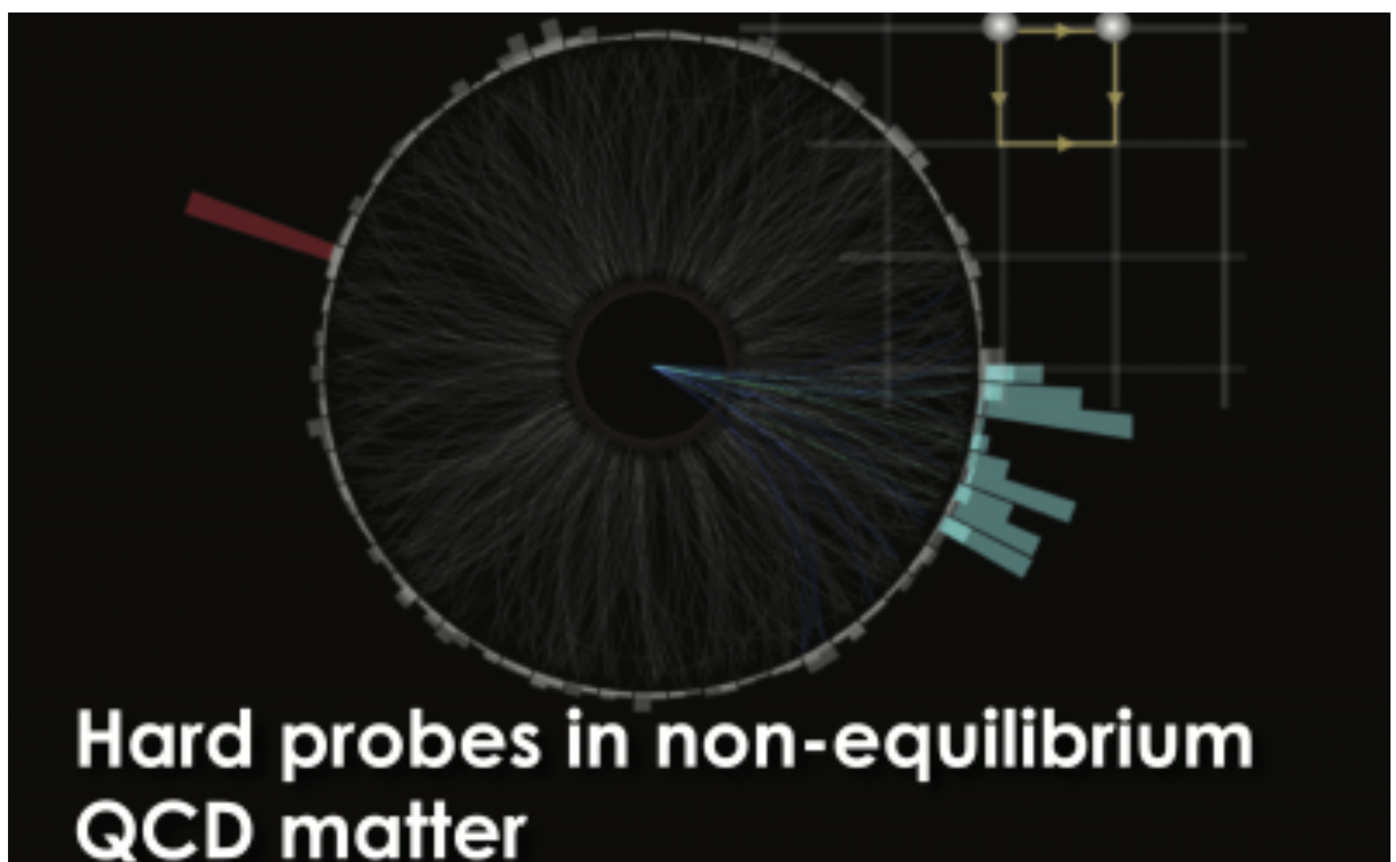
Jets at LHC and Future Colliders
The inevitable ML connection

Sanmay Ganguly
IIT-Kanpur
sanmayphy@github.io

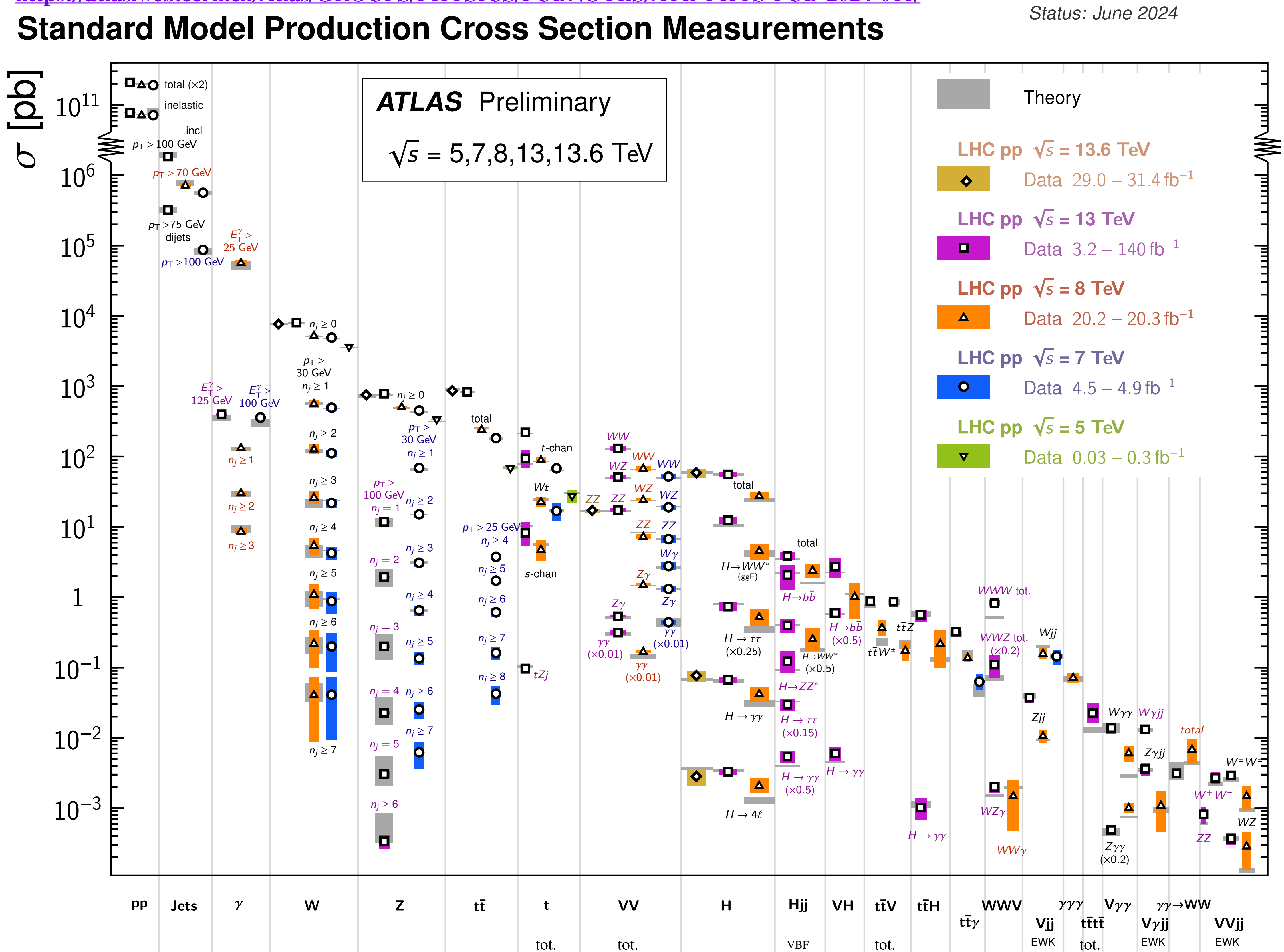
26th March, 2026

Part-1 : From Jets to ML4Jets

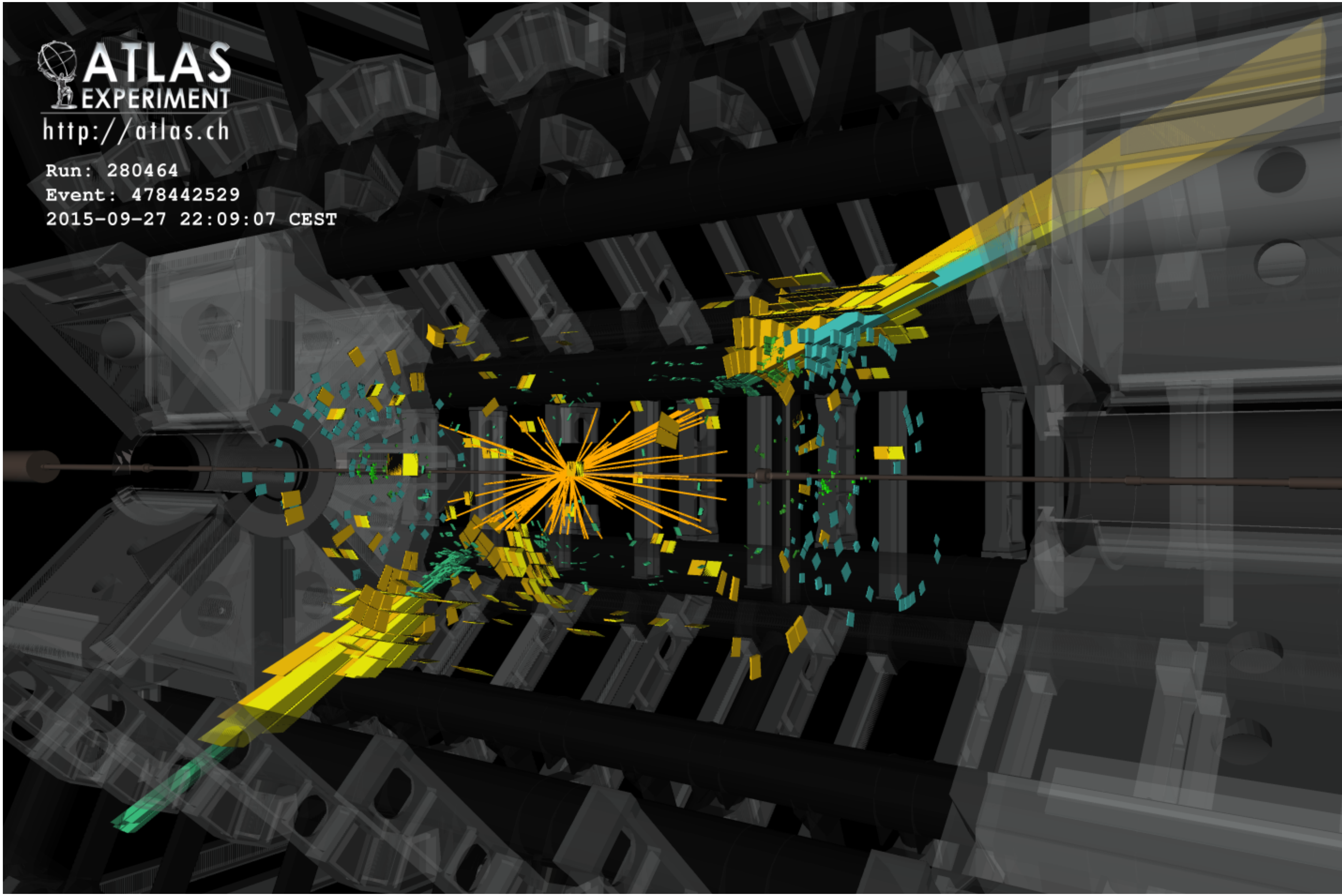


What does raw measurement reveals?

https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2024-011/



What precisely these “JET” events are?



What is our best explanation?

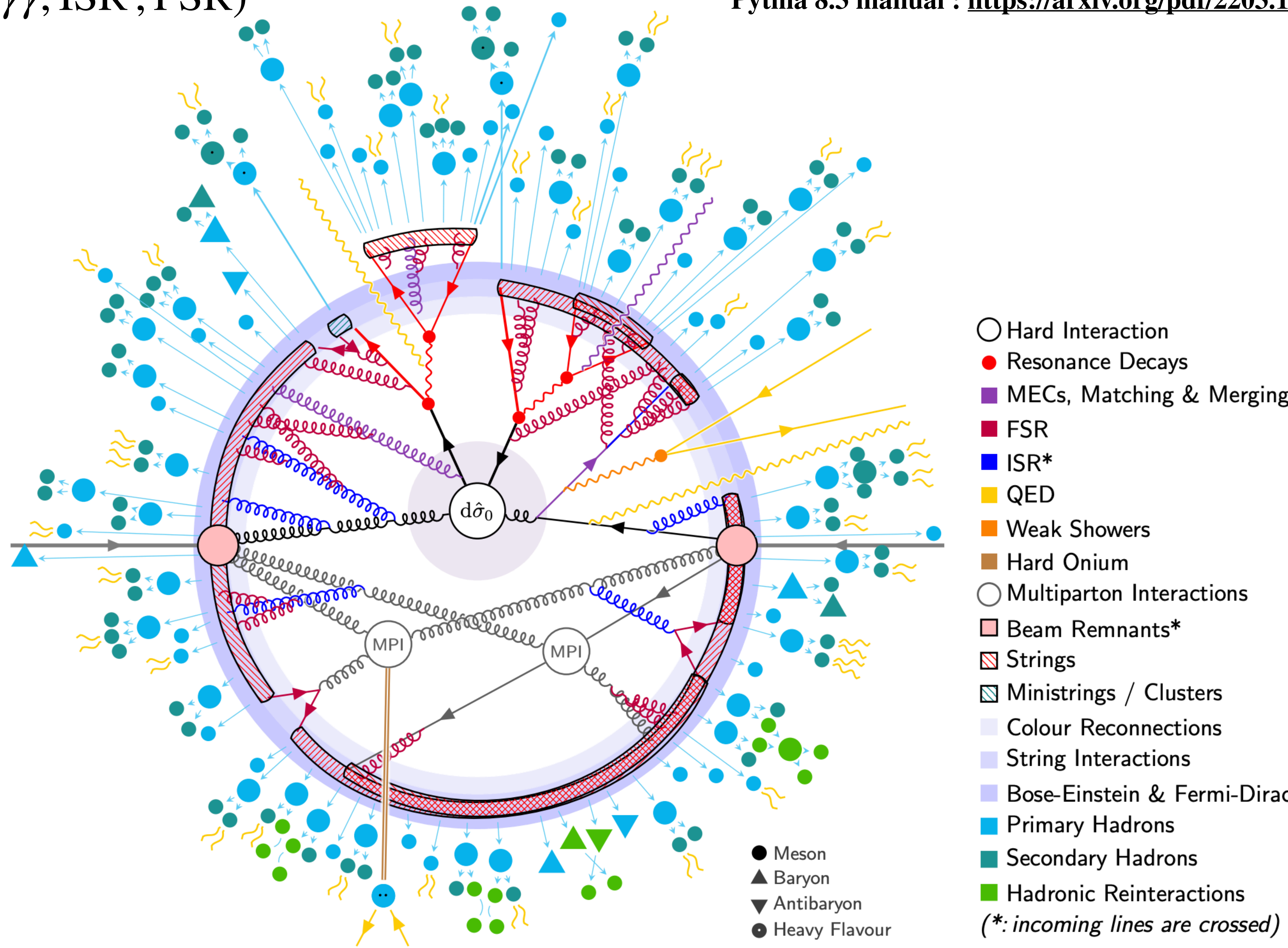
$\gamma\ (\pi^0 \to \gamma\gamma, \mathrm{ISR}\ , \mathrm{FSR})$

Pythia 8.3 manual : https://arxiv.org/pdf/2203.11601

$K_L\, , K_S$

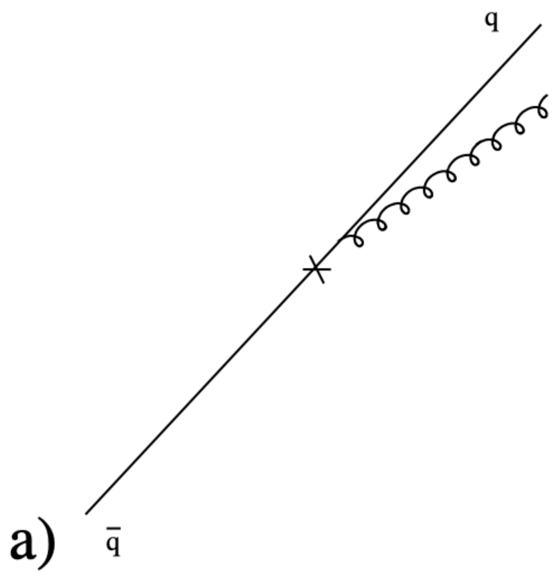
$\pi^{\pm}$

$l^{\pm}$

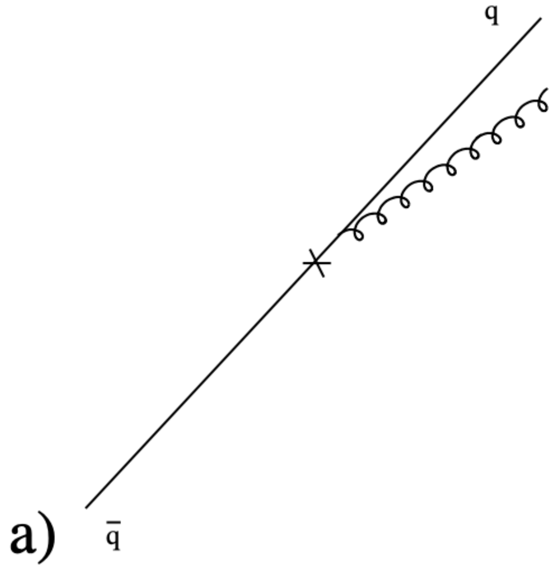


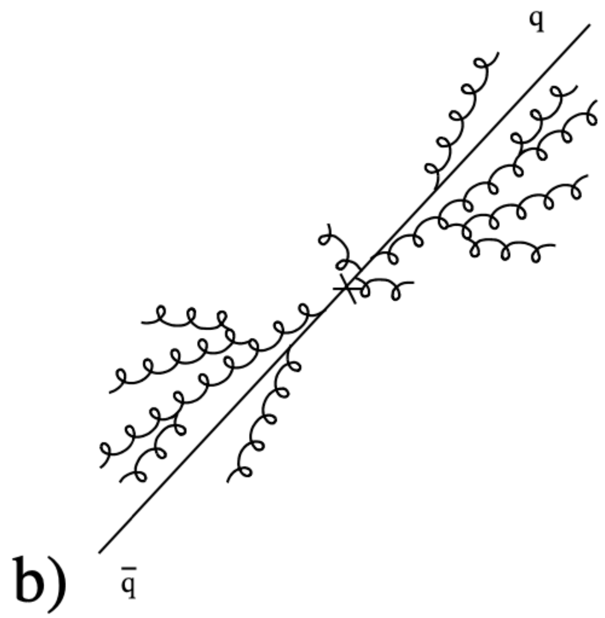
Let's just look into a back-to-back emission

Let's just look into a back-to-back emission

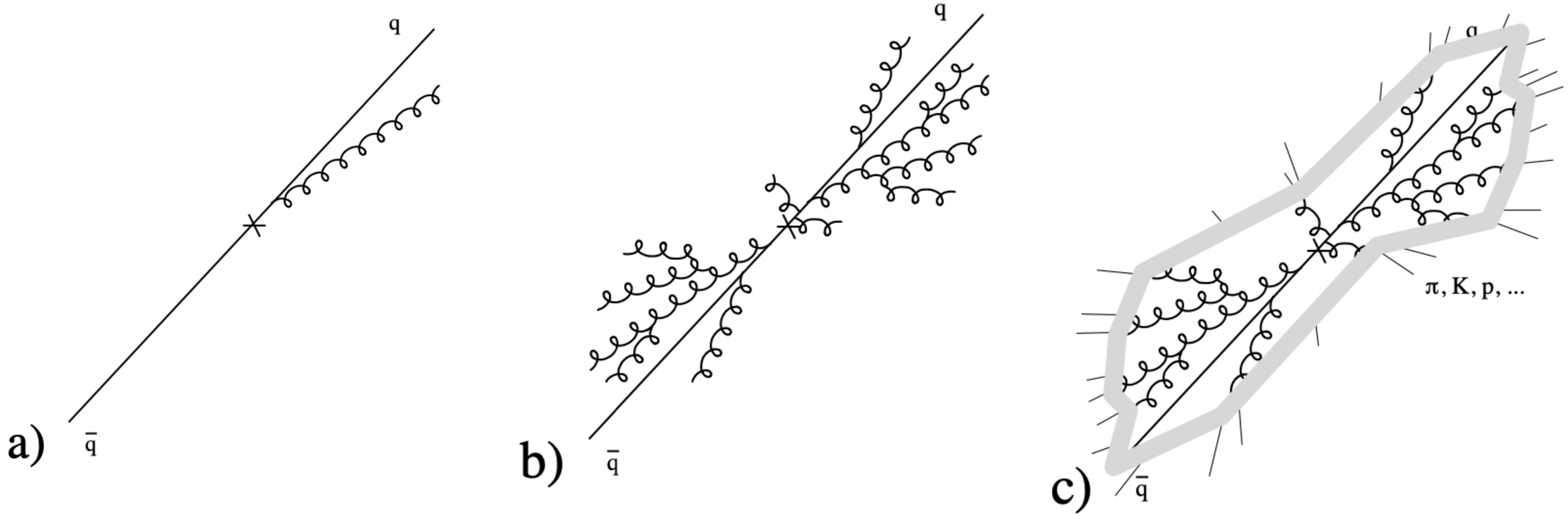


Let’s just look into a back-to-back emission

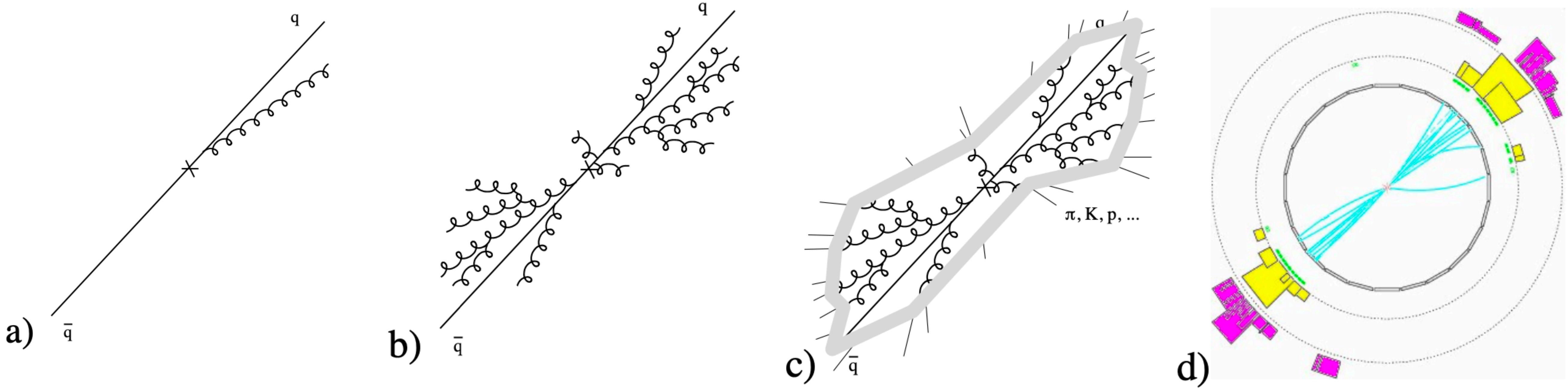




Let's just look into a back-to-back emission

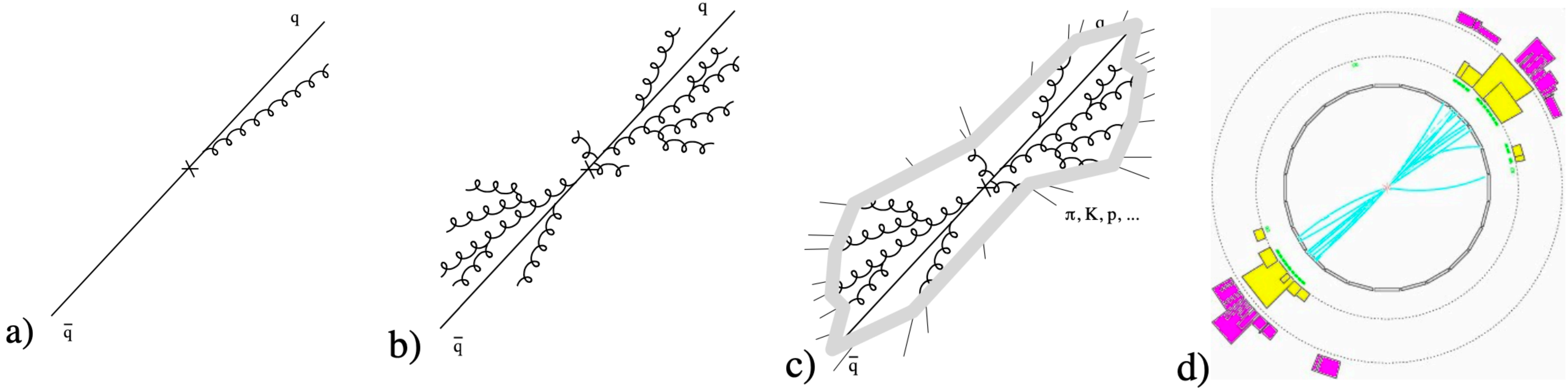


Let's just look into a back-to-back emission



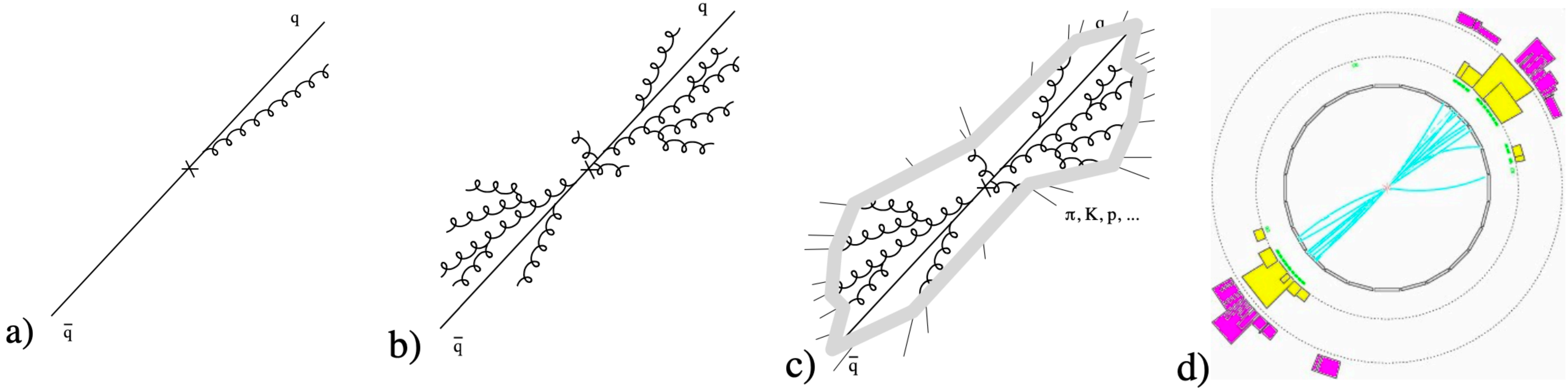
Let’s just look into a back-to-back emission

Figure source : Gavin Salam QCD lectures; 1011.5131



Let's just look into a back-to-back emission

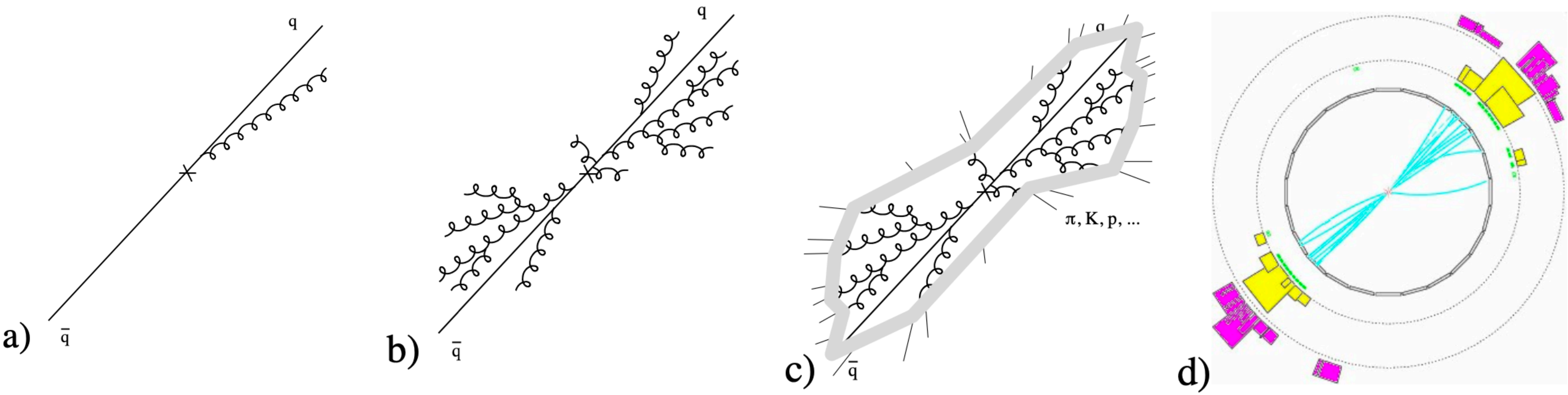
Figure source : Gavin Salam QCD lectures; 1011.5131



The fragmented partons tend to populate in the direction of the mother parton.

Let's just look into a back-to-back emission

Figure source : Gavin Salam QCD lectures; 1011.5131

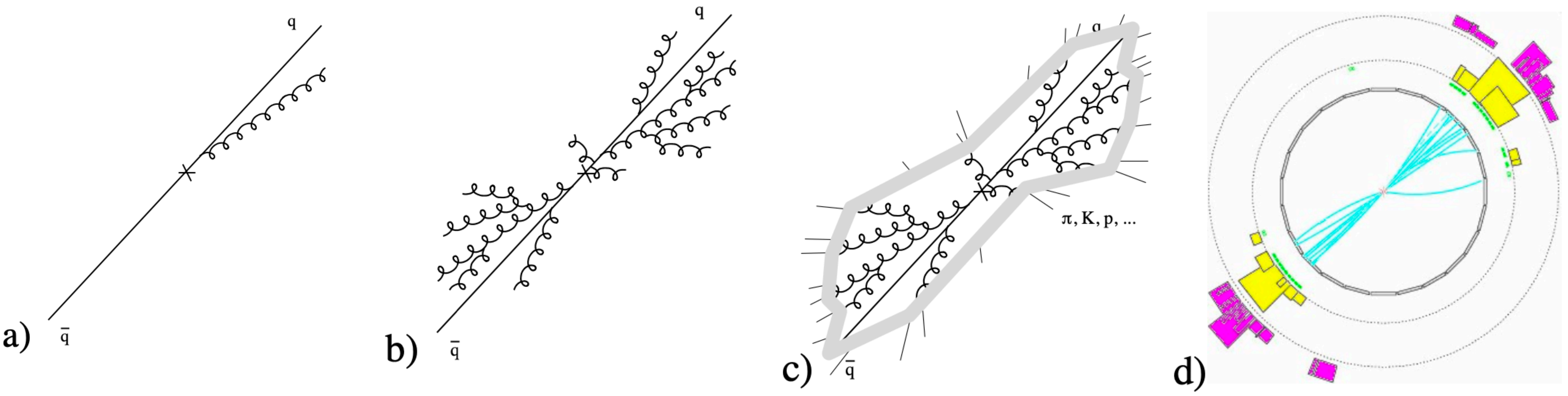


The fragmented partons tend to populate in the direction of the mother parton.

Thus we get a stream of particles, as an end state of fragmentation and hadronization, in the direction of the initial hard parton.

Let's just look into a back-to-back emission

Figure source : Gavin Salam QCD lectures; 1011.5131



The fragmented partons tend to populate in the direction of the mother parton.

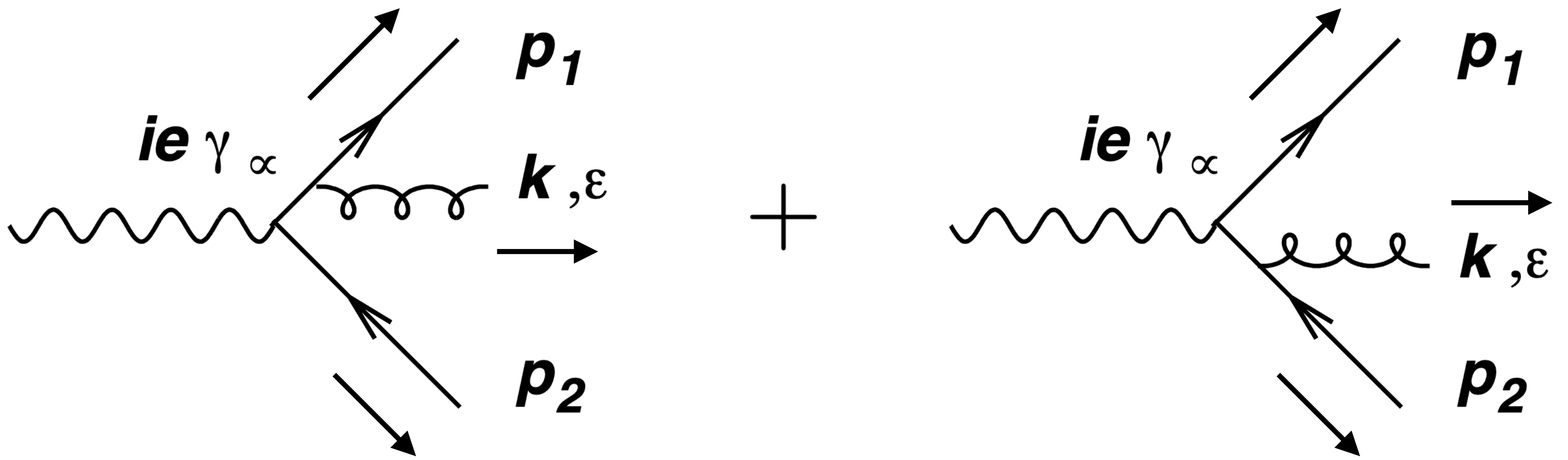
Thus we get a stream of particles, as an end state of fragmentation and hadronization, in the direction of the initial hard parton.

How do we explain this phenomena with our beloved lagrangian :

$$\mathscr{L} = -\frac{1}{4}\sum_{A=1}^{N_C^2-1} F^A_{\mu\nu}F^{A;\mu\nu} + \sum_{a,b=1}^{N_F} \bar{\psi}^a\left[i\gamma^\mu(\partial_\mu\delta_{ab} - g_s\sum_{C=1}^{N_C^2-1} T^C_{ab}A^C_\mu) - m\right]\psi^b$$

Let's do the analysis for a gluon radiation

$$\mathcal{M}_{q\bar{q}g} = $$

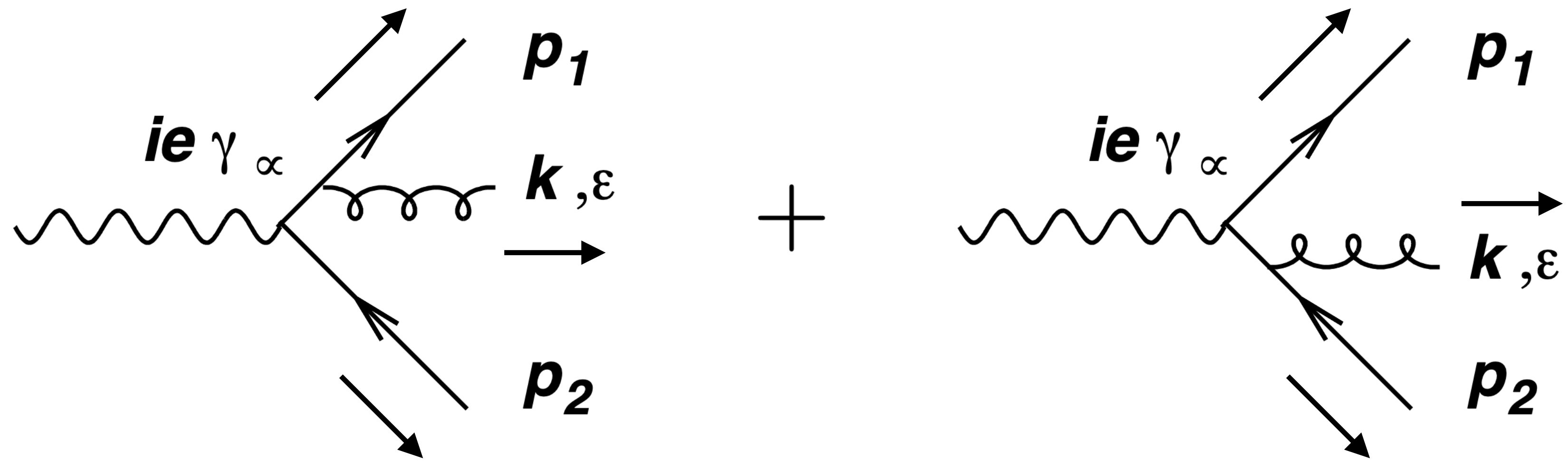


$$i\mathscr{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (ig_s T^a)\ \not{\epsilon}(k) * \left(\frac{i(\not{p}_1 + \not{k} + m)}{(p_1+k)^2 - m^2}\right) (-ie_q\gamma^\alpha)\ v(p_2)$$

$$i\mathscr{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (-ie_q\gamma^\alpha)\ \left(\frac{i(\not{p}_2 + \not{k} + m)}{(p_2+k)^2 - m^2}\right) (ig_s T^a)\ \not{\epsilon}(k) *\ v(p_2)$$

Let’s do the analysis for a gluon radiation

$\mathcal{M}_{q\bar{q}g} =$



$$i\mathcal{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (ig_s T^a)\ \not{\epsilon}(k)^* \left(\frac{i(\not{p}_1 + \not{k} + m)}{(p_1+k)^2 - m^2}\right) (-ie_q\gamma^\alpha)\ v(p_2)$$

$$i\mathcal{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (-ie_q\gamma^\alpha) \left(\frac{i(\not{p}_2 + \not{k} + m)}{(p_2+k)^2 - m^2}\right) (ig_s T^a)\ \not{\epsilon}(k)^*\ v(p_2)$$

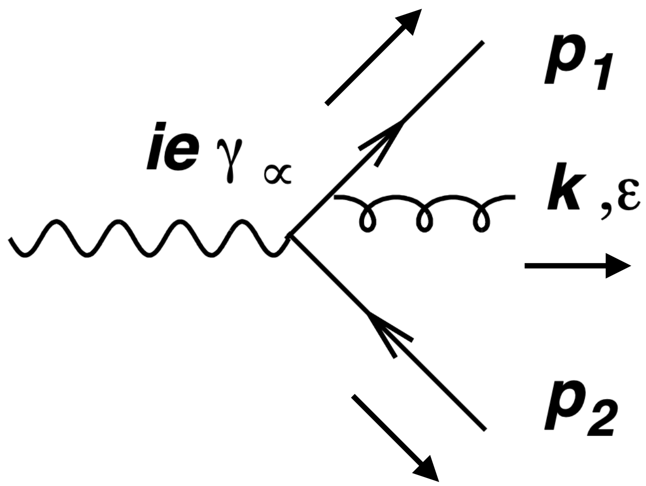
Let’s use the following identities to simplify the above amplitudes :

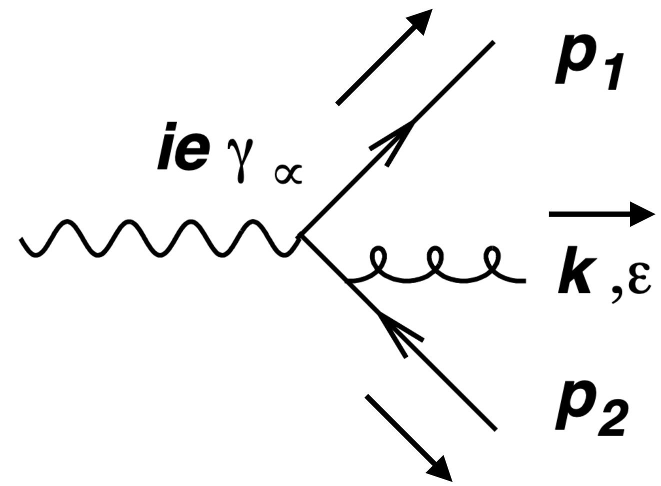
$$\bar{u}(p)\ \not{\epsilon}(k)^* (\not{p} + m) = 2\ \bar{u}(p)\ \epsilon_\mu(k)^*\ p^\mu\ ;\ (\not{p} + m)\ \ \not{\epsilon}(k)^*\ v(p) = 2\ \ \epsilon_\mu(k)^*\ p^\mu\ v(p)$$

$$\bar{u}(p)\ \not{\epsilon}(k)^* (\not{k}) = -\,i\ \bar{u}(p)\ \epsilon_\mu(k)^*\ \sigma^{\mu\nu}\ k_\nu\ ;\ \ \not{k}\ \ \not{\epsilon}(k)^*\ v(p) = -\,i\ \epsilon_\mu(k)^*\ \sigma^{\mu\nu}\ k_\nu\ v(p)$$

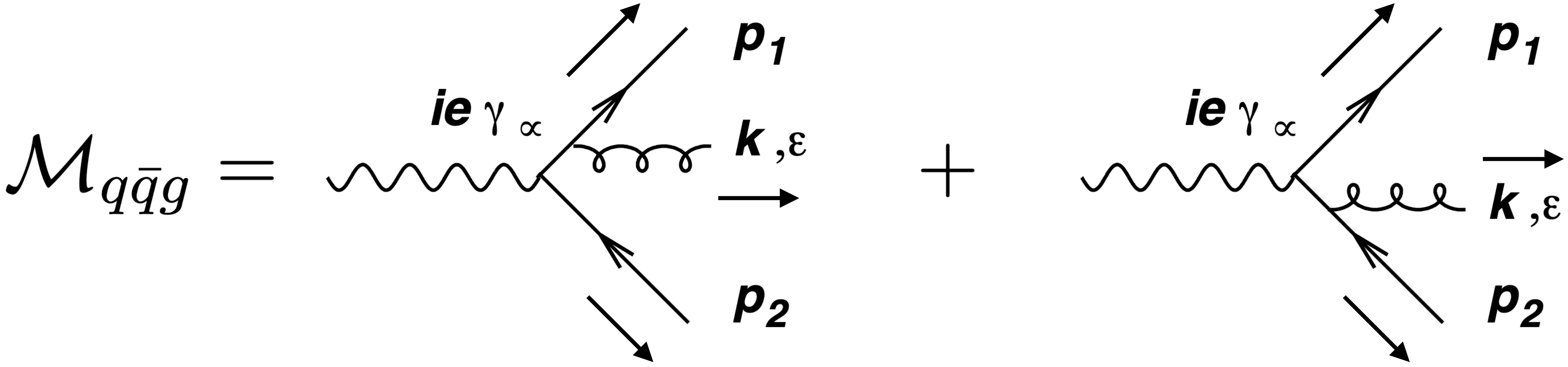
$$(p+k)^2 - m^2 = 2\ (p\,.\,k)$$

Let’s do the analysis for a gluon radiation

$\mathcal{M}_{q\bar{q}g} =$ 

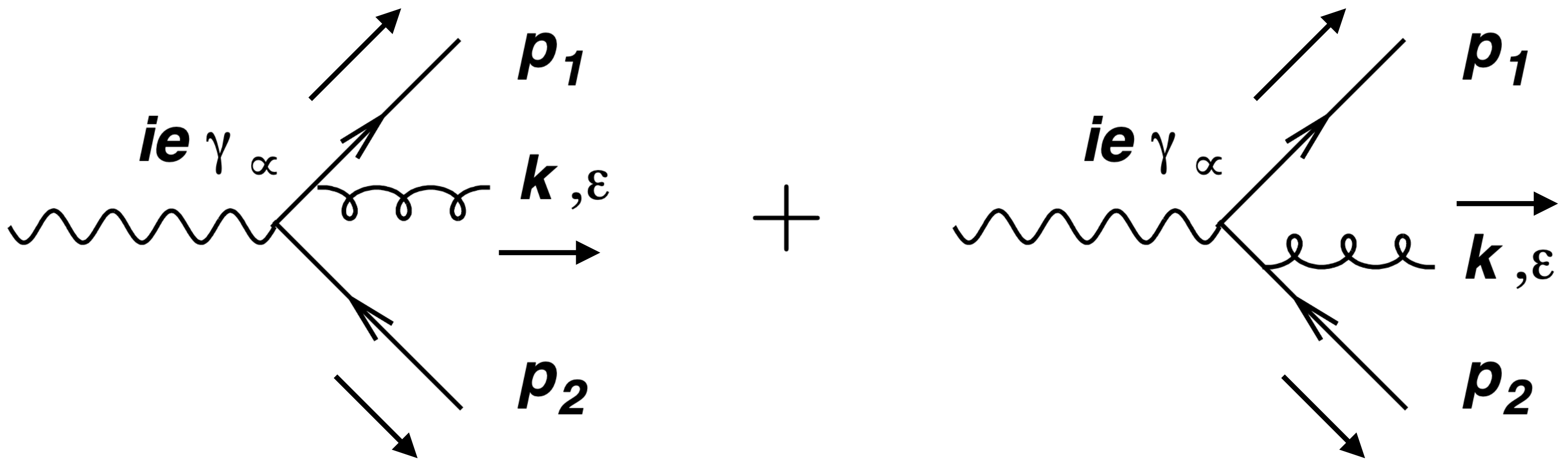
$+$ 

Let's do the analysis for a gluon radiation



$$\mathscr{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k) * \left(\frac{(2\ p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\ p_1 . k}\right)\ (e_q \gamma^\alpha)\ v(p_2)$$

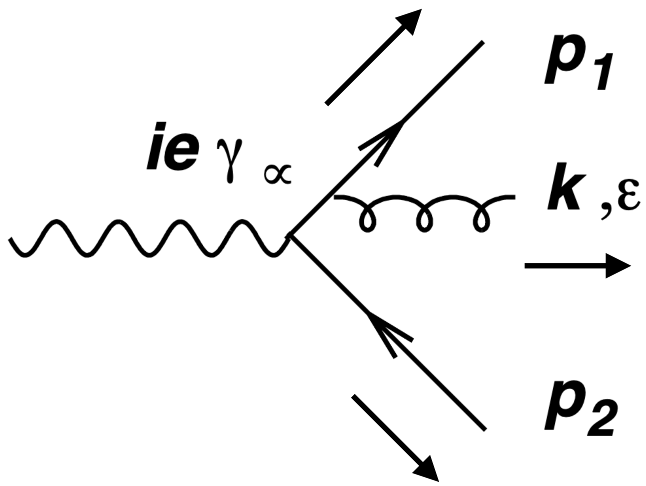
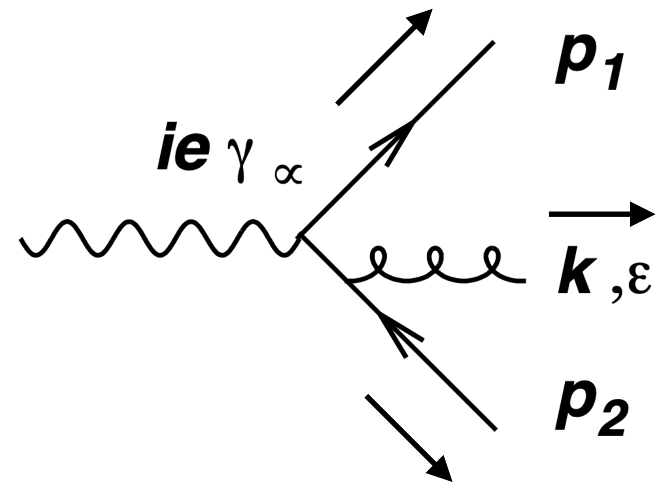
Let's do the analysis for a gluon radiation

$$\mathcal{M}_{q\bar{q}g} =$$ 

$$\mathscr{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k)^* \left(\frac{(2\ p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\ p_1 . k}\right) (e_q \gamma^\alpha)\ v(p_2)$$

$$\mathscr{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (e_q \gamma^\alpha) \left(\frac{(2\ p_2^\mu -\ i\ \sigma^{\mu\nu}\ k_\nu)}{2\ p_2 . k}\right) (g_s T^a)\ \epsilon_\mu(k)^*\ \ v(p_2)$$

Let's do the analysis for a gluon radiation

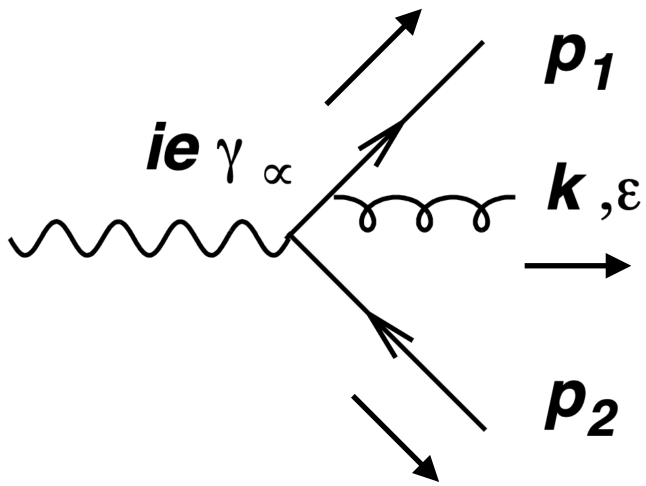
$\mathcal{M}_{q\bar{q}g} =$  $+$ 

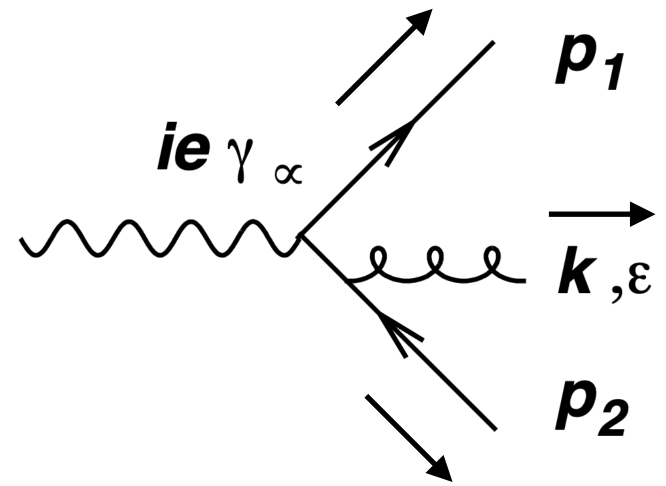
$$\mathcal{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k)^* \left(\frac{(2\ p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\ p_1 . k}\right) (e_q \gamma^\alpha)\ v(p_2)$$

$$\mathcal{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (e_q \gamma^\alpha) \left(\frac{(2\ p_2^\mu -\ i\ \sigma^{\mu\nu}\ k_\nu)}{2\ p_2 . k}\right) (g_s T^a)\ \epsilon_\mu(k)^*\ \ v(p_2)$$

In the limit $k \to 0$:

Let's do the analysis for a gluon radiation

$\mathcal{M}_{q\bar{q}g} =$ 

$+$ 

$$\mathcal{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k)^*\left(\frac{(2\,p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\,p_1\,.\,k}\right)(e_q\gamma^\alpha)\ v(p_2)$$

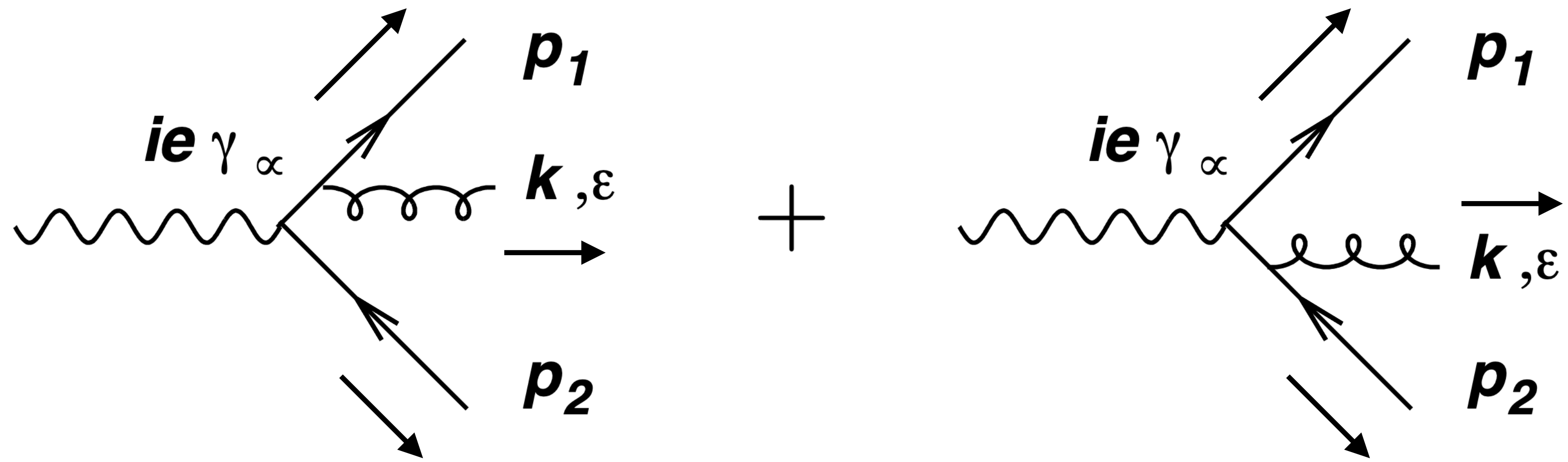
$$\mathcal{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (e_q\gamma^\alpha)\left(\frac{(2\,p_2^\mu -\ i\ \sigma^{\mu\nu}\ k_\nu)}{2\,p_2\,.\,k}\right)(g_s T^a)\ \epsilon_\mu(k)^*\ \ v(p_2)$$

In the limit $k \to 0$:

$$\mathcal{M}_{q\bar{q}g} = \bar{u}(p_1)\ (e_q\gamma^\alpha)\ (g_s T^a)\ v(p_2)\left(\frac{p_1\,.\,\epsilon_a(k)^*}{p_1\,.\,k} - \frac{p_2\,.\,\epsilon_a(k)^*}{p_2\,.\,k}\right)$$

Let's do the analysis for a gluon radiation

$\mathcal{M}_{q\bar{q}g} =$



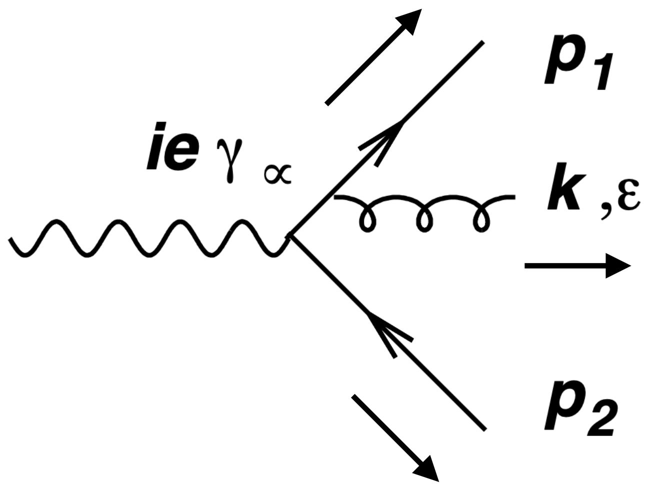
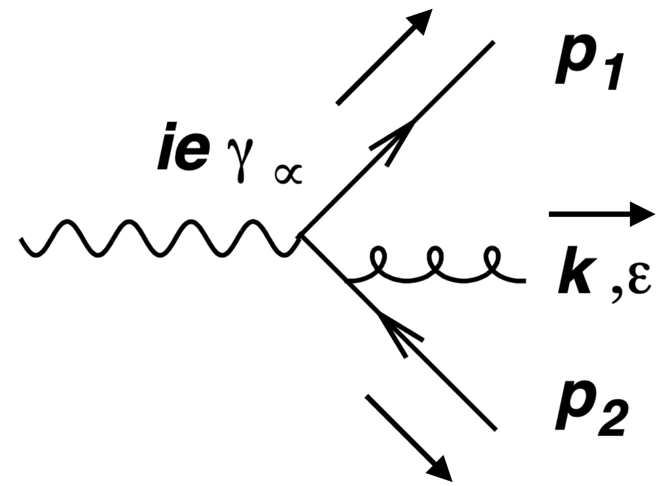
$$\mathscr{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k)^*\left(\frac{(2\,p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\,p_1 \,.\, k}\right)(e_q\gamma^\alpha)\ v(p_2)$$

$$\mathscr{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (e_q\gamma^\alpha)\left(\frac{(2\,p_2^\mu -\ i\ \sigma^{\mu\nu}\ k_\nu)}{2\,p_2 \,.\, k}\right)(g_s T^a)\ \epsilon_\mu(k)^*\ \ v(p_2)$$

In the limit $k \to 0$:
$$\mathscr{M}_{q\bar{q}g} = \bar{u}(p_1)\ (e_q\gamma^\alpha)\ (g_s T^a)\ v(p_2)\left(\frac{p_1 \,.\, \epsilon_a(k)^*}{p_1 \,.\, k} - \frac{p_2 \,.\, \epsilon_a(k)^*}{p_2 \,.\, k}\right)$$

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g} = -\ \bar{\mathscr{M}}^2{}_{q\bar{q}}\left(g_s^2\sum_a Tr_F\ T^aT^a\right)\left(\frac{p_1}{p_1 \,.\, k} - \frac{p_2}{p_2 \,.\, k}\right).\left(\frac{p_1}{p_1 \,.\, k} - \frac{p_2}{p_2 \,.\, k}\right)$$

Let’s do the analysis for a gluon radiation

$\mathcal{M}_{q\bar{q}g} =$  $+$ 

$$\mathcal{M}^1_{q\bar{q}g} = \bar{u}(p_1)\ (g_s T^a)\ \epsilon_\mu(k)^*\left(\frac{(2\,p_1^\mu - i\sigma^{\mu\nu}\ k_\nu)}{2\,p_1\,.\,k}\right)(e_q\gamma^\alpha)\ v(p_2)$$

$$\mathcal{M}^2_{q\bar{q}g} = -\ \bar{u}(p_1)\ (e_q\gamma^\alpha)\left(\frac{(2\,p_2^\mu -\ i\ \sigma^{\mu\nu}\ k_\nu)}{2\,p_2\,.\,k}\right)(g_s T^a)\ \epsilon_\mu(k)^*\ \ v(p_2)$$

In the limit $k \to 0$:

$$\mathcal{M}_{q\bar{q}g} = \bar{u}(p_1)\ (e_q\gamma^\alpha)\ (g_s T^a)\ v(p_2)\left(\frac{p_1\,.\,\epsilon_a(k)^*}{p_1\,.\,k} - \frac{p_2\,.\,\epsilon_a(k)^*}{p_2\,.\,k}\right)$$

$$\bar{\mathcal{M}}^2{}_{q\bar{q}g} = -\ \bar{\mathcal{M}}^2{}_{q\bar{q}}\left(g_s^2\sum_a Tr_F\ T^aT^a\right)\left(\frac{p_1}{p_1\,.\,k} - \frac{p_2}{p_2\,.\,k}\right).\left(\frac{p_1}{p_1\,.\,k} - \frac{p_2}{p_2\,.\,k}\right)$$

$$= \ \bar{\mathcal{M}}^2{}_{q\bar{q}}\ g_s^2\ N_cC_F\left[2\,\frac{p_1\,.\,p_2}{(p_1\,.\,k)\ (p_2\,.\,k)} - \frac{m^2}{(p_2\,.\,k)^2} - \frac{m^2}{(p_1\,.\,k)^2}\right],\ N_cC_F = \frac{1}{2}(N_C^2 - 1)$$

Let’s do the analysis for a gluon radiation

Let's do the analysis for a gluon radiation

For computing the x-section of the emission we need the 3-body phase space :

$$d\,\Phi_{q\bar{q}g} \approx d\,\Phi_{q\bar{q}}\,\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{2k^0}\,.$$

Let's do the analysis for a gluon radiation

For computing the x-section of the emission we need the 3-body phase space :

$$d\,\Phi_{q\bar{q}g} \approx d\,\Phi_{q\bar{q}}\,\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{2k^0}\,.$$

Hence the x-section of soft gluon emission, with $k^0 = E$:

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g}\,d\,\Phi_{q\bar{q}g} \approx \bar{\mathscr{M}}^2{}_{q\bar{q}}\,d\,\Phi_{q\bar{q}}\,\frac{2\alpha_S\,C_F}{\pi}\,E\,dE\,d(cos\theta)\,\frac{d\phi}{2\pi}\left(\frac{2\,p_1\,.\,p_2}{(p_1\,.\,k)\,(p_2\,.\,k)}\right).$$

Let's do the analysis for a gluon radiation

For computing the x-section of the emission we need the 3-body phase space :

$$d\,\Phi_{q\bar{q}g} \approx d\,\Phi_{q\bar{q}}\,\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{2k^0}\,.$$

Hence the x-section of soft gluon emission, with $k^0 = E$:

$$\bar{\mathcal{M}}^2{}_{q\bar{q}g}\,d\,\Phi_{q\bar{q}g} \approx \bar{\mathcal{M}}^2{}_{q\bar{q}}\,d\,\Phi_{q\bar{q}}\,\frac{2\alpha_S\,C_F}{\pi}\,E\,dE\,d(cos\theta)\,\frac{d\phi}{2\pi}\,\left(\frac{2\,p_1\,.\,p_2}{(p_1\,.\,k)\,(p_2\,.\,k)}\right)\,.$$

If we try to carry out the P.S. integral in $q\bar{q}$ COM frame :

$$\bar{\mathcal{M}}^2{}_{q\bar{q}g}\,d\,\Phi_{q\bar{q}g} \approx \bar{\mathcal{M}}^2{}_{q\bar{q}}\,d\,\Phi_{q\bar{q}}\,\frac{2\alpha_S\,C_F}{\pi}\,\frac{dE}{E}\,\frac{d\theta}{sin\theta}\,\frac{d\phi}{2\pi}\,.$$

Let's do the analysis for a gluon radiation

For computing the x-section of the emission we need the 3-body phase space :

$$d\ \Phi_{q\bar{q}g} \approx d\ \Phi_{q\bar{q}}\ \frac{d^3\vec{k}}{(2\pi)^3}\ \frac{1}{2k^0}\ .$$

Hence the x-section of soft gluon emission, with $k^0 = E$:

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g}\ d\ \Phi_{q\bar{q}g} \approx \bar{\mathscr{M}}^2{}_{q\bar{q}}\ d\ \Phi_{q\bar{q}}\ \frac{2\alpha_S\ C_F}{\pi}\ E\ dE\ d(cos\theta)\ \frac{d\phi}{2\pi}\ \left(\frac{2\ p_1 . p_2}{(p_1 . k)\ (p_2 . k)}\right) .$$

If we try to carry out the P.S. integral in $q\bar{q}$ COM frame :

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g}\ d\ \Phi_{q\bar{q}g} \approx \bar{\mathscr{M}}^2{}_{q\bar{q}}\ d\ \Phi_{q\bar{q}}\ \frac{2\alpha_S\ C_F}{\pi}\ \frac{dE}{E}\ \frac{d\theta}{sin\theta}\ \frac{d\phi}{2\pi}\ .$$

When a gluon is emitted from a quark line, in collinear limit, the additional piece in the x-section :

$$dS = \frac{2\alpha_S\ C_F}{\pi}\ \frac{dE}{E}\ \frac{d\theta}{sin\theta}\ \frac{d\phi}{2\pi} \approx \frac{2\alpha_S\ C_F}{\pi}\ \frac{dE}{E}\ \frac{d\theta}{\theta}\ .$$

Let's do the analysis for a gluon radiation

For computing the x-section of the emission we need the 3-body phase space :

$$d\,\Phi_{q\bar{q}g} \approx d\,\Phi_{q\bar{q}}\,\frac{d^3\vec{k}}{(2\pi)^3}\,\frac{1}{2k^0}\,.$$

Hence the x-section of soft gluon emission, with $k^0 = E$:

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g}\,d\,\Phi_{q\bar{q}g} \approx \bar{\mathscr{M}}^2{}_{q\bar{q}}\,d\,\Phi_{q\bar{q}}\,\frac{2\alpha_S\,C_F}{\pi}\,E\,dE\,d(cos\theta)\,\frac{d\phi}{2\pi}\left(\frac{2\,p_1\,.\,p_2}{(p_1\,.\,k)\,(p_2\,.\,k)}\right).$$

If we try to carry out the P.S. integral in $q\bar{q}$ COM frame :

$$\bar{\mathscr{M}}^2{}_{q\bar{q}g}\,d\,\Phi_{q\bar{q}g} \approx \bar{\mathscr{M}}^2{}_{q\bar{q}}\,d\,\Phi_{q\bar{q}}\,\frac{2\alpha_S\,C_F}{\pi}\,\frac{dE}{E}\,\frac{d\theta}{sin\theta}\,\frac{d\phi}{2\pi}\,.$$

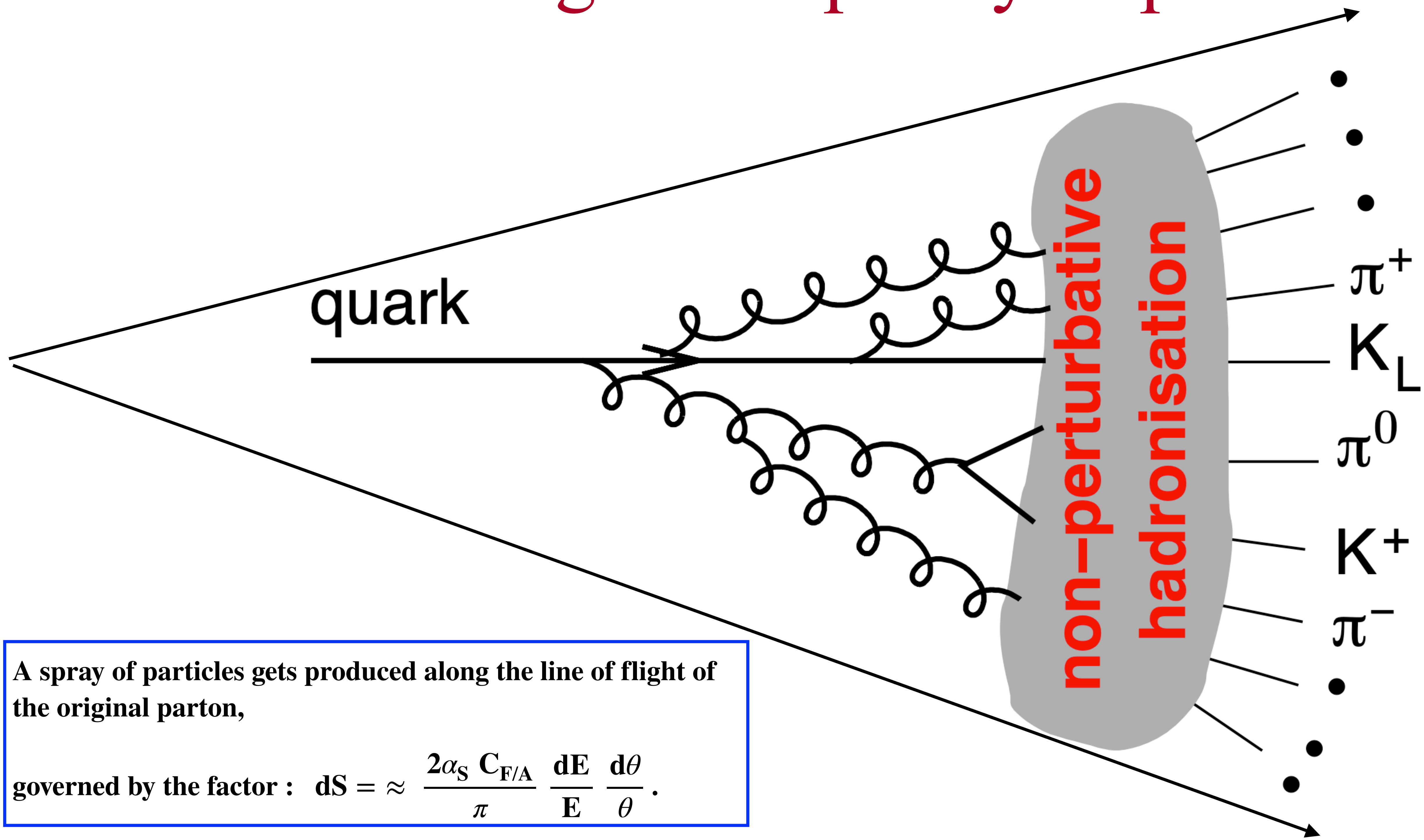
When a gluon is emitted from a quark line, in collinear limit, the additional piece in the x-section :

$$dS = \frac{2\alpha_S\,C_F}{\pi}\,\frac{dE}{E}\,\frac{d\theta}{sin\theta}\,\frac{d\phi}{2\pi} \approx \frac{2\alpha_S\,C_F}{\pi}\,\frac{dE}{E}\,\frac{d\theta}{\theta}\,.$$

When a gluon is emitted from a gluon line, in collinear limit,

the additional piece in the x-section : $dS = \approx \dfrac{2\alpha_S\,C_A}{\pi}\,\dfrac{dE}{E}\,\dfrac{d\theta}{\theta}\,.$

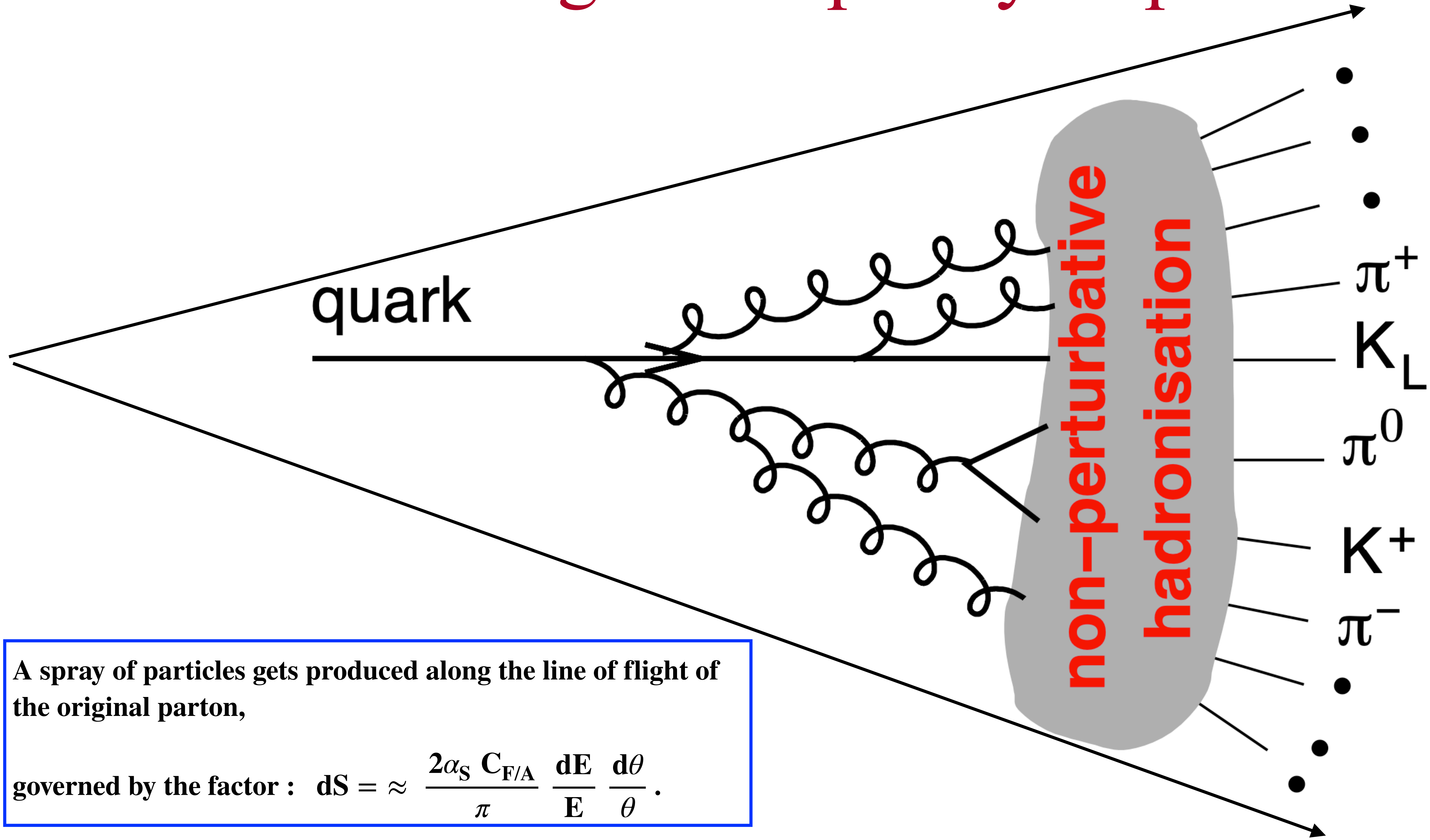
So we have a large multiplicity of particles



A spray of particles gets produced along the line of flight of the original parton,

governed by the factor : $\mathbf{dS} = \approx \frac{2\alpha_S \, C_{F/A}}{\pi} \frac{dE}{E} \frac{d\theta}{\theta}$.

So we have a large multiplicity of particles

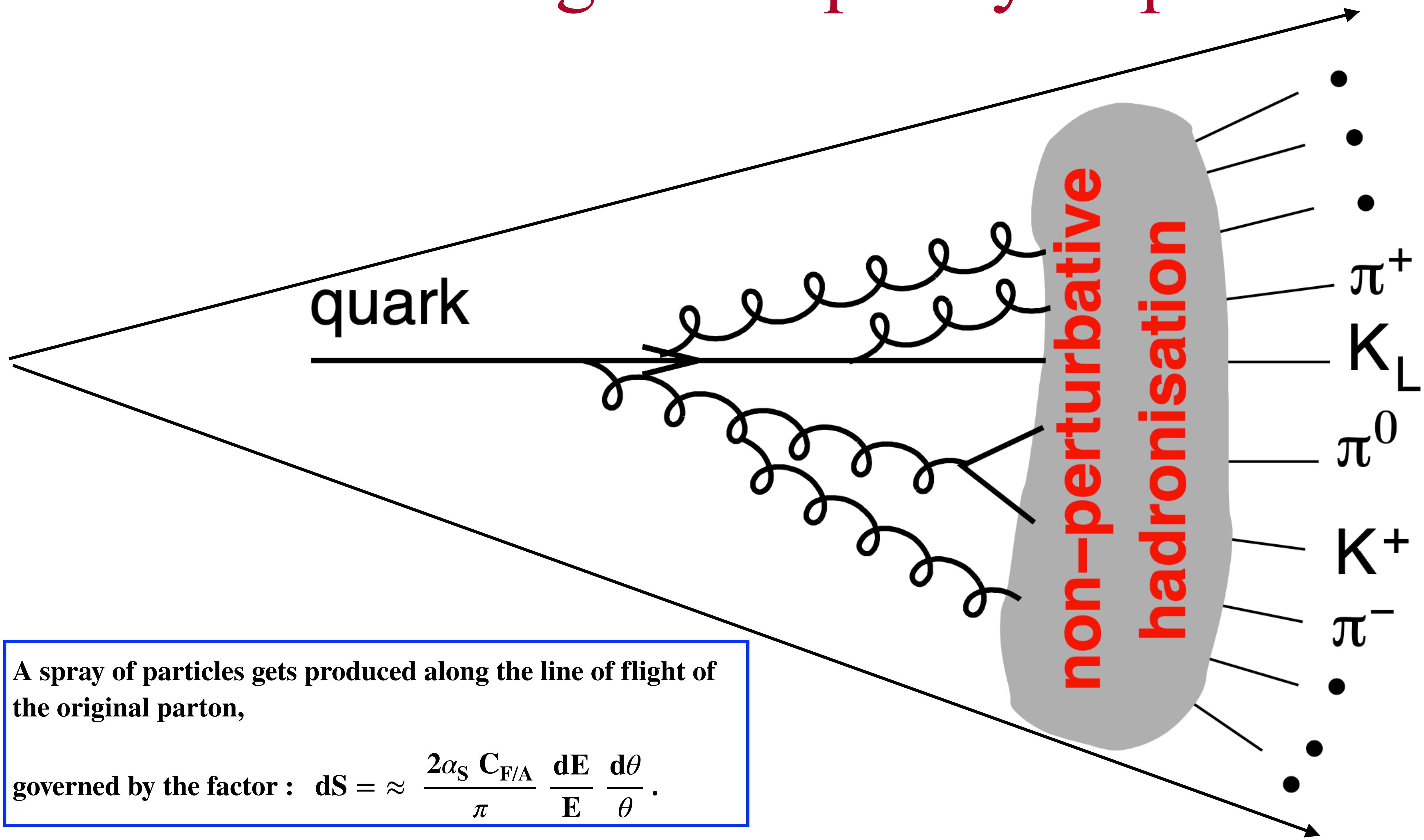


A spray of particles gets produced along the line of flight of the original parton,

governed by the factor : $\mathbf{dS} = \approx \frac{2\alpha_S\, C_{F/A}}{\pi} \frac{dE}{E} \frac{d\theta}{\theta}$.

$$\sigma_{tot} = \sigma_{q\bar{q}} \left(1 + \frac{2\alpha_s C_F}{\pi} \int \frac{dE}{E} \int \frac{d\theta}{\sin\theta} R\left(\frac{E}{Q}, \theta\right) - \frac{2\alpha_s C_F}{\pi} \int \frac{dE}{E} \int \frac{d\theta}{\sin\theta} V\left(\frac{E}{Q}, \theta\right)\right)$$

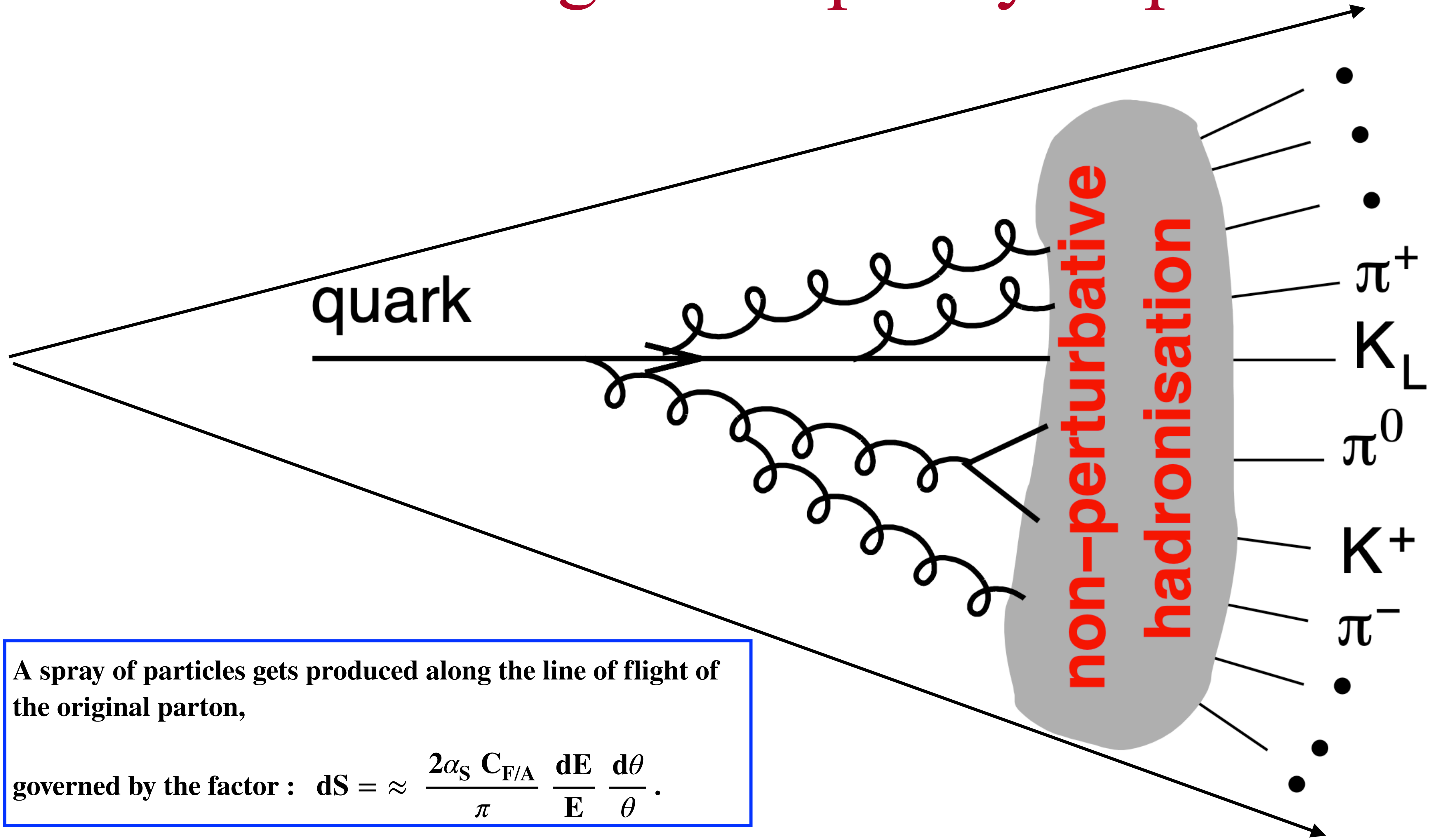
So we have a large multiplicity of particles



A spray of particles gets produced along the line of flight of the original parton,

governed by the factor : $\mathbf{dS} = \approx \frac{2\alpha_S\, C_{F/A}}{\pi} \frac{dE}{E} \frac{d\theta}{\theta}$.

So we have a large multiplicity of particles

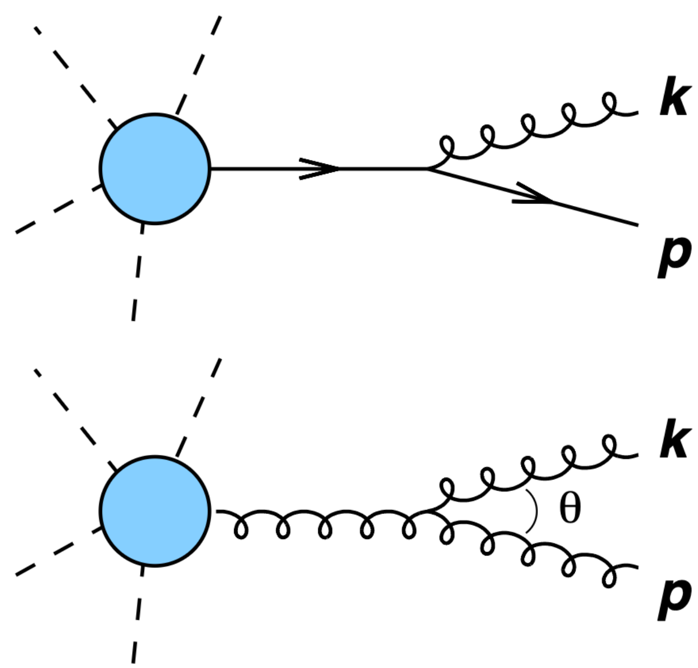


A spray of particles gets produced along the line of flight of the original parton,

governed by the factor : $\mathbf{dS} = \approx \frac{2\alpha_S\, C_{F/A}}{\pi} \frac{\mathbf{dE}}{\mathbf{E}} \frac{\mathbf{d}\theta}{\theta}$.

$$\sigma_{tot} = \sigma_{q\bar{q}} \left(1 + 1.045 \frac{\alpha_s(Q)}{\pi} + 0.94 \left(\frac{\alpha_s(Q)}{\pi} \right)^2 - 15 \left(\frac{\alpha_s(Q)}{\pi} \right)^3 + \cdots \right)$$

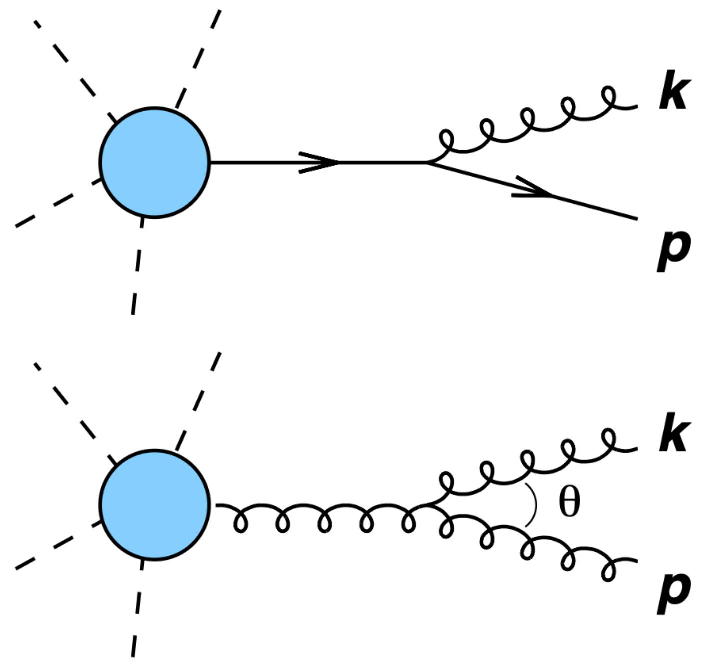
Shower multiplicity : basic understanding



$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

Shower multiplicity : basic understanding

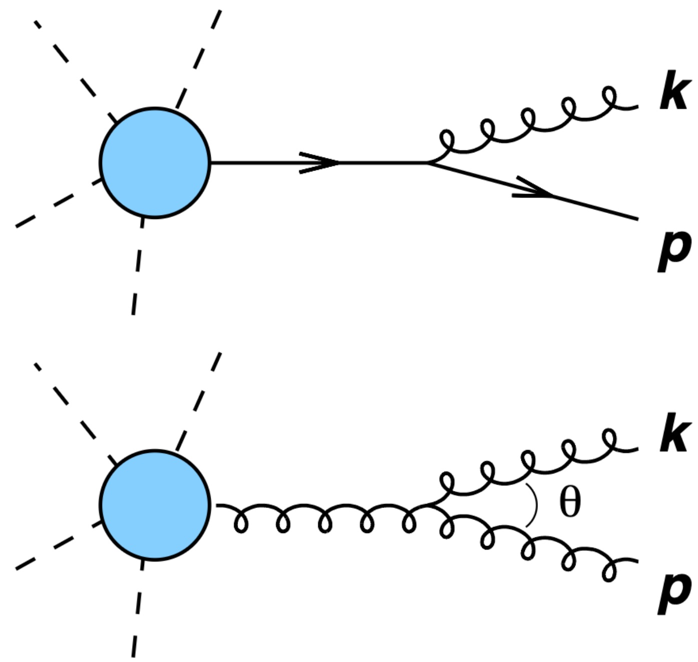


$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi}\int_{Q_0}^{Q}\frac{dE}{E}\int_{Q_0/E}^{\pi/2}\frac{d\theta}{\theta} \quad .$$

Shower multiplicity : basic understanding



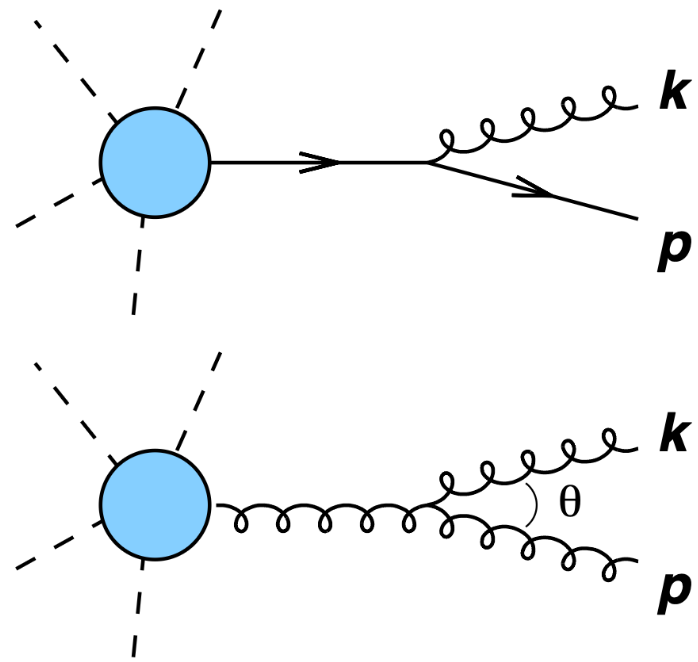
$$\simeq \frac{2\alpha_s C_F}{\pi} \frac{dE}{E} \frac{d\theta}{\theta}$$

$$\simeq \frac{2\alpha_s C_A}{\pi} \frac{dE}{E} \frac{d\theta}{\theta}$$

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi} \int_{Q_0}^{Q} \frac{dE}{E} \int_{Q_0/E}^{\pi/2} \frac{d\theta}{\theta} \; .$$

$$\langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi} \; ln^2(Q/Q_0) + \mathcal{O}\Big(\alpha_s \; ln(Q/Q_0)\Big) \, .$$

Shower multiplicity : basic understanding



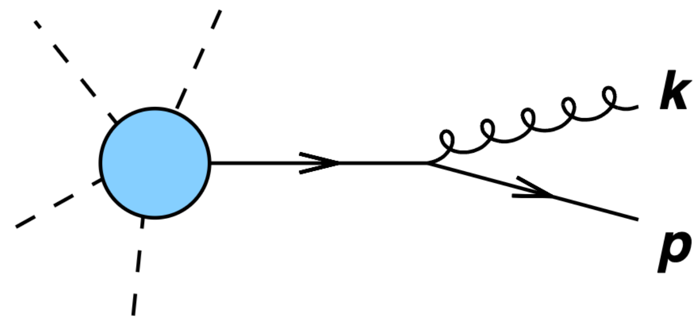
$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

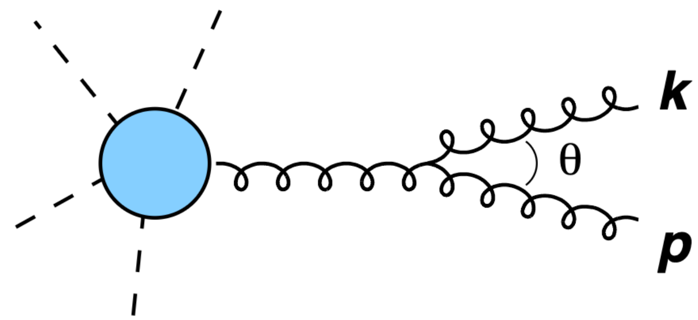
$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi}\int_{Q_0}^{Q}\frac{dE}{E}\int_{Q_0/E}^{\pi/2}\frac{d\theta}{\theta} . \qquad \langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi}\, ln^2(Q/Q_0) + \mathcal{O}\Big(\alpha_s\, ln(Q/Q_0)\Big) .$$

Now let's be crude and plug $\alpha_s = \alpha_s(Q) = \Big(2b_0\, ln(Q/\Lambda)\Big)^{-1}$, $b_0 = \dfrac{11C_A - 2N_f}{12\pi}$, *with* $C_A = N$.

Shower multiplicity : basic understanding



$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

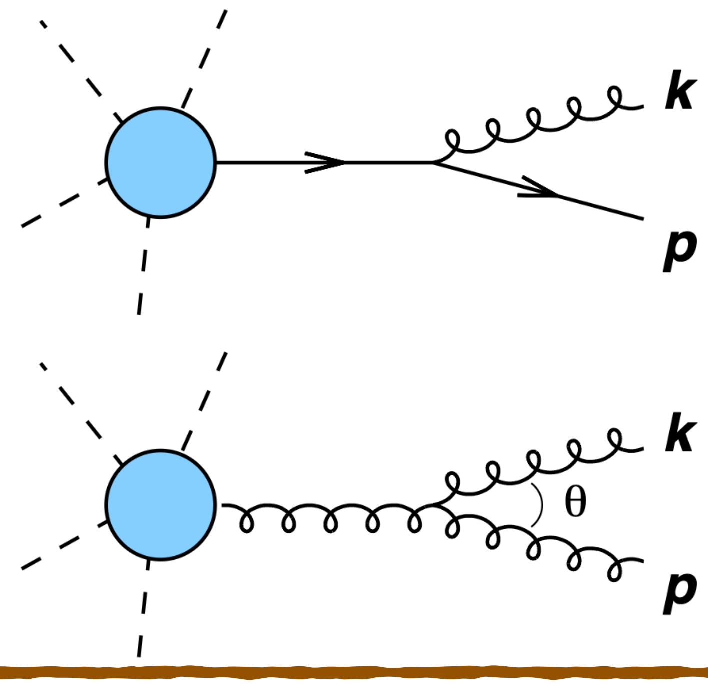


$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi}\int_{Q_0}^{Q}\frac{dE}{E}\int_{Q_0/E}^{\pi/2}\frac{d\theta}{\theta} \,. \qquad \langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi}\, ln^2(Q/Q_0) + \mathcal{O}\Big(\alpha_s \, ln(Q/Q_0)\Big) \,.$$

Now let's be crude and plug $\alpha_s = \alpha_s(Q) = \Big(2b_0 \, ln(Q/\Lambda)\Big)^{-1}$, $b_0 = \dfrac{11C_A - 2N_f}{12\pi}$, *with* $C_A = N$.

Shower multiplicity : basic understanding



$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

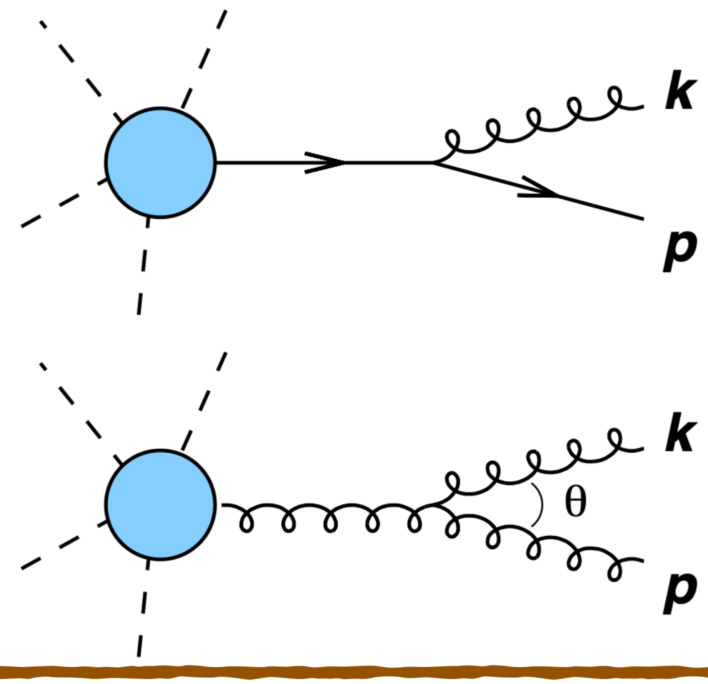
$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi}\int_{Q_0}^{Q}\frac{dE}{E}\int_{Q_0/E}^{\pi/2}\frac{d\theta}{\theta}\,. \qquad \langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi}\, ln^2(Q/Q_0) + \mathcal{O}\Big(\alpha_s\, ln(Q/Q_0)\Big)\,.$$

Now let's be crude and plug $\alpha_s = \alpha_s(Q) = \Big(2b_0\, ln(Q/\Lambda)\Big)^{-1}$, $b_0 = \dfrac{11C_A - 2N_f}{12\pi}$, *with* $C_A = N$.

$$\langle N_g \rangle \simeq \frac{C_F}{2b_0\pi} ln(Q/\Lambda) \simeq \frac{C_F}{4\ b_0^2\ \pi\ \alpha_s(Q)}$$

Shower multiplicity : basic understanding



$$\simeq \frac{2\alpha_s C_F}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\simeq \frac{2\alpha_s C_A}{\pi}\frac{dE}{E}\frac{d\theta}{\theta}$$

$$\langle N_g \rangle \simeq \frac{2\alpha_s C_F}{\pi}\int_{Q_0}^{Q}\frac{dE}{E}\int_{Q_0/E}^{\pi/2}\frac{d\theta}{\theta} \; . \qquad \langle N_g \rangle \simeq \frac{\alpha_s C_F}{\pi}\, ln^2(Q/Q_0) + \mathcal{O}\Big(\alpha_s\, ln(Q/Q_0)\Big) \, .$$

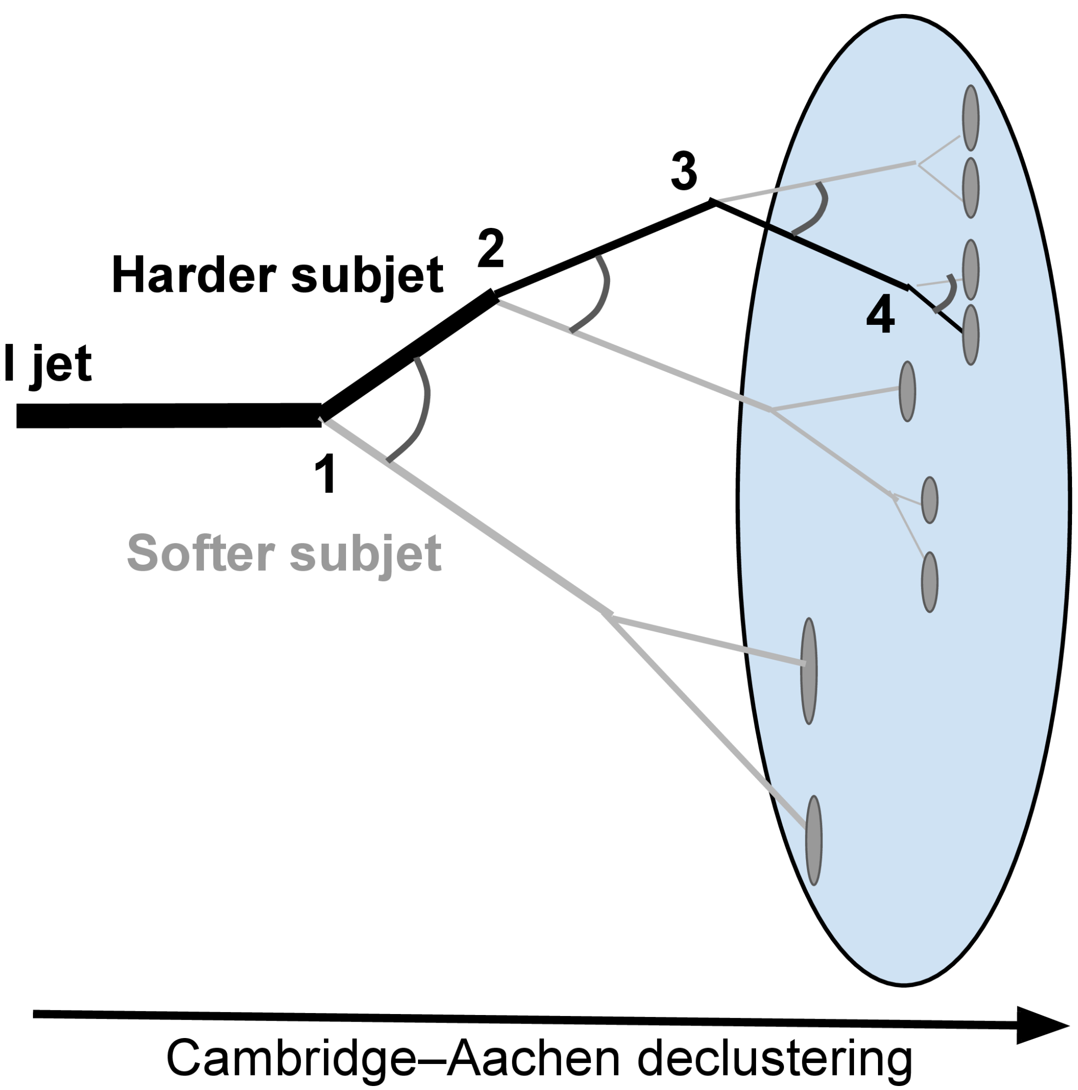
Now let's be crude and plug $\alpha_s = \alpha_s(Q) = \Big(2b_0\, ln(Q/\Lambda)\Big)^{-1}$, $b_0 = \dfrac{11C_A - 2N_f}{12\pi}$, *with* $C_A = N$.

$$\langle N_g \rangle \simeq \frac{C_F}{2b_0\pi} ln(Q/\Lambda) \simeq \frac{C_F}{4\; b_0^2\; \pi\; \alpha_s(Q)}$$

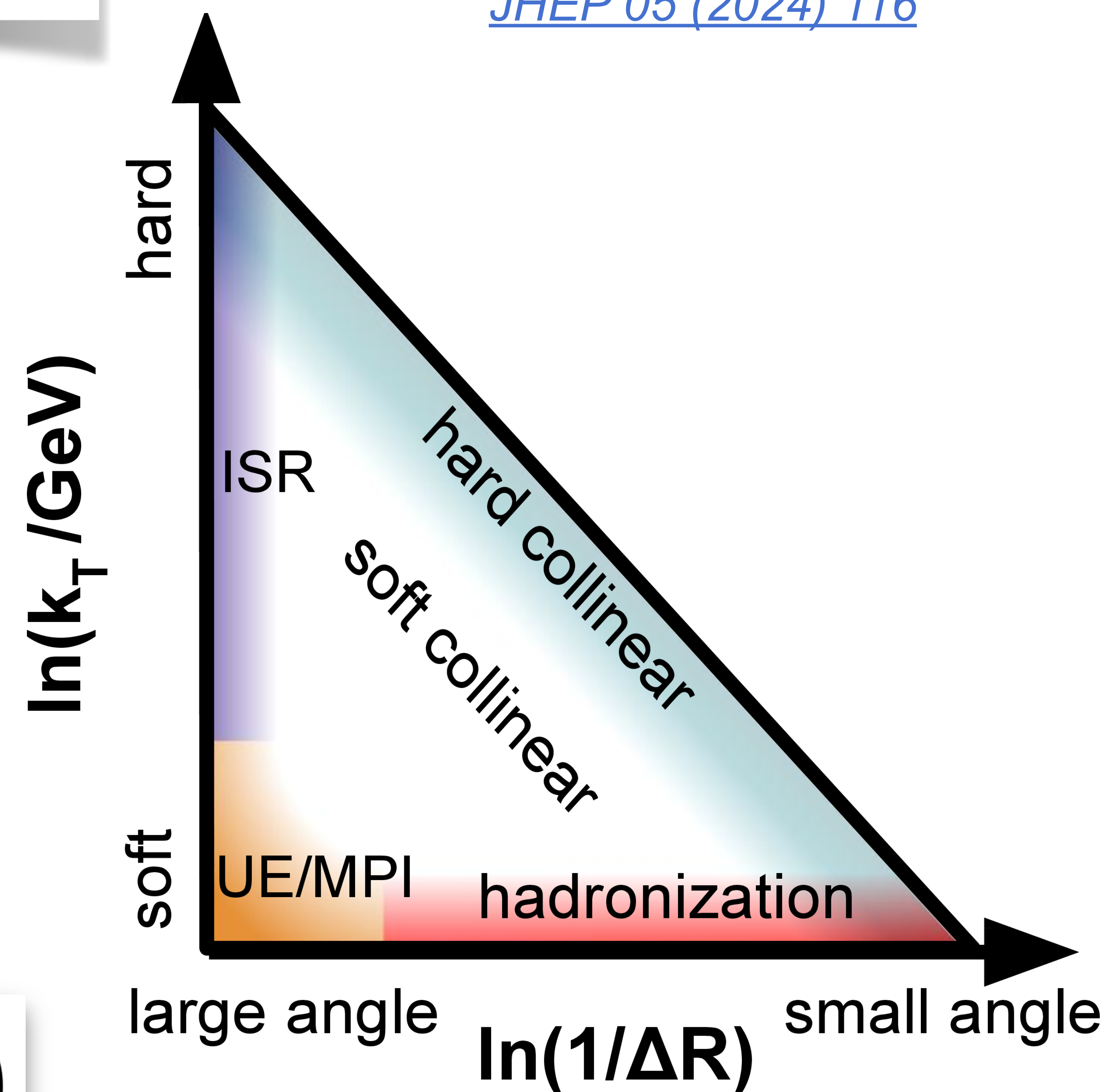
Jets have mass

Lund plane measurements

$$\rho(k_{\mathrm{T}}, \Delta R) \equiv \frac{1}{N_{\mathrm{jets}}} \frac{\mathrm{d}^2 N_{\mathrm{emissions}}}{\mathrm{d}\ln(k_{\mathrm{T}}/\,\mathrm{GeV})\,\mathrm{d}\ln(R/\Delta R)},$$



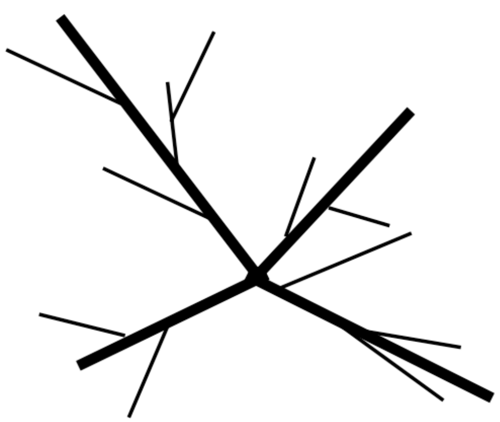
[JHEP 05 (2024) 116]



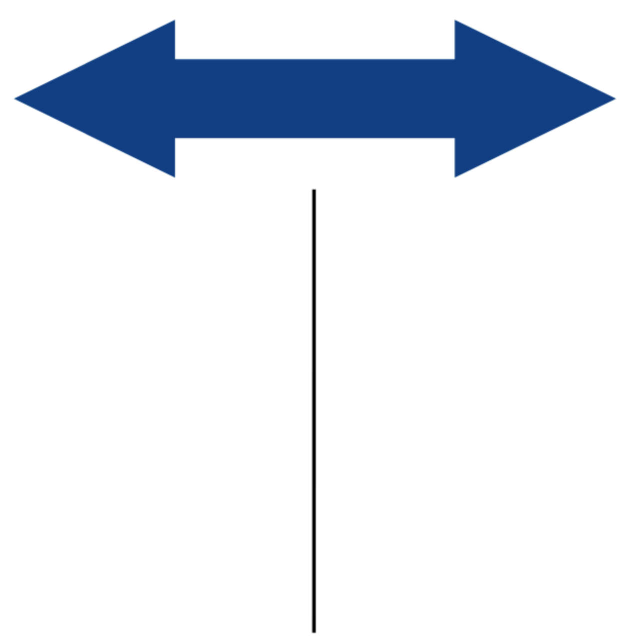
$$P_{q\to qg} = \frac{2\alpha_s C_F}{\pi}\frac{dz}{z}\frac{d\theta}{\theta} = \frac{2\alpha_s C_F}{\pi}\, d\left(\log\frac{1}{z}\right) d\left(\log\frac{1}{\theta}\right)$$

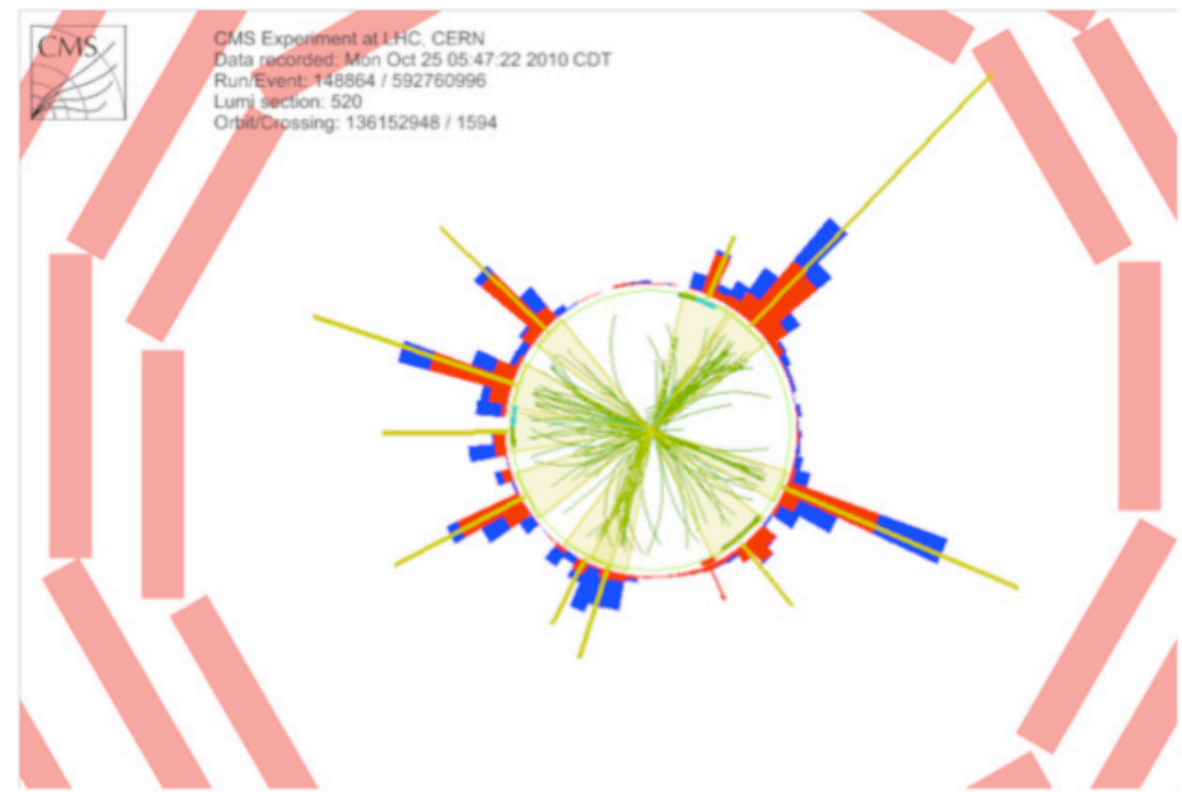
Particles to Parton mapping

Multileg + PS



??



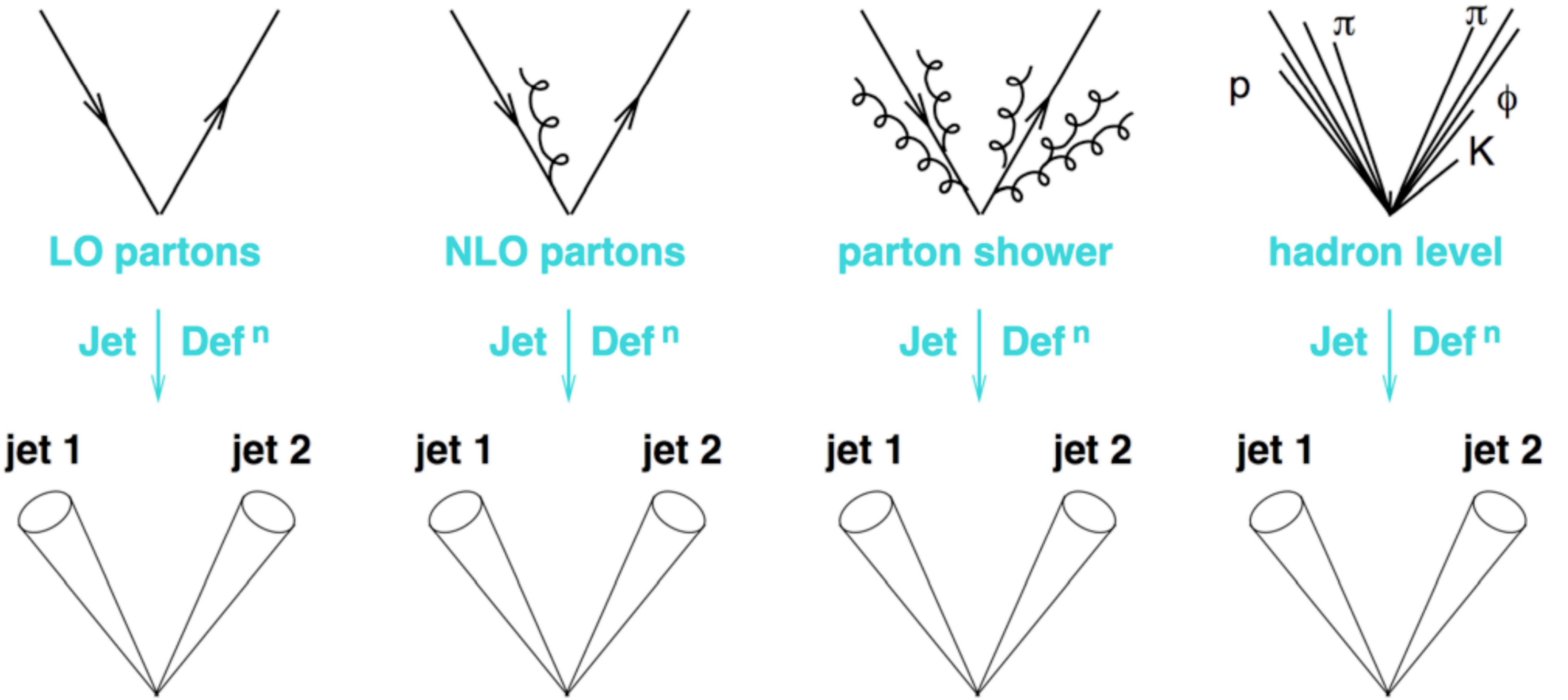


QCD predictions

Real data

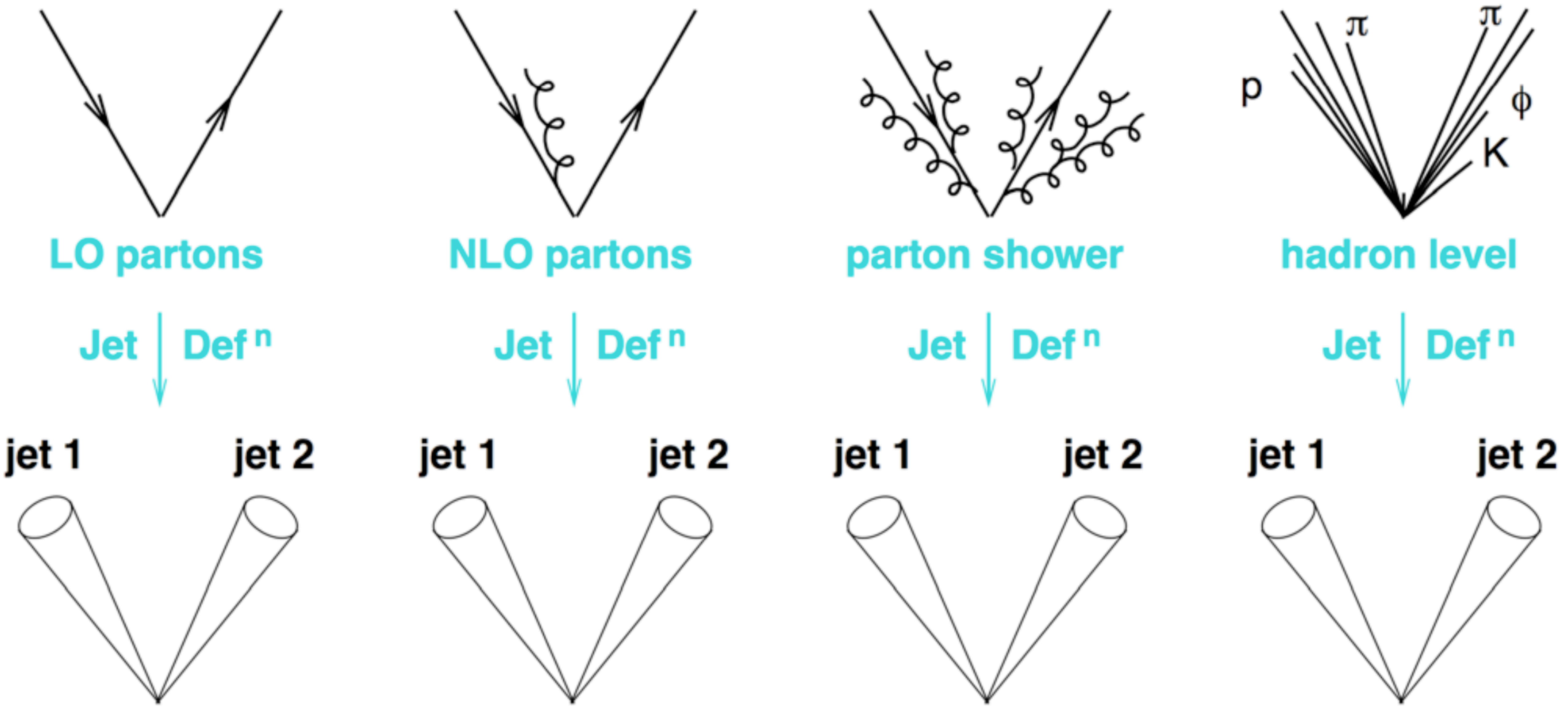
Jets

The requirement of a Jet algorithm



Projection to jets should be resilient to QCD effects

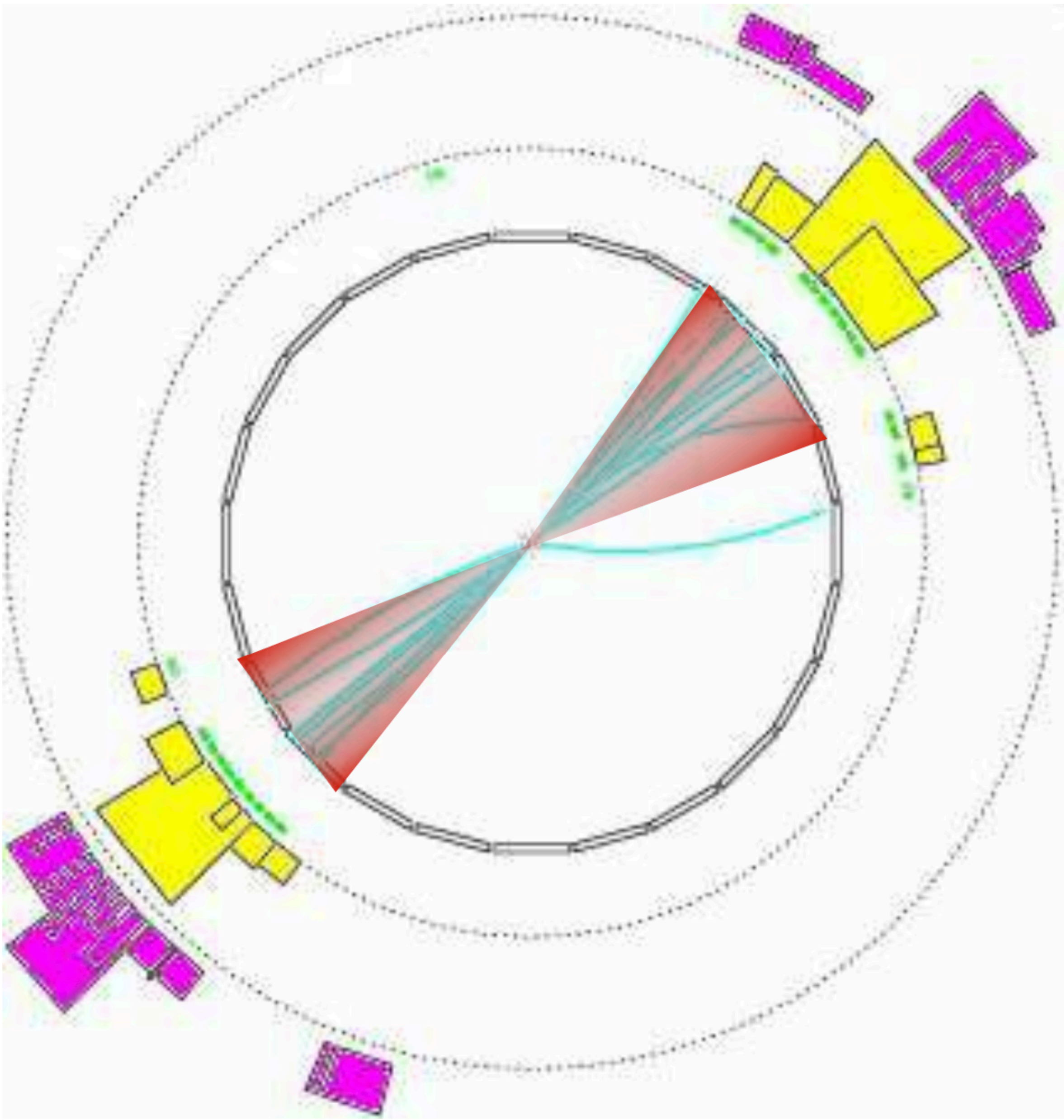
The requirement of a Jet algorithm



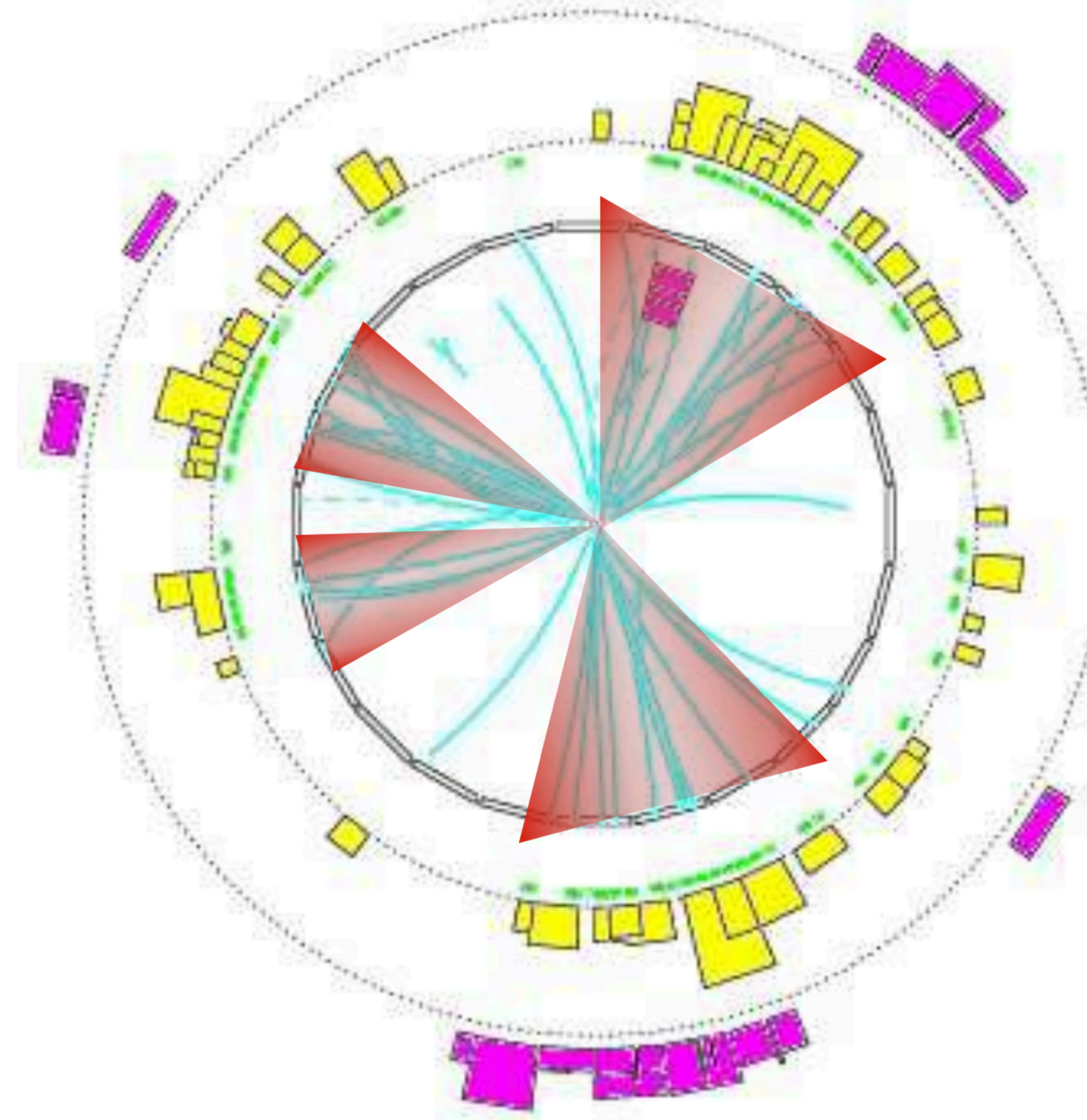
Projection to jets should be resilient to QCD effects

Statuary warning : A JET IS NOT SYNONYMOUS TO PARTON.

Jet reconstruction is a combinatorial puzzle



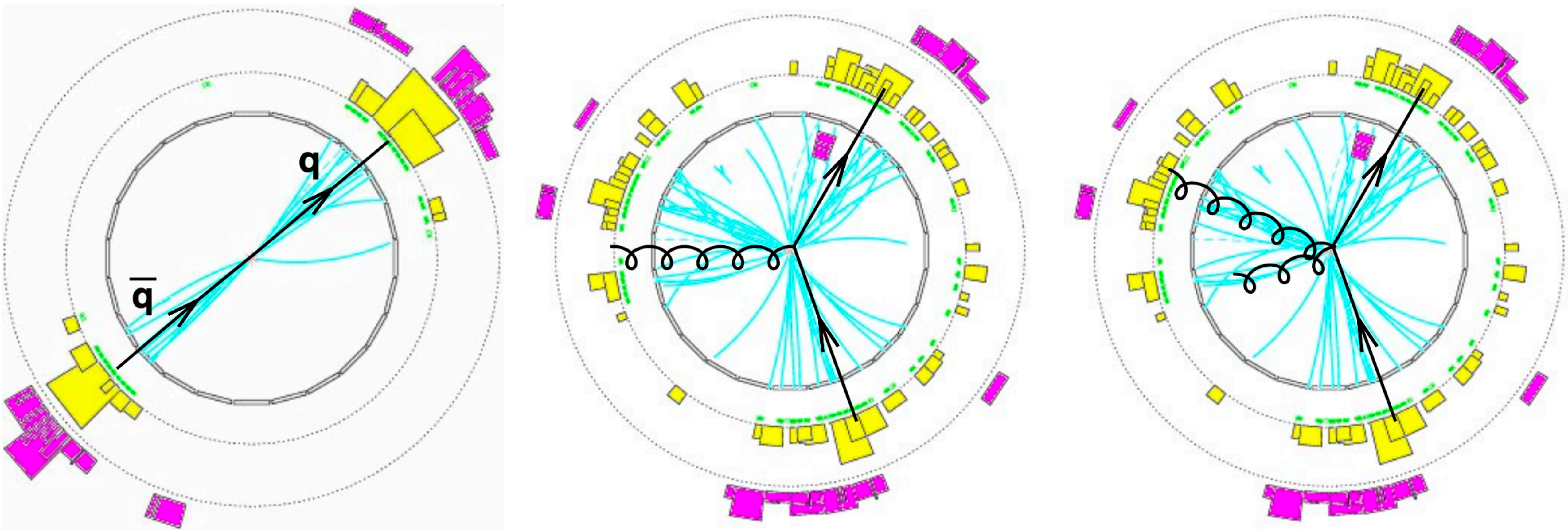
2 clear jets



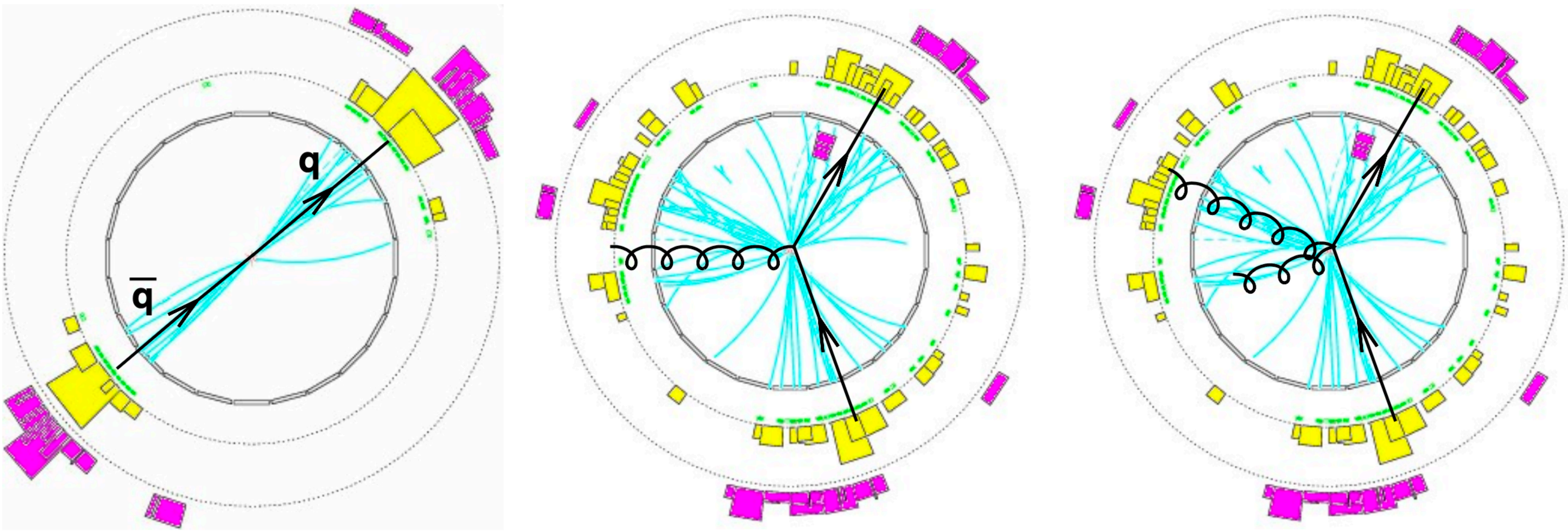
3 jets?
or 4 jets?

Jet reconstruction is a combinatorial puzzle

Jet reconstruction is a combinatorial puzzle



Jet reconstruction is a combinatorial puzzle



Close by energy clusters can also be from FSR of a hard parton.

How do we figure that out?

So what should be a jet definition?

$$\{p_i\} \xrightarrow{\text{jet definition}} \{j_k\}$$

particles,
4-momenta,
calorimeter towers, ...

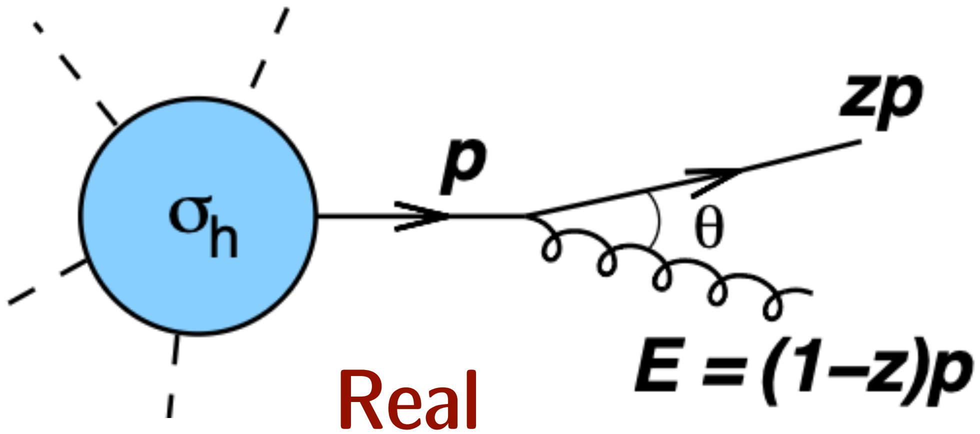
jets

- Which particles do you put together into a same jet?
- How do you recombine their momenta
  (4-momentum sum is the obvious choice, right?)

"Jet [definitions] are legal contracts between theorists and experimentalists"
-- MJ Tannenbaum

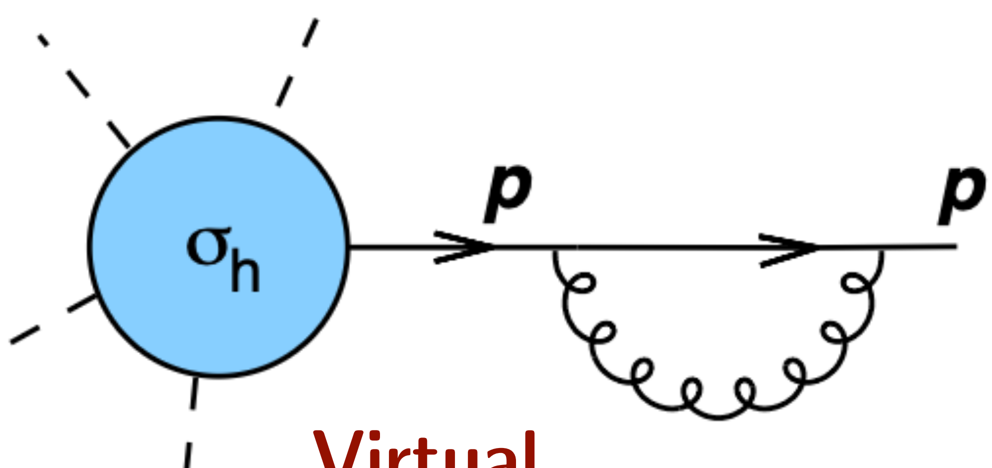
They're also a way of organising the information in an event
1000's of particles per events, up to 40.000,000 events per second

A jet definition better be IRC safe



Real

$$\sigma_{h+g} \simeq \sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2}$$



Virtual

$$\sigma_{h+V} \simeq -\sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2}$$

The IRC divergences from real emissions are cancelled by soft modes of loop integral.

If you are "**infrared and collinear safe**", i.e. your measurement doesn't care whether a soft/collinear gluon has been emitted then the **real and virtual divergences cancel.**

What's the point of jet construction

Jets can serve two purposes

- They can be ***observables***, that one can measure and calculate

- They can be ***tools***, that one can employ to extract specific properties of the final state

Different clustering algorithms have different properties and characteristics that can make them more or less appropriate for each of these tasks

IRC safety

An observable is **infrared and collinear safe** if, in the limit of a **collinear splitting**, or the **emission of an infinitely soft** particle, the observable remains **unchanged:**

$$O(X; p_1, \ldots, p_n, p_{n+1} \to 0) \to O(X; p_1, \ldots, p_n)$$

$$O(X; p_1, \ldots, p_n \parallel p_{n+1}) \to O(X; p_1, \ldots, p_n + p_{n+1})$$

This property ensures cancellation of **real** and **virtual** divergences in higher order calculations

If we wish to be able to calculate a jet rate in perturbative QCD the jet algorithm that we use must be IRC safe:
soft emissions and collinear splittings must not change the hard jets

First pointed out by Sterman & Weinberg

First pointed out by Sterman & Weinberg

VOLUME 39, NUMBER 23 PHYSICAL REVIEW LETTERS 5 DECEMBER 1977

Jets from Quantum Chromodynamics

George Sterman
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790

and

Steven Weinberg
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 July 1977)

First pointed out by Sterman & Weinberg

VOLUME 39, NUMBER 23 PHYSICAL REVIEW LETTERS 5 DECEMBER 1977

Jets from Quantum Chromodynamics

George Sterman
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790

and

Steven Weinberg
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 July 1977)

$$\sigma_a = (d\sigma/d\Omega)_0\,\Omega(g_E{}^2/3\pi^2)[-3\ln(E\delta/\mu) - 2\ln^2 2\epsilon - 4\ln(E\delta/\mu)\ln(2\epsilon) + \tfrac{17}{4} - \pi^2/3], \quad (2)$$

$$\sigma_b = (d\sigma/d\Omega)_0\,\Omega(g_E{}^2/3\pi^2)[2\ln^2(2\epsilon E/\mu) - \pi^2/6], \quad (3)$$

$$\sigma_c = (d\sigma/d\Omega)_0\,\Omega\{1 + (g_E{}^2/3\pi^2)[-2\ln^2(E/\mu) + 3\ln(E/\mu) - \tfrac{7}{4} + \pi^2/6]\}, \quad (4)$$

First pointed out by Sterman & Weinberg

VOLUME 39, NUMBER 23 PHYSICAL REVIEW LETTERS 5 DECEMBER 1977

Jets from Quantum Chromodynamics

George Sterman
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790

and

Steven Weinberg
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 July 1977)

$$\sigma_a = (d\sigma/d\Omega)_0\,\Omega(g_E{}^2/3\pi^2)[-3\ln(E\delta/\mu) - 2\ln^2 2\epsilon - 4\ln(E\delta/\mu)\ln(2\epsilon) + \tfrac{17}{4} - \pi^2/3], \quad (2)$$

$$\sigma_b = (d\sigma/d\Omega)_0\,\Omega(g_E{}^2/3\pi^2)[2\ln^2(2\epsilon E/\mu) - \pi^2/6], \quad (3)$$

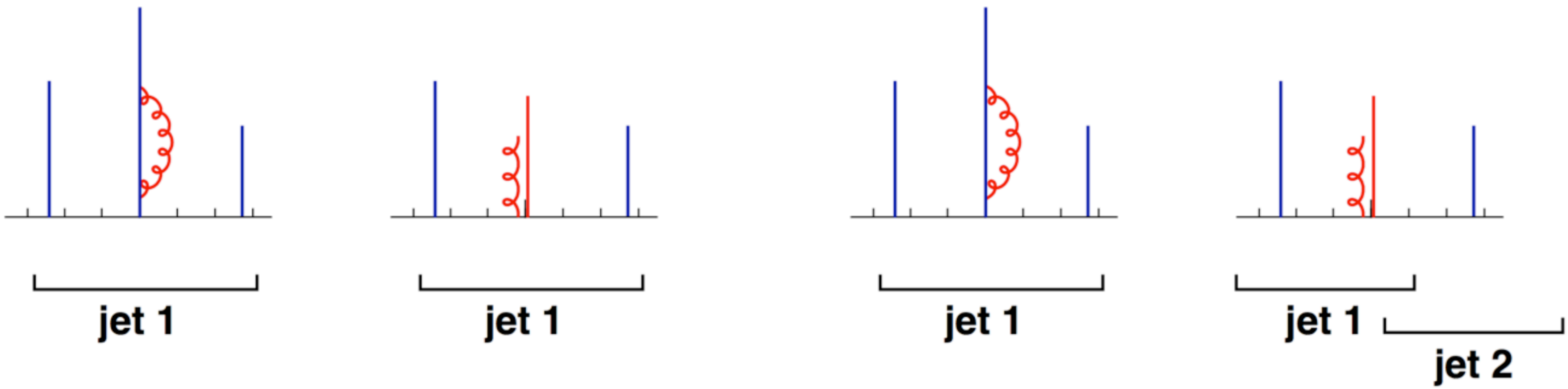
$$\sigma_c = (d\sigma/d\Omega)_0\,\Omega\{1 + (g_E{}^2/3\pi^2)[-2\ln^2(E/\mu) + 3\ln(E/\mu) - \tfrac{7}{4} + \pi^2/6]\}, \quad (4)$$

As expected, each separate contribution is singular for $\mu \to 0$, but cancellations[8] occur in the sum, and the final result is free of mass singularities:

$$\sigma(E, \theta, \Omega, \epsilon, \delta) = (d\sigma/d\Omega)_0\,\Omega[1 - (g_E{}^2/3\pi^2)(3\ln\delta + 4\ln\delta\ln 2\epsilon + \pi^2/3 - \tfrac{5}{2})]. \quad (6)$$

Jet defs : which one is legal?

Collinear Safe — **Collinear Unsafe**



$\alpha_s^n \times (-\infty)$ $\alpha_s^n \times (+\infty)$ $\alpha_s^n \times (-\infty)$ $\alpha_s^n \times (+\infty)$

Infinities cancel — **Infinities do not cancel**

Perturbative calculations of jet observable will only be possible with collinear (and infrared) safe jet definitions

Two main classes of algorithms

- ## Sequential recombination algorithms

Bottom-up approach: combine particles starting from **closest ones**

How? Choose a **distance measure**, iterate recombination until few objects left, call them jets

Works because of mapping closeness ⇔ QCD divergence

Examples: Jade, k_t, Cambridge/Aachen, anti-k_t, …..

Usually trivially made IRC safe, but their algorithmic complexity scales like N^3

- ## Cone algorithms

Top-down approach: find coarse regions of energy flow.

How? Find **stable cones** (i.e. their axis coincides with sum of momenta of particles in it)

Works because QCD only modifies energy flow on small scales

Examples: JetClu, MidPoint, ATLAS cone, CMS cone, SISCone…...

Can be programmed to be fairly fast, at the price of being complex and IRC unsafe

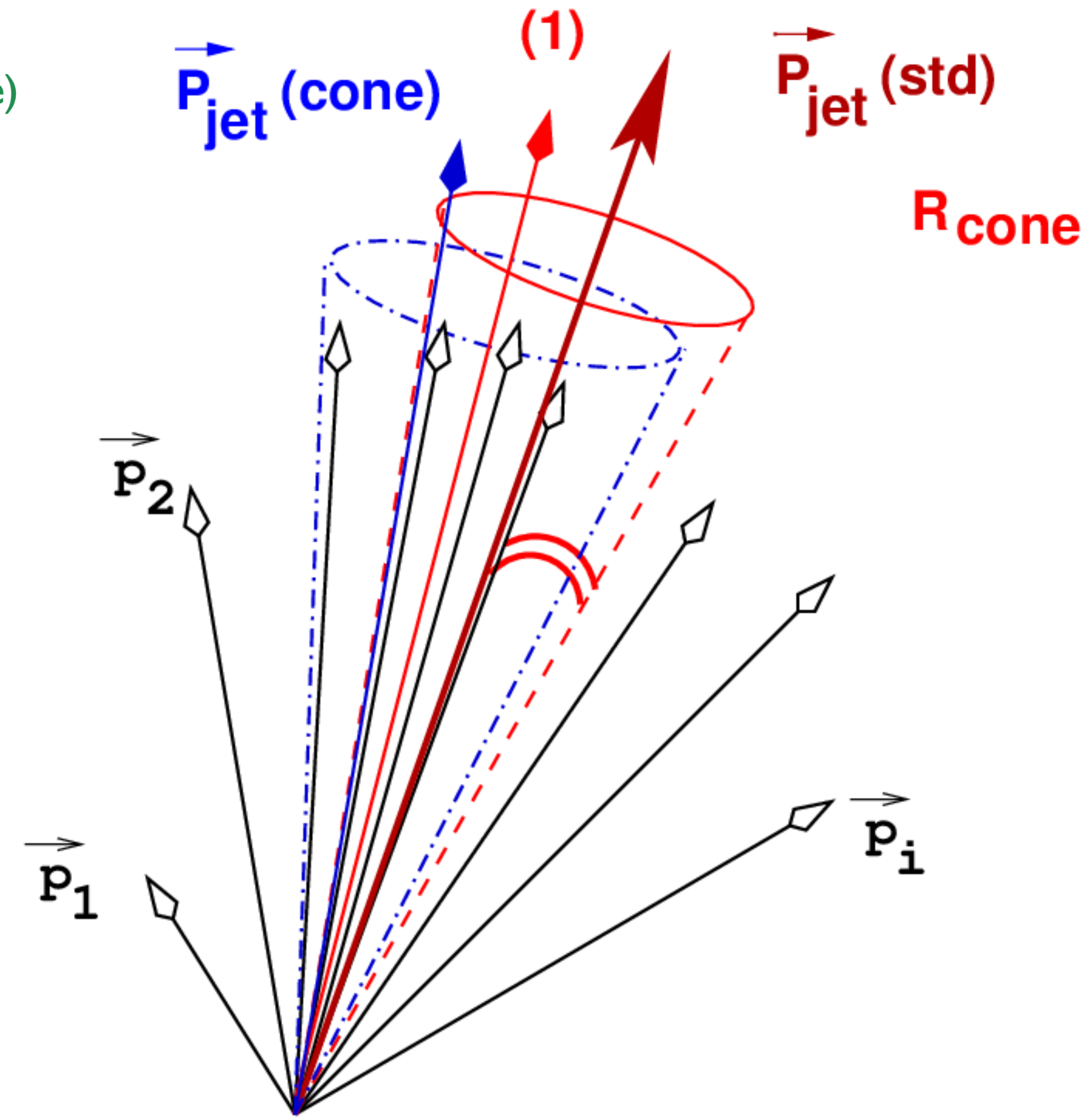
A word or two about cone algorithms

In partitional-type algorithms (i.e. cones), one wishes to find the **stable configurations**:
axis of cones coincides with sum of 4-momenta of the particles it contains

The 'safe' way of doing so is to test **all possible combinations** of N objects

The main sub-categories of cone algorithms are:

- ✱ **Fixed** cone with **progressive removal** (FC-PR) (PyJet, CellJet, GetJet)
- ✱ **Iterative** cone with **progressive removal** (IC-PR) (CMS iterative cone)
- ✱ **Iterative** cone with **split-merge** (IC-SM) (JetClu, ATLAS cone)
- ✱ **IC-SM** with **mid-points** (IC_{mp}-SM) (CDF MidPoint, D0 Run II)
- ✱ **IC_{mp}** with **split-drop** (IC_{mp}-SD) (PxCone)
- ✱ **Seedless** cone with **split-merge** (SC-SM) (SISCone)



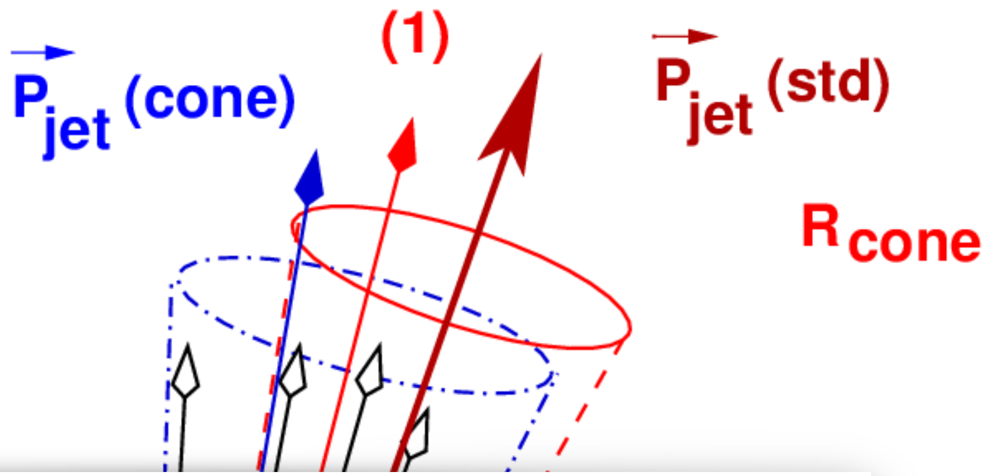
A word or two about cone algorithms

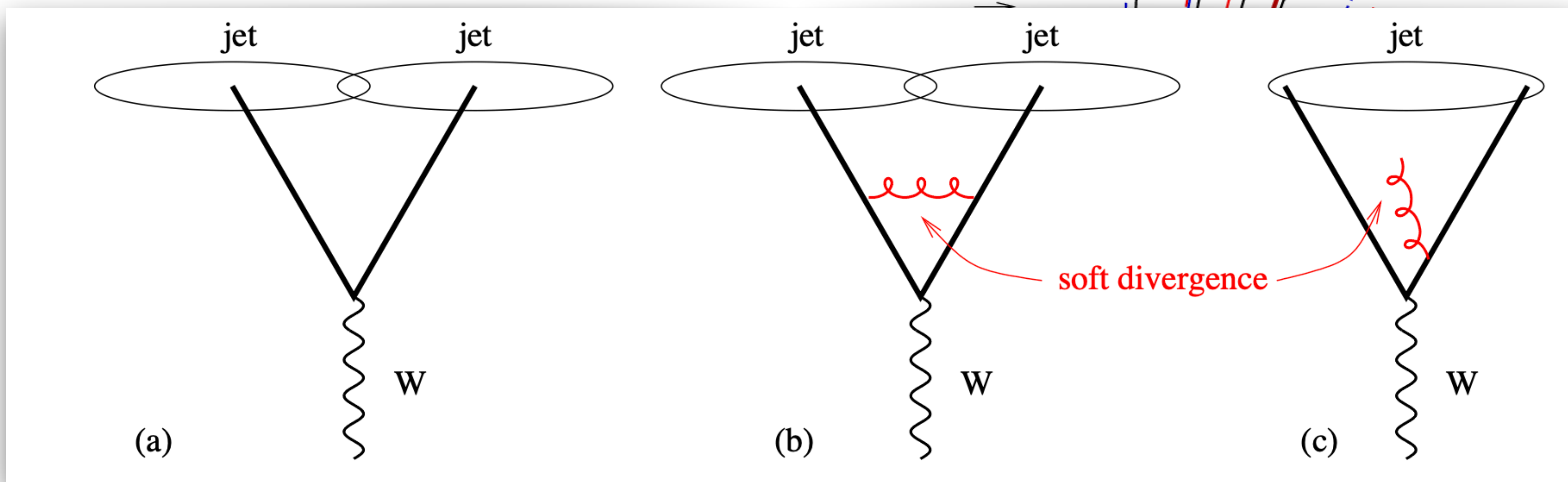
In partitional-type algorithms (i.e. cones), one wishes to find the **stable configurations**:
axis of cones coincides with sum of 4-momenta of the particles it contains

The 'safe' way of doing so is to test **all possible combinations** of N objects

The main sub-categories of cone algorithms are:

- ✱ **Fixed** cone with **progressive removal** (FC-PR) (PyJet, CellJet, GetJet)
- ✱ **Iterative** cone with **progressive removal** (IC-PR) (CMS iterative cone)
- ✱ **Iterative** cone with **split-merge** (IC-SM) (JetClu, ATLAS cone)
- ✱ **IC-SM** with **mid-points** (IC_{mp}-SM) (CDF MidPoint, D0 Run II)
- ✱ **IC_{mp}** with **split-drop** (IC_{mp}-SD) (PxCone)
- ✱ **Seedless** cone with **split-merge** (SC-SM) (SISCone)

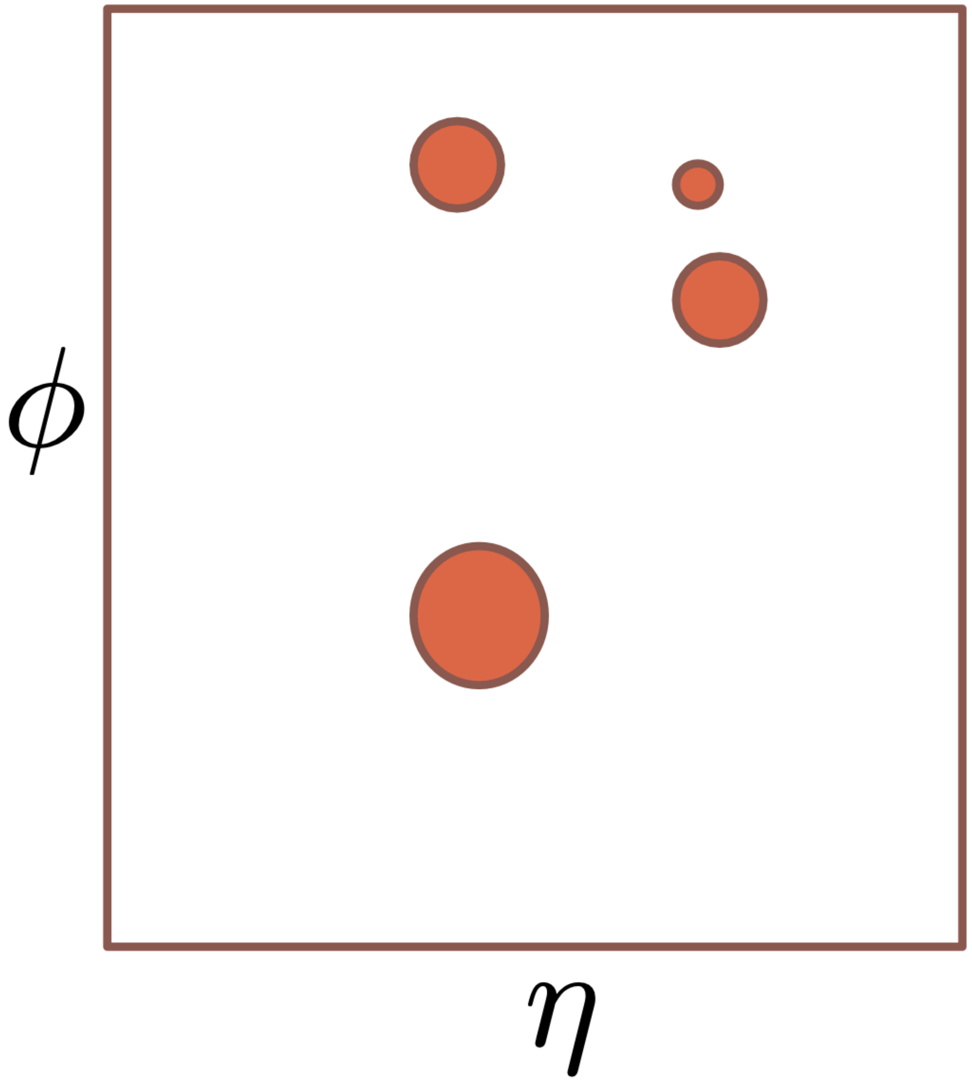




Recombination algorithm : basic intuitions

Start with a set of 4-vectors p^i

Represent the particles in η , ϕ plane



$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

Recombination algorithm : basic intuitions

Start with a set of 4-vectors p^i

Represent the particles in η , ϕ plane

$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

Recombination algorithm : basic intuitions

Start with a set of 4-vectors p^i

Represent the particles in η , ϕ plane

Calculate the pairwise distance :

$$d_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$

$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

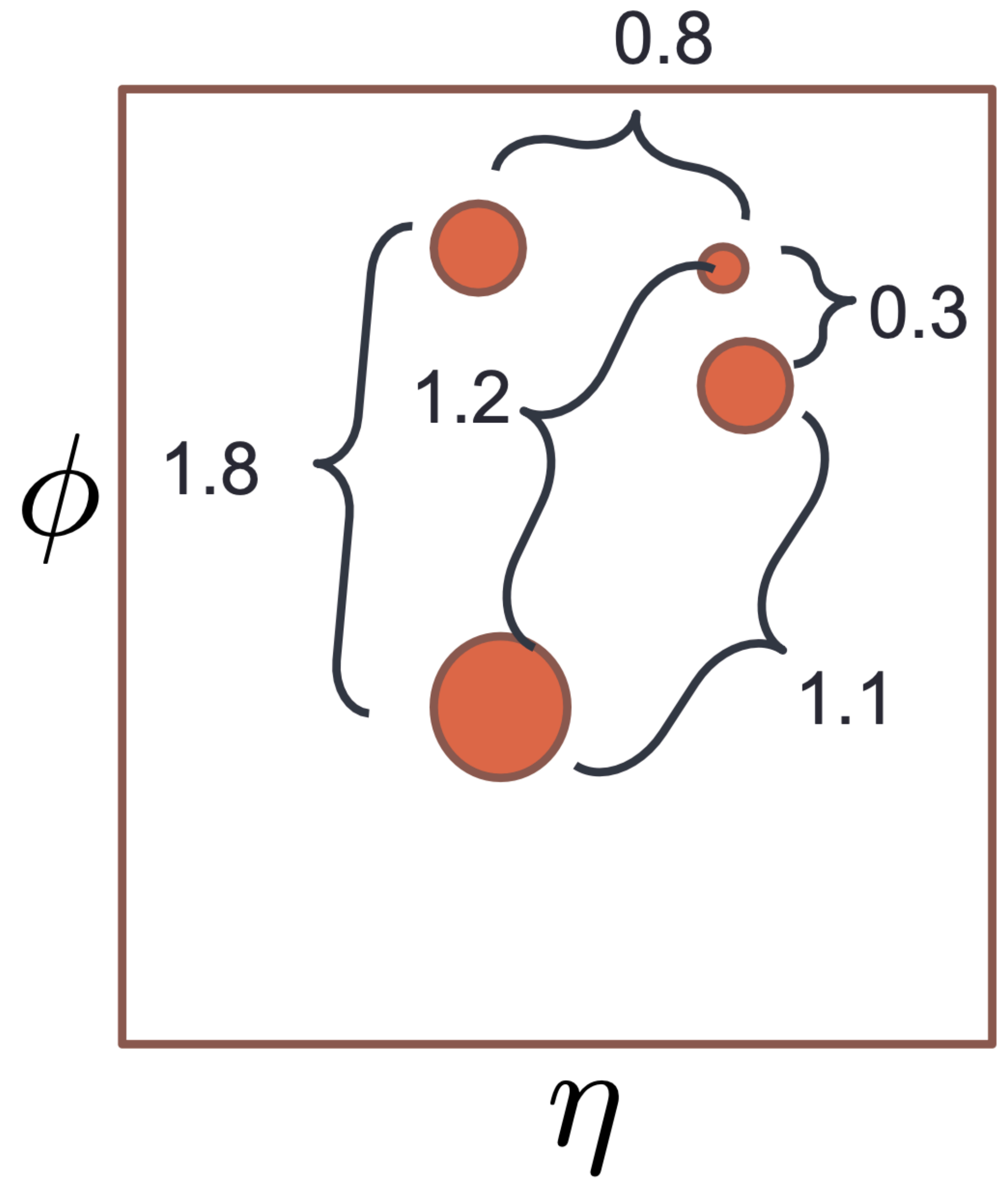
Recombination algorithm : basic intuitions

Start with a set of 4-vectors p^i

Represent the particles in η , ϕ plane

Calculate the pairwise distance :

$$d_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$



$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

Recombination algorithm : basic intuitions

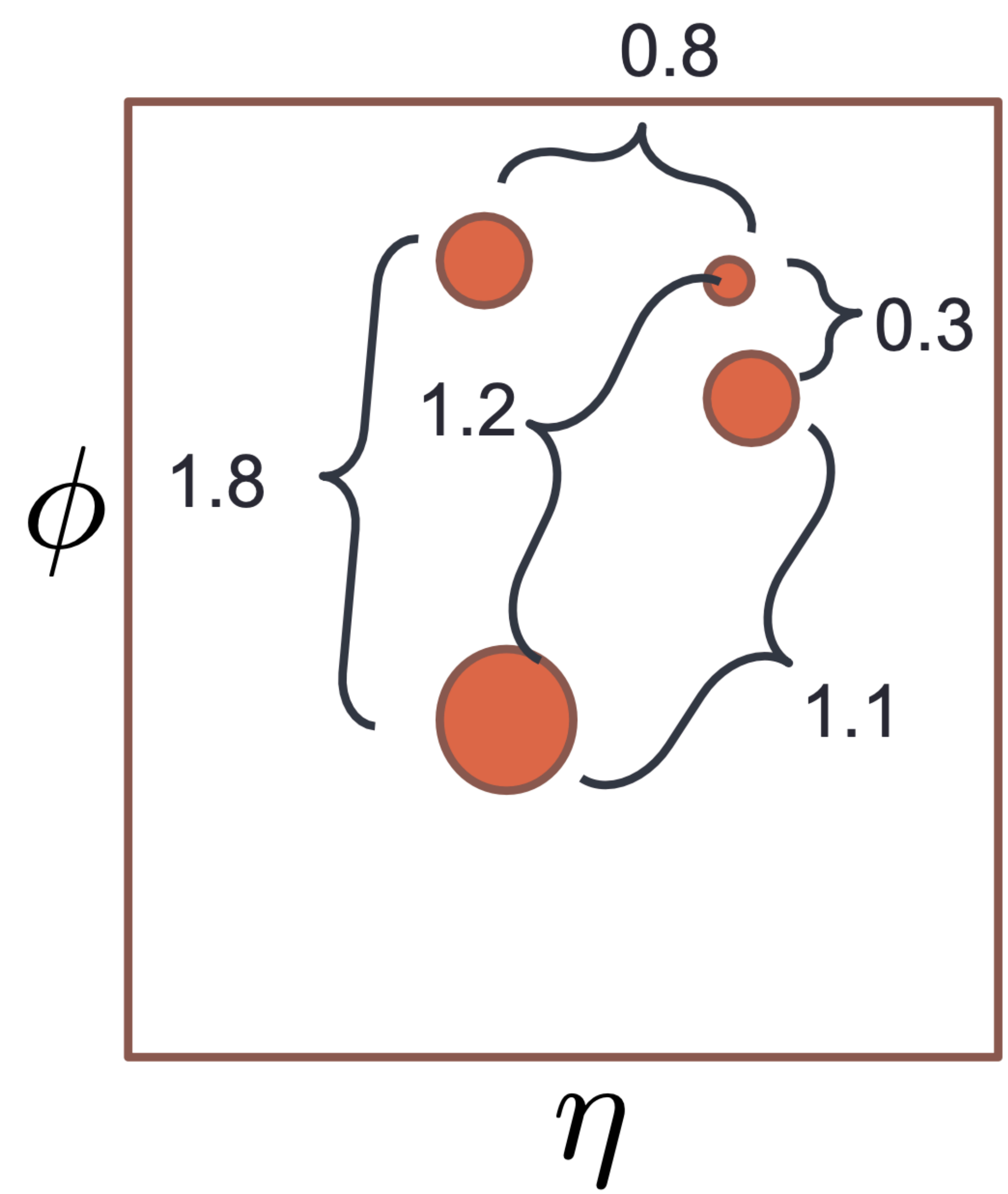
Start with a set of 4-vectors p^i

Represent the particles in η , ϕ plane

Calculate the pairwise distance :

$$d_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$

Merge the two closest particle.
i.e. replace the pair by a combined representation.



$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

Recombination algorithm : basic intuitions

Start with a set of 4-vectors p^i

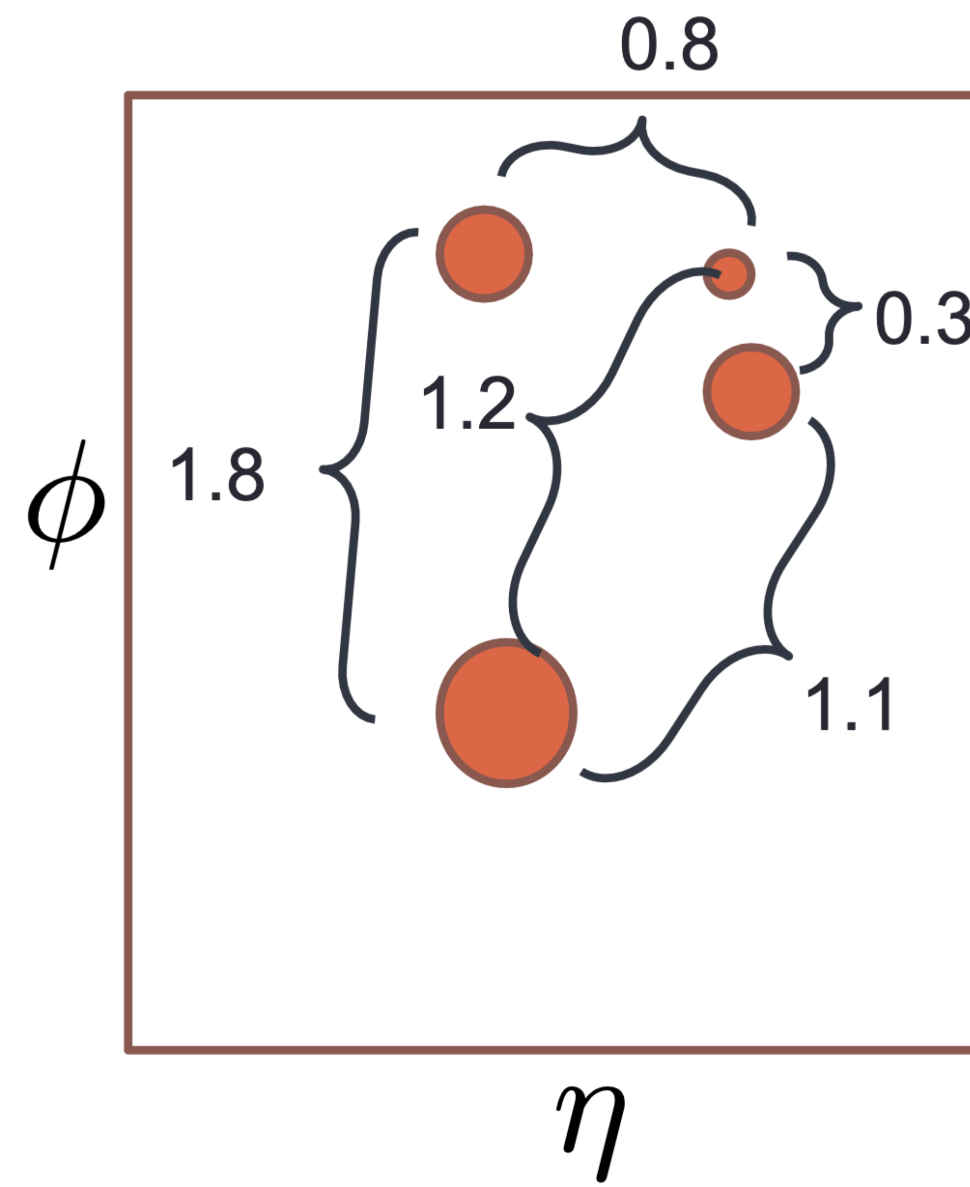
Represent the particles in η , ϕ plane

Calculate the pairwise distance :

$$d_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$

Merge the two closest particle.
i.e. replace the pair by a combined representation.

Continue until no two particles are closer than R



$$\eta = -\frac{1}{2} ln\left(tan\ \frac{\theta}{2}\right)$$

$$\phi = arctan\ \left(\frac{p_y}{p_x}\right)$$

e^+e^- k_T algorithm

Durham: $$y_{ij} = \frac{2\min(E_i^2, E_j^2)\,(1-\cos\theta_{ij})}{Q^2},$$

Geneva: $$y_{ij} = \frac{8}{9} \cdot \frac{2E_iE_j\,(1-\cos\theta_{ij})}{(E_i+E_j)^2},$$

Jade-E0: $$y_{ij} = \frac{2E_iE_j\,(1-\cos\theta_{ij})}{Q^2},$$

https://arxiv.org/pdf/1011.6247

e^+e^- k_T algorithm

Durham: $$y_{ij} = \frac{2\min(E_i^2, E_j^2)\,(1-\cos\theta_{ij})}{Q^2},$$

Geneva: $$y_{ij} = \frac{8}{9} \cdot \frac{2E_iE_j\,(1-\cos\theta_{ij})}{(E_i+E_j)^2},$$

Jade-E0: $$y_{ij} = \frac{2E_iE_j\,(1-\cos\theta_{ij})}{Q^2},$$

https://arxiv.org/pdf/1011.6247

1. Define a resolution parameter y_{cut}
2. For every pair (p_k, p_l) of final-state particles compute the corresponding resolution variable y_{kl}.
3. If y_{ij} is the smallest value of y_{kl} computed above and $y_{ij} < y_{cut}$ then combine (p_i, p_j) into a single jet ('pseudo-particle') with momentum p_{ij} according to a recombination prescription.
4. Repeat until all pairs of objects (particles and/or pseudo-particles) have $y_{kl} > y_{cut}$.
5. Objects with $E \geq E_{min}$ are called jets.

Generalized kT (recombination) algorithm

Generalized kT (recombination) algorithm

1. Take the particles in the event as our initial list of objects.
2. From the list of objects, build two sets of distances: an *inter-particle distance*

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2, \quad (3.1)$$

Generalized kT (recombination) algorithm

1. Take the particles in the event as our initial list of objects.
2. From the list of objects, build two sets of distances: an *inter-particle distance*

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2, \tag{3.1}$$

where p is a free parameter and ΔR_{ij} is the geometric distance in the rapidity-azimuthal angle plane (Eq. (2.34)), and a *beam distance*

$$d_{iB} = p_{t,i}^{2p} R^2, \tag{3.2}$$

with R a free parameter usually called the *jet radius*.

Generalized kT (recombination) algorithm

1. Take the particles in the event as our initial list of objects.
2. From the list of objects, build two sets of distances: an *inter-particle distance*

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2, \qquad (3.1)$$

where p is a free parameter and ΔR_{ij} is the geometric distance in the rapidity-azimuthal angle plane (Eq. (2.34)), and a *beam distance*

$$d_{iB} = p_{t,i}^{2p} R^2, \qquad (3.2)$$

with R a free parameter usually called the *jet radius*.

3. Iteratively find the smallest distance among all the d_{ij} and d_{iB}
   - If the smallest distance is a d_{ij} then objects i and j are removed from the list and recombined into a new object k (using the recombination scheme) which is itself added to the list.
   - If the smallest is a d_{iB}, object i is called a *jet* and removed from the list.

Generalized kT (recombination) algorithm

1. Take the particles in the event as our initial list of objects.
2. From the list of objects, build two sets of distances: an *inter-particle distance*

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2, \tag{3.1}$$

where p is a free parameter and ΔR_{ij} is the geometric distance in the rapidity-azimuthal angle plane (Eq. (2.34)), and a *beam distance*

$$d_{iB} = p_{t,i}^{2p} R^2, \tag{3.2}$$

with R a free parameter usually called the *jet radius*.

3. Iteratively find the smallest distance among all the d_{ij} and d_{iB}
   - If the smallest distance is a d_{ij} then objects i and j are removed from the list and recombined into a new object k (using the recombination scheme) which is itself added to the list.
   - If the smallest is a d_{iB}, object i is called a *jet* and removed from the list.

Go back to step-2

Generalized kT (recombination) algorithm

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2$$

Generalized kT (recombination) algorithm

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2$$

p = 1 k_t algorithm

S. Catani, Y. Dokshitzer, M. Seymour and B. Webber, Nucl. Phys. B406 (1993) 187
S.D. Ellis and D.E. Soper, Phys. Rev. D48 (1993) 3160

p = 0 Cambridge/Aachen algorithm

Y. Dokshitzer, G. Leder, S.Moretti and B. Webber, JHEP 08 (1997) 001
M.Wobisch and T.Wengler, hep-ph/9907280

p = -1 **anti-k_t algorithm**

MC, G. Salam and G. Soyez, arXiv:0802.1189

NB: in anti-kt pairs with a **hard** particle will cluster first: if no other hard particles are close by, the algorithm will give **perfect cones**

Generalized kT (recombination) algorithm

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2$$

Generalized kT (recombination) algorithm

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p})\Delta R_{ij}^2$$

p > 0

New **soft** particle ($p_t \to 0$) means that d → 0 ⇒ clustered first, no effect on jets

New **collinear** particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that d → 0 ⇒ clustered first, no effect on jets

p = 0

New **soft** particle ($p_t \to 0$) can be new jet of zero momentum ⇒ no effect on hard jets

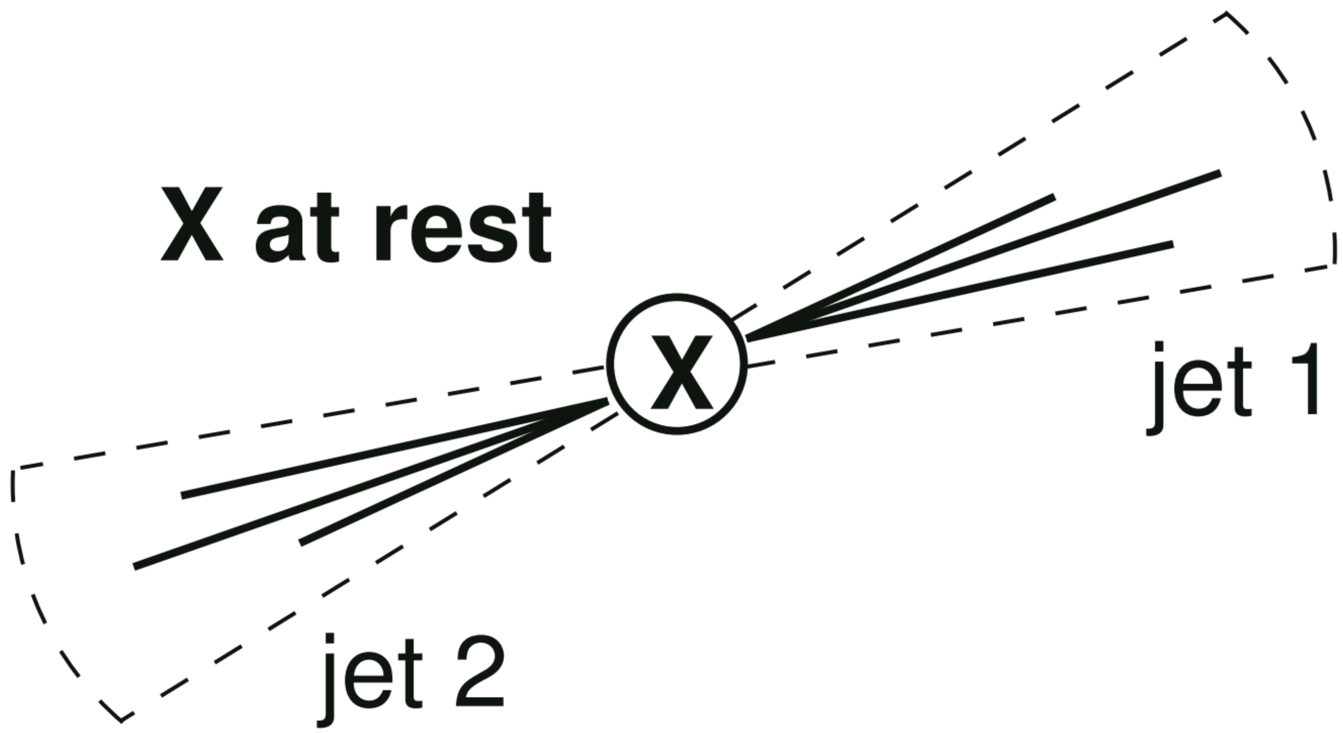
New **collinear** particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that d → 0 ⇒ clustered first, no effect on jets

p < 0

New **soft** particle ($p_t \to 0$) means d → ∞ ⇒ clustered last or new zero-jet, no effect on hard jets

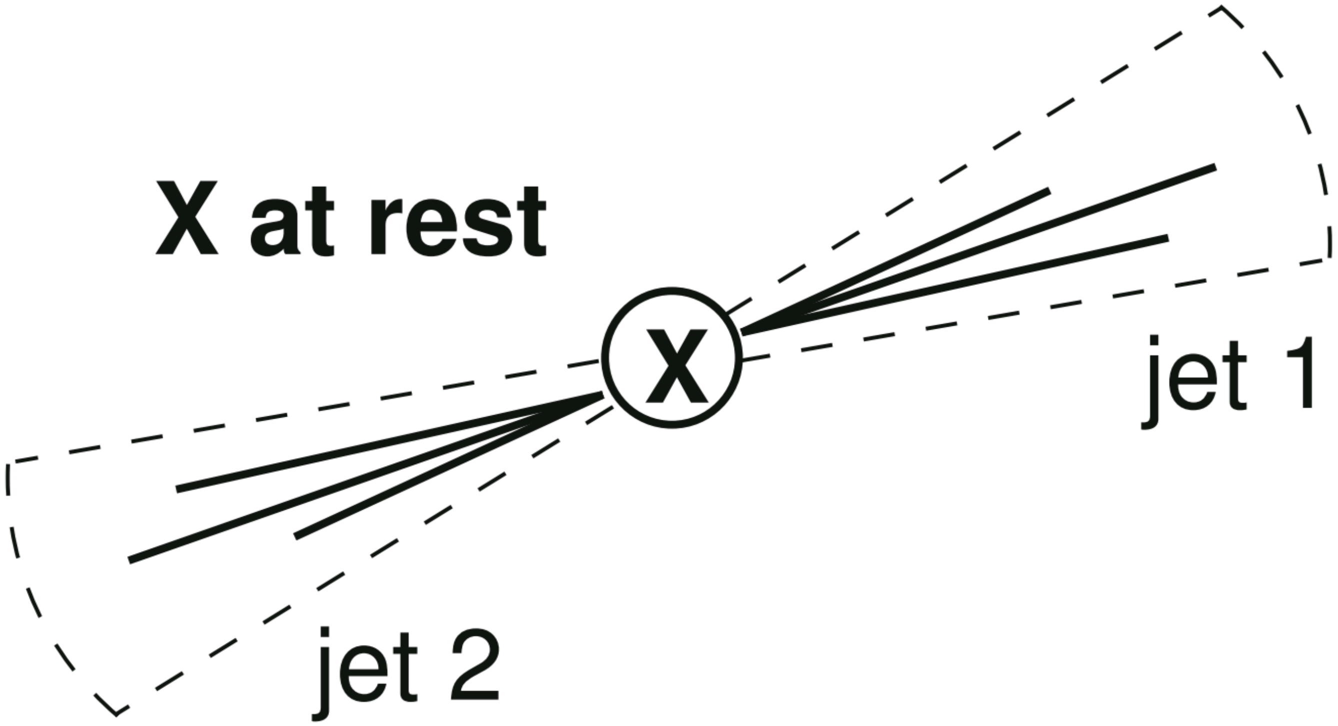
New **collinear** particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that d → 0 ⇒ clustered first, no effect on jets

Hadronic decay of a heavy particle

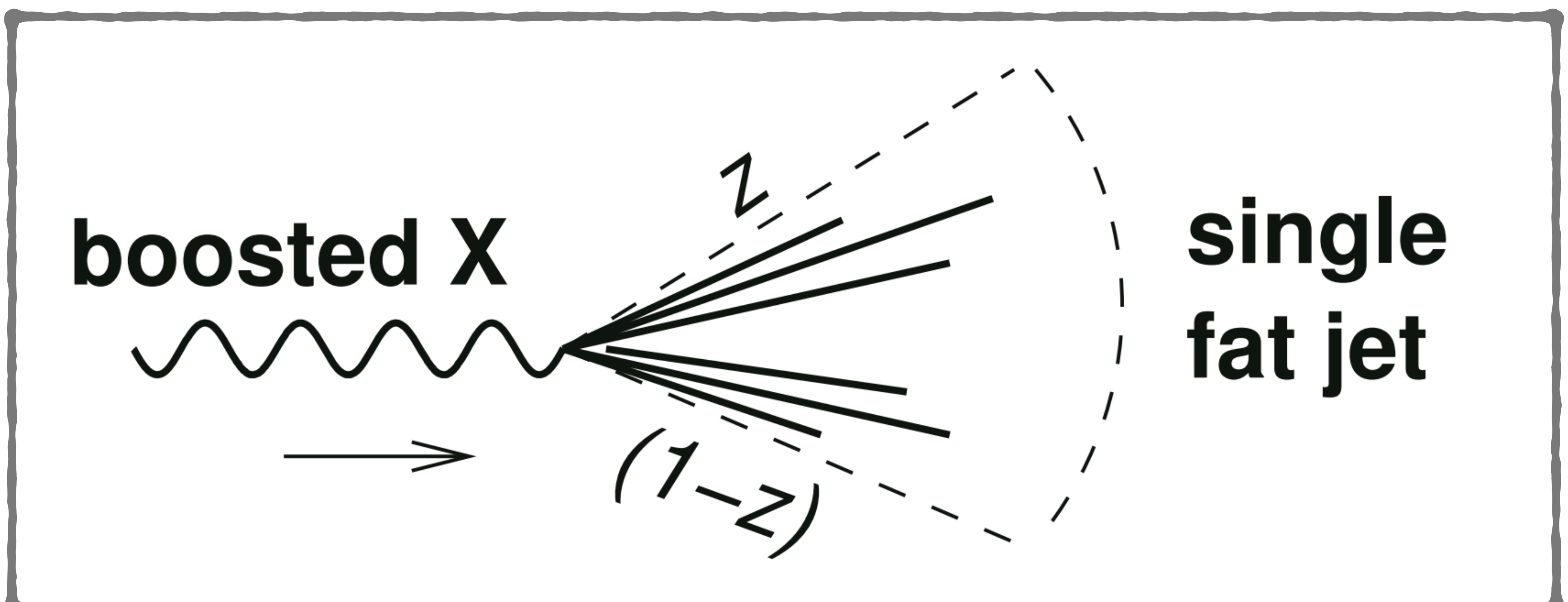


A heavy particle X decaying at rest

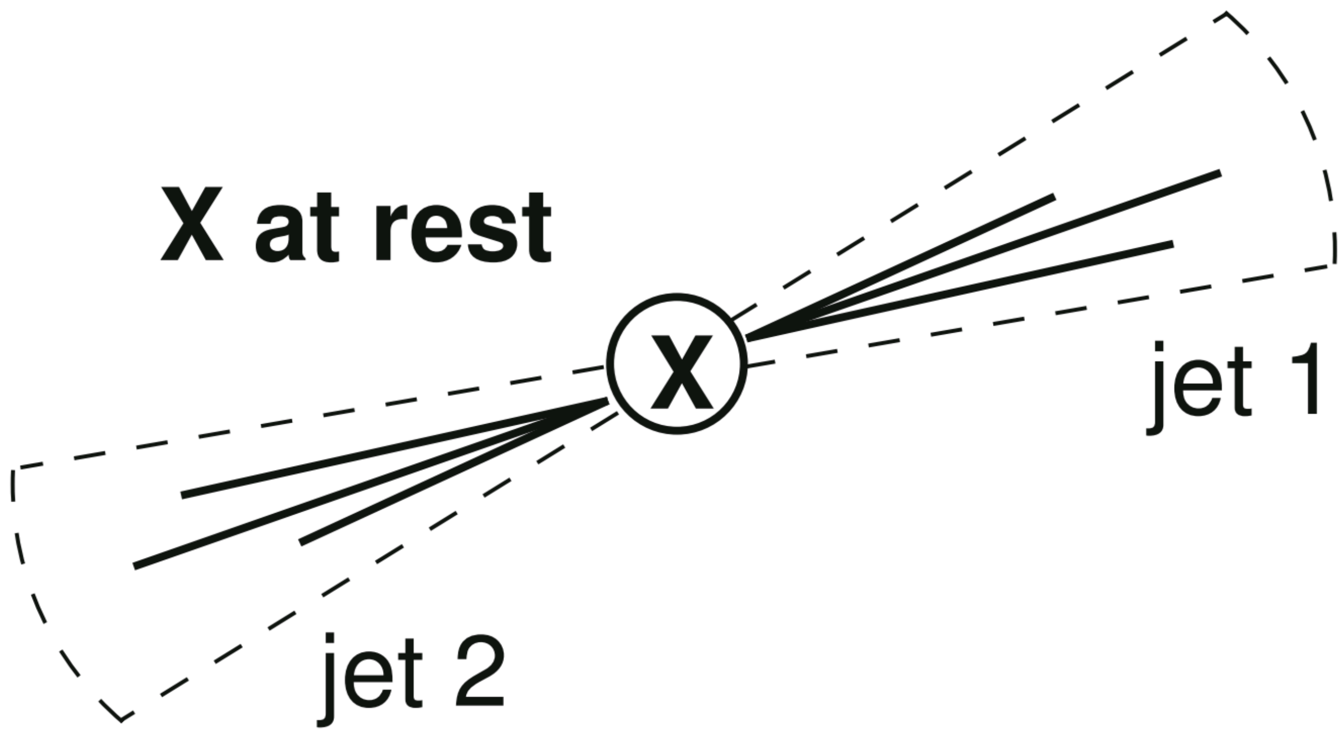
Hadronic decay of a heavy particle



A heavy particle X decaying at rest

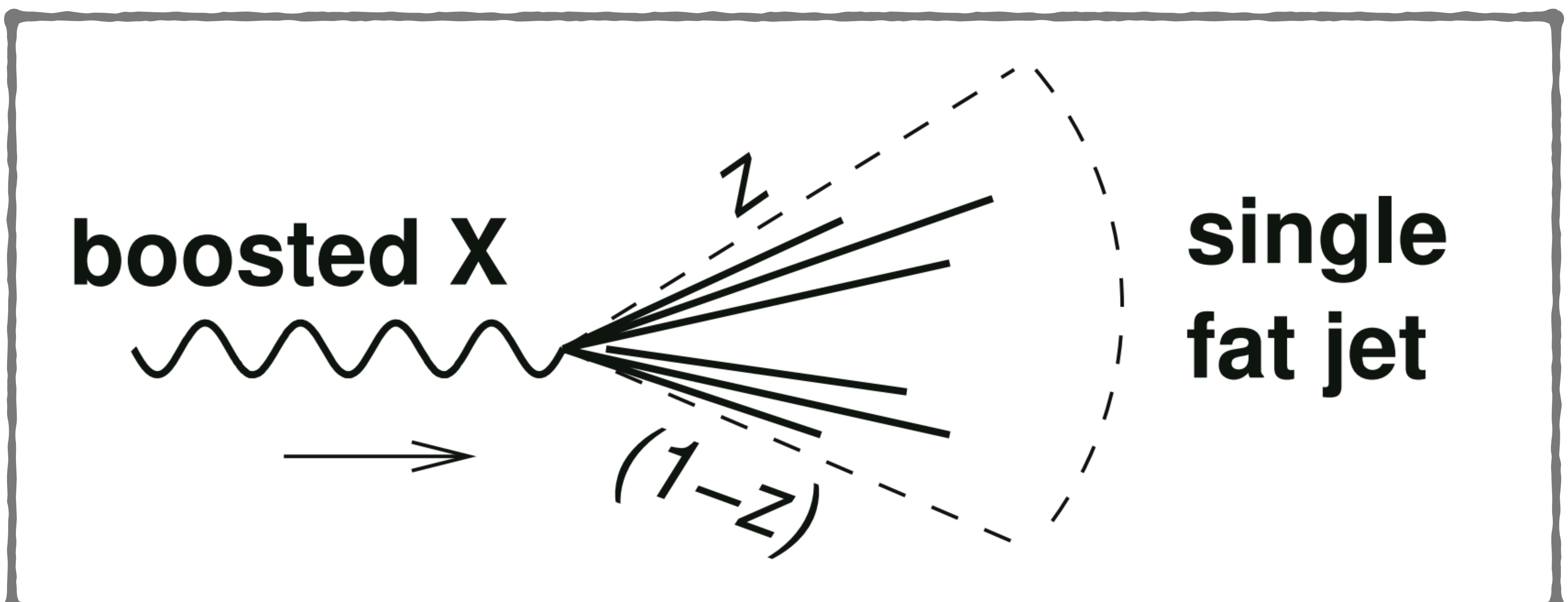


Hadronic decay of a heavy particle

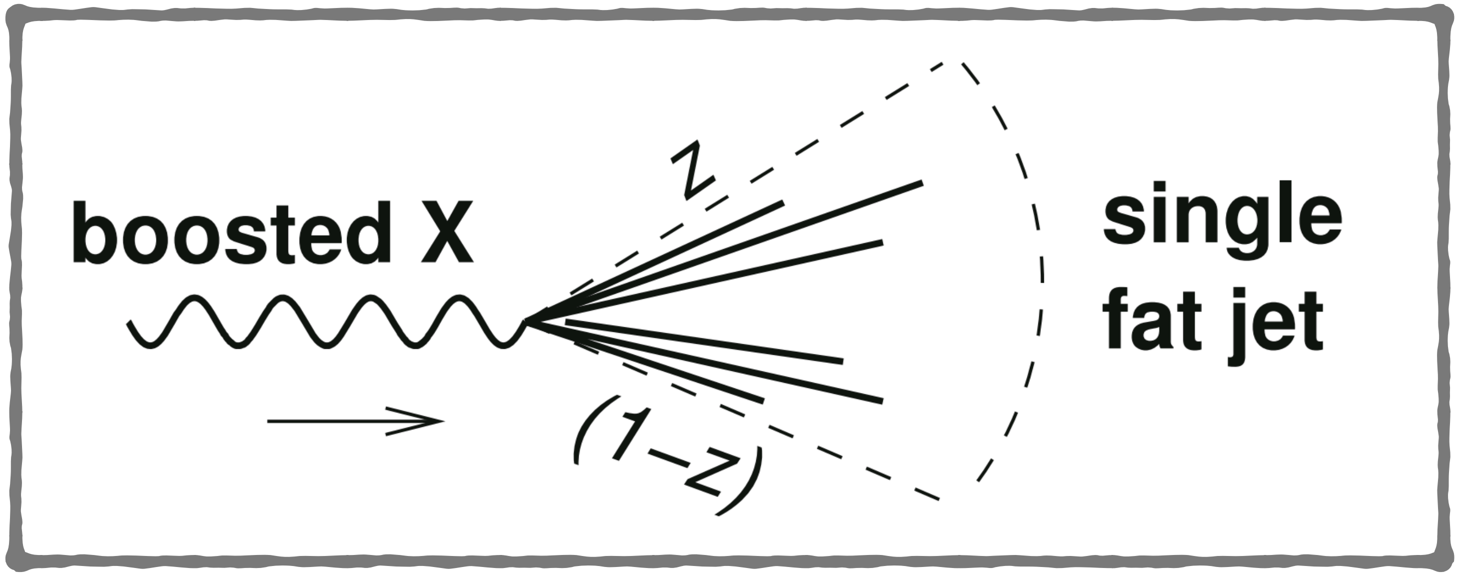


A heavy particle X decaying at rest

$$\Delta R \approx 2m/p_{\mathrm{T}}$$



Hadronic decay of a heavy particle



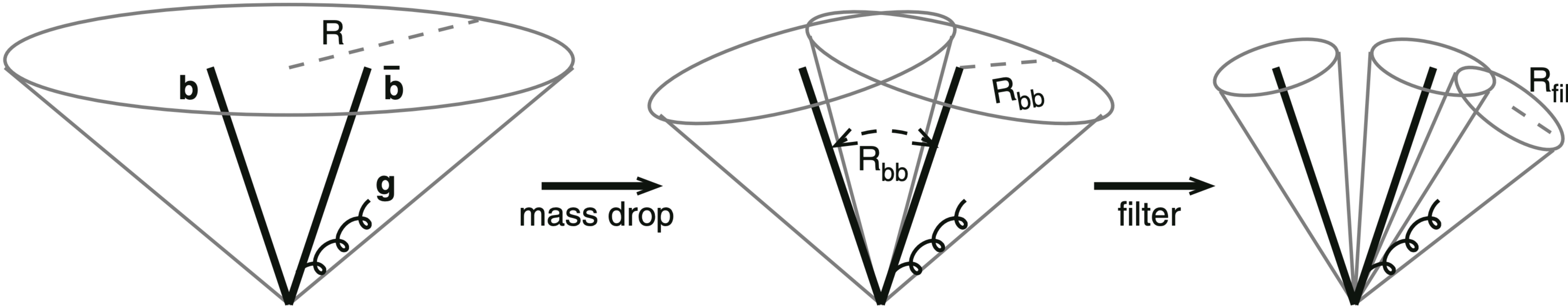
Electroweak particle X decaying hadronically into two jets.

$$m_X^2 = E_X^2 \; z(1-z) \; 2(1-cos\theta)$$

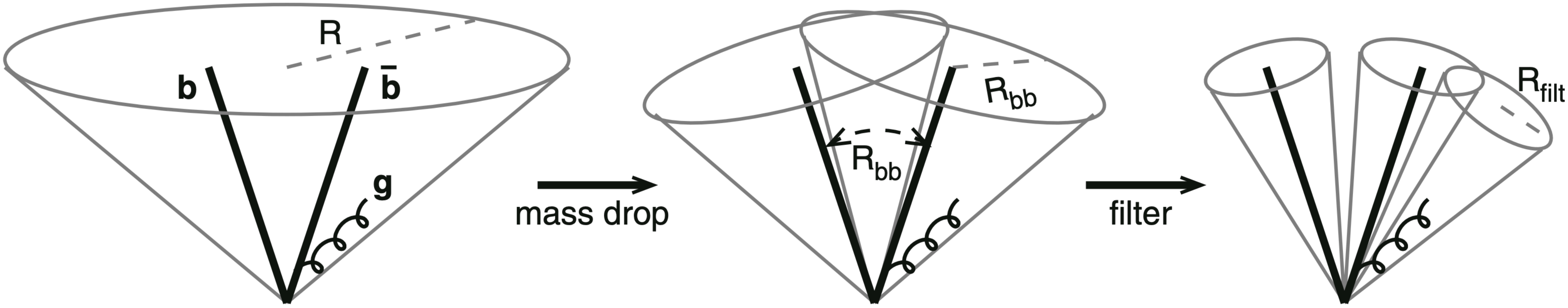
The two prongs end up in a single jet if :

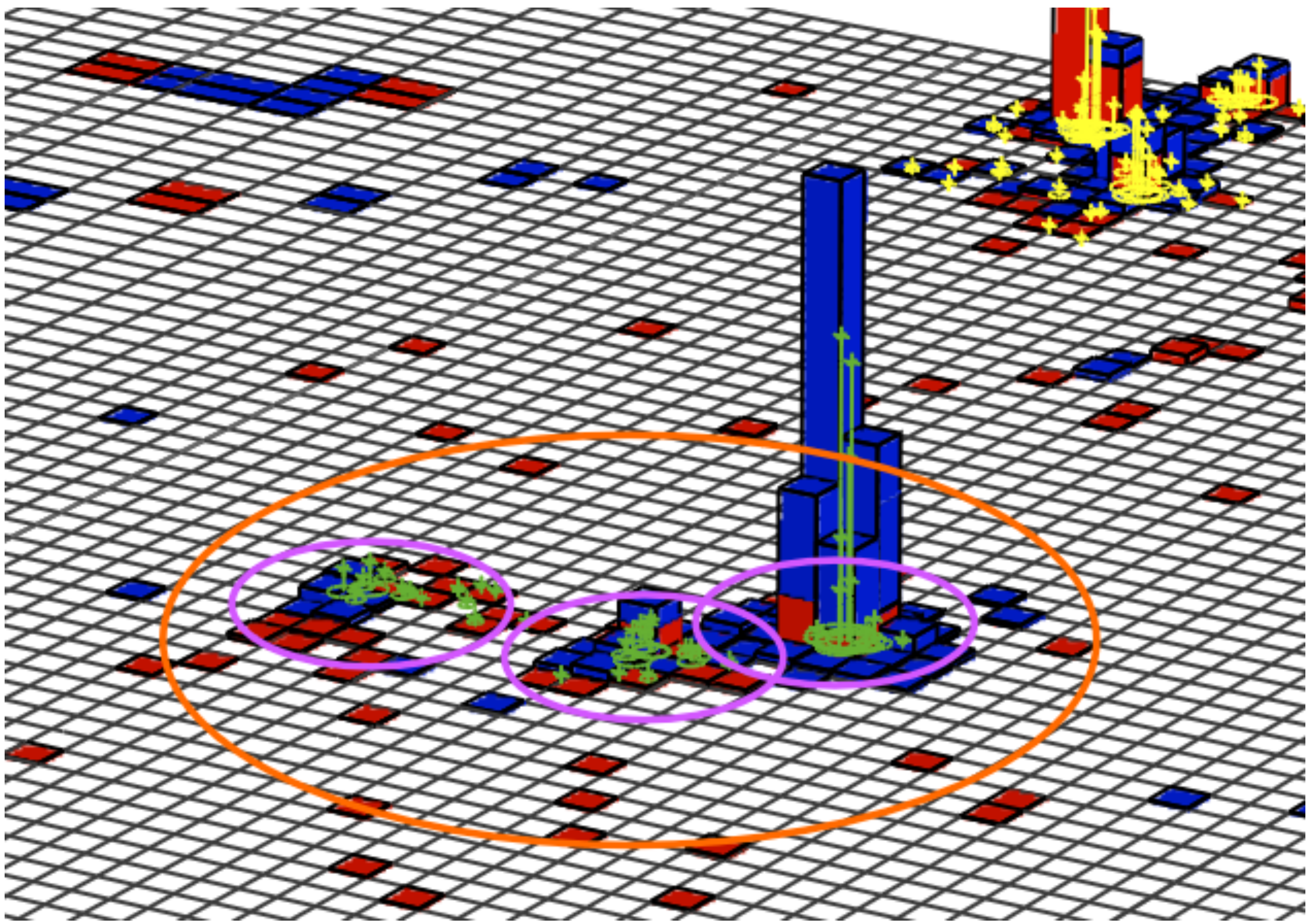
$$\Delta R \sim \frac{m}{p_T}\frac{1}{\sqrt{z(1-z)}} \sim \frac{2m}{p_T}$$

Initially in the context of $H \to b\bar{b}$ / $t \to bW$

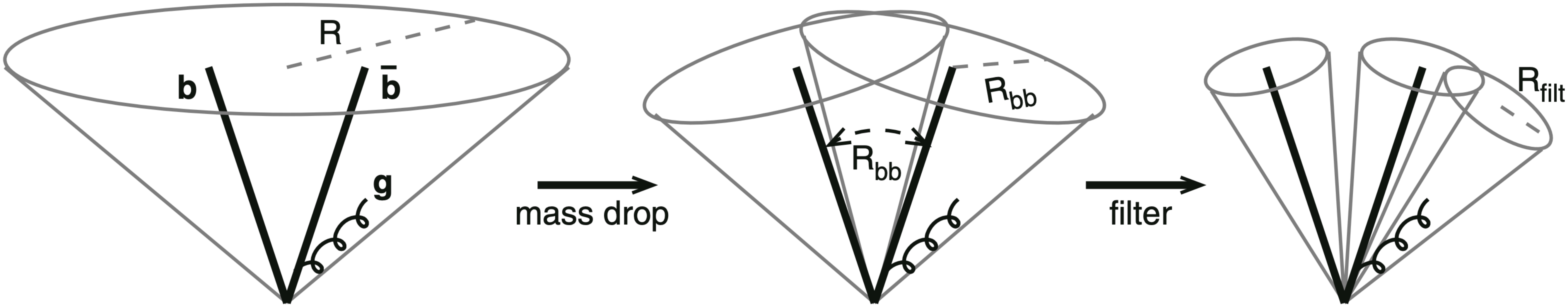


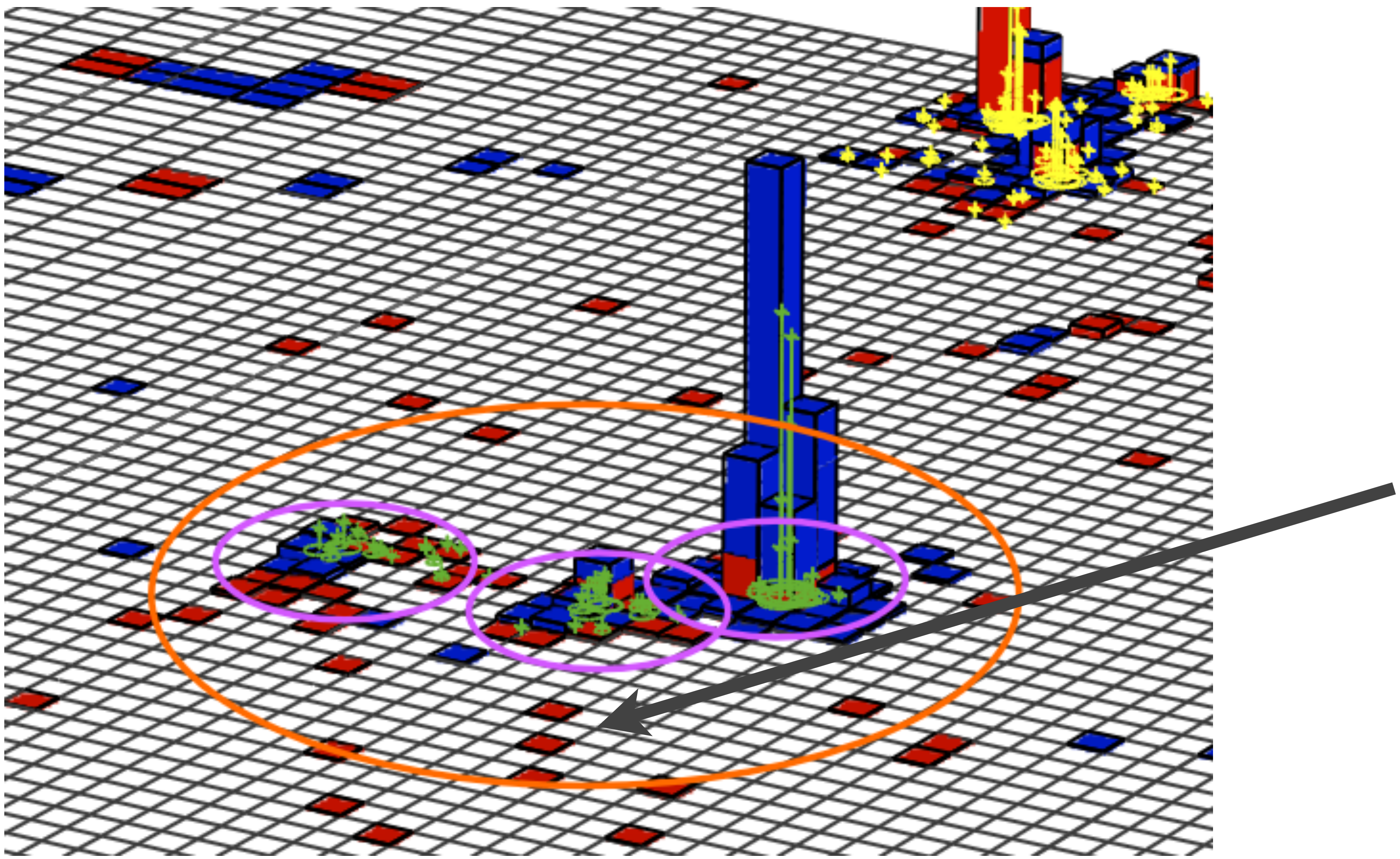
Initially in the context of $H \to b\bar{b}$ / $t \to bW$



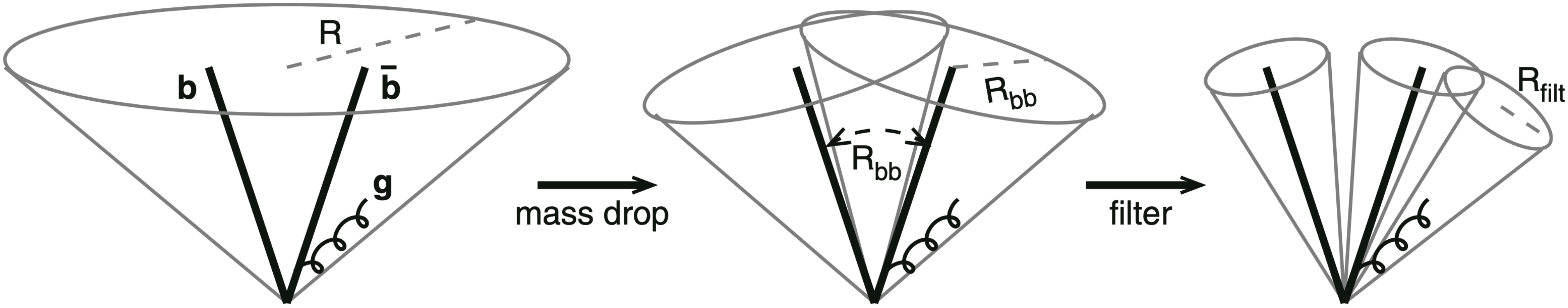


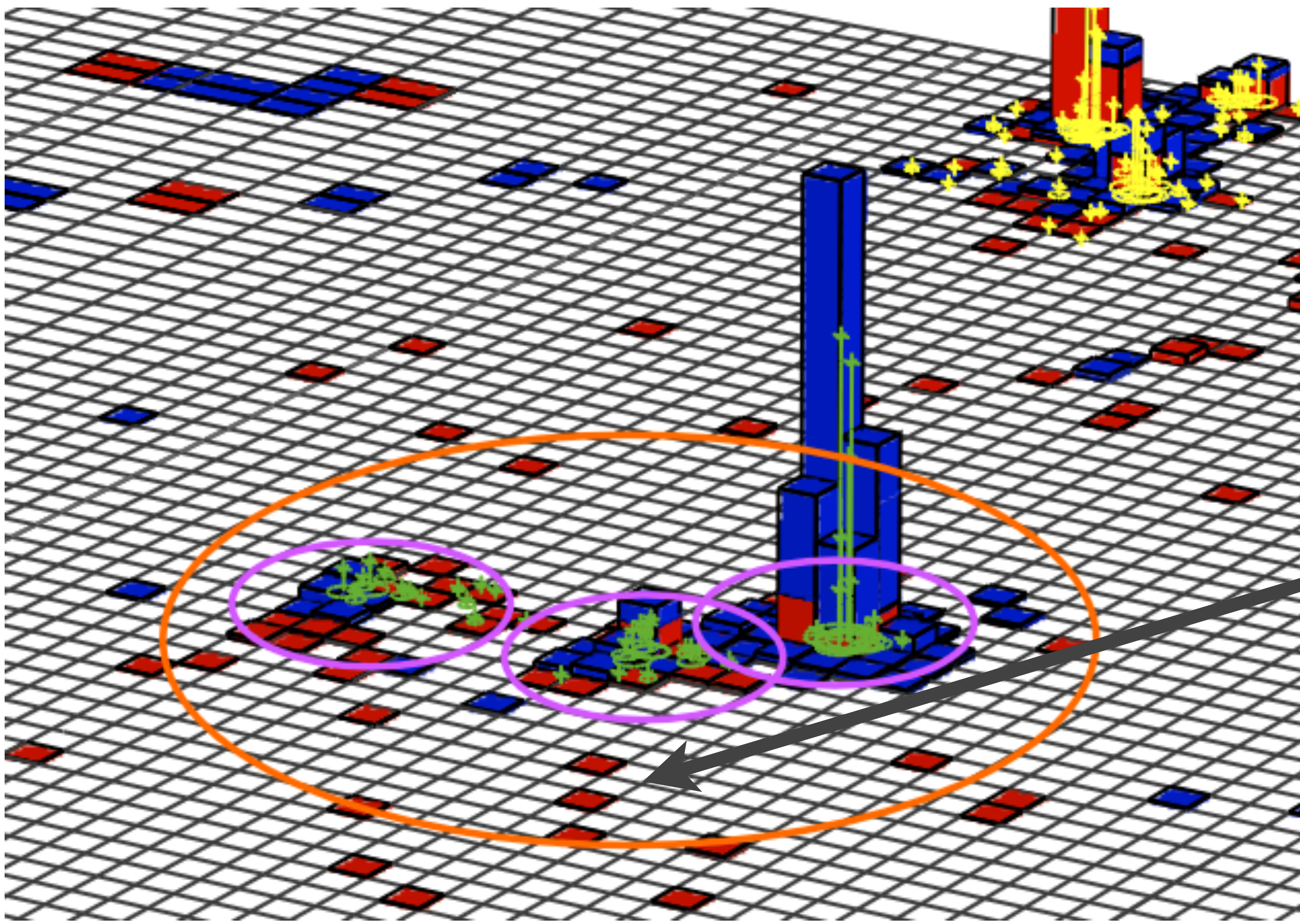
Initially in the context of $H \to b\bar{b}$ / $t \to bW$





Initially in the context of $H \to b\bar{b}$ / $t \to bW$

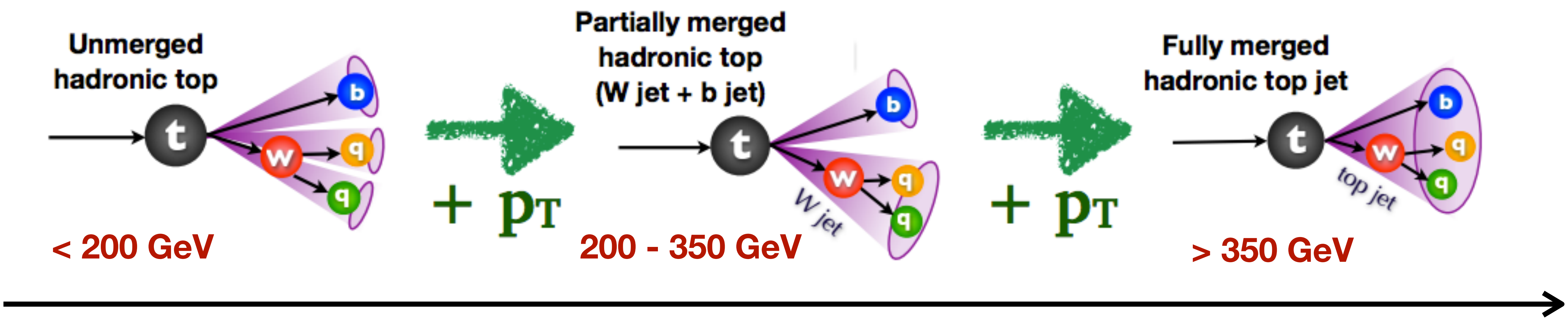


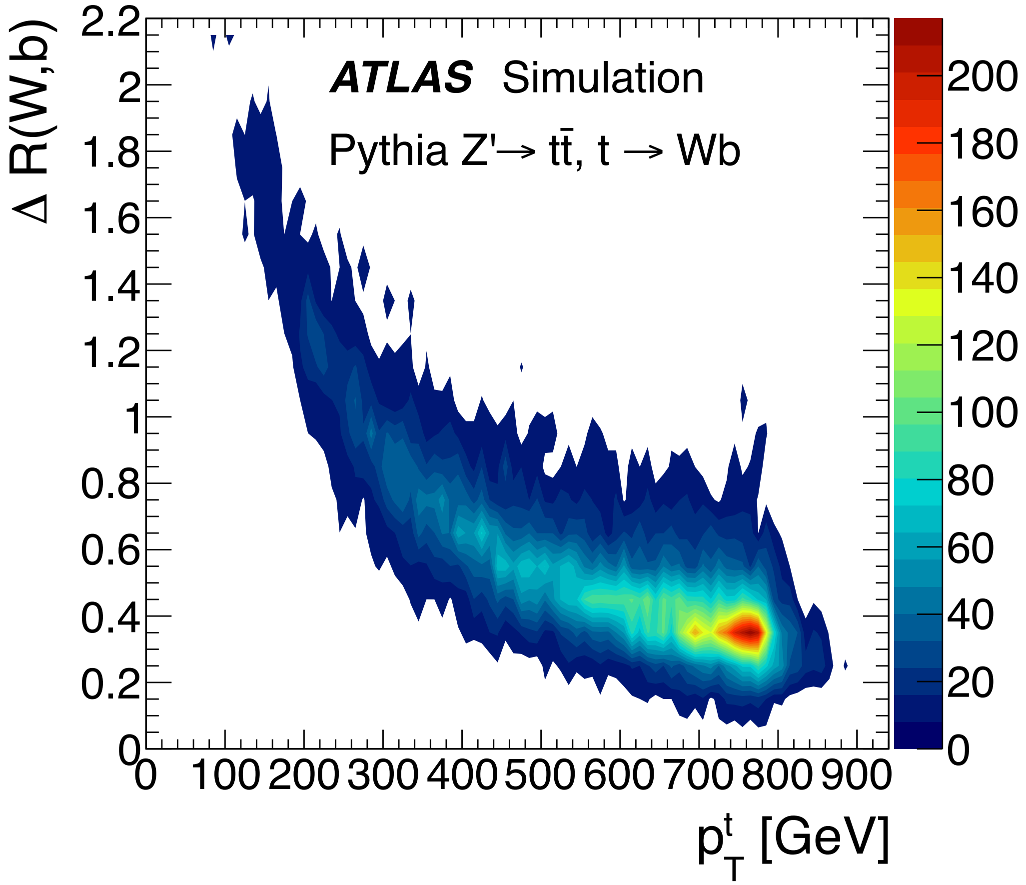


Better we get rid of this

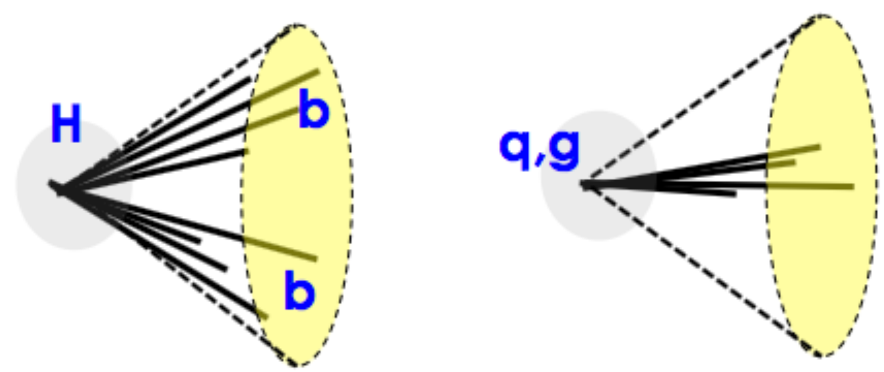
Feature of boosted particle decay

for $\Delta R \sim 1.0$





$\Delta R \sim 2m/p_T$

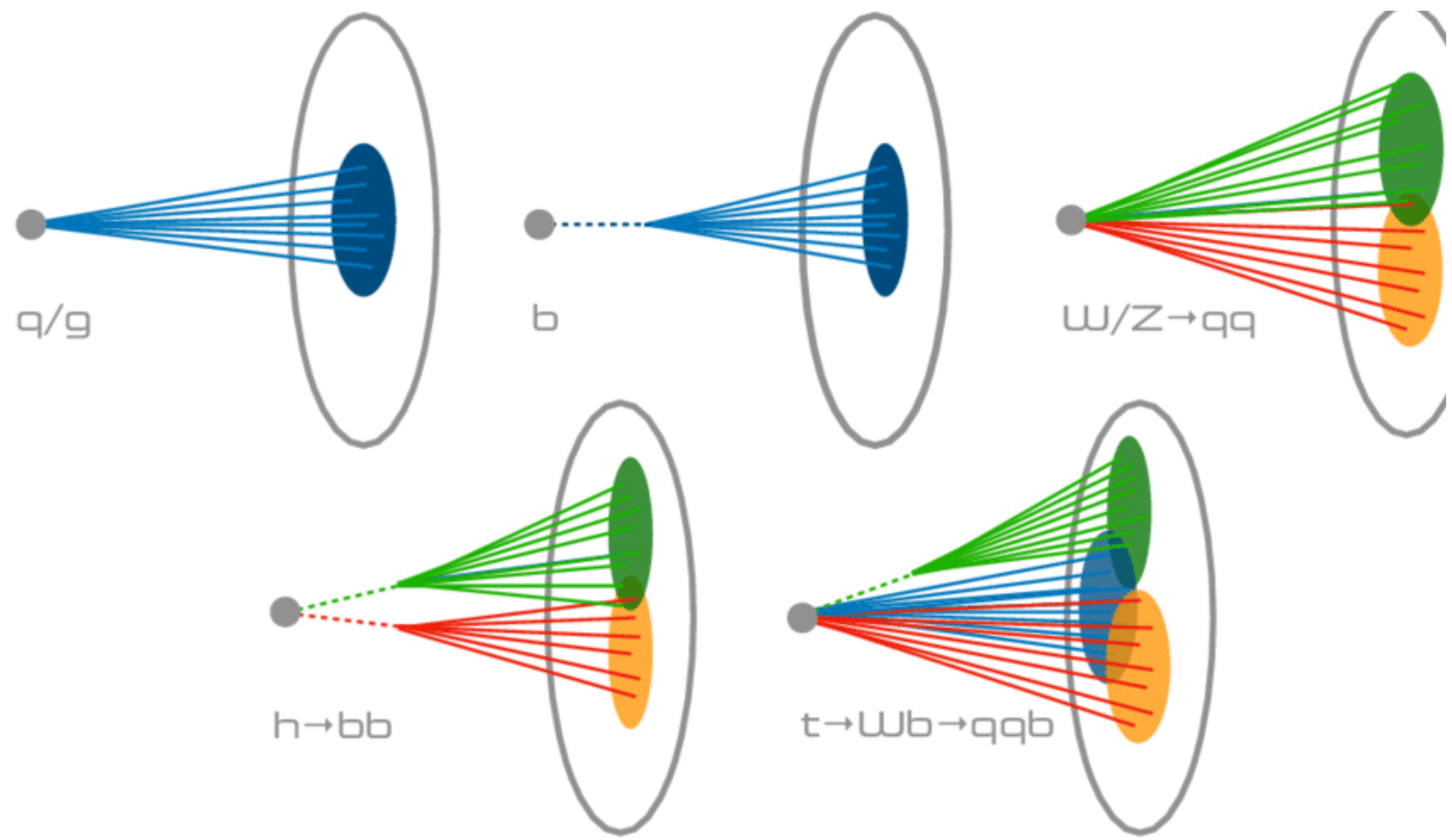


- Large R jets originating from decay of boosted heavy particle have different characteristics compared to fat jets from light flavor quarks or gluons.

- The major challenge with large R jets are heavy contamination due to pileup & large QCD background.

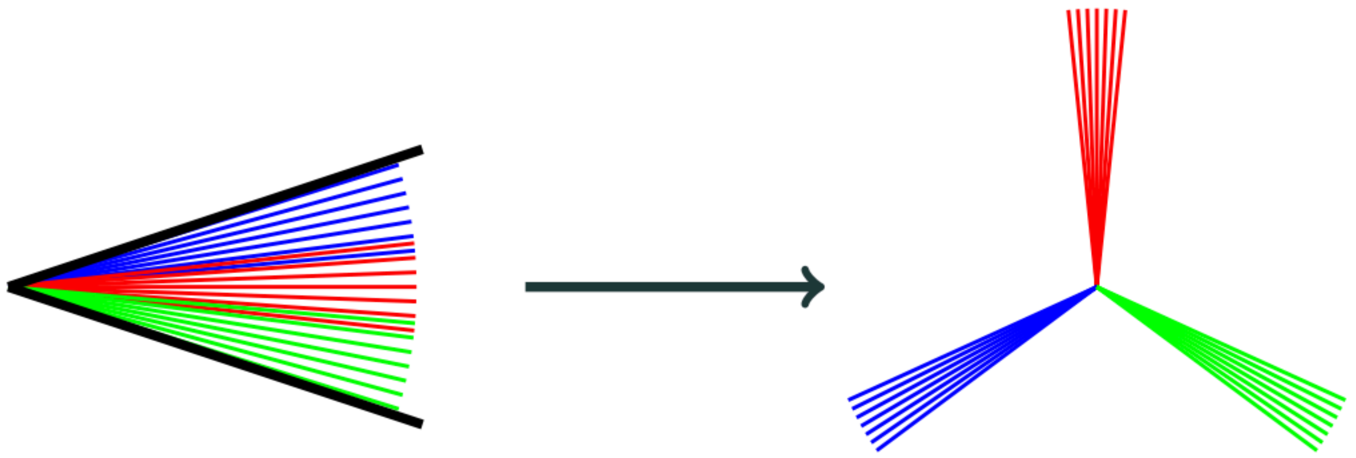
The generic class of substructures

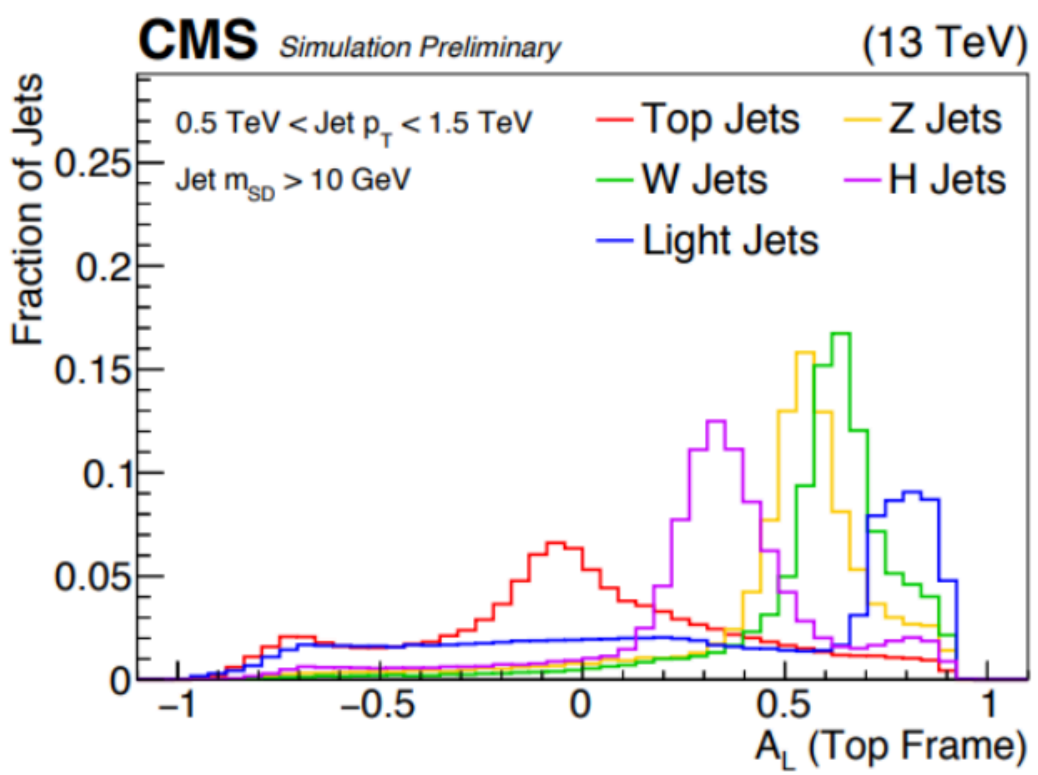
Category I: prong finders. Tools in this category exploit the fact that when a boosted massive object decays into partons, all the partons typically carry a sizeable fraction of the initial jet transverse momentum, resulting in multiple hard cores in the jet. Conversely, quark and gluon jets are dominated by the radiation of soft gluons, and are therefore mainly single-core jets. Prong finders therefore look for multiple hard cores in a jet, hence reducing the contamination from "standard" QCD jets. This is often used to characterise the boosted jets in terms of their "pronginess", i.e. to their expected number of hard cores: QCD jets would be 1-prong objects, W/Z/H jets would be two-pronged, boosted top



The generic class of substructures

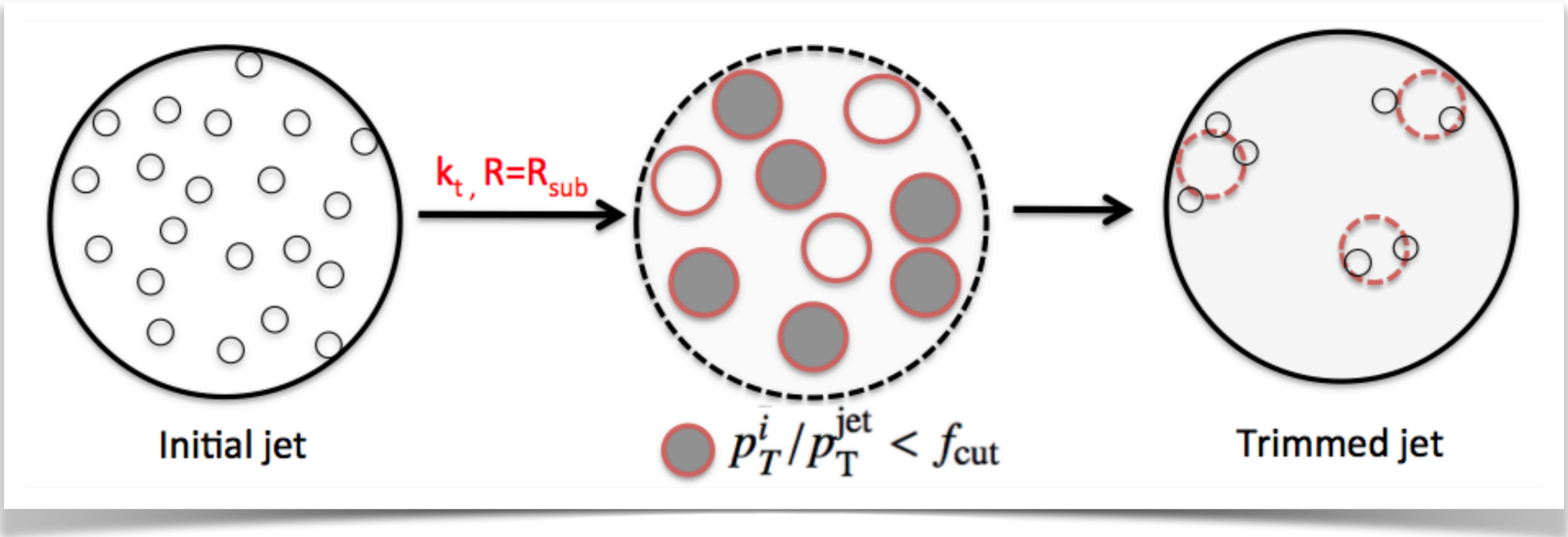
Category II: radiation constraints. The second main difference between signal and background jets is their colour structure. This means that signal and background jets will exhibit different soft-gluon radiation patterns. For example, QCD radiation associated with an EW-boson jet, which is colourless, it is expected to be less than what we typically find in a QCD jet. Similarly, quark-initiated jets are expected to radiate less soft gluons than gluon-initiated jets. Many jet shapes have been introduced to quantify the radiation inside a jet and hence separate signal jets from background jets.





$$A_L = \frac{\sum_{\text{jet}} p_z^{\text{jet}}}{\sum_{\text{jet}} p^{\text{jet}}}$$

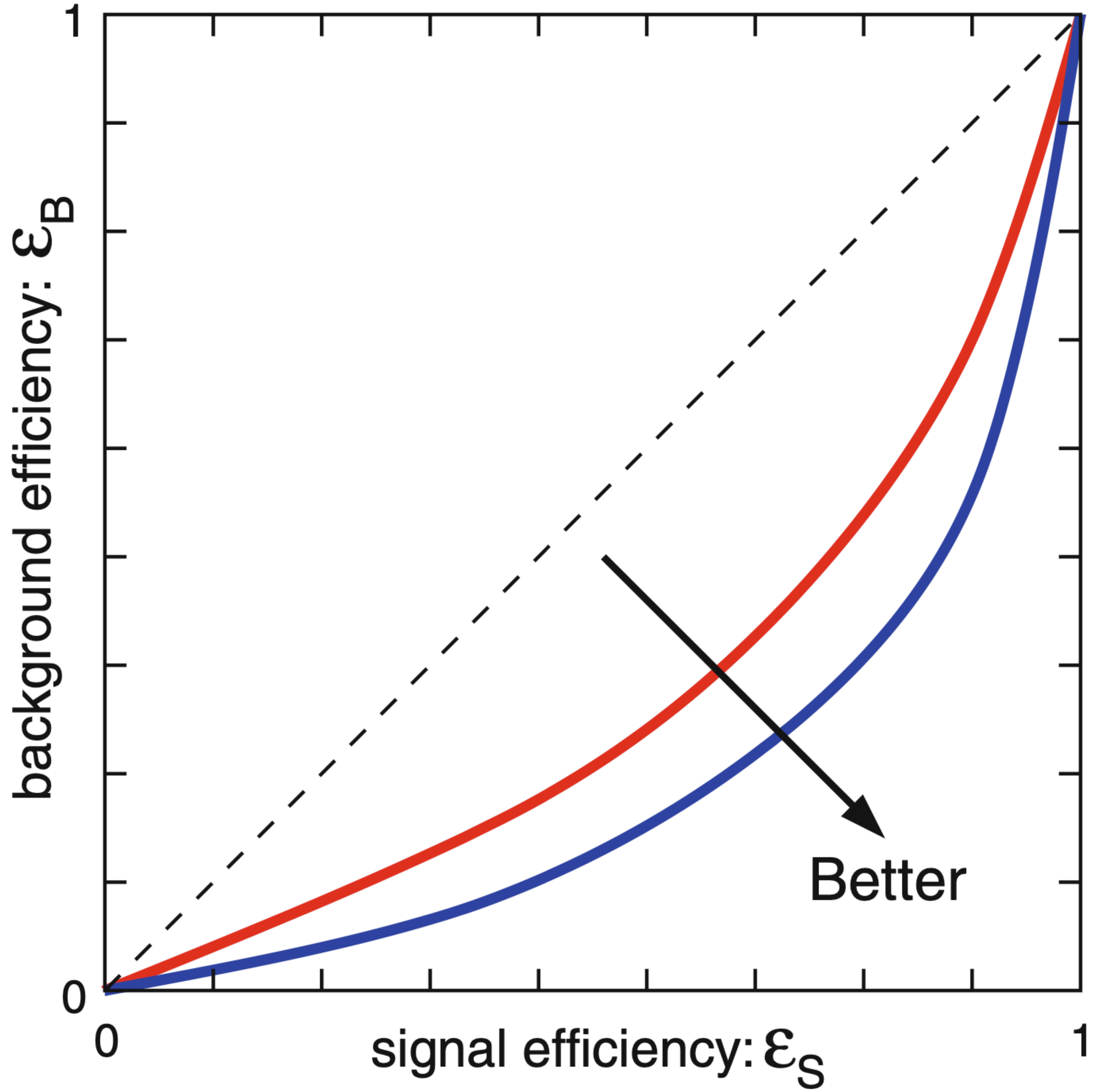
The generic class of substructures



Category III: groomers. There is a third category of widely-used tools related to the fact that one often use large-radius jets for substructure studies. As we have already discussed, because of their large area, these jets are particularly sensitive to soft backgrounds, such as the UE and pileup. "Grooming" tools have therefore been introduced to mitigate the impact of these soft backgrounds on the fat jets. These tools usually work by removing the soft radiation far from the jet axis, where it is the most likely to come from a soft contamination rather than from QCD radiation inside the jet. In many respects, groomers share similarities with prong finders, essentially due to the fact that removing soft contamination and keeping the hard prongs are closely related.

How do we quantify a tagger performance?

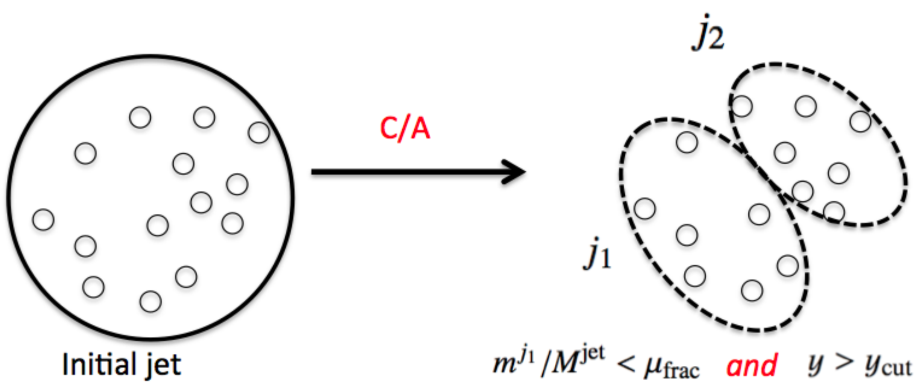
Receiver Operator Characteristics

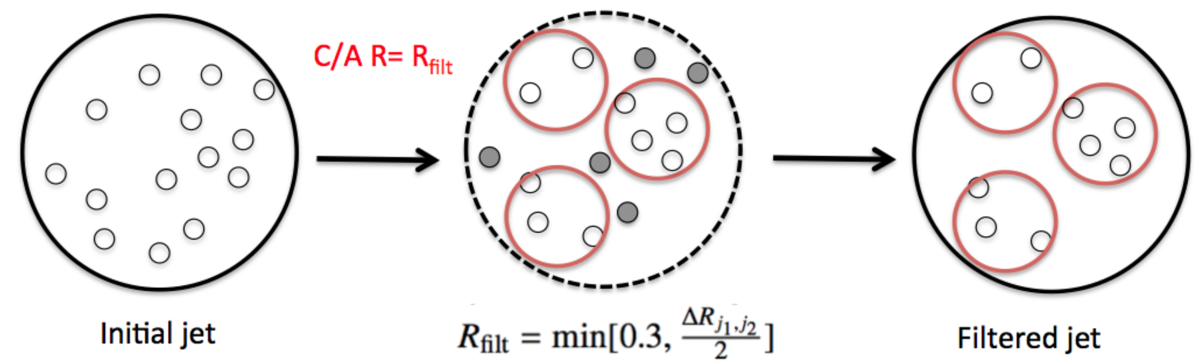


Cleaning large-R jets

The first step is to clean the hadronic jets via grooming : mass drop, trimming & pruning

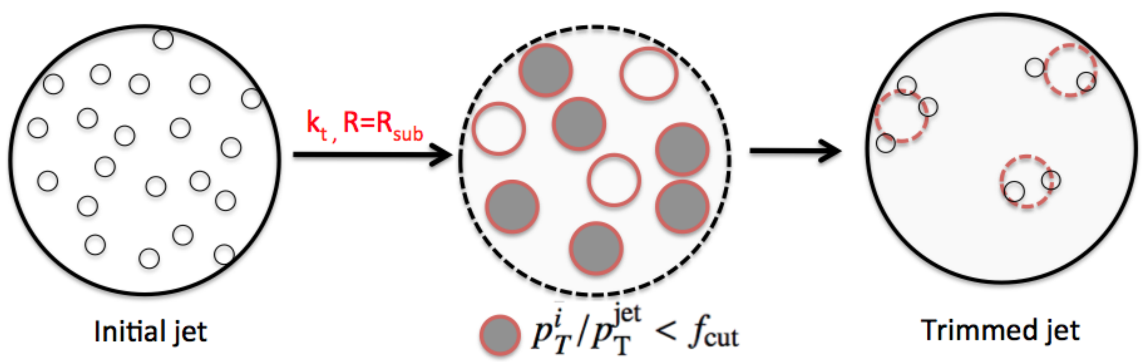
Mass Drop : It seeks to isolate concentrations of energy within a jet by identifying relatively symmetric sub-jets.



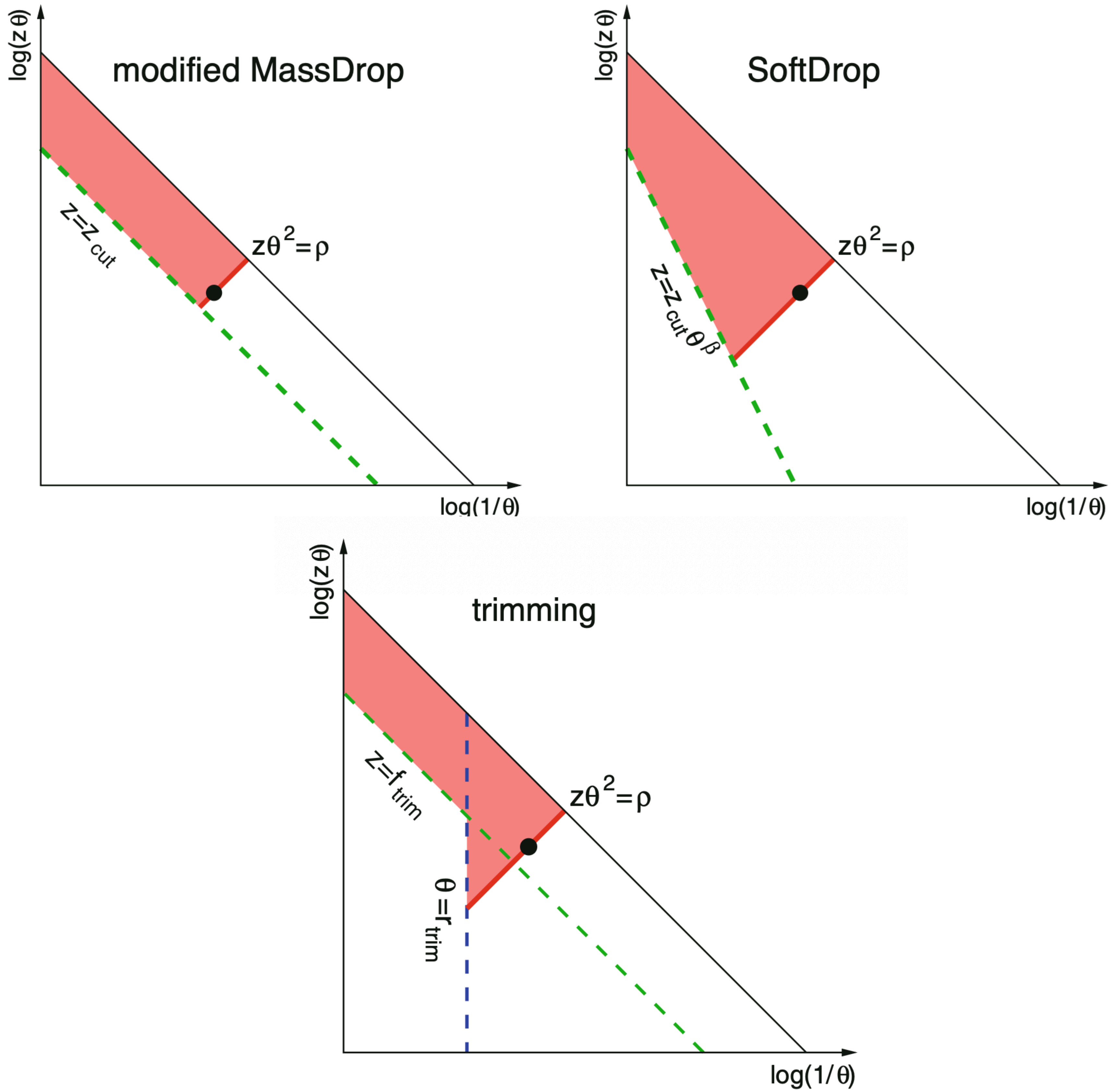


$$\mathbf{y = \frac{\min[(p_T^{j1})^2, (p_T^{j2})^2]}{(m_{jet})^2} \times \Delta R^2_{j1,j2}}$$

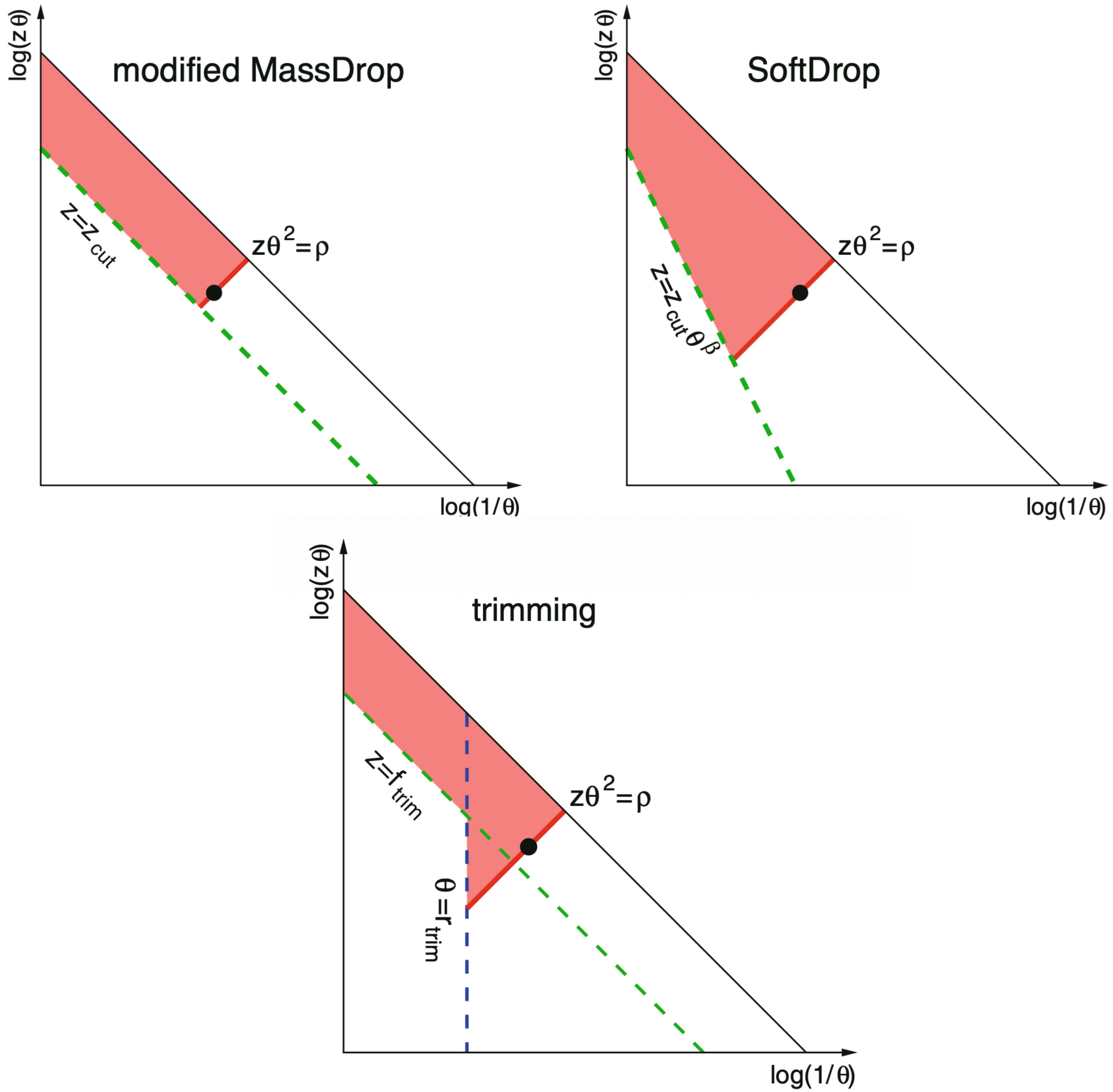
Trimming : Based on the idea that contamination from pileup, ISR, MPI in a jet should be much softer than the hard scattered constituents.



Each tagger kills a region in Lund plane



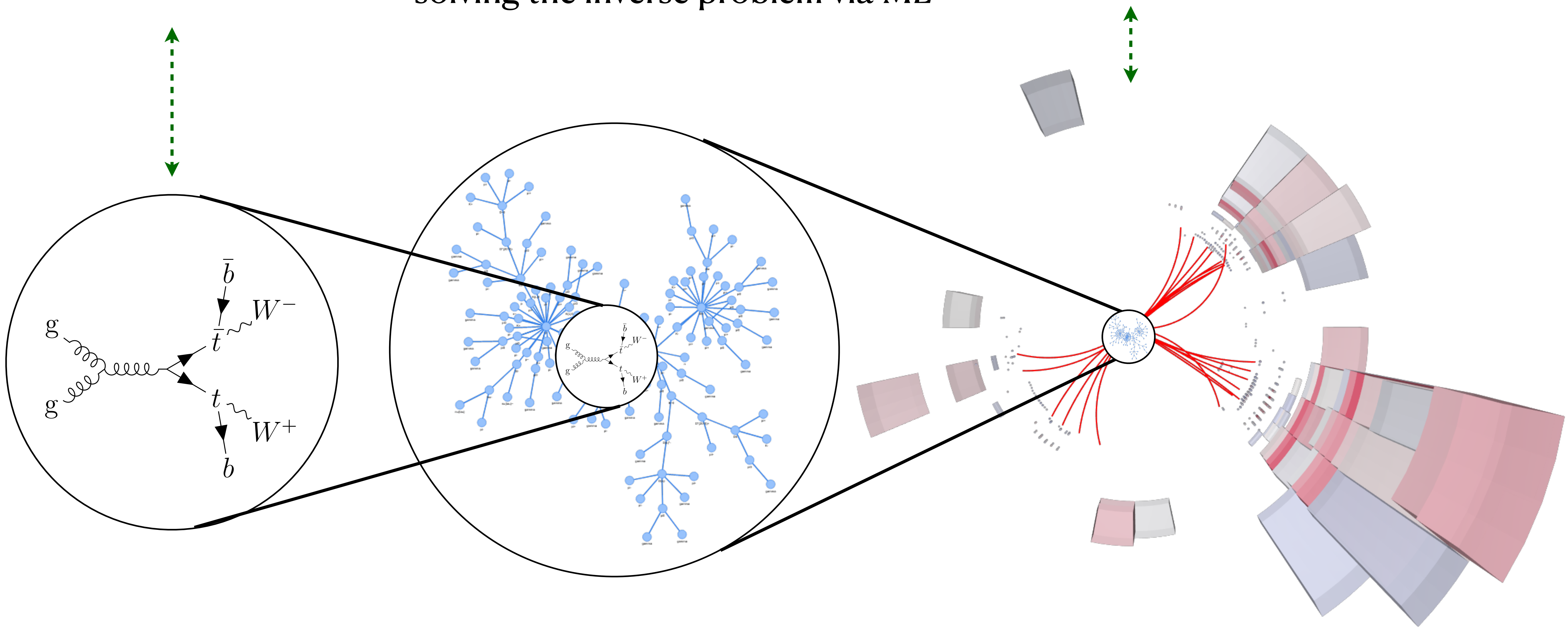
Each tagger kills a region in Lund plane



Why do we want to do ML with jets?

A broad strategy towards physics inference

Guess the Lagrangian ← solving the inverse problem via ML — $\frac{1}{\mathscr{L}}\frac{d^2N}{dp_T d\eta}$

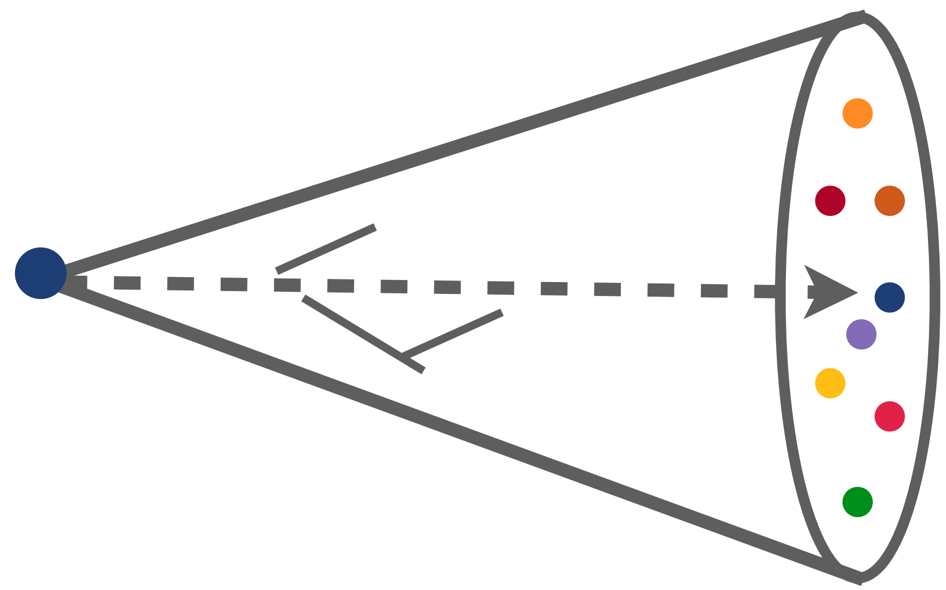


The physics at the core:
driven by the interaction between quantum fields, computed via perturbative or lattice techniques.

Final Particles
Produced via hadronization (no first principle analytic techniques are available)

Detector output/Readout
Produced via hardware or simulation

ML@Colliders : what's the broad task?

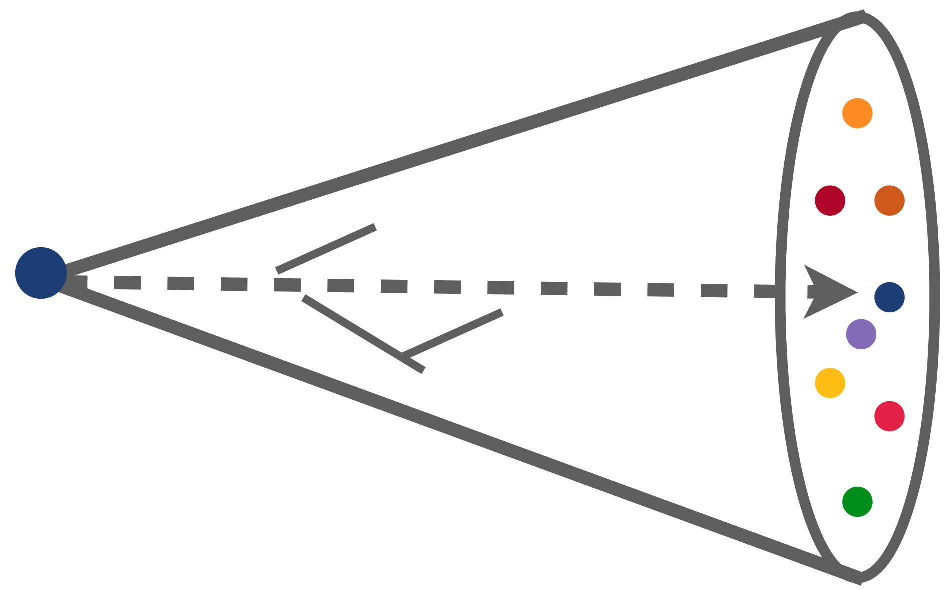
$f_{\{\theta\}}\Big($  $\Big) \longrightarrow \hat{y} = f_{\{\theta\}}(X)$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

ML@Colliders : what's the broad task?

1. Decide the right representation of the data (images/graphs/trees..)

$f_{\{\theta\}}\Big($



$\Big) \rightarrow \hat{y} = f_{\{\theta\}}(X)$

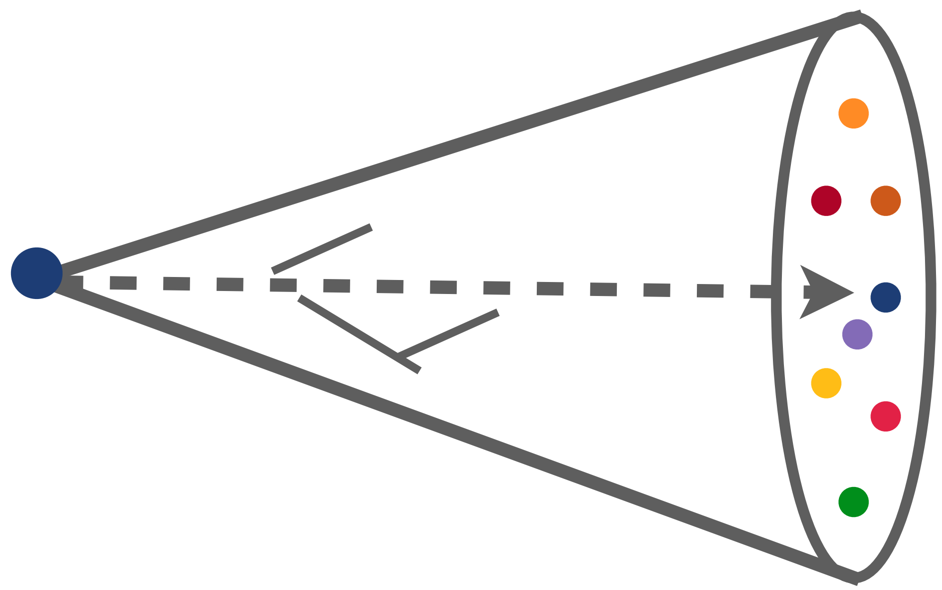
$$L(y, \hat{y}) \equiv L(\{\theta\})$$

ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

$$f_{\{\theta\}} \Big($$



$$\Big) \rightarrow \hat{y} = f_{\{\theta\}}(X)$$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

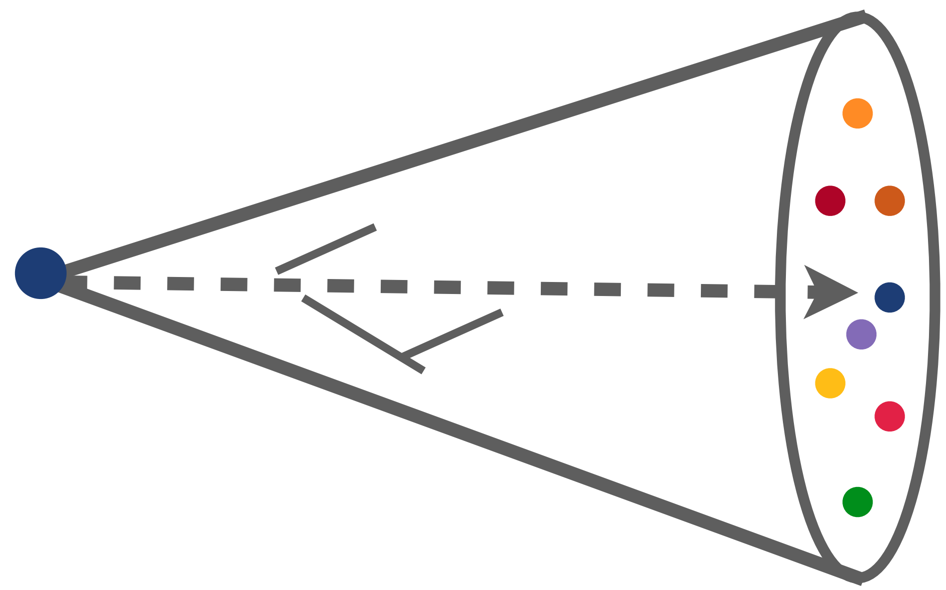
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$f_{\{\theta\}}$ (



) $\rightarrow$ $\hat{y} = f_{\{\theta\}}(X)$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

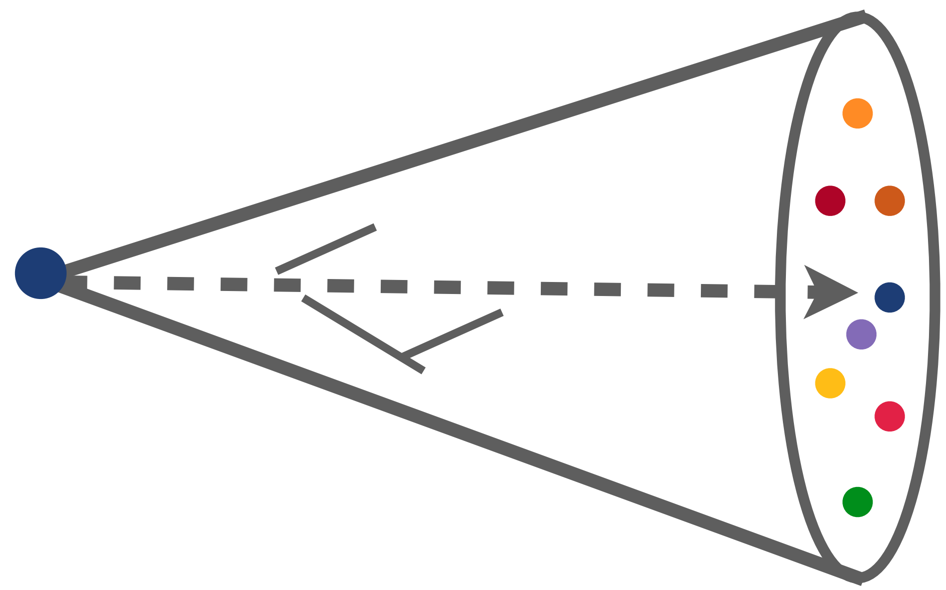
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$$f_{\{\theta\}}\Bigg($$



$$\Bigg) \rightarrow \hat{y} = f_{\{\theta\}}(X)$$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

Variation in data

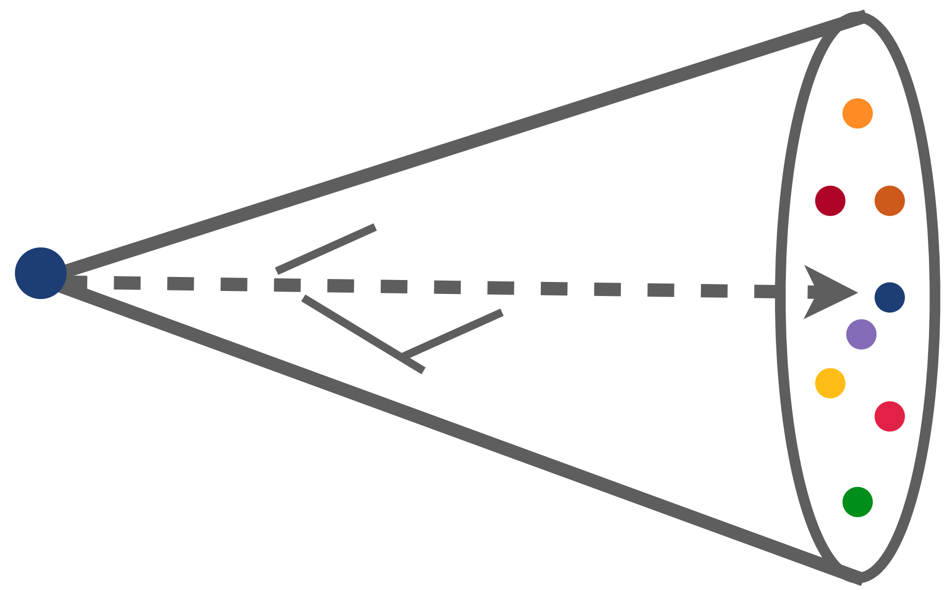
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$f_{\{\theta\}}$ (



) $\rightarrow$ $\hat{y} = f_{\{\theta\}}(X)$

$L(y, \hat{y}) \equiv L(\{\theta\})$

Variation in data

Unsupervised

No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE

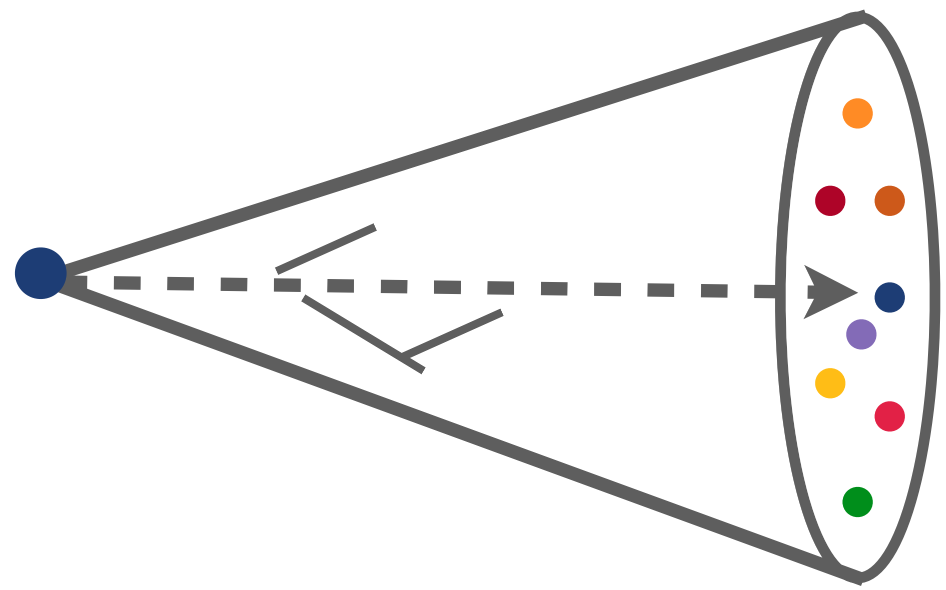
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$$f_{\{\theta\}} \Big($$



$$\Big) \rightarrow \hat{y} = f_{\{\theta\}}(X)$$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

Variation in data

Unsupervised

No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE

Semi-supervised

Noisy labels. estimate :

p(s-enriched)/p(s-depleted)

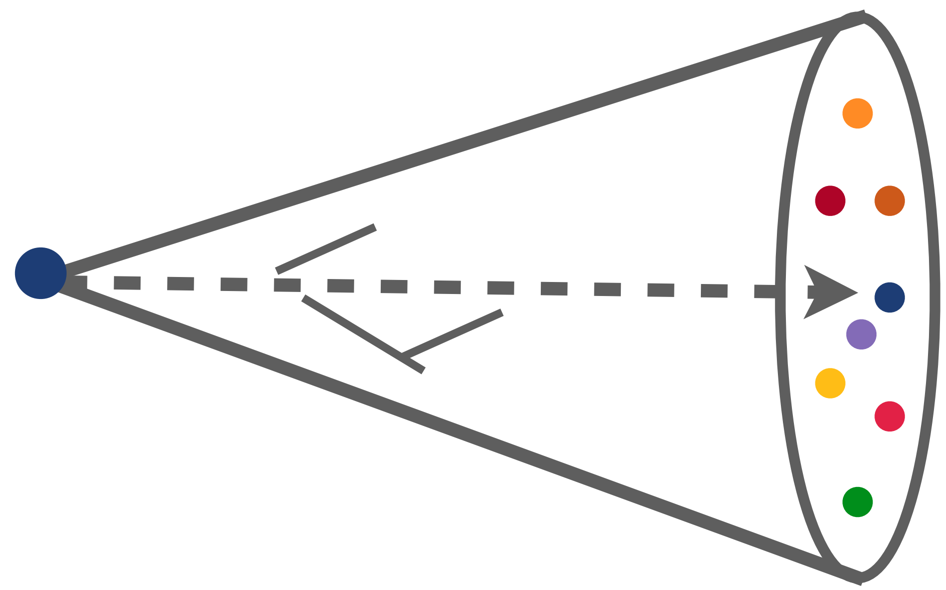
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$f_{\{\theta\}}$ (



) $\rightarrow$ $\hat{y} = f_{\{\theta\}}(X)$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

Variation in data

Unsupervised	Semi-supervised	Weakly-supervised
No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE	Noisy labels. estimate : p(s-enriched)/p(s-depleted)	Partial labels. e.g. simulating : SM bkg vs many NP signals.

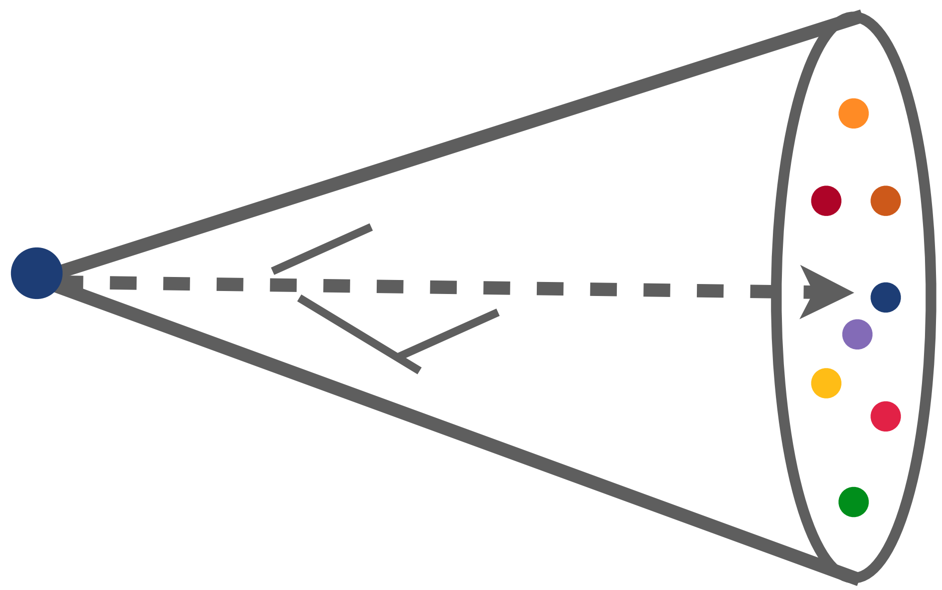
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$f_{\{\theta\}}$ (



) → $\hat{y} = f_{\{\theta\}}\,(X)$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

Variation in data

Unsupervised

No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE

Semi-supervised

Noisy labels. estimate :

p(s-enriched)/p(s-depleted)

Weakly-supervised

Partial labels. e.g. simulating : SM bkg vs many NP signals.

Supervised

Learning on all the well labeled data.

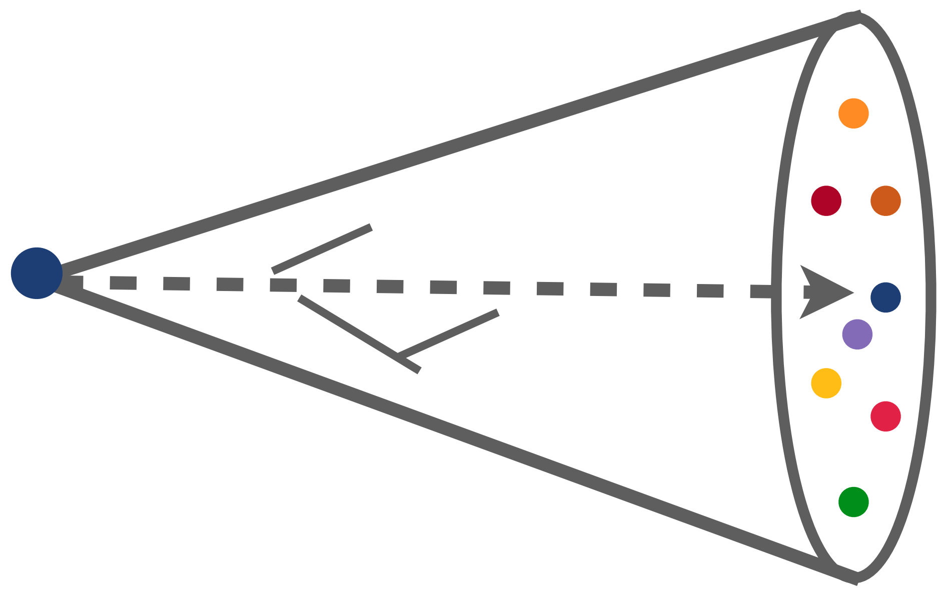
ML@Colliders : what's the broad task?

2. Choose a NN model (CNN/GNN/)

1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

$f_{\{\theta\}}$ (



) → $\hat{y} = f_{\{\theta\}}(X)$

$$L(y, \hat{y}) \equiv L(\{\theta\})$$

Self-supervised

Variation in data

Unsupervised

No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE

Semi-supervised

Noisy labels. estimate :

p(s-enriched)/p(s-depleted)

Weakly-supervised

Partial labels. e.g. simulating : SM bkg vs many NP signals.

Supervised

Learning on all the well labeled data.

Looking the problem through ML lens

2. Choose a NN model (CNN/GNN/)

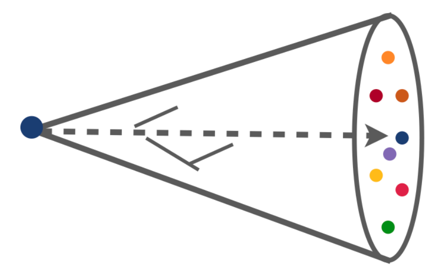
1. Decide the right representation of the data (images/graphs/trees..)

3. With a defined learning task, compute the loss function.

Uncertainty

Interpretability

$$f_{\{\theta\}} \Big(\quad \Big) \rightarrow \hat{y} = f_{\{\theta\}}(X)$$



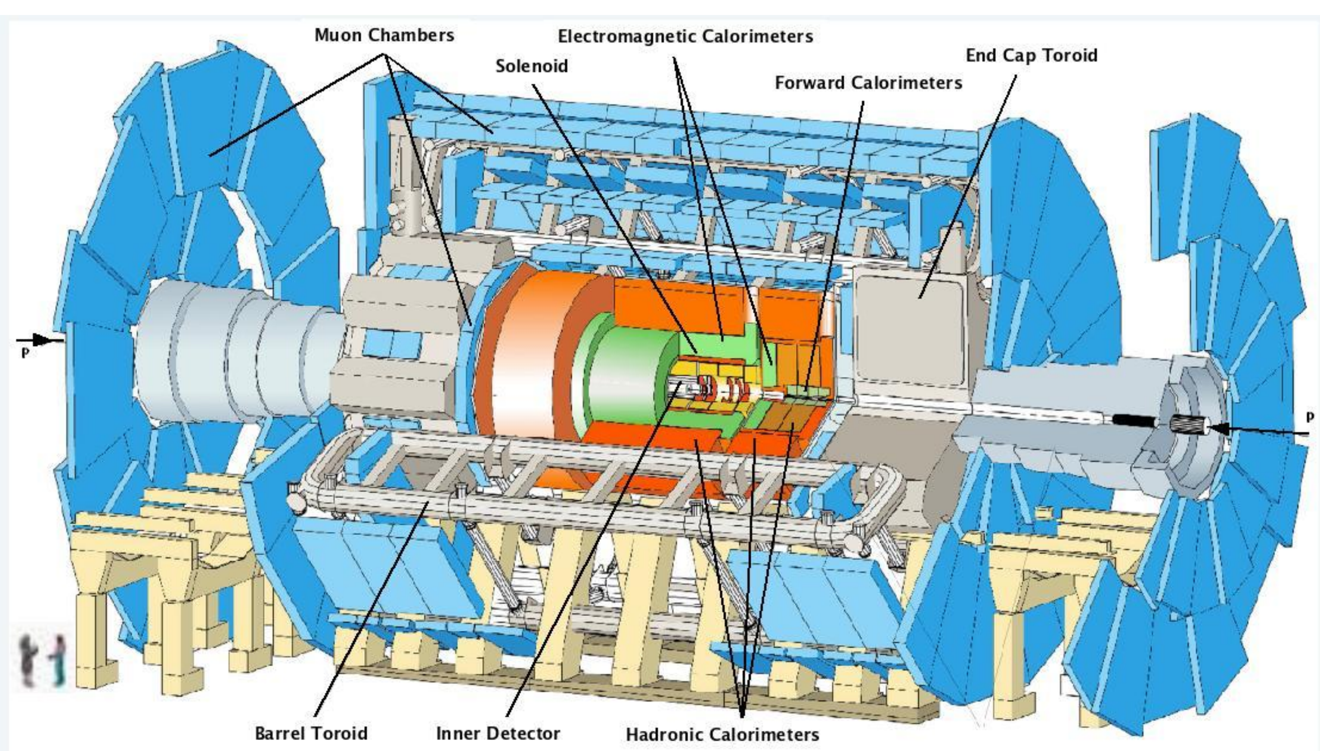
$$L(y, \hat{y}) \equiv L(\{\theta\})$$

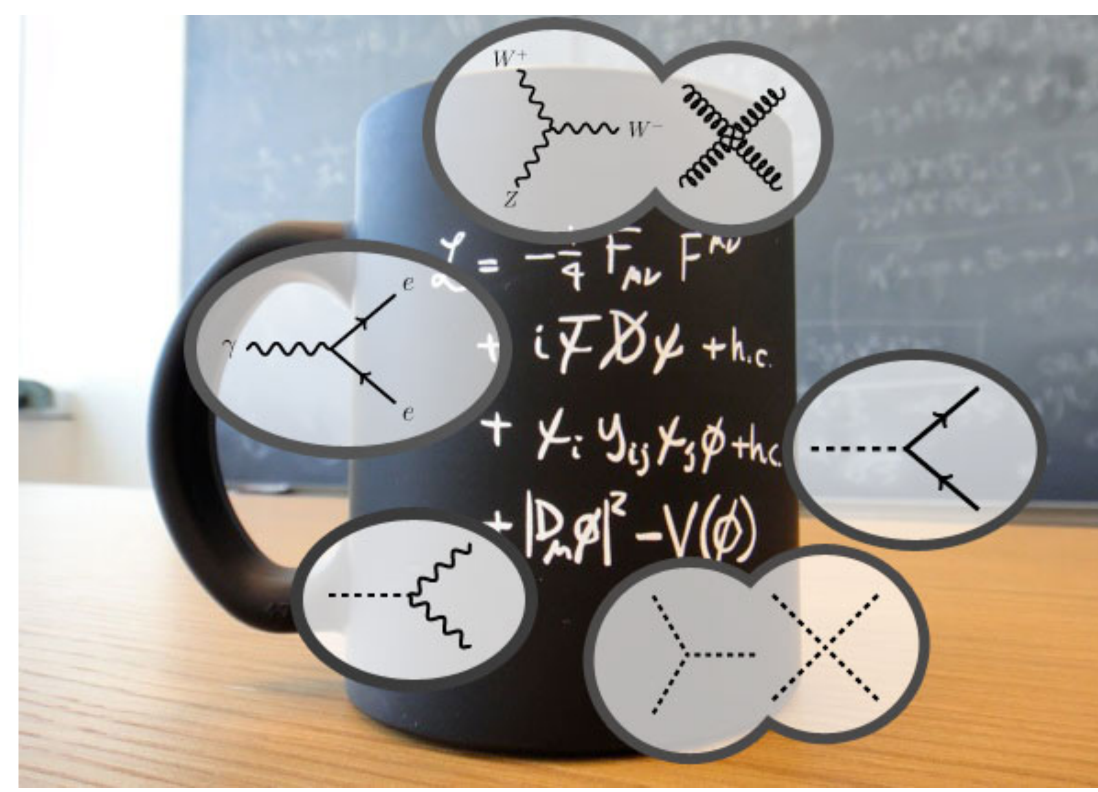
Self-supervised

Variation in data

Unsupervised	Semi-supervised	Weakly-supervised	Supervised
No-labels, the task is to figure out p(x) from which the data is drawn. e.g. VAE	Noisy labels. estimate : p(s-enriched)/p(s-depleted)	Partial labels. e.g. simulating : SM bkg vs many NP signals.	Learning on all the well labeled data.

Figure from : https://atlas-public.web.cern.ch/





ML 101

https://arxiv.org/pdf/2211.01421

$$x \to x^{(1)} \to x^{(2)} \cdots \to x^{(N)} \equiv f_\theta(x)$$

ML 101

https://arxiv.org/pdf/2211.01421

$$x \to x^{(1)} \to x^{(2)} \cdots \to x^{(N)} \equiv f_\theta(x)$$

$$x^{(n-1)} \to x^{(n)} := W^{(n)} x^{(n-1)} + b^{(n)}$$

ML 101

https://arxiv.org/pdf/2211.01421

$$x \to x^{(1)} \to x^{(2)} \cdots \to x^{(N)} \equiv f_\theta(x)$$

$$x^{(n-1)} \to x^{(n)} := W^{(n)} x^{(n-1)} + b^{(n)}$$

$$x^{(n-1)} \to x^{(n)} := \mathrm{ReLU}\left[W^{(n)} x^{(n-1)} + b^{(n)}\right]$$

ML 101

https://arxiv.org/pdf/2211.01421

$$x \to x^{(1)} \to x^{(2)} \cdots \to x^{(N)} \equiv f_\theta(x) \qquad x^{(n-1)} \to x^{(n)} := W^{(n)} x^{(n-1)} + b^{(n)}$$

$$x^{(n-1)} \to x^{(n)} := \mathrm{ReLU}\left[W^{(n)} x^{(n-1)} + b^{(n)}\right]$$

$$\begin{aligned}
\frac{d\mathcal{L}}{dW_{1j}^{(N)}} &= \frac{\partial \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right|^2}{\partial W_{1j}^{(N)}} \\
&= \frac{\partial \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right|^2}{\partial \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}]} \frac{\partial \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}]}{\partial [W_{1k}^{(N)} x_k^{(N-1)}]} \frac{\partial [W_{1k}^{(N)} x_k^{(N-1)}]}{\partial W_{1j}^{(N)}} \\
&= -2 \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right| \times 1 \times \delta_{jk} x_k^{(N-1)} \\
&\equiv -2\sqrt{\mathcal{L}}\, x_j^{(N-1)} \,,
\end{aligned}$$

ML 101

https://arxiv.org/pdf/2211.01421

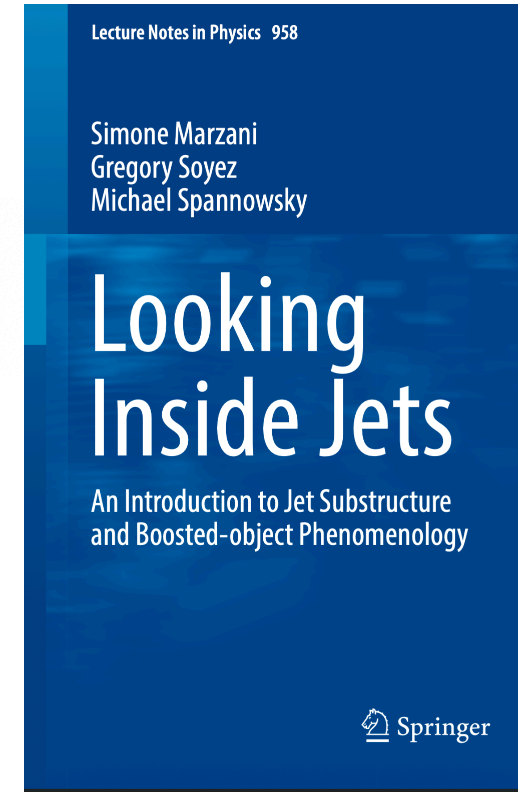
$$x \to x^{(1)} \to x^{(2)} \cdots \to x^{(N)} \equiv f_\theta(x) \qquad x^{(n-1)} \to x^{(n)} := W^{(n)} x^{(n-1)} + b^{(n)}$$

$$x^{(n-1)} \to x^{(n)} := \mathrm{ReLU}\left[W^{(n)} x^{(n-1)} + b^{(n)}\right]$$

$$\begin{aligned}
\frac{d\mathcal{L}}{dW_{1j}^{(N)}} &= \frac{\partial \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right|^2}{\partial W_{1j}^{(N)}} \\
&= \frac{\partial \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right|^2}{\partial \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}]} \frac{\partial \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}]}{\partial [W_{1k}^{(N)} x_k^{(N-1)}]} \frac{\partial [W_{1k}^{(N)} x_k^{(N-1)}]}{\partial W_{1j}^{(N)}} \\
&= -2 \left| f - \mathrm{ReLU}[W_{1k}^{(N)} x_k^{(N-1)}] \right| \times 1 \times \delta_{jk} x_k^{(N-1)} \\
&\equiv -2\sqrt{\mathcal{L}}\, x_j^{(N-1)} ,
\end{aligned}$$

$$\theta_j^{t+1} = \theta_j^t - \alpha \frac{\left\langle \dfrac{\partial \mathcal{L}^t}{\partial \theta_j} \right\rangle}{\epsilon + \sqrt{\left\langle \dfrac{\partial \mathcal{L}^t}{\partial \theta_j} \right\rangle^2}}$$

What do we want to tag?

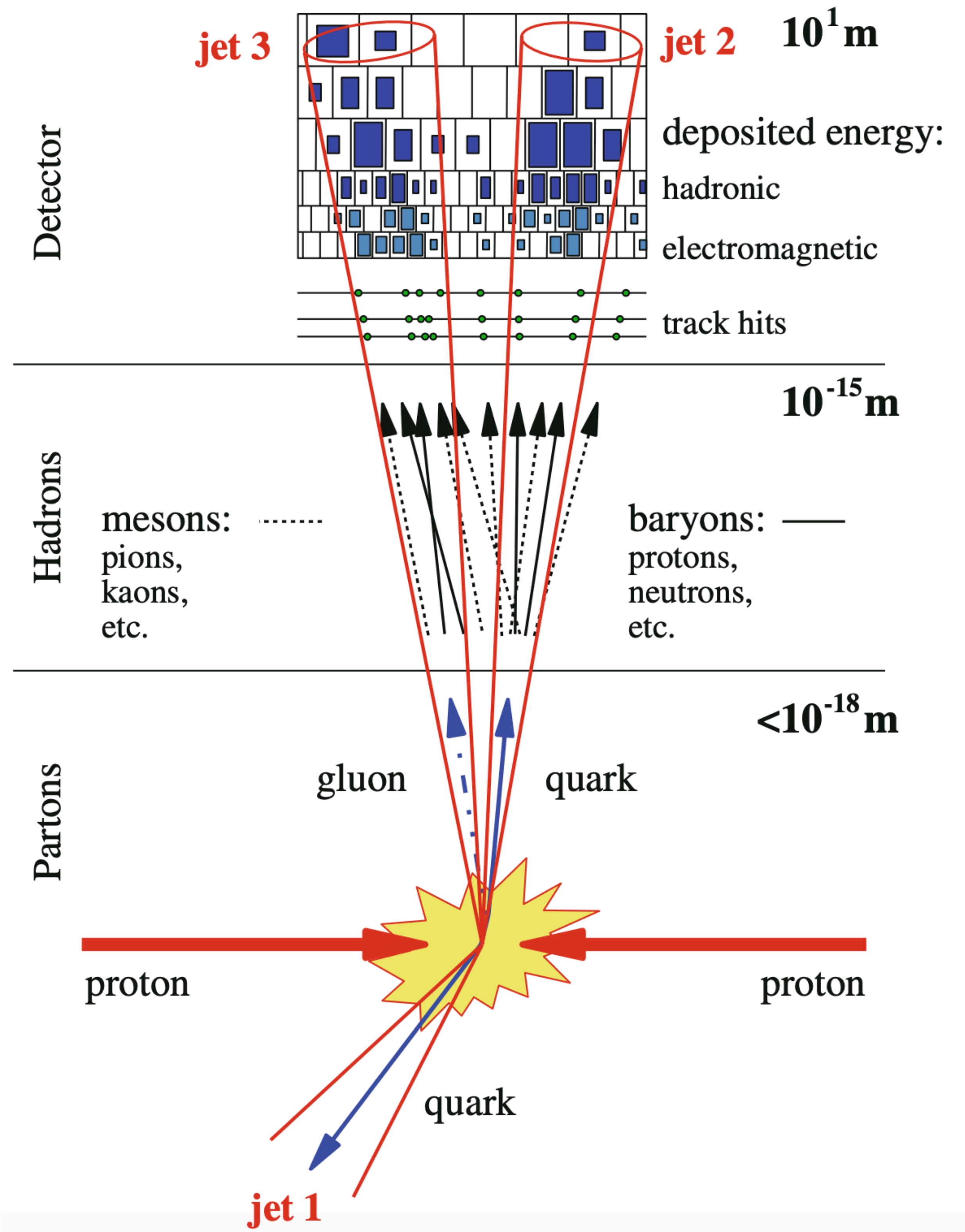


What is a Quark Jet?

From lunch/dinner discussions

Ill-Defined ↓ Well-Defined	What people sometimes think we mean ↑ What we mean	
		A quark parton
		A Born-level quark parton
		The initiating quark parton in a final state shower
	Quark as noun	An eikonal line with baryon number 1/3 and carrying triplet color charge
		A quark operator appearing in a hard matrix element in the context of a factorization theorem
		A parton-level jet object that has been quark-tagged using a soft-safe flavored jet algorithm (automatically collinear safe if you sum constituent flavors)
	Quark as adjective	A phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks (as interpreted by some suitable, though fundamentally ambiguous, criterion)

Why jet tagging is difficult?



The Large Hadron Collider
Editors: Thomas Schörner-Sadenius

$$\{p_1, p_2, \ldots, p_n\}$$

Jet Algorithm (for CA, kT, anti-kT)

$$\{j_1, j_2, \ldots, j_k\}$$

$$\{p_1, p_2, \ldots, p_n\} = F(q)$$

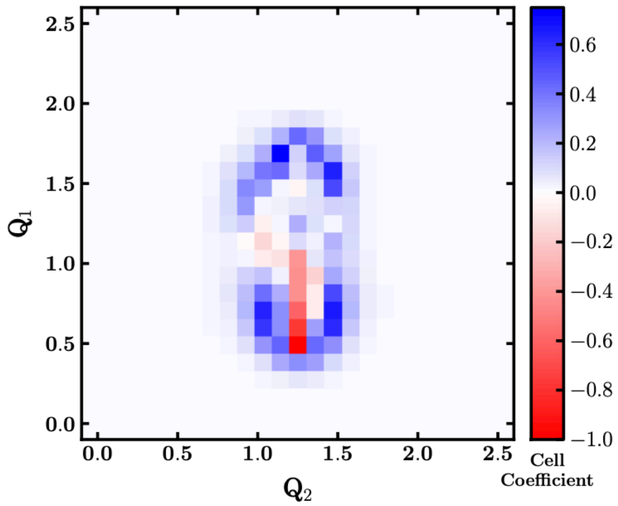
The forward problem is not computable from first principle

The question of jet tagging is how do we define the inverse problem?

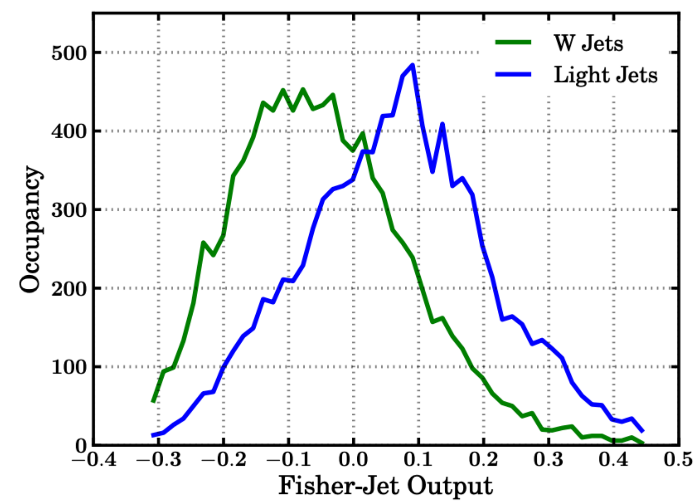
$$q = F^{-1}\Big(\{p_1, p_2, \ldots, p_n\}\Big) \ ?$$

Early jet tagging

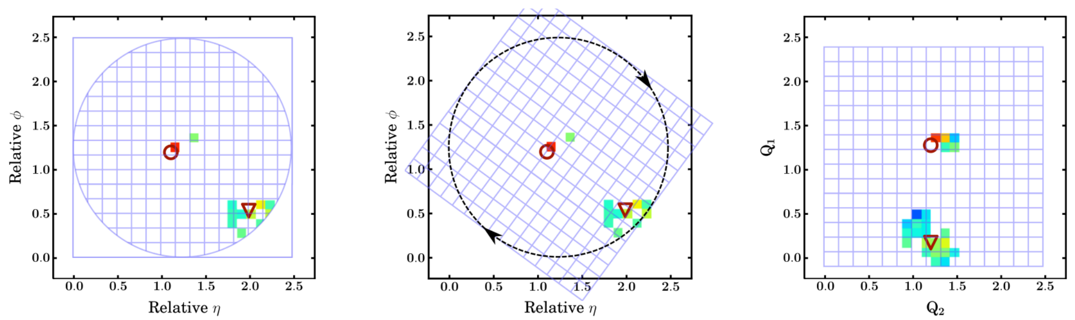
J. Cogan et-al JHEP 02 (2015) 118



(a) Fisher-Jet

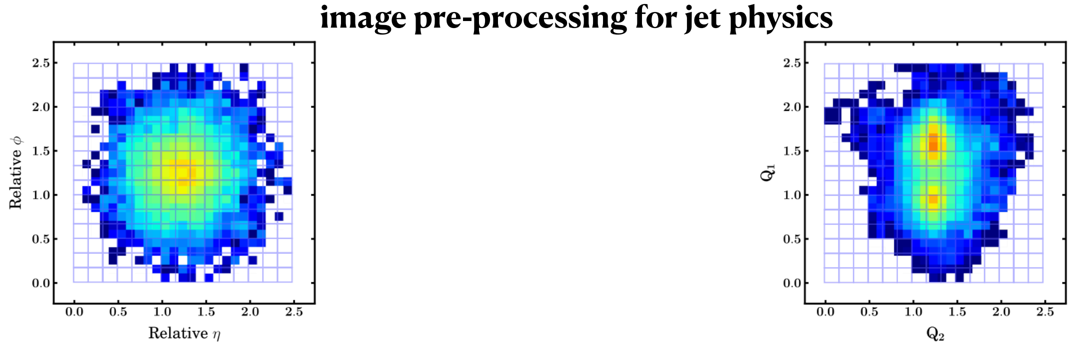


(b) Fisher-Jet Discriminant Output

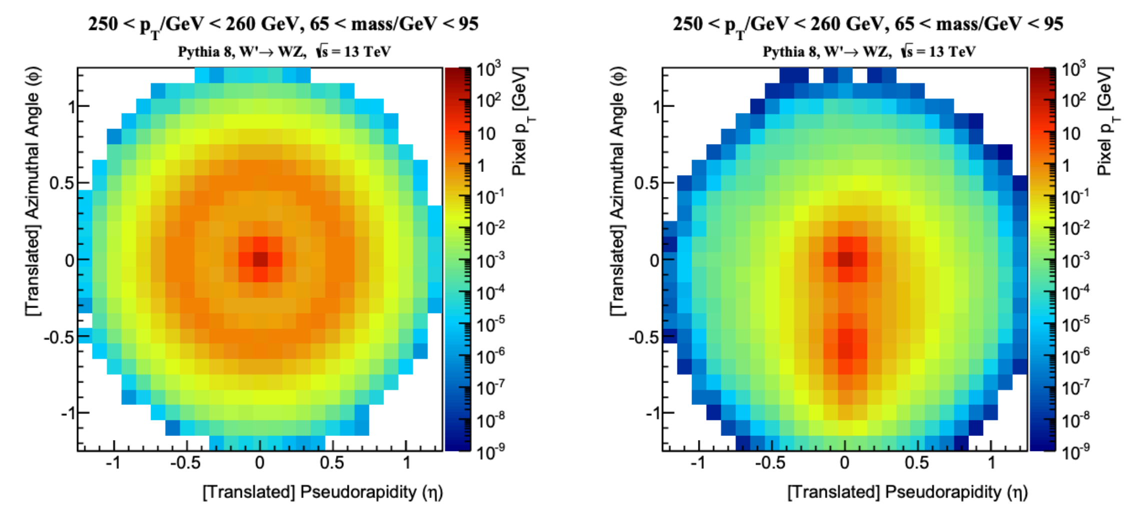


(a) Jet-image prior to rotation (b) Rotated pixel grid (c) Jet-image after projectio onto rotated grid, before transla tion

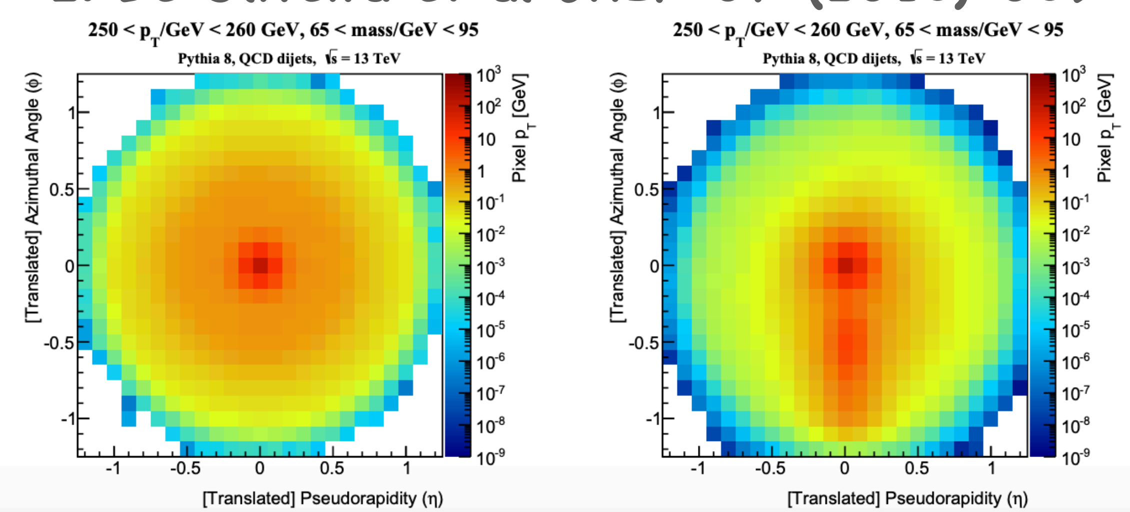
The first paper to discuss image pre-processing for jet physics

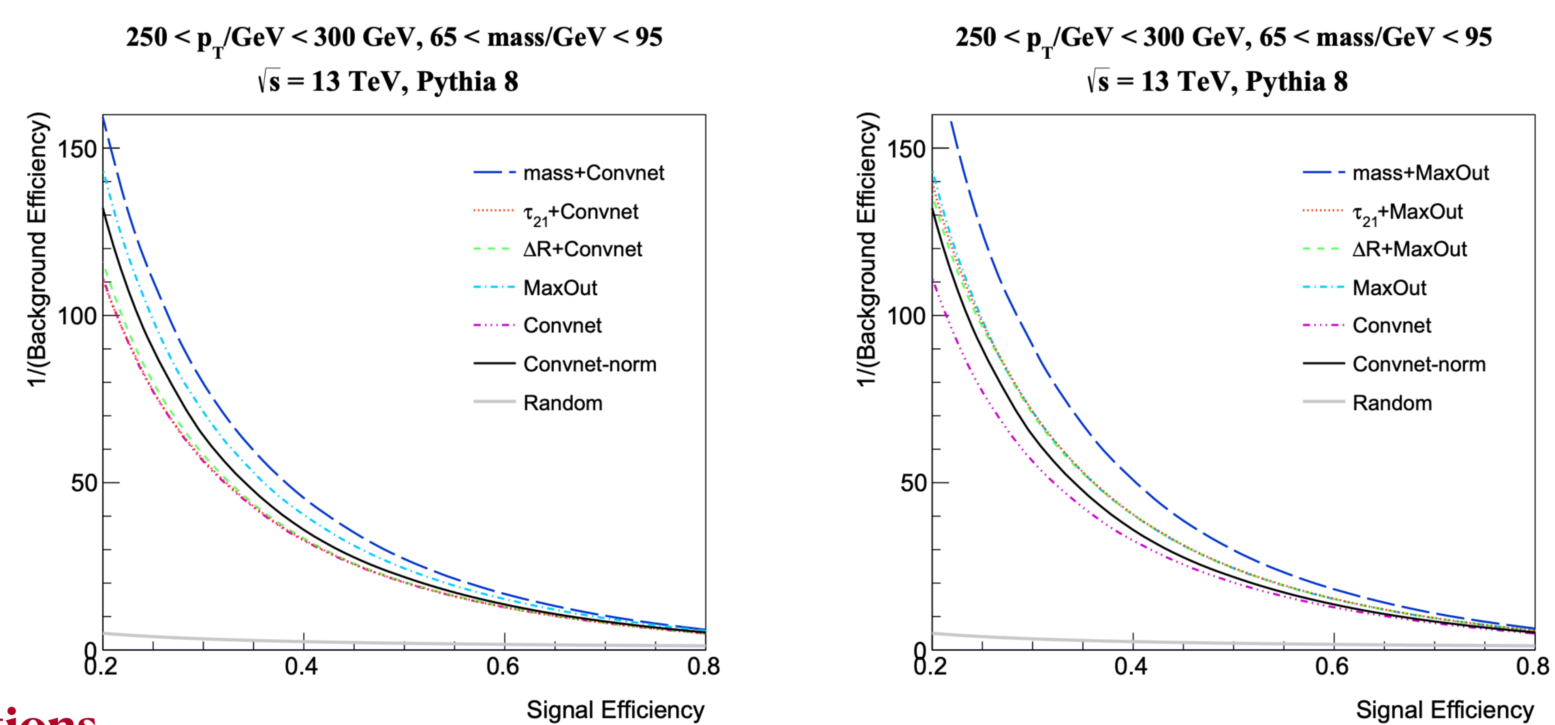


(d) Average jet-image, prior to rotation (e) Average jet-image, after pre processing



L. De Oliveira et-al JHEP 07 (2016) 069





Similar methods were applied for particle identifications

Different possible representations

The energy deposition pattern can be thought as multi-layer image.

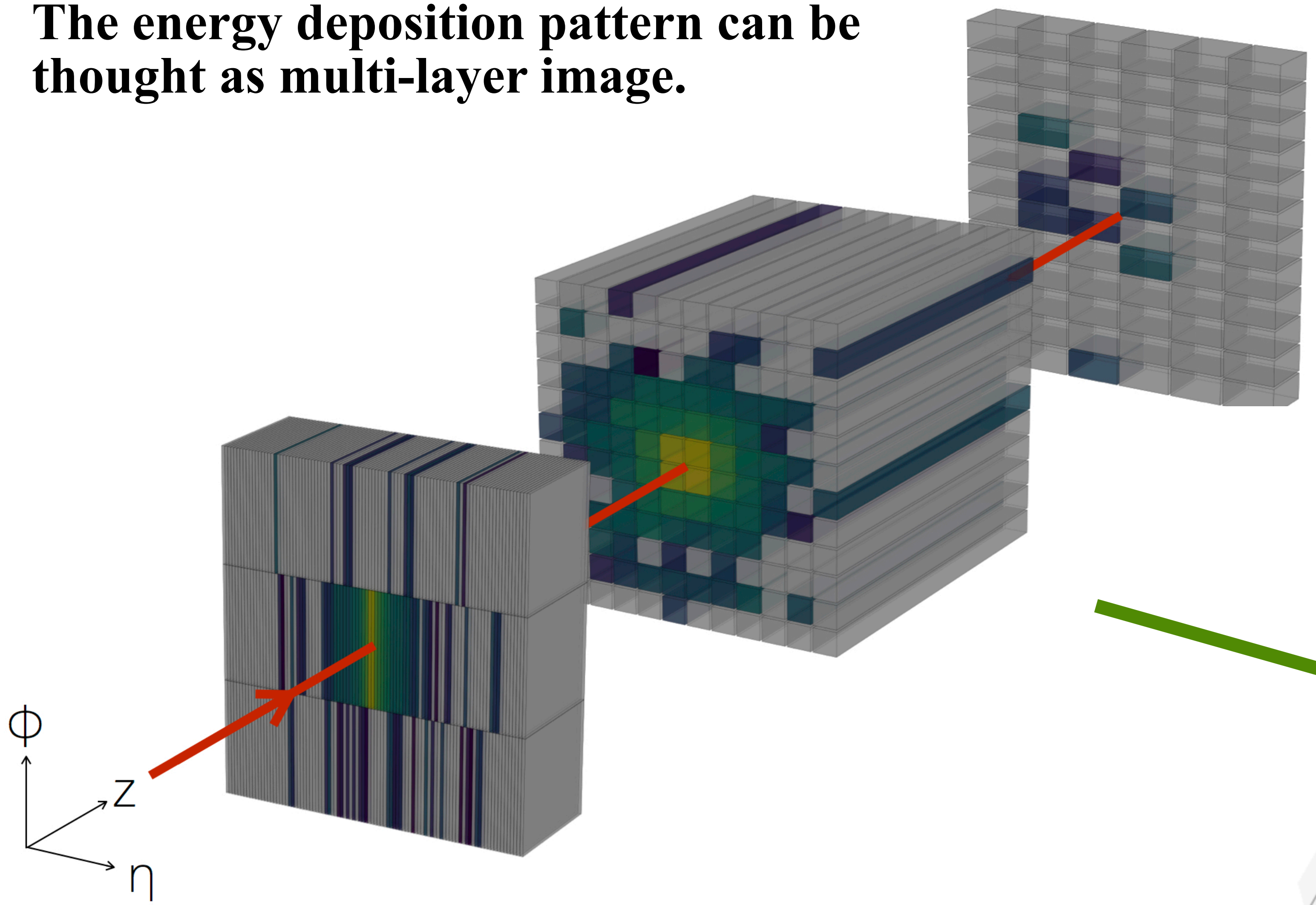
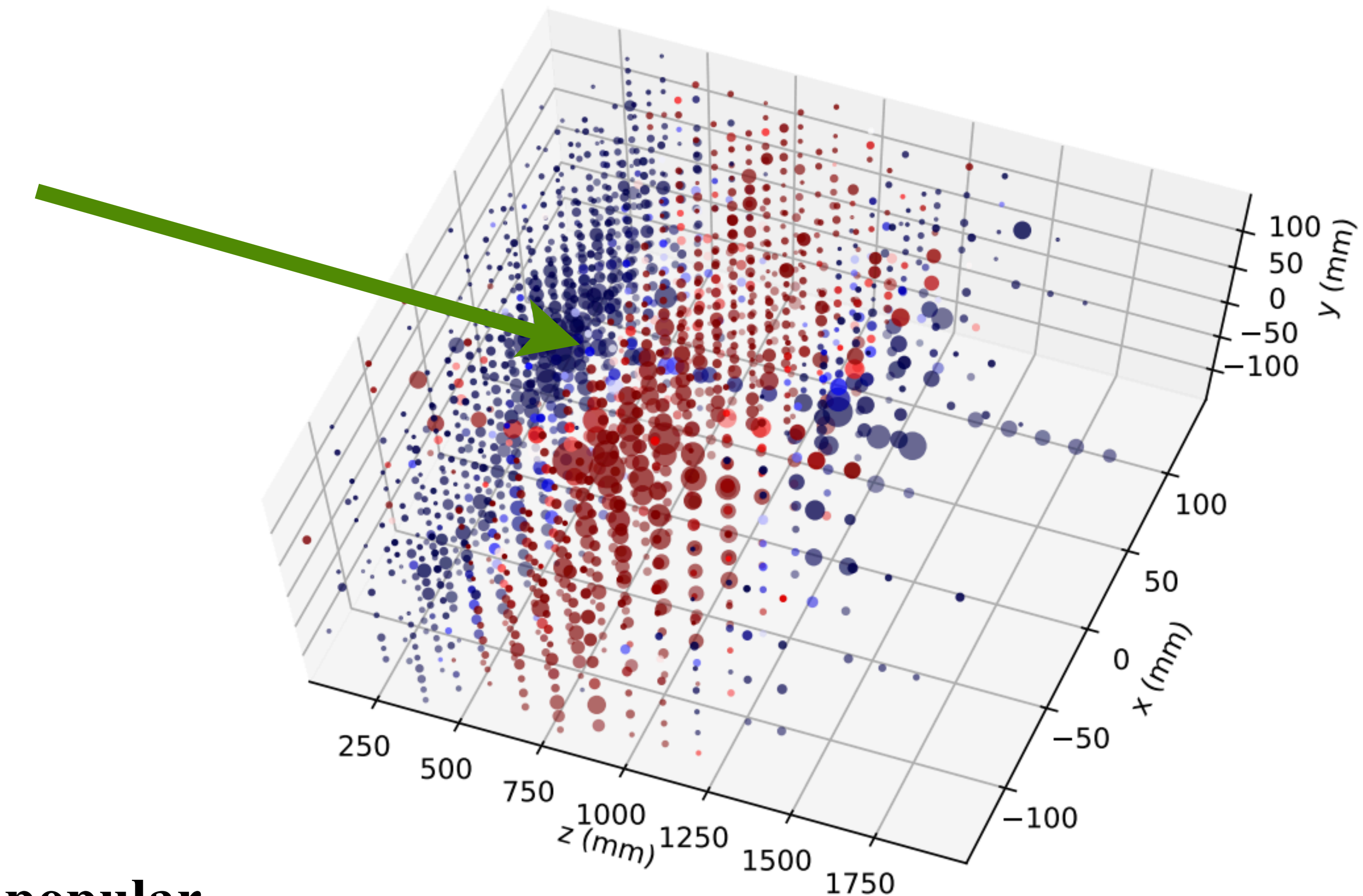


Image from 1705.02355

S. Qasim et-al EPJC 79 7 (2019) 608

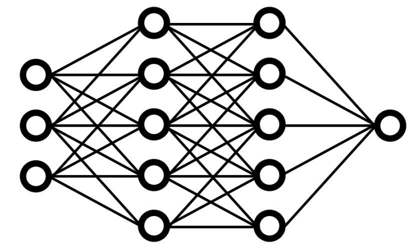


(a) Truth

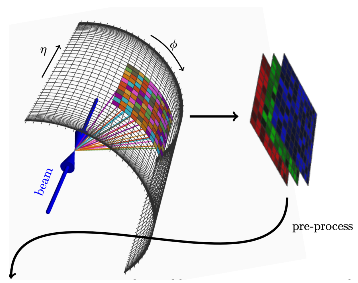
Soon point-cloud based methods became more popular.
Large sparsity in calorimeter shower were more efficient to treat.

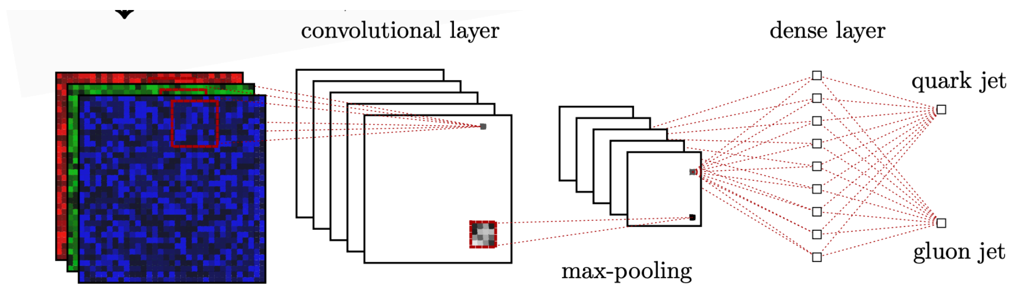
Data representation ⇔ NN correspondence

$$J = \{p_1^\mu, p_2^\mu, \dots\}$$

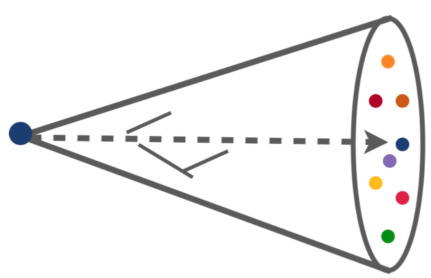


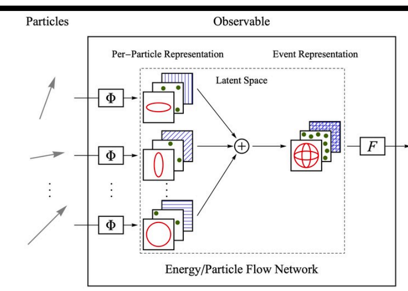
Ordered set
DNN



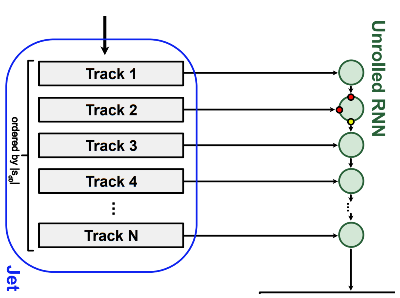


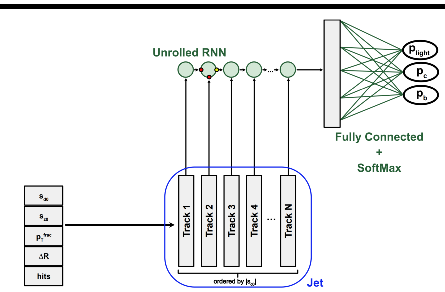
Grid
CNN



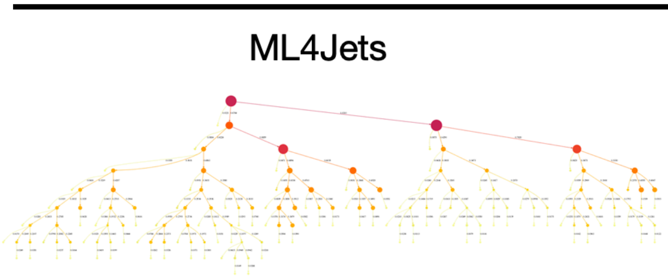


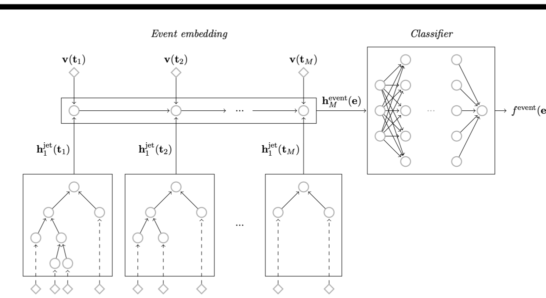
Unordered set
Deepest



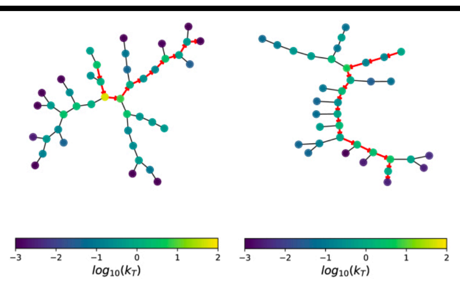


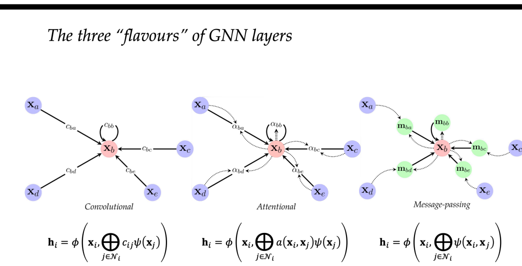
Sequential data
RNN





Tree structure
Deepset/GNN



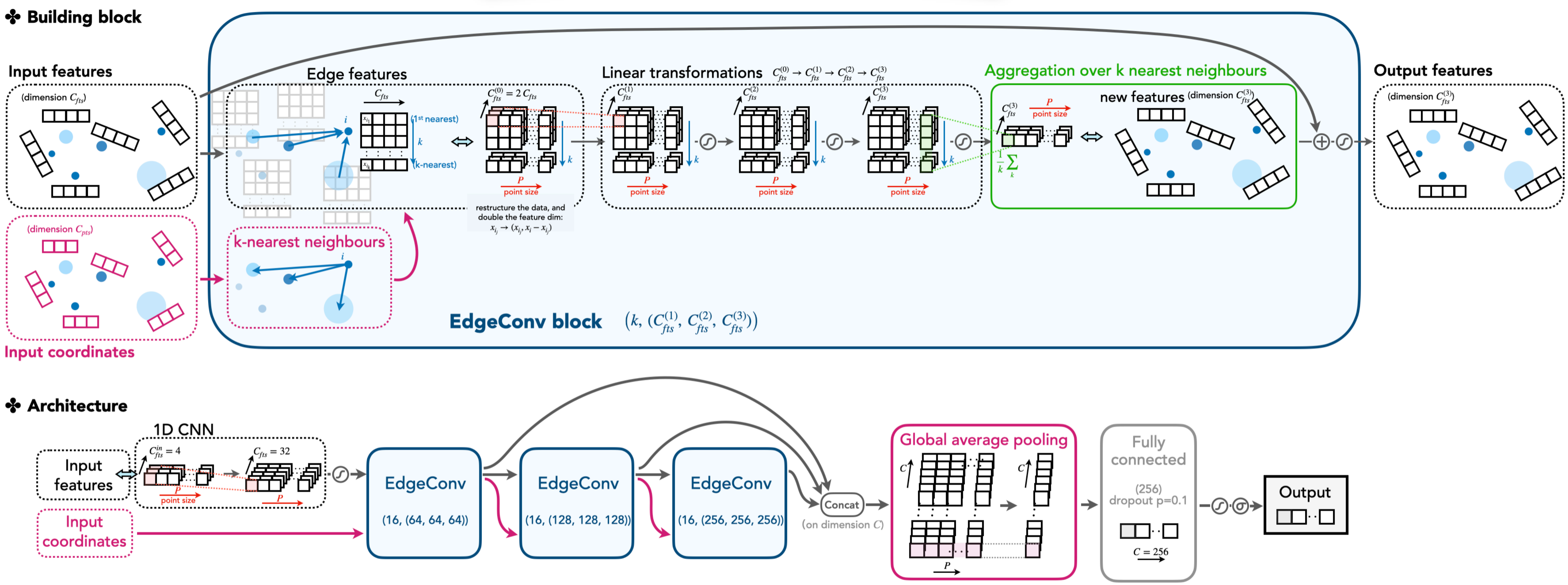


Graph
GNN

Object tagging

$$e'_{ijm} = \mathrm{ReLU}(\boldsymbol{\theta}_m \cdot (\mathbf{x}_j - \mathbf{x}_i) + \boldsymbol{\phi}_m \cdot \mathbf{x}_i),$$

Particle Net : 1902.08570
Huilin Qu, Loukas Gouskos



arXiv > cs > arXiv:1801.07829

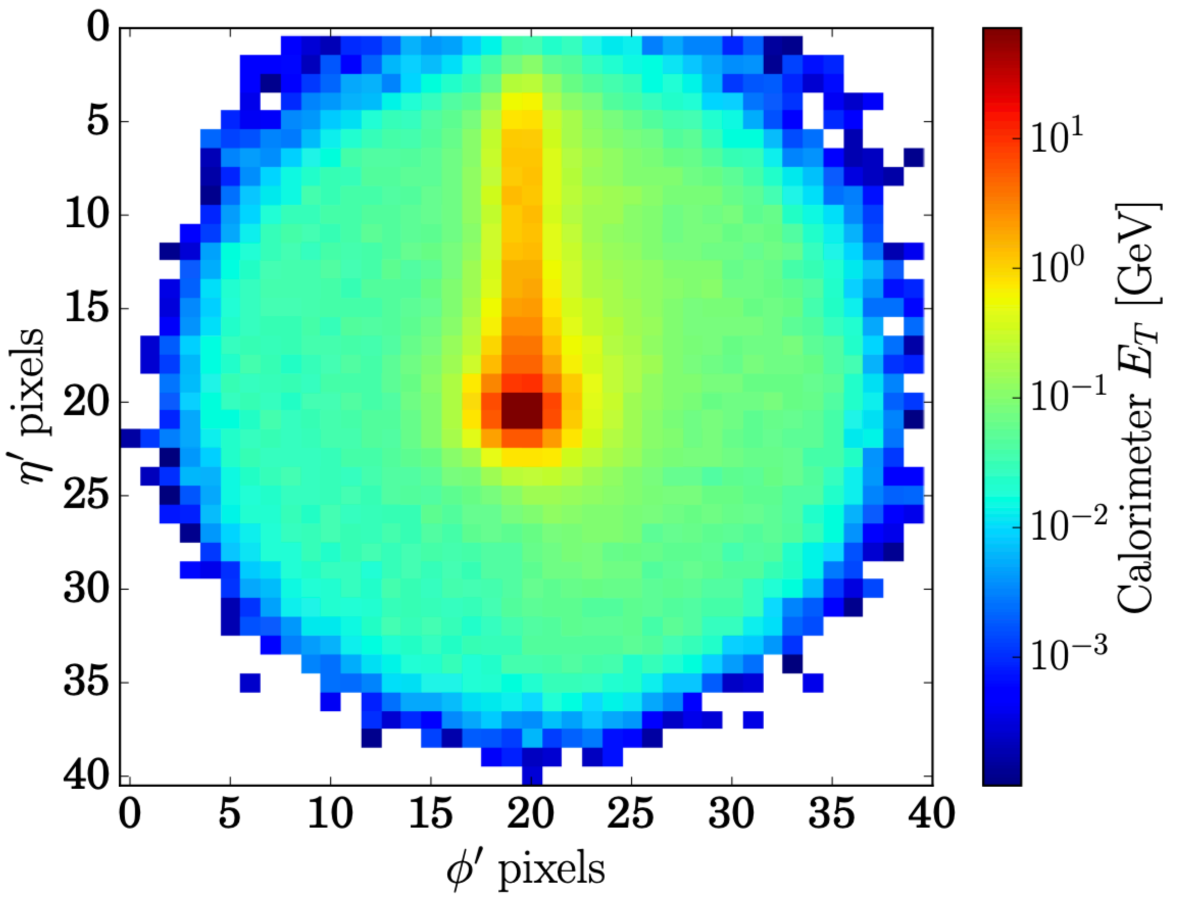
Computer Science > Computer Vision and Pattern Recognition

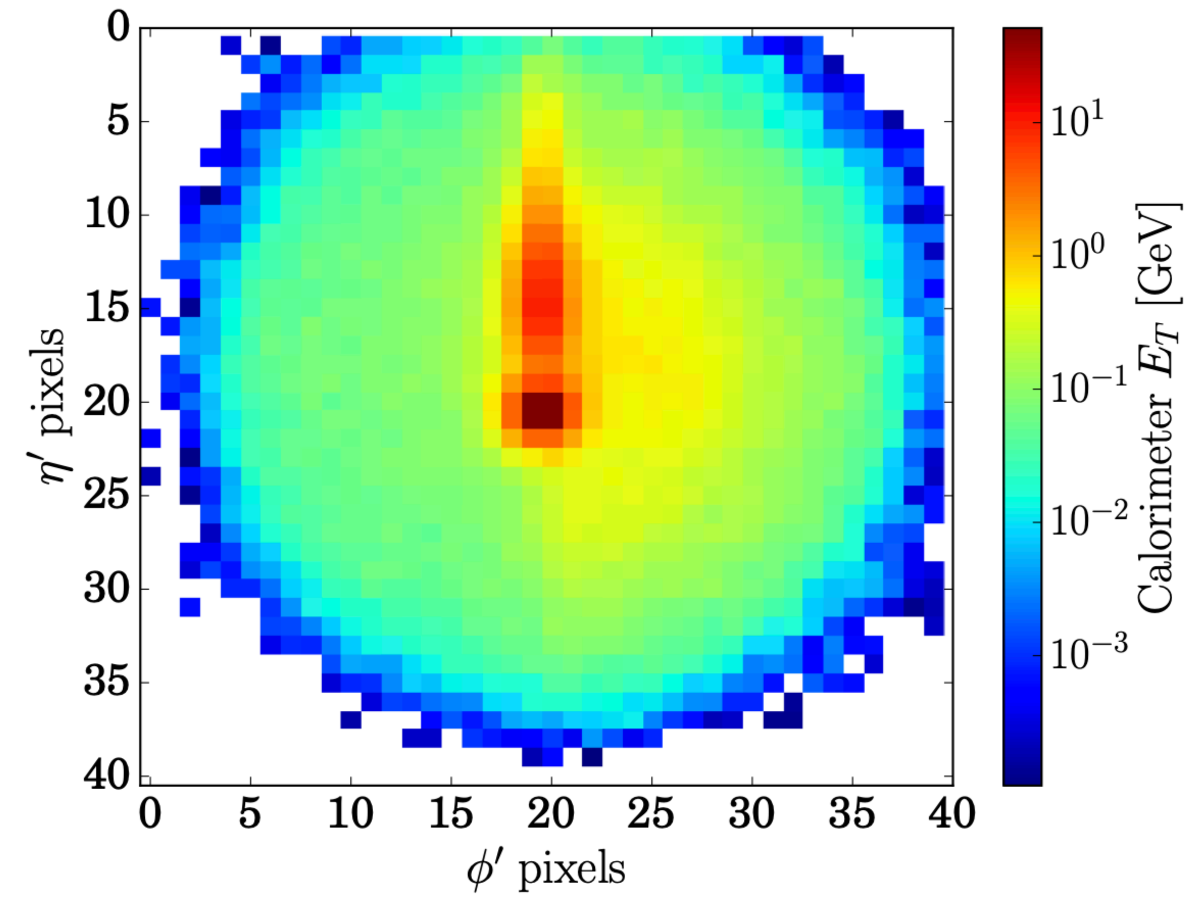
[Submitted on 24 Jan 2018 (v1), last revised 11 Jun 2019 (this version, v2)]

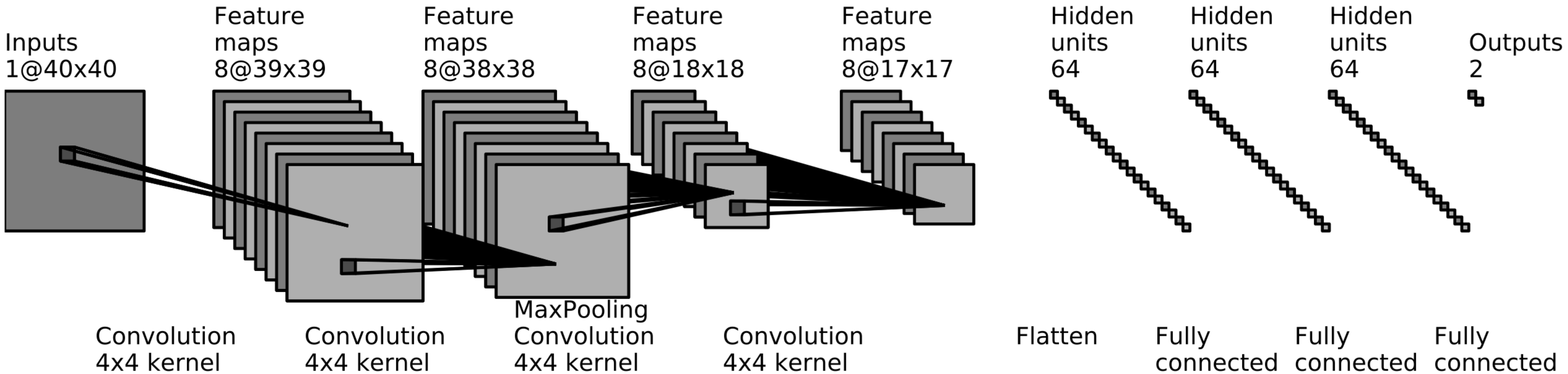
Dynamic Graph CNN for Learning on Point Clouds

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, Justin M. Solomon

Jet images & ML4Jets







Machine learning on sets

Let $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N} \in \mathbb{R}^{\mathbf{k}}$ be n pieces of data. This forms a set of cardinality N.

https://geometricdeeplearning.com/lectures/

Machine learning on sets

Let $\mathbf{x_1, x_2, \ldots, x_N} \in \mathbb{R}^{\mathbf{k}}$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

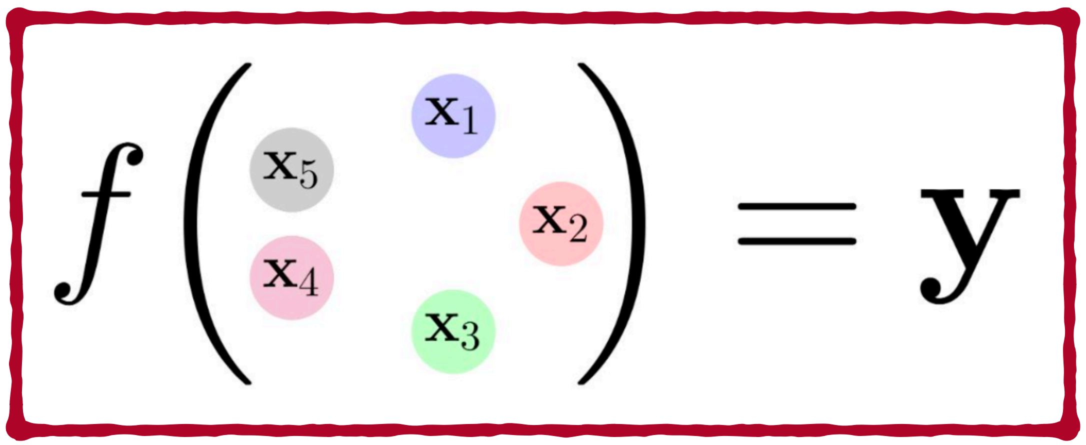
https://geometricdeeplearning.com/lectures/

Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

https://geometricdeeplearning.com/lectures/

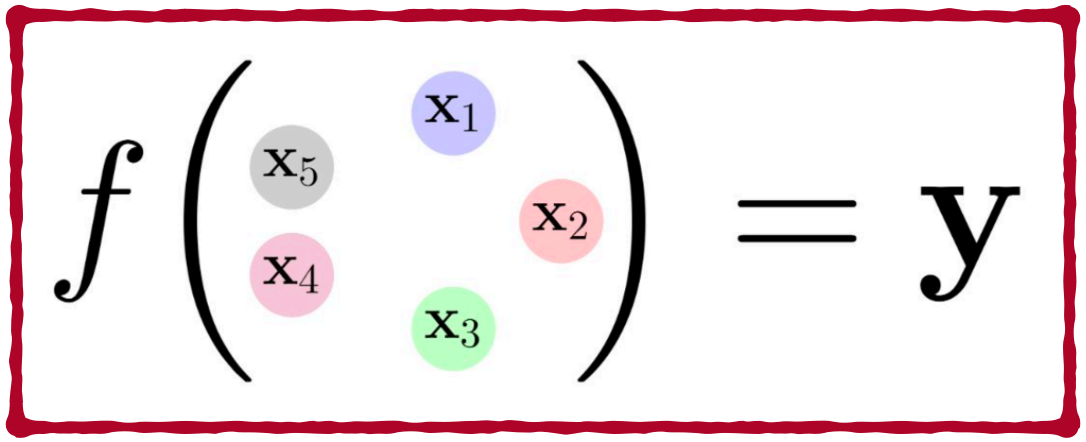


Machine learning on sets

Let $\mathbf{x_1, x_2, \ldots, x_N} \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

https://geometricdeeplearning.com/lectures/



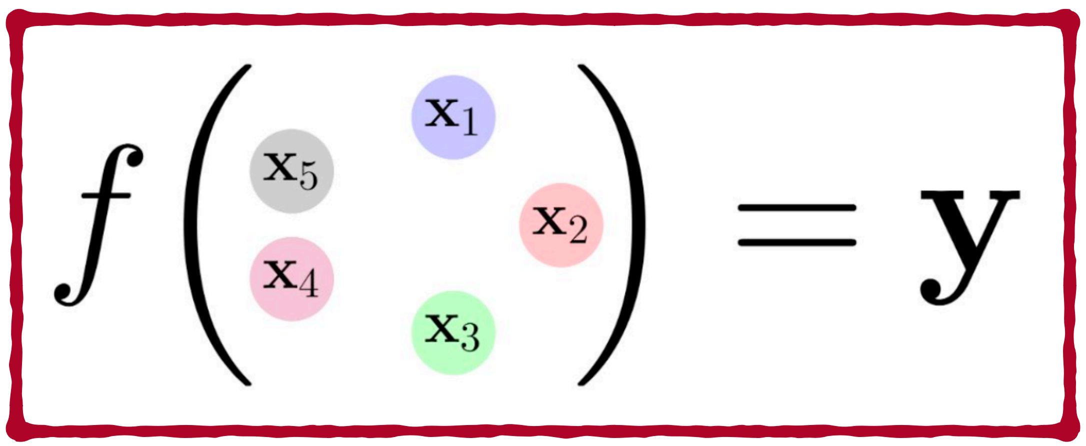
Basic required property : permutation invariance

Machine learning on sets

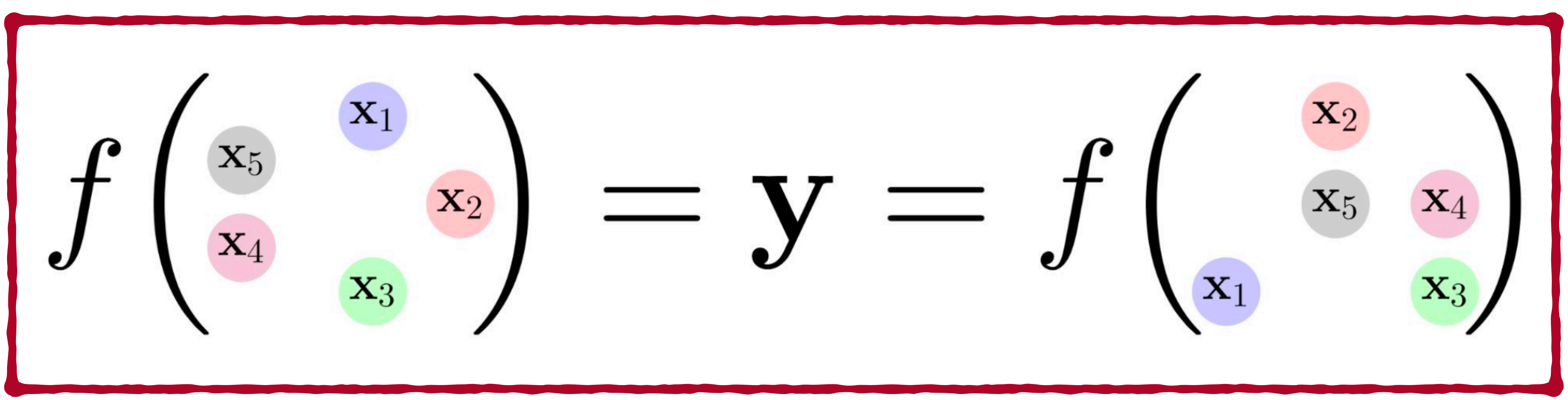
Let $x_1, x_2, \ldots, x_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

https://geometricdeeplearning.com/lectures/



Basic required property : permutation invariance



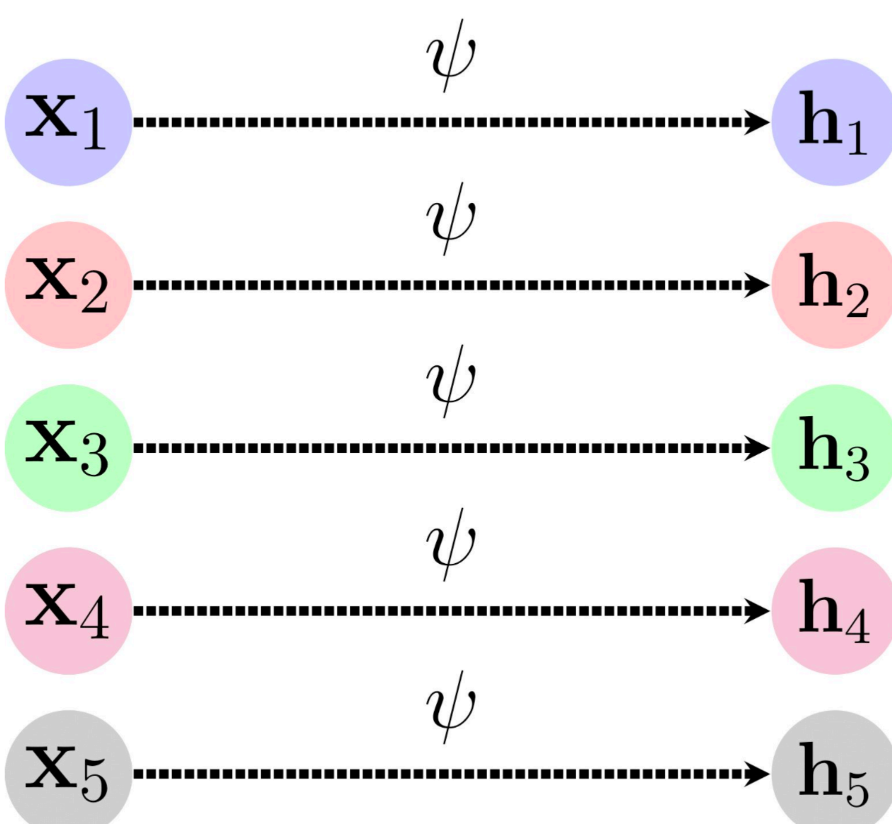
How the P.I. is achieved?

$$\mathbf{f}\left(\mathbf{PX}\right) = \mathbf{f}\left(\mathbf{X}\right)$$

Remember the permutation on a set?

$$\mathbf{P}_{(2,4,1,3)}\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} — & \mathbf{x}_1 & — \\ — & \mathbf{x}_2 & — \\ — & \mathbf{x}_3 & — \\ — & \mathbf{x}_4 & — \end{bmatrix} = \begin{bmatrix} — & \mathbf{x}_2 & — \\ — & \mathbf{x}_4 & — \\ — & \mathbf{x}_1 & — \\ — & \mathbf{x}_3 & — \end{bmatrix}$$

$$\mathbf{h_i} = \psi(\mathbf{x_i})$$



$$f\left(X\right) = \phi\left(\bigoplus_{i \in V} \psi(X_i)\right)$$

https://geometricdeeplearning.com/lectures/

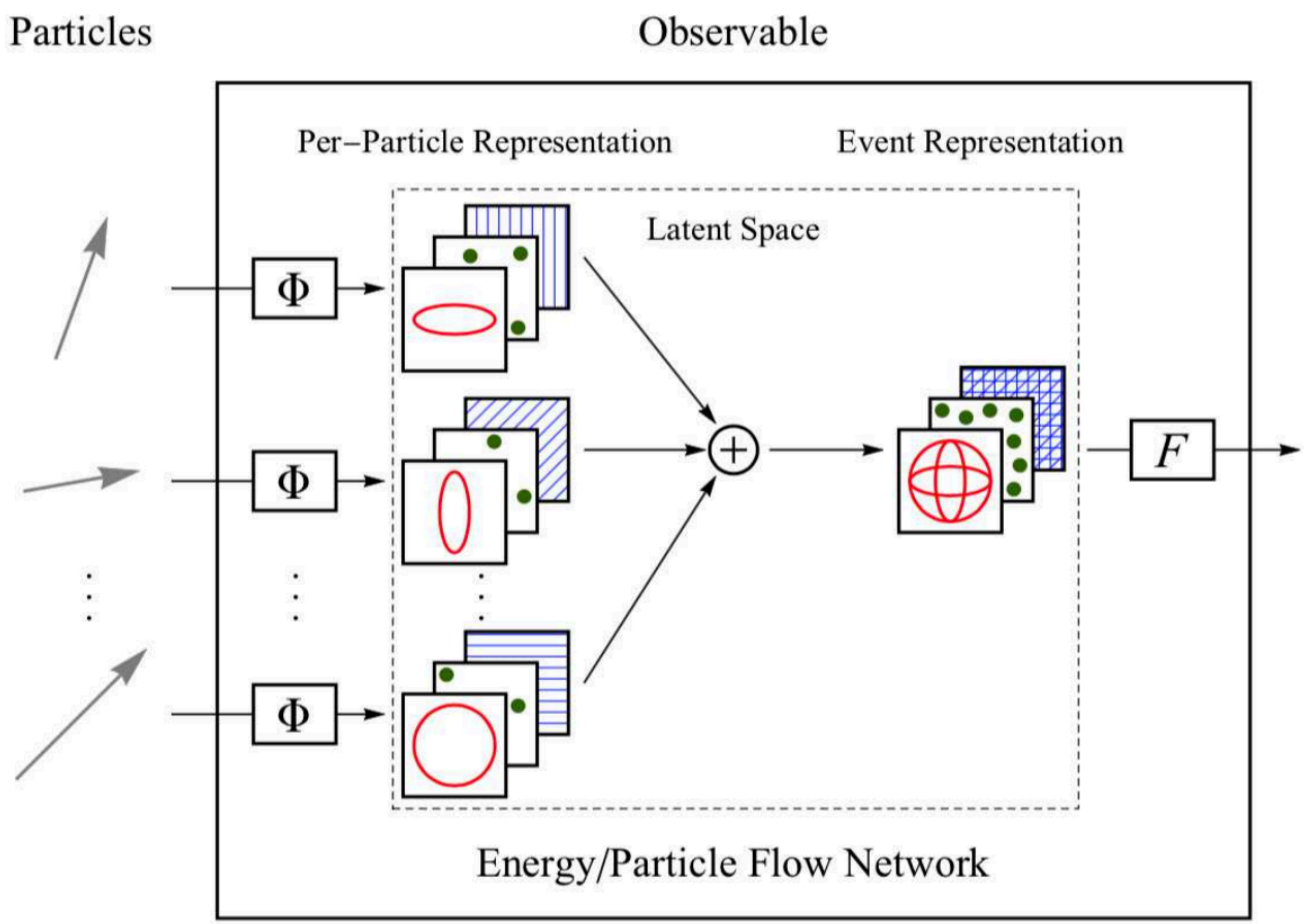
The permutation equivariant operation :

$$\Theta = \lambda \mathbf{I} + \gamma \left(\mathbf{1}\mathbf{1}^{\top}\right) \text{ for } \lambda, \gamma \in \mathbb{R}.$$

Example of deep-sets in HEP

arXiv:1810.05165

$$f(\{x_1, \ldots, x_M\}) = F\left(\sum_{i=1}^{M} \Phi(x_i)\right)$$



$$\text{EFN}\left(\{p_1^\mu, \ldots, p_M^\mu\}\right) = F\left(\sum_{i=1}^{M} z_i \Phi(\hat{p}_i)\right)$$

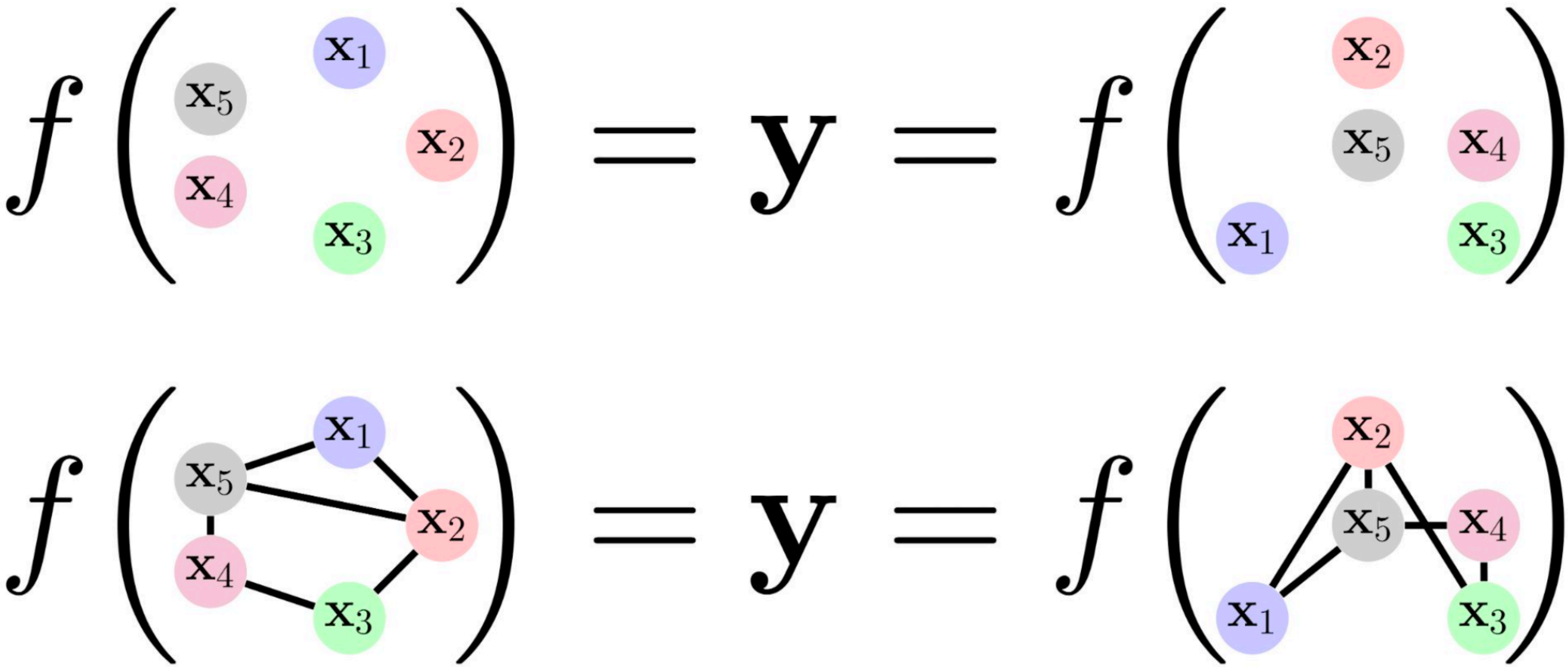
Manifestly IRC-safe latent space

$$\text{PFN}\left(\{p_1^\mu, \ldots, p_M^\mu\}\right) = F\left(\sum_{i=1}^{M} \Phi(p_i^\mu)\right)$$

Fully general latent space

What's the basic criteria of a GNN?

https://geometricdeeplearning.com/lectures/

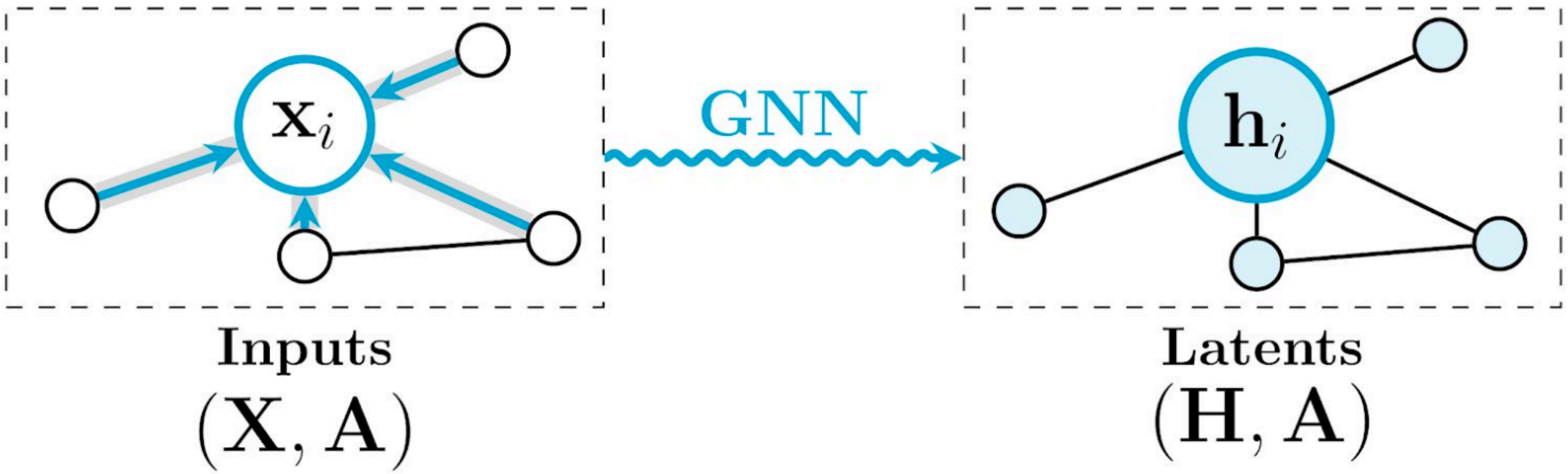


Invariance: $f(\mathbf{PX}, \mathbf{PAP}^\top) = f(\mathbf{X}, \mathbf{A})$

Equivariance: $f(\mathbf{PX}, \mathbf{PAP}^\top) = \mathbf{P}f(\mathbf{X}, \mathbf{A})$

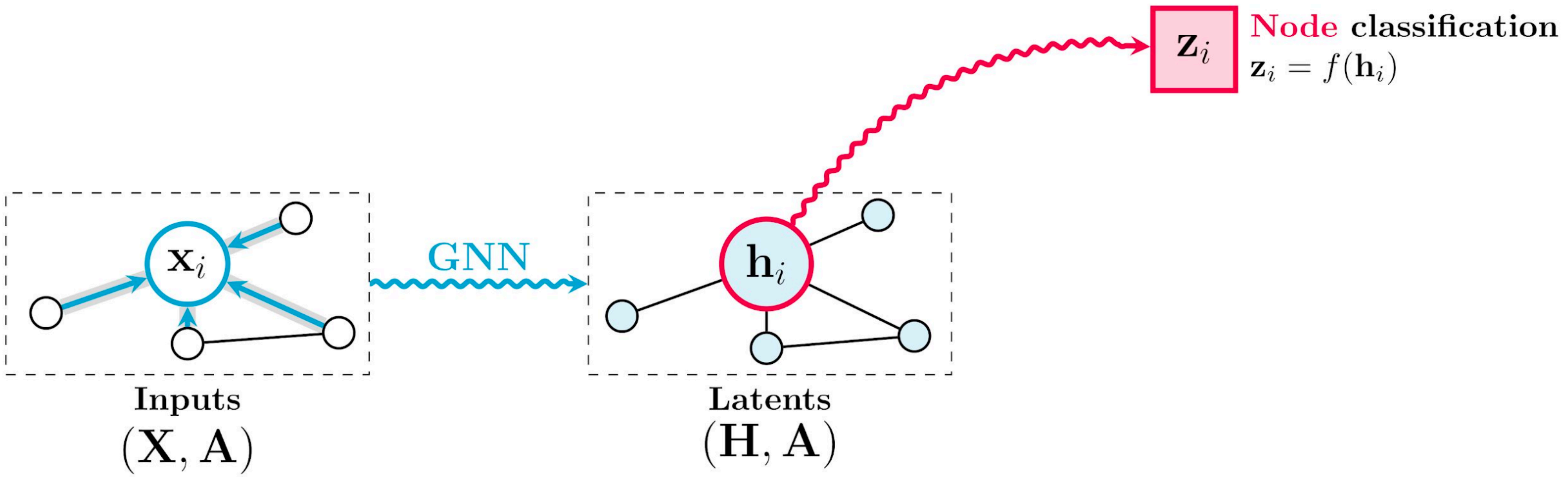
General methods of GNN

General methods of GNN



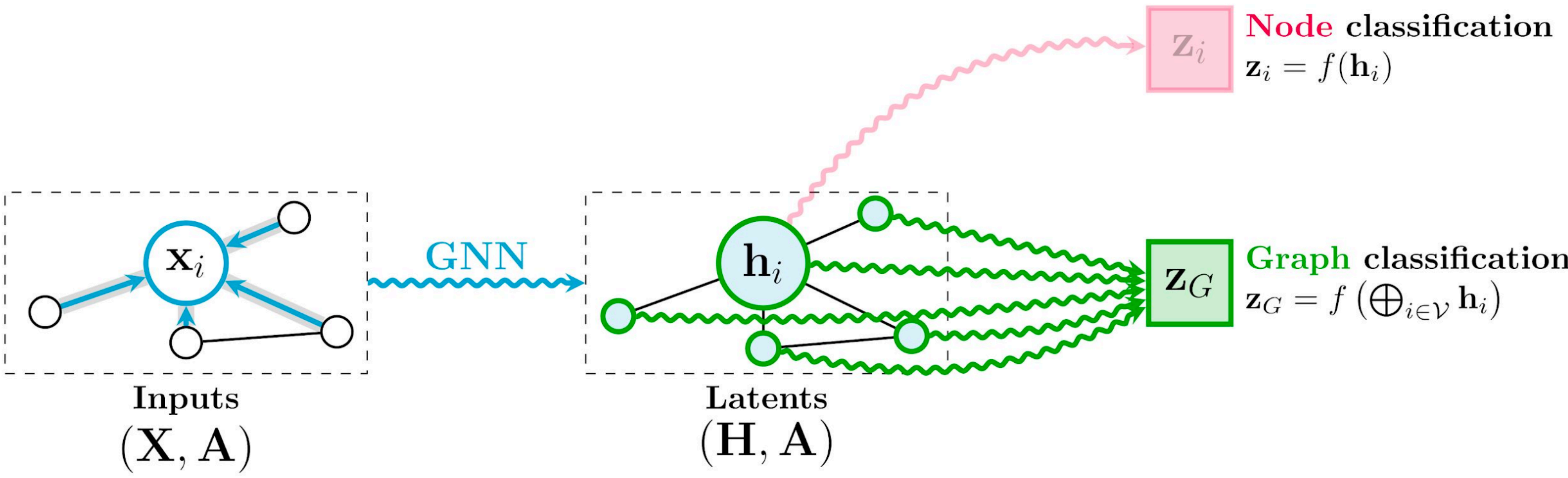
General methods of GNN

General methods of GNN



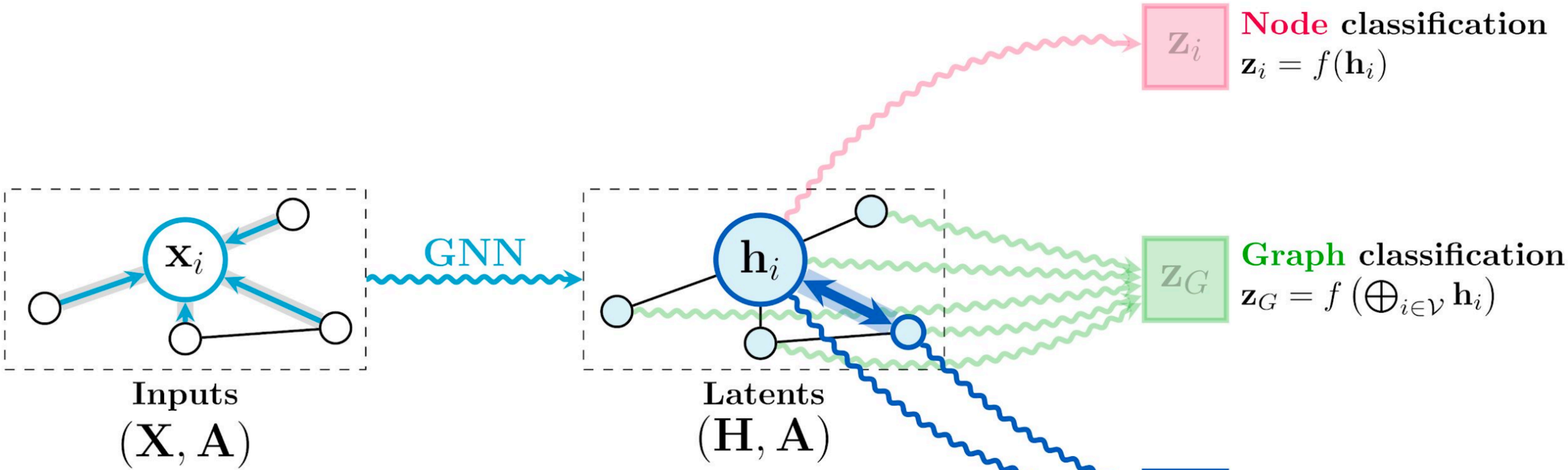
General methods of GNN

General methods of GNN



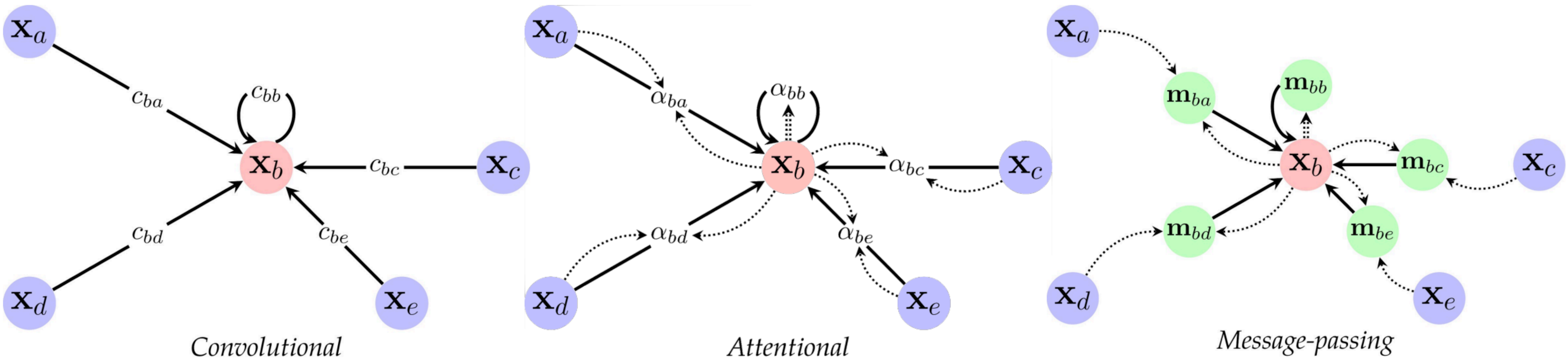
General methods of GNN

General methods of GNN



The different flavors of MPN

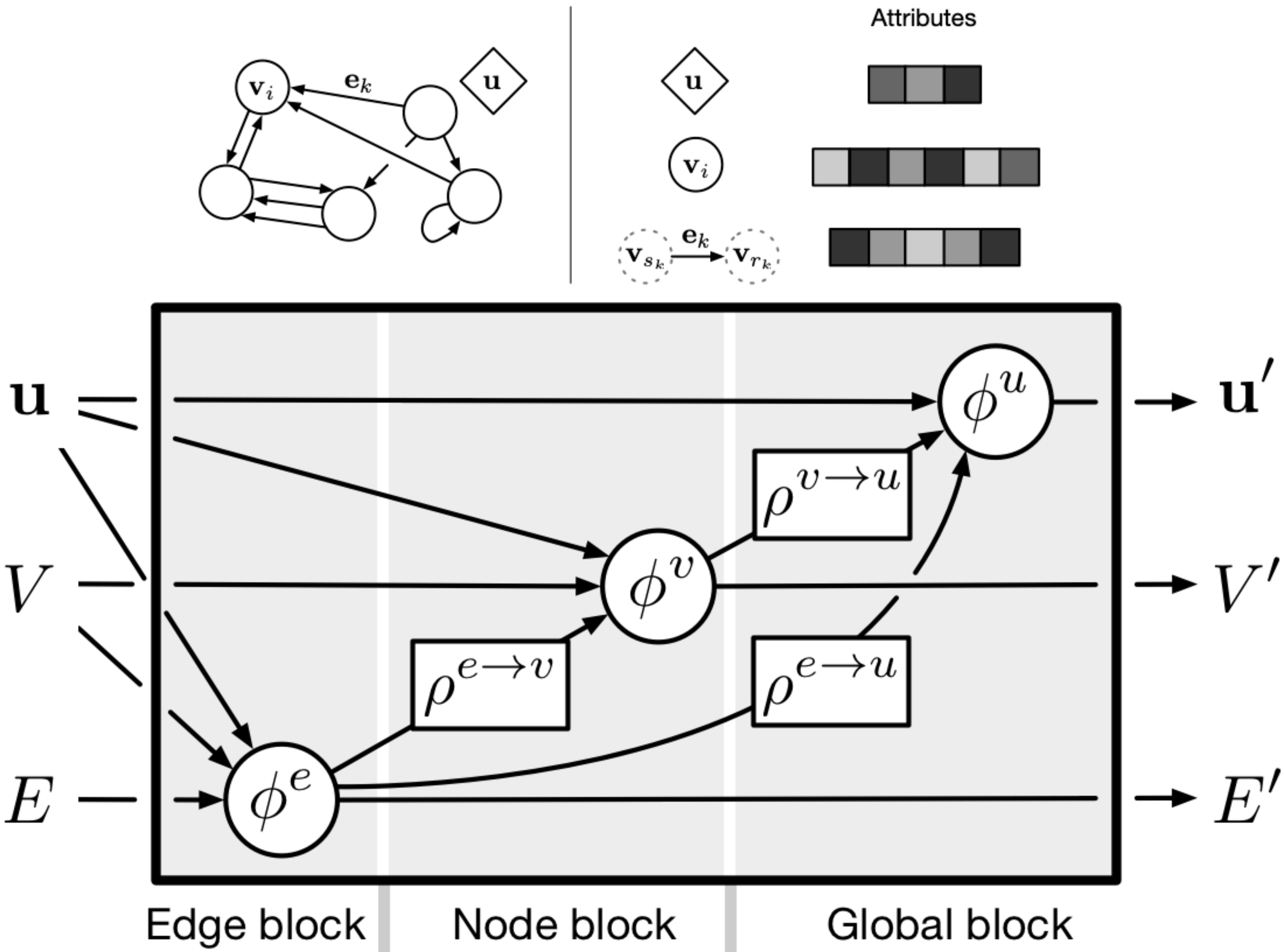
The three "flavours" of GNN layers



$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} c_{ij}\psi(\mathbf{x}_j)\right) \qquad \mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} a(\mathbf{x}_i,\mathbf{x}_j)\psi(\mathbf{x}_j)\right) \qquad \mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} \psi(\mathbf{x}_i,\mathbf{x}_j)\right)$$

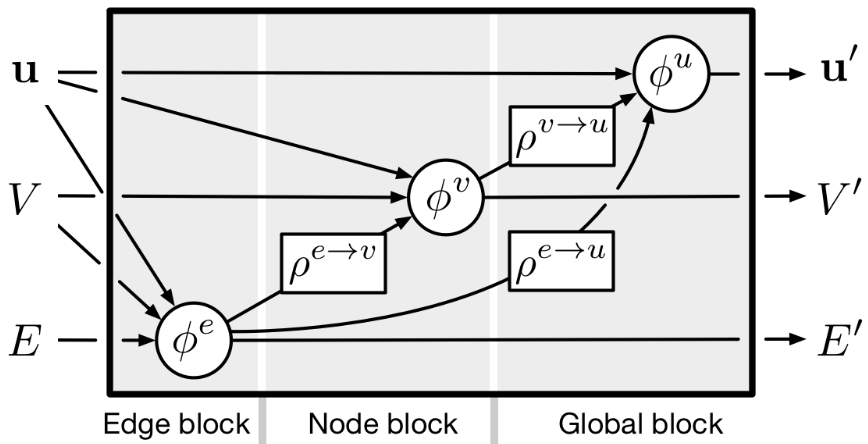
The general GNN

arXiv : 1806.01261

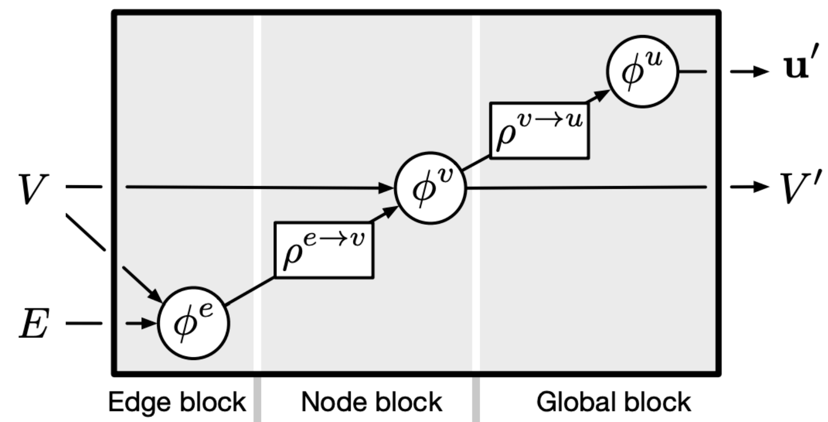


The general GNN

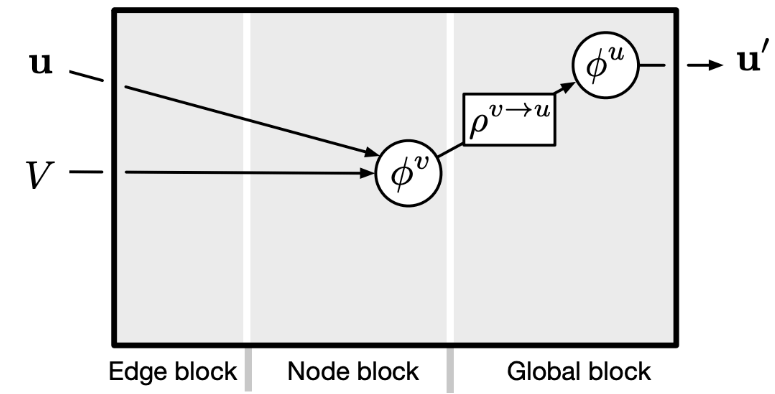
arXiv : 1806.01261



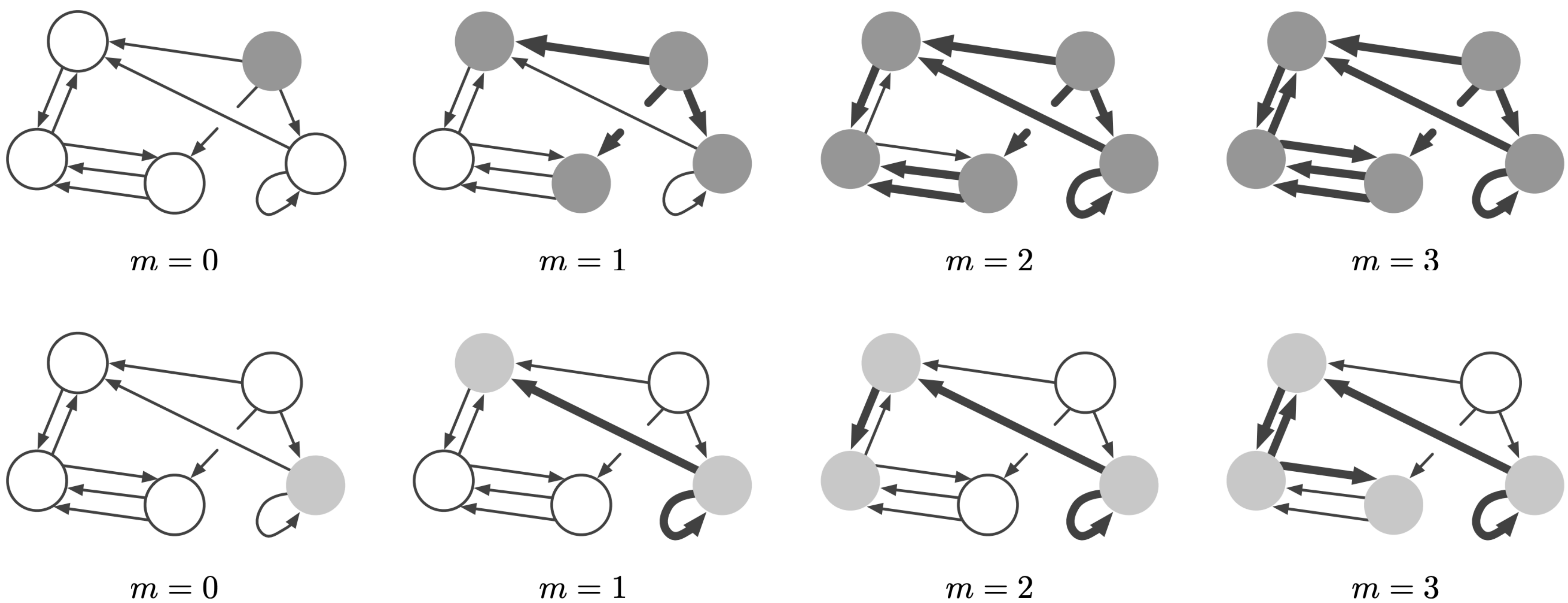
Full GN block



MPNN Layer

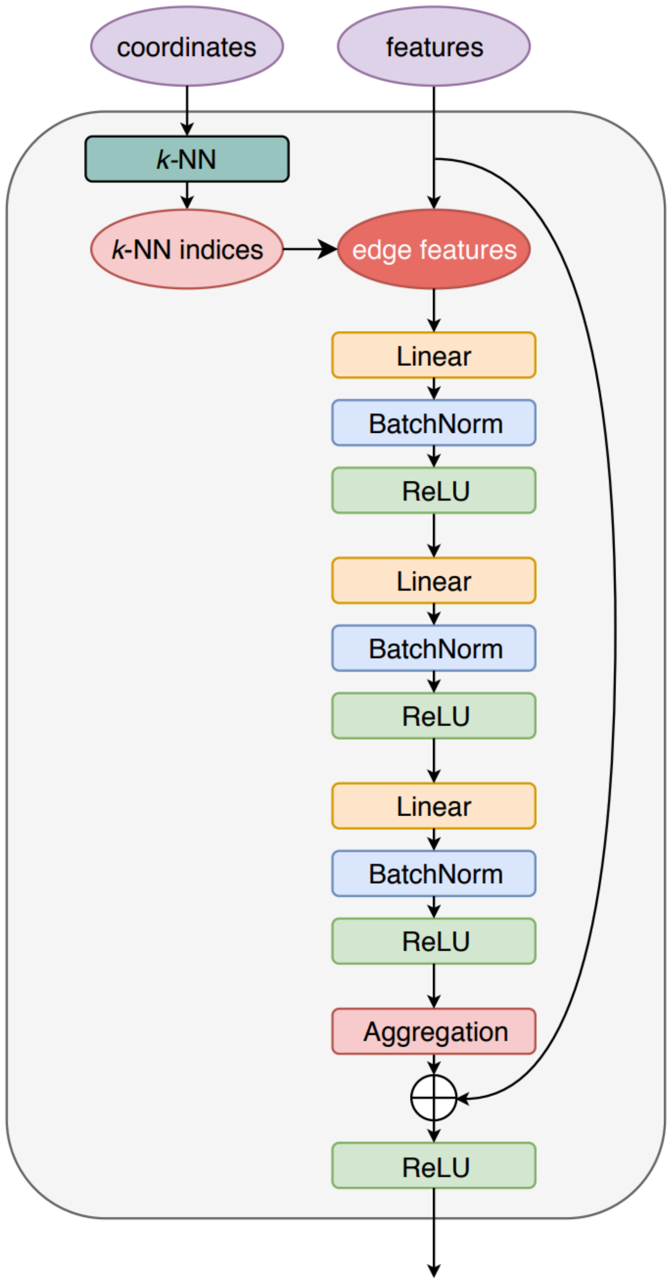


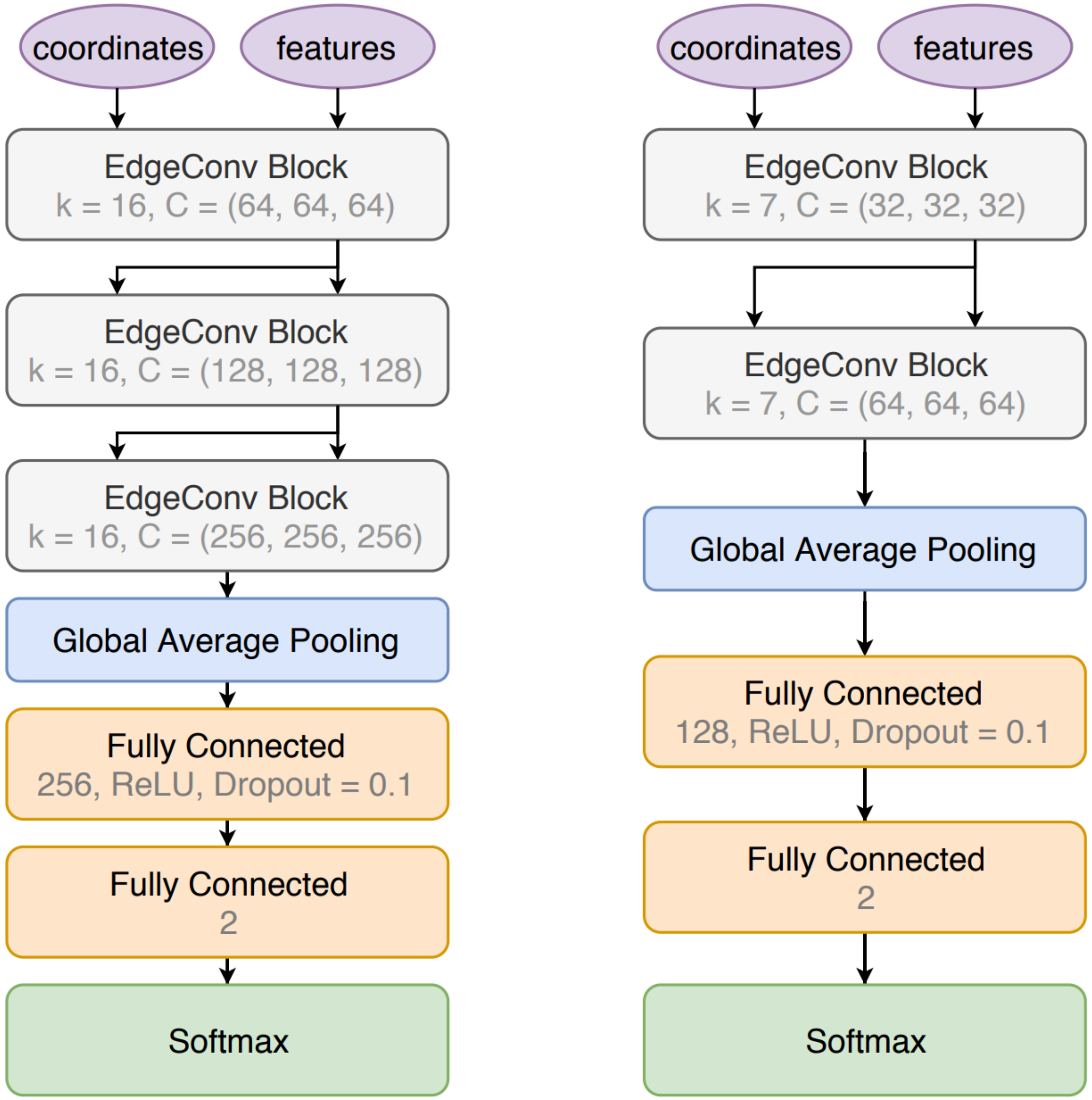
Deep-set layer



An example GNN in use

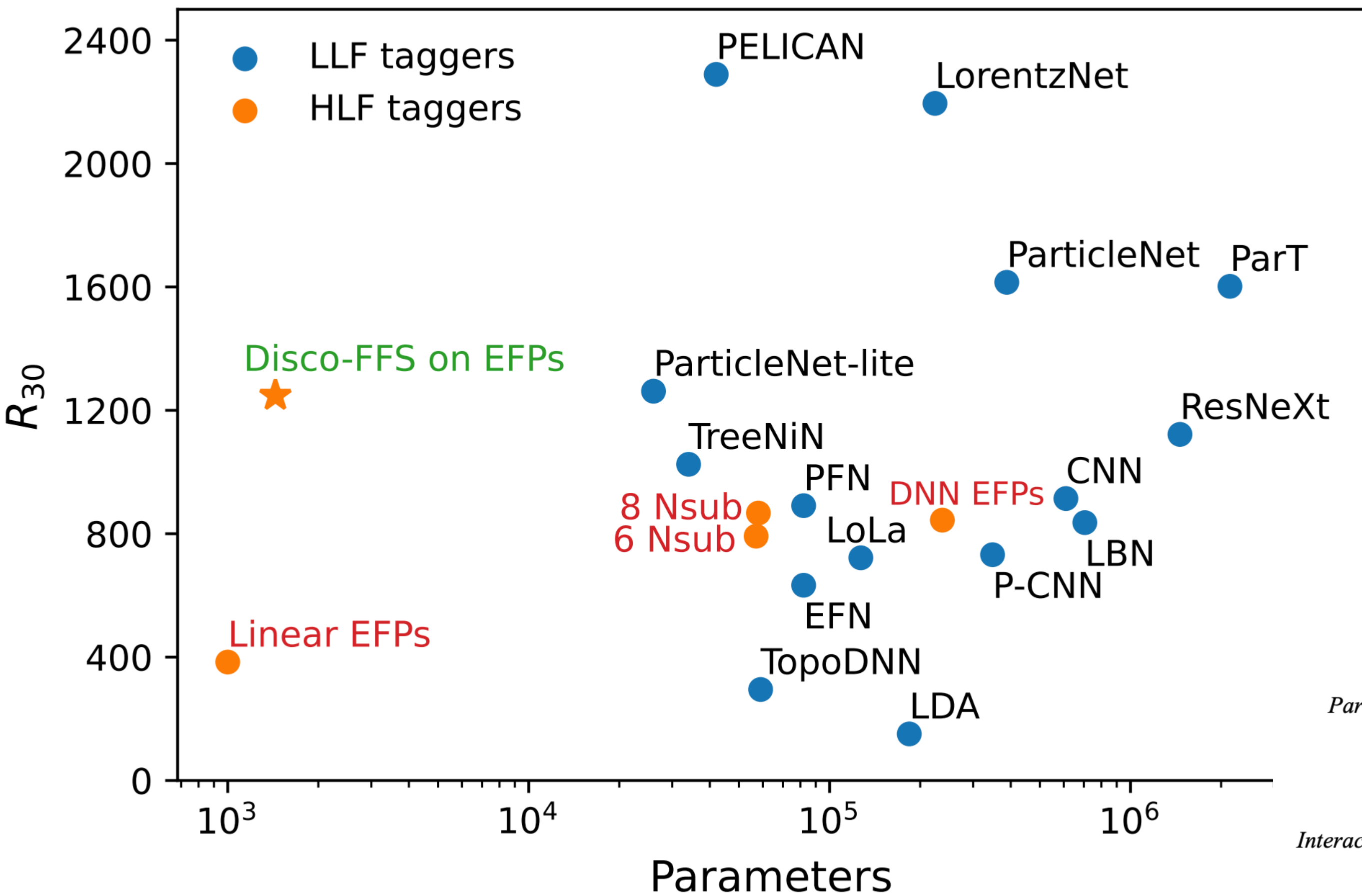
https://arxiv.org/pdf/1902.08570

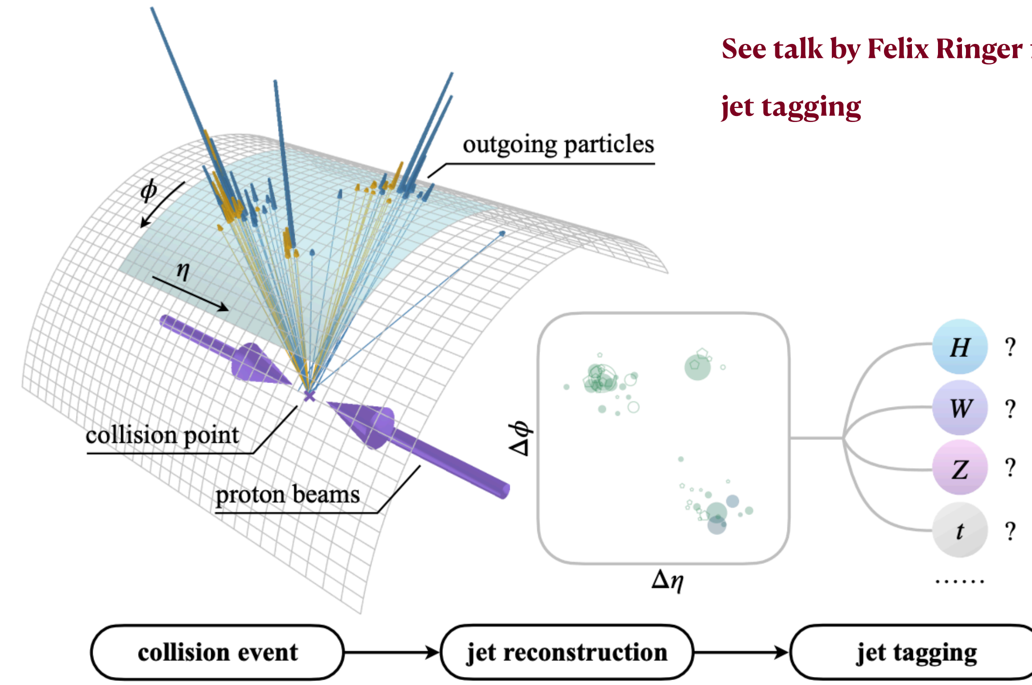




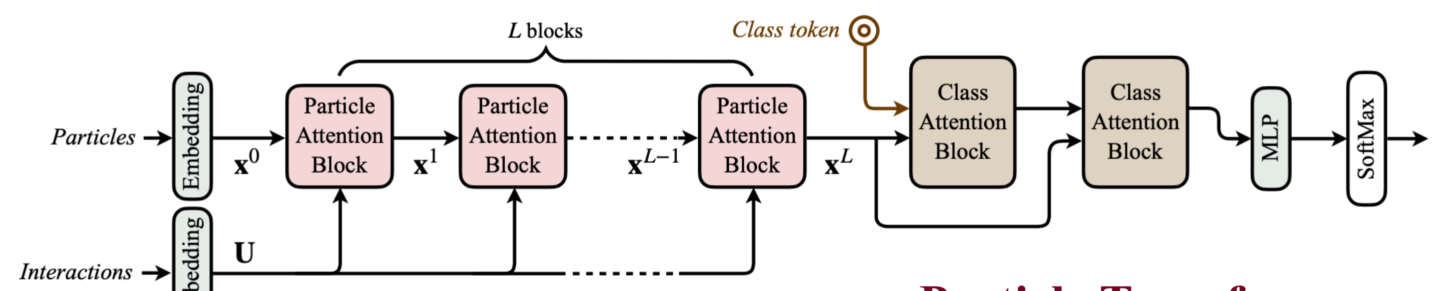
Object tagging

See talk by Felix Ringer for ML based EIC jet tagging



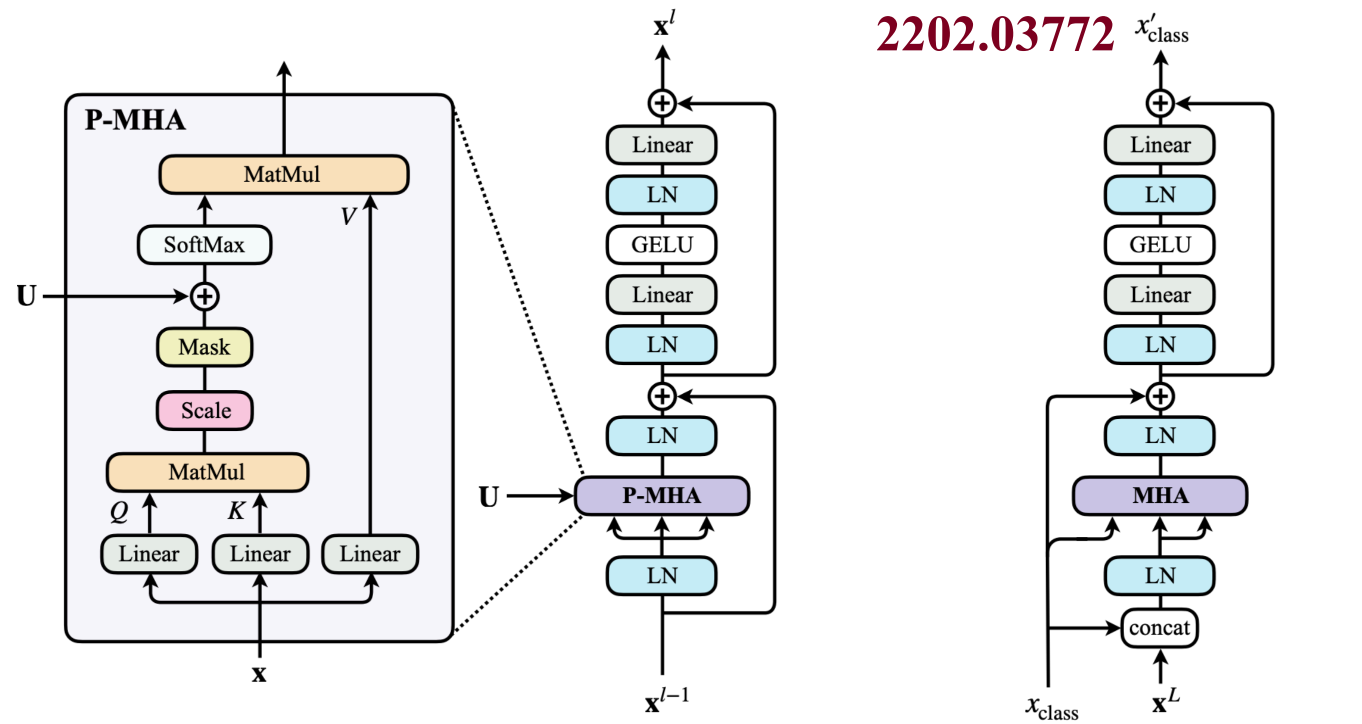


	Accuracy	AUC	$\mathrm{Rej}_{50\%}$	$\mathrm{Rej}_{30\%}$
P-CNN	0.930	0.9803	201 ± 4	759 ± 24
PFN	—	0.9819	247 ± 3	888 ± 17
ParticleNet	0.940	0.9858	397 ± 7	1615 ± 93
JEDI-net (w/ $\sum O$)	0.930	0.9807	—	774.6
PCT	0.940	0.9855	392 ± 7	1533 ± 101
LGN	0.929	0.964	—	435 ± 95
rPCN	—	0.9845	364 ± 9	1642 ± 93
LorentzNet	0.942	0.9868	498 ± 18	2195 ± 173
ParT	0.940	0.9858	413 ± 16	1602 ± 81
ParticleNet-f.t.	0.942	0.9866	487 ± 9	1771 ± 80
ParT-f.t.	**0.944**	**0.9877**	$\mathbf{691 \pm 15}$	$\mathbf{2766 \pm 130}$



(a) Particle Transformer

Particle Transformer : 2202.03772



(b) Particle Attention Block

(c) Class Attention Block

The takeaway from part-1

Jets are fascinating quantum mechanical objects, needs to be understood as best as possible from first principle analytics.

Measured jets are even more complex objects. Mapping them to Parton origin is hard.

We take help of modern machine learning methods to solve the existing problems.

The applications will be discussed in part-2

The takeaway from part-1

Jets are fascinating quantum mechanical objects, needs to be understood as best as possible from first principle analytics.

Measured jets are even more complex objects. Mapping them to Parton origin is hard.

We take help of modern machine learning methods to solve the existing problems.

The applications will be discussed in part-2

THANK YOU