MVA Techniques in ttH(γγ)

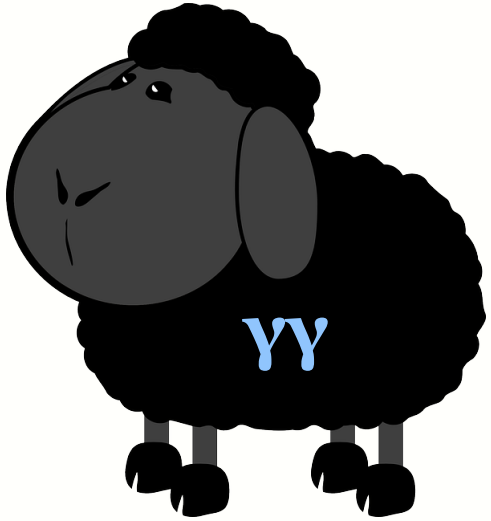
Rachel Hyneman[1]
Tom Schwarz[1]
Haichen Wang[2]

[1]University of Michigan, [2]LBNL

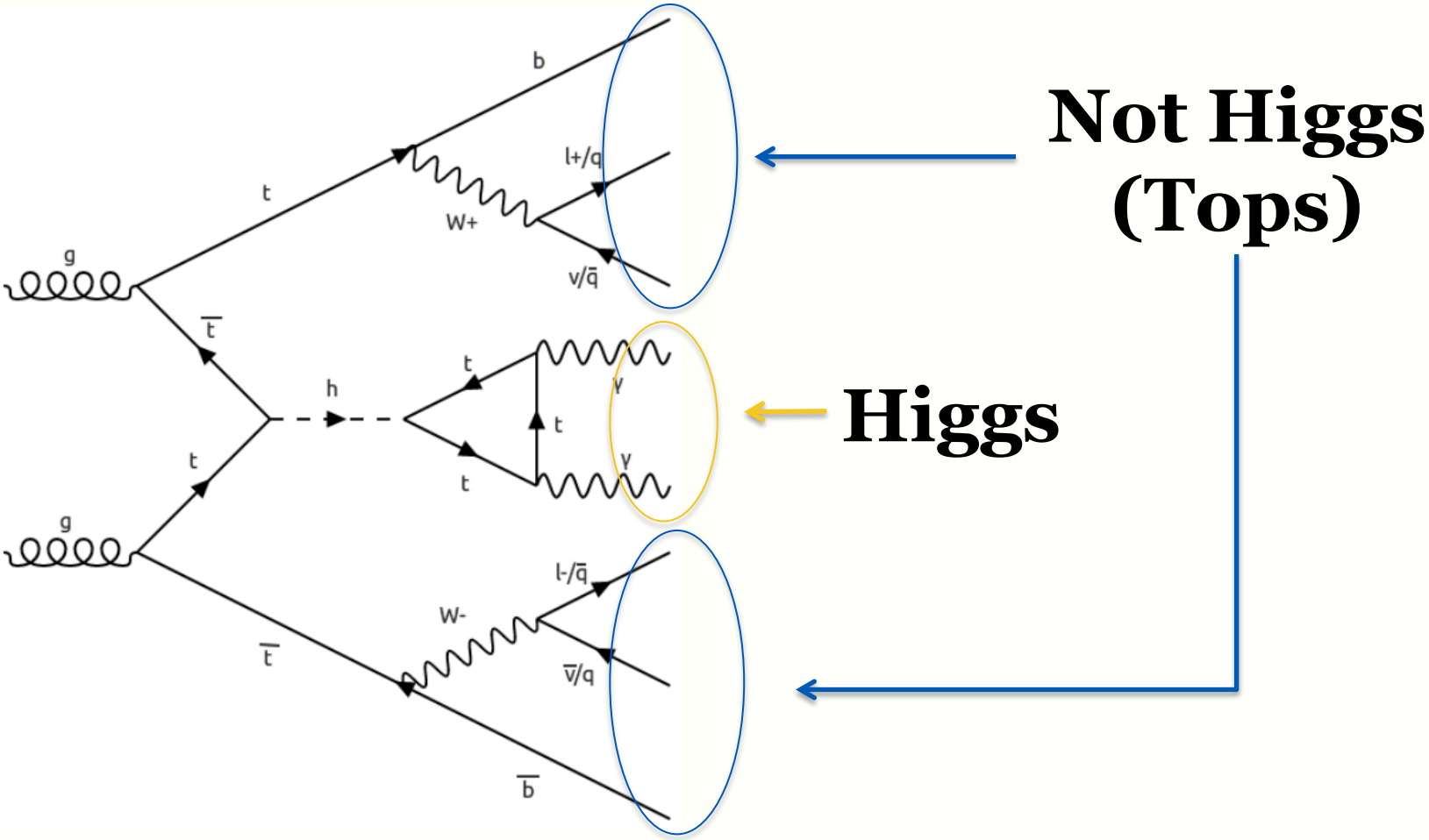
*Some of this work supported by NSF

Analysis Summary

- Reminder: this is a diphoton analysis with jets, not a top analysis with photons
  - Can reconstruct Higgs well from the two photons
  - Information on tt system extracted from jets, MET, and leptons
  - Base our background estimation on smoothly falling invariant diphoton mass spectrum
- BDT for event selection
- No explicit top reconstruction yet



What is Special About ttH(γγ)?



Not Higgs (Tops)

Higgs

We have no object ambiguity.

Possible BDT Approach

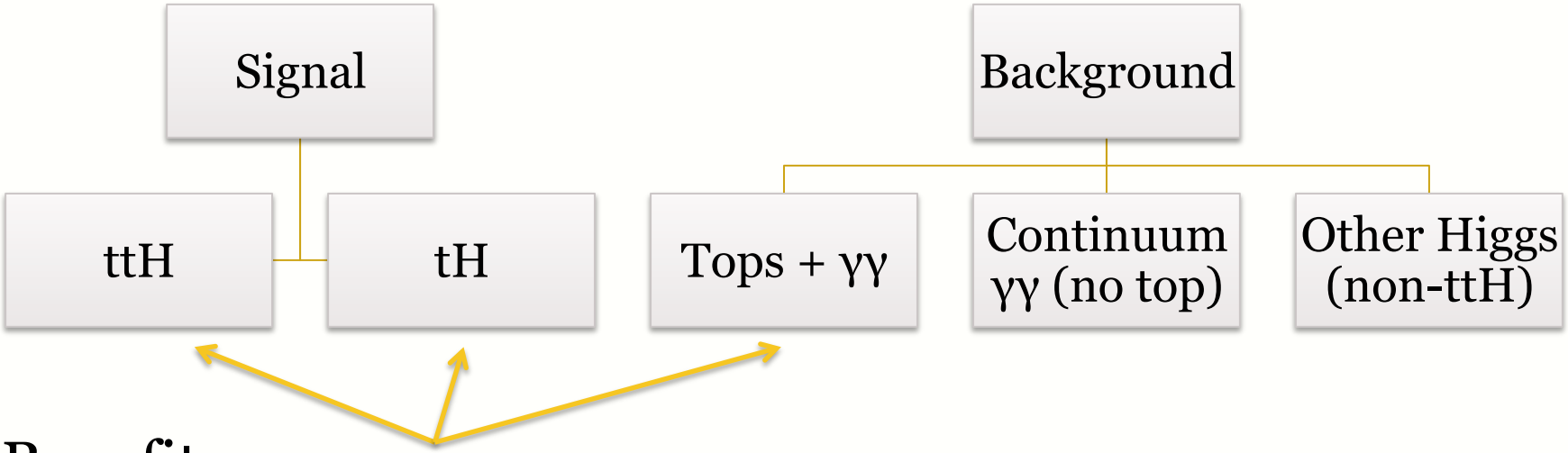
- Can use separate BDT for "hadronic" (no reconstructed leptons) and "leptonic" (1 or more reconstructed leptons) event types
- We could gain significance by training on both top-related (jet, MET, lepton) and higgs-related (photon) information as inputs into the same BDT
  - Train on MC ttH signal and data CR events (failing photon ID/Isolation cuts)
- Use expected significance to define BDT category boundaries

How Can We Do Better?

- We need a good control region to use for BDT training
  - Top-related and photon kinematics uncorrelated
  - Can veto photon variables (ID/Isolation) to get CR with very similar top-related kinematics
- Secondary concern: does training on photon variables bias the diphoton invariant mass?
  - Important, since we fit this to define our background in signal region
  - Note: can cross-check this easily

Possibility: Top Reconstruction

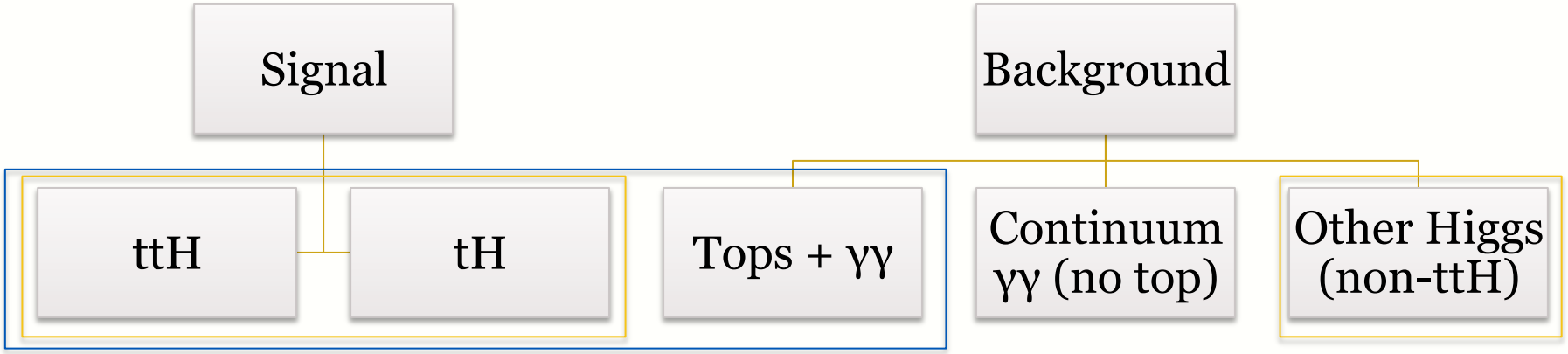
- We can borrow insights from our other top+Higgs friends



- Benefits:
  - Already done in many other analyses, can benefit from this work
  - Could be used to better understand tt kinematics
- Challenges:
  - Photon-based analyses don't save as much detailed jet info which could be used for top reconstruction

Possibility: 2D BDT Approach

- Idea: train one BDT on Higgs kinematics, and another on top kinematics



- "Top BDT"
  - Don't use photon variables here
- "Higgs BDT"
  - Can use Higgs signal and diphoton continuum MC
    - Note: γγ MC is not always the best, is this a problem?

Beyond Classification

- Another problem: we have a very poor understanding of our backgrounds

Can model this OK

Not so good at modeling this

- Most MVA techniques require decent training statistics and samples
  - Our MC is maybe not the best here… how can we get around any biases here?