Top quark spin and CP studies with L-GATr

[MM, H. Bahl, A. Petitjean, T. Plehn (in preparation)]

ML4Jets 2026

Marco Menen

16.09.2026

Leibniz Universität Hannover

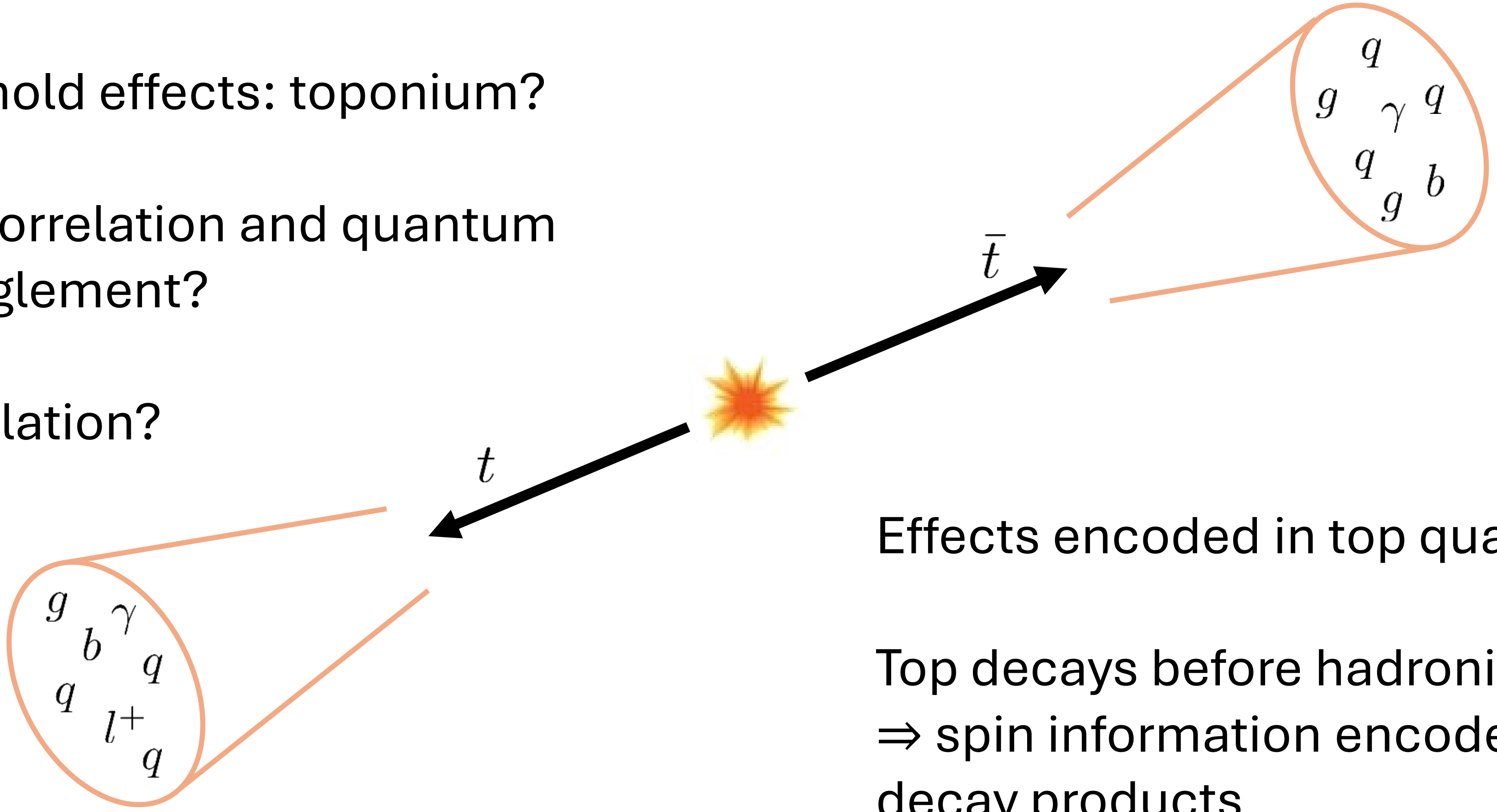
QF

technische universität dortmund

The $t\bar{t}$ production process

Many current research projects:

- Threshold effects: toponium?
- Spin correlation and quantum entanglement?
- CP violation?



Effects encoded in top quark spin

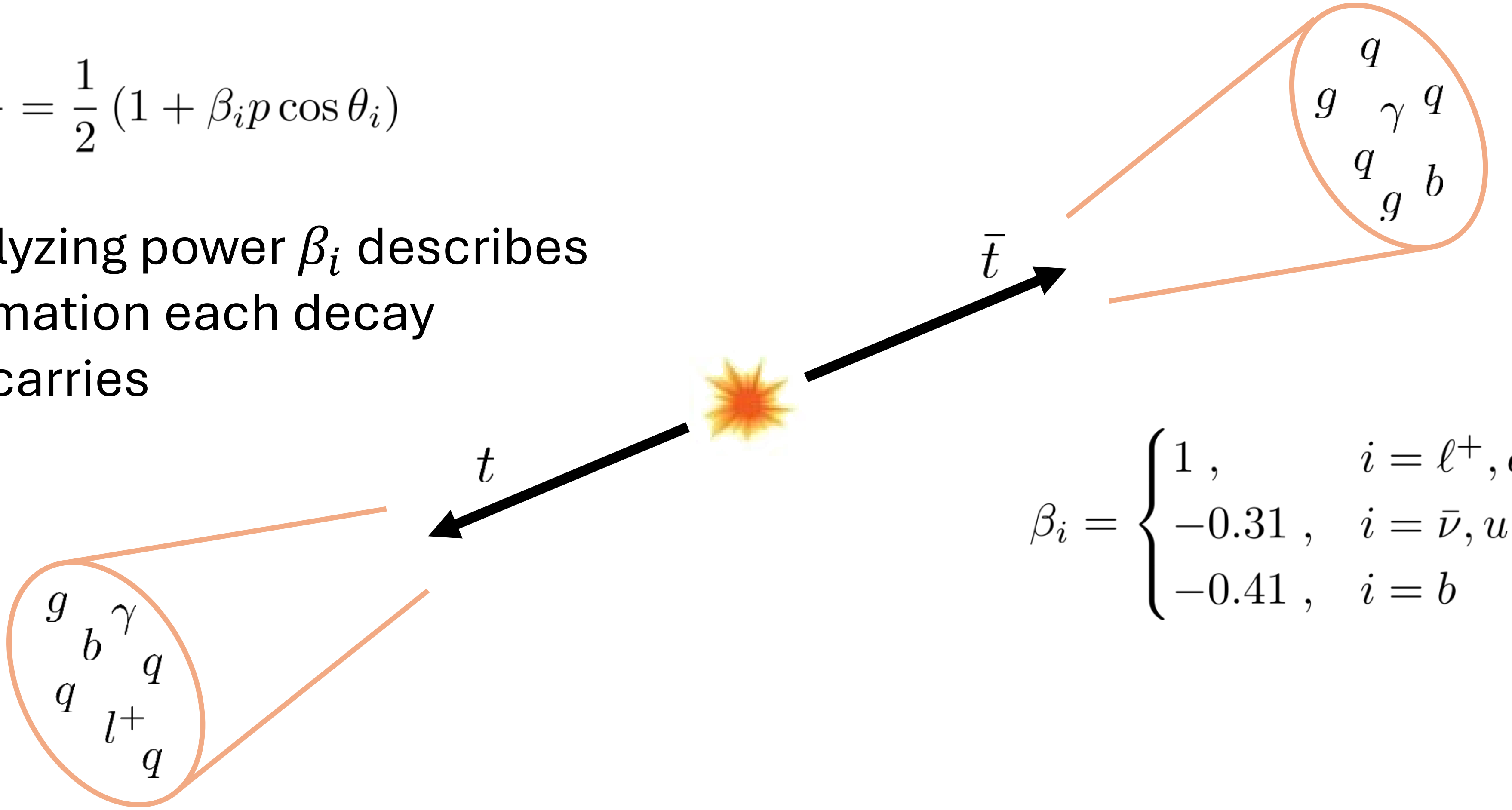
Top decays before hadronization
⇒ spin information encoded in decay products

The $t\bar{t}$ production process

Helicity angle probes top quark polarization:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}\left(1 + \beta_i p \cos\theta_i\right)$$

Spin analyzing power β_i describes the information each decay product carries



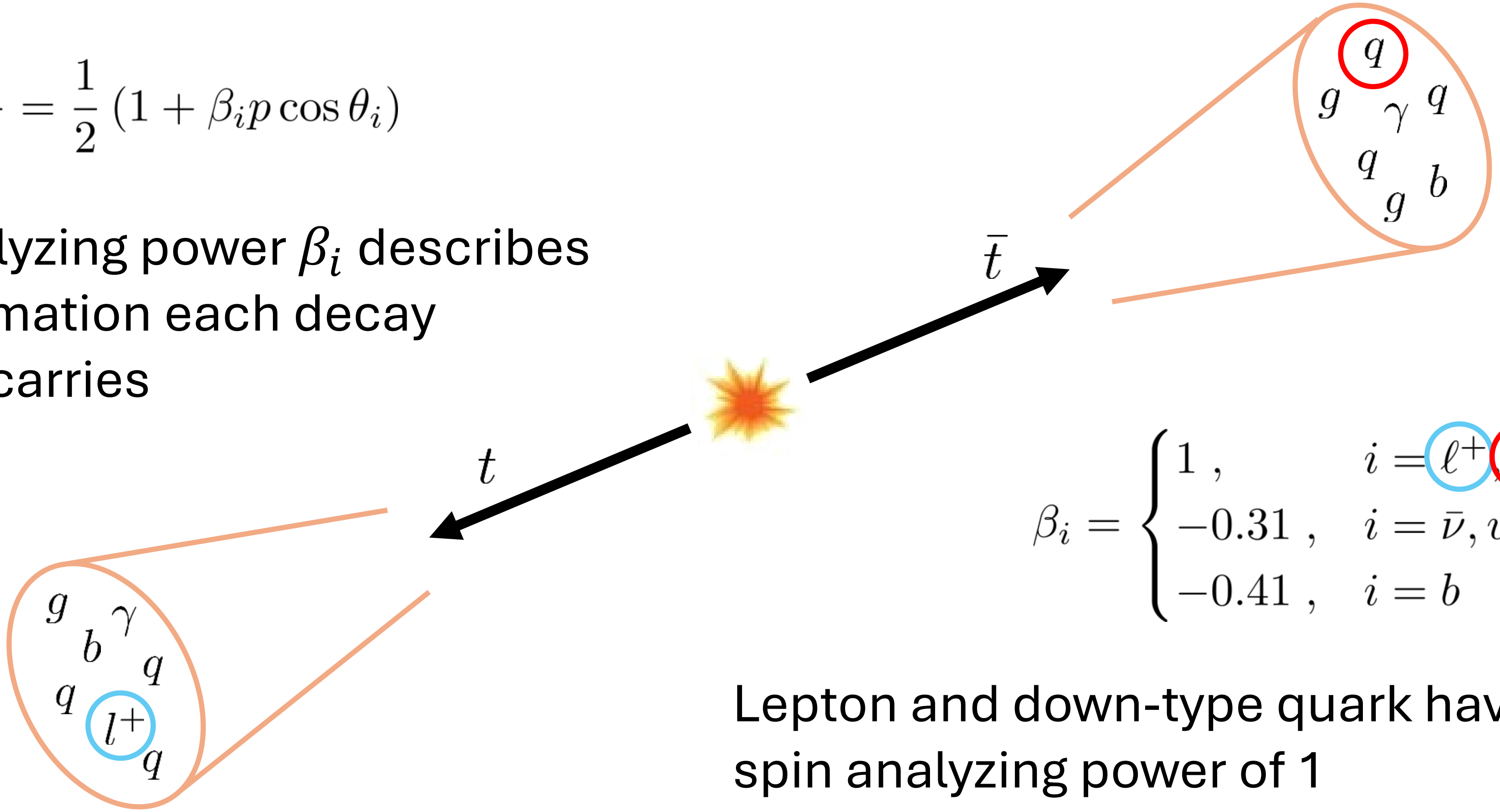
$$\beta_i = \begin{cases} 1\ , & i = \ell^+, \bar{d} \\ -0.31\ , & i = \bar{\nu}, u \\ -0.41\ , & i = b \end{cases}$$

The $t\bar{t}$ production process

Helicity angle probes top quark polarization:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}\left(1 + \beta_i p \cos\theta_i\right)$$

Spin analyzing power β_i describes the information each decay product carries



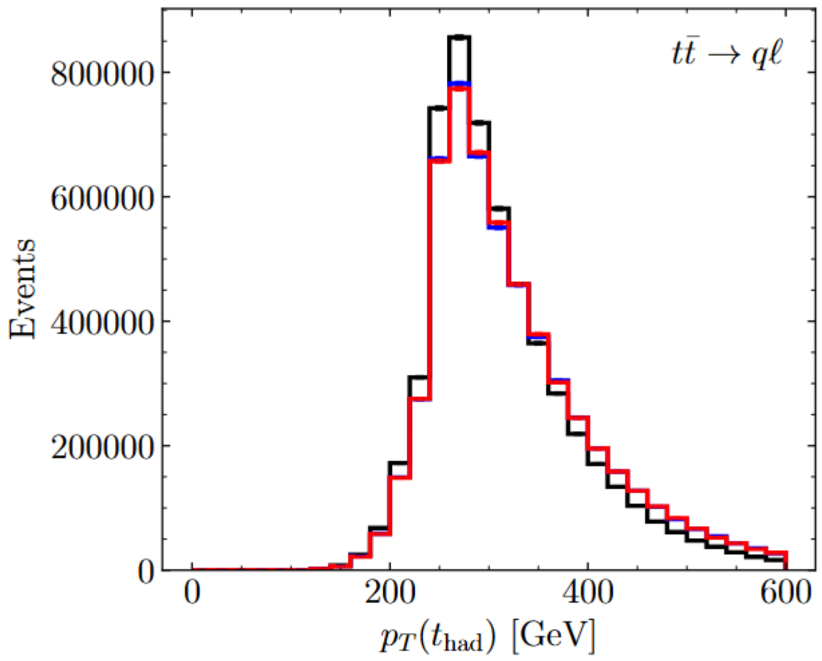
$$\beta_i = \begin{cases} 1\,, & i = \ell^+, \bar{d} \\ -0.31\,, & i = \bar{\nu}, u \\ -0.41\,, & i = b \end{cases}$$

Lepton and down-type quark have spin analyzing power of 1
⇒ Need to tag in respective decay

CP observables in $t\bar{t}$

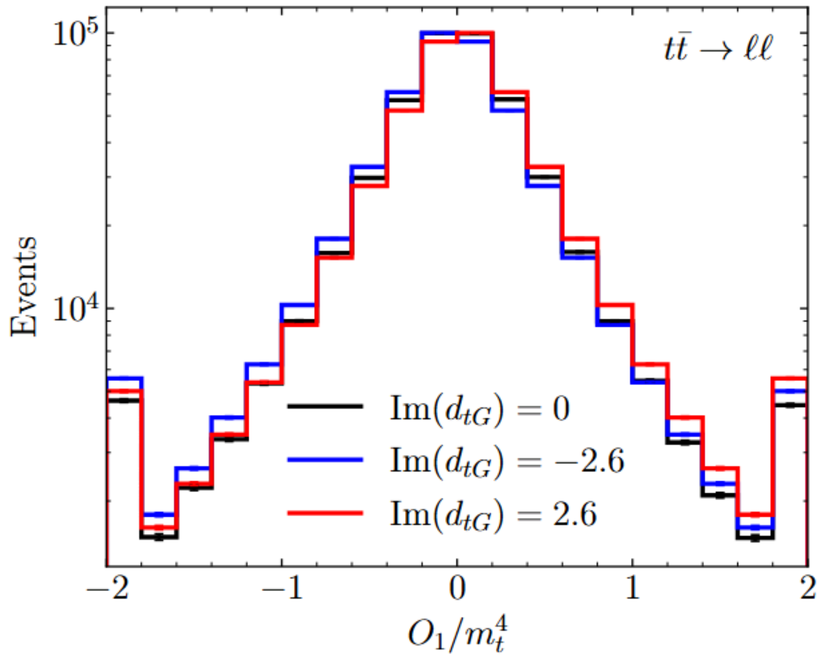
$$|\mathcal{M}|^2 = |\mathcal{M}_{\mathrm{SM}}|^2 + 2d_{tG}\mathrm{Re}\left(\mathcal{M}_{\mathrm{SM}}^\dagger \mathcal{M}_{\mathrm{BSM}}\right) + d_{tG}^2 |\mathcal{M}_{\mathrm{BSM}}|^2$$

CP-even observables:



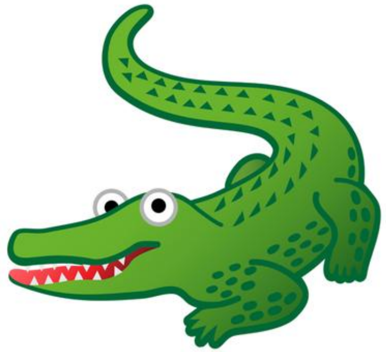
p_T is CP-sensitive but contains no information about sign of coupling

CP-odd observables:



Levi-Civita $\epsilon_{\mu\nu\rho\sigma} p_1^\mu p_2^\nu p_3^\rho p_4^\sigma$
needs top momenta and polarization

L-GATr

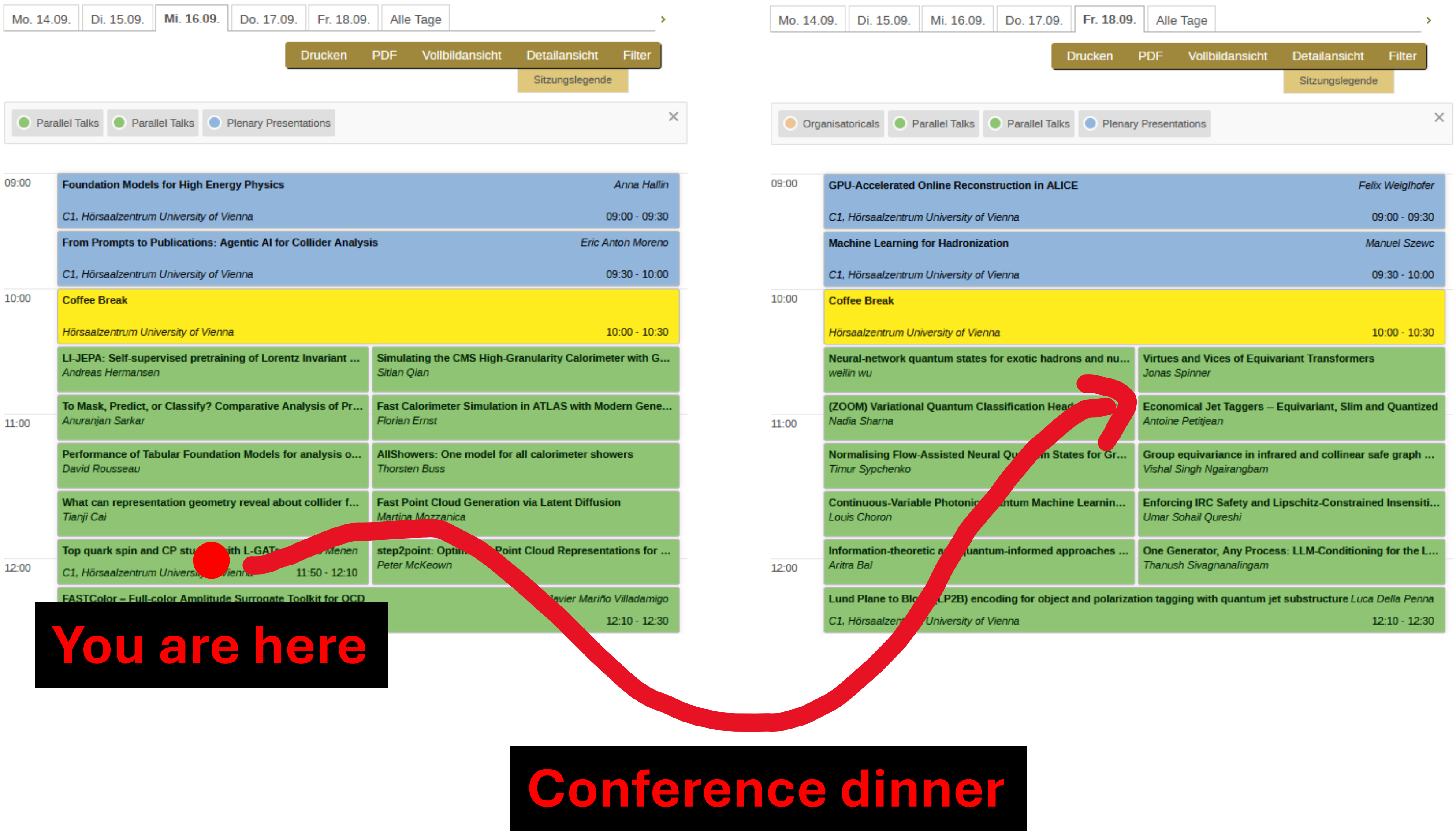


Lorentz-Equivariant
Geometric **A**lgebra
Transformer
[Bresó-Pla, Favaro, Petitjean, Spinner]

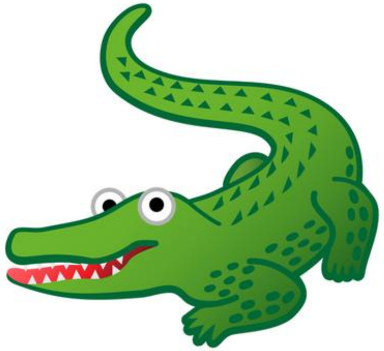
$$f(\Lambda x) = \Lambda f(x)$$

$\downarrow$

Each network layer
is equivariant



From L-GATr to L-GATr-slim



Lorentz-**E**quivariant
Geometric **A**lgebra
Transformer

[Bresó-Pla, Favaro, Petitjean, Spinner]

$$f(\Lambda x) = \Lambda f(x)$$

Each network layer
is equivariant

Input to L-GATr is multivector of spacetime algebra:

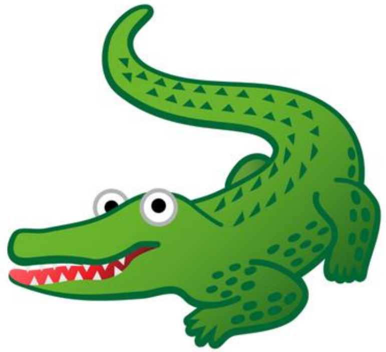
$$x^S = \mathrm{PID}$$
$$x_\mu^V = p_\mu$$
$$\vdots$$

$$\begin{pmatrix} x^S \\ x_\mu^V \\ x_{\mu\nu}^B \\ x_\mu^A \\ x^P \end{pmatrix} \in \mathbb{R}^{16}$$

[Brehmer et al. '24]

⇒ Can predict e.g. matrix element of P-odd transition via scalar + pseudoscalar outputs

From L-GATr to L-GATr-slim



Lorentz-Equivariant
Geometric **A**lgebra
Transformer

[Bresó-Pla, Favaro, Petitjean, Spinner]

$$f(\Lambda x) = \Lambda f(x)$$

Each network layer
is equivariant

Input to L-GATr is multivector of spacetime algebra:

$$\begin{pmatrix} x^S \\ x^V_\mu \\ x^B_{\mu\nu} \\ x^A_\mu \\ x^P \end{pmatrix} \in \mathbb{R}^{16}$$

$$x^S = \mathrm{PID}$$
$$x^V_\mu = p_\mu$$
$$\vdots$$

[Brehmer et al. '24]

⇒ Can predict e.g. matrix element of P-odd transition via scalar + pseudoscalar outputs

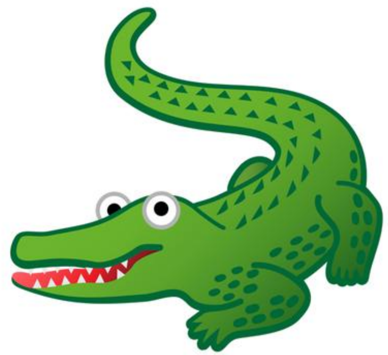
L-GATr-slim:
Reduce to scalar + vector

$$\begin{pmatrix} x^S \\ x^V_\mu \end{pmatrix} \in \mathbb{R}^5$$

[Petitjean et al. '25]

⇒ Sufficient for most applications, but much less computationally expensive

Recovering pseudoscalars in L-GATr-slim



Lorentz-Equivariant
Geometric **A**lgebra
Transformer
[Bresó-Pla, Favaro, Petitjean, Spinner]

$$f(\Lambda x) = \Lambda f(x)$$

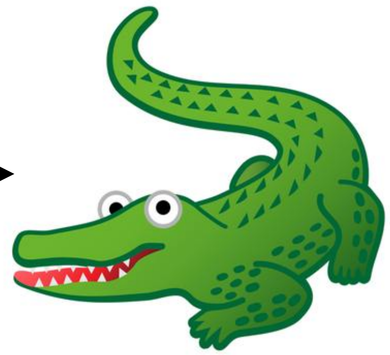
$\downarrow$

Each network layer
is equivariant

We are interested in CP-odd observables:
Give pseudoscalar back to L-GATr-slim

⇒ contains all information needed for $t\bar{t}$ spin and CP studies with a fraction of the cost for the full L-GATr

$$\begin{pmatrix} x^S \\ x^V_\mu \\ x^P \end{pmatrix} \in \mathbb{R}^6 \longrightarrow$$

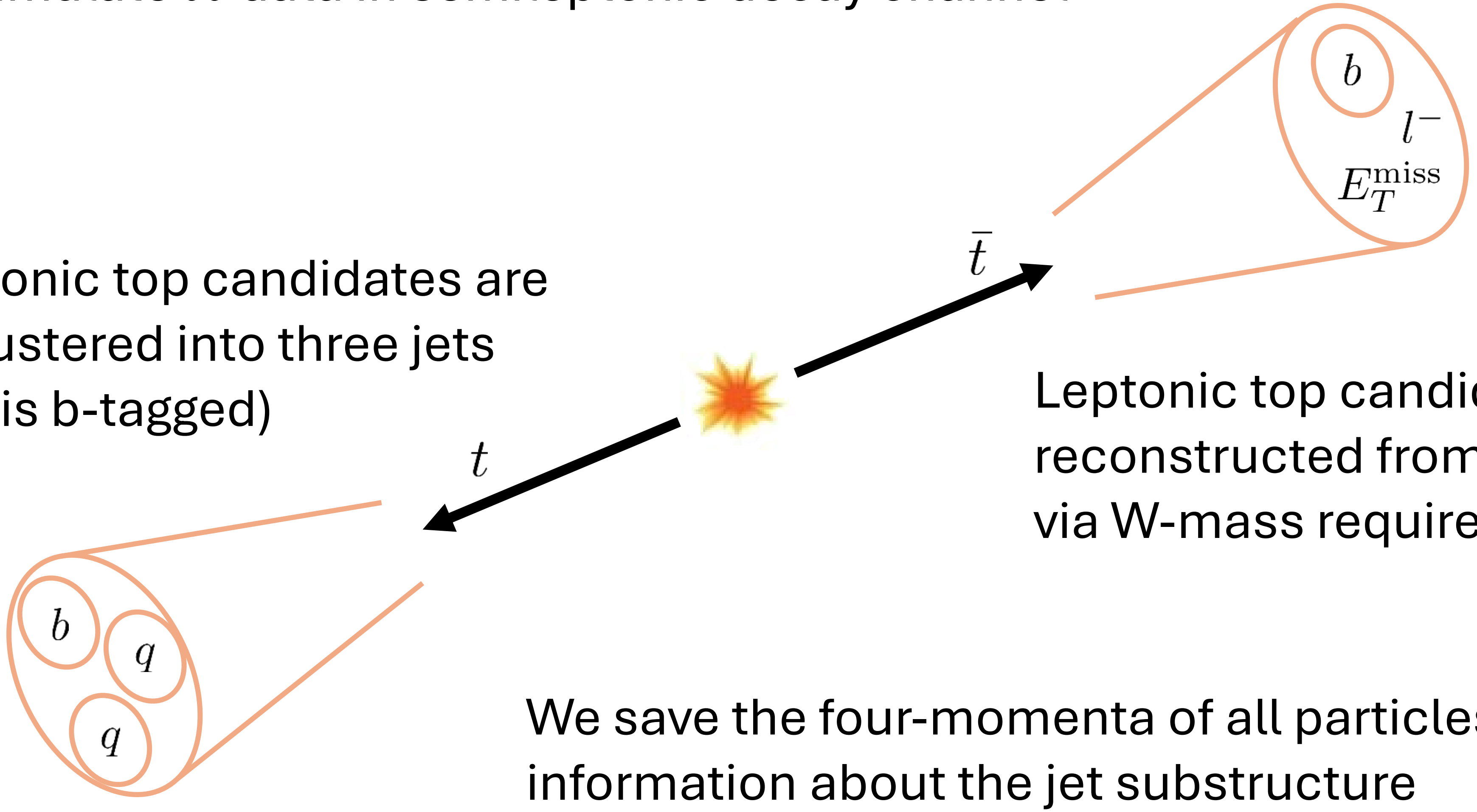


$$\longrightarrow \begin{pmatrix} x^S \\ x^V_\mu \\ x^P \end{pmatrix} \in \mathbb{R}^6$$

⇒ We can access the scalar and pseudoscalar information independently

Input data

We simulate $t\bar{t}$ data in semileptonic decay channel



Hadronic top candidates are declustered into three jets (one is b-tagged)

Leptonic top candidates are reconstructed from E_T^{miss} via W-mass requirement

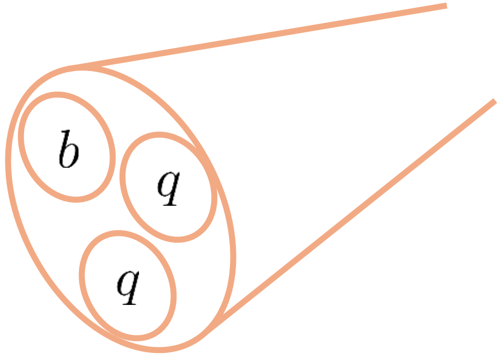
We save the four-momenta of all particles and all information about the jet substructure

Task 1: Hadronic spin reconstruction

Break down data into binary classification problem:
Identifying the down-type quarks from the two light quark jets

Hadronic spin proxy vector from weighted average:

$$q_d = p(d \to q_{\text{hard}}|\mathcal{O})\hat{q}_{\text{hard}} + p(d \to q_{\text{soft}}|\mathcal{O})\hat{q}_{\text{soft}}$$



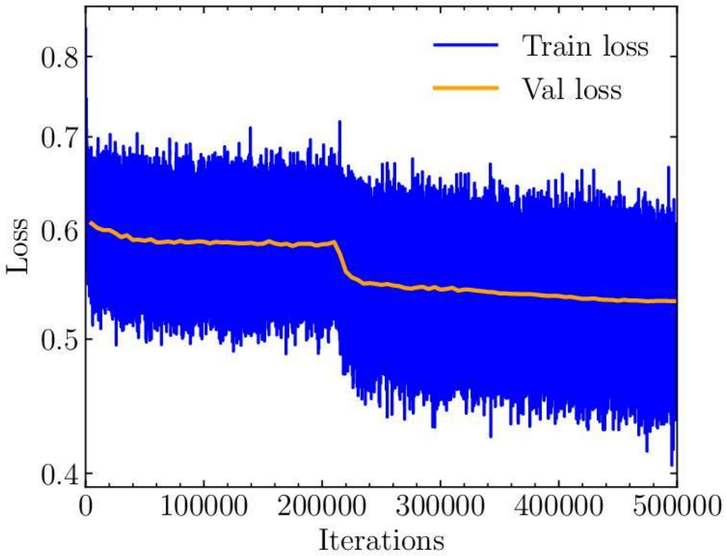
Spin analyzing power: $\beta_d = \langle |q_d| \rangle$

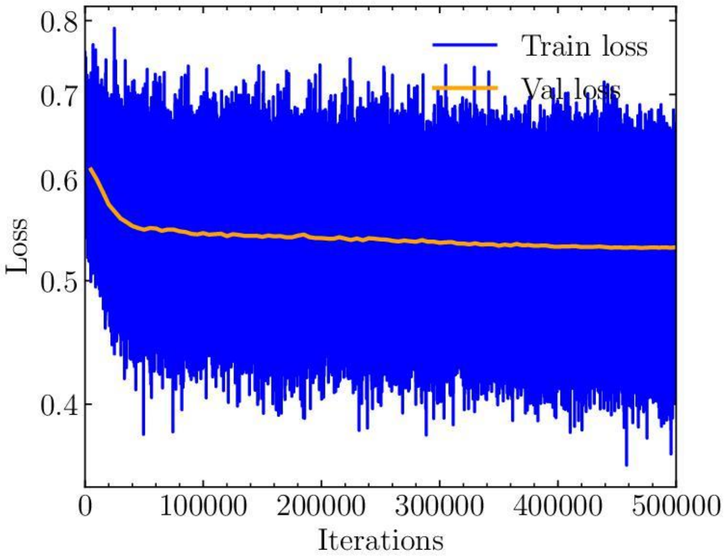
Previous approaches:

- [B. Tweedie '14] based on angular observables $\beta_d = 0.64$
- [Z. Dong et al. '24] based on ParticleNet $\beta_d = 0.68$

[Qu, Gouskos '19]

Task 1: Hadronic spin reconstruction



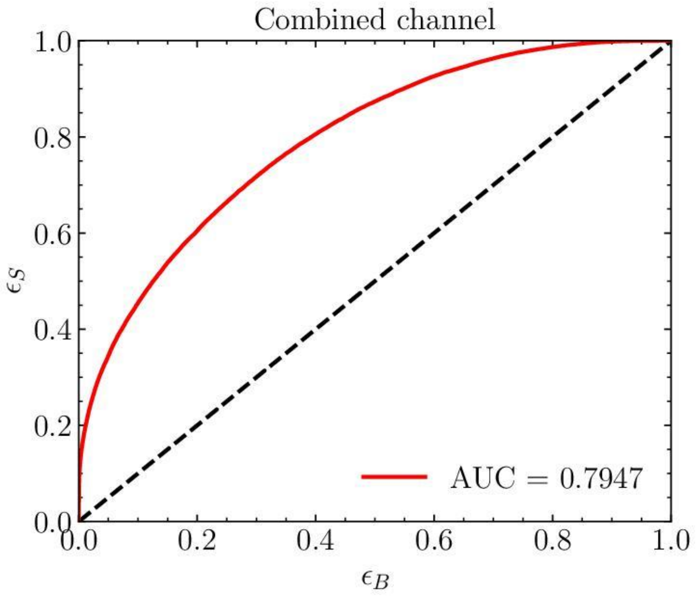


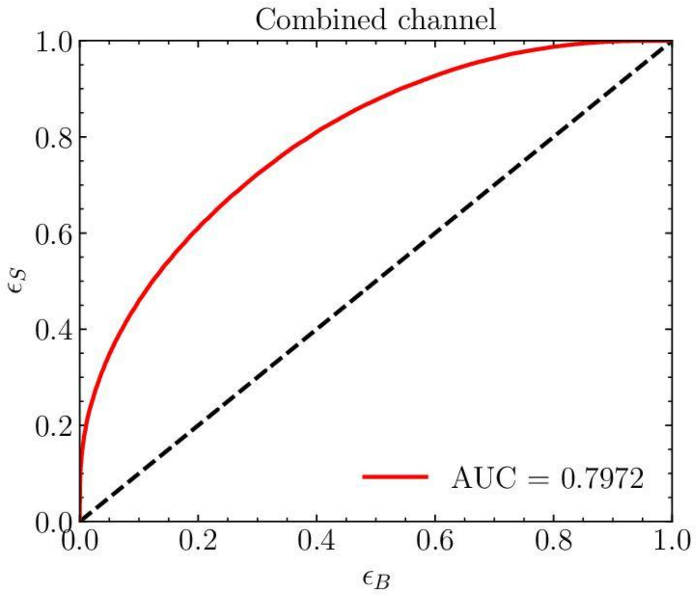
L-GATr vs. L-GATr-slim

Both models show similar performance

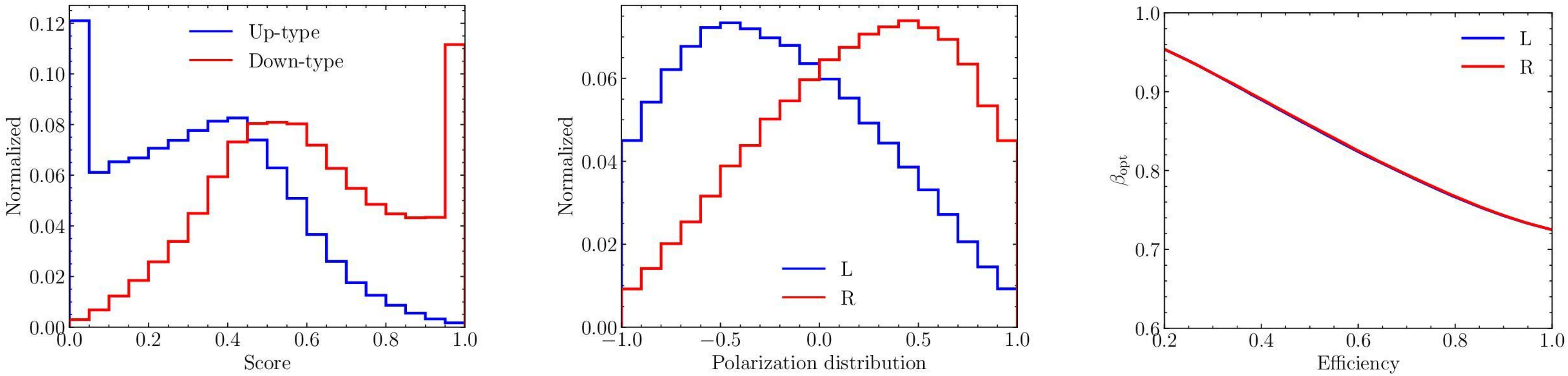
L-GATr-slim learns more smoothly

⇒ Only L-GATr-slim results shown from now on





Task 1: Hadronic spin reconstruction



Use score to build hadronic spin proxy $q_d = p(d \to q_{\text{hard}}|\mathcal{O})\hat{q}_{\text{hard}} + p(d \to q_{\text{soft}}|\mathcal{O})\hat{q}_{\text{soft}}$

$\Rightarrow$ Improve spin analyzing power to $\beta_d = 0.72$ ($\beta_d = 0.95$ for 20% best events)

Task 2: CP classification

$$|\mathcal{M}|^2 = |\mathcal{M}_{\mathrm{SM}}|^2 + 2d_{tG}\mathrm{Re}\left(\mathcal{M}^\dagger_{\mathrm{SM}}\mathcal{M}_{\mathrm{BSM}}\right) + d^2_{tG}\left|\mathcal{M}_{\mathrm{BSM}}\right|^2$$

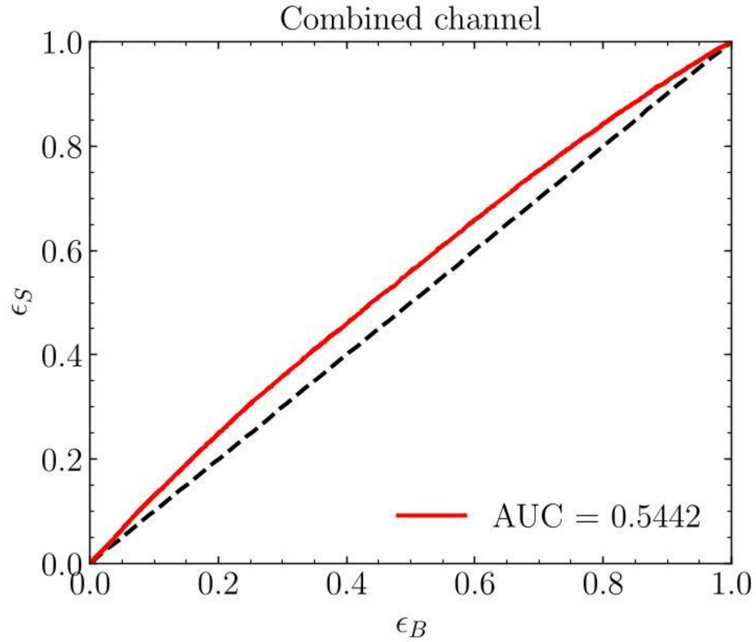
- Training to differentiate events with positive and negative interference produces an optimal CP-odd observable

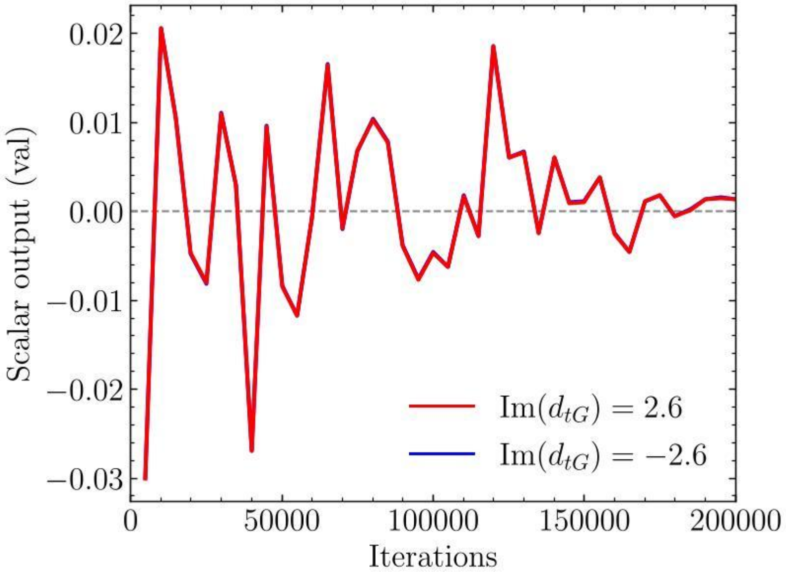
- Training to differentiate the SM from a non-zero d_{tG} contains both CP-even and CP-odd information

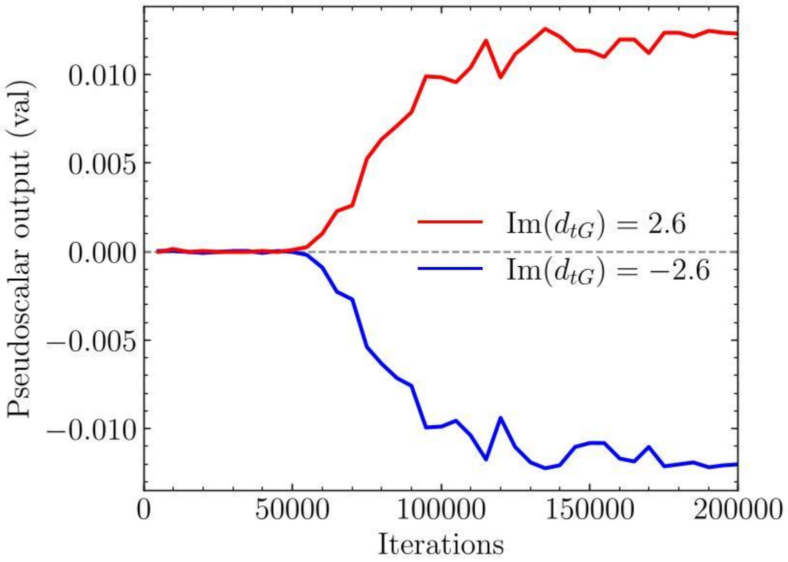
- Training is performed with full down quark information

- Evaluation is performed with the output from the down-tagging network

Task 2: CP classification: neg. vs. pos.





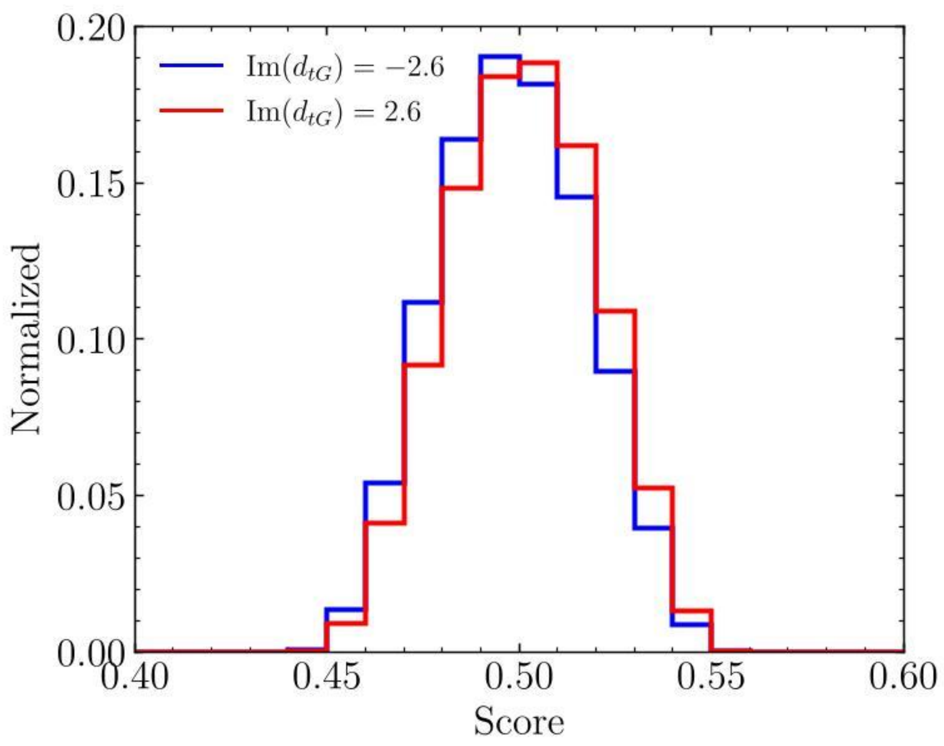


For the positive vs. negative training, all the separation power sits in the pseudoscalar channel

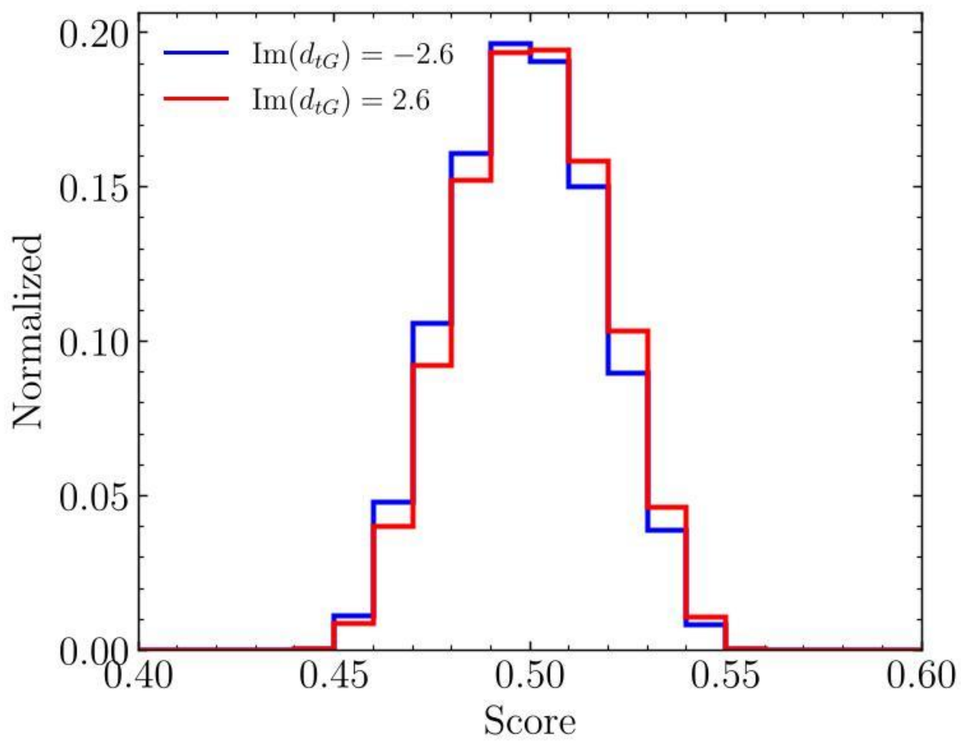
⇒ L-GATr-slim learns in the pseudoscalar channel and drives the scalar output to zero

Task 2: CP classification: neg. vs. pos.

Dileptonic channel (full spin information)

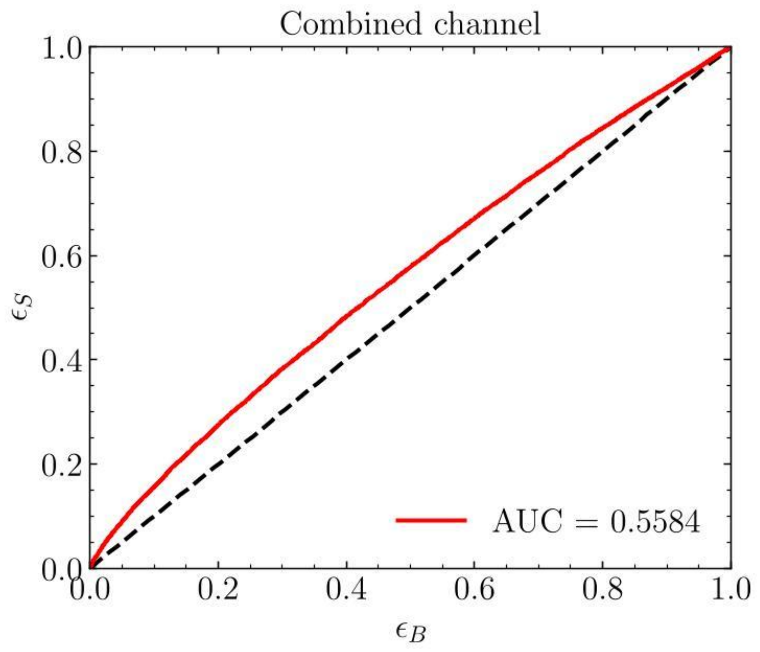


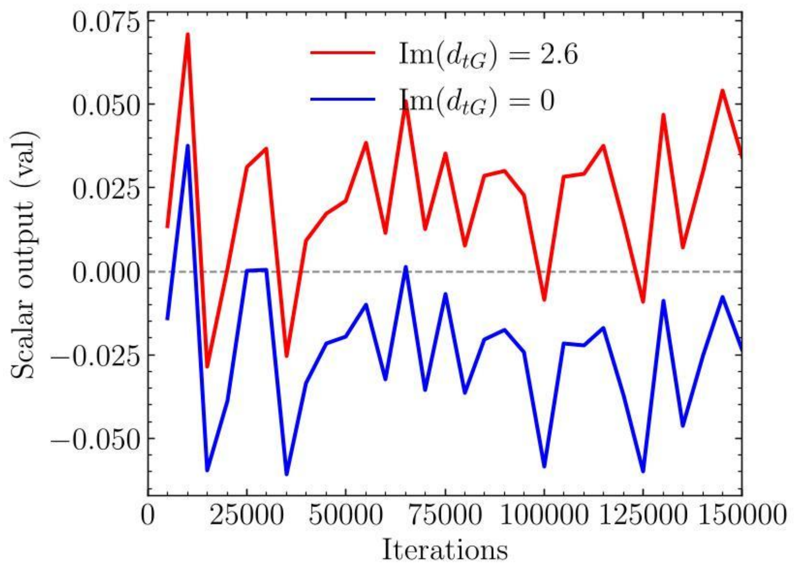
Semileptonic channel (down-tagging information)



Clear separation also in the semileptonic channel
Experimental sensitivity to be studied (number of events, backgrounds)

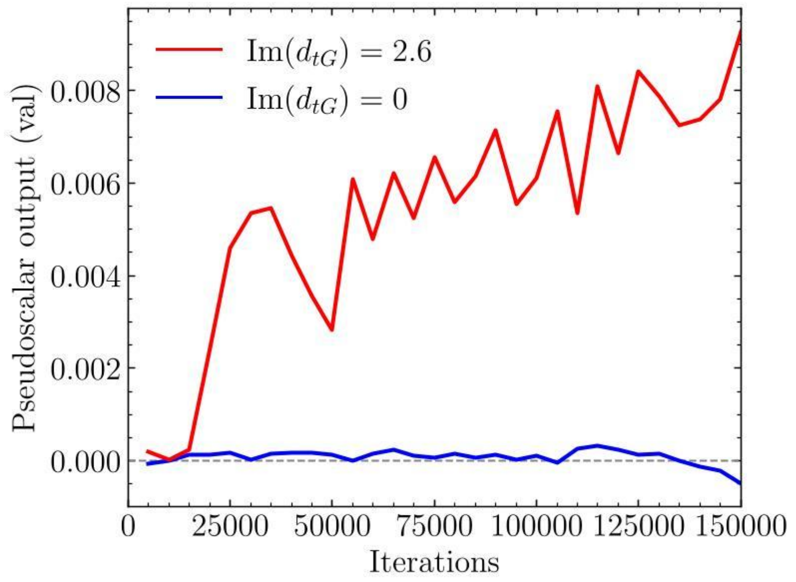
Task 2: CP classification: SM vs. pos.





Slightly better AUC due to the additional scalar separation power

The SM data stays at a pseudoscalar output of zero



Conclusions & Outlook

- L-GATr reaches state-of-the-art performance for hadronic top quark spin reconstruction via light quark tagging

- L-GATr-slim performs at least as well as the full architecture

- The Lorentz-equivariant design offers a clear separation of CP-violating data into its CP-even and CP-odd components for increased interpretability

Conclusions & Outlook

- L-GATr reaches state-of-the-art performance for hadronic top quark spin reconstruction via light quark tagging

- L-GATr-slim performs at least as well as the full architecture

- The Lorentz-equivariant design offers a clear separation of CP-violating data into its CP-even and CP-odd components for increased interpretability

- Outlook: Lunch