Systematically Searching for New Physics at the LHC

Daniel Whiteson
UC Irvine

UC Davis HEP Seminar, April 2014

Outline

I. Dark Matter

II. Topological Models

III. Deep networks

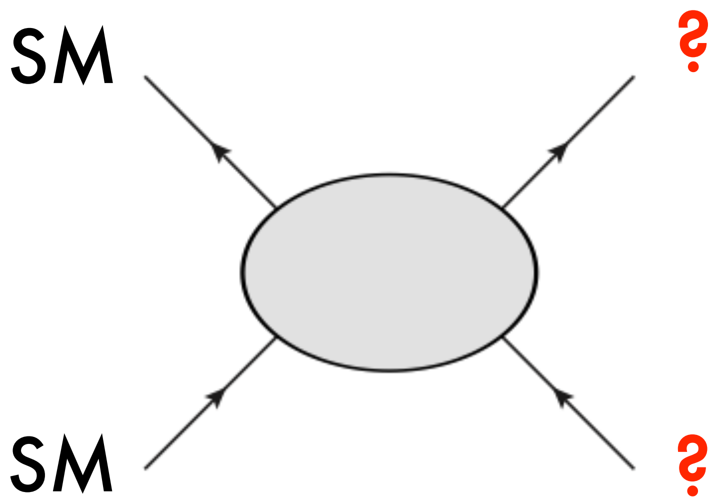
What do we know?

unknown unknown

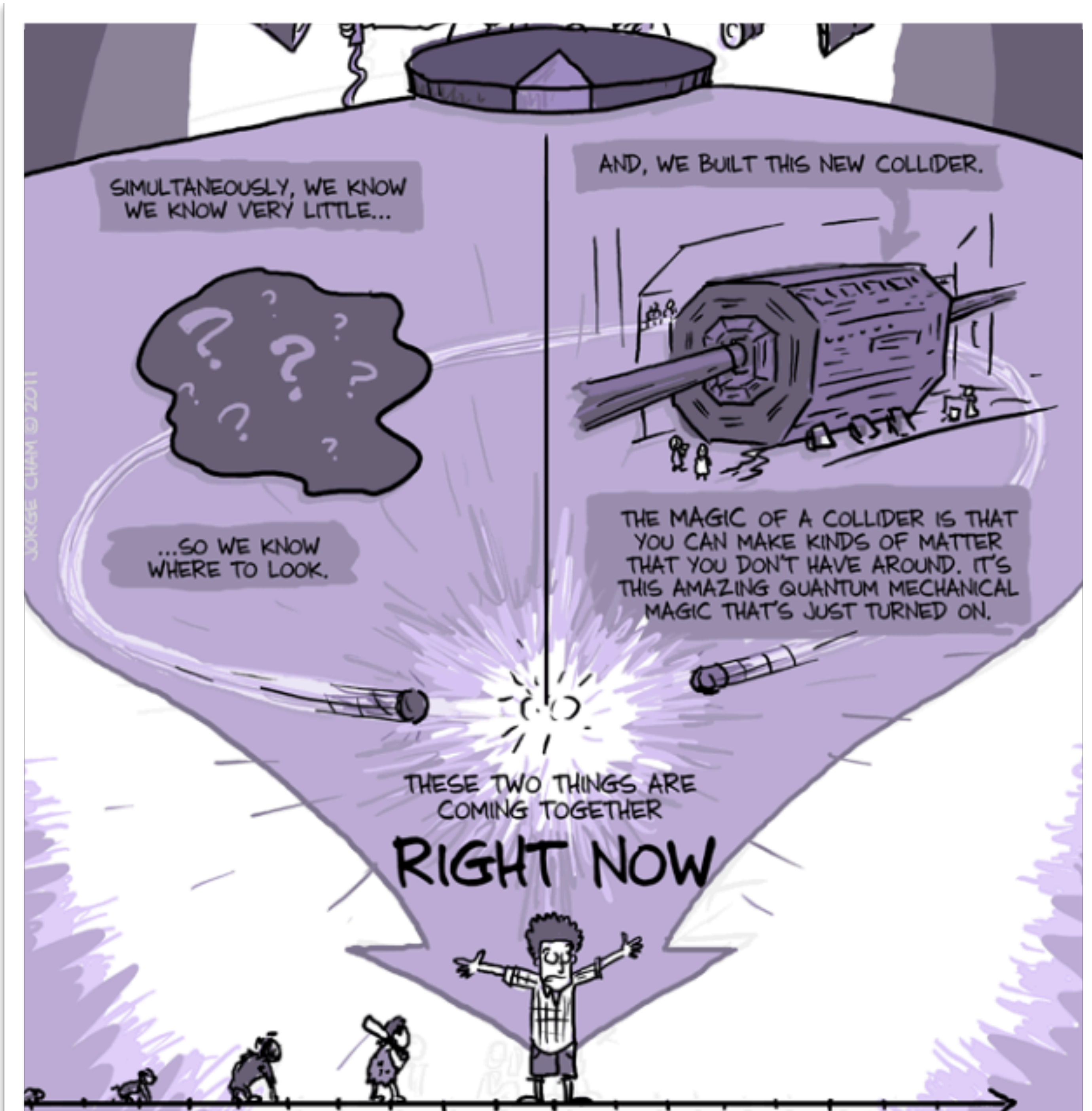
known unknown

known known

Colliders: true alchemy



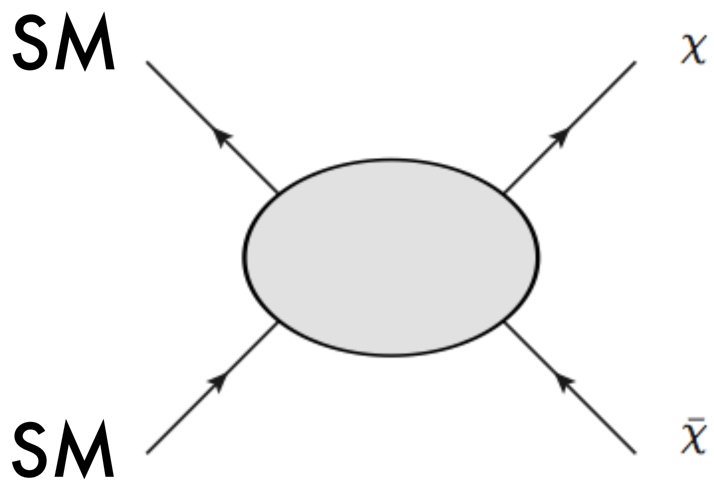
We can create new forms of matter, *even if we have* *little or no idea* *of what we are looking for!*


SIMULTANEOUSLY, WE KNOW WE KNOW VERY LITTLE...
...SO WE KNOW WHERE TO LOOK.
AND, WE BUILT THIS NEW COLLIDER.
THE MAGIC OF A COLLIDER IS THAT YOU CAN MAKE KINDS OF MATTER THAT YOU DON'T HAVE AROUND. IT'S THIS AMAZING QUANTUM MECHANICAL MAGIC THAT'S JUST TURNED ON.
THESE TWO THINGS ARE COMING TOGETHER
RIGHT NOW
JORGE CHAM © 2011

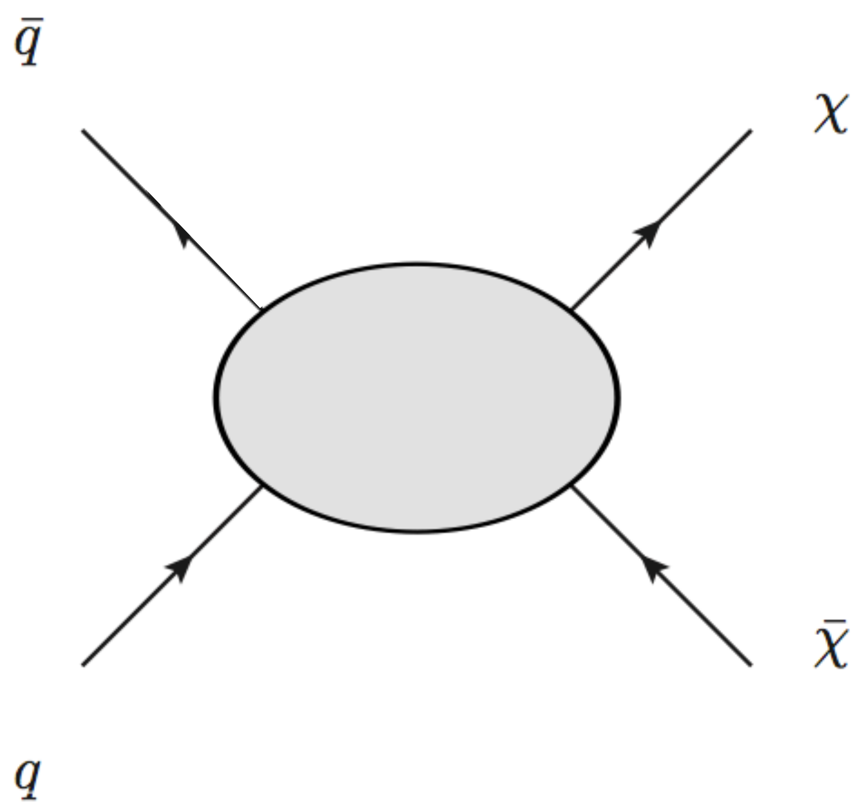
Interactions

<u>Important assumption:</u>

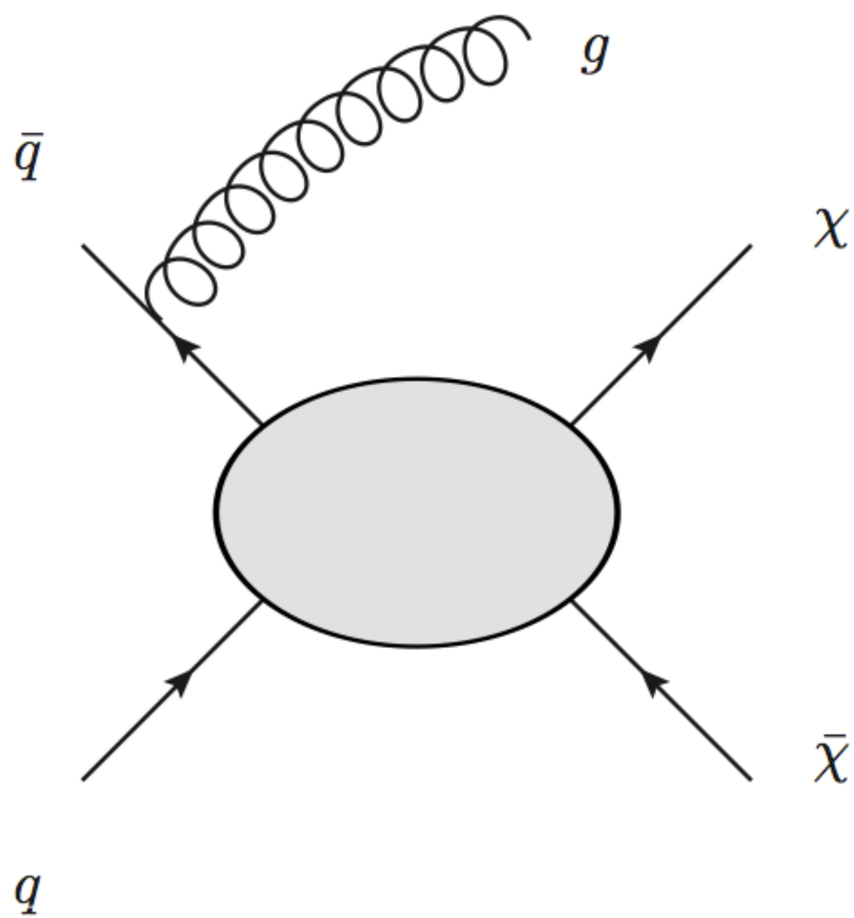
Requires **some** interaction with SM



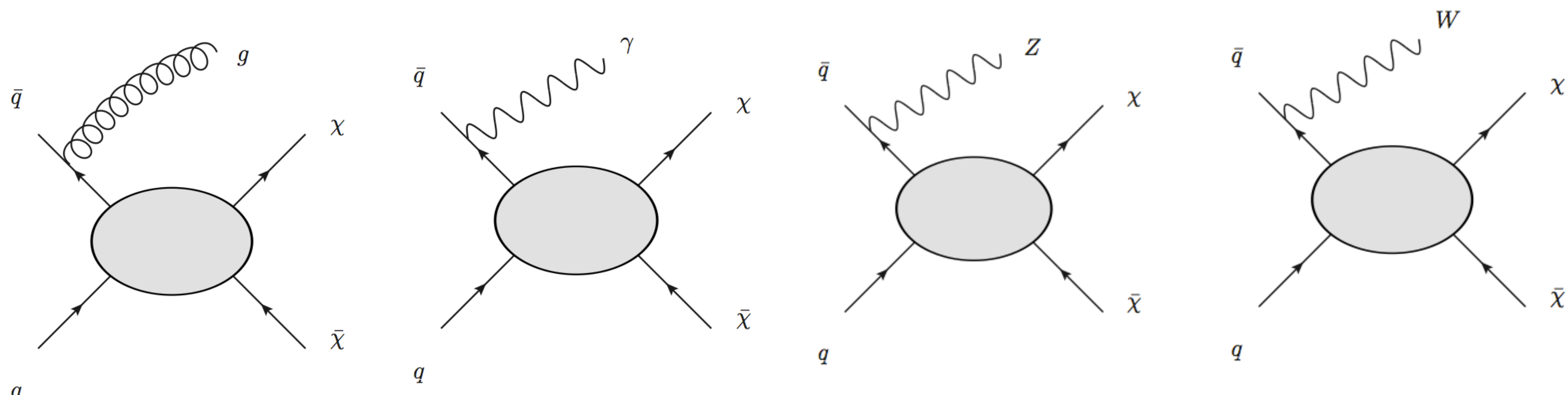
DM @ Colliders

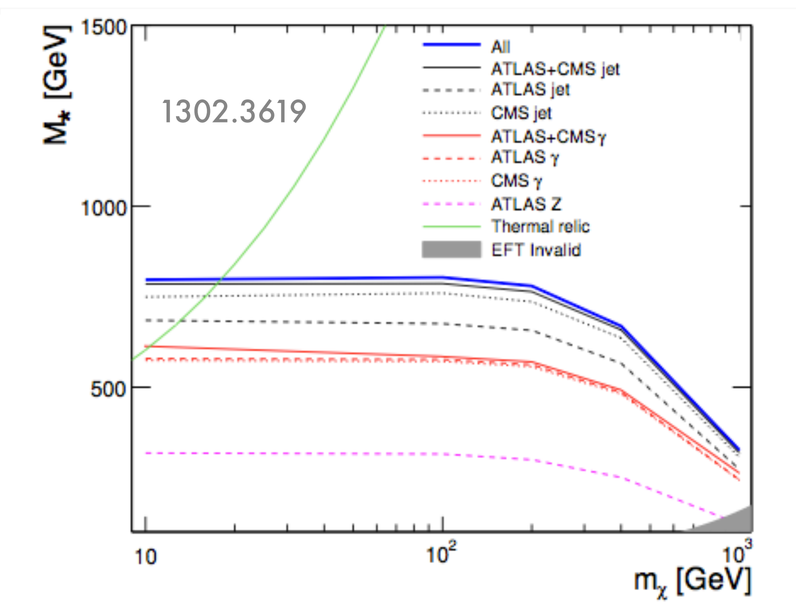


DM @ Colliders



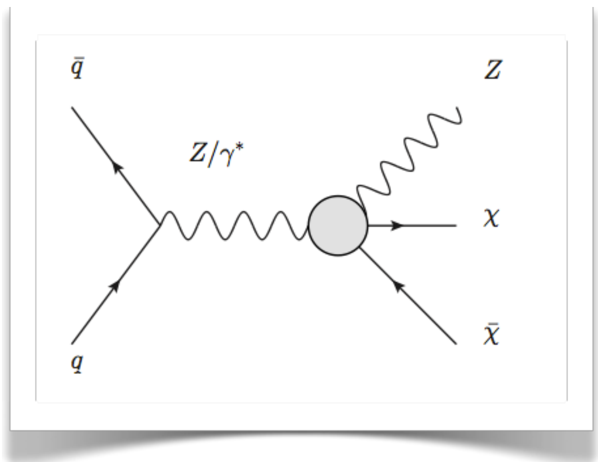
Look everywhere





Mono-jet most powerful for qqXX

Each mode has unique models where it is a discovery mode.



Outline

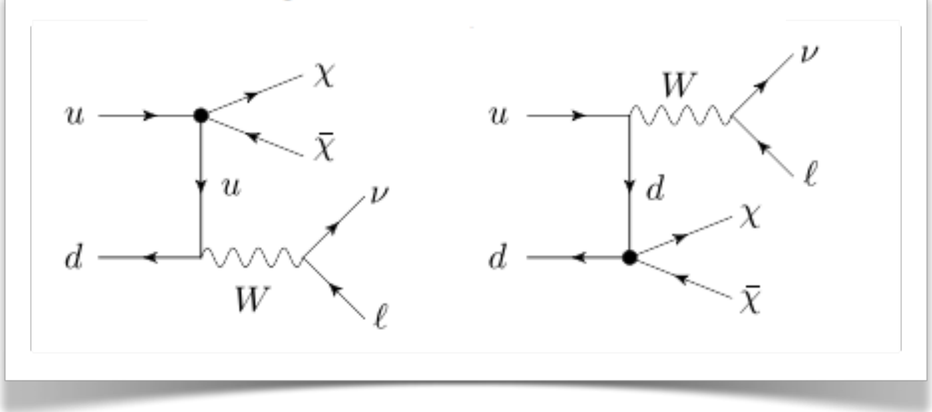
A. Mono-W
B. Mono-Z
C. Mono-Higgs

Mono-W

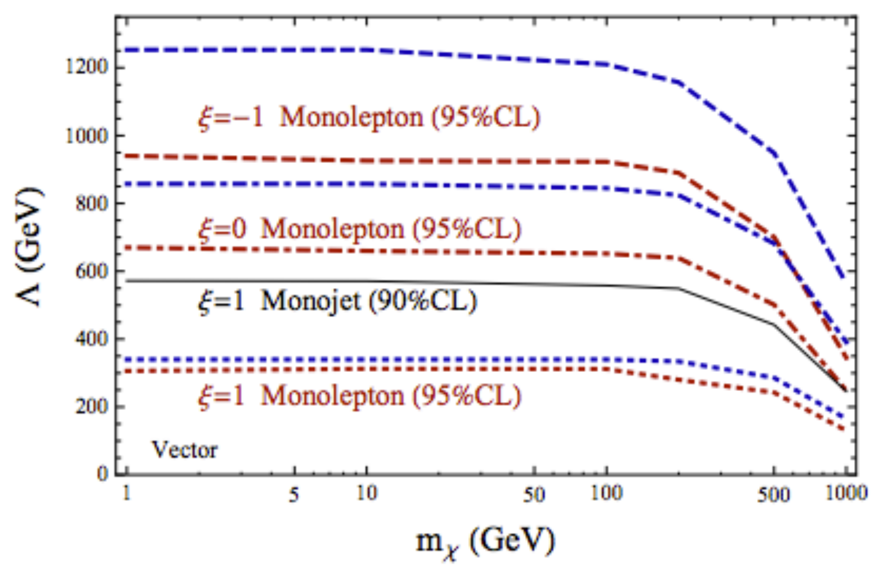
Mono-W theory

Searches with Mono-Leptons

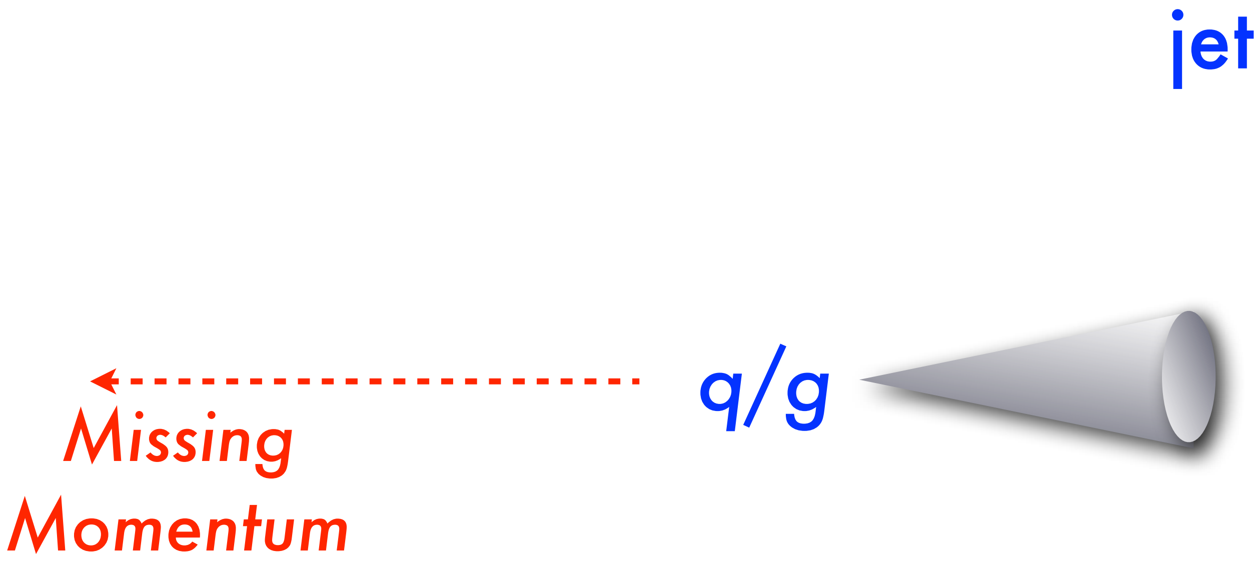
Yang Bai[a,b] and Tim M.P. Tait[c]



1208.4361

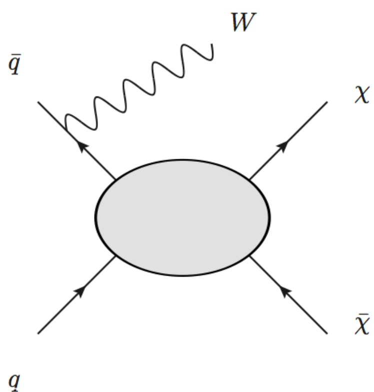


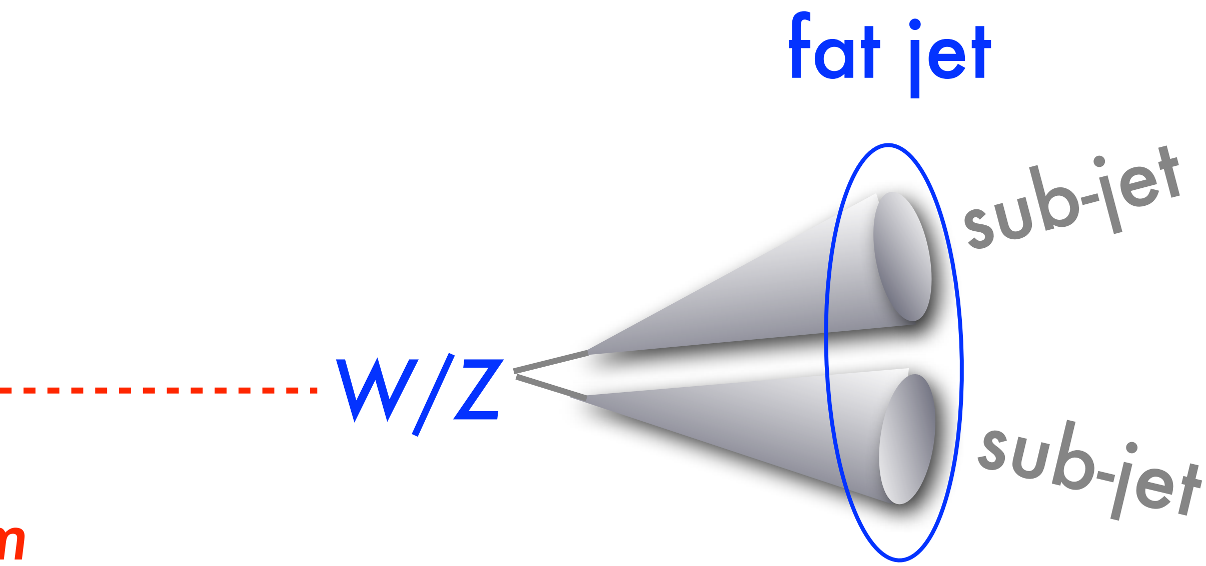
Mono-jet



Mono-heavy jet

1309.4017 (PRL)



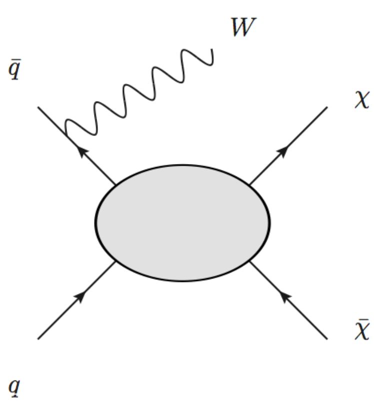


Missing
Momentum

 Ning Zhou, UCI

mono-W, etc

1309.4017 (PRL)



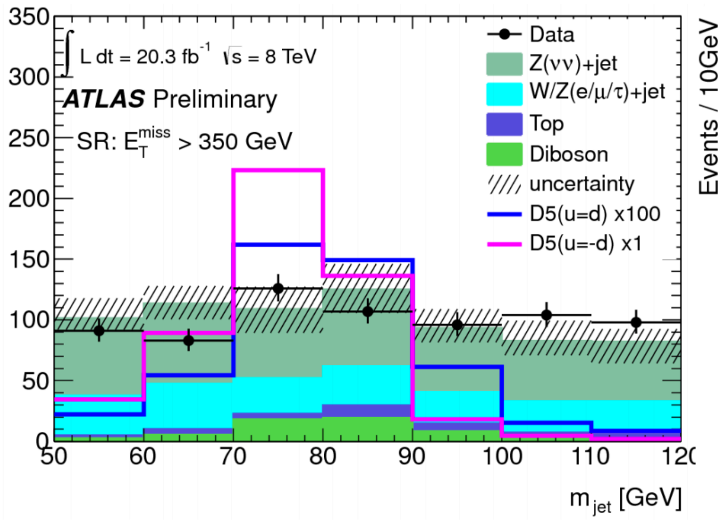
Fat jet p_T >250

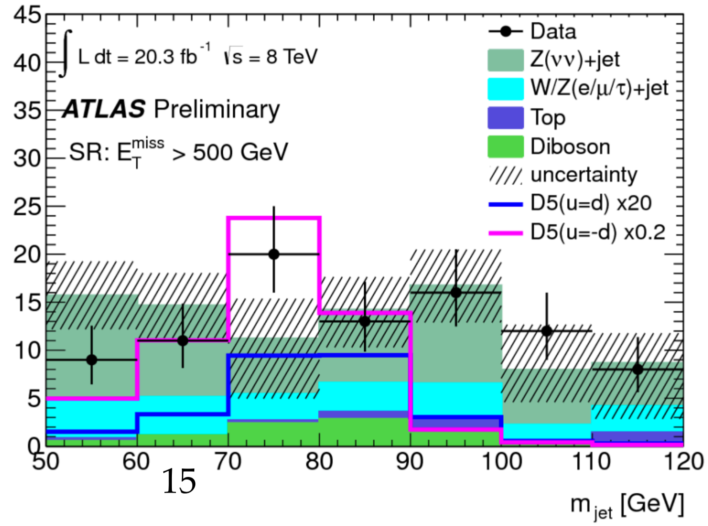
two subjets giving m_{jet} =[50,120]

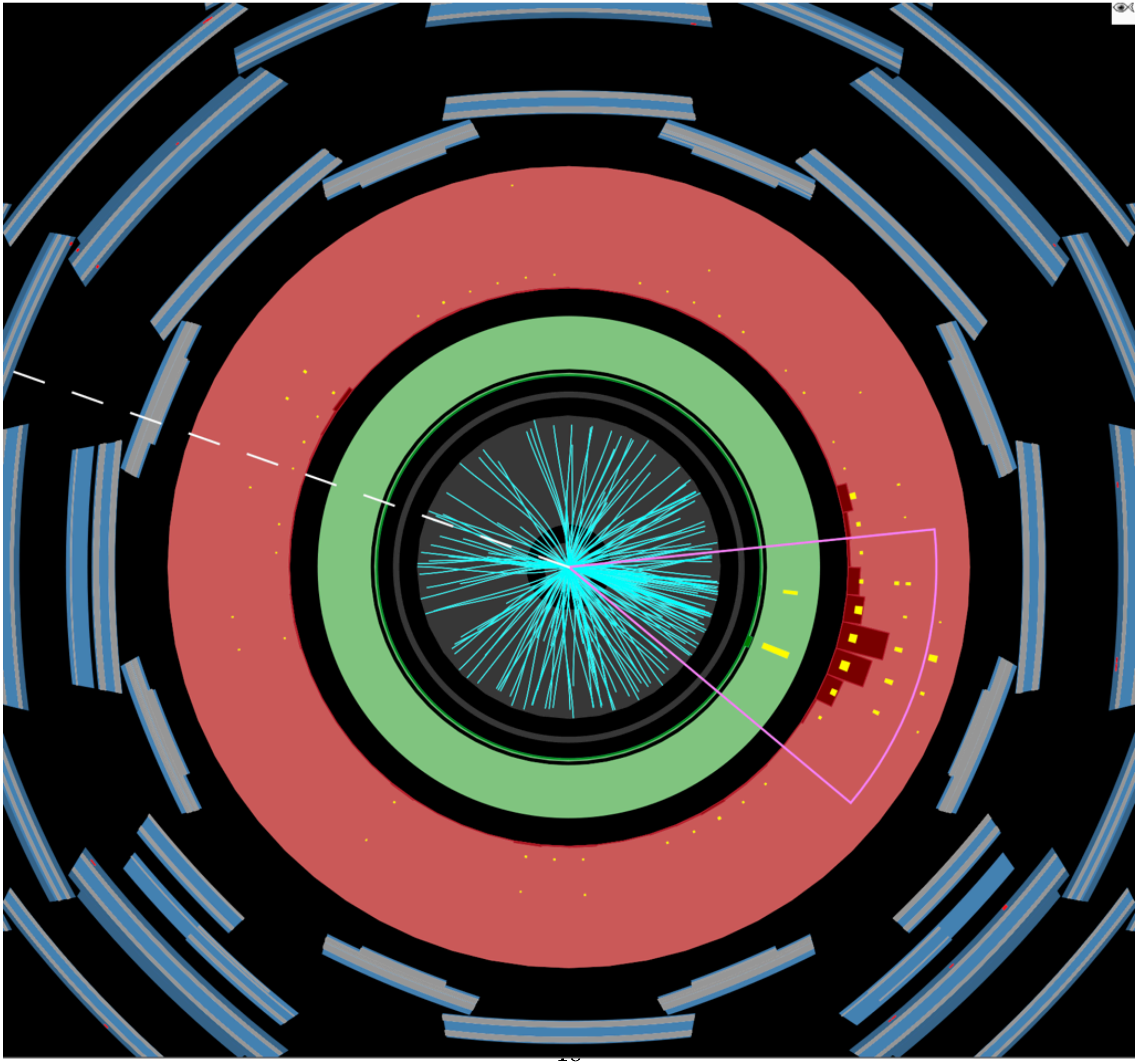
No e,mu,gamma

<= 1 additional narrow jets

MET >350 or 500

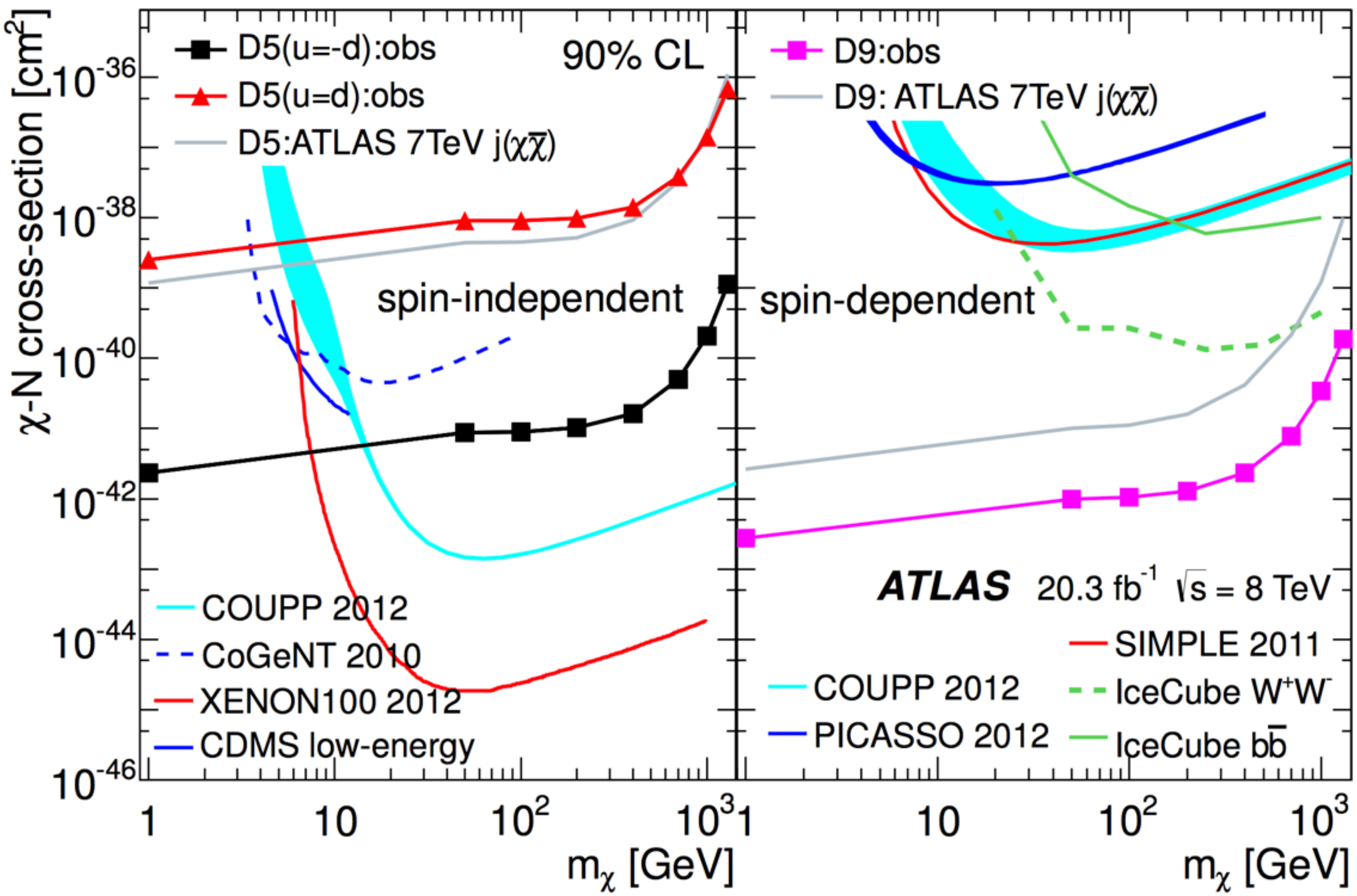




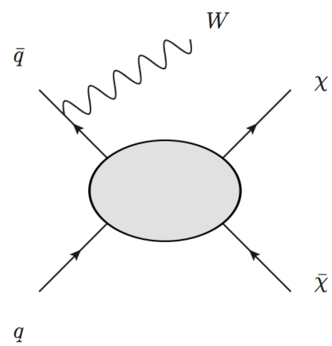


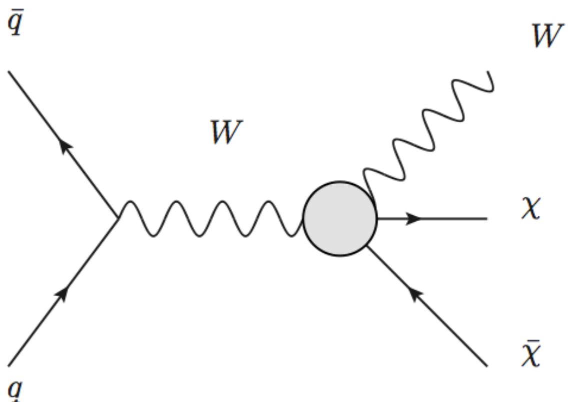
Limits

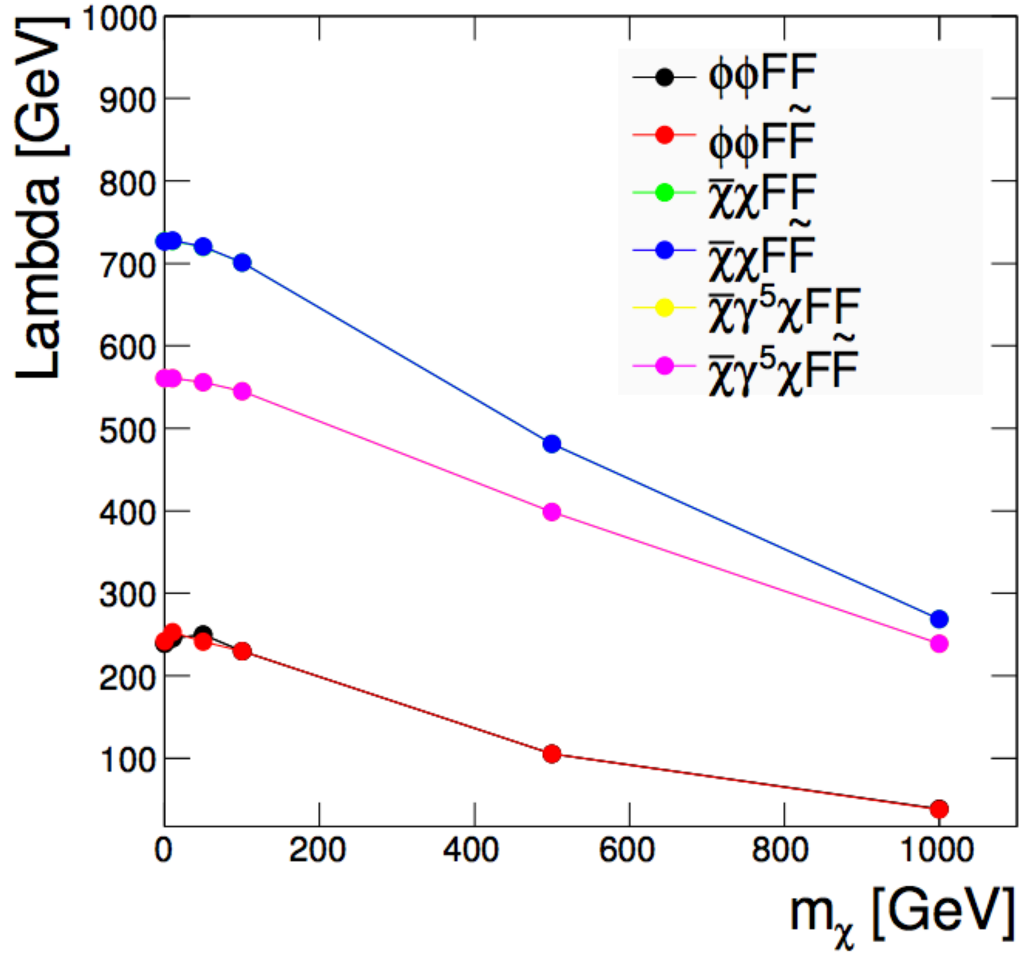
1309.4017 (PRL)



XX->WW



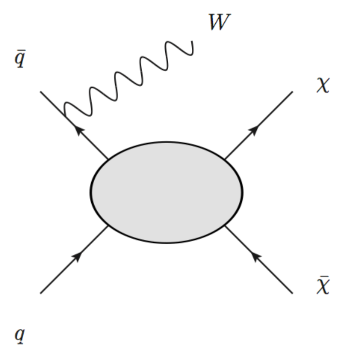


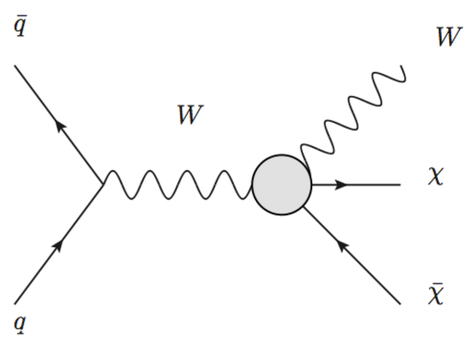


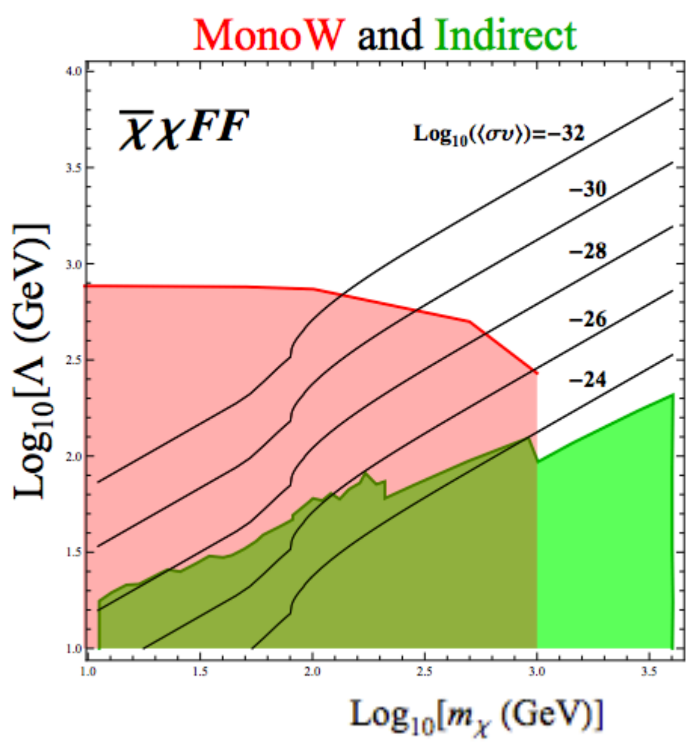
1403.6734

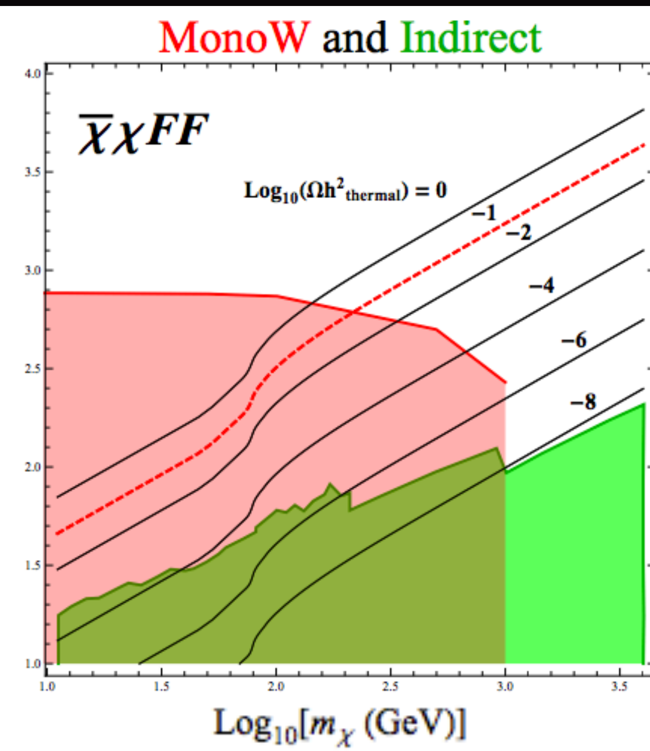
FIG. 4: Limits on Λ as a function of m_χ.

XX->WW









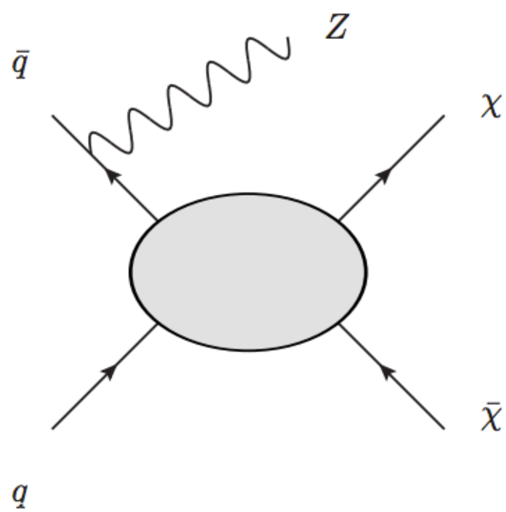
1403.6734

"Indirect" is an excluded region which is a combination of exclusions from the LAT line search, the LAT dwarf bounds and (at higher m_chi) the VERITAS Segue bounds. It is assumed that this DM makes up 100% of cosmological DM, no matter what its annihilation cross section is.

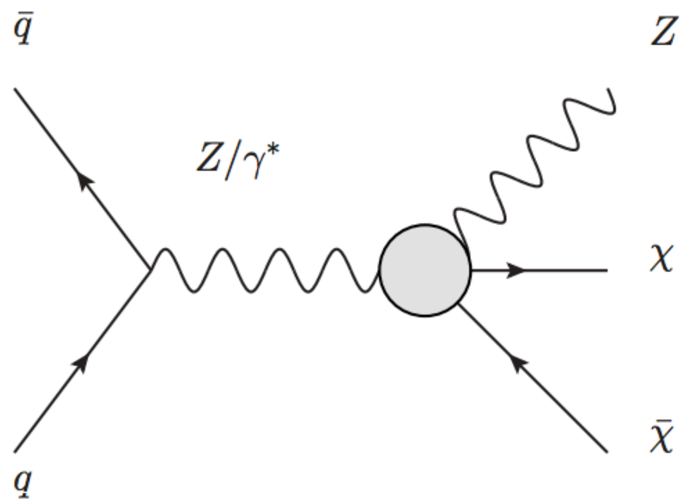
Mono-Z

1404.0051

EFTs



(a)Feynman diagram showing an ISR operator.

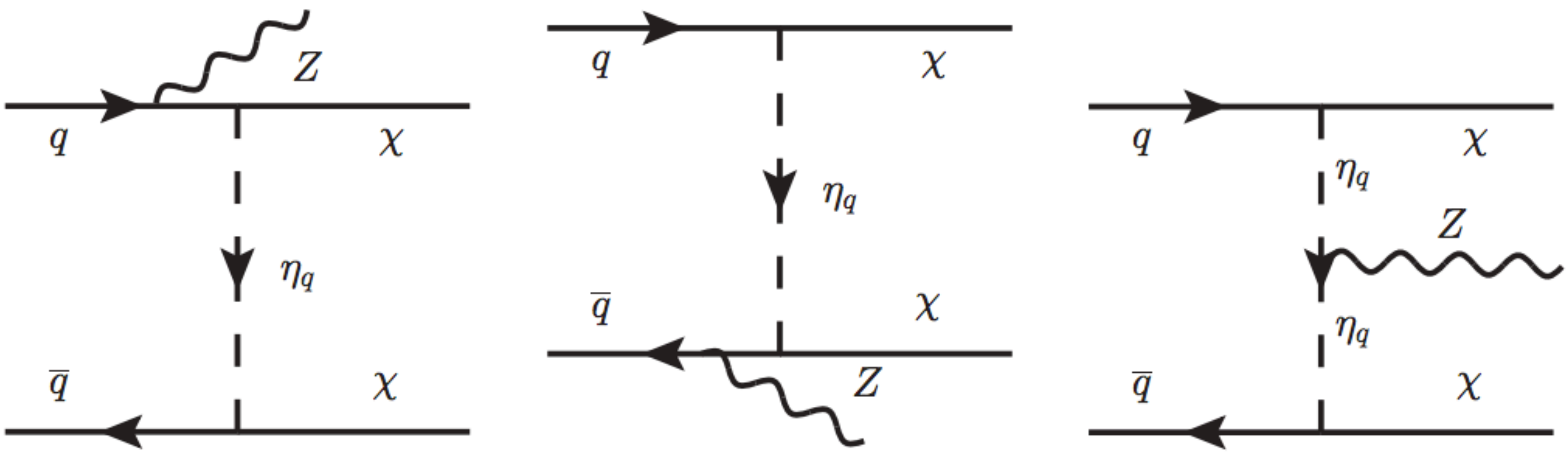


(b)Feynman diagram showing a $ZZ\chi\chi$ operator.

1404.0051

Andy Nelson, UCI

Simplified models



1404.0051

Selection

Process	E_T^{miss} threshold [GeV] 150	250	350	450
ZZ	41 ± 15	6.4 ± 2.4	1.3 ± 0.5	0.3 ± 0.1
WZ	8.0 ± 3.1	0.8 ± 0.4	0.2 ± 0.1	0.1 ± 0.1
WW, $t\bar{t}$, $Z \to \tau^+\tau^-$	1.9 ± 1.4	$0^{+0.7}_{-0.0}$	$0^{+0.7}_{-0.0}$	$0^{+0.7}_{-0.0}$
Z+jets	0.1 ± 0.1	–	–	–
W+jets	0.5 ± 0.3	–	–	–
Total	52 ± 18	7.2 ± 2.8	1.4 ± 0.9	$0.4^{+0.7}_{-0.4}$
Data	45	3	0	0

Selection

Two OS SF leps

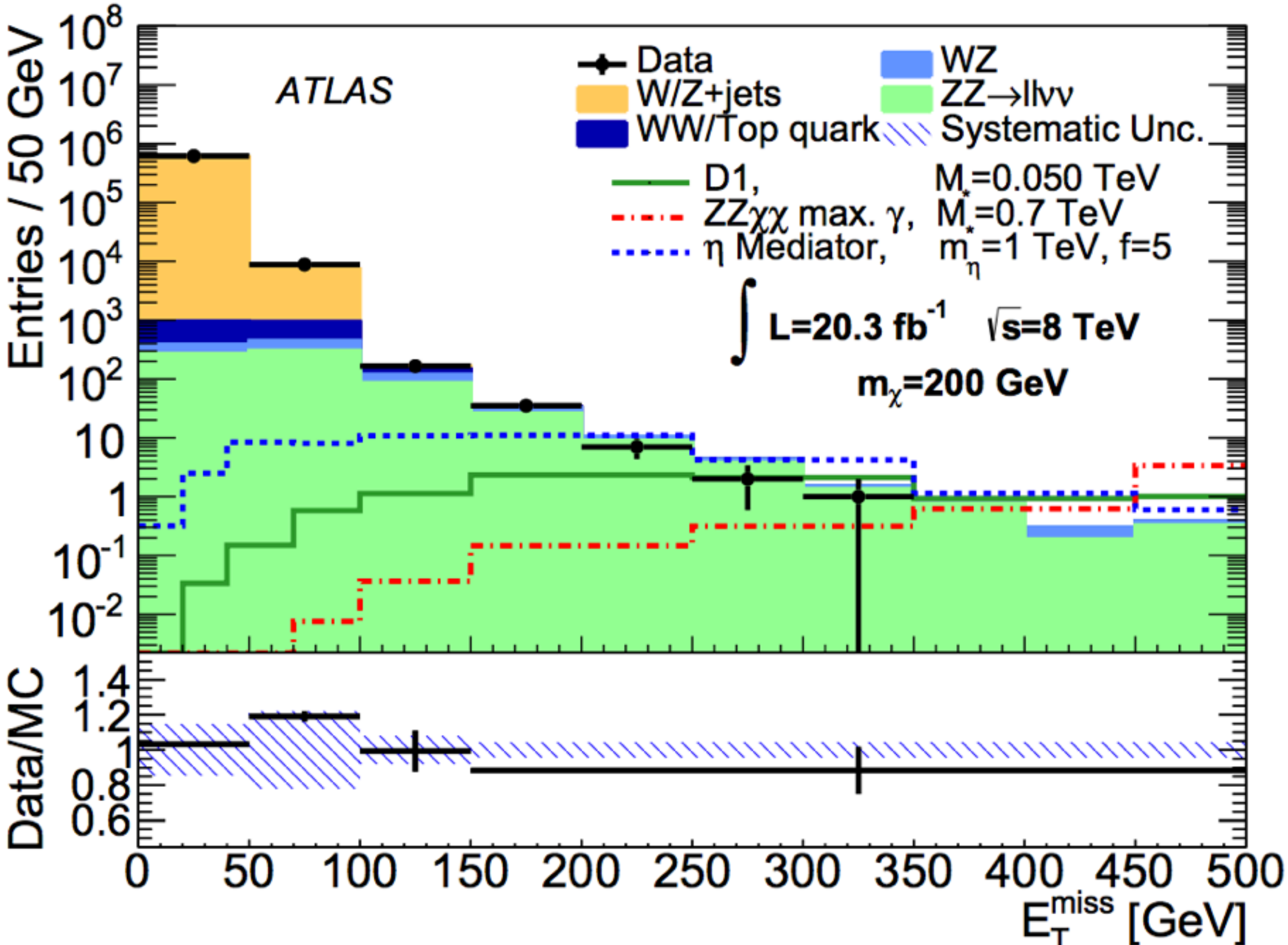
m_{ll} in [76,106]

veto jets, 3rd lep

MET angle cuts

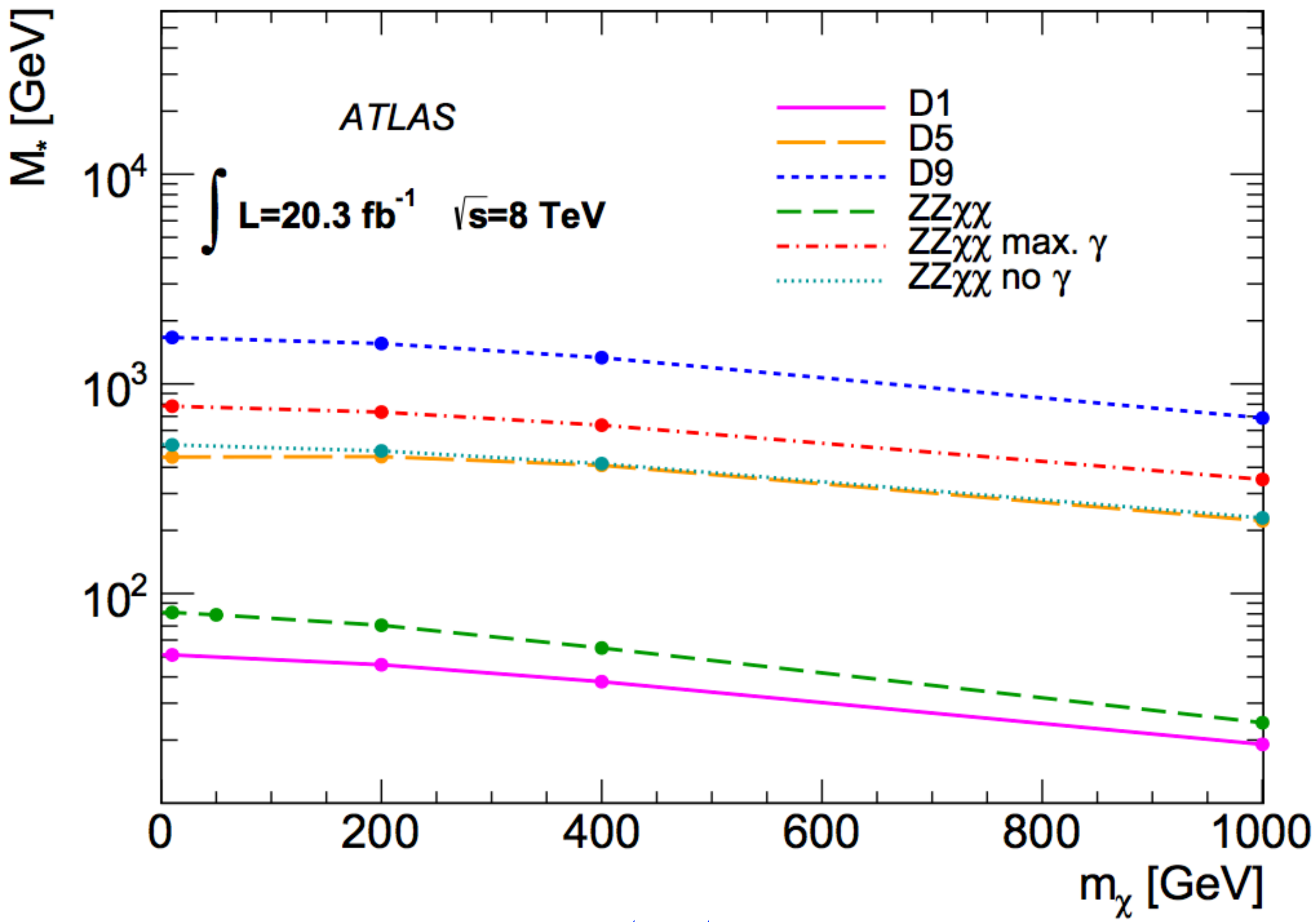
$|p_T^{\ell\ell} - E_T^{miss}|/p_T^{\ell\ell} < 0.5$

Data



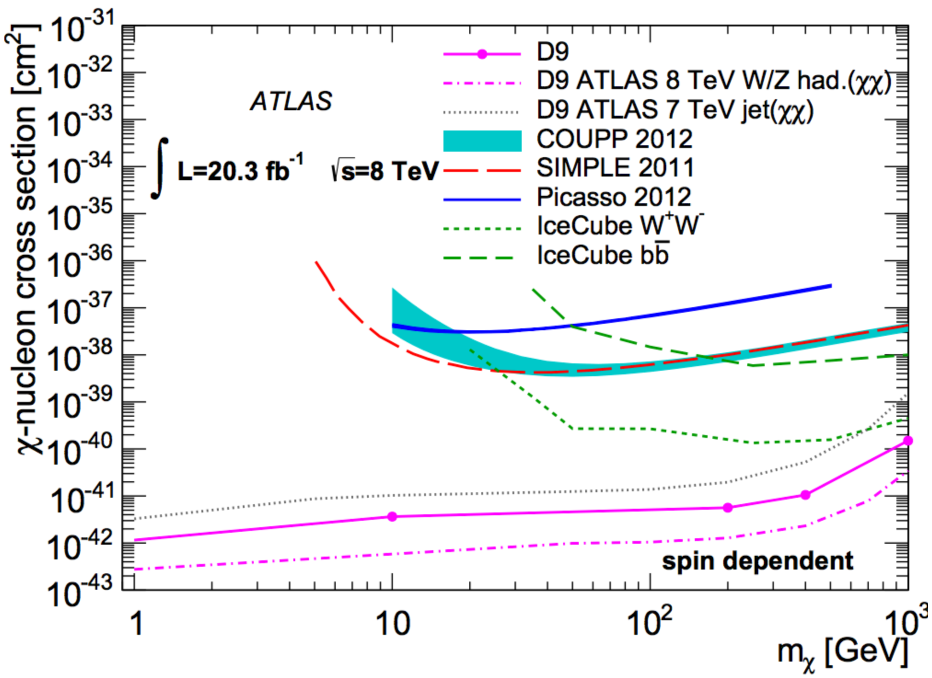
1404.0051

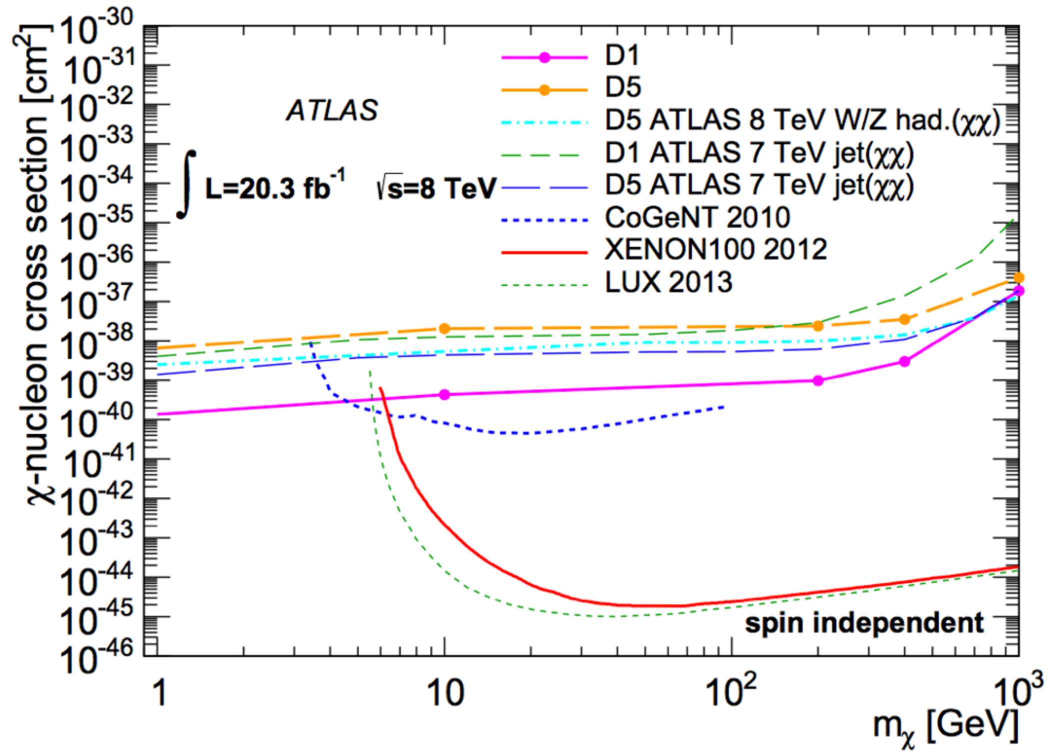
Limits....



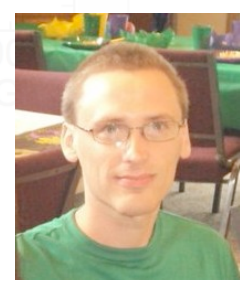
1404.0051

Limits....

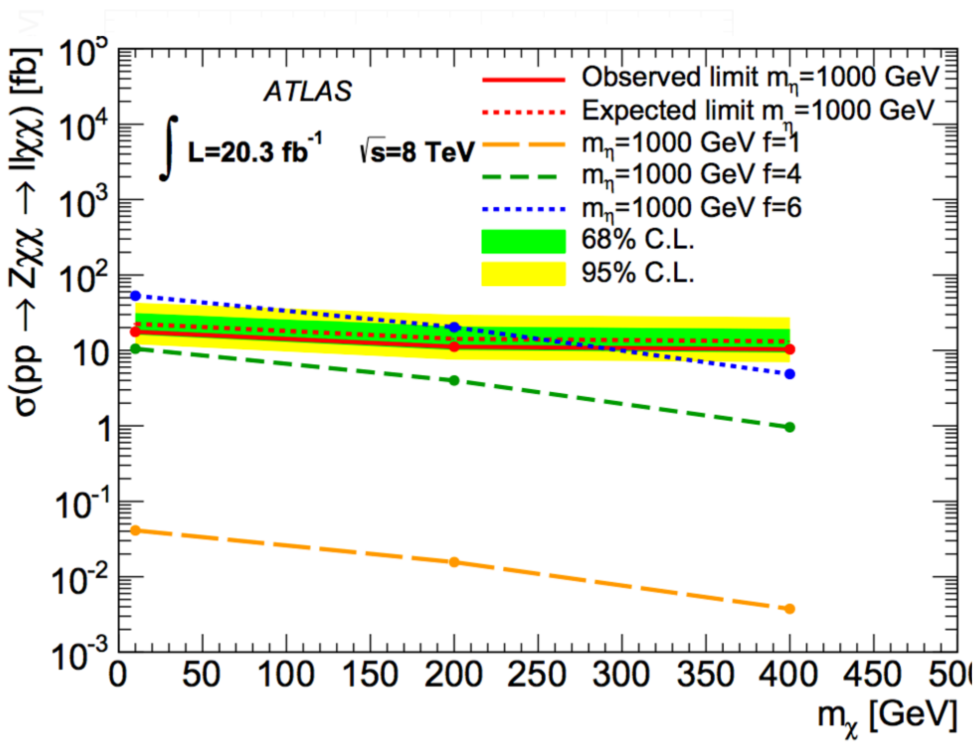


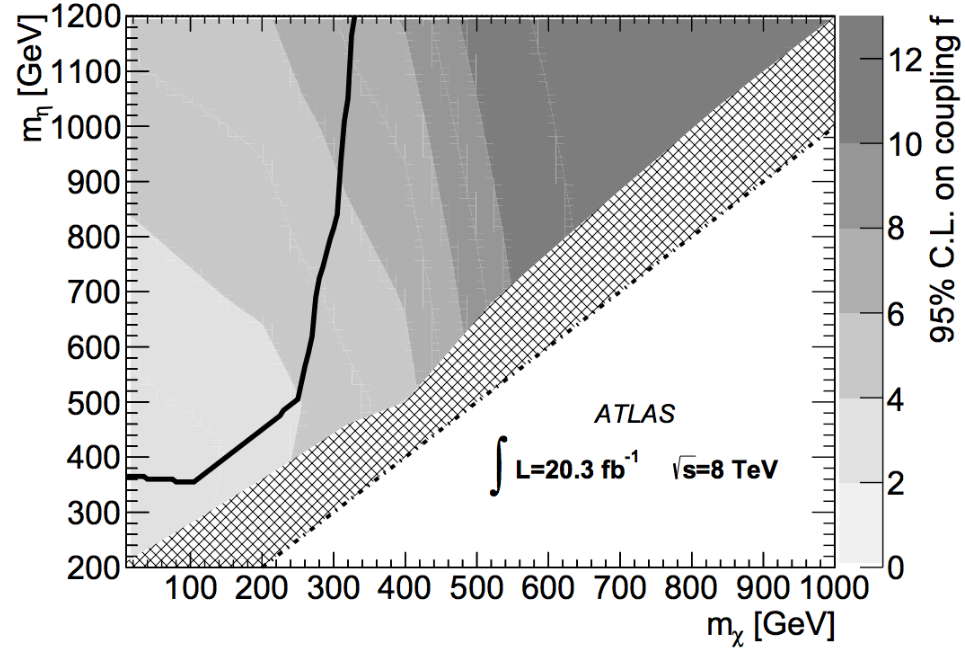


1404.0051



Limits....





1404.0051

Mono-Higgs

1312.2592

Models

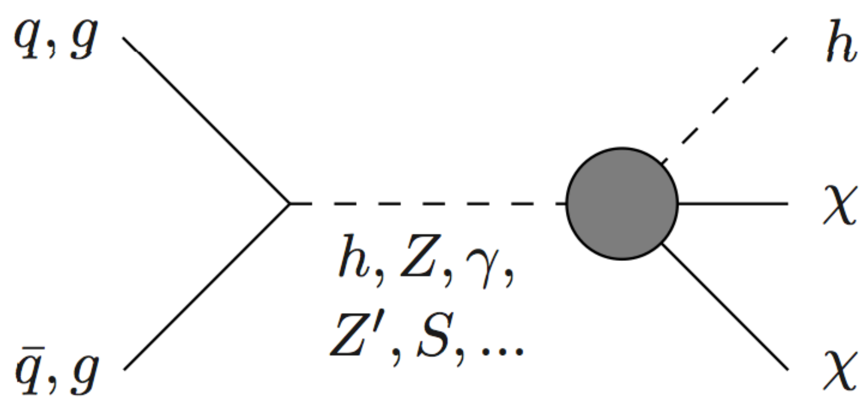


FIG. 1: Schematic diagram for mono-Higgs production in pp collisions mediated by electroweak bosons (h, Z, γ) or new mediator particles such as a Z' or scalar singlet S. The gray circle denotes an effective interaction between DM, the Higgs boson, and other states.

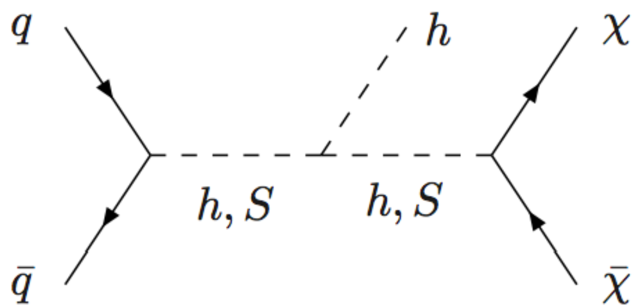
Models: EFT

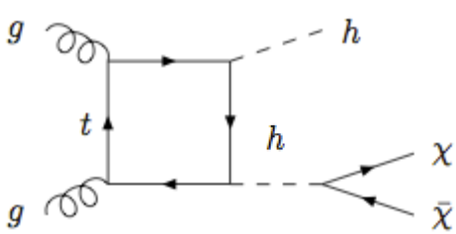
$\lambda|H|^2|\chi|^2$ Scalar wimp

$\frac{1}{\Lambda}|H|^2\bar{\chi}\chi\,, \quad \frac{1}{\Lambda}|H|^2\bar{\chi}i\gamma_5\chi$ Fermion wimp

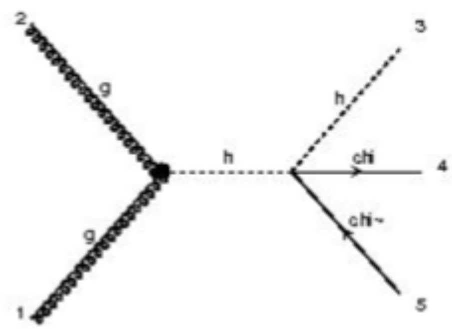
Vertices

di-Higgs

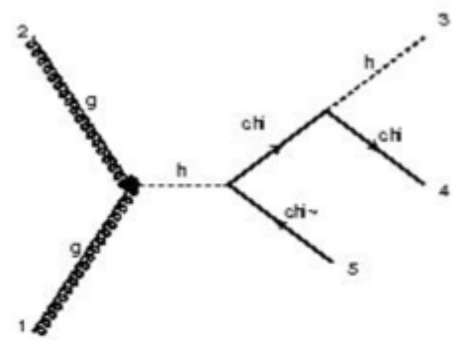




4-point vertex



Off-shell s-channel Higgs



(1) h->XX limited by invisible Higgs for mx<mh/2

(2)For large coupling, h->XX grows, suppresses SM H decays!

Other EFTs

Allow ZhXX-like vertices

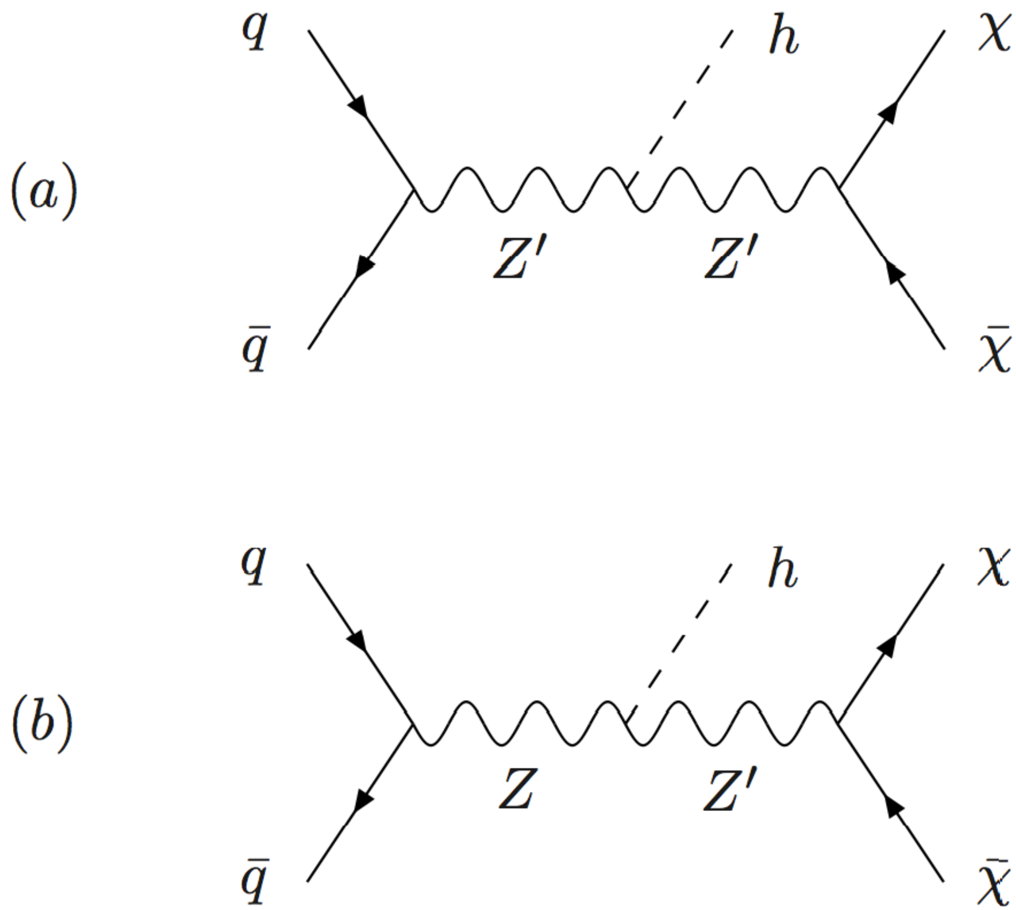
$$\frac{1}{\Lambda^2}\chi^\dagger i\overleftrightarrow{\partial^\mu}\chi H^\dagger iD_\mu H$$

Scalar wimp

$$\frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu\chi B_{\mu\nu}H^\dagger D^\nu H\,.$$

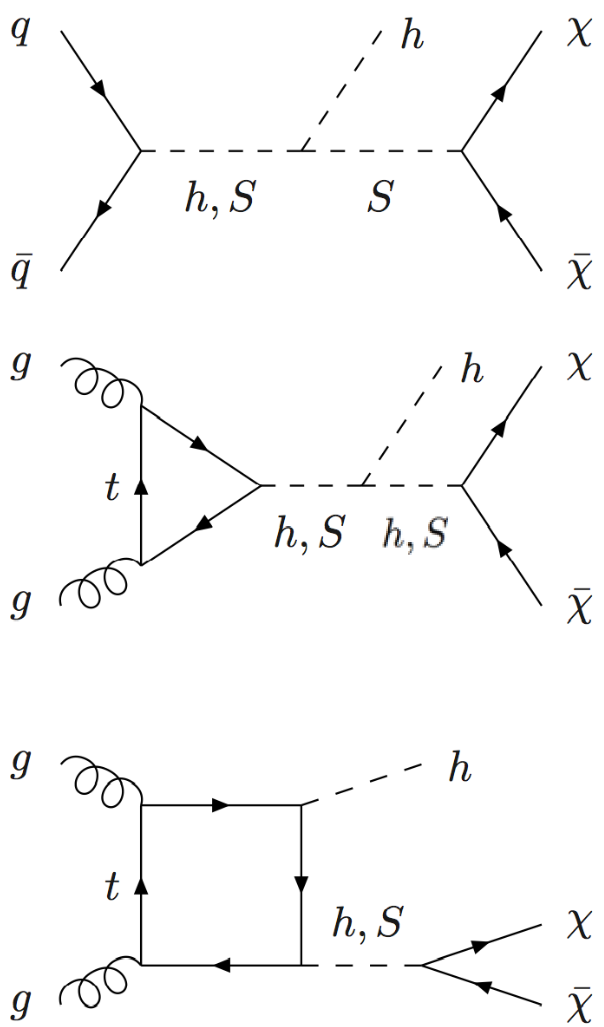
Fermion wimp

Simplified models: vector



with and without Z-Z′ mixing

Simplified models: scalar

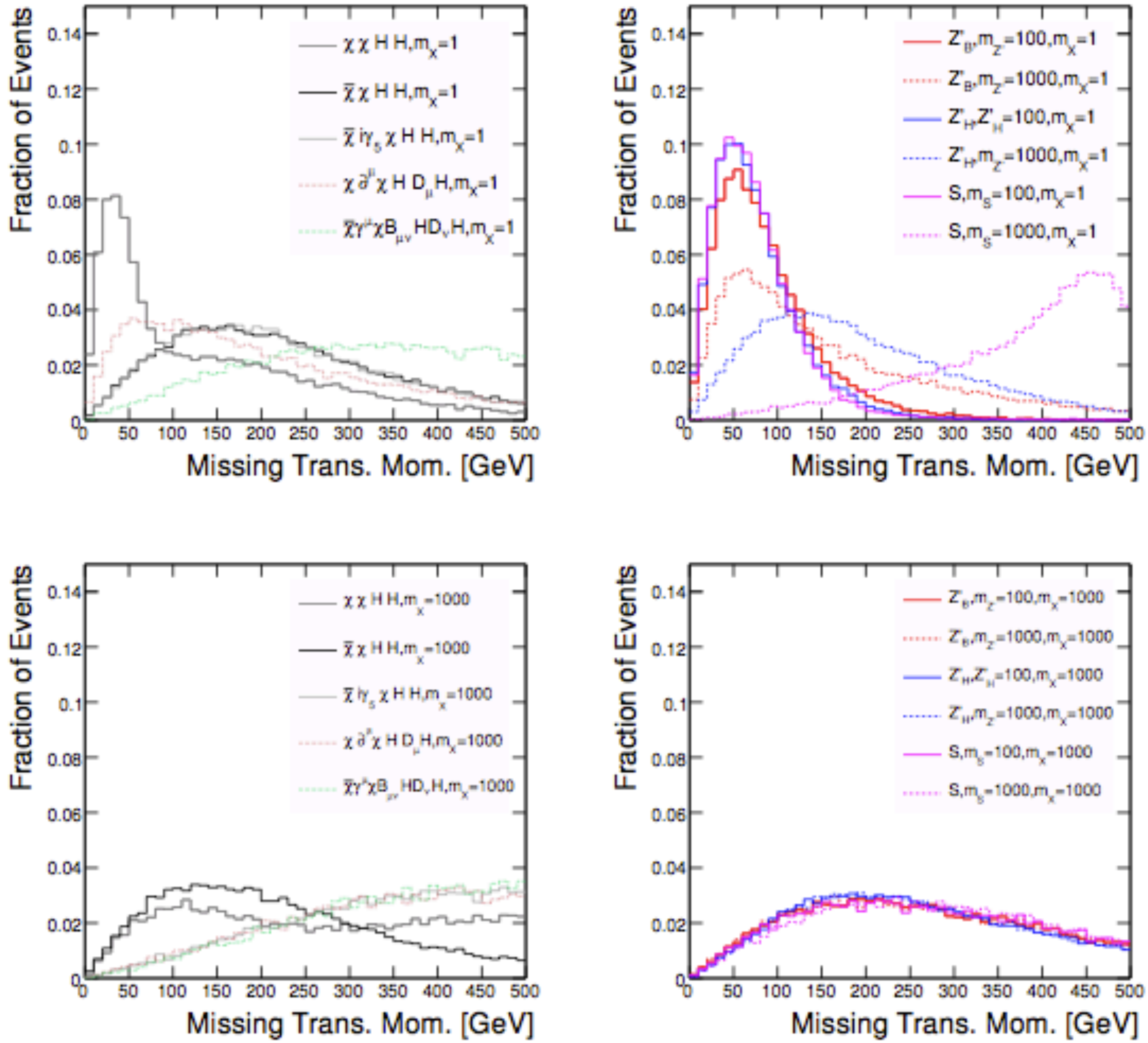


Box implemented as effective vertex in madgraph

MET

mx=1 GeV

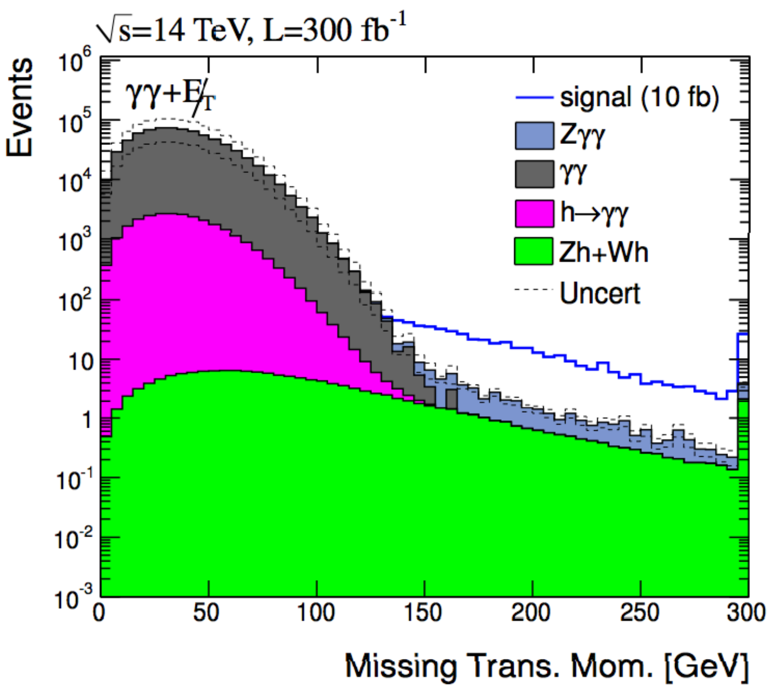
mx=1 TeV



EFTs

Simp. models

Gamma-gamma



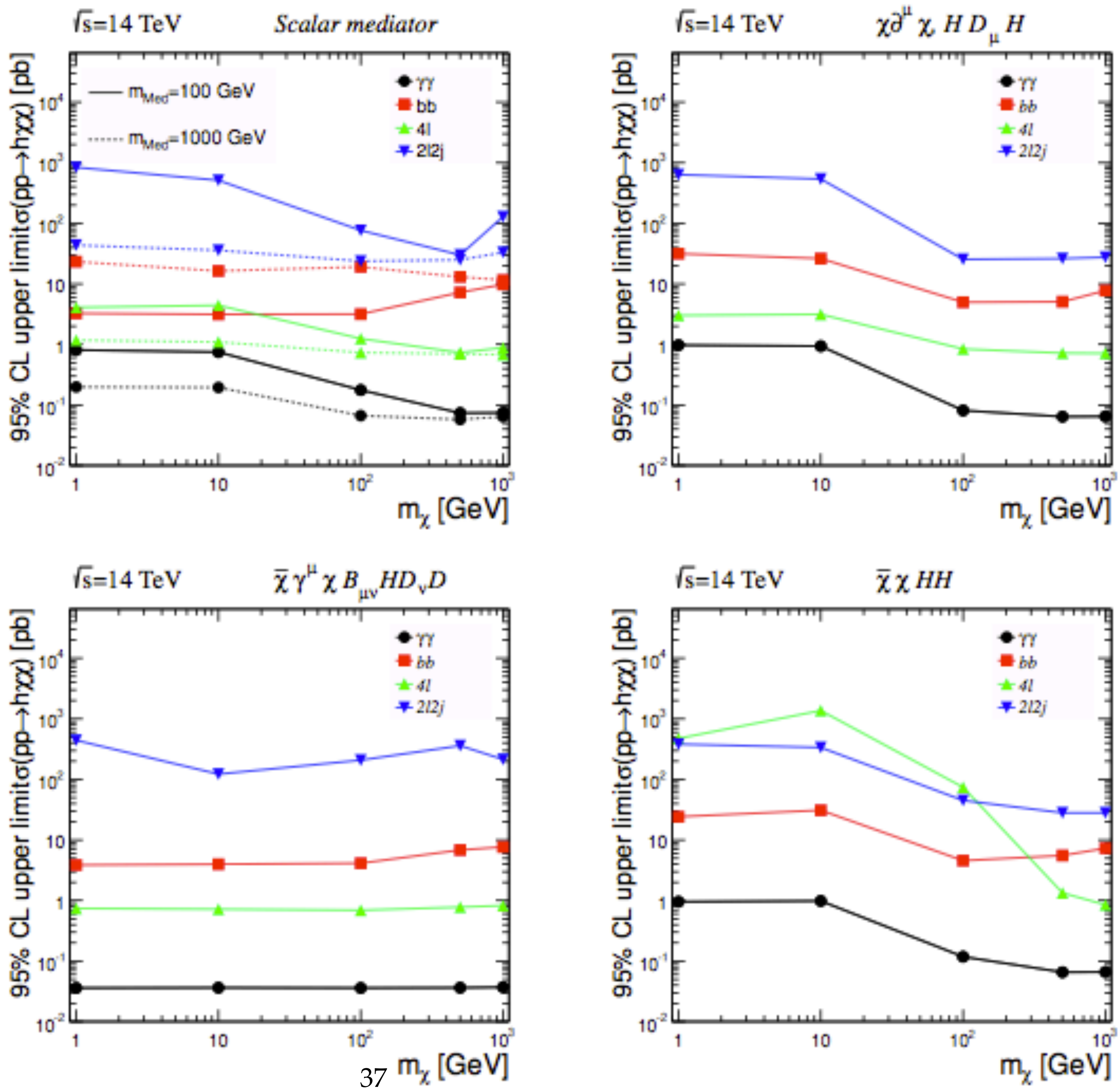
Selection

- two photons
- $m_{\gamma\gamma}$ in [110-130]
- MET > 100, 250 (8,14 TeV)

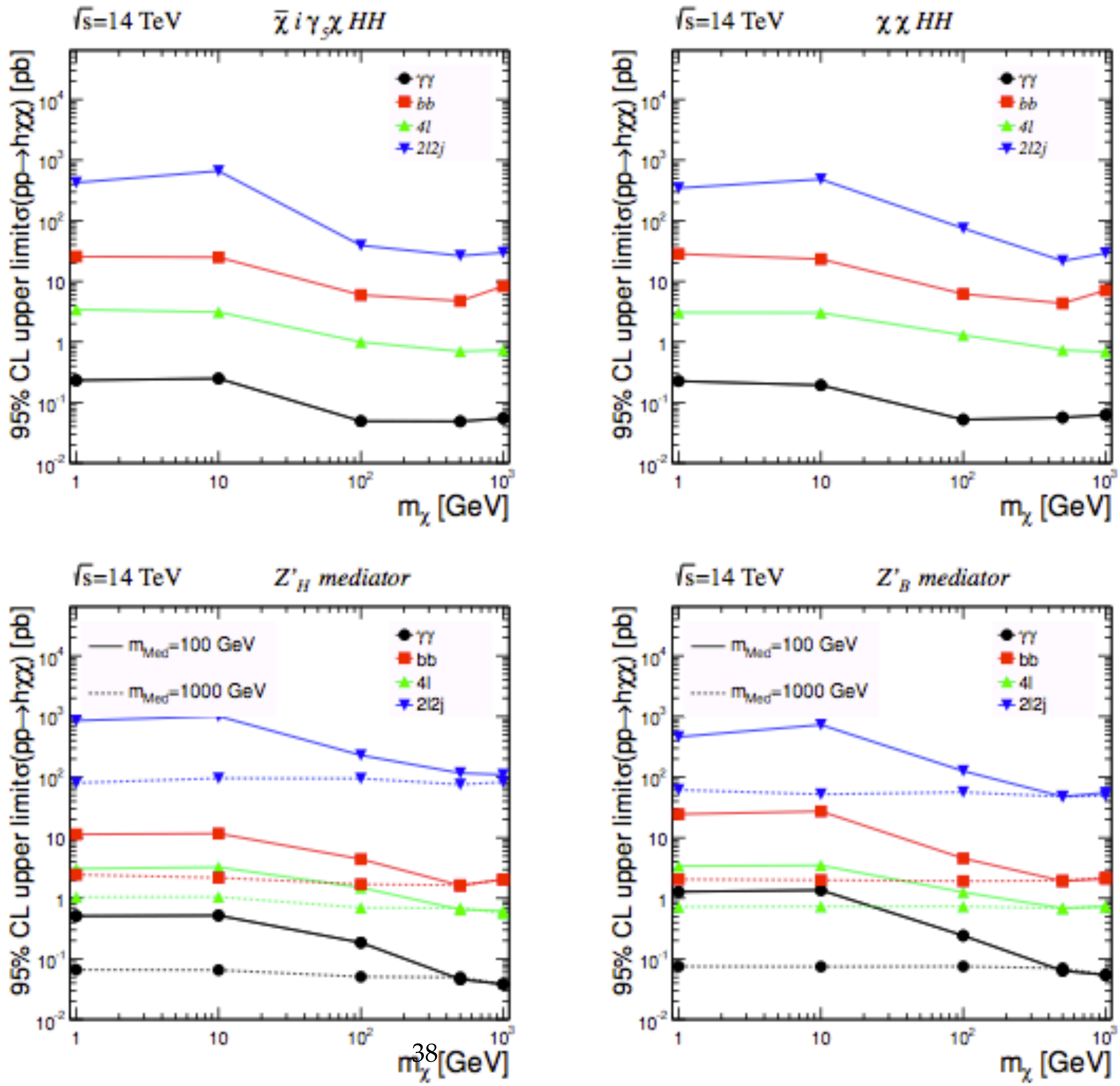
Backgrounds

- h→γγ + fake MET
- γγ + fake MET
- Zγγ, Z→νν
- Zh, Z→νν + Wh, W→lν

Assuming $h \rightarrow SM$ rates are unchanged



[redacted]



Assuming $h \rightarrow SM$ rates are unchanged

Parameter limits

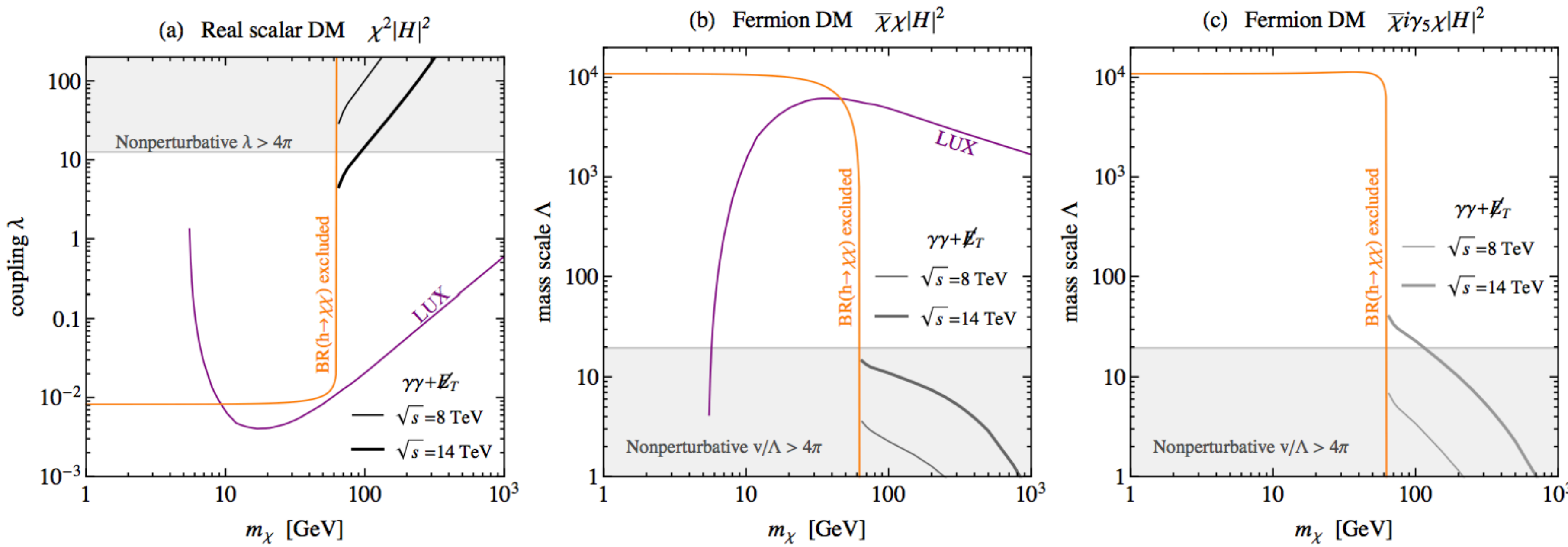


FIG. 20: Projected LHC mono-Higgs sensitivities at $\sqrt{s} = 8$ TeV (20 fb^{-1}) and 14 TeV (300 fb^{-1}), with $\gamma\gamma + \not{E}_T$ final states, on Higgs portal effective operators. All constraint contours exclude larger coupling λ or smaller mass scale Λ. Shaded region is excluded based on perturbativity arguments; orange contours denote limits from invisible h decays; purple contours are exclusion limits from LUX.

Note:

for $m_X < m_h/2$, no valid limits.

Large Lambda boosts h→XX, suppresses h→visible

Parameter limits

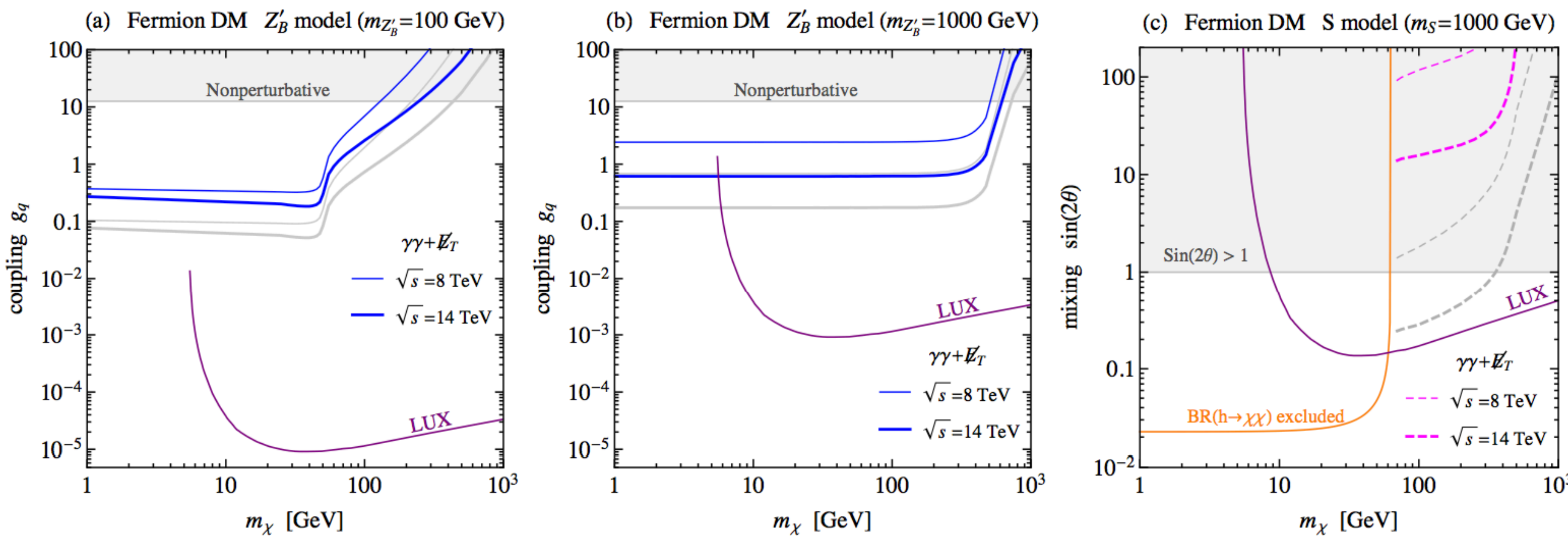
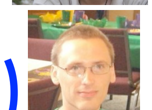
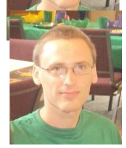


FIG. 22: Projected LHC mono-Higgs sensitivities at $\sqrt{s} = 8$ TeV (20 fb^{-1}) and 14 TeV (300 fb^{-1}), with $\gamma\gamma + \not{E}_T$ final states, on simplified models. All constraint contours exclude larger couplings or mixing angles. Shaded region is excluded based on perturbativity arguments or requiring $\sin\theta \leq 1$; orange contour denotes limit from invisible h decays; purple contours are exclusion limits from LUX.

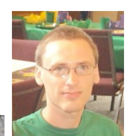
DM References + Plans

ATLAS


7 TeV γ+MET (1209.4625)

W→jj +MET (1309.4017)

Invisible Higgs (1402.3244)

Z+MET (1404.0051)

W→lν +MET (soon)

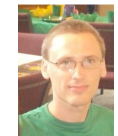

VBF Invisible Higgs (forthcoming)

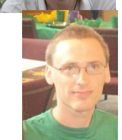
8 TeV γ+MET (forthcoming)

dijets (forthcoming)

Higgs+MET (forthcoming)

Pheno

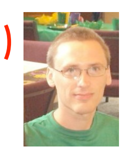

monoZ (1212.3352)

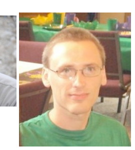
DM combo (1302.3619)

Fermi/LHC (1307.5064)

DM future (1307.5327)

H+MET (1312.2592)

Indirect WW (1403.6734)


Compressed spectra (forthcoming)
mono-Z′ (forthcoming)



Outline

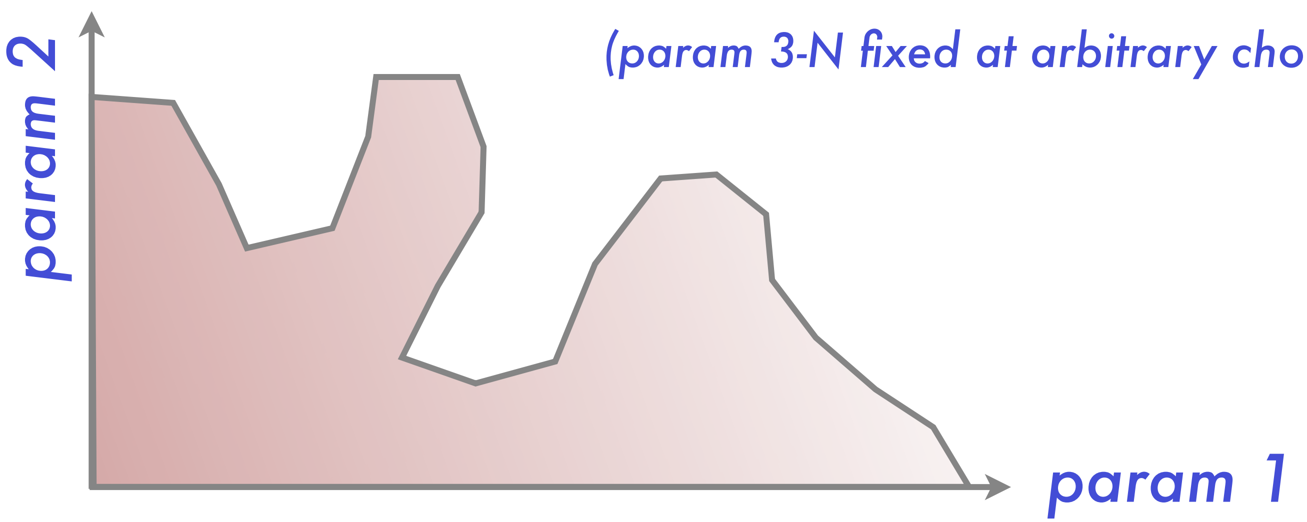
I. Dark Matter
II. Topological Models
III. Deep networks

Searching for new physics

Traditional approach

Bet on a specific theory

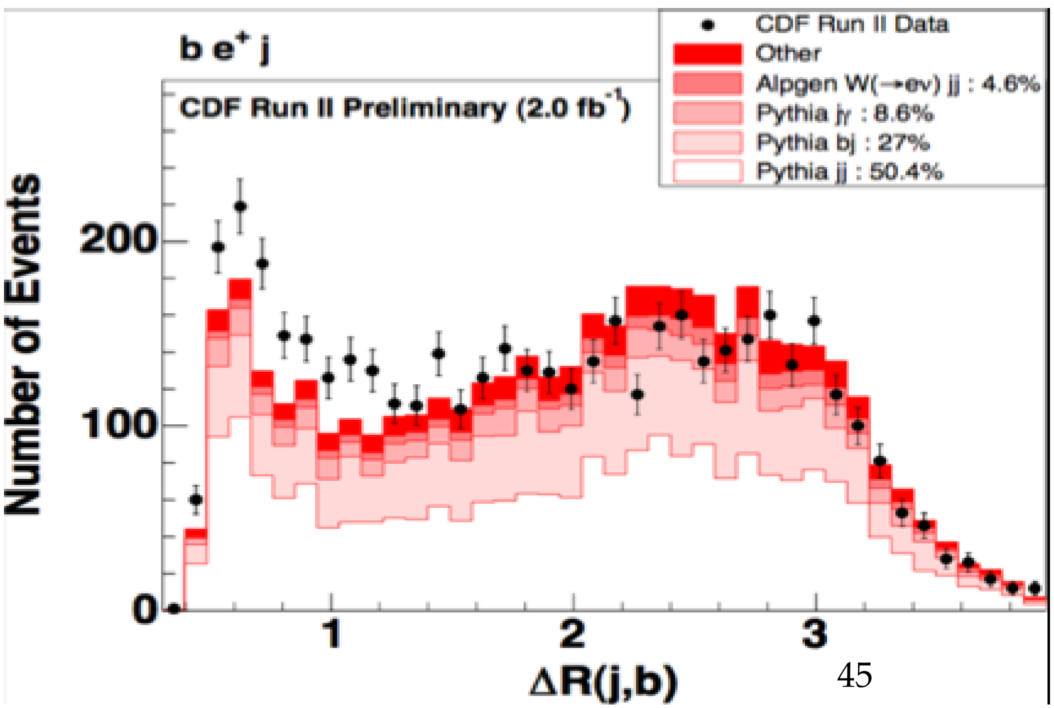
Optimize analysis to squeeze out maximal sensitivity to new physics.



Model independent search

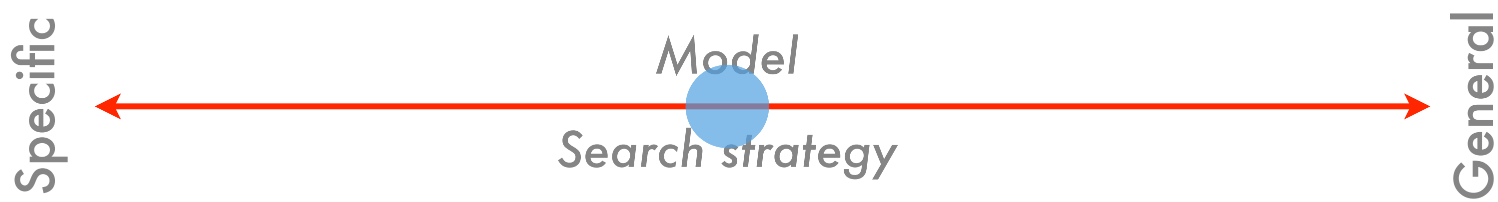
Discard the model

compare data to standard model



"Never listen to theorists.
Just go look for it."
–A. Pierce, 2010

Compromise



Admit the need for a model

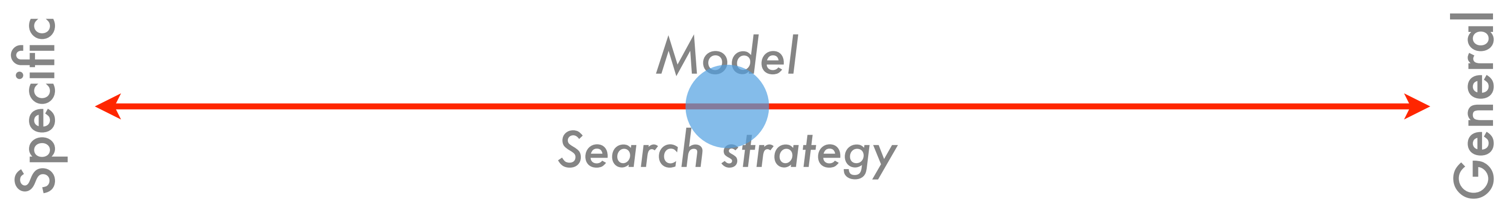
New signal requires a coherent physical explanation, even trivial or effective

Generalize your model

Construct simple models that describe classes of new physics which can be discovered at the LHC.

What are we good at discovering?

Compromise



<u>Admit the need for a model</u>

New signal requires a coherent physical explanation, even trivial or effective

<u>Generalize your model</u>

Construct simple models that describe classes of new physics which can be discovered at the LHC.

What are we good at discovering? Resonances!

Is this being done?

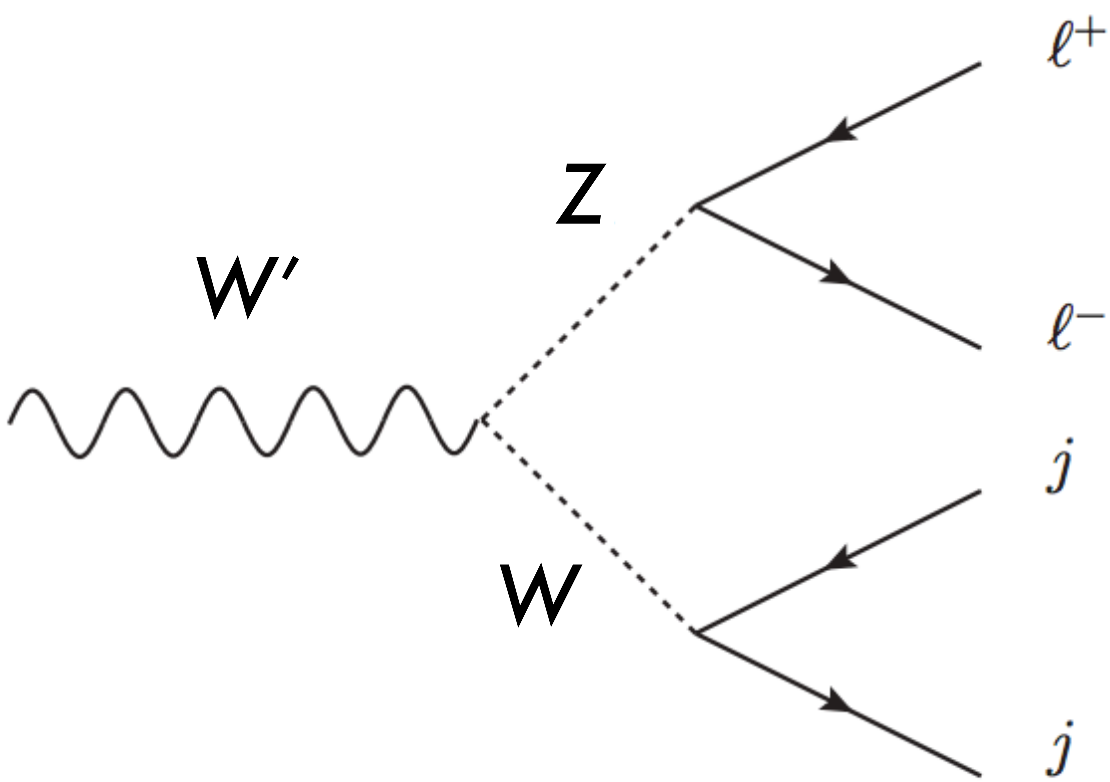
ℓ^+

ℓ^-

j

j

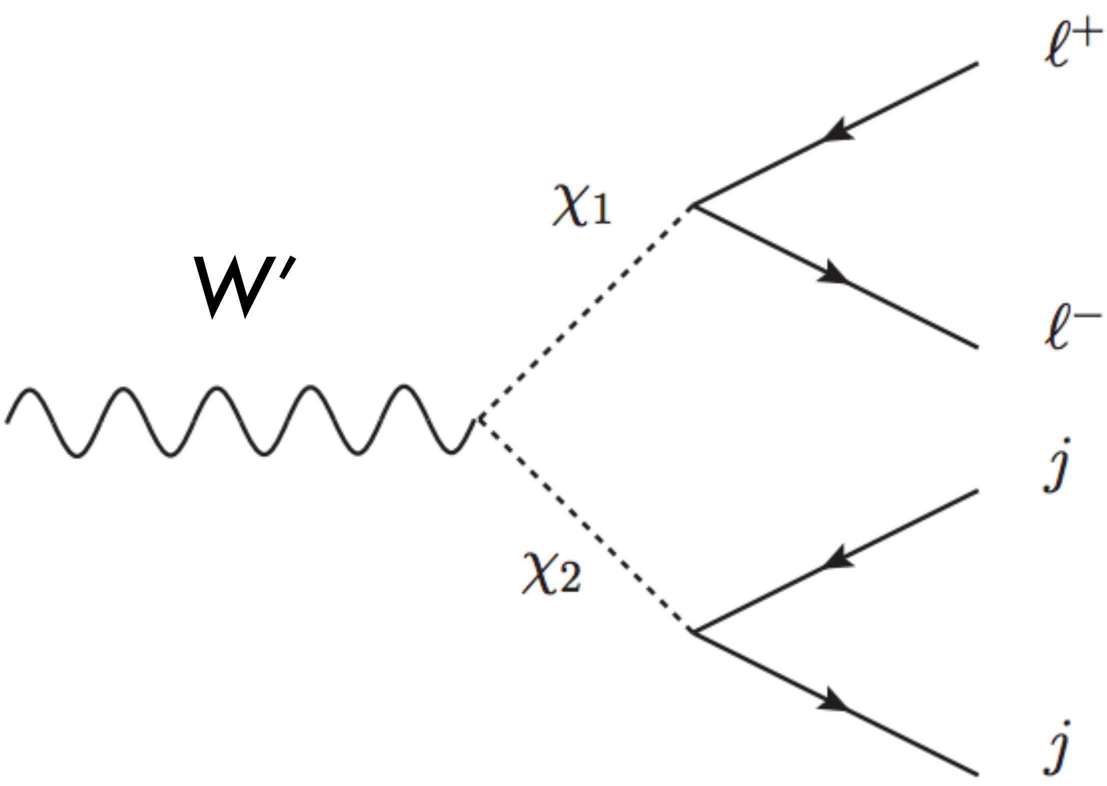
Is this being done?



$m_{ll} = m_Z$

$m_{jj} = m_W$

What about this?



m_{ll} != m_Z

m_{jj} != m_W

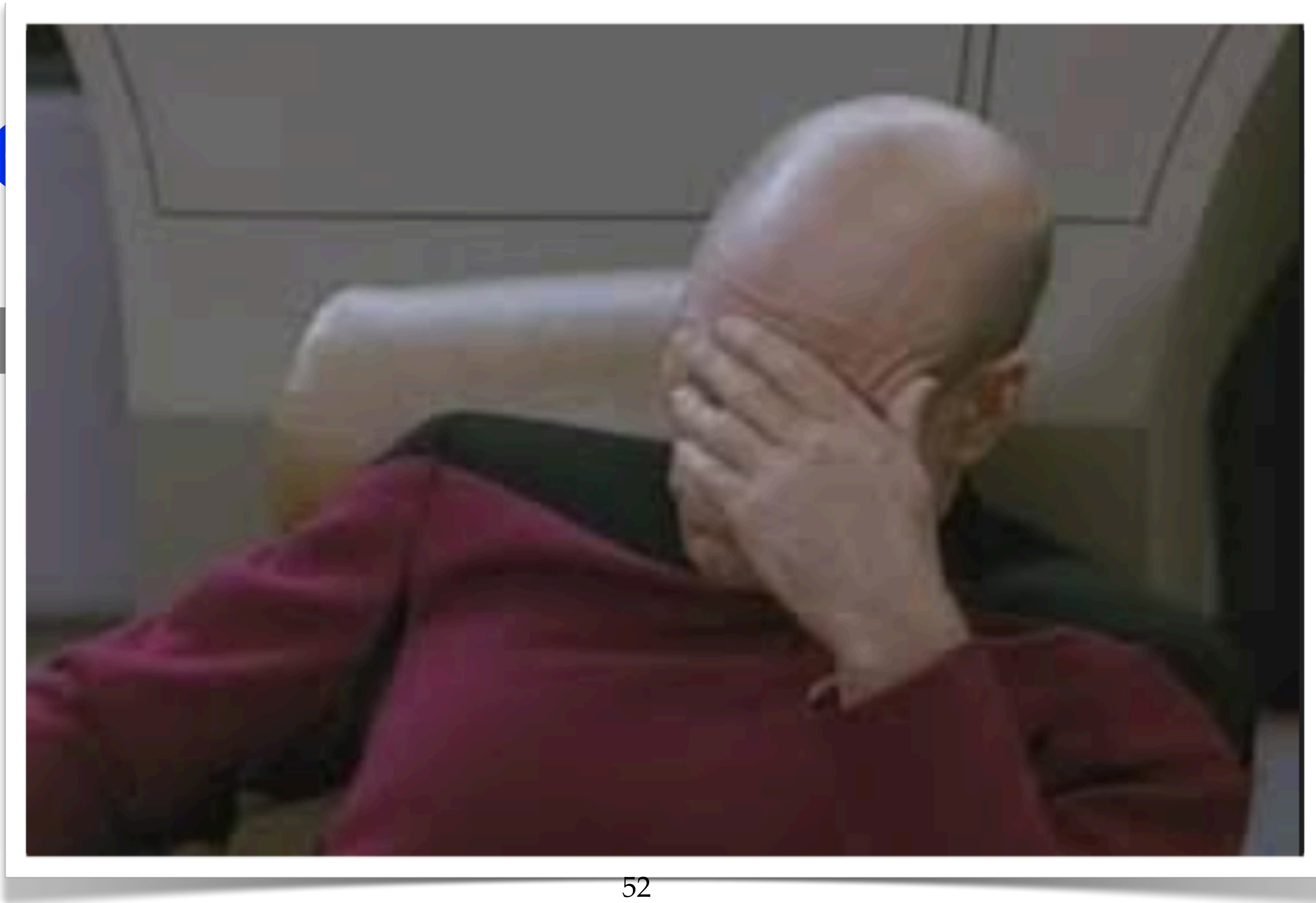
Missed resonances?

Easy-to-find resonances
may exist in our data and
nobody has looked!

Missed resonances?

E… s

ma… nd



Topological models

UCI Physics 247
Final project
arXiv: 1401.1462

FERMILAB-PUB-13-529-T

Systematically Searching for New Resonances at the Energy Frontier using Topological Models

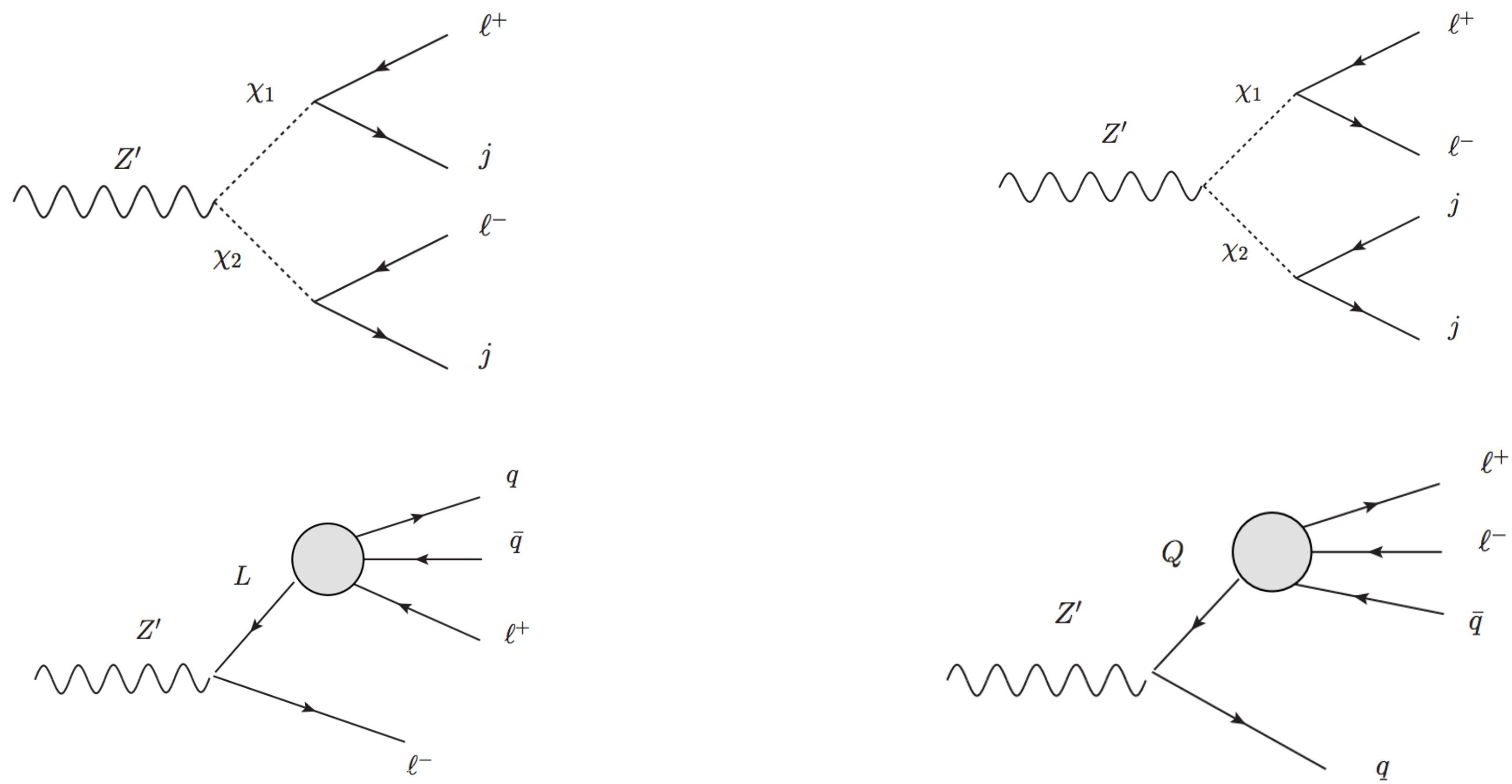
Mohammad Abdullah,[1] Eric Albin,[1] Anthony DiFranzo,[1] Meghan Frate,[1] Craig Pitcher,[1] Chase Shimmin,[1] Suneet Upadhyay,[1] James Walker,[1] Pierce Weatherly,[1] Patrick J. Fox,[2] and Daniel Whiteson[1]

[1]*Department of Physics and Astronomy, University of California, Irvine, CA 92697*
[2]*Fermi National Accelerator Laboratory, Batavia, IL 60615*

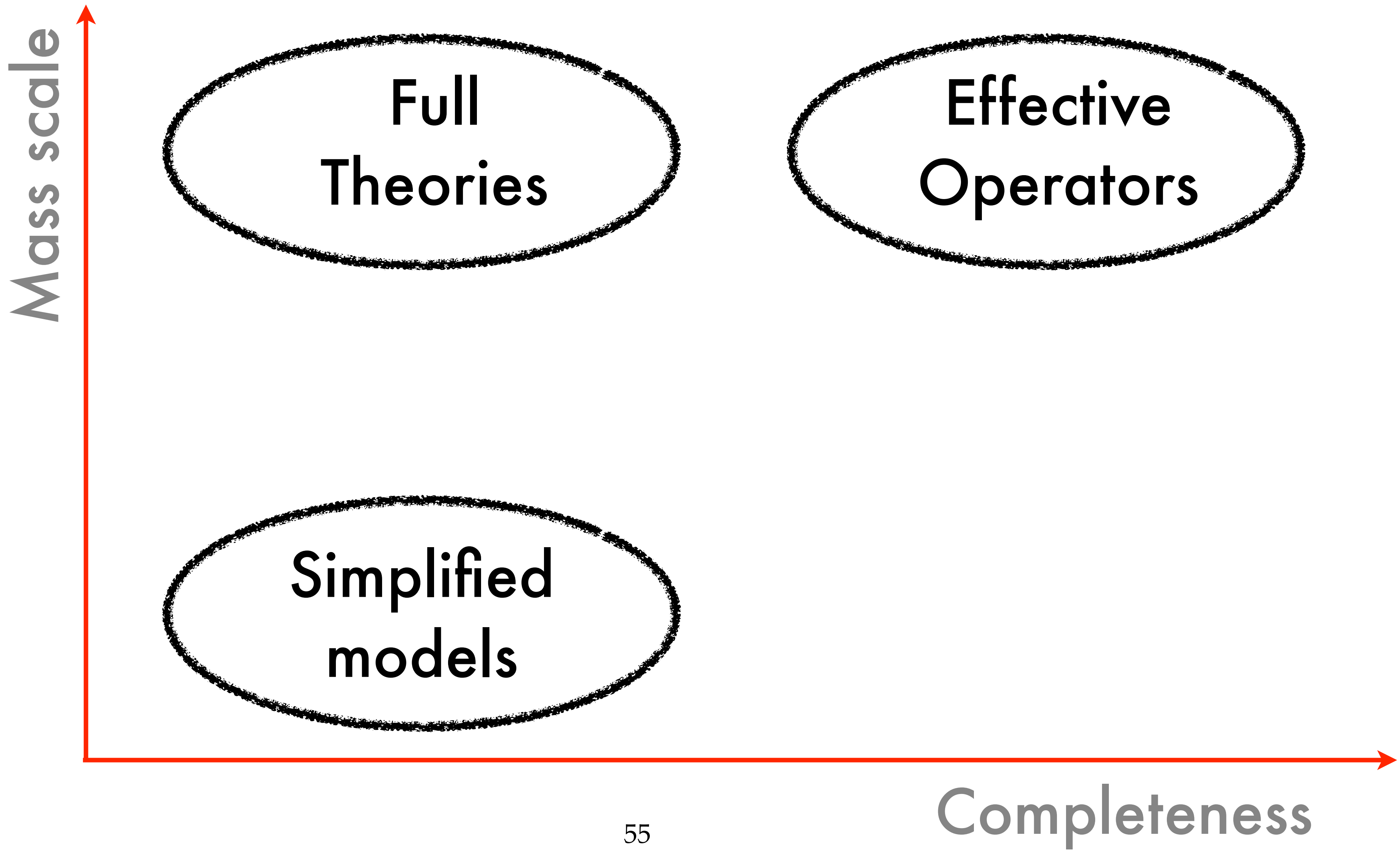
We propose a new strategy to systematically search for new physics processes in particle collisions at the energy frontier. An examination of all possible topologies which give identifiable resonant features in a specific final state leads to a tractable number of 'topological models' per final state and gives specific guidance for their discovery. Using one specific final state, $\ell\ell jj$, as an example, we find that the number of possibilities is reasonable and reveals simple, but as-yet-unexplored, topologies which contain significant discovery potential. We propose analysis techniques and estimate the sensitivity for pp collisions with $\sqrt{s} = 14$ TeV and $\mathcal{L} = 300$ fb^{-1}.

Topological models

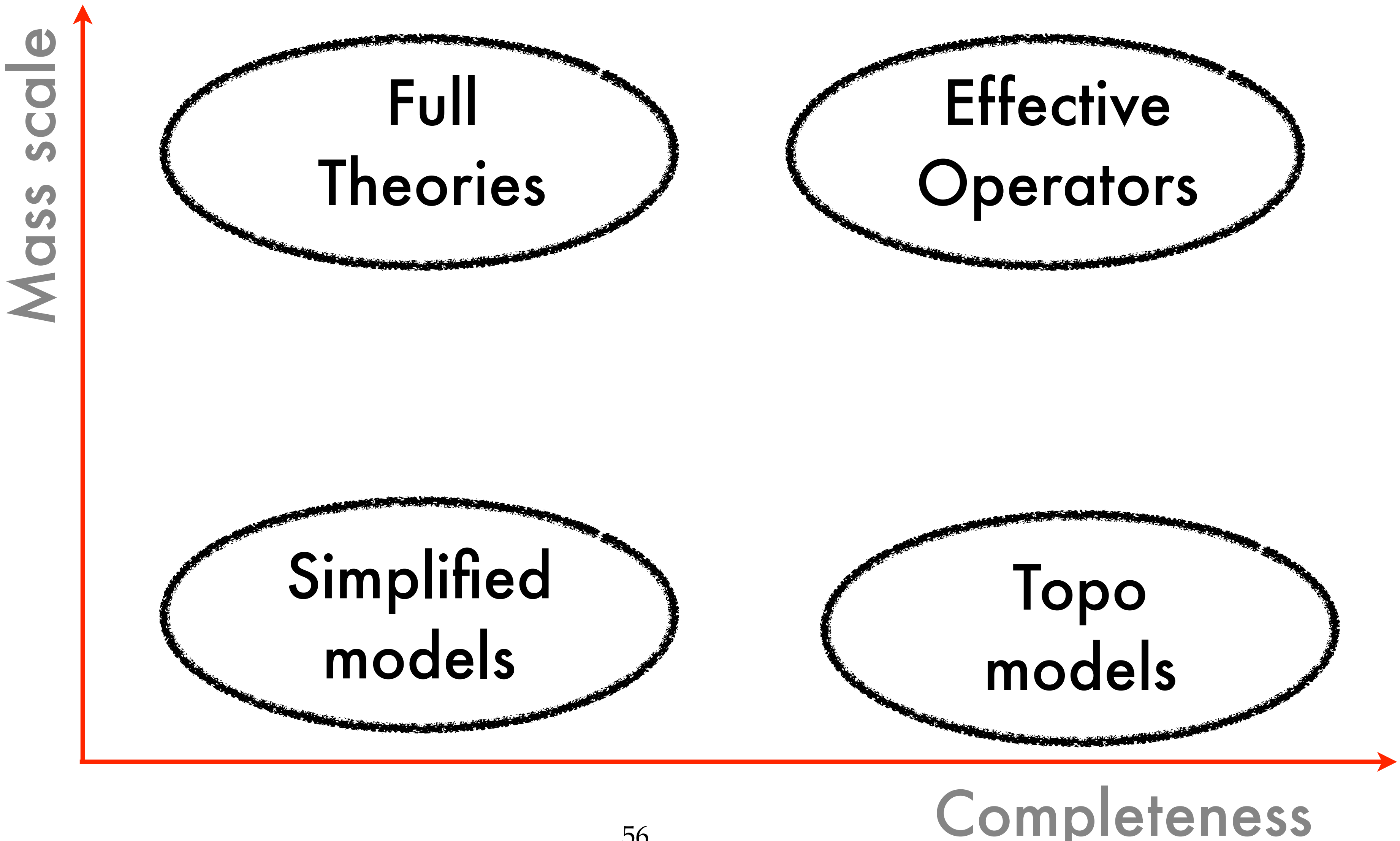
For a given final state (eg lljj) construct all models with resonances. Then look for them!



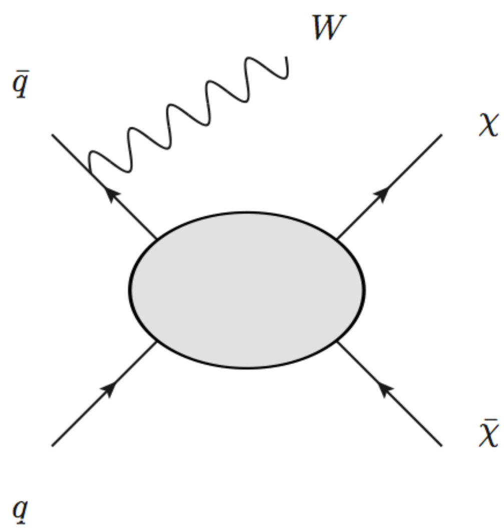
Connections to EFT, Simp. Models

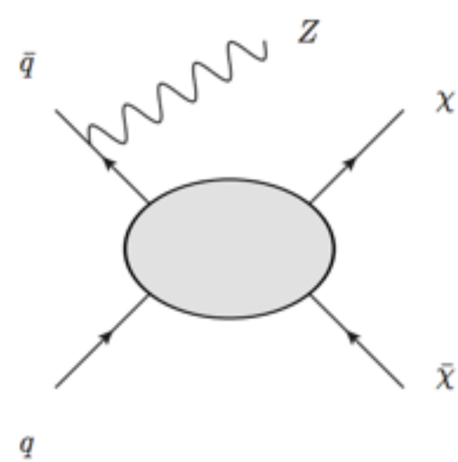


Connections to EFT, Simp. Models



Mono-Z'

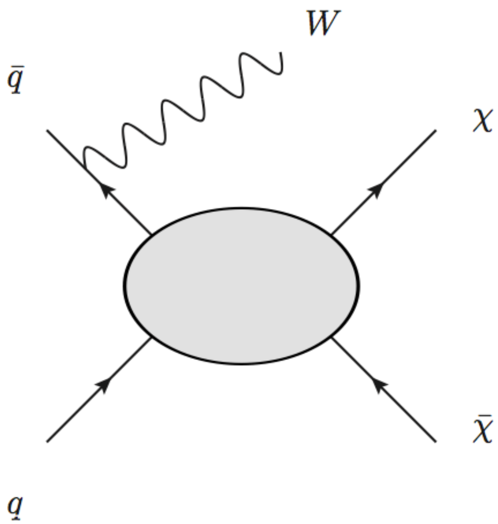


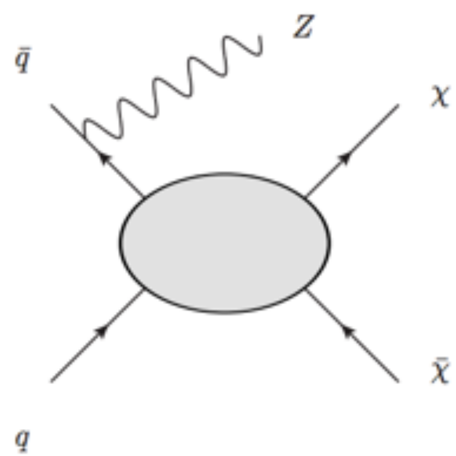


$m_{jj} = m_W$ or m_Z

$m_{ll} = m_Z$

Mono-Z'





$m_{jj} = m_W$ or m_Z

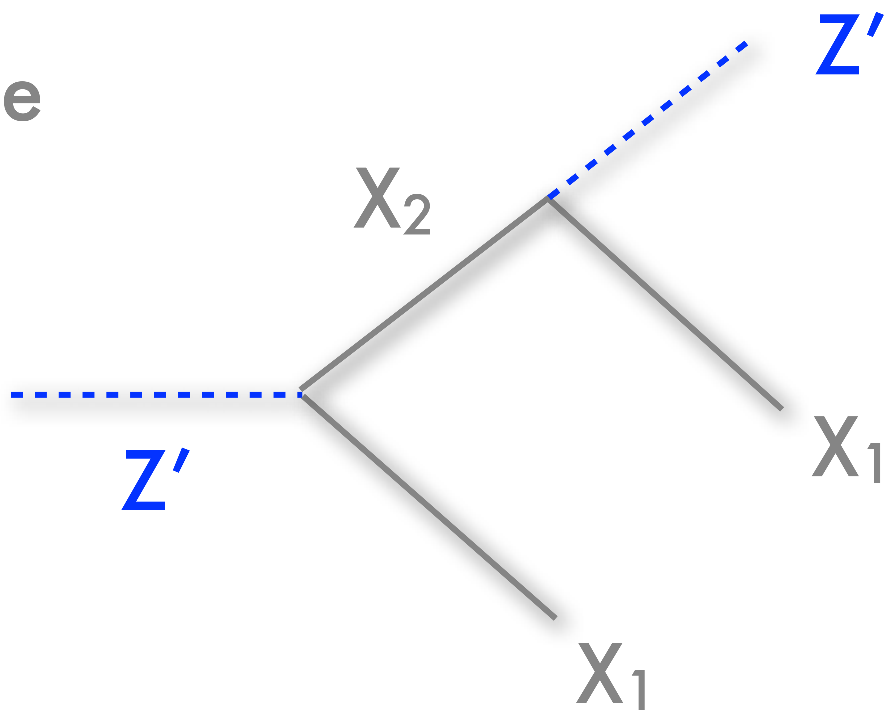
$m_{ll} = m_Z$

What about other values?

Mono-....

<u>Signature</u>

Heavy resonance
+ MET

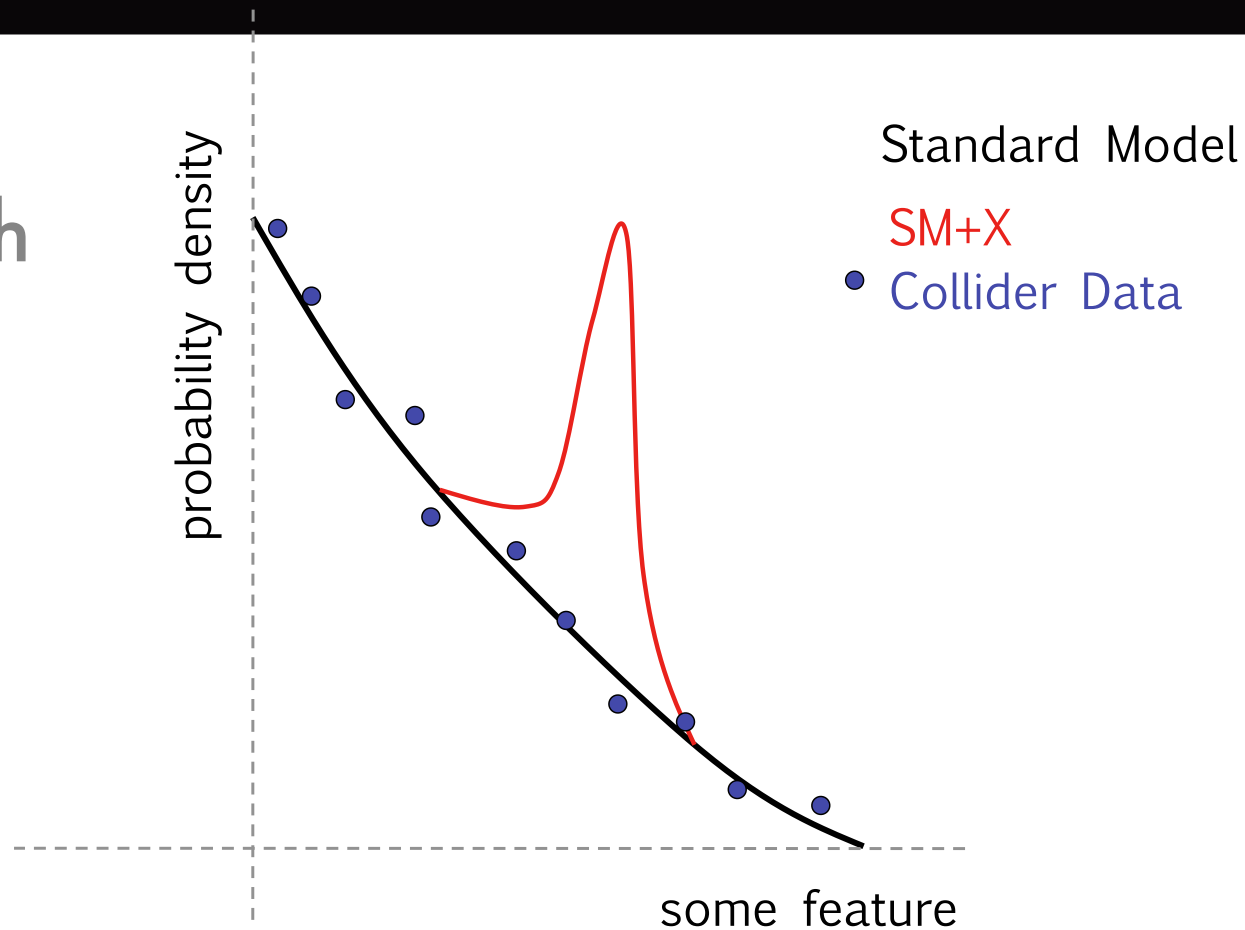


Outline

I. Dark Matter
II. Topological Models
III. Deep networks

How to find NP

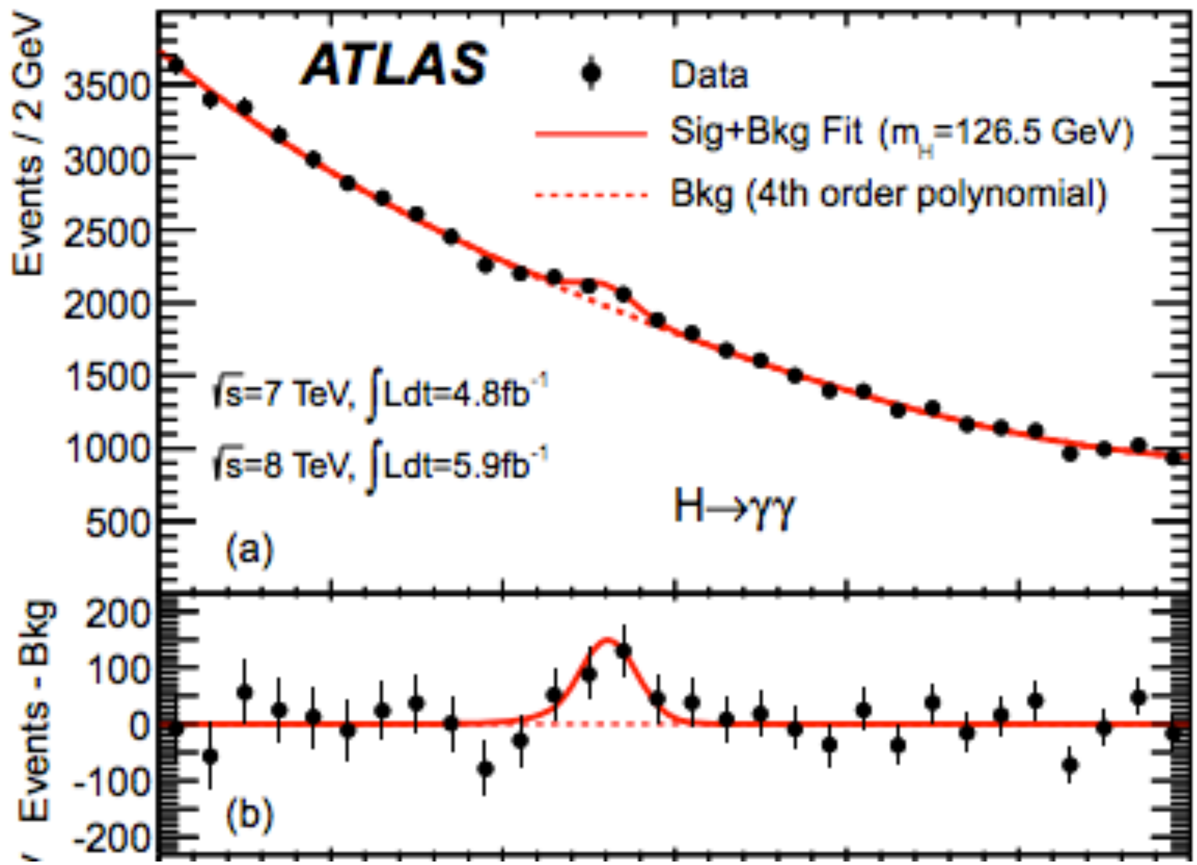
Isolate some feature in which two theories SM, SM+X can be best distinguished.



The data can tell us which hypothesis is preferred via a likelihood ratio:

$$\frac{L_{SM+X}}{L_{SM}} \qquad \frac{P(\text{data} \mid \text{SM+X})}{P(\text{data} \mid \text{SM})}$$

e.g.



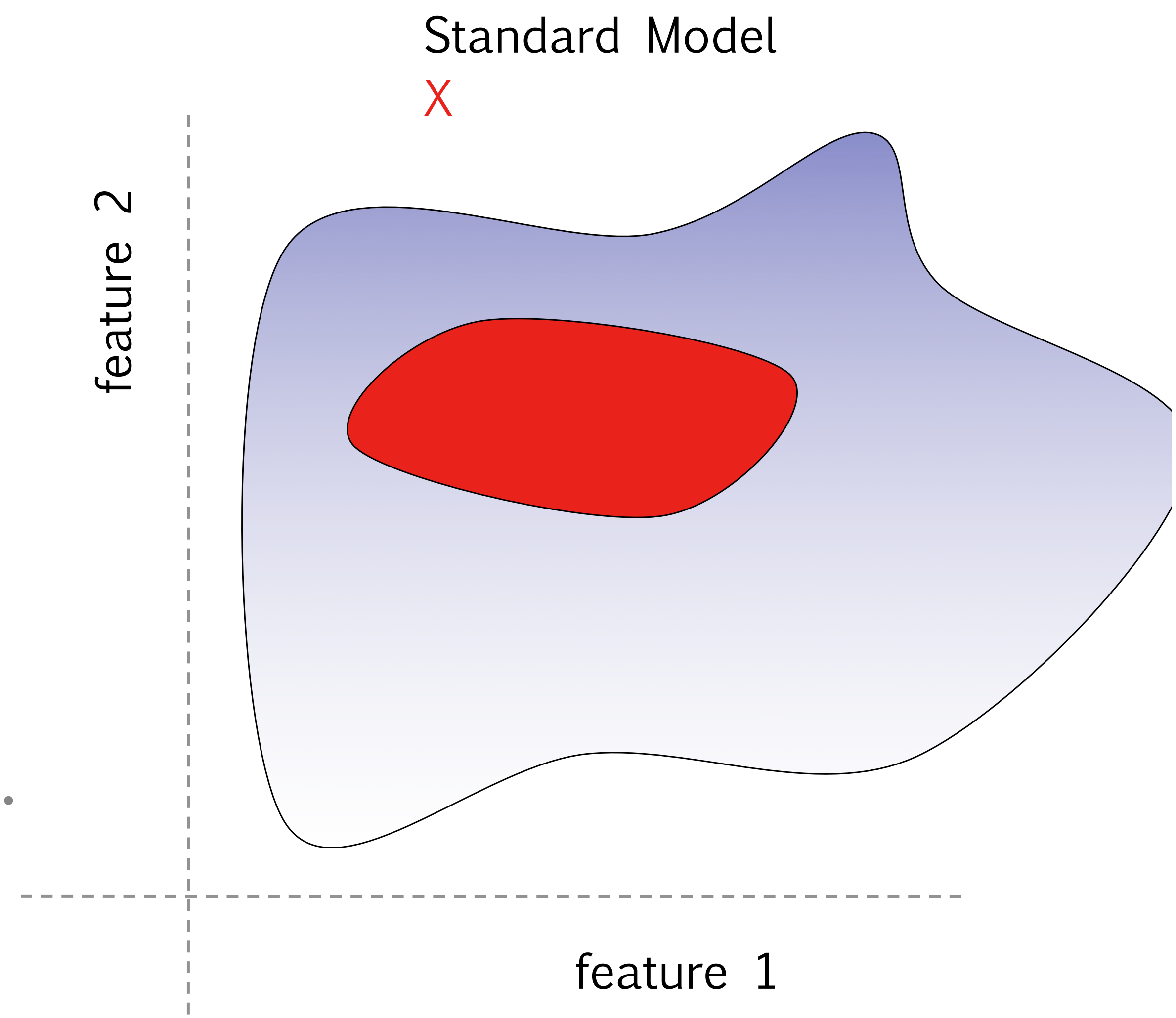
ATLAS
Data
Sig+Bkg Fit (m_H=126.5 GeV)
Bkg (4th order polynomial)
$\sqrt{s}$=7 TeV, $\int Ldt$=4.8fb^{-1}
$\sqrt{s}$=8 TeV, $\int Ldt$=5.9fb^{-1}
H→γγ
(a)
(b)
Events / 2 GeV
Events - Bkg
3500
3000
2500
2000
1500
1000
500
200
100
0
-100
-200

But...

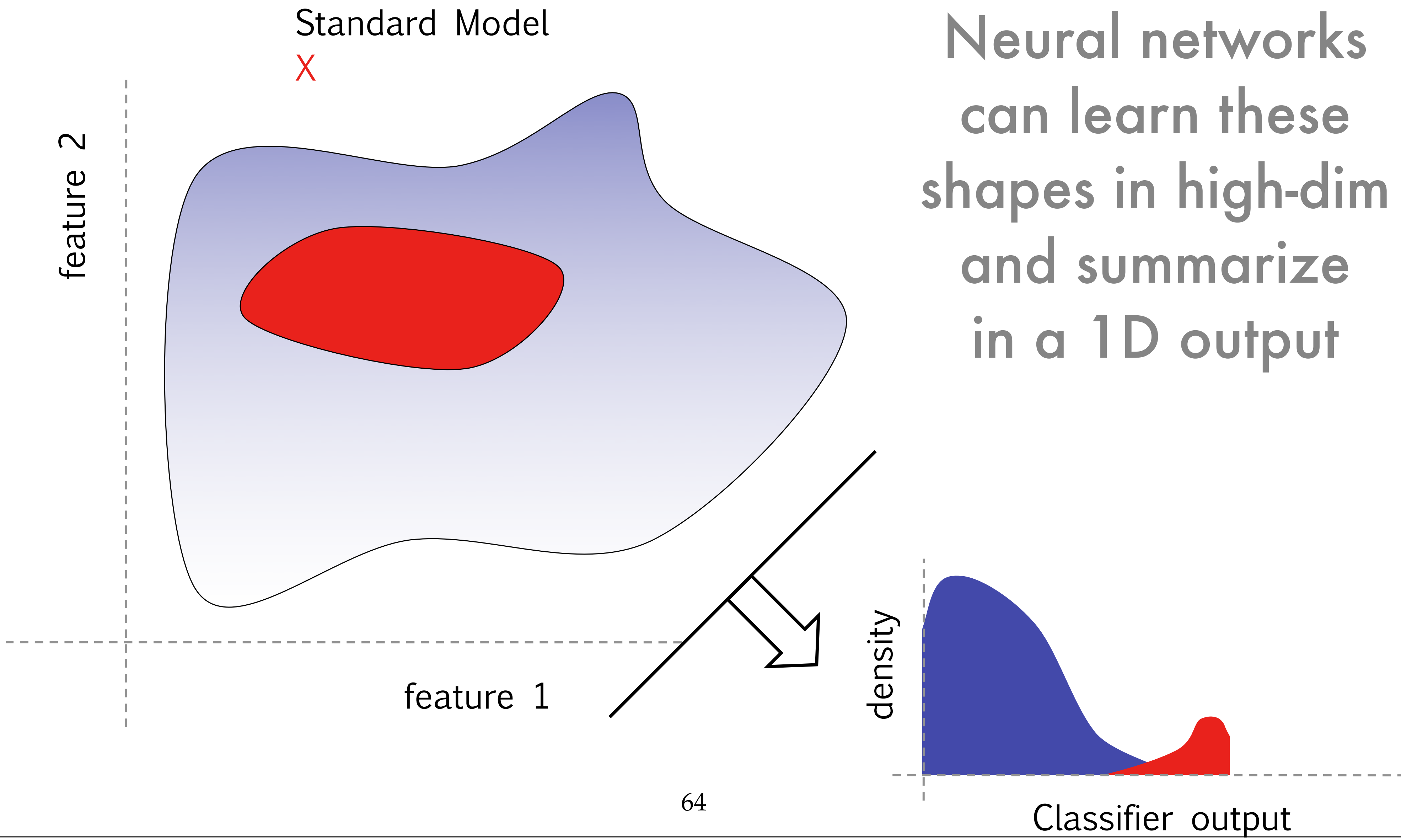
Reality is more complicated.

The full space can be very high dimensional.

Calculating likelihood in d-dimensional space requires ~100^d MC events.



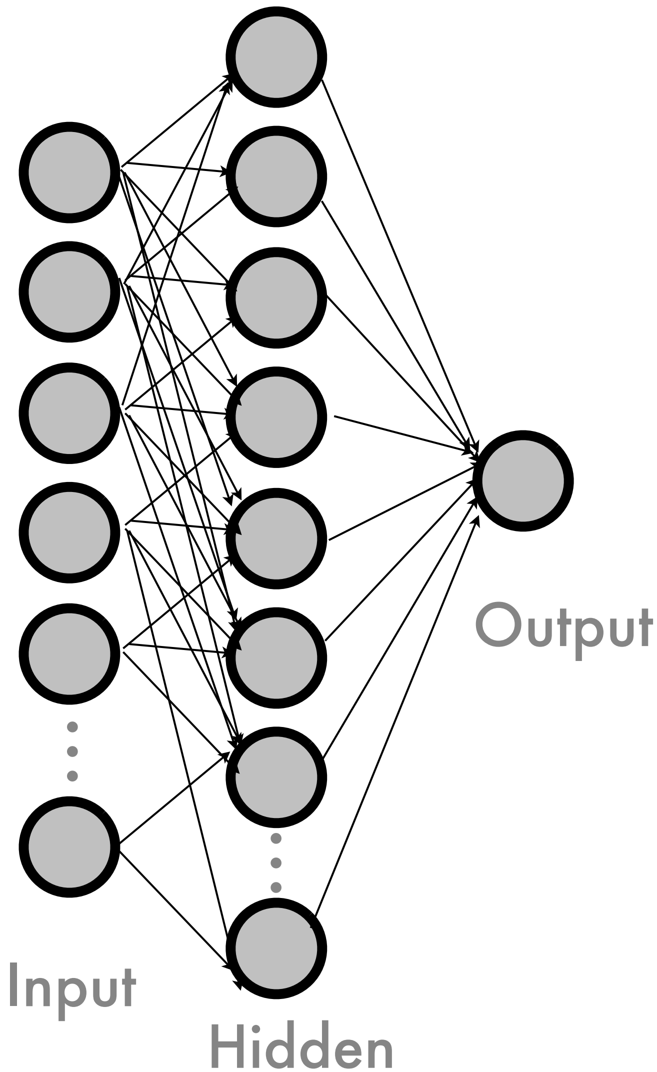
ML tools



Neural networks can learn these shapes in high-dim and summarize in a 1D output

Neural Networks

Essentially a functional fit with many parameters



Function

Each neuron's output is a function of the weighted sum of inputs.

Goal

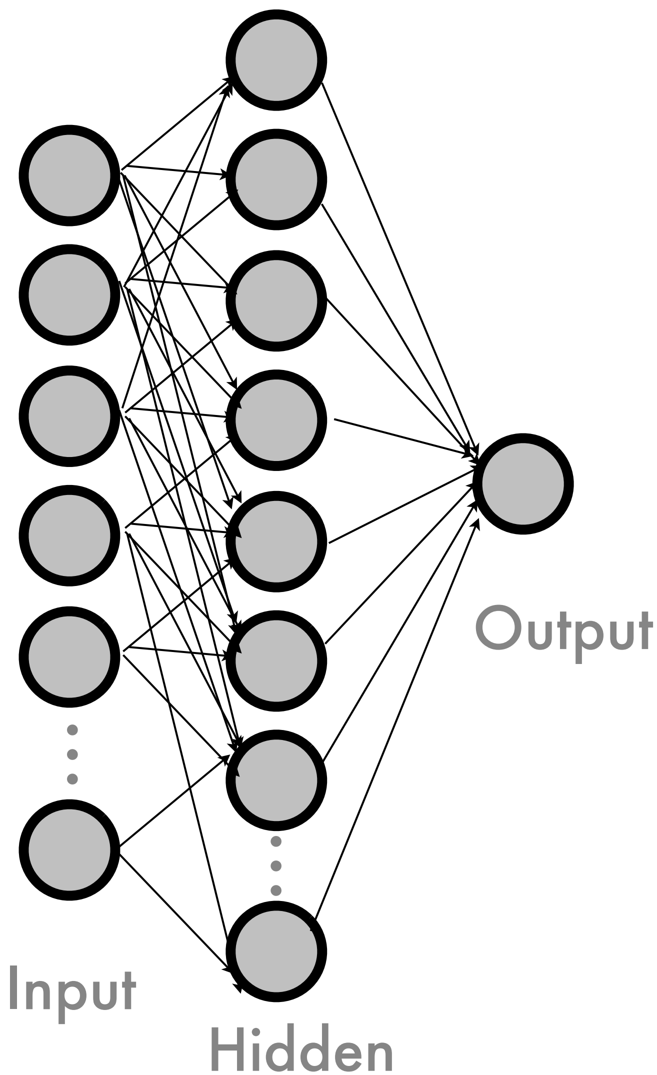
find set of weights which give most useful function

Learning

give examples, back-propagate error to adjust weights

Neural Networks

Essentially a functional fit with many parameters



Problem:

Networks with > 1 layer are very difficult to train.

Consequence:

Networks are not good at learning non-linear functions.

(like invariant masses!)

In short:

Can't just throw 4-vectors at NN.

Search for Input

ATLAS-CONF-2013-108

Can’t just use 4v

Can’t give it too many inputs

Painstaking search through input feature space.

Variable	VBF			Boosted		
	$\tau_{\rm lep}\tau_{\rm lep}$	$\tau_{\rm lep}\tau_{\rm had}$	$\tau_{\rm had}\tau_{\rm had}$	$\tau_{\rm lep}\tau_{\rm lep}$	$\tau_{\rm lep}\tau_{\rm had}$	$\tau_{\rm had}\tau_{\rm had}$
$m_{\tau\tau}^{\rm MMC}$	•	•	•	•	•	•
$\Delta R(\tau,\tau)$	•	•	•		•	•
$\Delta\eta(j_1,j_2)$	•	•	•			
m_{j_1,j_2}	•	•	•			
$\eta_{j_1}\times\eta_{j_2}$		•	•			
$p_{\rm T}^{\rm Total}$		•	•			
sum $p_{\rm T}$					•	•
$p_{\rm T}(\tau_1)/p_{\rm T}(\tau_2)$					•	•
$E_{\rm T}^{\rm miss}\phi$ centrality		•	•	•	•	•
$x_{\tau 1}$ and $x_{\tau 2}$						•
$m_{\tau\tau,j_1}$				•		
m_{ℓ_1,ℓ_2}				•		
$\Delta\phi_{\ell_1,\ell_2}$				•		
sphericity				•		
$p_{\rm T}^{\ell_1}$				•		
$p_{\rm T}^{j_1}$				•		
$E_{\rm T}^{\rm miss}/p_{\rm T}^{\ell_2}$				•		
$m_{\rm T}$		•			•	
$\min(\Delta\eta_{\ell_1\ell_2,\rm jets})$	•					
j_3 η centrality	•					
$\ell_1\times\ell_2$ η centrality	•					
ℓ η centrality		•				
$\tau_{1,2}$ η centrality			•			

Table 3: Discriminating variables used for each channel and category. The filled circles identify which variables are used in each decay mode. Note that variables such as $\Delta R(\tau,\tau)$ are defined either between the two leptons, between the lepton and $\tau_{\rm had}$, or between the two $\tau_{\rm had}$ candidates, depending on the decay mode.

Search for Input

ATLAS-CONF-2013-108

Can't just use 4v

Can't give it t[illegible] many inp[illegible]

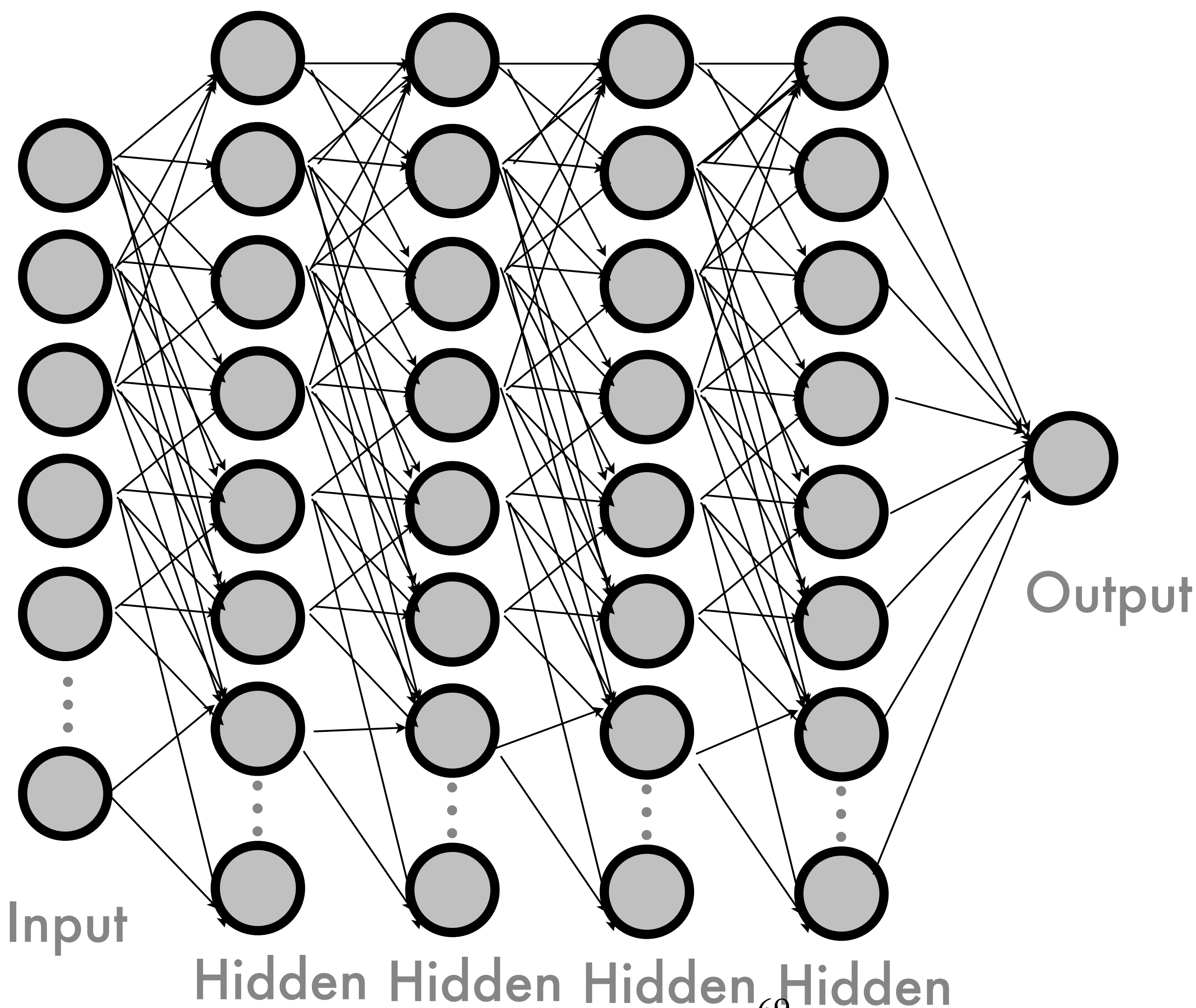
Painstaking[illegible] through inp[illegible] feature space.

Also true for BDTs, SVNs, etc

Variable	VBF $\tau_{lep}\tau_{lep}$	VBF $\tau_{lep}\tau_{had}$	VBF $\tau_{had}\tau_{had}$	Boosted $\tau_{lep}\tau_{lep}$	Boosted $\tau_{lep}\tau_{had}$	Boosted $\tau_{had}\tau_{had}$
$m_{\tau\tau}^{MMC}$	[illegible]	[illegible]	[illegible]	•	•	•
ΔR[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	•	•
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	•	•
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	•	•
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	•	•
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		•
[illegible]	[illegible]	[illegible]	[illegible]			
[illegible]	[illegible]	[illegible]	[illegible]	•		
[illegible]	[illegible]	[illegible]		•		
[illegible]	[illegible]			•		
[illegible]				•		
$p_T^{j_1}$				•		
$E_T^{miss}/p_T^{\ell_2}$				•		
m_T		•			•	
$\min(\Delta\eta_{\ell_1\ell_2,\mathrm{jets}})$	•					
j_3 η centrality	•					
$\ell_1 \times \ell_2$ η centrality	•					
ℓ η centrality		•				
$\tau_{1,2}$ η centrality			•			

Table 3: Discriminating variables used for each channel and category. The filled circles identify which variables are used in each decay mode. Note that variables such as $\Delta R(\tau,\tau)$ are defined either between the two leptons, between the lepton and τ_{had}, or between the two τ_{had} candidates, depending on the decay mode.

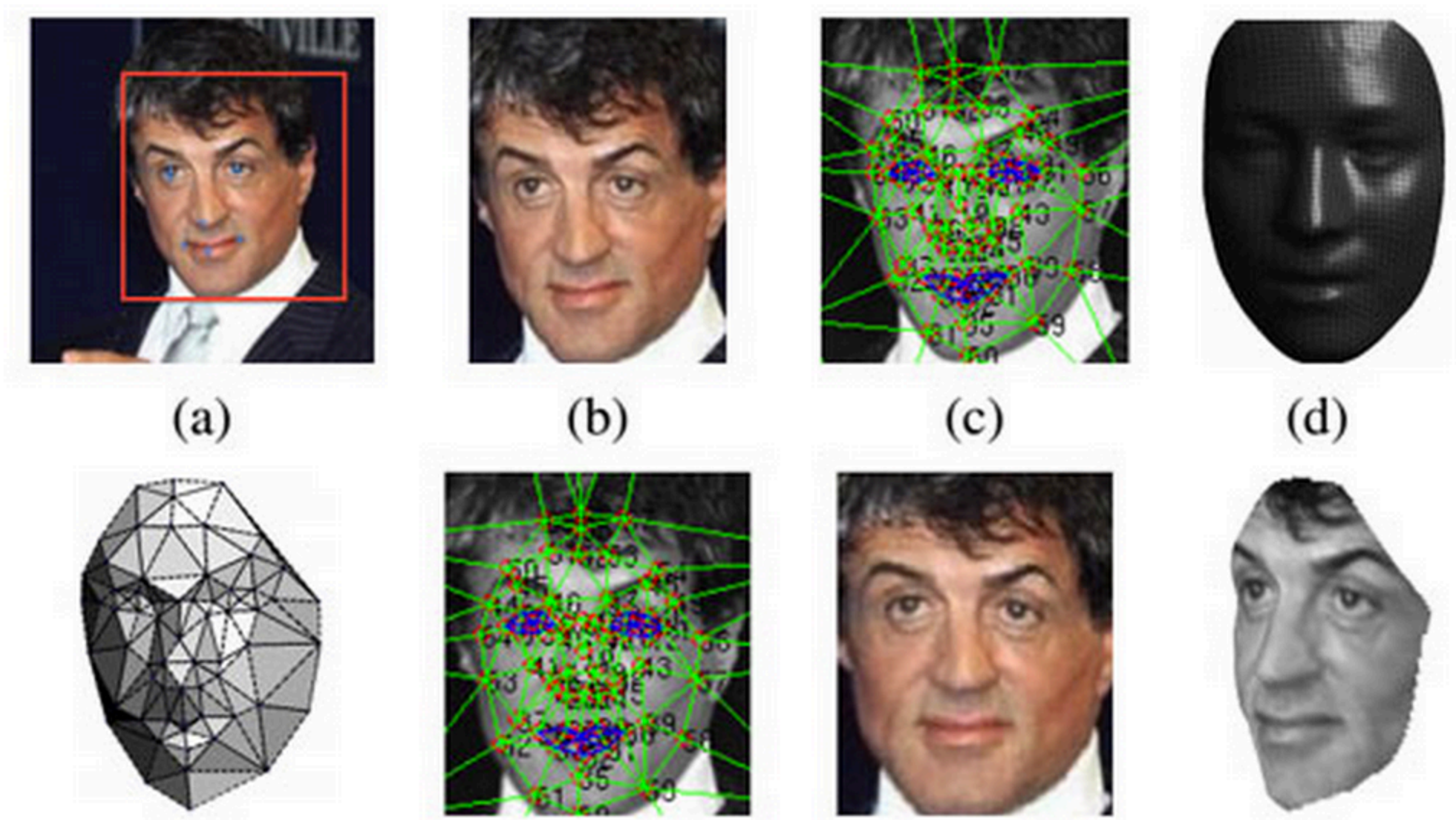
Deep networks



New tools let us train deep networks.

How well do they work?

Real world applications

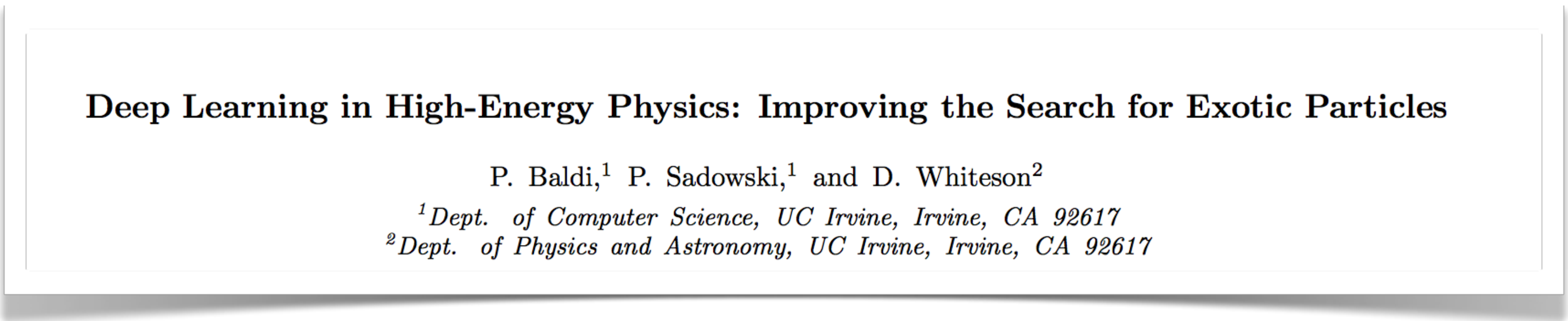


(a) (b) (c) (d)

(e) (f) (g) (h)

Head turn: DeepFace uses a 3-D model to rotate faces, virtually, so that they face the camera. Image (a) shows the original image, and (g) shows the final, corrected version.

Paper


Deep Learning in High-Energy Physics: Improving the Search for Exotic Particles

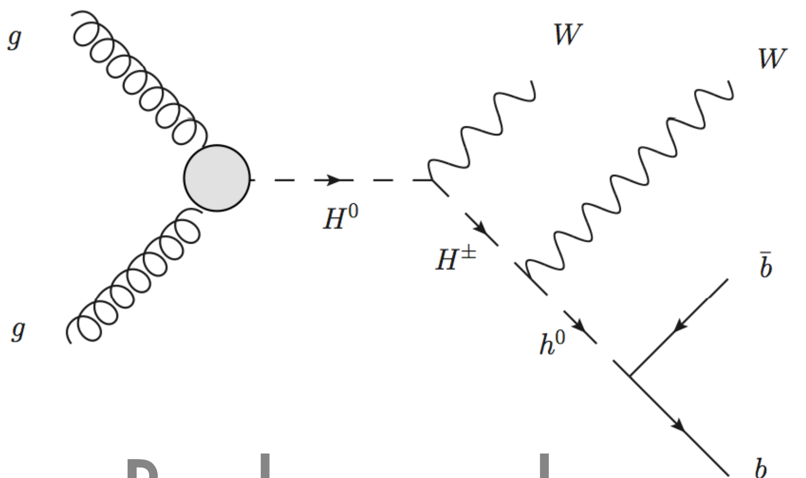
P. Baldi,[1] P. Sadowski,[1] and D. Whiteson[2]
[1]Dept. of Computer Science, UC Irvine, Irvine, CA 92617
[2]Dept. of Physics and Astronomy, UC Irvine, Irvine, CA 92617

arXiv: 1402.4735

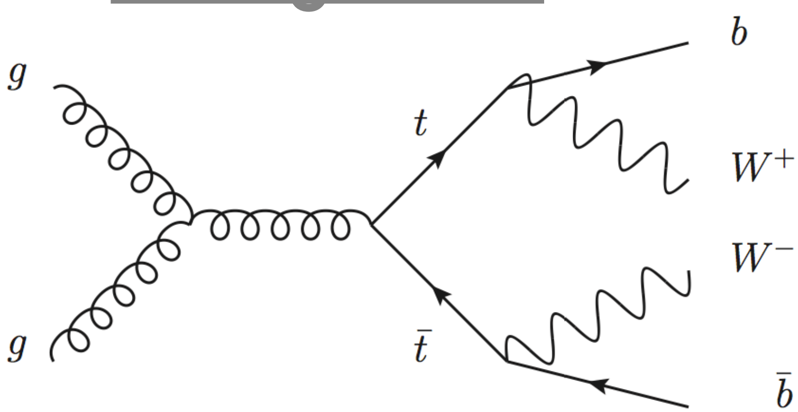
In revision at *Nature Comm.*

Benchmark problem

Signal



Background



Can deep networks automatically discover useful variables?

4-vector inputs

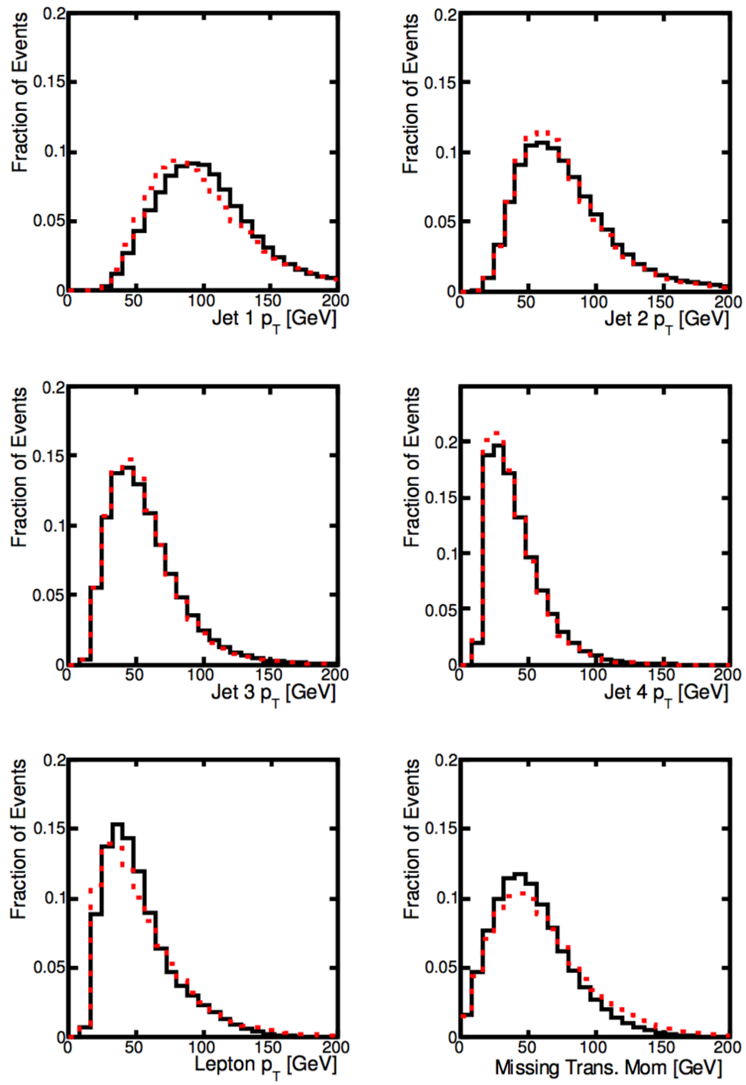
<u>21 Low-level vars</u>

jet+lepton mom. (3x5)

missing ET (2)

jet btags (4)

Not much separation visible in 1D projections

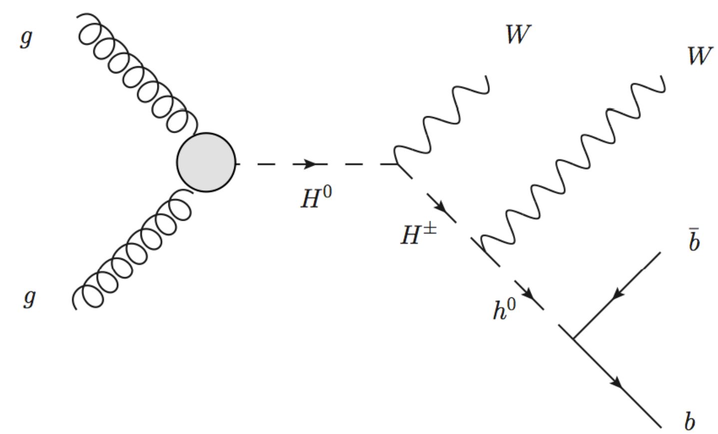


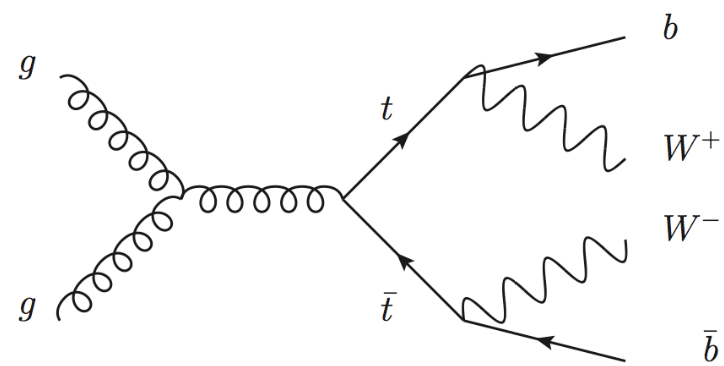
4-vector inputs

<u>7 High-level vars</u>

m(WWbb)
m(Wbb)
m(bb)

m(bjj)
m(jj)
m(lv)
m(blv)





4-vector inputs

7 High-level vars

m(WWbb)

m(Wbb)

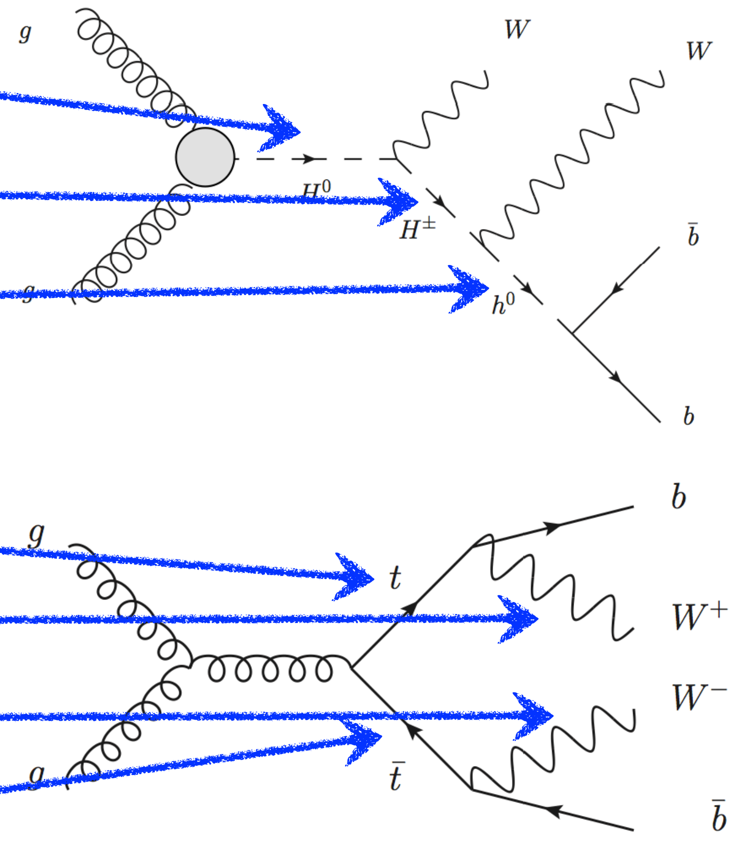
m(bb)

m(bjj)

m(jj)

m(lv)

m(blv)

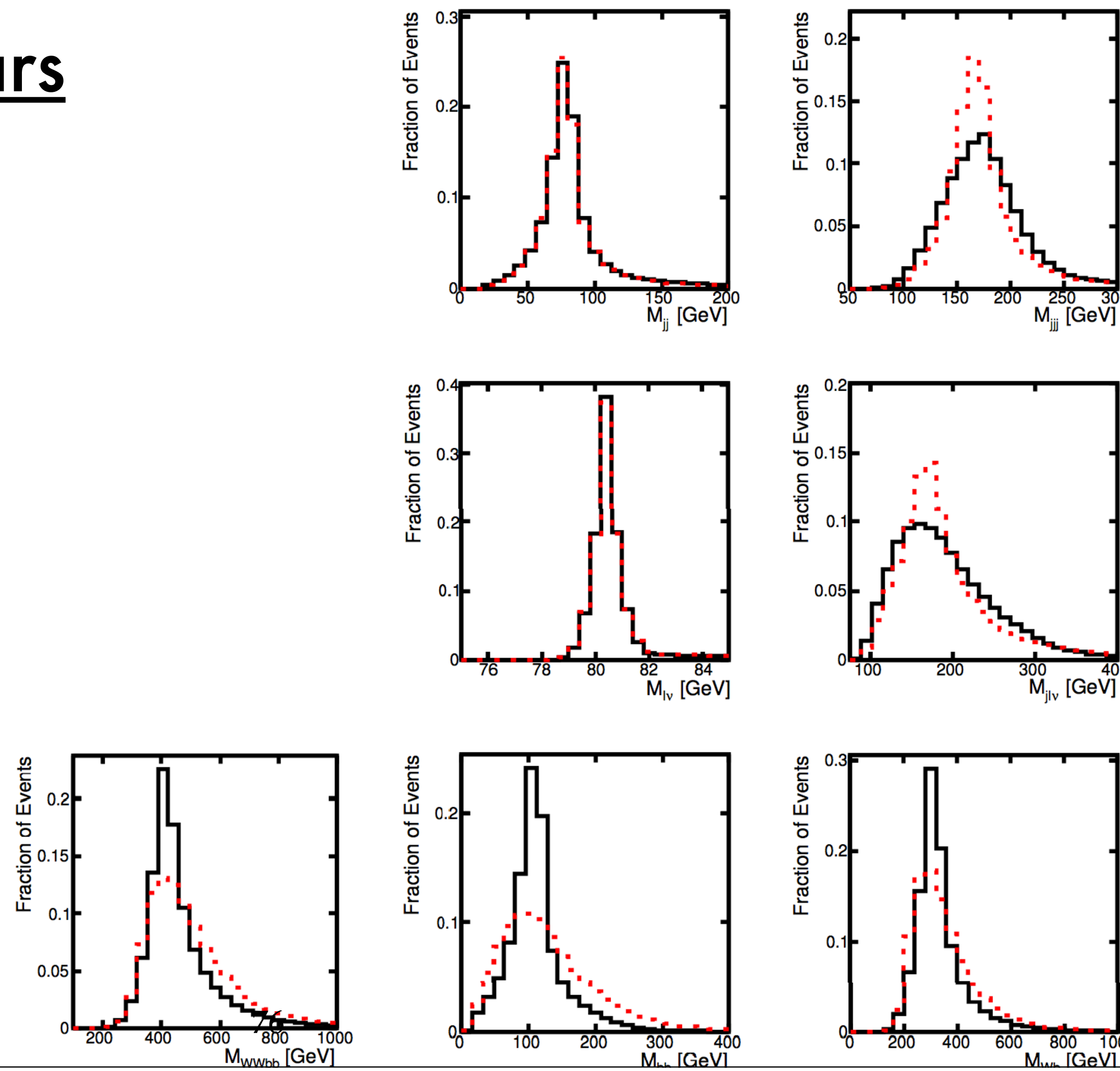


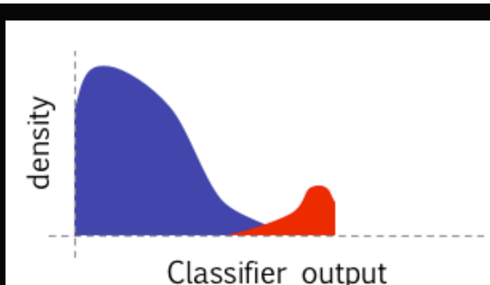
4-vector inputs

<u>7 High-level vars</u>

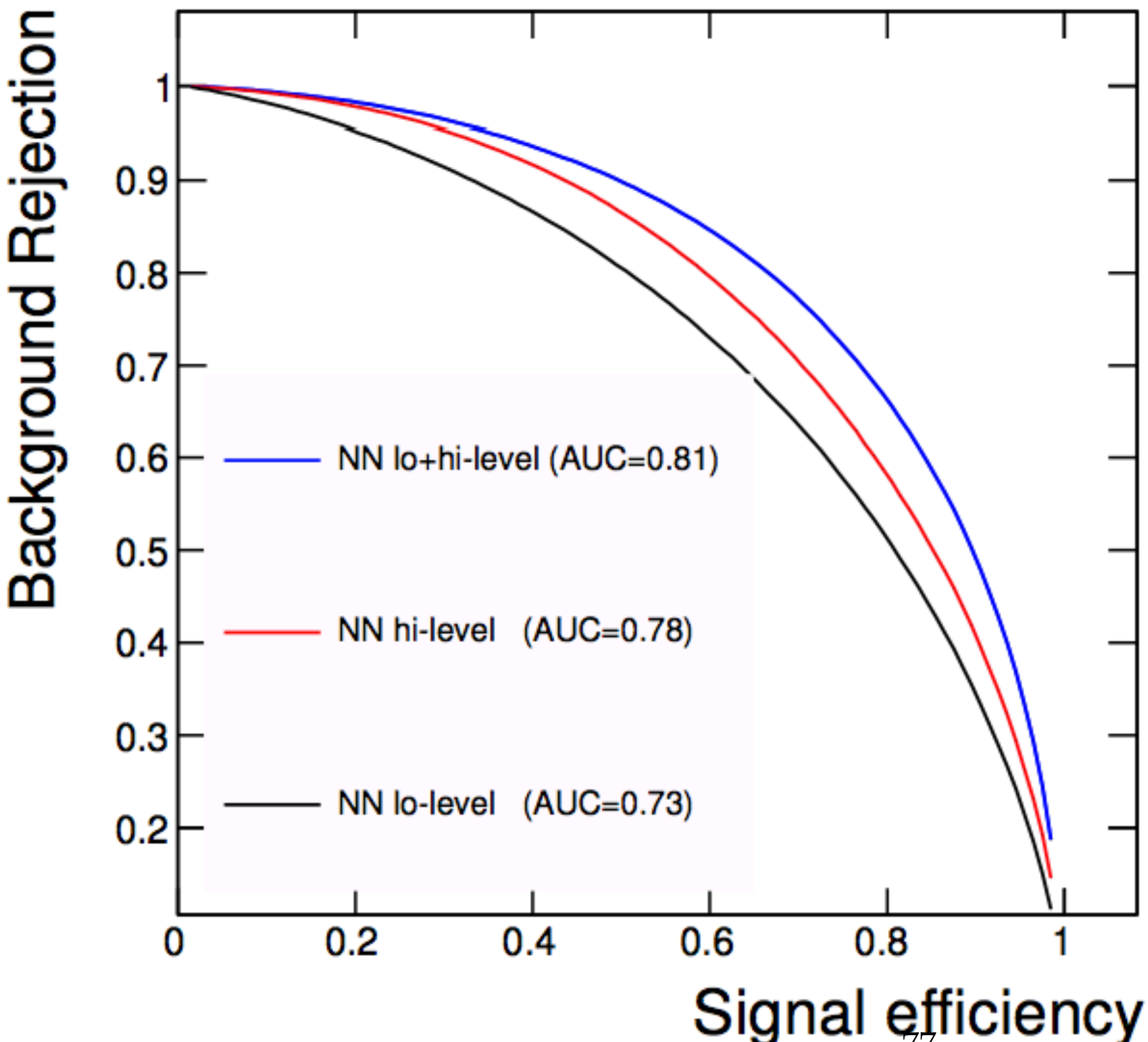
m(WWbb)
m(Wbb)
m(bb)

m(bjj)
m(jj)
m(lv)
m(blv)





Standard NNs



Results

Adding hi-level boosts performance

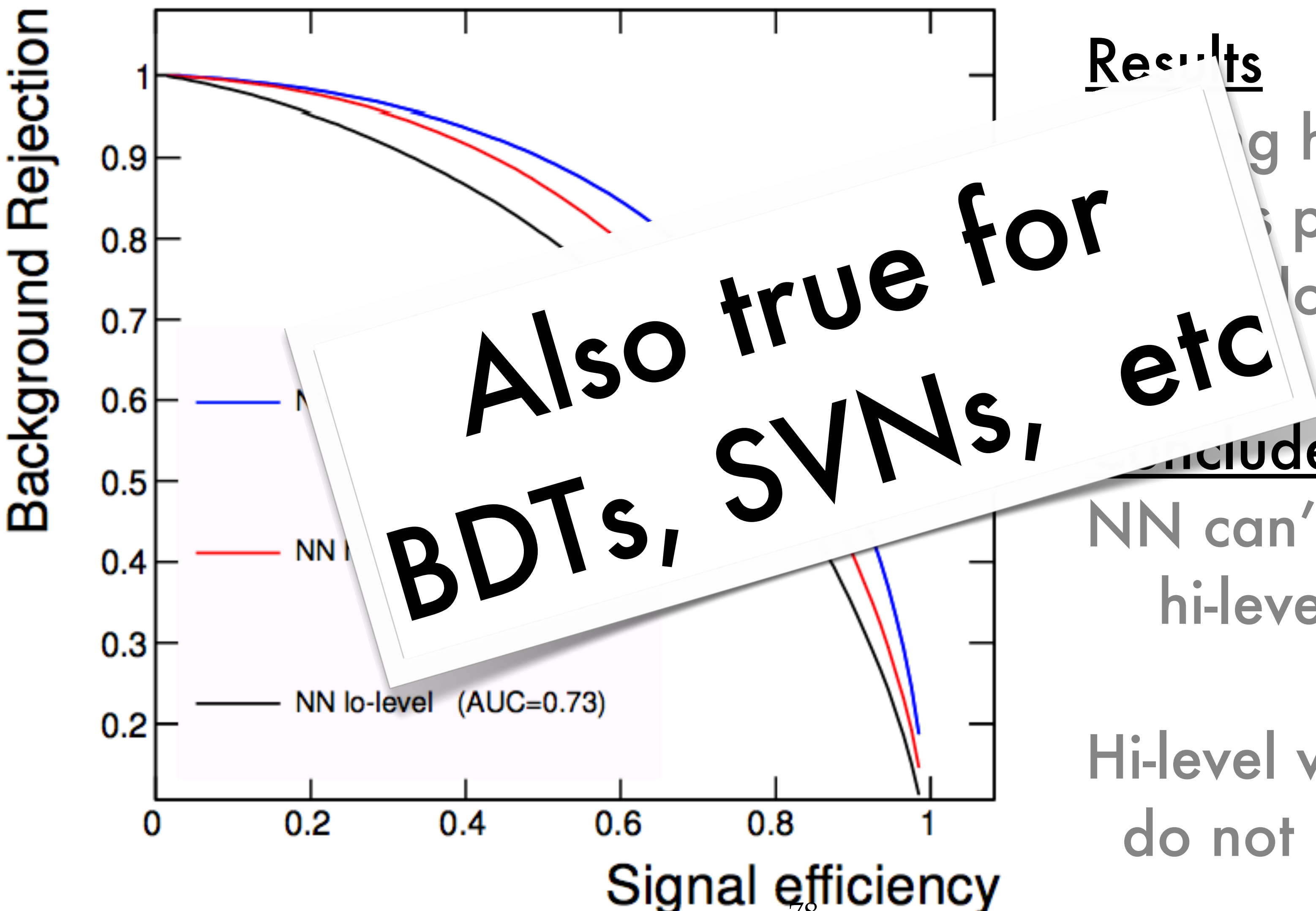
Better: lo+hi-level.

Conclude:

NN can't find hi-level vars.

Hi-level vars do not have all info

Standard NNs



Results

g hi-level

performance

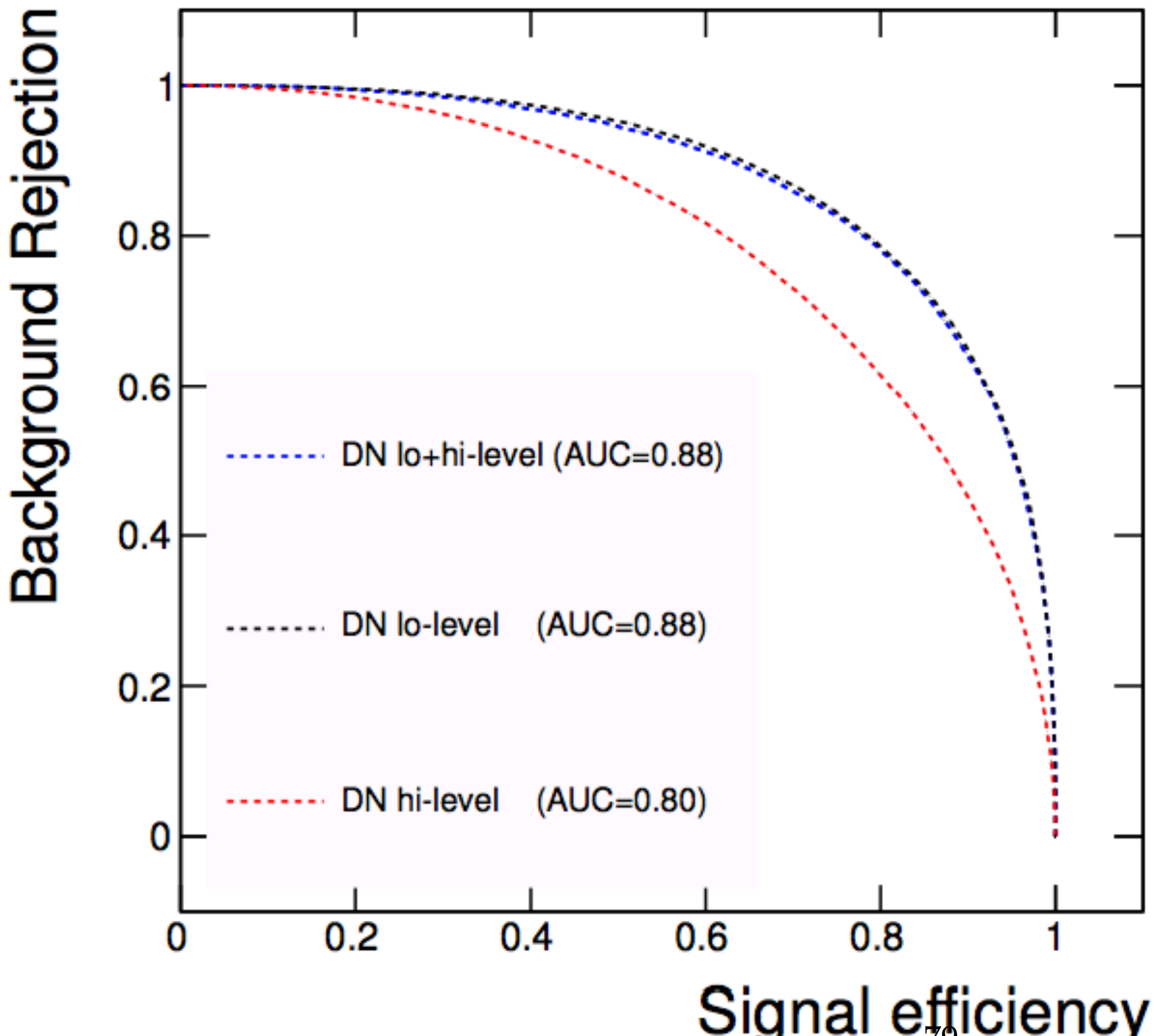
lo+hi-level.

nclude:

NN can't find hi-level vars.

Hi-level vars do not have all info

Deep Networks



Results

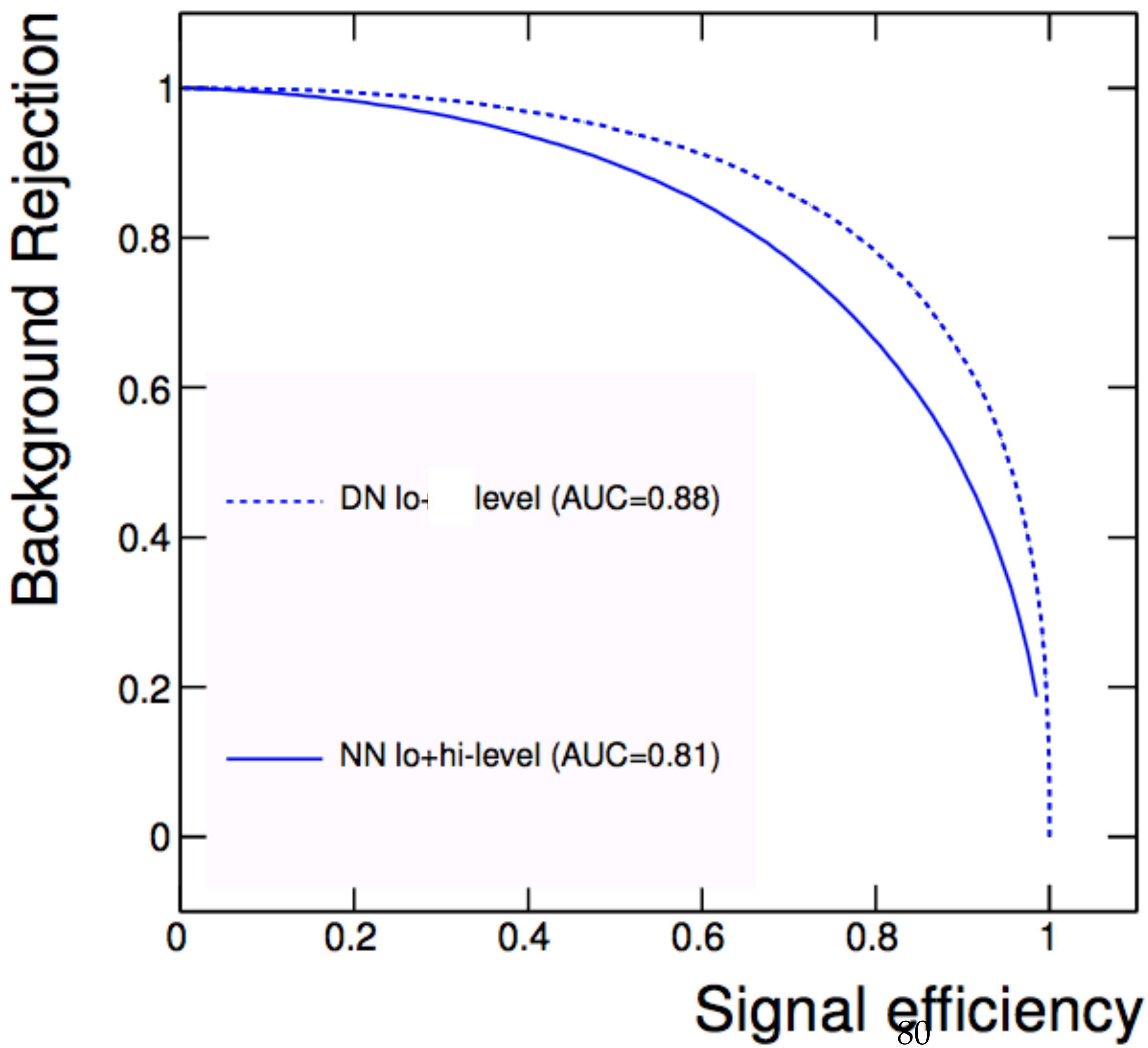
Lo+hi = lo.

Conclude:

DN can find hi-level vars.

Hi-level vars do not have all info are unnecessary

Deep Networks



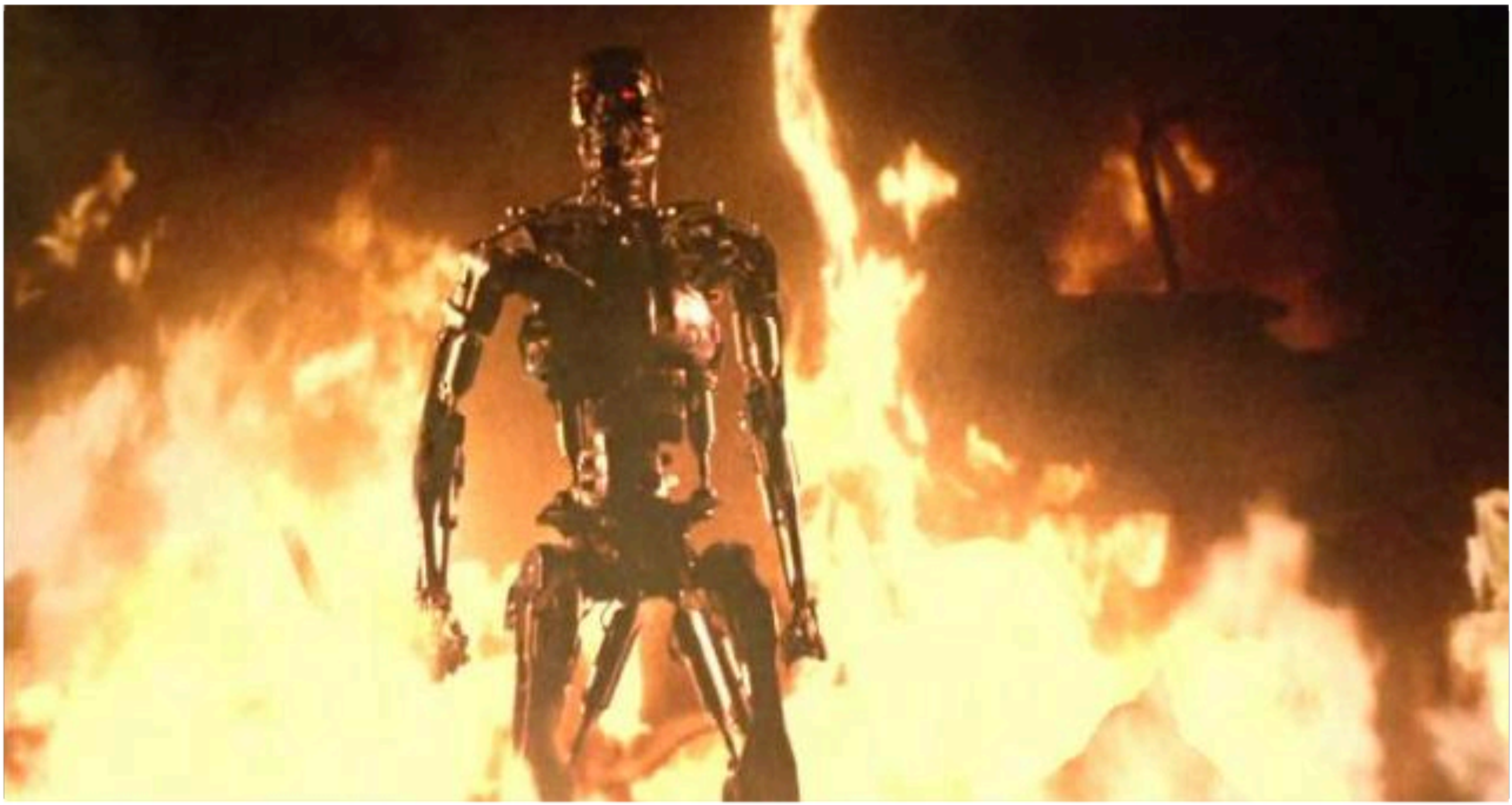
Results

DN > NN

Conclude:

DN does better than human assisted NN

The AIs win



Results

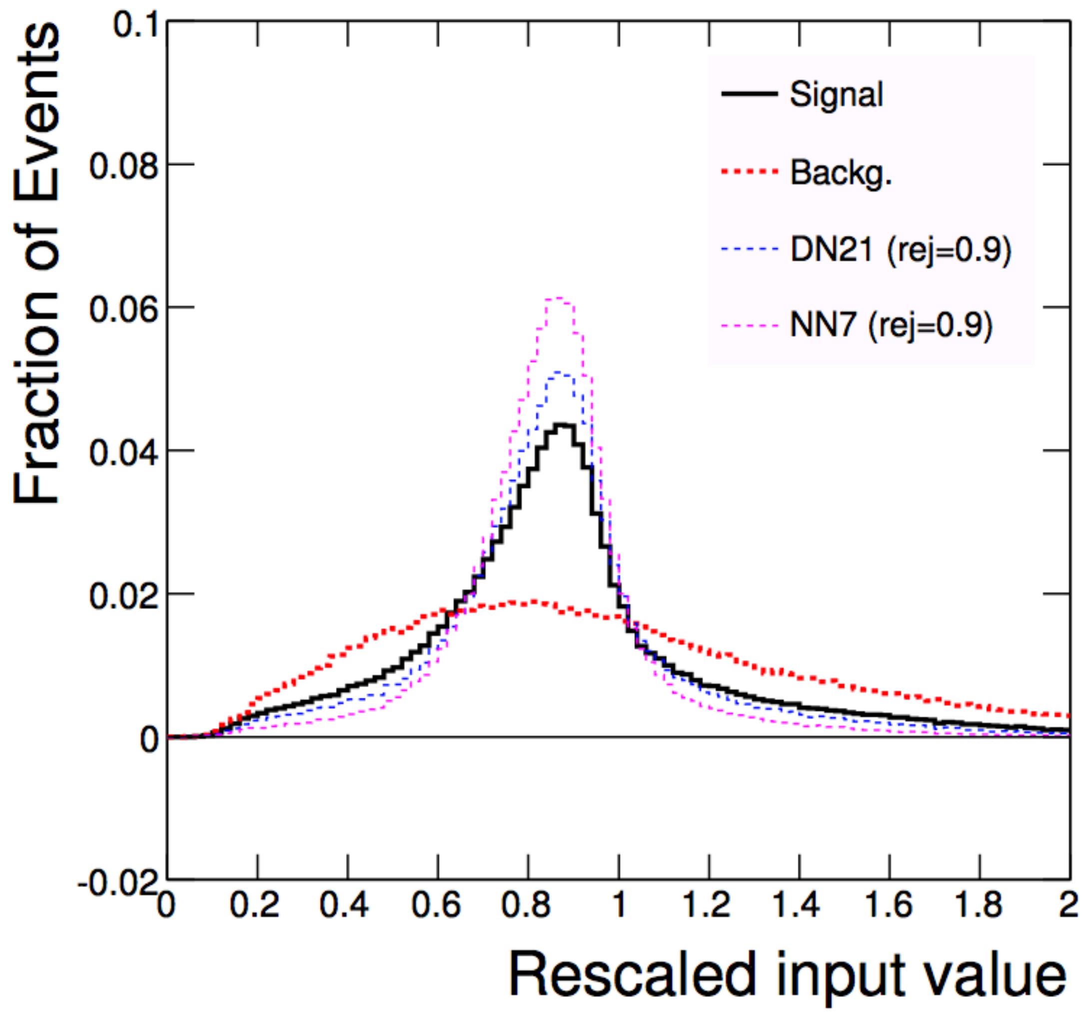
Identified example benchmark where traditional NNs fail to discover all discrimination power.

Adding human insight helps traditional NNs.

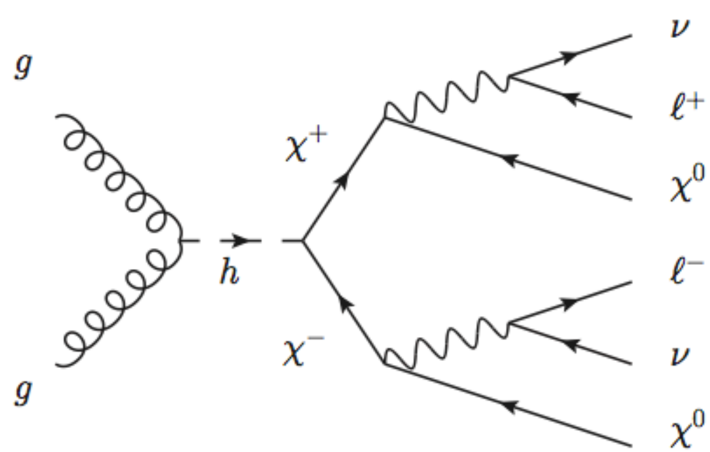
Deep networks succeed without human insight. Outperform human-boosted traditional NNs.

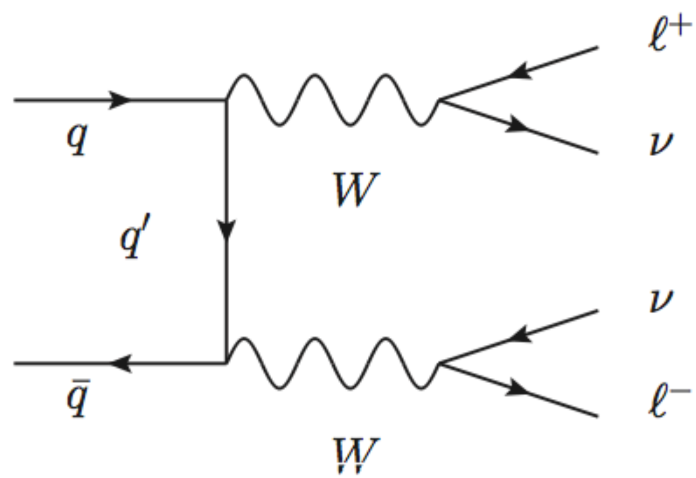
Why?

DN not as reliant on signal features. Cuts into background space.



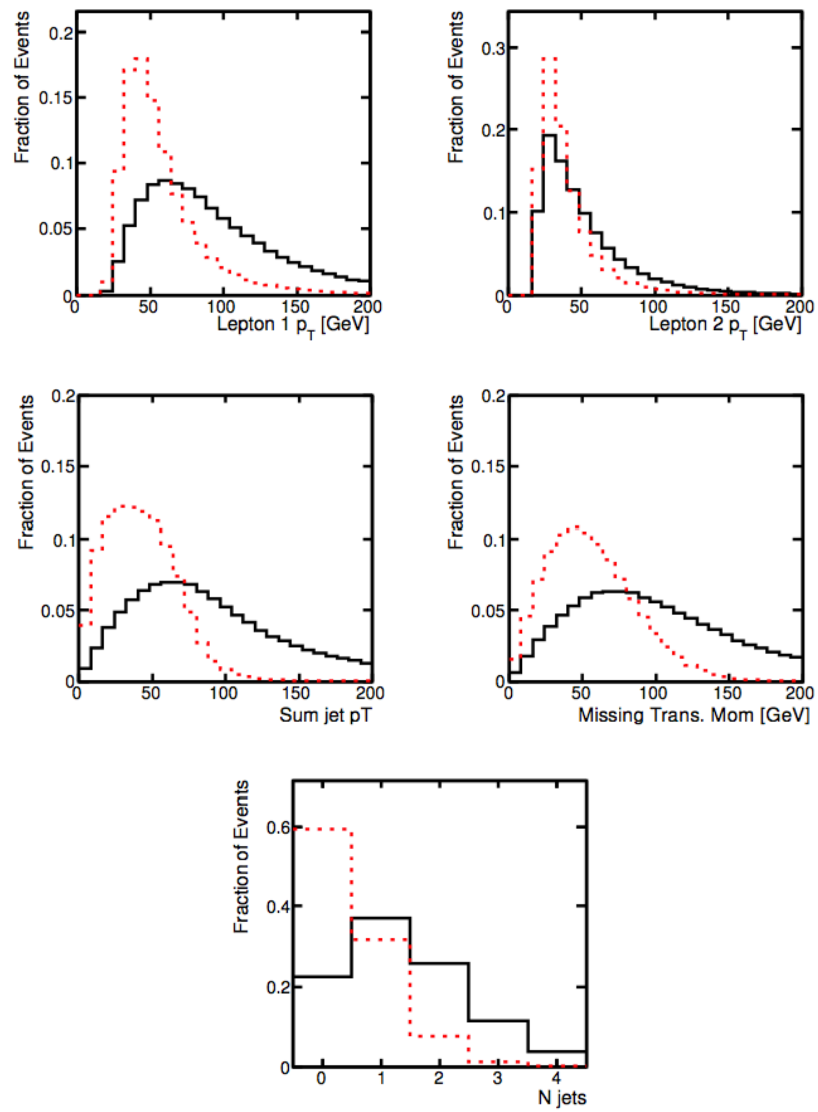
2nd case: SUSY





Can
Deep Networks
help us
find SUSY
in the data?

Low-level variables



High-level variables

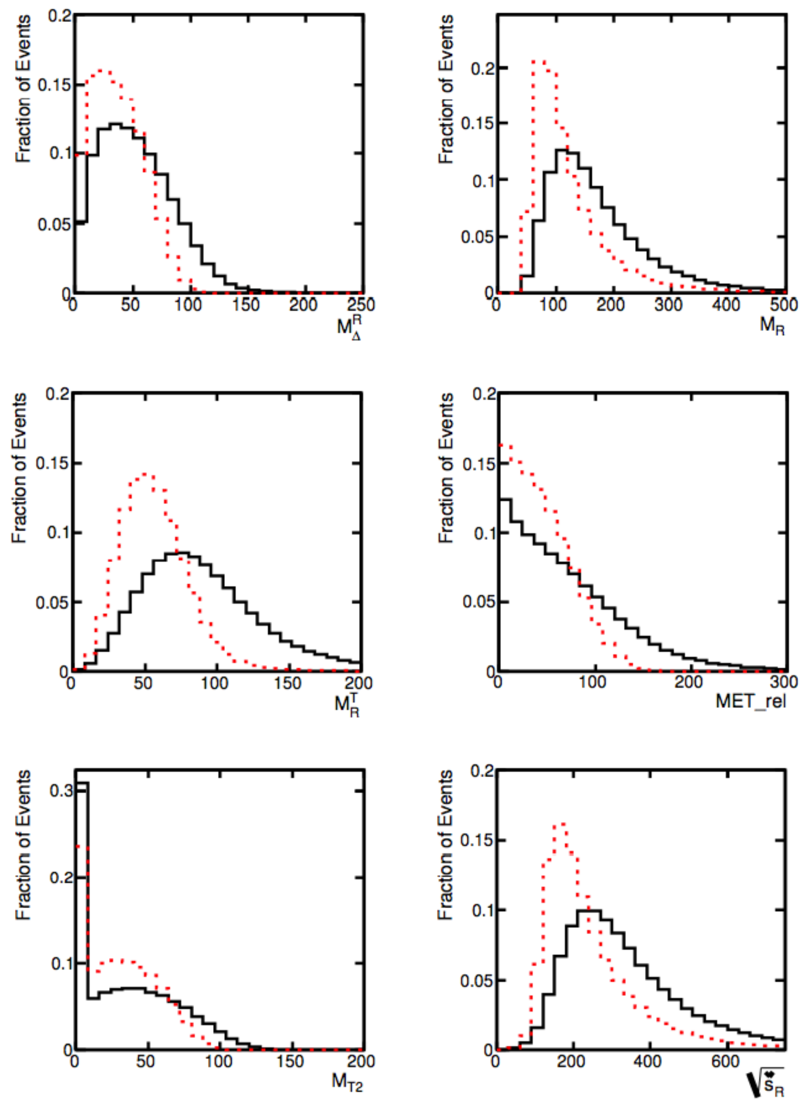
Axial-MET

Met-rel

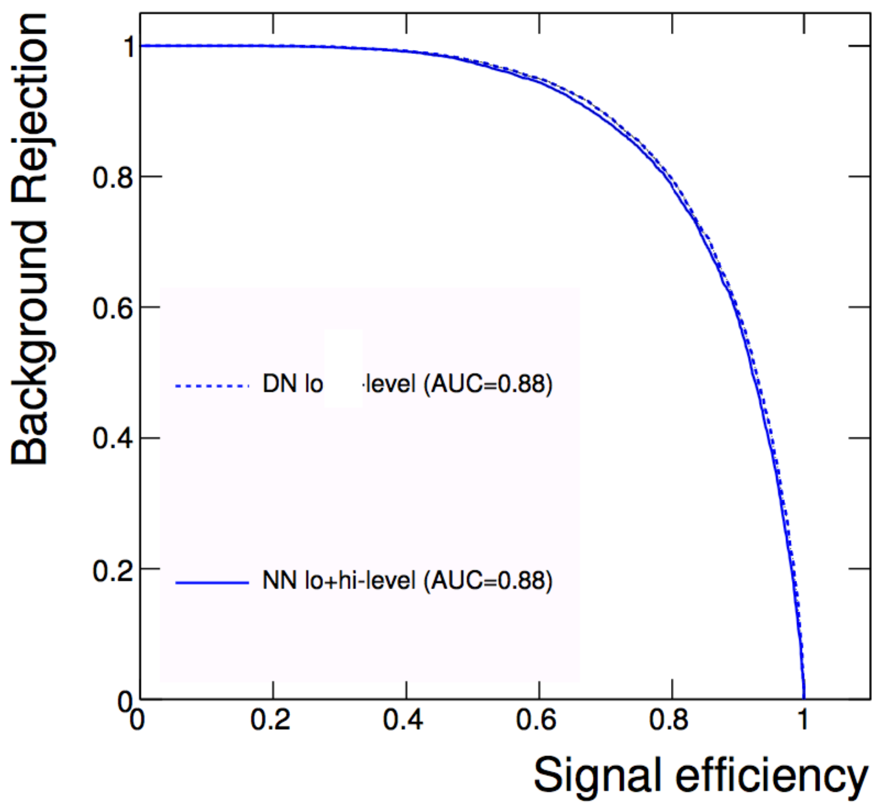
MT2

Razor

Super-razor



SUSY results



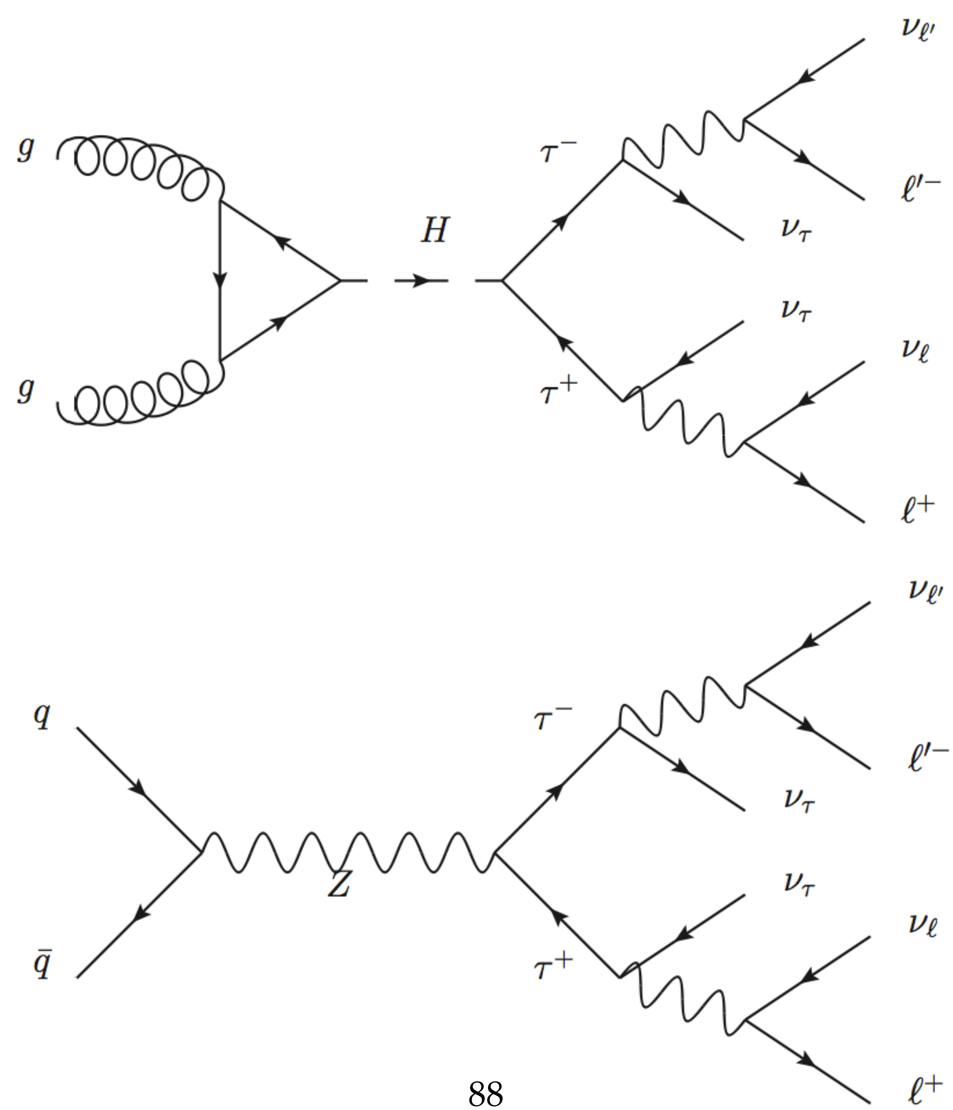
DN doesn’t need help

Outperforms human assisted NN

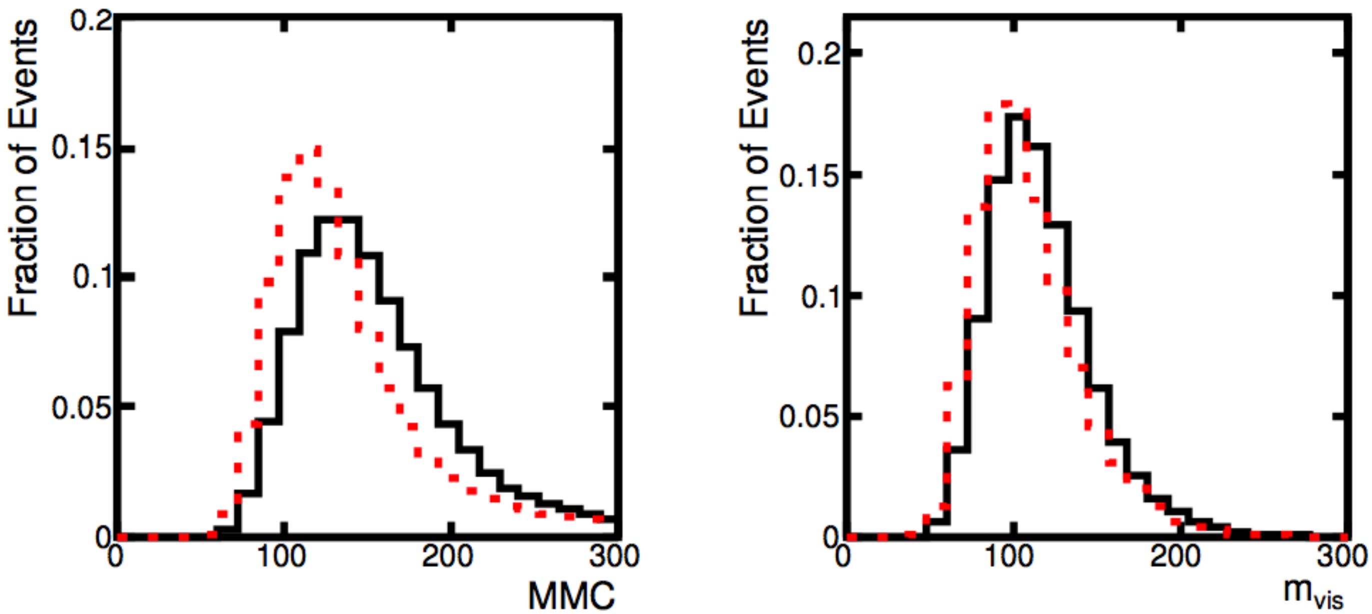
Margin is smaller
-> high level variables
are less helpful
and less needed!

Discovery significance			
Technique	Low-level	High-level	Complete
NN	6.5σ	6.2σ	6.9σ
DN	7.5σ	7.3σ	7.6σ

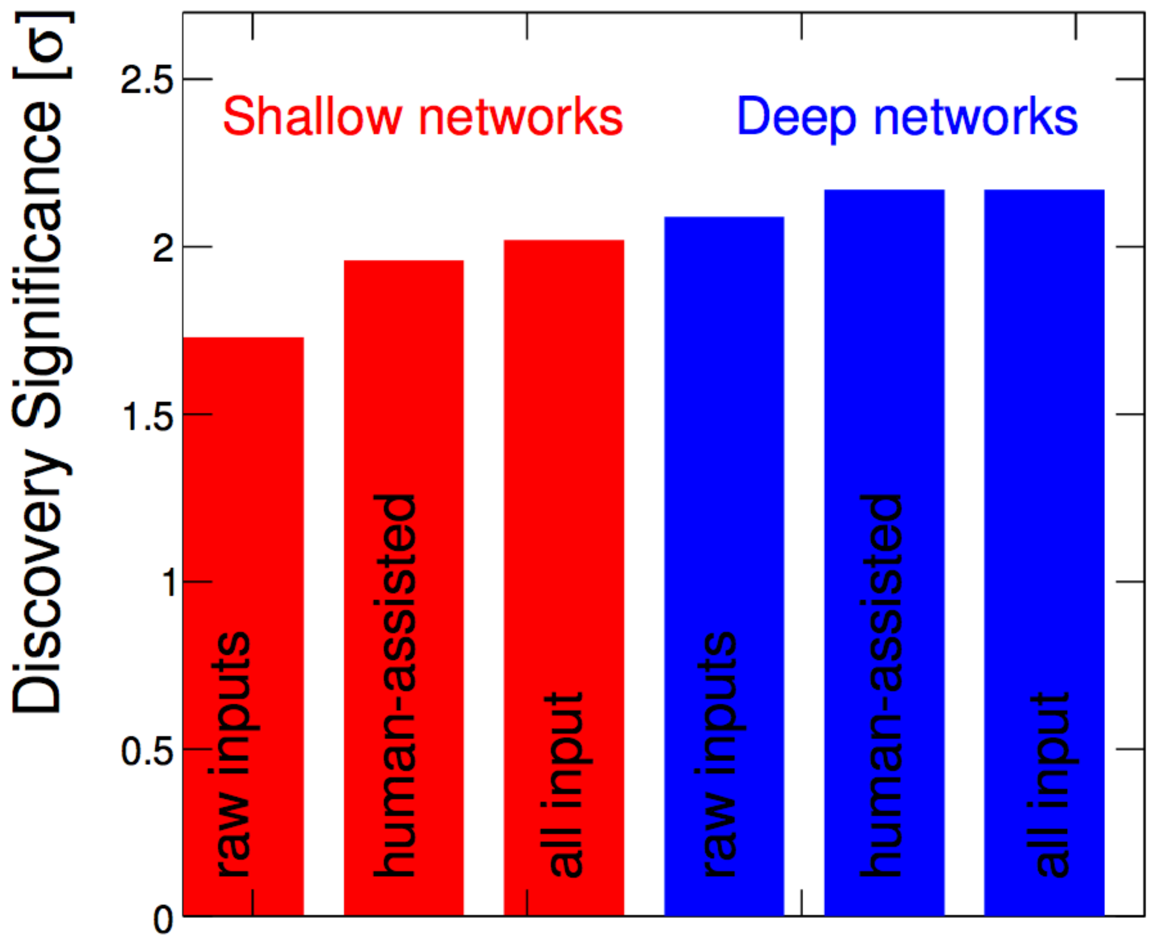
Preliminary: h->tautau



Preliminary: h->tautau



Preliminary: h->tautau



Summary

Dark matter:

broad-based attack on all LHC signals

Topological models:

Strategy to build complete set of models with discoverable resonances

Deep networks:

Networks can take 4-vectors, find powerful discriminants