4 Lectures on

QCD

Alexander Mitov
Cavendish Laboratory, Cambridge, UK

Lecture 3

The 2015 European School of High-Energy Physics
Bansko, Bulgaria, 2 – 15 September 2015
Lecture 3:

- Jets
- Decays of unstable particles
- Monte Carlo Integration
- Fixed Order Calculations with MC methods
- Parton Showers
- Matching of LO calculations with Parton Showers
Final states at hadron colliders (particles and jets)

- The particles observed at colliders are:
  - Strongly interacting ones (mesons, baryons)
  - Weakly interacting ones
    - The gauge bosons ($\gamma, Z, W$)
    - Higgs (which decays of course)
    - Leptons
- The problem is that there are many of them
- Even at high PT the number of identifiable tracks is in the dozens and hundreds.
- It would be very very hard to describe such large multiplicities down to the individual particle.
- Luckily, they tend to clump together along the direction of some hard parton (which we do not see directly) that initiated them.
- Such clumps of particles are called jets.
Jets at hadron colliders: an alternative way of thinking about hadron production

✓ Jets are not “Physical” objects: they are merely clusters of hadrons
✓ No two jets are the same!
✓ But Jets are natural at hadron colliders:
  ✓ Ex: describing the water molecules in Jet d'eau is hard
  ✓ The water jet itself depends on the dynamics among the constituent
  ✓ It is the natural thing to study when the detector is close to the water source.
✓ Have been measured at colliders since the late 1970-ies.

The process of jet formation:

More details: Gavin Salam, TASI lectures on Jets (2013)

Jets and infrared safety

- Jets are defined through some algorithm (or jet function) which tells us how the measured particles are grouped into jets.

- Many such definitions exist. A very popular one nowadays is the anti-KT algorithm of Cacciari, Salam and Soyez ’08.

- If you are interested in this subject, you want to read this paper http://arxiv.org/pdf/0802.1189v2.pdf

Infrared safety means that we work with observables that are not singular when soft/collinear emissions are made (Note: fragmentation functions are the opposite extreme since they are collinearly unsafe; there we collect all collinear singularities and put them into the hadron).

\[
\frac{d\sigma}{dO} = \sum_n \int |M^{(n)}|^2 \delta \left( O - F_O^{(n)}(p_1, \ldots, p_n) \right) d\Phi^{(n)}
\]

The ”Jet” observation function has to have the following property:

- \(F_O^{(n+1)}(p_1, \ldots, p_n, p_{n+1}) = F_O^{(n)}(p_1, \ldots, p_n), \text{ if } p_{n+1} \to 0\)
- \(F_O^{(n+1)}(p_1, \ldots, p_n, p_{n+1}) = F_O^{(n)}(p_1, \ldots, p_n + p_{n+1}), \text{ if } p_n || p_{n+1}\)
Jets: definition

- In a realistic LHC event one has to cluster around $10^3$-$10^4$ particles.
- Few particles are hard, most are soft (soft: i.e. momentum $>$ 0)
- For this we need:
  - A jet algorithm (IR safe one!) that will cluster all these partons into jets
  - Speed (i.e. fast algorithm)
- Some popular jet algorithms:
  - $K_T$
  - Cambridge-Aachen
  - Anti-$K_T$
- How does the clustering work?
- Define a distance function $d_{ij}$ between any two particles or proto-jets (proto-jet: a collection of particles that may not yet be a jet), as well as a distance $d_{iB}$ between each particle and the beam

\[
\begin{align*}
K_T\ \text{algorithm:} & \\
\quad d_{ij} &= \min(p_{ti}^2, p_{tj}^2) \Delta R_{ij}^2 / R^2 \\
\quad d_{iB} &= p_{ti}^2 \\
K_T\ \text{algorithm:} & \\
\quad d_{ij} &= \Delta R_{ij}^2 / R^2 \\
\quad d_{iB} &= 1. \\
\text{anti-}K_T\ \text{algorithm:} & \\
\quad d_{ij} &= \min(1/p_{ti}^2, 1/p_{tj}^2) \Delta R_{ij}^2 / R^2 \\
\quad d_{iB} &= 1/p_{ti}^2.
\end{align*}
\]

And: $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  \hspace{1cm} R: a parameter (cone size); $R=0.5-1$
Jets: definition

- The distance $\Delta R_{ij}$ is boost invariant
- The jet size $R$:
  - If $R$ is large, then the concept of jet loses its meaning because it becomes equivalent to the total cross-section.
  - If $R$ is very small, we have a problem: the jet is formally IR finite the the space where IR cancellation takes place becomes very tight and the IR cancellation becomes imperfect. As a result, in the limit $R \rightarrow 0$ we get terms like $\log[R]$.
  - Indeed, in the limit $R \rightarrow 0$ we must run into trouble since we conceptually go to the fragmentation function case we described previously (which was IR unsafe).

- Definition:
  - Construct the set of all measured momenta $p_i$
  - Compute all distances $d_{iB}$ and $d_{ij}$ (defined on previous page)
  - If $d_{iB}$ is the smallest on the list: call “I” a jet and remove from the list
  - If $d_{ij}$ is the smallest: then add $p_i$ and $p_j$ and replace them in the list with their sum.
  - Continue until all particle (proto-jets) momenta in the list are exhausted.
  - The resulting objects are our jets.
  - At the end, only jets above certain $p_T$ cut are used in the final analysis.

- Speed: for $N$ particles, the number of comparisons is $O(N^3)$. This is big.
- A library called FastJet exists, which reduces the time to $N \log[N]$. Moreover it provides common implementation and interface to many jet algorithms.
The shapes of Jets; addition of soft radiation

• Some jets are better at handling additional soft radiation (i.e. are less sensitive to it)

Figure 1: A sample parton-level event (generated with Herwig [8]), together with many random soft “ghosts”, clustered with four different jets algorithms, illustrating the “active” catchment areas of the resulting hard jets. For $k_t$ and Cam/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified.
Latest developments in Jets
For more info see the proceedings of the annual BOOST conference http://boost2015.uchicago.edu

- Boosted objects as jets.
  - Imagine $W$ decaying to jets ($W \rightarrow qq$). In cases the $W$ itself is very energetic its decay products will appear as a single jet. Same for decaying boosted tops.
  - The real motivation for considering such cases is searches for new physics:
    - Imagine a heavy resonance decaying to pair of tops (typical bSM possibility)
    - Each one of the tops will be highly boosted
    - The top decay products will be collimated.

- Jet substructure
  - A way of distinguishing normal QCD jets (they are not supposed to have any characteristic internal structure) from highly boosted decays is to try to identify the presence of sib-jets in a highly energetic and massive jets.
  - Explosion of interest and literature on this topic in the last few years. Many techniques developed: \textit{N-subjettiness, etc; Jet trimming, Jet filtering, Jet pruning}. For more info see http://arxiv.org/abs/1307.0007, http://arxiv.org/abs/1311.2708

- Fat jets and recent diboson 8TeV excess (ATLAS, CMS) (a search which is optimized towards heavy objects decaying to gauge bosons) See http://arxiv.org/abs/1506.00962

Backgrounds seem OK

Local excess after additional W,Z selection
Decay of unstable particles: narrow width approximation

For $\Gamma \ll m$

\[
\begin{array}{c}
\text{\includegraphics[width=0.2\textwidth]{diagram.png}}
\end{array}
\]

The propagator of an unstable particle (Breit-Wigner resonance) of momentum $q$, mass $m$ and width $\Gamma$ is:

\[
P(q, m, \Gamma) = \frac{1}{(q^2 - m^2)^2 + m^2\Gamma^2}
\]

(this is just the modulus square of the usual propagator for a particle of width $\Gamma$)

In the NWA we take the formal limit:

\[
\frac{1}{(q^2 - m^2)^2 + m^2\Gamma^2} \rightarrow \frac{\pi}{m\Gamma} \delta(q^2 - m^2)
\]

i.e. the decaying particle is placed on-shell. This way the Phase space for the complete $n$-body process factories into the product of the phase space of all particles but the decay products, times the decay of the unstable particle.

The NWA leads to drastic simplification:

\[
\sigma = \sigma_{\text{prod}} \times BR,
\]

\[
BR = \frac{\Gamma_{\text{partial}}}{\Gamma_{\text{tot}}}.\]

where $\Gamma_{\text{tot}}$ is the same as $\Gamma$ above. If there is only one decay mode then $BR = 1$.

Monte Carlo integration methods
Monte Carlo integration methods

• So far we discussed only analytical integration in our discussion of cross-sections.
• This is, of course fine, but as it turns out it is very restrictive given the realities of experimental analyses at colliders.

So, what’s the problem?

• The problem is that analytical integration is, by its very nature, inclusive.

Let’s look at an example:
  • f(x) is some probability density (we imagine it corresponds to some differential distribution)
  • Within the analytical integration approach a question we can ask is: what is the value: \[ F(a, b) = \int_{a}^{b} f(x) \, dx \]

• In effect this is a bin.
• Clearly this is well defined only if a=/= b
• Therefore we cannot ask, or predict, what will be the measured value of F in a single point \( F(a,a) \). In other words, within this approach, we cannot predict single events.

• Yet single events happen all the time at colliders.

• MC comes to the rescue!
Monte Carlo techniques are super useful for two very important reasons:

- Formally, they are an integration technique, i.e. we can use them to do integration numerically. Compare, for example with Gaussian integration.

Example: calculating the value of $\pi$

- Generate $N$ points randomly over the $(x,y)$ square
- Count the points inside the circle ($n$)
- Derive: $\pi/4 \approx n/N$
- A simple counting experiment. Error: $\sim \frac{1}{\sqrt{N}}$

- MC integration has smaller numerical precision in 1D,
- In higher dimensions MC has no competition.

The way the integration is done is by summing up discrete points in the continuous variable(s) being integrated. We interpret such discrete points as collider events.

Therefore, Monte Carlo integration offers the possibility to compute formally continuous distributions by summing up individual, discrete events, while being proper integration technique at the same time!

- A word of caution: although the interpretation of such MC events as the real-life collider events is absolutely tantalizing, one should be careful: this is only an integration technique which carries the inherent uncertainties of the underlying theoretical approximation.
Monte Carlo integration methods

• Note about probability density interpretation:

  Going beyond LO many of the contributions are not positive anymore.

• This is OK, since:
  • The observable is positive definite
  • It is a sum over partonic reactions, each of which is unphysical (their separation is scheme dependent)
  • The LO usually are positive but higher orders can individually be negative.
Fixed order calculations with MC techniques: LO

- A fully differential observable at LO is defined as:

\[
\frac{d\sigma}{dO} = \sum_{ij=g,u,\bar{u},\ldots} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M_{ij}(p_1, p_2 \to q_1, \ldots, q_n) \right|^2 \delta \left( O - F_O(q_1, \ldots, q_n) \right) d\Phi(q_1, \ldots, q_n) dx_1 dx_2
\]

- \( p_1 = x_1 P_1 \); \( p_2 = x_2 P_2 \)

- \( d\Phi(q_1, \ldots, q_n) = (2\pi)^4 \delta^4(p_1 + p_2 - q_1 \cdots - q_n) \frac{d^3 q_1}{(2\pi)^3 2E_1} \cdots \frac{d^3 q_{n-1}}{(2\pi)^3 2E_{n-1}} \)

- \( O \) labels some observable (say \( P_T \)) and \( F_O(\ldots) \) is its analytic representation through the final state momenta

- To integrate analytically, we first separate the pdf’s and the integrations over \( x_{1,2} \)
- Then perform analytically the integrations over the independent 3-momenta keeping \( x_{1,2} \) fixed

\[
\frac{d\sigma}{dO} = \sum_{ij=g,u,\bar{u},\ldots} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M_{ij}(p_1, p_2 \to q_1, \ldots, q_n) \right|^2 \delta \left( O - F_O(q_1, \ldots, q_n) \right) d\Phi(q_1, \ldots, q_n) dx_1 dx_2
\]

- Finally perform the remaining \( x_{1,2} \) integrations numerically.
- If the partonic x-section is simple enough we can get analytic expression and therefore compute the \( O \)-dependence as a smooth curve; no bins needed.
Fixed order calculations with MC techniques: LO

- A fully differential observable at LO is defined as:

\[
\frac{d\sigma}{dO} = \sum_{ij=g,u,\ldots} \int \frac{f_i(x_1)f_j(x_2)}{2s} M_{ij}(p_1,p_2 \rightarrow q_1,\ldots,q_n)^2 \delta\left(O - F_O(q_1,\ldots,q_n)\right) d\Phi(q_1,\ldots,q_n) dx_1 dx_2
\]

- Performing such integrations analytically is often not practical. In complicated cases it is hardly possible.

- MC integration is much simpler! Consider the above integral as a simultaneous integral over all final and initial states variables (i.e. consider $x_{1,2}$ on equal footing with the final state ones)

- Parameterize the final state momenta through independent variables $z_1,\ldots,z_{3(n-1)}$ such that they take values on the unit hypercube $0 \leq z_i \leq 1$:

\[
d\Phi(q_1,\ldots,q_n) = \frac{d^3q_1}{(2\pi)^3 2E_1} \cdots \frac{d^3q_{n-1}}{(2\pi)^3 2E_{n-1}} = J(z_1,\ldots,z_{3(n-1)}) dz_1 \cdots dz_{3(n-1)}
\]

- Note: the $z_i$’s are simply normalized energies and cos(angles)
- The matrix element depends on scalar products: $(p_{1,2}\cdot q_i)$ and $(q_i\cdot q_j)$. Rewrite them through $z_i$. The x-section now reads:

\[
\frac{d\sigma}{dO} = \int I(x_1, x_2, z_1, \ldots, z_{3(n-1)}) \delta\left(O - F_O(z_1,\ldots,z_{3(n-1)})\right) dx_1 dx_2 dz_1 \cdots dz_{3(n-1)}
\]

- Here is how we actually implement the MC integration:
Fixed order calculations with MC techniques: LO

\[
\frac{d\sigma}{dO} = \int w(x_1, x_2, z_1, \ldots, z_{3(n-1)}) \delta(O - F_O(z_1, \ldots, z_{3(n-1)})) \, dx_1 \, dx_2 \, dz_1 \ldots \, dz_{3(n-1)}
\]

- Attempting the above integral with a MC, as it is, is a bad idea:
  - Fix the value of \( O \)
  - Take a random point \((x_{1,2}, z_i)\)
  - \( F_O \) at this point will not be equal to the chosen \( O \)

- For MC integration we need to bin “events”. Therefore we need to replace the delta-function with a binning function:

\[
d\sigma_O = \int w(x_1, x_2, z_1, \ldots, z_{3(n-1)}) B\left(F_O(z_1, \ldots, z_{3(n-1)})\right) \, dx_1 \, dx_2 \, dz_1 \ldots \, dz_{3(n-1)}
\]

- The binning function is a set of theta functions; basically it takes values 1 or 0.

- Binning could be done simultaneously in several variables.
- Or even fully exclusively:
  1. Decide binning for each variable of interest
  2. Generate a point \((x_{1,2}, z_i)\). We call it “event”. Beware it is not exactly a physical event!
  3. Compute the value of \( B(\ldots) \) at this point. It is non-zero for only one bin.
  4. Compute the value of the weight \( w(\ldots) \) at this point. Add it to the bin determined in 3.
  5. Continue the process until sufficiently large number of “events” generated in each bin.
  6. Divide by the number of “events” in each bin (i.e. obtain the average \( w \) in each bin)
Fixed order calculations with MC techniques: LO

\[ d\sigma_O = \int w(x_1, x_2, z_1, \ldots, z_{3(n-1)}) B\left(F_O(z_1, \ldots, z_{3(n-1)})\right) dx_1 dx_2 dz_1 \ldots dz_{3(n-1)} \]

- Binned distributions are very easy to manipulate.
- One could even compute the “events” without regard of any binning!
  - Generate events
  - Save: event information \((x_{1,2,z})\) and weight \(w(...)\) for each event.
  - At a later point analyze and bin the events.

- Such approach allows unprecedented flexibility

- Without having to re-compute the matrix elements \(|M(...)|^2\), one could \textit{a posteriori}, after events are computed, change:
  - Value of the renormalization scale (recall \(|M()|^2 \sim \alpha_s^k(\mu_R)\); so divide by this an multiply by \(\alpha_s^k(\mu_R')\) evaluated a some different scale \(\mu_R' =/\mu_R\).
  - Value of factorization scale*
  - The pdf set*

- In practice computing at LO is fast enough; the above approach is very handy at NLO, and beyond.

* To do this, one has to save separately the contributions from all contributing partonic reactions, not just their sum, as implied by the above equation!
**Fixed order calculations with MC techniques: NLO**

- NLO x-sections can be computed following the LO MC methods described above.
- However, there are dramatic complications that arise at NLO, and which we describe next.

- At NLO we need to sum over all cuts (i.e. all different partonic final states) that contribute to the observable at hand.

- We expect this from our previous discussions of IR safety.

- Dijets as an example:
  
  1. If we want to have exactly 2 jets
     - then extra radiation has to be only Unresolved
  2. If we want at least 2 jets (i.e. could be 2, 3 or more)
     - Then extra radiation could be anything (i.e. Unresolved or Resolved)

- Putting it all together we get:
Fixed order calculations with MC techniques: NLO

- The so-called “Virtual” $2 \to n$ contribution:

$$
\frac{d\sigma(2 \to n; \varepsilon)}{dO} = \sum_{ij} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M_{ij}^{(1\text{Loop})}(2 \to n; \varepsilon) \right|^2 \delta \left( O - F_O(1, \ldots, n) \right) d\Phi(q_1, \ldots, q_n) dx_1 dx_2
$$

- The 1 Loop contribution is divergent; it contains explicit poles in epsilon.

$$
V \equiv \left| M_{ij}(2 \to n+1) \right|^2 = \left| M_{ij}^{(\text{Born})}(2 \to n+1) \right|^2 + \left| M_{ij}^{(1\text{Loop})}(2 \to n; \varepsilon) \right|^2 + \text{NNLO terms}
$$

- Phase space in d-dimensions. Note the phase space integration of this piece is regular but normally has to be performed in d-dim since terms $\sim \varepsilon$ can multiply poles from $V$ which results in finite contributions to the x-section.

- The so-called “Real” $2 \to n+1$ contribution:

$$
\frac{d\sigma(2 \to n+1; \varepsilon)}{dO} = \sum_{ij} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M_{ij}^{(\text{Born})}(2 \to n+1) \right|^2 \delta \left( O - F_O(1, \ldots, n+1) \right) d\Phi(q_1, \ldots, q_{n+1}) dx_1 dx_2
$$

- The amplitude is the Born one, so no poles in eps (still sub-leading terms in eps might have to be retained)

- The phase space is done in d-dimensions:
  - The integrand is finite
  - Upon integration over phase space eps poles are generated
Fixed order calculations with MC techniques: NLO

- There is another source of divergences: collinear divergences.
- They are simpler (of complexity of one-loop less, i.e. Born, but still have to be accounted for)
- Here is what happens, schematically, through NNLO:
  - After Real and Virtual corrections are added together the x-section is still divergent. Through NNLO we have (ρ stands for the relevant kinematic variable):
    \[
    \tilde{\sigma}(\varepsilon, \rho) = \tilde{\sigma}^{(0)}(\varepsilon, \rho) + \alpha_S \tilde{\sigma}^{(1)}(\varepsilon, \rho) + \alpha_S^2 \tilde{\sigma}^{(2)}(\varepsilon, \rho) + \ldots
    \]
  - Subtract collinear singularities for hadron colliders (i.e. factor them into the hadrons) as:
    \[
    \frac{\tilde{\sigma}_{ij}(\varepsilon, \rho)}{\rho} = \sum_{k,l} \left[ \frac{\hat{\sigma}_{kl}(x)}{x} \otimes \Gamma_{ki} \otimes \Gamma_{lj} \right](\rho)
    \]
  - From the above we derive the finite x-section
  - The process-independent collinear counter-terms are:

\[
\Gamma_{ij}(\varepsilon, x) = \delta_{ij} \delta(1 - x) + \alpha_S \Gamma_{ij}^{(1)}(\varepsilon, x) + \alpha_S^2 \Gamma_{ij}^{(2)}(\varepsilon, x),
\]

\[
\Gamma_{ij}^{(1)}(\varepsilon, x) = -\frac{1}{2\pi} \frac{P_{ij}^{(0)}(x)}{\epsilon},
\]

\[
\Gamma_{ij}^{(2)}(\varepsilon, x) = \left(\frac{1}{2\pi}\right)^2 \left\{ \frac{1}{2\epsilon^2} \left[ P_{ik}^{(0)} \otimes P_{kj}^{(0)}(x) + \beta_0 P_{ij}^{(0)}(x) \right] - \frac{1}{2\epsilon} P_{ij}^{(1)}(x) \right\}
\]
Fixed order calculations with MC techniques: NLO

- Adding Real and Virtual corrections is, unfortunately, highly non-trivial.

\[
\frac{d\sigma(2 \to n; \varepsilon)}{dO} = \sum_{ij} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M^{(1)\text{Loop}}_{ij}(2 \to n; \varepsilon) \right|^2 \delta(O - F_O(1, \ldots, n)) d\Phi(q_1, \ldots, q_n) dx_1 dx_2
\]

\[
\frac{d\sigma(2 \to n + 1; \varepsilon)}{dO} = \sum_{ij} \int \frac{f_i(x_1)f_j(x_2)}{2s} \left| M^{(\text{Born})}_{ij}(2 \to n+1) \right|^2 \delta(O - F_O(1, \ldots, n+1)) d\Phi(q_1, \ldots, q_{n+1}) dx_1 dx_2
\]

- Here are the problems:
  - The integration over the Real phase-space has to be done in d-dim.
  - This generates explicit poles and we have to control them analytically.
  - This is against the spirit of MC integrations where everything is done numerically.

- The next complication:
  - We want to have fully differential calculations.
  - This means Real and Virtual poles must cancel in every point (local cancellation).
  - And for any measurement function (recall that the measurement functions for the Real and Virtual corrections are different).
  - Therefore we have to ensure that poles cancel even before the observation functions have been specified.
  - It turns out it is sufficient to know their limiting behavior in soft/collinear limit (recall our discussion of jets)

\[
F_O^{(n+1)}(p_1, \ldots, p_n, p_{n+1}) = F_O^{(n)}(p_1, \ldots, p_n), \quad \text{if} \quad p_{n+1} \to 0
\]

\[
F_O^{(n+1)}(p_1, \ldots, p_n, p_{n+1}) = F_O^{(n)}(p_1, \ldots, p_n + p_{n+1}), \quad \text{if} \quad p_n \parallel p_{n+1}
\]
Fixed order calculations with MC techniques: NLO

• Two methods are available on the market.
  • Slicing method (older; not exact; being developed even to NNLO)
  • Subtraction method (newer; exact; has been extended to NNLO)

• To get a feeling for how the methods work we will simplify them;
  • Ignore the presence of collinear singularities (as the previous slide)
  • Hide the presence of pdf’s etc.
  • will consider only one kinematical variable (called x); think of it as the energy of the additional emitted gluon
  • The function $R(x)$ is finite for $x=0$.

\[
d\sigma_O = \int_0^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n)\]

• Recall that

$$F_O(n + 1; x = 0) = F_O(n) \; ; \; R(0) = V_{\text{pole}}$$

• In the phase-space slicing method we split the integration over x in two:

\[
d\sigma_O = \int_0^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n)
\]

\[
= \int_0^\delta \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \int_\delta^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n)
\]

• Take $\delta$ very small, $\delta<<1$. In the first term we can approximate $R(x)\approx R(0)$
• Set $\varepsilon=0$ in the second term (integration is now finite). Integrate it numerically
Fixed order calculations with MC techniques: NLO

• The slicing method:

\[ d\sigma_O = \int_0^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n) \]

\[ = \int_0^\delta \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \int_\delta^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n) \]

\[ \approx \left( -\frac{1}{\varepsilon} + \log(\delta) \right) V_{\text{pole}} F_O(n) + \int_\delta^1 \frac{dx}{x} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n) \]

\[ \approx \int_\delta^1 \frac{dx}{x} R(x) F_O(n + 1; x) + \log(\delta) V_{\text{pole}} F_O(n) + O(\delta) + V_{\text{fin}} F_O(n) \]

• When \( \delta \to 0 \) the approach becomes exact however the numerical integration becomes unstable.
• One has to show that the error due to finite \( \delta \) is small.

• The behavior for \( x \to 0 \) (ie. In the singular limit) can be predicted with resummation

  • \( Q_T \) resummation technique (Catani, Grazzini)

    • Used for computing all 2-to-2 reactions at NNLO where the Born final state is color singlet (WW, ZZ, WW). Now is being developed also for colorful final states like top-pair.

  • N-subjetettiness. New technique developed in the last year (Boughezal, Focke, Liu, Petriello)

    • NNLO corrections to (W+jet; Higgs+jet)
Fixed order calculations with MC techniques: NLO

- The subtraction method. Use the mathematical identity:

\[ x^{-1+\varepsilon} = \frac{1}{\varepsilon} \delta(x) + \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} \left[ \frac{\ln^n(x)}{x} \right] + \]

- Then rewrite (everything is exact)

\[
d\sigma_O = \int_{0}^{1} \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left( \frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}} \right) F_O(n)
\]

\[
= \int_{0}^{1} \frac{dx}{x} \left( R(x) F_O(n + 1; x) - R(0) F_O(n + 1; 0) \right) + V_{\text{fin}} F_O(n)
\]

\[
= \int_{0}^{1} \frac{dx}{x} \left( R(x) F_O(n + 1; x) - V_{\text{pole}} F_O(n) \right) + V_{\text{fin}} F_O(n)
\]

Vanishes for \(x \to 0\)

- The above integral is now finite
  - In any kinematical point
  - For any observable
  - Therefore, we can easily construct an MC (partonic MC)

- Method has been developed at NLO (basis for the MC@NLO); also at NNLO (Czakon). Used for top pair; Higgs +jet, top decay).
Fixed order calculations with MC techniques: NLO

\[
\begin{align*}
  d\sigma_O &= \int_0^1 \frac{dx}{x^{1+\varepsilon}} R(x) F_O(n + 1; x) + \left(\frac{V_{\text{pole}}}{\varepsilon} + V_{\text{fin}}\right) F_O(n) \\
  &= \int_0^1 \frac{dx}{x} (R(x) F_O(n + 1; x) - R(0) F_O(n + 1; 0)) + V_{\text{fin}} F_O(n) \\
  &= \int_0^1 \frac{dx}{x} \left( R(x) F_O(n + 1; x) - V_{\text{pole}} F_O(n) \right) + V_{\text{fin}} F_O(n)
\end{align*}
\]

- Although the method is exact and implementable in MC, it brings addt’l complications at NLO

- Finiteness is achieved through combination of events and counter-events; events have positive weight; counter-events – negative weight.
- Separately, they can be arbitrary large but added together they are finite.

- Events and counter-events are strongly correlated.
  - Generate event and its counter-event at the same time

- Two separate calculations are now needed (they are individually finite):
  1. Subtracted real contributions
  2. Finite Virtual term

- Can be extended to NNLO. Same idea – many more contributions
Parton showers and event generators (Leading Order)

Matching to LO calculations
What is a shower?

- Hard emissions (at some scale $Q$) are described well in fixed order perturbation theory.
- The probability for such emissions is suppressed by powers of $\alpha_s(Q) \ll 1$.
- Soft emissions (at some soft scale $S$) are ubiquitous because $\alpha_s(S) \approx 1$.
- In particular there can be many such emissions.

- Example: a typical event for $Z \rightarrow$ hadrons; down to $\sim$ GeV scales in average around 7 gluons are emitted.

- Here is how a typical hard collision event is developing:
  - After the collision hard radiation is possible; few very hard particles are produced. This part is well described by fixed order perturbation theory.
  - The produced partons are off-shell and can still radiate. Typically these are soft emissions (real and/or virtual).
  - As discussed previously such emissions are cheap and can be copious.
  - This stage is described by a “parton shower”, i.e. a calculator that simulates soft and/or collinear emissions. Due to their universality and factorizability, such calculations are much easier than full FO calculations.
  - Once the system is at very low scales $O(\text{GeV})$, perturbation theory completely breaks down. We enter the hadronization stage. Hadronization can be modeled “exclusively”.
  - Programs that do all steps above are called event generators (like HERWIG, PYTHIA).
Main hadronization models

Cluster

String (like string theory)

Credit: Ellis, Stirling, Webber
Main hadronization models

Cluster hadronization model

- Used in HERWIG
- Assumed that color singlet cluster are formed from neighboring q-qbar pairs
- These color cluster then decay into hadrons
- The mass of the clusters is few GeV
- How they decay to hadrons is model dependent. But a simple phase space-based model already works well.

- The model does not work very well with very massive cluster
- Problems with Baryons and heavy quarks.

String hadronization model

- Used in PYTHIA
- The color string formed from a quark pair breaks down into hadrons
- String is consistent with linear confining potential

- A – area
- Kinks – from gluon emissions
Why we need LO+PS matching?

- Multijet events are omnipresent at the LHC. QCD produces many of those; bSM too. To find bSM we need good understanding of the genuine QCD backgrounds.

![Background simulations](Image)

![Data vs. background](Image)

<table>
<thead>
<tr>
<th>Selection</th>
<th>( Z \rightarrow \nu\bar{\nu} )</th>
<th>( t\bar{t}/W \rightarrow e, \mu +X )</th>
<th>( t\bar{t}/W \rightarrow \tau_l +X )</th>
<th>QCD</th>
<th>Total background</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_{\text{jets}} )</td>
<td>( H_T ) [GeV]</td>
<td>( H_T ) [GeV]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5</td>
<td>500–800</td>
<td>200–300</td>
<td>1820±390</td>
<td>2210±450</td>
<td>1750±210</td>
<td>310±220</td>
</tr>
<tr>
<td>3–5</td>
<td>500–800</td>
<td>300–450</td>
<td>990±220</td>
<td>660±130</td>
<td>590±70</td>
<td>40±20</td>
</tr>
<tr>
<td>3–5</td>
<td>500–800</td>
<td>450–600</td>
<td>273±63</td>
<td>77±17</td>
<td>66.3±9.5</td>
<td>1.3±1.5</td>
</tr>
<tr>
<td>3–5</td>
<td>500–800</td>
<td>&gt;600</td>
<td>42±10</td>
<td>9.5±4.0</td>
<td>5.7±1.3</td>
<td>0.1±0.3</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥8</td>
<td>500–800</td>
<td>&gt;200</td>
<td>0.0±0.8</td>
<td>1.9±1.5</td>
<td>2.8±1.4</td>
<td>0.1±0.4</td>
</tr>
<tr>
<td>≥8</td>
<td>800–1000</td>
<td>&gt;200</td>
<td>0.6±0.6</td>
<td>4.8±2.9</td>
<td>2.3±1.2</td>
<td>0.5±0.9</td>
</tr>
<tr>
<td>≥8</td>
<td>1000–1250</td>
<td>&gt;200</td>
<td>0.6±0.5</td>
<td>1.4±1.5</td>
<td>2.9±1.3</td>
<td>0.7±0.7</td>
</tr>
<tr>
<td>≥8</td>
<td>1250–1500</td>
<td>&gt;200</td>
<td>0.0±0.9</td>
<td>5.1±3.5</td>
<td>1.4±0.9</td>
<td>0.5±0.9</td>
</tr>
<tr>
<td>≥8</td>
<td>&gt;1500</td>
<td>&gt;200</td>
<td>0.0±0.7</td>
<td>0.0±4.2</td>
<td>2.4±1.4</td>
<td>0.9±1.3</td>
</tr>
</tbody>
</table>

- Notice the limit of the simulation
- Notice the large number of jets that are actually measured. And this is for 8 TeV. The LHC now operates at 13 TeV!
Why we need LO+PS matching?

• The genuine shower programs cannot predict such events with any reasonable accuracy.
  
  • Pythia, for example, has only 2 → 1 and 2 → 2 processes genuinely built in. To generate many hard jets with a shower, one has to use the soft and collinear radiation of the shower well outside its intended “comfort” zone.
  
  • A warning – this can be achieved by playing with the scales – but would this be correct? And a more general warning: programs can produce any number. It is up to the user to make sense of produced results. The logic of “an imperfect number is better than no number” could be very useful but also very dangerous. One has to be very careful there!

• Large number of hard emissions are naturally described in fixed order perturbation theory. But these are single, colorful (typically massless) on-shell partons that look nothing like the jets we measure.

• Clearly, we need a combination of:
  
  • hard emissions generated by a complete fixed order calculation (these will give the proto-jets)
  
  • parton shower (builds the highly complex internal structure of the jets).

• Combining fixed order calculations with parton showers is a non-trivial task which is by now solved in many ways at LO. Doing this at NLO is still a very advanced problem. At NNLO this hasn’t even been seriously contemplated (not yet – but may not be far into the future!)
Before we match LO with PS: what’s a merging scheme?

• First, recall the distinction between inclusive and exclusive observable
  • Example: inclusive jets: typically, within QCD, this means two or more jets.
  • At LO however, we only include 2-to-2 diagrams. This way no final states with 3 or more particles (jets). At the same time a state with, e.g. 8 jets, also belongs to this inclusive observable.
  • Question is: how to account for such multi-jet events?
  • Exclusive event: one with a fixed number of final state particles (jets). For example:
    • Exactly 2
    • Exactly 3
    • ...
  • Note: at LO (and only at LO) the various final states are mutually exclusive, i.e. an inclusive sample is just a sum of exclusive ones. This absolutely doesn’t work at NLO and beyond!

• Thus, we arrive at the basic idea of merging samples at LO:
  • Introduce a separation measure between final states with n and n+1 partons.
  • Generate samples for all process with n final states, n<N_{max}. N_{max} \sim O(10) – see previous slide.
  • Add the samples. They are non-overlapping by construction, i.e. any double counting is avoided.
  • A question: this seems an easy thing to do. But then why do we need NLO calculations?
The difference between a merged LO sample and an NLO (or N^kLO) calculation

- Lets take as an example Higgs production:
  - Inclusive Higgs at LO: \( pp \rightarrow H \)
  - Inclusive Higgs at NLO: \( pp \rightarrow h; h+j \)
  - Inclusive Higgs at NNLO: \( pp \rightarrow h; h+j; h+jj \)

- A merged LO sample with \( N_{\text{max}} = 2 \) would cover all of the above final states.
- But not in the full kinematic range!

- For example, in the merged LO sample we are not allowed to make any two final state partons too close to each other. In fact, the result would diverge if we attempted to do that!

- Thus, the LO merged sample depends on the parameter that separates the different multiplicities.

- In contrast, in an NLO calculation one can take the extra emitted final state parton and make it as close as desired to any other parton. The divergence is compensated by the divergence in the loop virtual corrections that are absent in the merged sample!

- Similarly at NNLO: there one or two partons can become very close to any other parton. The divergences are much worse than at NLO but this is again compensated by the (now even more complicated) loop corrections.
A note of caution: the terms merging and matching are not always assigned the same meaning in the literature. Keep an open mind and all should eventually be clear from the context.

• There is no one “best” or unique way of doing this: the final result always contains ambiguities and dependence on unphysical scales as long as we work to finite orders in perturbation theory

• The main requirements for a good matching scheme are:
  • Avoid double counting (all emissions look the same: be they hard, or from the shower)
  • Avoid dead regions (i.e. kinematical regions unpopulated by radiation while they should be)
  • One scheme is better than another one if it is a better approximation (in the sense that both LO and NLO are imperfect, but NLO is clearly better than LO).
Common strategies for PS matching procedures

Follow the comparative study http://arxiv.org/pdf/0706.2569v2.pdf

1. A jet measure is defined and all relevant cross sections including jets are calculated for the process under consideration. I.e. for the production of a final state $X$ in pp-collisions, the cross sections for the processes $pp \rightarrow X + n$ jets with $n = 0, 1, \ldots, N = N_{\text{max}}$ are evaluated.

2. Hard parton samples are produced with a probability proportional to the respective total cross section, in a corresponding kinematic configuration following the matrix element.

3. The individual configurations are accepted or rejected with a dynamical, kinematics-dependent probability that includes both effects of running coupling constants and of Sudakov form factors. In case the event is rejected, step 2 is repeated, i.e. a new parton sample is selected, possibly with a new number of jets.

4. The parton shower is invoked with suitable initial conditions for each of the legs. In some cases, like, e.g. in the MLM procedure, this step is performed together with the step before, i.e. the acceptance/rejection of the jet configuration. In all cases the parton shower is constrained not to produce any extra jet; stated in other words: configurations that would fall into the realm of matrix elements with a higher jet multiplicity are vetoed in the parton shower step.

The matching procedures discussed below differ mainly in:

- the jet definition used in the matrix elements;
- how acceptance/rejection of jet configurations from the matrix element is performed;
- Details of, and the jet vetoing inside, the parton showering.
Restate the problem

• Let’s make it even more evident where’s the problem

  • We need a separation parameter at parton level $R_{\text{part}}$, i.e. any two partons must have $R > R_{\text{part}}$. It is needed, because if $R_{\text{part}} \to 0$ then the partonic $x$-section is IR divergent.
  • We also need a jet-level separation parameter (connected to jet definition, etc) $R_{\text{jet}}$ which separates jets from each other.

• Clearly, only $R_{\text{jet}}$ is physical because it is related to the measurement; not $R_{\text{part}}$.

• Yet, it is easy to see that an unmatched sample has strong dependence on the value of $R_{\text{part}}$: by taking smaller and smaller values for $R_{\text{part}}$ the $x$-section grows unbounded.

• Basically our prediction strongly depends on an unphysical parameter. This is a problem.

• $R_{\text{part}}$ should be smaller than $R_{\text{jet}}$ (because otherwise we will have unpopulated regions – or dead zones) which is undesirable.

• The goal of the FO+PS matching procedure is to minimize the dependence on this parton level cut

  Ideally it should be independent of it, but this is never the case.

• How to achieve this is not obvious. There are 3 main proposals.
Main LO + PS matching procedures

- Main algorithms:
  - MLM (Mangano ‘02)
  - CKKW (Catani, Kraus, Kuhn, Webber ‘01)
  - Dipole (Lonnblad’02)

- Their approaches are: if two partons are very close (can happen when $R_{\text{part}}$ is small) we somehow suppress or outright veto such event (a veto is a form of suppression).

- In MLM the parton level generation and shower are done without any intermediate checks. Only the final jets are checked for:
  - mutual separation
  - If each jet can be associated with one hard parton
  - all jets and partons can be paired
  - Any event where the above are not satisfied is vetoed.
  - The $R_{\text{part}}$ sensitivity is reduced because if two partons are very close they will produce jets that are close to each other and this is vetoed.

- In CKKW there are both parton-level and jet-level checks:
  - Associate a Sudakov factor at each vertex. This is an exponential which dampens parton-level events with small separation.
  - PS emissions which are hard (off-jet) are vetoed.
  - Important: t gives a prescription for how to choose the value of the renormalisation scale at each vertex (i.e. for each emission)!
CKKW merging procedure

- Choice of Renormalization and factorization scales (at each vertex)

Fig. courtesy of B. Webber