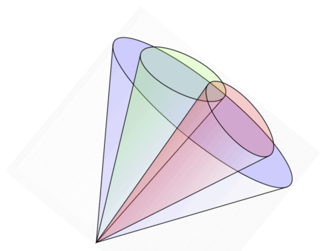
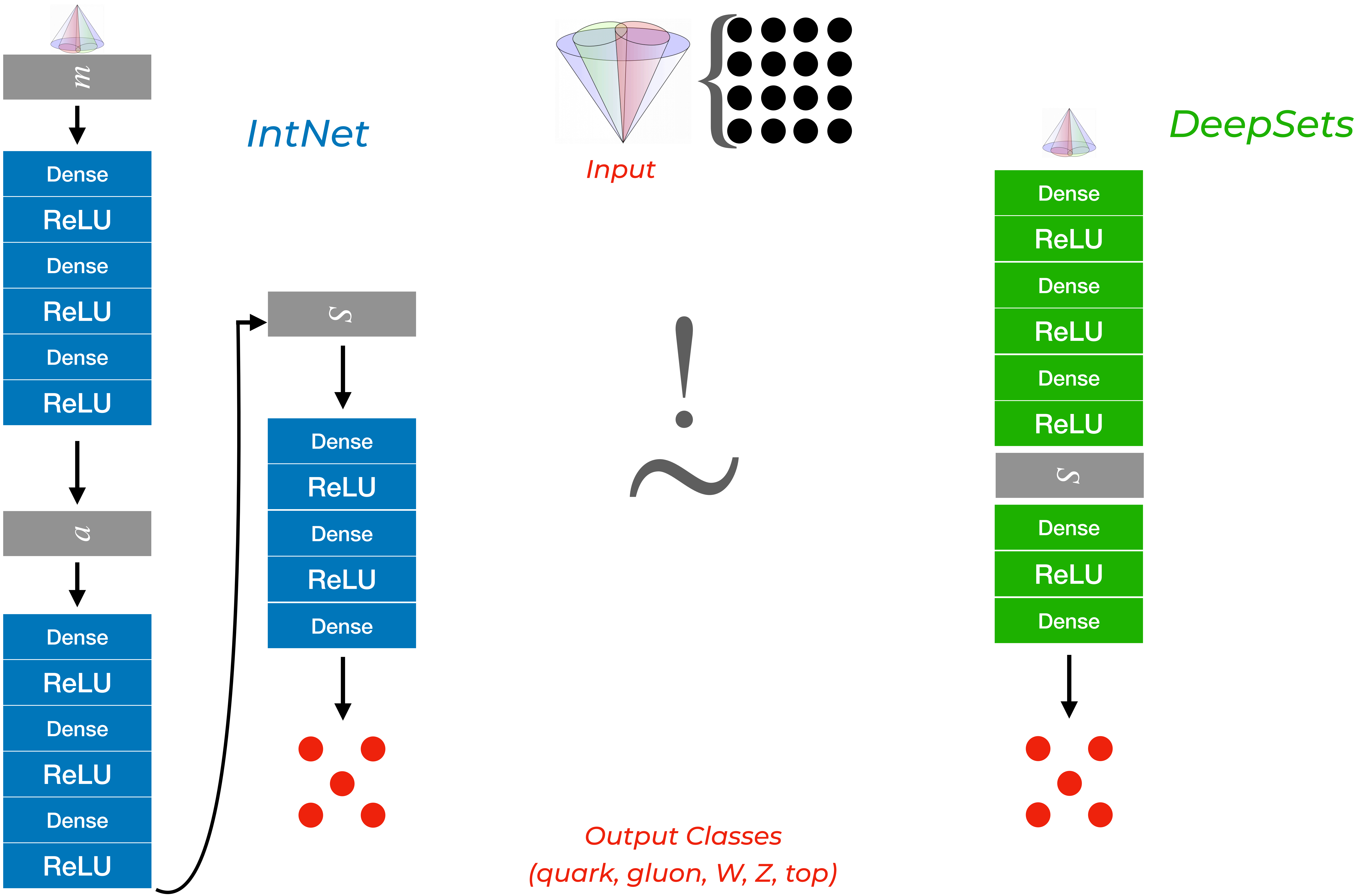


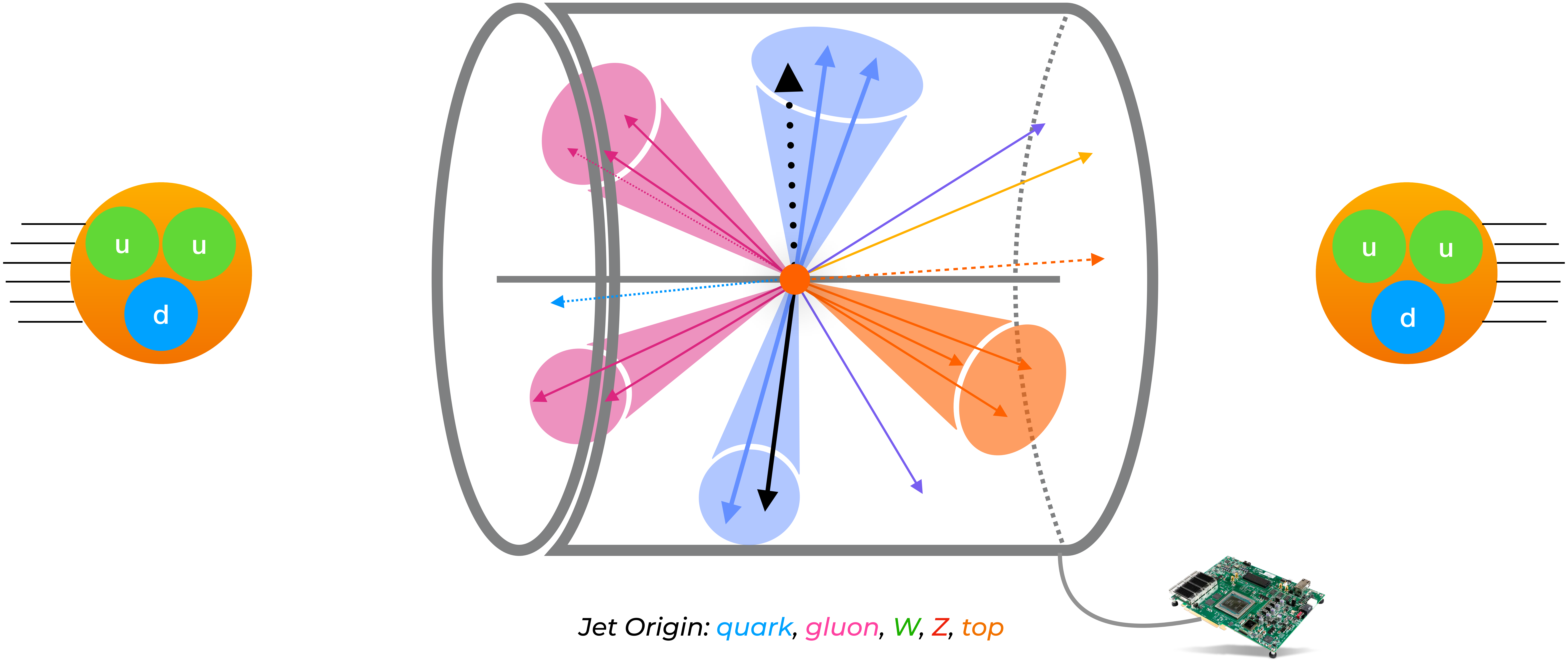
IntNets and DeepSets for Fast Jets

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

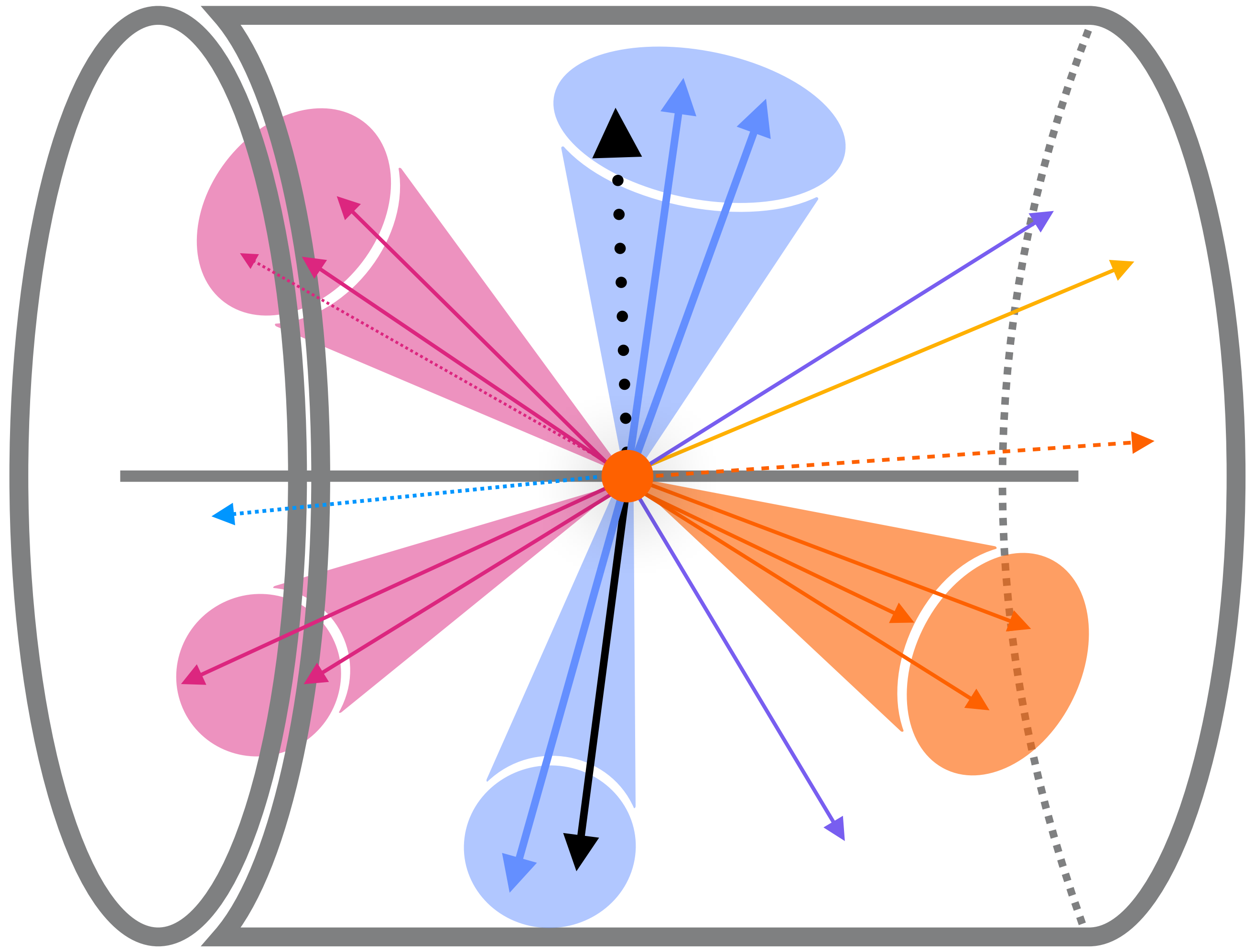

IntNet
Input
DeepSets
m
Dense
ReLU
Dense
ReLU
Dense
ReLU
a
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
ReLU
Dense
Output Classes
(quark, gluon, W, Z, top)

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



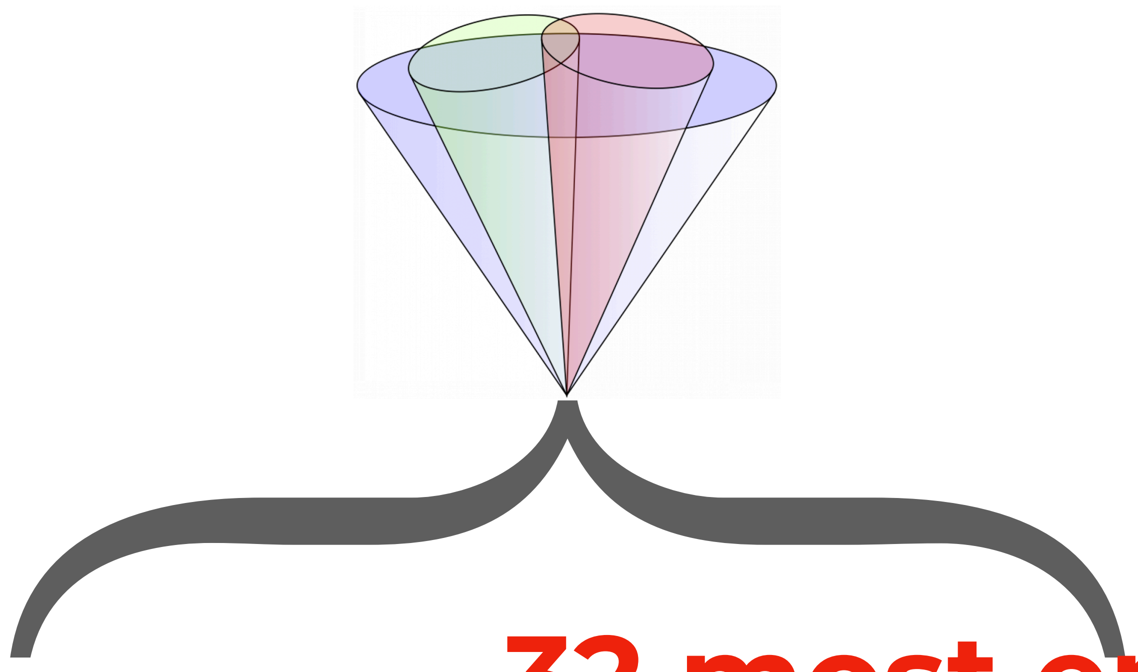
Jet Origin: quark, gluon, W, Z, top

*speed reasons



853 390 jets

https://zenodo.org/records/3602260



32 most energetic

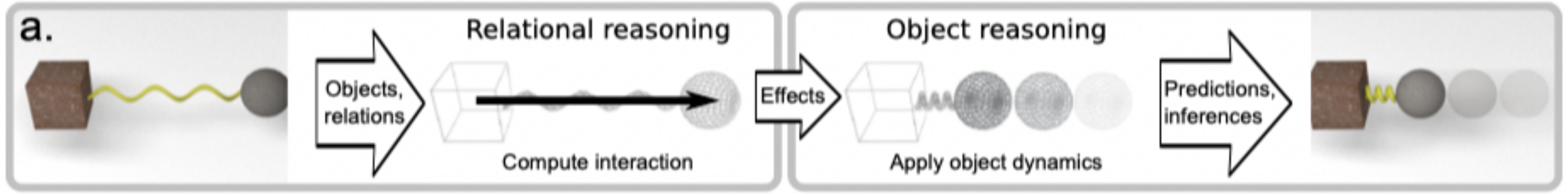
Up to ~~150~~ constituents

3

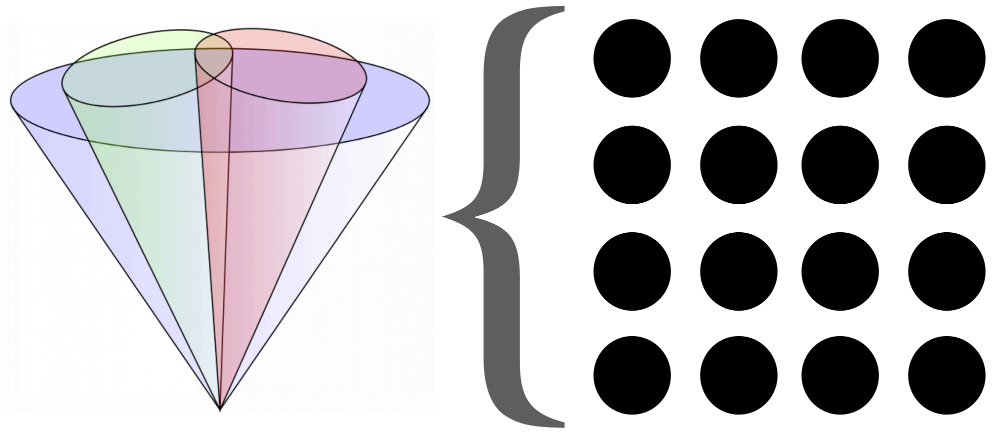
~~Up to 16~~ features:

~~$p_x, p_y, p_z, E, E_{\mathrm{rel}}$~~, p_T, ~~p_T^{rel}, η~~, η_{rel}, ~~η_{rot}~~, $\phi, \phi_{\mathrm{rel}}$, ~~$\phi_{\mathrm{rot}}$~~, ~~$\Delta R, \cos\theta, \cos\theta_{\mathrm{rel}}$~~

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.

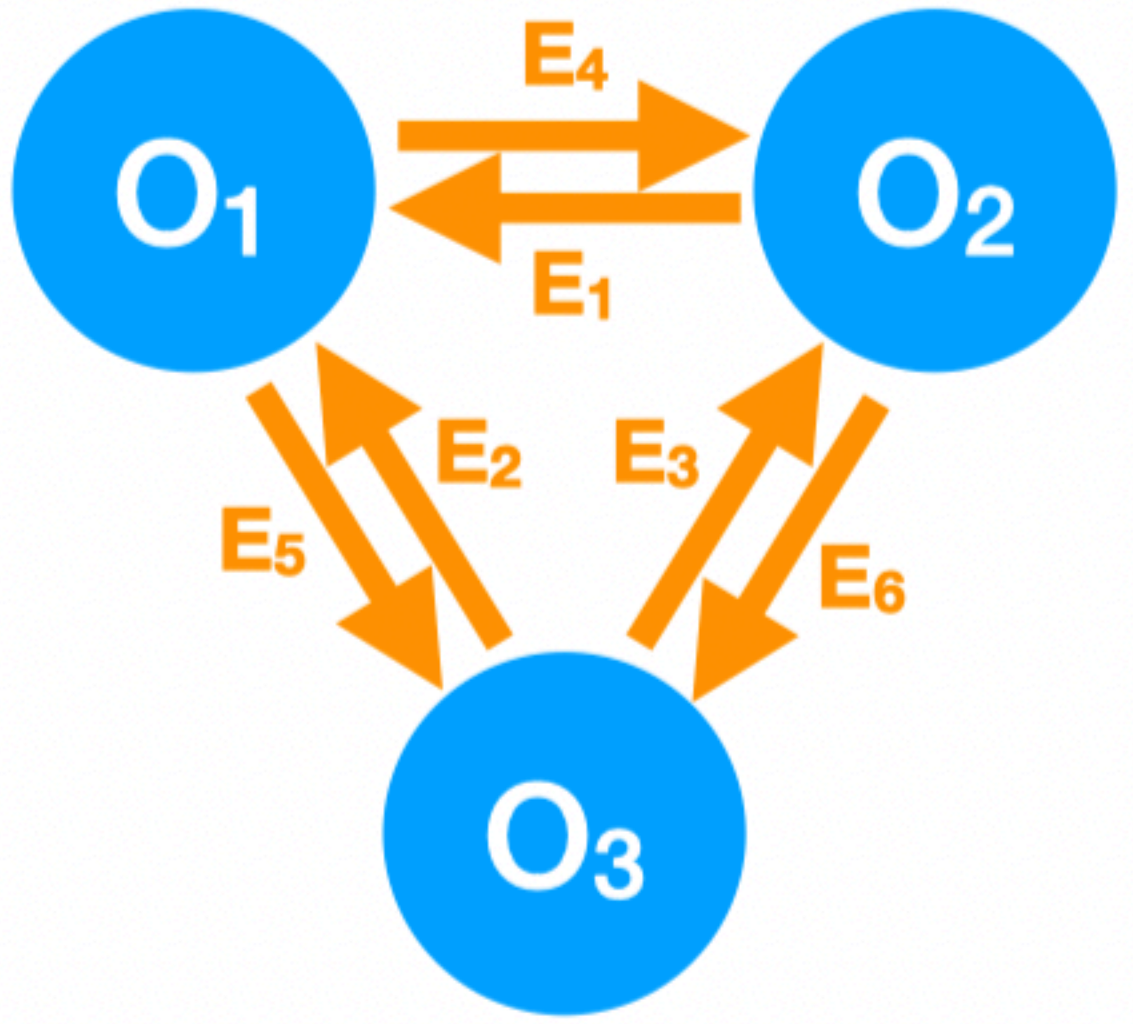


arXiv:1612.00222



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

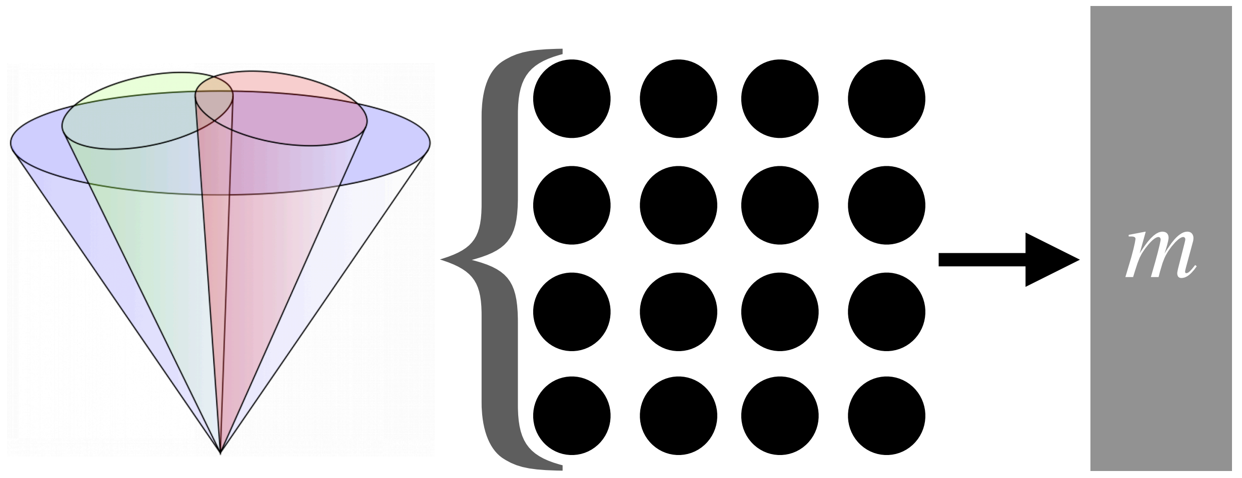
concat.



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

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

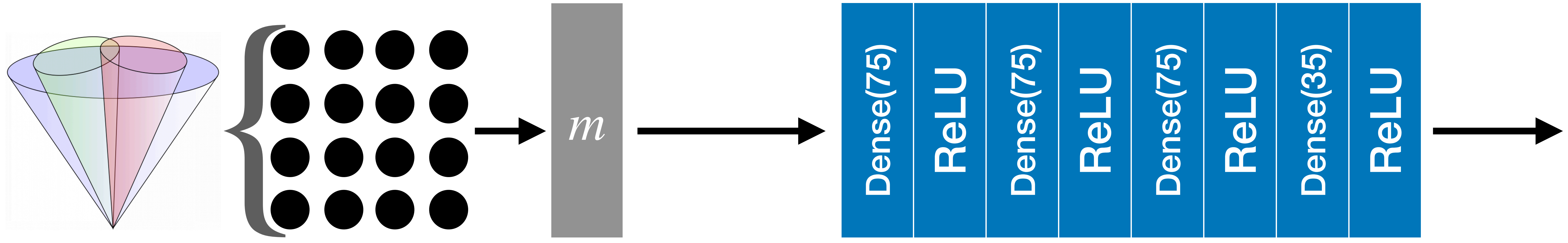
arXiv:1908.05318



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

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

arXiv:1908.05318



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

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

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

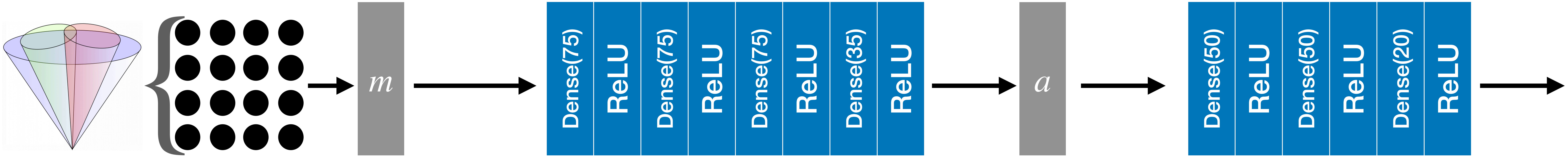
Cumulative effects of interaction received by a given vertex.

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

Concatenate to input.

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

arXiv:1908.05318



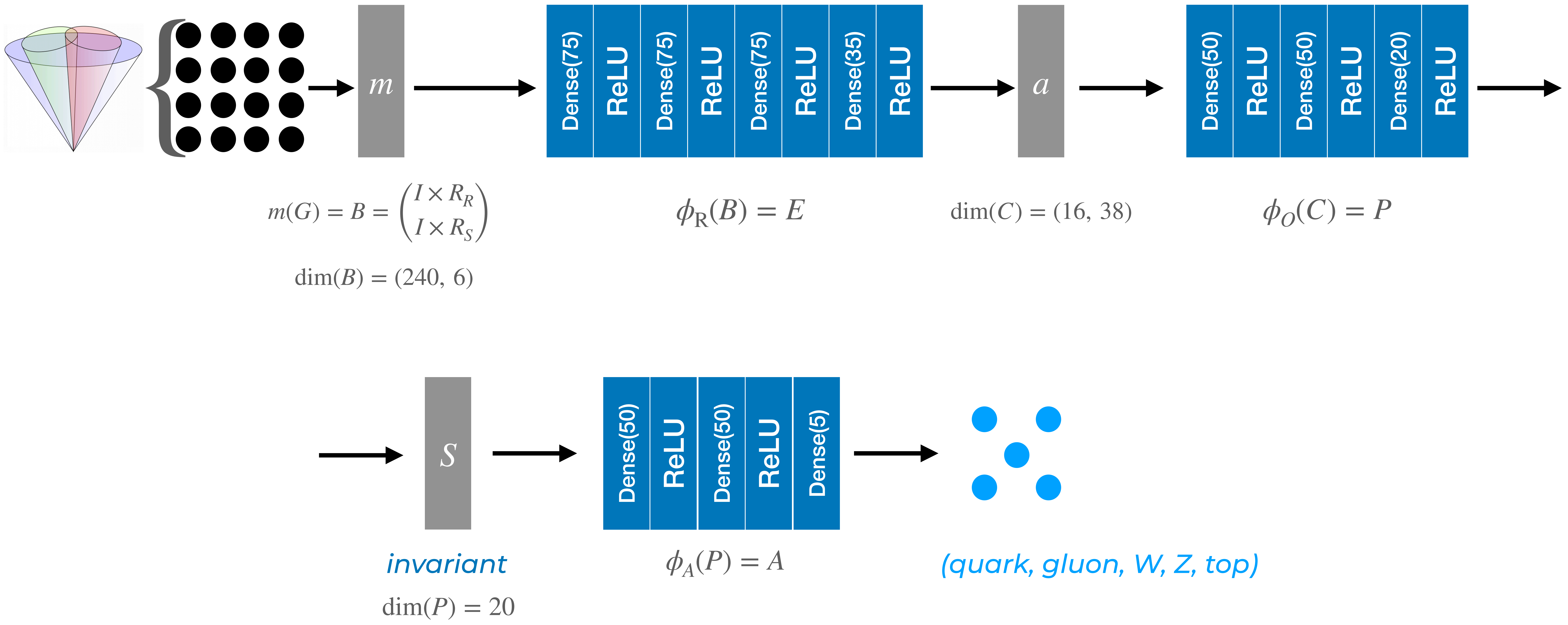
$$m(G) = B = \begin{pmatrix} I \times R_R \\ I \times R_S \end{pmatrix}$$

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

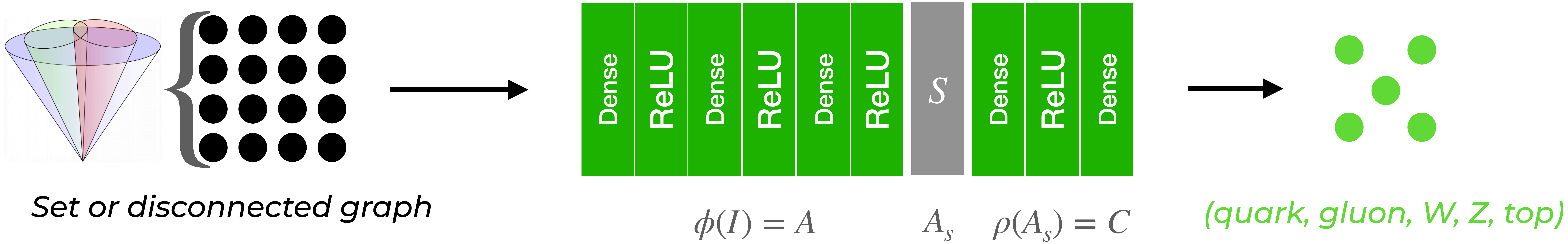
$$\phi_{\mathrm{R}}(B) = E$$

$$\dim(C) = (16,\ 38)$$

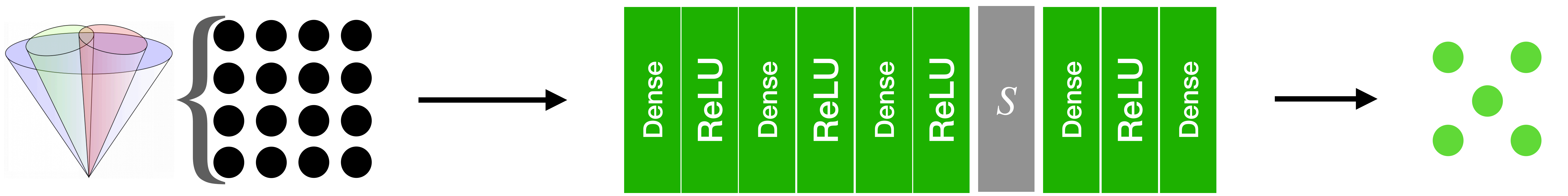
$$\phi_O(C) = P$$



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)

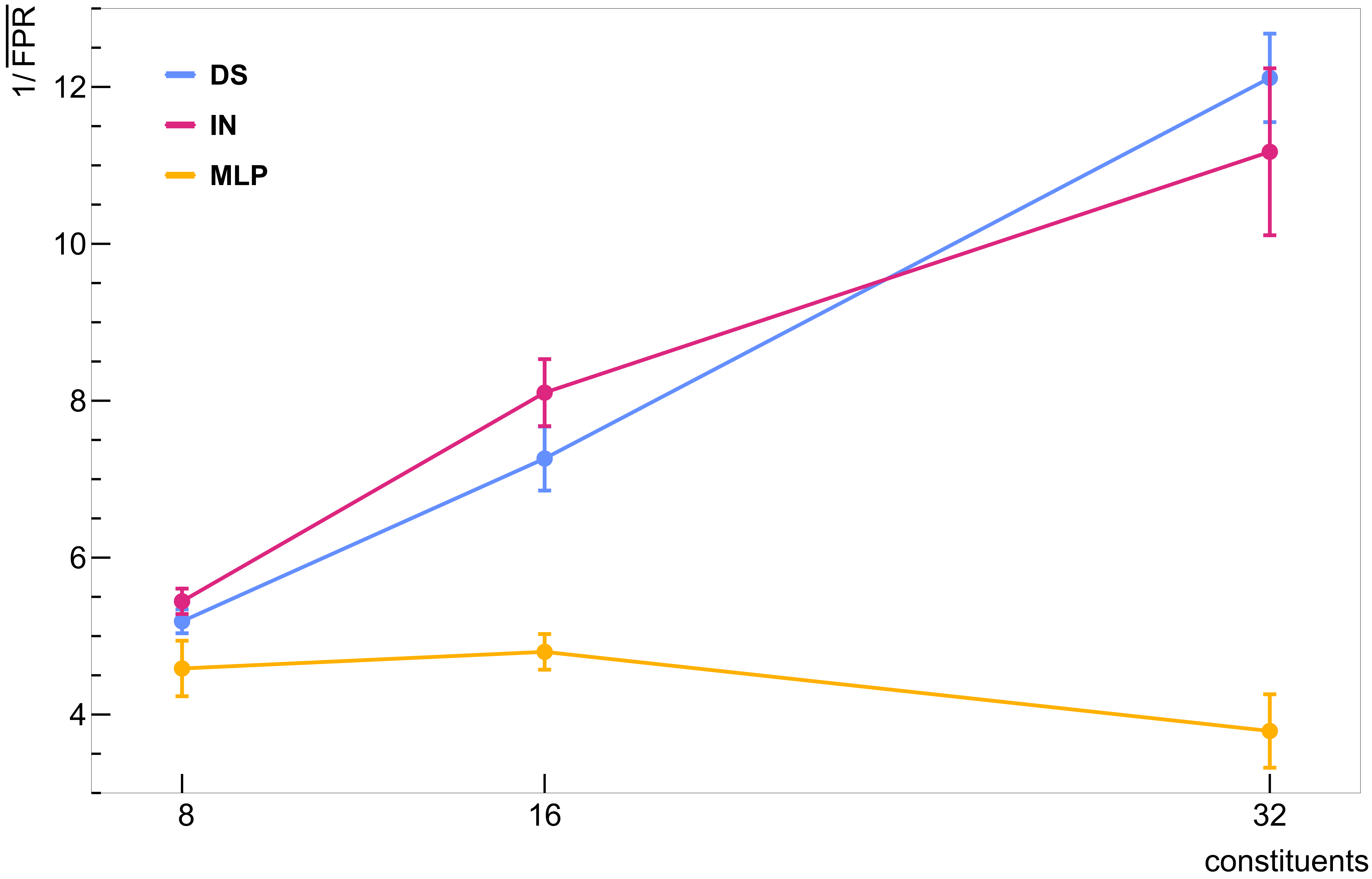
$$f : 2^{\mathfrak{X}} \to \mathcal{Y}$$

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

invariant

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

arXiv:1703.06114


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

Xilinx VU13P

Latencies **(ns)**

**8 bit quantised*

n particles	MLP	DS	IN
8	55	95	185
16	55	115	200
32	55	130	235

Xilinx VU13P

Resources **(LUTs)**

**8 bit quantised*

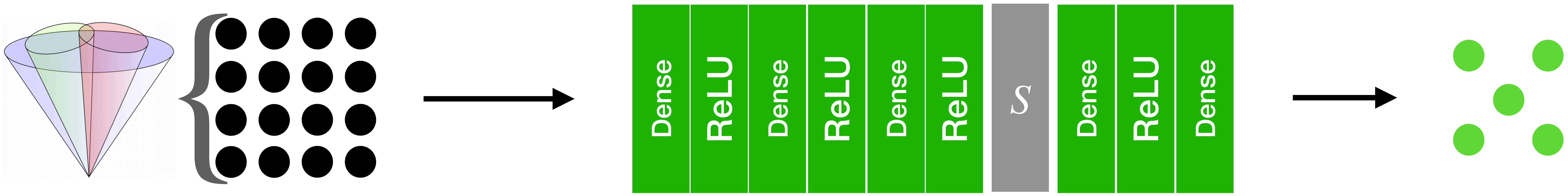
n particles	MLP	DS	IN
8	7%	31%	15%
16	7%	63%	36%
32	7%	76%	75%

- **Set representation ~ fully connected graph for this type of data!**
  - better representation?
  - other symmetries beyond permutation invariance to be exploited?

- **Incorporate symmetries bottom-up**
  - understand what symmetries are present in the data
  - test each separately

- **Geometrical models: smart, but can also be really FAST!**

Backup



Set or disconnected graph

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

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

(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

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

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

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

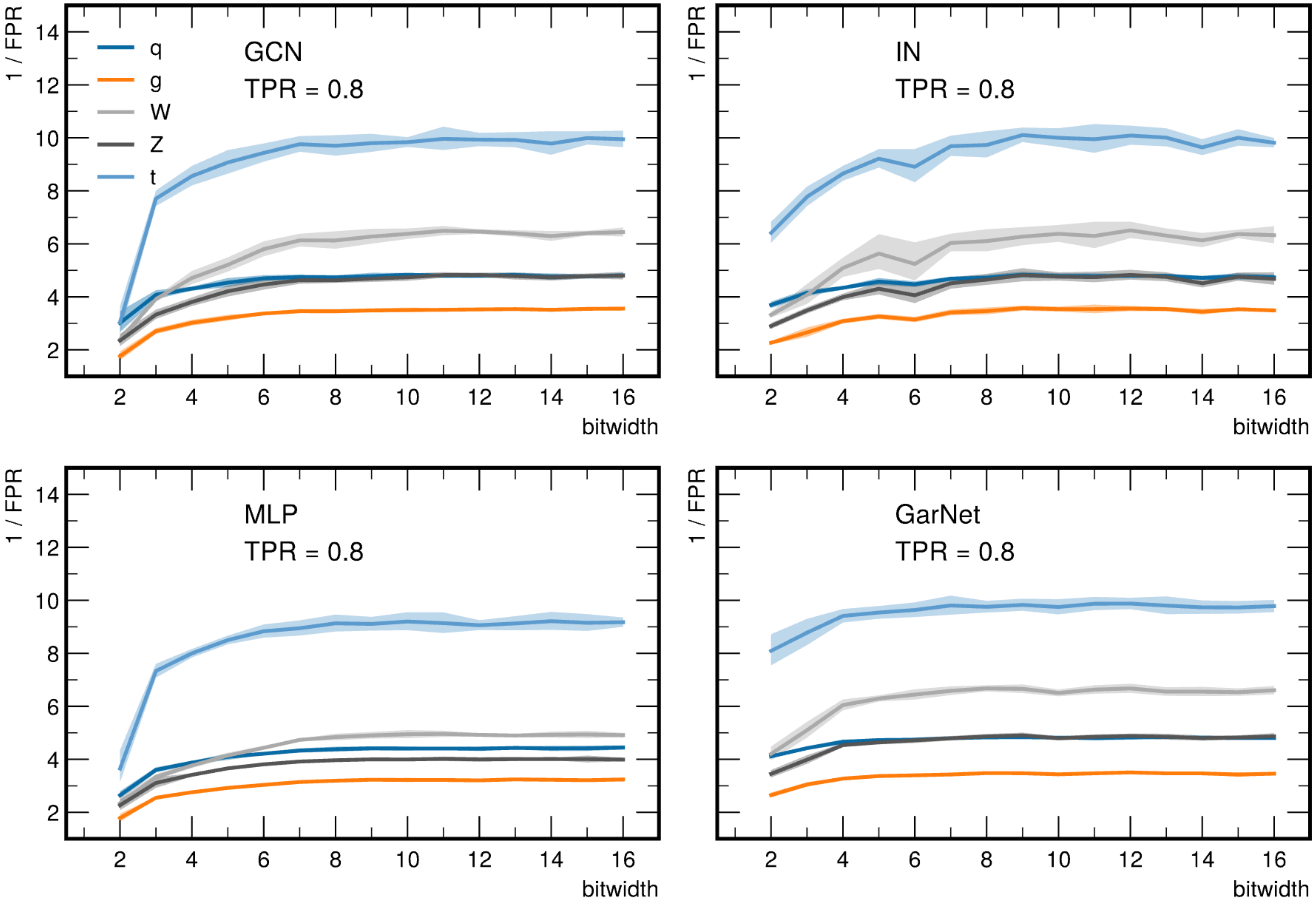
3. $a(G, X, E) \equiv C$

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.


GCN
TPR = 0.8
IN
TPR = 0.8
MLP
TPR = 0.8
GarNet
TPR = 0.8
q
g
W
Z
t
1 / FPR
bitwidth
2
4
6
8
10
12
14
16

Xilinx VU13P

Resources **(DSPs)**

**8 bit quantised*

n particles	**MLP**	**DS**	**IN**
8	2.4%	5.7%	6.4%
16	2.5%	4.5%	3.8%
32	2.4%	3.5%	7.3%

Xilinx VU13P

Resources **(FFs)**

**8 bit quantised*

n particles	MLP	DS	IN
8	0.6%	3.5%	6.7%
16	0.7%	6.9%	10.4%
32	0.7%	10.4%	32.7%

BRAMs are all 0.1% - softmax