Using the (Modified) Matrix Element Method to constrain $L_\mu - L_\tau$ Interactions

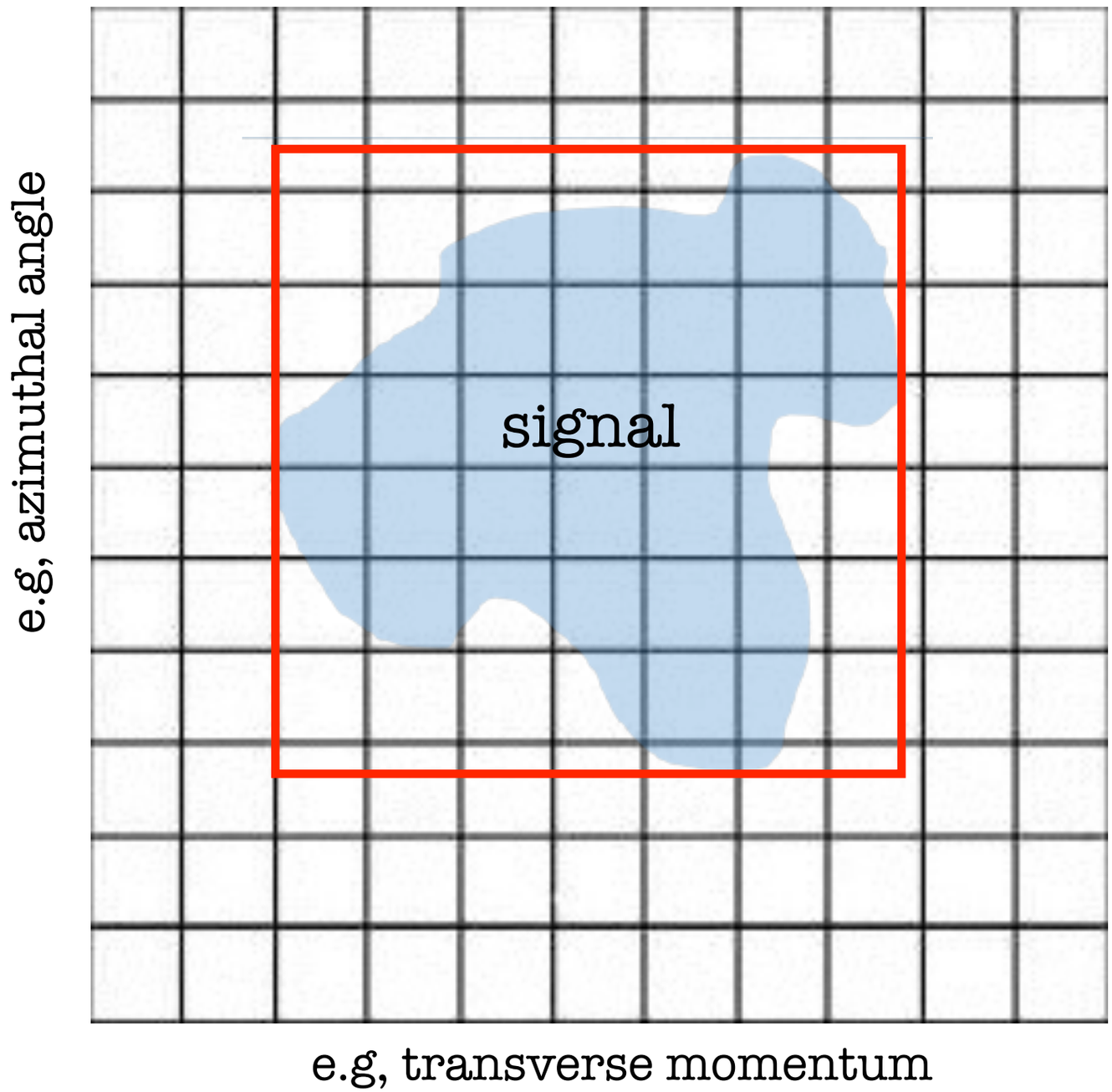
Fatemeh Elahi
University of Notre Dame
Pheno 2017

With Adam Martin
Appearing on arXiv tonight
May 8th, 2017

Optimizing our signal-background discrimination at colliders

At colliders, we measure the 4-momenta of (some) final state particles;
for example: transverse momentum, azimuthal angle, etc.

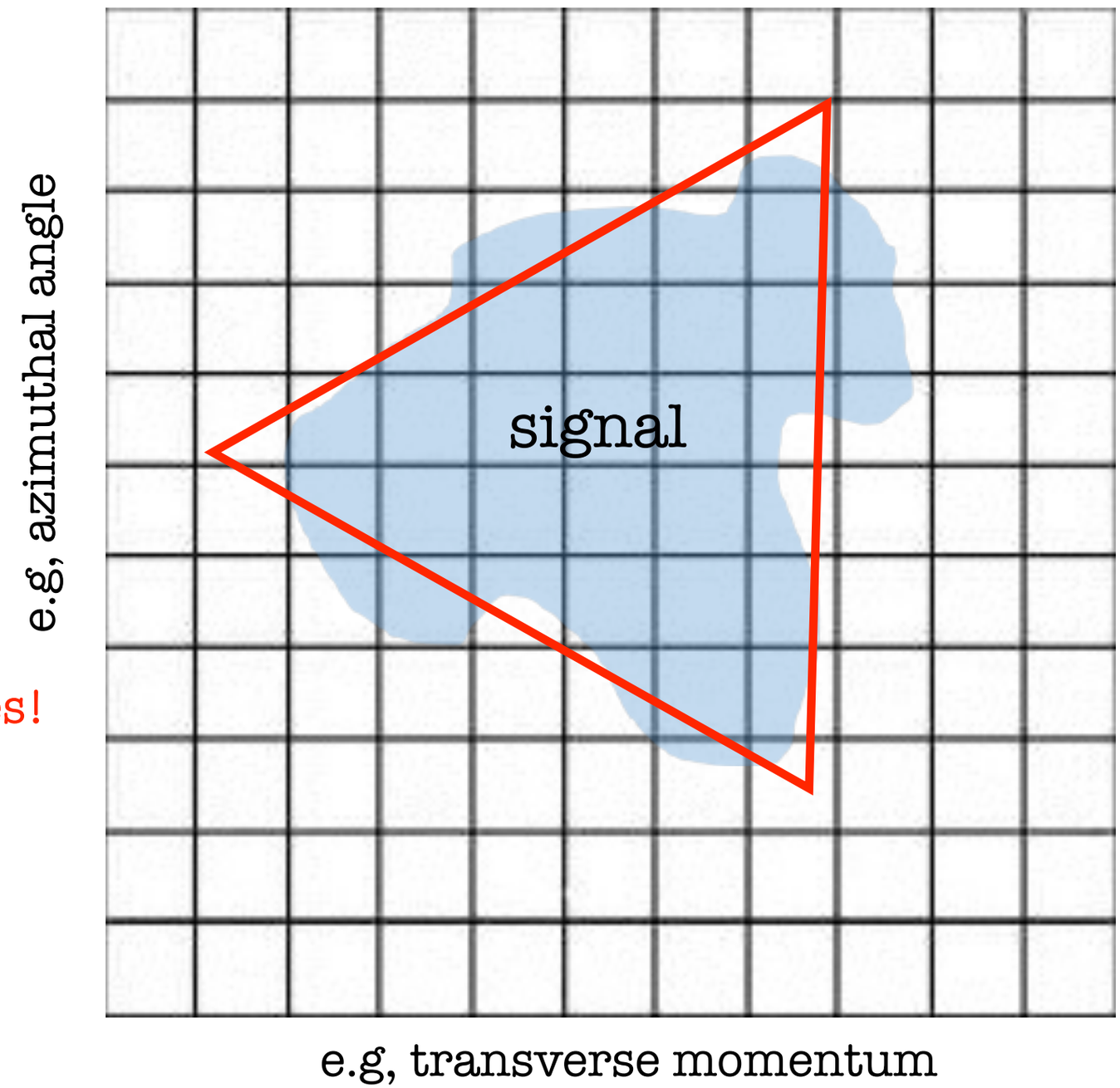
No correlation between variables!



Optimizing our signal-background discrimination at colliders

Use kinematic variables that include correlations between observables: invariant mass of some visible particles, angular separation between particles.

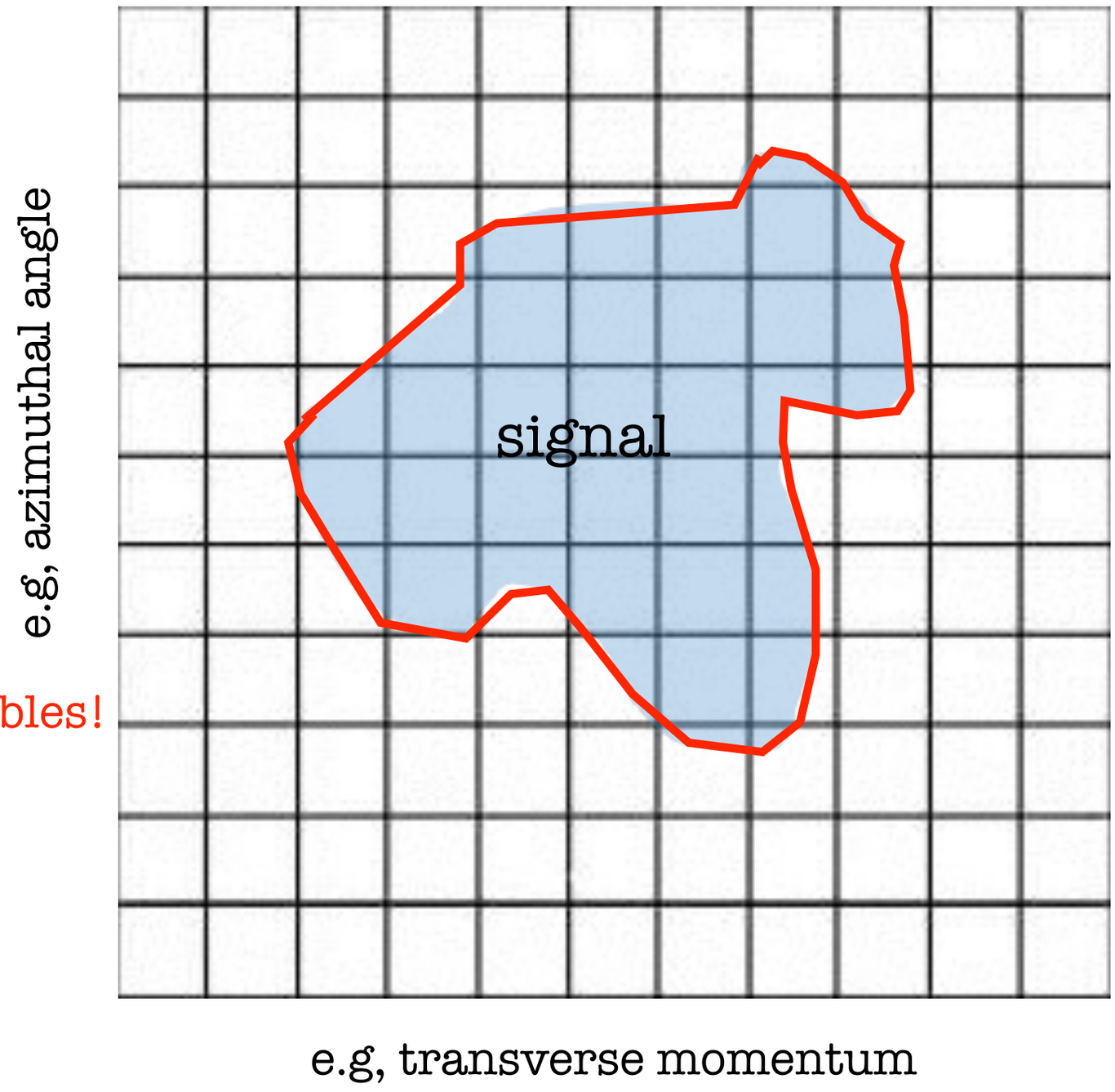
Some correlation between variables!



Optimizing our signal-background discrimination at colliders

Multivariant Analysis:
Neural Network,
BoostedDecision Tree,
Matrix Element Method.

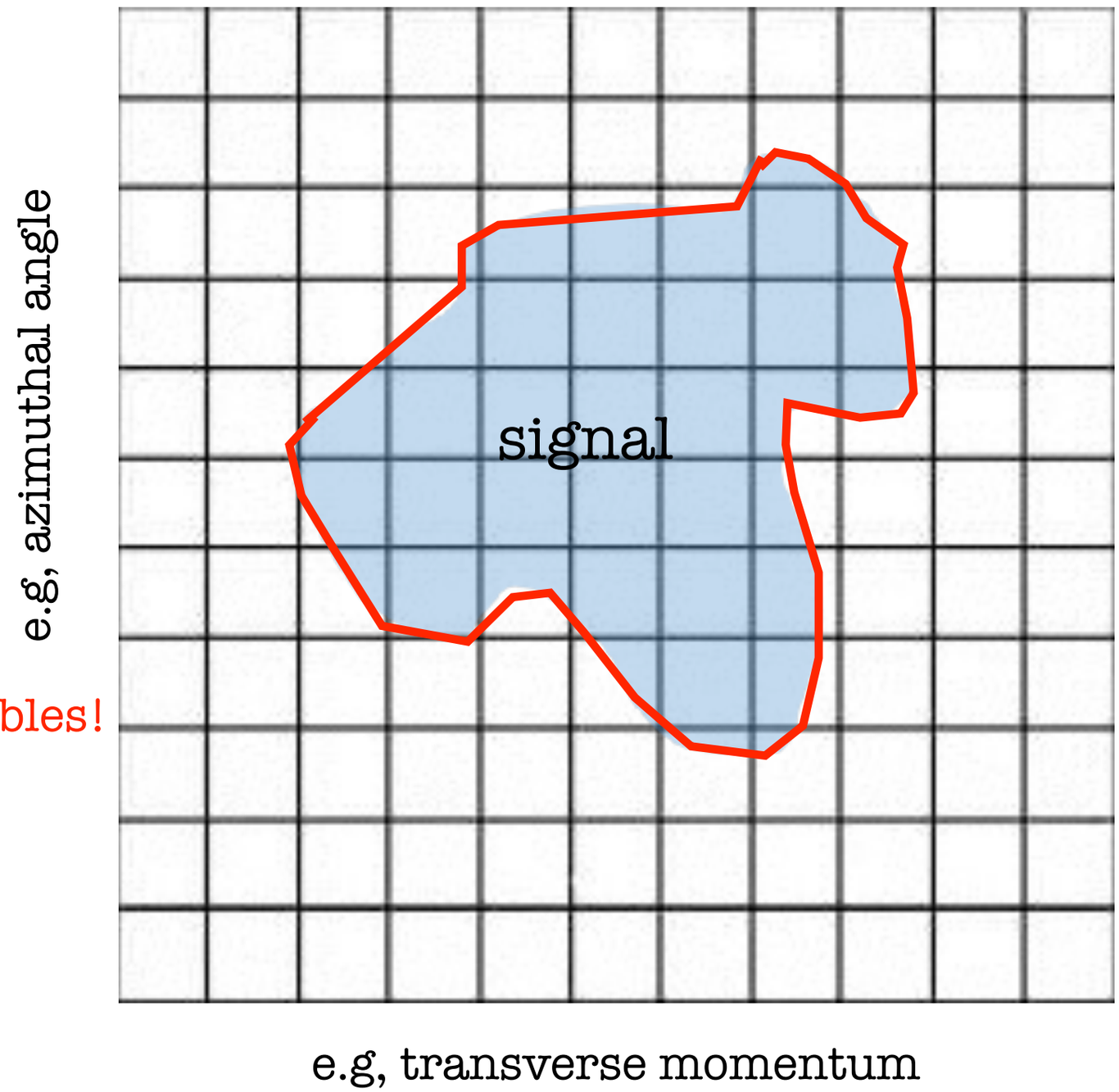
Optimizing correlation between variables!



Optimizing our signal-background discrimination at colliders

Multivariant Analysis:
Neural Network,
BoostedDecision Tree,
Matrix Element Method.

Optimizing correlation between variables!



Matrix Element Method (MEM)

Use matrix element of a process to distinguish signal from background.

$$|\mathcal{M}|^2 = |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2 \, \delta^4(p_i - p_f)$$

p_i, p_f : Initial and final state momenta

$p_f^{\rm inv}$: momenta of particles missed at the detector

α : model parameters (e.g, mass, couplings)

Matrix Element Method (MEM)

Use matrix element of a process to distinguish signal from background.

$$|\mathcal{M}|^2 = |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2 \, \delta^4(p_i - p_f)$$

p_i, p_f : Initial and final state momenta

$p_f^{\rm inv}$: momenta of particles missed at the detector

α : model parameters (e.g, mass, couplings)

The probability is defined as:

$$P(p^{\rm vis}|\alpha) = \frac{1}{\sigma} \int dp_i \, dp^{\rm inv} |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2 \delta^4(p_i - p_f)$$

σ : normalization factor to ensure $\int P(p^{\rm vis}|\alpha) dp^{\rm vis} = 1$

Matrix Element Method (MEM)

Use matrix element of a process to distinguish signal from background.

$$|\mathcal{M}|^2 = |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2\, \delta^4(p_i - p_f)$$

p_i, p_f : Initial and final state momenta

$p_f^{\rm inv}$: momenta of particles missed at the detector

α : model parameters (e.g, mass, couplings)

The probability is defined as:

$$P(p^{\rm vis}|\alpha) = \frac{1}{\sigma} \int dp_i\, dp^{\rm inv} |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2 \delta^4(p_i - p_f)$$

σ : normalization factor to ensure $\int P(p^{\rm vis}|\alpha) dp^{\rm vis} = 1$

For jets, there are other complications – not the discussion of this talk

Matrix Element Method (MEM)

Use matrix element of a process to distinguish signal from background.

$$|\mathcal{M}|^2 = |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2\, \delta^4(p_i - p_f)$$

p_i, p_f : Initial and final state momenta

$p_f^{\rm inv}$: momenta of particles missed at the detector

α : model parameters (e.g, mass, couplings)

The probability is defined as:

$$P(p^{\rm vis}|\alpha) = \frac{1}{\sigma}\int dp_i\, dp^{\rm inv} |\mathcal{M}(p_i, p_f^{\rm vis}, p_f^{\rm inv}; \alpha)|^2 \delta^4(p_i - p_f)$$

σ : normalization factor to ensure $\int P(p^{\rm vis}|\alpha) dp^{\rm vis} = 1$

Likelihood that a given event belongs to the signal:

$$\mathcal{L}(p^{\rm vis}|\alpha) = \frac{P(p^{\rm vis}|\alpha(\text{new physics}))}{P(p^{\rm vis}|\alpha(\text{SM}))}$$

Advantages

Uses ALL of the available kinematic information.

the most accurate measurement of top quark mass

Has a clear theoretical meaning, and does not require a phase of computer training like other MVAs.

Disadvantages

Computationally challenging:

Need to find the optimal value of $\alpha(\text{new physics})$.

Need to integrate for each unconstrained momenta p^{inv}.

Advantages

Uses ALL of the available kinematic information.

the most accurate measurement of top quark mass

Has a clear theoretical meaning, and does not require a phase of computer training like other MVAs.

Disadvantages

Computationally challenging:

Need to find the optimal value of $\alpha(\text{new physics})$.

Need to integrate for each unconstrained momenta p^{inv}.

Let's see how we can combat these challenges in a specific new physics example.

Example: $L_\mu - L_\tau$

An accidental global symmetry of SM.

Anomaly free! Can be gauged.

$L_\mu - L_\tau$ is not realized in nature (i.e. tau decay to electron), so it must be broken.

predicts a Z' that only couples to muon and tau family.

Currently not very constrained, because it does not couple at tree level to quarks (LHC) and electron (BaBar II).

Also motivated because of

$(g-2)_\mu$ anamolous magnetic moment of muon

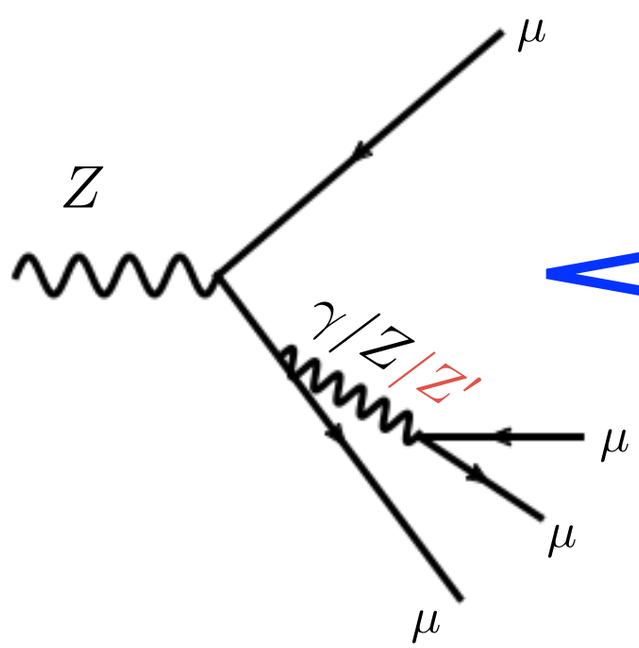
Also motivated because of the recent LHCb observation of

$$R_{K^{(*)}} = \frac{\mathrm{Br}(B \to K^{(*)}\mu^+\mu^-)}{\mathrm{Br}(B \to K^{(*)}e^+e^-)}$$

$L_\mu - L_\tau$ signature at the LHC:

$pp \to Z \to 4\mu$ is clean and well-measured.

$pp \to Z \to 4e$ and $pp \to Z \to 2e2\mu$ can be used as background control sample.



signal is defined as $SM + Z'$ to include interference.

12 independent observables strongly motivates the usage of an MVA.

The final states are fully reconstructable, and we have no unknown momenta.

But still need to determine $\alpha(\text{new physics}) = (M_{Z'}, g_{Z'}(\text{coupling}))$

Determining the model parameters to maximize signal likelihood

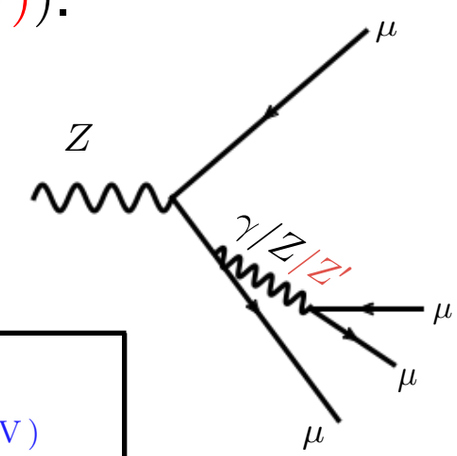
Maximize $\mathcal{L}(p^{\rm vis}|(M_{Z'}, g_{Z'})) = \dfrac{P(p^{\rm vis}|(M_{Z'}, g_{Z'}))}{P(p^{\rm vis})}$, for each event.

For a sample of N events, maximize $\sum_{i=1}^{\rm N\ events} \log \mathcal{L}(p_i^{\rm vis}|(M_{Z'}, g_{Z'}))$.

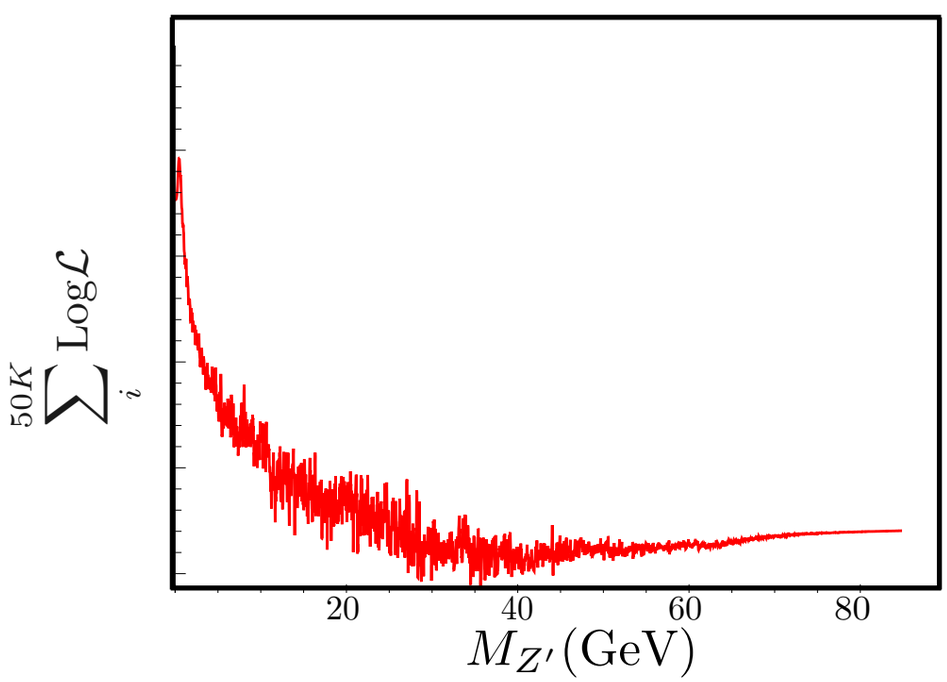
Determining the model parameters to maximize signal likelihood

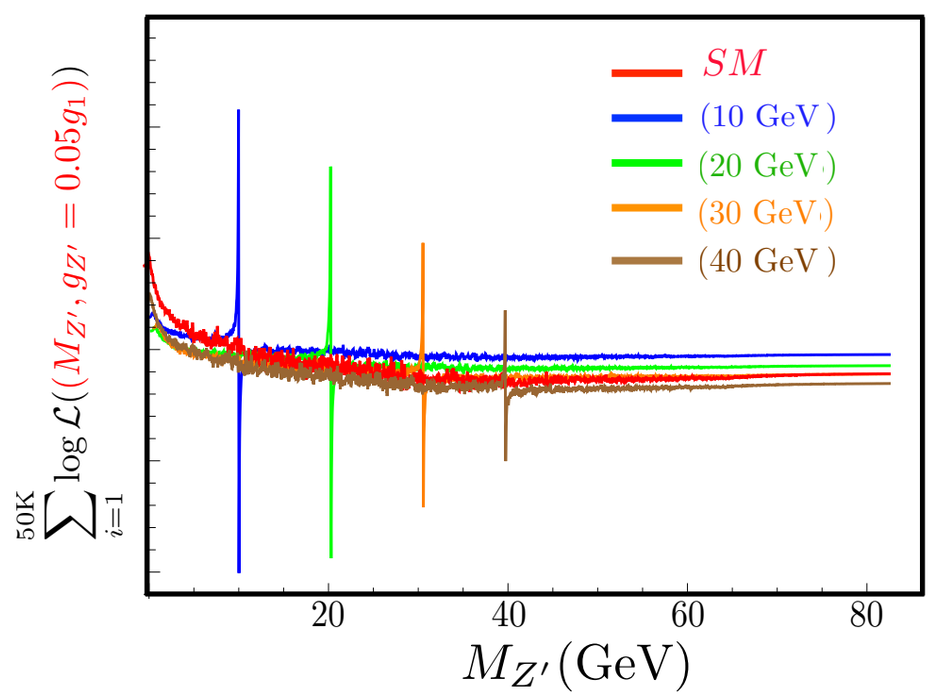
Maximize $\mathcal{L}(p^{\rm vis}|(M_{Z'}, g_{Z'})) = \dfrac{P(p^{\rm vis}|(M_{Z'}, g_{Z'}))}{P(p^{\rm vis})}$, for each event.

For a sample of N events, maximize $\sum_{i=1}^{\rm N\ events} \log \mathcal{L}(p_i^{\rm vis}|(M_{Z'}, g_{Z'}))$.



First, lets fix the coupling and find the optimal mass.





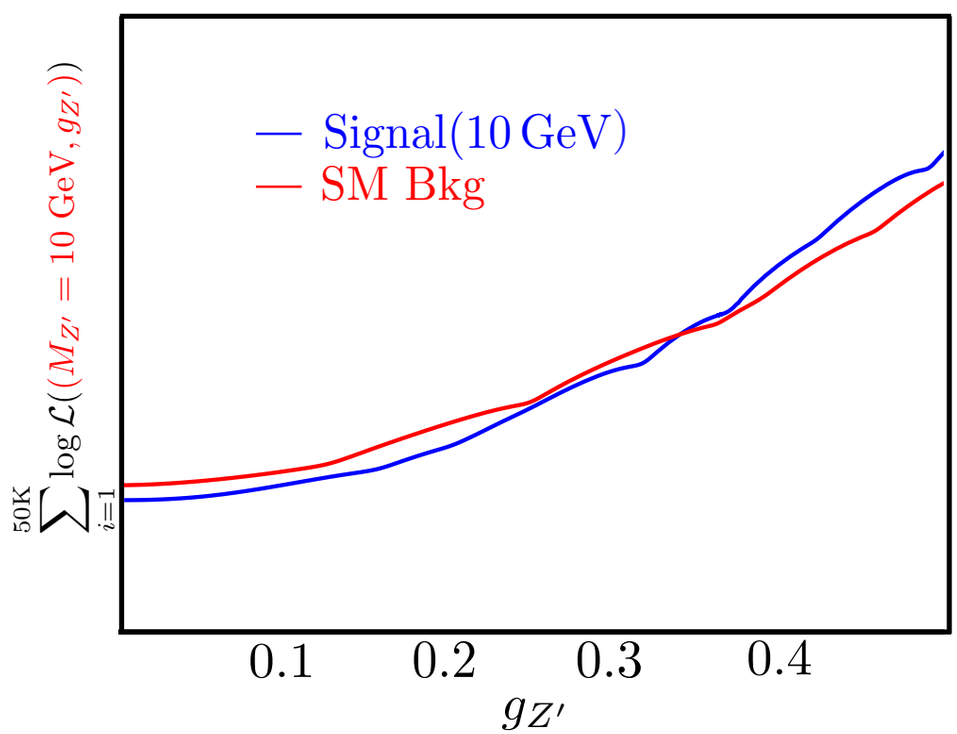
We can correctly determine the mass of Z', this way.

Determining the model parameters to maximize signal likelihood

Maximize $\mathcal{L}(p^{\rm vis}|(M_{Z'}, g_{Z'})) = \dfrac{P(p^{\rm vis}|(M_{Z'}, g_{Z'}))}{P(p^{\rm vis})}$, for each event.

For a sample of N events, maximize $\sum_{i=1}^{\rm N\ events} \log \mathcal{L}(p_i^{\rm vis}|(M_{Z'}, g_{Z'}))$.

Now lets fix the mass, and find the optimal value of the coupling:

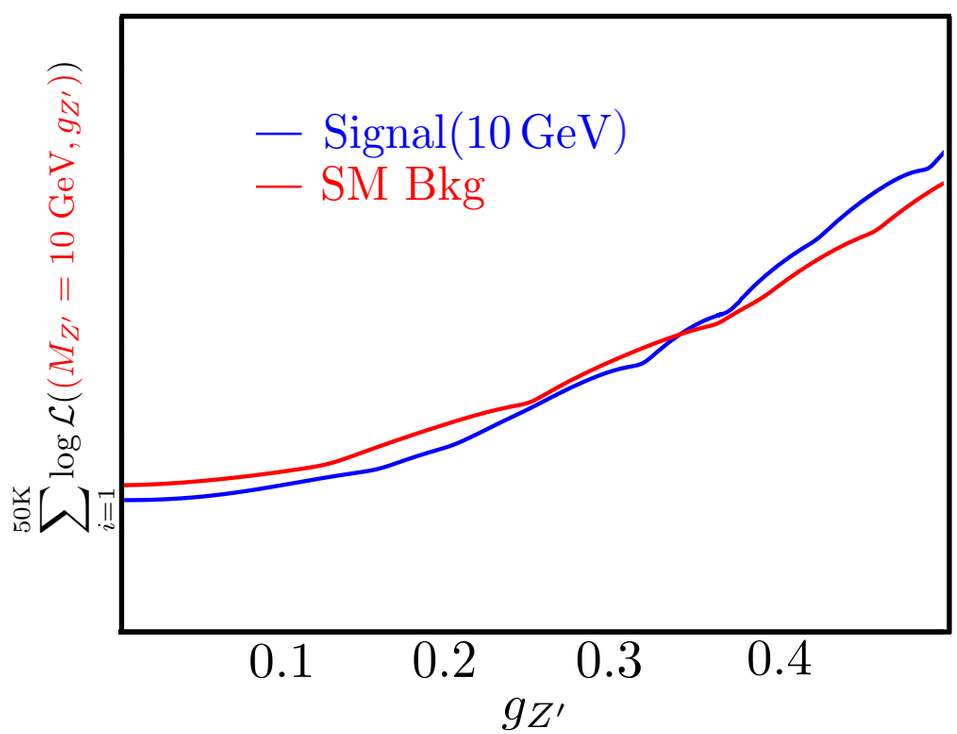


Determining the model parameters to maximize signal likelihood

Maximize $\mathcal{L}(p^{\rm vis}|(M_{Z'}, g_{Z'})) = \dfrac{P(p^{\rm vis}|(M_{Z'}, g_{Z'}))}{P(p^{\rm vis})}$, for each event.

For a sample of N events, maximize $\sum_{i=1}^{\rm N\ events} \log \mathcal{L}(p_i^{\rm vis}|(M_{Z'}, g_{Z'}))$.

Now lets fix the mass, and find the optimal value of the coupling:



Not successful

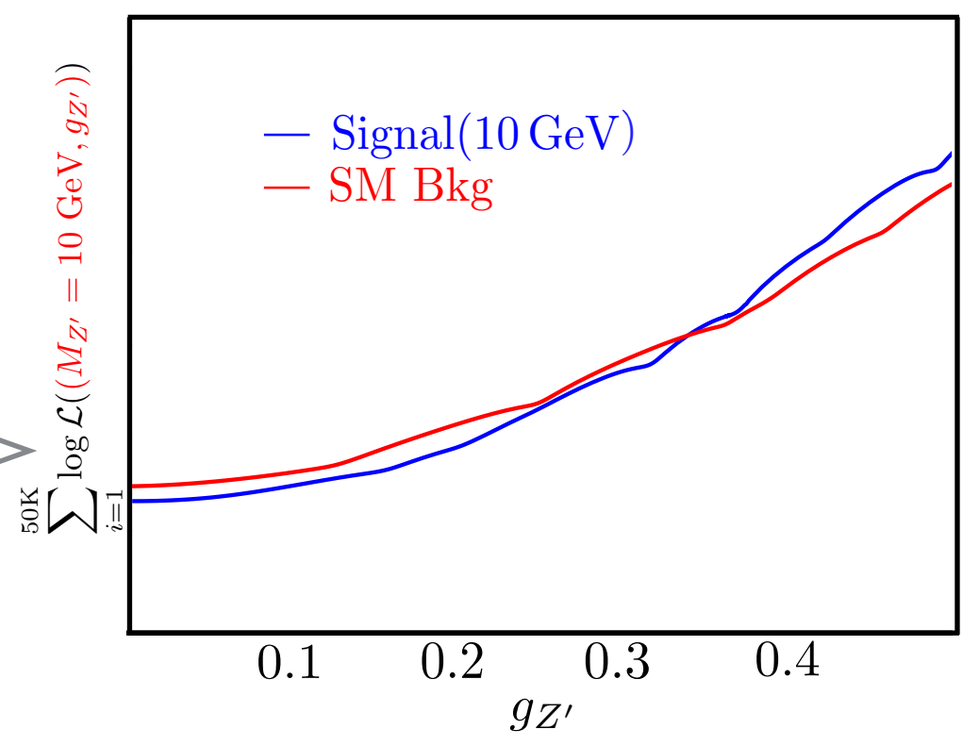
Determining the model parameters to maximize signal likelihood

Maximize $\mathcal{L}(p^{\rm vis}|(M_{Z'}, g_{Z'})) = \dfrac{P(p^{\rm vis}|(M_{Z'}, g_{Z'}))}{P(p^{\rm vis})}$, for each event.

For a sample of N events, maximize $\sum_{i=1}^{\rm N\ events} \log \mathcal{L}(p_i^{\rm vis}|(M_{Z'}, g_{Z'}))$.

Now lets fix the mass, and find the optimal value of the coupling:

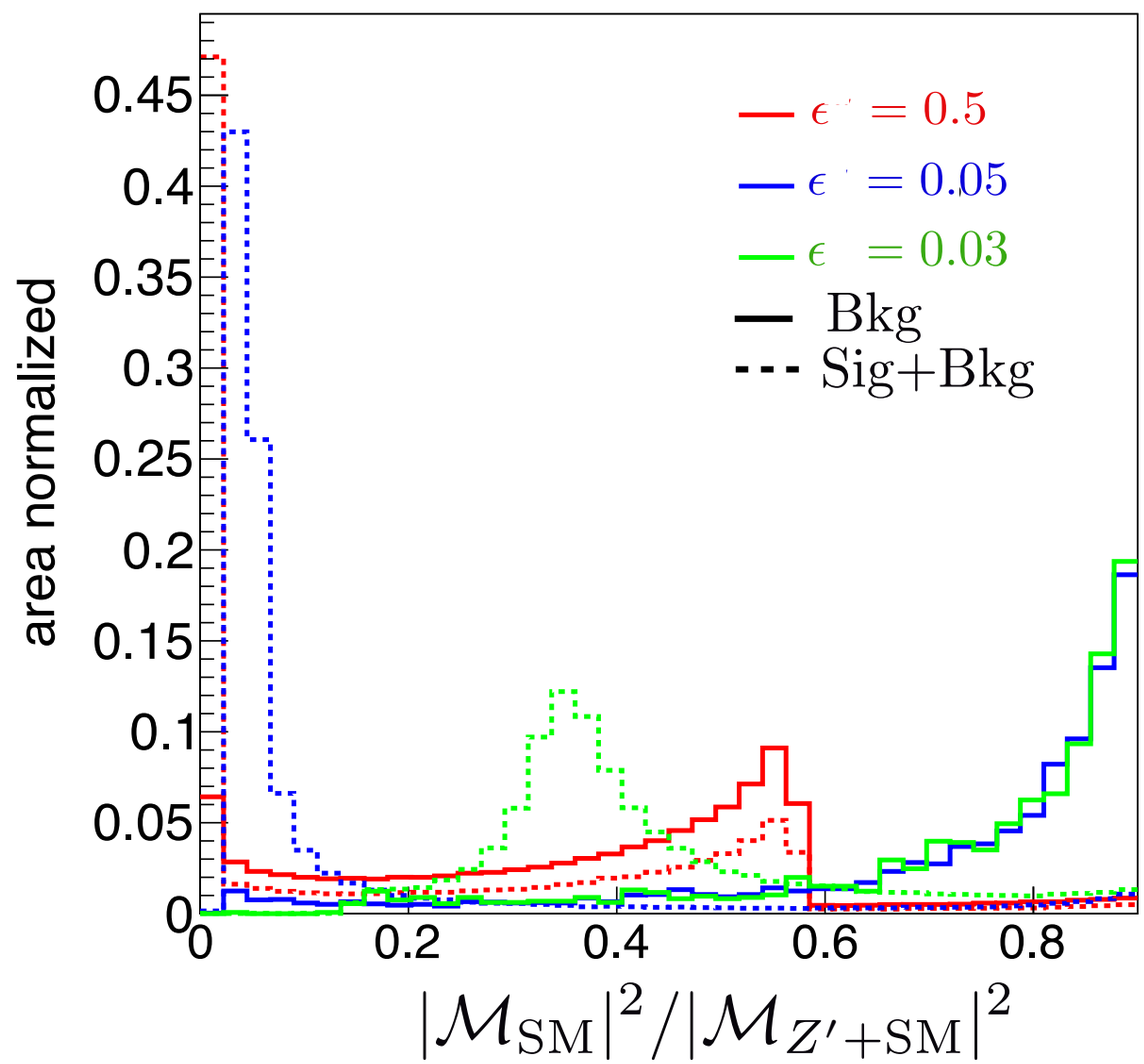
The problem is we are comparing only two numbers for each $g_{Z'}$, instead of the distribution of events.



Not successful

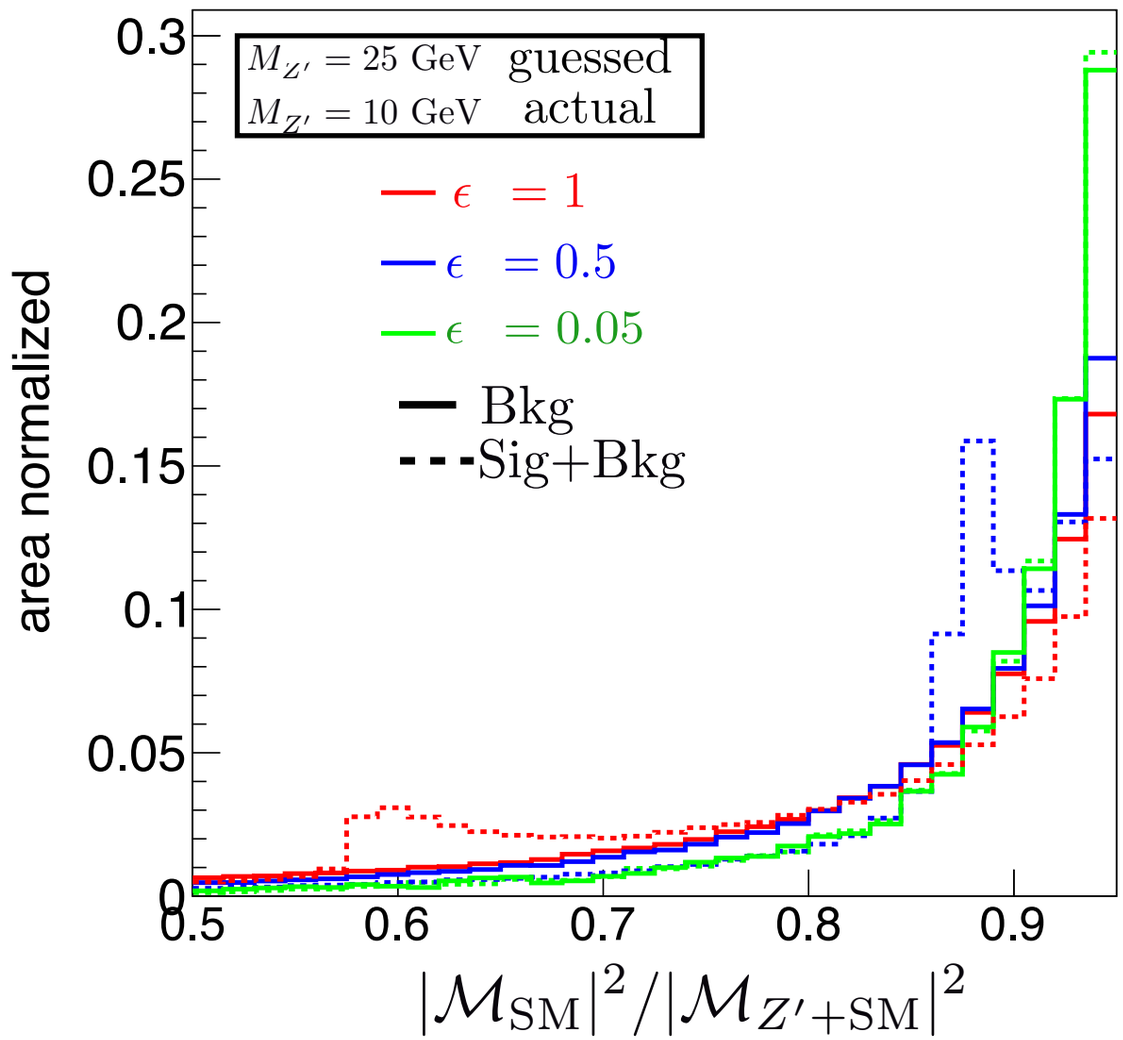
Looking at the distribution of events for various $g_{Z'} = \epsilon\, g_1$

The optimal value is $\epsilon = 0.05$, if we have guessed the mass of Z' correctly.



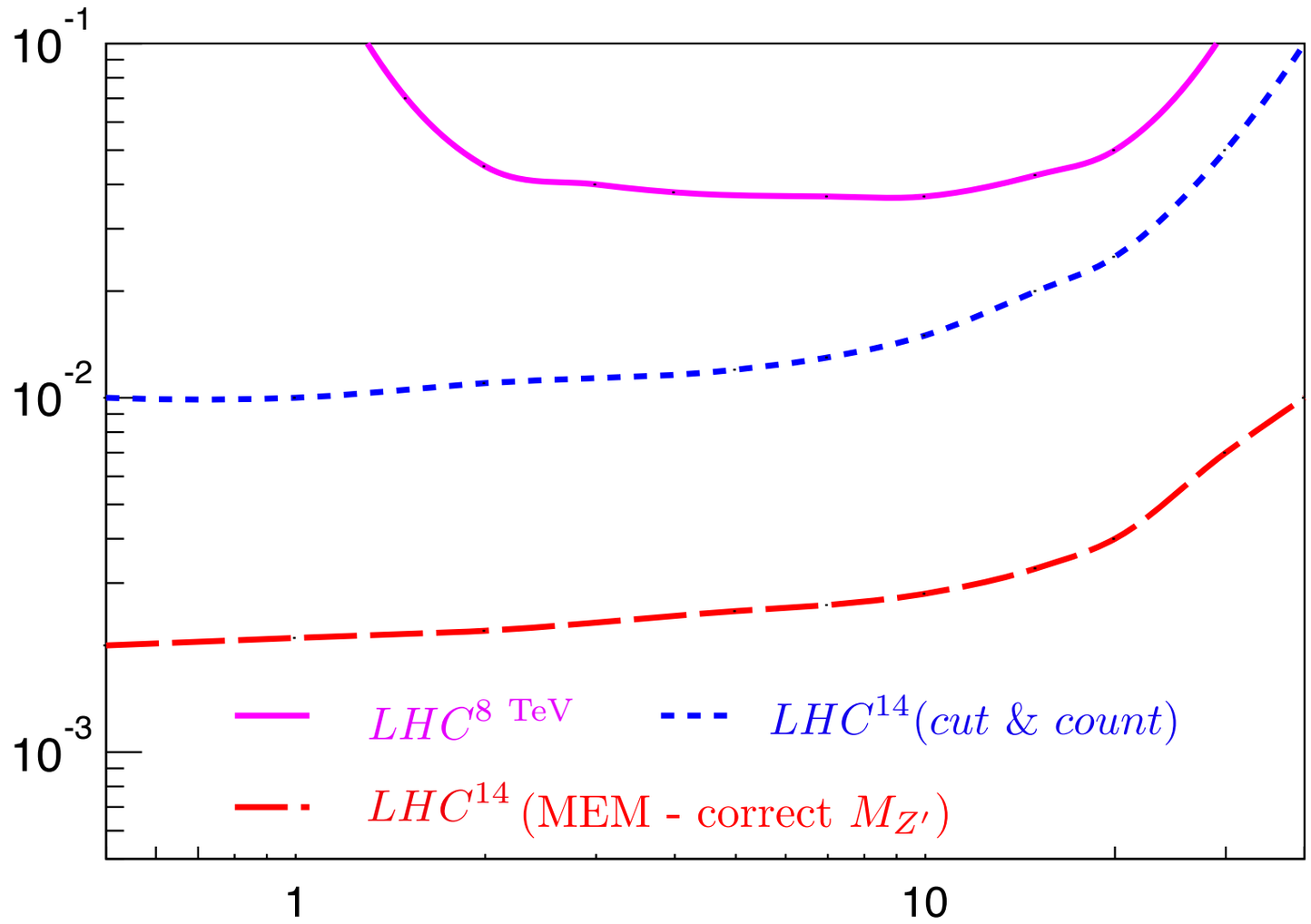
Looking at the distribution of events for various $g_{Z'} = \epsilon\, g_1$

The optimal value of ϵ is $\gtrsim 0.05$, if we guess the $M_{Z'}$ incorrectly.



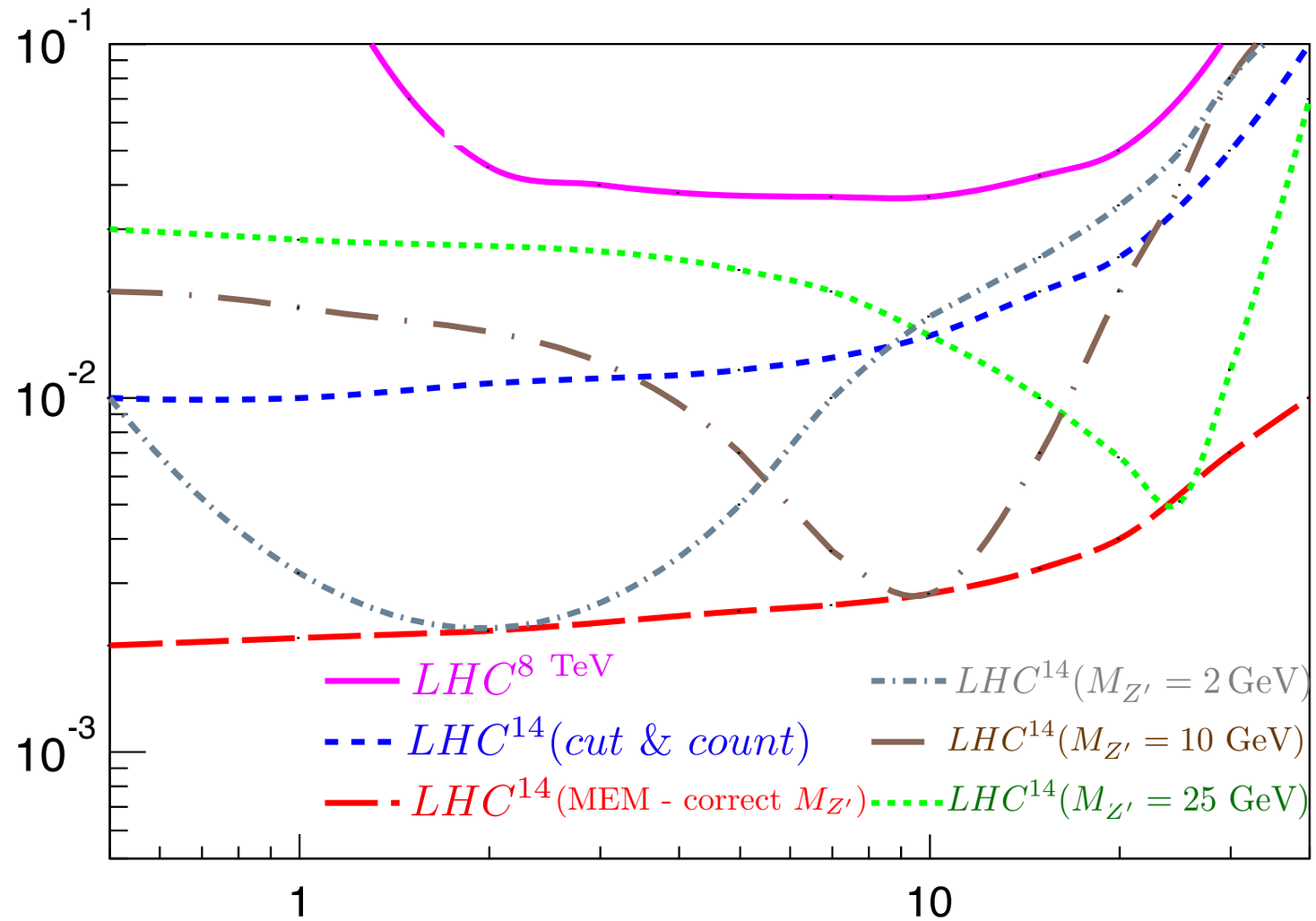
Exclusion Bounds compared to cut-and-count method

The exclusion bounds, where the luminosity of $LHC^{14\,\mathrm{TeV}}$ is assumed to be $300\,\mathrm{fb}^{-1}$.



Exclusion Bounds compared to cut-and-count method

The exclusion bounds, where the luminosity of $LHC^{14\,\mathrm{TeV}}$ is assumed to be $300\,\mathrm{fb}^{-1}$.



Contours if we fix the $M_{Z'}$ to a fixed value.

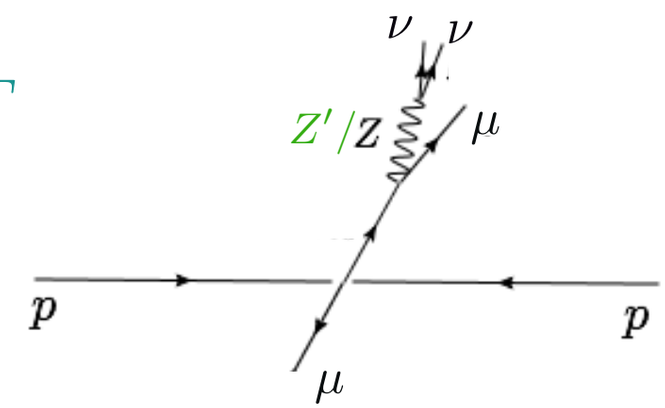
How about $M_{Z'} < 2m_\mu$?

Want Z' to be on-shell.

$$\mathrm{Br}(\mathrm{Z}' \to \nu_\ell \nu_\ell) = 100\%$$

no photon background.

$$pp \to \mu\mu \not{E}_T$$

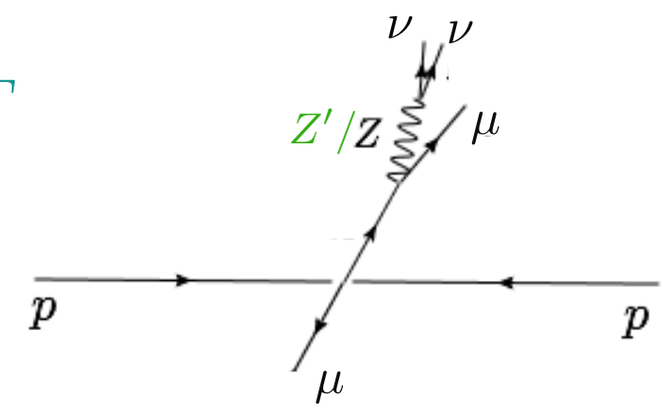


How about $M_{Z'} < 2m_\mu$?

Want Z' to be on-shell.

$$\mathrm{Br}(\mathrm{Z}' \to \nu_\ell \nu_\ell) = 100\%$$

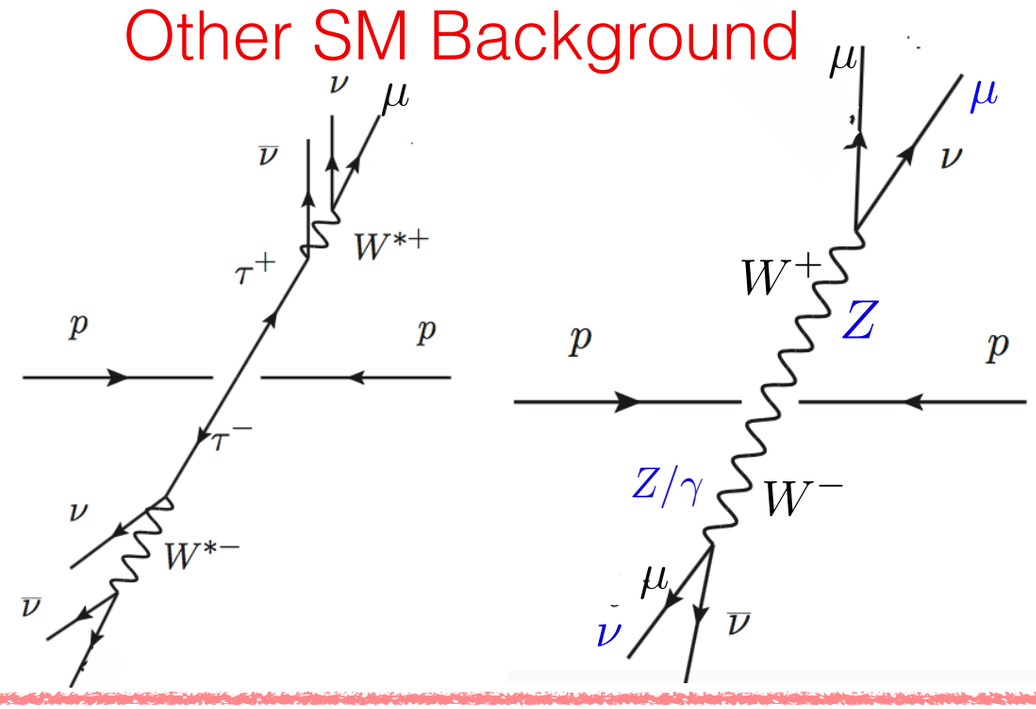
$$pp \to \mu\mu \not{E}_T$$



But new backgrounds emerge:

$$pp \to \tau^+\tau^-\big|_{\text{dilepton decay}}$$
$$pp \to W^{+*}W^{-*}\big|_{\text{dilepton decay}}$$
$$pp \to Z^*\,(Z^*/\gamma^*)$$
$$pp \to \mu^+\mu^- + \text{jets}$$

Other SM Background

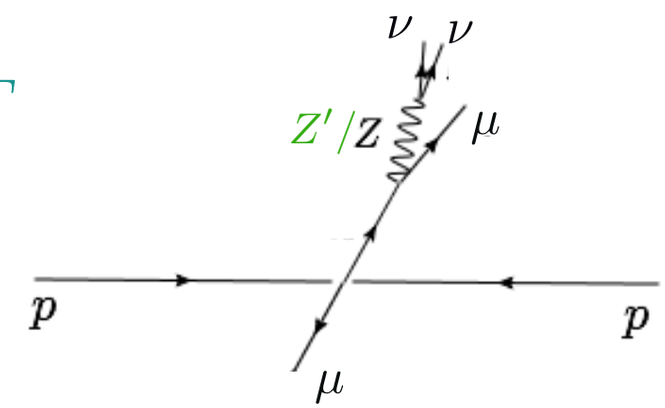


How about $M_{Z'} < 2m_\mu$?

Want Z' to be on-shell.

$$\mathrm{Br}(\mathrm{Z}' \to \nu_\ell \nu_\ell) = 100\%$$

$$pp \to \mu\mu \not{E}_T$$



But new backgrounds emerge:

$$pp \to \tau^+\tau^-\big|_{\text{dilepton decay}}$$

$$pp \to W^{+*}W^{-*}\big|_{\text{dilepton decay}}$$

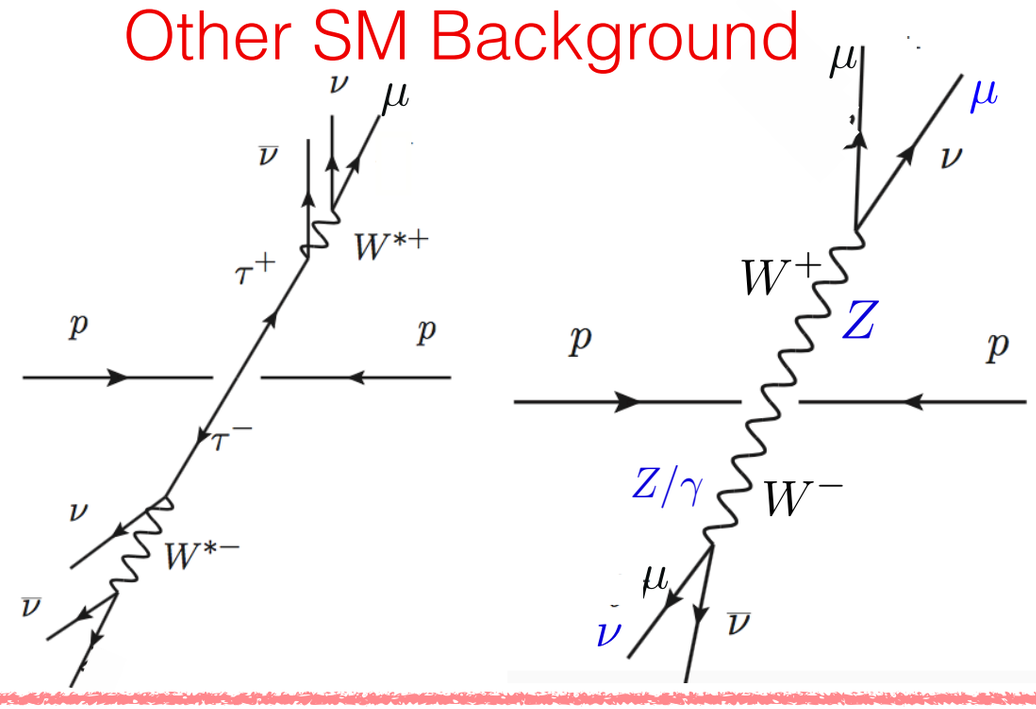
$$pp \to Z^*\,(Z^*/\gamma^*)$$

$$\boxed{pp \to \mu^+\mu^- + \text{jets}}$$

Require high MET

$$\not{E}_T > 20\,\text{GeV}$$

Other SM Background

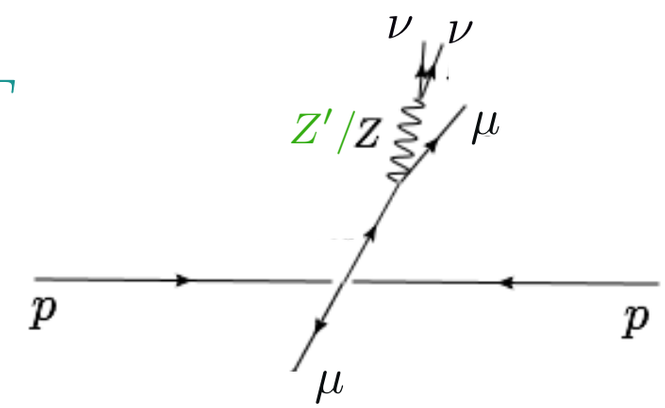


How about $M_{Z'} < 2m_\mu$?

Want Z' to be on-shell.

$$\mathrm{Br}(\mathrm{Z}' \to \nu_\ell \nu_\ell) = 100\%$$

$$pp \to \mu\mu \not{E}_T$$



But new backgrounds emerge:

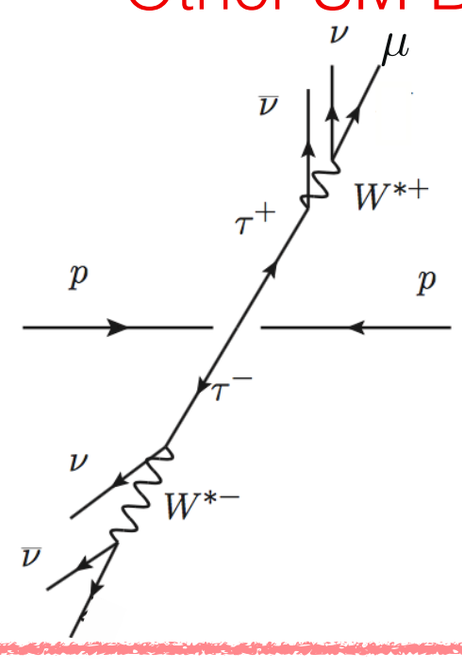
$$pp \to \tau^+\tau^-\big|_{\text{dilepton decay}}$$

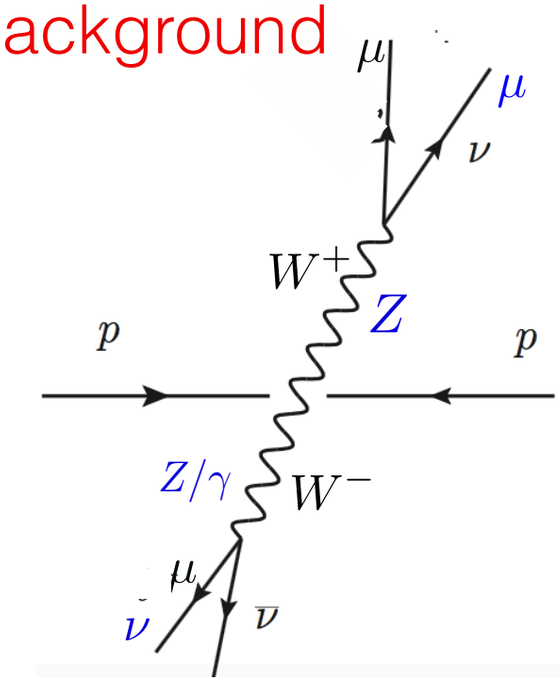
$$pp \to W^{+*}W^{-*}\big|_{\text{dilepton decay}}$$

$$pp \to Z^* \, (Z^*/\gamma^*)$$

$$pp \to \mu^+\mu^- + \text{jets}$$

Other SM Background



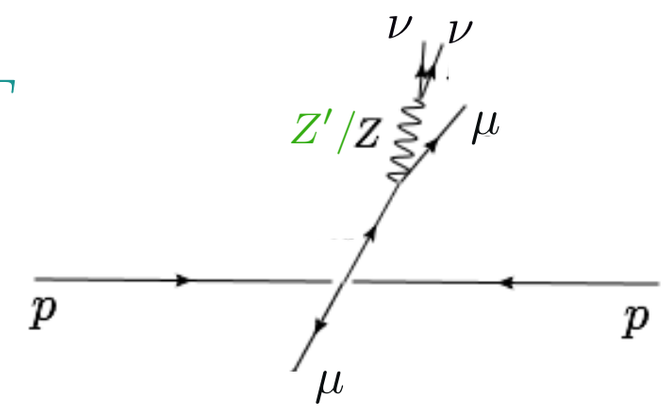


How about $M_{Z'} < 2m_\mu$?

Want Z' to be on-shell.

$$\mathrm{Br}(\mathrm{Z}' \to \nu_\ell \nu_\ell) = 100\%$$

$$pp \to \mu\mu \not{E}_T$$



But new backgrounds emerge:

$$pp \to \tau^+\tau^-\big|_{\text{dilepton decay}}$$

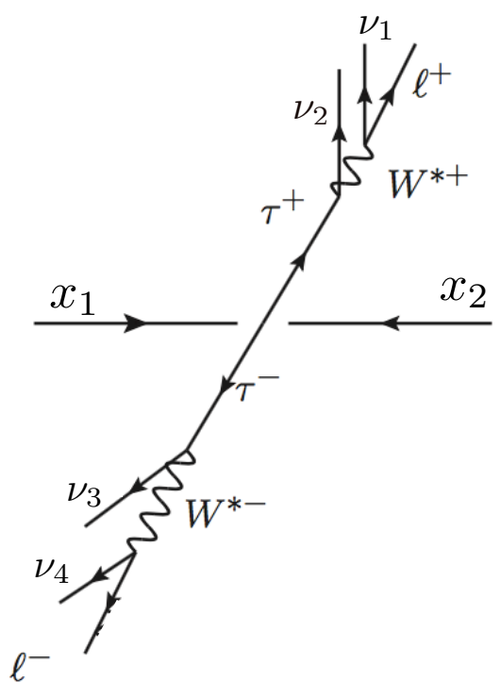
Let's use $|\mathcal{M}_{\tau\tau}|^2$ to separate $\tau^+\tau^-$ background from other processes.

Not using the likelihood ratio – computationally much faster!

Independent of $L_\mu - L_\tau$ new parameters.

Calculating $|\mathcal{M}_{\tau\tau}|^2$

Conventionally, we are supposed to integrate over un-constrained momenta.



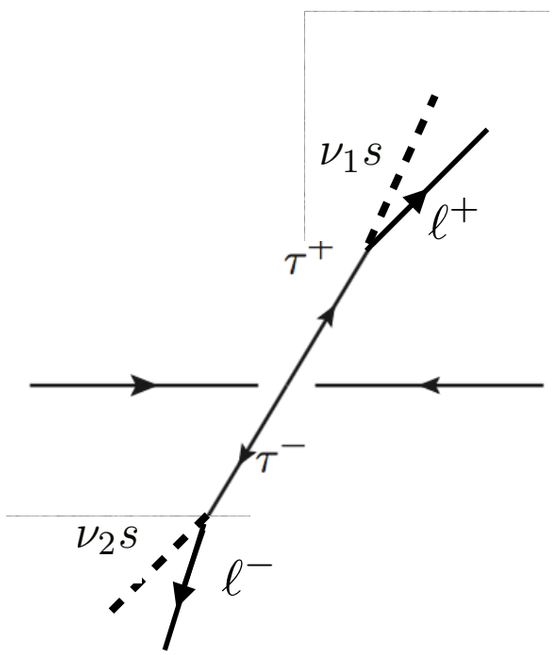
Unknown parameters	Constraints
3 momenta for ν_1	4: $\delta^4(p_i - p_f)$
3 momenta for ν_2	$(\ell^+ + \nu_1 + \nu_2)^2 = m_\tau^2$
3 momenta for ν_3	$(\ell^- + \nu_3 + \nu_4)^2 = m_\tau^2$
3 momenta for ν_4	
x_1	
x_2	

number of integrals: 8

Too many!!

Calculating *modified* $|\mathcal{M}_{\tau\tau}|^2$

Instead, let's look at



Unknown parameters	Constraints
4 momenta for $\nu_1 s$	4: $\delta^4(p_i - p_f)$
4 momenta for $\nu_2 s$	$2 : (\ell + \nu s)^2 = m_\tau$
x_1	$2 :$ *collinear* $\eta_{\nu s} = \eta_\ell$
x_2	$1 : m_{\tau\tau} = m_Z$
	$1 : p_T^{\ell^+} + p_T^{\nu_1 s} = p_T^{\ell^-} + p_T^{\nu_2 s}$

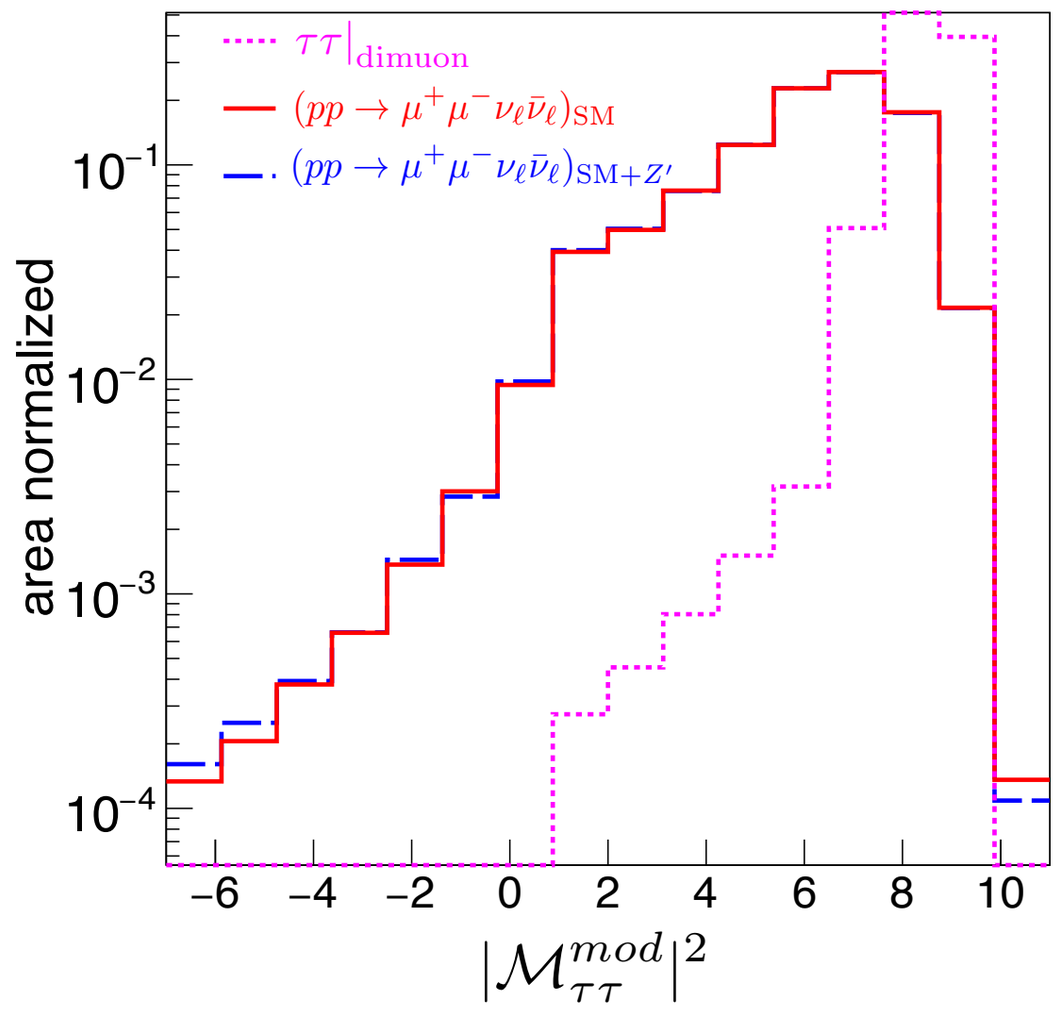
no integration

If the taus are produced from on-shell Z boson, they are very boosted, causing the neutrinos to be almost collinear with muons.

Eliminating the $\tau\tau$ background

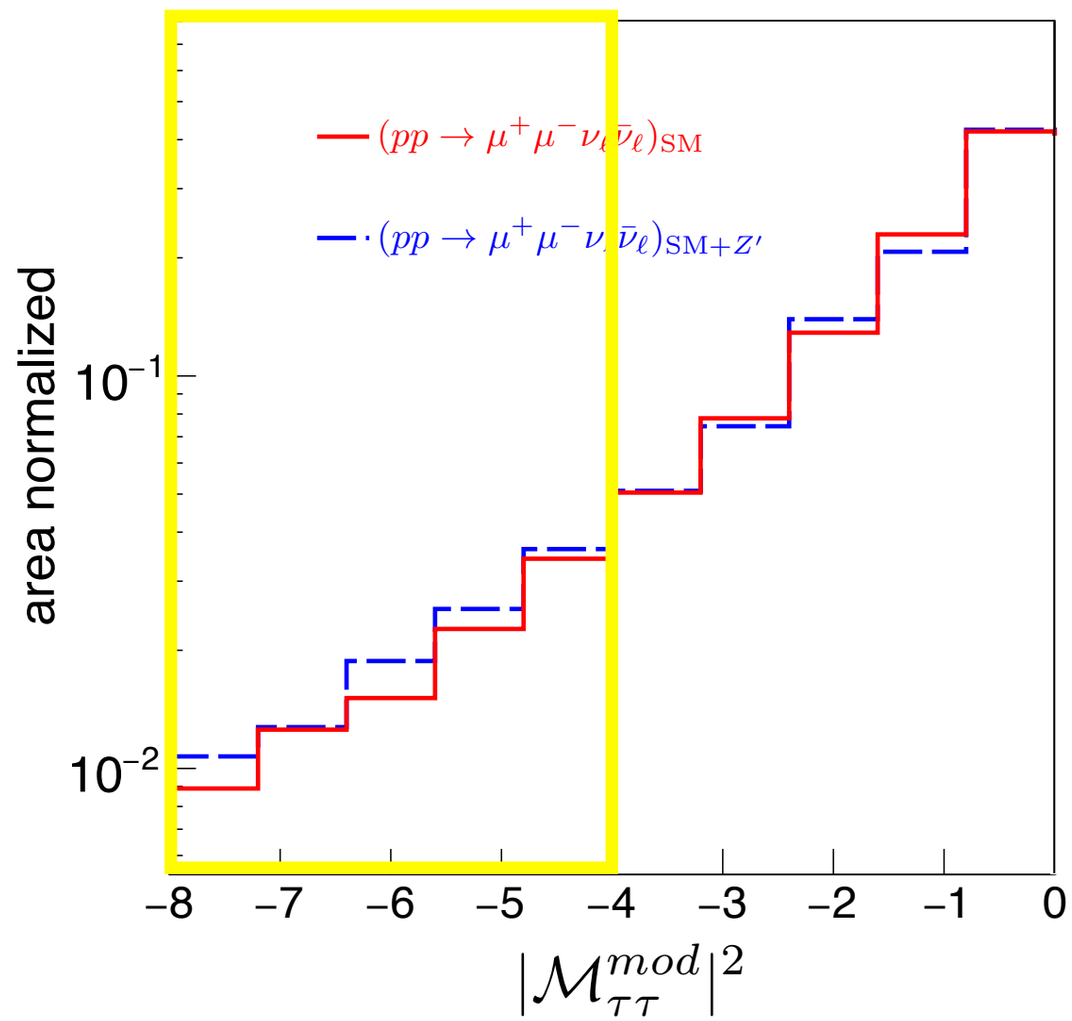
Because the assumptions were chosen based on $\tau^+\tau^-|_{\text{dimuon}}$ topology and were not reasonable in other processes, we have $|\mathcal{M}_{\tau\tau}^{mod}|^2 < 0$ for part of the distribution of other processes.

for example, we have have negative momentum



Zooming in $|\mathcal{M}_{\tau\tau}|^2 < 0$

There is an excess of the signal over background at at large, negative $|\mathcal{M}_{\tau\tau}^{mod}|^2$.



Understanding the excess

$$|\mathcal{M}_{\text{signal}}|^2 = \boxed{\text{only } Z' \text{ contribution}} + \text{ interference} + \text{ only SM contribution}$$

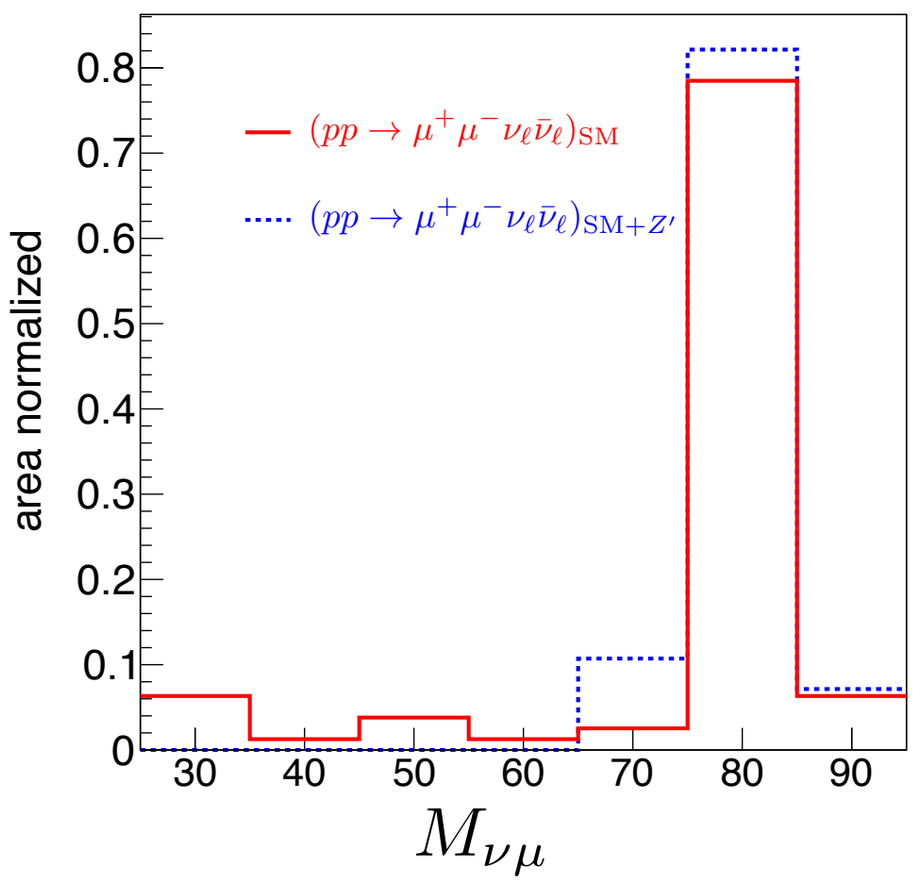
Similar to $\tau\tau$ process and thus is negligible after removing $\tau\tau$

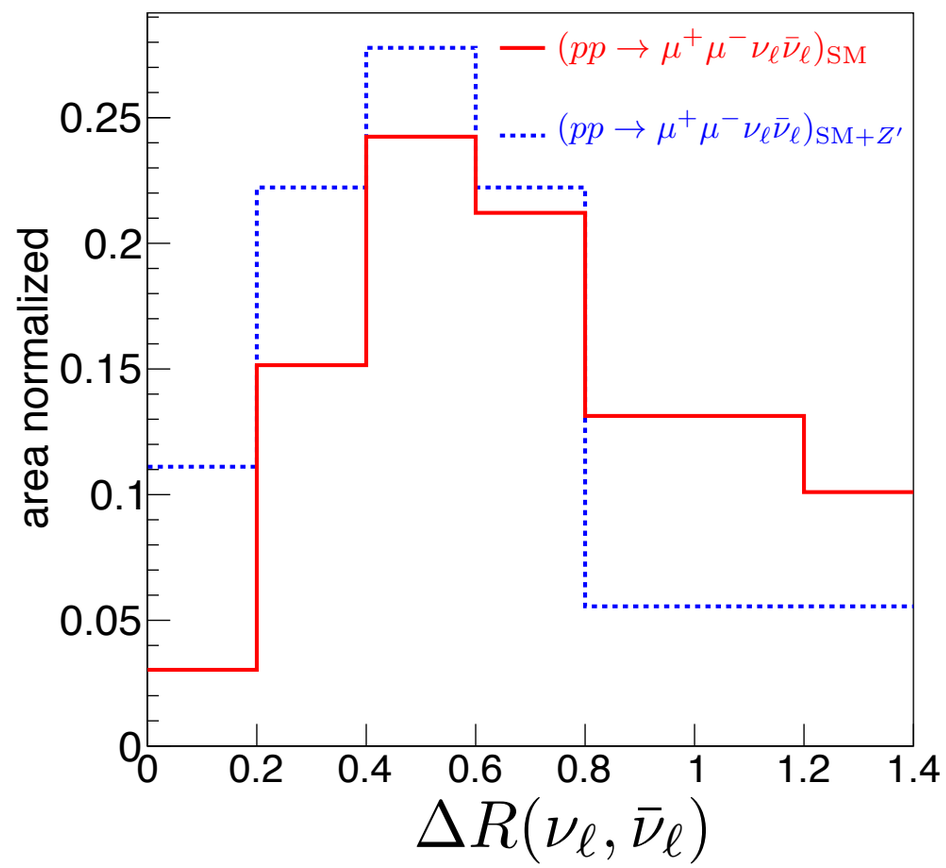
Understanding the excess

$$|\mathcal{M}_{\text{signal}}|^2 = \text{ only } Z' \text{ contribution } + \boxed{\text{interference}} + \text{ only SM contribution}$$

Important when SM piece or Z' piece or both are significant

The interference between WW background and Z'

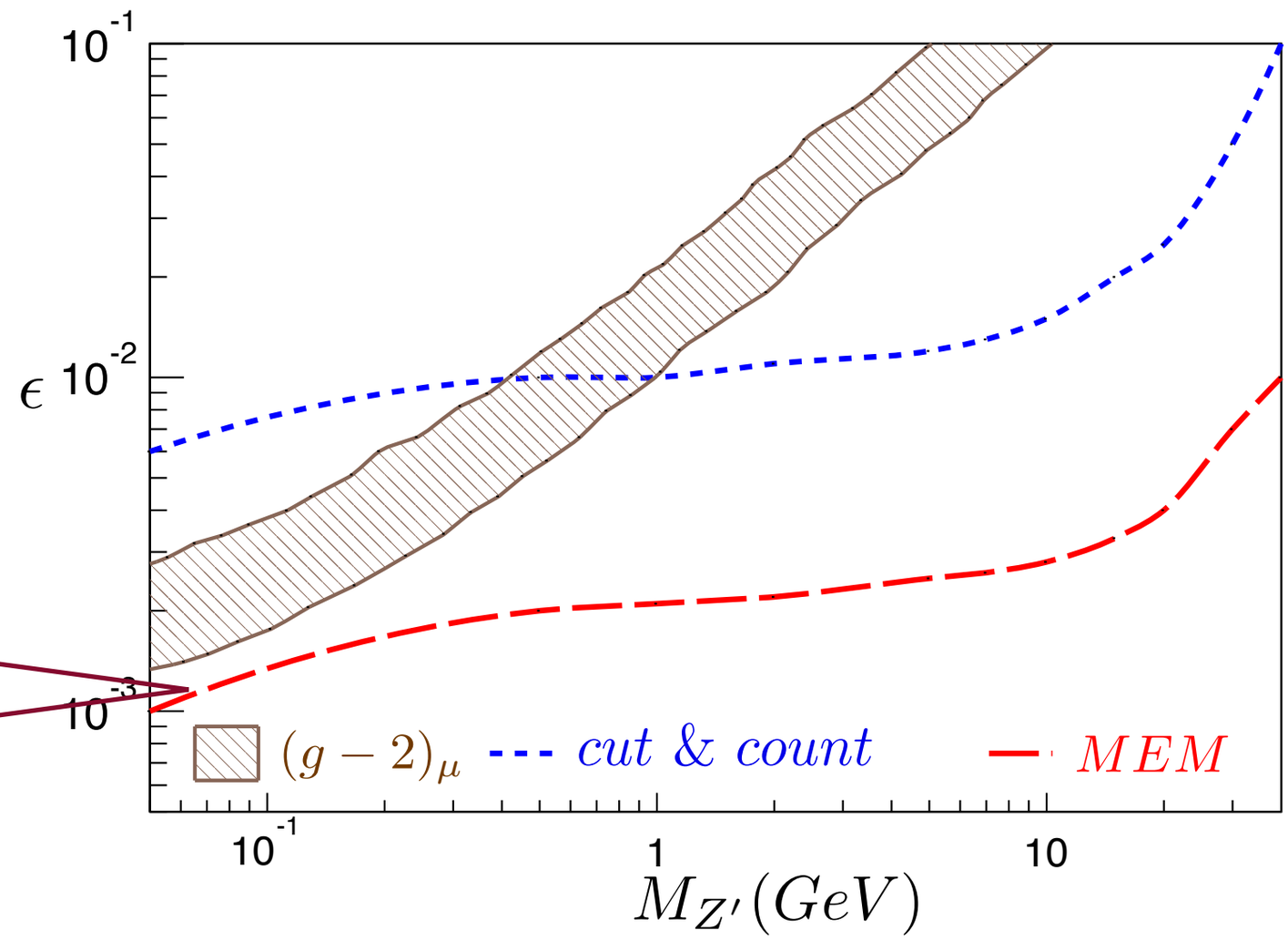




Exclusion Bounds compared to cut-and-count method

The exclusion bounds, with $300\,\mathrm{fb}^{-1}$ integrated luminosity and the $(g-2)_\mu$ band.

Independent of $L_\mu - L_\tau$ model parameters



Conclusion

- Matrix Element Method (MEM) uses all of the available kinematic information in the most optimized way.
- MEM is computationally challenging, and so it is not extensively used in BSM searches.
- We applied MEM to one simple and well-motivated extension of the SM $L_\mu - L_\tau$, where some of the challenges of the MEM are present.
- We can “modify” the MEM approach, making it computationally much faster, while maintaining a relatively high signal-background discrimination.
- Our sensitivity with MEM is enhanced by an order of magnitude compared with the cut-and-count method.

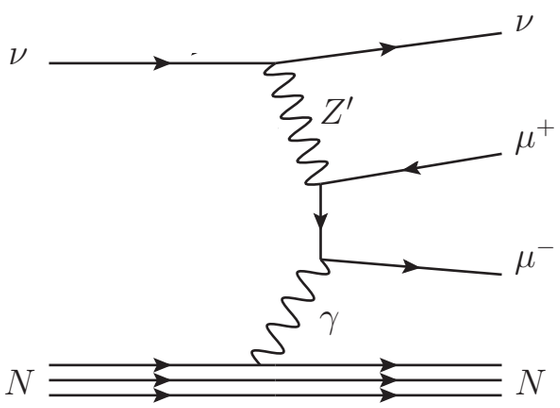
Conclusion

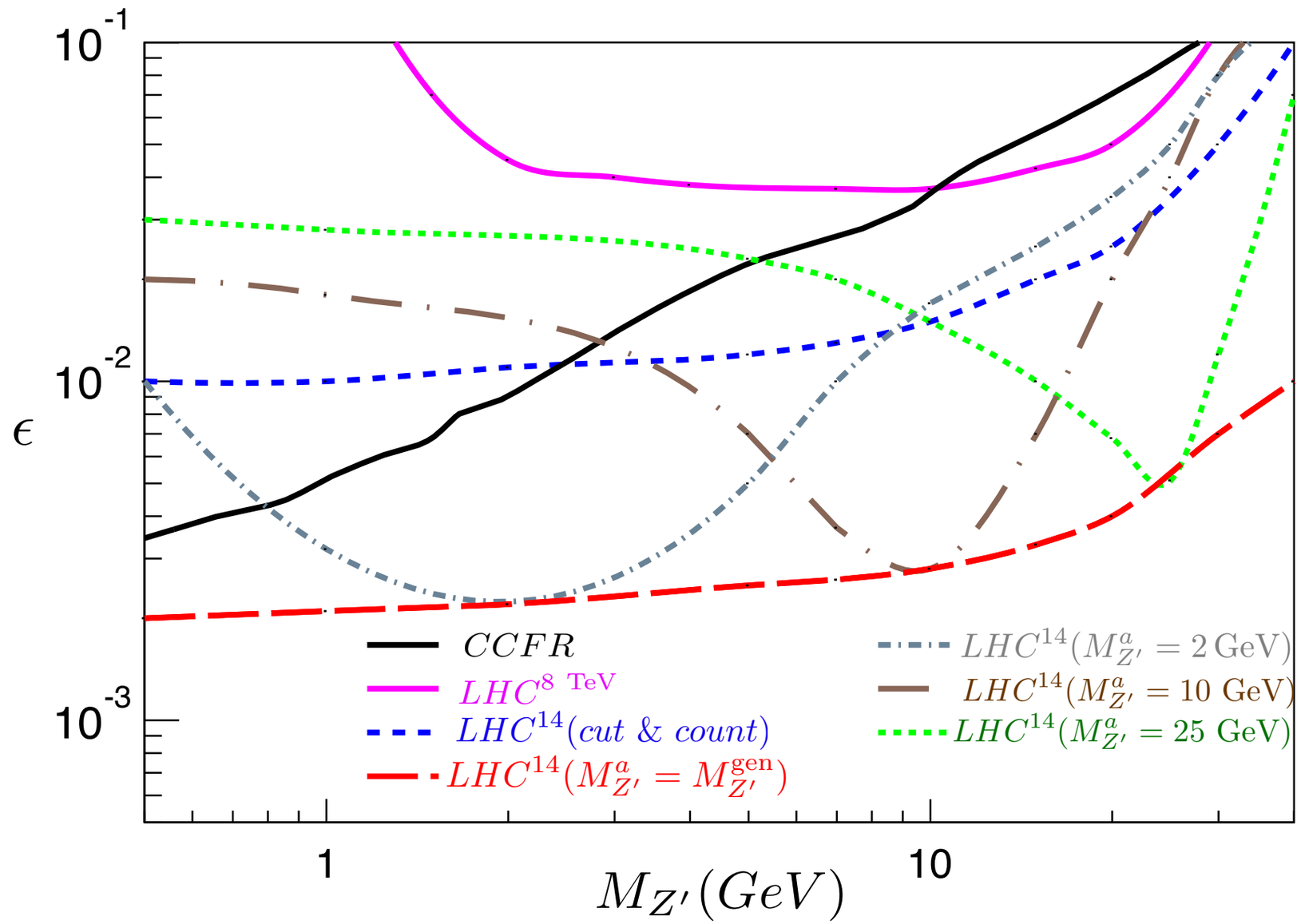
- Matrix Element Method (MEM) uses all of the available kinematic information in the most optimized way.
- MEM is computationally challenging, and so it is not extensively used in BSM searches.
- We applied MEM to one simple and well-motivated extension of the SM $L_\mu - L_\tau$, where some of the challenges of the MEM are present.
- We can “modify” the MEM approach, making it computationally much faster, while maintaining a relatively high signal-background discrimination.
- Our sensitivity with MEM is enhanced by an order of magnitude compared with the cut-and-count method.

Thank you!

back up

CCFR


ν
ν
Z'
μ^+
μ^-
γ
N
N


10^{-1}
10^{-2}
10^{-3}
ϵ
1
10
$M_{Z'}(GeV)$
$CCFR$
$LHC^{8\ \mathrm{TeV}}$
$LHC^{14}(cut\ \&\ count)$
$LHC^{14}(M^a_{Z'} = M^{\mathrm{gen}}_{Z'})$
$LHC^{14}(M^a_{Z'} = 2\,\mathrm{GeV})$
$LHC^{14}(M^a_{Z'} = 10\ \mathrm{GeV})$
$LHC^{14}(M^a_{Z'} = 25\ \mathrm{GeV})$