Variance and Convergence Analysis of Monte Carlo Line and Segment Sampling

Gurprit Singh   Bailey Miller   Wojciech Jarosz

Dartmouth College

Abstract

Recently researchers have started employing Monte Carlo-like line sample estimators in rendering, demonstrating dramatic reductions in variance (visible noise) for effects such as soft shadows, defocus blur, and participating media. Unfortunately, there is currently no formal theoretical framework to predict and analyze Monte Carlo variance using line and segment samples which have inherently anisotropic Fourier power spectra. In this work, we propose a theoretical formulation for lines and finite-length segment samples in the frequency domain that allows analyzing their anisotropic power spectra using previous isotropic *variance and convergence tools. Our analysis shows that judiciously oriented line samples not only reduce the dimensionality but also pre-filter C^0 discontinuities, resulting in further improvement in variance and convergence rates. Our theoretical insights also explain how finite-length segment samples impact variance and convergence rates only by pre-filtering discontinuities. We further extend our analysis to consider (uncorrelated) multi-directional line (segment) sampling, showing that such schemes can increase variance compared to unidirectional sampling. We validate our theoretical results with a set of experiments including direct lighting, ambient occlusion, and volumetric caustics using points, lines, and segment samples.*

CCS Concepts

•*Computing methodologies* → *Ray tracing;* •*Mathematics of computing* → *Stochastic processes; Computation of transforms;*

1. Introduction

While point sampling has dominated the view of Monte Carlo (MC) in graphics for years, MC integration need not always be a point sampling process. In fact, researchers have recently started employing MC-like estimators using *line samples* for rendering problems as diverse as anti-aliasing [JP00], motion blur [GDAM10], depth of field [TPD*12], hair rendering [BGAM12], as well as density estimation [JNSJ11] or path sampling [NNDJ12, GKH*13, KGH*14] in volumes. While these recent methods have shown considerable promise, little is currently known about the theoretical properties of such line-sampling estimators, and their extreme anisotropic nature makes them seemingly incompatible with the wealth of prior isotropic, point-based analyses. Fig. 1(a), for instance, shows the expected power spectrum of 3D line samples whose (offset) locations have a Poisson disk distribution. While this leads to an isotropic sampling power spectrum within the 2D subspace of offsets, the *radially averaged* power spectrum (Fig. 1(b), top) obscures this fact. More generally, for *d*-dimensional unidirectional line samples, the power spectrum resides within a (*d*-1)-dimensional subspace.

Lines and segment samples have already been studied in the frequency domain [SZG*13], however, this previous analysis did not provide any tools to analyze the variance and convergence rates

of MC integration using such samples. In this work, we perform a theoretical analysis of lines and segment (finite length) samples in the frequency domain which reveals that the expected power spectrum of lines and segment samples have an *isotropic* counterpart that allows us to analyze their inherently anisotropic power spectra using previous *isotropic* [PSC*15] variance and convergence tools. Our analysis further shows that judiciously oriented line samples can pre-filter C^0 discontinuities, thereby improving the variance and convergence rates beyond dimensionality reduction. However, with finite length segment samples we can only expect improvements due to pre-filtering of C^0 discontinuities, without any dimensionality reduction. We further extend our analysis to (uncorrelated) multi-directional line and segment samples. We validate our theoretical results with a set of experiments including direct lighting, ambient occlusion, and volumetric caustics using points, lines and segment samples.

2. Related work

Point sampling, variance, & convergence. Ever since Cook et al. [CPC84] introduced MC integration to graphics, researchers have noted that a careful arrangement of samples can impact the spectral distribution of error and dramatically reduce its overall
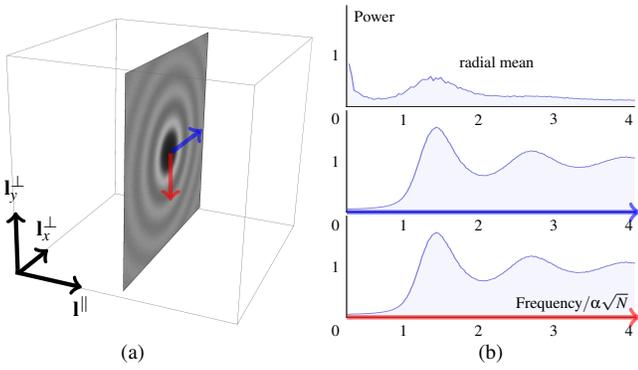
Figure 1: *(a) The expected power spectrum (DC at the center) of unidirectional parallel line samples (along $\mathbf{l}^{\parallel}$)—oriented orthogonal to the plane $(\mathbf{l}_x^{\perp}, \mathbf{l}_y^{\perp})$—is shown with Poisson disk line offset distributions. (b) The radial behavior of (a) is isotropic within the plane (shown for two directions with blue arrows in (a)) where the power spectrum is confined, however, the 3D radially averaged (radial mean) spectrum shown at the top in (b) does not reveal any Poisson disk characteristics. In (b), $\alpha > 0$ (more details in (21)).*

magnitude in numerical integration [DW85, Coo86, Mit91]. This has lead to extensive work on generating sample patterns which are stochastic, yet still maintain a low discrepancy [Shi91] or which exhibit so-called blue noise frequency spectra [Coo86, LD08]. Recent work [Dur11, SK13, PSC*15] has established a firm mathematical connection between the spectral properties of the sampling pattern and the *magnitude* of MC integration error. Moreover, careful sample placement—such as jittered [Coo86] and certain flavors of blue-noise sampling [BSD09, HSD13]—have now been shown to actually lead to asymptotically faster *convergence rates* [Mit96, RAMN12, SK13, SNJ*14, PSC*15]. We derive similar mathematical expressions governing variance and convergence rate, but for the case of stochastic sampling using line and segment samples which have inherently anisotropic expected power spectra.

Line sampling in rendering & related fields. While line sampling is relatively new in graphics (the idea being first applied to anti-aliasing by Jones and Perry [JP00]), it has been used for some time in related fields. A class of methods from neutron transport simulation known as "expected value estimators" and "track length estimators" [Spa66] essentially perform MC integration using line samples. These were independently developed and generalized in the graphics community in the form of "long beam" and "short beam" estimators, first for camera rays [JZJ08] and then for light rays [JNSJ11, SZLG10] in volumetric photon mapping, and later adapted to many-light methods [NNDJ12], path tracing approaches [GKH*13, KGH*14], and subsurface scattering [HCJ13]. Line samples have also cropped up for computing hemispherical visibility and motion blur [GDAM10, GBAM11], depth of field [TPD*12], visibility in hair [BGAM12], and masked environment lighting [NBMJ14]. Recently, Billen and Dutré [BD16] demonstrated improvements with line samples due to dimensionality reduction for direct illumination. While all of these approaches have demonstrated practical improvements for rendering, there is

currently little theoretical understanding of how such anisotropic sample patterns impact variance and convergence rate in the context of MC integration. We perform such an analysis, theoretically explaining and empirically validating the previously observed variance reduction properties of such line samples and segment samples.

Frequency analysis. Sun and colleagues [SZG*13] performed a frequency analysis of line and segment samples and showed that lines and finite length segments have inherently anisotropic power spectra. Their analysis was mainly focused on preserving the blue-noise properties of (uni- and multi-directional) samples to reduce noise and aliasing artifacts during *reconstruction*. We instead investigate the orthogonal problem of *integration*, presenting a frequency domain formulation of MC integration using line and segment samples. Our analysis reveals that these anisotropic sampling power spectra have isotropic counterparts, allowing us to leverage recently developed *isotropic* variance and convergence tools [PSC*15] to derive the variance convergence rates for lines and segment samples.

3. Preliminaries

We are interested in computing an integral of the form:

$$I = \int_{\mathbb{D}} f(\mathbf{x}) \, d\mathbf{x}, \tag{1}$$

where $\mathbb{D}$ is the unit d-dimensional Euclidean space.

Traditionally, Monte Carlo integration forms an approximation, $I_N \approx I$, by averaging evaluations of the integrand f at N *point* sample locations $\mathbf{p}_j$. For uniformly distributed samples, we can write:

$$I_N = \int_{\mathbb{D}} P_N(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}, \text{ with } P_N(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^{N} \delta(\|\mathbf{x} - \mathbf{p}_j\|), \tag{2}$$

where P_N is a normalized sampling function using N *points*.

In the frequency domain Φ, this integral takes the form:

$$I_N = \int_{\Phi} \mathcal{F}_{P_N}(\nu) \overline{\mathcal{F}_f(\nu)} \, d\nu, \text{ with } \mathcal{F}_{P_N}(\nu) = \frac{1}{N} \sum_{j=1}^{N} \mathcal{F}(\nu_j), \tag{3}$$

where $\overline{\mathcal{F}_f}$ is the complex conjugate of the integrand's Fourier spectrum, $\mathcal{F}_{P_N}$ is the spectrum of the normalized point sampling function, and $\mathcal{F}(\nu_j) = e^{-2\pi i \nu_j}$ with $\nu_j = \nu \cdot \mathbf{p}_j$ is the Fourier transform of a single point sample $\mathbf{p}_j$ at frequency ν. Prior work [Dur11, PSC*15] has shown that the variance of I_N can be expressed as:

$$Var(I_N) = \int_{\Theta} \langle \mathcal{P}_{P_N}(\nu) \rangle \mathcal{P}_f(\nu) \, d\nu, \tag{4}$$

where Θ includes all frequencies except DC, $\mathcal{P}_f(\nu) = \|\mathcal{F}_f(\nu)\|^2$ is the power spectrum of the integrand, and $\langle \mathcal{P}_{P_N}(\nu) \rangle$ is the expected power spectrum of the *homogenized*[†] normalized *point* sampling function. We build upon this knowledge to express Monte Carlo estimators using line and line segment samples, as well as their corresponding variances, in the Fourier domain.

[†] In the point processes literature [IPSS08], homogenization refers to *stationary* point processes for which the average number of points per some unit of extent such as length, area, or volume is constant depending on the underlying mathematical space.

4. Monte Carlo estimator for lines and segment samples

Sun and colleagues [SZG*13] derived the Fourier spectrum of line and line segment samples for the purposes of blue-noise sampling and reconstruction. We build on these definitions below to mathematically express Monte Carlo integration using such samples, both in the spatial domain and the frequency domain. We restrict ourselves initially to the *unidirectional* case where all lines or segments share the same direction. Once we start analyzing variance in Section 5, we will generalize this to the uncorrelated multidirectional case.

4.1. Line samples

We denote a d-dimensional parametric line as: $\mathbf{l}(t) = \mathbf{l}^\perp + \mathbf{l}^\parallel t$, where $\mathbf{l}^\parallel$ is a unit d-dimensional vector denoting the direction of the line, and $\mathbf{l}^\perp$ is the point on the line closest to the origin. We can express Monte Carlo integration using such line samples as:

$$I_N = \int_\mathbb{D} L_N(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}, \text{ where } L_N(\mathbf{x}) = \frac{1}{N}\sum_{j=1}^N \delta(\text{dist}(\mathbf{x}, \mathbf{l}_j)). \quad (5)$$

Compared to (2), this relies on a normalized sampling function L_N consisting of N uniformly distributed *lines* where $\text{dist}(\mathbf{x}, \mathbf{l}_j) = \|(\mathbf{l}_j^\perp - \mathbf{x}) + (\mathbf{x} \cdot \mathbf{l}_j^\parallel)\mathbf{l}_j^\parallel\|$ is the Euclidean distance between $\mathbf{x}$ and the j-th line sample $\mathbf{l}_j$.

Fourier Domain. In the frequency domain Φ, this integral takes an analogous form to (3) but where the point sampling spectrum $\mathcal{F}_{P_N}$ is replaced by the line sampling spectrum $\mathcal{F}_{L_N}$:

$$I_N = \int_\Phi \mathcal{F}_{L_N}(\nu)\overline{\mathcal{F}_f(\nu)}\, d\nu, \quad \mathcal{F}_{L_N}(\nu) = \frac{1}{N}\sum_{j=1}^N \mathcal{F}(\nu_j^\perp)\mathcal{K}_L(\nu_j^\parallel), \quad (6)$$

where $\nu_j^\perp = \nu \cdot \mathbf{l}_j^\perp$ denotes the frequency component in the offset plane, $\nu_j^\parallel = \nu \cdot \mathbf{l}_j^\parallel$ denotes the frequency component along the line samples, $\mathcal{F}(\nu_j^\perp) = e^{-2\pi i \nu_j^\perp}$ is the Fourier spectrum of the offset point $\mathbf{l}_j^\perp$, and $\mathcal{K}_L(\nu^\parallel) = \delta(\nu^\parallel)$ is non-zero only for frequency vectors that are orthogonal to the line sample direction. The power spectrum is simply $\mathcal{P}L_N(\nu) = \|\mathcal{F}_{L_N}(\nu)\|^2$.

Note that each line sample's frequency spectrum (each summand above) is that of a $(d\text{-}1)$-dimensional point spectrum in the coordinates perpendicular to the line $(\nu \cdot \mathbf{l}_j^\perp)$, and a delta impulse in the remaining coordinate $(\nu \cdot \mathbf{l}_j^\parallel)$ along the line. If all the lines share the same direction $\mathbf{l}_j^\parallel = \mathbf{l}^\parallel$, then the entire spectrum of the sample set is that of N $(d\text{-}1)$-dimensional points restricted to lie in a hyper-plane perpendicular to the lines. Fig. 1(a) illustrates this for $d = 3$ where parallel line samples are generated horizontally such that the power spectrum lies in a plane orthogonal to their direction.

4.2. Segment samples

Similar to lines, we denote a d-dimensional parametric segment as: $\mathbf{s}(\mathbf{p}, t) = \mathbf{p} + \mathbf{s}^\parallel t$, where $\mathbf{p}$ is the center of the segment, $\mathbf{s}^\parallel$ is a unit d-dimensional vector denoting the direction of the segment, and $t \in [-\lambda/2, \lambda/2]$ where λ is the length of the segment. The Monte

Carlo estimator for N such segment samples can be written as:

$$I_N = \int_\mathbb{D} S_N(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}, \text{ where } S_N(\mathbf{x}) = \frac{1}{N}\sum_{j=1}^N S(\mathbf{x}, \mathbf{s}_j) \quad (7)$$

where S_N is the sampling function using N *segments* with:

$$S(\mathbf{x}, \mathbf{s}_j) = \delta\left(\text{dist}(\mathbf{x}, \mathbf{s}_j)\right) H\left(\frac{\lambda}{2} - \left|(\mathbf{x} - \mathbf{p}_j) \cdot \mathbf{s}_j^\parallel\right|\right). \quad (8)$$

Here, H is the *Heaviside* function and $\text{dist}(\mathbf{x}, \mathbf{s}_j)$ is defined analogously to $\text{dist}(\mathbf{x}, \mathbf{l}_j)$ with $\mathbf{s}_j^\perp = \mathbf{p} - \mathbf{p} \cdot \mathbf{s}^\parallel$ being the point closest to the origin on the infinite line containing the segment.

Fourier Domain. In the frequency domain Φ, the integral in (7) takes an analogous form as (3) but where the point sampling spectrum $\mathcal{F}_{P_N}$ is replaced by the segment sampling spectrum $\mathcal{F}_{S_N}$:

$$I_N = \int_\Phi \mathcal{F}_{S_N}(\nu)\overline{\mathcal{F}_f(\nu)}\, d\nu, \quad \mathcal{F}_{S_N}(\nu) = \frac{1}{N}\sum_{j=1}^N \mathcal{F}(\nu_j)\mathcal{K}_S(\lambda, \nu_j^\parallel), \quad (9)$$

where $\mathcal{F}(\nu_j) = e^{-2\pi i \nu_j}$ is the Fourier spectrum of the *point* sample $\mathbf{p}_j$ at the center of the segment. Compared to (6) whose spectral kernel $\mathcal{K}_L$ is a Dirac delta function, the spectral kernel for segment samples, $\mathcal{K}_S(\lambda, \nu_j^\parallel) = \lambda \text{sinc}(\lambda \nu_j^\parallel)$, is a *sinc* function since a segment corresponds to a finite box filter. As a result, the frequency content of segments resides in the full d-dimensions, contrary to line samples, whose spectrum resides in a $(d\text{-}1)$-dimensional subspace. The power spectrum of segment samples is simply $\mathcal{P}S_N(\nu) = \|\mathcal{F}_{S_N}(\nu)\|^2$.

5. Variance formulation

As with Pilleboue et al.'s [PSC*15] formulation for point samples, the line and segment samples need to be homogenized (which is the same as performing Cranley-Patterson rotation [CP76]) to get the corresponding variance formulations. Samplers like white noise (random) and ones derived from white noise (Poisson disk [DW85], CCVT [BSD09], BNOT [dGBOD12]) are homogeneous by construction [PSC*15]. For other sampling strategies, the samples need to be homogenized. Homogenizing segments requires homogenizing the segment centers while line samples only require homogenizing the $(d\text{-}1)$ independent components of the line sample offset $\mathbf{l}^\perp$.

Given the sampling spectra for line (6) and segment (9) sampling, we could express the variance in the Fourier domain similar to that of point sampling (4):

$$Var(I_N) = \int_\Theta \langle \mathcal{P}_{L_N}(\nu) \rangle \mathcal{P}_f(\nu)\, d\nu \qquad \text{for lines, and} \quad (10)$$

$$Var(I_N) = \int_\Theta \langle \mathcal{P}_{S_N}(\nu) \rangle \mathcal{P}_f(\nu)\, d\nu \qquad \text{for segments,} \quad (11)$$

where $\langle \mathcal{P}_{L_N}(\nu) \rangle$ and $\langle \mathcal{P}_{S_N}(\nu) \rangle$ are the expected power spectra for N *homogenized* line and segment samples respectively.

Unfortunately, these expressions are not immediately useful since both expected sampling spectra are highly anisotropic. To gain further insights about how the variance of MC line and segment sampling relates to points, and to ultimately derive convergence rates, we will instead consider an alternate interpretation of these estimators.

5.1. Alternate interpretation of line and segment sampling

By expanding the sampling functions in (6) and (9), we can rewrite the Monte Carlo estimators in the Fourier domain as

$$I_N = \int_\Phi \overbrace{\frac{1}{N}\sum_{j=1}^N \mathcal{F}(\nu_j^\perp)}^{\text{Line sampling in } d}\underbrace{\mathcal{K}_L(\nu_j^\parallel)}\overbrace{\overline{\mathcal{F}_f(\nu)}}^{\text{Original integrand}}\,d\nu, \quad \text{for lines, and} \quad (12)$$

$$\underbrace{\text{Point sampling in } d\text{-1 subspace}}\quad\underbrace{\text{Convolved integrand}}$$

$$I_N = \int_\Phi \overbrace{\frac{1}{N}\sum_{j=1}^N \mathcal{F}(\nu_j)}^{\text{Segment sampling in } d}\underbrace{\mathcal{K}_S(\lambda,\nu_j^\parallel)}\overbrace{\overline{\mathcal{F}_f(\nu)}}^{\text{Original integrand}}\,d\nu \quad \text{for segments.} \quad (13)$$

$$\underbrace{\text{Point sampling in } d}\quad\underbrace{\text{Convolved integrand}}$$

The grouping specified by the over-braces is the original interpretation of the d-dimensional integration where evaluating each line (segment) sample involves integrating the original d-dimensional integrand along the line (segment). The grouping specified by the under-braces, however, shows that by premultiplying the integrand by the line or segment kernels, we can view this simply as Monte Carlo point sampling that operates on a *pre-filtered* integrand.

These two equivalent interpretations are analogous to the equivalence described by the Fourier slice theorem. While both views are equally valid, the second interpretation provides a clearer explanation for how such samples can provide benefits in Monte Carlo integration. Firstly, for both lines and segments, the convolution can potentially increase the effective integrand's smoothness and spectral decay, resulting in improved convergence. Moreover, for line samples, the sampling process is equivalent to point sampling in one dimension lower, which can provide faster convergence due to denser stratification. Line segments, on the other hand, correspond to point sampling in d dimensions, so no such dimensionality-reduction benefit will arise.

5.2. Alternate variance formulation

We can now express variance using the alternate pre-filtering interpretation of line and segment samples.

Line sampling. The integration domain becomes the $(d$-1)-dimensional subspace and we end up with point samples in this subspace. The variance (10) can therefore be rewritten:

$$Var(I_N) = \int_{\Theta^\perp} \langle \mathcal{P}_{P_N}(\tilde{\nu})\rangle \mathcal{P}_{\tilde{f}}^L(\tilde{\nu})\,d\tilde{\nu}, \quad (14)$$

where $\Theta^\perp$ is the $(d$-1)-dimensional Fourier integration domain in the plane of line offsets, $\langle \mathcal{P}_{P_N}(\tilde{\nu})\rangle$ is the expected *point* sampling power spectrum of line offsets in the $(d$-1) subspace, and $\mathcal{P}_{\tilde{f}}^L(\tilde{\nu})$ is the power spectrum of the effective *pre-filtered* integrand evaluated at a point $\tilde{\nu}$ in the $(d$-1) subspace.

Segment sampling. If we consider the integrand to be convolved along the segment directions, followed by point sampling, the vari-

ance (11) can be rewritten as:

$$Var(I_N) = \int_\Theta \langle \mathcal{P}_{P_N}(\nu)\rangle \mathcal{P}_{\tilde{f}}^S(\nu)\,d\nu, \quad (15)$$

where $\langle \mathcal{P}_{P_N}(\nu)\rangle$ is the expected point sampling power spectrum of the N segment centers, and $\mathcal{P}_{\tilde{f}}^S(\nu)$ is the power spectrum of the effective *pre-filtered* integrand. In contrast to line samples, segments do not reduce the dimensionality, so (15) remains in the original d dimensions. Only when the segment lengths span the entire domain—producing line samples—will we get dimensionality reduction.

Isotropic offset distributions. For line samples with isotropically distributed offsets, we get an isotropic spectrum in the $(d$-1)-dimensional subspace (Fig. 1(a)). Similarly, following the pre-filtering interpretation of line segments, we obtain an isotropic expected power spectrum if the segment centers are distributed isotropically. Consequently, we can further simplify equations (14) and (15) by radially averaged power spectra (following [PSC*15]):

$$Var(I_N) = \begin{cases} \mu_L \int_0^\infty \tilde{\rho}^{d-2}\langle \mathcal{P}_{P_N}(\tilde{\rho})\rangle \mathcal{P}_{\tilde{f}}^L(\tilde{\rho})\,d\tilde{\rho} & \text{lines} \\ \mu_S \int_0^\infty \rho^{d-1}\langle \mathcal{P}_{P_N}(\rho)\rangle \mathcal{P}_{\tilde{f}}^S(\rho)\,d\rho & \text{segments,} \end{cases} \quad (16)$$

where $\mu_L = \mu(\mathcal{S}^{d-2})$ and $\mu_S = \mu(\mathcal{S}^{d-1})$ denote the Lebesgue measures in the respective domains. ρ and $\tilde{\rho}$ are the radial frequencies in the original d-dimensional space and the $(d$-1)-dimensional subspace, respectively.

5.3. Multi-directional variance formulation

For multi-directional line sampling, if the sample *offsets* across different directions are *uncorrelated* then the estimators for different directions become uncorrelated random variables with additive variance. Averaging the uncorrelated estimators from m such directions with N_k samples each results in the average of variances from the individual estimators, weighted by their squared sample counts:

$$Var(I_N) = \sum_{k=1}^m \frac{N_k^2}{N^2}\int_{\Theta_k^\perp} \langle \mathcal{P}_{P_{N_k}}(\tilde{\nu}_k)\rangle \mathcal{P}_{\tilde{f}_k}^L(\tilde{\nu}_k)\,d\tilde{\nu}_k, \quad (17)$$

where $N = \sum_{k=1}^m N_k$ is the total number of line samples, $\mathcal{P}_{\tilde{f}_k}^L(\tilde{\nu}_k)$ is the power spectrum of the effective integrand $\tilde{f}_k$ which has been *pre-filtered* along the k-th direction, and $\langle \mathcal{P}_{P_{N_k}}(\tilde{\nu}_k)\rangle$ is the expected power spectrum of sampling the N_k line offsets $\mathbf{l}^{\perp_k}$ within the $(d$-1)-dimensional subspace $\Theta_k^\perp$ orthogonal to the k-th direction (see supplemental section 1 for a derivation in Fourier domain).

Similarly, for uncorrelated multi-directional line segment samples, the variance formulation from (15) can be rewritten as:

$$Var(I_N) = \sum_{k=1}^m \frac{N_k^2}{N^2}\int_\Theta \langle \mathcal{P}_{P_{N_k}}(\nu)\rangle \mathcal{P}_{\tilde{f}_k}^S(\nu)\,d\nu. \quad (18)$$

Isotropic offset distributions. For multiple directions, we can similarly generalize (16) for line and segment samples in the radial form

as follows:

$$Var(I_N) = \begin{cases} \mu_L \sum_{k=1}^{m} \frac{N_k^2}{N^2} \int_0^\infty \tilde{\rho}_k^{d-2} \langle \mathcal{P}_{P_{N_k}}(\tilde{\rho}_k) \rangle \mathcal{P}_{\tilde{f}_k}^L(\tilde{\rho}_k) \, d\tilde{\rho}_k & \text{lines} \\ \mu_S \sum_{k=1}^{m} \frac{N_k^2}{N^2} \int_0^\infty \rho_k^{d-1} \langle \mathcal{P}_{P_{N_k}}(\rho_k) \rangle \mathcal{P}_{\tilde{f}_k}^S(\rho_k) \, d\rho_k & \text{segments,} \end{cases} \quad (19)$$

where ρ_k and $\tilde{\rho}_k$ are the radial frequencies for the k-th direction in the original d- and $(d$-$1)$-dimensional subspace, respectively.

6. Convergence analysis

We use the variance formulations developed in the last section to study the convergence rates of line and segment samples. We first present equations in a unified manner which can directly apply to line and segment samples (as explained in the second half of this section). We start by writing down the expected sampling power spectra for random and jittered point samples for which the analytical form is known [Len66, DW85, DW92, GT04]:

$$\langle \mathcal{P}_{P_N}(\upsilon) \rangle = \begin{cases} \frac{1}{N} & \text{for random,} \\ \frac{1}{N} \left[1 - \prod_i^D \text{sinc}(\pi\upsilon)^2 \right] & \text{for jittered,} \end{cases} \quad (20)$$

where υ represents the frequency and D represents the dimensions. These analytic expressions can be directly fed into (14) for lines and (15) for segments with a simple change in variables (for υ and D) to study corresponding variance. Our analysis, however, is not restricted to only random and jittered sampling patterns. Thanks to the pre-filtering interpretations for lines (12) and segments (13), we can exploit the corresponding *isotropic*[‡] expected power spectra of line offsets and segment centers to leverage (19) for variance analysis. As a result, we can apply the sampling radial profiles and the convergence tool proposed by Pilleboue et al. [PSC*15]—which works for all sampling patterns with *isotropic* expected power spectra—as shown below:

$$\langle \mathcal{P}_{P_{N_k}}(r_k) \rangle = \begin{cases} \frac{\gamma_k}{N_k} \left(\frac{r_k}{\alpha_k N_k^{\frac{1}{D}}} \right)^{b_k} & r_k < \alpha_k N_k^{\frac{1}{D}} \\ \frac{\gamma_k}{N_k} & \text{otherwise} \end{cases}, \quad (21)$$

where r_k is the radial frequency for samples oriented along the k-th direction, b_k is the monomial degree, N_k is the number of samples along the k-th direction, $\gamma_k > 0$ and $\alpha_k \in \mathbb{R}^+/0$ is used to quantify the range of energy-free frequency (around the DC peak) with respect to the mean frequency.

To analyze convergence rates, we first restrict our *pre-filtered* integrands to integrable functions with smooth boundary (e.g. disk or sphere) [BCT01]. This can, however, be extended to arbitrary bounded convex regions [BHI03]. The worst case from this class of functions exhibits a power fall-off of order $\mathcal{O}\left(\rho^{-(d+1)}\right)$ where $\rho > 0$ is a radial frequency, and the best case is defined as a function

---

[‡] In our concurrent work [SJ17], we have shown that *jittered* samples have anisotropic expected power spectra; however, this mild anisotropy does not change the effective stratification asymptotically along any direction. It is therefore safe to *asymptotically* analyze jittered samples using isotropic convergence tools [PSC*15].

Table 1: *Variance convergence rates for N d-dimensional jittered and blue noise (CCVT [BSD09]) point samples [PSC*15].*

Samplers	d	$d = 1$	$d = 2$	$d = 3$
Jittered Best	$\mathcal{O}\left(N^{-1-\frac{2}{d}}\right)$	$\mathcal{O}\left(N^{-3}\right)$	$\mathcal{O}\left(N^{-2}\right)$	$\mathcal{O}\left(N^{-\frac{5}{3}}\right)$
Jittered Worst	$\mathcal{O}\left(N^{-1-\frac{1}{d}}\right)$	$\mathcal{O}\left(N^{-2}\right)$	$\mathcal{O}\left(N^{-1.5}\right)$	$\mathcal{O}\left(N^{-\frac{4}{3}}\right)$
CCVT Best	$\mathcal{O}\left(N^{-1-\frac{3}{d}}\right)$	$\mathcal{O}\left(N^{-4}\right)$	$\mathcal{O}\left(N^{-2.5}\right)$	$\mathcal{O}\left(N^{-2}\right)$
CCVT Worst	$\mathcal{O}\left(N^{-1-\frac{1}{d}}\right)$	$\mathcal{O}\left(N^{-2}\right)$	$\mathcal{O}\left(N^{-1.5}\right)$	$\mathcal{O}\left(N^{-\frac{4}{3}}\right)$

which has constant energy up to a certain radial frequency ρ_0. The overall variance convergence rates can be summarized in the following form:

$$Var(I_N) < \begin{cases} \mathcal{O}\left(N^{-1-\frac{b}{D}}\right) & \text{best-case} \\ \mathcal{O}\left(N^{-1-\frac{1}{D}}\right) & \text{worst-case} \end{cases}, \quad (22)$$

Note that, we can write out the best and worst case convergence rate by using $b = 3$ for CCVT [BSD09] and $b = 2$ for jittered in (22) [PSC*15]. Table 1 summarizes the convergence rates for jittered and CCVT in one, two and three dimensions for points, which is applicable in the case of lines and segments as discussed below. Note, however, that due to the convolution interpretation of lines (12) and segments (13), the integrand is *pre-filtered*, which can smooth out C_0 discontinuities, improving the spectral decay of the effective integrand and therefore its variance and convergence rate.

Line sampling. Following (20), we can write down the expected power spectrum in full d dimensions for line samples with random offset distribution as: $\langle \mathcal{P}_{L_N}(\upsilon) \rangle = \frac{1}{N} \delta(\upsilon \cdot \mathbf{l}^\perp)$, whereas, for jittered samples we can simply replace υ with $\upsilon \cdot \mathbf{l}^\perp$ and D by $(d$-$1)$ in (20). The product of $sinc(\cdot)$ goes over the $(d$-$1)$ dimensions spanning the hyperplane of possible line offsets $\mathbf{l}^\perp$. We illustrate these analytic power spectra for 3D in Fig. 2(a) for multi-directional line sampling where one direction uses randomly generated line offsets and the other direction uses jittered offsets. In the $(d$-$1)$ subspace, the same expected power spectrum can be simplified to the spectrum $\langle \mathcal{P}_{P_{N_k}}(\tilde{\rho}_k) \rangle$ of line offset distributions (for samples along the k-th direction) by simply replacing υ with $\tilde{\upsilon}$ and D with $(d$-$1)$ in (20) for each direction.

We can leverage the radial form of the expected power spectrum from (21) for samplers with unknown analytic form by replacing r_k with $\tilde{\rho}_k$ and D with $(d$-$1)$. Contrary to Pilleboue et al. [PSC*15]—who apply the radial averaging in the full d dimensions—for line samples the radial averaging needs to be performed only in the $(d$-$1)$ subspace. After plugging (21) back into (19), we obtain the convergence tool for line samples which would give the convergence rates shown in (22) for $D = (d$-$1)$ with the *pre-filtered* integrands.

Segment sampling. Similar to line samples, Eq. (21) can be used in (19) to analyze variance from segment samples by replacing r_k with ρ_k and D with d. The overall convergence rates for the best and the worst case can be obtained from (22) by replacing D with d for the *pre-filtered* integrands. Note that, segments samples have exactly the same equations as points, but using the pre-filtered integrand.
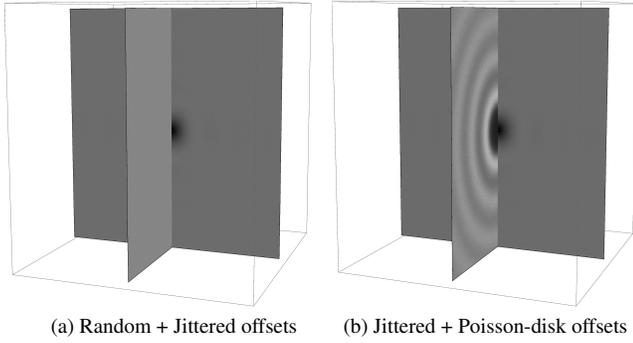
(a) Random + Jittered offsets   (b) Jittered + Poisson-disk offsets

Figure 2: *Bidirectional line sampling leads to expected power spectra constrained to the two planes of sample offsets (shown axis-aligned here for clarity). The line offset sampling process for each direction determines the spectrum within the corresponding plane (shown here for combinations of random, jittered, and Poisson-disk).*

Multi-directional line sampling. For line samples generated over a range of multiple directions, the variance convergence rate can be derived for each direction separately using our variance formulation (19). The overall variance behavior can be summarized as a sum of convergence rates from each $(d\text{-}1)$ subspace for each k-th direction. For example, for two *uncorrelated* line sampling directions shown in Fig. 2, the convergence rate can be derived from the individual 2D subspaces' power spectra in the following form:

$$Var(I_N) = \begin{cases} \mathcal{O}\left(N_1^{-2}\right) + \mathcal{O}(N_2) & \text{best} \\ \mathcal{O}\left(N_1^{-1.5}\right) + \mathcal{O}(N_2) & \text{worst} \end{cases}, \quad (23)$$

where, N_1 is the number of line sampling offsets in the 2D subspace having jittered expected power spectrum and N_2 correspond to the number of line offsets in the other 2D subspace containing random or Poisson disk line offset distribution (Fig. 2(b)). Depending on the integrand, the decay rate of variance can show good convergence
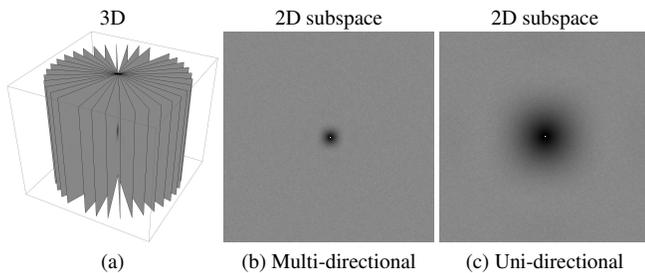

     3D   2D subspace   2D subspace

      (a)   (b) Multi-directional   (c) Uni-directional

Figure 3: *We analyze the multi-directional expected power spectrum (a) with $m = 16$ directions and $N_k = 256$ jittered line samples along each direction k, for a total of $N = 4096$ line samples in 3D. The dark region around DC for each 2D subspace (b) shrinks with more directions compared to the unidirectional case (c).*

for small N_1, N_2 value, but asymptotically we will see the worst of the two, i.e. $\mathcal{O}(N_2)$ convergence rate in both cases.

Also, by increasing the number of directions while keeping the total number of lines N fixed, the overall variance will increase. To explain this behavior, we consider $N = 4096$ line samples that are generated with $m = 16$ different *uncorrelated* directions in 3D. For a unidirectional case, all line offsets will be densely stratified in a 2D subspace orthogonal to the line samples' direction, resulting in the expected power spectrum shown in Fig. 3(c). If, however, multiple *uncorrelated* directions are used for a fixed N, the number of samples per direction—and therefore, the stratification density per direction—will decrease as we increase the number of directions. This would result in smaller dark region around the DC component Fig. 3(b). Since variance is the product of the integrand and sampling power spectra (17), the overall variance will increase.

7. Experiments

We now perform a set of experiments in 2D and 3D integration domains using different point, segment and line sampling patterns to validate our theoretical results. We study different isotropic samplers (random, jitter, CCVT, Poisson disk) and compare these results with other practical samplers like N-rooks, multijitter and 02Sequence[§]. All samplers are homogenized in the random number space. We generate the line sample offsets using 1D point samples for 2D integration problems, and 2D point samples for 3D integration problems. We start by validating the best and worst case variance convergence rates with analytical functions and later verify these convergence rates with realistic test scenes including direct illumination, ambient occlusion, and homogeneous participating medium. We perform all the renderings using PBRT-v3 [PJH16] and the corresponding variance analysis using the empirical error analysis code from Subr et al. [SSJ16].

Best and worst cases. To validate the best and the worst case convergence rates shown in Table 1 we consider a 3D ($d = 3$) integration domain $[0, 1)^3$ with a sphere as a worst case [PSC*15] and a 3D Gaussian function as a best case. While a Gaussian has a frequency spectrum with infinite extent, it is smooth enough (C^∞) to obtain the theoretical best-case convergence rates.

In Fig. 4, we first integrate using unidirectional line samples. For the sphere (a), jittered samples give the convergence rate of $\mathcal{O}\left(N^{-1.9}\right)$ instead of the 2D ($d = 2$) worst case convergence rate of $\mathcal{O}\left(N^{-1.5}\right)$ with line samples. This is because line samples *pre-filter* the C^0 discontinuities which results in further improvements on top of dimensionality reduction. For a Gaussian (b), the observed convergence rate is $\mathcal{O}\left(N^{-2}\right)$ which is the best case convergence rate in $d - 1$ (2D). Since a Gaussian has no C^0 discontinuities, pre-filtering does not further improve the rate of convergence.

[§] We use the non-scrambled version of 02Sequence sampler that directly ships with PBRT-v3, which is based on a paper by Kollig and Keller [KK02] and uses the first two dimensions derived from the Sobol sampler.
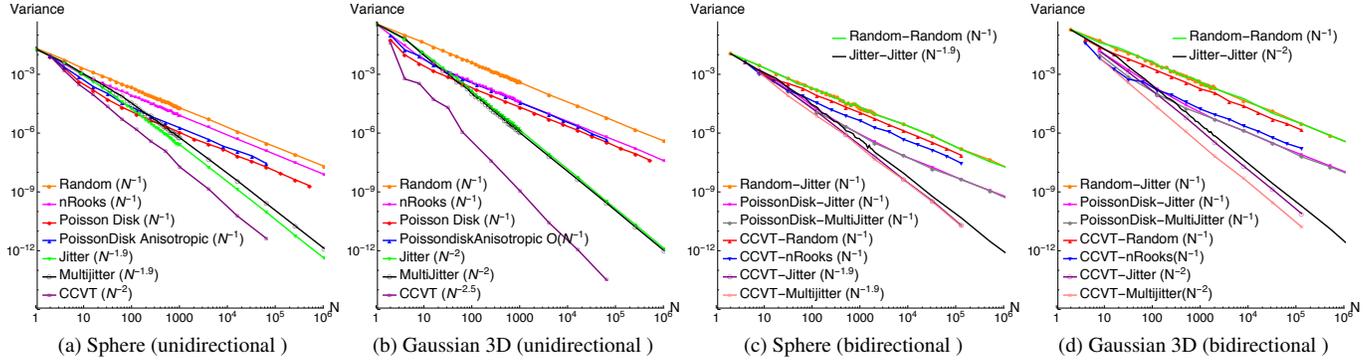
Figure 4: *Line samples 3D: Empirical variance convergence rates for the worst (Sphere) and the smooth (Gaussian 3D) case functions integrated using various unidirectional (a,b) and bidirectional (c,d) line sampling patterns. With jitter and CCVT, Gaussian (b) follows the convergence rate for d = 2 from Table 1 for this 3D integral. However, for the sphere (a), the experimental convergence rate is much better than the theoretical one shown in Table 1 for d = 2. This is because line samples not only reduce the dimensionality but also prefilter the C_0 discontinuities in the case of sphere. In (c,d), the convergence rate is dominated by the direction with samples having worst convergence.*
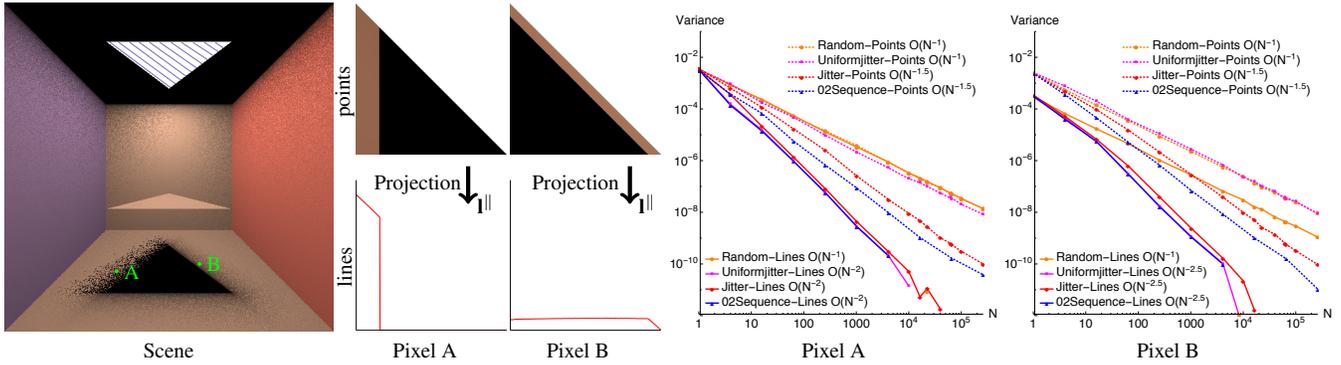

Figure 5: *Line orientation: We analyze the impact of line sample orientation on variance and convergence rate. Line samples on the light source are generated parallel to the edge above pixel A. Consequently, the regions below the left side of the triangle occluder see a step function (shown for Pixel A in second column) and benefit only from dimensionality reduction. The other two sides of the occluder benefit from both dimensionality reduction and* smoothing *of the integrand (shown for Pixel B in the second column).*

We further extend this experiment to the multi-directional setting with line samples generated from two directions. We use two *uncorrelated* samplers along the two directions and visualize the convergence behavior in Fig. 4(c,d). As predicted by our analysis, the overall variance convergence rate is ultimately determined by the line sampler with the worse convergence among the two directions. For example, even though CCVT gives good convergence rate for unidirectional samples, when coupled in the bidirectional case with random or Poisson-disk samples, the overall convergence rate drops to $\mathcal{O}\left(N^{-1}\right)$.

Direct illumination. We have rendered two direct illumination scenes (Figs. 5 and 7) following the approach by Billen and Dutré [BD16] which performs line sampling directly on an area light source and computes the visibility function analytically for diffuse and phong BRDFs. In Fig. 5, we consider a scene with a triangle light source and a triangle occluder of the same size and study the impact of line sampling orientation (shown as blue lines on the

light source). The first column shows the rendered scene with $N = 1$ line sample for each shading point. As Billen and Dutré [BD16] previously demonstrated, the amount of noise is different depending on the edges of the occluder. We are interested in examining how the orientation of lines impacts the effective integrand and therefore the convergence rate in MC integration. To this end, we show in the second column the original 2D integrands (as seen by traditional point sampling; top) and the corresponding projected 1D integrands (as seen by line samples that pre-filter vertically and sample in one dimension lower; bottom) for two different pixels.

For Pixel A, line samples are exactly parallel to the occluder edge which leads to a C^0 step discontinuity in the effective 1D integrand. This results in the worst-case variance convergence rate of $\mathcal{O}\left(N^{-2}\right)$ when using 1D jittered samples, as predicted by our theory by setting $d = 1$ in Table 1. This implies that, Pixel A benefits only due to dimensionality reduction.
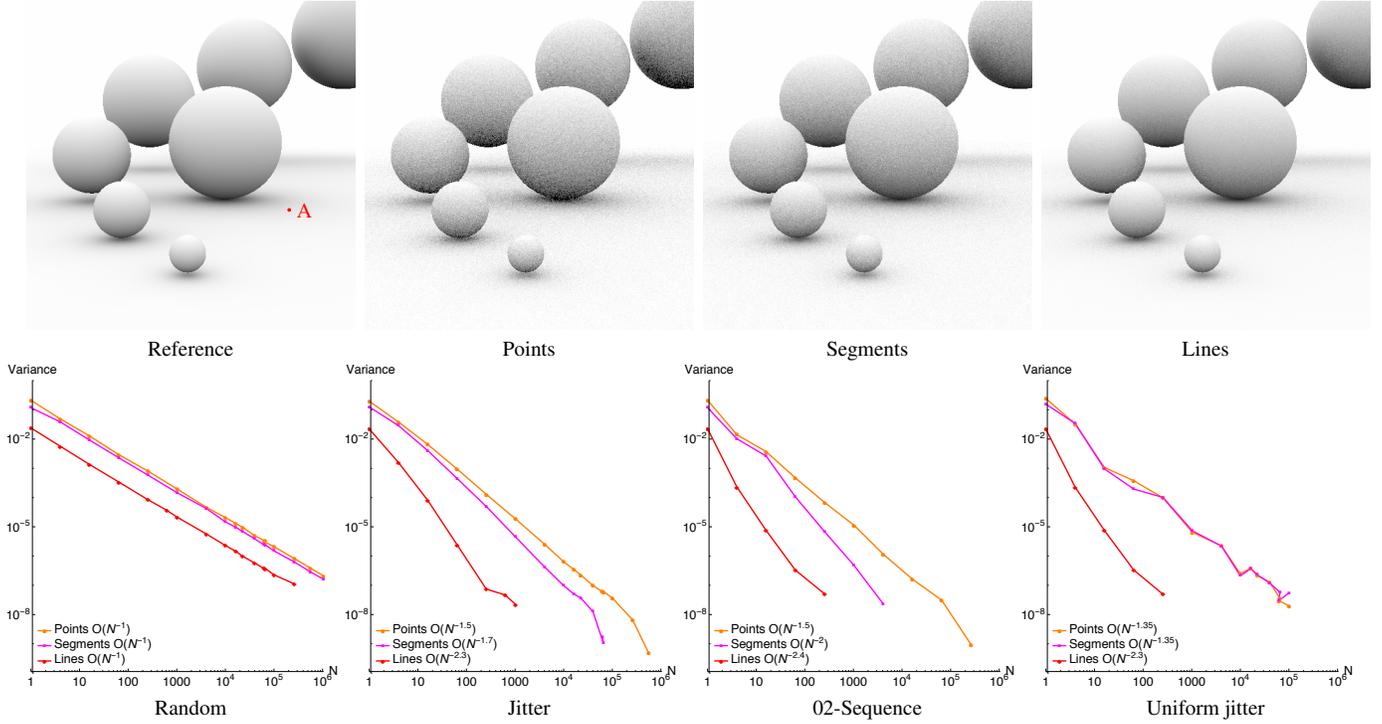
Figure 6: *Points, Segments & Lines: In the top row, we show the ambient occlusion scene which is rendered with N = 16 jittered points, segments (of length = 1 radian) and line samples. The reference image is generated with N = 1024 jittered line samples. The bottom row compares the improvements using segment and line samples over point samples for four different samplers at Pixel A (marked in red in the reference image). For more comparisons please refer to the supplemental Fig. 1.*
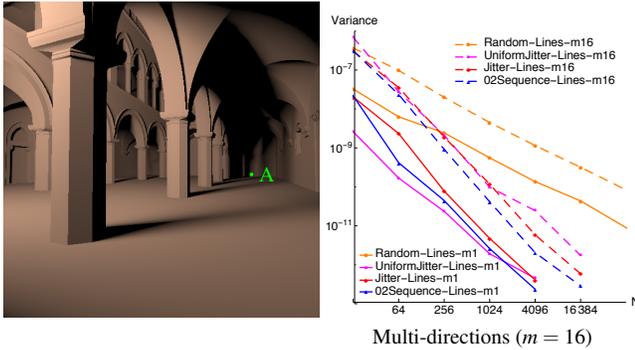

Figure 7: *Multi-directions: We generate 16 point samples per pixel and perform line sampling in a direction associated to each pixel sample with a fixed sample count.*

from Table 1. In Fig. 7, we compare variance in the Sponza scene for both unidirectional and multi-directional lines for a fixed number of jittered samples. We first sample each pixel with 16 point samples and then associate with each pixel sample a direction to perform line sampling on the light source. We keep the directions fixed over 100 trials while computing variance to avoid the impact of directional sampling on variance plots. For $m = 16$ directions, the corresponding variance for a given number of line samples N is much higher in the multi-directional case compared to the unidirectional case as explained in Sec. 6.

For Pixel B, however, the line samples are oriented in such a way that the *effective* integrand is pre-filtered, removing the C_0 discontinuity. This results in a convergence rate of $\mathcal{O}\left(N^{-2.5}\right)$ for this 2D integral, which is much better than Pixel A. Equation (6) provides a theoretical justification for this improvement. Similar results can be observed for other samplers. For point samples, however, the effective integrand is 2D and we observe 2D ($d = 2$) convergence rates

Ambient occlusion. We also implemented ambient occlusion integration using point, segment and line samples. Our approach resembles that of Gribel et al. [GBAM11] which uses line samples over the hemisphere to compute occlusion from triangle meshes, but our implementation only considers analytic spheres for simplicity. At each shade point we are interested in the integral of visibility across the upper hemisphere:

$$I = \int_{\mathcal{H}^2} \frac{V(\vec{\omega})\cos\theta}{\pi}\, d\vec{\omega} = \int_0^\pi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{V(\theta,\phi)\cos\theta\,|\sin\theta|}{\pi}\, d\theta\, d\phi. \quad (24)$$
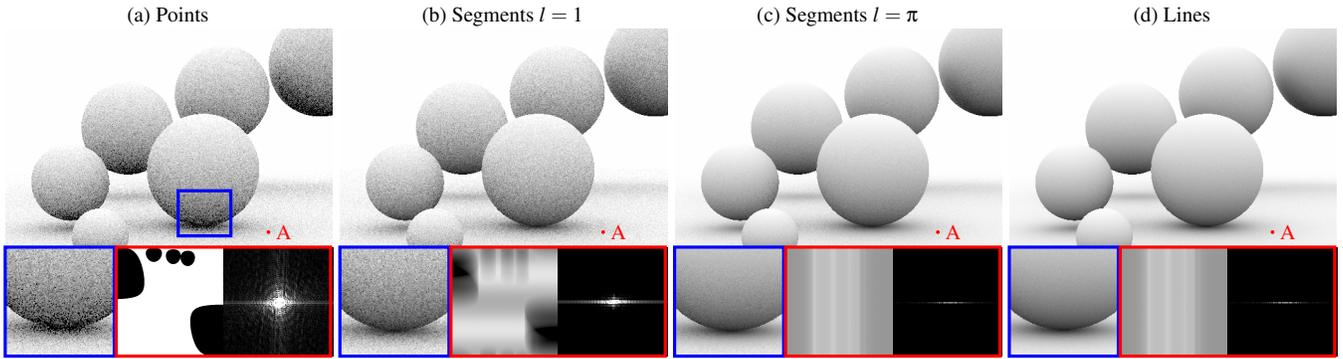
Figure 8: *Points, Segments & Lines: We compare the impact of segment length (l) on the variance for the ambient occlusion scene. All renderings are performed with N = 9 jittered samples. Inset in blue is the zoom-in region of the sphere at the center (marked in (a)). The insets in red visualize the integrand for Pixel A and the corresponding power spectrum. As the segment length increases, variation of the pre-filtered integrand (and its power spectrum) diminishes along the vertical axis due to the convolution by the segments. For length l = π, segments cover the full θ range like line samples, however, the variance due to these long segments is still higher compared to line samples. This is because, by construction, segment centers span a 2D domain while line offsets span a 1D domain, and can leverage denser 1D stratification. For more comparisons please refer to the supplemental Figs. 2, 3 & 4.*

The Monte Carlo estimators using points, lines and segments are:

$$I_N = \begin{cases} \dfrac{1}{\pi N} \sum_i^N \dfrac{V(\theta_i,\phi_i)\cos\theta_i\,|\sin\theta_i|}{\mathrm{pdf}(\theta_i,\phi_i)} & \text{for points,} \\[2.5ex] \dfrac{1}{\pi N} \sum_i^N \dfrac{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta,\phi_i)\cos\theta\,|\sin\theta|\,d\theta}{\mathrm{pdf}(\phi_i)} & \text{for lines, and} \\[2.5ex] \dfrac{1}{\pi N} \sum_i^N \dfrac{\int_{\theta_i^-}^{\theta_i^+} V(\theta,\phi)\cos\theta\,|\sin\theta|\,d\theta}{\mathrm{pdf}(\theta_i,\phi_i)\,\Delta\theta_i} & \text{for segments,} \end{cases} \quad (25)$$

where $\Delta\theta_i$ is the arc-length of the *i*-th segment, and $(\theta_i^-,\theta_i^+) = \theta \pm \Delta\theta_i$. In our implementation we ensure all segments have the same arc-length by wrapping them across the north pole and equator.

Fig. 6 compares the results from these three different types of estimators using a variety of different sampling patterns. We perform *cosine-weighted* sampling for point samples whereas for segments we uniformly sample in spherical coordinates (θ and φ). For lines we only need to choose $\phi \in [0,\pi)$, which we always sample uniformly. We can see that sampling using segments or lines improves convergence rate due to smoothing, and lines further improve convergence since their dimensionality reduction provides the opportunity for denser stratification. In Fig. 8, we compare variance due to different segment lengths for jittered samples. Note that for segments that span the full $\theta \in [-\pi/2,\pi/2]$ range (c), the underlying integrand at Pixel A (c,d) is the same for both segment and line samples; however, the corresponding variance is higher with segments. This is because segment centers span a 2D space, whereas line offsets span a 1D space and can therefore leverage a denser 1D stratification. However, if we use samplers that densely stratify both in 1D and 2D (e.g., multijitter), then these *long* segments (length = π) would perform exactly as line samples.

Participating medium. We also verified our predicted convergence rates for a simple participating medium rendering problem as shown in Fig. 9. In this setup, we have a point light source in a homogeneous medium and place a glass sphere (refractive index = 1.5) at the center of the domain, casting volumetric caustics. We use two different kinds of sampling strategies to simulate this illumination problem, corresponding to volumetric photon mapping using points [JZJ08], and the long-photon beam variant [JNSJ11, JNT*11] which solves transmittance along each direction analytically, forming line samples. We also consider both a 2D (flatland) light transport simulation, as well as a standard 3D simulation.

As we move from flatland to full 3D simulation, the variance convergence rates deteriorate as we would expect. Conversely, when we replace point samples with line samples in either domain, the convergence rates improve. It is interesting to note that the convergence rates in the 2D (flatland) sphere caustics scene rendered with photon points (Fig. 9(a)) is the same as the 3D sphere caustic scene rendered with beams/lines (Fig. 9(d)). This is because, as our theory predicts, the convergence rate is dependent on the sampling dimensionality (2D in both cases), and not the dimensionality of the original integrand. Using line samples for 3D light transport requires sampling the same number of dimensions (2) for the line offsets as the number of dimensions (2) required to simulated flatland transport with photon points. We do not see an additional improvement due to smoothing of the integrands because the smoothing is performed along (and not across) the visibility discontinuities. All scenarios therefore retain discontinuities in their integrands, leading to the *worst-case* convergence behavior for all samplers and domains.

8. Conclusion, limitations and future work

We expressed Monte Carlo integration using line and segment samples in both the spatial and frequency domains. By leveraging a dual, pre-filtered interpretation of the resulting estimators, we showed
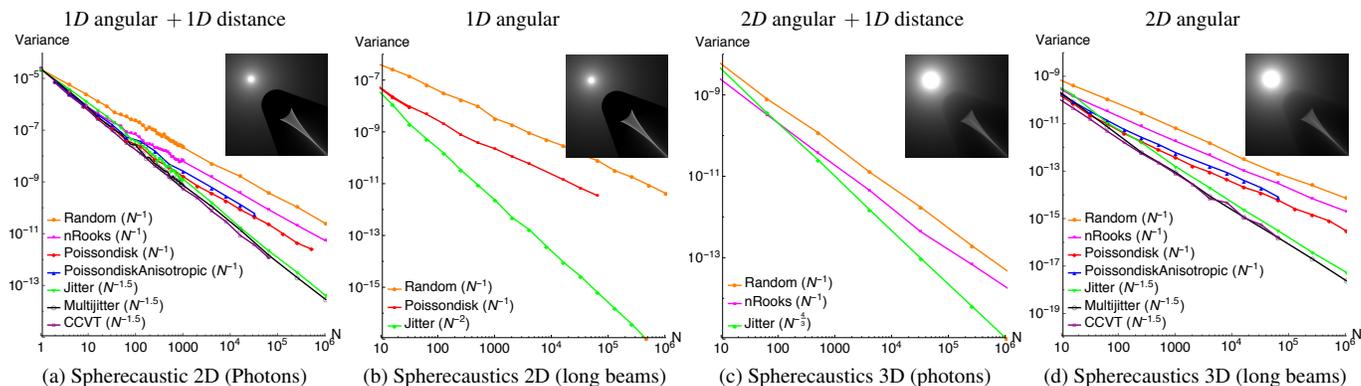
Figure 9: *Points & Lines: We analyze the variance convergence rates for a homogeneous participating medium which is rendered in 2D (flatland) and 3D using photons (a, c) and long beams (which are photon beams that analytically compute transmittance along each direction) in (b, d). On the top, we are showing the sampled dimensions for each case.*

how to leverage prior isotropic variance and convergence tools for point sampling to derive concrete expressions for the variance and convergence rates of a multitude of different line and segment sampling strategies. However, following the concurrent work by Singh and Jarosz [SJ17] it is now also possible to theoretically analyze line and segment samples directly in the original d-dimensions and in the $(d\text{-}1)$ subspaces using both the isotropic and anisotropic sample distributions (e.g. N-rooks, multijitter). Our theory reveals that we can expect convergence improvements from lines due to dimensionality reduction, but also improvements for both segments and lines due to their pre-filtering of the integrand. We validated this theory on a number of common rendering scenarios including ambient occlusion, direct illumination, and participating media rendering.

Correlated multi-directional sampling. A limitation of our multidirectional analysis is that it requires the samples from different directions to be uncorrelated. Particularly for segment samples it could be beneficial to sample directions and offsets in a joint fashion to obtain a combined sampling power spectrum with desirable variance/convergence properties. One possible realization of this might be to use multi-class blue-noise sampling as examined by Sun et al. [SZG*13]. More work is needed to evaluate such approaches and derive the variance of the corresponding, correlated multi-directional estimators.

Implications for importance sampling. Our pre-filtering interpretation also suggests that obtaining the most benefit from line or segment samples will require modifying the sampling PDF. We ran into this problem when implementing segment samples for the ambient occlusion experiment. Traditional point sampling typically importance samples the known terms in the integrand (i.e. $\cos\theta\sin\theta$ arising from foreshortening and the Jacobian of spherical integration), which allows those terms in the integrand to cancel out exactly with the sampling PDF, reducing variance. When using segment samples however, the effective integrand is filtered, so it no longer contains $\cos\theta\sin\theta$, but a convolved version of these terms. Using segment samples with the same importance sampling density as points can actually lead to increases in variance since these terms

no longer cancel out (see supplemental Figs. 3 and 5 for more details). We used a simple constant PDF in spherical coordinates when sampling segment centers to ensure that the sample weights have bounded variation, but this also suggests that new strategies need to be devised that sample proportional to the pre-filtered integrand to obtain maximum benefit.

Choosing sample directions. Our analysis also reveals that the choice of line/segment direction can impact whether we obtain variance/convergence improvements from smoothing. If the convolution induced by line/segment samples is parallel to discontinuities in the integrand, then this will not lead to further convergence improvements. We saw that choosing sample directions deterministically can lead to poor convergence in certain pixels, and distributing the sample budget across a multitude of directions deteriorates convergence to that of random sampling in the limit. In the absence of additional knowledge, using a single (but random) direction for all samples in a pixel is a reasonable strategy to reduce the risk of selecting a bad direction while ensuring the maximum possible stratification density. Intelligently choosing the sample directions (given some a priori knowledge of the integration problem) may be a promising strategy to extract further benefit from line and segment samples.

9. Acknowledgements

References

[BCT01] BRANDOLINI L., COLZANI L., TORLASCHI A.: Mean square decay of Fourier transforms in Euclidean and non Euclidean spaces. *Tohoku Math. J. (2) 53*, 3 (2001), 467–478. doi:10.2748/tmj/1178207421. 5

[BD16] BILLEN N., DUTRÉ P.: Line sampling for direct illumination. *Comp. Graph. Forum (EGSR) 35*, 4 (2016), 45–55. doi:10.1111/cgf.12948. 2, 7

[BGAM12] BARRINGER R., GRIBEL C. J., AKENINE-MÖLLER T.: High-quality curve rendering using line sampled visibility. *ACM Transaction on Graphics (SIGGRAPH Asia) 31*, 6 (Nov. 2012), 162:1–162:10. doi:10.1145/2366145.2366181. 1, 2

[BHI03] BRANDOLINI L., HOFMANN S., IOSEVICH A.: Sharp rate of average decay of the Fourier transform of a bounded set. *Geometric & Functional Analysis GAFA 13*, 4 (2003), 671–680. 5

[BSD09] BALZER M., SCHLÖMER T., DEUSSEN O.: Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Transaction on Graphics (SIGGRAPH) 28*, 3 (July 2009), 86:1–86:8. doi:10.1145/1531326.1531392. 2, 3, 5

[Coo86] COOK R. L.: Stochastic sampling in computer graphics. *ACM Transactions on Graphics 5*, 1 (Jan. 1986), 51–72. doi:10.1145/7529.8927. 2

[CP76] CRANLEY R., PATTERSON T. N. L.: Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis 13*, 6 (1976), 904–914. 3

[CPC84] COOK R. L., PORTER T., CARPENTER L.: Distributed ray tracing. *Computer Graphics (SIGGRAPH) 18*, 3 (1984), 137–145. doi:10.1145/964965.808590. 1

[dGBOD12] DE GOES F., BREEDEN K., OSTROMOUKHOV V., DESBRUN M.: Blue noise through optimal transport. *ACM Transaction on Graphics (SIGGRAPH Asia) 31* (Nov. 2012), 171:1–171:10. doi:10.1145/2366145.2366190. 3

[Dur11] DURAND F.: *A frequency analysis of Monte-Carlo and other numerical integration schemes.* Tech. Rep. TR-2011-052, MIT CSAIL, 2011. 2

[DW85] DIPPÉ M. A. Z., WOLD E. H.: Antialiasing through stochastic sampling. *Computer Graphics (SIGGRAPH)* (1985), 69–78. doi:10.1145/325334.325182. 2, 3, 5

[DW92] DIPPÉ M. A. Z., WOLD E. H.: Progress in computer graphics (vol. 1). Ablex Publishing Corp., Norwood, NJ, USA, 1992, ch. Stochastic Sampling: Theory and Application, pp. 1–54. 5

[GBAM11] GRIBEL C. J., BARRINGER R., AKENINE-MÖLLER T.: High-quality spatio-temporal rendering using semi-analytical visibility. *ACM Transaction on Graphics (SIGGRAPH) 30* (July 2011), 54:1–54:12. doi:10.1145/2010324.1964949. 2, 8

[GDAM10] GRIBEL C. J., DOGGETT M., AKENINE-MÖLLER T.: Analytical motion blur rasterization with compression. In *High-Performance Graphics* (2010), pp. 163–172. doi:10.2312/EGGH/HPG10/163-172. 1, 2

[GKH*13] GEORGIEV I., KŘIVÁNEK J., HACHISUKA T., NOWROUZEZAHRAI D., JAROSZ W.: Joint importance sampling of low-order volumetric scattering. *ACM Transaction on Graphics (SIGGRAPH Asia) 32*, 6 (Nov. 2013). doi:10.1145/2508363.2508411. 1, 2

[GT04] GABRIELLI A., TORQUATO S.: Voronoi and void statistics for superhomogeneous point processes. *Physical Review E 70*, 4 (2004), 041105. 5

[HCJ13] HABEL R., CHRISTENSEN P. H., JAROSZ W.: Photon beam diffusion: A hybrid monte carlo method for subsurface scattering. *Computer Graphics Forum (Proceedings of EGSR) 32*, 4 (June 2013). doi:10.1111/cgf.12148. 2

[HSD13] HECK D., SCHLÖMER T., DEUSSEN O.: Blue noise sampling with controlled aliasing. *ACM Transactions on Graphics 32*, 3 (July 2013), 25:1–25:12. doi:10.1145/2487228.2487233. 2

[IPSS08] ILLIAN J., PENTTINEN P., STOYAN H., STOYAN D.: *Statistical Analysis and Modelling of Spatial Point Patterns.* Statistics in Practice. Wiley, 2008. 2

[JNSJ11] JAROSZ W., NOWROUZEZAHRAI D., SADEGHI I., JENSEN H. W.: A comprehensive theory of volumetric radiance estimation using photon points and beams. *ACM Transactions on Graphics 30*, 1 (Feb. 2011), 5:1–5:19. doi:10.1145/1899404.1899409. 1, 2, 9

[JNT*11] JAROSZ W., NOWROUZEZAHRAI D., THOMAS R., SLOAN P.-P., ZWICKER M.: Progressive photon beams. *ACM Transaction on Graphics (SIGGRAPH Asia) 30*, 6 (Dec. 2011). doi:10.1145/2070781.2024215. 9

[JP00] JONES T. R., PERRY R. N.: Antialiasing with line samples. In *EGWR* (London, UK, 2000), Springer-Verlag, pp. 197–206. doi:10.1007/978-3-7091-6303-0_18. 1, 2

[JZJ08] JAROSZ W., ZWICKER M., JENSEN H. W.: The beam radiance estimate for volumetric photon mapping. *Comp. Graph. Forum (Eurographics) 27*, 2 (Apr. 2008), 557–566. doi:10.1111/j.1467-8659.2008.01153.x. 2, 9

[KGH*14] KŘIVÁNEK J., GEORGIEV I., HACHISUKA T., VÉVODA P., ŠIK M., NOWROUZEZAHRAI D., JAROSZ W.: Unifying points, beams, and paths in volumetric light transport simulation. *ACM Transaction on Graphics (SIGGRAPH) 33*, 4 (July 2014), 103:1–103:13. doi:10.1145/2601097.2601219. 1, 2

[KK02] KOLLIG T., KELLER A.: Efficient Multidimensional Sampling. *Computer Graphics Forum* (2002). doi:10.1111/1467-8659.00706. 6

[LD08] LAGAE A., DUTRÉ P.: A comparison of methods for generating Poisson disk distributions. *Comp. Graph. Forum 27*, 1 (2008), 114–129. doi:10.1111/j.1467-8659.2007.01100.x. 2

[Len66] LENEMAN O. A.: Random sampling of random processes: Impulse processes. *Information and Control 9*, 4 (1966), 347–363. 5

[Mit91] MITCHELL D.: Spectrally optimal sampling for distributed ray tracing. *Computer Graphics (SIGGRAPH) 25*, 4 (1991), 157–164. doi:10.1145/127719.122736. 2

[Mit96] MITCHELL D. P.: Consequences of stratified sampling in graphics. In *Computer Graphics (SIGGRAPH)* (1996), ACM, pp. 277–280. doi:10.1145/237170.237265. 2

[NBMJ14] NOWROUZEZAHRAI D., BARAN I., MITCHELL K., JAROSZ W.: Visibility silhouettes for semi-analytic spherical integration. *Comp. Graph. Forum 33*, 1 (Feb. 2014), 105–117. doi:10.1111/cgf.12257. 2

[NNDJ12] NOVÁK J., NOWROUZEZAHRAI D., DACHSBACHER C., JAROSZ W.: Virtual ray lights for rendering scenes with participating media. *ACM Transaction on Graphics 31*, 4 (July 2012). doi:10.1145/2185520.2185556. 1, 2

[PJH16] PHARR M., JAKOB W., HUMPHREYS G.: *Physically Based Rendering: From Theory To Implementation*, 3rd ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2016. 6

[PSC*15] PILLEBOUE A., SINGH G., COEURJOLLY D., KAZHDAN M., OSTROMOUKHOV V.: Variance analysis for Monte Carlo integration. *ACM Transaction on Graphics (SIGGRAPH) 34*, 4 (July 2015), 124:1–124:14. doi:10.1145/2766930. 1, 2, 3, 4, 5, 6

[RAMN12] RAMAMOORTHI R., ANDERSON J., MEYER M., NOWROUZEZAHRAI D.: A theory of Monte Carlo visibility sampling. *ACM Transactions on Graphics 31*, 5 (2012), 121:1–121:16. doi:10.1145/2231816.2231819. 2

[Shi91] SHIRLEY P.: Discrepancy as a quality measure for sample distributions. In *Proceedings of Eurographics* (1991), pp. 183–194. 2

[SJ17] SINGH G., JAROSZ W.: Convergence analysis for anisotropic Monte Carlo sampling spectra. *ACM Transaction on Graphics (SIGGRAPH) 36*, 4 (July 2017), 137:1–137:14. doi:10.1145/3072959.3073656. 5, 10

[SK13] SUBR K., KAUTZ J.: Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Transaction on Graphics (SIGGRAPH) 32*, 4 (July 2013), 128:1–128:12. doi:10.1145/2461912.2462013. 2

[SNJ*14] SUBR K., NOWROUZEZAHRAI D., JAROSZ W., KAUTZ J., MITCHELL K.: Error analysis of estimators that use combinations of stochastic sampling strategies for direct illumination. *Comp. Graph. Forum (EGSR) 33*, 4 (June 2014), 93–102. doi:10.1111/cgf.12416. 2

[Spa66] SPANIER J.: Two pairs of families of estimators for transport problems. *SIAM Applied Math. 14*, 4 (1966), 702–713. `doi:10.1137/0114059`. 2

[SSJ16] SUBR K., SINGH G., JAROSZ W.: Fourier analysis of numerical integration in Monte Carlo rendering: Theory and practice. In *ACM SIGGRAPH Courses* (New York, NY, USA, July 2016), ACM. `doi:10.1145/2897826.2927356`. 6

[SZG*13] SUN X., ZHOU K., GUO J., XIE G., PAN J., WANG W., GUO B.: Line segment sampling with blue-noise properties. *ACM Transaction on Graphics (SIGGRAPH) 32*, 4 (July 2013), 127:1–127:14. `doi:10.1145/2461912.2462023`. 1, 2, 3, 10

[SZLG10] SUN X., ZHOU K., LIN S., GUO B.: Line space gathering for single scattering in large scenes. *ACM Transaction on Graphics (SIGGRAPH) 29*, 4 (July 2010), 54:1–54:8. `doi:10.1145/1778765.1778791`. 2

[TPD*12] TZENG S., PATNEY A., DAVIDSON A., EBEIDA M. S., MITCHELL S. A., OWENS J. D.: High-quality parallel depth-of-field using line samples. In *High-Performance Graphics* (2012), pp. 23–31. `doi:10.2312/EGGH/HPG12/023-031`. 1, 2