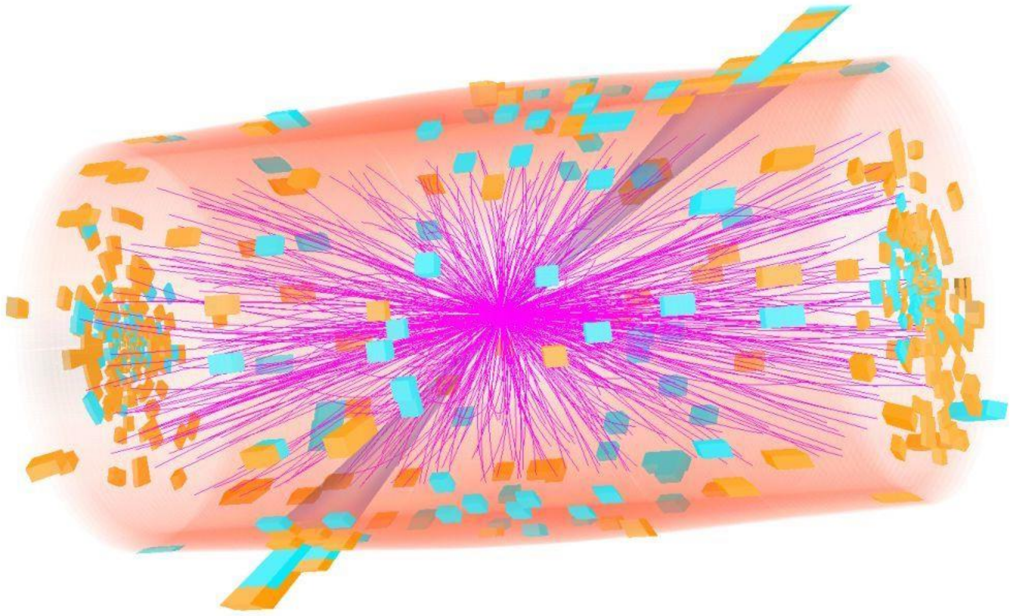
FAST AND PRECISE AMPLITUDE SURROGATES WITH SYMMETRY EQUIVARIANT NETWORKS

Víctor Bresó Pla

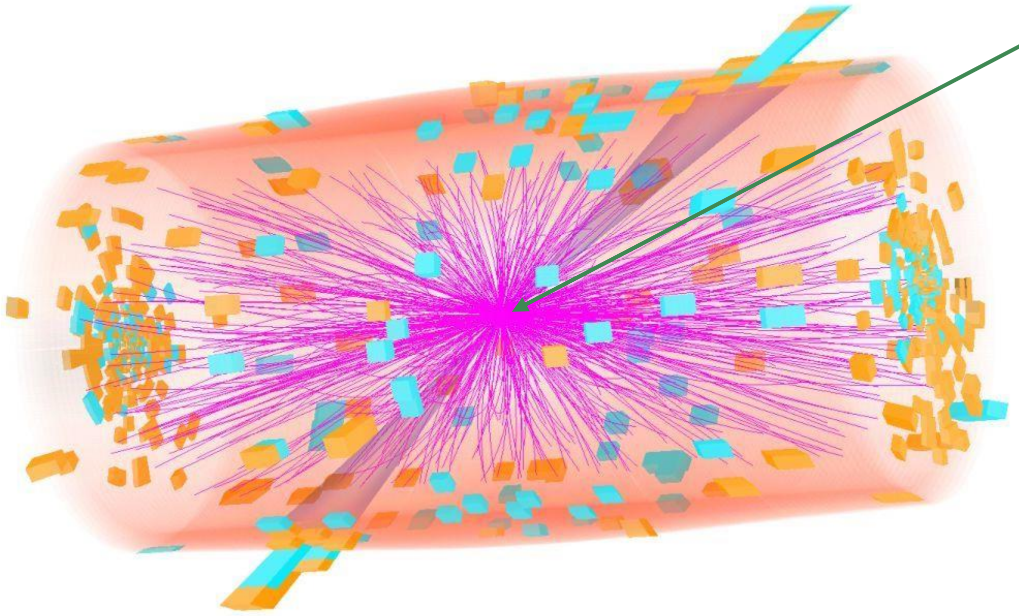
In collaboration with Jonas Spinner, Johann Brehmer, Pim de Haan, Tilman Plehn & Jesse Thaler

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

PROBLEMS IN COLLIDER PHYSICS



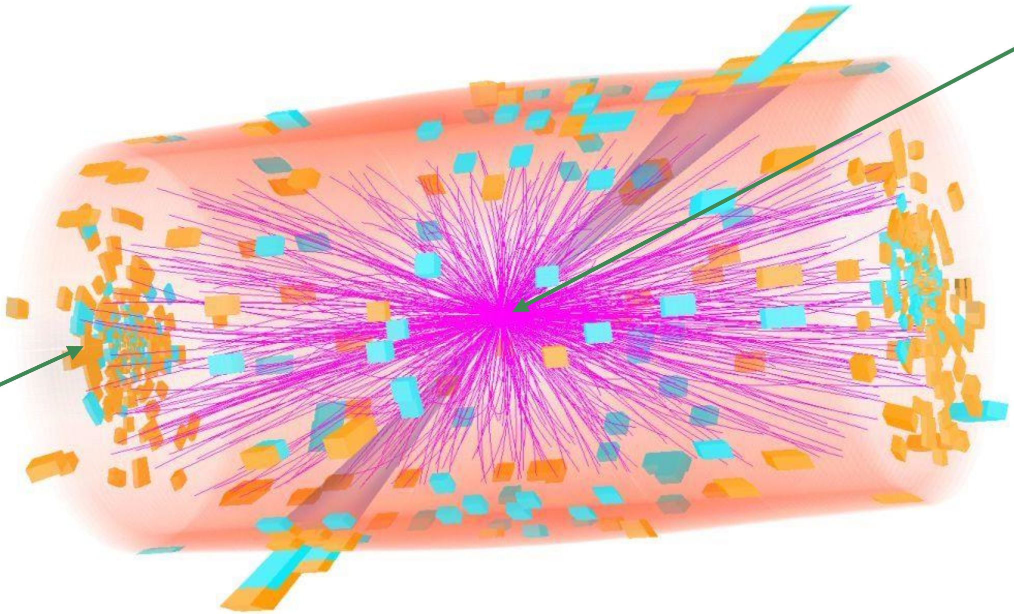
PROBLEMS IN COLLIDER PHYSICS



Predict the hard process interactions in a fast and precise way

PROBLEMS IN COLLIDER PHYSICS

Predict the hard process interactions in a fast and precise way

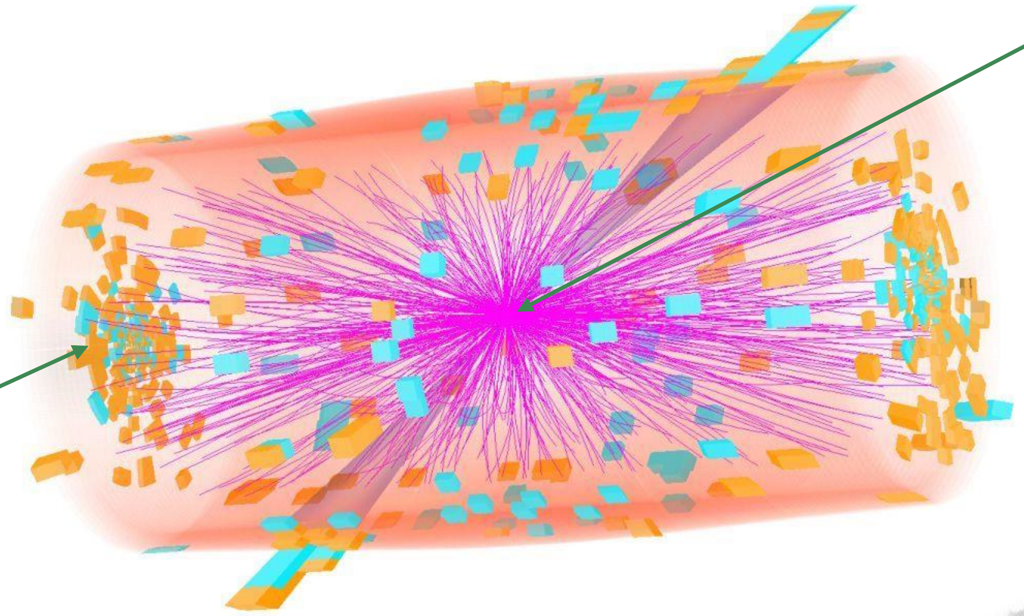


Identify particle structures at the detector level

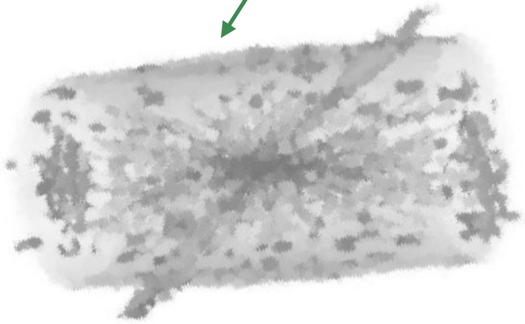
PROBLEMS IN COLLIDER PHYSICS

Predict the hard process interactions in a fast and precise way

Identify particle structures at the detector level

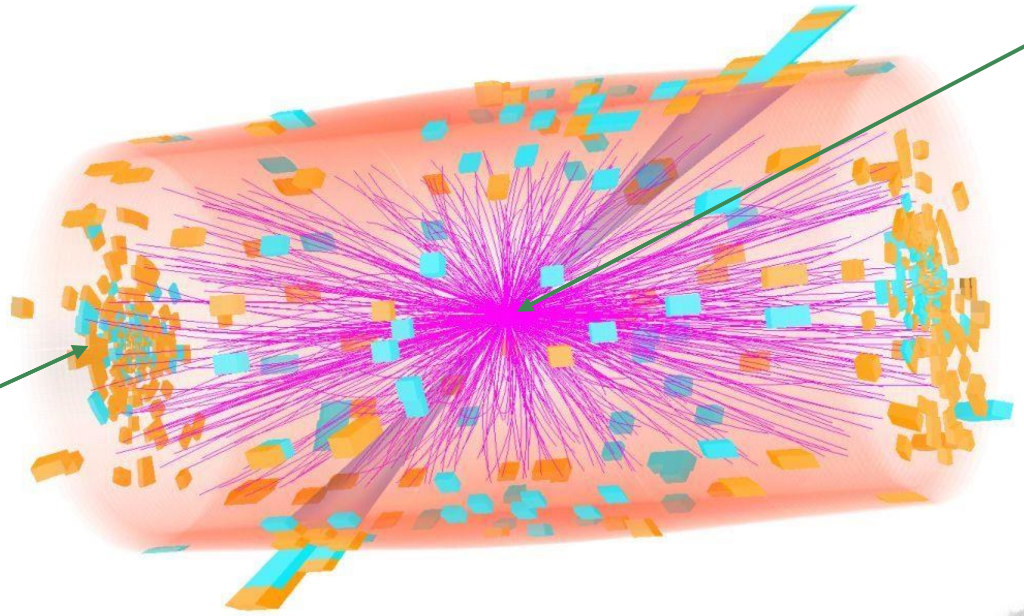


Simulate full events



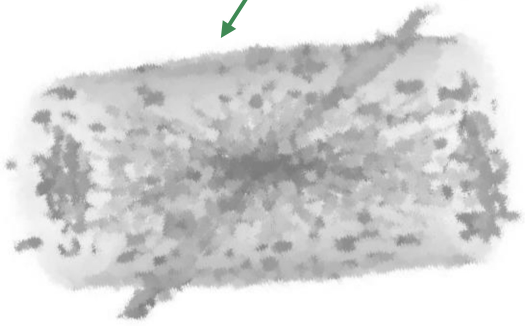
PROBLEMS IN COLLIDER PHYSICS

Predict the hard process interactions in a fast and precise way



Identify particle structures at the detector level

Simulate full events



All collider processes are ruled by strict and concrete **symmetry laws**

GEOMETRIC ALGEBRA TRANSFORMER

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias
- Symmetry awareness

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias
- Symmetry awareness
- Scalability and flexibility

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias
  - GEOMETRIC ALGEBRA
- Symmetry awareness
- Scalability and flexibility

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

$$x = x_s + x_0e_0 + x_1e_1 + x_2e_2 + x_3e_3 + x_{01}e_0e_1 + x_{02}e_0e_2 + x_{03}e_0e_3 + x_{12}e_1e_2 + x_{13}e_1e_3 + x_{23}e_2e_3 \\ + x_{012}e_0e_1e_2 + x_{013}e_0e_1e_3 + x_{023}e_0e_2e_3 + x_{123}e_1e_2e_3 + x_{0123}e_0e_1e_2e_3$$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

Scalars

$$x = \boxed{x_s} + x_0 e_0 + x_1 e_1 + x_2 e_2 + x_3 e_3 + x_{01} e_0 e_1 + x_{02} e_0 e_2 + x_{03} e_0 e_3 + x_{12} e_1 e_2 + x_{13} e_1 e_3 + x_{23} e_2 e_3 \\ + x_{012} e_0 e_1 e_2 + x_{013} e_0 e_1 e_3 + x_{023} e_0 e_2 e_3 + x_{123} e_1 e_2 e_3 + x_{0123} e_0 e_1 e_2 e_3$$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

Scalars

Lorentz vectors

$$x = x_s + x_0e_0 + x_1e_1 + x_2e_2 + x_3e_3 + x_{01}e_0e_1 + x_{02}e_0e_2 + x_{03}e_0e_3 + x_{12}e_1e_2 + x_{13}e_1e_3 + x_{23}e_2e_3$$
$$+ x_{012}e_0e_1e_2 + x_{013}e_0e_1e_3 + x_{023}e_0e_2e_3 + x_{123}e_1e_2e_3 + x_{0123}e_0e_1e_2e_3$$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

Scalars — Lorentz vectors — Lorentz bilinears

$$x = \underbrace{x_s}_{\text{Scalars}} + \underbrace{x_0 e_0 + x_1 e_1 + x_2 e_2 + x_3 e_3}_{\text{Lorentz vectors}} + \underbrace{x_{01} e_0 e_1 + x_{02} e_0 e_2 + x_{03} e_0 e_3 + x_{12} e_1 e_2 + x_{13} e_1 e_3 + x_{23} e_2 e_3}_{\text{Lorentz bilinears}}$$
$$+ x_{012} e_0 e_1 e_2 + x_{013} e_0 e_1 e_3 + x_{023} e_0 e_2 e_3 + x_{123} e_1 e_2 e_3 + x_{0123} e_0 e_1 e_2 e_3$$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

$$x = \underbrace{x_s}_{\text{Scalars}} + \underbrace{x_0 e_0 + x_1 e_1 + x_2 e_2 + x_3 e_3}_{\text{Lorentz vectors}} + \underbrace{x_{01} e_0 e_1 + x_{02} e_0 e_2 + x_{03} e_0 e_3 + x_{12} e_1 e_2 + x_{13} e_1 e_3 + x_{23} e_2 e_3}_{\text{Lorentz bilinears}}$$

$$+ \underbrace{x_{012} e_0 e_1 e_2 + x_{013} e_0 e_1 e_3 + x_{023} e_0 e_2 e_3 + x_{123} e_1 e_2 e_3}_{\text{Axial vectors}} + x_{0123} e_0 e_1 e_2 e_3$$

GEOMETRIC ALGEBRA

- A geometric algebra is a vector space that features an extra operation: the geometric product: $x, y \mapsto xy$
- A geometric algebra deals with multivectors, representing both objects and operators
- A general multivector in the Lorentz geometric algebra will be:

$$x = \underbrace{x_s}_{\text{Scalars}} + \underbrace{x_0 e_0 + x_1 e_1 + x_2 e_2 + x_3 e_3}_{\text{Lorentz vectors}} + \underbrace{x_{01} e_0 e_1 + x_{02} e_0 e_2 + x_{03} e_0 e_3 + x_{12} e_1 e_2 + x_{13} e_1 e_3 + x_{23} e_2 e_3}_{\text{Lorentz bilinears}}$$

$$+ \underbrace{x_{012} e_0 e_1 e_2 + x_{013} e_0 e_1 e_3 + x_{023} e_0 e_2 e_3 + x_{123} e_1 e_2 e_3}_{\text{Axial vectors}} + \underbrace{x_{0123} e_0 e_1 e_2 e_3}_{\text{Pseudoscalars}}$$

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias
- Symmetry awareness
- Scalability and flexibility

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias through geometric algebra representations
- Symmetry awareness
- Scalability and flexibility

GEOMETRIC ALGEBRA TRANSFORMER

- Geometric inductive bias through geometric algebra representations
- Symmetry awareness through Lorentz equivariant layers
- Scalability and flexibility

GEOMETRIC ALGEBRA TRANSFORMER

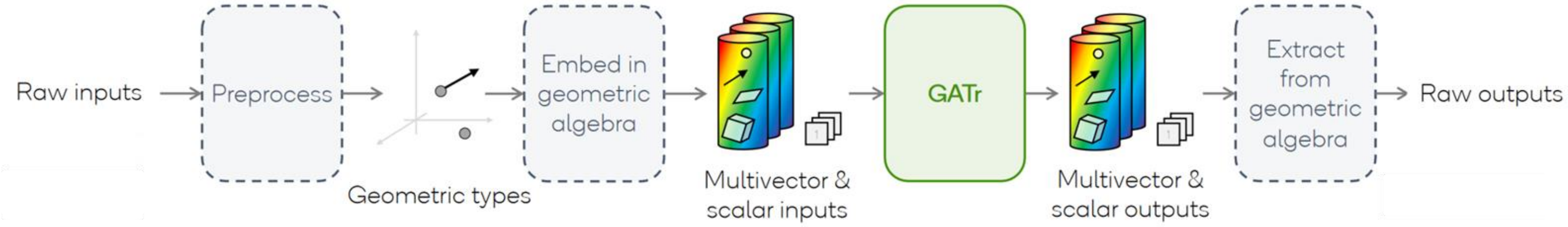
- Geometric inductive bias through geometric algebra representations
- Symmetry awareness through Lorentz equivariant layers
- Scalability and flexibility through dot-product attention

GEOMETRIC ALGEBRA TRANSFORMER

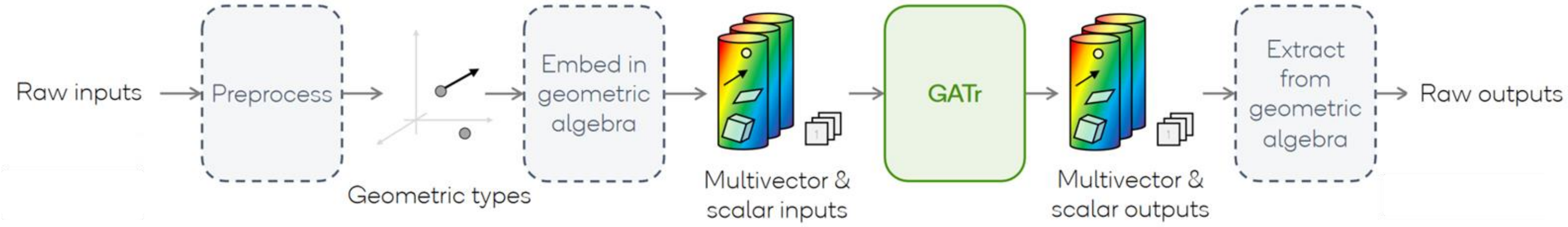
- Geometric inductive bias through geometric algebra representations
- Symmetry awareness through Lorentz equivariant layers
- Scalability and flexibility through dot-product attention

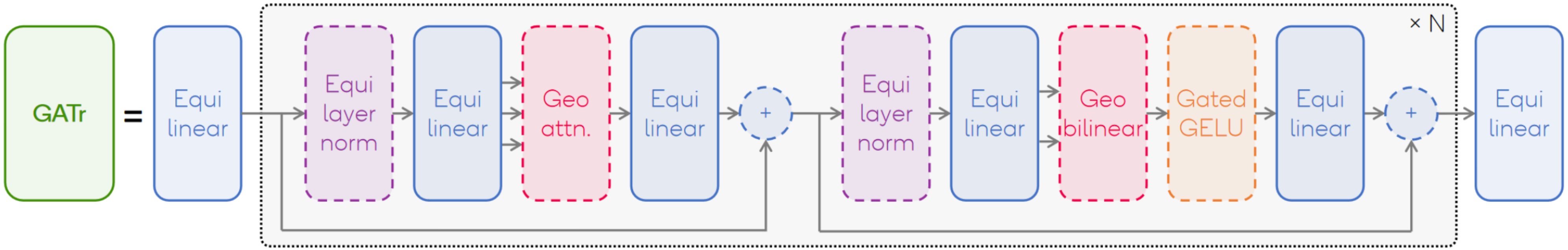
- Auxiliary scalar representations (for non-geometric data)
- Positional embeddings
- Axial attention

GEOMETRIC ALGEBRA TRANSFORMER

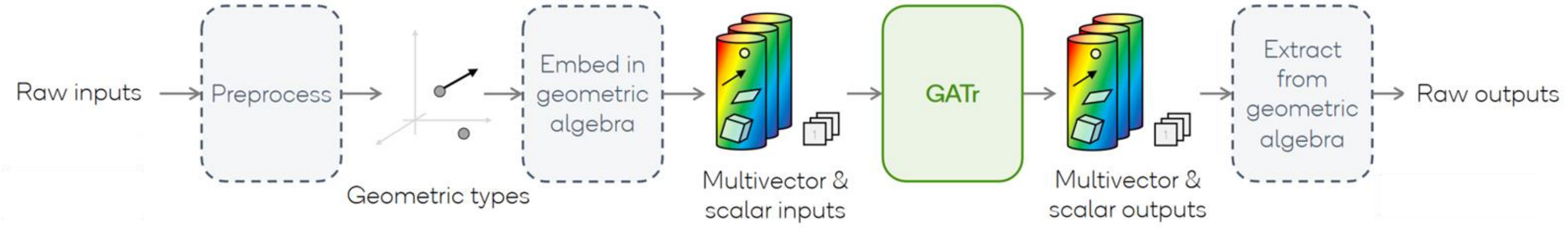


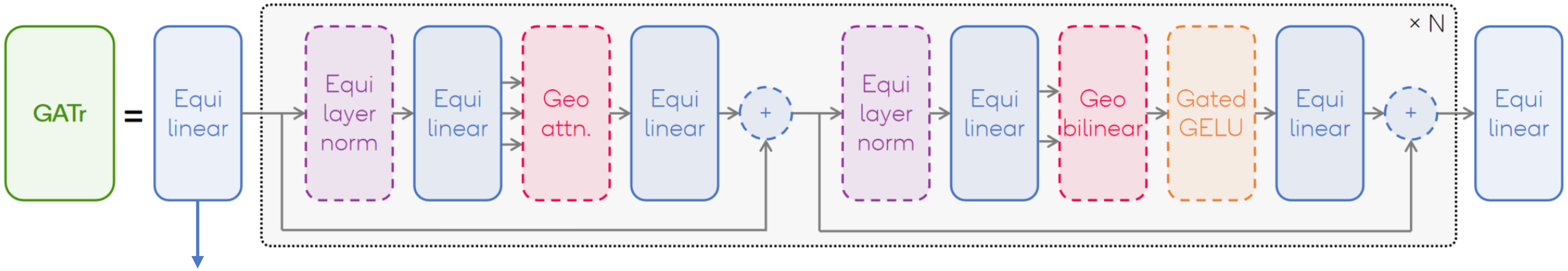
GEOMETRIC ALGEBRA TRANSFORMER





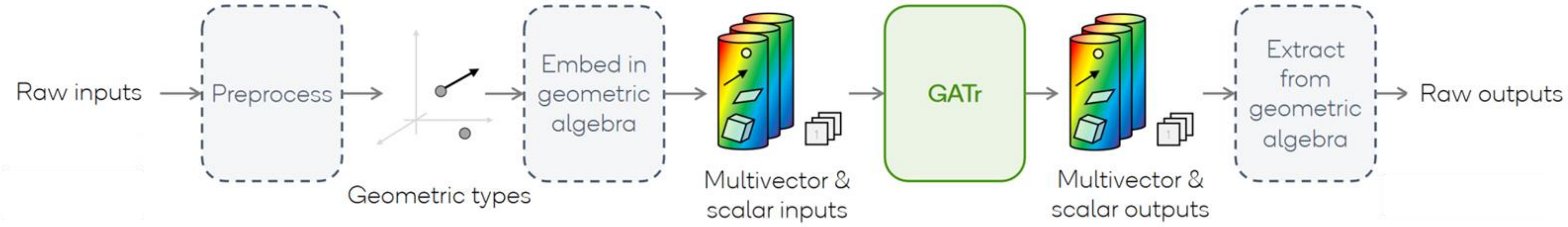
GEOMETRIC ALGEBRA TRANSFORMER



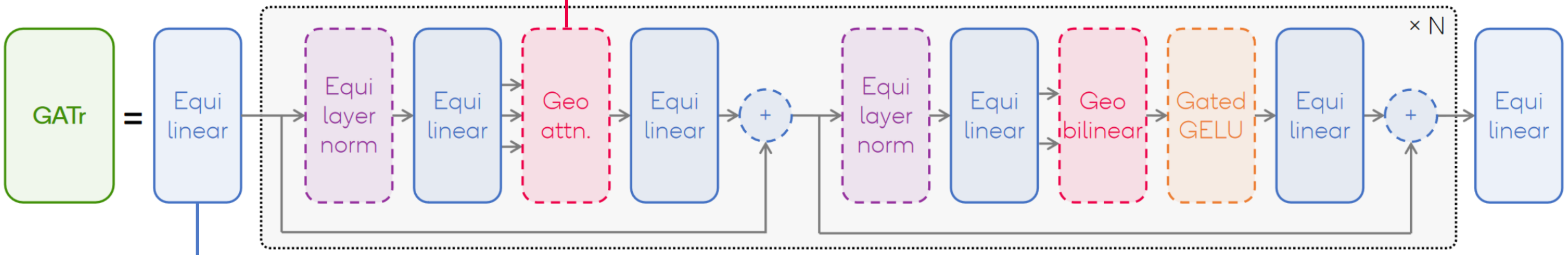


$$\phi(x) = \sum_{k=0}^{d} w_k \langle x \rangle_k$$

GEOMETRIC ALGEBRA TRANSFORMER

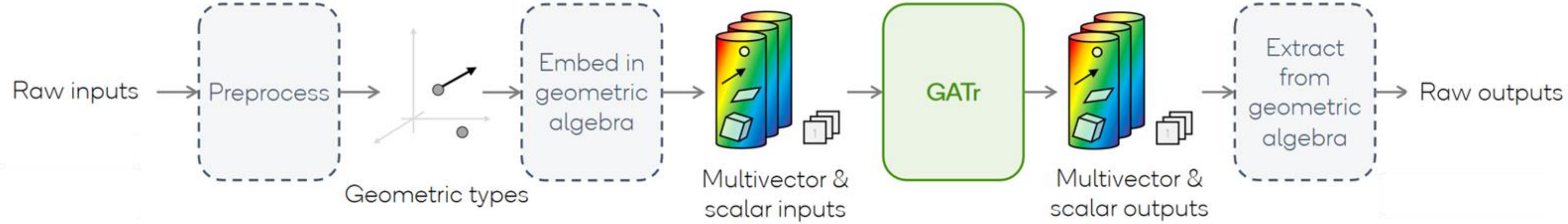


$$\text{Attention}(q,k,v)_{i'c'} = \sum_i \text{Softmax}_i \left(\frac{\sum_c \langle q_{i'c}, k_{ic} \rangle}{\sqrt{8 n_c}} \right) v_{ic'}$$

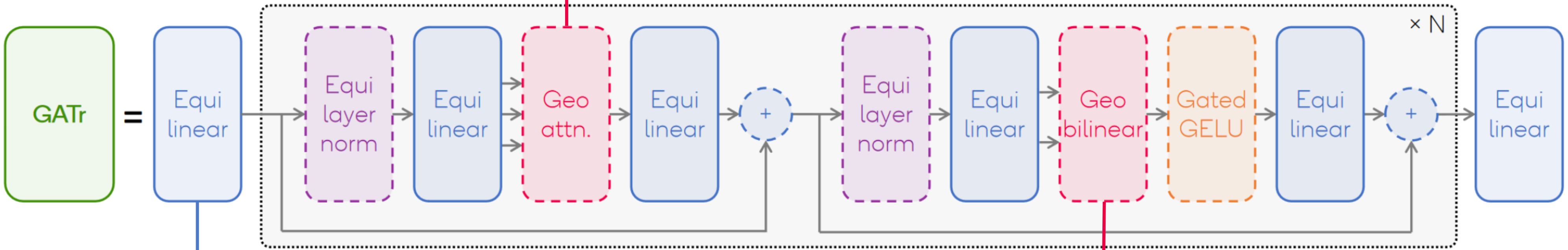


$$\phi(x) = \sum_{k=0}^{d} w_k \langle x \rangle_k$$

GEOMETRIC ALGEBRA TRANSFORMER



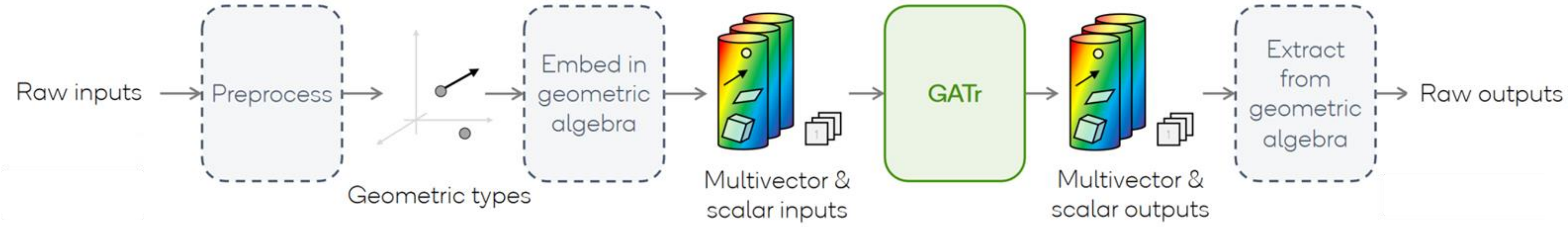
$$\text{Attention}(q,k,v)_{i'c'} = \sum_i \text{Softmax}_i \left(\frac{\sum_c \langle q_{i'c}, k_{ic} \rangle}{\sqrt{8 n_c}} \right) v_{ic'}$$



$$\phi(x) = \sum_{k=0}^{d} w_k \langle x \rangle_k$$

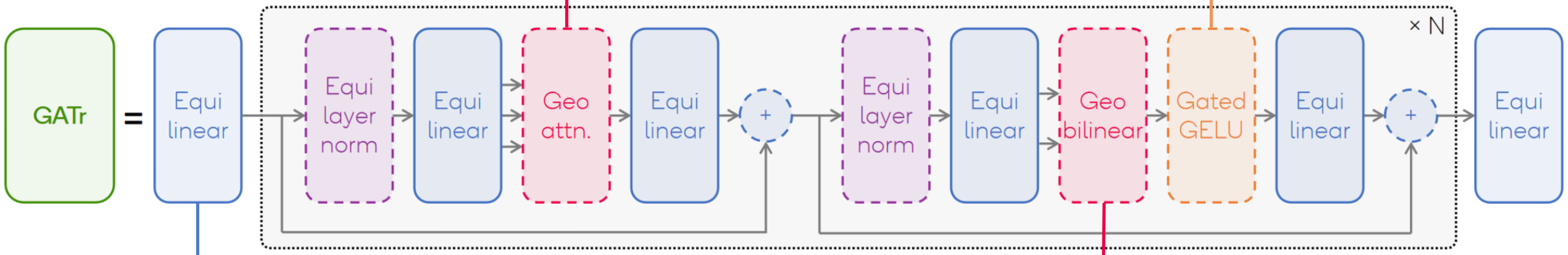
$$\text{Geometric}(x,y) = \text{Concatenate}_{\text{channels}}(xy)$$

GEOMETRIC ALGEBRA TRANSFORMER



$$\text{Attention}(q,k,v)_{i'c'} = \sum_i \text{Softmax}_i\left(\frac{\sum_c \langle q_{i'c}, k_{ic}\rangle}{\sqrt{8n_c}}\right) v_{ic'}$$

$$\text{GatedGELU}(x) = \text{GELU}(x_s)x$$



$$\phi(x) = \sum_{k=0}^{d} w_k \langle x\rangle_k$$

$$\text{Geometric}(x,y) = \text{Concatenate}_{\text{channels}}(xy)$$

AMPLITUDE SURROGATE APPLICATION

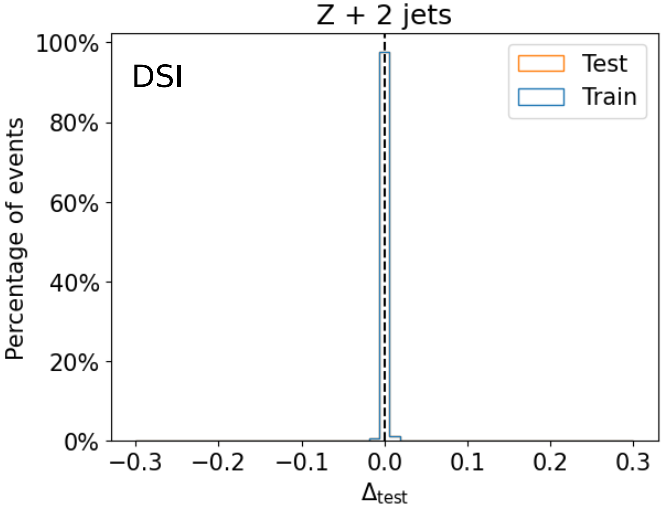
- Problem: Build an algorithm that can predict interaction amplitudes from phase space points for multiple processes

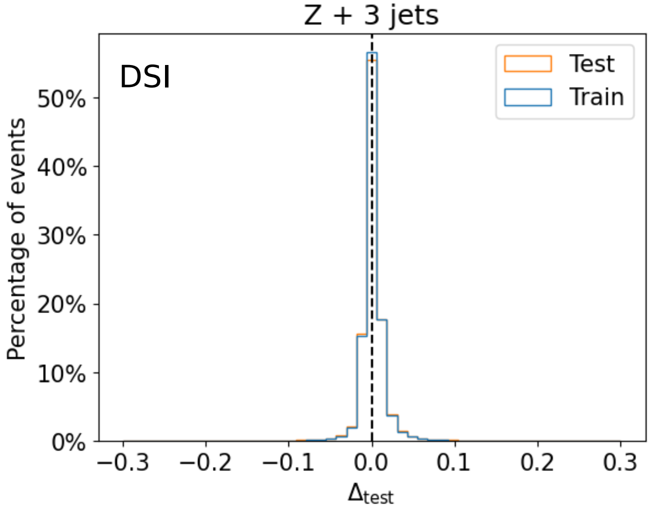
AMPLITUDE SURROGATE APPLICATION

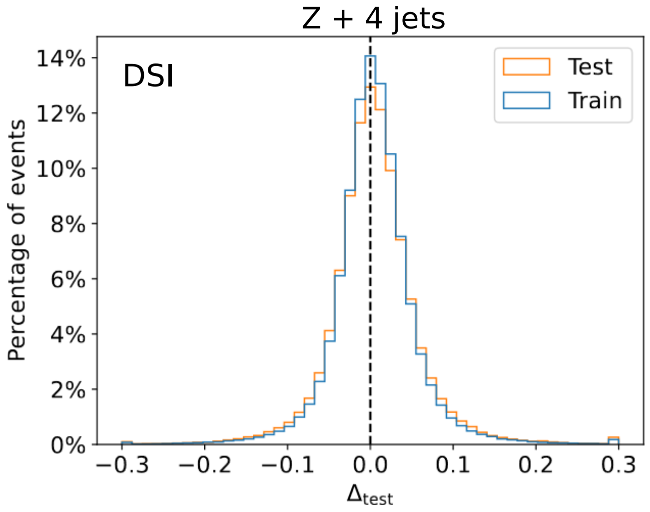
- Problem: Build an algorithm that can predict interaction amplitudes from phase space points for multiple processes
- Machine learning approaches struggle for 2 reasons:
  1. The output range covers a very large interval
  2. There exists a scaling problem. The more particles, the more complicated it is to estimate the amplitude

AMPLITUDE SURROGATE APPLICATION

- Problem: Build an algorithm that can predict interaction amplitudes from phase space points for multiple processes
- Machine learning approaches struggle for 2 reasons:
  1. The output range covers a very large interval
  2. There exists a scaling problem. The more particles, the more complicated it is to estimate the amplitude





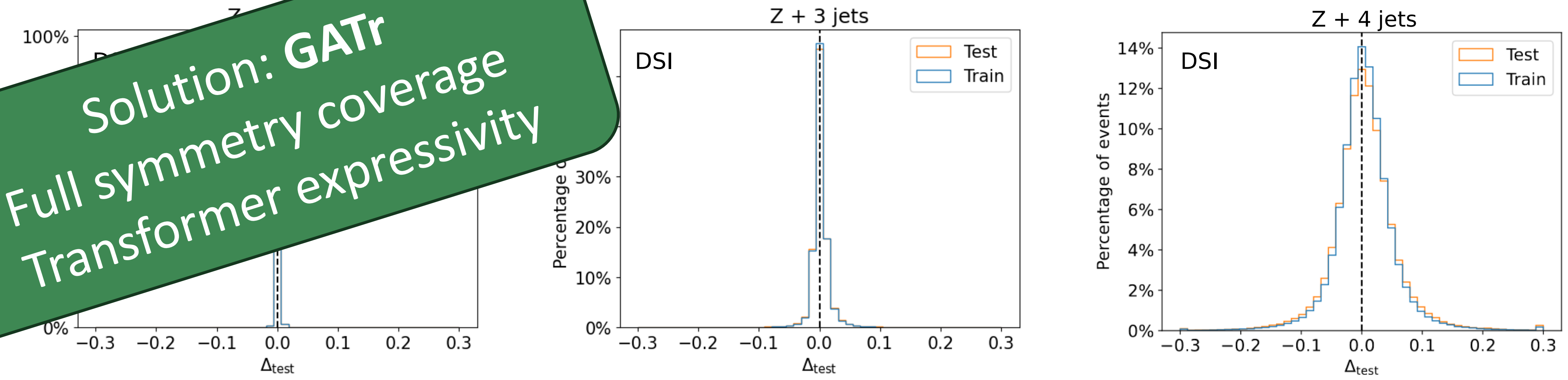


AMPLITUDE SURROGATE APPLICATION

- Problem: Build an algorithm that can predict interaction amplitudes from phase space points for multiple processes
- Machine learning approaches struggle for 2 reasons:
  1. The output range covers a very large interval
  2. There exists a scaling problem. The more particles, the more complicated it is to estimate the amplitude

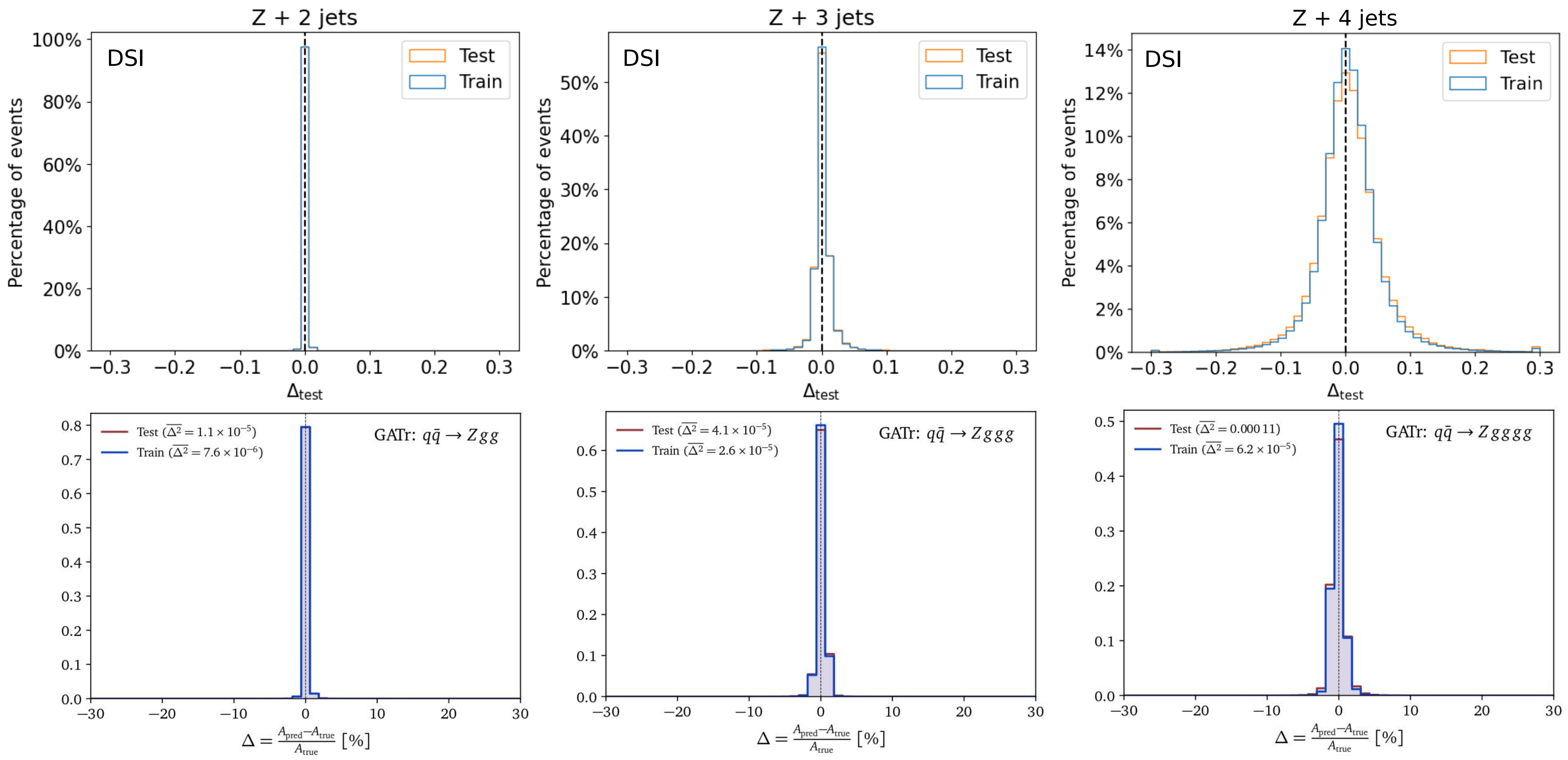
Solution: **GATr**

1. Full symmetry coverage
2. Transformer expressivity



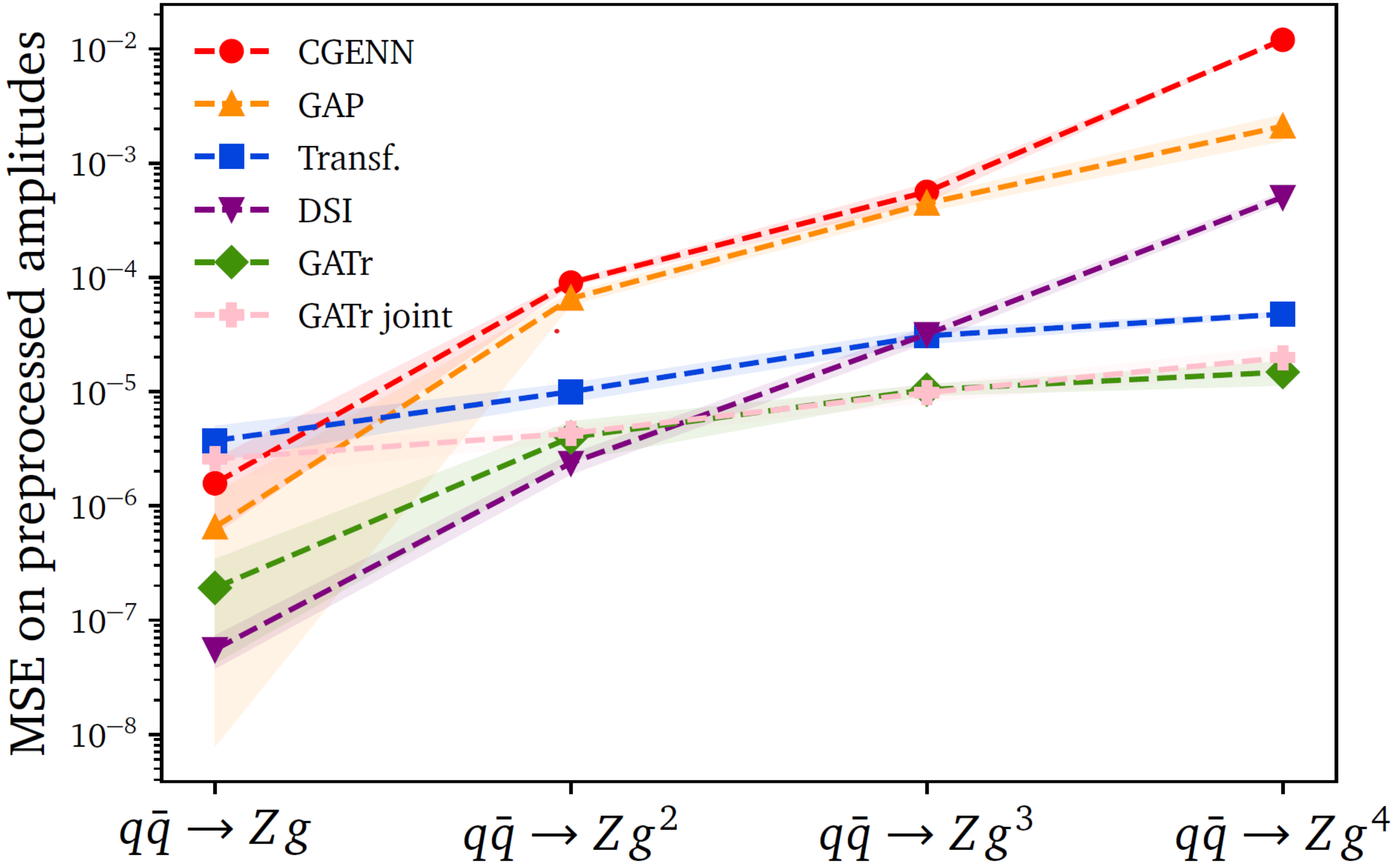
AMPLITUDE SURROGATE APPLICATION

Preliminary results!



AMPLITUDE SURROGATE APPLICATION

Preliminary results!

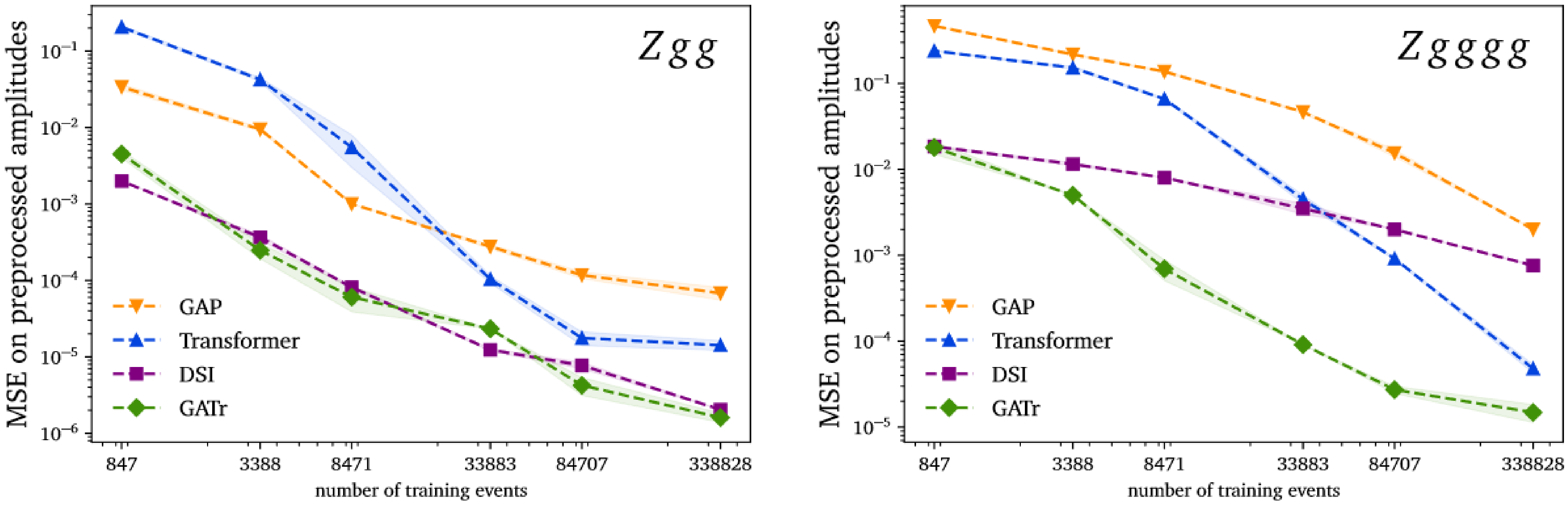


DSI: Deep Sets algorithm with momentum invariant inputs. Our main baseline.

GATr joint: GATr model trained with all data sets at the same time

AMPLITUDE SURROGATE APPLICATION

Preliminary results!



Excellent sample efficiency from **GATr**

CONCLUSIONS

- GATr is a Lorentz equivariant model that is able to produce excellent amplitude predictions for complex processes

CONCLUSIONS

- GATr is a Lorentz equivariant model that is able to produce excellent amplitude predictions for complex processes
- GATr scales much better than other geometric networks and displays great sample efficiency

CONCLUSIONS

- GATr is a Lorentz equivariant model that is able to produce excellent amplitude predictions for complex processes
- GATr scales much better than other geometric networks and displays great sample efficiency
- The joint setup allows for the same algorithm to be specialized in multiple processes without significant performance losses

CONCLUSIONS

- GATr is a Lorentz equivariant model that is able to produce excellent amplitude predictions for complex processes
- GATr scales much better than other geometric networks and displays great sample efficiency
- The joint setup allows for the same algorithm to be specialized in multiple processes without significant performance losses
- Outlook:
  - NLO amplitude regression
  - Other collider physics tasks