Monte Carlo Techniques and Optimisations for Physically Based Rendering

Eric Lafortune
Department of Computer Science, K.U.Leuven, Belgium
Program of Computer Graphics, Cornell University, USA
http://www.graphics.cornell.edu/~eric/
eric@graphics.cornell.edu

1. Monte Carlo integration

2. More samples

3. Stratified sampling

4. Antithetic variates

5. Control variates

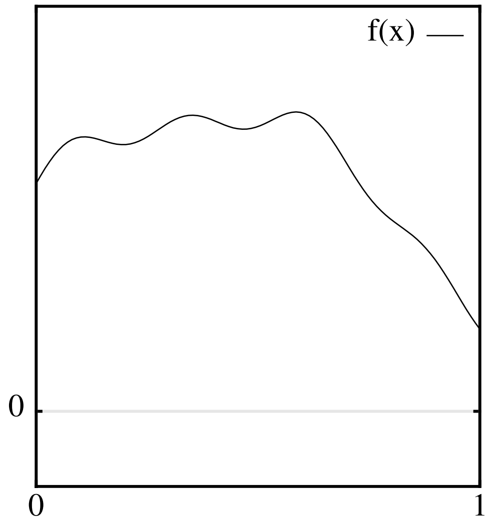
6. Importance sampling

7. Adaptive sampling

8. Russian roulette

9. Conclusions

1. Monte Carlo integration



$$I = \int_0^1 f(x)dx$$

Estimator:

$$\langle I \rangle = f(\bar{x})$$

Expected value:

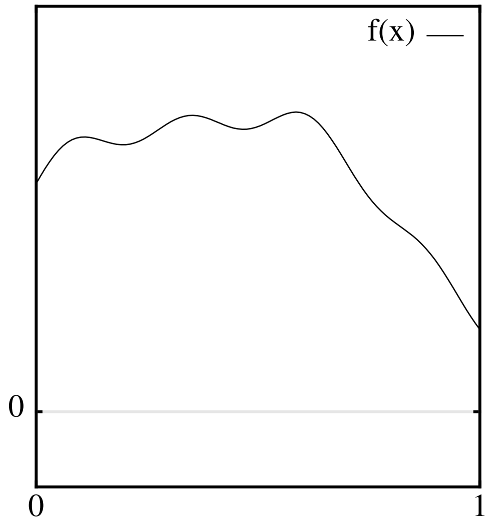
$$E(\langle I \rangle) = \int_0^1 f(x)dx = I$$

Variance:

$$\mathrm{var}(\langle I \rangle) = \int_0^1 (f(x) - I)^2 dx = \int_0^1 f^2(x)dx - I^2$$

Simple technique, even for complex problems.

2. More samples



$$I = \int_0^1 f(x)dx = \frac{1}{N}\left(\int_0^1 f(x)dx + \ldots + \int_0^1 f(x)dx\right)$$

Estimator:

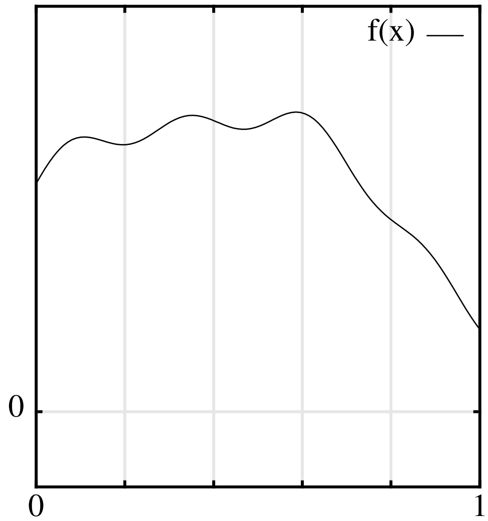
$$\langle I\rangle_N = \frac{1}{N}\sum_{i=1}^{N} f(\bar{x}_i)$$

Variance:

$$\operatorname{var}(\langle I\rangle_N) = \operatorname{var}(\langle I\rangle_1)/N$$

Slow convergence $\Rightarrow$ variance reduction.

3. Stratified sampling



$$I = \int_0^1 f(x)dx = \int_{X_1} f(x)dx + \ldots + \int_{X_N} f(x)dx$$

Estimator:

$$\langle I \rangle = \frac{1}{N}\sum_{i=1}^{N} f(\bar{x}_i)$$

Variance:

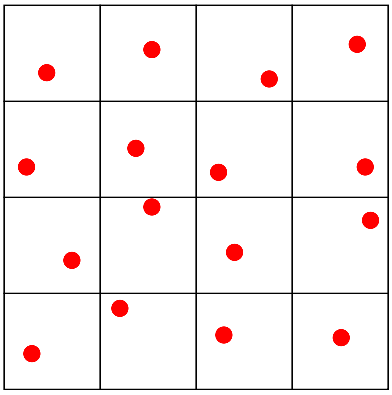
$$\mathrm{var}(\langle I \rangle) = \frac{1}{N}\int_0^1 f^2(x)dx - \sum_{i=1}^{N} I_i^2$$

Stratified samples are always better than
or at least as good as plain random samples.

For path tracing, on a set of test images:
reduction of the RMS error by 0...10...20...50 %.

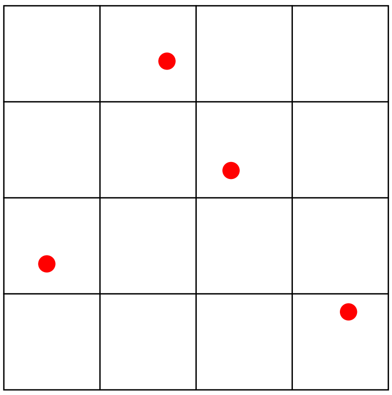
Higher dimensions:

Grid sampling:



-> N^d samples

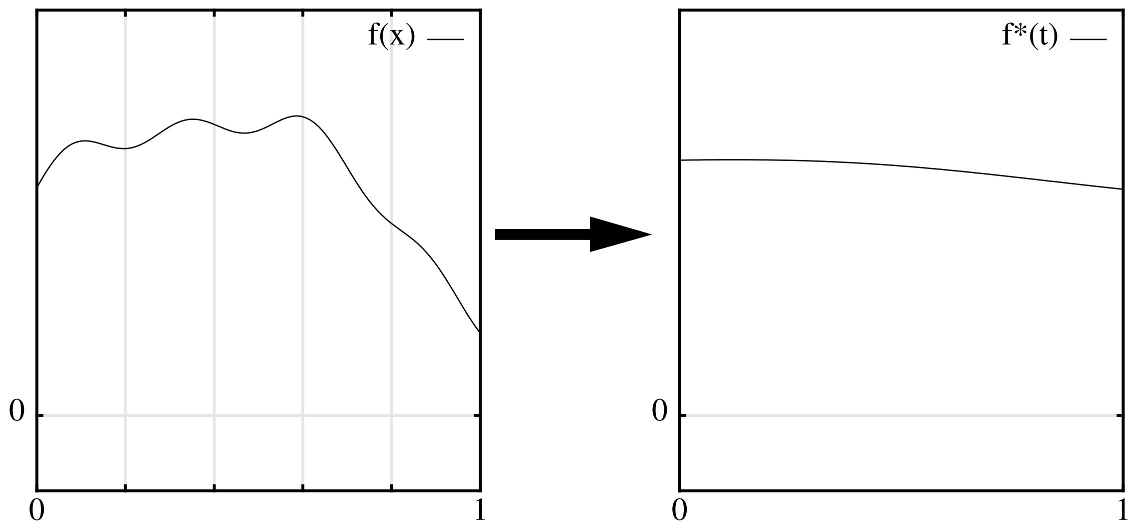
N-rooks sampling:



-> N samples

N-rooks sampling avoids the curse of dimensionality.

4. Antithetic variates / stratified sampling as a transformation



$$I = \int_0^1 f(x)dx = \frac{1}{N}\int_0^1 [f(x_1(t)) + \ldots + f(x_N(t))]dt$$

Estimator:

$$\langle I \rangle = \frac{1}{N}\sum_{i=1}^{N} f(x_i(\bar{t})) \text{ where e.g. } x_i(t) = \frac{i-1+t}{N}$$

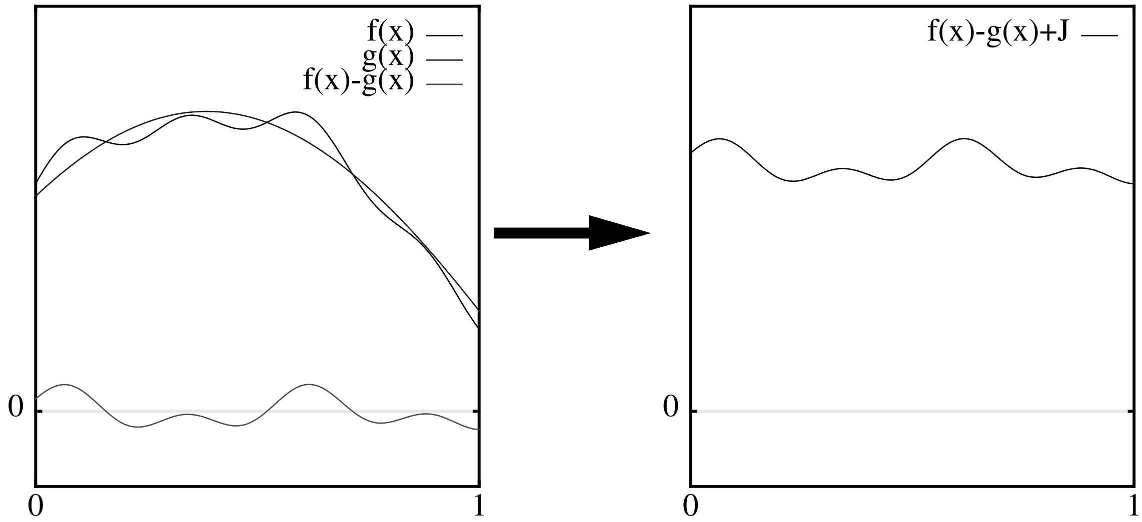
Variance:

$$\mathrm{var}(\langle I \rangle) = \frac{1}{N^2}\sum_{i,j}\int_0^1 f(x_i(t))f(x_j(t))dt - I^2$$

Note: this is not the usual and certainly not the most optimal way of using antithetic variates.

$\boxed{\text{Antithetic variates can produce a complete MC estimate based on a single random number.}}$

5. Control variates



$$I = \int_0^1 f(x)dx = \int_0^1 (f(x) - g(x) + J)dx$$

Estimator:

$$\langle I \rangle = f(\bar{x}) - g(\bar{x}) + J \quad \text{where } J = \int_0^1 g(x)dx$$

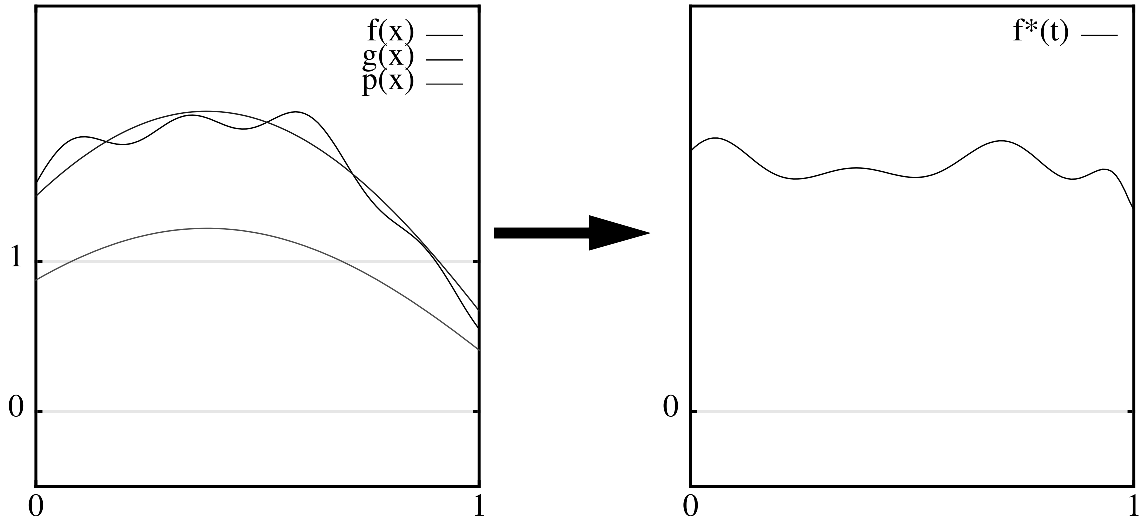
Variance:

$$\mathrm{var}(\langle I \rangle) = \int_0^1 (f(x) - g(x) + J)^2 dx - I^2$$

Control variates yield a lower variance
if the approximating function is good enough.

For path tracing, using the BRDF as a control variate:
reduction of the RMS error by 4...14 %.

6. Importance sampling



$$I = \int_0^1 f(x)dx = \int_0^1 \frac{f(P^{-1}(t))}{p(P^{-1}(t))}dt$$

Estimator:

$$\langle I \rangle = \frac{f(P^{-1}(\bar{t}))}{p(P^{-1}(\bar{t}))} = \frac{f(\bar{x})}{p(\bar{x})} \text{ where } \bar{x} = P^{-1}(\bar{t})$$

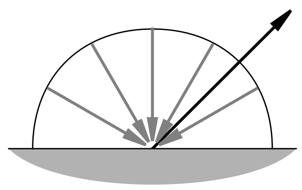
Variance:

$$\mathrm{var}(\langle I \rangle) = \int_0^1 \frac{f^2(P^{-1}(t))}{p^2(P^{-1}(t))}dt - I^2 = \int_0^1 \frac{f^2(x)}{p(x)}dx - I^2$$

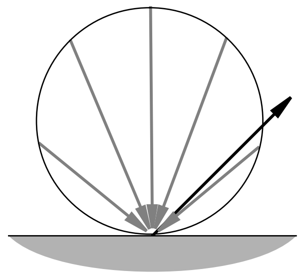
Importance sampling yields a lower variance
if the sampling function is good enough.

Application to path tracing, sampling the rendering equation:

- Uniform sampling of the hemisphere:

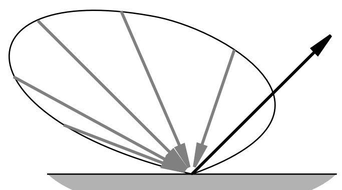


- Importance sampling, proportional to the cosine factor:



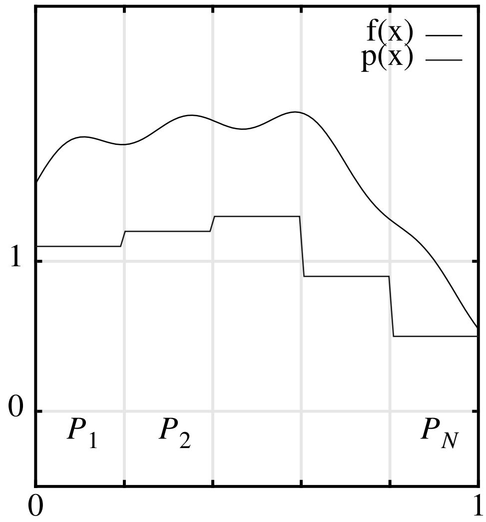
Reduction of the RMS error by 5...30 %.

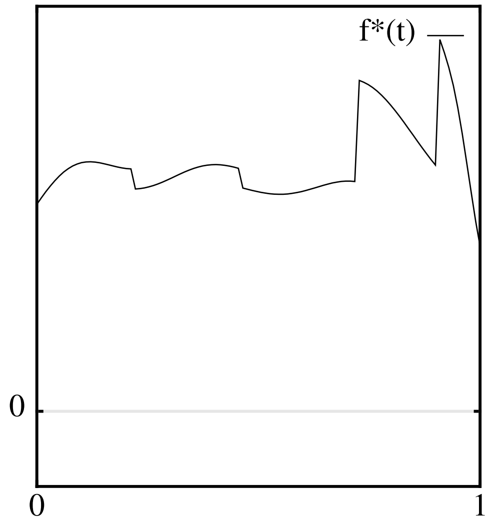
- Importance sampling, proportional to the cosine factor times the BRDF:



Reduction of the RMS error by 30...80 %.

7. Adaptive sampling





$$I = \int_0^1 f(x)dx = P_1 \int_{X_1} \frac{f(x)}{P_1} dx + \ldots + P_N \int_{X_N} \frac{f(x)}{P_N} dx$$

Estimator:

$$\langle I \rangle = \frac{f(\bar{x}_1)}{P_1} \text{ or } \frac{f(\bar{x}_2)}{P_2} \text{ or } \ldots \text{ or } \frac{f(\bar{x}_N)}{P_N}$$

Variance:

$$\operatorname{var}(\langle I \rangle) = \int_{X_1} \frac{f^2(x)}{2P} dx + \int_{X_2} \frac{f^2(x)}{2(1-P)} dx - I^2$$

Optimal probabilities P_i:

$$P_i = m_i / \sum_{j=1}^{N} m_j \quad \text{where:} \quad m_i = \sqrt{\int_{X_i} f^2(x)dx}$$

Limit cases:

- For estimate $f(x) \approx C$:

$$P_i = 1/N,$$ i.e. stratified sampling.

- For $N \to \infty$:

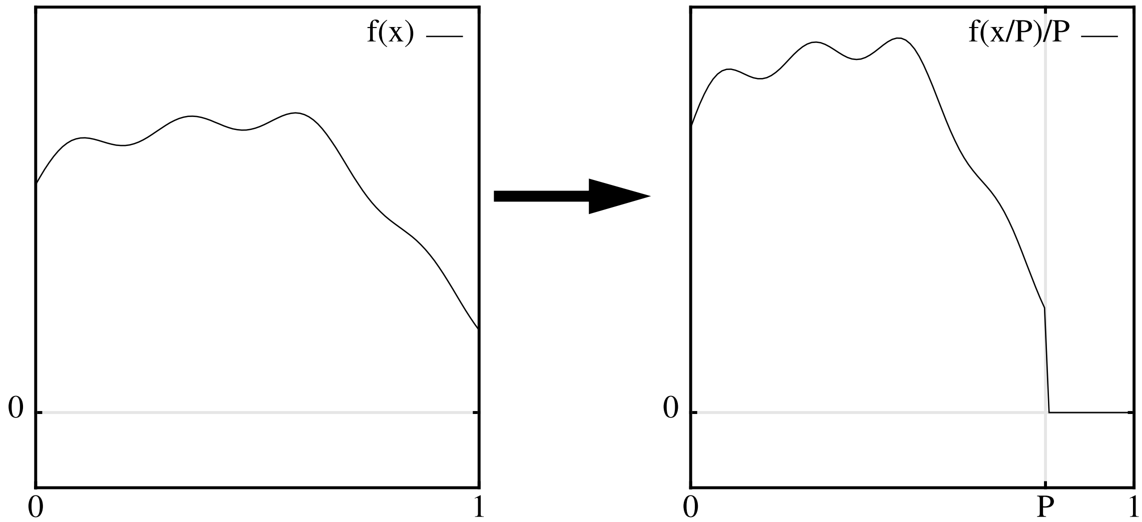
$$dP_i = \frac{\sqrt{\int_{dX_i} f^2(x)dx}}{\sum_{j=1}^{\infty} \sqrt{\int_{dX_j} f^2(x)dx}} \Rightarrow p(x) = \frac{f(x)}{\int_0^1 f(x)dx}$$

i.e. importance sampling.

Adaptive sampling only yields a lower variance if the sampling function is good enough.

For path tracing, using a 5D tree to store an approximation: reduction of the RMS error by 20...50 %.

8. Russian roulette



$$I = \int_0^1 f(x)dx = \int_0^1 \frac{1}{P} f\left(\frac{x}{P}\right) dx$$

Estimator:

$$\langle I \rangle = \frac{1}{P} f\left(\frac{\bar{x}}{P}\right)$$

Variance:

$$\mathrm{var}(\langle I \rangle) = \frac{1}{P}\int_0^1 f^2(x)dx - I^2$$

Russian roulette always yields a higher variance,
but it is unbiased.

9. Conclusions

Monte Carlo methods

- Simple and general solution, even for complex problems.
- Additional information about the problem
  $\Rightarrow$ variance reduction.
- Variance reduction can be regarded as a transformation of the integrand, to make it more flat and therefore easier to integrate.

Problems

- Slow convergence.
- Stochastic noise.
- No deterministic error bounds.
- Variance has to be estimated.

Ongoing research

- Variance reduction removes the stochastic element from Monte Carlo methods
  $\Rightarrow$ MC $\rightarrow$ Quasi-MC $\rightarrow$ classical quadrature ?
- Other 'biased' techniques.