

Sampling Techniques

Prof. Dr. Stefan Gumhold
Chair of Computer Graphics and
Visualization, TU Dresden



Overview

Sampling Techniques

- ◆ Generating Random Samplings
- ◆ Rejection Sampling
- ◆ Transformation Sampling
- ◆ Sampling Discrete Distribution
- ◆ Multiple Importance Sampling

Generating Random Samplings

Random Number Creation



- ◆ The basis for sampling according to a probability distributions form random number generators
- ◆ In C++ there is a wide support for this through the STL-header `<random>` with support for
 - ◆ Generators: Objects that generate uniformly distributed numbers.
 - ◆ Distributions: Objects that transform sequences of numbers generated by a generator into sequences of numbers that follow a specific random variable distribution, such as uniform, Normal or Binomial.
- ◆ One can use the function `std::random_shuffle()` from the `<algorithm>` header to bring a vector into random order (compare n-rooks sampling)

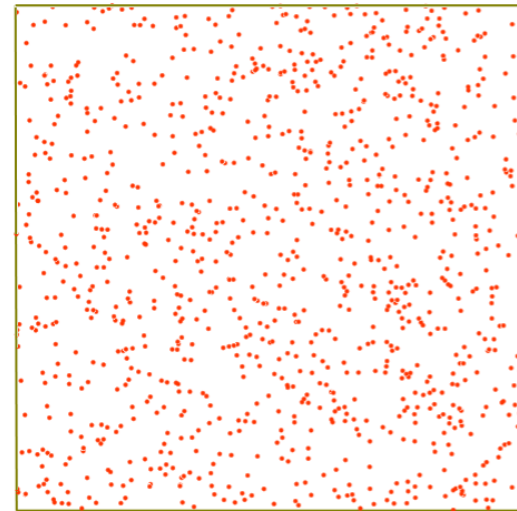
Generating Random Samplings

Example Sampling Code



```
#include <random>
#include <algorithm>

/// generate uniformly distributed random points in unit square
void random_sample(std::vector<P2D>& P, int n) {
    std::mt19937 rand_gen;
    std::uniform_real_distribution<double> uni_dist(0.0, 1.0);
    for (unsigned i=0; i<n; ++i) {
        double x = uni_dist(rand_gen);
        double y = uni_dist(rand_gen);
        P.push_back(P2D(x,y));
    }
}
```



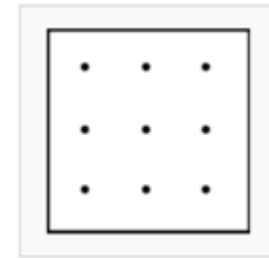
Output for n=1000

Generating Random Samplings

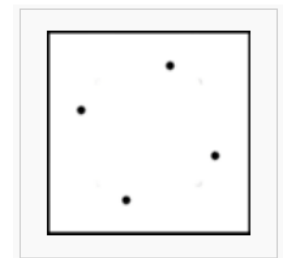
Further Sampling Strategies



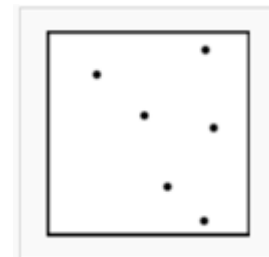
- samplings on **regular** and **rotated** regular **grids** are often used in real-time graphics.
- due to dependence between coordinates the problem of **dimensional core** remains
- Extending **uniform sampling** with rejection test, whether new sample is closer than ϵ to all other samples, yields **Possion disk** sampling and **avoids sample clustering**
- **Stratified grid** sampling extends regular grid sampling with uniform sampling on each grid cell
- **N-rooks** sampling performs stratified grid sampling in 1D per dimension and combines coordinate samples after random shuffle. In high dimensions n-rooks **covers each dimension much better**.



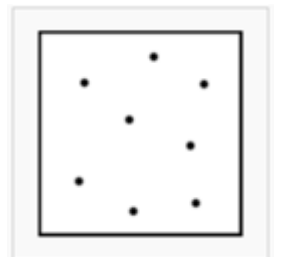
regular grid



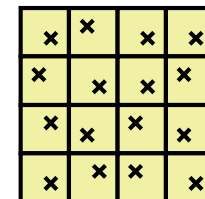
rotated grid



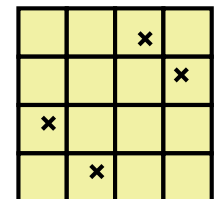
uniform
random



Poisson
disk



stratified
grid



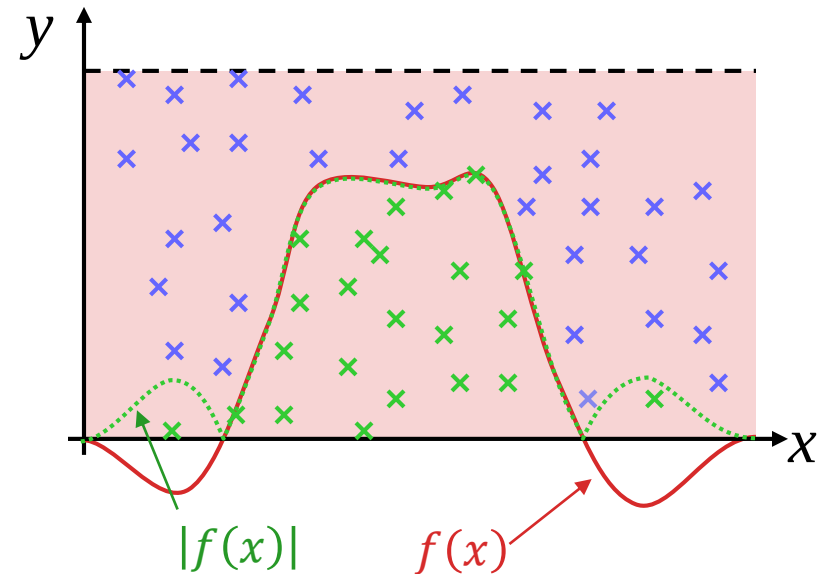
stratified
n-rooks

Rejection Sampling

- Given a function $f(x)$ over the domain Ω and an **upper bound** $f_{\max}$ with $\forall x \in \Omega: |f(x)| < f_{\max}$.
- Goal: sample from the distribution

$$p_f(x) = |f(x)| / \int_{\Omega} |f(x)| dx$$

- Idea: sample $\mathbf{z}_i = (x_i, y_i)$ uniformly from the volume $V = \Omega \times [0, f_{\max}]$ in a space with one dimension more and reject all samples above the graph of $|f(x)|$, i.e. if $y_i > |f(x_i)|$.
- This is only efficient if V is not significantly larger than $\int_{\Omega} |f(x)| dx$.



```

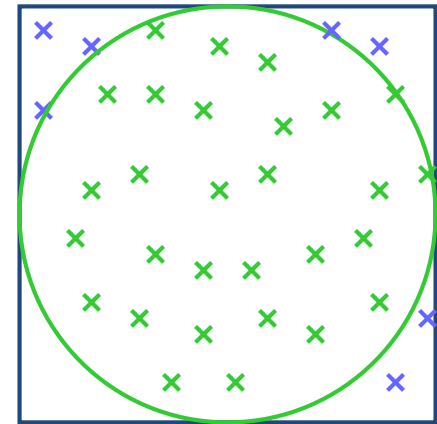
/// generate sample by rejection sampling
rejection_sample(x_min, x_max, f(x), f_max)
{
    do {
        x = uniform_rand(x_min, x_max);
        y = uniform_rand(0, f_max);
    } while (y > abs(f(x)));
    return x;
}

```

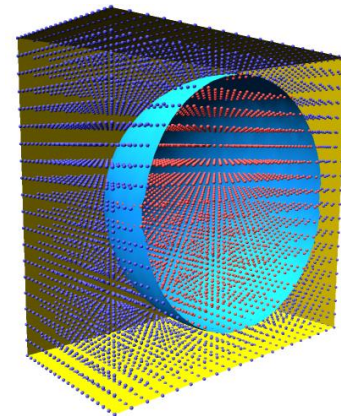
Rejection Sampling Shapes



- Rejection sampling can be used to sample uniformly inside a given shape (for example circle or sphere):
 - compute bounding box of shape
 - sample bounding box uniformly or with other sampling strategy
 - reject samples that are not inside
- Rejection sampling can be trivially combined with other sampling strategies
- There is no access to the probability density in rejection sampling



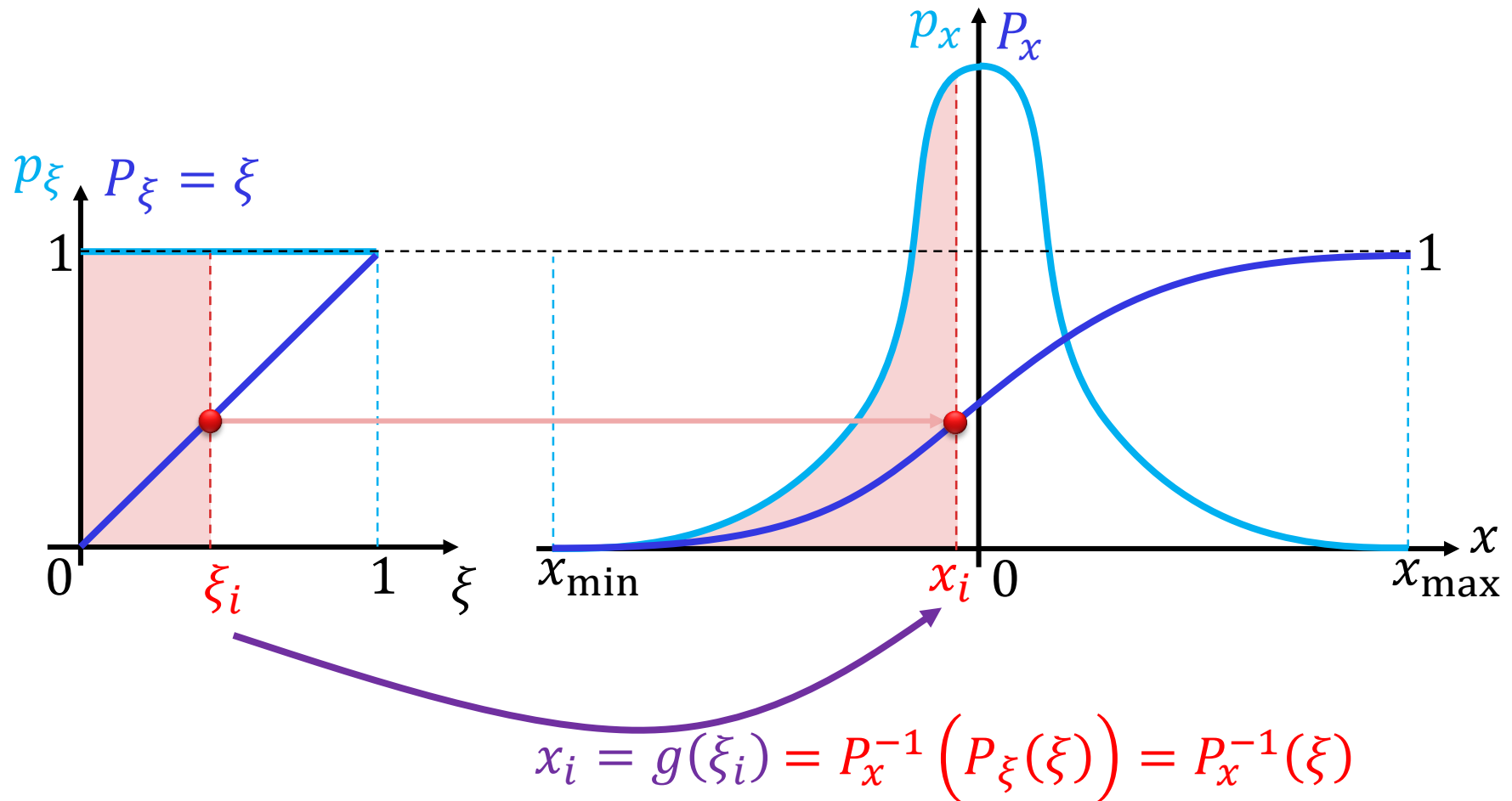
rejection sampling of circle



same works for sphere

Transformation Sampling

1D – Motivation





- ◆ Goal: we want to sample x_i according to a given distribution $p_x(x)$ defined over the interval $[x_{\min}, x_{\max}]$
- ◆ Idea: sample a ξ_i uniformly over $[0,1]$ and compute $x_i = g(\xi_i)$ with a suitable function $g(\xi)$.
- ◆ Here ξ is distributed according to $p_\xi(\xi) \equiv 1$ over $[0,1]$ and its cumulative probability function $P_\xi(\xi)$ is the identity:

$$P_\xi(\xi) = \int_0^\xi p_\xi(\xi') d\xi' = \xi$$

- ◆ For corresponding pairs of x and ξ the cum.prob.funcs must have the same value, i.e $\xi = P_\xi(\xi) = P_x(x)$, such that we get $g(\xi)$ from inversion of $P_x(x)$:

$$g(\xi) = P_x^{-1}(\xi)$$

Transformation Sampling

Example 1D



- Given: $p_x(x) = 3x^2/2$ over the interval $[-1,1]$
- First we need to integrate to get the cum.dist.func.

$$\begin{aligned} P_x(x) &= \int_{-1}^x p_x(\tilde{x}) d\tilde{x} = \frac{1}{2} \int_{-1}^x 3\tilde{x}^2 d\tilde{x} \\ &= \frac{1}{2} [\tilde{x}^3]_{-1}^x = (x^3 + 1)/2 \end{aligned}$$

- Next we need the inverse function, which we get from solving the following equation for x :

$$\xi = \frac{(x^3 + 1)}{2}$$

$$x^3 = 2\xi - 1$$

$$x = \sqrt[3]{2\xi - 1}$$

```
#include <cmath> // for cbrt
/// generate sample by transformation
transformation_sample()
{
    ξ = uniform_rand(0,1);
    return cbrt(2*ξ - 1);
}
```

Transformation Sampling

2D – marginalization

- ◆ Goal: we want to sample $\mathbf{x}_i = (x_i, y_i)$ according to a given distribution $p_{\mathbf{x}}(\mathbf{x})$ defined over the domain $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$
- ◆ Idea: sample a $\xi_i = (\xi_i, \eta_i)$ uniformly over $[0,1]^2$ and compute $\mathbf{x}_i = \mathbf{g}(\xi_i)$ with a suitable vector valued function $\mathbf{g}(\xi)$.
- ◆ Again we start with $p_{\xi}(\xi) \equiv 1$ over $[0,1]^2$.
- ◆ In 2D we construct $\mathbf{g}(\xi)$ in two steps.
- ◆ Firstly, we relate x and ξ in the function component $g_x(\xi)$ by marginalization over y :

$$p_x(x) = \int_{y_{\min}}^{y_{\max}} p_{\mathbf{x}}(x, y) dy \quad \text{and} \quad p_{\xi}(\xi) = \int_0^1 p_{\xi}(\xi, \eta) d\eta = 1$$

$$P_x(x) = \int_{x_{\min}}^x p_x(\tilde{x}) d\tilde{x} \quad \text{and} \quad P_{\xi}(\xi) = \int_0^{\xi} p_{\xi}(\tilde{\xi}) d\tilde{\xi} = \xi$$

- ◆ equating and solving for x :

$$P_x(x) = P_{\xi}(\xi) = \xi \quad \text{and} \quad x = P_x^{-1}(\xi)$$

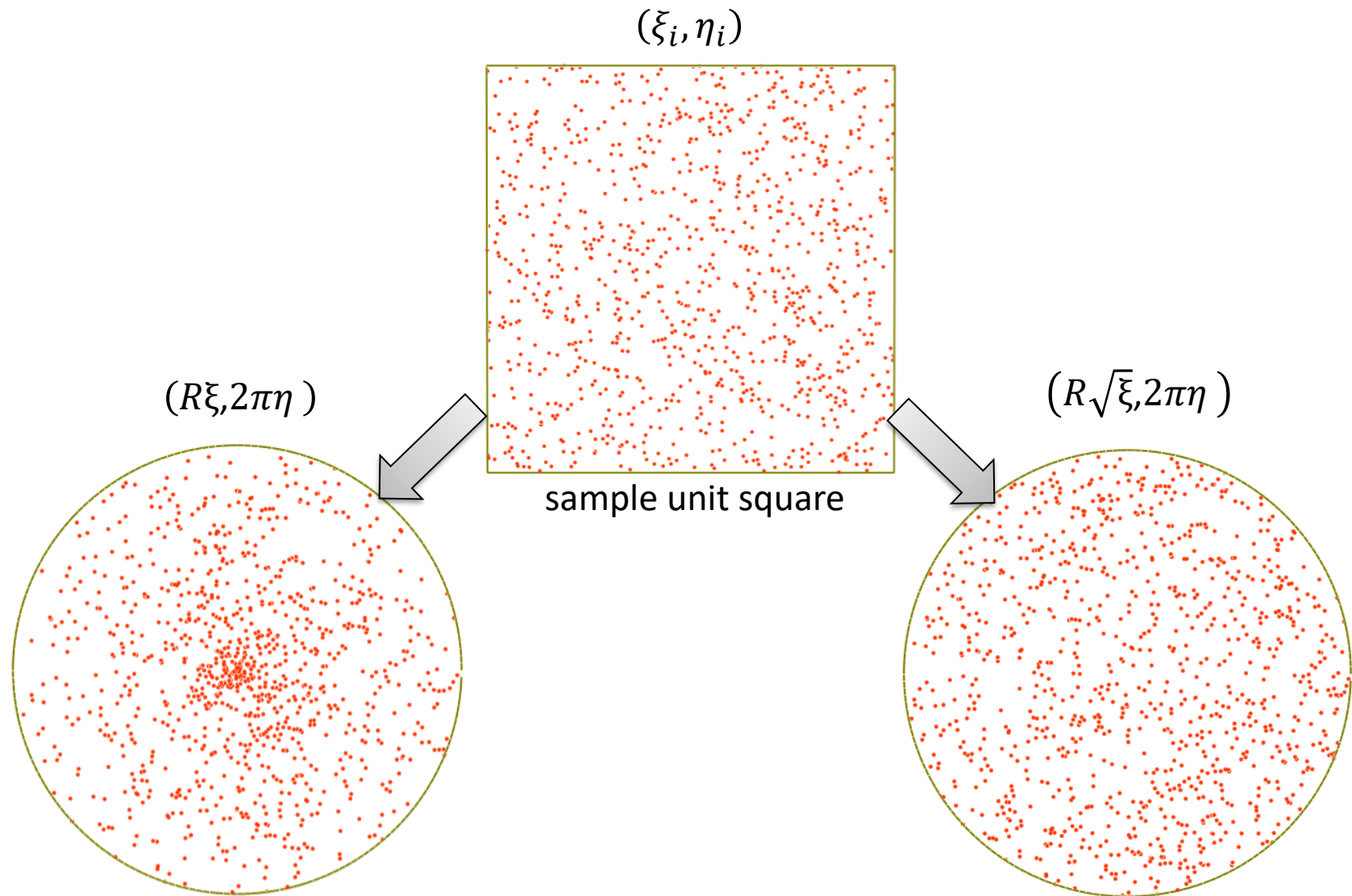
Transformation Sampling

2D – marginalization



- Secondly, we assume then ξ and x have been chosen and relate y and η in the function component $g_y(\xi, \eta)$ by computing the conditional probability pdfs and cdfs:
 $p_{y|x}(x, y) = p_x(x, y)/p_x(x)$, with $p_{\eta|\xi}(\xi, \eta) = 1$
 $P_{y|x}(x, y) = \int_{y_{\min}}^{\bar{y}} p_{y|x}(x, \tilde{y}) d\tilde{y}$, with $P_{\eta|\xi}(\xi, \eta) = P_{\eta}(\xi, \eta) = \eta$
- equating and solving for y :
 $P_{y|x}(x, y) = P_{\eta|\xi}(\eta) = \eta$, and $y = P_{y|x}^{-1}(x, \eta)$
- In summary we can define $\mathbf{g}(\xi)$ as
 $x = g_x(\xi) = P_x^{-1}(\xi)$
 $y = g_y(\xi, \eta) = P_{y|x}^{-1}(x, \eta)$
- And the sampling procedure as
 - $x_i = P_x^{-1}(\text{rand}(0,1))$
 - $y_i = P_{y|x}^{-1}(x_i, \text{rand}(0,1))$

Transformation Sampling Illustration in 2D for Disk





Transformation of Distribution

- we want to define uniform distribution on disk in Cartesian coordinates x, y and then transform it to radial coordinates r, θ
- Change of variable can be expressed by vector valued function $\mathbf{f}$:
$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{f} \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r \cdot \cos \theta \\ r \cdot \sin \theta \end{pmatrix}$$
- Theorem: a function $\mathbf{f}: \mathbf{A} \rightarrow \mathbf{B}: (r, \theta) \mapsto (x, y)$ relates distribution $p_A(r, \theta)$ on $\mathbf{A}$ with $p_B(x, y)$ on $\mathbf{B}$ via:
$$p_A(r, \theta) = |\mathbf{f}_{r, \theta}| \cdot p_B(x, y)$$
- where $|\mathbf{f}_{r, \theta}|$ is determinant of Jacobian matrix of $\mathbf{f}$.
- In case of uniform sampling of disk of radius R , we start with
$$p_B(x, y) = 1/\pi R^2$$
- next we compute Jacobian of $\mathbf{f}$:
$$\mathbf{f}_{r, \theta} = \begin{pmatrix} \cos \theta & -r \cdot \sin \theta \\ \sin \theta & r \cdot \cos \theta \end{pmatrix}$$
- determinant of Jacobian is
$$|\mathbf{f}_{r, \theta}| = r \cdot \cos^2 \theta + r \cdot \sin^2 \theta = r$$

Under the condition that $r > 0$
- This finally yields:
$$p_A(r, \theta) = r/\pi R^2$$

Transformation Sampling

Example 2D



- Given $p(r, \theta) = r/\pi R^2$.
- First we integrate over θ to get marginal distribution in r :

$$p_r(r) = \int_0^{2\pi} \frac{r}{\pi R^2} d\theta = \frac{2r}{R^2}$$

- Next we compute the cdf by integration over r :

$$P_r(r) = \int_0^r \frac{2r' dr'}{R^2} = \frac{r^2}{R^2}$$

- From $P_r(r) = \xi$ we get: $r = R\sqrt{\xi}$
- Next we compute conditional pdf for θ :

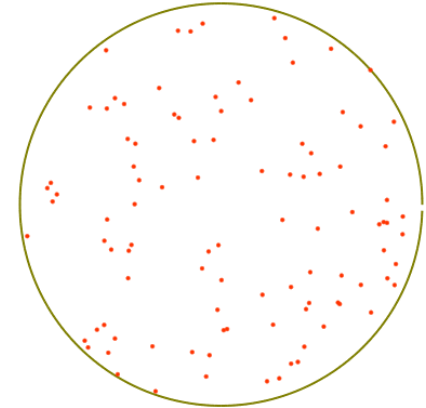
$$p_{\theta|r}(r, \theta) = \frac{p(r, \theta)}{p_r(r)} = \frac{r}{\pi R^2} / \frac{2r}{R^2} = \frac{1}{2\pi}$$

- Integration over θ yields $P_{\theta|r}(\theta) = \frac{\theta}{2\pi}$
- Inverting $P_{\theta|r}(\theta) = \eta$ yields: $\theta = 2\pi\eta$.

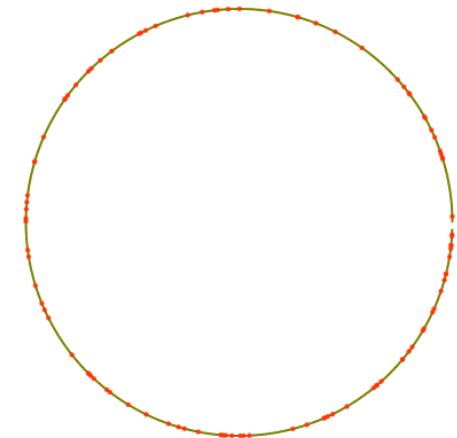
Transformation Sampling Projection



- ◆ To **sample manifolds** it is sometimes useful to sample in the **embedding space** and then project the samples onto the manifold
- ◆ example unit circle:
 - ◆ sample unit disk with rejection or transformation sampling
 - ◆ normalize sample
 - ◆ this yields uniform sampling of circle
- ◆ example sphere surface
 - ◆ sampling solid unit sphere with rejection/transf. sampling
 - ◆ normalize sample
 - ◆ this yields uniform sampling on sphere surface



uniform disk sampling



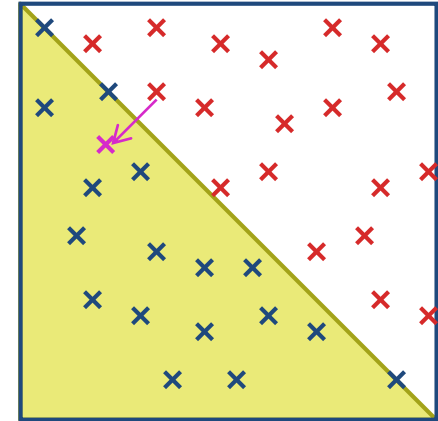
projected onto circle yields
uniform circle sampling

Transformation Sampling

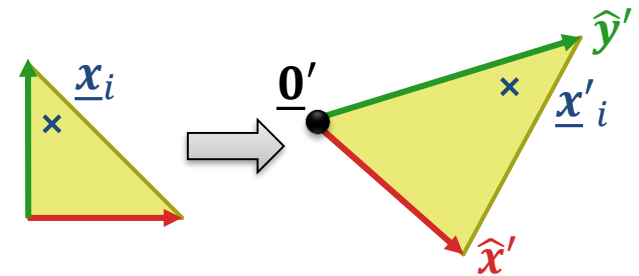
Double Coverage



- For some shapes it is easier to sample a shape of twice the size and then project the samples from wrong side back to the right side.
- example unit triangle:
 - extend triangle to unit square
 - sample unit square
 - reflect samples from upper right triangle back into the triangle
- sampling of unit triangle can be transformed to arbitrary triangle
- example hemisphere for given normal $\hat{n}$:
 - sample $\hat{\omega}$ from sphere surface
 - if $\langle \hat{n}, \hat{\omega} \rangle < 0$ negate $\hat{\omega}$



sampling of unit triangle



$$\underline{x}'_i = (\hat{x}' \quad \hat{y}' \quad \underline{o}') \begin{pmatrix} \underline{x}_i \\ 1 \end{pmatrix}$$

transformation from unit triangle to arbitrary triangle

Transformation Sampling a Diffuse BRDF in angle parameterization



- Given a constant BRDF ρ_{diff} , let probability density $p(\varphi, \theta)$ be proportional to brdf times cos term in angle parameteriz.
- normalization makes density independent of ρ_{diff} .
- Use marginalization approach:

$$p_{\theta}(\theta) = \int_{-\pi}^{\pi} \frac{\cos \theta \sin \theta}{\pi} d\varphi = 2 \cos \theta \sin \theta$$
$$\xi = P_{\theta}(\theta) = \int_0^{\theta} p_{\theta}(\tilde{\theta}) d\tilde{\theta} =$$
$$\left[-\cos^2 \tilde{\theta} \right]_0^{\theta} = (1 - \cos^2 \theta) = \sin^2 \theta$$

- solving for angles yields

$$\theta = \arcsin \sqrt{\xi}$$
$$\varphi = \pi(2\eta - 1)$$

$$p(\varphi, \theta) \propto \rho_{\text{diff}} \cos \theta \sin \theta$$

compare
script 02
slide 14



$$\int_{-\pi}^{\pi} \int_0^{\frac{\pi}{2}} \cos \theta \sin \theta d\theta d\varphi = \pi$$

$$p(\varphi, \theta) = \frac{\cos \theta \sin \theta}{\pi}$$

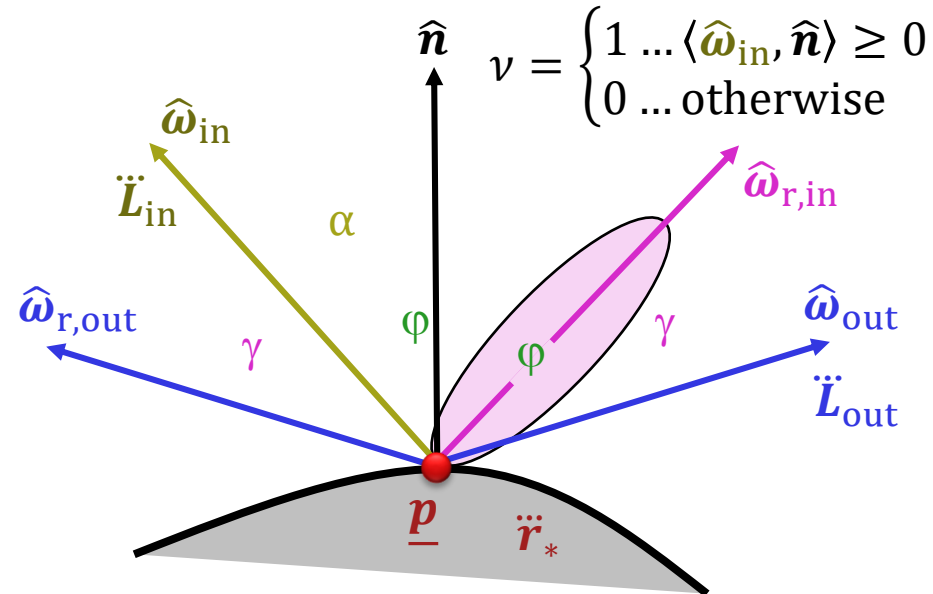
$$p_{\varphi|\theta}(\varphi, \theta) = \frac{p(\varphi, \theta)}{p_{\theta}(\theta)} = \frac{1}{2\pi}$$

$$\eta = P_{\varphi|\theta}(\varphi, \theta) = \int_{-\pi}^{\varphi} p_{\varphi|\theta}(\tilde{\varphi}, \theta) d\tilde{\varphi} =$$
$$\frac{[\tilde{\varphi}]_{-\pi}^{\varphi}}{2\pi} = \frac{\varphi + \pi}{2\pi}$$



$$\ddot{\mathbf{L}}_{\text{spec,Phong,mod}} = \nu \cdot \ddot{\mathbf{r}}_{\text{spec}} \otimes \ddot{\mathbf{L}}_{\text{in}} \cdot \langle \hat{\mathbf{w}}_{\text{out}}, \hat{\mathbf{w}}_{\text{r,in}} \rangle_+^m \cdot \langle \hat{\mathbf{w}}_{\text{in}}, \hat{\mathbf{n}} \rangle_+$$

- For physical plausibility the Phong shading model is modified by multiplying the cosine term to specular part
- This makes the mod. Phong-BRDF symmetric in $\hat{\omega}_{in}$ & $\hat{\omega}_{out}$.
- The angle γ is also found between $\hat{\omega}_{r,out}$ and $\hat{\omega}_{in}$.
- Therefore one uses $\hat{\omega}_{r,out}$ as north pole to sample $\hat{\omega}_{in}$ with γ replacing θ and φ rotating around $\hat{\omega}_{r,out}$ instead of $\hat{n}$.
- It is not possible anymore to incorporate ν and $\langle \hat{\omega}_{in}, \hat{n} \rangle_+$ into the probability density.



$$\begin{aligned}\ddot{\mathbf{f}}_{\text{spec,Phong,mod}} &= \nu \cdot \ddot{\mathbf{r}}_{\text{spec}} \cdot \langle \hat{\omega}_{\text{out}}, \hat{\omega}_{\text{r,in}} \rangle_+^m \\ &= \nu \cdot \ddot{\mathbf{r}}_{\text{spec}} \cdot \langle \hat{\omega}_{\text{r,out}}, \hat{\omega}_{\text{in}} \rangle_+^m\end{aligned}$$

$$p(\varphi, \gamma) \propto \cos^m \gamma \sin \gamma$$

Approximate Transformation

Sampling a Phong Lobe II



- normalization of density from integration over hemisphere.
- Using the marginalization approach:

$$\int_{-\pi}^{\pi} \int_0^{\frac{\pi}{2}} \cos^m \gamma \sin \gamma \, d\gamma d\varphi = \frac{2\pi}{m+1}$$
$$p(\varphi, \gamma) = \frac{m+1}{2\pi} \cos^m \gamma \sin \gamma$$

$$\xi = P_{\gamma}(\gamma) = \int_0^{\gamma} \int_{-\pi}^{\pi} \frac{m+1}{2\pi} \cos^m \tilde{\gamma} \sin \tilde{\gamma} \, d\varphi d\tilde{\gamma} = [-\cos^{m+1} \tilde{\gamma}]_0^{\gamma} = 1 - \cos^{m+1} \gamma$$

$$\eta = P_{\varphi|\gamma}(\gamma, \varphi) = \int_{-\pi}^{\varphi} \frac{1}{2\pi} d\tilde{\varphi} = \frac{[\varphi']_{-\pi}^{\varphi}}{2\pi} = \frac{\varphi + \pi}{2\pi}$$

- solving for angles yields
- ν can be incorporated by rejection sampling

$$\gamma = \arccos \sqrt[m+1]{1 - \xi}$$
$$\varphi = \pi(2\eta - 1)$$

Approximate Transformation

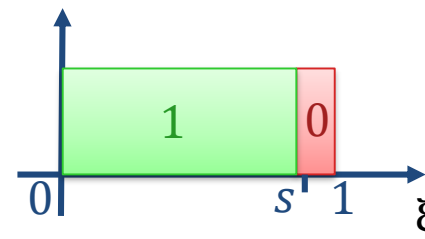
Sampling a Phong Lobe -Pseudocode



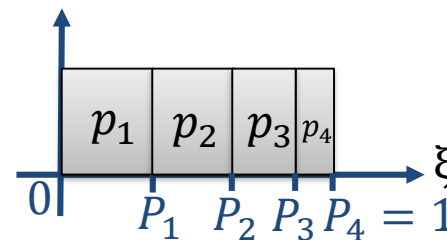
```
/// sample by transformation
sample_Phong_lobe( $\hat{\omega}_{out}$ ,  $\hat{n}$ , m)
{
    do {
        /// sample density
         $\xi$  = uniform_rand(0,1);
         $\gamma$  = arccos  $\sqrt[m+1]{1 - \xi}$ ;
         $\eta$  = uniform_rand(0,1);
         $\varphi$  =  $\pi(2\eta - 1)$ ;
        /// construct sampled direction
         $\hat{\omega}_{r,out}$  = reflect( $\hat{\omega}_{out}$ ,  $\hat{n}$ );
        extend_to_base( $\hat{\omega}_{r,out}$ ,  $\hat{x}$ ,  $\hat{y}$ );
         $\hat{\omega}_{in}$  = sphere_coords( $\varphi$ ,  $\gamma$ ,  $\hat{x}$ ,  $\hat{y}$ ,  $\hat{\omega}_{r,out}$ );
        /// reject as long as  $v = 0$ 
    } while ( $\langle \hat{\omega}_{in}, \hat{n} \rangle < 0$ )
    return  $\hat{\omega}_{in}$ ;
}
```

Sampling a Discrete Distribution

- Russian Roulette we need to sample only 2 states $\{0,1\}$ according to $p_1 = s$ and $p_0 = 1 - s$:
 - sample $\xi \in [0,1]$ uniformly
 - select 1 if $\xi \leq s$ and 0 otherwise
- More Discrete Distributions
 - for n light sources we define p_i proportional to power,
 - for triangle mesh with n triangles define p_i prop. to triangle area
 - for BRDF parts (D/S/M) define p_i prop. to hemispherical reflectance
- Sampling discrete distributions
 - compute cumulative distribution $P_i = \sum_{j=1}^i p_j$ with $P_0 = 0$
 - sample $\xi \in [0,1]$ uniformly
 - search primitive i with $P_{i-1} \leq \xi \leq P_i$

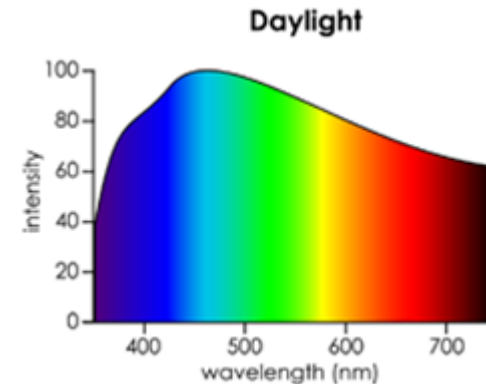


sampling from $\{0,1\}$ in ratio $s: 1 - s$



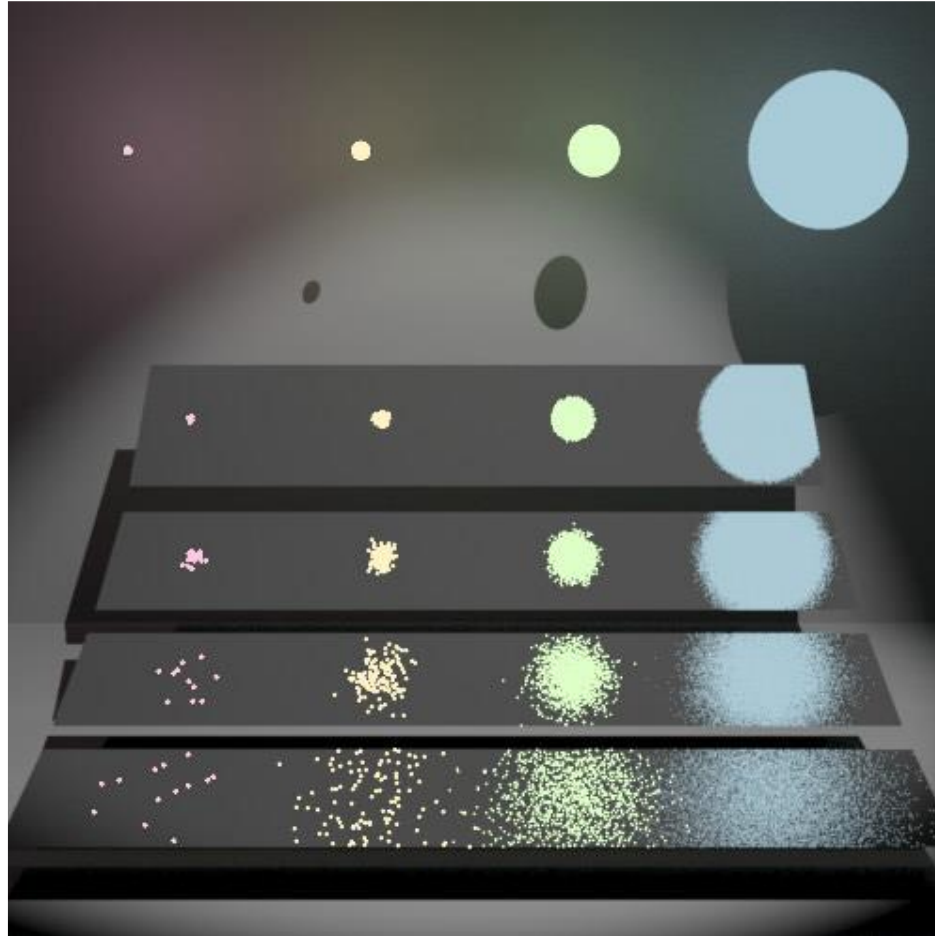
Sampling a measured Function

- the spectrum of certain light sources can be measured on discrete sample locations
- the emitted radiance of a point light source can be measured on a discrete sample of directions
- BRDFs can also be measured (see lecture on advanced materials)
- For each discretely sampled function we can derive a discrete distribution and sample it according to the previous slide



Multiple Importance Sampling

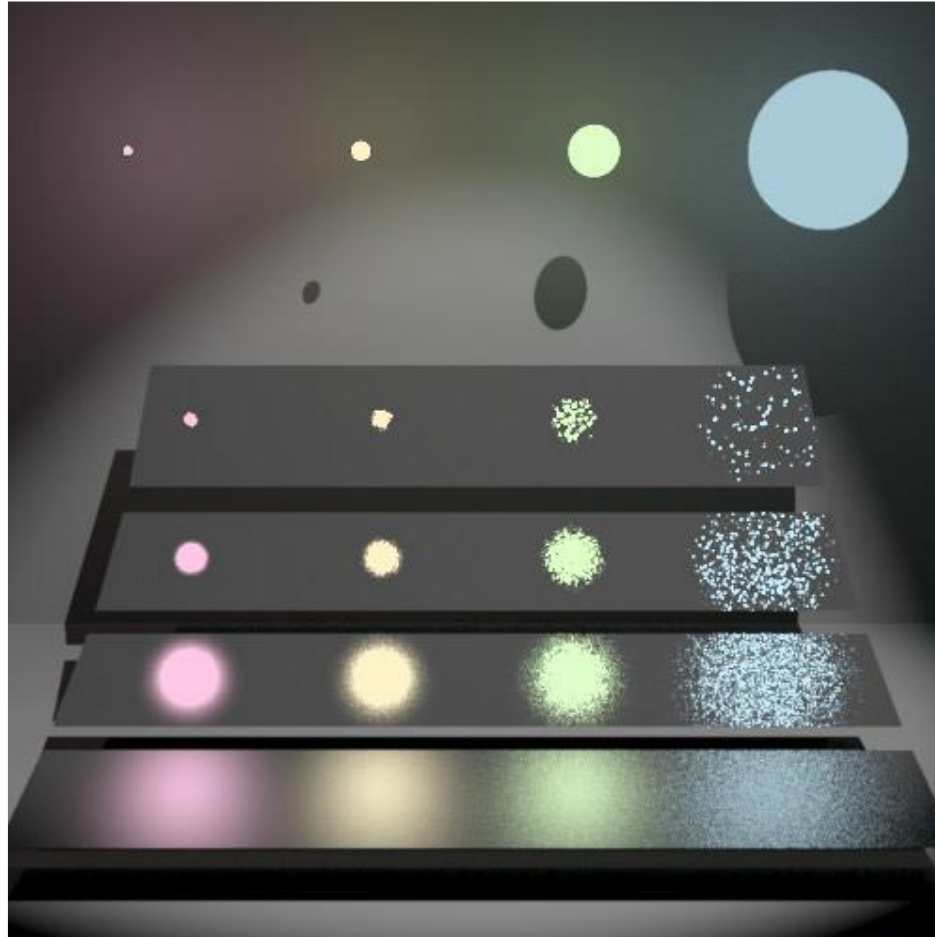
Motivation



BRDF sampling

Multiple Importance Sampling

Motivation



direct light sampling

Multiple Importance Sampling

Introduction

- Given m sampling strategies $p_i(x)$, **multiple importance sampling (MIS)** weights the strategies to **minimize overall variance**

- If integrand is product of several functions f_i

$$\int f(x) dx = \int \prod_{i=1}^m f_i(x) dx$$

per function sampling strategies p_i can be combined.

- For $N = \sum_{i=1}^m n_i$, we sample p_i with n_i samples $x_{i,j=1..n_i}$ and define **MIS-estimator** based on **weights $\omega_i(x)$** :

$$\hat{f}_N^{MIS} = \sum_{i=1}^m \frac{1}{n_i} \sum_{j=1}^{n_i} \omega_i(x_{i,j}) \frac{f(x_{i,j})}{p_i(x_{i,j})}$$

- note:** weights ω_i can depend on samples $x_{i,j}$.

Preliminaries

- The weights $\omega_i(x)$ need to be normalized if $f(x) \neq 0$ and zero where p_i becomes zero:

$$f(x) \neq 0 \Rightarrow \sum_{i=1}^m \omega_i(x) = 1 \quad \wedge \quad p_i(x) = 0 \Rightarrow \omega_i(x) = 0$$

Conclusions

- the second condition allows sampling techniques to ignore parts of the domain by setting p_i to zero.
- Veach shows in his PhD thesis that for any weight functions that fulfill the two conditions the MIS-estimator is **unbiased**:

$$E[\hat{f}_N^{MIS}] = \sum_{i=1}^m \frac{1}{n_i} \sum_{j=1}^{n_i} \int_{\Omega} \omega_i(x) \frac{f(x)}{p_i(x)} dx = \int_{\Omega} \sum_{i=1}^m \omega_i(x) f(x) dx = \int f(x) dx$$

Multiple Importance Sampling

Heuristics



Balance Heuristic

$$\hat{f}_N^{MIS} = \sum_i \frac{1}{n_i} \sum_{j=1}^{n_i} \omega_i(x_{i,j}) \frac{f(x_{i,j})}{p_i(x_{i,j})}$$

- Choose weights proportional to n_i and $p_i(x)$:

$$\omega_i(x) = \frac{n_i p_i(x)}{\sum_{k=1}^m n_k p_k(x)}$$

- As n_i and p_i cancel out the balance heuristic is equivalent to sample from the average density $\bar{p}(x) = \frac{1}{N} \sum_{k=1}^m n_k p_k(x)$:

$$\hat{f}_N^{MIS,Balance} = \frac{1}{N} \sum_i \sum_{j=1}^{n_i} \frac{f(x_{i,j})}{\bar{p}(x_{i,j})}$$

- Veatch shows in his PhD thesis that given the integral value F the balance heuristic has variance close to an optimal choice of weights with a maximal variance increase of

$$\left(\frac{1}{\min_{i \in \{1, \dots, m\}} n_i} - \frac{1}{\sum_{i=1}^m n_i} \right) F^2$$

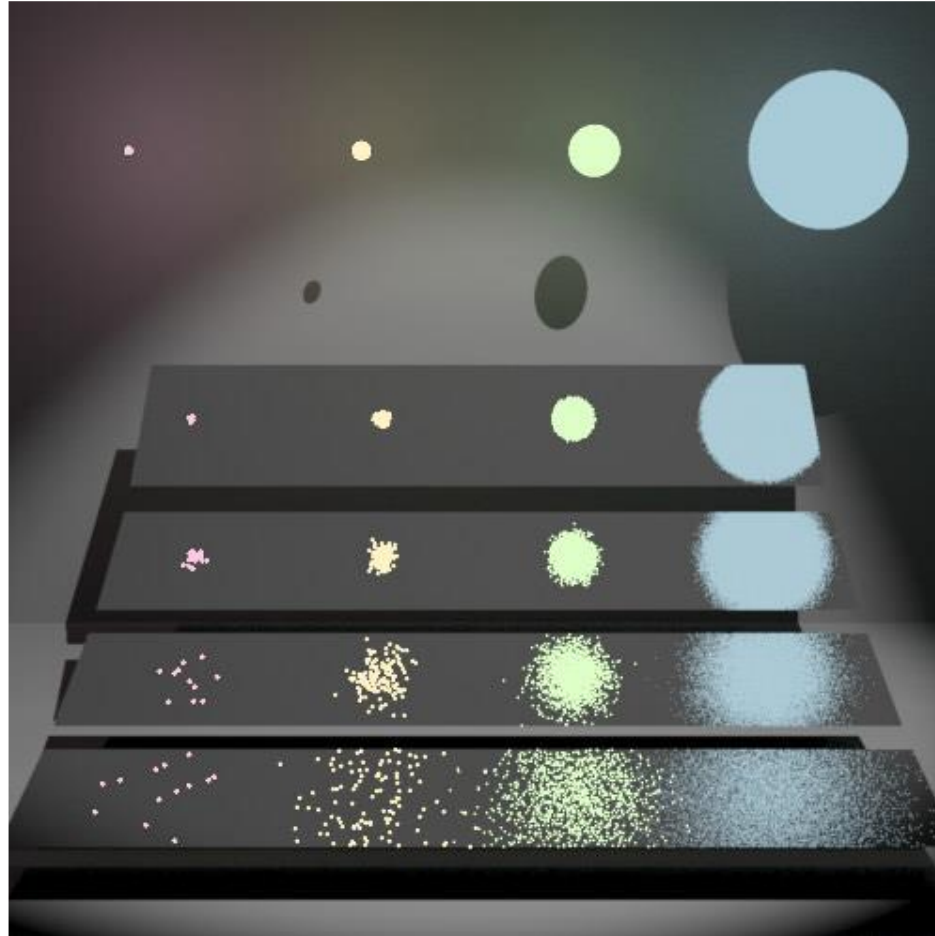
Power Heuristic

- ◆ The balance heuristic can be used most times
- ◆ If the variances of individual estimators get very low, the power heuristic can help to reduce variance further.
- ◆ Given the parameter β Veach defines in his PhD thesis

$$\omega_i^\beta(x) = \frac{(n_i p_i(x))^\beta}{\sum_k (n_k p_k(x))^\beta}$$

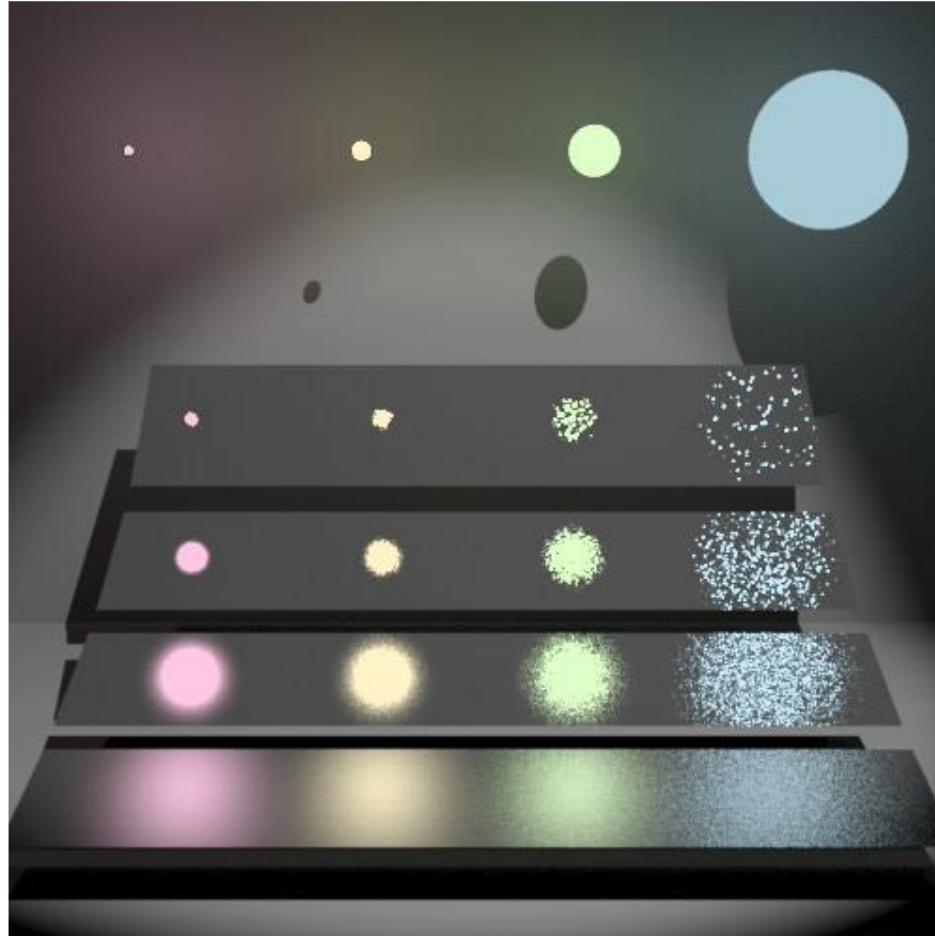
- ◆ Again Veach proves a bound for the power heuristic close to the optimal MIS weighting.
- ◆ In numerical experiments he finds a good choice for the exponent: $\beta = 2$.

Multiple Importance Sampling Comparison



BRDF sampling

Multiple Importance Sampling Comparison



direct light sampling

Multiple Importance Sampling Comparison



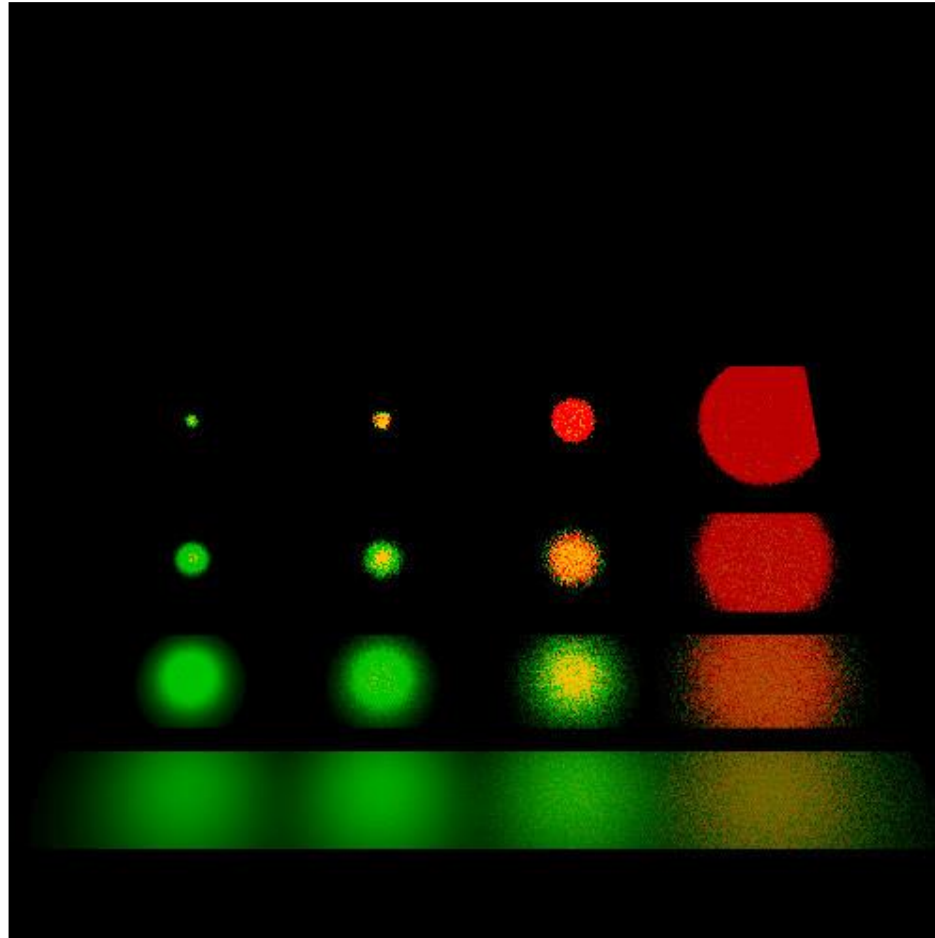
Balance Heuristic

Multiple Importance Sampling Comparison



Power Heuristic with $\beta = 2$

Multiple Importance Sampling Comparison



Weights of Power Heuristic color coded
with **red for BRDF** and **green for direct light sampling**



Summary

- ◆ For a given pdf there are a multitude of different sampling strategies from which one can choose
- ◆ Rejection sampling works even for unnormalized pdfs but in this case does not provide any means to evaluate the pdf at a sample
- ◆ Transformation sampling allows to use uniformly distributed samples and can be applied when the pdf is integrable and the cdf invertible.
- ◆ Pdfs resulting from the product of functions can be handled with multiple importance sampling, which samples each function and additionally weights the samples based on the function values.