Topology-Agnostic Event Reconstruction with Hierarchical Graph Neural Networks

Nathalie Soybelman, Nilotpal Kakati, Eilam Gross
Weizmann Institute of Science
Rehovot, Israel
{nathalie.soybelman, nilotpal.kakati, eilam.gross}@weizmann.ac.il

Francesco A. di Bello*
University of Genova
Genova, Italy
francesco.armando.di.bello@cern.ch

Abstract

We address the task of reconstructing unknown hierarchical structures from unordered sets of observations. To this end, we propose a hierarchical graph neural network that assembles such structures without relying on any topological priors. The model operates in stages: it first identifies low-level components and then infers higher-level assemblies. It also supports simultaneous set-level classification. Evaluated on a challenging particle physics benchmark, our method is the first in the field to be fully topology-agnostic, yet it matches the efficiency and achieves higher reconstruction purity than models constrained by predefined topologies.

1 Introduction

Solving inverse problems is a central challenge in high-energy physics. At the Large Hadron Collider [1], particle collisions produce unstable fundamental particles that decay almost instantly into secondary particles, which are captured by the detectors. To interpret the recorded data, an event reconstruction process consisting of two main stages is implemented: a low-level stage that infers final, stable particles from the raw detector response, followed by a high-level stage that reconstructs the initial unstable particles from these newly-formed objects.

While the low-level stage is itself a significant inverse problem where a variety of deep learning techniques have been successfully applied [2, 3, 4, 5], this report focuses on the subsequent high-level stage of reconstruction. As the number of unstable particles in an event increases, the combinatorial complexity of correctly assigning their decay products grows rapidly, rendering classical algorithms computationally intractable [6, 7, 8, 9, 10, 11]. Modern machine learning methods have been developed to tackle this combinatorial task, but most existing approaches hard-code a fixed, expected event topology into their architecture [6, 7, 12]. This reliance on a known structure, while effective in specific cases, fundamentally limits their generalizability to realistic scenarios where decay topologies may be partially observed, variable, or entirely unknown.

In this work, we propose a topology-agnostic approach based on a hierarchical graph neural network that learns to reconstruct decay chains without relying on predefined topologies. We demonstrate our model on a well-defined benchmark, the complex $t\bar{t}H$ process. This process involves the decay of top quarks to W-bosons and b-quarks; the W-bosons in turn decay either leptonically to a lepton and an

*Now at University of Pisa

Machine Learning and the Physical Sciences Workshop, NeurIPS 2025.

undetectable neutrino ($W \to l\nu$) or hadronically to a pair of quarks ($W \to qq'$); and the Higgs boson decays to a pair of b-quarks ($H \to b\bar{b}$). All quarks form jets during their decays, which are observed as cone-like objects. Crucially, and in contrast to other works, our approach does not require this specific decay process to be defined in advance. In that sense, our model is truly topology-agnostic. It is guided only by the typically underlying structure of sequential two-body decays, offering a flexible and scalable framework for event reconstruction in realistic scenarios where decay topologies are not known in advance.

2 Related Work

Deep learning approaches to event reconstruction in particle physics have largely been built on Graph Neural Networks (GNNs) and Transformer architectures. These models are well-suited to the task, as they naturally accommodate the permutation-invariant nature of particle sets and capture their underlying relational structure.

Graph-based models, such as HyPER [12] and Topograph [6], formulate event reconstruction as a graph-learning task where particles are nodes and the decay chain is inferred through learned edges. Despite differing in their methods for constructing hierarchical structures, both models are tailored to specific, fixed event topologies (e.g., top quark decay), which are hard-coded into their architectures and training objectives.

Transformer-based models, such as SPANet [7, 8, 13, 14], treat reconstruction as a set-assignment problem. They employ a symmetry-preserving attention mechanism to partition the input particles into groups, each corresponding to a parent particle. While the model architecture is flexible in principle, it still requires prior knowledge of the event topology, specifically the number and type of parent particles to reconstruct.

Hierarchical Graph Neural Networks are a natural fit for our task, reflecting the inherently hierarchical structure of particle decays. However, adapting standard approaches [15, 16] is non-trivial, as they typically generate coarse-grained graph representations via unsupervised or predefined pooling. In event reconstruction, by contrast, intermediate hierarchy levels correspond to physically meaningful particles (e.g., W or Higgs bosons), necessitating a supervised, object-centric construction of the hierarchy.

3 Model and Training

We propose TIGER (Topology-Independent Graph-based Event Reconstruction), a hierarchical graph neural network that dynamically assembles decay chains in two stages, guided solely by the physical

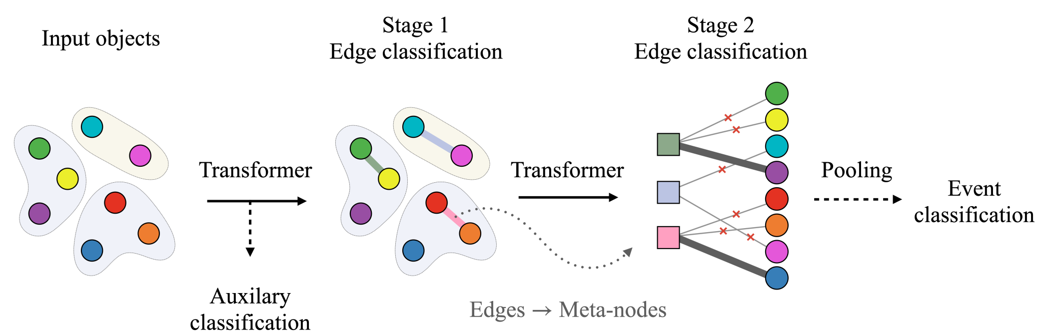

Figure 1: **Overview of the hierarchical graph neural network architecture.** The input set of objects is first encoded using a Transformer, followed by an auxiliary classification applied to the updated representations. Hierarchical reconstruction proceeds in two stages: in the first, edges in the fully connected graph corresponding to low-level components are identified; high-scoring pairs are promoted to meta-nodes. In the second stage, these meta-nodes are paired with original object nodes to form higher-level structures. Connections between meta-nodes and their own constituents are forbidden. An optional pooling step produces a global representation for event-level classification. Shaded bubbles indicate the ground-truth assignments.

principle of sequential two-body decays. The architecture supports simultaneous event reconstruction and classification. An overview is shown in Figure 1.

Encoding The model takes as input an unordered set of objects, primarily jets (collimated sprays of particles), leptons (e.g. electrons), and missing energy (indicating undetectable particles), treating them as nodes in a graph. The kinematic properties of jets and leptons are described by their momentum, direction, and mass, along with binary tags identifying the lepton type or whether a jet is a b-jet. Missing energy is characterised by its magnitude and direction. Each object is embedded into the same vector space using class-specific multilayer perceptrons (MLPs). A Diffusion Transformer (DiT) [17] then updates each object's representation based on the other objects in the event.

Auxiliary task To enhance the discriminative power of the representations, we include an auxiliary classification task that predicts the likely parent particle for each object based on its final embedding.

Hierarchical reconstruction The core of our method is a two-stage, bottom-up assembly process. In the first stage, a fully connected graph is constructed over all encoded objects, and a multi-class edge classification is performed to identify candidate intermediate particles (e.g., W and Higgs bosons). The resulting edge probabilities define potential pairings, which are instantiated as new *meta-nodes*. In the second stage, a separate DiT updates the representations of both the original object nodes and the meta-nodes, enabling information exchange. A bipartite graph is then formed between these two sets, and a final edge classification step predicts higher-level combinations–e.g., matching a W meta-node with a jet to form a top quark. To prevent error propagation and ensure training stability, our method diverges from classical hierarchical approaches. We explicitly retain all original nodes in the second stage, including those incorporated into meta-nodes.

Event classification To distinguish between signal and background events, the architecture performs event-level classification by pooling the final representations of all nodes and meta-nodes into a global event embedding, which is passed to a binary classifier.

The model is trained end-to-end with the loss being a weighted sum of the cross-entropy losses from the two stages of the hierarchical graph-learning $\mathcal{L}_{\text{stage}\,\{1,2\}}$, the auxiliary task following the encoding step $\mathcal{L}_{\text{aux}}$, and the event-level classification $\mathcal{L}_{\text{event}}$ with weights $\lambda_{\{1,2\}}$.

$$\mathcal{L} = (1 - 0.5\alpha)\mathcal{L}_{\text{stage 1}} + 0.5\alpha\mathcal{L}_{\text{stage 2}} + \lambda_1\mathcal{L}_{\text{aux}} + \lambda_2\mathcal{L}_{\text{event}} \tag{1}$$

The weights were optimised for each training to ensure all components contribute equally. For the pure $t\bar{t}H$ trainings we chose $\lambda_1 = \lambda_2 = 1/50$ and for the combined dataset we changed $\lambda_2 = 1$. The weight α is an epoch-dependent function ranging from 0 to 1: at the beginning of training, the network emphasises the first stage of graph learning for stability, and the contribution of the two stages is gradually balanced such that $\alpha = 1$ is reached at epoch 40. The network has a total of 1.1 million parameters, and all training was performed on a single NVIDIA RTX A6000 GPU. Each network was trained for 500 epochs, taking around 30 hours.

For inference, a greedy algorithm converts the network's probabilistic outputs into a concrete hierarchical structure. It iterates through candidate low-level components (e.g., W/Higgs bosons), sorted by their probabilities, and accepts those above a threshold whose constituent objects have not already been used. This bottom-up assembly process naturally avoids double-counting and, importantly, does not require a fixed number of target objects, making the method flexible and robust to variable or incomplete event topologies.

4 Experiments and Results

We evaluate our model on the simulated, public $t\bar{t}H$ dataset from [13]. The ground-truth process, which was introduced earlier, is specified in this dataset such that one top quark (t_h) decays via a hadronic W-boson and the other (t_l) via a leptonic one, while the Higgs boson produces two b-jets. Reconstruction is challenging due to the presence of up to 12 extra jets from unrelated collision fragments and the frequent absence of one or more components from the final decay products, in which case the target particle is marked as unreconstructable.

For our second study on event-level classification, we introduce the dominant $t\bar{t} + bb$ background process, which produces the same set of final objects but without the presence of a Higgs boson.

Performance is measured by the reconstruction efficiency, defined as the fraction of correctly reconstructed objects out of all reconstructable objects in the ground truth, and purity, the fraction of correctly reconstructed objects out of all predicted objects. Metrics are computed for each particle type separately as well as at the event level, where only events with all three unstable particles being reconstructed are considered. For a fair baseline comparison, the probability threshold in our inference algorithm is tuned to match the overall Higgs efficiency reported by SPANet.

The reconstruction results in Table 1 show that TIGER achieves comparable efficiency to SPANet, with a slight drop for t_h and full-event reconstruction, but substantially higher purity–reaching improvements of 5-12% for hadronic top quarks and Higgs bosons. The largest gain occurs for 6-jet events, where event-level purity more than doubles. This improvement arises from TIGER's topology-agnostic design, which avoids forcing a complete $t\bar{t}H$ reconstruction on events with missing components.

Table 1: Efficiency and purity comparison for $t\bar{t}H$ events. TIGER and the SPANet baseline are evaluated across varying jet multiplicities. Reconstruction targets include the full event ($t\bar{t}H$), hadronic top (t_h), leptonic top (t_l), and Higgs boson (H). SPANet values were taken from [13]. Underlined entries indicate efficiencies at which TIGER was aligned with SPANet for a fair purity comparison. Higher purity performance is indicated in bold.

	SPANet Eff. [%]				TIGER Eff. [%]				SPANet Pur. [%]				TIGER Pur. [%]			
	$t\bar{t}H$	t_h	t_l	H	$t\bar{t}H$	t_h	t_l	H	$t\bar{t}H$	t_h	t_l	H	$t\bar{t}H$	t_h	t_l	H
6 j	54	54	69	50	50	51	70	51	15	24	63	38	**32**	**35**	**64**	**50**
7 j	42	48	68	49	39	46	69	48	16	26	63	38	**27**	**34**	**64**	**47**
$\geq$8 j	33	42	68	47	30	40	67	45	14	26	**63**	37	**22**	**31**	62	**42**
All	45	49	69	<u>49</u>	40	46	69	<u>49</u>	15	25	63	38	**27**	**34**	**64**	**47**

Finally, we evaluate TIGER for event-level classification by training end-to-end on the combined $t\bar{t}H$ and $t\bar{t} + bb$ dataset. SPANet is tested both on a fine-tuned classifier on a pretrained SPANet backbone and in an end-to-end configuration. As shown in Figure 2, TIGER (AUC = 0.780) outperforms both the pretrained (AUC = 0.744) and fine-tuned (AUC = 0.771) SPANet models.

5 Conclusions and Outlook

In this work, we introduced TIGER, a topology-agnostic framework for event reconstruction. Its hierarchical graph architecture is motivated by the sequential nature of particle decays but does not rely on any predefined topology as an architectural prior or explicit input. This removes strong

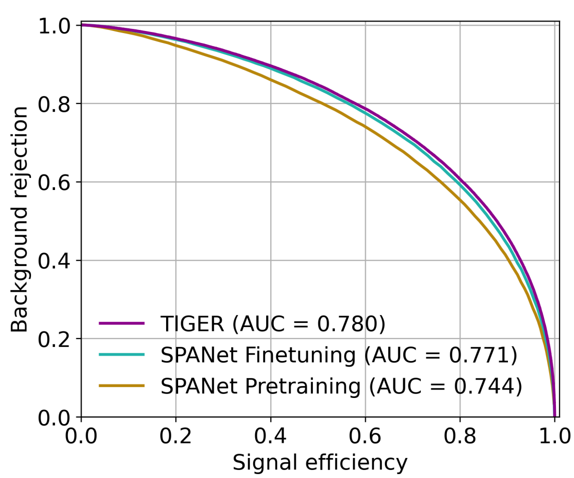

Figure 2: ROC curves comparing TIGER to two SPANet baselines on the event-level classification task of distinguishing signal from background events.

structural assumptions, making the learning problem more challenging but resulting in a more flexible and broadly applicable framework for realistic physics analyses.

On a challenging semileptonic $t\bar{t}H$ benchmark, TIGER achieves reconstruction efficiencies comparable to the baseline while delivering substantially higher purity. This gain stems directly from its topology-agnostic design, which avoids the bias of enforcing complete event reconstructions even when parts of the decay chain are missing. In addition, the architecture naturally supports event-level classification, outperforming a fine-tuned baseline in a signal-versus-background task without requiring multi-stage training.

These results establish TIGER as a versatile tool for event reconstruction at the LHC. The approach can be extended to analyses involving multiple signal and background topologies in parallel. For example, a single TIGER model could target rare Higgs decays like $H \to c\bar{c}$ [18, 19], or Di-Higgs searches [20] while simultaneously handling a wide range of competing backgrounds. Such capabilities point toward the long-term goal of a unified, general-purpose model for physics analyses at the LHC.

Code and data availability

The data was taken from `https://mlphysics.ics.uci.edu/data/2023_spanet/`. The code is publicly available at `https://github.com/nathaliesoy/tiger`.

Acknowledgments

We thank Alberto Ibanez, Alexander Shmakov, and Hideki Okawa for insightful discussions and clarifications concerning the baseline model used in our comparison. NS, NK, and EG are supported by the Minerva Stiftung under grant number 715027, and the Weizmann Institute for Artificial Intelligence grant program Ref 151676.

References

[1] Lyndon Evans and Philip Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, aug 2008.

[2] HEP ML Community. A Living Review of Machine Learning for Particle Physics.

[3] Nilotpal Kakati, Etienne Dreyer, Anna Ivina, Francesco Armando Di Bello, Lukas Heinrich, Marumi Kado, and Eilam Gross. HGPflow: Extending Hypergraph Particle Flow to Collider Event Reconstruction. 10 2024.

[4] Joosep Pata, Javier Duarte, Farouk Mokhtar, Eric Wulff, Jieun Yoo, Jean-Roch Vlimant, Maurizio Pierini, and Maria Girone. Machine learning for particle flow reconstruction at cms. *Journal of Physics: Conference Series*, 2438(1):012100, February 2023.

[5] Spandan Mondal and Luca Mastrolorenzo. Machine learning in high energy physics: a review of heavy-flavor jet tagging at the lhc. *The European Physical Journal Special Topics*, 233(15–16):2657–2686, July 2024.

[6] Lukas Ehrke, John Andrew Raine, Knut Zoch, Manuel Guth, and Tobias Golling. Topological reconstruction of particle physics processes using graph neural networks. *Physical Review D*, 107(11), June 2023.

[7] Michael James Fenton, Alexander Shmakov, Ta-Wei Ho, Shih-Chieh Hsu, Daniel Whiteson, and Pierre Baldi. Permutationless many-jet event reconstruction with symmetry preserving attention networks. *Physical Review D*, 105(11), June 2022.

[8] Alexander Shmakov, Michael James Fenton, Ta-Wei Ho, Shih-Chieh Hsu, Daniel Whiteson, and Pierre Baldi. Spanet: Generalized permutationless set assignment for particle physics using symmetry preserving attention. *SciPost Physics*, 12(5), May 2022.

[9] ATLAS Collaboration. Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8$ tev with the atlas detector. *Journal of High Energy Physics*, 2017(9), September 2017.

[10] ATLAS Collaboration. Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{\mathrm{s}}$ = 13 tev using the atlas detector. *Journal of High Energy Physics*, 2021(1), January 2021.

[11] Johannes Erdmann, Stefan Guindon, Kevin Kroeninger, Boris Lemmer, Olaf Nackenhorst, Arnulf Quadt, and Philipp Stolte. A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework. *Nucl. Instrum. Meth. A*, 748:18–25, 2014.

[12] Callum Birch-Sykes, Brian Le, Yvonne Peters, Ethan Simpson, and Zihan Zhang. Reconstructing short-lived particles using hypergraph representation learning. *Phys. Rev. D*, 111:032004, Feb 2025.

[13] Michael James Fenton, Alexander Shmakov, Hideki Okawa, Yuji Li, Ko-Yang Hsiao, Shih-Chieh Hsu, Daniel Whiteson, and Pierre Baldi. Reconstruction of unstable heavy particles using deep symmetry-preserving attention networks. *Communications Physics*, 7(1), April 2024.

[14] Haoyang Li, Marko Stamenkovic, Alexander Shmakov, Michael Fenton, Darius Shih-Chieh Chao, Kaitlyn Maiya White, Caden Mikkelsen, Jovan Mitic, Cristina Mantilla Suarez, Melissa Quinnan, Greg Landsberg, Harvey Newman, Pierre Baldi, Daniel Whiteson, and Javier Duarte. Reconstruction of boosted and resolved multi-higgs-boson events with symmetry-preserving attention networks, 2025.

[15] Zhen Zhang, Jiajun Bu, Martin Ester, Jianfeng Zhang, Chengwei Yao, Zhi Yu, and Can Wang. Hierarchical graph pooling with structure learning, 2019.

[16] Zhitao Ying, Jiaxuan You, Christopher Morris, Xiang Ren, Will Hamilton, and Jure Leskovec. Hierarchical graph representation learning with differentiable pooling. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[17] William S. Peebles and Saining Xie. Scalable diffusion models with transformers. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4172–4182, 2022.

[18] Search for Higgs boson decay to a charm quark-antiquark pair via ttH production. Technical report, CERN, Geneva, 2025.

[19] ATLAS Collaboration. Measurements of wh and zh production with higgs boson decays into bottom quarks and direct constraints on the charm yukawa coupling in 13 tev pp collisions with the atlas detector. *Journal of High Energy Physics*, 2025(4), April 2025.

[20] ATLAS Collaboration. Combination of Searches for Higgs Boson Pair Production in pp Collisions at s=13 TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 133(10):101801, 2024.