Introduction to MadGraph/MadEvent 1

Rikkert Frederix
University of Zurich

2011 IPMU-YITP School on Monte Carlo Tools for the LHC, Kyoto, Japan, September 5-9, 2011

MadGraph's philosofy

- T. Edison:
  Genius is 1% inspiration and 99% perspiration.

- MadGraph's goal:
  ok, guys, let's improve on the 99%.

Aims for these lectures

- Get you acquainted with the concepts and techniques used in event generation with MadGraph/ MadEvent
- Give you hands-on experience with matrix element generation, event generation and analysis, as well as new physics implementation
- Answer as many questions from you as I can... (So, please ask questions!)

OUTLINE

- MadGraph/Madevent beginner
  - MadGraph
  - MadEvent
- Advanced user

Outline

- MadGraph/Madevent beginner
  - MadGraph
  - MadEvent
- Advanced user

Master equation for hadron colliders

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\rm FS}\, f_a(x_1,\mu_F) f_b(x_2,\mu_F)\, \hat\sigma_{ab\to X}(\hat s,\mu_F,\mu_R)$$

Phase-space integral | Parton density functions | Parton-level cross section

- Parton-level cross section from matrix elements: model and process dependent
- Parton density (or distribution) functions: process independent
- Differences between colliders given by parton luminosities

In practice

- To compute a cross section (or make distributions) the following steps have to be taken

MadGraph

1. Define all subprocess contributions to a given process (u u~ ➞ t t~ g, g g ➞ t t~ g, d g ➞ t t~ d, etc.) *Easy*

2. Write down all the amplitudes (e.g. with Feynman diagrams) *Straight-forward*

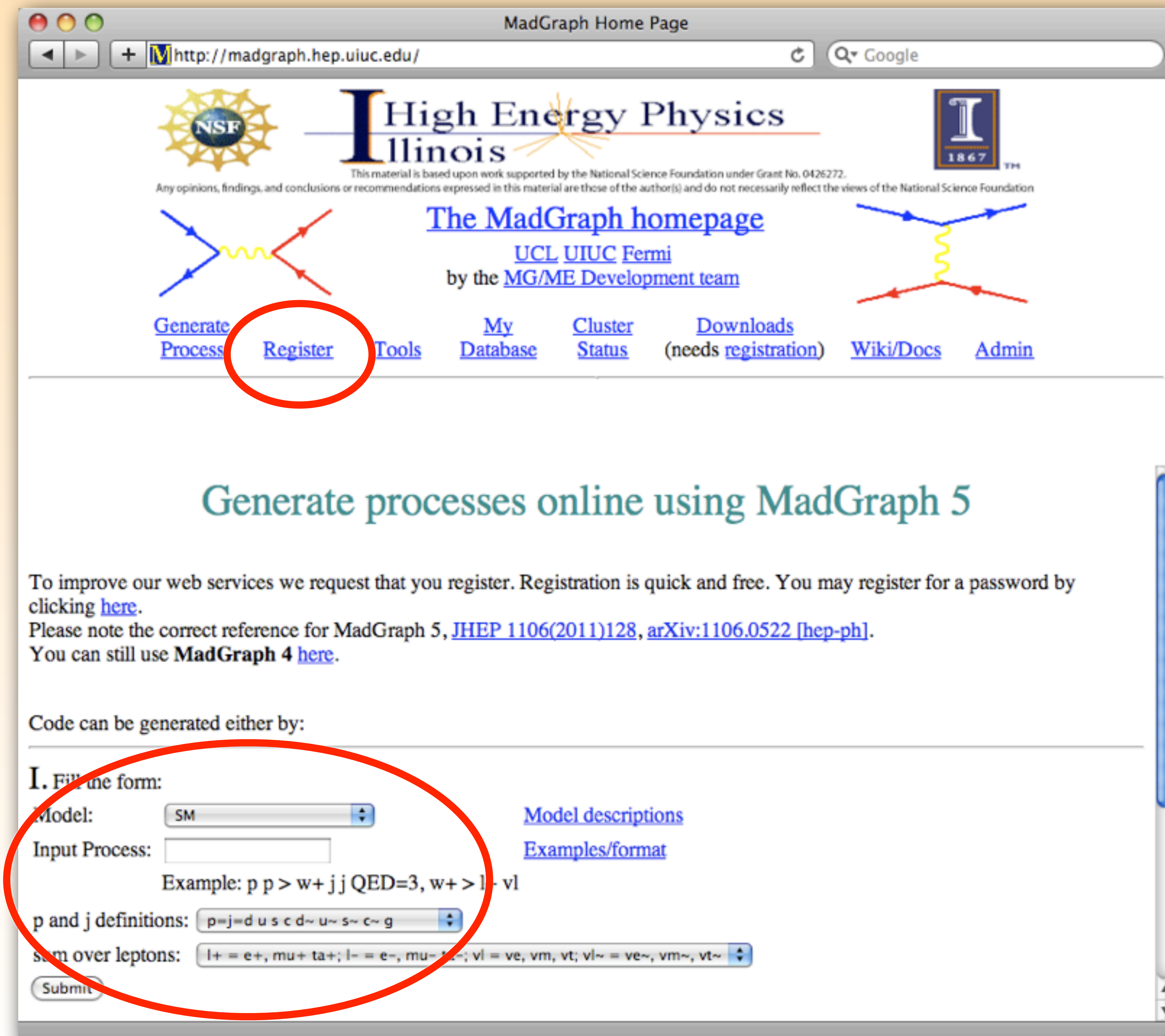
MadEvent

3. Perform the phase-space integration (using MC techniques) & generate (unweighted) events (that can be showered (and put through detector simulation)) *Difficult*

- This is what MadGraph/MadEvent does for you in an automatic way (at leading order in perturbation theory)

MadGraph

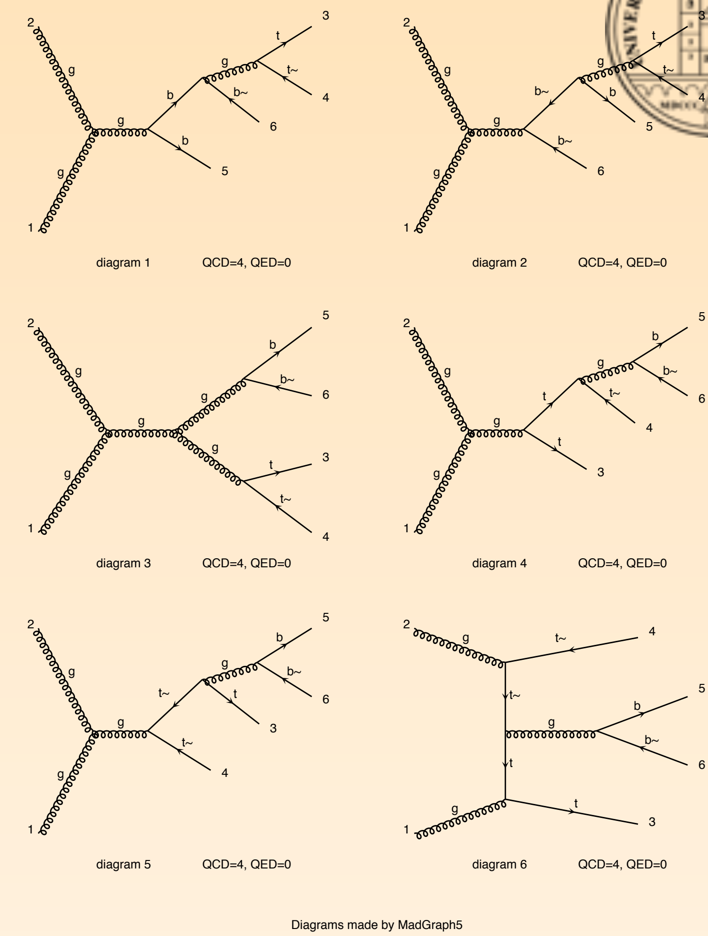
- How to start?
  - Open your favorite web brower and go to http://madgraph.hep.uiuc.edu/ or any of the other publicly available MadGraph computer clusters


MadGraph Home Page
http://madgraph.hep.uiuc.edu/
High Energy Physics
Illinois
This material is based upon work supported by the National Science Foundation under Grant No. 0426272.
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation
The MadGraph homepage
UCL UIUC Fermi
by the MG/ME Development team
Generate Process
Register
Tools
My Database
Cluster Status
Downloads (needs registration)
Wiki/Docs
Admin
Generate processes online using MadGraph 5
To improve our web services we request that you register. Registration is quick and free. You may register for a password by clicking here.
Please note the correct reference for MadGraph 5, JHEP 1106(2011)128, arXiv:1106.0522 [hep-ph].
You can still use MadGraph 4 here.
Code can be generated either by:
I. Fill the form:
Model: SM
Model descriptions
Input Process:
Examples/format
Example: p p > w+ j j QED=3, w+ > l+ vl
p and j definitions: p=j=d u s c d~ u~ s~ c~ g
sum over leptons: l+ = e+, mu+ ta+; l- = e-, mu- ta-; vl = ve, vm, vt; vl~ = ve~, vm~, vt~
Submit

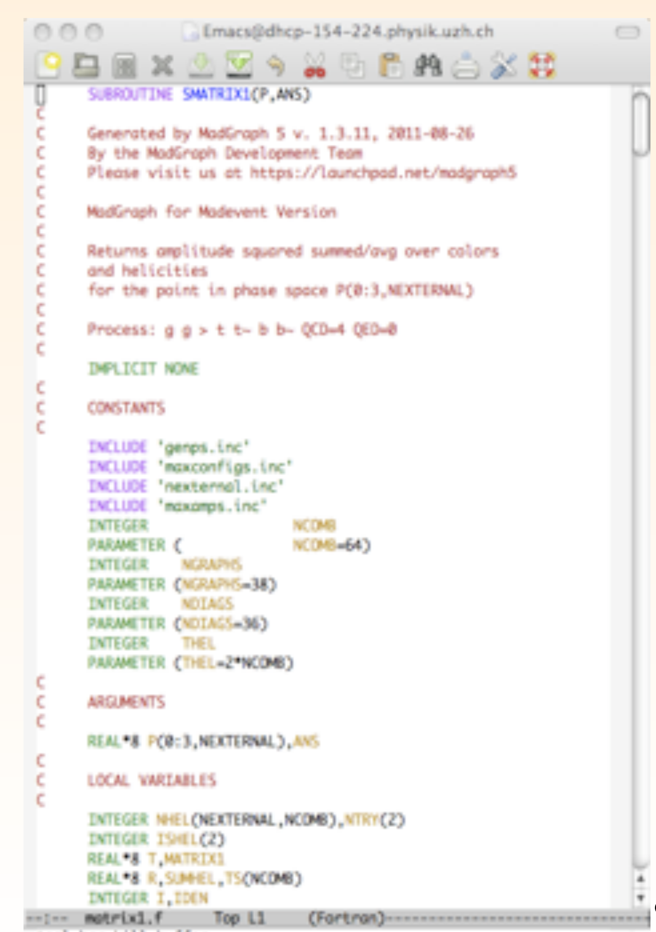
MadGraph

- User requests:
  - Physics model (e.g. SM)
  - Process (e.g. p p > t t~ b b~ QCD=4 QED=0)

MadGraph



- MadGraph returns
  - Feynman diagram
  - Self-contained Fortran77 code for $|M|^2$ (including code for phase-space integration)

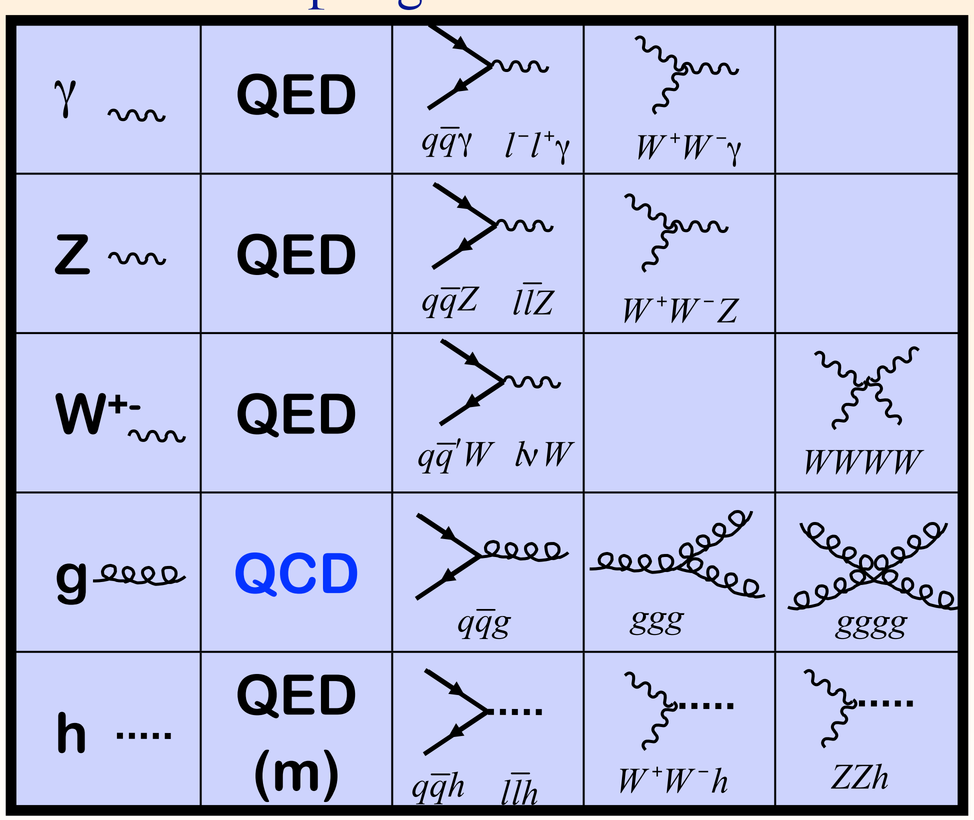


PROCESS SYNTAX

particle definitions for SM

Fermions				Bosons	
type	1st gen	2nd gen	3rd gen	name	symbol
quarks	u (u~)	c (c~)	t (t~)	gluon	g
	d (d~)	s (s~)	b (b~)	photon	a
leptons	ve (ve~)	vm (vm~)	vt (vt~)	weak bosons	w+ (w-), z
	e- (e+)	mu- (mu+)	ta- (ta+)	higgs	h

coupling orders of SM



- The syntax for the process line is as follows
  - ☞ particles are separated by spaces
  - initial and final states are separated by a ">"
  - use "x x > z > y y" to require particle z to appear as an s-channel intermediate state
  - use "x x > y y /z" to exclude particle z
  - use "x x > y y $z" to exclude particle z from an s-channel
  - ☞ use comma's and brackets to define decay chains: "x x > y6 z1 z3, (z1 > y1 z2, z2 > y2 y3), z3 > y4 y5"
  - ☞ if no coupling orders given, MadGraph will minimize QED and maximize QCD couplings

☞ changed from MG 4 to MG 5

MadGraph v5

- MadGraph v5 was officially released in June this year
- It has been a major rewrite of the MadGraph v4, which is going to be phased-out
  - There will be no more new features implemented in v4 (apart from the sporadic bug-fix)
- Everybody should be moving to MadGraph v5
  - Based on similar principles as MG4, but implemented in a totally different way
    - Simpler code (high-level language)
    - More versatile (UFO → Aloha → matrix elements)
    - Faster (better algorithms & improved PS integration)

Diagram generation

- Written in python (high-level language which uses extremely-optimized libraries (mostly in C) for common manipulations)
- This makes the diagram generation faster than the old MG 4 (which was written in Fortran)
- Improved algorithms allow for 5, 6, 7, ...-point interactions (for higher-dimensional effective interactions)
- Extensive test suite, to check each function & feature

Process	MadGraph 4	MadGraph 5	Subprocesses	Diagrams
$pp \to jjj$	29.0 s	25.8 s	34	307
$pp \to jjl^+l^-$	341 s	103 s	108	1216
$pp \to jjje^+e^-$	1150 s	134 s	141	9012
$u\bar{u} \to e^+e^-e^+e^-e^+e^-$	772 s	242 s	1	3474
$gg \to ggggg$	2788 s	1050 s	1	7245
$pp \to jj(W^+ \to l^+\nu_l)$	146 s	25.7 s	82	304
$pp \to t\bar{t}$+full decays	5640 s	15.7 s	27	45
$pp \to \tilde{q}/\tilde{g}\,\tilde{q}/\tilde{g}$	222 s	107 s	313	475
7 particle decay chain	383 s	13.9 s	1	6
$gg \to (\tilde{g} \to u\bar{u}\tilde{\chi}_1^0)(\tilde{g} \to u\bar{u}\tilde{\chi}_1^0)$	70 s	13.9 s	1	48
$pp \to (\tilde{g} \to jj\tilde{\chi}_1^0)(\tilde{g} \to jj\tilde{\chi}_1^0)$	—	251 s	144	11008

Exercises I

- Log on to http://madgraph.hep.uiuc.edu and generate some processes. Guess which diagrams will appear, and think about which subprocesses will contribute:
  - u u~ > t t~ (with and without QED vertices)
  - g g > t t~
  - g g > t t~ h
  - p p > t t~ b b~
  - p p > t t~ j j
- Exercises available at the madgraph wiki page
  http://cp3wks05.fynu.ucl.ac.be/twiki/bin/view/Main/IPMUYITPMCSchool2011

Outline

- MadGraph/MadEvent beginner
  - MadGraph
  - MadEvent
- Advanced user

MadEvent

- Let's go back to the process that we just generated on-line...

MadGraph Home Page

http://madgraph.hep.uiuc.edu/

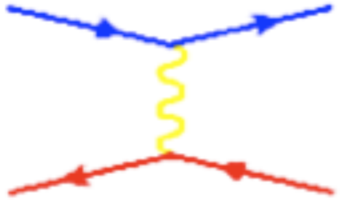
High Energy Physics Illinois

This material is based upon work supported by the National Science Foundation under Grant No. 0426272.
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

The MadGraph homepage

UCL UIUC Fermi

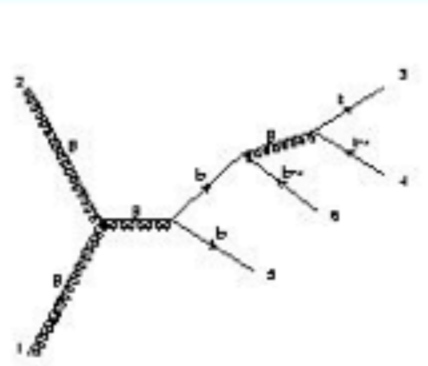
by the MG/ME Development team



Generate Process | Register | Tools | My Database | Cluster Status | Downloads (needs registration) | Wiki/Docs | Admin

---

Process: **p p > t t~ b b~**
QCD=4
QED=0
Model: **sm**



Links	Status
Process Information	Generation Complete
Code Download	Available
On-line Event Generation	Available (access restricted)
Results and Event Database	No runs available

Notes:

Last Update: Tue Aug 30 02:56:13 CDT 2011

MadEvent on-line

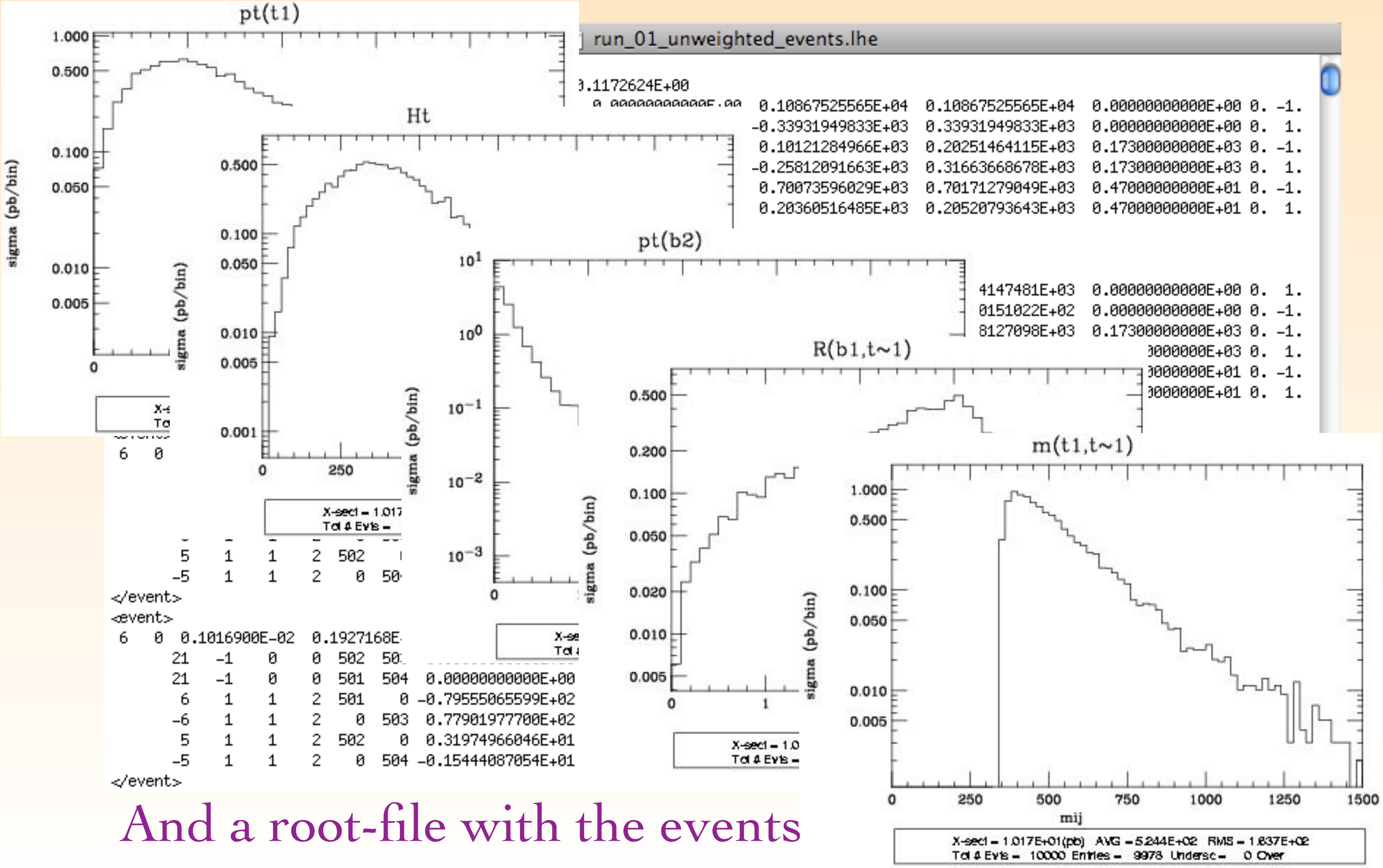
- Send me an e-mail (with your username), and I can give you access for running on our clusters
- User needs to provide (starting from a process generated by MadGraph)
  - model parameters (such as masses and coupling strenghts)
  - collider parameters (such as collision energy)
  - and, possibly, basic acceptance cuts (to avoid IR divergences)

MadEvent on-line

run_01_unweighted_events.lhe


```
<event>
 6   0  0.1016900E-02  0.1808590E+03  0.7957747E-01  0.1172624E+00
       21   -1    0    0  502  503  0.00000000000E+00  0.00000000000E+00  0.10867525565E+04  0.10867525565E+04  0.00000000000E+00 0. -1.
       21   -1    0    0  501  504  0.00000000000E+00  0.00000000000E+00 -0.33931949833E+03  0.33931949833E+03  0.00000000000E+00 0.  1.
        6    1    1    2  501    0 -0.59994840294E+01  0.28339815394E+02  0.10121284966E+03  0.20251464115E+03  0.17300000000E+03 0. -1.
       -6    1    1    2    0  503 -0.52184916050E+02 -0.31306841917E+02 -0.25812091663E+03  0.31663668678E+03  0.17300000000E+03 0.  1.
        5    1    1    2  502    0  0.34697961305E+02  0.11996490016E+02  0.70073596029E+03  0.70171279049E+03  0.47000000000E+01 0. -1.
       -5    1    1    2    0  504  0.23486438774E+02 -0.90294634933E+01  0.20360516485E+03  0.20520793643E+03  0.47000000000E+01 0.  1.
</event>
<event>
 6   0  0.1016900E-02  0.1907016E+03  0.7957747E-01  0.1163799E+00
       -2   -1    0    0    0  501  0.00000000000E+00  0.00000000000E+00  0.75114147481E+03  0.75114147481E+03  0.00000000000E+00 0.  1.
        2   -1    0    0  503    0  0.00000000000E+00  0.00000000000E+00 -0.72550151022E+02  0.72550151022E+02  0.00000000000E+00 0. -1.
        6    1    1    2  502    0  0.89414634842E+02  0.27670123043E+02  0.17000190772E+03  0.25998127098E+03  0.17300000000E+03 0. -1.
       -6    1    1    2    0  502 -0.60587947037E+02 -0.22019518840E-01  0.32320293992E+03  0.37156431500E+03  0.17300000000E+03 0.  1.
        5    1    1    2  503    0 -0.31412647870E+02 -0.12181207871E+02  0.13139831888E+03  0.13573041103E+03  0.47000000000E+01 0. -1.
       -5    1    1    2    0  501  0.25859600652E+01 -0.15466895654E+02  0.53988157268E+02  0.56415628826E+02  0.47000000000E+01 0.  1.
</event>
<event>
 6   0  0.1016900E-02  0.2489952E+03  0.7957747E-01  0.1121328E+00
       21   -1    0    0  501  503  0.00000000000E+00  0.00000000000E+00  0.10901548395E+04  0.10901548395E+04  0.00000000000E+00 0. -1.
       21   -1    0    0  502  504  0.00000000000E+00  0.00000000000E+00 -0.24528828154E+03  0.24528828154E+03  0.00000000000E+00 0. -1.
        6    1    1    2  501    0 -0.63410753496E+02 -0.19976077611E+03  0.10538070994E+03  0.29147793289E+03  0.17300000000E+03 0. -1.
       -6    1    1    2    0  503  0.69371980485E+02  0.15921290804E+03  0.86585020836E+03  0.89988155058E+03  0.17300000000E+03 0. -1.
        5    1    1    2  502    0  0.77836222535E+00 -0.18989098793E+01 -0.81916051073E+02  0.82076434681E+02  0.47000000000E+01 0. -1.
       -5    1    1    2    0  504 -0.67395892140E+01  0.42446777948E+02 -0.44448309234E+02  0.62007202926E+02  0.47000000000E+01 0. -1.
</event>
<event>
 6   0  0.1016900E-02  0.1927168E+03  0.7957747E-01  0.1162065E+00
       21   -1    0    0  502  503  0.00000000000E+00  0.00000000000E+00  0.29430190167E+03  0.29430190167E+03  0.00000000000E+00 0. -1.
       21   -1    0    0  501  504  0.00000000000E+00  0.00000000000E+00 -0.20865975429E+03  0.20865975429E+03  0.00000000000E+00 0.  1.
        6    1    1    2  501    0 -0.79555065599E+02 -0.35049716038E+02 -0.58346309673E+02  0.20221469509E+03  0.17300000000E+03 0.  1.
       -6    1    1    2    0  503  0.77901977700E+02  0.32147748008E+02  0.17960544706E+03  0.26322863151E+03  0.17300000000E+03 0. -1.
        5    1    1    2  502    0  0.31974966046E+01  0.30254032900E+01 -0.24811167865E+02  0.25633203085E+02  0.47000000000E+01 0.  1.
       -5    1    1    2    0  504 -0.15444087054E+01 -0.12343525999E+00 -0.10805822141E+02  0.11885126279E+02  0.47000000000E+01 0. -1.
</event>
```

MadEvent on-line



And a root-file with the events

Exercises II

- Generate events for some processes for Tevatron and LHC energies. Does the cross section scale as you expected?
- Compare cross sections with processes of other participants. Do they differ as you would expect?
- Look at the generated plots. Are the distributions as you expect? Discuss with your neighbors.
- Example processes:
  - p p > t t~ QED=0
  - p p > l+ vl
  - p p > e+ e-
  - g g > H, H > e+ ve mu- vm~ (in the HiggsEFT model)
  - p p > W+ W-, W+ > e+ ve, W- > mu- vm~
  - ...

On-line vs. Off-line

- It is recommended to run the code on our clusters:
  - Always the latest version
  - Uses many CPUs, so results are often quickly obtained
  - No need to install the code
  - Personal database
  - Works also from your smartphone :-)
- However, sometimes running locally works as well
  - Gives more flexibility (more models, other cuts or event-by-event renormalization and factorization scales, etc.)

Phase-space integration

- Calculations of cross sections or decay widths involve integrations over high-dimensional phase space with complicated integration bounds ("cuts") of very peaked functions:

$$\sigma = \frac{1}{2s}\int |\mathcal{M}|^2 d\Phi(n)$$

$$Dim[\Phi(n)] \sim 3n$$

- General and flexible method is needed

Monte Carlo Integration

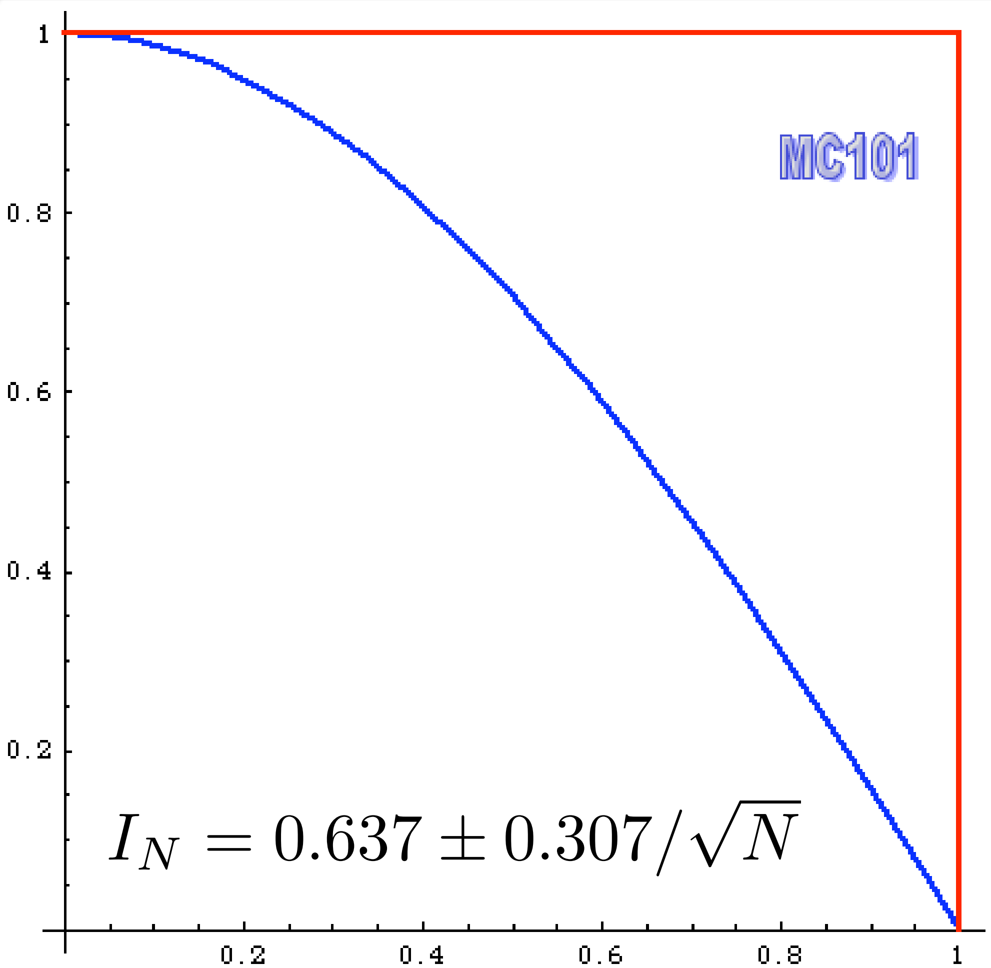
$$I = \int_{x_1}^{x_2} f(x)dx \quad \Rightarrow \quad I_N = (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x)$$

$$V = (x_2 - x_1)\int_{x_1}^{x_2} [f(x)]^2 dx - I^2 \quad \Rightarrow \quad V_N = (x_2 - x_1)^2 \frac{1}{N}\sum_{i=1}^{N}[f(x)]^2 - I_N^2$$

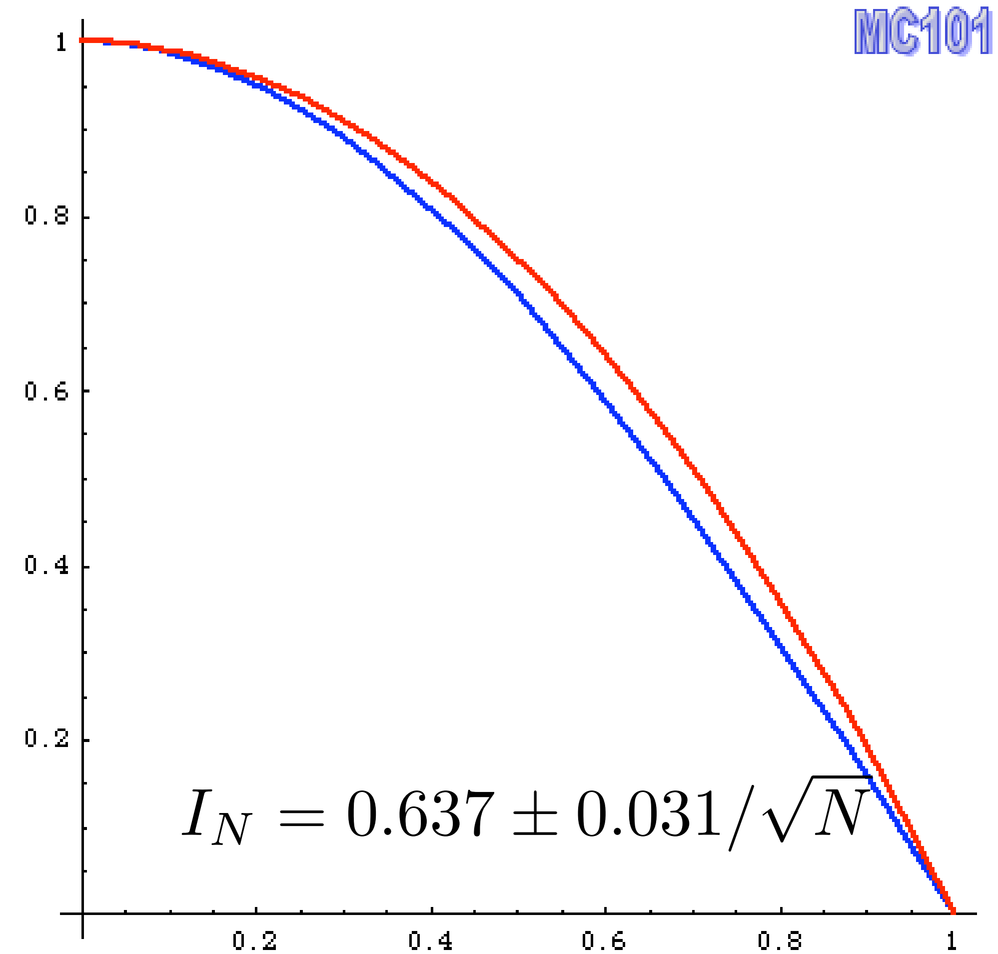
$$I = I_N \pm \sqrt{V_N/N}$$

- Convergence is slow, but can be estimated easily
- Error does not depend on the number of dimensions
- Improvement by minimizing V_N
- Optimal case: $f(x)$=Constant $\Rightarrow V_N=0$

Importance sampling



$$I = \int_0^1 dx \cos\frac{\pi}{2}x$$

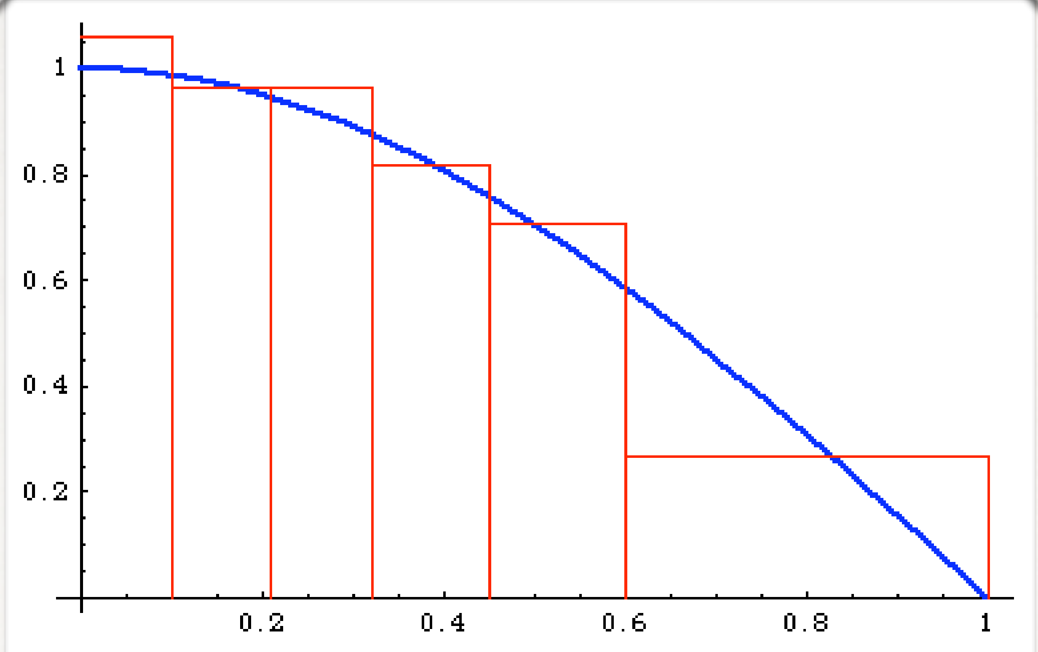


$$I = \int_0^1 dx(1-x^2)\frac{\cos\frac{\pi}{2}x}{1-x^2}$$

$$= \int_{\xi_1}^{\xi_2} d\xi \frac{\cos\frac{\pi}{2}x[\xi]}{1-x[\xi]^2} \Rightarrow \simeq 1$$

Importance sampling

- But this only works if you know a lot (preferably the functional form) of $f(x)$
- Alternative: learn during the run and build a step function $p(x)$ that approximates $f(x)$ $\Rightarrow$ VEGAS algorithm



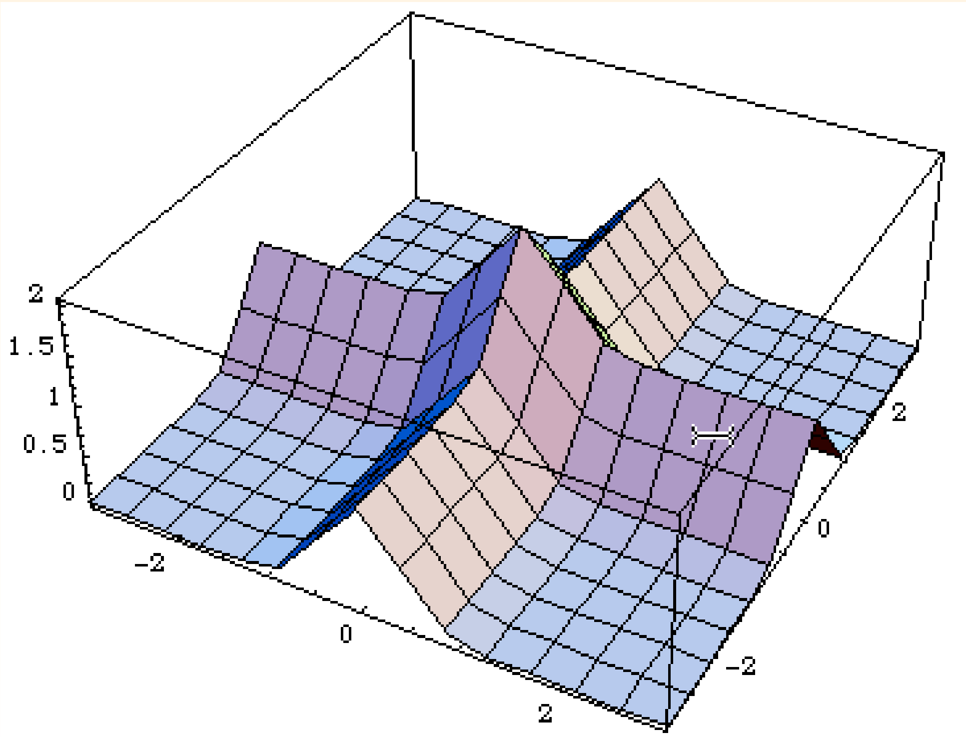
Many bins where $f(x)$ is fluctuating $\Rightarrow$ more points thrown where $f(x)$ is large

In more dimensions

- Importance sampling can be generalized to more dimensions

$$p(x) = p_a(x_1) \cdot p_b(x_2) \cdot p_c(x_3) \cdot \ldots$$

- But the peaks of $f(x)$ need to be "aligned" to the axis!



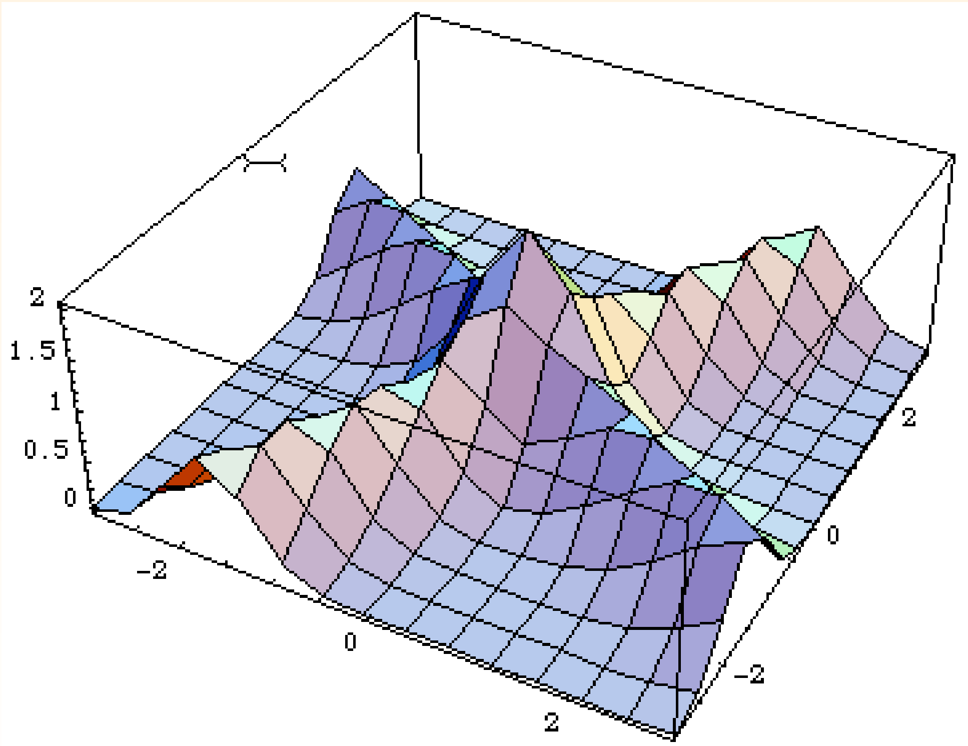
This is okay

In more dimensions

- Importance sampling can be generalized to more dimensions

$$p(x) = p_a(x_1) \cdot p_b(x_2) \cdot p_c(x_3) \cdot \ldots$$

- But the peaks of $f(x)$ need to be "aligned" to the axis!



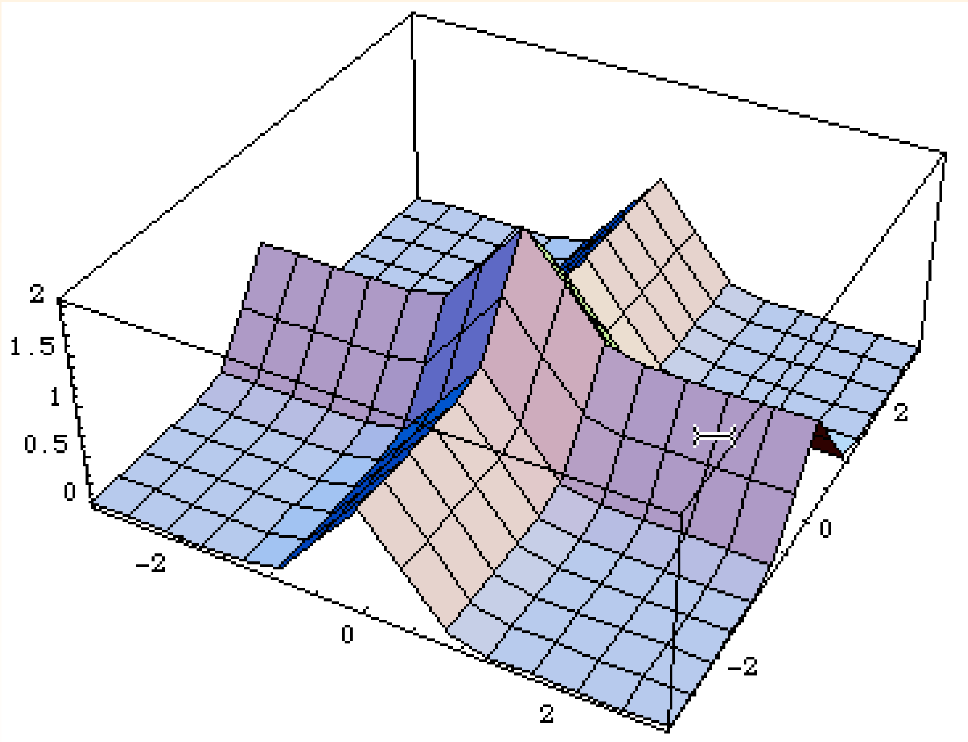
This is not okay

In more dimensions

- Importance sampling can be generalized to more dimensions

$$p(x) = p_a(x_1) \cdot p_b(x_2) \cdot p_c(x_3) \cdot \ldots$$

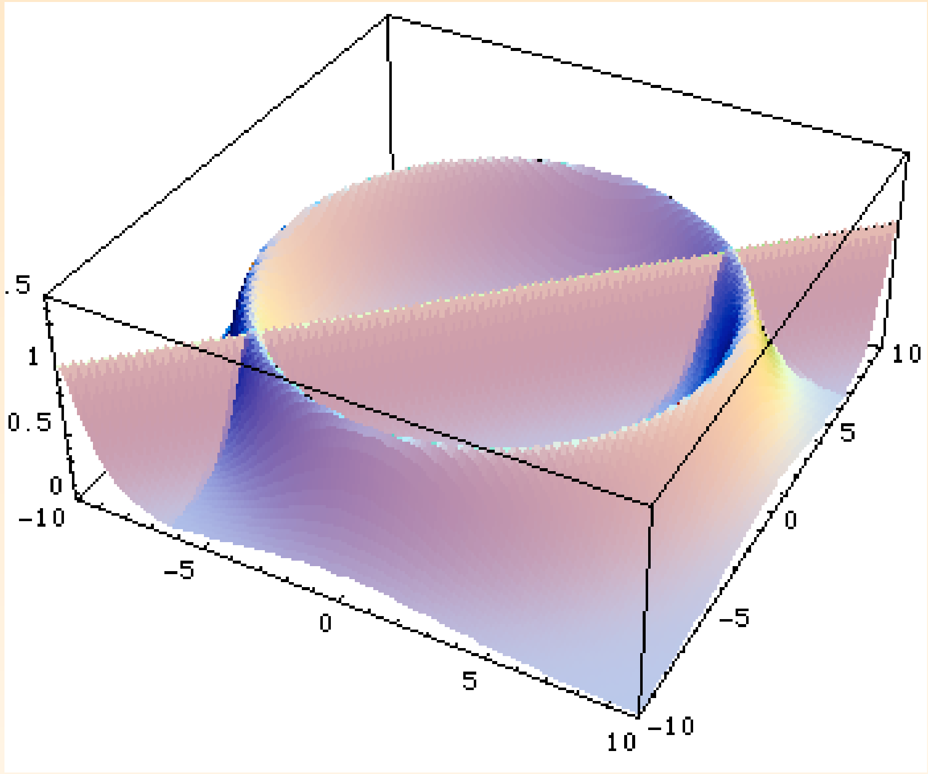
- But the peaks of $f(x)$ need to be "aligned" to the axis!



This is not okay

But a simple rotation (change of variables) brings it back to this

MULTI-CHANNEL INTEGRATION



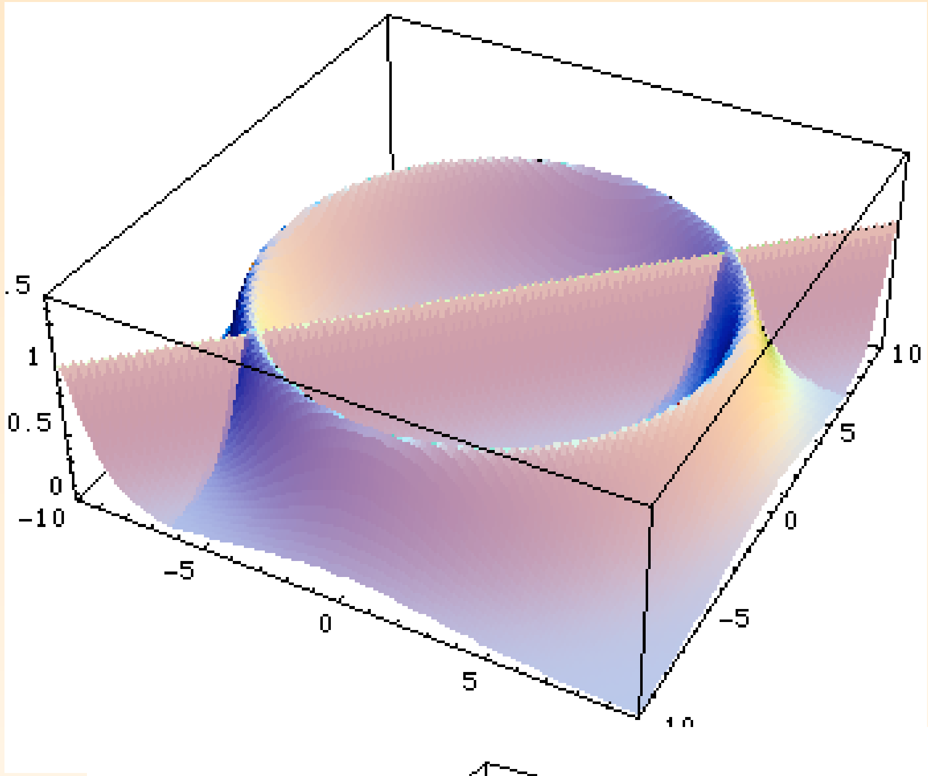
- Some integrals are considerably more complicated and a single change of variables does not help: VEGAS is bound to fail

- Solution: use different transformations for each of the "structures" = channels

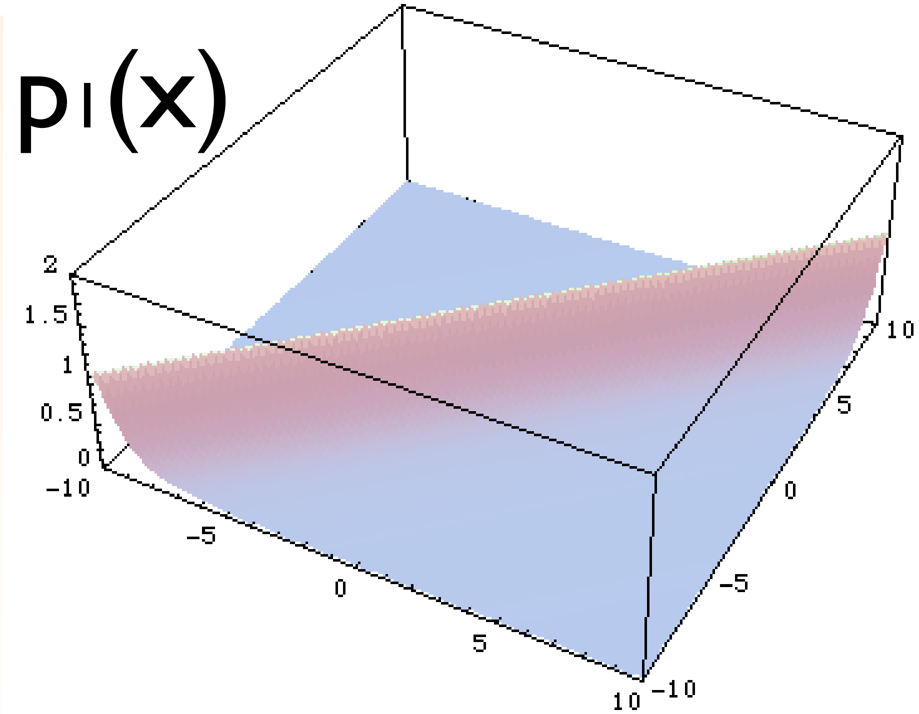
$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x) \qquad \text{with} \qquad \sum_{i=1}^{n} \alpha_i = 1$$

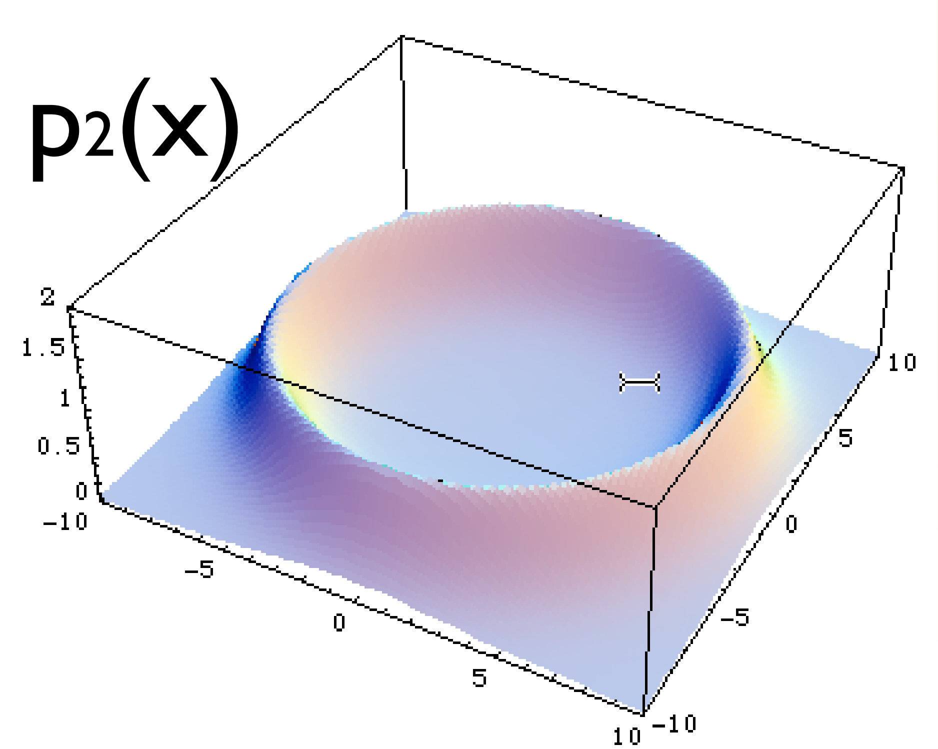
- with each $p_i(x)$ taking care of one structures at the time

MULTI-CHANNEL INTEGRATION

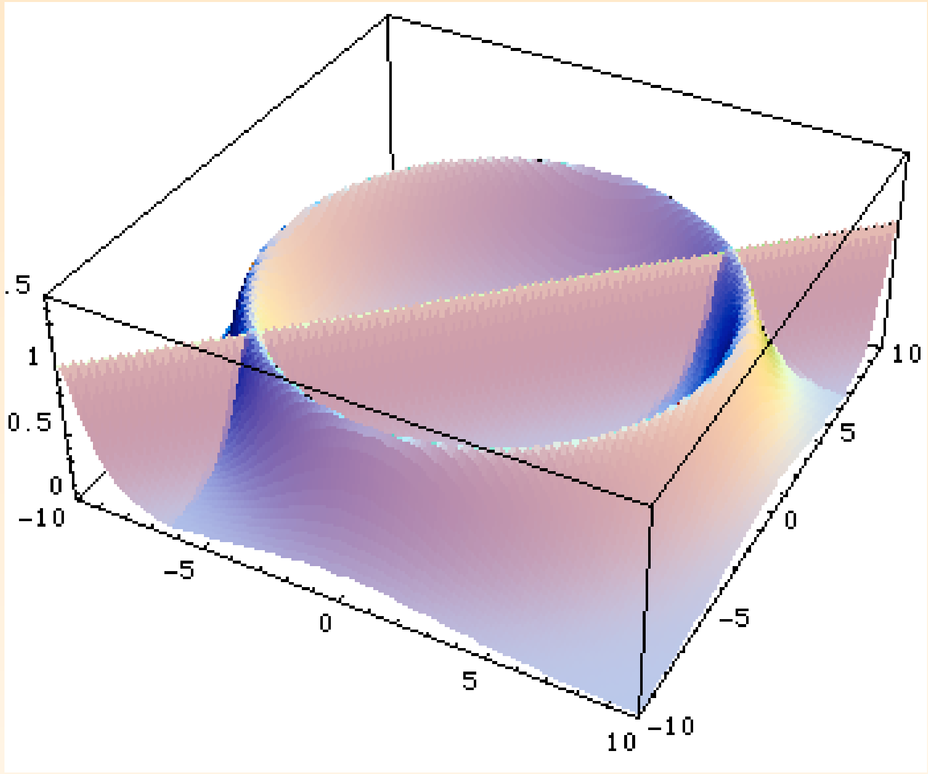


- Some integrals are considerably more complicated and a single change of variables does not help: VEGAS is bound to fail





Multi-channel integration



- Some integrals are considerably more complicated and a single change of variables does not help: VEGAS is bound to fail

- Solution: use different transformations for each of the "structures" = channels

$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x) \qquad \text{with} \qquad \sum_{i=1}^{n} \alpha_i = 1$$

$$I = \int f(x) dx = \sum_{i=1}^{n} \alpha_i \int \frac{f(x)}{p(x)} p_i(x) dx$$

Problem

- We still need to know a lot about the integrand: We need to know where the possible peaks are, how to disentangle them in separate channels and how to map each channel so that peaks are aligned to integration variables
- MadEvent's solution:
  - use the Feynman diagrams themselves

Single diagrams enhanced multi-channel

- Consider the PS integration of an amplitude $|M|^2$ at the tree level with lots of diagrams contributing to it. If there were a basis of functions
  $f(x) = \sum_i f_i(x)$
  such that
  - We know how to integrate each f_i
  - They describe all possible peaks

  Then we would have solved the problem
- Such a basis exists!

$$f_i = \frac{|A_i|^2}{\sum_j |A_j|^2} |A_{\rm tot}|^2$$

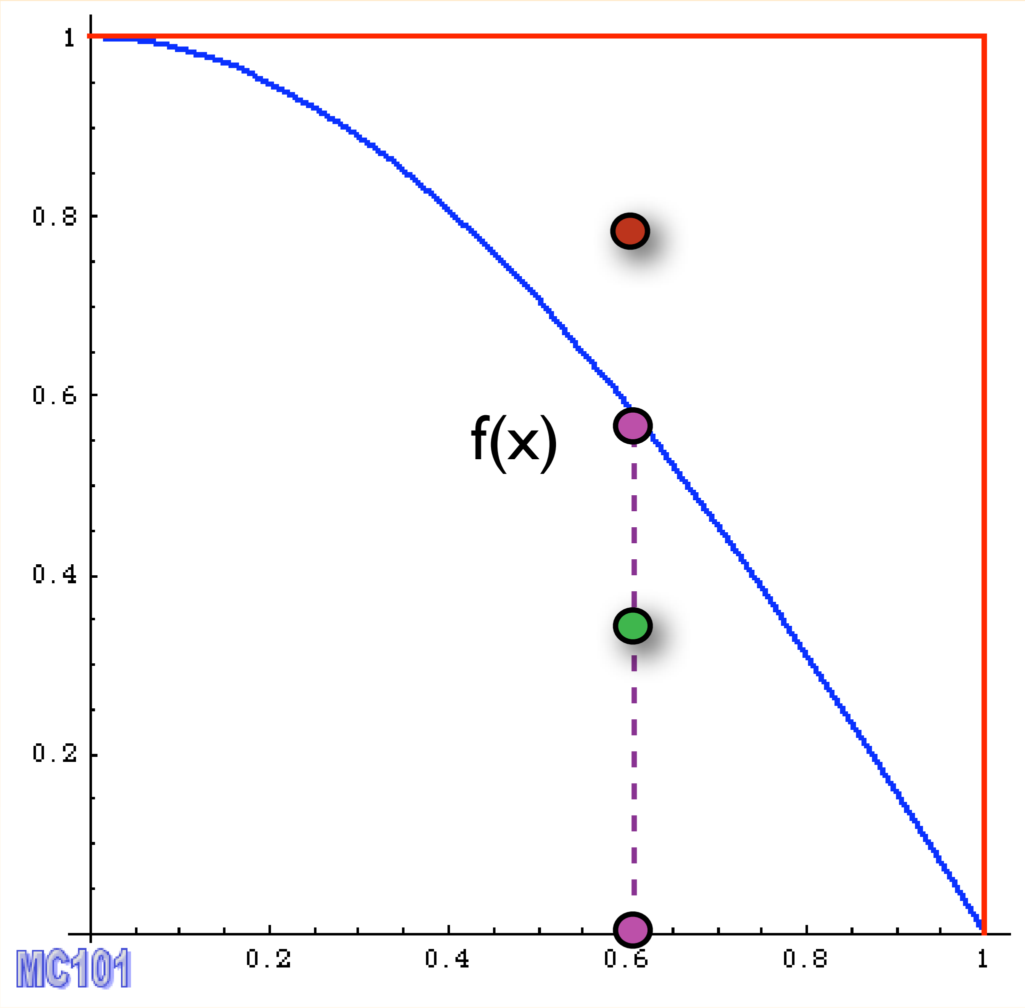
$$f_i = \frac{|A_i|^2}{\sum_j |A_j|^2} |A_{\rm tot}|^2$$

- Key idea:
  - Any single diagram is easy to integrate (peak structure based on propagators)
  - Divide integration into pieces, based on single diagrams
- Get N independent integrals
  - Error added in quadrature (so no extra cost)
  - No need to calculate "weight" function from other channels
  - Can optimize # of points for each independently
  - Parallel in nature
- What about interference?
  - Never creates "new" peaks, so we're okay.

Event generation

- Every phase-space point computed in this way, can be seen as an event (=collision) in a detector
- However, they still carry the "weight" of the matrix elements:
  - events with large weights where the cross section is large
  - events with small weights where the cross section is small
- In nature, the events don't carry a weight:
  - more events where the cross section is large
  - less events where the cross section is small
- How to go from weighted events to unweighted events?

Event generation



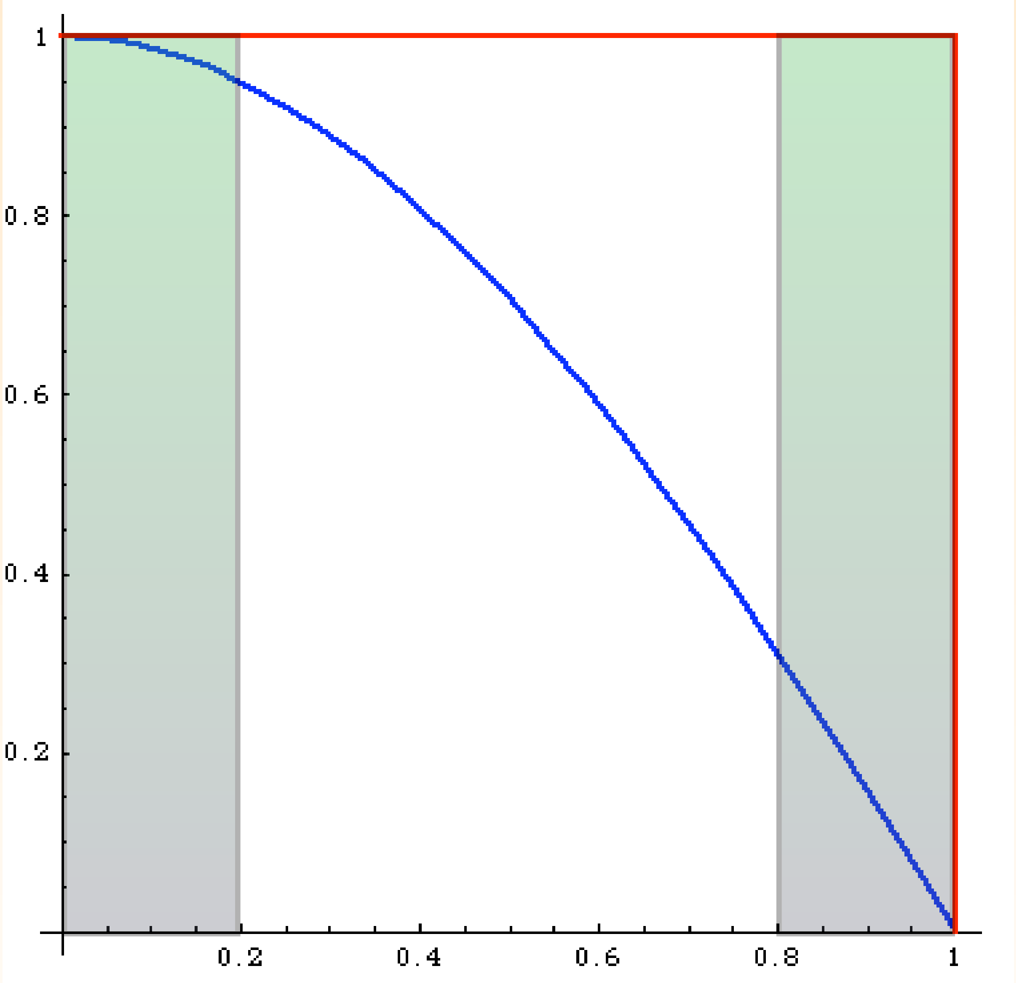
- pick an x at random
- calculate $f(x)$
- pick a y at random $0<y<y_{max}$
- if $f(x)>y$, accept the event, otherwise throw it away.

$$\text{Efficiency: } \frac{\text{Accepted}}{\text{Total tries}}$$

UNWEIGHTED EVENTS

- What's the difference?

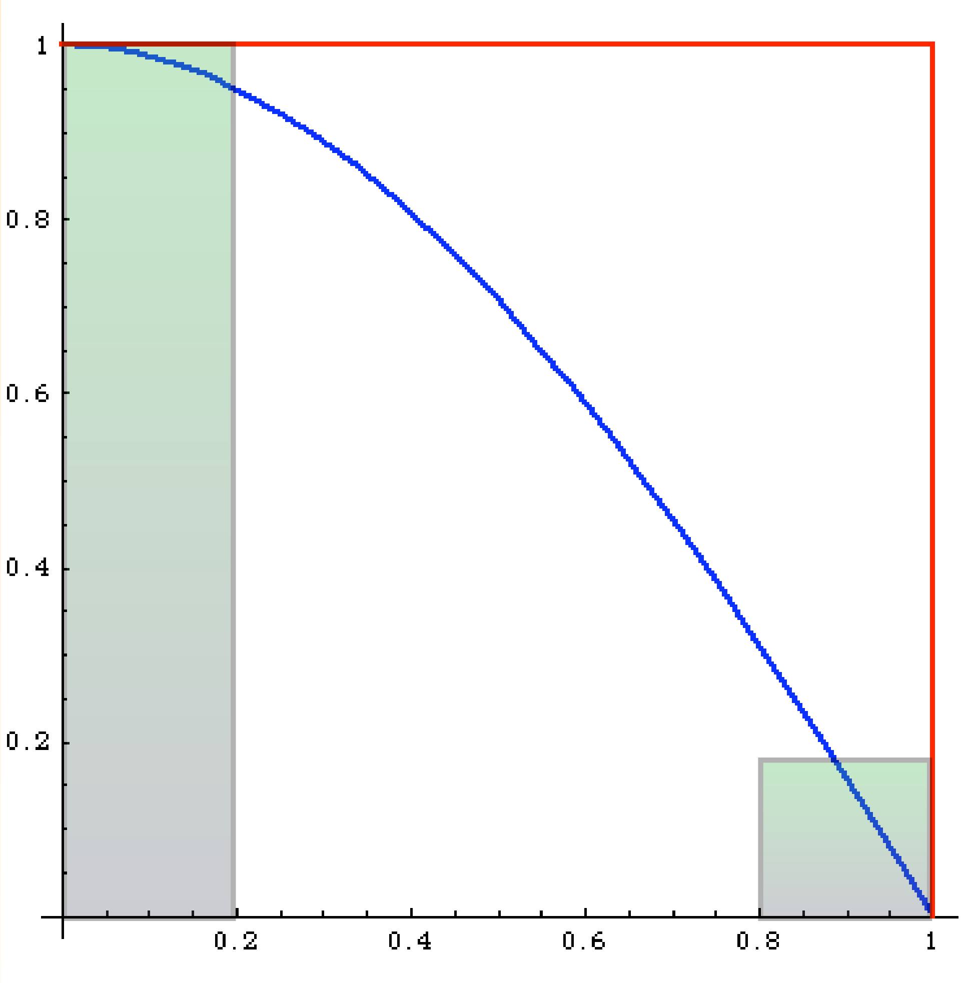
- Before:
  Same number of events in areas of phase-space with different probabilities: events must have different weights



UNWEIGHTED EVENTS

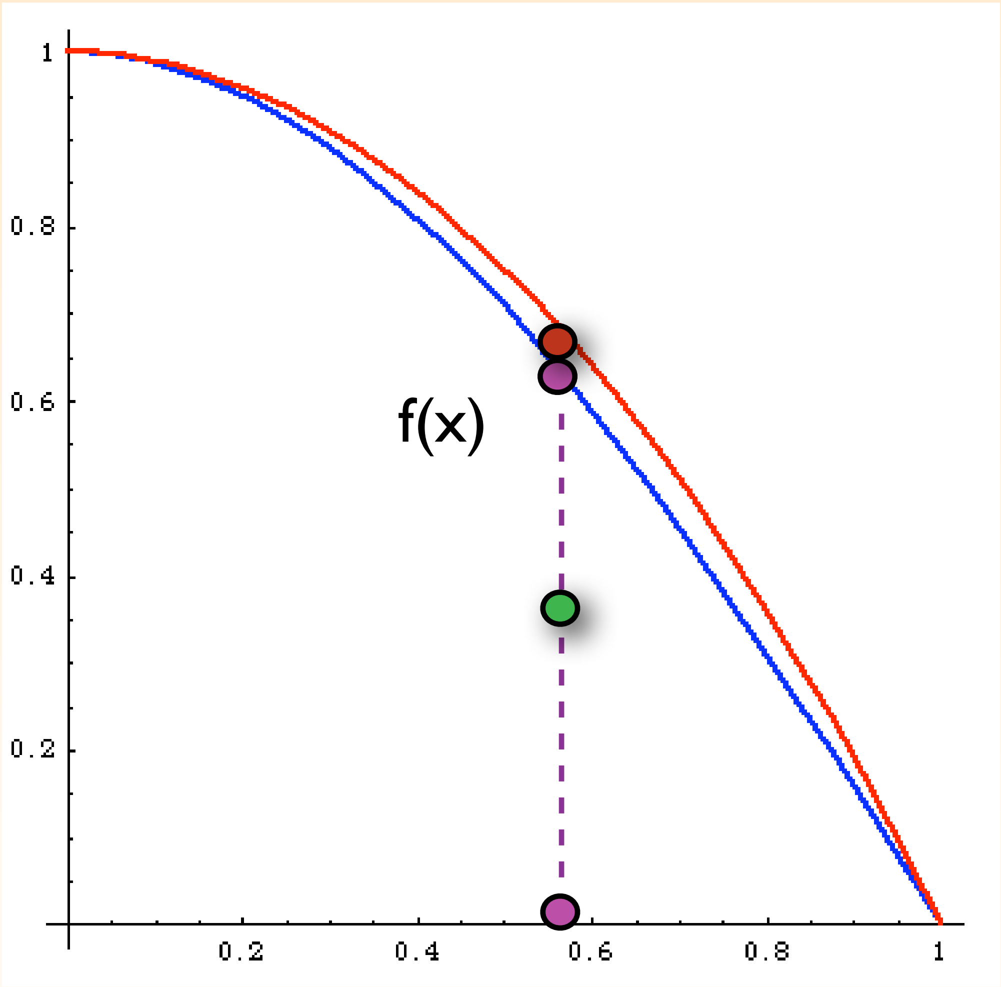
- What's the difference?

- After:
  Number of events is proportional to the probability of areas of phase space: events have all the same weight ("unweighted")



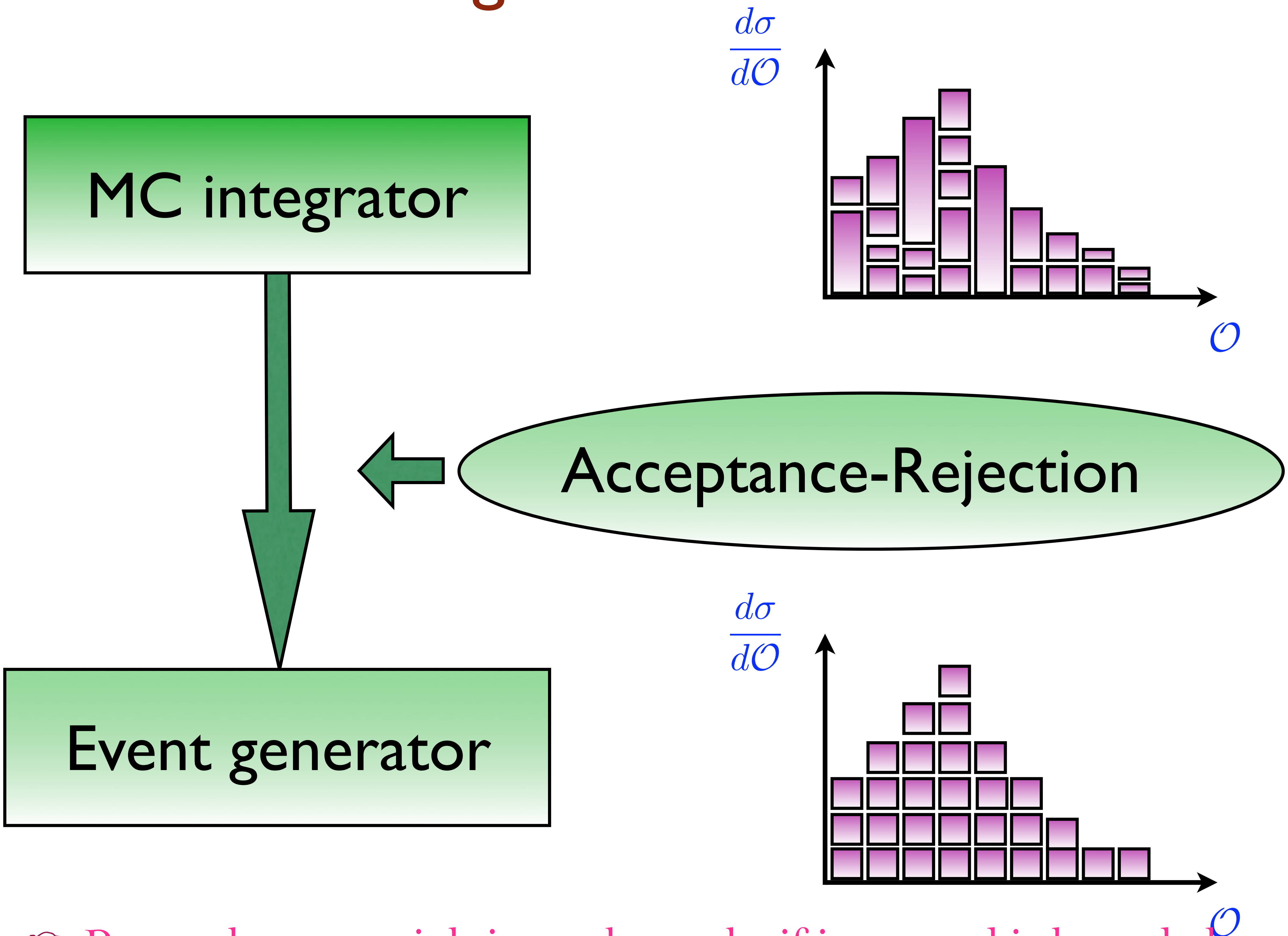
Events distributed as in Nature

IMPROVED EFFICIENCY



- Like before, the more you know about the integrand, the higher the efficiency
- Can use the same techniques as before, i.e. adaptive phase-space integration (VEGAS)

Event generation



☞ Remember, unweighting only works if integrand is bounded

Exercises III

- Download the code for a process generated on-line, untar it in an new, empty directory and look at the ./Cards/param_card.dat and ./Cards/run_card.dat. Do you recognize the inputs?
- Generate some events. Open the (unweighted) events file and try to understand what is in it.
- Go to the 'Tools' section on the webpage. Upload your event file to make some plots and distribution. Try changing the default 'ma_card.dat' to improve your analysis
- Or, download the MadAnalysis package from the download page. Read the in header of the 'ma_card.dat' file and try to generate some plots. (To generate plots from the 'plots.top' file, topdrawer can be used (see the link on the Downloads page to the wiki))

Summary

- How to run MadGraph and generate the diagrams
- How to run MadEvent on-line and get plots and an event-file
- How the phase-space integration is done to get this event-file

- In my next lecture, we'll have a look at
  - some details on the algorithms used to generate the diagrams
  - new physics implementations
  - what future versions will bring us...
- Please, have a look at the exercises!