A Theory of Monte Carlo Visibility Sampling

RAVI RAMAMOORTHI
University of California, Berkeley
JOHN ANDERSON and MARK MEYER
Pixar Animation Studios
and
DEREK NOWROUZEZAHRAI
Disney Research Zurich and University of Montreal

Soft shadows from area lights are one of the most crucial effects in high-quality and production rendering, but Monte-Carlo sampling of visibility is often the main source of noise in rendered images. Indeed, it is common to use deterministic uniform sampling for the smoother shading effects in direct lighting, so that all of the Monte Carlo noise arises from visibility sampling alone. In this article, we analyze theoretically and empirically, using both statistical and Fourier methods, the effectiveness of different nonadaptive Monte Carlo sampling patterns for rendering soft shadows.

We start with a single image scanline and a linear light source, and gradually consider more complex visibility functions at a pixel. We show analytically that the lowest expected variance is in fact achieved by uniform sampling (albeit at the cost of visual banding artifacts). Surprisingly, we show that for two or more discontinuities in the visibility function, a comparable error to uniform sampling is obtained by "uniform jitter" sampling, where a constant jitter is applied to all samples in a uniform pattern (as opposed to jittering each stratum as in standard stratified sampling). The variance can be reduced by up to a factor of two, compared to stratified or quasi-Monte Carlo techniques, without the banding in uniform sampling.

We augment our statistical analysis with a novel 2D Fourier analysis across the pixel-light space. This allows us to characterize the banding frequencies in uniform sampling, and gives insights into the behavior of uniform jitter and stratified sampling. We next extend these results to planar area light sources. We show that the best sampling method can vary, depending on the type of light source (circular, Gaussian, or square/rectangular). The correlation of adjacent "light scanlines" in square light sources can reduce the effectiveness of uniform jitter sampling, while the smoother shape of circular and Gaussian-modulated sources preserves its benefits—these findings are also exposed through our frequency analysis. In practical terms, the theory in this article provides guidelines for selecting visibility sampling strategies, which can reduce the number of shadow samples by 20–40%, with simple modifications to existing rendering code.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Shading; raytracing*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Monte Carlo sampling, stratified sampling, raytracing, soft shadows

ACM Reference Format:

Ramamoorthi, R., Anderson, J., Meyer, M., and Nowrouzezahrai, D. 2012. A theory of Monte Carlo visibility sampling. ACM Trans. Graph. 31, 5, Article 121 (August 2012), 16 pages.
DOI = 10.1145/2231816.2231819.
http://doi.acm.org/10.1145/2231816.2231819.

1. INTRODUCTION

Soft shadows from area lights are a key visual feature in high-quality and production rendering, and crucial to set the mood of a scene. Accurate soft shadows are most commonly rendered using Monte Carlo sampling of visibility to the area light at each pixel. For complex shadows, there can be considerable Monte Carlo noise, with hundreds of visibility samples required for acceptable images. Because of this expense, a number of alternative approximate techniques have been proposed and even used in production, but nearly all have artifacts or need significant manual parameter turning.

Soft shadow rendering would be considerably more practical if efficient Monte Carlo sampling strategies for visibility could be found and analyzed. In this article, we conduct a theoretical study, backed by empirical results, of different (nonadaptive) sampling strategies. We obtain new insights based on both the statistical properties of visibility as a binary function, and a Fourier analysis in the combined pixel-light space. We differ from previous work in focusing specifically on visibility, rather than Monte Carlo rendering in general.

To develop the theory, we first consider a linear light source and image scanline, for a 2D visibility field in the pixel-light space. (This setup introduces the basic insights and methodology, which we then extend to more general planar area light sources.) While a similar configuration would apply canonically to flatland visibility, our only assumption at this stage is a linear light, and we consider a full 3D scene. For each pixel in this case, the 1D visibility function over the linear light is binary, switching between 0 and 1 at a finite set of discontinuities. Our theoretical contributions include the following.

Theory of Visibility Sampling for Linear Lights. First, consider a single discontinuity, so visibility is a simple Heaviside step function

J. Anderson is currently affiliated with Google Inc.
Author's addresses: R. Ramamoorthi (corresponding author), Department of Electrical Engineering and Computer Science, 525 Soda Hall, University of California, Berkeley, CA, 94720; email: ravir@eecs.berkeley.edu; J. Anderson, M. Meyer, Pixer Animation Studios; D. Nowrouzezahrai, Disney Research Zurich and University of Montreal, Canada.

DOI 10.1145/2231816.2231819
http://doi.acm.org/10.1145/2231816.2231819

(or its complement) (Section 4.1). In stratified nonadaptive sampling, we show that the lowest expected error is obtained by placing samples at the center of the stratum for uniform sampling, even though it is well-known that uniform sampling leads to visual banding artifacts. Some quasi-Monte Carlo methods perform even worse than stratified sampling in this case, since they often place samples at the end-points of the stratum.

Uniform Jitter Sampling. Now, consider the case of two discontinuities (Section 4.2), so visibility is a box or rect function, starting at 0 (blocked), moving to 1 (visible), and returning to 0 (blocked)—the converse case of starting at 1 (visible) is equivalent. In this case, if the two discontinuities are in different strata, we derive a surprising result—error comparable to uniform sampling, but without banding artifacts, can be achieved by "uniform jitter" sampling, where we jitter a uniform grid, rather than jittering each stratum separately. In essence, the uniform jitter correlates the sample location at both discontinuities, leading to lower expected variance, by a factor of two, than stratified jittered sampling. While the uniform jitter method has been reported before in the literature (for example, the jittered offset quadratures in Ouellette and Fiume [2001]), it is very rarely used to our knowledge, and our article argues that it has much greater practical significance.

2D Pixel-Light Fourier Analysis. Of course, uniform sampling is not used in practice, as it leads to banding. We conduct a novel 2D Fourier analysis in the pixel-light space (Section 5), showing how banding arises from amplification of select spatial frequencies. We also show, through Fourier analysis diagrams, how uniform jitter sampling alleviates the banding concerns in pure uniform sampling. Our analysis is based on the sheared shape of the Fourier spectrum for visibility, as recently derived in Egan et al. [2011].

Planar Area Light Sources. We next seek to extend these linear light results to planar area lights (Section 6). Interestingly, the best sampling pattern depends on the type of light source (circular, Gaussian, square), as shown in Figure 1. For square or rectangular lights, uniform jitter sampling can lead to undesirable correlations between different "light scanlines" in the sampling pattern. On the other hand, for smoother falloffs like Gaussian or circular lights, uniform jitter sampling outperforms standard stratified sampling, or extensions like quasi-Monte Carlo sampling, and the method for circular sources [Shirley and Chiu 1997]. These results can also be understood with a Fourier analysis of the Monte Carlo sampling strategy.

We validate all of our theoretical results with a number of empirical simulations and results on actual scenes of varying complexity. While this article is focused primarily on a theoretical analysis of visibility sampling, our results also have practical implications. All of the Monte Carlo sampling strategies discussed here are very simple, which indicates that 20%–40% fewer samples can be used with minor modifications of existing code. In fact, we have implemented all of the methods as simple shaders for both offline rendering using Renderman (which is what the results in this article report on), and for interactive ray-tracing using NVIDIA's Optix (as shown in the accompanying video). While we focus on still image rendering, similar ideas apply to video, since each frame can simply be rendered with the appropriate sampling method.

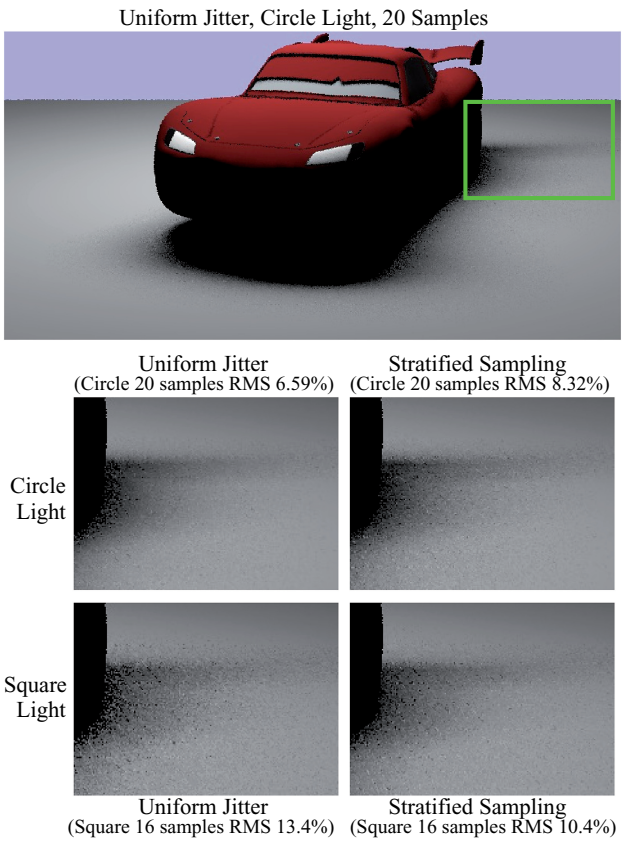


Fig. 1. This article shows that different sampling strategies perform better for different light source types. For circular lights (image and top row), uniform jitter sampling, that jitters a uniform grid at each pixel, has less noise than pure stratified sampling. (While the sample in each stratum is jittered independently in stratified sampling, only a single jitter per pixel is used in uniform jitter, and applied to each stratum.) For square lights (bottom row), the converse is true, with stratified sampling performing better. These images are shown with a relatively small number of samples throughout the article, to allow the reader to better assess the noise characteristics. Model courtesy Juan Buhler and Pixar, used with permission. ©Disney/Pixar

2. PREVIOUS WORK

This article draws on a rich literature in Monte Carlo and adaptive sampling, analysis of visibility, and signal-processing methods.

Monte Carlo Sampling. Different strategies for Monte Carlo sampling have been studied since Cook's seminal work on stochastic sampling [Cook 1986] and earlier studies on antialiasing and ray-tracing [Dippe and Wold 1985; Lee et al. 1985; Purgathofer 1986]. Mitchell [1987] has conducted some notable work on optimal sampling patterns. It is clear that purely random sampling is usually suboptimal, and better results can be obtained with stratified sampling [Mitchell 1996], where the sample location in each stratum is jittered (also known as jittered stratified sampling). A more uniform distribution of samples can sometimes be achieved by quasi-Monte Carlo methods [Niederreiter 1992; Heinrich and Keller 1994]. A frequency analysis indicates that the blue noise pattern [Cook 1986] can be desirable, and many recent works have focused on efficient Poisson disk sampling patterns [Wei 2008; 2010]. A comprehensive empirical study of Monte Carlo sampling strategies for linear lights is conducted in Ouellette and Fiume [2001].

Most strategies discussed in our article are based on stratified sampling, focusing on where in the stratum to place each sample. We also briefly discuss quasi-Monte Carlo strategies. We are inspired by the previous literature, but differ in focusing purely on visibility,

where its particular form allows us to derive further insights into the best sampling patterns.

Adaptive Sampling. Greater efficiency can sometimes be obtained with a variety of adaptive sampling techniques [Mitchell 1991; Guo 1998; Hachisuka et al. 2008; Overbeck et al. 2009]. There are also methods specialized to exploit shadow coherence [Agrawala et al. 2000; Ben-Artzi et al. 2006], and recent techniques for sheared reconstruction [Egan et al. 2011]. While these methods show the promise for great speedups in the future, many of them can sometimes have undesirable artifacts in complex regions, and in some cases involve manual parameter tuning.

This article focuses on (nonadaptive) standard Monte Carlo sampling, since that is still the most common method for accurate evaluation of soft shadows. We believe the approach in this article is largely orthogonal to adaptive sampling methods, which can leverage our insights separately at each refinement level. Moreover, shadow coherence methods like Agrawala et al. [2000] and Ben-Artzi et al. [2006] are independent of the base sampling scheme, and are therefore also likely to benefit from our theoretical analysis.

Analysis of Visibility. The structure of the "visibility skeleton" has been analyzed in previous work [Durand et al. 1997], but these results have been employed primarily for finite element radiosity methods, rather than Monte Carlo rendering. We also draw on frequency and first-order shadow analyses in Ramamoorthi et al. [2005, 2007], and Lanman et al. [2008], but note that these works focus on different theoretical questions. Our assumptions on the visibility field are more lightweight, essentially assuming that it is a binary function, and exploiting basic properties of its Fourier spectrum.

Signal-Processing. Many of our insights are derived from a Fourier analysis of visibility and Monte Carlo integration. In particular, we conduct a 2D Fourier analysis in pixel-light space. We are inspired by the space-angle Fourier analysis in Durand et al. [2005], and its application to understanding the sheared visibility spectrum [Egan et al. 2011]. A Fourier analysis or signal-processing view of Monte Carlo integration is rare in the literature, but we are inspired by the frequency periodograms in Ouellette and Fiume [2001], and the similar analysis in Durand [2011]. We extend these works in considering the full 2D pixel-light space, combining Fourier analysis with statistical properties of the visibility function, considering area light sources, and by analyzing different sampling patterns specifically for visibility.

3. BACKGROUND

We now introduce the basic background and notation for the article. The direct lighting from an area source can be written as

$$B(\mathbf{x}) = \int_{[0...1]\times[0...1]} L(\mathbf{y})T(\mathbf{x},\mathbf{y})V(\mathbf{x},\mathbf{y})\,d\mathbf{y}, \tag{1}$$

where B is the reflected radiance or image intensity, L is the lighting, V is the visibility, and the image is parameterized by $\mathbf{x}$, while the light source is parameterized by $\mathbf{y}$ (these will in general be 2D quantities that we map onto the unit square, although we start by analyzing the case of 1D linear light sources). T is the transport term that includes the BRDF, as well as the geometric form factor from the surface to the light source.

Our main focus is on diffuse soft shadows, since a different reflection technique (and layer) is often used for glossy highlights in production applications. One common simplification in previous work [Soler and Sillion 1998; Agrawala et al. 2000; Egan et al. 2011] is that the transport is sufficiently low frequency that it can be factored out of the integral, and substituted by its average value $\bar{T}$,

$$B(\mathbf{x}) \approx \bar{T}(\mathbf{x}) \int_{[0...1]\times[0...1]} L(\mathbf{y})V(\mathbf{x},\mathbf{y})\,d\mathbf{y}. \tag{2}$$

A more accurate version of this approximation uses stratified sampling for evaluating the visibility integral, where the sample locations in each stratum are chosen randomly to evaluate V, but deterministically (typically at the center of the stratum) to evaluate the lower-frequency T. This approach does introduce some bias, but the reduced variance is usually an acceptable trade-off. Indeed, the variance is now *only* in soft shadow regions, due to visibility alone.

In this article, we focus on computing

$$\hat{V}(\mathbf{x}) = \int_{[0...1]\times[0...1]} L(\mathbf{y})V(\mathbf{x},\mathbf{y})\,d\mathbf{y}, \tag{3}$$

where $\hat{V}(\mathbf{x})$ is the net visibility. An important special case occurs for a uniform light source, where we can assume $L(\mathbf{y}) = 1$,

$$S(\mathbf{x}) = \int_{[0...1]\times[0...1]} V(\mathbf{x},\mathbf{y})\,d\mathbf{y}, \tag{4}$$

where $S(\mathbf{x})$ is the fractional visibility or attenuation, and we assume we have normalized, so we are integrating over a unit area.

In the rest of this article, we will consider direct, nonadaptive Monte Carlo evaluation of Eqs. (3) and (4). In parts of the analysis we will also assume characteristic features of V, namely that it is a binary indicator function (either 1 or 0), with a fixed (usually sparse) set of discontinuities, where its value changes. These specialized forms, that abstract and isolate the key issues involved in computing soft shadows, allow us to derive new insights regarding various Monte Carlo sampling strategies.

4. SAMPLING LINEAR LIGHT SOURCES

We first analyze the simpler case of linear light sources, also focusing on a single image scanline, so x and y are scalars. The visibility function $V(x, y)$ can now be visualized as a 2D quantity. This analysis introduces our key statistical methodology, which we extend in the next sections to Fourier analysis and planar area lights.

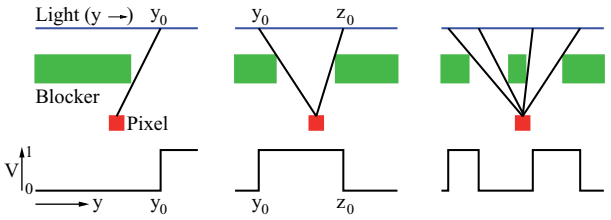
To proceed further, we assume the visibility function is binary (0 or 1), as is the case for opaque objects. Now, we can characterize the visibility at each pixel $V(\cdot, y)$ solely by the number and location of the discontinuities, where it toggles between 0 and 1, as shown in Figure 2. We will gradually consider more complex assumptions about visibility, which increase the number of discontinuities. We consider Eq. (4) in this section; Section 5.3 will briefly generalize the results to a varying light source in Eq. (3).

4.1 Single Discontinuity—Heaviside Function

The case where visibility is a constant at a pixel, with no discontinuities (completely blocked or fully visible) is not interesting, since any Monte Carlo sampling scheme for Eq. (4) will have zero variance. Therefore, we first consider the simplest interesting case, where for each pixel, $V(\cdot, y)$ has a single discontinuity (Figure 2(a)),

$$S = \int_0^1 H(y - y_0)\,dy, \tag{5}$$

where H is the Heaviside step function, and we omit the variable x since we are focusing on a single pixel. This assumes the visibility transitions from 0 to 1 at y_0 (we could also use $1-H$ for transitioning from 1 to 0 with similar insights).



(a) Single Discontinuity: Heaviside Function (b) Two Discontinuities: Boxcar Function (c) Multiple Discontinuities: Multiple Box Functions

Fig. 2. Schematic of visibility for a single pixel $V(\cdot, y)$ for a linear light source. Visibility is a binary indicator function, and can be completely characterized by the structure of the discontinuities where it transitions between 0 and 1. From left to right: (a) a single discontinuity: visibility is a Heaviside step function, (b) two discontinuities: visibility is a box function, and (c) multiple discontinuities which are grouped into multiple box functions.

It is clear that the correct value for the fractional attenuation is $S = 1 - y_0$. Our goal is to analyze the error in evaluating Eq. (5) by Monte Carlo sampling, keeping in mind that we do not know a priori that the visibility has a single discontinuity, nor what y_0 is,

$$\bar{S} = \frac{1}{N} \sum_{k=1}^{N} H(y_k - y_0), \tag{6}$$

where y_k are the N evaluation points for Monte Carlo sampling.

4.1.1 *Error Analysis of Sampling Schemes.* Clearly, the error of Eq. (6) depends on the number of $y_k < y_0$. Calling this number $A(y_0; y_k)$, we have that $\bar{S} = 1 - A(y_0; y_k)/N$, with error

$$|S - \bar{S}| = \left| \frac{A(y_0; y_k)}{N} - y_0 \right| \leq D_N^*(y_k). \tag{7}$$

The worst-case error is the highest over all possible values for y_0, and is known as the star discrepancy $D_N^*(y_k)$ of the Monte Carlo sampling pattern [Niederreiter 1992].

In 1D, a lower bound on the star discrepancy is given by the relation $D_N^*(y_k) \geq 1/(2N)$. It is easy to see intuitively that the bound is met by a *uniform sampling* pattern, that has

$$y_k = \left[\frac{1}{2N}, \frac{3}{2N}, \ldots, \frac{2N-1}{2N} \right]. \tag{8}$$

Figure 3(a) shows how the error varies with the location of y_0 for uniform sampling. The maximum error is $1/(2N)$ as expected (it will vary between $-1/(2N)$ and $+1/(2N)$). It is also easy to see that the average error across all y_0 is minimized by uniform sampling (it must be, assuming a uniform probability distribution for y_0, since no "clumpy" configuration for samples can be preferred).

Uniform and Stratified Sampling. We note that uniform sampling can be viewed as a special case of stratified sampling, with the strata boundaries at $(1/N, 2/N, \ldots, (N-1)/N, 1)$, and the samples always placed at the center of each stratum. In contrast, standard jittered stratified sampling places samples at a random location within the stratum. The errors in evaluation of S arise only from the stratum in which the discontinuity y_0 lies, which without loss of generality can be taken as the first stratum, in the range $[0, 1/N]$.

In this case, the error depends only on the relative location of y_0 and the Monte Carlo sample in the first stratum y_1, that is, whether $y_1 > y_0$ (stratum value 1) or $y_1 < y_0$ (stratum value 0). Plugging

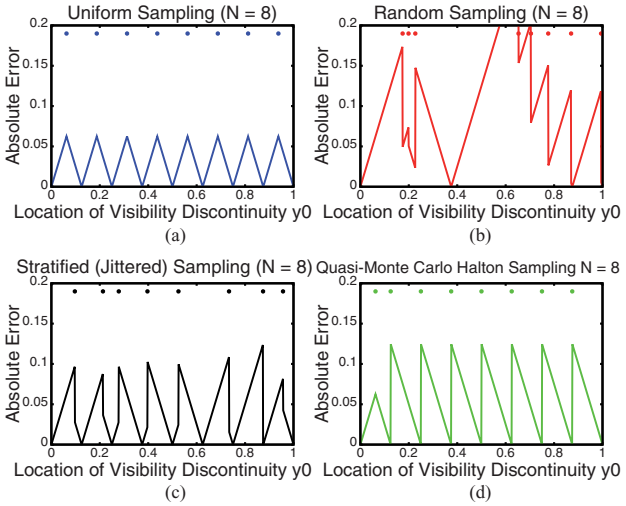


Fig. 3. Absolute error as a function of the discontinuity location y_0, for uniform sampling, random sampling, stratified sampling with jitter, and quasi-Monte Carlo (Halton base 2). These plots use $N = 8$, and the sample points are shown on the top of each plot. Uniform sampling has maximum error = $1/(2N) = 0.0625$, while stratified and Halton have a maximum error of $1/N = 0.125$, and (unstratified) random sampling has much higher errors. Note that Halton performs worse than stratified, because samples are often placed at the ends of a stratum.

into Eq. (7) (and remembering that $y_0 \in [0, 1/N]$),

$$S - \bar{S} = \frac{H(y_0 - y_1)}{N} - y_0 = \frac{1}{N} \left[H(\hat{y_0} - \hat{y_1}) - \hat{y_0} \right], \tag{9}$$

where it is convenient to define a normalized $\hat{y_0} = N y_0$, as well as $\hat{y_1} = N y_1$, which lie in the range $\hat{y_0}, \hat{y_1} \in [0, 1]$. The error is

$$\begin{aligned} |S - \bar{S}| &= \tfrac{1}{N}(1 - \hat{y_0}) \;\; : \hat{y_1} < \hat{y_0} \\ &= \tfrac{1}{N}\hat{y_0} \;\; : \hat{y_1} > \hat{y_0}, \end{aligned} \tag{10}$$

which is easily verified as the error in sampling a Heaviside function with one sample (and where the factor of $1/N$ accounts for the localization of the discontinuity to a single stratum).

In the uniform sampling case, $\hat{y_1} = \frac{1}{2}$, and the maximum error is $1/(2N)$ as expected. For jittered stratified sampling, $\hat{y_1}$ can lie anywhere in the range from 0 to 1, and the maximum possible error is twice that, $1/N$. It is also instructive to look at the expected variance (and corresponding standard deviation or RMS error) over all $\hat{y_0}$, given by

$$\begin{aligned} \left\langle (S - \bar{S})^2 \right\rangle &= \frac{1}{N^2} \left(\int_0^{\hat{y_1}} \hat{y_0}^2 \, d\hat{y_0} + \int_{\hat{y_1}}^1 (1 - \hat{y_0})^2 \, d\hat{y_0} \right) \\ &= \frac{1}{3N^2} \left(\hat{y_1}^3 + (1 - \hat{y_1})^3 \right). \end{aligned} \tag{11}$$

This expression makes clear that the expected error depends (only) on where in the stratum the Monte Carlo samples are taken. The error is clearly minimized when $\hat{y_1} = 1 - \hat{y_1}$, that is, for uniform sampling when $\hat{y_1} = \frac{1}{2}$. In this case, $\left\langle (S - \bar{S})^2 \right\rangle_{\text{UNIF}} = 1/(12N^2)$. For jittered stratified sampling, computing the expected error requires a

Sample Location in Stratum	Worst-Case	RMS Error	Variance
End-Points of Stratum	$\frac{1}{N}$	$\frac{1}{\sqrt{3}N}$	$\frac{1}{3N^2}$
Random (Jittered)	$\frac{1}{N}$	$\frac{1}{\sqrt{6}N}$	$\frac{1}{6N^2}$
Center (Uniform)	$\frac{1}{2N}$	$\frac{1}{\sqrt{12}N}$	$\frac{1}{12N^2}$

Fig. 4. Errors in stratified sampling of linear lights with one visibility discontinuity. Different sample placement strategies are shown, including end-points (worst), standard jittered, and center of stratum (uniform sampling).

further averaging over $\hat{y_1}$, and we obtain

$$\begin{aligned}\left\langle (S-\bar{S})^2 \right\rangle_{\text{STRAT}} &= \frac{1}{3N^2}\int_0^1 \left({\hat{y_1}}^3 + (1-\hat{y_1})^3 \right) d\hat{y_1} \\ &= \frac{1}{6N^2}. \end{aligned} \tag{12}$$

These results are summarized in Figure 4. Uniform sampling has the lowest expected error, with both worst-case and RMS twice as good as the extreme of placing samples at one end of the stratum. Stratified jittered sampling would require $\sqrt{2} \approx 1.4$ times as many samples to achieve the same error as uniform sampling.

Inspite of these advantages, uniform sampling is rarely used, as it can lead to banding in the image. This is due to the correlation in sampling pattern for different pixels. In Section 5, we develop a Fourier analysis that gives important new insights into both the optimal error nature of uniform sampling, and the banding artifacts. In Section 4.2, we show how a simple modification of uniform sampling can potentially achieve the same error performance for more complex visibility functions, without banding artifacts.

Other Sampling Strategies. We also briefly touch on Quasi-Monte Carlo (QMC) methods. In essence, these are techniques for generating low-discrepancy patterns that may have close to ideal blue-noise spectra, and are superior in many cases to stratified jittered sampling. Figure 3 includes results using the QMC Halton sequence [Keller 1997]. Surprisingly, the performance for this application is worse than stratified sampling, and just marginally better than the worst-case result (end-points of stratum) in Figure 4.

Some intuition is provided by considering the first four points of the Halton sequence with base 2, [1/2, 1/4, 3/4, 1/8]. In terms of our discussion of stratified sampling, the strata boundaries lie at [1/4, 1/2, 3/4, 1]. Three of the four Halton points (1/2, 1/4, 3/4) lie at the end-points of a stratum, and only one (1/8) lies at the stratum center. Thus, while the Halton sequence does stratify its samples well, the placement largely at the boundary of a stratum is not ideal, and is in fact a poor choice of $\hat{y_1}$ in Eq. (11).

Much recent attention has focused on sampling and reconstruction using compressive sensing [Candes 2006; Candes et al. 2006; Candes and Tao 2006], and the method has been applied to Monte Carlo rendering [Sen and Darabi 2010]. While a Heaviside function for visibility is inherently sparse in a wavelet or other basis, the benefits of compressive sensing arise largely from using random measurement patterns, not from point sampling. It is possible to use compressive sensing with a randomly chosen set of point samples, but the signal needs to be sparse in an incoherent basis, typically the Fourier series. Since visibility has sharp discontinuities, it is sparse only in a wavelet or other localized representation, not in the Fourier basis. Indeed, for any nonadaptive sampling scheme, the errors are bounded by Eq. (11), and cannot be improved by compressive methods.

For smooth integrands, higher-order methods such as Simpson's rule and Gaussian quadratures can give significantly faster convergence. However, visibility discontinuities correspond to the canonical nonsmooth indicator function, and therefore mitigate most benefits, as compared to the simple mid-point rule on which uniform sampling is based. In fact, Simpson's rule can be seen as approximately the same as uniform sampling, but with different weights for different strata, which will somewhat increase RMS error.

4.2 Two Discontinuities—Boxcar Function

We now consider the case where the visibility at a pixel $V(\cdot, y)$ is a boxcar function (also known as a box or rect function), with two discontinuities (Figure 2(b)). Formally (compare to Eq. (5),

$$S = \int_0^1 (H(y-y_0) - H(y-z_0))\, dy, \tag{13}$$

where the two discontinuities are noted as y_0 (visibility goes from 0 to 1) and z_0 (visibility goes from 1 to 0), and we assume that we have $z_0 > y_0$. Such structures are common for many canonical configurations as shown in Figure 2(b). The true value of the integral is $S = z_0 - y_0$, and Monte Carlo sampling evaluates it as

$$\bar{S} = \frac{1}{N}\sum_{k=1}^{N} H(y_k - y_0) - H(y_k - z_0). \tag{14}$$

4.2.1 *Error Analysis of Sampling Schemes.* We now make the crucial assumption that the discontinuities in y_0 and z_0 lie in different strata. This assumption is generally valid for most scenes, since different geometric regions are involved (for example, the left and right parts of a blocker that would be well-separated in angle). Indeed, the purpose of stratification is to group different visibility events or regions in different strata.[1] Under this assumption, like in the single-discontinuity case, the error for any stratified sampling scheme is determined entirely by the sample placement in the two strata of interest. In fact, Eq. (9) can be generalized to

$$\begin{aligned} S - \bar{S} &= \frac{1}{N}\left[(H(\hat{y_0} - \hat{y_1}) - \hat{y_0}) + (\hat{z_0} - H(\hat{z_0} - \hat{y_2})) \right] \\ &= \frac{1}{N}\left[(\hat{z_0} - \hat{y_0}) + (H(\hat{y_0} - \hat{y_1}) - H(\hat{z_0} - \hat{y_2})) \right]. \end{aligned} \tag{15}$$

It is important to understand that without loss of generality, we have placed the two discontinuities in the first and second strata, with $\hat{y_1}$ corresponding to $\hat{y_0}$ and $\hat{y_2}$ to $\hat{z_0}$. The normalized values of all variables lie between [0, 1], corresponding to the fraction of the stratum they are in. In particular, $\hat{z_0}$ is the fractional value along the stratum, not the global value of z_0. A schematic is in Figure 5.

Now, define $\triangle = \hat{z_0} - \hat{y_0} \in [-1, 1]$. There are four cases of interest, analogous to Eq. (10),

$$\begin{aligned} S - \bar{S} &= \frac{\triangle}{N} &&: \hat{y_0} < \hat{y_1}\ ,\ \hat{z_0} < \hat{y_2} \\ &= -\frac{1-\triangle}{N} &&: \hat{y_0} < \hat{y_1}\ ,\ \hat{z_0} > \hat{y_2} \\ &= \frac{1+\triangle}{N} &&: \hat{y_0} > \hat{y_1}\ ,\ \hat{z_0} < \hat{y_2} \\ &= \frac{\triangle}{N} &&: \hat{y_0} > \hat{y_1}\ ,\ \hat{z_0} > \hat{y_2}. \end{aligned} \tag{16}$$

Since $\triangle \in [-1, 1]$, the maximum error in preceding lines 2 and 3 can be $2/N$. This is expected (the worst-case error is twice that for

[1] In cases of thin fibers or other tiny geometry, stratification itself often has little benefit, and the analysis in this article doesn't apply. Indeed, pure random sampling performs as well as most other methods, as shown in Figure 21.

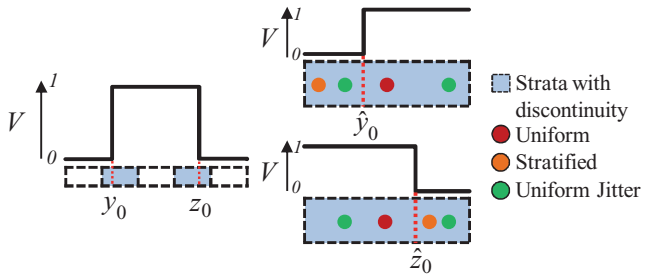


Fig. 5. Left: Schematic for boxcar function visibility (two discontinuities). Right: Assuming the discontinuities can be localized into distinct strata, the results depend only on the discontinuity and sample location within those strata. We show the sampling schemes for uniform (red, at center of stratum), uniform jitter (two possibilities shown in green, the sample location is the same for both strata, but jittered over the full pixel), and for stratified (one possibility shown in orange; jittering is independent for both strata). The orange stratified case shown here is the one of highest variance (00 in both strata). Uniform jitter eliminates this high-variance case (it can only be 01, 11 or 10, not 00 for this example).

sampling a single discontinuity). If we assume without loss of generality that $\Delta \geq 0$, then the maximum error is in the third preceding line, where $|S - \bar{S}| = (1 + \Delta)/N > 1/N$. We will now show that uniform sampling, and a variant we call uniform jittered sampling, eliminate this "bad" case (when $\hat{y}_0 > \hat{y}_1$, $\hat{z}_0 < \hat{y}_2$, $\hat{z}_0 > \hat{y}_0$), and hence have lower worst-case and expected variance. This is also shown schematically in Figure 5.

First, consider uniform sampling, so $\hat{y}_1 = \hat{y}_2 = \frac{1}{2}$. Now, assume $\Delta > 0$ so that $\hat{z}_0 > \hat{y}_0$. But in this case, if $\hat{y}_0 > \frac{1}{2}$, we must also have $\hat{z}_0 > \frac{1}{2}$. Therefore, the third line in Eq. (16) can never occur, and the worst-case error is only $1/N$, or half that for general stratified sampling, as in the case of a single discontinuity.

Uniform Jittered Sampling. In fact, it is easy to see that the same benefits are obtained by a *uniform jittered sampling*, which combines many of the benefits of uniform and stratified sampling. While this scheme is similar to the jittered offset quadratures reported in Ouellette and Fiume [2001], it is rarely used to our knowledge, and we believe it deserves more attention. In particular, we set $\hat{y}_1 = \hat{y}_2 = \gamma$ as in uniform sampling, but where γ is a randomly chosen variable as in stratified sampling (rather than $\frac{1}{2}$ in uniform sampling). In other words, we take the uniform sampling pattern, and jitter it by a constant offset at each shading location (image pixel). Each stratum has its sample jittered by the same amount, as opposed to stratified sampling that jitters each stratum separately. We will see that this approach preserves much of the variance reduction from uniform sampling for shadows from linear light sources, while eliminating the banding artifacts (γ is different for each pixel).

As for uniform sampling, if we assume $\Delta > 0$ so that $\hat{z}_0 > \hat{y}_0$, the fact that $\hat{y}_0 > \gamma$ implies that $\hat{z}_0 > \gamma$, so the bad case in the third line of Eq. (16) is eliminated. In fact, we can now reformulate Eq. (16) in the general case (with Δ positive or negative) as

$$\begin{aligned} |S - \bar{S}| &= \tfrac{|\Delta|}{N} \quad : \hat{y}_0 < \gamma \ , \ \hat{z}_0 < \gamma \\ &= \tfrac{1-|\Delta|}{N} \quad : \hat{y}_0 < \gamma \ , \ \hat{z}_0 > \gamma \\ &= \tfrac{1-|\Delta|}{N} \quad : \hat{y}_0 > \gamma \ , \ \hat{z}_0 < \gamma \\ &= \tfrac{|\Delta|}{N} \quad : \hat{y}_0 > \gamma \ , \ \hat{z}_0 > \gamma. \end{aligned} \tag{17}$$

Sampling Method	Worst-Case	RMS Error	Variance
Stratified	$\frac{2}{N}$	$\frac{1}{\sqrt{3}N}$	$\frac{1}{3N^2}$
Uniform Jitter	$\frac{1}{N}$	$\frac{1}{\sqrt{6}N}$	$\frac{1}{6N^2}$
Uniform	$\frac{1}{N}$	$\frac{1}{\sqrt{6}N}$	$\frac{1}{6N^2}$

Fig. 6. Errors in stratified sampling of linear lights with two discontinuities (visibility is a boxcar function). Uniform jittered sampling achieves the same expected variance as uniform sampling, and reduces error by a factor of two compared to stratified sampling.

Expected Variance. Having considered the worst-case error, we now compute the expected variance for different sampling schemes, assuming a uniform distribution in the discontinuity locations $\hat{y}_0$ and $\hat{z}_0$. By doing the calculations explicitly in Eq. (15), using a symbolic integrator, we obtain a rather simple result that generalizes Eq. (11).

$$\boxed{\left\langle (S - \bar{S})^2 \right\rangle = \frac{1}{N^2} \left[\frac{1}{6} + (\hat{y}_1 - \hat{y}_2)^2 \right].} \tag{18}$$

For the two-discontinuity case, the expected variance depends only on the difference in the sample locations for the two strata, $\hat{y}_1 - \hat{y}_2$. This is clearly minimized in the uniform sampling case when $\hat{y}_1 = \hat{y}_2 = 1/2$. However, the same variance reduction is achieved by jittered uniform sampling for any choice of γ, where $\hat{y}_1 = \hat{y}_2 = \gamma$, and γ is chosen randomly for each pixel (but then offsets the sample in each stratum by the same amount). The expected variance is now $1/(6N^2)$. This is twice as much as the expected variance for uniform sampling for one discontinuity, as expected. The variance for stratified sampling is obtained by considering $\hat{y}_1$ and $\hat{y}_2$ as uniform variables, and is given by $\left\langle (S - \bar{S})^2 \right\rangle_{\text{STRAT}} = 1/(3N^2)$, also twice that for a single discontinuity.

Perhaps most interestingly, the error for uniform jittered sampling ($1/(6N^2)$ as for uniform sampling) is the same as in the single-discontinuity case (in that case, uniform jitter is the same as standard stratified sampling, since we only care about the sample location in the one stratum of interest). For a single discontinuity, uniform jitter cannot help, but with two discontinuities, the error at both strata is decorrelated by the sampling pattern, to an extent that the net variance remains the same. These results are summarized in Figure 6, and should be compared to Figure 4.

4.2.2 *Multiple Discontinuities.* A more complex visibility function with multiple discontinuities can be treated by grouping the discontinuities into individual box functions (reducing to the two-discontinuity case given before, with possibly an additional single odd discontinuity), as shown in Figure 2(c). Therefore, this case reduces[2] to the analysis earlier in this section, with each pair or box function treated separately. Of course, the variance will increase linearly with the number of discontinuities.

Figure 7(a) shows the variance as a function of the complexity of the visibility function, for the common case of up to 7 discontinuities. We normalize the variance, multiplying by N^2 to allow direct numerical comparison with the theory in Figures 4 and 6. As expected, the variance from uniform sampling is the lowest (given from Figure 4 by 1/12 times the number of discontinuities), and that for stratified sampling is twice as large (1/6 per discontinuity).

[2]It is theoretically possible to construct scenes that have some correlations between the pairs of discontinuities, that must also be taken into account. However, we have not found this to occur in practice for general scenes.

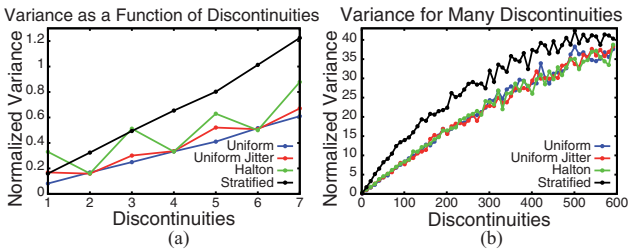


Fig. 7. Graphs showing the expected variance (over 1000 trials, with the number of samples $N = 256$) as a function of the number of discontinuities in visibility for a number of different simulated sampling schemes on linear lights. The left plot shows the common case of up to 7 discontinuities, while the right plot shows more extreme behavior, when we have very complex visibility with possibly multiple discontinuities in a single stratum; in this limit the performance of all schemes becomes comparable to stratified sampling. The variance is normalized by multiplying by N^2.

Uniform jitter has an error similar to stratified and twice as large as uniform for a single discontinuity (1/6 versus 1/12). However, the error does not change as we go to two discontinuities (Figure 6) and is now comparable to uniform. Indeed, this holds for any even number of visibility discontinuities, since they can be grouped into independent box functions. If there is an additional odd discontinuity, the variance increases by a constant 1/12. The results for Halton sampling are even more interesting, and nonmonotonic between odd and even numbers of visibility changes. For even discontinuities, errors behave similarly to uniform jitter and uniform sampling. However, the errors are larger for odd discontinuities, since samples are often placed at the end of a stratum (Figure 4).

Figure 7(b) is the extreme case, with very large numbers of discontinuities. Since the errors of Halton and uniform jitter are only a constant offset over uniform for odd numbers of discontinuities, all three graphs look very similar at this scale. However, stratified sampling still has about twice as much variance even up to 100 discontinuities. However, for extremely complex (essentially random) visibility, where there are multiple discontinuities in a stratum, our theory no longer applies and the variance of all methods approaches that of stratified sampling.

4.2.3 *Results*. Finally, we verify the analysis with the scene in Figure 8 using 16 visibility samples. The grids lead to complex soft shadows that are a good test for the theory. Uniform sampling leads to banding, even though the numerical error (an average of 3.97%) is lowest. Uniform jitter achieves almost the same quality (average error of 4.21%) but without banding. Both methods perform better than stratified (average error of 5.36%); Figure 8(c) is clearly noisier than Figure 8(b). Our theory predicts the variance (square of the standard deviation or error) should be twice as much for stratified as for uniform sampling. There is good agreement, with the variance over the whole image being 83% higher for stratified (some parts of the scene have no visibility events).

In summary, for linear light sources, the extremely simple step of switching from stratified to uniform jitter sampling for visibility can reduce the number of samples needed by about 30% in practice (the maximum theoretical benefit is $\sqrt{2}$ or approximately 40%; moreover, uniform jitter almost never increases the error relative to stratified). While this 30% benefit is small, it is concrete, and comes at minimal implementation or visual cost (no banding), while also being grounded in a careful theoretical analysis.

5. 2D PIXEL-LIGHT FOURIER ANALYSIS

Section 4 has considered the variance at each pixel separately. Further insight can be obtained from a Fourier analysis that simultaneously considers an image scanline (spatial) and the linear light (angular). We refer to this as 2D pixel-light Fourier analysis, since the spatial and angular dimensions interact [Durand et al. 2005].

5.1 Preliminaries: Fourier Analysis of Monte Carlo and Visibility Spectrum

To proceed with a frequency analysis, we must first introduce two important concepts—the Fourier analysis of Monte Carlo integration, and the form of the visibility spectrum. While Monte Carlo methods are usually studied statistically, using concepts like mean and variance, it is known that a Fourier analysis can convey some insights [Ouellette and Fiume 2001; Durand 2011]. Our contribution is to extend these ideas to a combined pixel-light analysis, and consider the effects of various sampling schemes (uniform, stratified, uniform jitter). To do so, we will need to know the form of the visibility Fourier spectrum. We draw on the recent results of Egan et al. [2011], who show that the spectrum is a wedge in pixel-light space for blockers in a range of depths. For simplicity, we consider blockers at a fixed depth, in which case the spectrum is a line, and we can derive considerable intuition.

First, consider Monte Carlo evaluation of Eq. (4),

$$\begin{aligned} S(x) = \int_0^1 V(x,y)\,dy &\approx \frac{1}{N}\sum_{k=1}^{N} V(x,y_k) \\ &= \int_0^1 V(x,y)\left(\frac{1}{N}\sum_{k=1}^{N}\delta(y-y_k)\right)dy \\ \bar{S}(x) &= \int_0^1 V(x,y)P(x,y)\,dy, \end{aligned} \tag{19}$$

where we make clear that the discrete Monte Carlo summation can also be considered as integration against a sampling pattern $P(x, y)$ that has delta functions at the chosen samples y_k. We have made explicit the dependence of P on x, since for a given sampling scheme, y_k may depend on x (in our case, they are either the same for all x for uniform or different for stratified sampling).

Now, Eq. (19) can be transformed into the Fourier domain, where we use bold letters for the Fourier transforms, and Ω for the frequencies. The multiplication of V and P in the primal domain becomes a Fourier domain convolution

$$\begin{aligned} Q(x,y) &= V(x,y)P(x,y) \qquad \bar{S}(x) = \int_0^1 Q(x,y)\,dy \\ \mathbf{Q}(\Omega_x,\Omega_y) &= \mathbf{V}(\Omega_x,\Omega_y)\otimes\mathbf{P}(\Omega_x,\Omega_y), \end{aligned} \tag{20}$$

while the integration in Eq. (19) restricts to the $\Omega_y = 0$ line

$$\bar{\mathbf{S}}(\Omega_x) = \mathbf{Q}(\Omega_x, 0) = (\mathbf{V}\otimes\mathbf{P})_{\Omega_y=0}\,. \tag{21}$$

Next, we need the form of the visibility spectrum $\mathbf{V}$, for which we use the results in Soler and Sillion [1998] and Egan et al. [2011]. A schematic of the setup is shown in Figure 9(a). We will further need to define $G()$ for the visibility function at the blocker plane, with Fourier transform $\mathbf{G}$, and $\beta > 1$ as a measure of depth; β is the ratio of distances between source and receiver to source and blocker ($\beta = 1$ when the blocker is at the receiver, and is large for blockers close to the light). It is easy to see from basic trigonometry that

$$V(x,y) = G\left(x\frac{1}{\beta} + y\left(1-\frac{1}{\beta}\right)\right). \tag{22}$$

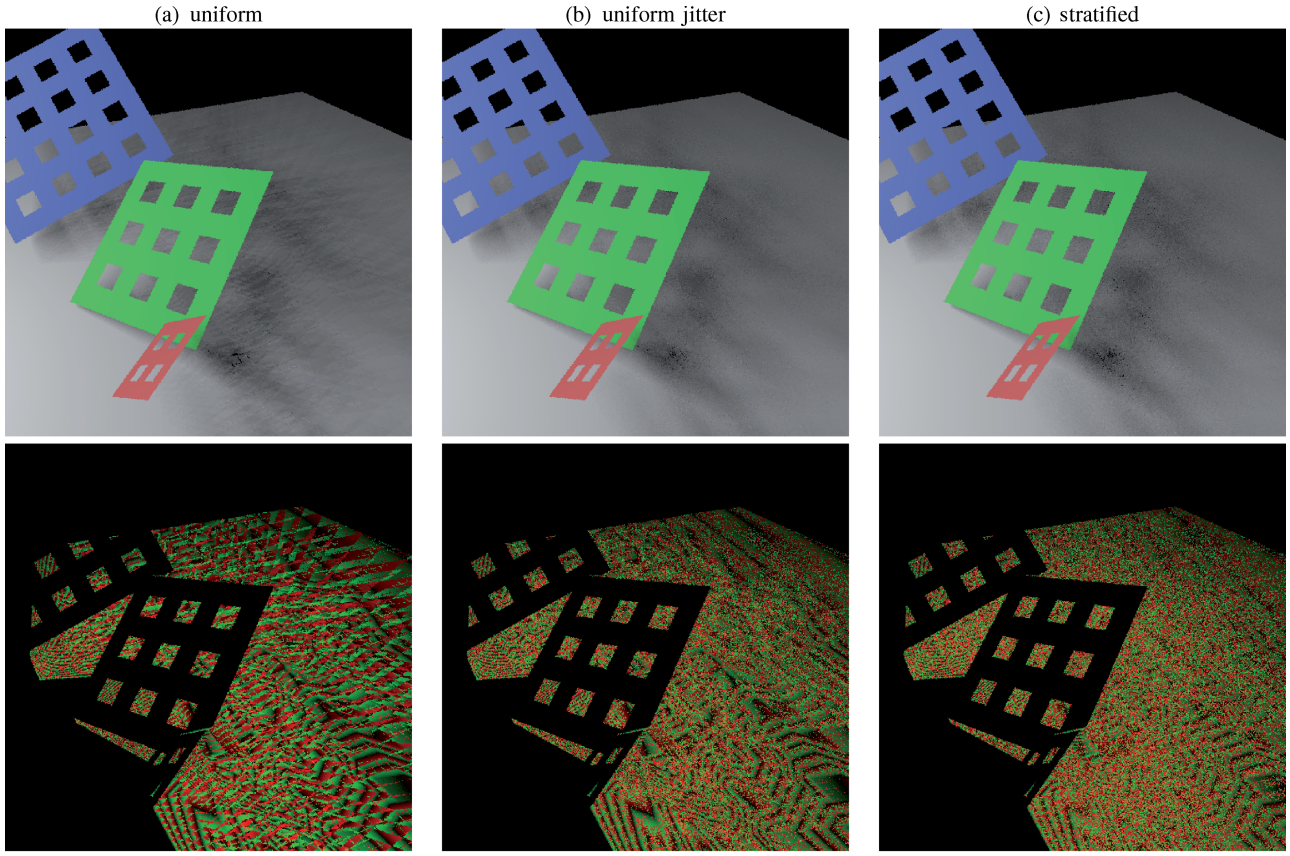


Fig. 8. Images with a linear light source and 16 visibility samples. The grids create interesting soft shadowing patterns with multiple visibility discontinuities at a pixel. The top row shows images and the bottom row error plots (red and green are different signs for the error, and intensity is proportional to magnitude; the errors are scaled up for display). (a) Uniform has the lowest error but leads to banding. This is also clear in the bottom row, from the structured nature of the error image. (b) Uniform Jitter has almost as low an error without banding, and considerably less noise than (c) Stratified. The reader may wish to flip between the supplementary images to see the differences most closely.

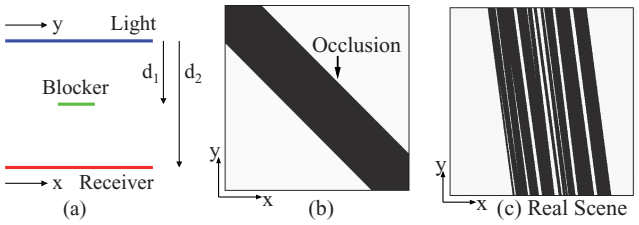


Fig. 9. Canonical visibility. (a): Schematic of setup; the blocker is in between light and receiver; the slope of the Fourier spectrum is determined by the ratio $\beta = d_2/d_1$. *(b):* Visibility function $V(x, y)$. Note that the occlusion is constant along lines with slope $(1-\beta)^{-1}$. *(c):* Similar occlusion function for a real scene. While more complicated, it has a similar sheared form. The Fourier spectrum of *(b)* is analyzed in Figure 10.

The occlusion pattern is shown in Figure 9(b), and remains constant on lines of slope $(1-\beta)^{-1}$. In essence, the occlusion pattern in G is sheared based on the depth of the blockers. For a real scene, as seen in Figure 9(c), the occlusion pattern may be more complicated but still has the same basic form, largely being a shear applied to G.[3] Now, we can proceed to consider the Fourier spectra for Figure 9(b), shown in Figure 10. From Eq. (22), it is possible to derive (see Eq. (7) in Egan et al. [2011])

$$\mathbf{V}(\Omega_x, \Omega_y) = \beta \mathbf{G}(\beta\Omega_x)\delta((\beta-1)\Omega_x - \Omega_y). \tag{23}$$

The spectrum is shown in Figure 10(a) and the energy is concentrated on a line with slope $\beta - 1$, as seen from the previous expression. Note that the spectrum on that line is also modulated by the original blocker spectrum $\mathbf{G}$ per the preceding Equation

In practice, because the light and receiver are finite, and the Fourier transform is not computed on a toroidal domain, there will also be some energy along the axes, in particular the $\Omega_y = 0$ line. Indeed, it is this line in the Fourier diagram that corresponds to the final result, since integration in the primal domain corresponds to restricting the Fourier spectrum to the horizontal axis $\Omega_y = 0$. In our

[3]While the occluder depth, and hence amount of shearing, may not be constant in the whole scene, the range of depths locally is often bounded. This observation has been exploited, for example, by Egan et al. [2011].

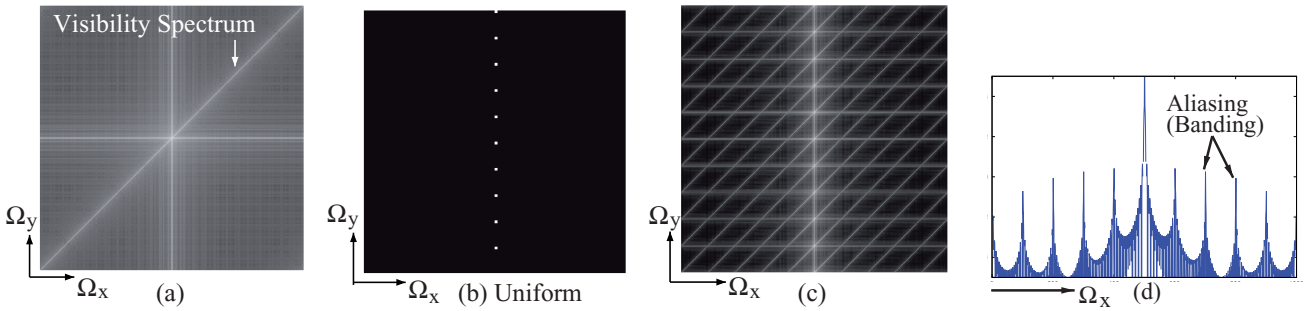


Fig. 10. Fourier analysis of Monte Carlo visibility sampling for the canonical uniform sampling case. (a): Fourier spectrum for visibility in Figure 9(b). Most energy lies on a line in the frequency domain (and also on the axes; the horizontal axis is the actual integrated visibility). (b): Fourier spectrum for uniform sampling. We have shown frequencies on a 1000×1000 grid with 100 samples, so there are 10 "aliases" all lying on the vertical axis. (c): Fourier spectrum for visibility multiplied by sampling pattern, which is a convolution in the frequency domain, leading to vertical replicas of the sheared visibility spectrum. Inaccuracies and aliasing occur when these replicas touch the final signal on the horizontal axis. (d): Fourier spectrum of net visibility. Notice the spikes (seen as banding) at select spatial frequencies, corresponding to the sampling pattern.

analysis of sampling schemes, numerical error will be introduced when aliases (due to sampling) of the sheared spectrum of Eq. (23) influence the horizontal line $\Omega_y = 0$.

Summary. Ours is a Fourier view of Monte Carlo sampling, wherein numerical error corresponds in frequency space to aliasing at the zero frequency. Also note that unlike traditional analyses of aliasing, we only care about aliasing to the $\Omega_y = 0$ line, since that is all that is preserved by Monte Carlo integration; aliasing elsewhere has no effect on the final result. We now proceed to analyze various sampling schemes in this context.

5.2 Analysis of Sampling Schemes

5.2.1 *Uniform Sampling.* We start with an analysis of uniform sampling, providing intuition about both the low numerical error, as well as artifacts from banding. Fourier spectra for the various steps for uniform sampling, corresponding to Eqs. (20) and (21), are shown in Figure 10.

We first consider the frequency spectrum $\mathbf{P}$ of the sampling pattern, as shown in Figure 10(b). Formally, for uniform sampling,

$$P(x, y) = \frac{1}{N} \sum_{k=1}^{N} \delta\left(y - \frac{k - \frac{1}{2}}{N}\right). \tag{24}$$

This is essentially a comb pattern in the y dimension, as is common in sampling theory, while there is no x dependence (hence the Fourier transform will be restricted to the vertical axis with $\Omega_x = 0$). The Fourier transform is also a comb along the vertical or y axis, and is given by

$$\mathbf{P}(\Omega_x, \Omega_y) \sim \delta(\Omega_x) \sum_{k=-\infty}^{\infty} \delta(\Omega_y - kN), \tag{25}$$

where we have omitted the exact phase information in $\mathbf{P}$ for simplicity. The spectrum for $\mathbf{P}$ is shown in Figure 10(b).

Now, the multiplication of visibility and sampling pattern to obtain Q is the standard sampling operation, which leads to replicas in the Fourier domain. In other words, since $\mathbf{Q} = \mathbf{V} \otimes \mathbf{P}$ from Eq. (20), we will have replicas of the visibility spectrum along the vertical axis, at frequencies a multiple of the sampling rate N. This effect can clearly be seen in Figure 10(c), where the sheared shadow spectrum is repeated vertically. Note that there are no horizontal repetitions, since the sampling pattern has a Fourier spectrum only along the vertical Ω_y axis.

Unlike traditional sampling analysis, this aliasing is in itself not a problem, since integration will restrict us to the line $\Omega_y = 0$ in the Fourier spectrum. Hence, it is only aliasing onto this line that will affect the final answer. To proceed further, we note that the convolution for $\mathbf{Q} = \mathbf{V} \otimes \mathbf{P}$ can be turned into a finite sum, because $\mathbf{P}$ is a comb, composed of a sum of delta functions for each k. For each visibility replica k, we consider the error introduced by aliasing onto the $\Omega_y = 0$ line, and sum these errors. Each replica of offset kN from Eq. (25) only interesects the line $\Omega_y = 0$ at $\Omega_x = kN/(\beta - 1)$, per Eq. (23). It can be seen that the numerical error is

$$\bar{\mathbf{S}} - \mathbf{S} \sim \sum_{k \neq 0} \beta \mathbf{G}\left(kN \frac{\beta}{\beta - 1}\right) \delta\left(\Omega_x - \frac{kN}{\beta - 1}\right). \tag{26}$$

Often, the error will be dominated by the first (or first few) terms. The aliases/bands can be seen as large numerical error, leading to spikes in the Fourier spectrum in Figure 10(d).

Discussion. Eq. (26) provides considerable insight. First, it gives a Fourier domain explanation for banding from uniform sampling. Error is concentrated on select spatial frequencies, which is what is perceptually distracting, even though the actual numerical error may be small. This can clearly be seen in Figure 10(d). In fact, we will see that uniform jitter sampling resolves banding by spreading the error through all spatial frequencies.

Moreover, we have explicitly found the spatial frequencies, $\Omega_x = kN/(\beta - 1)$ that band, relating them to the angular sampling rate, and the depth of the blockers. This allows for a number of insights. First, increasing the sampling rate N pushes the banding to higher frequencies. Assuming the initial spectrum $\mathbf{G}$ is decaying, the numerical error and perceptual artifacts will be less, as expected. Although it is not the focus of the article, Eq. (26) also leads to some potentially intriguing possibilities. If some prior knowledge of $\mathbf{G}$ is available, it may be possible to choose a sampling rate so that the $\mathbf{G}(kN\beta/(\beta - 1))$ terms are small or vanish, or to combine information from different sampling rates with different banding frequencies in a postprocess. (Our preliminary tests indicate this is indeed possible for very simple scenes, but is difficult in more complicated real-world situations with a range of depths, which is why we do not discuss it further in this article.)

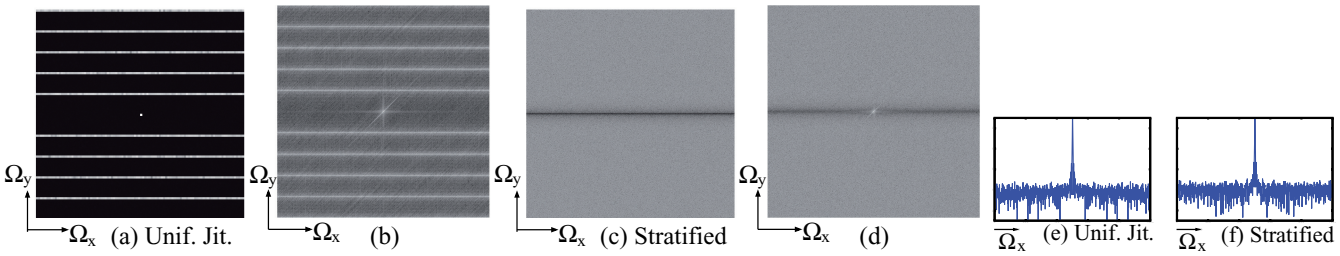


Fig. 11. Fourier analysis of visibility sampling for uniform jitter and stratified sampling (the visibility spectrum is shown in Figure 10(a)). (a): Sampling pattern for uniform jitter, which has the same vertical Ω_y frequences as in Figure 10(b), but with energy distributed over all Ω_x frequencies. (b): Fourier spectrum for product of uniform jitter sampling and visibility. The energy on the $\Omega_y = 0$ line is now distributed, rather than having discrete replicas as in Figure 10(c). (c): Sampling pattern for stratified sampling, which replaces structure with noise. Notice a "band-gap" at the center, that is, however, much smaller than that for uniform or uniform jitter. (d): Fourier spectrum for product of visibility and stratified sampling. There is no discrete aliasing, but it is replaced by noise. (e) and (f): Final results (Fourier spectrum) for uniform jitter and stratified sampling. In both cases, errors due to banding are replaced by noise.

Finally, the preceding analysis also provides considerable intuition into why uniform sampling should lead to the lowest numerical error. Banding simply concentrates the error in select spatial frequencies, as opposed to noise, but does not affect the overall numerical values. More importantly, the separation of replicas in the uniform sampling pattern in Figures 10(b) and (c) is the largest that can be obtained, as compared to alternatives like stratified or quasi-Monte Carlo sampling (this is also verified numerically in Ouellette and Fiume [2001]). Consider the visibility for a single pixel, that is, $V(y)$: As seen in Eq. (5), visibility is a Heaviside function in the simplest case, and it has a Fourier transform whose overall magnitude decays with the inverse of the frequency. Hence, numerical error or aliasing comes from higher-frequency regions of the visibility function that will be expected to have lower magnitude. However, uniform sampling leads to banding effects that make it unacceptable for most graphics applications. Indeed, uniform jitter is an attempt at a compromise, to get low numerical error without banding.

5.2.2 *Uniform Jitter Sampling.* We now turn our attention to a Fourier analysis of uniform jitter sampling. Recall that we now have a uniform sampling grid, but one that is jittered at every pixel

$$P(x, y) = \frac{1}{N} \sum_{k=1}^{N} \delta \left(y - \frac{k - \gamma(x)}{N} \right), \tag{27}$$

where $\gamma(x)$ is a uniform random variable chosen for each pixel or spatial location x between $[0 \ldots 1]$. For pure uniform sampling, we simply set $\gamma(x) = 1/2$ for all pixels.

The Fourier transform can be easily obtained by considering the transform first along y and then along x. The final spectrum is actually separable, since the $\gamma(x)$ factor just introduces a spatially-dependent phase, but the basic form of the spectrum is the same at each pixel. In particular, the y transform is the same as in uniform sampling, with the comb function leading to a comb in the Fourier domain. However, the phase depends on $\gamma(x)$ and is different at each pixel, unlike in uniform sampling where it is the same. Since $\gamma(x)$ is drawn from a random (white noise) distribution, so too will the Fourier transform along Ω_x and we obtain (compare with Eq. (25)),

$$\mathbf{P}(\Omega_x, \Omega_y) \sim \Gamma(\Omega_x) \sum_{k=-\infty}^{\infty} \delta(\Omega_y - kN), \tag{28}$$

where Γ is a random (white noise) variable. The magnitude of the resulting Fourier spectrum is seen in Figure 11(a), and is nonzero at the same Ω_y frequencies as for uniform sampling, but is distributed over the entire Ω_x axis instead of being concentrated at $\Omega_x = 0$.

Now, convolving visibility with the sampling pattern will diffuse error through all spatial frequencies, since replicas are created at all horizontal shifts (but only at discrete vertical frequencies). This is seen in Figure 11(b). As before, we care about aliasing only onto the $\Omega_y = 0$ line. It is easy to see that the same frequencies in $\mathbf{G}$ as for uniform sampling will affect the result, but the numerical error will now be diffused into all spatial frequencies (since there are replicas at all horizontal shifts). Analogous to Eq. (26),

$$\bar{\mathbf{S}} - \mathbf{S} \sim \sum_{k \neq 0} \beta \mathbf{G} \left(kN \frac{\beta}{\beta - 1} \right) \Gamma \left(\Omega_x - \frac{kN}{\beta - 1} \right). \tag{29}$$

Discussion. Figure 8 compares images rendered with uniform and uniform jitter sampling, along with error images. We see that uniform sampling has a distinctive structure corresponding to banding, while uniform jitter does not, while having comparable error.

An interesting point is that the $\mathbf{G}$ term in Eq. (29) indicates that the same frequency in the original blocker visibility is diffused as error to all spatial frequencies. This error term can vary in an oscillatory fashion over the image as β changes, and could in principle lead to some structure in the noise patterns. However, this structure in noise is much less distracting than banding structure in the actual image, and we have not observed it at all in most realistic scenes (including Figure 8 and most of our other examples).

5.2.3 *Stratified Sampling.* The sampling pattern for standard stratified sampling is similar to Eq. (27), but with $\gamma(x, k)$ now being random for each stratum in each pixel,

$$P(x, y) = \frac{1}{N} \sum_{k=1}^{N} \delta \left(y - \frac{k - \gamma(x, k)}{N} \right). \tag{30}$$

Analytic formulae for the resulting Fourier transforms are harder to derive, since they are no longer separable. However, the key insights are clear from Figure 11. For uniform and uniform jitter sampling (Figure 11(a)), there is a vertical "gap" of $\Omega_y = N$, corresponding to frequencies until the first replica. For stratified sampling in Figure 11(c), there is a smaller vertical gap.[4] These observations are also consistent with the frequency periodograms of

[4]A more sophisticated sampling approach like blue noise would make this gap larger, but still less than that for uniform sampling; we make some observations on Poisson Disk and quasi-Monte Carlo methods later in Section 8.

Ouellette and Fiume [2001]. Because of the randomness in stratified sampling, there is no banding, and error is diffused to all spatial frequencies.

In summary, a Fourier analysis indicates the key reason why stratified sampling can have higher numerical error than uniform—the frequency gap for the aliases is smaller, and for a decaying visibility Fourier spectrum, numerical error will come from higher-amplitude regions. Uniform jitter addresses this by incorporating the same separation of replicas as in uniform sampling, but diffusing the error among all spatial frequencies like in stratified sampling, avoiding the banding inherent in pure uniform sampling.

5.3 Varying Light Source

We now briefly consider a light source that has spatially-varying intensity. As discussed earlier, our focus is on visibility. We can treat smooth lighting and shading variation in the standard way, by using uniform sampling for the lighting terms and Monte Carlo only for visibility. In this way, our preceding analysis still holds.

Nevertheless, it is also instructive to consider directly applying our theory to a varying light source.[5] Combining Eqs. (3) and (19),

$$\bar{V}(x) = \int_0^1 V(x, y)L(y)P(x, y)\,dy, \tag{31}$$

where we have added the lighting term $L(y)$ to the integral, and we now use $\bar{V}$ to signify the modulated visibility, instead of just the average attenuation $\bar{S}$. This is a multiplication of three terms inside the integral (analogous in some ways to the triple product integral in Ng et al. [2004]), and not as amenable to convolution-based frequency analysis. Instead, considerable insight that leverages our earlier results can be obtained by grouping the lighting term L with either visibility V or the sampling pattern P.

First, consider a single pixel (so we fix x), and group the lighting with the visibility, defining $V'(y) = V(y)L(y)$. Now,

$$\bar{V} = \int_0^1 V'(y)P(y)\,dy \;=\; (\mathbf{V}' \otimes \mathbf{P})_0$$
$$\mathbf{V}' \;=\; \mathbf{V} \otimes \mathbf{L}. \tag{32}$$

As a canonical example of a varying light source, we consider one with a Gaussian distribution of intensity. In the Fourier domain, $\mathbf{V}'$ is then obtained by Gaussian-filtering or smoothing the Fourier spectrum of the original visibility $\mathbf{V}$. Figure 12 shows the original and smoothed visibility. Note that because of the nature of the visibility spectrum, arising from discontinuities, the original visibility spectrum is oscillatory, in addition to the inverse frequency decay. However, the filtered version behaves more smoothly. This smoother behavior is more amenable to frequency analysis (without consideration that a particular frequency or phase will skew the results), and we have noticed that Gaussian[6] or other smooth light sources often show greater advantages of uniform jitter over stratified sampling than square or constant lights.

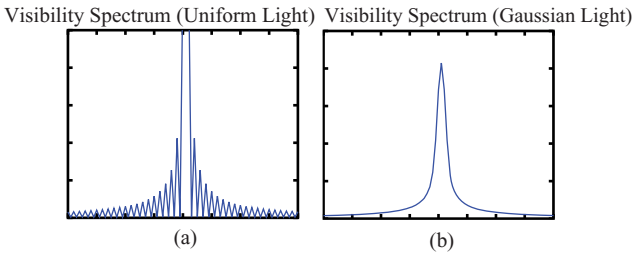


Fig. 12. Visibility Fourier spectrum for one pixel (for a single discontinuity on a linear light) for a constant source (left) and a Gaussian light (right). The RMS error in (a) is the same for uniform jitter and stratified as expected for a single discontinuity, while in (b) the smoother spectrum enables uniform jitter to perform essentially as well as uniform sampling.

For the 2D pixel-light analysis, it may be more insightful to group the lighting term with the sampling pattern, defining $P'(x, y) = P(x, y)L(y)$, and the Fourier transform $\mathbf{P}' = \mathbf{P} \otimes \mathbf{L}$. For uniform sampling and uniform jitter sampling, this diffuses or filters the Fourier spectrum vertically along the Ω_y axis. Banding still occurs with uniform sampling, but is less concentrated in specific discrete frequencies (instead diffusing to frequency bands). Similarly, the error in uniform jitter averages the magnitude of several frequencies, rather than coming from particular discrete frequencies in the visibility spectrum. More insights, including an analysis of different light source shapes, will follow in our discussion of planar area lights in the next section.

6. PLANAR AREA LIGHTS

In this section, we discuss a generalization from linear to planar area light sources. With any kind of stratification, the samples on the light are effectively divided into "scanlines", and we can directly apply the statistical analysis for linear light sources from Section 4 to each "light scanline" independently; similar arguments apply to the Fourier analysis in Section 5.[7] In this sense, the extension to planar area lights is straightforward. However, there are some important caveats, and the shape of the light source can make an important difference. Indeed, much of the research in this article was motivated by a surprising observation shown in Figure 13. For a circular light source, uniform jitter sampling performs better than stratified sampling, as expected. However, for a square light source, the opposite is true!

A schematic of stratified and uniform jitter sampling on a single pixel with a canonical visibility pattern is shown in Figure 14. For 2D planar area lights, visibility discontinuities typically become lines rather than single points for 1D linear lights. When these line discontinuities are (close to) aligned with the sampling grid, the integrals or sampling patterns for different "light scanlines" are highly correlated for uniform or uniform jitter sampling and do not further reduce variance, diminishing the benefits. Note that this is the other side of the coin in terms of correlation; the decorrelation

[5]Although we do not consider it in this article, these results are also easily extended to the case when shading effects like Lambertian cosine falloff or inverse square falloff are taken into account, since we can include all of these effects in the "effective" light source modulation function $L(x, y)$.

[6]Of course, in practice, importance sampling may be used within the Gaussian that effectively reduces it back to the case of a constant intensity light within each stratum. The irregular sampling patterns from importance sampling do not fit well within our theoretical framework and we do not consider this issue further in our article. However, we will see in the next section that for planar area lights, the benefits of Gaussian lights are also achieved by shapes like circular sources, where importance sampling is not possible.

[7]While a single horizontal light "scanline" might have each sample at a different vertical location if we jitter, the visibility function at that scanline is still a 1D function, being binary 0 or 1 at each sample, and with discontinuities where it moves from blocked to visible. In this sense, we are still considering a 1D function on each light "scanline", even though samples are distributed in 2D on the light even for a single row or column.

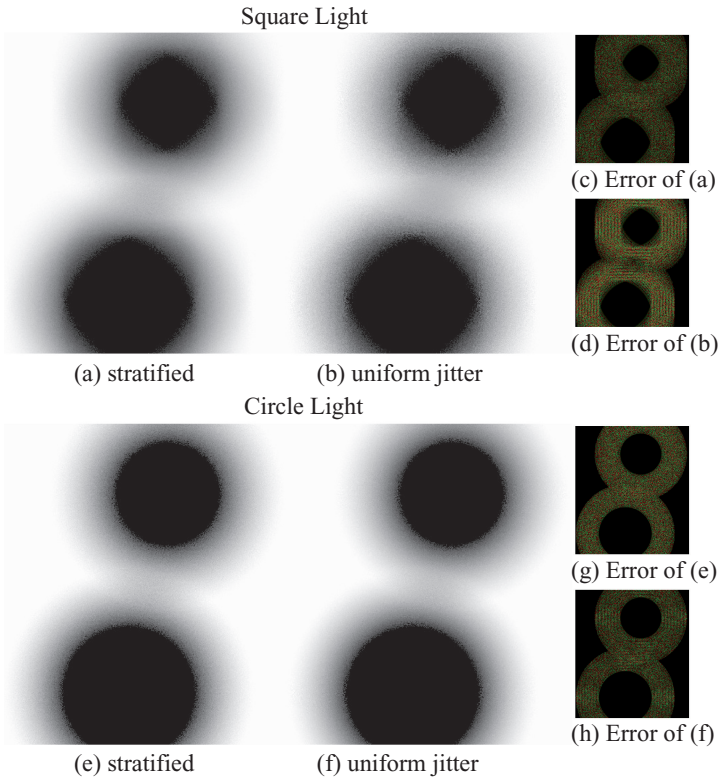


Fig. 13. Comparison of stratified (left) and uniform jitter (right) sampling, for a square light (top) and circular light (bottom). These images show a simple scene with two spherical occluders above a ground plane, rendered with 50 samples. The error images shown on the top and bottom right-hand side (where red is positive and green is negative) are multiplied by 10× for visualization. Curiously, uniform jitter performs better than stratified for circular lights (bottom row), but not on square lights (top row).

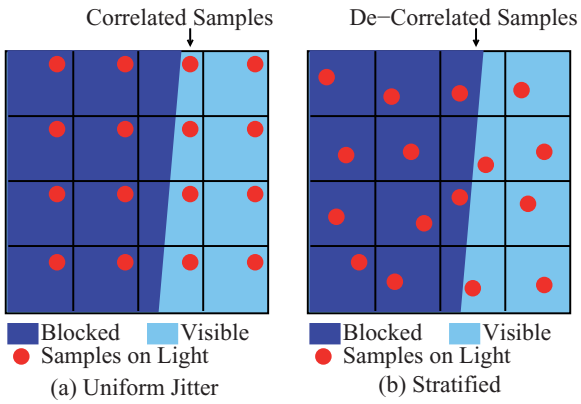


Fig. 14. Canonical visibility configuration for uniform jitter (left) and stratified (right). The correlation among the different "light scanlines" for uniform jitter can lead to poor results with square lights, since each "light scanline" obtains the same result, without further decreasing variance.

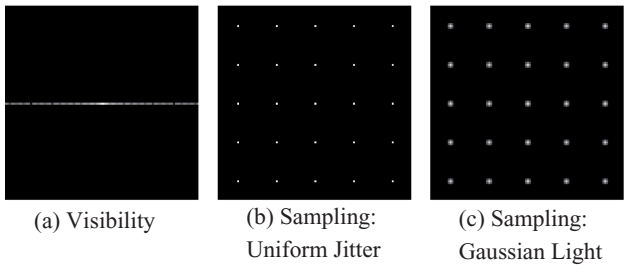


Fig. 15. (a): Fourier spectrum (at a single pixel) for the visibility on the square light, for a vertical discontinuity, similar to Figure 14. Discontinuities are typically lines or curves in light space, as well as in Fourier space. In this case, the spectrum is horizontal. (b): Spectrum of uniform jitter sampling, that is a comb, with energy concentrated at discrete frequencies. Large errors can result, since the frequencies on the horizontal axis overlap the visibility spectrum. (c): A Gaussian light is equivalent to convolving the sampling spectrum with a Gaussian. The energy is now diffused away from discrete frequencies, reducing the aliasing or overlap with visibility.

at multiple discontinuities for linear lights leads to the benefits of uniform jitter in the first place. We also experimented with a number of variants, such as the more closely packed quincunx pattern, but did not find any significant difference.

This effect can also be analyzed in the Fourier domain, where we limit our attention to a single pixel (it is still 2D because of the 2D light, but should be distinguished from the 2D pixel-light analysis in Section 5). We know in this case that $\bar{S} = (\mathbf{V} \otimes \mathbf{P})_0$, and error is obtained by aliasing into the zero or DC frequency. Pictorially, if we overlay the visibility and sampling spectra, and they have strong overlap outside the DC component, the numerical error will be large. In terms of the figures, imagine Figure 15(a) overlayed

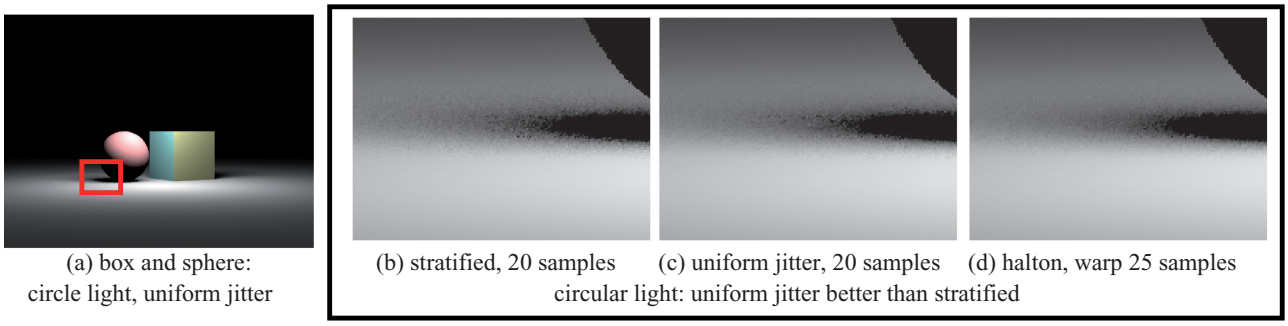


Fig. 16. Comparison of sampling methods for a simple scene with a sphere and a box, shown in (a,c) rendered with 20 uniform jitter samples and a circular light source. Next, we show an equal sample (b) and equal quality (d) comparison with stratified and quasi-Monte Carlo sampling. Note that (d) uses the Halton sequence, and the Shirley-Chiu [1997] square to disc warp, that we empirically determined performs best among the stratified sampling schemes.

on Figure 15(b); they will overlap strongly along the horizontal aliases in Figure 15(b). In constrast, if we overlay Figure 15(a) and Figure 15(c), they will still overlap on the horizontal line, but the energy in Figure 15(c) is no longer concentrated in delta functions at fixed locations on that line, but diffused, so the overlap is less pronounced.

More formally, for a uniform or uniform jitter pattern, since the primal domain pattern is a comb, the Fourier domain pattern is a comb as well, with energy concentrated at discrete frequencies. On the other hand, the visibility spectrum for edge or line discontinuties also lies along a line. Thus, energy is concentrated in this line, and if it overlaps strongly with any of the discrete frequencies in the sampling spectrum, high numerical error can result. In 1D, this is unlikely, but for 2D lights, as shown in Figure 15, the visibility spectrum could be horizontal or vertical, leading to high numerical error. This is simply a Fourier interpretation of Figure 14. Note that even for other edge directions, the line will often overlap somewhere with the discrete frequencies in the sampling spectrum.

For varying light sources (Section 5.3), we know that the sampling pattern can be regarded as convolved with the light source. Thus, the sampling energy is diffused or filtered away from a discrete frequency to a band, and the numerical error from overlap of the visibility spectrum line is less severe. We show this for a Gaussian light source in Figure 15(c) and a similar analysis applies to circular lights (which have a frequency spectrum corresponding to the Airy disk). Hence, uniform jitter often performs better than stratified sampling for circular, Gaussian, or other smooth light sources, while it performs worse for square or rectangular lights (see Figure 13).

7. RESULTS

So far, we have developed the theory of visibility sampling, showing simple canonical examples to verify the key ideas. We now present results on more realistic scenes that demonstrate the theory holds broadly in practice. In particular, we show that uniform jitter sampling outperforms stratified sampling for circular light sources. To focus on visibility, we render these scenes with one sample per pixel without antialiasing, with simple diffuse shading.

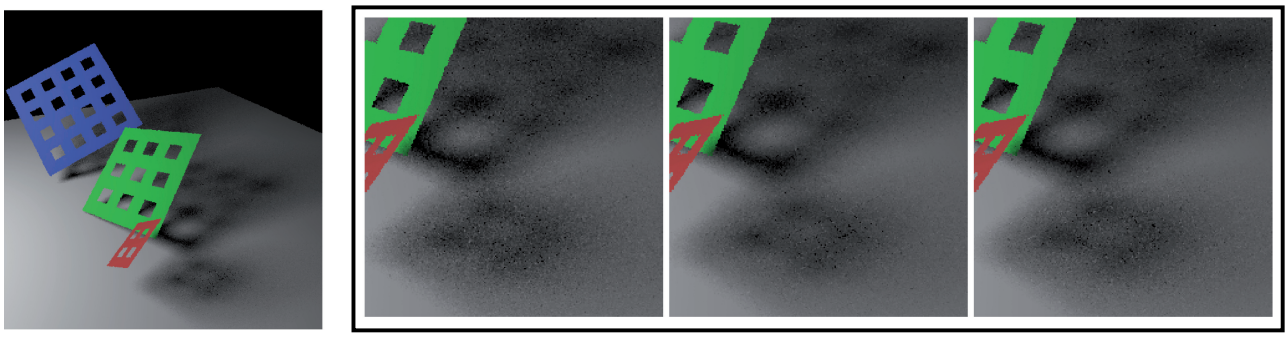
We have implemented two separate experimentation platforms: one using the (offline) RenderMan shading language, and another using NVIDIA's Optix GPU (real-time) ray-tracing API. The former is used for all of the images in this section, and allowed us to easily prototype different light source types and sampling patterns, as well as being readily applicable to standard RIB scene files. The supplementary video shows a few of the scenes using the Optix platform. The sampling patterns were easily integrated into a single `closest hit` kernel, allowing for interactive manipulation for visualization and comparison.

First, we show a simple scene with a sphere and a box in Figure 16(a) (courtesy Christophe Hery). In Figures 16(b) and (c), closeups compare uniform jitter and stratified sampling with 20 samples[8] for a circular light. We clearly see lower error for uniform jitter. Numerically, the errors are 3.5% for uniform jitter sampling, and 4.5% for stratified. As an equal quality comparison, in Figure 16(d) we show 25 samples using quasi-Monte Carlo with the Halton sequence, and the Shirley-Chiu warping transformation [1997]. We found empirically that this was the best competing method (which is verified in the literature), but uniform jitter sampling still performs approximately 25% better. As predicted by the theory, uniform jitter outperforms stratified and quasi-Monte Carlo sampling for linear, circular, and Gaussian light sources, but is worse for square lights (Figure 18).

Next, we show a more complicated shadowing pattern from three grids (this and the tentacles scene are courtesy Florian Hecht, from Egan et al. [2011]). Figure 17 shows the image with a circular light source. Uniform jitter with 20 samples in Figures 17(a) and (c) has substantially less noise than stratified with the same number of samples in Figure 17(b), and even than quasi-Monte Carlo sampling with the Shirley-Chiu transformation and 25 samples in Figure 17(d). This is a clear benefit of approximately 30% over conventional sampling methods. Similar benefits hold for a light with a Gaussian falloff (Figure 19(d), compare to (c)). Surprisingly, uniform jitter is even slightly better for square lights in this scene (Figure 19(b) has less noise than (a)). This is because of the complexity and irregularity of the shadows, as compared to some of the simpler scenes. The benefit for linear lights has already been shown in Figure 8. Finally, Figure 20 shows a graph of the error as a function of the number of samples, indicating a clear benefit for uniform jitter sampling on both circular and Gaussian lights.

In Figure 1, we show that our theory holds, even with large geometric models (scene composition courtesy Juan Buhler). The main image is shown with a circular light source and 20 samples (as with all of the examples in this article, a lower sampling count is used to allow the reader to better judge the noise patterns). Closeups and overall numerical RMS errors indicate a clear benefit for uniform

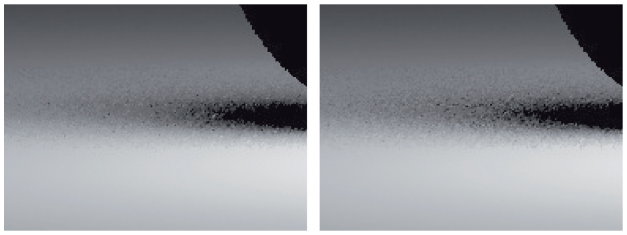
[8] In both cases, this is an average of 20 samples per pixel, and we use rejection sampling, with 25 original samples.



(a) grids: circle light, uniform jitter

(b) stratified, 20 samples (c) uniform jitter, 20 samples (d) halton, warp 25 samples
circular light: uniform jitter better than stratified

Fig. 17. Comparison of sampling methods for a scene with complex shadows from three grids (same geometry as in Figure 8) and a circular light source. As in Figure 16, images (a,c) are rendered with 20 uniform jitter samples. An equal sample comparison with stratified sampling (b) clearly indicates the benefit of uniform jitter (c), which is also somewhat better than even 25 samples with the Halton sequence and Shirley-Chiu [1997] warping (d).



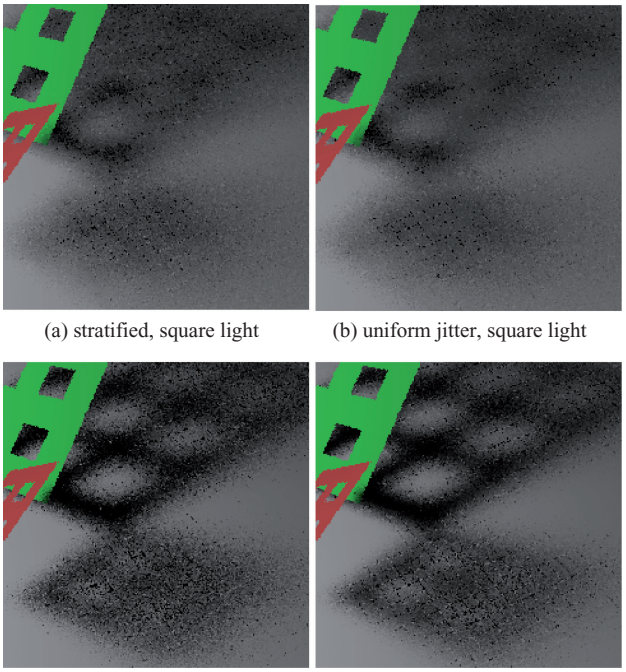
(a) stratified, 16 samples (b) uniform jitter, 16 samples
square light, uniform jitter worse than stratified

Fig. 18. Comparison of stratified and uniform jitter for a square light source for the sphere and box scene in Figure 16. For a square light, uniform jitter sampling performs worse than stratified sampling, as expected.

jitter over stratified sampling for circular lights (top row), while the situation is reversed for square lights (bottom row), as expected.

We also include a supplemental video that compares the different sampling strategies for different light source types, for different sample counts. These clips were captured in real time using our (unoptimized) Optix demo. These results confirm the still images shown in this section, and in particular the benefits of uniform jitter for circular lights. Temporal stability is maintained in a comparable fashion for all sampling methods as expected, since each frame is rendered separately: per-pixel random jitter buffers are reset when either the lighting, sampling scheme, or view are manipulated.

Finally, we show a limitation of the theory, with the very complex shadows from the tentacles scene in Figure 21 (also shown in the supplementary video). In this case, there are multiple visibility discontinuities within a stratum, and so the assumption of isolated discontinuities in Section 4 does not hold. In fact, stratification itself gives only a minor benefit for such complex cases [Mitchell 1996], and we also compare to pure random sampling (see graphs in Figures 21(e) and (f)). The closeups and graphs validate that while uniform jitter is marginally better, all methods including purely random sampling perform almost the same. In these very difficult cases, a more complicated approach of exploiting interpixel and shadow light field coherence as in Egan et al. [2011] seems to be the only possibility for further speedups. We emphasize, however, that such complex visibility patterns are not common for most parts of ordinary indoor and most outdoor scenes.



(a) stratified, square light (b) uniform jitter, square light

(c) stratified, gaussian falloff (d) uniform jitter, gaussian falloff

Fig. 19. The same insets as in Figure 17, this time for square lights without (a,b) and with (c,d) a Gaussian falloff. Uniform jitter is clearly better for Gaussian lights (d is better than c), and in this example with its complex geometry and irregular shadows, surprisingly even for square lights (b has less noise than a). All of these examples use 16 samples per pixel.

8. DISCUSSION AND FUTURE WORK

In this article, we focus on stratified sampling patterns, since those are very commonly used, and are shown by the theory to provide optimal results in many cases. Sampling methods based on blue noise [Yellot 1983] are attractive alternatives. These include both the Poisson Disk methods [Dippe and Wold 1985; Cook 1986], for which many recent fast sample generation schemes have been proposed [Dunbar and Humphreys 2006; Wei 2008], as well as

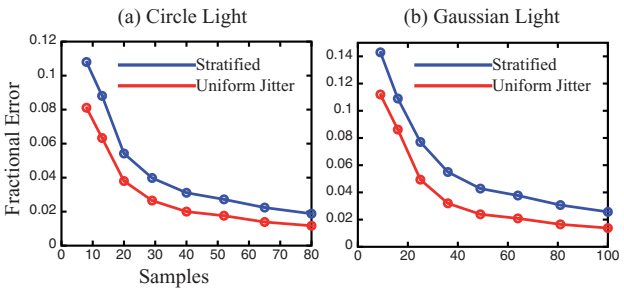


Fig. 20. Graphs of fractional error vs. number of samples for the grids scene, for a circular light (a) and a Gaussian light (b). Uniform jitter provides a clear benefit in both cases over stratified sampling.

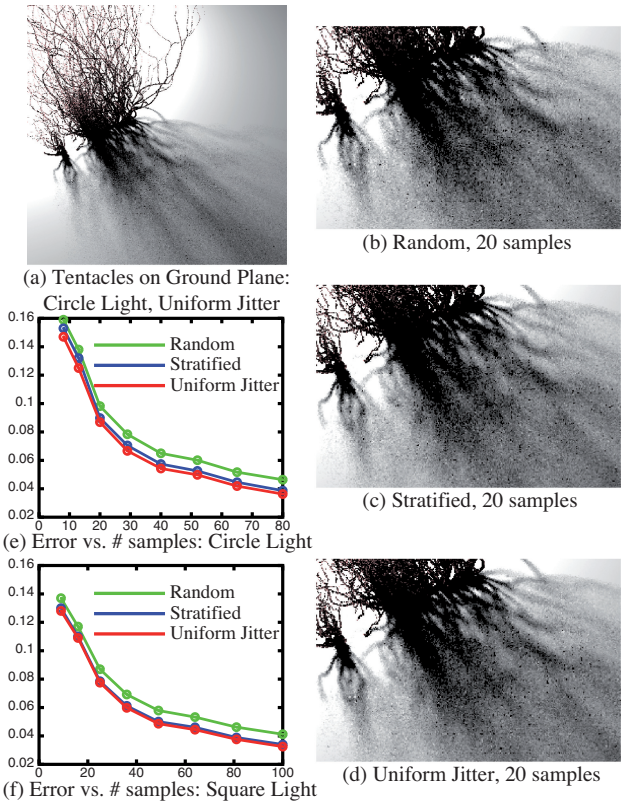


Fig. 21. A scene with very complex shadows from the thin geometry of the tentacles (a). The closeups shown on the right (for a circular light, with 20 samples) are comparable for purely random sampling, stratified, and uniform jitter (b-d). The graphs (e-f) show that essentially any sampling scheme performs not much better than random sampling, given this complexity in shadows. Fortunately, this case is not typical in common scenes.

Quasi-Monte Carlo (QMC) approaches [Niederreiter 1992; Keller 1997] that exhibit blue-noise properties [Ouellette and Fiume 2001].

In the limit that the minimum Poisson Disk sample spacing is set to the stratum length, these methods essentially reduce to uniform jitter sampling (technically with a hexagonal packing in 2D rather than a grid), with very similar properties. In practice, however, the relative radius is typically set to $\rho = 0.75$ of this value (Lagae and Dutre [2008] recommend $\rho = [0.65 \ldots 0.85]$) in order to avoid regular configurations. For linear lights, the advantages of uniform jitter stem from correlated sample placements in the strata where visibility discontinuities occur. Poisson disk sampling does not guarantee this correlation, given the more irregular distribution of samples. Moreover, since $\rho = 0.75 < 1$, the "band-gap" in Figure 11 around the DC term will be smaller than for uniform (jitter) sampling, even though it will be larger than for stratified sampling. Moreover, in our initial tests, we found that the more uneven spacing and stratification caused by $\rho < 1$ leads to larger errors than the sampling strategies (both stratified and uniform jitter) considered in this article. Further investigation of these insights is worthwhile.

We have briefly touched on QMC approaches like the Halton sequence earlier in the article; there are also many other sequences described in the literature. While the canonical Halton sequence performs poorly in Figure 7, one can jitter the offset independently at each pixel for rendering much like uniform jitter, in which case Halton will perform very similarly to uniform jitter sampling. This leaves open the possibility that another suitably chosen QMC sequence could be competitive with uniform jitter even for linear and circular lights; initial results in this direction are encouraging.

More generally, the theory of this article indicates that uniform sampling often has low numerical errors. Uniform jitter seeks to preserve the numerical properties while removing banding. But both methods can give poor performance in 2D for square lights. Techniques like QMC and blue-noise sampling offer some of the advantages of both worlds, with near-uniform sample spacing, while removing some of the correlation effects; however, this article indicates they are not always competitive. Further research is needed to see if modifications of these approaches can provide the best performance for visibility sampling. Moreover, all of our work has looked at the constants involving errors in different sampling methods. The convergence rate itself is the same, and not affected (for example, variance decays as $1/N^2$ in all cases in Figures 4 and 6); we do not believe this can be improved.

Finally, we have provided only initial guidance about different sampling patterns for square and circular lights. It remains an interesting avenue of work to consider a broader class of light source shapes and develop a general theory of when different sampling methods are expected to work better. Moreover, while we have tested the predictions with a range of different blockers, as seen in our figures, it is still possible the shape of the occluder also plays some role.

9. CONCLUSIONS

In this article, we have derived a comprehensive theory of Monte Carlo visibility sampling. Instead of considering a general shading integral, we focus on the hardest component, the binary visibility function. This allows us to derive novel statistical insights. Starting with the simplest case of a linear light source, we show how the number of discontinuities in the visibility pattern can affect the efficiency of sampling, and that the new approach of uniform jitter sampling is often preferred for more complex visibility.

Perhaps the most important contribution of this article is the development of new analytical tools for analyzing sampling. Besides the statistical analysis in Section 4, we introduce a new Fourier approach in Section 5 for analyzing Monte Carlo integration in *both* the spatial (pixel) and angular (light) domain together. This analysis precisely identifies the nature of banding in uniform sampling, by focusing error in specific spatial frequencies. We also see how the error is diffused out to alleviate banding for uniform jitter sampling. Finally, we show in Section 6 that the extension to planar area lights surprisingly depends on the shape of the light source, with uniform jitter sampling preferred for linear, circular, and Gaussian lights,

but not rectilinear sources. While the benefits of different sampling patterns are relatively modest, we show practically that gains of 20–40% can be achieved with very minimal modifications in code. This understanding will allow more efficient rendering algorithms targeted to the specific scene and light source.

An interesting observation is that the shape of the light source has not so far been a major factor in shadow algorithms based on Monte Carlo sampling. Perhaps the most common approach is to use a square light source by default. However, our analysis has shown that the sampling strategy can differ significantly for different lights, and indicates that linear and circular lights may therefore be preferable to square lights in some applications.

In the future, we believe a thorough analysis of different sampling patterns can shed key insights on Monte Carlo integration for other shading problems. We believe our insights on 2D pixel-light Fourier analysis, and considering the shape and spectrum of the light source, can have many implications for other challenging topics like motion blur, glossy highlights, and depth of field.

ACKNOWLEDGMENTS

We acknowledge many inspiring early discussions with Rob Cook, especially on formulating the pixel-light 2D frequency domain view of sampling. Fredo Durand first inspired us with his notes on the frequency analysis of Monte Carlo integration. The anonymous reviewers provided extensive helpful comments on a very complicated manuscript that has allowed us to be clearer about many of the issues. Li-Yi Wei and Christophe Hery provided input for the discussion of blue noise and general sampling schems. Some of the scenes used in this article are courtesy of Kevin Egan, Florian Hecht, Christophe Hery and Juan Buhler.

REFERENCES

AGRAWALA, M., RAMAMOORTHI, R., HEIRICH, A., AND MOLL, L. 2000. Efficient image-based methods for rendering soft shadows. In *Proceedings of the ACM SIGGRAPH 00 Conferrence*. 375–384.

BEN-ARTZI, A., RAMAMOORTHI, R., AND AGRAWALA, M. 2006. Efficient shadows for sampled environment maps. *J. Graph. Tools 11,* 1, 13–36.

CANDES, E. 2006. Compressive sampling. In *Proceedings of the International Congress of Mathematics*. Number 3, 1433–1452.

CANDES, E., ROMBERG, J., AND TAO, T. 2006. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math. 59,* 8, 1207–1223.

CANDES, E. AND TAO, T. 2006. Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inf. Theory 52,* 12, 5406–5425.

COOK, R. 1986. Stochastic sampling in computer graphics. *ACM Trans. Graph. 5,* 1, 51–72.

DIPPE, M. AND WOLD, E. 1985. Antialiasing through stochastic sampling. In *Proceedings of the ACM SIGGRAPH 85 Conference*. 69–78.

DUNBAR, D. AND HUMPHREYS, G. 2006. A spatial data structure for fast poisson-disk sample generation. *ACM Trans. Graph. 25,* 3, 503–508.

DURAND, F. 2011. A frequency analysis of monte-carlo and other numerical integration schemes. Tech. rep. MIT-CSAIL-TR-2011-052 http://hdl.handle.net/1721.1/67677, MIT CSAIL.

DURAND, F., DRETTAKIS, G., AND PUECH, C. 1997. The visibility skeleton: A powerful and efficient multi-purpose global visibility tool. In *Proceedings of the ACM SIGGRAPH 97 Conference*. 89–100.

DURAND, F., HOLZSCHUCH, N., SOLER, C., CHAN, E., AND SILLION, F. 2005. A frequency analysis of light transport. *ACM Trans. Graph. 25,* 3, 1115–1126.

EGAN, K., HECHT, F., DURAND, F., AND RAMAMOORTHI, R. 2011. Frequency analysis and sheared filtering for shadow light fields of complex occluders. *ACM Trans. Graph. 30,* 2.

GUO, B. 1998. Progressive radiance evaluation using directional coherence maps. In *Proceedings of the ACM SIGGRAPH 98 Conference*. 255–266.

HACHISUKA, T., JAROSZ, W., WEISTROFFER, R., DALE, K., HUMPHREYS, G., ZWICKER, M., AND JENSEN, H. 2008. Multidimensional adaptive sampling and reconstruction for ray tracing. *ACM Trans. Graph. 27,* 3.

HEINRICH, S. AND KELLER, A. 1994. Quasi-Monte carlo methods in computer graphics. Tech. rep. 242/3, University of Kaiserslautern.

KELLER, A. 1997. Instant radiosity. In *Proceedings of the ACM SIGGRAPH 97 Conference*. 49–56.

LAGAE, A. AND DUTRE, P. 2008. A comparison of methods for generating poisson disk patterns. *Comput. Graph. Forum 27,* 1, 114–129.

LANMAN, D., RASKAR, R., AGRAWAL, A., AND TAUBIN, G. 2008. Shield fields: modeling and capturing 3D occluders. *ACM Trans. Graph. 27,* 5.

LEE, M., REDNER, A., AND USELTON, S. 1985. Statistically optimized sampling for distributed ray tracing. In *Proceedings of the ACM SIGGRAPH 85 Conference*. 61–68.

MITCHELL, D. 1987. Generating antialiased images at low sampling densities. In *Proceedings of the ACM SIGGRAPH 87 Conference*. 65–72.

MITCHELL, D. 1991. Spectrally optimal sampling for distribution ray tracing. In *Proceedings of the ACM SIGGRAPH 91 Conference*. 157–164.

MITCHELL, D. 1996. Consequences of stratified sampling in graphics. In *Proceedings of the ACM SIGGRAPH 96 Conference*. 277–280.

NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2004. Triple product wavelet integrals for all-frequency relighting. *ACM Trans. Graph. 23,* 3, 475–485.

NIEDERREITER, H. 1992. *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM.

OUELLETTE, M. AND FIUME, E. 2001. On numerical solutions to one-dimensional integration problems with application to linear light sources. *ACM Trans. Graph. 20,* 4, 232–279.

OVERBECK, R., DONNER, C., AND RAMAMOORTHI, R. 2009. Adaptive Wavelet Rendering. *ACM Trans. Graph. 28,* 5.

PURGATHOFER, W. 1986. A statistical model for adaptive stochastic sampling. In *Proceedings of the Eurographics Conference*. 145–152.

RAMAMOORTHI, R., KOUDELKA, M., AND BELHUMEUR, P. 2005. A Fourier theory for cast shadows. *IEEE Trans. Pattern Anal. Mach. Intell. 27,* 2, 288–295.

RAMAMOORTHI, R., MAHAJAN, D., AND BELHUMEUR, P. 2007. A first order analysis of lighting, shading, and shadows. *ACM Trans. Graph. 26,* 1.

SEN, P. AND DARABI, S. 2010. Compressive estimation for signal integration in rendering. *Comput. Graph. Forum 29,* 4, 1355–1363.

SHIRLEY, P. AND CHIU, K. 1997. A low distortion map between disk and square. *J. Graph. Tools 2,* 3, 45–52.

SOLER, C. AND SILLION, F. 1998. Fast calculation of soft shadow textures using convolution. In *Proceedings of the ACM SIGGRAPH 98 Conference*. 321–332.

WEI, L. 2008. Parallel poisson disk sampling. *ACM Trans. Graph. 27,* 3.

WEI, L. 2010. Multi-Class blue noise sampling. *ACM Trans. Graph. 29,* 4.

YELLOT, J. 1983. Spectral consequences of photoreceptor sampling in the rhesus retina. *Science 221*, 382–385.

Received August 2011; revised December 2011; accepted February 2012