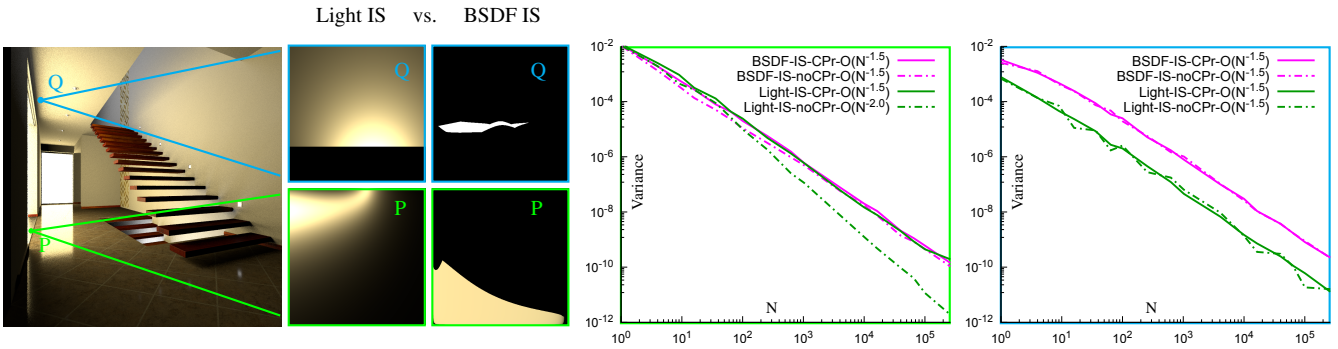
Volume 38 (2019), Number 1 pp. 1–12 COMPUTER GRAPHICS *forum*

Fourier Analysis of Correlated Monte Carlo Importance Sampling

Gurprit Singh[1,4] Kartic Subr[2] David Coeurjolly[3] Victor Ostromoukhov[3] Wojciech Jarosz[4]

[1]Max-Planck Institute for Informatics, Saarbrücken, Germany
[2]University of Edinburgh, UK
[3]Université de Lyon, CNRS/LIRIS, France
[4]Dartmouth College, USA



(a) Reference (b) Underlying pixel functions (c) Pixel P (d) Pixel Q

Figure 1: *The convergence rate of Monte Carlo integration with correlated (jittered) samples depends on the importance function (light vs. BSDF). We analyze two Pixels P & Q directly illuminated by a square area light source in (**a**). Pixel P is fully visible from the light source, which results in a smooth integrand when performing light source surface area sampling (visualized in bottom-left **b**) and a convergence rate of $\mathcal{O}(N^{-2})$ (**c**: dot-dashed green curve, all plots are in log-log scale). C_0 discontinuities in the integrand result in $\mathcal{O}(N^{-1.5})$ convergence, which can happen: when using homogenization or Cranley-Patterson rotation [CP76] (CPr, solid green curve) since it introduces boundary discontinuities; when the light source is partially occluded (Pixel Q); or when sampling the BSDF (in magenta) since this treats the boundary of the light as a C_0 discontinuity even when the light is fully visible (visualized in the second column of **b**).*

Abstract

Fourier analysis is gaining popularity in image synthesis as a tool for the analysis of error in Monte Carlo (MC) integration. Still, existing tools are only able to analyze convergence under simplifying assumptions (such as randomized shifts) which are not applied in practice during rendering. We reformulate the expressions for bias and variance of sampling-based integrators to unify non-uniform sample distributions (importance sampling) as well as correlations between samples while respecting finite sampling domains. Our unified formulation hints at fundamental limitations of Fourier-based tools in performing variance analysis for MC integration. At the same time, it reveals that, when combined with correlated sampling, importance sampling (IS) can impact convergence rate by introducing or inhibiting discontinuities in the integrand. We demonstrate that the convergence of multiple importance sampling (MIS) is determined by the strategy which converges slowest and propose several simple approaches to overcome this limitation. We show that smoothing light boundaries (as commonly done in production to reduce variance) can improve (M)IS convergence (at a cost of introducing a small amount of bias) since it removes C_0 discontinuities within the integration domain. We also propose practical integrand- and sample-mirroring approaches which cancel the impact of boundary discontinuities on the convergence rate of estimators.

CCS Concepts

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

1. Introduction

Monte Carlo (MC) and other sampling-based integration techniques have become the cornerstone of modern rendering algorithms due to their ability to approximate high-dimensional integrals. The strategy used to generate the samples has a crucial impact on approximation error, and the literature is rich with strategies to mitigate error. Modern image synthesis algorithms typically combine several strategies to form unified estimators in hopes of exploiting each of their strengths. While combined estimators are common in practice, their analysis remains an open problem. In this paper, we analyze importance sampling when used in conjunction with correlated sampling patterns.

Importance sampling (IS) [Coc63] is a strategy that mitigates error by drawing samples from a tailored parent distribution (importance function). The variance of the IS estimator is proportional to the variance of the *effective integrand*: the original integrand divided by the importance function. In situations where the qualitative behavior of the integrand can be predicted, importance functions are designed, a priori, to reduce the variance of the effective integrand. The resulting approximation error depends on the choice of importance function, relative to the integrand, and the benefit due to IS is independent of the number of samples used. Consequently, despite IS, the *variance convergence rate* for pseudorandomly distributed samples remains at $\mathcal{O}(N^{-1})$ like naïve MC sampling.

Correlated sampling (CS) is another error-mitigation strategy, which operates by introducing correlations within each realization of N sample points, e.g. in jittered sampling, correlations are induced by the relative positions of the cells positioned on a regular grid. Despite each of the N samples being drawn pseudorandomly within a single cell, jittered sampling improves convergence rate [Hab70]. Some other strategies, viz. Poisson Disk sampling or Quasi-Monte Carlo (QMC) sampling, explicitly impose correlations by construction to improve equidistribution [Zar72].

Tools for analyzing sampling-based integrators can be broadly classified into three categories: Analysis in the primal domain, the Fourier domain, and number theoretic approaches. Error analysis of IS has predominantly been in the primal domain while CS analyses have favored the latter two categories. The analysis of the combined estimator requires a unified formulation.

Our formulation builds upon the early work by Sloan and Joe [SJ94] which analyzed the approximation error—of smooth periodic functions that have absolute convergent Fourier series representations—using *deterministic* lattice methods. We reformulate their expression for error in terms of bias and variance to analyze *stochastic* sampling-based integrators. Our formulation unifies non-uniform sample distributions (importance sampling) as well as correlations between (stochastic) samples while respecting finite sampling domains. Our formulation further emphasizes the role of phase information to maximize variance and convergence gains during correlated and/or importance sampling of C_0-discontinuous integrands, which has previously been ignored by assuming uniform and random translation/rotation of samples [RAMN12, Ö16].

In this paper we take two steps to tackle the problem of analyzing and improving combined correlated and importance-sampled MC estimators:

(1) We resolve important limitations of existing Fourier analyses (Sec. 4), which have previously assumed a constant importance function, homogenization, and an infinite sampling domain. Under these assumptions, the MC integration error can be expressed as the sum of two terms. Unfortunately, none of these simplifying assumptions are true in practical rendering scenarios, making the theory mispredict error and convergence rates in practice. We derive a third (previously ignored) term which is needed to properly predict error once any one of these assumptions is violated, revealing the reason for prior disagreements between theory and practice. Unfortunately, our derivations also reveal that *directly* leveraging this Fourier-based analysis is likely impractical in the general setting due to the curse of dimensionality: the complexity of the previously missing third term grows exponentially with the dimensionality of the problem.

(2) Our Fourier analysis, however, provides key insights which we leverage *indirectly* for practical improvement for the simplified problem of direct illumination with multiple importance sampling (MIS) [Vea98]. Our analysis reveals that the finite sampling domain, homogenization and importance sampling all interplay to introduce or inhibit C_0 discontinuities in the integrand, and that this can have a dramatic impact on not just error, but also *convergence rate*. We also show that MIS is not able to overcome these detrimental effects as it inherits the convergence properties of the worst of the constituent strategies. With these interactions in mind, we show that (M)IS convergence improves by 1) removing certain discontinuities (on the integration domain boundary thanks to a mirroring strategy or within the domain by integrand filtering) or 2) by pushing them to the boundary of the integration domain. While pre-filtering introduces some bias, it allows us to provably retain (in certain situations) the improved convergence rate of correlated sampling when MISing BSDF and light sampling for direct illumination.

2. Related work

Sampling-based integrators construct estimates of integrals using weighted averages of the integrand evaluated at sample locations. Their error characteristics are dictated by the locations of the samples and the weights associated with each sample. For uniformly distributed pseudorandom samples with equal weights, a general strategy used to improve convergence of the estimator is to introduce correlations between samples. Another approach is to tailor the distribution of the samples a priori while accordingly adjusting the weights to maintain desirable properties in the resulting estimator.

Correlated Sampling: Neyman [Ney34] adopts a divide-and-conquer strategy whereas Cook [Coo86] proposed to impose distance constraints between samples. The former, or stratification, is a classical strategy yielding estimators with convergence rates of $\mathcal{O}\left(N^{-1-b/d}\right)$ [Hab70], where d is the dimensionality of the sampling domain and $b \in \mathbb{R}^+$ depends on the smoothness properties of the integrand. Distance-based constraints appear in a variety of contexts such as hard-core point processes [IPSS08] and blue noise sampling [Uli87]. In computer graphics, several techiques have been proposed for analysis [DW92, Mit91, Shi91, OF01, LD08, OG12, WW11] as well as synthesis [Coo86, BSD09, dGBOD12, KTBV16, WPC*14, CSHD03, KCODL06, AHD15] of these types of samples. Yet another class of methods sacrifices stochasticity for equidistribution, and in-

duces favorable correlations by minimizing a measure known as *discrepancy*. Such low discrepancy sampling methods [Nie78, KPR12] and the resulting QMC estimators are considered state of the art in rendering [PJH16]. They preserve desirable equidistribution properties without requiring knowledge of the number of samples a priori and lead to estimators with superior convergence rates for low and medium dimensional spaces when integrands respect certain smoothness criteria. C_0 discontinuities in the integrand are also known to adversely affect convergence rates of jittered sampling estimators [Mit96].

Importance Sampling: The scattering integral over the incident sphere (or hemisphere) of directions includes radiance, visibility, the scattering function (BSDF or phase function) and a cosine. The literature contains numerous methods that use different combinations of these components [ARBJ03, GKH*13, Hd14] as parent distributions either by multiplying them [CJAMJ05, CAM08, JCJ09] or by averaging them using multiple importance sampling (MIS) [Vea98]. A comprehensive study can be found in PBRT [PJH16]. Importance sampling has also been combined with pre-filtering [KC08] to suppress errors.

Error analysis of integrators: The error due to point distributions [Rip77, IPSS08] have been well studied in the statistical literature [OZ00, Owe13, Hes03]. In computer graphics, various measures such as spatial discrepancy [Shi91], statistical tests of hypothesis [SA07], and pair correlations [Ö16] have been proposed for quantifying the error characteristics of sampling-based estimators. Since correlations impact the spectral distribution, many researchers [Coo86, DW85, Mit91, Lem09] have analyzed these approaches in the Fourier domain.

Durand [Dur11], Subr and Kautz [SK13] rewrote MC estimators in the frequency domain and derived expressions for bias and variance. Pilleboue et al. [PSC*15] extended these formulations for correlated samples with isotropic spectral distirbutions allowing in-depth convergence analysis. Singh and Jarosz [SMJ17, SJ17] later enhanced these formulations for anisotropic power spectra. Sloan and Joe [SJ94] performed similar error analyses for *deterministic* samples using the Fourier series to account for the finite domain sampling. Our analysis also uses the Fourier series due to finite sampling domains but in contrast to Sloan and Joe [SJ94] we focus on *stochastically* generated samples with non-uniform probability density functions (PDFs) and propose theoretical closed-form expressions for both bias and variance in terms of the spectral properties of both the sampler and the integrand involved.

Error due to discontinuities has been well studied in the Fourier domain. The localization of the Fourier transform around a discontinuity has received considerable attention in mathematics [Hör90]. Recently, Lessig [Les17] address visibility-induced aliasing by exploiting the fact that the Fourier transform of a discontinuity decays slowly only in the normal direction. In rendering, Ramamoorthi et al. [RAMN12] performed a comprehensive analysis of errors due to partial visibility while integrating in pixel-light space. This work presented a statistical and Fourier analysis of the impact of discontinuities on estimation error (while ignoring phase information), and demonstrated that it is possible to mitigate error by randomly shifting a regular grid. In a similar vein, Oztirelli [Ö16] recently demonstrated improvements by uniformly rotating strata boundaries along the discontinuities. Our formulation shows that phase information is crucial for variance and convergence improvement which is lost with uniform and random shifting. We further reveal that convergence rates can be severely impacted due to simplifying assumptions (e.g. toroidal wrapping and Cranley-patterson rotation [CP76] or homogenization [PSC*15]) made in sample generation algorithms and sampling strategies during integration.

3. Review of Fourier analyses of MC

Consider an N-sample-estimator $\mu_N \approx I$ where

$$I := \int_{[0,1]^d} f(\mathbf{x})\, \mathrm{d}\mathbf{x}, \quad \text{and} \quad \mu_N := \frac{1}{N}\sum_{k=1}^{N} \alpha_k f(\mathbf{x}_k). \tag{1}$$

Here $\mathbf{x} \in [0,1]^d$ is the spatial integration variable and $\mathbf{x}_k \in [0,1]^d$ is the k^{th} sample along with its associated *sampling weight* $\alpha_k \in \mathbb{R}^+$. For an unbiased MC estimator, the expected value $\langle \mu_N \rangle = I$. This can be achieved by choosing sampling weights as the reciprocal of the probability density evaluated at the sampled locations [SK13].

In previous work [Dur11, SK13, PSC*15], the estimator was rewritten as an inner product of the integrand f and a sampling function $S(\mathbf{x}) := \frac{1}{N}\sum_{k=1}^{N} \alpha_k \delta(\mathbf{x} - \mathbf{x}_k)$ where δ denotes the Dirac delta distribution. Further, since the Fourier transform preserves inner products, the estimator was written as the inner product of the complex conjugate of the Fourier transform of the integrand $\mathcal{F}_f$ and the Fourier transform of the sampling function $\mathcal{F}_S$. This can be written as an integral over an infinite domain,

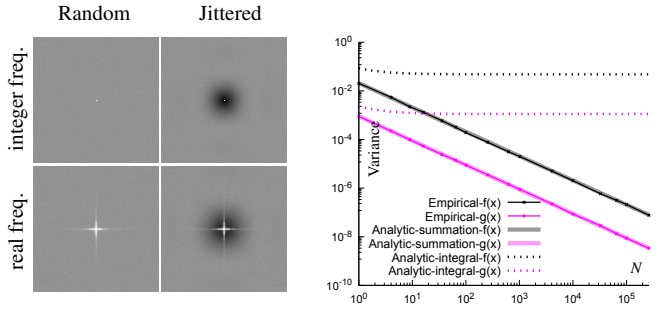
$$\mu_N = \int_{\mathbb{R}^d} \mathcal{F}_f(\nu)^*\, \mathcal{F}_S(\nu)\, \mathrm{d}\nu, \tag{2}$$

where the frequency variable ν is continuous. Since the Fourier transform operates on an infinite domain, it is unclear how to handle finite domains.

One possibility is to multiply the integrand (or the sampling pattern) by a box [SK13] (or its higher dimensional equivalent), which results in the Fourier spectrum being convolved by a Sinc (or its high-dimensional analog). Subr and Kautz [SK13] expressed the variance of the estimator in terms of the expected Fourier spectra of the sampling pattern and integrand. In the context of computer graphics, power spectra (real numbers) are often encountered rather than Fourier spectra (complex numbers). Pilleboue et al. [PSC*15] derived the variance in terms of sampling $\mathcal{P}_S(\cdot)$ and integrand $\mathcal{P}_f(\cdot)$ power spectra, under certain assumptions of stationarity:

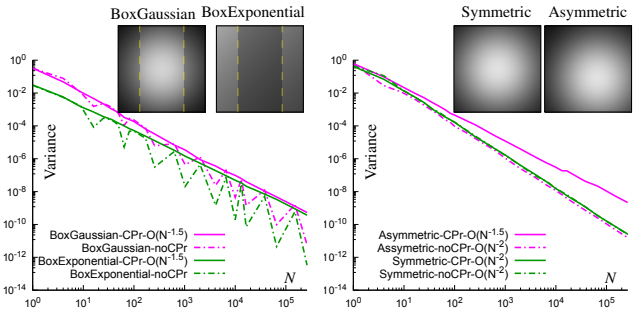
$$\mathrm{Var}\,(\mu_N) = \int_{\mathbb{R}^d} \langle \mathcal{P}_S(m) \rangle\, \mathcal{P}_f(m)\, \mathrm{d}m \; - \; I^2, \tag{3}$$

where $\langle \cdot \rangle$ represents the expectation operator. Without special handling of the finite domain, the d-dimensional expected sampling power spectrum gets polluted by $\prod_i^d \mathrm{Sinc}(\pi m_i)^2$ term at non-integer frequencies (see Fig. 2a) resulting in erroneous prediction of variance (dotted curves in Fig. 2b). We analytically derive expressions for random and jittered samples' expected power spectra in the supplemental Sec. 5.2 (eq. 51) & Sec. 6.3 (eq. 86).



(a) Expected power spectra (b) Convergence plots

Figure 2: *The expected power spectra in (**a**) for random and jittered samples is polluted by a $\mathrm{Sinc}(\cdot)^2$ term at non-integer frequencies due to the finite sampling domain (visible as a white cross in the bottom row). This cross pollutes variance prediction when the continuous formulation (3) is used to compute variance (dotted curves in **b**) since the integral is computed at all* real *frequencies. This results in a non-converging behavior. Our Fourier series based variance formulation (8) ensures that empirical and analytic computations match correctly (solid curves with and without markers). The plots (in log-log scale) show the convergence for a tent function $f(x)$ and a smooth function $g(x)$ for which variance can be computed analytically (details in suppl. Sec. 7.3).*



(a) Strata alignment (b) Boundary discontinuities

Figure 3: *(**a**) Jittered sampling shows good convergence (dotted curve) when the strata boundaries align with the step discontinuities introduced at $x = 0.25$ and $x = 0.75$ in an otherwise smooth gaussian and an exponential function. However, homogenization (CPr) destroys this alignment resulting in bad convergence (shown in log-log scale) behavior overall (solid curves). (**b**) A gaussian function showing best-case convergence from Pilleboue et al. [PSC*15] behaves badly (solid magenta) after homogenization when slightly off-centered (details in Sec. 5.2).*

If the integral is calculated only at integer frequencies, however, prediction using the same equation matches empirical results on both test integrands (thick curves in Fig. 2b). Pilleboue et al. managed to get correct error bounds due to the underlying assumption of toroidal domain in their experimental setup [KSP*15]. This highlights the first gap in existing methods, which is that the formulations as derived *do not handle the finite domain of I* and must be modified to match empirical results.

In order to extend the original equations [Dur11, SK13] to correlated samples, Pilleboue et al. [PSC*15] transformed estimators to be shift-invariant by assuming each realization of N samples is translated by a unique random offset. Unfortunately this *homogenization* step, which was vital for their analysis, significantly alters the behavior of the estimators leading to incorrect prediction. Fig. 3 demonstrates this disagreement in predicted and observed behavior due to homogenization for two cases: 1) homogenization destroys alignment of strata boundaries with the step discontinuities resulting in worst convergence overall (Fig. 3a), and 2) homogenization introduces boundary discontinuities for asymmetric functions as shown for a translated Gaussian in Fig. 3b (Sec. 5.2).

Furthermore, a major shortcoming of this assumption is that it precludes analyses of non-uniform samples in the Fourier domain. One could imagine using the same formulation (3) for IS by warping the integrand back in the primary sample space, however, integrand smoothness would be affected due to homogenization resulting in bad convergence behavior (Fig. 3). Consequently, existing formulations *are ill-suited for analyzing IS* in the Fourier domain.

To summarize, previous work on Fourier analysis of Monte Carlo schemes do not account for non-uniform distributions. They handle correlated samples implicitly since the expected power spectrum of the sampling function encodes this information. However, these correlations as well as the relative alignment of the importance function with the integrand are lost due to homogenization; they are not handled in the equations for variance in the Fourier domain. In the following section, we derive a generalized expression for the error of sampling-based integrators in the Fourier domain, which accounts for finite domains, includes non-uniform samples and incorporates correlations between samples.

4. Fourier domain analysis

4.1. New formulation based on Fourier series

We assume that the sampling domain is toroidal. Under this assumption, the integrand and sampling function may be seen as periodic functions with periodicity given by the extent of the domain (1 along each axis for integration over the hypercube). This leads to a discrete spectrum characterized by Fourier series coefficients (Fig. 4).

Finite domain: A d-dimensional, unit-periodic function $g(x)$ can be expanded in terms of its Fourier series coefficients as:

$$g(\mathbf{x}) \quad := \quad \sum_{m\in\mathbb{Z}^d} \mathbf{g}_m \, \mathrm{e}^{i2\pi(m\cdot\mathbf{x})}, \tag{4}$$

where the m^{th} coefficient $\mathbf{g}_m \in \mathbb{C}^d$ is

$$\mathbf{g}_m \quad := \quad \int_{[0,1]^d} g(\mathbf{x}) \, \mathrm{e}^{-i2\pi(m\cdot\mathbf{x})} \, \mathrm{d}\mathbf{x}. \tag{5}$$

The N-sample estimator μ_N can be written in terms of Fourier series coefficients of the integrand ($\mathbf{f}_m \in \mathbb{C}^d$) and those of the sampling function ($\mathbf{S}_m \in \mathbb{C}^d$). For simplicity, we will henceforth focus on $d = 1$, without loss of generalization. See Fig. 5 for a geometric interpretation of the first two coefficients (uniform sampling) and the second coefficient in the case of IS.

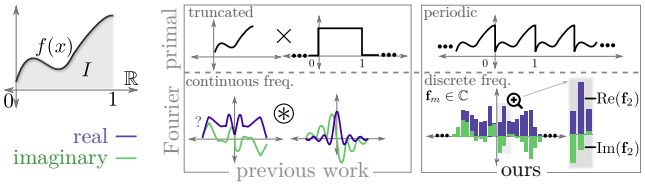


Figure 4: *Previous work applied the Fourier transform, which results in a continuous frequency domain representation of functions. However, since the sampling domain is finite, the continuous spectra gets blurred (convolved) with a Sinc(·). In this work, we rectify this by safely assuming that functions over finite domains are periodic, and therefore we represent them using their discrete Fourier spectra (Fourier series).*

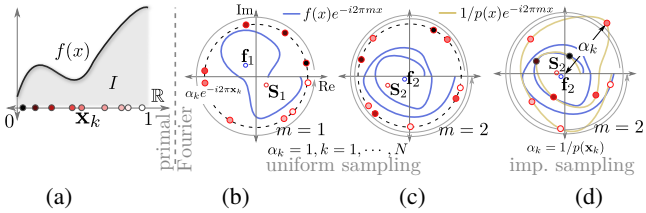


Figure 5: *Samples are shown in the primal (circles with black outline in **a**) and in the complex plane (circles with red outline in **b,c & d**) for $m = 1, 2$. For constant PDF, Fourier counterpart of each sample stays on a unit circle in the complex plane (dashed circles in b & c) whereas after IS these samples lie along the orange curve (Fourier transformed weighting function). Fourier coefficients $\mathbf{S}_1$ and $\mathbf{S}_2$ are obtained as the sum of the samples in the complex plane. The Fourier transformed integrand is shown in blue.*

Substituting the Fourier series expansions of f and S in the estimator $\mu_N = \int_0^1 f(x)S(x)\,\mathrm{d}x$, we obtain (derivation in suppl. Sec. 2):

$$\boxed{\mu_N = \sum_{m\in\mathbb{Z}} \mathbf{f}_m^* \mathbf{S}_m} \tag{6}$$

where z^* denotes the complex conjugate of $z \in \mathbb{C}$ and the m^{th} coefficient $\mathbf{S}_m := \frac{1}{N}\sum_{k=1}^{N} \alpha_k e^{-i2\pi(m\cdot \mathbf{x}_k)}$ is obtained by substituting the definition of S (in place of $g(x)$) in eq. 5. Since $\mathbf{S}_m$ is a random variable, μ_N is a random variable as expected. If $\Delta := I - \mu_N$ is the error of a single realization of μ_N, we analyze bias and variance by expanding $\langle\Delta\rangle$ and $\mathrm{Var}(\Delta)$ respectively as in previous work.

Bias: Substituting the rhs from eq. 6 in the definition of bias, $\langle\Delta\rangle = I - \langle\mu_N\rangle$, followed by separation of the summation into DC ($m = 0$) and other coefficents (see Sec. 3 of supplemental)

$$\langle\Delta\rangle \quad = \quad \mathbf{f}_0^* \left(1 - \langle\mathbf{S}_0\rangle\right) \quad - \sum_{m\in\mathbb{Z}, m\neq 0} \mathbf{f}_m^* \langle\mathbf{S}_m\rangle . \tag{7}$$

This is the discretized version (due to periodicity in the primal) of the form proposed by prior work [SK13], but our formulation accounts for the finite domain. Consider the case of random sampling with sampling weights $\alpha_k = 1$. The first term in eq. 7 vanishes since the expected DC value of the sampling spectrum is unity. As for the second term, an intuitive way to visualize $\langle\mathbf{S}_m\rangle$ is as the average of points $\mathbf{S}_m$ in the complex plane. When all $\alpha_k = 1$, these points lie on the unit circle. For random sampling and any $m \in \mathbb{Z}$, the points are distributed uniformly on the circle since $\mathbf{x}_k \in [0,1]$. Their average is zero, and so the second term in eq. 7 also vanishes for this case. If the samples are distributed non-uniformly according to a parent distribution $p(\mathbf{x})$, one way of mitigating bias would be to ensure that each of the terms[†] goes to zero. We show that choosing $\alpha_k = 1/p(\mathbf{x}_k)$ causes both terms to vanish, verifying that importance sampling is unbiased (see Sec. 3.1 in the supplemental).

Variance: For the general case of non-uniform samples with correlations, applying the variance operator to the rhs of eq. 6 results in a summation of covariance terms between all pairs of frequency variables $l, m \in \mathbb{Z}$ (see supplemental Sec. 4 for derivation). We factor this nested summation into three groups, resulting in

$$\mathrm{Var}(\Delta) \quad = \quad I^2\,\mathrm{Var}(\mathbf{S}_0) \quad + \sum_{\substack{m\in\mathbb{Z}\\ m\neq 0}} \mathbf{f}_m^*\mathbf{f}_m \left\langle \mathbf{S}_m^*\mathbf{S}_m \right\rangle$$

$$\boxed{+ \sum_{m\in\mathbb{Z}} \sum_{\substack{l\in\mathbb{Z}\\ l\neq m}} \mathbf{f}_m^*\mathbf{f}_l \left\langle \mathbf{S}_m\mathbf{S}_l^* \right\rangle .} \tag{8}$$

The first term contains the DC contribution, the second contains the variance terms (diagonal elements if it were a finite covariance matrix) and the third term contains all other terms (off-diagonal elements if it were a finite covariance matrix). We choose to express the variance in this form since the first two terms are exactly the discrete versions of previous work [PSC*15]. When homogenization is applied, the third term vanishes, thus verifying the result of Pilleboue et al. We now briefly explain each of the three terms in the rhs of eq. 8.

The first term is the variance of the DC component of the sampling spectrum, scaled by the square of the integral. For uniformly random samples with unit weights, $\mathbf{S}_0 = 1$ and therefore zero variance. In the case of importance sampling, $\mathbf{S}_0 = \frac{1}{N}\sum_{k=1}^{N}\alpha_k$ is the sum of the weights and therefore non-zero in general. One way to force the DC term to vanish, while performing IS, would be to set each weight to be $\alpha_k \leftarrow \alpha_k/\mathbf{S}_0$. This case, where the sampling weights for every secondary IS estimate add up to 1 is commonly known as normalized importance sampling [Owe13].

The first two terms together represent the discretized version of eq. 3, where $\mathbf{f}_m^*\mathbf{f}_m$ and $\mathbf{S}_m^*\mathbf{S}_m$ are the power spectra of the integrand and sampling function respectively. Since correlations between samples are encoded in the power spectra, the second term can be used to explain some of the benefits due to correlated samples. For example, the expected power spectra of jittered and blue-noise sampling patterns contain less energy at low-frequencies, leading to variance reduction. The covariance terms, which explicitly represent sample correlations in the frequency domain, have not been studied before.

The third term in eq. 8 is novel and crucial to explain IS and CS without homogenization. If the samples are homogenized, then

[†] It may also be possible that neither term is zero but their difference is zero, but such an estimator would be more difficult to tailor.

the third term vanishes. This can be proven by first defining $\mathbf{S}_m^\tau = \mathbf{S}_m e^{i2\pi\tau m}$ as the translated Fourier coefficient with τ as the translation vector, which allows rewriting the translated product of $\mathbf{S}_m\mathbf{S}_l^*$ as:

$$\mathbf{S}_m^\tau \mathbf{S}_l^{\tau *} = \mathbf{S}_m \mathbf{S}_l^* \, e^{i2\pi\tau m} e^{-i2\pi\tau l} \tag{9}$$

$$= \mathbf{S}_m \mathbf{S}_l^* \, e^{i2\pi\tau(m-l)}. \tag{10}$$

Following Pilleboue et al. [PSC*15], homogenization is equivalent to studying the translated version of each realization and averaging over all translations, which can be compactly written as:

$$\left\langle \mathbf{S}_m \mathbf{S}_l^* \right\rangle = \int_{\mathcal{T}} \mathbf{S}_m \mathbf{S}_l^* \, e^{i2\pi\tau(m-l)} \, d\tau, \tag{11}$$

where $\mathcal{T}$ is the toroidal domain. Note that, in the above integral, the exponential part is non-zero only for $m = l$. Since, $m \neq l$ in the third summation part of the variance equation in eq. 8, it renders the third term zero. This term also quantifies alignment of the samples with the integrand (Table 1) which accounts for the phase information. We study this in detail in the next section (Fig. 6). Previous analyses [RAMN12, PSC*15, Ö16] safely ignore the phase by relying on uniform and random translations/rotations of the stratification grid.

4.2. Unifying IS and CS

The expectation of the cross-terms $\langle \mathbf{S}_m \mathbf{S}_l^* \rangle$ in (8) can be simplified (see Sec. 4.2 of suppl. material) to

$$\left\langle \mathbf{S}_m \mathbf{S}_l^* \right\rangle = \frac{1}{N} \int_0^1 \frac{1}{p(x)} e^{-i2\pi(m-l)x} \, dx \tag{12}$$

$$+ \frac{1}{N^2} \sum_{k=1}^{N} \sum_{j \neq k}^{N} \int_0^1 \int_0^1 \frac{e^{-i2\pi(mx_1 - lx_2)}}{p(x_1)\, p(x_2)} \rho(x_1, x_2) \, dx_1 \, dx_2,$$

where the joint PDF $\rho(x_1, x_2)$ (related to the pair correlation function [IPSS08, OG12]) explicitly specifies correlations. The double integral in eq. 12 is zero when samples are independent (uncorrelated) since $\rho(x_1, x_2) = p(x)_1 \, p(x)_2$, leading to $\langle \mathbf{S}_m \mathbf{S}_l^* \rangle = \mathbf{W}_{m-l}/N$ where $\mathbf{W}_m \in \mathbb{C}$ is the m^{th} Fourier series coefficient of the weighting function $w(x) = 1/p(x)$, i.e. $\alpha_k = w(\mathbf{x}_k)$ (see Sec. 5 of suppl. material). We study the covariance terms by considering non-uniform distribution and correlations independently.

Perfect IS (1D): Zero variance is achieved in the academic case of the effective integrand being a constant, i.e. the sampling distribution is a scaled version of the integrand. Although this is not useful in practice, it serves as a simple case to highlight the impact of the third term in eq. 8. Consider a 1D integration problem where the integrand and the sampling distribution are step functions, with the discontinuities aligned. The samples are independent. In this case, we calculated all three terms in eq. 8 analytically (see Sec. 8.2 in the supplemental) and found that the first two terms (known in previous work) sum to a value of $0.140625/N$ while the third term is $-0.140625/N$. Due to the non-uniform distribution, prior work is unable to predict the correct variance while the first and third terms in our formulation are crucial in correcting for the non-uniformity (top half of Table 1).

Table 1: *Here we analyze all three terms from (8) using a 1D step function for two cases. Case I: Perfect importance sampling using random samples. Case II: Uniform sampling with jittered samples while keeping the step discontinuity aligned with one of the strata boundary for all N. Variance reduction due to IS or strata boundary alignment with a step is only captured by the third term. Numerically approximating the third term results in slight deviation from the expected variance of zero.*

	N	1st Term	2nd Term	3rd Term (real, imaginary)	Variance (empirical)	Variance (expected)
Case I	4	0.0178056	0.0175754	(-0.0353513, 0.0)	0.00002966	0.0
	16	0.0044613	0.0044005	(-0.0088477, 0.0)	0.00001420	0.0
	64	0.0010782	0.0010973	(-0.0021738, 0.0)	0.00000172	0.0
Case II	4	0.0	0.00519631	(-0.0051782, 0.0)	0.00001808017	0.0
	16	0.0	0.00032349	(-0.0003256, 0.0)	-0.00000214325	0.0
	64	0.0	0.00001980	(-0.0000182, 0.0)	0.00000152890	0.0

Correlations due to jittered sampling: Consider integrating a step function in 1D by sampling a constant PDF but using jittered samples. If the discontinuity in the integrand aligns perfectly with a stratum boundary, then the integral can be estimated with zero variance. Even though the samples are drawn according to a constant PDF, previous work is unable to predict the correct variance because the correlations between jittered samples and the integrand are lost due to homogenization. In our case, the third term cancels with the second term correctly predicting zero variance when a stratum is aligned with the discontinuity (bottom half of Table 1). These calculations require knowledge of the joint PDF, which can be easily derived for jittered samples in 1D [Ö16] (see also suppl. Sec. 6.1).

The role of phase: Fig. 6 visualizes the components of the third term in eq. 8 for a step integrand and jittered sampling pattern in 1D. Each image represents the 2D space spanned by frequency variables l (horizontal) and m (vertical), with colors indicating the value of the Fourier coefficients. To avoid numerical issues, we derived the

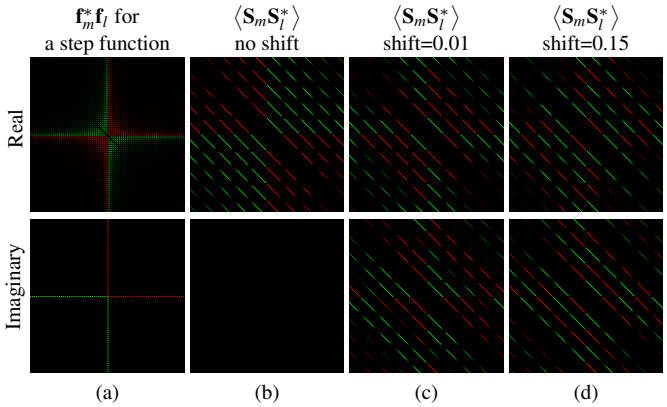
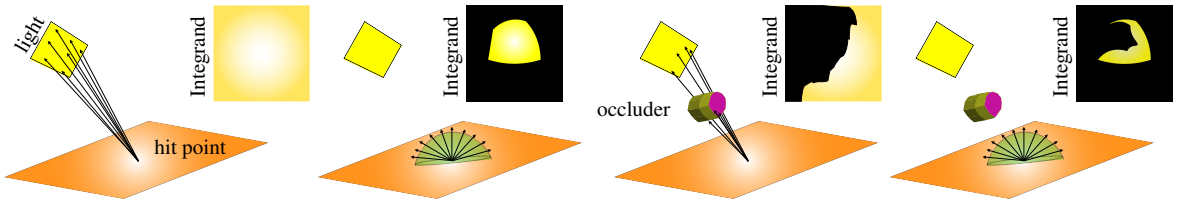


Figure 6: *To emphasize the role of phase in variance reduction, we visualize the third term from (8) for a 1D step function (**a**) and 1D jittered samples (for $N = 16$ in **b,c,d**) when all the strata are shifted as indicated above each column. Green, black, and red indicate positive, zero, and negative values, respectively.*



(a) Light IS (b) BSDF IS (c) Light IS with occluder (d) BSDF IS with occluder

Figure 7: *We illustrate the integrands (top-right in each column) for a shade point directly illuminated by an area light source for light IS and BSDF IS. (**a**) Light IS has a smooth underlying integrand when the light source is fully visible, unlike BSDF sampling (**b**), which sees light boundary as a C_0 discontinuity. Partially occluded shade points, on the other hand, always have a C_0 discontinuity (**c,d**).*

values analytically (see suppl. Sec. 6.4 & 7.1). Despite the third term being real (Table 1, proof in suppl. Sec. 4.1) the sub-terms within are complex. This necessitates a careful analysis of phase and highlights the challenges of non-shift-invariant Fourier analysis.

Higher dimensions: The space of $\langle \mathbf{S}_m \mathbf{S}_l^* \rangle$ is spanned by the frequency variables l and m. In our 1D case, this space is $\mathbb{C}^2$ which we visualized using 2D images for the real and complex components (Figure 6). For a d-dimensional integration domain, this space is $2d$-dimensional which then gets projected down to a single number. Even for a 2D domain, this means that correlations need to be analyzed in a 4D space.

4.3. Summary

In this section, we presented the first Fourier domain formulation of error of sampling-based integrators which accounts for the finite domains, non-uniformly distributed samples and correlated sampling. We expressed estimator variance as a sum of terms where one of them explicitly captures correlations between samples as well as alignment of the samples with the integrand. Our derivations suggest that the latter plays a key role in reducing error. Unfortunately, our analysis also reveals that Fourier analysis is not well suited to assess alignment since the correlations need to be tracked in a high ($2d$)-dimensional space first before projection along all dimensions.

5. Experiments

Our analysis suggests that the improved convergence of correlated samplers are limited by the presence of discontinuities. In this section, we first show that the use of importance functions can compensate for these discontinuities and can improve convergence rates of the resulting combined estimators. Although it is impractical to expect all discontinuities to be known a priori, some discontinuities may be avoided by changing the domain of integration (Sec. 5.1). To counter the effects of boundary discontinuities (discussed in Sec. 5.2), we propose a novel approach of *integrand mirroring* (Sec. 5.2.1). We further propose smooth boundary area light sources to avoid certain discontinuities (Sec. 5.2.2). We verified this in different situations by extending existing empirical tools [SSJ16]. All the scenes are rendered using PBRT [PJH16]. To facilitate the analysis, we restrict the dimensionality of the underlying integrands to 2D by only considering the directly illuminated components of all rendered scenes from one area light source. However, this analysis directly generalizes to higher dimensions without any loss of generality. To perform variance analysis, we vary the number of samples over the area light and the visible hemisphere while shooting only one primary ray from the center of each pixel.

5.1. Improving convergence rates using IS and CS

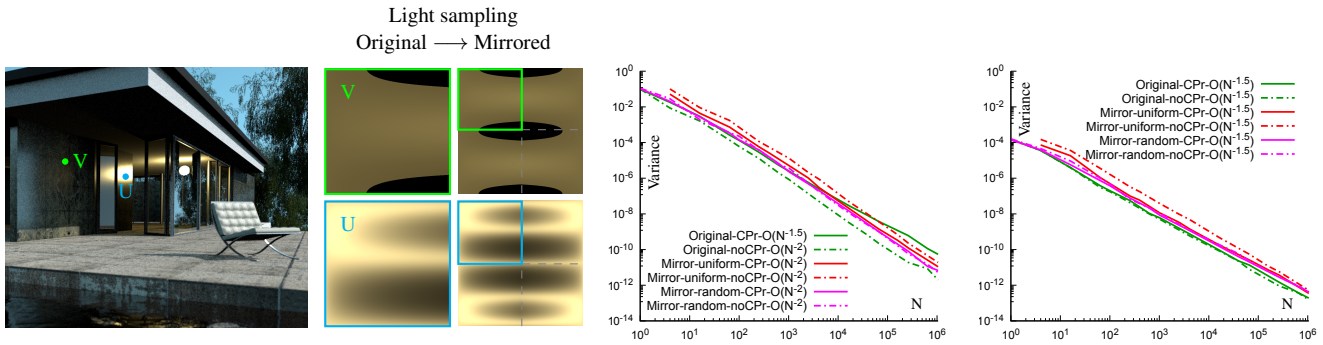
Consider the problem of estimating direct illumination (illustrated in Fig. 7) where sampling incident directions over the visible hemisphere manifests the light source boundaries as C_0 discontinuities (Fig. 7b) within the underlying integrand. These discontinuities, however, can be avoided by sampling the light sources instead (Fig. 7(a)). We analyze Pixel P which has a smooth underlying integrand (bottom-left Fig. 1b) when importance sampled using light PDF. This results in a convergence rate of $\mathcal{O}(N^{-2})$ (dash-dotted green curve in Fig. 1c). On the other hand, pixels that are partially occluded from a light source (Pixel Q), or which are importance sampled using the BSDF, always have a C_0 discontinuity (see second column Fig. 1b), which restricts their convergence rate to $\mathcal{O}(N^{-1.5})$ (Fig. 1b and d). Consequently, light importance sampling can also be seen as shifting the discontinuities within the domain to the boundaries of the integration domain resulting in a smooth *effective integrand* and an improved convergence rate. Our analysis provides a novel reasoning for why this classical solution improves the convergence of the estimator and not just the variance.

5.2. Boundary-value discontinuities due to finite domain

Our Fourier series formulation assumes that functions are periodic and that samples are generated toroidally in the domain. One of the consequences of the toroidal (mod) wrapping is that functions that appear smooth within the entirety of the domain might still introduce a "boundary discontinuity" because the values across the toroidally wrapped boundary do not coincide (first column in Fig. 9a). We performed experiments to confirm that such "invisible" discontinuities in otherwise smooth integrands adversely affect convergence.

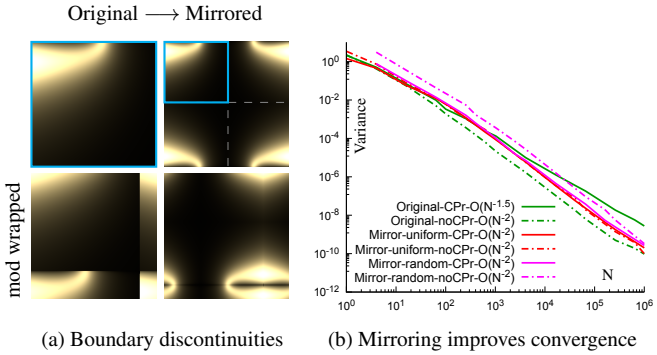
Toroidally wrapping the samples after homogenization or Cranley-Patterson rotation (CPr) introduces boundary discontinuities within the integrand (bottom-left Fig. 9a) due to the asymmetry across the boundary. This degrades convergence (see subplot b) from $\mathcal{O}(N^{-2})$ (dash-dotted curve) to $\mathcal{O}(N^{-1.5})$ (solid-green curve).

Incidentally, enforcing shift-invariance via random offsets [RAMN12, PSC*15] has a similar effect as toroidal wrapping



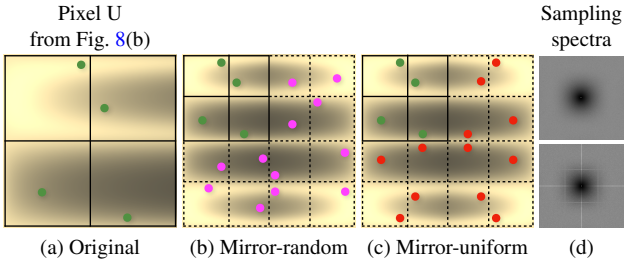
(a) Reference (b) Underlying pixel functions (c) Pixel U (d) Pixel V

Figure 8: *We perform integrand mirroring to avoid boundary discontinuities due to homogenization or Cranley-Patterson rotation (CPr). Two pixel functions in (**b**)—importance sampled from a spherical area light—are shown before and after mirroring. For Pixel U, convergence rate (shown in log-log scale) after mirroring improves to be $\mathcal{O}(N^{-2})$ due to symmetric boundary values despite homogenization. However, a discontinuous integrand (Pixel V) does not benefit from mirroring (**d**). The original underlying sampler is jittered.*



(a) Boundary discontinuities (b) Mirroring improves convergence

Figure 9: *Toroidally wrapping a sampling pattern results in introducing discontinuities within an otherwise smooth function (Pixel P from Fig. 1) as shown in bottom-left (**a**). However, this can be avoided by first mirroring the original integrand (top-right). Consequently, mod wrapping this mirrored version does not introduce discontinuities (bottom-right), and maintains the good convergence rate of $\mathcal{O}(N^{-2})$ (in log-log scale) despite homogenization (CPr). Note that the relative variance is still affected by the sampling strategy (details in Sec. 5.2.1). The underlying sampler used is jittered.*



(a) Original (b) Mirror-random (c) Mirror-uniform (d)

Figure 10: *After integrand mirroring (**a** → **b**,**c**), the integration domain quadruples (in 2D). (**b**) We keep the original (green) samples and equally stratify the rest of the domain to generate jittered samples (magenta). (**c**) It is also possible to simply mirror the (green) samples over the full domain (in red) along with the integrand. (**d**) This, however, introduces some high energy streaks in the expected power spectrum (bottom) which are not present for mirror-random samples (top).*

and therefore also suffers from degraded convergence when integrands have unequal values at wrapping boundaries. In Fig. 3(b), we observe that slightly shifting a 2D Gaussian integrand so that the values at the boundaries do not match causes a deterioration in the convergence rate from $\mathcal{O}(N^{-2})$ to $\mathcal{O}(N^{-1.5})$. We confirmed similar behavior for direct illumination estimators in real scenes (see Fig. 8 and 12). To counter the effect of boundary discontinuities we propose *integrand mirroring*.

5.2.1. Integrand mirroring

Since the boundary discontinuities are caused by the asymmetric values across the integration domain, we propose to mirror the integrand horizontally and vertically which results in symmetric values across the boundaries (top-right Fig. 9a and bottom-right Fig. 8b). In 2D, this process quadruples the integration domain, e.g. from $[0,1)^2$ to $[0,2)^2$.

To sample this domain, we adopt two simple strategies. An illustration is shown in Fig. 10 with (green) samples overlaid over the original integrand. After integrand mirroring, one approach is to span the stratification grid across the mirrored domain and generate random samples directly within each stratum (Fig. 10b). This is equivalent to having four random samples per stratum in the original integration domain. We call this the *mirror-random* sampling strategy. Another approach is to simply mirror the green samples in the original domain (in the spirit of antithetic samples) along with the integrand (Fig. 10b). In this case, the benefit is that no new samples need to be generated. We call this strategy *mirror-uniform* sampling. The corresponding power spectra show that *mirror-uniform* sampling imposes certain regularity which is visible as high energy streaks in the resultant expected power spectrum.

To see the benefits of integrand mirroring, we consider Pixel U from Fig. 8 which has a smooth underlying integrand with asymmetric boundary values. After mirroring, we observe a convergence rate

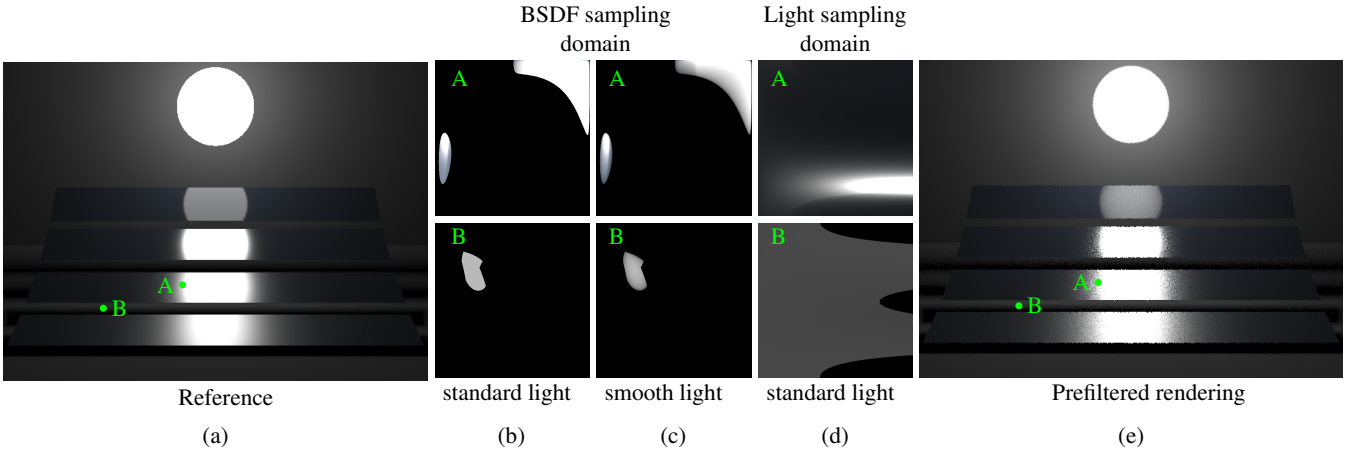
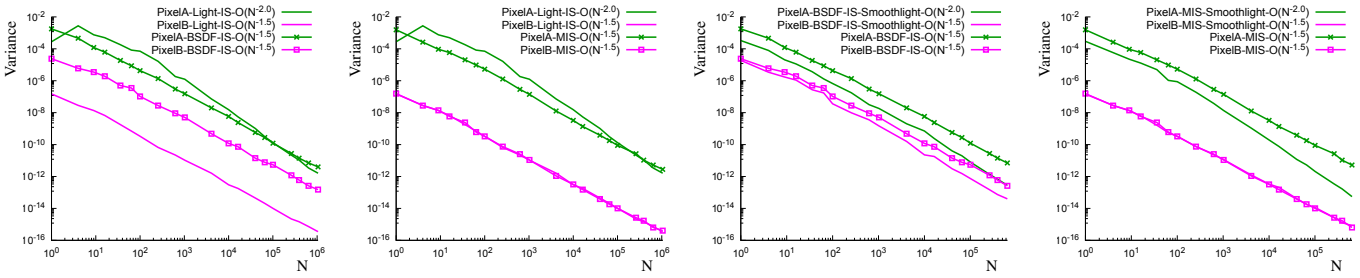


Figure 11: *Underlying pixel functions for A & B are shown in **b, c & d** for a directly illuminated scene in **a**. The comparison is drawn between standard spherical lighting vs. smoothed boundary spherical light source (prefilitered). This smoothing is performed by scaling the radiance using the dot product of the ray direction and the light source normal. Full rendering is shown in (e).*



(a) Light vs BSDF (standard) (b) Light vs MIS (standard) (c) BSDF (standard vs smooth light) (d) MIS (standard vs smooth light)

Figure 12: *Convergence rates for Pixel A (unoccluded) & B (partially occluded) from Fig. 11 are plotted (in log-log scale) for jittered samples. Light IS has faster convergence (**a,b**) compared to BSDF or MIS for Pixel A. MIS and BSDF IS convergence can be improved by light boundary smoothing (**c,d**).*

of $\mathcal{O}(N^{-2})$ with and without homogenization (CPr) with both mirror sampling strategies (red and magenta curves in Fig. 9b and 8c). However, with *mirror-random* the relative variance is slightly higher in Fig. 9b (magenta curves) whereas with *mirror-uniform* variance is higher in Fig. 8c. Integrands with C_0 discontinuities (Fig. 8, Pixel V) do not benefit from integrand mirroring as shown in Fig. 8d.

5.2.2. Smooth boundary light source

The above analyses show benefits of *integrand mirroring* with light IS. The convergence behavior due to BSDF IS does not improve with mirroring since it sees the finite area light source as a C_0 discontinuity within the underlying integrand (Fig. 7b). This also affects the convergence rate during multiple importance sampling (MIS). To validate this, we extend Veach's [Vea98] analysis of light vs. BSDF IS by investigating convergence rates rather than variance.

The scene in Fig. 11(a) is rendered with light IS by first shooting one primary ray from the center of the pixel followed by 1024 jittered light (or shadow) ray samples. Pixel A is unoccluded and is directly illuminated wheras Pixel B is occluded by one of the metal plates in the scene. As shown in Fig. 12(a) and (b), light PDF sampling would converge at a rate of $\mathcal{O}(N^{-2})$ compared to BSDF sampling (shown with dotted lines) which converges at $\mathcal{O}(N^{-1.5})$ but only for pixels with unoccluded light rays. However, for occluded Pixel B, the convergence rate is always $\mathcal{O}(N^{-1.5})$. In Fig. 11(b) we visualize this behavior for individual pixels for both BSDF and light sampling.

We also compare light PDF sampling with multiple importance sampling (MIS) using a power heuristic in Fig. 12(b). For both pixels A and B, MIS converges with $\mathcal{O}(N^{-1.5})$. This validates that MIS can be seen as a combination of both light and BSDF sampling convergence rates, with the expected convergence to be the worst of the two [SJ17]. Our analysis also identifies correctly that *the convergence rate exhibited by area light sampling would be at least as good as that achieved by BSDF sampling.*

The above reasoning can also be applied to distant illumination from environment maps. The lack of discontinuities can only be exploited at points which receive unoccluded illumination, since otherwise the discontinuities due to occlusion would dominate. Unoccluded points will enjoy convergence benefits due to light PDF sampling. However, in practice, since rendering software such as

PBRT [PJH16] use the luminance channel as the importance function, the effective integrand in each of the color channels may not entirely be free of discontinuities. So, for a multichannel environment map, importance sampling each channel will improve convergence at unoccluded pixels.

Improving MIS & BSDF IS: Based on our convergence analysis we observe that discontinuities can be introduced or removed by importance sampling. However, MIS still suffers the worst of the two convergence rates since the BSDF sampling strategy always treats an area light source as a discontinuous binary function. We therefore analyze the convergence by smoothing the light boundary to improve upon this bottleneck. To demonstrate our hypothesis, we smooth an area light source (which is a common practical trick for variance reduction) by multiplying the outgoing power by the dot product of the ray direction and the normal of the spherical area light source. The resulting directly illuminated scene is shown in Fig. 11(c) which is rendered with 16 jittered samples per pixel. The corresponding convergence plots in Fig. 12(c and d) shows improved convergence in the case of both BSDF sampling and MIS for unoccluded hit points at the cost of introducing negligible bias.

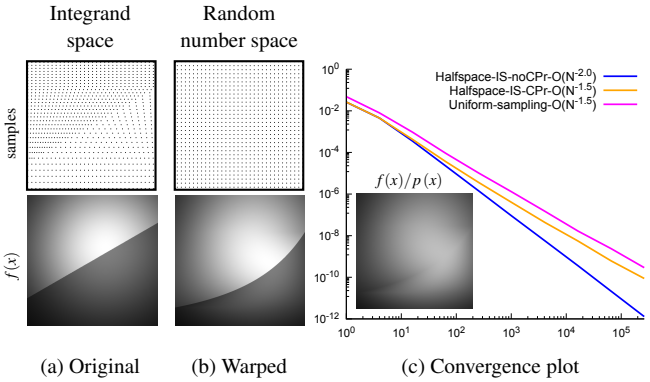


Figure 13: *$f(x)$ is a product of an asymmetric Gaussian and a Halfspace function. Importance sampling the Halfspace shows improvement in convergence rate (blue plot in **c**) as discontinuities are importance sampled, resulting in a smooth effective integrand (inset in **c**). Top row (in **a**) shows importance samples drawn from an Halfspace PDF, which when warped back to the random number space (**b**) modifies the integrand accordingly.*

6. Conclusion and Future work

We presented a complete formulation for error analysis of sampling-based integrators in the Fourier domain. Our formulation elegantly handles finite domains, non-uniform sampling as well as correlations between samples. We showed using rigorous derivations as well as empirical evidence that our formulation predicts variance correctly in many scenarios where existing theory fails. We also identified the role of discontinuities in integrands in adversely affecting the convergence rates of sampling-based estimators. We showed that the convergence rate of the combined estimator (importance sampling and correlated sampling) can be controlled by changing the importance function alone. Our analysis highlights the importance of the integrand discontinuities and how these relate to the sampling PDF and stratification boundaries. Using these insights, we showed how to preserve the improved convergence rate of correlated sampling when using MIS for direct lighting by smoothing out the boundary of the light source. To counter boundary discontinuities, we propose *integrand mirroring* and tailored two simple mirror sampling strategies to sample this mirrored domain. Since the idea of mirroring is inspired from antithetic sampling [HM56a, HM56b], it would be an interesting future direction to study the effect of leveraging antithetic samples for these strategies.

Importance sampling the discontinuities: Our analyses and experimental results suggest several potential directions for future research. Our results challenge the conventional view of designing the importance function to merely be proportional to the value of the integrand, especially when perfect IS is not possible (as is mostly the case). Instead, it suggests that it may sometimes be better to choose the importance function (in conjunction with correlated samples) so that discontinuities that appear in the correlation structure align with the integrand. The analysis by Oztireli [Ö16], Singh and Jarosz [SJ17] already indicate benefits due to alignment but for uniform samples.

It is, however, also possible to importance sample the discontinuities present within the integrand (not shown in our proposed analysis). We show a simple setup in Fig. 13 where we estimate the product of a shifted 2D Gaussian and a halfspace (2D step edge) function (Fig. 13a) using jittered samples distributed according to the halfspace function as the importance function (Fig. 13b shows the corners of the strata). Importance sampling the non-axis-aligned halfspace function induces a mapping between the unit square where random numbers are drawn, which we call the *random number space* and the halfspace as viewed in Fig. 13a, which we call the *integrand space*. The integrand evaluated in the random number space is shown in Fig. 13b. It appears warped due to the mapping.

The combined (IS and jittered) estimator exhibits a convergence rate of $\mathcal{O}(N^{-2})$ (blue curve in Fig. 13c) compared to $\mathcal{O}(N^{-1.5})$ without IS (magenta curve) or with IS but also with homogenization (orange curve). The improvement in convergence rate must be due to the diminished role of the discontinuity in the effective integrand $f(x)/p(x)$ (inset in Fig. 13c). Although the effective integrand is the same for the estimator with homogenization, its convergence rate suffers due to the boundary discontinuity. In most practical situations, it is unreasonable to expect the discontinuities to be known a priori, therefore, developing this idea requires further investigation.

Multiple-lights sampling: Our proposed analysis is performed for only one light source. For many-light scenarios, current sampling strategies uses one sample coordinate to randomly select the light source and uniformly distribute the samples among different light sources without increasing the dimensionality of the problem. Although simple, this strategy destroys the underlying correlations within the samples (e.g., stratification) resulting in slower convergence. It would be interest to extend our analysis and design sampling strategies for multiple light sources.

Integrating a pulse train: The covariance (third) term in our variance formulation suggests to preserve the phase for variance/convergence improvements which can be interpreted as careful

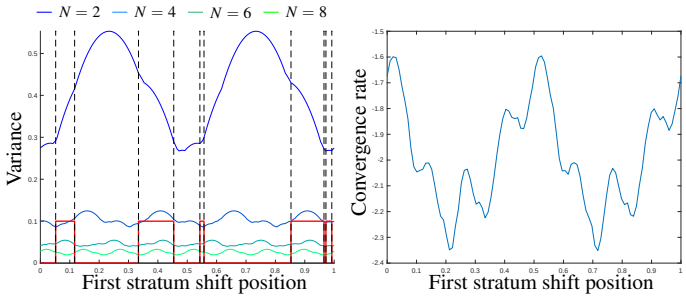


Figure 14: *With jittered samples, variance and convergence rate for a train of pulse (in red) depends on the location of the strata boundary w.r.t. the sample location. In the plots, the sampling strata are shifted along the horizontal axis by a constant $u \in [0,1]$. The dashed vertical lines represent the C_0 discontinuity locations.*

shifting of the strata boundaries w.r.t. the discontinuities. We demonstrate a simple setup of train of pulses (red curve shown in Fig. 14) composed of shifted Heavisides $H(x-u)$. Shifting the position of the first stratum boundary by an offset $\phi \in [0,1]$ is equivalent to shifting the integrand to $H(x-u+\phi)$. Thus, as the offset ϕ is varied, for a given N, the continuously changing variance of the corresponding estimator for a Heaviside function is: Var $\left(\mu_{js,N,\phi}\right) = v_\phi(1-v_\phi)/N^2$, where $v_\phi = N(u-\phi) - \lfloor N(u-\phi) \rfloor$. The variance of the N-sample jittered sampling estimator integrating a pulse-train simply amounts to adding the Var $\left(\mu_{js,N}\right)$ corresponding to each of the Heavisides. We calculated these variances analytically and plotted them in Fig. 14. The plot on the left is of predicted variance vs ϕ for different values of N and the plot on the right shows the best-fit convergence rates obtained for different values of ϕ. For any given N, the minimum variance may not necessarily correspond to a ϕ where a discontinuity is aligned with the strata boundary. Similarly, the best possible convergence rate of $O(N^{-2.35})$ is obtained when the strata are offset by $\phi = 0.21$ or $\phi = 0.71$. This simple experiment demonstrates that both error and convergence rate might be controlled by tailoring the relative offsets between sampling patterns and discontinuities in the integrand. Any arbitrarily discontinuous function can be analyzed as a pulse train multiplied by a smooth function.

One strategy could be to *optimize* the position of the strata boundaries in jittered sampling w.r.t. the integrand discontinuities (third term in (8)). While this could potentially improve variance and convergence rate (Fig. 14), in practice this would require knowledge of all discontinuities, which may be intractable, especially in higher dimensions. We instead showed that importance sampling according to different PDFs can move *some* discontinuities to the boundary of the integration domain, where their presence does not impact convergence rate. Simultaneously, we can pre-filter the discontinuities for other sampling strategies, preserving good convergence rates in the MIS combination for pixels where only these discontinuities are present in the integrand. While we show improvements only due to pre-filtered light source boundary, more general pre-filtering of integrand discontinuities could be a fruitful avenue of future research to leverage improved convergence rates for global illumination.

7. Acknowledgements

We are grateful to all the anonymous reviewers for their constructive remarks. This work was partially supported by the Fraunhofer and Max Planck cooperation program within the German pact for research and innovation (PFI) and NSF grant CNS-1205521. Kartic Subr was supported by a Royal Society University Research Fellowship and Wojciech Jarosz was partially supported by NSF grant ISS-1812796.

References

[AHD15] AHMED A. G. M., HUANG H., DEUSSEN O.: AA patterns for point sets with controlled spectral properties. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 34*, 6 (Oct. 2015), 212:1–212:8. 2

[ARBJ03] AGARWAL S., RAMAMOORTHI R., BELONGIE S., JENSEN H. W.: Structured importance sampling of environment maps. *ACM Transactions on Graphics 22*, 3 (July 2003), 605–612. 3

[BSD09] BALZER M., SCHLÖMER T., DEUSSEN O.: Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Trans. Graph. (Proc. SIGGRAPH) 28*, 3 (2009). 2

[CAM08] CLARBERG P., AKENINE-MOELLER T.: Practical product importance sampling for direct illumination. *Comp. Graph. Forum (Proc. Eurographics)* (2008). 3

[CJAMJ05] CLARBERG P., JAROSZ W., AKENINE-MÖLLER T., JENSEN H. W.: Wavelet importance sampling: Efficiently evaluating products of complex functions. *ACM Trans. Graph. 24*, 3 (July 2005), 1166–1175. 3

[Coc63] COCHRAN W. G.: *Sampling Techniques, 2nd Edition.* John Wiley, 1963. 2

[Coo86] COOK R. L.: Stochastic sampling in computer graphics. *ACM Transactions on Graphics 5*, 1 (Jan. 1986). 2, 3

[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. 1, 3

[CSHD03] COHEN M., SHADE J., HILLER S., DEUSSEN O.: Wang tiles for image and texture generation. *ACM Trans. Graph. (Proc. SIGGRAPH) 22*, 3 (2003). 2

[dGBOD12] DE GOES F., BREEDEN K., OSTROMOUKHOV V., DESBRUN M.: Blue noise through optimal transport. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 31* (2012). 2

[Dur11] DURAND F.: *A frequency analysis of Monte-Carlo and other numerical integration schemes*. Tech. Rep. TR-2011-052, MIT CSAIL, 2011. 3, 4

[DW85] DIPPÉ M. A. Z., WOLD E. H.: Antialiasing through stochastic sampling. *Computer Graphics (Proc. SIGGRAPH)* (1985). 3

[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. 2

[GKH*13] GEORGIEV I., KŘIVÁNEK J., HACHISUKA T., NOWROUZEZAHRAI D., JAROSZ W.: Joint importance sampling of low-order volumetric scattering. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 32*, 6 (Nov. 2013). 3

[Hab70] HABER S.: Numerical evaluation of multiple integrals. *SIAM Review 12*, 4 (1970), 481–526. 2

[Hd14] HEITZ E., D'EON E.: Importance sampling microfacet-based BSDFs using the distribution of visible normals. *Comp. Graph. Forum (Proc. EGSR) 33*, 4 (July 2014), 103–112. 3

[Hes03] HESTERBERG T.: *Advances in Importance Sampling*. PhD thesis, Stanford, CA, USA, 2003. 3

[HM56a] HAMMERSLEY J. M., MAULDON J. G.: General principles of antithetic variates. *Mathematical Proceedings of the Cambridge Philosophical Society 52*, 3 (1956), 476–481. 10

[HM56b] HAMMERSLEY J. M., MORTON K. W.: A new Monte Carlo technique: antithetic variates. *Mathematical Proceedings of the Cambridge Philosophical Society 52*, 3 (1956), 449–475. 10

[Hör90] HÖRMANDER L.: *The analysis of linear partial differential operators: Distribution theory and Fourier analysis*. Springer Study Edition. Springer-Verlag, 1990. 3

[IPSS08] ILLIAN J., PENTTINEN P. A., STOYAN H., STOYAN D.: *Statistical analysis and modeling of spatial point patterns*. Statistics in Practice. Wiley, 2008. 2, 3, 6

[JCJ09] JAROSZ W., CARR N. A., JENSEN H. W.: Importance sampling spherical harmonics. *Comp. Graph. Forum (Proc. Eurographics) 28*, 2 (Apr. 2009), 577–586. 3

[KC08] KŘIVÁNEK J., COLBERT M.: Real-time shading with filtered importance sampling. In *Proceedings of the Nineteenth Eurographics Conference on Rendering* (Aire-la-Ville, Switzerland, Switzerland, 2008), Comp. Graph. Forum (Proc. EGSR), Eurographics Association, pp. 1147–1154. 3

[KCODL06] KOPF J., COHEN-OR D., DEUSSEN O., LISCHINSKI D.: Recursive wang tiles for real-time blue noise. *ACM Trans. Graph. (Proc. SIGGRAPH) 25*, 3 (2006). 2

[KPR12] KELLER A., PREMOZE S., RAAB M.: Advanced (quasi) Monte Carlo methods for image synthesis. In *SIGGRAPH '12 Courses* (New York, USA, 2012), ACM. 3

[KSP*15] KAZHDAN M., SINGH G., PILLEBOUE A., COEURJOLLY D., OSTROMOUKHOV V.: Variance analysis for Monte Carlo integration: A representation-theoretic perspective. *ArXiv e-prints* (May 2015). 4

[KTBV16] KAILKHURA B., THIAGARAJAN J. J., BREMER P.-T., VARSHNEY P. K.: Stair blue noise sampling. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 35*, 6 (Nov. 2016), 248:1–248:10. 2

[LD08] LAGAE A., DUTRÉ P.: A comparison of methods for generating Poisson disk distributions. *Comp. Graph. Forum 27*, 1 (2008). 2

[Lem09] LEMIEUX C.: *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Series in Statistics. Springer New York, 2009. 3

[Les17] LESSIG C.: Controlling and Sampling Visibility Information on the Image Plane. In *Eurographics Symposium on Rendering: Experimental Ideas & Implementations* (2017), Zwicker M., Sander P., (Eds.), The Eurographics Association. 3

[Mit91] MITCHELL D. P.: Spectrally optimal sampling for distributed ray tracing. *Computer Graphics (Proc. SIGGRAPH) 25*, 4 (1991). 2, 3

[Mit96] MITCHELL D. P.: Consequences of stratified sampling in graphics. In *Annual Conference Series (Proc. SIGGRAPH)* (1996), ACM. 3

[Ney34] NEYMAN J.: On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society 97*, 4 (1934), 558–625. 2

[Nie78] NIEDERREITER H.: Quasi-Monte Carlo methods and pseudo-random numbers. *Bull. Amer. Math. Soc. 84*, 6 (11 1978), 957–1041. 3

[Ö16] ÖZTIRELI A. C.: Integration with stochastic point processes. *ACM Transactions on Graphics 35*, 5 (Aug. 2016), 160:1–160:16. 2, 3, 6, 10

[OF01] OUELLETTE M. J., FIUME E.: On numerical solutions to one-dimensional integration problems with applications to linear light sources. *ACM Transactions on Graphics 20*, 4 (Oct. 2001), 232–279. 2

[OG12] ÖZTIRELI A. C., GROSS M.: Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 31*, 6 (2012). 2, 6

[Owe13] OWEN A. B.: *Monte Carlo theory, methods and examples*. https://statweb.stanford.edu/ owen/mc/ (accessed April 09, 2018). 2013. 3, 5

[OZ00] OWEN A., ZHOU Y.: Safe and effective importance sampling. *Journal of the American Statistical Association 95*, 449 (2000), 135–143. 3

[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. 3, 7, 10

[PSC*15] PILLEBOUE A., SINGH G., COEURJOLLY D., KAZHDAN M., OSTROMOUKHOV V.: Variance analysis for Monte Carlo integration. *ACM Trans. Graph. (Proc. SIGGRAPH) 34*, 4 (July 2015). 3, 4, 5, 6, 7

[RAMN12] RAMAMOORTHI R., ANDERSON J., MEYER M., NOWROUZEZAHRAI D.: A theory of Monte Carlo visibility sampling. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 5 (2012). 2, 3, 6, 7

[Rip77] RIPLEY B. D.: Modelling spatial patterns. *Journal of the Royal Statistical Society. Series B (Methodological) 39*, 2 (1977), 172–212. 3

[SA07] SUBR K., ARVO J.: Statistical hypothesis testing for assessing Monte Carlo estimators: Applications to image synthesis. *Proc. of Pacific Graphics* (Oct 2007), 106–115. 3

[Shi91] SHIRLEY P. S.: Discrepancy as a quality measure for sample distributions. In *Proc. Eurographics* (1991). 2, 3

[SJ94] SLOAN I., JOE S.: *Lattice methods for multiple integration*. Oxford science publications. Clarendon Press, 1994. 2, 3

[SJ17] SINGH G., JAROSZ W.: Convergence analysis for anisotropic Monte Carlo sampling spectra. *ACM Trans. Graph. (Proc. SIGGRAPH) 36*, 4 (July 2017), 137:1–137:14. 3, 9, 10

[SK13] SUBR K., KAUTZ J.: Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. Graph. (Proc. SIGGRAPH) 32*, 4 (July 2013). 3, 4, 5

[SMJ17] SINGH G., MILLER B., JAROSZ W.: Variance and convergence analysis of Monte Carlo line and segment sampling. *egsr08 36*, 4 (2017), 79–89. 3

[SSJ16] SUBR K., SINGH G., JAROSZ W.: Fourier analysis of numerical integration in Monte Carlo rendering: Theory and practice. In *ACM SIGGRAPH 2016 Courses* (New York, NY, USA, 2016), ACM, pp. 10:1–10:40. 7

[Uli87] ULICHNEY R.: *Digital Halftoning*. MIT Press, Cambridge, MA, USA, 1987. 2

[Vea98] VEACH E.: *Robust Monte Carlo methods for light transport simulation*. PhD thesis, Stanford, CA, USA, 1998. AAI9837162. 2, 3, 9

[WPC*14] WACHTEL F., PILLEBOUE A., COEURJOLLY D., BREEDEN K., SINGH G., CATHELIN G., DE GOES F., DESBRUN M., OSTROMOUKHOV V.: Fast tile-based adaptive sampling with user-specified Fourier spectra. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (2014). 2

[WW11] WEI L.-Y., WANG R.: Differential domain analysis for non-uniform sampling. *ACM Trans. Graph. 30*, 4 (July 2011). 2

[Zar72] ZAREMBA S.: *Applications de la théorie des nombres à l'analyse numérique:*. Academic Press Rapid Manuscript Reproduction. Academic Press, 1972. 2