Data Analysis Using Kinematic Variables

Doojin Kim

UNIVERSITY of FLORIDA

LPC "Data Challenge" Workshop, FNAL, May 21 2015

Introduction

Model

- Already introduced at MC4BSM-6@Cornell 2012 (see also arXiv:1209.0297)
- Four new particles: U, E, ϕ_1, ϕ_2

$$\mathcal{L}_{\text{s.m.}} = -\frac{m_1^2}{2}\phi_1^2 - \frac{m_2^2}{2}\phi_2^2 - m_{12}^2\phi_1\phi_2$$

$$\mathcal{L}_{\text{f.m.}} = M_U\bar{U}U + M_E\bar{E}E$$

$$\mathcal{L}_{\text{Yuk}} = \lambda_1\phi_1\bar{U}P_R u + \lambda_2\phi_2\bar{U}P_R u + \lambda'_1\phi_1\bar{E}P_R e + \lambda'_2\phi_2\bar{E}P_R e$$

$M_U = 400$ GeV

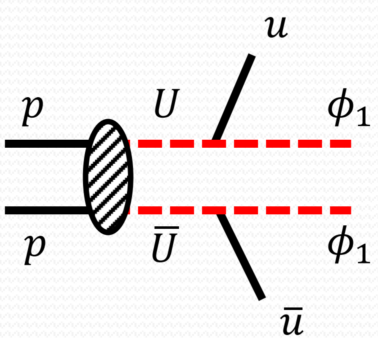
$m_2 = 300$ GeV

$M_E = 250$ GeV

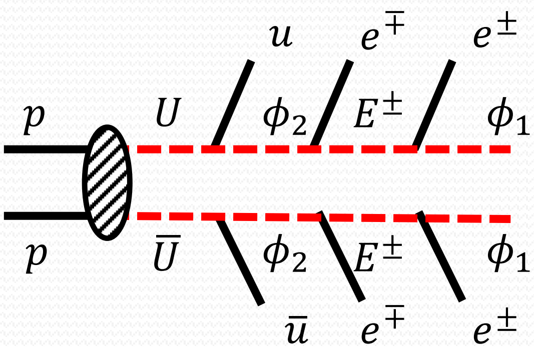
$m_1 = 200$ GeV

Introduction

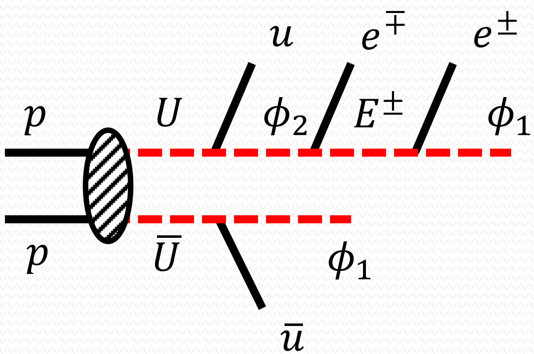
Collider signatures



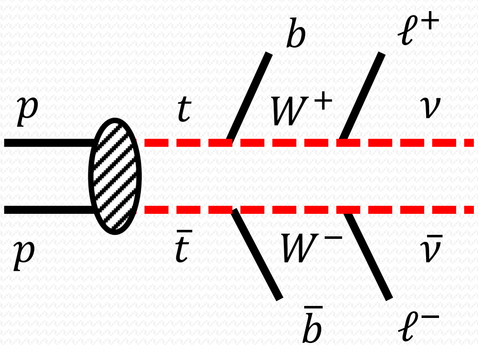
- Tough (e.g., $\tilde{q}$ search)



- Easy (e.g., electroweakino search)



- Challenging

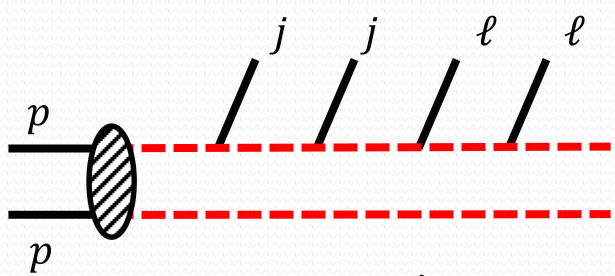


- Major background

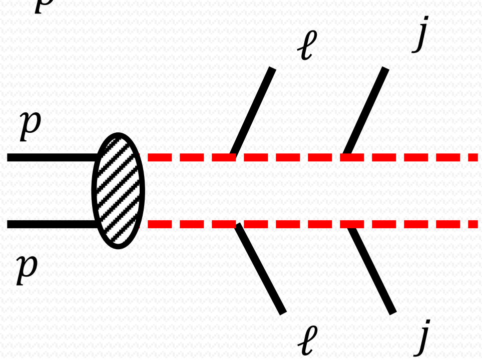
Introduction

Other possibilities

- The chosen model is the "official" toy model for MC4BSM 2012/2015
- There are other possible models to give rise to the same collider signature



- Production of a heavier particle associated with a dark matter particle [Bai and Cheng '10]

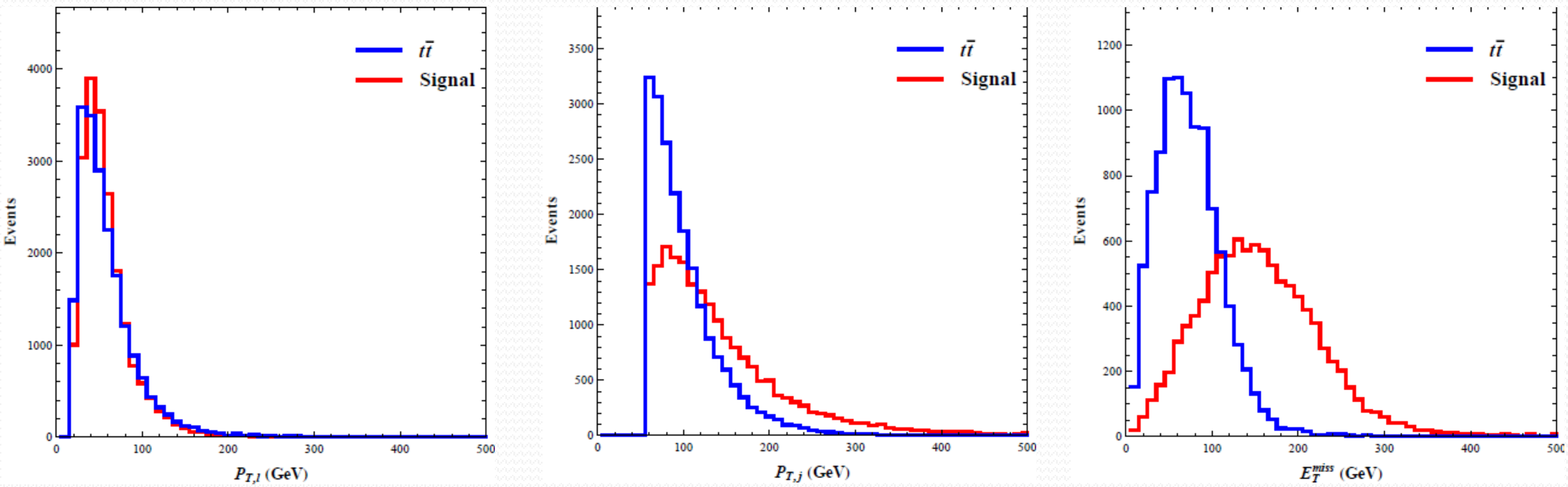


- Jets and leptons are switched (with respect to event topology of $t\bar{t}$) [**Cho, Gainer, DK, Matchev, Moortgat, Pape and Park '14**]

Event Selection

Event selection

- 2 jets + 2 opposite-signed leptons + missing transverse momentum
- Checking standard observables such as p_T^j, E_T^{miss} etc.

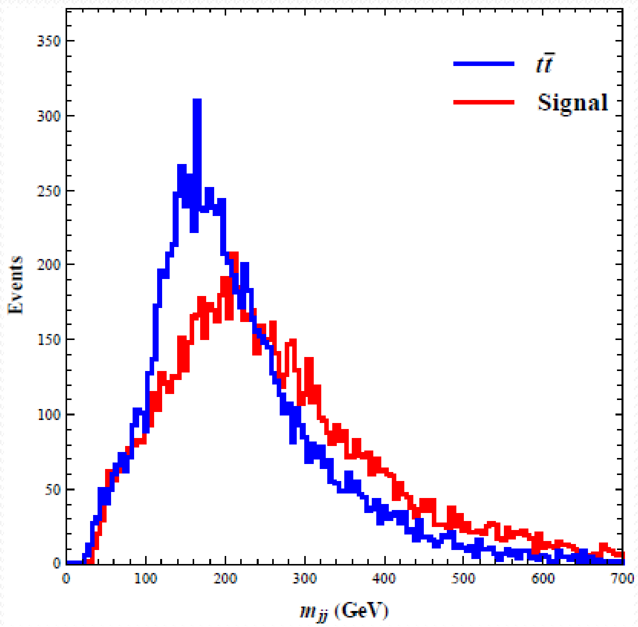


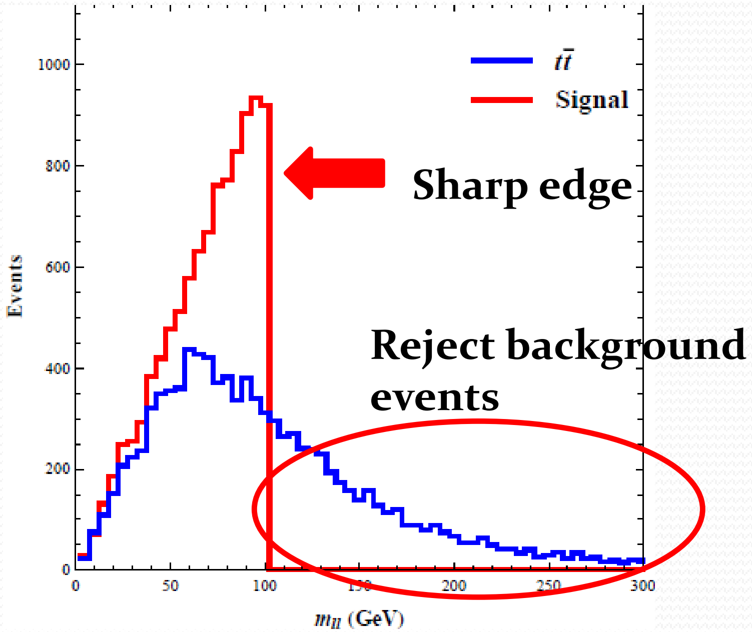
- They are very similar to each other. → Signal distribution can be easily buried in the corresponding background distribution (due to huge production cross section of $t\bar{t}$).

Data Analysis with Kinematic Variables

2-body invariant masses

- Dijet/dilepton invariant masses: NO combinatorial issue





- Signal and background distributions are very similar to each other
- Signal distribution is distinctive

Data Analysis with Kinematic Variables

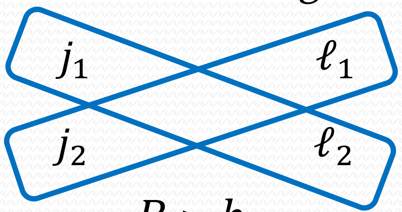
2-body invariant masses

- Invariant masses formed by one jet and one lepton → there is a combinatorial issue so that some prescription is needed

- Partitioning I

$A > a$

- Partitioning II



$B > b$

- Six possible orderings

(1) $\boldsymbol{a \quad b \quad A \quad B}$

(2) $\boldsymbol{a \quad b \quad B \quad A}$

(3) $\boldsymbol{b \quad a \quad A \quad B}$

(4) $\boldsymbol{b \quad a \quad B \quad A}$

(5) $\boldsymbol{a \quad A \quad b \quad B}$

(6) $\boldsymbol{b \quad B \quad a \quad A}$

- Four possible variables

$\mathbf{min}\{\boldsymbol{a}, \boldsymbol{b}\} =$ **Red**

$\mathbf{max}\{\boldsymbol{a}, \boldsymbol{b}\} =$ **Blue**

$\mathbf{min}\{\boldsymbol{A}, \boldsymbol{B}\} =$ **Green**

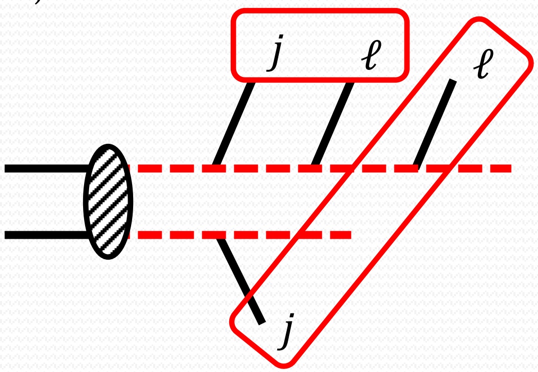
$\mathbf{max}\{\boldsymbol{A}, \boldsymbol{B}\} =$ **Black**

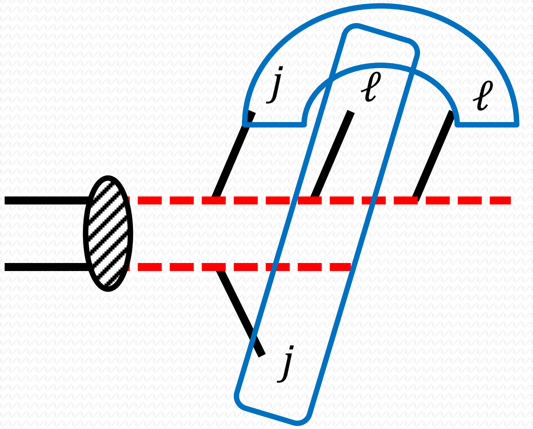
Data Analysis with Kinematic Variables

2-body invariant masses

- Assuming that for $t\bar{t}$, partitioning I is correct for convenience,

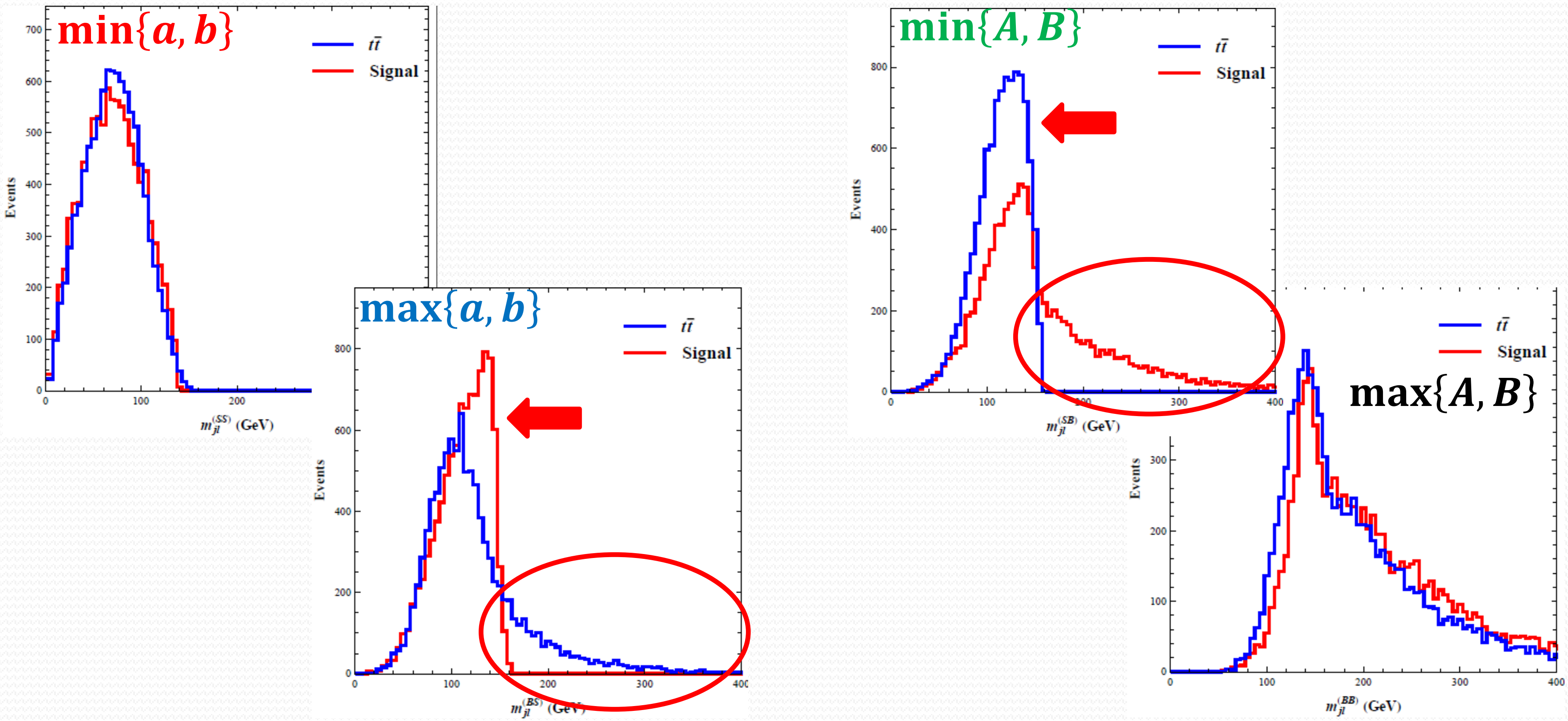
	Signal	Background
$\min\{a, b\}$	E	E
$\max\{a, b\}$	E	T (one out of 6 cases)
$\min\{A, B\}$	T	E
$\max\{A, B\}$	T	T (three out of 6 cases)





Data Analysis with Kinematic Variables

- 2-body invariant masses

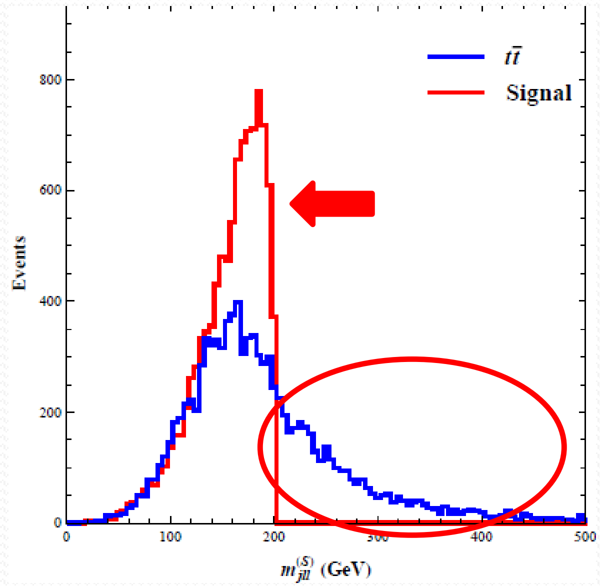


Data Analysis with Kinematic Variables

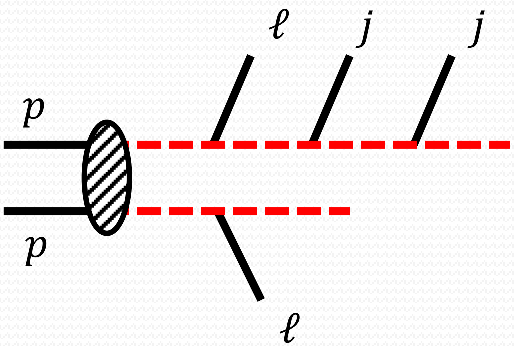
3-body invariant masses

- There exist combinatorial issues

	Signal	Background
Smaller m_{jll}	$\boxed{\text{E}}$	T
Larger m_{jll}	T	T
Smaller m_{jjl}	T	T
Larger m_{jjl}	T	T



- C.f.

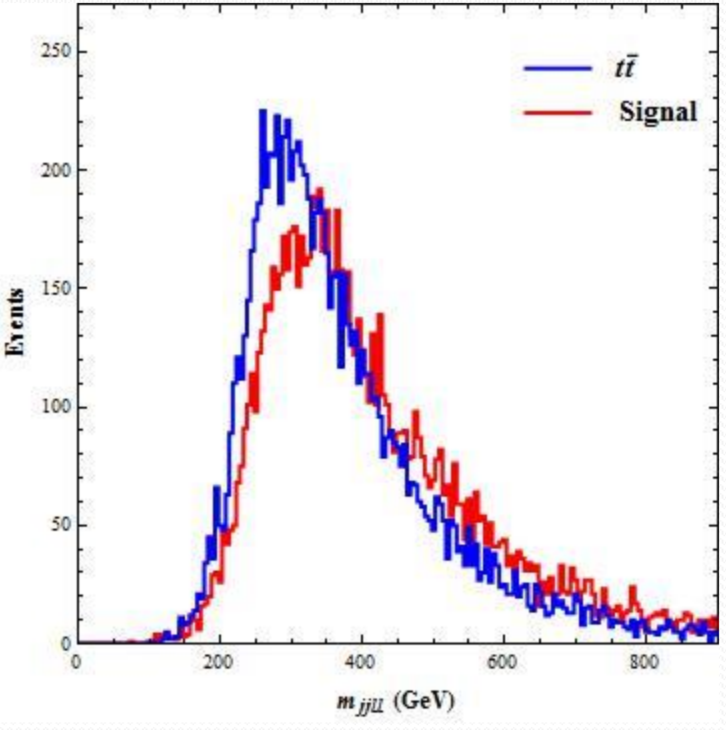


: Signal distribution in smaller m_{jjl} will be bounded above

Data Analysis with Kinematic Variables

4-body invariant masses

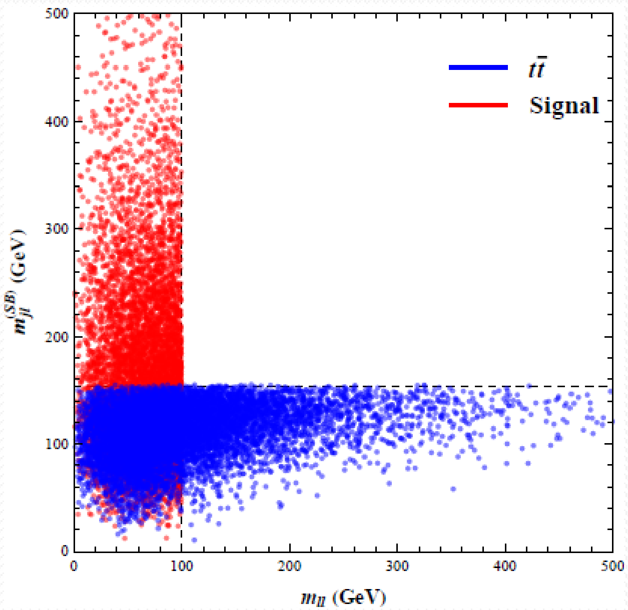
- Both signal and background distributions develop a tail structure so that they are similar to each other, but there is no combinatorial issue



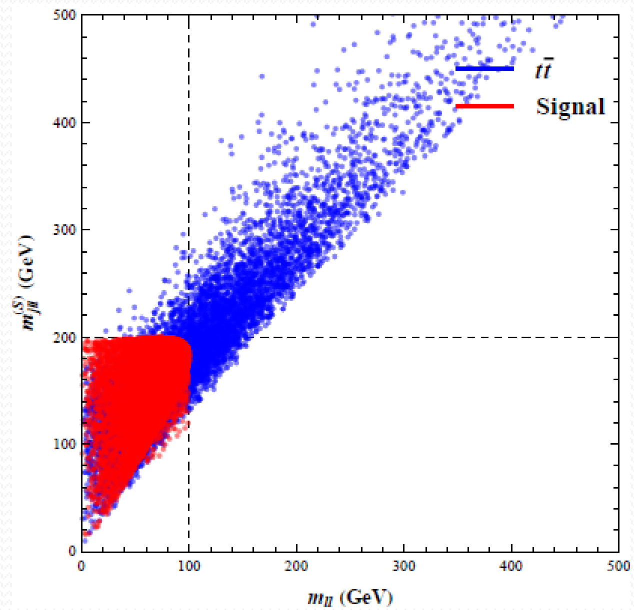
Data Analysis with Kinematic Variables

Scatter plots between invariant masses

- Available is more information which is typically lost when the relevant phase space is projected out to one dimensional space



- m_{ll} vs. $\min\{A, B\}$



- m_{ll} vs. smaller m_{jll}

Data Analysis with Kinematic Variables

Exact(?) solution

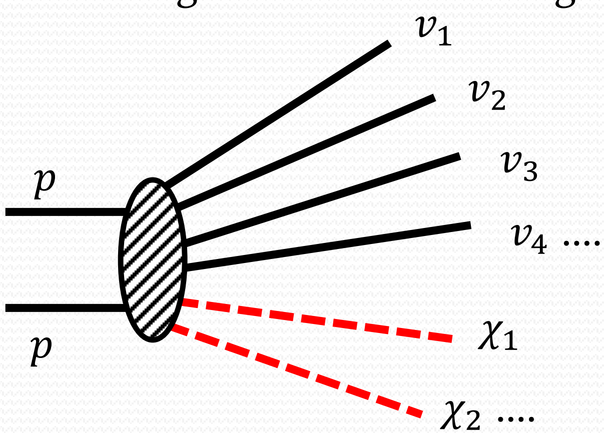
- The existence of invisible particles in the final state makes it hard to reconstruct the full decay process
- However, if we assume $t\bar{t}$, there are 8 unknowns (4 momenta of two neutrinos) and 8 constraints (2 MET conditions + 3 on-shell mass conditions per decay side) → **Exactly solvable**!
  - ✓ Background: several (real/physical) solutions [Sonnenschein '06]
  - ✓ Signal: some cases (if the relevant phase space has overlap with the phase space for $t\bar{t}$

- Practical issues
  - ✓ Resolution: even $t\bar{t}$ might not have solutions
  - ✓ Combinatorics: wrong combination would result in non-physical solutions only

Data Analysis with Kinematic Variables

$\hat{s}_{\min}$

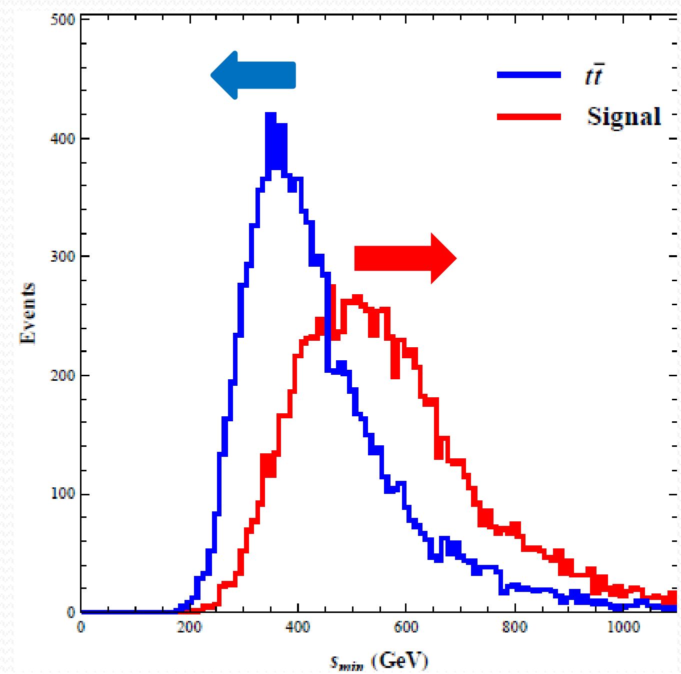
- Estimating the overall scale of given events [Konar, Kong and Matchev '08]



$$\vec{p}_{iT} = \frac{m_i}{M_{inv}} \vec{P}_T^{miss}, \; p_{iz} = \frac{m_i P_z}{\sqrt{E^2 - P_z^2}} \sqrt{1 + \left(\frac{P_T^{miss}}{M_{inv}}\right)^2}$$

$$M_{inv} = \sum_{i=1}^{n_{inv}} m_i$$

$$\hat{s}_{min} = \left(E + \sum_{i=1}^{n_{inv}} \sqrt{m_i^2 + \vec{p}_{iT}^{\,2} + p_{iz}^2}\right)^2 - \left(P_z + \sum_{i=1}^{n_{inv}} p_{iz}\right)^2$$



Data Analysis with Kinematic Variables

M_{T2} or M_2 variables

- Instead of exact solution, we make some model assumptions and find an ansatz, e.g., M_{T2} or M_2 variables
- M_{T2}: a generalization of transverse mass to the case where mother particles are pair-produced and each of them decays into an invisible particle along with visible state [Lester and Summers '99]

$$M_{T2}(\tilde{m}) \equiv \min_{\vec{q}_{1T},\vec{q}_{2T}} \left\{\max\left[M_{TP_1}(\vec{q}_{1T},\tilde{m}),\ M_{TP_2}(\vec{q}_{2T},\tilde{m})\right]\right\}$$
$$\vec{q}_{1T} + \vec{q}_{2T} = \not{\vec{P}}_T$$

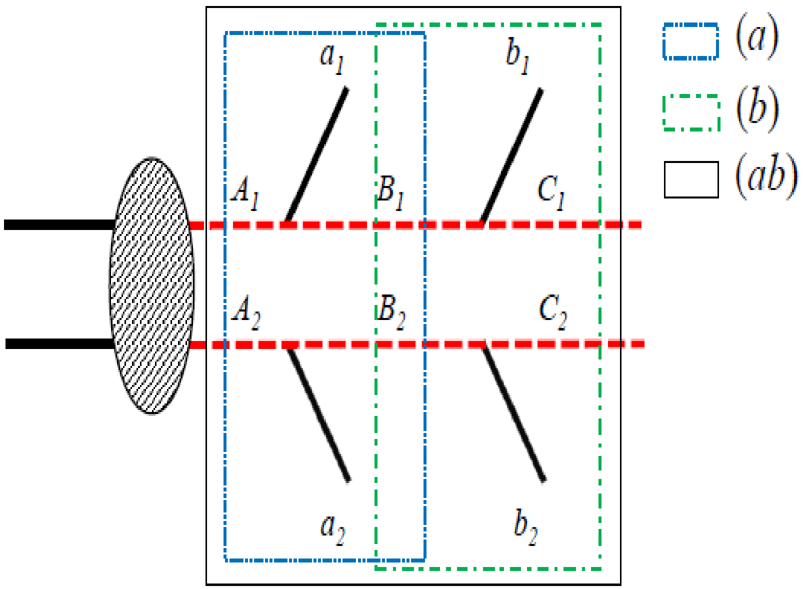
- M_2: (3+1) dimensional analogue of (2+1) dimensional M_{T2} variable [Barr, Khoo, Konar, Kong, Lester, Matchev and Park '11; Mahbubani, Matchev and Park '12; **Cho, Gainer, DK, Matchev, Moortgat, Pape and Park '14**]

$$M_2(\tilde{m}) \equiv \min_{\vec{q}_1,\vec{q}_2} \left\{\max\left[M_{P_1}(\vec{q}_1,\tilde{m}),\ M_{P_2}(\vec{q}_2,\tilde{m})\right]\right\}$$
$$\vec{q}_{1T} + \vec{q}_{2T} = \not{\vec{P}}_T + \text{(optional) on-shell constraints}$$

Data Analysis with Kinematic Variables

Subsystems

- Assuming $t\bar{t}$-like decay topology, we have
  - ✓ Subsystem (ab): minimizing masses of A_i, treating C_i as invisible particles
  - ✓ Subsystem (a): minimizing masses of A_i, treating B_i as invisible particles. Visible particles b_i considered as downstream particles
  - ✓ Subsystem (b): minimizing masses of B_i, treating C_i as invisible particles. Visible particles a_i considered as upstream particles



Data Analysis with Kinematic Variables

Subsystem (ab)

- Introducing the hypothesized mass for C_i as $\tilde{m}$, we have
  - ✓ Missing transverse momentum conditions: $\vec{q}_{1T} + \vec{q}_{2T} = \vec{P}_T^{miss}$
  - ✓ On-shell constraint for A_i: $M(\vec{q}_1; \vec{p}_{a_1}, \vec{p}_{b_1}, \tilde{m}) = M(\vec{q}_2; \vec{p}_{a_2}, \vec{p}_{b_2}, \tilde{m})$
- Then, 3 unknowns are left, and the relevant minimization is performed over those unknowns $\longrightarrow M_{2CX}$
- Equivalence theorem [Barr, Khoo, Konar, Kong, Lester, Matchev and Park '11; Mahbubani, Matchev and Park '12; **Cho, Gainer, DK, Matchev, Moortgat, Pape and Park '14**]
  - ✓ $M_{2CX} = M_{T2}$ event-by-event

Data Analysis with Kinematic Variables

Subsystem (ab)

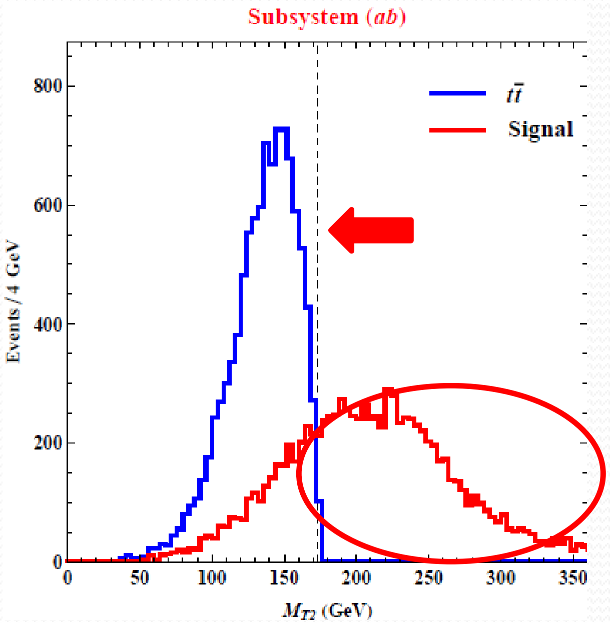
- Analytic formula for M_{T2} is available

$$(M_{T2}^{bal})^2 = \tilde{m}^2 + A_T + \sqrt{\left(1 + \frac{4\tilde{m}^2}{2A_T - m_{1v}^2 - m_{2v}^2}\right)(A_T^2 - m_{1v}^2 m_{2v}^2)}$$

$$M_{T2}^{unbal} = \tilde{m} + m_{iv} \ (i = 1,2)$$

where $A_T = E_{1T}E_{2T} + \vec{p}_{1T} \cdot \vec{p}_{2T}$

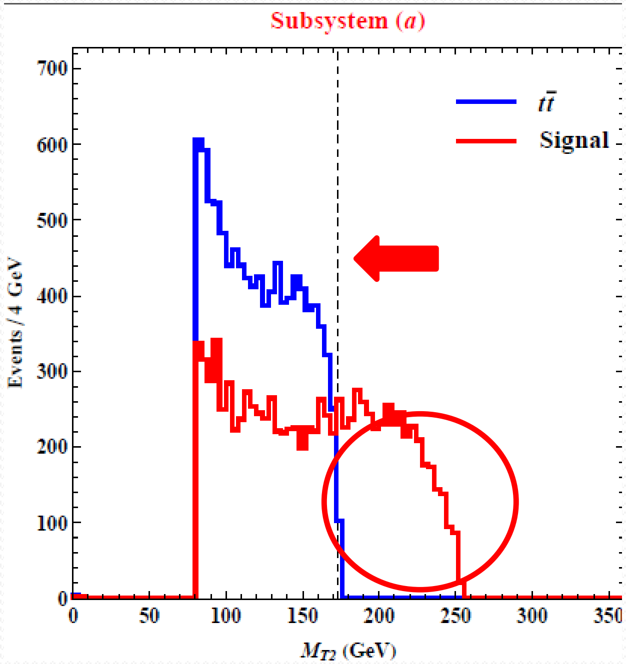
- The hypothesized mass is typically assumed to be zero as per neutrino mass



Data Analysis with Kinematic Variables

Subsystem (a)

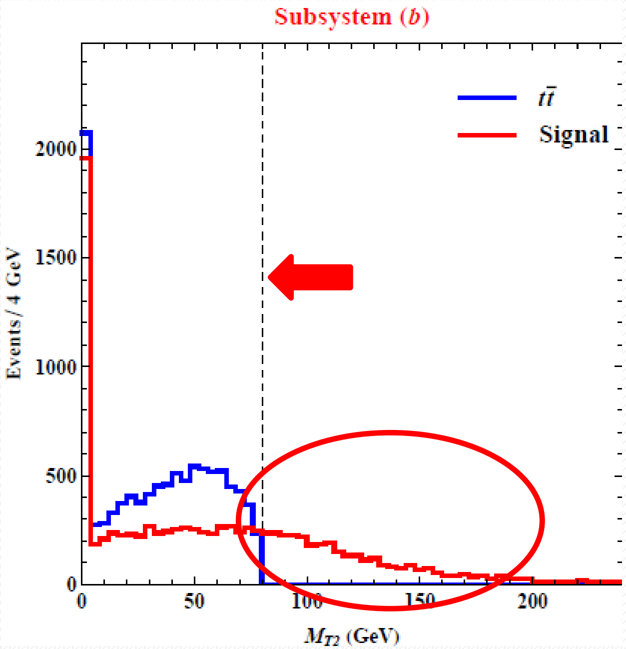
- Analytic formula for M_{T2} in the (ab) subsystem is reusable accordingly.
- The hypothesized mass is typically assumed to be 80 GeV as per W gauge boson mass



Data Analysis with Kinematic Variables

Subsystem (b)

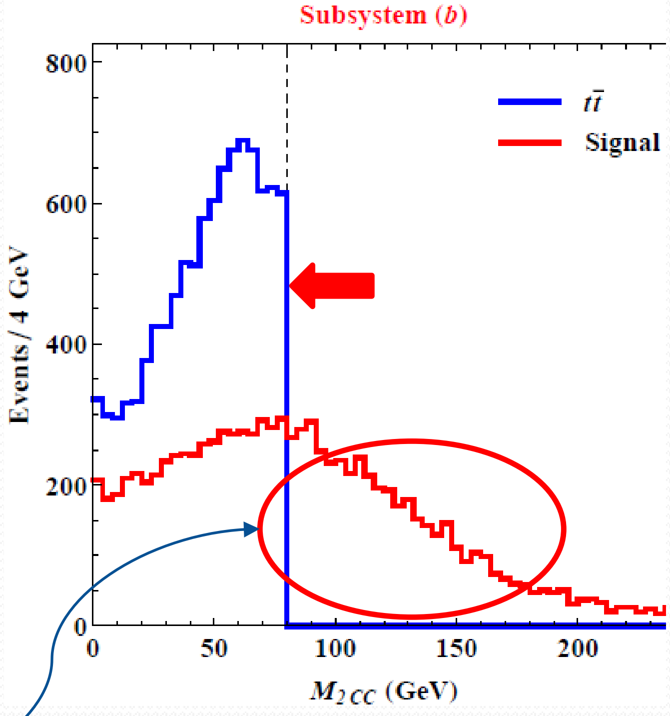
- In the presence of upstream momentum, analytic formula for M_{T2} is NOT available unlike the other two subsystems → numerical minimization is needed (See my talk at MC4BSM 2015 as well, e.g., *http://www.phys.ufl.edu/~cho/Optimized_Mass/OptM_introduction.html*)
- The hypothesized mass is typically assumed to be 0 GeV as per neutrino mass



Data Analysis with Kinematic Variables

More constrained M_2 variables

- For example, in subsystem (ab), one more on-shell constraint for B_i is possible: $M(\vec{q}_1; \vec{p}_{b_1}, \tilde{m}) = M(\vec{q}_2; \vec{p}_{b_2}, \tilde{m})$
- Then, 2 unknowns are left, and the relevant minimization is performed over those unknowns → M_{2CC}
- Numerical minimization is needed (e.g., *http://www.phys.ufl.edu/~cho/Optimized_Mass/OptM_introduction.html*)



More signal events migrate beyond the kinematic endpoint for $t\bar{t}$

Advertisement

Library of Optimized Masses $\bar{M}$ for Event Topologies

Code of the Optimized Mass, $\bar{M}$

http://www.phys.ufl.edu/~cho/Optimized_Mass/OptM_introduction.html

be found in "**Description of the $\bar{M}$**". More detailed class reference and member function/variables for each optimized mass can be found in "**Dictionary of $\bar{M}$**".

Contact

Links

The constrained minimization process for the optimization of the $\bar{M}$ is based on the algorithm of the augmented Lagrangian method, realized with the power of the unconstrained minimization algorithms of **MINUIT** for the case of the M_2 variable. In particular, the constrained-M_2 variable, which is a good example of the optimized mass, had been surveyed in the series of papers:

Guide to transverse projections and mass-constraining variables A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, M. Park

On-shell constrained M_2 variables with applications to mass measurements and topology disambiguation, W. S. Cho, J. S. Gainer, D. Kim, K. T. Matchev, F. Moortgat, L. Pape and M. Park, arXiv:1401.1449 [hep-ph].

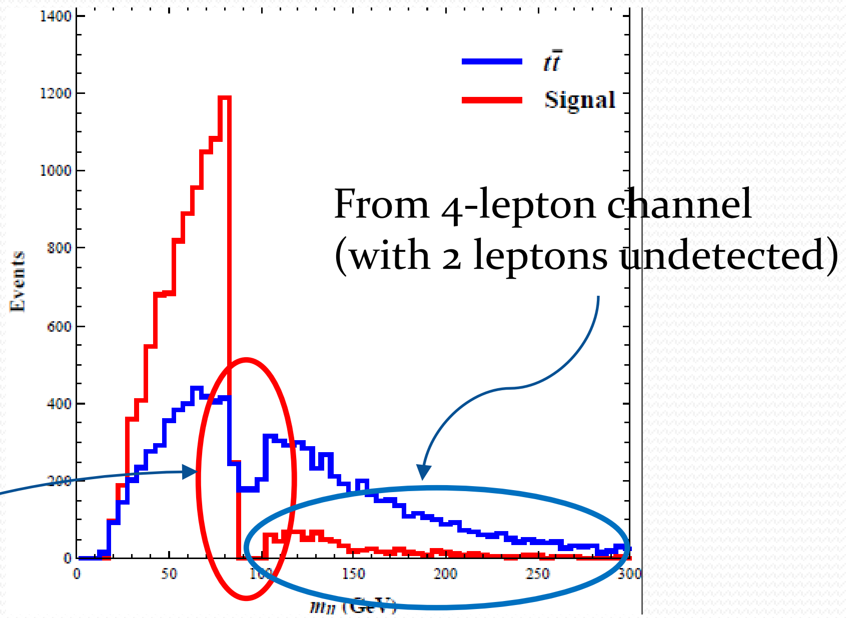
As optimized in a topology-by-topology basis, our code provides the routine for various optimized masses, especially for the well-known background/signal events of the Standard Model/Beyond the Standard Model. The list of the embodied $\bar{M}$ can be found in the "**Description of**

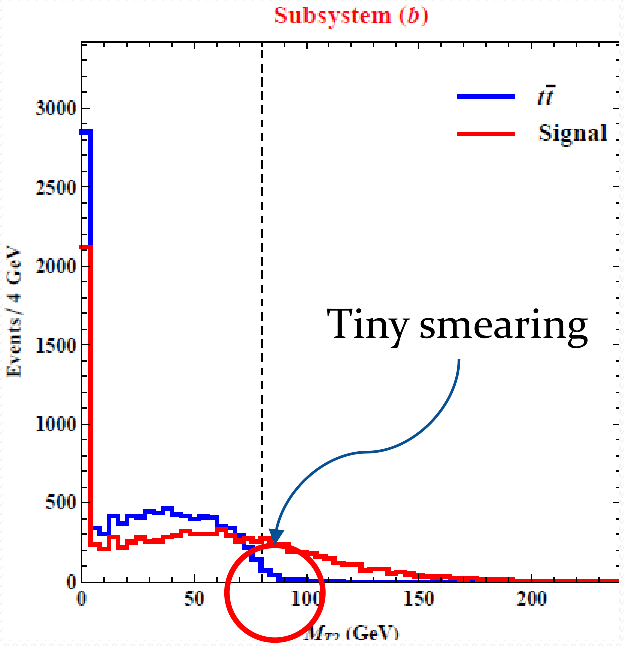
Thank you!

Results with Detector-level Events

Working observables as in parton-level case

- Leptons are clean even at the detector level → Expected to have very similar results, e.g., m_{ll} and leptonic M_{T2}



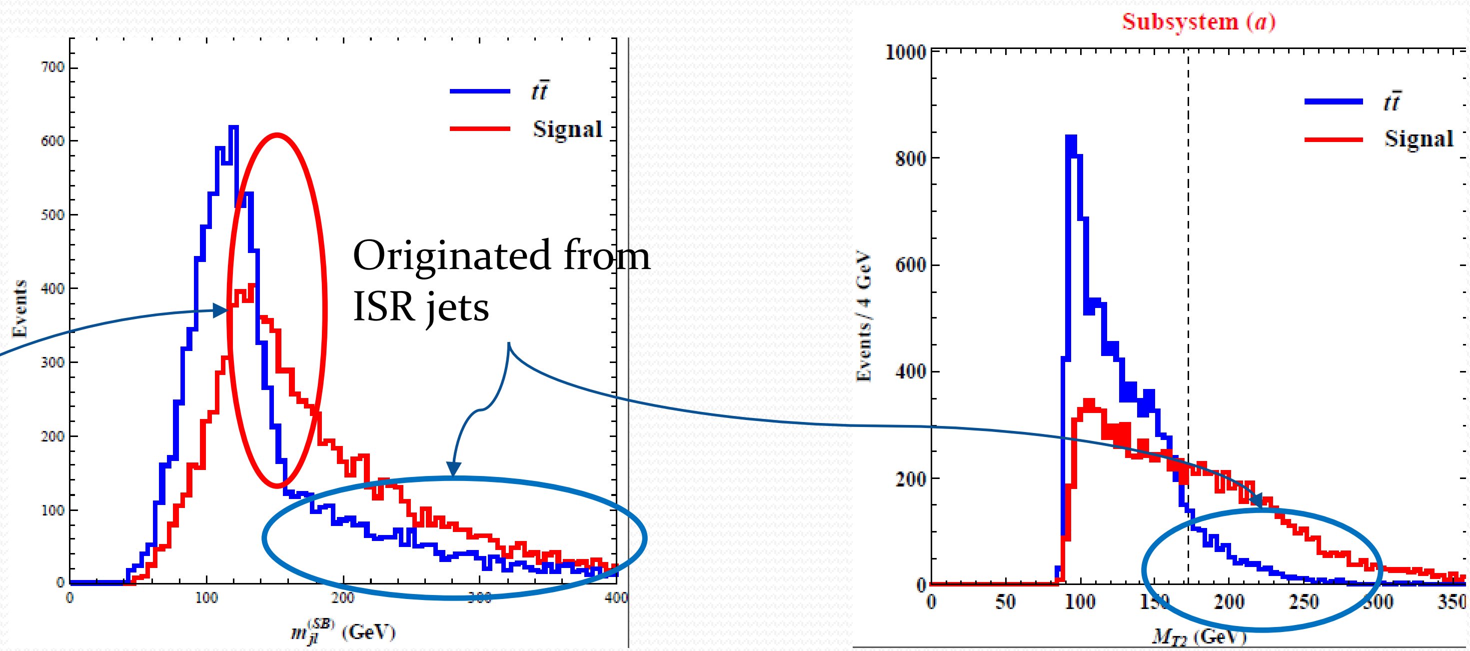


Z mass window for same flavor channel

Results with Detector-level Events

Challenging observables

- (ISR/FSR) Jets are our "problem children" → Better understanding is needed!!



Endpoint structure is still there