Hunting and identifying colored resonances in four-top events with machine learning

Thomas Flacke

KIAS

T. Flacke, Jeong Han Kim, Pyungwon Ko, M. Kunkel, Jun Seung Pi, W. Porod, L. Schwarze [JHEP 11 (2023) 009]

T. Flacke, Jeong Han Kim, M. Kunkel, Jun Seung Pi, W. Porod [arXiv:2506xxxx]

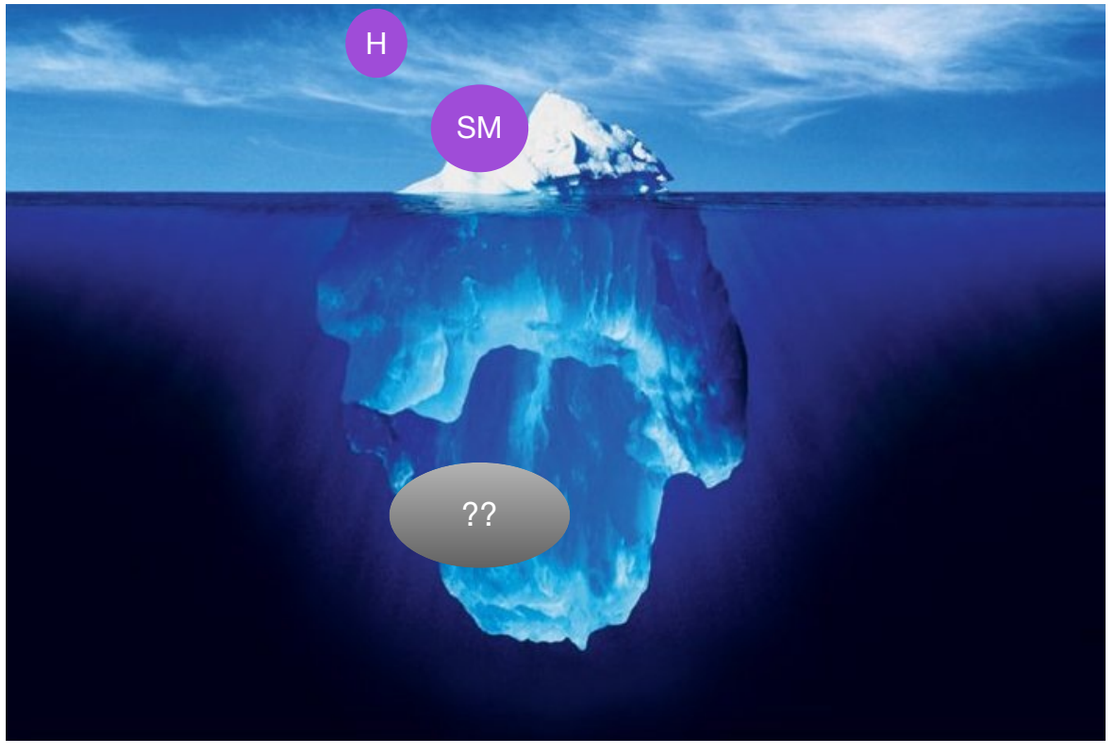
2025 KIAS CAINS Workshop,
Park Roche Resort, Jeongseon 2025/05/29

Outline

- Introduction: The Standard Model of particle physics and particle colliders.

- Searching for new colored particles which decay to 4 tops with ML
  - Identifying signal events
  - Distinguishing different BSM candidates

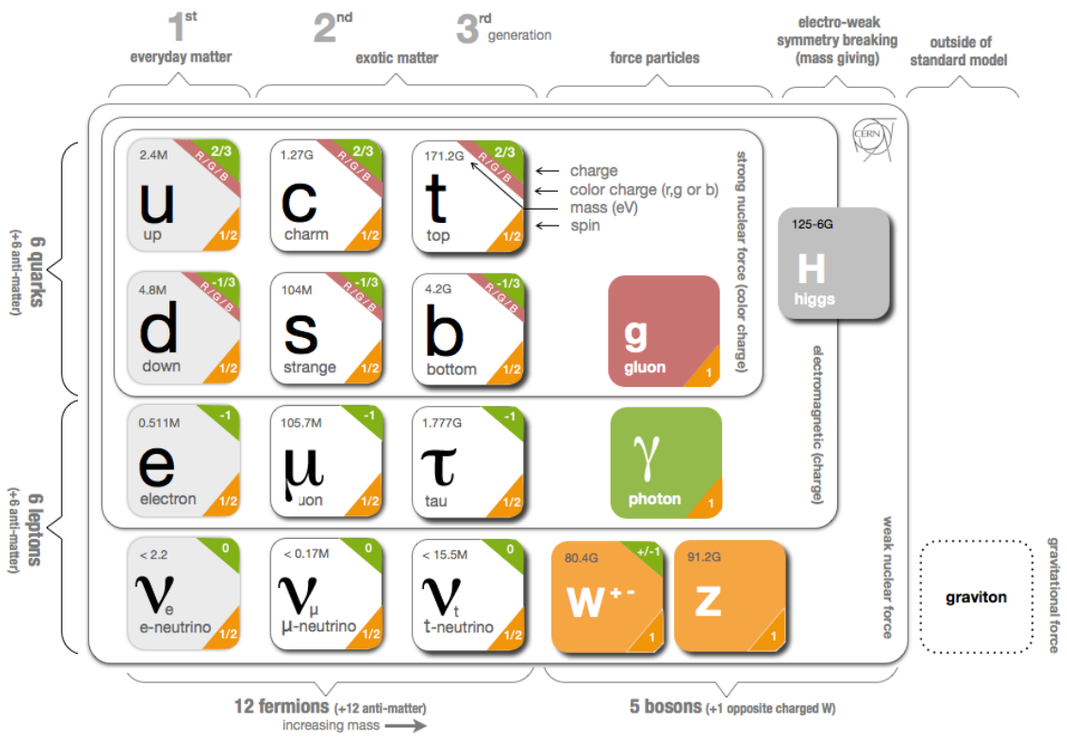
- Conclusions



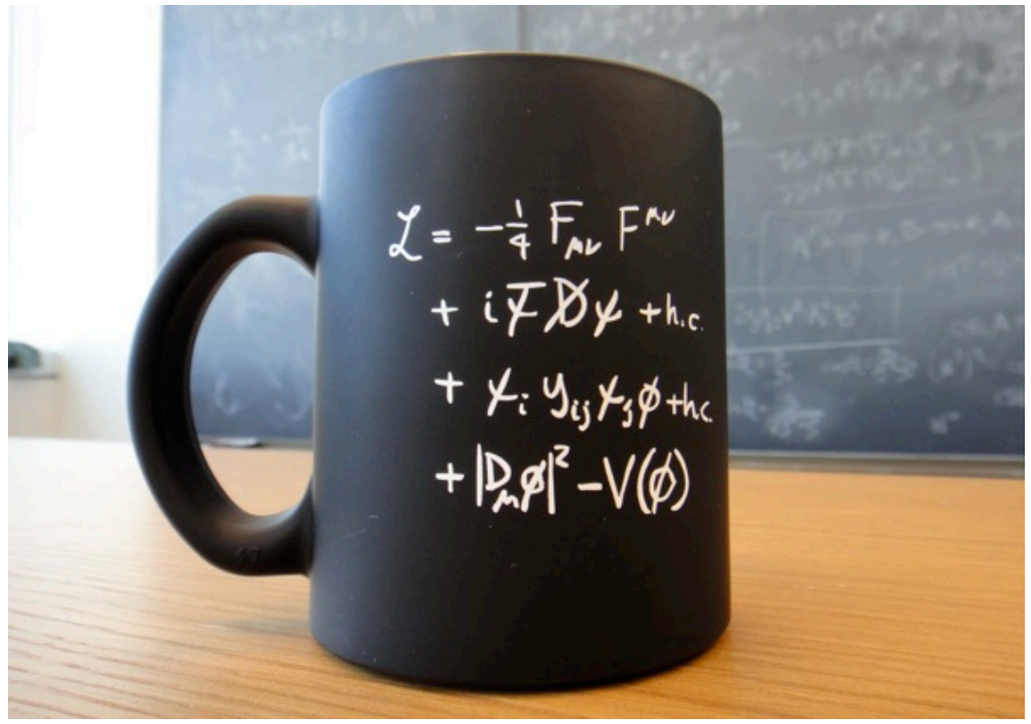
The Standard Model (SM) of particle physics

field content ↔ particles

characterized by masses and charges

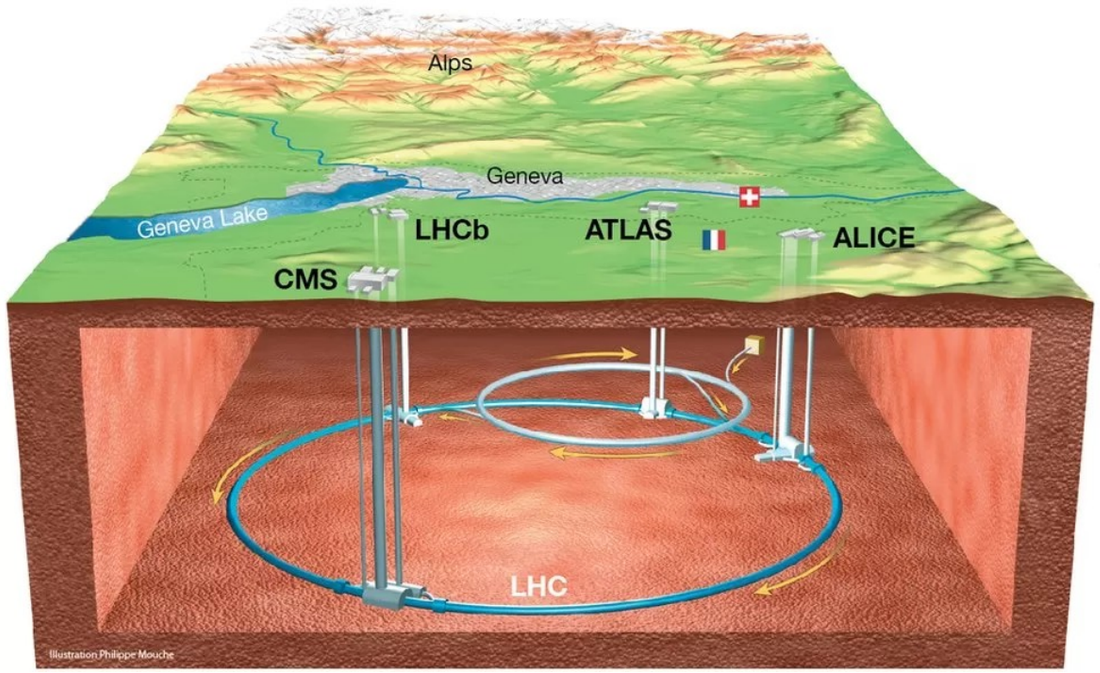


Interactions described by a QFT Lagrange density

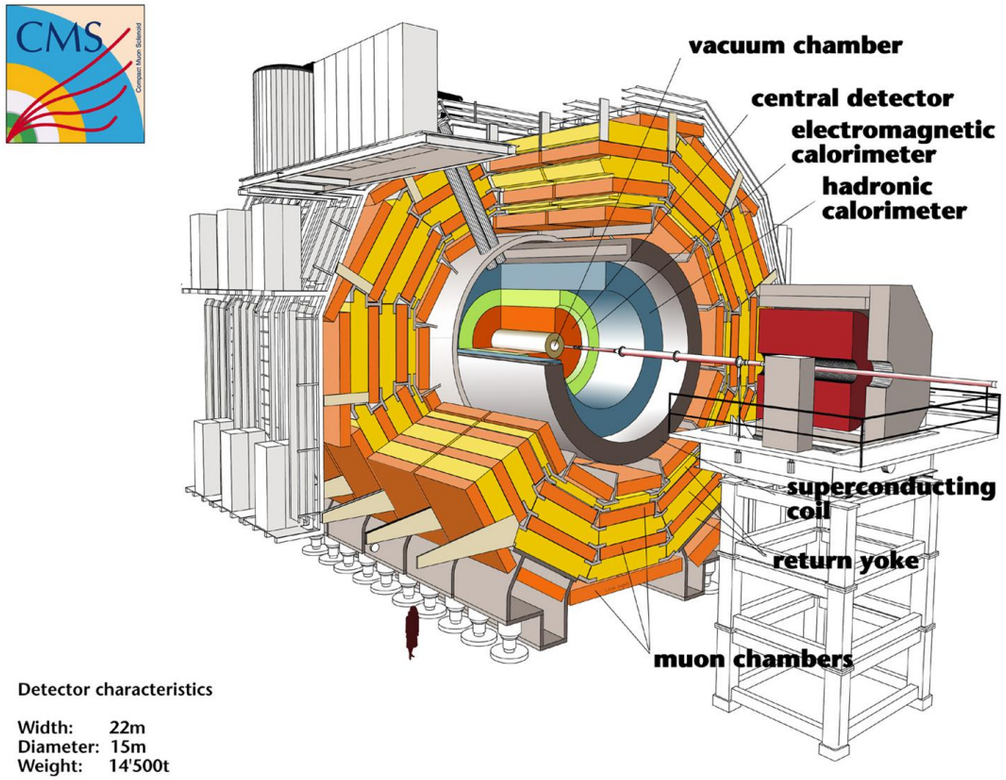


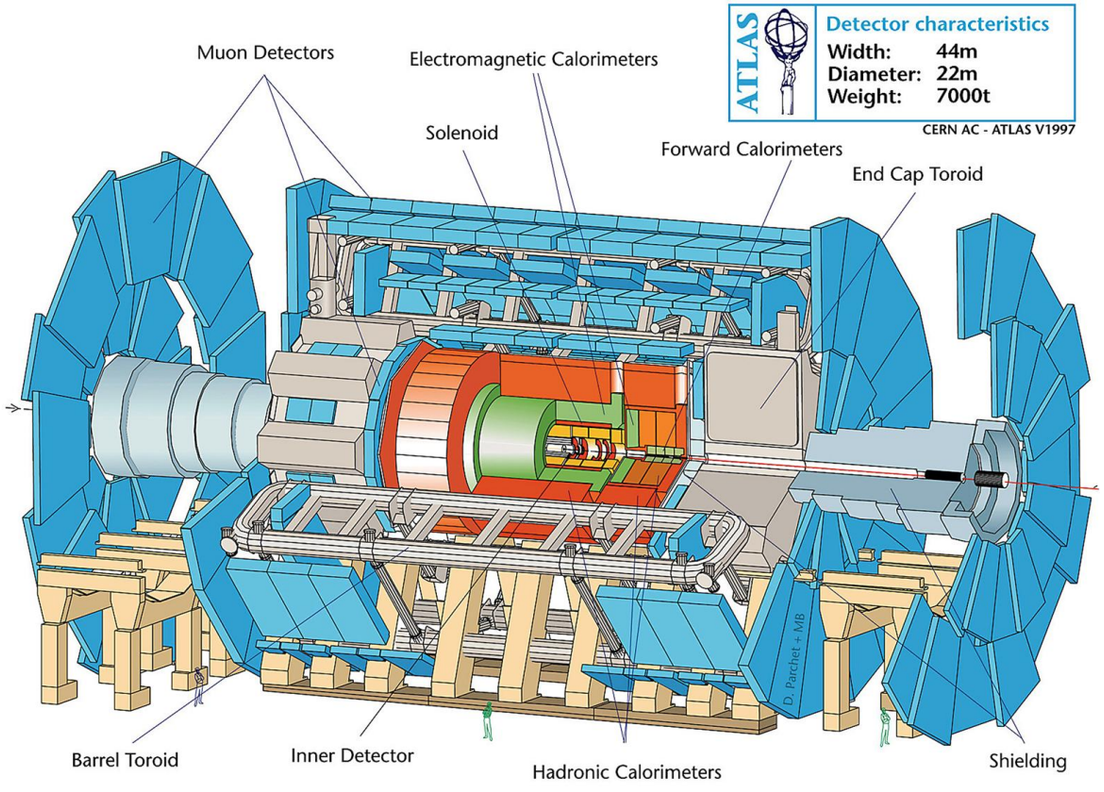
- The SM extremely successful (tested at the Large Hadron Collider (LHC) and many precision experiments — and still not falsified!)
- The SM does not explain Dark Matter, neutrino masses, baryon asymmetry, the smallness of the Higgs mass, the strong CP problem, …
- Extensions of the SM must be formulated and tested in experiments.

Particle collisions in practice: LHC

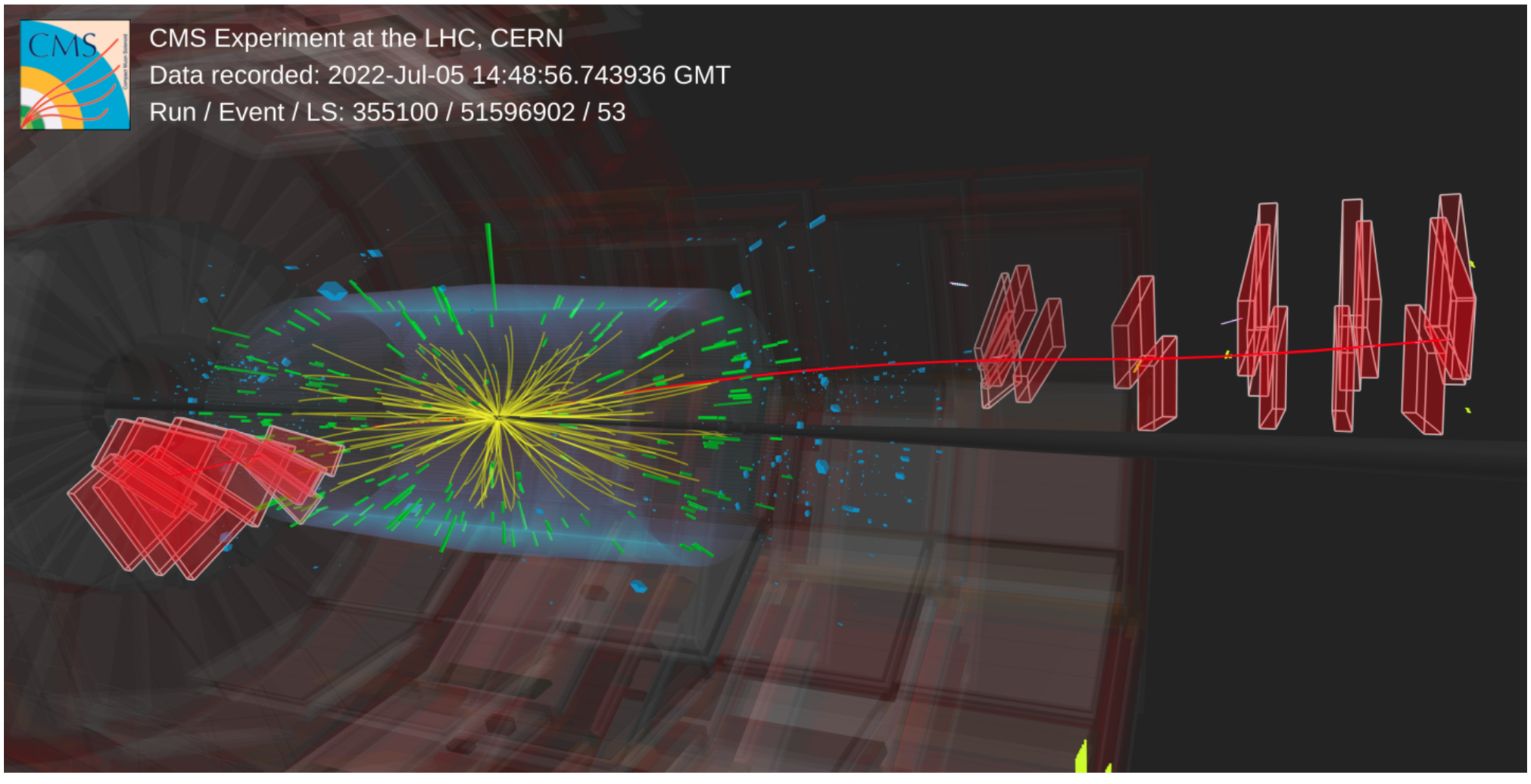


- Energy: 13.6 TeV (run3)
- collisions : ~$3 \cdot 10^7$/s (run3)
- typical event file size: 1 MB
  (ref: ATLAS fact sheet)
- stored data: 10^4 TB/year
  (ref: ATLAS fact sheet)
- Note I: events have to be highly pre-selected through triggers to allow storage.
- Note II: "Interesting" physics events are a tiny subset of the data. Example: one Higgs is produced in every ~10^{10} collisions





How an event looks like
(in practice; reconstructed CMS event)



Production and decay of the colored resonances at hadron colliders

Colored resonances in various color representations which dominantly couple to the top sector of the Standard Model are predicted in Composite Higgs models (and also in other Standard Model extensions).

The interaction Lagrangians we use read:

$$\mathcal{L}_{\text{int},8} = \lambda_8 \, \bar{t} \, S_8 \, t$$

$$\mathcal{L}_{\text{int},6} = \lambda_6 \, \bar{t} \, S_6 P_L \, t^c + \text{h.c.}$$

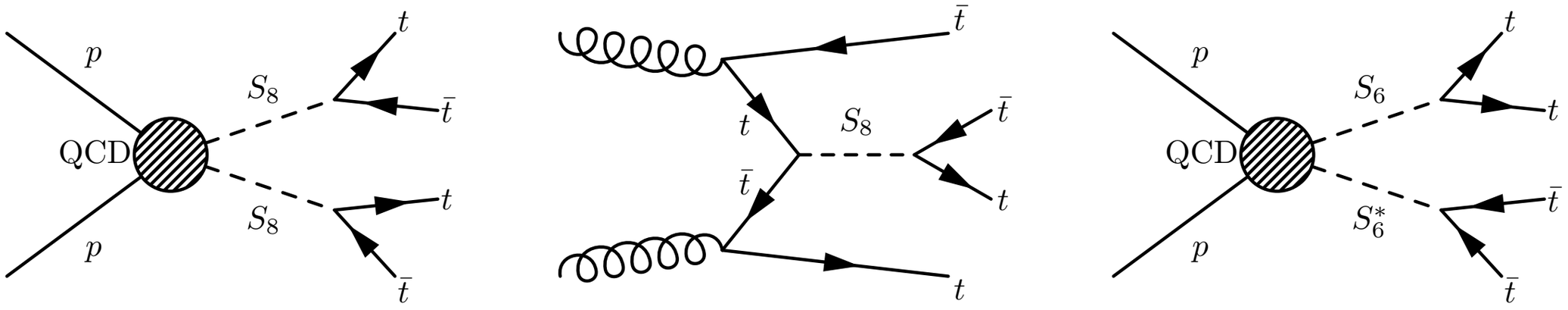


Figure 1. Production of four top quarks via the QCD pair productions and the single production of colored scalars.

Current bounds from LHC searches

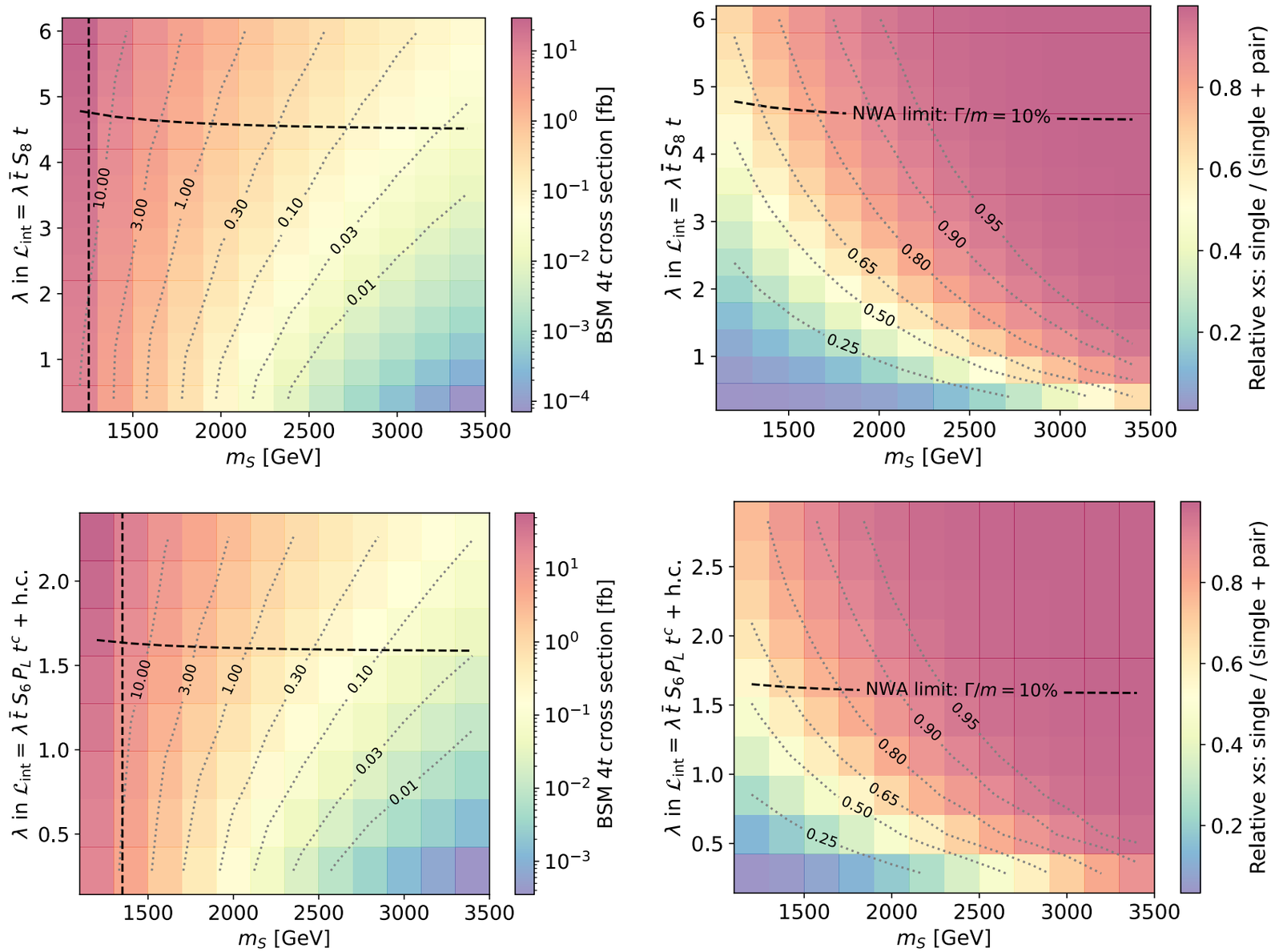


Figure 2. Comparison of single production and pair production of (top) colour octet scalars S_8 and (bottom) colour sextet scalars S_6 with (left) total and (right) relative cross sections calculated at LO. The vertical and horizontal dashed lines indicate the limits from recasts and validity of the narrow width approximation, respectively. Since the single and pair production cannot be cleanly separated, we use $\sigma(pp \to 4t; S_8) - \sigma(pp \to S_8 S_8)$ as a proxy for single production, where "; S_8" indicates that at least one S_8 has to occur (analogous for S_6). This is only a good approximation if interference effects are small.

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances

Event simulation and pre-selection:

- Simulation chain: Feynrules → Madgraph5 → Pythia8 → Delphes3.4.1 →Fastjet3.3.1
- #events
  Signal: 6m events per benchmark mass (1.2 TeV - 2.5 TeV in 100 GeV steps)
  Background: 4t (4.2m), tth (70m), ttV (150m), ttVV (4.3m), VVV (48m)
- basic selection cuts:
  - exactly 2 same-sign leptons
  - at least 3 b-tagged jets
  - at least 3 (more) jets
  - mild missing p_T cut (20 GeV)
  - mild S_T cut (400 GeV)
  - standard lepton isolation and rapidity criteria (ATLAS config)

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances

Data & data-pre-processing

Kinematic data:

We demand 2 leptons, 3 jets, 3 b-jets and construct from them 51 kinematic observables

$$\mathcal{K} = \bigcup_{i \neq j} M_{ij} \cup \bigcup_{i \neq j} \Delta R_{ij} \cup \bigcup_{i} p_{Ti} \cup \{\not{E}_T, S_T\}$$

Jet images:

For each event, we determine an angular maps in the following way:

1. Set the center of the (η,ϕ) plane as the midpoint between the two same-sign leptons.

2. Determine the (η,ϕ) map of the p_T of (a) charged "jets", ((b) neutral "jets",) (c) di-leptons by binning objects of the respective class in a 50x50 grid and and summing the p_T in each bin to obtain the pixel intensity

→ $V_{\text{image}}^{(C,\ell)} = (2 \times 50 \times 50)$ or $V_{\text{image}}^{(C,N,\ell)} = (3 \times 50 \times 50)$

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances

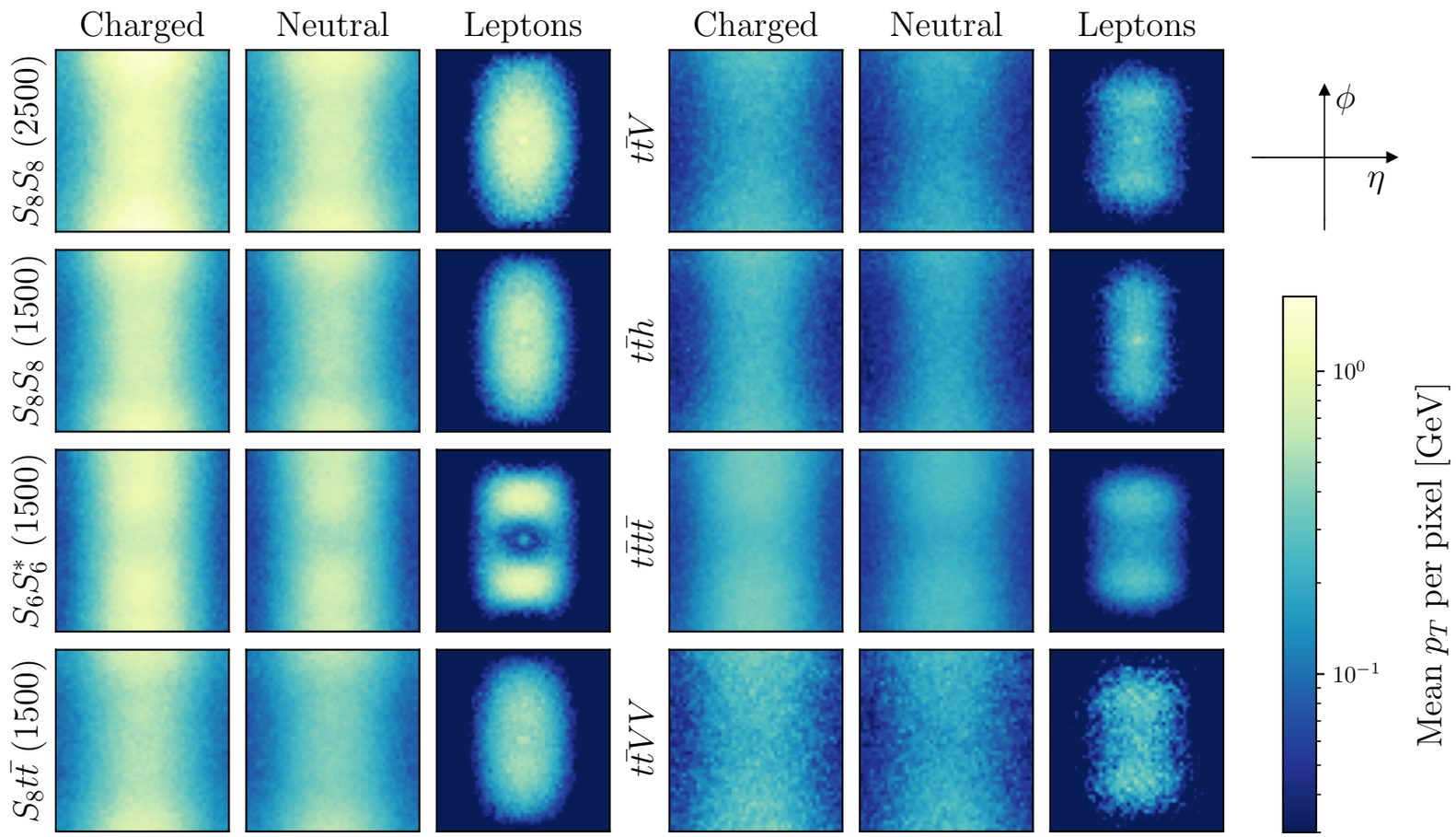
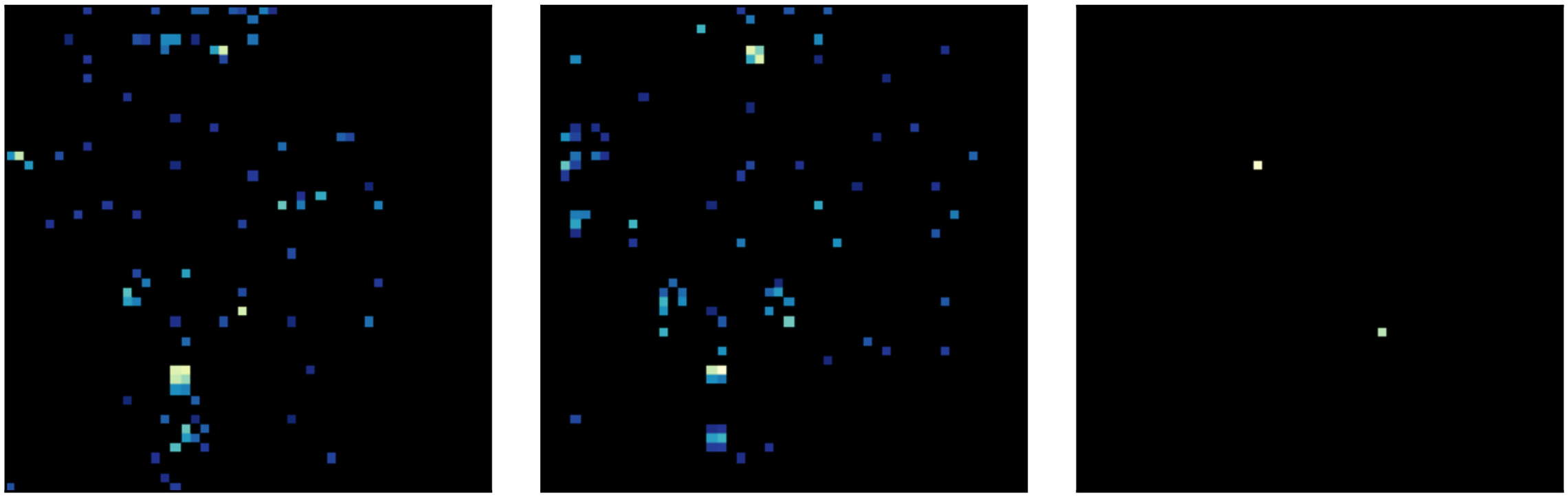


Figure 3. Jet images for background (right) and signal processes (left) where the parentheses indicate the scalar mass in GeV. The three columns show the images of charged hadrons, neutral hadrons, and the two isolated leptons. Each panel shows the average distribution of particles taken over all simulated events.

overlaid jet images (main background and signal classes)

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances



single jet images are sparse

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances

Network: Convolutional Neural Network (CNN) combined with a simple multilayer perceptron (MLP)

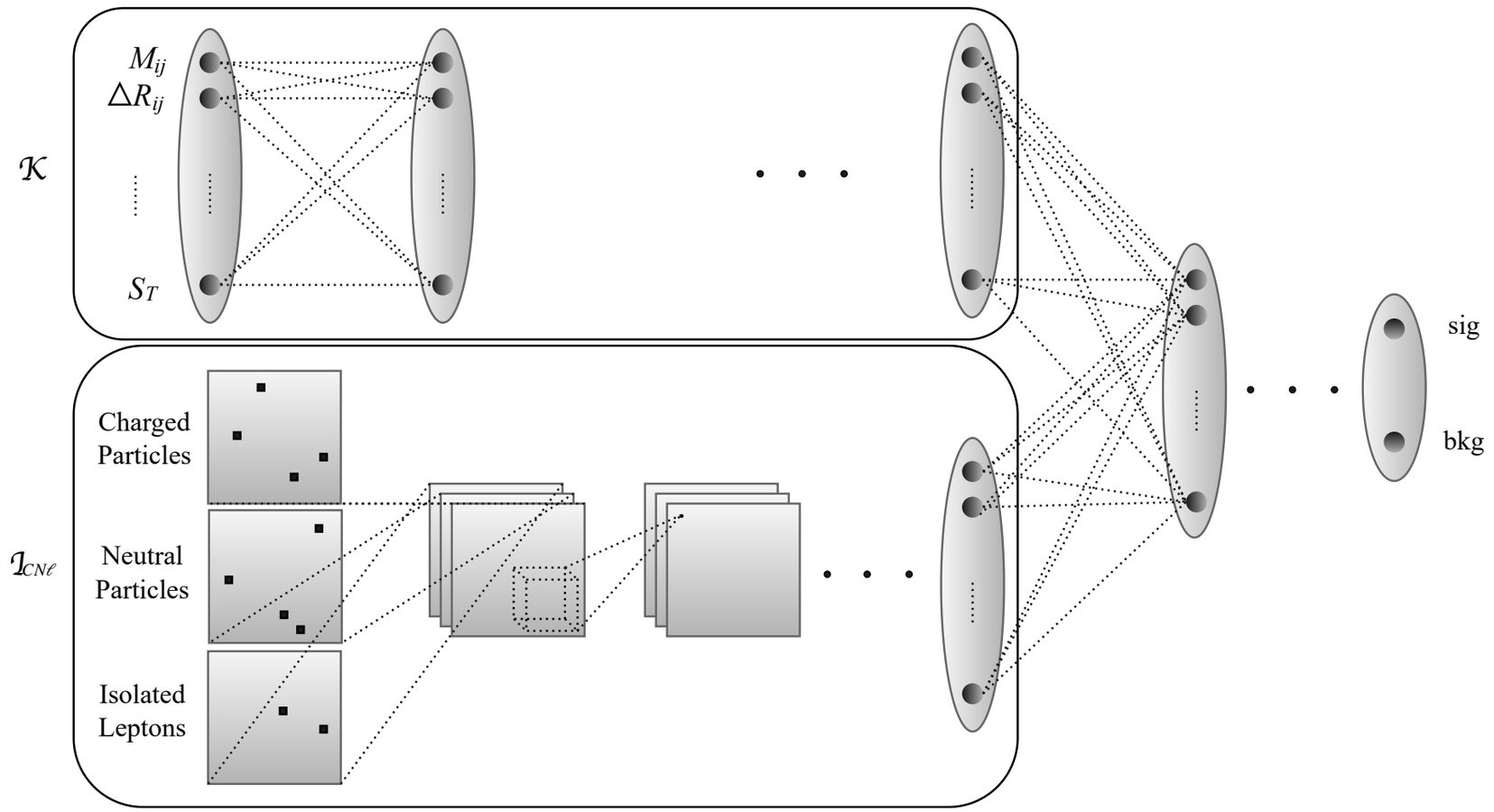


Figure 5. A schematic architecture of the neural networks used in this paper. The separate DNN chain in the upper panel is used only when the kinematic variables are included.

Machine Learning applied to BSM searches: 4-top events from BSM colored resonances

Network: Convolutional Neural Network (CNN) combined with a simple multilayer perceptron (MLP) in more detail

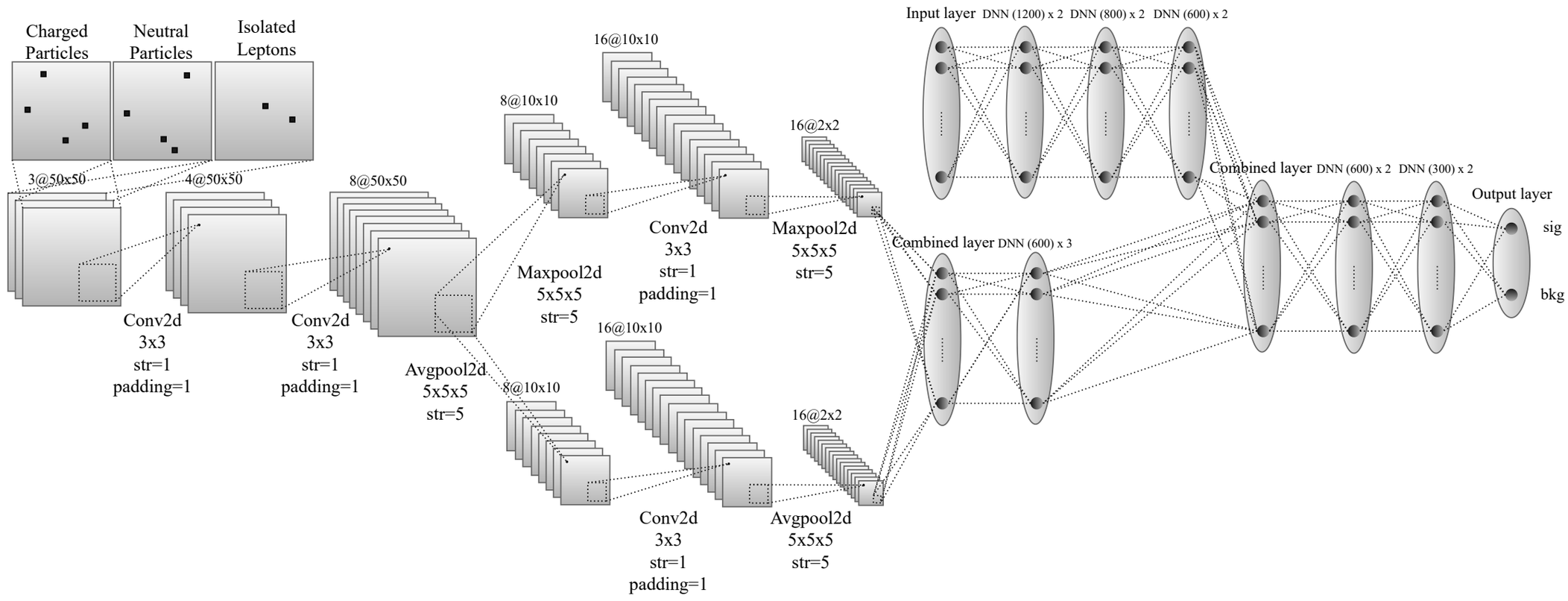
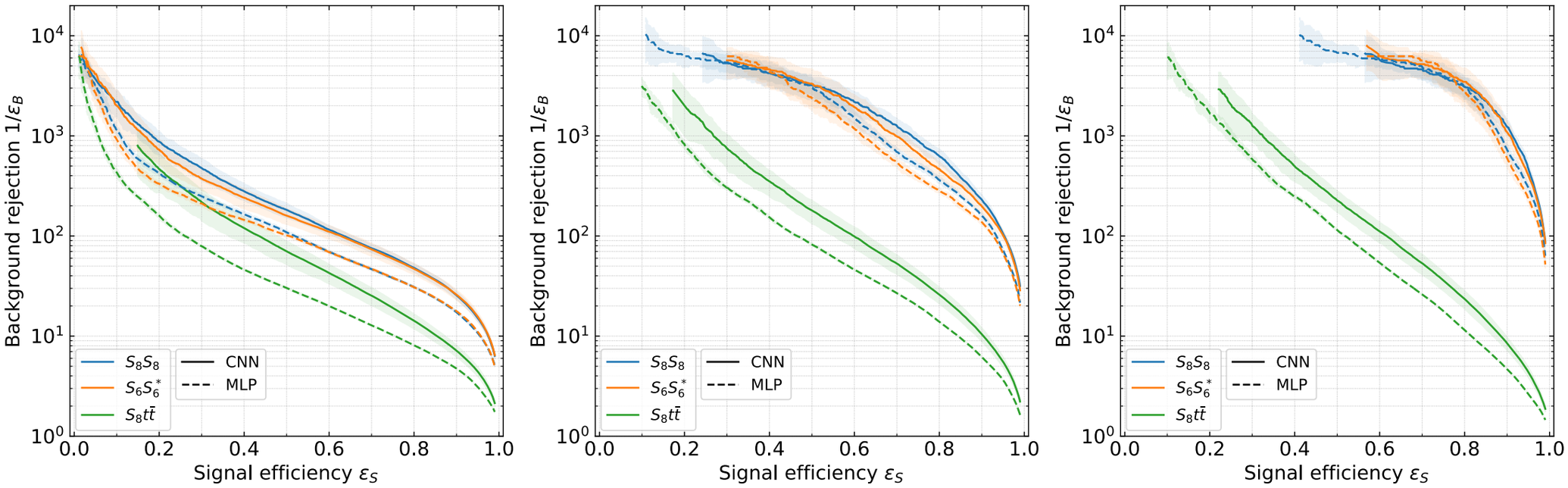


Figure 11. A schematic CNN architecture used in this article. The separate MLP in the right-upper corner is used only when kinematic variables are included.

Task 1: Distinguishing signal from background



(a) $m_S = 1200$ GeV (b) $m_S = 1800$ GeV (c) $m_S = 2300$ GeV

Figure 5. Receiver operator characteristic (ROC) curves comparing network performances for different processes and mass points.

As a working point, we put a fixed NN score cut, demanding 5 background events to pass.

Passing this cut are S signal events and B=5 background events. From these we calculate the significances for discovery and exclusion by

$$Z_{\text{dis}} \equiv \sqrt{-2\,\ln\left(\frac{L(B|S+B)}{L(S+B|S+B)}\right)}, \qquad Z_{\text{exc}} \equiv \sqrt{-2\,\ln\left(\frac{L(S+B|B)}{L(B|B)}\right)}$$

with the Poisson likelihood $L(x|n) = \frac{x^n}{n!}e^{-x}$. When calculating the 2σ exclusion bound

To obtain exclusion and discovery bounds on the cross-section, we resale the cross-section until we reach $Z_{\text{dis}} \geq 5, \qquad Z_{\text{exc}} \geq 1.64$.

Task 1: Distinguishing signal from background

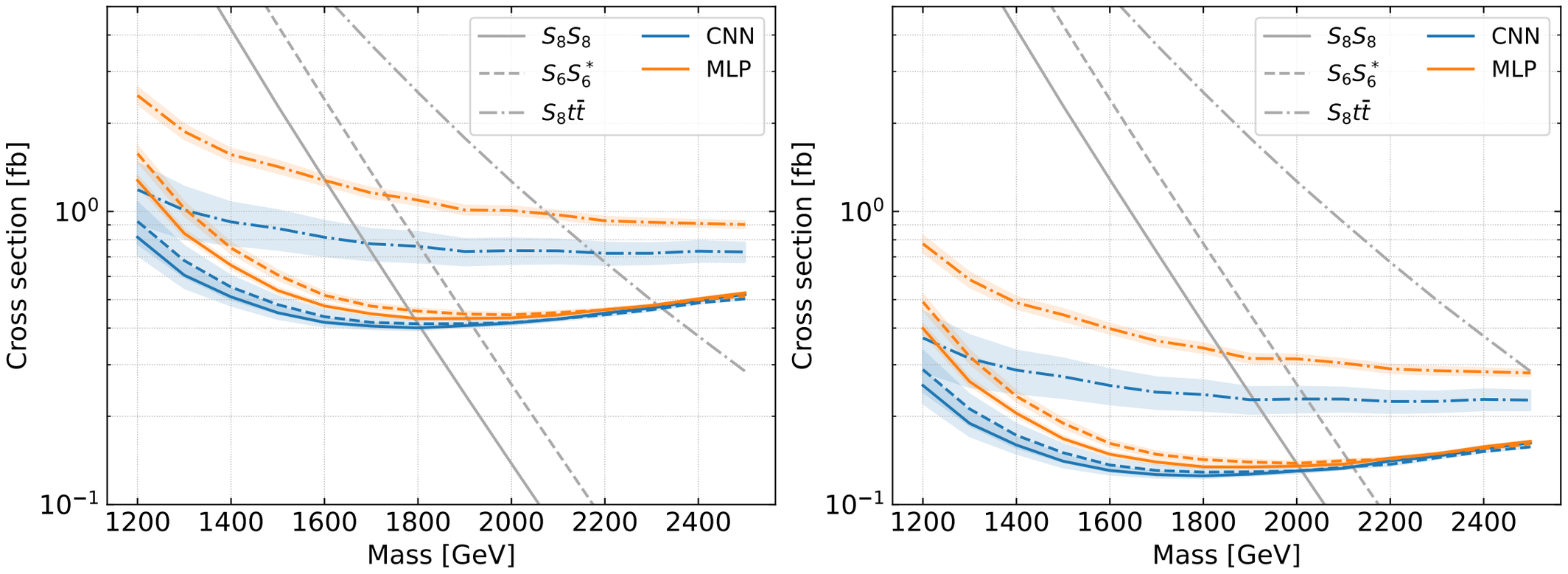


Figure 6. Discovery reach (left) and exclusion limit (right) for different processes and networks. The process $pp \to ttS_8$ is dominated by a single production but also includes a portion of pair production.

Task 2: Distinguishing sextet pair production, octet pair production, and single production

Suppose LHC finds a 5σ excess. Can one tell different models apart?

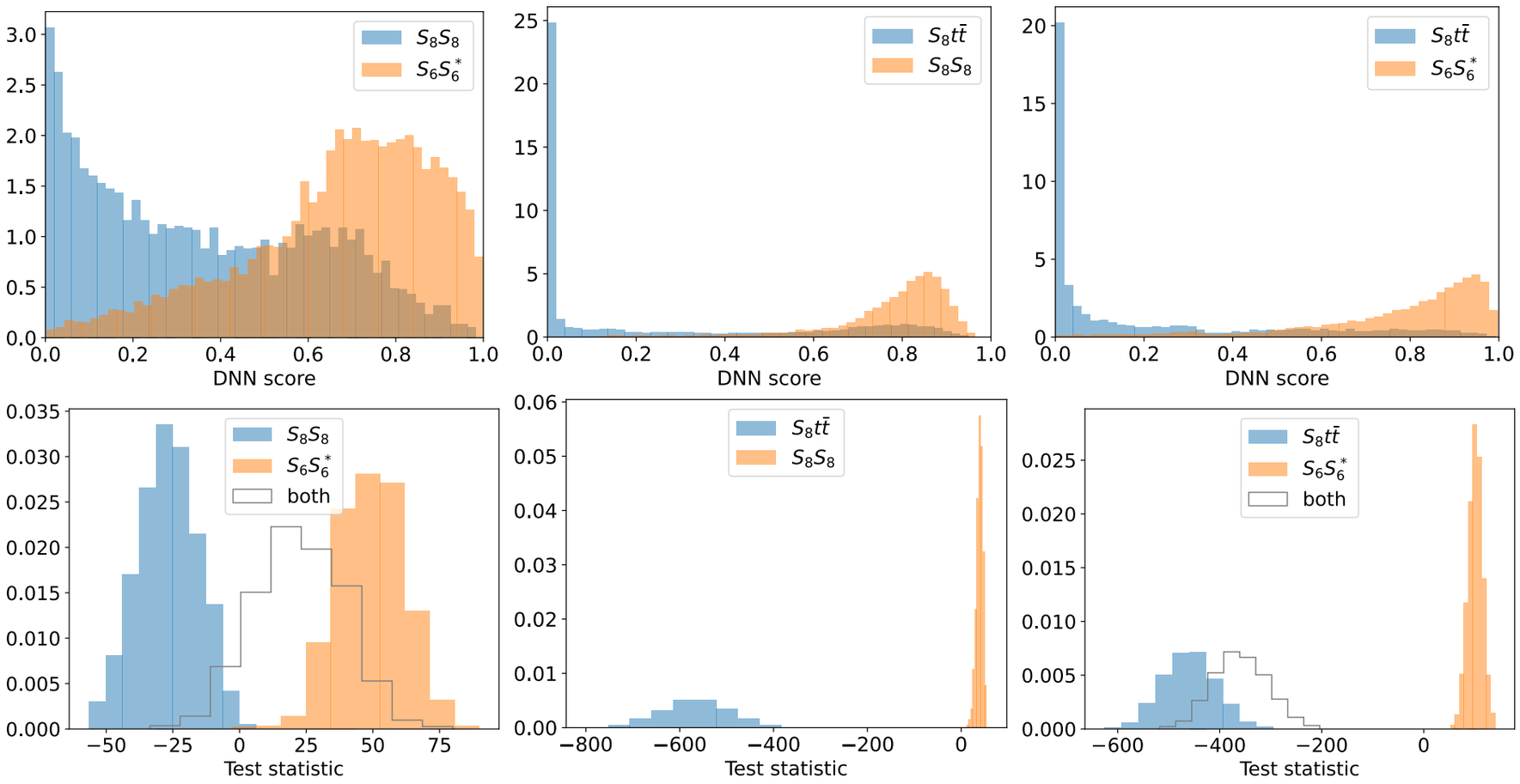
We re-train the same network architecture for model discrimination between any combination of two models out of (octet pair production, octet single production, sextet pair pair production). For each pair we obtain the DNN score distributions.

We then draw samples x_i from this distribution where $i = 1, \ldots, N_{5\sigma}$ with $N_{5\sigma}$ the number of events needed for discovery. The goal is to determine whether they come from distribution $j = 1$ or 2. To this end we construct the test statistic t from a log-likelihood ratio,

$$t = -2 \ln \frac{\mathcal{L}_1(\{x_i\})}{\mathcal{L}_2(\{x_i\})}, \qquad \mathcal{L}_j(\{x_i\}) = \prod_{i=1}^{N_{5\sigma}^j} f_j(x_i)$$

for two hypotheses f_1 and f_2. By drawing many such sample sets and calculating the test statistic for each, we obtain the distribution of t

Task 2: Distinguishing sextet pair production, octet pair production, and single production



(a) S_8S_8 vs $S_6S_6^*$ (b) $S_8t\bar{t}$ vs S_8S_8 (c) $S_8t\bar{t}$ vs $S_6S_6^*$

Figure 7. Separating two signal processes for a scalar mass of 1.8 TeV. The top row shows the NN score distribution obtained from training the CNN to separate two signal processes. The second row shows the distributions of the corresponding test statistics as defined in Eq. (5.3).

Summary

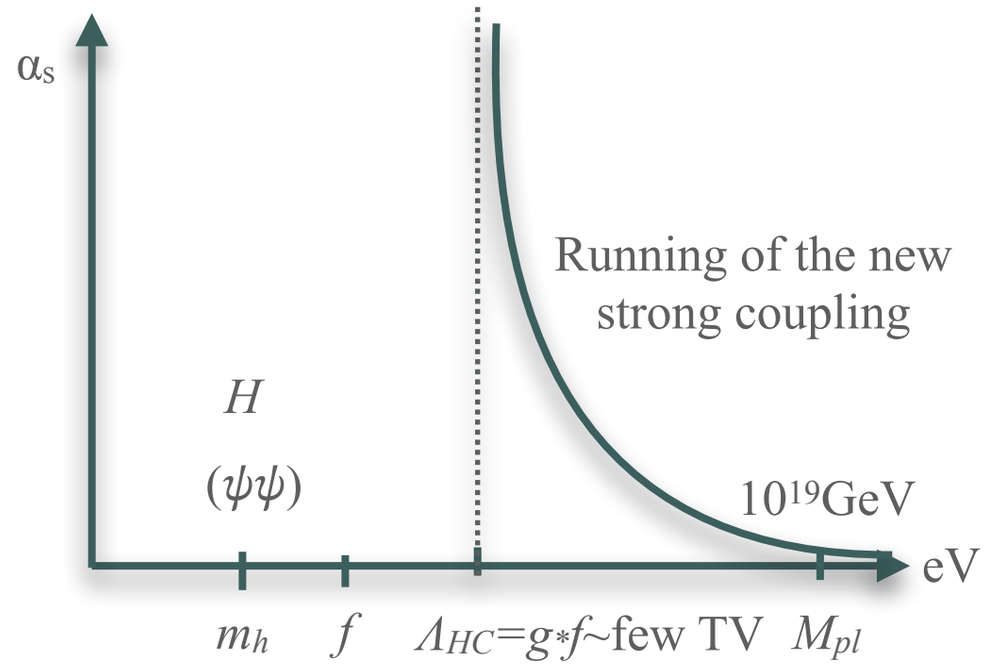
- Many BSM models predict BSM colored states (in various color representations) which can decay into top quarks.
- Current LHC searches for 4-top events constrain color octets (sextets) to be heavier than 1.2 TeV (1.3 TeV).
- The studied combination of a DNN for kinematic information and a convolutional neural network for jet images provide excellent performance on 4-top final states with 2 same-sign leptons.
  For LHC at 3000 fb^{-1} we find a discovery potential up to m=1.8 TeV (m=1.92 TeV) for pair produced octet (sextet) scalars and an exclusion potential up to m=2.02 TeV (m=2.14 TeV).
- We demonstrated that with our network architecture the number of events needed for discovery are sufficient to discriminate octet pair-production, sextet pair-production and octet single-production.

Backup

Motivation for a composite Higgs

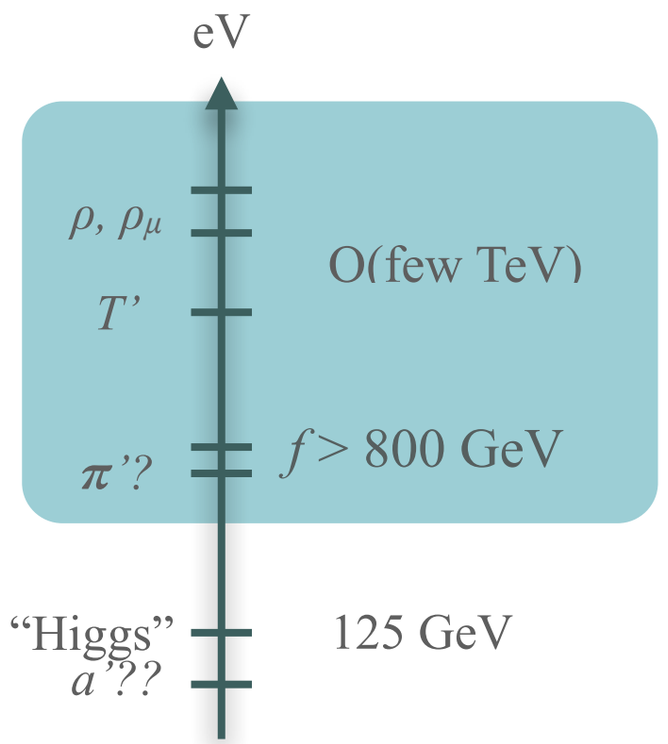
An alternative solution to the hierarchy problem:

- Generate a scale $\Lambda_{HC} << M_{pl}$ through a new confining gauge group.
- Interpret the Higgs as a pseudo-Nambu-Goldstone boson (pNGB) of a spontaneously broken global symmetry of the new strong sector.
  [Georgi, Kaplan (1984)]



The price to pay:

- additional resonances around Λ_{HC} (vectors, vector-like fermions, scalars),
- additional light scalars (pNGBs).
- deviations of the Higgs couplings from their SM values of O(v/f).



Composite Higgs Models: Towards underlying models

A wish list to construct and classify candidate models:

Underlying models of a composite Higgs should

[Gherghetta etal (2015), Ferretti etal (2014), PRD 94 (2016) no 1, 015004, JHEP 1701, 094]

- contain no elementary scalars (to not re-introduce a hierarchy problem),
- have a simple hyper-color group,
- have a Higgs candidate amongst the pNGBs of the bound states,
- have a top-partner amongst its bound states (for top mass via partial compositeness),

The resulting models have several common features:

- All models contain hyper baryons beyond the top partners.
- All models predict SM neutral, electroweak and colored pNGBs beyond the Higgs multiplet.

List of "minimal" CHM UV embeddings

$G_{\rm HC}$	ψ	χ	Restrictions	$-q_\chi/q_\psi$	Y_χ	Non Conformal	Model Name
	Real	Real	$\mathrm{SU}(5)/\mathrm{SO}(5)\times\mathrm{SU}(6)/\mathrm{SO}(6)$				
$SO(N_{\rm HC})$	$5\times\mathbf{S}_2$	$6\times\mathbf{F}$	$N_{\rm HC}\geq 55$	$\frac{5(N_{\rm HC}+2)}{6}$	1/3	/	
$SO(N_{\rm HC})$	$5\times\mathbf{Ad}$	$6\times\mathbf{F}$	$N_{\rm HC}\geq 15$	$\frac{5(N_{\rm HC}-2)}{6}$	1/3	/	
$SO(N_{\rm HC})$	$5\times\mathbf{F}$	$6\times\mathbf{Spin}$	$N_{\rm HC}=7,9$	$\frac{5}{6}$, $\frac{5}{12}$	1/3	$N_{\rm HC}=7,9$	M1, M2
$SO(N_{\rm HC})$	$5\times\mathbf{Spin}$	$6\times\mathbf{F}$	$N_{\rm HC}=7,9$	$\frac{5}{6}$, $\frac{5}{3}$	2/3	$N_{\rm HC}=7,9$	M3, M4
	Real	Pseudo-Real	$\mathrm{SU}(5)/\mathrm{SO}(5)\times\mathrm{SU}(6)/\mathrm{Sp}(6)$				
$Sp(2N_{\rm HC})$	$5\times\mathbf{Ad}$	$6\times\mathbf{F}$	$2N_{\rm HC}\geq 12$	$\frac{5(N_{\rm HC}+1)}{3}$	1/3	/	
$Sp(2N_{\rm HC})$	$5\times\mathbf{A}_2$	$6\times\mathbf{F}$	$2N_{\rm HC}\geq 4$	$\frac{5(N_{\rm HC}-1)}{3}$	1/3	$2N_{\rm HC}=4$	M5
$SO(N_{\rm HC})$	$5\times\mathbf{F}$	$6\times\mathbf{Spin}$	$N_{\rm HC}=11,13$	$\frac{5}{24}$, $\frac{5}{48}$	1/3	/	
	Real	Complex	$\mathrm{SU}(5)/\mathrm{SO}(5)\times\mathrm{SU}(3)^2/\mathrm{SU}(3)$				
$SU(N_{\rm HC})$	$5\times\mathbf{A}_2$	$3\times(\mathbf{F},\overline{\mathbf{F}})$	$N_{\rm HC}=4$	$\frac{5}{3}$	1/3	$N_{\rm HC}=4$	M6
$SO(N_{\rm HC})$	$5\times\mathbf{F}$	$3\times(\mathbf{Spin},\overline{\mathbf{Spin}})$	$N_{\rm HC}=10,14$	$\frac{5}{12}$, $\frac{5}{48}$	1/3	$N_{\rm HC}=10$	M7
	Pseudo-Real	Real	$\mathrm{SU}(4)/\mathrm{Sp}(4)\times\mathrm{SU}(6)/\mathrm{SO}(6)$				
$Sp(2N_{\rm HC})$	$4\times\mathbf{F}$	$6\times\mathbf{A}_2$	$2N_{\rm HC}\leq 36$	$\frac{1}{3(N_{\rm HC}-1)}$	2/3	$2N_{\rm HC}=4$	M8
$SO(N_{\rm HC})$	$4\times\mathbf{Spin}$	$6\times\mathbf{F}$	$N_{\rm HC}=11,13$	$\frac{8}{3}$, $\frac{16}{3}$	2/3	$N_{\rm HC}=11$	M9
	Complex	Real	$\mathrm{SU}(4)^2/\mathrm{SU}(4)\times\mathrm{SU}(6)/\mathrm{SO}(6)$				
$SO(N_{\rm HC})$	$4\times(\mathbf{Spin},\overline{\mathbf{Spin}})$	$6\times\mathbf{F}$	$N_{\rm HC}=10$	$\frac{8}{3}$	2/3	$N_{\rm HC}=10$	M10
$SU(N_{\rm HC})$	$4\times(\mathbf{F},\overline{\mathbf{F}})$	$6\times\mathbf{A}_2$	$N_{\rm HC}=4$	$\frac{2}{3}$	2/3	$N_{\rm HC}=4$	M11
	Complex	Complex	$\mathrm{SU}(4)^2/\mathrm{SU}(4)\times\mathrm{SU}(3)^2/\mathrm{SU}(3)$				
$SU(N_{\rm HC})$	$4\times(\mathbf{F},\overline{\mathbf{F}})$	$3\times(\mathbf{A}_2,\overline{\mathbf{A}}_2)$	$N_{\rm HC}\geq 5$	$\frac{4}{3(N_{\rm HC}-2)}$	2/3	$N_{\rm HC}=5$	M12
$SU(N_{\rm HC})$	$4\times(\mathbf{F},\overline{\mathbf{F}})$	$3\times(\mathbf{S}_2,\overline{\mathbf{S}}_2)$	$N_{\rm HC}\geq 5$	$\frac{4}{3(N_{\rm HC}+2)}$	2/3	/	

[JHEP 2202, 208]

[Ferretti (2014)]

[JHEP1511,201]

[Vecchi (2015)]

Additional model: SU(3) with 8x(F,F) [Appelquist,Ingoldby,Piai (2021)]

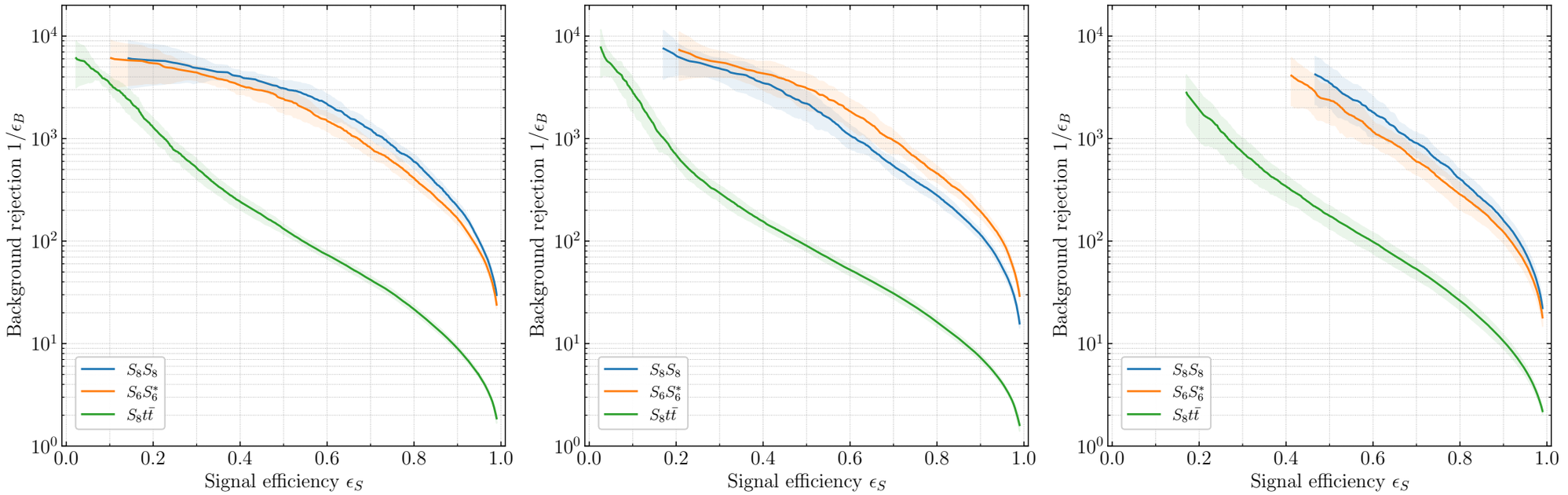
[JHEP1701,094]

Classification of colored resonances in these models

	Models	χ (R, Y, B)	π	$\mathcal{V}^\mu$	$\mathcal{A}^\mu$	Ψ	di-quark
C1	M1-2	$(\mathrm{R}, -\frac{1}{3}, \frac{1}{6})$	8_0, $6_{-2/3}$	8_0, 1_0, $3_{2/3}$	8_0, $6_{-2/3}$	8, 1, 3, 6	none
C2	M3-4, M8-11	$(\mathrm{R}, \frac{2}{3}, \frac{1}{3})$	8_0, $6_{4/3}$	8_0, 1_0, $3_{-4/3}$	8_0, $6_{4/3}$	3	$\pi_6, \mathcal{V}_3^\mu, \mathcal{A}_6^\mu$
C3	M5	$(\mathrm{Pr}, -\frac{1}{3}, \frac{1}{6})$	8_0, $3_{2/3}$	8_0, 1_0, $6_{-2/3}$	8_0, $3_{2/3}$	8, 1, 3, 6	none
C4	M6-7	$(\mathrm{C}, -\frac{1}{3}, \frac{1}{6})$	8_0	8_0, 1_0	8_0	8, 1, 3, 6	none
C5	M12	$(\mathrm{C}, \frac{2}{3}, \frac{1}{3})$	8_0	8_0, 1_0	8_0	3	none

Table 2: Properties of the spin-0 (π), spin-1 ($\mathcal{V}^\mu$, $\mathcal{A}^\mu$) and spin-1/2 (Ψ) lightest resonances in the 12 models, grouped in 5 classes. Each class is determined by the properties of the χ species, listed in the second column by irrep type (R for real, Pr for pseudo-real and C for complex). For the resonances, the colours indicate the baryon numbers, with black for $B = 0$, red for $B = \pm 1/3$ and blue for $B = \pm 2/3$. In the last column we indicate the bosons that can decay into a di-quark state (tt).

[2404.02198]



(a) S_8S_8 (b) $S_6S_6^*$ (c) $S_8t\bar{t}$

Figure 9. Receiver operator characteristic (ROC) curves comparing networks trained on S_8S_8, $S_6S_6^*$, $S_8t\bar{t}$, tested across different signal types.