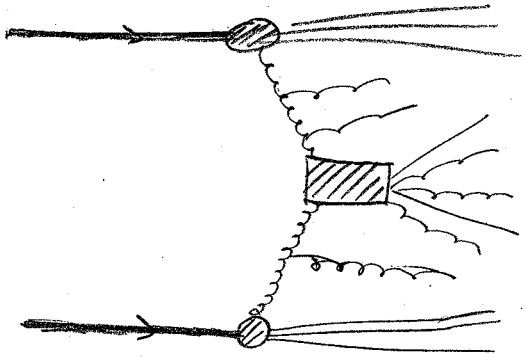
QCD in pp collisions

We shall mostly be concerned by QCD at the Tevatron and the LHC.

The most important difference between them is that at the energy of the LHC and for the type of signals we want to discover, an extremely good knowledge of the background will be required.

From the following picture of a pp event



this means we need to control the following parts

1) the PDF i.e. constrain them + understand their errors
2) the matrix elements : enormous amount of work on calculating multileg & higher-order processes.
3) the parton shower : i.e. • match with matrix elements (avoid double-counting)
   • tune the Monte-Carlo
4) the final-state analysis : defining jets / cats / ... efficiently

Beside this precision requirements, the physics is quite similar for studies at the Tevatron and the LHC as far as QCD is concerned (a difference is e.g. boosted top) Hence, many things in this Chapter apply to both cases. (We will discuss more differences later)

However, when compared to e^+e^- or DIS collisions, numbers of complications come in :

- final-state partons reinteract with the hadronic beams.
  The net effect is to produce extra soft particles known as the UNDERLYING EVENT.

- The initial state is made of QCD partons $\Rightarrow$ the matrix elements are often harder to compute.

* Kinematics

- Hard scattering in pp collisions is typically a collision between two partons carrying fractions x_1 and x_2 of the incoming hadrons.
  Contrarily to the case of e^+e^- collisions considered earlier, since $x_1 \neq x_2$ in general, the LAB frame for this collision is not the Centre-of-Mass frame of the two partons. So, we want boost-invariant variables.
- For this reason, the angle θ is replaced by the RAPIDITY

$$y = \frac{1}{2} \log\left(\frac{E + p_z}{E - p_z}\right)$$

A 4-vector is thus usually defined through

p_t its transverse momentum (boost-invariant)
ϕ its azimuthal angle
y its rapidity (sometimes also m, its mass)

$$p^\mu \equiv (p_t \cos(\phi), p_t \sin(\phi), p_t \sinh(y), p_t \cosh(y)) \quad \text{for the massless case}$$

(or)

$$p^\mu \equiv (p_t \cos(\phi), p_t \sin(\phi), \sqrt{p_t^2 + m^2} \sinh(y), \sqrt{p_t^2 + m^2} \cosh(y)) \quad \text{for the massive case}$$

Rapidity differences are boost invariant

- Other variables
  - pseudo-rapidity $\eta = -\log[\tan(\theta/2)]$ — η differences are boost-invariant.

$$\tan(\theta) = \frac{p_t}{p_z} = \frac{2\tan(\theta/2)}{1 - \tan^2(\theta/2)} \Rightarrow t = \tan(\theta/2),\ t^2 + \frac{2p_z}{p_t} t - 1 = 0$$

$$\Rightarrow t = \frac{-p_z}{p_t} + \sqrt{1 + \frac{p_z^2}{p_t^2}} = \frac{\sqrt{E^2 + m^2} - p_z}{p_t}$$

for a massless particle, $\tan(\theta/2) = \frac{E - p_z}{p_t} = \frac{E - p_z}{\sqrt{E^2 - p_z^2}} = \sqrt{\frac{E - p_z}{E + p_z}} \Rightarrow y = \eta$

  - transverse energy. Since E is not boost invariant, one often uses

$$E_t^2 = E^2 - p_z^2 = p_t^2 + m^2$$

* 2→2 scattering

• kinematics

- invariant mass $(k_1+k_2)^2 = m_1^2 + m_2^2 + 2\,k_1.k_2$

$$= m_1^2 + m_2^2 + 2\,E_{t1}E_{t2}\cosh(\Delta y) - 2\,p_{t1}p_{t2}\cos(\Delta\phi)$$

$\Rightarrow$ for massless particles, $\Pi_{12}^2 = 2\,E_{t1}E_{t2}\left[\cosh(\Delta y) - \cos(\Delta\phi)\right]$

$$= 2\,p_{t1}p_{t2}\left[\cosh(\Delta y) - \cos(\Delta\phi)\right]$$

- incoming partons: $p_1^\mu \equiv x_1\left(0,0,\frac{\sqrt{s}}{2},\frac{\sqrt{s}}{2}\right)$

$p_2^\mu \equiv x_2\left(0,0,-\frac{\sqrt{s}}{2},\frac{\sqrt{s}}{2}\right)$ $\Rightarrow \hat{s} = x_1 x_2 s$

outgoing partons: $k_1^\mu \equiv E_t\left(\cos(\phi),\sin(\phi),\sinh(y_1),\cosh(y_1)\right)$

$k_2^\mu \equiv E_t\left(+\cos(\phi),-\sin(\phi),\sinh(y_2),\cosh(y_2)\right)$

we have

$$E_t\left[\sinh(y_1)+\sinh(y_2)\right] = \frac{\sqrt{s}}{2}(x_1 - x_2)$$

$$E_t\left[\cosh(y_1)+\cosh(y_2)\right] = \frac{\sqrt{s}}{2}(x_1 + x_2)$$

$\Rightarrow$

$$x_1 = \frac{2E_t}{\sqrt{s}}\left(e^{y_1}+e^{y_2}\right)$$

$$x_2 = \frac{E_t}{\sqrt{s}}\left(e^{-y_1}+e^{-y_2}\right)$$

the invariant mass is

$$\hat{s} = \Pi_{12}^2 = 2E_t^2\left[\cosh(\Delta y)+1\right] \geq 4E_t^2$$

which implies, $x_1 x_2 \geq 4E_t^2/s$ (with $0 \leq x_1, x_2 \leq 1$)

Note that we can invert the eq giving $x_{1,2}$ to get

$$y_1 = \log\left[\frac{x_1}{2\varepsilon}\left(1 \pm \sqrt{1-\frac{4\varepsilon^2}{x_1 x_2}}\right)\right] = -\log\left[\frac{x_2}{2\varepsilon}\left(1 \mp \sqrt{1-\frac{4\varepsilon^2}{x_1 x_2}}\right)\right]$$

$$y_2 = \log\left[\frac{x_1}{2\varepsilon}\left(1 \mp \sqrt{1-\frac{4\varepsilon^2}{x_1 x_2}}\right)\right] = -\log\left[\frac{x_2}{2\varepsilon}\left(1 \pm \sqrt{1-\frac{4\varepsilon^2}{x_1 x_2}}\right)\right]$$

with $\varepsilon = E_t/\sqrt{s}$

The smallest accessible x is thus $\frac{4E_t^2}{s}$ (with the other x being 1 and $|y_1| = |y_2| = \log\left(\frac{2E_t}{\sqrt{s}}\right)$)

>0 if $x_2 = 1$
<0 if $x_1 = 1$

↳ maximal rapidity at a given E_t

These relations constrain the actual available phase-space (see Figure)

For the other Mandelstam variables, we have

$$t = (p_1 - k_1)^2 = -2p_1.k_1 = -E_t\sqrt{s}\,e^{-y_1}x_1 = -E_t\sqrt{s}\,e^{+y_2}x_2 = -E_t^2\left(1+e^{y_2-y_1}\right)$$

$$u = -E_t\sqrt{s}\,e^{-y_2}x_2 = -E_t\sqrt{s}\,e^{y_1}x_1 = -E_t^2\left(1+e^{y_1-y_2}\right)$$

One also introduces the rapidity of the "partonic system" $y = \frac{1}{2}\log\left(\frac{x_1}{x_2}\right)$

from which we have

$$\begin{cases} x_1 = \frac{M}{\sqrt{s}}\,e^{y} \\ x_2 = \frac{M}{\sqrt{s}}\,e^{-y} \end{cases}$$

• Drell-Yan: lepton pair production in pp collisions

If M^2 is the invariant mass (squared) of the lepton pair, one has (at lowest order)

$$\frac{d\sigma}{dM^2} = \int dx_1 \int dx_2 \sum_q \left[q(x_1, M^2)\,\bar{q}(x_2, M^2) + (1 \rightleftarrows 2) \right] \frac{d\hat{\sigma}}{dM^2}$$

with

$$\frac{d\hat{\sigma}}{dM^2} = \frac{1}{N_c^2} \cdot (e_q^2 N_c)\, \frac{4\pi\alpha^2}{3M^2}\, \delta(M^2 - x_1 x_2 s)$$

This matrix element is the same as for $e^+e^- \to q\bar{q}$ except for the factor $1/N_c^2$ that accounts for the average over the colour of the incoming quarks.

One also often uses the double-differential cross-section including the dependence on the rapidity of the lepton pair:

$$\frac{d^2\sigma}{dM^2 dy} = \int dx_1\, dx_2 \sum_q \left[q(x_1, M^2)\bar{q}(x_2, M^2) + (1 \rightleftarrows 2) \right] \frac{e_q^2}{N_c} \frac{4\pi\alpha^2}{3M^2}\, \delta(M^2 - x_1 x_2 s)\, \delta\left(y - \tfrac{1}{2}\log(x_1/x_2)\right)$$

$$= \int dx_1 \sum_q \left[q(x_1, M^2)\,\bar{q}\left(\frac{M^2}{s x_1}, M^2\right) + (1 \rightleftarrows 2) \right] \frac{4\pi\alpha^2 e_q^2}{3M^2 N_c} \cdot \frac{1}{x_1 s}\, \delta\left(y + \log(M/\sqrt{s}) - \log(x_1)\right)$$

$$= \frac{4\pi\alpha^2}{3N_c} \frac{1}{M^2 s} \sum_q e_q^2 \left[q\left(\frac{M}{\sqrt{s}} e^{y}, M^2\right) \bar{q}\left(\frac{M}{\sqrt{s}} e^{-y}, M^2\right) + (y \rightleftarrows -y) \right]$$

Alternatively, the cross-section is also sometimes given in bins of M and x_F, the Feynman-x of the lepton pair defined by

$$x_F = \frac{2}{\sqrt{s}} (p_{l^+} + p_{l^-})_z = x_1 - x_2$$

(the last equality only stands at leading order)

One easily finds

$$\frac{d^2\sigma}{dM^2\, dx_F} = \frac{4\pi\alpha^2}{3N_c} \frac{1}{M^2 s} \sum_q e_q^2 \left[q\left(\frac{\sqrt{x_F^2 + \frac{4M^2}{s}} + x_F}{2}, M^2\right) \bar{q}\left(\frac{\sqrt{x_F^2 + \frac{4M^2}{s}} - x_F}{2}\right) + (x_F \rightleftarrows -x_F) \right]$$

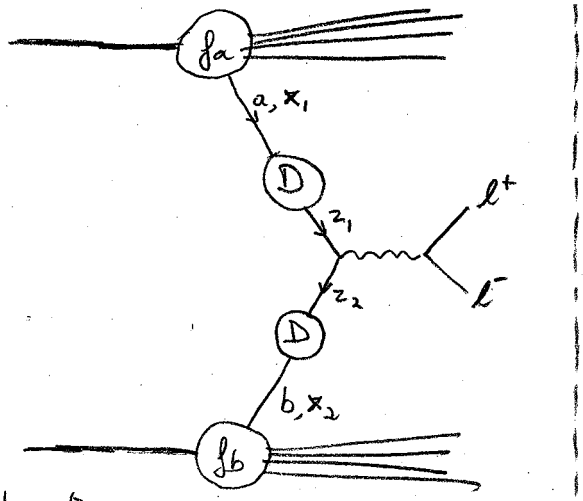
Factorisation has actually been proven for DY at any order of perturbation theory

$$\frac{d\sigma}{dM^2} = \int dx_1\, dx_2 \int dz_1\, dz_2 \sum_{a,b} f_a(x_1, M^2)\, f_b(x_2, M^2)\, D_{ab}(z_1/x_1, z_2/x_2)\, \frac{d\hat{\sigma}}{dM^2}(z_1, z_2; M^2)$$

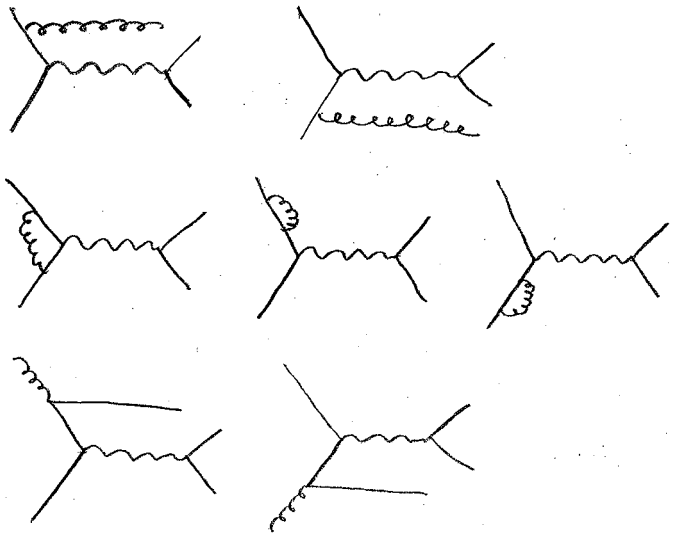
where the sum now goes over all partons ($q, \bar{q}$ & g) and the coefficients D_{ab} are computed order-by-order in perturbation theory.

(See Figure on the next page)

The fact that the PDF entering DY processes are the same as the DIS one is a remarkable property.



Schematic representation of factorisation for Drell Yan processes



matrix elements at NLO

To illustrate that point, let us give the NLO expression : ($z_1 = z x_1$, $z_2 = x_2$ or $z_2 = z x_2$, $z_1 = x_1$)

$$\frac{d\sigma}{d\Pi^2} = \int dx_1\, dx_2\, dz \;\; \frac{4\pi\alpha^2}{N_c \Pi^2}\, \delta(\Pi^2 - x_1 x_2 z s)$$

$$\left(\sum_q e_q^2 \left[q(x_1,\Pi^2)\,\bar{q}(x_2,\Pi^2) + (1 \rightleftarrows 2) \right] D_q(z) \right.$$

$$\left. + \sum_q e_q^2 \left\{ \left[q(x_1,\Pi^2) + \bar{q}(x_1,\Pi^2) \right] g(x_2,\Pi^2) + (1 \rightleftarrows 2) \right\} D_g(z) \right)$$

with

$$D_q(z) = \delta(1-z) + \frac{\alpha_s}{2\pi} D_q^{(1)}(z)$$

$$D_g(z) = \frac{\alpha_s}{2\pi} D_g^{(1)}(z)$$

* PDF for the LHC

Perturbatively, we usually ASSUME factorisation

$$\sigma = \sum_{ab} \int dx_1\, dx_2\; f_a(x_1, \mu_F^2)\; f_b(x_2, \mu_F^2)\; \sigma_0^{ab}$$

(it is only proven for a limited amount of specific observables like Drell-Yan).
This explicitly show that we need both the PDF and the matrix elements.

Let us first consider the PDF part. There are different methods to extract the errors on the PDF ($\Delta\chi^2$, Hessian, ...) but we won't discuss this part here. These errors come from different sources:

- error on the data
- choice of initial condition (μ_F^2, α_s^2, x-dependence)
- treatment of heavy quarks
- ...

The errors on PDF only account for the first part of these (see Figures)
Some of the errors can easily be traced. Small-x gluons, for example, only enter in the Q^2 slope of F_2, and since the available Q^2 range at small x is small, the error gets large.
Conversely, at large x, gluons are much smaller than quarks, and mostly constrained by large-p_t jets at the Tevatron, so their error is bigger too.

However, when it comes to computation of the errors on specific observables, it sometimes appears that various PDF are incompatible, even taking their respective errors into account.
A typical example is the prediction for the W/Z cross-section (a candidate for luminosity measurements) where the error on PDF is the dominant one.

The W/Z cross-section also shows a difference between the Tevatron and the LHC.
Because of the larger energy, the LHC will probe smaller-x partons, which in turn means more gluons than quarks. While dominated by qquarks at Tevatron, W/Z production at the LHC will be dominated by gluons.

• Jets at lowest order

At lowest order, the jet cross-section is given by $2\to 2$ matrix elements. Since the 2 outgoing partons will be back-to-back, we will have exactly 2 jets (regardless of the algorithm).

The 2-particle phase-space is (in the partonic system)

$$\int d\hat{\Phi}_2 = \frac{d^4k_1}{(2\pi)^4}\frac{d^4k_2}{(2\pi)^4}(2\pi)\delta(k_1^2)(2\pi)\delta(k_2^2)(2\pi)^4\delta^{(4)}(k_1+k_2-\text{“}\hat{s}\text{”})$$

$$= \frac{1}{16\pi^2}\int k_{t_1}\,dk_{t_1}\,dy_1\,d\phi_1\,k_{t_2}\,dk_{t_2}\,dy_2\,d\phi_2\;\delta\big(k_{t_1}\cos(\phi_1)+k_{t_2}\cos(\phi_2)\big)$$

$\int dk_{t_2}\int d\phi_2$, $k_{t_1}\to k_t$:

$$\delta\big(k_{t_1}\sin(\phi_1)+k_{t_2}\sin(\phi_2)\big)\,\delta\big(k_{t_1}\sinh(y_1)+k_{t_2}\sinh(\phi_2)\big)\,\delta\big(k_{t_1}\cosh(y_1)+k_{t_2}\cosh(y_2)-\sqrt{\hat{s}}\big)$$

$$= \frac{1}{16\pi^2}\int k_t\,dk_t\,d\phi_1\,dy_1\,dy_2\,\delta\big(k_t\sinh(y_1)+k_t\sinh(y_2)\big)\,\delta\big(k_t\cosh(y_1)+k_t\cosh(y_2)-\sqrt{\hat{s}}\big)$$

$\int d\phi_1\int dy_2$, $y_1\to\hat{y}$:

$$= \frac{1}{16\pi}\int\frac{dk_t}{k_t}\frac{dy}{|\sinh(y)|}\,\delta\Big(\cosh(\hat{y})-\frac{\sqrt{\hat{s}}}{2k_t}\Big)$$

$$= \frac{1}{16\pi}\int\frac{dk_t}{k_t}\frac{d\hat{y}}{|\cosh(y)\sinh(y)|}\Big[\delta\Big(\hat{y}-\operatorname{arccosh}\Big(\frac{\sqrt{\hat{s}}}{2k_t}\Big)\Big)+\delta\Big(\hat{y}+\operatorname{arccosh}\Big(\frac{\sqrt{\hat{s}}}{2k_t}\Big)\Big)\Big]$$

$$= \frac{1}{4\pi}\int\frac{k_t\,dk_t\,d\hat{y}}{\sqrt{\hat{s}}\sqrt{\hat{s}-4k_t^2}}\Big[\delta\Big(\hat{y}-\operatorname{arccosh}\Big(\sqrt{\hat{s}}/2k_t\Big)\Big)+\delta\Big(\hat{y}+\operatorname{arccosh}\Big(\sqrt{\hat{s}}/2k_t\Big)\Big)\Big]$$

Note: $\operatorname{arccosh}(x) = \log\big(x\pm\sqrt{x^2-1}\big)$ $\Rightarrow$

$$\hat{y} = \log\Big(\frac{\sqrt{\hat{s}}}{2k_t}\pm\sqrt{\frac{\hat{s}}{4k_t^2}-1}\Big)$$

$$= \log\Big[\frac{\sqrt{x_1x_2s}}{2k_t}\Big(1\pm\sqrt{1-\frac{4k_t^2}{x_1x_2s}}\Big)\Big]$$

$$= \frac{1}{2}\log\Big(\frac{x_2}{x_1}\Big)+\log\Big[\frac{x_1\sqrt{s}}{2k_t}\Big(1\pm\sqrt{1-\frac{4k_t^2}{x_1x_2s}}\Big)\Big]$$

$$\Rightarrow y_{1,2} = y\pm\hat{y}$$

- at a given k_t ($\le\sqrt{s}/2$), the phase-space for (x_1,x_2) is constrained by $x_1x_2\ge\frac{4k_t^2}{s}$

The jet cross-section will depend on a combination of all possible $2\to 2$ scattering processes. At LO, we have

$$\frac{d^2\sigma}{dk_t\,dy} = \int dx_1\,dx_2 \left\{ \delta\left(y - \left[\tfrac{1}{2}\log(x_1/x_2) + \hat{y}\right]\right) + \delta\left(y - \left[\tfrac{1}{2}\log(x_1/x_2) - \hat{y}\right]\right)\right\}$$

$$\sum_{q\neq q'} \left[q(x_1)+\bar{q}(x_1)\right]\left[q'(x_2)+\bar{q}'(x_2)\right]\left(\frac{d\hat{\sigma}}{dk_t}\Big|_{qq'\to qq'} + \frac{d\hat{\sigma}}{dk_t}\Big|_{qq'\to q'q}\right)$$

$$+ \sum_q \left[q(x_1)\,q(x_2)+\bar{q}(x_1)\,\bar{q}(x_2)\right] \frac{d\hat{\sigma}}{dk_t}\Big|_{qq\to qq}$$

$$+ \left[q(x_1)\,\bar{q}(x_2)+\bar{q}(x_1)\,q(x_2)\right]\left(\frac{d\hat{\sigma}}{dk_t}\Big|_{q\bar{q}\to q'\bar{q}'} + \frac{d\hat{\sigma}}{dk_t}\Big|_{q\bar{q}\to gg} + \frac{d\hat{\sigma}}{dk_t}\Big|_{q\bar{q}\to \bar{q}q}\right)$$

$$+ \left\{\left[q(x_1)+\bar{q}(x_1)\right]g(x_2) + \left[q(x_2)+\bar{q}(x_2)\right]g(x_1)\right\}\left(\frac{d\hat{\sigma}}{dk_t}\Big|_{qg\to qg} + \frac{d\hat{\sigma}}{dk_t}\Big|_{qg\to gq}\right)$$

$$+ g(x_1)\,g(x_2)\left(\frac{d\hat{\sigma}}{dk_t}\Big|_{gg\to q\bar{q}} + \frac{d\hat{\sigma}}{dk_t}\Big|_{gg\to \bar{q}q} + \frac{d\hat{\sigma}}{dk_t}\Big|_{gg\to gg}\right)$$

Notes: - when applicable, sums over final-state flavours is understood

- $\frac{d\hat{\sigma}}{dk_t}\Big|_{ab\to cd}$ are proportional to matrix elements $|\overline{\mathcal{M}}|^2_{ab\to cd}$ up to a factor $\frac{1}{2\hat{s}}\,\frac{1}{16\pi\,\mathrm{ch}(\hat{y}_0)\,\mathrm{sh}|\hat{y}_0|\,k_t}$ $\qquad \hat{y}_0 = \pm\,\mathrm{arccosh}\left(\sqrt{\hat{s}}/2k_t\right)$

$$|\mathcal{M}|^2_{qq\to q'q'} = g^4\,\frac{4}{9}\,\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}$$

$$|\overline{\mathcal{M}}|^2_{q\bar{q}\to q\bar{q}} = g^4\left[\frac{4}{9}\left(\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2} + \frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}\right) - \frac{8}{27}\,\frac{\hat{u}^2}{\hat{s}\hat{t}}\right]$$

$$|\overline{\mathcal{M}}|^2_{qq\to qq} = g^4\left[\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2} - \frac{4}{9}\,\frac{\hat{u}^2+\hat{s}^2}{\hat{s}\hat{u}}\right]$$

$$|\overline{\mathcal{M}}|^2_{gg\to gg} = g^4\,\frac{9}{2}\left(3 - \frac{\hat{u}\hat{t}}{\hat{s}^2} - \frac{\hat{u}\hat{s}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2}\right)$$

with

$$\begin{cases} \hat{s} = x_1x_2 s = 4k_t^2\,\mathrm{ch}^2(y_0) \\ \hat{t} = -k_t^2\left(1+e^{-2y_0}\right) \\ \hat{u} = -k_t^2\left(1+e^{2y_0}\right) \end{cases}$$

and the other matrix elements are easily obtained by crossing symmetry. (Exercise)

- because of the $\hat{t}\rightleftarrows\hat{u}$ $(\equiv \hat{y}_0 \rightleftarrows -\hat{y}_0)$ symmetry, we can assume $\hat{y}_0 > 0$ and add a factor 2.

Since

$$x_{1,2} = \frac{m}{\sqrt{s}} e^{\pm y}$$

with

$$y = \frac{1}{2}\log(x_1/x_2) \qquad \text{and} \qquad m = \sqrt{2k_t^2[\cosh(2\hat{y})+1]} = 2k_t\cosh(\hat{y}) \qquad (\hat{y} = \hat{y}_0)$$

we have

$$dx_1 dx_2 = \frac{4k_t^2}{s} \cdot 2\cosh(\hat{y})\sinh(\hat{y})\, dy\, d\hat{y}$$

Hence

$$\frac{d^2\sigma}{dk_t\, dy} = \int dy\, d\hat{y} \sum_{ab} f_a(x_1) f_b(x_2) |\bar{m}|^2_{ab} \cdot \frac{1}{2x_1x_2s} \frac{1}{16\pi k_t} \frac{8k_t^2}{s} 2\left[\delta(y-(y+\hat{y})) + \delta(y-(y-\hat{y}))\right]$$

Realising that $k_t^2 = \frac{x_1x_2s}{4\cosh(\hat{y})}$, this can be rewritten

$$\boxed{k_t^3 \frac{d^2\sigma}{dk_t\, dy} = \frac{1}{16\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} \frac{d\hat{y}}{\cosh^4(\hat{y})} \delta(y-(y+\hat{y})) \sum_{ab} x_1 f_a(x_1)\, x_2 f_b(x_2) |\bar{m}|^2_{ab}}$$

where $x_{1,2} = \frac{2k_t}{\sqrt{s}} \cosh(\hat{y}) e^{\pm y}$

So far we have not specified the scale appearing in the PDF (and the, potentially different, scale of α_s). One typically set $\mu_F = \mu_R = k_t$ though only a higher-order computation would give something more precise.

Finally, note that, to a good accuracy, one has

$$k_t^3 \frac{d^2\sigma}{dk_t\, dy} \approx \frac{1}{16\pi} \int_{-\infty}^{\infty} dy \frac{d\hat{y}}{\cosh^4(\hat{y})} \delta(y-(y+\hat{y})) \; x_1 F(x_1)\, x_2 F(x_2) \; |\bar{m}|^2_{gg\to gg}$$

with the effective parton density

$$F(x) = g(x) + \frac{4}{9} \sum_q [q(x) + \bar{q}(x)]$$

* Multilegs / NLO Matrix elements

For the level of precision that one needs at the LHC (both for QCD measures and QCD backgrounds), LO computations are usually not sufficient.

There is thus a huge effort to

- compute important processes at NLO
- use them in computer-base programs / MC (MCFM, MC@NLO, MadGraph (tree level), NLOJet, ...)

The aim of the second step is double: first one wants to have the Matrix Elements (ME) in a form that one can use on a computer, maybe even have part of the computation automated. Then we want to use this as a Monte Carlo: either at fixed order, or embedded in a parton shower + hadronisation + UE. The technical point in this last step is to avoid double-counting between the ME and the parton shower.

As far as the computation of the ME itself is concerned, the generic strategy goes as follows:

$$\sigma = \sigma_{Born} + \sigma_{NLO}$$

if the LO (Born) piece has n final-state partons, $\sigma_{Born} = \int_n d\sigma_{Born}$, the NLO contribution will have a real and a virtual part:

$$\sigma_{NLO} = \int_{n+1} d\sigma_{real} + \int_n d\sigma_{virtual}$$

both divergent, but such that the sum is finite. The idea is to introduce an approximation $d\sigma_{approx}$ which has the same singularities and can play the role of a counterterm:

$$\sigma_{NLO} = \int_{n+1} \left(d\sigma_{real} - d\sigma_{approx} \right) + \int_n \left(d\sigma_{virtual} + \int_1 d\sigma_{approx} \right)$$

All integrations can then be performed in 4-dimensional space.
(+ This can be put in a weighted Monte Carlo).

Note: at LO, roughly everything is known / automated / implemented up to a large number of external legs (e.g. MadGraph: up to 9). The technicality here coming from the large number of graphs contributing when the number of legs increases.

In practice, many processes are already available at NLO. Nearly all $2 \to 2$ processes and a long list of $2 \to 3$ processes. Only a few $2 \to 4$ processes are known at NLO, and only a few processes are known at NNLO. This is thus still an important/active field of research.

See the next pages for examples of processes available from MCFM (extract of MCFM-5.3 manual p 12-15)
One typically has (1,2 or 3 W/Z/H, 1 or 2 t) + jets.
Note also that decays involving c/b quarks usually take their non-zero mass into account.

See also the excerpt from the Les Houches 2007 NLO multileg Working Group Report for the recently computed processe and a recent wishlist.

Wishlist : ($V \equiv W|Z|\gamma$) (VBF = Vector Boson Fusion)

$pp \to t\bar{t}\,b\bar{b}$		} for $t\bar{t}H$
$pp \to t\bar{t} + 2j$		
$pp \to VV\,b\bar{b}$		} for VBF $\to H \to VV$
$pp \to VV + 2j$		
$pp \to V + 3j$		new physics
$pp \to b\bar{b}\,b\bar{b}$		H + new physics
$gg \to WW$	at $O(\alpha_{ew}^2 \alpha_s^3)$	H background
$pp \to t\bar{t}$	at NNLO	benchmark (normalisation)
VBF	at NNLO	} benchmark + H couplings
$Z/\gamma + j$	at NNLO	
W/Z	at NNLO QCD + NLO Ew	precision for benchmark

nproc	$f(p_1)+f(p_2)\to\ldots$	Order
1	$W^+(\to \nu(p_3)+e^+(p_4))$	NLO
6	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))$	NLO
11	$W^+(\to \nu(p_3)+e^+(p_4))+f(p_5)$	NLO
12	$W^+(\to \nu(p_3)+e^+(p_4))+\gamma(p_5)$	NLO
13	$W^+(\to \nu(p_3)+e^+(p_4))+\bar{c}(p_5)$	NLO
14	$W^+(\to \nu(p_3)+e^+(p_4))+\bar{c}(p_5)$ [massless]	NLO
16	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+f(p_5)$	NLO
17	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+\gamma(p_5)$	NLO
18	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+c(p_5)$	NLO
19	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+c(p_5)$ [massless]	NLO
20	$W^+(\to \nu(p_3)+e^+(p_4))+b(p_5)+\bar{b}(p_6)$ [massive]	LO
21	$W^+(\to \nu(p_3)+e^+(p_4))+b(p_5)+\bar{b}(p_6)$	NLO
22	$W^+(\to \nu(p_3)+e^+(p_4))+f(p_5)+f(p_6)$	NLO
23	$W^+(\to \nu(p_3)+e^+(p_4))+f(p_5)+f(p_6)+f(p_7)$	LO
24	$W^+(\to \nu(p_3)+e^+(p_4))+b(p_5)+\bar{b}(p_6)+f(p_7)$	LO
25	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+b(p_5)+\bar{b}(p_6)$ [massive]	LO
26	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+b(p_5)+\bar{b}(p_6)$	NLO
27	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+f(p_5)+f(p_6)$	NLO
28	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+f(p_5)+f(p_6)+f(p_7)$	LO
29	$W^-(\to e^-(p_3)+\bar{\nu}(p_4))+b(p_5)+\bar{b}(p_6)+f(p_7)$	LO
31	$Z^0(\to e^-(p_3)+e^+(p_4))$	NLO
32	$Z^0(\to 3\times(\nu(p_3)+\bar{\nu}(p_4)))$	NLO
33	$Z^0(\to b(p_3)+\bar{b}(p_4))$	NLO
41	$Z^0(\to e^-(p_3)+e^+(p_4))+f(p_5)$	NLO
42	$Z_0(\to 3\times(\nu(p_3)+\bar{\nu}(p_4)))-$ [sum over 3 ν] $+f(p_5)$	NLO
43	$Z^0(\to b(p_3)+\bar{b}(p_4))+f(p_5)$	NLO
44	$Z^0(\to e^-(p_3)+e^+(p_4))+f(p_5)+f(p_6)$	NLO
45	$Z^0(\to e^-(p_3)+e^+(p_4))+f(p_5)+f(p_6)+f(p_7)$	LO
48	$Z^0(\to e^-(p_3)+e^+(p_4))+\gamma(p_5)$	NLO
49	$Z^0(\to 3\times(\nu(p_3)+\bar{\nu}(p_4)))-$ [sum over 3 ν] $+\gamma(p_5)$	NLO
50	$Z^0(\to e^-(p_3)+e^+(p_4))+b(p_5)+b(p_6)$ [massive]	LO
51	$Z^0(\to e^-(p_3)+e^+(p_4))+b(p_5)+\bar{b}(p_6)$	NLO
52	$Z_0(\to 3\times(\nu(p_3)+\bar{\nu}(p_4)))+b(p_5)+\bar{b}(p_6)$	NLO
53	$Z^0(\to b(p_3)+\bar{b}(p_4))+b(p_5)+\bar{b}(p_6)$	NLO
54	$Z^0(\to e^-(p_3)+e^+(p_4))+b(p_5)+\bar{b}(p_6)+f(p_7)$	LO
56	$Z^0(\to e^-(p_3)+e^+(p_4))+c(p_5)+\bar{c}(p_6)$	NLO

nproc	$f(p_1) + f(p_2) \to \ldots$	Order
61	$W^+(\to \nu(p_3) + e^+(p_4)) + W^-(\to e^-(p_5) + \bar{\nu}(p_6))$	NLO
62	$W^+(\to \nu(p_3) + e^+(p_4)) + W^-(\to q(p_5) + \bar{q}(p_6))$	NLO
63	$W^+(\to q(p_3) + \bar{q}(p_4)) + W^-(\to e^-(p_5) + \bar{\nu}(p_6))$	NLO
64	$W^+(\to \nu(p_3) + e^+(p_4)) + W^-(\to e^-(p_5) + \bar{\nu}(p_6))$ [no pol]	NLO
66	$W^+(\to q(p_3) + \bar{q}(p_4)) + W^-(\to e^-(p_5) + \bar{\nu}(p_6)) + f(p_7)$	NLO
71	$W^+(\to \nu(p_3) + \mu^+(p_4)) + Z^0(\to e^-(p_5) + e^+(p_6))$	NLO
72	$W^+(\to \nu(p_3) + \mu^+(p_4)) + Z^0(\to \nu_e(p_5) + \bar{\nu}_e(p_6))$	NLO
73	$W^+(\to \nu(p_3) + \mu^+(p_4)) + Z^0(\to b(p_5) + \bar{b}(p_6))$	NLO
76	$W^-(\to \mu^-(p_3) + \bar{\nu}(p_4)) + Z^0(\to e^-(p_5) + e^+(p_6))$	NLO
77	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + Z^0(\to \nu(p_5) + \bar{\nu}(p_6))$	NLO
78	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + Z^0(\to b(p_5) + \bar{b}(p_6))$	NLO
81	$Z^0(\to \mu^-(p_3) + \mu^+(p_4)) + Z^0(\to e^-(p_5) + e^+(p_6))$	NLO
82	$Z^0(\to e^-(p_3) + e^+(p_4)) + Z^0(\to 3 \times (\nu(p_5) + \bar{\nu}(p_6)))$	NLO
83	$Z^0(\to e^-(p_5) + e^+(p_6)) + Z^0(\to b(p_3) + \bar{b}(p_4))$	NLO
84	$Z^0(\to b(p_3) + \bar{b}(p_4)) + Z^0(\to 3 \times (\nu(p_5) + \bar{\nu}(p_6)))$	NLO
85	$Z^0(\to e^-(p_3) + e^+(p_4)) + Z^0(\to 3 \times (\nu(p_5) + \bar{\nu}(p_6))) + f(p_7)$	LO
86	$Z^0(\to \mu^-(p_3) + \mu^+(p_4)) + Z^0(\to e^-(p_5) + e^+(p_6))$ [no γ^*]	NLO
87	$Z^0(\to e^-(p_3) + e^+(p_4)) + Z^0(\to 3 \times (\nu(p_5) + \bar{\nu}(p_6)))$ [no γ^*]	NLO
88	$Z^0(\to e^-(p_3) + e^+(p_4)) + Z^0(\to b(p_5) + \bar{b}(p_6))$ [no γ^*]	NLO
89	$Z^0(\to b(p_3) + \bar{b}(p_4)) + Z^0(\to 3 \times (\nu(p_5) + \bar{\nu}(p_6)))$ [no γ^*]	NLO
91	$W^+(\to \nu(p_3) + e^+(p_4)) + H(\to b(p_5) + b(p_6))$	NLO
92	$W^+(\to \nu(p_3) + e^+(p_4)) + H(\to W^+(\nu(p_5)e^+(p_6)) + W^-(e^-(p_7)\bar{\nu}(p_8)))$	NLO
96	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + H(\to b(p_5) + \bar{b}(p_6))$	NLO
97	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + H(\to W^+(\nu(p_5)e^+(p_6)) + W^-(e^-(p_7)\bar{\nu}(p_8)))$	NLO
101	$Z^0(\to e^-(p_3) + e^+(p_4)) + H(\to b(p_5) + b(p_6))$	NLO
102	$Z^0(\to 3 \times (\nu(p_3) + \bar{\nu}(p_4))) + H(\to b(p_5) + \bar{b}(p_6))$	NLO
103	$Z^0(\to b(p_3) + \bar{b}(p_4)) + H(\to b(p_5) + \bar{b}(p_6))$	NLO
106	$Z^0(\to e^-(p_3) + e^+(p_4)) + H(\to W^+(\nu(p_5)e^+(p_6)) + W^-(e^-(p_7)\bar{\nu}(p_8)))$	NLO
107	$Z^0(\to 3 \times (\nu(p_3) + \bar{\nu}(p_4))) + H(\to W^+(\nu(p_5)e^+(p_6)) + W^-(e^-(p_7)\bar{\nu}(p_8)))$	NLO
108	$Z^0(\to b(p_3) + \bar{b}(p_4)) + H(\to W^+(\nu(p_5)e^+(p_6)) + W^-(e^-(p_7)\bar{\nu}(p_8)))$	NLO
111	$H(\to b(p_3) + b(p_4))$	NLO
112	$H(\to \tau^-(p_3) + \tau^+(p_4))$	NLO
113	$H(\to W^+(\nu(p_3) + e^+(p_4)) + W^-(e^-(p_5) + \bar{\nu}(p_6)))$	NLO
114	$H(\to Z^0(\mu^-(p_3) + \mu^+(p_4)) + Z^0(e^-(p_5) + e^+(p_6))$	NLO
115	$H(\to Z^0(3 \times (\nu(p_3) + \bar{\nu}(p_4))) + Z^0(e^-(p_5) + e^+(p_6))$	NLO
116	$H(\to Z^0(\mu^-(p_3) + \mu^+(p_4)) + Z^0(b(p_5) + \bar{b}(p_6))$	NLO

nproc	$f(p_1) + f(p_2) \to \ldots$	Order
141	$H(\to b(p_3) + \bar{b}(p_4)) + b(p_5)(+g(p_6))$	NLO
142	$H(\to b(p_3) + \bar{b}(p_4)) + \bar{b}(p_5)(+b(p_6))$	NLO
143	$H(\to b(p_3) + \bar{b}(p_4)) + b(p_5) + \bar{b}(p_6)$ [both observed]	NLO
151	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{b}(p_6)) + e^-(p_7) + \bar{\nu}(p_8))$	LO
152	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{b}(p_6)) + q(p_7) + \bar{q}(p_8))$	LO
156	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{\nu}(p_7) + e^-(p_8) + b(p_6)) + g(p_9)$	LO
157	$t\bar{t}$ [for total Xsect]	NLO
158	$b\bar{b}$ [for total Xsect]	NLO
159	$c\bar{c}$ [for total Xsect]	NLO
160	$t\bar{t} + g$ [for total Xsect]	LO
161	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + q(p_6)$ [t-channel]	NLO
162	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + q(p_6)$ [decay]	NLO
166	$\bar{t}(\to e^-(p_3) + \bar{\nu}(p_4) + \bar{b}(p_5)) + q(p_6)$ [t-channel]	NLO
167	$\bar{t}(\to e^-(p_3) + \bar{\nu}(p_4) + \bar{b}(p_5)) + q(p_6)$ [decay]	NLO
171	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{b}(p_6))$ [s-channel]	NLO
172	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{b}(p_6))$ [decay]	NLO
176	$\bar{t}(\to e^-(p_3) + \bar{\nu}(p_4) + \bar{b}(p_5)) + b(p_6))$ [s-channel]	NLO
177	$\bar{t}(\to e^-(p_3) + \bar{\nu}(p_4) + \bar{b}(p_5)) + b(p_6))$ [decay]	NLO
180	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + t(p_5)$	NLO
181	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + t(\nu(p_5) + e^+(p_6) + b(p_7))$	NLO
182	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + t(\nu(p_5) + e^+(p_6) + b(p_7))$ [decay]	NLO
185	$W^+(\to \nu(p_3) + e^+(p_4)) + \bar{t}(p_5)$	NLO
186	$W^+(\to \nu(p_3) + e^+(p_4)) + \bar{t}(e^-(p_5) + \bar{\nu}(p_6) + \bar{b}(p_7))$	NLO
187	$W^+(\to \nu(p_3) + e^+(p_4)) + \bar{t}(e^-(p_5) + \bar{\nu}(p_6) + \bar{b}(p_7))$ [decay]	NLO
190	$t(p_3) + \bar{t}(p_4) + H(p_5)$	LO
191	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{\nu}(p_7) + e^-(p_8) + \bar{b}(p_6)) + H(p_9 + p_{10})$	LO
196	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{\nu}(p_7) + e^-(p_8) + b(p_6)) + Z(e^-(p_9), e^+(p_{10}))$	LO
197	$t(\to \nu(p_3) + e^+(p_4) + b(p_5)) + \bar{t}(\to \bar{\nu}(p_7) + e^-(p_8) + \bar{b}(p_6)) + Z(b(p_9), \bar{b}(p_{10}))$	LO
201	$H(\to b(p_3) + \bar{b}(p_4)) + f(p_5)$ [full mt dep.]	LO
202	$H(\to \tau^-(p_3) + \tau^+(p_4)) + f(p_5)$ [full mt dep.]	LO
203	$H(\to b(p_3) + \bar{b}(p_4)) + f(p_5)$	NLO
204	$H(\to \tau^-(p_3) + \tau^+(p_4)) + f(p_5)$	NLO
206	$A(\to b(p_3) + \bar{b}(p_4)) + f(p_5)$ [full mt dep.]	LO
207	$A(\to \tau^-(p_3) + \tau^+(p_4)) + f(p_5)$ [full mt dep.]	LO
208	$H(\to W^+(\to \nu(p_3) + e^+(p_4)) + W^-(\to e^-(p_5) + \bar{\nu}(p_6))) + f(p_7)$	NLO

nproc	$f(p_1) + f(p_2) \to \ldots$	Order
211	$H(\to b(p_3) + b(p_4)) + f(p_5) + f(p_6)$ [WBF]	NLO
212	$H(\to \tau^-(p_3) + \tau^+(p_4)) + f(p_5) + f(p_6)$ [WBF]	NLO
213	$H(\to W^+(\nu(p_3)e^+(p_4)) + W^-(e^-(p_5)\bar{\nu}(p_6))) + f(p_7) + f(p_8)$ [WBF]	NLO
216	$H(\to b(p_3) + \bar{b}(p_4)) + f(p_5) + f(p_6) + f(p_7)$ [WBF+jet]	NLO
217	$H(\to \tau^-(p_3) + \tau^+(p_4)) + f(p_5) + f(p_6) + f(p_7)$ [WBF+jet]	NLO
221	$\tau^-(\to e^-(p_3) + \bar{\nu}_e(p_4) + \nu_\tau(p_5)) + \tau^+(\to \bar{\nu}_\tau(p_6) + \nu_e(p_7) + e^+(p_8))$	LO
261	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5)$	NLO
262	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5)$	NLO
263	$Z^0(\to e^-(p_3) + e^+(p_4)) + \bar{b}(p_5) + b(p_6)$ [1 b-tag]	NLO
264	$Z^0(\to e^-(p_3) + e^+(p_4)) + \bar{c}(p_5) + c(p_6)$ [1 c-tag]	NLO
266	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5)(+\bar{b}(p_6))$	NLO
267	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5)(+\bar{c}(p_6))$	NLO
271	$H(b(p_3) + b(p_4)) + f(p_5) + f(p_6)$ [in heavy top limit]	LO
272	$H(\tau^-(p_3) + \tau^+(p_4)) + f(p_5) + f(p_6)$ [in heavy top limit]	LO
273	$H(b(p_3) + \bar{b}(p_4)) + f(p_5) + f(p_6) + f(p_7)$ [in heavy top limit]	LO
274	$H(\tau^-(p_3) + \tau^+(p_4)) + f(p_5) + f(p_6) + f(p_7)$ [in heavy top limit]	LO
311	$f(p_1) + b(p_2) \to W^+(\to \nu(p_3) + e^+(p_4)) + b(p_5) + f(p_6)$	LO
316	$f(p_1) + b(p_2) \to W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + b(p_5) + f(p_6)$	LO
321	$f(p_1) + b(p_2) \to W^+(\to \nu(p_3) + e^+(p_4)) + c(p_5) + f(p_6)$	LO
326	$f(p_1) + b(p_2) \to W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + c(p_5) + f(p_6)$	LO
331	$W^+(\to \nu(p_3) + e^+(p_4)) + c(p_5) + f(p_6)$ [c-s interaction]	LO
336	$W^-(\to e^-(p_3) + \bar{\nu}(p_4)) + c(p_5) + f(p_6)$ [c-s interaction]	LO
341	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5) + f(p_6)[+f(p_7)]$	NLO
342	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5) + f(p_6)[+b\ (p_7)]$	(REAL)
346	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5) + f(p_6) + f(p_7)$	LO
347	$Z^0(\to e^-(p_3) + e^+(p_4)) + b(p_5) + f(p_6) + b\ (p_7)$	LO
351	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5) + f(p_6)[+f(p_7)]$	NLO
352	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5) + f(p_6)[+c\ (p_7)]$	(REAL)
356	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5) + f(p_6) + f(p_7)$	LO
357	$Z^0(\to e^-(p_3) + e^+(p_4)) + c(p_5) + f(p_6) + c\ (p_7)$	LO
902 –	[Internal consistency checks]	–

Table 5: Processes indicated by choice of the variable nproc.

Process ($V \in \{Z, W, \gamma\}$)	Comments
Calculations completed since Les Houches 2005	
1. $pp \to VV$jet	WWjet completed by Dittmaier/Kallweit/Uwer [3]; Campbell/Ellis/Zanderighi [4] and Binoth/Karg/Kauer/Sanguinetti (in progress)
2. $pp \to$ Higgs+2jets	NLO QCD to the gg channel completed by Campbell/Ellis/Zanderighi [5]; NLO QCD+EW to the VBF channel completed by Ciccolini/Denner/Dittmaier [6, 7]
3. $pp \to VVV$	ZZZ completed by Lazopoulos/Melnikov/Petriello [8] and WWZ by Hankele/Zeppenfeld [9]
Calculations remaining from Les Houches 2005	
4. $pp \to t\bar{t}b\bar{b}$	relevant for $t\bar{t}H$
5. $pp \to t\bar{t}$+2jets	relevant for $t\bar{t}H$
6. $pp \to VVb\bar{b}$,	relevant for VBF $\to H \to VV$, $t\bar{t}H$
7. $pp \to VV$+2jets	relevant for VBF $\to H \to VV$ VBF contributions calculated by (Bozzi/)Jäger/Oleari/Zeppenfeld [10–12]
8. $pp \to V$+3jets	various new physics signatures
NLO calculations added to list in 2007	
9. $pp \to b\bar{b}b\bar{b}$	Higgs and new physics signatures
Calculations beyond NLO added in 2007	
10. $gg \to W^*W^*$ $\mathcal{O}(\alpha^2\alpha_s^3)$	backgrounds to Higgs
11. NNLO $pp \to t\bar{t}$	normalization of a benchmark process
12. NNLO to VBF and Z/γ+jet	Higgs couplings and SM benchmark
Calculations including electroweak effects	
13. NNLO QCD+NLO EW for W/Z	precision calculation of a SM benchmark

Table 1: The updated experimenter's wishlist for LHC processes

VBF = Vector Boson Fusion

3.16

* Jets (general approach)

• from e^+e^- to pp collisions

The fact that the beams are hadronic themselves have a few effects on how one has to treat jet clustering.

1) e^+e^- jet events are fairly simple/quiet (see Figures): outside of the few "clusters" that make the jets, there is not much activity in the detector.

For pp collisions, since the beams are hadronic, we have additional (non-perturbative) (strong) interactions, between the beams, the beam remnants and the initial and final-state particles (INITIAL AND FINAL-STATE INTERACTIONS). The net effect is to produce additional soft particles more-or-less uniformly all over the detector. This is know as the UNDERLYING EVENT. (UE).
Typically, the multiplicity goes from a few 10 to a few 100 (Note: in the case of heavy ion collisions we can have up to 40.000 particles).
Monte-Carlo use models for the UE. The parameters are used to describe Tevatron data, Extrapolation to LHC energies can give uncertainties of a factor $\simeq 2$, which are the subject of dedicated studies.

2) As already discussed, because of boost-invariance requirements, one uses the rapidity (y) instead of the polar angle (θ). The jet definitions will have to be adapted to this new variable.

3) The hard scattering has incoming partons i.e. we will also have perturbative radiation from the initial state.

If the 2 partons are $p_1^\mu \equiv x_1(0,0,\frac{\sqrt{s}}{2},\frac{\sqrt{s}}{2})$
$p_2^\mu \equiv x_2(0,0,-\frac{\sqrt{s}}{2},\frac{\sqrt{s}}{2})$ $\quad p_1\cdot p_2 = x_1 x_2 \frac{s}{2}$

the radiation (from the antenna formula) will be

$$\frac{\alpha_s}{2\pi}\ \frac{p_1\cdot p_2}{p_1\cdot k\ p_2\cdot k}\ d^4(k)\ \delta(k^2)$$

if

$$k^\mu \equiv \left(k_t\cos(\phi);\ k_t\sin(\phi), E_t\sinh(y), E_t\cosh(y)\right)$$

$$\Rightarrow d^4k\ \delta^+(k^2) = k_t\,dk_t\,d\phi\ E_t\,dE_t\,dy\ \delta^+(E_t^2-k_t^2) = \frac{1}{2}k_t\,dk_t\,d\phi\,dy$$

$$\Rightarrow p_2\cdot k = x_2 k_t\frac{\sqrt{s}}{2}\left[\cosh(y)-\sinh(y)\right] \Rightarrow (p_1\cdot k)(p_2\cdot k) = x_1x_2\frac{s}{4}k_t^2$$

Hence the radiation is

$$\frac{\alpha_s}{\pi}\ \frac{dk_t}{k_t}\ d\phi\ d\eta$$

i.e. uniform & boost invariant (independent on x_1 & x_2)

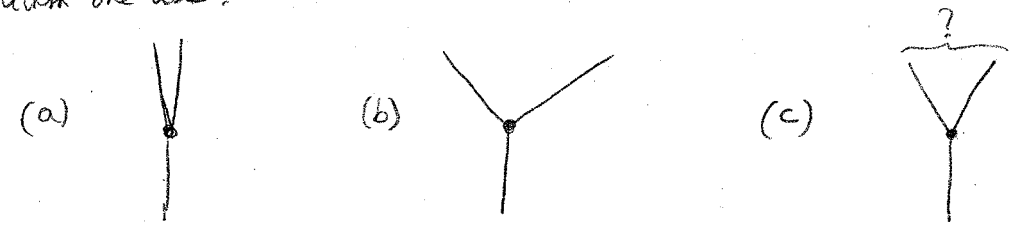
4) Because of the high luminosity LHC will operate at, many pp interactions can occur per bunch crossing. At high lumi, this corresponds to 0.25 mb^{-1} per bunch crossing i.e. typically 25 interactions (per bunch crossing). A hard event will thus be seen together with many others, typically soft. These extra interactions, known as PILEUP (PU) bring another set of a few thousand particles to the event, reasonably uniformly over the detector.

Note: At the Tevatron, PU was already present, but with a much smaller effect.

• (see next page)

• Defining jets for pp colliders

As we have already seen for e^+e^- collisions, the "jets" in an event depend on the clustering algorithm one uses.



(a) should be classified as 2 jets, (b) as 3 jets, but (c) is somehow unclear.

At the Tevatron in 1990, one has agreed on a set of rules that any jet algorithm should satisfy (the SNOWMASS Accords):

1) Simple to implement in an experimental analysis
2) Simple to implement in theoretical calculations
3) Defined at any order of perturbation theory
4) Yields finite cross sections at any order of perturbation theory
5) Yields a cross section that is relatively insensitive to hadronisation.

This means that the algorithm can be used
- by theoreticians i.e. well-defined, finite perturbatively
- by experimentalists i.e. well-defined, fast enough, as insensitive as possible to the UE (& PU)

We will see later that those fundamental requirements have not always been fulfilled.

- What are the jets used for? (an incomplete list!)
  - inclusive cross-section: constraint on the PDF
  - W/Z + jets
  - $t \to W^+ + b \to q\bar{q}b \to 3$ jets (Note: for boosted tops, the 3 jets are very close ⇒ difficult to disentangle)
  - diffraction / forward jets / Mueller-Navelet jets
  - jet shapes
  - $H \to b\bar{b}$
  - $Z' \to q\bar{q}, \ldots$ } peak in the dijet invariant mass spectrum.
  - Many applications in SUSY searches:
    4/8 jets + missing E_t
    no jets above $p_t = \ldots$

  - ---

- Jets at different levels
  - perturbative level: $2 \to n$ matrix element computed exactly at $O(\alpha_s^p)$
  - parton level: also including parton showers
    This typically resums collinear divergences $O(\alpha_s^p \log^p(k_t^2/\mu_F^2))$
  - hadron level: since partons are not directly observed, one need a model to transform the final-state partons into hadrons
  - multiple interactions / UE: includes interactions with the beam (remnants) that give the underlying event
  - detector level: take into account the response of the detector (efficiencies, calorimeter towers, ...)
  - pile-up: add multiple pp interactions together (Poisson distribution of average $\sigma_{tot}\mathcal{L}_b$, with $\mathcal{L}_b$ the luminosity per bunch crossing)

$$\mathcal{L} = \frac{n \cdot N_A N_B}{t\, S} \simeq 10^{34}\, cm^{-2} s^{-1}$$

$$\mathcal{L}_b = \frac{N_A N_B}{S} = \frac{t}{n}\mathcal{L} \simeq 2510^{-26}\, cm^{-2} \simeq 0.25\, mb^{-1}$$

$N_A = N_B = 1.15\, 10^{11}$ particles per bunch
$n = 2808$ # bunches/beam
$t = n \cdot 25\, 10^{-9}$ s revolution time
$S = 4\pi\sigma^*$
$\sigma^* = 16.6\, \mu m$ beam size

For $\sigma_{tot} \simeq 100$ mb, this means around 25 simultaneous pp interactions

- the k_t & Cambridge/Aachen algorithm.

the k_t algorithm introduced in e^+e^- collisions, can easily be adapted to pp using new distances in ϕ and y i.e.,

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2)\left(\Delta\phi_{ij}^2 + \Delta y_{ij}^2\right)$$

and stop the clustering when $(d_{ij})_{min} < d_{cut}$. This is known as the exclusive k_t algorithm.

However, one usually prefers the (longitudinally invariant) inclusive k_t algorithm which introduces a k_t distance between 2 objects

$$d_{ij}^{(k_t)} = \frac{\min(k_{ti}^2, k_{tj}^2)}{R^2}\left(\Delta y_{ij}^2 + \Delta\phi_{ij}^2\right)$$

as well as a beam distance for every object

$$d_{iB}^{(k_t)} = \frac{k_{ti}^2}{R^2}$$

Clustering proceeds as for the exclusive version, by merging the 2 objects corresponding to the smallest distance, except that, when the smallest distance is a beam distance, the object is said to recombine with the beam and is promoted as a jet.

The Cambridge/Aachen algorithm works the same way with distances

$$d_{ij}^{C/A} = \frac{1}{R^2}\left(\Delta y_{ij}^2 + \Delta\phi_{ij}^2\right) \quad \text{and} \quad d_{iB}^{C/A} = 1$$

In both cases, one has introduced a parameter R (instead of d_{cut} for the exclusive case). At lowest order, both algorithms will recombine two particles if the (geometrical y,ϕ) distance between them is $\leq R$.
(this is more complicated with more particles and depends on the algorithm)
The parameter R controls therefore the "size" of the jets.

Pro: nice perturbative behaviour (theoretically) (remember soft gluons resummation in e^+e^-)

Cons: In practice: large sensitivity to the UE

• The cone algorithm

Most modern cone algorithms start with the concept of a stable cone: a cone centred on (y_C, ϕ_C) and of half opening angle (i.e. radius) R, is said STABLE if its total 4-momentum

$$P_C \equiv \sum_{i \in C} p_i$$

points in the same direction as the centre itself, i.e. if p_C has rapidity y_C and azimuth ϕ_C.

A stable cone defines a dominant direction of energy flow and is thus well-suited to define jets. We thus want to define

$$\{\text{jets}\} \equiv \{\text{stable cones}\}$$

However, stable cones can overlap, so one has to address that in order to get the jets.

Before doing so, let us mention that, usually, stable cones are found by iterative methods (SEEDED cones).

1) start from a given centre (SEED)
2) for that seed, compute the cone contents 4-momentum
3) if it is stable, we have a stable cone
   if it is not, use this new direction as the new seed
   and go back to 2 (i.e. iterate)

Notes: • one usually imposes a maximal number of iterations to avoid endless loops

• Ratcheting: one sometimes imposes that while iterating, the cone centre remains within the initial cone

We are now in position to specify the 2 main classes of cone algorithms.

Cone with split-merge:

1) Find **all** stable cones (of radius R). We will call them protojets

1') If some particles do not belong to any stable cone,
rerun (recursively) the stable-cone search on those only.

2) Split-merge: take the 2 hardest protojets (in p_t or E_t) that overlap.
if (the 4-momentum of) their overlap has a p_t (or E_t) above a fraction f of the 2nd hardest jet p_t (or E_t), merge the 2 protojets.
Otherwise split them by associating each particle to the protojet to which centre it is closer.
f is called the overlap threshold
Repeat until no overlapping protocones are left.
The remaining protocones are called the jets

Parameters: the radius R
the overlap threshold f
optional: a number of pass n_{pass} (for 1' multi-pass search)
a $p_{t,min}$ threshold for stable cones entering split-merge
a maximal number of iterations for stable-cone search
an option for ratcheting
a seed threshold

Cone with progressive removal

1) Start with the hardest (remaining) particle as a seed
2) Iterate until stable, remove the particles and call that a jet.
3) go back to 1.

Parameters: the radius R
optional: a maximal number of iterations for stable-cone search
a minimal threshold for the seed

- Practically, the cone algorithm turns out to be less sensitive to the underlying event, so it is used by many of the pp experiments. Though the cone with split-merge is usually preferred, the cone with progressive removal has the feature that the hardest jets are fully circular i.e. soft resilient. This property is known to ease the calibration of jets.

- We have already mentionned that stable cones are most of the time found by iterating from seeds. But so far, we have not specified the set of seeds. The simplest option is to start from

$$\{\text{seeds}\} = \{\text{particles in the event, potentially over a } p_{t,\min} \text{ threshold}\}$$

CDF JetClu and the ATLAS Cone are such examples.

- <u>Midpoint</u> and <u>JetClu / ATLAS Cone IR unsafety</u>

The previous set of seed is troublesome.
The failure comes from the requirement that a jet algorithm need to give finite results in perturbative theory.
As we already saw in e^+e^- jets, this means that the algorithm has to be infrared and collinear safe (IRC safe) (otherwise the cancellation between real and virtual diagrams would not take place order-by-order in the perturbative series).

In the case of the cone algorithms, this noticeably means that the set of stable cone must be IRC safe i,e. cannot change if one adds an infinitely soft gluon to the event, or split one particle in 2 collinear ones.

However, if one considers the following event with 2 hard particles distant by more than R but less than 2R

Starting with the particles as seeds, one finds 2 stable cones that do not overlap, hence we find 2 jets.
But, if one adds a soft gluon between the 2 hard particles, a 3rd stable cone taking all particles is found.

The overlap between this new stable cone and the 2 previously found is 100% of the latter. Therefore, protojets will be merged (whatever f is) and we will find only 1 jet.

Thus, CDF JetClu and the ATLAS Cone are IR unsafe!

Note also that a seed threshold would solve that problem. Unfortunately, because collinear splitting can lower the p_t below the threshold, leading to stable cones being missed, this situation would be collinear unsafe!

To solve that problem, the set of seeds needs to be extended:

$$\{\text{seeds}\} = \{\text{particles in the event, potentially over a } p_{t,\min} \text{ threshold}\} + \{\text{midpoints between the stable cones found with the previous set}\}$$

This obviously solves the problem that we just discussed with JetClu, as in the first event, iterating from the midpoint between the 2 first stable cone found will find the one we had missed.
Note however that the collinear unsafety remains when a seed threshold is applied!

The cone algorithm with this new set of seed is usually called MidPoint.

Here is a summary of the algorithms we have so far

k_t	recombination, $d_{ij} = \min(k_{ti}^2, k_{tj}^2)(\Delta y^2 + \Delta\phi^2)$, $d_{iB} = R^2 k_{ti}^2$
Cambridge/Aachen	recombination, $d_{ij} = \Delta y^2 + \Delta\phi^2$ $\quad d_{iB} = R^2$
CDF JetClu ATLAS Cone	seeded cone, particles as seeds
CDF MidPoint DO runII Cone Px Cone	seeded cone, particles as seeds + midpoints
CMS Iterative Cone	cone with progressive removal

- Recombination scheme:

one thing we have not yet addressed is how to sum particle's momenta when recombining particles or checking the stability of a cone.

Mainly because they do not have access to the full 4-vector, experimentalists have often use the p_t-scheme recombination which uses a p_t-weighted sum:

$$p_t = p_{ta} + p_{tb}$$

$$y = \frac{1}{p_t}(p_{ta} y_a + p_{tb} y_b)$$

$$\phi = \frac{1}{p_t}(p_{ta} \phi_a + p_{tb} \phi_b) \quad \text{(modulo some care about the periodicity)}$$

(note: one can invariantly use p_t or E_t). Experimentally, one measures the polar angle θ and get $y = -\log(\tan(\theta/2))$ and $E_t = E \sin(\theta)$, E being the measured energy.

Note that with that scheme, the objects remain massless after recombination.

For that reason, and also because it is theoretically better justified, recent jet analysis have switched to the E-scheme recombination which directly sums 4-vectors.

- Recent progress: I - speeding up the k_t algorithm.

Q: How to implement the k_t clustering ?

Solution 1: KtJet

(i) compute all d_{ij}, d_{iB}, find the minimum $\sim N^2$
(ii) recombine the closest pair (or) get a new jet ~ 1
(iii) restart from (i) with the remaining particles $\times N$

$O(N^3)$

This implementation is $O(N^3)$ i.e. rather slow.
Some people have complained about that speed problem and switched to the cone algorithm (JetClu : $O(N^2)$, Midpoint : $O(N^3)$ with a coefficient depending on the seed threshold).

BUT : we can do better (Cacciari, Salam : 06)

Idea/Lemma : if d_{ij} is the minimal distance and $k_{ti} < k_{tj}$, then

$$(\Delta y^2 + \Delta\phi^2)_{ij} < (\Delta y^2 + \Delta\phi^2)_{il} \qquad \forall l \neq i,j$$

Proof : assume it is not true. Then it exists p s.t. $(R^2_{ab} = \Delta y^2_{ab} + \Delta\phi^2_{ab})$

$$R^2_{ip} < R^2_{ij}$$

This implies

$$\begin{aligned} d_{ip} &= \min(k_{ti}^2, k_{tp}^2)\, R^2_{ip} \\ &\leq k_{ti}^2\, R^2_{ip} \\ &\leq k_{ti}^2\, R^2_{ij} \\ &\leq \min(k_{ti}^2, k_{tj}^2)\, R^2_{ij} \equiv d_{ij} \end{aligned}$$

which is inconsistent with the hypothesis.

This results means that, if $n(i)$ is the nearest (geometrical) neighbour of i, then the minimal d_{ij} is among the $d_{i,n(i)}$.

That leads to a $O(N^2)$ implementation of the k_t algorithm (see next page)

Note that this is valid for a broader class of algorithm of the form

$$d_{ij} = \min\left[f(k_{ti}), f(k_{tj})\right] R^2_{ij}$$

which includes the Cambridge/Aachen algorithm.

Implementation 2 : $O(N^2)$

(i) for every particle, compute its nearest neighbour (NN) i.e. compute the $d_{i,n(i)}$, and the d_{iB}. $\sim N^2$

(ii) find the minimum distance among the $d_{i,n(i)}$, d_{iB} $\sim N$

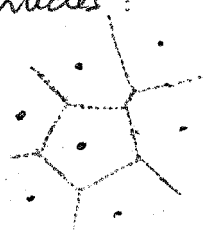
(iii) recombine i with n(i) or with the beam ~ 1

(iv) if necessary, recompute the NN of the new object, recompute the NN of the particles that had i and (if merged) n(i) as NN. $\sim 1 \times N$

(v) go back to (ii) $\times N$

This can actually be further improved using technique such as tiling and is presently the fastest method up to $N \sim 5000$.

For larger N, we can do even better. This uses yet another trick: the Voronoi diagram which is made of the bissector of all pairs of particles :



It has the property that the NN of a point is in one of the $O(1)$ surrounding cells. With appropriate techniques, building the Voronoi diagram is $O(N \log(N))$ (and allows step (i) to be $O(N)$). Dynamically updating the Voronoi diagram when one adds/removes a point is $O(\log(N))$ allowing to reduce step (iv) to $O(\log(N))$ (repeated N times).

Finally, using a priority queue (e.g a binary tree) that allow to keep the $d_{i,n(i)}, d_{iB}$ sorted at a cost of $O(N \log(N))$ for the initial build and $O(\log(N))$ for adding or removing an object, step (ii) becomes $O(1)$ (NB: we add $O(\log(N))$ to step (iv)).

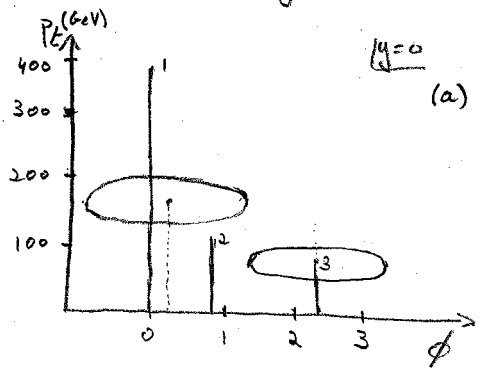
So, at the end, we have an implementation which is $O(N \log(N))$.

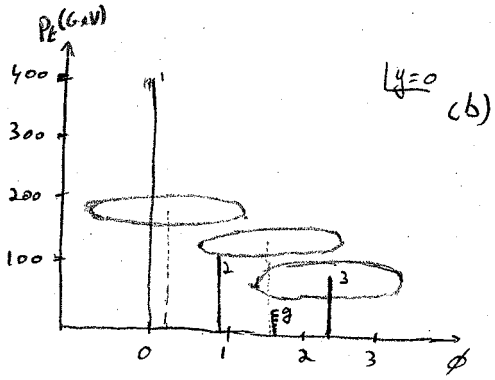
Note that the coefficient in front of the $N \log(N)$ is a bit large so that this only becomes faster than the $O(N^2)$ implementation for $N \gtrsim 5000$.

• Recent progress : II - IR unsafety of the MidPoint algorithm.

We have already seen that MidPoint has been introduced to solve an IR unsafety problem of the Jet Clu algorithm (when 2 particles were in a nearby vicinity).

The example below shows that all this do is to push the problem one order further in the perturbative expansion i.e. when 3 particles are in a common vicinity.





(a) no soft gluon : iterating from 1 gives a first stable cone (1+2)
2 gives the first stable cone again (1+2)
3 gives a second stable cone (3)
the midpoint between the 2 stable cones gives the first stable cone once more.

$\Rightarrow$ 2 jets $\{1,2\}$ and $\{3\}$

(b) soft gluon : iterating from 1,2,3 and the associated midpoint gives the two stable cones found previously
g gives a NEW stable cone (2+3)
the new midpoints come back to (1+2) or (3).

$\Rightarrow$ 3 stable cones : $\{1,2\}$, $\{2,3\}$ and $\{3\}$

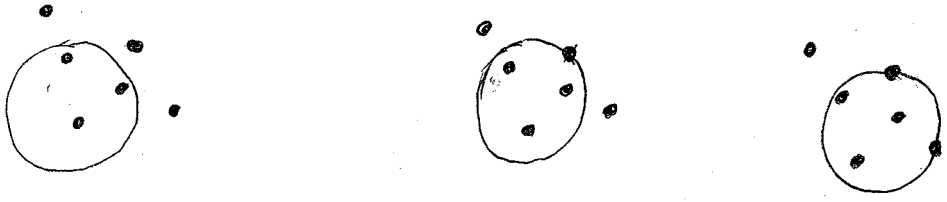
for $f = 0.5$, $\{1,2\}$ and $\{2,3\}$ will be merged into $\{1,2,3\}$
then $\{1,2,3\}$ and $\{3\}$ will be merged

$\Rightarrow$ 1 jet $\{1,2,3\}$

$\Rightarrow$ MidPoint is IR unsafe at the level of 3 particles in a common vicinity.

the solution?

- First, note that the IR unsafety comes from the fact that we have missed a stable cone. i.e. in the previous example, the set $\{2,3\}$ was already stable in the case (a) without the additional soft gluon and the problem comes from the fact that we have missed it (while the additional seed has allowed to find it).
- Adding combinations of triplets of stable cones (or similar ideas), beside making the algorithm way slower, would not solve the problem.
- Since the problem comes from the fact that the seeded/iterative implementation of the stable-cone search misses stable cones, the solution is rather logic: we need a method to find (provably) <u>all</u> stable cones.
- Naive solution: for every possible subset of the N initial particles, test if it forms a stable cone.
  Since there is 2^N subsets, this is $O(N.2^N)$ i.e. extremely slow.
  Thus, we need to find something better
- <u>Note</u>: for MidPoint, the complexity is $O(N^3)$:
  for a given seed, computing the cone contents is $O(N)$.
  checking if it is stable is $O(N)$ (= particle contents from the new direction)
  iteration multiplies by a constant
  ⇒ stable cone search from 1 seed is $O(N)$
  # seeds? from the initial N particles (i.e. $O(N)$ seeds), we find $O(N)$ stable cones
  This means $O(N^2)$ midpoints
  Iterating from this $O(N^2)$ midpoints gives the $O(N^3)$ complexity.
- <u>Idea</u>:



ANY circular enclosure in the (y,ϕ) plane (of fixed radius) can be
(i) translated into a random direction until it touches a point
(ii) rotated around that point until it touches a second point (*)
in such a way that the contents of the enclosure does not change.

(*) except if there is no other points within a distance $2R$ of the first point in which case it forms a stable cone on its own.

- As a consequence, by enumerating all possible pairs of points (distant by at most $2R$) + considering the 2 circles of radius R tangent to them + considering the 4 possible inclusion/exclusion states of the 2 points in the circle, we enumerate all possible enclosures.
- This gives a $O(N^3)$ recipe to find all stable cones:
  (i) get all enclosures by enumerating pairs of points $\rightarrow O(N^2)$
  (ii) for each enclosure, test stability $\times O(N)$
  $\overline{O(N^3)}$

- We can actually do slightly better. The idea is to first proof that, on average, there is $O(N)$ stable cones (exercise), hence we may hope getting some improvement by not performing $O(N^2)$ stability check.
  Then, note that performing a stability test is 2 things (both $O(N)$):
  (a) computing the cone contents to get the direction of its 4-momentum
  (b) test that the contents computed from this direction is the same as the initial one.
  We will use 3 tricks to improve this
  (i) for a parent particle, get all children particles it can form a pair with for each of these particles, there are 2 possible centres for the tangent circles. The first trick is to order all these centres associated with a given parent in angle. That way, when going from one to the next one, the contents changes at most by 2 particles (the 2 children associated with the 2 centres). That reduces (a) to $O(1)$.
  (ii) instead of making a full stability check for the $O(N^2)$ cone contents, just check the stability w.r.t. the parent and child (i.e. $O(1)$). This will leave $O(N)$ candidates at the end (instead of $O(N^2)$) for which we can carry a full stability check.
  (iii) During the process of successively computing the stability of each enclosure w.r.t. their (parent + child) pairs, we want to keep track of what has been already tested. To do that, we associate a q-bit tag to every particle and, by logical xor, to any set of particles. Contents are then compared by checkxor.
  This introduces a probability of collision that one can control by choosing q. (in practice, $q=96$ is more than enough for LHC purposes).

So the final algorithm goes as follows:

(i)	for all particle i ($i=1 \to N$) as "parent"	
(ii)	get the list of all particles in a distance $2R$ from i and the corresponding 2 centres for each of them (if there is no such particle, i forms a stable cone on its own)	$O(N)$
(iii)	Sort the centres found in (ii) by angle around the parent i	$O(N\log(N))$
(iv)	Take the first centre, compute its contents	$O(N)$
(v)	repeat	
(vi)	for each of the 4 cases of edge points being in/out	
(vii)	if the cone has not been found, add it to the list of distinct enclosures	$O(1)$
(viii)	if the cone is not labelled unstable, test its stability w.r.t. its parent & child particle. If not stable, label it as unstable.	$O(1)$
(ix)	end for	
(x)	Move to the next circle, different from the previous by at most 2 particles. The contents is updated by xor.	$O(1)$
(xi)	until all centres have been considered	$O(N\times 1 = N)$
(xii)	end for	$O(N^2\log(N))$
(xiii)	for each cone not labelled unstable	
(xiv)	check its stability. If stable, add it to the list of stable cones	$O(N)$
(xv)	end for	$O(N^2)$
		$O(N^2\log(N))$

Note: the split-merge that one has to run afterwards is $O(N^2)$ (Exercise)

So, we have managed to

① Solve the IR unsafety problem of MidPoint by finding ALL stable cones

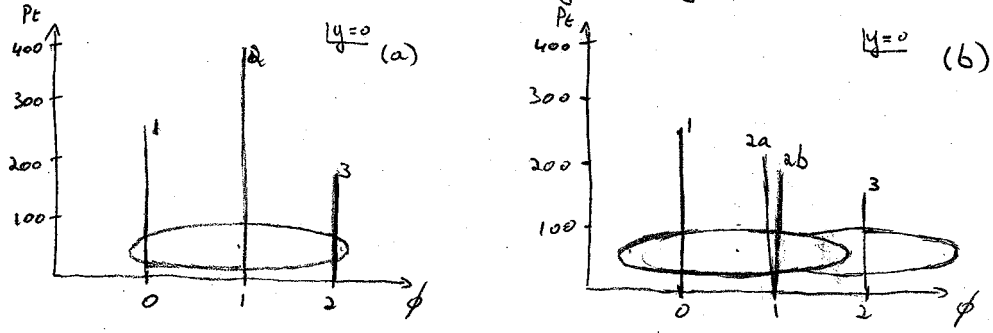
② in $O(N^2\log(N))$ instead of $O(N^3)$

The resulting cone algorithm is named SISCone

(for Seedless Infrared-Safe Cone algorithm)

• Recent progress: III : Collinear unsafety of the CDS Iterative Cone and the anti-k_t algorithm.

- As previously to show the IR unsafety of MidPoint, let us consider the following 3-particle event (in a nearby vicinity)



(a) starting from the hardest seed (2), we find a stable cone containing {1,2,3}

⇒ 1 jet {1,2,3}

(b) after collinear splitting of 2 into 2a and 2b, the hardest seed is 1. Starting from that seed, the stable cone found is {1,2a,2b} ≡ {1,2}
We are then left with 3 which is its own stable cone

⇒ 2 jets {1,2} & {3}

⇒ the CDS Iterative Cone is collinear unsafe.

- Replacement candidate: The distance used for recombination algorithms so far can be put under the form

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\left(\Delta y_{ij}^2 + \Delta\phi_{ij}^2\right) \qquad d_{iB} = R^2 k_{ti}^{2p}$$

with $p=1$ for k_t and $p=0$ for the Cambridge/Aachen.
We introduce the anti-k_t algorithm corresponding to $p=-1$.

The reason why this works as the iterative cone is the following: if a particle has a large k_t, it will be associated a small distance. Thus the hard particles will aggregate the softer in their surrounding and the hard jets will be circular (and soft-resilient, in the sense that a soft particle will not modify their shape)
And this is exactly the characteristic of the Iterative Cone.

Notes: • $p>0$ ($p<0$) has a behaviour similar to k_t (anti-k_t)
• anti-k_t has a fast implementation using the same techniques as for k_t.

• Recent progress: IV - a final note on IRC unsafety of MidPoint and the Iterative Cone.

We have mentionned that both the IR unsafety of the MidPoint algorithm and the collinear unsafety of the Iterative Cone start when we have 3 particles in a nearby vicinity, plus a 4th one (QCD or electroweak) to balance momentum.

Let us thus consider the case of W + jets. These 3-particle configurations are of order $\alpha_{ew}\alpha_s^2$. The divergence due to the unsafety will appear at the next order and should be regularised by non-perturbative effects at a scale of the order of Λ_{QCD}. Thus, the collinear splitting or soft emission will be of order $\alpha_{ew}\alpha_s^3 \log(P_{t,jet}/\Lambda_{QCD})$. Since $\alpha_s \log(P_{t,jet}/\Lambda_{QCD})$ is of order 1, we end with a contribution of order $\alpha_{ew}\alpha_s^3$. This means that the previous computation at order $\alpha_{ew}\alpha_s^3$ is of no meaning (as the IRC unsafety will give a correction of the same order).

Initially, one could have thought that because of collinear splitting or soft gluon emission requires one power of α_s, the problem appear one order further.

Actually, the IRC unsafety appear at the first order where MidPoint misses a stable cone (or the Iterative Cone is sensitive to collinear split.)

This is summarised in the next table

process	1st cone missed	order	last meaningful order
inclusive cross-section	α_s^4	NNLO	NLO
W/Z + jets	$\alpha_{ew}\alpha_s^3$	NNLO	NLO
3-jet events	α_s^4	NLO	LO
W/Z + 2 jets	$\alpha_{ew}\alpha_s^3$	NLO	LO
masses in 3-jet events	α_s^4	LO	none!

Notes: • some of these are already accessible by theory computations

• JetClu and the ATLAS Cone do not use the midpoint seeds i.e. the last meaningful order is 1 order lower than in the table. Note that some Tevatron comparisons of W/Z + jets with NLO QCD have been done using JetClu.

- Jet definitions in the experiments

  - The TEVATRON

    because of its sensitivity to the UE, the Tevatron experiments have not use k_t very extensively (they used it sometimes)

    JetClu, though superseeded by MidPoint, has been used quite a lot (probably still used sometimes)

    Midpoint is (by far) the most extensively used. Both CDF and DO use their implementations of the algorithm (they are slightly different)

  - The LHC

    Access to a large number of algorithms:
    - k_t, Cambridge/Aachen, anti-k_t
    - SISCone
    - MidPoint (probably the CDF implementation)
    - The ATLAS Cone or the CMS Iterative Cone

    Note: used at various levels:
    - (low-level) triggers, mostly to reduce rates (speed is a real issue here!)
    - analysis

  - There are other (less used) jet algorithms not presented here!

- A final "word of caution": jets depend on the definition one uses to cluster the event. To have a consistent comparison (data vs. theory for example) it is important to use the same definition in both cases. This means that, whenever one uses jets, one needs to specify:

  (i) the precise algorithm (e.g. "the cone algorithm with $R=0.4$" is not sufficient! MidPoint? SISCone?)

  (ii) the parameters (R, f, ...)