

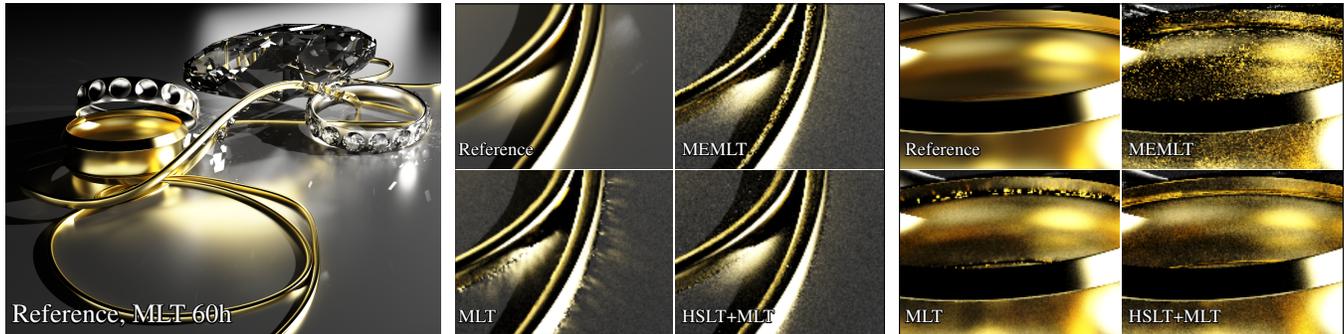
# The Natural-Constraint Representation of the Path Space for Efficient Light Transport Simulation

Anton S. Kaplanyan

Johannes Hanika

Carsten Dachsbacher

Karlsruhe Institute of Technology



**Figure 1:** Equal-time comparison of the JEWELRY scene with difficult glossy paths. Left: Reference computed with the original Metropolis Light Transport (MLT) in 60 hours. The two closeups show 10 minute renders using original MLT, Manifold Exploration (MEMLT) and MLT with our proposed mutation (HSLT+MLT). Note that BDPT cannot efficiently explore indirect caustics even at high sample counts. All MLT variants used 1 million initial samples, MEMLT used  $\lambda = 50$ . There was no special handling of direct illumination. The average number of samples per pixel: MEMLT 621, MLT 1424, HSLT+MLT 612. Our new mutation estimates the optimal sampling density of difficult specular and glossy transport and excels at capturing hard features like thin highlights along highly curved rims.

## Abstract

The path integral formulation of light transport is the basis for (Markov chain) Monte Carlo global illumination methods. In this paper we present *half vector space light transport (HSLT)*, a novel approach to sampling and integrating light transport paths on surfaces. The key is a partitioning of the path space into subspaces in which a path is represented by its start and end point constraints and a sequence of generalized half vectors. We show that this representation has several benefits. It enables importance sampling of all interactions along paths in between two endpoints. Based on this, we propose a new mutation strategy, to be used with Markov chain Monte Carlo methods such as Metropolis light transport (MLT), which is well-suited for all types of surface transport paths (diffuse/glossy/specular interaction). One important characteristic of our approach is that the Fourier-domain properties of the path integral can be easily estimated. These can be used to achieve optimal correlation of the samples due to well-chosen mutation step sizes, leading to more efficient exploration of light transport features. We also propose a novel approach to control stratification in MLT with our mutation strategy.

**CR Categories:** I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing

**Keywords:** global illumination, Metropolis light transport

**Links:** [DL](#) [PDF](#) [WEB](#) [CODE](#)

©ACM, 2014. This is the authors version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in ACM Transactions on Graphics, 33, 4, July 2014. doi.acm.org/10.1145/2601097.2601108

## 1 Introduction

Rendering photorealistic images has been one of the main goals of computer graphics. Physically-based simulation of light transport plays a crucial role in this endeavor. And indeed, we can observe significant progress in the large number of global illumination methods developed over time. However, the inherent complexity of light transport creates challenges that have not yet been fully mastered. For example, computing light transport for complex scenes with glossy materials remains nontrivial due to (nearly) non-samplable specular-diffuse-specular paths and difficult visibility.

There are two main strategies to address the challenges of light transport simulation. The first is to strive to improve the underlying mathematical integration method. Second, there is a demanding effort to provide proper level of detail (LoD) for scene description data, such as geometry, textures, and even acceleration structures. In this paper, we focus on the mathematical integration of the transport problem which is particularly challenging in scenes with complex illumination, geometry and materials.

The path integral formulation of light transport [Veach 1998] and numerical integration and sampling techniques, such as (Markov chain) Monte Carlo methods, form the basis of nowadays global illumination methods. In all their diversity, these methods share the concept of stochastically creating paths, connecting the sensor to the lights. With limited computational budgets, i.e., finite sample (path) count, the sampling strategy is of utmost importance. Metropolis light transport (MLT) [Veach and Guibas 1997] was the first application of Markov chain Monte Carlo (MCMC) sampling to this problem. This class of approaches samples the path space by mutating paths and concentrates on samples with high contributions to the image. MLT is able to locally explore regions of the path space that are otherwise difficult to sample, e.g. pure glossy interactions that are important for natural images.

In this paper we show how light transport on surfaces (in the absence of participating media) can be solved in a domain that is better suited for these difficult phenomena. Consider for example a perfect specular reflection: in this case the trajectory of light is dic-

tated by Fermat’s fundamental principle, intuitively the angle of incidence is equal to the angle of reflection. The configuration can likewise be expressed using the half vector between these two directions which then coincides with the surface normal. This geometric configuration is always met for specular reflections and we thus call it a *natural constraint*.<sup>1</sup>

These constraints have long been used in optical design, e.g., in ABCD matrix analysis or transfer/system matrices, and also were applied to image formation with pencil tracing [Shinya et al. 1987]. Recently, these geometric considerations of Fermat’s principle led to an improved sampling technique for specular interactions with manifold exploration [Jakob 2013] in MLT.

In light transport simulation, a path is typically represented as a list of vertices (locations of interactions), or by its start vertex and a sequence of directions. In this paper, we propose a different representation, where the relation of incident and outgoing directions at interactions is specified by the direction of the (generalized) half vector [Sommerfeld and Runge 1911]. This representation is a local parametrization and unique only within a subspace of the path space (intuitively, paths via the same surface patches, Fig. 2).

For light transport simulation, we need to sample the entire path space. However, this representation, which we detail further in Sec. 4, has a number of beneficial properties when considering light transport *within* each subspace:

- many practical bidirectional scattering distribution functions (BSDFs) can be naturally [Walter et al. 2007] or more compactly [Rusinkiewicz 1998] expressed with half vectors,
- we show that it enables importance sampling both geometry (e.g. curvature) and BSDFs at all interactions along paths,
- we can directly generate valid paths by just perturbing the half vectors (Fig. 2), even with glossy and specular BSDFs, curved surfaces and small light sources.

In addition to this representation, our main contributions are:

- a novel global illumination method using this representation to robustly handle complex *inter-surface* light transport,
- a symbolic and numerical analysis of the transformation between our representation and standard measurement spaces previously used for integration in the path space (Sec. 4),
- a new mutation strategy for bidirectional light transport that operates on entire paths without explicit connections (Sec. 5),
- a novel stratification scheme for MCMC rendering (Sec. 6),
- and we show how to employ recent results on the Fourier analysis of Monte Carlo simulation [Subr and Kautz 2013] to further improve sampling (Sec. 6.3).

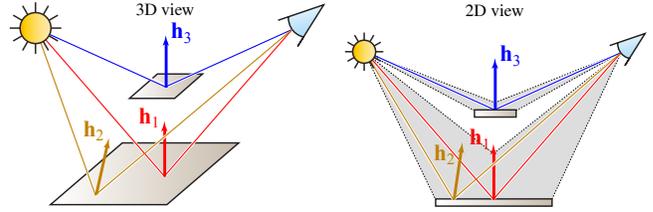
## 2 Background and Previous Work

**Path Integral** Light transport methods compute a solution to the path integral defined for the  $j$ -th pixel measurement  $I_j$  over *measurement contribution function*  $f_j(\mathbf{X})$  [Veach 1998, p.223] on the product surface area measure  $d\mathbf{X}$  as

$$I_j = \int_{\Omega} f_j(\mathbf{X}) d\mathbf{X}, \quad \text{with } d\mathbf{X} \equiv d\mu(\mathbf{X}) = \prod_{i=0}^k dA(\mathbf{x}_i), \quad (1)$$

where  $\Omega$  denotes the path space, i.e., the space of all valid paths  $\mathbf{X} = (\mathbf{x}_0, \dots, \mathbf{x}_k)$  of all lengths  $k = (1..∞)$  that connect a light

<sup>1</sup>Snell’s law is a special case of Fermat’s principle for refractions where the *generalized* half vector (see appendix) aligns with the surface normal.



**Figure 2:** The new representation of light transport paths partitions the path space into submanifolds in which a path can uniquely be represented by its natural constraints (see Sec. 5.2 for a formal definition). Here two submanifolds (right, gray) with paths of length 2 via one of the surfaces are shown:  $\mathbf{h}_1$  and  $\mathbf{h}_2$  specify two paths in the same subspace, while the red and blue paths have the same half vector  $\mathbf{h}_1 = \mathbf{h}_3$ , but are in different subspaces. Our method enables to easily sample these submanifolds.

source to the sensor via interactions at vertex positions  $\mathbf{x}_i$ . Hereafter we omit the dependence of  $f(\mathbf{X})$  on the pixel index  $j$ , as the pixel filter can be applied at the sample accumulation stage. We denote the corresponding outgoing and incident directions from  $\mathbf{x}_i$  to  $\mathbf{x}_{i+1}$  with  $\mathbf{o}_i = -\mathbf{i}_{i+1}$ , the projected solid angle measure  $d\sigma^+(\mathbf{o}) \equiv d\mathbf{o}^\perp$  and the surface area measure  $d\mathbf{x} \equiv dA(\mathbf{x})$ . Later, we will also consider half vectors and projected half vectors which are denoted as  $\mathbf{h}_i$  and  $\mathbf{h}_i^\perp$ , respectively. Table 1 and Fig. 3 summarize our notation.

**Monte Carlo Methods** Stochastic path generation in MC methods can either start from the camera (path tracing [Kajiya 1986]), from the light sources (light tracing [Arvo 1986]), or from both sides with bidirectional path tracing (BDPT) [Veach and Guibas 1994; Lafortune and Willems 1993] involving deterministic connections of subpaths. Many-lights methods [Keller 1997] reuse light subpaths for connections by interpreting their vertices as virtual light sources illuminating the scene. Often these methods employ clustering and importance sampling (e.g. [Walter et al. 2005]), and multiple importance sampling [Walter et al. 2012]. We refer the reader to recent surveys on this topic [Dachsbacher et al. 2013; Krivánek et al. 2013]. Bidirectional methods are based on deterministic connection of path vertices, which implies that bidirectional scattering distribution functions (BSDF) at two connected vertices cannot be importance-sampled.

**Photon Mapping** Another popular family of Monte Carlo methods are derived from photon mapping [Shirley et al. 1995; Jensen 1996] where a large number of light subpaths are generated and the vertices stored as photons. Light subpaths are then connected to camera subpaths using kernel estimation, trading variance for bias. These methods recently gained more attention since the advent of progressive methods [Hachisuka et al. 2008; Knaus and Zwicker 2011; Kaplanyan and Dachsbacher 2013a] and unified frameworks coupling photon mapping and unbiased Monte Carlo methods [Georgiev et al. 2012; Hachisuka et al. 2012]. Compared to pure (MC)MC methods, these approaches introduce visible bias and still need significant amounts of memory for storing photons, even though the progressive methods partly alleviate this problem.

**Markov Chain Monte Carlo Methods** The Metropolis light transport (MLT) method [Veach and Guibas 1997] employs Markov chain Monte Carlo (MCMC) sampling (using the Metropolis-Hastings algorithm [Metropolis et al. 1953; Hastings 1970]) for computing the path integral by iteratively sampling new paths  $\mathbf{X}_{t+1}$  from a current path  $\mathbf{X}_t$ . The process is bootstrapped with an ini-

tial path  $\mathbf{X}_0$  obtained from bidirectional path tracing and subsequently the path space is sampled using MCMC with specifically designed mutation strategies. For the Metropolis-Hastings algorithm these strategies need to compute the transition probability  $T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})$  for a state change to happen. To ensure that the Markov chain converges to the intended stationary distribution, a newly proposed mutation is accepted only with a certain *acceptance probability*  $\min(1, a)$  with

$$a = \frac{f(\mathbf{X}_{t+1})/T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})}{f(\mathbf{X}_t)/T(\mathbf{X}_{t+1} \rightarrow \mathbf{X}_t)} := \frac{R_{t+1}}{R_t}. \quad (2)$$

MLT shares path samples  $\mathbf{X}_i$  among all pixels, i.e., for estimating all pixel integrals at once, effectively computing an image histogram. For this, each sample is projected to the image plane and its contribution is accumulated accordingly.

Several variants of MLT and applications of MCMC sampling techniques have been proposed in computer graphics literature. Kelemen et al. [2002] perform mutations in primary sample space, i.e., on random number vectors that are used to generate the paths. As a consequence, the performance of this method also depends on the path sampling technique (how random numbers map to paths).

MCMC sampling has also been used with photon mapping where subpaths are mutated [Fan et al. 2005; Chen et al. 2011; Hachisuka and Jensen 2011]. In contrast, we importance-sample all interactions of a path at once, which is not the case when a path is connected to a photon location where the BSDF has to be evaluated for already given directions.

Energy redistribution path tracing (ERPT) [Cline et al. 2005] uses the original set of MLT mutations [Veach and Guibas 1997], but a large number of short Markov chains to stratify the image plane sampling. In contrast, our method employs explicit stratification within the chain by computing the necessary changes to path configuration to distribute samples evenly. This makes explicit filtering, as in the original ERPT paper [Cline et al. 2005], unnecessary.

Also, there exist various other extensions of MLT which do not directly change path sampling and can thus be used together with our method, such as Population Monte Carlo [Cappé et al. 2004] and gradient domain MLT [Lehtinen et al. 2013]. The latter computes image gradients plus a coarse approximation of an image. The final image is then obtained thereof by solving a Poisson equation.

**Manifold Exploration** Most related to our work is manifold exploration (ME) [Jakob and Marschner 2012; Jakob 2013], a mutation strategy for MLT specifically designed to improve the sampling of specular and highly glossy paths. It is based on the observation that a specular vertex entails a Dirac delta in the integrand (due to the BSDF), which effectively removes this dimension from the integration. This essentially reduces the domain of integration to a lower-dimensional sub-manifold enclosed in path space. Jakob and Marschner [2012] propose to implement this as a deterministic *manifold walk* procedure which enforces Fermat’s principle as a constraint on specular vertices, after performing a perturbation on any of the other vertices. To render a caustic for example, the algorithm performs a *lens perturbation* [Veach and Guibas 1997] by tracing from the sensor towards a perturbed direction until reaching the caustic. The rest of the path up to the light source is then updated respecting Snell’s law for refraction at the interfaces. While producing excellent results for caustics caused by specular interactions, this technique can be suboptimal for highly glossy transport. We discuss this method in more detail in Sec. 5.1. Note that unbiased (MC)MC methods, including manifold exploration, cannot sample all transport paths, which requires regularization [Kaplanian and Dachsbacher 2013b].

Term	Description
$\mathbf{v}, \mathbf{v}^\perp$	Vector and its planar projection onto tangent frame
$\ \mathbf{v}\ , \bar{\mathbf{v}}$	$\mathcal{L}_2$ length of a vector, unit (normalized) vector
$d\mathbf{v}, \Delta\mathbf{v}$	Differential of vector, finite difference
$\mathbf{X}$	Ordered set of points, also light transport path
$M,  M $	Matrix $M$ , its determinant
$\left  \frac{d(\mathbf{b}_0 \dots \mathbf{b}_n)}{d(\mathbf{a}_0 \dots \mathbf{a}_n)} \right $	Jacobian determinant of transformation from space $\mathbf{A}$ to $\mathbf{B}$
$\mathbf{x}, \mathbf{i}, \mathbf{o}$	Surface point, incident and outgoing direction
$(\mathbf{u}, \mathbf{v}, \mathbf{n})$	Tangent frame at surface point
$\mathbf{h}, \mathbf{h}^\perp$	Half vector and projected half vector
$f(\mathbf{X})$	Measurement contribution function for path $\mathbf{X}$
$f_s(\mathbf{i}_t \rightarrow \mathbf{o}_t) = f_s(\mathbf{x}_t)$	Bidirectional scattering distribution function (BSDF) at $\mathbf{x}_t$

Table 1: Notation used throughout our paper.

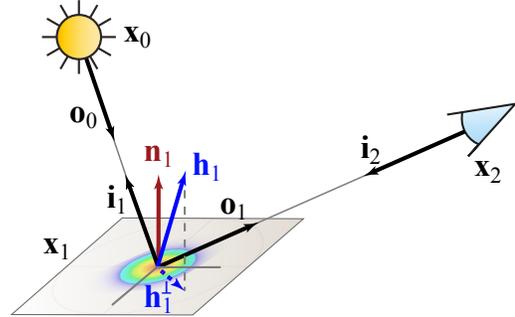


Figure 3: Illustration of our notation on a simple path. The vertex indices are starting from the light. Please also see Table 1 for an overview of the notation.

### 3 Overview

We first introduce the space of natural constraints and its relation to the standard measurement spaces otherwise used in global illumination (Sec. 4). Along this way, we analyze the transformation between the spaces and discuss important special cases. Throughout the paper, we ignore visibility and assume the BSDFs are well-approximated in half vector domain (leaving out some parameters, like Fresnel terms).

That followed, we introduce the basis for our new mutation strategy which makes use of the new sampling domain and enables its use with Metropolis light transport methods (Sec. 5). This mutation strategy perturbs half vectors along the path individually while keeping the first point on the light source and the last point on the sensor. That is, it explores the subspace containing the current path (see Fig. 2) by constraint mutation.

In Sec. 6 we detail how this mutation strategy can be controlled by balancing the three aspects of importance sampling: exploring the BSDFs, enforcing stratification on the image plane, and sampling according to a frequency analysis. We show that these goals can be conveniently achieved at the same time with our theory.

## 4 Integration in Space of Natural Constraints

### 4.1 Integration in Path Space

In order to compute the path integral (Eq. 1) we have to generate paths to sample the path space  $\Omega$ , and compute the paths’ flux densities, i.e., the contributions to the image. However, this path sampling is typically not done in the measurement space, but in the *sampling domain of outgoing directions* (or projected solid angle)

where Monte Carlo methods operate: eye subpaths and/or light subpaths are generated by sampling a direction at an interaction location (and possibly these sub-paths are deterministically connected to form complete paths). To compute the flux densities of paths we still need to perform a change of variables from the sampling domain of directions at a vertex  $\mathbf{x}_i$  to the surface area measure at the next vertex  $d\mathbf{o}_i^\perp \mapsto d\mathbf{x}_{i+1}$  which is represented by the following Jacobian (called *geometric term*):

$$G(\mathbf{x}_i, \mathbf{x}_{i+1}) = \left| \frac{d\mathbf{o}_i^\perp}{d\mathbf{x}_{i+1}} \right| = \left| \frac{d\mathbf{o}_i \langle \mathbf{o}_i, \mathbf{n}_i \rangle}{d\mathbf{x}_{i+1}} \right| = \frac{|\langle \mathbf{i}_{i+1}, \mathbf{n}_{i+1} \rangle \langle \mathbf{o}_i, \mathbf{n}_i \rangle|}{\|\mathbf{x}_i - \mathbf{x}_{i+1}\|^2}.$$

Note that these geometric terms only appear explicitly in next event estimation in path tracing or in the subpath connection in BDPT as they otherwise cancel out when sampling outgoing directions.

Unfortunately, these geometric terms can introduce difficulties in radiance measurements as they are not always well-behaved on geometric junctions, e.g. at corners. Nevertheless, the measurement space is the only space for a comparison of the flux density and the sampling probability of two paths and thus a very important space for techniques that require such comparisons, e.g. multiple importance sampling and Metropolis light transport [Veach 1998].

On the other hand, the sampling domain of outgoing directions is best-suited for sampling scattering interactions. Note, a path is not explicitly represented in this space (e.g. as an ordered set of vertices) and has to be raytraced to determine the respective next scattering locations. As a consequence, altering directions can lead to changes in the path topology, e.g., when hitting a different surface, or a drastic change of an endpoint position. This makes it an inconvenient space for directly working on paths for point-to-point transport (connection from a light source to the sensor), since the trajectory cannot be easily controlled.

Although paths are easy to sample in this domain, it shares one limitation with most path sampling techniques: at some step a path is completed, i.e., it forms a point-to-point connection, and then the values of some BSDFs have to be evaluated post factum, which can lead to unpredictable changes of flux carried by such a path. That is, the goal is to importance-sample a path with all BSDF interactions.

## 4.2 The Space of Natural Constraints

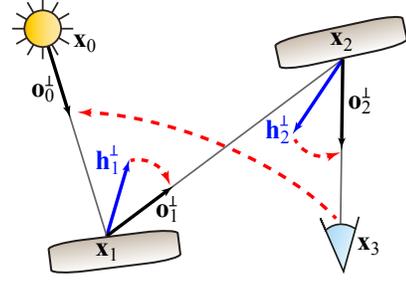
For light transport simulation we need to (1) sample the path, and (2) compute the measurement contribution of the path which, for the domain of outgoing directions, potentially involves a change of variables introducing singularities. We propose to use a different space for sampling which does not suffer from the aforementioned singularities and further has the desirable property that it works on entire paths, enabling us to importance-sample all interactions along paths in between two endpoints (e.g., a pixel on the sensor and a point on a light source), that is, we keep  $\mathbf{x}_0$  and  $\mathbf{x}_k$  fixed.

This *space of natural constraints* is based on the observation that a complete path—in a subspace of the path space—can be represented by an ordered set of vertices,  $\mathbf{X} = (\mathbf{x}_0, \dots, \mathbf{x}_k)$ , but it can also be uniquely represented using the two endpoints  $\mathbf{x}_0$  and  $\mathbf{x}_k$  and a sequence of projected half vectors at the interactions (see Fig. 3). More precisely we introduce the *space of natural constraints*, or the *projected half vector domain*, as

$$\mathbf{H}^\perp = (\mathbf{x}_0, \mathbf{h}_1^\perp, \mathbf{h}_2^\perp, \dots, \mathbf{h}_{k-1}^\perp, \mathbf{x}_k) \in \Omega(\mathbf{H}^\perp) \subset \Omega.$$

The term half vector domain is more intuitive when considering a single constraint (half vector). We use the terms interchangeably.

Our sampling domain has two key properties that we will introduce step-by-step in the next sections:



**Figure 4:** The simplified product of Jacobians first transforms the displacement at the last vertex  $\mathbf{x}_3$  to the outgoing direction at the first one ( $\mathbf{x}_0$ ), allowing to track the changes in flux density throughout the path. Then a transformation from projected half vectors to projected outgoing direction at every scattering vertex is applied.

- the change of variables, related to the sampling in the space of natural constraints, entails that the integrand becomes flatter and decomposed into highly decorrelated subintegrals,
- even more importantly, the domain of constraints enables us to importance-sample an entire subspace  $\mathbf{H}^\perp$ , i.e., all paths of the same topology in the same submanifold.

## 4.3 Analysis of the New Domain

We are concerned with the integration over transport paths to compute the incident flux on a sensor pixel due to emission at a light source (point-to-point transport). Next, we introduce the transformation of such complete paths sampled in the half vector domain to the measurement space and analyze the properties of this mapping.

The transformation is obtained from the Jacobian of the half vector constraint matrix (see appendix or [Jakob 2013]) from projected half vector measure  $d\mathbf{H}^\perp$  into surface area measure  $d\mathbf{X}$ :

$$\int_{\Omega(\mathbf{X}_0)} f(\mathbf{X}) d\mathbf{X} = \int_{\Omega(\mathbf{H}_0^\perp)} f(\mathbf{X}) \left| \frac{d\mathbf{X}}{d\mathbf{H}^\perp} \right| d\mathbf{H}^\perp, \quad (3)$$

where  $\Omega(\mathbf{X}_0)$  and  $\Omega(\mathbf{H}_0^\perp)$  denote the same subspace of the path space where the local parameterization of the submanifold around the path  $\mathbf{X}_0$  is valid. This validity means that the transformation Jacobian is smooth on the submanifold, which in practice requires a region of  $C^2$  continuity of the surfaces around the vertices. We will detail in Sec. 5 how this requirement is practically met.

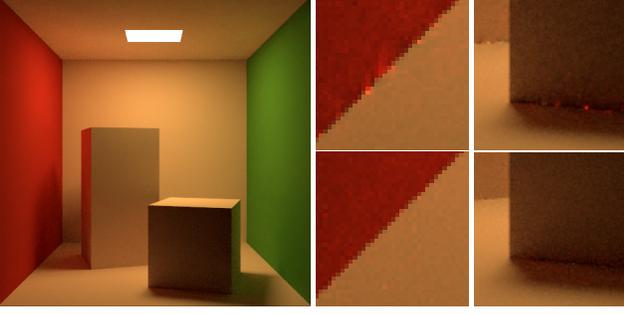
The complete measurement contribution then includes both Jacobians, the Jacobian of projected half vectors  $d\mathbf{x}_i \mapsto d\mathbf{h}_i^\perp$  and the product of geometric terms  $d\mathbf{o}_{i-1}^\perp \mapsto d\mathbf{x}_i$ :

$$f(\mathbf{X}) \left| \frac{d\mathbf{X}}{d\mathbf{H}^\perp} \right| = \rho(\mathbf{X}) \left( \prod_{i=0}^{k-1} G(\mathbf{x}_i, \mathbf{x}_{i+1}) \right) \left| \frac{d(\mathbf{x}_1 \dots \mathbf{x}_{k-1})}{d(\mathbf{h}_1^\perp \dots \mathbf{h}_{k-1}^\perp)} \right|. \quad (4)$$

Here  $\rho(\mathbf{X})$  denotes the product of BSDFs  $f_s$  at each surface vertex, the sensor responsivity  $W_c$  at the end point  $\mathbf{x}_k$ , and emitter radiance  $L_e$  at the start point  $\mathbf{x}_0$ . These components of  $\rho(\mathbf{X})$  are straightforward to evaluate and common to all global illumination methods. We will thus focus on the product of the two Jacobians.

### 4.3.1 Properties of the Product of Jacobians

In this section, we analyze the most common case of point-to-point transport, where constraints are two endpoints and the projected half vectors (on-surface microfacets) for interactions in between.



**Figure 5:** The CORNELL BOX (256 samples/pixel) with diffuse surfaces rendered using MLT with the original set of mutations (equivalent to manifold exploration for diffuse surfaces) in on-surface area measure (closeups top row), and with our method (bottom row). Note how the singularities in the corners cancel out thanks to the well-behaved product of Jacobians with the half vector domain.

For a given complete path  $\mathbf{x}_0, \dots, \mathbf{x}_k$  we analyze properties of the transformation in Eq. 4:

$$\begin{aligned}
 (\text{Eq. 4})/\rho(\mathbf{X}) &= \left| \frac{d(\mathbf{o}_0^\perp \dots \mathbf{o}_{k-1}^\perp)}{d(\mathbf{x}_1 \dots \mathbf{x}_k)} \right| \left| \frac{d(\mathbf{x}_1 \dots \mathbf{x}_{k-1})}{d(\mathbf{h}_1^\perp \dots \mathbf{h}_{k-1}^\perp)} \right| \\
 &= \left| \frac{d(\mathbf{o}_0^\perp, \mathbf{o}_1^\perp \dots \mathbf{o}_{k-1}^\perp)}{d(\mathbf{x}_k, \mathbf{x}_1 \dots \mathbf{x}_{k-1})} \right| \left| \frac{d(\mathbf{x}_1 \dots \mathbf{x}_{k-1})}{d(\mathbf{h}_1^\perp \dots \mathbf{h}_{k-1}^\perp)} \right| \\
 &= \left| \frac{d\mathbf{o}_0^\perp}{d\mathbf{x}_k} \right| \prod_{i=1}^{k-1} \left| \frac{d\mathbf{o}_i}{d\mathbf{h}_i} \right| \left| \frac{\langle \mathbf{o}_i, \mathbf{n}_i \rangle}{\langle \mathbf{h}_i, \mathbf{n}_i \rangle} \right|, \quad (5)
 \end{aligned}$$

where we first swapped matrix columns (property of the Jacobian determinant) and then applied the chain rule in this rearrangement. The first term of Eq. 5 can be computed using the method of *transfer matrices* [Shinya et al. 1987] that has recently been used as *generalized geometric term* for perfect specular chains [Jakob 2013]:

$$\left| \frac{d\mathbf{o}_0^\perp}{d\mathbf{x}_k} \right| = \left| \frac{d\mathbf{o}_0^\perp}{d\mathbf{x}_1} \right| \left| \frac{d\mathbf{x}_1}{d\mathbf{x}_k} \right| = G(\mathbf{x}_0, \mathbf{x}_1) |T_1|. \quad (6)$$

The transfer matrix  $T_1$  contains the differential transformation of an on-surface tangent patch at  $\mathbf{x}_1$  to  $\mathbf{x}_k$  based on the differential geometry. That is, it tracks the changes of area density along the light trajectory from the first on-surface vertex after the light source  $\mathbf{x}_1$  to the last vertex  $\mathbf{x}_k$  on the camera sensor.

Thus, Eq. 5 can intuitively be viewed as first transforming the area density as if considering a specular chain using transfer matrices (Fig. 4). Thereafter the individual transformations from half vectors to outgoing directions are applied to account for glossy transport.

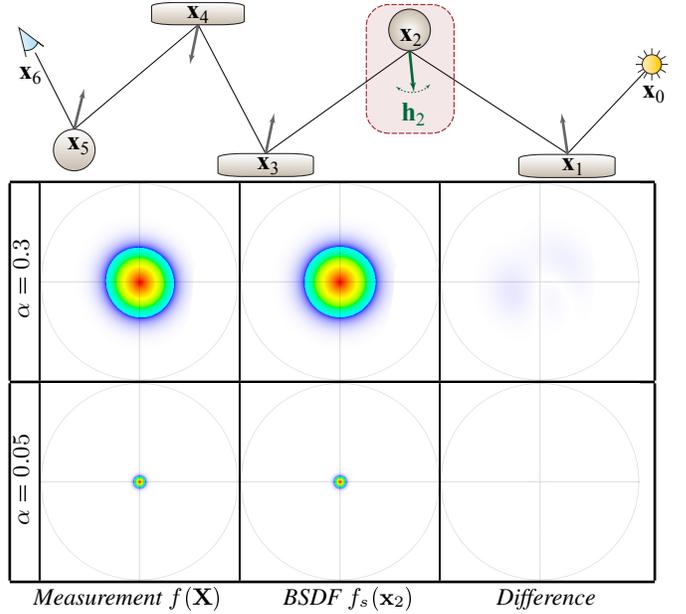
As an example, let us consider only surface reflections (as in Fig. 4). If we insert the reflection Jacobians  $|d\mathbf{h}_i/d\mathbf{o}_i| = 1/|4 \cdot \langle \mathbf{h}_i, \mathbf{o}_i \rangle|$  for transforming half vector density to outgoing direction (see [Walter et al. 2007] and Eq. 10 in the appendix) into Eq. 5, we obtain:

$$= \left| \frac{d\mathbf{o}_0^\perp}{d\mathbf{x}_k} \right| \prod_{i=1}^{k-1} 4 \cdot \left| \frac{\langle \mathbf{o}_i, \mathbf{n}_i \rangle \langle \mathbf{o}_i, \mathbf{h}_i \rangle}{\langle \mathbf{h}_i, \mathbf{n}_i \rangle} \right|. \quad (7)$$

Note that all singularities caused by geometric terms except for  $G(\mathbf{x}_0, \mathbf{x}_1)$  cancel out, which is a desirable property for integration.

### 4.3.2 Qualitative and Quantitative Analysis

The simplified product of Jacobians in Eq. 5 is numerically more robust than evaluating geometric terms for difficult geometric con-



**Figure 6:** Changes of the measurement contribution  $f(\mathbf{X})$  (left) and the Blinn-Phong BSGF  $f_s(\mathbf{x}_2)$  (center) caused by perturbing only the half vector at the highlighted vertex. We visualize normalized values in false-color inside the domain of the projected half vector (the unit circle). The differences (right) between the normalized distributions show that the changes of the full integrand are mostly due to the changes in the local BSGF at the perturbed half vector (Eq. 4).

figurations, e.g. corners, as transfer matrices in these configurations degenerate to identities compared to singularities (Fig. 5).

Another advantage of the new domain is that the integrand dimensions are highly decorrelated from each other, i.e., we can assume that changing the value of one constraint (e.g., a projected half vector) causes the change in the measurement contribution that mostly comes from the local change of the corresponding BSGF. Note that we do not take into account visibility and terms that are not expressible in our domain, such as the Fresnel coefficient, directional emission distribution of the light source and directional sensor responsiveness, for our importance sampling. In practice, we did not notice significant loss of sampling efficiency, and rely on Metropolis sampling to account for these terms. Thus, modifications to a projected half vector  $\mathbf{h}_i^\perp$  mostly influences the BSGF at a single vertex  $\mathbf{x}_i$ . The remainder of the integrand (Eq. 5) depends on the differential geometry, which is well-behaved within a submanifold.

By analyzing this remainder (Eq. 7 shows an example), we can conclude that all components of this product vary smoothly: the first term contains one initial geometric term and the determinant of one transfer matrix. The product consists of cosines that are known to introduce smooth low-frequency changes in the half vector domain.

To support these claims, we provide numerical validation. In Fig. 6 (more data provided in the supplemental material), we analyze how the product of Jacobians influences the shape of the integrand in half vector domain. Note that the distributions are very similar, which confirms that the product of Jacobians is relatively flat. This also demonstrates our observations about the decorrelated dimensions in the new domain. It also shows that the effect of the Jacobians in the case of glossy transport is negligible. Note that even for almost diffuse transport the product remains well-behaved.

#### 4.4 Specular Interactions in the New Domain

Specular interactions are always considered as a special case in light transport. In fact, as we will see, one of the terms in the product in Eq. 5 cancels out with a specular BSDF. Note that this reduces the dimensionality of the integration and entails that the respective constraint (half vector) should not be perturbed when sampling paths.

This cancellation is observed when the specular BSDF  $f_s(\mathbf{i} \rightarrow \mathbf{o})$  with its delta distribution is regarded in the half vector domain. The respective transformation from outgoing direction domain to the domain of half vectors appends the Jacobian  $d\mathbf{o} \mapsto d\mathbf{h}$  as

$$f_s = \kappa \frac{\delta_{\mathbf{o}}(\mathbf{s})}{|\langle \mathbf{o}, \mathbf{n} \rangle|} = \kappa \frac{\delta_{\mathbf{h}^\perp}(\mathbf{0})}{|\langle \mathbf{o}, \mathbf{n} \rangle|} \left| \frac{d\mathbf{h}}{d\mathbf{o}} \right|, \quad (8)$$

where  $\mathbf{s}$  is the outgoing direction of perfect specular interaction and  $\kappa$  folds all other factors like the reflection/transmission coefficient and the Fresnel term; we directly omitted the dot product  $|\langle \mathbf{h}, \mathbf{n} \rangle| = 1$  in the numerator caused by the projected half vector.

Note that Eq. 8 contains the inverse Jacobian,  $|d\mathbf{h}/d\mathbf{o}|$ , to the respective Jacobian in Eq. 5. For the special case of a purely specular chain in a path, we can observe this cancellation in the resulting sub-Jacobian, i.e., the corresponding terms in Eq. 5 are equal to Eq. 6: it results in the generalized geometric term used in manifold exploration [Jakob 2013]. In practice this means that we can directly evaluate materials with delta distributions in our framework (by combining Eq. 8 with Eq. 10) as they are deduced to the canonical form. That is, they do not require treatment in a special domain, e.g. like the “discrete measure” used in the Mitsuba renderer [2010].

#### 4.5 Discussion

The aforementioned properties enable a path sampling strategy to importance-sample all interactions of a path inbetween endpoints at once. For example, it is easy to estimate the change in path measurement due to altering the value of a single half vector constraint. Note that many BSDFs can easily be expressed [Walter et al. 2012], represented [Rusinkiewicz 1998], or approximated (e.g., by the Beckmann roughness equivalent [Jakob 2010]) in this domain.

The major advantage of integrating in this space is that we can modify individual constraints while preserving others, e.g. the endpoints of a path. This is in contrast to other measures: with the measure of outgoing directions, the endpoint position depends on chosen outgoing directions; or in the on-surface area measure, moving one vertex affects the directions to two adjacent vertices.

In this section we detailed only the most common case of point-to-point transport, where two endpoints are fixed. Other types of boundary constraints follow the same derivation, e.g., directional constraints due to a directional light source or an orthographic camera model which replace the positional constraints. The only difference is that the positional constraint at the corresponding endpoint is replaced by a directional constraint (the same as in manifold exploration, please see [Jakob 2013, p.70] for more details).

With Eq. 5, we compute the path measurement contribution  $f(\mathbf{X})$  using transfer matrices (or system matrices), which were prior used only for specular transport paths [Shinya et al. 1987]. Thus, we employ the transfer matrices for the general light transport framework with arbitrary scattering (such as glossy and diffuse) along the path. In other words, the path is always treated as a pure specular transport going through the specular microfacets defined by a set of preselected (in our case, sampled) half vectors (see Fig. 4).

## 5 Half Vector Space Light Transport

As mentioned before, our representation considers subspaces of the path space, and only within a subspace the Jacobians described in the previous section are valid. To sample from each subspace, we employ a Markov chain method and explore the space of natural constraints by using small perturbations to the half vectors. In practice, similarly to original MLT or ERPT [Cline et al. 2005], we sample a new path with another method, such as bidirectional mutation or BDPT, and transform it into the half vector representation. In this form, we can then perturb the constraints of the path (the endpoints as well as the half vectors) and by this efficiently explore the subspace. These two steps, new path sampling and perturbation, enable us to sample the entire path space.

### 5.1 Half Vector Space Mutation

The key of this exploration is a new mutation strategy, detailed in this section, which adds to the MLT framework. Starting with the current path in the Markov chain, we compute the geometric derivatives of the constraints in the tangent spaces of all vertices on the path. This step is analogous to manifold exploration and results in a block-tridiagonal matrix  $J$  which expresses the change of projected half vectors with respect to vertex positions in their respective tangent spaces (Eq. 11, described in more detail in Appendix A and [Jakob 2013]). In this process we also compute and store the projected half vectors of the current path.

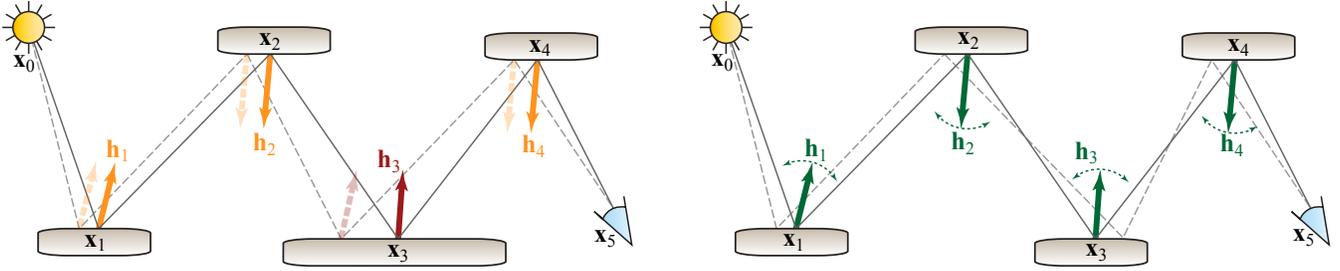
We then proceed by mutating the path by perturbing all half vector constraints simultaneously, e.g. by sampling on a small disk around each individual half vector of the current sample. Note that all segments of a path can be affected even if only one half vector changes. We first focus on how paths are mutated in general, and next describe our sampling strategy for constraints which considers surface roughness as well as stratification on the image plane (Sec. 6).

**Obtaining Path Vertices** After perturbing the half vectors, we need to determine the locations of the new path vertices. These locations are required, for example, to test for visibility or to read out the surface properties. To this end, we use a modified predictor-corrector Newtonian method from manifold exploration [Jakob and Marschner 2012]. In principle this works by first *predicting* new vertex positions using  $J$  as the Jacobian in a Newtonian method. And second, *correcting* these positions by using ray tracing to project the predicted points back onto the scene geometry.

If we applied the original correction scheme of manifold exploration to project vertex positions, it would retrace the path starting at one of the endpoints, e.g., from  $\mathbf{x}_0$  towards  $\mathbf{x}_1$ , and enforce the newly sampled half vectors at every interaction. This results in a displaced end vertex  $\mathbf{x}'_k$  and the correction  $T_1 \cdot (\mathbf{x}'_k - \mathbf{x}_k)$  can be used to iteratively refine the path vertices  $\mathbf{x}'_i$ . This original algorithm converges when the error  $(\mathbf{x}'_k - \mathbf{x}_k)$  is not too large; this approach typically works very well for purely specular surfaces.

However, we use another approach, which first explicitly predicts the change of all vertex positions at once using  $\Delta \mathbf{X} = J^{-1} \Delta \mathbf{H}^\perp$  (see Eq. 11 in the Appendix) and then project the new vertex positions onto the surfaces. The above prediction-correction steps have to be iterated until the half vectors of the resulting path match the sampled constraints. Note that, due to the sparsity of the matrix, a full inversion can be avoided (Sec. 7). This strategy, although a little more costly to evaluate, typically has better convergence to desired half vectors, e.g., at grazing interaction angles.

We select the projection direction (either from the sensor or from the light source) by computing  $|d\mathbf{x}_1/d\mathbf{x}_{k-1}| = |T_1| \cdot |T_{k-1}|^{-1}$ ,



**Figure 7:** *Manifold exploration (left) fixes the half vectors (orange) for two selected subpaths by stochastically classifying vertices as specular. It then perturbs the direction at one of the end points and tries to construct a new path. Unfortunately, the scattering at the vertex between these two subchains (here:  $\mathbf{x}_3$ , half vector depicted in red) is not importance-sampled. In contrast, our new strategy takes all BSDF interactions along the path into account. Moreover, we importance-sample all half vectors at once, compared to uncontrollably changing only one constraint (red) in manifold exploration. This reduces the sample correlation and leads to faster exploration of illumination features.*

which is the first-order estimation of the movement of  $\mathbf{x}_1$  when moving  $\mathbf{x}_{k-1}$ . If this ratio is less than one, we start projecting vertices with  $\mathbf{x}_{k-1}$ ,  $\mathbf{x}_{k-2}$  etc. (from the sensor), as this minimizes the projection error. Otherwise, we project starting from the light ( $\mathbf{x}_1$ ,  $\mathbf{x}_2$ , ...). This allows us to improve the convergence (see Table 2).

**Acceptance Probability** To compute the acceptance probability as  $\min(1, a)$  (Eq. 2), both transition probabilities  $T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})$  and  $f(\mathbf{X})$  are computed in product vertex area measure  $d\mathbf{X}$ . When mutating in constraint space  $\Omega(\mathbf{H}^\perp)$  we can easily compute the transition probability  $T(\mathbf{H}_t^\perp \rightarrow \mathbf{H}_{t+1}^\perp)$ , as the joint probability is just the product of the probabilities of the individual perturbations  $p(\mathbf{h}_{t+1}^\perp | \mathbf{h}_t^\perp)$ . To compute  $a$  in product vertex area measure, we transform  $T(\mathbf{H}_t^\perp \rightarrow \mathbf{H}_{t+1}^\perp)$  back to  $\Omega(\mathbf{X})$  and evaluate

$$R_{t+1} = f(\mathbf{X}_{t+1}) / \left( T(\mathbf{H}_t^\perp \rightarrow \mathbf{H}_{t+1}^\perp) \left| \frac{d\mathbf{H}^\perp}{d\mathbf{X}} \right| \right) = f(\mathbf{X}_{t+1}) \left| \frac{d\mathbf{X}}{d\mathbf{H}^\perp} \right| / T(\mathbf{H}_t^\perp \rightarrow \mathbf{H}_{t+1}^\perp),$$

where the second line shows the interpretation of integration in half vector space, as described in Sec. 4.3. This is a big advantage when implementing this formula, since it is easy to compute  $T(\mathbf{H}_t^\perp \rightarrow \mathbf{H}_{t+1}^\perp)$  and the numerator can take advantage of the simplified measurement computation (Eq. 5).

Recall that in order to guarantee convergence with MLT, by achieving ergodicity, the set of mutations has to cover the whole path space. The proposed mutation does not meet this requirement, e.g. it does not mutate path lengths. To this end, we always use our mutation together with the bidirectional mutation [Veach 1998]. Note

	From light	From eye	Proposed
Successful generation	47.8%	51.2 %	52.8%
Acceptance rate	45.3%	46.8%	47.4%
Average #iterations	2.20	2.19	2.18

**Table 2:** *Statistics for different direction of projection for the corrector in our predictor-corrector scheme in the KITCHEN scene. The statistics was gathered on the same number of paths (10.84M) rendered with the same initial seed paths. The left column is for the ray-traced projection that always starts from the direction of the light source. The middle column is for projecting from the sensor (eye); and the rightmost column shows the advantages of our direction selection heuristic.*

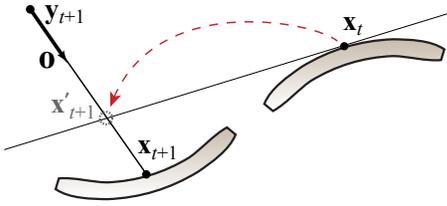
that, similar to manifold exploration, our method does not introduce visible bias in practice, even though paths are constructed up to a limited numerical precision.

**Comparison to Manifold Exploration** Since the formulation used to derive manifold exploration (ME) is based on removing Dirac deltas from the integration, it is conceptually limited to specular surfaces (i.e., the special case described in Sec. 4.4). The extension of ME to glossy materials stochastically classifies all but one glossy interaction as specular, and keeps the half vectors for these “quasi-specular” interactions fixed. In every mutation a different half vector is non-fixed. The BSDF at this interaction is evaluated when the path construction is complete. Note that this can lead to uncontrollable change in path measurement contribution  $f(\mathbf{X})$  due to the lack of importance sampling this BSDF. Fig. 7 (left) shows an example, where vertex  $\mathbf{x}_3$  is classified as non-specular, which means that ME will not importance-sample the BSDF at  $\mathbf{x}_3$ , nor try to keep the half vector constant. Note that because the half vectors at the other vertices are kept constant, they might be insufficiently resampled by the method, which can lead to visible correlation artifacts. For example, the scratch-like artifacts in the gold ring closeup in Fig. 1 and in the reflection on the wall in Fig. 13 appear because the manifold exploration often proposes significant changes to the chain of glossy interactions. Also, treating glossy vertices this way still requires to account for the Jacobian determinant, which in turn requires a costly full matrix inversion in the case when specular and glossy vertices are mixed [Jakob 2013].

In contrast, mutating in the half vector domain does not depend on classifying vertices as specular or non-specular, and can directly explore a glossy caustic by importance-sampling all BSDF interactions at once. Specular interactions do not require special treatment, as their half vectors always remain at a fixed position (at zero), effectively reducing the dimensionality of the integration. Note that the simplified measurement contribution in half vector space (Eq. 5) treats all types of interactions and BSDFs (diffuse, glossy, and specular) equally, without a need to handle any type differently.

## 5.2 Transitioning between Subspaces

As noted in Sec. 4.3, the integration in the domain of half vectors is valid only on small subspaces  $\Omega(\mathbf{H}_t^\perp)$  (Fig. 2). To be more precise, we assume that within such subspace of the path space the implicit function theorem holds [Spivak 1965]. That is, this subspace is a region of continuity of the constraints Jacobian in Eq. 11. This implies, for example,  $\mathbb{C}^2$  smoothness of the surfaces in such subspaces (which in turn provides the continuity of constraint derivatives). Manifold exploration is also more efficient with continuity of the



**Figure 8:** When jumping from one submanifold to another, first the point  $\mathbf{x}'_{t+1}$  is predicted by the constraint Jacobian matrix  $J$  evaluated at point  $\mathbf{x}_t$ . Then the projection step traces a ray from the point of the previous vertex of the proposal  $\mathbf{y}_{t+1}$  to find the projected point  $\mathbf{x}_{t+1}$  on the new subspace.

constraints when performing the walk [Jakob 2013, p.94]. However, such ideal requirements are practically impossible to achieve on arbitrary scenes.

Our mutation works directly in the space of constraints and relies on the derivatives of the constraints at every vertex along the path. Thus the discontinuities in these derivatives, e.g. in normal derivatives, become more apparent (see Fig. 9, left). The remainder of this section is devoted to this problem and introduces an important extension to our random walk, which enables it to switch between different subspaces, leading to easier exploration of the path space.

Analogous to manifold exploration, we test whether the walk with our projection method from a current path  $\mathbf{H}_t^\perp$  to a proposal path  $\mathbf{H}_{t+1}^\perp$  is reversible, i.e., if we can also walk back from  $\mathbf{H}_{t+1}^\perp$  to  $\mathbf{H}_t^\perp$ . Note that our predictor-corrector scheme might potentially find a new path that satisfies the proposed constraints  $\mathbf{H}_{t+1}^\perp$  but lies in a different subspace. In this case, if we then perform a reverse walk, we might end up with a path that satisfies the original constraints  $\mathbf{H}_t^\perp$  but still lies in this different subspace. We determine such cases by testing whether the vertex positions match the positions of the original path and reject the proposal based on that. This ensures the detailed balance requirement of the Metropolis-Hastings algorithm.

In case of a non-reversible walk, manifold exploration rejects the proposal path. In contrast, in our method, if the walk fails (i.e., when our predictor-corrector scheme does not converge to the given new constraints  $\mathbf{H}_{t+1}^\perp$ ), we reject the proposal but attempt to use the predicted path differently instead of immediately failing the walk. For that, we test whether the first order prediction of the new path given by the Jacobian of the constraints results in a valid path, i.e., with non-zero throughput after projection; note that this path is the prediction of the first iteration of our scheme. If this path is valid, we make sure it is non-reversible and take it as a *suggested* perturbation of the path. We interpret this case as an indication of a jump from one subspace to another and apply different transition rules (described below). This extension allows for better exploration of fragmented and highly displaced geometry (see Fig. 9).

For performing a random walk, we compute the new transition probability density  $T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})$  of sampling this suggested path in the on-surface measure. This probability density consists of the probability densities of (1) predicting each vertex  $\mathbf{x}'_{t+1}$  and (2) projecting  $\mathbf{x}'_{t+1}$  to obtain  $\mathbf{x}_{t+1}$ . The former is the probability density of sampling the point  $\mathbf{x}'_{t+1}$  on the tangent plane of the current path's vertex  $\mathbf{x}_t$ . This simple change of the domain is computed using the constraint Jacobian as

$$p(\mathbf{X}'_{t+1}|\mathbf{X}_t) = p(\mathbf{H}_{t+1}^\perp|\mathbf{H}_t^\perp) \left| \frac{d\mathbf{X}}{d\mathbf{H}^\perp} \right| = p(\mathbf{H}_{t+1}^\perp|\mathbf{H}_t^\perp) |J_{\mathbf{H}_t^\perp}|^{-1}.$$

The correction step obtains the projected vertex  $\mathbf{x}_{t+1}$  by casting a ray from the previous vertex  $\mathbf{y}_{t+1}$  towards the outgoing direction



**Figure 9:** SALAD BOWL: the exploration across geometric edges is difficult as a walk cannot easily jump to another subspace (left); with our improvement transition, the walk can more easily jump between submanifolds leading to more uniform exploration of the illumination (center); right: false colors showing where the walk attempted to jump (red) and where the jump was successful (green).

$\mathbf{o} = \overline{\mathbf{y}_{t+1}\mathbf{x}'_{t+1}}$  (Fig. 8). Note that  $\mathbf{x}'_{t+1}$  might be different from  $\mathbf{x}_{t+1}$ , i.e., it might not lie on a surface. The second probability density accounts for this and converts the on-surface probability density of the predicted vertex from the virtual surface (defined by the tangent plane at  $\mathbf{x}_t$ ) to the density at the new projected vertex  $\mathbf{x}_{t+1}$  on the actual surface. As the projection is done by tracing a ray, we transform the density using the shared domain of fixed outgoing direction  $\mathbf{o}$  (Fig. 8), which leads to a simple ratio of geometric terms for the points  $\mathbf{x}'_{t+1}$  and  $\mathbf{x}_{t+1}$ :

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t) = p(\mathbf{x}'_{t+1}|\mathbf{x}_t) \left| \frac{d\mathbf{o} d\mathbf{x}'_{t+1}}{d\mathbf{x}_{t+1} d\mathbf{o}} \right| = p(\mathbf{x}'_{t+1}|\mathbf{x}_t) \frac{G(y, \mathbf{x}_{t+1})}{G(y, \mathbf{x}'_{t+1})}.$$

The transition probability of the full proposal path  $T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})$  is then computed sequentially as a joint product of probabilities at every vertex along this path, starting from the same direction as the projection step. Note that this transition probability should satisfy the detailed balance. Practically, we achieve this by dividing by the current transition probabilities for both ways [Hastings 1970] when computing the acceptance probability  $a$  as in Eq. 2, i.e., we compute both  $T(\mathbf{X}_t \rightarrow \mathbf{X}_{t+1})$  and the reverse probability  $T(\mathbf{X}_{t+1} \rightarrow \mathbf{X}_t)$ . Moreover, to make this random walk symmetric, we test upfront that the Newtonian method cannot converge *both ways* guaranteeing that the walk cannot be done from the suggested path back to the current path using our regular Newtonian method (we denote such walks as *bidirectionally invalid*).

**Discussion** In other words, if a walk is not possible, we try to take a path suggested by the constraints derivatives and then treat it as a path generated by perturbing directions, as in other mutation strategies. This enables jumping between different subspaces.

Note that we use this procedure only if a walk is bidirectionally invalid, which indicates the boundary of a subspace. The fraction of non-reversible walks depends on the structure and smoothness of the scene geometry and is slightly higher for our method compared to manifold exploration (see Table 3) due to higher dependence on the differential geometry around a path.

## 6 Sampling Improvements

In the last sections, we described the principle of perturbing a path in half vector space. To make the perturbation efficient for computing images, the following sections describe how to adapt the mutation step sizes to the individual BSDFs, stratify samples in image space, and consider the bandwidth of the signal to decide upon the strength of perturbation of the individual dimensions.

	Ours	ME
Average # iterations	2.18/6.63	2.61/2.98
Generation rate	58.8%	46.2/59.2%
Non-reversible	3.3%	2.1%
- Submanifold jump	88.3% (of 3.3%)	N/A
Acceptance rate	47.1%	35.7%
Mutations	453.6 M	488.6 M

**Table 3:** Statistics for our mutation strategy and the corresponding numbers for manifold exploration for the equal-time renders in Fig. 1. The number of iterations is counted per successful proposal and total. Note that our method can create almost the same total number of mutations as manifold exploration.

### 6.1 Optimal Expected Step Size for Interactions

We can optimize the integration for individual interactions, as the space of half vectors has the property of decorrelated dimensions. That is, changing one half vector  $\mathbf{h}_i^\perp$  mostly influences the value of the BSDF at the vertex  $\mathbf{x}_i$  only (see Sec. 4.3.2). This *local* analysis can be applied to every constraint independently to improve the sampling of paths as a whole.

In the following we assume that all BSDFs can express their roughness as Beckmann equivalent  $\alpha$ , i.e., the microfacet (or half vector) distribution of a BSDF can be approximated by a Beckmann distribution. As Walter et al. [2012] noticed, most microfacet-based BSDFs are easily expressible in this way. This assumption is also practical for a renderer and is, e.g., used in Mitsuba Renderer [2010].

If we consider a single interaction, e.g. a glossy reflection, we can estimate the optimal expected value of the mutation step size. Projected onto the parallel plane at distance one to the tangent plane, a Beckmann lobe is a 2D Gaussian around the center with variance  $\sigma^2 = \alpha^2/2$ . Note that we use the projected half vector  $\mathbf{h}^\perp$  which is related to this plane-plane domain as  $\mathbf{h}_{p-p} = \mathbf{h}/\sqrt{1 - |\mathbf{h}^\perp|^2}$ .

The optimal MCMC acceptance rate for such a 2D Gaussian is about 35% [Bremaud 1999] and the optimal expected step size for a Gaussian proposal is  $s = (2/\pi)|\Sigma|$ , where  $\Sigma$  is the covariance matrix of the BSDF's 2D Gaussian. We can derive the expected step size in plane-plane for a single half vector dimension as

$$s_{p-p, \max} = \alpha/\sqrt{\pi}.$$

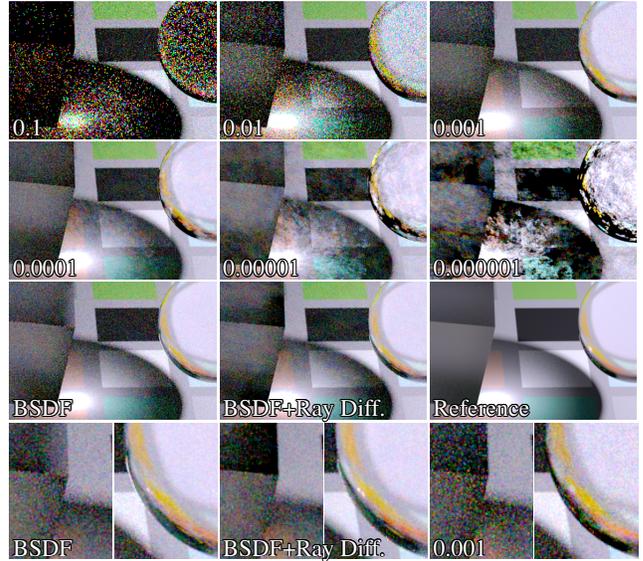
When mutating a half vector constraint, the expected value of the change to the half vector (i.e., the distance between the current and perturbed half vector on the parallel plane) should have an expected value of  $s_{p-p, \max}$  to achieve the fastest mixing of the Markov chain (we describe the actual perturbation after analyzing stratification in Sec. 6.2). This upper bound for an optimal step size is designed to achieve the maximum exploration of a 2D Gaussian.

Because visibility is ignored in our estimation, it can narrow the actual support when paths are mutated. We account for this by reducing the step size as

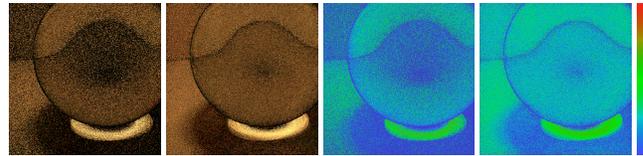
$$s_{p-p} = \gamma \cdot s_{p-p, \max}, \quad \text{with } \gamma \in (0, 1].$$

The scaling parameter  $\gamma$  has only a slight effect on the sampling efficiency. We found that  $\gamma = 0.1$  (10% of the upper bound) gives the best results for difficult visibility cases in practice and thus has been used for all rendered images in this work.

Note that  $s_{p-p}$  is defined in the plane-plane domain. It can be converted to half vector domain using the projected half vector and the relation provided above. However, in practice we use a numerically more stable Möbius transformation described later in Sect. 7 and in



**Figure 10:** A simple scene with a diffuse ground plane, and a sphere and a cube with a rough dielectric with different roughnesses (0.01 and 0.1), rendered with 128 samples per pixel and different fixed step sizes as indicated (two top rows). The 3rd row shows our optimal step sizes for BxDF bandwidth and the same combined with ray differentials, as well as a reference render. The insets in the bottom row show that our automatic BxDF estimation is comparable to the best manually chosen parameter. The large step sizes for the BxDF maximize the exploration of the submanifold, but are limited by the ray differentials in favor of stratification on the image plane. For an uncropped version, see the supplemental material.



**Figure 11:** False color visualization of sample counts in the CORNELL BOX with a glass sphere rendered with 8 spp without (left) and with (right) the proposed stratification.

the supplementary. We demonstrate the importance of the optimal step size selected for each individual interaction in Fig. 10.

### 6.2 Stratification over the Image Plane

The second aspect of a good mutation strategy is to ensure stratification of the samples on the image plane. We show how to achieve this by computing the required perturbation to the half vectors from the desired stratification on the image plane, and how to combine these perturbations with the BSDF perturbations derived earlier.

Given a path to mutate, in order to map changes in image space to half vector space, we first compute the ray differentials [Igehy 1999] from the current position on the sensor  $\mathbf{x}_k$  to the first vertex in the scene (which requires no tracing of the actual differentials). Then we project the ray differential vectors along the  $\mathbf{u} \times \mathbf{v}$  basis of the image plane ( $\mathbf{u}$  and  $\mathbf{v}$  are on the image plane with a length of 1 pixel horizontally and vertically) onto the tangent frame of the first vertex in the scene  $\mathbf{x}_{k-1}$ . This gives us the pair of offset vectors  $\Delta \mathbf{x}_{\mathbf{u}, k-1}$  and  $\Delta \mathbf{x}_{\mathbf{v}, k-1}$ . Veach [1998] suggested to scale these vectors by 5% of the image size and use this as the desired change



differential lengths  $\|\Delta\mathbf{h}_u\|$  and  $\|\Delta\mathbf{h}_v\|$  along the respective axes). Then we apply a Möbius transformation [Hanika 2011, p.40] to move the center of this distribution to the current half vector  $\mathbf{h}^\perp$ . This has the full disk as support and never samples outside it.

## 8 Results, Discussion and Limitations

We implemented our techniques in Mitsuba and computed all our equal-time renders using this framework on an Intel i7-3930K hex-core CPU at 3.20GHz using 12 threads. To facilitate future work and comparisons, our full implementation is available.

We compare our method, i.e., half vector mutations with seeding by bidirectional mutation (see Sect. 5.1), to MLT with manifold explorations (ME) and Kelemen’s variant of MLT perturbing in the primary sample space (PSSMLT) [Kelemen et al. 2002] and we additionally show results of bidirectional path tracing (BDPT).

**Parameters** We use the default parameters in the Mitsuba renderer which are the recommended optimal parameters from the corresponding publications. According to our experiments, they also yield the best results on our test scenes. We use 1.0 for the average number of chains and 100 mutations per chain for ERPT. We use  $\lambda = 50$  for manifold exploration (results with different  $\lambda$ -values are provided in the supplementary).

**Scenes** Fig. 1 show a scene with a difficult mixture of glossy and purely specular surfaces, lit by a point spot light and area lights. Fig. 13 shows a setup with glossy transport and complex distant illumination coming from the outside. Most of the illumination in the image is due to a caustic reflected off the glossy floor. When using a path tracer as mapping of primary space samples to path space in PSSMLT, BSDFs (and thus half vectors) are importance-sampled, but in contrast to our method, in the domain of outgoing directions. As a consequence, it always misses the point light source. If a bidirectional path sampler is used, a deterministic connection will be performed, evaluating a BSDF instead of importance-sampling it. ME explores such scenes by keeping the light vertex  $\mathbf{x}_0$  fixed, but cannot importance-sample all BSDFs simultaneously, leading to visible correlation in the rendered highlights.

We would like to point out that manifold exploration and our method are the only perturbations supporting the exploration of reflected caustics (specular-diffuse-specular paths); finding them is an orthogonal problem [Kaplanyan and Dachsbacher 2013b].

Scenes with complex occlusion are challenging for our mutation strategy, as we did not specifically design it to work with difficult visibility. Unfortunately, representing the path integral in half vector domain will incur that all visibility edges are now present in all decorrelated 2D slices of the integration domain (i.e., it appears in each BSDF interaction as the one depicted in Fig. 6). Please refer to the supplementary material where we show further numerical examples for scenes with high occlusion. The AJAR DOOR scene (Fig. 14) belongs to this category of difficult scenes. Nonetheless, our mutation strategy handles this case well and even performs on par with ME in some regions of the image, e.g. on the back wall which is explored very uniformly. Although our method worked well on the scenes we tested, we anticipate that treating visibility more efficiently will be important future work.

Although our implementation leaves room for optimization, our mutation strategy involves more costly computation than other strategies in MLT. Note that we show equal-time comparisons not images with an equal number of samples. That is, the more costly yet higher quality samples with our method ultimately pay off.

We demonstrated that our mutation can perform very well on complete paths. However, it is also practical to use it as a drop-in replacement for the deterministic manifold walk procedure used in manifold exploration for connecting two points.

In our implementation used to generate the results, we always retrace the entire path to reduce the correlation of the generated samples. Diffuse interactions, however, are easier explored by only retracing small parts of a path and then performing a connection. This contrasts the problem that keeping a part of a path fixed increases the correlation between samples (see Fig. 13), which is a well-known disadvantage of MCMC methods. A combination with more classical mutation strategies remains an interesting direction for future work.

Another interesting direction would be to explore more convenient parameterizations for half vectors, e.g., a plane-plane parameterization as used for light fields, where the Beckmann lobe can be better expressed and sampled.

Our work focuses on surface rendering. In presence of participating media, the dimensionality of the problem increases, since there are no explicitly defined differential manifolds as in the case of surfaces. We can imagine that our framework could be extended by additional constraints for participating media.

## 9 Conclusion

In this paper we presented a novel approach to sampling and integrating light transport paths. For this we introduced the space of natural constraints and this domain proved being an attractive alternative space to compute and analyze light transport. We derived how to convert from this domain to previously used measurement spaces and examined the resulting transformation Jacobian determinant. Our new representation has several benefits and enabled us to design a mutation strategy for Metropolis light transport which can importance sample all BSDF interactions along a path, while controlling image space stratification and incorporate recent findings of frequency analysis of light transport at the same time. It has been tailored for glossy and specular interactions, but likewise handles diffuse transport robustly.

**Acknowledgments** We would like to thank the anonymous reviewers (esp. #10) for their detailed comments, Wenzel Jakob for Mitsuba [Jakob 2010], Miika Aittala, Samuli Laine, and Jaakko Lehtinen for the AJAR DOOR, and Alex Telford for the JEWELRY scene.

## A Half Vectors and Transformations

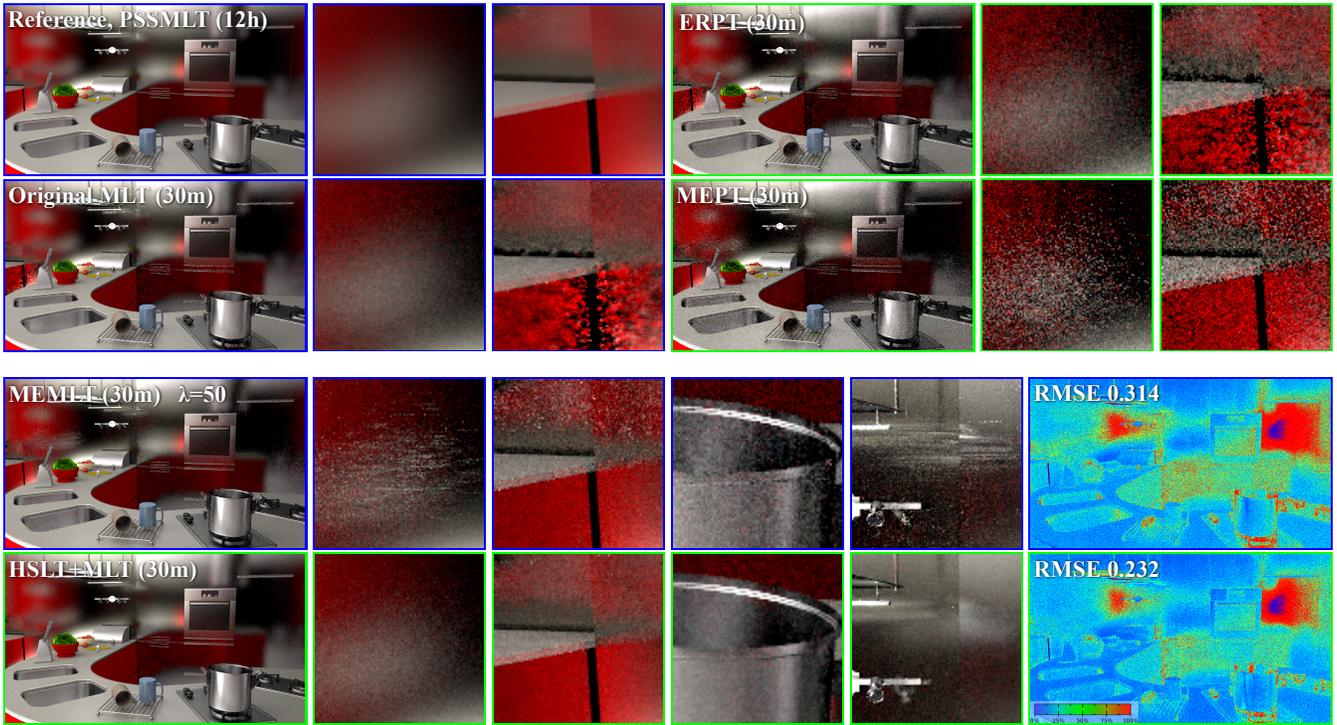
The (non-projected) half vector  $\mathbf{h}$  is defined as the *generalized half vector* [Sommerfeld and Runge 1911; Walter et al. 2007], as

$$\mathbf{h} = (\mathbf{i} + \eta\mathbf{o}) / \|\mathbf{i} + \eta\mathbf{o}\|.$$

The term  $\eta_i$  is the ratio of indices of refraction at vertex  $i$ , corresponding to the previous and next edge of the path. It is 1 for reflection and  $\eta_o/\eta_i$  for transmission events, where  $\eta_o$  corresponds to the edge along the outgoing direction  $\mathbf{o}$ . We can transform the half vector density  $d\mathbf{h}$  to the density of an outgoing direction  $d\mathbf{o}$  by applying Jacobians from BSDF theory [Walter et al. 2007]:

$$\left| \frac{d\mathbf{h}}{d\mathbf{o}} \right| = \frac{\eta^2 |\langle \mathbf{h}, \mathbf{o} \rangle|}{(\langle \mathbf{h}, \mathbf{i} \rangle + \eta \langle \mathbf{h}, \mathbf{o} \rangle)^2} \quad (10)$$

Another useful Jacobian for the change of variables from vertex positions density  $d\mathbf{x}_i$  in local tangent space to the density of projected



**Figure 13:** Equal-time rendering of the KITCHEN scene with difficult glossy paths. The top row shows BDPT, MLT, ERPT and MEPT for comparison. The two bottom rows compare manifold exploration (MEMLT) (237 spp, acceptance rate 42.7%) to our method (HSLT) (229 spp, acceptance rate 67.3%). Note how efficient the new mutation can estimate the optimal sampling density of difficult regions and handle hard features such as glossy caustics (please see the supplementary for more results).

half vectors  $d\sigma^\perp(\mathbf{h}_i) \equiv d\mathbf{h}_i^\perp = d\mathbf{h}_i |\langle \mathbf{h}_i, \mathbf{n}_i \rangle|$  which is written as

$$J = \frac{d\mathbf{H}^\perp}{d\mathbf{X}} = \frac{d(\mathbf{h}_1^\perp \dots \mathbf{h}_{k-1}^\perp)}{d(\mathbf{x}_1 \dots \mathbf{x}_{k-1})} = \begin{pmatrix} B_1 & C_1 & & & & \\ A_2 & B_2 & C_2 & & & \\ & \dots & \dots & \dots & & \\ & & A_{k-2} & B_{k-2} & C_{k-2} & \\ & & & A_{k-1} & B_{k-1} & \end{pmatrix}, \quad (11)$$

which has a structure of a block tridiagonal matrix with  $2 \times 2$  blocks  $A_i, B_i, C_i$ . More details on this matrix can be found in Jakob's PhD thesis [2013, Sec.4.3]. It can also be used as a first order approximation that maps finite changes of path vertices  $\Delta\mathbf{X} = (\Delta\mathbf{x}_1 \dots \Delta\mathbf{x}_{k-1})$  to offsets  $\Delta\mathbf{H}^\perp = (\Delta\mathbf{h}_1^\perp \dots \Delta\mathbf{h}_{k-1}^\perp)$  in projected half vector space as  $\Delta\mathbf{H}^\perp = J \cdot \Delta\mathbf{X}$ .

## References

- ARVO, J. 1986. Backward ray tracing. In *SIGGRAPH Course Notes*, 259–263.
- ASHIKHMIN, M., AND SHIRLEY, P. 2000. An anisotropic Phong BRDF model. *Journal of Graphics Tools* 5, 2, 25–32.
- BREMAUD, P. 1999. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer. ISBN 0-387-98509-3.
- CAPPÉ, O., GUILLIN, A., MARIN, J.-M., AND ROBERT, C. P. 2004. Population Monte Carlo. *Journal of Computational and Graphical Statistics* 13, 4.
- CHEN, J., WANG, B., AND YONG, J.-H. 2011. Improved stochastic progressive photon mapping with Metropolis sampling. *Computer Graphics Forum* 30, 4, 1205–1213.
- CLINE, D., TALBOT, J., AND EGBERT, P. K. 2005. Energy redistribution path tracing. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 24, 3, 1186–1195.
- DACHSBACHER, C., KRIVÁNEK, J., HASAN, M., ARBREE, A., WALTER, B., AND NOVÁK, J. 2013. Scalable realistic rendering with many-light methods. *Computer Graphics Forum DOI: 10.1111/cgf.12256*.
- FAN, S., CHENNEY, S., AND CHI LAI, Y. 2005. Metropolis photon sampling with optional user guidance. In *Proc. Eurographics Symposium on Rendering*, 127–138.
- GEORGIEV, I., KRIVANEK, J., DAVIDOVIC, T., AND SLUSALLEK, P. 2012. Light transport simulation with vertex connection and merging. *ACM Trans. on Graphics (Proc. SIGGRAPH Asia)* 31, 6, 192:1–192:10.
- HACHISUKA, T., AND JENSEN, H. W. 2011. Robust adaptive photon tracing using photon path visibility. *ACM Trans. on Graphics* 30, 5, 114:1–114:11.
- HACHISUKA, T., OGAKI, S., AND JENSEN, H. W. 2008. Progressive photon mapping. *ACM Trans. on Graphics (Proc. SIGGRAPH Asia)* 27, 5, 130:1–130:8.
- HACHISUKA, T., PANTALEONI, J., AND JENSEN, H. W. 2012. A path space extension for robust light transport simulation. *ACM Trans. on Graphics (Proc. SIGGRAPH Asia)* 31, 6, 191:1–191:10.
- HANIKA, J. 2011. *Spectral light transport simulation using a precision-based ray tracing architecture*. PhD thesis, Ulm University. VTS-ID/7539.



**Figure 14:** The AJAR DOOR rendered in 30 minutes. Left: reference (rendered with PSSMLT in 12 hours). Middle: using MEMLT with the original set of mutations, RMSE 0.244. Right: with HSLT (only our mutation and bidirectional mutation), RMSE 0.176. Note that although this scene is challenging for our method (all half vectors have limited variation due to difficult visibility), it is still handled well.

- HASTINGS, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 1, 97–109.
- IGEY, H. 1999. Tracing ray differentials. *Proc. SIGGRAPH*, 179–186.
- JAKOB, W., AND MARSCHNER, S. 2012. Manifold exploration: a Markov chain Monte Carlo technique for rendering scenes with difficult specular transport. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 31, 4, 58:1–58:13.
- JAKOB, W., 2010. Mitsuba renderer. <http://www.mitsuba-renderer.org>.
- JAKOB, W. 2013. *Light transport on path-space manifolds*. PhD thesis, Cornell University.
- JENSEN, H. W. 1996. Global illumination using photon maps. In *Proc. Eurographics Workshop on Rendering*, 21–30.
- KAJIYA, J. T. 1986. The rendering equation. *Computer Graphics (Proc. SIGGRAPH)*, 143–150.
- KAPLANYAN, A. S., AND DACHSBACHER, C. 2013. Adaptive progressive photon mapping. *ACM Trans. on Graphics* 32, 2, 16:1–16:13.
- KAPLANYAN, A. S., AND DACHSBACHER, C. 2013. Path space regularization for holistic and robust light transport. *Computer Graphics Forum (Proc. of Eurographics)* 32, 2, 63–72.
- KELEMEN, C., SZIRMAY-KALOS, L., ANTAL, G., AND CSONKA, F. 2002. A simple and robust mutation strategy for the Metropolis light transport algorithm. *Computer Graphics Forum* 21, 3, 531–540.
- KELLER, A. 1997. Instant radiosity. *Proc. SIGGRAPH*, 49–56.
- KNAUS, C., AND ZWICKER, M. 2011. Progressive photon mapping: A probabilistic approach. *ACM Trans. on Graphics* 30, 3, 25:1–25:13.
- KRIVÁNEK, J., GEORGIEV, I., KAPLANYAN, A., AND CANADA, J. 2013. Recent advances in light transport simulation: Theory and practice. In *ACM SIGGRAPH Courses*.
- LAFORTUNE, E. P., AND WILLEMS, Y. D. 1993. Bi-directional path tracing. In *Proc. of COMPUGRAPHICS*, 145–153.
- LEHTINEN, J., KARRAS, T., LAINE, S., AITTALA, M., DURAND, F., AND AILA, T. 2013. Gradient-domain Metropolis light transport. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 32, 4, 95:1–95:12.
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., AND TELLER, E. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* 21, 6, 1087–1092.
- RUSINKIEWICZ, S. 1998. A new change of variables for efficient BRDF representation. In *Proc. Eurographics Workshop on Rendering*, 11–22.
- SHINYA, M., TAKAHASHI, T., AND NAITO, S. 1987. Principles and applications of pencil tracing. *Computer Graphics (Proc. SIGGRAPH)*, 45–54.
- SHIRLEY, P., WADE, B., HUBBARD, P. M., ZARESKI, D., WALTER, B., AND GREENBERG, D. P. 1995. Global illumination via density-estimation. In *Proc. Eurographics Workshop on Rendering*, 219–230.
- SOMMERFELD, A., AND RUNGE, J. 1911. Anwendung der Vektorrechnung auf die Grundlagen der geometrischen Optik. *Annalen der Physik* 340, 277–298.
- SPIVAK, M. 1965. *Calculus on manifolds*. Addison-Wesley. ISBN 978-0-81-334612-0.
- SUBR, K., AND KAUTZ, J. 2013. Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 32, 4, 128:1–128:12.
- VEACH, E., AND GUIBAS, L. 1994. Bidirectional estimators for light transport. In *Proc. Eurographics Workshop on Rendering*, 147–162.
- VEACH, E., AND GUIBAS, L. J. 1997. Metropolis light transport. *Proc. SIGGRAPH*, 65–76.
- VEACH, E. 1998. *Robust Monte Carlo methods for light transport simulation*. PhD thesis, Stanford University. AAI9837162.
- WALTER, B., FERNANDEZ, S., ARBREE, A., BALA, K., DONIKIAN, M., AND GREENBERG, D. P. 2005. Lightcuts: a scalable approach to illumination. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 24, 3, 1098–1107.
- WALTER, B., MARSCHNER, S., LI, H., AND TORRANCE, K. 2007. Microfacet models for refraction through rough surfaces. In *Proc. Eurographics Symposium on Rendering*, 195–206.
- WALTER, B., KHUNGURN, P., AND BALA, K. 2012. Bidirectional lightcuts. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 31, 4, 59:1–59:11.