Reweighting for Godot

Jamie Gainer
MC4BSM— May 20, 2015

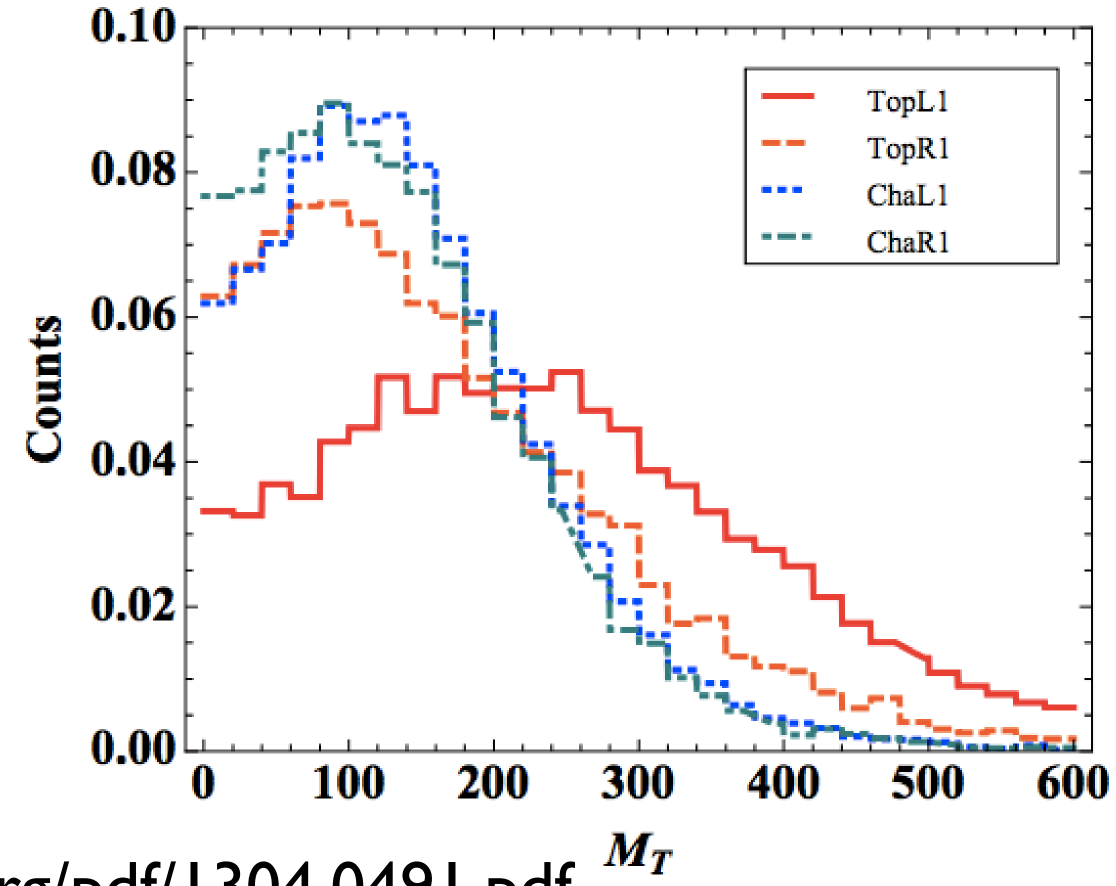
JG, Lykken, Matchev, Mrenna, and Park: JHEP 1410 (2014) 78

- I’m going to advocate reweighting events for Monte Carlo studies of new physics signals

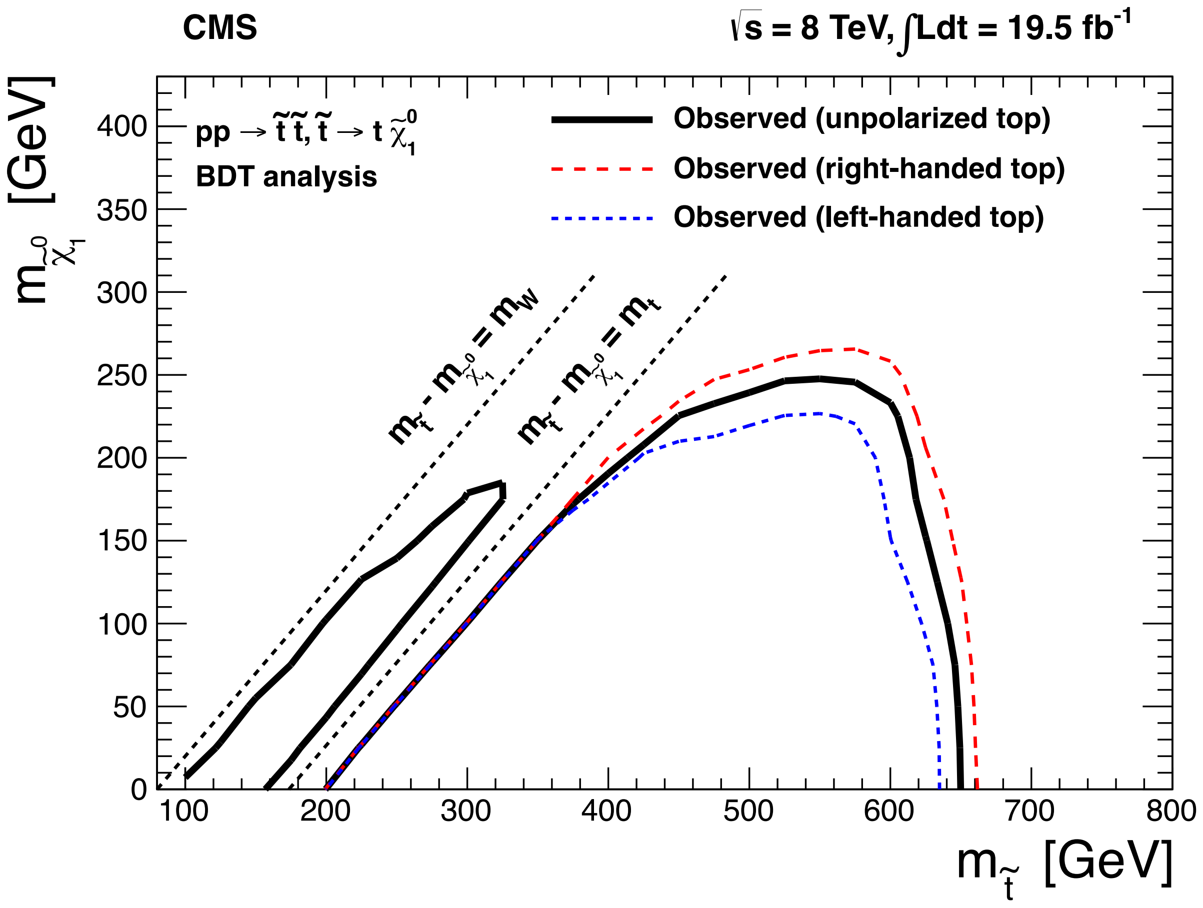
- **Reweighting happens already!!!**

Event Re-weighting

- We decided not to generate any feature of the model (e.g. squark mixing).
- Instead, generate flat ME and then re-weight events for specific cases (e.g. LH/RH top).
- In case of stop production, worked out re-weighing procedure as top polarization was found to have a discernible impact through the lepton spectrum.
- Such re-weighting generally not considered though. Need to find a generic way to do this.



http://arxiv.org/pdf/1304.0491.pdf

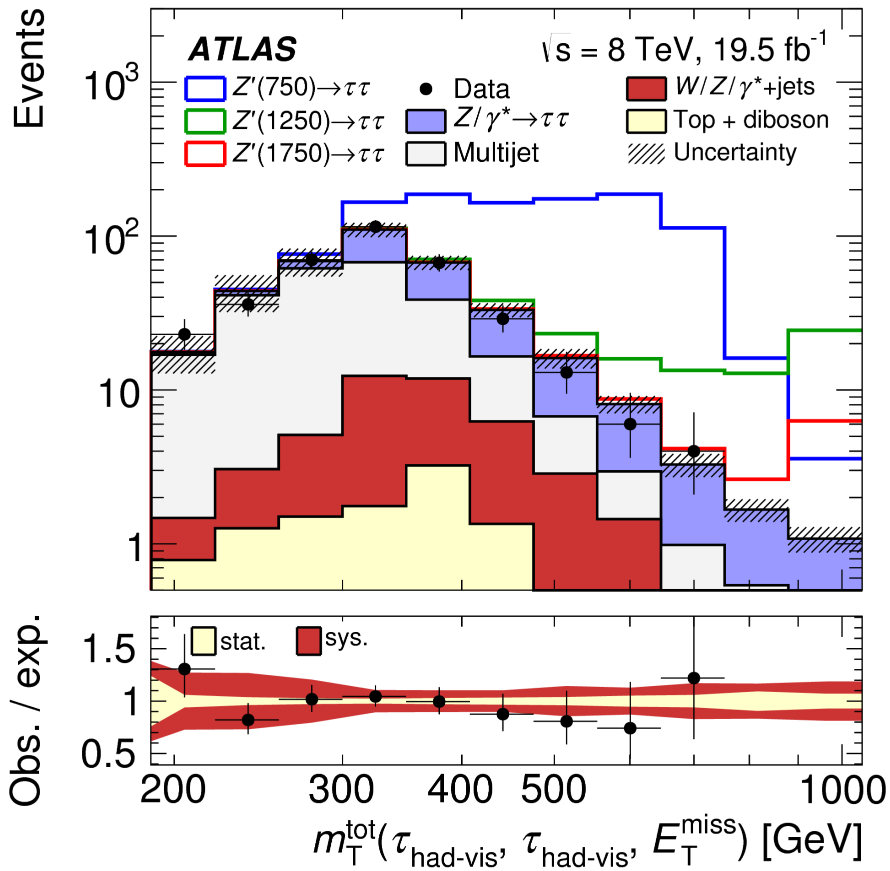


SUS-13-011
arXiv:1308.1586

Ditau Resonances

arXiv: 1502.07177

- **Heavy gauge bosons with non-universal coupling to 3rd gen fermions**
- **Z' → $\tau^+\tau^-$ τ reco with leptonic decays / 1- or 3-prong hadronic decays**



Main event generators / PDF sets	
signal	background
Z': Pythia 8 DY reweighted w/ TauSpinner	**DY**: Pythia 8 / CTEQ6L1 HOC: FEWZ / MSTW2008NNLO
	tt, Wt: MC@NLO+Herwig / CT10
	Diboson: Herwig++ + Herwig / CTEQ6L1

Model / 95% CL lower mass limits	
Z' SSM	2.02 TeV
G(221)	1.3—2.1 TeV

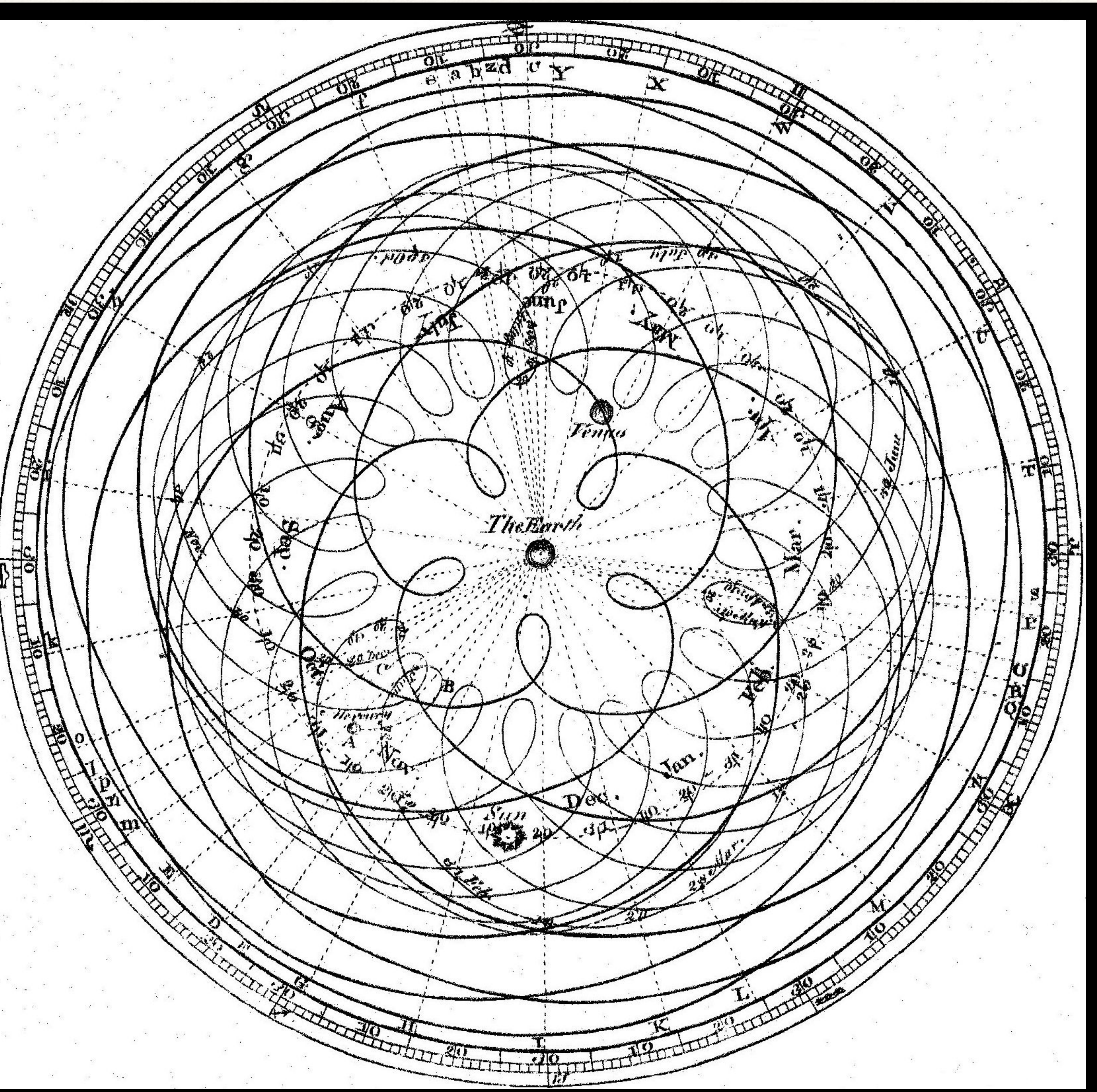
- **Pythia8 DY @ LO reweighted for various signals**
- **TauSpinner correctly accounts for spin effects in tau decays**
- **Higher order corrections (HOC) include QCD, EW effects**

- (See also the MadGraph talks on Monday)

- **My main goals are**

  1. To argue that studying the multi-parameter theory spaces may be more feasible than we think.
  2. To get folks thinking about how to study these theory spaces in better ways (tools?)

New physics searches are using
and should use
models with many parameters.



epicycles?

No.

- **MSSM**: 105 parameters beyond those of the SM.
(Dimopoulos and Sutter 1995, Martin 2000)
- Historically, studied **mSUGRA** instead to handle the number of parameters
- Now: **pMSSM**, **general gauge mediation**, etc. involve more parameters than mSUGRA
- Push toward **NMSSM** to explain heavy Higgs masses
- "**General**" NMSSM (no Z_3 symmetry imposed) adds still more parameters but may have useful features
(Cahill-Rowley, JSG, Hewett, Rizzo 2014)

Not Just a SUSY Problem

- Even in effective field theory approaches, one can have a lot of parameters
- Example: (Goodman, Ibe, Rajaraman, Shepherd, Tait, and Yu, 2010) considers 24 operators that couple WIMPS to SM particles
  - 14 for Dirac fermion WIMPs
  - 6 for complex scalars
  - 4 for real scalars

Not Just a SUSY Problem

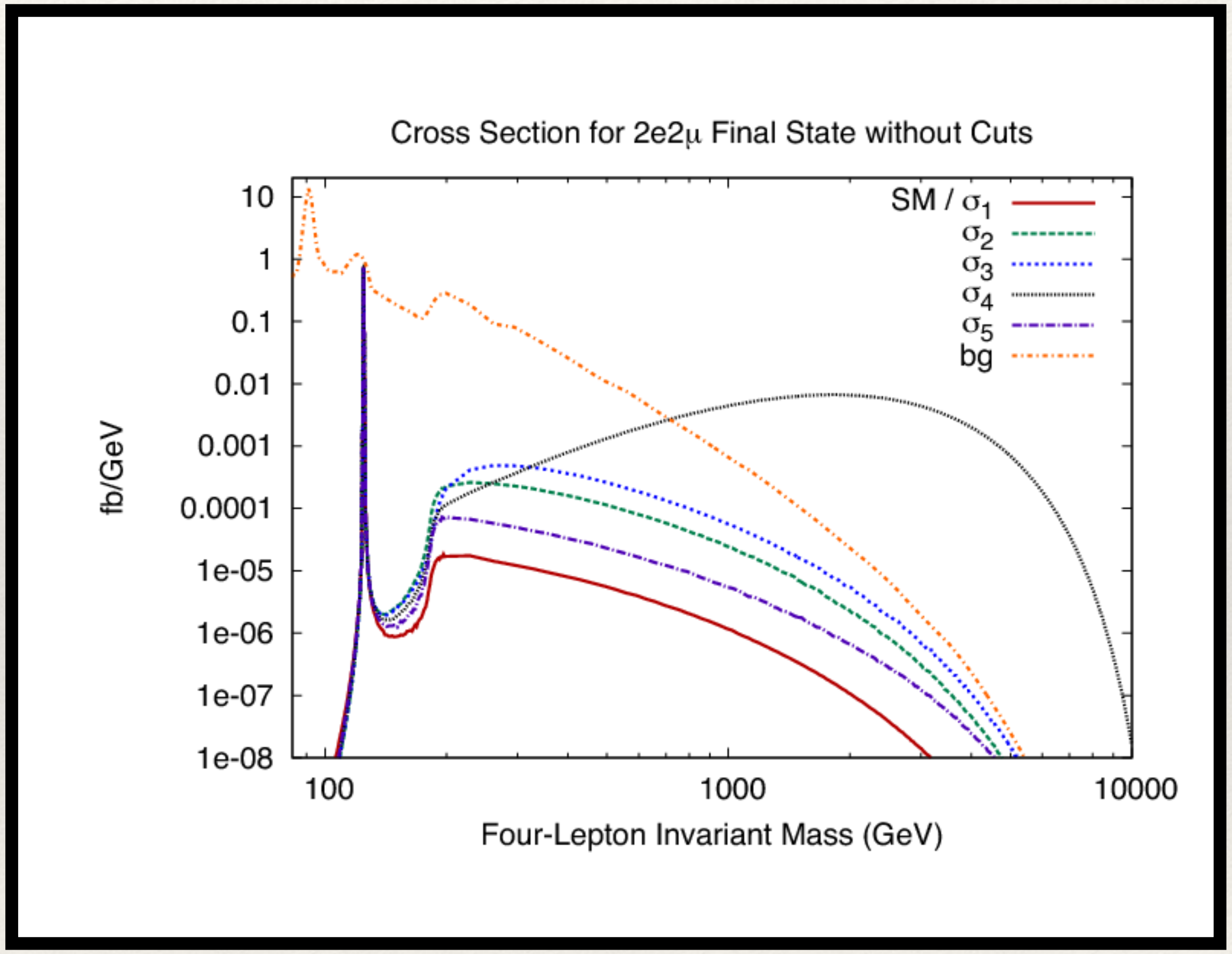
- The specific situation that got us thinking about this was X (H) → 4ℓ

$$\mathcal{L} \supset -\kappa_{1,ZZ}\frac{M_Z^2}{v}XZ_\mu Z^\mu - \frac{\kappa_{2,ZZ}}{2v}XZ_{\mu\nu}Z^{\mu\nu} - \frac{\kappa_{3,ZZ}}{2v}XZ_{\mu\nu}\tilde{Z}^{\mu\nu}$$
$$+\frac{\kappa_{4,ZZ}M_Z^2}{M_X^2 v}\Box XZ_\mu Z^\mu + \frac{2\kappa_{5,ZZ}}{v}XZ_\mu \Box Z^\mu - \frac{\kappa_{2,Z\gamma}}{2v}XZ_{\mu\nu}F^{\mu\nu}$$
$$-\frac{\kappa_{3,Z\gamma}}{2v}XZ_{\mu\nu}\tilde{F}^{\mu\nu} + \frac{\kappa_{4,Z\gamma}M_Z^2}{M_X^2 v}XZ_\mu \partial_\nu F^{\mu\nu} - \frac{\kappa_{2,\gamma\gamma}}{2v}XZ_{\mu\nu}F^{\mu\nu}$$
$$-\frac{\kappa_{3,\gamma\gamma}}{2v}XZ_{\mu\nu}\tilde{F}^{\mu\nu} - \kappa_{1,ZZ'}\frac{M_Z^2}{v}XZ_\mu Z'^\mu$$

Many couplings/ parameters to describe potential
X → $V^{(*)}V^{(*)}$ → 4ℓ !!!

Not Just a SUSY Problem

- **These parameters matter!**
- Often-neglected HZZ operators can dramatically change off-shell cross section (JG, Lykken, Matchev, Mrenna, Park, 2014)
- Light Z′ (Curtin, Essig, Gori, Jaiswal, Katz, Liu, Liu, McKeen, Shelton, Strassler, Surujon , Tweedie, Zhong, 2013) could indicate a hidden valley
- CP-violation in $h\gamma\gamma$ can be probed in four-leptons (Chen, Harnik, and Vega-Morales 2014)
- …



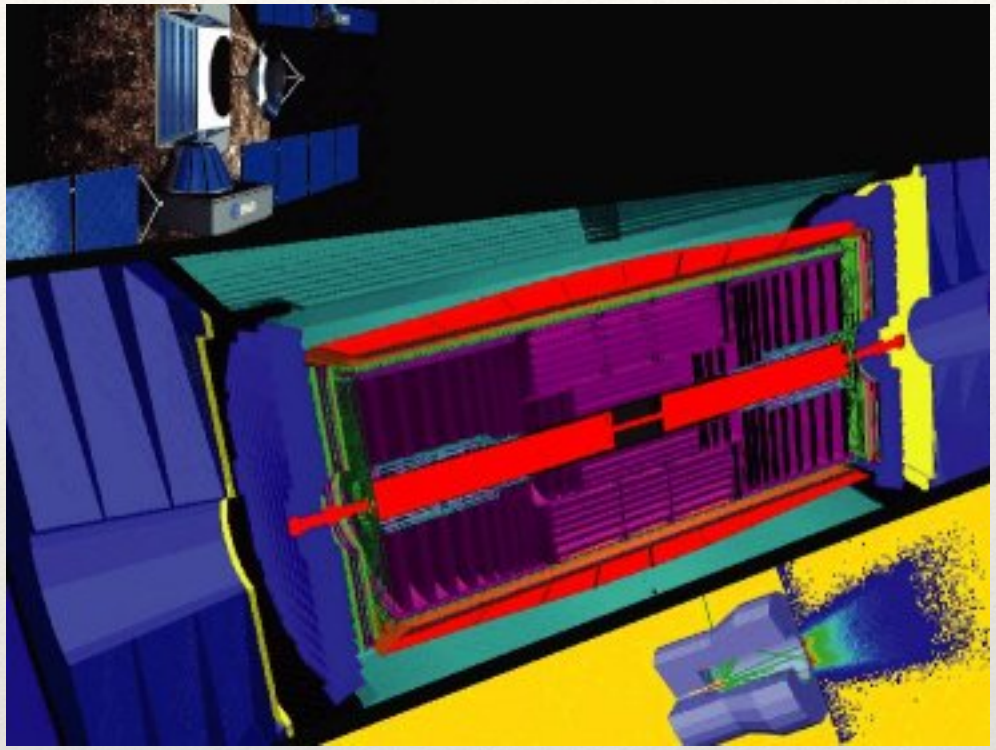
- (JG, Lykken, Matchev, Mrenna, Park, 2014)

The Problem with Parameters

- Need large signal MC samples to evaluate sensitivity
- Here MC = “fullsim” MC: run through high-powered detector simulator (Geant4)
- Time consuming!
    - ~event per minute

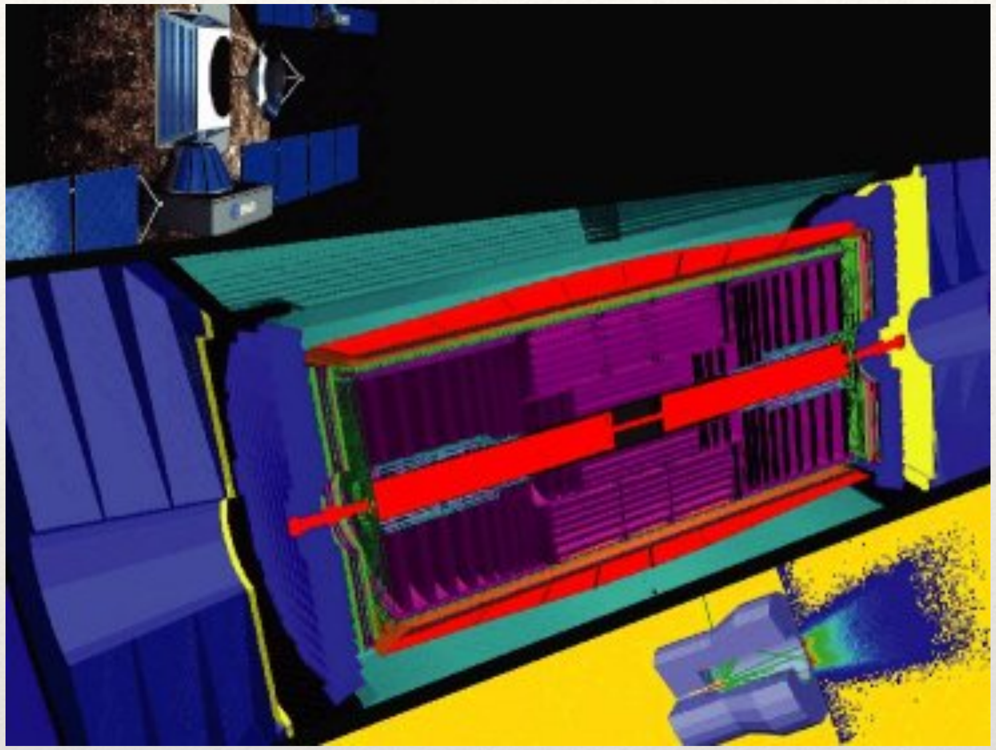
      vs.

    - many events per second



- Makes it challenging to study large parameter spaces in full generality, especially with computationally intensive variables (e.g., MEM)

The Problem with Parameters



- Need large signal MC samples to evaluate sensitivity
- Here MC = “fullsim” MC: run through high-powered detector simulator (Geant4)
- Time consuming!
  - ~event per minute

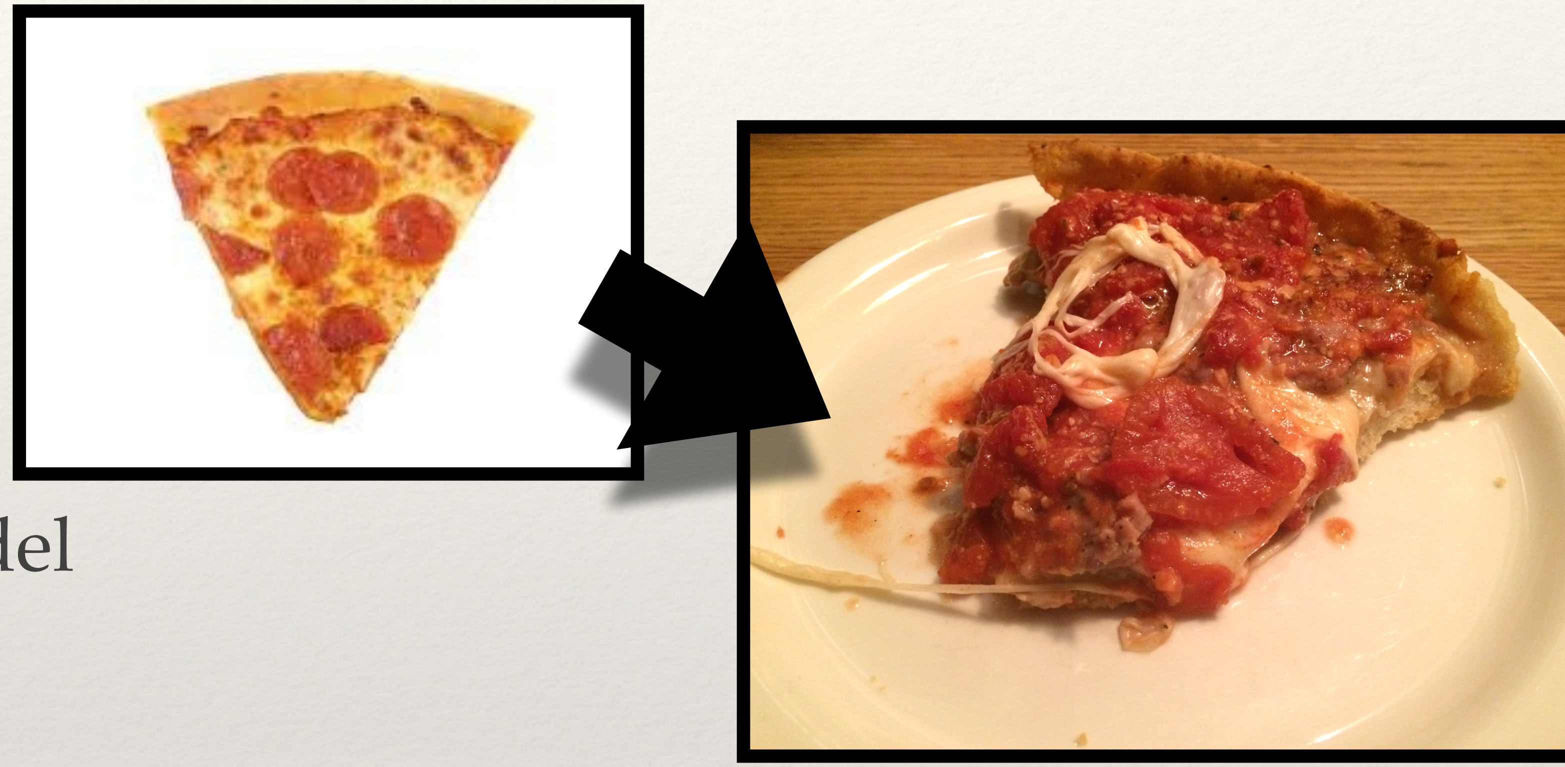
    vs.

  - many events per second
- Makes it challenging to study large parameter spaces in full generality, especially with computationally intensive variables (e.g., MEM)

or does it???

Reweighting

- Our proposed response is to reweight events to cover potentially large theory parameter spaces



- We know how to reweight!
- pdf-reweighting used to model
  - other pdf sets
  - other beam energies

reweighting pizza:
a Chicagoland tradition

- And there's reweighting in other contexts too (see early slides, also e.g. MadSpin, etc.) However, I think this is under-appreciated, especially as a tool for studying multi-parameter theory spaces.

What is (Full) MC Simulation?

Basic Picture

- "true" = hard-process parton-level event, "actual" momenta
- "objects" = jets, electrons, muons, etc.

- A theorist's view: $$\boldsymbol{x}_{i,\text{objects}} = T(\boldsymbol{x}_{i,\text{objects}}, \boldsymbol{x}_{i,\text{true}})\boldsymbol{x}_{i,\text{true}}$$

What is (Full) MC Simulation?

- **T** is a (generalized) transfer function mapping our easy-to-simulate parton-level events to detector-simulated events
- **T** maps every parton-level level event to a detector-level event
- **T** is *almost always* independent of the new physics we are considering
  - One exception: looking for quirks in the underlying event (Harnik and Wizansky 2008)

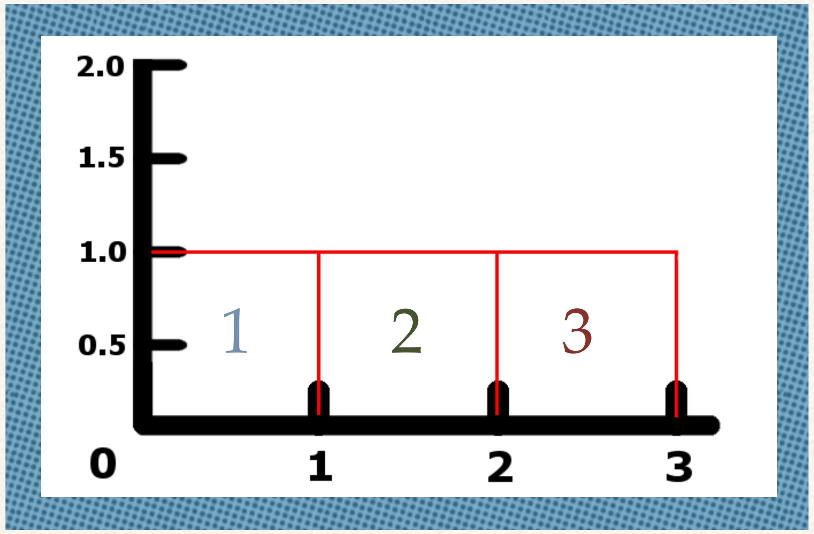
How to Reweight

- We're interested in distributions (p_T, E_T, H_T, M_{T2}, M_2, MEM/ BDT/ NN discriminant)
- Histogram for a given variable in model (**B**, **β**)*
  can be found from events generated for model (**A**, **α**) by weighting each event by R:

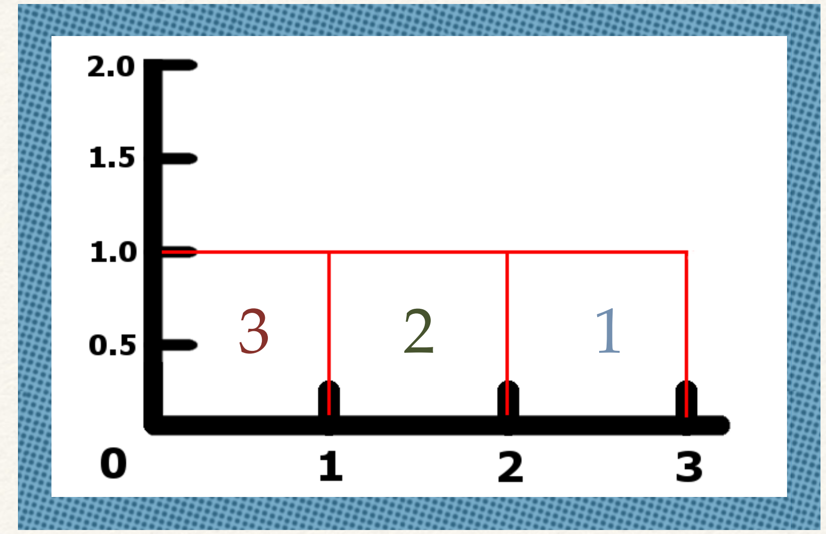
$$R(A, \boldsymbol{\alpha}, B, \boldsymbol{\beta}) \equiv \left(\frac{d\sigma(B, \boldsymbol{\beta}; \boldsymbol{x}_{\text{true}})}{d\boldsymbol{x}_{\text{true}}} \Big/ \frac{d\sigma(A, \boldsymbol{\alpha}; \boldsymbol{x}_{\text{true}})}{d\boldsymbol{x}_{\text{true}}} \right)$$

- This is for MC: we can use truth values.
  No expensive integration over invisible particles/ transfer functions, etc.
- New physics independence of T gives us this simple form for R

*my weird notation for discrete and continuous parameters x in model A



y in model A



Unweighted events events generated for model A

1

parton level: (0.43, 2.55)
detector level: (0.50, 2.50)
$d^2\sigma/dxdy$ (model A): 0.9
$d^2\sigma/dxdy$ (model B): 0.9

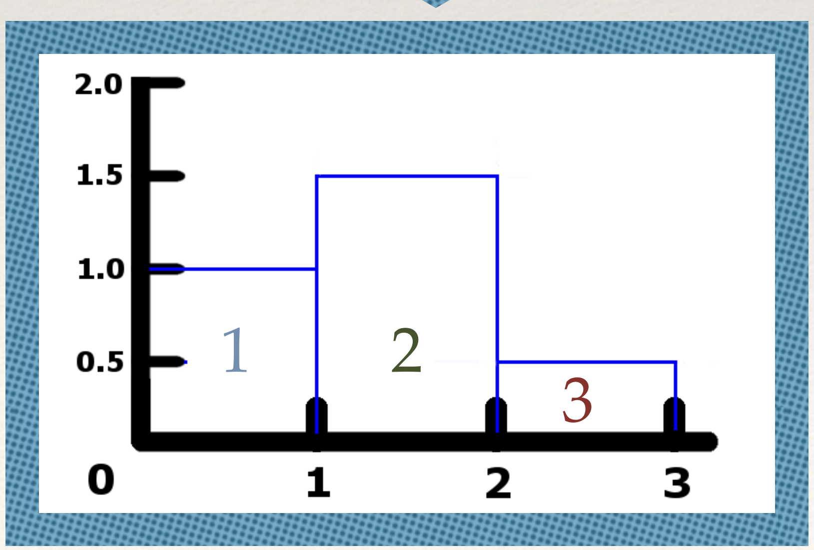
2

parton level: (1.47, 1.41)
detector level: (1.50, 1.50)
$d^2\sigma/dxdy$ (model A): 1.2
$d^2\sigma/dxdy$ (model B): 1.8

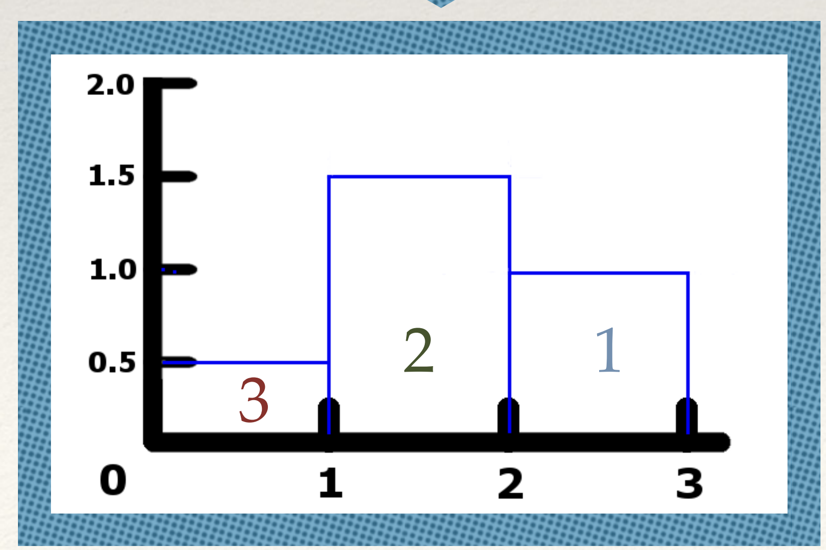
3

parton level: (2.53, 0.55)
detector level: (2.50, 0.50)
$d^2\sigma/dxdy$ (model A): 1.1
$d^2\sigma/dxdy$ (model B): 0.55

x in model B

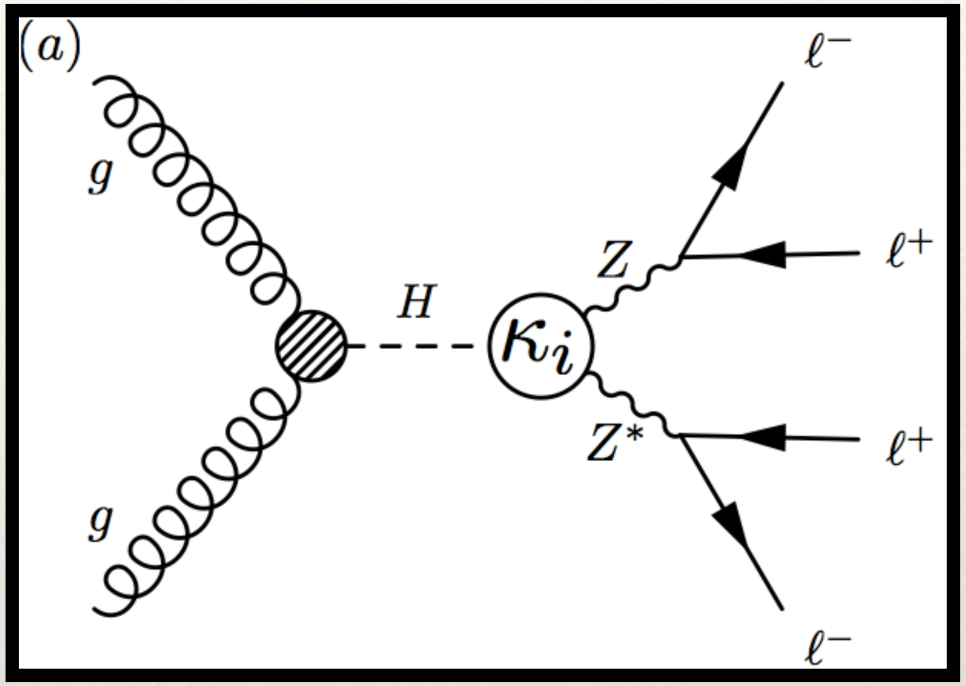


y in model B



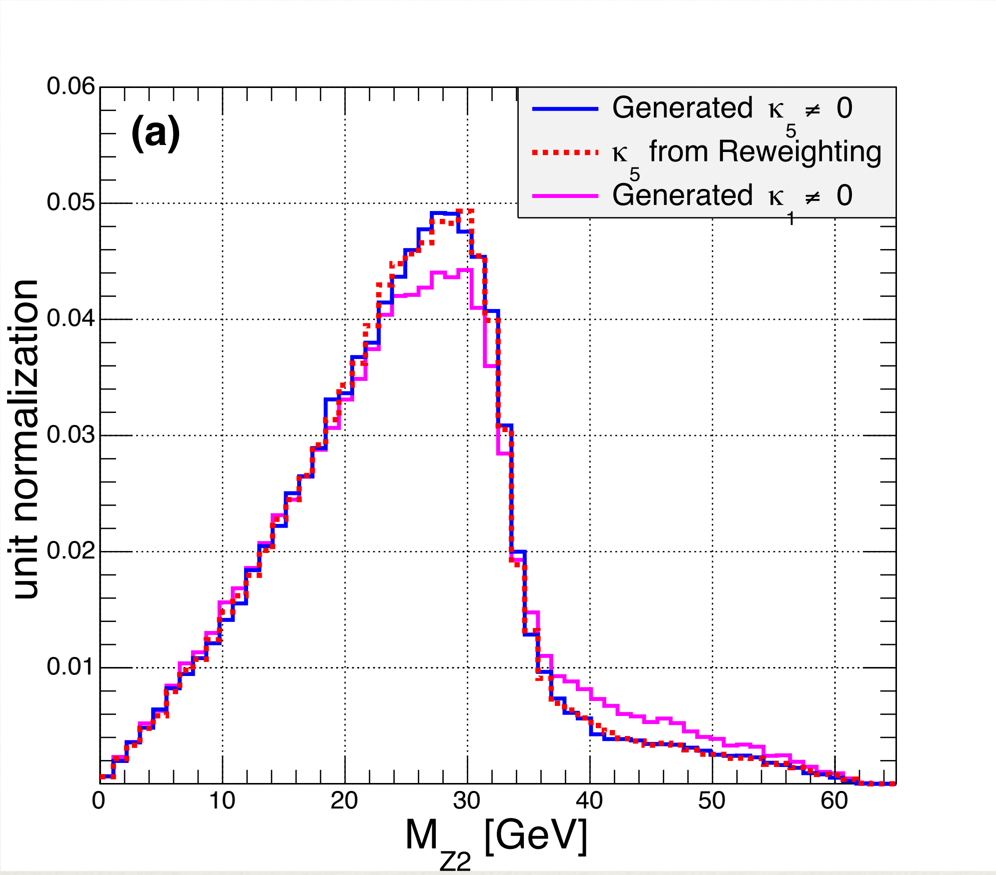
Example: Higgs to Four Leptons

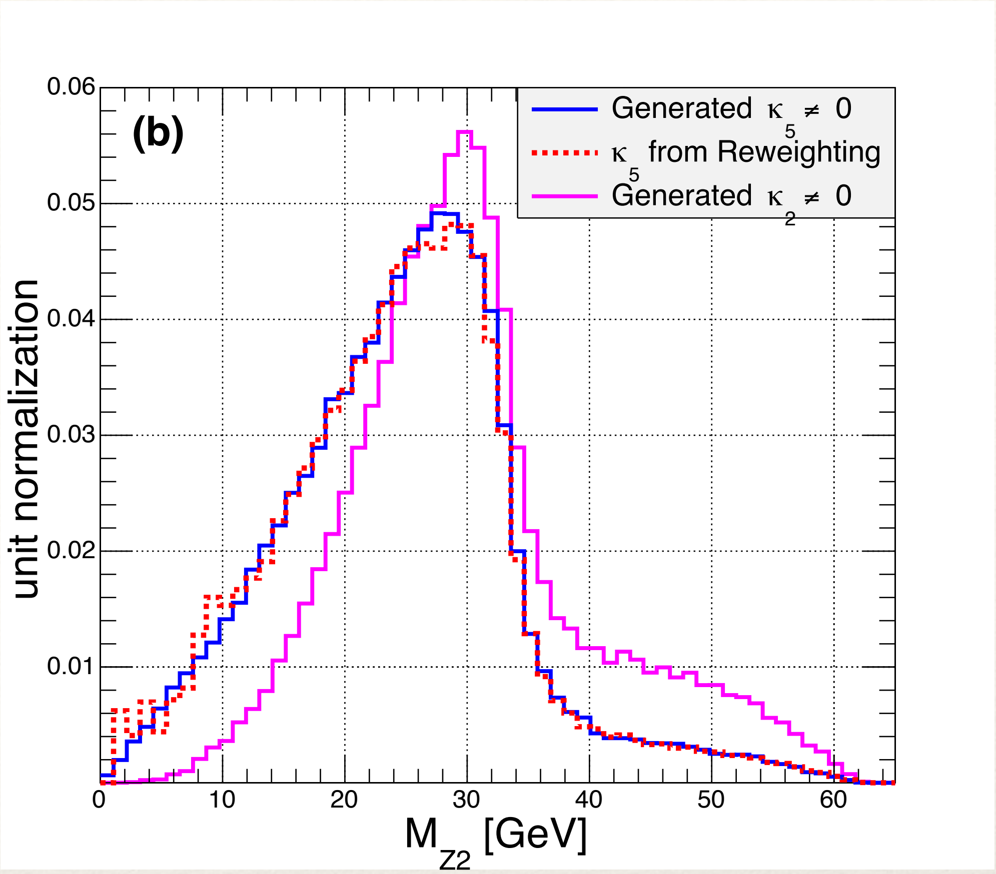
- We start with four-lepton events generated using
pure κ_1 (tree-level SM),
pure κ_2 (CP-even, higher dimensional)
and κ_3 (CP-odd) couplings
- Reweight to obtain distributions for κ_5



$$\mathcal{L} \supset \sum_{i=1}^{5} \kappa_i \mathcal{O}_i = -\kappa_1 \frac{M_Z^2}{v} H Z_\mu Z^\mu - \frac{\kappa_2}{2v} H F_{\mu\nu} F^{\mu\nu} - \frac{\kappa_3}{2v} H F_{\mu\nu} \tilde{F}^{\mu\nu} + \frac{\kappa_4 M_Z^2}{M_X^2 v} \Box H Z_\mu Z^\mu + \frac{2\kappa_5}{v} H Z_\mu \Box Z^\mu.$$

- (JG, Lykken, Matchev, Mrenna, Park, 2013, 2014)

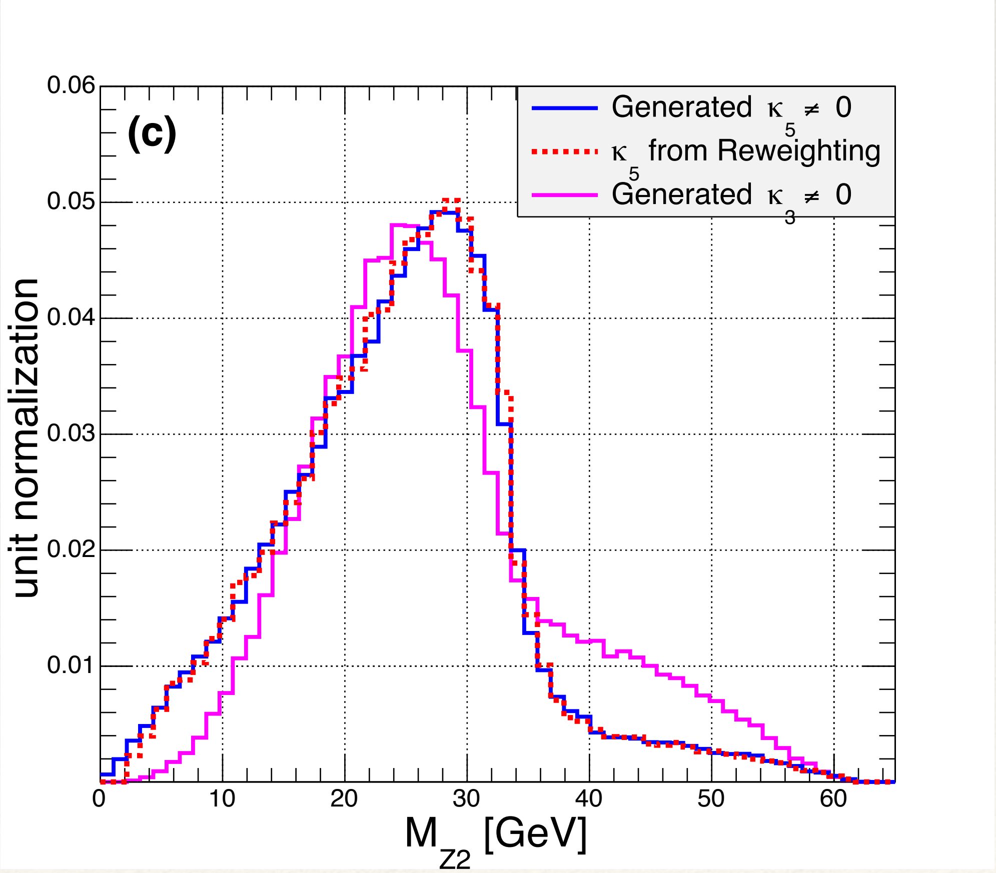




M2 distribution for pure k1, k2, k3 events

M2 distribution for pure k5 events

M2 distribution for pure k5 events obtained by reweighting pure k1, k2, k3 events



Uncertainties

- A particular bin of an weighted histogram has value

$$T = \sum w_k = N\langle w\rangle$$

- And error

$$\delta = \sqrt{\sum w_k^2} = \sqrt{N}\sqrt{\langle w^2\rangle}$$

- Thus

$$\frac{\delta}{T} = \frac{1}{\sqrt{N}}\sqrt{1 + \frac{\sigma_w^2}{\langle w\rangle^2}}$$

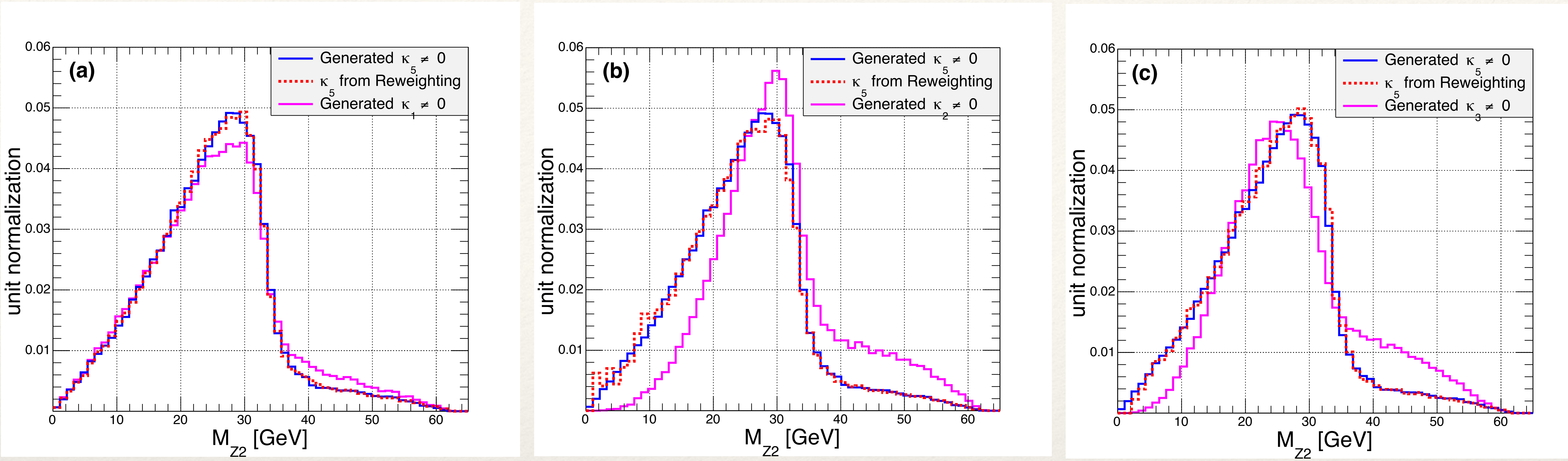
- Spread in weights increases error on bin values

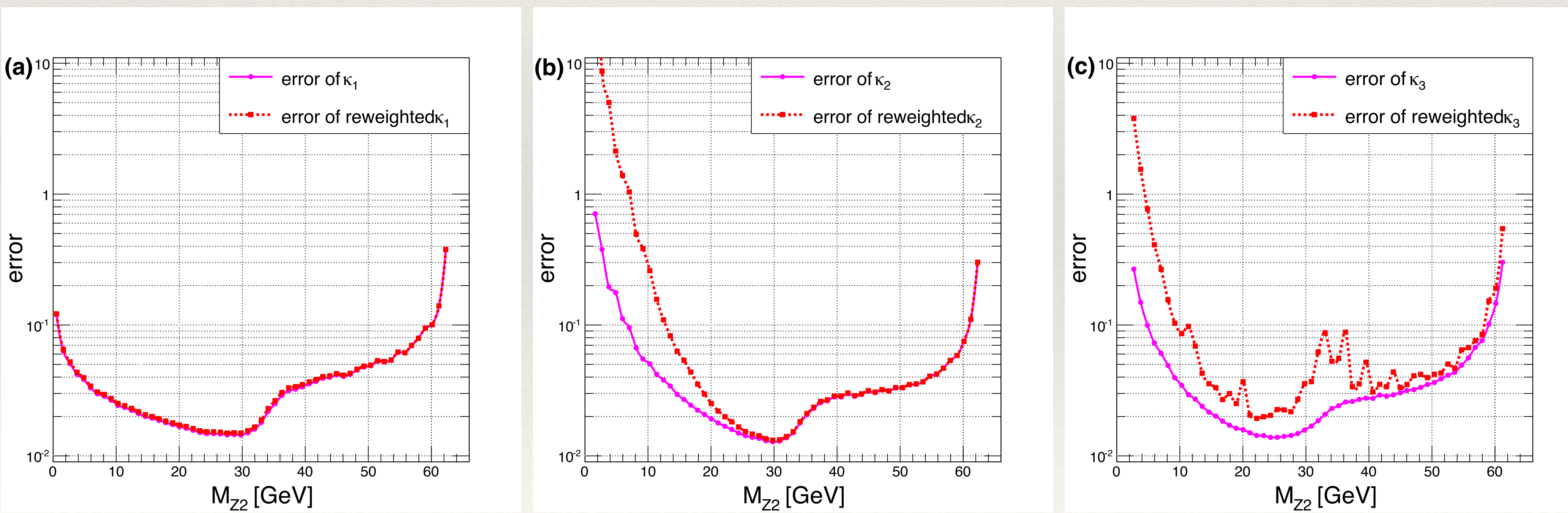
Uncertainties

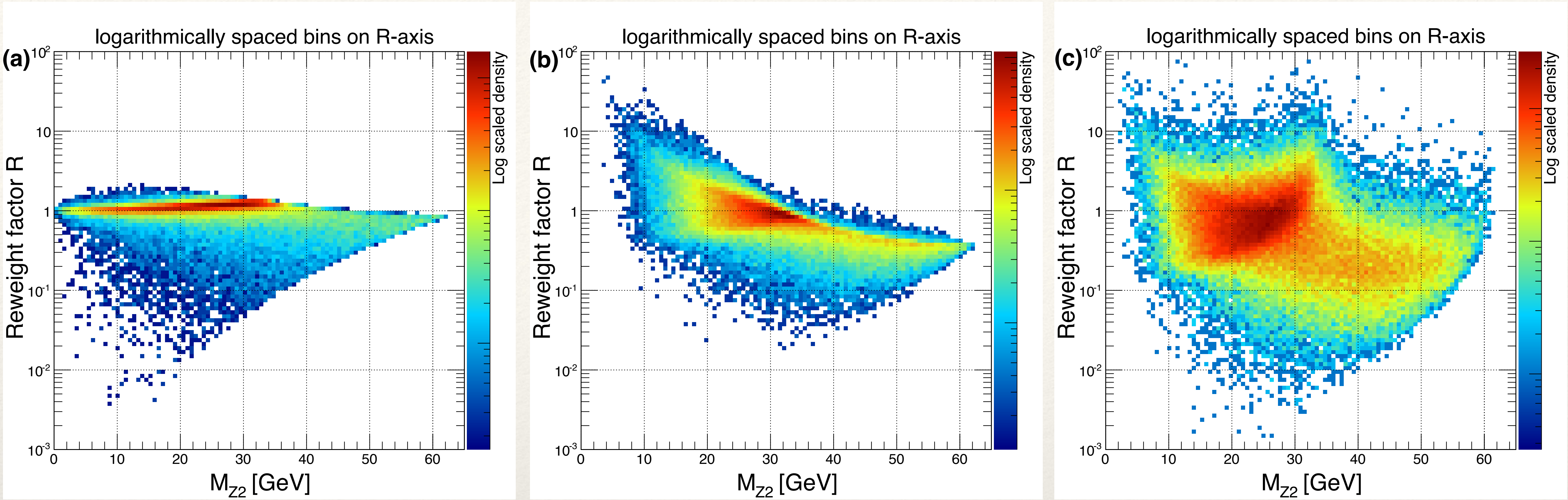
- The error in a weighted histogram for model divided by the error for an unweighted histogram for model is

$$\left(\frac{\delta}{T}\right) / \left(\frac{1}{\sqrt{T}}\right) = \sqrt{\langle w \rangle} \frac{1}{\sqrt{N}} \sqrt{1 + \frac{\sigma_w^2}{\langle w \rangle^2}}$$

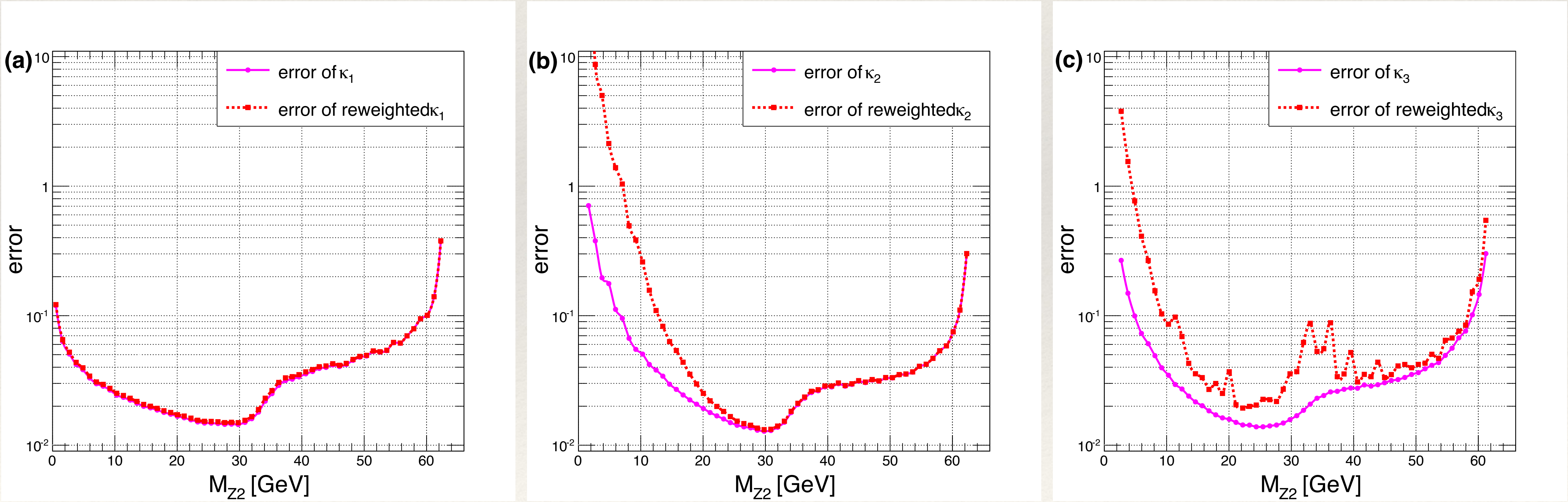
- We can have large uncertainties on bin values in the weighed histogram due to
  - Large weights
  - Large spread of weights


(a)
Generated κ5 ≠ 0
κ5 from Reweighting
Generated κ1 ≠ 0
(b)
Generated κ5 ≠ 0
κ5 from Reweighting
Generated κ2 ≠ 0
(c)
Generated κ5 ≠ 0
κ5 from Reweighting
Generated κ3 ≠ 0
unit normalization
MZ2 [GeV]


(a)
error of κ1
error of reweightedκ1
(b)
error of κ2
error of reweightedκ2
(c)
error of κ3
error of reweightedκ3
error
MZ2 [GeV]



logarithmically spaced bins on R-axis
(a)
(b)
(c)
Reweight factor R
Log scaled density
M_{Z2} [GeV]



(a)
(b)
(c)
error
error of κ_1
error of reweightedκ_1
error of κ_2
error of reweightedκ_2
error of κ_3
error of reweightedκ_3
M_{Z2} [GeV]

Moving Forward…

- Goal is to obtain histograms with low errors (especially in regions of interest) throughout parameter space while running as few events as possible through detector simulation

- We can quantify how well we are doing by looking at the values of and the variance in the weights (R) used to obtain reweighed histograms

- Potential approaches to covering parameter space with low uncertainties:
  - Generate additional events only when uncertainties on bins in particular histograms exceed a certain threshold
  - Generate unweighted events for several benchmark points; reweigh to an arbitrary point from a nearby benchmark
  - Possibly determine some optimal distribution for the unweighted events

Conclusions

- While I like the pun,
  I have mixed feelings about referring to the title of a play where everyone talks about a "Godot" who never shows up.

- Hopefully, at Run 2 we will stop "Waiting for BSM!"

- Using reweighting to study large BSM parameter spaces may play a part in ending the wait!