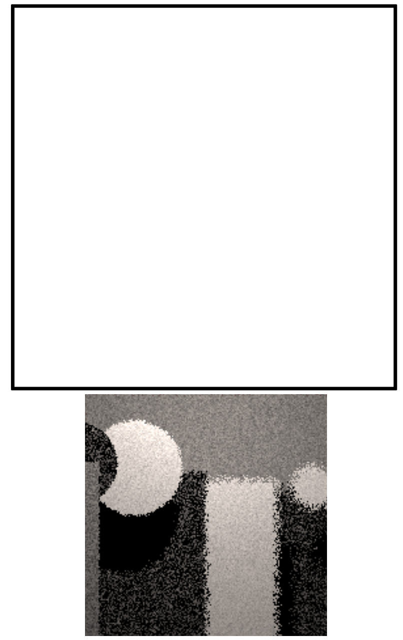
Error Analysis of Common Sampling Stratgies

Gurprit Singh

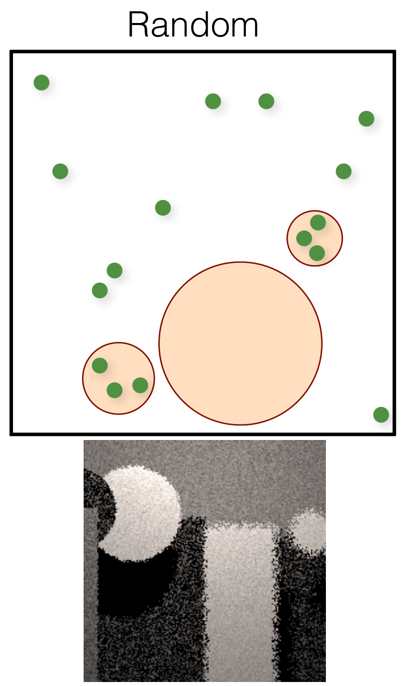
gsingh@mpi-inf.mpg.de

max planck institut
informatik

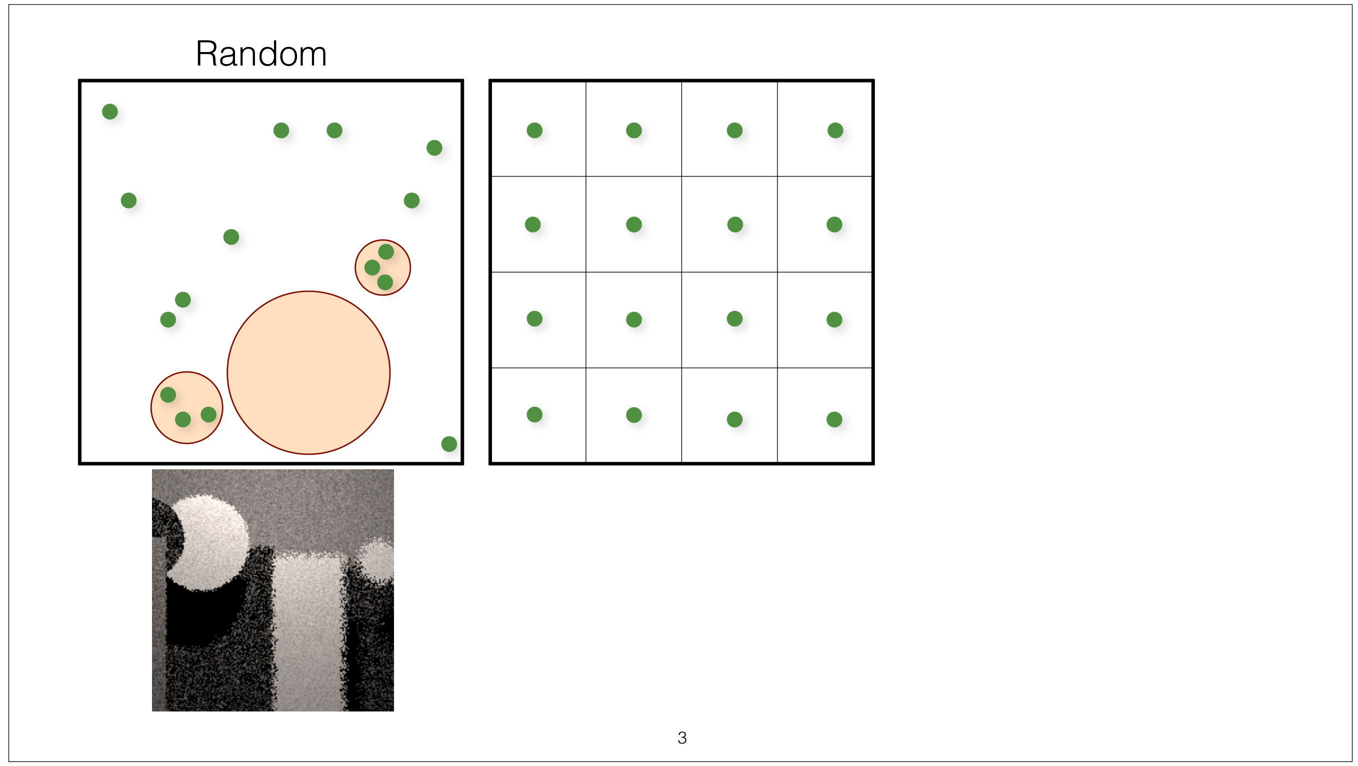
Random



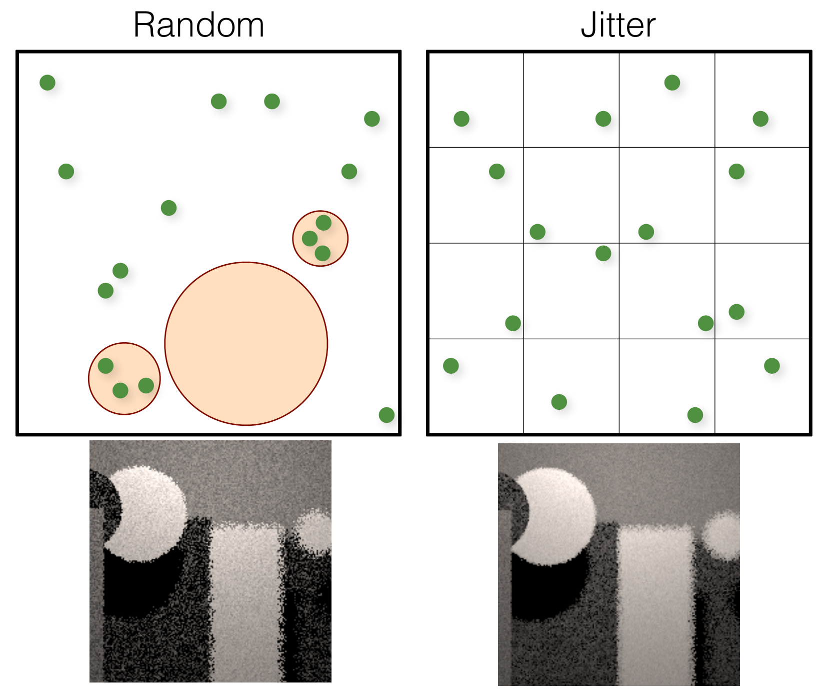
We start with random sampling [CLICK] which involves randomly generating samples in the integration domain, the noise level is significantly high. We can improve this sample distribution by...



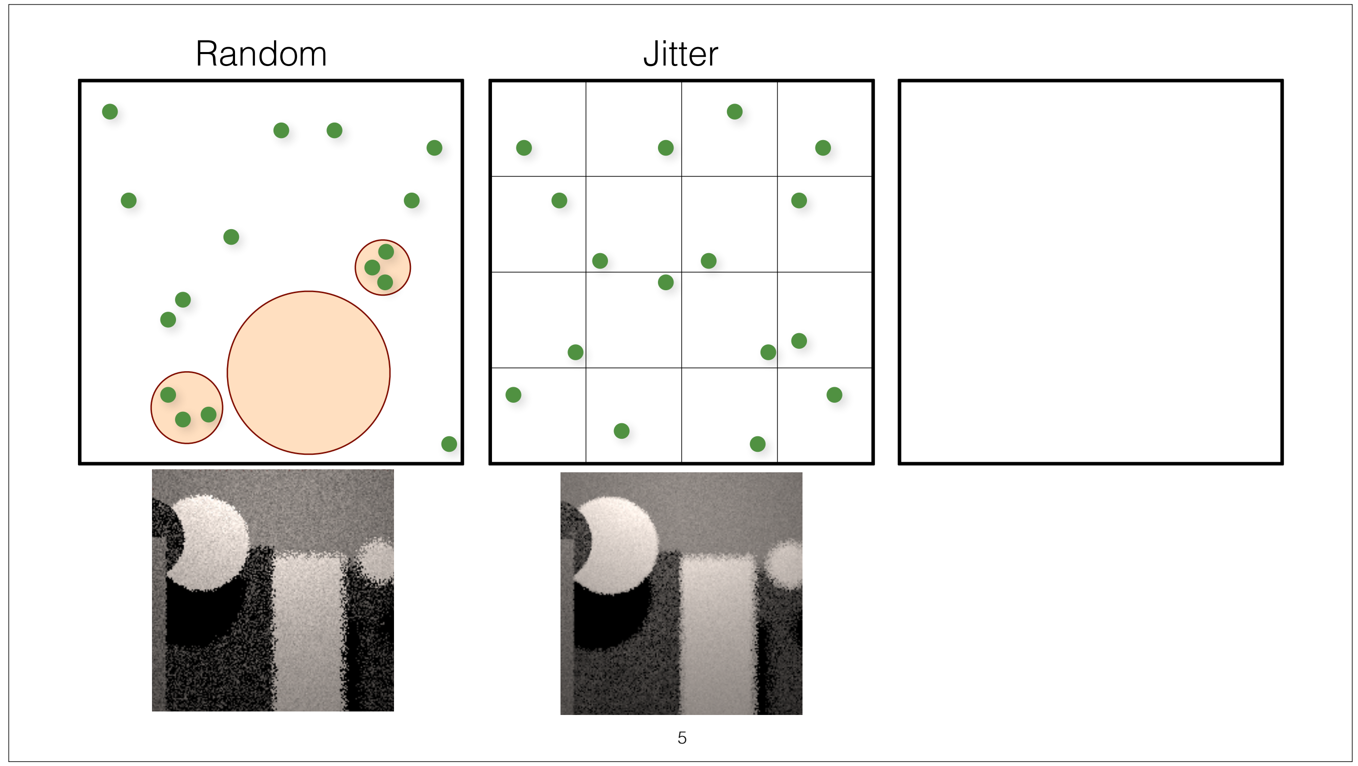
We start with random sampling [CLICK] which involves randomly generating samples in the integration domain, the noise level is significantly high. We can improve this sample distribution by...



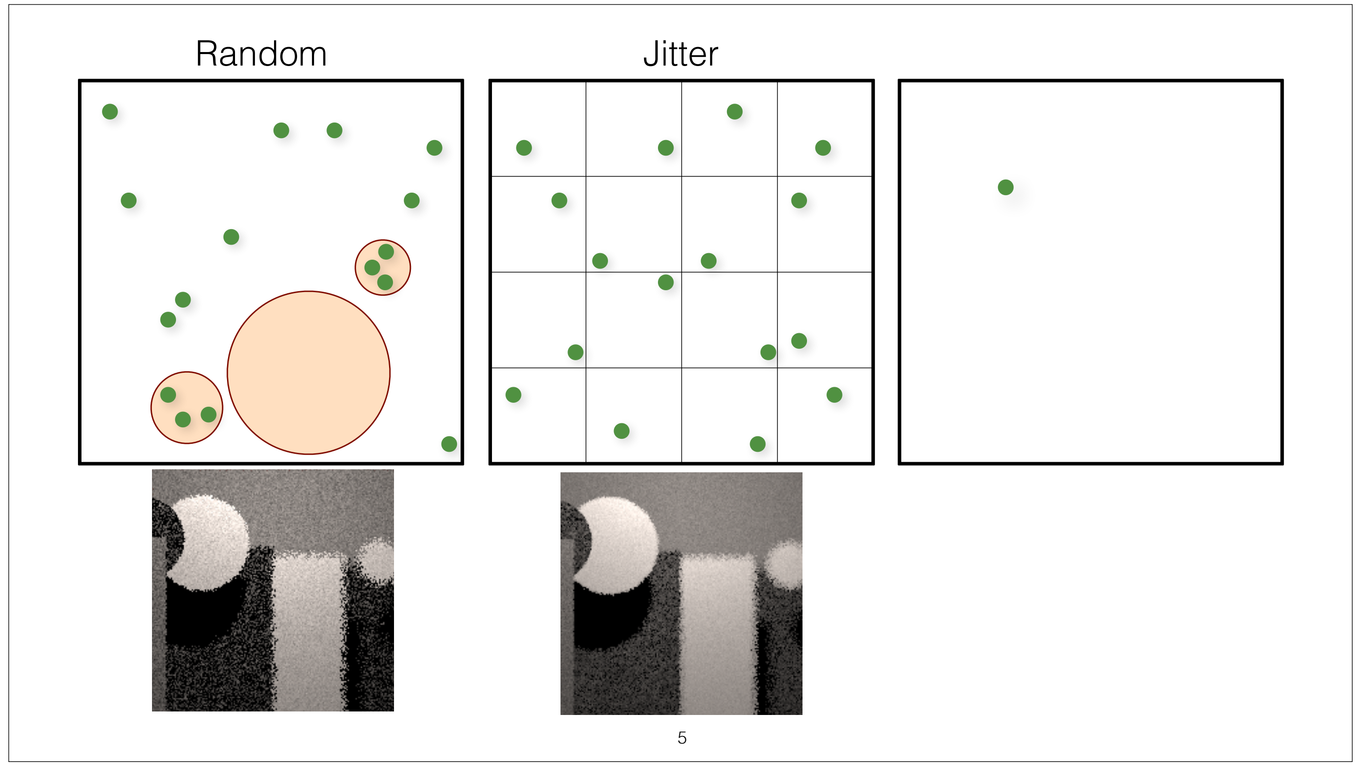
...dividing the domain in equal strata and placing the samples at the center of each stratum. The regularity has been known to cause some aliasing effects which can be easily avoided by randomly jittering...



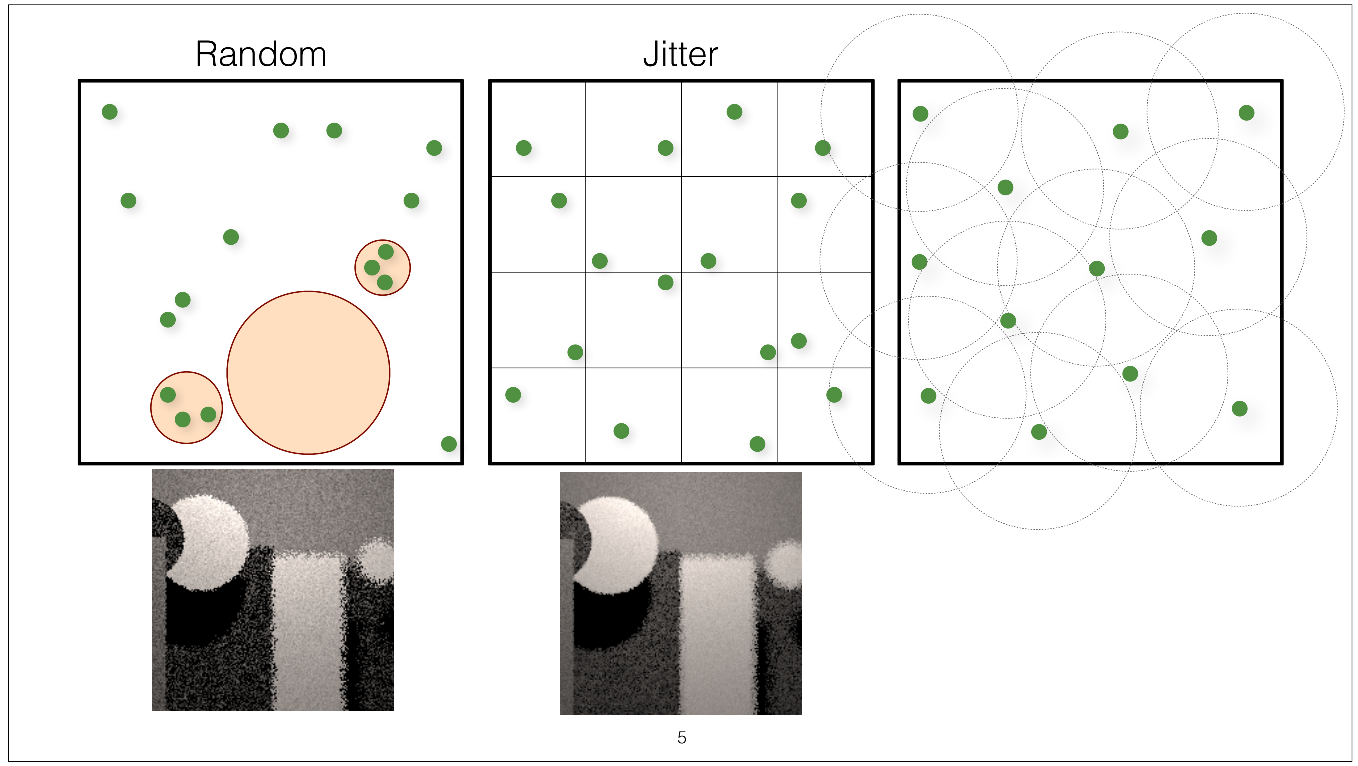
...each sample independently in each stratum. This is called jittered sampling and as you might notice the noise level has already improved for the same sample count. We can further improve the uniformity of samples by, say,...



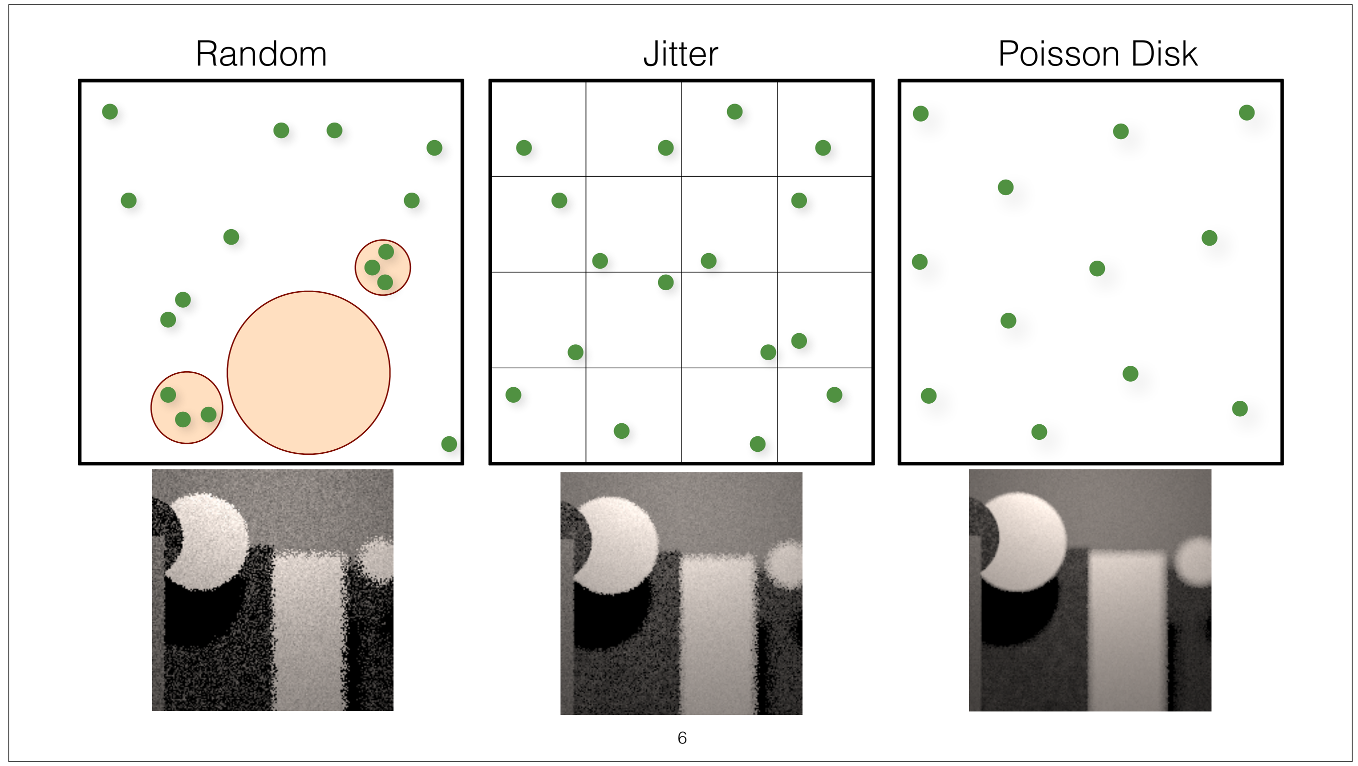
...generating a sample [CLICK] in the domain following a naive dart throwing approach where a radius is assigned to each sample and a new sample is only accepted if it falls outside the disk radius. This gives us Poisson disk samples...



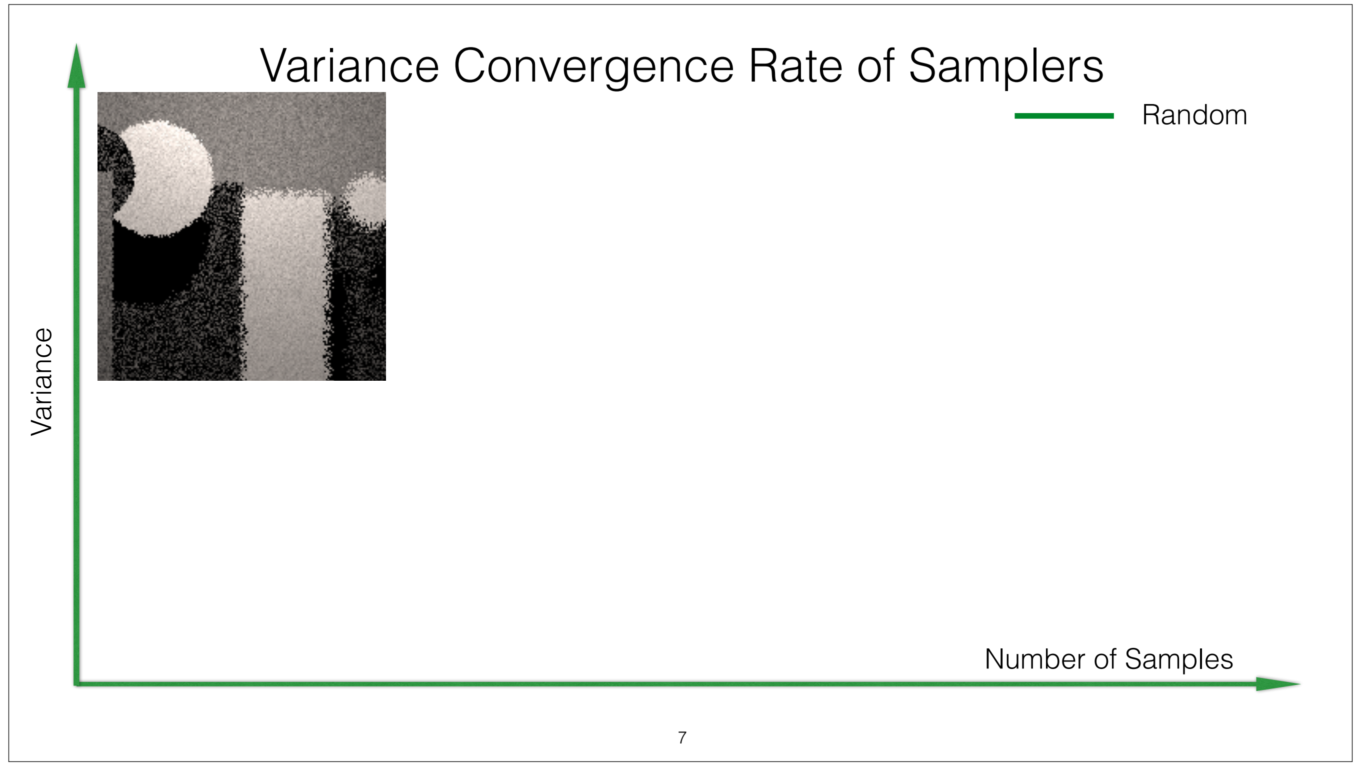
...generating a sample [CLICK] in the domain following a naive dart throwing approach where a radius is assigned to each sample and a new sample is only accepted if it falls outside the disk radius. This gives us Poisson disk samples...



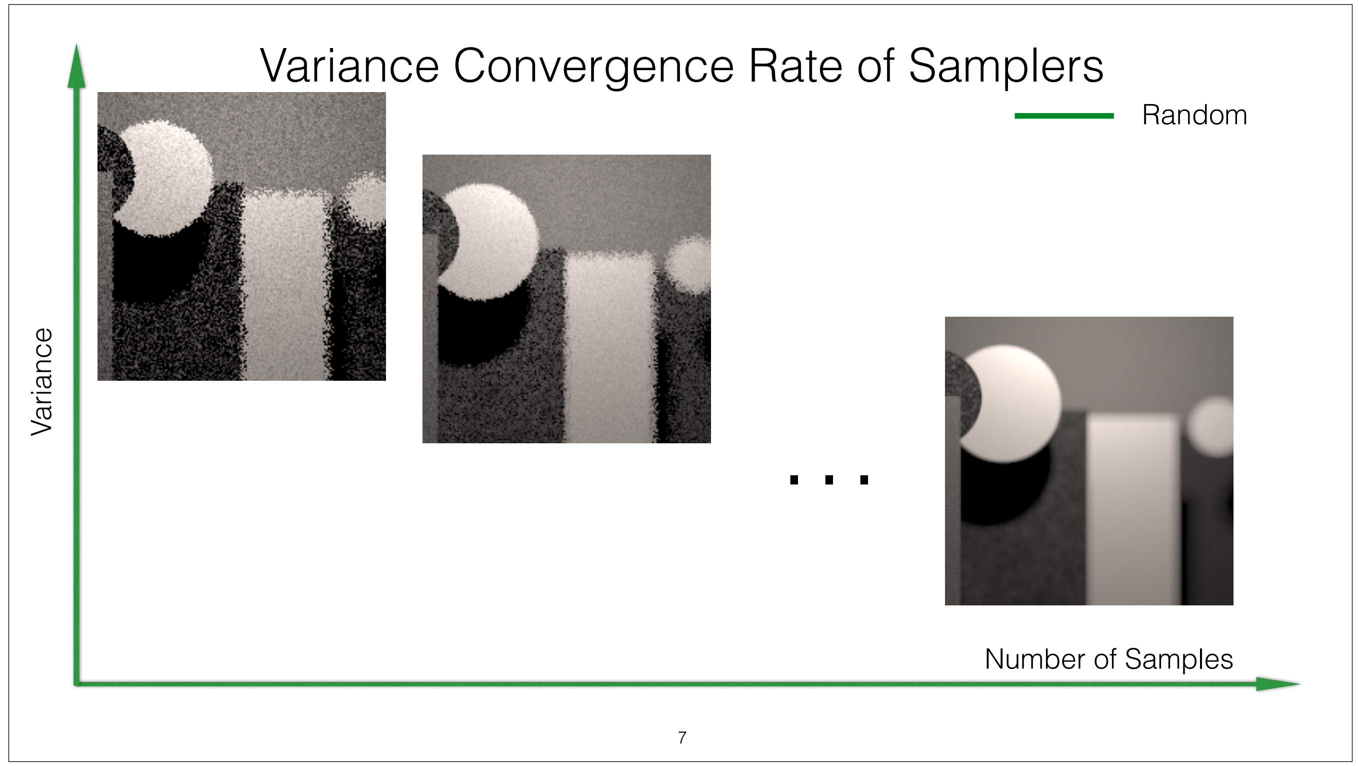
...generating a sample [CLICK] in the domain following a naive dart throwing approach where a radius is assigned to each sample and a new sample is only accepted if it falls outside the disk radius. This gives us Poisson disk samples...



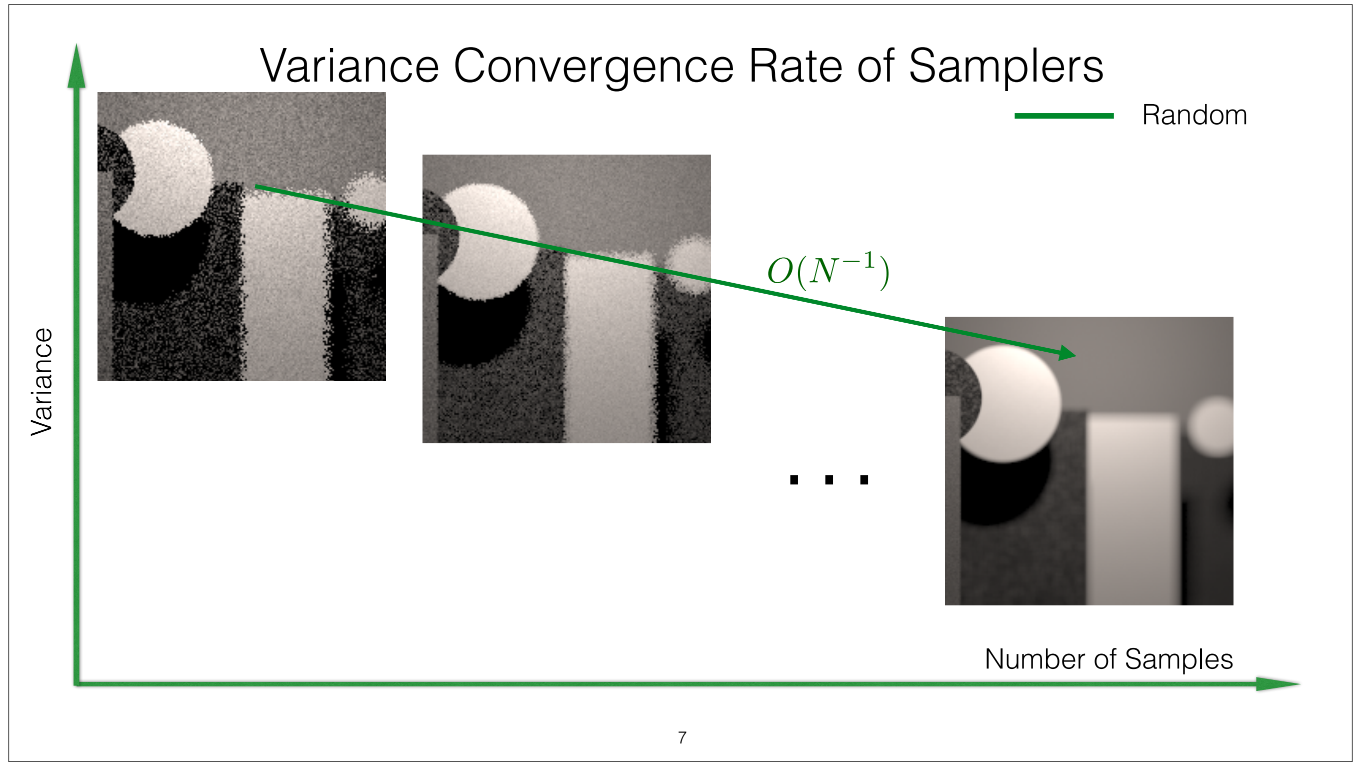
... which is well distributed and the corresponding noise in the image has also gone down for the same number of samples. Another way to decrease variance is to keep...



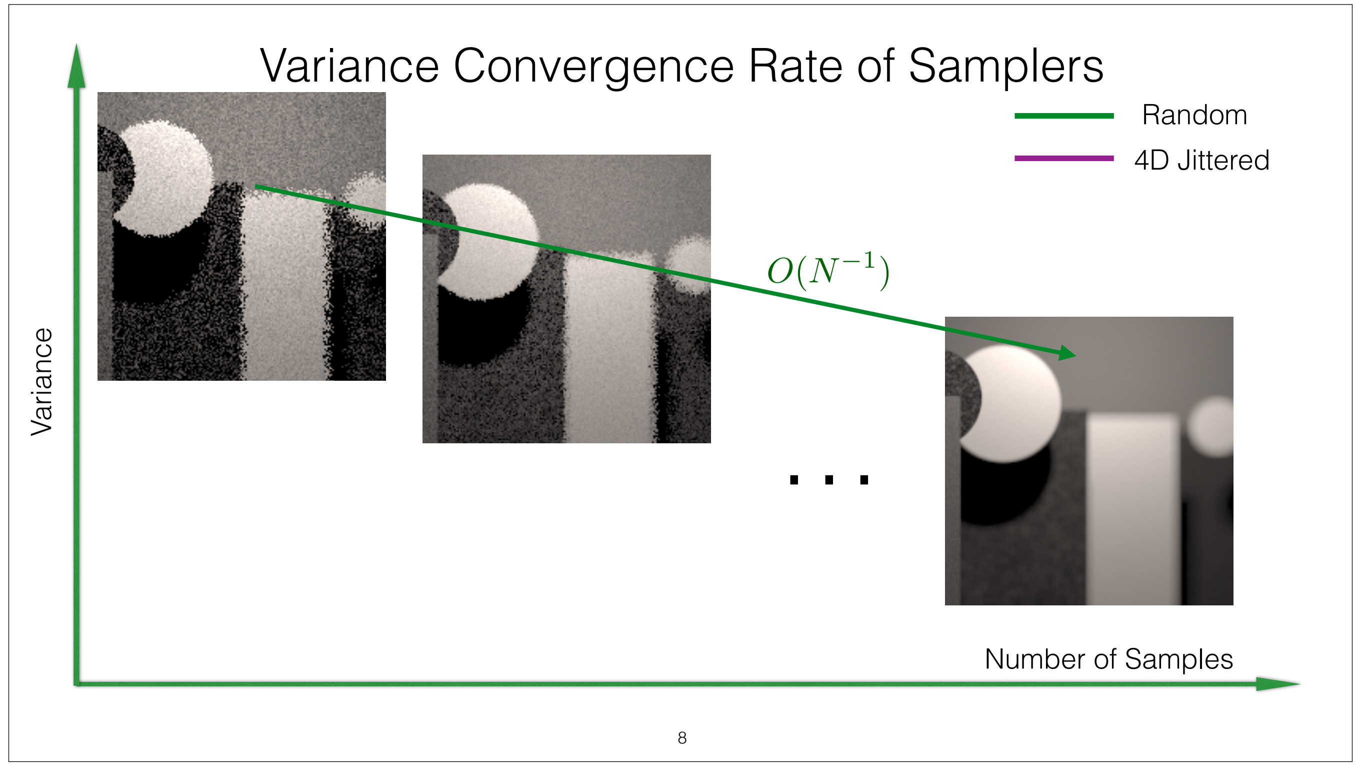
…increase the sample count till the image becomes noise-free (converge). The rate at which this image converges depend on the underlying sampling pattern used. For example, with random samples [CLICK] the convergence rate is always O(N^-1), 4D jittered samples [CLICK], we would obtain a 4D convergence rate of O(N^-1.25) whereas with Poisson disk…



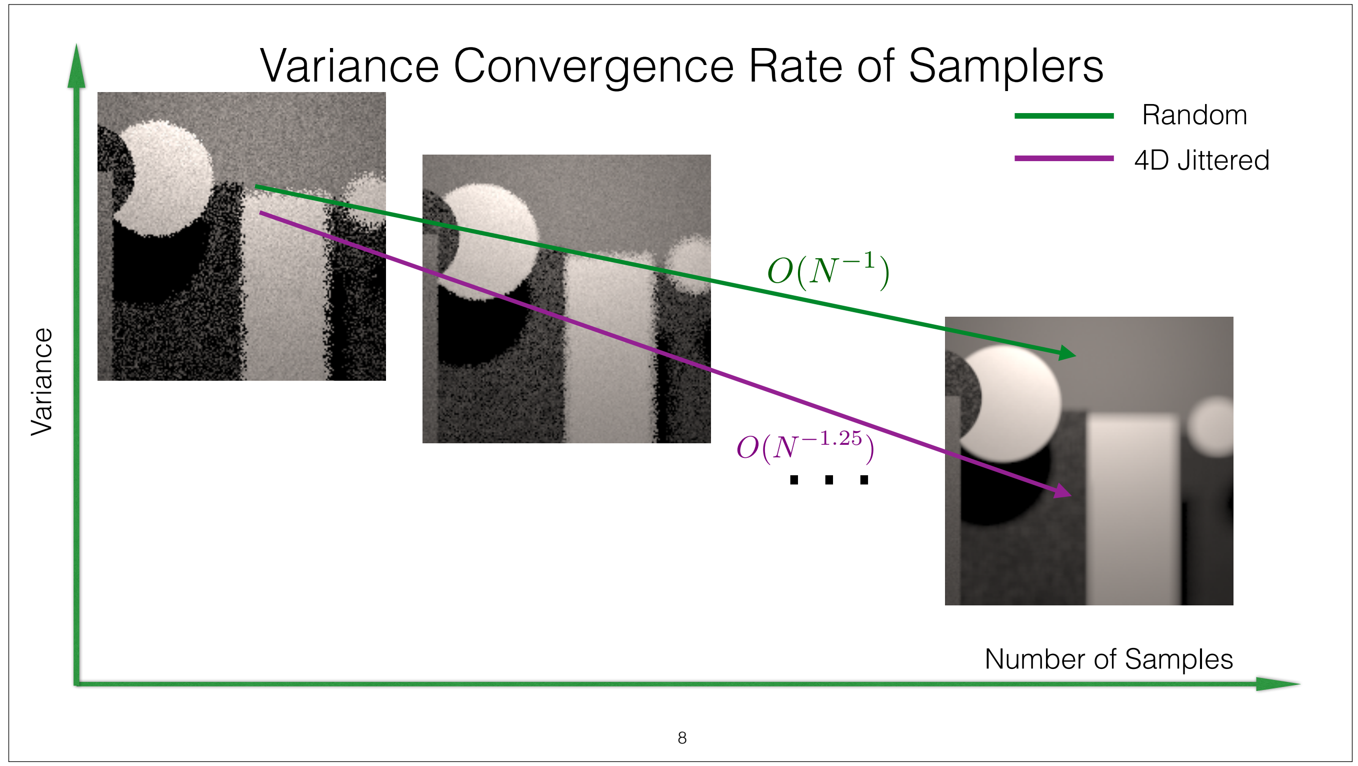
…increase the sample count till the image becomes noise-free (converge). The rate at which this image converges depend on the underlying sampling pattern used. For example, with random samples [CLICK] the convergence rate is always O(N^-1), 4D jittered samples [CLICK], we would obtain a 4D convergence rate of O(N^-1.25) whereas with Poisson disk…



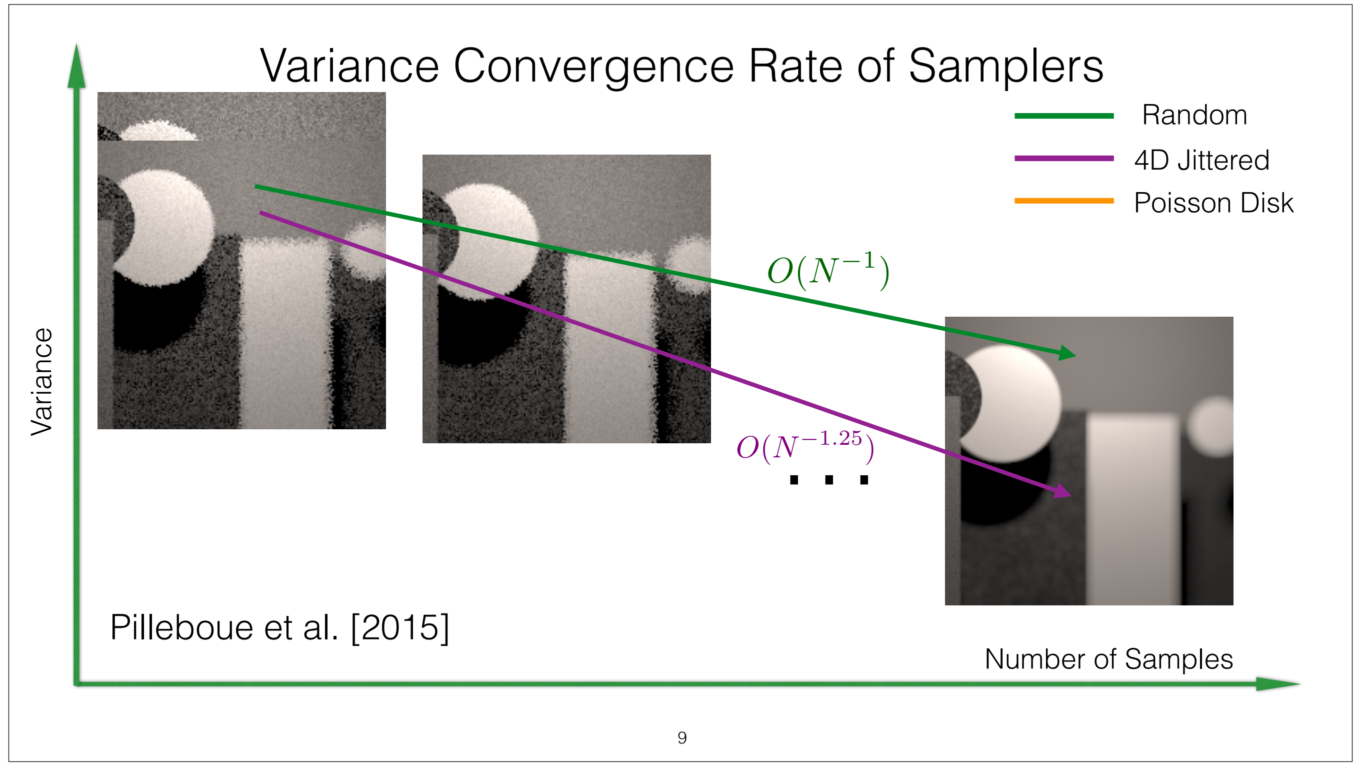
…increase the sample count till the image becomes noise-free (converge). The rate at which this image converges depend on the underlying sampling pattern used. For example, with random samples [CLICK] the convergence rate is always O(N^-1), 4D jittered samples [CLICK], we would obtain a 4D convergence rate of O(N^-1.25) whereas with Poisson disk…



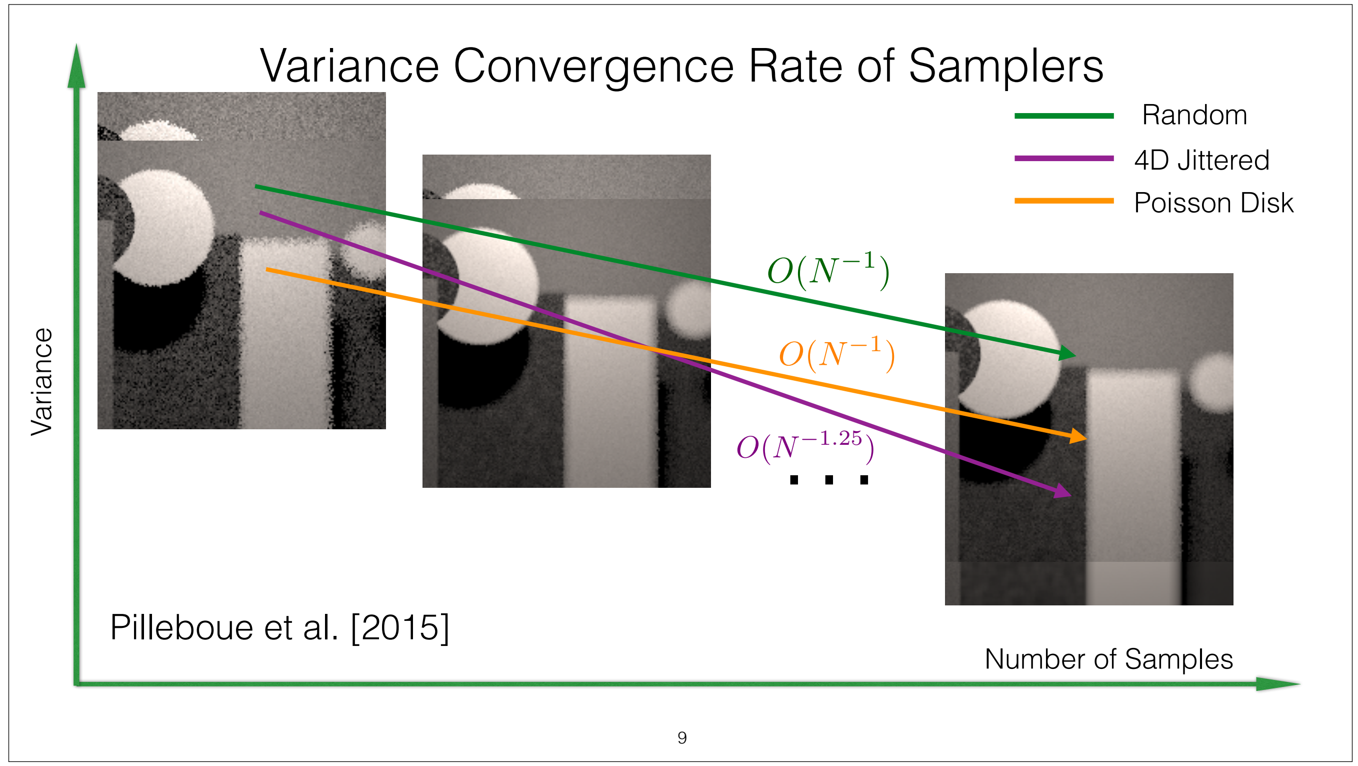
…increase the sample count till the image becomes noise-free (converge). The rate at which this image converges depend on the underlying sampling pattern used. For example, with random samples [CLICK] the convergence rate is always O(N^-1), 4D jittered samples [CLICK], we would obtain a 4D convergence rate of O(N^-1.25) whereas with Poisson disk…



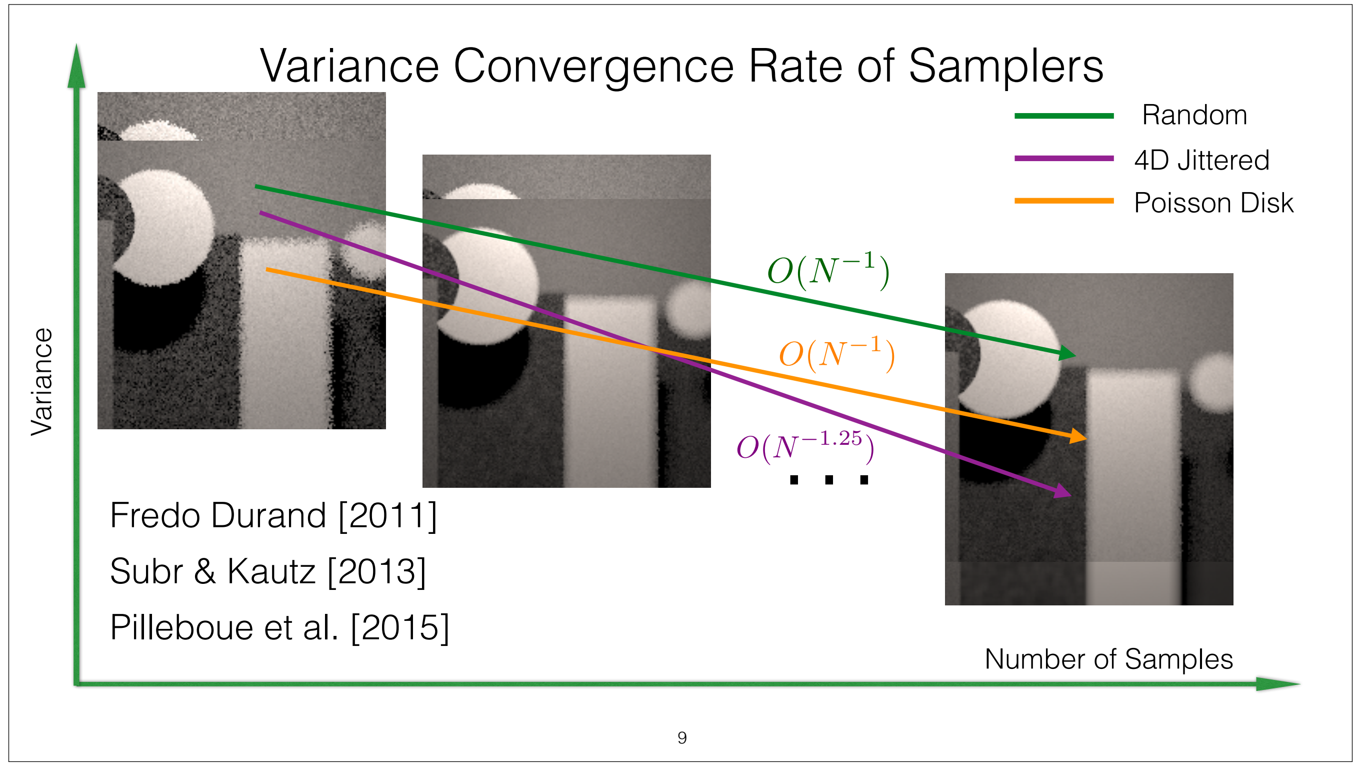
…increase the sample count till the image becomes noise-free (converge). The rate at which this image converges depend on the underlying sampling pattern used. For example, with random samples [CLICK] the convergence rate is always O(N^-1), 4D jittered samples [CLICK], we would obtain a 4D convergence rate of O(N^-1.25) whereas with Poisson disk…



… we can obtain less noisy images at small sample count but [CLICK] as we increase the samples the convergence rate obtained is O(N^-1).
This reflects that if the sampling budget is small (which is the case in many interactive applications) it is best to use Poisson disk samplers, whereas for large sampling budget, for example in offline rendering for movie frames, we should consider samplers with good convergence.
These convergence rates can be empirically computed using the sample variance. We proposed a mathematical convergence tool in the Fourier domain, which allows to theoretically derive these convergence rates for blue noise samples. Now, lets try to understand what characteristics of these samples are affecting the quality of the rendered images.

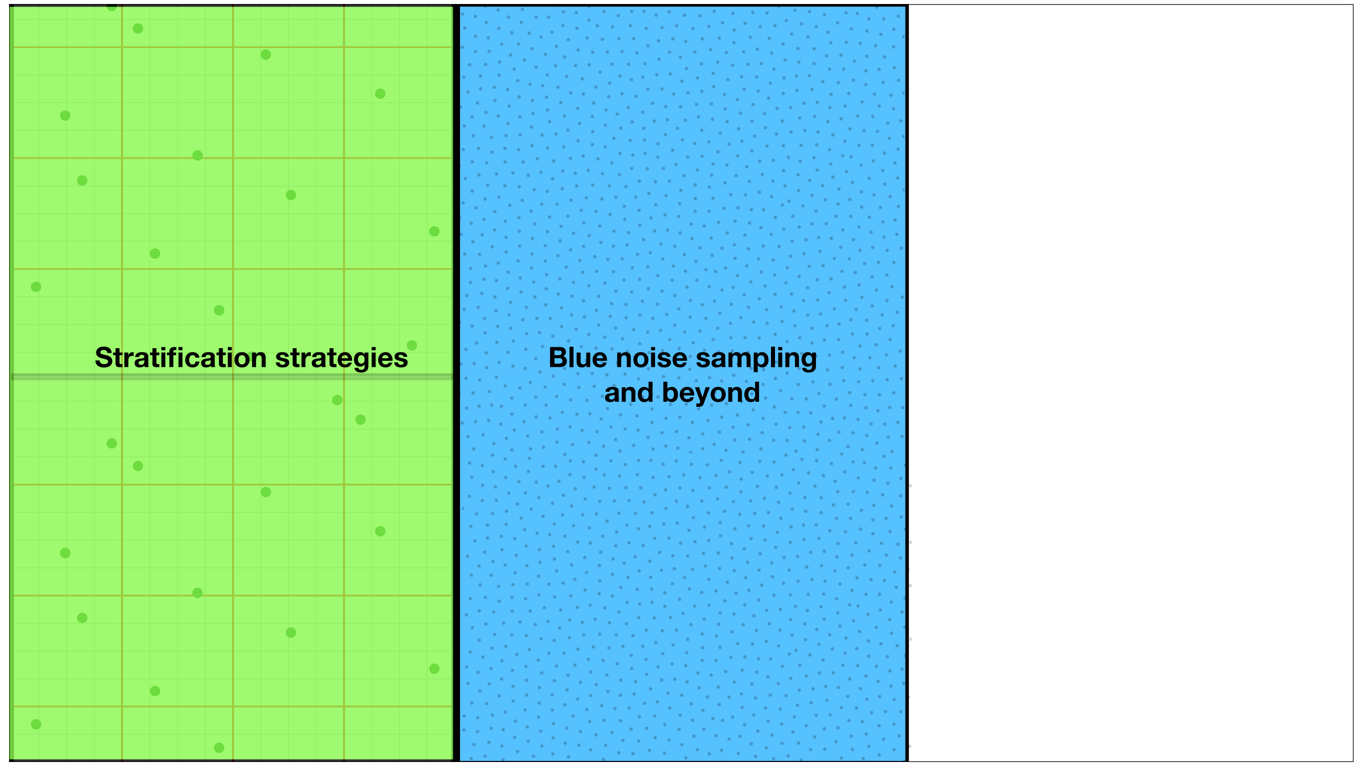


… we can obtain less noisy images at small sample count but [CLICK] as we increase the samples the convergence rate obtained is O(N^-1).
This reflects that if the sampling budget is small (which is the case in many interactive applications) it is best to use Poisson disk samplers, whereas for large sampling budget, for example in offline rendering for movie frames, we should consider samplers with good convergence.
These convergence rates can be empirically computed using the sample variance. We proposed a mathematical convergence tool in the Fourier domain, which allows to theoretically derive these convergence rates for blue noise samples. Now, lets try to understand what characteristics of these samples are affecting the quality of the rendered images.

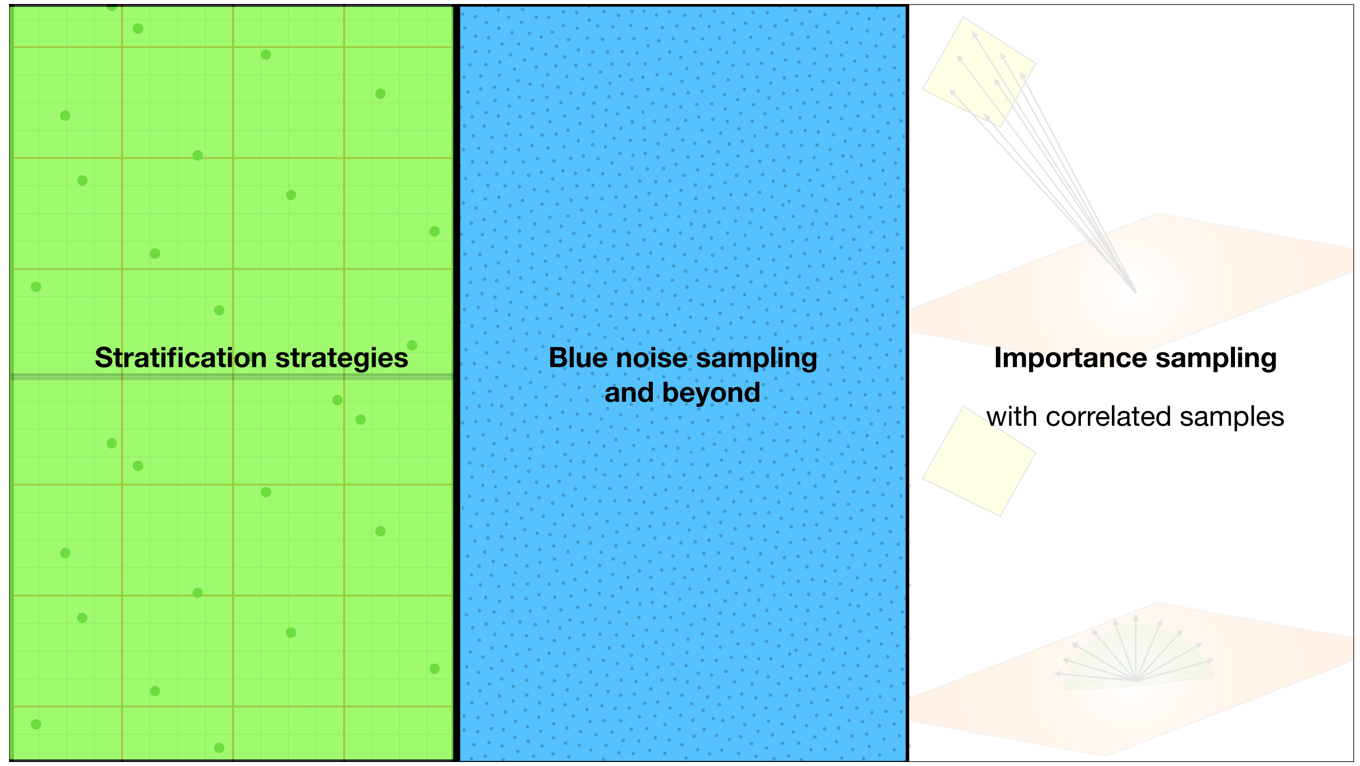


… we can obtain less noisy images at small sample count but [CLICK] as we increase the samples the convergence rate obtained is O(N^-1).
This reflects that if the sampling budget is small (which is the case in many interactive applications) it is best to use Poisson disk samplers, whereas for large sampling budget, for example in offline rendering for movie frames, we should consider samplers with good convergence.
These convergence rates can be empirically computed using the sample variance. We proposed a mathematical convergence tool in the Fourier domain, which allows to theoretically derive these convergence rates for blue noise samples. Now, lets try to understand what characteristics of these samples are affecting the quality of the rendered images.

We start by looking into error introduced by different stratification strategies
[CLICK] We then analyze error due to blue noise distributions and how this idea is extended beyond to introduce discrepancy related measures
[CLICK] In the end, we will briefly look at how to analyze importance sampling using the presented Fourier tools and how correlated samples could affect the convergence properties of these importance sampling strategies.



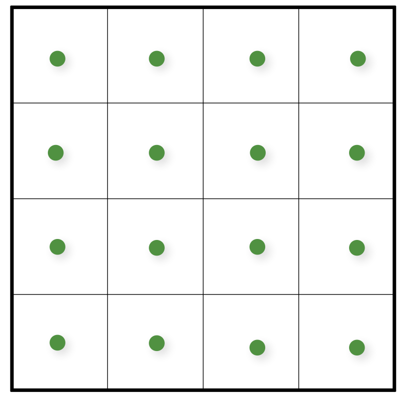
We start by looking into error introduced by different stratification strategies
[CLICK] We then analyze error due to blue noise distributions and how this idea is extended beyond to introduce discrepancy related measures
[CLICK] In the end, we will briefly look at how to analyze importance sampling using the presented Fourier tools and how correlated samples could affect the convergence properties of these importance sampling strategies.



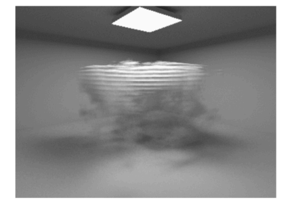
We start by looking into error introduced by different stratification strategies
[CLICK] We then analyze error due to blue noise distributions and how this idea is extended beyond to introduce discrepancy related measures
[CLICK] In the end, we will briefly look at how to analyze importance sampling using the presented Fourier tools and how correlated samples could affect the convergence properties of these importance sampling strategies.

Regular grid samples

Pauly et al. [2000]



Regular grid

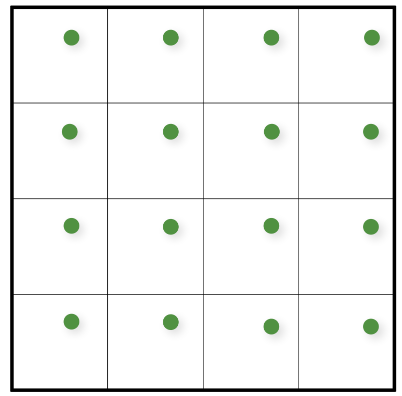


Regular

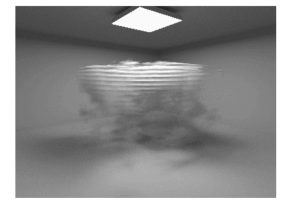
The regular artifacts due to regular grid sampling can be avoided by simply shifting...

Uniformly jittered regular grid

Pauly et al. [2000]



Uniform jitter

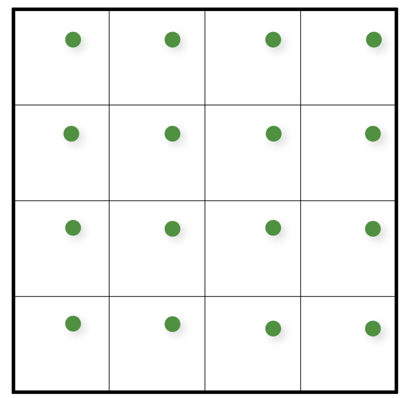


Regular

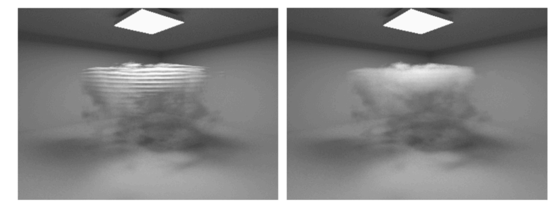
... the samples randomly within the strata, which is called uniform jitter sampling.

Uniformly jittered regular grid

Pauly et al. [2000]



Uniform jitter

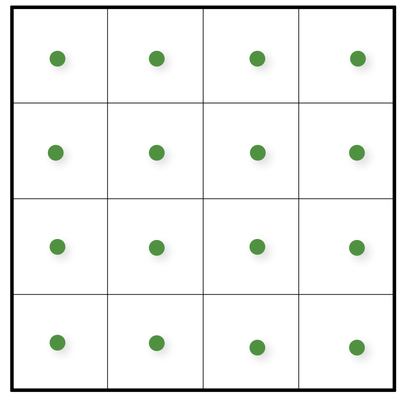


Regular Uniform jitter

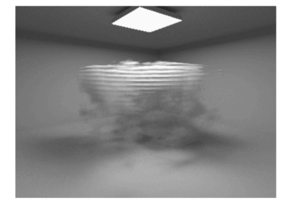
... the samples randomly within the strata, which is called uniform jitter sampling.

Randomly jittered samples

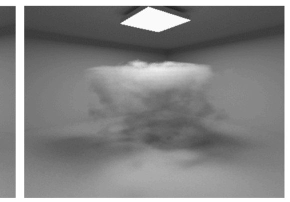
Pauly et al. [2000]



Random jitter

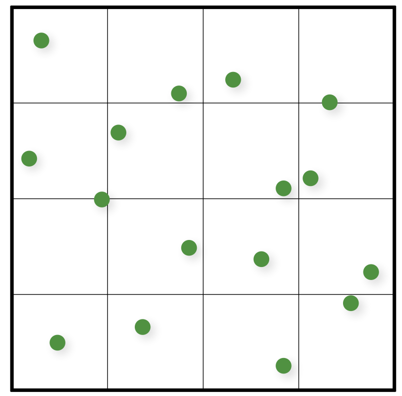


Regular



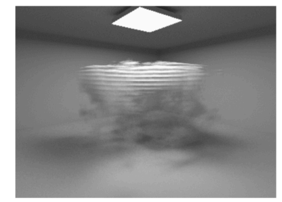
Uniform jitter

Randomly jittered samples

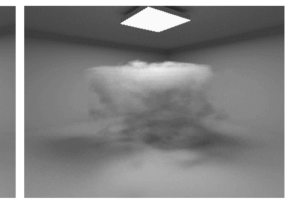


Random jitter

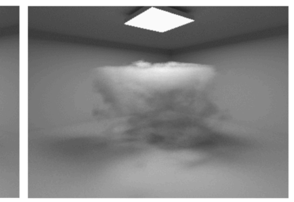
Pauly et al. [2000]



Regular



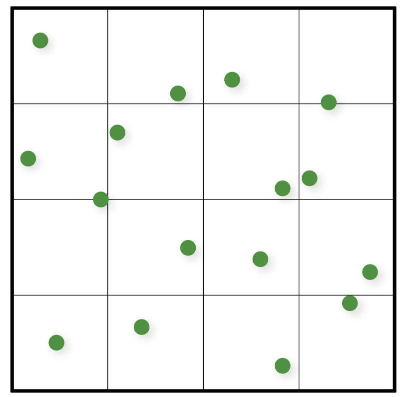
Uniform jitter



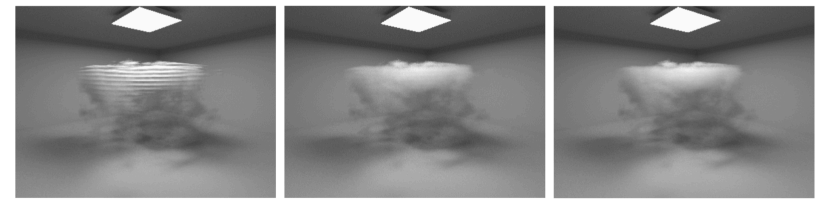
Random jitter

Another approach to avoid these artifacts is to randomly generate samples within each stratum that is called random jittering

Randomly jittered samples

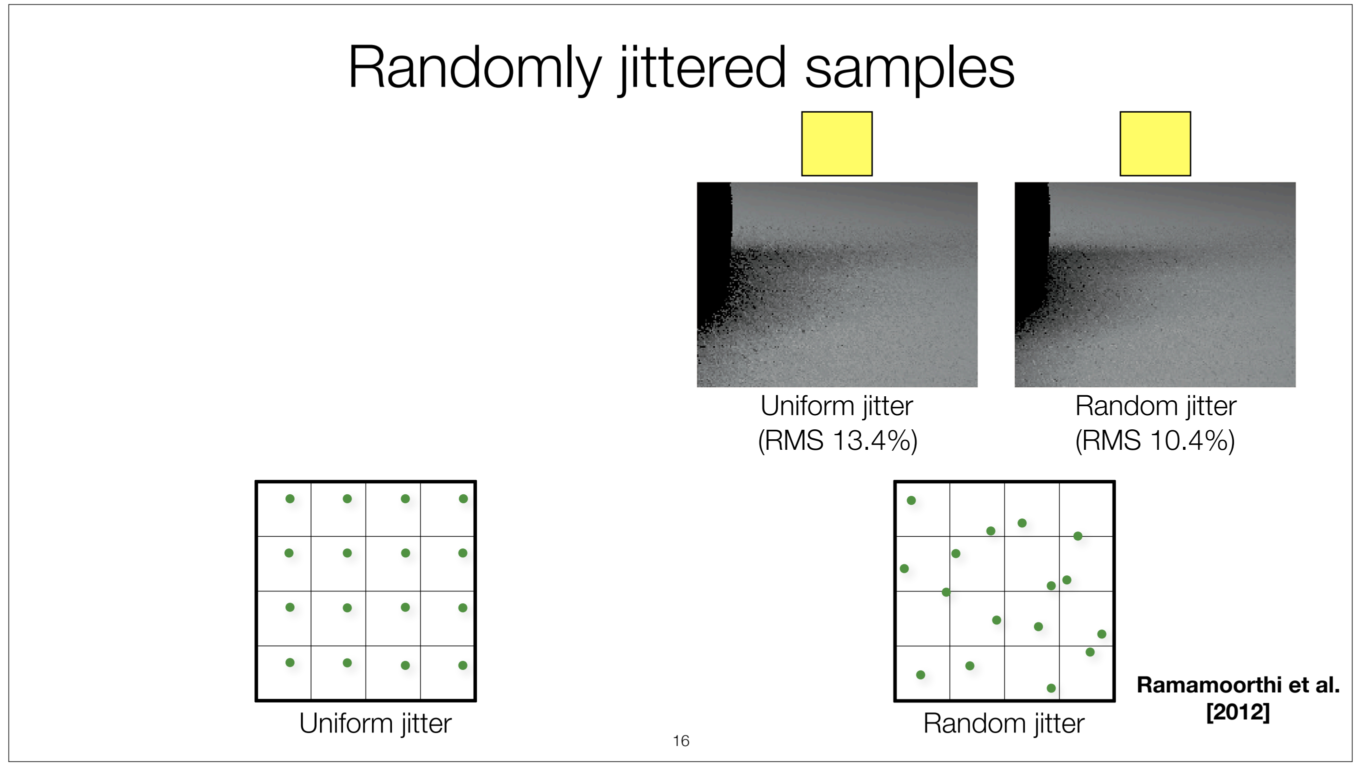


Pauly et al. [2000]

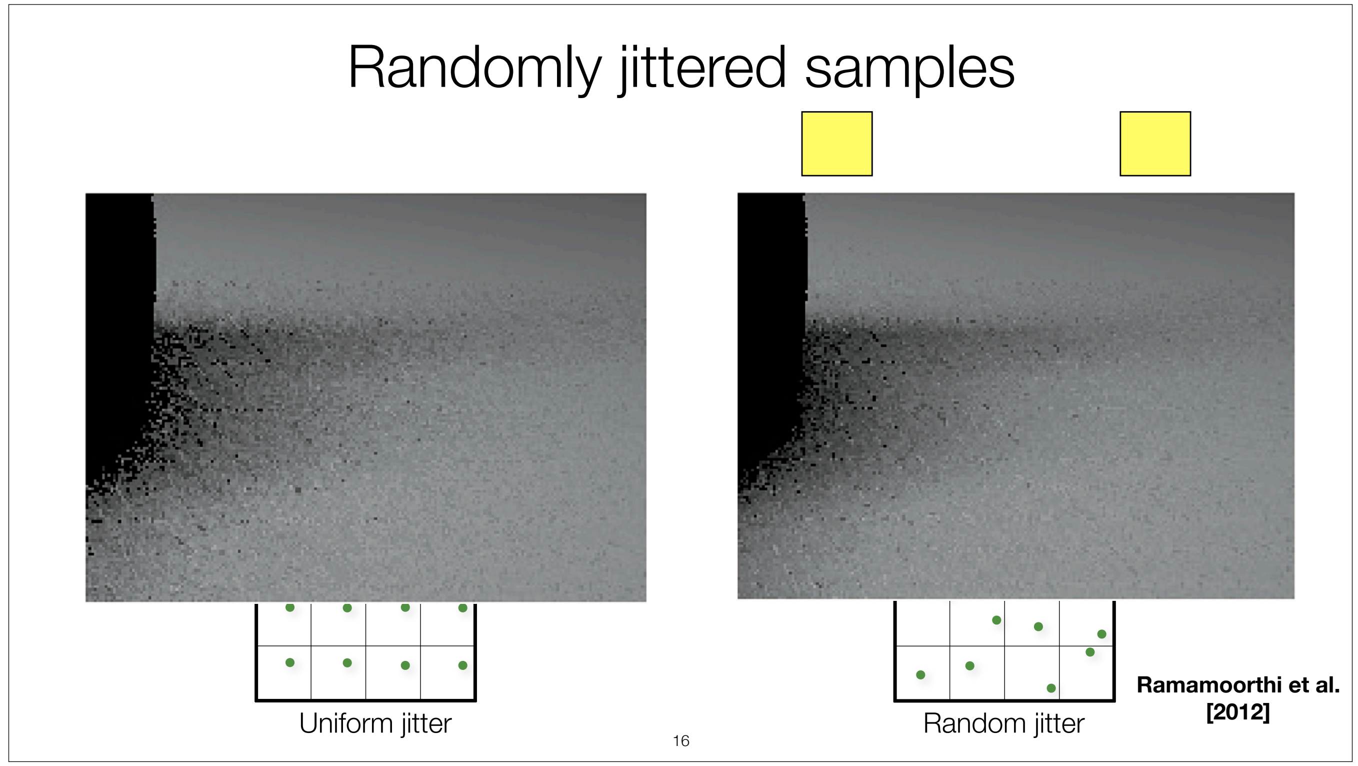


Regular   Uniform jitter   Random jitter

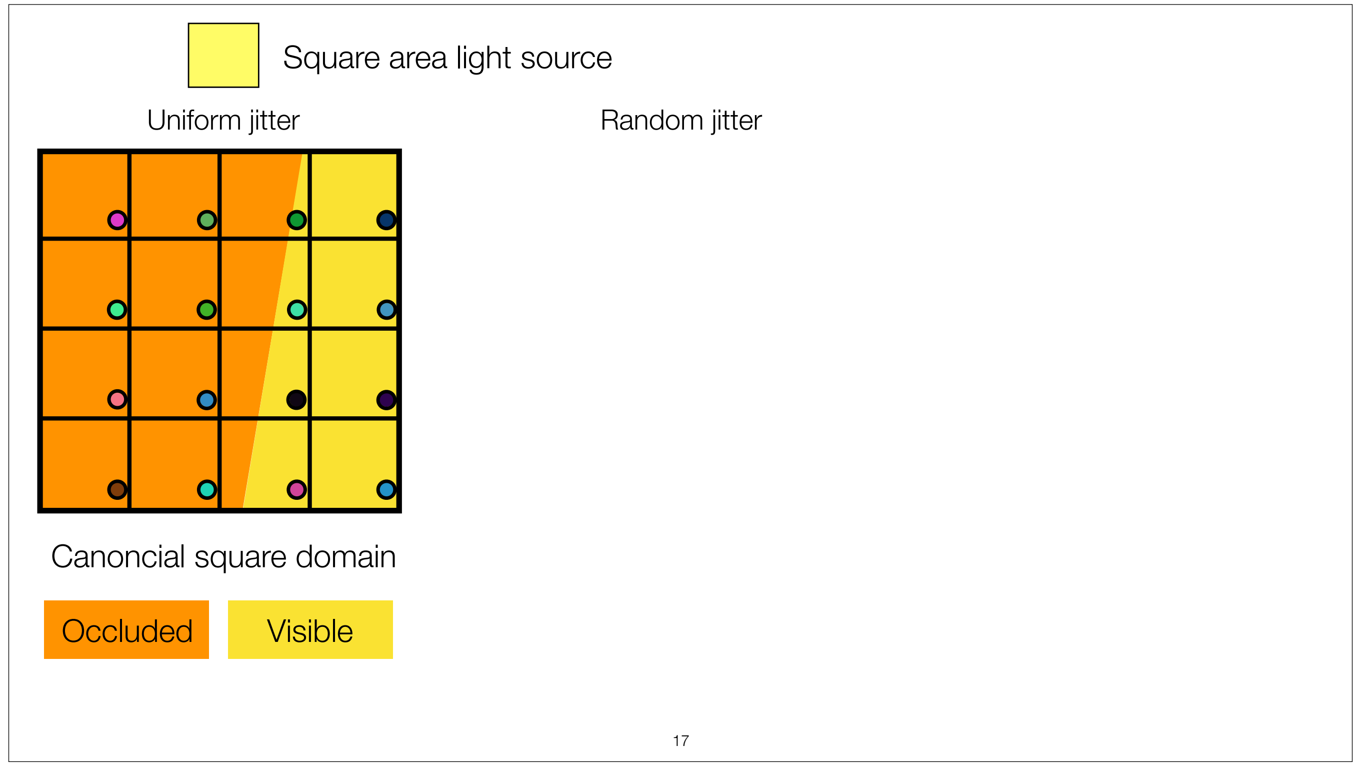
We would like to understand how these two different strategies affect the error during rendering.



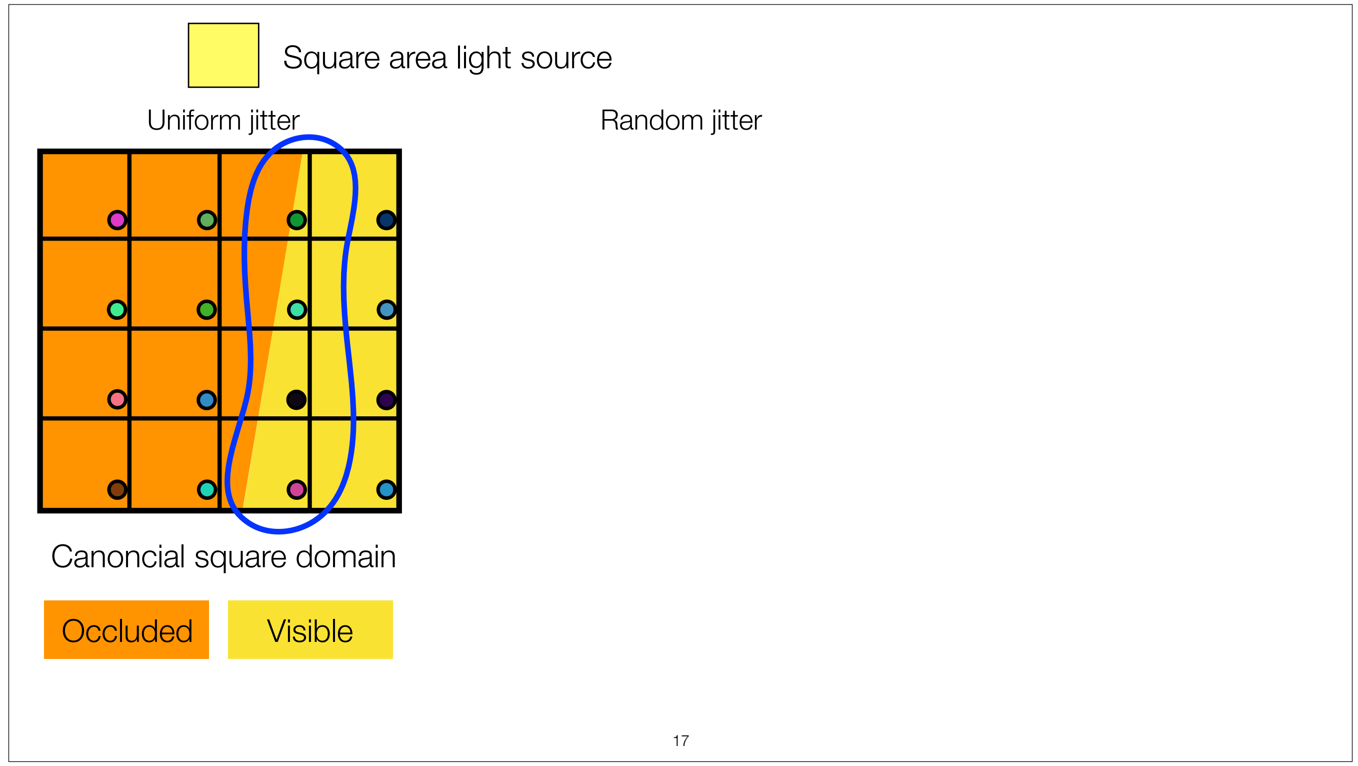
Here we show one example, where for Quad (square) area light source, random jittered sampling [CLICK] gives less noisy image compared to uniform jittering. Let's try to understand why is that.



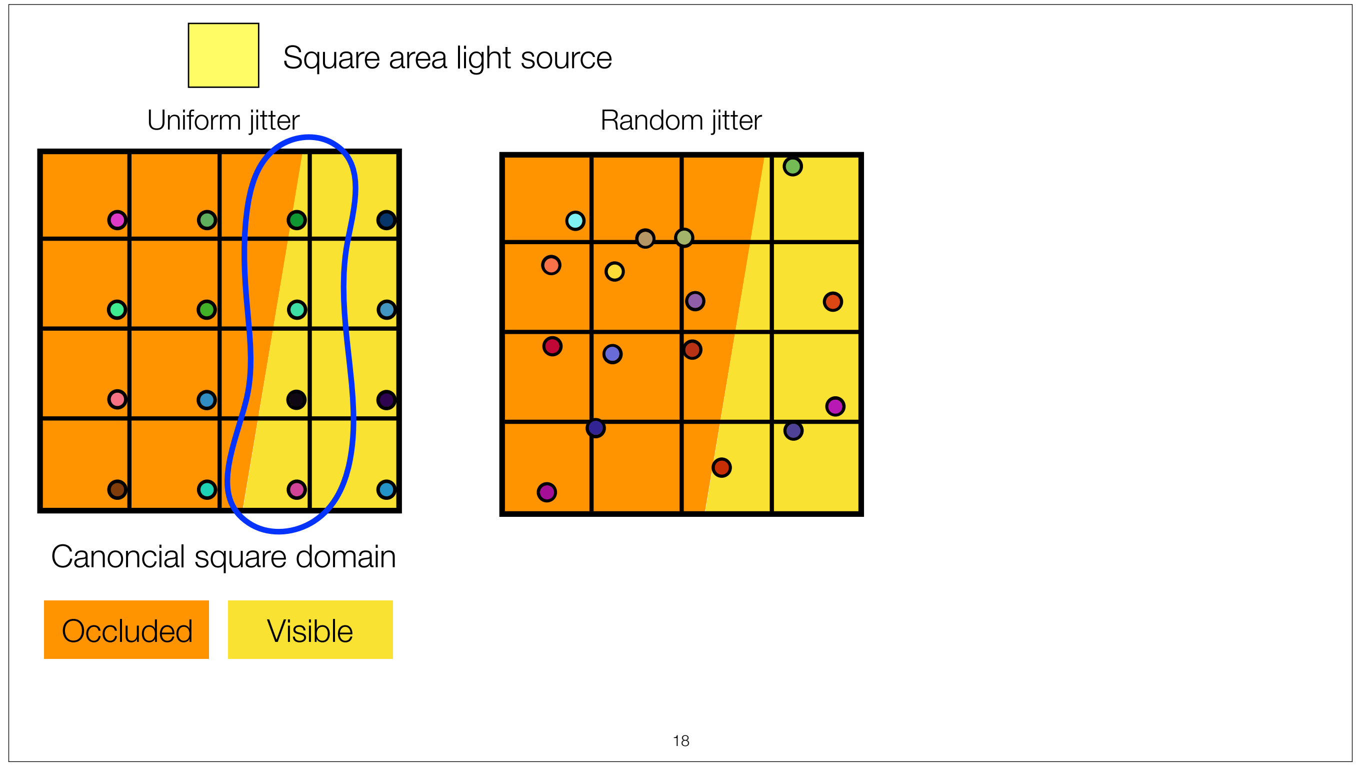
Here we show one example, where for Quad (square) area light source, random jittered sampling [CLICK] gives less noisy image compared to uniform jittering. Let's try to understand why is that.



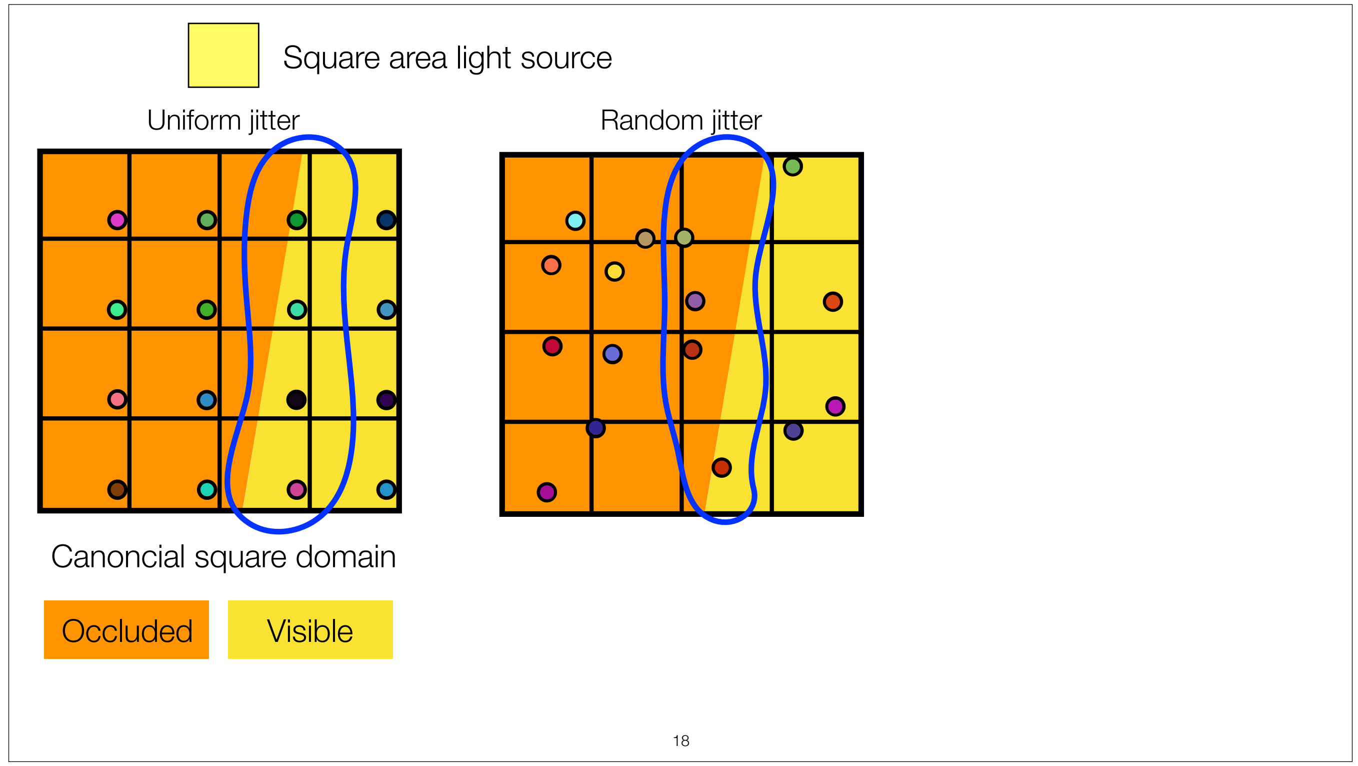
Here we are looking at a light source sampled using uniformly jittered samples. The orange part is occlude whereas the yellow part is visible from the light source. [CLICK] If we look at these samples near the discontinuity, uniform jittering shifts all samples on one side of the discontinuity creating some kind of positive correlation with respect to the discontinuity. However, with randomly jittered samples...



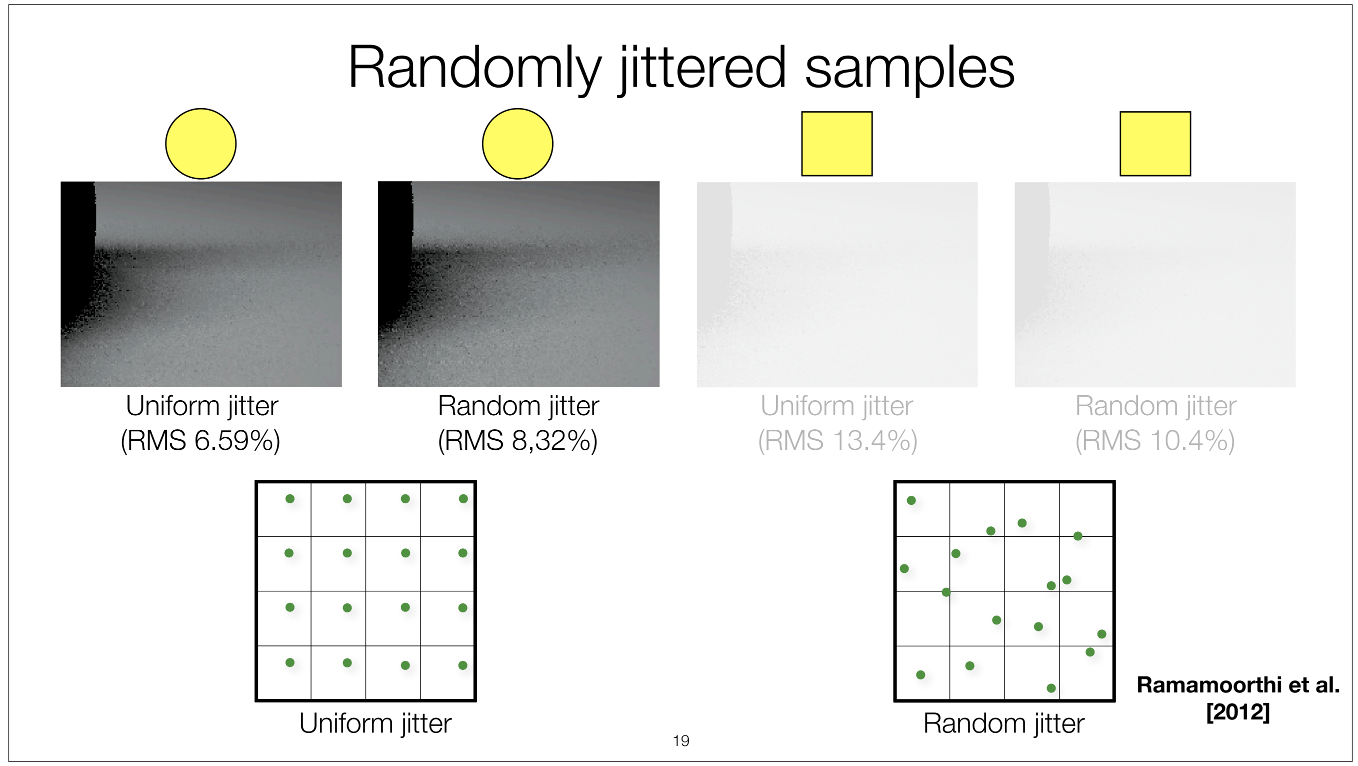
Here we are looking at a light source sampled using uniformly jittered samples. The orange part is occlude whereas the yellow part is visible from the light source. [CLICK] If we look at these samples near the discontinuity, uniform jittering shifts all samples on one side of the discontinuity creating some kind of positive correlation with respect to the discontinuity. However, with randomly jittered samples...



...if we look at the set of pixels with the discontinuity [CLICK], since the samples were generated randomly jittered, samples can be easily found on either side of the discontinuity. This decorrelates randomly jittered samples w.r.t. the discontinuity resulting in less noise.

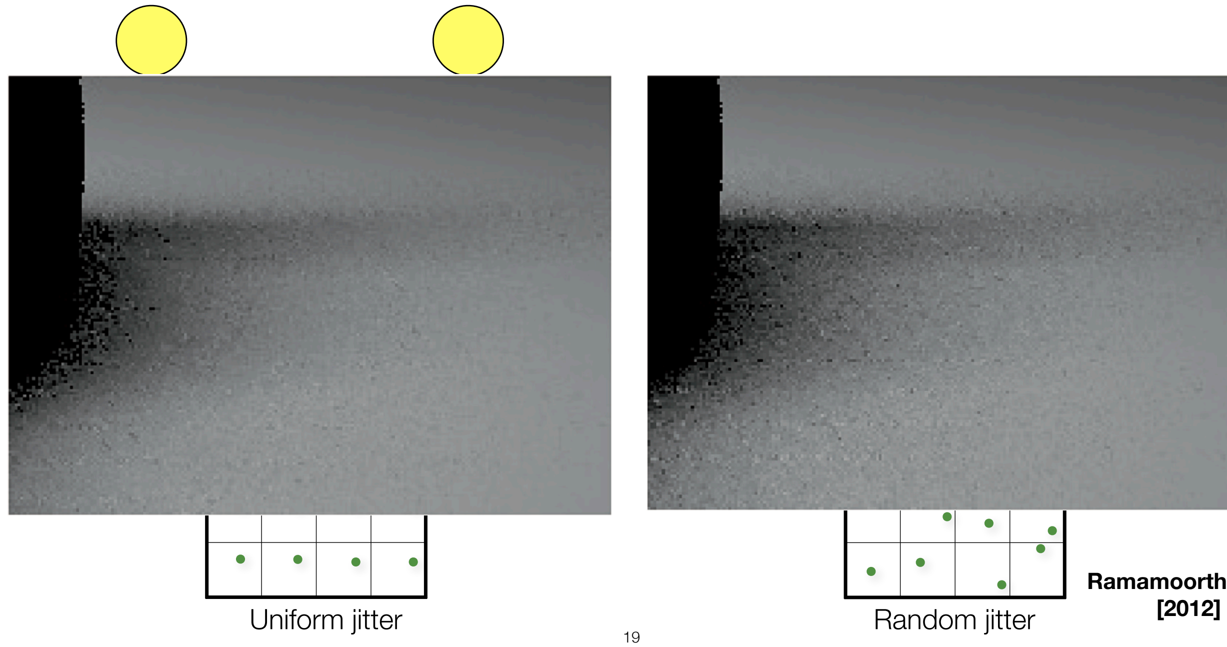


...if we look at the set of pixels with the discontinuity [CLICK], since the samples were generated randomly jittered, samples can be easily found on either side of the discontinuity. This decorrelates randomly jittered samples w.r.t. the discontinuity resulting in less noise.



Now, if we change the shape of our light source to a Disk light, we see that [CLICK] uniform jitter is far better than random jittering. Furthermore,...

Randomly jittered samples



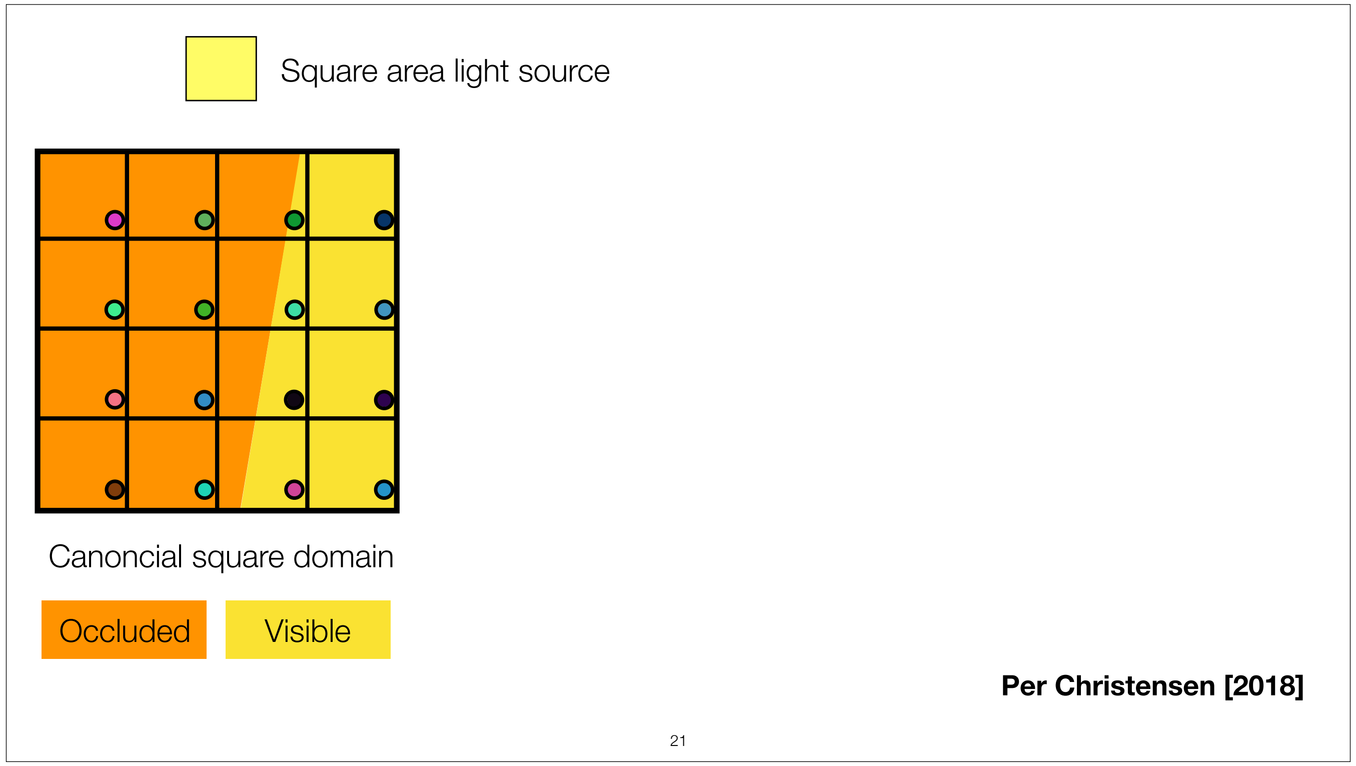
Ramamoorthi et al. [2012]

Now, if we change the shape of our light source to a Disk light, we see that [CLICK] uniform jitter is far better than random jittering. Furthermore,...

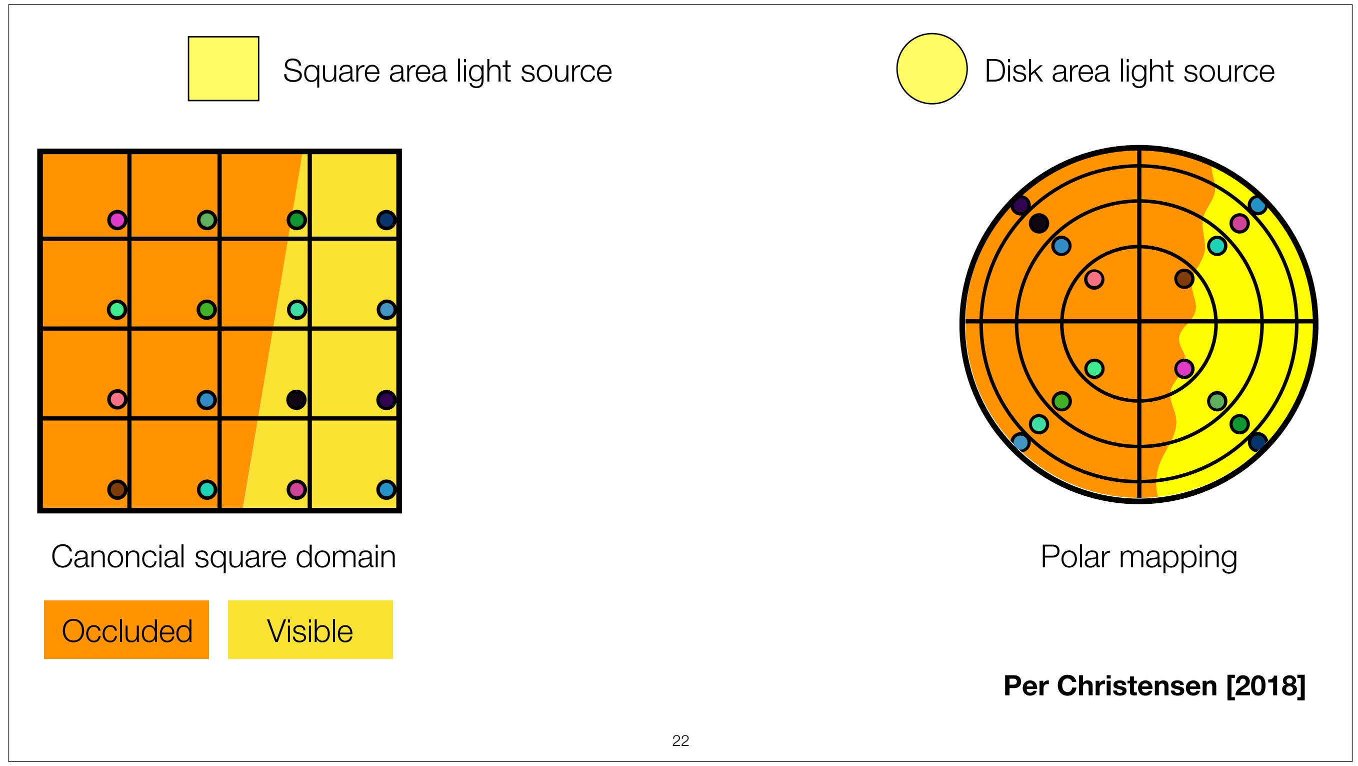
Polar mapping performs **better** for some samplers compared to **concentric mapping**

observed by **Andrew Kensler [2013]**

It was observed that for circular light source, the way these samples are mapped also affect the quality. Surprisingly, polar mapping performs better than concentric mapping in some cases. Let's first look at these mappings.

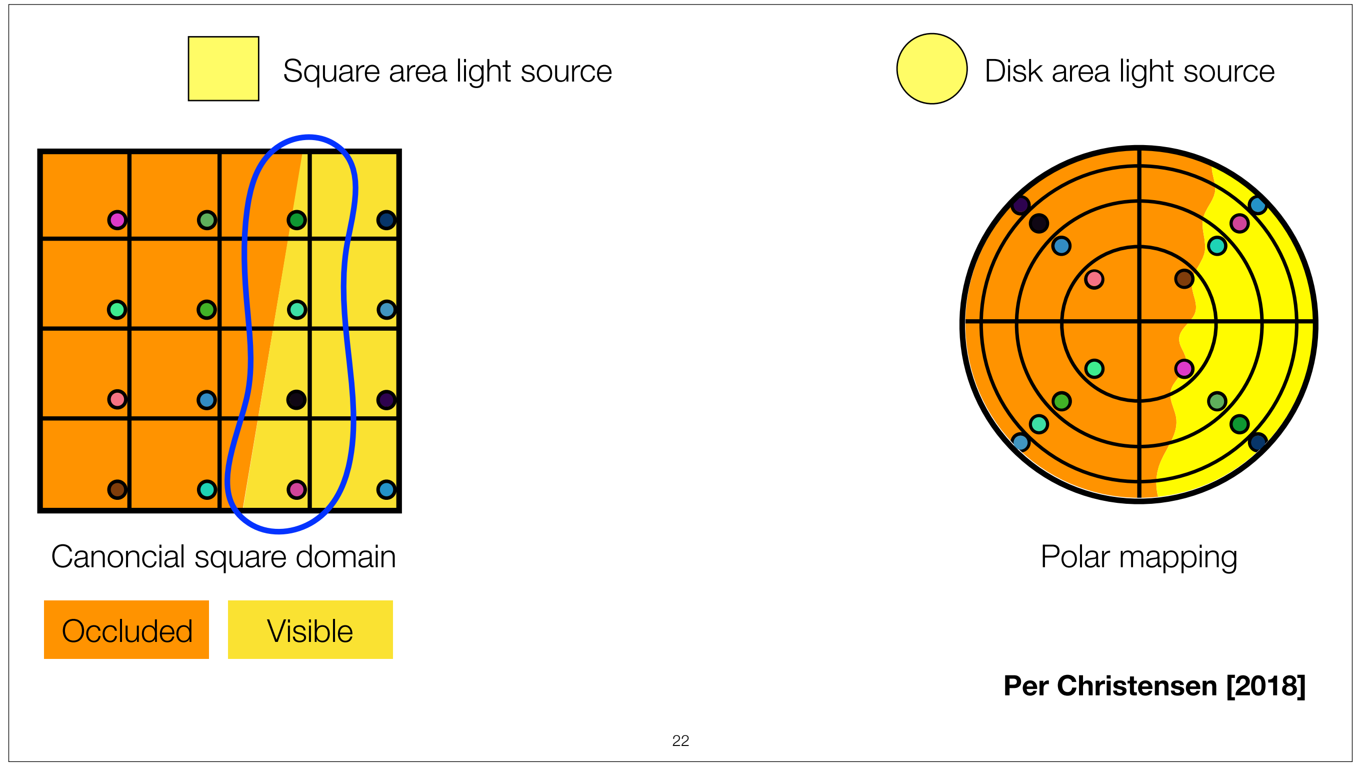


Samples in a circular domain or Disk are generated by warping the samples from the square domain using, say, polar mapping, which ...

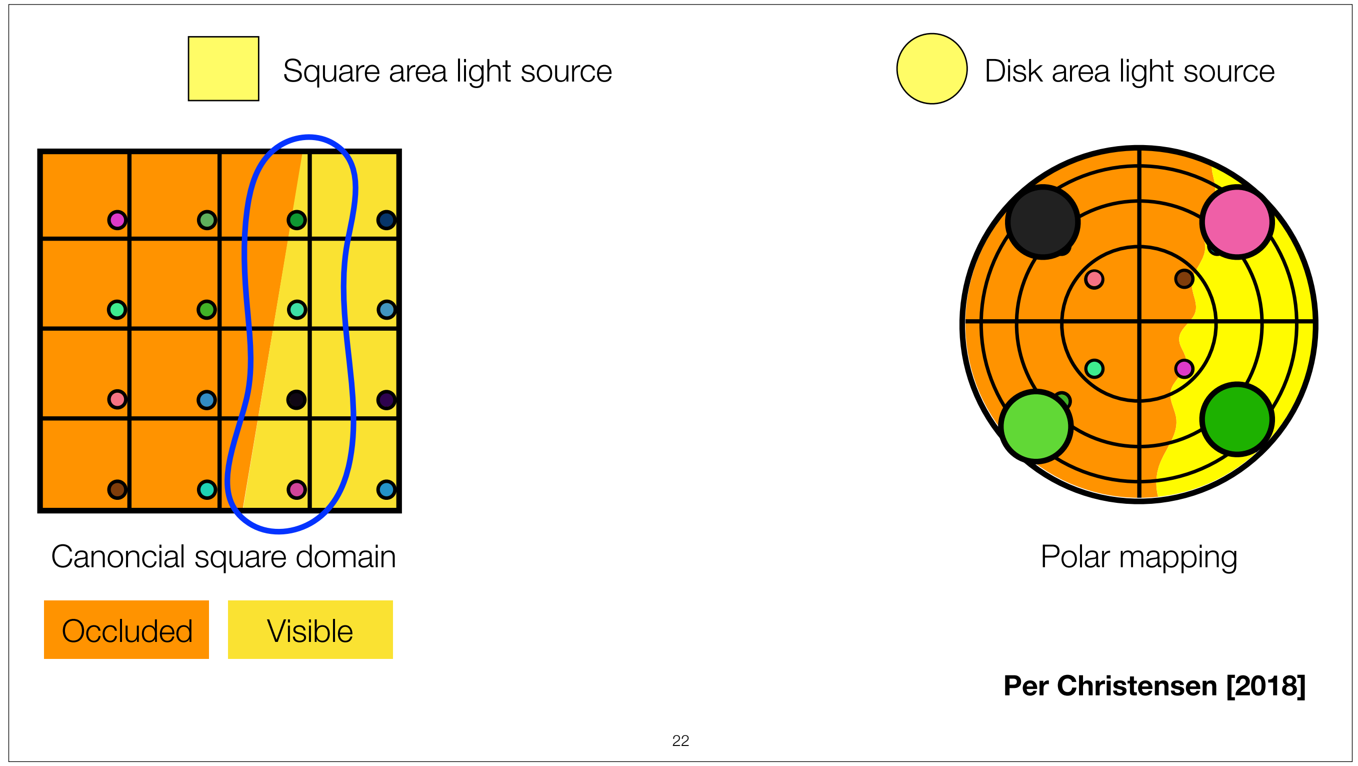


... which distributed samples in this manner.

[CLICK] If we look at the samples near the discontinuity, on the Disk [CLICK] they are placed on either side of the domain in a concentric ring. Here the strata are of uneven shape due to the distortion by the polar mapping which is not considered good, but it was reported that this might explain why polar mapping behaves better for correlated multi-jittered samples (by Kensler) then concentric mapping. In concentric mapping, ...

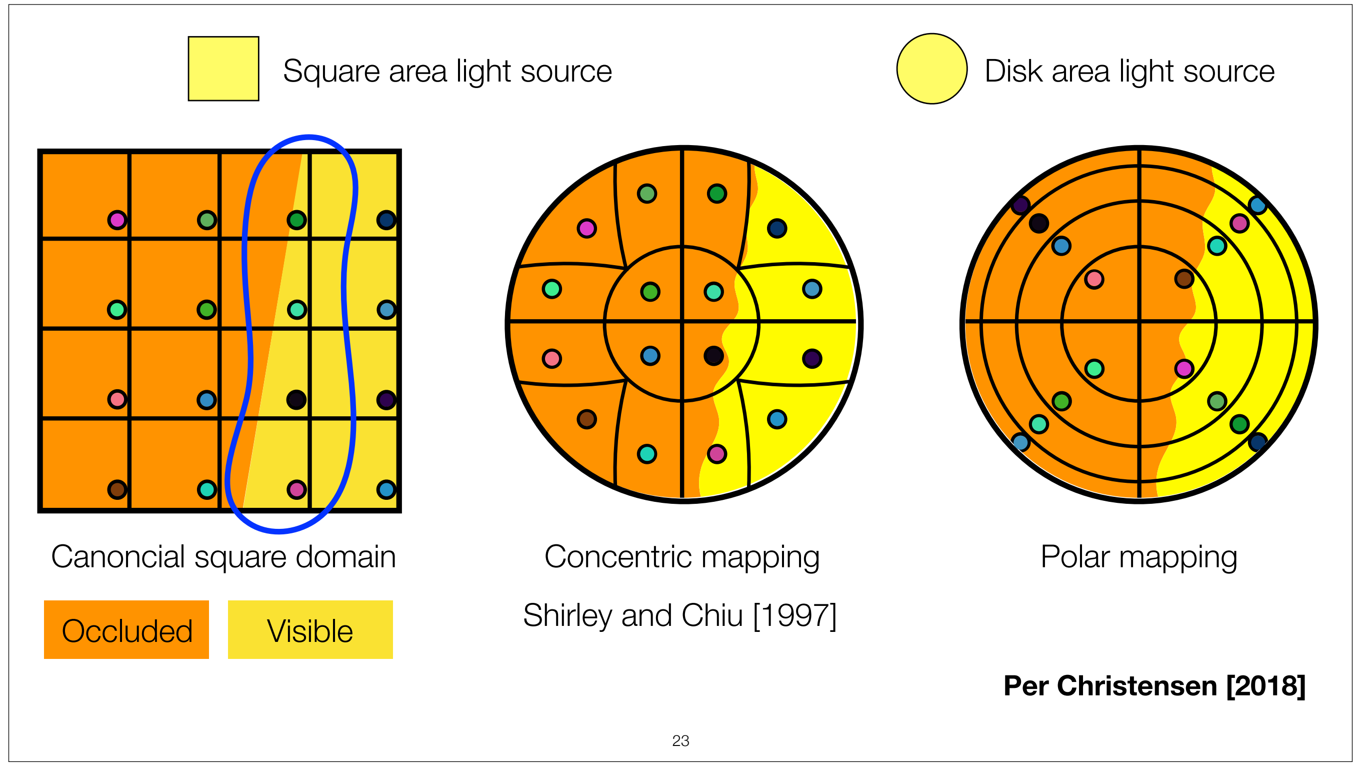


... which distributed samples in this manner.
[CLICK] If we look at the samples near the discontinuity, on the Disk [CLICK] they are placed on either side of the domain in a concentric ring. Here the strata are of uneven shape due to the distortion by the polar mapping which is not considered good, but it was reported that this might explain why polar mapping behaves better for correlated multi-jittered samples (by Kensler) then concentric mapping. In concentric mapping, ...

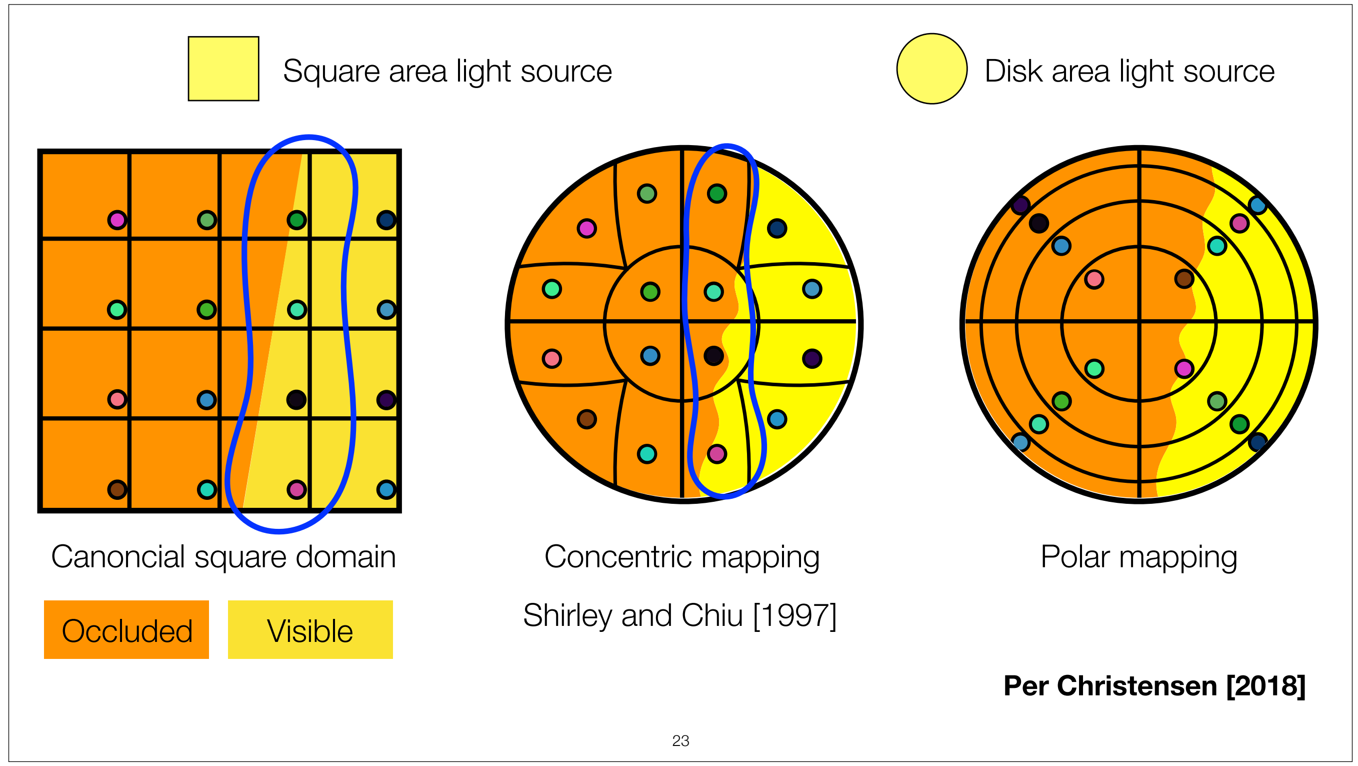


... which distributed samples in this manner.

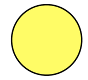
[CLICK] If we look at the samples near the discontinuity, on the Disk [CLICK] they are placed on either side of the domain in a concentric ring. Here the strata are of uneven shape due to the distortion by the polar mapping which is not considered good, but it was reported that this might explain why polar mapping behaves better for correlated multi-jittered samples (by Kensler) then concentric mapping. In concentric mapping, ...



... the strata have a very similar shape compared to the square strata and
if we look again at these samples near the discontinuity, they remain [CLICK] in almost the same vicinity.

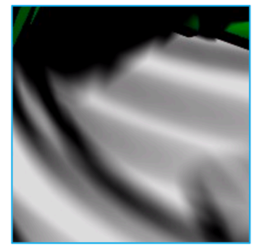


... the strata have a very similar shape compared to the square strata and
if we look again at these samples near the discontinuity, they remain [CLICK] in almost the same vicinity.



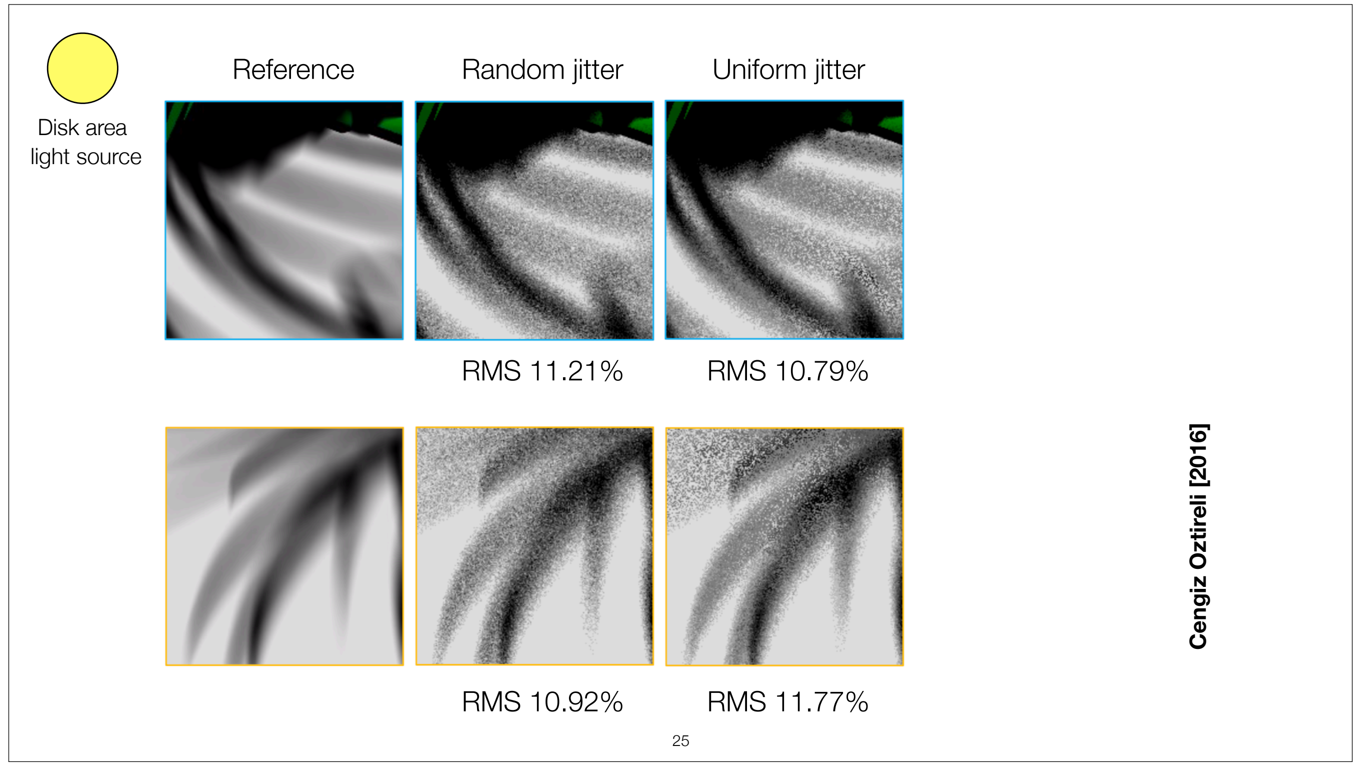
Disk area
light source

Reference

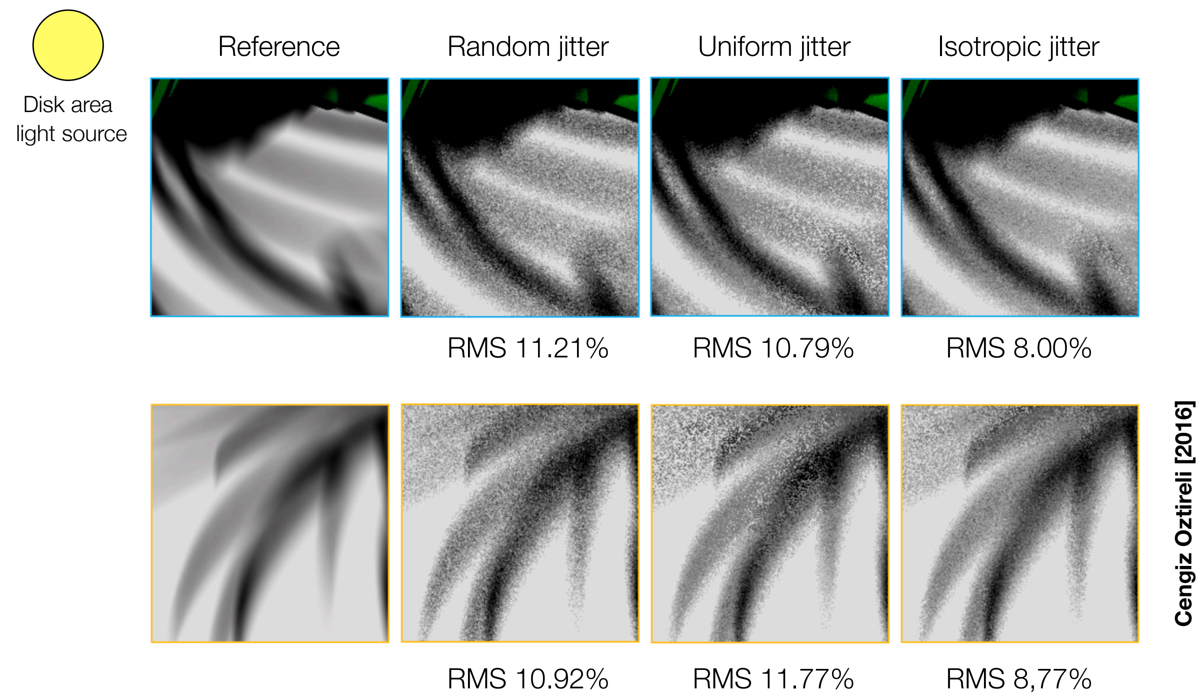


Cengiz Oztireli [2016]

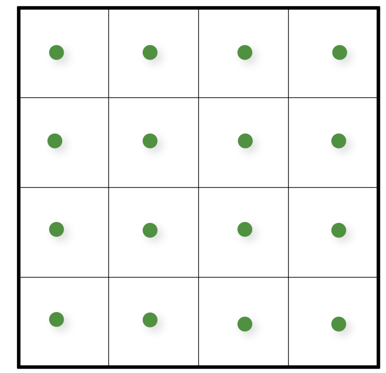
Cengiz analyzed this further and observed that with Disk light sources...



...sometimes uniform jitter is good and sometimes bad compared to random jittering. He proposed another variant named isotropic jittering.

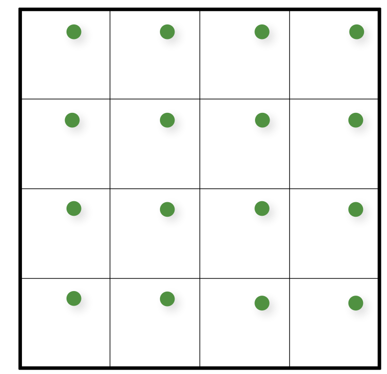

Disk area
light source
Reference
Random jitter
Uniform jitter
Isotropic jitter
RMS 11.21%
RMS 10.79%
RMS 8.00%
RMS 10.92%
RMS 11.77%
RMS 8,77%
Cengiz Oztireli [2016]

Isotropic jitter = uniform jitter + random rotation



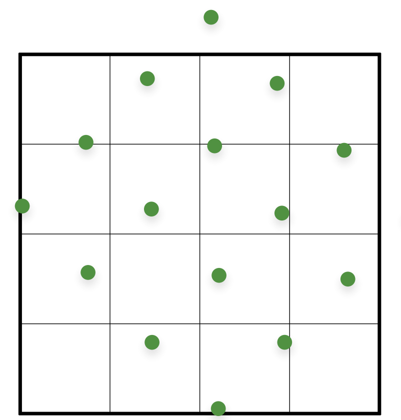
The idea is,
[CLICK] you first randomly shift the samples and then,
[CLICK] randomly rotate them.

Isotropic jitter = uniform jitter + random rotation



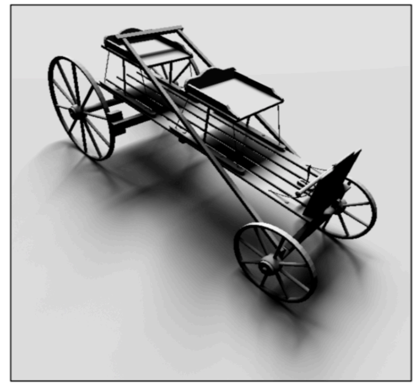
The idea is,
[CLICK] you first randomly shift the samples and then,
[CLICK] randomly rotate them.

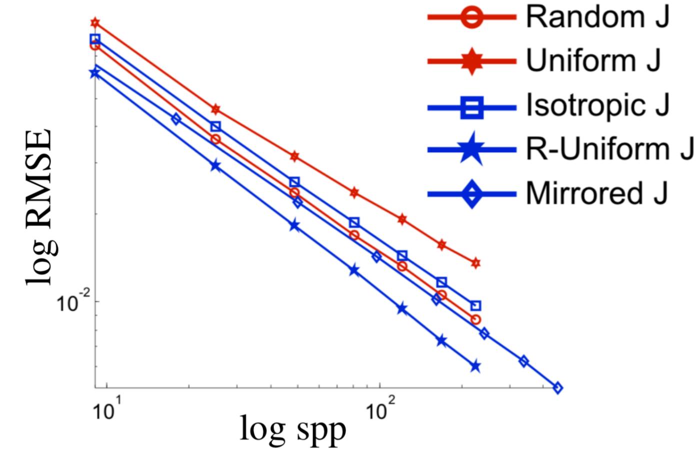
Isotropic jitter = uniform jitter + random rotation



The idea is,
[CLICK] you first randomly shift the samples and then,
[CLICK] randomly rotate them.

Rotated uniform jitter better for not too complex shadows

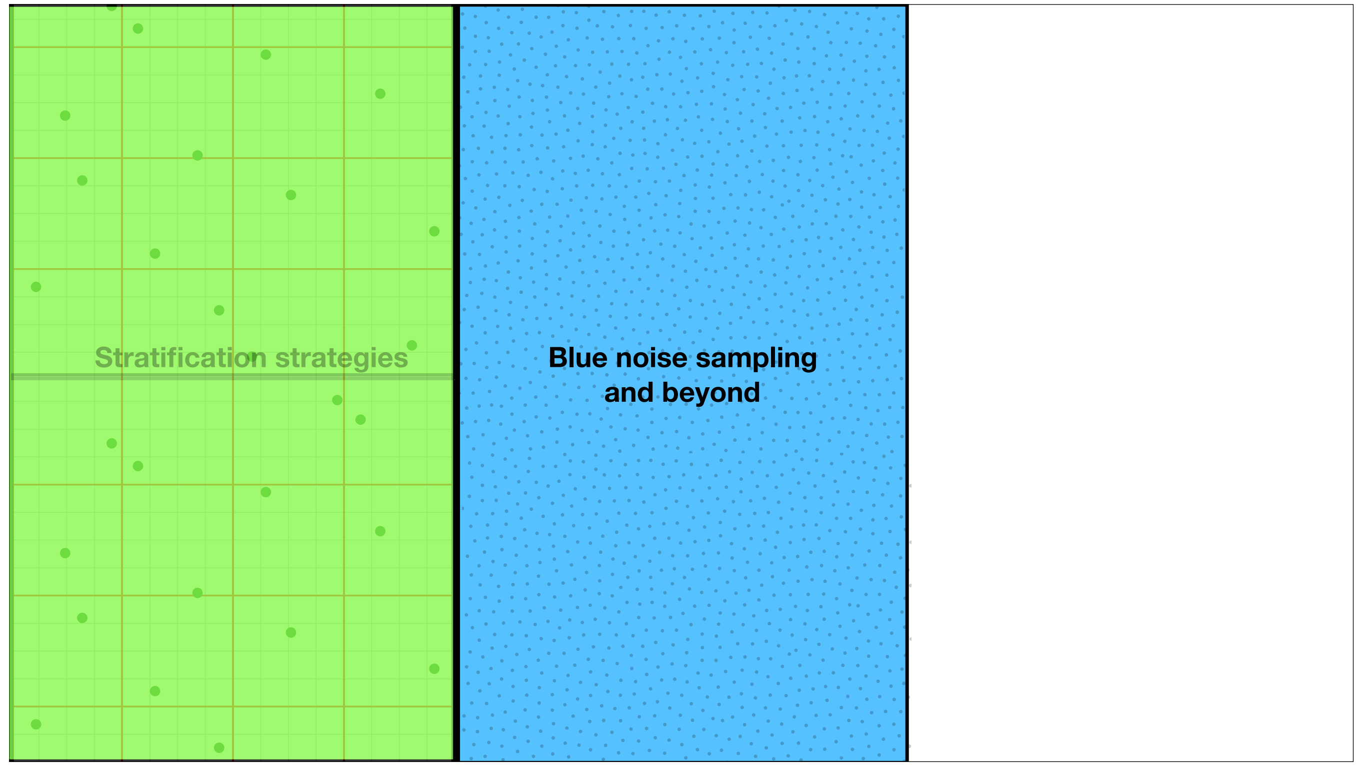




Cengiz Oztireli [2016]

This showed improvements for not too complex shadows, especially when this rotation is done with some knowledge of the orientation of the occlusion boundaries.

This shows that these simple correlations introduced by different jittering variants can favorably affect the error.
[CLICK[Let's now look at how other correlations, namely blue noise samples, affect the error in Monte Carlo estimation.



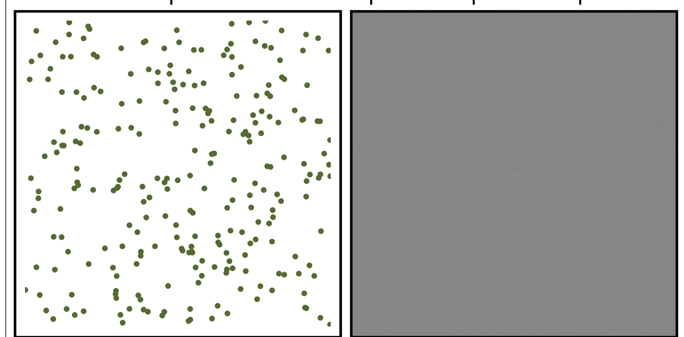
This shows that these simple correlations introduced by different jittering variants can favorably affect the error.
[CLICK[Let's now look at how other correlations, namely blue noise samples, affect the error in Monte Carlo estimation.

Fourier analysis of sample correlations

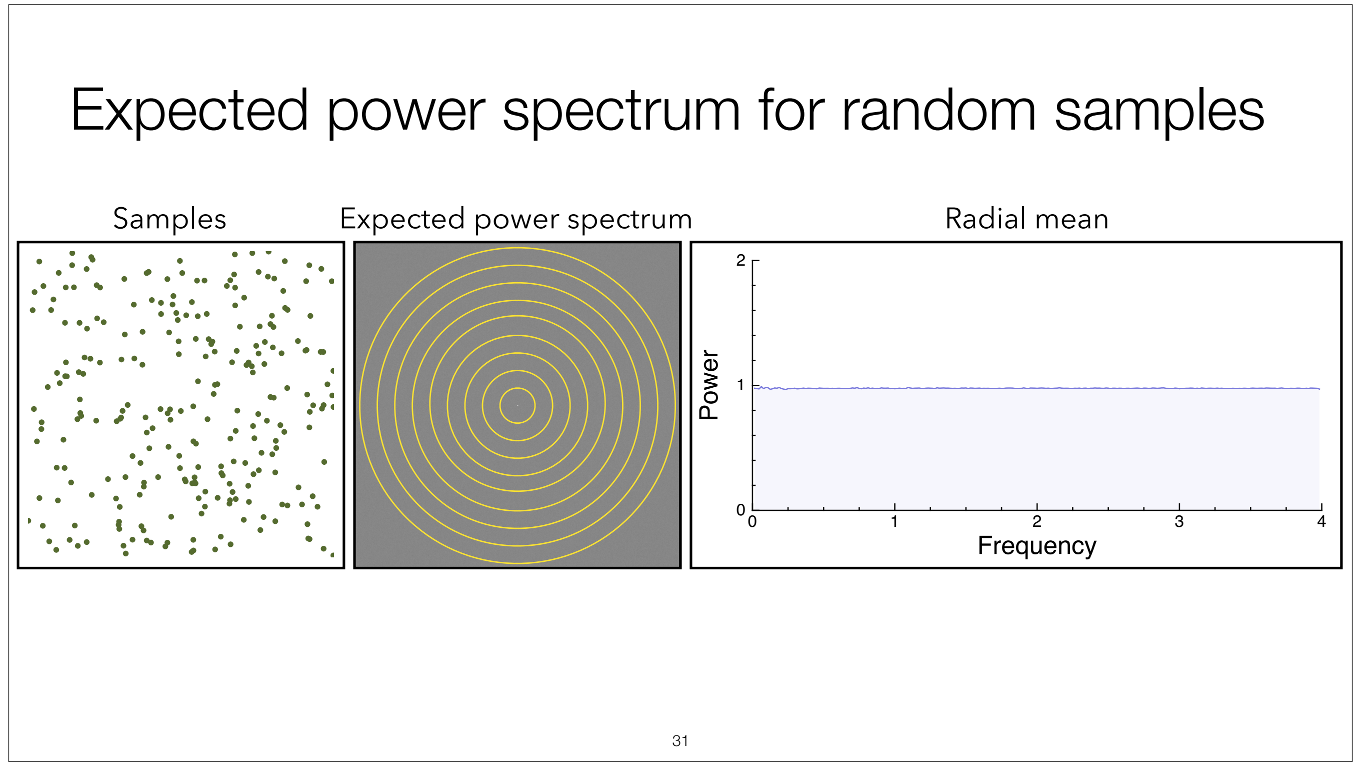
Fourier tools are often used to understand the characteristics of sample correlations

Expected power spectrum for random samples

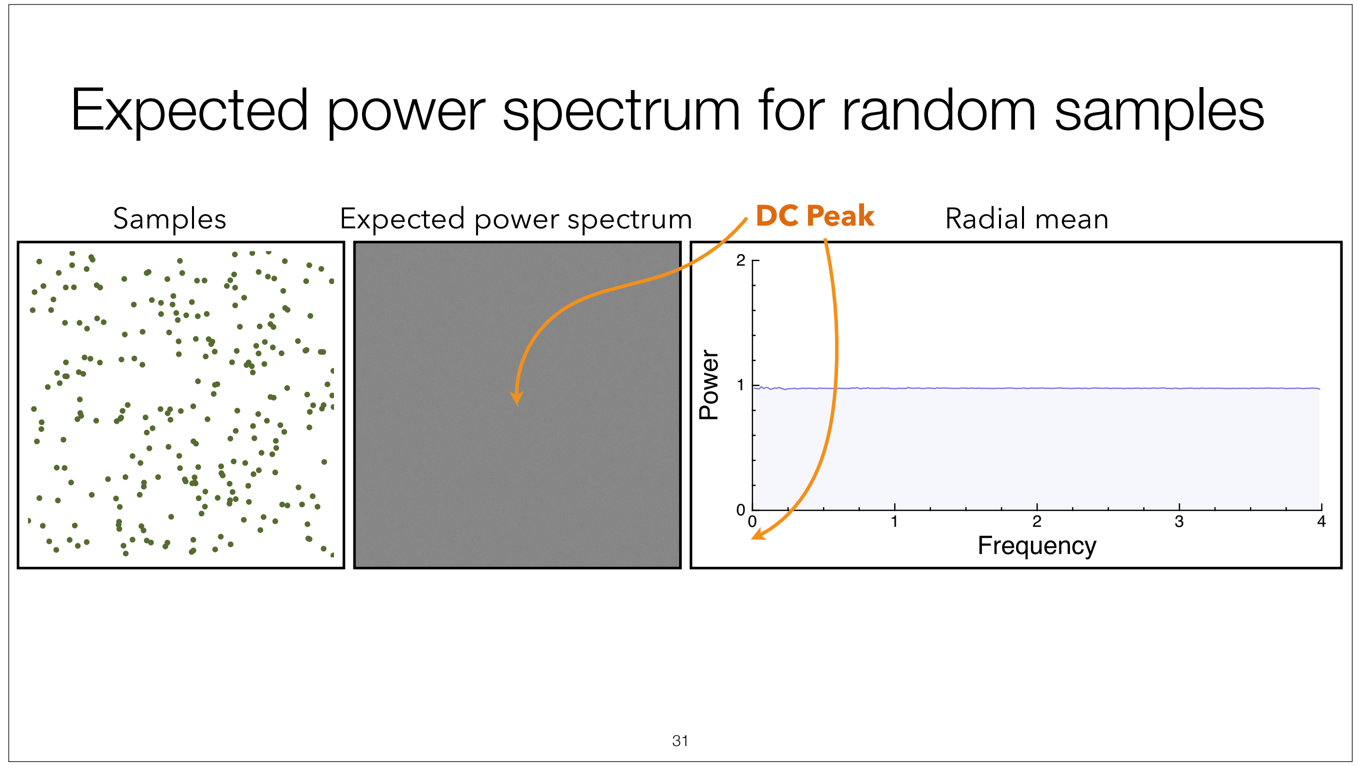
Samples Expected power spectrum Radial mean



For example, we can compute the expected power spectrum of random samples, which is a flat gray image. This spectrum can be [CLICK] radially averaged to get a 1D radial version of this spectrum. Here, [CLICK] the center of the 2D spectrum is the DC peak (or the zero frequency), which represents the starting frequency of the radially averaged spectrum.

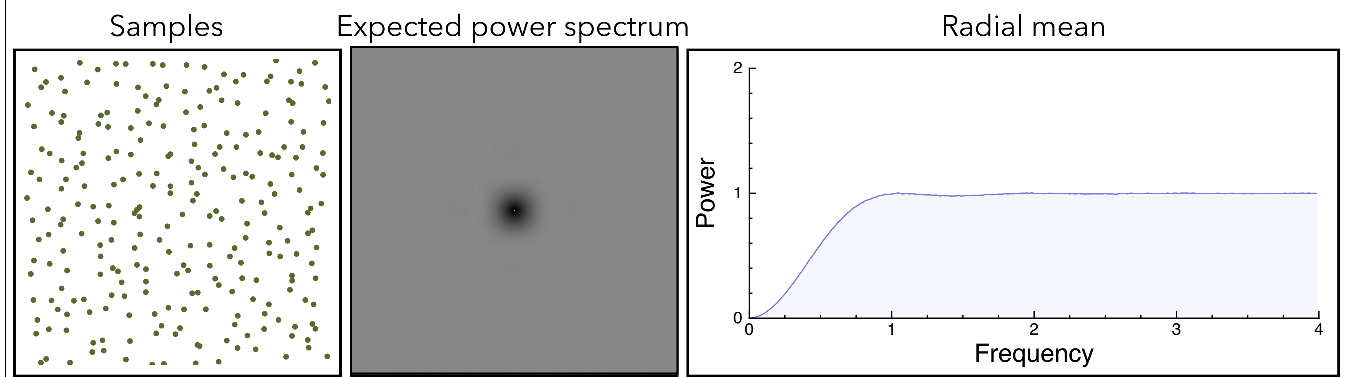


For example, we can compute the expected power spectrum of random samples, which is a flat gray image. This spectrum can be [CLICK] radially averaged to get a 1D radial version of this spectrum. Here, [CLICK] the center of the 2D spectrum is the DC peak (or the zero frequency), which represents the starting frequency of the radially averaged spectrum.

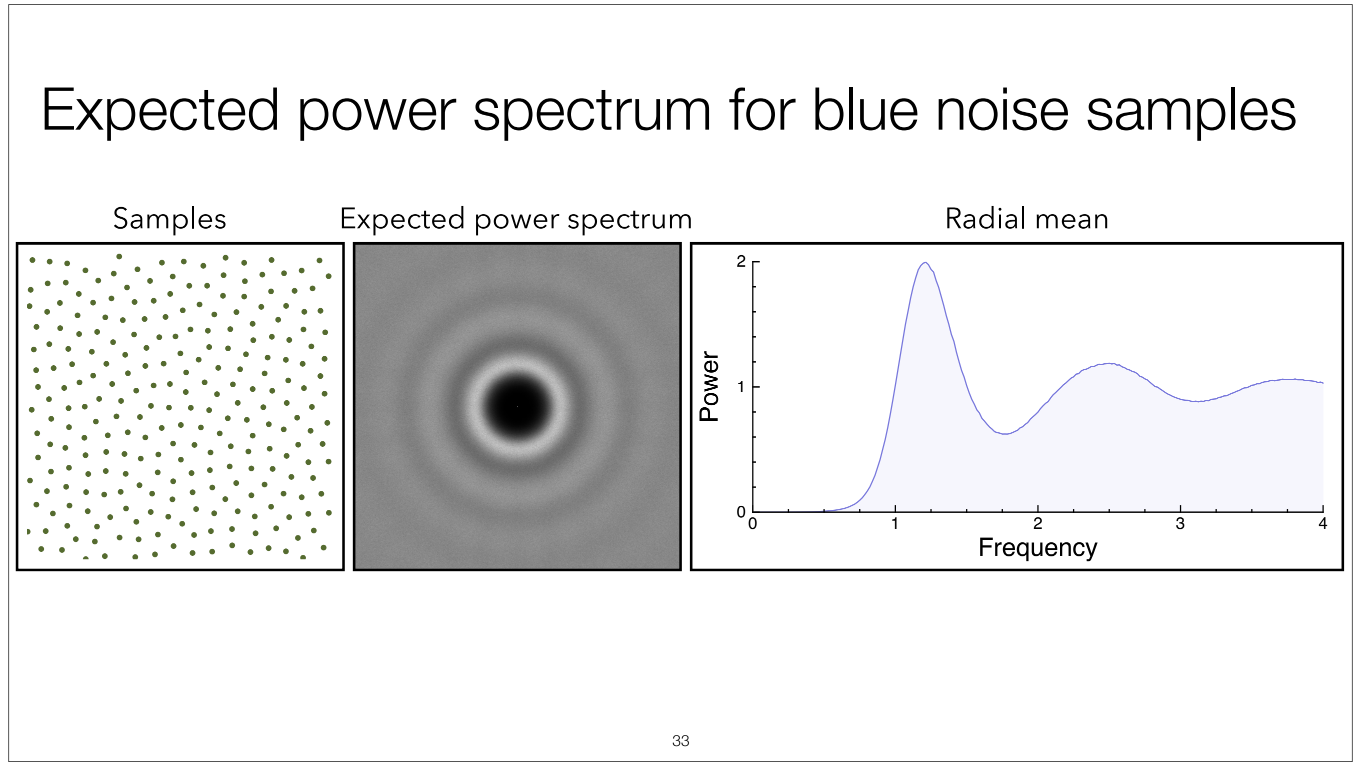


For example, we can compute the expected power spectrum of random samples, which is a flat gray image. This spectrum can be [CLICK] radially averaged to get a 1D radial version of this spectrum. Here, [CLICK] the center of the 2D spectrum is the DC peak (or the zero frequency), which represents the starting frequency of the radially averaged spectrum.

Expected power spectrum for jittered samples

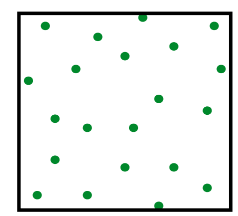


For jittered samples, the spectrum changes and gets some dark region in the low frequency region around the DC peak which is also well captured in the radially averaged profile between the range [0,1]. Note that the horizontal axis in the radial profile represents a normalized frequency (m/sqrt[N] value for an m-th frequency).

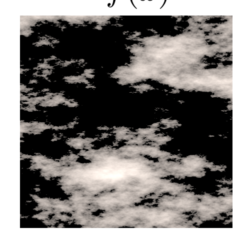


For blue noise samplers, this energy free low frequency region becomes larger.

Variance in terms of power spectra



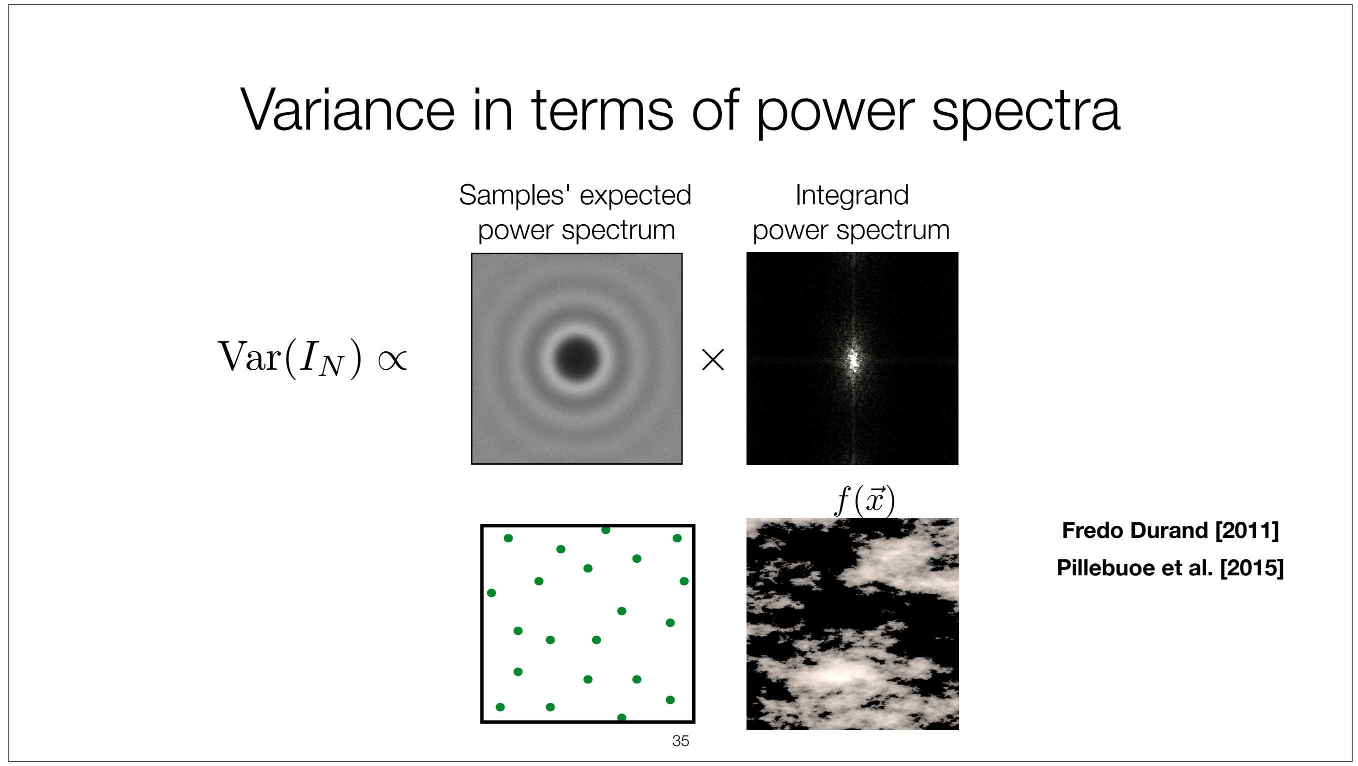
$f(\vec{x})$



Fredo Durand [2011]

Pillebuoe et al. [2015]

When we sample a given function f(x), the variance during Monte Carlo integration due to these samples is...

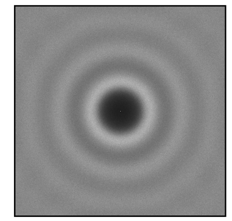


... proportional to the product of the integrand power spectrum and the samples' expected power spectrum.

Variance in terms of power spectra

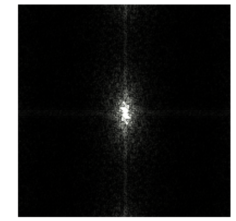
$$\mathrm{Var}(I_N) \propto$$

Samples' expected power spectrum



$\times$

Integrand power spectrum



Fredo Durand [2011]

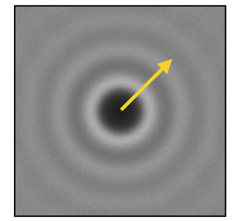
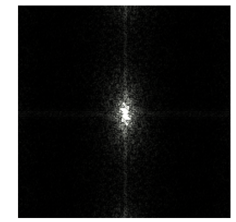
Pillebuoe et al. [2015]

For samples with [CLICK] isotropic power spectra, [CLICK] that is having same energy distribution for a given radial distribution, the corresponding variance expression can be simplified to...

Variance in terms of power spectra

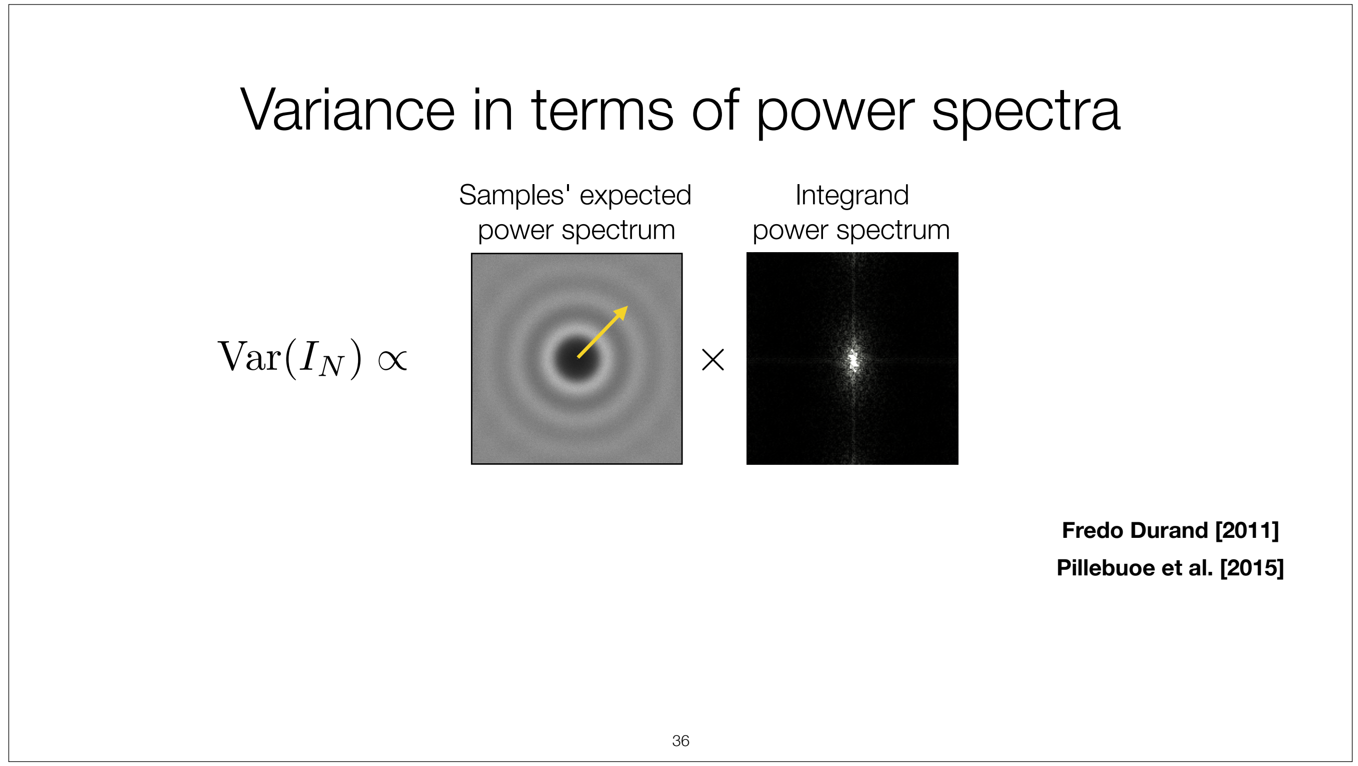
Samples' expected power spectrum

Integrand power spectrum

$$\mathrm{Var}(I_N) \propto$$  $\times$ 

Fredo Durand [2011]
Pillebuoe et al. [2015]

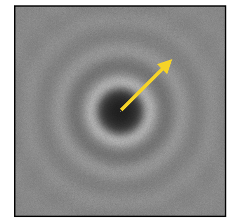
For samples with [CLICK] isotropic power spectra, [CLICK] that is having same energy distribution for a given radial distribution, the corresponding variance expression can be simplified to...



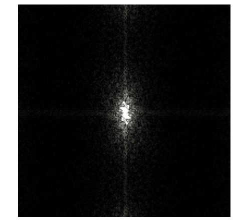
For samples with [CLICK] isotropic power spectra, [CLICK] that is having same energy distribution for a given radial distribution, the corresponding variance expression can be simplified to...

Variance in terms of power spectra

Samples' expected power spectrum



Integrand power spectrum



$$\mathrm{Var}(I_N) \propto [\text{Samples' expected power spectrum}] \times [\text{Integrand power spectrum}]$$

Fredo Durand [2011]

Pillebuoe et al. [2015]

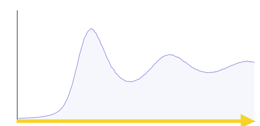
For samples with [CLICK] isotropic power spectra, [CLICK] that is having same energy distribution for a given radial distribution, the corresponding variance expression can be simplified to...

Variance in terms of power spectra

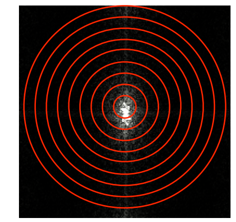
Samples' expected power spectrum

Integrand power spectrum

$$\mathrm{Var}(I_N) \propto$$



$\times$

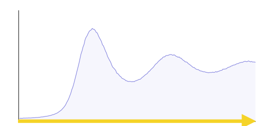
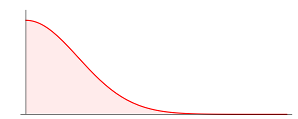


Pillebuoe et al. [2015]

...the radial counterpart. By taking the radial average of the corresponding integrand spectrum...

Variance in terms of power spectra

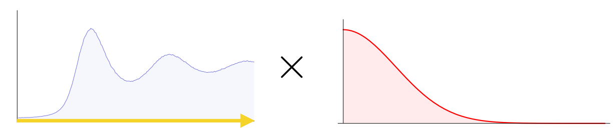
Samples' expected power spectrum | Integrand power spectrum

$$\mathrm{Var}(I_N) \propto$$  $\times$ 

Pilleboue et al. [2015]

...the correponding variance can be represented as function of these 1D radial profiles.

Convergence rate depends on the low frequency region



$$\mathrm{Var}(I_N) \propto \;\; \times$$

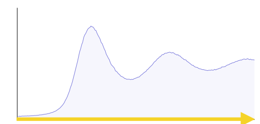
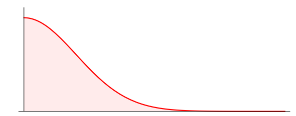
Samplers	Worst Case	Best Case
Random		
Poisson Disk		
Jitter		
CCVT		

Pilleboue et al. [2015]

We used these radial profiles [CLICK] to derive convergence rates for different sampling patterns.
The surprising result noted here is the convergence behavior of Poisson disk which belongs to the blue noise class with jittered samples, which has much better convergence.

Convergence rate depends on the low frequency region

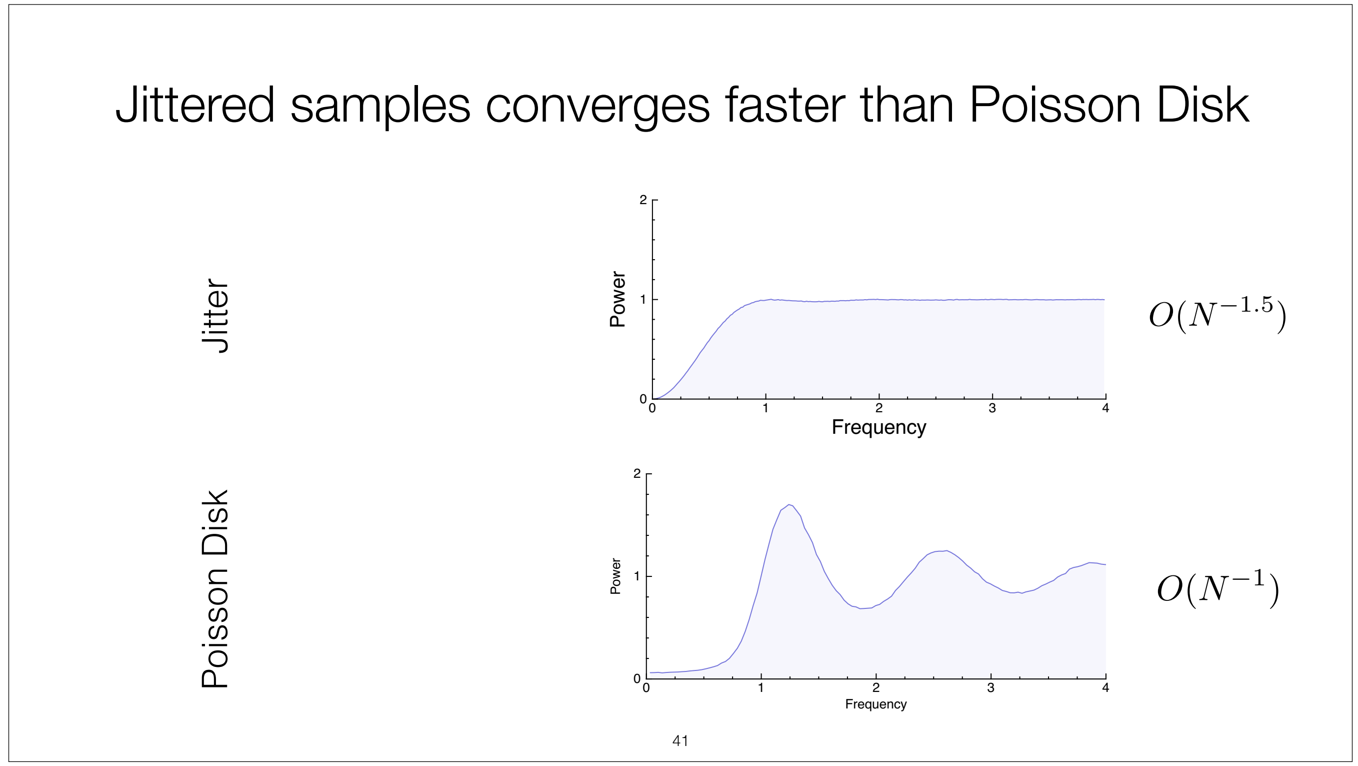
Samples' expected power spectrum | Integrand power spectrum

$\mathrm{Var}(I_N) \propto$  $\times$ 

Samplers	Worst Case	Best Case
Random	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Poisson Disk	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Jitter	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-2})$
CCVT	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-3})$

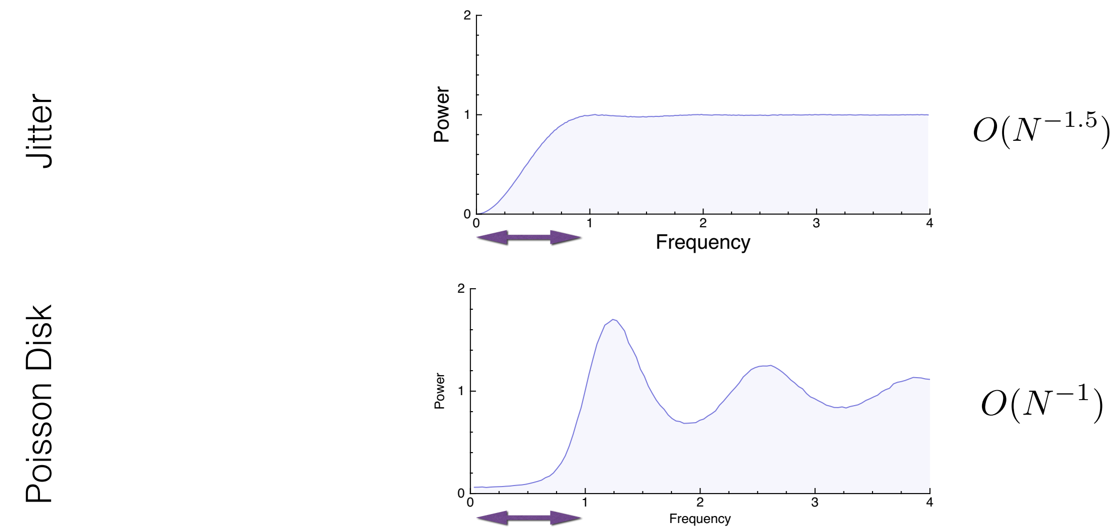
Pilleboue et al. [2015]

We used these radial profiles [CLICK] to derive convergence rates for different sampling patterns.
The surprising result noted here is the convergence behavior of Poisson disk which belongs to the blue noise class with jittered samples, which has much better convergence.

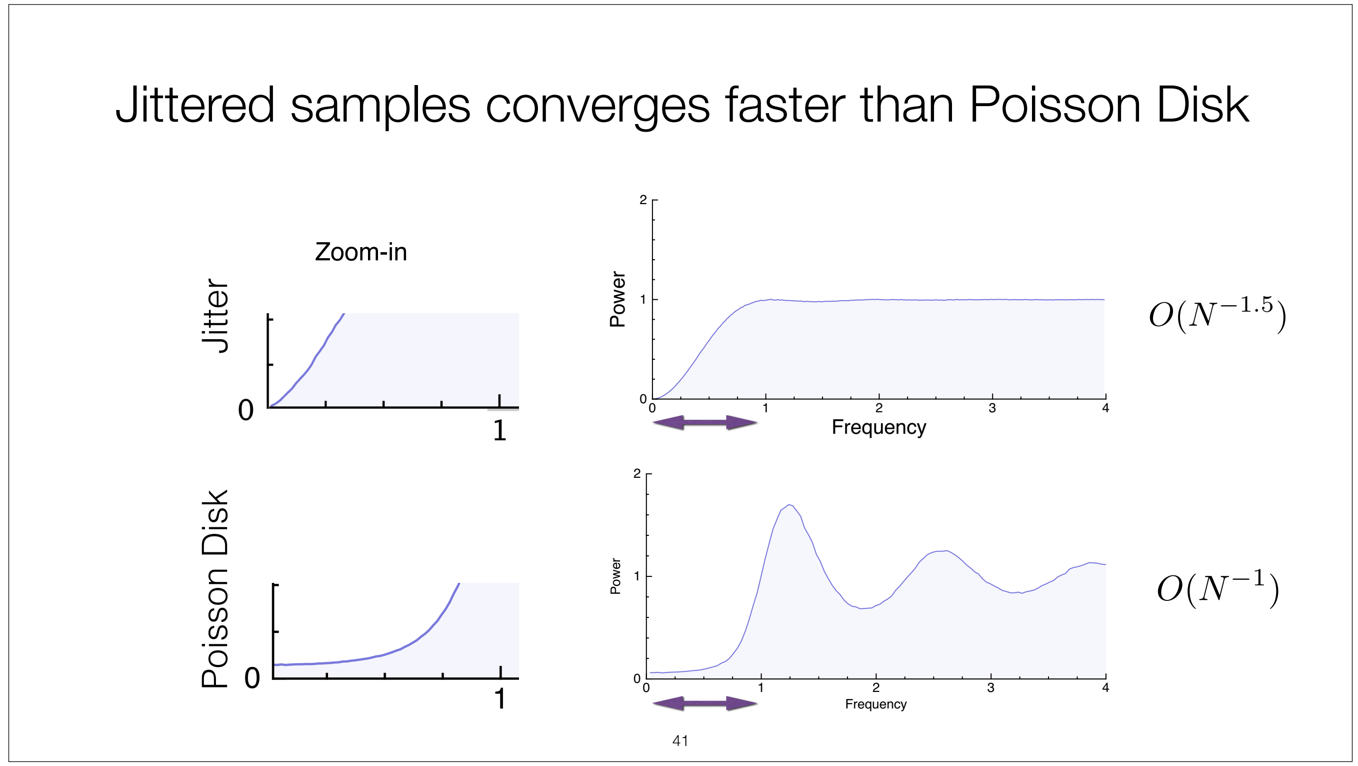


This can be explained by [CLICK] looking at the low-frequency region of these radial profiles.
[CLICK] For jittered samples, the radial profiles goes to zero near the DC (zero) frequency, whereas for Poisson disk samples the radial profile has an offset, which translates into it's bad convergence.

Jittered samples converges faster than Poisson Disk



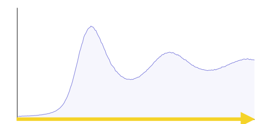
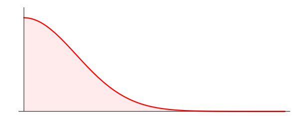
This can be explained by [CLICK] looking at the low-frequency region of these radial profiles.
[CLICK] For jittered samples, the radial profiles goes to zero near the DC (zero) frequency, whereas for Poisson disk samples the radial profile has an offset, which translates into it's bad convergence.



This can be explained by [CLICK] looking at the low-frequency region of these radial profiles.
[CLICK] For jittered samples, the radial profiles goes to zero near the DC (zero) frequency, whereas for Poisson disk samples the radial profile has an offset, which translates into it's bad convergence.

Convergence rate depends on the low frequency region

Samples' expected power spectrum	Integrand power spectrum

$$\mathrm{Var}(I_N) \propto$$  $$\times$$ 

Samplers	Worst Case	Best Case
Random	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Poisson Disk	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Jitter	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-2})$
CCVT	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-3})$

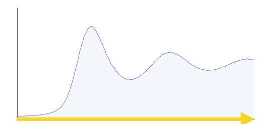
Pilleboue et al. [2015]

These convergence tools were developed for samplers with [CLICK] isotropic power spectra. However, the samplers used in practice have anisotropic energy distribution in their power spectrum. For example, ...

Convergence rate depends on the low frequency region

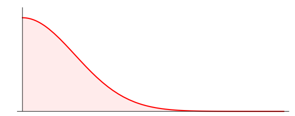
$$\mathrm{Var}(I_N) \propto$$

Samples' expected power spectrum



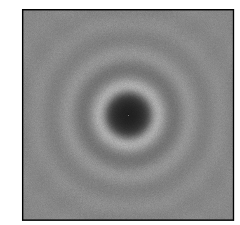
$\times$

Integrand power spectrum



Samplers	Worst Case	Best Case
Random	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Poisson Disk	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Jitter	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-2})$
CCVT	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-3})$

Isotropic Spectrum
Poisson Disk



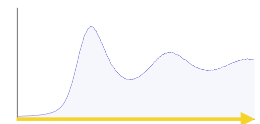
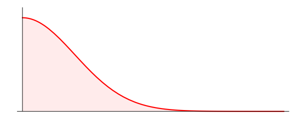
Pilleboue et al. [2015]

These convergence tools were developed for samplers with [CLICK] isotropic power spectra. However, the samplers used in practice have anisotropic energy distribution in their power spectrum. For example, ...

Convergence rate depends on the low frequency region

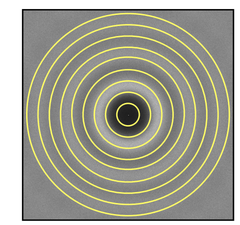
Samples' expected power spectrum

Integrand power spectrum

$$\mathrm{Var}(I_N) \propto$$  $\times$ 

Samplers	Worst Case	Best Case
Random	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Poisson Disk	$\mathcal{O}(N^{-1})$	$\mathcal{O}(N^{-1})$
Jitter	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-2})$
CCVT	$\mathcal{O}(N^{-1.5})$	$\mathcal{O}(N^{-3})$

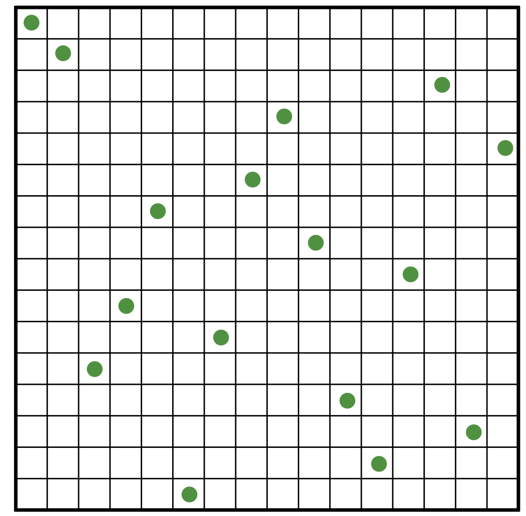
Isotropic Spectrum
Poisson Disk



Pilleboue et al. [2015]

These convergence tools were developed for samplers with [CLICK] isotropic power spectra. However, the samplers used in practice have anisotropic energy distribution in their power spectrum. For example, ...

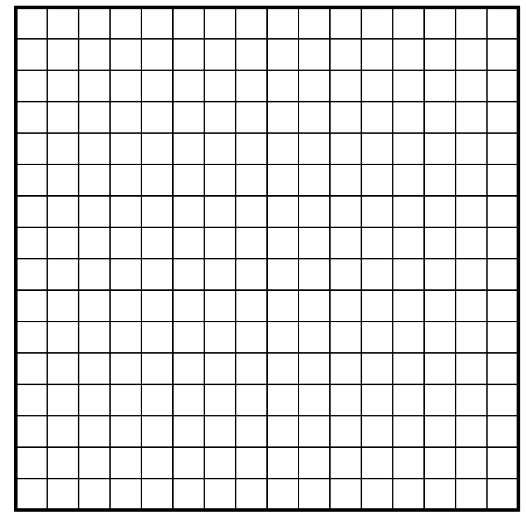
Latin Hypercube Sampler (N-rooks)



Slide after Wojciech Jarosz

…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.

Latin Hypercube Sampler (N-rooks)

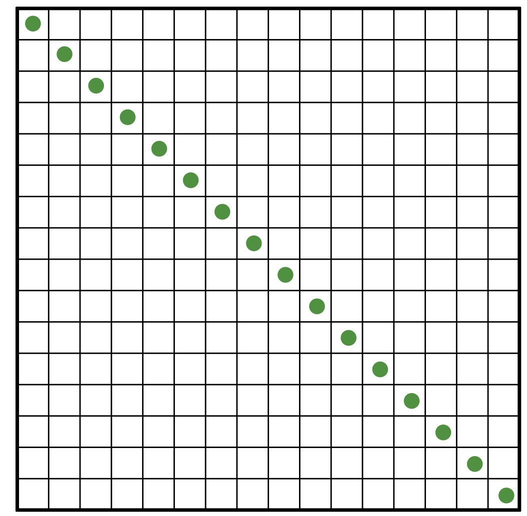


44 Slide after Wojciech Jarosz

…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.

Latin Hypercube Sampler (N-rooks)

Initialize



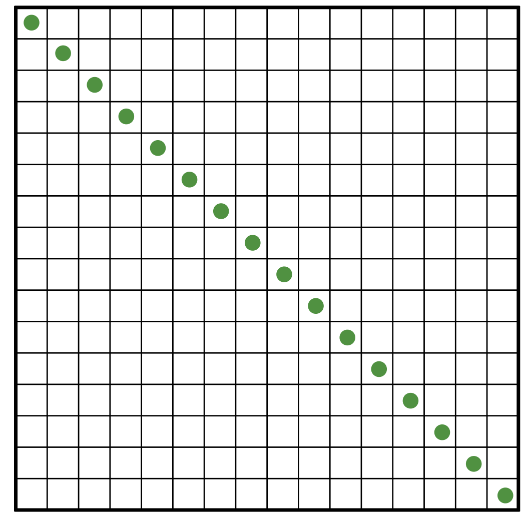
44

Slide after Wojciech Jarosz

…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.

Latin Hypercube Sampler (N-rooks)

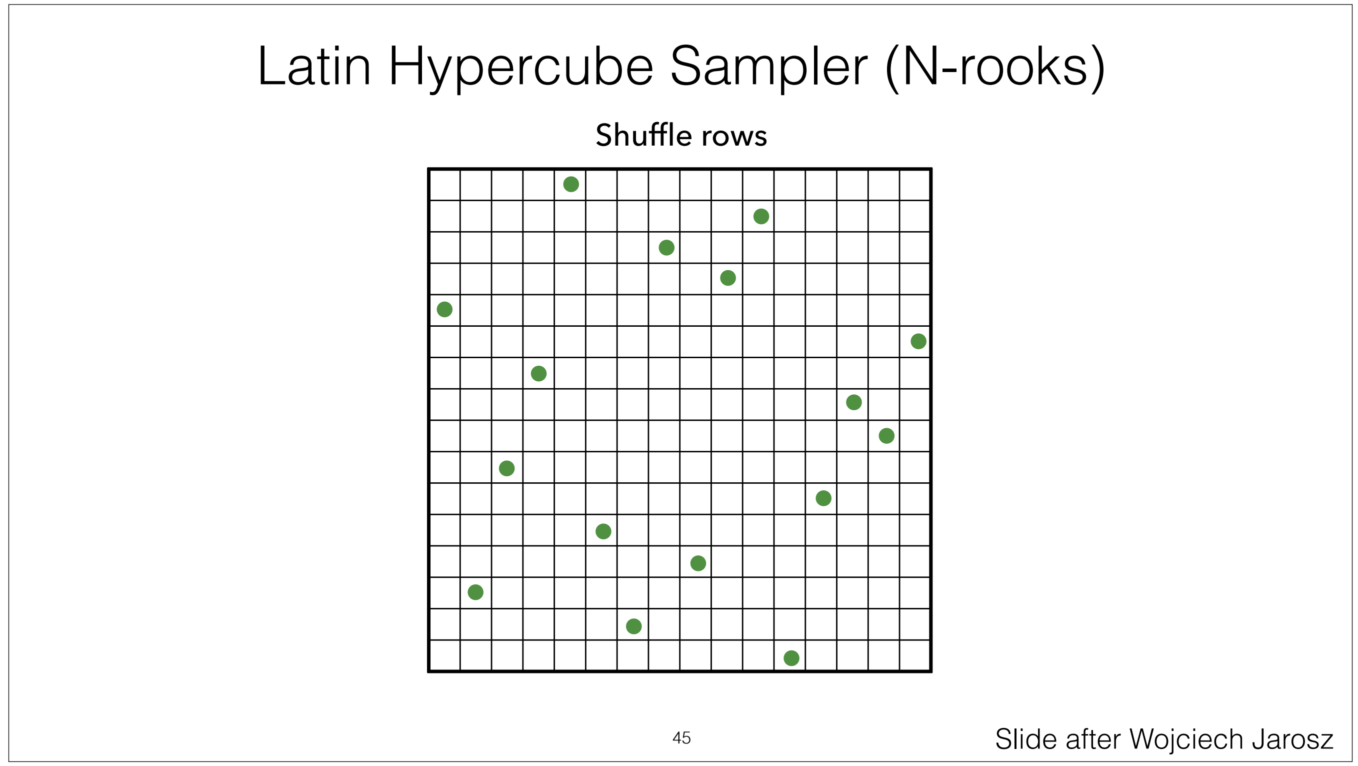
Shuffle rows



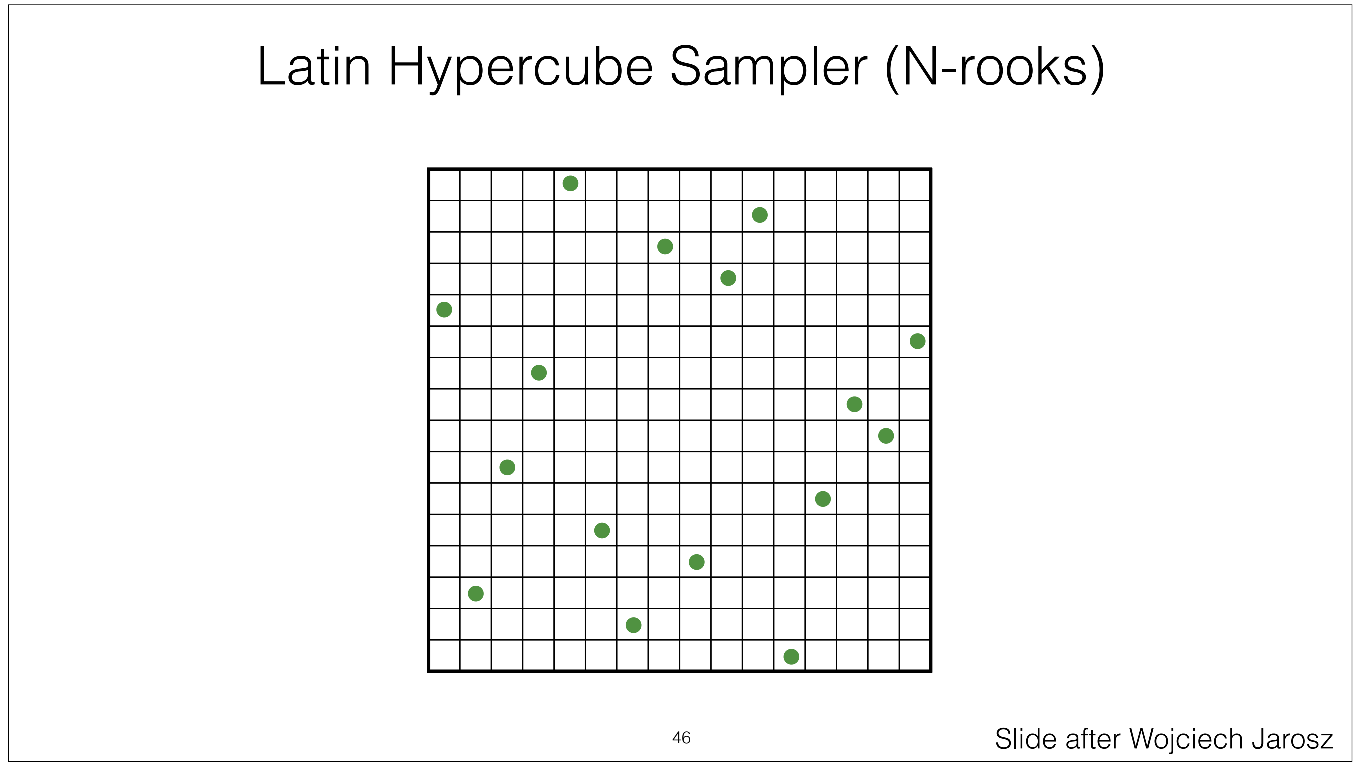
45

Slide after Wojciech Jarosz

…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.

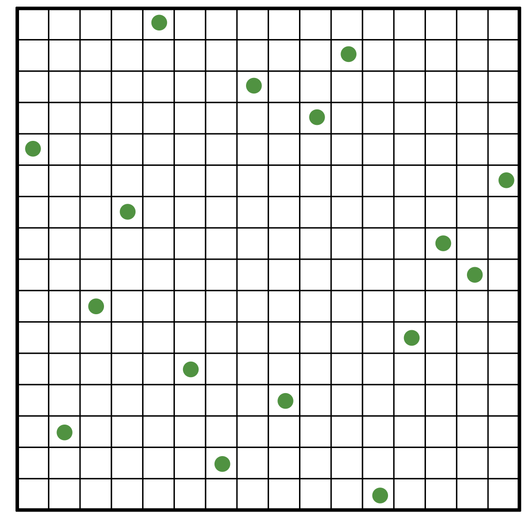


…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.



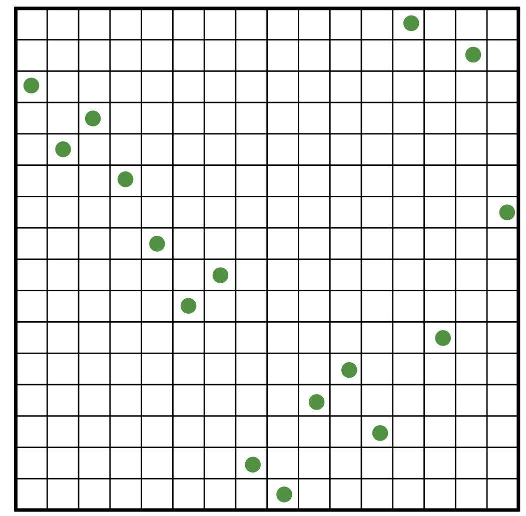
…a Latin hypercube sampler, that is constructed by first generating samples along a diagonal followed by randomly shuffling along the rows and then along the columns.

Latin Hypercube Sampler (N-rooks)



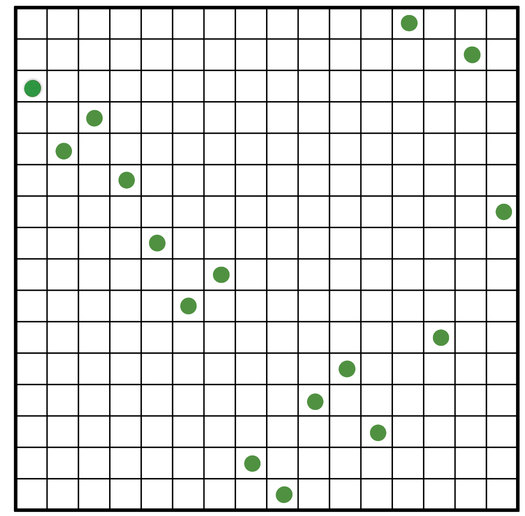
Slide after Wojciech Jarosz

Latin Hypercube Sampler (N-rooks)



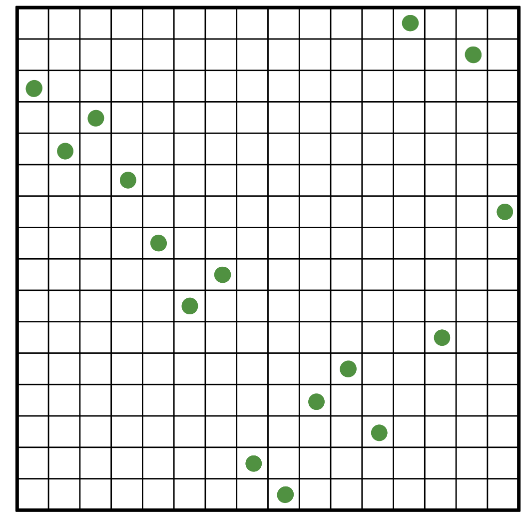
Slide after Wojciech Jarosz

Latin Hypercube Sampler (N-rooks)



Slide after Wojciech Jarosz

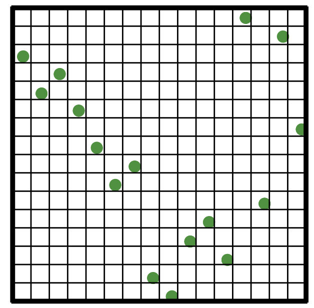
Latin Hypercube Sampler (N-rooks)



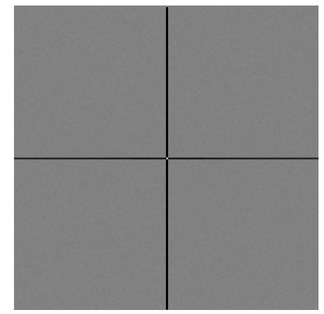
Slide after Wojciech Jarosz

As a result, the underlying power spectrum has an...

Anisotropic Sampling Power Spectra



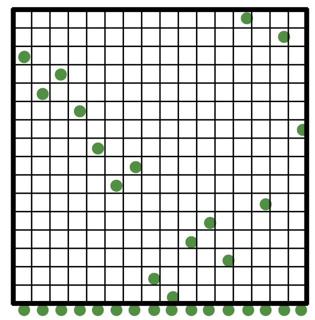
N-rooks /
Latin Hypercube



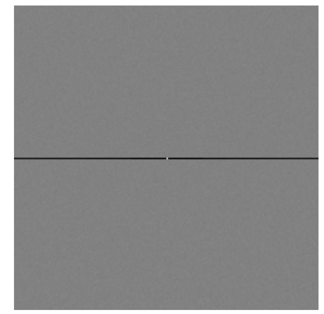
N-rooks
Spectrum

… anisotropic power spectrum with hairline structures visible as a dark cross in the middle. These hairline anisotropies are there due to the denser stratification along the X…

Anisotropic Sampling Power Spectra



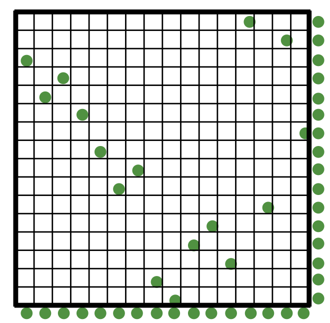
N-rooks /
Latin Hypercube



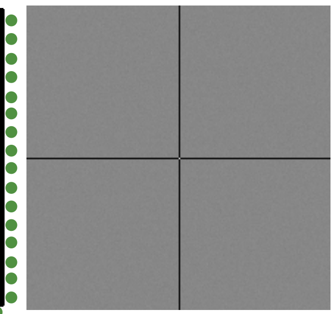
Spectrum

… anisotropic power spectrum with a dark cross in the middle. These hairline anisotropies are there due to the denser stratification along the X…

Anisotropic Sampling Power Spectra



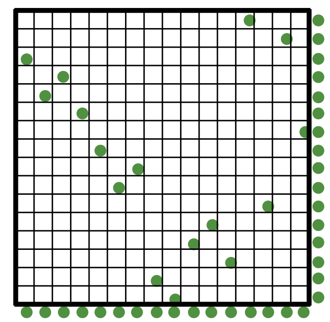
N-rooks /
Latin Hypercube



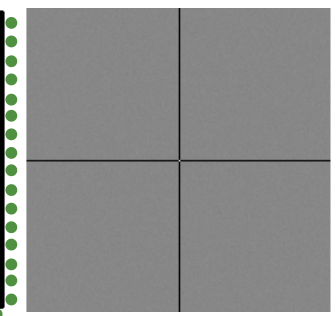
N-rooks
Spectrum

…and the Y-axis. It is also possible to directly obtain good 2D stratified samples…

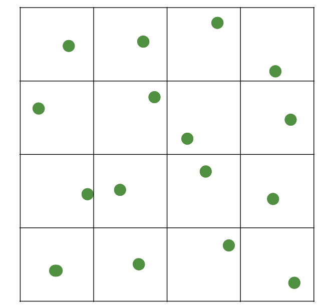
Anisotropic Sampling Power Spectra



N-rooks /
Latin Hypercube



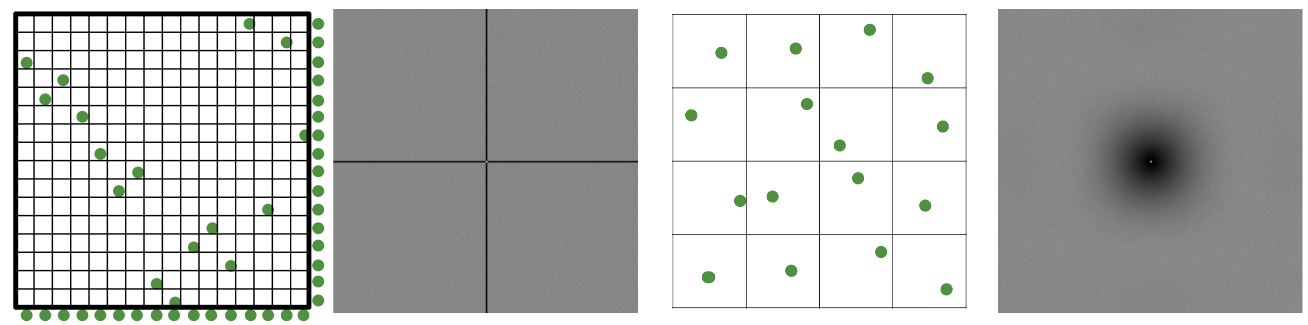
N-rooks
Spectrum



Jitter

…which has a power spectrum [CLICK] with a dark region around the center. Chiu and colleagues, found a better construction for these samples…

Anisotropic Sampling Power Spectra



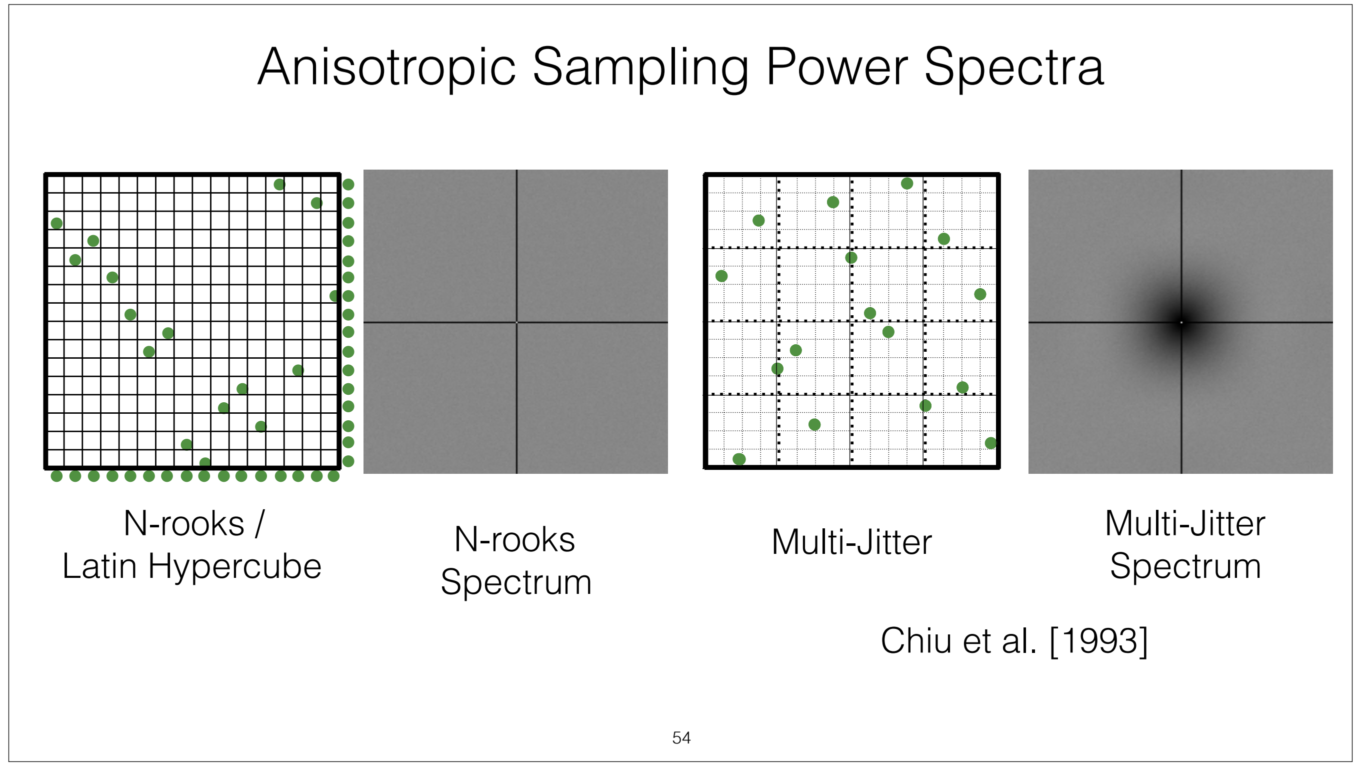
N-rooks / Latin Hypercube

N-rooks Spectrum

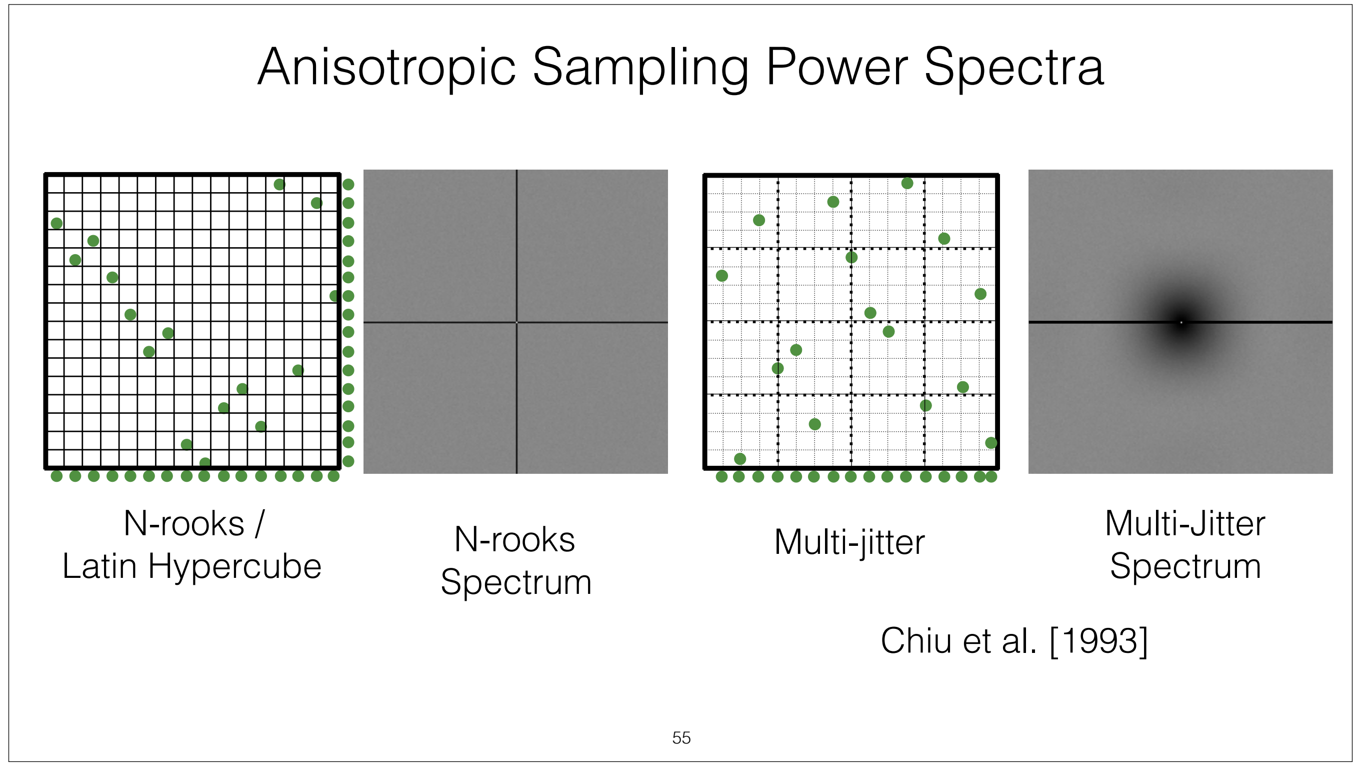
Jitter

Jitter Spectrum

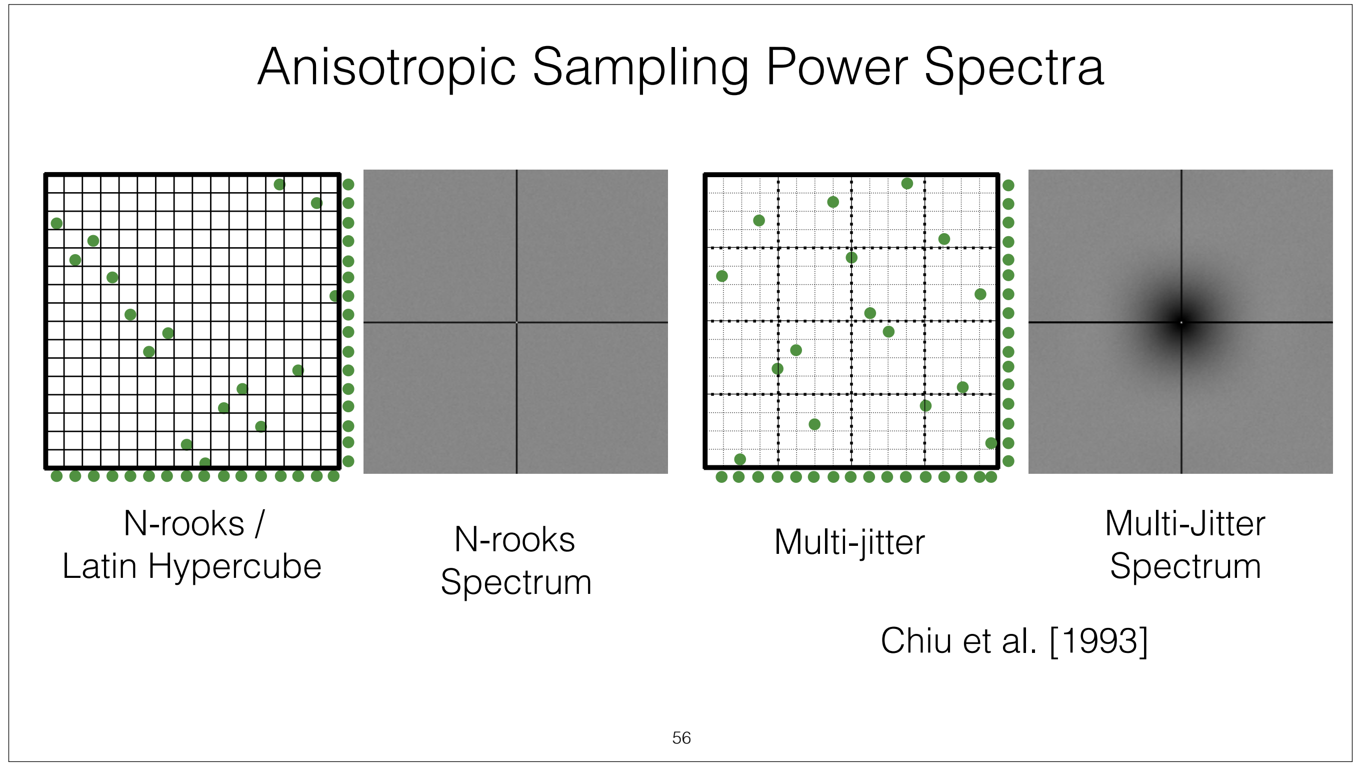
…which has a power spectrum [CLICK] with a dark region around the center. Chiu and colleagues, found a better construction for these samples…



…to obtain denser stratification…



…along the horizontal…

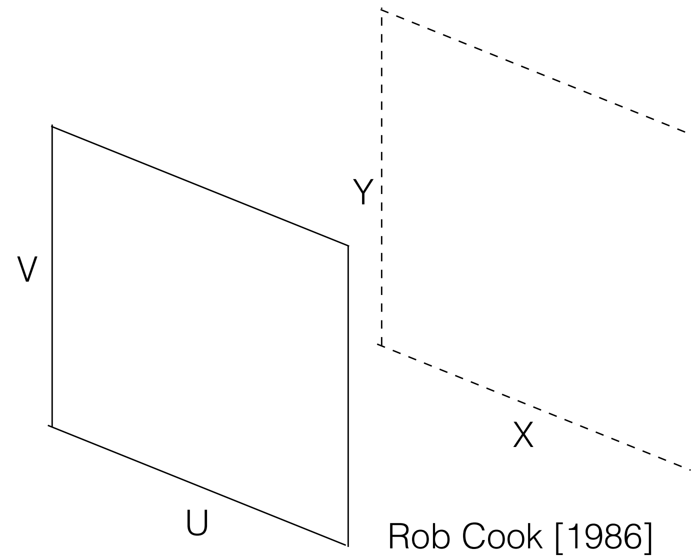


…and vertical axis, on top of 2D stratification, which results in multi-jittered samples with a hairline anisotropy along the canonical axes that is visible as a cross in the middle of it’s spectrum. The same ideas extend to…

Sampling in Higher Dimensions

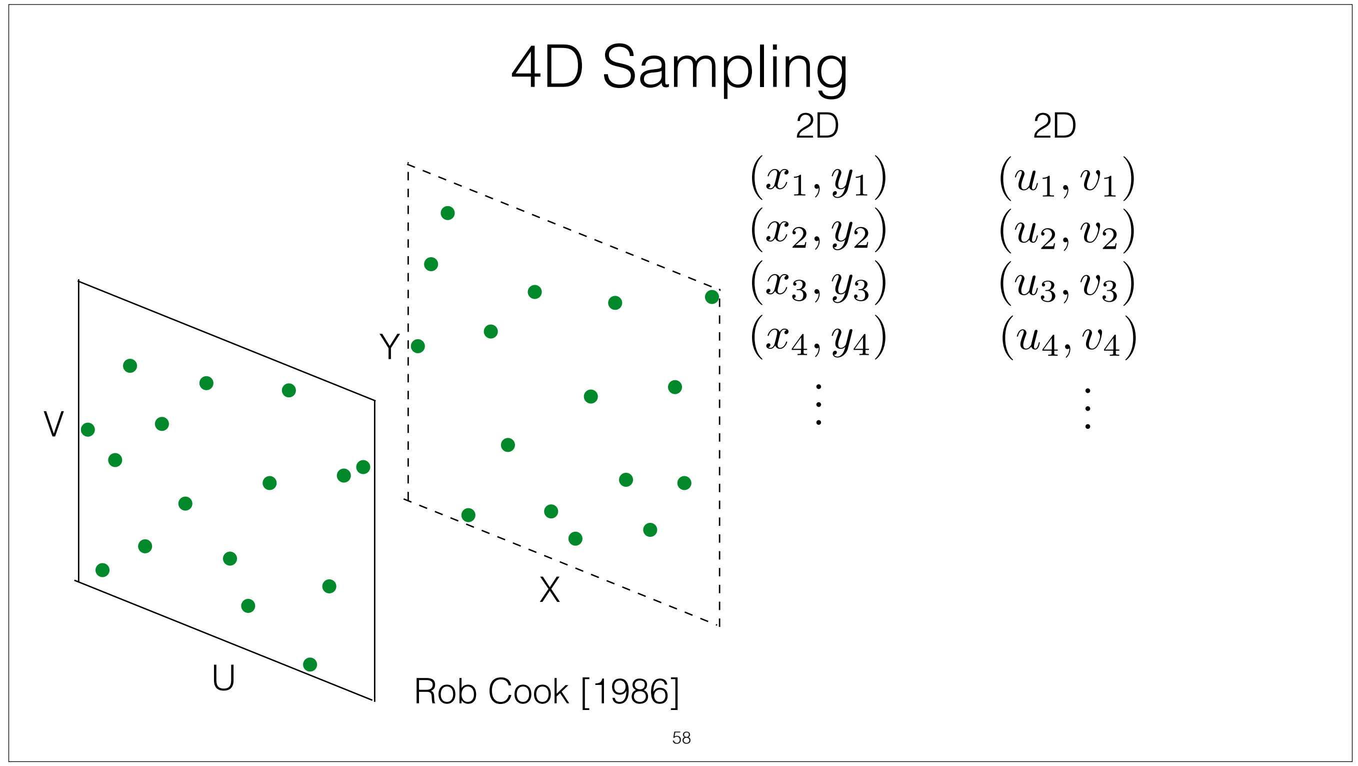
…to higher dimensions. For example, in 4D…

4D Sampling



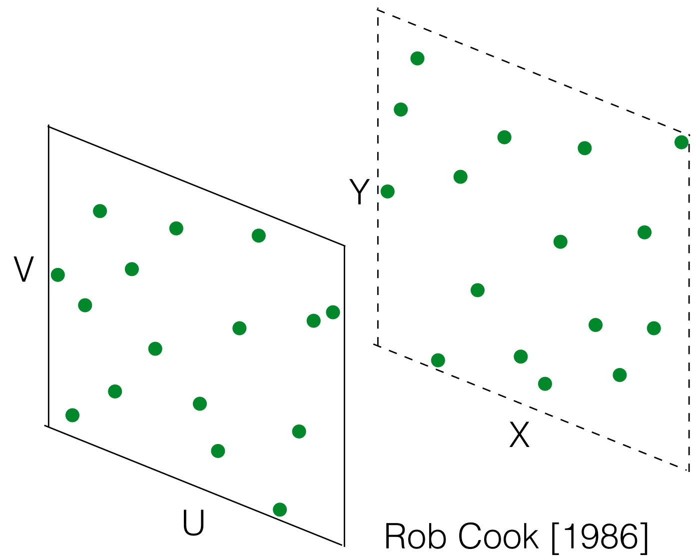
Rob Cook [1986]

…instead of directly sampling the full 4D space, Rob Cook in [1986] proposed to sample [CLICK] the lower 2D subspaces first, UV and XY here, and then randomly permute these 2D samples to form [CLICK] 4D tuples, which can then be used to evaluate an underlying 4D integrand. In practice…



…instead of directly sampling the full 4D space, Rob Cook in [1986] proposed to sample [CLICK] the lower 2D subspaces first, UV and XY here, and then randomly permute these 2D samples to form [CLICK] 4D tuples, which can then be used to evaluate an underlying 4D integrand. In practice…

4D Sampling



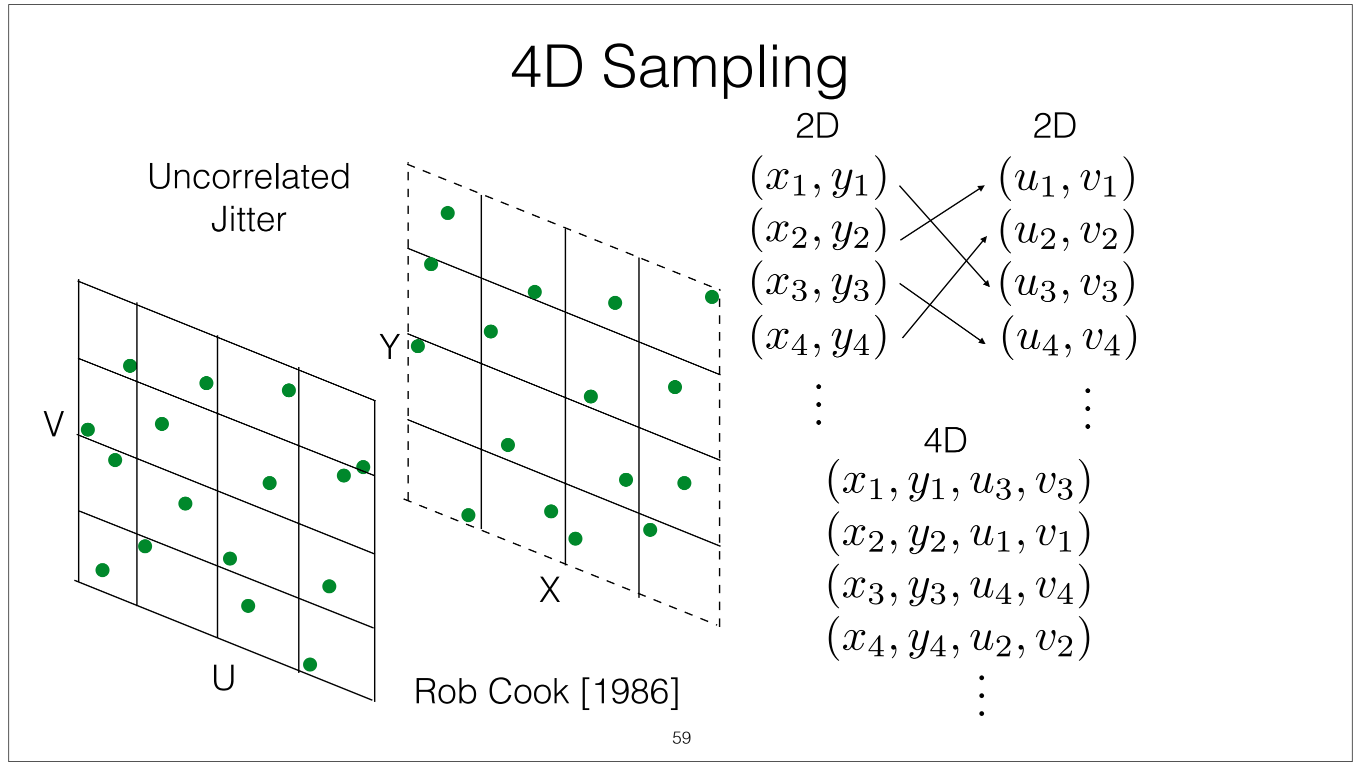
Rob Cook [1986]

2D		2D
(x_1, y_1)		(u_1, v_1)
(x_2, y_2)		(u_2, v_2)
(x_3, y_3)		(u_3, v_3)
(x_4, y_4)		(u_4, v_4)
$\vdots$		$\vdots$

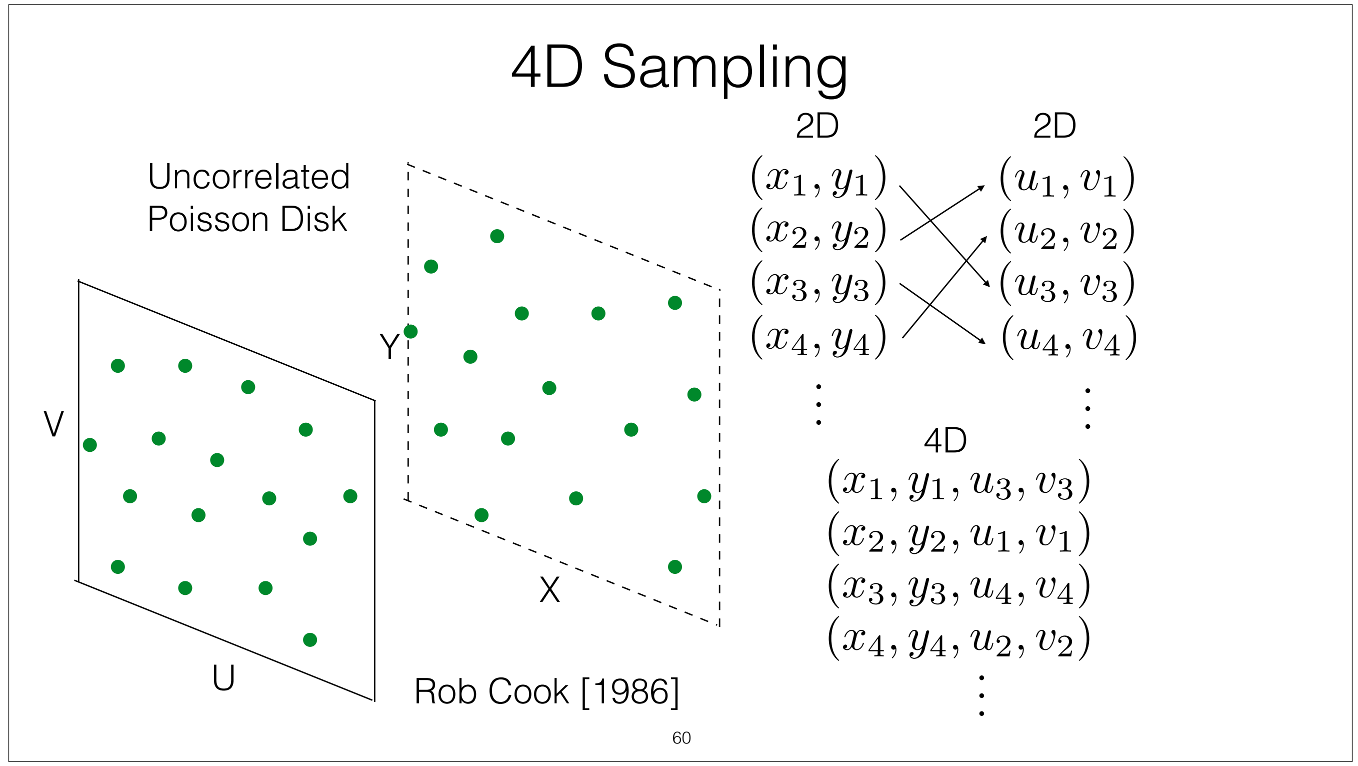
4D

$$(x_1, y_1, u_3, v_3)$$
$$(x_2, y_2, u_1, v_1)$$
$$(x_3, y_3, u_4, v_4)$$
$$(x_4, y_4, u_2, v_2)$$
$$\vdots$$

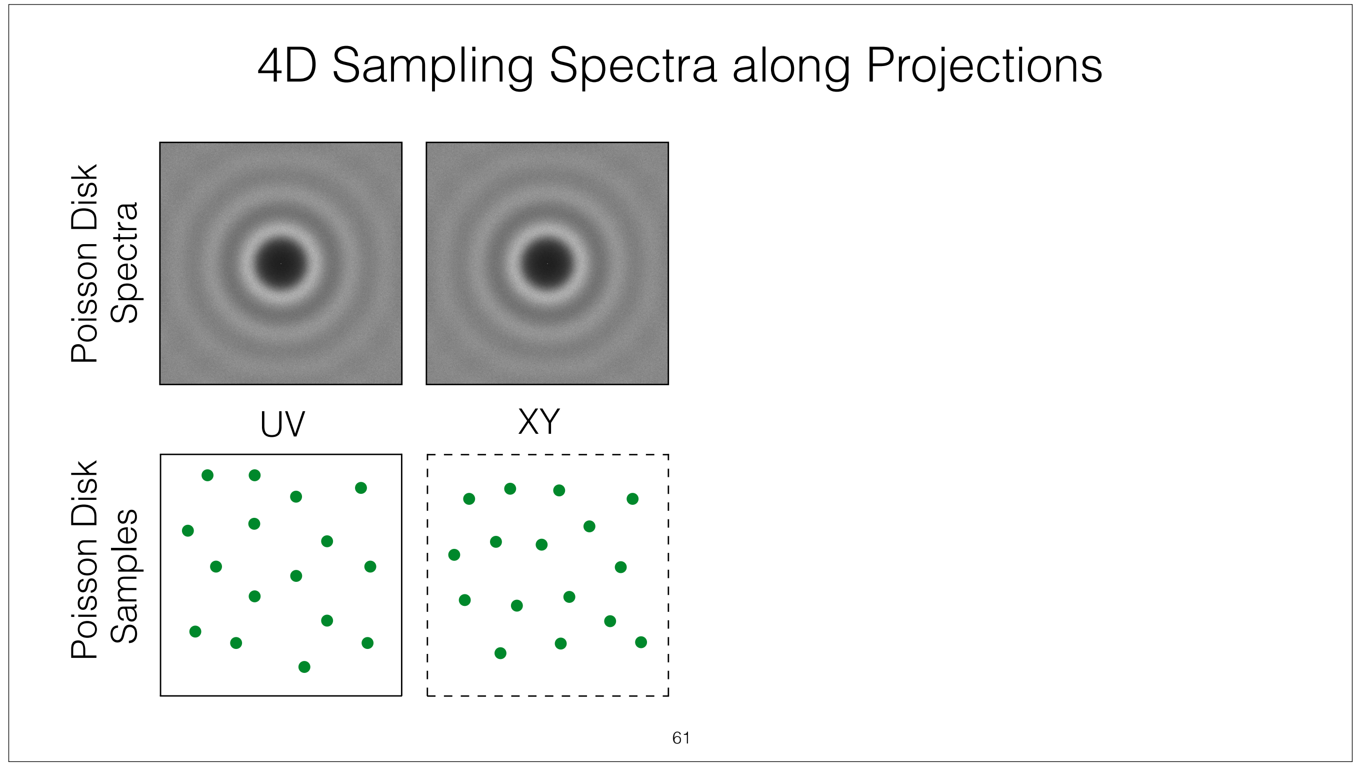
…instead of directly sampling the full 4D space, Rob Cook in [1986] proposed to sample [CLICK] the lower 2D subspaces first, UV and XY here, and then randomly permute these 2D samples to form [CLICK] 4D tuples, which can then be used to evaluate an underlying 4D integrand. In practice…



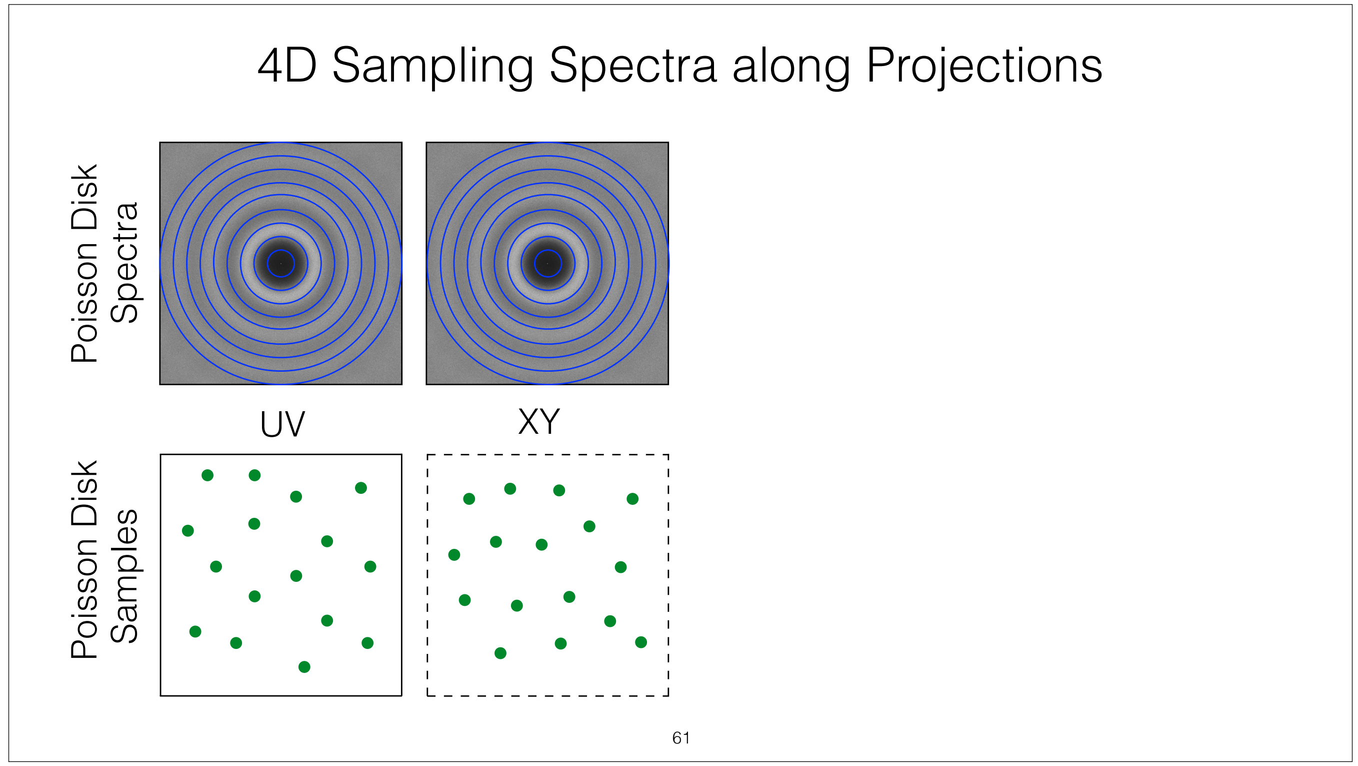
…rendering systems tend to use jittered samples on these 2D subspaces. It is, however, beneficial to use more sophisticated samplers like Poisson disk which are more uniform. If we look at the power spectra for these underlying 2D projections…



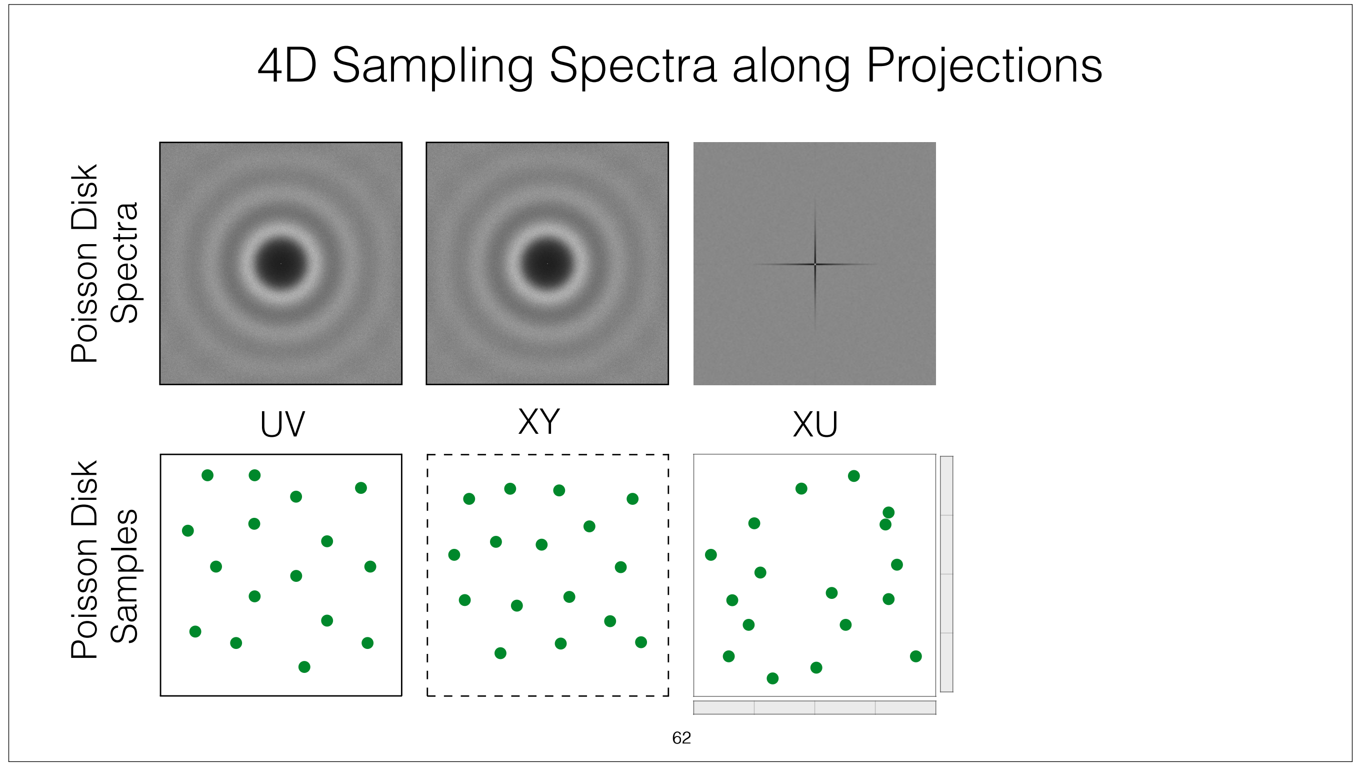
…rendering systems tend to use jittered samples on these 2D subspaces. It is, however, beneficial to use more sophisticated samplers like Poisson disk which are more uniform. If we look at the power spectra for these underlying 2D projections…



…they are [CLICK] isotropic in nature. However, if look at the mixed 2D projections XU…

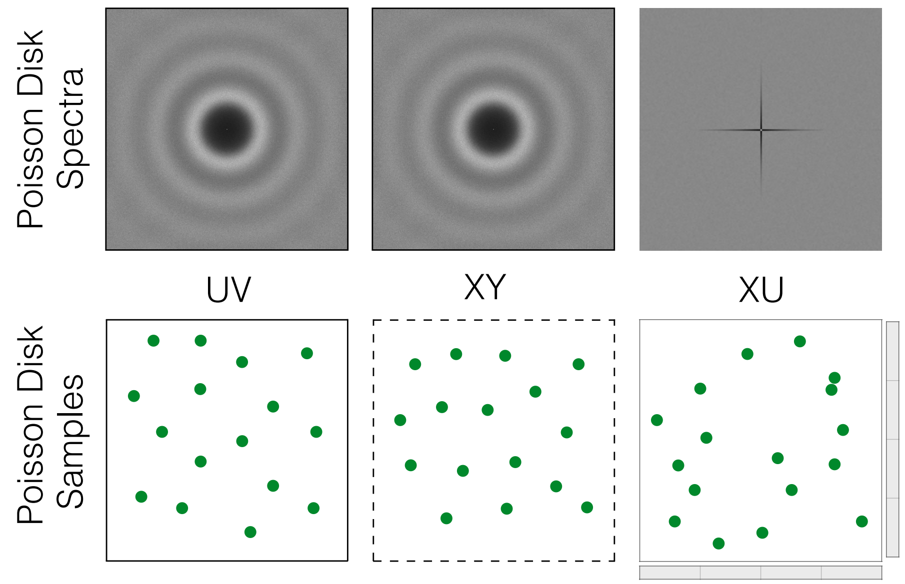


…they are [CLICK] isotropic in nature. However, if look at the mixed 2D projections XU...

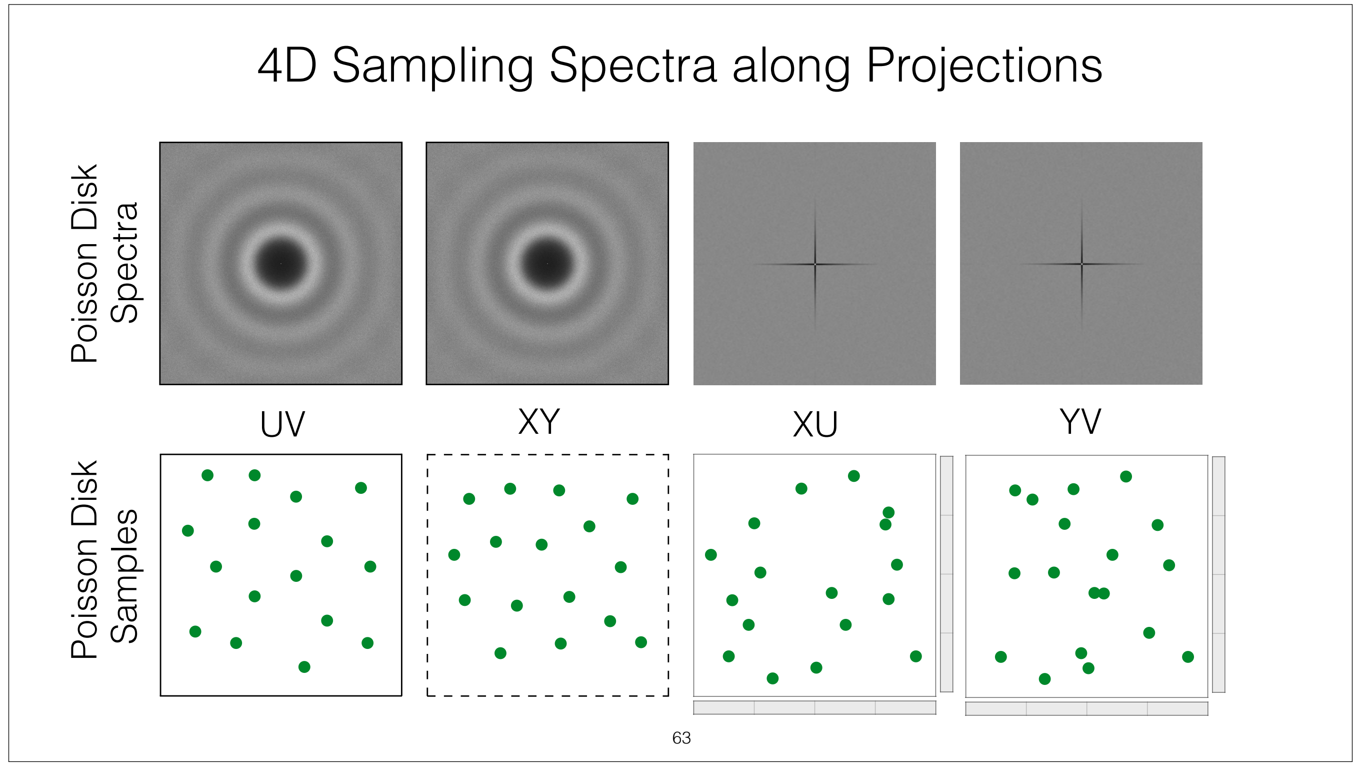


... it has hairline anisotropy along the [CLICK] axes. [CLICK] The same is true for YV projections.

4D Sampling Spectra along Projections



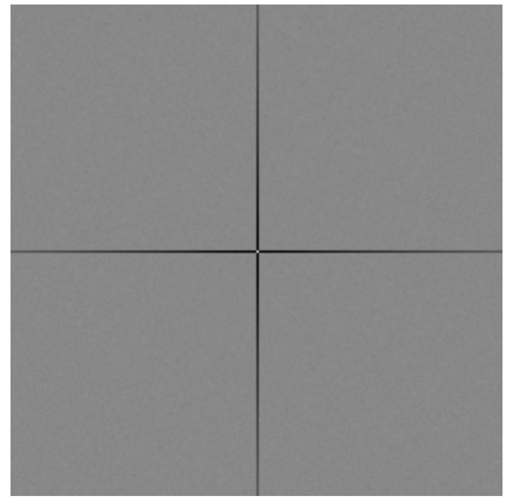
... it has hairline anisotropy along the [CLICK] axes. [CLICK] The same is true for YV projections.



Let's look at the power spectrum of an N-rooks expected power spectrum to understand the effect of these anisotropic structures on the variance convergence rate.

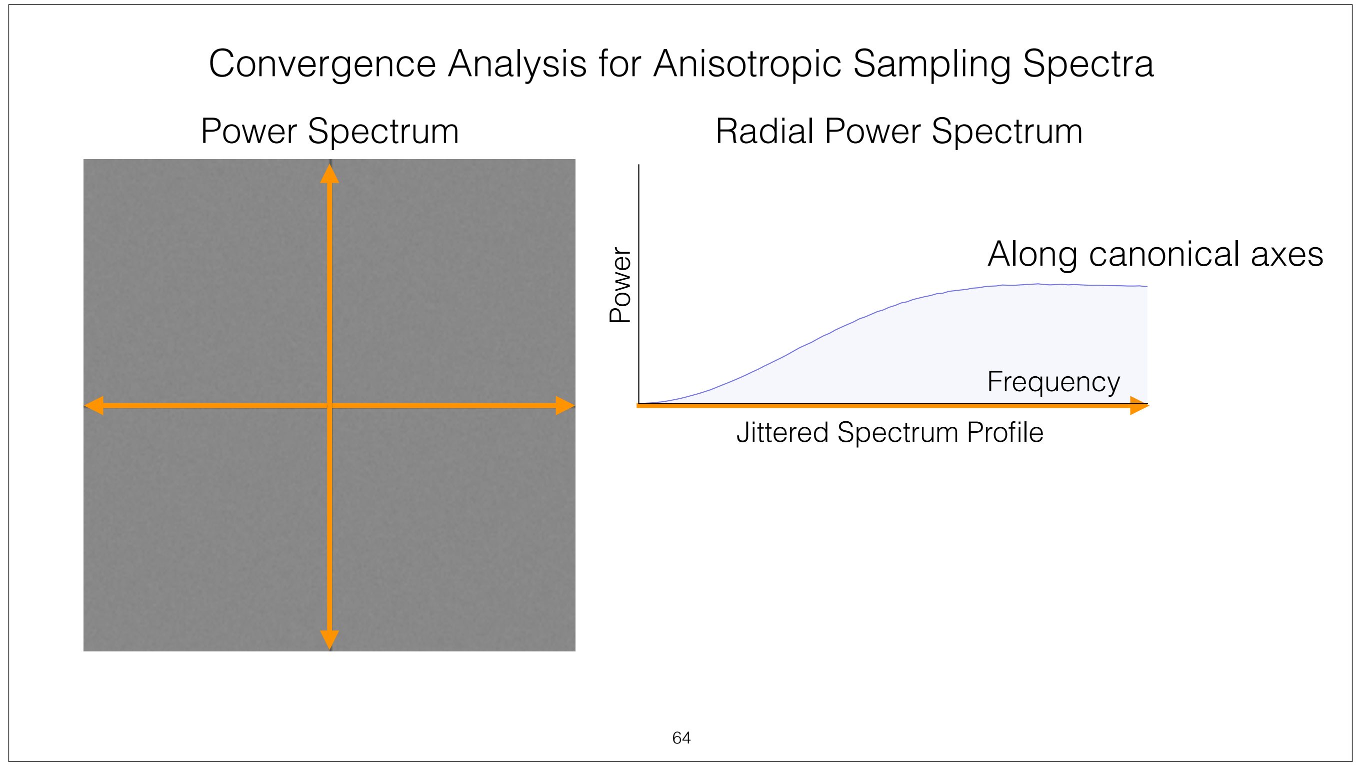
Convergence Analysis for Anisotropic Sampling Spectra

Power Spectrum

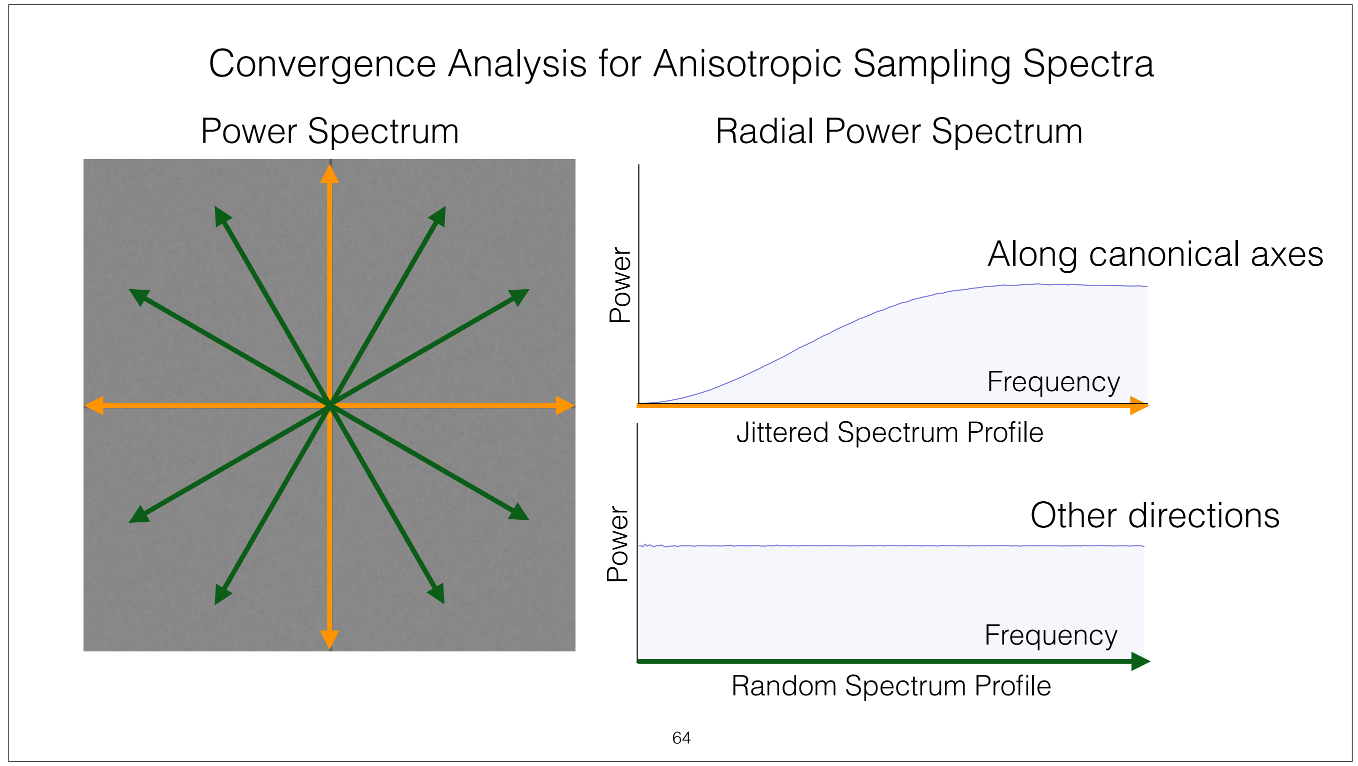


Radial Power Spectrum

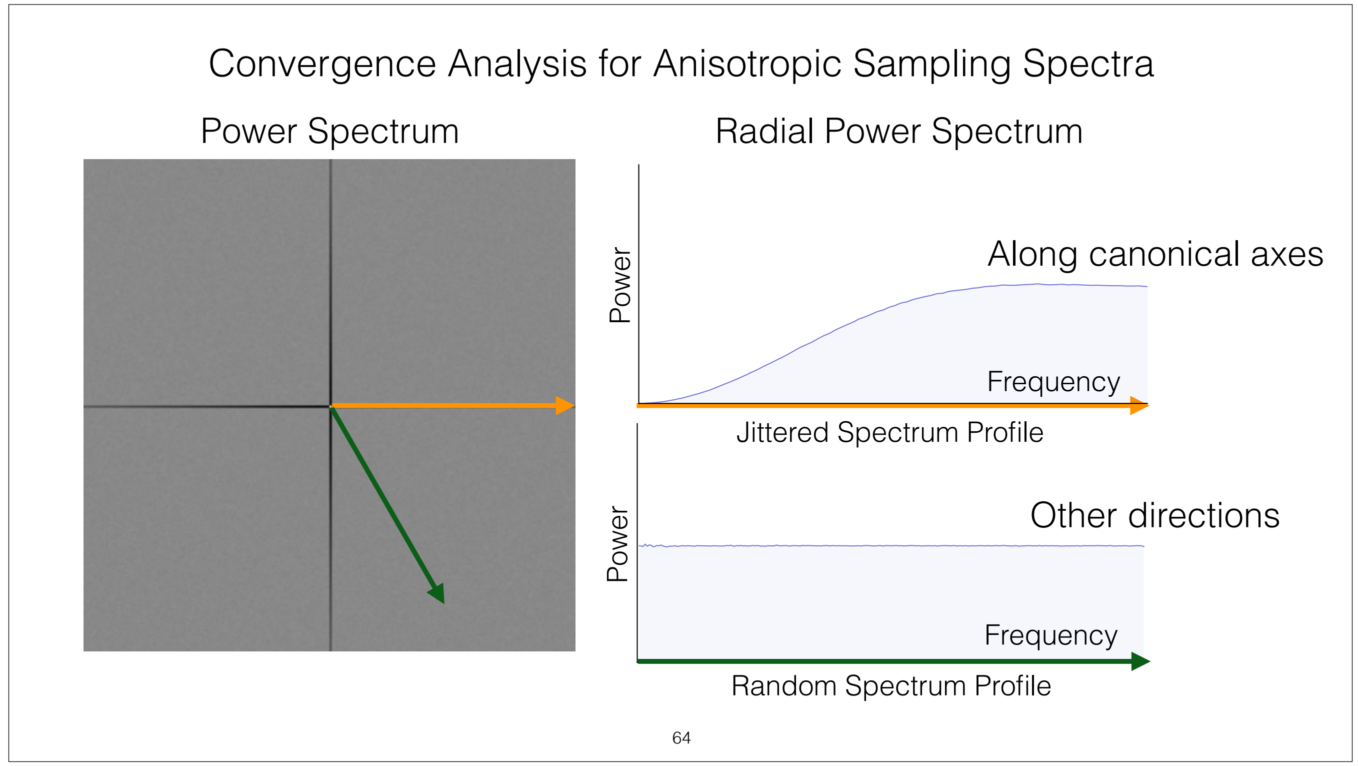
N-rooks expected spectrum has [CLICK] the same radial profile along the canonical axes, and [CLICK] a constant radial profile along all other directions. For the convergence rate, we only need…



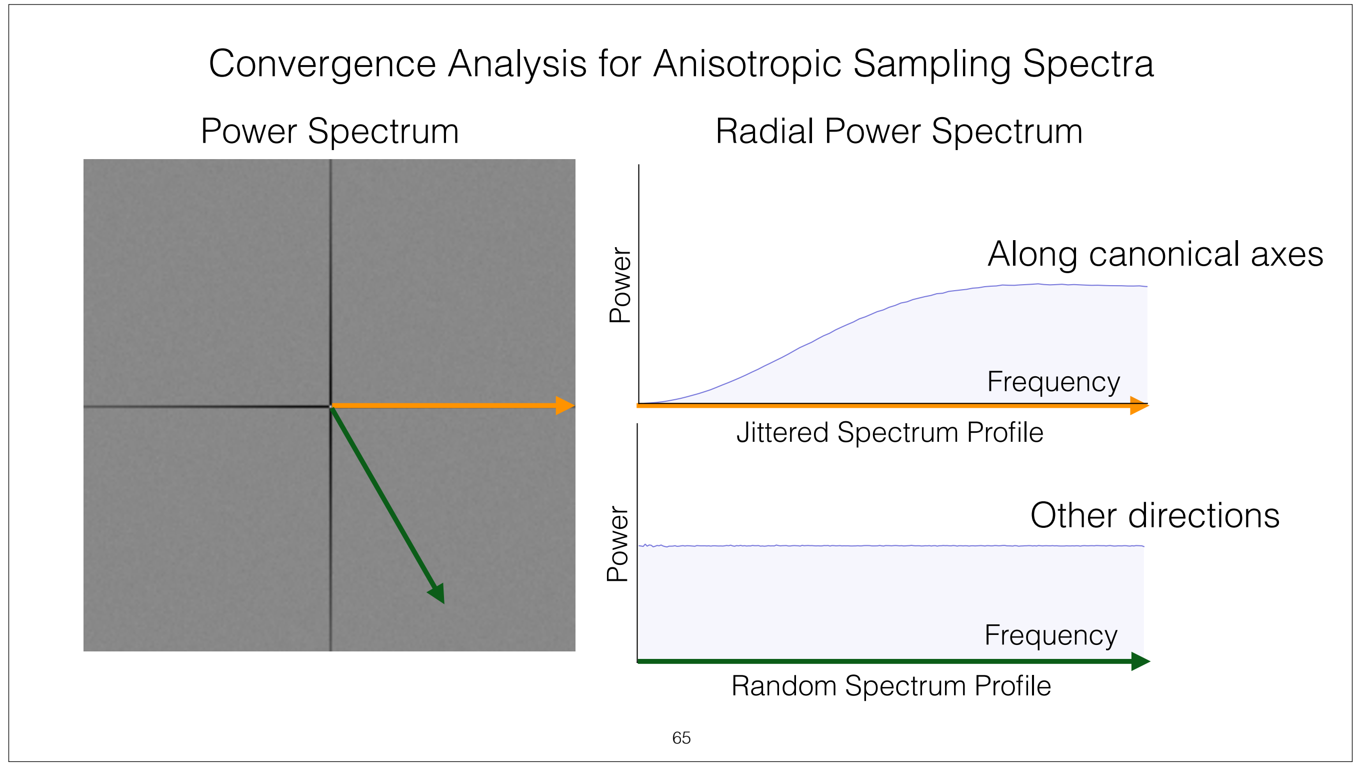
N-rooks expected spectrum has [CLICK] the same radial profile along the canonical axes, and [CLICK] a constant radial profile along all other directions. For the convergence rate, we only need…



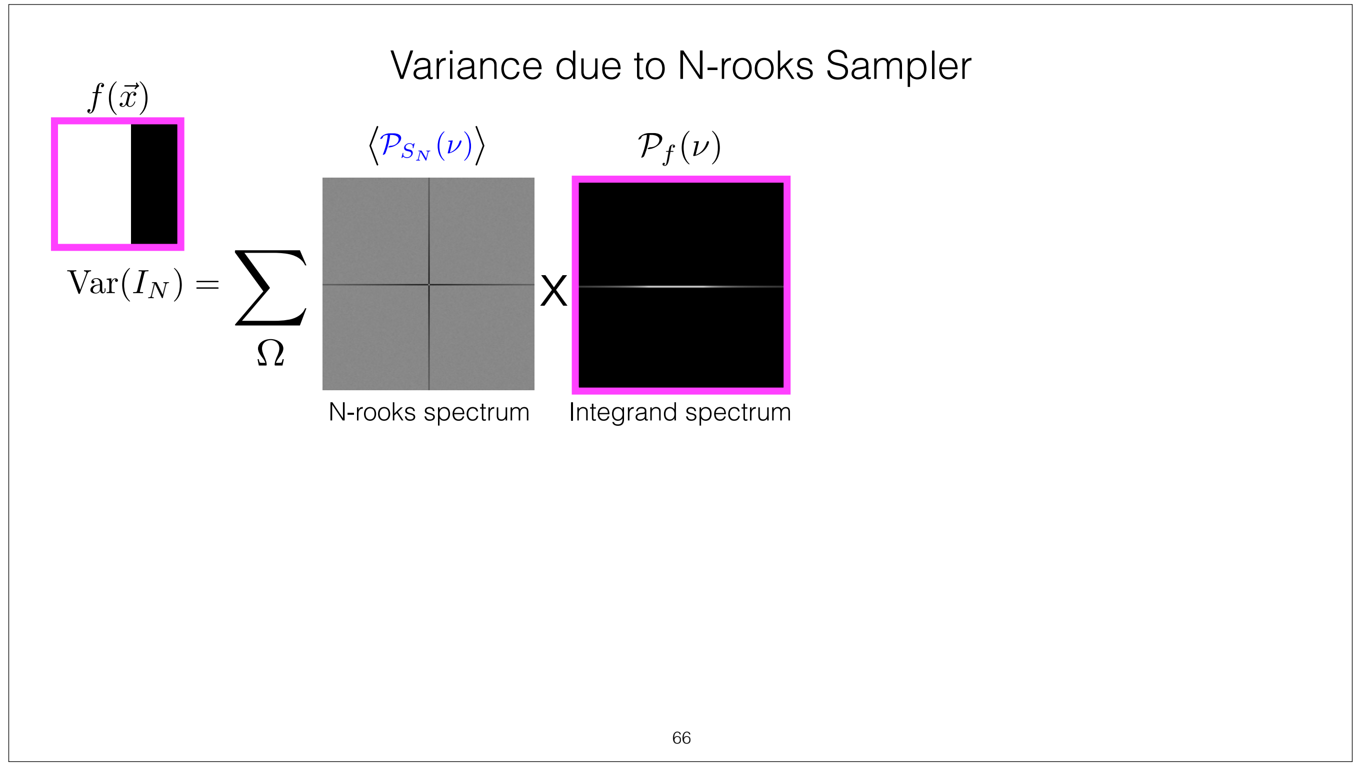
N-rooks expected spectrum has [CLICK] the same radial profile along the canonical axes, and [CLICK] a constant radial profile along all other directions. For the convergence rate, we only need…



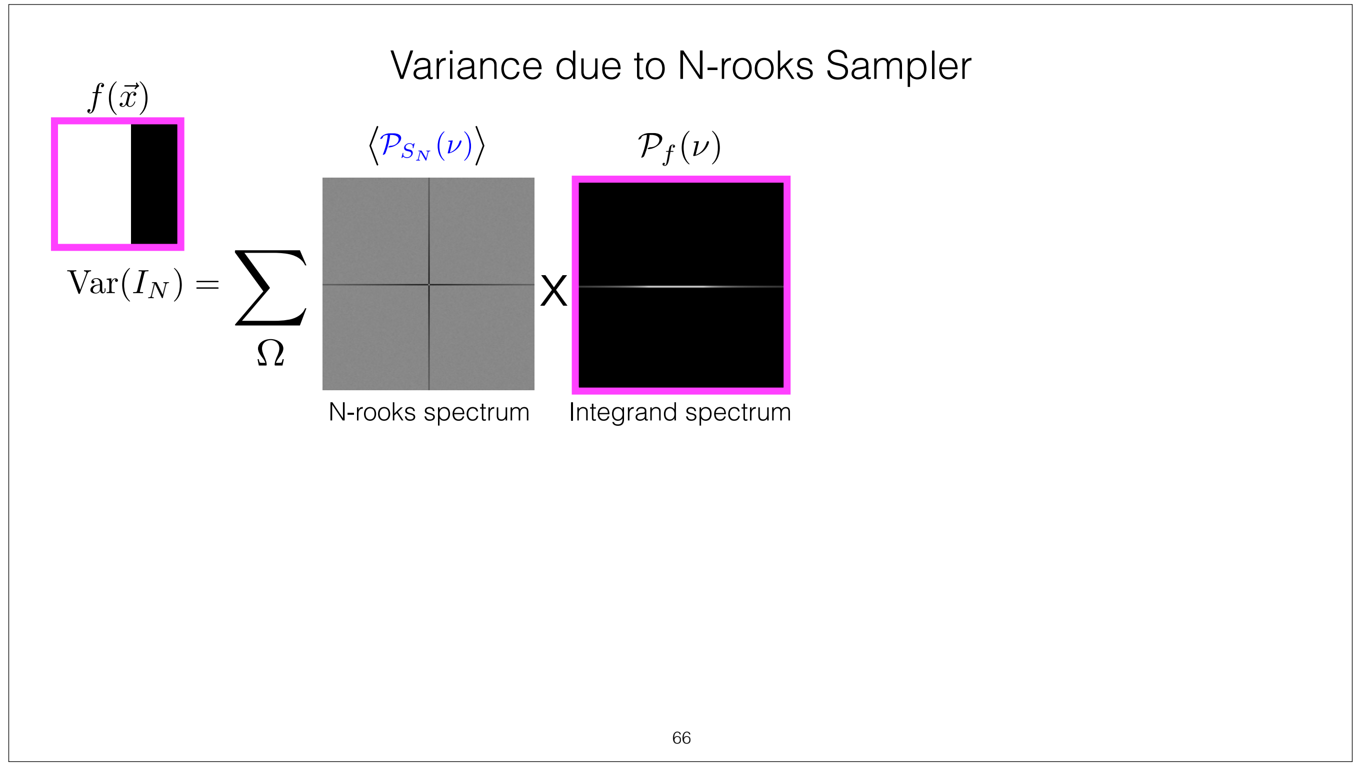
N-rooks expected spectrum has [CLICK] the same radial profile along the canonical axes, and [CLICK] a constant radial profile along all other directions. For the convergence rate, we only need…



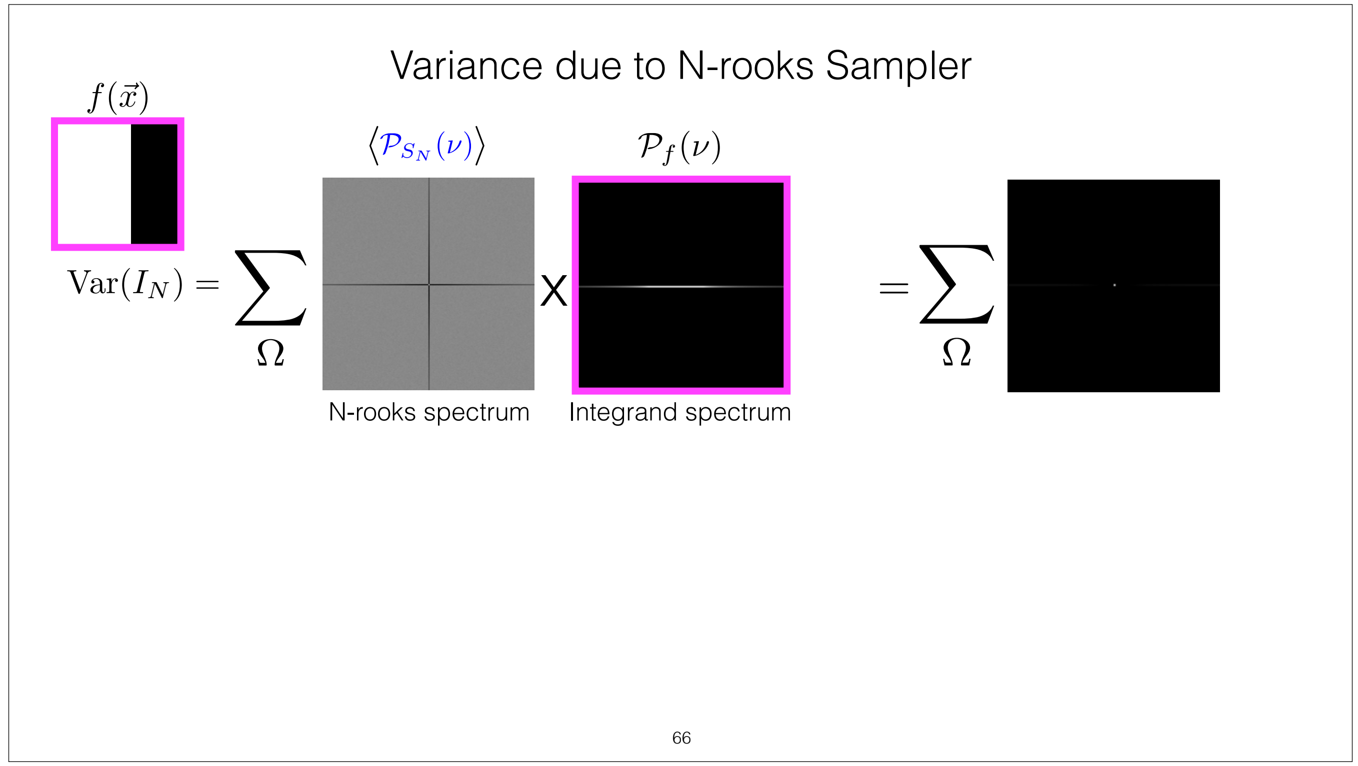
… one of the canonical direction (shown in orange) and one of the direction from the rest of the spectrum (shown in green) since the behavior is the same in all other directions. Now, depending on the integrands we can observe different convergence rates for the same sampler. For example,…



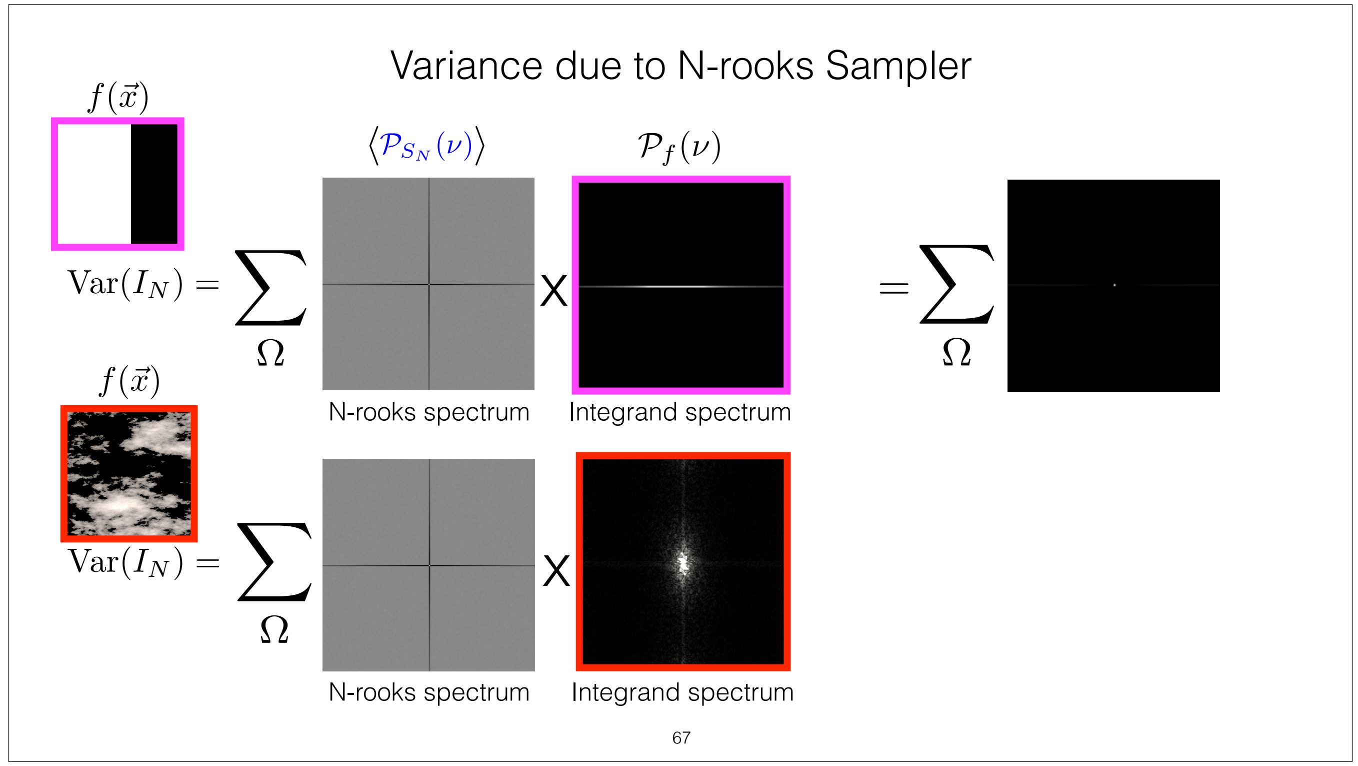
...for a step function with all the energy in the spectrum along the horizontal direction. Due to the dark hairline anisotropy [CLICK] present in the sampling spectrum, their [CLICK] product goes down very quickly, resulting in huge variance reduction and good asymptotic convergence. However, for the second pixel…



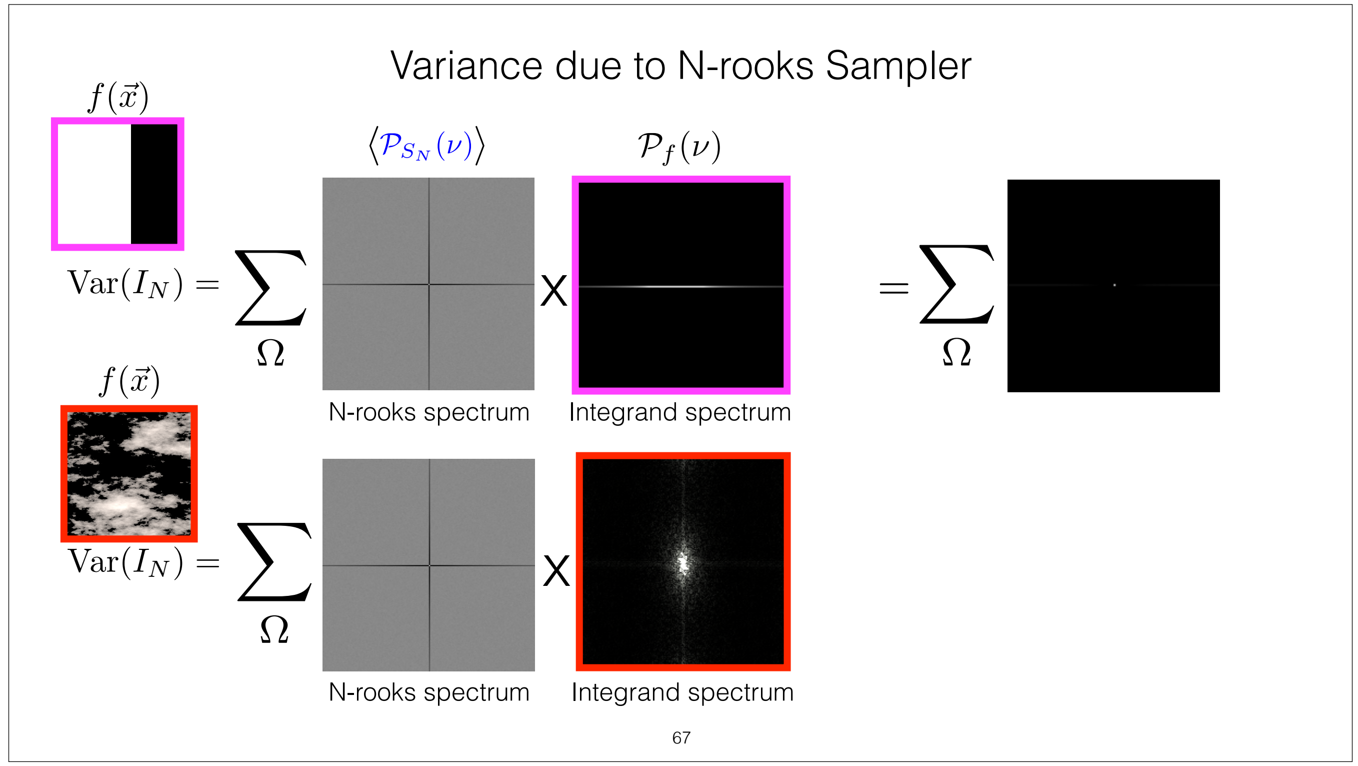
...for a step function with all the energy in the spectrum along the horizontal direction. Due to the dark hairline anisotropy [CLICK] present in the sampling spectrum, their [CLICK] product goes down very quickly, resulting in huge variance reduction and good asymptotic convergence. However, for the second pixel…



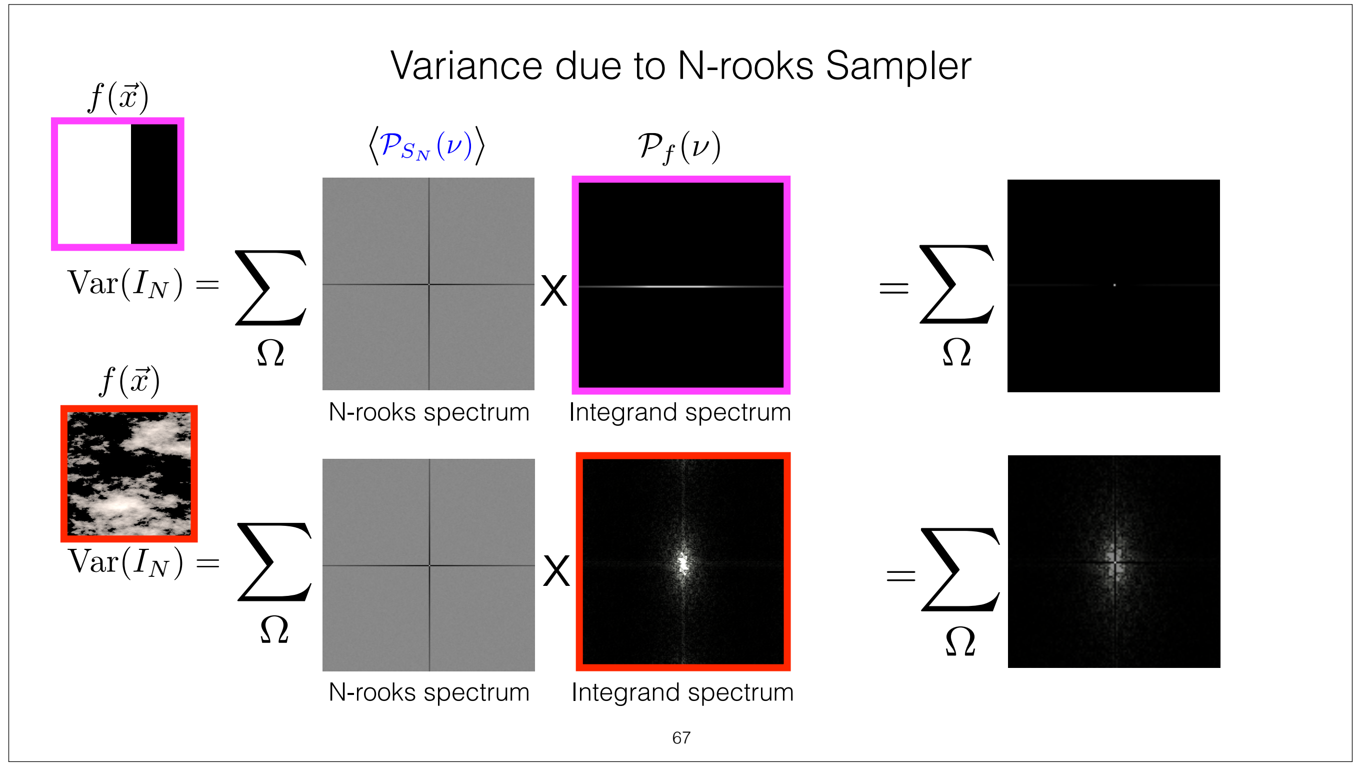
...for a step function with all the energy in the spectrum along the horizontal direction. Due to the dark hairline anisotropy [CLICK] present in the sampling spectrum, their [CLICK] product goes down very quickly, resulting in huge variance reduction and good asymptotic convergence. However, for the second pixel…



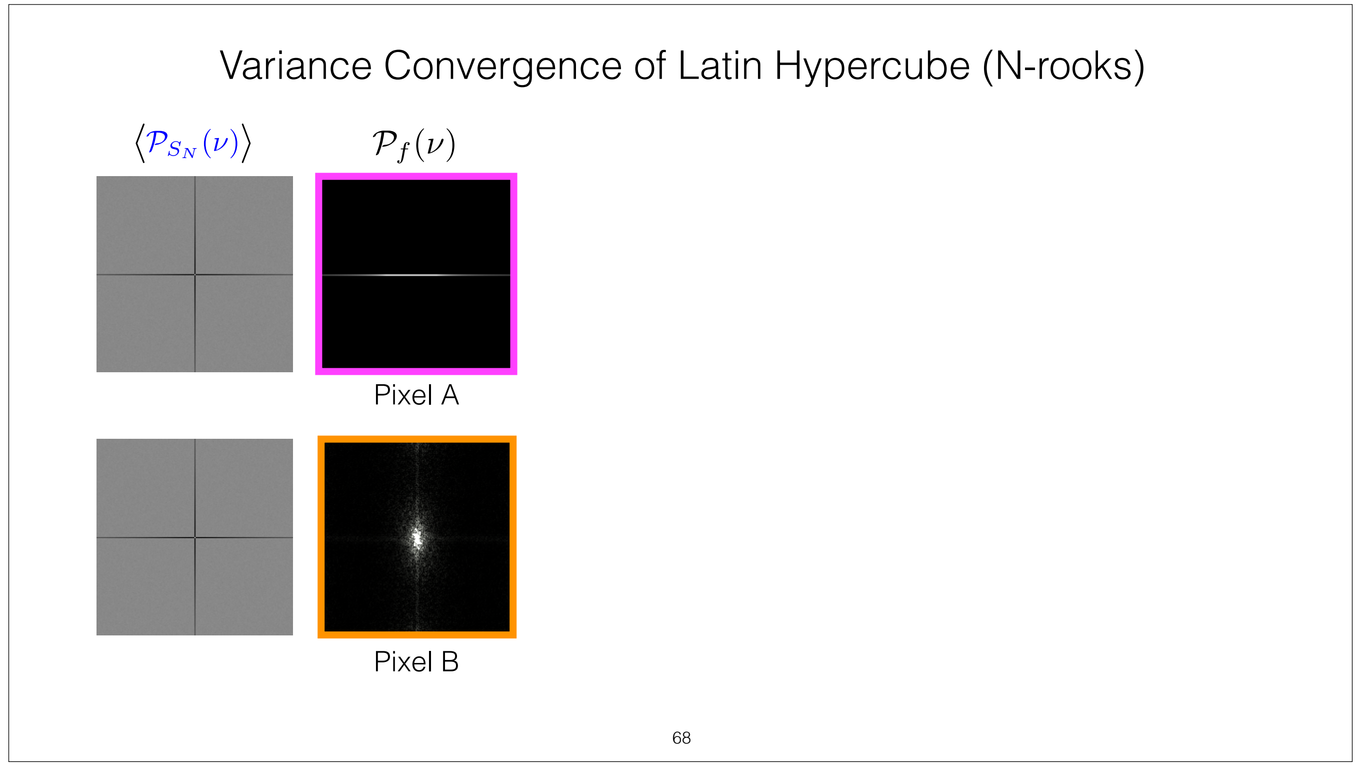
… since the integrand spectrum has energy spread over all the directions, the [CLICK] hairline anisotropy of the sampling spectrum [CLICK] does not significantly reduce the product, resulting in higher variance. We further verified this..



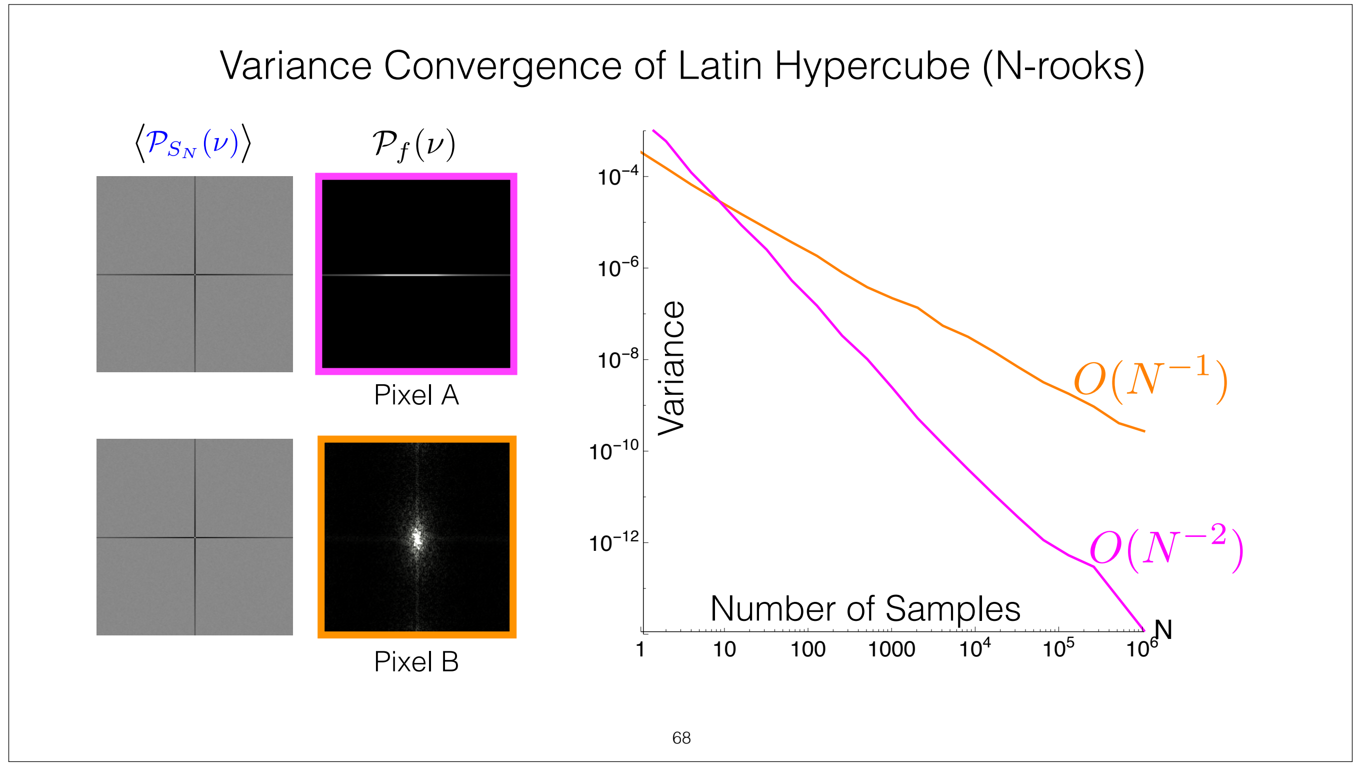
… since the integrand spectrum has energy spread over all the directions, the [CLICK] hairline anisotropy of the sampling spectrum [CLICK] does not significantly reduce the product, resulting in higher variance. We further verified this..



… since the integrand spectrum has energy spread over all the directions, the [CLICK] hairline anisotropy of the sampling spectrum [CLICK] does not significantly reduce the product, resulting in higher variance. We further verified this..



…experimentally, where we plot variance with increasing sample count. This shows that if we can align the anisotropic structures of the sampling spectrum Ps with that of the integrand spectrum Pf, we can gain huge variance reductions, as shown with the magenta curve. But in most scenarios…



…experimentally, where we plot variance with increasing sample count. This shows that if we can align the anisotropic structures of the sampling spectrum Ps with that of the integrand spectrum Pf, we can gain huge variance reductions, as shown with the magenta curve. But in most scenarios…

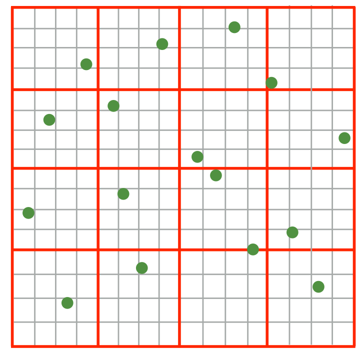
Non-Axis Aligned Integrand Spectra

$\mathcal{P}_f(\nu)$

Integrand Spectrum

…the underlying integrand spectrum has arbitrary orientation. If we choose to sample this function…

Non-Axis Aligned Integrand Spectra

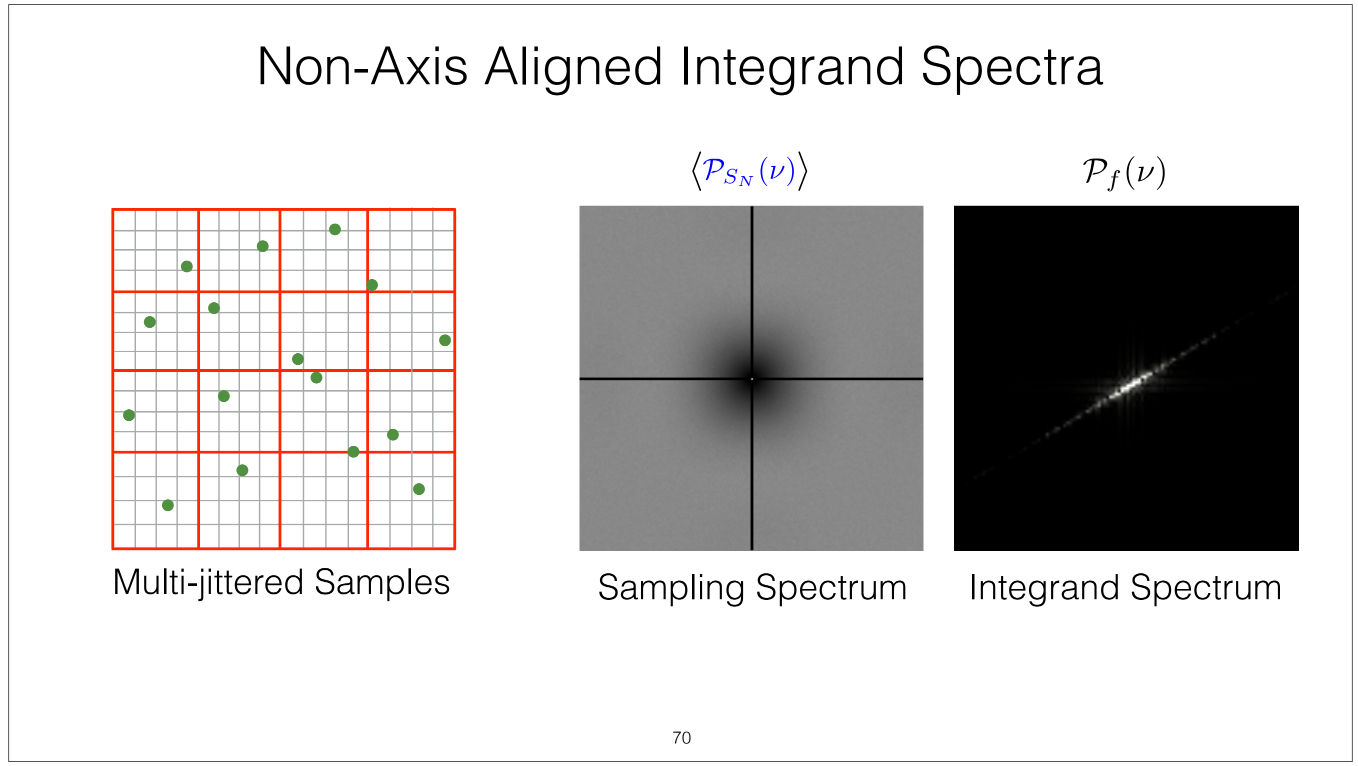


Multi-jittered Samples

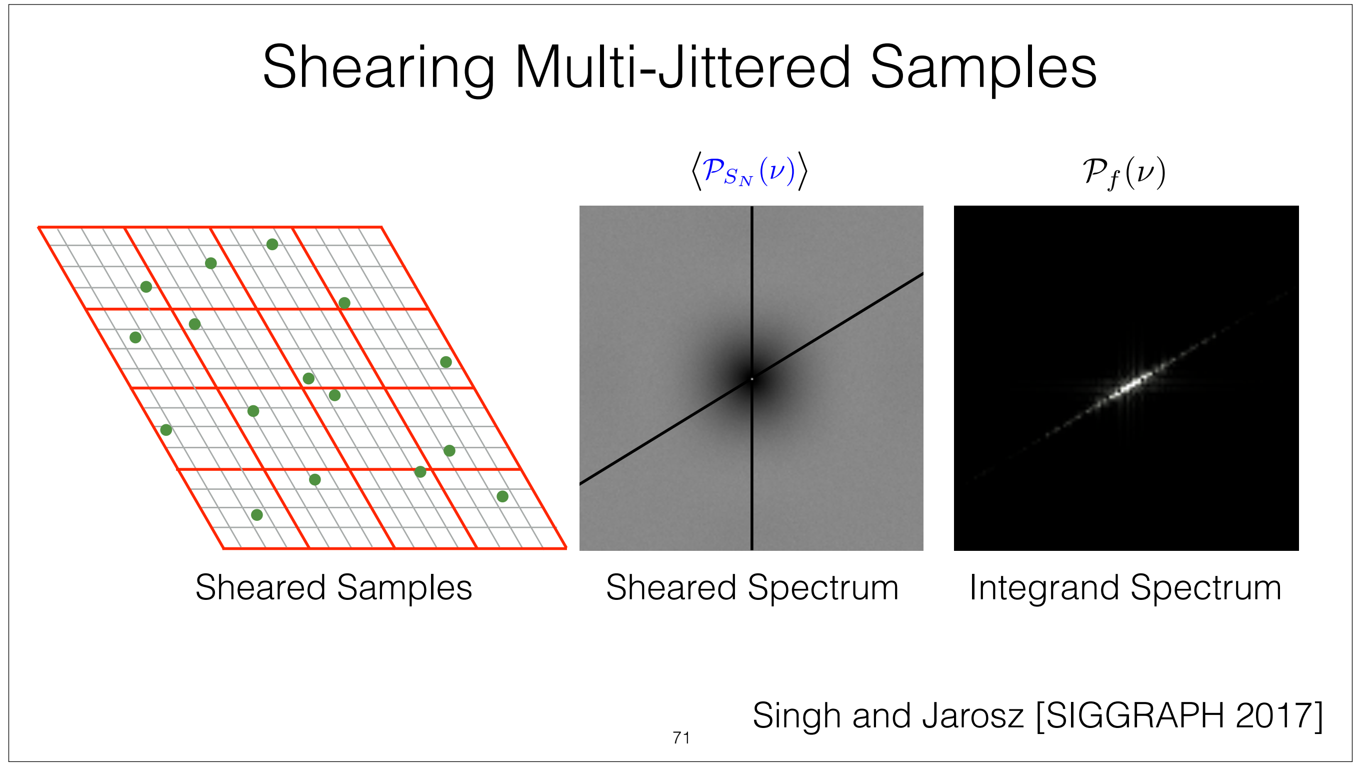
$\mathcal{P}_f(\nu)$

Integrand Spectrum

…with multi-jittered samples which has [CLICK] the following power spectrum, we won’t be able to benefit from these hairline anisotropic structures since they are axis-aligned. To solve this issue, we propose to shear…



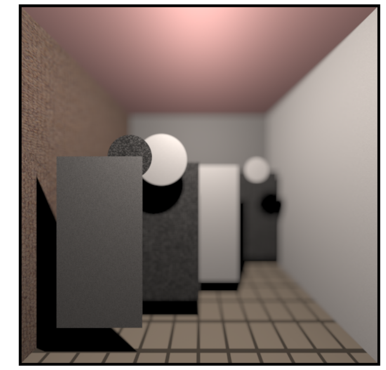
…with multi-jittered samples which has [CLICK] the following power spectrum, we won't be able to benefit from these hairline anisotropic structures since they are axis-aligned. To solve this issue, we propose to shear…



…the samples in such a way that we can align the sampling spectrum with that of the integrand spectrum. Now, to show the improvements due to shearing, lets look at a rendering example.

Variance Heatmap

Uncorrelated Multi-jittered

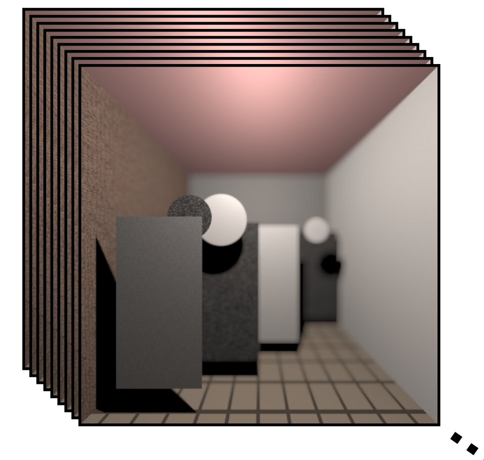


We visualize the variance heatmap; For this, we render the same cornellbox scene [CLICK] multiple times generating 100 images with uncorrelated-multijittered samples followed by computing the variance of each pixel over these 100 images. The resulting variance heatmap [CLICK] is shown as a gray scale image where the brighter pixels means high variance. After shearing the samples in the XU and YV subspaces…

Variance Heatmap

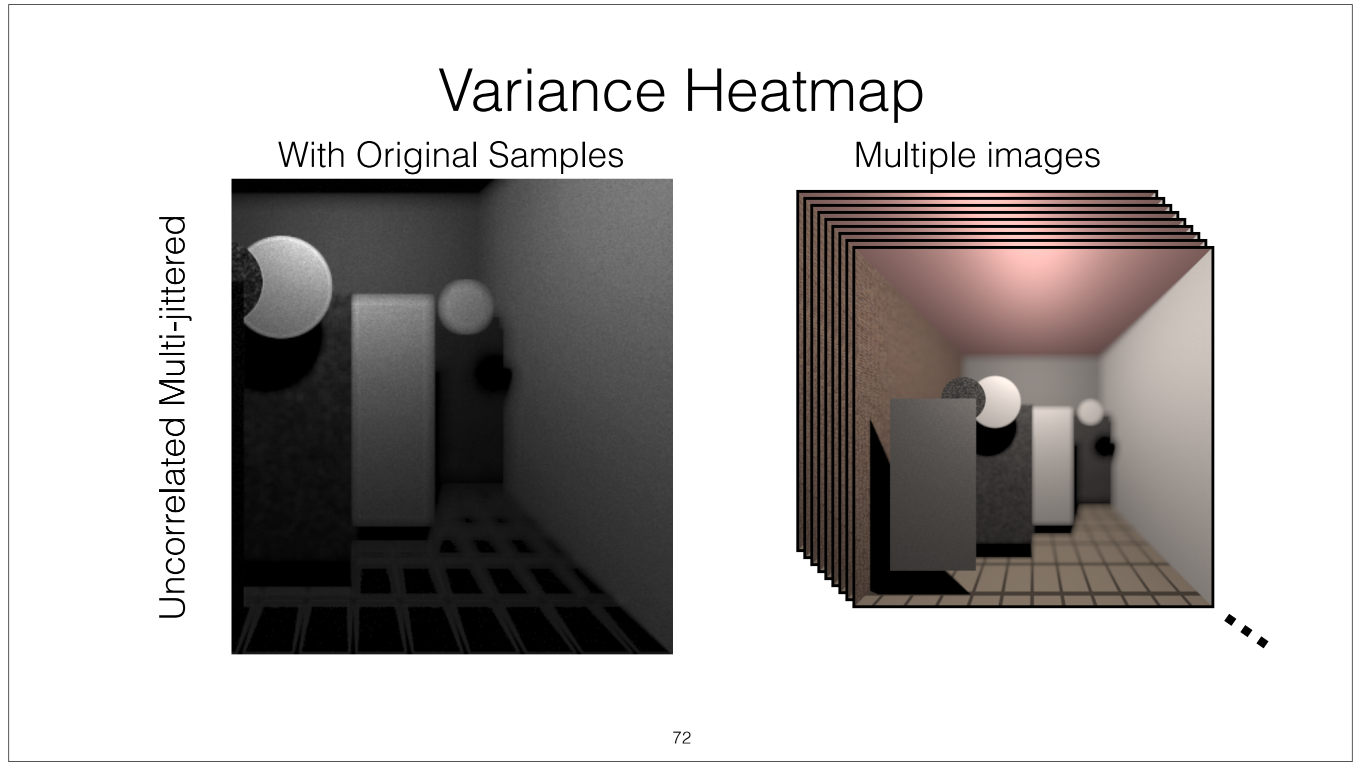
Multiple images

Uncorrelated Multi-jittered



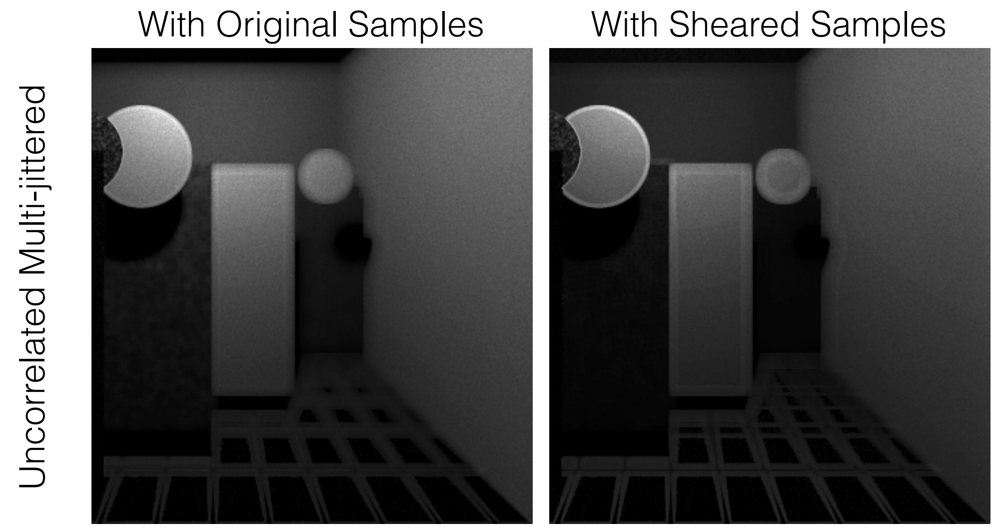
72

We visualize the variance heatmap; For this, we render the same cornellbox scene [CLICK] multiple times generating 100 images with uncorrelated-multijittered samples followed by computing the variance of each pixel over these 100 images. The resulting variance heatmap [CLICK] is shown as a gray scale image where the brighter pixels means high variance. After shearing the samples in the XU and YV subspaces…



We visualize the variance heatmap; For this, we render the same cornellbox scene [CLICK] multiple times generating 100 images with uncorrelated-multijittered samples followed by computing the variance of each pixel over these 100 images. The resulting variance heatmap [CLICK] is shown as a gray scale image where the brighter pixels means high variance. After shearing the samples in the XU and YV subspaces…

Variance Heatmap



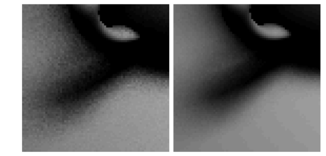
…we observe dramatic variance reduction in pixels with no occluders and modest improvement at pixels with discontinuities. The problem here is that improvements come after a really large N. The major reason for this limited improvement is that existing samplers have only hairline anisotropic structures..

So far...

Blue noise samplers can have better convergence compared to stratified samples

Denser stratification can lead to anisotropic spectra which improves convergence

What properties we desire in a sampler?



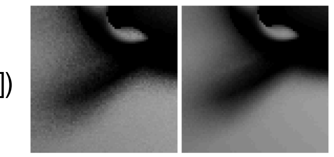
Progressivity

High speed (millions of samples per second)

Extension to dimensions beyond 2D
(Spoke dart throwing, Mitchell [2018])

What properties we desire in a sampler?

Progressivity ✓ (Ahmed et al. [2017], Christensen et al. [2018])

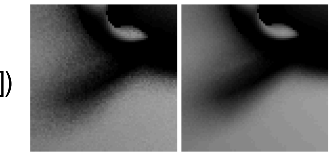


High speed (millions of samples per second)

Extension to dimensions beyond 2D
(Spoke dart throwing, Mitchell [2018])

What properties we desire in a sampler?

Progressivity ✓ (Ahmed et al. [2017], Christensen et al. [2018])

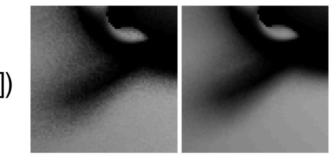


High speed (millions of samples per second) ✓

Extension to dimensions beyond 2D
(Spoke dart throwing, Mitchell [2018])

What properties we desire in a sampler?

Progressivity ✓ (Ahmed et al. [2017], Christensen et al. [2018])



non-progressive progressive

High speed (millions of samples per second) ✓

Extension to dimensions beyond 2D ✗
(Spoke dart throwing, Mitchell [2018])

Low-Discrepancy Sampling

Deterministic sets of points specially crafted to be evenly distributed (have low discrepancy).

Entire field of study called Quasi-Monte Carlo (QMC)

The Van der Corput Sequence

- Radical Inverse Φ_b in base 2
- Subsequent points "fall into biggest holes"

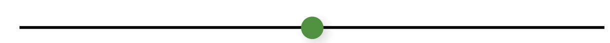
k	Base 2	Φ_b

best discrepancy for infinite sequence

The Van der Corput Sequence

- Radical Inverse Φ_b in base 2
- Subsequent points “fall into biggest holes”

k	Base 2	Φ_b
1	1	.1 = 1/2



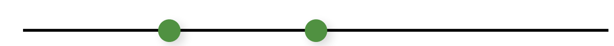
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points “fall into biggest holes”

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4



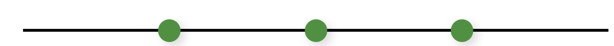
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points “fall into biggest holes”

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4



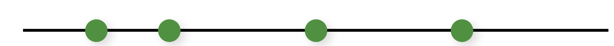
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points "fall into biggest holes"

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4
4	100	.001 = 1/8



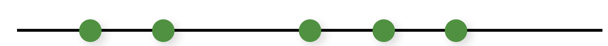
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points “fall into biggest holes”

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4
4	100	.001 = 1/8
5	101	.101 = 5/8



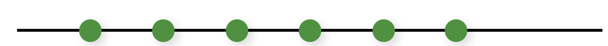
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points "fall into biggest holes"

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4
4	100	.001 = 1/8
5	101	.101 = 5/8
6	110	.011 = 3/8



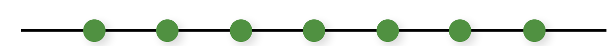
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points "fall into biggest holes"

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4
4	100	.001 = 1/8
5	101	.101 = 5/8
6	110	.011 = 3/8
7	111	.111 = 7/8



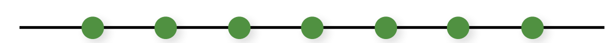
best discrepancy for infinite sequence

The Van der Corput Sequence

Radical Inverse Φ_b in base 2

Subsequent points “fall into biggest holes”

k	Base 2	Φ_b
1	1	.1 = 1/2
2	10	.01 = 1/4
3	11	.11 = 3/4
4	100	.001 = 1/8
5	101	.101 = 5/8
6	110	.011 = 3/8
7	111	.111 = 7/8
...		



best discrepancy for infinite sequence

Halton and Hammersley Points

Halton: Radical inverse with different base for each dimension:

$$\vec{x}_k = (\Phi_2(k), \Phi_3(k), \Phi_5(k), \ \ldots \ , \Phi_{p_n}(k))$$

Halton and Hammersley Points

Halton: Radical inverse with different base for each dimension:

$$\vec{x}_k = (\Phi_2(k), \Phi_3(k), \Phi_5(k), \ \ldots\ , \Phi_{p_n}(k))$$

- The bases should all be relatively prime.

Halton and Hammersley Points

Halton: Radical inverse with different base for each dimension:

$$\vec{x}_k = (\Phi_2(k), \Phi_3(k), \Phi_5(k), \ldots, \Phi_{p_n}(k))$$

- The bases should all be relatively prime.
- Incremental/progressive generation of samples

Halton and Hammersley Points

Halton: Radical inverse with different base for each dimension:

$$\vec{x}_k = (\Phi_2(k), \Phi_3(k), \Phi_5(k), \ldots, \Phi_{p_n}(k))$$

- The bases should all be relatively prime.
- Incremental/progressive generation of samples

Hammersley: Same as Halton, but first dimension is k/N:

$$\vec{x}_k = ({}^k/_N, \Phi_2(k), \Phi_3(k), \Phi_5(k), \ldots, \Phi_{p_n}(k))$$

Halton and Hammersley Points

Halton: Radical inverse with different base for each dimension:

$$\vec{x}_k = (\Phi_2(k), \Phi_3(k), \Phi_5(k), \ldots, \Phi_{p_n}(k))$$

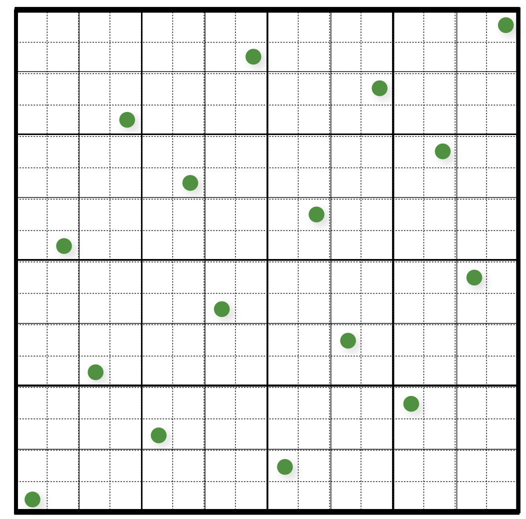
- The bases should all be relatively prime.
- Incremental/progressive generation of samples

Hammersley: Same as Halton, but first dimension is k/N:

$$\vec{x}_k = ({}^k/_N, \Phi_2(k), \Phi_3(k), \Phi_5(k), \ldots, \Phi_{p_n}(k))$$

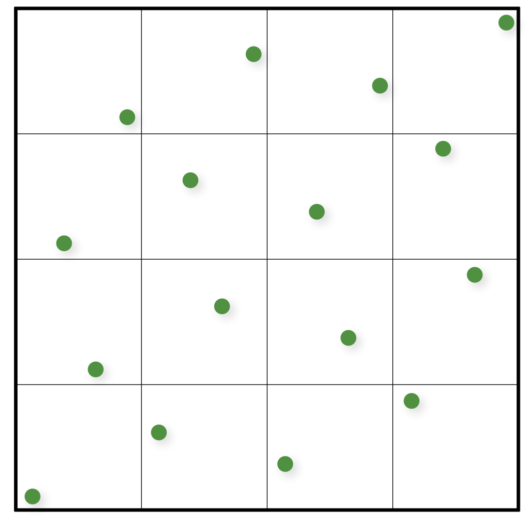
- Not incremental, need to know sample count, N, in advance

The Hammersley Sequence



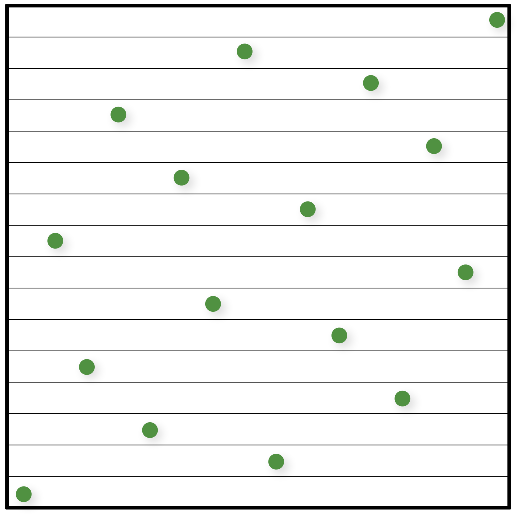
1 sample in each “elementary interval”

The Hammersley Sequence



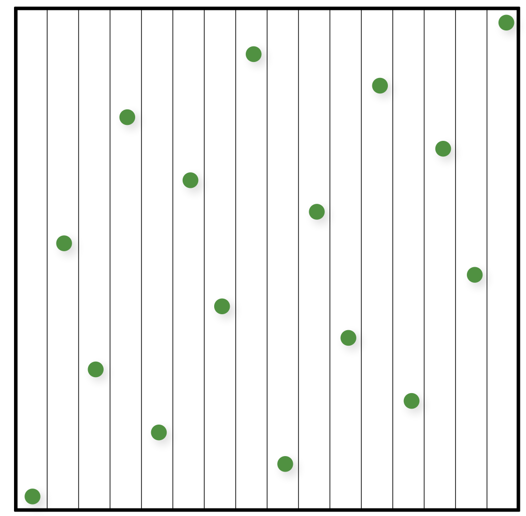
1 sample in each “elementary interval”

The Hammersley Sequence



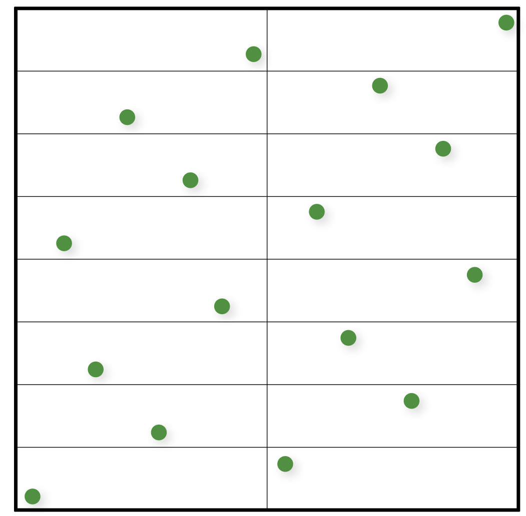
1 sample in each “elementary interval”

The Hammersley Sequence



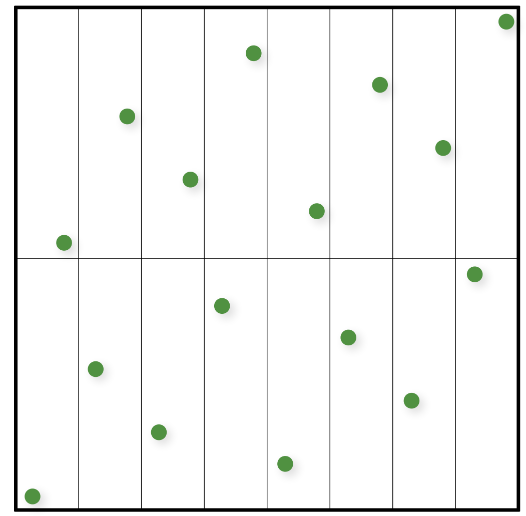
1 sample in each "elementary interval"

The Hammersley Sequence



1 sample in each “elementary interval”

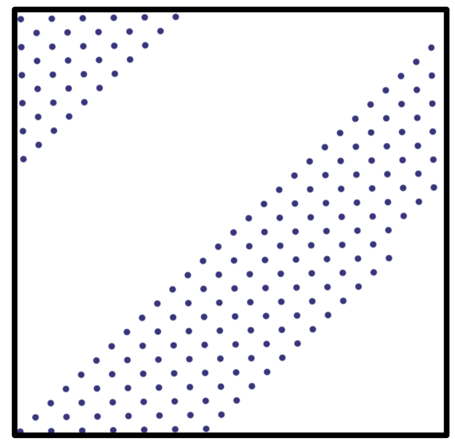
The Hammersley Sequence



1 sample in each “elementary interval”

Why do we need to scramble?

Halton Projection (29, 31)

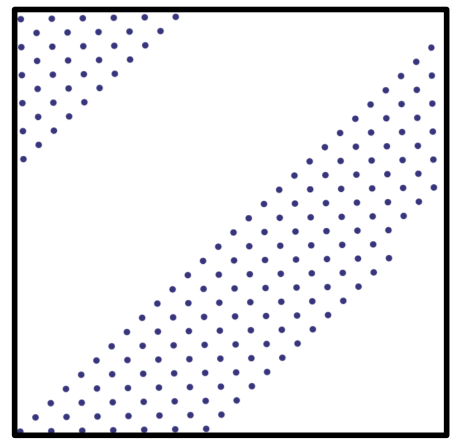


85

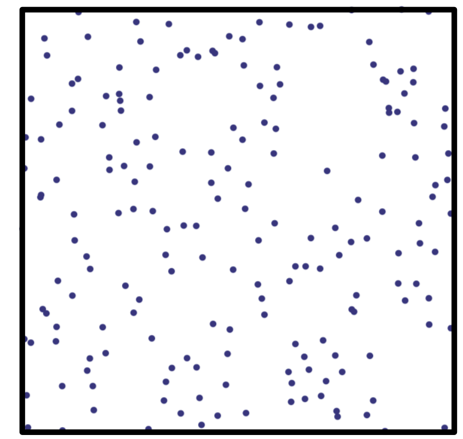
These deterministic samplers can have voids in the domain when the sample count is not exactly dyadic.
[CLICK] We can scramble these samples to fill the domain. Let's see some rendering results.

Why do we need to scramble?

Halton Projection (29, 31)

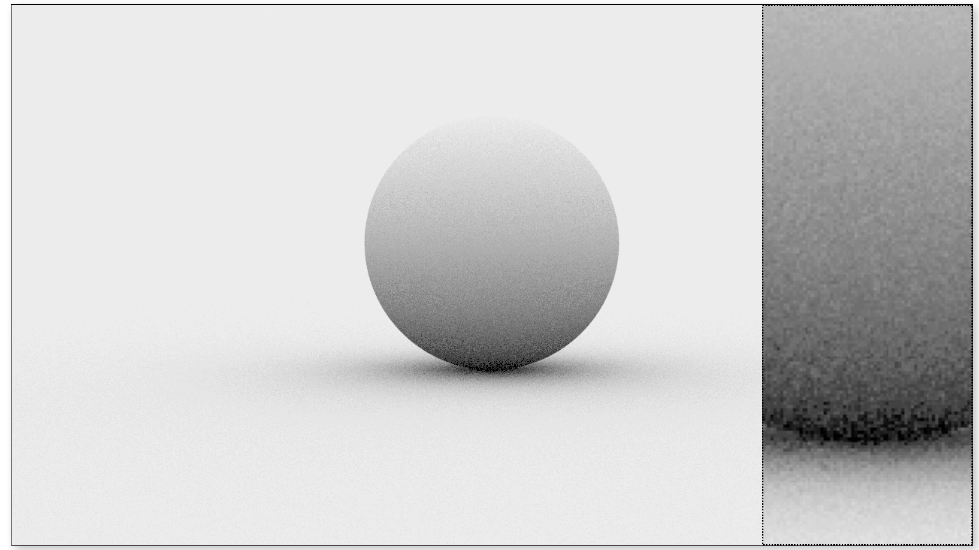


Scrambled Halton Projection (29, 31)



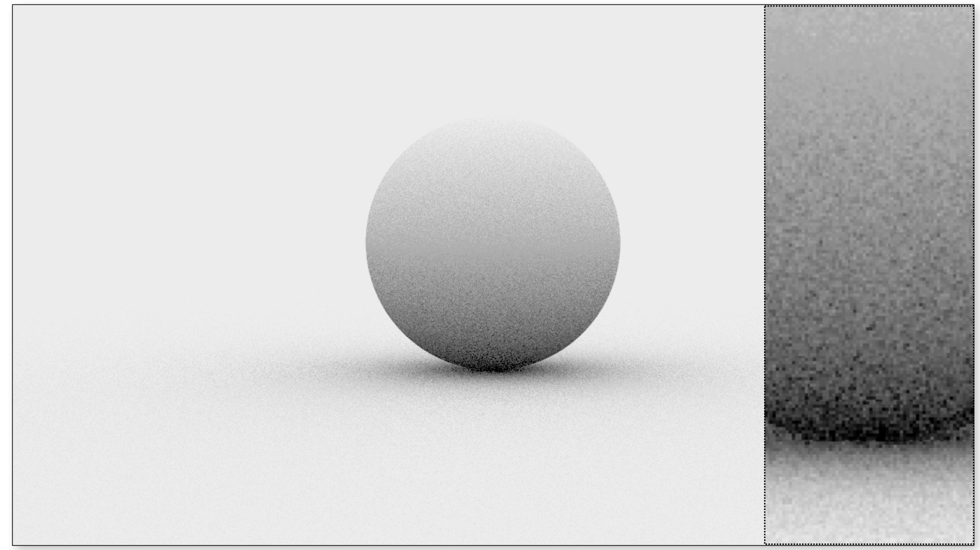
These deterministic samplers can have voids in the domain when the sample count is not exactly dyadic.
[CLICK] We can scramble these samples to fill the domain. Let's see some rendering results.

Scrambled Low-Discrepancy Sampling

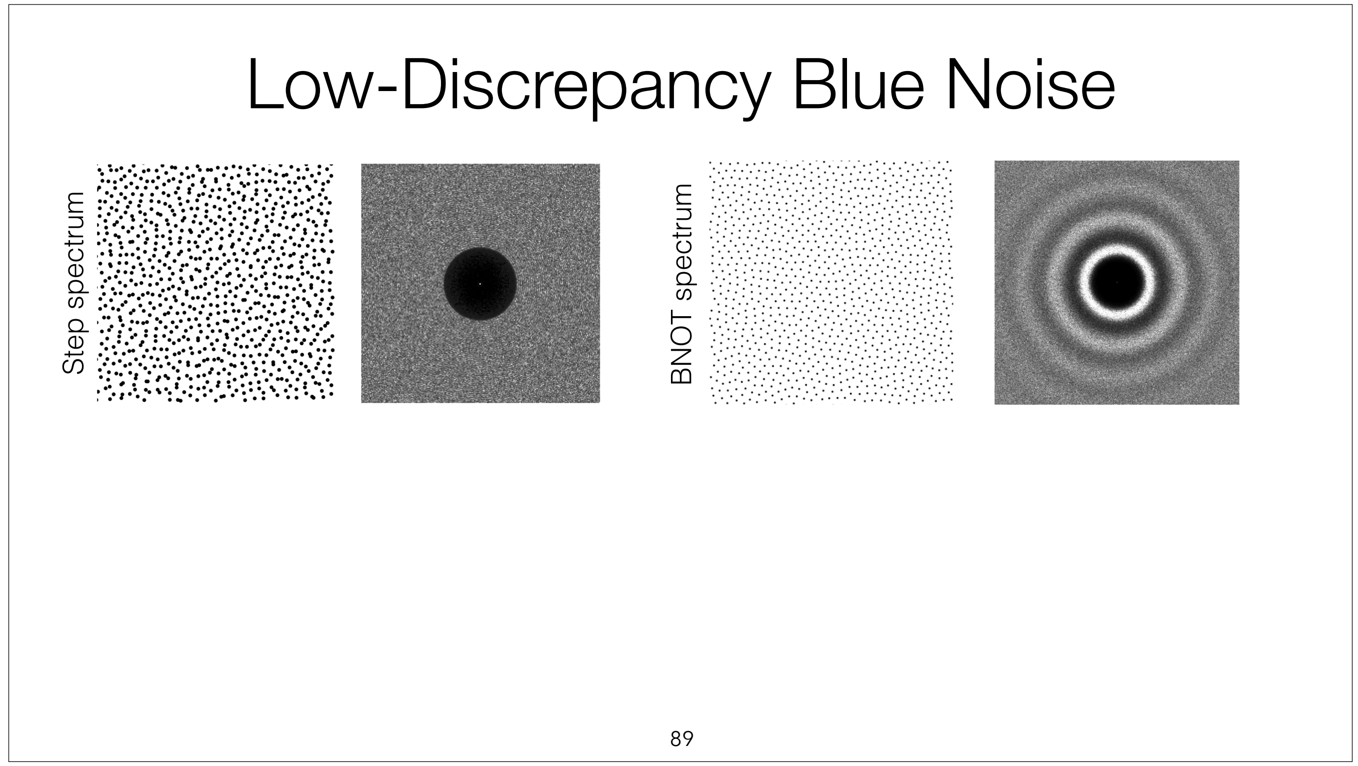


Low discrepancy samplers show less noise compared to randomly jittered samples.

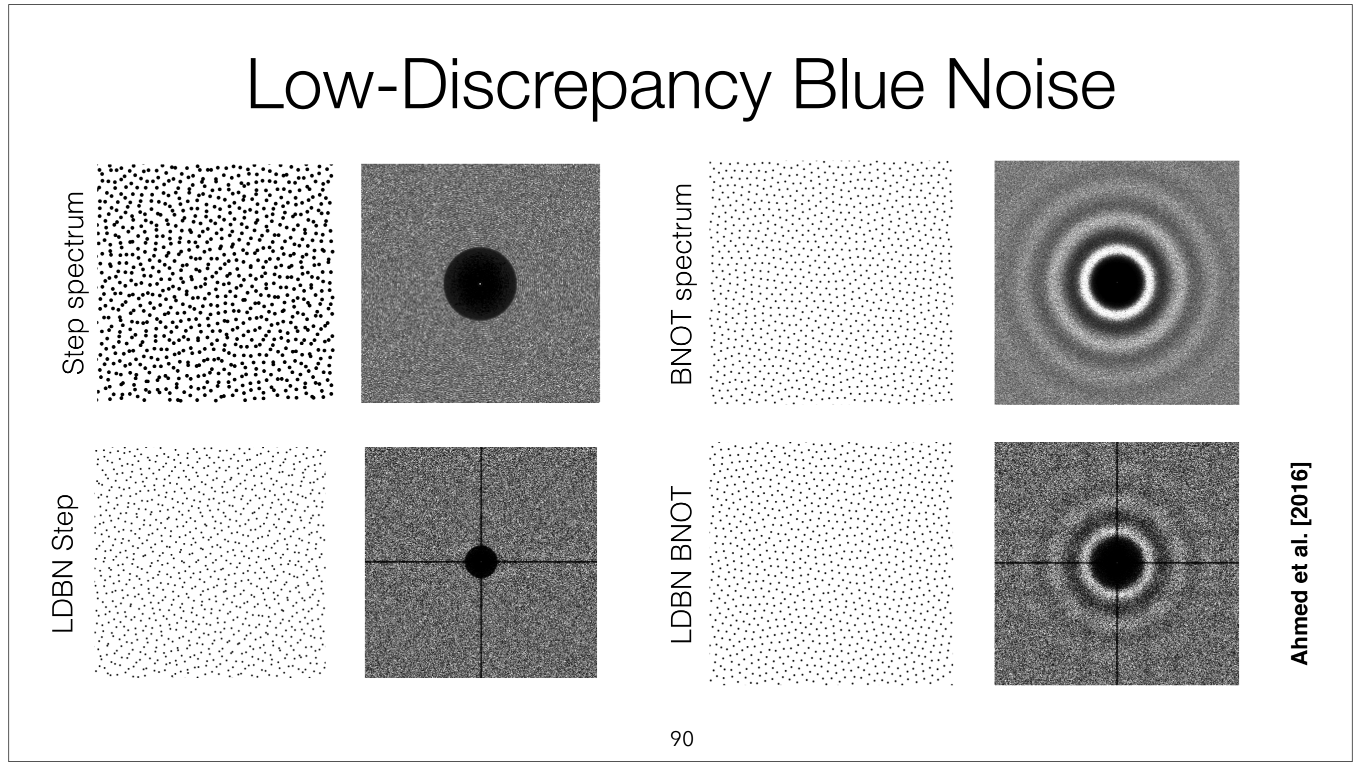
Monte Carlo (16 jittered samples)



Can we combine blue noise properties with low discrepancy?

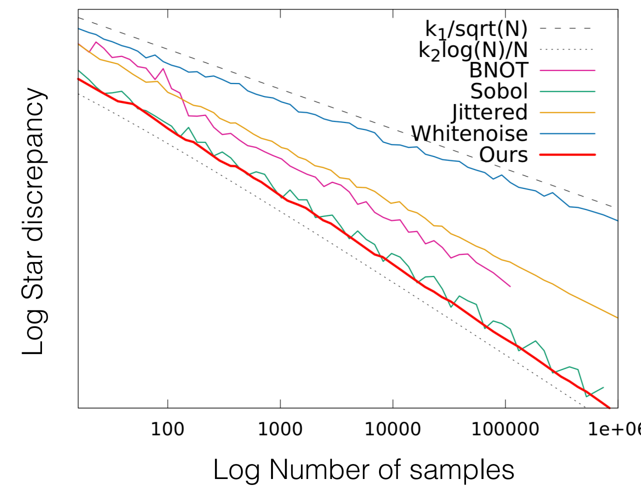


Recently, couple of papers were published that incorporate certain stratification properties to different blue noise targets resulting in lowering their discrepancy properties.



Recently, couple of papers were published that incorporate certain stratification properties to different blue noise targets resulting in lowering their discrepancy properties.

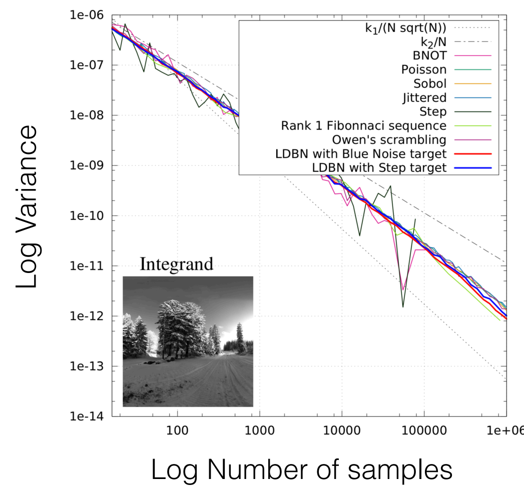
Low-Discrepancy Blue Noise

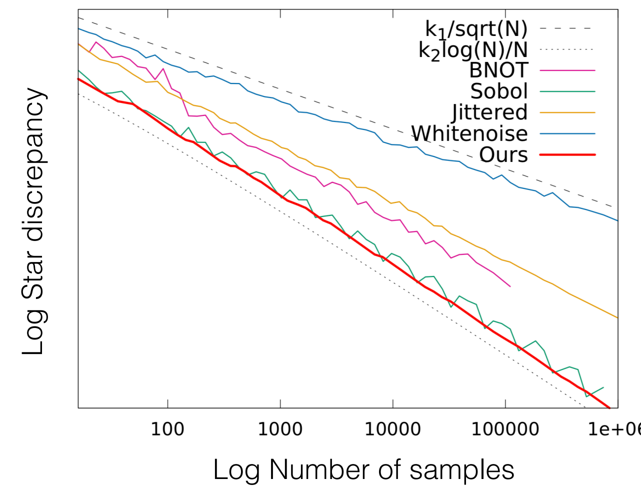


Ahmed et al. [2016]

The plots are shown where the discrepancy is comparable to Sobol.

Low-Discrepancy Blue Noise

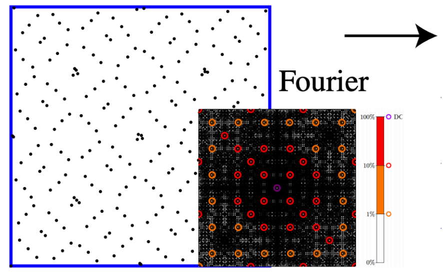




Ahmed et al. [2016]

The corresponding variance convergence is also comparable.

Low-Discrepancy Blue Noise 2D-Projections

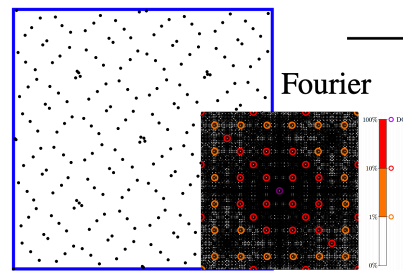


Sobol

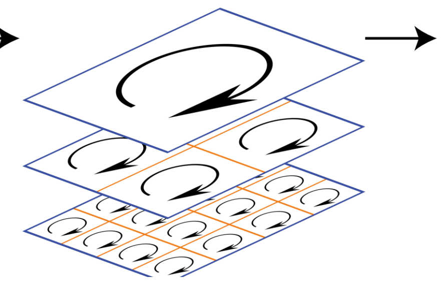
Perrier et al. [2018]

Perrier and colleagues took a step further and developed a smart scrambling strategy that introduces blue noise properties directly into the well known sobol sequences.

Low-Discrepancy Blue Noise 2D-Projections



Sobol

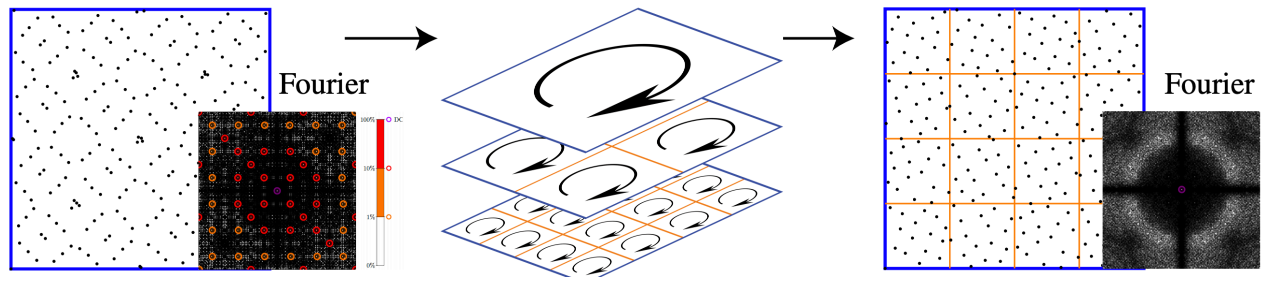


Special scrambling

Perrier et al. [2018]

Perrier and colleagues took a step further and developed a smart scrambling strategy that introduces blue noise properties directly into the well known sobol sequences.

Low-Discrepancy Blue Noise 2D-Projections

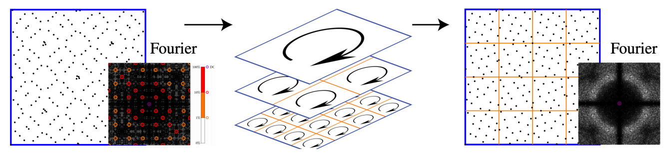


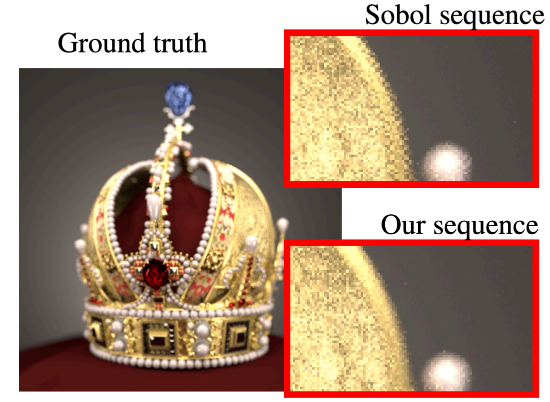
Sobol | Special scrambling | Blue noise characteristics

Perrier et al. [2018]

Perrier and colleagues took a step further and developed a smart scrambling strategy that introduces blue noise properties directly into the well known sobol sequences.

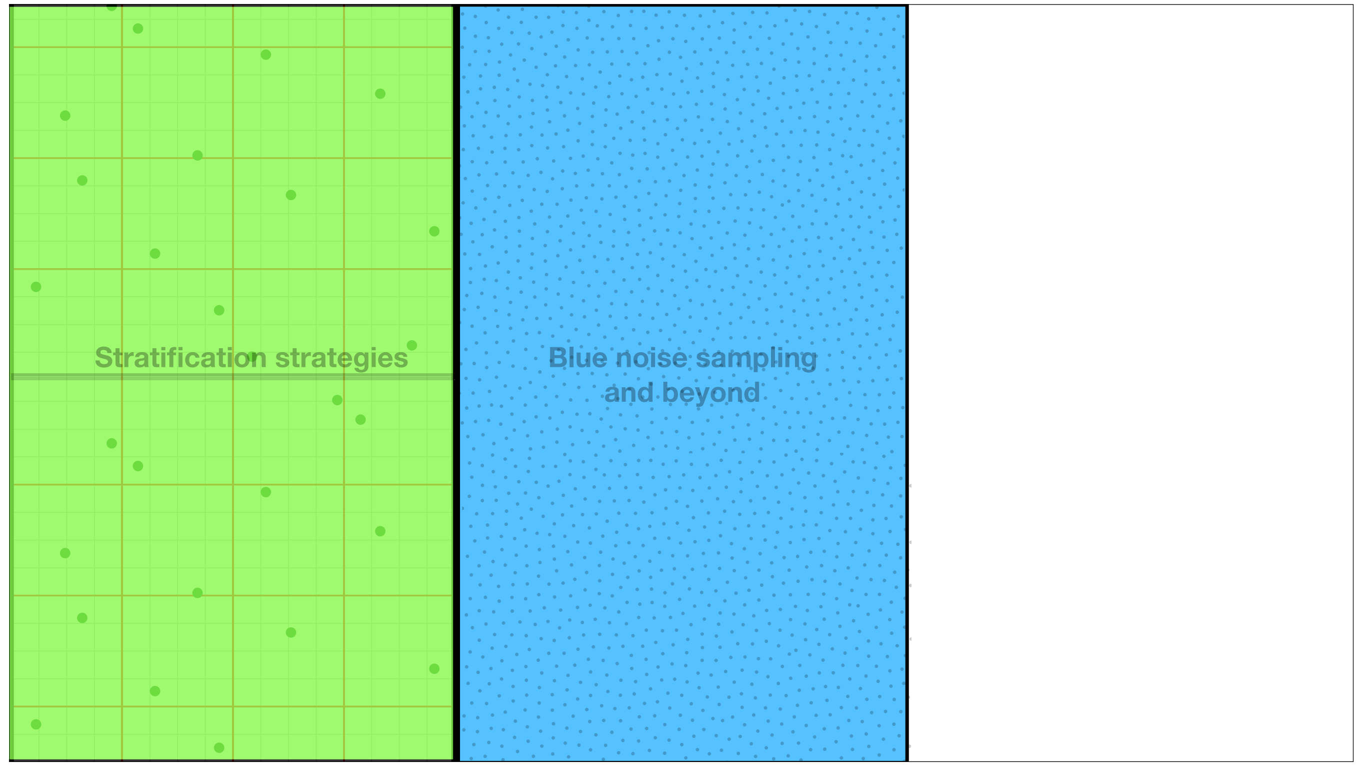
Low-Discrepancy Blue Noise 2D Sobol Projections



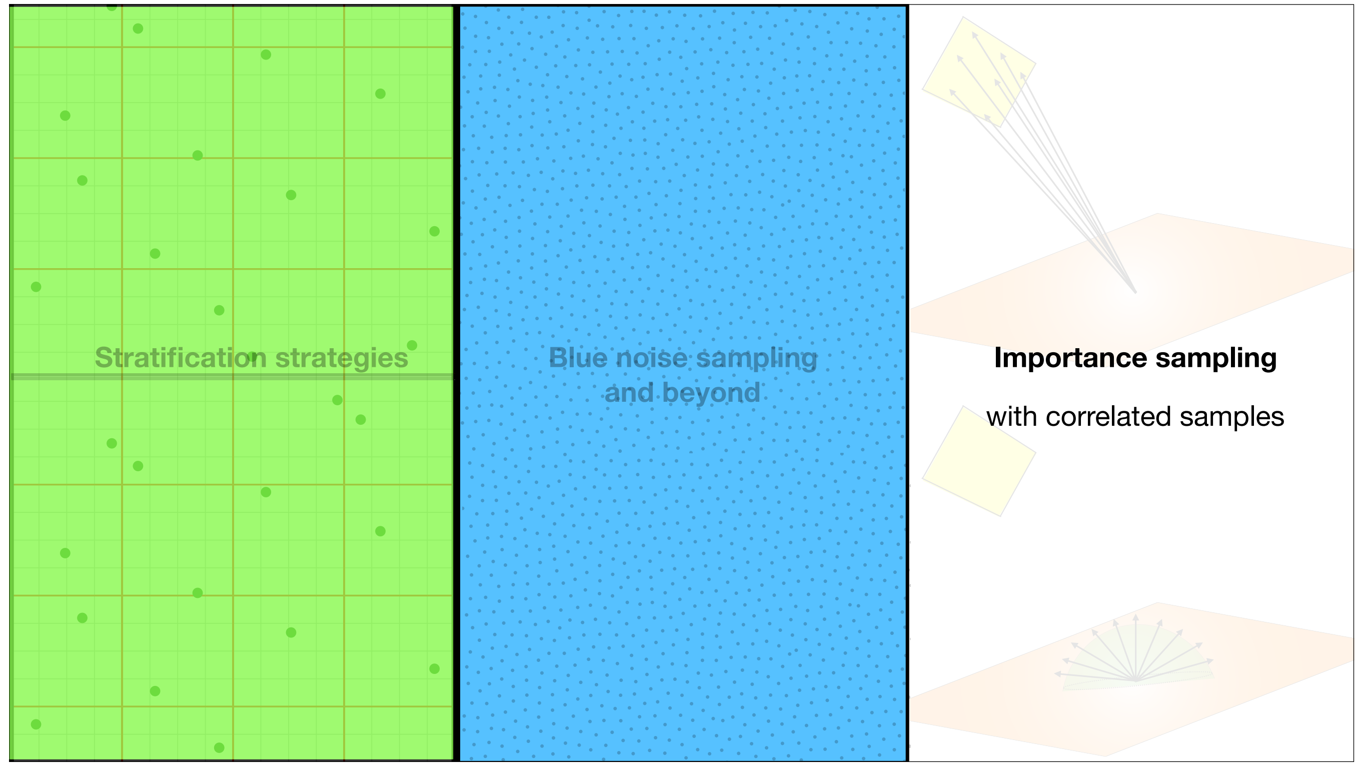


Perrier et al. [2018]

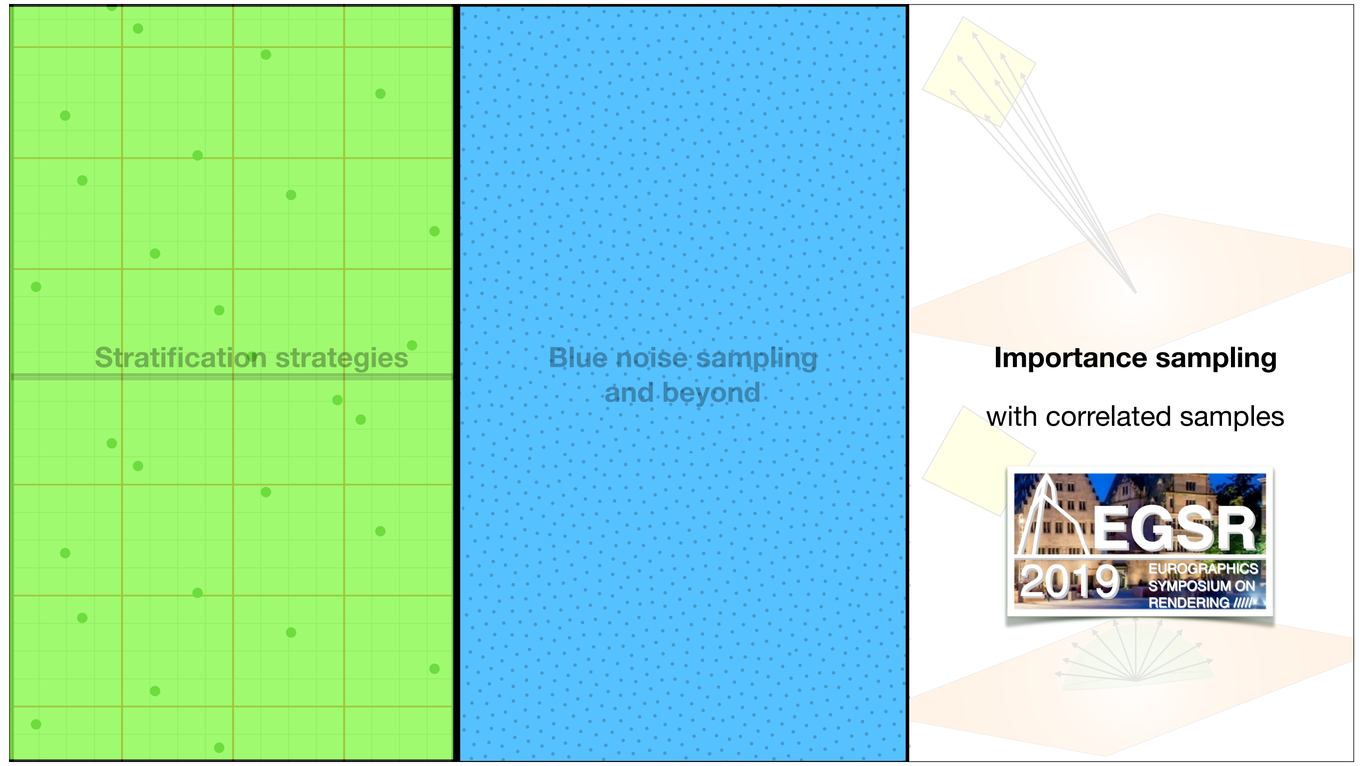
This also shows improvements in rendering quality near the edges of the scene.



In the end, we now briefly touch upon correlated importance sampling which is recently published at CGF and we are hoping to present it at EGSR this year.

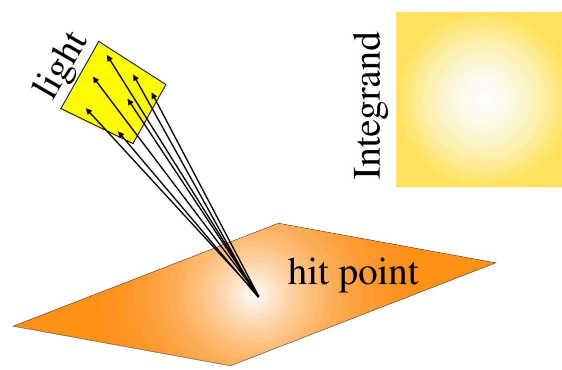


In the end, we now briefly touch upon correlated importance sampling which is recently published at CGF and we are hoping to present it at EGSR this year.

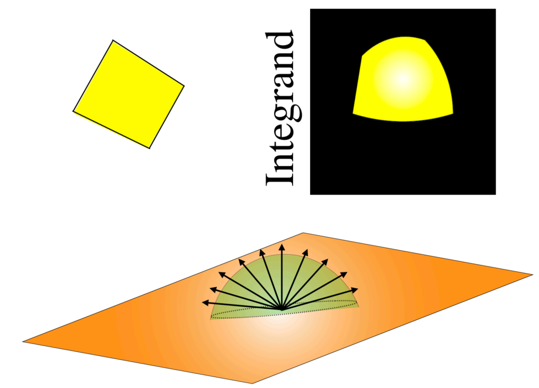


In the end, we now briefly touch upon correlated importance sampling which is recently published at CGF and we are hoping to present it at EGSR this year.

Light IS vs BSDF IS



Light Importance Sampling



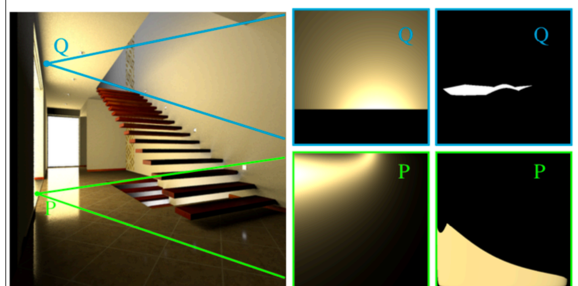
BSDF Importance Sampling

Singh et al. [2019]

Two common importance sampling (IS):
light IS: generate samples on the light source and generate shadow rays to the hit point. This results in a smooth looking underlying integrand at that hit point.
BSDF IS: When the visible hemisphere is sampled, the samples see the light source boundary as a discontinuity, making the underlying integrand C_0 discontinuous.

Scene illuminated by area direct lighting

Light IS BSDF IS



Reference

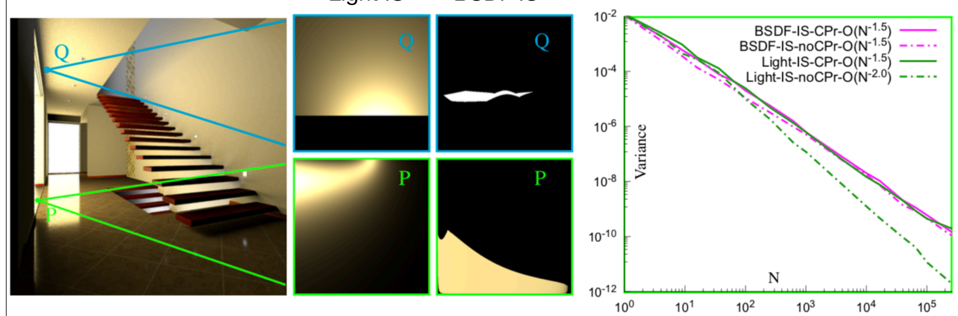
Underlying pixel functions

Singh et al. [2019]

This directly reflected into the corresponding convergence rates. We focus on two pixels: Pixel P (the hit point is directly visible from the light source) and Pixel Q (partially occluded from the light source).

Unoccluded pixels' convergence benefit from Light IS

Light IS BSDF IS

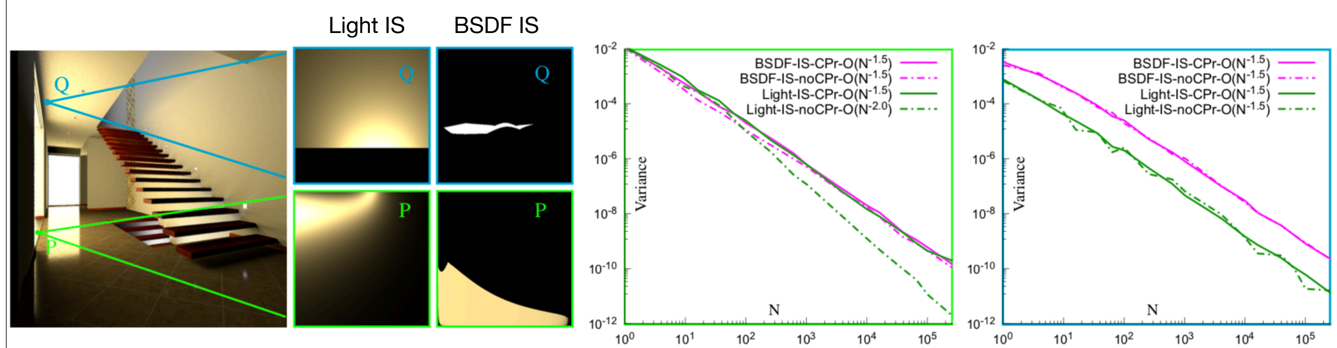


Reference | Underlying pixel functions | Pixel P

Singh et al. [2019]

light IS for unoccluded hit points (Pixel P with no occluders) shows good convergence rate compared to partially occluded hit points. BSDF IS however does not improve any convergence behavior irrespective of whether the hit point is occluded or not.

Occluded pixels (no improvement in convergence)



Reference

Underlying
pixel functions

Pixel P

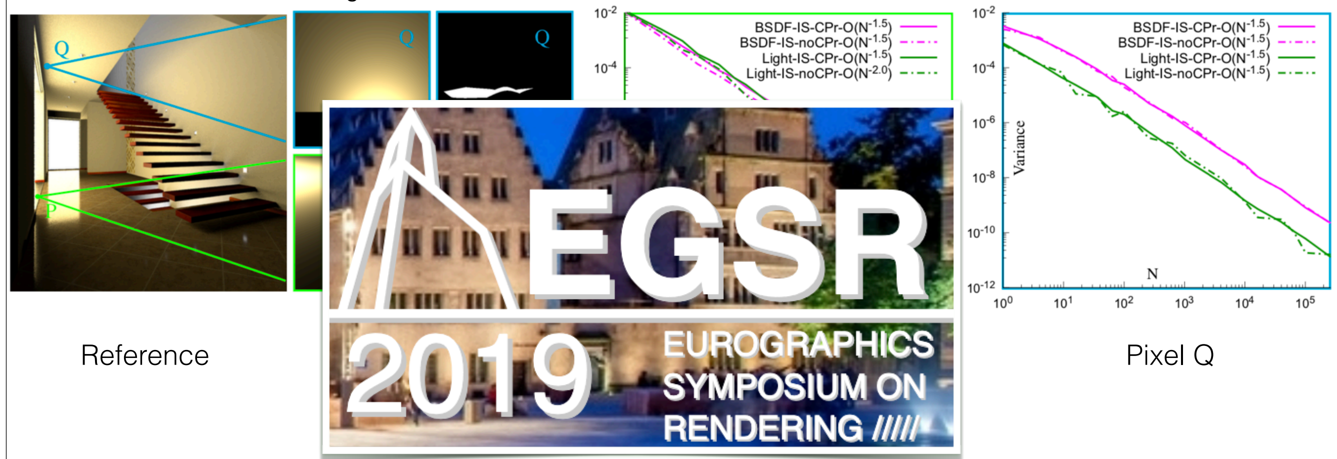
Pixel Q

Singh et al. [2019]

More at EGSR ;)

Occluded pixels (no improvement in convergence)

Light IS BSDF IS



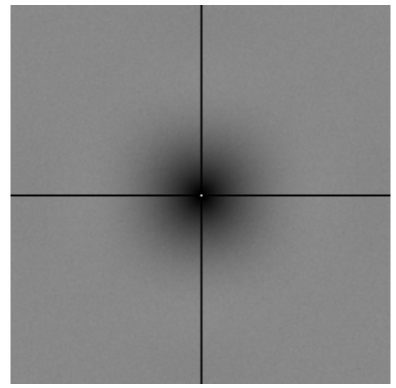
Reference Pixel Q

Singh et al. [2019]

More at EGSR ;)

Futuristic sampling target spectrum

Multi-jittered



Future design

Singh and Jarosz [2017]

Future directions:
Might be interesting to generate samples with correlations that can have wider anisotropic structures.

Future research directions

Direct link between spatial and Fourier statistics needs further investigation

Progressive samplers in higher dimensions

Adapting sample correlations w.r.t. the underlying integrand in high dimensions

Acknowledgments

Some slides borrowed from Wojciech Jarosz and Kartic Subr

All the anonymous reviewers who helped shape this survey paper into its final form

Acknowledgments

Some slides borrowed from Wojciech Jarosz and Kartic Subr

All the anonymous reviewers who helped shape this survey paper into its final form