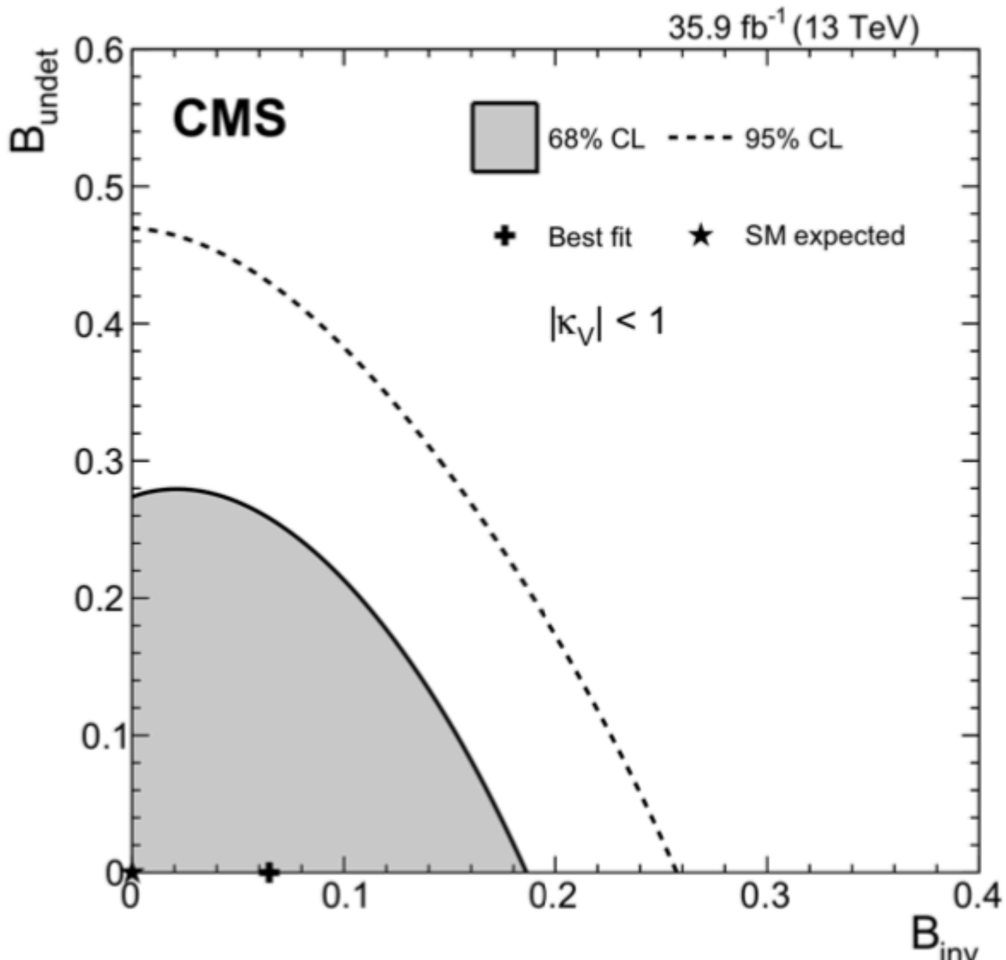
Search for a Light Pseudoscalar Higgs Boson with Boosted Topologies at CMS

Grace Haza
University of California, Davis

New Perspectives @ Fermilab
20/21 July 2020

search for exotic Higgs decays

- meta-analysis of SM Higgs results by CMS
- remaining uncertainties in measured BFs leave room for BSM couplings
- direct searches for exotics and Higgs decays to exotics well motivated



CMS: EPJ C 79, article 421

2 Higgs Doublet Model

- Add additional doublet to SM Higgs sector (eg. MSSM)
- Lightest scalar in 2HDM compatible with H(125)
- Characterize model via doublet (Φ_1 and Φ_2) couplings to fermions

	Type-1	Type-2	Type-3 (lepton-specific)	Type-4 (flipped)
Up-type quarks	Φ_2	Φ_2	Φ_2	Φ_2
Down-type quarks	Φ_2	Φ_1	Φ_2	Φ_1
Charged leptons	Φ_2	Φ_1	Φ_1	Φ_2

CMS: arXiv:1701.02032

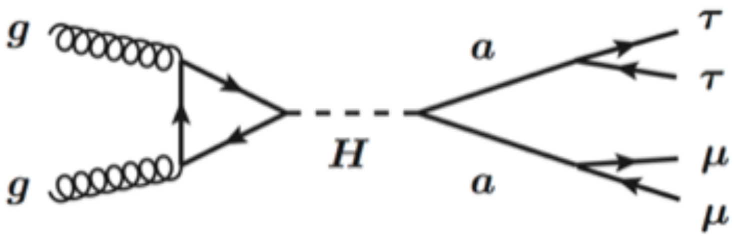
- Type 3 - enhanced couplings to leptons especially at high tan β
- tan β= v_2/v_1 where v_1 and v_2 are the vacuum expectation values for Φ_1 and Φ_2

2HDM+singlet

- extend 2HDM to add additional singlet field: 2HDM+S
  - theoretically motivated (NMSSM)
  - neutral scalars (h_1, h_2, h_3)
  - neutral pseudoscalars (a_1, a_2)
  - charged $H^{\pm}$

- in this case, SM-like H(125) is one of the neutral scalars
- naturally light pseudoscalars
- can search for direct production of h -> aa (h_1 -> a_1 a_1)
- possibility of Higgs bosons with masses > 125 GeV

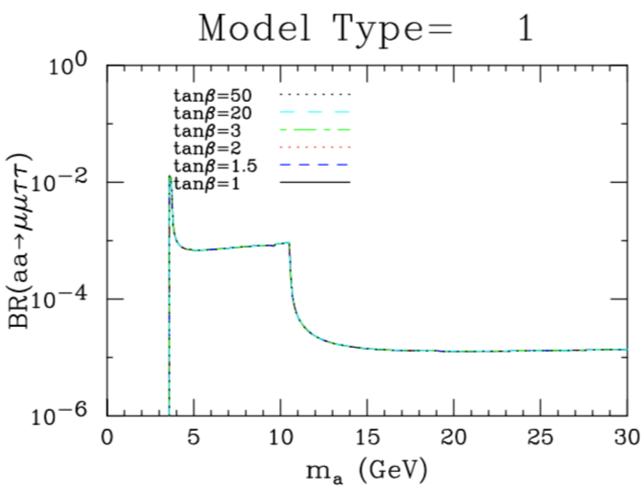
boosted topology

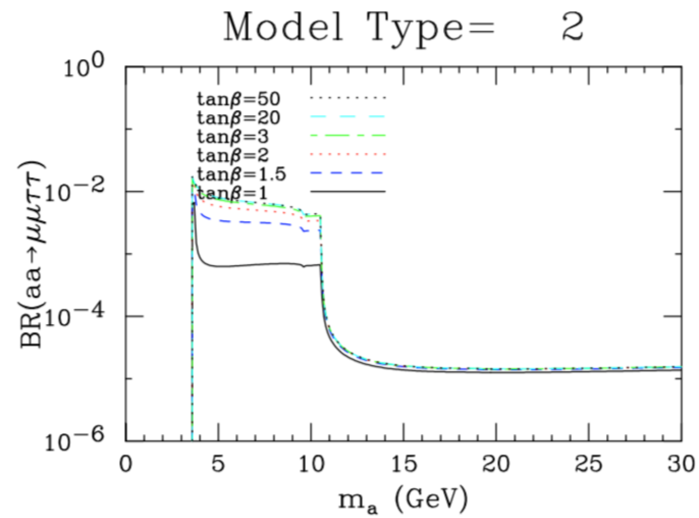


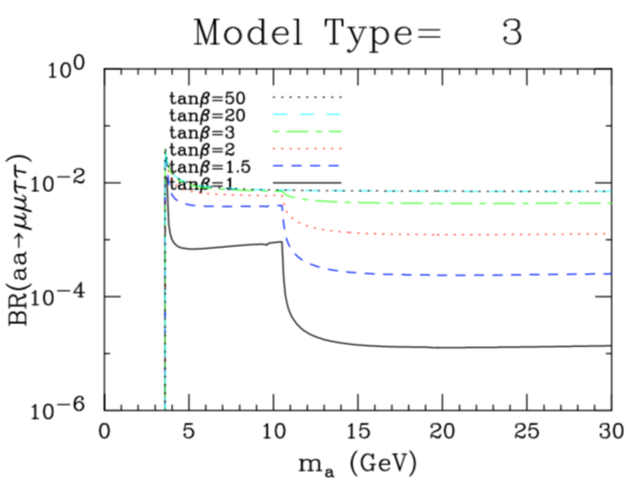
H mass (GeV)	a mass (GeV)
125	4 - 21
250	5 - 20
500	5 - 25
750	10 - 30
1000	10 - 50

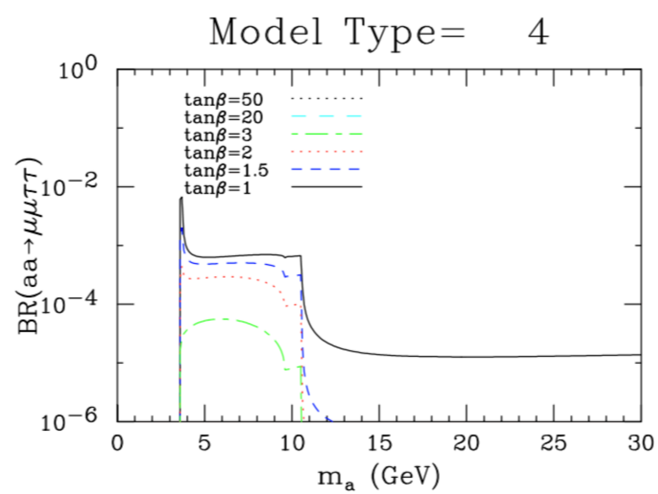
- difference between H mass and a mass gives boosted topology
- leads to collimated tau decays
- μμττ channel provides high branching fraction (ττ) and high efficiency (μμ)

branching fractions



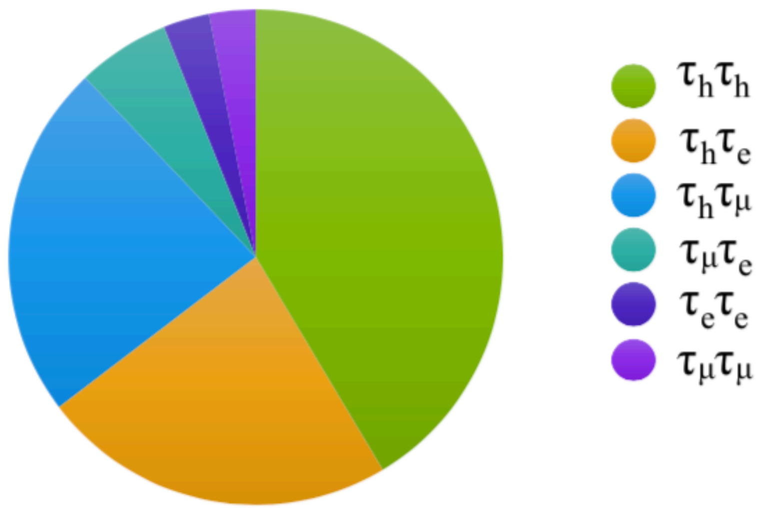




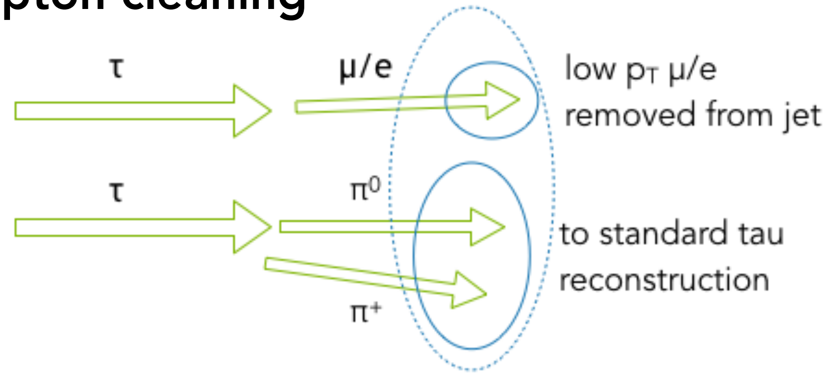


plots by John Gunion, UC Davis

channels



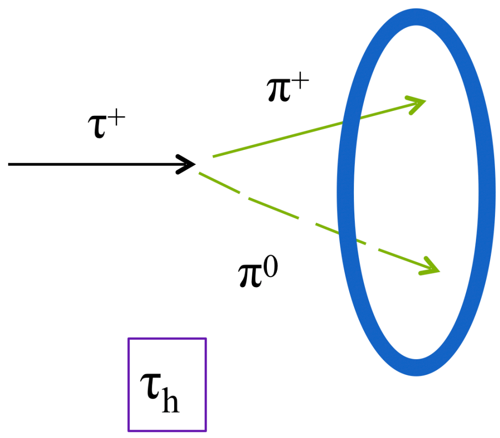
Lepton cleaning

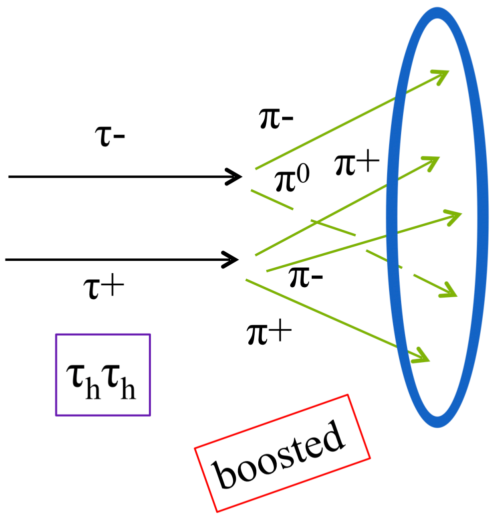


- $\tau_\mu\tau_{had}$ and $\tau_e\tau_{had}$: cleaning lepton object from tau cone
- $\tau_{had}\tau_{had}$: machine learning techniques being applied
- $\tau_\mu\tau_\mu$, $\tau_e\tau_e$, $\tau_\mu\tau_e$: use standard e/μ ID
- 2016 analysis: $\tau_\mu\tau_{had}$ channel only

one and two hadronic tau decays

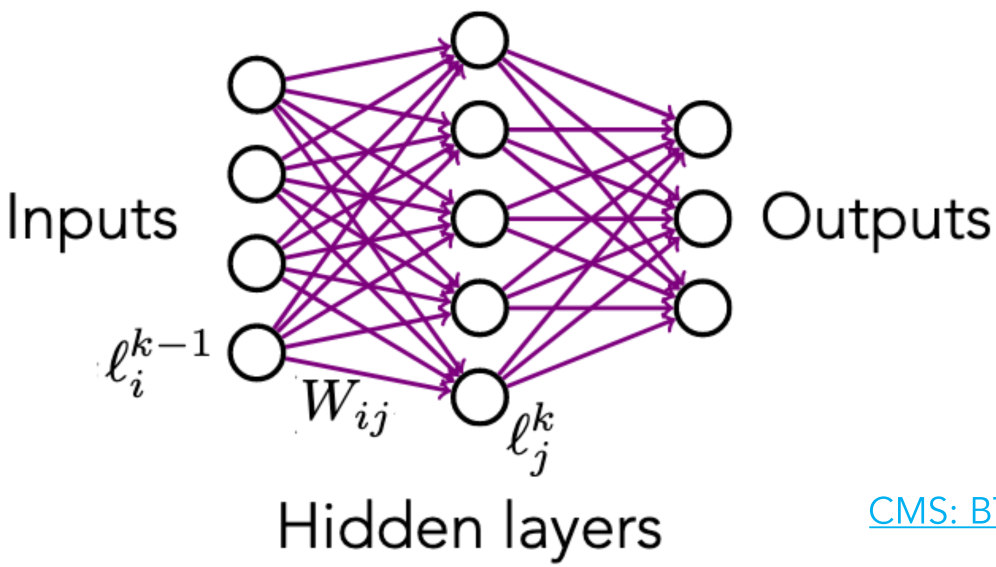
- $\tau_{had}\tau_{had}$ channel important but high level of hadronic backgrounds





machine learning: neural net

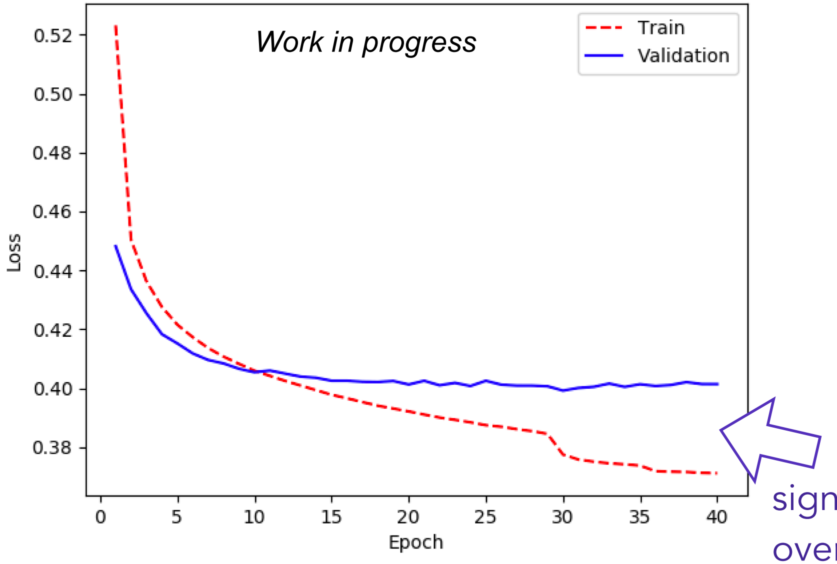
- inputs: jet, charged hadron, neutral hadron, electron, muon, photon variables
- outputs: number related to probability of being a certain type of jet (classification among overlapping di tau, light jet, or b jet)
- W_{ij} are the weights updated during training (iterative process where neural net is improved based on success and failure on representative sample)



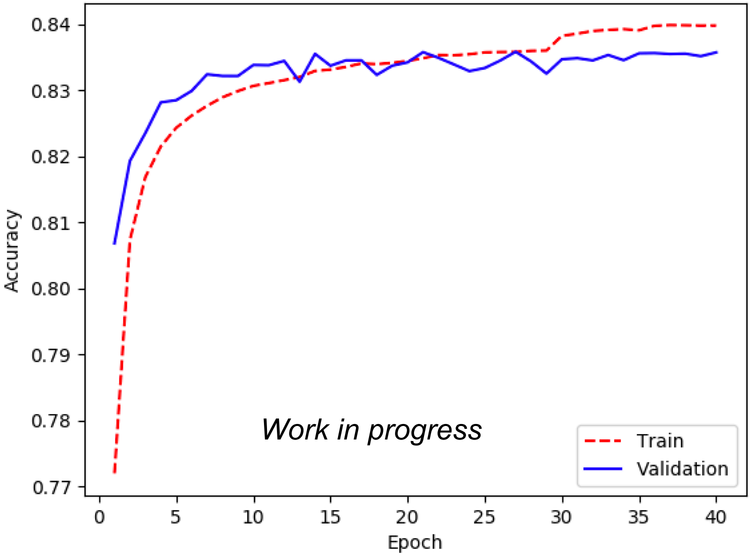
CMS: BTV-16-002

machine learning procedure

- during training, monitor loss function (penalizes false positives and false negatives, weighted by confidence of model) and minimize
- validation: part of training sample set aside and not actually trained on, monitored to avoid divergence of validation and training performance

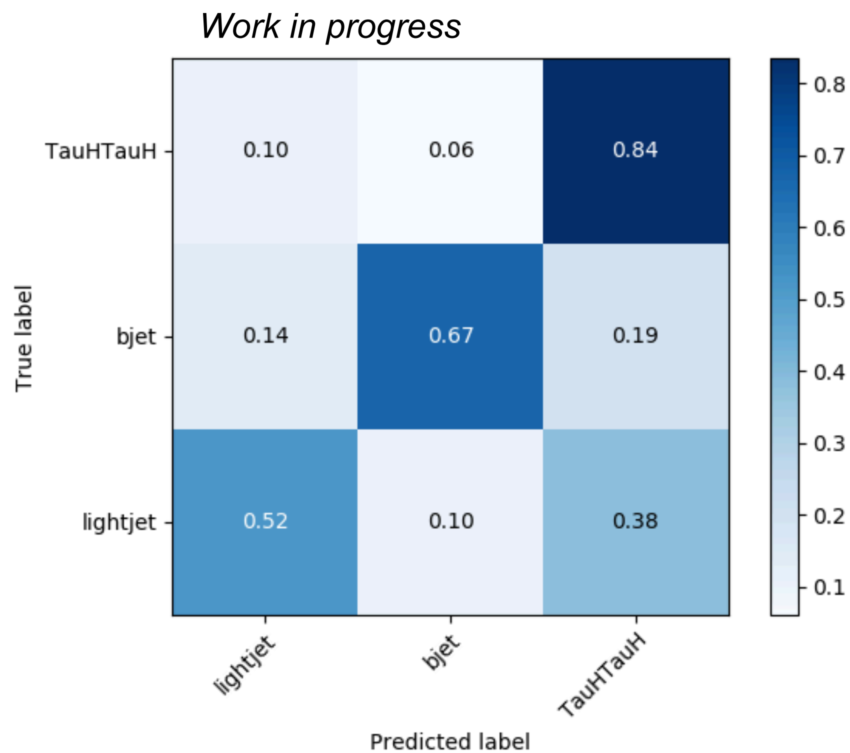


sign of overtraining

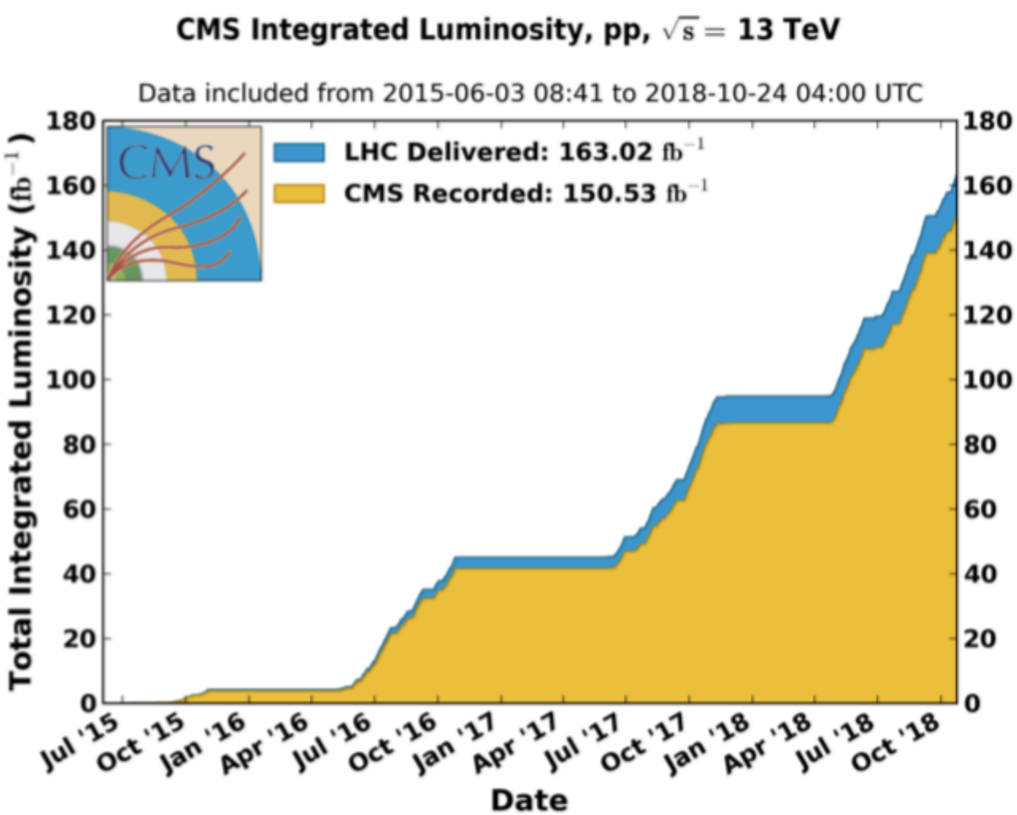


machine learning: mass decorrelation

- searching for a range of exotic Higgs and pseudoscalar masses
- Monte Carlo samples available for a finite number of discrete Higgs and pseudoscalar mass pairs
- mass decorrelation: prevent the net from "learning" to look for specific masses



outlook



- search for H -> aa -> μμττ
- implement mass decorrelation in $\tau_{had}\tau_{had}$ channel
- working on other channels in parallel including cleaning technique
- result on Full Run 2 dataset to discover new particle (or improve observed limits)

backup

theoretical motivation

- BSM final state likely to have “non-negligible” branching fraction
- 2HDM: if $m_a < m_h /2$ then 2HDM, potential must be tuned
- MSSM: pseudoscalar, but it must be > 95GeV
- NMSSM: two pseudoscalar higgs bosons, one (likely) lighter than 125 GeV
- NMSSM is special case of 2HDM+S

pseudoscalar decays

- branching ratios of a depends on lepton masses

$$\frac{\Gamma(\mathrm{a} \to \mu^+\mu^-)}{\Gamma(\mathrm{a} \to \tau^+\tau^-)} = \frac{m_\mu^2\sqrt{1-(2m_\mu/m_\mathrm{a})^2}}{m_\tau^2\sqrt{1-(2m_\tau/m_\mathrm{a})^2}}.$$

CMS: arxiv:1701.02032

m (τ) = 1.776 GeV

m (μ) = 0.106 GeV

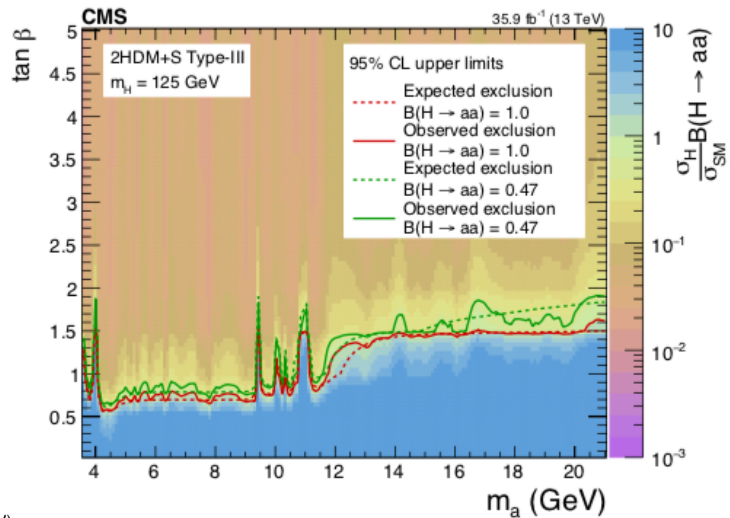
= 0.0077 for m_a = 4 GeV

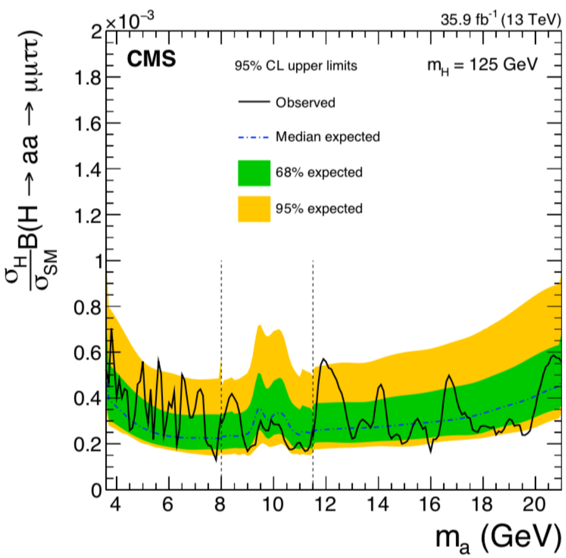
- decay to τ favored

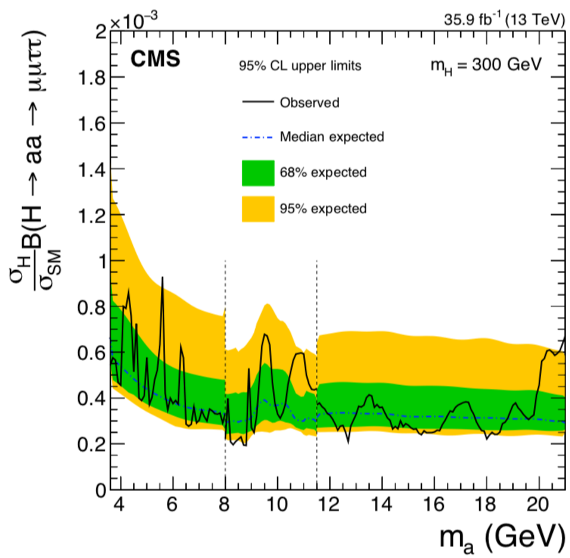
= 0.0036 for m_a = 20 GeV

previous results

- 2016 data only
- $\tau_{\mu}\tau_{had}$ channel only

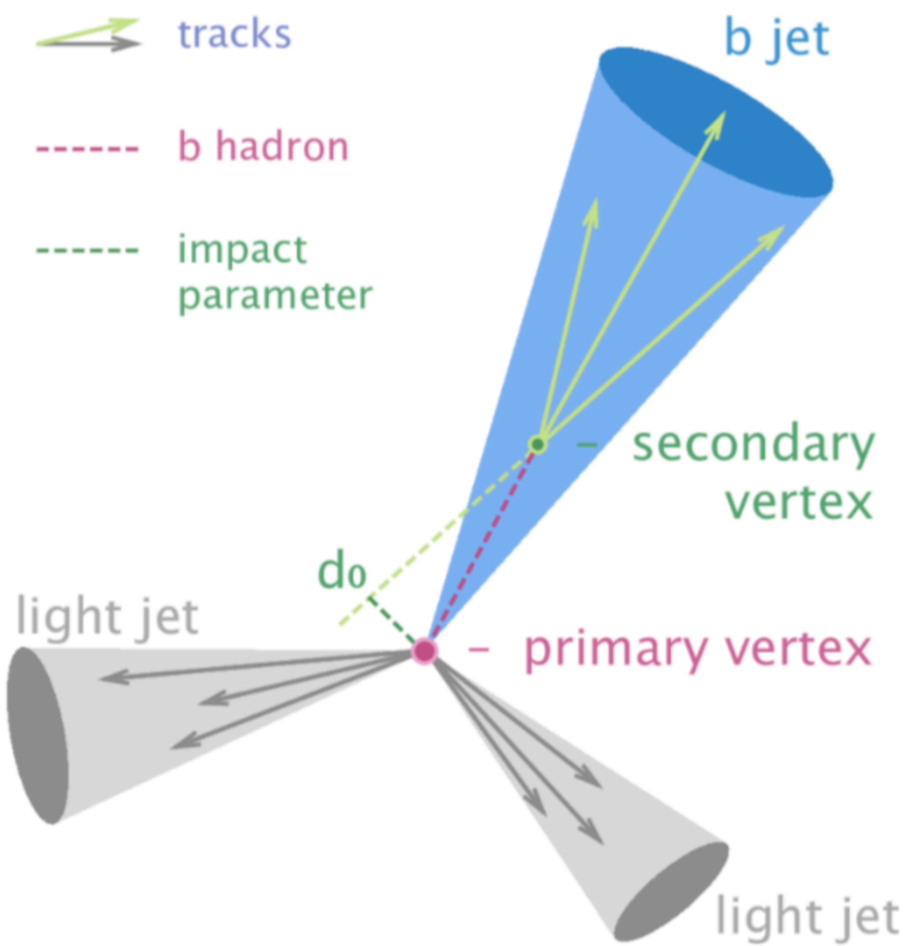




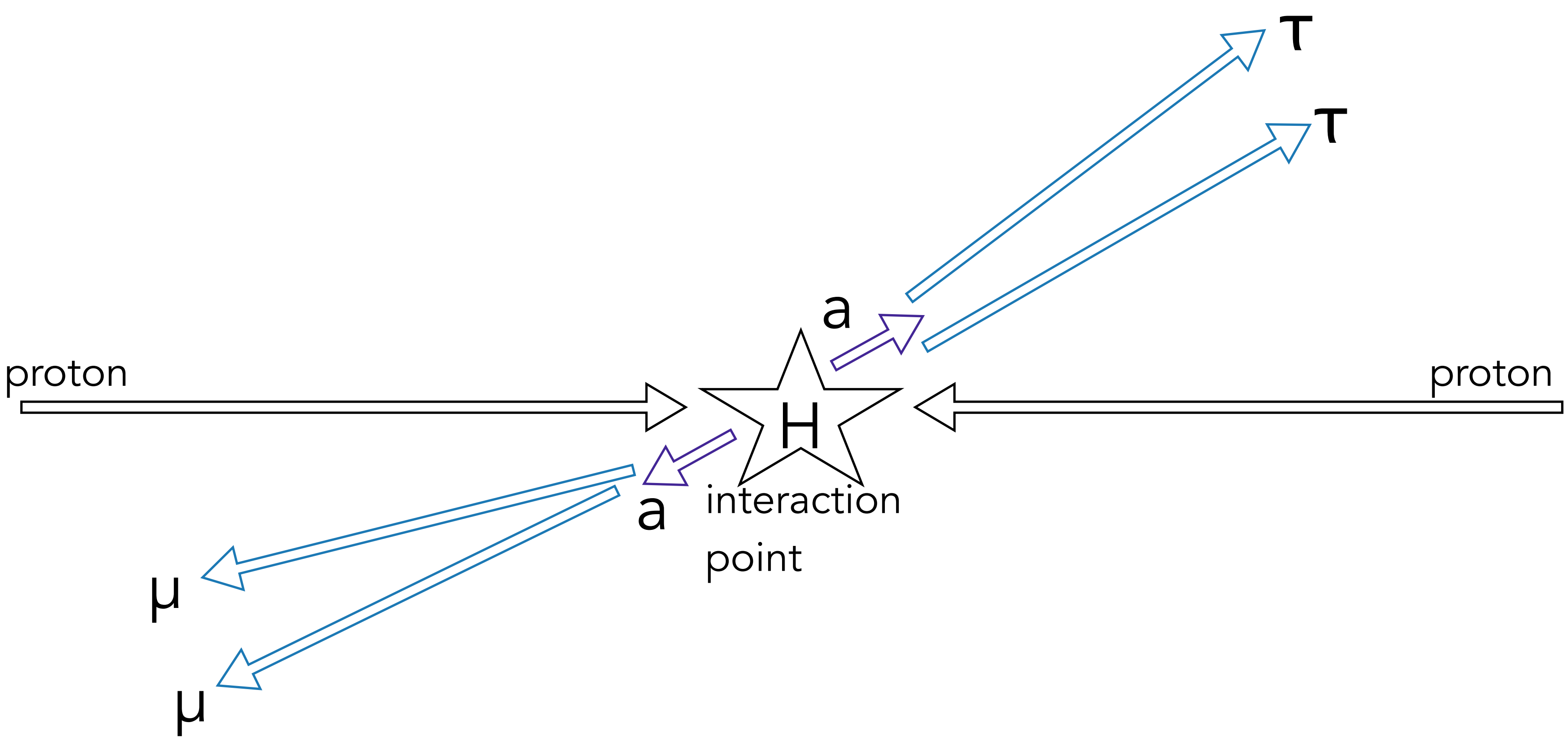


CMS: arXiv:2005.08694v1

light flavor and heavy flavor jets

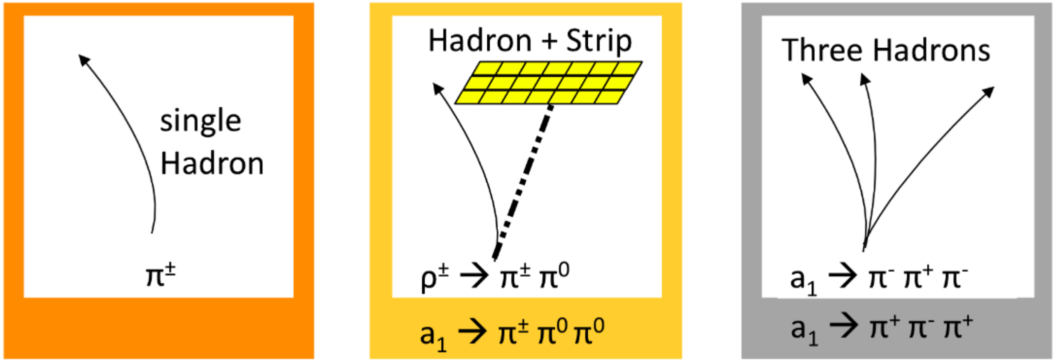


event in detector



Hadron Plus Strips

- begins with particle flow constituents in AK4 jet
- strips in electromagnetic calorimeter are used to reconstruct π^0 -> γγ
- strips combined with charged particle flow hadrons to determine decay mode (number of charged particles in final state)

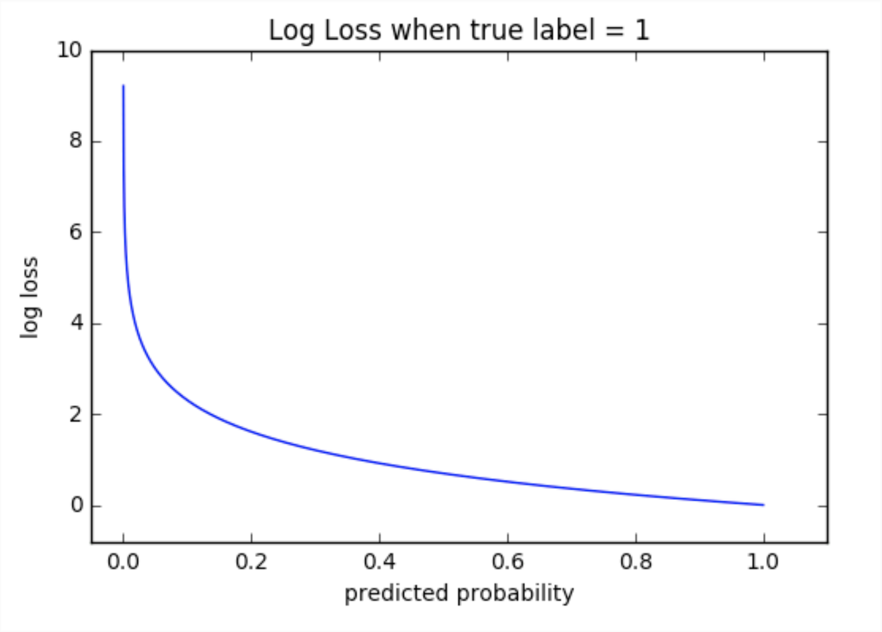


machine learning: hyperparameters

- batch size = number of training samples processed before parameters of neural net (weights) updated

- epochs = number of complete passes through the entire training set

cross entropy loss function and accuracy



Loss for each category label per observation:

$$-\sum_{c=1}^{M} y_{o,c} \log(p_{o,c})$$

- M = number of categories
- y = 0 or 1 if class label *c* is correct for sample *o*
- p = predicted possibility *o* is label *c*

source

- accuracy = (number of true positives + number of true negatives) / total number of predictions