Volumetric Multi-View Rendering

Basile Fraboni[1,2], Antoine Webanck[2], Nicolas Bonneel[2], and Jean-Claude Iehl[2]

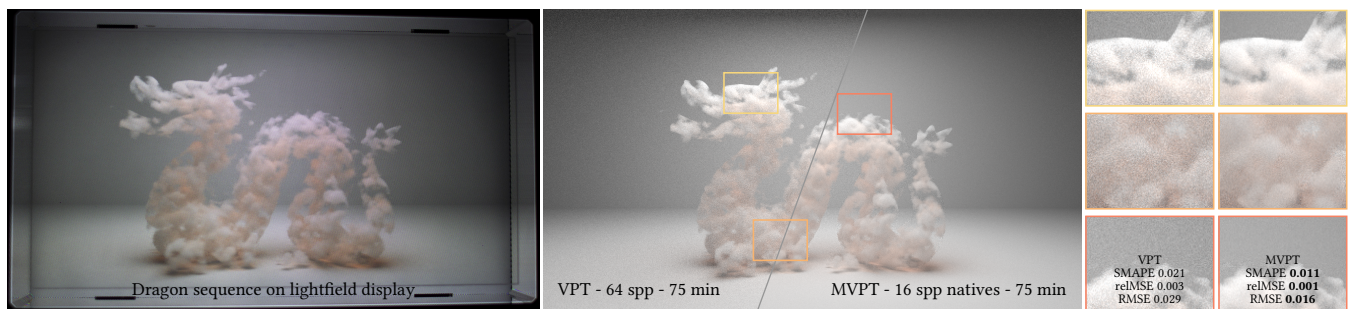[1]INSA Lyon, CNRS/LIRIS, France   [2]Univ Lyon, CNRS/LIRIS, France

Figure 1: *We render a sequence of 45 frames of the Dragon cloud scene including heterogeneous medium to be displayed on a lightfield screen. Instead of rendering the sequence frame by frame using volumetric path tracing (VPT), our method (MVPT) jointly renders the sequence at once reusing some computation during the simulation. We observe a significant variance reduction at equal time when rendering the whole sequence – or reach a desired quality at lower time. In addition the correlation of reused paths between nearby frames almost makes the flickering caused by dissimilar random sampling dissapear. This makes the sequences pleasier to the eye even at low sample count.*

Abstract

Rendering photo-realistic images using Monte Carlo path tracing often requires sampling a large number of paths to reach acceptable levels of noise. This is particularly the case when rendering participating media, that complexify light paths with multiple scattering events. Our goal is to accelerate the rendering of heterogeneous participating media by exploiting redundancy across views, for instance when rendering animated camera paths, motion blur in consecutive frames or multi-view images such as lenticular or light-field images. This poses a challenge as existing methods for sharing light paths across views cannot handle heterogeneous participating media and classical estimators are not optimal in this context. We address these issues by proposing three key ideas. First, we propose new volume shift mappings to transform light paths from one view to another within the recently introduced null-scattering framework, taking into account changes in density along the transformed path. Second, we generate a shared path suffix that best contributes to a subset of views, thus effectively reducing variance. Third, we introduce the multiple weighted importance sampling estimator that benefits from multiple importance sampling for combining sampling strategies, and from weighted importance sampling for reducing the variance due to non contributing strategies. We observed significant reuse when views largely overlap, with no visible bias and reduced variance compared to regular path tracing at equal time. Our method further readily integrates into existing volumetric path tracing pipelines.

CCS Concepts
• Computing methodologies → Computer graphics; Ray tracing; Visibility;

Keywords: Path tracing, Path reusing, Path sharing, MIS

1. Introduction

Taking advantage of spatial or temporal coherence by reusing the same samples in similar rendering configurations (stereo, temporal sequences, camera arrays, etc.) has proved to drastically re-

duce rendering times in the context of offline Monte Carlo integration [Sch19, FINB19, HBGM11, MFSSK06, HDMS03, AH95, AH93]. Indeed such techniques share parts of paths that are difficult to obtain among several views to amortize the cost of sampling and evaluating the entire paths. To the best of our knowledge reusing paths has never been proposed for heterogeneous volumetric light transport. Building upon recent work for rendering heterogeneous

participating media [MGJ19], we introduce a novel way of transforming and reusing null scattering chains to leverage volumetric path reusing in the context of rendering multiple views.

Multi-view rendering consists in finding some point within the scene visible from one camera, which we refer to as *pivot* point, and reconnecting it towards other cameras for which the pivot is also visible. The chain of events connecting the camera to the pivot point forms the path *prefix*. The key idea of the technique is to mutualize the construction and the evaluation of the subpath following the pivot – the path *suffix*. As the number of cameras increases, the construction cost of the path suffix is amortized as it is massively reused among cameras.

Once a *base* prefix has been sampled from one camera we express the new *shifted* prefixes, that connect the pivot to other cameras, as deterministic transformations (i.e. *shift mappings*) of the base prefix. Existing solutions to multi-view rendering require knowing the exact probability density function (pdf) of the path prefixes in order account for the associated change in density. We make use of a closed-form pdf expressed by Miller et al. for heterogeneous participating media [MGJ19], and design a new way of transforming chains of delta tracking events [NGHJ18] from one camera to another.

Classical unbiased *multiple importance sampling* (MIS) can be used to accumulate and weight all contributions of these base and shifted paths. However, these estimators are not optimal for variance reduction when the pdfs of transformed prefixes are very different from the base ones (see Sec. 4). To further reduce the variance, we hence combine MIS with *weighted importance sampling* (WIS) estimators [BSW00] to obtain a new *multiple weighted importance sampling* (MWIS) estimator. This estimator allows our technique to remain consistent, with a rapidly vanishing bias that is imperceptible in practice, even at low sample count.

Our pipeline consists of three main contributions. First, we introduce a path transformation between multiple views – along with its Jacobian – that benefits from Miller et al. null scattering framework for heterogeneous participating media [MGJ19]. Second, we importance sample path suffixes in a way that they best contribute to multiple views. Third, we develop the MWIS estimator to combine contributions from each view in a consistent way with lower variance than MIS. Our overall pipeline achieves significant reuse, in that the number of samples reused in one frame linearly increases with the number of views. We showcase our technique on videos of dynamic scenes with motion blur, lenticular images, lightfield images and printed holographic stereograms (see Fig. 7). Our pipeline also achieves significant speedups for wider baselines, which we demonstrate with animated camera paths on static scenes (see Fig. 8). Finally, the noise within our renderings is consistent across views which improves temporal consistency.

2. Related work

Light transport in participating media. The light transport integral [Kaj86, VG97], its extension to participating media [PKK00] and null scattering [MGJ19] formulate the estimate of a pixel j as

the integral of all light contributions reaching the pixel:

$$I_j = \int_\Omega f_j(\bar{\mathbf{x}}) \mathrm{d}\mu(\bar{\mathbf{x}}) \tag{1}$$

where Ω is the union of all light paths $\bar{\mathbf{x}}$ of finite lengths, $\mathrm{d}\mu$ is the product of the measures of the differential area/volume elements at each path vertex, f_j is the *measurement contribution function* (MCF) of the path. The *MCF* is the product of the filter response W_j, the transmittance T, the generalized scattering function f_s (bidirectional scattering distribution function or phase function), the geometric terms G and the emitted radiance L_e between a camera vertex x_0 and a light vertex x_k:

$$f_j(\bar{\mathbf{x}}) = W_j(x_0, x_1) G(x_0, x_1) T(x_0, x_1) L_e(x_k, x_{k-1}) \cdot$$
$$\left[\prod_{i=1}^{k-1} f_s(x_{i-1}, x_i, x_{i+1}) G(x_i, x_{i+1}) T(x_i, x_{i+1}) \right]. \tag{2}$$

Notations can be found in the work of Novak et al. [NGHJ18], and are reproduced in Appendix A for completeness.

Monte Carlo estimation. Monte Carlo integration consists in numerically evaluating the integral of a function f by sampling it at random locations $\{\bar{\mathbf{x}}_i\}_i$ where the $\bar{\mathbf{x}}_i$ follow some probability density function (pdf) p. The resulting estimator is $\hat{I}_n = \frac{1}{n} \sum_i^n \frac{f(\bar{\mathbf{x}}_i)}{p(\bar{\mathbf{x}}_i)}$, where n is the total number of samples. This tool has been made popular within computer graphics for its ability to evaluate the light transport integral, via methods such as path tracing.

Such a stochastic approach results in noisy estimates, unless a large number of samples is used. A challenge is thus to reduce the variance of this estimator, effectively reducing noise, notably for rendering purposes.

Reusing paths. Several authors proposed solutions to reuse parts of paths across several pixels integrals, by transforming them, to amortize the cost of construction and evaluation of the reused subpaths. Given a base path $\bar{\mathbf{x}}$ with joint pdf $p(\bar{\mathbf{x}})$ and a bijective transformation S – the *shift mapping* – the base path is transformed such that $\bar{\mathbf{x}}' = S(\bar{\mathbf{x}})$. The pdf of a transformed sample $\bar{\mathbf{x}}'$ takes into account the change of density through the Jacobian determinant of the transformation, $p(\bar{\mathbf{x}}') = p(\bar{\mathbf{x}}) \left| \det\left(\frac{\mathrm{d}S(\bar{\mathbf{x}})}{\mathrm{d}\bar{\mathbf{x}}} \right) \right|^{-1} = p(\bar{\mathbf{x}}) \left| S'(\bar{\mathbf{x}}) \right|^{-1}$.

These transformations have been used in several applications, such as *discrete path reusing* [BSH02, XS07, BPE17, BWP*20] and *path space filtering* [KDB14, BFK19, WGGH20], that share several paths suffixes for a set of prefixes using *path reconnection*. Gradient domain rendering approaches sample and estimate gradients as well as an estimate of the image and perform a Poisson reconstruction [GHV*18, BPE17, MKD*16, KMA*15, LKL*13, HGP*19]. The construction of correlated gradients paths uses *pixel shifts* – i.e. that transforms film positions in the image, specular chains shifts such as *manifold exploration* [JM12] or *half vector copy* [KMA*15] as well as path reconnections. Another class of methods additionally uses *lens shifts* – i.e. that transforms the lens interaction of a prefix – to either supersample depth of field on a single camera [Sch19], or to connect the prefix to another camera lens in the context of multiple view rendering of surfaces [HBGM11, MFSSK06, HDMS03, AH95, AH93, FINB19, Sch19]. For camera and

scene animation, paths can be reused along the time dimension [Sch19, MKD*16] using *time shifts*. Schwarzhaupt reevaluates occlusions due to object or camera motion, but not the shading to remain tractable [Sch19]. More recently, a method for real-time sub-path re-use has been proposed for volume rendering [LWY21], but the proposed estimator is not consistent or precludes combining direct illumination techniques which is possible with the framework from Miller et al. [MGJ19]. Another approach makes use of primary space control variates that approximate the integrand with polynomials for path tracing [CJMn21], but this is limited to low-dimensional integrals. Leimkühler et al. propose a warping approach to efficiently generate multiple views but this technique is limited to surfaces, do not re-evaluate view dependent shading and work best with very close views [LSR17].

Null-scattering framework. Until recently, no path formulation provided the closed-form pdf associated with the *free-flight-distance sampling* – the probability density of light interacting with the medium after a certain distance along the line of flight – and transmittance evaluation of ordinary heterogeneous participating media (see the survey of Novak et al. [NGHJ18]). Miller et al. addressed this issue within a *null scattering framework* that takes into account fictitious (null) particles in the medium [MGJ19]. They render heterogeneous media with the help of an upper bound $\bar{\mu}$ of the extinction coefficient μ_t, and the transmittance $\bar{T}$ of the combined homogenous medium which has a trivial closed-form expression, $\bar{T}(x, y) = \exp(-\bar{\mu} \cdot \|y - x\|)$.

We propose new shift mappings that transforms these null particles in path space from one base path traversing media towards other shifted paths. We demonstrate their benefit by extendind multi-view rendering to support dense heterogeneous media, by combining base and transformed paths within a single pixel estimator.

Combining sampling techniques. Several sampling strategies, with pdfs $\{p_k\}_k$, can be used and combined to form a single estimator – using the multiple importance sampling framework [VG95]. Weighting these strategies to obtain a single value of lowest possible variance is a difficult task [KVG*19] – in fact, poor weights can actually increase variance. A simple choice, the *balance heuristic*, weights each sample by the relative contribution of its own pdf, $w_k(x_i) = \frac{n_k p_k(x_i)}{\sum_l n_l p_l(x_i)}$, with n_k the number of samples for the strategy k of pdf p_k, which is optimal among positive weights for *one-sample* estimators [VG95]. Kondapaneni et al. proposed optimal weights that are not constrained to be positive and that provably minimize the variance of the *multi-sample* estimators [KVG*19] of the form $\sum_k \sum_i w_k(x_i) \frac{f(x_i)}{n_k p_k(x_j)}$. However, evaluating these optimal weights requires $m \times m$ integrations for m strategies.

In the context of multiple view rendering, early approaches designed for surfaces and diffuse materials made assumptions potentially introducing bias and/or variance [HBGM11, MFSSK06, HDMS03, AH95, AH93]. Recently, both Fraboni et al. [FINB19] and Schwarzhaupt [Sch19] propose to reduce these artifacts by rejecting degenerated transformations, thus limiting Jacobian variations and the variance introduced in the MIS combination. Still, these approaches do not produce the best possible variance reduction, due to the use of a multi-sample estimator. In our context,

each pixel is considered as a potential sampling strategy, making the optimal MIS approach [KVG*19] intractable with millions of such strategies. The stochastic variant of continuous MIS (SMIS) [WGGH20] is also not applicable for multi-view rendering due to the random selection of strategies, which will fail at selecting pixels that are likely to share paths.

An additional source of variance in the MIS combination comes from the fact that shifted paths do not follow the base pdf – i.e. the pdf of sampling the path from the pixel directly – which is in fact the ideal pdf regarding the terms of the integrand. Fortunately, a biased but consistent technique called *weighted importance sampling* (WIS) has been proposed to simulate the use of a known pdf p, the *target* pdf, that is good regarding the integrand, with a set of samples drawn from an arbitrary pdf q, the *source* pdf. This technique has been used in rendering applications [HHM18, BSW00, Kel96]. Bitterli et al. showed that the *resampled importance sampling* (RIS) estimator [TCE05] is a bias corrected version of WIS [BWP*20]. However resampling prefix candidates would be intractable in our context as paths are unbounded.

We combine MIS and WIS weights to build a self normalizing multi-strategy estimator that benefits from both techniques. Therefore, a sample considered poor with a low source pdf value could in fact be good for the integrand with a high target pdf value. This escapes Veach's theoretical bounds for optimal weights as poor samples do not increase the variance of the estimator anymore if p is well chosen.

3. Multiple View Path Construction

We first generate a ray by importance sampling a camera shutter, lens and film. If a medium boundary is found, delta tracking is used to sample interactions distances across the medium and event sampling to determine the interactions types [NGHJ18, MGJ19]. This method samples a homogenized version of the participating medium. The majorant extinction coefficient of the homogenized medium is an upper-bound of the real one expressed as the sum of the real extinction and a fictious complementary one. The sampled interaction is then considered *null* with probability equal to the ratio between fictious and homogenized extinctions. If a null interaction is sampled, we continue the sampling routine until either we leave the medium or a real interaction is found.

Finally, the complete path prefix is a *chain* of random events: film interaction, lens interaction, $\{0, \dots, m\}$ null interactions in the medium, and a last interaction either on a surface or a real medium interaction. This last interaction, which we will refer to as the *pivot*, is the key to our work since it can be visible from other points of view. Before constructing the path suffix, we connect every camera to the pivot by shifting (i.e. transforming) all the events composing this *base* prefix to form *shifted* prefixes, as we describe next.

3.1. Prefix shift

Once a base prefix is generated and the pivot is found, we try to build shifted prefixes that connect the pivot to other cameras, the *targets* as depicted in Fig. 2. The shifted chain must have the same number of random events as the base one, in order to be defined

on the same probability space. We start by shifting the base lens interaction to the lens of a target camera. We then check that the pivot is not occluded and falls within the target camera view frustum and exposure time. Next, the remaining sampled events along the base prefix are shifted to build another prefix connecting a new lens event to the pivot interaction. A ray is traced from the target lens to the pivot to collect the target medium sections. If only one of either the base or shifted prefix crosses medium sections, we cannot construct a valid shift because the number of random events differs between the two chains. This can happen on the volume silhouettes and may lead to undersampling in these areas.

If neither of the prefixes cross any medium section, the shift is complete as there is no medium interaction, but only a surface interaction at the pivot that remains unchanged.

Our goal is then to define a bijective transformation of the base medium interactions, such that the shifted medium interactions fall within the target medium sections of the target segment. The medium interaction depths can thus be mapped back and forth to unique and valid interactions on base and target prefixes.

We also need the Jacobian of the transformation in closed form to correctly account for the change in density. All required transformations are described next. In the following, index i denotes the pixel used to generate the base prefix and index j denotes another pixel and its shifted prefix.
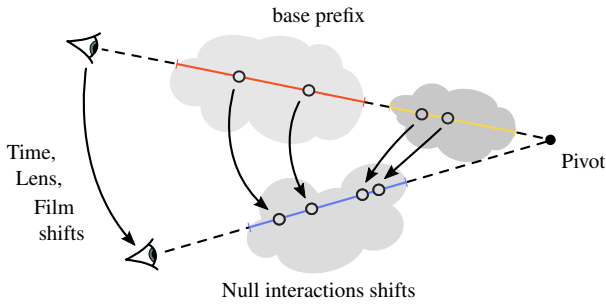

Figure 2: *Prefix chains shift configuration*

Shutter time shift. We construct paths at fixed time since dealing with object motion would require re-evaluating occlusions and shading on the complete path due to moving geometry – the cost of the time shift would be equivalent to constructing a full path. Hence the base shutter time sampled within the camera shutter interval is reused $x'_{\text{time}} = x_{\text{time}}$, and the Jacobian of the shift is $\left| S'_{i \to j}(x_{\text{time}}) \right| = 1$.

Lens shift. The lens event is shifted by reusing the random numbers that sampled the base lens interaction [KSKAC02, HGP*19], used similarly in the work of Schwarzhaupt [Sch19] and Fraboni et al. [FINB19]. It consists in transforming the point on the camera lens aperture (a 2D shape) towards another camera lens aperture. Assuming that both apertures are disk-shaped, and given a base lens interaction x_{lens}, this results in the simple transformation:

$$x'_{\text{lens}} = \frac{r'}{r} \cdot x_{\text{lens}} \quad \text{with Jacobian} \quad \left| S'_{i \to j}(x_{\text{lens}}) \right| = \frac{r^2}{r'^2} \quad (3)$$

where r and r' are base and target aperture radii.

Film shift. After shifting the lens event, the film position is then constrained by the refraction of the ray connecting the lens to the pivot. Hence, we have to evaluate the position of the shifted film interaction x'_{film} using x'_{lens} and x'_{pivot}, as in light tracing or bidirectional methods. Finally, the change of measure between the base film interaction and the shifted one is accounted with the following Jacobian [CTE05, LKL*13]:

$$\left| S'_{i \to j}(x_{\text{film}}) \right| = \frac{g(x_{\text{lens}}, x_{\text{pivot}})}{g(x_{\text{lens}}, x_{\text{film}})} \cdot \frac{g(x'_{\text{lens}}, x'_{\text{pivot}})}{g(x'_{\text{lens}}, x'_{\text{film}})} \quad (4)$$

where $g(a,b) = \frac{D(a,b)}{\|b-a\|^2}$ and $D(a,b) = |\vec{\omega}_{ab} \cdot \vec{n}_b|$ if b is on a surface else $D(a,b) = 1$.

Null interactions shift. Shifting medium interactions consists in transforming the depth t of an interaction sampled using *delta tracking* [NGHJ18, MGJ19] on the base prefix segment, to another valid depth location t' on the shifted segment. A valid shift location should ensure that the pdf of the shifted interaction is well defined – i.e. inside a medium.

Our null interactions shift accounts for the *majorant optical depth* $\bar{\tau}$ along the ray segment between camera and pivot. The majorant optical depth is normalized along the base ray segment and reused on the target ray segment. Thus, by construction, the shifted interactions are ensured to fall within the target medium sections. On the contrary, simpler shifts such as random numbers replay generate interactions that may fall outside the medium. Note that we use the majorant optical depth because the true optical depth does not in general have a closed form expression, or has one that is too complex to inverse. On the contrary, the majorant optical depth is piecewise linear on the medium sections and is thus always invertible in closed form.

The majorant optical depth $\bar{\tau}$ is obtained by integrating the majorant extinction coefficient $\bar{\mu}$ from the medium entry to the interaction position:

$$\bar{\tau}(t) = \int_0^t \bar{\mu}(x)\mathrm{d}x. \quad (5)$$

As the majorant coefficient is constant over medium sections and zero outside, the majorant optical depth is piecewise linear. It thus can be obtained by summing over the medium sections crossed along the prefix until reaching the interaction at depth t:

$$\bar{\tau}(t) = \sum_k^{n_t} \bar{\mu}_k \left(\min\left(t, t_{\max,k} \right) - t_{\min,k} \right)$$
$$= \bar{\mu}_{n_t} \left(t - t_{\min,n_t} \right) + \sum_k^{n_t - 1} \bar{\mu}_i \left(t_{\max,k} - t_{\min,k} \right) \quad (6)$$

where $t_{\min,k}$ and $t_{\max,k}$ are the entry and exit depths of the k-th volume section and n_t is the index of the volume section containing depth t. In case of overlapping media, we replace the unique majorant by the sum of majorants in overlapping sections.

The mapping is then defined as the copy of the normalized majorant optical depths along the base and target rays segments:

$$\frac{\bar{\tau}'(t')}{\bar{\tau}'_{\max}} = \frac{\bar{\tau}(t)}{\bar{\tau}_{\max}} \quad \Leftrightarrow \quad \bar{\tau}'(t') = \frac{\bar{\tau}'_{\max}}{\bar{\tau}_{\max}} \bar{\tau}(t) \quad (7)$$

where $\bar{\tau}_{\max}$ and $\bar{\tau}'_{\max}$ are the total majorant optical depths along the base and shifted prefixes. From Eq. 7 we can express the distance t' on the target segment as a function of t the distance on the base segment:

$$t' = \frac{\tau'_{\max}}{\tau_{\max}} \cdot \frac{\bar{\mu}_{n_t}}{\bar{\mu}'_{n_{t'}}} \cdot t + R \qquad (8)$$

where R is a residual that does not depend on t nor t', and n_t is the volume section index containing depth t on the base segment (resp. $n_{t'}$ for depth t' on the target segment), see Appendix B for the complete derivation. The Jacobian determinant of the transformation is:

$$\left|S'_{i\to j}(t)\right| = \frac{\tau'_{\max}}{\tau_{\max}} \cdot \frac{\bar{\mu}_{n_t}}{\bar{\mu}'_{n_{t'}}}. \qquad (9)$$

We apply this transformation for each base null interaction. In the presence of a single medium section on the base and target segments, this mapping corresponds to a linear scale of the majorant optical depth. We discuss two other volume shift mappings in supplementary materials.

Pivot shift / copy. The last interaction where the base and the shifted prefix join can either be on a surface or in a medium but its location stays fixed in space, hence $x'_{\text{pivot}} = x_{\text{pivot}}$ and $\left|S'_{i\to j}(x_{\text{pivot}})\right| = 1$.

Jacobian and pdf evaluation. The complete shift of base prefix $\bar{\mathbf{x}}_i$ from pixel i to pixel j generates a new chain $\bar{\mathbf{x}}_j$ of m interactions that connects the camera to the pivot. The Jacobian of the chain transformation is the product of each independent shift Jacobian:

$$\left|S'_{i\to j}(\bar{\mathbf{x}}_i)\right| = \prod_{k=1}^{m} \left|S'_{i\to j}(x_{i,k})\right|. \qquad (10)$$

Finally the pdf of the path $\bar{\mathbf{x}}_i$ from pixel i transformed to path $\bar{\mathbf{x}}_j$ for pixel j is given by:

$$p_{i\to j}(\bar{\mathbf{x}}_j) = c_i p_i(\bar{\mathbf{x}}_i) \left|S'_{i\to j}(\bar{\mathbf{x}}_i)\right|^{-1}. \qquad (11)$$

where c_i denotes the relative number of samples for strategy i. We additionnally compute the ideal pdf – i.e. as if the path $\bar{\mathbf{x}}_j$ was sampled directly from pixel j, which reads:

$$p_{j\to j}(\bar{\mathbf{x}}_j) = c_j p_j(\bar{\mathbf{x}}_j) \underbrace{\left|S'_{j\to j}(\bar{\mathbf{x}}_j)\right|^{-1}}_{=1} = c_j \prod_{k=1}^{m} p_j(x_{j,k}). \qquad (12)$$

Both pdfs are further used to compute combination weights accounting for multiple strategies (see Sec. 4).

3.2. Computing path suffix

After finding a base pivot point and constructing the possible prefixes by shifting all events along the base prefix, we can evaluate direct and indirect illumination at once for all cameras with successful shifts.

Direct illumination. The direct part is simple in the sense that sampling light sources does not generally depends on prefix directions, using for example *next event estimation*. We hence apply regular light source sampling: we generate a light sample, compute

the common part of the contribution for all prefixes, and finally multiply it by the scattering function and the path throughput associated to each prefix j independently. We can further improve the results by combining multiple direct illumination sampling techniques [MGJ19].

Indirect illumination sampling. Indirect illumination is more complex. We need to sample a single scattering direction to start building the shared path suffix, but the scattering distribution function at the pivot interaction is generally view-dependent. This is notably the case for mostly glossy and perfect specular materials, and highly anisotropic medium phase functions.

We could sample the indirect direction w.r.t the base prefix scattering function or w.r.t to a uniform mixture of the available prefixes scattering functions. Sampling from the base prefix only, without accounting for other cameras for which the pivot point is visible [HDMS03, MFSSK06, Sch19] tends to produce high variance spikes in the images in our experiments. Instead, we choose the second approach that considers all the cameras that could have generated the suffix path, which proved to be more robust in practice.

The next section discusses two ways to combine the base and shifted prefixes contributing to a pixel.

4. Combining base and shifted samples

All base and shifted paths can be combined with a standard unbiased MIS multi-sample estimator as summarized in Alg. 1. This is a multi-sample estimator which incrementally accumulates weighted base and shifted contributions for each pixel.

Algorithm 1 Multi-view path tracing with classical MIS. Note that line 16 accumulates multi-sample MIS weighted contributions in the pixel, but the estimator is not progressive. A full pass over all pixels of all cameras is required to get a complete pixel value.

1: **for** each pixel i **do**
2:     sample base prefix $\bar{\mathbf{x}}_i$ (Sec. 3)
3:     $S_r \leftarrow 0$
4:     **for** each camera k **do**
5:         *// shift the base path $\bar{\mathbf{x}}_i$ to camera k through pixel j*
6:         $\bar{\mathbf{x}}_j \leftarrow \text{shift}(\bar{\mathbf{x}}_i, k)$ $\triangleright$ eval the shifted prefix $\bar{\mathbf{x}}_j$ (Sec. 3.1)
7:         $p_{i\to j}(\bar{\mathbf{x}}_j) \leftarrow \frac{c_i \cdot p_i(\bar{\mathbf{x}}_i)}{|S'_{i\to j}|}$ $\triangleright$ eval the shifted pdf (Eq. 11)
8:         $p_{j\to j}(\bar{\mathbf{x}}_j) \leftarrow c_j \cdot p_j(\bar{\mathbf{x}}_j)$ $\triangleright$ eval the ideal pdf (Eq. 12)
9:         $S_r \leftarrow S_r + \frac{p_{j\to j}(\bar{\mathbf{x}}_j)}{p_{i\to j}(\bar{\mathbf{x}}_j)}$
10:     **end for**
11:     sample light sources and compute direct (Sec. 3.2)
12:     sample a path suffix and compute indirect (Sec. 3.2)
13:     **for** each prefix $\bar{\mathbf{x}}_j$ associated with pixel j **do**
14:         *// eval MIS weights and accumulate contributions*
15:         $w_i(\bar{\mathbf{x}}_j) \leftarrow \frac{1}{S_r}$ $\triangleright$ eval the weight
16:         $I_j \leftarrow I_j + w_i(\bar{\mathbf{x}}_j) \cdot \frac{f_j(\bar{\mathbf{x}}_j)}{p_{i\to j}(\bar{\mathbf{x}}_j)}$
17:     **end for**
18: **end for**

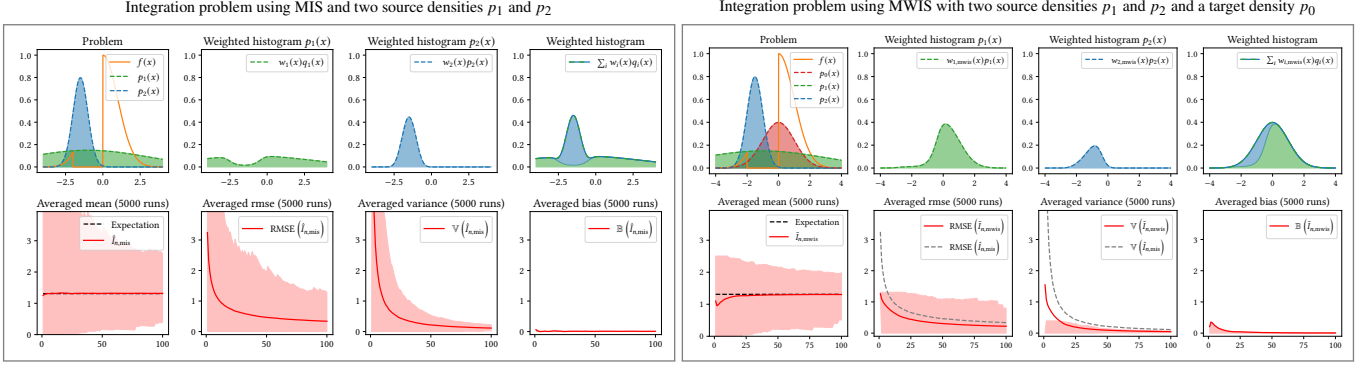However, it suffers from two main sources of variance.

Figure 3: *Difference between MIS (left) and MWIS (right) to integrate f, shown in orange, using two distributions p_1 and p_2. Given an a-priori distribution p_0 (red) that better fits the integrand, MWIS reweights samples such that the weighted sample histogram perfectly reproduce the target density p_0. This results in a reduction of the error and variance, at the expense of a slight bias that vanishes in $O(n^{-1})$. We exploit MWIS to reweight shifted samples as if they were directly sampled from the pixel they contribute to.*

First, the distribution of shifted prefixes does not follow the ideal importance sampling for a given pixel. Hence, the mixture density resulting from MIS estimators is not optimal in term of variance reduction. Fig. 4 shows example scenes with high re-use between cameras. In this case, the MIS estimator does not produce the best possible variance reduction. Variance can be reduced with a self-normalized estimator (see next) that reweights samples w.r.t to a target pdf closer to the integrand, which in our case is the base pdf.

As a toy example, consider two pixels A and B and their respective integrands f_A and f_B. We sample pixel A with a single sample a following pdf $p_A \propto f_A$, and B with a sample b following $p_B \propto f_B$. When integrating the function f_A over A:

- the MIS balance estimator reads:

$$I_A = \frac{f_A(a)}{p_A(a) + p_B(a)} + \frac{f_A(b)}{p_A(b) + p_B(b)}$$

- the estimator that we introduce in Sec. 4.1 reweights sampling strategies as:

$$I_A = \left(w_A(a) \frac{f_A(a)}{p_A(a)} + w_B(b) \frac{f_A(b)}{p_A(b)} \right) \cdot \frac{1}{w_A(a) + w_B(b)}$$

where w_A and w_B are such that the integral is consistent. In this case, we are summing up values of the form $\frac{f_A}{p_A}$, which is ideal when p_A is (near) proportional to f_A. This reweighting produces less variance than MIS that sums contributions of the form $\frac{f_A}{p_A + p_B}$, which is not ideal when p_B is far from p_A (see also Fig. 3).

Second, reusing all paths in all situations is another source of variance, due to the potentially poor importance sampling of view-dependent scattering functions. Fig. 5 shows examples of such cases on glossy materials and anisotropic phase functions where MIS exhibits spikes and higher noise than the baseline. Addressing this issue implies selecting cameras that are more likely to share suffix paths, i.e. which can jointly benefit from importance sampling. However selecting a subset of similar prefixes within the MIS framework is non trivial, as the number of samples per strategy has to be fixed beforehand and can not be modified on the fly without biasing the result. MIS allows nonetheless to either keep all prefixes, or discard all of them (deterministically or stochastically). In

constrast WIS estimators are self normalized and allow more flexibility in choosing paths to re-use and to select similar shifted prefixes w.r.t to the base prefix. This selection is further discussed in Sec. 4.3. However, WIS is not designed to handle multiple source strategies.

4.1. Multiple weighted importance sampling

We extend standard WIS [PS66, Spa79, BSW00] to combine several sampling strategies of *source* pdfs $q_1, \ldots, q_k$ and reweight the samples using a *target* pdf p, that is chosen to be a good strategy w.r.t the integrand.

Weighted importance sampling. Given two sampling strategies with pdf p and q, the Monte Carlo integral and expectation w.r.t p can be written as:

$$I = \int_{\Omega} f(\bar{\mathbf{x}}) d\mu(\bar{\mathbf{x}}) = \mathbb{E}\left[\frac{f(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})}\right] = \mathbb{E}\left[\frac{f(\bar{\mathbf{x}})}{q(\bar{\mathbf{x}})} \frac{q(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})}\right] \quad (13)$$

which can be further approximated as follows:

$$I \approx \frac{\mathbb{E}\left[\frac{f(\bar{\mathbf{x}})}{q(\bar{\mathbf{x}})}\right]}{\mathbb{E}\left[\frac{p(\bar{\mathbf{x}})}{q(\bar{\mathbf{x}})}\right]} = \frac{\mathbb{E}\left[w(\bar{\mathbf{x}}) \frac{f(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})}\right]}{\mathbb{E}[w(\bar{\mathbf{x}})]} \quad \text{with} \quad w(\bar{\mathbf{x}}) = \frac{p(\bar{\mathbf{x}})}{q(\bar{\mathbf{x}})}. \quad (14)$$

where samples x are drawn according to q. The latter equation corresponds to the WIS [PS66, Spa79, BSW00] or *self-normalized importance sampling* (SNIS) [Owe13] integral and can be evaluated with the following *ratio estimator*:

$$\tilde{I} = \frac{\sum_i^n w(\bar{\mathbf{x}}_i) \frac{f(\bar{\mathbf{x}}_i)}{p(\bar{\mathbf{x}}_i)}}{\sum_i^n w(\bar{\mathbf{x}}_i)}. \quad (15)$$

The WIS estimator is biased (as $\mathbb{E}\left[\frac{X}{Y}\right] \geq \frac{\mathbb{E}[X]}{\mathbb{E}[Y]}$) but consistent since the bias tends to zero with rate $O(n^{-1})$ (i.e. it is asymptotically unbiased) [Spa79, PS66]. Note that the WIS weight w can be known up to a constant factor.
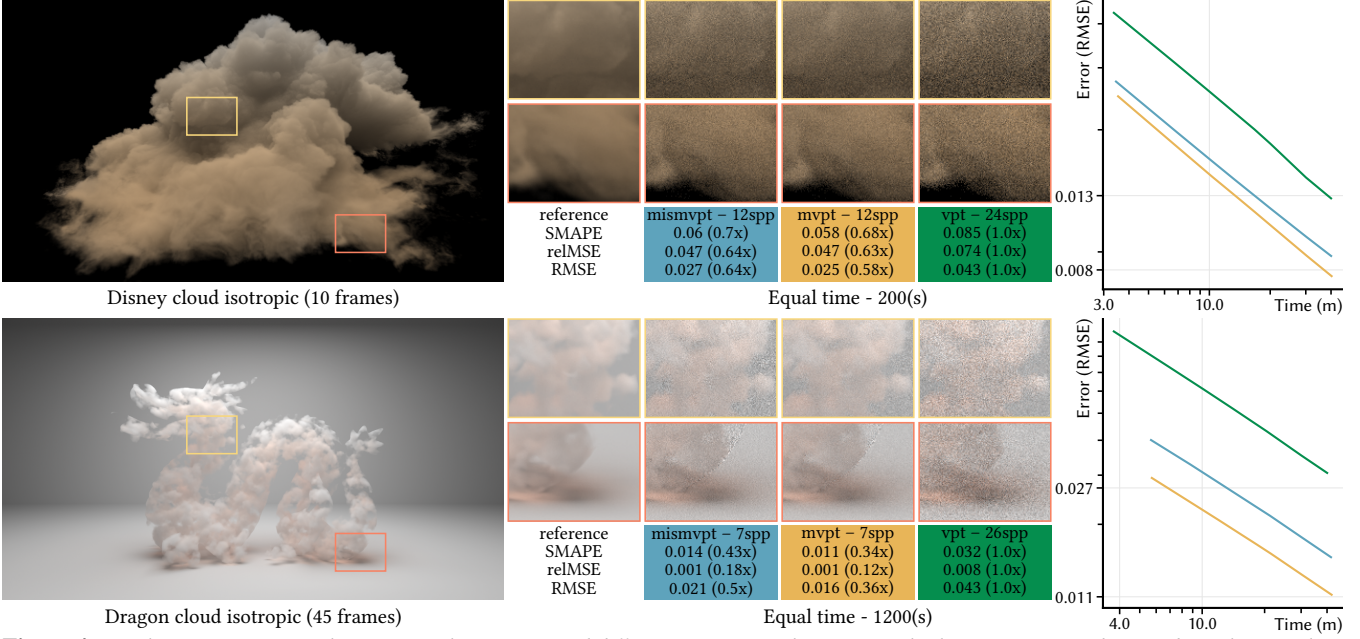
Figure 4: *We show two toy example scenes with isotropic and diffuse scattering only. Hence only the estimator used to combine the sampling strategies makes a difference in term of variance reduction. Even though multi-view path tracing using an MIS estimator (mismvpt) reduces the variance w.r.t the baseline (vpt) at equal rendering time, another family of consistent WIS estimators (mvpt) allows to further reduce the variance of combined samples (base and shifted). Even directly visible surfaces (dragon – bottom crop) see their variance reduced.*
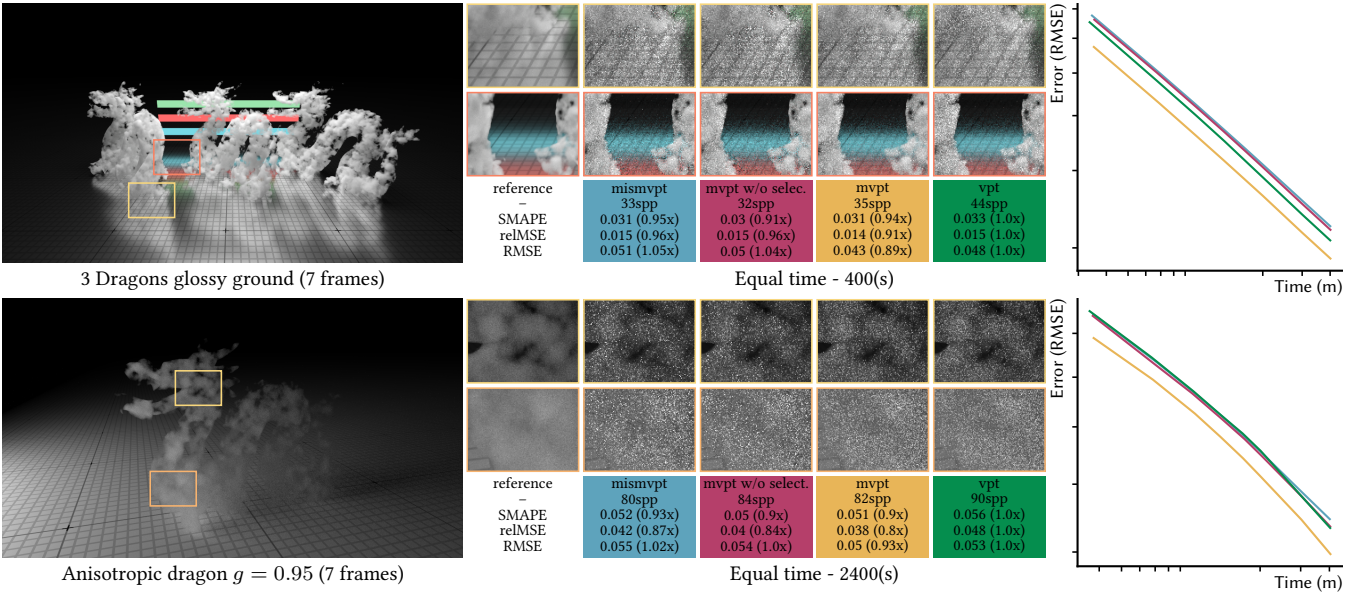

Figure 5: *The 3 Dragons scene includes a large glossy surfaces and the anisotropic dragon scene includes a very strong forward scattering medium lit from the side, for which we expect less or no re-use. Forcing to reuse samples between cameras (mismvpt and mvpt w/o selection) in such situations increases the variance, resulting in higher noise than the baseline (vpt) at equal time. In comparison, selecting subsets of prefixes that are likely to share suffix paths like in our (mvpt) removes these poor contributions to match at worst the baseline (i.e. no reuse).*

Our estimator. We rewrite the WIS integral and modify the weighting functions in Eq. 14 to both take into account multiple source strategies and the chosen target strategy.

The MWIS integral $I_{\mathrm{mwis}} = \sum_{i=1}^{k} I_{i,\mathrm{wis}} \cdot W_i$ consists of a weighted sum of WIS integrals associated with each strategy, $I_{i,\mathrm{wis}}$, with respective weight W_i. Denoting $w_{i,\mathrm{mwis}}(\bar{\mathbf{x}})$ the weighting function associated with strategy i (we postpone the discussion of the choice

of the ideal $w_{i,\text{mwis}}$), we have:

$$
I_{\text{mwis}} = \sum_{i=1}^{k} \frac{\mathbb{E}_{q_i}\left[w_{i,\text{mwis}}(\bar{\mathbf{x}})\frac{f(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})}\right]}{\mathbb{E}_{q_i}\left[w_{i,\text{mwis}}(\bar{\mathbf{x}})\right]} \cdot \frac{\mathbb{E}_{q_i}\left[w_{i,\text{mwis}}(\bar{\mathbf{x}})\right]}{\sum_{j=1}^{k}\mathbb{E}_{q_j}\left[w_{j,\text{mwis}}(\bar{\mathbf{x}})\right]}
$$
$$
= \frac{\sum_{i=1}^{k}\mathbb{E}_{q_i}\left[w_{i,\text{mwis}}(\bar{\mathbf{x}})\frac{f(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})}\right]}{\sum_{i=1}^{k}\mathbb{E}_{q_i}\left[w_{i,\text{mwis}}(\bar{\mathbf{x}})\right]} \tag{16}
$$

where $\mathbb{E}_{q_i}[x] = \int x q_i(x)dx$. Estimating this integral via Monte Carlo leads to the following ratio estimator:

$$
\tilde{I}_{\text{mwis}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} w_{i,\text{mwis}}(\bar{\mathbf{x}}_{ij})\frac{f(\bar{\mathbf{x}}_{ij})}{p(\bar{\mathbf{x}}_{ij})}}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} w_{i,\text{mwis}}(\bar{\mathbf{x}}_{ij})} \tag{17}
$$

where strategy i is selected to generate n_i samples x_{ij}. Note that if the sets of samples used in the numerator and the denominator are identical, the MWIS estimator turns into a weighted mean of samples. Although this is not required for MWIS to work, we use the same samples for practical reasons.

Weighting heuristics. For each sample x, the function $w_{i,\text{mwis}}(\bar{\mathbf{x}})$ serves two purposes. First it reweights the source density w.r.t the target density – i.e. comparably to WIS. This is desirable since strategy i has been sampled with density q_i but the contribution is evaluated as if the sample was drawn from the target density p. Second, it appropriately combines the multiple strategies – i.e. comparably to MIS – since strategy i has been selected among k available strategies to generate the sample $\bar{\mathbf{x}}$. Our MWIS weighting functions thus takes the following form:

$$
w_{i,\text{mwis}}(\bar{\mathbf{x}}) = w_{i,\text{mis}}(\bar{\mathbf{x}}) \cdot w_{i,\text{wis}}(\bar{\mathbf{x}}) \tag{18}
$$

which is a valid weighting heuristic within the MWIS framework thanks to the self-normalization of the estimator. For this estimator to converge, it is sufficient that $p(\bar{\mathbf{x}}) > 0$ whenever $f(\bar{\mathbf{x}}) > 0$ and $\exists i, q_i(\bar{\mathbf{x}}) > 0$ whenever $p(\bar{\mathbf{x}}) > 0$. The WIS-related part of the weight is constrained by the source and target pdf used, hence:

$$
w_{i,\text{wis}}(\bar{\mathbf{x}}) = \frac{p(\bar{\mathbf{x}})}{c_i q_i(\bar{\mathbf{x}})} \tag{19}
$$

where q_i is the density that generated the sample $\bar{\mathbf{x}}$, and c_i is the probability of selecting the i-th strategy – i.e. the relative number of samples per strategy. The MIS-related part on contrary can be any heuristic following the MIS weighting conditions. Hence, the MIS balance heuristic [VG95] is a good choice that takes into account the local repartition of samples:

$$
w_{i,\text{mis}}(\bar{\mathbf{x}}) = \frac{c_i q_i(\bar{\mathbf{x}})}{\sum_{j=1}^{k} c_j q_j(\bar{\mathbf{x}})}. \tag{20}
$$

Finally, using the balance heuristic the MWIS weighting function reduces to the simple form:

$$
w_{i,\text{mwis}}(\bar{\mathbf{x}}) = \frac{p(\bar{\mathbf{x}})}{\sum_{j=1}^{k} c_j q_j(\bar{\mathbf{x}})}. \tag{21}
$$

The latter equation enlightens the fact that, within balance MWIS, we reweight the samples using a weighted mixture of distributions – i.e. the denominator – which is also the combined distribution of MIS balance estimators. Still, MIS and MWIS weighting functions are different and serve different purposes.

Error and bias. Denoting the *bias* β and the *mean squared error* (mse) ε^2 of the ratio estimator from Eq. 17, the ratio $\frac{\beta^2}{\varepsilon^2}$ is upper bounded, and the bound tends to zero as the number of samples n tends to infinity [PS66]. The bound depends on the variance of the weighting functions, hence as their variance decreases the bias decreases, that is when the source density is closer to the target density. For sufficiently large n, and assuming that the weighting functions are bounded and normalized, the mse is approximately:

$$
\varepsilon^2 \approx \frac{1}{n}\sum_i \int_\Omega w_{i,\text{mwis}}^2(\bar{\mathbf{x}})\left(\frac{f(\bar{\mathbf{x}})}{p(\bar{\mathbf{x}})} - I\right)^2 q_i(\bar{\mathbf{x}})d\bar{\mathbf{x}} \tag{22}
$$

which is also equal to the variance as the bias vanishes. This estimator is related to the multi-sample RIS estimator [BWP*20,TCE05], as a biased but consistent self-normalizing version. We illustrate our MWIS estimator in Fig. 3 in 1D.

4.2. MWIS for muti-view rendering

After generating path prefixes, their associated Jacobians and pdfs, we compute MWIS weights for each of them. We compute the weights at the first visible interaction once and use them to accumulate a single contribution for both direct and indirect illumination. To avoid precision issues we re-write MWIS weights in term of ratios of pdfs and Jacobians. In practice, we compute these ratios event per event as a chain is transformed, since the ratios of final pdfs may suffer precision issues due to large pdfs in dense heterogeneous volumes. Denoting $w_{i \to j}(\bar{\mathbf{x}}_j)$ the weight for path $\bar{\mathbf{x}}_j$ in pixel j that results from transforming path $\bar{\mathbf{x}}_i$ from pixel i, and using Eq. 11, Eq. 12 and the chain rule, we obtain a more numerically stable expression:

$$
w_{i \to j}(\bar{\mathbf{x}}_j) = \frac{\frac{c_j q_j(\bar{\mathbf{x}}_j)}{c_i q_i(\bar{\mathbf{x}}_i)} \cdot |S'_{i \to j}(\bar{\mathbf{x}}_i)|}{\sum_k \frac{c_k q_k(\bar{\mathbf{x}}_k)}{c_i q_i(\bar{\mathbf{x}}_i)} \cdot |S'_{i \to k}(\bar{\mathbf{x}}_i)|} = \frac{\frac{p_{j \to j}(\bar{\mathbf{x}}_j)}{p_{i \to j}(\bar{\mathbf{x}}_j)}}{\sum_k \frac{p_{k \to k}(\bar{\mathbf{x}}_k)}{p_{i \to k}(\bar{\mathbf{x}}_k)}} \tag{23}
$$

see the supplementary document for the complete derivation. We thus efficiently compute and sum up the ratios of source and target pdf multiplied by the Jacobian during the prefix shift step for each candidate pixel. We finally compute MWIS weights when adding sample contributions to the image buffer normalizing the ratios by the sum of ratios. The complete process is summarized in Alg. 2.

4.3. Prefix selection

The path suffix sampling process described in Sec. 4 should not increase the variance of the base pixel, nor of the target pixels. We chose to proceed similarly to Fraboni et al. [FINB19] with stochastic acceptance or rejection of shifted prefixes. We accept a candidate shifted prefix if the associated scattering distribution at the pivot point is similar to that of the base prefix (Fig. 6). Doing so requires to compare scattering distributions, which is a difficult problem. While Fraboni et al. use a heuristic that cannot be used for phase functions, we use an acceptance probability based on a discrete total variation (TV) approximation between pairs of scattering distributions. This TV approximation is evaluated on the fly and produces a well-defined distance for general distributions, including phase functions. We evaluate the TV distance on a small set of directions $\{\omega_i\}_i$, and their density for both distributions, before

Algorithm 2 Multi-view path tracing with MWIS pseudocode. Note that line 17 progressively computes the estimator's sum in (Eq. 17).

1: **for** each pixel i **do**
2:     sample base prefix $\bar{\mathbf{x}}_i$ (Sec. 3)
3:     $S_r \leftarrow 0$
4:     **for** each camera k **do**
5:         *// shift the base path $\bar{\mathbf{x}}_i$ to camera k through pixel j*
6:         $\bar{\mathbf{x}}_j \leftarrow \text{shift}(\bar{\mathbf{x}}_i, k)$ $\triangleright$ eval the shifted prefix $\bar{\mathbf{x}}_j$ (Sec. 3.1)
7:         $p_{i \rightarrow j}(\bar{\mathbf{x}}_j) \leftarrow \frac{c_i \cdot p_i(\bar{\mathbf{x}}_i)}{|S'_{i \rightarrow j}|}$ $\triangleright$ eval the shifted pdf (Eq. 11)
8:         $p_{j \rightarrow j}(\bar{\mathbf{x}}_j) \leftarrow c_j \cdot p_j(\bar{\mathbf{x}}_j)$ $\triangleright$ eval the ideal pdf (Eq. 12)
9:         $S_r \leftarrow S_r + \frac{p_{j \rightarrow j}(\bar{\mathbf{x}}_j)}{p_{i \rightarrow j}(\bar{\mathbf{x}}_j)}$
10:     **end for**
11:     sample light sources and compute direct (Sec. 3.2)
12:     sample a path suffix and compute indirect (Sec. 3.2)
13:     **for** each prefix $\bar{\mathbf{x}}_j$ associated with pixel j **do**
14:         *// eval MWIS weights and accumulate contributions*
15:         $w_{i \rightarrow j}(\bar{\mathbf{x}}_j) \leftarrow \frac{\frac{p_{j \rightarrow j}(\bar{\mathbf{x}}_j)}{p_{i \rightarrow j}(\bar{\mathbf{x}}_j)}}{S_r}$ $\triangleright$ eval the weight (Eq. 21 - 23)
16:         $W_j \leftarrow W_j + w_{i \rightarrow j}(\bar{\mathbf{x}}_j)$ $\triangleright$ denominator (Eq. 17)
17:         $I_j \leftarrow I_j + \frac{w_{i \rightarrow j}(\bar{\mathbf{x}}_j)}{W_j} \cdot \left(\frac{f_j(\bar{\mathbf{x}}_j)}{p_{j \rightarrow j}(\bar{\mathbf{x}}_j)} - I_j \right)$ $\triangleright$ eval (Eq. 17)
18:     **end for**
19: **end for**

Figure 6: *At the pivot point, the scattering function of all prefixes should be similar to the base one to avoid increasing the variance. We use an approximate measure of the shared volume of both distributions (hatched region) to accept or reject the prefix. Our approximation handles both surface and volume scattering distributions.*

normalizing both discrete pdf sets and computing the approximate TV distance as:

$$\text{TV}(p, q, \{\omega_i\}_i) = \frac{1}{2} \sum_i \left| \frac{p(\omega_i)}{\sum_k p(\omega_k)} - \frac{q(\omega_i)}{\sum_k q(\omega_k)} \right|. \quad (24)$$

This approximation is inexpensive for a small number of directions. In practice, we use two directions chosen according to the type of scattering distribution at the pivot interaction. For surface interactions, we use the mirror directions of both viewing directions. For medium interactions, we use the forward directions of both viewing directions if the phase is forward scattering, and the backward directions in case of backward scattering. This approximation overestimates the true TV distance but gives correct estimates for rough single and bi-layered materials (i.e. rough plastics, conductors, dielectrics, coatings) and phase functions with varying anisotropy.

We compare every candidate prefix to the base prefix and reject those with large TV distances. The acceptance probability of a candidate prefix reads: $P_{i \rightarrow j} = 1 - \text{TV}(p_i, p_j, \{\omega_k\}_k)$. In practice, this random selection does not introduce noticeable fireflies, but rather reduces the variance since paths are rejected when distributions mismatch. To avoid unnecessary computations, this test can be performed before the null interactions shift (detailed in section 3.1). To finalize the suffix, a direction is sampled through the uniform mixture of scattering distributions associated to the accepted prefixes, and the remaining part of the suffix is built by regular path tracing as it is view-independent.

5. Applications and results

We implemented our method in a custom renderer that includes a state of the art volumetric integrator using the spectral heteroge-
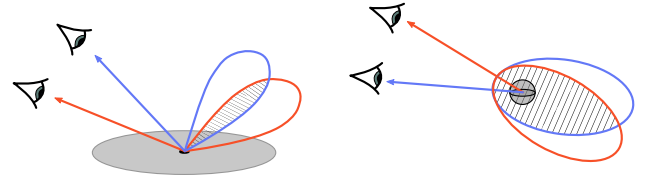
neous volumetric path sampling approach of Miller et al. [MGJ19], with MIS between direct illumination techniques (i.e. light sampling+ratio tracking and bsdf sampling+delta tracking). Our multi-view integrator decouples the first bounce, to compute the prefixes and MWIS weights, from the rest of the volumetric path tracing loop. We store a single image buffer per camera to reduce memory use. For each sample we write its respective contributions in each framebuffer that it contributes to. We provide our C++ implementation in supplementary materials. We compare several multi-view rendering variants: (mvpt) our multi-view rendering with MWIS and prefix selection, (mvpt w/o selection) our multi-view rendering with MWIS and full re-use, (amvpt) our multi-view rendering with MWIS, prefix selection and adaptive sampling, (mismvpt) multi-view rendering with MIS and full re-use as used in [Sch19], (vpt) frame by frame rendering as the baseline, and finally (Fraboni et al.) multi-view rendering adapted from [FINB19].

We illustrate our technique on several applications, including the rendering of lenticular and lightfield images, holographic stereograms, virtual walkthrough in static scenes, and videos of dynamic scenes with motion blur. We designed several test scenes with various materials and camera setups to evaluate our algorithm under different conditions. Our results can be better appreciated in our supplementary video, better showing animated and parallax effects that are not reproduced in print. In the following examples except the Disney cloud, all volumes are procedural and made of ridged multifractal or simpler functions [MPPW94].

Lenticular images. We print lenticular images that consist of 10 views with small baseline and horizontal parallax (Fig. 7, top). A lenticular sheet of 60 lenticles per inch was used, images were printed at 600 dpi, leaving 10 views per lenticle.

Lightfield images. Our lightfield display – a 8K Looking Glass display – renders horizontal parallax light fields consisting of 45 views (Fig. 7, bottom).

Holographic stereograms. Holographic stereograms are holograms printed using many ordinary renderings of the same scene (as opposed to rendering an interference pattern). Cameras are typically positioned extremely close to the object, and baseline is very small. Each rendered image is displayed on a spatial light modulator traversed by a laser light beam.

Figure 7: *Top row. We print lenticular images using 10 views. Bottom row. Our lightfield display renders images consisting of 45 views. We demonstrate our method on 3 lightfield images.*

A system of lenses prints its optical Fourier transform on a tiny part of a sensitive film, called hogel. A monochrome hologram can be seen in the right inset, generated using a prototype printer. Here, the grid of hogels is composed of 114×171 views. Rendering $114 \times 171 \approx 20k$ views at once would be intractable due to memory limits, so we rendered tiles of 7×7 views.
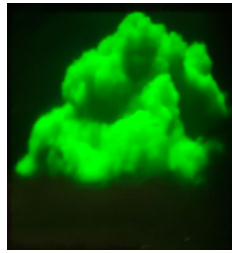
Virtual walkthroughs. Camera paths in static scenes can also benefit from our method. Our supplementary video showcases a walkthrough over the dragon scene with 100 frames, each considered as one view.

Video motion blur. We demonstrate the benefit of our method on a sequence of an animated cloud. In that case, the temporal filtering (i.e., motion blur) reuses paths from views that are temporally adjacent. We use a filter size of 5 frames.

Results. Fig. 9 illustrates how rendering errors decrease with the number of views. As expected, as the number of views increases, the number of reused samples increases and the variance is effectively reduced in the images. However, the overhead of the shifts and the fact that reused paths are short in these setups, break the theoretical bound of $\sqrt{N}$ times less error for N views.

Full render statistics and equal-time comparisons for all our results are provided in Table 1, while rendered images with comparison to single-view path-tracing (vpt) can be seen in Fig. 8. Our method significantly reduces rendering errors at equal time.

We illustrate in Fig. 4 the variance reduction due to the MWIS (mvpt) estimator in comparison to MIS (mismvpt). MIS suffers from the transformed densities that are not ideal on surfaces and in volumes due to the Jacobians of shifts. In contrast, MWIS use the knowledge of the base pdf to improve the results significantly. We further show in Fig. 5 that re-using all paths (mismvpt and mvpt w/o selection) may increase the variance with view-dependent scattering functions (BSDF or phase functions). On the contrary our selective re-use (mvpt), that is made possible through our self-normalized MWIS estimator, does not degrades the result in such cases.

Comparisons of our (mvpt) variants to the previous method of Fraboni et al. [FINB19] (updated for volumes) and the equivalent of [Sch19] (updated for volumes – mismvpt) are shown in Fig. 10. The scene is composed of cloud cubes of varying anisotropy (decreasing from back to front, left to right), our methods (mvpt and mvpt w/o selection) produce the best variance reduction.

6. Discussion and limitations

Our approach favorably reduces variance in comparison to classical path tracing at equal time. However, while our method works well in a number of useful situations, it also has limitations.

As we can see in Fig. 4, multi-view approaches may lead to uneven sample repartition in the images (bottom crop – dragon silhouette), hence leading to uneven noise levels near volume silhouettes. Our method is also useful only on non truly specular surfaces, since no reuse at the first real interaction can occur in such situations which also results in undersampled regions. A simple multi-pass adaptive sampling of the image plane, is sufficient to mitigate this issue as illustrated in Fig. 11.

Solutions have been explored to transform constrained specular paths, such as specular manifold exploration [JM12, ZGJ20] or techniques for direct lighting through specular chains [HDF15, WHD17] and specular shifts (i.e. half vector copy) for gradient domain rendering [HGP*19]. Our shift mappings for null scattering chains could be combined with such techniques, to handle specular manifold exploration, independently from our contributions. Such combination could further reduce the sample inhomogeneity in multi-view images. Similarly, image-based denoising techniques have been proposed to further reduce noise in the context of spatio-temporal sequences [VRM*18, ZRJ*15] for surfaces and volumes [IGMM21]. While they often require exporting numerous additional feature layers per image, our technique could feed these denoisers with cleaner inputs for improved results.

Finally, our research prototype is designed as proof of concept and is not representative of the true performance that an optimized version could reach. Due to the high number of image buffers, using adequate structures to maintain coherency while writing contributions could further improve the results of our method in equal time comparisons.

7. Conclusion

We introduced new shift mappings in heterogeneous participating media for path reusing purposes. We showed that our method allows to render complex scattering in participating media in an efficient way by sharing the construction of difficult paths across views, in several multi-view configurations. We improve upon classical frame by frame volumetric path tracing when the scene benefits from reuse by an order of magnitude. Our shifts volumetric mappings could be used in other path reusing applications, such as gradient domain rendering.

We further introduced a new MWIS estimator that is convenient when classical MIS is not and that effectively reduces the variance of a set of arbitrary techniques (mixed poor and good). Even though
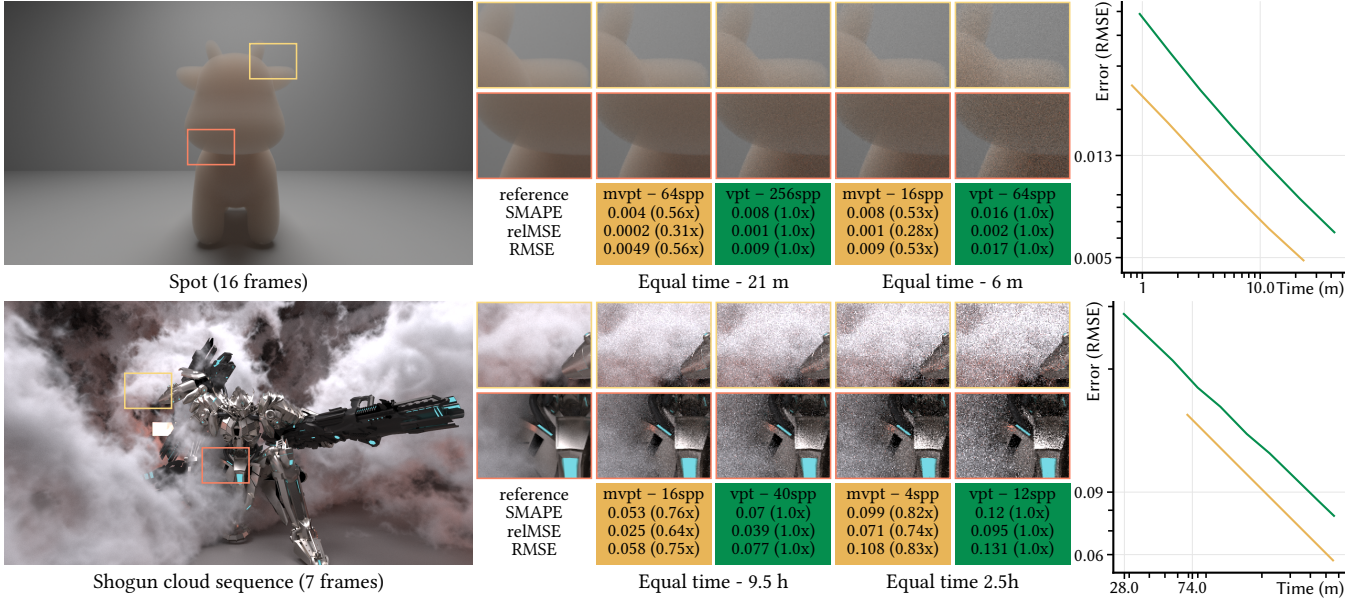
Figure 8: *Side by side comparison of our technique (mvpt) against independent volumetric path tracing (vpt) on different sequences.*
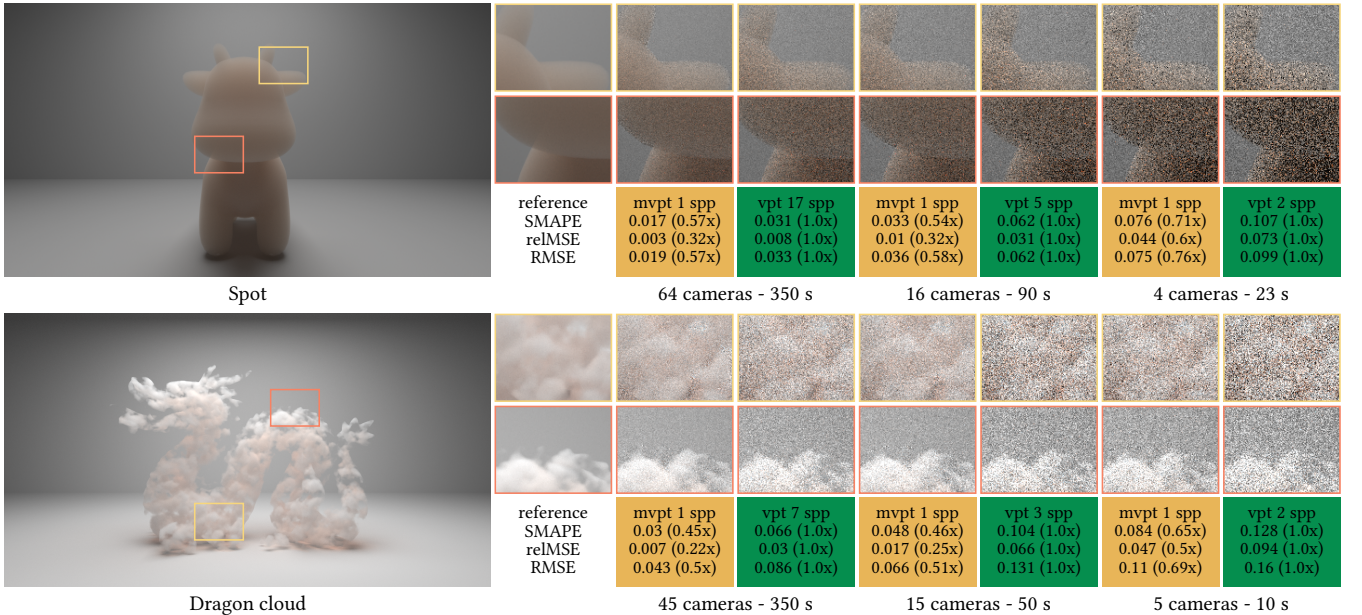

Figure 9: *Comparisons with variable number of cameras. Increasing the number of views results in lower error at equal rendering time.*

our estimator is biased but consistent, we believe that the combination of MIS and WIS is a powerful tool that may improve estimates in other applications. A formal study and an unbiased version of our estimator, that does not require stochastic resampling nor additional random walk, are also interesting research directions.

8. Acknowledgements

References

[AH93] ADELSON S. J., HODGES L. F.: Stereoscopic ray-tracing. *The Visual Computer 10*, 3 (1993), 127–144. 1, 2, 3

[AH95] ADELSON S. J., HODGES L. F.: Generating exact ray-traced animation frames by reprojection. *IEEE Computer Graphics and Applications 15* (1995), 43–52. 1, 2, 3

Scene				Multi-View (mvpt)				Single-View (vpt)		
Title	Figures	Resolution	Views	relMSE	Time	Native SPP	Mean SPP	relMSE	Time	SPP
Shogun † (7 frames)	8	1280x720	7	**0.025**	9.5h	16	94	0.039	9.5h	40
Shogun † (lenticular)	7	1280x720	10	–	15.5h	16	122	–	–	–
Shogun † (video)	–	1280x720	50	–	55h	8	–	–	–	–
Spot ‡ (4 cameras)	9	1280x720	4	**0.044**	23s	1	3	0.073	23s	2
Spot ‡ (16 cameras)	8 and 9	1280x720	16	**0.01**	90s	1	15	0.031	90s	5
Spot ‡ (light field)	7 and 9	1280x720	64	**0.003**	350s	1	52	0.008	350s	17
Spot ‡ (lenticular)	7	1920x1080	10	**0.001**	7m	16	–	0.002	7m	64
Dragon ‡ (5 cameras)	9	1280x720	5	**0.047**	10s	1	5	0.094	10s	2
Dragon ‡ (15 cameras)	9	1280x720	15	**0.017**	50s	1	13	0.066	50s	3
Dragon † (light field)	1, 4, 7 and 9	1280x720	45	**0.00047**	42m	14	–	0.0039	40m	52
Anisotropic dragon ‡	5	1280x720	7	**0.040**	40m	82	–	0.0479	40m	90
3 Dragons glossy ground ‡	5 and 11	1280x720	7	**0.0045**	40m	206	–	0.0049	40m	266
Cloud ‡ (lenticular)	4 and 7	1920x1080	10	**0.015**	40m	137	–	0.025	40m	217
Cloud ‡ (light field)	7	1280x720	45	**0.023**	350m	16	290	0.029	350m	242
Cloud § (hologram)	section 5	1920x1080	19494	**0.009**	201d	4	80	0.011	201d	64
Motion blur cloud † (video)	–	1280x720	20	–	21h	16	48	–	21h	20
Dragon flyover ‡ (video)	–	1024x640	100	**0.0011**	96m	4	–	0.0014	96m	62
Cloubes ‡	10	1280x720	7	**0.0018**	40m	262	1153	0.0045	40m	351
Mixed cubes ‡	11	1280x720	7	**0.034**	20m	364	1552	0.0305	20m	645

Table 1: *Equal time comparison for various scene setups. Compute setups:* † *Desktop computer 1 (AMD Ryzen 7 2700x 8 cores – 16Go RAM),* ‡ *desktop computer 2 (AMD Ryzen 9 3900x 12 cores – 32Go RAM),* § *compute cluster IN2P3 with 424 jobs (8 cores – 32Go RAM each) of 7x7 camera tiles.*

[BFK19] BINDER N., FRICKE S., KELLER A.: Massively parallel path space filtering. *ArXiv* (2019). 2

[BPE17] BAUSZAT P., PETITJEAN V., EISEMANN E.: Gradient-domain path reusing. *ACM Trans. on Graphics (SIGGRAPH Asia) 36*, 6 (2017). 2

[BSH02] BEKAERT P., SBERT M., HALTON J.: Accelerating path tracing by re-using paths. In *Workshop on Rendering* (2002), Eurographics. 2

[BSW00] BEKAERT P., SBERT M., WILLEMS Y. D.: Weighted importance sampling techniques for Monte Carlo radiosity. In *Workshop on Rendering Techniques* (2000), Springer-Verlag, pp. 35–46. 2, 3, 6

[BWP*20] BITTERLI B., WYMAN C., PHARR M., SHIRLEY P., LEFOHN A., JAROSZ W.: Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting. *ACM Trans. on Graphics (SIGGRAPH) 39*, 4 (2020). 2, 3, 8

[CJMn21] CRESPO M., JARABO A., MUÑOZ A.: Primary-space adaptive control variates using piecewise-polynomial approximations. *ACM Trans. on Graphics (TOG) 40*, 3 (2021). 3

[CTE05] CLINE D., TALBOT J., EGBERT P.: Energy redistribution path tracing. *ACM Trans. on Graphics (TOG) 24*, 3 (2005). 4

[FINB19] FRABONI B., IEHL J.-C., NIVOLIERS V., BOUCHARD G.: Adaptive Multi-view Path Tracing. In *Symposium on Rendering - DL-only and Industry Track* (2019), Eurographics. 1, 2, 3, 4, 8, 9, 10, 13

[GHV*18] GRUSON A., HUA B.-S., VIBERT N., NOWROUZEZAHRAI D., HACHISUKA T.: Gradient-domain volumetric photon density estimation. *ACM Trans. on Graphics (TOG) 37* (2018). 2

[HBGM11] HENRICH N., BAERZ J., GROSCH T., MÜLLER S.: Accelerating path tracing by eye-path reprojection. In *International Congress on Graphics and Virtual Reality (GRVR)* (2011). 1, 2, 3

[HDF15] HANIKA J., DROSKE M., FASCIONE L.: Manifold Next Event Estimation. *Computer Graphics Forum* (2015). 10

[HDMS03] HAVRAN V., DAMEZ C., MYSZKOWSKI K., SEIDEL H.-P.: An efficient spatio-temporal architecture for animation rendering. In *SIGGRAPH Sketches & Applications* (2003), ACM. 1, 2, 3, 5

[HGP*19] HUA B.-S., GRUSON A., PETITJEAN V., ZWICKER M., NOWROUZEZAHRAI D., EISEMANN E., HACHISUKA T.: A survey on gradient-domain rendering. In *Computer Graphics Forum* (2019), vol. 38, Eurographics, pp. 455–472. 2, 4, 10

[HHM18] HEITZ E., HILL S., MCGUIRE M.: Combining analytic direct illumination and stochastic shadows. In *SIGGRAPH Symposium on Interactive 3D Graphics and Games* (2018), ACM, p. 2. 3

[IGMM21] IGLESIAS-GUITIAN J. A., MANE P., MOON B.: Real-time denoising of volumetric path tracing for direct volume rendering. *arXiv* (2021). 10

[JM12] JAKOB W., MARSCHNER S.: Manifold exploration: a Markov Chain Monte Carlo technique for rendering scenes with difficult specular transport. *ACM Trans. on Graphics (SIGGRAPH) 31*, 4 (2012). 2, 10

[Kaj86] KAJIYA J. T.: The rendering equation. In *SIGGRAPH* (1986), ACM, pp. 143–150. 2

[KDB14] KELLER A., DAHM K., BINDER N.: Path space filtering. In *SIGGRAPH Talks* (2014), ACM. 2

[Kel96] KELLER A.: Quasi-Monte Carlo radiosity. In *Workshop on Rendering Techniques* (1996), Eurographics. 3

[KMA*15] KETTUNEN M., MANZI M., AITTALA M., LEHTINEN J., DURAND F., ZWICKER M.: Gradient-domain path tracing. *ACM Trans. on Graphics (SIGGRAPH) 34*, 4 (2015), 123:1–123:13. 2

[KSKAC02] KELEMEN C., SZIRMAY-KALOS L., ANTAL G., CSONKA F.: A simple and robust mutation strategy for the Metropolis light transport algorithm. *Computer Graphics Forum 21*, 3 (2002), 531–540. 4

[KVG*19] KONDAPANENI I., VÉVODA P., GRITTMANN P., SKŘIVAN T., SLUSALLEK P., KŘIVÁNEK J.: Optimal multiple importance sampling. *ACM Trans. on Graphics (SIGGRAPH) 38*, 4 (2019), 1–14. 3

[LKL*13] LEHTINEN J., KARRAS T., LAINE S., AITTALA M., DURAND F., AILA T.: Gradient-domain Metropolis light transport. *ACM Trans. on Graphics (SIGGRAPH) 32*, 4 (2013), 95:1–95:12. 2, 4

[LSR17] LEIMKÜHLER T., SEIDEL H.-P., RITSCHEL T.: Minimal warping: Planning incremental novel-view synthesis. *Computer Graphics Forum 36*, 4 (2017), 1–14. 3

[LWY21] LIN D., WYMAN C., YUKSEL C.: Fast volume rendering with spatiotemporal reservoir resampling. *ACM Trans. on Graphics (TOG) 40*, 6 (2021). 3
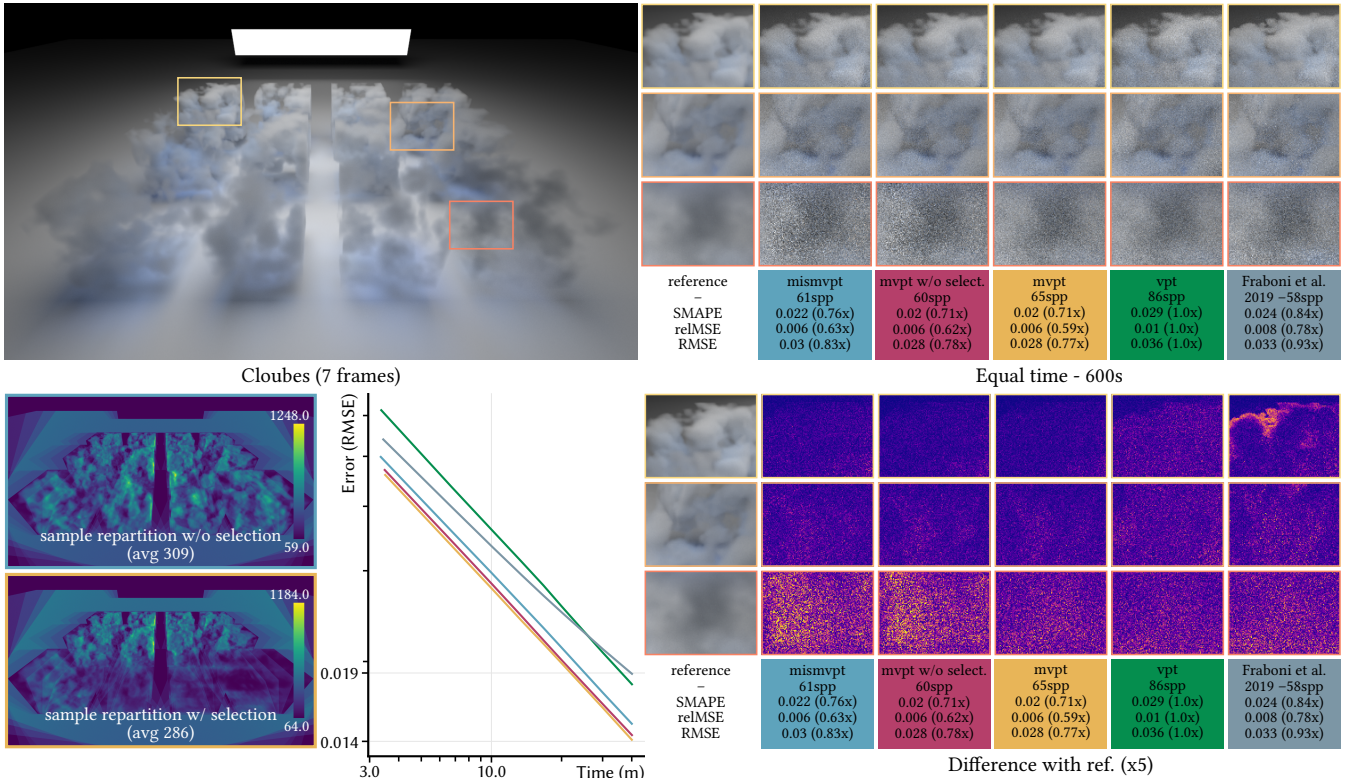
Figure 10: *We compare our methods (mvpt) that use our MWIS estimator with and without prefix selection, to (mismvpt) that uses MIS and no selection, and state of the art multiple view rendering estimators [FINB19] that uses a biased MIS variant and a different prefix selection taking into account both Jacobian variations and scattering functions (updated with our shifts and scattering function similarity). Reusing all prefixes disregarding the scattering components (mismvpt and mvpt w/o selection) results in noisier estimates than the baseline (vpt) in anisotropic volumes (bottom crops). Although the selection proposed by Fraboni et al. (gray – blue curve and insets) limits the degradations due to large variations of Jacobians, it still exhibits fireflies and a visible bias that is not consistent as the number of samples increases. Our multi-view (mvpt) solutions exhibit lower levels of noise and no visible bias (nor in the curve slopes) nor artifacts even at low sample count, and performs better than independent path tracing (vpt) and (mismvpt).*

[MFSSK06] MÉNDEZ FELIU À., SBERT M., SZIRMAY-KALOS L.: Reusing frames in camera animation. *WSCG 14*, 1-3 (2006). 1, 2, 3, 5

[MGJ19] MILLER B., GEORGIEV I., JAROSZ W.: A null-scattering path integral formulation of light transport. *ACM Trans. on Graphics (SIGGRAPH) 38*, 4 (2019). 2, 3, 4, 5, 9

[MKD*16] MANZI M., KETTUNEN M., DURAND F., ZWICKER M., LEHTINEN J.: Temporal gradient-domain path tracing. *ACM Trans. on Graphics (SIGGRAPH Asia) 35*, 6 (2016), 246:1–246:9. 2, 3

[MPPW94] MUSGRAVE F. K., PEACHEY D., PERLIN K., WORLEY S.: *Texturing and modeling: a procedural approach*. Academic Press Professional, 1994. 9

[NGHJ18] NOVÁK J., GEORGIEV I., HANIKA J., JAROSZ W.: Monte Carlo methods for volumetric light transport simulation. *Computer Graphics Forum (Eurographics State of the Art Reports) 37*, 2 (2018). 2, 3, 4, 14

[Owe13] OWEN A. B.: *Monte Carlo theory, methods and examples*. Standford, 2013. URL: https://statweb.stanford.edu/~owen/mc/. 6

[PKK00] PAULY M., KOLLIG T., KELLER A.: Metropolis light transport for participating media. In *Workshop on Rendering Techniques* (2000), Springer-Verlag, pp. 11–22. 2

[PS66] POWELL M. J., SWANN J.: Weighted uniform sampling: a Monte Carlo technique for reducing variance. *IMA Journal of Applied Mathematics 2*, 3 (1966), 228–236. 6, 8

[Sch19] SCHWARZHAUPT J.: Path tracing in production: Part 1: Modern path tracing chapter 4: Finding good paths. In *SIGGRAPH Courses* (2019), ACM. 1, 2, 3, 4, 5, 9, 10

[Spa79] SPANIER J.: A new family of estimators for random walk problems. *IMA Journal of Applied Mathematics 23*, 1 (1979), 1–31. 6

[TCE05] TALBOT J., CLINE D., EGBERT P.: Importance Resampling for Global Illumination. In *Symp. on Rendering* (2005), Eurographics. 3, 8

[VG95] VEACH E., GUIBAS L. J.: Optimally combining sampling techniques for Monte Carlo rendering. In *SIGGRAPH* (1995), ACM, pp. 419–428. 3, 8

[VG97] VEACH E., GUIBAS L. J.: Metropolis light transport. In *SIGGRAPH* (1997), ACM, pp. 65–76. 2

[VRM*18] VOGELS T., ROUSSELLE F., MCWILLIAMS B., RÖTHLIN G., HARVILL A., ADLER D., MEYER M., NOVÁK J.: Denoising with kernel prediction and asymmetric loss functions. *ACM Trans. on Graphics (TOG) 37*, 4 (2018). 10

[WGGH20] WEST R., GEORGIEV I., GRUSON A., HACHISUKA T.: Continuous multiple importance sampling. *ACM Trans. on Graphics (TOG) 39*, 4 (2020). 2, 3

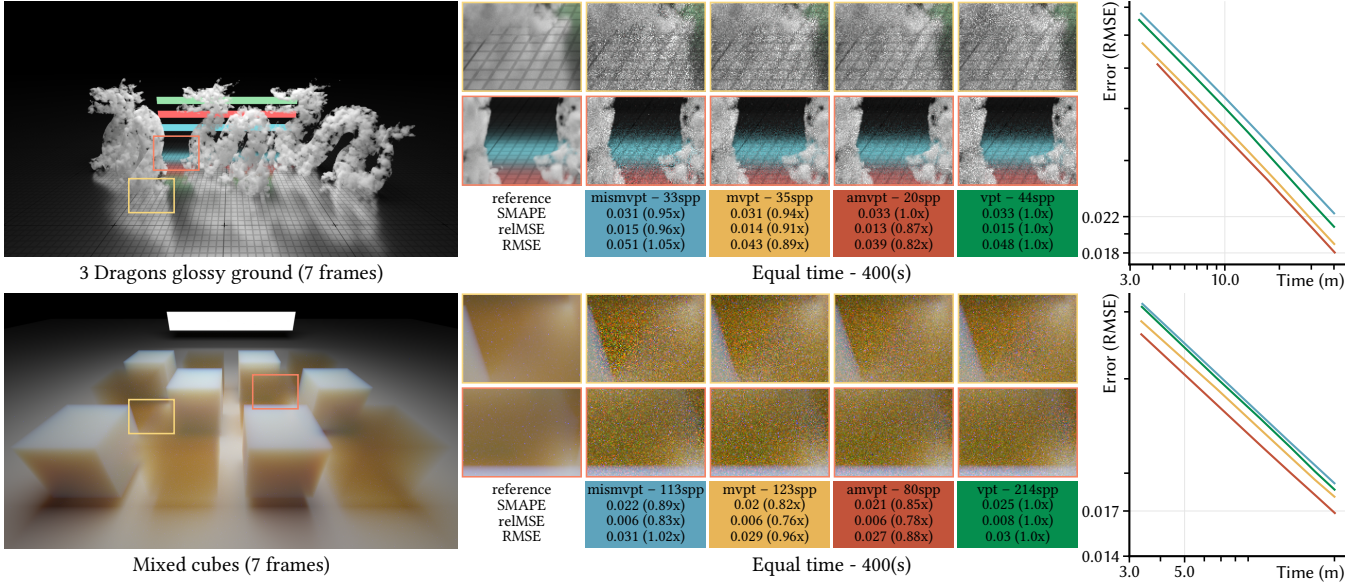[WHD17] WEBER P., HANIKA J., DACHSBACHER C.: Multiple ver-

Figure 11: *Adaptive refinement. The top row shows a scene with isotropic phase function and a glossy ground. The bottom row shows mixed isotropic and strongly forward anisotropic volumes ($g = 0.98$). As we expect less re-use with strongly view-dependent scattering functions (mvpt), the adaptive variant of multi view rendering (amvpt) helps distribute samples in the undersampled areas. The adaptive variant has a small overhead due to error and sample maps evaluation, and distributes less samples in regions where the multi-view already performs well.*

tex next event estimation for lighting in dense, forward-scattering media. *Computer Graphics Forum* (2017). 10

[XS07] XU Q., SBERT M.: A new way to re-using paths. In *Computational Science and Its Applications* (2007), pp. 741–750. 2

[ZGJ20] ZELTNER T., GEORGIEV I., JAKOB W.: Specular manifold sampling for rendering high-frequency caustics and glints. *ACM Trans. on Graphics (SIGGRAPH) 39*, 4 (2020). 10

[ZRJ*15] ZIMMER H., ROUSSELLE F., JAKOB W., WANG O., ADLER D., JAROSZ W., SORKINE-HORNUNG O., SORKINE-HORNUNG A.: Path-space motion estimation and decomposition for robust animation filtering. *Computer Graphics Forum (EGSR) 34*, 4 (2015), 131–142. 10

Appendix A: Definitions

We explicit notations and definitions for the measurement contribution function in volumetric light transport (Eq. 2) as defined by Novak et al. [NGHJ18] for the manuscript to remain self-contained.

$$G(x,y) = \frac{D(x,y)D(y,x)}{\|x-y\|^2} \quad \text{where}$$

$$D(x,y) = \begin{cases} |n(x) \cdot \omega_{x \to y}| & \text{if } x \text{ is on a surface,} \\ 1 & \text{if } x \text{ is in a medium.} \end{cases}$$

$$T(x,y) = e^{-\int_{t_x}^{t_y} \mu_t(x_t) dt} \quad \text{where} \quad x_t = x + t \cdot \omega_{x \to y}$$

$$L_e(x,y) = \begin{cases} L_e(x, \omega_{x \to y}) & \text{if } x \text{ is on a surface,} \\ \mu_a(x) L_e(x, \omega_{x \to y}) & \text{if } x \text{ is in a medium.} \end{cases}$$

$$f_s(x,y,z) = \begin{cases} f_r(\omega_{y \to x}, \omega_{y \to z}) & \text{if } y \text{ is on a surface,} \\ \mu_s(y) f_p(\omega_{y \to x}, \omega_{y \to z}) & \text{if } y \text{ is in a medium.} \end{cases}$$

Appendix B: Majorant optical depth shift derivation

We complete the derivation of our shift mapping of medium interaction depths using the majorant optical depth. Starting from Eq. 7 and noting $C = \frac{\bar{\tau}'_{\max}}{\bar{\tau}_{\max}}$, we have:

$$\bar{\tau}'(t') = C \cdot \bar{\tau}(t) \tag{25}$$

Plugging the piecewise constant definition of the majorant optical depth (Eq. 6) into the latter equation leads to:

$$\sum_{i=1}^{n_{t'}} \bar{\mu}'_i \left(\min\left(t', t'_{\max,i}\right) - t'_{\min,i} \right) = C \cdot \sum_{j=1}^{n_t} \bar{\mu}_j \left(\min\left(t, t_{\max,j}\right) - t_{\min,j} \right) \tag{26}$$

where n_t is the volume section index containing depth t on the base segment (and resp. $n_{t'}$ on the target segment for depth t'), $t_{\min,j}$, $t_{\max,j}$ and $\bar{\mu}_j$ are the entry, the exit and the constant majorant extinction coefficient of the j-eth volume section of the base segment (and resp. $t'_{\min,i}$, $t'_{\max,i}$ and $\bar{\mu}'_i$ on the target segment).

We can rearrange the above equation as follows:

$$(26) \Leftrightarrow \quad \bar{\mu}'_{n_{t'}} \cdot t' + a_1 = C \cdot [\bar{\mu}_{n_t} \cdot t + a_2]$$
$$\Leftrightarrow \quad t' = C \cdot \frac{\bar{\mu}_{n_t}}{\bar{\mu}'_{n_{t'}}} \cdot t + R \tag{27}$$

where $a_1 = \sum_{i=1}^{n_{t'}-1} \bar{\mu}'_i \left(t'_{\max,i} - t'_{\min,i} \right) - \bar{\mu}'_{n_{t'}} \cdot t'_{\min,n_{t'}}$, $a_2 = \sum_{j=1}^{n_t-1} \bar{\mu}_j \left(t_{\max,j} - t_{\min,j} \right) - \bar{\mu}_{n_t} \cdot t_{\min,n_t}$ and $R = \frac{C \cdot a_2 - a_1}{\bar{\mu}'_{n_{t'}}}$. Replacing the constant C and rewriting the majorant extinction coefficients as function of depth, as $\bar{\mu}_{n_t} = \bar{\mu}(t)$, leads to the final form of Eq. 8.