# Introduction to Event Generators

Lecture 1 of 4

Peter Skands
Monash University
(Melbourne, Australia)

# MELBOURNE?

Australia's

**4**

deadliest animals:

Horses (7/yr)

Cows (3/yr)

Dogs (3/yr)

Roos (2/yr)



Melbourne

**Monash University:**

70,000 students

(Australia's largest uni)

~ 20km SE of Melbourne City Centre

**School of Physics & Astronomy**;

4 HEP theorists

+ post docs & students


Physics Lab

# DISCLAIMER

## This course covers:

Lecture 1: Foundations of MC Generators

Lecture 2: Parton Showers

Lecture 3: Jets and Confinement

Lecture 4: Physics at Hadron Colliders

Supporting Lecture Notes (~80 pages): *"Introduction to QCD"*, arXiv:1207.2389
+ MCnet Review: *"General-Purpose Event Generators"*, Phys.Rept.504(2011)145

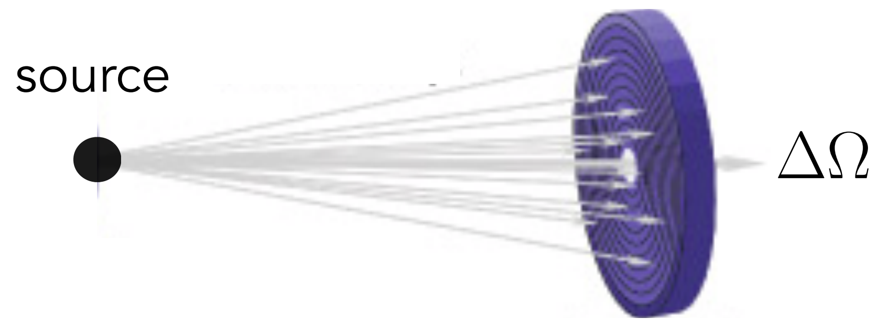
## It does not cover:

**Simulation of BSM physics** → Lectures by V Hirschi
**Matching and Merging** → Lectures by S Höche
**Heavy Ions and Cosmic Rays** → Lectures by K Werner
**Event Generator Tuning** → Lecture by H Schulz
+ many other (more specialised) topics such as: heavy quarks, hadron and τ decays, exotic hadrons, lattice QCD, spin/polarisation, low-x, elastic, …

# CONTENTS



**Scattering Experiments:**




LHC detector
Cosmic-Ray detector
Neutrino detector
X-ray telescope
…

→ Integrate differential **cross sections** over specific **phase-space** regions

Predicted number of counts
= integral over solid angle

$$N_{\mathrm{count}}(\Delta\Omega) \propto \int_{\Delta\Omega} \mathrm{d}\Omega \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$$

$$\mathrm{d}\Omega = \mathrm{d}\cos\theta\,\mathrm{d}\phi$$

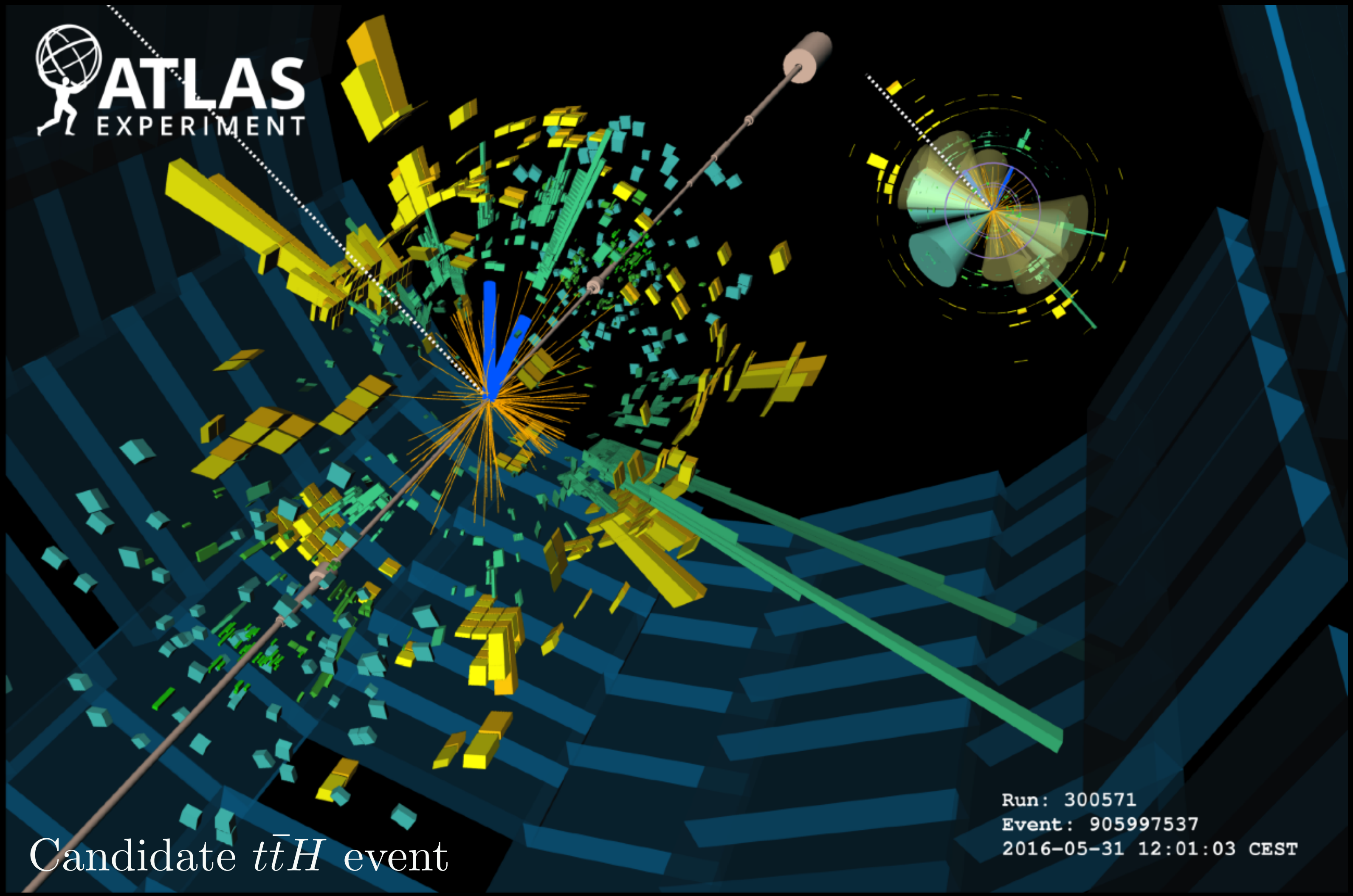
**In particle physics:**
Integrate over all quantum histories
(+ interferences)

# d$\sigma$/d$\Omega$; how hard can it be?

If event generators could talk:

Someone hold my drink while I approximate the amplitude (squared) for this …

(to all orders, + non-perturbative effects)




ATLAS-PHOTO-2016-014-13

… integrate it over a ~300-dimensional phase space

… and estimate the detector response

# Q C D

in Event Generators

$$\mathcal{L} = \bar{\psi}_q^i (i\gamma^\mu)(D_\mu)_{ij}\psi_q^j - m_q \bar{\psi}_q^i \psi_{qi} - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}$$

## Quark fields

$$\psi_q^j = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}$$

L invariant under

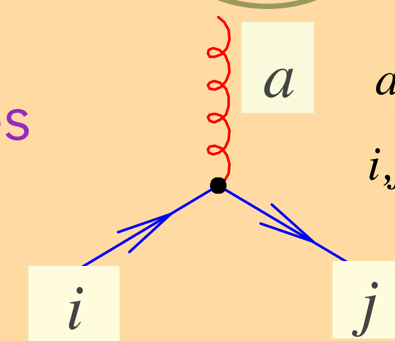
$\psi \to U\psi$

SU(3)
Local Gauge
Symmetry

## Covariant Derivative

$$D_{ij}^\mu = \delta_{ij}\partial^\mu - i g_s t^a_{ij} A^\mu$$

$\Rightarrow$ Feynman rules




$a \in [1,8]$

$i,j \in [1,3]$

## Gell-Mann Matrices *($t^a = ½\lambda^a$)*

(Traceless and Hermitian)

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{-2}{\sqrt{3}} \end{pmatrix}$$

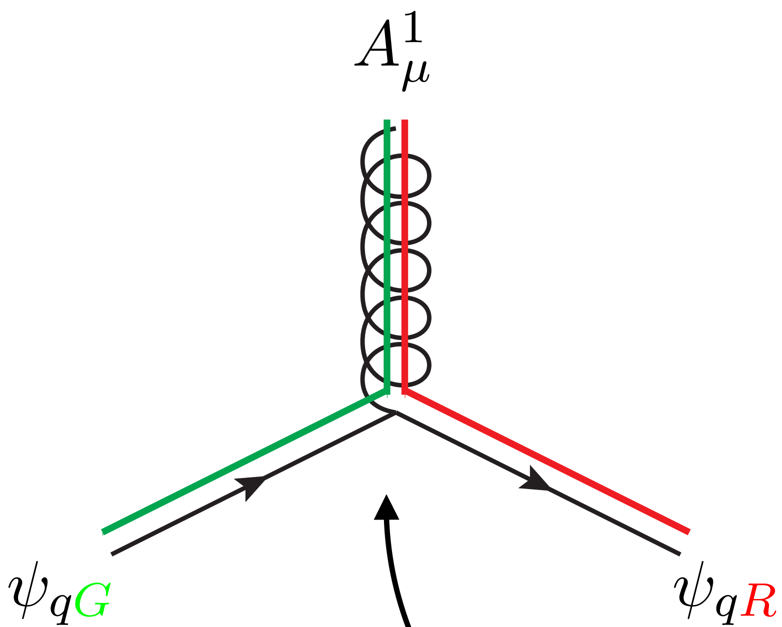
# INTERACTIONS IN COLOUR SPACE

## A quark-gluon interaction

(= one term in sum over colours)

$$\bar{\psi}^i_q(i\gamma^\mu)(D_\mu)_{ij}\psi^j_q$$

$$D^\mu_{ij} = \delta_{ij}\partial^\mu - i g_s t^a_{ij} A^\mu$$




$$\propto \quad -\tfrac{i}{2}g_s \qquad \bar{\psi}_{qR} \qquad \lambda^1 \qquad \psi_{qG}$$

$$= \quad -\tfrac{i}{2}g_s \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

gluon (adjoint) colour index ∈ [1,8]

gluon Lorentz-vector index ∈ [0,3]

$$-i\, g_s\, t^1_{ij}\, \gamma^\mu_{\alpha\beta}\, A^1_\mu \quad -i\, g_s\, t^2_{ij}\, \gamma^\mu_{\alpha\beta}\, A^2_\mu \quad - \ldots$$

fermion colour indices ∈ [1,3]

fermion spinor indices ∈ [1,4]

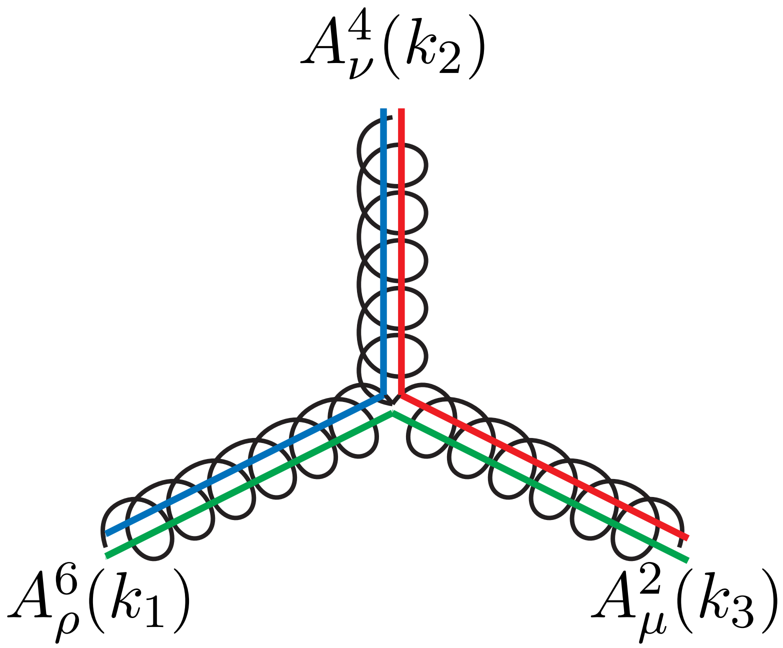
Amplitudes Squared summed over colours → traces over *t* matrices
→ ***Colour Factors*** (see literature, or back of these slides)

## A gluon-gluon interaction

(no equivalent in QED)

$$-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}$$

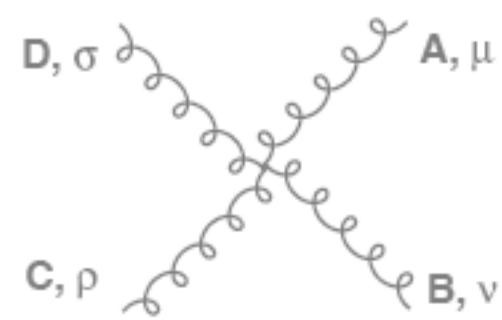
$$F^a_{\mu\nu} = \underbrace{\partial_\mu A^a_\nu - \partial_\nu A^a_\mu}_{\text{Abelian}} + \underbrace{g_s f^{abc} A^b_\mu A^c_\nu}_{\text{non-Abelian}}$$




$$\propto \quad -g_s\, f^{246} \quad [(k_3-k_2)^\rho g^{\mu\nu} + (k_2-k_1)^\mu g^{\nu\rho} + (k_1-k_3)^\nu g^{\rho\mu}]$$

$$i f^{abc} = 2\mathrm{Tr}\{t^c[t^a,t^b]\}$$

(there is also a 4-gluon vertex, proportional to $g_s^2$)




$-ig_s^2 f^{XAC} f^{XBD}[g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\gamma}] + (C,\gamma) \leftrightarrow (D,\rho) + (B,\nu) \leftrightarrow (C,\gamma)$

**Structure Constants of SU(3)**

$$f_{123} = 1 \tag{14}$$

$$f_{147} = f_{246} = f_{257} = f_{345} = \frac{1}{2} \tag{15}$$

$$f_{156} = f_{367} = -\frac{1}{2} \tag{16}$$

$$f_{458} = f_{678} = \frac{\sqrt{3}}{2} \tag{17}$$

Antisymmetric in all indices

All other $f_{abc} = 0$

Amplitudes Squared summed over colours → traces over *t* matrices

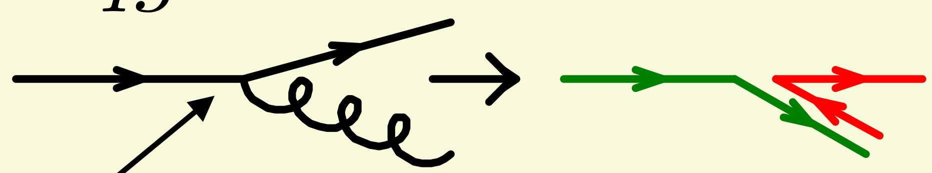
→ ***Colour Factors*** (see literature, or back of these slides)

## MC generators use a simple set of rules for "colour flow"

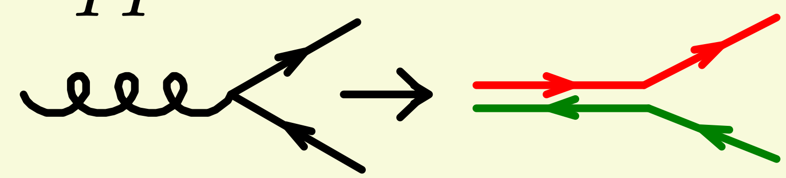
Based on **"Leading Colour"** $8 = 3 \otimes \overline{3} \ominus 1$ ($\Rightarrow$ valid to ~ $1/N_C^2$ ~ 10%)

**LC: represent gluons as outer products of triplet and antitriplet**

$q \to qg$

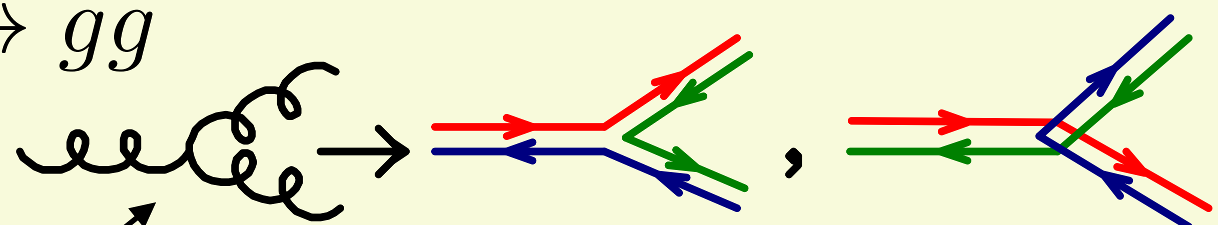


$g \to q\bar{q}$



Illustrations from PDG Review on *MC Event Generators*

$g \to gg$



➜ Lecture 2

"Strong Ordering", $\alpha_s(p_\perp)$, "Coherence", "Recoils" [(E,p) cons.]

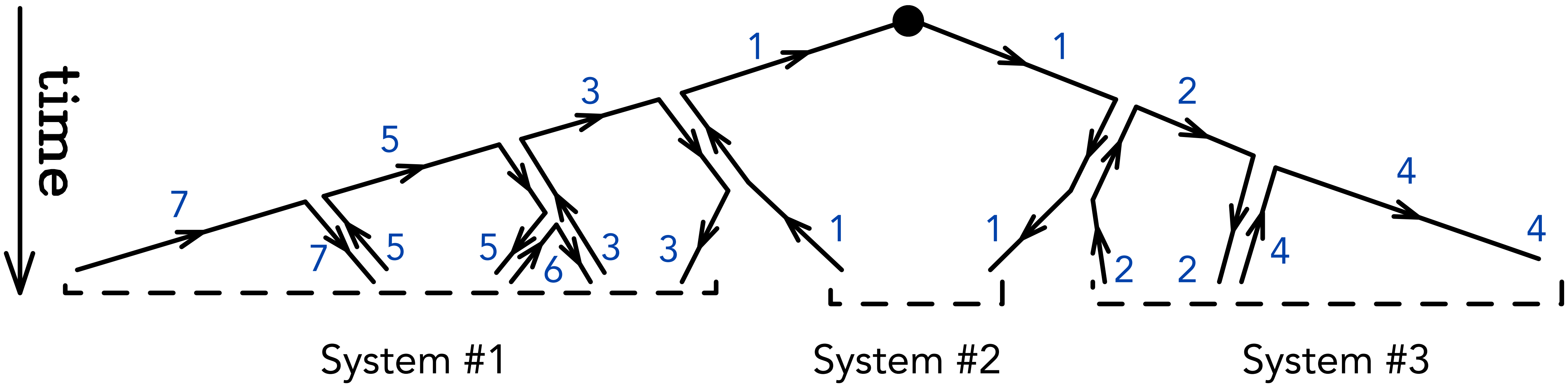
+ Mass effects: *t, b, (c?) quarks, coloured resonances;* Spin effects *(J cons; polarisation; spin correlations);* Corrections beyond LC (or LL)

# COLOUR FLOW

Showers (can) generate lots of partons, $\mathcal{O}$(10-100).

Colour Flow used to determine ***between which partons confining potentials arise***

Example: $Z^0 \rightarrow q\bar{q}$




Coherence of pQCD cascades → suppression of "overlapping" systems
→ Leading-colour approximation pretty good

(LEP measurements in $e^+e^- \rightarrow W^+W^- \rightarrow$ hadrons confirm this (at least to order 10% ~ $1/N_c^2$))

**Note**: (much) more color getting kicked around in hadron collisions.
Signs that LC approximation is breaking down? → Lecture 4

## Bjorken scaling:

To first approximation, QCD is SCALE INVARIANT (a.k.a. conformal)

Jets inside jets inside jets …

Fluctuations (loops) inside fluctuations inside fluctuations …

If the strong coupling didn't "run", this would be absolutely true (e.g., N=4 Supersymmetric Yang-Mills)

Since $\alpha_s$ only runs slowly (logarithmically) → can still gain insight from fractal analogy (→ lecture 2 on showers)

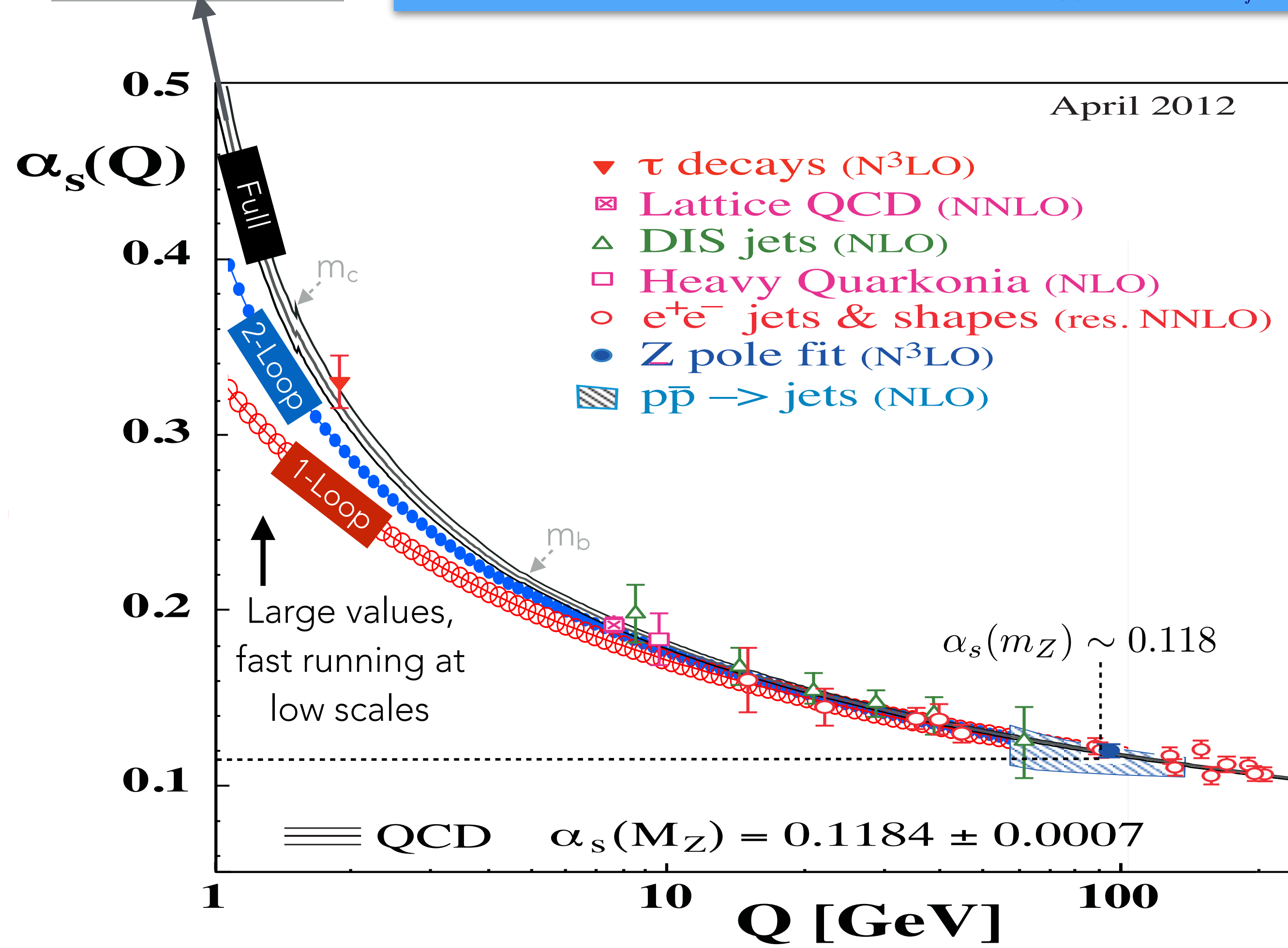
$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \beta(\alpha_s)$$

$$= -\alpha_s^2 (\boxed{b_0} + b_1 \alpha_s + b_2 \alpha_s^2 + \ldots)$$

Asymptotic Freedom

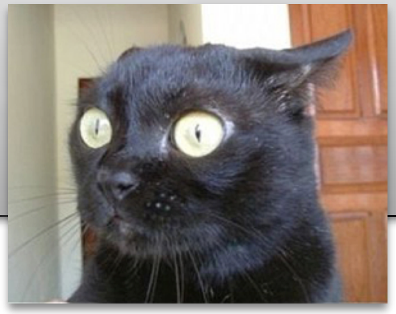
1-Loop β function coefficient: $b_0 = \frac{11 C_A - 2 n_f}{12\pi} > 0$ for $n_f \leq 16$

Landau Pole at $\Lambda_{QCD}$~200 MeV




Note: I use the terms "conformal" and "scale invariant" interchangeably
Strictly speaking, conformal (angle-preserving) symmetry is more restrictive than just scale invariance

# MANY WAYS TO SKIN A CAT

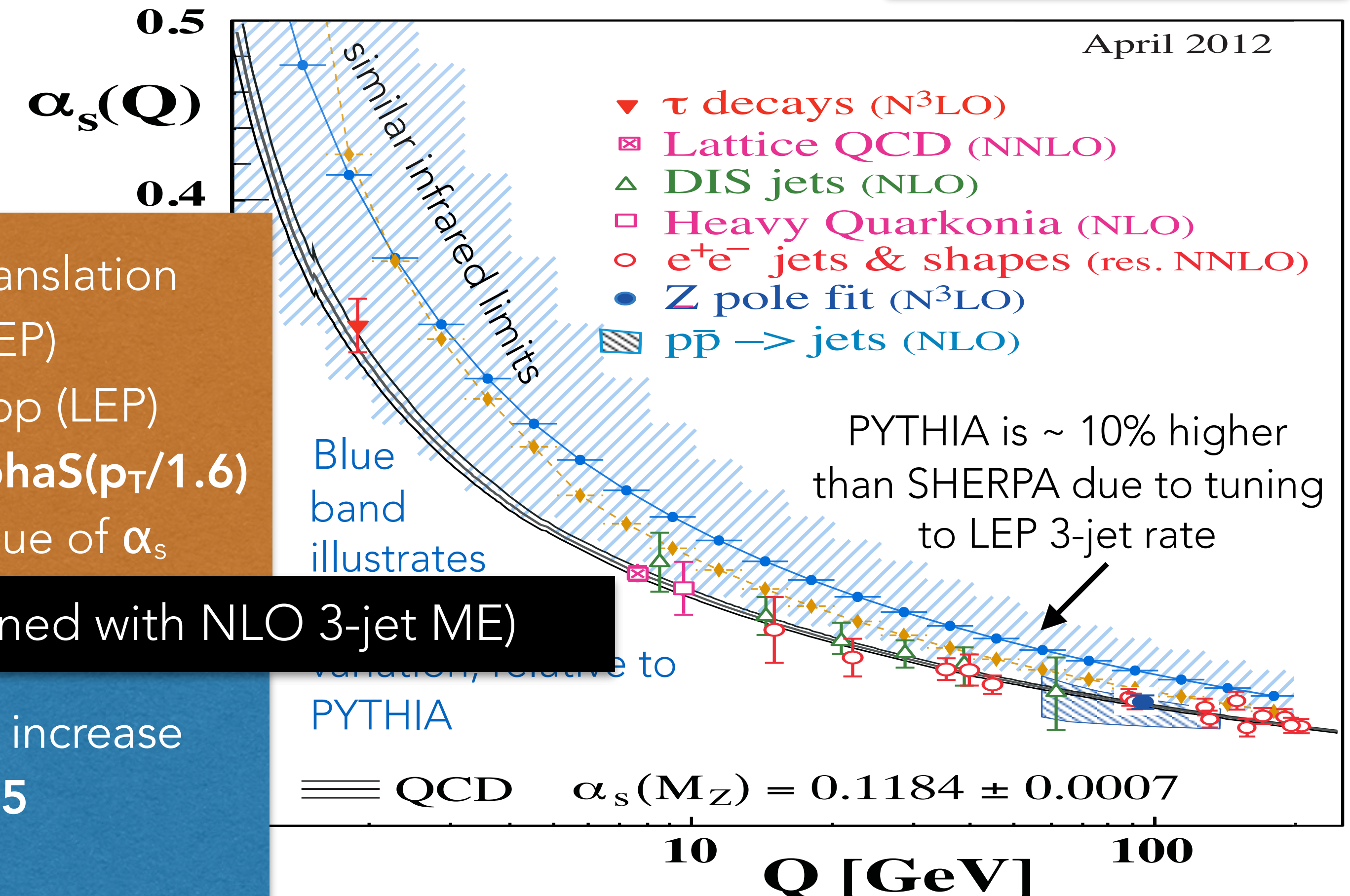


The strong coupling is (one of) the main perturbative parameter(s) in event generators. It controls:

- The overall amount of QCD initial- and final-state radiation
- Strong-interaction cross sections (and resonance decays)
- The rate of (mini)jets in the underlying event

MCs: get value from: PDG? PDFs? Fits to data (tuning)?

**Example (for Final-State Radiation):**




(also note: definitions of $Q=p_T$ not exactly the same)

**SHERPA** : uses PDF or PDG value, with "CMW" translation
**alphaS($m_Z$)** default = **0.118** (pp) or 0.1188 (LEP)
**running order:** default = **3-loop** (pp) or 2-loop (LEP)
**CMW scheme translation:** default **use ~ alphaS($p_T$/1.6)**
→ roughly 10% increase in the effective value of $\alpha_s$

will undershoot LEP 3-jet rate by ~ 10% (unless combined with NLO 3-jet ME)

**PYTHIA** : tuning to LEP 3-jet rate; requires ~ 20% increase
**TimeShower:alphaSvalue** default = **0.1365**
**TimeShower:alphaSorder** default = **1**
**TimeShower:alphaSuseCMW** default = **off**

Agrees with LEP 3-jet rate "out of the box"; but no guarantee tuning is universal.

## Scale variation ~ uncertainty; why?

Scale dependence of calculated orders must be canceled by contribution from uncalculated ones (+ non-pert)

$$\alpha_s(Q^2) = \alpha_s(m_Z^2)\frac{1}{1 + b_0\;\alpha_s(m_Z)\ln\frac{Q^2}{m_Z^2} + \mathcal{O}(\alpha_s^2)} \qquad b_0 = \frac{11N_C - 2n_f}{12\pi}$$

→ $\alpha_s(Q_1^2) - \alpha_s(Q_2^2) \; = \; \alpha_s^2\, b_0 \ln(Q_2^2/Q_1^2) \; + \; \mathcal{O}(\alpha_s^3)$

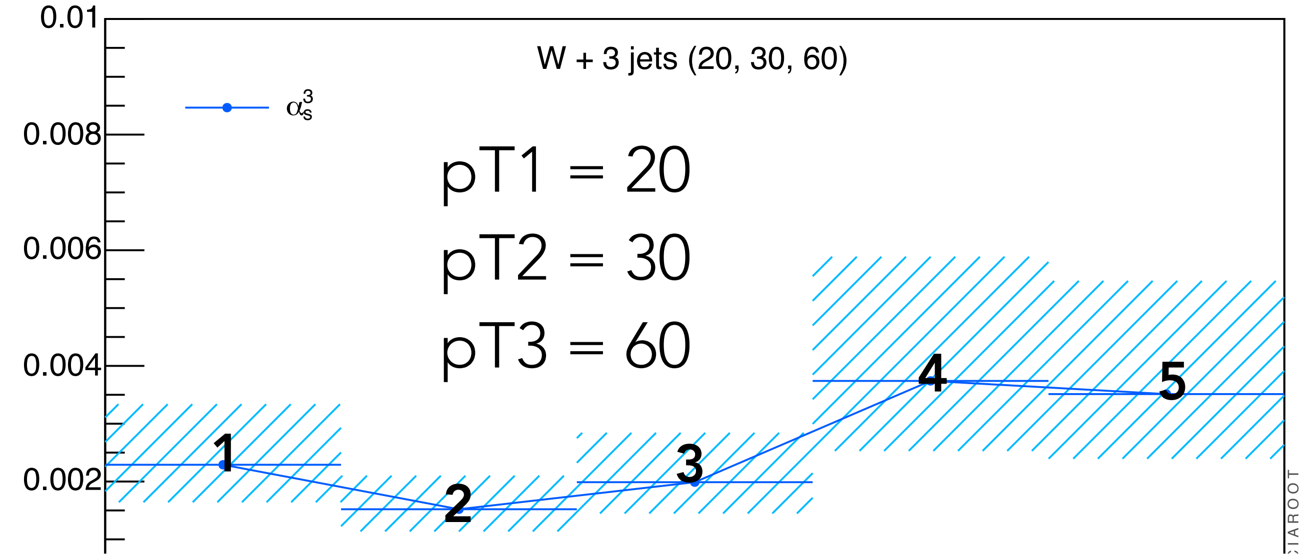
→ Generates terms of higher order, proportional to what you already have ($|M|^2$)→ a first naive* way to estimate uncertainty

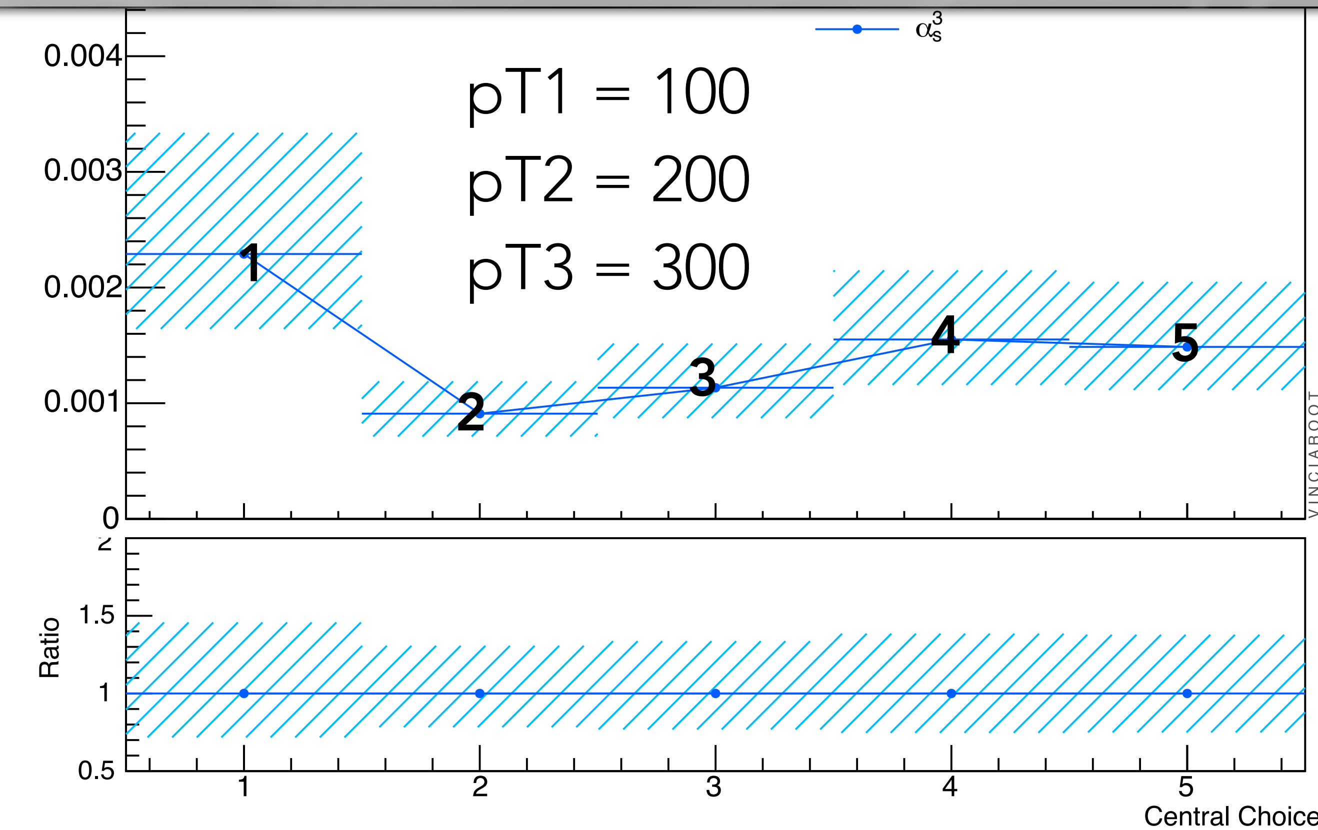
*warning: some believe it is the only way … but be agnostic! Really a lower limit. There are other things than scale dependence …

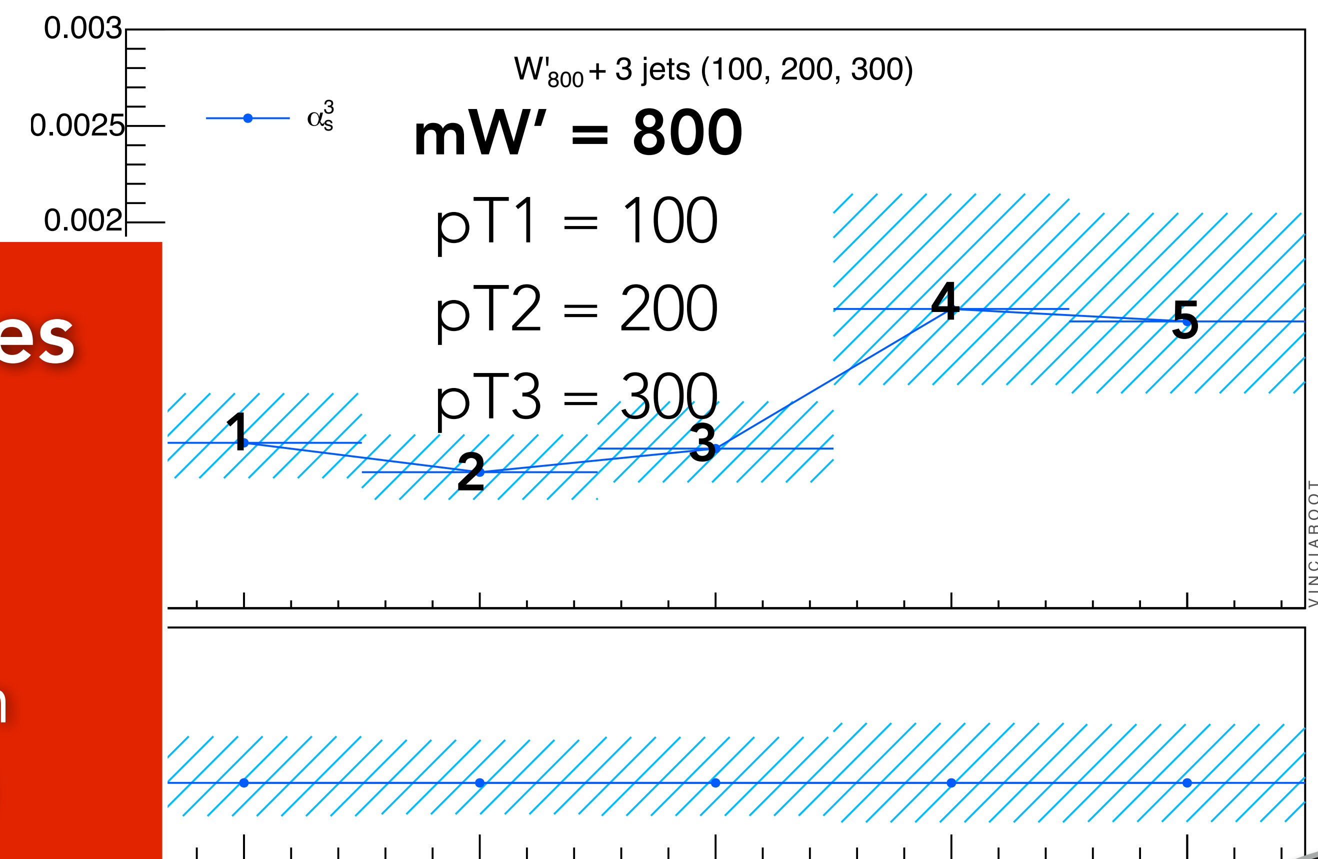
# Example: pp → W + 3 jets

Some choices for $\mu_R$

1: MW
2: MW + Sum(|pT|)
3: -"- (quadratically)
4: Geometric mean pT (~shower)
5: Arithmetic mean pT




**If you have multiple QCD scales**

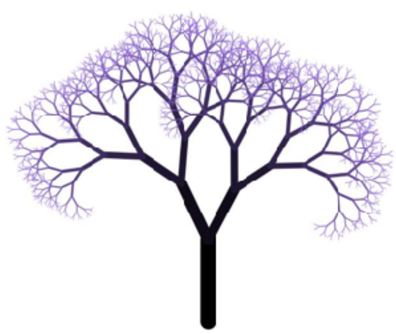
→ variation of $\mu_R$ by factor 2 in each direction not exhaustive!

Also consider functional dependence on each scale (+ N$^{(n)}$LO → some compensation)

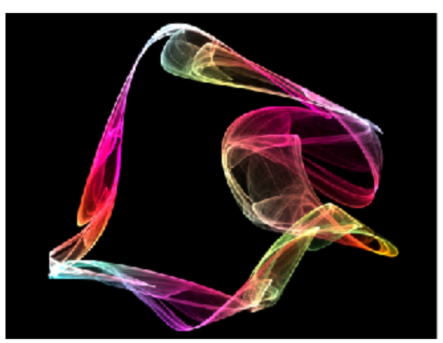
# BEYOND FIXED ORDER

QCD is more than just a perturbative expansion in $\alpha_s$

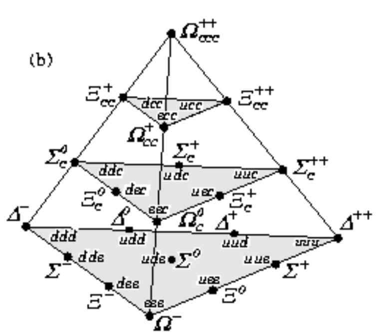
(and Perturbation theory is more than Feynman diagrams)



**Jets** $\longleftrightarrow$ amplitude structures $\longleftrightarrow$ fundamental quantum field theory / gauge theory. Precision jet (structure) studies. ➜ Lecture 2



**Strings** (strong gluon fields) $\longleftrightarrow$ quantum-classical correspondence. String physics. Dynamics of confinement / hadronisation phase transition. ➜ Lecture 3



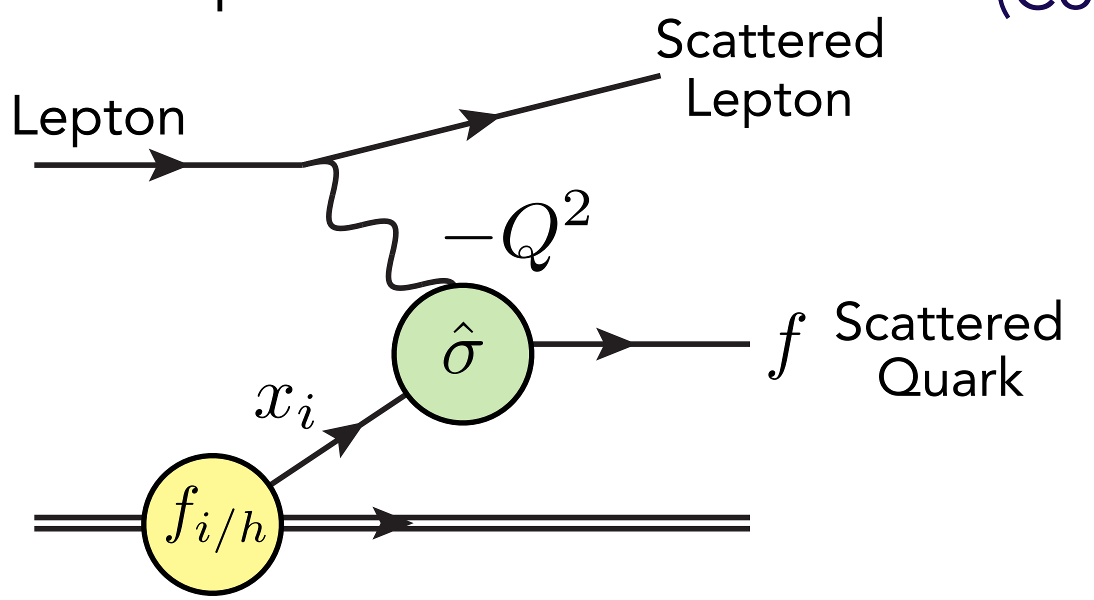
**Hadrons** $\longleftrightarrow$ Spectroscopy (incl excited and exotic states), lattice QCD, (rare) decays, mixing, light nuclei. Hadron beams → MPI, diffraction, … ➜ Lecture 4

**Factorisation** ⇛ Fixed-order cross sections still useful.

In DIS, there is a formal proof (Collins, Soper, 1987)

Deep Inelastic Scattering (DIS)

(By "deep", we mean $Q^2 >> M_h^2$)




Note: Beyond LO, $f$ can be more than one parton

→ We really *can* write the cross section in factorised form :

$$\sigma^{\ell h} = \sum_i \sum_f \int dx_i \int d\Phi_f \, f_{i/h}(x_i, Q_F^2) \, \frac{d\hat{\sigma}^{\ell i \to f}(x_i, \Phi_f, Q_F^2)}{dx_i \, d\Phi_f}$$

Sum over Initial (i) and final (f) parton flavors

$\Phi_f$ = Final-state phase space

$f_{i/h}$ = PDFs Assumption: $Q^2 = Q_F^2$

Differential partonic Hard-scattering Matrix Element(s)

# A PROPOS FACTORISATION

**Why do we need PDFs, parton showers / jets, etc.?**
**Why are Fixed-Order QCD matrix elements not enough?**

F.O. QCD requires **Large scales** ⇛ $\alpha_s$ small enough to be perturbative
(⇢ cannot be used to address intrinsically soft physics such as minimum-bias or diffraction, but still OK for high-scale/hard processes)

F.O. QCD requires **No scale hierarchies** ⇛ $\alpha_s \ln(Q_i/Q_j)$ small
In the presence of scale hierarchies, propagator singularities integrate to logarithms (tomorrow's lecture) which ruin fixed-order expansion.

**But!!!** we collide - and observe - hadrons, with *non-perturbative structure*, that participate in hard processes, whose scales are *hierarchically greater* than $m_{had} \sim 1$ GeV.

→ A Priori, no perturbatively calculable observables in QCD

MCnet

# FACTORISATION ⇛ WE CAN STILL CALCULATE!

**Why is Fixed Order QCD not enough?**

: It requires all resolved scales ≫ $\Lambda_{QCD}$ AND no large hierarchies

**PDFs:** connect incoming hadrons with the high-scale process

**Fragmentation Functions:** connect high-scale process with final-state hadrons

(each is a non-perturbative function modulated by initial- and final-state radiation)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}X} = \sum_{a,b}\sum_{f}\int_{\hat{X}_f} f_a(x_a,Q_i^2) f_b(x_b,Q_i^2) \frac{\mathrm{d}\hat{\sigma}_{ab\to f}(x_a,x_b,f,Q_i^2,Q_f^2)}{\mathrm{d}\hat{X}_f} D(\hat{X}_f \to X, Q_i^2, Q_f^2)$$

PDFs: needed to compute inclusive cross sections

FFs: needed to compute (semi-)exclusive cross sections

In MCs: made exclusive as **initial-state radiation** + non-perturbative hadron (beam-remnant) structure (+ multiple parton-parton interactions)

In MCs: **resonance decays, final-state radiation**, hadronisation, hadron decays (+ final-state interactions?)

Resummed pQCD: All resolved scales ≫ $\Lambda_{QCD}$ **AND** $X$ Infrared Safe

*)pQCD = perturbative QCD

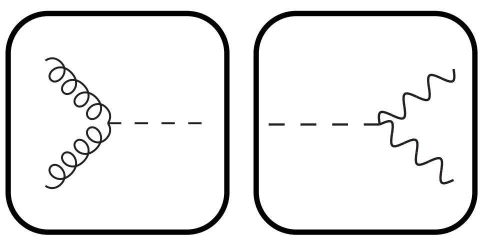
Will take a closer look at both PDFs and final-state aspects (jets and showers) in the next lectures

# ORGANISING THE CALCULATION

**Divide and Conquer** → Split the problem into many (nested) pieces

+ Quantum mechanics → Probabilities → Random Numbers

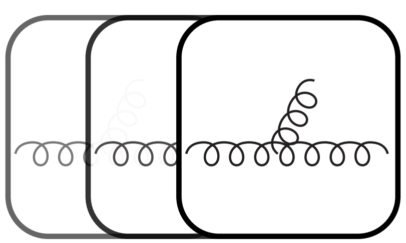
$$\mathcal{P}_{\mathrm{event}} = \mathcal{P}_{\mathrm{hard}} \otimes \mathcal{P}_{\mathrm{dec}} \otimes \mathcal{P}_{\mathrm{ISR}} \otimes \mathcal{P}_{\mathrm{FSR}} \otimes \mathcal{P}_{\mathrm{MPI}} \otimes \mathcal{P}_{\mathrm{Had}} \otimes \ldots$$



## Hard Process & Decays:

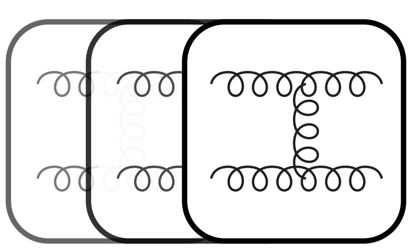
Use process-specific (N)LO matrix elements (e.g., $gg \to H^0 \to \gamma\gamma$)

→ Sets "hard" resolution scale for process: $Q_{MAX}$



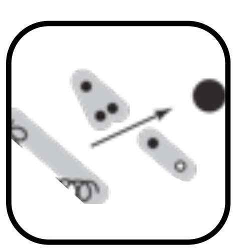
## ISR & FSR (Initial- & Final-State Radiation):

Driven by differential (e.g., DGLAP) evolution equations, $dP/dQ^2$, as function of resolution scale; from $Q_{MAX}$ to $Q_{HAD}$ ~ 1 GeV



## MPI (Multi-Parton Interactions)

Protons contain lots of partons → can have additional (soft) parton-parton interactions → Additional (soft) "Underlying-Event" activity



## Hadronisation

Non-perturbative modeling of partons → hadrons transition

# THE MAIN WORKHORSES

## PYTHIA (begun 1978)

Originated in hadronisation studies: Lund String model

Still significant emphasis on soft/non-perturbative physics

## HERWIG (begun 1984)

Originated in coherence studies: angular-ordered showers

Cluster hadronisation as simple complement

## SHERPA (begun ~2000)

Originated in ME/PS matching (CKKW-L)

Own variant of cluster hadronisation

## + Many more specialised:

Matrix-Element Generators, Matching/Merging Packages, Resummation packages, Alternative QCD showers, Soft-QCD MCs, Cosmic-Ray MCs, Heavy-Ion MCs, Neutrino MCs, Hadronic interaction MCs (GEANT/FLUKA; for energies below $E_{CM}$ ~ 10 GeV), (BSM) Model Generators, Decay Packages, …

# → MONTE CARLO

***MC:*** **any technique that makes use of random sampling** (to provide numerical estimates)

Prescribed for cases of complicated integrands/boundaries in high dimensions

# → MONTE CARLO

***MC:*** **any technique that makes use of random sampling** (to provide numerical estimates)

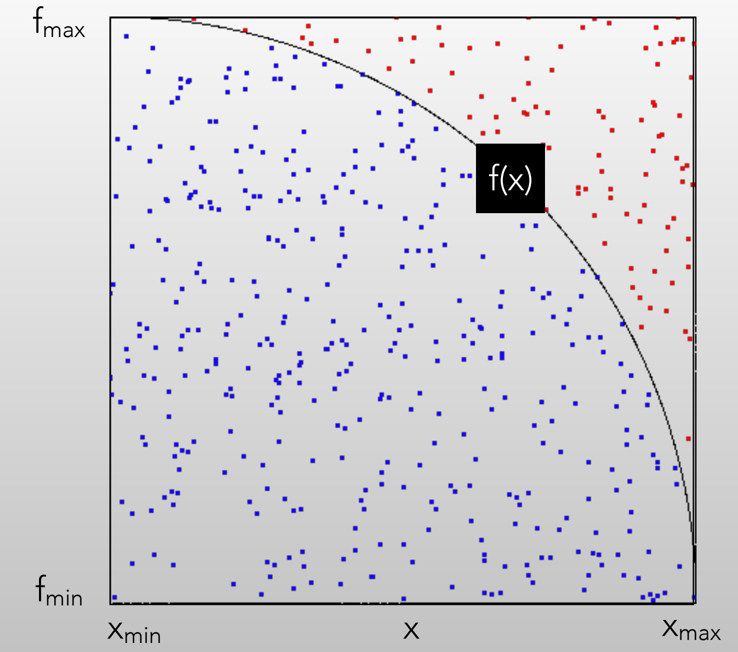
Prescribed for cases of complicated integrands/boundaries in high dimensions

***Example: Integrate f(x)***

1. Compute area of box (you can do it!)
2. Throw random (x,y) points uniformly inside box
3. If y < f(x) : accept **(blue)**; else reject **(red)**
4. After $N_{tot}$ throws, you have an estimate

$$\int_{x_{\mathrm{min}}}^{x_{\mathrm{max}}} f(x)\mathrm{d}x \sim A_{\mathrm{box}}\, N_{\mathrm{blue}}/N_{\mathrm{tot}}$$

5. Central limit theorem ⇛ converges to $A_{blue}$




## Recap Convergence:

**Calculus:** {A} converges to B
if *n* exists for which $|A_{i>n} - B| < \varepsilon$, for any $\varepsilon > 0$

**Monte Carlo:** {A} converges to B
if *n* exists for which
**the probability for** $|A_{i>n} - B| < \varepsilon$,
**is > P, for any P[0<P<1]** for any $\varepsilon > 0$

"This risk, that convergence is only given with a certain probability, is inherent in Monte Carlo calculations and is the reason why this technique was named after the world's most famous gambling casino." [F. James, MC theory and practice]

***MC:*** **any technique that makes use of random sampling** (to provide numerical estimates)

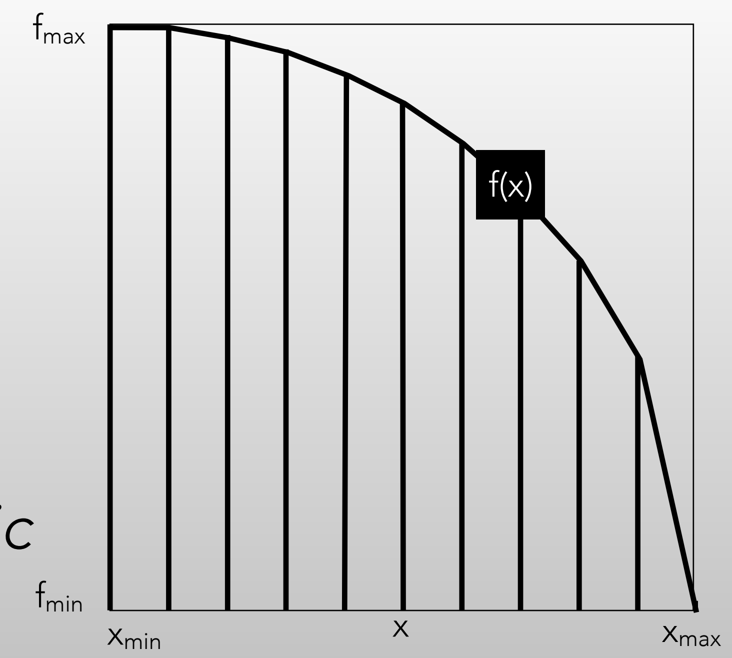
Prescribed for cases of complicated integrands/boundaries in **high dimensions**

### *Example: Integrate f(x)*

Could also have used standard 1D num. int.
(e.g., "Fixed-Grid": Trapezoidal rule, Simpson's rule ...)
→ typically faster convergence in 1D

but few general optimised methods in 2D; none beyond 3D & convergence rate becomes worse ...

*The convergence rate of MC remains the stochastic $1/\sqrt{n}$ independent of dimension* !*




**) You still need to worry about **variance**; physics has lots of peaked/singular functions → adaptive sampling (or stratification)*

| **Numerical Integration:** Relative Uncertainty (after n function evaluations) | $n_{eval}$ / bin | One Dimension Conv. Rate | D Dimensions Conv. Rate |
|---|---|---|---|
| Trapezoidal Rule (2-point) | $2^D$ | $1/n^2$ | $1/n^{2/D}$ |
| Simpson's Rule (3-point) | $3^D$ | $1/n^4$ | $1/n^{4/D}$ |
| Monte Carlo | 1 | $1/n^{1/2}$ | $1/n^{1/2}$ |

*+ optimisations (stratification, adaptation),* ***iterative solutions (Markov-Chain Monte Carlo)***

# JUSTIFICATION:

## MC CAN PROVIDE PERFECT ACCURACY, WITH STOCHASTIC PRECISION

### 1. Law of large numbers (MC is accurate)

For a function, $f$, of random variables, $x_i$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(x_i) = \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x$$

Monte Carlo Estimate — The Integral

**For infinite n:** Monte Carlo is a consistent estimator

(note: in real world, we only deal with *approximations* to Nature's f(x) → less than perfect accuracy)

### 2. Central limit theorem (MC precision is stochastic: 1/√n)

The sum of n independent random variables (of finite expectations and variances) is asymptotically Gaussian

(no matter how the individual random variables are distributed)

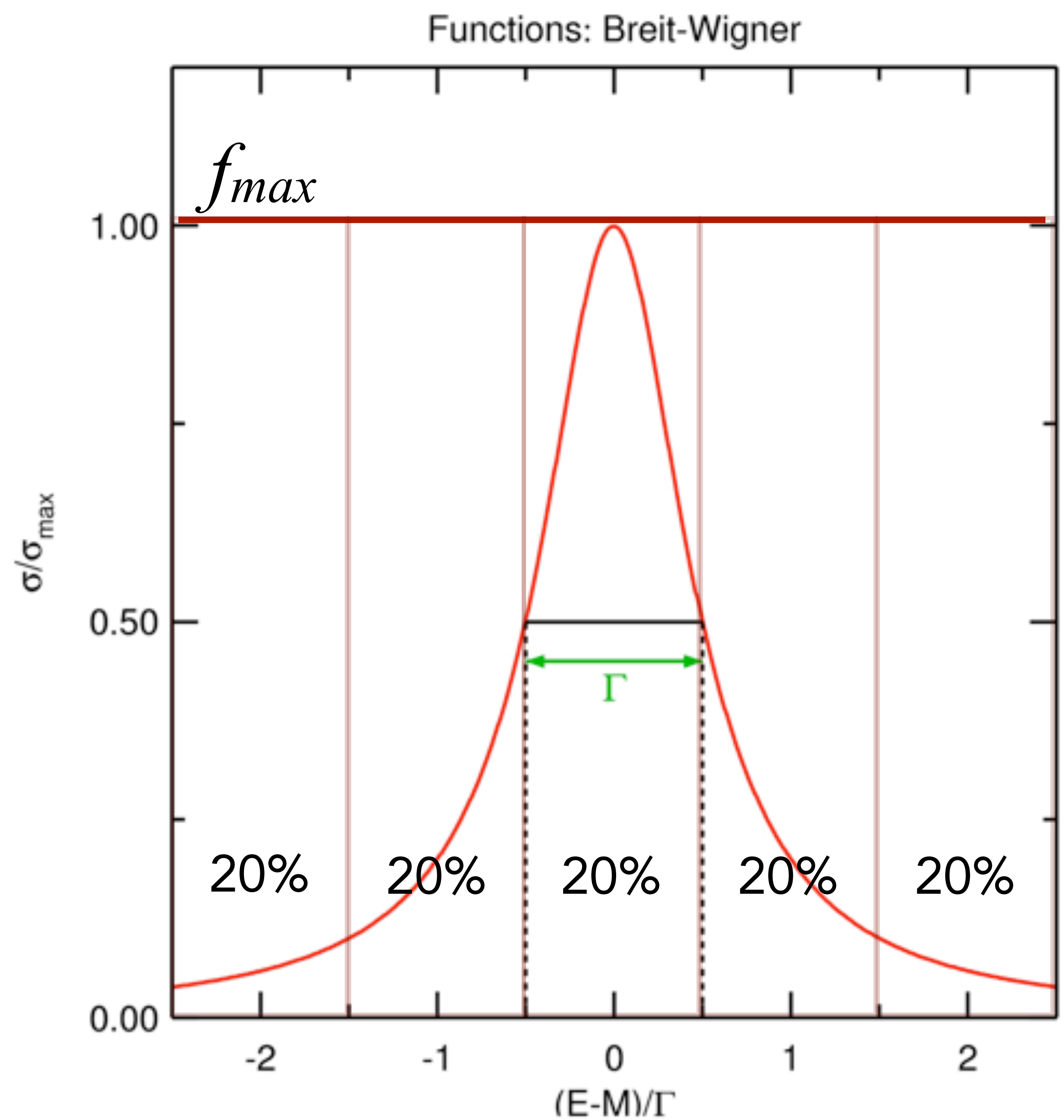
**For finite n:**

The Monte Carlo estimate is Gauss distributed around the true value → with 1/√n precision

In other words: MC stat unc same as for data

Variance $V = (x_2 - x_1)\int_{x_1}^{x_2}\left[f(x)\right]^2 dx - \left[\int_{x_1}^{x_2} f(x)\, dx\right]^2$

# PEAKED FUNCTIONS

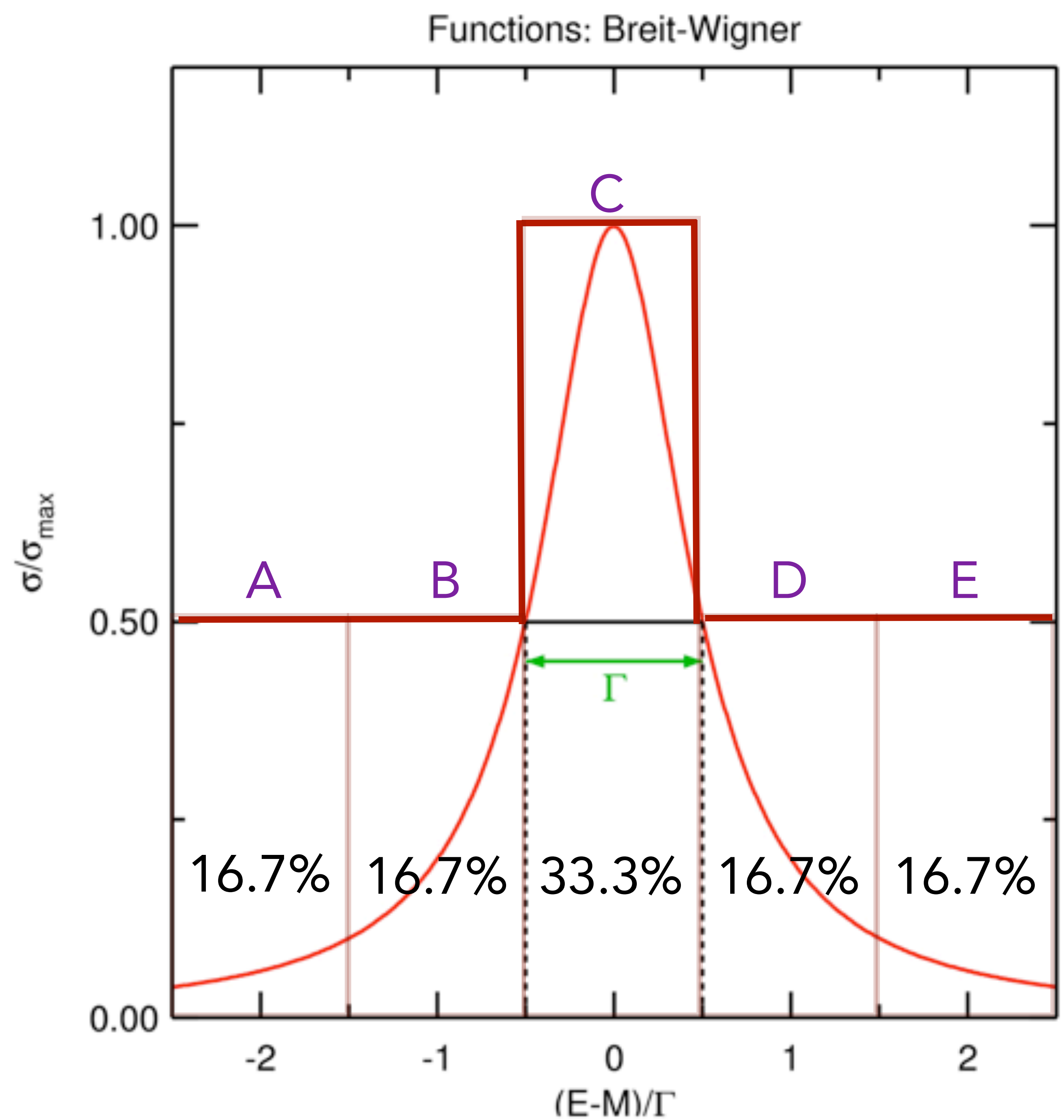



Precision on integral dominated by the points with $f \approx f_{max}$ *(i.e., peak regions)*

→ slow convergence if high, narrow peaks

Variance $V = (x_2 - x_1)\int_{x_1}^{x_2}\left[f(x)\right]^2 dx - \left[\int_{x_1}^{x_2} f(x)\, dx\right]^2$

# STRATIFIED SAMPLING




→ Make it twice as likely to throw points in the peak

Choose:

For:

$6*R_1 \in$ [0,1] → Region A

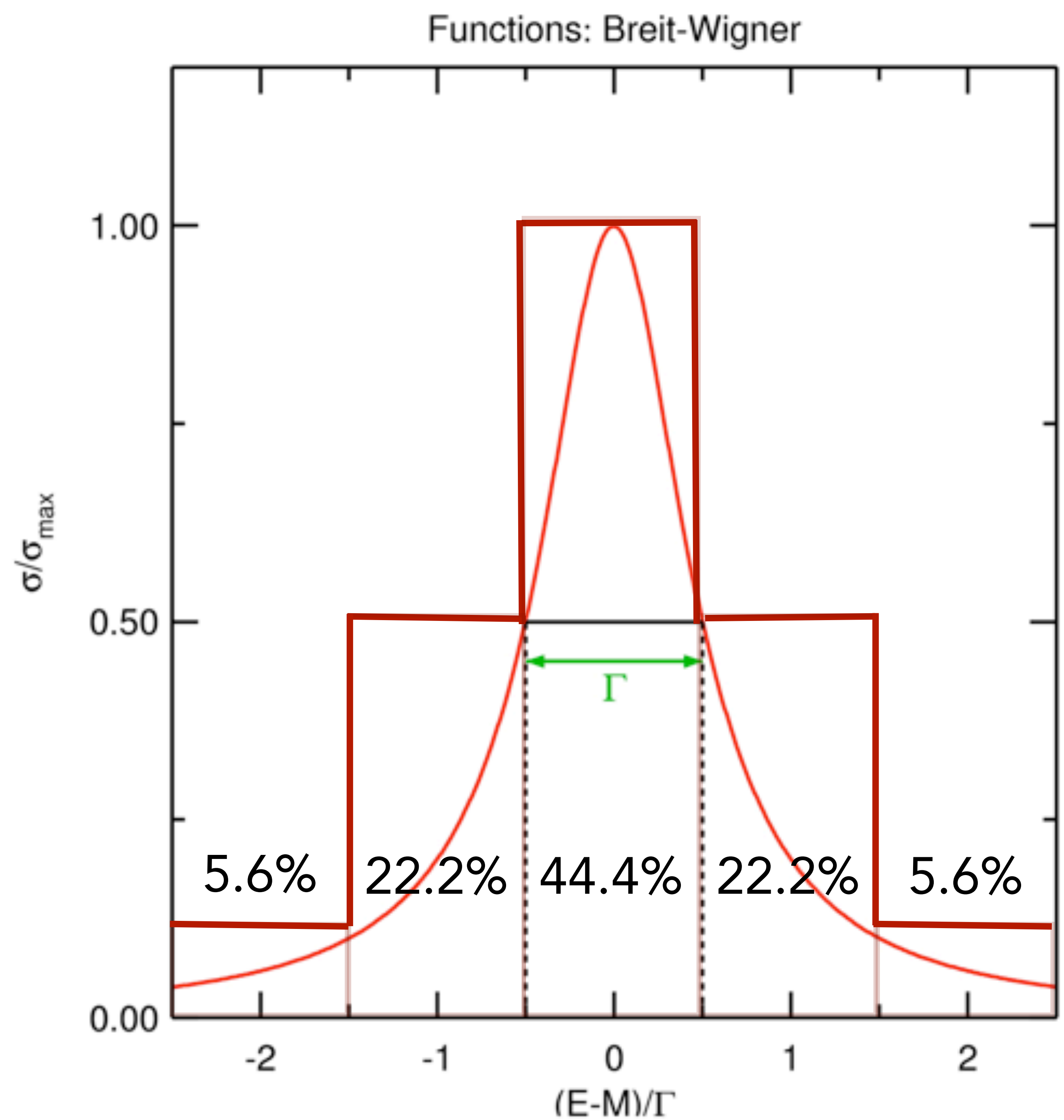
[1,2] → Region B

[2,4] → Region C

[4,5] → Region D

[5,6] → Region E

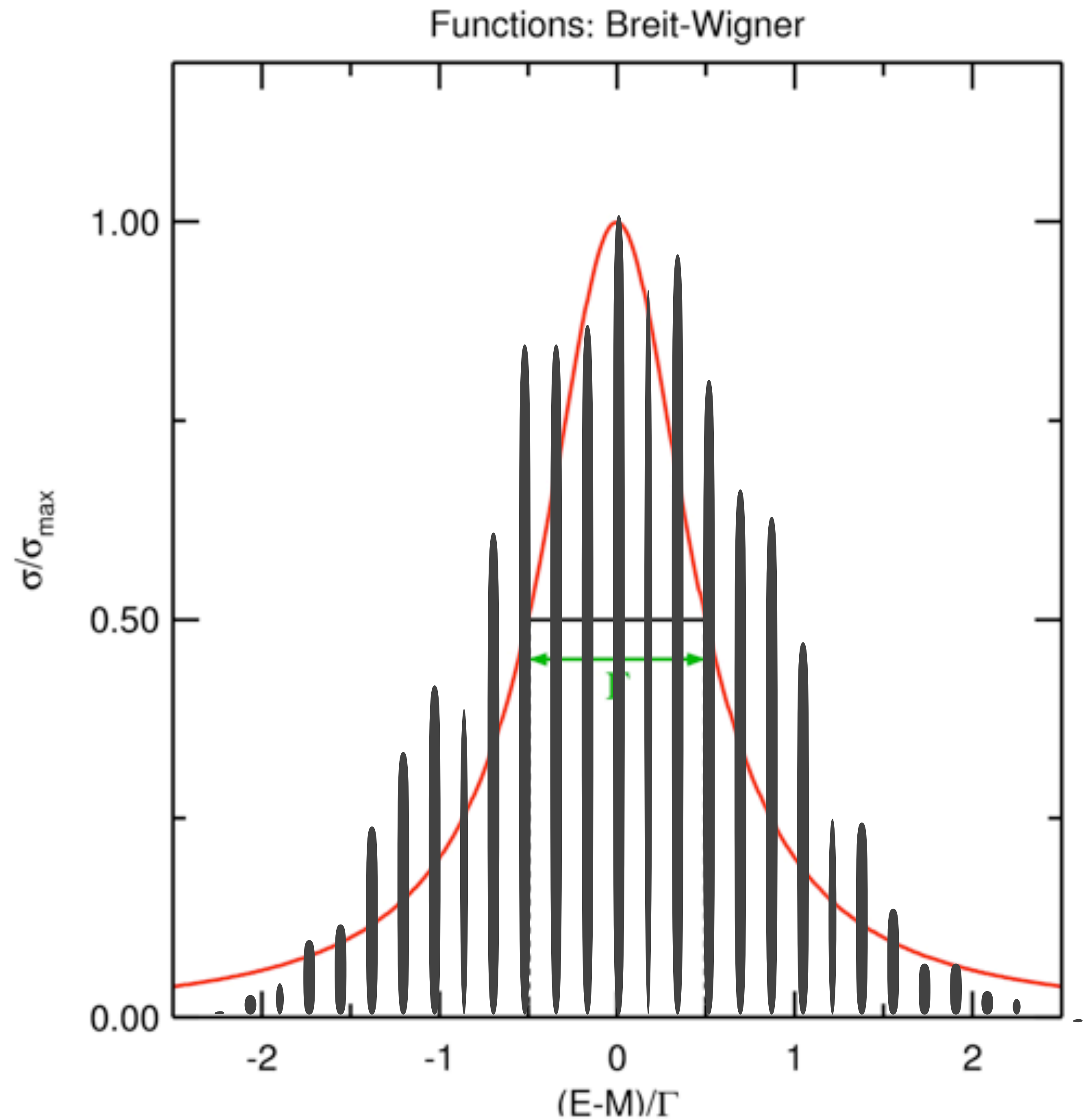
→ faster convergence for same number of function evaluations




→ Can even design algorithms that do this automatically as they run (not covered here)

→ Adaptive sampling

# IMPORTANCE SAMPLING




→ or throw points according to some smooth peaked function for which you have, or can construct, a random number generator (here: Gauss)

Any MC generator contains LOTS of examples of this.

(+ some generic algorithms though generally never as good as dedicated ones: e.g., VEGAS algorithm)

Note: if several peaks: do **multi-channel importance sampling** (~ competing random processes)

1) You are inputting knowledge: obviously need to know where the peaks are to begin with … (say you know, e.g., the location and width of a resonance or singularity)

2) Stratified sampling increases efficiency by combining n-point quadrature with the MC method, with further gains from adaptation

3) Importance sampling:

$$\int_a^b f(x)\mathrm{d}x = \int_a^b \frac{f(x)}{g(x)}\mathrm{d}G(x)$$

Flat sampling in x → Flat sampling in G(x)

Effectively does flat MC with changed integration variables

Fast convergence if $f(x)/g(x) \approx 1$

# SIMPLE MC EXAMPLE

## NUMBER OF PEDESTRIANS (IN LUND) WHO WILL GET HIT BY A CAR THIS WEEK

**Complicated Function:**

**Time-dependent**

Traffic density during day, week-days vs week-ends

(I.E., NON-TRIVIAL TIME EVOLUTION OF SYSTEM)

**No two pedestrians are the same**

Need to compute probability for each and sum

(SIMULATES HAVING SEVERAL DISTINCT TYPES OF “EVOLVERS”)

**(Multiple outcomes** (ignored for today)**:)**

Hit → keep walking, or go to hospital?

Multiple hits = Product of single hits, or more complicated?

# MONTE CARLO APPROACH

## Approximate Traffic

Simple overestimate:

highest recorded density

of most careless drivers,

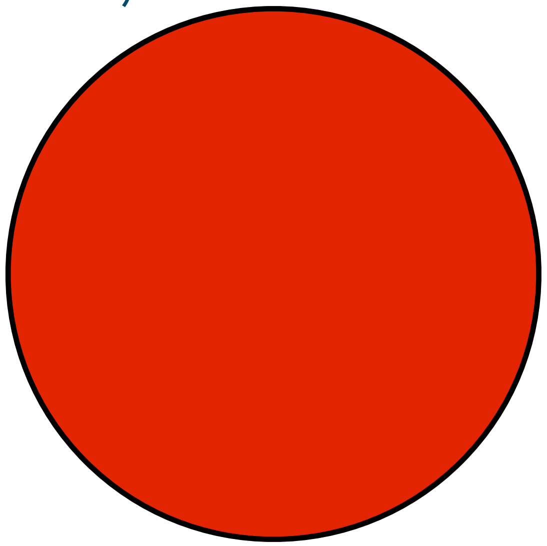
driving at highest recorded speed

…

## Approximate Pedestrian

by most completely reckless and accident-prone person (e.g., MCnet student wandering the streets lost in thought after these lectures …)

This extreme guess will be the equivalent of a simple area (~integral) we can calculate:



## Off we go…

basically a special application of importance sampling; transforming a uniform distribution to a non-uniform one

Throw random accidents according to:

Solve for *t(R)*

Uniformly distributed random number ∈ [0,1] →

$$R = \int_{t_0}^{t} \mathrm{d}t' \int_{\mathrm{Area}} \mathrm{d}^2x \sum_{i=1}^{n_{\mathrm{ped}}} \alpha_i(x,t')\, \rho_i(x,t')\, \rho_c(x,t')$$

Sum over Pedestrians

Pedestrian-Car interaction

Density of Pedestrian i

Density of Cars

*$t_0$ : starting time*
*$t$ : time of accident*

Too Difficult

↓

Simple Overestimate

Larger trial area with simple boundary (in this case, circle)

$$R_{\mathrm{trial}} = (t_{\mathrm{trial}} - t_0)\,(\pi r_{\mathrm{max}}^2)\;\alpha_{\mathrm{max}}\, n_{\mathrm{ped}}\, \rho_{c\mathrm{max}}$$

Solve for $t_{trial}(R_{trial})$

Hit rate for most accident-prone pedestrian with worst driver

Rush-hour density of cars

(Also generate trial *x*, e.g., uniformly in circular area around Lund)

(Also generate trial *i*; a random pedestrian gets hit)

(note: this generator is *unordered*; not asking whether that pedestrian was already hit earlier…)

# ACCEPT OR REJECT TRIAL

Now you have a trial. Veto the trial if generated *x* is outside desired physical boundary. If inside, accept trial hit *(i,x,t)* with probability

(exactly equivalent to when we coloured points **blue** [accept] or **red** [reject] )

$$\text{Prob(accept)} = \frac{\alpha_i(x,t)\ \rho_i(x,t)\ \rho_c(x,t)}{\alpha_{\max}\ \rho_{\mathrm{cmax}}}$$

*Using the following:*

$\rho_c$ : *actual density of cars at location x at time t*

$\rho_i$ : *actual density of student i at location x at time t*

$\alpha_i$ : *The actual "hit rate" (OK, not really known, but could fit to past data: "tuning")*

→ True number = number of accepted hits

(caveat: we didn't really treat multiple hits …

→ Sudakovs & Markov Chains; tomorrow)

# SUMMARY: HOW WE DO MONTE CARLO

Take your system

Generate a "trial" (event/decay/interaction/… )

Not easy to generate random numbers distributed according to exactly the right distribution?

May have complicated dynamics, interactions …

→ use a simpler "trial" overestimating distribution

Flat with some stratification

Or importance sample with simple overestimating function (for which you can ~ easily generate random numbers)

# SUMMARY: HOW WE DO MONTE CARLO

Take your system

Generate a "trial" (event/decay/interaction/… )

Accept trial with probability f(x)/g(x)

- f(x) contains all the complicated dynamics
- g(x) is the simple trial function

If accept: replace with new system state

If reject: keep previous system state

no dependence on g(x) in final result - only affects convergence rate

**And keep going:** generate next trial …

# SUMMARY: HOW WE DO MONTE CARLO

Take your system

Generate a "trial" (event/decay/intera

- Accept trial with probability f(x)/g(x)
  - f(x) contains all the complicated dynamics
  - g(x) is the simple trial function
- If accept: replace with new system stat
- If reject: keep previous system state

no dependence on g
final result - only af
convergence rate

Sounds deceptively simple, but …

**with it, you can integrate**
arbitrarily complicated functions *(and chains of nested functions)*,
over arbitrarily complicated regions, in arbitrarily many dimensions …

**And keep going:** generate next trial …

## A Psychological Tip

*Whenever you're called on to make up your mind,*
*and you're hampered by not having any,*
*the best way to solve the dilemma, you'll find,*
*is simply by spinning a penny.*

*No -- not so that chance shall decide the affair*
*while you're passively standing there moping;*
*but the moment the penny is up in the air,*
*you suddenly know what you're hoping.*

PSYKOLOGISK
HUSRÅD

Stejler man foran et vanskeligt valg
og vil ha det afgjort prompte,
er det et såre fornuftigt princip
at platte og krone om det.

Ikke at valget skal ske pr. hazard,
imens man selv sidder og måber,
men: lige når mønten er kastet til vejrs,
så véd man præcis, hvad man håber.

Piet Hein
GRUK

[Piet Hein, Danish scientist, poet & friend of Niels Bohr]

# Extra Slides

# IF YOU WANT TO PLAY WITH RANDOM NUMBERS

I will not tell you how to *write* a Random-number generator. (For that, see the references in the writeup.)

Instead, I *assume* that you can write a computer code and link to a random-number generator, from a library

E.g., ROOT includes one that you can use if you like.

PYTHIA also includes one

From the PYTHIA 8 HTML documentation, under "Random Numbers":

Random numbers `R` uniformly distributed in `0 < R < 1` are obtained with

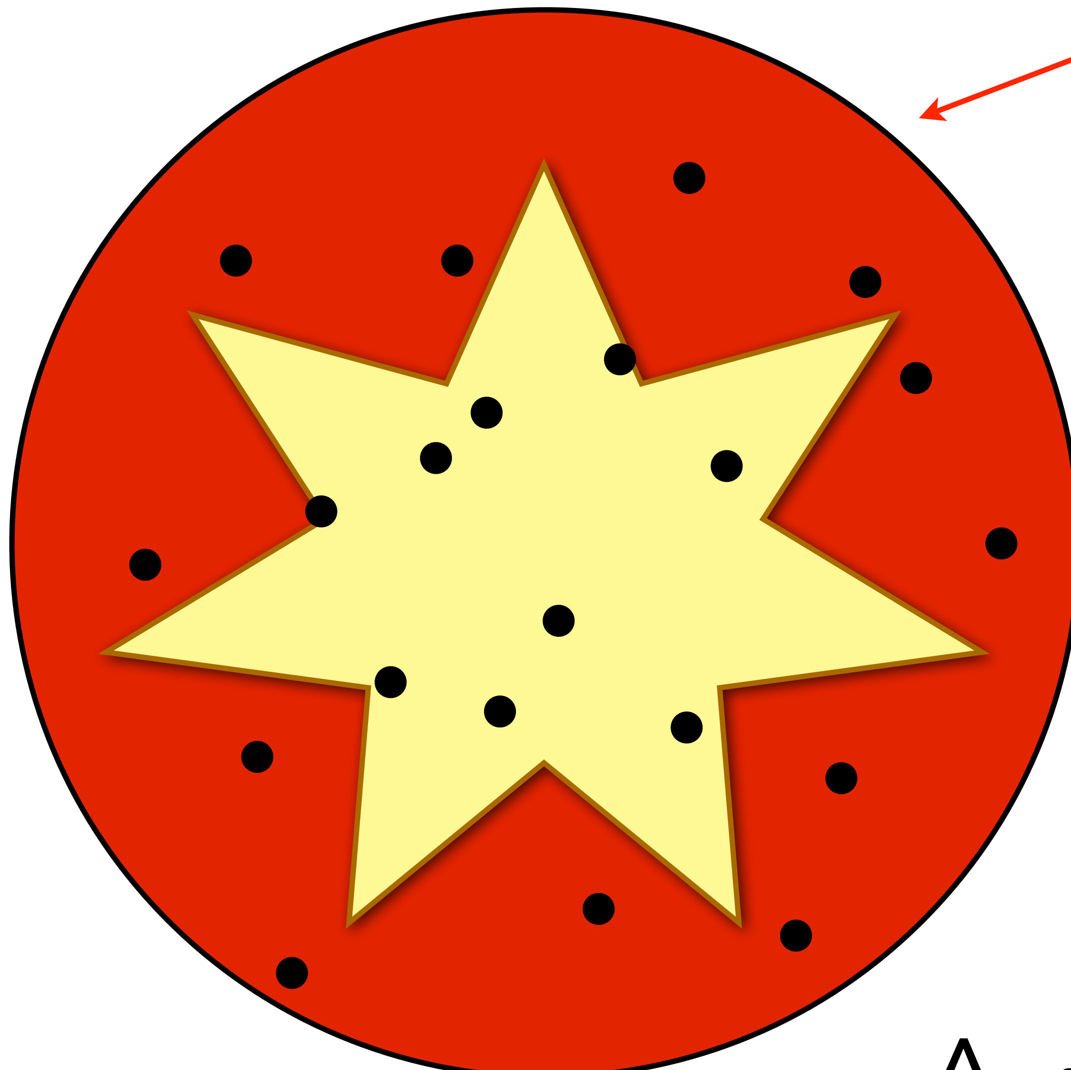
```
Pythia8::Rndm::flat();
```

+ Other methods for exp, x*exp, 1D Gauss, 2D Gauss.

Example 1: simple function (=constant); complicated boundary

Now get a few friends, some balls, and throw random shots inside the circle
(PS: be careful to make your shots truly random)

Count how many shots hit the shape inside and how many miss



Assume you know the area of <u>this</u> shape:
$\pi R^2$
(an overestimate)

Earliest Example of MC calculation: Buffon's Needle (1777) to calculate $\pi$

G. Leclerc, Comte de Buffon (1707-1788)

$$A_{\star} \approx N_{hit}/N_{miss} \times \pi R^2$$

## Colour Factors

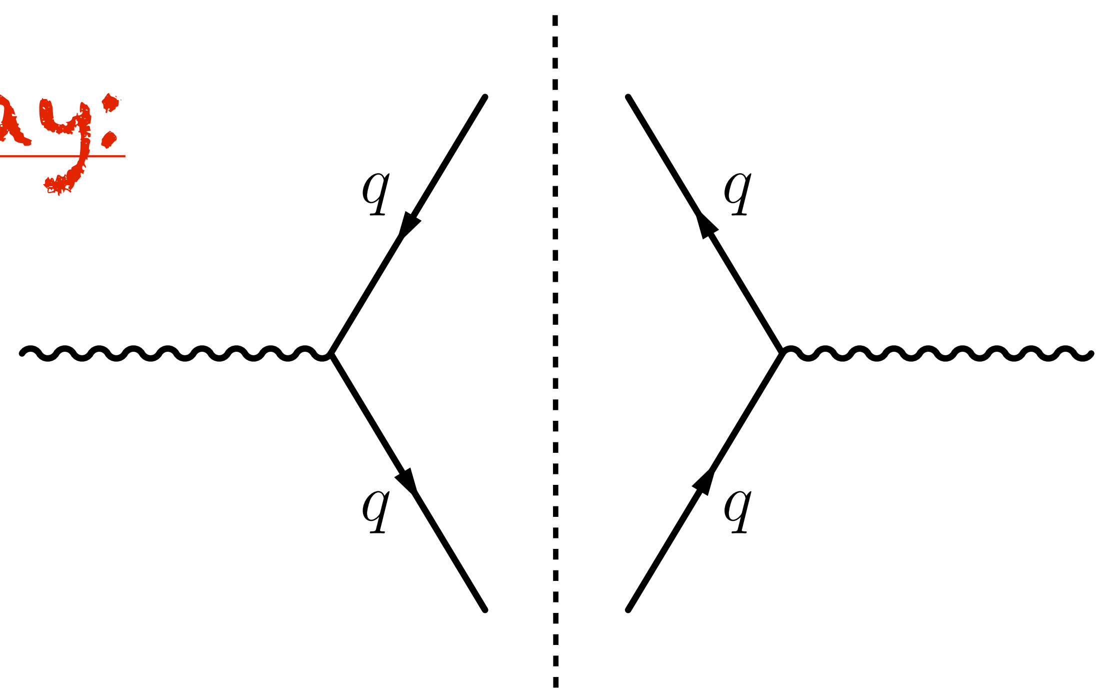
Processes involving coloured particles have a "colour factor".

It counts the enhancement from the sum over colours.

(*average* over incoming colours → can also give suppression)

Z Decay:

$$\sum_{\text{colours}} |M|^2 =$$



## Colour Factors

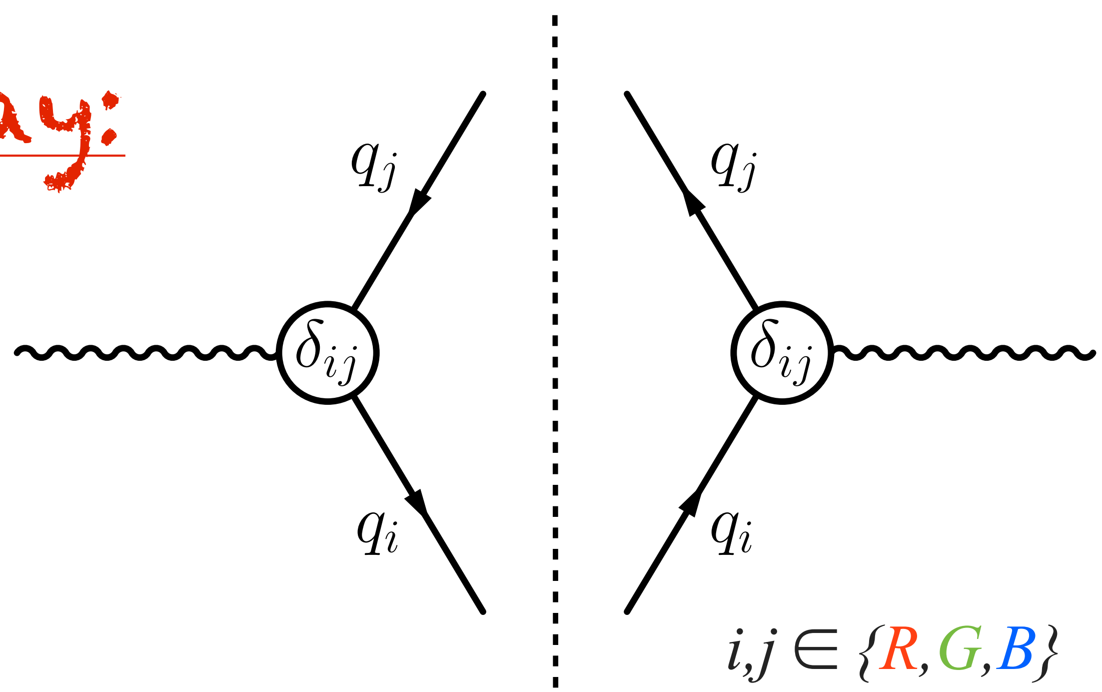
Processes involving coloured particles have a "colour factor".

It counts the enhancement from the sum over colours.

(*average* over incoming colours → can also give suppression)

**Z Decay:**

$$\sum_{\text{colours}} |M|^2 =$$




$$\propto \delta_{ij}\delta^*_{ji} = \mathrm{Tr}[\delta_{ij}] = N_C$$

# INTERACTIONS IN COLOUR SPACE

## Colour Factors

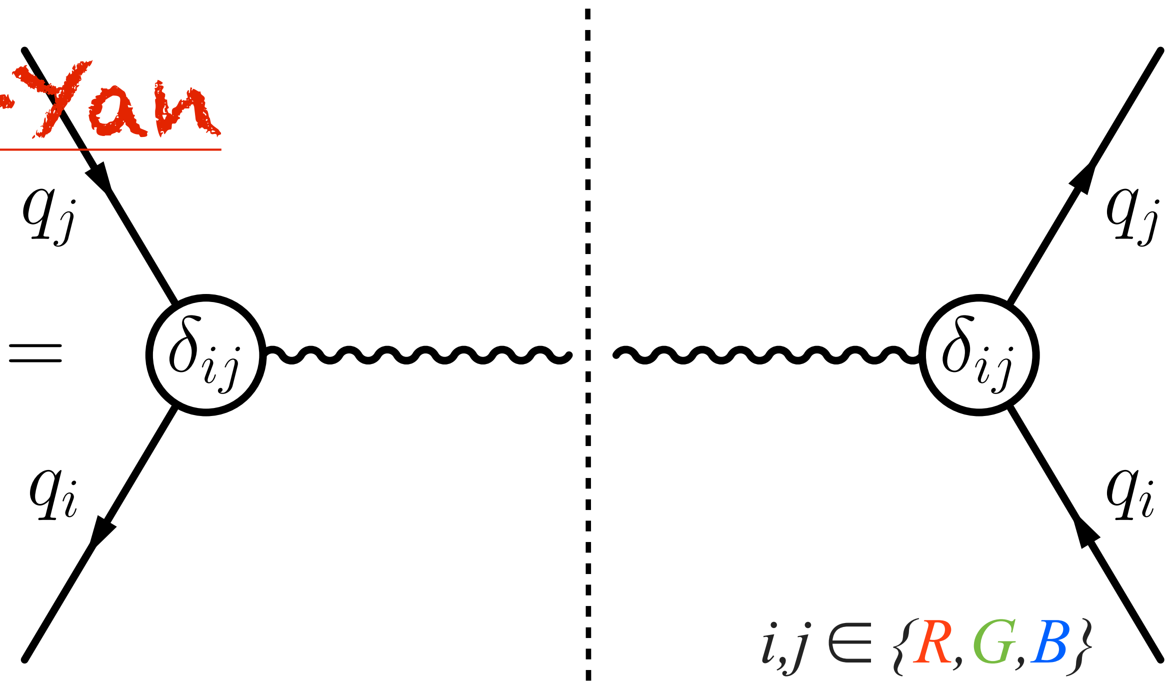
Processes involving coloured particles have a "colour factor".

It counts the enhancement from the sum over colours.

(*average* over incoming colours → can also give suppression)

Drell-Yan

$$\frac{1}{9}\sum_{\text{colours}} |M|^2 =$$

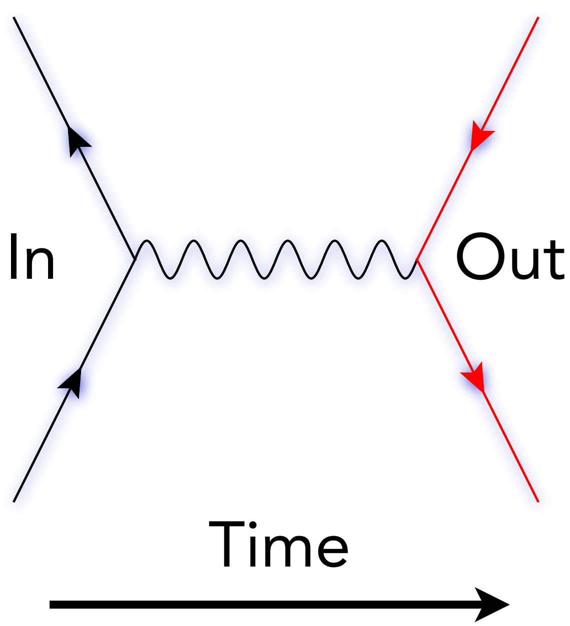



$$\propto \delta_{ij}\delta^*_{ji}\,\frac{1}{N_C^2}$$
$$= \mathrm{Tr}[\delta_{ij}]\,\frac{1}{N_C^2}$$
$$= 1/N_C$$

# CROSSINGS

$e^+e^- \to \gamma^*/Z \to q\bar{q}$

(Hadronic Z Decay)

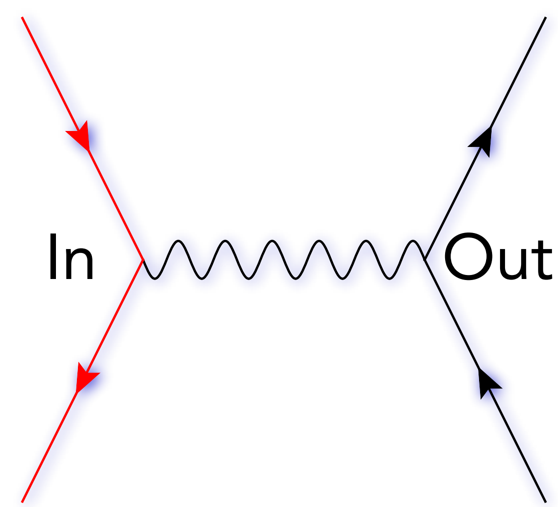



Color Factor:

$$\mathrm{Tr}[\delta_{ij}] = N_C$$

$q\bar{q} \to \gamma^*/Z \to \ell^+\ell^-$

(Drell & Yan, 1970)

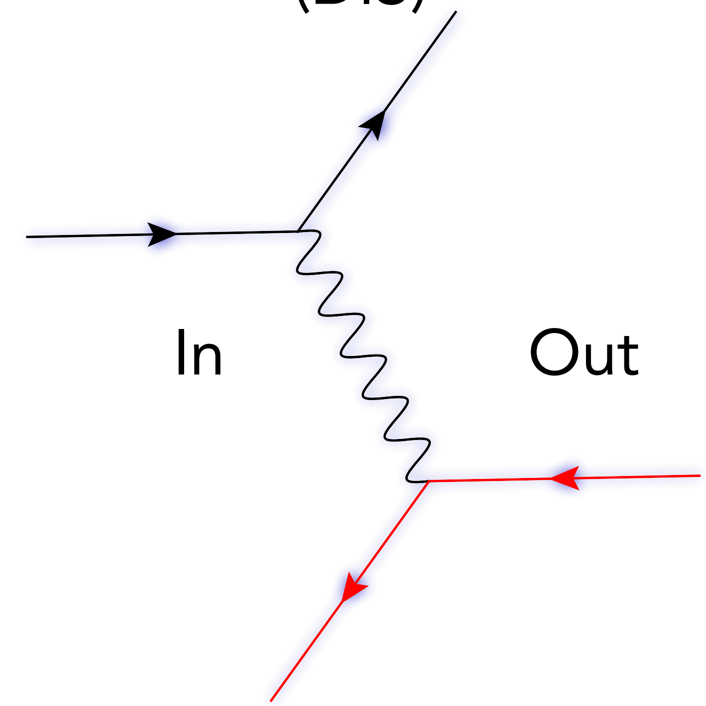



Color Factor:

$$\frac{1}{N_C^2}\mathrm{Tr}[\delta_{ij}] = \frac{1}{N_C}$$

$\ell q \stackrel{\gamma^*/Z}{\to} \ell q$

(DIS)




Color Factor:

$$\frac{1}{N_C}\mathrm{Tr}[\delta_{ij}] = 1$$

## Colour Factors

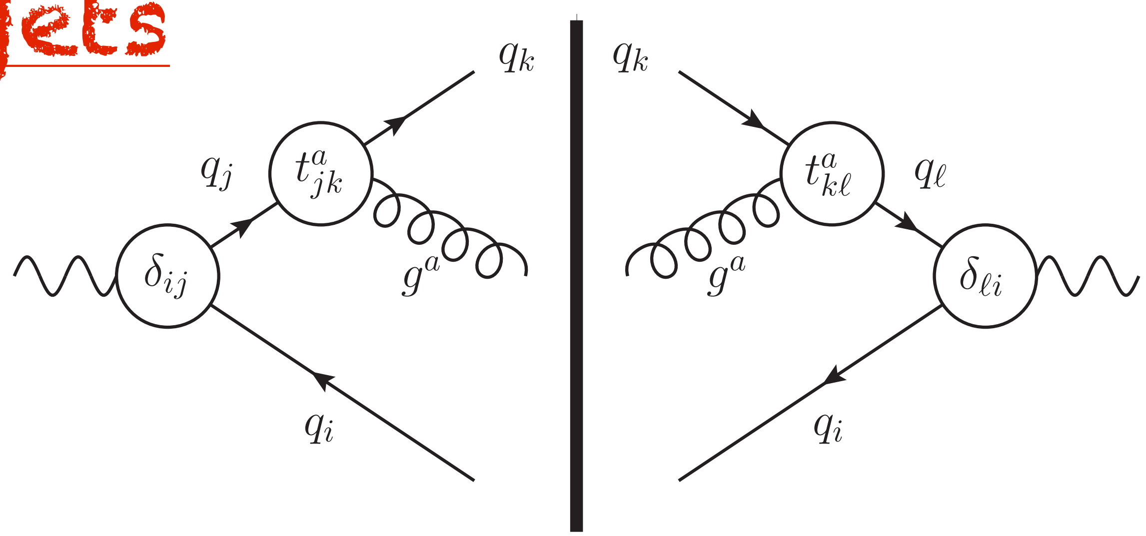
Processes involving coloured particles have a "colour factor".

It counts the enhancement from the sum over colours.

(*average* over incoming colours → can also give suppression)

Z→3 jets

$$\sum_{\text{colours}} |M|^2 =$$



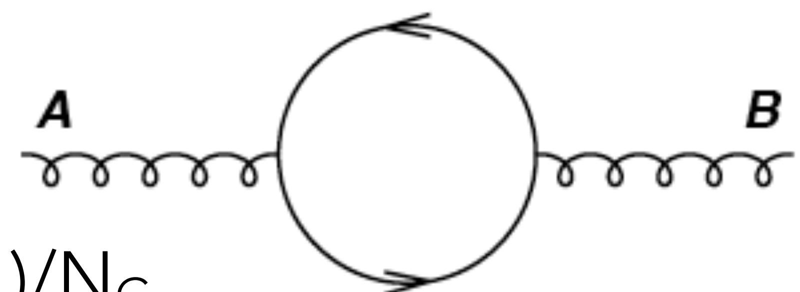

$$\begin{aligned} &\propto \delta_{ij} t^a_{jk} t^a_{k\ell} \delta_{\ell i} \\ &= \mathrm{Tr}\{t^a t^a\} \\ &= \frac{1}{2}\mathrm{Tr}\{\delta\} = 4 \end{aligned}$$

$i,j \in \{R,G,B\}$
$a \in \{1,\ldots,8\}$

## Colour factors squared produce traces

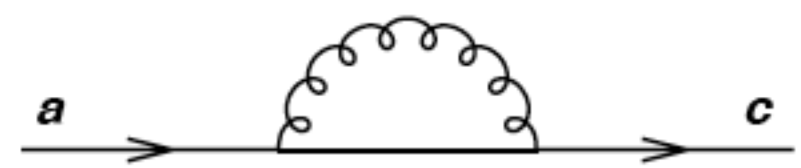
| Trace Relation | Example Diagram |
| --- | --- |

$$\mathrm{Tr}(t^A t^B) = T_R \delta^{AB}\,, \quad T_R = \tfrac{1}{2}$$

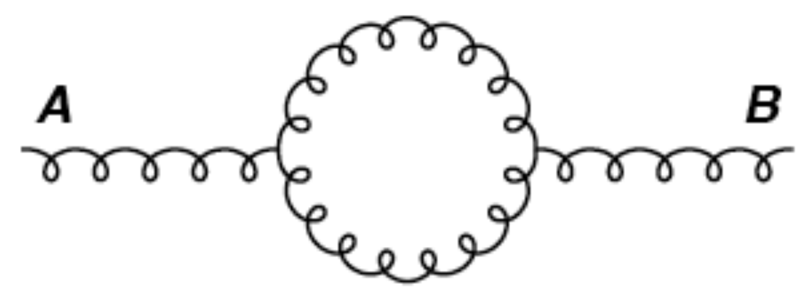



$T_R(N_c^2\text{-}1)/N_C$

$$\textstyle\sum_A t^A_{ab} t^A_{bc} = C_F \delta_{ac}\,, \quad C_F = \frac{N_c^2 - 1}{2N_c} = \frac{4}{3}$$

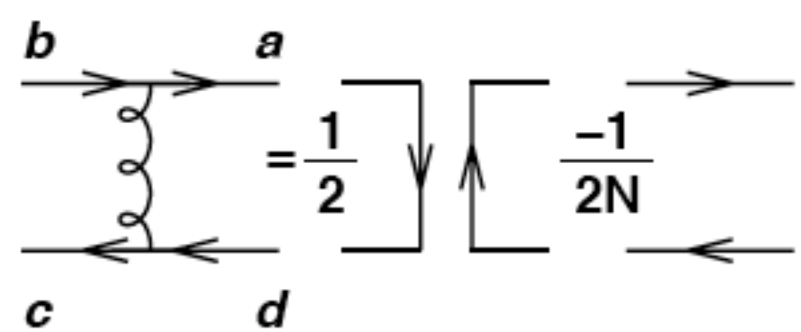



$$\textstyle\sum_{C,D} f^{ACD} f^{BCD} = C_A \delta^{AB}\,, \quad C_A = N_c = 3$$




$$t^A_{ab} t^A_{cd} = \frac{1}{2}\delta_{bc}\delta_{ad} - \frac{1}{2N_c}\delta_{ab}\delta_{cd} \quad \text{(Fierz)}$$

$T_R$ $\quad$ $T_R/N_C$




(from ESHEP lectures by G. Salam)

## Real QCD isn't conformal

The coupling runs logarithmically with the energy scale

$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \beta(\alpha_s) \qquad \beta(\alpha_s) = -\alpha_s^2(b_0 + b_1\alpha_s + b_2\alpha_s^2 + \ldots) \ ,$$

$$b_0 = \frac{11C_A - 2n_f}{12\pi} \qquad b_1 = \frac{17C_A^2 - 5C_A n_f - 3C_F n_f}{24\pi^2} = \frac{153 - 19n_f}{24\pi^2}$$

1-Loop β function coefficient

2-Loop β function coefficient

$$b_2 = \frac{2857 - 5033n_f + 325n_f^2}{128\pi^3}$$

$$b_3 = \text{known}$$

**Asymptotic freedom** in the ultraviolet

**Confinement (IR slavery?)** in the infrared

# Multi-Scale Exercise

Skands, TASI Lectures, arXiv:1207.2389

If needed, can convert from multi-scale to single-scale

$$
\begin{aligned}
\alpha_s(\mu_1)\alpha_s(\mu_2)\cdots\alpha_s(\mu_n) &= \prod_{i=1}^{n}\alpha_s(\mu)\left(1+b_0\,\alpha_s\ln\left(\frac{\mu^2}{\mu_i^2}\right)+\mathcal{O}(\alpha_s^2)\right) \\
&= \alpha_s^n(\mu)\left(1+b_0\,\alpha_s\ln\left(\frac{\mu^{2n}}{\mu_1^2\mu_2^2\cdots\mu_n^2}\right)+\mathcal{O}(\alpha_s^2)\right)
\end{aligned}
$$

by taking geometric mean of scales

# Phase Space Generation

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

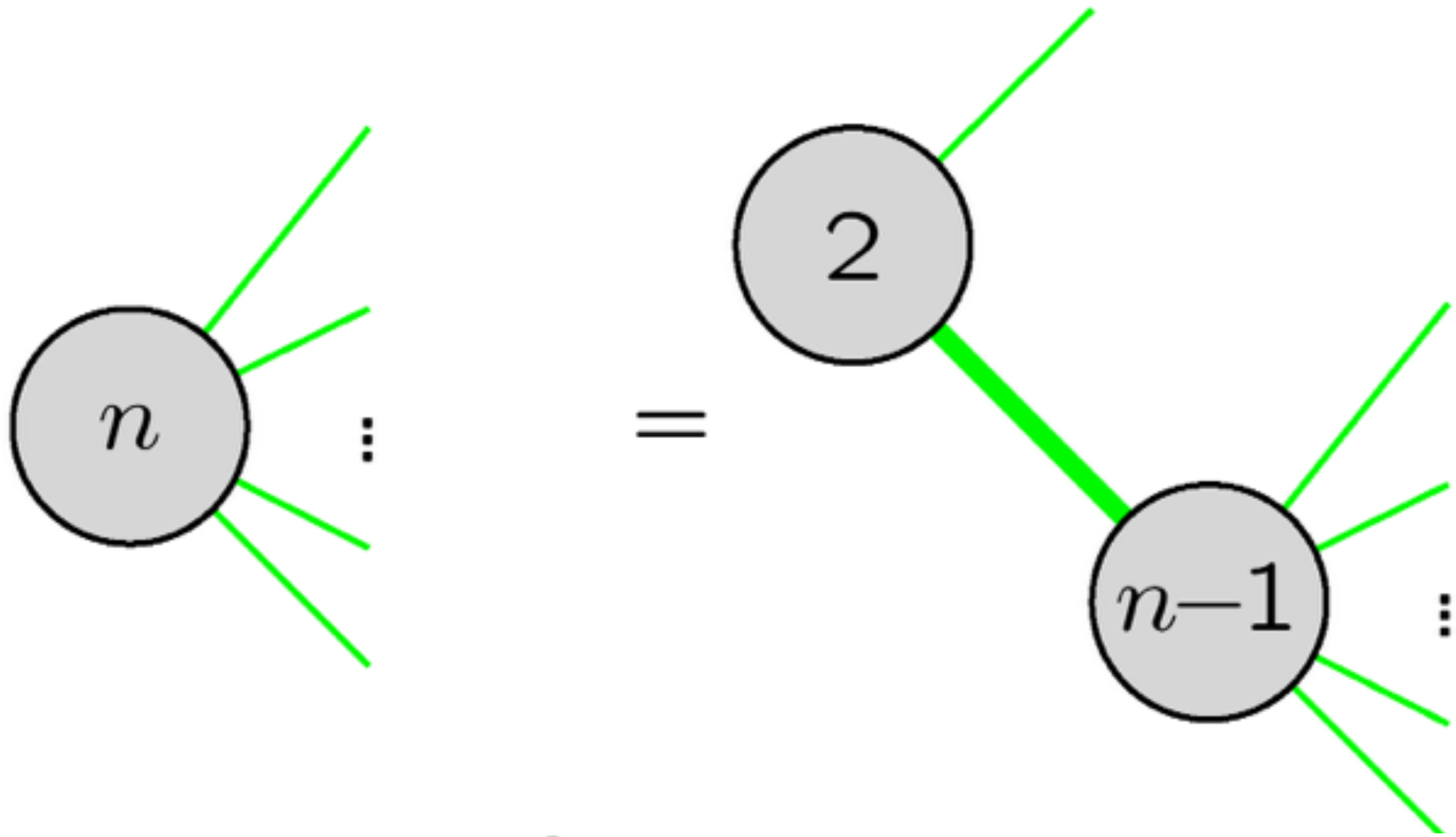
- Phase space:

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

- Two-body easy:

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

- Other cases by recursive subdivision:



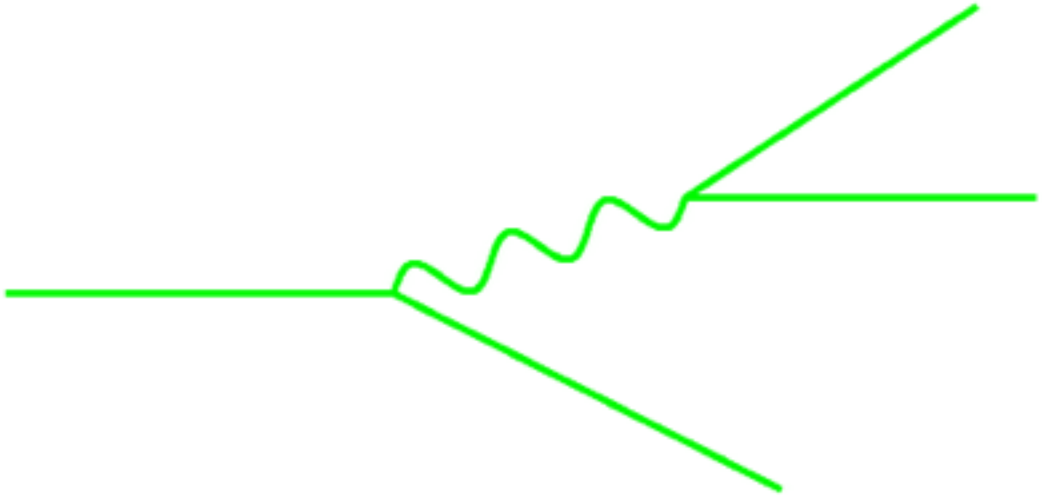

$$d\Pi_n(M) = \frac{1}{2\pi}\int_0^{(M-m)^2} dm_x^2 \, d\Pi_2(M) \, d\Pi_{n-1}(m_x)$$

- 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 - must be removed by importance sampling:

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