



LUND UNIVERSITY

CTEQ-MCnet School 2010
Lauterbad, Germany
26 July - 4 August 2010

Introduction to Monte Carlo Event Generators

Torbjörn Sjöstrand
Lund University

1. (today) Introduction and Overview; Monte Carlo Techniques
 2. (today) Matrix Elements; Parton Showers I
 3. (tomorrow) Parton Showers II; Matching Issues
 4. (tomorrow) Multiple Parton–Parton Interactions
 5. (Wednesday) Hadronization and Decays; Generator Status

Disclaimer 1

These lectures will *not* cover:

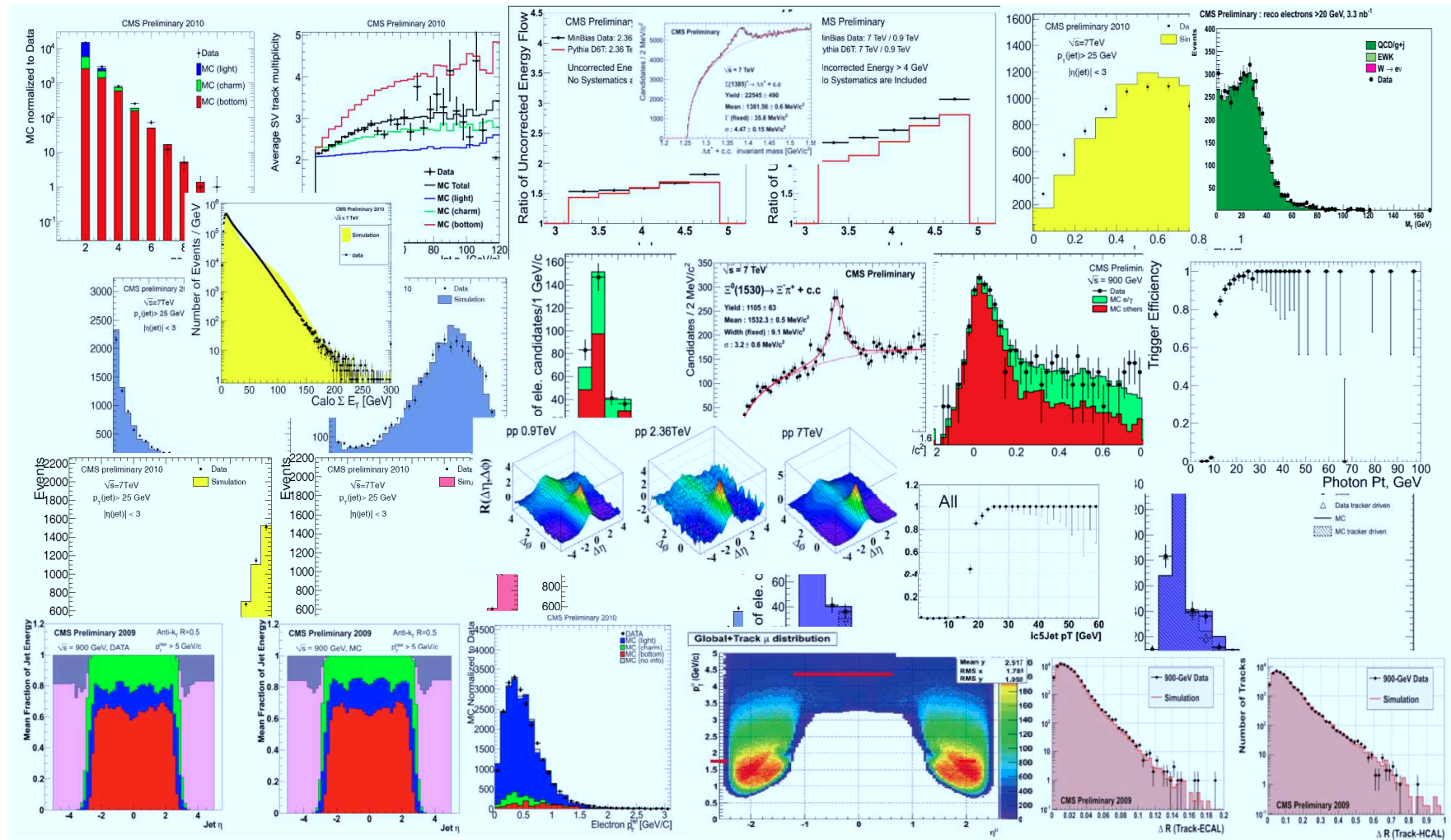
- ★ Heavy-ion physics:
 - without quark-gluon plasma formation, or
 - with quark-gluon plasma formation.
- ★ Specific physics studies for topics such as
 - B production,
 - Higgs discovery,
 - SUSY phenomenology,
 - other new physics discovery potential.

They *will* cover the “normal” physics that will be there in (essentially) all LHC pp events, from QCD to exotics:

- ★ the generation and availability of different processes,
- ★ the addition of parton showers,
- ★ the addition of an underlying event,
- ★ the transition from partons to observable hadrons, plus
- ★ the status and evolution of general-purpose generators.

Disclaimer 2

ICHEP is on in Paris, with many new LHC results announced.



At this school there will be four experimental talks on first LHC results. My lectures will help to give background, but show very few LHC plots.

Read More

These lectures (and more):

<http://home.thep.lu.se/~torbjorn/> and click on “Talks”

Peter Skands, European School of High Energy Physics, June 2010:

<http://home.fnal.gov/~skands/slides/>

Many presentations at the MCnet Summer School, Lund, July 2009:

<http://conference.ippp.dur.ac.uk/>

conferenceOtherViews.py?view=ipp&confId=264#2009-07-01

Many presentations at the CTEQ–MCnet Summer School, Aug 2008:

<http://conference.ippp.dur.ac.uk/>

conferenceOtherViews.py?view=ipp&confId=156

Bryan Webber, MCnet school, Durham, April 2007:

<http://www.hep.phy.cam.ac.uk/theory/webber/>

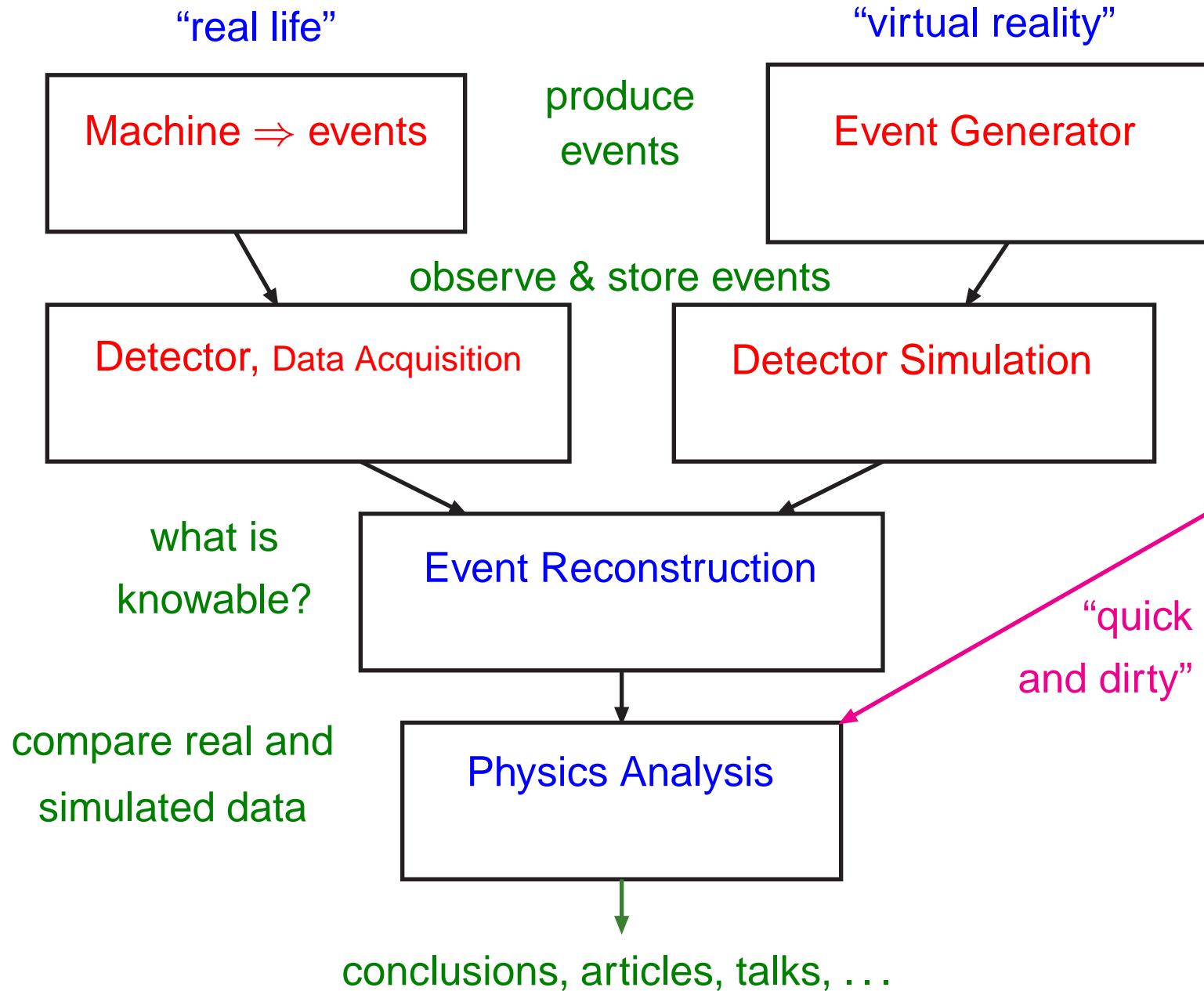
Peter Richardson, CTEQ Summer School lectures, July 2006:

<http://www.ippp.dur.ac.uk/~richardn/talks/>

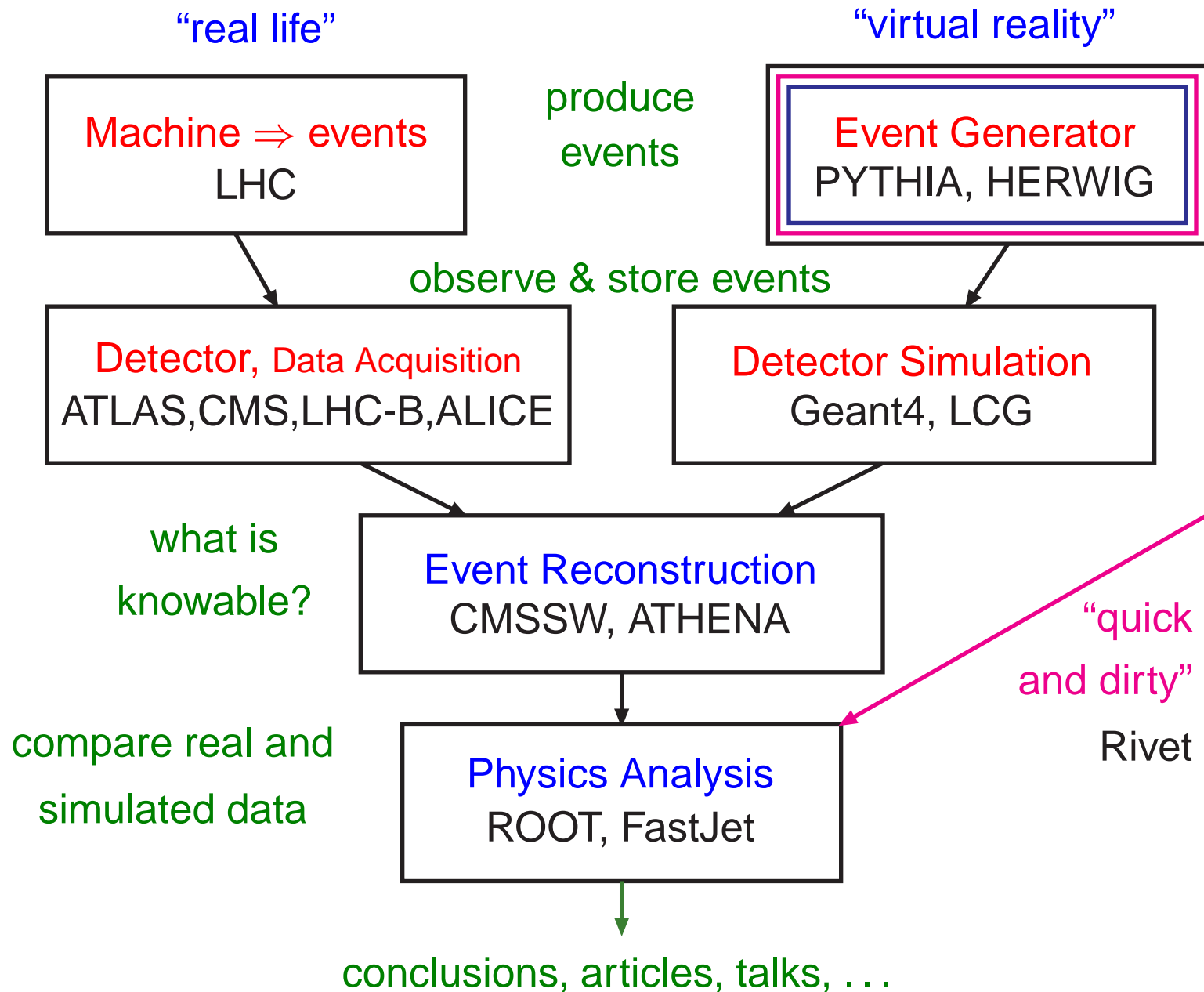
The “Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics”, hep-ph/0403045

<http://arxiv.org/pdf/hep-ph/0403045>

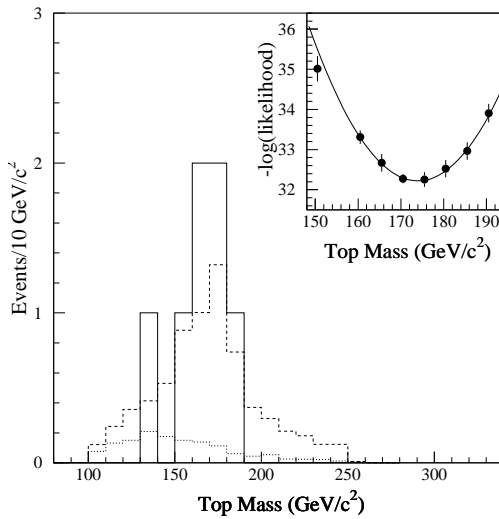
Event Generator Position



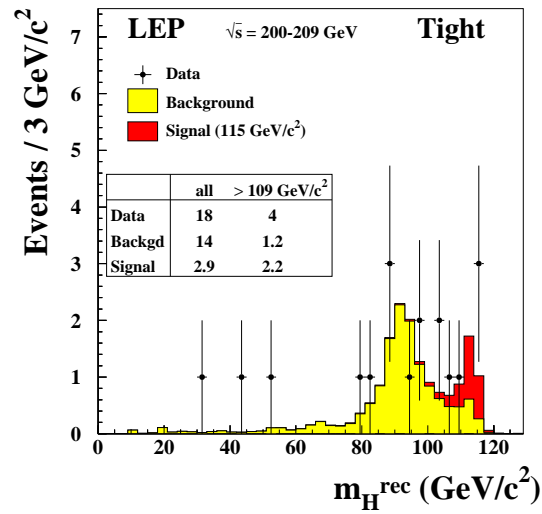
Event Generator Position



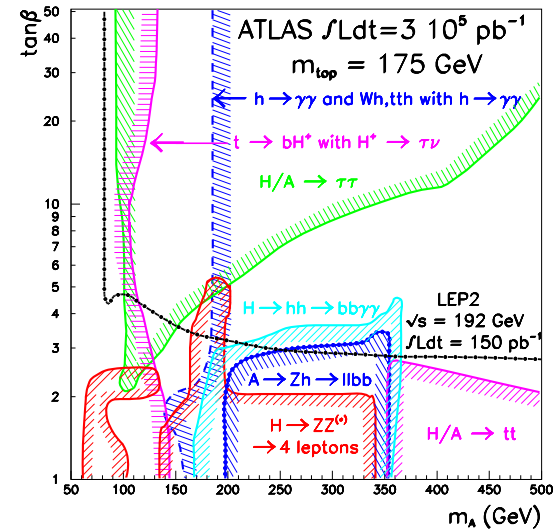
Why Generators? (I)



top discovery
and mass
determination



Higgs (non)
discovery



Higgs and
supersymmetry
exploration

not feasible without generators

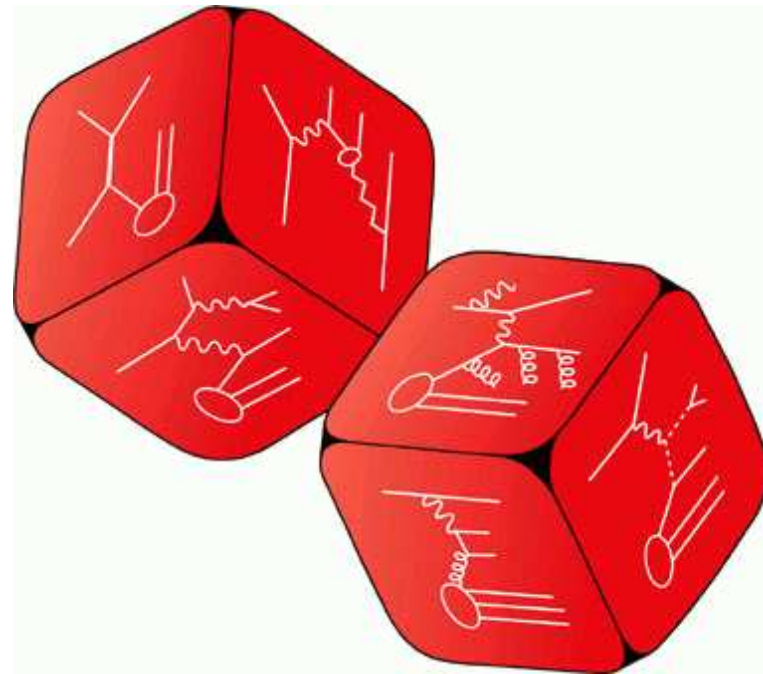
Why Generators? (II)

- Allow theoretical and experimental studies of *complex* multiparticle physics
- Large flexibility in physical quantities that can be addressed
 - Vehicle of ideology to disseminate ideas from theorists to experimentalists

Can be used to

- predict event rates and topologies
⇒ can estimate feasibility
- simulate possible backgrounds
⇒ can devise analysis strategies
 - study detector requirements
⇒ can optimize detector/trigger design
 - study detector imperfections
⇒ can evaluate acceptance corrections

A tour to Monte Carlo



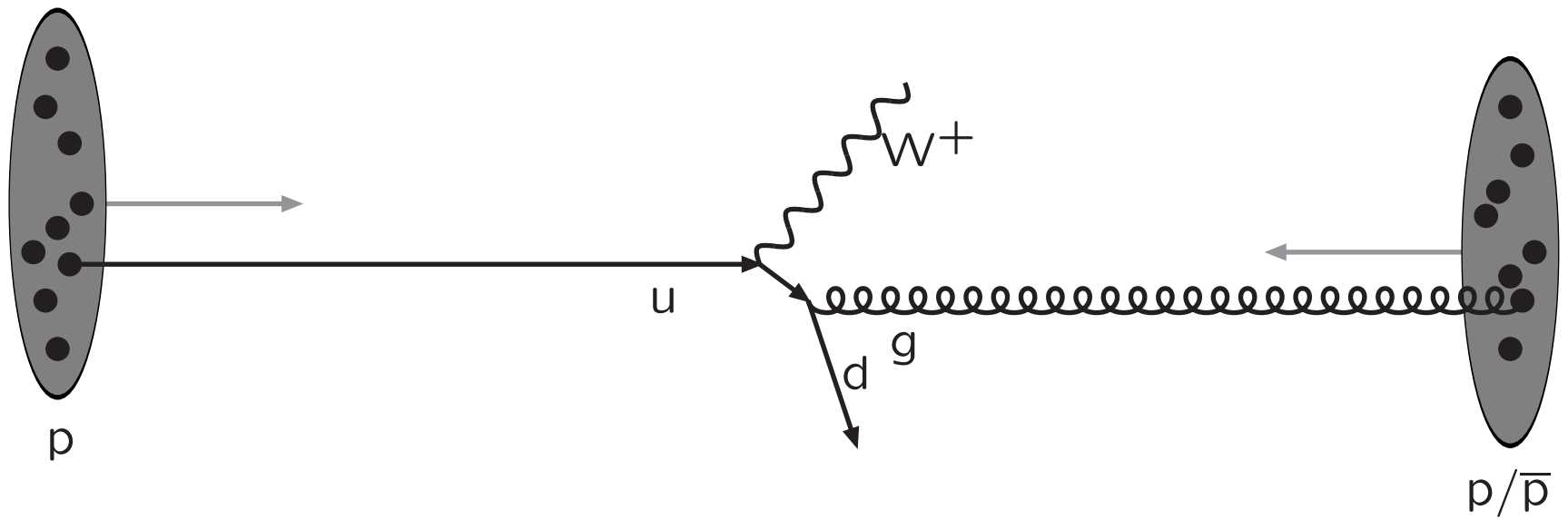
... because Einstein was wrong: God does throw dice!
Quantum mechanics: amplitudes \implies probabilities
Anything that possibly can happen, will! (but more or less often)

The structure of an event

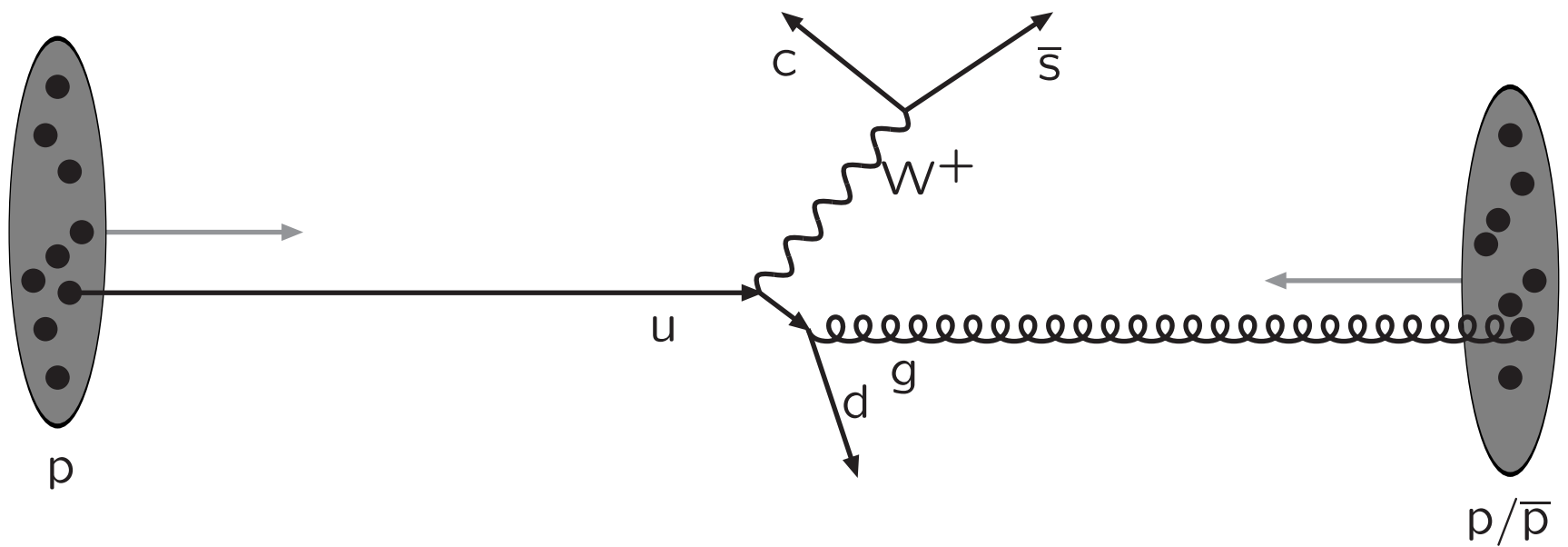
Warning: schematic only, everything simplified, nothing to scale, ...



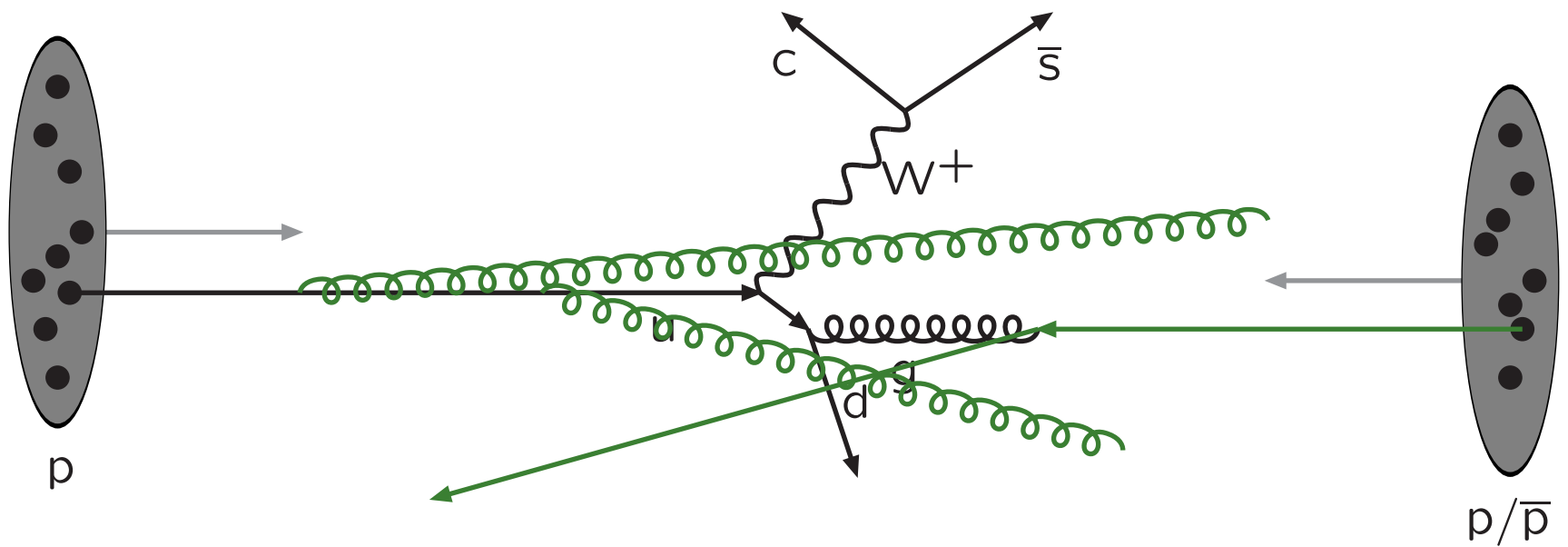
Incoming beams: parton densities



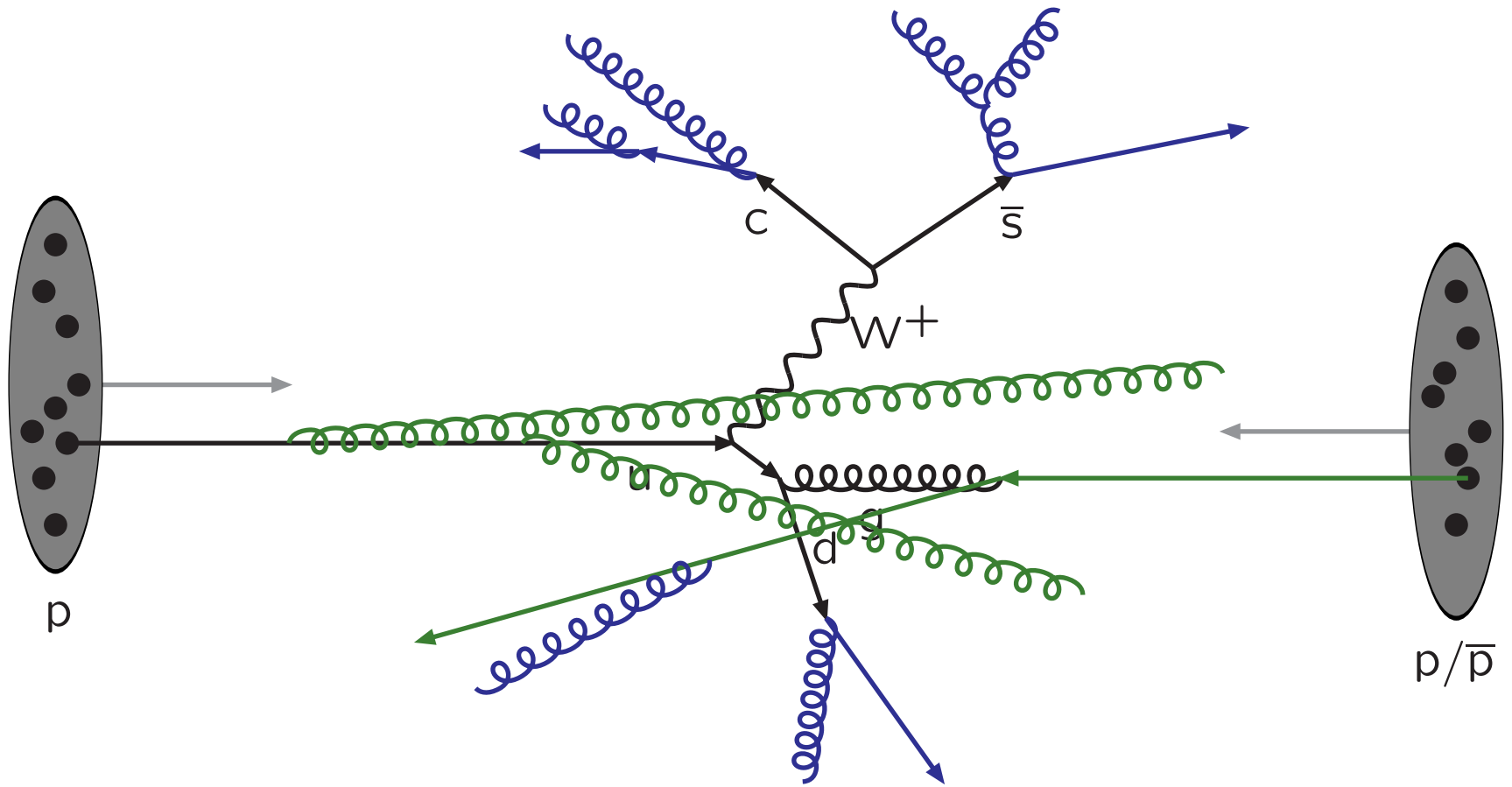
Hard subprocess: described by matrix elements



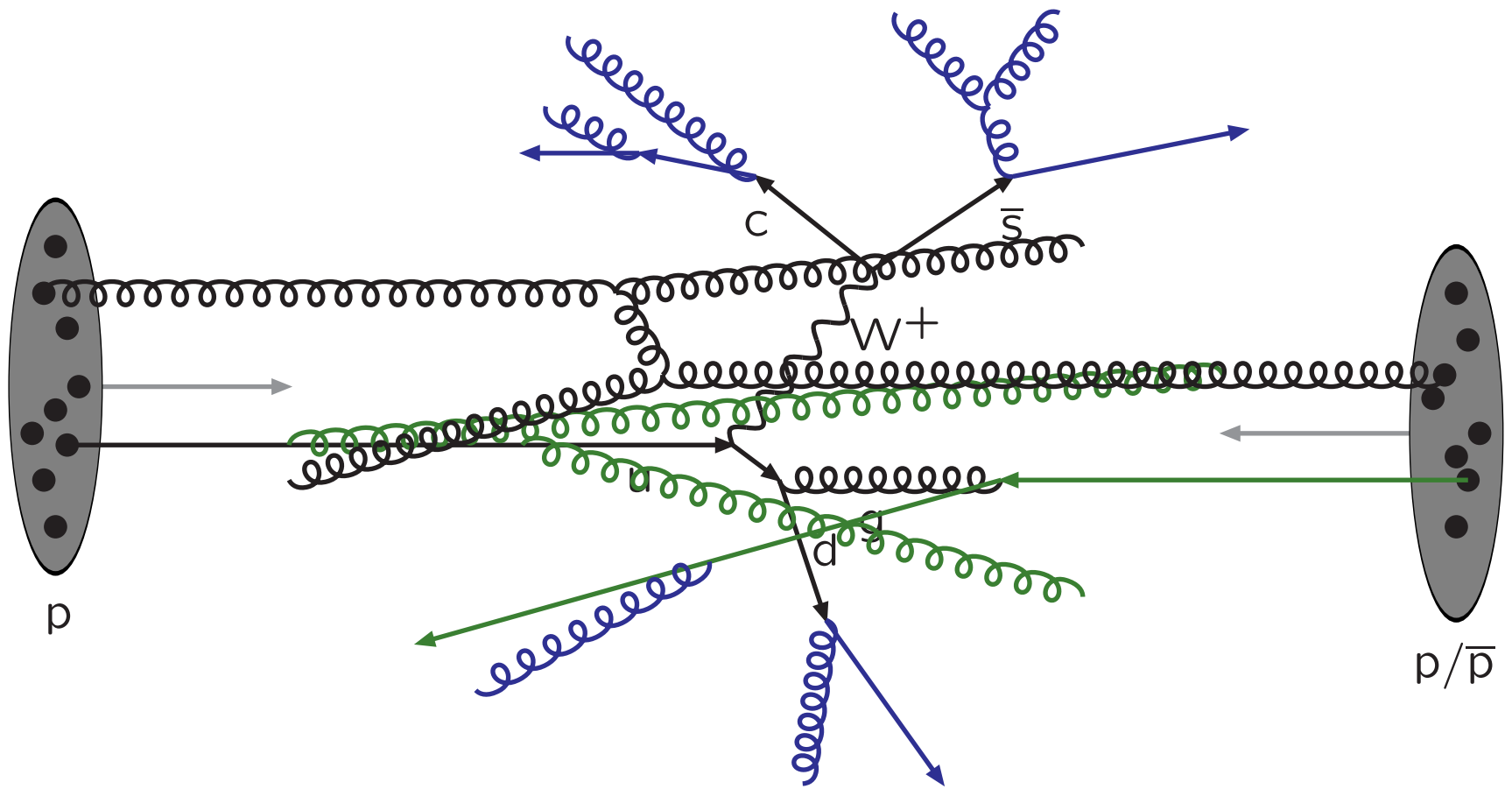
Resonance decays: correlated with hard subprocess



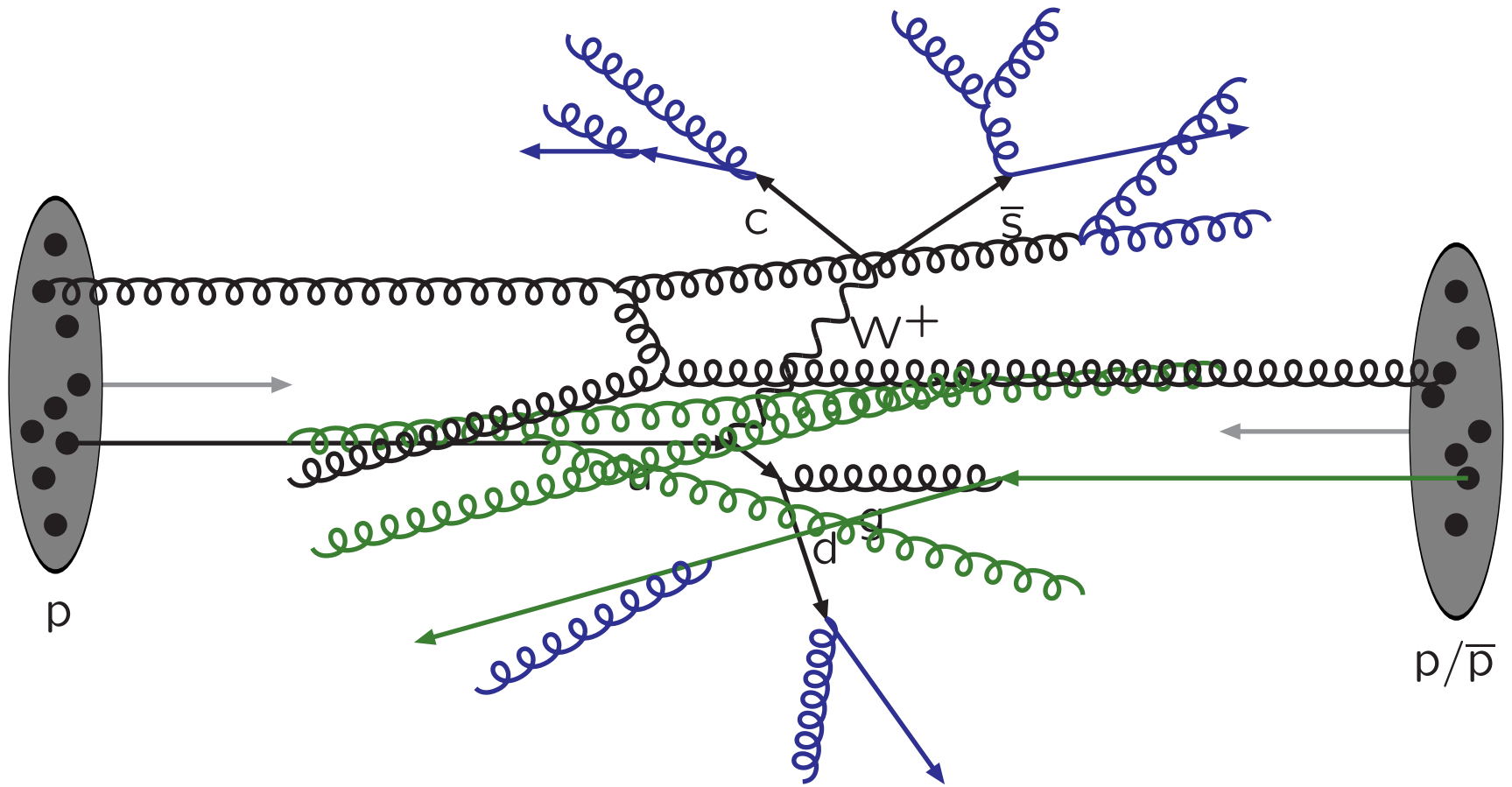
Initial-state radiation: spacelike parton showers



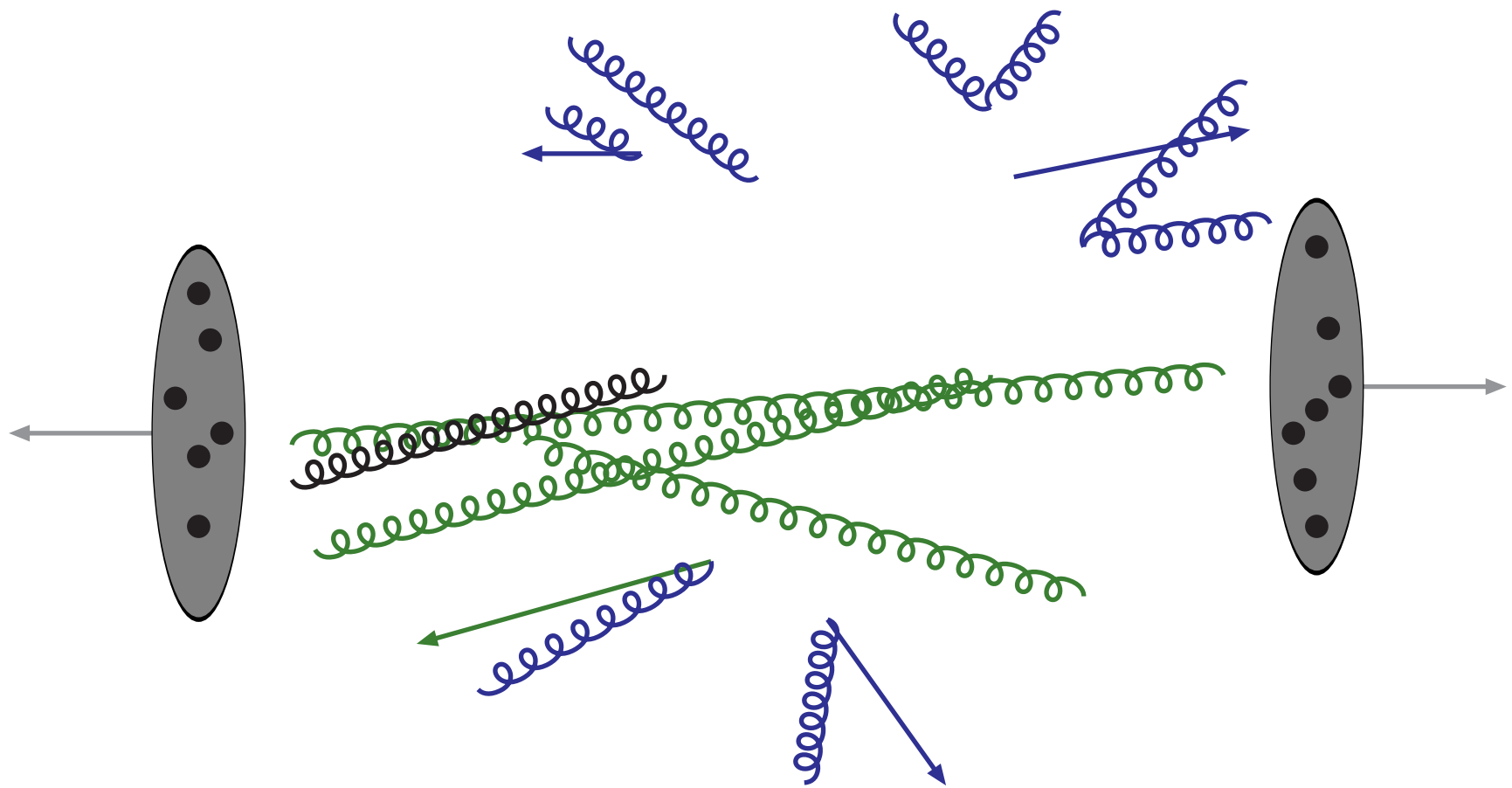
Final-state radiation: timelike parton showers



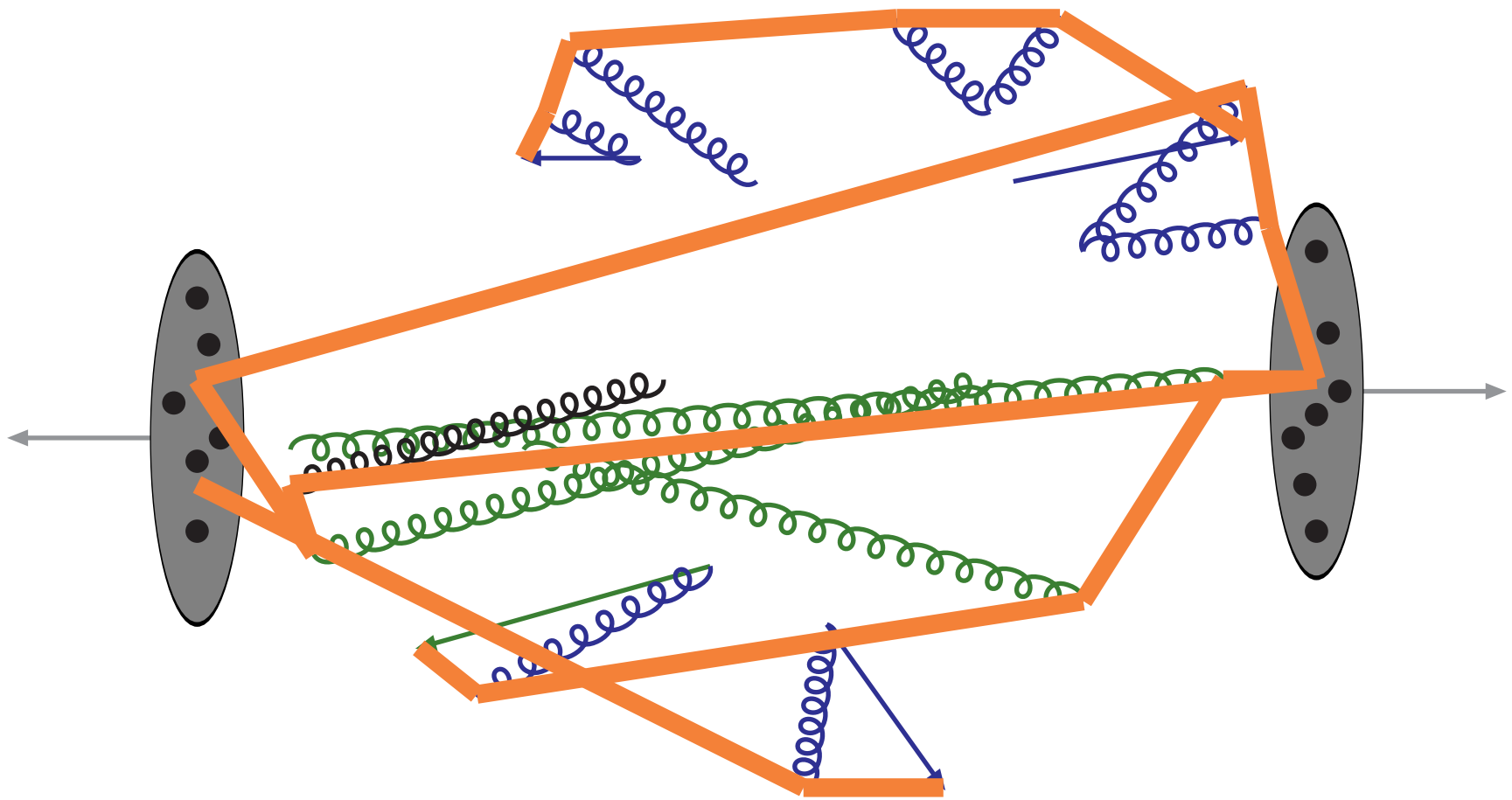
Multiple parton-parton interactions ...



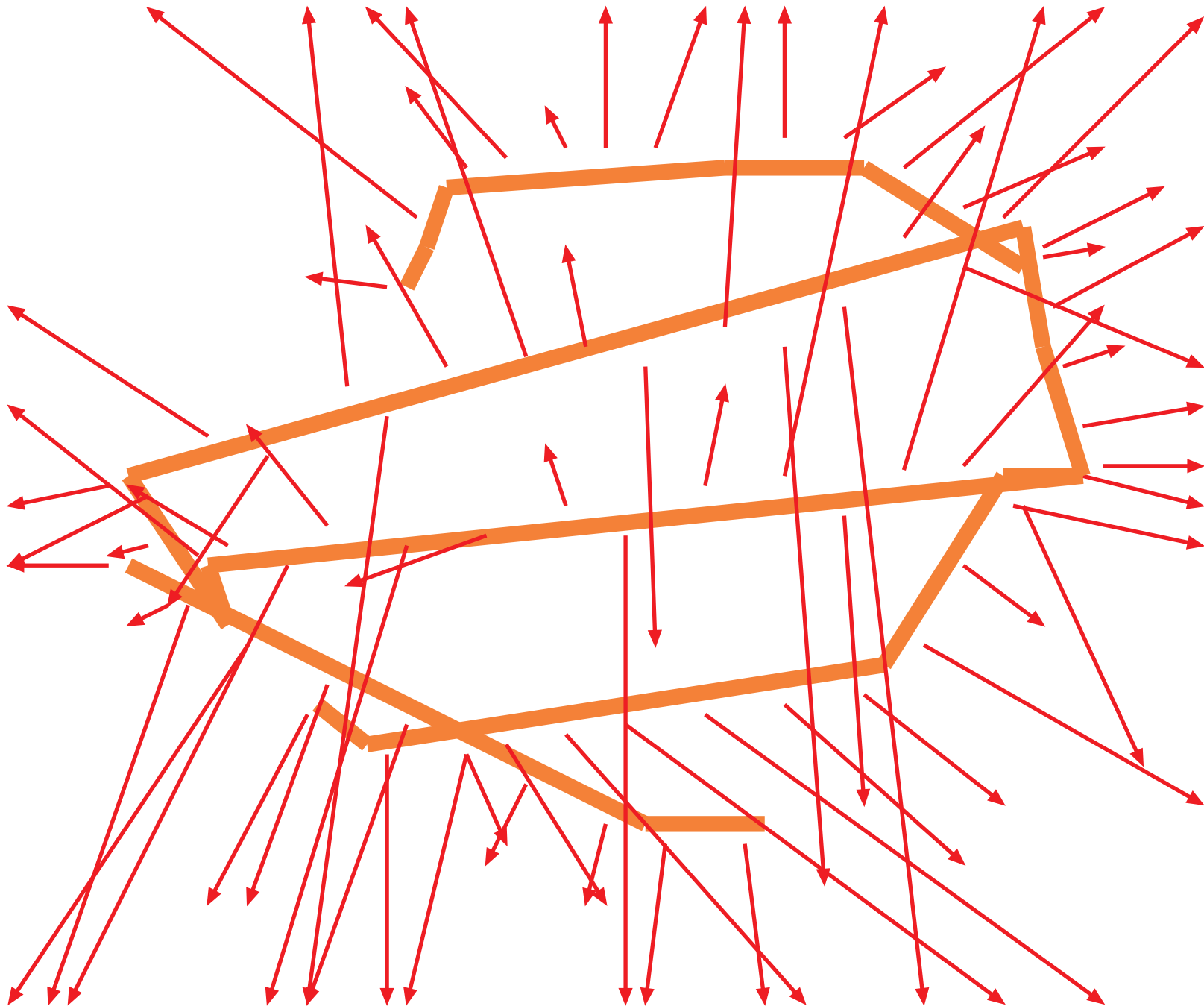
... with its **initial-** and **final-**state radiation



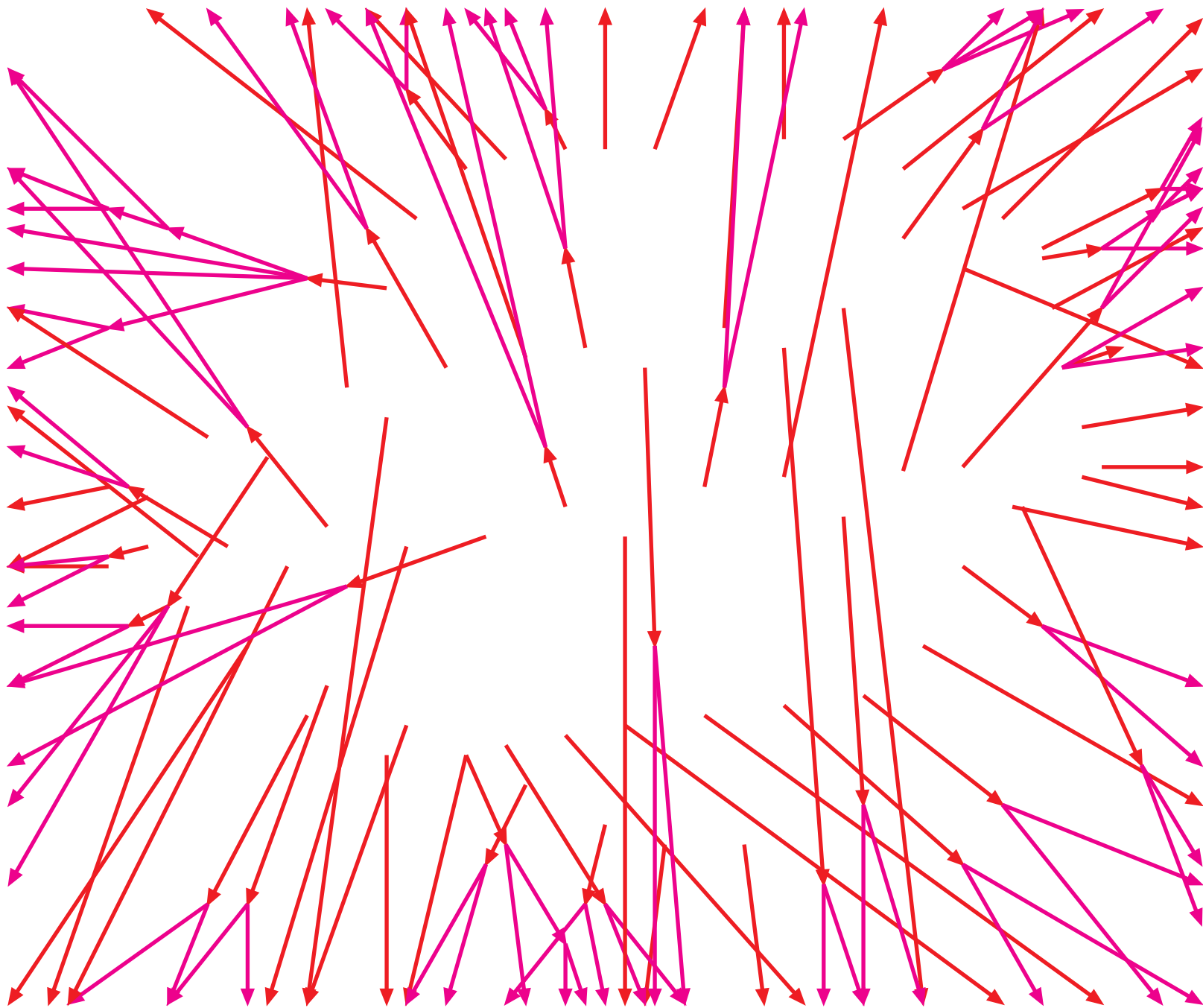
Beam remnants and other outgoing partons



Everything is connected by colour confinement strings
Recall! Not to scale: strings are of hadronic widths

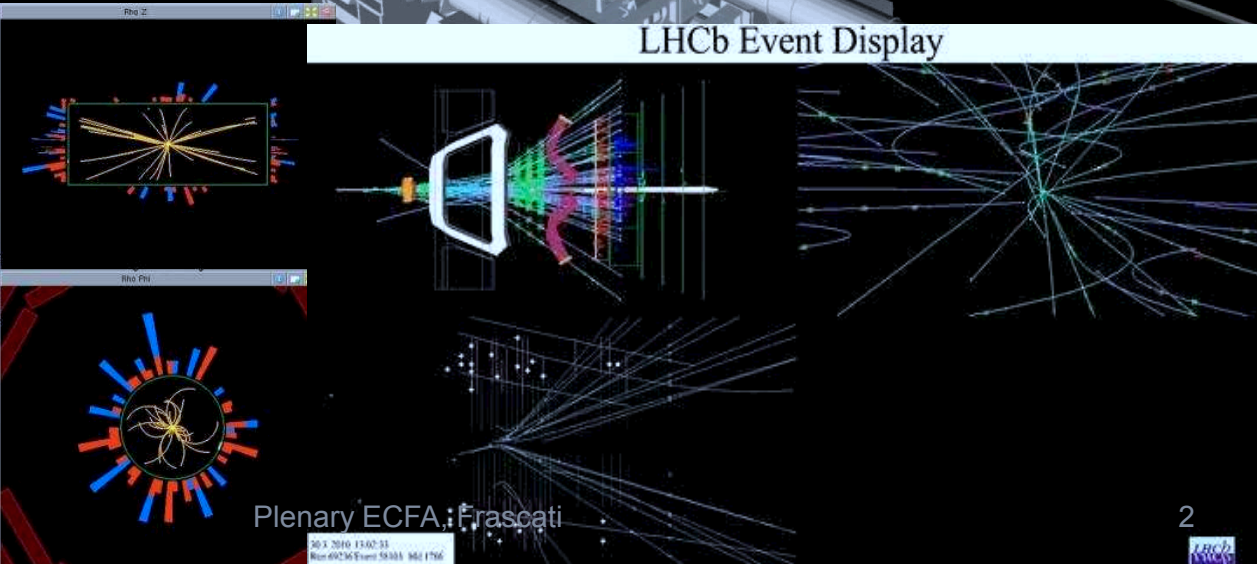
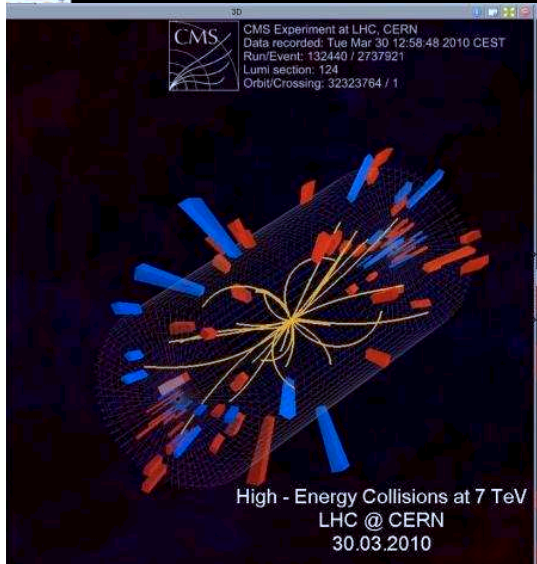
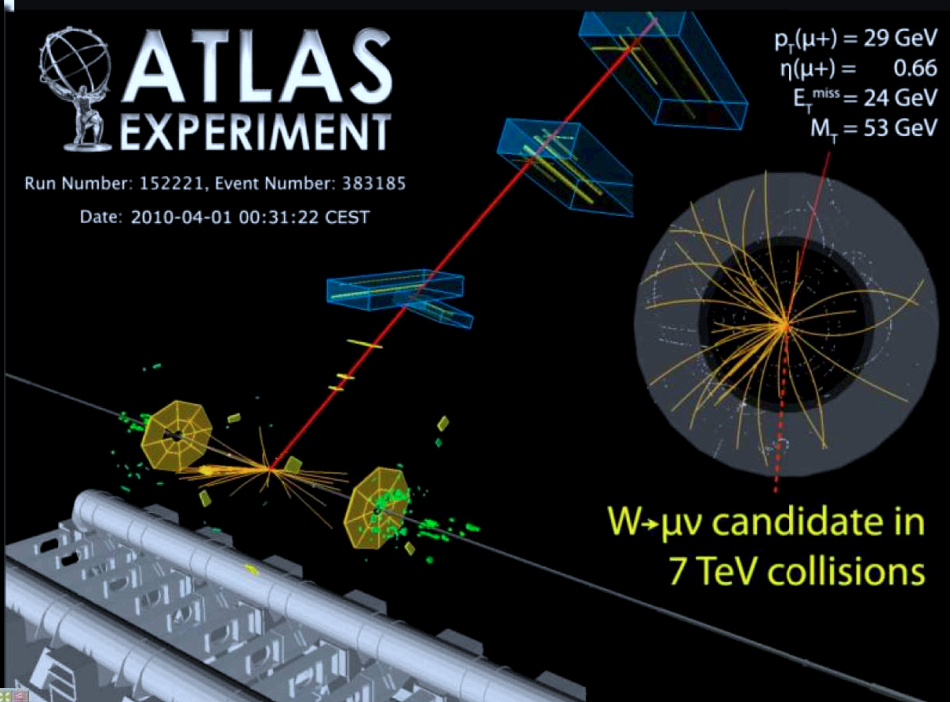
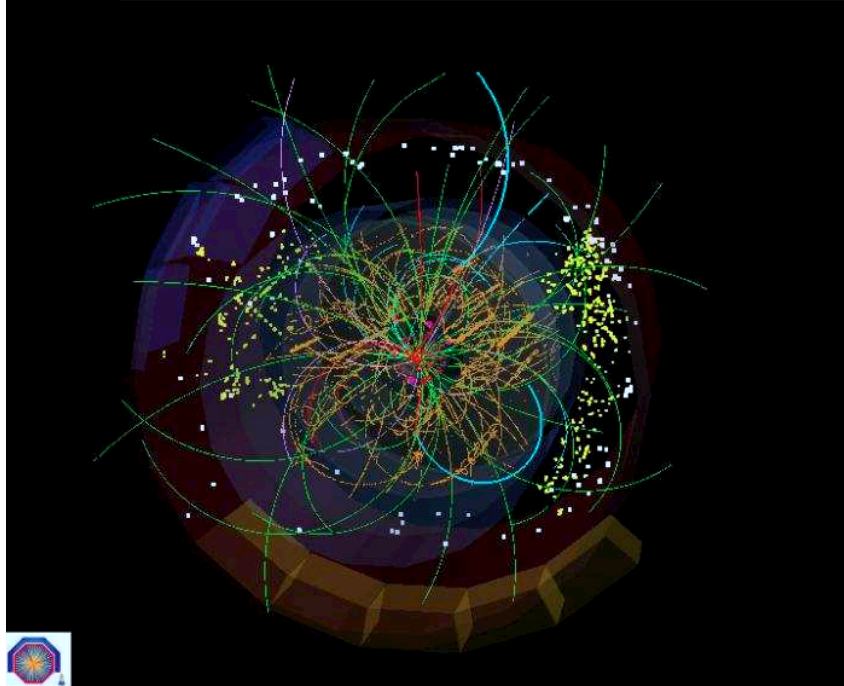


The strings fragment to produce primary hadrons



Many hadrons are unstable and decay further

March 30, 2010 7 TeV!



These are the particles that hit the detector

The Monte Carlo method

Want to generate events in as much detail as Mother Nature

⇒ get average *and* fluctuations right

⇒ make random choices, \sim as in nature

$$\sigma_{\text{final state}} = \sigma_{\text{hard process}} \mathcal{P}_{\text{tot,hard process} \rightarrow \text{final state}}$$

(appropriately summed & integrated over non-distinguished final states)

where $\mathcal{P}_{\text{tot}} = \mathcal{P}_{\text{res}} \mathcal{P}_{\text{ISR}} \mathcal{P}_{\text{FSR}} \mathcal{P}_{\text{MI}} \mathcal{P}_{\text{remnants}} \mathcal{P}_{\text{hadronization}} \mathcal{P}_{\text{decays}}$

with $\mathcal{P}_i = \prod_j \mathcal{P}_{ij} = \prod_j \prod_k \mathcal{P}_{ijk} = \dots$ in its turn

⇒ **divide and conquer**

an event with n particles involves $\mathcal{O}(10n)$ random choices,

(flavour, mass, momentum, spin, production vertex, lifetime, ...)

LHC: ~ 100 charged and ~ 200 neutral (+ intermediate stages)

⇒ several thousand choices

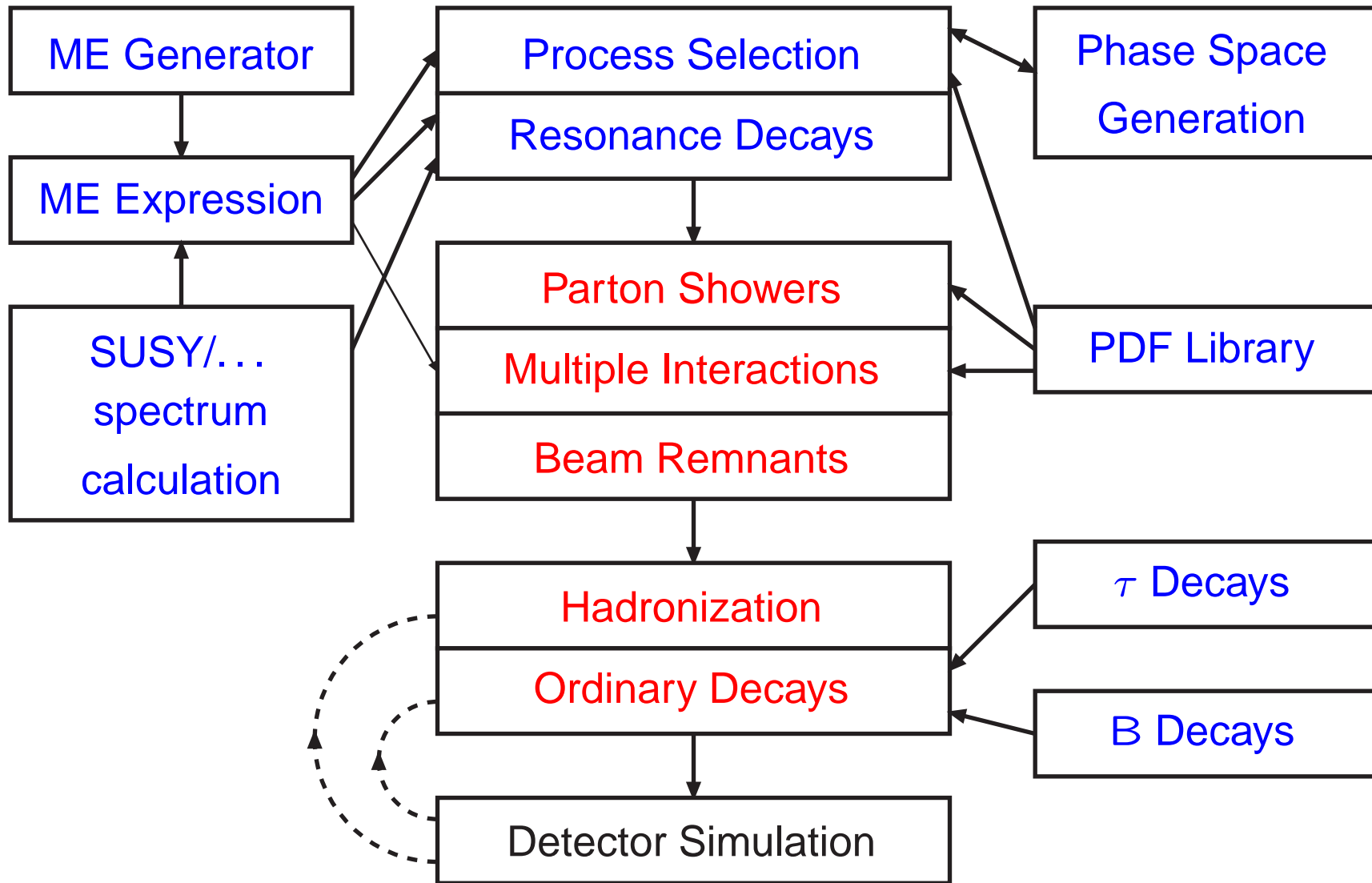
(of $\mathcal{O}(100)$ different kinds)

Generator Landscape

	General-Purpose	Specialized
Hard Processes	HERWIG PYTHIA SHERPA 	a lot
Resonance Decays		HDECAY, ...
Parton Showers		Ariadne/LDC, VINCIA, ...
Underlying Event		PHOJET/DPMJET
Hadronization		none (?)
Ordinary Decays		TAUOLA, EvtGen

specialized often best at given task, but need General-Purpose core

The Bigger Picture



need standardized interfaces (LHA/LHEF, LHAPDF, SUSY LHA, HepMC, ...)

PDG Particle Codes

A. Fundamental objects

1	d	11	e^-	21	g						add – sign for antiparticle, where appropriate + diquarks, SUSY, technicolor, ...
2	u	12	ν_e	22	γ	32	Z'^0				
3	s	13	μ^-	23	Z^0	33	Z''^0				
4	c	14	ν_μ	24	W^+	34	W'^+				
5	b	15	τ^-	25	h^0	35	H^0	37	H^+		
6	t	16	ν_τ			36	A^0	39	Graviton		

B. Mesons

$100 |q_1| + 10 |q_2| + (2s + 1)$ with $|q_1| \geq |q_2|$
 particle if heaviest quark u, \bar{s} , c, \bar{b} ; else antiparticle

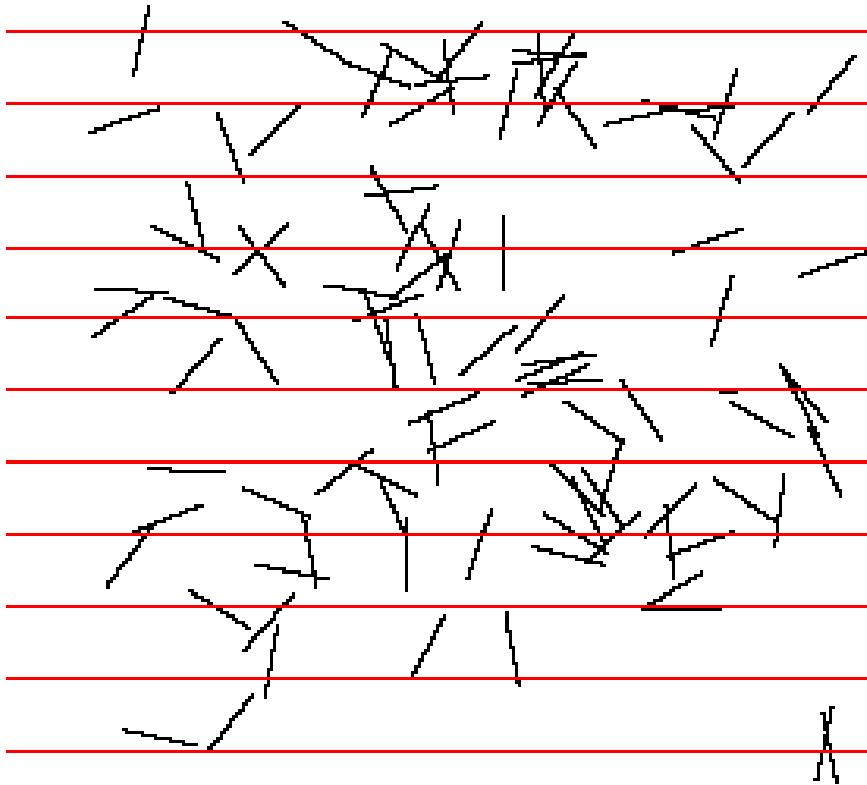
111	π^0	311	K^0	130	K_L^0	221	η^0	411	D^+	431	D_S^+
211	π^+	321	K^+	310	K_S^0	331	η'^0	421	D^0	443	J/ ψ

C. Baryons

$1000 q_1 + 100 q_2 + 10 q_3 + (2s + 1)$
 with $q_1 \geq q_2 \geq q_3$, or Λ -like $q_1 \geq q_3 \geq q_2$

2112	n	3122	Λ^0	2224	Δ^{++}	3214	Σ^{*0}
2212	p	3212	Σ^0	1114	Δ^-	3334	Ω^-

Monte Carlo Techniques



- Random Numbers
- Spatial Problems & Methods
- Temporal Problems & Methods

Buffon's needles

Random Numbers

Monte Carlo assume access to a good random number generator R :

(i) inclusively R is uniformly distributed in $0 < R < 1$

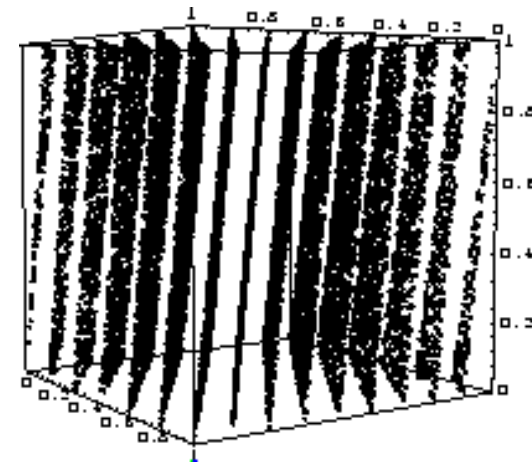
(ii) there are no correlations between R values along sequence

Radioactive decay \Rightarrow true random numbers

Computer algorithms \Rightarrow pseudorandom numbers

Many (in)famous pitfalls:

- short periods
- Marsaglia effect: multiplets along hyperplanes
 \Rightarrow do not trust “standard libraries” with compiler



Recommended:

- Marsaglia–Zaman–Tsang (RANMAR), improved by Lüscher (RANLUX):
can pick $\sim 900,000,000$ different sequences, each with period $> 10^{43}$
but state is specified by 100 words (97 double precision reals, 3 integers)
- l'Ecuyer (RANECU):
can pick 100 different sequences, each with period $> 10^{18}$, by two seeds

Spatial vs. Temporal Problems

“Spatial” problems: no memory

- 1) What is the land area of your home country?
+ Pick a point at random, with equal probability on this area.
- 2) What is the integrated cross section of a process?
+ Pick an event at random, according to the differential cross section.

“Temporal” problems: has memory

- 1) Traffic flow: What is probability for a car to pass a given point at time t , given traffic flow at earlier times?
Lumping from red lights, antilumping from finite size of cars!
- 2) Radioactive decay: what is the probability for a radioactive nucleus to decay at time t , given that it was created at time 0?
- 3) What is the probability for a parton to branch at a “virtuality” scale Q , given that it was created at a scale Q_0 ?

In particle physics normally combined;

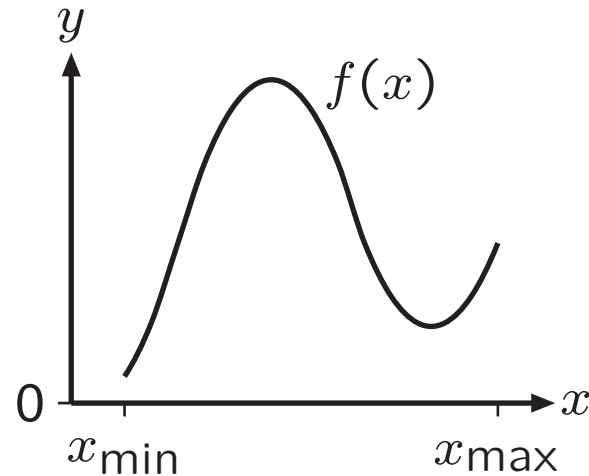
temporal evolution, but with spatial integral at each time:

What is the probability for a parton to branch at Q ,

with daughters sharing the mother momentum some specific way?

Spatial Methods

Assume function $f(x)$,
studied range $x_{\min} < x < x_{\max}$,
where $f(x) \geq 0$ everywhere
(in practice x is multidimensional)



Two standard tasks:

1) Calculate (approximatively)

$$\int_{x_{\min}}^{x_{\max}} f(x') dx'$$

usually: integrated cross section from differential one

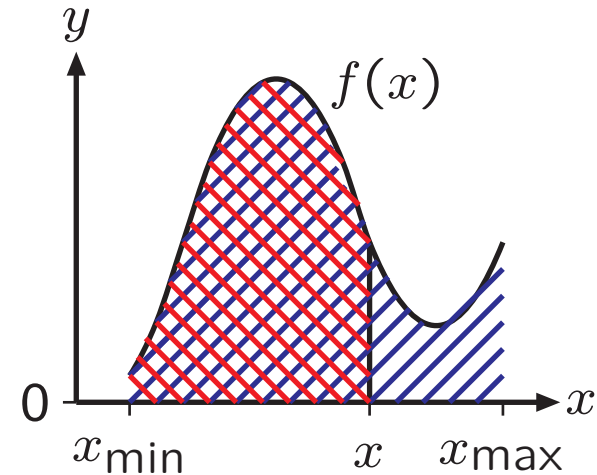
2) Select x at random according to $f(x)$

usually: probability distribution from quantum mechanics,
normalization to unit area implicit

Note n -dimensional integration $\equiv n + 1$ -dimensional volume:

$$\int f(x_1, \dots, x_n) dx_1 \dots dx_n \equiv \int \int_0^{f(x_1, \dots, x_n)} 1 dx_1 \dots dx_n dx_{n+1}$$

Selection of x according to $f(x)$
 is equivalent to uniform selection of (x, y) in the area
 $x_{\min} < x < x_{\max}$, $0 < y < f(x)$
 since $\mathcal{P}(x) \propto \int_0^{f(x)} 1 \, dy = f(x)$



Therefore

$$\int_{x_{\min}}^x f(x') \, dx' = R \int_{x_{\min}}^{x_{\max}} f(x') \, dx'$$

Method 1: Analytical solution

If know primitive function $F(x)$ and know inverse $F^{-1}(y)$ then

$$\begin{aligned} F(x) - F(x_{\min}) &= R (F(x_{\max}) - F(x_{\min})) = R A_{\text{tot}} \\ \implies x &= F^{-1}(F(x_{\min}) + R A_{\text{tot}}) \end{aligned}$$

Proof:

introduce $z = F(x_{\min}) + R A_{\text{tot}}$. Then

$$\frac{d\mathcal{P}}{dx} = \frac{d\mathcal{P}}{dR} \frac{dR}{dx} = 1 \frac{1}{\frac{dx}{dR}} = \frac{1}{\frac{dx}{dz} \frac{dz}{dR}} = \frac{1}{\frac{dF^{-1}(z)}{dz} \frac{dz}{dR}} = \frac{\frac{dF(x)}{dx}}{\frac{dz}{dR}} = \frac{f(x)}{A_{\text{tot}}}$$

Example 1:

$$f(x) = 2x, 0 < x < 1, \implies F(x) = x^2$$

$$F(x) - F(0) = R(F(1) - F(0)) \implies x^2 = R \implies x = \sqrt{R}$$

Example 2:

$$f(x) = e^{-x}, x > 0, F(x) = 1 - e^{-x}$$

$$1 - e^{-x} = R \implies e^{-x} = 1 - R = R \implies x = -\ln R$$

Method 2: Hit-and-miss

If $f(x) \leq f_{\max}$ in $x_{\min} < x < x_{\max}$

use **interpretation as an area**

1) select $x = x_{\min} + R(x_{\max} - x_{\min})$

2) select $y = R f_{\max}$ (new R !)

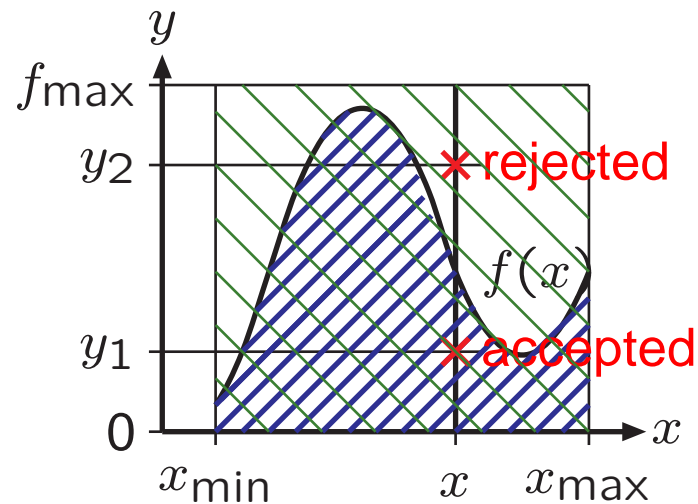
3) while $y > f(x)$ cycle to 1)

Integral as by-product:

$$I = \int_{x_{\min}}^{x_{\max}} f(x) dx = f_{\max} (x_{\max} - x_{\min}) \frac{N_{\text{acc}}}{N_{\text{try}}} = A_{\text{tot}} \frac{N_{\text{acc}}}{N_{\text{try}}}$$

Binomial distribution with $p = N_{\text{acc}}/N_{\text{try}}$ and $q = N_{\text{fail}}/N_{\text{try}}$, so error

$$\frac{\delta I}{I} = \frac{A_{\text{tot}} \sqrt{pq/N_{\text{try}}}}{A_{\text{tot}} p} = \sqrt{\frac{q}{p N_{\text{try}}}} = \sqrt{\frac{q}{N_{\text{acc}}}} \rightarrow \frac{1}{\sqrt{N_{\text{acc}}}} \text{ for } p \ll 1$$



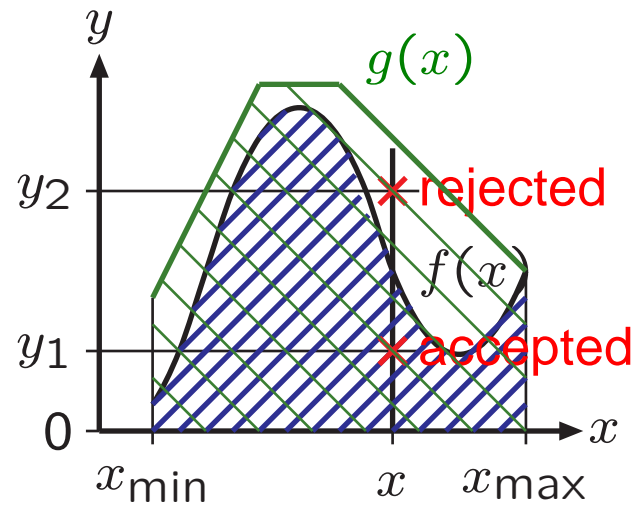
Method 3: Improved hit-and-miss (importance sampling)

If $f(x) \leq g(x)$ in $x_{\min} < x < x_{\max}$

and $G(x) = \int g(x') dx'$ is simple

and $G^{-1}(y)$ is simple

- 1) select x according to $g(x)$ distribution
- 2) select $y = R g(x)$ (new $R!$)
- 3) while $y > f(x)$ cycle to 1)



Example 3:

$$f(x) = x e^{-x}, x > 0$$

Attempt 1: $F(x) = 1 - (1 + x) e^{-x}$ not invertible

Attempt 2: $f(x) \leq f(1) = e^{-1}$ but $0 < x < \infty$

Attempt 3: $g(x) = N e^{-x/2}$

$$\frac{f(x)}{g(x)} = \frac{x e^{-x}}{N e^{-x/2}} = \frac{x e^{-x/2}}{N} \leq 1$$

for rejection to work, so find maximum:

$$\frac{d}{dx} \left(\frac{f(x)}{g(x)} \right) = \frac{1}{N} \left(1 - \frac{x}{2} \right) e^{-x/2} = 0 \implies x = 2$$

Normalize so $g(2) = f(2) \implies N = 2/e$

$$G(x) \propto 1 - e^{-x/2} = R$$

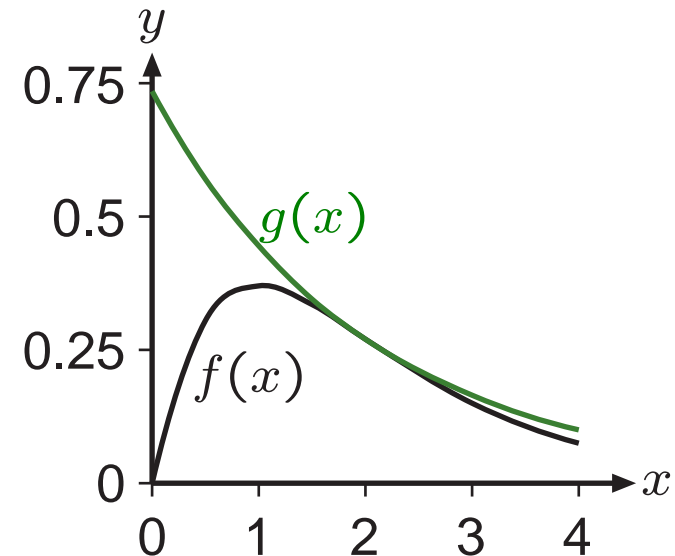
$$\implies x = -2 \ln R \text{ so}$$

$$1) \text{ select } x = -2 \ln R$$

$$2) \text{ select } y = R g(x) = R 2e^{-(1+x/2)}$$

$$3) \text{ while } y > f(x) = x e^{-x} \text{ cycle to 1)}$$

$$\text{efficiency} = \frac{\int_0^\infty f(x) dx}{\int_0^\infty g(x) dx} = \frac{e}{4}$$



Attempt 4: pull the rabbit ...

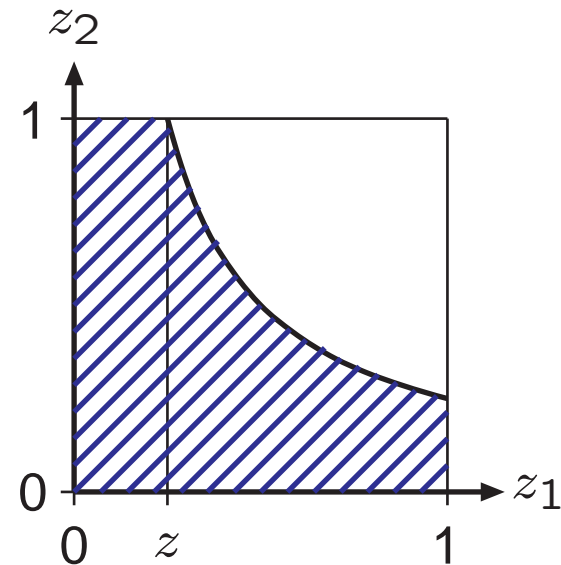
$$x = -\ln(R_1 R_2)$$

$$\text{since with } z = z_1 z_2 = R_1 R_2$$

$$\begin{aligned} F(z) &= \int_0^z f(z') dz' \\ &= \int_0^z 1 dz_1 + \int_z^1 \frac{z}{z_1} dz_1 \\ &= z - z \ln z \end{aligned}$$

$$\text{and using that } x = -\ln z \iff z = e^{-x}$$

$$F(x) = 1 - F(z = e^{-x}) = 1 - e^{-x} + e^{-x} (-x) \implies f(x) = x e^{-x}$$



Method 4: Multichannel

If $f(x) \leq g(x) = \sum_i g_i(x)$,

where all g_i “nice” (but $g(x)$ not)

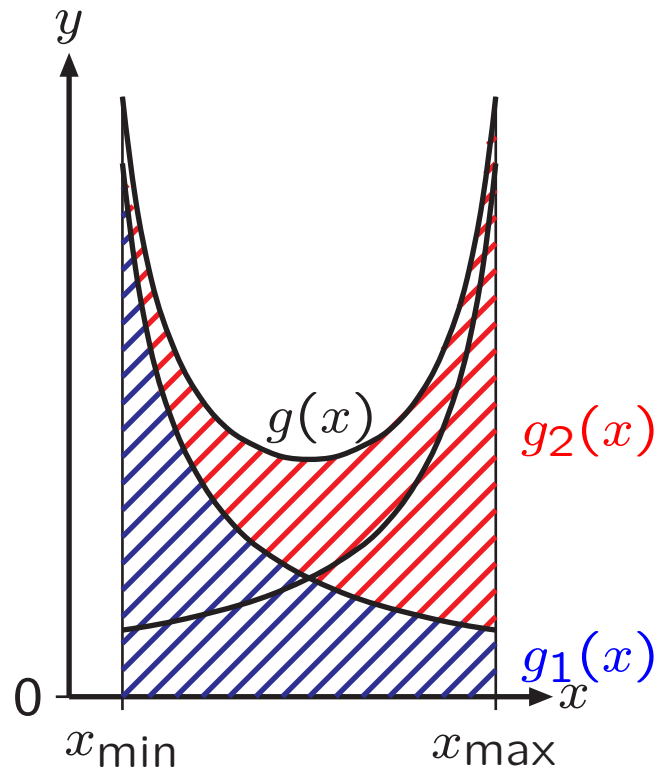
1) select i with relative probability

$$A_i = \int_{x_{\min}}^{x_{\max}} g_i(x') dx'$$

2) select x according to $g_i(x)$

3) select $y = R g(x) = R \sum_i g_i(x)$

4) while $y > f(x)$ cycle to 1)



Example 4:

$$f(x) = \frac{1}{\sqrt{x(1-x)}}, \quad 0 < x < 1$$

$$g(x) = \frac{1}{\sqrt{x}} + \frac{1}{\sqrt{1-x}} = \frac{\sqrt{x} + \sqrt{1-x}}{\sqrt{x(1-x)}}, \quad \frac{1}{\sqrt{2}} \leq \frac{f(x)}{g(x)} \leq 1$$

1) if $R < 1/2$ then $g_1(x)$ else $g_2(x)$

2) $g_1: G_1(x) = 2\sqrt{x} = 2R \implies x = R^2$

$g_2: G_2(x) = 2(1 - \sqrt{1-x}) = 2R \implies x = 1 - R^2$

Method 5: Variable transformations

- map to finite x range
- map away singular/peaked regions

Method 6: Special tricks

e.g. $f(x) \propto e^{-x^2}$ is not integrable, but

$$\begin{aligned} f(x) dx f(y) dy &\propto e^{-(x^2+y^2)} dx dy \\ &= e^{-r^2} r dr d\phi \propto e^{-r^2} dr^2 d\phi \\ F(r^2) = 1 - e^{-r^2} &\implies r^2 = -\ln R_1 \\ x &= \sqrt{-\ln R_1} \cos(2\pi R_2) \\ y &= \sqrt{-\ln R_1} \sin(2\pi R_2) \end{aligned}$$

Comment:

In practice almost always multidimensional integrals

$$\int_V f(\mathbf{x}) d\mathbf{x} = V \frac{1}{N_{\text{try}}} \sum_i f(\mathbf{x}_i) \quad \text{or} \quad = \int_V g(\mathbf{x}) d\mathbf{x} \frac{N_{\text{acc}}}{N_{\text{try}}}$$

gives error $\propto 1/\sqrt{N}$ irrespective of dimension

whereas trapezium rule error $\propto 1/N^2 \rightarrow 1/N^{2/d}$ in d dimensions,

and Simpson's rule error $\propto 1/N^4 \rightarrow 1/N^{4/d}$ in d dimensions

Temporal methods: The Veto Algorithm

Consider “radioactive decay”:

$N(t)$ = number of remaining nuclei at time t

but normalized to $N(0) = 1$ instead, so equivalently

$N(t)$ = probability that (single) nucleus has not decayed by time t

$P(t) = -dN(t)/dt$ = probability for decay at time t

Normally $P(t) = cN(t)$, with c constant, but assume time-dependence:

$$P(t) = -\frac{dN(t)}{dt} = f(t)N(t) ; f(t) \geq 0$$

Standard solution:

$$\frac{dN(t)}{dt} = -f(t)N(t) \iff \frac{dN}{N} = d(\ln N) = -f(t) dt$$

$$\ln N(t) - \ln N(0) = -\int_0^t f(t') dt' \implies N(t) = \exp\left(-\int_0^t f(t') dt'\right)$$

$$F(t) = \int_0^t f(t') dt' \implies N(t) = \exp(-(F(t) - F(0)))$$

$$N(t) = R \implies t = F^{-1}(F(0) - \ln R)$$

What now if $f(t)$ has no simple $F(t)$ or F^{-1} ?

Hit-and-miss not good enough, since for $f(t) \leq g(t)$, g "nice",

$$t = G^{-1}(G(0) - \ln R) \implies N(t) = \exp\left(-\int_0^t g(t') dt'\right)$$

$$P(t) = -\frac{dN(t)}{dt} = g(t) \exp\left(-\int_0^t g(t') dt'\right)$$

and hit-or-miss provides rejection factor $f(t)/g(t)$, so that

$$P(t) = f(t) \exp\left(-\int_0^t g(t') dt'\right)$$

where it ought to have been

$$P(t) = f(t) \exp\left(-\int_0^t f(t') dt'\right)$$

Correct answer is:

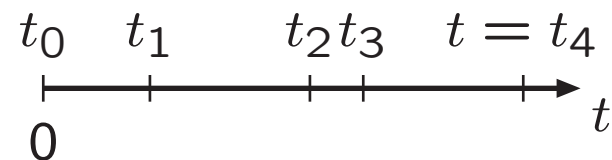
0) start with $i = 0$ and $t_0 = 0$

1) $++i$ (i.e. increase i by one)

2) $t_i = G^{-1}(G(t_{i-1}) - \ln R)$, i.e. $t_i > t_{i-1}$

3) $y = R g(t)$

4) while $y > f(t)$ cycle to 1)



Proof:

define $S_g(t_a, t_b) = \exp\left(-\int_{t_a}^{t_b} g(t') dt'\right)$

$$P_0(t) = P(t = t_1) = g(t) S_g(0, t) \frac{f(t)}{g(t)} = f(t) S_g(0, t)$$

$$\begin{aligned} P_1(t) &= P(t = t_2) = \int_0^t dt_1 g(t_1) S_g(0, t_1) \left(1 - \frac{f(t_1)}{g(t_1)}\right) g(t) S_g(t_1, t) \frac{f(t)}{g(t)} \\ &= f(t) S_g(0, t) \int_0^t dt_1 (g(t_1) - f(t_1)) = P_0(t) I_{g-f} \end{aligned}$$

$$\begin{aligned} P_2(t) &= \dots = P_0(t) \int_0^t dt_1 (g(t_1) - f(t_1)) \int_{t_1}^t dt_2 (g(t_2) - f(t_2)) \\ &= P_0(t) \int_0^t dt_1 (g(t_1) - f(t_1)) \int_0^t dt_2 (g(t_2) - f(t_2)) \theta(t_2 - t_1) \\ &= P_0(t) \frac{1}{2} \left(\int_0^t dt_1 (g(t_1) - f(t_1)) \right)^2 = P_0(t) \frac{1}{2} I_{g-f}^2 \end{aligned}$$

$$\begin{aligned} P(t) &= \sum_{i=0}^{\infty} P_i(t) = P_0(t) \sum_{i=0}^{\infty} \frac{I_{g-f}^i}{i!} = P_0(t) \exp(I_{g-f}) \\ &= f(t) \exp\left(-\int_0^t g(t') dt'\right) \exp\left(\int_0^t dt_1 (g(t_1) - f(t_1))\right) \\ &= f(t) \exp\left(-\int_0^t f(t') dt'\right) \end{aligned}$$

Temporal methods: The Winner Takes It All

Assume “radioactive decay” with two possible decay channels 1 & 2

$$P(t) = -\frac{dN(t)}{dt} = f_1(t)N(t) + f_2(t)N(t)$$

Alternative 1: use normal veto algorithm with $f(t) = f_1(t) + f_2(t)$.
Once t selected, pick decays 1 or 2 in proportions $f_1(t) : f_2(t)$.

Alternative 2: pick t_1 according to $P_1(t_1) = f_1(t_1)N_1(t_1)$
and t_2 according to $P_2(t_2) = f_2(t_2)N_2(t_2)$.
If $t_1 < t_2$ then pick decay 1, while if $t_2 < t_1$ decay 2.

Proof:

$$\begin{aligned} P_1(t) &= (f_1(t) + f_2(t)) \exp\left(-\int_0^t (f_1(t') + f_2(t')) dt'\right) \frac{f_1(t)}{f_1(t) + f_2(t)} \\ &= f_1(t) \exp\left(-\int_0^t (f_1(t') + f_2(t')) dt'\right) \\ &= f_1(t) \exp\left(-\int_0^t f_1(t') dt'\right) \exp\left(-\int_0^t f_2(t') dt'\right) \end{aligned}$$

Especially convenient when temporal and/or spatial dependence of f_1 and f_2 are rather different.

Summary Lecture 1

- Event generators indispensable ●
- Quantum Mechanics \implies probabilities ●
 - ★ Divide and conquer ★
 - Main physics components: ●
 - ★ Hard processes and resonance decays ★
 - ★ Initial- and final-state radiation ★
 - ★ Multiple parton–parton interactions and beam remnants ★
 - ★ Hadronization and decays ★
- Monte Carlo Techniques: ●
 - ★ Use good random number generator ★
 - ★ Monte Carlo = selection *and* integration ★
 - ★ Adapt Monte Carlo approach to problem at hand ★
 - ★ Multichannel and Veto algorithms common ★