Detecting Kinematic Boundary surfaces in Phase Space and Particle Mass Measurement

Dipsikha Debnath

University of Florida

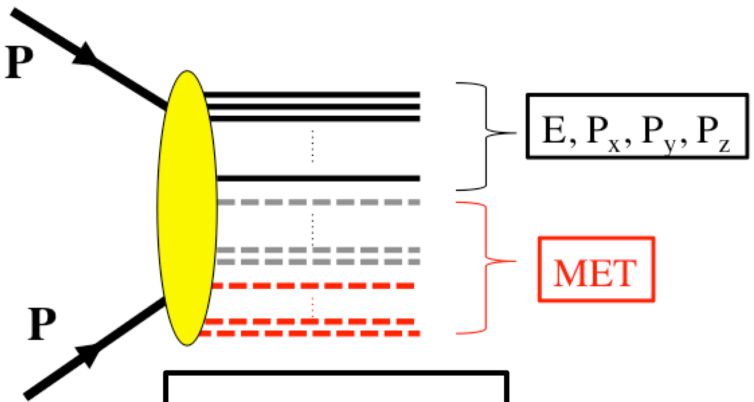
Work with James Gainer, Can Kilic, Doojin Kim, Konstantin Matchev, Yuan-Pao Yang

Based on arXiv: 1611.04487 [hep-ph]

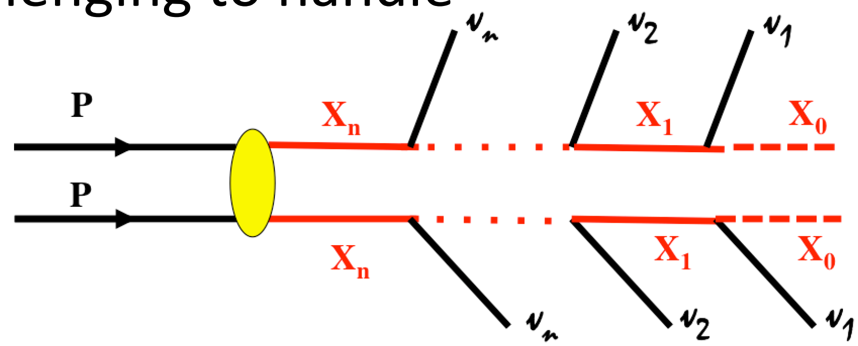
New Perspectives 2017

MET events

MET events are challenging to handle



Experiment



Theory

- Only SM particles v_i's can be seen in the detector
  - -How do you identify the two chains?
- X_0's (WIMP) are invisible, i.e, momenta unknown (except P_T sum)
  - What about SM neutrinos among the x_i' s?
  - How well can you reconstruct the WIMP momenta?

Measuring the unknown particle masses is a difficult problem.

Investi-Gators

Mass measurement in MET events

What to do about unknown invisible momenta?

Ignore

Focus on the visible sector only

-kinematic endpoint method

[Hinchliffe et al. 1997]

Guess

Use an ansatz which optimizes (min/max) a suitable function of the momenta

-MT2, M2 assisted method

[Cho,Choi,Kim,Park 2008]
[Kim, Matchev, Moortgat, Pape 2017]

Compute

Use on-shell relations and the MET constraint to compute exactly

-Polynomial method

[Cheng et al. 2008]

Mass measurement in MET events

What to do about unknown invisible momenta?

Ignore

Focus on the visible sector only

-kinematic endpoint method

[Hinchliffe et al. 1997]

Guess

Use an ansatz which optimizes (min/max) a suitable function of the momenta

-MT2, M2 assisted method

[Cho,Choi,Kim,Park 2008]
[Kim, Matchev, Moortgat, Pape 2017]

Compute

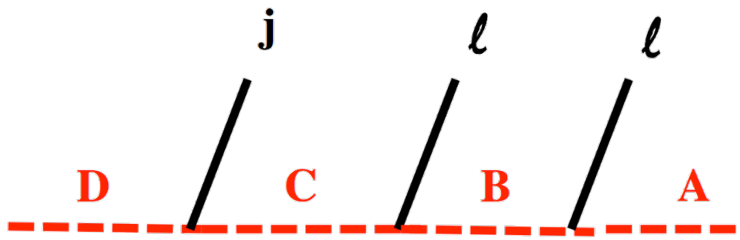
Use on-shell relations and the MET constraint to compute exactly

-Polynomial method

[Cheng et al. 2008]

The classic endpoint method

- Identify a sub-chain and form all possible invariant mass distributions of different pairs of visible particles



A, B, C , D : unknown particles
j, l : visible known particle

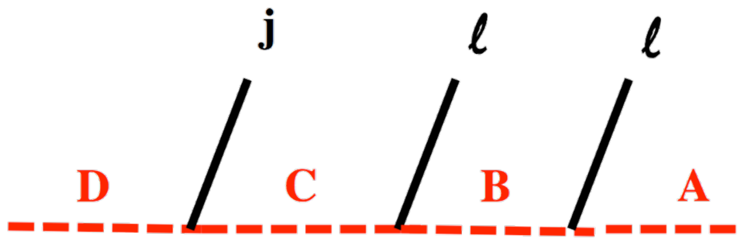
- Measure the endpoints to solve for 4 masses $\{m_D, m_C, m_B, m_A\}$.
- There are 5 endpoints:

$$\{m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}, m_{jl(hi)}^{max}, m_{jll(\theta>\frac{\pi}{2})}^{min}\}$$

- 5 measurements, 4 unknowns. Should be more than sufficient. Not so fast!

The classic endpoint method

- Identify a sub-chain and form all possible invariant mass distributions of different pairs of visible particles



A, B, C , D : unknown particles
j, l : visible known particle

- Measure the endpoints to solve for 4 masses $\{m_D, m_C, m_B, m_A\}$.
- There are 5 endpoints:

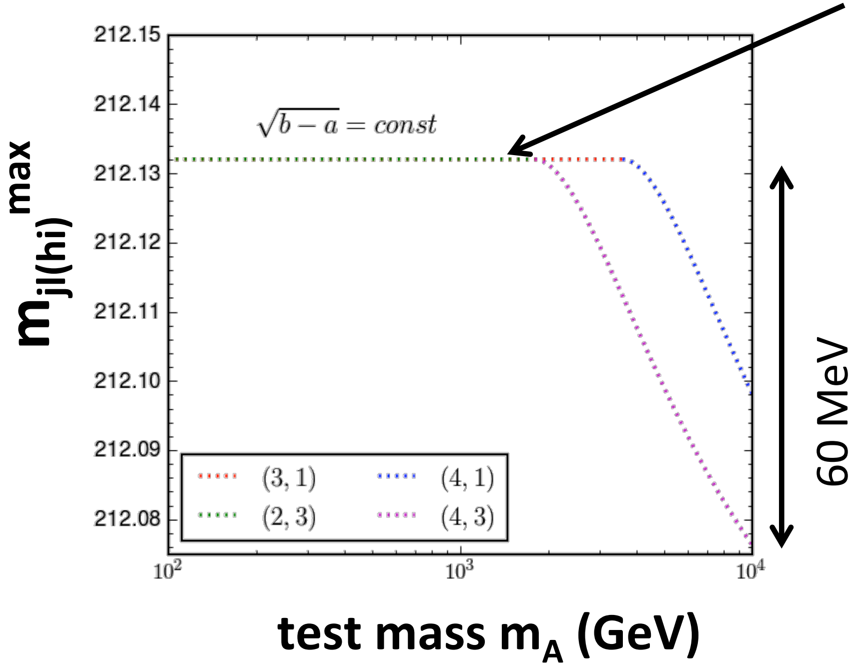
$$\{\underbrace{m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}}_{m_B,\ m_C,\ m_D}, m_{jl(hi)}^{max}, m_{jll(\theta>\frac{\pi}{2})}^{min}\}$$

? ?

m_A

- 5 measurements, 4 unknowns. Should be more than sufficient. Not so fast!

Endpoint $m_{jl(hi)}^{max}$

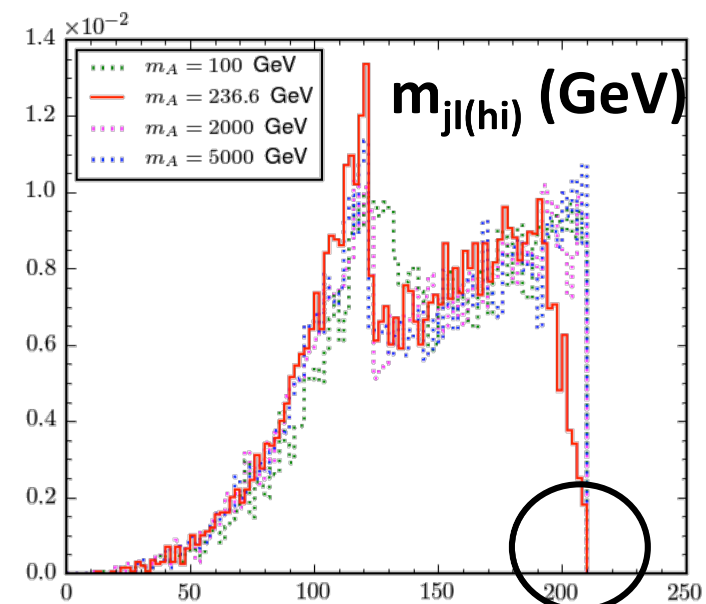


- This endpoint measurement is not independent.

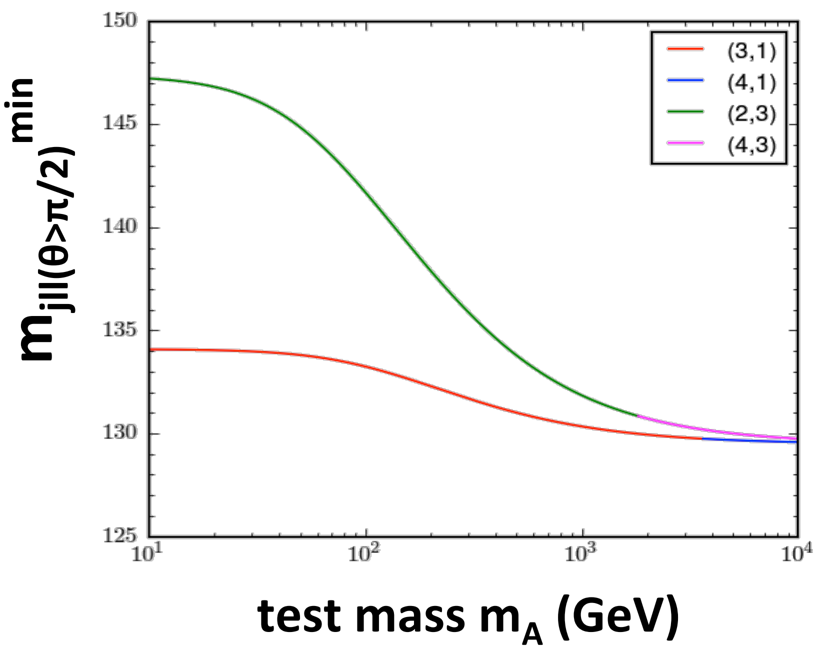
$$(m_{jll}^{max})^2 = (m_{jl(hi)}^{max})^2 + (m_{ll}^{max})^2$$

- $m_{jl(hi)}^{max}$ is not that sensitive to the test mass.

The variation of endpoint is unnoticeable even with large statistics and no background.

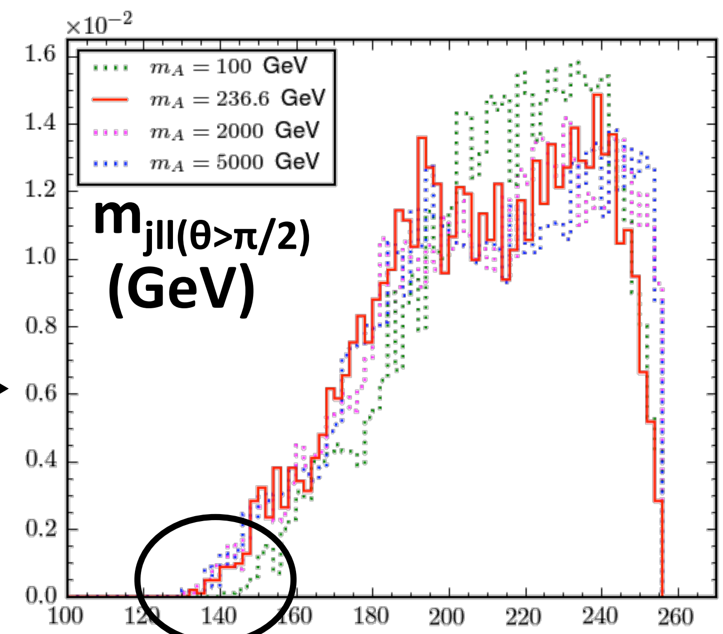


Endpoint $m_{jll(\theta>\pi/2)}^{min}$



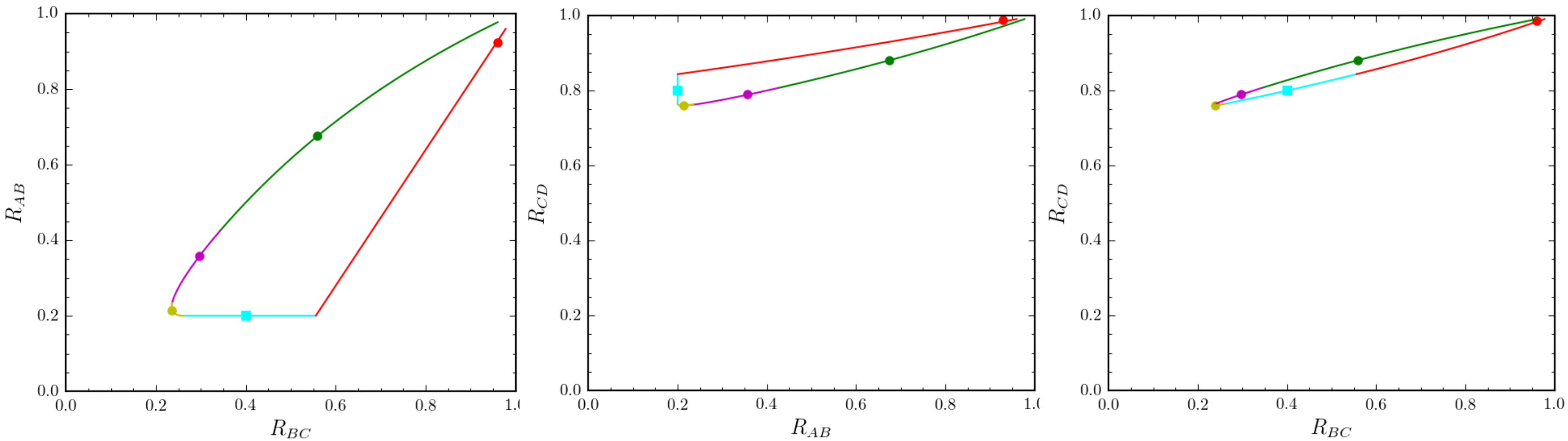
- $m_{jll(\theta>\pi/2)}^{min}$ is not sensitive to test mass.
- Not easy to spot lower endpoint as background would dominate.

The signal distribution is very poorly populated near its lower endpoint



Trajectory in mass parameter space

- The two trajectories in mass parameter space leading to the same endpoints m_{ll}^{max}, m_{jll}^{max}, $m_{jl(lo)}^{max}$ and mass of particle A.



Colors are different regions in mass parameter space

$$R_{AB} = \frac{m_A^2}{m_B^2}, R_{BC} = \frac{m_B^2}{m_C^2}, R_{CD} = \frac{m_C^2}{m_D^2}$$

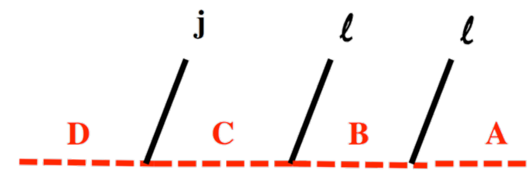
Cyan dot is the true study point

The flat direction ($m_B(m_A)$, $m_C(m_A)$, $m_D(m_A)$, m_A) in mass parameter space can not be identified with the endpoint method.

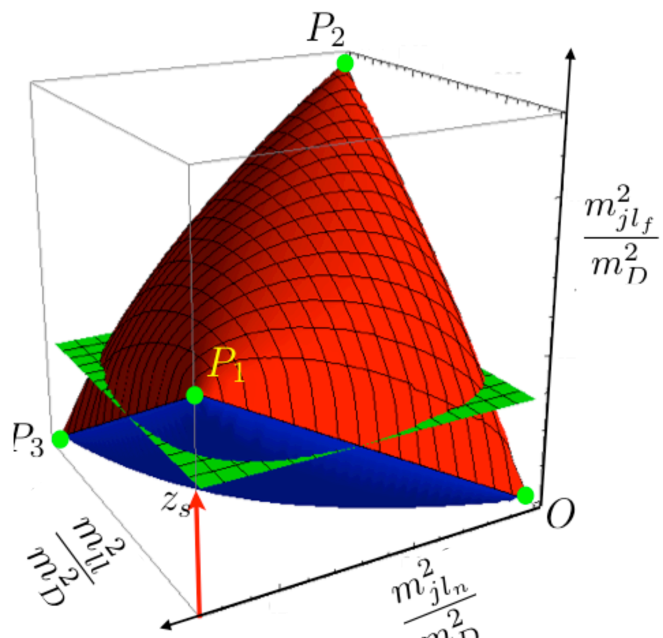
Try alternative method

The phase space

- Our proposal: measure the phase space, do not project to lower dimension.



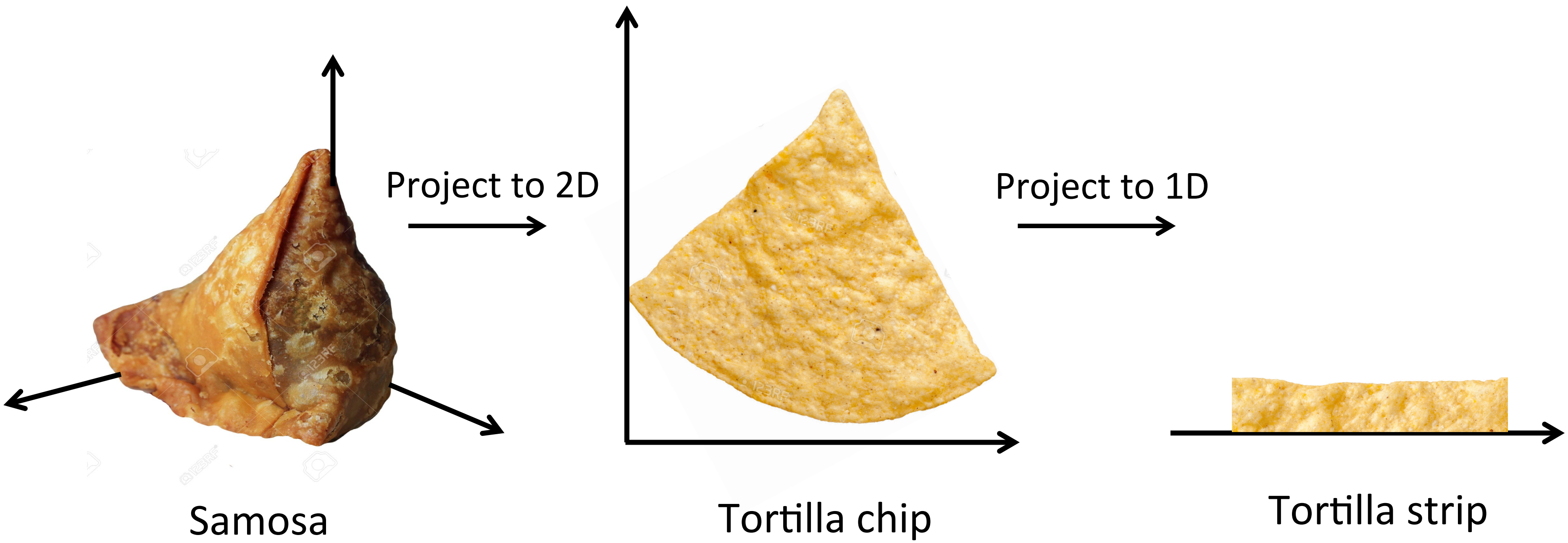
The Signal region is compact, looks like "Samosa"



[Kim, Park, Matchev 2015]

- Advantage: uses the maximal amount of information

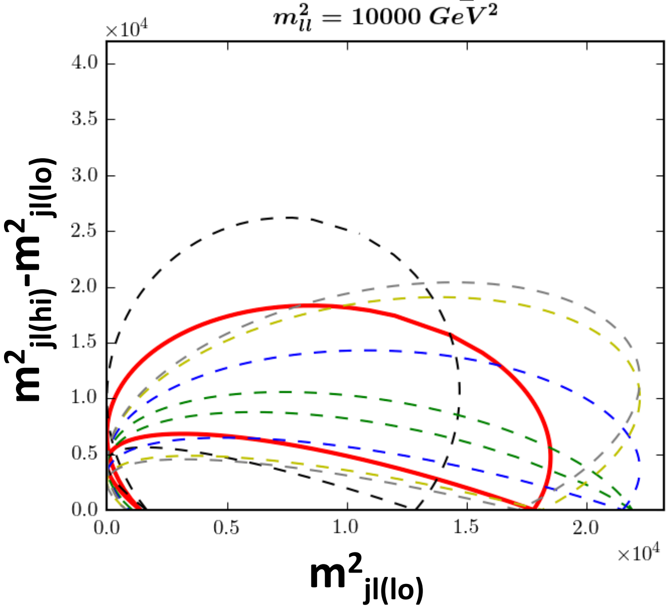
Do not ignore a samosa
It has filling too….



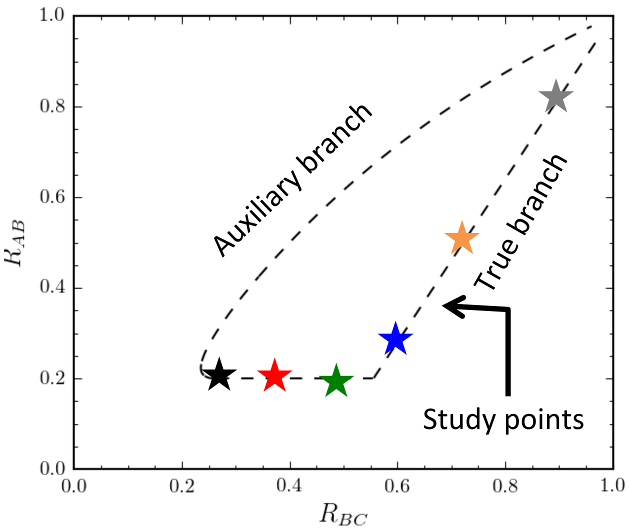
Samosa boundaries

- The boundary of samosa is a 2D surface which depends on input masses.

$$\mathcal{S}: \quad \frac{m^2_{j\ell_f}}{(m^{max}_{j\ell_f})^2} = \left[\sqrt{\frac{m^2_{\ell\ell}}{(m^{max}_{\ell\ell})^2}\left(1-\frac{m^2_{j\ell_n}}{(m^{max}_{j\ell_n})^2}\right)} \pm \frac{m_B}{m_C}\sqrt{\frac{m^2_{j\ell_n}}{(m^{max}_{j\ell_n})^2}\left(1-\frac{m^2_{\ell\ell}}{(m^{max}_{\ell\ell})^2}\right)}\right]^2$$



2D slice of samosa



Trajectory in mass parameter space

$$m_{jl(lo)} \equiv \min\left\{m_{jl_n}, m_{jl_f}\right\}$$
$$m_{jl(hi)} \equiv \max\left\{m_{jl_n}, m_{jl_f}\right\}$$

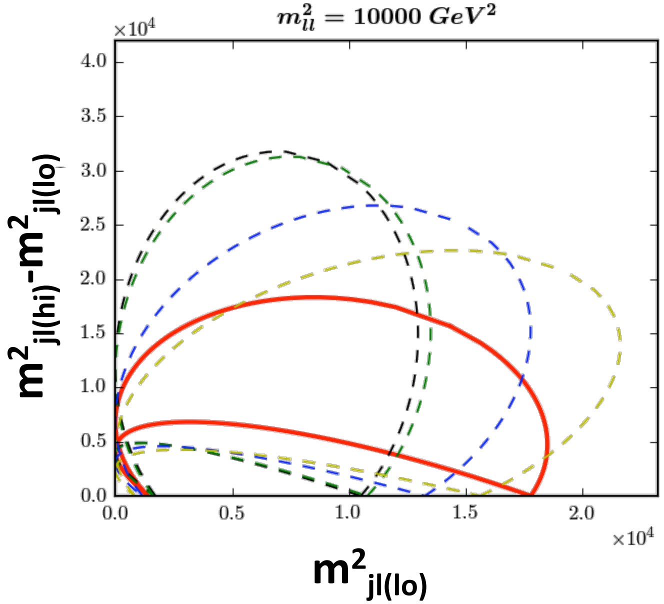
The shape and size of samosa vary significantly with input masses. But projection on to one dimension loses that information.

Red solid line : true study point

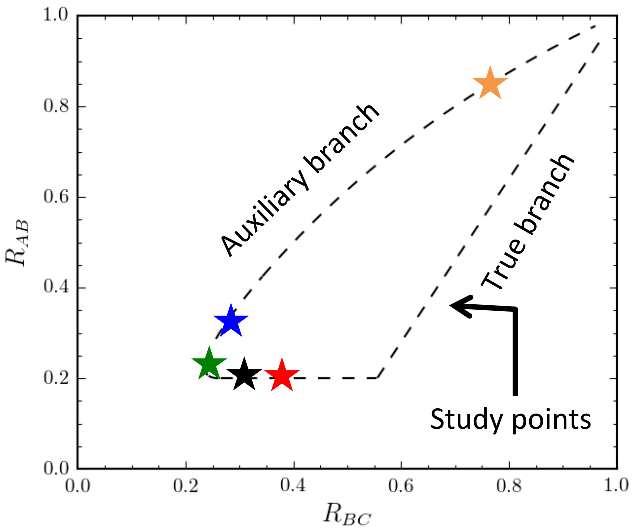
Samosa boundaries

- The boundary of samosa is a 2D surface which depends on input masses.

$$\mathcal{S}: \quad \frac{m^2_{j\ell_f}}{(m^{max}_{j\ell_f})^2} = \left[\sqrt{\frac{m^2_{\ell\ell}}{(m^{max}_{\ell\ell})^2}\left(1-\frac{m^2_{j\ell_n}}{(m^{max}_{j\ell_n})^2}\right)} \pm \frac{m_B}{m_C}\sqrt{\frac{m^2_{j\ell_n}}{(m^{max}_{j\ell_n})^2}\left(1-\frac{m^2_{\ell\ell}}{(m^{max}_{\ell\ell})^2}\right)}\right]^2$$



2D slice of samosa



Trajectory in mass parameter space

$$m_{jl(lo)} \equiv \min\{m_{jl_n}, m_{jl_f}\}$$
$$m_{jl(hi)} \equiv \max\{m_{jl_n}, m_{jl_f}\}$$

The shape and size of samosa vary significantly with input masses. But projection on to one dimension loses that information.

Red solid line : true study point

Surface fitting

- Design a global variable for surface fitting:

$$\bar{\Sigma}(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D) \equiv \frac{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da\, |\vec{\nabla}\rho(\vec{r})|}{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da}$$

The gradient is largest at the signal boundary

- In order to compute the gradient or find signal boundary, we take use of geometric properties of Voronoi tessellations.

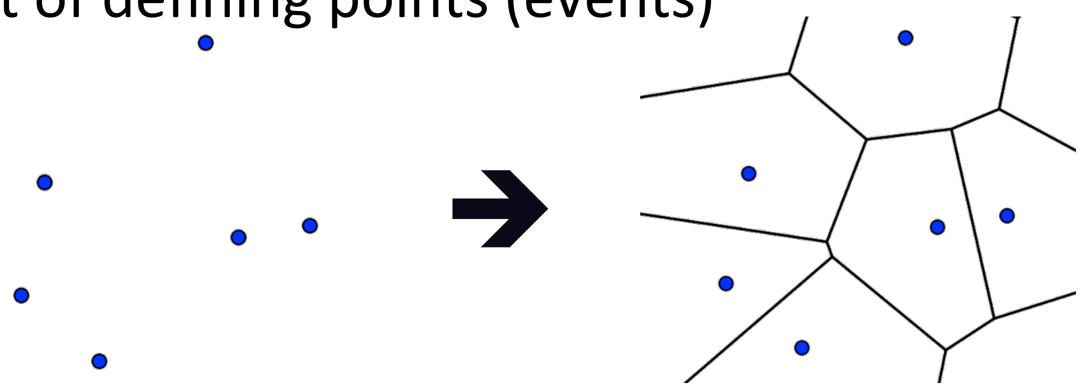
Surface fitting

- Design a global variable for surface fitting:

$$\bar{\Sigma}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D) \equiv \frac{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da\,|\vec{\nabla}\rho(\vec{r})|}{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da}$$

The gradient is largest at the signal boundary

- In order to compute the gradient or find signal boundary, we take use of geometric properties of Voronoi tessellations.
- **Voronoi (Dirichlet) Tessellation:** a tessellation where each tile is defined as the set of points closest to one of the points in a discrete set of defining points (events)



Each Voronoi cell has only 1 data point

First use of VT in HEP was in Fermilab by D0 collaboration [hep-ex/006011]

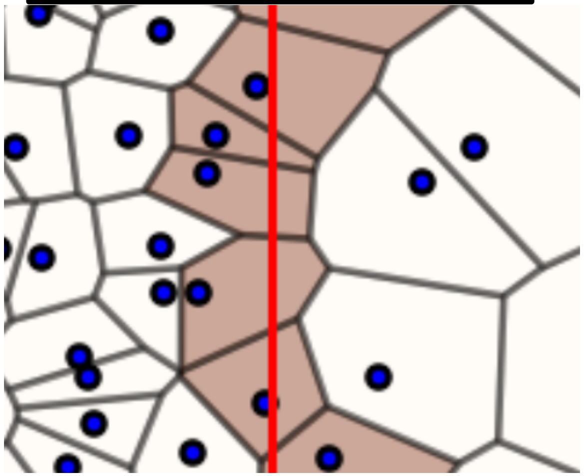
Find good variable

- Use Voronoi tessellations in 3D to spot the kinematic boundary.
  - -Find a test statistic for signal (boundary cells) to be well-separated from that of the background (non-boundary cells)

Find good variable

- Use Voronoi tessellations in 3D to spot the kinematic boundary.
  -Find a test statistic for signal (boundary cells) to be well-separated from that of the background (non-boundary cells)
- A good variable: Relative standard deviation (RSD) of neighboring cell volume.

Example: Flat boundary



NP+SM NP+SM SM

$$\bar{\sigma}_i \equiv \frac{1}{\langle v(N_i)\rangle} \sqrt{\sum_{j \in N_i} \frac{\left(v_j - \langle v(N_i)\rangle\right)^2}{|N_i| - 1}}$$

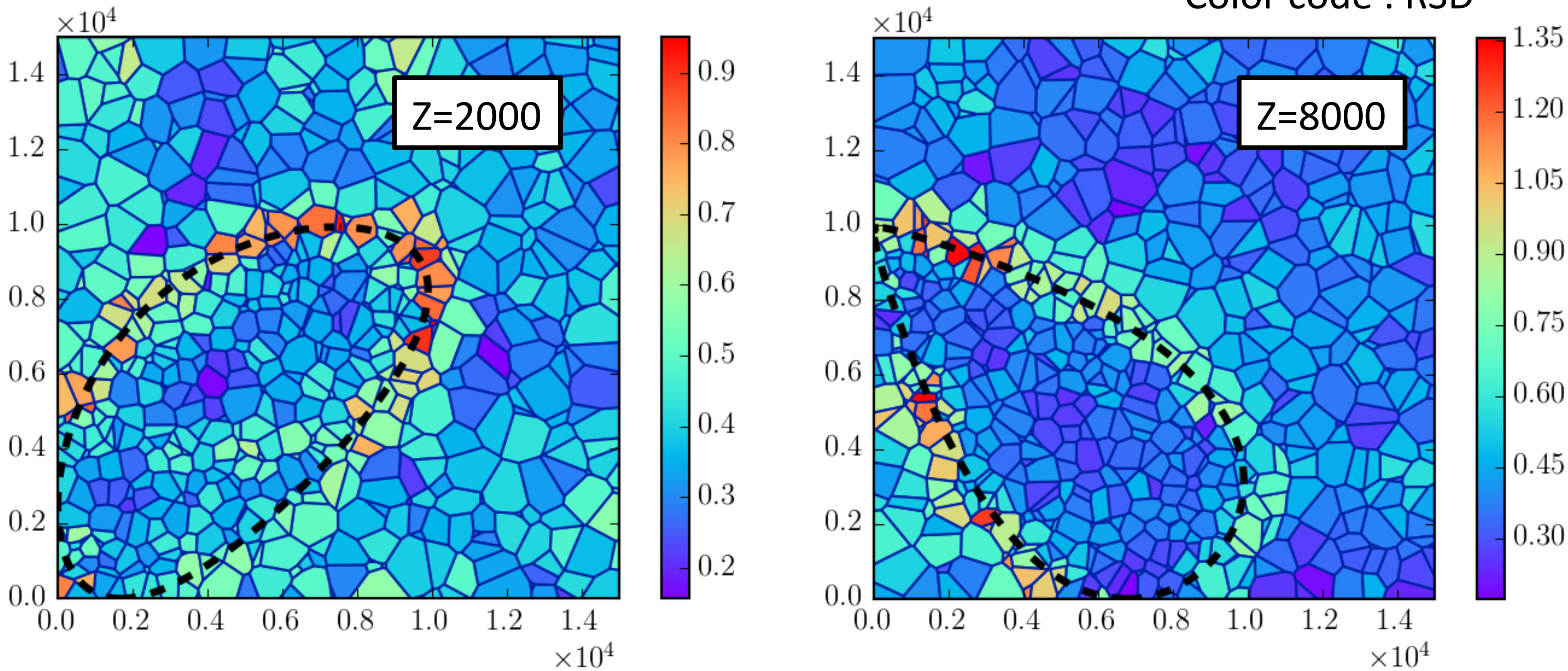
Boundary cells have a big spread in neighboring cell volume.

Relative Standard Deviation plots

2D slices of 3D Voronoi tessellations of Signal+Background

Along x axis m_{ll}^2, y axis m_{jln}^2, z axis m_{jlf}^2

Color code : RSD



RSD is relatively larger for boundary cells (the black dashed line).
Thus RSD is a good variable to pick out the boundary cells.

Surface fitting

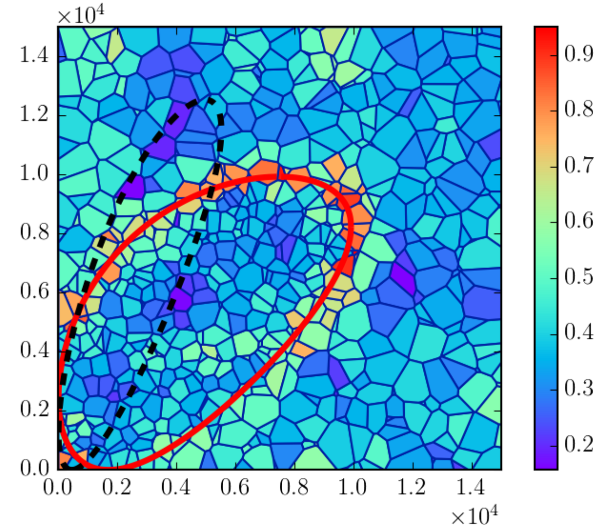
- In order to stay model independent we relate the gradient in the surface fitting variable with the Relative Standard Deviation (RSD) of Voronoi cells.

$$\bar{\Sigma}(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D) \equiv \frac{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da \, |\vec{\nabla}\rho(\vec{r})|}{\int_{\tilde{S}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da}$$

- Replace the gradient estimator function by

$|\vec{\nabla}\rho(\vec{r})| \longrightarrow \bar{\sigma}_i$ ≅ RSD for i-th Voronoi cell

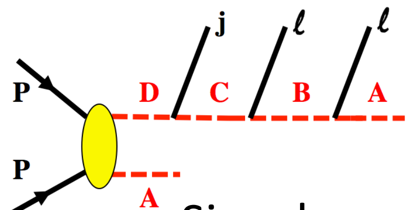
- The maximum of this variable is expected to occur at the correct choice of masses (m_A, m_B, m_C, m_D).



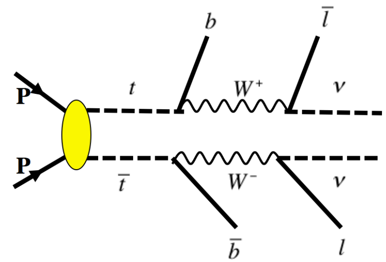
Red: correct guess of masses
Black: wrong guess of masses

Mass parameter scan

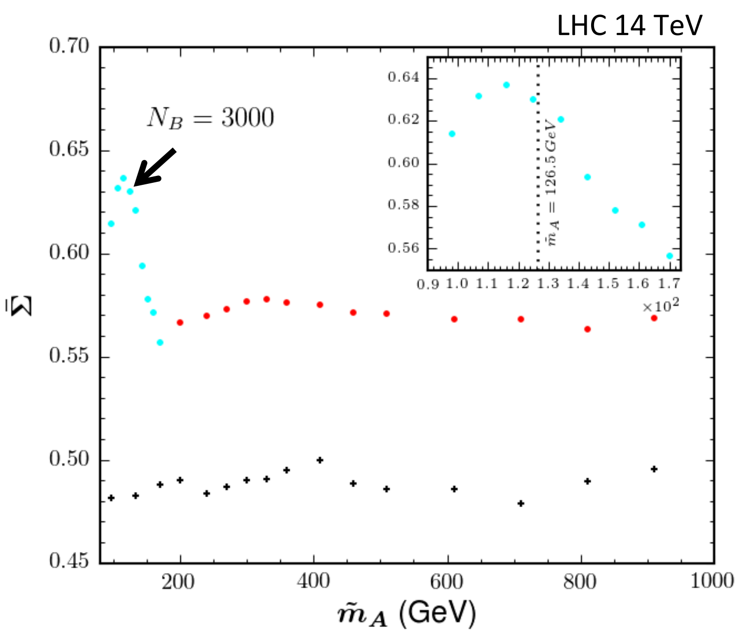
- The mass parameter scan for $\bar{\Sigma}(\tilde{m}_A, \tilde{m}_B(\tilde{m}_A), \tilde{m}_C(\tilde{m}_A), \tilde{m}_D(\tilde{m}_A))$

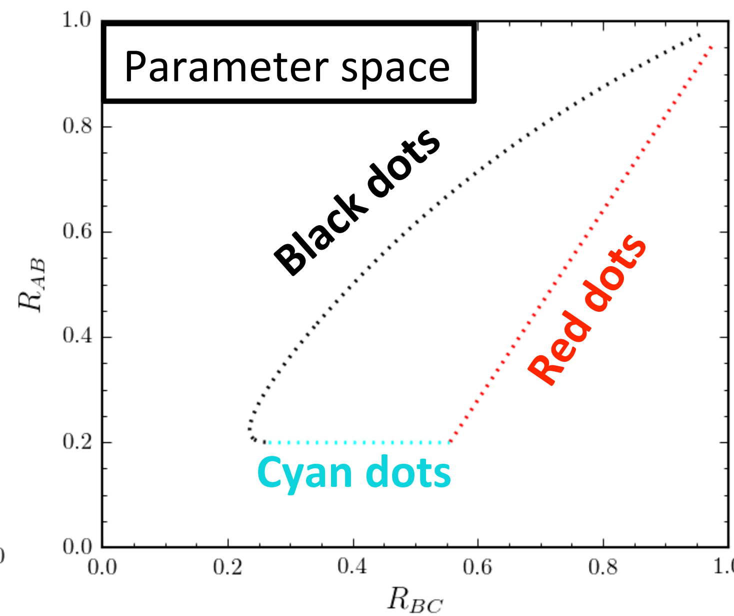


Signal



Background

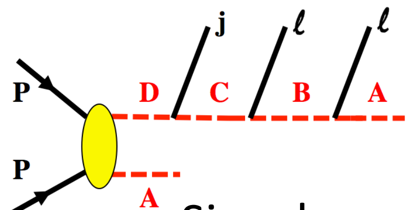




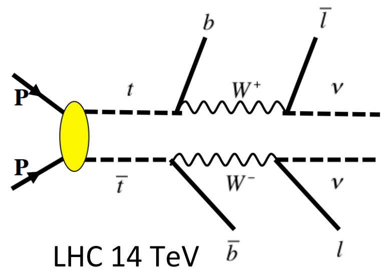
The function peaks near true value of the particle mass

Detector effect

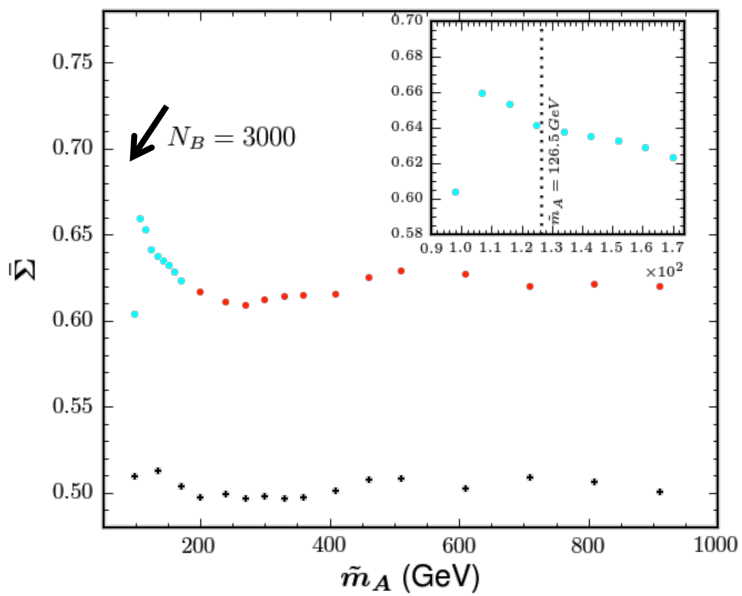
- Repeat with detector resolution



Signal

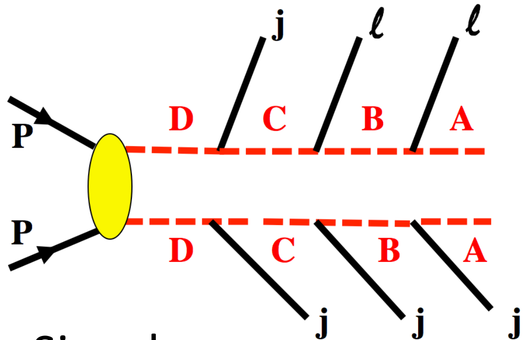


Background

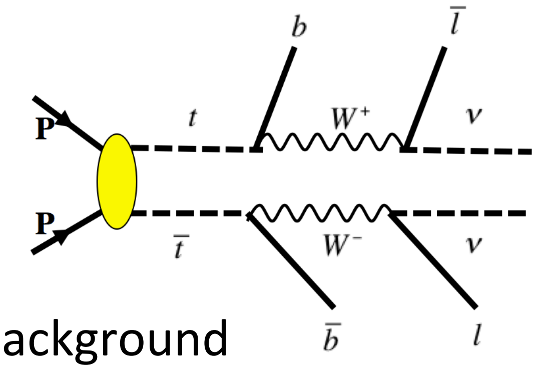


The peak structure is preserved but degraded due to detector resolution

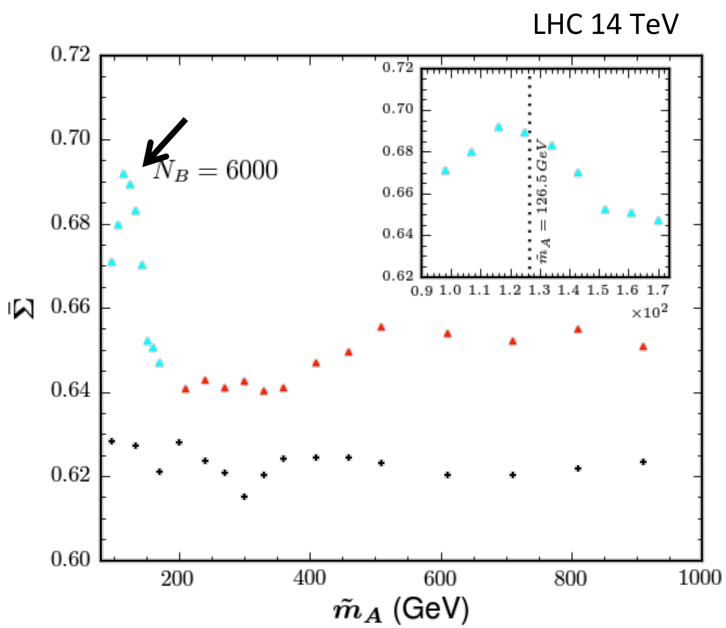
Combinatorics effect



Signal



Background



Include the two jets which give smallest jet-lepton-lepton invariant masses i.e, 2 fold ambiguity

The peak structure is still preserved near the true mass

Closing remarks

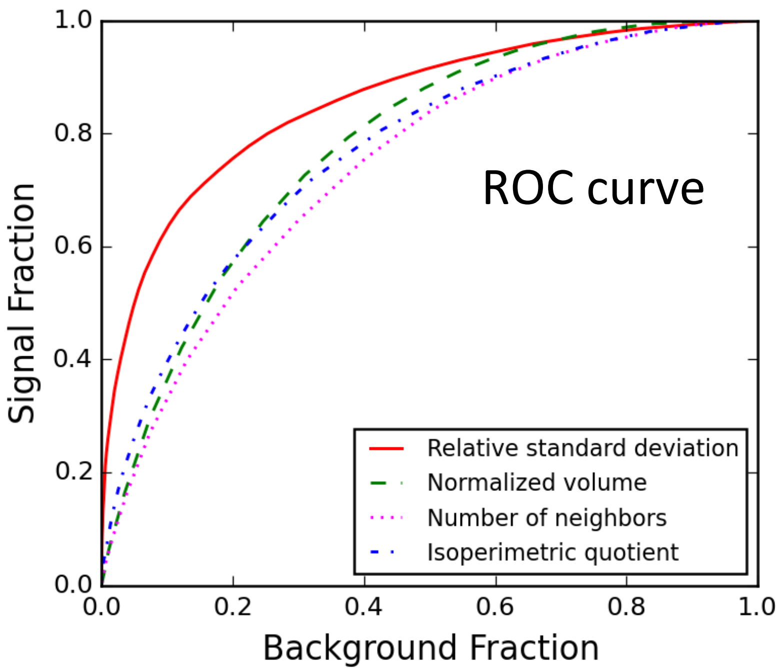
- We have examined the classic endpoint method for particle mass determination
  -some measurements are not independent
  -some are affected by experimental resolution
- Our proposed method takes advantage of full dimensional information.
- We have tested our Voronoi-based algorithm for detecting the 2D boundary surface and demonstrated that it can be usefully applied in order to lift the degeneracy along the flat direction.
- Future direction-
  - estimate the statistical precision of the mass measurement method.
  - generalize to longer decay chain with more visible particles.

Thanks for your attention!

Question?

RSD as a good variable

- Use Voronoi tessellations in 3D to spot the kinematic boundary.
  -Find a test statistic for signal (boundary cells) to be well-separated from that of the background (non-boundary cells)
- A good variable: Relative standard deviation (RSD) of neighboring cell volume.



Signal: Boundary cells
Background: Non-boundary cells

The red solid line shows that RSD is a good variable to pick out the boundary cells, i.e the signal boundary