135.54	>1790	>200	>1358	>200	>1844	>200	>2783	>200
Ours	29	6.23	62	8.63	27	5.73	337	107.26	70	18.93	829	121.32

is able to handle all the light paths well, resulting in slow convergence over time. In contrast, our method can effectively sample the $L_{DD}(S|D)^*SDS^*E$ path while allowing sufficient time for other advantageous sampling strategies from SPCBPT to contribute to the indirect illumination, ensuring stable convergence. While LT can handle paths with type $L_{DD}SDE$, like the reflected light on the wall in *Breakfast*. However, our method achieves better performance than SPCBPT with LT, demonstrating its efficiency in sampling these paths. The *Breakfast* scene is typically considered manageable with a pure LT algorithm. As the BDPT baseline, LVCBPT+LT generally shows poorer performance than SPCBPT+LT, as shown in Figure 8

and Table 3. In *Projector*, the subspace-based probabilistic sampling will struggle with specular paths, and LVCBPT+LT exhibits a slight advantage due to faster iteration speeds. However, in other test scenes, results show that SPCBPT+LT outperforms LVCBPT+LT, which is consistent with the experimental results reported in Su et al. [2022]. Our method shows the best performance in all scenes.

To further evaluate our algorithm, we conducted an additional experiment comparing LT with our approach, as shown in Figure 9. Remarkably, our method outperformed LT, delivering superior rendering results in just 60s, while LT required 180s but with much



Fig. 9. Comparison between Light Tracing (LT) and our approach in *Breakfast*. Our approach delivers superior rendering results in 60s, whereas LT takes 180s. This showcases the efficiency of our method in efficiently sampling challenging paths and producing high-quality caustic patterns in significantly less time than LT.

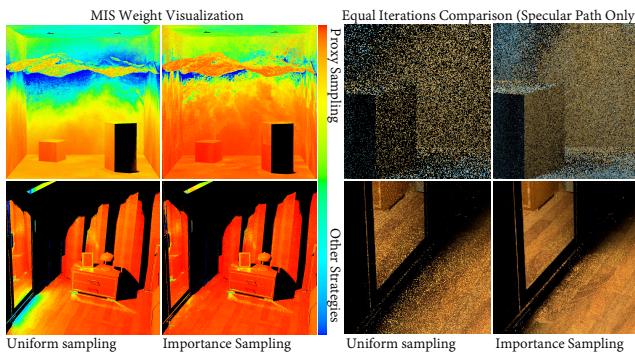


Fig. 10. MIS weight visualization and equal iterations (32 iterations for *Bedroom*, 128 iterations for *Water*) comparison for our method with/without subspace-based importance sampling for the incomplete light sub-path. Overall MAPE for uniform sampling vs. importance sampling is 69.13% vs. 54.98% in *Water* and 20.96% vs. 19.74% in *Bedroom*.

noise. This highlights the efficiency of our method in effectively sampling challenging paths and producing high-quality caustic patterns in significantly less time than LT.

We also report the statistics of time cost and iterations required to reach the specified MAPE in Table 3. Our approach demonstrates higher efficiency over SPCBPT, SPCBPT with LT, and PT, although the improvement varies across different scenes. For instance, in scenarios illuminated by large light sources like *Kitchen*, tracing the direct illumination of specular light sub-path is relatively easier for PT. So, our improvement is less impressive. In contrast, in a challenging scene where the light source is very small, it becomes extremely difficult for unidirectional path tracing to find the light source. Our approach is particularly advantageous.

6.3 Sampling in Subspace

In our approach, subspace-based probabilistic sampling is used to select the incomplete light sub-path for proxy sampling. We learn the effect of subspace sampling through equal-iteration comparison and visualizing the MIS weight function for proxy sampling. The comparison for our method with/without subspace-based probabilistic sampling is shown in Figure 10. We focus on the sampling of

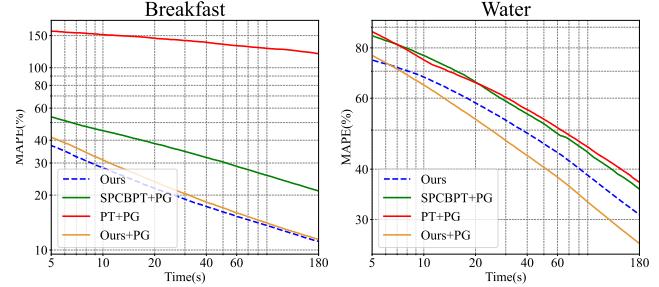


Fig. 11. Convergence over time in terms of MAPE. We compare the performance of PT+Path Guiding (PG), SPCBPT+PG, our approach+PG, and our approach solely. 60s is used for *Breakfast* and 180s for *Water*. Path guiding is pre-trained for 30s, and the pretraining time cost is not included. Our approach outperforms SPCBPT+PG significantly in *Breakfast*, and achieves speedup when combined with PG in *Water* scene.

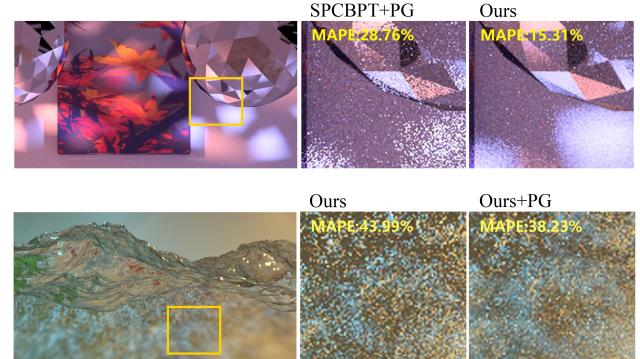


Fig. 12. Rendering results obtained by equal time (60s for *Breakfast* and 180s for *Water*) corresponds to Figure 11. The upper row demonstrates the superiority of our approach over SPCBPT+PG; the lower row highlights the improvement of our algorithm through combination with PG in this *Water* scene.

the specular-involved $L_{DD}(S|D)^*SDS^*E$ path and disable the rendering for the other paths. MIS weight is assigned based on the quality of proxy sampling, and the red regions indicate the advantage area for our method. Generally, the efficiency of proxy sampling is highest when the diffuse surface is far from the specular surface; thus, the improvement of subspace-based probabilistic sampling is relatively small in this case (overall MAPE 20.96% → 19.74% in *Bedroom*). However, in a more complicated scene setting where the diffuse surface is close to the specular surface, the probabilistic connections method can help find the appropriate incomplete path for connection and provide efficient enhancement for the specular path rendering (overall MAPE 69.13% → 54.98% in *Water*), thus, extending the capability of proxy sampling.

6.4 Combination and Comparison with Path Guiding

Our approach can be integrated well with the sophisticated path guiding technique. We integrate the path guiding method based on the SD-tree Müller et al. [2017] with our method and test its

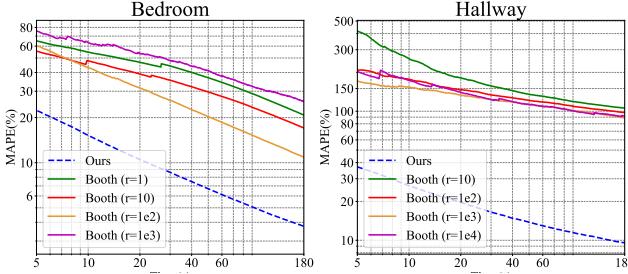


Fig. 13. Convergence over time in terms of MAPE. We perform an ablation study to assess the performance with and without our RR technique. The base algorithm is Booth [2007] with parameter r set at four different values. The results demonstrate the effectiveness and superiority of our RR technique.

performance at *Breakfast* and *Water*. For each scene, we spend 30s to train the SD-tree for path guiding. We show the equal-time rendering with convergence plot in Figure 11, and visual result in Figure 12, time cost for pre-training not included.

Since our approach can handle the specular-involved path in *Breakfast* exceptionally well, it converges fast within just 60s. Therefore, the speedup from path guiding is not significant and may even slightly slow down the performance of our approach, as shown in the *Breakfast* scene. We also observed that our approach significantly outperforms SPCBPT+PG, possibly due to the small light source and the diffuse surface being far from the specular surface. In such case, a 30s training may not be sufficient for path guiding to help SPCBPT find the location of the light source quickly. In the scenario of *Water*, the performance of our proxy sampling declines when the diffuse surface is very close to the specular surface (as indicated by the blue region in the MIS visualization of Figure 10). Because in this case the specular vertex y_{s-1} of the appropriate incomplete sub-path is usually very close to the eye sub-path and harder to select in the incomplete sub-path sampling, making our proxy sampling ineffective. In contrast, random direction sampling has a high probability of performing a good $D \rightarrow S$ bounce when the distance between diffuse surface D and specular surface S is very close. In such cases, combining our method with path guiding can gain benefits from both sides and improve performance, leading to even better results for this complicated setting.

6.5 Ablation Study on RR Technique

To assess the effectiveness of our RR technique, we conducted an ablation study, as shown in Figure 13. Due to the integration challenges with the method proposed by Blanchet et al. [2015], we opted for the framework described by Booth [2007] as our baseline. Their RR strategy requires a manually set hyperparameter, r , and also necessitates an upper bound, B , which was the same in their experiments and ours. Our experimental findings reveal that the performance of the algorithm generally improves with an increase in r up to a certain point, after which it diminishes. Specifically, the turning point occurs at an order of magnitude of 100 in the Bedroom scenario and at 1000 in the Hallway scenario, indicating that the optimal range for r varies between different scenes. our

results significantly outperform the baseline, irrespective of the r 's setting, suggesting that our RR method can effectively enhance the performance of reciprocal estimation.

6.6 Comparison with SMIS/MMIS

SMIS/MMIS can serve as a substitution for our reciprocal estimation in our proxy tracing approach. Therefore, we have integrated SMIS/MMIS into our proxy tracing framework and conducted a comparative analysis with our reciprocal estimation-based approach, using scenes such as *Bedroom*, *Hallway*, and *Projector*. For this comparison, We only render the paths enhanced with proxy tracing, as the other parts remain consistent across methods. The n in SMIS n refers to the number of techniques combined in SMIS. We mainly compare with $n = 4$ and $n = 8$ cases here since a larger n would be costly due to the $O(n^2)$ cost in the MIS process. Figure 14 illustrates the convergence over time in terms of MAPE in these scenes. Our reciprocal estimator outperforms SMIS/MMIS significantly. Reflecting on the limitations of SMIS/MMIS outlined in Sec. 4.3, it becomes evident that these constraints are noticeable when applied to scenes like *Hallway* and *Projector*.

In *Hallway*, there are four light sources including an environment light. The scene is predominantly illuminated by two lights within glass lamps. It corresponds to the second failure case described in Sec. 4.3, where different lights have distinct visibilities. Figure 15 reveals the bias inherent in SMIS n . In SMIS2, SMIS4, and SMIS8, the illuminated areas appear noticeably darker compared to our unbiased approach. Increasing n reduces this bias, making SMIS16 visually closer to our technique. However, note that there are only four light sources in this scene. In more complex scenes with hundreds or thousands of light sources, SMIS would require an impractically large n to get good results. Additionally, we examined MMIS4 in Figure 15. MMIS, which integrates SMIS with unbiased techniques like PT, mitigates the bias of SMIS. Specifically, the hot pixels in the output indicate other techniques' contribution. Nevertheless, as no technique effectively samples these paths, the overall efficiency remains sub-optimal. In *Projector*, the primary source of illumination is a projector that comprises a textured light source within a box and a convex lens. This setup aligns with the first failure case for SMIS, characterized by multiple specular vertices following the light vertex. In Figure 15, we demonstrate the inability of SMIS/MMIS in this scenario. Conversely, in *Bedroom*, with its single light source and absence of consecutive specular surfaces, conditions are seemingly favorable for SMIS/MMIS. Despite this, our approach still significantly outperforms SMIS/MMIS, underscoring its high efficiency even in advantageous settings for the comparative methods.

6.7 Reciprocal Estimation for CMIS Framework

Our reciprocal estimation method can be applied to the CMIS framework since it can provide unbiased and efficient estimation for CMIS weight $\tilde{w}(t, x)$ in Equation 6. Here, we apply our method to the photon plane algorithm [Deng et al. 2019], which is an algorithm where CMIS excels.

Photon plane is introduced by Deng et al. [2019] for single scattering volumetric rendering. The sampling method is shown in

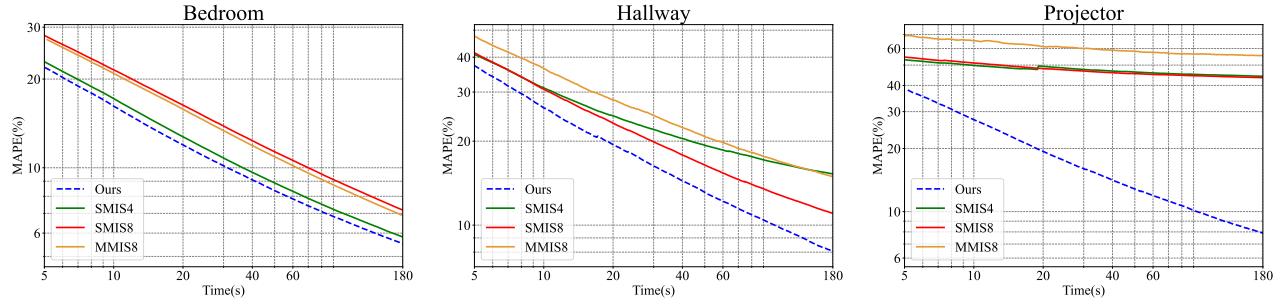


Fig. 14. Convergence over time in terms of MAPE. We compare the performance of SMIS4, SMIS8, MMIS8, and our approach. 120s is used for each scene. Our approach outperforms SMIS/MMIS significantly in *Hallway* and *Projector*, and even in *Bedroom* where SMIS/MMIS was considered to be advantageous.



Fig. 15. Comparison between SMIS/MMIS and our approach based on reciprocal estimation. We focus exclusively on rendering paths that are enhanced with proxy tracing, as the other components remain identical. (a) In *Hallway*, SMIS n displays obvious bias (darker) when n is small. MMIS4 mitigates SMIS4's bias by combining unbiased techniques, the contribution of which appears as hot pixels. (b) In *Projector*, SMIS/MMIS completely fails to capture the specular-involved difficult paths (see left), while our approach can handle this scene efficiently (see right).

Figure 16(a). A photon plane is determined by a segment s in the light source and an outgoing direction ω_l . When an eye sub-path hits the photon plane, a light source vertex can be found by projecting the intersection vertex in direction $-\omega_l$, then a full path \bar{x} can be sampled by connecting the camera, the intersection and the projected vertex on the light source.

Computing $p(\bar{x})$ is a challenge since the same \bar{x} can be generated from different photon plane. West et al. [2020] indicated that the sampling of the segment s can be split into the sampling of rotation angle α and bias v , and α can be treated as the technique identifier in technique space $\mathcal{T} = [0, \pi]$ for CMIS framework, as shown in

Figure 16(b). The CMIS weight $\tilde{\omega}(\alpha, \bar{x}) = \frac{p(\alpha, \bar{x})}{\int_{\mathcal{T}} p(\alpha, x) d\alpha}$ involves the reciprocal of integral which is handled by SMIS in West et al. [2020]. We replace SMIS by our reciprocal estimation method. The efficiency of reciprocal estimation depends on the supporting distribution for sampling α . The optimal supporting distribution $q^*(\alpha|\bar{x})$ depends on ω_e as well as the position of the vertex on the light source. Therefore, we cache the piece-wise approximation of $q^*(\alpha|\bar{x})$ for multiple \bar{x} , and select the appropriate $q(\alpha|\bar{x})$ based on ω_e and the vertex on the light source of current \bar{x} . The memory cost of the cache is within 16MB.

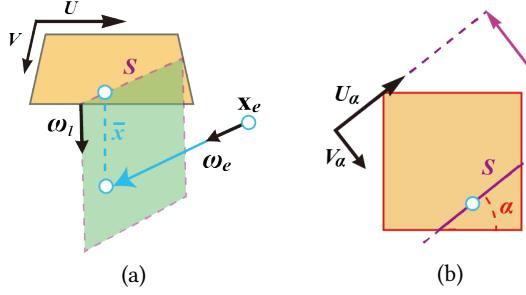


Fig. 16. (a) Sampling method of photon plane algorithm [Deng et al. 2019]. Photon plane is generated by sampling a segment s in the light source and an outgoing direction ω_l . When an eye sub-path hits the photon plane, the camera, the intersection position, and the ω_l -projection of the intersection position on the light source make the full path \bar{x} . (b) A segment s can be sampled by first sampling a rotation angle α and then sampling the bias v .

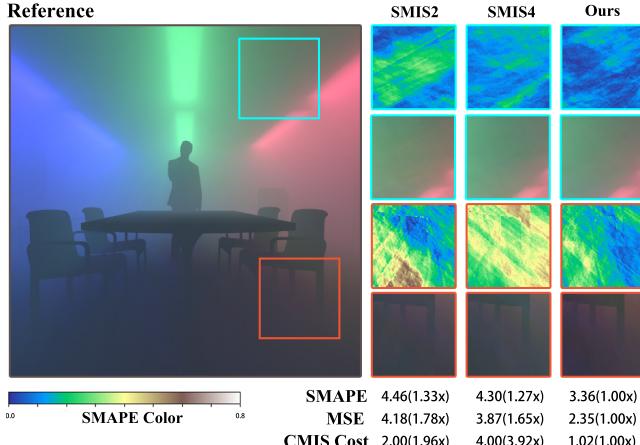


Fig. 17. Equal-sample comparison between estimators for SMIS [West et al. 2020] and our method, specifically in volumetric scenarios rendered by photon plane algorithm [Deng et al. 2019]. SMAPE is visualized by heat map and MSE below is magnified by 10^4 . CMIS Cost refers to the average time cost of $p(\alpha, \bar{x})$ used for CMIS weight estimation.

The equal-sample comparison between SMIS2/SMIS4 and our method is shown in Figure 17. Note that SMIS and our method incur different computational costs for evaluating $p(\alpha, \bar{x})$ to estimate the CMIS weight, $\tilde{\omega}(\alpha, \bar{x})$. Our method achieves more efficient CMIS weight estimation with lower computation cost, showing the advantages of our reciprocal estimation method. In terms of SMAPE (Symmetric Mean Absolute Percentage Error) and MSE, our approach also shows superior performance over the alternative approaches.

7 CONCLUSION, LIMITATION, AND FUTURE WORK

In this paper, we have proposed a novel method for tracing difficult specular-involved paths based on efficient reciprocal estimation. By discarding problematic segments of the original path and tracing a proxy path that can meet specular constraint, we enable a

probabilistic connection to handle specular paths well. To ensure unbiased estimations, we introduce a sub-optimal RRS function that significantly boosts efficiency. Additionally, we propose an optimal setting of B , minimizing the upper bound on the cost of our reciprocal estimation. Specifically, This optimized setting not only aids in the reciprocal estimation required for tracing incomplete sub-paths but also employs a light sub-path reuse strategy to diminish the overhead associated with reciprocal estimations. We have developed a method that allows for replacing small segments of the path, resulting in significant improvements. We believe this approach serves as a powerful tool, enabling the seamless substitution of problematic paths with preferable alternatives while maintaining the unbiased nature of the Monte Carlo estimation of the path integral.

To provide an intuitive understanding, our method improves connection efficiency by broadening the narrow lobe of the BSDF on specular surfaces to encompass a wider range. Specifically, for the $L_{DD}S^*$ type light sub-path, this expanded range correlates with the projection area of an area light source at the specular vertex. This enhancement facilitates the inclusion of an eye sub-path within the valid range more easily, particularly when the eye sub-path is far from the specular vertex or when the light source is large. However, challenges arise when the specular surface is far from a small light source or when the diffuse eye vertex is near the specular light vertex. Under such conditions, our connection strategy may struggle to be effective. Despite the assistance of subspace-based probabilistic connections, the efficiency of our method may still face limitations. This is a fundamental challenge of connection-based sampling strategies, highlighting an area ripe for future research endeavors.

Another limitation of our method pertains to the sampling efficiency of proxy vertices, which yields optimal results when these vertices are readily samplable. Constructing an efficient supporting distribution for reciprocal estimation presents significant challenges. To mitigate these issues, we have incorporated path guiding techniques, as demonstrated in the *Water* scene depicted in Figure 12, specifically to address scenarios involving close specular-diffuse interactions. Furthermore, it is feasible to integrate path guiding with our method to develop a more efficient supporting distribution. However, the specifics of such an integration require further detailed exploration.

In addition to enhancing the rendering of paths involving specular interactions, our proposed proxy sampling method holds promise for a range of applications. It allows for the retracing of problematic segments of a full path while preserving most existing vertices, complementing current sampling methodologies. Identifying appropriate scenarios for its application and reducing the computational overhead associated with reciprocal estimation are challenges we aim to tackle in our future research. Furthermore, our method has the potential to become a state-of-the-art BDPT algorithm. It can be effectively integrated with the recent Progressive Photon Mapping (PPM) algorithm to form a highly efficient Vertex Connection and Merging (VCM) technique [Lin et al. 2023]. This integration could significantly speed up the VCM technique. Exploring this integration represents a promising avenue for our future research.

ACKNOWLEDGMENTS

We extend our gratitude to all the anonymous reviewers for their helpful suggestions. This work is supported by the National Key R&D Program of China (No. 2023YFF0905103) and NSFC of China (No. 62172013). We also wish to thank all the test scenes providers: MrChimp2313 (House), Wig42 (Hallway), SlykDrako (Bedroom), Jay-Artist (Kitchen), Yaoyi Bai et al. (Projector), Benedikt Bitterli (Water). We would like to thank Mr. Jierui Ren and Mr. Xiaobai Chen for their assistance in making parts of the illustrations.

REFERENCES

- Sai Bangaru, Tzu-Mao Li, and Frédéric Durand. 2020. Unbiased Warped-Area Sampling for Differentiable Rendering. *ACM Trans. Graph.* 39, 6 (2020), 245:1–245:18.
- Jose H. Blanchet, Nan Chen, and Peter W. Glynn. 2015. Unbiased Monte Carlo computation of smooth functions of expectations via Taylor expansions. In *2015 Winter Simulation Conference (WSC)*. 360–367. <https://doi.org/10.1109/WSC.2015.7408178>
- Thomas E. Booth. 2007. Unbiased Monte Carlo Estimation of the Reciprocal of an Integral. *Nuclear Science and Engineering* 156, 3 (2007), 403–407. <https://doi.org/10.13182/NSE07-A2707>
- Walt Disney Animation Studios Brent Burley. 2012. Physically-based shading at disney. In *ACM SIGGRAPH*. 1–27.
- D. Chandler. 1987. *Introduction To Modern Statistical*. Introduction To Modern Statistical.
- T. Davidović, J. Krivanek, M. Hasan, and P. Slusallek. 2014. Progressive Light Transport Simulation on the GPU. *ACM Trans. Graph. (TOG)* 33, 3 (05 2014), 29:1–19.
- X. Deng, S. Jiao, B. Bitterli, and W. Jarosz. 2019. Photon surfaces for robust, unbiased volumetric density estimation. *ACM Trans. Graph. (TOG)* 38, 4 (07 2019), 46:1–12.
- I. Georgiev, J. Krivanek, T. Davidovic, and P. Slusallek. 2012. Light Transport Simulation with Vertex Connection and Merging. *ACM Trans. Graph. (TOG)* 31, 6 (11 2012), 192:1–192:10.
- Toshiya Hachisuka and Henrik Wann Jensen. 2009. Stochastic progressive photon mapping. *ACM Transactions on Graphics (TOG)* 28, 5 (2009), 141.
- T. Hachisuka, S. Ogaki, and H. W. Jensen. 2008. Progressive Photon Mapping. *ACM Trans. Graph. (TOG)* 27 (12 2008), 130.
- T. Hachisuka, J. Pantaleoni, and H. W. Jensen. 2012. A Path Space Extension for Robust Light Transport Simulation. *ACM Trans. Graph. (TOG)* 31, 6 (11 2012), 191:1–10.
- Johannes Hanika, Marc Droske, and Luca Fuscione. 2015. Manifold Next Event Estimation. *Comput. Graph. Forum* 34, 4 (07 2015), 87–97.
- H. W. Jensen. 1996. Importance Driven Path Tracing Using the Photon Map. *Eurographics Rendering Workshop* (09 1996), 326–335.
- H. W. Jensen and N. Christensen. 1995. Photon Maps in Bidirectional Monte Carlo Ray Tracing of Complex Objects. *Computers & Graphics* 19 (03 1995), 215–224.
- Ivo Kondapaneni, Petř Vévoda, Pascal Grittmann, Tomáš Skřiván, Philipp Slusallek, and Jaroslav Krivánek. 2019. Optimal multiple importance sampling. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 1–14.
- E. P. Lafourcade and Y. D. Willems. 1999. A 5D Tree to Reduce the Variance of Monte Carlo Ray Tracing. In *Rendering Techniques*.
- 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*. 1–8.
- Zehui Lin, Chenxiao Hu, Jinzhu Jia, and Sheng Li. 2023. Hypothesis Testing for Progressive Kernel Estimation and VCM Framework. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–15. <https://doi.org/10.1109/TVCG.2023.3274595>
- Z. Lin, S. Li, X. Zeng, C. Zhang, J. Jia, G. Wang, and D. Manocha. 2020. CPPM: chi-squared progressive photon mapping. *ACM Trans. Graph. (TOG)* 39, 6 (11 2020), 240:1–12. <https://doi.org/10.1145/3414685.3417822>
- T. Müller, M. Gross, and J. Novák. 2017. Practical Path Guiding for Efficient Light-Transport Simulation. *Computer Graphics Forum* 36, 4 (07 2017), 91–100.
- K. Nabata, K. Iwasaki, and Y. Dobashi. 2020. Resampling-aware Weighting Functions for Bidirectional Path Tracing Using Multiple Light Sub-Paths. *ACM Trans. Graph. (TOG)* 39, 2 (03 2020), 15:1–11.
- Jan Novák, Iliyan Georgiev, Johannes Hanika, and Wojciech Jarosz. 2018. Monte Carlo methods for volumetric light transport simulation. *Computer Graphics Forum (Proceedings of Eurographics - State of the Art Reports)* 37, 2 (2018). <https://doi.org/10/gd2jqq>
- S. G. Parker, J. Bigler, A. Dietrich, H. Friedrich, J. Hoberock, D. Luebke, D. McAllister, M. McGuire, K. Morley, A. Robison, et al. 2010. Optix: a general purpose ray tracing engine. *ACM Trans. Graph. (TOG)* 29, 4 (07 2010), 66:1–13.
- S. Popov, R. Ramamoorthi, F. Durand, and G. Drettakis. 2015. Probabilistic Connections for Bidirectional Path Tracing. *Computer Graphics Forum* 34, 4 (07 2015), 75–86.
- Hao Qin, Xin Sun, Qiming Hou, Baining Guo, and Kun Zhou. 2015. Unbiased Photon Gathering for Light Transport Simulation. *ACM Trans. Graph.* 34, 6, Article 208 (nov 2015), 14 pages. <https://doi.org/10.1145/2816795.2818119>
- A. Rath, P. Grittmann, Sebastian H., P. Vévoda, P. Slusallek, and J. Krivánek. 2020. Variance-Aware Path Guiding. *ACM Trans. Graph. (TOG)* 39, 4 (07 2020), 151:1–12.
- Alexander Rath, Pascal Grittmann, Sebastian Herholz, Philippe Weier, and Philipp Slusallek. 2022a. EARS: Efficiency-Aware Russian Roulette and Splitting. *ACM Trans. Graph. (TOG)* 41, 4, Article 81 (jul 2022), 14 pages. <https://doi.org/10.1145/3528223.3530168>
- Alexander Rath, Pascal Grittmann, Sebastian Herholz, Philippe Weier, and Philipp Slusallek. 2022b. EARS: Efficiency-Aware Russian Roulette and Splitting. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2022)* 41, 4, Article 81 (jul 2022), 14 pages. <https://doi.org/10.1145/3528223.3530168>
- F. Su, S. Li, and G. Wang. 2022. SPCBPT: Subspace-based Probabilistic Connections for Bidirectional Path Tracing. *ACM Trans. Graph. (TOG)* 41, 4 (07 2022), 77:1–14.
- Y. Tokuyoshi and T. Harada. 2018. Bidirectional path tracing using backward stochastic light culling. 1–2.
- Y. Tokuyoshi and T. Harada. 2019. Hierarchical russian roulette for vertex connections. *ACM Trans. Graph. (TOG)* 38, 4 (07 2019), 36:1–12.
- Eric Veach. 1998. *Robust monte carlo methods for light transport simulation*. Ph.D. Dissertation. Stanford, CA, USA. Advisor(s) Guibas, Leonidas J. AAI9837162.
- E. Veach and L. Guibas. 1995a. Bidirectional estimators for light transport. In *Photorealistic Rendering Techniques*. 145–167.
- E. Veach and L. Guibas. 1995b. Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*. 419–428.
- J. Vorba, O. Karlík, M. Šík, T. Ritschel, and J. Krivanek. 2014. On-line Learning of Parametric Mixture Models for Light Transport Simulation. *ACM Trans. Graph. (TOG)* 33, 4 (07 2014), 101:1–11.
- Rex West, Iliyan Georgiev, Adrien Gruson, and Toshiya Hachisuka. 2020. Continuous Multiple Importance Sampling. *ACM Trans. Graph.* 39, 4, Article 136 (aug 2020), 12 pages. <https://doi.org/10.1145/3386569.3392436>
- Rex West, Iliyan Georgiev, and Toshiya Hachisuka. 2022. Marginal Multiple Importance Sampling. In *SIGGRAPH Asia 2022 Conference Papers (SA '22)*. Association for Computing Machinery, New York, NY, USA, Article 42, 8 pages. <https://doi.org/10.1145/3550469.3555388>
- Shanhe Wu. 2008. Some improvements of Aczél's inequality and Popoviciu's inequality. *Computers & Mathematics with Applications* 56, 5 (2008), 1196–1205. <https://doi.org/10.1016/j.camwa.2008.02.021>
- M Zackary, B Benedikt, G Iliyan, and J Wojciech. 2022. Unbiased and consistent rendering using biased estimators. *ACM Trans. Graph. (TOG)* 41, 4 (7 2022), 48:1–13.
- Tizian Zeltner, Iliyan Georgiev, and Wenzel Jakob. 2020. Specular Manifold Sampling for Rendering High-Frequency Caustics and Glints. *Transactions on Graphics (Proceedings of SIGGRAPH)* 39, 4 (July 2020). <https://doi.org/10.1145/3386569.3392408>

APPENDIX

We provide a detailed description of our reciprocal estimator, along with the convergence condition, expectation, variance analysis, and optimal settings.

A OUR ESTIMATOR

Given:

- Target function $f(x) > 0$
- Sample distribution $p(x)$ s.t. $p(x) > 0$ when $f(x) > 0$
- Sample $X \sim p(x)$
- RRS function $r(x|\bar{z}) > 0$
- Constant $B > 0$

Here \bar{z} refers to the precedent samples of the estimation.

We give our estimator \tilde{I}' , corresponding to $\tilde{I}(\bar{x})$ in Equation 17 of the text, that estimates $\frac{1}{\int_R f(x) dx} \cdot \int_R f(x) dx$.

Note that the target of $\tilde{I}'(\bar{z})$ is to estimate $\frac{1}{\int_R f(x) dx} - \frac{1}{B}$ whatever the value of the president samples \bar{z} and \tilde{z} is irrelevant to $g(x)$. It is obvious that an optimal RRS function $r^*(x|\emptyset)$ for $I'(\emptyset)$ is also optimal for any $I'(\bar{z})$ of president samples \bar{z} . Therefore, we simply ignore the existence of \bar{z} and discuss the optimization of an RRS function $r(x)$ that only depends on the current sample x in the rest of our derivation.

Algorithm 1 Estimator \tilde{I}

```

1: procedure  $\tilde{I}$ 
2:   return  $\tilde{I}'(\emptyset) + 1/B$ 
3: end procedure
4: procedure  $\tilde{I}'(\bar{z})$ 
5:    $x_0 \leftarrow p.sample()$ 
6:    $g_0 \leftarrow 1 - f(x_0)/(B * p(x_0))$ 
7:    $r_0 \leftarrow r(x_0|\bar{z})$ 
8:    $r'_0 \leftarrow \lfloor r_0 \rfloor$ 
9:    $ans \leftarrow g_0/B$ 
10:  if uniform(0, 1).sample() <  $r_0 - r'_0$  then
11:     $r'_0 \leftarrow r'_0 + 1$ 
12:  end if
13:  for  $i = 0$  to  $r'_0$  do
14:     $ans \leftarrow ans + g * \tilde{I}'(\bar{z} + x_0)/r_0$ 
15:  end for
16:  return  $ans$ 
17: end procedure

```

To facilitate our derivation, we present another estimator I corresponding to \tilde{I}_{sign} in the text. This estimator is almost identical to \tilde{I} , except for the different value added to ans . It's feasible to discuss I although our real estimator is \tilde{I} , because:

- (1) Our derivation for I also works for \tilde{I} with an additional approximation.
- (2) \tilde{I} is always better than I , thus decreasing the I 's variance will also decrease an upper bound for \tilde{I} 's variance. We will show this in Appendix G.

In the following steps, we will first discuss the convergence condition, expectation, and variance of I , then discuss the optimal RRS function $r(x)$ and B . Finally, we will show \tilde{I} is always better than I , i.e., has less variance.

Algorithm 2 Estimator I

```

1: procedure I
2:    $x_0 \leftarrow p.sample()$ 
3:    $g_0 \leftarrow 1 - f(x_0)/(B * p(x_0))$ 
4:    $r_0 \leftarrow r(x)$ 
5:    $r'_0 \leftarrow \lfloor r_0 \rfloor$ 
6:    $ans \leftarrow 1/B$ 
7:   if uniform(0, 1).sample() <  $r_0 - r'_0$  then
8:      $r'_0 \leftarrow r'_0 + 1$ 
9:   end if
10:  for  $i = 0$  to  $r'_0$  do
11:     $ans \leftarrow ans + g * I() / r_0$ 
12:  end for
13:  return  $ans$ 
14: end procedure

```

B CONVERGENCE CONDITION

Let $g(x) = 1 - \frac{f(x)}{Bp(x)}$ and $\{I_i\}(i \in \mathbb{N})$ be a sequence of independently and identically distributed random variables with respect to I , we

have

$$I = \frac{1}{B} + \frac{g(X)}{r(X)} \sum_i^{r(X)} I_i. \quad (28)$$

Here $\sum_i^r I_i = \sum_i^{\lfloor r \rfloor} I_i + I^*$, with

$$I^* = \begin{cases} I_{\lceil r \rceil} & \text{probability } r - \lfloor r \rfloor \\ 0 & \text{probability } \lceil r \rceil - r \end{cases}$$

As mentioned in the text, our estimator evaluates a Taylor expansion, namely

$$\begin{aligned} \frac{1}{\int_R f(x)dx} &= \frac{1}{B} \frac{1}{1 - E[g(X)]} \\ &= \frac{1}{B} (1 + E[g(X)] + E^2[g(X)] + \dots). \end{aligned}$$

According to the convergence domain of the Taylor expansion, I converges when $E[g(X)] \in (-1, 1)$. Since $E[g(X)] = 1 - \frac{1}{B} \int_R f(x)dx$, the convergence condition for I is

$$\frac{1}{B} \int_R f(x)dx \in (0, 2). \quad (29)$$

When $E[g(x)] \in (-1, 1)$, $E[I]$ is convergent. However, $E[I^2]$ or $V[I]$ are harder to converge and need a stricter condition. We will next prove that in order to let $E[I^2]$ converge, we at least require

$$E[|g(X)|] < 1. \quad (30)$$

We will prove this by contradiction. If $E[|g(x)|] \geq 1$ and $E[I^2]$ still converges, it holds that

$$\begin{aligned} E[(I - \frac{1}{B})^2] &= E[\frac{g^2(X)}{r^2(X)} E[(\sum_i^{r(X)} I_i)^2]] \\ &\geq E[\frac{g^2(X)}{r^2(X)} r(X) E[I^2]] \\ &= E[\frac{g^2(X)}{r(X)}] E[I^2] \\ &\geq E[\frac{g^2(X)}{r(X)}] E[(I - \frac{1}{B})^2]. \end{aligned}$$

Here $E[I^2] \geq E[(I - \frac{1}{B})^2]$, because

$$E[(I - \frac{1}{B})^2] - E[I^2] = \frac{1}{B} (\frac{1}{B} - \frac{2}{\int_R f(x)dx}) \leq 0,$$

which is assured by the condition (Equation 29).

Now we have

$$E[(I - \frac{1}{B})^2] \geq E[\frac{g^2(X)}{r(X)}] E[(I - \frac{1}{B})^2]. \quad (31)$$

Since $E[|g(X)|] \geq 1$, with Cauchy's inequality, it's easy to prove

$$E[\frac{g^2(X)}{r(X)}] > \frac{E^2[|g(X)|]}{E[r(X)]} > 1.$$

The equality can not hold because $E[r(x)] < 1$ is required for a legal RRS function. This raises the contradiction with (Equation 31) as $E[(I - \frac{1}{B})^2] > 0$. Therefore, we conclude that when $E[|g(x)|] \geq 1$, $E[I^2]$ is not convergent, and so is $V[I]$.

C EXPECTATION

The expectation is as:

$$\begin{aligned} E[I] &= \frac{1}{B} + E\left[\frac{g(X)}{r(X)} \sum_i^{r(X)} I_i\right] \\ &= \frac{1}{B} + E[g(X)]E(I) \\ &= \frac{1}{B(1 - E[g(X)])} \\ &= \frac{1}{\int_R f(x)dx}. \end{aligned}$$

Therefore, we have

$$E[I] = \frac{1}{\int_R f(x)dx}. \quad (32)$$

D VARIANCE

Recalling previously defined $\sum_i^r I_i$, when r is an integer, it's easy to derive

$$E\left[\left(\sum_i^r I_i\right)^2\right] = rE[I^2] + (r^2 - r)E^2[I]. \quad (33)$$

We adopt an approximation² and use this result when r is a positive real number. With this approximation, we derive the variance of I :

$$\begin{aligned} V[I] &= V[I - \frac{1}{B}] \\ &= E[(I - \frac{1}{B})^2] - E^2[I - \frac{1}{B}] \\ &= E\left[\frac{g^2(X)}{r^2(X)} \left(\sum_i^{r(X)} I_i\right)^2\right] - (E[I] - \frac{1}{B})^2 \\ &\approx E\left[\frac{g^2(X)}{r(X)} (E[I^2] + (r(X) - 1)E^2[I])\right] - (E[I] - \frac{1}{B})^2 \\ &= E\left[\frac{g^2(X)}{r(X)}\right]V[I] + E[g^2(X)]E^2[I] - (E[I] - \frac{1}{B})^2. \end{aligned}$$

Let $c_0 = E[g^2(X)]E^2[I] - (E[I] - \frac{1}{B})^2$, note that c_0 is irrelevant with $r(x)$, then

$$V[I] = E\left[\frac{g^2(X)}{r(X)}\right]V[I] + c_0.$$

Therefore, with approximation (Equation 33), we have

$$V[I] = \frac{c_0}{1 - E\left[\frac{g^2(X)}{r(X)}\right]}. \quad (34)$$

E OPTIMAL FORMULATION OF R(X)

Next, we will prove that in order to achieve max efficiency, the optimal formulation of $r(x) \equiv |g(x)|$. We first define efficiency as the product of cost and variance, namely

$$Eff = \frac{1}{C \times V[I]}. \quad (35)$$

²This approximation is also adopted in Rath et al. [2022b]

In a single iteration of I , the expectation of the number of new iterations is $E[r(X)]$. Therefore, $C = \frac{1}{1 - E[r(X)]}$, which is the expected number of runs of I . Obviously, we need C to be a finite positive number, so $E[r(x)] \in (0, 1)$ is required. Combining the previously obtained variance of I , we have

$$\begin{aligned} Eff &= \frac{1}{c_0} (1 - E[r(X)]) (1 - E\left[\frac{g^2(X)}{r(X)}\right]) \\ &= \frac{1}{c_0} (1 - \int_R r(x)p(x)dx) (1 - \int_R \frac{g^2(x)}{r(x)} p(x)dx). \end{aligned}$$

From the integral form of Aczél's inequality³, we derive

$$Eff \leq \frac{1}{c_0} (1 - \int_R |g(x)|p(x)dx)^2. \quad (36)$$

The equality holds if and only if $r(x) \equiv |g(x)|$.

In order to achieve this optimality, $E[|g(X)|] \in (0, 1)$ is required. This is not a restriction as we have proved in Appendix B that $E[|g(x)|] < 1$ is required for $V[I]$'s convergence.

F OPTIMAL VALUE FOR B

We now assume $E[|g(X)|] \in (0, 1)$ and $r(x)$ takes the optimal form $|g(x)|$, the following discussion explores the optimal value for B .

Since $r(x) \equiv |g(x)|$, now we have

$$\begin{aligned} Eff &= \frac{1}{c_0} (1 - \int_R |g(x)|p(x)dx)^2 \\ &= \frac{(1 - E[|g(X)|])^2}{E[g^2(X)]E^2[I] - (E[I] - \frac{1}{B})^2} \\ &= \frac{(1 - E[|g(X)|])^2}{E[g^2(X)]E^2[I] - E^2[g(X)]E^2[I]} \\ &= \frac{(1 - E[|g(X)|])^2}{V[g(X)]} \times \frac{1}{E^2[I]}. \end{aligned}$$

In the derivation above, we used a relation between $E[g(X)]$ and $E[I]$: $E[g(X)]E[I] = E[I] - \frac{1}{B}$. It's easy to get this relation since $E[g(X)] = 1 - \frac{1}{BE[I]}$.

Note that $g(x) = 1 - \frac{f(x)}{Bp(x)}$, therefore $V[g(X)] = V[g(X) - 1] = \frac{1}{B^2}V\left[\frac{f(X)}{p(X)}\right]$, thus

$$Eff = (B - B \times E[|g(X)|])^2 \times \frac{1}{E^2[I]V\left[\frac{f(X)}{p(X)}\right]}.$$

Let $c_1 = \frac{1}{E^2[I]V\left[\frac{f(X)}{p(X)}\right]}$, note that c_1 is irrelevant with B , then

$$\begin{aligned} Eff &= c_1(B - \int_R |B - \frac{f(x)}{p(x)}|p(x)dx)^2 \\ &= c_1(\int_R (B - |B - \frac{f(x)}{p(x)}|)p(x)dx)^2 \\ &\leq c_1(\int_R f(x)dx)^2. \end{aligned}$$

³The Aczél's inequality states that if $a_i, b_i (i = 1, 2, \dots, n)$ are positive numbers such that $a_1^2 - \sum_{i=2}^n a_i^2 > 0$ or $b_1^2 - \sum_{i=2}^n b_i^2 > 0$, then $(a_1^2 - \sum_{i=2}^n a_i^2)(b_1^2 - \sum_{i=2}^n b_i^2) \leq (a_1 b_1 - \sum_{i=2}^n a_i b_i)^2$. The equality holds only if $\frac{a_i}{b_i} = \frac{a_1}{b_1}$ for any i . More details in Wu [2008].

The equality condition is $B \geq \max\{\frac{f(x)}{p(x)}\}$. We assume $p(x)$ to be a feasible sampling distribution for $f(x)$, i.e. $\max\{\frac{f(x)}{p(x)}\} < +\infty$ so the equality condition can be achieved.

As B becomes greater than $\max\{\frac{f(x)}{p(x)}\}$, the efficiency remains constant. However, this conclusion was derived under approximation (Equation 33). When we assume $B \geq \max\{\frac{f(x)}{p(x)}\}$, I degenerates into a simple form that is similar to the geometric distribution, and the computation of $V[I]$ no longer needs approximation. We review our derivation under this context.

With $B \geq \max\{\frac{f(x)}{p(x)}\}$, $g(x) = 1 - \frac{f(x)}{Bp(x)} \in (0, 1)$, I degenerates into a simple form:

$$I = \frac{1}{B} + \begin{cases} I & \text{probability } g(X) \\ 0 & \text{probability } 1 - g(X) \end{cases}$$

Let $q = \int_R g(x)p(x)dx$, which assembles the continue probability in geometric distribution, then

$$C = \frac{1}{1-q}, V[I] = \frac{q}{(B-Bq)^2} = \frac{q}{E^2[I]}.$$

We have

$$Eff = \frac{1}{C \times V[I]} = (\frac{1}{q} - 1)E^2[I].$$

Therefore, as q decreases, the efficiency increases. However, q is a monotonically increasing function of B . Taking into account the condition $B \geq \max\{\frac{f(x)}{p(x)}\}$, we consider $B = \max\{\frac{f(x)}{p(x)}\}$ as the optimal value.

G COMPARISON BETWEEN I AND \tilde{I}

While we have been using I for the proof previously, our derivation also apply to \tilde{I} using an additional approximation, namely

$$\begin{aligned} V[\tilde{I}] &= V\left[\frac{g(X)}{B} + \frac{g(X)}{r(X)} \sum_i \tilde{I}_i\right] \\ &\approx V\left[\frac{g(X)}{B}\right] + V\left[\frac{g(X)}{r(X)} \sum_i \tilde{I}_i\right]. \end{aligned}$$

which assumes the first part and the second part in RHS is irrelevant. With this approximation, the previous conclusion $r(x) \equiv g(x)$ and $B \geq \max\{\frac{f(x)}{p(x)}\}$ still holds true, but the last part in Appendix F that removes approximation (Equation 33) can not be easily applied to \tilde{I} , as $V[\tilde{I}]$ becomes a complex function of B and simple monotonicity analysis is no longer feasible. But we will prove $B = \max\{\frac{f(x)}{p(x)}\}$ is still sub-optimal for \tilde{I} in the end.

We first show \tilde{I} 's variance is less than I . Recall that

$$I = \frac{1}{B} + \frac{g(X)}{r(X)} \sum_i I_i$$

and

$$\tilde{I}' = \frac{g(X)}{B} + \frac{g(X)}{r(X)} \sum_i \tilde{I}'_i.$$

Let $I' = I - \frac{1}{B}$, we have

$$I' = \frac{g(X)}{r(X)} \sum_i (I'_i + \frac{1}{B}).$$

Note that $V[I] = V[I']$ and $V[\tilde{I}] = V[\tilde{I}']$, so we can just compare $V[I']$ and $V[\tilde{I}']$. In addition, with $E[I'] = E[\tilde{I}'] = \frac{1}{\int_R f(x)dx} - \frac{1}{B}$, the comparison between $V[I]$ and $V[\tilde{I}]$ eventually turns into the comparison between $E[I'^2]$ and $E[\tilde{I}'^2]$:

For I ,

$$E[\tilde{I}'^2] = E\left[\frac{g^2(X)}{r^2(X)}\left(\frac{r(X)}{B} + \sum_i \tilde{I}_i\right)^2\right]$$

and for I' ,

$$\begin{aligned} E[I'^2] &= E\left[\frac{g^2(X)}{r^2(X)}\left(\sum_i (I_i + \frac{1}{B})\right)^2\right] \\ &= E\left[\frac{g^2(X)}{r^2(X)}\left(\frac{\lfloor r(X) \rfloor}{B} + \frac{\overline{r(X)}}{B} + \sum_i I_i\right)^2\right] \end{aligned}$$

where

$$\overline{r} = \begin{cases} 1 & \text{probability } r - \lfloor r \rfloor \\ 0 & \text{probability } \lceil r \rceil - r \end{cases}$$

Let $\widetilde{Z}' = \frac{r}{B} + \sum_i \tilde{I}_i$ and $Z' = \frac{\lfloor r \rfloor}{B} + \frac{\overline{r}}{B} + \sum_i I_i$. When r is an integer, $E[Z'^2] = E[\widetilde{Z}'^2]$. Otherwise, $\overline{r} \neq 0$, we can prove $E[\widetilde{Z}'^2] < E[Z'^2]$, and $E[\tilde{I}'^2] < E[I'^2]$. Therefore, we eventually prove $V[\tilde{I}] \leq V[I]$, which means \tilde{I} is a better estimator.

As we have derived in Appendix F, $B = \max\{\frac{f(x)}{p(x)}\}$ is optimal for I . Although we only attain $B \geq \max\{\frac{f(x)}{p(x)}\}$ for \tilde{I} , since $V[I]$ is an upper bound for $V[\tilde{I}]$, $B = \max\{\frac{f(x)}{p(x)}\}$ can still be considered sub-optimal for \tilde{I} .