Energy-weighted Message-passing Networks
-- An IRC safe prescription for Jets

Anomalies 2021, IITH

Vishal Ngairangbam
Physical Research Laboratory, Ahmedabad

Based on: arxiv: 2109.14636
(Partha Konar, VN, Michael Spannowsky)

Outline

- Jet-substructure & IRC safe observables
- Graphs and message-passing neural networks(MPNNs)
- Energy-Weighted Message-passing
    - Building an IRC safe graph
    - IRC safe message-passing
- Results and discussions

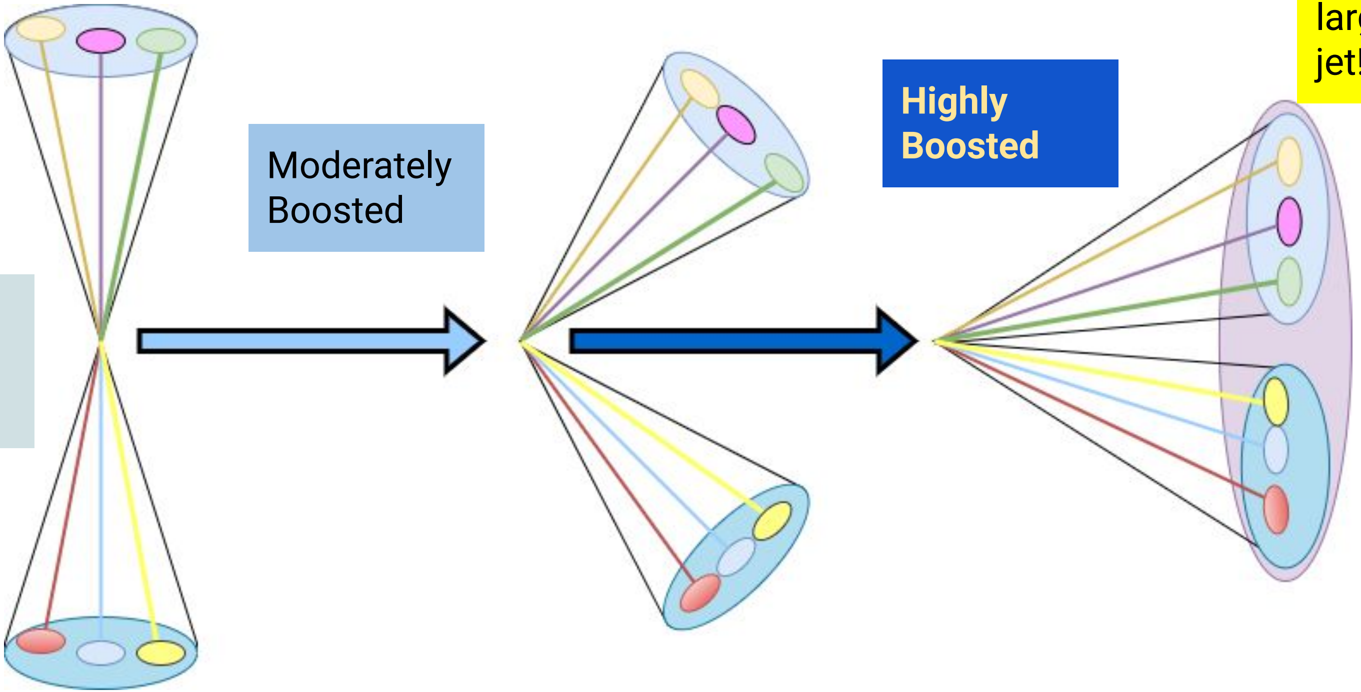
Jet Substructure at LHC

Decay Products captured in a large-radius(fat) jet!!

Two-body decay at Rest

Moderately Boosted

Highly Boosted



EW scale particle : decaying to quarks (dijet like signature) ⇒ Merged large-radius-jet in boosted regime

Jet Substructure at LHC

z_i is relative hardness

for hadronic colliders: $z_i = \frac{p_T^i}{\sum_{j=1}^{n} p_T^j}$

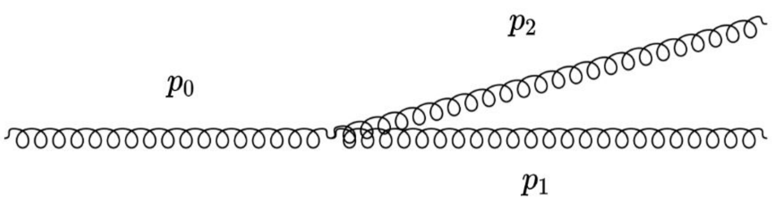
$p_0 = (z_0, \hat{p}_0)$

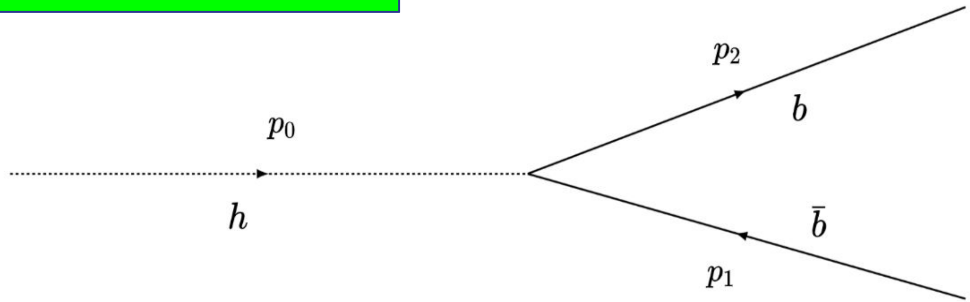
$p_1 = (z_1, \hat{p}_1)$

$p_2 = (z_2, \hat{p}_2)$

$p_0 = p_1 + p_2$

First splitting/decay at Parton-level





Dominantly Soft or collinear

$(z_2 << z_1 \text{ or } \theta_{12} \to 0)$

Democratic splitting

$(z_2 \sim z_1)$

QCD jets(Background)

Higgs jets (2-prong Signal)

Infra-red and Collinear (IRC) Safe observables

For an observable $\mathcal{O}_n$ defined on n particles.

$$\mathcal{O}_{n+1}(p_a,..,p_b,p_r,p_s,p_c,.) \rightarrow \mathcal{O}_n(p_a,..,p_b,p_q,p_c,.)$$

In the infra-red $(z_r \rightarrow 0 \text{ or } z_s \rightarrow 0)$ or collinear limits $(\Delta_{rs} \rightarrow 0)$

Calculable in pQCD!!

For a splitting: $q \rightarrow r + s$

$$p_q = p_r + p_s$$

$$p_q = (z_q, \hat{p_q})$$

$$p_r = (z_r, \hat{p_r})$$

$$p_s = (z_s, \hat{p_s})$$

Infra-red and Collinear (IRC) Safe observables

For an obser

$\mathcal{O}_{n+1}(p_a,$ … $, p_q, p_c, \cdot)$

In the infra-red (

How can we make the features learned by Message-passing neural networks IRC safe?

$p_q = (z_q, \hat{p}_q)$

For a splitting: $q \to r + s$

$p_r = (z_r, \hat{p}_r)$

$p_s = (z_s, \hat{p}_s)$

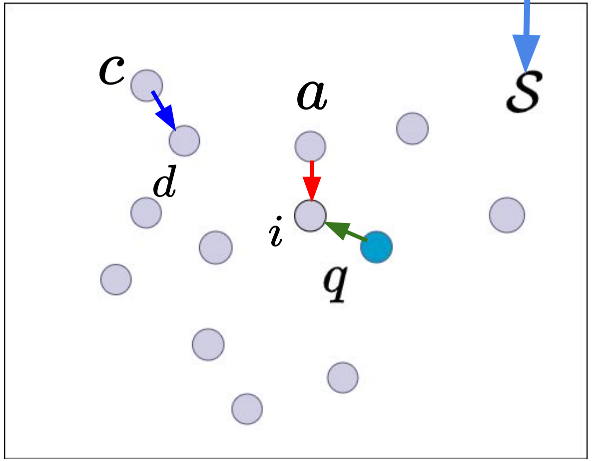
$$p_q = p_r + p_s$$

Calculable in pQCD!!

Graphs: Compact efficient data structures

$\mathcal{S} = \{a, b, i, q, ..\}$

Node Set: all particles within a jet



$\mathcal{E} = \{(i, a), (i, q), (d, c), \ldots\}$

Edge set

A graph $G(\mathcal{S}, \mathcal{E})$ defined on a set $\mathcal{S}$, with edge-set $\mathcal{E}$

Node-features: $\{\mathbf{h}_a, \mathbf{h}_b, \mathbf{h}_c, \ldots\}$

Four-momenta, charge, etc,

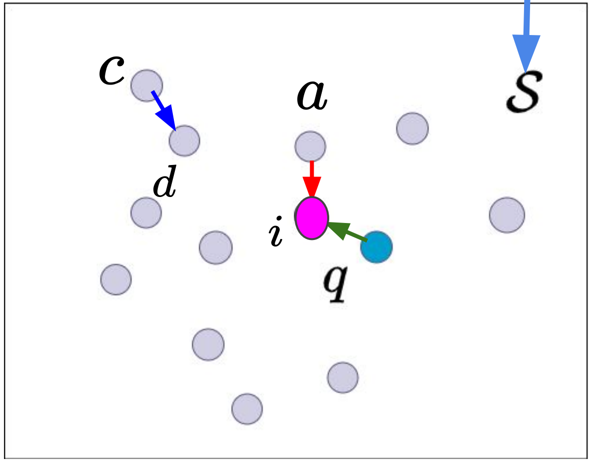
Edge-features: $\{\mathbf{e}_{ia}, \mathbf{e}_{iq}, \mathbf{e}_{dc}, \ldots\}$

$m_{ia}, \Delta R_{ia} etc,$

Message-passing neural networks(Graph Neural Networks)

$\mathcal{S} = \{a, b, i, q, ..\}$

Node Set: all particles within a jet



$\mathcal{E} = \{(i, a), (i, q), (d, c), \ldots\}$

Edge set

Message passing operation

Message-passing (each edge) $\Phi(\mathbf{h}_i, \mathbf{h}_q)$

Node readout (each node) $\mathbf{H}_i = \sum_{j \in \mathcal{N}[i]} \Phi(\mathbf{h}_i, \mathbf{h}_j)$

$\mathcal{N}[i] =$ Set of all nodes with incoming connections to i

Graph-readout (full graph)

$$\mathbf{g} = \sum_{j \in \mathcal{S}} \mathbf{H}_j$$

To Dense network

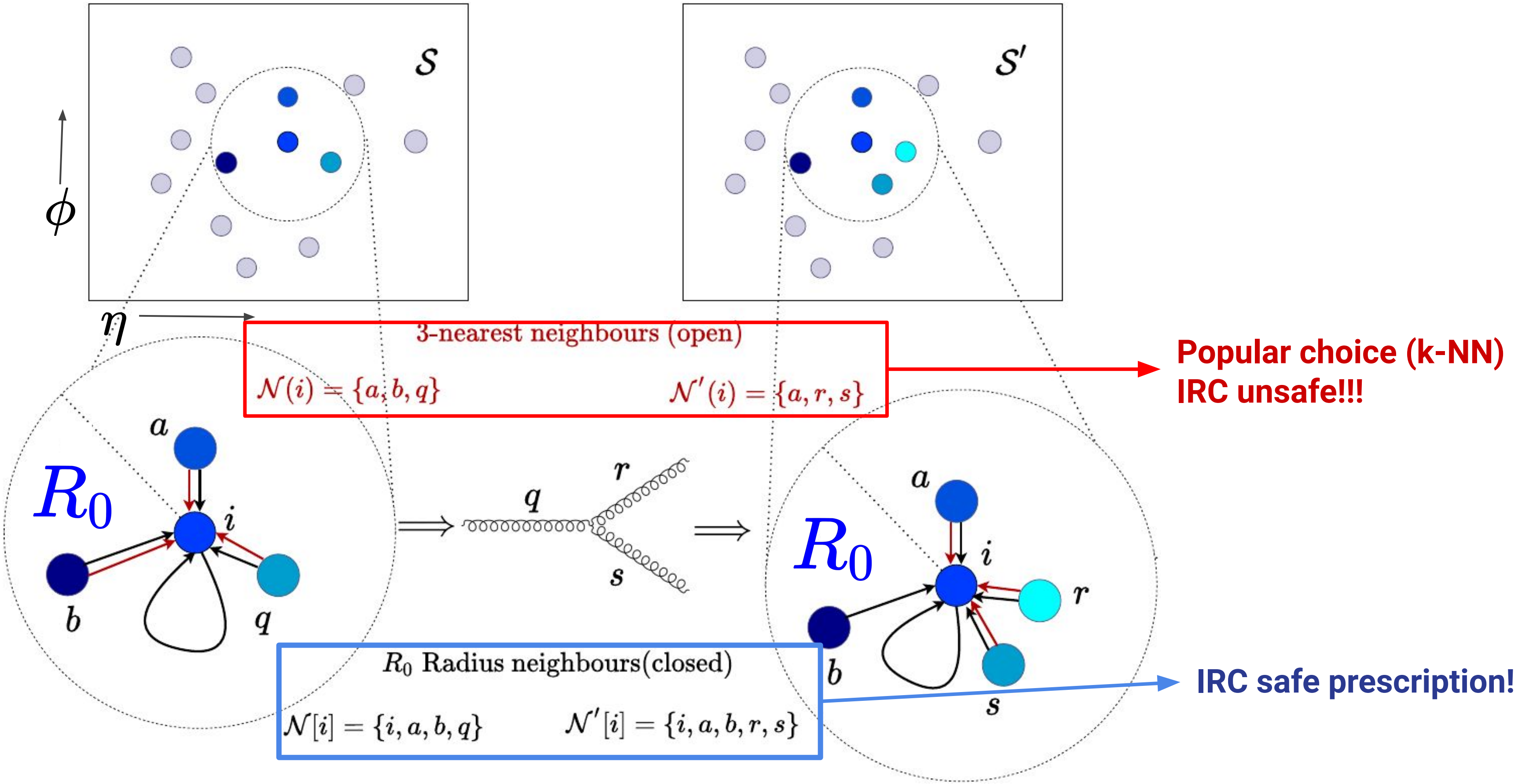
Deep-sets vs Message-passing neural networks(MPNN)

Deep-sets per-particle map: $\Phi(p_i)$	**MPNNs** Message-function: $\Phi(p_i, p_j)$
Cannot extract inter-particle correlations	Can extract **inter-particle correlations**
Only single particle information	**Graph construction algorithm** controls information extraction at first layer(via node-readout)
Iterative application has no additional complexity on feature extraction, except functional composition $\Phi'(\Phi(p_i))$	Gradual increase in information in node-features, **after each iteration**
No such control	**Number of iterations** control the scope of information contained in the final node-feature

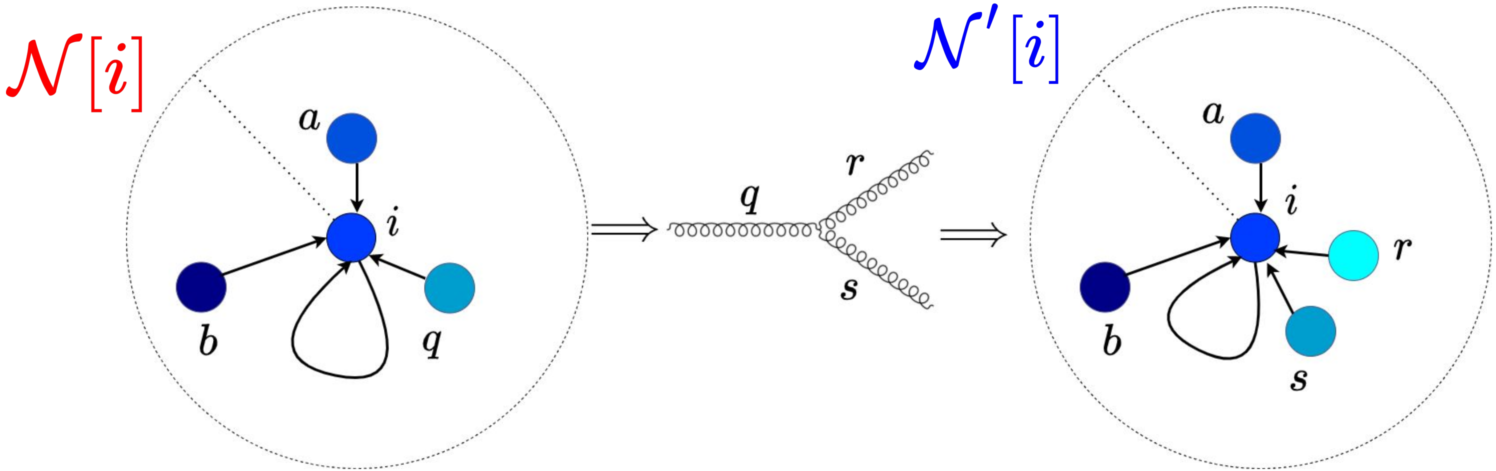
Energy Flow Networks(EFNs): IRC safe deep-sets framework

JHEP 01 (2019) 121, Komiske, Metodiev, Thaler

IRC safe jet-graphs



Energy-weighted Message-passing(EMPN)

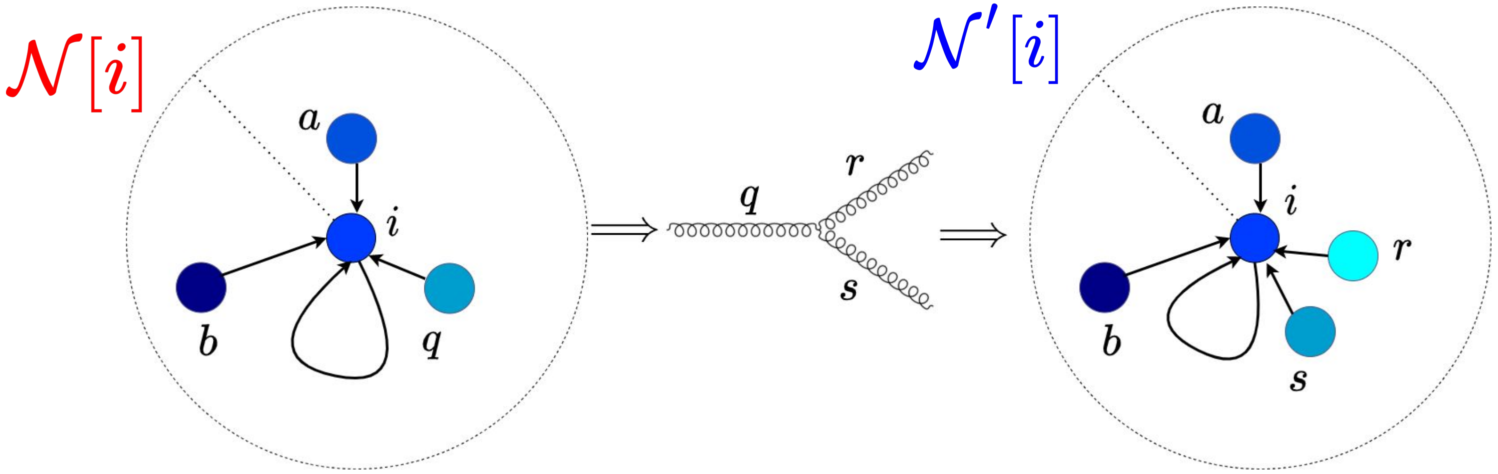


Message-passing operation

$$\mathbf{H}_i = \sum_{j \in \mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \hat{\Phi}(\hat{p}_i, \hat{p}_j)$$

$$\omega_j^{(\mathcal{N}[i])} = \frac{p_T^j}{\sum_{k \in \mathcal{N}[i]} p_T^k}$$

Energy-weighted Message-passing(EMPN)



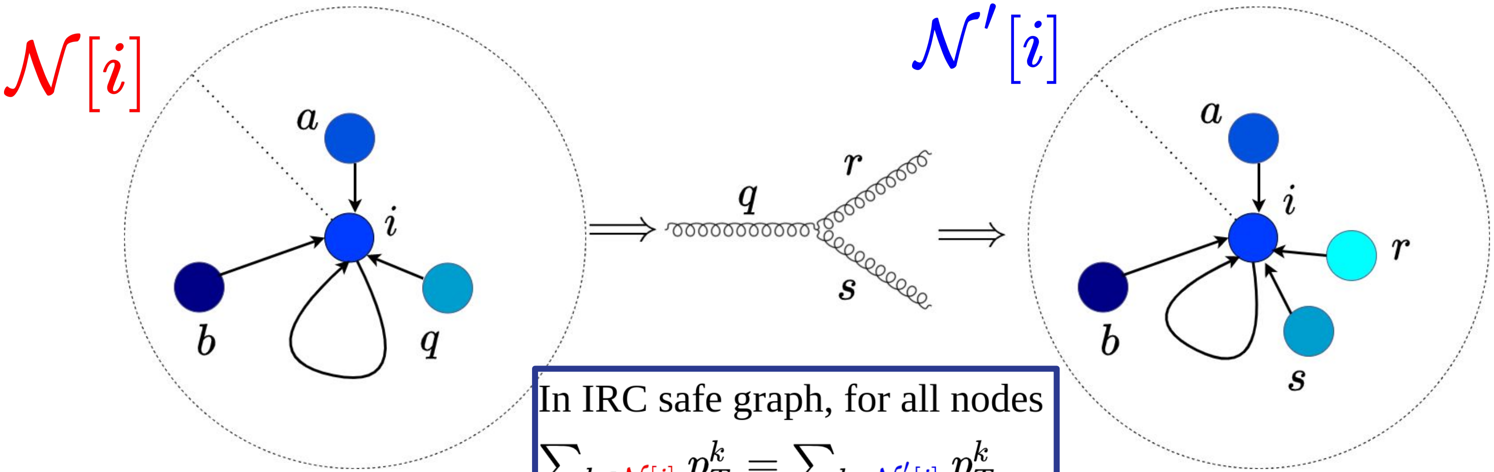
Message-passing operation

$$\mathbf{H}_i = \sum_{j \in \mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \hat{\Phi}(\hat{p}_i, \hat{p}_j)$$

$$\omega_j^{(\mathcal{N}[i])} = \frac{p_T^j}{\sum_{k \in \mathcal{N}[i]} p_T^k}$$

IR Safety: $z_j \to 0 \Rightarrow \omega_j^{(\mathcal{K})} \to 0$ for any $\mathcal{K}$

Energy-weighted Message-passing(EMPN)



In IRC safe graph, for all nodes

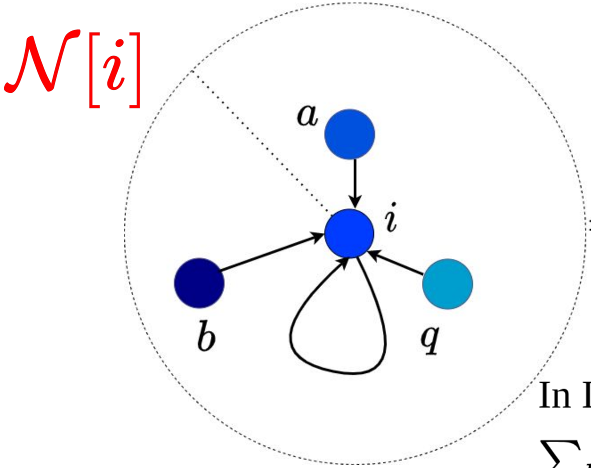
$$\sum_{k\in\mathcal{N}[i]} p_T^k = \sum_{k\in\mathcal{N}'[i]} p_T^k$$

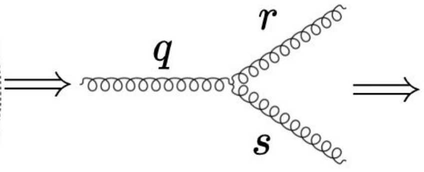
Message-passing operation

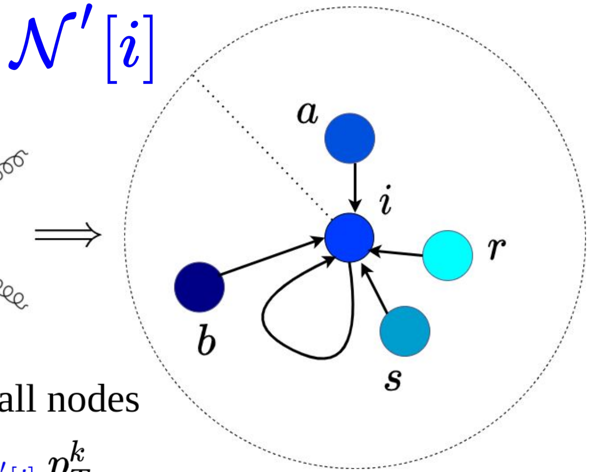
$$\mathbf{H}_i = \sum_{j\in\mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \hat{\Phi}(\hat{p}_i, \hat{p}_j)$$

$$\omega_j^{(\mathcal{N}[i])} = \frac{p_T^j}{\sum_{k\in\mathcal{N}[i]} p_T^k}$$

Energy-weighted Message-passing(EMPN)







In IRC safe graph, for all nodes

$$\sum_{k\in\mathcal{N}[i]} p_T^k = \sum_{k\in\mathcal{N}'[i]} p_T^k$$

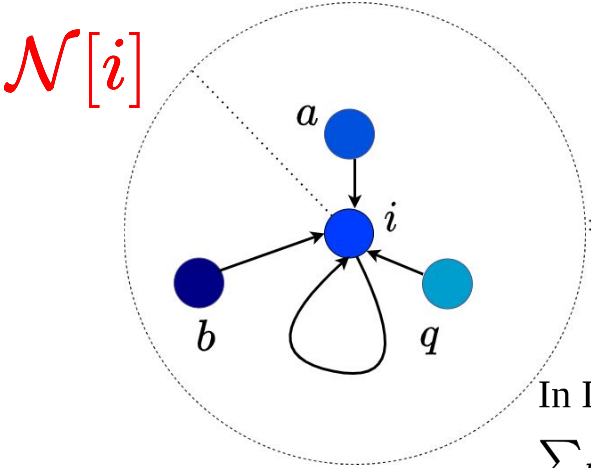
$$p_T^q = p_T^r + p_T^s \Rightarrow \omega_q^{(\mathcal{N}[i])} = \omega_r^{(\mathcal{N}'[i])} + \omega_s^{(\mathcal{N}'[i])}$$

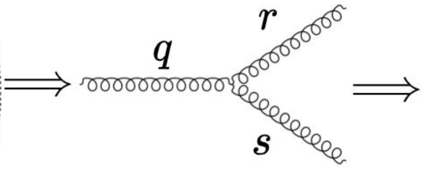
Message-passing operation

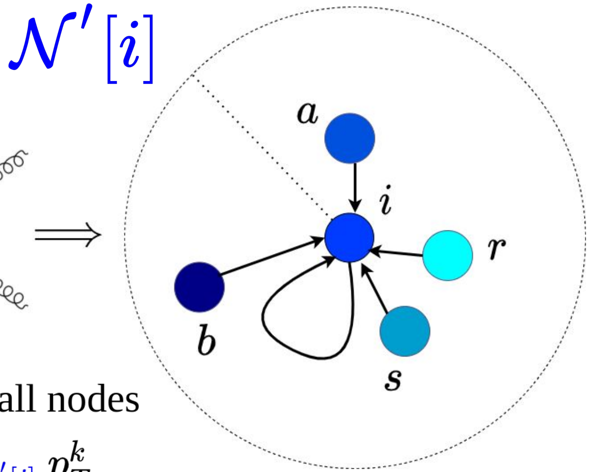
$$\mathbf{H}_i = \sum_{j\in\mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \hat{\Phi}(\hat{p}_i, \hat{p}_j)$$

$$\omega_j^{(\mathcal{N}[i])} = \frac{p_T^j}{\sum_{k\in\mathcal{N}[i]} p_T^k}$$

Energy-weighted Message-passing(EMPN)







In IRC safe graph, for all nodes

$$\sum_{k\in\mathcal{N}[i]} p_T^k = \sum_{k\in\mathcal{N}'[i]} p_T^k$$

Message-passing operation

$$\mathbf{H}_i = \sum_{j\in\mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \, \hat{\Phi}(\hat{p}_i, \hat{p}_j)$$

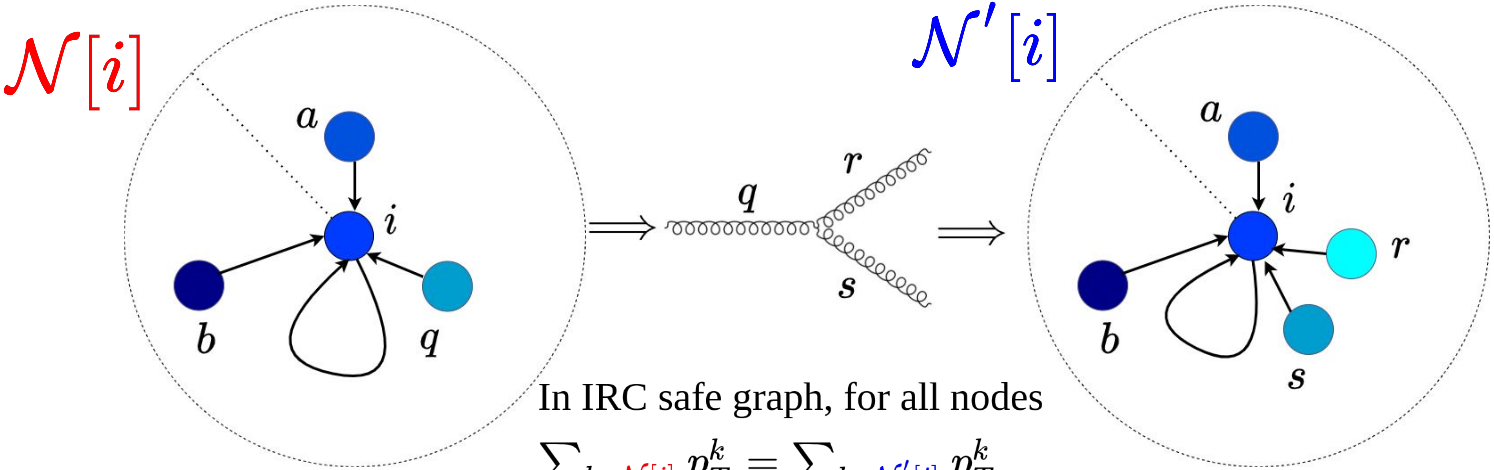
$$\omega_j^{(\mathcal{N}[i])} = \frac{p_T^j}{\sum_{k\in\mathcal{N}[i]} p_T^k}$$

$$p_T^q = p_T^r + p_T^s \Rightarrow \omega_q^{(\mathcal{N}[i])} = \omega_r^{(\mathcal{N}'[i])} + \omega_s^{(\mathcal{N}'[i])}$$

C Safety: $\hat{p}_q = \hat{p}_r = \hat{p}_s$

$$\omega^{(\mathcal{N}[i])} \, \hat{\Phi}(\hat{p}_i, \hat{p_q}) = \omega^{(\mathcal{N}'[i])} \, \hat{\Phi}(\hat{p}_i, \hat{p}_r) + \omega^{(\mathcal{N}[i])} \, \hat{\Phi}(\hat{p}_i, \hat{p}_s)$$

Energy-weighted Message-passing(EMPN)



In IRC safe graph, for all nodes

$$\sum_{k\in\mathcal{N}[i]} p_T^k = \sum_{k\in\mathcal{N}'[i]} p_T^k$$

Message-passi

$\mathbf{H}_i = \sum_{j\in\mathcal{N}$ … $\mathcal{N}'[i]) + \omega_s^{(\mathcal{N}'[i])}$

… $\omega^{(\mathcal{N}[i])}\ \hat{\Phi}(\hat{p}_i, \hat{p}_s)$

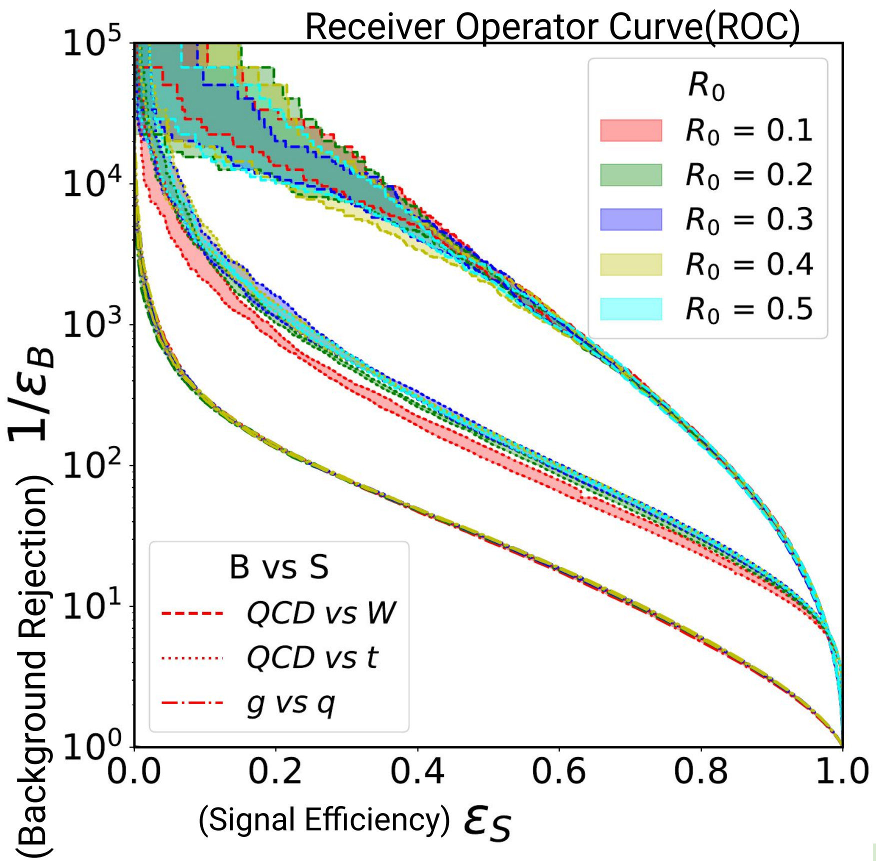
$\sum_{k\in\mathcal{N}[i]} p_T$ …

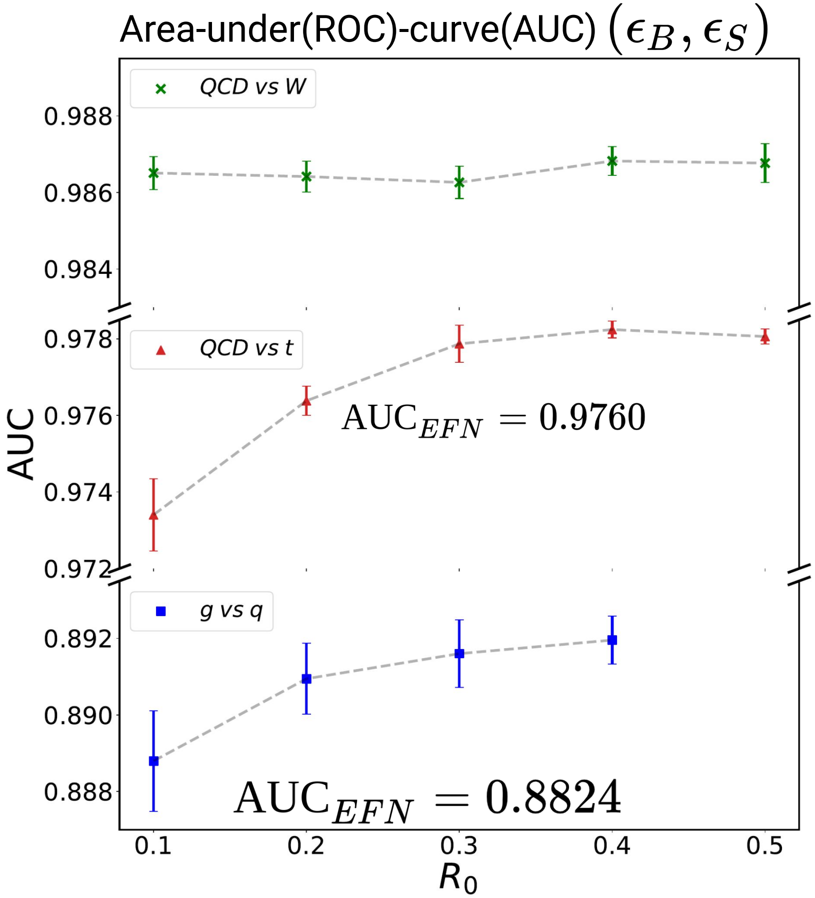
IRC safe updated node-features:

$$\mathbf{H}_i = \mathbf{H}'_i$$

Results

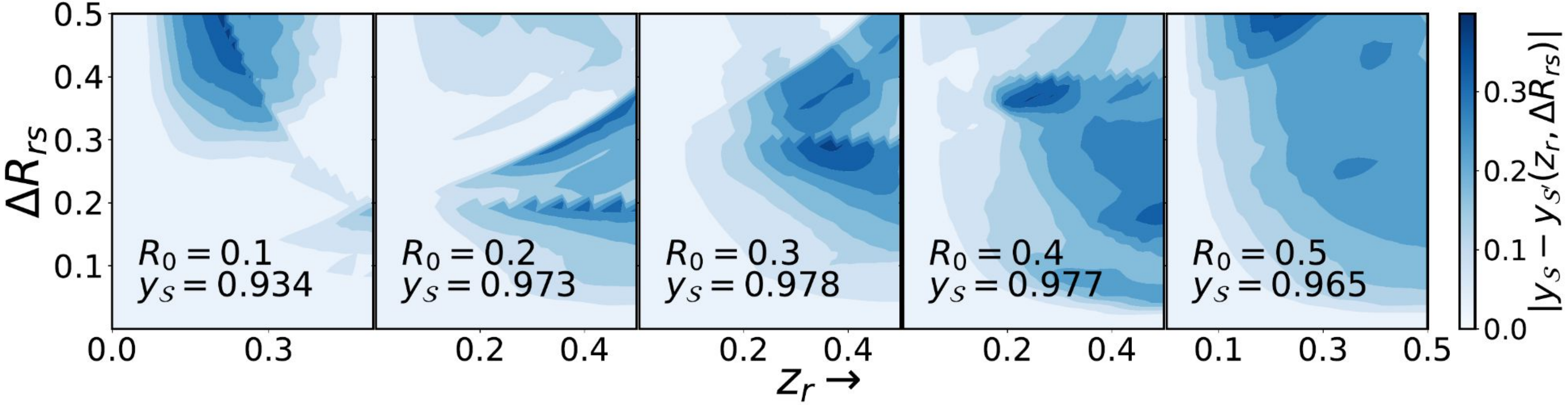
Network Performance





Higher AUC = better discrimination

Examining IRC Safety



Split the hardest constituent in a jet and vary z_r and ΔR_{rs}

Network Output: $y_S \quad y_{S'}(z_r, \Delta R_{rs})$

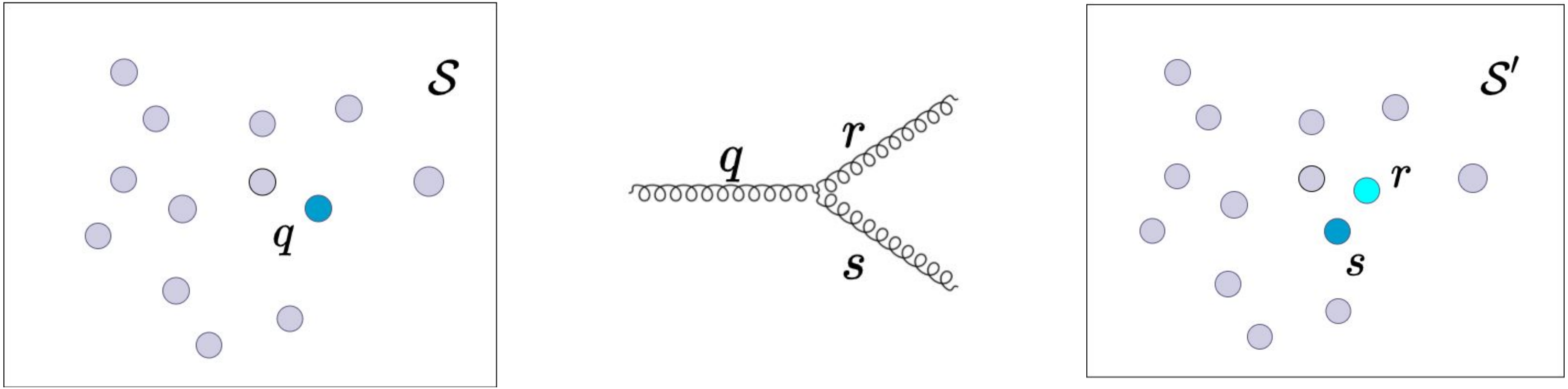
Increasing R_0 **decreases stability** of network output to **additional emissions**

Conclusions

- Generalised Energy Flow Networks to extract local correlations via message-passing operations
- Single Energy-weighted message passing improves upon EFNs
- Iterative application does not spoil IRC safety, further room for improvement
- Devised generic graph construction algorithms which give invariant graph structure in the ***deletion*** of a soft or collinear vertex
- Possibility to structure graphs and networks with highly intuitive physics input
- **General enough to study inclusive event shapes**
- Can we understand the extracted features within pQCD?

Back-up

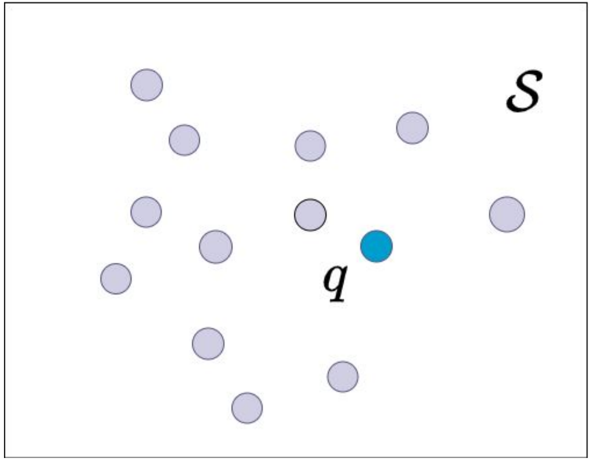
EMPN: Iterative application

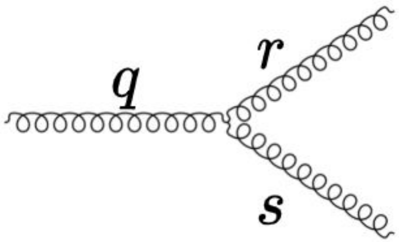


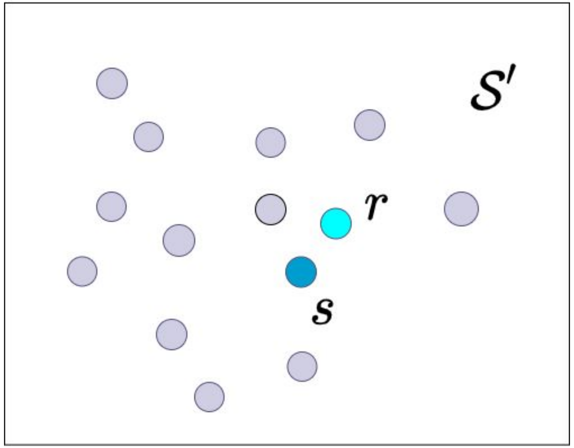
C limit:

$$\mathbf{h}_q^{(1)} = \mathbf{h}_r^{(1)} = \mathbf{h}_s^{(1)}$$

EMPN: Iterative application







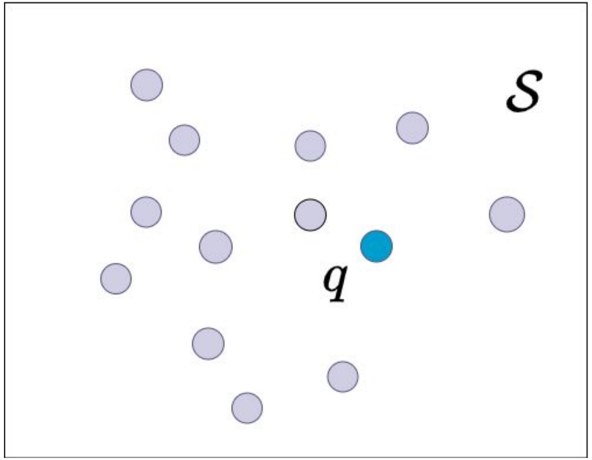
C limit:

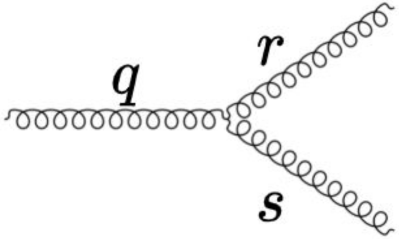
$$\mathbf{h}_q^{(1)} = \mathbf{h}_r^{(1)} = \mathbf{h}_s^{(1)}$$

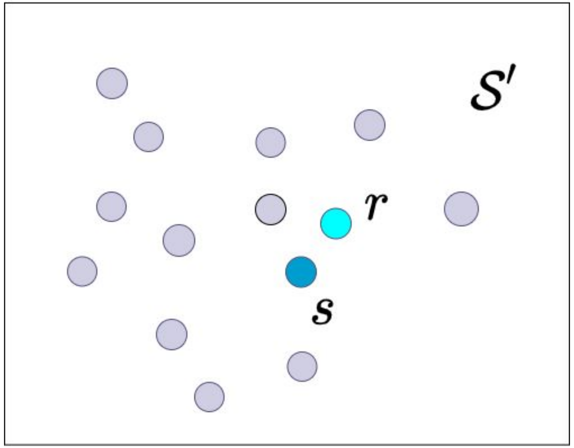
$$\mathbf{h}_i^{(2)} = \sum_{j \in \mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])} \hat{\mathbf{\Phi}}^{(1)}(\mathbf{h}_i^{(1)}, \mathbf{h}_j^{(1)})$$

IRC Safe!!!

EMPN: Iterative application







Iterative Application is IRC safe!

C limit:

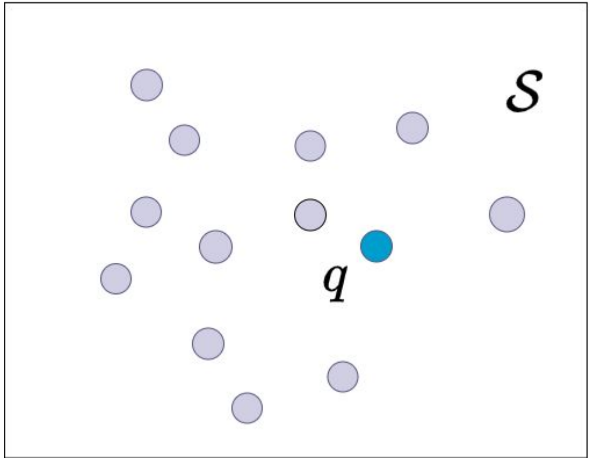
$$\mathbf{h}_q^{(l)} = \mathbf{h}_r^{(l)} = \mathbf{h}_s^{(l)}$$

$$\mathbf{h}_i^{(l+1)} = \sum_{j\in\mathcal{N}[i]} \omega_j^{(\mathcal{N}[i])}\, \hat{\Phi}^{(l)}\left(\mathbf{h}_i^{(l)}, \mathbf{h}_j^{(l)}\right)$$

IRC Safe!!!

$$\mathbf{h}_i^{(0)} = \hat{p}_i$$

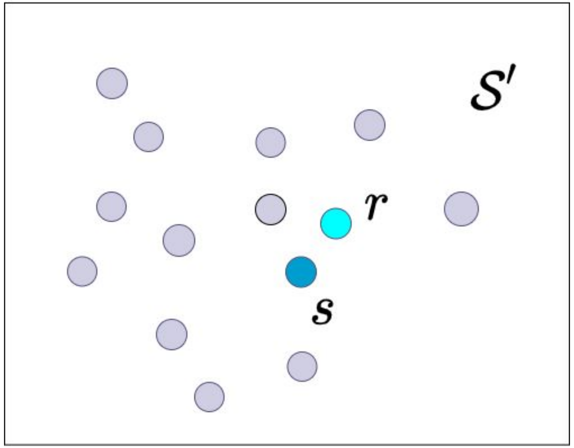
EMPN: Graph-readout



<u>Graph readout</u>

$$\mathbf{g} = \sum_{j \in \mathcal{S}} \omega_j^{(\mathcal{S})} \, \mathbf{h}_j^{(L)}$$

Representation of the full jet is IRC safe



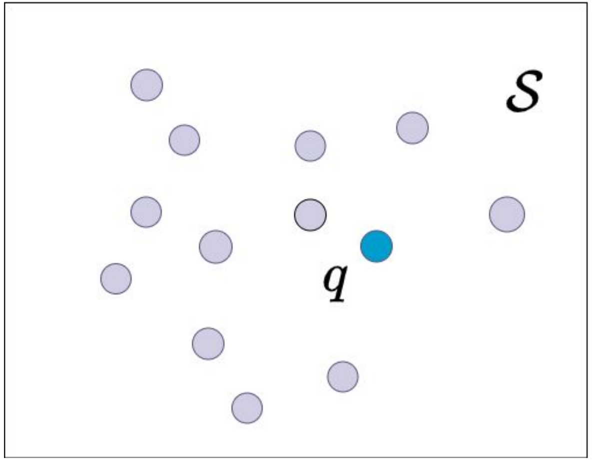
Graph on $\mathcal{S}'$ contain $\mathbf{h}_r^{(L)}$ and $\mathbf{h}_s^{(L)}$

C limit:

$$\mathbf{h}_q^{(L)} = \mathbf{h}_r^{(L)} = \mathbf{h}_s^{(L)}$$

$L =$ num. iterations

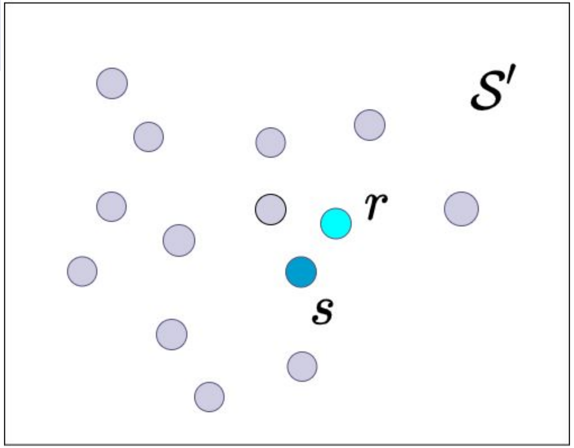
EMPN: Graph-readout



Information content determined by R_0 and L

Graph readout

$$\mathbf{g} = \sum_{j \in \mathcal{S}} \omega_j^{(\mathcal{S})} \, \mathbf{h}_j^{(L)}$$



Graph on $\mathcal{S}'$ contain $\mathbf{h}_r^{(L)}$ and $\mathbf{h}_s^{(L)}$

C limit:

$$\mathbf{h}_q^{(L)} = \mathbf{h}_r^{(L)} = \mathbf{h}_s^{(L)}$$

$L =$ num. iterations

Dataset Details

Sl. No	Jet Class	Parton-level	MPI	Detector Simulation	Jet Radius (anti-kT)	Transverse momentum [GeV]	Classification Scenario
1.	Gluon	Pythia8	Yes	No	0.4	[500,550]	Gluon vs Quark
2.	Quark	Pythia8	Yes	No	0.4	[500,550]	Gluon vs Quark
3.	QCD	Pythia8	No	Yes	0.8	[550,650]	QCD vs Top/W
4.	Top	Pythia8	No	Yes	0.8	[550,650]	QCD vs Top
5.	W	Madgraph5	No	Yes	0.8	[550,650]	QCD vs W

[1-2] Publicly available q/g dataset [Komiske *et.al*] (used in EFNs)

[3,4] Publicly available top tagging dataset [Kasieczka *et.al*] (used in EFNs)

[5] Generated with same specifications as [3,4]

Energy Flow Networks: A Special case of EMPNs

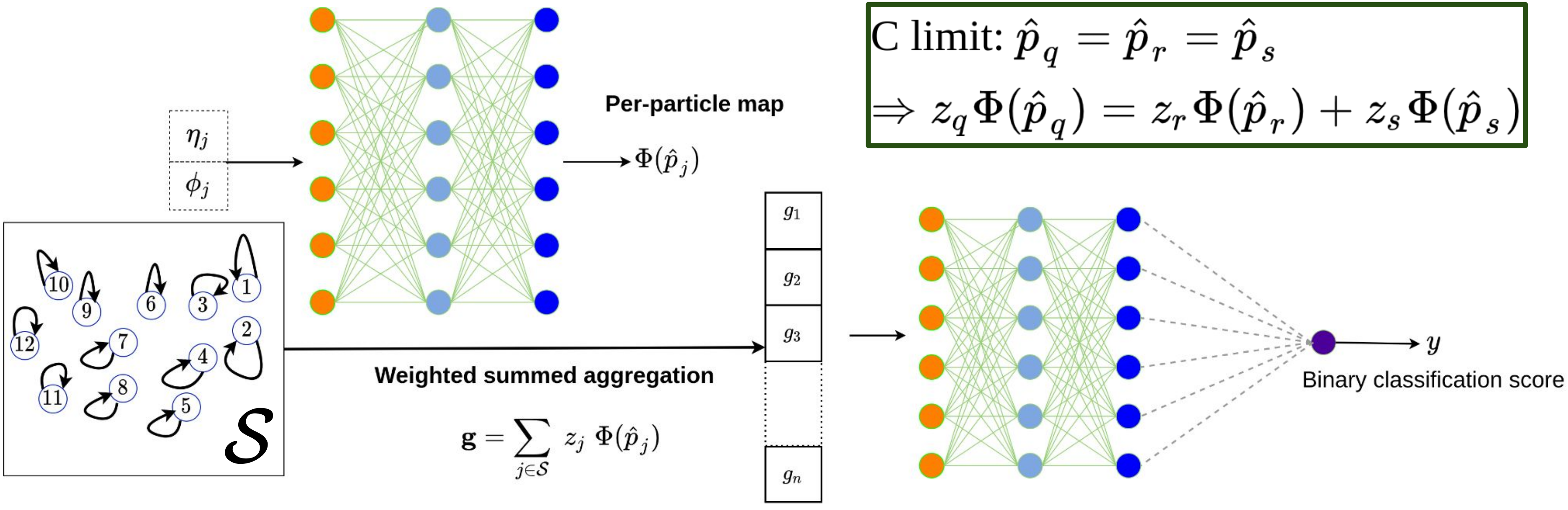
JHEP 01 (2019) 121, Komiske, Metodiev, Thaler

Per-particle map is a special message function constant for the second argument!

IR limit: $z_r = 0 \Rightarrow z_r\ \Phi(\hat{p}_r) = 0$

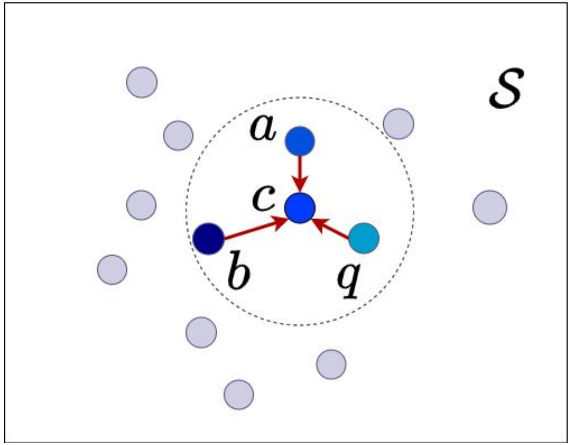
C limit: $\hat{p}_q = \hat{p}_r = \hat{p}_s$

$\Rightarrow z_q \Phi(\hat{p}_q) = z_r \Phi(\hat{p}_r) + z_s \Phi(\hat{p}_s)$

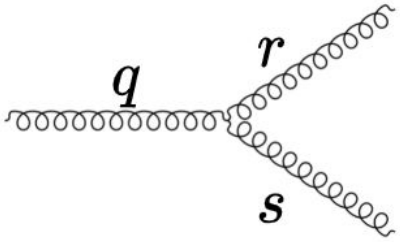


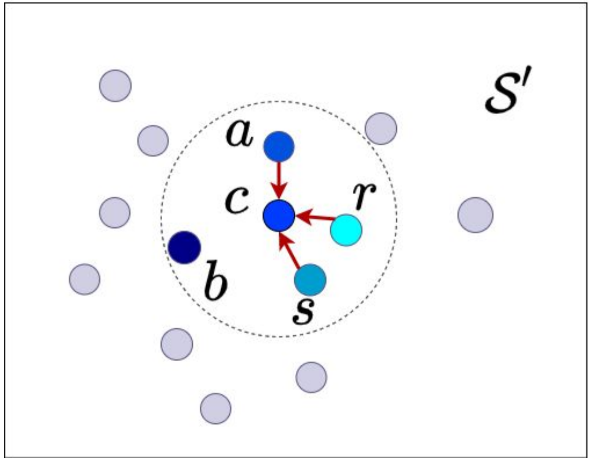
Jet-graph: k-nearest neighbour

(η, ϕ)-plane



k-NN graph structurally IRC unsafe!





$$\mathcal{N}(c) = \{a, b, q\}$$

$$\mathcal{N}'(c) = \{a, r, s\}$$

$$\mathbf{h}_i^{(l+1)} = \square_{j \in \mathcal{N}(i)}^{local} {}^{i}\mathbf{m}_j$$

$$\lim_{z_r \to 0} \mathbf{h}_c^{\prime(l+1)} \neq \mathbf{h}_c^{(l+1)}$$

Similar in the collinear limit