

Fourier Analysis of Stochastic Sampling Strategies for Assessing Bias and Variance in Integration

Kartic Subr*

University College London

Jan Kautz†

University College London

Abstract

Each pixel in a photorealistic, computer generated picture is calculated by approximately integrating all the light arriving at the pixel, from the virtual scene. A common strategy to calculate these high-dimensional integrals is to average the estimates at stochastically sampled locations. The strategy with which the sampled locations are chosen is of utmost importance in deciding the quality of the approximation, and hence rendered image.

We derive connections between the spectral properties of stochastic sampling patterns and the first and second order statistics of estimates of integration using the samples. Our equations provide insight into the assessment of stochastic sampling strategies for integration. We show that the amplitude of the expected Fourier spectrum of sampling patterns is a useful indicator of the bias when used in numerical integration. We deduce that estimator variance is directly dependent on the variance of the sampling spectrum over multiple realizations of the sampling pattern. We then analyse Gaussian jittered sampling, a simple variant of jittered sampling, that allows a smooth trade-off of bias for variance in uniform (regular grid) sampling. We verify our predictions using spectral measurement, quantitative integration experiments and qualitative comparisons of rendered images.

Keywords: stochastic sampling, Fourier analysis, Monte Carlo sampling

1 Introduction

Image synthesis requires the integration of multi-dimensional signals. This is commonly achieved by averaging the values of the signal, stochastically sampled at a number of discrete locations. The quality of the final output critically depends on the location and weights of these samples. The sampling strategy affects the accuracy and precision of the estimates for the integral

Designing effective sampling strategies for integration is crucial for efficient image synthesis. It is desirable to analyze stochastic sampling patterns directly, in order to judge their suitability for numerical integration; i.e., a sampling pattern should lead to accurate, low-variance estimates with as few samples as possible. Unfortunately their analysis is challenging, since the efficacy of a sampling pattern depends on an intricate combination of three factors: the distribution of the sample locations; the weighting used to accumulate sampled contributions; and the arrangement and structure of samples.

For example, importance sampling approaches carefully tailor the distribution and weights to reduce variance of the resulting estimates. Stratified sampling [Neyman 1934; Keller 2002; Ostromoukhov 2007] and Poisson-disk sampling [Matérn 1960; Matérn 1986; Cook 1986; Lagae and Dutré 2008] are examples of approaches that reduce variance by enforcing structural constraints (partitioning and minimum-radius resp.) on the sampling pattern.

In this paper, we abstract away these choices and focus on the properties of the Fourier spectrum of the sampling function. We derive expressions for the bias and variance of the integration scheme as a function of the Fourier spectrum of the sampling pattern. We demonstrate the utility of our theory by analyzing a minor modification of the standard jittered-sampling algorithm.

Much work has been done on the spectral analysis of sampling patterns for reducing the effects of aliasing in reconstruction [Dippe and Wold 1985; Mitchell 1987; Mitchell and Netravali 1988]. Cook argued that visual artifacts in image synthesis were not a result of point sampling [1986], rather, that they were results of the sampling being regular. Many approaches have been proposed [Hachisuka et al. 2008; Egan et al. 2009; Soler et al. 2009; Lehtinen et al. 2012] to adapt the sampling distributions and reconstruction filters to the underlying continuous signal. Much of the work in image reconstruction relies on each sample, of the continuous signal, containing no error. However, each evaluation of the signal requires a multi-dimensional integration (time, lens, reflectance, lighting, occlusion, etc.) that is typically realized through stochastic sampling. Recent work [Ramamoorthi et al. 2012] analyses errors in Monte Carlo visibility sampling, in specific contexts, and provides insight using a Fourier analysis of integration. In this paper we analyze the error of a general integrator using stochastic sampling, *over multiple realizations of sampling patterns* drawn using the same strategy.

We distinguish between the manifestation of error in the form of bias and variance, and express each in terms of the frequency spectra of the sampling function and integrand. We express the bias of the integration scheme in terms of the expected Fourier spectrum of the sampling signal, over multiple realizations of the sampling pattern, and the Fourier spectrum of the integrand. We derive that the variance in the integration depends on the variance of the sampling spectrum modulated by the power spectral density of the integrand. Our equations reveal that the commonly used *periodogram is not ideally suited to analyze sampling patterns* if the goal is to assess the bias and variance when the samples are used in integration. Instead, we show that the amplitude of the expected sampling spectrum provides direct insight into predicting bias.

We apply our theory to analyse a simple sampling strategy: Jittering each sample on a regular grid using a Normal distribution. As our analysis shows, this simple variant of jittered sampling, which easily generalizes to higher dimensions, is effective and can be used to smoothly transition from a regular sampling pattern (yields a biased estimate with zero variance) to one that mitigates bias at the cost of adding variance.

1.1 Related work

Assessing sampling patterns Many tools exist to evaluate the samples themselves. For example, statistics such as Ripley's K and L statistics [Ripley 1977] are commonly used to model point distributions, while tools such as the variogram and autocorrelation are used for assessing distributions. In computer graphics, methods such as frequency domain analysis [Dippe and Wold 1985; Cook 1986; Mitchell 1991], a measure of spatial discrepancy [Shirley 1991], statistical tests of hypothesis [Subr and Arvo 2007], and point correlation [Öztireli and Gross 2012] have been

*e-mail: kartic@gmail.com

†e-mail: j.kautz@ucl.ac.uk

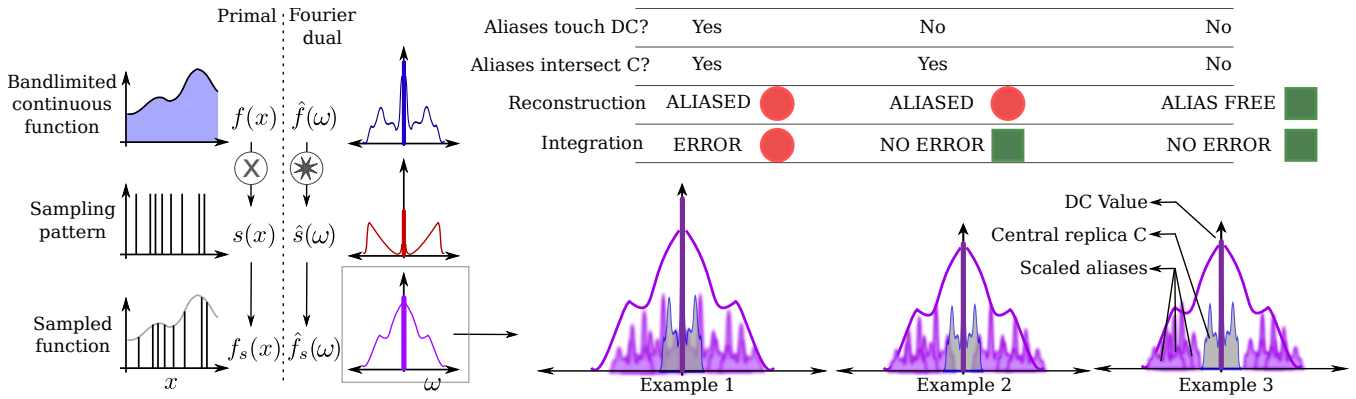


Figure 1: Illustration of errors in reconstruction and integration introduced by the sampling spectrum. Only the amplitude spectra are shown, for illustration. The sampled function f_s is the product of the bandlimited function f and the sampling pattern s . In the Fourier domain, the spectrum of f_s is obtained by convolving the spectra (blue and red) of f and s . If the sampling pattern has low enough frequencies so that the aliased copies of \hat{f} (blurry magenta) pollute the DC ($\hat{f}_s(0)$ in example 1), they introduce an error in the numerical integration [Durand 2011]. For reconstruction, the aliased copies need to be pushed even more (factor of 2) to the higher frequencies, or they will intersect the central replica of \hat{f} and produce an aliased reconstruction (example 2). In example 3, \hat{s} contains energy in frequencies greater than the Nyquist rate. Phase plays an important role: that is, a shift in the sampling pattern (primal) may change the error dramatically.

proposed. Recently, Wei and Wang [2011] proposed a reformulation of Fourier domain analysis for analyzing adaptive sampling [Wei and Wang 2011]. Bowers et al [2010] proposed a novel analysis of samples on surfaces, by projecting them onto the Eigenvectors of the mesh Laplacian. However, we conjecture that the measure of quality of a sampling strategy is intertwined with its application. For example, good quality might correspond to aesthetic appeal for dithering and stippling, estimators with low variance and bias for use in integration, various error metrics when used for signal reconstruction or robustness in physical simulations [Fienup 1997; Machiraju et al. 1995]. In contrast, we focus on assessing the bias and variance of integrators using stochastic sampling.

Low discrepancy patterns Shirley [1991] introduced a measure of spatial discrepancy to assess the quality of a sample set. Mitchell [1992] then proposed sampling strategies with the goal of minimizing this measure idea by trading randomness for low-discrepancy. Multidimensional integration using quasi-random patterns converge more rapidly [Kollig and Keller 2002], trading bias for variance. Much of the work in Quasi-Monte Carlo integration in computer graphics can be seen as specific instances of classes discussed in a seminal book by Niederreiter [Niederreiter 1992]. In a connection with Poisson-disk sampling, patterns that expose a large minimum toroidal distance [Larcher and Pillichshammer 2001] are useful for quasi-random Monte Carlo integration in computer graphics. Sampling using low-discrepancy sequences typically yield low-variance integration estimates. However, their regularity causes reconstructions to be more susceptible to aliasing [Pharr and Humphreys 2010]. This is typically avoided by combining regular and irregular sampling [Keller and Heidrich 2001; Kollig and Keller 2002]. We study a variant of jittered sampling which can be viewed as such a combination.

Spectral analysis in rendering Mitchell analysed the problem of reconstructing images [Mitchell 1991] in the frequency domain. In accordance with his thesis, that sampling patterns containing higher frequencies tended to reduce perceptible artifacts, he proposed a scanning sampling algorithm to optimally sample higher dimensions for application in distributed ray tracing. Ramamoorthi and Hanrahan [2004] studied reflection of light and expressed it as a bandlimiting operation in the frequency domain. Durand et al [Durand et al. 2005] derived transport operators in the Fourier domain for emission, transmission, reflection and occlusion. Further work led to efficient algorithms for sampling the lens aperture [Soler

et al. 2009] as well as time [Egan et al. 2009] domains. Belcour et al [2012] present an algorithm to propagate approximate spectral information of the 5D local light field. These approaches together provide approximate information about the Fourier spectrum of the integrand. In this paper, we study the impact of the spectra of the integrand as well as the sampling patterns on the quality of numerical integration.

Spectral analysis of sampling patterns The distribution of energy of a signal in the different frequencies has been shown to be an important consideration for designing sampling strategies for reconstruction [Mitchell 1987; Mitchell and Netravali 1988; Dippe and Wold 1985], digital half-toning and color separation [Amidror et al. 1994] as well as integration. The periodogram has become a standard tool, and is available in point sampling libraries [Schlömer and Deussen 2011] for analyzing non-adaptive sampling patterns spectrally. Wei and Wang [2011] proposed a Fourier domain analysis for analyzing adaptive sampling. The goal of their formulation is to enable comparison of quality, of the arrangement of points, across non-uniformly sampled sets. In contrast, our analysis predicts the quality of numerical integration using the samples.

Spectral analysis of sampling for integration It is well known that sampling below the Nyquist rate is prone to aliased signal reconstruction of bandlimited signals. For integration of bandlimited signals, spectral aliases introduce error in the estimates when the sampling is below half this rate [Křivánek and Colbert 2008; Durand 2011; Belcour et al. 2012]. Figure 1 illustrates the difference between errors in reconstruction and integration. We study this error introduced by the spectral aliases of the sampling pattern. Our analysis is inspired by approaches that analyze formulae used in numerical integration [Luchini 1994; Durand 2011] and recent work by Ramamoorthi et al. [2012], which assesses errors due to various sampling strategies for estimating visibility. However, our approach differs from related work in three important ways: we consider phase information, we assess secondary estimators and we derive explicit equations for bias and variance in terms of the sampling spectra.

Jittered sampling In standard *jittered sampling*, the center of each cell of a regular grid is uniformly randomly perturbed, a.k.a. *jittered*, within the cell. This produces a uniformly distributed set of samples (stratified sampling with proportional allocation) whose merits and demerits are well understood [Mitchell 1996; Arvo

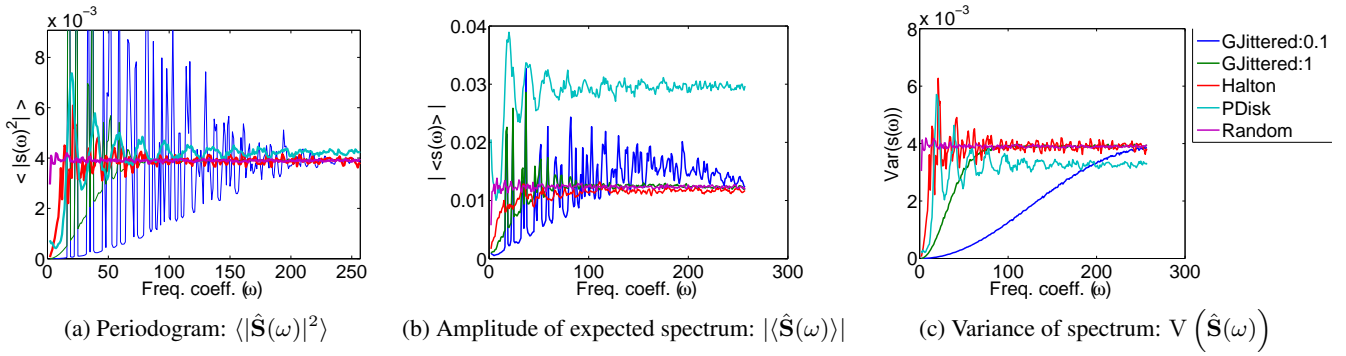


Figure 2: A comparison of three spectral properties of multiple sampling strategies. The periodogram is computed by averaging power spectra of multiple realizations of the sampling pattern $\mathbf{S}(x)$. The amplitude of the expected spectrum, on the other hand, first averages the realizations of the spectra (including phase) and then computes its amplitude. The variance of the sampling spectrum measures the uncertainty in the spectrum, over the realizations, at each frequency. (b) is more informative than (a)

2001]. In Monte Carlo path tracing, stratified sampling provides image antialiasing, at no extra cost, while radiance estimates are integrated in path space [Gallaher 1973; Cook 1986; Mitchell 1987]. Jittered sampling is simple, extends naturally to higher dimensions, and is straightforward to parallelize. Standard jittered sampling yields a radially anisotropic distribution of samples around centers of the grid cells. We study a variant of jittered sampling in which the jitter is radially symmetric. We derive equations for the bias and variance of this sampling scheme.

Signal processing We are inspired by work in signal processing that involve point processes [Bartlett 1964]. Spectral characteristics of trains of randomly occurring delta functions with randomly varying intensities, or impulse processes [Leneman 1966], were used to study communication systems with discrete pulse modulated controls. We seek to study the spectral profile of sampling signals, as Brémaud et al. [2003] did for common point processes and shot noise. There is a key difference. Much of the digital signal processing literature analyses a *sampled* signal, trying to recover the underlying continuous signal. In this paper, our goal is to design and assess a sampling signal for numerical integration of continuous signals which can be sampled at will at chosen locations. This can cause confusion in terminology. For example, Cook [1986] misinterprets the notion of jitter in an older paper [Balakrishnan 1962]. Jitter in the latter referred to uncertainty in the periodicity of the sampling function. That is, while dealing with real acquisition systems, jitter is undesirable but unavoidable. In computer graphics, jitter is typically used to describe an intentional perturbation of the sampling function, for improved reconstruction or integration.

We focus on the analysis of error in numerical integration using stochastic sampling, as a function of the underlying sampling function. We derive the bias and variance of such integrators (section 2) in terms of the expected Fourier spectrum of the sampling function and its variance respectively. We study the bias and variance of a simple sampling algorithm (section 3).

2 Bias and variance of secondary estimators

Our goal is to analyze the quality of numerical integration, within the interval $[0, T]$, of a function $f(x) : \mathbb{R} \mapsto \mathbb{R}^+$. For example, a *primary* Monte Carlo estimator

$$I \equiv \frac{1}{T} \int_0^T f(x) dx \approx f(\mathbf{X}_i)$$

is well known to be unbiased if the single sample $\mathbf{X}_i \in [0, T]$ is drawn from a constant probability distribution function (pdf) within

the domain $[0, T]$. However, this simple estimator is rarely used in practice since it tends to have a high variance. A typically more useful *secondary* estimator is obtained by averaging N primary estimates at different \mathbf{X}_i . This process scales the variance down by a factor of N .

A general estimator for I is obtained by weighted averaging,

$$I \approx \frac{1}{N} \sum_{i=1}^N \alpha_i f(\mathbf{X}_i),$$

where $\alpha_i \in \mathbb{R}^+$ are normalized weights. Substituting $\alpha_i = 1/N$ and using randomly distributed \mathbf{X}_i in $[0, T]$ yields the secondary Monte Carlo estimator described above. While it is certainly possible to indiscriminately choose weights, α_i , this is also counter-productive since it potentially biases the estimator beyond use. The weighted average can be written as an integral

$$\frac{1}{N} \sum_{i=1}^N \alpha_i f(\mathbf{X}_i) = \int_0^T f(x) \mathbf{S}(x) dx, \quad (1)$$

where $\mathbf{S}(x) = \frac{1}{NT} \sum \alpha_i \delta(x - \mathbf{X}_i)$, $i = 1..N$ is the general, weighted sampling pattern. When the elements of the set $\{\mathbf{X}_i\}$ are obtained using some stochastic sampling strategy, $\mathbf{S}(x)$ is a random variable¹. Each such integral (rhs of eq. 1) is also a random variable that provides a single estimate of I .

The quality of these estimates varies across sampling strategies. Specifically, the quality depends on an intricate combination of the *distribution* from which $\{\mathbf{X}_i\}$ are drawn (eg. importance sampling), the *weights* α_i in the weighted average and a potentially beneficial *correlated arrangement* of the samples (eg. Poisson disk sampling). As an example of the effect of arrangement, consider stratified sampling and random sampling. The two strategies produce estimators that converge differently although the distribution and weighting of the samples is the same over the domain. However, the spectral characteristics of jittered sampling, as a result of the partitioning constraint induced by stratification, are different from those of random sampling. We express the bias and variance of stochastic integration schemes in terms of the spectral statistics of the integrand and sampling function, thus accounting for the combination of all three factors without explicitly studying each of them.

Notation The overscript hat denotes Fourier transforms and ω is the frequency variable. We denote random variables using capi-

¹a function of random variables is a random variable.

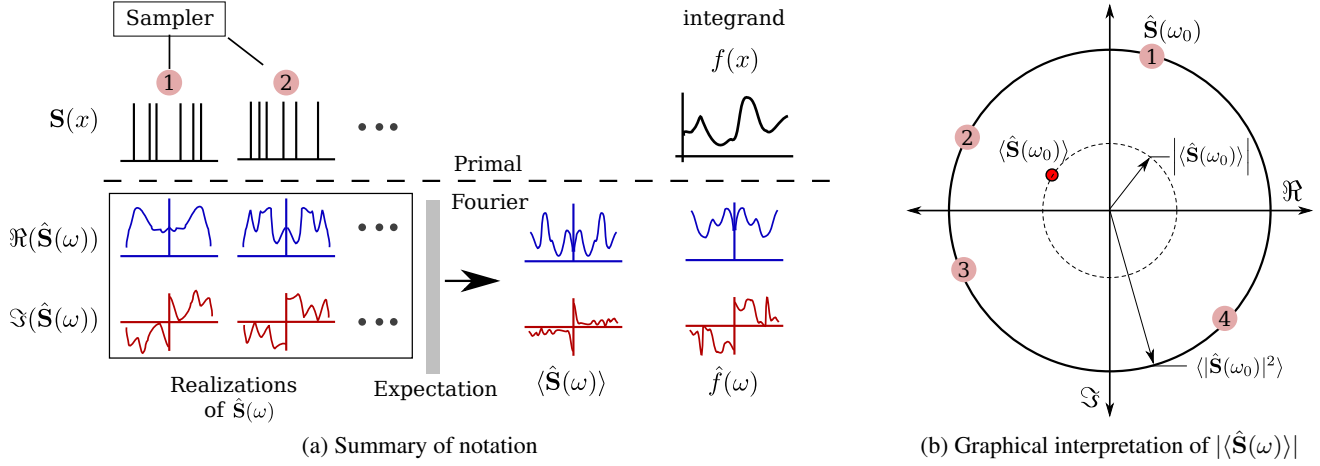


Figure 3: We study the statistics of the Fourier spectrum $\hat{\mathbf{S}}(\omega)$ over multiple realizations of the sampling function $\mathbf{S}(x)$. (a) illustrates our notation. (b) Shows 4 realizations (numbered pink circles) of $\hat{\mathbf{S}}(\omega)$ in the complex plane, evaluated at a given frequency ω_0 . The red circle is the mean of the 4 instances. The solid circle is the unit circle and the radius of the dotted circle is the amplitude of the mean value of the realizations. $V(\hat{\mathbf{S}}(\omega))$ describes the uncertainty in the location of the red dot. For this particular example, the periodogram is $\langle |\hat{\mathbf{S}}(\omega_0)|^2 \rangle = 1$ regardless of where, on the unit circle, the numbered pink circles are located.

talized bold letters. $\langle \cdot \rangle$ denotes the expectation operator and $V(\cdot)$ denotes the variance operator on random variables.

Realizations of the sampling function Each set of samples used to obtain a secondary estimate, forms a single realization of the sampling function $\mathbf{S}(x)$ that may be evaluated over the domain. For example, consider an image rendered with N samples $\{\mathbf{X}_i\}_{i=1..N}$ per pixel. The incident radiance, through each pixel, is a secondary estimate. Multiple realizations of the radiance estimate use different sets $\{\mathbf{X}_i\}$, drawn using the same strategy. Assuming that neighboring pixels have similar integrands and that they sample using the same strategy, a high variance manifests as noise in the image. We derive the influence of the statistics of $\hat{\mathbf{S}}(\omega)$, over multiple realizations of $\{\mathbf{X}_i\}$, on the bias and variance of the estimator.

2.1 Error in a single estimate

A single estimate \mathbf{F} , obtained using a single realization of $\mathbf{S}(x)$, is

$$\mathbf{F} = \int_0^T f(x) \mathbf{S}(x) dx = \hat{f}_{\Pi}(\omega) \otimes \hat{\mathbf{S}}(\omega)|_{\omega=0} \quad (2)$$

We use the shorthand $f_{\Pi}(x) \equiv f(x)\Pi_0^T(x)$ and $|_{\omega=0}$ to denote the DC term, which is the integral we seek. $\Pi_0^T(x)$ denotes the boxcar function which evaluates to unity when $x \in [0, T]$ and zero otherwise. $\hat{f}_{\Pi}(\omega)$ and $\hat{\mathbf{S}}(\omega)$ denote the Fourier spectra of the integrand² and sampling function respectively.

The error $\Delta \equiv I - \mathbf{F}$, of a single estimate (primary or secondary) can be written in terms of the Fourier spectra of the sampling function and integrand as

$$\Delta = \hat{f}_{\Pi}(0) - \int \hat{\mathbf{S}}(\omega) \hat{f}_{\Pi}(-\omega) d\omega \quad (3)$$

The error is a real value, despite $\mathbf{S}(x)$ and $\hat{f}_{\Pi}(-\omega)$ being complex. Since the integrand and sampling functions are real, the integral

² $\hat{f}_{\Pi}(\omega)$ is the convolution of $\hat{f}(\omega)$ with a sinc (Fourier transform of the box function).

over positive frequencies and negative frequencies have the same amplitude but their phases differ in sign. Adding the two integrals yields a real value. This error closely resembles (up to the handling of the domain) the expression derived in [Ramamoorthi et al. 2012; Durand 2011].

We study two properties of this error over different realizations of $\mathbf{S}(x)$. First, its expected value $\langle \Delta \rangle$ indicates whether the estimator yields the correct result, on average — that the estimator is *unbiased*. Second, its *variance*, $V(\Delta)$, provides information about the coherence of the estimates.

2.2 Bias

The bias of an estimator is the expected error

$$\langle \Delta \rangle = \hat{f}_{\Pi}(0) - \int \langle \hat{\mathbf{S}}(\omega) \rangle \hat{f}_{\Pi}(-\omega) d\omega, \quad (4)$$

due to linearity of the expectation operator. $\langle \hat{\mathbf{S}}(\omega) \rangle$ is the expectation of the Fourier spectrum of the sampling function over multiple realizations of the estimator. One way to obtain an unbiased estimator is to ensure that $\langle \hat{\mathbf{S}}(\omega) \rangle = \delta(\omega)$ in eq. 4. That is, construct a sampling strategy where the expected sampling spectrum only contains a DC term.

Example 1: Uniform jitter is unbiased In uniform jitter [Ouellette and Fiume 2001; Ramamoorthi et al. 2012] each secondary estimate uses a regular grid with a constant offset. This differs from regular jittered sampling in that each realization of the sampling pattern is still a regular grid, but across realizations, the entire grid is jittered randomly. It can be shown that the expected spectrum of a uniformly jittered regular grid contains only DC (see app. C), and therefore yields unbiased estimates.

Example 2: Eliminating bias in non-uniform sampling Consider a primary estimator using a weighted sampling function $\mathbf{S}(x) = \alpha(\mathbf{X}_i) \delta(x - \mathbf{X}_i)$ with $\mathbf{X}_i \sim g(x)$, where $g(x) : [0, T] \mapsto \mathbb{R}^+$ is

the pdf from which \mathbf{X}_i are drawn. Its expected Fourier spectrum is

$$\begin{aligned} \langle \hat{\mathbf{S}}(\omega) \rangle &= \langle \alpha(\mathbf{X}_i) (\cos(2\pi\omega\mathbf{X}_i) + \iota \sin(2\pi\omega\mathbf{X}_i)) \rangle \\ &= \int \alpha(x) (\cos(2\pi\omega x) + \iota \sin(2\pi\omega x)) g(x) dx \end{aligned}$$

An obvious weighting scheme to obtain $\langle \hat{\mathbf{S}}(\omega) \rangle = \delta(\omega)$ is to set $\alpha(x) = 1/g(x)$. That is, *importance sampling is unbiased* when the importance sample is positive everywhere (by the definition of $g(x)$). Random sampling is a particular case where $g(x)$ is a constant. Primary and secondary Monte Carlo estimators, using random sampling, are unbiased if the \mathbf{X}_i are drawn randomly from a constant pdf. Other weighting schemes are possible. The weights may be chosen to trade variance for bias.

2.3 Variance

The variance of the estimator is equal to the variance of the error since the integrand is constant across realizations of the sampling function. Furthermore, the realizations are mutually independent and hence

$$\boxed{V(\Delta) = \int V(\hat{\mathbf{S}}(\omega)) (\hat{f}_{\Pi}(-\omega))^2 d\omega.} \quad (5)$$

Equation 5 suggests that, for the resulting estimator to have a low variance, the variance of the sampling spectrum (across multiple realizations of $\{\mathbf{X}_i\}$) needs to be low where the integrand's power spectral density is large. Such a complementary spectral distribution of energy is even more important for variance, than for bias, due to its modulation by the power (as opposed to amplitude) spectrum of the integrand.

By the definition of the variance of a complex value, the spectrum of the sampling pattern is the sum of its real and imaginary parts:

$$\begin{aligned} V(\hat{\mathbf{S}}(\omega)) &= \langle (\hat{\mathbf{S}}(\omega) - \langle \hat{\mathbf{S}}(\omega) \rangle) (\hat{\mathbf{S}}(\omega) - \langle \hat{\mathbf{S}}(\omega) \rangle)^* \rangle \\ &= V(\Re(\hat{\mathbf{S}}(\omega))) + V(\Im(\hat{\mathbf{S}}(\omega))) \end{aligned} \quad (6)$$

z^* denotes the complex conjugate and $\Re(z)$ and $\Im(z)$ denote the real and imaginary parts of $z \in \mathbb{C}$. Since the variance of the estimator is the inner product of this spectral variance and the power spectrum of the integrand, studying $V(\hat{\mathbf{S}}(\omega))$ provides clues about classes of integrands that may be estimated with low variance. Unlike bias, the variance of the estimator does not depend on phase information in the integrand's Fourier spectrum.

2.4 Approximation: Amplitude of expected sampling spectrum

Our expression for bias (eq. 4) relies on knowledge of phase information of the integrand. It is impractical to expect knowledge of phase information of the integrand (or else I would be known). We now develop a conservative approximation for bias by ignoring the phase information in the integrand's spectrum.

The normalization of the sampling function results in the expected sampling spectrum's DC being unity. So the rhs of eq. 4 reduces to

$$\int |\langle \hat{\mathbf{S}}(\omega) \rangle| |\hat{f}_{\Pi}(-\omega)| e^{\Phi(\langle \hat{\mathbf{S}}(\omega) \rangle)} e^{\Phi(\hat{f}_{\Pi}(-\omega))} d\omega, \quad (7)$$

(excluding at $\omega = 0$) using the notation $z = |z|e^{\Phi(z)}$ to express each function in terms of its amplitude as well as phase. Ignoring the phases $e^{\Phi(\langle \hat{\mathbf{S}}(\omega) \rangle)}$ and $e^{\Phi(\hat{f}_{\Pi}(-\omega))}$ amounts to assuming that

the bias is the sum of the products of the magnitudes of the sampling and integrand spectra at all non-zero frequencies. This is a reasonable assumption when the phase of the integrand is not available, that leads to a worst case analysis of the bias. A large value for $|\langle \hat{\mathbf{S}}(\omega) \rangle|$ at any frequency ω necessarily causes an increase in the bias of the resulting estimator, unless the corresponding $|\hat{f}(\omega)| = 0$. When nothing is known about the integrand, the amplitude $|\langle \hat{\mathbf{S}}(\omega) \rangle|$ may therefore be used to predict the degree of bias for certain spectral properties of the integrand.

Example Figure 2(b) plots the amplitude of the expected spectrum for multiple sampling strategies. For example, we can predict that Poisson disk sampling (cyan), if used in integration, will result in the largest bias (on average). This is a conservative approximation, yet provides more information to assess potential bias than the periodogram. In addition to the phase of the integrand, the periodogram $\langle |\hat{\mathbf{S}}(\omega)|^2 \rangle$ also ignores the mutual phases across realizations of the sampling spectrum (see fig. 3(b) for an illustration). The periodograms of uniform jitter and regular grid sampling are identical, suggesting that the *periodogram is an unreliable tool* for assessing bias and variance of stochastic sampling for integration.

2.5 Summary

We presented equations to express the bias and variance of numerical integrators in terms of the Fourier spectra of the sampling function and integrand. The phase information present in each instance of the sampling pattern is essential for accurate prediction of bias and variance. Knowledge of the amplitude of the expected sampling spectrum is useful for bias prediction. Knowledge of the variance of the real and imaginary parts of the sampling spectrum are useful for predicting variance. Next, we present a simple variant of stratified sampling, and derive the expected value of its spectrum $\langle \hat{\mathbf{S}}(\omega) \rangle$ and the variance of its magnitude spectrum $V(|\hat{\mathbf{S}}(\omega)|)$.

3 Case study: Gaussian jittered sampling

We analyze a variant of jittered sampling, where samples on a regular grid are perturbed by independently drawn random variables from a radially symmetric Normal distribution. We call this Gaussian jitter and choose to analyze this strategy since Gaussians can be radially symmetric and generalize naturally to higher dimensions. In addition, our quantitative experiments reveal that Gaussian jitter has surprisingly good convergence properties (fig. 5). Formally, we perturb each \mathbf{X}_i to be $\mathbf{X}_i + \psi$ where the ψ are drawn independently for each i , from a Normal distribution with a zero mean and a variance of σ^2 .

Let $\hat{\mathbf{S}}'(\omega)$ be the Fourier transform of the sampling function with the Gaussian perturbation. We assume that the \mathbf{X}_i are fixed³ across instances of jittered sample sets, and derive the expected value and variance of the Fourier transform of Gaussian jittered samples to be

$$\langle \hat{\mathbf{S}}'(\omega) \rangle \approx_2 \left(1 - \frac{(2\pi\omega\sigma)^2}{2}\right) \frac{\hat{1}_{\perp N}(\omega)}{N} \quad \text{and} \quad (8)$$

$$V(\hat{\mathbf{S}}'(\omega)) \approx_2 \frac{(2\pi\omega\sigma)^2}{N} (1 + 2(\pi\omega\sigma)^2) \quad (9)$$

respectively. We use \approx_2 to denote second order approximation. The derivations are shown for the more general case, in appendices A and B, where the sampling spectrum is some general $\hat{\mathbf{S}}(\omega)$ (before adding jitter).

³we derive variance for fixed \mathbf{X}_i (Appendix B), but we derive the expected spectrum for potentially stochastic \mathbf{X}_i (Appendix A)

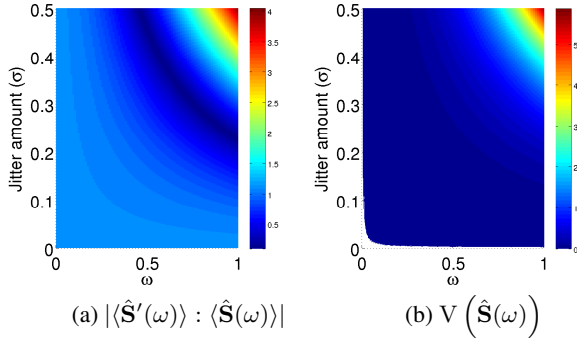


Figure 4: Heat-maps showing the effect of the jitter parameter σ . (a) At each ω , σ corresponding to the light blue region (ratio is 1) implies that the bias is not modified. The dark blue region is where bias is reduced, and the yellow region introduces severe bias. (b) Choosing σ to be in the dark blue region contains the variance at the corresponding frequency. Best viewed on screen.

The first observation is that $\langle \hat{S}'(\omega) \rangle = \delta(\omega)$ when $\langle \hat{S}(\omega) \rangle = \delta(\omega)$. That is, Gaussian jitter does not introduce bias in the estimator if applied on samples that originally had pure DC as their expected spectrum. The second observation is that, if applied on a biased estimator, the σ parameter can be used to modify the degree of bias.

In general, setting $\sigma = 0$ implies no jitter and therefore zero variance. As σ is increased, the variance increases. Figure 4(b) shows a heat-image of the variance as a function of the jitter parameter and frequency.

For bias, however, the first term on the rhs of eq. 8 suggests a less intuitive dependence on σ . Figure 4(a) shows $|\langle \hat{S}'(\omega) \rangle| : |\langle \hat{S}(\omega) \rangle|$ as a heat image, over the jitter parameter and frequency. The light blue area (where the ratio is unity) is the range of σ where the jitter does not modify the sampling spectrum at the corresponding frequency. The dark blue region, is where the corresponding σ reduces the bias. Beyond the yellow region, the jitter increases bias. If the integrand is known to contain energy in the higher frequencies, then σ must be chosen to be small enough to fall in the blue region, to reduce bias.

In all our experiments, we report σ as a fraction of the average sampling rate along each dimension. The actual parameter used in the experiments is the reported σ divided by \sqrt{N} (for 2D sampling).

4 Results

4.1 Spectral measurement

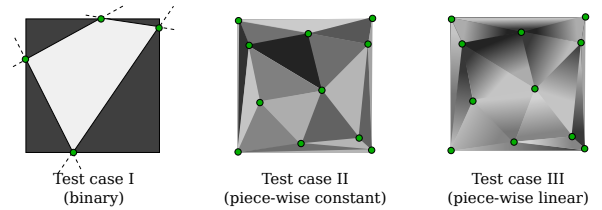
Figure 2 plots three statistics of the spectrum of sampling patterns, over multiple realizations of the sampling function. We drew 256 2D samples using 5 different strategies, computed the statistics over 20 iterations on a 512×512 grid, and plotted their radially averaged values. Figure 2(a) shows the periodogram, or averaged power spectrum over the iterations, against frequency. Figure 2(b) shows the amplitude of the expected sampling spectrum averaged with phase over the iterations, against frequency. Figure 2(c) is of the variance of the sampling spectrum, computed as shown in eq. 6, against frequency. The periodogram is often considered sufficient for the analysis of sampling patterns. Not only do the latter two statistics reveal rich information about the sampling, but they also provide intuition about the bias and variance if the samples were to be used in numerical integration. For example, if the integrand is known to contain high frequencies, say between the 100th and 200th coefficients, Figure 2(b) and (c) coupled with eqs. 4 and 5

reveal that Poisson disk sampling (cyan) and Gaussian jitter with $\sigma = 0.1$ (blue) are more likely to increase bias but reduce variance.

4.2 Quantitative tests

We performed quantitative measurements by estimating the integral of hypothetical pixels using the test cases. In each case, we estimated the integrals of simple functions using secondary estimators using various sampling strategies. We plotted the normalized variance of the estimators against the error of the averaged secondary estimators for each strategy (figs. 5a, b and c). Each sampling strategy results in a point on this plot, where the vertical axis corresponds to averaged variance and the horizontal axis is the averaged deviation from the analytically computed reference value. With Gaussian jitter (brown triangles), increasing σ decreases the bias at the cost of increased variance. However, the resulting variance is still comparable to quasi-random sequences. Furthermore, we plotted the variance as a function of the number of primary estimates N used in each realization of the secondary estimate (figs. 5d, e and f). A larger negative slope on this plot indicates faster convergence of the secondary estimator (see fig. 6 for a summary of observed convergence rates).

The data for all plots were generated by averaging errors over 10 random instances of each of the test cases. For each instance of the test case, 50 iterations of the secondary estimator were used with up to 1024 primary samples. To allow comparison across instances of the test cases, we computed relative errors: Mean of the secondary estimates divided by the analytical mean; and variance of the secondary estimates divided by their squared mean. We used the method of Gamito and Maddock [Gamito and Maddock 2009] for the Poisson disk sampling and Matlab’s implementation for all the low-discrepancy samples (with scrambling).



Test case I We integrated occlusion due to four half-planes within a square pixel. To allow simple computation of the occlusion area analytically, we constructed the half-planes so that they do not overlap in the pixel. That is, we drew four random samples (green points), one on each side of the square pixel, and connected them to form a quad. Each evaluation of the function returns 0 if the location lies in the quad and 1 otherwise. As reference, we computed the shaded area analytically.

Test case II We integrated a piecewise constant function within a square pixel. We obtained the function by sparse, random sampling of the square domain, Delaunay triangulation of the samples and random allocation of weights to each of the triangles. Each evaluation of the function returns the weight of the triangle in which the query point is located. The analytical integral was obtained as a weighted sum of the triangles.

Test case III We integrated a piecewise linear function within a square pixel. This case is similar to II, with the difference that the random weights were allocated to the vertices and linearly interpolated within the triangles.

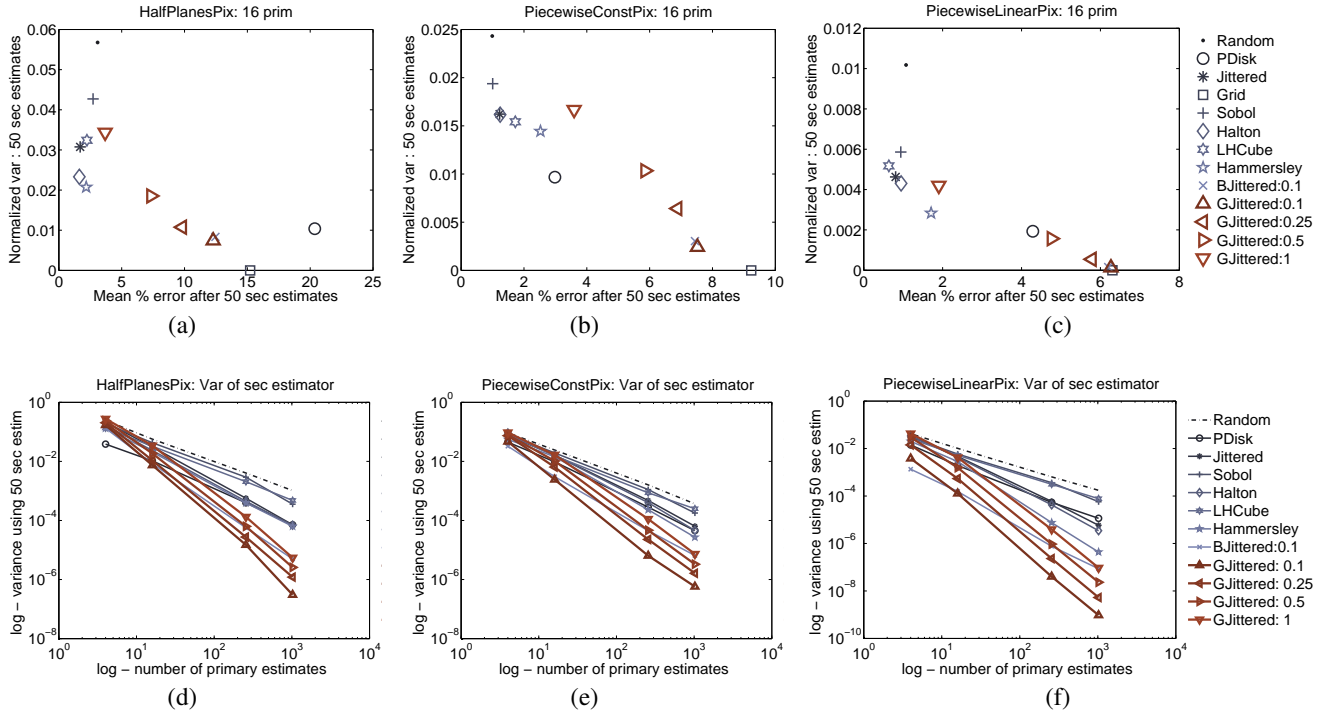


Figure 5: Plots depicting results from our quantitative experiments. The three columns correspond to the test cases described in sec. 4. The plots on the first row chart different strategies over the space of variance against accuracy (indicator of bias). Gaussian jittered sampling (brown triangles) can be adapted using σ to mitigate the bias due to regular sampling at the cost of adding variance. The plots on the second row compare the variances of the various strategies over 50 realizations, for different values of N . Gaussian jitter (brown lines) converge faster (larger slope) than other techniques. “LHCube” is latin hypercube sampling, “BJittered:0.1” is box-jitter within 0.1 times the width of each stratum and “GJittered:0.25” denotes Gaussian jitter with $\sigma = 0.25$.

	I: binary	II: p/w const	III: p/w linear
Random	$O(N^{-1.0020})$	$O(N^{-0.9745})$	$O(N^{-1.0098})$
Poisson disk	$O(N^{-1.0605})$	$O(N^{-1.1902})$	$O(N^{-1.3629})$
Jittered	$O(N^{-1.4171})$	$O(N^{-1.1877})$	$O(N^{-1.5003})$
Sobol	$O(N^{-1.0200})$	$O(N^{-1.0374})$	$O(N^{-1.0879})$
Halton	$O(N^{-1.4233})$	$O(N^{-1.2927})$	$O(N^{-1.5890})$
LHCube	$O(N^{-1.0112})$	$O(N^{-1.1255})$	$O(N^{-1.0903})$
Hammersley	$O(N^{-1.3813})$	$O(N^{-1.2324})$	$O(N^{-1.7772})$
BJittered:0.1	$O(N^{-1.8969})$	$O(N^{-1.6161})$	$O(N^{-1.7370})$
GJittered:0.1	$O(N^{-2.2478})$	$O(N^{-2.1373})$	$O(N^{-2.6781})$
GJittered:0.25	$O(N^{-2.1368})$	$O(N^{-1.9015})$	$O(N^{-2.5759})$
GJittered:0.5	$O(N^{-1.9464})$	$O(N^{-1.7070})$	$O(N^{-2.3969})$
GJittered:1	$O(N^{-1.7542})$	$O(N^{-1.5311})$	$O(N^{-2.0955})$

Figure 6: Observed empirical convergence rates for the three test cases (from figs. 5d, e and f).

4.3 Qualitative tests: Rendering

We incorporated Gaussian jittered sampling, in 1D and 2D, within PBRT [Pharr and Humphreys 2010]. We tested integration over one and two dimensional domains. In one dimension, we tested Integration over time for motion blur 9. In 2D we performed integration over light sources 8 and the lens aperture for defocus 7.

For low sampling rates (first rows of 7 and 9), as expected, the grid introduces obvious aliasing artifacts while random sampling has high variance. Low discrepancy sampling has lower variance than random sampling, but the noise is still prohibitive. Gaussian

jitter allows a smooth transition, using σ , from noise to banded artifacts. Uniform jitter shows some banding because, with few realizations, the influence of the regular grid is high.

5 Discussion

Trading bias for variance The regular grid yields a simple but biased estimator for integration. Since the estimator is consistent, it is useful when a large number of samples is used. At low sampling rates, our theory predicts that gaussian jitter may be used to reduce bias at the cost of increased variance, using a single parameter σ .

Variance of the sampling spectrum Spectral assessment of sampling patterns is typically performed using averaged power spectra (periodograms). Our theory derives that the variance of the sampling spectrum, over these iterations, provides a more reliable measure of the quality of integrators built using the sampling patterns. For example, consider the periodogram of a random offset and random rotation applied to a 2D regular grid. The spectrum, on average, is a set of concentric circles with peaks located close to the peaks of traditional “blue noise” distributions in computer graphics. The average radial energy distribution and anisotropy, according to the periodogram, might seem appealing. Yet each instance of the pattern is perfectly regular!

Spectrally-motivated sampling for integration Equation 5 is a preliminary step towards spectrally-motivated sampling for numerical integration. For integrating a specific integrand, given partial information (say its amplitude or power spectrum), a desirable spectral distribution for the sampling pattern may be worked backwards from eq. 5. Then recent algorithms that can generate point samples

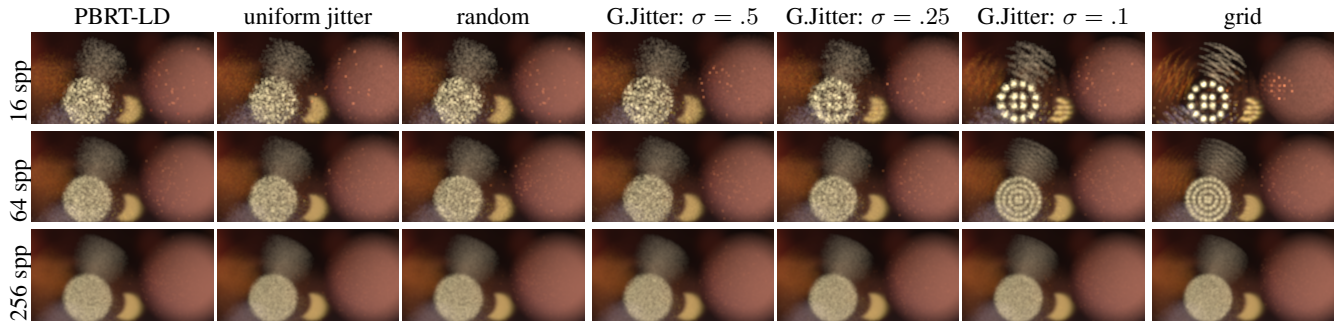


Figure 7: Comparison of bokeh resulting from different sampling strategies over a combination of the image and lens, using path tracing.

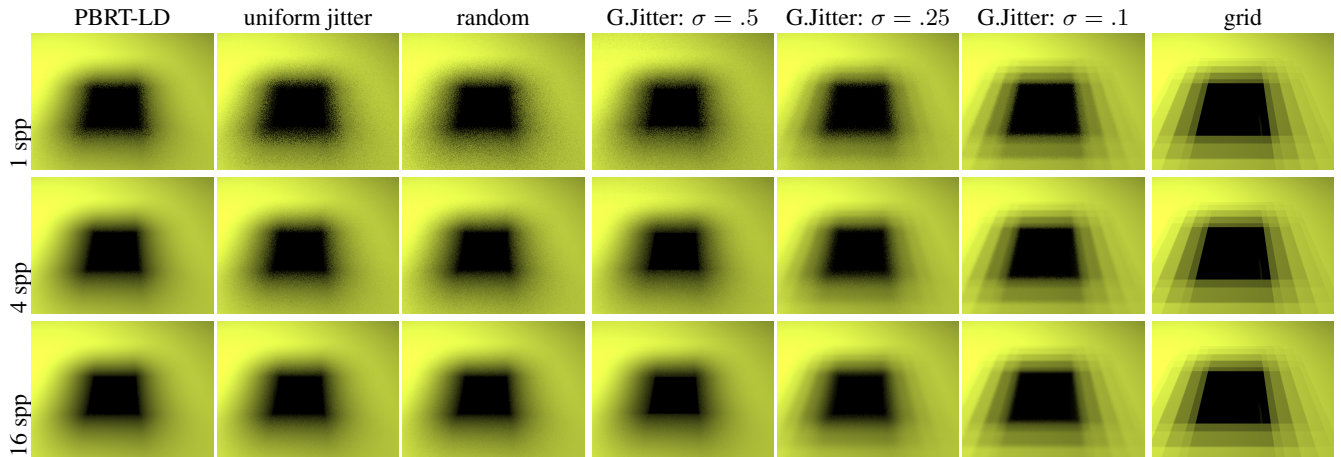


Figure 8: Soft shadows computed using 1, 4 and 16 secondary estimates per pixel (rows). A constant $N = 16$ visibility rays were used towards the light source (disc) placed above a square occluder. σ can be used to provide a continuum between the biased structure due to the grid and the variance of PBRT's low-discrepancy sampler.

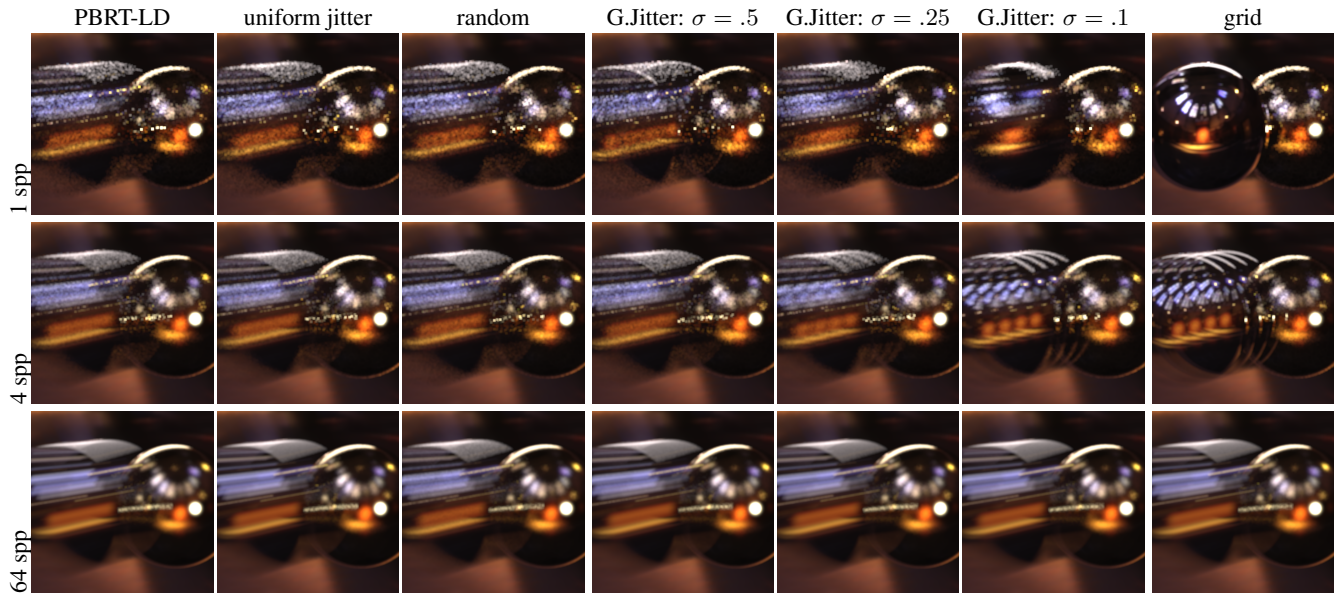


Figure 9: A comparison of the estimated motion blur integrated using different sampling strategies, using path tracing. At 4 spp, the bias of the regular grid creates objectionable artifacts. Gaussian jitter allows its trade-off for increased variance. This variance is comparable qualitatively to that of low-discrepancy sampling. At 64 spp, the result using a regular grid is qualitatively comparable to the result using low-discrepancy sampling.

with a specified noise spectrum [Zhou et al. 2012] may be used to actually generate the samples for integration.

Bounded domain sampling Two approaches may be taken to apply Fourier analysis to a bounded domain $[0, T]$. The first is truncation, and assumes that the function is zero outside $[0, T]$. Our absorption of the boxcar function in the integrand (eq. 2) takes this path. This approach runs the risk of introducing spurious frequencies due to the truncation, particularly if the domain is comparable to the scale of variation in the signal. The second is repetition, which assumes that the function is periodic with period T . We use this approach to derive that the expected spectrum of uniform jitter contains only DC (see Appendix C).

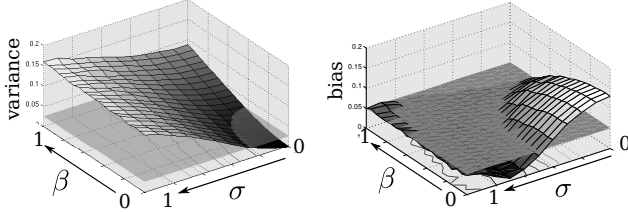


Figure 10: Combining uniform jitter with offset parameter β and our Gaussian jitter with parameter σ . Plots of variance and bias when the combined estimator was used to integrate the zone plate function. When $\alpha = 0$, $\sigma^2 = 0$ (deterministic sampling, regular grid) the variance is zero and the bias is maximum. When $\alpha = 1$, $\sigma^2 = 0$ (uniform jitter), the bias vanishes. In this example, the bias is still always lower than with regular grid sampling. The gray planes show the variance and bias of Poisson-disk sampling (similar number of samples).

Connection with uniform jitter Each realization of uniform jitter is a regular grid, with an overall offset. However, in our case, each realization contains samples that are independently jittered, Normally, from the grid centers. We experimented with combining our algorithm with uniform jitter. That is, for each realization, we apply a Gaussian-jitter on each sample obtained using uniform jitter. We varied the offset β (for uniform jitter) applied to the grid, and measured the average bias and variance while integrating a zone plate (see fig. 10).

Connection with box-jitter A simple alternative to Gaussian jitter is to simply jitter with a rectangular distribution where the distribution is narrower than the size of each stratum. For example, half-jittered sampling [Shirley 1991] only samples half the area of each stratum. We derive the expected spectrum and variance of the spectrum due to box jitter as

$$\langle \hat{S}'(\omega) \rangle \approx_2 \left(1 - \frac{(2\pi\omega\sigma)^2}{6} \right) \frac{\hat{1}1_N(\omega)}{N} \quad \text{and} \quad (10)$$

$$V(\hat{S}'(\omega)) \approx_2 \frac{(2\pi\omega\sigma)^2}{N} \left(\frac{1}{3} + \frac{4}{45}(\pi\omega\sigma)^2 \right), \quad (11)$$

by proceeding exactly as for Gaussian jitter (starting from eq. 12), except that we use the first, second and fourth order central moments of the rectangular distribution in $[-\sigma, \sigma]$ instead of those of the Normal distribution. While the variance of box-jitter is lower, its bias is high in several settings. The bias comparison plot (fig. 11) compares the bias introduced by box and Gaussian jittered sampling for various values of σ .

Lattice structure Recent work [Keller et al. 2006] suggests that using other lattice structures than a square grid is beneficial for an-

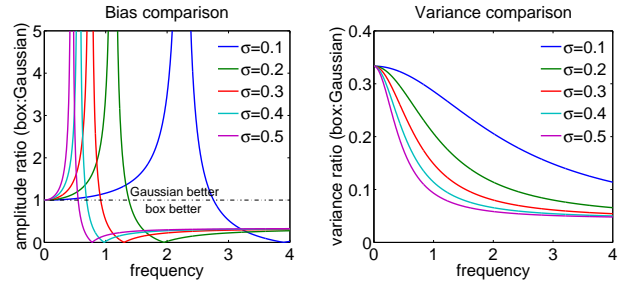


Figure 11: Comparison of box-jitter and Gaussian jitter. Left: plots of the ratio of the rhs of equations 10 and 8 for various σ . Box jitter is more biased if the integrand contains energy at frequencies where the ratio is greater than 1. Right: plot of the ratio of the rhs of equations 11 and 9 for various σ .

tialiasing. Our algorithm for generating the samples does not require a square lattice. The analysis and use of the samples to design a Monte Carlo estimator will depend on the lattice used.

Anisotropy Anisotropy can be interpreted in two ways. First, there is anisotropy of a given set of samples. This is a useful measure since it describes propensity of the samples to structure. There is also the notion of anisotropy of a particular strategy, obtained by averaging the power spectrum over multiple runs (periodograms). The latter is less informative. Again, consider a regular grid with a uniformly random rotation. Its anisotropy is high for a given set of samples, but vanishes when averaged over multiple realizations.

6 Conclusion

We studied the spectral characteristics of stochastic sampling patterns over multiple realizations of the patterns according to a fixed strategy. We derived formulae for the bias and variance of stochastic integration in terms of the the Fourier spectrum of the underlying stochastic sampling strategy.

We presented two new measures of the quality of sampling strategies in terms of their suitability to application in integration: The amplitude of the expected sampling spectrum and the variance of the sampling spectrum. We used these measures to assess Gaussian jittered sampling and compared it with box-jitter. Both algorithms controllably enable the reduction in bias due to uniform (grid) sampling in exchange for slight variance. We verified our results using quantitative experiments that suggest that Gaussian jitter has better convergence than a number of strategies that are currently used.

7 Acknowledgements

Thanks to Frédo Durand and Sylvain Paris for insightful discussions that inspired this analysis. Thanks to Cyril Soler for his many suggestions on reviewing an early draft. Kartic Subr acknowledges the Newton International Fellowship (2011-2012) from the Royal Academy of Engineering. Finally, thanks to the anonymous reviewers for providing useful feedback towards improving the paper.

References

- AMIDROR, I., HERSCH, R. D., AND OSTROMOUKHOV, V. 1994. Spectral analysis and minimization of moiré patterns in color separation. *J. Electron. Imaging* 3, 295–317.
- ARVO, J. 2001. Stratified sampling of 2-manifolds. *SIGGRAPH 2001 Course Notes* 29, 2.

- BALAKRISHNAN, A. 1962. On the problem of time jitter in sampling. *Information Theory, IRE Transactions on* 8, 3 (April), 226–236.
- BARTLETT, M. S. 1964. The spectral analysis of two-dimensional point processes. *Biometrika* 51, 299–311.
- BELCOUR, L., SOLER, C., SUBR, K., HOLZSCHUCH, N., AND DURAND, F. 2012. 5d covariance tracing for efficient defocus and motion blur. Tech. Rep. MIT-CSAIL-TR-2012-034, MIT, MA, November 2012.
- BOWERS, J., WANG, R., WEI, L.-Y., AND MALETZ, D. 2010. Parallel poisson disk sampling with spectrum analysis on surfaces. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 29, 6, 166:1–166:10.
- BRÉMAUD, P., MASSOULIÉ, L., AND RIDOLFI, A. 2003. Power spectra of random spike fields & related processes. *Journal of Applied Probability* 2002, 1116–1146.
- COOK, R. L. 1986. Stochastic sampling in computer graphics. *ACM Transactions on Graphics* 5, 1 (Jan.), 51–72.
- DIPPE, M. A. Z., AND WOLD, E. H. 1985. Antialiasing through stochastic sampling. In *ACM SIGGRAPH '85*, B. A. Barsky, Ed., 69–78.
- DURAND, F., HOLZSCHUCH, N., SOLER, C., CHAN, E., AND SILLION, F. X. 2005. A frequency analysis of light transport. *ACM Trans. Graph. (Proc. SIGGRAPH)* 24, 3 (July), 1115–1126.
- DURAND, F. 2011. A frequency analysis of Monte-Carlo and other numerical integration schemes. Tech. Rep. MIT-CSAIL-TR-2011-052, CSAIL, MIT, MA.
- EGAN, K., TSENG, Y.-T., HOLZSCHUCH, N., DURAND, F., AND RAMAMOORTHY, R. 2009. Frequency analysis and sheared reconstruction for rendering motion blur. *ACM Trans. Graph.* 28, 3 (July), 93:1–93:13.
- FIENUP, J. R. 1997. Invariant error metrics for image reconstruction. *Appl. Opt.* 36, 32 (Nov), 8352–8357.
- GALLAHER, L. J. 1973. A multidimensional Monte Carlo quadrature with adaptive stratified sampling. *Commun. ACM* 16, 1 (Jan.), 49–50.
- GAMITO, M. N., AND MADDOCK, S. C. 2009. Accurate multidimensional Poisson-disk sampling. *ACM Transactions on Graphics* 29, 1 (Dec.), 8:1–8:19.
- HACHISUKA, T., JAROSZ, W., WEISTROFFER, R. P., DALE, K., HUMPHREYS, G., ZWICKER, M., AND JENSEN, H. W. 2008. Multidimensional adaptive sampling and reconstruction for ray tracing. *ACM Trans. Graph. (Proc. SIGGRAPH)* 27, 3 (Aug.), 33:1–33:10.
- KELLER, A., AND HEIDRICH, W. 2001. Interleaved sampling. In *Rendering Techniques*, 269–276.
- KELLER, A., HEINRICH, S., AND NIEDERREITER, H. 2006. *Monte Carlo and Quasi-Monte Carlo methods*. Springer.
- KELLER, A. 2002. *Stratification by Rank-1-Lattices*. Interner Bericht. Universität Kaiserslautern, Fachbereich Informatik.
- KOLLIG, T., AND KELLER, A. 2002. Efficient multidimensional sampling. *Comput. Graph. Forum (Proc. Eurographics)* 21, 3, 557–557.
- KŘIVÁNEK, J., AND COLBERT, M. 2008. Real-time shading with filtered importance sampling. *Computer Graphics Forum (Proc. Eurographics Symposium on Rendering)* 27, 4, 1147–1154.
- LAGAE, A., AND DUTRÉ, P. 2008. A comparison of methods for generating poisson disk distributions. *Comput. Graph. Forum* 27, 1, 114–129.
- LARCHER, G., AND PILLICHSHAMMER, F. 2001. Walsh series analysis of the L2-discrepancy of symmetrized point sets. *Monatshefte für Mathematik* 132, 1, 1–18.
- LEHTINEN, J., AILA, T., LAINE, S., AND DURAND, F. 2012. Reconstructing the indirect light field for global illumination. *ACM Trans. Graph. (Proc. SIGGRAPH)* 31, 4 (July), 51:1–51:10.
- LENEMAN, O. A. 1966. Random sampling of random processes: Impulse processes. *Information and Control* 9, 4, 347–363.
- LUCHINI, P. 1994. Fourier analysis of numerical integration formulae. *Computer Physics Communications* 83, 23, 227–235.
- MACHIRAJU, R., SWAN, E., AND YAGEL, R. 1995. Spatial domain characterization and control of reconstruction errors. In *Proceedings of the 6th Eurographics Workshop on Rendering*, 33–44.
- MATÉRN, B. 1960. Spatial variation. *Meddelanden fran Statens Skogsforskningsinstitut* 49, 1–144.
- MATÉRN, B. 1986. *Spatial Variation*, 2nd ed. Springer Verlag.
- MITCHELL, D. P., AND NETRAVALI, A. N. 1988. Reconstruction filters in computer-graphics. *SIGGRAPH Comput. Graph.* 22 (June), 221–228.
- MITCHELL, D. P. 1987. Generating antialiased images at low sampling densities. M. C. Stone, Ed., vol. 21, 65–72.
- MITCHELL, D. P. 1991. Spectrally optimal sampling for distribution ray tracing. T. W. Sederberg, Ed., vol. 25, 157–164.
- MITCHELL, D. P. 1992. Ray Tracing and Irregularities of Distribution. In *Third Eurographics Workshop on Rendering*, 61–69.
- MITCHELL, D. 1996. Consequences of stratified sampling in graphics. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 277–280.
- NEYMAN, J. 1934. 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, 558–625.
- NIEDERREITER, H. 1992. *Quasi-Monte Carlo Methods*. John Wiley & Sons, Ltd.
- OSTROMOUKHOV, V. 2007. Sampling with polyominoes. *ACM Trans. Graph. (Proc. SIGGRAPH)* 26, 3 (July), 78:1–78:6.
- OUELLETTE, M. J., AND FIUME, E. 2001. On numerical solutions to one-dimensional integration problems with applications to linear light sources. *ACM Trans. Graph.* 20, 4 (Oct.), 232–279.
- ÖZTIRELI, A. C., AND GROSS, M. 2012. Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 31, 6 (Nov.), 170:1–170:10.
- PHARR, M., AND HUMPHREYS, G. 2010. *Physically Based Rendering, Second Edition: From Theory To Implementation*, 2nd ed. Morgan Kaufmann Publishers Inc.

- RAMAMOORTHI, R., AND HANRAHAN, P. 2004. A signal-processing framework for reflection. *ACM Trans. Graph.* 23, 4 (Oct.), 1004–1042.
- RAMAMOORTHI, R., ANDERSON, J., MEYER, M., AND NOWROUZSAHRAI, D. 2012. A theory of monte carlo visibility sampling. *ACM Trans. Graph.* 31, 5 (Sept.), 121:1–121:16.
- RIPLEY, B. 1977. Modelling spatial patterns. *J. Roy. Statist. Soc. B* 39, 172–212.
- SCHLÖMER, T., AND DEUSSEN, O. 2011. Accurate spectral analysis of two-dimensional point sets. *Journal of Graphics, GPU, and Game Tools* 15, 3, 152–160.
- SHIRLEY, P. 1991. Discrepancy as a quality measure for sampling distributions. In *Proc. Eurographics '91*, 183–194.
- SOLER, C., SUBR, K., DURAND, F., HOLZSCHUCH, N., AND SILLION, F. 2009. Fourier depth of field. *ACM Trans. Graph.* 28, 2 (May), 18:1–18:12.
- SUBR, K., AND ARVO, J. 2007. Statistical hypothesis testing for assessing Monte Carlo estimators: Applications to image synthesis. In *Pacific Graphics 2007*, 106–115.
- WEI, L.-Y., AND WANG, R. 2011. Differential domain analysis for non-uniform sampling. *ACM Trans. Graph. (Proc. SIGGRAPH)* 30, 4, 50:1–50:10.
- ZHOU, Y., HUANG, H., WEI, L.-Y., AND WANG, R. 2012. Point sampling with general noise spectrum. *ACM Trans. Graph. (Proc. SIGGRAPH)* 31, 4 (July), 76:1–76:11.

A Gaussian jitter of stochastic samples: Expected spectrum

Consider a sampling signal with just one sample \mathbf{X}_i . That is $\mathbf{S}(x) = \delta(x - \mathbf{X}_i)$, $i = 1$. If $\mathbf{S}'(x)$ is the sampling function after a Gaussian jitter by ψ , then the expected spectrum after jitter is

$$\hat{\mathbf{S}}'(\omega) = \cos(2\pi\omega(\mathbf{X}_i + \psi)) + i \sin(2\pi\omega(\mathbf{X}_i + \psi))$$

Let \mathbf{a}_i and \mathbf{b}_i denote its real and imaginary parts. The Taylor's series expansion about $\langle \psi \rangle = 0$, of the imaginary coordinate $\mathbf{b}_i(\psi) = \sin(2\pi\omega(\mathbf{X}_i + \psi))$ can be written as

$$\mathbf{b}_i(\psi) \approx \sin(2\pi\omega\mathbf{X}_i) + 2\pi\omega \cos(2\pi\omega\mathbf{X}_i)\psi - \frac{(2\pi\omega)^2 \sin(2\pi\omega\mathbf{X}_i)}{2!} \psi^2 - \frac{(2\pi\omega)^3 \cos(2\pi\omega\mathbf{X}_i)}{3!} \psi^3 + \dots \quad (12)$$

Applying the expectation operator on each side of the equation, truncating and substituting $\sin(2\pi\omega\mathbf{X}_i) = \Im(\langle \hat{\mathbf{S}}(\omega) \rangle)$,

$$\langle \mathbf{b}_i \rangle \approx_2 \left(1 - \frac{(2\pi\omega\sigma)^2}{2}\right) \Im(\langle \hat{\mathbf{S}}(\omega) \rangle).$$

Deriving the real part of the expected spectrum in similar fashion, and extending to N samples, $\mathbf{S}'(x) = \delta(x - \mathbf{X}_i - \psi)$, $i = 1..N$ to contain N impulses, its expected Fourier transform is

$$\langle \hat{\mathbf{S}}'(\omega) \rangle \approx_2 \frac{1 - 2(\pi\omega\sigma)^2}{N} \langle \hat{\mathbf{S}}(\omega) \rangle. \quad (13)$$

B Gaussian jitter of fixed-location samples: Variance of spectrum

When the \mathbf{X}_i are fixed and the jitter amount, ψ , is the only variable, we can apply the variance operator to eq. 12. Then

$$V(\mathbf{b}_i) \approx_2 (2\pi\omega\sigma)^2 \left(\cos^2(2\pi\omega\mathbf{X}_i) + \frac{(2\pi\omega\sigma)^2 \sin^2(2\pi\omega\mathbf{X}_i)}{2} \right)$$

and similarly, the variance of the real part of $\langle \hat{\mathbf{S}}'(\omega) \rangle$

$$V(\mathbf{a}_i) \approx_2 (2\pi\omega\sigma)^2 \left(\sin^2(2\pi\omega\mathbf{X}_i) + \frac{(2\pi\omega\sigma)^2 \cos^2(2\pi\omega\mathbf{X}_i)}{2} \right)$$

Adding $V(\sum \mathbf{a}_i/N) + V(\sum \mathbf{b}_i/N)$ yields the result in eq. 9.

C Spectrum for uniform jitter (1D)

In this case, the single random sample, periodically repeated, forms a Dirac comb (Shah function) with period T and an offset of x_1 . Its Fourier transform is also a Dirac comb, with period $1/T$ and a phase offset of $e^{-i2\pi x_1 \omega}$. That is,

$$\begin{aligned} \mathbf{S}(x) &= \sum_{j \in \mathbb{Z}} \delta(x + jT - x_1) \quad \text{and} \\ \hat{\mathbf{S}}(\omega) &= e^{-i2\pi x_1 \omega} \sum_{j \in \mathbb{Z}} e^{-i2\pi(-jT)\omega}. \end{aligned} \quad (14)$$

Since $\mathbf{S}(x)$ is periodic, $\hat{\mathbf{S}}(\omega)$ can only be non-zero at $\omega = k/T$, $k \in \mathbb{Z}$. Substituting this into eq. 14, we obtain

$$\lim_{\omega \rightarrow k/T} \hat{\mathbf{S}}(\omega) = e^{-i2\pi k x_1 / T} \lim_{\omega \rightarrow k/T} \delta(\omega - k/T)$$

since the exponential term in the summation is unity for each $\omega = k/T$. If x_1 is uniformly randomly distributed in $[0, T]$ (random sampling),

$$\begin{aligned} \langle \lim_{\omega \rightarrow k/T} \hat{\mathbf{S}}(\omega) \rangle &= \langle e^{-i2\pi k \varsigma} \rangle \lim_{\omega \rightarrow k/T} \delta(\omega - k/T) \\ &= \begin{cases} \delta(\omega) & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (15)$$

where $\langle \cdot \rangle$ denotes the expected value operator and $\varsigma = x_1/T$ is a random variable uniformly distributed in $[0, 1]$. Equation 15 shows that, for a randomly offset Dirac comb, the amplitude of the expected Fourier spectrum is zero everywhere but at $\omega = 0$.