How to perform collider studies
Part I

Ben O'Leary

Julius-Maximilians-Universität Würzburg

German-Egyptian School of Particle Physics,
Zewail City of Science and Technology, Giza,
February 25th, 2013

How to perform collider studies

Part I:

Motivation and planning

Outline

Definitions

What is "a collider study" and how do you "perform" one?

- For my purposes, a "collider study" is generating an appropriate simulation of a particle collider and analyzing the generated events, using appropriate tools.

You need a plan

Performing a collider study starts *long* before installing `WHIZARD/CalcHEP/...` - you need to **plan**!

(1) Why are you generating a Monte Carlo (MC) simulation?

(2) What exactly are you going to simulate?

(3) How will you create the simulation?

Why are you generating an MC simulation?

- Compare *theory* to *experiment*
- Compare predicted particle properties to measured properties:
  - mass m_X
  - charge q_X
  - spin s_X
  - lifetime τ_X / decay width $\Gamma_X = \hbar/\tau_X$
  - branching ratios $\mathrm{BR}(X \to ...)$
  - production cross-sections $\sigma(X, ...)$

- τ_X long enough $\rightarrow$
  - m_X, q_X, s_X, τ_X, BRs already measured
  - calculate directly, compare to observed value, no need to simulate
  - *e.g.* define model, input (maybe via RGEs) $\rightarrow$ prediction for $\mathrm{BR}(B_s \rightarrow \mu\bar{\mu})$, compare to $1.1 \times 10^9 < \mathrm{BR}(B_s \rightarrow \mu\bar{\mu}) < 6.4 \times 10^9$

- τ_X too short $\rightarrow$
  - (binned) σ_i(various signals, cuts) as functions of m_X, q_X, s_X, Γ_X, various couplings, ...
  - generally need to simulate

Typical aims of collider studies:

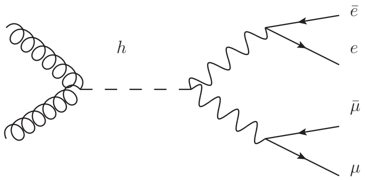
- Is this model compatible with observed excesses or non-observations?
- What cuts maximize observability of this signal?
- What is the distribution of cross-section against this kinematic variable?

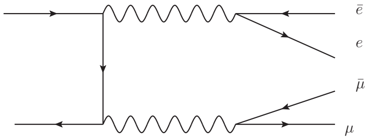
Is this model compatible with observed excesses or non-observations?

- Often not obvious if a model or parameter space region of a model can explain an observed excess.
- Also not obvious if model is incompatible with exclusions.
- Usually sufficient to simulate model's excess compared to SM contribution and compare to experiment's excess compared to SM.

What cuts maximize observability of this signal?

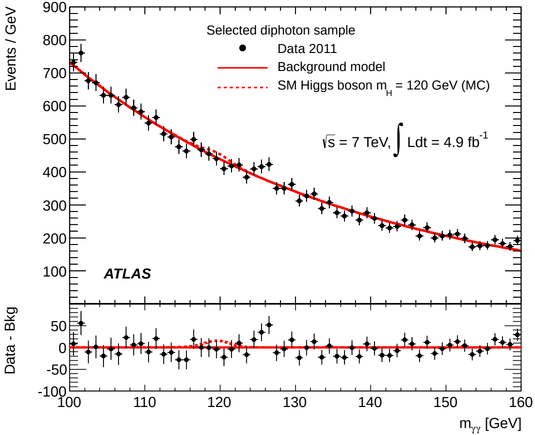
- Minimizing backgrounds while cutting as little signal as possible is very important.
- Background needs to be simulated as well as signal, obviously.





What is the distribution of cross-section against this kinematic variable?

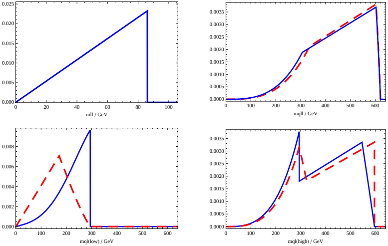
- Classic bump-hunting, *e.g.* $\mathrm{d}\sigma/\mathrm{d}m_{\gamma\gamma}$



(Credit to Atlas for the picture!)

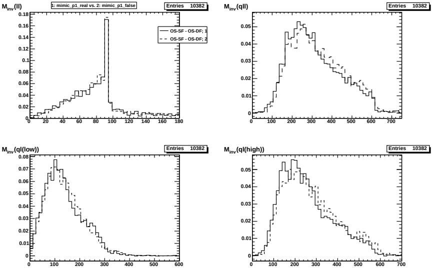
- Other cuts to reduce background can affect distribution (from work with Jonas Lindert):

$m_{\ell\ell}$ $m_{q\ell\ell}$



$m_{q\ell}$(low) $m_{q\ell}$(high)

Theoretical shapes



MC, just SUSY, perfect reconstruction with hardest jet

Why do we need MC rather than grad students + pencils + paper?

- We need to compare model's $\sigma_{\text{theo}}(\{m_i\}, \{g_j\}, \{Y_k\}, ...)$ to measured σ_{exp}.
- Theory should be *predictive* and map to *physical observables*.
- Observables from parameters often not easy, even at lowest orders - *e.g.* 4×4 mass matrix (MSSM neutralinos).
- Colliders bring their own complications...

Complication 1

Complication 1: multiparticle phase spaces

- $1 \to n$ decays and $2 \to n$ scatterings require integrations over n-body phase spaces
- $\Gamma(a \to n) = \int \mathrm{d}P_n \frac{|\mathcal{M}|^2}{2m_a}$
- $\sigma(a, b \to n) = \int \mathrm{d}P_n \frac{|\mathcal{M}|^2}{4|E_a p_b^z - E_b p_a^z|}$
- $\mathrm{d}P_n = (2\pi)^4 \delta^4(\sum p_{\mathrm{in}} - \sum p_{\mathrm{out}}) \Pi_i \frac{\mathrm{d}\mathbf{p}}{(2\pi)^3 (2E_i)}$

Two-body phase space

$n = 2$, center-of-momentum frame:

$$\mathrm{d}P_2 = (2\pi)^4\delta^3(\mathbf{p}_1+\mathbf{p}_2)\delta\left(\sum E_{\mathrm{in}}-E_1-E_2\right)\frac{\mathrm{d}\mathbf{p}_1}{(2\pi)^3(2E_1)}\frac{\mathrm{d}\mathbf{p}_2}{(2\pi)^3(2E_2)}$$

$$\int |\mathcal{M}(p_1,p_2)|^2\mathrm{d}P_2 = \int \frac{|\mathcal{M}(p_1,p_2)|^2}{(2\pi)^2(2E_2)}\delta\left(\sum E_{\mathrm{in}} - E_1 - E_2\right)\frac{\mathrm{d}\mathbf{p}_1}{(2E_1)}$$

where

$$\mathbf{p}_2 = -\mathbf{p}_1, E_2 = \sqrt{|\mathbf{p}_1|^2 + m_2^2}, E_1 = \sqrt{|\mathbf{p}_1|^2 + m_1^2}$$

$$\mathrm{d}\mathbf{p}_1 = |\mathbf{p}_1|^2 \mathrm{d}|\mathbf{p}_1| \mathrm{d}\Omega_1 = |\mathbf{p}_1| E_1 \mathrm{d}E_1 \mathrm{d}\Omega_1$$

where $|\mathbf{p}_1| = \sqrt{E_1^2 - m_1^2}$
$\delta(\sum E_{\mathrm{in}} - E_1 - E_2) = \frac{\delta(E_1 - E_1^{\mathrm{sol.}})}{|1 + (\mathrm{d}E_2/\mathrm{d}E_1)|}$

$E_2 = \sqrt{|\mathbf{p}_1|^2 + m_2^2} = \sqrt{E_1^2 - m_1^2 + m_2^2}$
$\delta(\sum E_{\mathrm{in}} - E_1 - E_2) = \frac{E_2 \delta(E_1 - E_1^{\mathrm{sol.}})}{\sum E_{\mathrm{in}}}$

$$\int |\mathcal{M}(p_1, p_2)|^2 \mathrm{d}P_2 =$$
$$\int \frac{|\mathcal{M}(p_1,p_2)|^2}{(2\pi)^2(2E_2)} \delta\left(\sum E_{\text{in}} - E_1 - E_2\right) \frac{\mathrm{d}\mathbf{p}_1}{(2E_1)}$$

$$= \int \frac{|\mathcal{M}(p_1,p_2)|^2}{16\pi^2} \delta\left(E_1 - E_1^{\text{sol.}}\right) \frac{|\mathbf{p}_1|}{\sum E_{\text{in}}} \mathrm{d}E_1 \mathrm{d}\Omega_1$$
$$= \int \frac{|\mathcal{M}(p_1,p_2)|^2}{16\pi^2} \frac{|\mathbf{p}_1|}{\sum E_{\text{in}}} \mathrm{d}\Omega_1$$

with all energies and momenta fixed by 4-momentum conservation.

Three-body phase space

$n = 3$, center-of-momentum frame:

$$\int |\mathcal{M}(p_1, p_2, p_3)|^2 \mathrm{d}P_3 =$$
$$\int \frac{|\mathcal{M}|^2}{2^8 \pi^5 E_1 E_2 E_3} \delta(\sum E_{\text{in}} - E_1 - E_2 - E_3)$$
$$\delta^3(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \mathrm{d}\mathbf{p}_1 \mathrm{d}\mathbf{p}_2 \mathrm{d}\mathbf{p}_3$$
$$= \int \frac{|\mathcal{M}|^2}{2^8 \pi^5 E_1 E_2 E_3} \delta(\sum E_{\text{in}} - E_1 - E_2 - E_3)$$
$$|\mathbf{p}_1|^2 \mathrm{d}|\mathbf{p}_1| \mathrm{d}\Omega_1 |\mathbf{p}_2|^2 \mathrm{d}|\mathbf{p}_2| \mathrm{d}\Omega_2$$

where E_3 is a function of $\mathbf{p}_1$ and $\mathbf{p}_2$, and $\mathbf{p}_3 = -\mathbf{p}_1 - \mathbf{p}_2$

$$E_3^2 - m_3^2 = |\mathbf{p}_1 + \mathbf{p}_2|^2 =$$
$$|\mathbf{p}_1|^2 + |\mathbf{p}_2|^2 + |\mathbf{p}_1||\mathbf{p}_2| \cos(\theta_{12})$$

$$E_3 \mathrm{d}E_3 = |\mathbf{p}_1||\mathbf{p}_2| \mathrm{d}[\cos(\theta_{12})]$$
$$E_2 \mathrm{d}E_2 = |\mathbf{p}_2| \mathrm{d}|\mathbf{p}_2|$$
$$E_1 \mathrm{d}E_1 = |\mathbf{p}_1| \mathrm{d}|\mathbf{p}_1|$$
$$\mathrm{d}\Omega_2 = \mathrm{d}\phi_{12} \mathrm{d}[\cos(\theta_{12})]$$

$\int |\mathcal{M}(p_1, p_2, p_3)|^2 \mathrm{d}P_3 =$
$\int \frac{|\mathcal{M}|^2}{2^8 \pi^5} \delta(\sum E_{\text{in}} - E_1 - E_2 - E_3) \mathrm{d}\Omega_1 \mathrm{d}\phi_{12} \mathrm{d}E_1 \mathrm{d}E_2 \mathrm{d}E_3$
$= 2^{-8} \pi^{-5} \int |\mathcal{M}(p_1, p_2, p_3)|^2 \mathrm{d}\Omega_1 \mathrm{d}\phi_{12} \mathrm{d}E_1 \mathrm{d}E_2$
where p_1, p_2, p_3 are functions of E_1, E_2, angles.

Decay into massless particles:

$p_1 \cdot p_2 = m(E_1 + E_2 - (m/2))$
$p_2 \cdot p_3 = m((m/2) - E_1)$
$p_3 \cdot p_1 = m((m/2) - E_2)$
E_2 range: $(m/2) - E_1$ to $(m/2)$
E_1 range: 0 to $(m/2)$

Four-body phase space

- $n = 4$? Too much work for $n = 3$ already!
- Massive decay products already make $n = 3$ integral limits very complicated.
- We could build $1 \to 4$ out of $1 \to 2$ followed by subsequent $1 \to 2$ decays – assuming that interference is negligible! (NWA: later.)

Complication 2

Complication 2: kinematic cuts

- Real detectors do not cover full solid angle (*e.g.* beam pipes).
- Real detectors do not trigger on arbitrarily soft particles.
- Often want to remove backgrounds with typical kinematic configuration.

$\int |\mathcal{M}(p_1, p_2)|^2 \mathrm{d}P_2 \Rightarrow$
$\int |\mathcal{M}(p_1, p_2)|^2 \Theta(\theta_{\mathrm{max}} - \theta)\Theta(\theta - \theta_{\mathrm{min}}) \mathrm{d}P_2$

$$= \int \frac{|\mathcal{M}(p_1,p_2)|^2}{16\pi^2} \frac{|\mathbf{p}_1|}{\sum E_{\mathrm{in}}} \Theta(\theta_{\mathrm{max}} - \theta)\Theta(\theta - \theta_{\mathrm{min}}) \mathrm{d}\Omega_1$$

$$= \int_{\cos(\theta_{\mathrm{min}})}^{\cos(\theta_{\mathrm{max}})} \mathrm{d}[\cos(\theta)] \frac{|\mathcal{M}(p_1,p_2)|^2}{8\pi} \frac{|\mathbf{p}_1|}{\sum E_{\mathrm{in}}}$$

assuming azimuthal symmetry.

- More than 2 final-state particles $\rightarrow$ intractible phase-space integrals, even approximating with $1 \rightarrow 2$ multistage decays
- Limiting phase space to region where *e.g.* $|m_{e^+e^-} - m_Z| > 2\Gamma_Z$, integrating over other particles is hopeless. (Even writing $m_{e^+e^-}$ in terms of p_i is often a non-starter.)

Complication 3: invisible particles

- Neutrinos escape detection at colliders.
- Dark matter candidates also escape detection.
- *Sometimes* we can reconstruct *one* invisible particle.
- Lack of reconstruction $\Rightarrow$ difficulty in measuring things even independent of σ complications.

Sometimes we can deal with invisible particles:

e.g. can we measure m_h from $h \to \tau^+\tau^-$?

- $m_h^2 = (p_{\tau^+} + p_{\tau^-})^2$
- Unfortunately, τ leptons always decay to a final state with at least 1 neutrino...

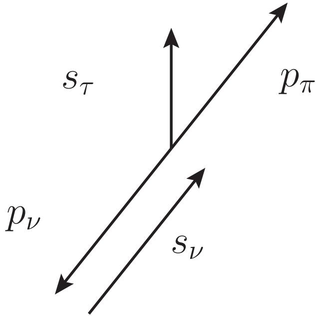
Visible $m_{\tau\tau}$

Consider $h \to \tau^+\tau^- \to \pi^+\bar{\nu}\pi^-\nu$

- $|\mathbf{p}_\pi| < |\mathbf{p}_\tau|$
- $m_h^2 = (p_{\tau^+} + p_{\tau^-})^2 \geq (p_{\pi^+} + p_{\pi^-})^2 = m_{\pi\pi}^2$
- m_h given by endpoint of $\mathrm{d}\Gamma/\mathrm{d}m_{\pi\pi}$!

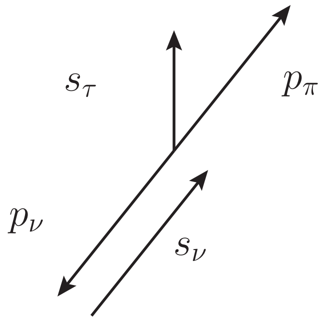
Simple τ decay

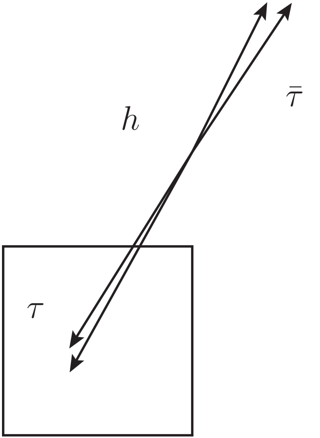
Assume massless π, ν:



In τ rest frame,
$|\mathbf{p}_\pi| = m_\tau/2$

In h rest frame, $\mathbf{p}_\pi$ very parallel to $\mathbf{p}_\tau$





Assume massless π, ν:

$p_\pi = (m_\tau/2)(\gamma(1+\beta z), \gamma(\beta + z), ..., ...)$

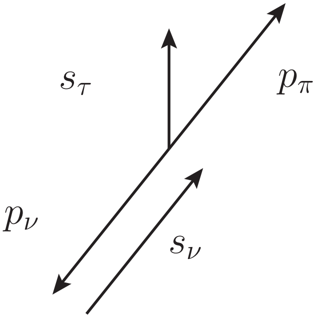
$=$

$((1+z)m_h/4)(1, 1, \mathcal{O}(m_\tau/m_h), \mathcal{O}(m_\tau/m_h))$

for $m_h \gg m_\tau$ $[\gamma = m_h/(2m_\tau), \beta \to 1]$

z is cosine of angle of $\mathbf{p}_\pi$ to $\mathbf{p}_\tau$ in τ rest frame.

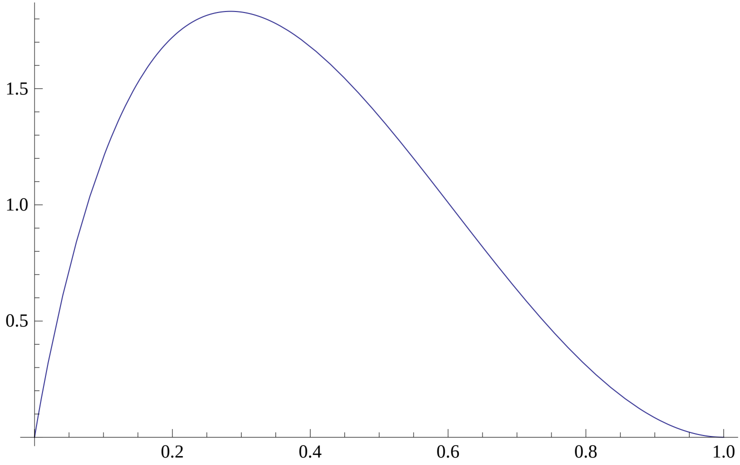
$m_{\pi\pi}^2 = (m_h^2/4)(1+z^+)(1+z^-)$



Left-handed $\tau^\pm \to$ probability of $z^\pm = (1 \mp z^\pm)/2$
Right-handed $\tau^\pm \to$ probability of $z^\pm = (1 \pm z^\pm)/2$
Spin-0 $h \to$ both τ^+, τ^- same helicity.

$\mathrm{d}\Gamma(m_{\pi\pi}^2 = rm_h^2) =$
$N\int_{-1}^{+1}\mathrm{d}z^- \int_{-1}^{+1}\mathrm{d}z^+$
$\delta((1 + z^+)(1 + z^-) - (4/r))$
$(1 - z^-)(1 + z^+)/4 +$ opposite helicities
$\mathrm{d}\Gamma(m_{\pi\pi}^2 =$
$2Nr\int_{-1}^{+1}\mathrm{d}z \left(1 - \left(\frac{4r}{1+z} - 1\right)\right) \Theta \left(1 - \left(\frac{4r}{1+z} - 1\right)\right)$
[remembering $\delta(f(x)) = \delta(x - x^{\mathrm{sol}})/f'(x)$
and $\int_a^b \mathrm{d}x\delta(x - c) = \Theta(c - a)\Theta(b - c)$]

$$\mathrm{d}\Gamma(m_{\pi\pi}^2)/\mathrm{d}r = (4/9)r(1 - r + r\ln r)$$



- That was a lot of work!
- Needed collinear approximation.
- In general, invisible particles make analytic work impossible.

Complication 4

Complication 4: bound states and showering

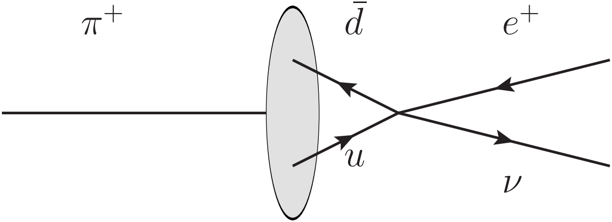
a.k.a.: QCD hates you!

Complication 4a

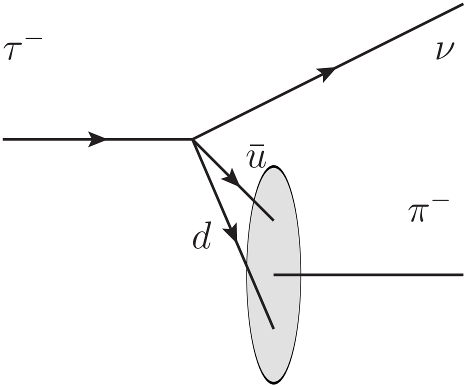
Complication 4a: bound states

- Field operator on bra or ket:
  $e^- | e^-(p) \rangle \to u(p)$
- Free colored particles are not valid final states.
- What field operator do we match to an initial-state proton?

Sometimes we can work with mesons!



- $\bar{d}\gamma^\mu\gamma^5 u|\pi^+(p)\rangle \to iF_\pi p^\mu$
- Measure F_π from $\pi^+ \to e^+\nu$: predict $\pi^+ \to \mu^+\nu$!

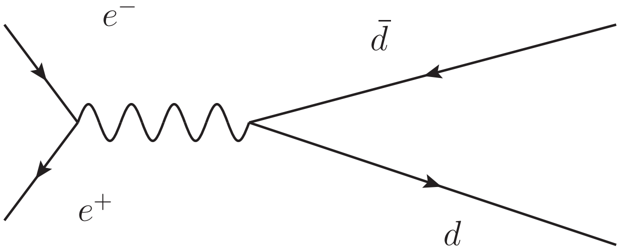


- Predict $\tau^- \to \pi^- \nu$!

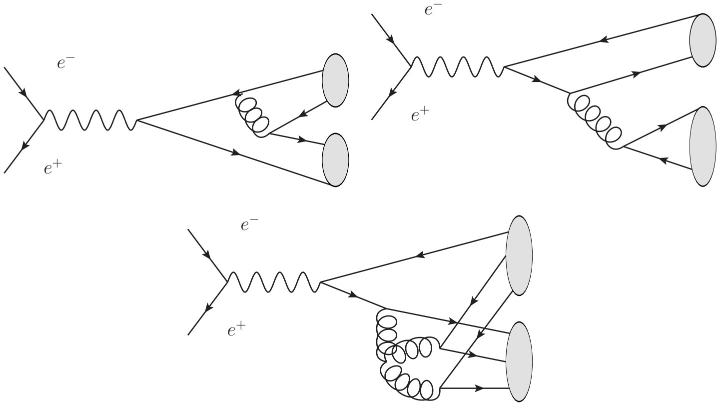
- ... But that is about it, apart from variations on meson decay constants.
- Protons are complicated...

- $q|p^{+}(p)\rangle \to f(x) \times u(x \times p)$ (similar for gluons).
- $f(x)$ is parton distribution function (PDF).
- PDF measured from deep inelastic scattering.
- Color field not balanced, beam remnant still there!
- Center of momentum of partons unknown! Only *transverse momentum* can be balanced.

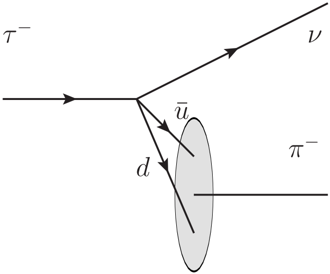
Colored objects do not exist as free final states.



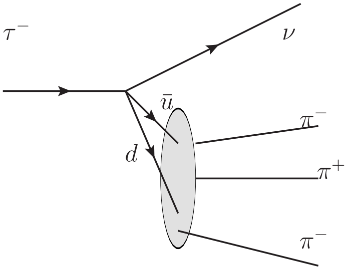
Hadronization: color-singlet objects need to form



Hadronization is not even dominated by minimal amount of bound states!



$\mathrm{BR}(\tau^- \to \pi^- \nu) = 10.8\%$



$\mathrm{BR}(\tau^- \to \pi^- \pi^+ \pi^- \nu) = 9.0\%$

Huge number of final-state particles + hadronization not being easy $\Rightarrow$ MC.

Complication 4b

Complication 4b: showering

- Free colored particles are still not valid final states.
- Timescales for hadronization $\sim 1/\Lambda_{\mathrm{QCD}} \sim 10^{-23}s \gg 1/TeV$.
- Color field has to react to violently-accelerated color-charged particles.
- High-energy q, g radiate energy and color until hadrons form.
- Such *parton showering* is well-suited to MC.

Complication 4c

Complication 4c: jets

- Most parton showering is very collinear to radiating parton.
- Disentangling individual hadrons from spray hitting detector is generally impossible.
- Resolving individual hadrons is not usually particularly desired anyway – reconstructing initiating parton usually more important.
- Burst of hadrons clustered together hitting detector is referred to as a *jet*.

- Unfortunately it's almost impossible to tell if a jet was initiated by g, u, d, s, c.
- *Sometimes* bottom quarks can be identified (b-*tagging*: about 60%)
- Also distinguishing 2 close jets from 1 wide jet more of an art.
- There are several algorithms (Cambridge-Aachen, k_T, anti-k_T, ...), all with advantages; however, probably best to stick to default of whichever software you use unless you have a good reason to change it.

- Parton showering cannot be ignored!
- Many searches at the LHC involve vetoing events with too few hard jets.
- Hard process might only produce 2 jets, but veto requires 3 or 4.
  - *E.g.* $pp \to \tilde{q}\tilde{q} \to j\tilde{\chi}j\tilde{\chi}$ passes 4-jet cut up to 50% of the time for very heavy squarks!
- In principle, one could sum up Feynman diagrams which include parton radiation, but that is impractical.

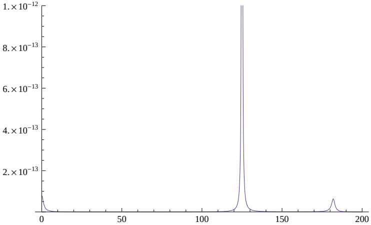
Complication 4d

Complication 4d: the rest of the protons!

- $gg \to h \to \gamma\gamma$ – but $g \neq p$!
- Rest of protons (*beam remnant*) usually splatters down beam pipe, but...
- ... another pair of gluons might interact! (*Multiple interactions.*)
- ... the beam remnant might radiate a bit into the forward part of the detector.
- Luckily all this is taken care of in `Pythia`, `Herwig`, *etc.*!

Intermediate particles on-shell: the narrow width approximation and cascade decays

- In principle we need to integrate the full $|\mathcal{M}|^2$ over the n-body phase space.
- "Off-shellness" of propagators suppresses momentum configurations.

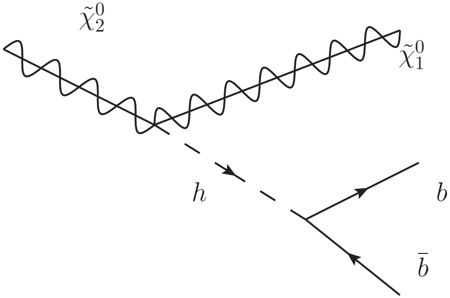


Phase space is very sparse and "spikey".

The narrow width approximation

$|\text{propagator}|^2 = |(p^2 - m^2 + i\Gamma m)^{-1}|^2 = ((p^2 - m^2)^2 + \Gamma^2 m^2)^{-1}$
$\int_{-\infty}^{\infty} |\text{propagator}|^2 \mathrm{d}[p^2] = \pi/(\Gamma m)$

$\Rightarrow$ replace $|\text{propagator}|^2$ with $\pi\delta(p^2 - m^2)/(\Gamma m)$
Pretty good if $\Gamma \ll m$ – the *narrow width approximation* (NWA).



h $|\text{propagator}|^2 \to \pi\delta(p_h^2 - m_h^2)/(\Gamma_h m_h)$:
$|\text{this diagram}|^2 \Rightarrow \Gamma(\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 h) \times (\Gamma(h \to b\bar{b})/\Gamma_h)$
$= \Gamma(\tilde{\chi}_2^0 \to \tilde{\chi}_1^0 h) \times \text{BR}(h \to b\bar{b})$

- NWA very useful!
- Can build long *cascade decay chains*
- NWA can preserve spin correlations: *e.g.* $\eta_{\mu\nu}\eta^{\rho\sigma}((p^2 - m^2)^2 + \Gamma^2 m^2)^{-1} \to \pi\eta_{\mu\nu}\eta^{\rho\sigma}\delta(p^2 - m^2)/(\Gamma m)$
- Remember that it's only valid for *narrow* decay widths!
- Often a trade-off of accuracy for speed.

- One consequence of NWA: can normalize cross-sections to better calculations.
- Most MC is at LO, but many production σs are known to NLO or better.
- Can scale complicated signal by ratio of NLO to LO for $2 \to 2$ process: the K factor.
- K factor is only an approximation, but works out well in practice normally.
- *e.g.* can use `Prospino` to get K factors for processes beginning with gluino pair production, squark pair production, *etc.*; `LHC-FASER` has tabulated colored sparticle production over the masses relevant to the LHC. Can work out full NLO with `FormCalc` if practical.

Accuracy

- You don't want to compare theory predictions with large statistical fluctuations to experiment.
- Ideally you simulate ab^{-1} of data, and scale to the luminosity reported in the experiment's analysis.
- If you want about 1% accuracy, you need to simulate roughly $(1\%)^{-2} = 10^4$ events.

How to prepare to perform
a collider study!

Preparation plan

- I hope that now you are motivated to use Monte Carlo methods to compare theory to experiment.
- I hope that you are aware of various issues caused by QCD:
  - Quarks and gluons are not directly detectable: only jets. Quark flavor is not observable in general.
  - Jets radiate more jets: a multijet veto may trigger on hard processes with only a few final-state partons.
  - At the LHC, only transverse momentum can be balanced.
- I hope that you understand the NWA and know when it is appropriate.

Simulation plan

- Computers are not infinitely fast: you will need to be selective.
- Most generators give you the option of choosing what processes to simulate: if you have good reason to believe a process is not relevant, don't bother simulating it (*e.g.* production of particles with masses of 5 TeV.)

Software setup

- Now you can decide which MC codes to install.
- I won't advertize any generator in particular: there will be enough of that soon enough. Choose the generator(s) most suited to your needs.
- You will need to decide what processes to simulate to best approximate the model: quicker approximations → more statistics...
- I now assume that you will successfully install and run the software. After coffee we will discuss what to do with the output!

Checklist

- Do you know all the channels leading to the signal?
  - If in doubt, put it in.
- Do you know the normalizations of the channels, and how many events in each channel to simulate to achieve the required accuracy?
- Do you want to make any trade-offs of accuracy for speed?
- Do you know the limitations of your tools?