

Binning Bonanza: Metric Design for Quantum Observables and Efficient Dimensionality Reduction

Mohit Srivastav
for the MiLoMerge Authors

[Pheno 2026](#)

May 12, 2026



Supporting Material

- What will be presented is shown in [ArXiv:2601.10822](https://arxiv.org/abs/2601.10822)
“*Maximizing Returns: Optimizing Experimental Observables at the LHC*”

• **Not possible
without
everyone!**

arXiv > hep-ph > arXiv:2601.10822

High Energy Physics - Phenomenology

[Submitted on 15 Jan 2026]

Maximizing Returns: Optimizing Experimental Observables at the LHC

Jeffrey Davis, Andrei V. Gritsan, Lucas S. Mandacaru Guerra, Lucas Kang, Michalis Panagiotou, Jeffrey Roskes, Mohit Srivastav

We introduce a framework that integrates both analytical and machine-learning approaches for calculating observables optimal for EFT and broader applications at the LHC. A new metric for evaluating the performance of these approaches has been introduced. In addition, we demonstrate how the majority of relevant information can be effectively stored in a limited number of bins, allowing for efficient data analysis, data preservation, and global data combination, while also providing tools to achieve these benefits. A key feature of this approach is the reduction in the dimensionality of the observable information, which enhances both the effectiveness and practicality of the data analysis while maximizing gains within limited resources. These features have been demonstrated through simulated analyses of the Higgs boson production and decay processes at the LHC.

Subjects: **High Energy Physics - Phenomenology (hep-ph)**; High Energy Physics - Experiment (hep-ex)

Cite as: [arXiv:2601.10822](https://arxiv.org/abs/2601.10822) [hep-ph]

(or [arXiv:2601.10822v1](https://arxiv.org/abs/2601.10822v1) [hep-ph] for this version)

<https://doi.org/10.48550/arXiv.2601.10822>

Submission history

From: Mohit Srivastav [[view email](#)]

[v1] Thu, 15 Jan 2026 19:48:17 UTC (744 KB)



Interference in Collider Physics

- Searches in EFT (and some SM measurements) rely heavily on interference between processes



Interference in Collider Physics

- Searches in EFT (and some SM measurements) rely heavily on interference between processes
- Pure EFT effects suppressed by $\frac{1}{\Lambda^2}$



Interference in Collider Physics

- Searches in EFT (and some SM measurements) rely heavily on interference between processes
- Pure EFT effects suppressed by $\frac{1}{\Lambda^2}$

- **Interference with the SM only suppressed by $\frac{1}{\Lambda}$**
- Need to be sensitive to interference in the EFT limit ($\sigma_{BSM} \rightarrow 0$)



How do We Quantify Sensitivity?



How do We Quantify Sensitivity?

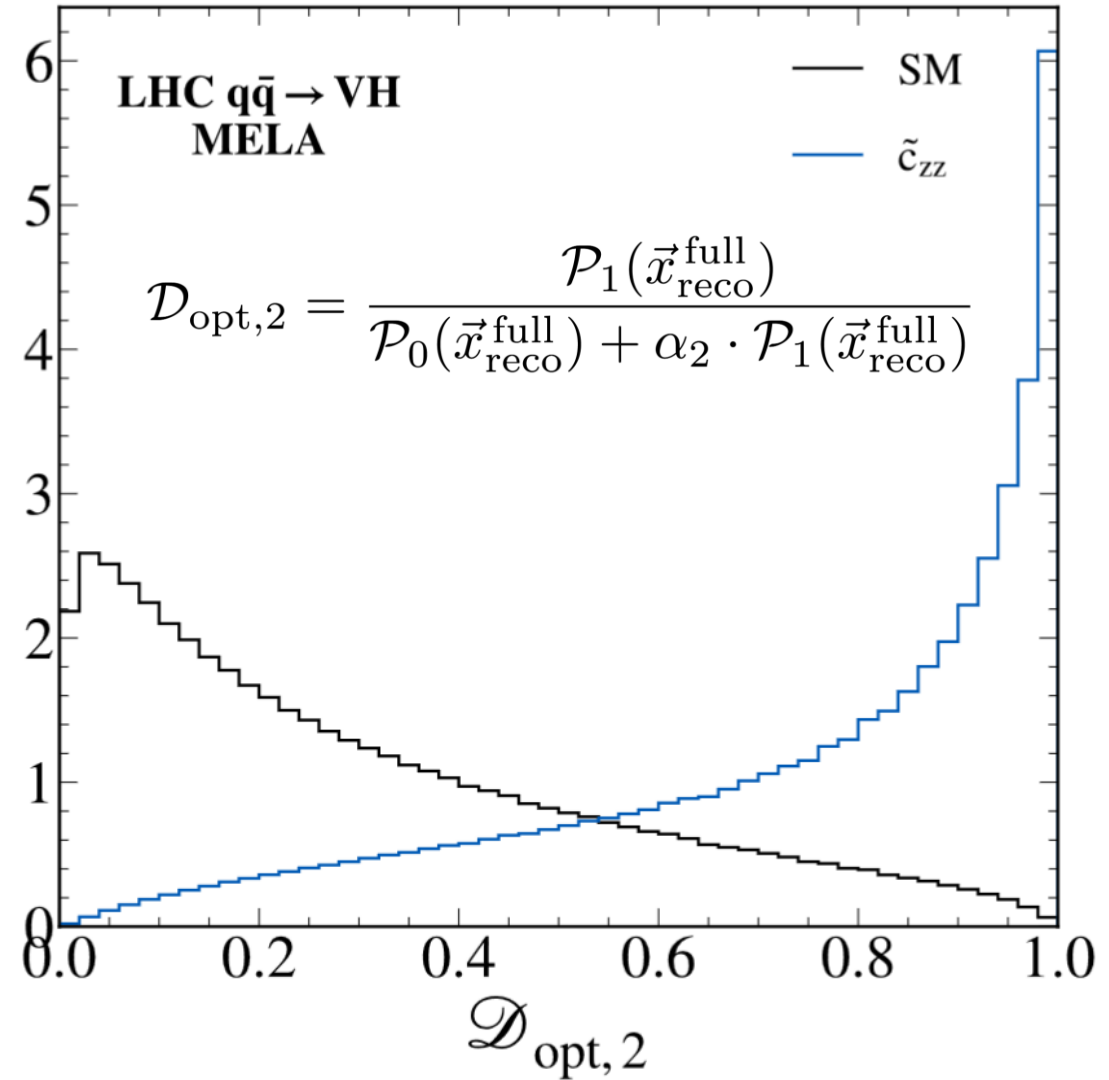
- In 4-lepton, we know that analytic matrix-elements are “optimal” observables to separate hypotheses.
 - Separating “Pure terms”
 - Separating 2 “Pure terms” from interference

$$\mathcal{D}_{\text{opt},2} = \frac{\mathcal{P}_1(\vec{x}_{\text{reco}}^{\text{full}})}{\mathcal{P}_0(\vec{x}_{\text{reco}}^{\text{full}}) + \alpha_2 \cdot \mathcal{P}_1(\vec{x}_{\text{reco}}^{\text{full}})}$$
$$\mathcal{D}_{\text{opt},1}^{(\alpha)} = \frac{2\mathcal{P}_{01}(\vec{x}_{\text{reco}}^{\text{full}})}{\mathcal{P}_0(\vec{x}_{\text{reco}}^{\text{full}}) + \alpha \cdot \mathcal{P}_1(\vec{x}_{\text{reco}}^{\text{full}})}$$



How do We Quantify Sensitivity?

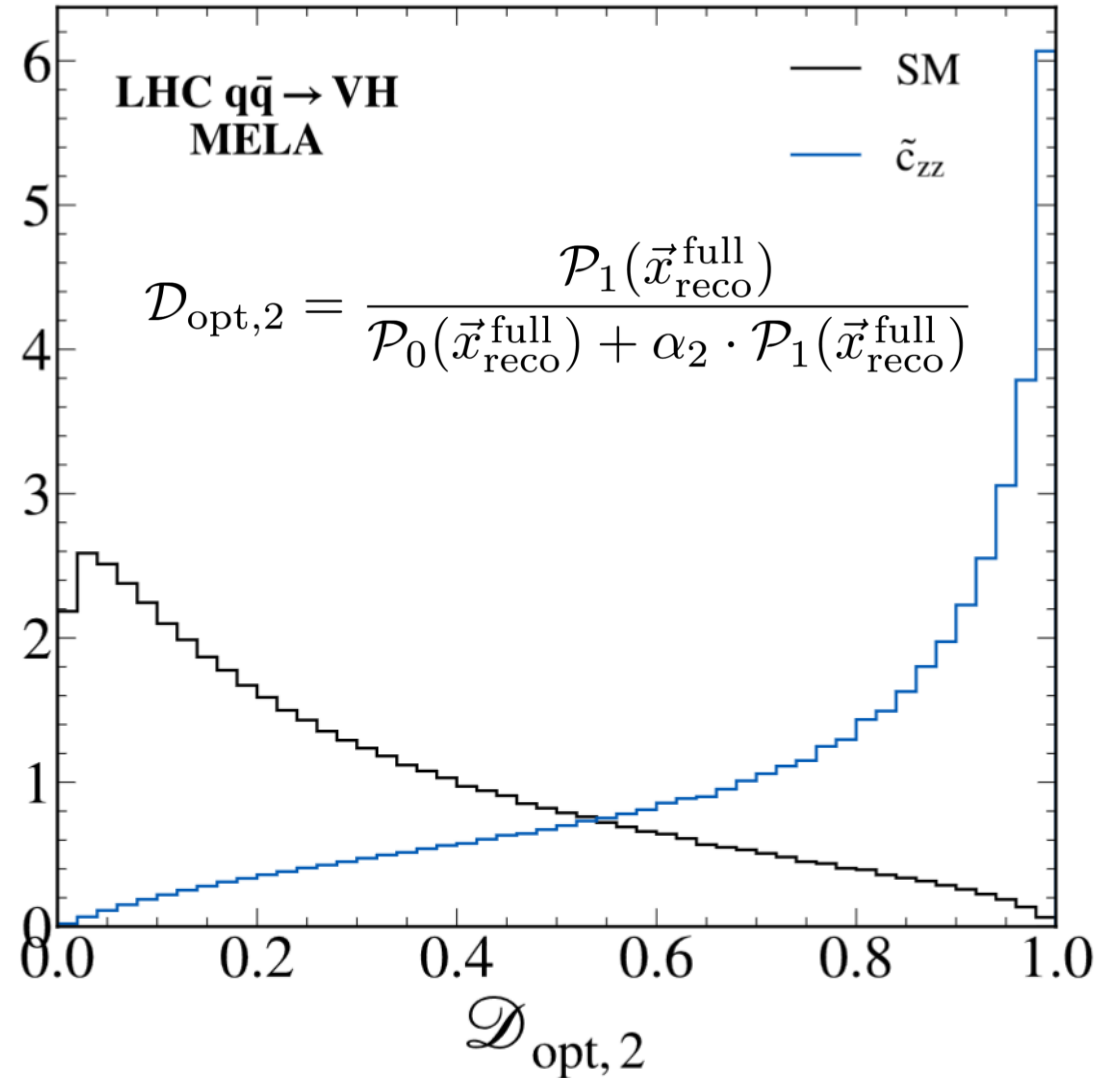
- In 4-lepton, we know that analytic matrix-elements are “optimal” observables to separate hypotheses.
 - Separating “Pure terms”
 - Separating 2 “Pure terms” from interference





How do We Quantify Sensitivity?

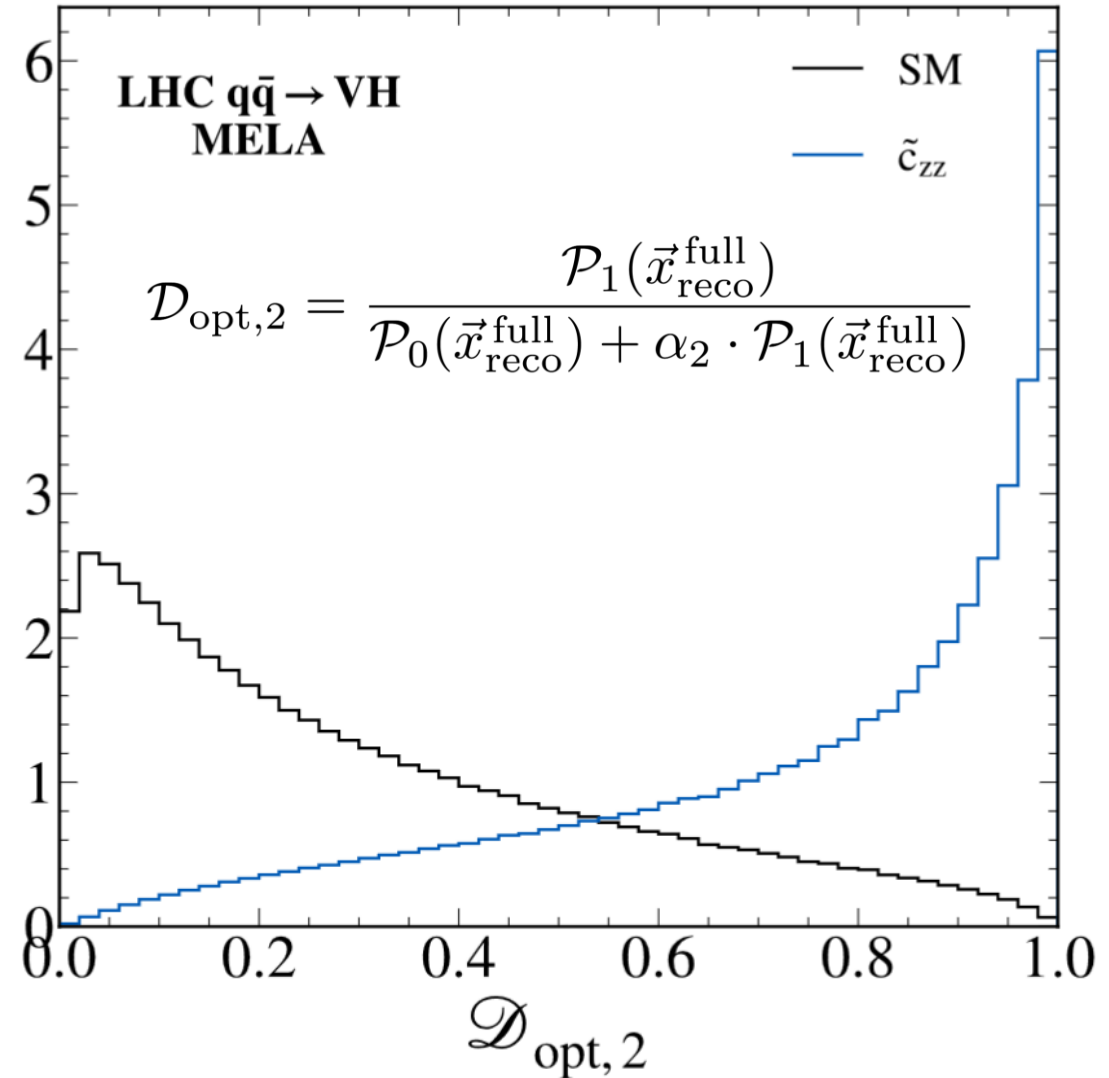
- In 4-lepton, we know that analytic matrix-elements are “optimal” observables to separate hypotheses.
 - Separating “Pure terms”
 - Separating 2 “Pure terms” from interference
- However, these are not feasible for many analyses.
- We need metrics to measure sensitivity to create/compare novel observables





How do We Quantify Sensitivity?

- In 4-lepton, we know that analytic matrix-elements are “optimal” observables to separate hypotheses.
 - Separating “Pure terms”
 - Separating 2 “Pure terms” from interference
- However, these are not feasible for many analyses.
- We need metrics to measure sensitivity to create/compare novel observables
- Traditionally, one would use a ROC curve to measure the strength of a discriminator...

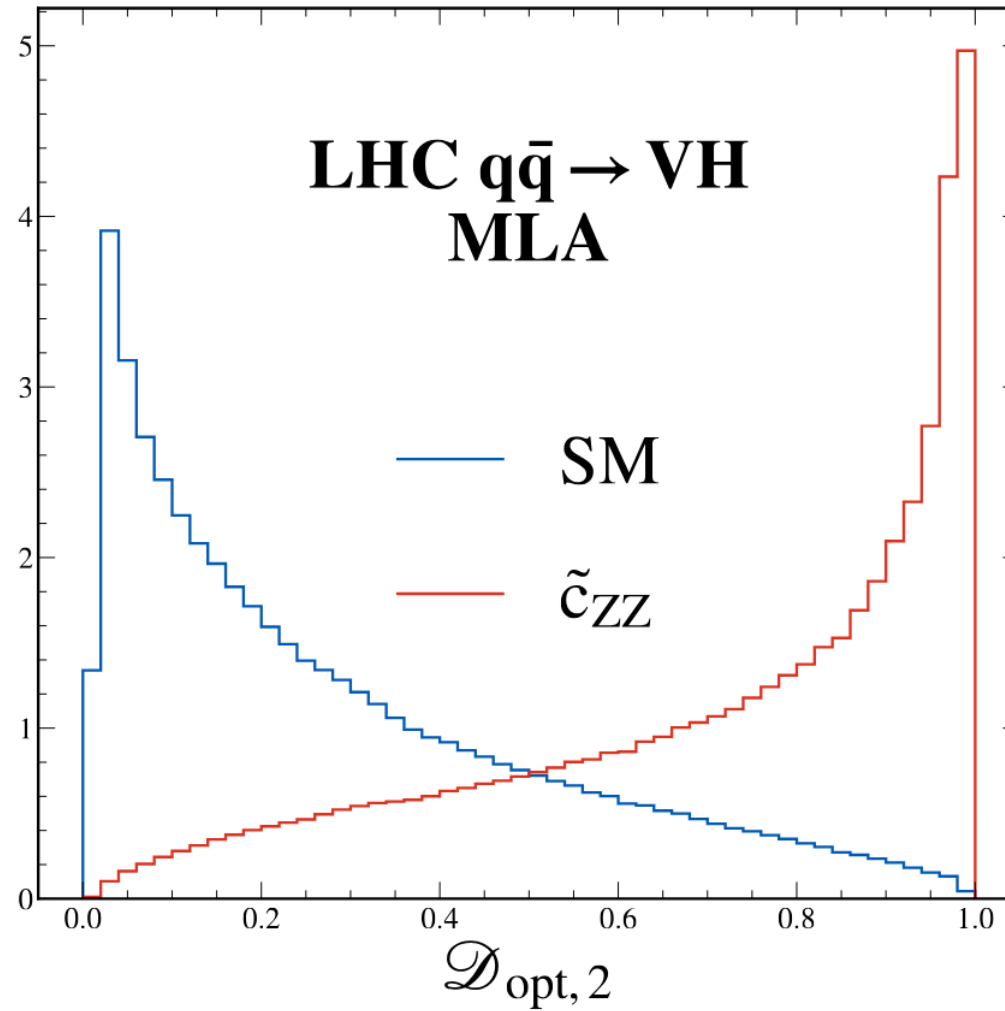




The ROC Curve

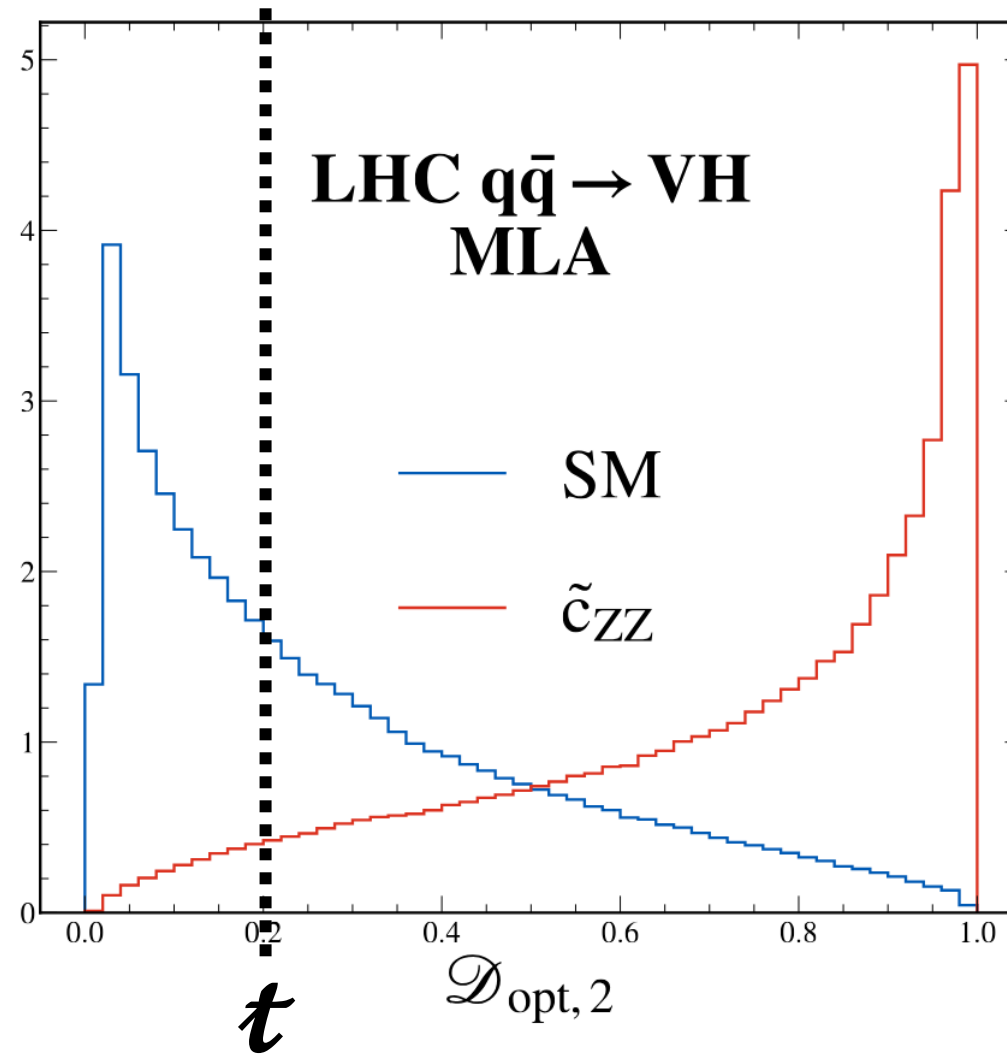


The ROC Curve





The ROC Curve

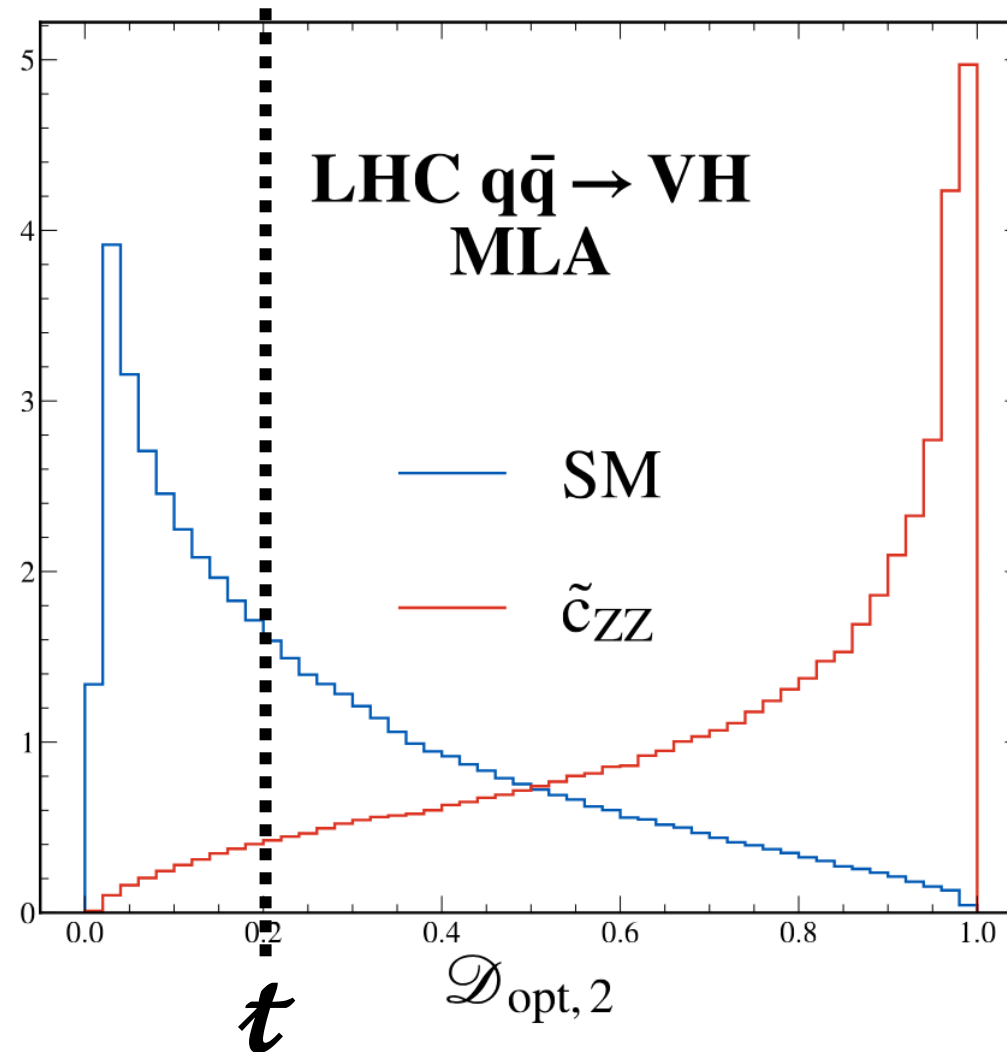




The ROC Curve

$$Y(\beta) = TPR = \frac{1}{|A|} \sum_{x \in A} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

$$X(\beta) = FPR = \frac{1}{|B|} \sum_{x \in B} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

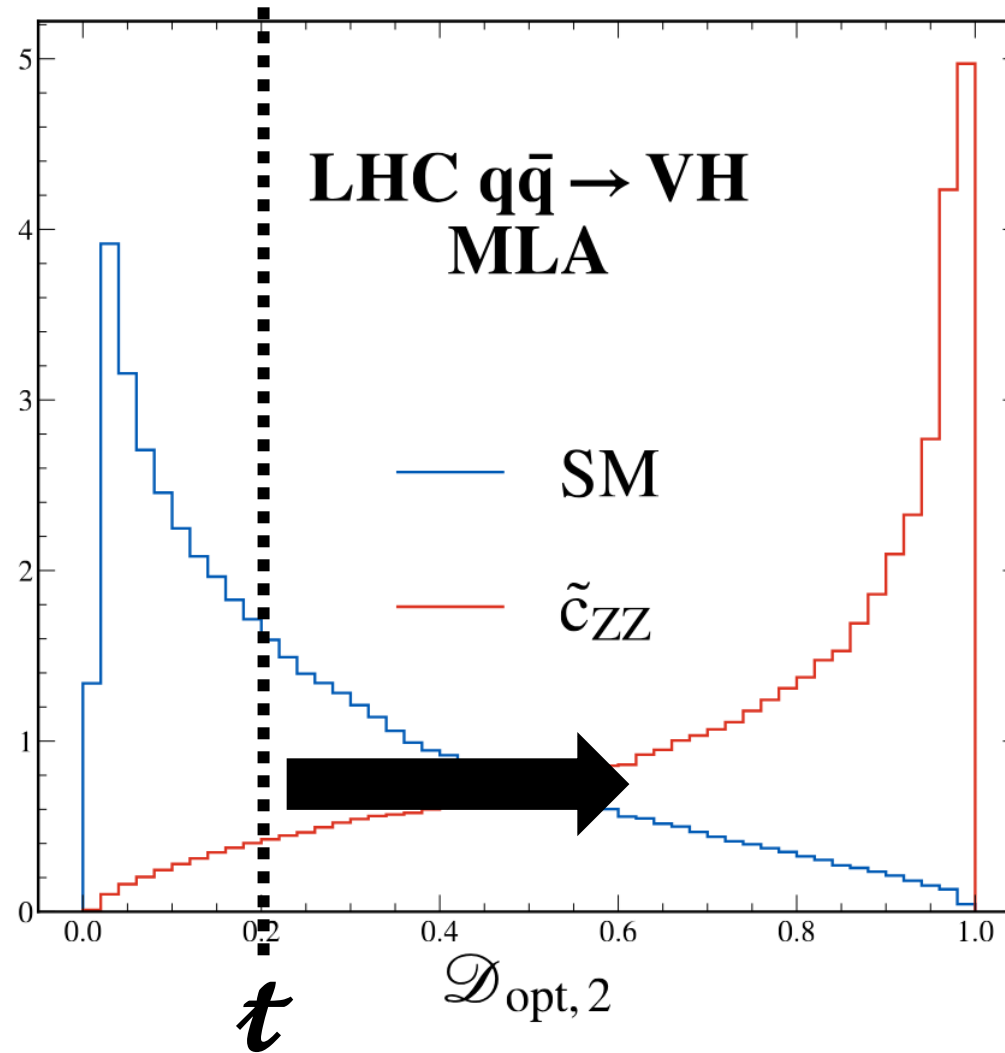




The ROC Curve

$$Y(\beta) = TPR = \frac{1}{|A|} \sum_{x \in A} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

$$X(\beta) = FPR = \frac{1}{|B|} \sum_{x \in B} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$





The ROC Curve

[J. I. Marcum, A Statistical Theory of Target Detection by Pulsed Radar: Mathematical Appendix \(1948\)](#)

[W. Peterson, T. Birdsall and W. Fox, "The theory of signal detectability" \(1954\)](#)

$$Y(\beta) = TPR = \frac{1}{|A|} \sum_{x \in A} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

$$X(\beta) = FPR = \frac{1}{|B|} \sum_{x \in B} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

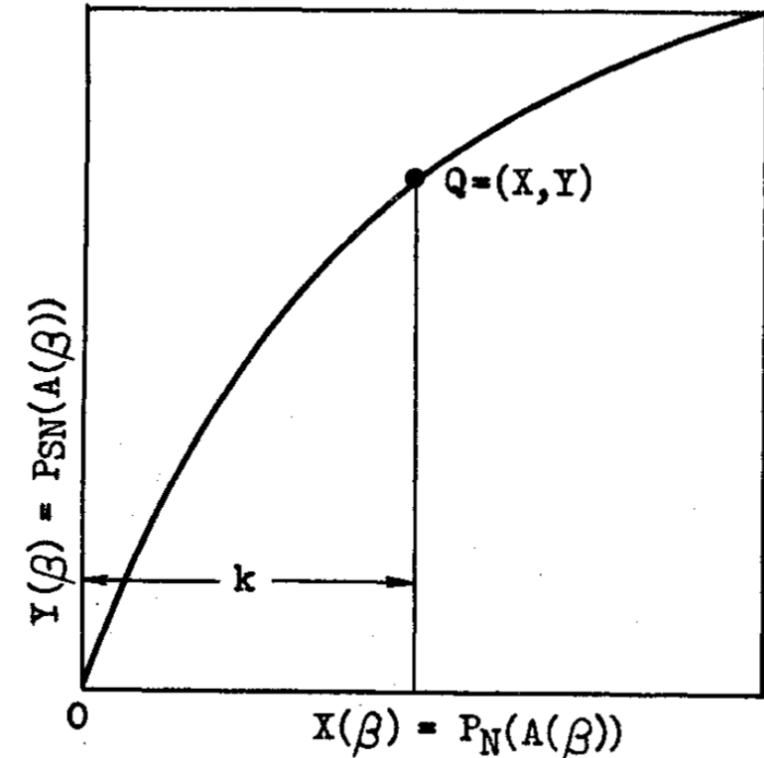


FIG. 1. TYPICAL ROC CURVE



The ROC Curve

[J. I. Marcum, A Statistical Theory of Target Detection by Pulsed Radar: Mathematical Appendix \(1948\)](#)

[W. Peterson, T. Birdsall and W. Fox, "The theory of signal detectability" \(1954\)](#)

$$Y(\beta) = TPR = \frac{1}{|A|} \sum_{x \in A} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

$$X(\beta) = FPR = \frac{1}{|B|} \sum_{x \in B} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

Effectiveness of the observable is then quantified by AUC (area under the curve).

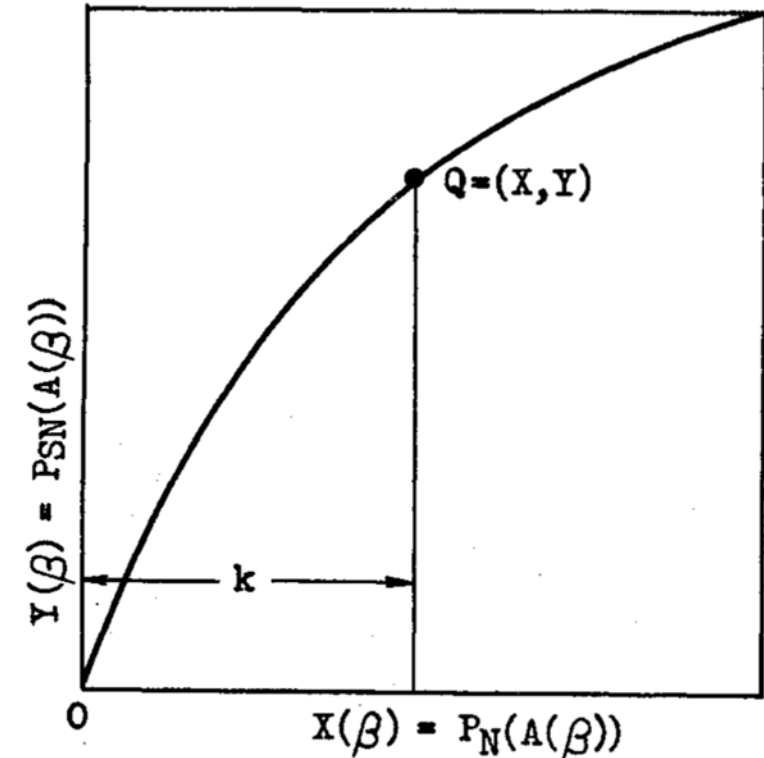


FIG. 1. TYPICAL ROC CURVE



How do We Quantify Sensitivity?



How do We Quantify Sensitivity?

Conventional methods like ROC curves assume that each distribution represents a positive probability distribution



How do We Quantify Sensitivity?

Conventional methods like ROC curves assume that each distribution represents a positive probability distribution

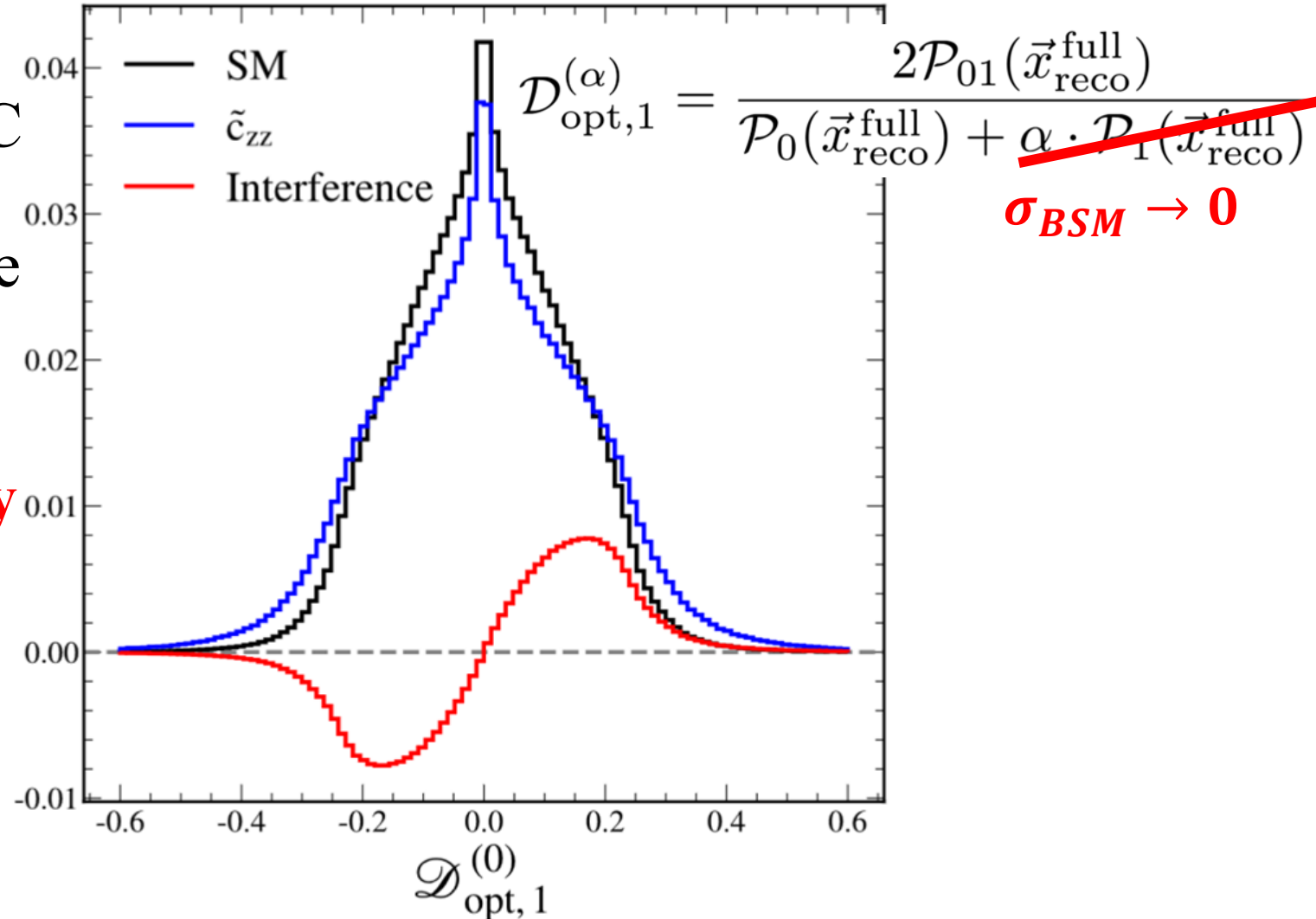
Interference does not abide by that assumption

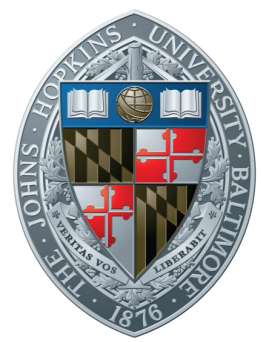


How do We Quantify Sensitivity?

Conventional methods like ROC curves assume that each distribution represents a positive probability distribution

Interference does not abide by that assumption



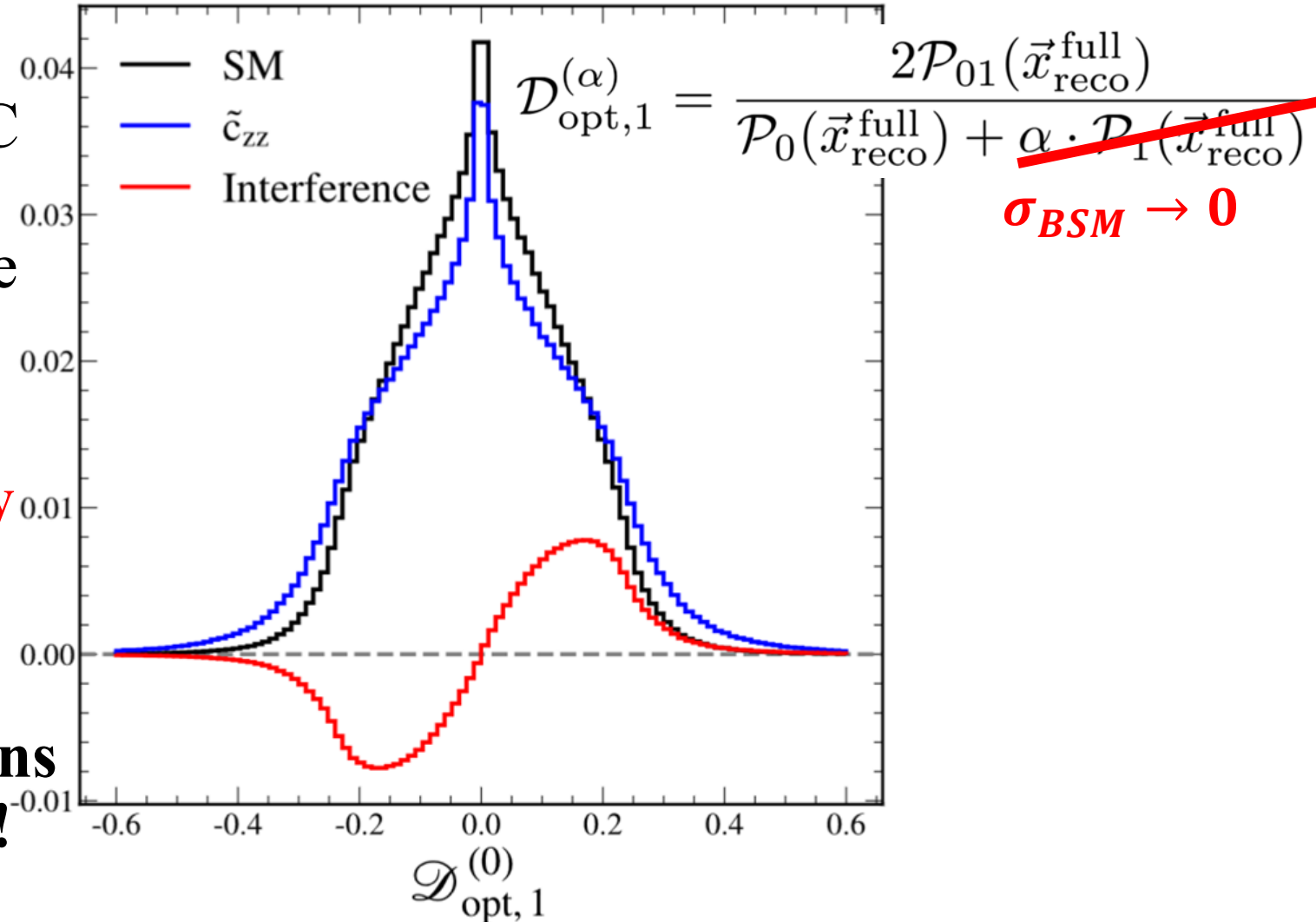


How do We Quantify Sensitivity?

Conventional methods like ROC curves assume that each distribution represents a positive probability distribution

Interference does not abide by that assumption

We need to separate distributions for analysis, not classify them!





How do We Quantify Sensitivity?



How do We Quantify Sensitivity?

Sensitivity \approx Separability



How do We Quantify Sensitivity?

Sensitivity \approx Separability

Separability \neq Classification Power



How do We Quantify Sensitivity?

Sensitivity \approx Separability

Separability \neq Classification Power

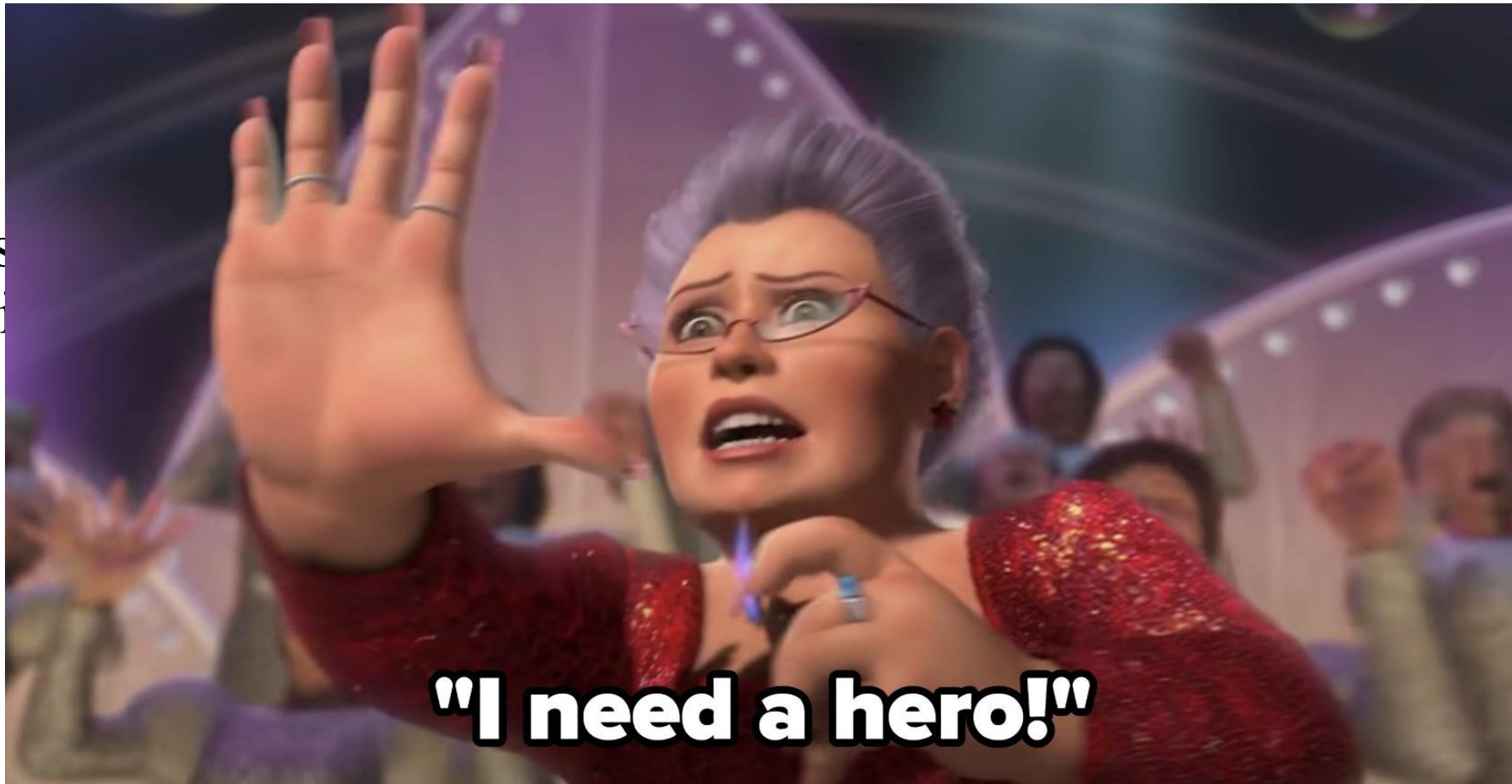
We need something that is **easy to calculate** from a histogram that quantifies a “distance” between 2 distributions akin to a ΔNLL in an analysis



How do We Quantify Sensitivity?

We need s
“d

ntifies a
S



"I need a hero!"



LOC Curves!



LOC Curves!

Likelihood Operating Characteristic



LOC Curves!

Likelihood Operating Characteristic

- Instead of AUC, use the length of the curve itself (LOC)



LOC Curves!

Likelihood Operating Characteristic

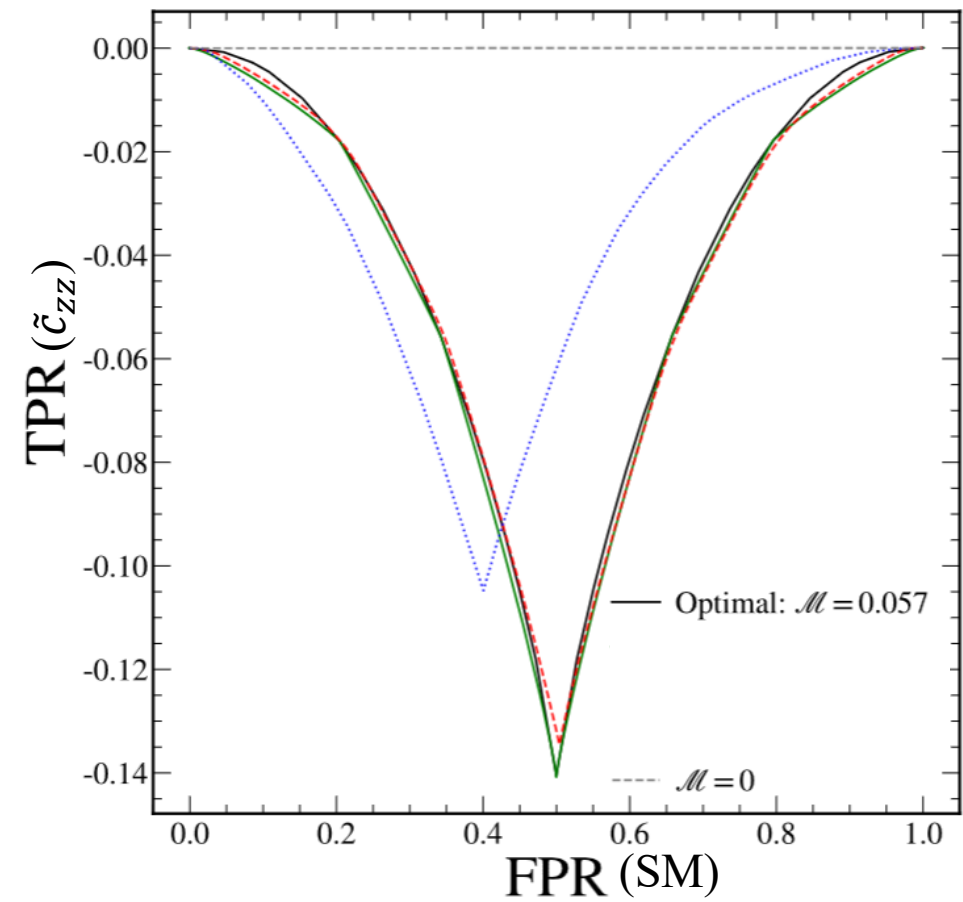
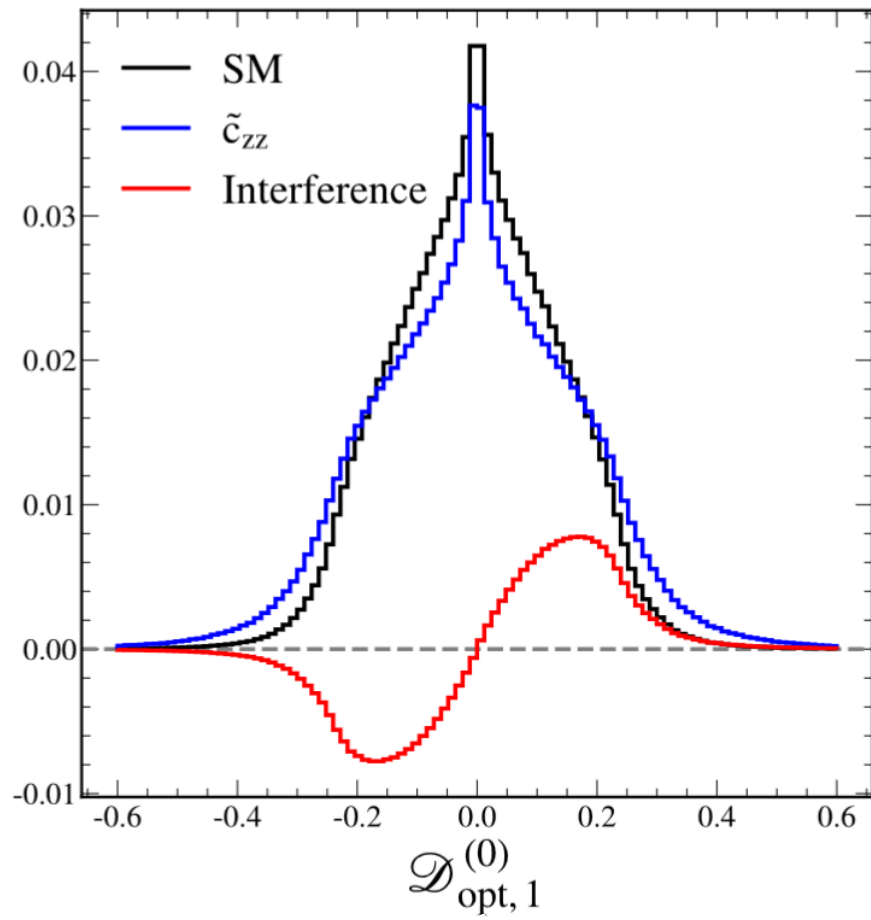
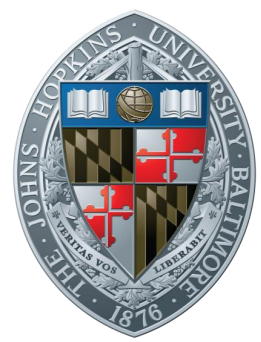
- Instead of AUC, use the length of the curve itself (LOC)
- Constructed the same as a ROC curve, but now hypothesis B is now normalized relative to hypothesis A
- As LOC goes up, so does ΔNLL

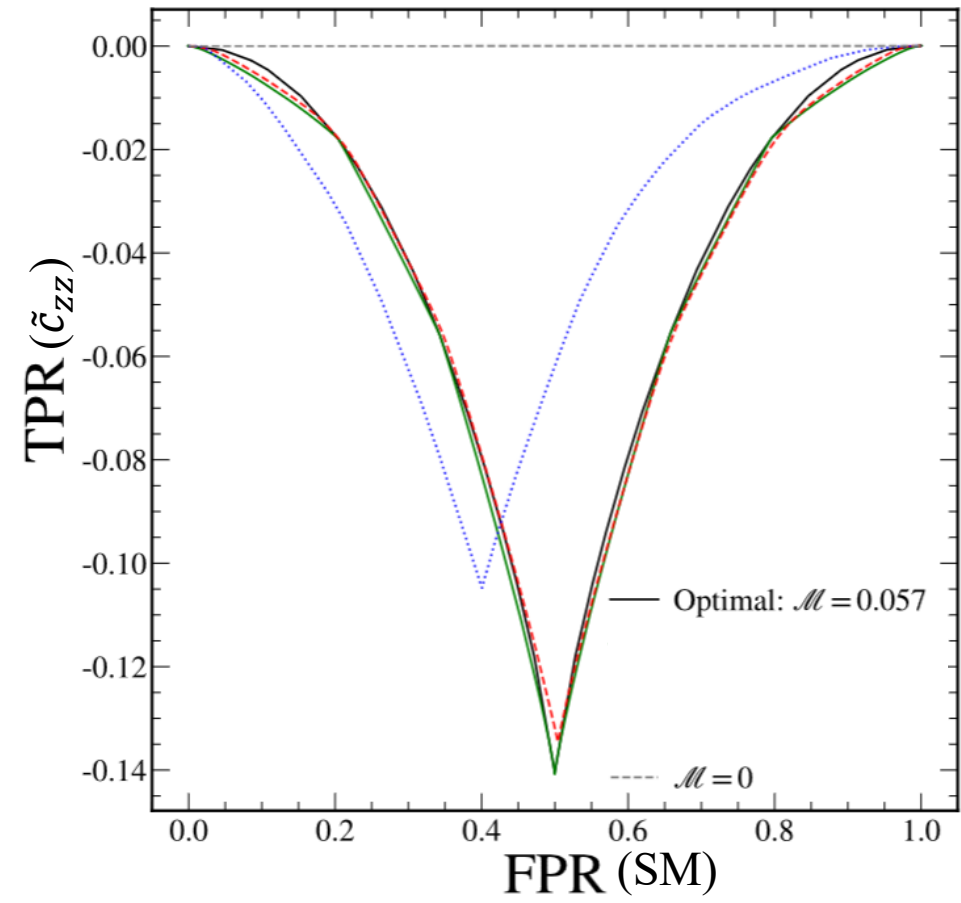
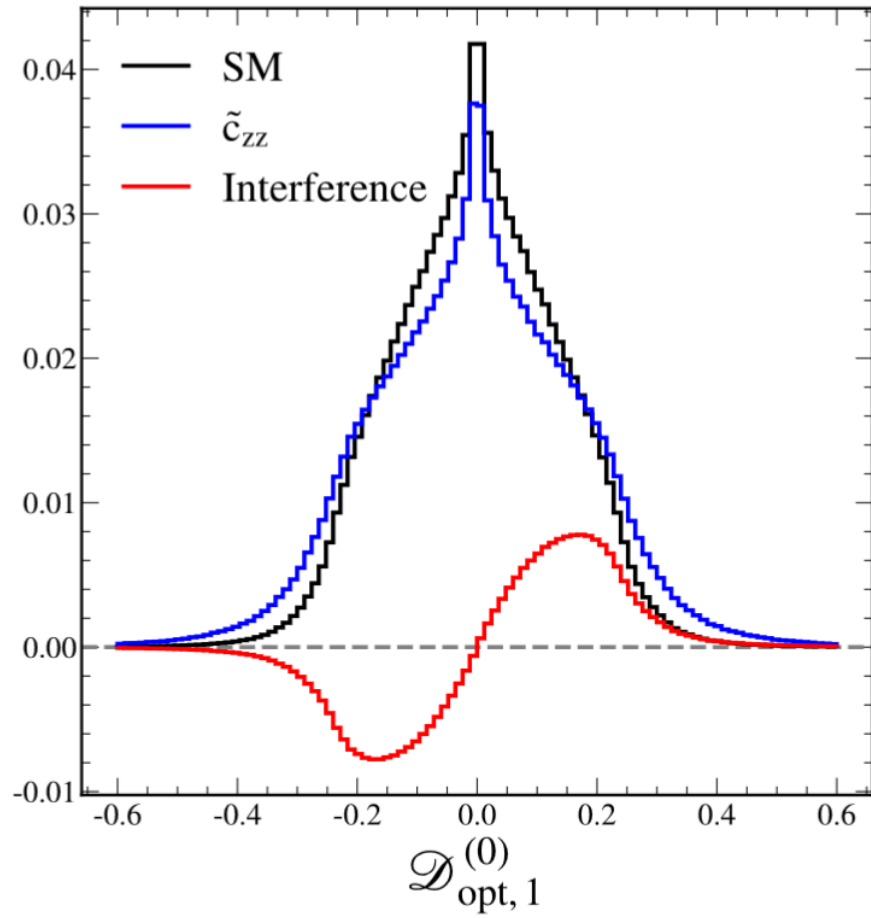
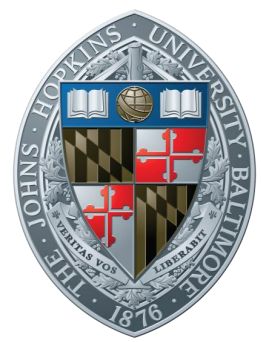
LOC Curves!

Operating Characteristic

- Instead of ΔNLL , the curve itself (LOC)
- Constructed the same as ΔNLL , but hypothesis B is now normalized relative to hypothesis A
- As LOC goes up, so does ΔNLL

“0 separation” is now the length of a straight line!





- We can now compare the sensitivity one would get in an analysis for different distributions
- That got us thinking...

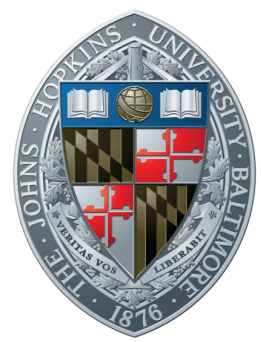


Finding an Optimal Solution



Finding an Optimal Solution

- Having too many bins makes fits unstable
- Having too few makes fit sensitivity go down



Finding an Optimal Solution

- Having too many bins makes fits unstable
- Having too few makes fit sensitivity go down
- What if we could merge bins together to ensure fit stability in a way that minimizes loss of sensitivity?
 - Start from a fine binning and merge down to allow for non-trivial shapes in phasespace



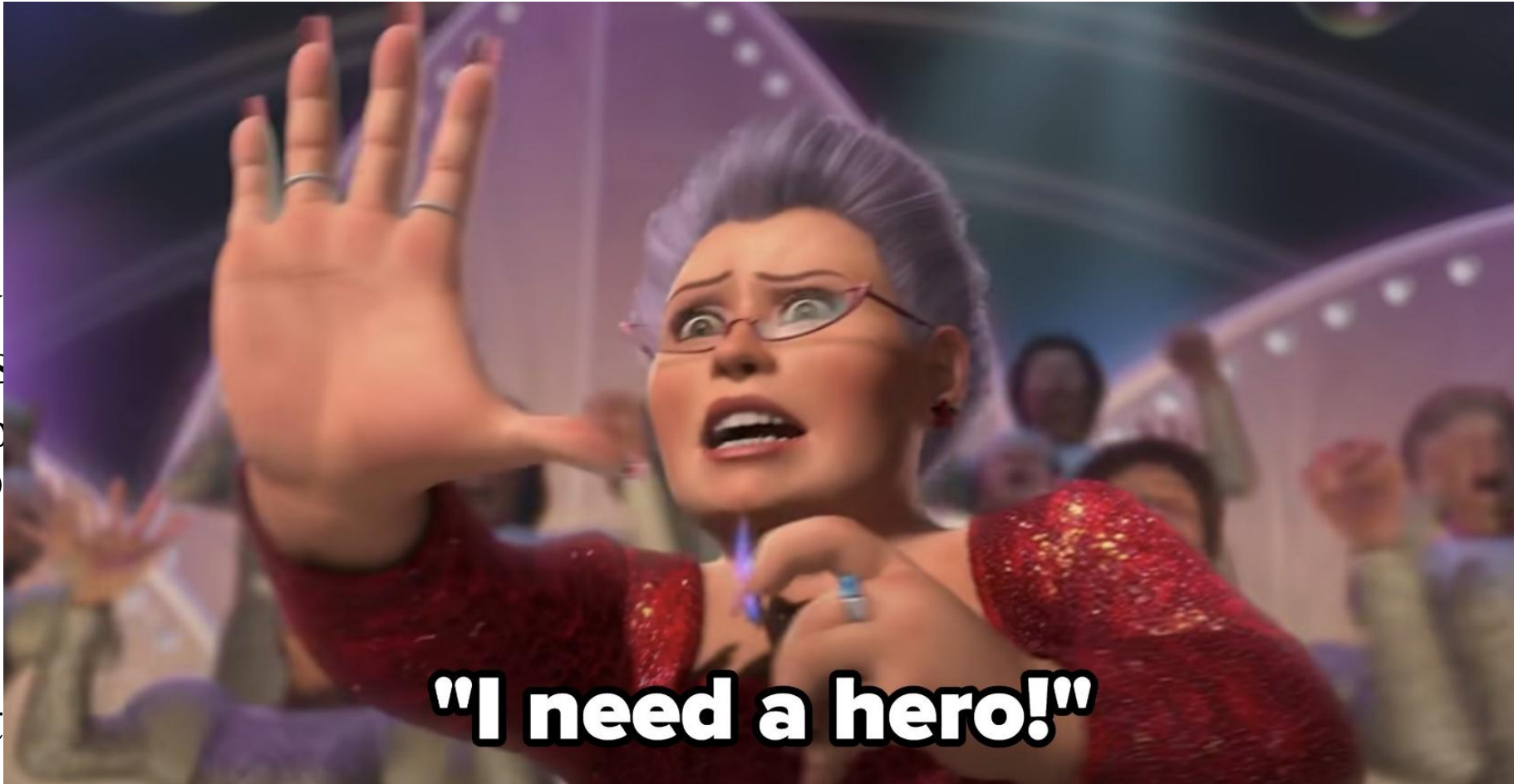
Finding an Optimal Solution

- Having too many bins makes fits unstable
- Having too few makes fit sensitivity go down
- What if we could merge bins together to ensure fit stability in a way that minimizes loss of sensitivity?
 - Start from a fine binning and merge down to allow for non-trivial shapes in phasespace
- **Naively, this would take $\mathcal{O}(2^n)$ time**



Finding an Optimal Solution

- Having to
- Having to
- What if we
minimizes
 - Start from
phasesp
- Naively, t





The MiLoMerge Framework



The MiLoMerge Framework

1. The length of the curve (LOC) is order-invariant, unlike the AUC



The MiLoMerge Framework

1. The length of the curve (LOC) is order-invariant, unlike the AUC
2. Minimizing the loss of length minimizes the loss of performance per bin-merge



The MiLoMerge Framework

1. The length of the curve (LOC) is order-invariant, unlike the AUC
2. Minimizing the loss of length minimizes the loss of performance per bin-merge
3. One can then craft a “merging score” for a pair of bins i, j that showcases the separating power of that bin
 - Done iteratively, takes $O(N^2)$ time

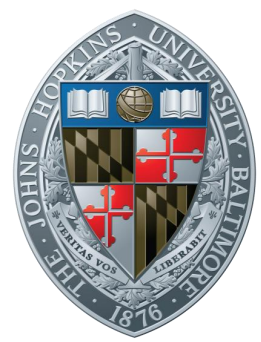


MiLoMerge Merging Score



MiLoMerge Merging Score

$$\mathcal{D}_{ij} = \sum_h \sum_{h' > h} w_{h'h} \left(H_i^{(h)} H_j^{(h')} - H_i^{(h')} H_j^{(h)} \right)^2$$



MiLoMerge Merging Score

$$D_{ij} = \sum_h \sum_{h' > h} w_{h'h} \left(H_i^{(h)} H_j^{(h')} - H_i^{(h')} H_j^{(h)} \right)^2$$

- $H_i^{(h)}$ represents a bin from index i for hypothesis h
- $w_{h'h}$ represents a weight applied between the 2 hypotheses
- Take the bins with the smallest value of D_{ij} and merge them iteratively



MiLoMerge Merging Score

$$D_{ij} = \sum_h \sum_{h' > h} w_{h'h} \left(H_i^{(h)} H_j^{(h')} - H_i^{(h')} H_j^{(h)} \right)^2$$

- $H_i^{(h)}$ represents a bin from index i for hypothesis h
- $w_{h'h}$ represents a weight applied between the 2 hypotheses
- Take the bins with the smallest value of D_{ij} and merge them iteratively

Now the time taken when storing values in memory is:

$$N^2 + \sum_{i=1}^{i=k \leq N-2} (N - i) \rightarrow O(N^2)$$

MiLoMerge Merging Score

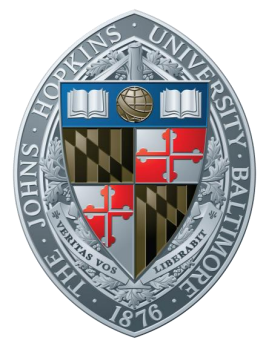
$$\sum_h \sum_{h'} w_{h'h} \left(H_i^{(h)} H_j^{(h')} - H_i^{(h')} H_j^{(h)} \right)^2$$

Let's see this in action!

- $w_{h'h}$ for hypothesis h
- Take the bins with the 2 hypotheses
- Merge them iteratively

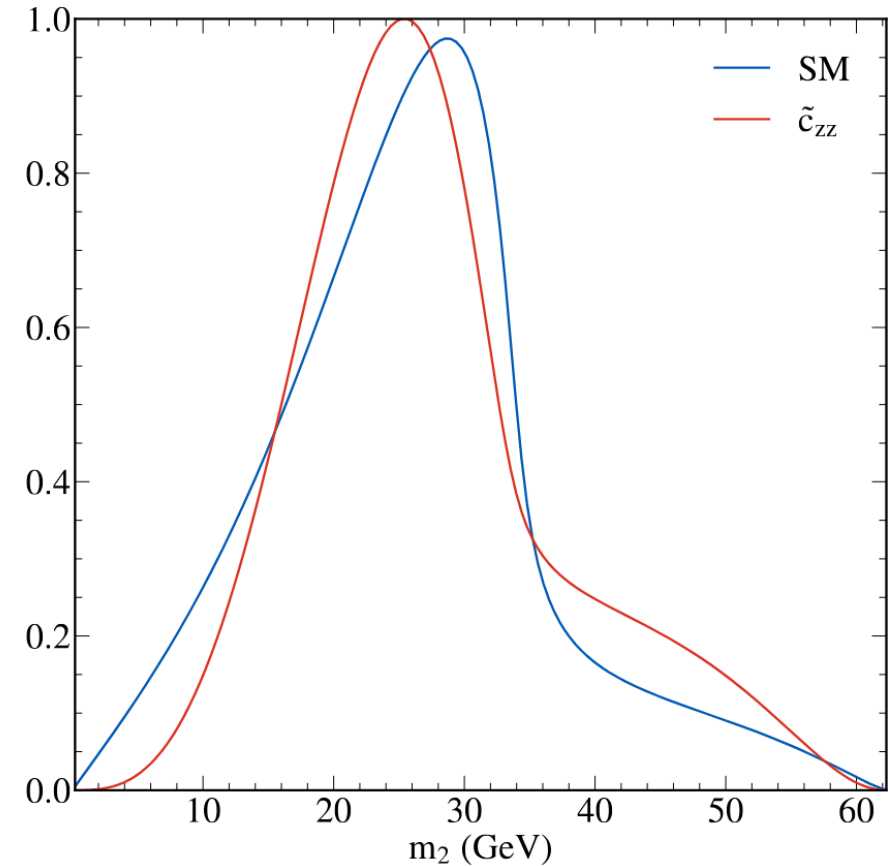
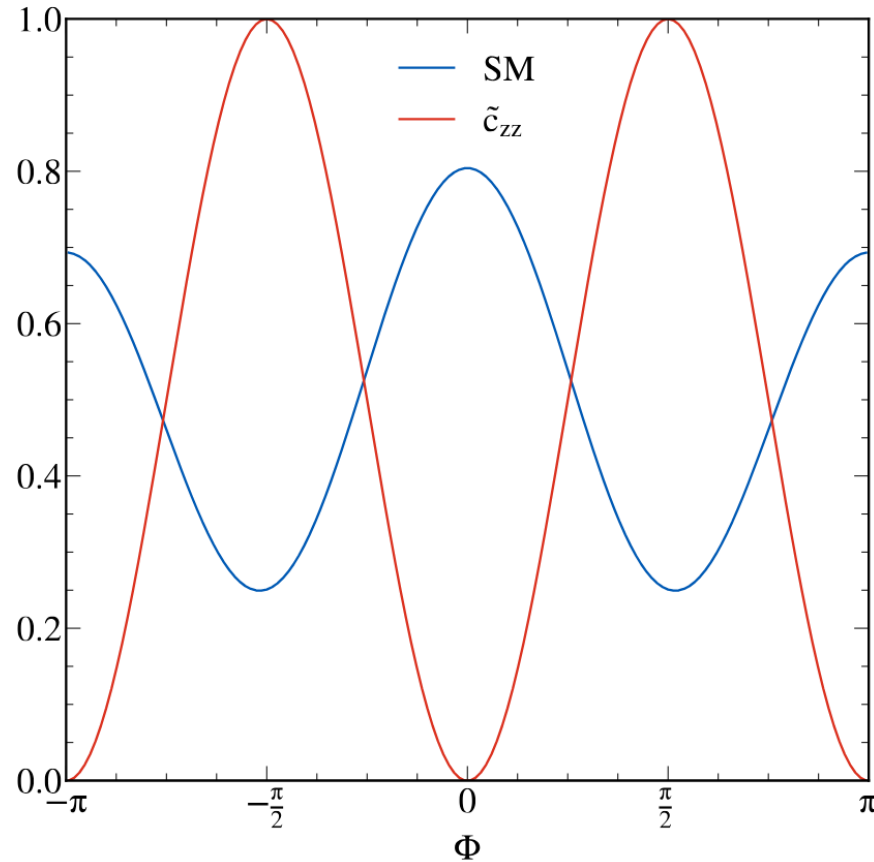
Now the time taken when storing values

$$N^2 + \sum_{i=1}^{i=k \leq N-2} (N - i) \rightarrow O(N^2)$$



The MiLoMerge Framework

- Take distributions of m_2 and ϕ , two observables in $H \rightarrow 2e2\mu$ decays
- Separating SM and CP-odd BSM hypotheses of equal size





The MiLoMerge Framework

Integrating out
all other
dependence,
one can craft
an analytic
discriminant

$$\mathcal{D}(m_2, \phi)$$

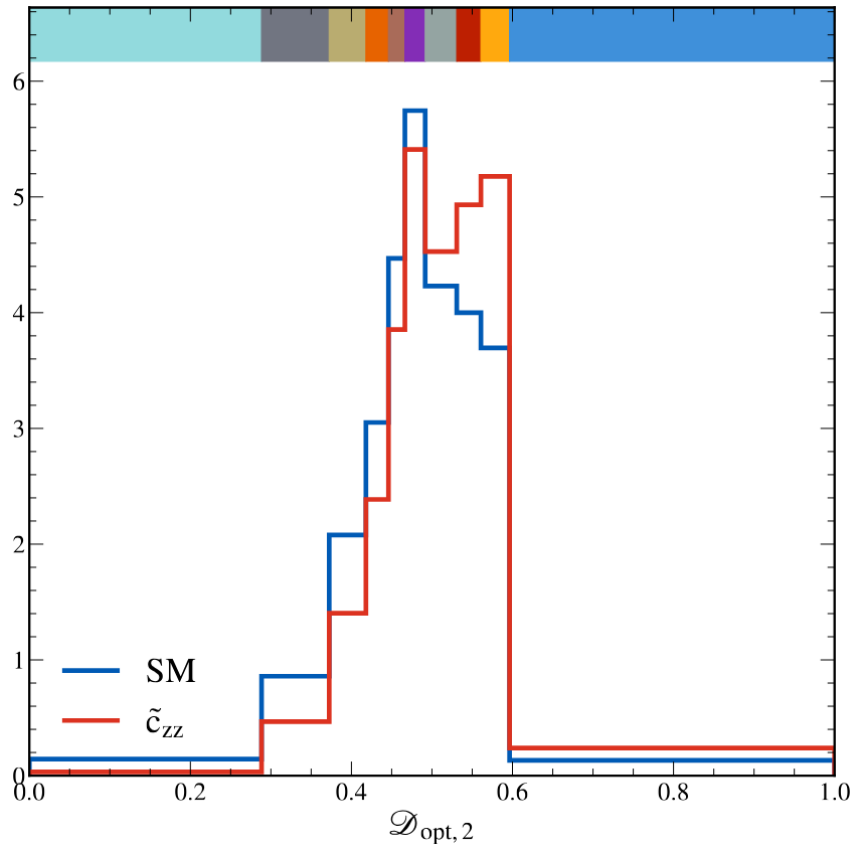
[On the spin and parity of a single-produced resonance at the LHC](#) (2014)



The MiLoMerge Framework

Integrating out all other dependence, one can craft an analytic discriminant

$$\mathcal{D}(m_2, \phi)$$



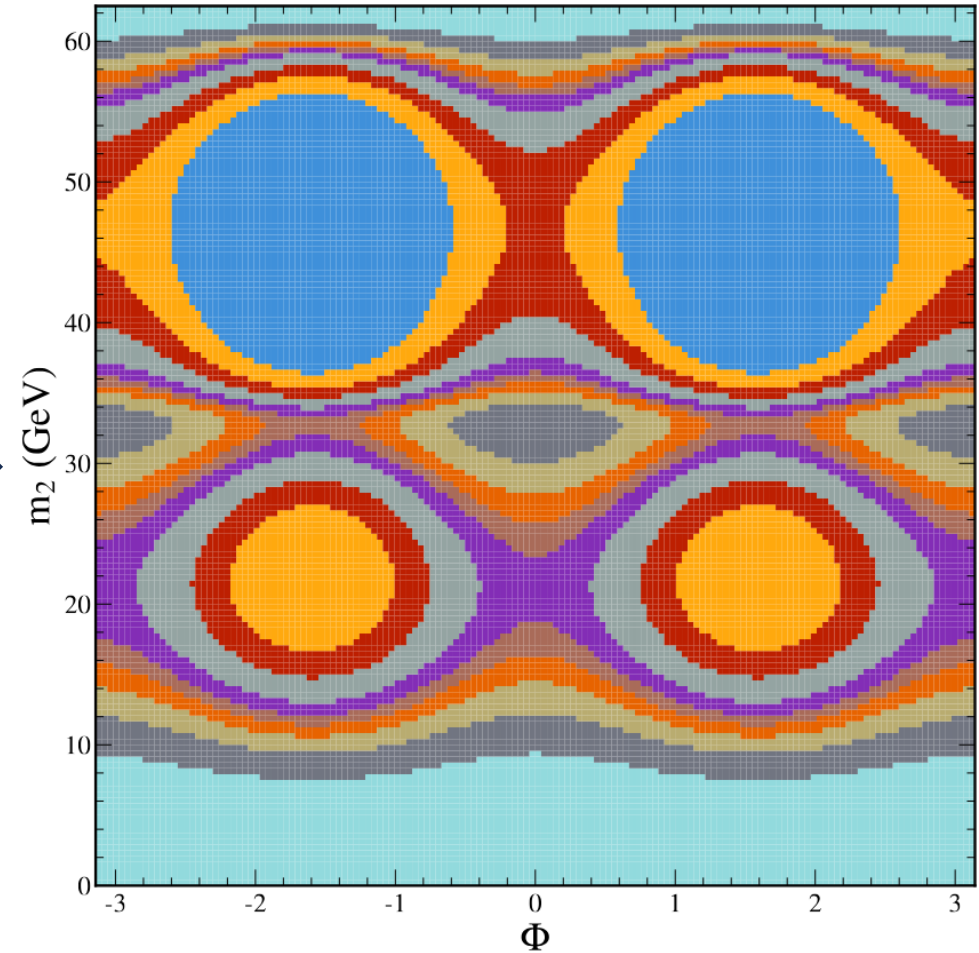
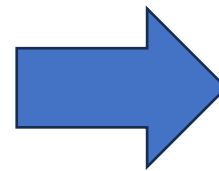
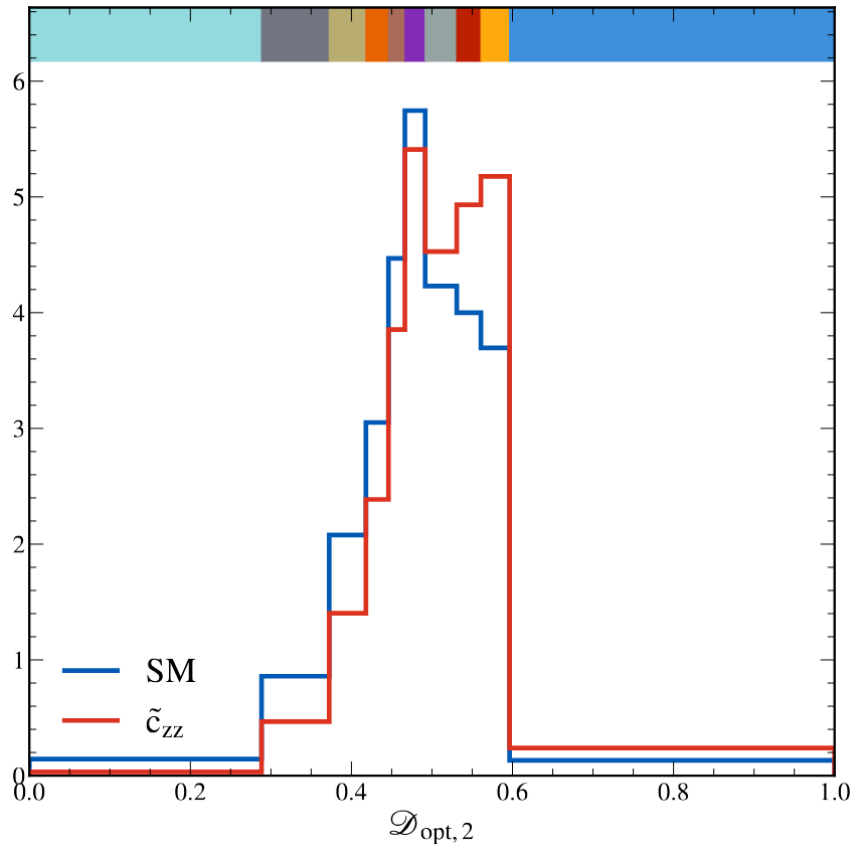
[On the spin and parity of a single-produced resonance at the LHC \(2014\)](#)



The MiLoMerge Framework

Integrating out all other dependence, one can craft an analytic discriminant

$$\mathcal{D}(m_2, \phi)$$



[On the spin and parity of a single-produced resonance at the LHC \(2014\)](#)



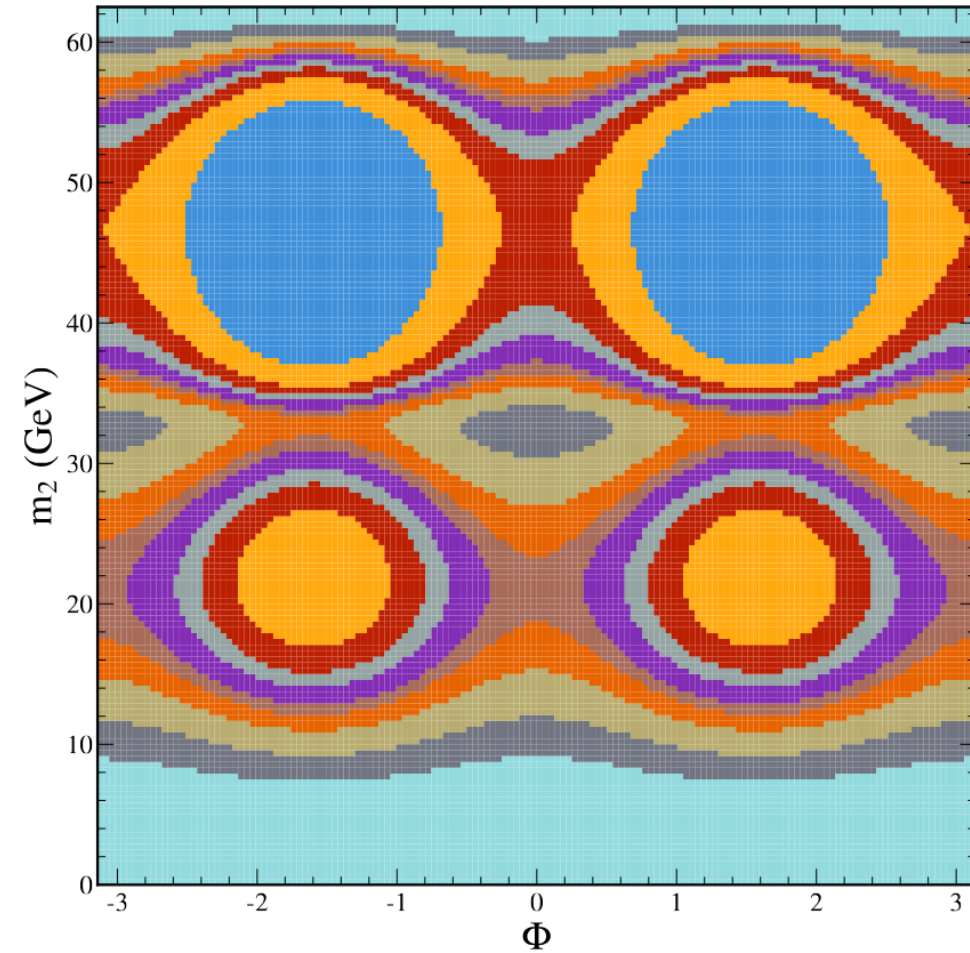
The MiLoMerge Framework

Now run the
bin-merging
procedure and
compare



The MiLoMerge Framework

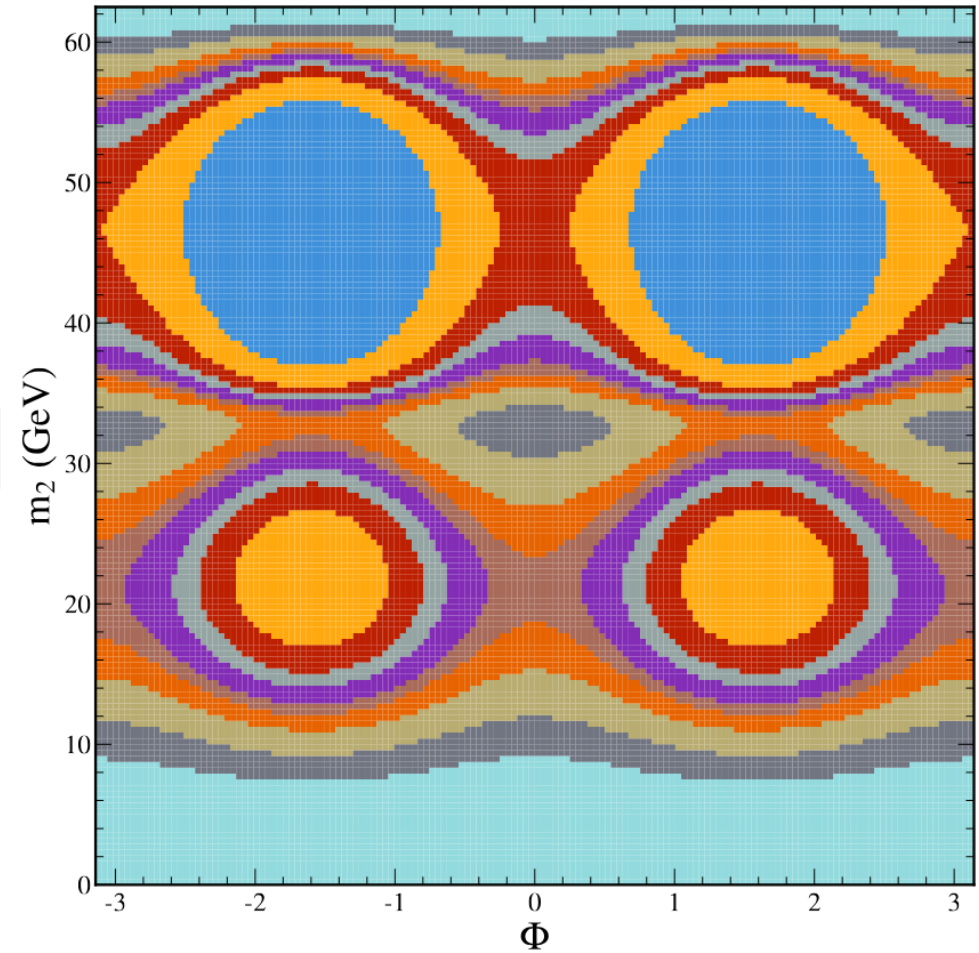
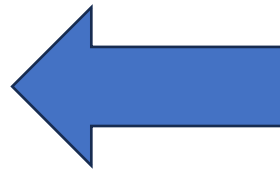
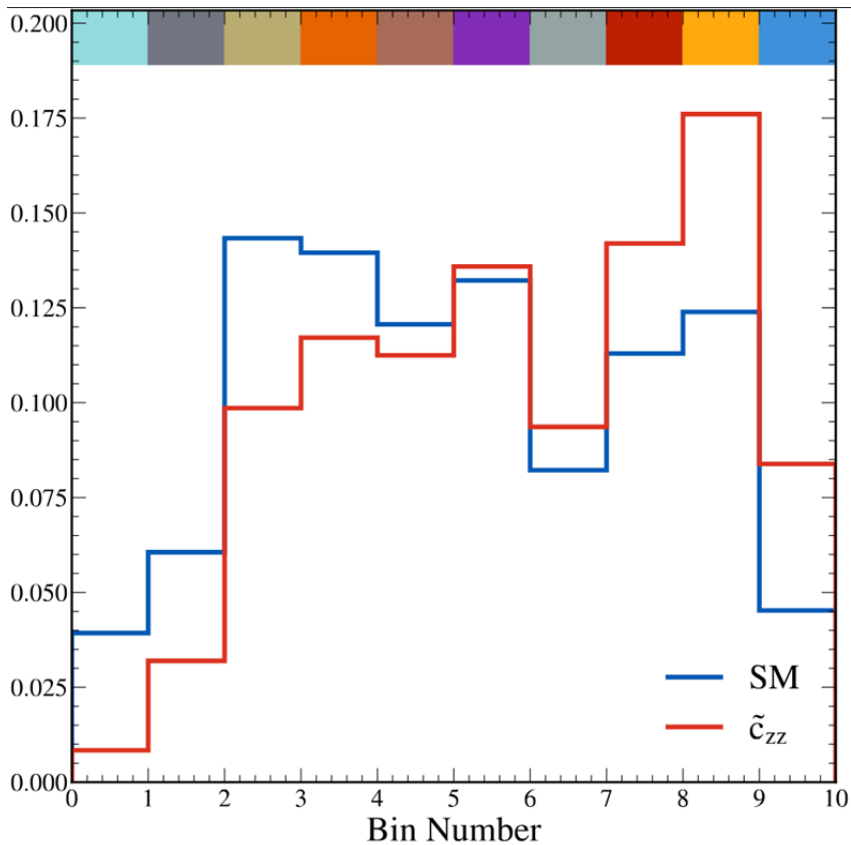
Now run the
bin-merging
procedure and
compare





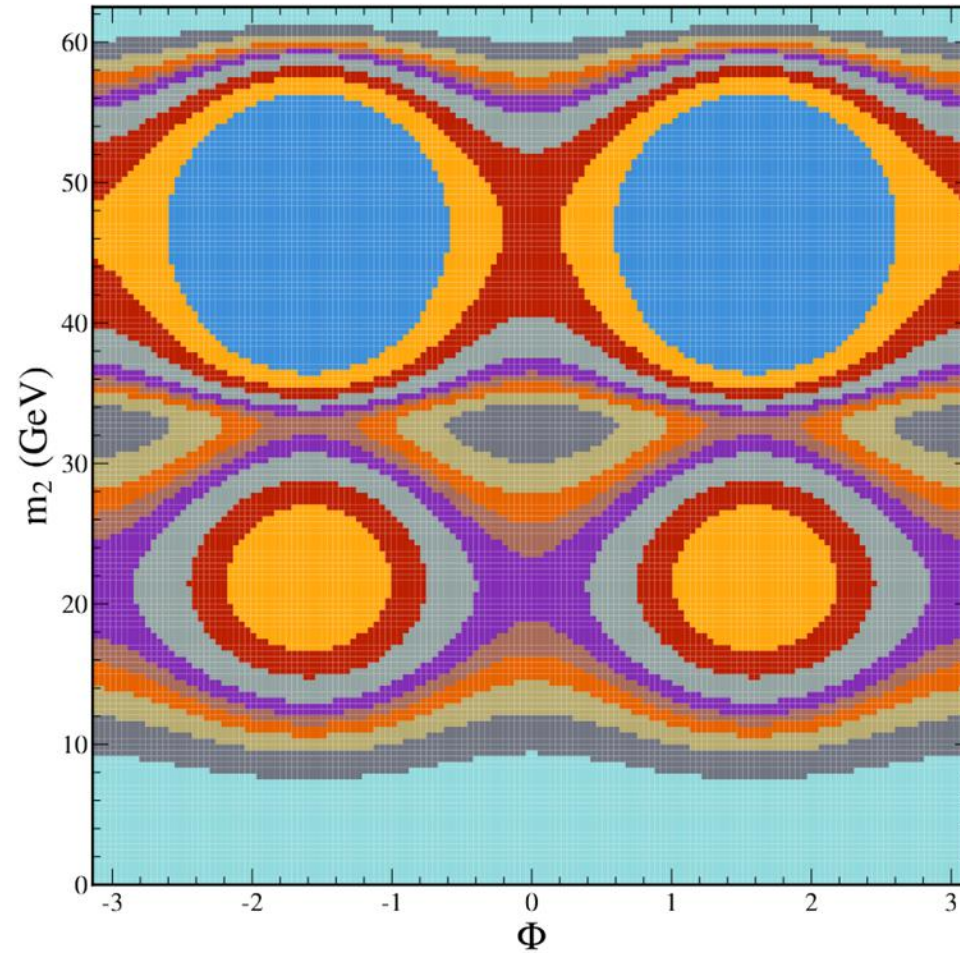
The MiLoMerge Framework

Now run the bin-merging procedure and compare



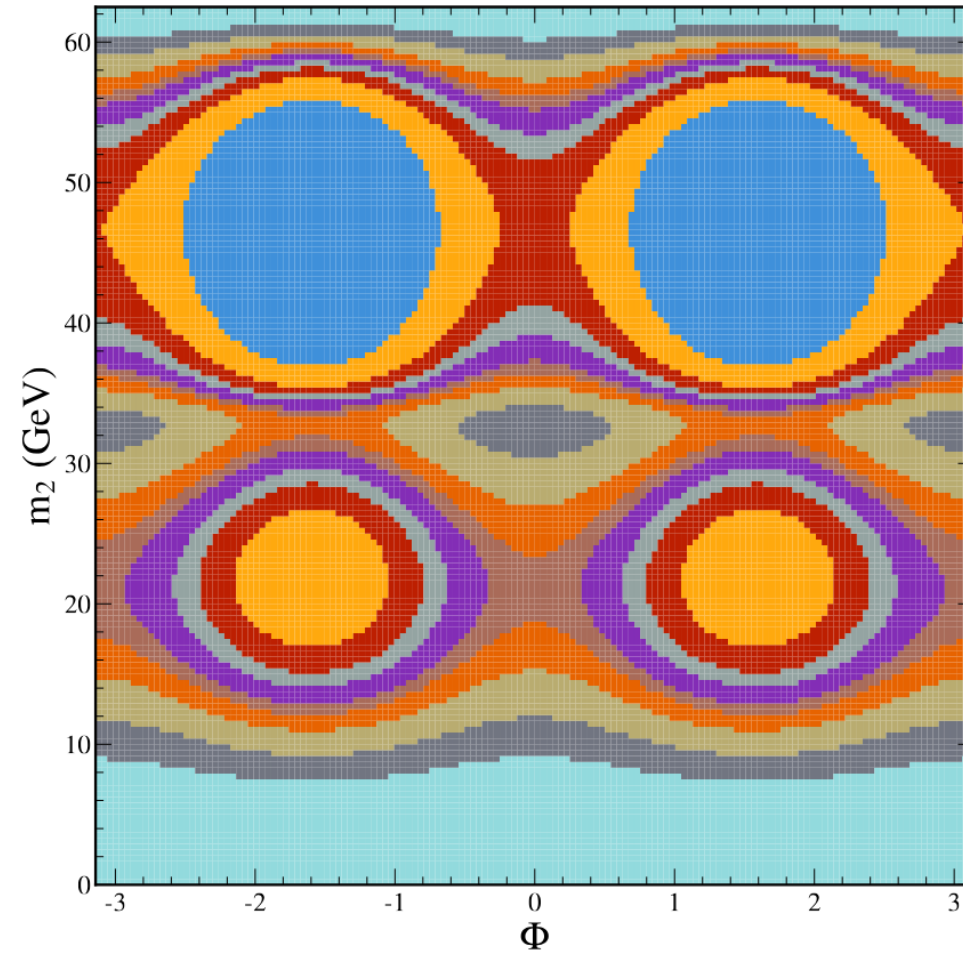


Flip-o-Rama



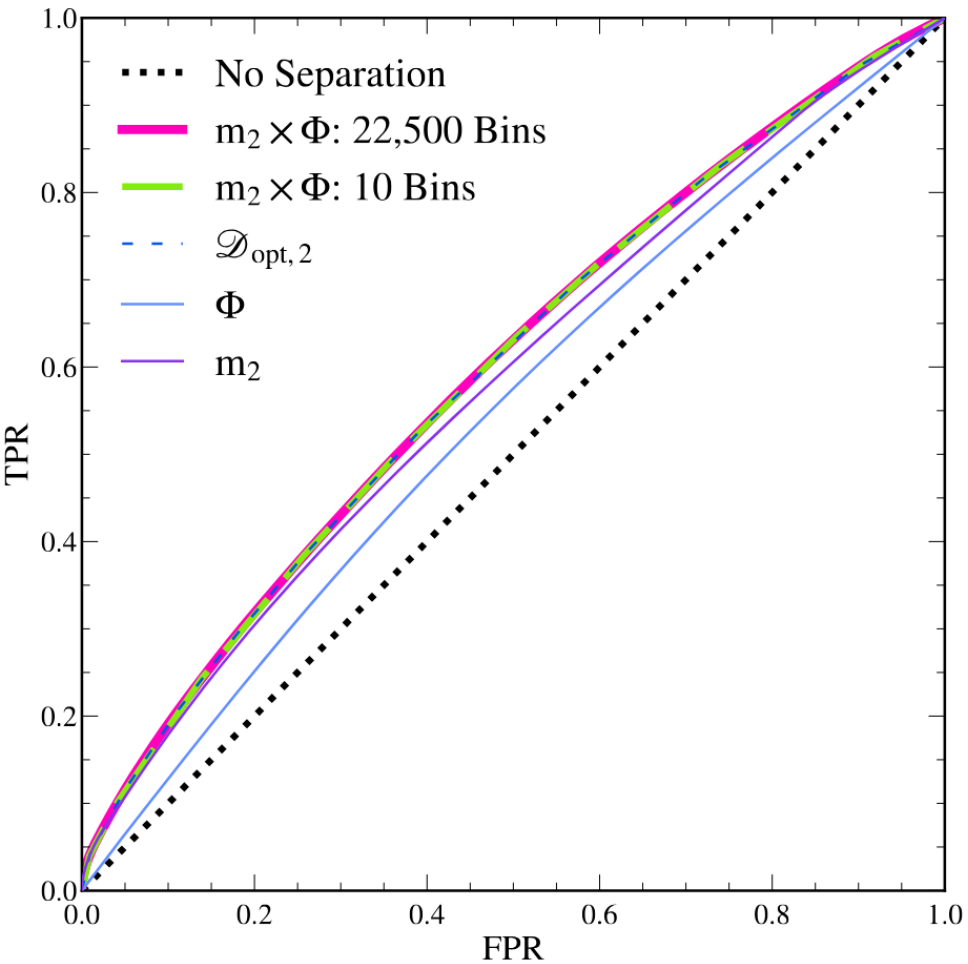


Flip-o-Rama



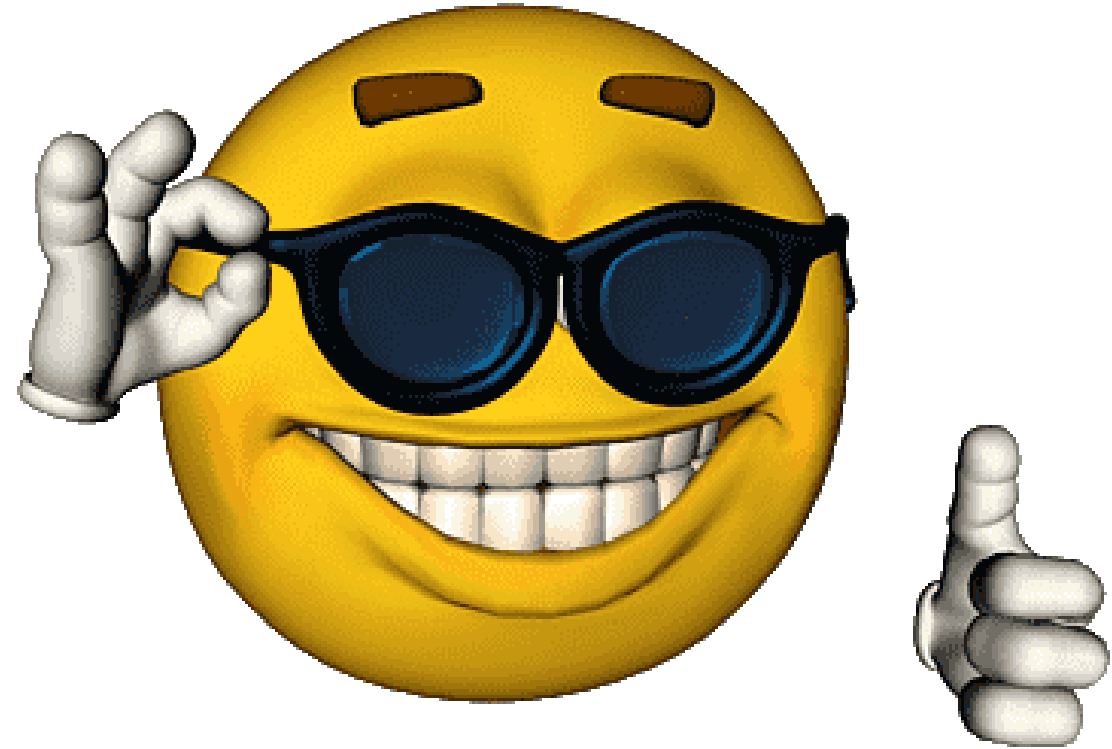
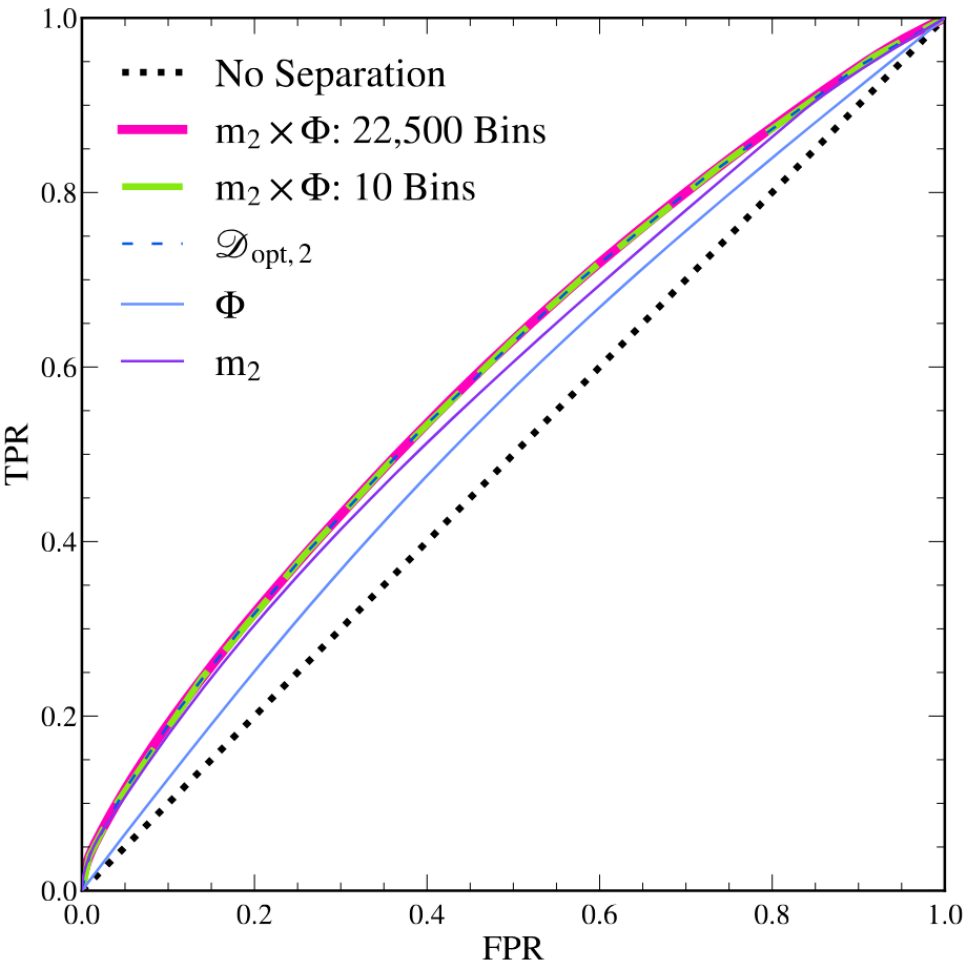


The MiLoMerge Framework





The MiLoMerge Framework





But Wait, There's More



But Wait, There's More

- The curve generated using the LoC Curve convention is not limited to pairwise comparisons!



But Wait, There's More

- The curve generated using the LoC Curve convention is not limited to pairwise comparisons!

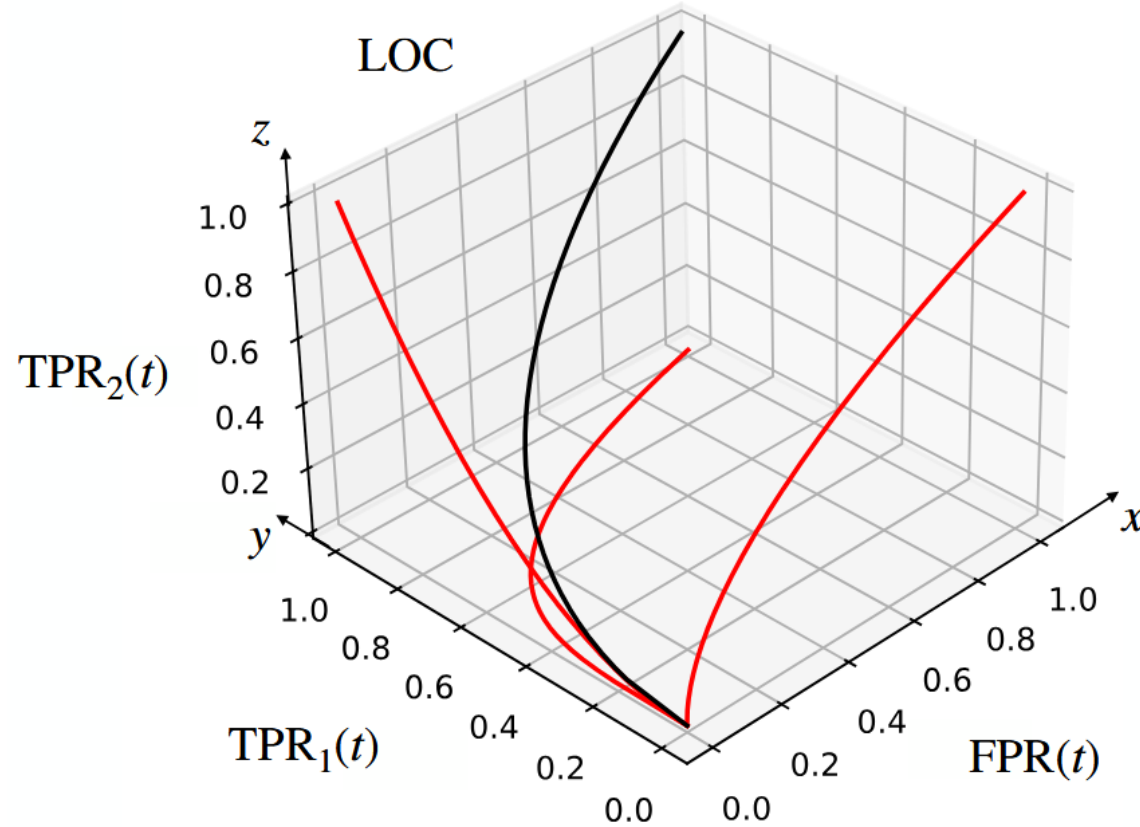


FIG. 11: Illustration of the LOC curve in 3D space, featuring two $TPR_h(t)$ dimensions and one $FPR(t)$. The complete LOC curve is depicted in black, while the individual curve projections are shown in red.



But Wait, There's More

- The curve generated using the LoC Curve convention is not limited to pairwise comparisons!
- We can minimize the length of this N-dimensional curve utilizing arbitrary hypotheses to get the best binning for hypotheses against each other

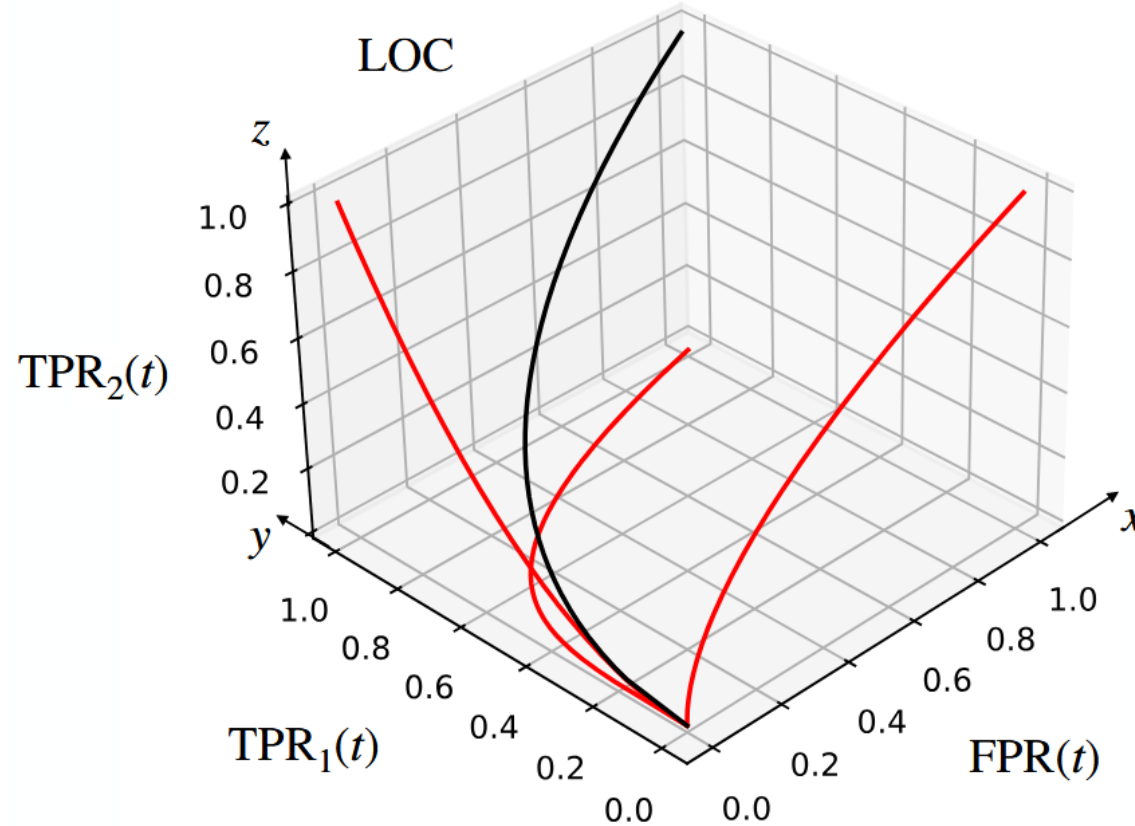
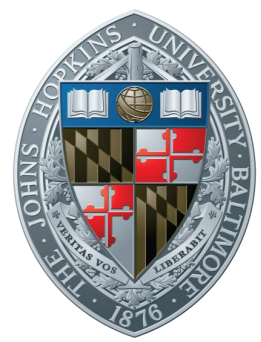


FIG. 11: Illustration of the LOC curve in 3D space, featuring two $TPR_h(t)$ dimensions and one $FPR(t)$. The complete LOC curve is depicted in black, while the individual curve projections are shown in red.



Analysis Use Case

- We can utilize MiLoMerge to model an analysis for 7 different SMEFT coefficients floated simultaneously at the HVV vertex in $H \rightarrow ZZ^* \rightarrow 4\ell$ decays at LHE-level



Analysis Use Case

- We can utilize MiLoMerge to model an analysis for 7 different SMEFT coefficients floated simultaneously at the HVV vertex in $H \rightarrow ZZ^* \rightarrow 4\ell$ decays at LHE-level
- **Take 7 observables, each sensitive to a single EFT coupling, and put them together to generate ~27k bins, then merge down to 150!**
 - Here we only care about comparing SM vs c_i

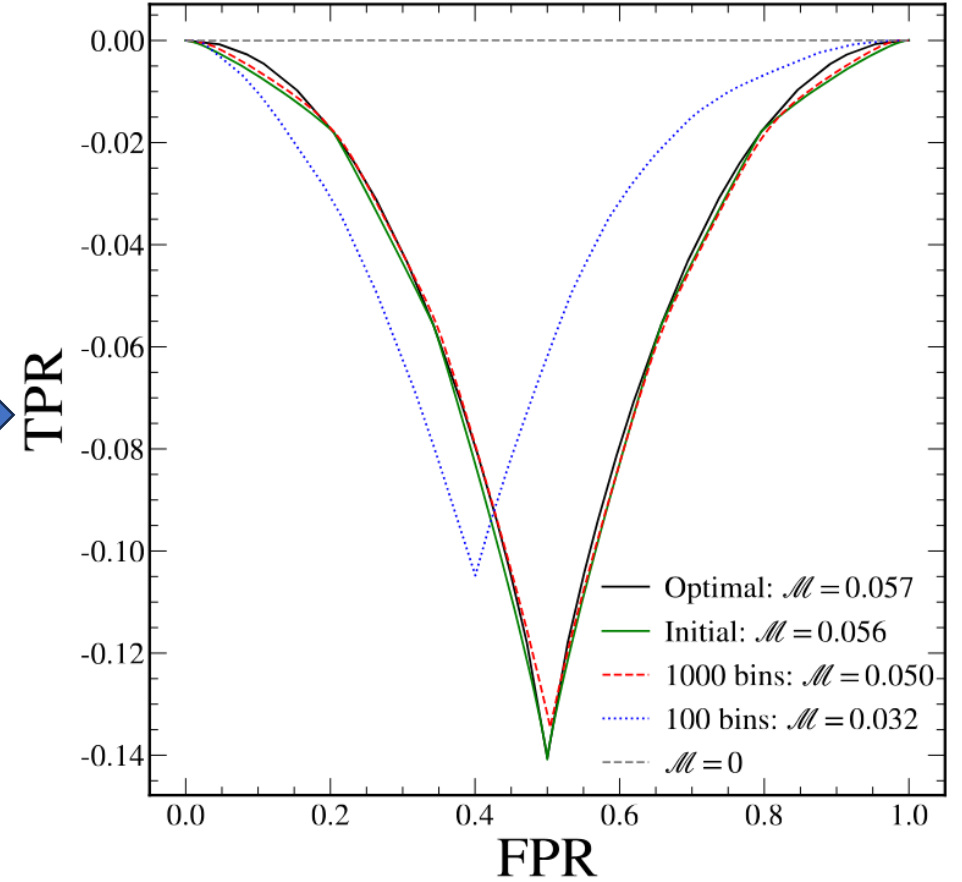
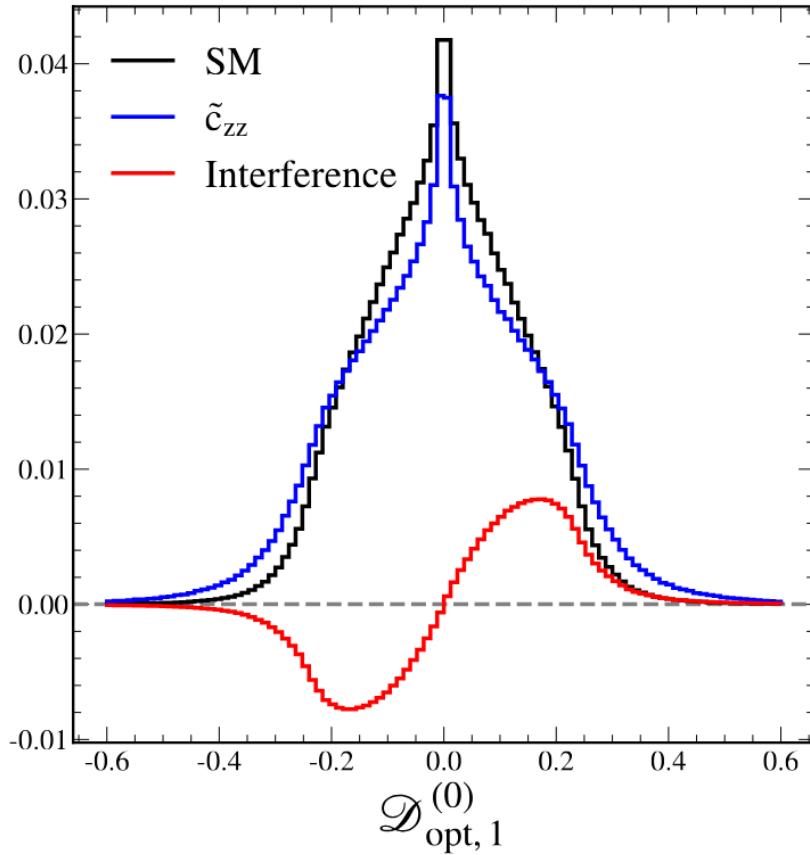


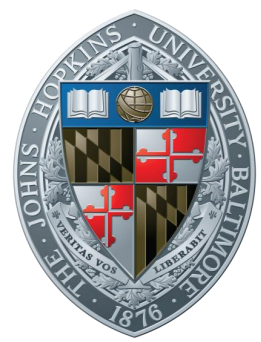
Analysis Use Case

- We can utilize MiLoMerge to model an analysis for 7 different SMEFT coefficients floated simultaneously at the HVV vertex in $H \rightarrow ZZ^* \rightarrow 4\ell$ decays at LHE-level
- **Take 7 observables, each sensitive to a single EFT coupling, and put them together to generate ~27k bins, then merge down to 150!**
 - Here we only care about comparing SM vs c_i
- The binning produced at LHE-level was the basis for the same SMEFT analysis on CMS
 - Results presented at [Higgs 2025](#) by Jeffrey Davis, [~45 minutes ago](#) by Nicholas Pinto, and mentioned in the [plenary this morning](#)

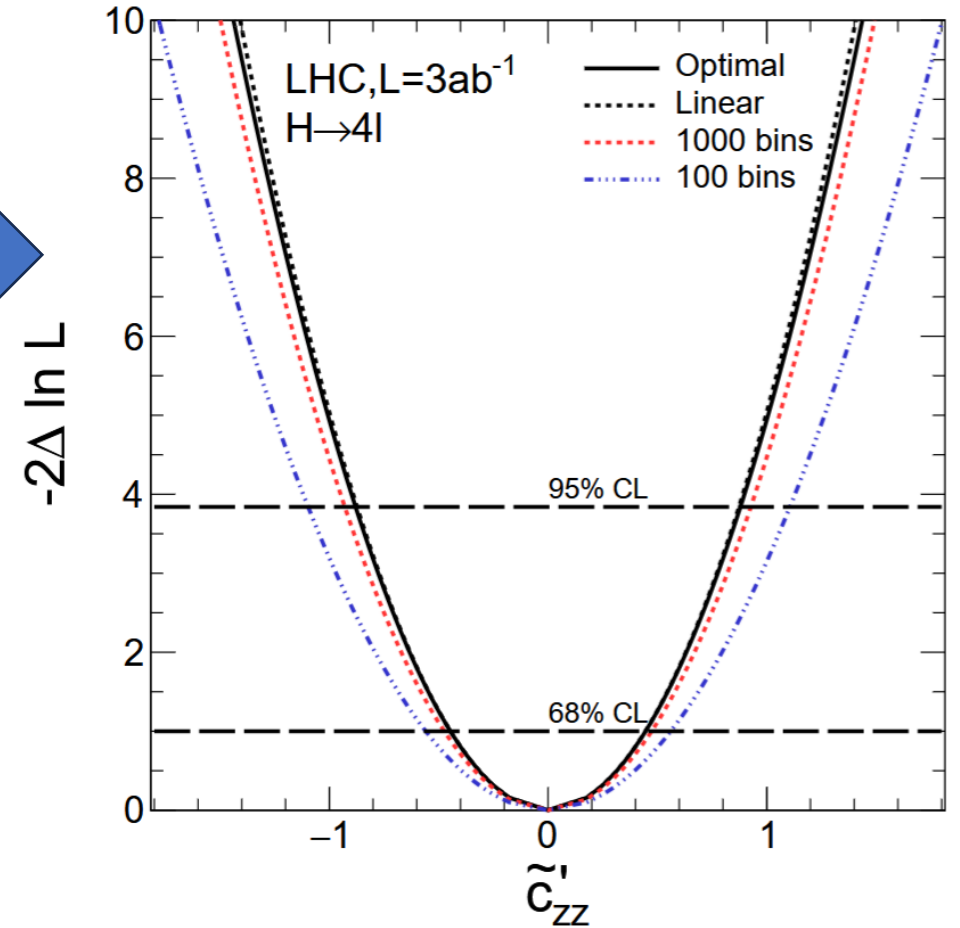
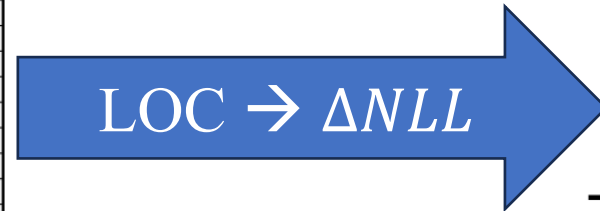
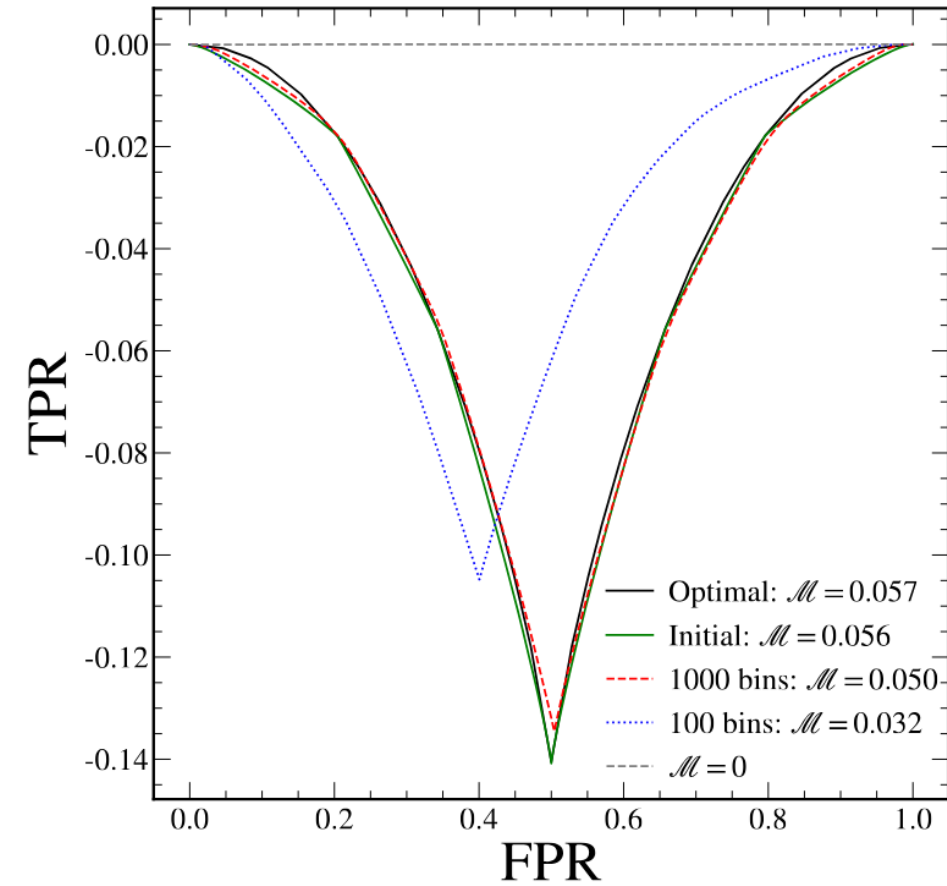


Analysis Use Case



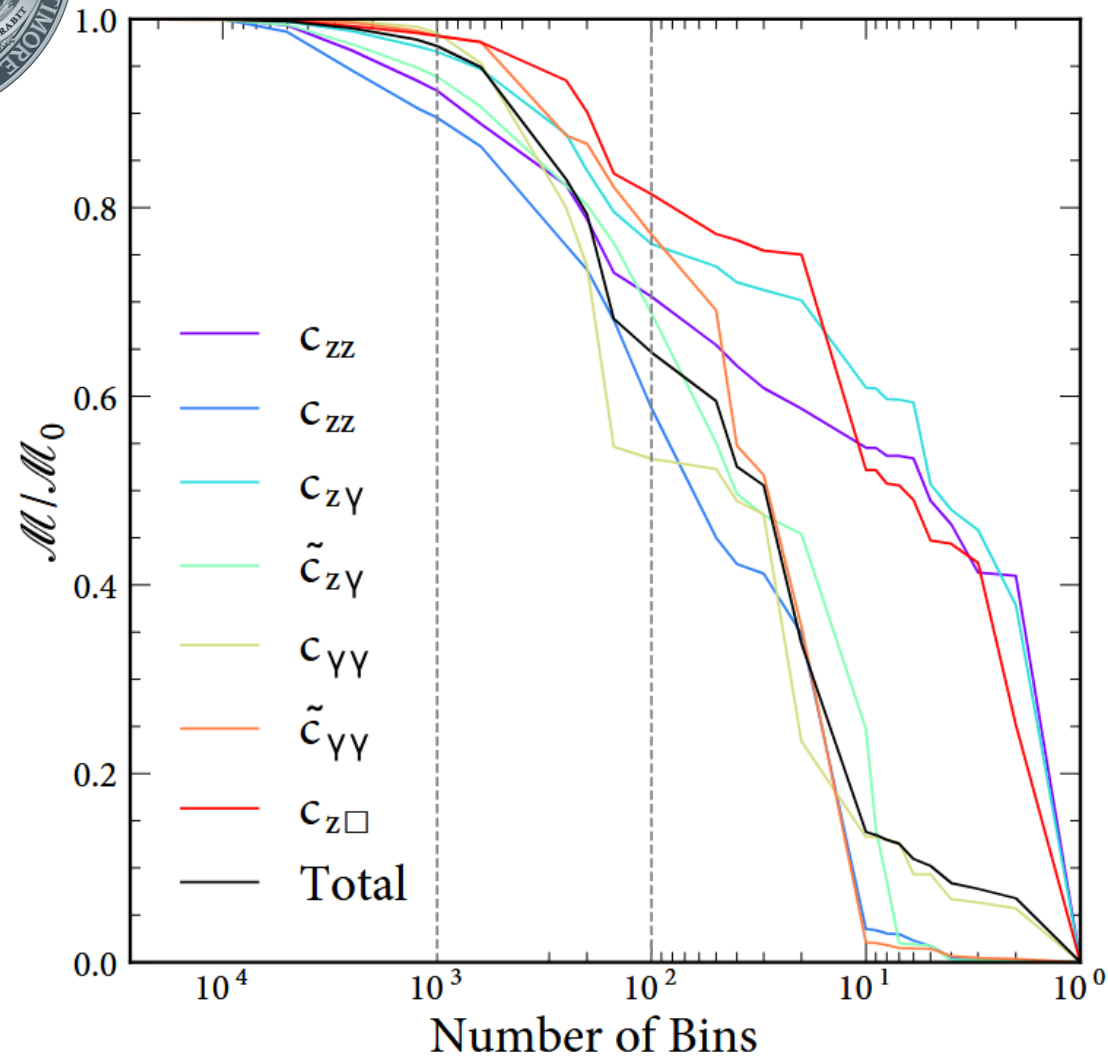


Analysis Use Case





Analysis Use Case



- Plot of ratio of the LOC for a given bin-number vs the initial binning
- 2 orders of magnitude fewer bins down to 100 with at most $\sim 50\%$ degradation, often much less!



Package Release

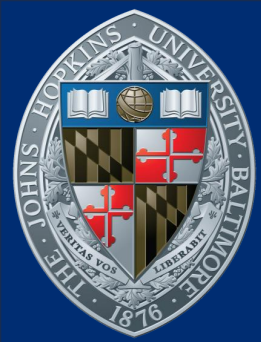
- We have released the entire framework on how to do this merging in the **MiLoMerge** package, which is available in Python for wide usage
- Available on the [JHUGen website](https://www.jhugeneering.com/)
- Available on PyPi at <https://pypi.org/project/MiLoMerge/>
- Documentation and some code examples available at <https://spin.pha.jhu.edu/MiLoMerge/>

A screenshot of the MiLoMerge 1.0.1 documentation page. The page has a dark theme. On the left is a sidebar with a search bar and a "CONTENTS" section containing links for "Package Documentation", "MiLoMerge Tutorials", and "Changelog". The main content area has a title "MiLoMerge Documentation" and an "Overview" section. The overview text states: "This documentation is provided to assist in the usage of the MiLoMerge package for 'Minimal Loss Merging'. The MiLoMerge package is designed to provide an interface for merging bins together such that the loss in separability between N distributions is minimized." Below this, it says "MiLoMerge is based off of Maximizing Returns: Optimizing Experimental Observables at the LHC, which is currently available on ArXiv." There is also a "Relevant URLs" section with text about the JHUGen website and GitHub, and a "PyPi page" link. An "Installation" section follows, stating "You can install MiLoMerge with pip, conda, or directly from the homepage." At the bottom of the installation section is a code block:

```
pip install MiLoMerge
```

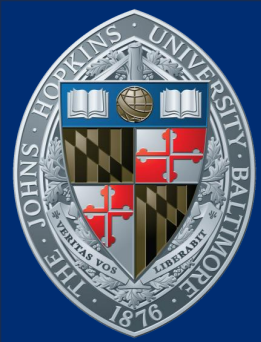
. On the right side of the page, there is a "ON THIS PAGE" sidebar with links for "Overview", "Relevant URLs", "Installation", "Papers and Talks", "Papers", "Talks", "Citing MiLoMerge", and "Frequently Asked Questions".

A lot of work has gone in on the development side to ensure that the standards are up-to-date with those in [Scikit-HEP](#) (docs, unit testing, CI/CD, etc.)



Conclusion

- The MiLoMerge framework utilizes a novel interpretation of the ROC curve, the LOC curve, to quantify the performance of observables, including those sensitive to quantum interference
- The MiLoMerge framework also carries the capability to merge bins together for an arbitrary number of hypothesis to generate an optimized binning for a given N
- The package is available on PyPi and has been used in multiple CMS analyses to date



I WANT MORE

LPC EFT Workshop



Hosted by the Center of
Fundamental Interactions


JOHNS HOPKINS
UNIVERSITY

June 1-2, 2026

3701 San Martin Dr.
Baltimore, MD



Bloomberg Center for Physics and Astronomy, photo credit: Flux Studio Ltd (2014)

- The [LPC EFT workshop](#) will have talks about SMEFT and more
- Contributions from phenomenologists, AMO and HEP experimentalists, etc.

Organizing committee:

Jonathan Wilson (Baylor)
Lindsey Gray (Fermilab)
Nick Smith (Fermilab)
Vasilije Perovic (Fermilab)
Kelci Mohrman (Florida)
Andrei Gribsan (JHU)
Jeffrey Davis (JHU)
Surjeet Rajendran (JHU)
Michael Schmitt (Northwestern)
Kevin Lannon (Notre Dame)
Andrea Piccinelli (Notre Dame)
Adam Martin (Notre Dame)
Saptaparna Bhattacharya (SMU)

Local organizing committee:

Andrei Gribsan (JHU)
Jeffrey Davis (JHU)
Mohit Srivastav (JHU)
Zhiyuan Huang (JHU)
Nick Pinto (JHU)
Michalis Panagiotou (JHU)
Ruoxi Wang (JHU)

LPC Events committee

Gabriele Benelli (Brown U., Co-Chair)
Jonathan Wilson (Baylor U., Liaison)

LPC Coordinators:

Jim Hirshauer (Fermilab)
Isobel Ojalvo (Princeton)

Mohit Srivastav- Johns Hopkins University

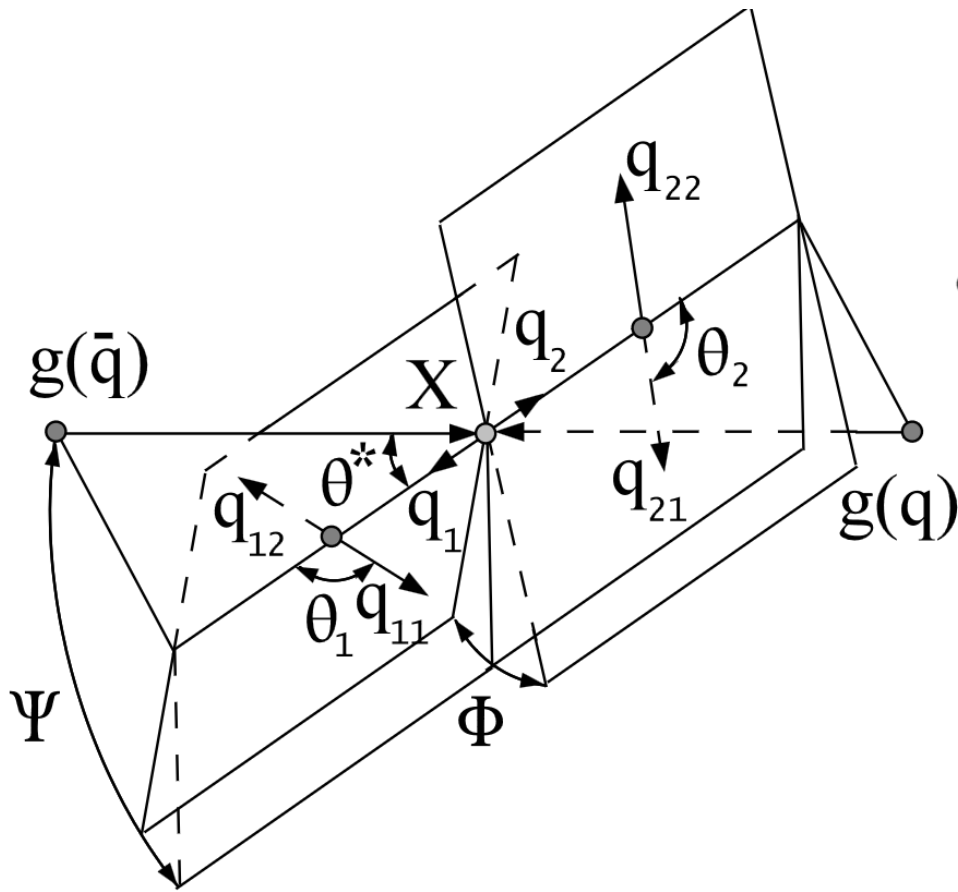


BACKUP



Angle Definitions

[ArXiv:1208:4018](https://arxiv.org/abs/1208.4018)



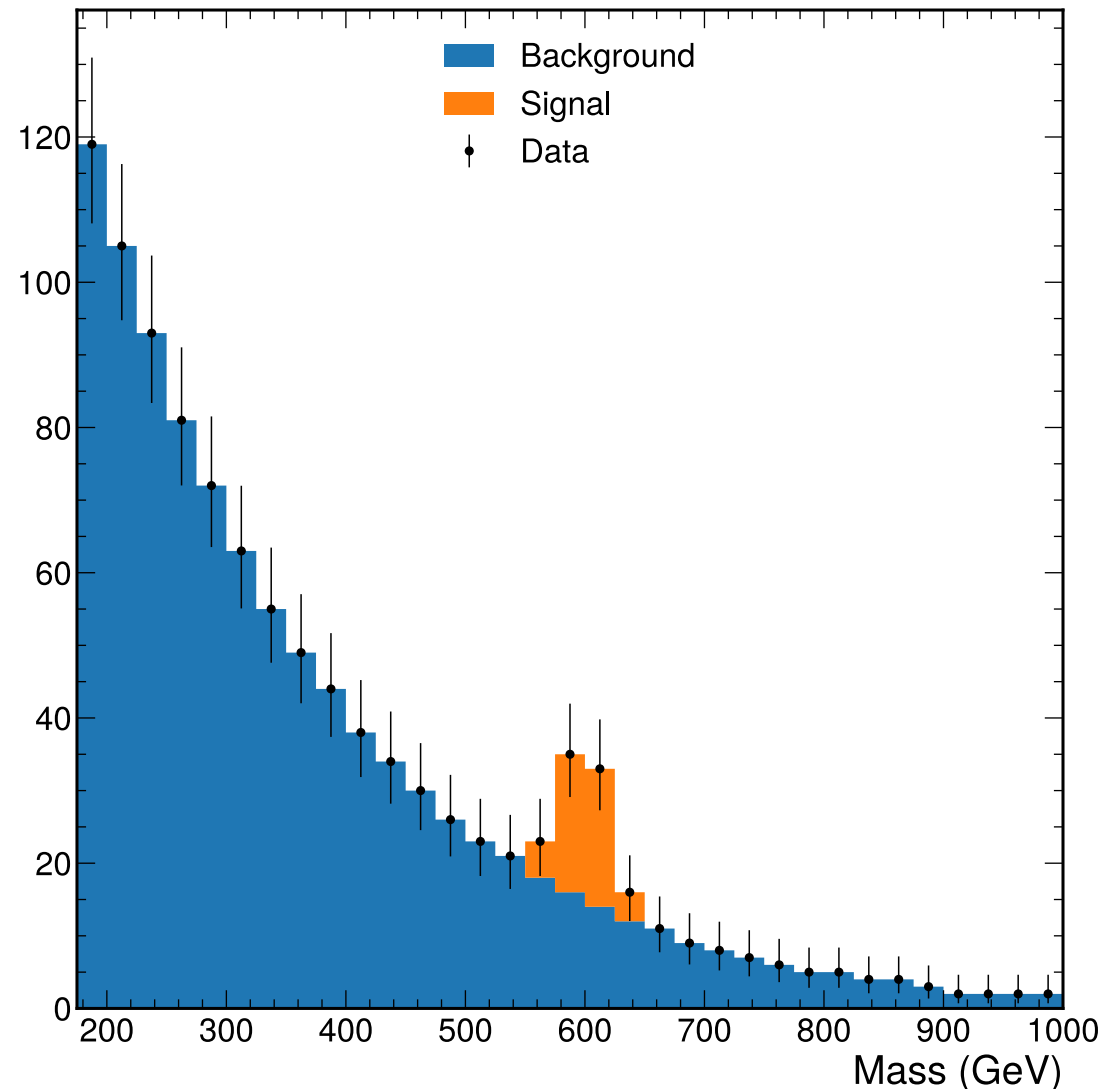
$$\Phi = \frac{\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)}{|\mathbf{q}_1 \cdot (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2)|} \times \cos^{-1} (-\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2)$$

$$\hat{\mathbf{n}}_1 = \frac{\mathbf{q}_{11} \times \mathbf{q}_{12}}{|\mathbf{q}_{11} \times \mathbf{q}_{12}|}, \quad \hat{\mathbf{n}}_2 = \frac{\mathbf{q}_{21} \times \mathbf{q}_{22}}{|\mathbf{q}_{21} \times \mathbf{q}_{22}|}$$



Separability

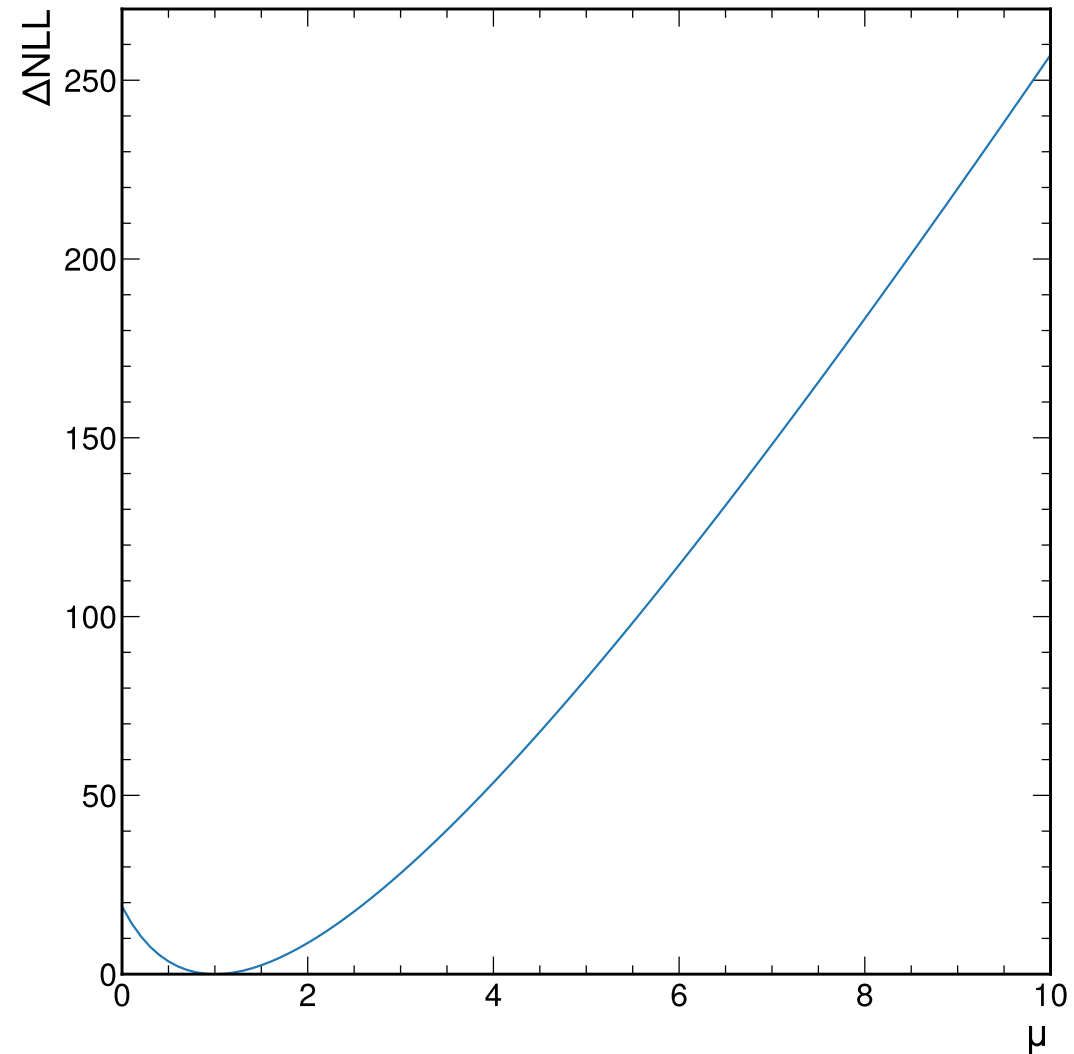
- Take a simple model that we can make of signal and background overlaid with “data”
- Take a simple scaling function where the signal is scaled by a dimensionless μ , and see the ΔNLL to measure the separability between our “data” and different values of μ

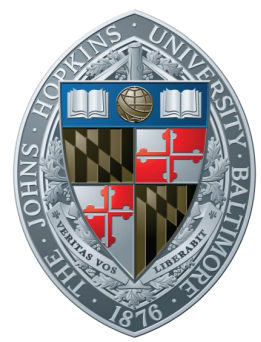




Separability

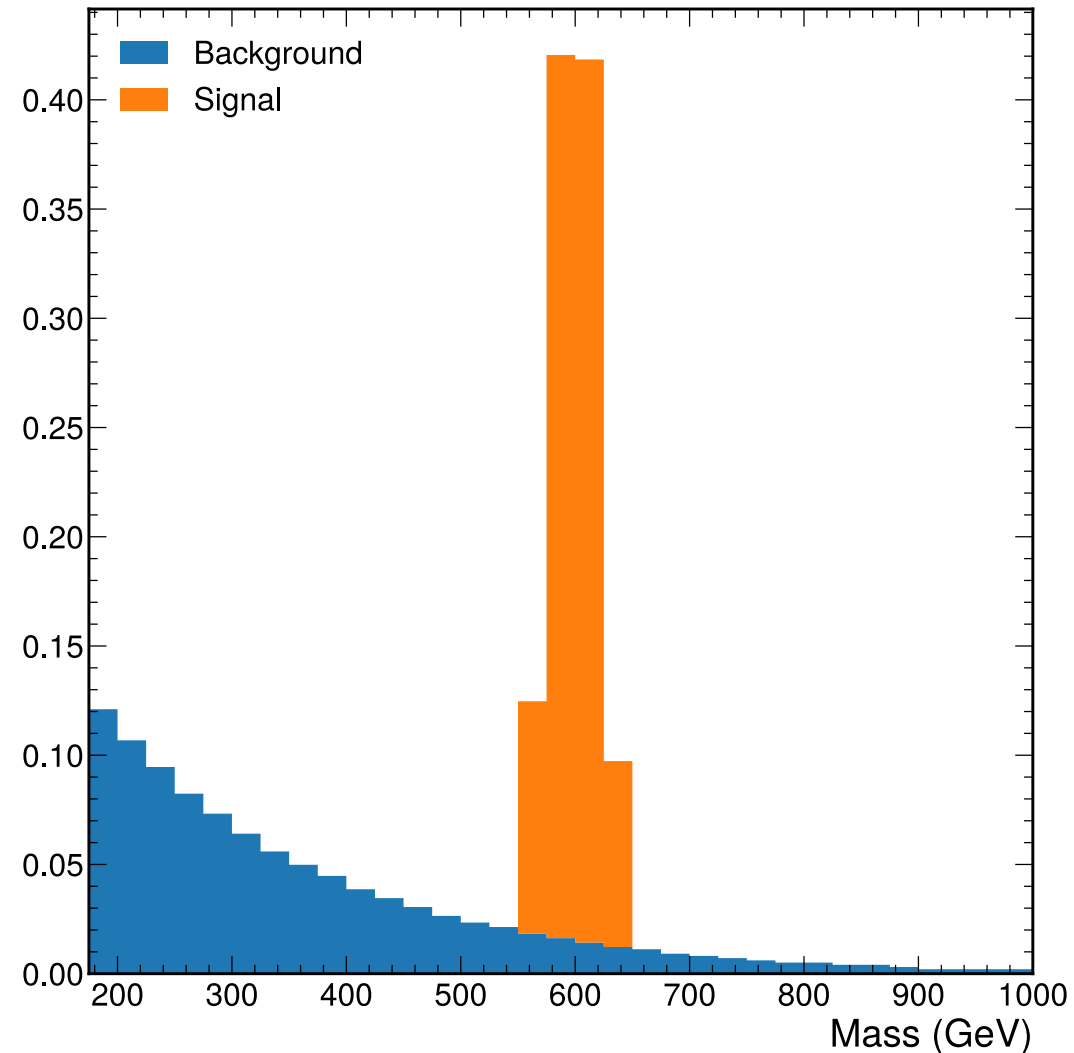
- Take a simple model that we can make of signal and background overlaid with “data”
- Take a simple scaling function where the signal is scaled by a dimensionless μ , and see the ΔNLL to measure the separability between our “data” and different values of μ





Separability

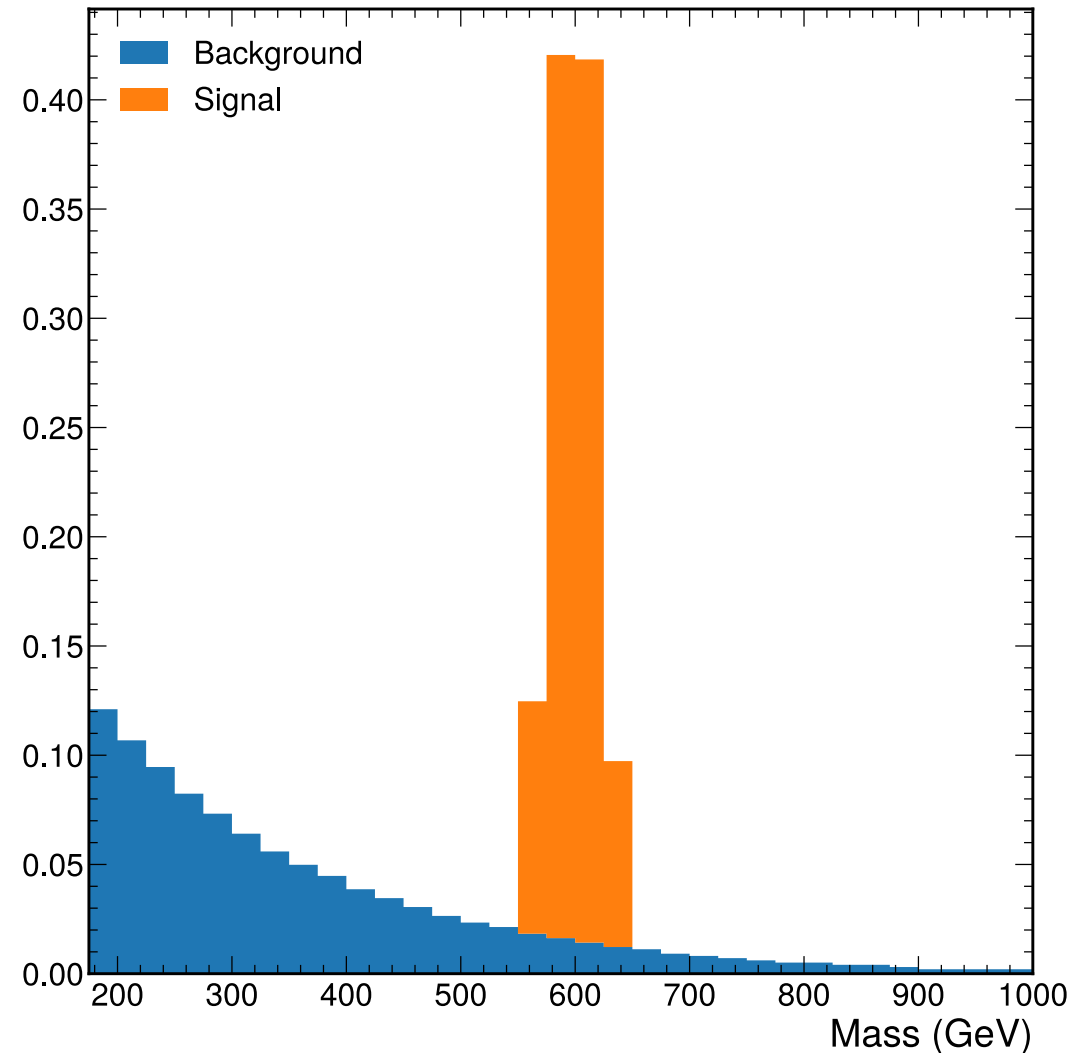
- **What does the ROC curve deal in?**
Because the distributions are normalized such that $\Sigma = 1$ for each hypothesis, this is what the metric is really “seeing”
- No matter the value of $\mu \neq 0$, due to this normalization, it will showcase the same distribution





Separability

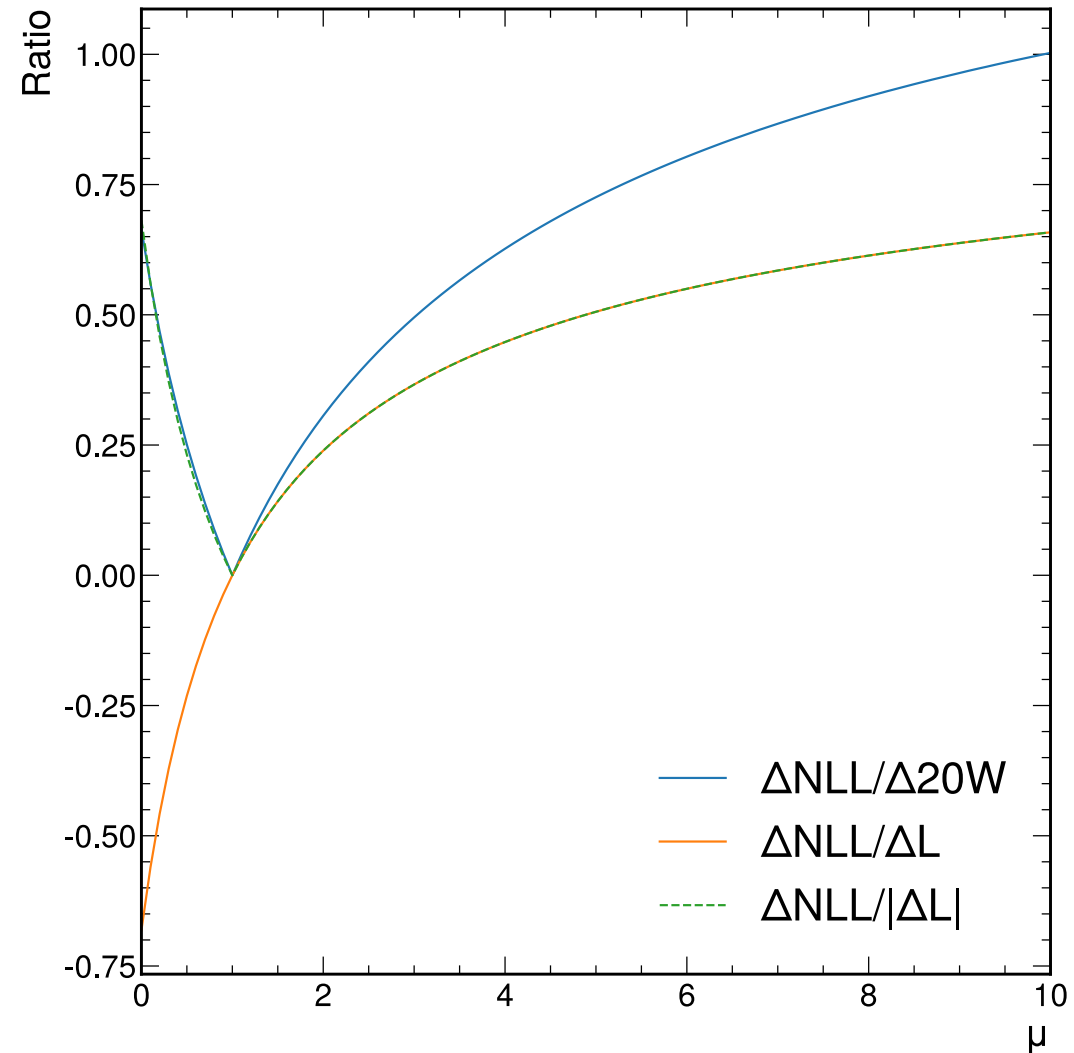
- **What does the ROC curve deal in?**
Because the distributions are normalized such that $\Sigma = 1$ for each hypothesis, this is what the metric is really “seeing”
- No matter the value of $\mu \neq 0$, due to this normalization, it will showcase the same distribution
- **Utilizing this to quantify a loss is not nearly as useful!**





Separability

- Comparing the length to the ΔNLL , we can see that there is a clear monotonic relationship between that and the difference in length, ΔL
- Can compare to other common measures of “distance” between distributions (i.e., Wasserstein metric, ΔW)

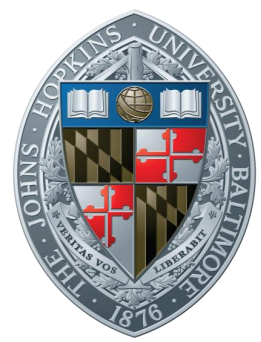




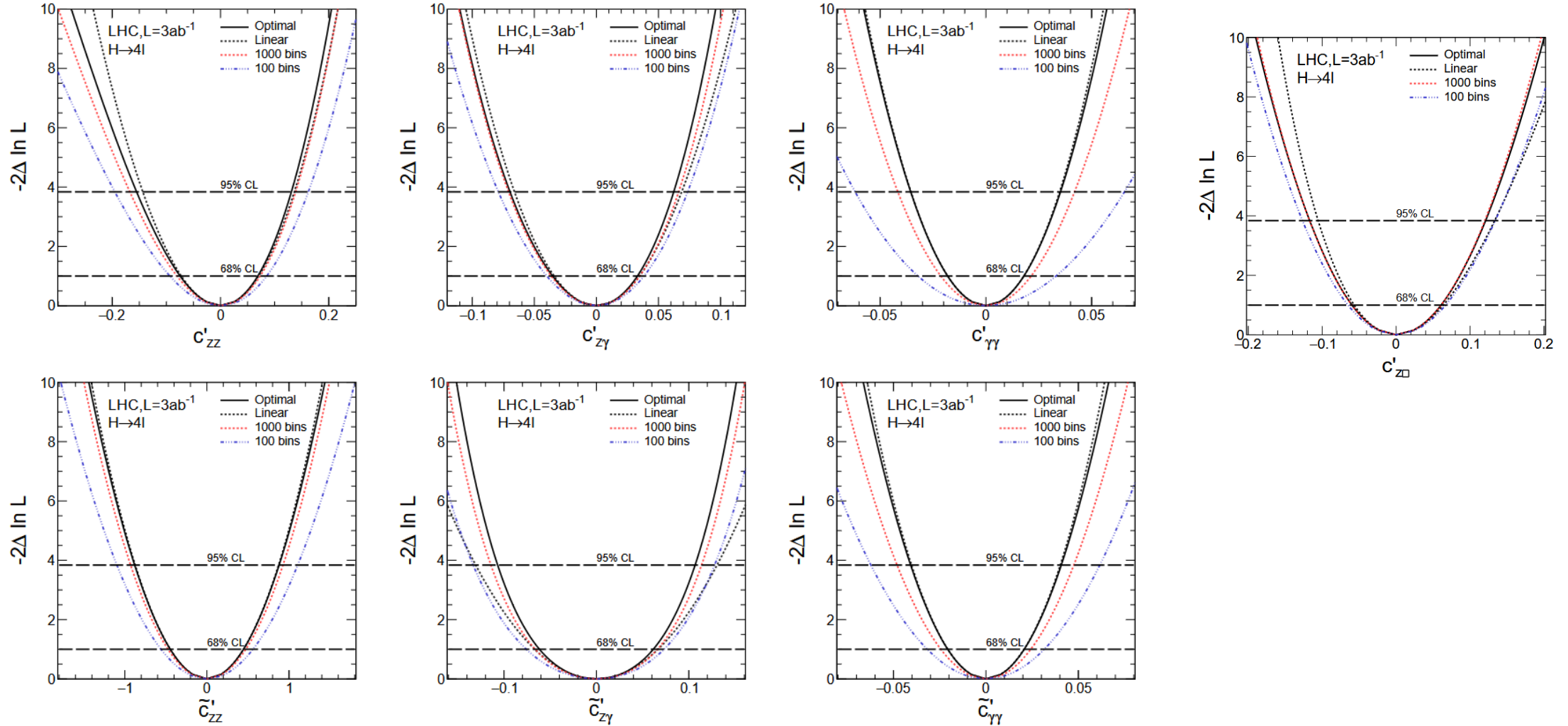
LOC Curve Formula

$$FPR = \frac{1}{|A|} \sum_{x \in A} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$

$$TPR_h = \frac{1}{|A|} \sum_{x \in B_h} \begin{cases} 1 & \text{if } s(x) > t \\ 0 & \text{o.w.} \end{cases}$$



Overall Results





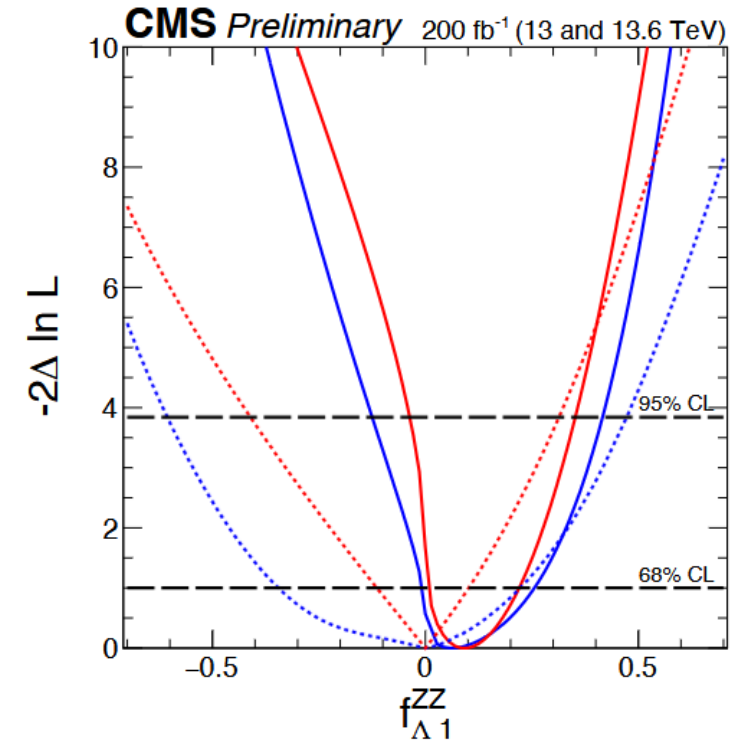
CMS Results

[CMS-PAS-HIG-25-011](#)

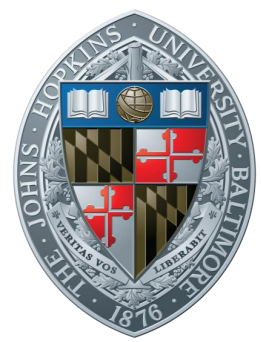


Effective Field Theory: Results

- Perform maximum-likelihood scan of all 7 anomalous couplings.
- Scans performed with all other AC simultaneously floated (blue), and all other AC fixed (red).
- Dashed = Expectation from Monte Carlo.
- Solid = Observation from CMS Data.



[Study of Spin Correlations in Higgs boson decays to four leptons at CMS \(Pheno 2026\)](#)



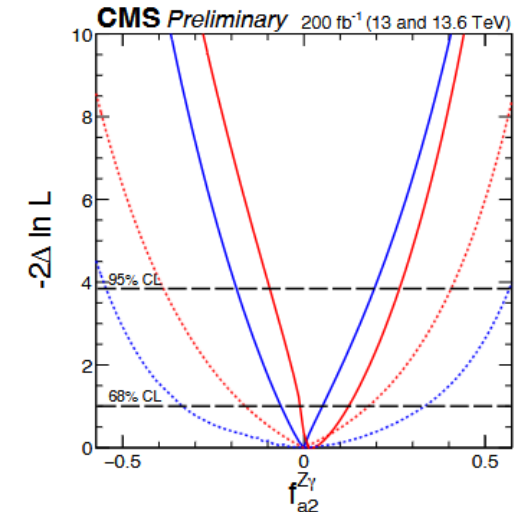
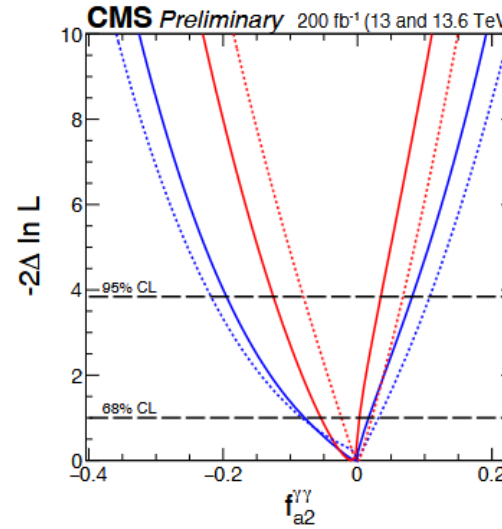
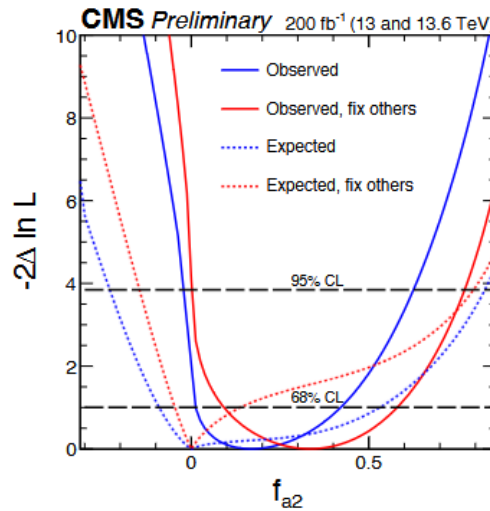
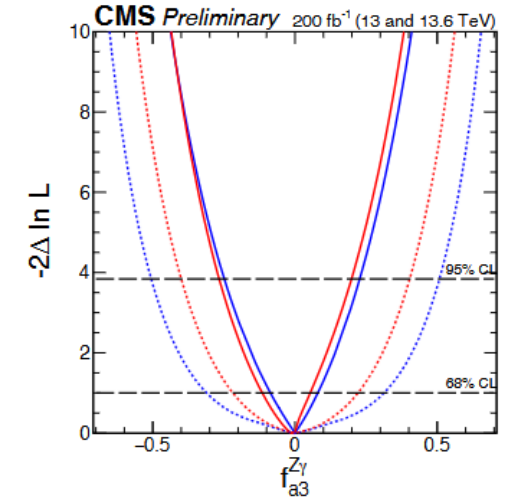
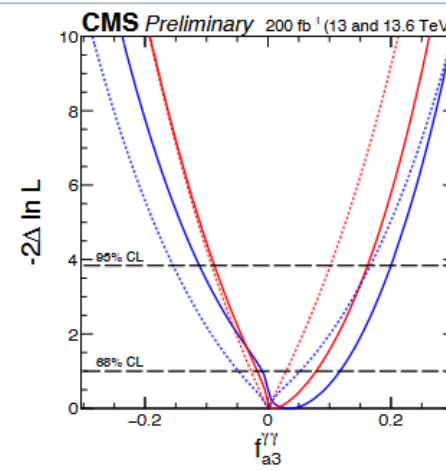
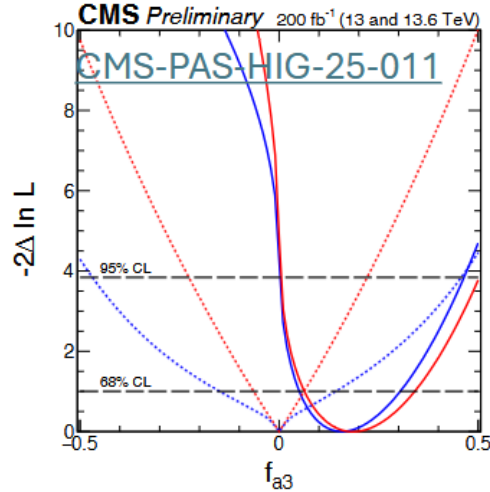
CMS Results

CMS-PAS-HIG-25-011



Effective Field Theory: Results

[Study of Spin Correlations in Higgs boson decays to four leptons at CMS \(Pheno 2026\)](#)



Nicholas Pinto - Phenomenology 2026 Symposium

16