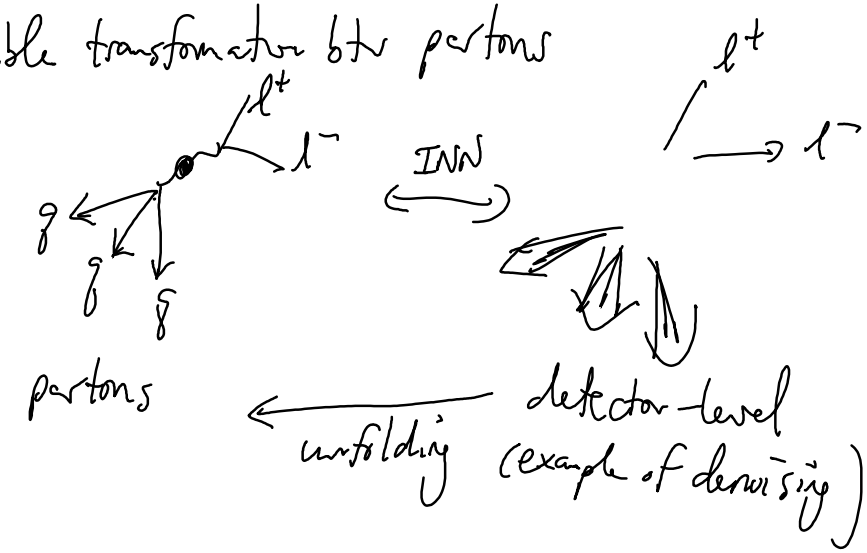
Friday, April 23, 2021 3:21 PM

lecture 25

Reminder: Please fill out the course survey!!

HEP applications of Normalizing Flows

- Anomaly Detection → see next lecture
- Unfolding (Heidelberg group): learn invertible transformation btw partons (?)



- Phase space integration/sampling (part of event generation)

Friday, April 23, 2021 3:34 PM

- 2001.10028 + Hoeche, Schulz
- 2001.05486 "iflow" Gao, Isaacson & Krause
- 2001.05478 Bothmann et al (similar idea)

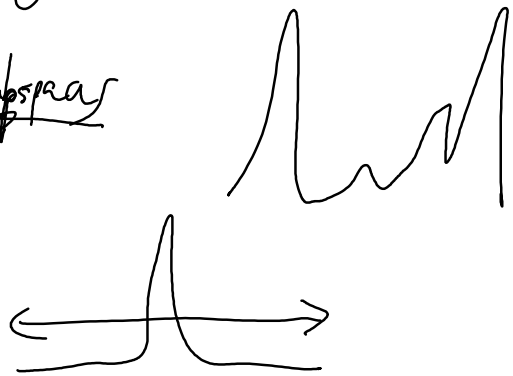
monte-carlo integration in high dimensions is expensive/inefficient

have to sample a lot if $f(x)$ is sharply peaked/localized on lower dim'l subspace and inefficiently

$\int d^d x\, f(x)$

multimodal

sparsity



idea: if $\int p(x)\, f(x)\, d^d x$

p ← probability distribution ← $\frac{d\sigma}{d\Gamma}$ ← phase space.

learn $p(x)$ to fit true $p(x)$

Importance sampling: $\frac{1}{N}\sum_{x_i \sim p(x)} f(x_i)$

Friday, April 23, 2021 3:58 PM

New (final) Topic: Anomaly Detection @ LHC

↕

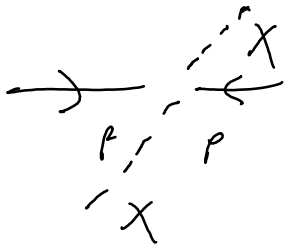
"Model-Independent Searches for NP @ LHC"

LHC — Huge dataset

How do we find anomalies (new physics) if we don't know what we're looking for?

If we do know what we're looking for

Suppose X is heavy, charged, stable

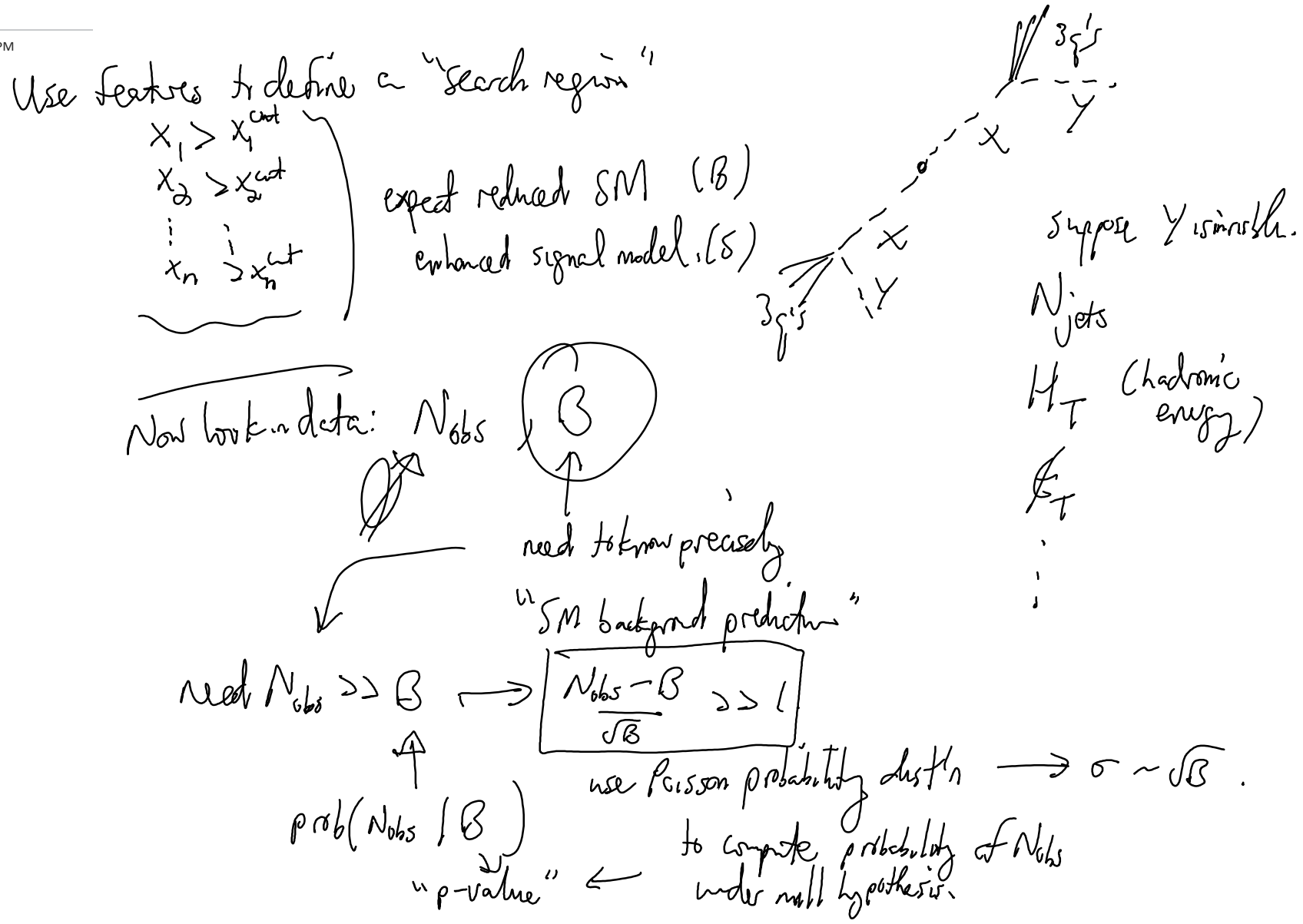


maybe know $\frac{d\sigma}{d\eta d\Omega}$ ... → signature in detector

characteristic feature of X event

e.g. hits in tracker

Friday, April 23, 2021 4:06 PM



Friday, April 23, 2021 4:16 PM

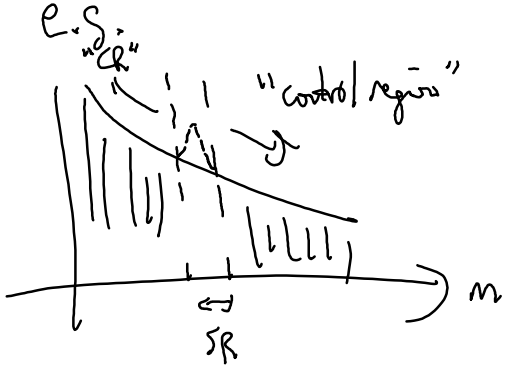
How to predict B?

→ rely on simulations
calculating SM $\frac{d\sigma}{d\vec{p}}$'s
generating events
simulating detector
...

generally not precise enough...

→ data-driven (don't rely on simulations)

"sideband bg estimation"



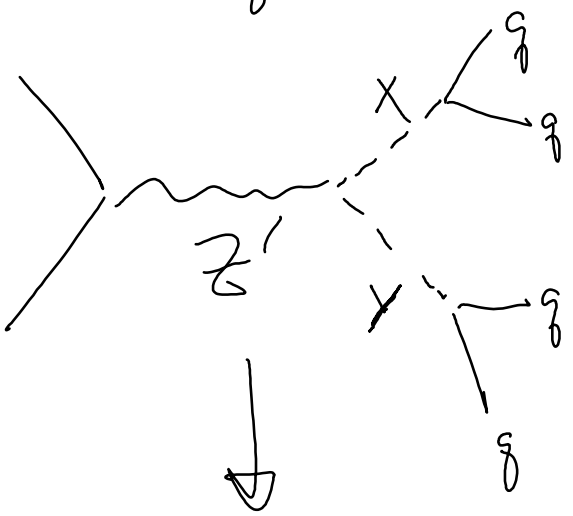
if m dist'n smooth for bg
and localized for signal
↓
fit m dist'n in CR
interpolate to get B in SR.

most analyses
combine data-driven
& simulation.

→ 99.99...% all searches for NP @ LHC
are like this.

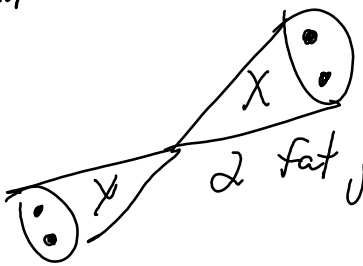
Friday, April 23, 2021 4:22 PM

Example of a signature that has no explicit search for it yet:



$m_{Z'} >> m_{X,Y}$

so X, Y decay products are collimated
are highly boosted



2 fat jets w/ 2-prong substructure.

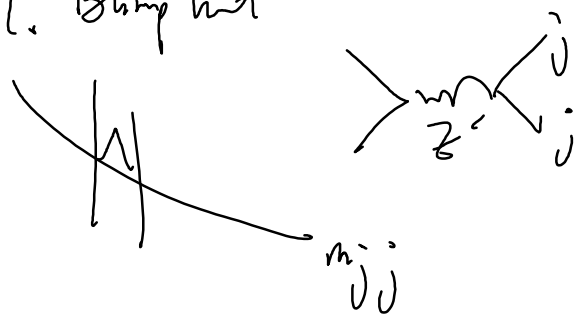
could be sitting in data w/ $>5\sigma$ significance

Q: Could we design a model-agnostic search that would discover this model?

Friday, April 23, 2021 4:27 PM

Existing model indep searches @ LHC — 2 kinds

1. Bump hunt



Almost fully model-independent
- Assume signal localized in m_{jj}
- Assume bg is smooth.

2. "general search" (ATLAS), "MUSIC" (CMS)

bin data into many categories (exclusive cuts) (thousands)

Compare against SM simulation in each bin

Compute p-values...

pros - signal model independent

cons - huge look elsewhere effect w/ so many bins
- heavily simulation dependent, systematics challenging...

Idea: train classifier on data (1) vs. simulation (0)
optimal for general search!

optimal classifier b/w 2 hypothesis is $\frac{p(x|H_1)}{p(x|H_0)}$ → $\frac{p(x|\text{data})}{p(x|SM)}$ optimal test for deviation from SM