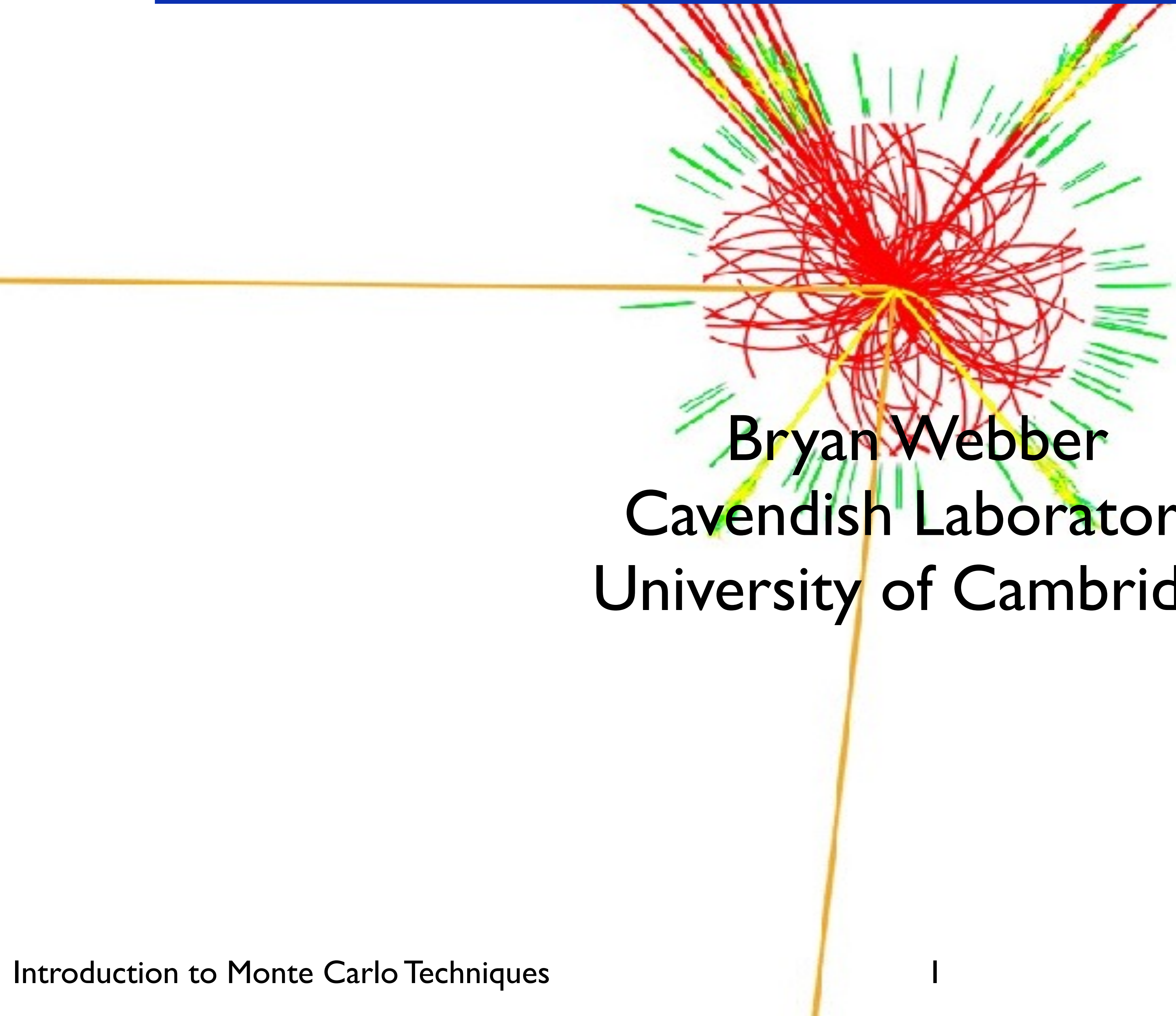
Introduction to Monte Carlo Techniques



Bryan Webber
Cavendish Laboratory
University of Cambridge

Introduction to Monte Carlo

- Lecture 1: The Monte Carlo method
  - theoretical foundations and limitations
  - parton-level event generation
- Lecture 2: Hadron-level event generation
  - parton showering
  - hadronization and underlying event
  - sample of results

Why Monte Carlo?

- Something to do with gambling?
- Not a place but a method …

Monte Carlo Event Generation

Monte Carlo Event Generation

- Aim is to produce simulated (particle-level) datasets like those from real collider events
  - i.e. lists of particle identities, momenta, ...
  - simulate quantum effects by (pseudo)random numbers
- Essential for:
  - Designing new experiments and data analyses
  - Correcting for detector and selection effects
  - Testing the SM and measuring its parameters
  - Estimating new signals and their backgrounds

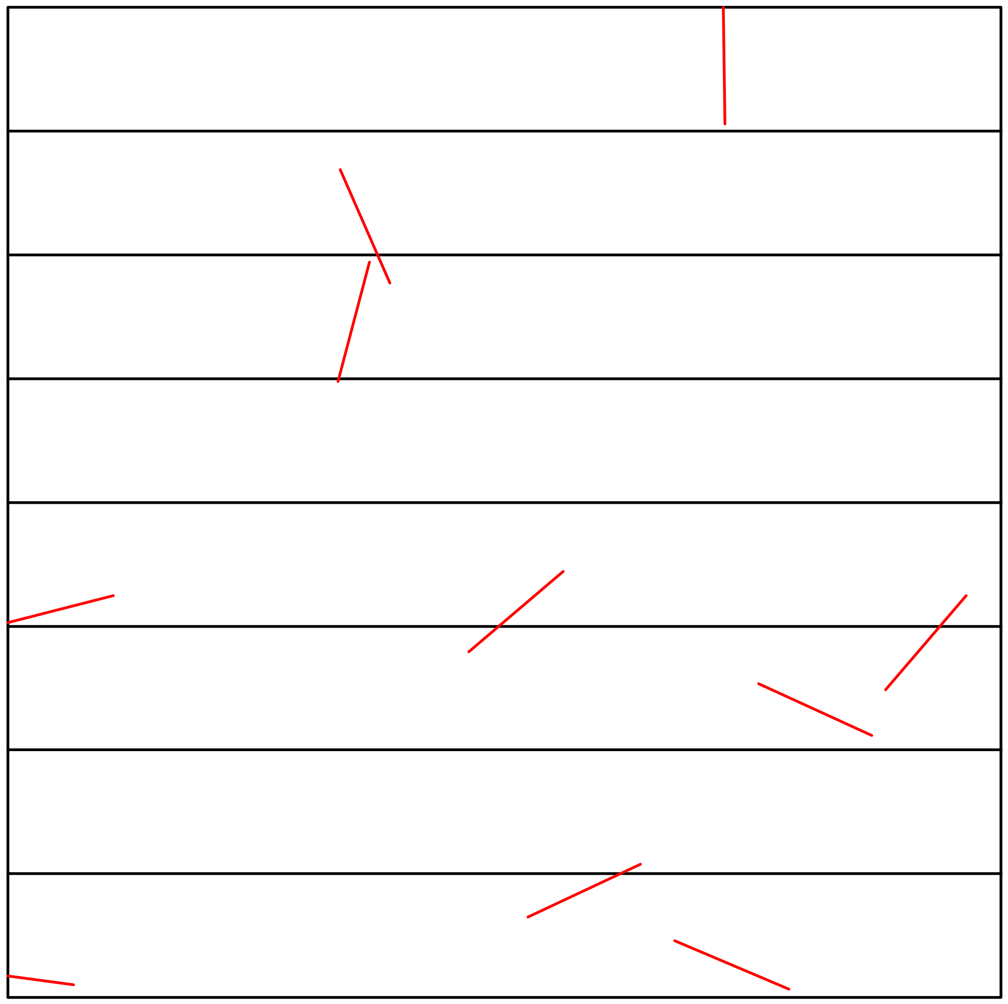
References

- A. Buckley et al., "General-purpose event generators for LHC physics", Phys.Rept. 504 (2011) 145 (MCNET-11-01, arXiv: 1101.2599)
- M.H. Seymour & M. Marx, "Monte Carlo Event Generators", MCNET-13-05, arXiv: 1304.6677
- A. Siódmok, "LHC event generation with general-purpose Monte Carlo tools", Acta Phys. Polon. B44 (2013)1587

Monte Carlo Method

Buffon's needle

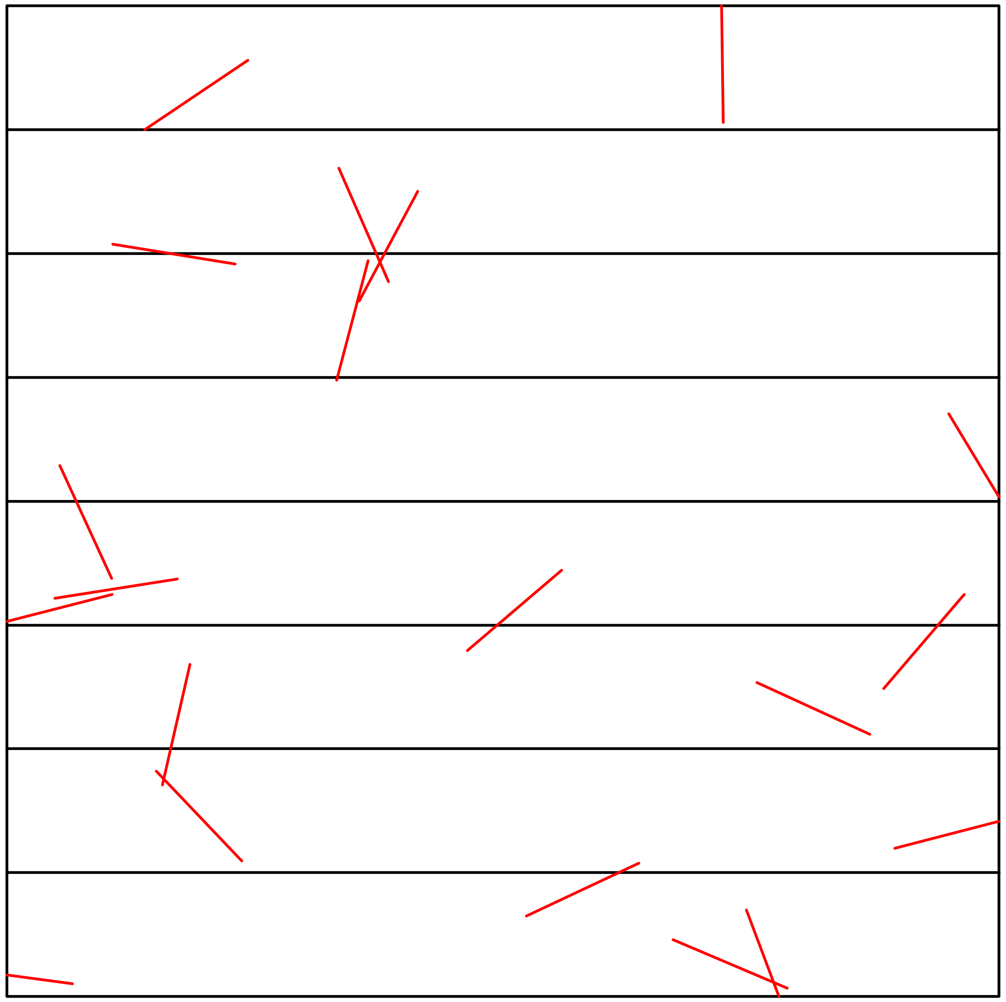
G-L Leclerc, Comte de Buffon, 1707-1788



2x10/5 = 4.000

Buffon’s needle

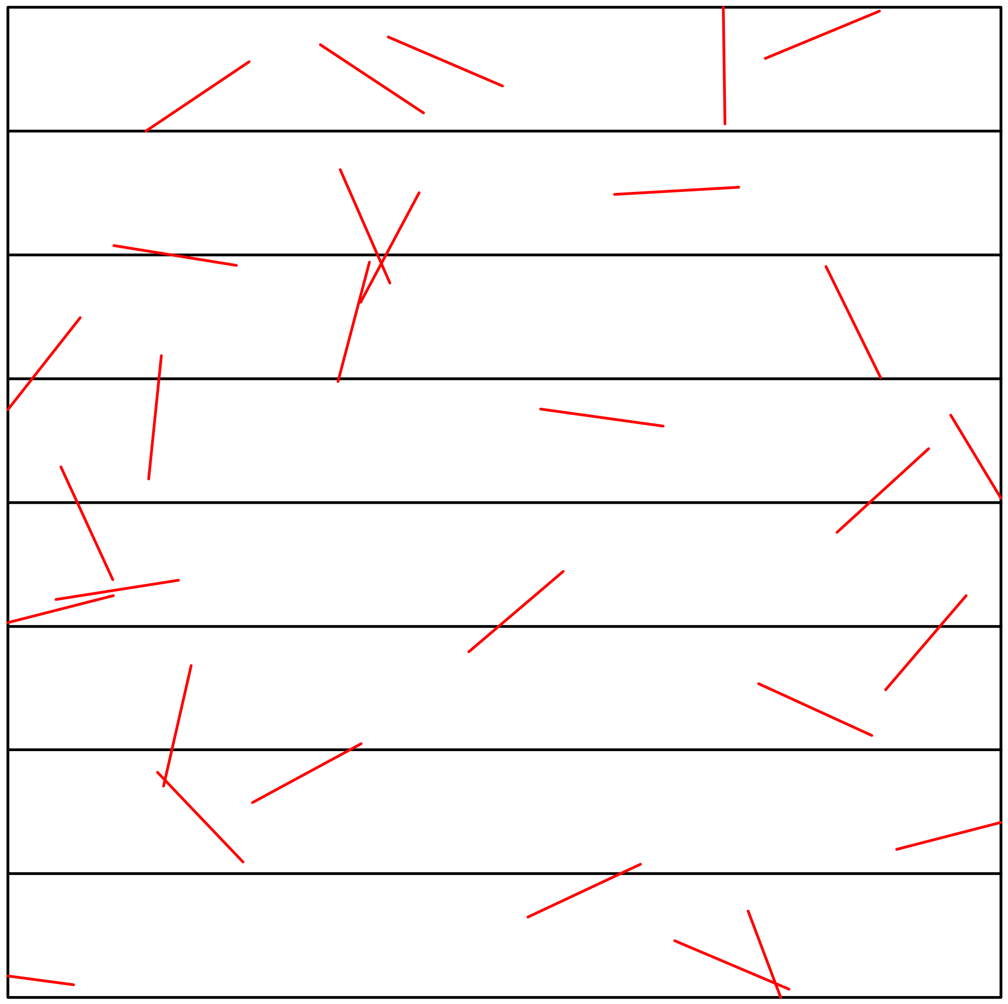
G-L Leclerc, Comte de Buffon, 1707-1788



2x20/12 = 3.333

Buffon's needle

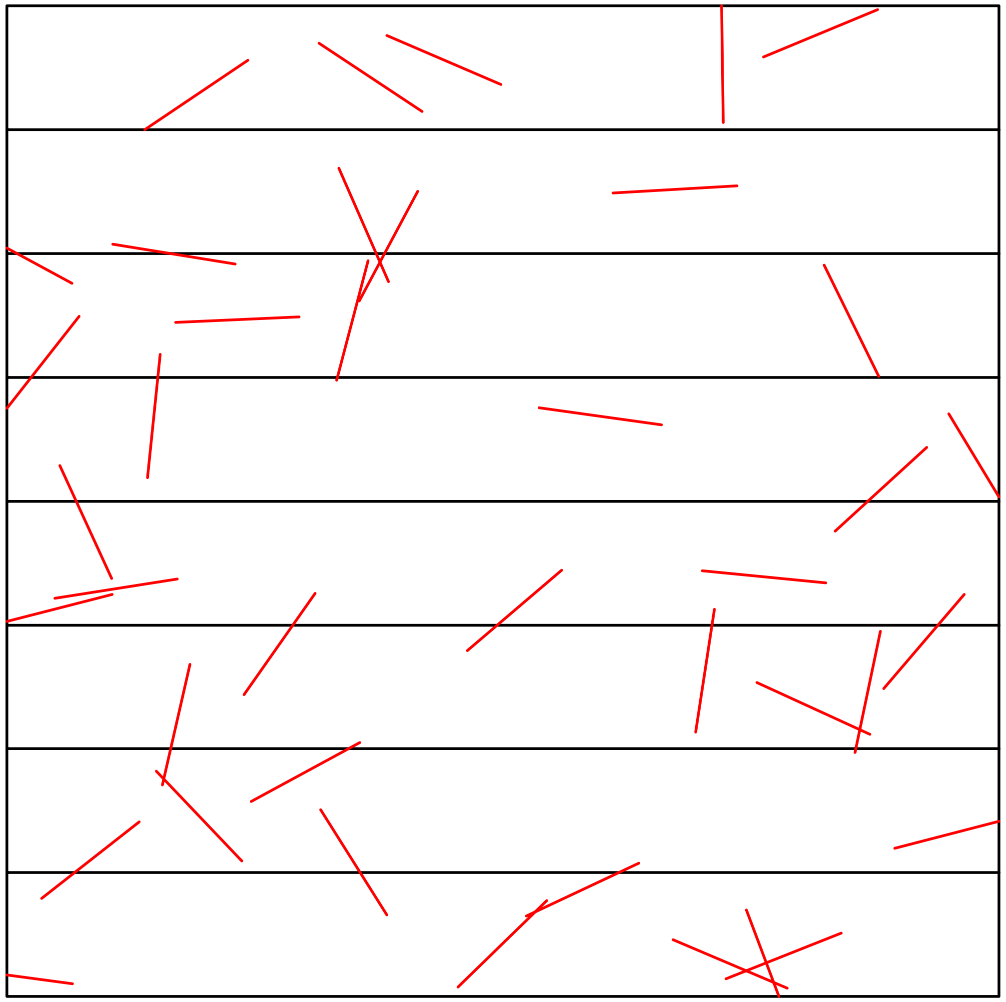
G-L Leclerc, Comte de Buffon, 1707-1788



2x30/16 = 3.750

Buffon’s needle

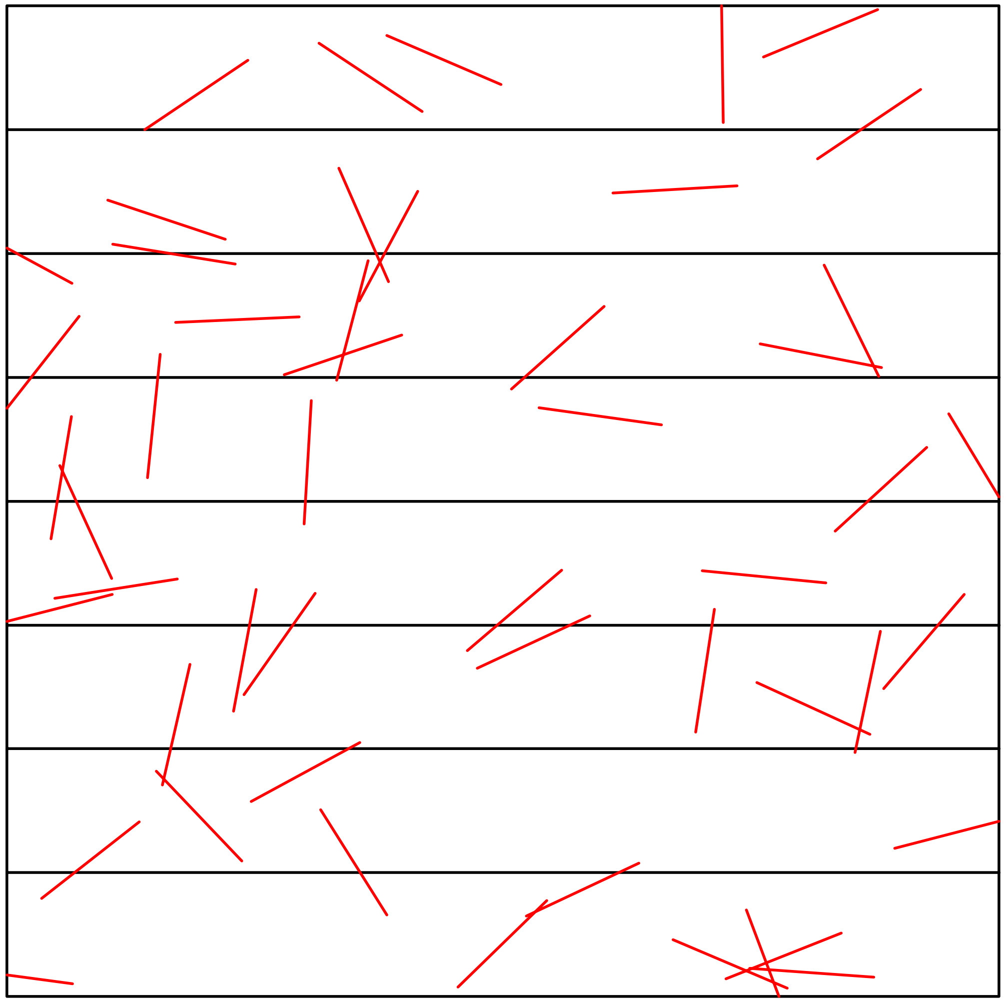
G-L Leclerc, Comte de Buffon, 1707-1788



2x40/22 = 3.636

Buffon’s needle

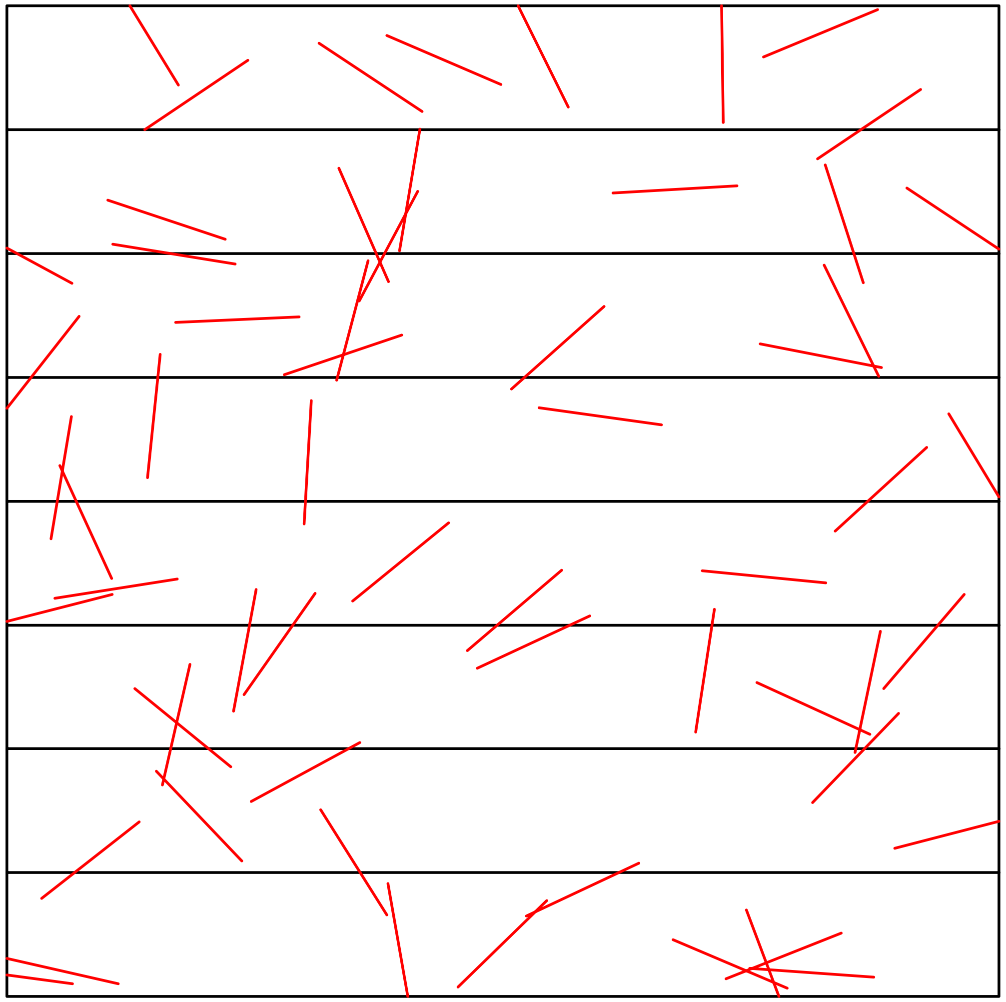
G-L Leclerc, Comte de Buffon, 1707-1788



2x50/28 = 3.571

Buffon’s needle

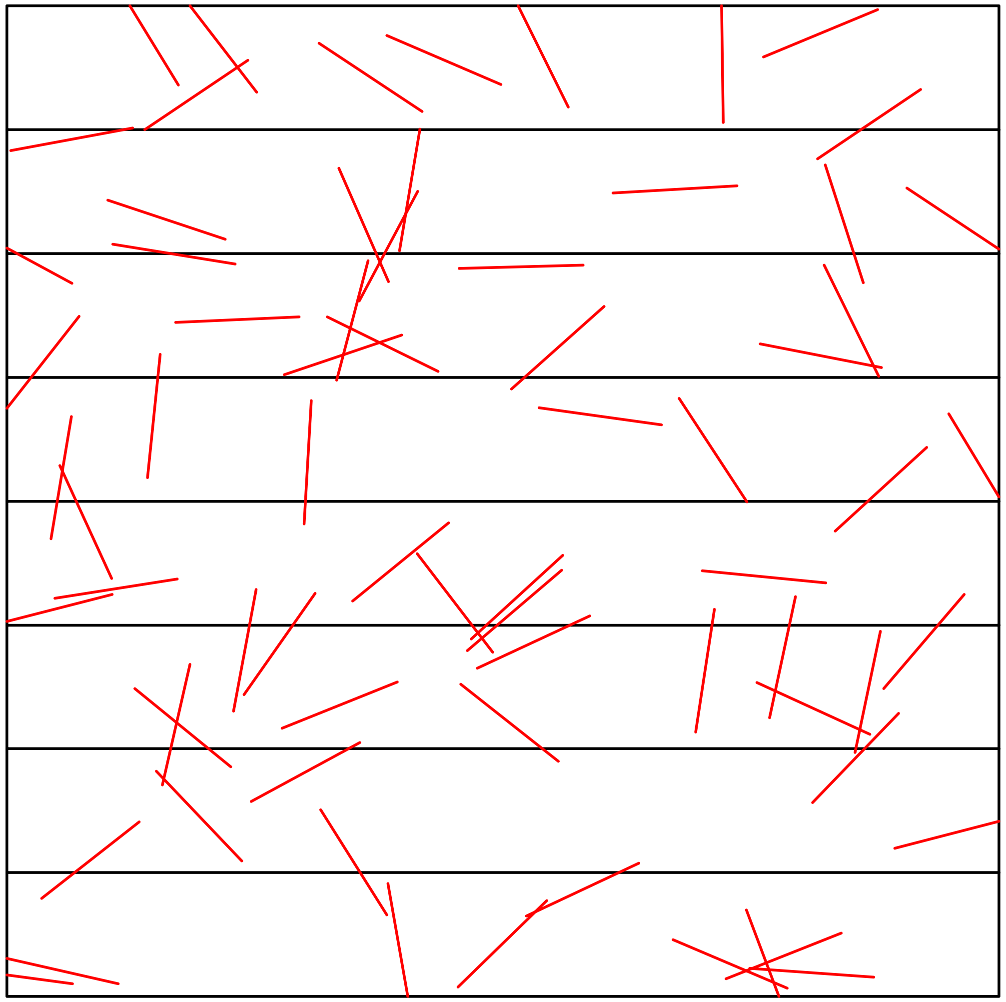
G-L Leclerc, Comte de Buffon, 1707-1788



2x60/34 = 3.529

Buffon's needle

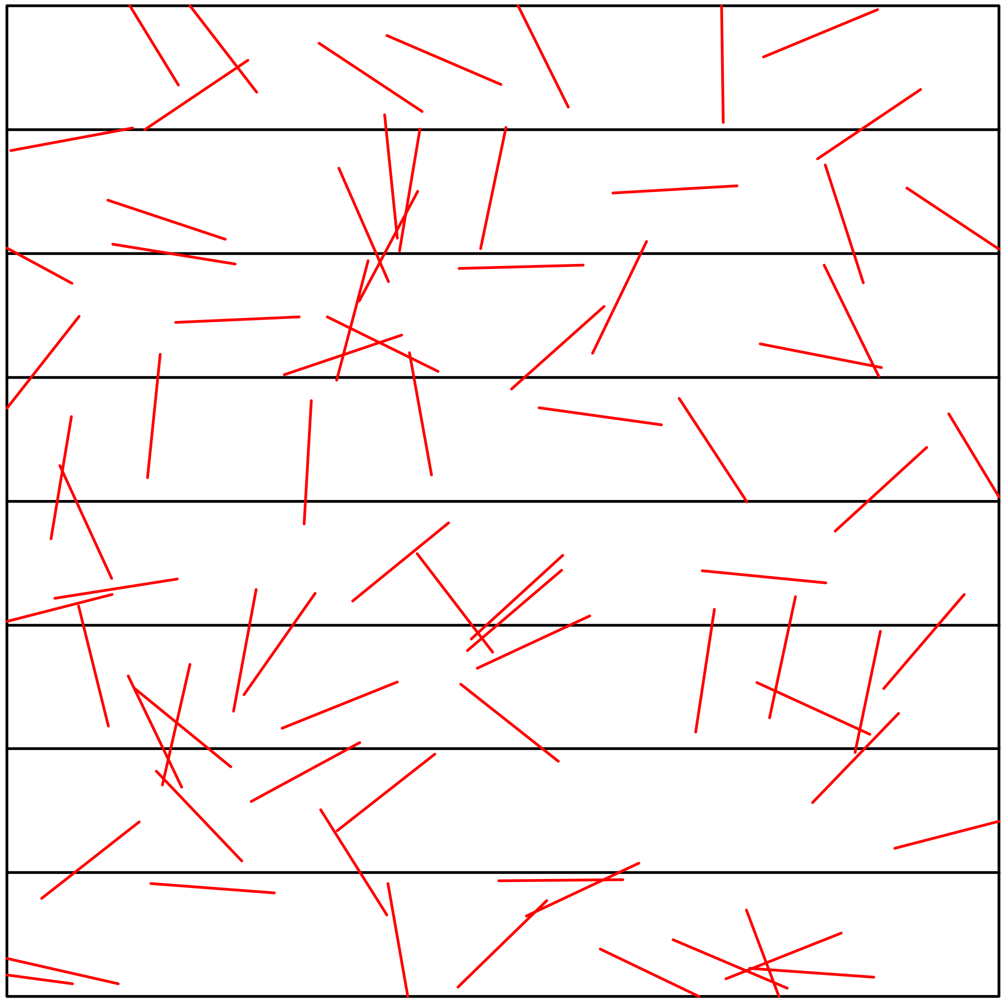
G-L Leclerc, Comte de Buffon, 1707-1788



$$2\text{x}70/40 = 3.500$$

Buffon’s needle

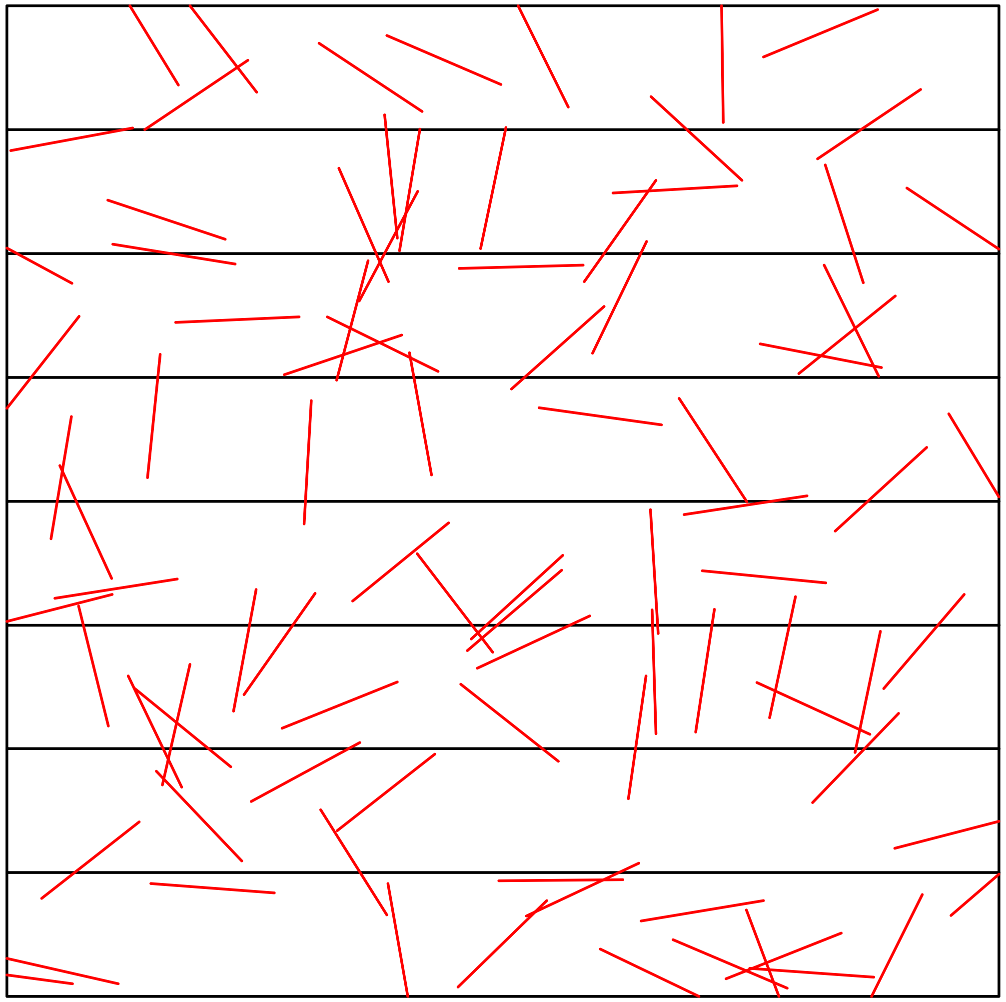
G-L Leclerc, Comte de Buffon, 1707-1788



2x80/47 = 3.404

Buffon's needle

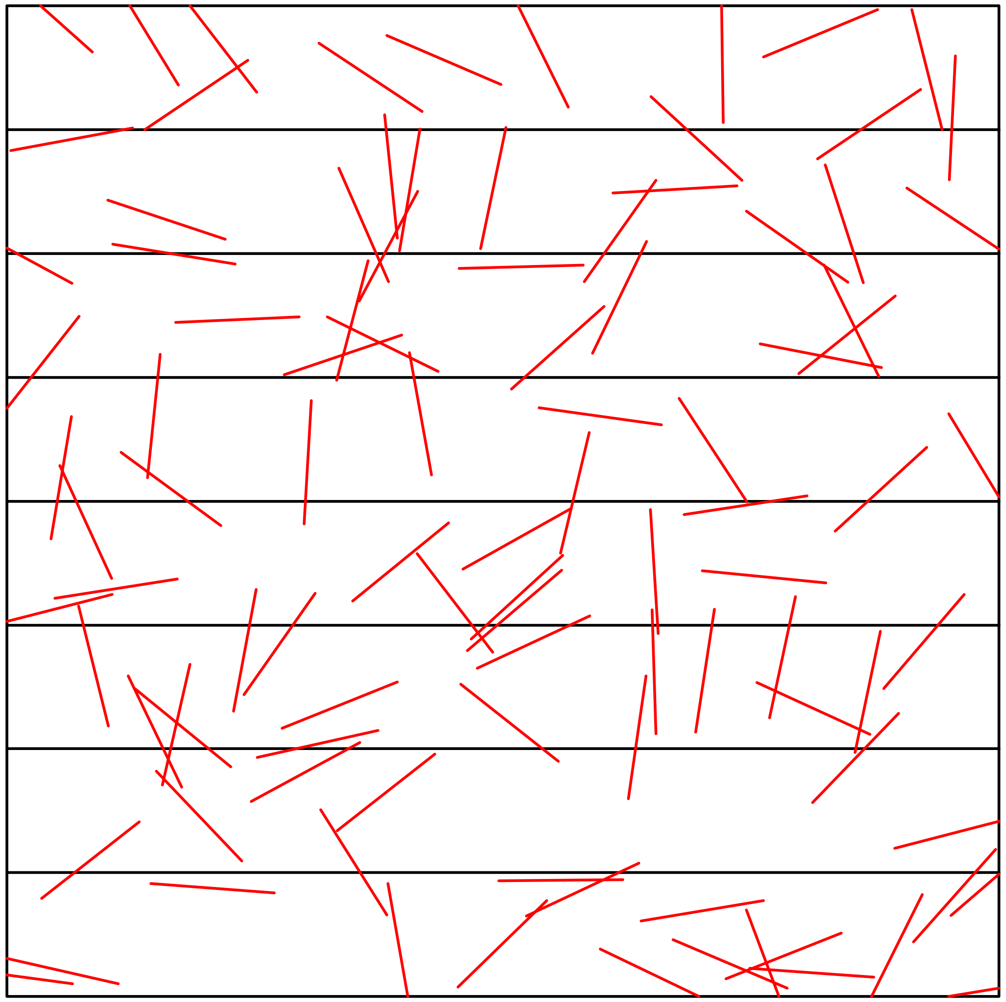
G-L Leclerc, Comte de Buffon, 1707-1788



2x90/55 = 3.273

Buffon’s needle

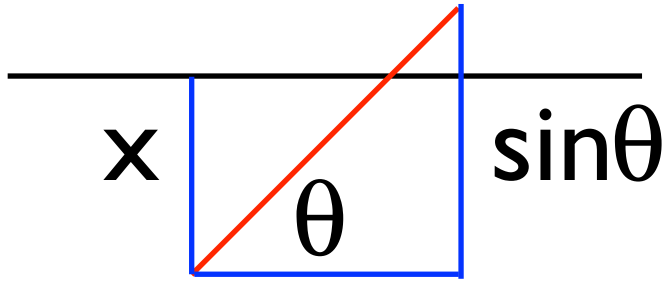
G-L Leclerc, Comte de Buffon, 1707-1788

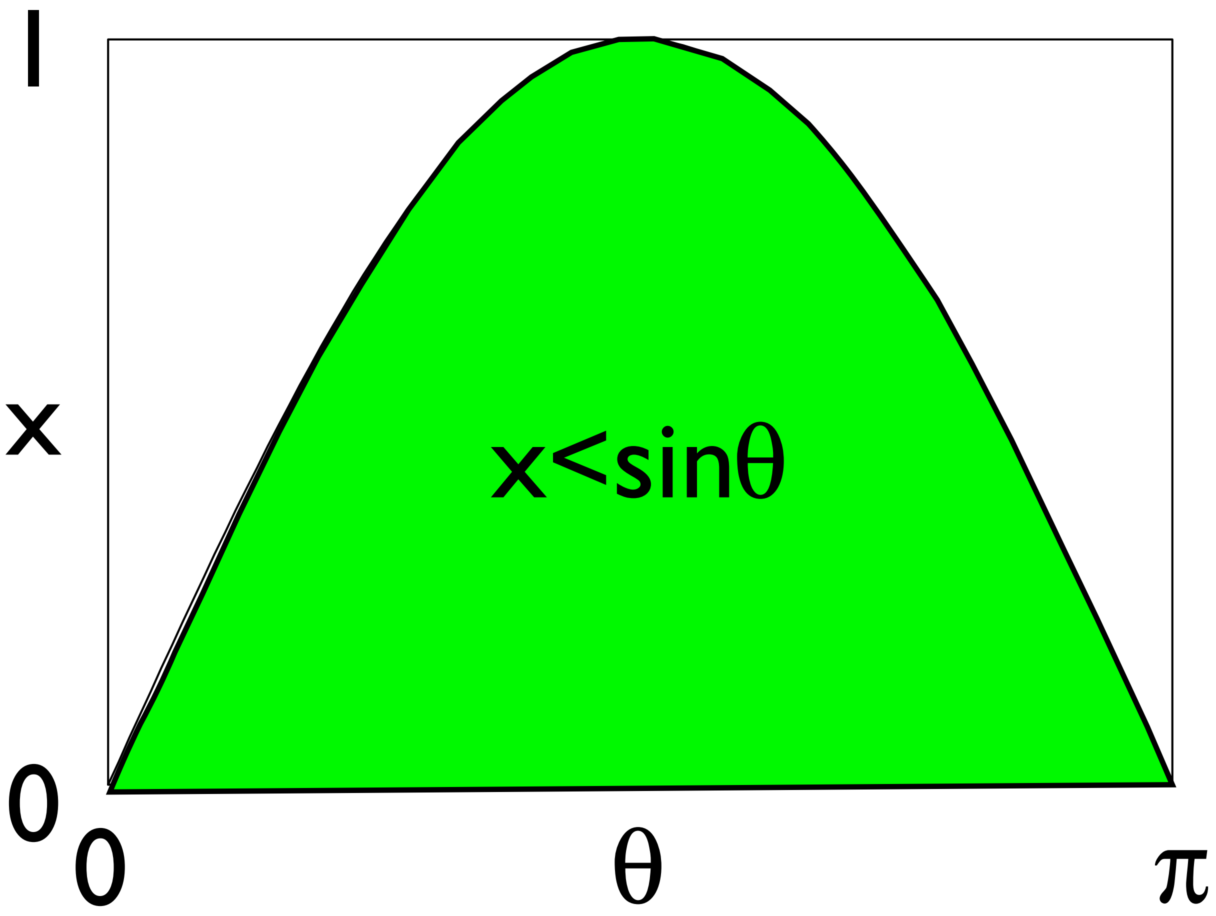


2x100/63 = 3.175

Buffon's needle

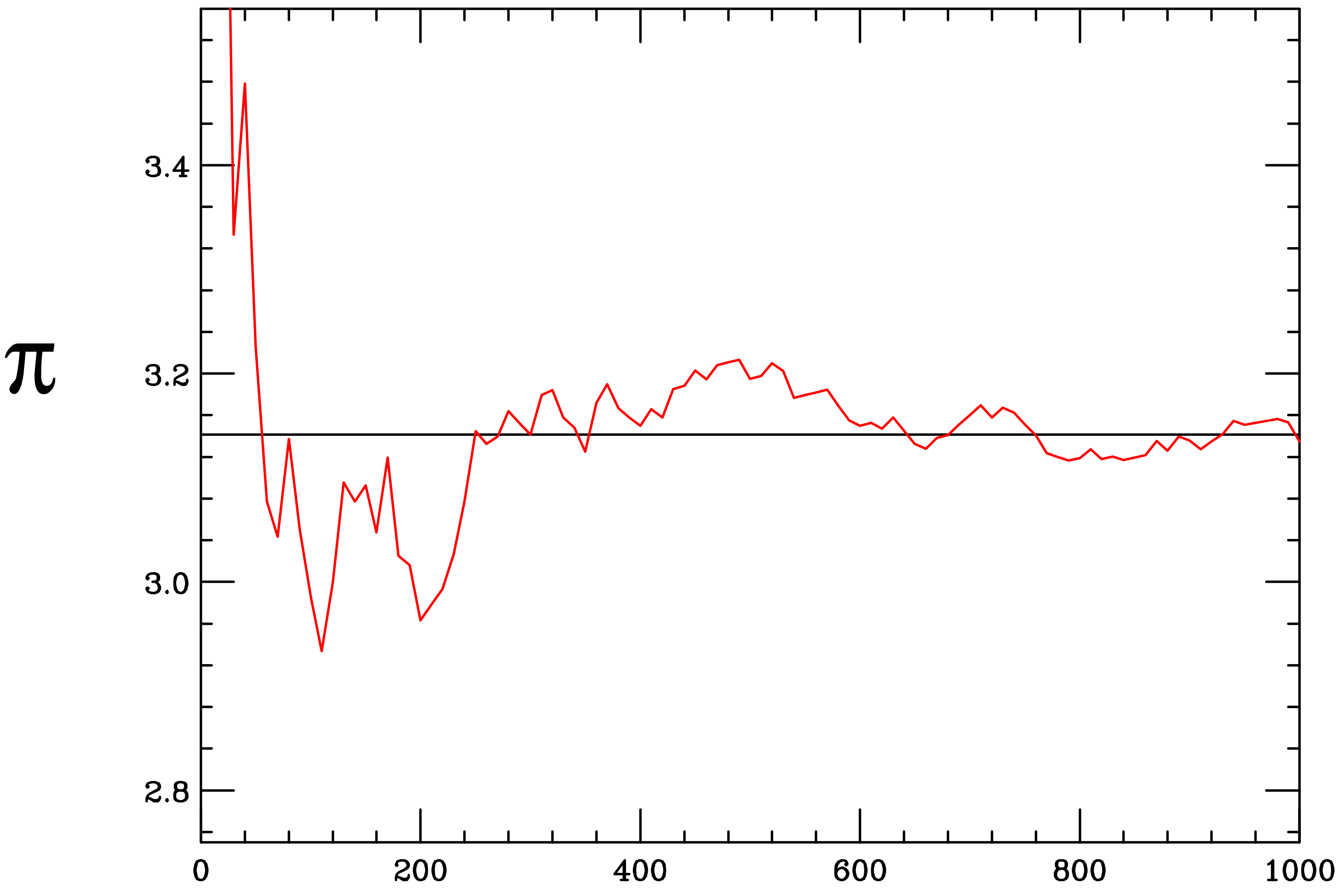
Events (needle drops) are represented by random points in (θ,x) phase space





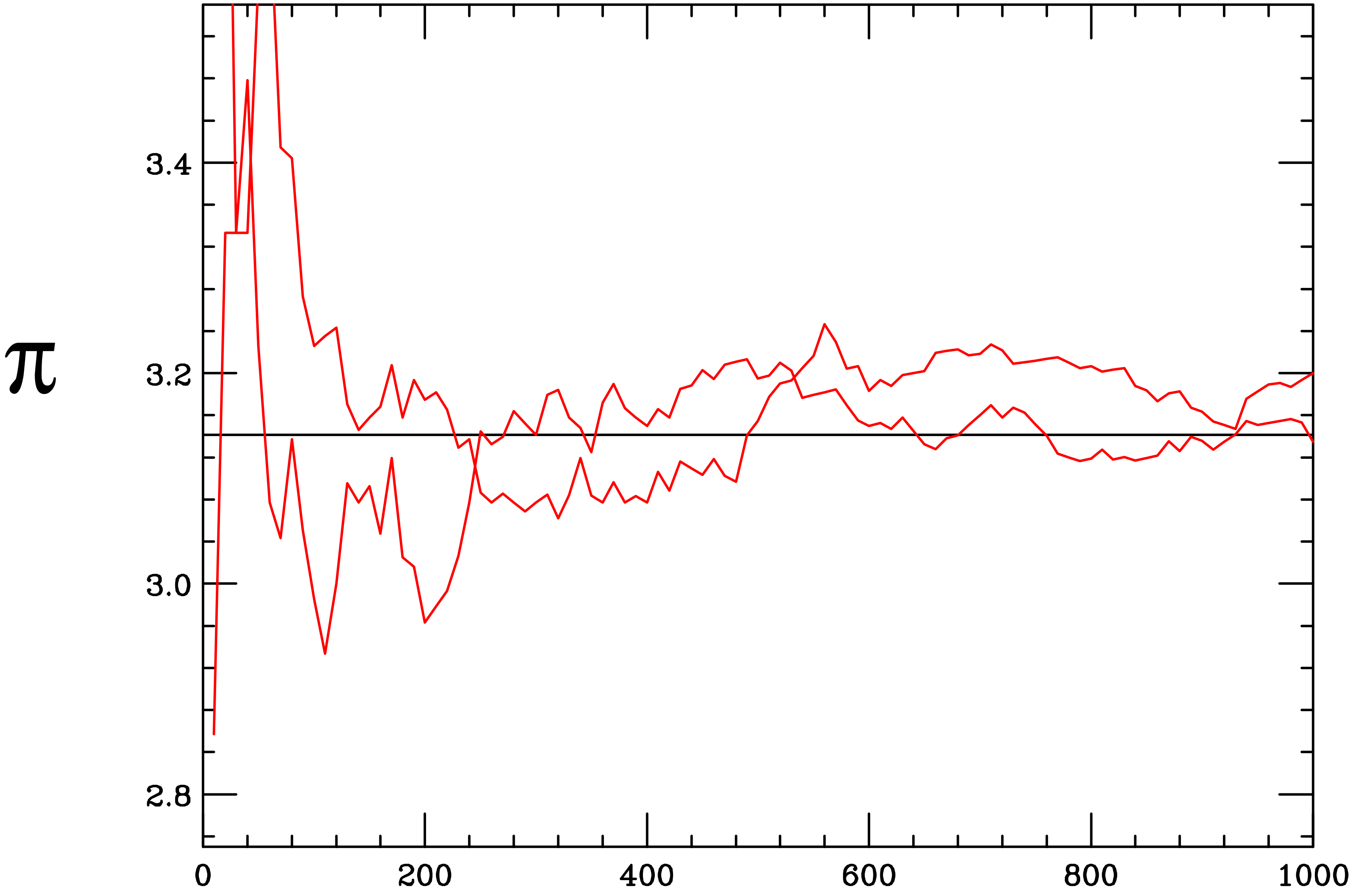
P(x<sinθ)=2/π

Buffon’s needle



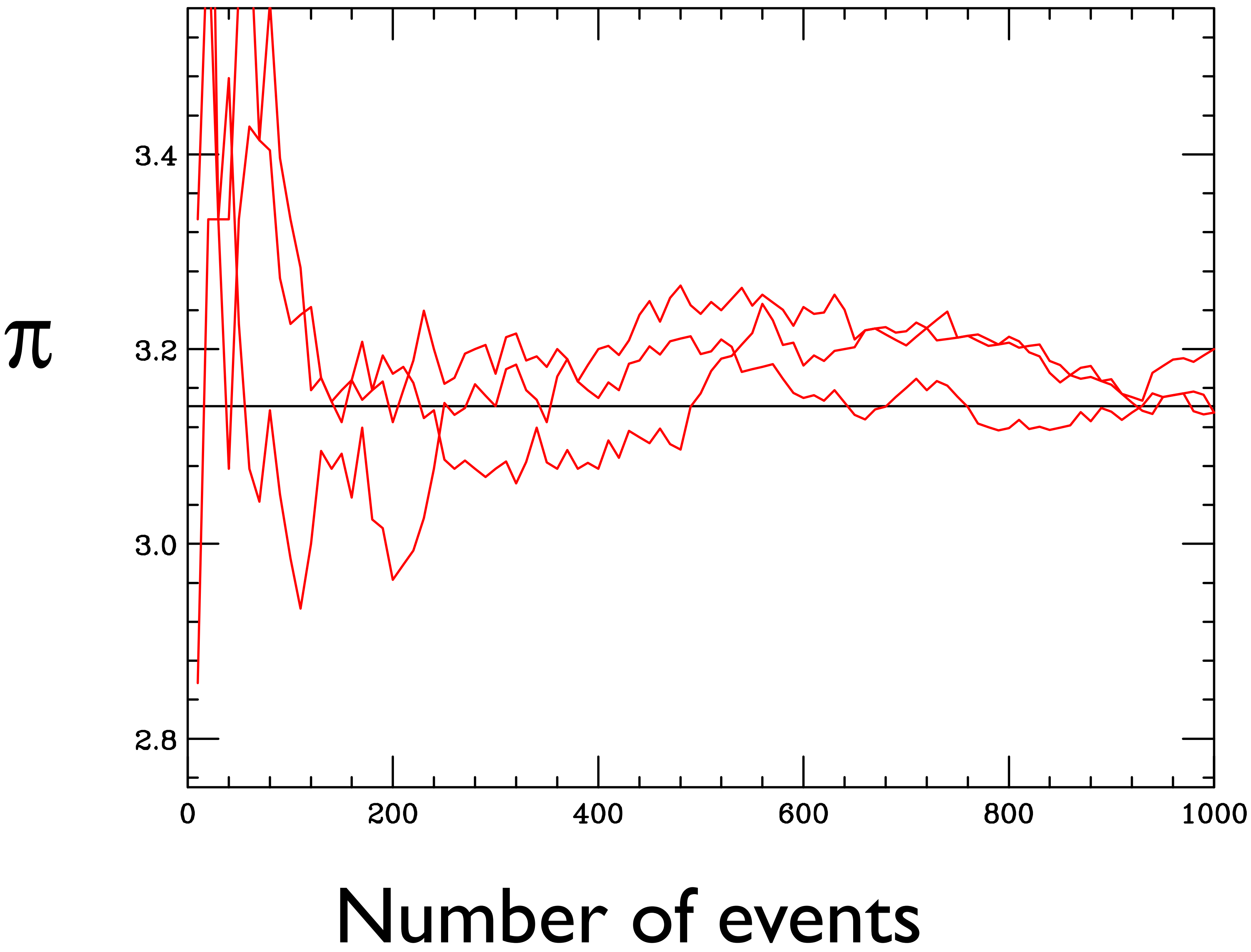
Number of events

Buffon's needle

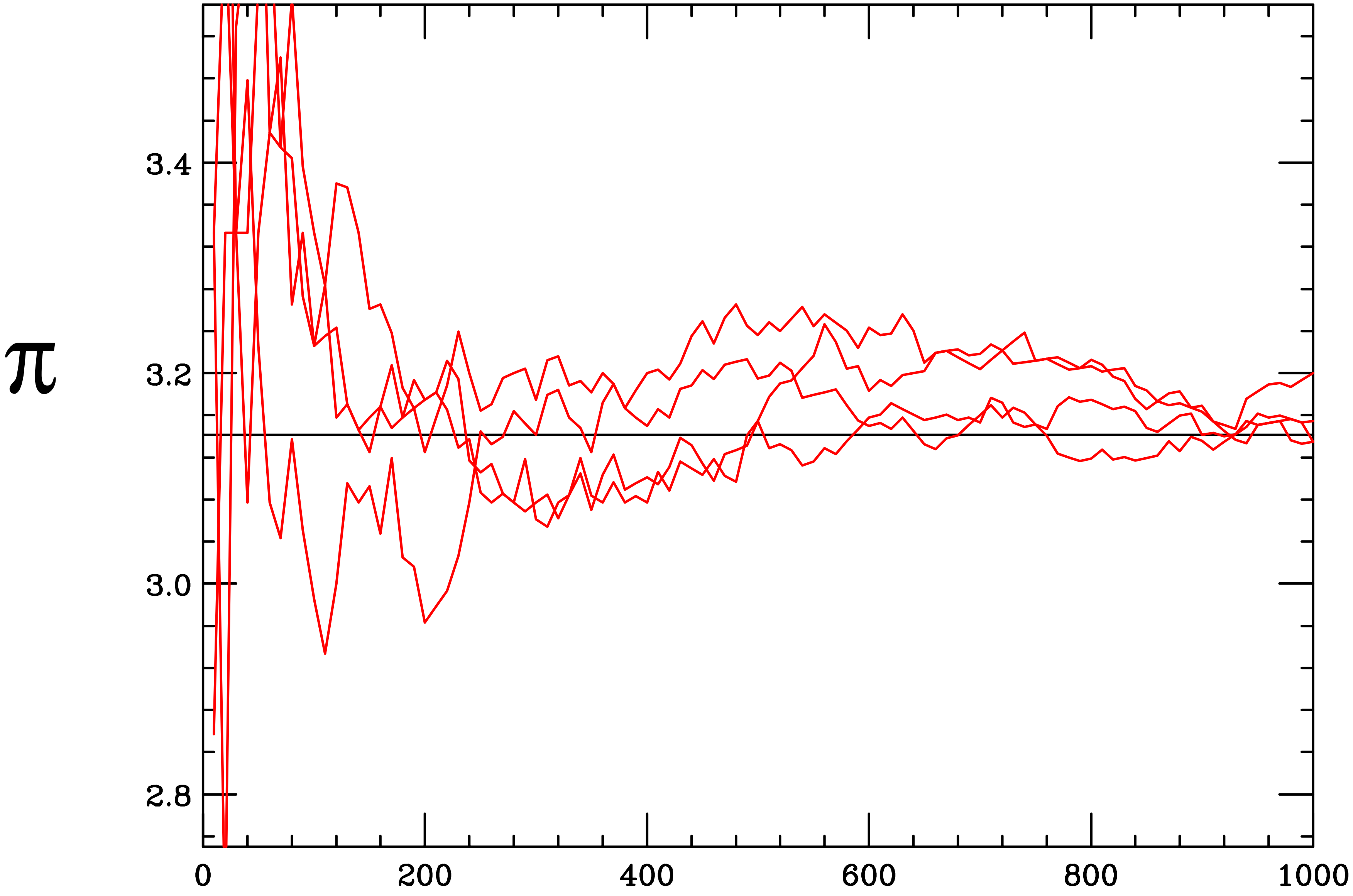


Number of events

Buffon’s needle



Buffon's needle



Number of events

Some statistics

- Expected value of a discrete random variable x = probability-weighted sum over possible outcomes = expected mean of a large number of independent trials
  - $E[x] = x_1p_1 + x_2p_2 + x_3p_3 + \ldots$
  - Here $x_1=1$ (needle on line, $p_1=2/\pi$) or else $x_2= 0$ (needle off line). Hence $E[x]=2/\pi$
- Variance = Mean square deviation
  - $Var[x] = E[(x-E[x])^2] = E[x^2]-(E[x])^2$
  - Here $Var[x] = p_1-p_1^2 = 2/\pi(1-2/\pi)$
- (RMS) Standard deviation $\sigma_x = \sqrt{Var[x]} = 0.481$ here

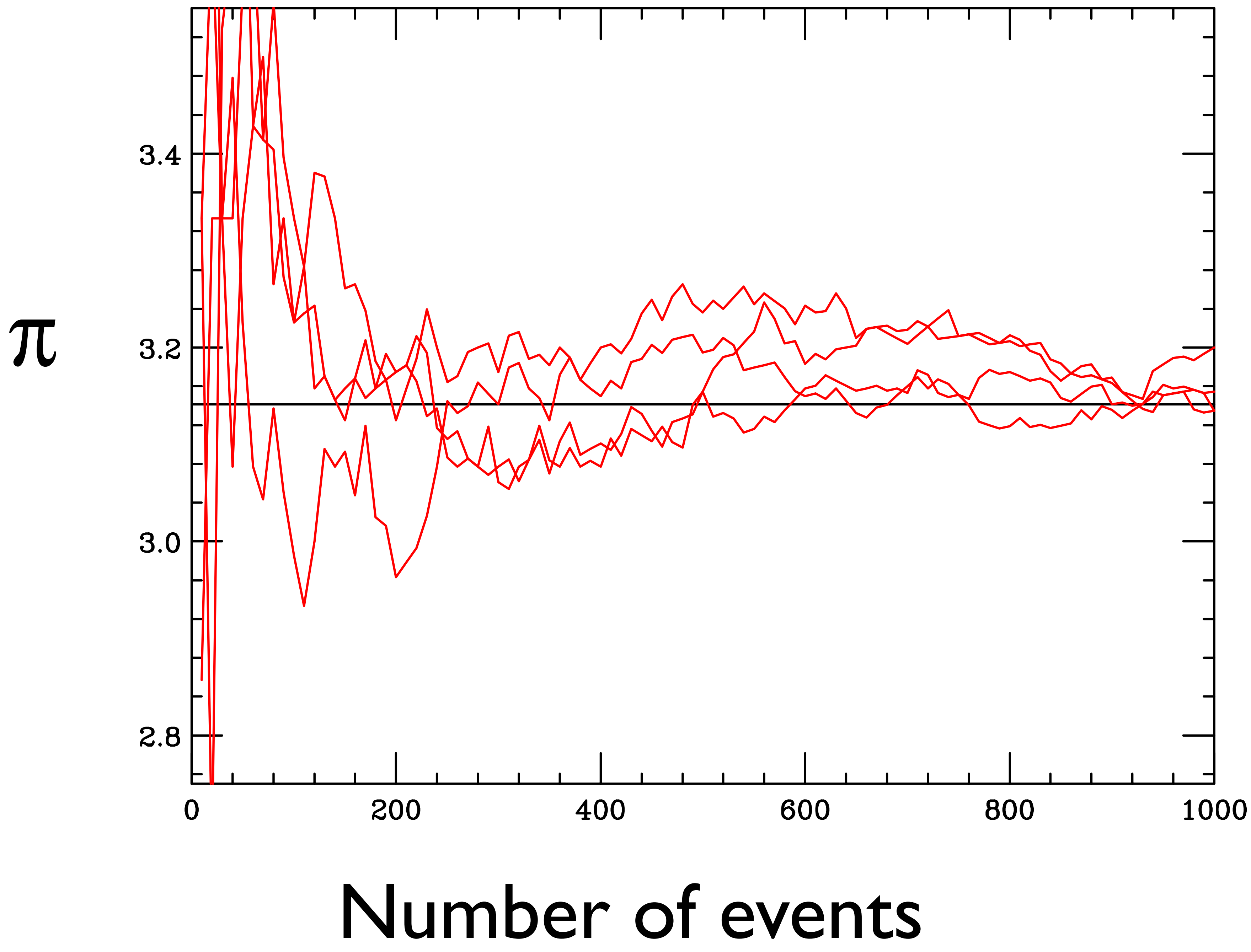
Statistics (cont'd)

- Variances for uncorrelated random variables are additive
  - $\mathrm{Var}[x_1+x_2] = \mathrm{Var}[x_1] + \mathrm{Var}[x_2] + 2\cancel{(E[x_1x_2]-E[x_1]E[x_2])}$
- For N identical independent trials and $S_N=x_1+\ldots+x_N$,
  - $E[S_N] = N\,E[x]$, $\mathrm{Var}[S_N] = N\,\mathrm{Var}[x]$, $\sigma_{SN} = \sqrt{N}\,\sigma_x$
- Here, for N needles, $E[S_N]= 2N/\pi$, so $\sigma_\pi/\pi = \sigma_{SN}/E[S_N] = \sqrt{N}\,\sigma_x/(2N/\pi)$, i.e. standard deviation in estimate of π is
  - $\sigma_\pi = \pi^2\sigma_x/(2\sqrt{N}) = 2.37/\sqrt{N}$
- $2N/S_N = \pi \pm \underbrace{2.37/\sqrt{N}}_{1\sigma}$

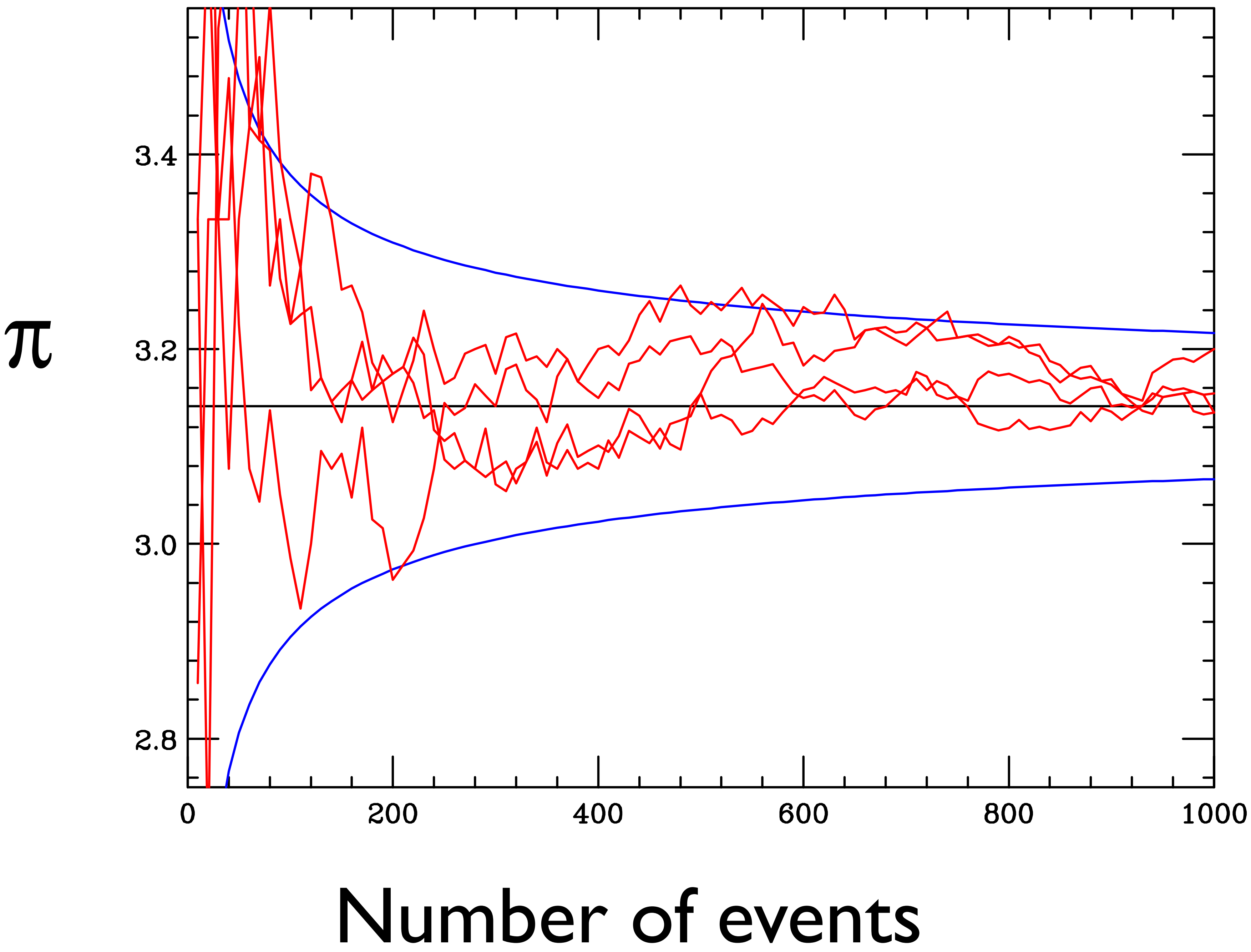
Central limit theorem:

$$P(< 1\sigma) = 68\%$$

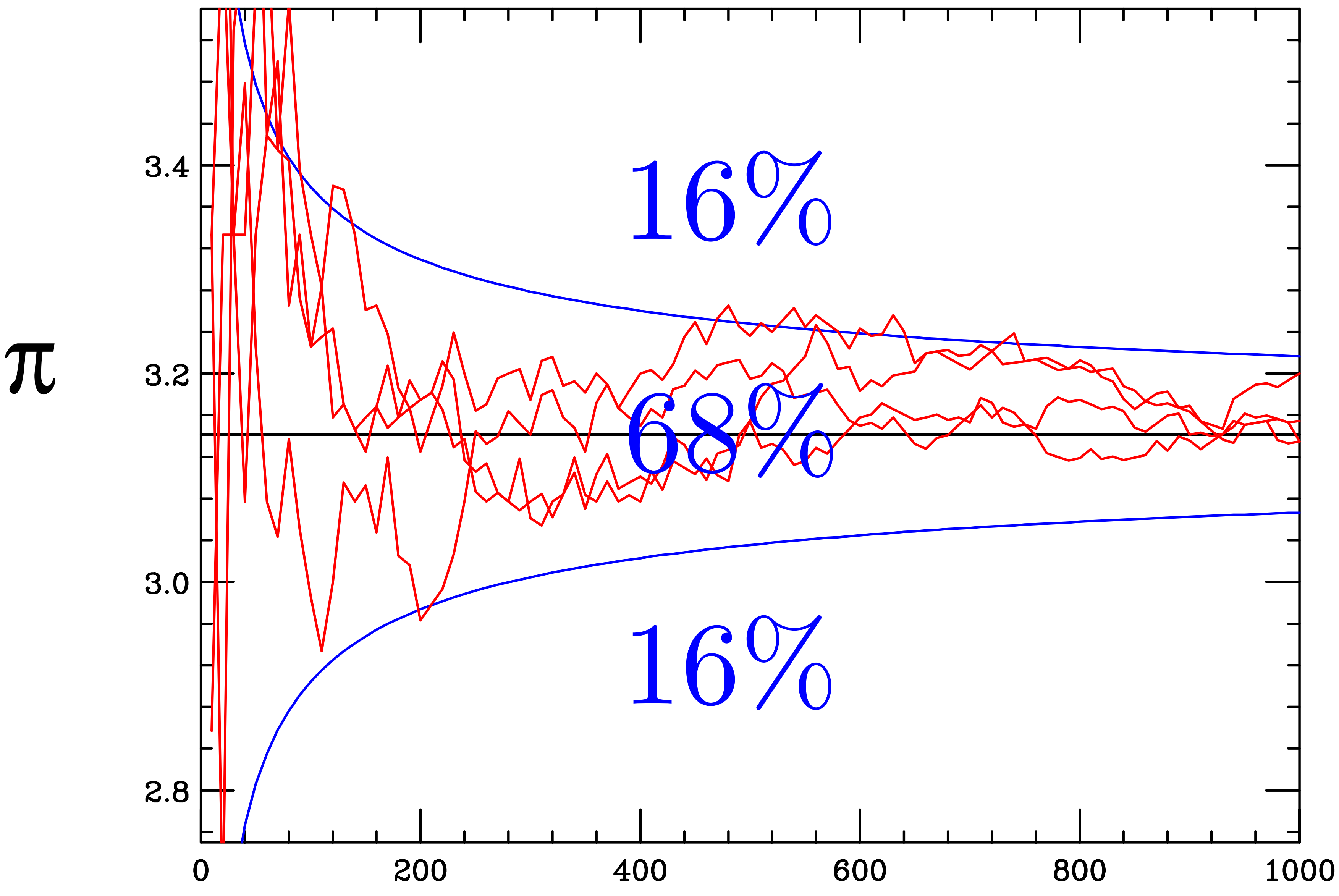
Buffon’s needle



Buffon's needle



Buffon’s needle



Number of events

Monte Carlo Integration

- Basis of all Monte Carlo methods:

$$I = \int_a^b f(x)\,dx \approx \frac{1}{N}\sum_{i=1}^{N} \underbrace{(b-a)f(x_i)}_{\text{weight } w_i} \equiv I_N$$

where x_i are randomly (uniformly) distributed on [a,b].

- Then $I = \lim_{N\to\infty} I_N = E[w]\ ,\quad \sigma_I = \sqrt{\mathrm{Var}[w]/N}$

where $\mathrm{Var}[w] = E[(w - E[w])^2] = E[w^2] - (E[w])^2$

$$= (b-a)\int_a^b [f(x)]^2 dx - [\int_a^b f(x)\,dx]^2 \equiv V$$

$$\boxed{I = I_N \pm \sqrt{V/N}}$$

Central limit theorem:
$P(< 1\sigma) = 68\%$

Convergence

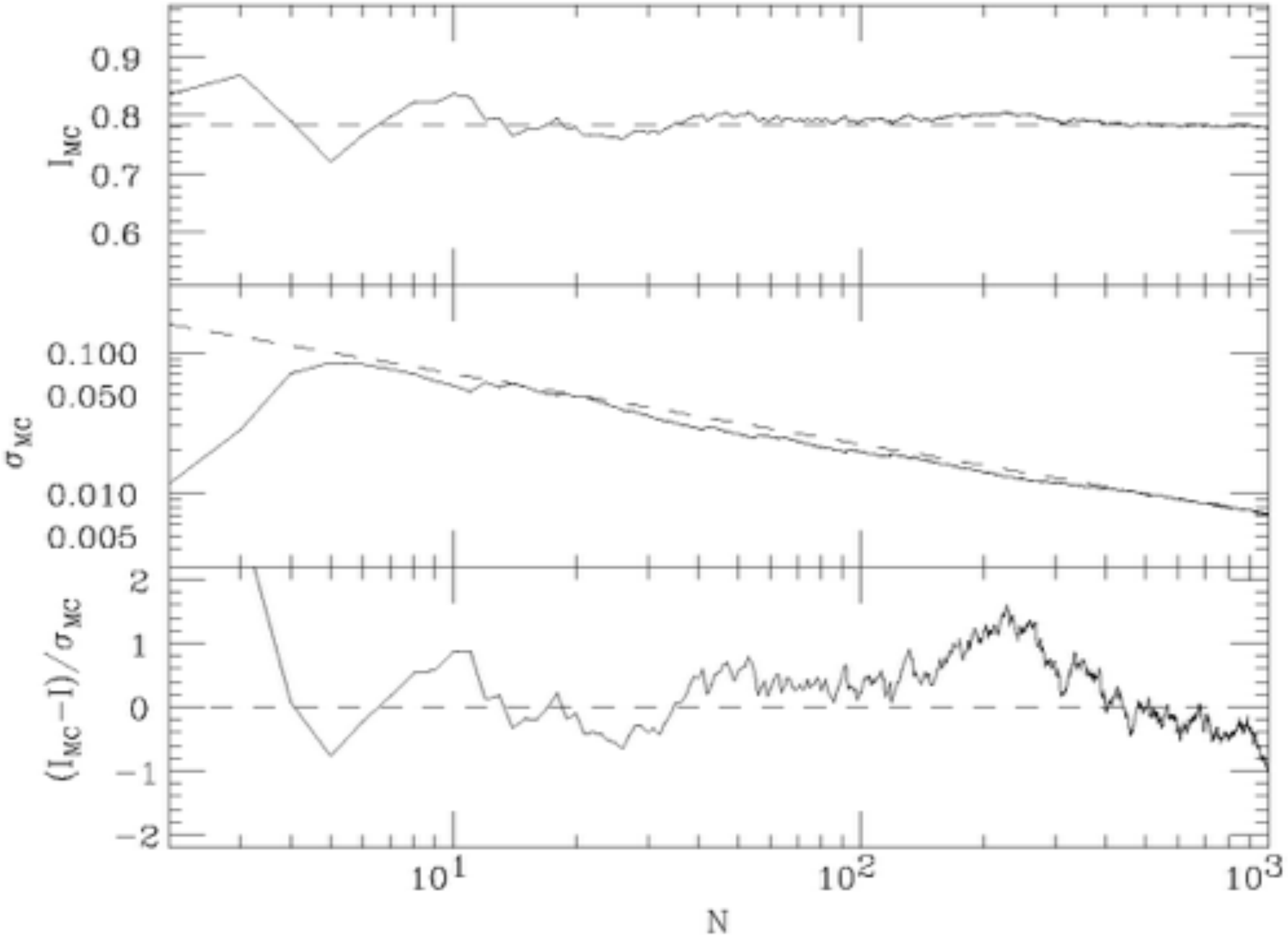
- Monte Carlo integrals governed by Central Limit Theorem: error $\propto 1/\sqrt{N}$

c.f. trapezium rule $\propto 1/N^2$

Simpson's rule $\propto 1/N^4$

$\Bigg\}$ only if derivatives exist and are finite, e.g.

$\sqrt{1-x^2} \sim 1/N^{3/2}$



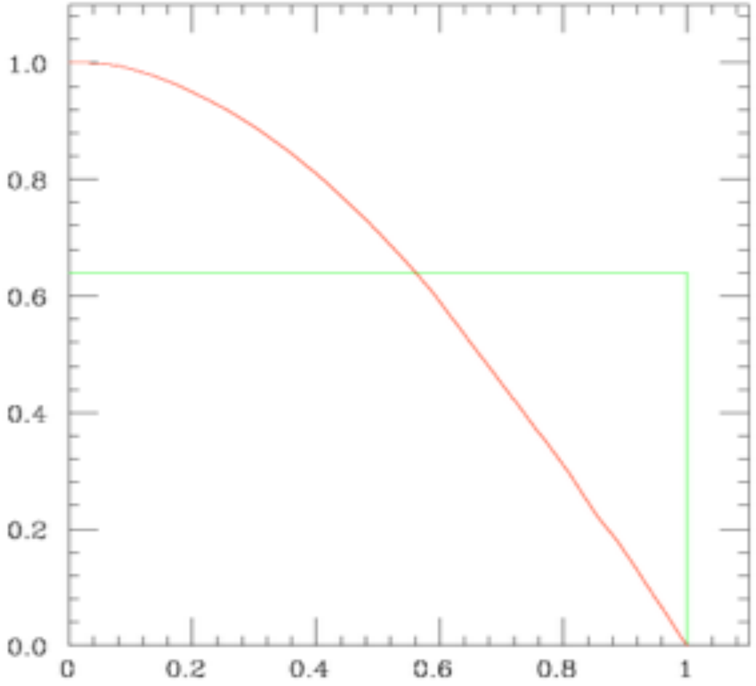
$$I = \int_0^1 \sqrt{1-x^2}\,dx = \frac{\pi}{4} = 0.785$$

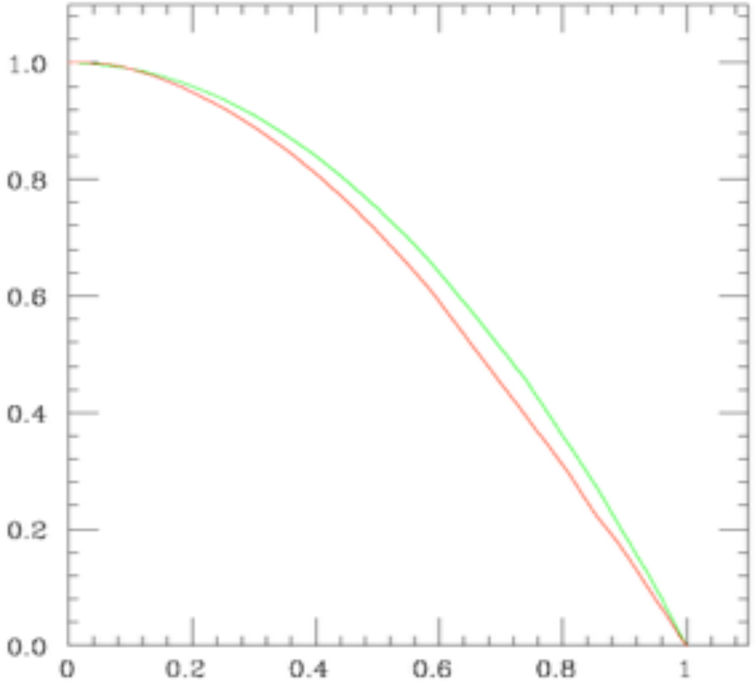
$$\sqrt{V} = \sqrt{\frac{2}{3} - \frac{\pi^2}{16}} = 0.223$$

$$I = 0.785 \pm \frac{0.223}{\sqrt{N}}$$

Importance Sampling

- Convergence is improved by putting more points in regions where integrand is largest
- Corresponds to a Jacobian transformation
- Variance is reduced (weights "flattened")



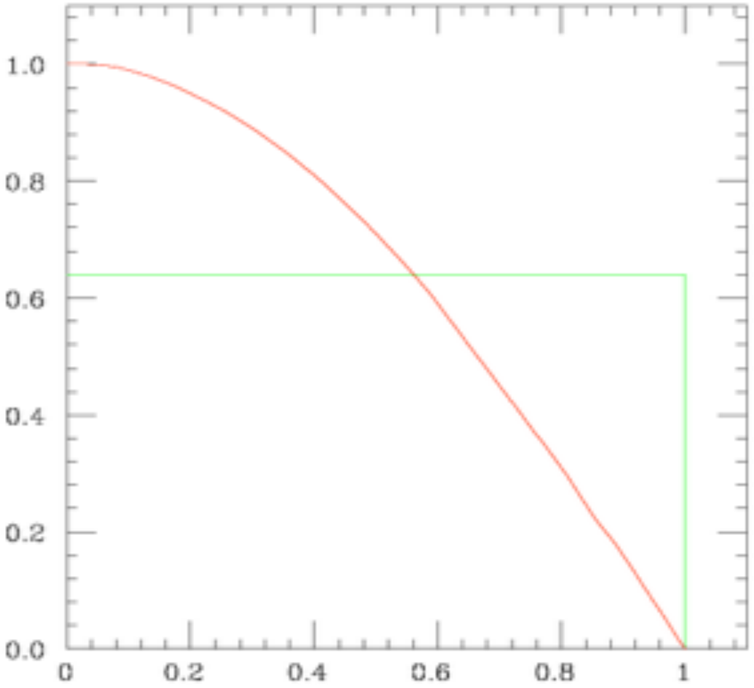


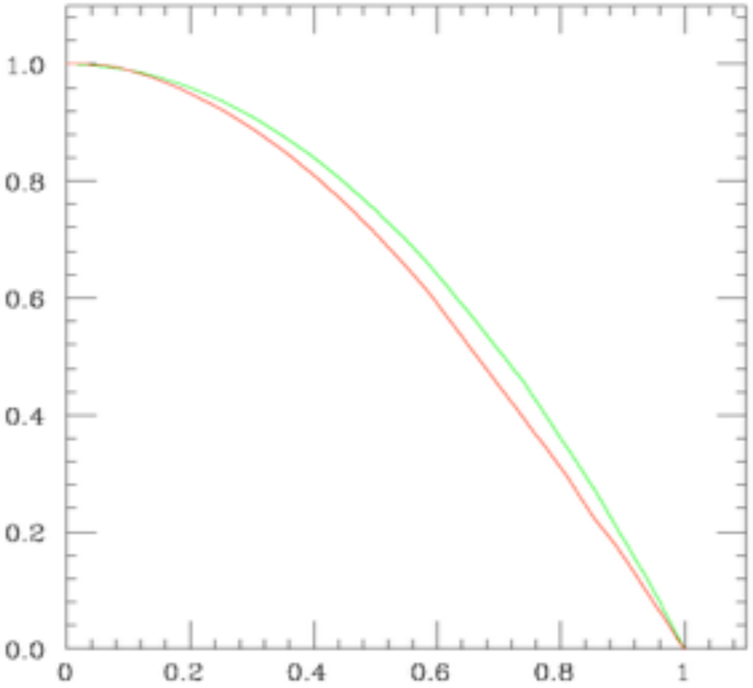
$$\begin{aligned} I &= \int_0^1 dx \cos\tfrac{\pi}{2}x \\ &= 0.637 \pm 0.308/\sqrt{N} \end{aligned}$$

$$\begin{aligned} I &= \int_0^1 dx(1-x^2)\,\frac{\cos\frac{\pi}{2}x}{1-x^2} \\ &= \int d\rho\,\frac{\cos\frac{\pi}{2}x}{1-x^2}[x(\rho)] \\ &= 0.637 \pm 0.037\,/\sqrt{N} \end{aligned}$$

Hit-and-Miss

- Accept points with probability = w_i/w_{max} (provided all $w_i \geq 0$)
- Accepted points are distributed like real events (cf. Buffon's needle)
- MC efficiency $\varepsilon_{MC} = E[w]/w_{max}$ improved by importance sampling





$$\varepsilon_{\rm MC} = 1/I = 2/\pi = 64\%$$

$$\sigma = \sqrt{\frac{\varepsilon_{\rm MC}(1-\varepsilon_{\rm MC})}{N}} = \frac{0.48}{\sqrt{N}}$$

$$\varepsilon_{\rm MC} = \int_0^1 dx(1-x^2)/I = 3/\pi = 95\%$$

$$\sigma = \sqrt{\frac{\varepsilon_{\rm MC}(1-\varepsilon_{\rm MC})}{N}} = \frac{0.21}{\sqrt{N}}$$

Multi-dimensional Integration

- Formalism extends trivially to many dimensions
- Particle physics: very many dimensions,

  e.g. phase space = 3 dimensions per particles,

  LHC event ~ 250 hadrons.
- Monte Carlo error remains $\propto 1/\sqrt{N}$
- Trapezium rule $\propto 1/N^{2/d}$
- Simpson's rule $\propto 1/N^{4/d}$

Monte Carlo: Summary

Disadvantages of Monte Carlo:

- Slow convergence in few dimensions.

Advantages of Monte Carlo:

- Fast convergence in many dimensions.
- Arbitrarily complex integration regions (finite discontinuities not a problem).
- Few points needed to get first estimate ("feasibility limit").
- Every additional point improves accuracy ("growth rate").
- Easy error estimate.
- Hit-and-miss allows unweighted event generation, i.e. points distributed in phase space just like real events.

Phase Space Generation

$$\sigma = \frac{1}{2s}\int |\mathcal{M}|^2 \, d\Pi_n(\sqrt{s})$$

$$\Gamma = \frac{1}{2M}\int |\mathcal{M}|^2 \, d\Pi_n(M)$$

- Phase space:

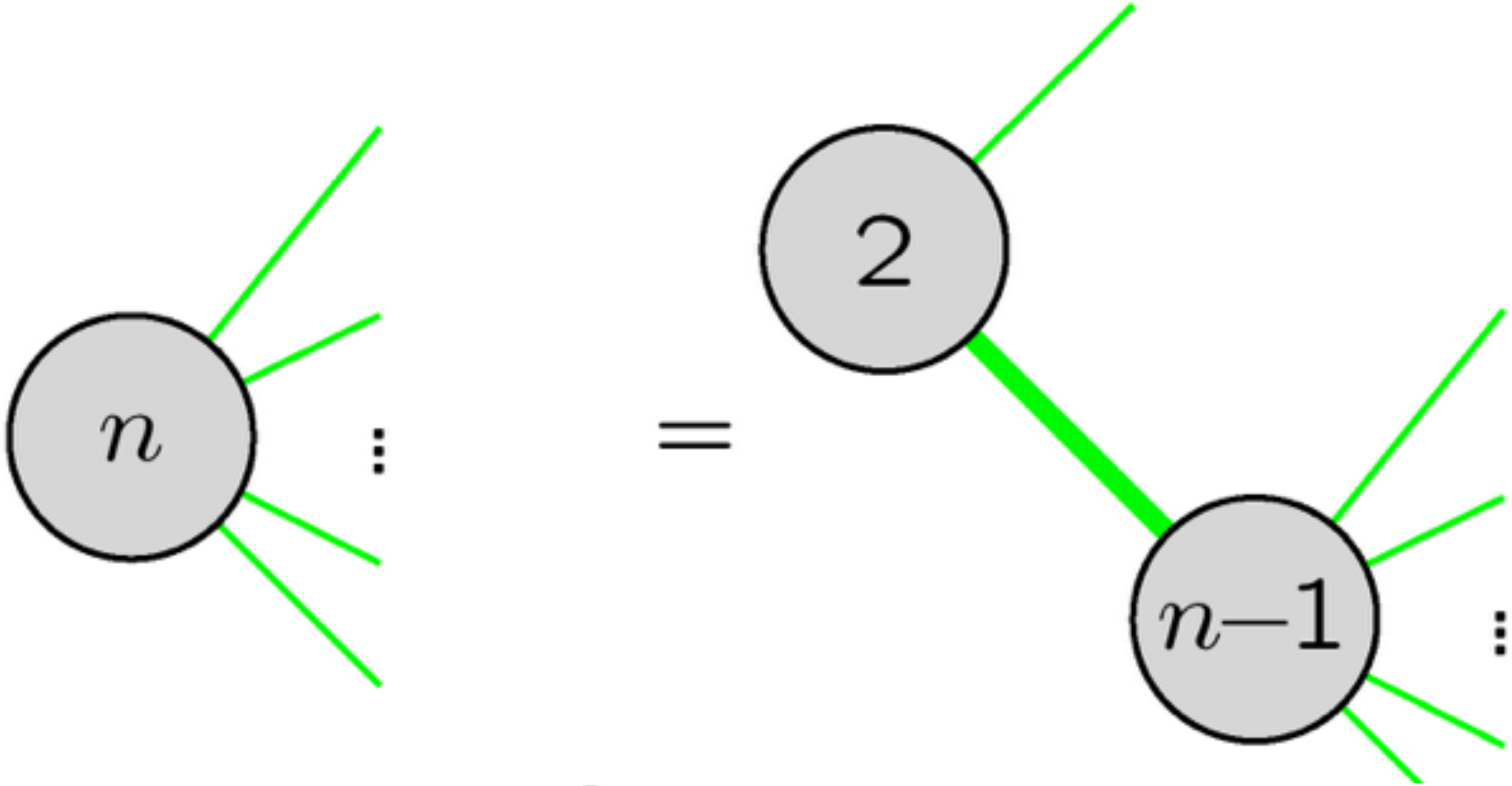
$$d\Pi_n(M) = \left[\prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 (2E_i)}\right] (2\pi)^4 \delta^{(4)}\left(p_0 - \sum_{i=1}^{n} p_i\right)$$

- Two-body easy:

$$d\Pi_2(M) = \frac{1}{8\pi}\,\frac{2p}{M}\,\frac{d\Omega}{4\pi}$$

Phase Space Generation

- Other cases by recursive subdivision:



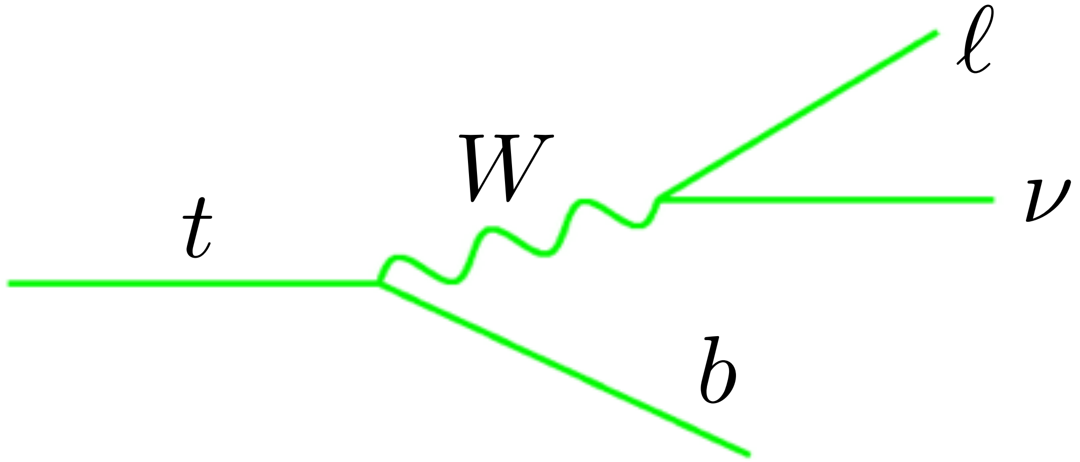
$$d\Pi_n(M) = \frac{1}{2\pi}\int_0^{(M-m)^2} dm_x^2\, d\Pi_2(M)\, d\Pi_{n-1}(m_x)$$

- Or by ‘democratic’ algorithms: RAMBO, MAMBO
  Can be better, but matrix elements rarely flat.

Particle Decays

- Simplest example

  e.g. top quark decay:



$$|\mathcal{M}|^2 = \frac{1}{2}\left(\frac{8\pi\alpha}{\sin^2\theta_w}\right)^2 \frac{p_t \cdot p_\ell \, p_b \cdot p_\nu}{(m_W^2 - M_W^2)^2 + \Gamma_W^2 M_W^2}$$

$$\Gamma = \frac{1}{2M}\frac{1}{128\pi^3}\int |\mathcal{M}|^2 dm_W^2 \left(1 - \frac{m_W^2}{M^2}\right)\frac{d\Omega}{4\pi}\frac{d\Omega_W}{4\pi}$$

Breit-Wigner peak of W very strong - but can be removed by importance sampling:

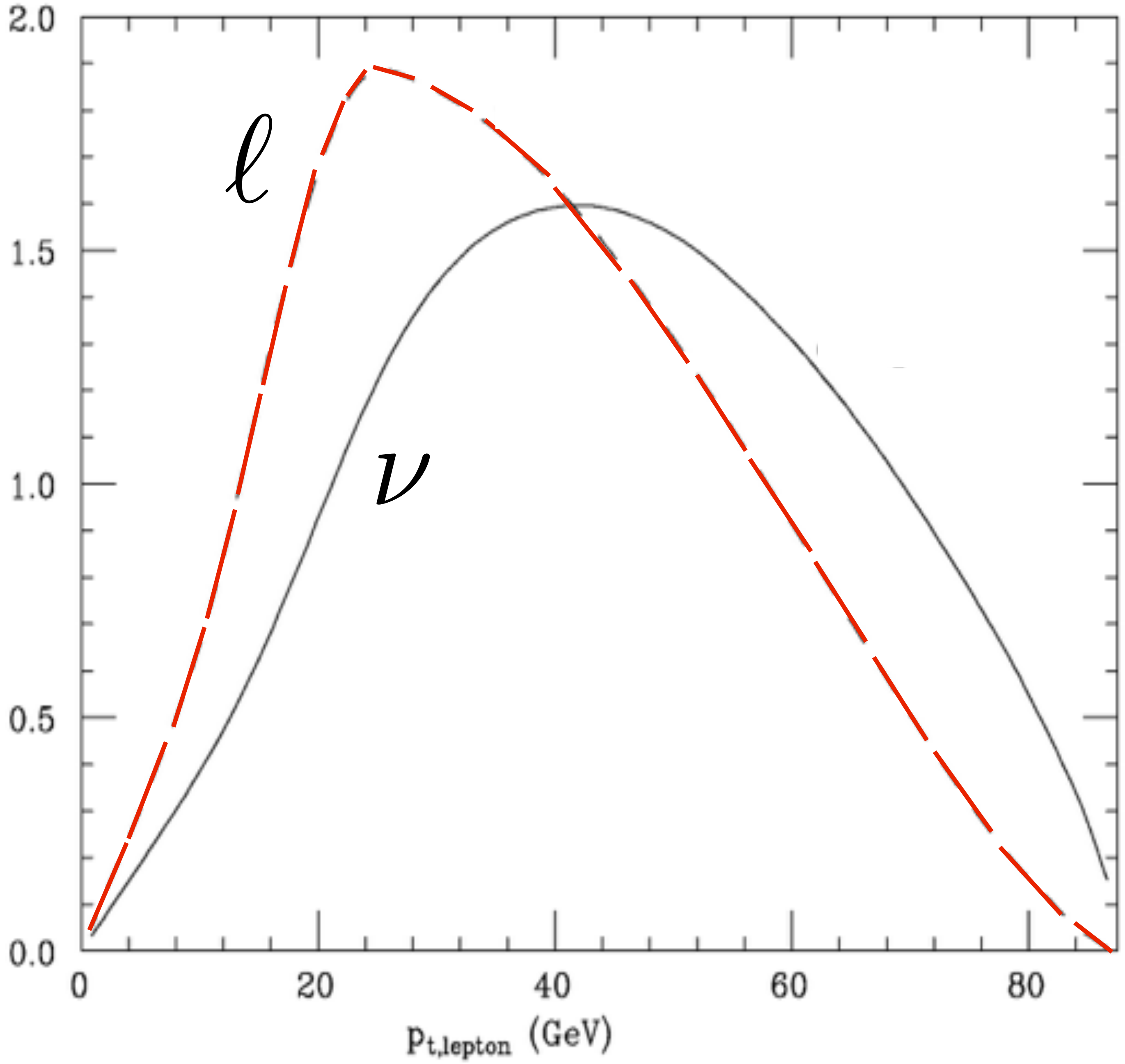
$$m_W^2 \to \arctan\left(\frac{m_W^2 - M_W^2}{\Gamma_W M_W}\right)$$ (prove it!)

Associated Distributions

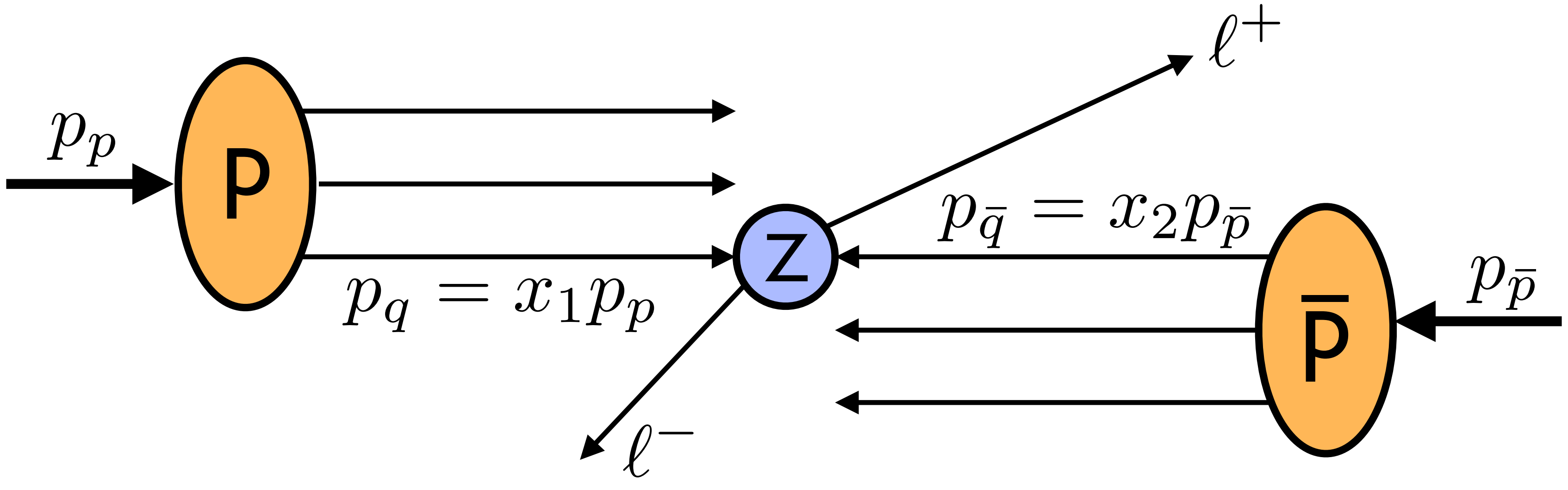
Big advantage of Monte Carlo integration:

- Simply histogram any associated quantities.
- Almost any other technique requires new integration for each observable.
- Can apply arbitrary cuts/ smearing.

e.g. lepton momentum in top decays:



Hadron-Hadron Cross Sections



$s = E_{\text{cm}}^2 = 4p_p p_{\bar{p}}$

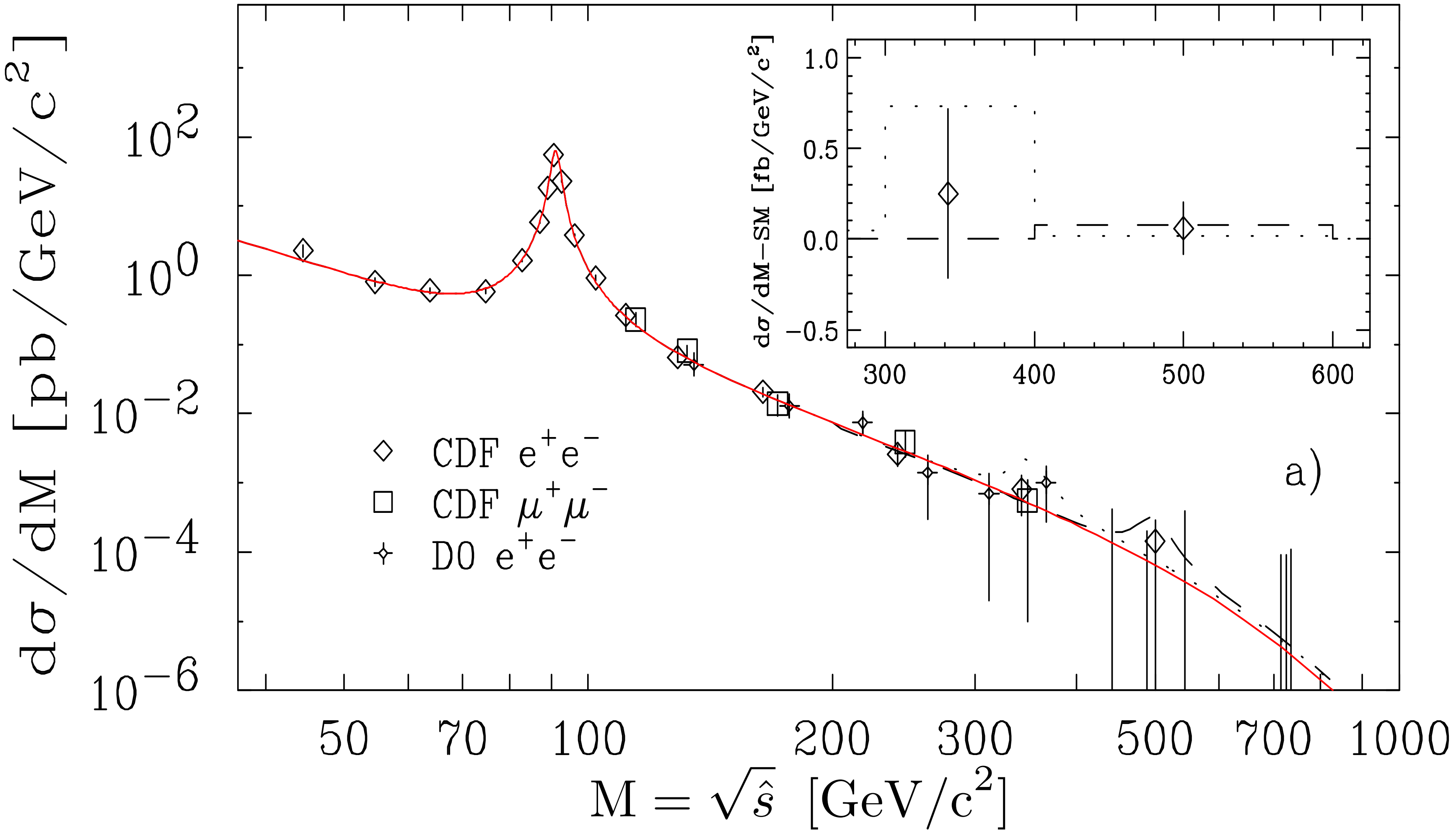
$\hat{s} = 4p_q p_{\bar{q}} = x_1 x_2 s$

- Consider e.g. $p\bar{p} \to Z^0 \to \ell^+\ell^-$
- Integrations over incoming parton momentum distributions:

$$\sigma(s) = \int_0^1 dx_1 f(x_1) \int_0^1 dx_2 f(x_2)\, \hat{\sigma}(x_1 x_2 s)$$

- Hard process cross section $\hat{\sigma}(\hat{s})$ has strong peak, due to Z^0 resonance: needs importance sampling (like W in top decay)

$p\bar{p} \to \ell^+\ell^-$cross section



$$\hat{\sigma}_{q\bar{q}\to Z^0\to\ell^+\ell^-} = \frac{4\pi\hat{s}}{3M_Z^2}\frac{\Gamma_\ell\Gamma_q}{(\hat{s}-M_Z^2)^2+\Gamma_Z^2M_Z^2}$$

- “Background” is $q\bar{q} \to \gamma^* \to \ell^+\ell^-$

Parton-Level Monte Carlo Calculations

Now we have everything we need to make parton-level cross section calculations and distributions

Can be largely automated…

- MADGRAPH
- GRACE
- COMPHEP
- AMEGIC++
- ALPGEN

But…

- Fixed parton/jet multiplicity
- No control of large higher-order corrections
- Parton level

→ Need hadron level event generators

Summary of Lecture 1

- Monte Carlo is a very convenient numerical integration method.
- Well-suited to particle physics: difficult integrands, many dimensions.
- Integrand non-negative → hit-and-miss event generation.
- Hard process: use parton-level generator.
- Next: parton showers and hadron-level event generation