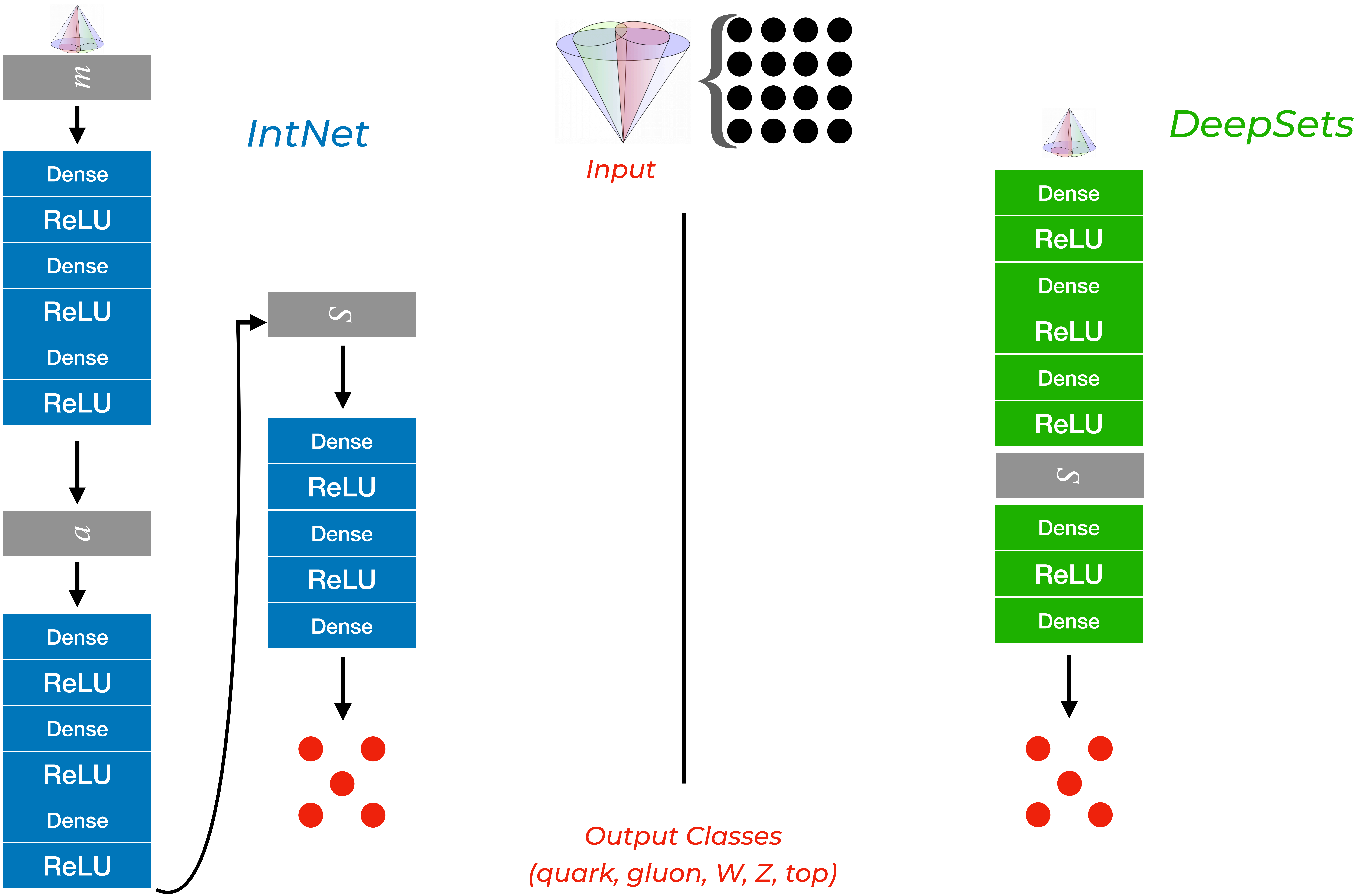
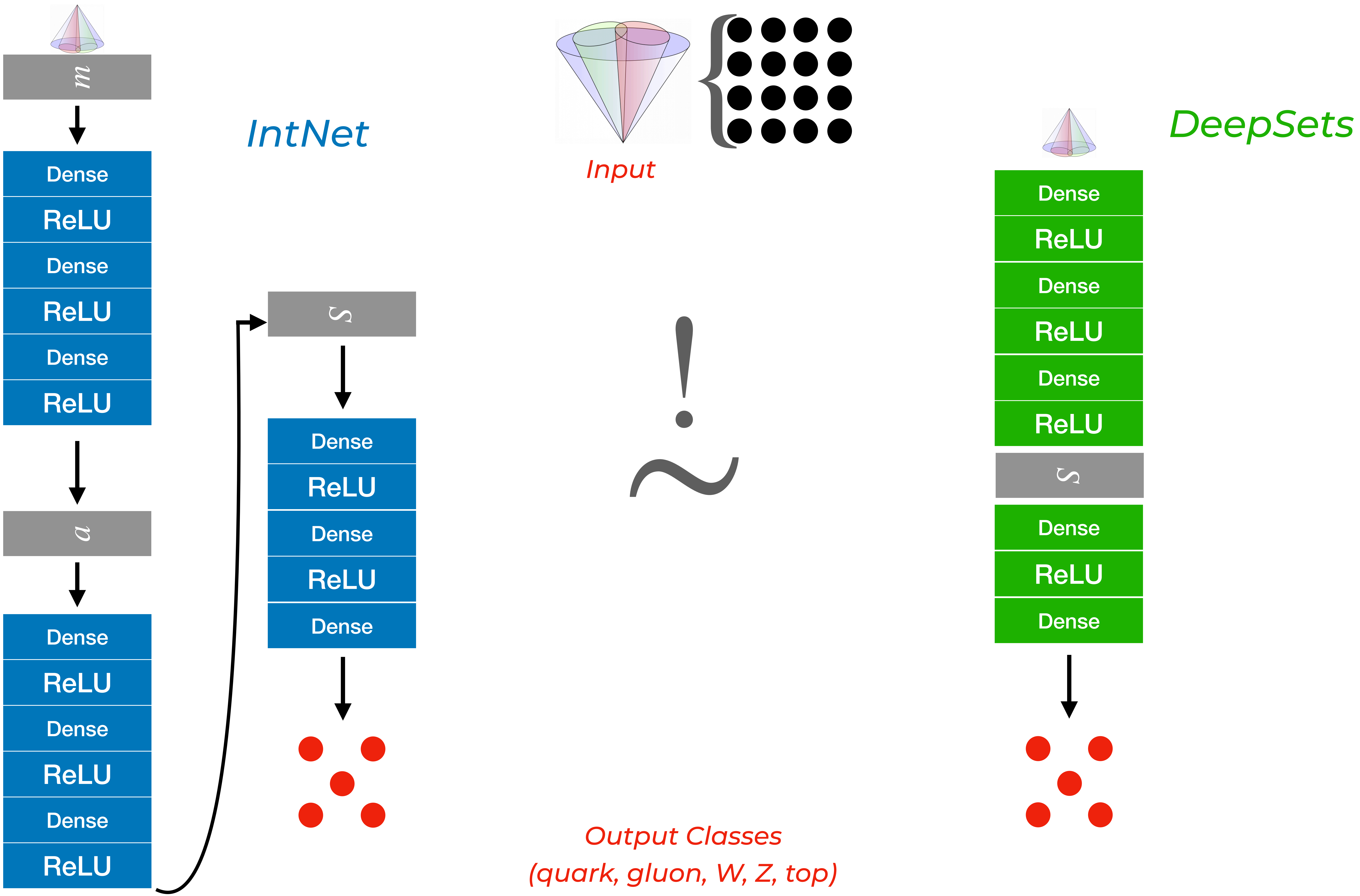
IntNets and DeepSets for Fast Jets

Patrick Odagiu

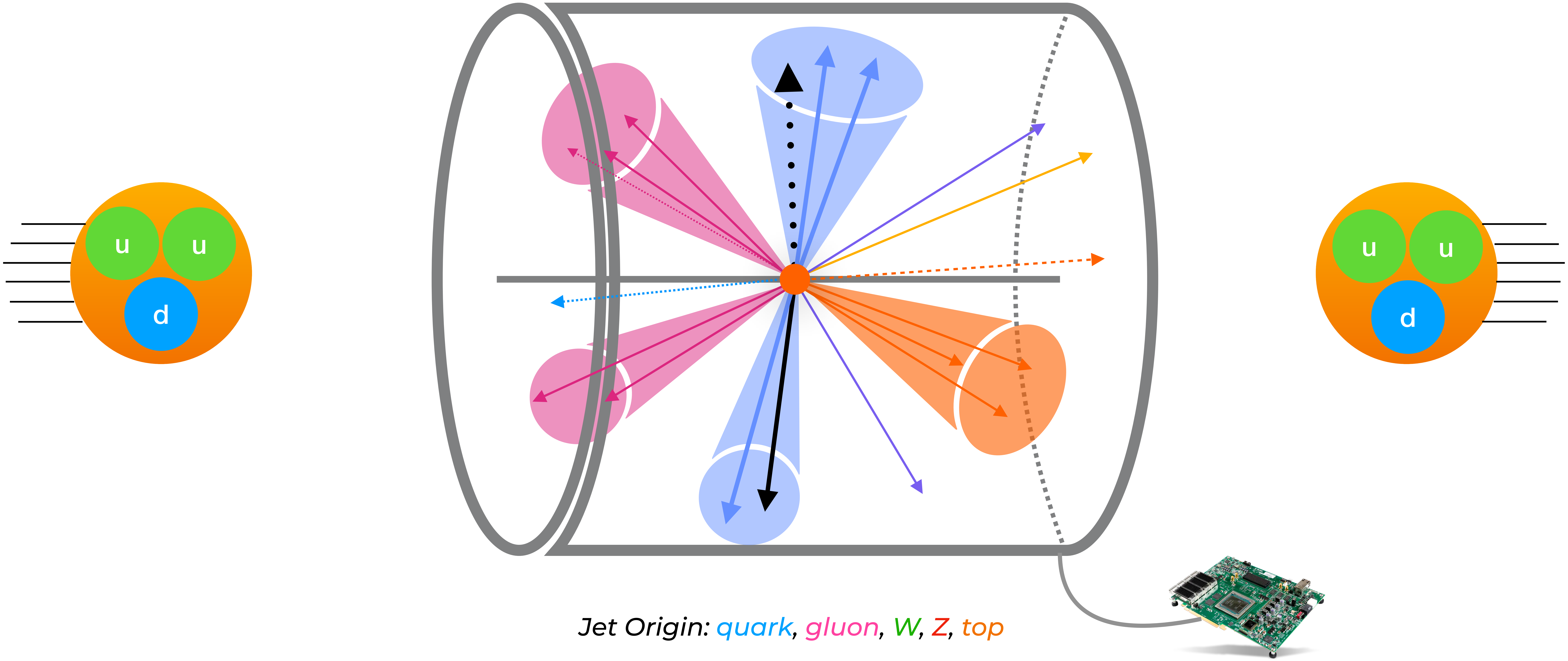
Andre Szjnader, Zhiqiang Que, Javier Duarte, Thea Aarrestad, Vladimir Loncar, Jennifer Ngadiuba, Philipp Rincke, Johannes Haller, Gregor Kasieczka, Ihor Komarov, Finn Labe, Artur Lobanov, Matthias Schroder, Maurizio Pierini, Arpita Seksaria, Wayne Luk, and Sioni Summers


m
IntNet
Dense
ReLU
Dense
ReLU
Dense
ReLU
a
Dense
ReLU
Dense
ReLU
Dense
ReLU
S
Dense
ReLU
Dense
ReLU
Dense
Input
Output Classes
(quark, gluon, W, Z, top)
DeepSets
Dense
ReLU
Dense
ReLU
Dense
ReLU
S
Dense
ReLU
Dense

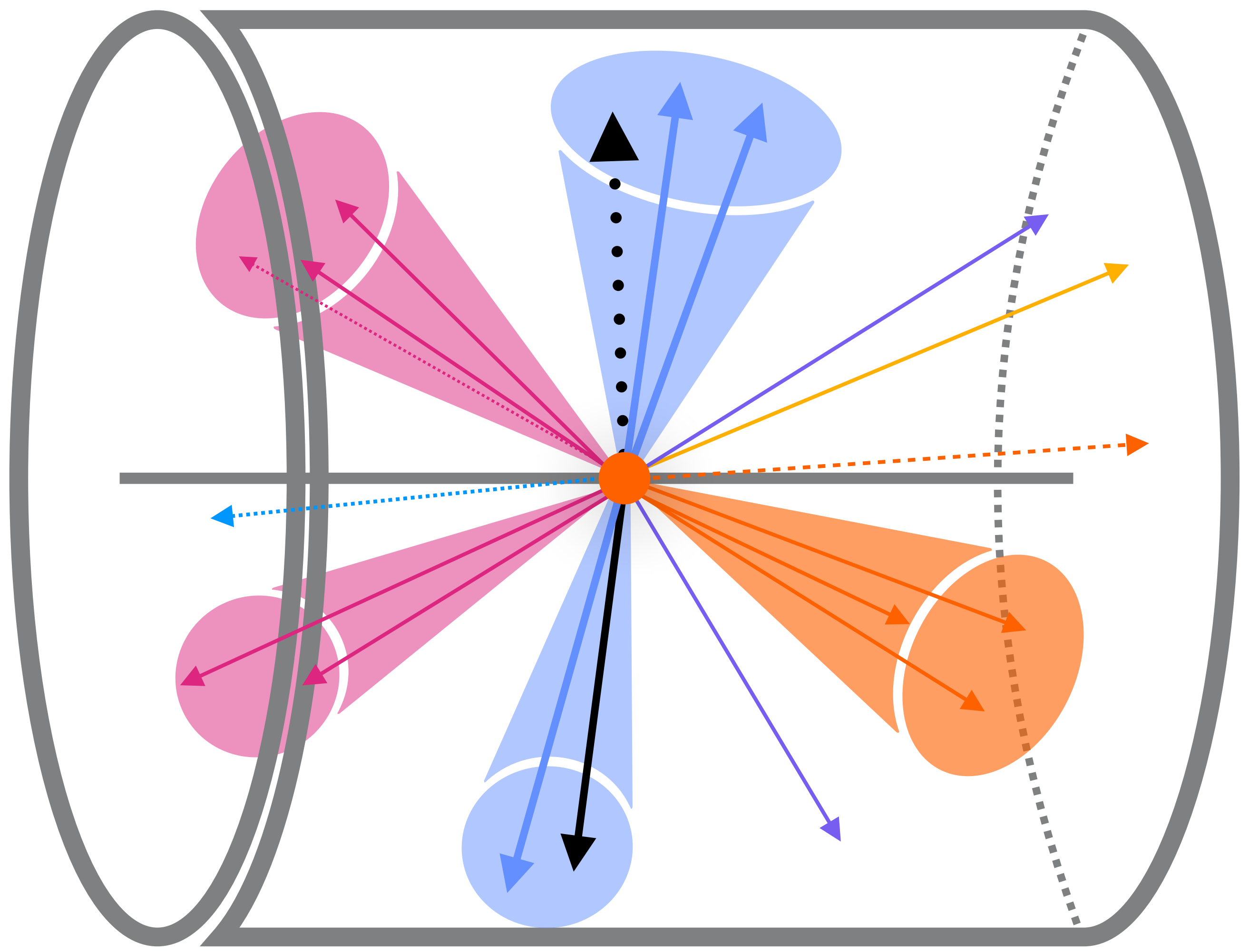


IntNet
Input
m
Dense
ReLU
Dense
ReLU
Dense
ReLU
a
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
DeepSets
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
Output Classes
(quark, gluon, W, Z, top)

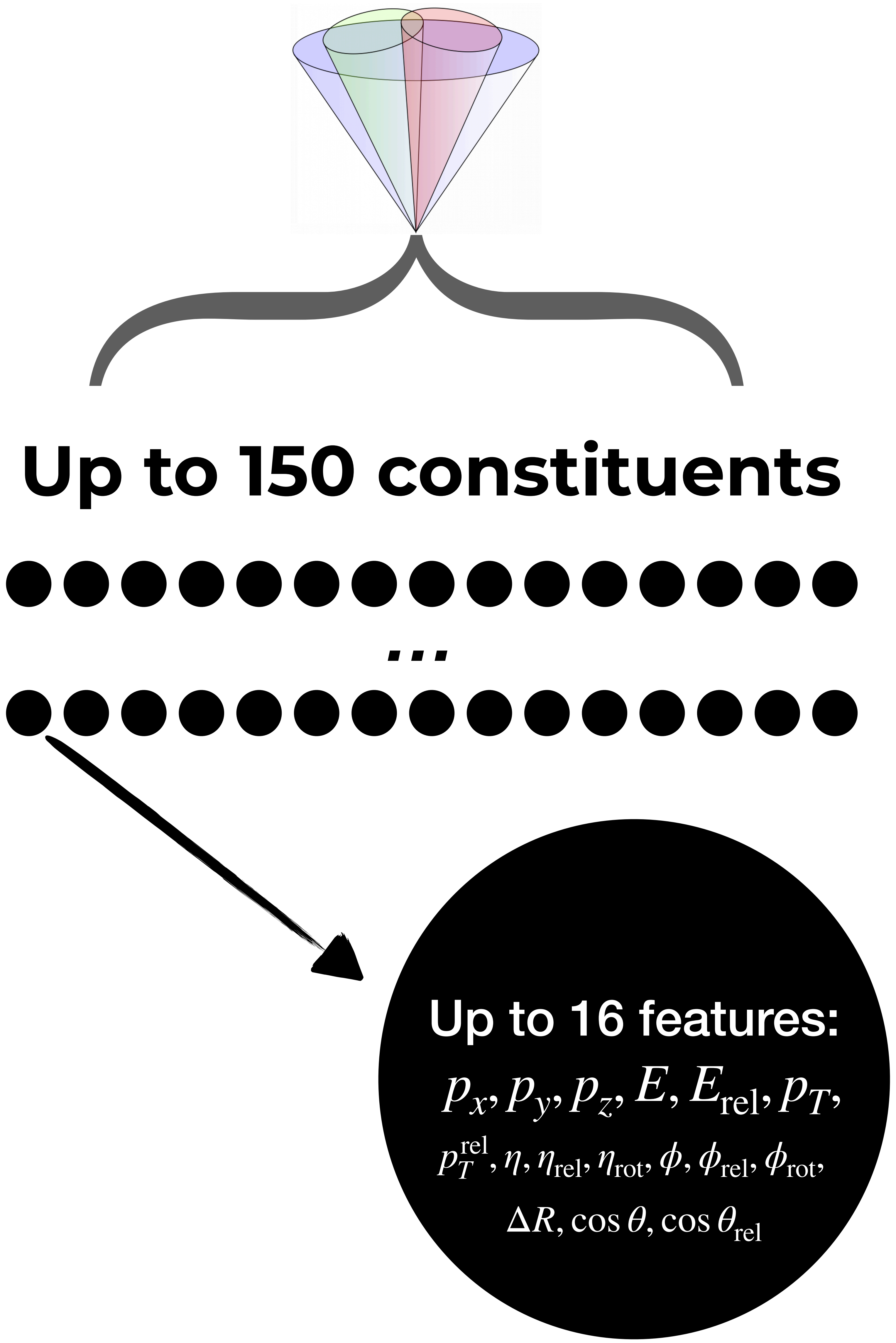
Collision every 25 ns: ~60 Tb/s data influx



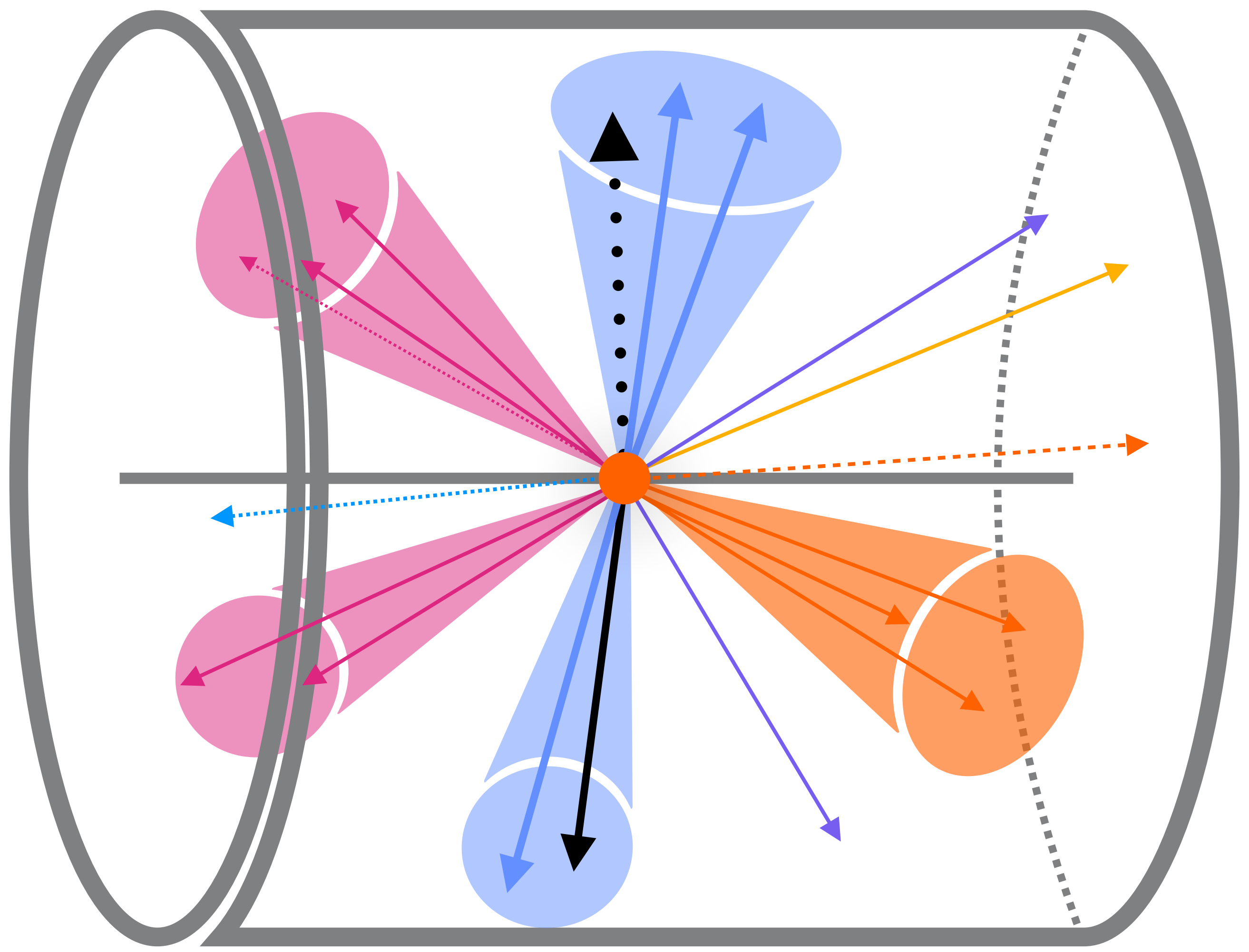
Jet Origin: quark, gluon, W, Z, top



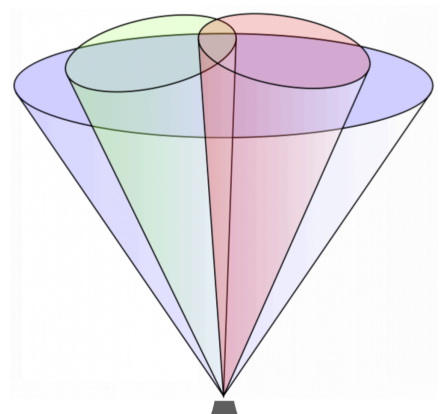
853 390 jets


Up to 150 constituents
...
Up to 16 features:
$p_x, p_y, p_z, E, E_{\mathrm{rel}}, p_T,$
$p_T^{\mathrm{rel}}, \eta, \eta_{\mathrm{rel}}, \eta_{\mathrm{rot}}, \phi, \phi_{\mathrm{rel}}, \phi_{\mathrm{rot}},$
$\Delta R, \cos\theta, \cos\theta_{\mathrm{rel}}$

*speed reasons

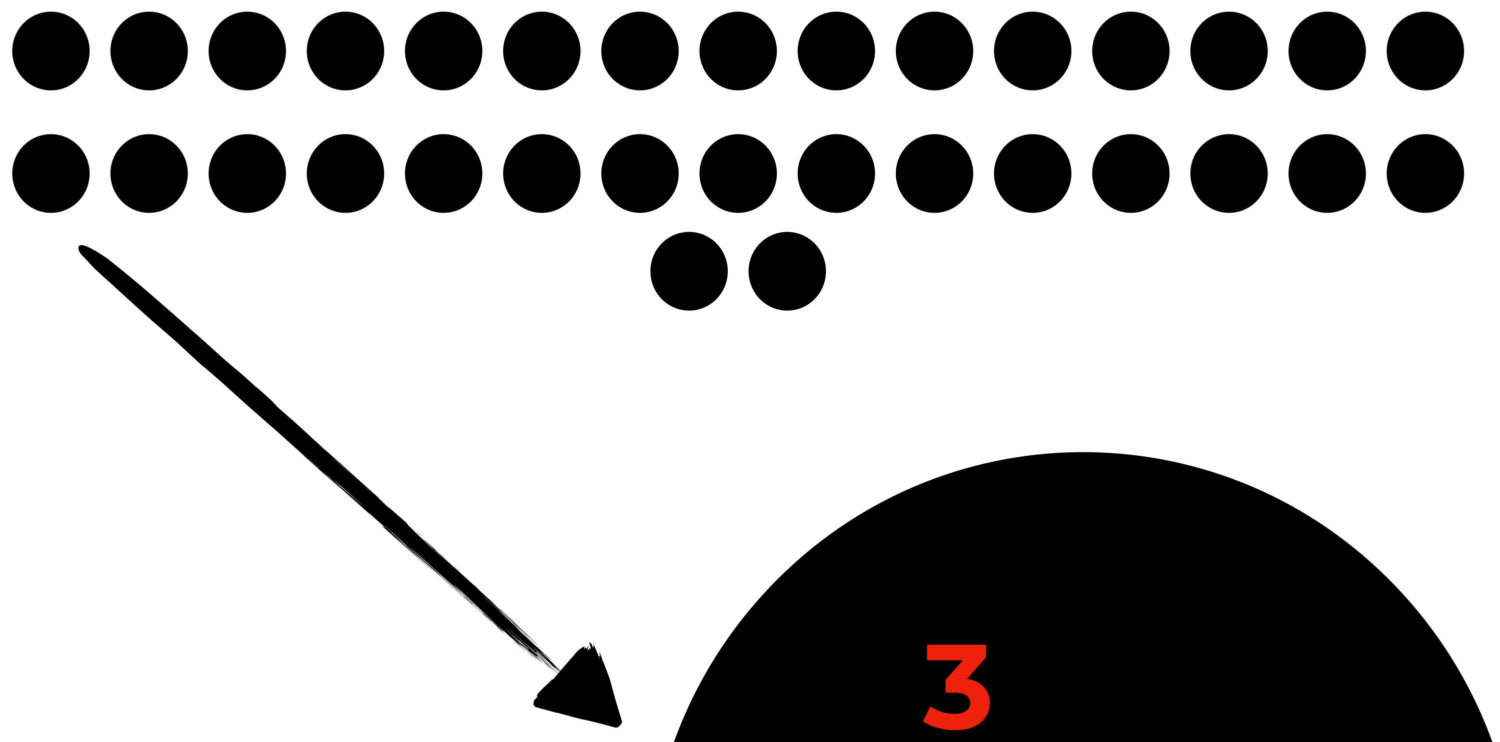


853 390 jets



32 most energetic

Up to ~~150~~ constituents



3

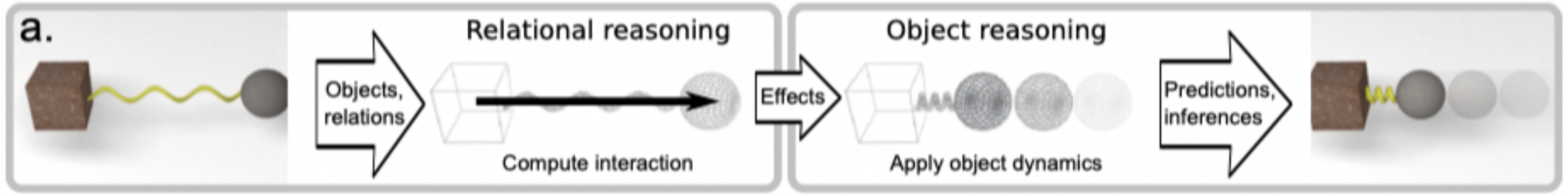
~~Up to 16~~ features:

~~$p_x, p_y, p_z, E, E_{\mathrm{rel}}$~~, p_T,

~~p_T^{rel}, η~~, η_{rel}, ~~$\eta_{\mathrm{rot}}, \phi$~~, ϕ_{rel}, ~~ϕ_{rot}~~

~~$\Delta R, \cos\theta, \cos\theta_{\mathrm{rel}}$~~

An interaction network looks at a complex system in terms of objects and relations, and then in terms of the dynamics and interactions of these elements.



arXiv:1612.00222

$$\text{IN} = \phi_{\text{O}}(a(G, X, \phi_{\text{R}}(m(G))))$$

1. $m(G) \equiv \langle o_i, o_j, r_k \rangle \equiv b_k$

2. $\phi_{\text{R}}(\langle o_i, o_j, r_k \rangle) = e_k$

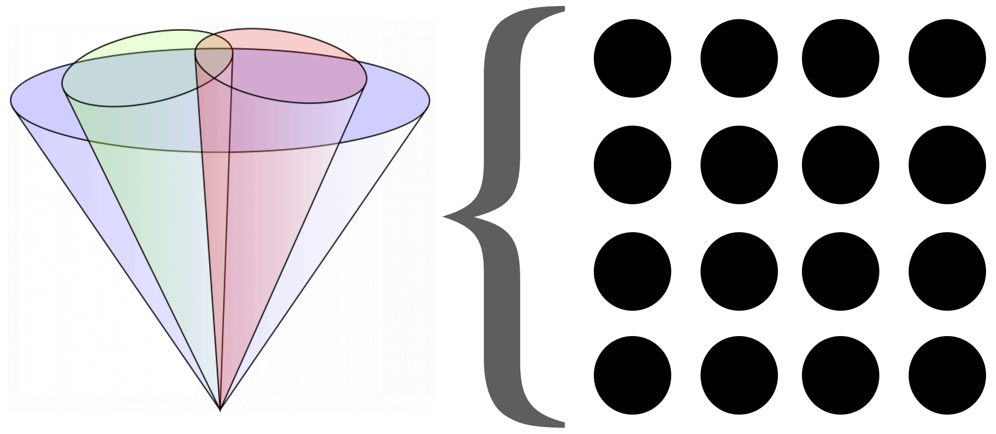
3. $a(G, X, E) \equiv C$

4. $\phi_O(C) = P$

In simulation, can be interpreted as state of the constituents of G at (t+1).

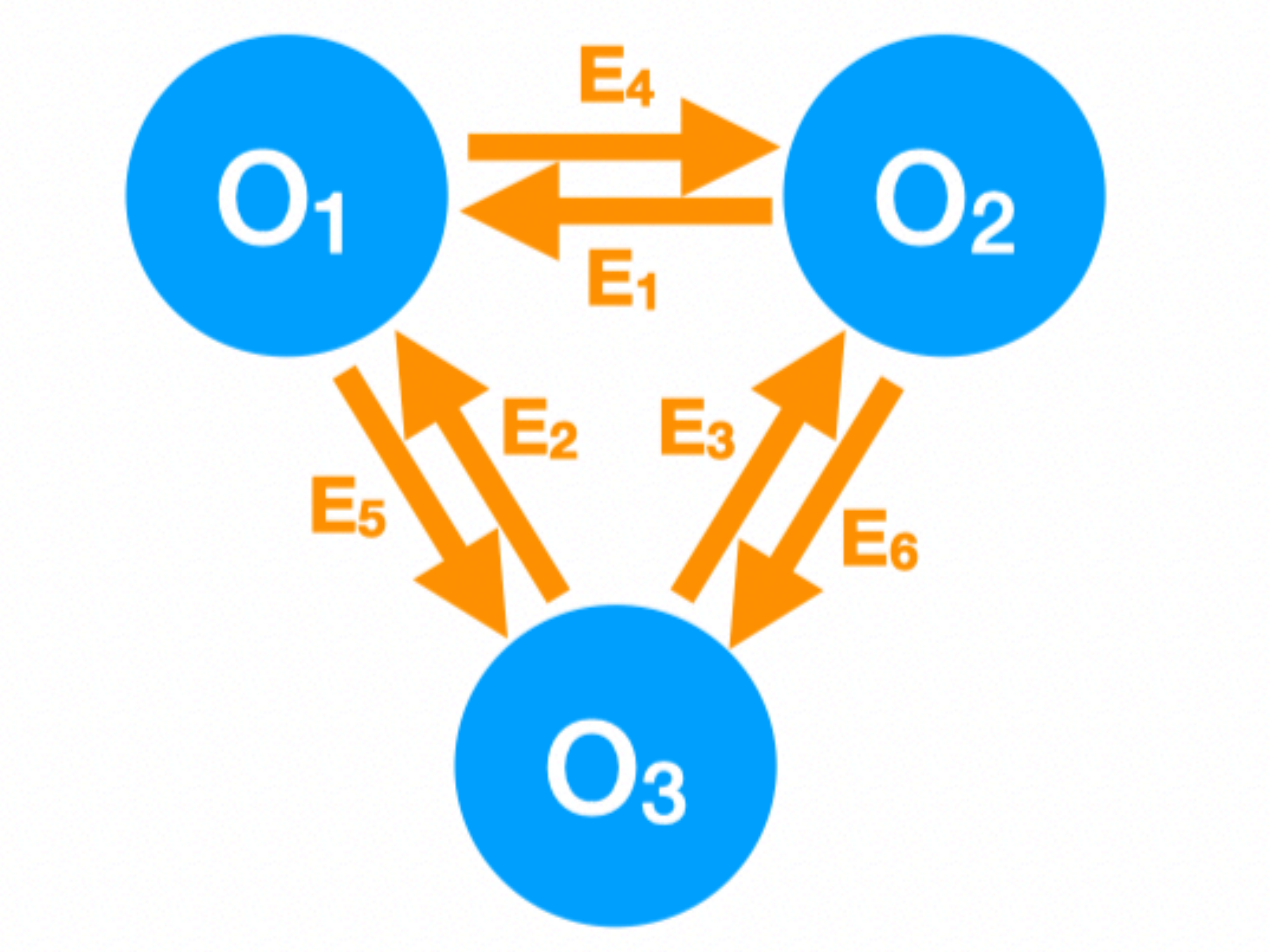
5. $\phi_A(P) = A$

Can learn order of P.



$$I \equiv \begin{pmatrix} p_T^1 & \dots & p_T^{16} \\ \eta^1 & \dots & \eta^{16} \\ \phi^1 & \dots & \phi^{16} \end{pmatrix} \times \left.\begin{matrix} R_S \equiv \begin{pmatrix} 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \\ \dim(R_S) = \dim(R_R) = N_O \times N_E \\ R_R \equiv \begin{pmatrix} 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 0 \end{pmatrix} \end{matrix}\right\} m(G) = B = \begin{pmatrix} I \times R_R \\ I \times R_S \end{pmatrix}$$

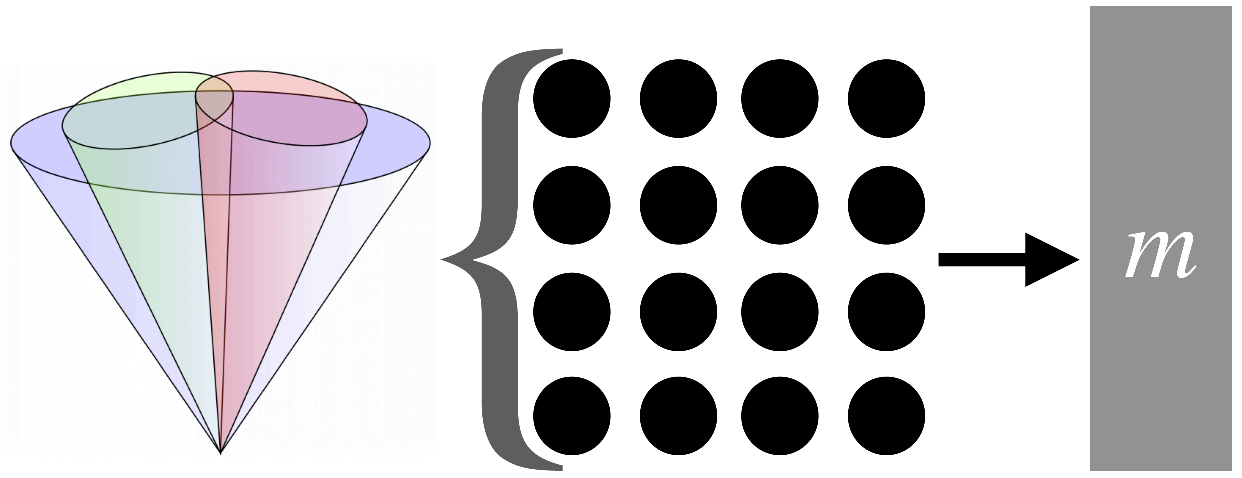
concat.



$$R_S = \begin{array}{c} \\ O_1 \\ O_2 \\ O_3 \end{array} \begin{array}{c} \begin{matrix} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

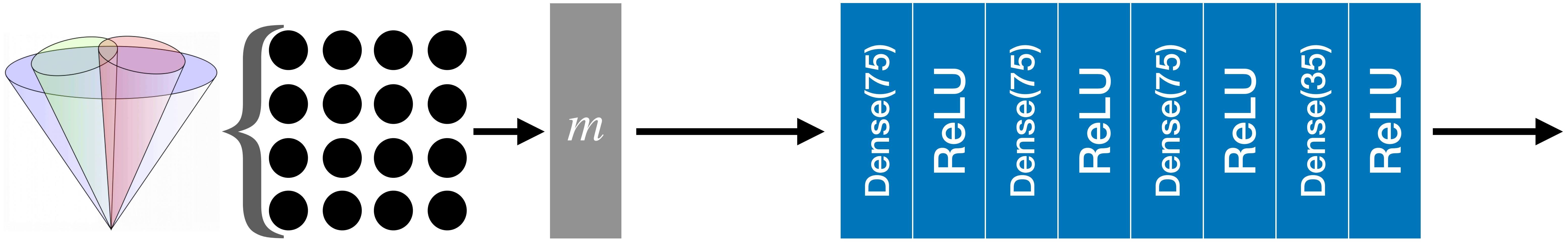
$$R_R = \begin{array}{c} \\ O_1 \\ O_2 \\ O_3 \end{array} \begin{array}{c} \begin{matrix} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 \end{matrix} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}.$$

arXiv:1908.05318



$$m(G) = B = \begin{pmatrix} I \times R_R \\ I \times R_S \end{pmatrix}$$

$$\dim(B) = (240,\ 6)$$



$$m(G) = B = \begin{pmatrix} I \times R_R \\ I \times R_S \end{pmatrix}$$

$$\dim(B) = (240,\ 6)$$

$$\phi_{\mathrm{R}}(B) = E$$

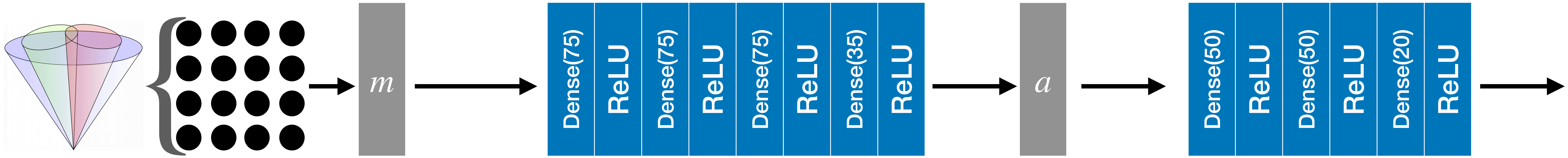
Cumulative effects of interaction received by a given vertex.

$$\bar{E} = ER_R^\top$$

Concatenate to input.

$$C = \begin{pmatrix} I \\ \bar{E} \end{pmatrix}$$

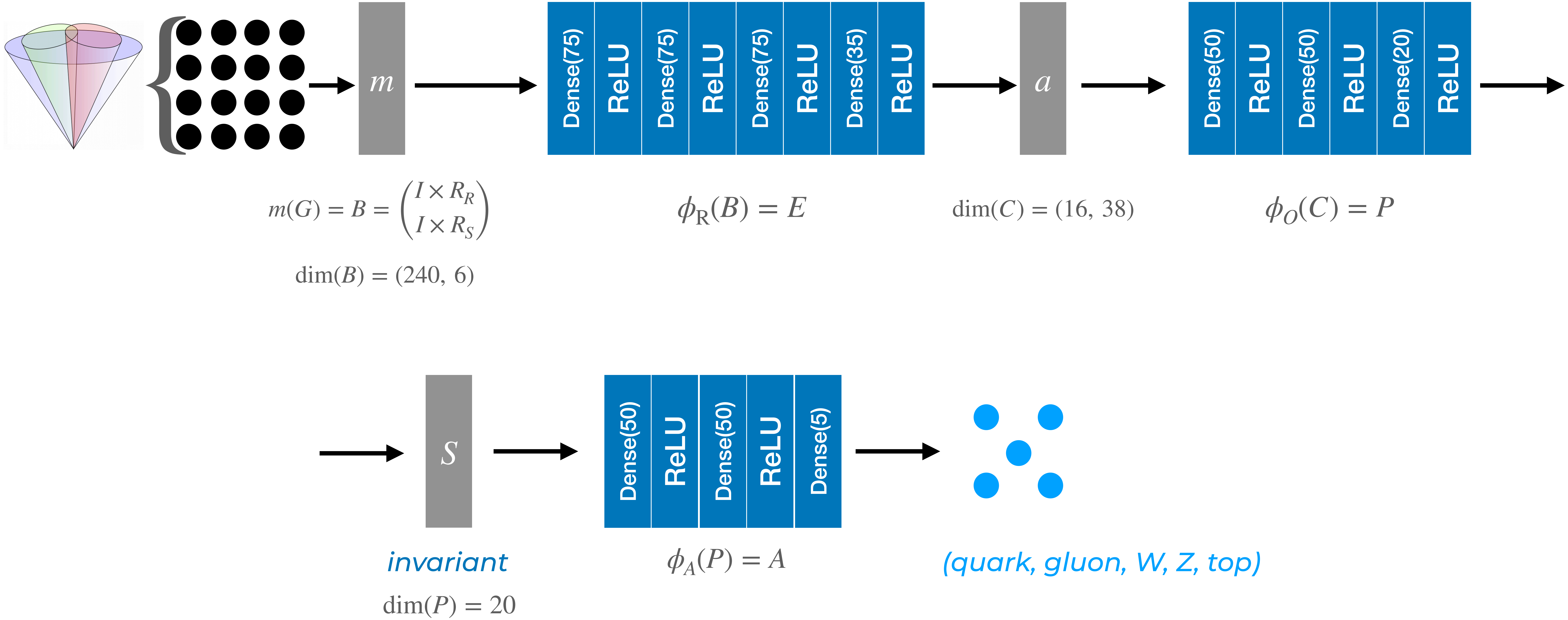
arXiv:1908.05318



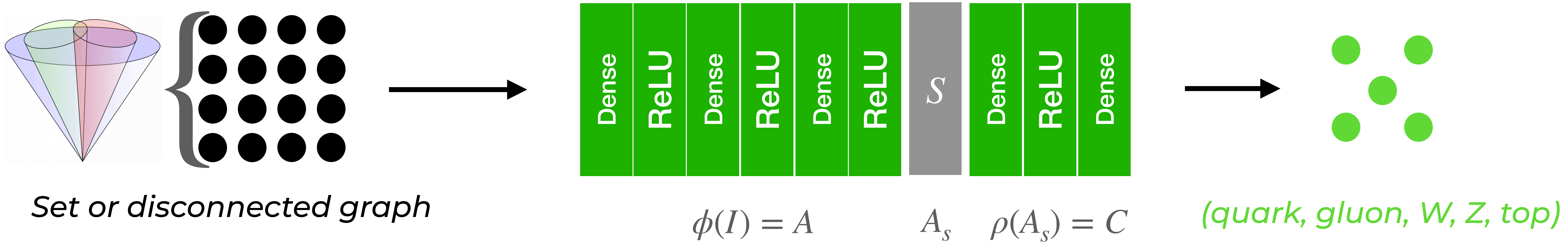
$$m(G) = B = \begin{pmatrix} I \times R_R \\ I \times R_S \end{pmatrix} \qquad \phi_{\mathrm{R}}(B) = E \qquad \dim(C) = (16,\ 38) \qquad \phi_O(C) = P$$

$$\dim(B) = (240,\ 6)$$

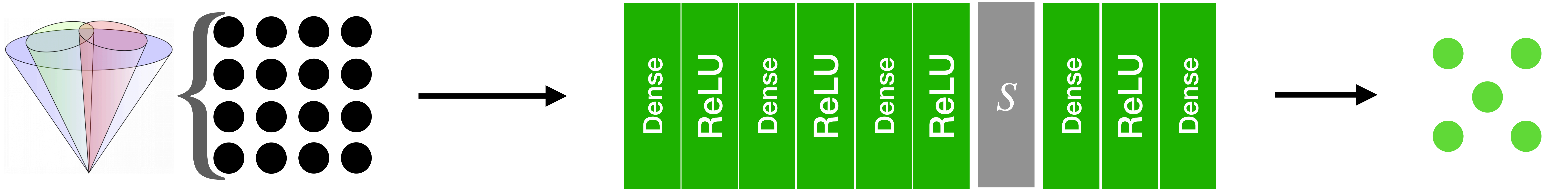
arXiv:1908.05318



arXiv:1908.05318



arXiv:1703.06114



Set or disconnected graph

$X = \{x_1, \ldots, x_M\},\ x_m \in \mathfrak{X}$

$\phi(I) = A$ $\quad A_s \quad$ $\rho(A_s) = C$

(quark, gluon, W, Z, top)

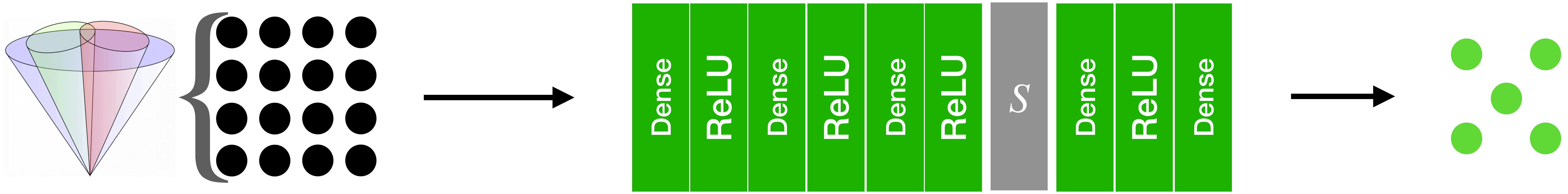
$$f : 2^{\mathfrak{X}} \rightarrow \mathcal{Y}$$

$$f(\{x_1, \ldots, x_M\}) = f(\{x_{\pi(1)}, \ldots, x_{\pi(M)}\})$$

invariant

$$\rho\left(\sum_{x \in X} \phi(x) \right)$$

arXiv:1703.06114



Set or disconnected graph

$X = \{x_1, \ldots, x_M\},\ x_m \in \mathfrak{X}$

$\phi(I) = A$ $\quad A_s \quad$ $\rho(A_s) = C$

(quark, gluon, W, Z, top)

$c : \mathfrak{X} \rightarrow \mathbb{N}$

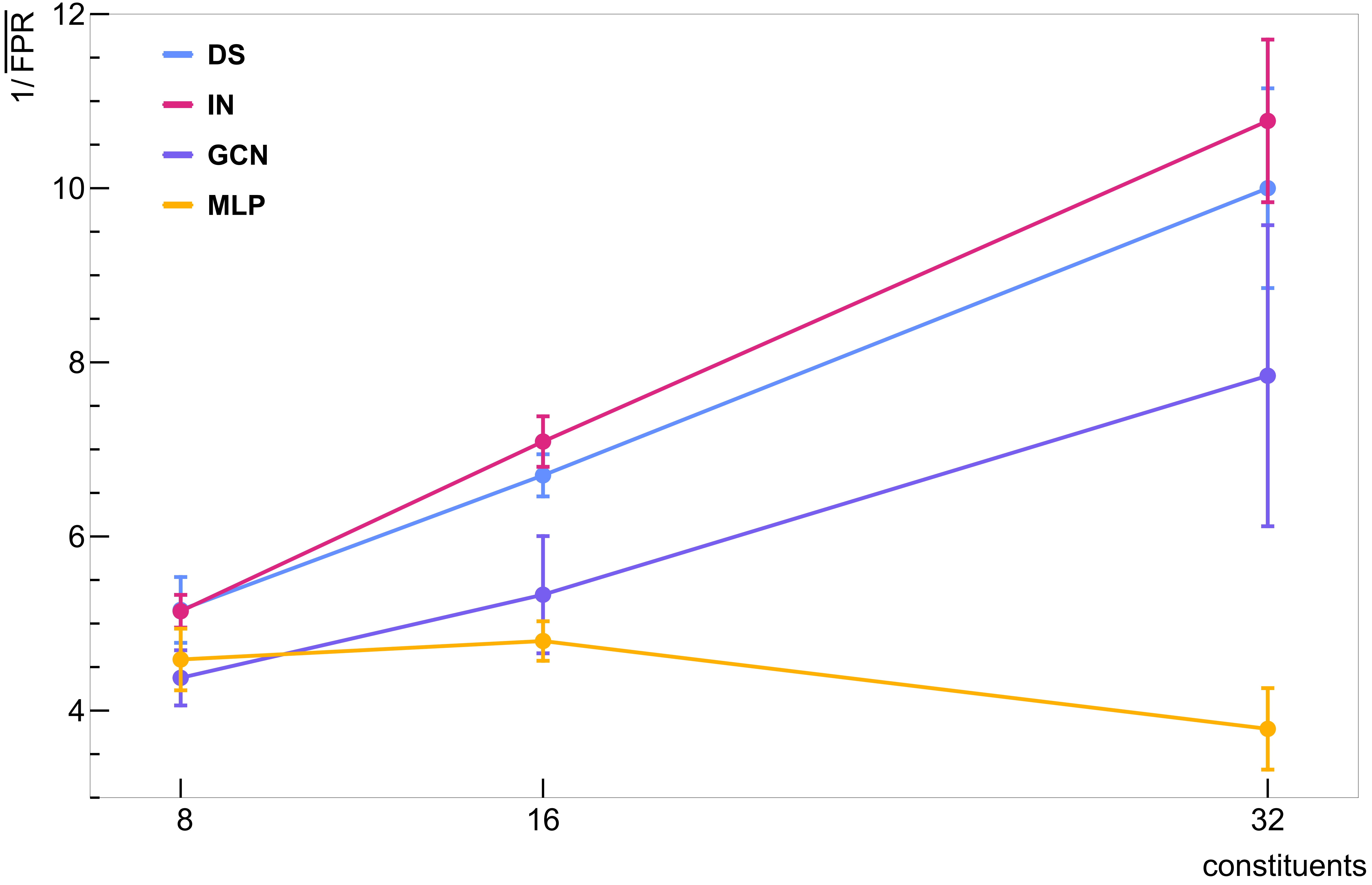
$\phi(x) = 4^{-c(x)}$

then $\sum_{x \in X} \phi(x)$ *is a unique representation for every* $X \in 2^{\mathfrak{X}}$

similarly construct $\rho : \mathbb{R} \rightarrow \mathbb{R}$ s.t. $f(X) = \rho\left(\sum_{x \in X} \phi(x)\right)$

**Countable Sets*

arXiv:1703.06114



$1/\overline{FPR}$
12
10
8
6
4
DS
IN
GCN
MLP
8
16
32
constituents

Xilinx VU9P

Latencies
(ns)

*8 bit quantised

n particles	MLP	DS	IN
8	55	75	175
16	55	95	230
32	55	130	265

Xilinx VU9P

Resources
(LUTs)

**8 bit quantised*

n particles	MLP	DS	IN
8	7.6%	31%	30.8%
16	7.5%	63%	57.6%
32	7.2%	118%	97.2%

- ★ **Set representation ~ fully connected graph (for this data!)**
  - ✴ better representation yet to be found?
  - ✴ other symmetries to be exploited? e.g. rotational invariance
- ★ **Incorporate symmetries bottom-up**
  - ✴ gain intimate understanding of symmetries in data
  - ✴ test each hypothesis separately
- ★ **Geometrical models: smart but more importantly, FAST!**