Monte Carlo methods in numerical integration

Robert Rüger

Institut für **T**heoretische **P**hysik
Goethe-Universität Frankfurt

February 10th, 2011

Contents

The goal ...

Numerical evaluation of definite integrals
for high-dimensional functions.

$$\int_{a_1}^{b_1} \mathrm{d}x_1 \cdots \int_{a_D}^{b_D} \mathrm{d}x_D \; f(x_1, \ldots, x_D)$$

Contents

Numerical quadrature

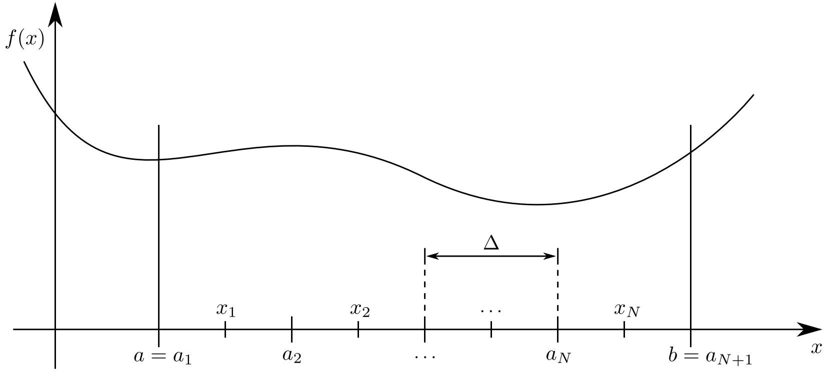
... what is the idea behind it?

Don't integrate f directly!

Find another function $\tilde{f}$ that resembles f
and is easily integrable!

Numerical quadrature

... subdivision of the integration interval



Numerical quadrature

... Taylor expansion in every subinterval

$$
\begin{aligned}
&\int_a^b f(x)\,\mathrm{d}x \\
&= \sum_{n=1}^{N} \int_{a_n}^{a_n+\Delta} f(x)\,\mathrm{d}x \\
&\approx \sum_{n=1}^{N} \int_{a_n}^{a_n+\Delta} \left(f(x_n) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_n} (x - x_n) + \frac{1}{2} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_n} (x - x_n)^2 \right) \mathrm{d}x \\
&= \underbrace{\Delta \sum_{n=1}^{N} f(x_n)}_{I(N)} + \underbrace{\frac{\Delta^3}{24} \sum_{n=1}^{N} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_n}}_{E(N)}
\end{aligned}
$$

The error $E(N)$ of the approximation $I(N)$ is determined by the curvature of f.

$$E(N) = \frac{\Delta^3}{24} \sum_{n=1}^{N} \left. \frac{\mathrm{d}^2 f}{\mathrm{d}x^2} \right|_{x_n}$$

For a well behaved f, the curvature will be bounded.

$$\exists\, C \in \mathbb{R} \mid C \geq \left| \frac{\mathrm{d}^2 f}{\mathrm{d}x^2} \right| \;\forall\, x \in [a, b]$$

$$\Rightarrow \quad |E(N)| \leq \frac{C N \Delta^3}{24} = \frac{C(b-a)^3}{24} \frac{1}{N^2}$$

$$\int_a^b f(x)\,\mathrm{d}x = \Delta \sum_{n=1}^{N} f(x_n) + \mathcal{O}\left(N^{-2}\right)$$

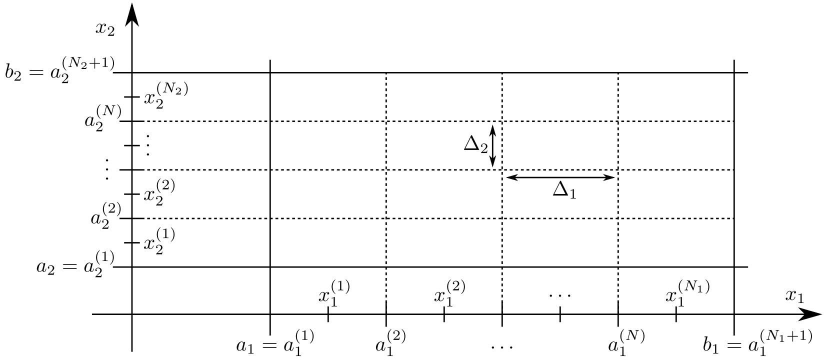
Substitution of Δ reveals another interpretation of $I(N)$.

$$\Delta \sum_{n=1}^{N} f(x_n) = (b-a)\frac{\sum_{n=1}^{N} f(x_n)}{N} = (b-a)\langle f\rangle$$

$\langle f\rangle$ is calculated as the arithmetic mean of the integrand's values at **equidistantly** distributed sample points.

Multidimensional numerical quadrature

... subdivision of the integration hypercube



Multidimensional numerical quadrature

... Taylor expansion in every subhypercube

$$\int_{\{a_d\}}^{\{b_d\}} f(\{x_d\})\,\mathrm{d}^D x = \sum_{\{n_d=1\}}^{\{N_d\}} \int_{\{a_d^{(n_d)}\}}^{\{a_d^{(n_d)}+\Delta_d\}} f\big(\{x_d\}\big)\,\mathrm{d}^D x$$

$$\approx \sum_{\{n_d=1\}}^{\{N_d\}} \int_{\{a_d^{(n_d)}\}}^{\{a_d^{(n_d)}+\Delta_d\}} \Bigg[f\left(\{x_d^{(n_d)}\}\right) + \sum_{i=1}^{D} \left.\frac{\partial f}{\partial x_i}\right|_{\{x_d^{(n_d)}\}} \left(x_i - x_i^{(n_i)}\right) + \frac{1}{2}\sum_{i=1}^{D} \left.\frac{\partial^2 f}{\partial x_i^2}\right|_{\{x_d^{(n_d)}\}} \left(x_i - x_i^{(n_i)}\right)^2 + \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{D} \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\{x_d^{(n_d)}\}} \left(x_i - x_i^{(n_i)}\right)\left(x_j - x_j^{(n_j)}\right) \Bigg]\,\mathrm{d}^D x$$

Second and fourth term vanish due to their linearity ...

Multidimensional numerical quadrature

... Taylor expansion in every subhypercube

$$\int_{\{a_d\}}^{\{b_d\}} f\left(\{x_d\}\right) \,\mathrm{d}^D x$$

$$\approx \underbrace{\sum_{\{n_d=1\}}^{\{N_d\}} f\left(\left\{x_d^{(n_d)}\right\}\right) \prod_{j=1}^{D} \Delta_j}_{I(N)} + \underbrace{\frac{\prod_{j=1}^{D} \Delta_j}{24} \sum_{\{n_d=1\}}^{\{N_d\}} \sum_{i=1}^{D} \left.\frac{\partial^2 f}{\partial x_i^2}\right|_{\{x_d^{(n_d)}\}} \Delta_i^2}_{E(N)}$$

Multidimensional numerical quadrature

... further analysis of the error

$$E(N) = \frac{\prod_{j=1}^{D} \Delta_j}{24} \sum_{\{n_d=1\}}^{\{N_d\}} \sum_{i=1}^{D} \left.\frac{\partial^2 f}{\partial x_i^2}\right|_{\{x_d^{(n_d)}\}} \Delta_i^2$$

We again assume that the curvature of f is bounded by C.

$$\exists\, C \in \mathbb{R} \mid C \geq \left|\frac{\partial^2 f}{\partial x_i^2}\right| \forall\, x_i \in [a_i, b_i]$$

$$\Rightarrow \quad |E(N)| \leq \frac{CN \prod_{j=1}^{D} \frac{b_j - a_i}{N_i}}{24} \sum_{i=1}^{D} \Delta_i^2 = \frac{C \prod_{j=1}^{D} (b_j - a_i)}{24} \sum_{i=1}^{D} \Delta_i^2$$

Multidimensional numerical quadrature

... further analysis of the error

One can easily express the error as a function of the total number of samples, if $N_i = N_j \; \forall \, i, j$.

$$|E(N)| \leq \frac{C \prod_{j=1}^{D}(b_j - a_j)}{24} \sum_{i=1}^{D} \left(\frac{b_i - a_i}{N_i} \right)^2$$

$$= \frac{C \sum_{i=1}^{D}(b_i - a_i)^2 \prod_{j=1}^{D}(b_j - a_j)}{24} \frac{1}{N^{\frac{2}{D}}}$$

Multidimensional numerical quadrature
... putting it all together

$$\int_{\{a_d\}}^{\{b_d\}} f(\{x_d\})\, \mathrm{d}^D x = \sum_{\{n_d=1\}}^{\{N_d\}} f\left(\{x_d^{(n_d)}\}\right) \prod_{j=1}^{D} \Delta_j + \mathcal{O}\left(N^{-\frac{2}{D}}\right)$$

For multidimensional integration the order of the error's decay is **dependent** on the dimensionality of the integration!

Multidimensional numerical quadrature

… another interpretation of $I(N)$

$$\sum_{\{n_d=1\}}^{\{N_d\}} f\left(\{x_d^{(n_d)}\}\right) \prod_{j=1}^{D} \frac{b_j - a_j}{N_j}$$

$$= \frac{\sum_{\{n_d=1\}}^{\{N_d\}} f\left(\{x_d^{(n_d)}\}\right)}{N} \prod_{j=1}^{D} (b_j - a_j)$$

$$= V\langle f\rangle$$

$\langle f\rangle$ is calculated as the arithmetic mean of the integrand's values at sample points that form a high-dimensional **lattice**.

Monte Carlo integration using simple sampling

... what is it?

We have seen that in order to integrate, one has to calculate the mean of a function.

$$\int_V f(\vec{x})\,\mathrm{d}^D x = V\langle f\rangle$$

$\langle f\rangle$ is approximated as the arithmetic mean of sample values.

It is **not** necessary to choose equidistant samples. It is sufficient to use uniformly distributed samples.

$$\lim_{N\to\infty}\frac{N'}{N} = \frac{V'}{V}$$

Random samples are uniformly distributed!

Monte Carlo integration using simple sampling

... the integral estimator

$$I(N) = V \frac{\sum_{n=1}^{N} f(\vec{x}_n)}{N}$$

$\vec{x}_n$ is a random point within the integration region.

In contrast to equidistant sampling it is important to note that $I(N)$ is not a definite value for a finite number of samples N. While equidistant sampling has an **error**, simple random sampling has an **uncertainty**.

$$\mathrm{Var}\big(I(N)\big) = \frac{V^2}{N^2}\sum_{n=1}^{N}\mathrm{Var}\big(f(\vec{x}_n)\big) = \frac{V^2}{N}\mathrm{Var}\big(f(\vec{x}_n)\big)$$
$$= \frac{V^2}{N}\frac{N}{N-1}\Big(\langle f^2\rangle - \langle f\rangle^2\Big) \approx \frac{V^2}{N}\Big(\langle f^2\rangle - \langle f\rangle^2\Big)$$

The uncertainty of the integral estimator is then calculated as its standard deviation.

$$\sigma_I(N) = \sqrt{\mathrm{Var}\big(I(N)\big)} = V\sqrt{\frac{\langle f^2\rangle - \langle f\rangle^2}{N}}$$

Monte Carlo integration using simple sampling

... the integral estimator and its uncertainty

$$\int_V f(\vec{x}_n)\,\mathrm{d}V = \underbrace{V\frac{\sum_{n=1}^{N} f(\vec{x}_n)}{N}}_{I(N)} \pm \underbrace{V\sqrt{\frac{\langle f^2\rangle - \langle f\rangle^2}{N}}}_{\sigma_I(N)}$$

$$= V\frac{\sum_{n=1}^{N} f(\vec{x}_n)}{N} + \mathcal{O}\left(N^{-\frac{1}{2}}\right)$$

It is important to note, that the order of the uncertainty's decay is **independent** of the integration's dimensionality.

Monte Carlo integration using importance sampling

... the idea behind it

We have already seen, that the uncertainty of simple sampling MC integration is dependent on the variance of $f(\vec{x}_n)$.

$$\sigma_i^2(N) = \mathrm{Var}\big(I(N)\big) = \frac{V^2}{N}\mathrm{Var}\big(f(\vec{x}_n)\big)$$

In other words: The uncertainty is small for 'flat' functions and large for 'mountainous' functions!

Idea: Apply a transformation $f \to \tilde{f}$ that does not change the integral but makes the integrand 'flat'!

Monte Carlo integration using importance sampling

... transformation of the integrand

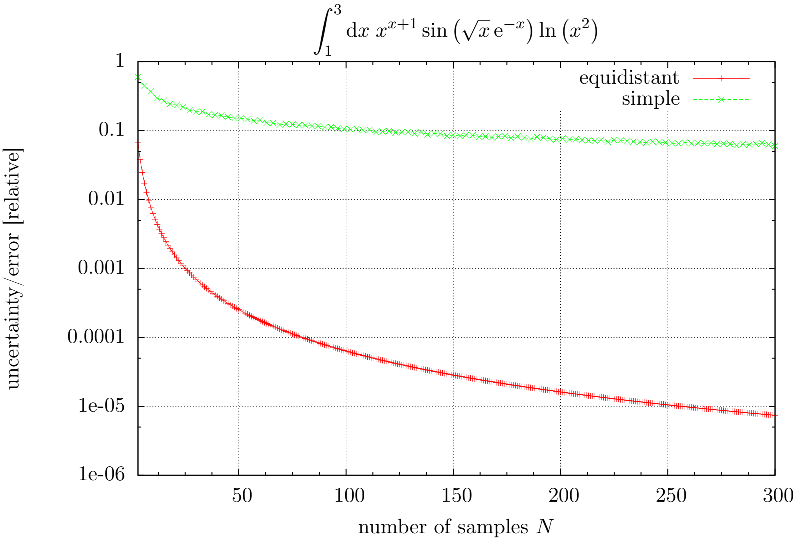
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b \frac{f(x)}{\rho(x)}\,\rho(x)\mathrm{d}x \approx \frac{1}{N}\sum_{n=1}^{N}\frac{f(x_n)}{\rho(x_n)}\,\underbrace{\rho(x_n)\Delta}_{\text{weighting}}$$

Instead of taking evenly distributed samples and weighting them with $\rho(x)$, one could take samples distributed as $\rho(x)$ and weight them evenly.

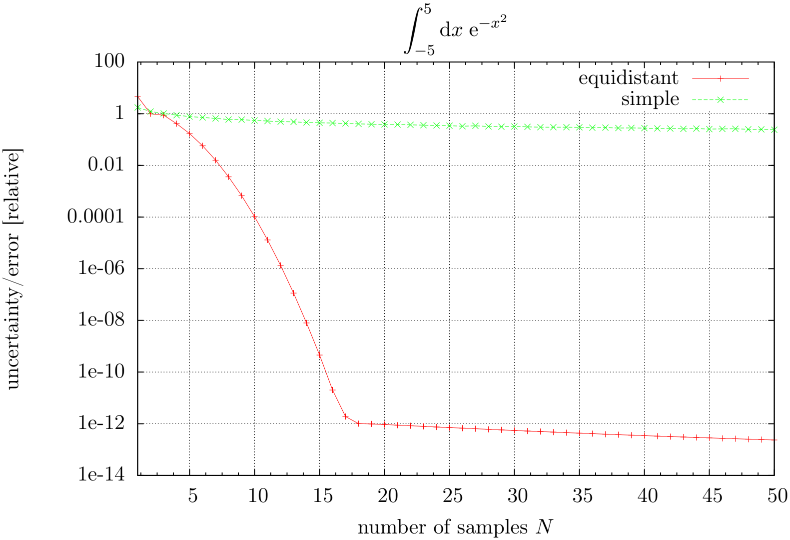
$$\int_a^b f(x)\,\mathrm{d}x \approx \frac{1}{N}\sum_{n=1}^{N}\frac{f\left(x_n^{(\rho)}\right)}{\rho\left(x_n^{(\rho)}\right)}$$

Contents

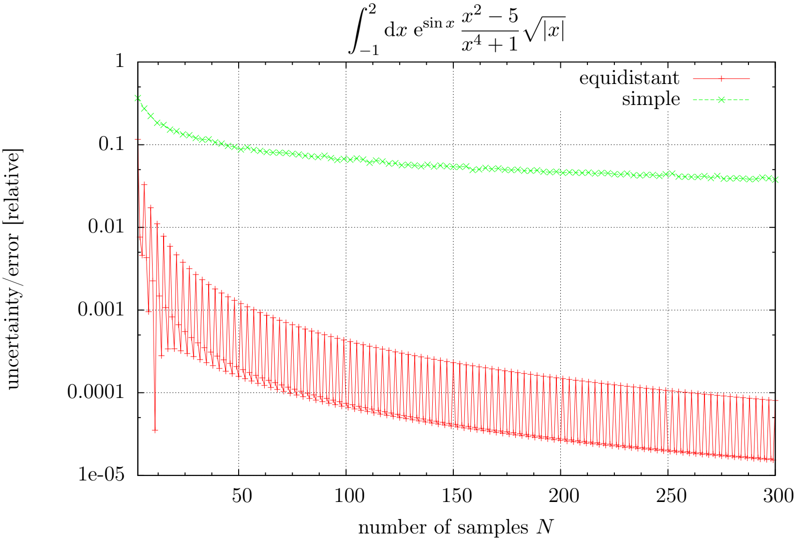
Equidistant sampling vs. simple sampling



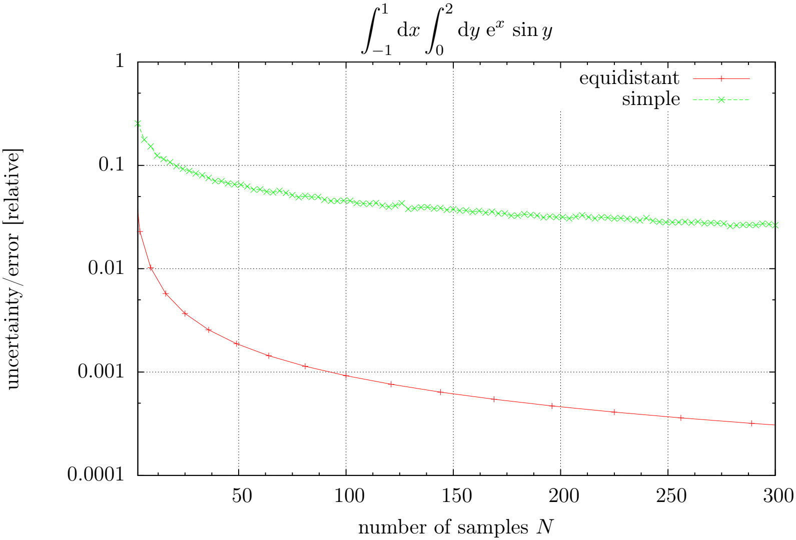
Equidistant sampling vs. simple sampling



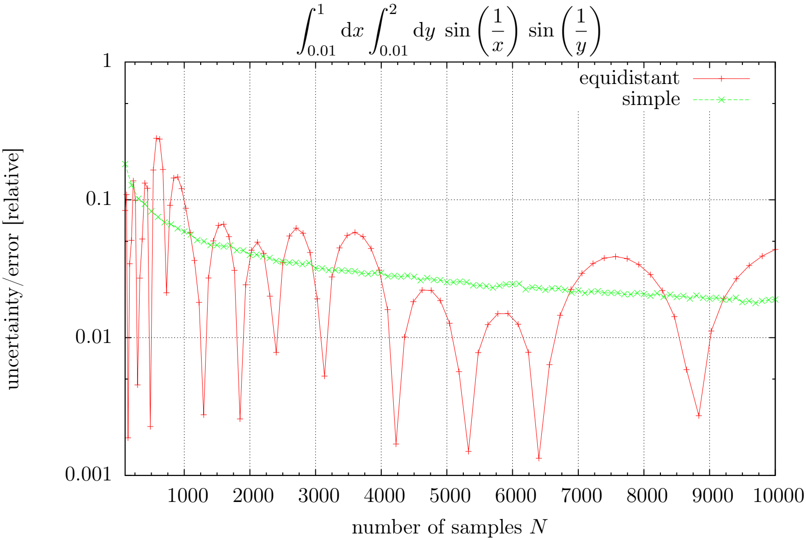
Equidistant sampling vs. simple sampling



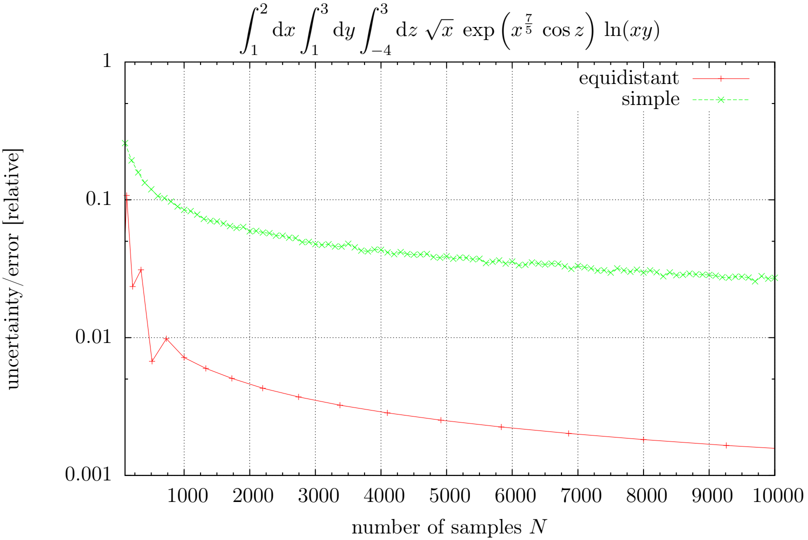
Equidistant sampling vs. simple sampling



Equidistant sampling vs. simple sampling

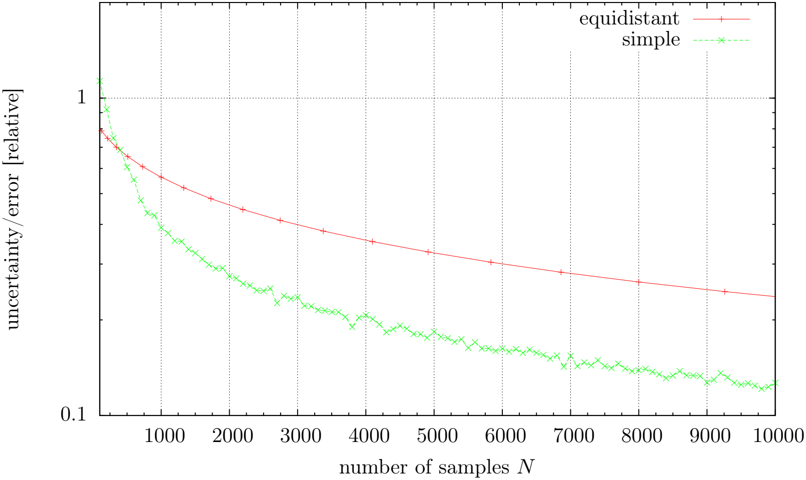


Equidistant sampling vs. simple sampling

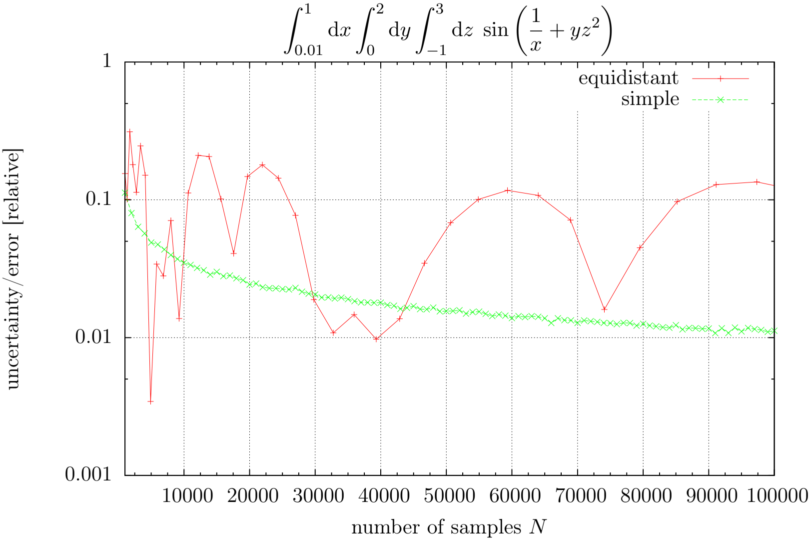


Equidistant sampling vs. simple sampling

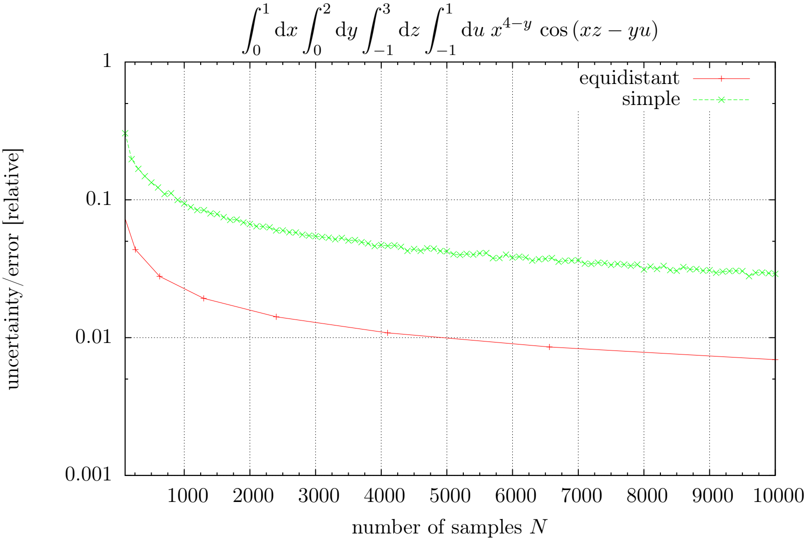
$$\int_{-1}^{2} \mathrm{d}x \int_{1}^{3} \mathrm{d}y \int_{-1}^{0} \mathrm{d}z \;\exp\left(xy\,|z| - x^3 y^2\right)$$



Equidistant sampling vs. simple sampling

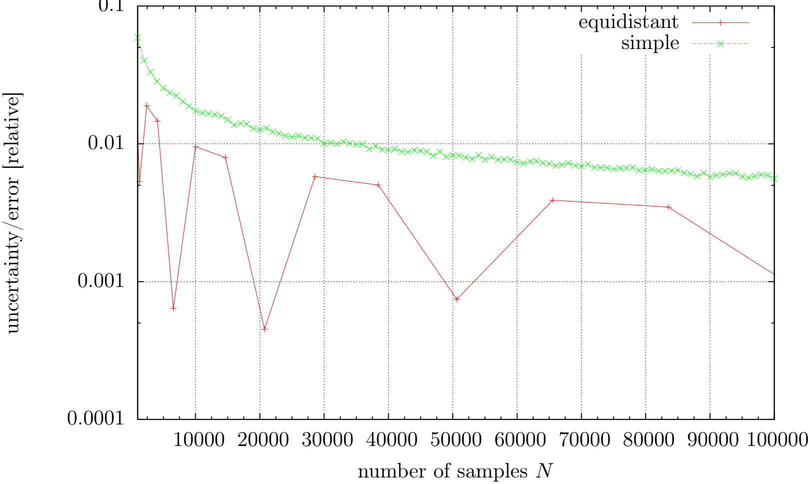


Equidistant sampling vs. simple sampling

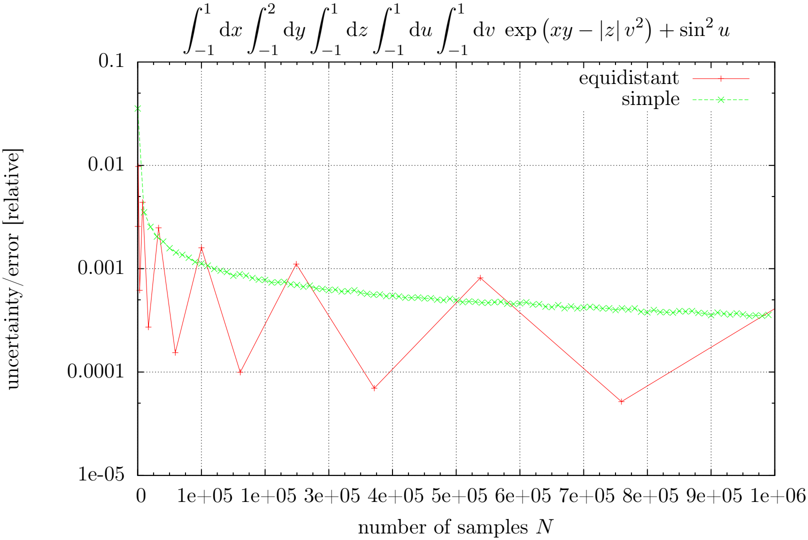


Equidistant sampling vs. simple sampling

$$\int_0^1 \mathrm{d}x \int_0^2 \mathrm{d}y \int_{-1}^3 \mathrm{d}z \int_{-1}^1 \mathrm{d}u \;\; (1+x)^u \; \left(u^2-7\right) \; \sqrt{|y|} \; (z-3)^2$$

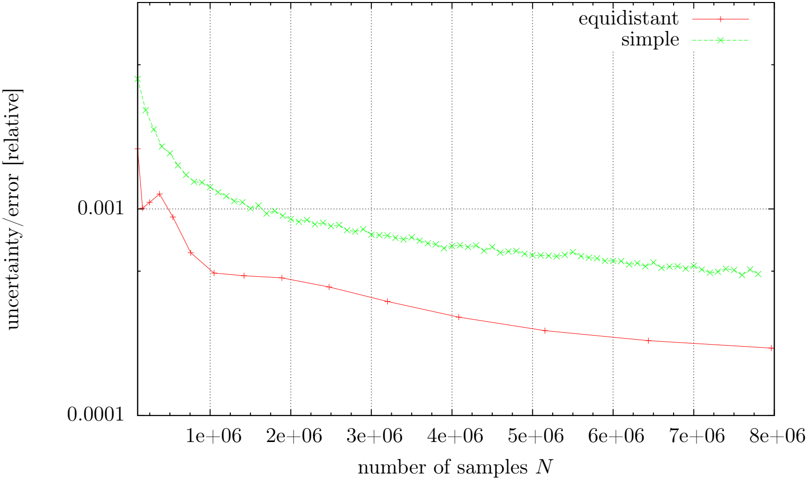


Equidistant sampling vs. simple sampling



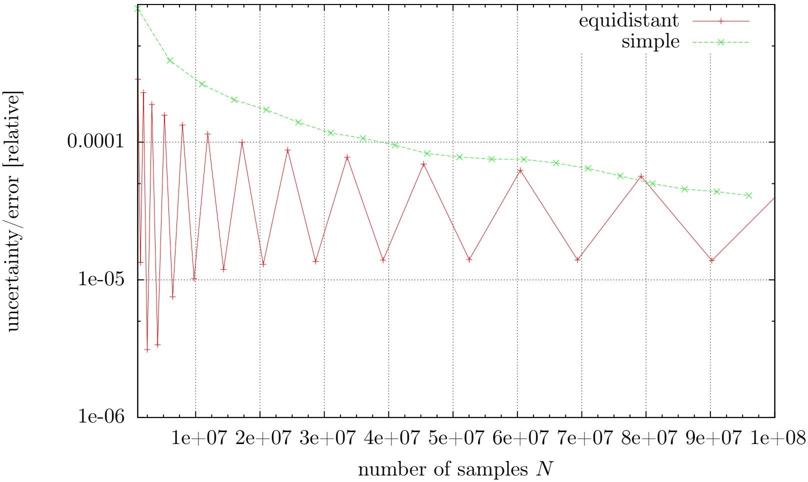
Equidistant sampling vs. simple sampling

$$\int_{-1}^{2} \mathrm{d}x \int_{-1}^{1} \mathrm{d}y \int_{-1}^{2} \mathrm{d}z \int_{-1}^{1} \mathrm{d}u \int_{1}^{3} \mathrm{d}v \;\cos(xv)\,\sin\left(\exp\left(xy+zu^2\right)\right)\,\left(v^3-x^2u\right)\,\ln(v)$$



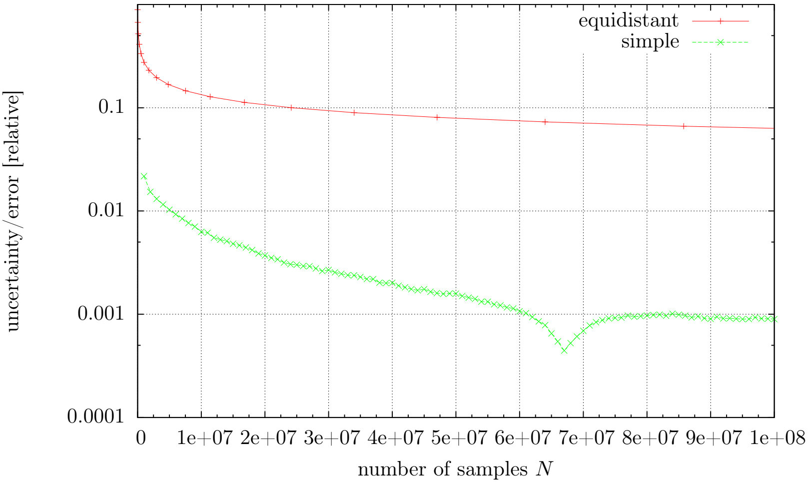
Equidistant sampling vs. simple sampling

$$\int_{-1}^{2} \mathrm{d}x \int_{-1}^{1} \mathrm{d}y \int_{-1}^{2} \mathrm{d}z \int_{-1}^{1} \mathrm{d}u \int_{1}^{3} \mathrm{d}v \;\; \cos(xv)\, \sin(z+1)\, \exp\left(xy + \sqrt{|u|}\right) + \sqrt{|v|}$$

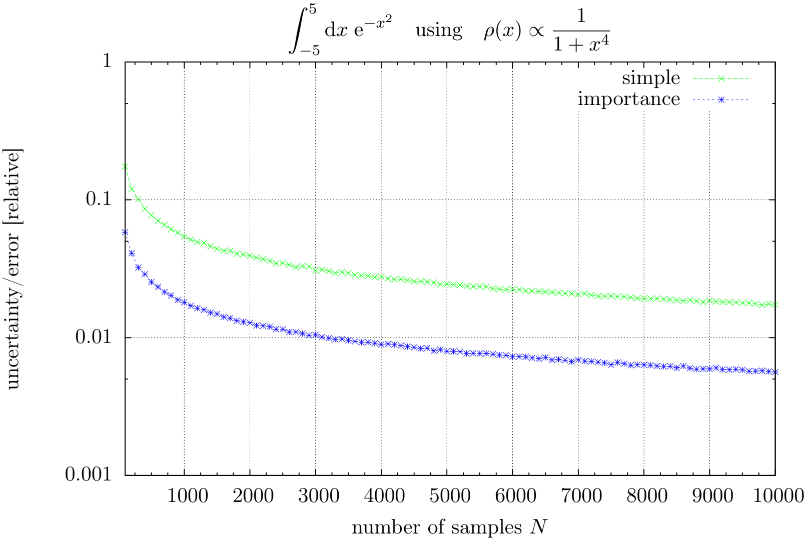


Equidistant sampling vs. simple sampling

$$\int_{-1}^{2} \mathrm{d}x \int_{-1}^{1} \mathrm{d}y \int_{1}^{3} \mathrm{d}z \int_{-1}^{1} \mathrm{d}u \int_{1}^{3} \mathrm{d}v \int_{1}^{2} \mathrm{d}w \;\sin\left(x + y^2 + \sqrt{z}\right)\,\exp\left(x^2 + uv - w^3\right)$$



Simple sampling vs. importance sampling



Conclusion

- Monte Carlo integration is less dependent on the shape of the function and converges faster than equidistant sampling for $D > 4$.
- Importance sampling converges faster than simple sampling, but it is difficult to find a suitable ρ.