Improving mass measurement in cascade decay with Voronoi tessellations

Yuan-Pao Yang

University of Texas at Austin

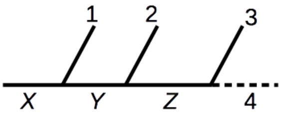
Based on:
1611.04487, 1606.02721
with
D. Debnath, J. S. Gainer, C. Kilic,
D. Kim, and K. T. Matchev

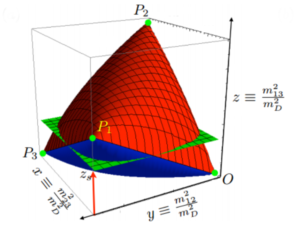
MC4BSM, 2017

Outline

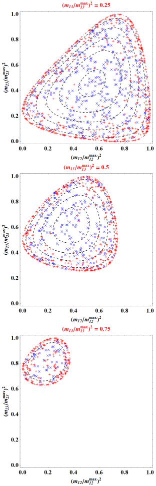
Outline

Appendix to Matt's talk...





From 1512.02222 (Kim, Matchev, Park)

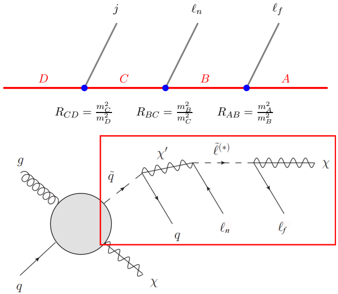


- "2+2+2" topology
- 3-D phase space $(m_{12}^2, m_{13}^2, m_{23}^2)$
- Enhancement on the boundary

Boundary
$\Updownarrow$
Mass measurement

Why we need boundary

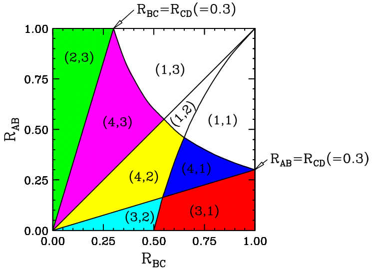
We want to do a realistic mass measurement!



- "2+2+2" topology!
- Traditional way
  - $m_{j\ell_n}^{max}$, $m_{j\ell_f}^{max}$, $m_{\ell\ell}^{max}$, $m_{j\ell\ell}^{max}$
  - 4 unknowns, 4 equations
- Combinatorial problem
  - Cannot tell ℓ_n and ℓ_f
  - $m_{j\ell(hi)} \equiv max\{m_{j\ell_n}, m_{j\ell_f}\}$
  - $m_{j\ell(lo)} \equiv min\{m_{j\ell_n}, m_{j\ell_f}\}$
  - $m_{\ell\ell}^{max}$, $m_{j\ell\ell}^{max}$, $m_{j\ell(hi)}^{max}$, $m_{j\ell(lo)}^{max}$
  - 4 unknowns, 4 equations, again
- What goes wrong with classical method?

Why we need boundary

- Piecewise-defined formulae!
- In $(2,3)$, $(3,1)$, $(3,2)$:
  - $(m_{j\ell\ell}^{max})^2 = (m_{\ell\ell}^{max})^2 + (m_{j\ell(hi)}^{max})^2$
  - 4 unknowns, 3 equations
  - Classical method doesn't work!



Study point	P_{32}	P_{31}
m_A (GeV)	126.5	5000.0
m_B (GeV)	282.8	5207.4
m_C (GeV)	447.2	5324.2
m_D (GeV)	500.0	5372.1
$m_{\ell\ell}^{max}$ (GeV)	309.8	
$m_{j\ell\ell}^{max}$ (GeV)	368.8	
$m_{j\ell(hi)}^{max}$ (GeV)	149.1	
$m_{j\ell(lo)}^{max}$ (GeV)	200.0	

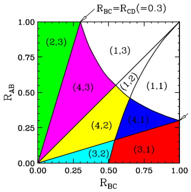
- For a set of given edge $\rightarrow$ 1-parameter family!

Why we need boundary

edges and end-points formulae

$$\left(m_{ll}^{max}\right)^2 = m_D^2\, R_{CD}\,(1-R_{BC})\,(1-R_{AB})$$

$$\left(m_{j\ell\ell}^{max}\right)^2 = \begin{cases} m_D^2(1-R_{CD})(1-R_{AC}), & R_{CD} < R_{AC}, \quad \text{case } (1,-) \\ m_D^2(1-R_{BC})(1-R_{AB}R_{CD}), & R_{BC} < R_{AB}R_{CD}, \quad \text{case } (2,-) \\ m_D^2(1-R_{AB})(1-R_{BD}), & R_{AB} < R_{BD}, \quad \text{case } (3,-) \\ m_D^2\left(1-\sqrt{R_{AD}}\right)^2, & \text{otherwise}, \quad \text{case } (4,-) \end{cases}$$



$$\left(m_{j\ell(lo)}^{max}\right)^2 = \begin{cases} m_D^2\,(1-R_{CD})\,(1-R_{BC}), & (2-R_{AB})^{-1} < R_{BC} < 1, \quad \text{case } (-,1) \\ m_D^2\,(1-R_{CD})\,(1-R_{AB})\,(2-R_{AB})^{-1}, & R_{AB} < R_{BC} < (2-R_{AB})^{-1}, \quad \text{case } (-,2) \\ m_D^2\,(1-R_{CD})\,(1-R_{AB})\,(2-R_{AB})^{-1}, & 0 < R_{BC} < R_{AB}, \quad \text{case } (-,3) \end{cases}$$

$$\left(m_{jl(hi)}^{max}\right)^2 = \begin{cases} m_D^2\,(1-R_{CD})\,(1-R_{AB}), & (2-R_{AB})^{-1} < R_{BC} < 1, \quad \text{case } (-,1) \\ m_D^2\,(1-R_{CD})\,(1-R_{AB}), & R_{AB} < R_{BC} < (2-R_{AB})^{-1}, \quad \text{case } (-,2) \\ m_D^2\,(1-R_{CD})\,(1-R_{BC}), & 0 < R_{BC} < R_{AB}, \quad \text{case } (-,3) \end{cases}$$

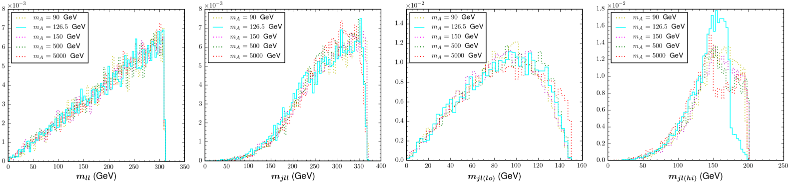
From 0903.4371 (Burns, Matchev, Park)

Why we need boundary

The 1-parameter family

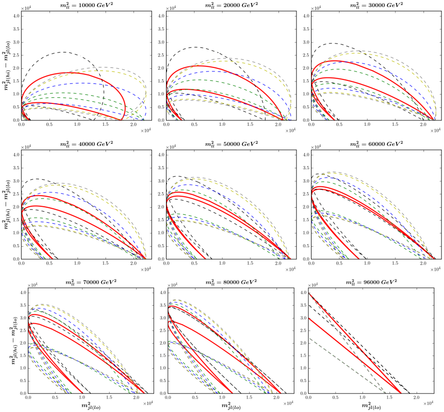
For case of $(m_{j\ell\ell}^{max})^2 = (m_{\ell\ell}^{max})^2 + (m_{j\ell(hi)}^{max})^2$, the best thing we can do is to parametrize this one-parameter family as:

$$\begin{cases} m_D = m_D(m_A; m_{\ell\ell}^{max}, m_{j\ell\ell}^{max}, m_{j\ell(lo)}^{max}), \\ m_C = m_C(m_A; m_{\ell\ell}^{max}, m_{j\ell\ell}^{max}, m_{j\ell(lo)}^{max}), \\ m_B = m_B(m_A; m_{\ell\ell}^{max}, m_{j\ell\ell}^{max}, m_{j\ell(lo)}^{max}), \\ m_A \end{cases}$$



Why we need boundary

The 1-parameter family



Where $\tilde{m}_A$ is

- 100 GeV
- 126.5 GeV*
- 173 GeV
- 500 GeV
- 2000 GeV
- 4000 GeV

How we get the boundary

- Disadvantage of likelihood method:
  - it relies on interior information, *i.e.* matrix element.
  - background (mainly $t\bar{t}$) events populates outside the signal boundary
- Let's try the boundary enhancement!
- Voronoi tessellation partitions the phase space into cells
  - Each point gets a cell
  - Cells have geometry properties:
    - volume
    - neighbor
    - And many others!
  - They are helpful for boundary determination!

Outline

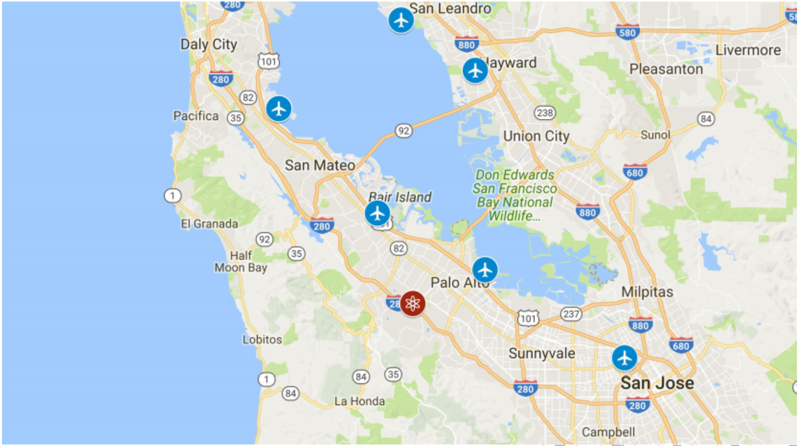
Definition

Optional Subtitle

"The Voronoi diagram for a set of points in a given space $\mathbb{R}^d$ is the partitioning of that space into regions such that all locations within any one region are closer to the generating point than to any other."

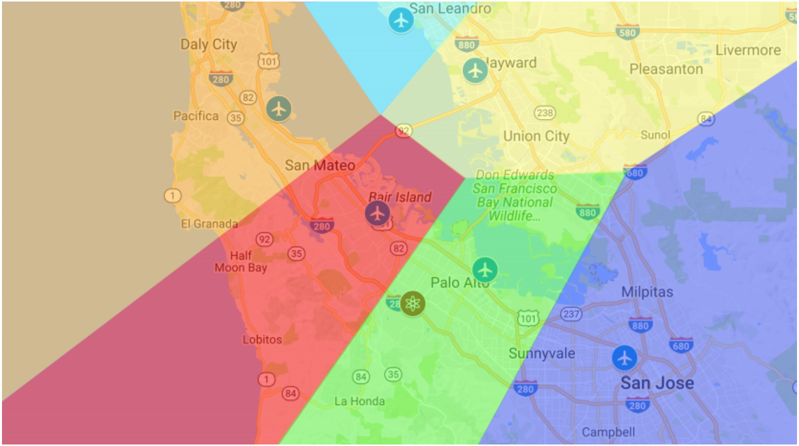
From Voro Wiki

Bay Area Airports

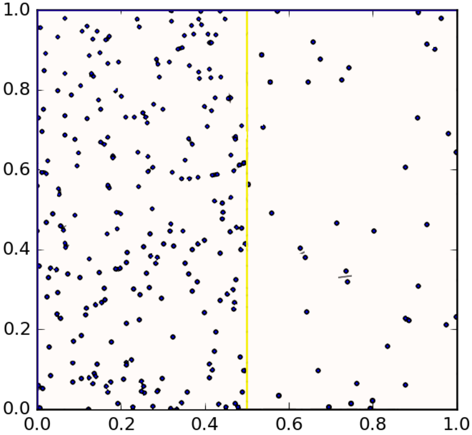


Bay Area Airports

with Voronoi

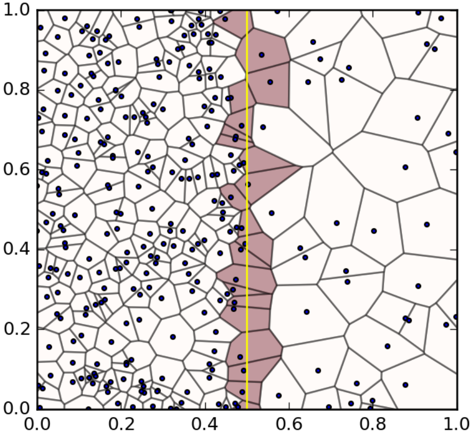


Box



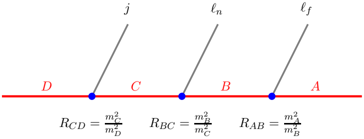
Box

with Vonoroi



Outline

What do we want in the toy model?

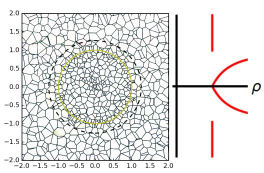


Toy Background

- Uniform!

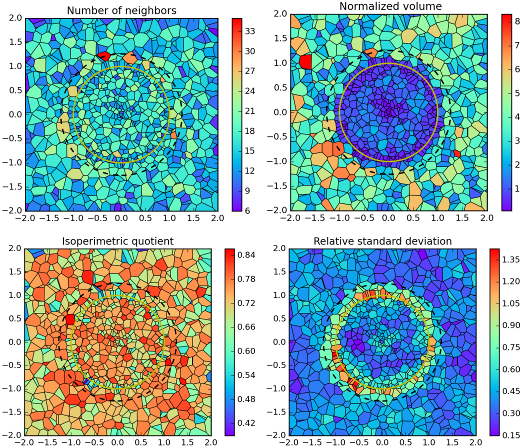
Toy Signal

- 3-D
- Enhancement on the boundary

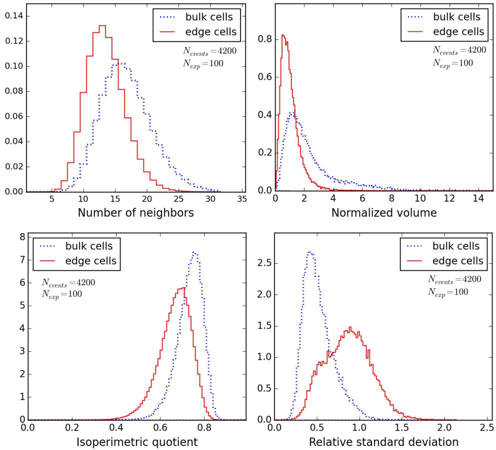


Geometry Properties

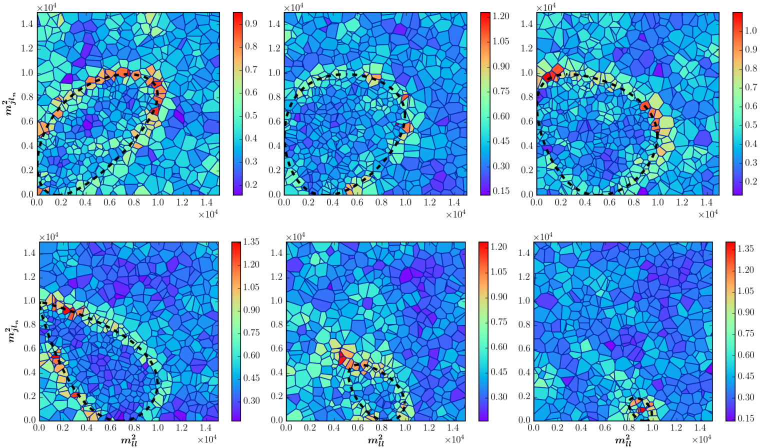
- No. of neighbors
- volume
- isoperimetric q: $\dfrac{\text{volume}}{(\text{surface area})^{3/2}}$
- relative std: $\dfrac{\text{std of neighbor's volume}}{\text{average of neighbor's volume}}$



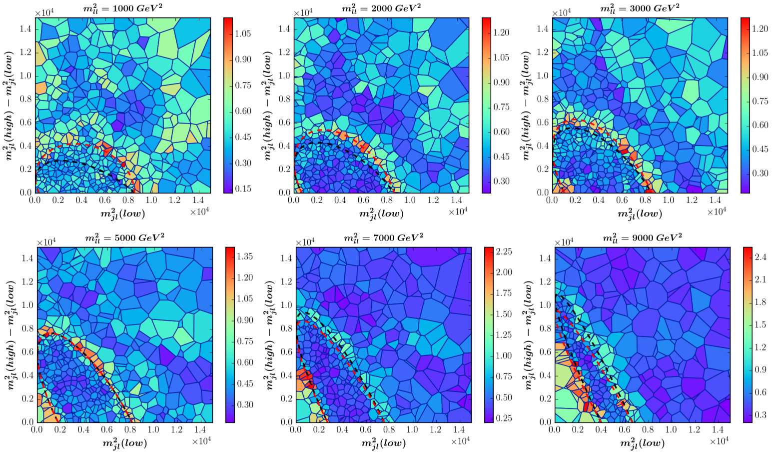
Geometry Properties



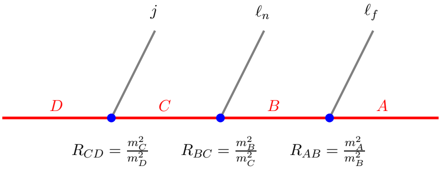
Simple Case

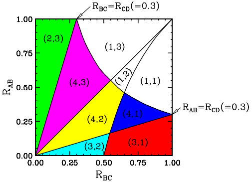

×10^4
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4
$m^2_{jl_n}$
m^2_{ll}
0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2
1.20 1.05 0.90 0.75 0.60 0.45 0.30 0.15
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2
1.35 1.20 1.05 0.90 0.75 0.60 0.45 0.30
1.20 1.05 0.90 0.75 0.60 0.45 0.30
1.35 1.20 1.05 0.90 0.75 0.60 0.45 0.30 0.15

With Combinatorics



Review of the problem





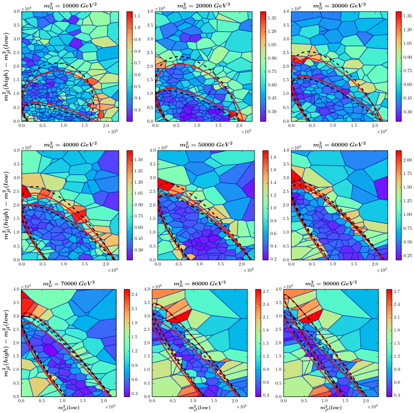
Study Case: (m_A, m_B, m_C, m_D)=(126.5, 282.8, 447.2, 500) GeV

- Inside (3, 2) region $((m_{\ell\ell}^{max})^2 + (m_{j\ell(hi)}^{max})^2 = (m_{j\ell\ell}^{max})^2)$
- Cannot solve the 4 equations after measuring the 4 edges (4 unknowns, 3 ind. equations)
- Can express as $(\tilde{m}_A, \tilde{m}_B(\tilde{m}_A), \tilde{m}_C(\tilde{m}_A), \tilde{m}_D(\tilde{m}_A))$ 1-parameter family

Scan the 1-parameter family

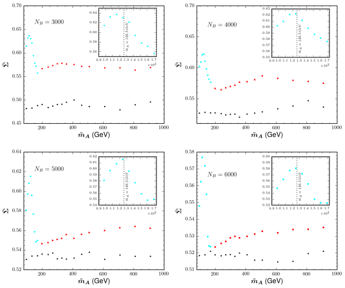
1. We can explore the 1-parameter family by varying $\tilde{m}_A$
2. For a given $\tilde{m}_A$, we can know the cells laying on the boundary
3. We can calculate the average RSD per unit area of boundary cells
4. Expect the true spectrum will maximize that average

Scan the 1-D family



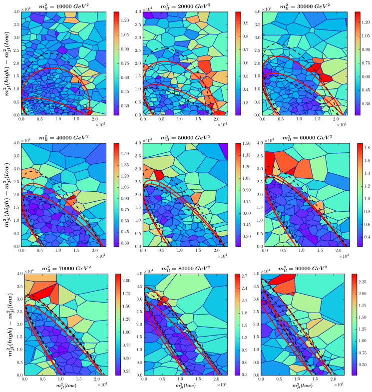
N_S	3000
m_A	126.5

N_B	3000	4000	5000	6000
$\tilde{m}_A$	116	125	125	125

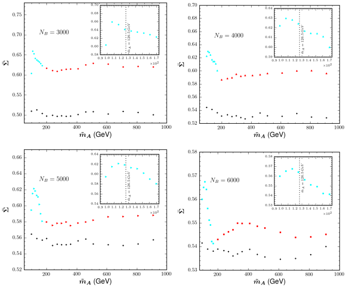


Scan the 1-D family

With detector effect



N_S	3000			
m_A	126.5			
N_B	3000	4000	5000	6000
$\tilde{m}_A$	107	107	116	116



Summary

- Have done...
  - With Voronoi tessellation, determined the boundary of "2+2+2" topology signal under $t\bar{t}$ background, combinatorics, and detector effect.
- Potential extension:
  - different topology
  - discovery!

Backup

Energy resolution

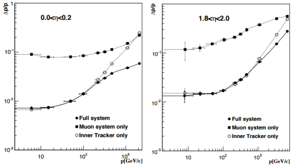
- Hadronic calorimeter resolution

$$\frac{\sigma}{E} = \left(\frac{1}{\sqrt{E}} \right)$$

- Eletromagnetic calorimeter resolution

$$\left(\frac{\sigma}{E} \right)^2 = \left(\frac{0.0363}{\sqrt{E}} \right)^2 + \left(\frac{0.124}{E} \right)^2 + 0.0026^2$$

- Muon momentum resolution



From "CMS physics: Technical design report," CERN-LHCC-2006-001, CMS-TDR-008-1