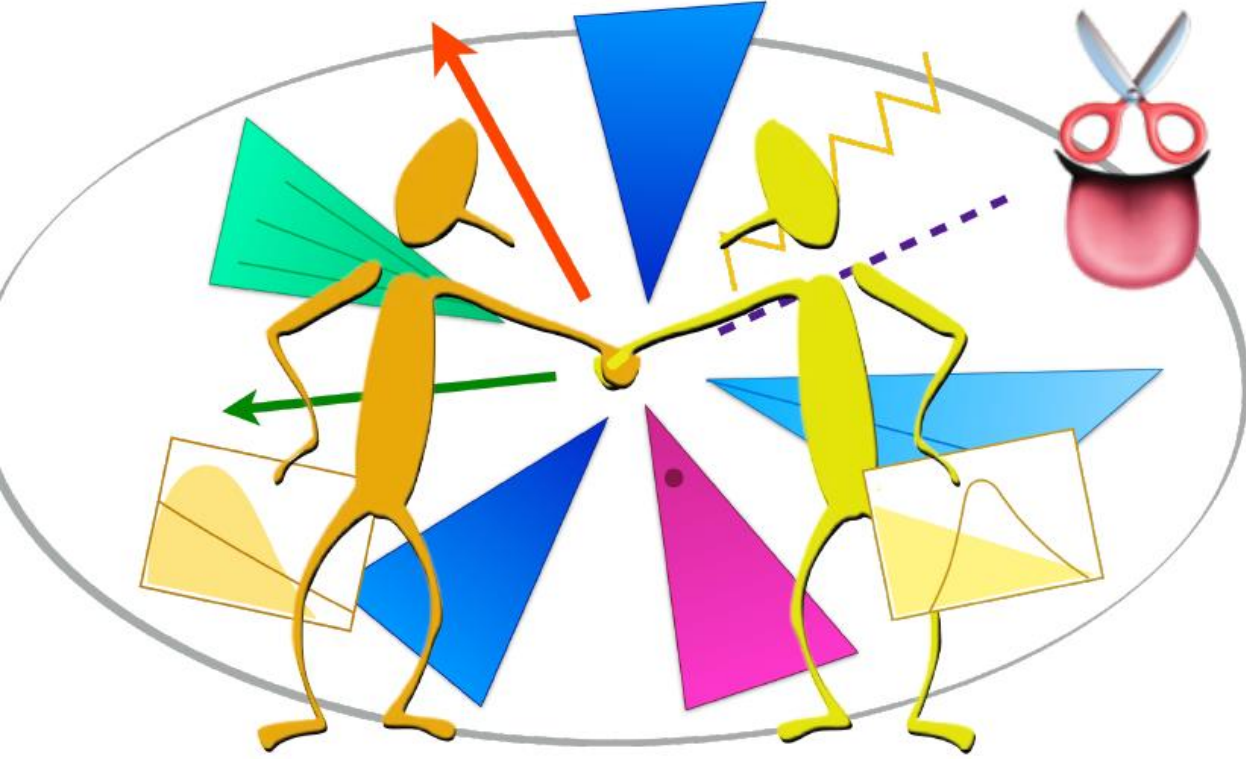


ADL/CutLang ile Fizik Odaklı YEF Analizi



Aytül Adıgüzel (İstanbul Üniversitesi)
Sina Aktaş (Boğaziçi Üniversitesi)
Burak Dağlı (TOBB ETÜ)
Berare Göktürk (Boğaziçi Üniversitesi)
Ümit Kaya (İstinye Üniversitesi)
Erkcan Özcan (Boğaziçi Üniversitesi)
Ahmetcan Sansar (İstanbul Üniversitesi)
Sezen Sekmen (Kyungpook Nat. U.)
Saleh Sultansoy (TOBB ETÜ)
Burak Şen (ODTÜ)
Gökhan Ünel (UC Irvine)

cern.ch/adl

YEFİST 2022
25 Eylül 2022



Analiz Betimleme Dili (ADL) & CutLang

Mevcut analiz çerçevelerinde, veri analizleri C++ ve Python gibi genel amaçlı dillere dayanan analiz yazılım çerçeveleri kullanarak yapılırlar:

- Bu çerçevelerde **fizik içeriği ve teknik işlemciler birbirine geçmişlerdir** ve birlikte yürütülürler.
- Fizik içeriğinin bu dağınık ve standarttan uzak ifade tarzı analizleri geliştirmeyi, anlamayı, içeriği iletmeyi ve yeni modeller ile yorumlamayı zorlaştırır.
- Farklı çerçevelerde çalışmak ve öğrenmek zaman alır.

Önerdiğimiz alternatif sistem:

- **Fizik algoritmasını tamamen teknik detaylardan ayırır.**
- Veriyle daha doğrudan bir etkileşime izin verir.

Analiz Betimleme Dili & CutLang geleneksel olmayan bir analiz sistemi sağlar:

- YEF analizinin fizik içeriğini tanımlayan **alana özgü, bildirimsel** bir dil.
- Analiz ile ilgili işler için dili çalıştırılabilir hale getiren araçlar.



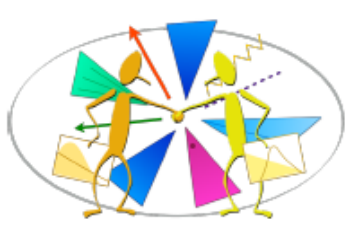
BHÇ Fiziği için Analiz Betimleme Dili

Analiz Betimleme Dili (ADL) bir YEF analizinin fizik içeriğini standart ve açık bir şekilde ifade eden alana özgü, bildirimsel bir dildir (DSL).

- **Harici DSL:** Analize özgü kavramları ifade etmek için özel olarak tasarlanmış sözdizimi. Parçacık fizikçilerinin kavramsal akıl yürütmesini yansıtır. Programlamaya değil fiziğe odaklanır.
- **Bildirimsel:** Yapılması gerekeni açıklar, ancak nasıl yapılacağını tanımlamaz.
- **Kolay okunabilir:** Açık, kendi kendini tanımlayan sözdizimi kuralları.
- **Herkes için tasarlandı:** deneyciler, fenomenoloji yapanlar, öğrenciler, ilgilenen diğer herkese açık...

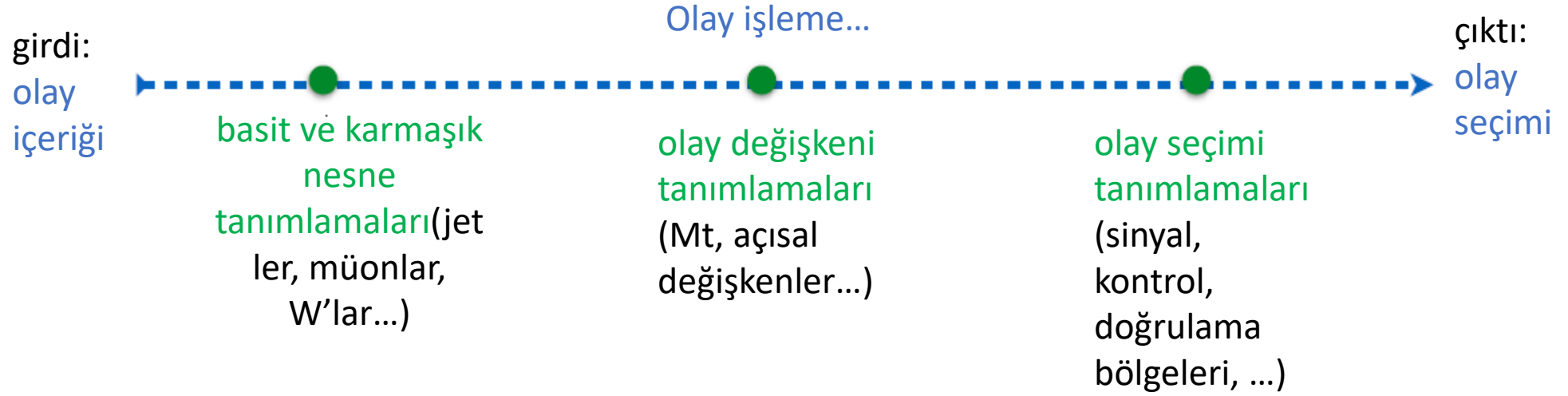
ADL çerçevelerden bağımsızdır --> ADL'yi tanıyan herhangi bir çerçevede analiz koşturulabilir.

- **Çok amaçlı kullanım:** Belirli bir amaç için en uygun GPL / çerçeve otomatik olarak çevrilebilir veya dahil edilebilir. Örneğin deney, analiz, (yeniden)yorumlama, analiz sorguları...
- Gruplar arası **kolay iletişim:** Deneysel, fenomenoloji, hakemler, öğrenciler.
- Analizin fizik içeriğinin **korunumunu sağlar.**



ADL Kapsamı

- Olay işleme: Öncelikli odak noktamız.



- Analiz sonuçları, örneğin sayımlar ve belirsizlikler: Mevcut.
- Histogramlama: Mevcut.
- Sistemik belirsizlikler: Devam ediyor.
- Seçilen olaylarla işlemler, örneğin aralan tahmini, ölçek faktörü türetme: Henüz kapsam dahilinde değil.



ADL Yapısı

ADL şunlardan oluşur:

- Açık sözdizimi kurallarına sahip okunması kolay bir DSL kullanarak analiz algoritmasını açıklayan bir **düz metin ADL dosyası**.
- ADL sözdizimi ile ifade edilmesi basit olmayan değişkenleri (ör. MT2, ML algoritmaları) içine alan **bağımsız işlevlerden oluşan bir kütüphane**. Dahili veya harici (kullanıcı) fonksiyonları.

- Bir **ADL dosyası** nesne, değişken ve olay seçimi tanımlarını ayıran bloklardan oluşur. Bloklar **anahtar kelime-talimat** yapısına sahiptir.
 - **anahtar kelimeler**, analiz kavramlarını ve işlemlerini belirtir.

```
blocktype blockname  
keyword1 instruction1  
keyword1 instruction2  
keyword3 instruction3 # comment
```

- **Sözdizimi**, matematiksel ve mantıksal işlemleri, karşılaştırma ve optimizasyon operatörlerini, indirgeyicileri, 4 vektör cebirini ve YEF'e özgü fonksiyonları($d\phi$, dR , ...) içerir.



ADL Sözdizimi: ana bloklar, anahtar kelimeler, işlemciler

Blokun amacı	Anahtar kelime
Nesne tanımlama blokları	object
Olay seçme blokları	region
Analiz veya ADL bilgisi	info
Tablo	table

Anahtar kelime amacı	Anahtar kelime
değişkenleri ve sabitleri tanımlama	define
nesne veya olay seçimi	select
nesne veya olay reddetme	reject
ana nesneyi tanımlama	take
ağırlıkları uygulama	weight
değişkenleri yazdırma	print
nesneleri sıralama	sort
histogramları tanımlama	histo
değişkenleri kaydetme	save

İşlem	İşlemci
Karşılaştırma işlemcileri	> < => =< == != [] (dahil) [] (hariç)
Matematiksel işlemciler	+ - * / ^
Mantıksal işlemciler	and or not
Koşul işlemcisi	koşul ? doğru : yanlış
Optimizasyon işlemcileri	~= (en yakın) ~! (en uzak)
Lorentz vektörleri toplama	LV1 + LV2 LV1 LV2

ADL sözdizimi kuralları ve kullanım örnekleri:

<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ADL>



ADL Sözdizimi: fonksiyonlar

Standart/dahili fonksiyonlar: Yeterince genel matematik ve YEF işlemleri dilin ve onu yorumlayan herhangi bir aracın parçası olabilir.

- Matematik fonksiyonları: `abs()`, `sqrt()`, `sin()`, `cos()`, `tan()`, `log()`, ...
- Collection reducers: `size()`, `sum()`, `min()`, `max()`, `any()`, `all()`, ...
- YEF spesifik fonksiyonlar: `dR()`, `dphi()`, `deta()`, `m()`, ...
- Nesne ve koleksiyon işleme: `union()`, `comb()`, ...

Harici/kullanıcı fonksiyonları: Mevcut işlemcileri kullanarak ifade edilemeyen değişkenler ve standart fonksiyonlar, bir ADL dosyasından çağrılabilen bağımsız fonksiyonlarla ifade edilebilir ve bir veri tabanı aracılığıyla derleyiciler tarafından erişilebilir.

- Basit olmayan algoritmalara sahip değişkenler: MT2, aplanarity, razor değişkenleri, ...
- Analitik olmayan değişkenler: nesne ve tetikleme verimlilikleri, değişken/ML ile hesaplanan değişkenler ve verimlilikler, ...

CutLang yorumlayıcısı ve çerçeve



CutLang çalışma zamanı yorumlayıcısı:

- Analiz içeriği derlenmez. ADL dosyasını olaylar üzerinden doğrudan çalıştırır.
- CutLang'in kendisi C++ ile yazılmıştır, herhangi bir modern Unix ortamında çalışır.
- Lorentz Vektör işlemleri ve histogramları için ROOT sınıflarını temel alır.
- ADL sözdizimi işleme: Lex & Yacc gereçleri ile.

CutLang çerçevesi: yorumlayıcı + araçlar
Olay ve harici fonksiyon girdileri:

- Olay girdileri: ROOT dosyaları aracılığıyla.
 - Birden çok girdi tipi: Delphes, CMS, NanoAOD, ATLAS/CMS Open Data, LVL0, FCC. Daha fazlası kolaylıkla eklenebilir.
- Tüm olay tipleri çalışma zamanı yorumlayıcısı tarafından tanınacak önceden tanımlanmış parçacık nesne tiplerine dönüştürülür --> aynı ADL dosyası farklı girdi tipleri ile çalıştırılabilir.

CutLang Github: <https://github.com/unelg/CutLang>

CutLang Yayınları: [arXiv:1801.05727](https://arxiv.org/abs/1801.05727), [arXiv:1909.10621](https://arxiv.org/abs/1909.10621)
[arXiv:2101.09031](https://arxiv.org/abs/2101.09031)

CutLang yorumlayıcısı ve çerçeve



CutLang çerçevesi: yorumlayıcı + araçlar

Çalıştırma çıktısı:

- **Sonuç çıktıları:** ROOT dosyaları aracılığıyla.
 - Her seçim bölgesi için cutflow, binler ve kullanıcı tarafından tanımlanmış histogramların bulunduğu bir TDirectory.
- **Analizin herhangi bir aşamasında olay içeriğinin kaydedilmesi** (ör. ML için)
 - Seçilen nicelikler olay seçimlerinden sonra CSV dosyaları halinde kaydedilebilir.

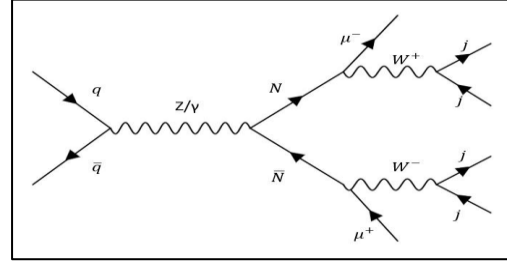
Yardımcı araçlar:

- Verisetleri, tesir kesitler, örneğe özel ağırlıklar için **metin tabanlı konfigürasyon dosyaları**.
- **Çıktı dosyalarını birleştirmek ve grafik çizimi için betikler de mevcut.**

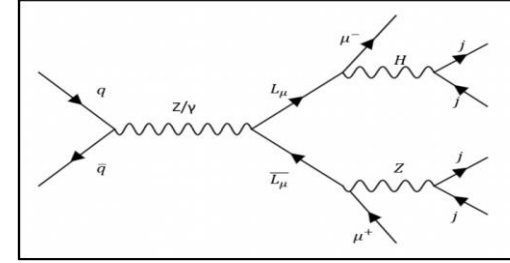
Bunların hepsi CutLang çerçevesine özgüdür ve ADL'nin kapsamı dışındadır.
Devam ediyor.

ADL/CutLang ile Vektör Benzeri Ağır Lepton Analizi

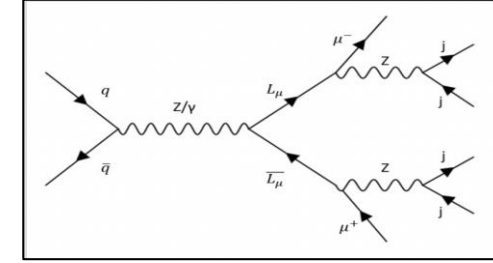
Büyük Hadron Çarpıştırıcısı tarafından 13 TeV kütle merkezi enerjisinde 139 fb⁻¹ toplam ışınlılığa karşılık gelen, ATLAS algıcının 2015-2018 yılları arasında kaydettiği veride ADL/CutLang kullanarak E6 Büyük Birleşme Kuramı'nın öngördüğü vektör benzeri ağır leptonları (VLL) araştırılmaktadır. Bu çalışmada ağır leptonun çift üretimine ait son durumda zıt yüklü 2 lepton içeren bozunum kanalları sunulmuştur.



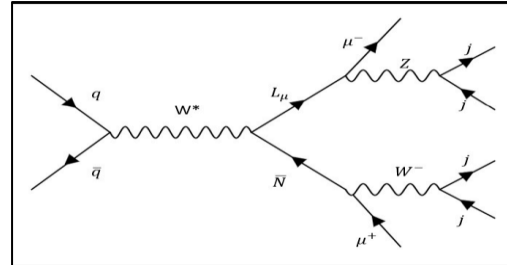
$pp \rightarrow N\bar{N} \rightarrow W^+\mu^-W^-\mu^+ \rightarrow 2\mu 4j$



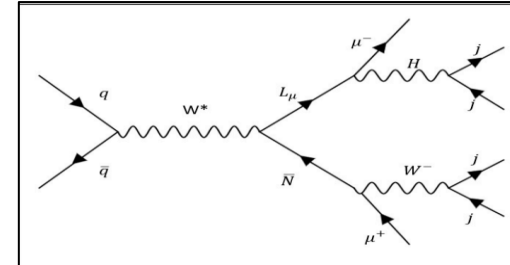
$pp \rightarrow L\bar{L} \rightarrow H\mu^-Z\mu^+ \rightarrow 2\mu 4j$



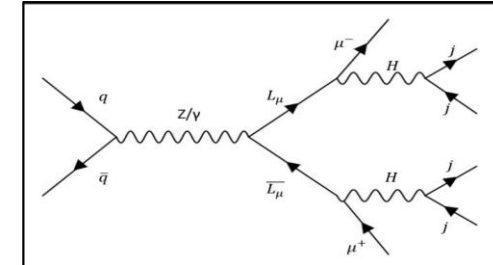
$pp \rightarrow L\bar{L} \rightarrow Z\mu^-Z\mu^+ \rightarrow 2\mu 4j$



$pp \rightarrow L\bar{N} \rightarrow Z\mu^-W^-\mu^+ \rightarrow 2\mu 4j$



$pp \rightarrow L\bar{N} \rightarrow H\mu^-W^-\mu^+ \rightarrow 2\mu 4j$



$pp \rightarrow L\bar{L} \rightarrow H\mu^-H\mu^+ \rightarrow 2\mu 4j$

- Son durumda zıt yüklü 2 lepton içeren sinyal ve doğrulama bölgelerine ait seçimler:

	OS leptons	Pt(leps) > 20 GeV	Tight ID and FC Looselsolation for electrons	Medium ID and FC Looselsolation for muons	Tau veto	Z veto	1Z	nJets >= 2	Mll > 15 GeV
SR	✓	✓	✓	✓	✓	✓	✗	✓	✓
Zjets VR	✓	✓	✓	✓	✓	✗	✓	✓	✓
ttbar VR	✓	✓	✓	✓	✓	✗	✗	✓	✓

2LOS Sinyal Bölgesi için ADL Yapısı

Nesne tanımlamaları

```
#OBJECT SELECTIONS
object goodMuo : MUO
  select Pt(MUO) > 20
  select absEta(MUO) < 2.5
  select Medium(MUO) == 1
  select Loose(MUO) == 1

object goodEle : ELE
  select Pt(ELE) > 20
  select absEta(ELE) < 2.4
  reject absEta(ELE) [] 1.37 1.52
  select Tight(ELE) == 1
  select Loose(ELE) == 1

object goodJet : JET
  select Pt(JET) >= 15
  select absEta(JET) < 5.5

#EVENT VARIABLE DEFINITIONS
define ll01 : goodMuo[0] goodMuo[1]
define Zcutmumu : abs(m(ll01) - 91.2)
```

Olay değişkeni tanımlamaları

Olay ağırlıkları

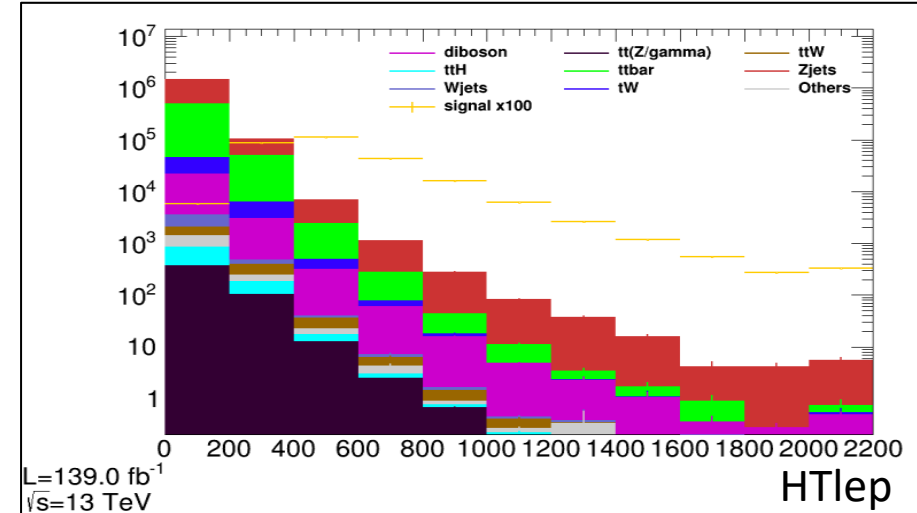
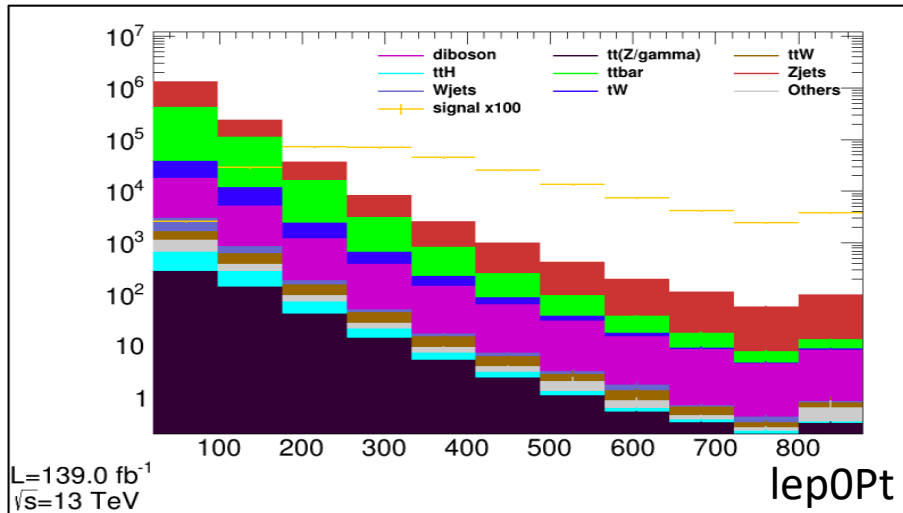
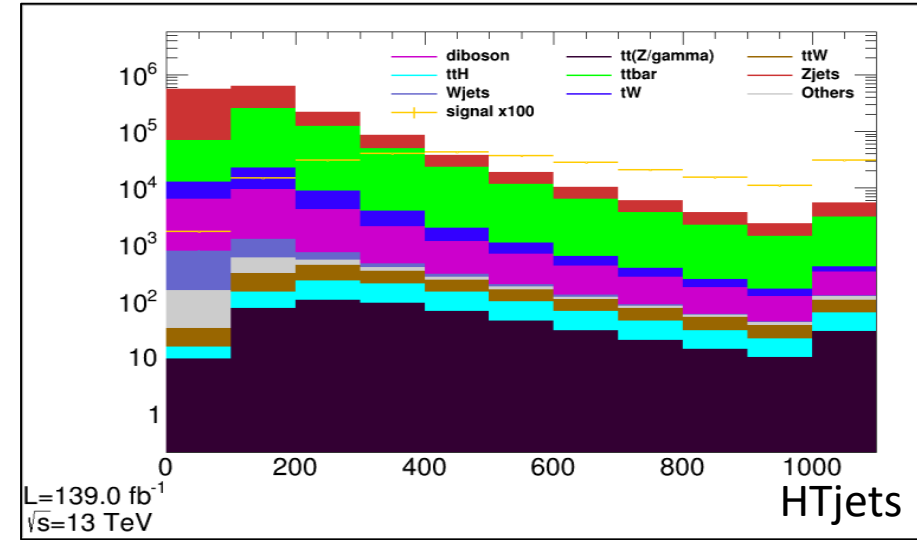
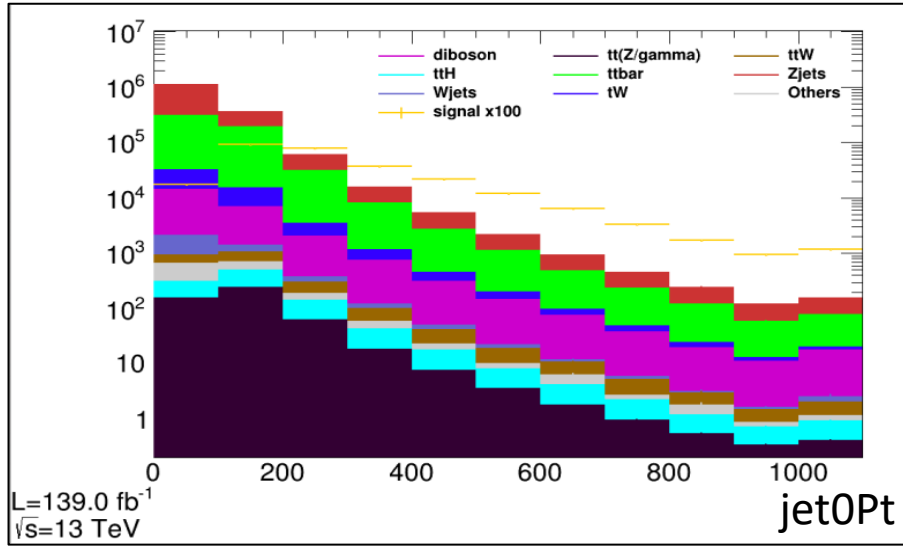
Histogram tanımlamaları

```
#EVENT SELECTION
region SR2LOSmm
select ALL
select XSLumiCorrSF
weight Wyears YearEff
weight VVweight VVNJET
weight TTweight ttZweight
weight TTweight ttbarweight
select bTagSF
select LEPSf
select Size(goodJet) >= 2
select Size(goodMuo) == 2
select Size(goodEle) == 0
select Size(TAU) == 0
select q(ll01) == 0
select Zcutmumu > 10
select m(ll01) > 15
select dileptype == 1
histo jetsize, "Njets", 10, 0, 10, Size(goodJet)
histo bjetsize, "Nbjets", 10, 0, 10, Size(BJET)
histo jet0Pt, "jet0Pt", 10, 0, 1000, Pt(goodJet[0])
histo HTjets, "HT jets", 10, 0, 1000, fHT(goodJet)
histo hMET, "MET (GeV)", 10, 0, 1000, MET
histo HTlep, "HT lep", 10, 0, 2000, fHT(goodMuo)
histo hMll01, "Mll01", 10, 0, 1000, m(ll01)
histo lep0Pt, "lep0Pt", 10, 20, 800, Pt(goodMuo[0])
```

Olay seçimleri

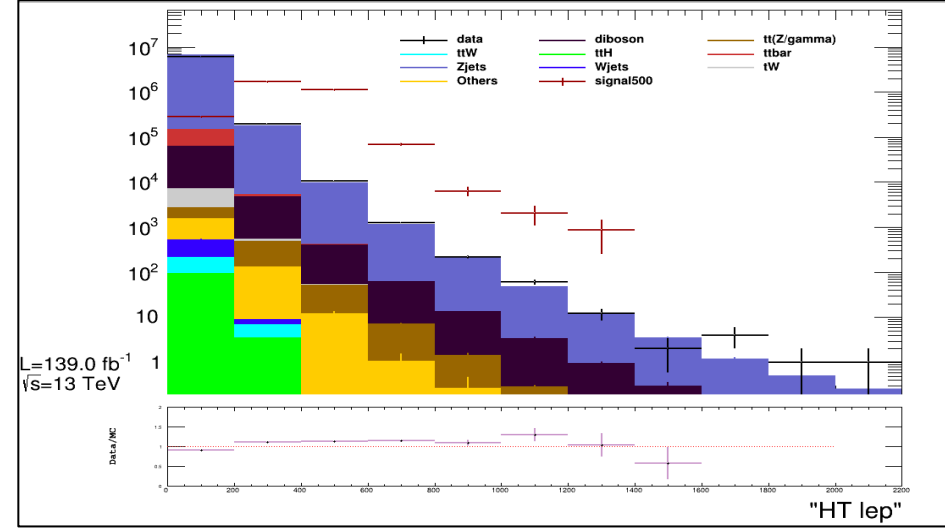
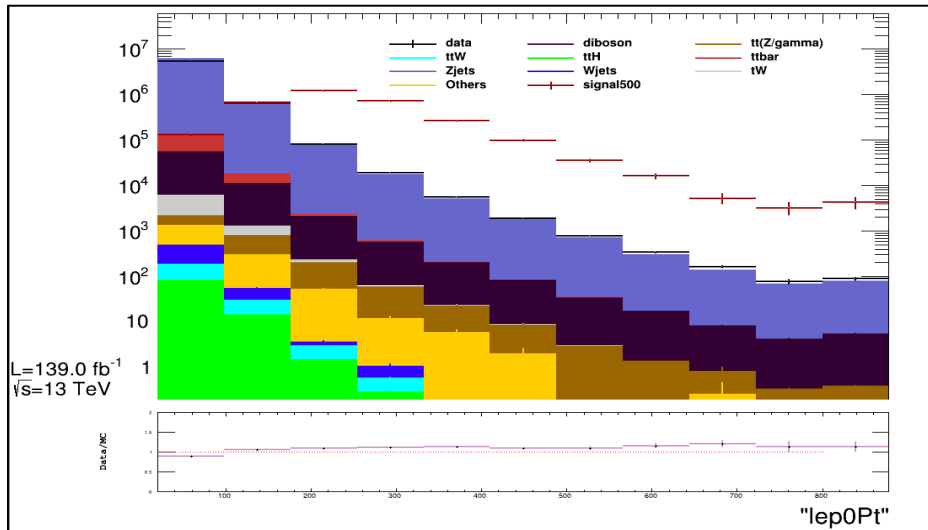
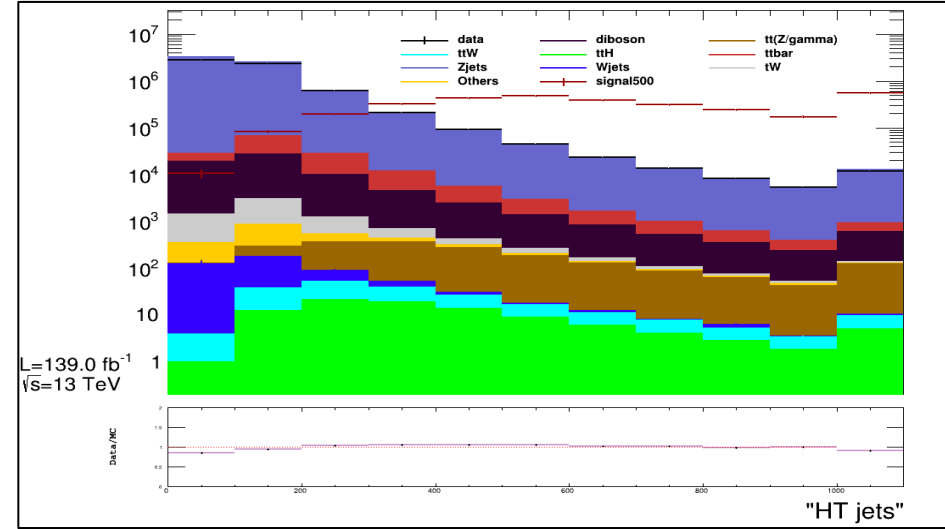
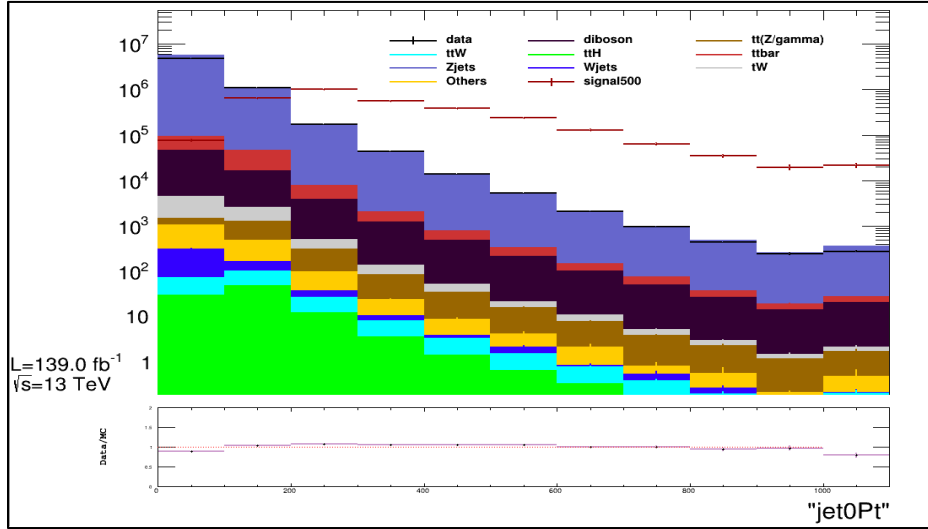
Sonuçlar

- 2LOS sinyal bölgesi için elde edilen kinematik dağılımlar.



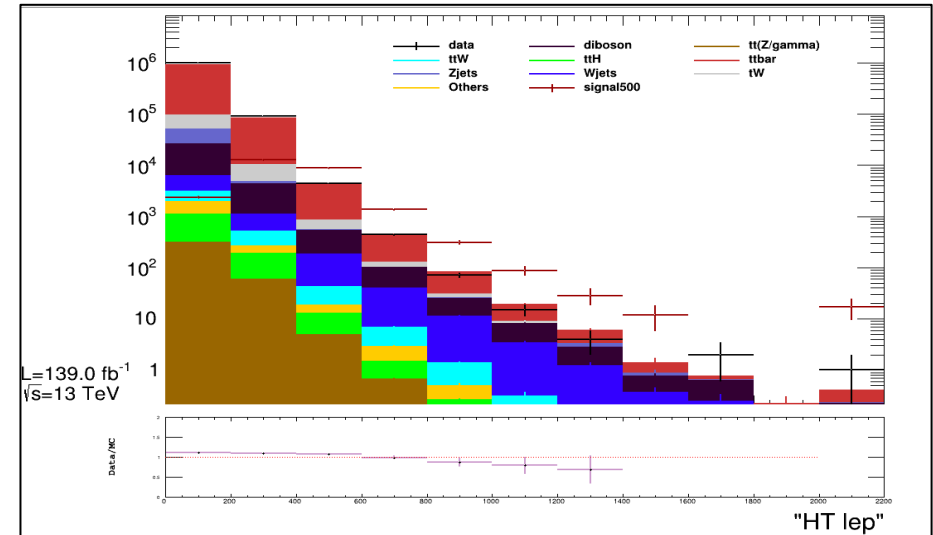
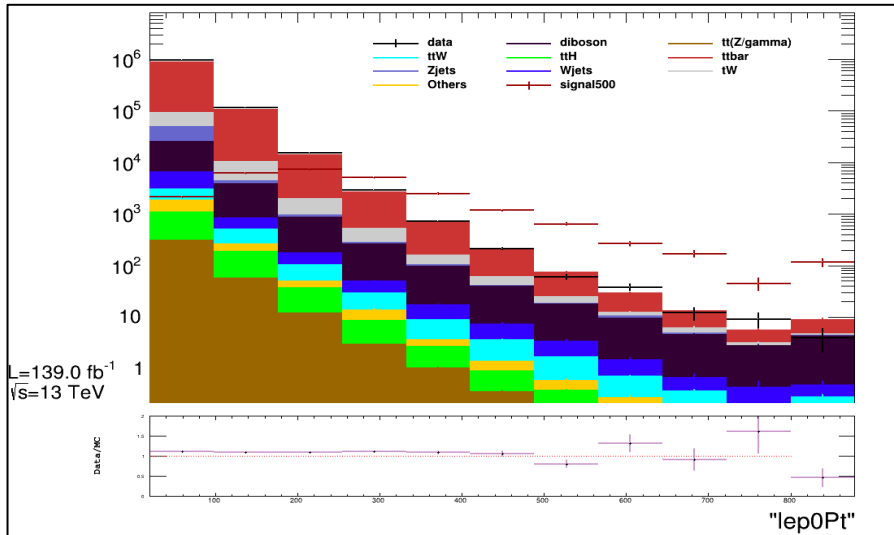
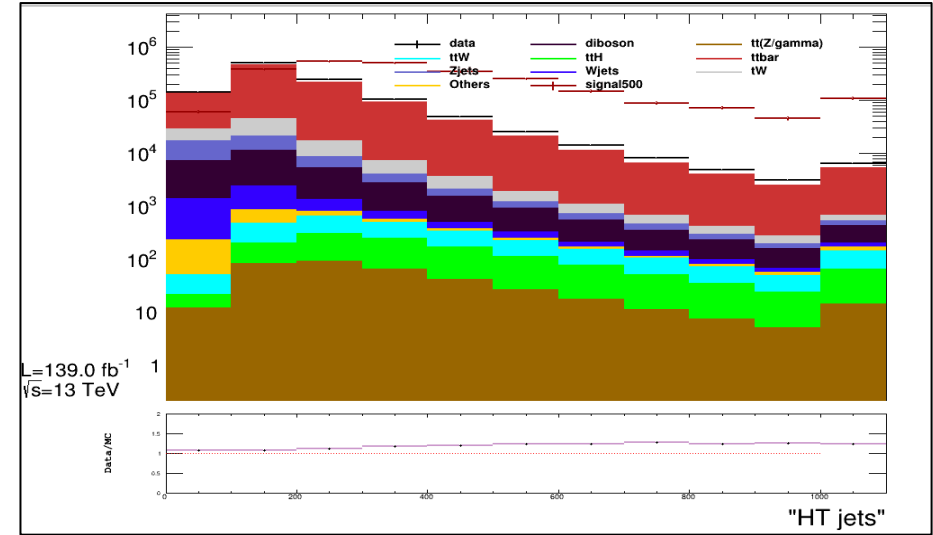
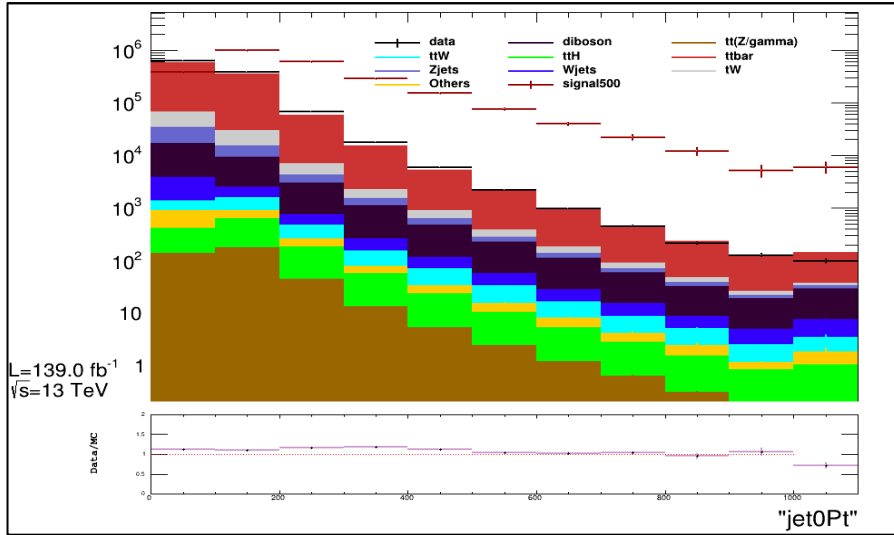
Sonuçlar

- Zjets doğrulama bölgesi için elde edilen kinematik dağılımlar:



Sonuçlar

- $t\bar{t}$ doğrulama bölgesi için elde edilen kinematik dağılımlar.



Hadronik Bozon ve Ağır Leptonların Yeniden Yapılandırılması

- Son durumda zıt yüklü 2 lepton içeren bozunum kanallarında hadronik bozonlar ve ağır leptonlar, $M_L = 500 \text{ GeV}$ kütleli sinyal Monte Carlo örneği kullanılarak yeniden yapılandırılmıştır. En iyi adayları bulmak için χ^2 algoritması kullanıldı.

$$\chi^2 = \frac{(m_{LH}^{reco} - m_{LZ}^{reco})^2}{(\sigma_{LH})^2 + (\sigma_{LZ})^2} + \frac{(m_H^{reco} - m_H^{real})^2}{(\sigma_H)^2} + \frac{(m_Z^{reco} - m_Z^{real})^2}{(\sigma_Z)^2}$$

#EVENT VARIABLE DEFINITIONS

```
define jetA : goodJet[-1]
define jetB : goodJet[-2]
define jetC : goodJet[-3]
define jetD : goodJet[-4]
define muA : goodMuo[-1]
define muB : goodMuo[-2]
```

Herhangi ikişer
jetten oluşan Higgs
ve Z bozonlarının
tanımlanması

```
define Zhad2 : jetC jetD
define Hhad1 : jetA jetB
```

Ağır leptonların
tanımlanması

```
define Lz2 : Zhad2 muB
define Lh1 : Hhad1 muA
```

```
define HZchi : (m(Hhad1) - 125)^2 / 10^2 + (m(Zhad2) - 91)^2 / 8^2
define totchiHZ : (m(Lh1) - m(Lz2))^2 / (27.4^2 + 27.5^2) + HZchi
```

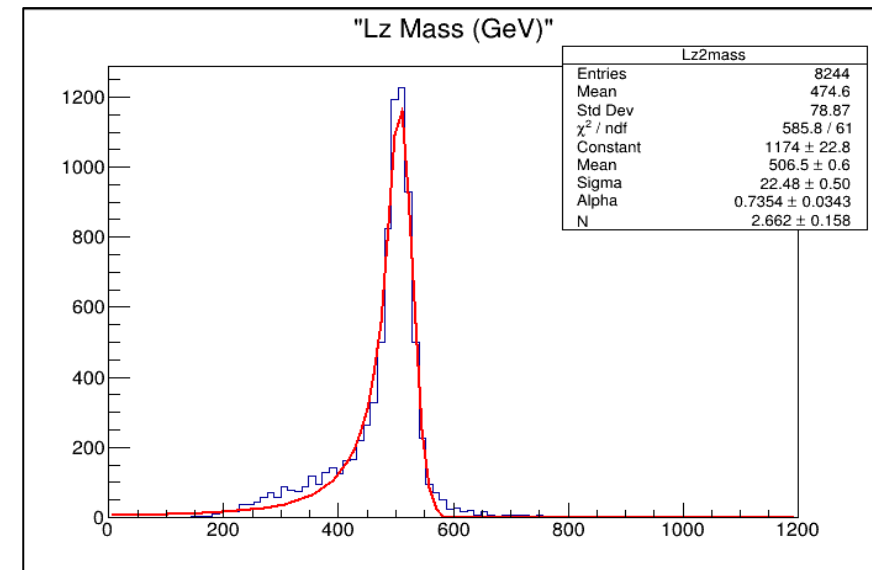
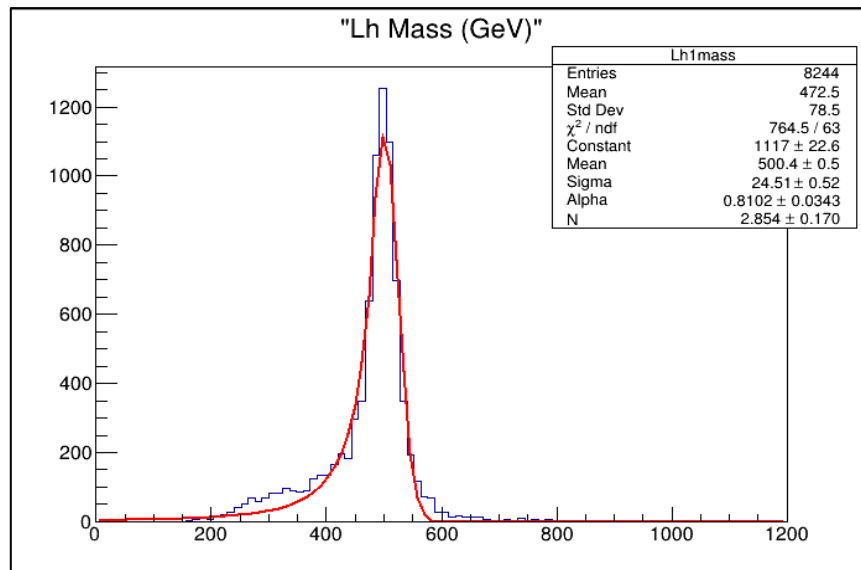
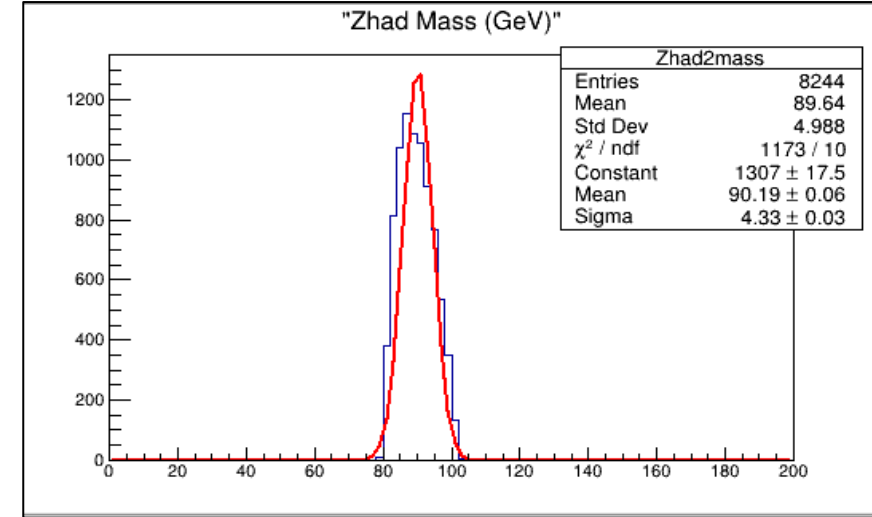
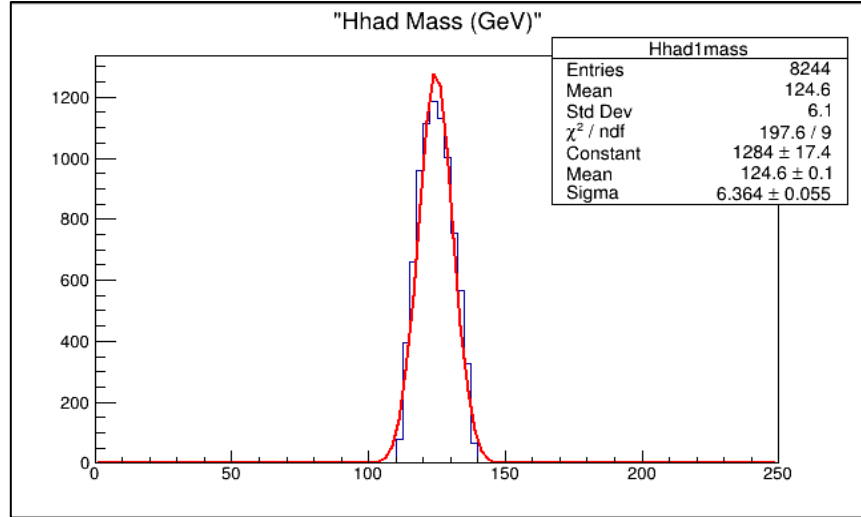
χ^2 algoritması

```
region HZ2LOS
preselection
select HZchi ~= 0
select totchiHZ ~= 0
histo Hmassvschi, "Higgs Mass vs Chi2", 200, 0, 500, 100, 0, 250, totchiHZ, m(Hhad1)
select totchiHZ < 3
histo Hhad1mass, "Hhad Mass (GeV)", 100, 0, 250, m(Hhad1)
histo Zhad2mass, "Zhad Mass (GeV)", 100, 0, 200, m(Zhad2)
histo Lh1mass, "Lh Mass (GeV)", 100, 0, 1200, m(Lh1)
histo Lz2mass, "Lz Mass (GeV)", 100, 0, 1200, m(Lz2)
```

χ^2 minimizasyonu

Hadronik Bozon ve Ağır Leptonların Yeniden Yapılandırılması

- Yeniden yapılandırılan Higgs, Z bozonlarının ve ağır leptonların kütle dağılımları ($M_L = 500 \text{ GeV}$)



Hedefler

ADL/CutLang:

- **Nesne kombinasyonları için geliştirilmiş betimleme:** Basit kombinasyonlar ve optimizasyonlar yapılabilir. (ör. en yüksek p_T 'ye sahip leptonlardan oluşan Z veya en yakın kütleyi veren herhangi iki leptondan oluşan Z, top kuark kütle optimizasyonu, $H \rightarrow ZZ, \dots$). Ancak geliştirilmiş ve temiz bir sözdizimi gerekli.
- **Sistemik belirsizlikler:** Sona yaklaşıldı.
- **Fizik veri tipleri ve araçlarıyla otomatize edilmiş arayüz:** Farklı veri tiplerini okuma ve ADL'deki temsili ile eşleştirilmesini otomatikleştirmek. Harici fonksiyonların okunmasının otomatikleştirmek.

VLL Analizi:

- **Zıt yüklü 2 lepton** son durumları için yapılan çalışmaların **3 lepton ve 4 lepton** durumları için de tekrarlanıp **data/monte carlo** oranı elde edilmesi.
- Standart Model'in öngörülerinden bir sapma görülmediği durumda **VLL kütlesi üzerine %95 güvenilirlik seviyesinde dışarlama limiti konulması.**

İlave Slaytlar

Analizde Kullanılan Deney Verisi ve Monte Carlo Örnekleri:

- **Deney verileri:** ATLAS algıcı tarafından 2015-2018 yılları arasında kaydedilen veri (139 fb^{-1})
- **Ardalanlar:** Z+jets, W+jets, Diboson, ttbar, Fourtop, Threetop, Singletop, ttZ, tZ, tW, ttW, ttH, tWH, tWZ, ttWW, tHjb, ggZZ
- **Sinyal:** $M_L = 500 \text{ GeV}$