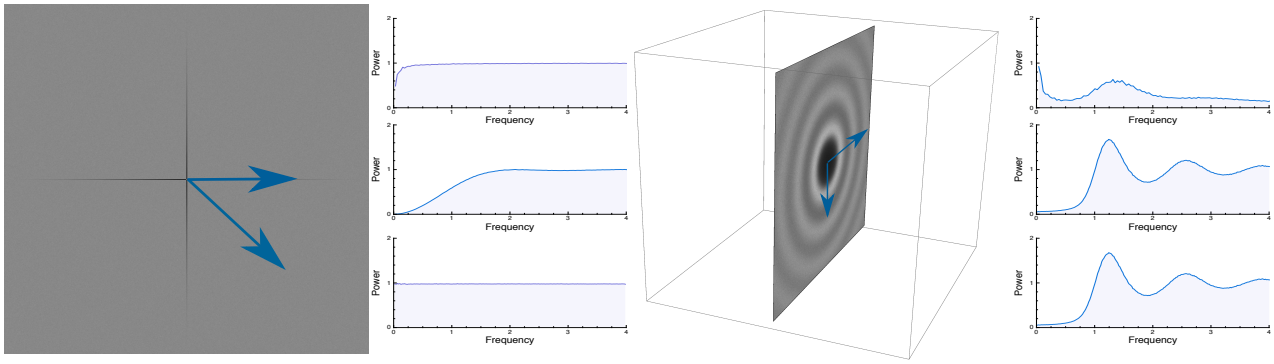
Monte Carlo Convergence Analysis for Anisotropic Sampling Power Spectra

Gurprit Singh Wojciech Jarosz
Dartmouth College
Dartmouth Computer Science Technical Report TR2016-816



(a) n-rooks (2D point sampling power spectra) (b) Poisson disk (3D line sampling power spectra)

Figure 1: *Illustration of (a) an expected anisotropic Fourier power spectrum for 2D n-rook ($N = 256$) point samples (**left**). The accompanying radial power spectra correspond to the radial average (top), radial behavior along the horizontal arrow (middle), and radial behavior along the diagonal arrow (bottom) of the 2D n-rooks power spectrum. Similarly, on the **right**, (b) we show the expected Fourier power spectrum for 3D parallel line samples (generated horizontally, orthogonal to the plane containing the power) using isotropic Poisson disk samples ($N = 4096$), with radial power spectra corresponding to radial averaging (top) within the 3D domain (power scaled for visualization) and radial behavior along two directions (middle and bottom) within the subspace containing power.*

Abstract

Traditional Monte Carlo (MC) integration methods use point samples to numerically approximate the underlying integral. This approximation introduces variance in the integrated result, and this error can depend critically on the sampling patterns used during integration. Most of the well known samplers used for MC integration in graphics, e.g. jitter, Latin hypercube (n-rooks), multi-jitter, are anisotropic in nature. However, there are currently no tools available to analyze the impact of such anisotropic samplers on the variance convergence behavior of Monte Carlo integration. In this work, we propose a mathematical tool in the Fourier domain that allows analyzing the variance, and subsequently the convergence rate, of Monte Carlo integration using any arbitrary (anisotropic) sampling power spectrum. We apply our analysis to common anisotropic *point* sampling strategies in Monte Carlo integration, and extend our analysis to recent Monte Carlo approaches relying on line samples which have inherently anisotropic power spectra. We validate our theoretical results with several experiments using both point and line samples.

Keywords: Stochastic Sampling, Monte Carlo Integration, Line Sampling, Global Illumination, Rendering

1 Introduction

Since being introduced to graphics by Cook and collegues [1984], Monte Carlo (MC) integration has become the cornerstone of most modern rendering algorithms. Historically, MC integration in rendering has involved sampling a function at various stochastically placed *points* to approximate an integral, e.g. the radiance through a pixel. This estimation is error-prone, however, and many researchers [Cook 1986; Dippé and Wold 1985; Subr and Kautz 2013; Subr et al. 2014; Pilleboue et al. 2015] have therefore investigated how the properties of the integrand and sample points impact the error and convergence rate of this estimation. These analyses have provided important theoretical insights and have lead to tangible improvements in rendering. Many of these analyses have leveraged the Fourier domain to better understand the underlying characteristics of different sampling patterns. Within this domain, the *radially averaged* Fourier power spectrum [Ulichney 1987] has been perhaps the most essential tool to analyze point samples, characterizing various stochastic sampling patterns ranging from white noise to blue noise, and more recently being used to derive variance convergence rates of various stochastic samplers [Pilleboue et al. 2015].

While *radial averaging* is appropriate for analyzing *isotropic* Fourier power spectra, many of the stochastic point sampling strategies used in rendering—such as N-rooks [Shirley 1991] or even jittered sampling [Cook 1986]—are in fact *anisotropic*. For anisotropic sampling power spectra, radial averaging can be less informative, or worse, misleading. For example, in Fig. 1(a), the 2D N-rooks sampling pattern has *radial* behavior of a jittered sampling power spectrum along the canonical axes, but a flat, *white noise* radial behavior in other directions. This information is lost in the *radially averaged* power spectrum shown at the top of the radial plots.

While point sampling has dominated the view of MC in graphics for years, Monte Carlo integration need not always be a point sampling process. In fact, researchers have recently started employing Monte Carlo-like estimators using *line samples* for rendering problems as diverse as anti-aliasing [Jones and Perry 2000], motion blur [Gribel et al. 2010], depth of field [Tzeng et al. 2012], hair rendering [Barringer et al. 2012], as well as density estimation [Jarosz et al. 2011a] or path sampling [Novák et al. 2012; Georgiev et al. 2013; Křivánek et al. 2014] 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(b), for instance, shows line samples whose locations have a Poisson disk distribution (with isotropic power spectrum within the subspace), but the *radially averaged* power spectrum (top) does not reveal this fact.

In this paper, we study such inherently anisotropic sampling processes in Monte Carlo integration and the impact they have on variance and convergence rate. Our primary contribution is a mathematical derivation of variance which enables Monte Carlo convergence rate analysis of arbitrary (isotropic or anisotropic) sampling power spectra. We apply our formulation to Monte Carlo integration using anisotropic point sampling, deriving convergence rates which can inform design principles for applying these sampling strategies in practical rendering problems. We also leverage our anisotropic variance analysis formulation to the inherently anisotropic properties of Monte Carlo integration using line samples. We derive expressions predicting the convergence rate of Monte Carlo line sampling, relating this to the sampling strategy used to place each line sample. Our theoretical results and empirical validations reveal that line sampling can not only *decrease variance*, but it can also *improve convergence rate*.

2 Related Work

Point sampling, variance, & convergence. Since the introduction of MC to graphics [Cook et al. 1984], researchers have noted that a careful arrangement of samples can impact the spectral distribution and dramatically reduce the overall magnitude of error in numerical integration [Dippé and Wold 1985; Cook 1986; Mitchell 1991]. This has lead to extensive work on generating sample patterns which are stochastic, yet still maintain a low discrepancy [Shirley 1991] or which exhibit so-called blue noise frequency spectra [Cook 1986; Lagae and Dutré 2008]. Recent work [Durand 2011; Subr and Kautz 2013; Pilleboue et al. 2015] has established a firm mathematical connection between the spectral properties of the sampling pattern and the *magnitude* of MC integration error. Moreover, carefully sample placement—such as jittered [Cook 1986] and certain flavors of blue-noise sampling [Balzer et al. 2009; Heck et al. 2013]—have now been shown to actually lead to asymptotically faster *convergence rates* [Mitchell 1996; Ramamoorthi et al. 2012; Subr and Kautz 2013; Subr et al. 2014; Pilleboue et al. 2015]. We derive similar mathematical expressions governing variance and convergence rate, but for the case of stochastic placement and evaluation of point or line samples that might have an arbitrary anisotropic expected power spectrum.

Generation/analysis of anisotropic sampling patterns. Many sampling methods used in graphics are inherently anisotropic (e.g. jitter, n-rooks, multi-jitter), and anisotropic variants [Feng et al. 2008; Li et al. 2010; Wachtel et al. 2014] of popular isotropic sampling patterns also exist. Such approaches have been shown to be more suitable for certain geometry processing applications [Alliez et al. 2003; Lévy and Liu 2010], or for instancing anisotropic geometric primitives [Li et al. 2010]. Sun et al. [2013] generate line (segment) samples with blue-noise properties, and projective relaxation [Reinert et al. 2016] enforces blue noise properties under multiple planar projections of a point sampling pattern, both of which result in anisotropic sampling spectra. Isotropic Fourier tools are ill-equipped to analyze the error arising from Monte Carlo integration using such sampling patterns.

A few analysis approaches have been developed, however, to handle specific forms of anisotropy outside the context of Monte Carlo variance analysis. For anisotropic distributions that arise from a *global*, invertible warp, one can warp the samples back to the uniform domain to perform isotropic spectral analysis [Li et al. 2010]. Wei and Wang [2011] proposed a more general framework that operates on the inter-sample distances in order to analyze non-uniform/adaptive sampling. With knowledge of the warping function's Jacobian, they can approximately eliminate differential anisotropy (i.e. local non-uniform scaling or sheering) of the samples for analysis. While these approaches were concerned with spectral analysis, they do not seek to establish how such anisotropy affects variance and convergence rate in Monte Carlo integration.

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 [2000]), it has been used for some time in related fields. A class of Monte Carlo methods from neutron transport simulation known as "expected value estimators" and "track length estimators" [Spanier 1966] essentially perform Monte Carlo 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 [Jarosz et al. 2008] and then for light rays [Jarosz et al. 2011a; Sun et al. 2010] in volumetric photon mapping, and later adapted to many-light methods [Novák et al. 2012] and path tracing approaches [Georgiev et al. 2013; Křivánek et al. 2014]. Line samples have also cropped up for computing hemispherical visibility and motion blur [Gribel et al. 2010; Gribel et al. 2011], depth of field [Tzeng et al. 2012], visibility in hair [Barringer et al. 2012], and masked environment lighting [Nowrouzezahrai et al. 2014]. 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 Monte Carlo integration.

3 Preliminaries

We are interested in computing an integral of the form:

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

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

Monte Carlo integration using point samples. Monte Carlo integration forms an approximation, I_N, of I by evaluating the integrand f at N sample locations $\mathbf{s}_j$ uniformly distributed over the domain $\mathbb{D}$. This sampling process can be expressed in continuous form by multiplying the original integrand f with a normalized sampling function S consisting of delta responses:

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

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

$$I_N = \int_{\Phi} \mathcal{F}_S(\nu)\overline{\mathcal{F}_f(\nu)}\,\mathrm{d}\nu, \text{ with } \mathcal{F}_S(\nu) = \frac{1}{N}\sum_{j=1}^{N} e^{-2\pi i(\nu\cdot\mathbf{s}_j)}, \tag{3}$$

where $\overline{\mathcal{F}_f}$ is the complex conjugate of the integrand's spectrum, and $\mathcal{F}_S$ is the Fourier spectrum of the normalized sampling function where each summand is the Fourier transform of a single point.

Variance formulation. Prior work [Durand 2011; Pilleboue et al. 2015] has shown that the variance of I_N depends on the power spectrum, $\mathcal{P}(\nu) = \|\mathcal{F}(\nu)\|^2$, of the integrand and the expected

power spectrum of the *homogenized*[1] sampling pattern:

$$Var\left(I_N\right) = \int_{\Theta} \langle \mathcal{P}_S(\nu) \rangle \, \mathcal{P}_f(\nu) \, \mathrm{d}\nu \,, \tag{4}$$

where Θ includes all frequencies except DC, and $\langle \mathcal{P}_S(\nu) \rangle$ is the expected power spectrum of the normalized[2] sampling function.

Prior work [Leneman 1966; Dippé and Wold 1985; Dippé and Wold 1992] has derived analytic expected power spectra for common point sampling patterns, for instance [Gabrielli and Torquato 2004]:

$$\langle \mathcal{P}_S(\nu) \rangle = \begin{cases} \frac{1}{N} & \text{for random, and} \\ \frac{1}{N}\left(1 - \prod_i^d \mathrm{Sinc}(\pi\nu_i)^2\right) & \text{for jittered} \end{cases} \tag{5}$$

where ν_i is the i-th dimension of frequency vector ν.

To analyze the variance and convergence rate of specific sampling patterns, Pilleboue et al. [2015] further simplify Eq. (4) by going to polar coordinates and collapsing the integrand's power spectrum $\mathcal{P}_f$ and the expected sampling power spectrum $\langle \mathcal{P}_S(\cdot) \rangle$, under the assumption of isotropic sampling power spectra, into their radial averages $\breve{\mathcal{P}}(\cdot)$, arriving at:

$$Var\left(I_N\right) = \int_0^{\infty} \rho^{d-1} \, \breve{\mathcal{P}}_S(\rho) \breve{\mathcal{P}}_f(\rho) \, \mathrm{d}\rho. \tag{6}$$

With this simplification, their primary contribution was showing that if the radially averaged sampling power spectra can be expressed analytically, then the corresponding variance convergence rates can be derived for a given class of functions. To more easily apply this idea to complex radial power spectra, they showed that it is often sufficient to piecewise bound the radial mean power spectrum using a monomial in the low-frequency region and a constant for high frequencies, with the degree of the low-frequency monomial bound ultimately determining the convergence rate *(for more details please refer to Appendix A)*. Unfortunately, by relying on the radially averaged power spectra, Pilleboue et al.'s analysis only truly applies to isotropic point sampling spectra.

4 Generalized Variance Formulation

In order to obtain a variance formulation that works for anisotropic sampling spectra, we avoid relying on the *radially averaged* power spectra (6), and instead seek a formulation that allows analyzing the *radial* behavior along any direction. We start by rewriting Eq. (4) in polar coordinates:

$$Var\left(I_N\right) = \int_0^{\infty} \rho^{d-1} \int_{\mathcal{S}^{d-1}} \langle \mathcal{P}_S(\rho\mathbf{n}) \rangle \, \mathcal{P}_f(\rho\mathbf{n}) \, \mathrm{d}\mathbf{n} \, \mathrm{d}\rho \tag{7}$$

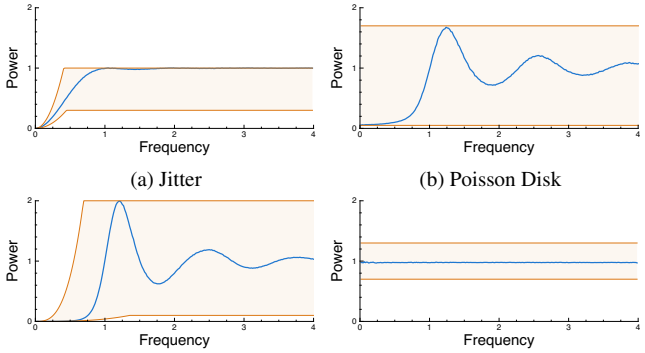
where ρ represents the radial component and $\mathbf{n}$ is a unit-length vector residing on the $(d-1)$-dimensional sphere $\mathcal{S}^{d-1}$ representing the angular component of the frequency vector $\omega = \rho\mathbf{n}$.

Generalization: We start by swapping the order of integration:

$$Var\left(I_N\right) = \int_{\mathcal{S}^{d-1}} \int_0^{\infty} \rho^{d-1} \, \langle \mathcal{P}_S(\rho\mathbf{n}) \rangle \, \mathcal{P}_f(\rho\mathbf{n}) \mathrm{d}\rho \, \mathrm{d}\mathbf{n}. \tag{8}$$

[1] *Homogenization* of sampling patterns ensures that the sampling Fourier coefficients are uniformly distributed in the complex plane resulting in an unbiased estimator [Subr and Kautz 2013]. In the point processes literature, 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.

[2] Pilleboue et al. [2015] use an unnormalized sampling function, so their expression for variance includes an additional $1/N$ normalization factor. We instead fold this factor directly into the normalized sampling function.



(a) Jitter (b) Poisson Disk (c) CCVT [Balzer et al. 2009] (d) Random

Figure 3: Radial *power spectra for common samplers can be approximated (bounded) by simple monomial profiles. Here, monomials of degree (a) $b = 2$ (quadratic), (b) $b = 0$ (constant), (c) $b = 3$ (cubic), and (d) $b = 0$ (constant) bound the radial power spectra.*

This equation can be seen as an integral over a hypersphere with $f(\cdot)$ Lebesgue integrable, which makes the integral bounded. If we restrict our formulation to only well-behaved $\mathcal{P}_f$ and $\mathcal{P}_S$—having finite bounded discontinuities [Apostol 1974]—the above variance integral (Eq. (8)) can be written as a limit of sums which corresponds to subdividing the integral hypersphere into similar-sized (not necessarily equal-sized) cones of any base-shape. These cones can be used to approximate the volume of this hypersphere, that allows rewriting Eq. (8) in the following form:

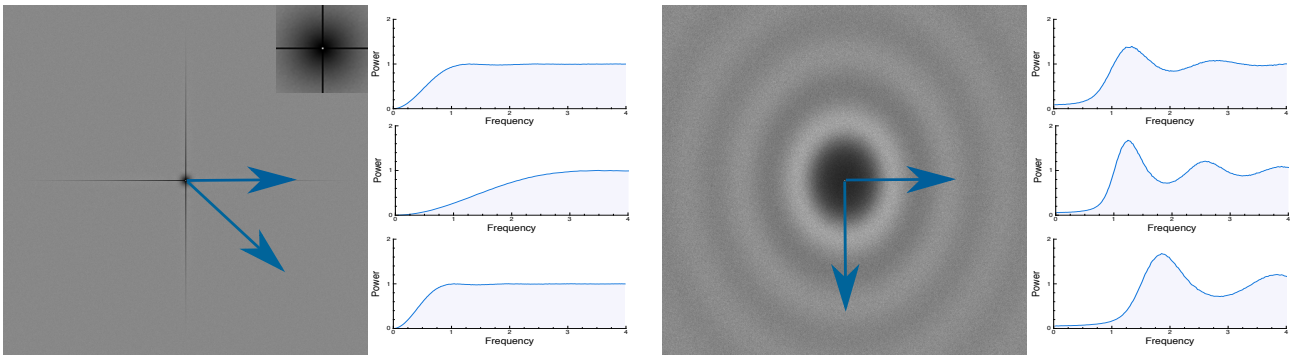
$$\boxed{Var\left(I_N\right) = \lim_{m\to\infty} \sum_{k=1}^{m} \int_0^{\infty} \rho^{d-1} \, \langle \mathcal{P}_S(\rho\mathbf{n}_k) \rangle \, \mathcal{P}_f(\rho\mathbf{n}_k) \mathrm{d}\rho \Delta\mathbf{n}_k} \tag{9}$$

This is a valid representation in the context of true infinitesimal calculus [Keisler 2012]. In the above formulation, $\Delta\mathbf{n}_k$ is a differential volume of the k-th cone. In the limit, we assume no angular variation of power in this k-th differential cone which allows us to consider a single direction corresponding to each cone. Here, $\Delta\mathbf{n}_k$ is a constant that approaches to zero as m tends to infinity.

From Eq. (9), variance of Monte Carlo integration can be obtained by summing the radially integrated terms along each individual direction k. This implies that, irrespective of whether our expected sampling power spectra $\langle \mathcal{P}_S(\cdot) \rangle$ is isotropic or not, we can analyze each k-th direction independently to know the overall behavior of the underlying sampler. Doing this in practice is difficult, however, since i) there are infinitely many possible directions, and ii) computing variance along any direction would require solving the integral analytically which is only possible if we know the analytical form of our sampling and integrand power spectra.

Since we are only interested in asymptotic convergence rates, it is sufficient for us to analytically bound the radial sampling power spectrum for any direction $\mathbf{n}_k$. Like prior work [Pilleboue et al. 2015], we bound the radial behaviour, along each direction $\mathbf{n}_k$, with a *monomial* of degree $b_k \geq 0$ in the low frequency region $(0, \nu]$, up to a certain radial frequency ν (as shown in Fig. 3), and a flat spectrum beyond. Instead of doing this for the *radially averaged power spectrum*, however, we consider this independently for each angular direction $\mathbf{n}_k$. Assuming a highly anisotropic sampling power spectrum ($\langle \mathcal{P}_S(\rho\mathbf{n}_k) \rangle$), we may require a monomial of different degree b_k along each direction $\mathbf{n}_k$ to bound the radial behavior.

In Sec. 6, we will ultimately express the radial behavior, for each summand of Eq. (9), within the two regions using monomials that depend on the particular sampling pattern, and prove that the overall variance convergence rate of the underlying sampling pattern can be deduced by studying the power spectrum along a single *partic-*



(a) Muljitter power spectrum

(b) Anisotropic Poisson disk power spectrum

Figure 2: *Illustration of a 2D point sampling power spectra for (a) multijitter, with the zoomed-in inset for the low frequency region is at the top right and (b) the anisotropic Poisson disk sampling patterns. Compared to Fig. 1(b), in 2D the radially averaged Poisson disk power spectrum still reveals the underlying Poisson disk behaviour since the anisotropy is introduced only due to skewing of points.*

ular direction. We will also consider analytic best and worst case integrands.

5 Line Samples

We now show how the generalized variance formulation we presented in Sec. 4 applies not only to anisotropic point sampling, but also to the inherently anisotropic process induced by line sampling. To accomplish this we first provide some mathematical background that will allow us to express the expected power spectrum of Monte Carlo line sampling.

Monte Carlo Estimator 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.

Monte Carlo integration using uniformly distributed line samples can be expressed as in Eq. (2), but using a line sampling function:

$$I_N = \int_{\mathbb{D}} L(\mathbf{x}) f(\mathbf{x})\, \mathrm{d}\mathbf{x}, \text{ where } L(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^{N} \delta(\mathrm{dist}(\mathbf{x}, \mathbf{l}_j)), \quad (10)$$

and $\mathrm{dist}(\mathbf{x}, \mathbf{l}_j) = \|(\mathbf{l}_j^{\perp} - \mathbf{x}) + (\mathbf{x} \cdot \mathbf{l}_j^{\parallel})\mathbf{l}_j^{\parallel}\|$ 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 as Eq. (3) where the point sampling spectrum $\mathcal{F}_S$ is replaced by the line sampling spectrum $\mathcal{F}_L$:

$$I_N = \int_{\Phi} \overline{\mathcal{F}_f(\nu)} \mathcal{F}_L(\nu)\, \mathrm{d}\nu, \quad (11)$$

The Fourier transform of a line sample can be easily derived [Santaló 1953; Gel 1966] for arbitrary dimensions. Sun and colleagues [2013] presented the formulation in 2D for the purposes of blue-noise sampling. We build on these formulations to express the frequency and power spectra of Monte Carlo line sampling in d dimensions as:

$$\mathcal{F}_L(\nu) = \frac{1}{N} \sum_{j=1}^{N} \delta(\nu \cdot \mathbf{l}_j^{\parallel}) e^{-2\pi i (\nu \cdot \mathbf{l}_j^{\perp})}, \quad (12)$$

where the power spectrum is simply $\mathcal{P}_L(\nu) = \|\mathcal{F}_L(\nu)\|^2$. Note that, for a given direction, each line sample's frequency spectrum (each summand above) is that of a $(d-1)$-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 $(d-1)$-dimensional points restricted to lie in a hyper-plane perpendicular to the lines. Fig. 1(b) illustrates this for $d = 3$ where parallel line samples are generated horizontally such that the power spectrum lies in a subspace orthogonal to the direction of line samples.

Dual interpretations. Inserting Eq. (12) into Eq. (11) and slightly rearranging terms provides further insights into what is happening:

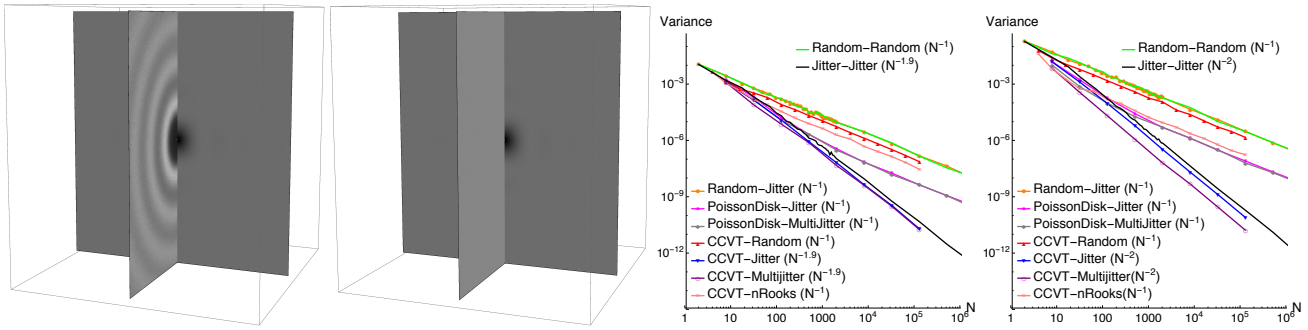
$$I_N = \int_{\Phi} \overbrace{\underbrace{\frac{1}{N} \sum_{j=1}^{N} e^{-2\pi i (\nu \cdot \mathbf{l}_j^{\perp})}}_{\text{Point sampling spectrum in } d-1} \underbrace{\delta(\nu \cdot \mathbf{l}_j^{\parallel}) \overline{\mathcal{F}_f(\nu)}}_{\text{Prefiltered integrand } \tilde{f}}}^{\text{Line sampling spectrum in } d \quad \text{Original } d\text{-dimensional integrand}} \mathrm{d}\nu. \quad (13)$$

The grouping specified by the over-braces is the original interpretation of the d-dimensional integration where evaluating each line sample involves integrating the original d-dimensional integrand along the line. The grouping specified by the under-braces shows that we can also interpret this as first collapsing (integrating) the original integrand's Fourier spectrum $\mathcal{F}_f$ along the direction $\mathbf{l}_j^{\parallel}$ into a prefiltered $(d-1)$-dimensional integrand spectrum $\mathcal{F}_{\tilde{f}} = \delta \overline{\mathcal{F}_f}$, followed by point sampling the remaining $(d-1)$ dimensions in the hyperplane spanned by $\mathbf{l}_j^{\perp}$. 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 line samples can improve convergence rate: i) the sampling process is equivalent to point sampling one dimension lower, which can provide faster convergence due to denser stratification and ii) the integrand has been pre-integrated, which can potentially increase its smoothness and spectral decay, resulting in further convergence improvement.

Expected Power Spectra. From the expressions for point sampling (Equation 5), and the connection between line sampling in d and point sampling in $(d-1)$ dimensions discussed above, it follows that the expected power spectrum for N line samples distributed over the span of $\mathbf{l}^{\perp}$ is:

$$\langle \mathcal{P}_L(\nu) \rangle = \begin{cases} \frac{1}{N} \delta\left(\nu \cdot \mathbf{l}^{\perp}\right) & \text{for random,} \\ \frac{1}{N} \left[1 - \prod_i^{d-1} \mathrm{Sinc}(\pi \nu_i \mathbf{l}_i^{\perp})^2\right] & \text{for jittered} \end{cases} \quad (14)$$

line offset distributions. The product goes over the $d-1$ dimensions spanning the hyperplane of possible line offsets $\mathbf{l}^{\perp}$. We illustrate



(a) Poissondisk-Jitter (b) Random-Jitter (c) Variance anaysis Sphere (d) Variance analysis Gaussian 3D

Figure 4: ***Multi-directional line samples:*** *Power spectrum of 3D line samples generated (axis-aligned for clarity) with: (a) a Poisson disk line sample distribution along one direction and a jittered line sample distribution along the other, and (b) random–jitter. In (c) and (d) we plot the variance convergence rates for Sphere and Gaussian 3D integrands, respectively, using a variety of multi-directional line sampling pairs denoted* pattern1–pattern2 *for the sampling patterns used to generate lines along each direction.*

these analytic power spectra for 3D in Fig. 4(b) for multi-directional line sampling where one direction uses randomly generated line samples and the other direction uses jittered line samples.

Homogenized line samples Samplers like white noise (random) and ones derived from white noise (Poisson disk [Dippé and Wold 1985], CCVT [Balzer et al. 2009], BNOT [de Goes et al. 2012]) are homogeneous by construction [Pilleboue et al. 2015]. Since the frequency content of parallel line samples in d-D is equivalent to that of point samples in $(d-1)$-D, homogenizing line samples requires homogenizing the $(d-1)$ independent components of the line sample offset $\mathbf{l}^{\perp}$.

5.1 Variance formulation

To derive a generalized variance formulation analogous to Eq. (9) for the case of lines, we can simply start with either interpretation of Eq. (13) and insert it into Eq. (4) in place of Eq. (3) to obtain an equivalent version of Eq. (4) for line sampling. We can then express this in polar coordinates (7), swap the order of integration (8), and convert the spherical integral into a discrete sum of directions. At the end of these steps, we obtain two equivalent versions of Eq. (9) corresponding to the two interpretations of line sampling discussed in Eq. (13):

$$Var\left(I_N\right) = \lim_{m\to\infty} \sum_{k=1}^{m} \int_0^{\infty} \rho^{d-1} \left\langle \mathcal{P}_L(\rho\mathbf{n}_k) \right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\, \mathrm{d}\rho\Delta\mathbf{n}, \quad (15a)$$

$$Var\left(I_N\right) = \lim_{m\to\infty} \sum_{k=1}^{m} \int_0^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k) \right\rangle \mathcal{P}_{\tilde{f}}(\rho\mathbf{n}_k)\, \mathrm{d}\rho\Delta\mathbf{n}. \quad (15b)$$

Eq. (15a) is directly analogous to Eq. (9) but using the expected sampling power spectrum for d-dimensional lines $\langle\mathcal{P}_L\rangle$ instead of d-dimensional points $\langle\mathcal{P}_S\rangle$. Eq. (15b), however, arises by interpreting line samples in d dimensions as point samples in $(d-1)$ dimensions, but with an effective *prefiltered* integrand $\tilde{f}$. In both cases, since the product of the power spectra will lie only in the hyperplane perpendicular to the lines, we only need to consider the radial behavior of the variance along the frequency directions $\mathbf{n}_k \in \mathcal{S}^{d-2}$ on a one-dimension-lower hypersphere.

6 Theoretical convegence analysis

To derive theoretical convergence rates we first restrict our integrands to integrable functions of the form $f(x)\chi_\Omega(x)$ with $f(x)$ is defined in Ω, a bounded domain, with smooth boundary (where $\chi_\Omega(x)$ is a characteristic function of Ω) [Brandolini et al. 2001]. This can, however, be extended to arbitrary bounded convex regions [Brandolini et al. 2003]. The worst case from this class of functions exhibits the power fall- off of the order $\mathcal{O}\left(\rho^{-(d+1)}\right)$ where $\rho > 0$ is a radial frequency.

Best and Worse case: To obtain best and worse case convergence rate for various samplers we first fix our integrand power spectra $\mathcal{P}_f(\rho\mathbf{n}_k)$, along each direction $\mathbf{n}_k$, to have the form:

$$\mathcal{P}_f^W(\rho\mathbf{n}_k) = \begin{cases} c_f & \rho < \rho_0, \\ c_f \rho^{-d-1} & \text{otherwise} \end{cases} \text{ and } \mathcal{P}_f^B(\rho\mathbf{n}_k) = \begin{cases} c_f & \rho < \rho_0, \\ 0 & \text{otherwise.} \end{cases}, \quad (16)$$

where, $c_f > 0$ is a constant and $\rho_0 \in \mathbb{R}^+/0$ is finite. Note that, in the rest of the paper we may drop the superscripts for brevity and refer implicitly to the case (best or worst) we are studying. This will allow us to analytically derive convergence rates for any sampling pattern with anisotropic power spectra using our variance formulation (Eq. (17)), which requires solving the integral analytically along each k-th direction.

6.1 Convergence from a single direction

In this section, we prove that convergence rate of an anisotropic power specrtum is dependent on a single *particular* direction. We first split the radial integral in Eq. (9) into separate regions $(0,\nu)$ and (ν,∞):

$$Var\left(I_N\right) = \lim_{m\to\infty} \sum_{k=1}^{m} \left[\int_0^{\nu} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k) \right\rangle \mathcal{P}_f(\rho\mathbf{n}_k) \mathrm{d}\rho + \int_{\nu}^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k) \right\rangle \mathcal{P}_f(\rho\mathbf{n}_k) \mathrm{d}\rho \right] \Delta\mathbf{n}_k. \quad (17)$$

Note that, this decomposition will potentially be different for every direction. Since the sampling power spectra scales with the number of samples N, anisotropy along a direction would potentially impact the *effective* number of samples (N_k) along that direction. As a result, ν would depend on the the *effective* N_k number of samples that could potentially impact that particular direction.

This implies that, the radial behaviour along each direction, that can be approximated by a *monomial* of degree b_k, depends directly on

N_k:

$$\langle \mathcal{P}_S(\rho \mathbf{n}_k) \rangle = \begin{cases} \frac{\gamma_k}{N} \left(\frac{\rho}{\alpha_k N_k} \right)^{b_k} & \rho < \alpha_k N_k \\ \frac{\gamma_k}{N} & \text{otherwise} \end{cases}, \quad (18)$$

where $\nu = \alpha_k N_k$. Note that, contrary to Pilleboue et al. [2015] this radial profile formulation is applicable to both the point and line samples with arbitrary power spectra. Here, γ_k is a positive constant, N_k represents *effective* number of samples along the k-th direction, N is the total number of samples, d is the dimension and $\alpha_k \in \mathbb{R}^+/0$ is used to quantify the range of energy-free frequency with respect to the mean frequency along each direction.

Since variance convergence rate directly depends on the *degree of the monomial* [Pilleboue et al. 2015] used to bound the radial curve, we can derive convergence rates along each direction $\mathbf{n}_k$ separately in terms of the order $\mathcal{O}(\cdot)$ of N_k samples. Asymptotically, the summand from Eq. (17) showing the worst convergence rate among all would dominate the overall convergence rate of the underlying anisotropic samples. This implies that, ultimately, the convergence rate would be decided by the direction $\mathbf{n}_k$ that would exhibit the worst convergence behaviour. A mathemtically more rigorous proof of this statement is given in Appendix B. *This proves our claim that the variance convergence rate of an underlying sampling pattern in Monte Carlo integration can be deduced from one particular direction*. □

For isotropic samplers: To obtain convergence rates for stochastic samplers with *isotropic* power spectra, we can simply use $N_k = \sqrt[d]{N}$ in Eq. (18) for point samples—which would make the previously proposed convergence tool [Pilleboue et al. 2015] a special case to our *generalised version*—and $N_k = \sqrt[d-1]{N}$ for line samples (more details are given in Appendix A).

In the rest of this section, we will use the monomial *profiles* proposed by Pilleboue and colleagues. For example, for jittered pattern $b_k = 2$ (*quadratic profile*) and for random and Poisson disk samples $b_k = 0$ *(constant profile) (Fig. 3)*.

6.2 Anisotropic point sampling power spectra

As shown in Sec. 4, to study the variance convergence behavior of point sampling patterns with anisotropic power spectra it is sufficient to analyze the radial behavior along the direction that could potentially give the worst case convergence. In practice, the anisotropic sampling patterns we encounter show anisotropy mainly along a finite set of directions (Fig. 2). Therefore, we only need to consider the variance formulation (Eq. (17)) along these few directions based on the variation we observe in the power spectra. Note, however, that to obtain *variance* we would need to consider all the directions.

We can now directly use Eq. (16) and (18) in Eq. (17) to analytically derive best and worst case variance convergence rates for any sampler.

Latin hypercube sampling (LHS): In d-dimensions, to generated N Latin hypercube point samples, we first generate N $1D$ jittered point samples along each dimension and then randomly permute them to make a d-dimensional tuple. This implies that, we have $N_k = N$ *effective* jittered samples along each of the canonical axes in d-dimensions. In Fig. 1(a), we are showing expected Fourier power spectrum of a 2D instance of LHS sampling, which is known as *n-rooks* [Shirley 1991]. This figure reveals that along the canonical axes, we have radial behaviour of jittered sampling pattern, which is what we expect. By substituting Eq. (16) and (18) in Eq. (17), we obtain the convergence rate as follows:

Worst case along the canonical axes: Since $N_k = N$ along the canonical axes, we get:

$$\begin{aligned} \mathit{Var}\,(I_N) &< \int_0^{\alpha N} \rho^{d-1} \frac{\gamma}{N} \left(\frac{\rho}{\alpha N} \right)^2 c_f \rho^{-d-1} \,\mathrm{d}\rho + \\ &\quad \int_{\alpha N}^{\infty} \rho^{d-1} \frac{\gamma}{N} c_f \rho^{-d-1} \,\mathrm{d}\rho\,, \quad (19) \\ &< \frac{c_f \gamma}{\alpha^2 N^3} \int_0^{\alpha N} \mathrm{d}\rho + \frac{c_f \gamma}{N} \int_{\alpha N}^{\infty} \rho^{-2} \mathrm{d}\rho\,, \quad (20) \\ &< \mathcal{O}\big(N^{-2}\big) \quad (21) \end{aligned}$$

Best case convergence along the canonical axes:

$$\begin{aligned} \mathit{Var}\,(I_N) &< \int_0^{\rho_0} \rho^{d-1} \frac{\gamma}{N} \left(\frac{\rho}{\alpha N} \right)^2 c_f \,\mathrm{d}\rho\,, \quad (22) \\ &< \mathcal{O}\big(N^{-3}\big) \quad (23) \end{aligned}$$

However, due to the random permutation the rest of the power spectrum exhibits flat radial behaviour. This implies that across all directions, other than the canonical axes, the convergence rate would be $\mathcal{O}\big(N^{-1}\big)$. Therefore, from the variance Eq. (17), we can summarise the overall best and worst case convergence rates in the two regions (each of the $d-1$ hyper-planes + the rest of the region) as follows:

$$\mathit{Var}\,(I_N) < \begin{cases} \mathcal{O}\big(N^{-2}\big) + \mathcal{O}\big(N^{-1}\big) & \text{worst-case} \\ \mathcal{O}\big(N^{-3}\big) + \mathcal{O}\big(N^{-1}\big) & \text{best-case}\ . \end{cases} \quad (24)$$

Since the addition of two $\mathcal{O}(\cdot)$ functions would asymptotically be dominated by the one with worse convergence, the overall convergence rate for most of the integrands sampled and integrated using Latin hypercube sampling would not be better than $\mathcal{O}\big(N^{-1}\big)$. Integrands that vary only along one of the dimensions, however, would benefit from the stratification and converge much faster. For integrands that vary along all dimensions, the constants in front of each of the summands in the convergence rate would be dictated by the relative amount of variation of the integrand in each dimension. This means that for finite sample counts, the Monte Carlo estimator may initially follow the slope of the faster convergence rate, but with increasing samples the slower convergence rate will dominate any constant factors.

Multi-jitter: For multi-jitter sampling pattern, the power spectrum (Fig. 2(a)) exhibits radial behaviour of jittered samples in all the directions in the power spectrum. The only difference is that, since the stratification along the canonical axes is exactly equal to the number of samples N, the *effective* number of samples along each dimension would be $N_k = N$ in d dimensions, whereas, across all other directions it would be $N_k = \sqrt[d]{N}$. This explains why the jittered radial profile is scaled along the canonical axes. Note that, similar to LHS, for multi-jitter sampling, the convergence rate along the canonical axes would be the same as we derived for LHS. For other directions in the power spectrum, the radial behaviour is that of jitter, with $\sqrt[d]{N}$ *effective* samples, for which the worst and best case convergence rates has been already derived [Pilleboue et al. 2015]. Here we summarize the overall variance convergence rates:

$$\mathit{Var}\,(I_N) < \begin{cases} \mathcal{O}\big(N^{-2}\big) + \mathcal{O}\Big(N^{-1-\frac{1}{d}}\Big) & \text{worst-case} \\ \mathcal{O}\big(N^{-3}\big) + \mathcal{O}\Big(N^{-1-\frac{2}{d}}\Big) & \text{best-case} \end{cases} \quad (25)$$

As before, the $\mathcal{O}(\cdot)$ function with the worse asymptotic behavior among the two summands in the best and the worst cases would dominate the overall convergence rate. This shows that *multi-jitter*

samples would have convergence similar to that of jittered sampling patterns. The only difference would be for the integrands that would only vary along one of the canonical axes for which the convergence rate would be way faster. As before, integrands that varying to differing degrees in each dimension may exhibit convergence slopes that change for finite sample counts before settling at the worse of the two asymptotic convergence rates.

Uncorrelated jitter: Generating uncorrelated jitter samples in different subspaces to solve a d-dimensional integral is one of the most highly recommended sampling strategy in rendering [Pharr and Humphreys 2010]. The idea was first suggested by Cook [1986] where he proposed to randomly combine different samplers for higher-dimensional integration[3]. Latin hypercube sampling can be seen as a special case of uncorrelated jitter sampling where 1D jittered samples are generated for each single dimension, which are later randomly permuted to solve the d-dimensional integral. Depending on the dimensionality of the subspaces on which jittered samples are generated, the overall convergence can get affected. Here we discuss one of the variants of uncorrelated jitter, other variants can similarly follow the procedure to obtain the respective convergence rates.

We consider a d-dimensional uncorrelated jitter pattern for N samples, where we generate 1D, 2D and $(d-3)$D uncorrelated jittered subspaces, which are then randomly permuted to form a d-dimensional tuple. Similar to the LHS (Eq. (24)) and multi-jitter (Eq. (25)) case, following Eq. (18) we can easily obtain the variance convergence rate for this setting in the form:

$$Var\left(I_N\right) < \tag{26}$$
$$\begin{cases} \overbrace{\mathcal{O}\left(N^{-2}\right)}^{\text{1D}} \overbrace{+\mathcal{O}\left(N^{-1.5}\right)}^{\text{2D}} \overbrace{+\mathcal{O}\left(N^{-1-\frac{1}{d-3}}\right)}^{(d-3)\text{D}} + \mathcal{O}\left(N^{-1}\right) & \text{worst-case} \\ \mathcal{O}\left(N^{-3}\right) + \mathcal{O}\left(N^{-2}\right) + \mathcal{O}\left(N^{-1-\frac{2}{d-3}}\right) + \mathcal{O}\left(N^{-1}\right) & \text{best-case .} \end{cases}$$

In the above equation, we are show that if the integrand will be varying only within one of the subspaces, which contain uncorrelated jitter samples (in 1D, 2D, $(d-3)$D) subspaces), we would obtain the corresponding convergence rates as shown in Eq. (26). However, if the integrand is varying in-between these subspaces, we will fall back to the, much worse, Monte Carlo convergence rate of $\mathcal{O}\left(N^{-1}\right)$. This would happen because of the random permutation step that is used to combine these subspaces to make a d- dimensional tuple. Note that, there are many ways to perform uncorrelated jittering, and depending on the dimensionality of the subspaces containing jittered samples, the convergence expression from Eq. (26) would change.

Anisotropic blue noise samplers: There exist many blue noise samplers that could allow generation of anisotropic blue noise power spectra. We have shown one illustration of anisotropic Poisson disk sampling power spectrum in Fig. 2(b), where the shape of the low frequency region is that of axis-aligned *ellipse*. Note that, this kind of anisotropy only scale the radial behaviour by a constant, and therefore, would not affect the asymptotic convergence rate in any direction. This is different from what we have seen in the case of stratification methods (LHS, multi-jitter and uncorrelated-jitter) where the anisotropy introduces different convergence rates along different directions. This difference can be explained from the fact that, these stratification methods change the *effective* number of samples along some particular directions whereas this is not the case with anisotropic blue noise samplers that are generated by applying a Jacobian to match some characteristics of the underlying density function [Li et al. 2010; Wei and Wang 2011].

However, if the anisotropic blue noise samplers would somehow change the *effective* number of samples in a particular direction or region in the power spectrum, while keeping the underlying blue noise characteristics, the convergence rate along that set of directions could be highly improved. This would be the case with the recent projective blue noise sampling approach [Reinert et al. 2016] that ensures that all d-dimensional sample points also form denser blue-noise sampling spectra along lower-dimensional projections.

6.3 Line sampling power spectra

From the knowledge gained from Sec. 5 and Sec. 6, we can now easily derive the convergence rates for line samples. Using the fact that, the expected power spectra of line samples is confined within the $d-1$ subspace (as illustrated in Fig. 1(b)), we could directly apply the convergence rates we derive in Sec. 6.2 in the $d-1$ dimensions along the directions within the $\mathcal{S}^{d-2}$ angular domain. The same can also be easily derived by substituting Eq. (18) into the variance formulation Eq. (15b) for the best and the worst cases (Eq. (16)).

For an *isotropic* sampling power spectra in the $d-1$ subspace for lines, the overall convergence rates can be summarised as follows:

$$Var\left(I_N\right) < \begin{cases} \mathcal{O}\left(N^{-\frac{d}{d-1}}\right) & \text{worst-case} \\ \mathcal{O}\left(N^{-\frac{b+d-1}{d-1}}\right) & \text{best-case ,} \end{cases} \tag{27}$$

which is exactly what we would obtain, if we use $d-1$ dimensional point samples in the d-dimensional integration domain. For line samples with offsets following an *anisotropic* sampling power spectrum, all the results we derived in Sec. 6.2 for anisotropic point sampling spectra, e.g. LHS, would directly apply to the case of line samples, but for multi-jitter, with one dimension reduced (only for the dimension dependent terms). Note that, we can write out the best and worst case convergence rate by using $b=3$ for CCVT and $b=2$ for jitter and multi-jitter in Eq. (27).

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[4] Eq. (15b). In Fig. 4(a,b), we illustrate two examples where line sampling is performed in two separate directions using different sampling patterns along each direction. In the next section, we will look at some concerete multi-directional line sampling strategies and will analyse their variance convergence plots.

7 Experiments

We now perform a set of 2D and 3D integration experiments using different point- and line-sampling patterns to validate our theoretical results. To analyze the variance convergence rate of anisotropic point samples, we implement jitter, multi-jitter, n-rooks/Latin hypercube, and anisotropic Poisson disk samples. For line samples, we leverage all of these techniques as well as CCVT [Balzer et al. 2009] and Poisson disk (isotropic) samplers to specify the positions of the lines. We generate the line sample offsets using 1D point samples for 2D integration problems, and 2D point samples for 3D integration problems. We plan to release all our C++ analysis code publicly.

[3]For example, to solve a 4D integral, instead of generating a 4-dimensional jittered pattern, Cook suggested to first generate two *uncorrelated* 2D jittered patterns and then randomly combine both to solve the 4D integral.

[4]If we have two uncorrelated random variables (MC line estimators in different directions using uncorrelated sample locations), then their variance is additive, and therefore the convergence rate will be the worse of the two.

To perform our worst case analysis (16), we use a binary disk function in 2D and a binary sphere function in 3D, both of which have a worst case power decay rate [Pilleboue et al. 2015]. We use a multivariate Gaussian for our best case convergence analysis since it is smooth enough (C^∞) to obtain the theoretical best case convergence rates despite having a spectrum with infinite extent. We additionally validate these analytically verified convergence rates with realistic test scenes including an environment map, ambient occlusion, and homogeneous participating media.

7.1 Point Samples

We start by analyzing multi-jitter and n-rooks point samples for 2D integration. For multi-jitter, our theoretical convergence derivation (25) predicts best- and worst-case convergence rates of $\mathcal{O}\left(N^{-2}\right)$ and $\mathcal{O}\left(N^{-1.5}\right)$ respectively, and $\mathcal{O}\left(N^{-1}\right)$ for both best and worst case with n-rook samples (24). The variance plots for the Disk, Gaussian and the environment map integration in Fig. 5 agree with these predictions. For n-rooks sampling, we further analyze how the two best-case convergence rates derived in Eq. (24) would vary as the integrand varies along a particular direction. We choose an anisotropic Gaussian function (Fig. 5(d)) and scale its standard deviation σ_x along the X-axis to induce different constant factors in the predicted convergence rates (24). We can see that variance exhibits a slope corresponding to the better $\mathcal{O}\left(N^{-3}\right)$ convergence rate up to some finite number of samples, but ultimately switches to the asymptotic convergence rate of $(\mathcal{O}\left(N^{-1}\right))$ with more samples. The exact location where this transition happens is influenced by how much of the integrand's power falls along the canonical axes vs. the remainder of the frequency domain. Multi-jitter exhibits similar behavior, starting with a variance slope of $\mathcal{O}\left(N^{-3}\right)$ before transitioning to the asymptotic convergence of $\mathcal{O}\left(N^{-2}\right)$. *While the primary motivation for our derivations was on asymptotic convergence rate, they also provide (to our knowledge) the first principled explanation for the multi-slope convergence rates of certain sampler–integrand combinations.*

Ambient Occlusion. We further confirm convergence rates for multi-jitter and n-rooks point sampling patterns when used to render ambient occlusion Fig. 8(a,b). Here we plot mean squared error (MSE) vs. the number of secondary rays per pixel used to perform hemispherical visibility integration. We generate hemispherical samples using 2D multi-jitter and n-rooks point samples using an area preserving polar mapping and compute the reference image using $N = 32K$ jittered secondary rays per pixel.

7.2 Line Samples

In this section, we empirically compare variance convergence rates for line samples. Our theory predicts that analytically evaluating an integrand using line samples will not only reduce the dimensionality of the integrand but may also smooth out discontinuities. An illustration is shown in Fig. 7, where we show that for disk (2D) and sphere (3D) integrands, line samples would prefilter the true binary integrand into a smooth function, ultimately improving the convergence rate beyond dimensionality reduction. For integrands that are already sufficiently smooth, however, the improvement in convergence rate will be only due to dimensionality reduction. In Fig. 6, we validate these findings by plotting variance convergence rates for the disk and Gaussian in 2D and 3D for different sampling patterns.

Comparison to point samples. For the disk function (worst case), multi-jitter point samples show (Fig. 5(a)) a convergence rate of $\mathcal{O}\left(N^{-1.5}\right)$, while multi-jitter line samples exhibit (Fig. 6(a)) an improved convergence rate of $\mathcal{O}\left(N^{-2.85}\right)$. Note that, if there

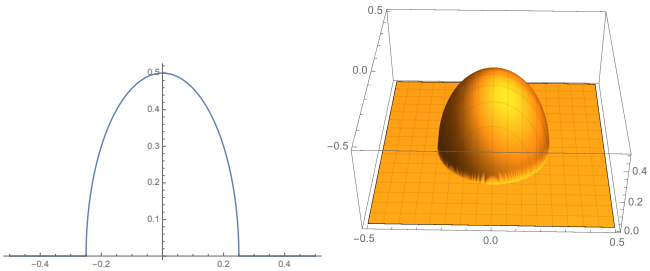


Figure 7: *While the disk and sphere act as binary functions for point samples, analytic line sample integration smooths these out to effective integrands (left: disk, right: sphere) which return the length of the lines through the original function.*

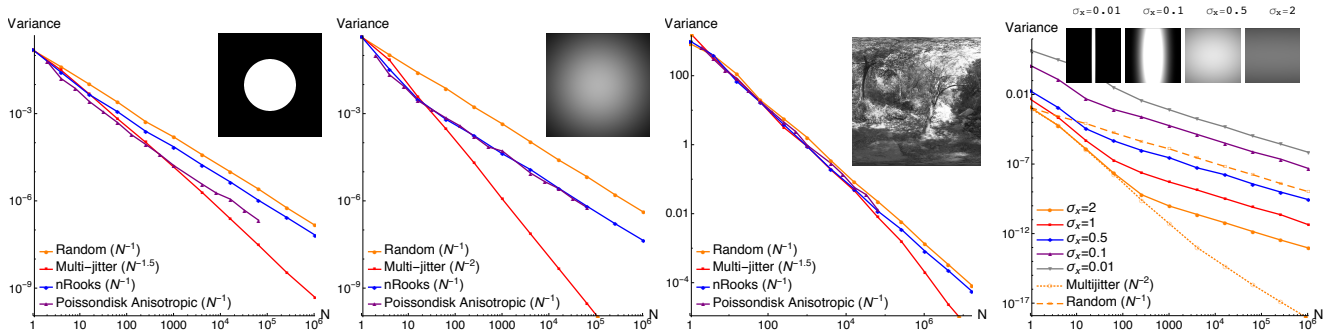
would have been only convergence improvement due to dimensionality reduction, we should obtain a convergence rate of $\mathcal{O}\left(N^{-2}\right)$ (as listed in Table 1). Since analytic integration using line samples smooths out the discontinuities, we actually obtain a convergence rate which is close to the best case convergence in 1D. For the Gaussian function in 2D, we obtain the convergence rate of $\mathcal{O}\left(N^{-3}\right)$ using line samples which is exactly what we should get for point samples in one dimension lower (in 1D). This also confirms that sufficiently smooth functions only get convergence benefits due to dimensionality reduction when using line samples.

In 3D, we generate line samples using 2D point samples, for which different point sampling stratification methods would have different characteristics. We compare three variants of jitter with different blue noise sampling patterns in Fig. 6(c,d). Similar to the 2D case Fig. 6(a,b), jitter and multi-jitter sampling patterns in 3D converge almost at the same rate for both the disk and the Gaussian function. N-rooks, however, give a convergence rate of $\mathcal{O}\left(N^{-1}\right)$ in 2D and 3D, matching our prediction (24). Poisson disk (isotropic and anisotropic) also converges with $\mathcal{O}\left(N^{-1}\right)$ as predicted, whereas CCVT sampling patterns converge faster than both jitter and Poisson disk. The best case (Gaussian 3D, Fig. 6(d)) convergence for CCVT matches exactly with Eq. (27) for $b = 3$ (Fig. 3).

Multi-directional line sampling. We also perform multi-directional line sampling convergence analysis, where we sample two orthogonal directions[5] using different sampling patterns. As shown in Fig. 4(c,d), among the two samplers used to generate line samples along each direction, the one with the worse convergence behavior ultimately decides the overall variance convergence rate. For example, in Fig. 6(c,d) we have shown that CCVT outperforms all samplers (along one or more directions) presented in this paper in terms of convergence. However, during multi-directional analysis, if one of the direction would be sampled by CCVT and the other by n-rooks, or random or Poisson disk, the overall convergence rate would fall back to naive Monte Carlo convergence $\mathcal{O}\left(N^{-1}\right)$.

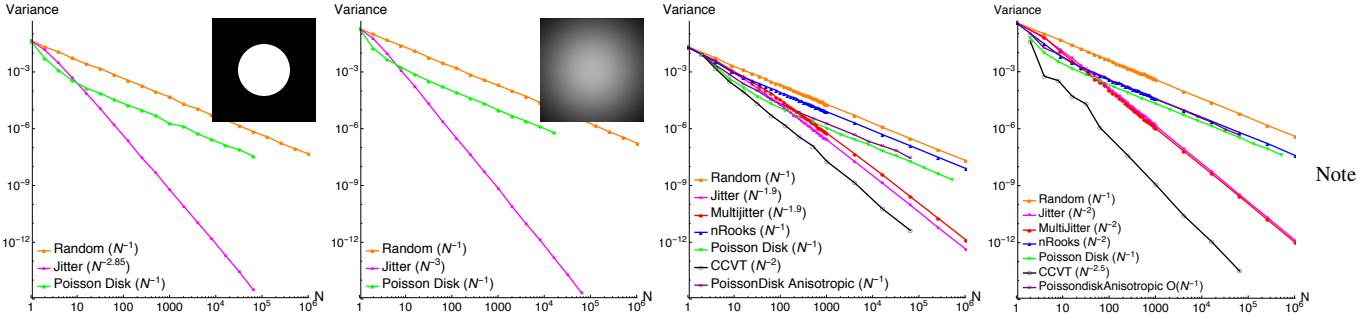
Participating media. We also verified our predicted convergence rates for a simple participating media rendering problem as shown in Fig. 9. In this setup, we have a point light source in a homogeneous medium with a glass sphere (refractive index = 1.5) placed at the center of the domain which cast volumetric caustics. We use two different kinds of sampling strategies to simulate this illumination problem, corresponding to volumetric photon mapping using points [Jarosz et al. 2008], and the long-photon beam variant [Jarosz et al. 2011a; Jarosz et al. 2011b] which solves transmittance along each direction analytically, forming line samples.

[5] Our analysis is equally valid for randomly oriented line samples.



(a) Disk (b) Gaussian 2D (c) Environment map (d) Anisotropic Gaussian with N-Rooks

Figure 5: ***Point samples 2D****: Illustrating variance convergence rate using 2D point samples (with anisotropic power spectra) for the worst (Disk) and the simple (Gaussian) case integrands. Random sampling curve is there for reference. (d) We also analyse the change in convergence by scaling an Gaussian along one direction using σ_x. Random sampling convergence plot is shown in dashes as reference for $\sigma_x = 2$, with multi-jitter in dotted orange format.*



Disk Gaussian 2D Sphere Gaussian 3D

Note that, since 2D line samples are generated using 1D point samples all stratifications (jitter, multi-jitter and n-rooks) would behave the same.

Figure 6: ***Line samples 2D & 3D:*** *Empirical variance convergence rates for the worst (Disk, sphere) and the simple (Gaussian 2D & 3D) case functions integrated using various line sampling patterns. For 2D, we only list jittered line samples since the 1D line positions generated by jitter, multi-jitter and n-rooks produce identical lines and behavior.*

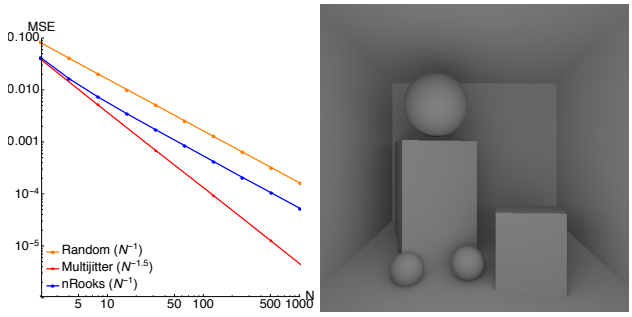
As we go higher in dimension, the variance convergence rate deteoriates. However, if we compare the 2D spherecaustics scene rendered with photons (Fig. 9(a)), to the 3D spherecaustic scene rendered with beams (Fig. 9(d)), we see that the convergence rates are the same. Initially, this behavior for the scene rendered with beams in 3D seems surprising since we may expect to see improvement in convergence due to smoothing of the underlying integrand, as discussed in Sec. 5. The underlying sampling domain in both of these cases (Fig. 9(a,d)), however, is actually 2D (denoted on top of these images). The integral in Fig. 9(d) is not a smoothed out version of light transport, but a discontinuous version over the 2D angular domain, which would ultimately give the 2D worst case convergence for all the samplers. This explains, why we get the same *worst case* convergence rates in both Fig. 9(a) and 9(d) despite using long beams in the latter.

8 Discussion

Just like line samples, line segment samples also have a highly anisotropic Fourier power spectrum [Sun et al. 2013]. In this work, we have focused our convergence analysis to line samples mainly

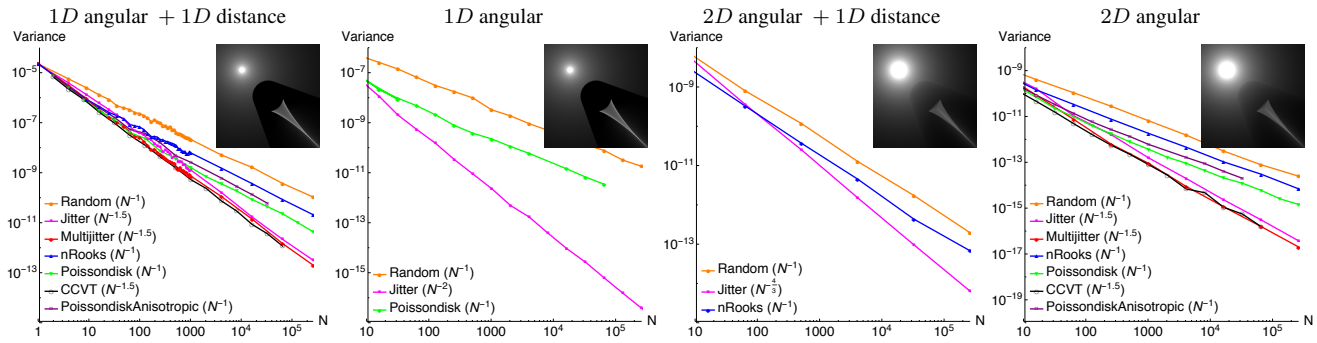
Table 1: *Variance convergence for N, d-dimensional jittered point samples [Pilleboue et al. 2015].*

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



(a) Convergence plot for (b) (b) Ambient Occlusion

Figure 8: *(a) Convergence analysis for a cornell box scene with Ambient occlusion (b) using multi-jitter and n-rooks point samples (Random samples are there for reference).*



(a) Spherecaustic 2D (Photons) (b) Spherecaustics 2D (long beams) (c) Spherecaustics 3D (photons) (d) Spherecaustics 3D (long beams)

Figure 9: ***Points & Lines:*** *Illustrating variance convergence rate for a homogeneous participating media rendered in 2D 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.*

because all the rendering applications, that we are aware of, uses line samples and not segments, to perform Monte Carlo integration. However, our analysis can be easily extended to line segment samples. To illustrate that, we first write down the Fourier transform of line segment samples that can be easily derived in the form:

$$\mathcal{F}_L(\nu) = \frac{1}{N}\sum_{j=1}^{N} e^{-2\pi i(\nu\cdot\mathbf{l}_j^{\perp})}\mathrm{lSinc}(\nu\cdot\mathbf{l}_j^{\parallel})\,, \tag{28}$$

Now, we provide a dual interpretation of line segment samples, similar to line samples in Eq. (13):

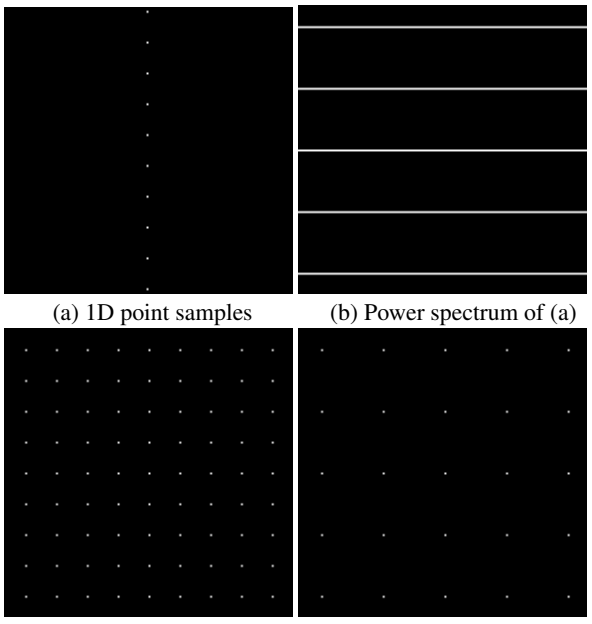
$$I_N = \int_{\Phi} \overbrace{\underbrace{\frac{1}{N}\sum_{j=1}^{N} e^{-2\pi i(\nu\cdot\mathbf{l}_j^{\perp})}}_{\text{Point sampling spectrum in } d}\mathrm{lSinc}(\nu\cdot\mathbf{l}_j^{\parallel})}^{\text{Segment sampling spectrum in } d}\overbrace{\overline{\mathcal{F}_f(\nu)}}^{\text{Original } d\text{-dimensional integrand}}\,\underbrace{\mathrm{d}\nu}_{\text{Prefiltered integrand } \tilde{f}}. \tag{29}$$

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 segment. The grouping specified by the under-braces shows that we can also interpret this as first integrating the original integrand's Fourier spectrum $\mathcal{F}_f$ along the direction $\mathbf{l}_j^{\parallel}$, within finite extent of the segment, into a prefiltered d-dimensional integrand spectrum, followed by point sampling the d dimensions. These two equivalent interpretations are very similar to what we observed in the case of line samples Eq. (13). However, the major difference here is that there is no dimensionality reduction on integration along the line segments. As a result, the convergence improvement that we would obtain using line segment samples would be solely due to smoothing of dicontinuities.

9 Appendix

A Monomial approximation of radial power spectra

We start from the very basic and look at the *regular* sampling power spectrum more closely. As shown in Fig. 10, for N *regular* point samples in $1D$, the corresponding power spectrum (Fig. 10(b)) contains replicas of the *DC line*, which is passing through the center of Fig. 10(b)). These replicas (aliases) are N distance (frequency) away from the DC peak, along the vertical axis. Now, if we go one dimension higher, in $2D$, the power spectrum would contain only high energy peaks, with the horizontal or vertical distance between the DC (component) peak and the first peak would be exactly $\sqrt{N}$. As we go higher in the dimension, the first peak would be at a distance $\sqrt[d]{N}$ distance away from the DC peak along the canonical



(a) 1D point samples (b) Power spectrum of (a)

(c) 2D point samples (d) Power spectrum of (c)

Figure 10: *Illustrating how far away is the first peak from the DC component, which in (d) is located at the center of the power spectrum and in (b) represents the horizontal line passing through the center of the power spectrum.*

axes, in d dimensions. This explains the fact why as we keep on increasing the number of point samples, the regular sampling power spectrum would contain less and less peaks near the low frequency region around the DC peak.

This notion can be generalised to jitter and other stochastic (blue noise) samplers, where the first peak in the jitter and/or blue noise samplers can be scaled at a distance $\alpha\sqrt[d]{N}$ from the DC (component) peak, where $\alpha > 0$. Note that, the α variable can be used to align all different radial sampling power spectra such that in the low frequency region they all can be categorised, in an unbiased fashion, according to their radial behaviour.

Now, for our analysis, we normalise the frequency components (ν_i) of the d-dimensional power spectrum with $\sqrt[d]{N}$, which allows us to write the $\mathcal{P}_S(\nu)$ in the normalised frequency form $\mathcal{P}_S\left(\frac{\nu}{\alpha\sqrt[d]{N}}\right)$. As a result, we can represent the low frequency region, of the Power spectrum for different samplers in a unit square, illustrated in Fig. 11, irrespective of the number of point samples N used to compute the

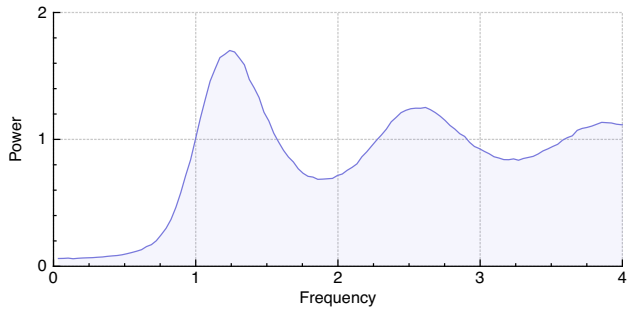


Figure 11: *Radial curve is scaled using α parameter so that the low frequency region could be well confined within the unit radial-power axes. Note that, once the α value is adjusted, for any number of samples, the corresponding radial power spectrum would not scale with N, given that we normalize the radial frequency variable ρ by $\alpha \sqrt[d]{N}$.*

power spectrum. In general, higher the number of samples used to compute the samping power spectrum, the more accurate this representation would be, for the power spectra.

A.1 Point Samples

To obtain the generalised polynomial *profile* for point samples, we first consider the case of jittered point samples. The *unnormalised* power spectrum for jittered sampling power spectra can be written as:

$$\langle \mathcal{P}_S(\omega) \rangle = \begin{cases} N^2 & \text{DC component} \\ N\left(1 - \prod_i^d \text{Sinc}(\pi\nu_i)^2\right) & \text{otherwise} \end{cases}, \tag{30}$$

whereas, the fully normalised power spectrum can be written as:

$$\langle \mathcal{P}_S(\omega) \rangle = \begin{cases} 1 & \text{DC component} \\ \frac{1}{N}\left(1 - \prod_i^d \text{Sinc}(\pi\nu_i)^2\right) & \text{otherwise} \end{cases}. \tag{31}$$

Note that, the fully normalised sampling power spectra would have it's DC component value equal to 1, as shown in Eq. (31). Since the variance does not depend on the DC component, we are only interested in the following term of the jittered power spectrum (Eq. (31)):

$$\langle \mathcal{P}_S(\omega) \rangle = \frac{1}{N}\left(1 - \prod_i^d \text{Sinc}(\pi\nu_i)^2\right) \tag{32}$$

where ν_i is the i-th dimension of frequency vector ν. Pilleboue and colleagues [2015] proposed a theoretical tool that allows to write the $\left(1 - \prod_i^d \text{Sinc}(\pi\nu_i)^2\right)$ term from Eq. (32), in the following radially averaged form:

$$= \begin{cases} \frac{\gamma\rho^b}{\alpha^b \sqrt[d]{N^b}} & \nu < \alpha \sqrt[d]{N} \\ \gamma & \text{otherwise} \end{cases}. \tag{33}$$

where $\rho > 0$ is the radial frequency and $b = 2$ for jitter sampling pattern [Gabrielli and Torquato 2004; Torquato et al. 2006]. Substituting Eq. (33) for $\left(1 - \prod_i^d \text{Sinc}(\pi\nu_i)^2\right)$ in Eq. (32), we get the generalised expected power spectrum form, which in radial frequency terms can be written as follows:

$$\left\langle \breve{\mathcal{P}}_S(\rho) \right\rangle = \begin{cases} \frac{\gamma\rho^b}{\alpha^b \sqrt[d]{N^b} N} & \nu < \alpha \sqrt[d]{N} \\ \frac{\gamma}{N} & \text{otherwise} \end{cases}, \tag{34}$$

where $\breve{\mathcal{P}}_S(\cdot)$ represents the radial power spectrum. Note that, the extra $\frac{1}{N}$ factor in Eq. (34) is due to the factor $1/N$ in Eq. (32). This generalised polynomial form of the power spectra is the same as the one proposed by Pilleboue and colleagues [2015], with the only exception that in this case (Eq. (34)) we are using a fully normalised power spectrum that gives an extra $\frac{1}{N}$ factor in this form.

A.2 Analytic Line Samples

In the case of line samples, we have shown in the paper that the analytic jittered line sampling power spectrum can be written as:

$$\langle \mathcal{P}_L(\omega) \rangle = \left\{ \frac{1}{N}\left[1 - \prod_i^{d-1} \text{Sinc}(\pi\nu_i \mathbf{l}_i^{\perp})^2\right] \quad \text{for jittered} \right. \tag{35}$$

As a result, the above expression in Eq. (34), for the case of line samples can be given as follows:

$$\left\langle \breve{\mathcal{P}}_L(\rho) \right\rangle = \begin{cases} \frac{\gamma\rho^b}{\alpha^b N^{\frac{b}{d-1}} N} & \rho < \alpha N^{\frac{1}{d-1}} \\ \frac{\gamma}{N} & \text{otherwise} \end{cases}. \tag{36}$$

where $\breve{\mathcal{P}}_L(\rho)$ represents the radial power spectrum, $b = 2$ for jitter sampling pattern and N represents the number of line samples. Note that, compared to Eq. (34), in Eq. (36) we have $d-1$ instead of d, which arises due to the line sampling power spectrum spanning only a $(d-1)$-dimensional subspace. As a result, the line sampling power spectrum in this $d-1$-dimensional subspace would have the form of the *point sampling* power spectrum, with zero power along the d-th dimension. For example, in $3D$ with line samples along the Z-axis, we would obtain a $2D$ sampling power spectrum on the X-Y space which would be the same as having a $2D$ sampling power spectrum due to $2D$ point samples.

B Proof

To show that the overall convergence rate of a sampling pattern can be obtained from a particular direction of its sampling power spectrum, we start by assuming that only along one direction the sampling power spectrum has constant energy over the whole radial domain, i.e, it is bounded by a constant profile ($b_k = 0$), whereas, for all other directions we have a non-zero monomial behavior ($b_k > 0$) in the low frequency region. We rewrite Eq. (9) in-parts as follows:

$$Var\left(I_N\right) < \sum_k^{\substack{m-1 \\ \text{directions}}} \left[\int_0^\infty \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k) \right\rangle \mathcal{P}_f(\rho\mathbf{n}_k) \mathrm{d}\rho\right] \Delta\mathbf{n}_k \tag{37}$$

$$+ \int_0^\infty \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_r) \right\rangle \mathcal{P}_f(\rho\mathbf{n}_r) \mathrm{d}\rho\,, \tag{38}$$

where, we have separated one of the direction $1 \leq r \leq m$ (with constant profile) from all the $m-1$ directions. We divide the $m-1$ directions in the low and high frequency regions while keeping the

r-th direction as it is:

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \Bigg[\int_0^{\nu} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho + \int_{\nu}^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho\Bigg]\Delta\mathbf{n}_k \quad (39)$$

$$+ \int_0^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_r)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_r)\mathrm{d}\rho \quad (40)$$

Since the r-th direction has a constant profile ($\mathcal{P}_S(\omega) = \gamma/N$) we can rewrite the whole integral as:

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \Bigg[\int_0^{\nu} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho + \int_{\nu}^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho\Bigg]\Delta\mathbf{n}_k \quad (41)$$

$$+ \int_0^{\infty} \rho^{d-1} \frac{\gamma}{N} \mathcal{P}_f(\rho\mathbf{n}_r)\mathrm{d}\rho \quad (42)$$

Here, the integral over the r-th direction doesn't depend on N irrespective of what the function f is, therefore, we can rewrite it in $\mathcal{O}(\cdot)$ notation as follows:

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \Bigg[\int_0^{\nu} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho + \int_{\nu}^{\infty} \rho^{d-1} \left\langle \mathcal{P}_S(\rho\mathbf{n}_k)\right\rangle \mathcal{P}_f(\rho\mathbf{n}_k)\mathrm{d}\rho\Bigg]\Delta\mathbf{n}_k \quad (43)$$

$$+ \mathcal{O}\left(\frac{1}{N}\right) \quad (44)$$

To solve the integrals over the other $m-1$ direrctions, we chose the worst and the best cases (Eq. (16)) for a given class of functions:

Worst Case:

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \Bigg[\int_0^{\rho_0} \rho^{d-1} \frac{\gamma_k}{N} \left(\frac{\rho}{\alpha_k N_k}\right)^{b_k} c_f \mathrm{d}\rho \quad (45)$$

$$+ \int_{\rho_0}^{\nu} \rho^{d-1} \frac{\gamma_k}{N} \left(\frac{\rho}{\alpha_k N_k}\right)^{b_k} c_f \rho^{-d-1} \mathrm{d}\rho \quad (46)$$

$$+ \int_{\nu}^{\infty} \rho^{d-1} \frac{\gamma_k}{N} c_f \rho^{-d-1} \mathrm{d}\rho\Bigg]\Delta\mathbf{n}_k + \mathcal{O}\left(\frac{1}{N}\right)$$

$$< \frac{\gamma_k\, c_f}{N} \sum_{k}^{\substack{m-1\\ \text{directions}}} \Bigg[\int_0^{\rho_0} \frac{\rho^{b_k+d-1}}{(\alpha_k N_k)^{b_k}} \mathrm{d}\rho + \int_{\rho_0}^{\nu} \frac{\rho^{b_k-2}}{(\alpha_k N_k)^{b_k}} \mathrm{d}\rho + \int_{\nu}^{\infty} \rho^{-2} \mathrm{d}\rho\Bigg]\Delta\mathbf{n}_k + \mathcal{O}\left(\frac{1}{N}\right) \quad (47)$$

Here, $\nu = \alpha_k N_k$, which after plugging in above equation gives:

$$Var\left(I_N\right) < \begin{cases} \sum_{k}^{\substack{m-1\\ \text{directions}}} \mathcal{O}\left(\frac{1}{N {N_k}^{b_k}}\right) + \mathcal{O}\left(\frac{1}{N}\right) & 0 < b_k < 1 \\ \sum_{k}^{\substack{m-1\\ \text{directions}}} \mathcal{O}\left(\frac{1}{N N_k}\right) + \mathcal{O}\left(\frac{1}{N}\right) & b_k \geq 1 \end{cases} \quad (48)$$

Since, the sum of $\mathcal{O}(\cdot)$ notations is asymptotically dominated by the worst $\mathcal{O}(\cdot)$ behaviour. We can write the overall convergence rate, using Eq. (48), in the following form:

$$\boxed{Var\left(I_N\right) < \mathcal{O}\left(\frac{1}{N}\right)} \quad (49)$$

This shows that, the asymptotic convergence rate of a sampling pattern, with anisotropic sampling power spectra, depends on the direction that has the $\mathcal{O}(1/N)$ convergence rate. However, if none of the directions has a constant profile behaviour, the overall convergence rate can be summarized from Eq. (48) in the following form:

$$\boxed{Var\left(I_N\right) < \begin{cases} \mathcal{O}\left(\frac{1}{N {N_k}^{b_k}}\right) & \underset{\text{infimum}}{0<b_k<1} \\ \mathcal{O}\left(\frac{1}{N N_k}\right) & b_k \geq 1 \end{cases}} \quad (50)$$

where, the overall convergence rate would be dominated by the k-th direction having the minimum value of b_k, for $b_k \in (0, 1)$.

Best Case:

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \left[\int_0^{\rho_0} \rho^{d-1} \frac{\gamma_k}{N} \left(\frac{\rho}{\alpha_k N_k}\right)^{b_k} c_f \mathrm{d}\rho\right]\Delta\mathbf{n}_k + \mathcal{O}\left(\frac{1}{N}\right)$$

$$< \frac{\gamma_k\, c_f}{N} \sum_{k}^{\substack{m-1\\ \text{directions}}} \left[\int_0^{\rho_0} \frac{\rho^{b_k+d-1}}{(\alpha_k N_k)^{b_k}} \mathrm{d}\rho\right]\Delta\mathbf{n}_k + \mathcal{O}\left(\frac{1}{N}\right) \quad (51)$$

$$Var\left(I_N\right) < \sum_{k}^{\substack{m-1\\ \text{directions}}} \mathcal{O}\left(\frac{1}{N {N_k}^{b_k}}\right) + \mathcal{O}\left(\frac{1}{N}\right), \quad (52)$$

where, $b_k > 0$. Again, since the sum of $\mathcal{O}(\cdot)$ notations is asymptotically dominated by the worst $\mathcal{O}(\cdot)$ behaviour. The overall convergence rate is:

$$\boxed{Var\left(I_N\right) < \mathcal{O}\left(\frac{1}{N}\right)} \quad (53)$$

However, from Eq. (52), if none of the directions has a constant profile behaviour, the overall convergence rate would be dominated by the k-th direction having the minimum value of b_k, and can be written in the following form:

$$\boxed{Var\left(I_N\right) < \left\{\mathcal{O}\left(\frac{1}{N {N_k}^{b_k}}\right) \quad \underset{\text{infimum}}{\forall\, b_k>0}\right.} \quad (54)$$

□

References

ALLIEZ, P., COHEN-STEINER, D., DEVILLERS, O., LÉVY, B., AND DESBRUN, M. 2003. Anisotropic polygonal remeshing. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 22*, 3 (July).

APOSTOL, T. 1974. *Mathematical Analysis*. Addison-Wesley series in mathematics. Addison-Wesley Publishing Company.

BALZER, M., SCHLÖMER, T., AND DEUSSEN, O. 2009. Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Trans. Graph. (Proc. SIGGRAPH) 28*, 3.

BARRINGER, R., GRIBEL, C. J., AND AKENINE-MÖLLER, T. 2012. High-quality curve rendering using line sampled visibility. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 31*, 6 (Nov.).

BRACEWELL, R. 2000. *The Fourier Transform and Its Applications*. Electrical engineering series. McGraw Hill.

BRANDOLINI, L., COLZANI, L., AND TORLASCHI, A. 2001. Mean square decay of Fourier transforms in euclidean and non euclidean spaces. *Tohoku Math. J. (2) 53*, 3.

BRANDOLINI, L., HOFMANN, S., AND IOSEVICH, A. 2003. Sharp rate of average decay of the fourier transform of a bounded set. *Geometric & Functional Analysis GAFA 13*, 4.

COOK, R. L., PORTER, T., AND CARPENTER, L. 1984. Distributed ray tracing. *Computer Graphics (Proc. SIGGRAPH) 18*, 3 (Jan.).

COOK, R. L. 1986. Stochastic sampling in computer graphics. *ACM Transactions on Graphics 5*, 1 (Jan.).

DE GOES, F., BREEDEN, K., OSTROMOUKHOV, V., AND DESBRUN, M. 2012. Blue noise through optimal transport. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 31*.

DIPPÉ, M. A. Z., AND WOLD, E. H. 1985. Antialiasing through stochastic sampling. *Computer Graphics (Proc. SIGGRAPH)*.

DIPPÉ, M. A. Z., AND WOLD, E. H. 1992. Progress in computer graphics (vol. 1). Ablex Publishing Corp., Norwood, NJ, USA, ch. Stochastic Sampling: Theory and Application.

DURAND, F. 2011. A frequency analysis of Monte-Carlo and other numerical integration schemes. Tech. Rep. TR-2011-052, MIT CSAIL.

FENG, L., HOTZ, I., HAMANN, B., AND JOY, K. I. 2008. Anisotropic noise samples. *Visualization and Computer Graphics, IEEE Transactions on 14*, 2.

GABRIELLI, A., AND TORQUATO, S. 2004. Voronoi and void statistics for superhomogeneous point processes. *Physical Review E 70*, 4.

1966. Generalized functions. In *Integral Geometry and Representation Theory*, G. I.M., G. M.I., and V. N.Ya., Eds. Academic Press.

GEORGIEV, I., KŘIVÁNEK, J., HACHISUKA, T., NOWROUZEZAHRAI, D., AND JAROSZ, W. 2013. Joint importance sampling of low-order volumetric scattering. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 32*, 6 (Nov.).

GRIBEL, C. J., DOGGETT, M., AND AKENINE-MÖLLER, T. 2010. Analytical motion blur rasterization with compression. In *Proc. of HPG*.

GRIBEL, C. J., BARRINGER, R., AND AKENINE-MÖLLER, T. 2011. High-quality spatio-temporal rendering using semi-analytical visibility. *ACM Trans. Graph. (Proc. SIGGRAPH) 30* (Aug.).

HECK, D., SCHLÖMER, T., AND DEUSSEN, O. 2013. Blue noise sampling with controlled aliasing. *ACM Transactions on Graphics 32*, 3.

JAROSZ, W., ZWICKER, M., AND JENSEN, H. W. 2008. The beam radiance estimate for volumetric photon mapping. *Comp. Graph. Forum (Proc. Eurographics) 27*, 2 (Apr.).

JAROSZ, W., NOWROUZEZAHRAI, D., SADEGHI, I., AND JENSEN, H. W. 2011. A comprehensive theory of volumetric radiance estimation using photon points and beams. *ACM Transactions on Graphics 30*, 1 (Feb.).

JAROSZ, W., NOWROUZEZAHRAI, D., THOMAS, R., SLOAN, P.-P., AND ZWICKER, M. 2011. Progressive photon beams. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 30*, 6 (Dec.).

JONES, T. R., AND PERRY, R. N. 2000. Antialiasing with line samples. In *Rendering Techniques (Proc. EG Workshop on Rendering)*, Springer-Verlag, London, UK.

KEISLER, H. 2012. *Elementary Calculus: An Infinitesimal Approach*. Dover Books on Mathematics Series. Dover Publications.

KŘIVÁNEK, J., GEORGIEV, I., HACHISUKA, T., VÉVODA, P., ŠIK, M., NOWROUZEZAHRAI, D., AND JAROSZ, W. 2014. Unifying points, beams, and paths in volumetric light transport simulation. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (July).

LAGAE, A., AND DUTRÉ, P. 2008. A comparison of methods for generating Poisson disk distributions. *Comp. Graph. Forum 27*, 1.

LENEMAN, O. A. 1966. Random sampling of random processes: Impulse processes. *Information and Control 9*, 4.

LÉVY, B., AND LIU, Y. 2010. Lp centroidal voronoi tessellation and its applications. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 29*, 4 (July).

LI, H., WEI, L.-Y., SANDER, P. V., AND FU, C.-W. 2010. Anisotropic blue noise sampling. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 29*, 6 (Dec.).

MITCHELL, D. 1991. Spectrally optimal sampling for distributed ray tracing. *Computer Graphics (Proc. SIGGRAPH) 25*, 4.

MITCHELL, D. P. 1996. Consequences of stratified sampling in graphics. In *Annual Conference Series (Proc. SIGGRAPH)*, ACM.

NOVÁK, J., NOWROUZEZAHRAI, D., DACHSBACHER, C., AND JAROSZ, W. 2012. Virtual ray lights for rendering scenes with participating media. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 4 (July).

NOWROUZEZAHRAI, D., BARAN, I., MITCHELL, K., AND JAROSZ, W. 2014. Visibility silhouettes for semi-analytic spherical integration. *Comp. Graph. Forum 33*, 1 (Feb.).

PHARR, M., AND HUMPHREYS, G. 2010. *Physically Based Rendering: From Theory To Implementation*, 2nd ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

PILLEBOUE, A., SINGH, G., COEURJOLLY, D., KAZHDAN, M., AND OSTROMOUKHOV, V. 2015. Variance analysis for Monte Carlo integration. *ACM Trans. Graph. (Proc. SIGGRAPH) 34*, 4 (July).

RAMAMOORTHI, R., ANDERSON, J., MEYER, M., AND NOWROUZEZAHRAI, D. 2012. A theory of Monte Carlo visibility sampling. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 5.

REINERT, B., RITSCHEL, T., SEIDEL, H., AND GEORGIEV, I. 2016. Projective Blue Noise Sampling. *Comp. Graph. Forum*.

SANTALÓ, L. 1953. *Introduction to integral geometry*. Actualités scientifiques et industrielles. Hermann.

SHIRLEY, P. 1991. Discrepancy as a quality measure for sample distributions. In *Proc. Eurographics*.

SPANIER, J. 1966. Two pairs of families of estimators for transport problems. *SIAM Journal on Applied Mathematics 14*, 4.

SUBR, K., AND KAUTZ, J. 2013. Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. Graph. (Proc. SIGGRAPH) 32*, 4 (July).

SUBR, K., NOWROUZEZAHRAI, D., JAROSZ, W., KAUTZ, J., AND MITCHELL, K. 2014. Error analysis of estimators that use combinations of stochastic sampling strategies for direct illumination. *Comp. Graph. Forum (Proc. EGSR) 33*, 4 (June).

SUN, X., ZHOU, K., LIN, S., AND GUO, B. 2010. Line space gathering for single scattering in large scenes. *ACM Trans. Graph. (Proc. SIGGRAPH) 29*, 4 (July).

SUN, X., ZHOU, K., GUO, J., XIE, G., PAN, J., WANG, W., AND GUO, B. 2013. Line segment sampling with blue-noise properties. *ACM Trans. Graph. (Proc. SIGGRAPH) 32*, 4 (July).

TORQUATO, S., UCHE, O., AND STILLINGER, F. 2006. Random sequential addition of hard spheres in high euclidean dimensions. *Physical Review E 74*, 6.

TZENG, S., PATNEY, A., DAVIDSON, A., EBEIDA, M. S., MITCHELL, S. A., AND OWENS, J. D. 2012. High-quality parallel depth-of-field using line samples. In *Proc. of HPG*.

ULICHNEY, R. 1987. *Digital Halftoning*. MIT Press, Cambridge, MA, USA.

WACHTEL, F., PILLEBOUE, A., COEURJOLLY, D., BREEDEN, K., SINGH, G., CATHELIN, G., DE GOES, F., DESBRUN, M., AND OSTROMOUKHOV, V. 2014. Fast tile-based adaptive sampling with user-specified Fourier spectra. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4.

WEI, L.-Y., AND WANG, R. 2011. Differential domain analysis for non-uniform sampling. *ACM Trans. Graph. 30*, 4 (July).