In this notes, we study the requirements for the simplest toy generative model that would resemble a full generative model for jets.

We consider a 2-body decay. In the parent rest frame, we have the parent momentum $p_p^\mu = p_1^\mu + p_2^\mu = (\sqrt{s}, 0, 0, 0)$. From requiring 4-momentum conservation, the children energies are given by

$$\begin{aligned} E_1 &= \frac{\sqrt{s}}{2}\left(1 + \frac{m_1^2}{s} - \frac{m_2^2}{s}\right) \\ E_2 &= \frac{\sqrt{s}}{2}\left(1 + \frac{m_2^2}{s} - \frac{m_1^2}{s}\right) \end{aligned} \tag{1}$$

and the magnitude of their 3-momentum by

$$|\vec{p}| = \frac{\sqrt{s}}{2}\bar{\beta} = \frac{\sqrt{s}}{2}\sqrt{1 - \frac{2(m_1^2 + m_2^2)}{s} + \frac{(m_1^2 - m_2^2)^2}{s^2}} \tag{2}$$

1 Case with $m_1 = m_2 = m$

In this section we will specialize to the case where $m_1 = m_2 = m$ (we will comment on the full general case in the next section).

$$\begin{aligned} E_1 = E_2 &= \frac{\sqrt{s}}{2} = E \\ |\vec{p}| &= E\sqrt{1 - \frac{m^2}{E^2}} \end{aligned} \tag{3}$$

If we apply a boost to the lab frame, with factor $\beta\hat{n}$, we obtain for the children 3-momentum

$$\begin{aligned} \vec{p'}_1 &= -\gamma E\beta\hat{n} \,+\, E\sqrt{1 - \frac{m^2}{E^2}}\ [\hat{r} + (\gamma - 1)(\hat{r}\cdot\hat{n})\hat{n}] \\ \vec{p'}_2 &= -\gamma E\beta\hat{n} \,-\, E\sqrt{1 - \frac{m^2}{E^2}}\ [\hat{r} + (\gamma - 1)(\hat{r}\cdot\hat{n})\hat{n}] \end{aligned} \tag{4}$$

with $\hat{r}$ a unit vector that specifies the direction of the children momentum in the parent rest frame. For the parent we get

$$\vec{p'}_p = -2\gamma E\beta\hat{n} \tag{5}$$

From (5), we can rewrite (4) as:

$$
\begin{aligned}
\vec{p'}_1 &= \frac{1}{2}\vec{p'}_p + \vec{\Delta} \\
\vec{p'}_2 &= \frac{1}{2}\vec{p'}_p - \vec{\Delta}
\end{aligned}
\tag{6}
$$

with

$$
\vec{\Delta} = E\sqrt{1-\frac{m^2}{E^2}}\ \left[\hat{r} + (\gamma-1)(\hat{r}\cdot\hat{n})\hat{n}\right] \tag{7}
$$

Let's study the dependence of $|\vec{\Delta}| = \Delta(E, m, \hat{r}, \gamma, \hat{n})$. We will work in the (y,z) plane.

- $\gamma = \frac{E_p}{m_p}$ or $\gamma\beta = |\vec{p}_p|/m_p$
- If we draw an angle $\phi \in \{-\pi, \pi\}$ from a uniform distribution, we get in the $(\hat{y}, \hat{z})$ plane, $\hat{r} = (\sin\phi, \cos\phi)$.
- At each step, we draw ϕ and m, and boost the children to the lab frame. Next, we promote each child to a parent and repeat the process. Thus, the unit vector $\hat{n}$ is fixed by the previous step and given by $\hat{n} = (\sin\theta_p, \cos\theta_p)$ with $\theta_p = \tan^{-1}\left(\frac{p'^p_y}{p'^p_z}\right)$ the parent angle with respect to the $\hat{z}$ axis in the lab frame.
- At each step, $E_1 = E_2 = E = \frac{\sqrt{s}}{2} = \frac{m_p}{2}$ is fixed by the parent mass.

As a result, we obtain $|\Delta| = \Delta(m, \phi)$ given by

$$
\begin{aligned}
\Delta_y &= \frac{m_p}{2}\sqrt{1-4\frac{m^2}{m_p^2}}\left[\sin\phi + \left(\frac{m_p}{2m}-1\right)\cos(\phi-\theta_p)\sin\theta_p\right] \\
\Delta_z &= \frac{m_p}{2}\sqrt{1-4\frac{m^2}{m_p^2}}\left[\cos\phi + \left(\frac{m_p}{2m}-1\right)\cos(\phi-\theta_p)\cos\theta_p\right]
\end{aligned}
\tag{8}
$$

1.1 Minimum requirements for a toy generative model for jets

The traditional clustering algorithms are based on a measure given by:

$$
d_{ij} = \min(p_{\mathrm{T}i}^{2\alpha}, p_{\mathrm{T}j}^{2\alpha})\ \frac{\Delta R_{ij}^2}{R^2} \tag{9}
$$

where $\Delta R_{ij} = (\theta_i - \theta_j)$, with θ_i the angle of particle i with respect to the $\hat{z}$ axis in the lab frame. Also, $\alpha = \{-1, 0, 1\}$ defines the {anti-kt, CA and kt} algorithms respectively.

Thus for a meaninful definition of d_{ij} the toy model should be at least 2D. Also, in our 2D model, we identify $p_{\mathrm{T}} = |p_y| = |\vec{p'}|\ |\sin\theta|$.

At each splitting, we define the children momentum following (10):

$$
\begin{aligned}
\vec{p'}_1 &= \frac{1}{2}\vec{p'}_p + \vec{\tilde{\Delta}} \\
\vec{p'}_2 &= \frac{1}{2}\vec{p'}_p - \vec{\tilde{\Delta}}
\end{aligned}
\tag{10}
$$

Next, it remains to define $\tilde{\Delta}$ in the toy model. We consider two options:

1. Set a starting value $\vec{\tilde{\Delta}}_0$ (and $\vec{p'}^{\,0}_p$). At each step j we define:

$$\vec{\tilde{\Delta}}_{j+1} = |\tilde{\Delta}_j|\, r\, (\sin\phi, \cos\phi) \tag{11}$$

   where we draw ϕ from a uniform distribution and r is drawn from an exponential distribution $f(x,\lambda) = \lambda e^{-\lambda x}$. [1]

   This is the model we currently agreed to build.

2. We could define $\tilde{\Delta}$ as in (7):

$$\tilde{\Delta} = E\sqrt{1 - \frac{m^2}{E^2}}\ \left[\hat{r} + (\gamma - 1)(\hat{r}\cdot\hat{n})\hat{n}\right] = \tilde{\Delta}(m, \phi) \tag{12}$$

   We draw ϕ from a uniform distribution. We can get m in a way to resemble the sudakov factor approach as:

$$m_{j+1} = \frac{m_j}{2}\ r \tag{13}$$

   where r is drawn from an exponential distribution $f(x,\lambda) \propto e^{-\lambda x}$ for $x \in (0,1)$. The prescription of (13) solves one of the problems of the traditional parton showers given that it satisfies $m_L + m_R \leq m_p$[2]. Then, we could think of m_{j+1} as the off-shell mass value, to avoid the required reshuffling. This results in leaves where each pair of siblings have the same mass, all different among pairs of siblings.

   This case would be closer to the real parton shower but more complex and time consuming.

[1] Ideally, this distribution should be bounded between 0 and 1. The pyro distribution I am using is for $x \in [0,\infty)$. My understanding is that we could still use the pyro distribution and accept only the values for $x \in [0,1)$, which would account to a global normalization factor ?

[2] Traditional parton showers draw m_L and m_R independently, so they should check the constraint is satisfied.

2 General case, with $m_1 \neq m_2$

We could also build a model for this case, adding extra features. In this case, (10) becomes

$$
\begin{aligned}
\vec{p'}_1 &= \frac{E_1}{E_p}\vec{p'}_p + E_1 \; [\hat{r} + (\gamma - 1)(\hat{r} \cdot \hat{n})\hat{n}] \\
\vec{p'}_2 &= \frac{E_2}{E_p}\vec{p'}_p + E_2 \; [\hat{r} + (\gamma - 1)(\hat{r} \cdot \hat{n})\hat{n}]
\end{aligned}
\tag{14}
$$

Also, we should replace (13) by

$$
\begin{aligned}
m_1^{j+1} &= m_p^j \; r_1 \\
m_2^{j+1} &= (m_p^j - m_1^{j+1}) \; r_2
\end{aligned}
\tag{15}
$$

where r_1 and r_2 are independently drawn from an exponential distribution $f(x, \lambda) \propto e^{-\lambda x}$ for $x \in (0, 1)$.

3 Previous toy models and their issues