Toy Generative Model for Jets

July 26, 2019

In this notes, first we study the requirements for the simplest toy generative model that would resemble a full generative model for jets. Next, we describe how to build such a model.

1 Kinematic considerations

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

$$E_L = \frac{\sqrt{s}}{2}\left(1 + \frac{m_L^2}{s} - \frac{m_R^2}{s}\right)$$
$$E_R = \frac{\sqrt{s}}{2}\left(1 + \frac{m_R^2}{s} - \frac{m_L^2}{s}\right) \tag{1}$$

and the magnitude of their 3-momentum by

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

1.1 Simplest case: $m_L = m_R = m$

In this section we will specialize to the case where $m_L = m_R = m$ (we will comment on the general case in the next subsection).

$$E_L = E_R = \frac{\sqrt{s}}{2} = E$$
$$|\vec{p}| = E\sqrt{1 - \frac{m^2}{E^2}} \tag{3}$$

1

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

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

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

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

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

$$
\begin{aligned}
\vec{p}_{\mathrm{L}}^{\,l} &= \frac{1}{2}\vec{p}_p^{\,l} - \vec{\Delta} \\
\vec{p}_{\mathrm{R}}^{\,l} &= \frac{1}{2}\vec{p}_p^{\,l} + \vec{\Delta}
\end{aligned}
\tag{6}
$$

with

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

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 $\hat{n} = (\sin\theta_{\mathrm{p}}, \cos\theta_{\mathrm{p}})$ with $\theta_{\mathrm{p}} = \tan^{-1}\left(\frac{(p_{\mathrm{p}})_y}{(p_{\mathrm{p}})_z}\right)$ the parent angle with respect to the $\hat{z}$ axis in the lab frame.

- At each step, $E_{\mathrm{L}} = E_{\mathrm{R}} = E = \frac{\sqrt{s}}{2} = \frac{m_{\mathrm{p}}}{2}$ is fixed by the parent mass.

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

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

2

1.2   General case, with $m_{\mathbf{L}} \neq m_{\mathbf{R}}$

In this case, (6) becomes

$$\vec{p}_{\mathrm{L}}^{\,l} = \frac{E_{\mathrm{L}}}{E_{\mathrm{p}}}\vec{p}_{\mathrm{p}}^{\,l} - |\vec{p}| \; [\hat{r} + (\gamma - 1)(\hat{r} \cdot \hat{n})\hat{n}]$$
$$\vec{p}_{\mathrm{R}}^{\,l} = \frac{E_{\mathrm{R}}}{E_{\mathrm{p}}}\vec{p}_{\mathrm{p}}^{\,l} + |\vec{p}| \; [\hat{r} + (\gamma - 1)(\hat{r} \cdot \hat{n})\hat{n}]$$

$$(9)$$

where $|\vec{p}|$ is as in (2):

$$|\vec{p}| = \frac{\sqrt{s}}{2}\bar{\beta} = \frac{\sqrt{s}}{2}\sqrt{1 - \frac{2(m_{\mathrm{L}}^2 + m_{\mathrm{R}}^2)}{s} + \frac{(m_{\mathrm{L}}^2 - m_{\mathrm{R}}^2)^2}{s^2}}$$

2   Toy Model for Jets

In this section we describe the model we built as well as the previous versions and their issues. We also mention other possibilities for a toy model for jets.

2.1   Minimum requirements

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

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

$$(10)$$

where $\Delta R_{ij} = (\phi_i - \phi_j)$ in a 2D model, with ϕ_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 meaningful definition of d_{ij} the toy model should be at least 2D.

2.2   Current model: Model 4

We build a 2D model in the (y, z) plane, where $\hat{z}$ is the direction of the beam axis and $\hat{y}$ the transverse direction. We define $p_{\mathrm{T}} = |p_y|$.

2.2.1   Splitting of a node

Given a parent node with momentum $\vec{p}_{\mathrm{p}}$, and a scalar value Δ_{p} that sets the scale of the splitting, we define the splitting function as follows:

3

1. We draw a value ϕ_p for the angle in the (y,z) plane, from a uniform distribution in the range $\{0, 2\pi\}$ and define

$$\vec{\Delta}_\mathrm{p} = \Delta_\mathrm{p} \left(\sin\phi_\mathrm{p}, \cos\phi_\mathrm{p} \right) \tag{11}$$

2. We define the L,R children nodes momentum following (6),

$$\begin{aligned}
\vec{p}_\mathrm{L} &= \frac{1}{2}\vec{p}_\mathrm{p} - \vec{\Delta}_\mathrm{p} \\
\vec{p}_\mathrm{R} &= \frac{1}{2}\vec{p}_\mathrm{p} + \vec{\Delta}_\mathrm{p}
\end{aligned} \tag{12}$$

3. We separately draw r_L and r_R from an exponential distribution

$$f(r, \lambda) = \lambda e^{-\lambda r} \tag{13}$$

and generate

$$\begin{aligned}
\Delta_\mathrm{L} &= \Delta_\mathrm{p}\, r_\mathrm{L} \\
\Delta_\mathrm{R} &= \Delta_\mathrm{p}\, r_\mathrm{R}
\end{aligned} \tag{14}$$

2.2.2 Generative process

Input parameters:

- $\vec{p}_0$: momentum of the jet. This will be the input value for the root node of the tree.

- λ: decaying rate for the exponential distribution.

- Δ_0: Initial scale (units of energy).

- Δ_cut: cut-off scale to stop the showering process (units of energy).

We start with the root of the tree as the parent node and build the jet binary tree recursively as follows:

- We split the parent node, and get $\vec{p}_\mathrm{L}$, $\vec{p}_\mathrm{R}$, Δ_L, Δ_R.

- If $\Delta_\mathrm{L/R} > \Delta_\mathrm{cut}$, we promote the L/R node to a parent node (i.e. we promote $\vec{p}_\mathrm{L/R}$ to $\vec{p}_\mathrm{p}$ and $\Delta_\mathrm{L/R}$ to Δ_p), and split again.

2.2.3   Heavy resonance vs QCD like jet

To model a jet coming from a heavy resonance X decay, e.g. a W boson jet, we use $\Delta_{\mathrm{p}} = \frac{m_X}{2}$ to split the root node.

To model a QCD like jet, we split the root node with $\Delta_{\mathrm{p}} = \Delta_0\, r_{\mathrm{root}}$, where r_{root} is drawn from the exponential distribution (13).

2.2.4   Reconstruct $\{\vec{p}_{\mathbf{p}}, \Delta_{\mathbf{p}}, \phi_{\mathbf{p}}, r_L, r_R\}$ from $\{\Delta_L, \Delta_R, \vec{p}_{\mathbf{L}}, \vec{p}_{\mathbf{R}}\}$

We show how to reconstruct $\{\vec{p}_{\mathrm{p}}, \Delta_{\mathrm{p}}, \phi_{\mathrm{p}}, r_L, r_R\}$ from a bottom-up approach for each splitting. From $\{\Delta_L, \Delta_R, \vec{p}_{\mathrm{L}}, \vec{p}_{\mathrm{R}}\}$ we get:

$$\vec{p}_{\mathrm{p}} = \vec{p}_{\mathrm{L}} + \vec{p}_{\mathrm{R}}$$
$$\vec{\Delta}_{\mathrm{p}} = \frac{1}{2}(\vec{p}_{\mathrm{R}} - \vec{p}_{\mathrm{L}}) \tag{15}$$

Next, we get

$$\Delta_{\mathrm{p}} = |\vec{\Delta}_{\mathrm{p}}|$$
$$\phi_{\mathrm{p}} = \tan^{-1}\frac{(\Delta_{\mathrm{p}})_y}{(\Delta_{\mathrm{p}})_z}$$
$$r_{\mathrm{L}} = \frac{\Delta_{\mathrm{L}}}{\Delta_{\mathrm{p}}}$$
$$r_{\mathrm{R}} = \frac{\Delta_{\mathrm{R}}}{\Delta_{\mathrm{p}}} \tag{16}$$

If the L/R node is a leaf, then we set

$$\Delta_{\mathrm{L/R}} = \Delta_{\mathrm{cut}} \tag{17}$$

2.3   Previous toy models and their issues

2.3.1   Model 1

This was a 1D model where d_{ij} in (10) was not well defined.

2.3.2   Model 2

In this model we defined the splitting of (6) as

$$\vec{p}_{\mathrm{L}} = \vec{p}_{\mathrm{p}} - \vec{\Delta}_{\mathrm{p}}$$
$$\vec{p}_{\mathrm{R}} = \vec{p}_{\mathrm{p}} + \vec{\Delta}_{\mathrm{p}} \tag{18}$$

5

which gives

$$\vec{p}_\text{p} = \frac{1}{2}(\vec{p}_\text{L} + \vec{p}_\text{R}) \tag{19}$$

As a result, we would get jets with different momentum $\vec{p}$ depending on the clustering algorithm, e.g. k_t, anti-k_t, etc. The reason is that the jet momentum in this case is not permutation invariant with respect to the order in which we cluster its constituents.[1]

2.3.3 Model 3

This model differs from Model 4 in the way we do the splitting of a node, which also changes the generative process. In this case, given a parent node with momentum $\vec{p}_\text{p}$, and a scalar value Δ_gp that comes from the grandparent node, we define the splitting function as follows:

1. We draw r_p from the exponential distribution of (13) and calculate

$$\Delta_\text{p} = \Delta_\text{gp}\, r_\text{p} \tag{20}$$

2. We draw a value ϕ_p for the angle in the (y,z) plane, from a uniform distribution in the range $\{0, 2\pi\}$ and define as in (11)

$$\vec{\Delta}_\text{p} = \Delta_\text{p}\, (\sin\phi_\text{p}, \cos\phi_\text{p})$$

3. If $\Delta_\text{p} > \Delta_\text{cut}$ we split the parent node by defining the L,R children nodes momentum following (6),

$$\vec{p}_\text{L} = \frac{1}{2}\vec{p}_\text{p} - \vec{\Delta}_\text{p}$$
$$\vec{p}_\text{R} = \frac{1}{2}\vec{p}_\text{p} + \vec{\Delta}_\text{p}$$

Next, we promote the L and R nodes to a parent node each (i.e. we promote $\vec{p}_\text{L/R}$ to $\vec{p}_\text{p}$ and Δ_p to Δ_gp), and repeat the splitting process.

The problem with this model is that we cannot reconstruct Δ_gp from a bottom-up approach. As a result, we cannot derive from (20) the value r_p drawn from the distribution.

---

[1]E.g. when clustering 3 jet constituents with momentum $\{a, b, c\}$ we see $\frac{1}{2}(\frac{a+b}{2}) + \frac{c}{2} \neq \frac{1}{2}(\frac{a+c}{2}) + \frac{b}{2}$

2.4 Other ideas for Toy Generative Models

We could define $\vec{\Delta}$ as in (7):

$$\vec{\Delta} = E\sqrt{1 - \frac{m^2}{E^2}} \ [\hat{r} + (\gamma - 1)(\hat{r} \cdot \hat{n})\hat{n}] = \vec{\Delta}(m, \phi) \tag{21}$$

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

$$m_{\text{L/R}} = \frac{m_{\text{p}}}{2} \ r_{\text{L/R}} \tag{22}$$

where $r_{\text{L/R}}$ is drawn from an exponential distribution $f(x, \lambda) \propto e^{-\lambda x}$ for $x \in (0,1)$. The prescription of (22) solves one of the problems of the traditional parton showers given that it satisfies $m_{\text{L}} + m_{\text{R}} \leq m_{\text{p}}$[2]. Then, we could think of $m_{\text{L/R}}$ 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 a physics parton shower but more complex and time consuming.

2.4.1 General case, with $m_{\textbf{L}} \neq m_{\textbf{R}}$

We could also build a model for this case, adding extra features. We should replace (22) by

$$\begin{aligned} m_{\text{L}} &= m_{\text{p}} \ r_{\text{L}} \\ m_{\text{R}} &= (m_{\text{p}} - m_{\text{L}}) \ r_{\text{R}} \end{aligned} \tag{23}$$

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

---

[2]Traditional parton showers draw m_{L} and m_{R} independently, so they should check the constraint is satisfied.