The Search for Contact Interactions in Dilepton Channels at the Compact Muon Solenoid

Daniel Nicolas Salazar-Gallegos

University of Illinois Urabana-Champaign

Fermi National Accelerator Laboratory

Summer Internships in Science and Technology

Supervisors:

Dr. Pushpalatha Bhat

Dr. Leonard Spiegel

Fermi National Accelerator Laboratory

August 17, 2017

Abstract

Presently, the Standard Model (SM) states that there exists three generations of quarks and leptons. Each generation is distinguished by the fermion rest mass where higher generations have higher masses. Under the SM, $q\bar{q} \rightarrow \mu\bar{\mu}$ follow the Drell-Yan (DY) process. If the process were to deviate from the SM and agree with a Contact Interaction (CI) model, it would be the first evidence for quark/lepton compositeness. This implies quarks and leptons are composed of yet more elementary particles, which are sometimes called preons. Using Monte Carlo simulation we can demonstrate the differences between the DY and CI models. Specifically, the invariant mass spectra of dilepton events and the Collins-Soper angle distribution show significant variation between the two models. The simulation results for the models can be compared to data and, in the absence of evidence for CI, higher limits can be set for the energy threshold for compositeness. In this paper, we present Monte Carlo simulation studies and effects of some systematic measurement uncertainties as the results.

I. Introduction

According to the SM and is verified in past experiments, the most elementary matter particles are the quarks and leptons. As far as data has shown, these particles appear to be point like with no current evidence of compositeness. However, the higher energies at the LHC for 2016 allow for a reevaluation of this conclusion. If there exists particles that are more elementary than quarks and leptons, hereto referred to as preons, evidence of this substructure would begin to emerge in the detectors at the LHC.

One important question compositeness theories seeks to answer is why there are multiple generations of quarks and leptons. For just about every measure, every generation after the first in the SM is similar. Charges, and spins do not vary between generations. However, mass does increase after the first generation. When this phenomenon was observed with hadrons, their compositeness was proposed and later discovered as quarks. In the case of quark and lepton compositeness, each generation would then be an excited state of the first generation. Additionally, the discovery of additional generations may also strongly suggest the existence of preons.

The CMS detector is capable of detecting signals for a range of Contact Interaction models if they exist. The $q\bar{q} \rightarrow \mu^+\mu^-$ process specifically has a signal that arises notably from the Drell-Yan background. Other backgrounds are smaller and are not studied here. For the SM, this process is characterized by the two quarks annihilating to a photon or Z-boson, then decaying into a lepton-antilepton pair. The CI model has the same starting and final products, however we know from previous analysis that the propagator must be incredibly massive. This energy value, referred to as Λ, is the energy scale where the CI signal clearly shows deviation of data from the Drell-Yan background. The cross section of our signal can be described as

$$\frac{d\sigma^{CI/DY}}{dM_{\mu\mu}} = \frac{d\sigma^{DY}}{dM_{\mu\mu}} - \eta_{ij}\frac{I}{\Lambda^2} + \eta_{ij}^2\frac{C}{\Lambda^4} \tag{1}$$

where the first term is the Drell-Yan background and the last two terms are CI terms. η is a multiplier with values of either -1 or +1 to include the possibilities of contact interactions constructively or destructively interfering with the Drell-Yan background. The final term is a contact interaction only term that will dominate at higher collision energies. We can also observe from this equation that as Λ approaches an arbitrarily large value, the signal approaches the SM.

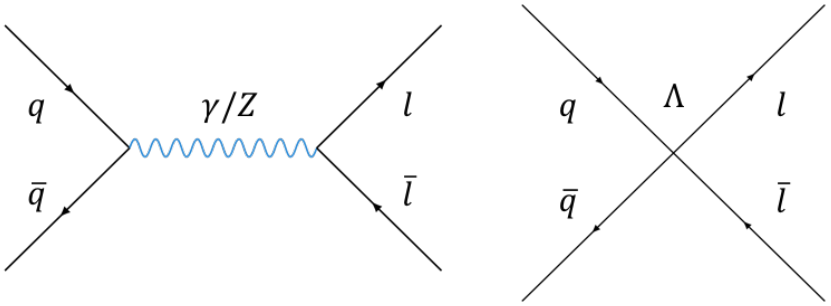

Figure 1: (Left) SM, Drell-Yan process of the channel being investigated. It is the primary background and defined by the annihilation of a quark-antiquark pair into a photon/Z-boson then decaying into a lepton-antilepton pair. (Right) Contact Interaction. The energy scale Λ is large and the mediator so massive that by approximation the propagator can be ignored. The value for Λ is where new physics emerges! [**?**]

The interactions from the helicities of the incoming quarks must also be considered in the CI[**?**]. The most general description of which is

$$\begin{aligned}\mathcal{L}_{ql} =& \frac{g_0^2}{\Lambda^2}[\eta_{LL}(\bar{q}_L\gamma^\mu q_L)(\bar{\mu}_L\gamma_\mu\mu_L) \\ &+ \eta_{LR}(\bar{q}_L\gamma^\mu q_L)(\bar{\mu}_R\gamma_\mu\mu_R) \\ &+ \eta_{RL}(\bar{u}_R\gamma^\mu u_R)(\bar{\mu}_L\gamma_\mu\mu_L) \\ &+ \eta_{RL}(\bar{d}_R\gamma^\mu d_R)(\bar{\mu}_L\gamma_\mu\mu_L) \\ &+ \eta_{RR}(\bar{u}_R\gamma^\mu u_R)(\bar{\mu}_R\gamma_\mu\mu_R) \\ &+ \eta_{RR}(\bar{d}_R\gamma^\mu d_R)(\bar{\mu}_R\gamma_\mu\mu_R)]\end{aligned} \tag{2}$$

Helicity is defined as the projection of angular momentum in the particles direction of motion. A right handed particle, distinguished by subscript R, is a particle where the projection of spin is in the direction of motion. A left handed particle, distinguished by subscript L, is a particle where the projection of spin is opposite the direction of motion. In this study, the helicity states of the quarks and letptons that are probed are: Left-Left (LL), Right-Left (RL), and Right-Right (RR).

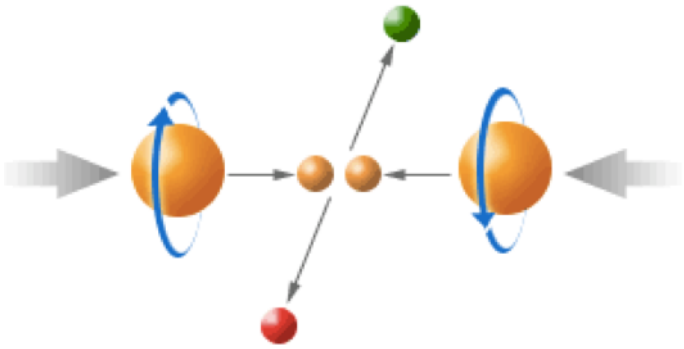

Figure 2: Example of a Left-Left handed particle collision

Presently, there is no evidence for contact interactions. CMS has produced limits on Λ from Run 1 and 2 which are as follows.

Item	Run I	Run II
Luminosity	5.7fb^{-1}	36.3fb^{-1}
θ_{CS}	No	Yes

Dilepton	Λ_C (TeV)	Λ_D (TeV)
e^+e^-	18.3	13.5
$\mu^+\mu^-$	15.2	12.0

Table 1: Current Limits on Λ according to CMS using data from Run 1 and 2.[**?**]

The lack of evidence may be due to the LHC and other particle colliders not having reached high enough energies in order to observe contact effects. In order to observe contact interactions, collisions must approach the binding energies of the preons within quarks and leptons. However,

it is possible that, again, no evidence will be found for contact interactions. If this is the case, higher limits will be placed on Λ.

This year's summer project was focused on developing the analysis tools and workflow that will be used for the Contact Interaction analysis. This includes producing Monte Carlo (MC) samples for $\Lambda = 22$ TeV, developing and validating the analyzers that extract the data we need from the MC, and producing histograms of the data. Additionally, statistical studies of data and detector resolutions and acceptance were done on the simulations in order to develop our combination tools so that once the MC samples are prepared, data analysis can begin.

II. Generation of Monte Carlo Samples

I. Monte Carlo Simulation Workflow

In order to have a thorough analysis, simulation of the events must be done. PYTHIA 8 is the Monte Carlo event generator used. It uses the theory of the SM and of CI physics and produces events of the specific sample. These are summarized in Table 2. These events are then passed into a detector simulator which simulates the CMS detector and pileup that occurs within. Due to the higher luminosities in the LHC, for each simulated collision, there can be several proton-proton collisions per event. This is what is simulated in pileup. After pileup, the event files are funneled into AOD then MiniAOD to reduce the size of the files.

Afterward, they are ntupled with the relevant information needed for the analysis. Next, the data is extracted and put into histograms. Lastly, the bin counts are extracted and funneled into a combine tool that will use Bayesian statistics to determine if there is a signal, or, if not, set new limits on Λ.

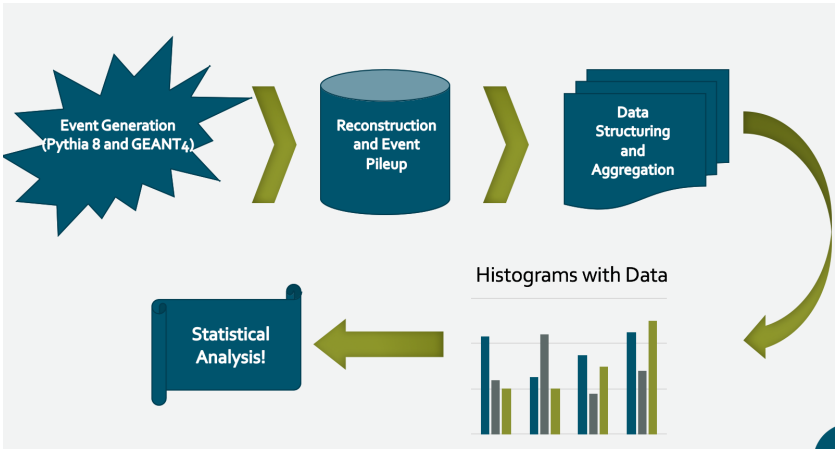

Figure 3: Flow chart demonstrating the analysis work flow. The primary analysis software used was ROOT. The Higgs Combine Tool will be used for statistical analysis.

Parameter type	Parameter value
Particle	e, μ
Helicities	LL, LR, RR
Λ (TeV)	10, 16, 22, 28, 100k
η	± 1
Mass cuts (GeV)	300, 800, 1300, 2000*

Table 2: Summary of samples that were privately generated. Each permutation of parameters contains 50,000 events. The 2000 GeV mass cut samples were generated only for the Left-Left constructive samples.

II. MC Samples and Distributions

II.1 Dilepton Invariant Mass Spectrum

The particular channel that will be probed has a lepton-antilepton pair in the final state. Therefore it makes sense to make observations of the invariant mass of the two leptons. In the Drell-Yan process, the invariant mass will contain a resonance peak around the rest mass of the Z-boson and a large cross section below that peak. If there exists another propagator with a rest mass of Λ, then there should be an additional resonance peak beyond the Z peak. However, if there is compositeness, then there will be a broad exess of events over the DY background at high lepton masses. Monte Carlo simulations demonstrate that for Contact Interactions and low Λ, the invariant mass spectrum will rise noticeably above the SM background.

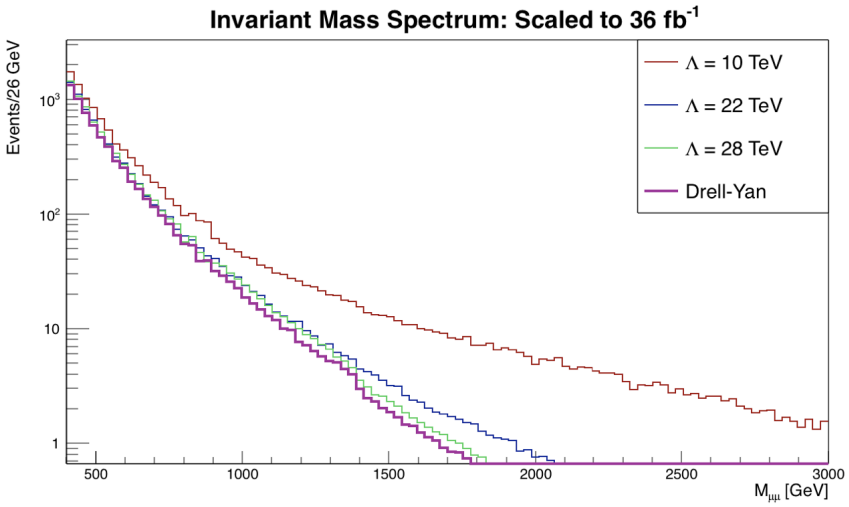

Figure 4: Dilepton invariant mass MC simulation of various Λ values. Displayed for CI are samples taken with Left-Left helicities with constructive interference with DY. For smaller Λ the distribution rises well above the DY background at high mass. This difference decreases as Λ becomes larger.

II.2 Collins-Soper Frame Angle

The SM, Drell-Yan process also predicts a generally non-isotropic distribution of the Collins-Soper frame angle. The Collins-Soper frame angle (θ_{CS}) is defined as the angle the negatively charged lepton make with the beam line in center of mass frame of the dileptons. Demonstrated in Figure 5. For most helicities, this is also the case. The Left-Right CI is the only model that will produce a more isotropic distribution of θ_{CS}. The cosine of this frame angle is defined as **[?]**

$$cos(\theta_{CS}) = \frac{k_z(l^+l^-)}{|k_z(l^+l^-)|} \frac{2(k_1^+k_2^- - k_1^-k_2^+)}{m(l^+l^-)\sqrt{m(l^+l^-)^2 + k_T(l^+l^-)^2}} \tag{3}$$

where $\vec{k}$ is the momentum 4-vector of one of the leptons. k_z is therefore defined as the momentum of the lepton in the z direction. $k_{1,2}^{\pm}$ is defined as

$$k_{1,2}^{\pm} \equiv \frac{E_{1,2} \pm k_{(1,2)z}}{\sqrt{2}} \tag{4}$$

k_1 corresponds to the negative lepton and k_2 corresponds to the positive lepton. These defined quantities allow the θ_{CS} to be expressed in terms of easy to measure kinematic variables. This is because the quantity $(k_1^+k_2^- - k_1^-k_2^+)$ is invariant under relativistic z boosts. Figure 6 demonstrates the simulated distributions of the Collins-Soper angle in terms of the different helicity models.

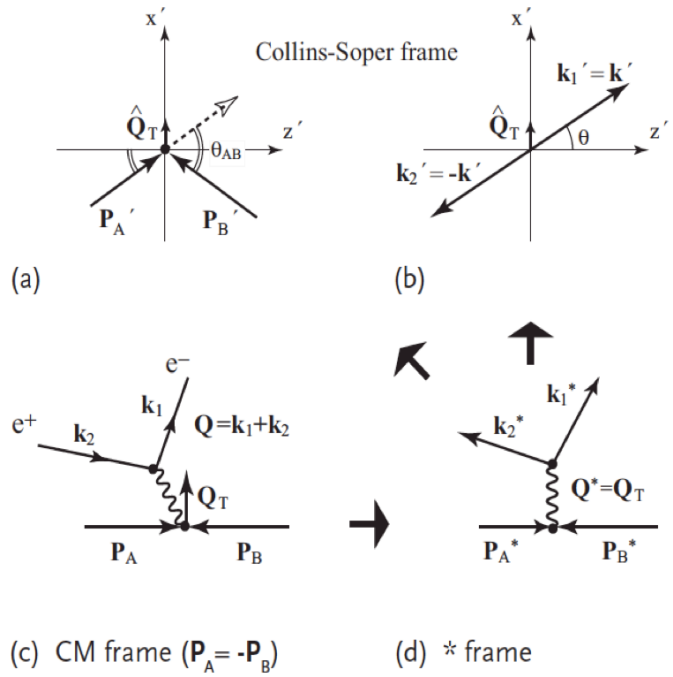

Figure 5: (a)(b): Collins-Soper frame with respect to the incoming protons and outgoing leptons respectively. θ_{CS} shown in (b) as the angle the leptons make with the z-axis. (c) Center of mass frame of the protons or the Lab frame. (d) Intermediate frame. Used for simplifying calculations in order to express the θ_{CS} in terms of lab variables. **[?]**

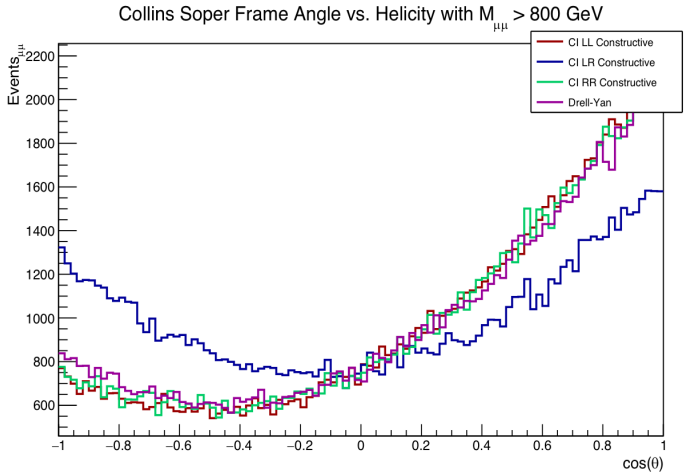

Figure 6: Collins-Soper frame angle's dependance on helicity for a minimum mass cut of 800 GeV for $\Lambda = 22$ TeV. Lower masses are dominated by the SM and will produce a non-isotropic distribution regardless of the helicity. Each CI sample was using a constructive model here. From the histogram, it is apparent that the DY is non-isotropic as well as the LL and RR models. The LR CI model produces a slightly more isotropic distribution distinct from DY.

III. Reconstructed Quantities

Once the event generator was verified, reconstruction of the events in the CMS detector was performed. The sample distributions plotted show the same expected behavior. Merged histograms were scaled according to their cross section with a luminosity of 36 fb^{-1}. Histograms in figures 7 and 8 contain the reconstructed distributions.

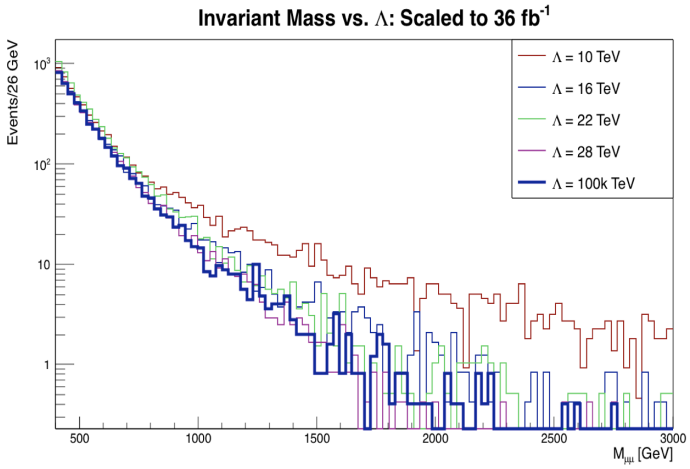

Figure 7: Reconstructed events for several Λ. 100k TeV is used as an approximation for the SM Drell-Yan process. Again, the distribution for Λ rise higher above the background than for larger values.

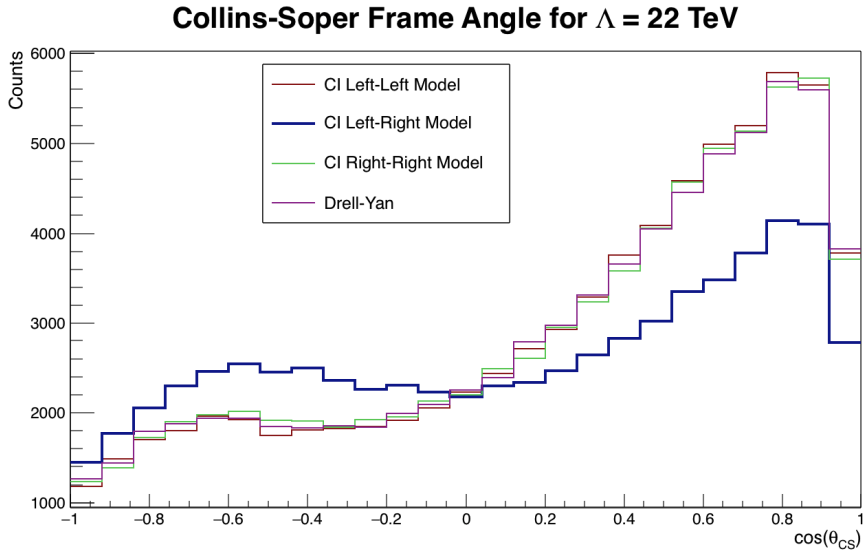

Figure 8: $cos(\theta_{CS})$ for reconstructed events for several helicities. This is similar to what was demonstrated in figure 6. The dramatic decline in events at ± 1 is a result a decrease of acceptance near the beam line. Detectors can not be placed at angles too close to the beam line.

IV. Statistical Studies

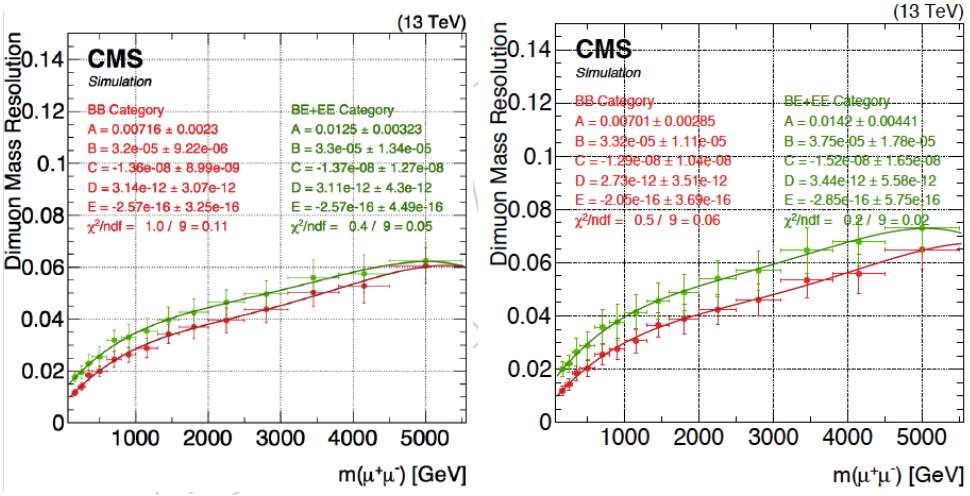

Figure 9: Figure taken CMS analysis note: [**?**].The errors of each point include both statistical and systematic uncertainties. The left plots shows the pure MC resolution and the right one shows the smeared resolution to take into account the data resolution.

In addition to development and validation of the workflow, statistical studies were done in order to understand the effect of the systematic errors in the CMS detectors. One of which was

the dilepton mass resolution as reported by CMS[**?**]. This resolution was modeled by adding these uncertainties to the pre-reconstructed MC samples. According to the analysis note, the fitting function used was defined as a function of invariant mass

$$\sigma_{\mu^+\mu^-}(m) = A + B \cdot m + C \cdot m^2 + D \cdot m^3 + E \cdot m^4 \tag{5}$$

Parameters for the BB category were used for a smeared resolution that takes into account the data resolution[**?**]. This uncertainty was applied to the generated events with the outcomes shown on figures 10 and 11.

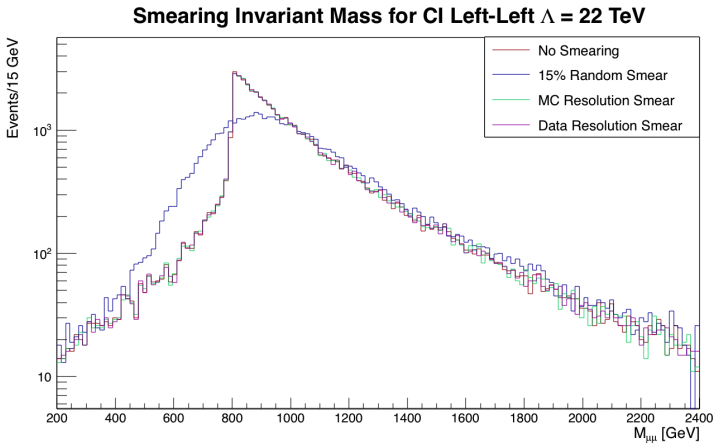

Figure 10: 10% gaussian momentum smearing with a log y scale

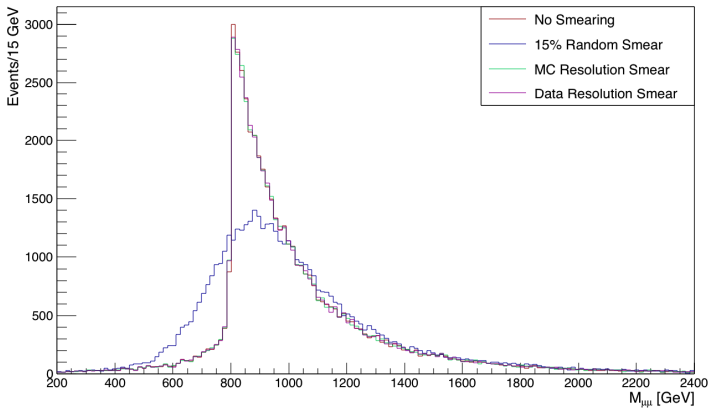

Figure 11: 10% gaussian momentum smearing with a linear y scale

These histograms are then entered into our Higgs Combine Tool as a way to validate it in preparation for data analysis, and to asses the effects of momentum resolution.

An additional study performed was an acceptance times migration simulation study. This was obtained by the simple expression for each bin.

$$A \times M = \frac{N_{\text{reco}}}{N_{\text{gen}}} \tag{6}$$

This produced the following plot for the $\Lambda = 22$ TeV sample.

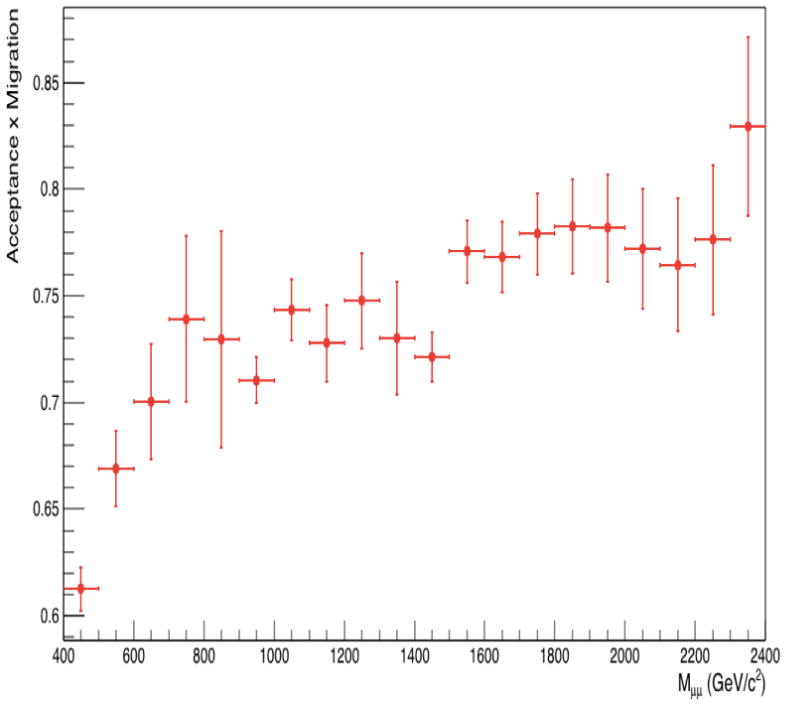

Figure 12: Acceptance x Migration as a function of dilepton invariant mass.

The larger acceptance at higher dilepton mass is sensible. The systematic uncertainties means that events at lower mass bins will migrate to higher mass bins

V. Conclusions

To assist in the search for contact interactions and quark/lepton compositness at CMS, Monte Carlo simulation studies were made. The analysis tools developed will become useful as the workflow is consolidated to just a few steps. As demonstrated, the Monte Carlo simulations are behaving as expected. Expected limits can be calculated once all the samples are fully generated and ntupled. Additionally, a large base of scripts and code was developed to assist in automating the workflow for future students and contributors to the analysis.

VI. Acknowledgements

Thanks to my supervisors Pushpa Bhat and Leonard Spiegel for accepting me as their intern and providing me with challenging, yet meaningful work. I also want to thank Peter Dong, Shawn Zaleski, and Sudeshna Banerjee for guiding me and for their patience in answering my many, many questions. Lastly, thank you to Fermilab and the SIST program: Judy Nunez, Sandra Charles, and Elliot McCrory for giving me the opportunity of a lifetime!

References

[1] E. EICHTEN, I. HINCHLIFFE, K. LANE, C. QUIGG, 1984. Supercollider Physics: VIII. Composite Quarks and Leptons. Batavia, U.S.A.: FNAL

[2] Kottachchi, Chamath. "SEARCH FOR CONTACT INTERACTIONS USING THE DIMUON MASS SPECTRUM IN P-P COLLISIONS AT ? S = 8 TeV AT CMS." Diss. Wayne State U, 2014.

[3] Nagashima, Yorikiyo. 2013. Elementary Particle Physics Vol.2: Foundations of the Standard Model. Weinheim, Germany: WILEY-VCH.

[4] GEANT4 Collaboration, GEANT4-a simulation toolkit, Nucl. Instrum. Meth. A 506 (2003) 250

[5] The CMS Collaboration. 2013. Search for contact interactions in $\mu^+\mu^-$ events in pp collisions at $\sqrt{s} = 7$ TeV. Eur. Phys (2013)

[6] The CMS Collaboration. 2016. Search for High-Mass Resonances Decaying to Muon Pairs in pp Collisions at $\sqrt{s}$ = 13 TeV with the full 2016 data set of 37 fb and combination with 2015 result. Eur. Phys. (2016)

VII. Appendix

I. Collins-Soper frame angle derivation

First some useful Quantities We first define some angle Θ which we will call pseudorapidity (left). Generic lorentz boost (right)

$$\begin{aligned}\Theta &= \frac{1}{2}\ln\frac{E+p_zc}{E-p_zc}\\ &= \frac{1}{2}\ln\frac{\gamma mc^2+\gamma mv_zc}{\gamma mc^2-\gamma mv_zc}\\ &= \frac{1}{2}\ln\frac{1+\beta}{1-\beta}\\ 1+\beta &= e^{2\Theta}(1-\beta)\\ e^{2\Theta}-1 &= \beta(e^{2\Theta}+1)\\ \frac{e^{2\Theta}-1}{e^{2\Theta}+1} &= \tanh(\Theta)=\beta\end{aligned}$$

$$\Lambda^{\alpha}_{\beta}p_{\mu} = p'_{\mu}$$

$$\begin{pmatrix}\gamma & -\beta\gamma & 0 & 0\\ -\beta\gamma & \gamma & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}E\\ p_xc\\ p_yc\\ p_zc\end{pmatrix} = \begin{pmatrix}\gamma E-\beta\gamma p_xc\\ \gamma p_xc-\beta\gamma E\\ p_yc\\ p_z\end{pmatrix}$$

We can then express the energy and momentum of the boosted frame as...

$$E' = \gamma E - \beta\gamma p_xc$$

$$p'_xc = \gamma p_xc - \beta\gamma E$$

Lastly, the derivation requires boosting from one rest frame to another. We can assume then that the momentum in one frame will be 0. We can generalize this to transverse momentum. The generic lorentz boosts then become, first we suppose

$$p_Tc = 0 \text{ then...}$$

$E = \frac{-p'_Tc}{\beta\gamma}$ Because it is in the rest frame, E is also the invariant mass, and

$$\beta\gamma = \frac{|p'_Tc|}{E} = \frac{|Q_T|}{Q}\text{. and also}$$

$$\gamma = \frac{\sqrt{Q^2+Q_T^2}}{Q}$$

For the CS angle, we will start from the center of mass frame and move to the lab frame. We define the variables in this frame as

$$k = \frac{Q}{2} \qquad k'_1 = \begin{pmatrix}k\\ k\sin(\theta)\\ 0\\ k\cos(\theta)\end{pmatrix} \qquad k'_2 = \begin{pmatrix}k\\ -k\sin(\theta)\\ 0\\ -k\cos(\theta)\end{pmatrix}$$

Then boost in the transverse plane

$$k^*_1 = \begin{pmatrix}\gamma k+\gamma\beta k\sin(\theta)\\ \gamma\beta k+\gamma k\sin(\theta)\\ 0\\ k\cos(\theta)\end{pmatrix} \qquad k^*_2 = \begin{pmatrix}\gamma k-\gamma\beta k\sin(\theta)\\ \gamma\beta k-\gamma k\sin(\theta)\\ 0\\ k\cos(\theta)\end{pmatrix}$$

Then define the quantities

$$k^{\pm}_{1,2} \equiv \frac{E_{1,2} \pm k_{(1,2)z}}{\sqrt{2}}$$

We want to solve for the quantity $(k_1^+ k_2^- - k_1^- k_2^+)$. The expanded defined terms are then.

$$\begin{aligned}
k_1^{*+} &= \frac{k}{\sqrt{2}}\big(\gamma(1+\beta\sin(\theta)) + \cos(\theta)\big) \\
k_1^{*-} &= \frac{k}{\sqrt{2}}\big(\gamma(1+\beta\sin(\theta)) - \cos(\theta)\big) \\
k_2^{*+} &= \frac{k}{\sqrt{2}}\big(\gamma(1-\beta\sin(\theta)) - \cos(\theta)\big) \\
k_2^{*-} &= \frac{k}{\sqrt{2}}\big(\gamma(1-\beta\sin(\theta)) + \cos(\theta)\big)
\end{aligned}$$

Plug this into the expression. S and C are used here instead of cosine and sine.

$$\begin{aligned}
(k_1^+ k_2^- - k_1^- k_2^+) =& \frac{k^2}{2}\big[(\gamma(1+\beta S) + C)(\gamma(1-\beta S) + C) \\
& - (\gamma(1+\beta S) - C)(\gamma(1-\beta S) - C)\big] \\
=& \frac{k^2}{2}\big[\gamma^2(1-\beta^2 S^2) + \gamma C(1+\beta S) + \gamma C(1-\beta S) + C^2 \\
& - (\gamma^2(1-\beta^2 S^2) - \gamma C(1+\beta S) - \gamma C(1-\beta S) + C^2)\big] \\
=& \frac{k^2}{2}(4k\gamma C) = 2k^2\gamma\cos(\theta) \\
=& \frac{2\sqrt{Q^2+Q_T^2}}{Q}\Big(\frac{Q}{2}\Big)^2\cos(\theta) \\
\therefore \cos(\theta_{CS}) =& 2\frac{(k_1^{*+}k_2^{*-} - k_1^{*-}k_2^{*+})}{Q\sqrt{Q^2+Q_T^2}}
\end{aligned}$$

We then do a boost in the z direction. Luckily, this quantity is invariant under z boosts. Let's prove that this is the case. This time we will use the pseudorapidity instead of lorentz factors

$$\begin{aligned}
E^* =& E\cosh(\Theta) - k_z\sinh(\Theta) \\
k_z^* =& -E\sinh(\Theta) + k_z\cosh(\Theta)
\end{aligned}$$

$$\begin{aligned}
k_1^+ =& \frac{1}{\sqrt{2}}(E_1 + k_{z1})(\cosh(\Theta) - \sinh(\Theta)) \\
k_1^- =& \frac{1}{\sqrt{2}}(E_1 - k_{z1})(\cosh(\Theta) + \sinh(\Theta)) \\
k_2^+ =& \frac{1}{\sqrt{2}}(E_2 + k_{z2})(\cosh(\Theta) - \sinh(\Theta)) \\
k_2^- =& \frac{1}{\sqrt{2}}(E_2 - k_{z2})(\cosh(\Theta) + \sinh(\Theta))
\end{aligned}$$

Then some algebra.

$$k_1^+ k_2^- = \frac{1}{2}(E_1 + k_{z1})^*(E_2 - k_{z2})^*(\cosh^2(\Theta) - \sinh^2(\Theta)) = (k_1^+ k_2^-)^*$$
$$k_1^- k_2^+ = \frac{1}{2}(E_1 - k_{z1})^*(E_2 + k_{z2})^*(\cosh^2(\Theta) - \sinh^2(\Theta)) = (k_1^- k_2^+)^*$$

Therefore it is invariant! and

$$(k_1^+ k_2^- - k_1^- k_2^+) = (k_1^+ k_2^- - k_1^- k_2^+)^*$$

$$\therefore \cos(\theta_{CS}) = 2\frac{(k_1^+ k_2^- - k_1^- k_2^+)}{Q\sqrt{Q^2 + Q_T^2}}$$

QED