PELICAN

PERMUTATION- AND LORENTZ-EQUIVARIANT NETWORKS FOR PARTICLE PHYSICS

Alexander Bogatskiy, Timothy Hoffman,
Xiaoyang Liu, David W. Miller, Jan T. Offermann
ML4Jets: November 4th, 2023

THE UNIVERSITY OF CHICAGO DATA SCIENCE INSTITUTE
FLATIRON INSTITUTE
efi ENRICO FERMI INSTITUTE
Miller Lab

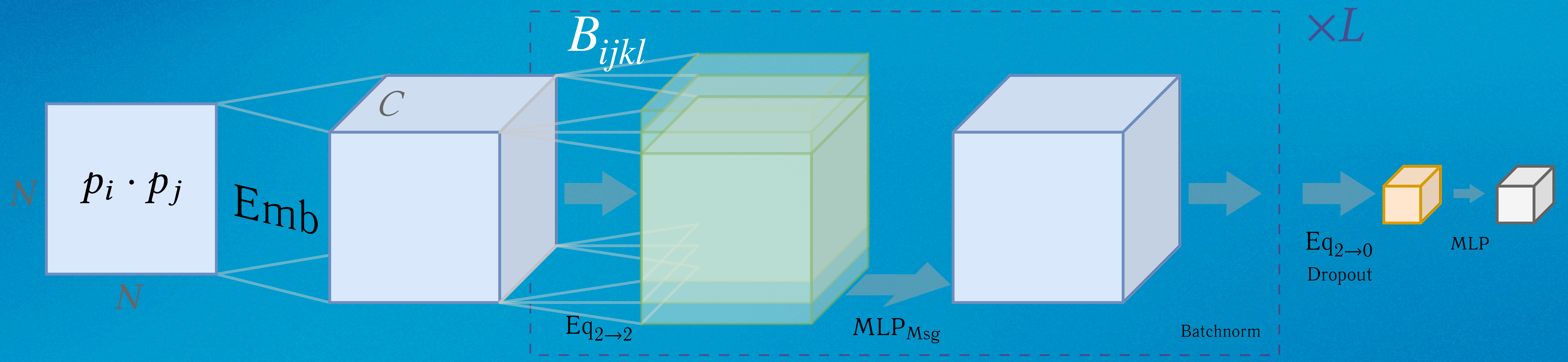
Overview

- Architecture Summary
- Tagging and Vector Reconstruction
- Small Parameter Limit

THE UNIVERSITY OF CHICAGO DATA SCIENCE INSTITUTE
FLATIRON INSTITUTE
ENRICO FERMI INSTITUTE
Miller Lab

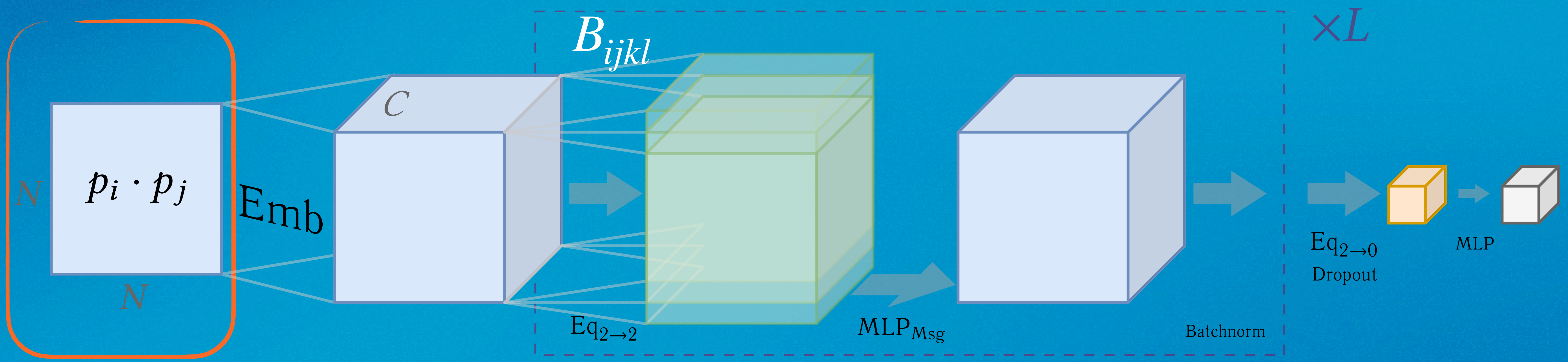
Symmetries as a Fundamental Orientation

- Lorentz and permutation symmetries are fundamentally important to jet physics
- Begin with a jet J, and its constituent four-vectors p_i
- Use the full set of natural Lorentz invariants $p_i \cdot p_j$ as building blocks
- Use particle-index permutation-equivariance tensors B_{ijkl} as fundamental network operations



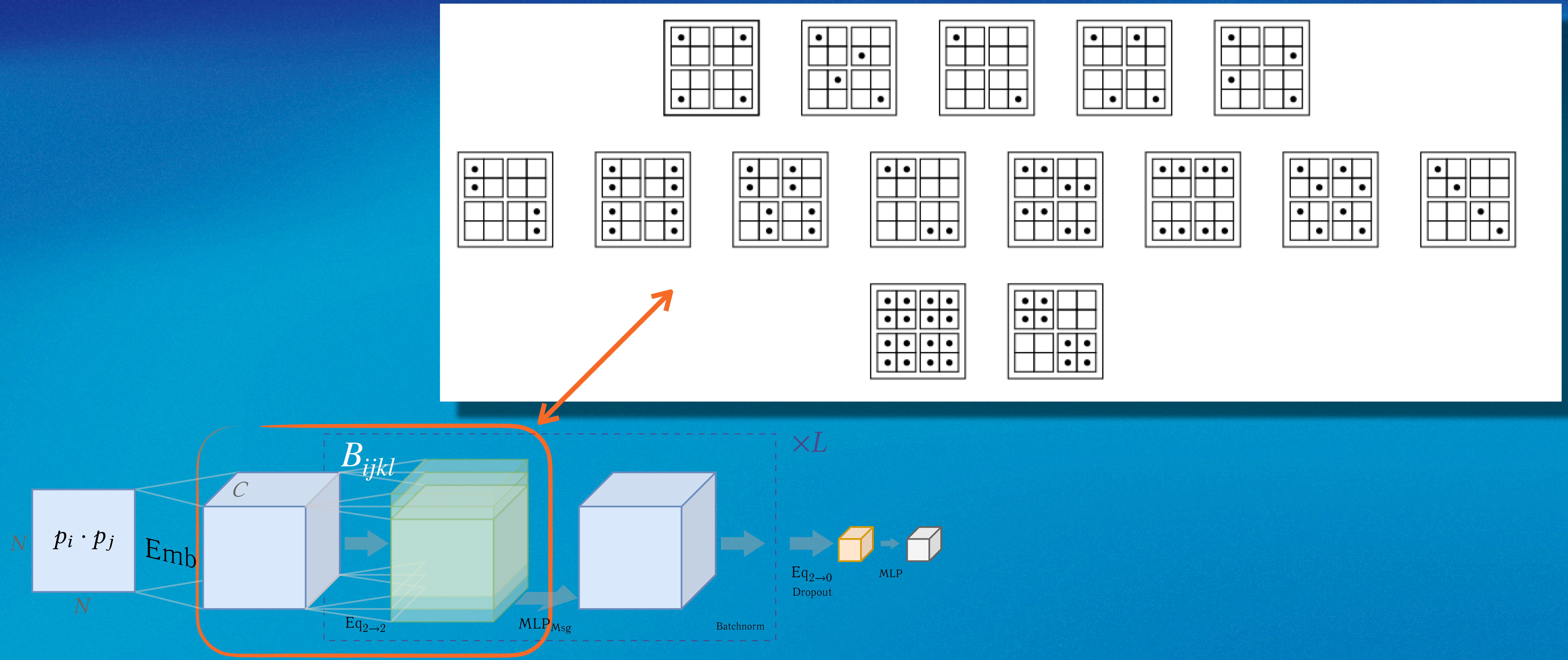
Lorentz Equivariance

- Construct Lorentz-scalars (particle ID) and vectors (intermediate particles)
- Target Lorentz-scalars: $f(\{p_i\}) = I(\{p_i \cdot p_j\})$
- Target Lorentz-vectors: $f^\mu(\{p_i\}) = \sum_k I_k(\{p_i \cdot p_j\}) p_k^\mu$



Permutation Equivariance

- 15 permutation equivariant linear maps of matrices into matrices
- 5 maps of matrices into vectors
- 2 maps of matrices into scalars

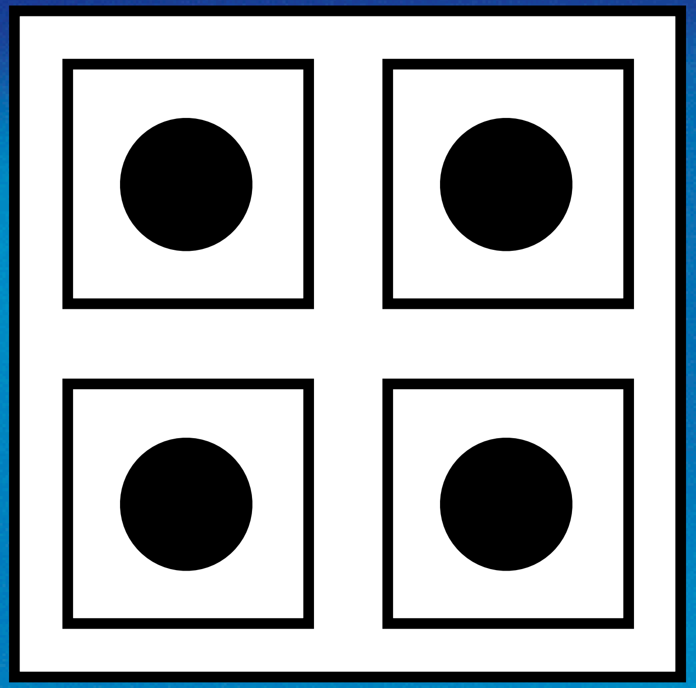


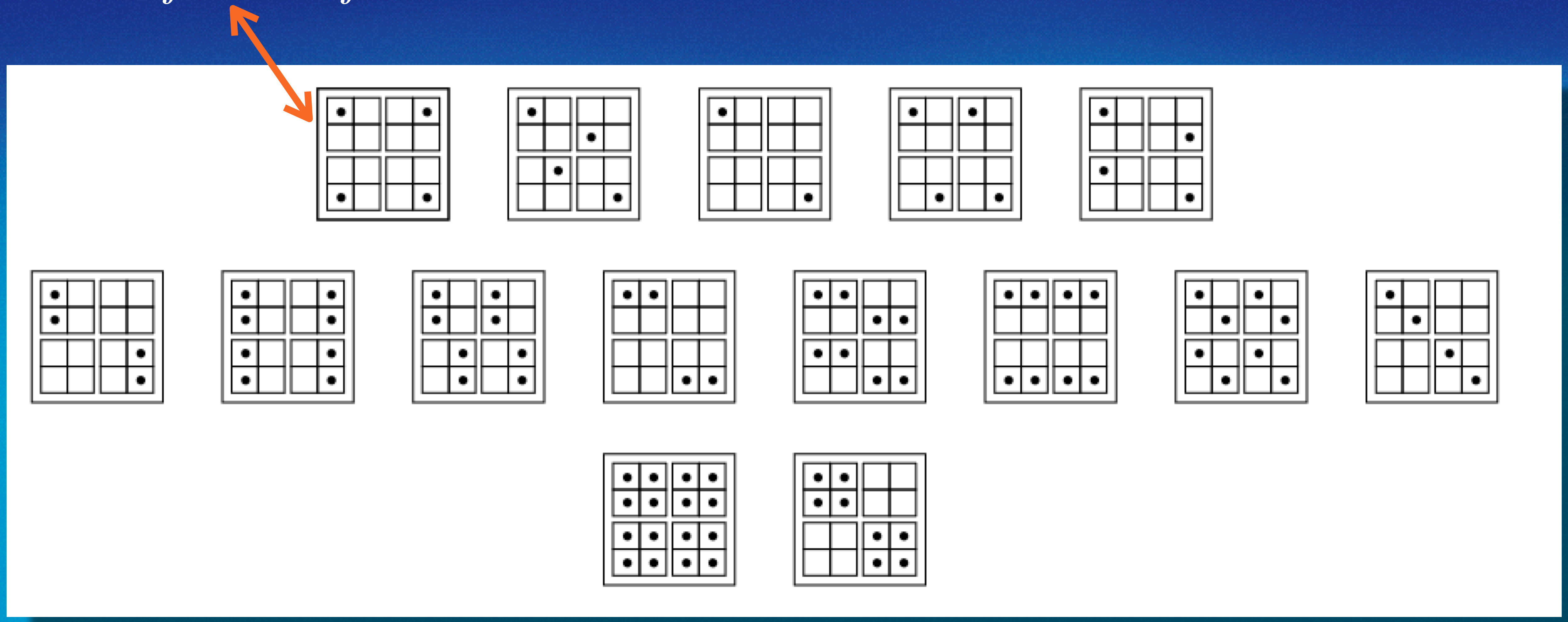
Permutation Equivariance

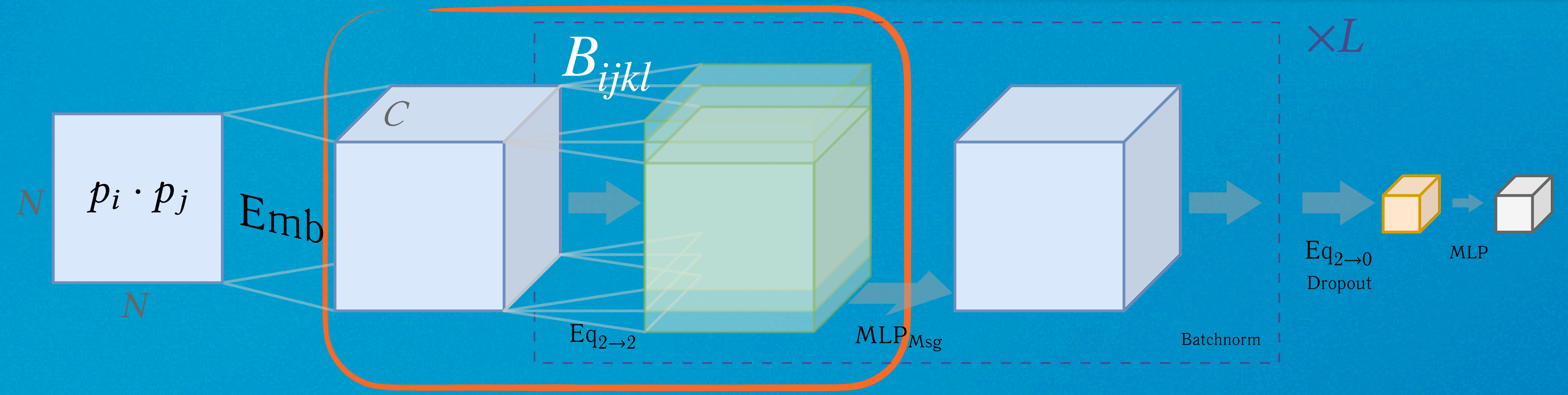
- 15 permutation equivariant linear maps of matrices into matrices
- 5 maps of matrices into vectors
- 2 maps of matrices into scalars

Identity

$\delta_{ik}\delta_{jl}(p_i \cdot p_j) = p_k \cdot p_l$

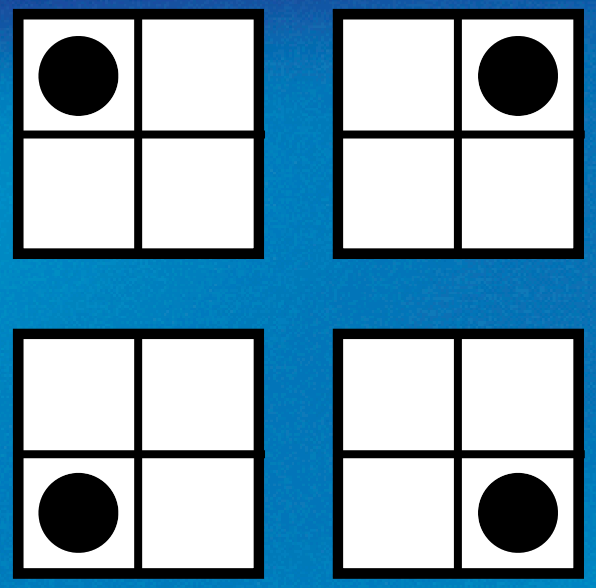






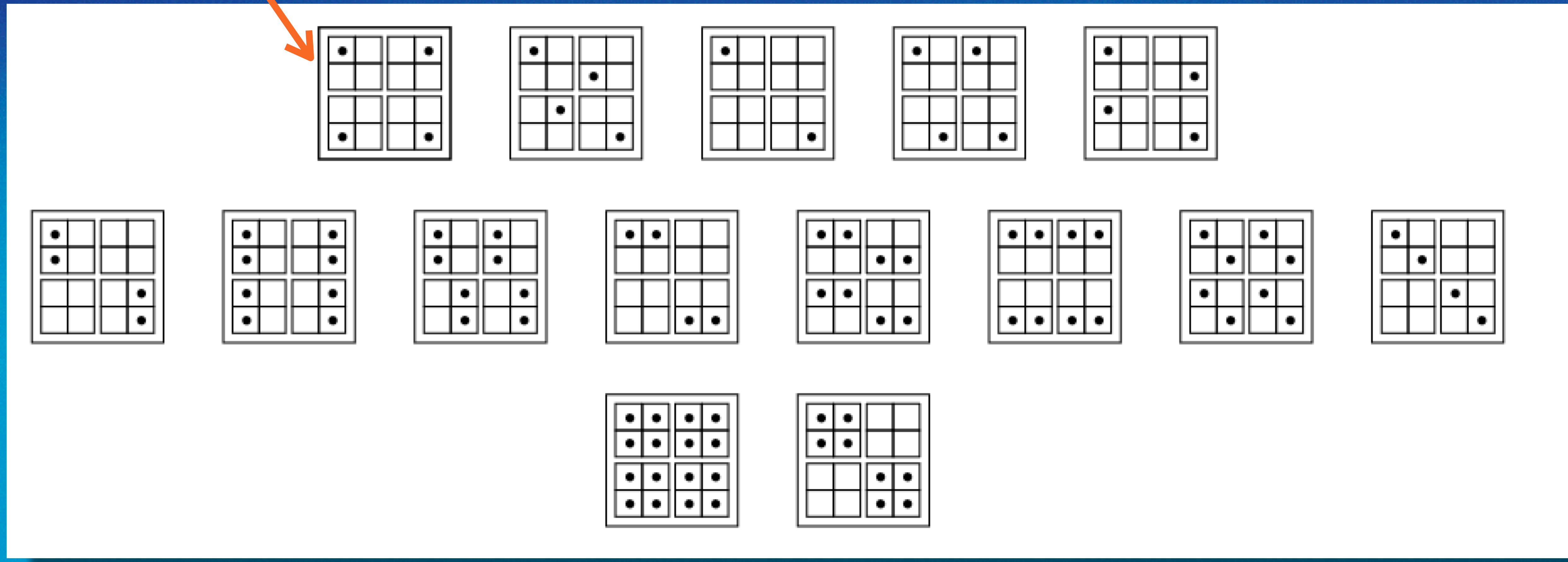
Permutation Equivariance

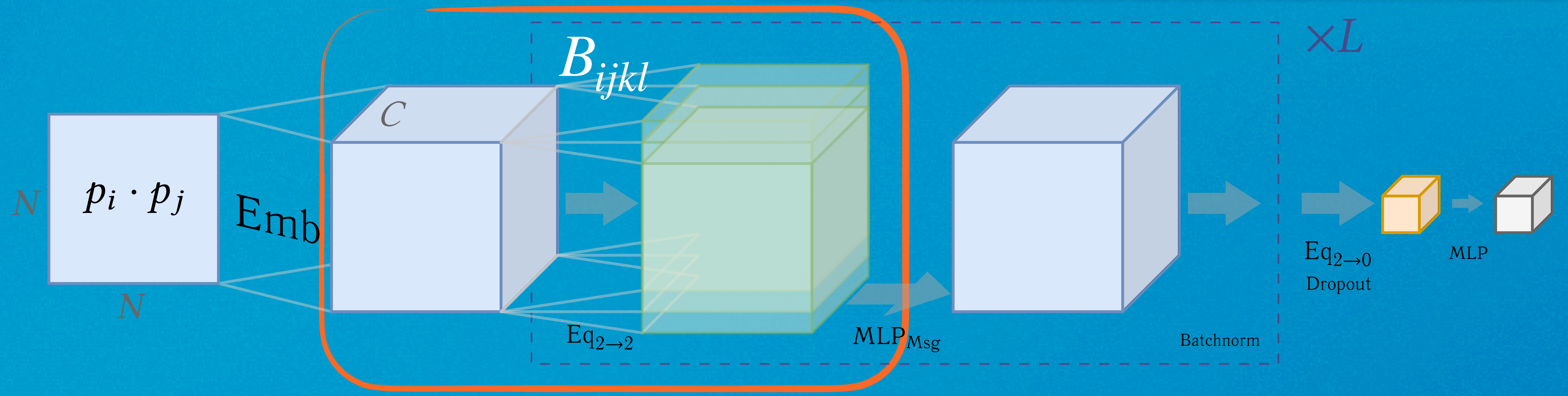
- 15 permutation equivariant linear maps of matrices into matrices
- 5 maps of matrices into vectors
- 2 maps of matrices into scalars



Identity

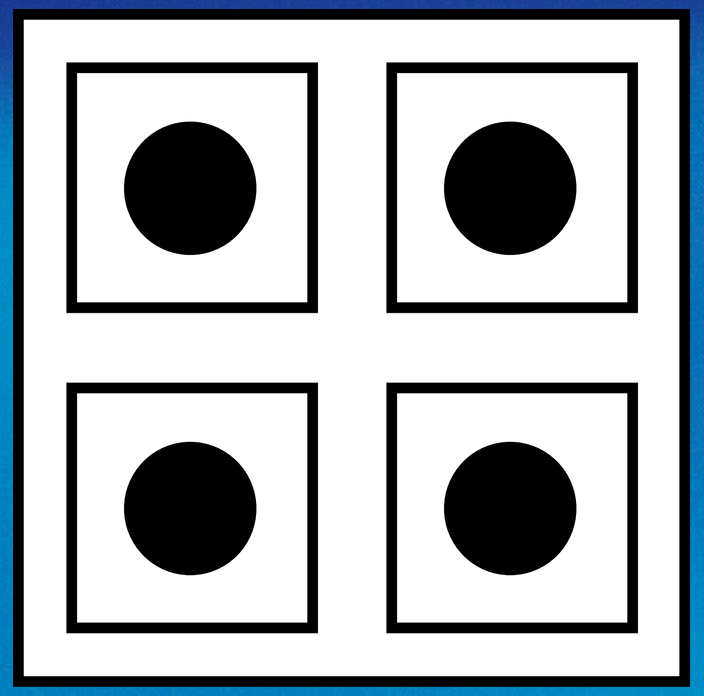
$$\delta_{ik}\delta_{jl}(p_i \cdot p_j) = p_k \cdot p_l$$





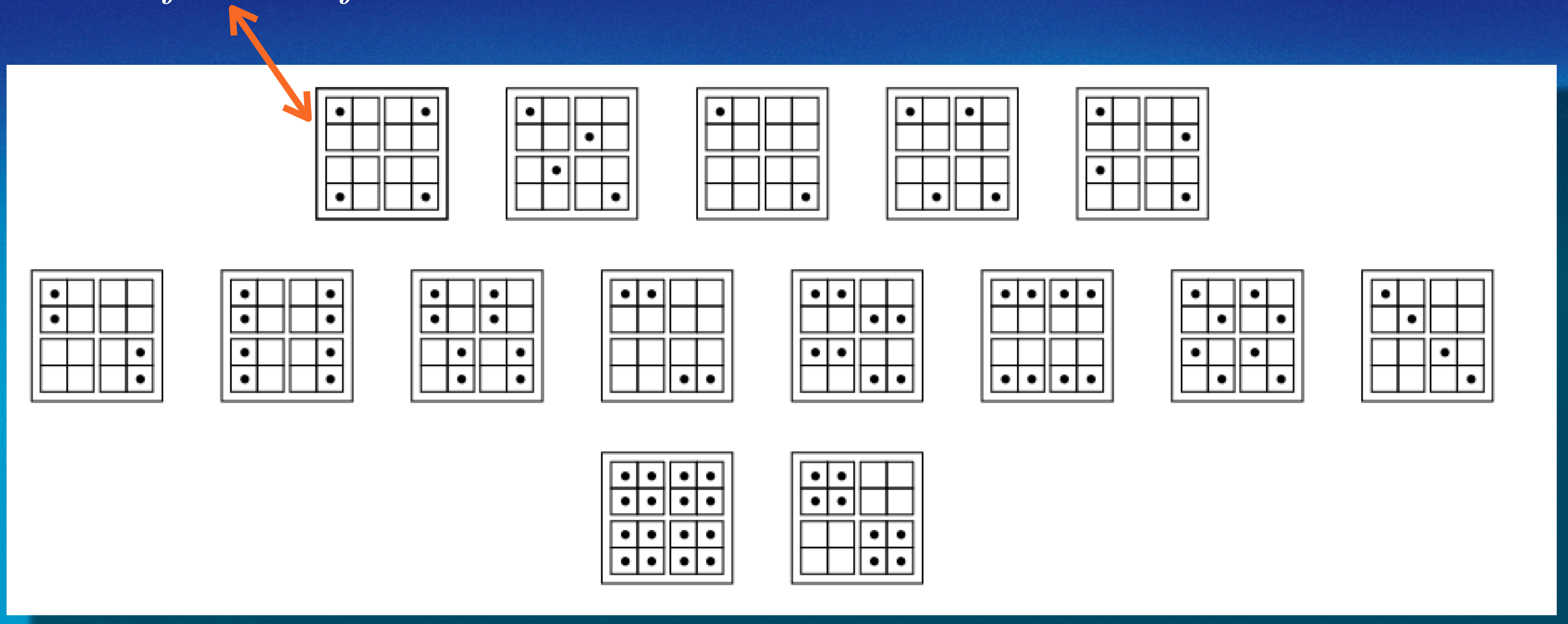
Permutation Equivariance

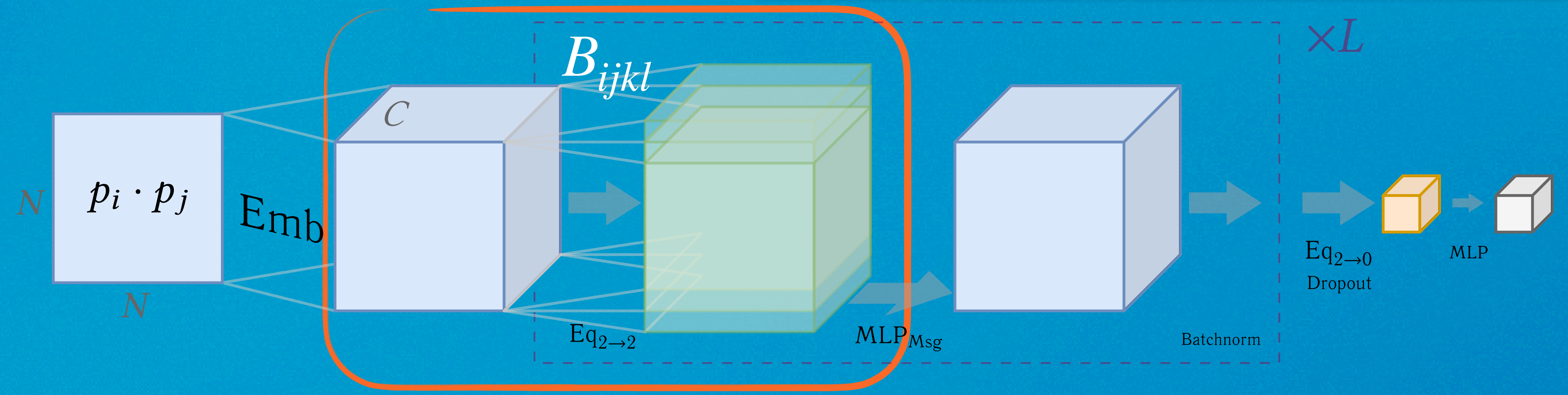
- 15 permutation equivariant linear maps of matrices into matrices
- 5 maps of matrices into vectors
- 2 maps of matrices into scalars



Identity

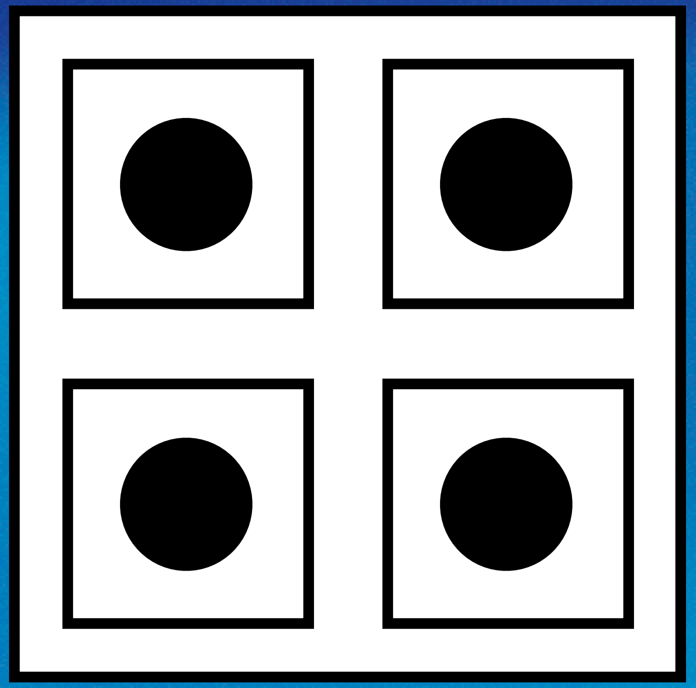
$$\delta_{ik}\delta_{jl}(p_i \cdot p_j) = p_k \cdot p_l$$

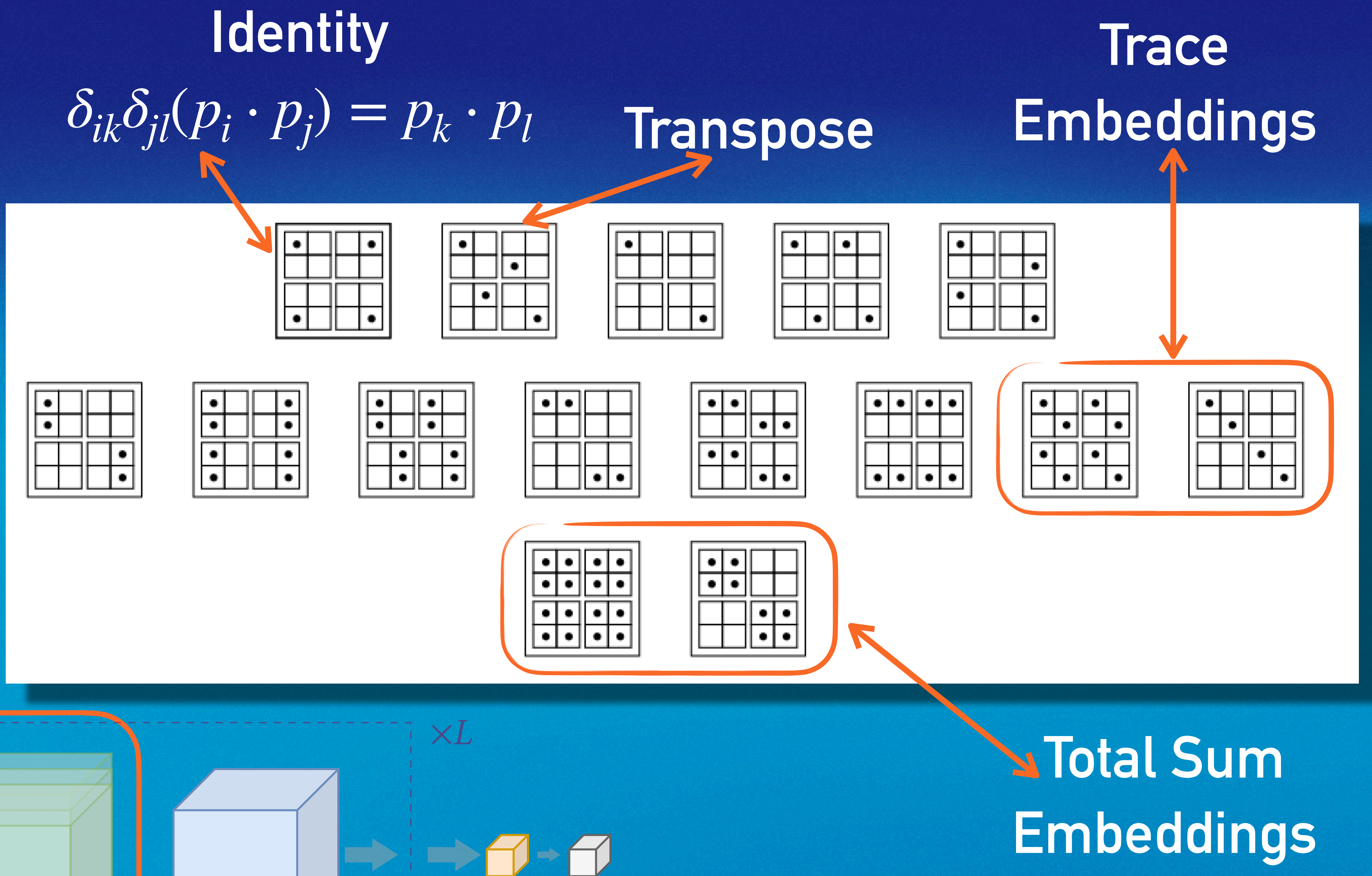


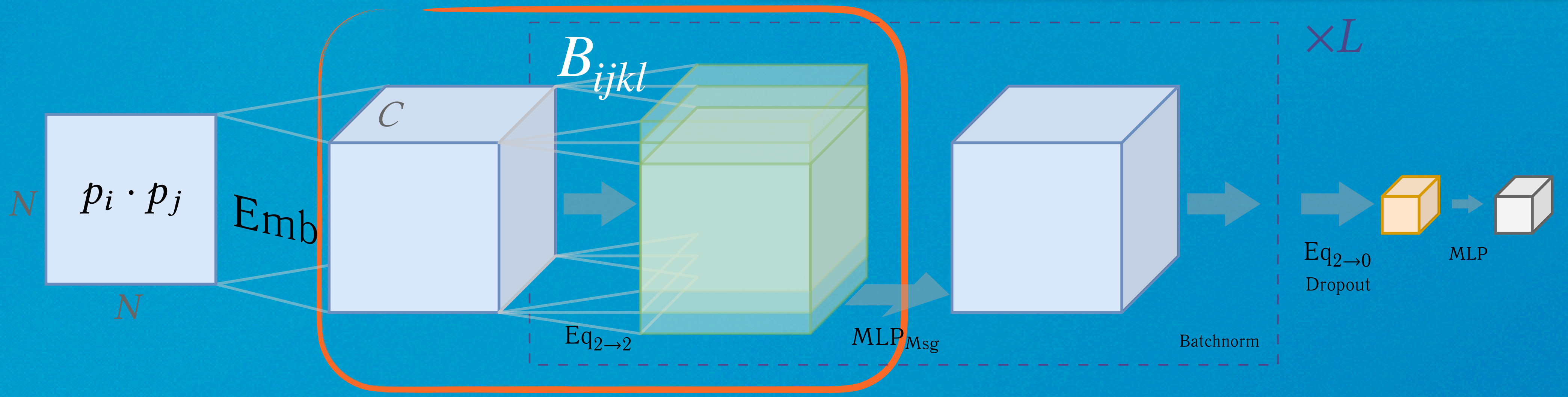


Permutation Equivariance

- 15 permutation equivariant linear maps of matrices into matrices
- 5 maps of matrices into vectors
- 2 maps of matrices into scalars



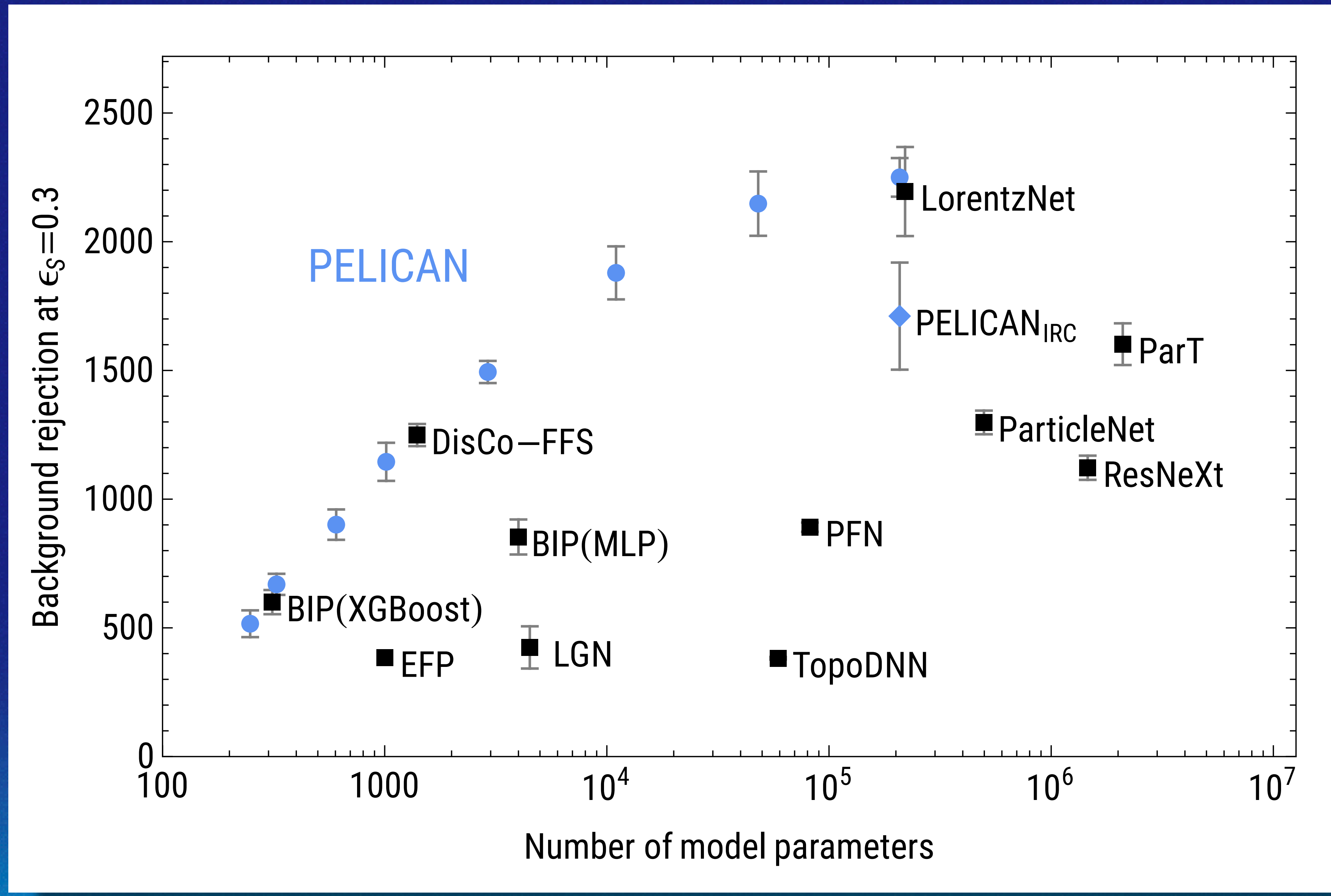




Tagging

- Target Lorentz-scalars: $f(\{p_i\}) = I(\{p_i \cdot p_j\})$
- Hadronic top-jet vs. QCD background jets
- SOTA performance at all orders of parameters!

Toptag dataset



Multiclass jet dataset (q, g, W, Z, t)

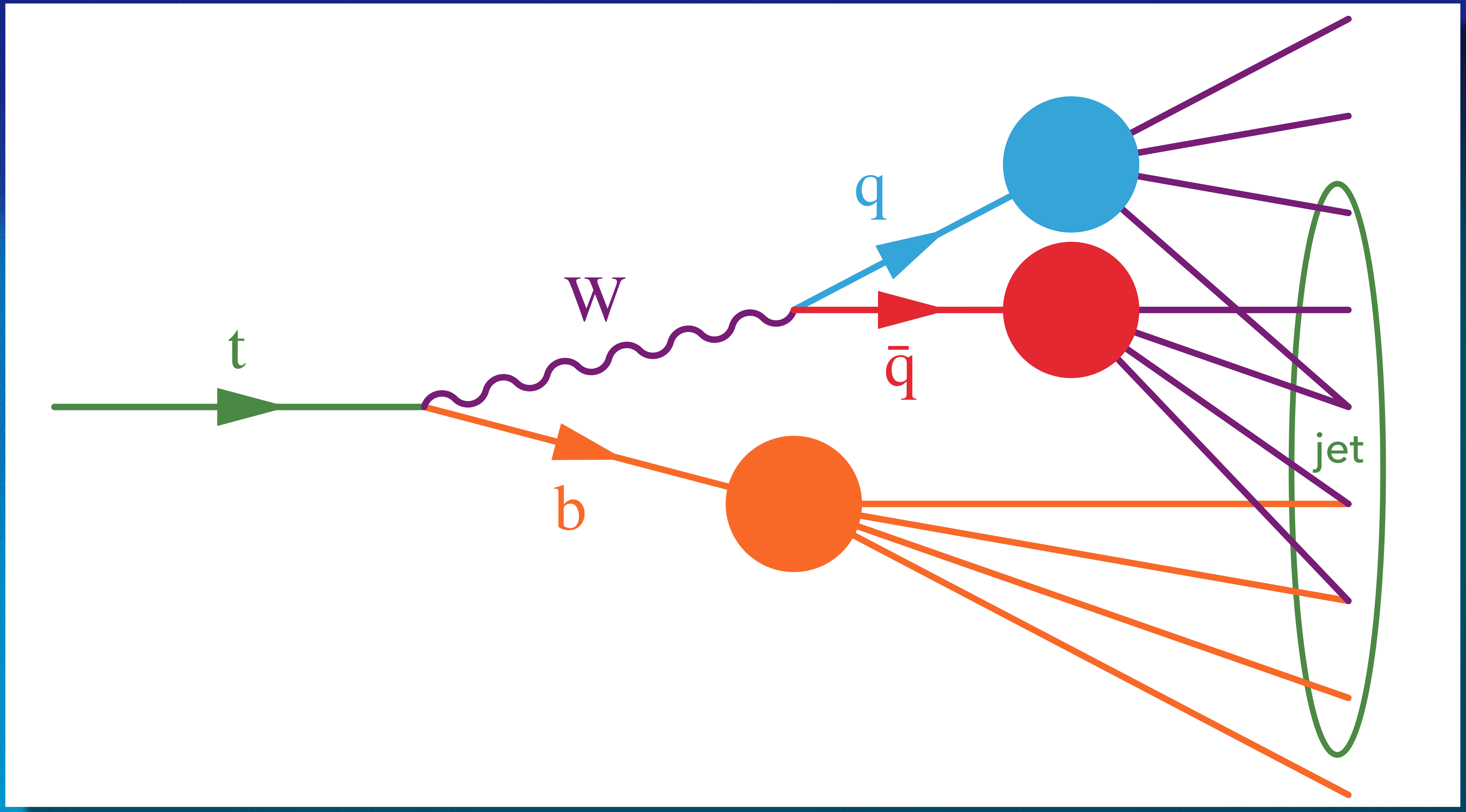
Architecture	Gluon	Light quark	W-boson	Z-boson	Top quark	# Params
			AUC			
JEDI-net	0.9529	0.9301	0.9739	0.9679	0.9683	34k
PCT	0.9623	0.9414	0.9789	0.9814	0.9757	193k
LorentzNet	0.9681(3)	0.9479(4)	0.9837(2)	0.9813(3)	0.9793(3)	224k
PELICAN	**0.9693(1)**	**0.9493(1)**	**0.9840(1)**	**0.9816(1)**	**0.9803(1)**	208k

Quark-gluon dataset

Architecture	Accuracy	AUC	$1/\epsilon_B$ ($\epsilon_S = 0.3$)	$1/\epsilon_B$ ($\epsilon_S = 0.5$)	# Params
Not IRC-safe, w/ PID					
PFN-ID[24]	–	0.9052(7)	–	37.4 ± 0.7	82k
ParticleNet-ID[50]	0.840	0.9116	98.6 ± 1.3	39.8 ± 0.2	498k
ABCNet[26]	0.840	0.9126	118.2 ± 1.5	–	230k
LorentzNet[26]	0.844	0.9156	110.2 ± 1.3	42.4 ± 0.4	220k
ParT$_{full}$ [35]	0.849	0.9203	129.5 ± 0.9	47.9 ± 0.5	2.1M
PELICAN$_{PID}$	0.8555(2)	0.9247(3)	134.8 ± 1.8	51.3 ± 0.7	211k

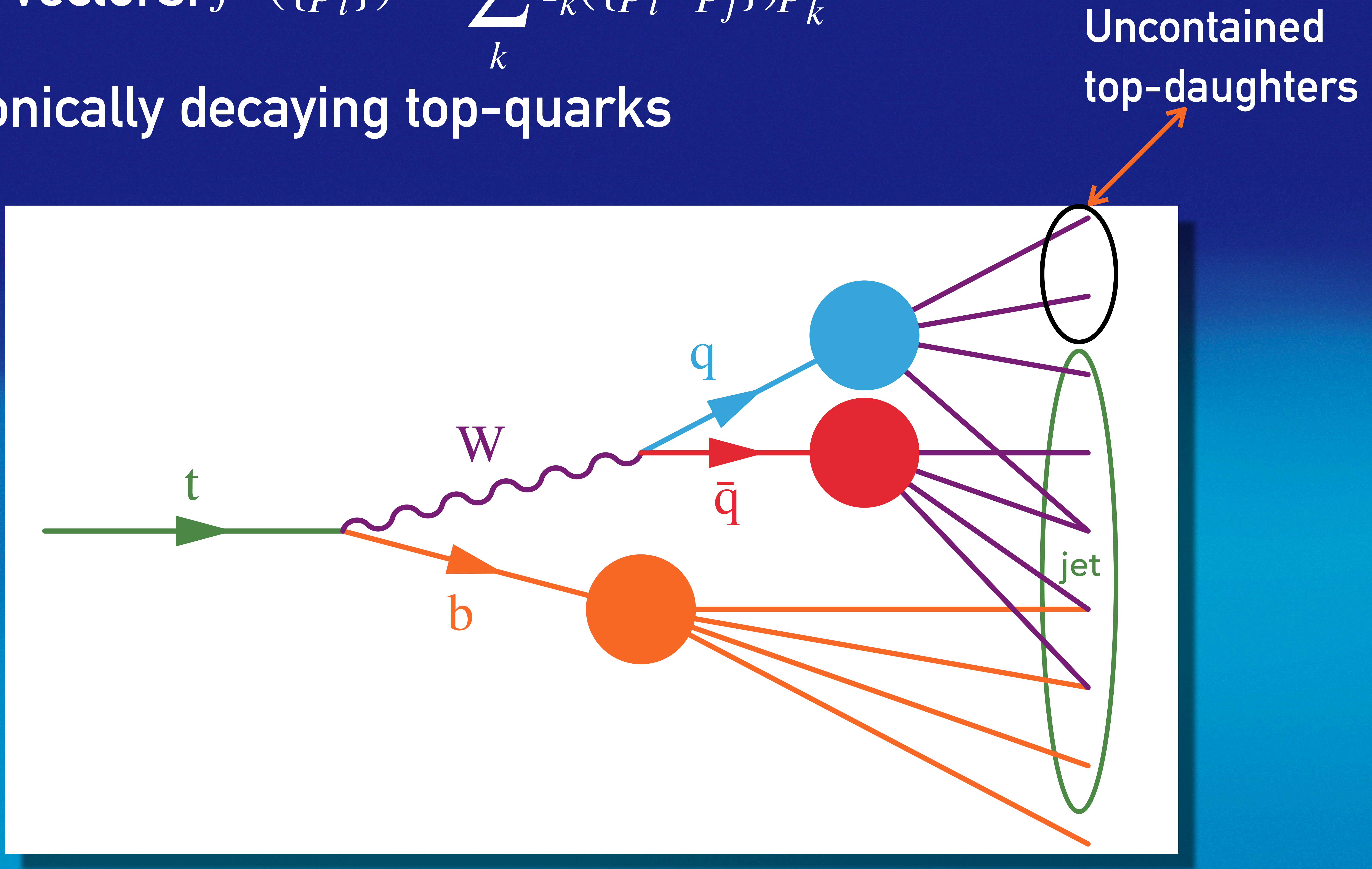
Four-Vector Reconstruction

- Target Lorentz-vectors: $f^{\mu}(\{p_i\}) = \sum_k I_k(\{p_i \cdot p_j\}) p_k^{\mu}$
- Consider hadronically decaying top-quarks



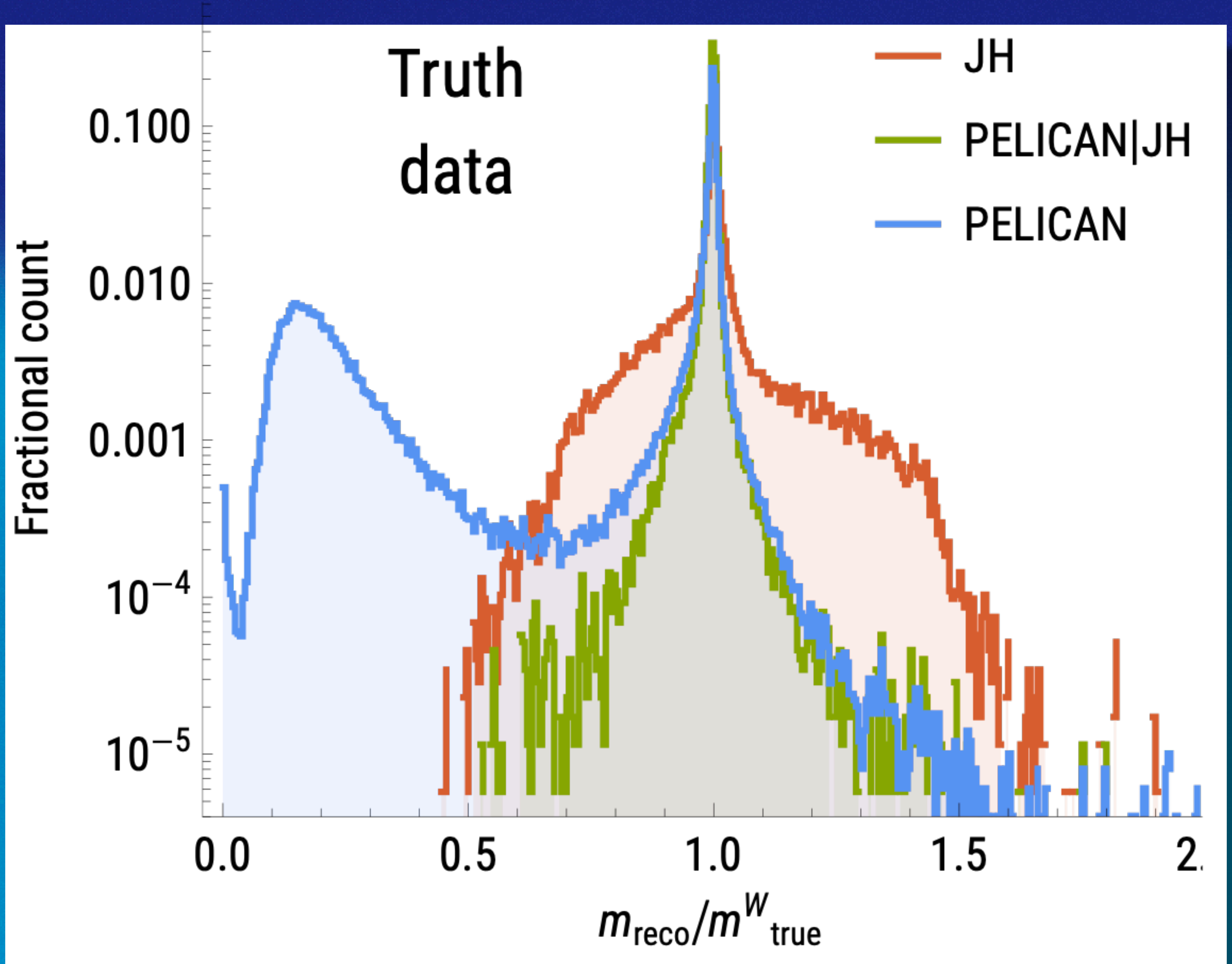
Four-Vector Reconstruction

- Target Lorentz-vectors: $f^{\mu}(\{p_i\}) = \sum_k I_k(\{p_i \cdot p_j\})p_k^{\mu}$
- Consider hadronically decaying top-quarks

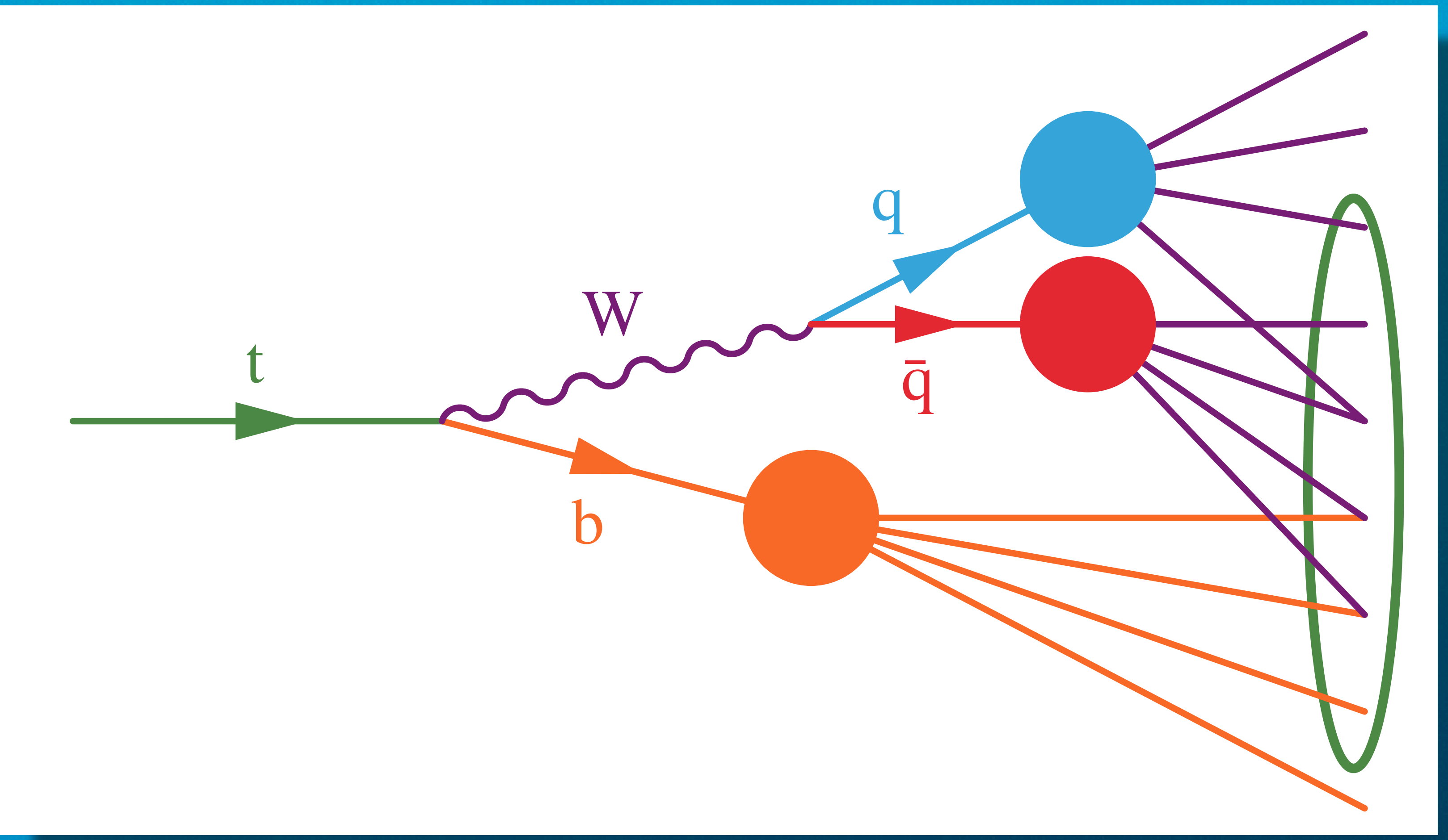


Four-Vector Reconstruction

- Target Lorentz-vectors: $f^{\mu}(\{p_i\}) = \sum_k I_k(\{p_i \cdot p_j\}) p_k^{\mu}$
- Compare to John-Hopkins W-boson reconstruction



	Method	σ_{p_T} (%)	σ_m (%)	$\sigma_{\Delta R}$ (centirad)
DELPHES	JH	9.8 %	8.3 %	9.6
	PELICAN|JH	3.6 %	2.8 %	3.1
	PELICAN	6.2 %	39.6 %	5.6



Reconstruction as a Tagger

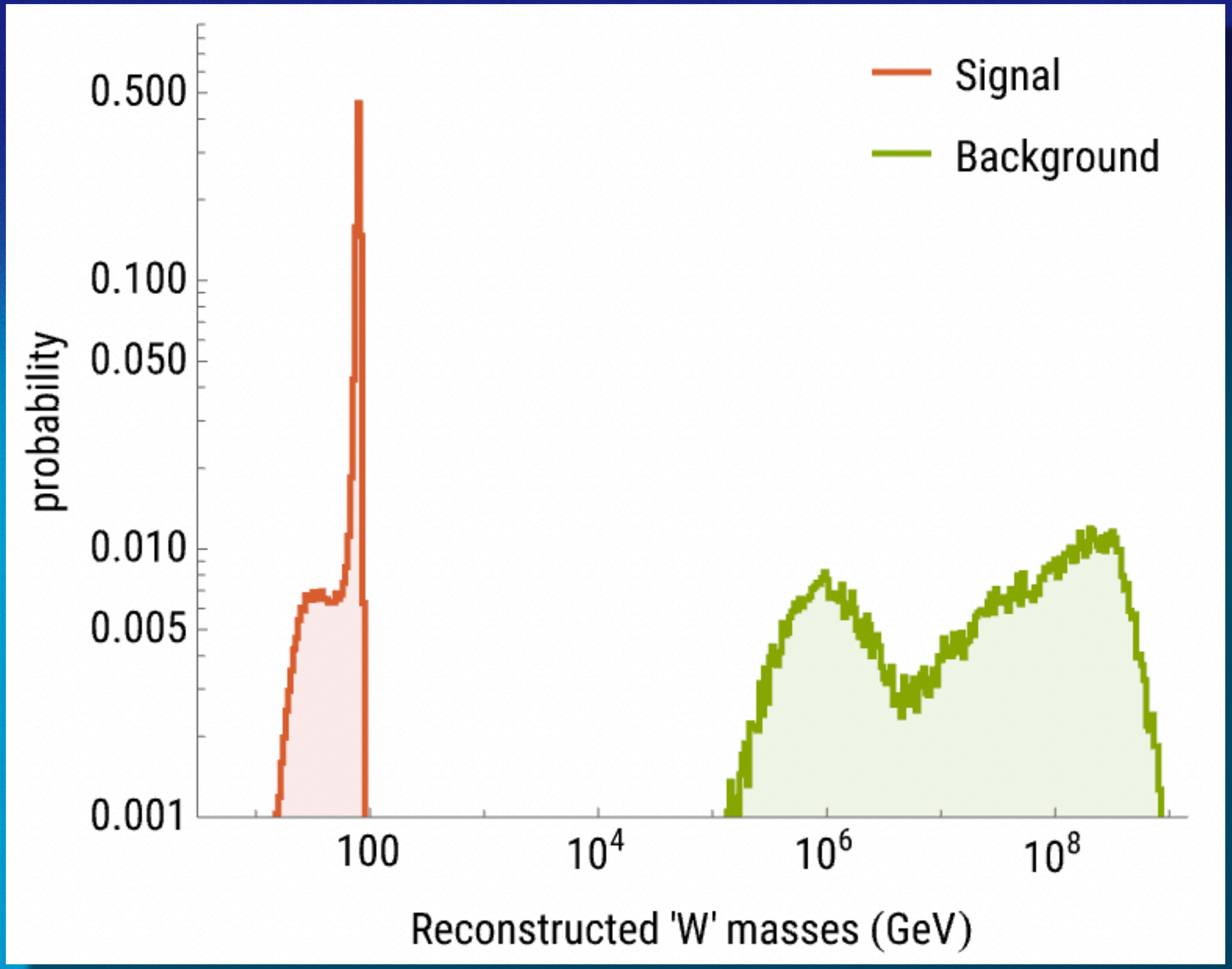
- Start with mixed bag of top-quark and QCD jets: classify → vector reconstruction

	True Signal	True Background
Predicted Signal	88.7%	11.2%
Predicted Background	11.3%	88.8%

Reconstruction as a Tagger

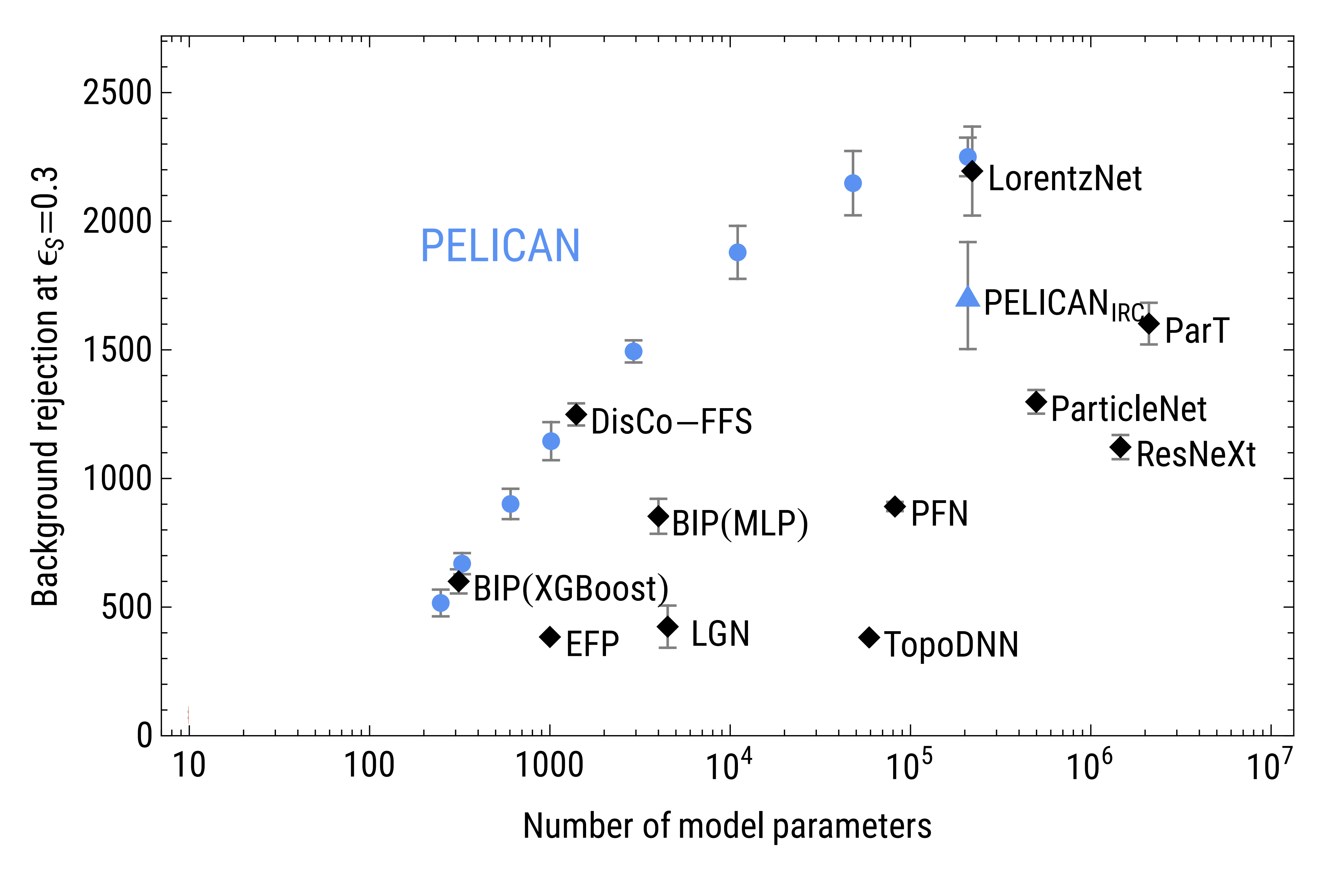
- Start with mixed bag of top-quark and QCD jets: classify → vector reconstruction
- Can cut on assumed reconstructed "W-boson" as a tagger!

	True Signal	True Background
Predicted Signal	88.7%	11.2%
Predicted Background	11.3%	88.8%



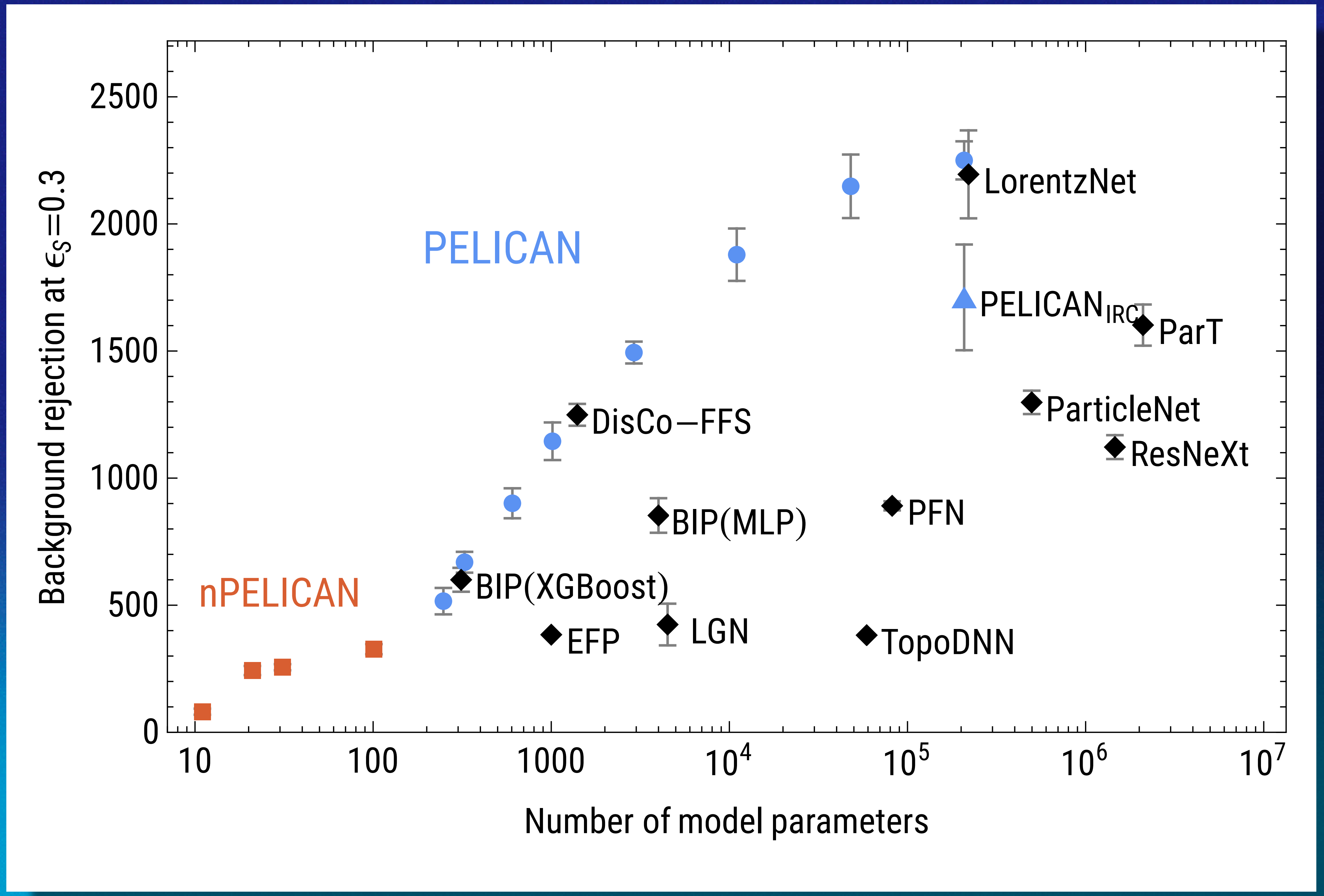
Small Model Limit: nanoPELICAN

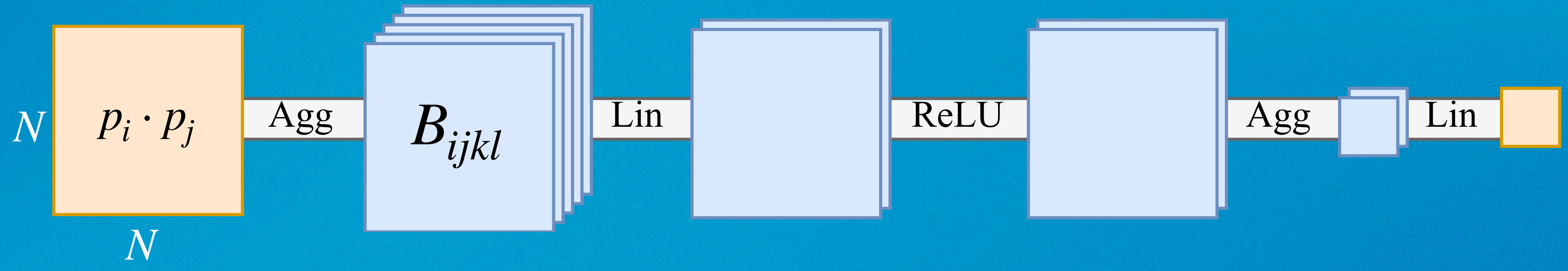
- Interpretability is crucial in particle physics
- Construct smallest viable models



Small Model Limit: nanoPELICAN

- Interpretability is crucial in particle physics
- Construct smallest viable models
- Reduce to a single hidden layer
- $p_i \cdot p_j$ is traceless and symmetric
- Only 6 aggregators remain

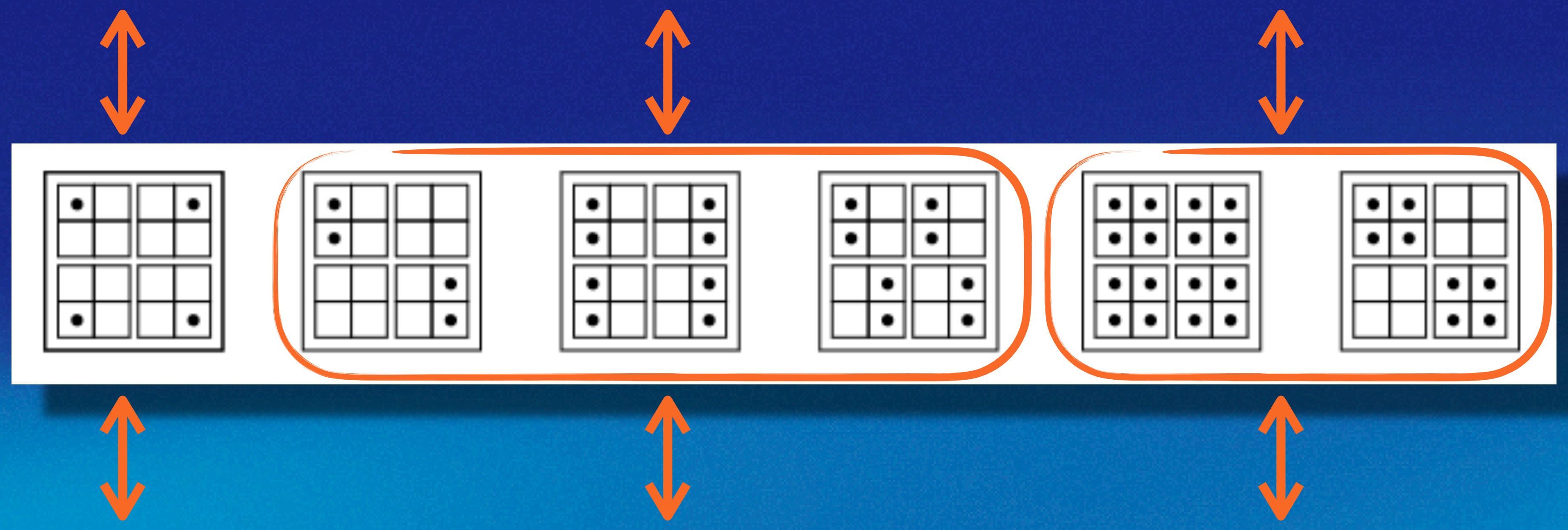




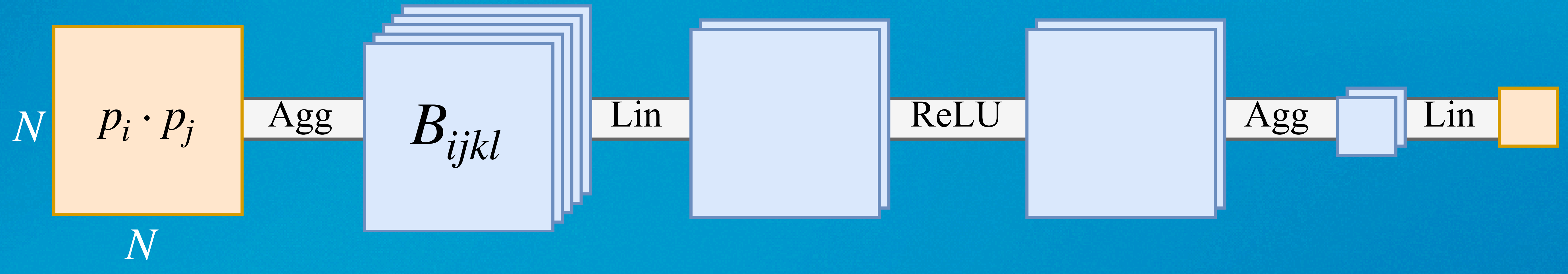
nanoPELICAN

- Only 6 aggregators remain

Identity | Row Aggregation | Total Aggregation

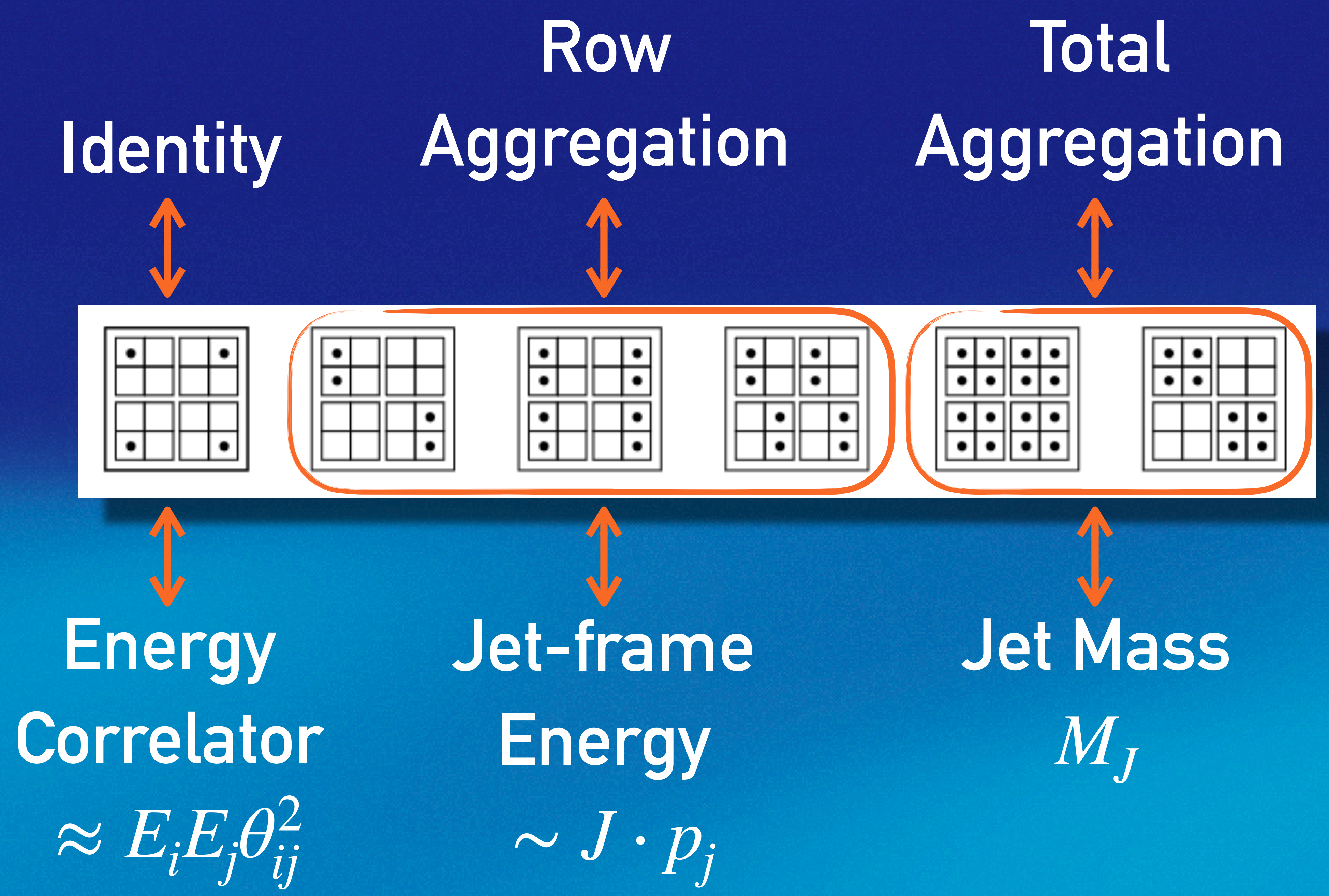


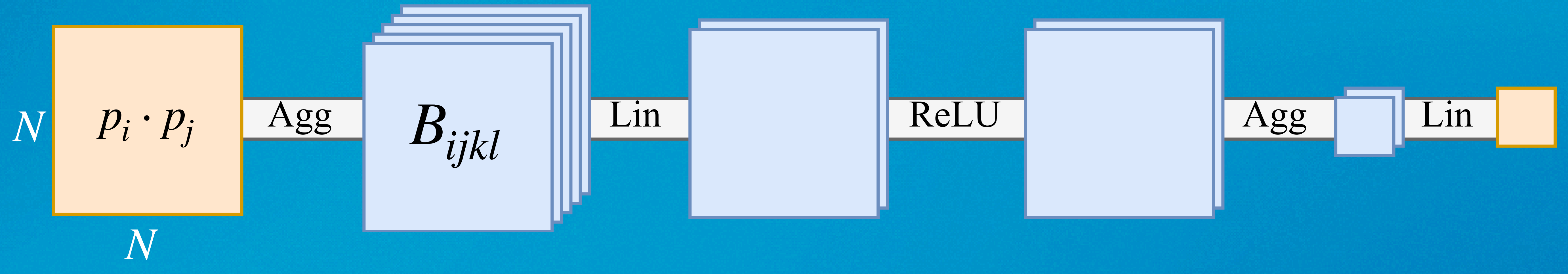
Energy Correlator $\approx E_i E_j \theta_{ij}^2$ | Jet-frame Energy $\sim J \cdot p_j$ | Jet Mass M_J



nanoPELICAN

- Only 6 aggregators remain

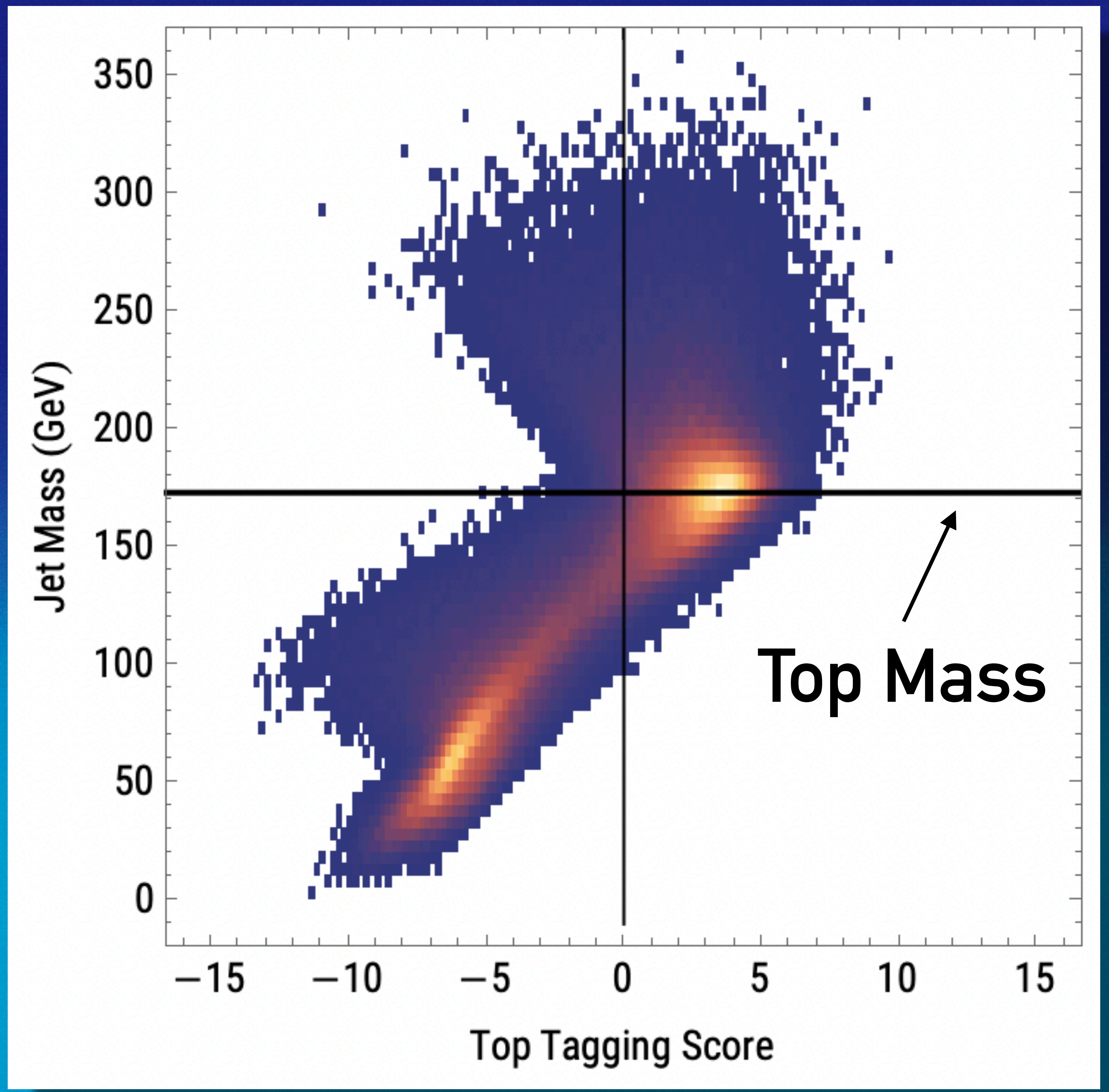




Accuracy: 91.8%

AUC: 97.2%

Parameters: 19



Conclusions

- PELICAN achieves SOTA performance with explainable outputs on Lorentz-scalar and Lorentz-vector targets
- Relatively performant even with O(10) parameters!
    - Gives hope for full PELICAN interpretability
- Many prospects for future applications!
    - Fast tagging and other online analyses
    - Explainable energy calibration
    - Particle helicity measurements
    - High-precision jet-containment measurements for offline analysis
    - Track reconstruction
    - Astrophysics applications

Thanks!

PELICAN paper

NanoPELICAN paper

PELICAN codebase

nanoPELICAN codebase

Data generation codebase

Alexander Bogatskiy   Timothy Hoffman   David W. Miller   Xiaoyang Liu   Jan T. Offermann

THE UNIVERSITY OF CHICAGO DATA SCIENCE INSTITUTE   FLATIRON INSTITUTE   ENRICO FERMI INSTITUTE   Miller Lab

Backup

More Classification

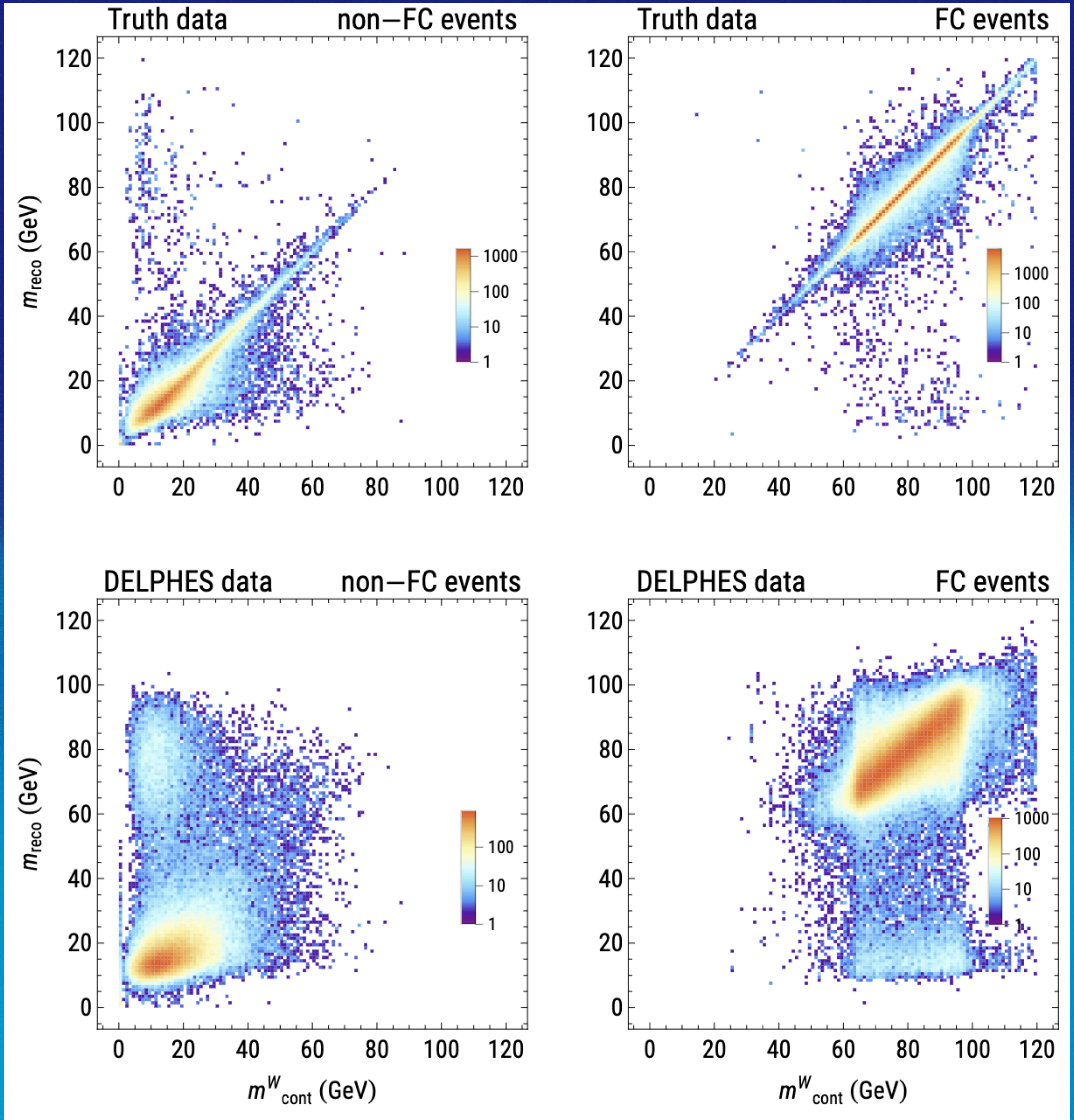
Multiclass

Architecture	Gluon	Light quark	W-boson	Z-boson	Top quark	# Params
			AUC			
JEDI-net	0.9529	0.9301	0.9739	0.9679	0.9683	34k
PCT	0.9623	0.9414	0.9789	0.9814	0.9757	193k
LorentzNet	0.9681(3)	0.9479(4)	0.9837(2)	0.9813(3)	0.9793(3)	224k
PELICAN	**0.9693(1)**	**0.9493(1)**	**0.9840(1)**	**0.9816(1)**	**0.9803(1)**	208k
			TPR at FPR=0.10			
JEDI-net	0.878(1)	0.822(1)	0.938(1)	0.910(1)	0.930(1)	34k
PCT	0.891(1)	0.833(1)	0.932(1)	**0.946(1)**	0.941(1)	193k
LorentzNet	0.912(1)	0.855(1)	0.952(1)	0.939(1)	0.949(1)	224k
PELICAN	**0.916(1)**	**0.860(1)**	**0.953(1)**	0.940(1)	**0.951(1)**	208k
			TPR at FPR=0.01			
JEDI-net	0.485(1)	0.302(1)	0.704(1)	0.769(1)	0.633(1)	34k
PCT	0.513(2)	0.298(2)	**0.834(1)**	0.781(1)	0.700(3)	193k
LorentzNet	0.557(4)	0.319(2)	0.800(3)	0.850(3)	0.753(3)	224k
PELICAN	**0.567(1)**	**0.320(1)**	0.804(1)	**0.850(1)**	**0.761(1)**	208k

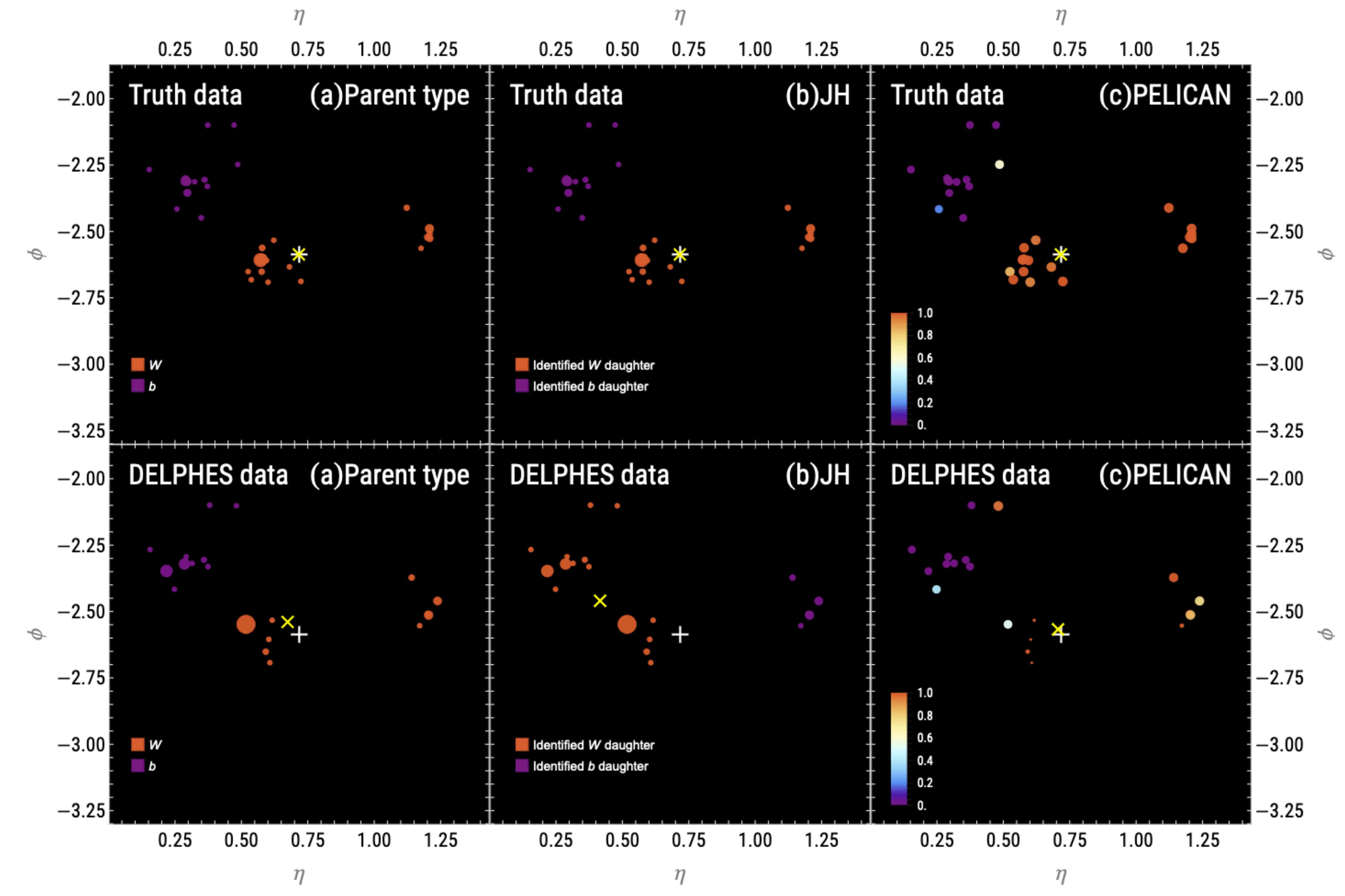
Quark-gluon

Architecture	Accuracy	AUC	$1/\epsilon_B$ $(\epsilon_S = 0.3)$	$1/\epsilon_B$ $(\epsilon_S = 0.5)$	# Params
Not IRC-safe, w/ PID					
PFN-ID[24]	–	0.9052(7)	–	37.4 ± 0.7	82k
ParticleNet-ID[50]	0.840	0.9116	98.6 ± 1.3	39.8 ± 0.2	498k
ABCNet[26]	0.840	0.9126	118.2 ± 1.5	–	230k
LorentzNet[26]	0.844	0.9156	110.2 ± 1.3	42.4 ± 0.4	220k
$\mathrm{ParT_{full}}$ [35]	0.849	0.9203	129.5 ± 0.9	47.9 ± 0.5	2.1M
$\mathrm{PELICAN_{PID}}$	0.8555(2)	0.9247(3)	134.8 ± 1.8	51.3 ± 0.7	211k
Not IRC-safe, w/o PID					
PFN[24]	–	0.8911(8)	–	30.8 ± 0.4	82k
ParticleNet[50]	0.828	0.9014	85.4	33.7	498k
PELICAN	0.8342(2)	0.9059(8)	88.9 ± 0.5	36.0 ± 0.2	209k
IRC-safe					
EFN[24]	–	0.8824(5)	–	28.6 ± 0.3	82k
EFP[18]	–	0.8919	–	29.7	1k
EMPN[55]	–	0.8932(6)	–	30.8 ± 0.2	~110k
$\mathrm{PELICAN_{IRC}}$	0.8299(3)	0.8955(18)	85.7 ± 1.2	33.8 ± 0.2	209k

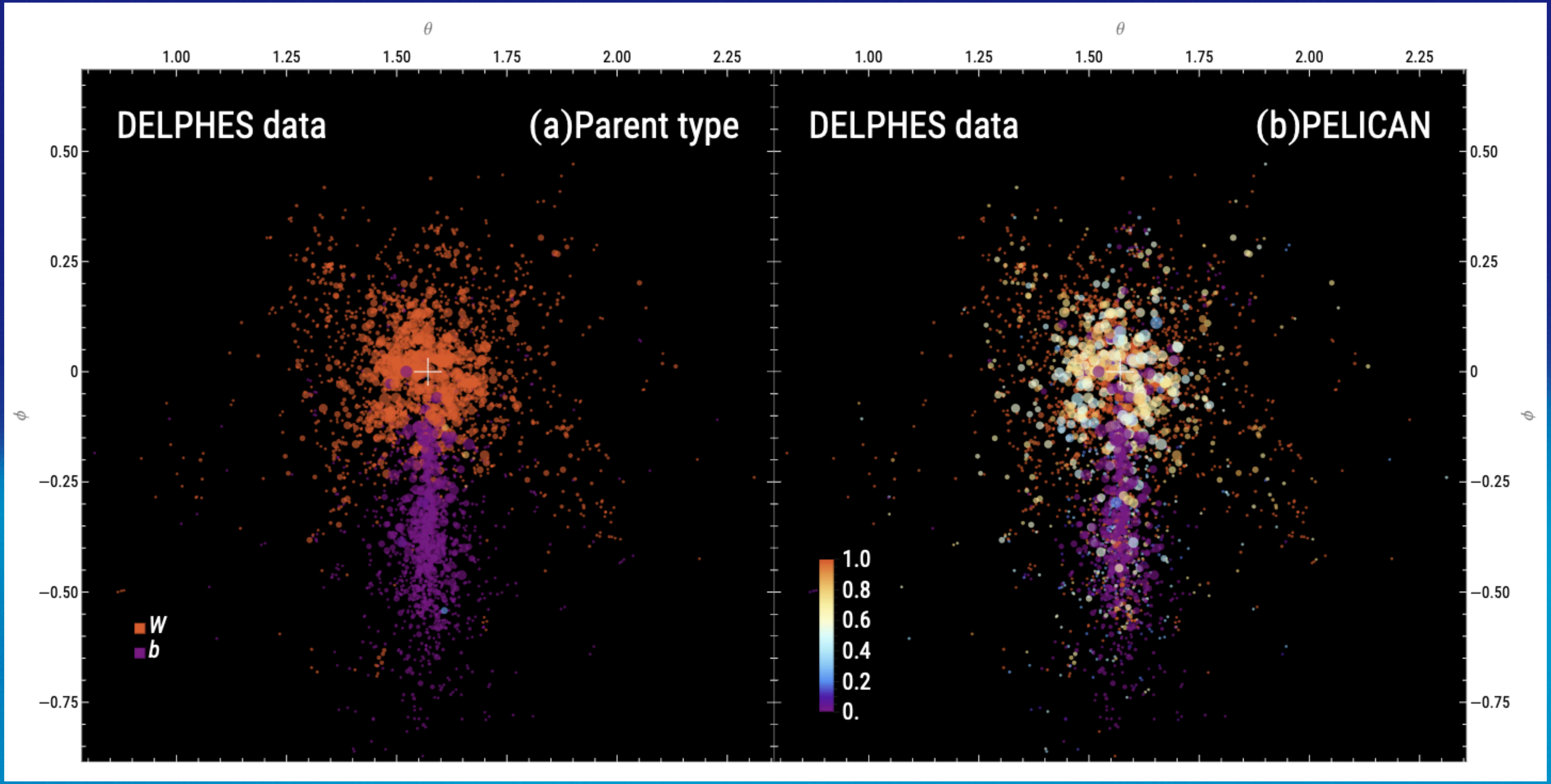
Vector Reconstruction



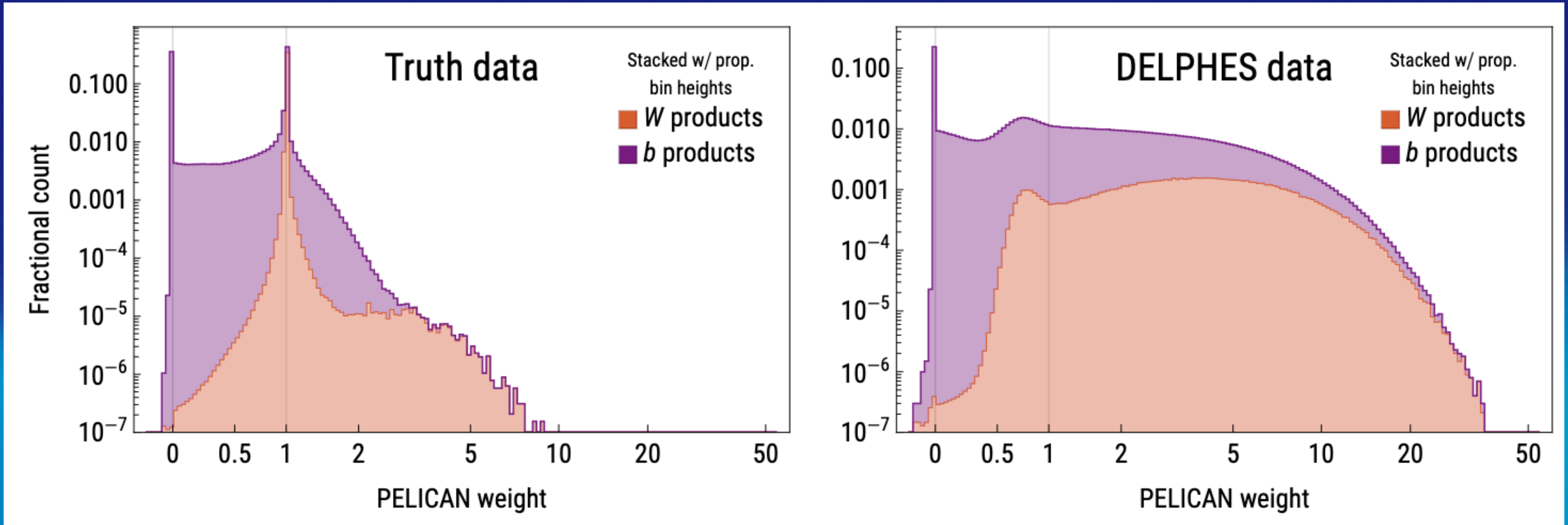
Vector Reconstruction



Vector Reconstruction



Vector Reconstruction Weights



Dataset Details