

# Новый алгоритм реконструкции струй от быстрых $W$ - и $Z$ -бозонов

О.В.Зенин

ИФВЭ

# Мотивация

- Поиск состояний  $WW/ZZ/WZ$  с инвариантной массой  $M_{VV} = O(1 \text{ ТэВ})$ .
- $\text{Br}(W \rightarrow \text{jets}) \simeq 70\% \Rightarrow$  увеличение статистики в разы по сравнению с  $W \rightarrow \ell\nu$ .
- При  $p_T^{W,Z} \sim M_{VV} \sim 1 \text{ ТэВ}$  распад  $W/Z \rightarrow \text{jets}$  сколлимирован в конусе  $\Delta R \sim M_{W,Z}/M_{VV} \sim 0.1$  вокруг  $\vec{p}_{W,Z}$
- Реконструкция стандартными алгоритмами AntiKt0.4(0.6) неэффективна (слипание струй, выпадение из конуса части более мягкой струи от  $V \rightarrow jj$ , засорение pileup/UE).
- В ATLAS используется реконструкция алгоритмами Aachen–Cambridge, (Anti)Kt с избыточным конусом  $R = 0.6 - 1.2$  с последующей очисткой и выявлением субструктуры  $J \rightarrow jj$  (см., например, [CERN-PH-EP-2015-204](#)). Алгоритмы очистки (*pruning*, *trimming*, *grooming*) имеют произвольные параметры, оптимизируемые для максимального подавления фона от КХД-струй при фиксированной эффективности идентификации  $W/Z \rightarrow J$ .

**Задача:** построить алгоритм реконструкции струи

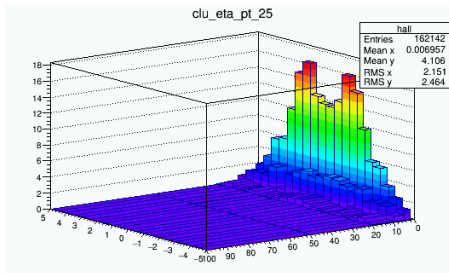
- с минимумом произвольных параметров
- без фиксированного конуса
- автоматически подавляющий pileup
- не требующий дополнительной очистки реконструированной струи
- не хуже использующихся в АТЛАС алгоритмов (см. [CERN-PH-EP-2015-204](#)) по соотношению эффективности для сигнала  $W/Z \rightarrow J$  ( $= 50\%$ ) / подавление КХД-фона  $\sim 1/50$  при  $p_T(W/Z \rightarrow J) > 200 \text{ GeV}$ ,  $\mu = 0 - 30$ .

## Алгоритм

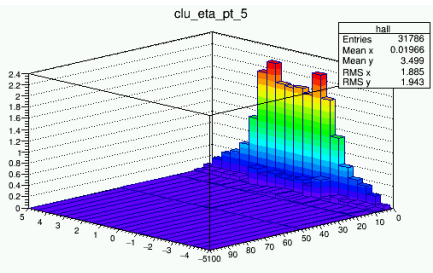
- Используем CaloCalTopCluster'ы с  $E_T^{clus} > 0.5, 1, 2 \text{ GeV}$ .  
Информация из Inner Detector не используется (кроме числа реконструированных вершин  $type=1,3$ ).
- В кластеризации участвуют 4-импульсы двух входящих партонов, оцененные по  $\Sigma \vec{p}^{clus}$  в калориметре.
- Для каждой пары объектов сравнивается вероятность прихода хотя бы одного из  $pileup/UE$  (1), и вероятность прихода одного из них из сигнала с излучением второго объекта (2). Если (1)>(2), пара игнорируется, если (2)>(1) – добавляется к списку кандидатов на объединение. Пара с максимальной вероятностью (2) объединяется, при этом для объединения вероятность прихода из сигнала наследуется от констит்யюента с максимальной вероятностью прихода из сигнала. Объекты с массой  $> Q_{max} = const \cdot M_W$  замораживаются (свободный параметр  $const = 0.1 - 1$ )
- Вероятность прихода кластера из  $pileup$  для событий с данным  $N_{type1 vtx} - 1$  вычисляется из событий MinBias/ZeroBias.

# Алгоритм (оценка pileup)

- Разбиение фазового объема по  $p_T, \eta$  произвольно (см. рис.) и предполагается изотропным по  $\phi$ . Разбиение по  $\phi$ : находим плоскость реакции,  $\phi_0 = \sum_{clus} p_T^{clus} (\phi_{clus} - (\phi_Z + \pi)) / \sum_{clus} p_T^{clus}$ , и ширину ведущей струи по  $\phi$ ,  $\Delta(\phi - \phi_0)^2 = \sum_{clus} p_T^{clus} (\phi_{clus} - (\phi_0))^2 / \sum_{clus} p_T^{clus}$ . Узкий бин по  $\phi - [\phi_0 - \Delta\phi, \phi_0 + \Delta\phi]$ , остальное делим на три бина с равной шириной.



$$N_{vtx1} = 25$$



$$N_{vtx1} = 5$$

## Алгоритм

Если в бине  $[p_T, \eta, \phi]$  ожидается  $b$  кластеров из pileup/UE, то в событии  $Z + X$  при  $n$  кластерах в этом бине, вероятность для одного кластера прийти из pileup/UE

$$w_b = \frac{b \cdot 1 + b + \dots + b^{n-1} / (n-1)!}{n \cdot 1 + b + \dots + b^n / n!}$$

а вероятность прийти из сигнала  $w_s = 1 - w_b$ .

Вероятность излучения частицы 2,  $(1+2) \rightarrow 1+2$ ,  $Q = p_1 + p_2$  в бин  $[\Delta p_{T2}, \Delta \eta_2, \Delta \phi_2]$ :

$$w_{rad} \sim \max[w_{s1}, w_{s2}] \times \frac{\alpha_S}{\pi} \cdot C \frac{Q_0}{p_0^{(2)}} \frac{p_T^{(2)} \Delta p_T^{(2)}}{Q^2} \Delta \eta^{(2)} \frac{\Delta \phi^{(2)}}{2\pi}$$

Если  $w_{rad} > 1 - w_{s1}w_{s2}$ , то добавляем пару 1+2 в список кандидатов на объединение, если  $w_{rad} < 1 - w_{s1}w_{s2}$  – игнорируем.

$\frac{\alpha_S}{\pi} \times C$  – свободный параметр, фиксируем его на  $0.1 \sim \frac{\alpha_S(M_Z)}{\pi} \cdot (C_A, C_F)$

**При реконструкции не используются условия на  $p_T$  и форму струи**

Кинематика из [CERN-PH-EP-2015-204](#):  $M_{W'} = 1000$  GeV,  
 $W' \rightarrow W(J)Z(\mu^+\mu^-)$ :

- Сигнал  $W'(1000 \text{ GeV}) \rightarrow WZ$ :  $350 < p_{truth W}^T < 500$  GeV,  
 $|y_{truth W}| < 1.2$

`mc15_13TeV.302221.MadGraphPythia8EvtGen_A14NNPDF23L0_HVT_AgV1_VcWZ_1lqq_m1000`

- Фон  $Z + jets$ :  $350 < |\Sigma \vec{p}_{q,g}^T| < 500$  GeV,  $|y_{\Sigma q,g}| < 1.2$ ,  $|y_{q,g}| < 2.5$

`mc15_13TeV.361405.Sherpa_CT10_Zmumu_Pt280_500_CVetoBVeto`

`mc15_13TeV.361406.Sherpa_CT10_Zmumu_Pt280_500_CFilterBVeto`

`mc15_13TeV.361407.Sherpa_CT10_Zmumu_Pt280_500_BFilter`

- Реконструкция:  $300 < p_{reco J}^T < 500$  GeV,  $|y_{reco J}| < 1.2$

$M_{W'} = 500 \text{ GeV}$ :

- Сигнал  $W'(500 \text{ GeV}) \rightarrow WZ$ :  $200 < p_{truth\ W}^T < 250 \text{ GeV}$ ,  
 $|y_{truth\ W}| < 1.2$

mc15\_13TeV.302216.MadGraphPythia8EvtGen\_A14NNPDF23LO\_HVT\_Agv1\_VcWZ\_1lqq\_m0500

- Фон  $Z + jets$ :  $200 < |\Sigma \vec{p}_{q,g}^T| < 250 \text{ GeV}$ ,  $|y_{\Sigma q,g}| < 1.2$ ,  $|y_{q,g}| < 2.5$

mc15\_13TeV.361402.Sherpa\_CT10\_Zmumu\_Pt140\_280\_CVetoBVeto

mc15\_13TeV.361403.Sherpa\_CT10\_Zmumu\_Pt140\_280\_CFilterBVeto

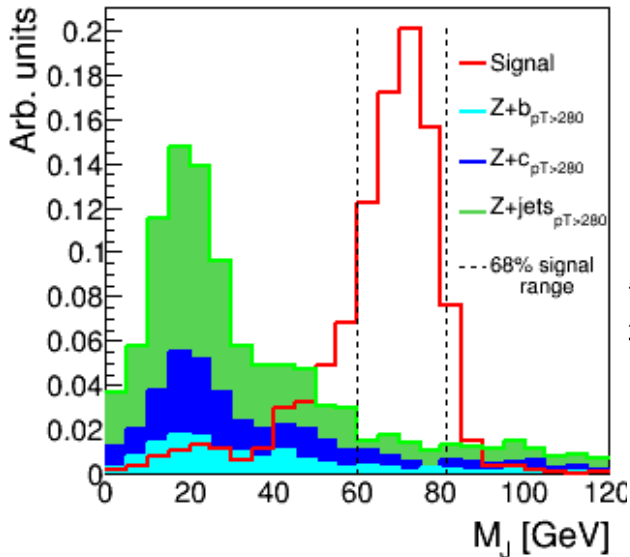
mc15\_13TeV.361404.Sherpa\_CT10\_Zmumu\_Pt140\_280\_BFilter

- Реконструкция:  $150 < p_{reco\ J}^T < 250 \text{ GeV}$ ,  $|y_{reco\ J}| < 1.2$

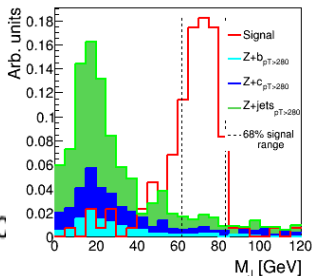


- Для сигнала подбирается минимальное по ширине окно  $M_J$ , содержащее 68% событий, отобранных по  $p_J^T$ ,  $y_J$  струи с максимальным в событии  $p_T$ .
- Для оценки эффективности подавления КХД-струй из  $Z + jets$  спектр  $p_{\Sigma q, g}^T, |y_{q, g}| < 2.5$  перевзвешивается к распределению  $p_{truth\ W}^T$  сигнала  $W' \rightarrow WZ$  (близко к процедуре [CERN-PH-EP-2015-204](#)).
- Дальнейшие обрезания по структурным переменным подбираются так, что эффективность для сигнала  $\epsilon = 50\%$ .
- Результаты без перевзвешивания фона  $Z + jets$  тоже показаны.

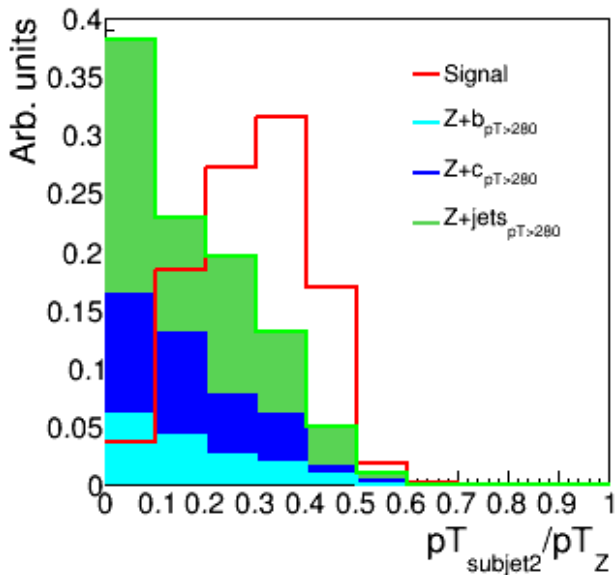
$350 < p_T^J < 500$  GeV,  $|y| < 1.2$  (cf. CERN-PH-EP-2015-204)



68% сигнала между двумя вертикальными штриховыми линиями. Слева  $\bar{\mu} = 15$ , при  $\bar{\mu} \simeq 25$  картинка практически не меняется:



$350 < p_J^T < 500$  GeV,  $|y| < 1.2$ , обрезания:  $x_2$

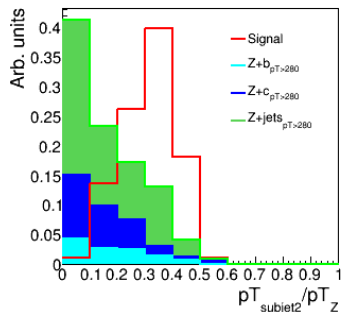


Основное подавление фона в сигнальном окне

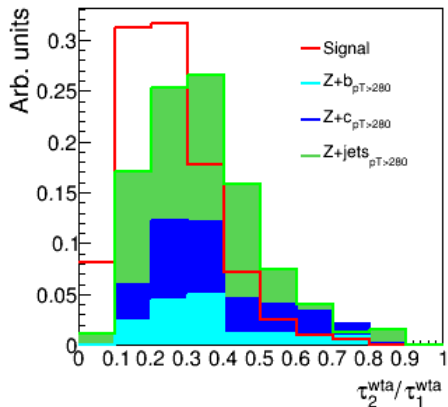
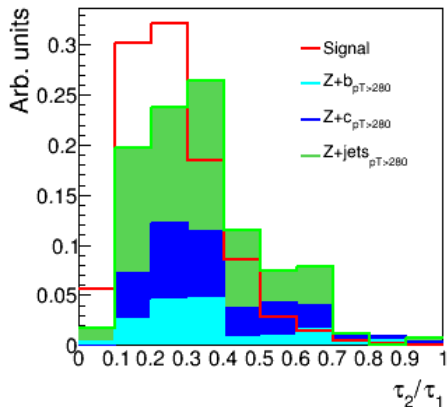
$M_J$  за счет

$$x_2 = p_T^{\text{subjet2}} / p_T^Z < 0.1$$

$$\bar{\mu} \sim 25:$$



$350 < p_T^J < 500$  GeV,  $|y| < 1.2$ , структ. переменные:  $\frac{\tau_2}{\tau_1}$



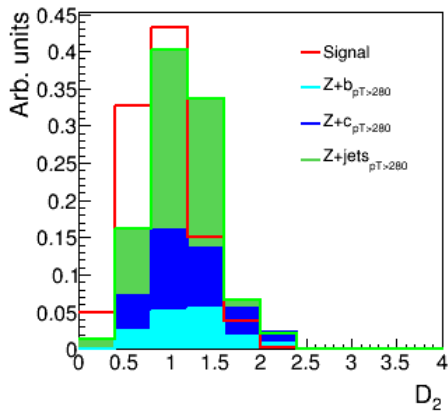
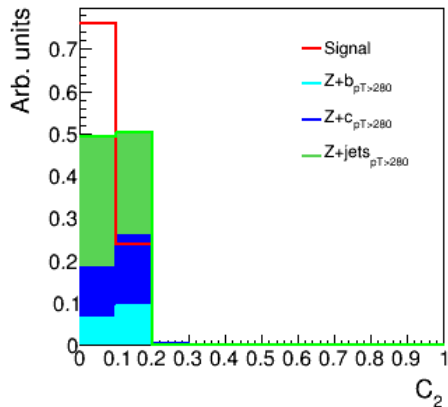
$\tau_2/\tau_1$  характеризуют наличие в массивной струе двух узких суб-струй

► см. определения  $\tau_{1,2}, C_2, D_2$

$\tau_2^{wta}/\tau_1^{wta} < 0.37$  дает  $\epsilon_{signal} = 50\%$

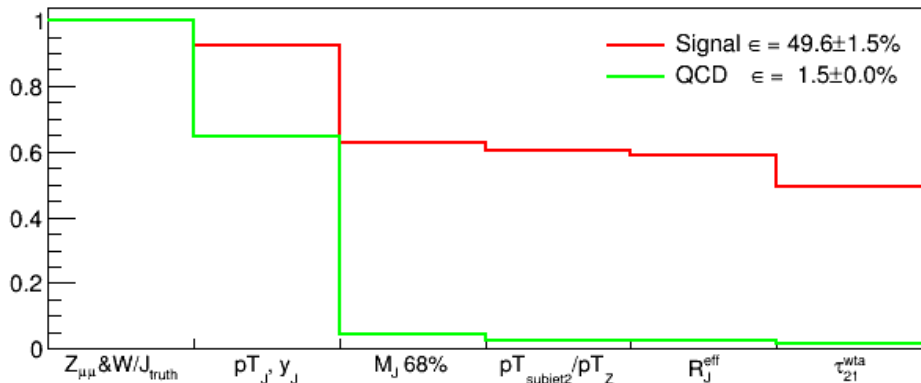
(можно обрезать и по другим переменным:  $C_2, D_2$  – см. след. слайд)

$350 < p_T^J < 500$  GeV,  $|y| < 1.2$ , структурные переменные:  
 $C_2, D_2$



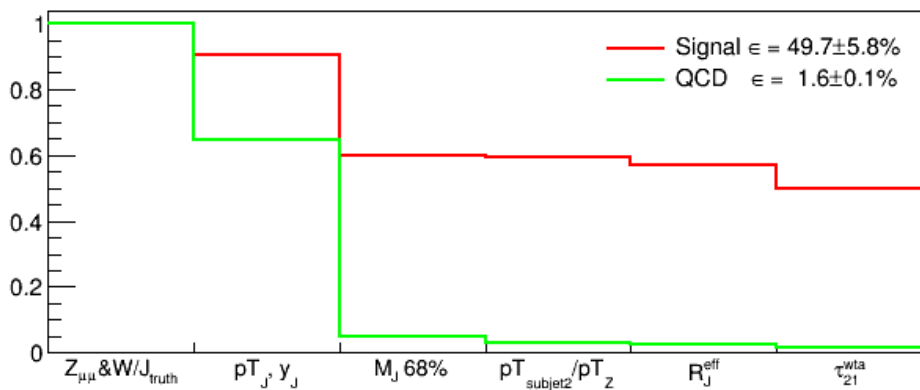
$350 < p_T^J < 500 \text{ GeV}, |y| < 1.2$

Эффект обрезаний по  $M_J, x_2, \tau_2^{wta} / \tau_1^{wta}$

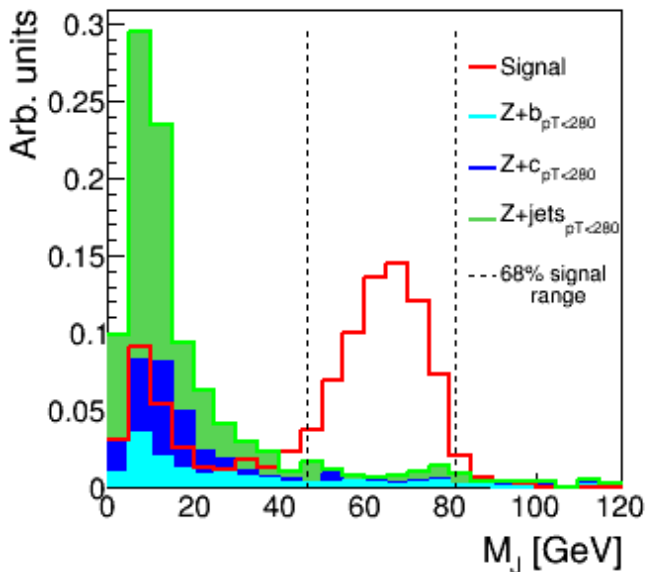


Ср. [CERN-PH-EP-2015-204](#) (Fig. 14): в той же кинематике при  $\epsilon_{\text{signal}} = 50\%$  лучшее подавление фона  $\epsilon_{\text{QCD}} \simeq 2\%$  дает алгоритм Cambridge/Aachen  $R = 1.0$ , pruned ( $R_{\text{cut}} = 0.5, Z_{\text{cut}} = 0.15$ ), обрезание по  $C_2^{\beta=1}$ . ▶ СМ.  $\epsilon_{\text{signal}}$  VS  $\epsilon_{\text{bkg}}$

То же для  $\mu > 20$ ,  $\bar{\mu} \simeq 25$  (обрезания те же)



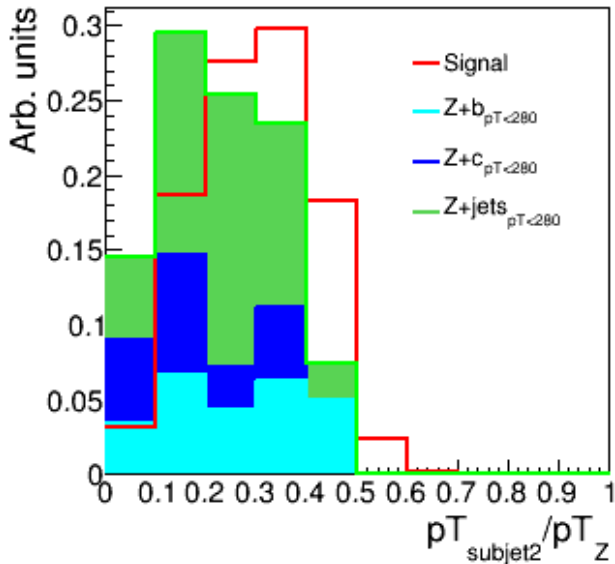
$200 < p_T^J < 250$  GeV,  $|y| < 1.2$  (cf. CERN-PH-EP-2015-204)



68% сигнала между двумя вертикальными штриховыми линиями. Здесь  $\bar{\mu} = 15$ , при  $\bar{\mu} \simeq 25$  разрешение для сигнала ухудшается.

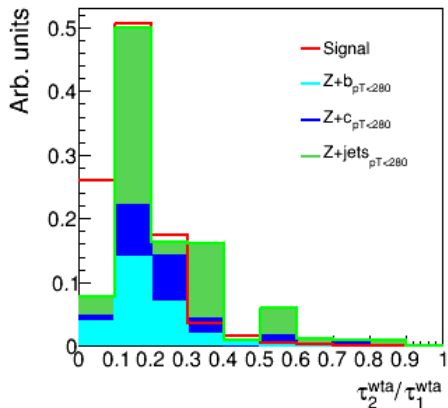
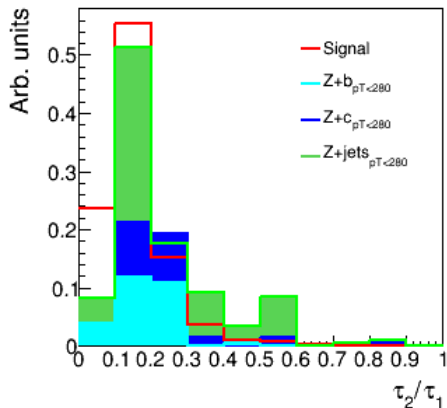


$200 < p_J^T < 250$  GeV,  $|y| < 1.2$ , обрезания:  $x_2$



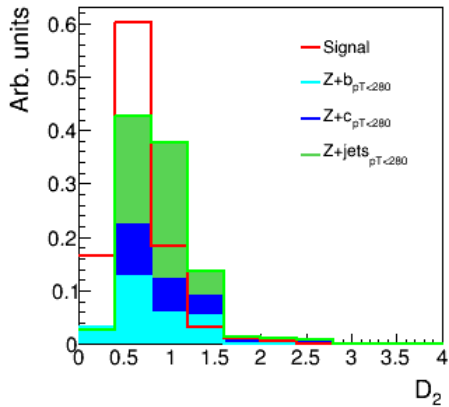
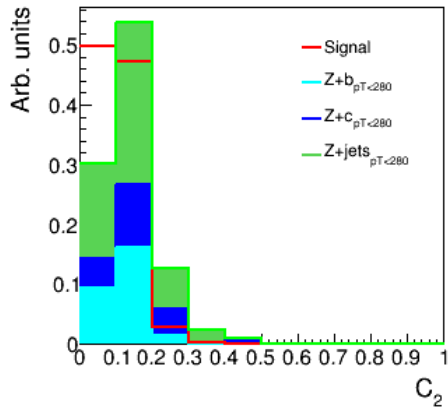
Основное подавление фона в сигнальном окне  $M_J$  за счет  $x_2 = p_T^{\text{subjet2}} / p_T^Z < 0.18$

$200 < p_T^J < 250$  GeV,  $|y| < 1.2$ , структ. переменные:  $\frac{\tau_2}{\tau_1}$



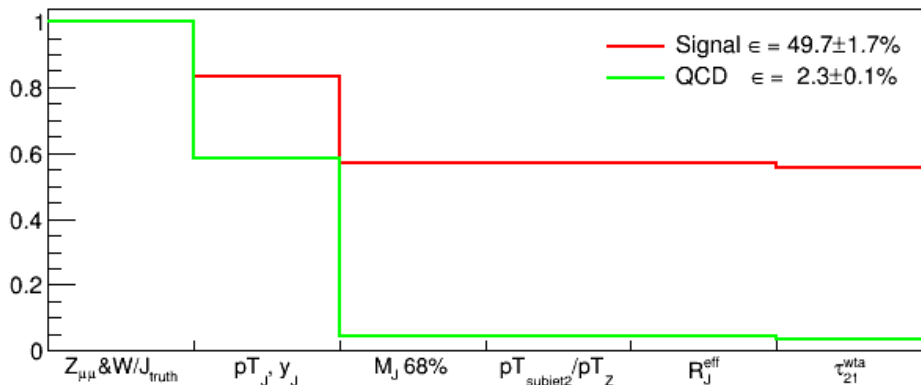
$\tau_2^{wta}/\tau_1^{wta} < 0.21$  дает  $\epsilon_{signal} = 50\%$ , можно обрезать и по другим переменным:  $C_2$ ,  $D_2$

$350 < p_T^J < 500$  GeV,  $|y| < 1.2$ , структурные переменные:  
 $C_2, D_2$



$200 < p_T^J < 250 \text{ GeV}, |y| < 1.2$

Эффект обрезаний по  $M_J, x_2, \tau_2^{wta} / \tau_1^{wta}$

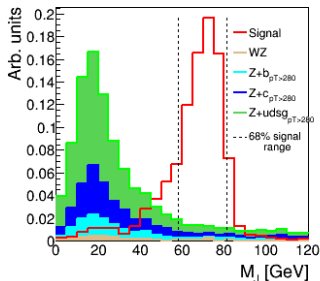
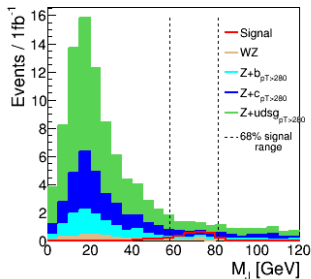


Ср. CERN-PH-EP-2015-204 (Fig. 15): в той же кинематике при  $\epsilon_{\text{signal}} = 50\%$  лучшее подавление фона  $\epsilon_{\text{QCD}} \simeq 2\%$  дает алгоритм anti- $k_T$   $R = 1.0$ , trimmed ( $f_{\text{cut}} = 5\%$ ,  $R_{\text{sub}} = 0.2$ ), обрезание по  $D_2^{\beta=1}$ .

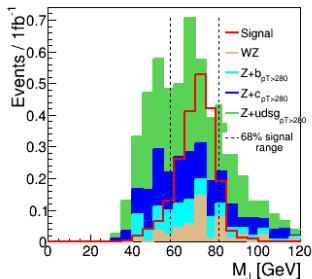
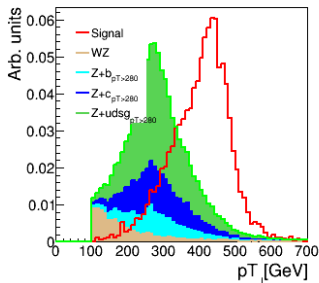
► СМ.  $\epsilon_{\text{signal}} \text{ VS } \epsilon_{\text{bkg}}$

$W'(1000\text{GeV}) \rightarrow W(J)Z(\mu^+\mu^-)$ , фон  $Z + jets$  и  $ZW$  без перевзвешивания

После ре-  
конструкции  
 $W \rightarrow J$ :

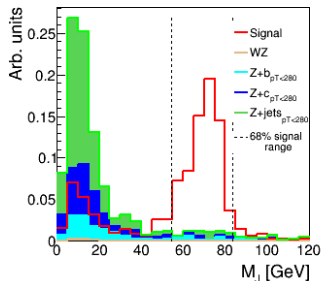
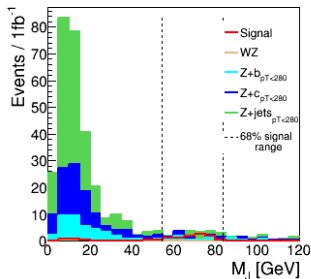


После обреза-  
ний:

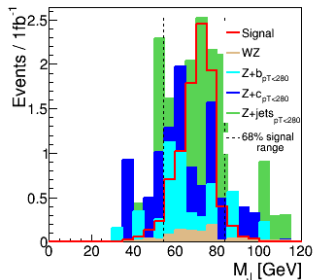
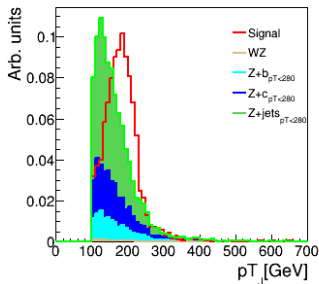


$W'(500\text{GeV}) \rightarrow W(J)Z(\mu^+\mu^-)$ , фон  $Z + jets$  и  $ZW$  без перевзвешивания

После ре-  
конструкции  
 $W \rightarrow J$ :



После обре-  
заний:



## Выводы, вопросы

- Метрика построения струи с “pileup-фильтром” дает результаты не хуже, чем использующиеся в АТЛАСе сложные алгоритмы A/C, AntiKt ( $R=0.6-1.2$ ) + “jet grooming” (cf. [CERN-PH-EP-2015-204](#))
- Новый алгоритм имеет меньше свободных параметров, чем используемые в АТЛАСе, и имеет прозрачный физический смысл.

Вопросы [TODO]:

- Resolved vs. unresolved  $V \rightarrow jj$  с  $Q_{max}/M_V \sim 0.1 - 0.2$ ?
- Другие структурные переменные?
- Калибровка  $E_J$ ,  $M_J$ : пик  $W$  у нас выглядит как узкий резонанс + “радиационный хвост”, так как теряется часть более мягкой суб-струи. Форму хвоста (то есть вероятность потери мягкого  $g$  в  $W \rightarrow qqg$ ) можно получить по известному распределению  $g$  в  $W \rightarrow qqg$  и известной плотности pileup  $\Rightarrow$  поправка к реконструированному импульсу и массе струи.
- Доложить в группе jet substructure.

# Backup



## Структурные переменные

$$R_J^{eff} = \sum_i p_{T_i} \Delta R_{i,J} / \sum_i p_{T_i}$$

$$\tau_0 = \sum_i p_{T_i}, \quad \tau_1 = \frac{1}{\tau_0} \sum_i p_{T_i} \Delta R(i, J)$$

$$\tau_1^{wta} = \frac{1}{\tau_0} \sum_i p_{T_i} \Delta R(i, \text{hardest in } J)$$

$$\tau_2^{wta} = \frac{1}{\tau_0} \sum_i p_{T_i} \min \{ \Delta R(i, \text{hardest in } \text{subject}_1), \Delta R(i, \text{hardest in } \text{subject}_2) \}$$

$$\tau_{21}^{wta} = \tau_2^{wta} / \tau_1^{wta}$$

## Структурные переменные: корреляторы

$$E_{CF} 0(\beta) = 1,$$

$$E_{CF} 1(\beta) = \sum_{i \in J} p_{T_i},$$

$$E_{CF} 2(\beta) = \sum_{i < j \in J} p_{T_i} p_{T_j} (\Delta R_{ij})^\beta,$$

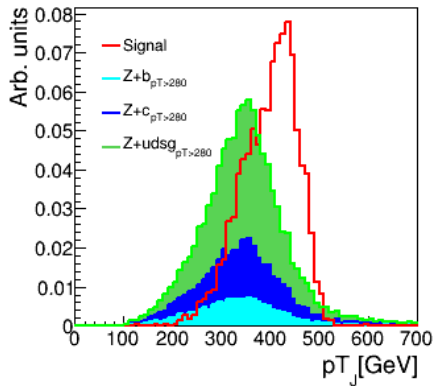
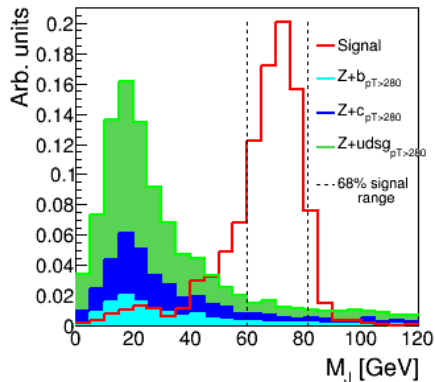
$$E_{CF} 3(\beta) = \sum_{i < j < k \in J} p_{T_i} p_{T_j} p_{T_k} (\Delta R_{ij} \Delta R_{ik} \Delta R_{jk})^\beta,$$

$$e_2^{(\beta)} = \frac{E_{CF} 2}{E_{CF} 1^2}, \quad C_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^2},$$

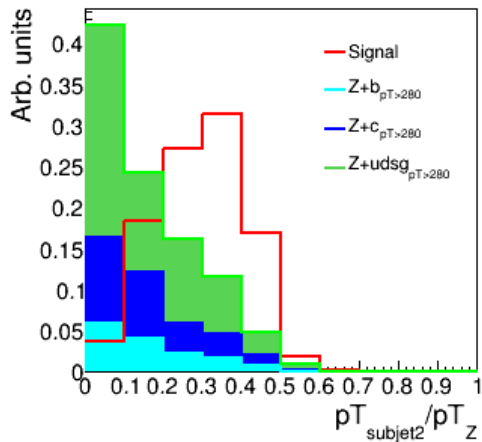
$$e_3^{(\beta)} = \frac{E_{CF} 3}{E_{CF} 1^3}, \quad D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$$

# Без перевзвешивания $Z + jets$

$350 < p_T^J < 500$  GeV,  $|y| < 1.2$  (cf. CERN-PH-EP-2015-204)

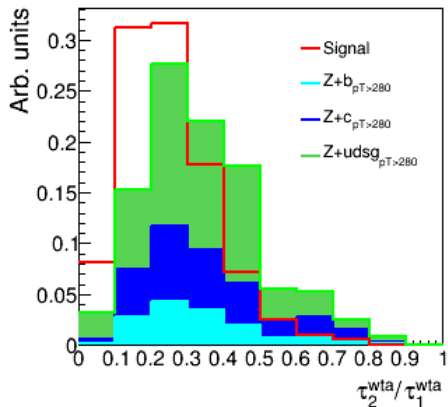
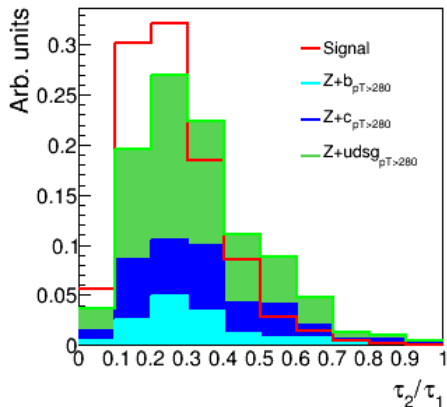


$350 < p_J^T < 500$  GeV,  $|y| < 1.2$ , обрезания:  $x_2$



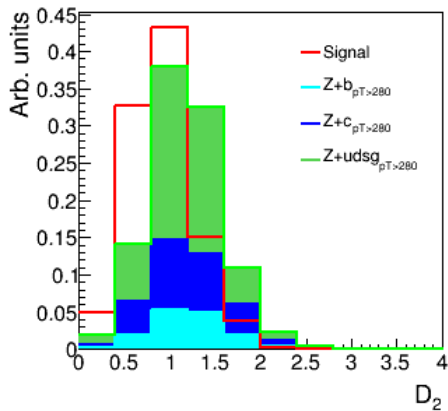
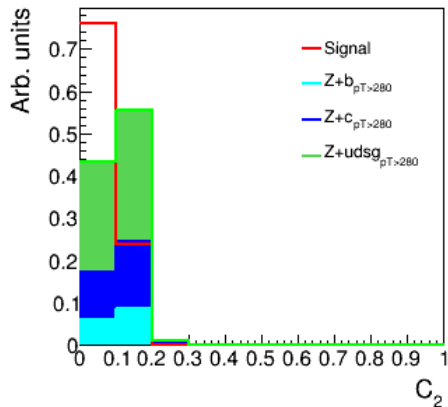
Основное подавление фона в сигнальном окне  $M_J$  за счет  $x_2 = p_T^{\text{subject2}} / p_T^Z < 0.1$

$350 < p_T^J < 500$  GeV,  $|y| < 1.2$ , структ. переменные:  $\frac{\tau_2}{\tau_1}$



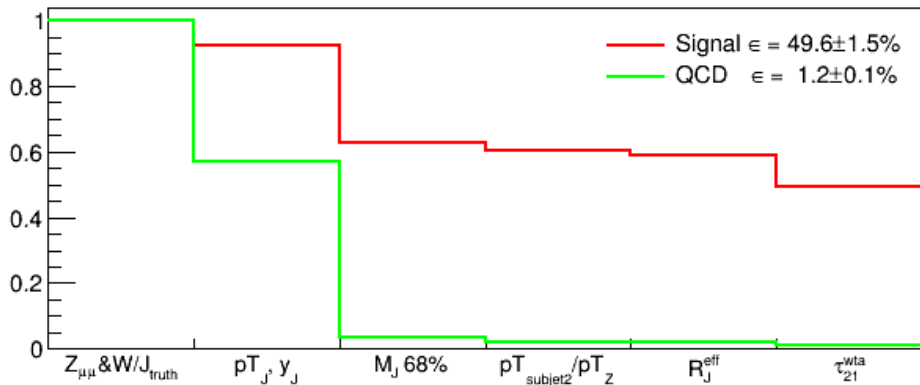
$\tau_2^{wta}/\tau_1^{wta} < 0.37$  дает  $\epsilon_{signal} \simeq 50\%$

$350 < p_T^J < 500$  GeV,  $|y| < 1.2$ , структурные переменные:  
 $C_2, D_2$



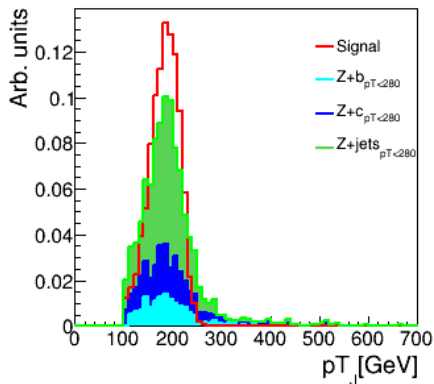
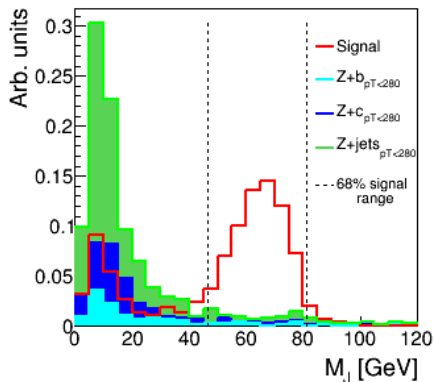
$350 < p_T^J < 500 \text{ GeV}, |y| < 1.2$

Эффект обрезаний по  $M_J, x_2, \tau_2^{wta} / \tau_1^{wta}$

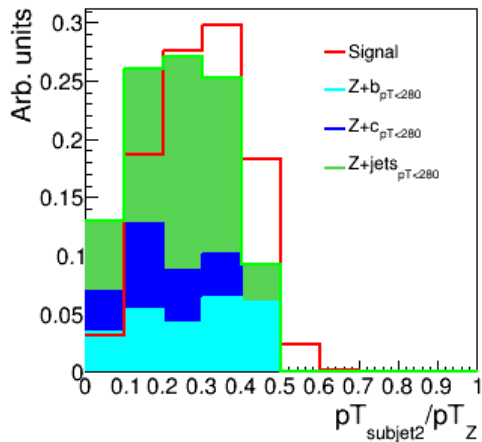




$200 < p_T^J < 250$  GeV,  $|y| < 1.2$  (cf. CERN-PH-EP-2015-204)

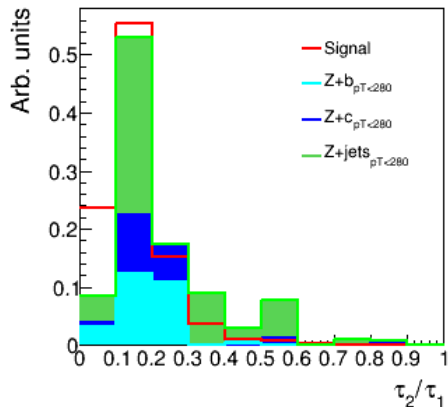


$200 < p_J^T < 250$  GeV,  $|y| < 1.2$ , обрезания:  $x_2$

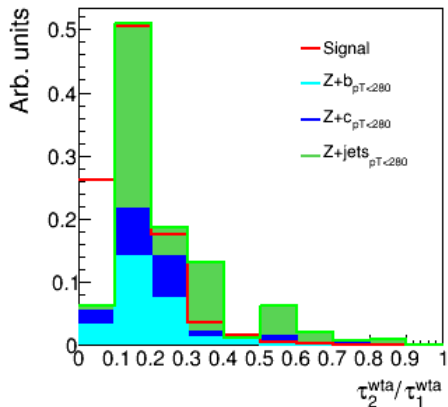


Основное подавление фона в сигнальном окне  $M_J$  за счет  $x_2 = p_T^{\text{subjet2}} / p_T^Z < 0.18$

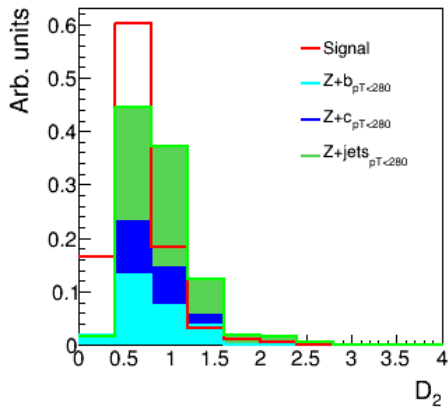
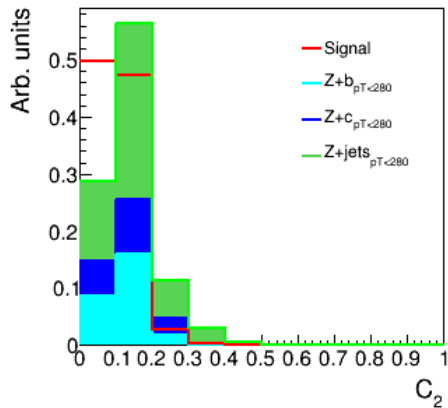
$200 < p_T^J < 250$  GeV,  $|y| < 1.2$ , структ. переменные:  $\frac{\tau_2}{\tau_1}$



$\tau_2^{wta}/\tau_1^{wta} < 0.24$  дает  $\epsilon_{signal} \simeq 50\%$

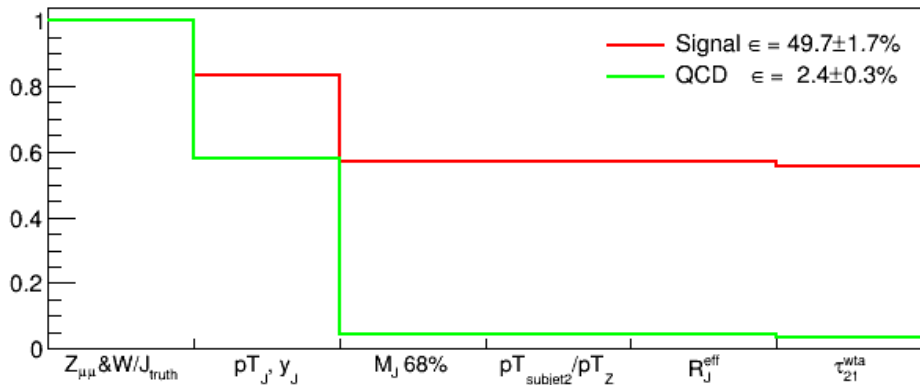


$200 < p_T^J < 250$  GeV,  $|y| < 1.2$ , структурные переменные:  
 $C_2, D_2$

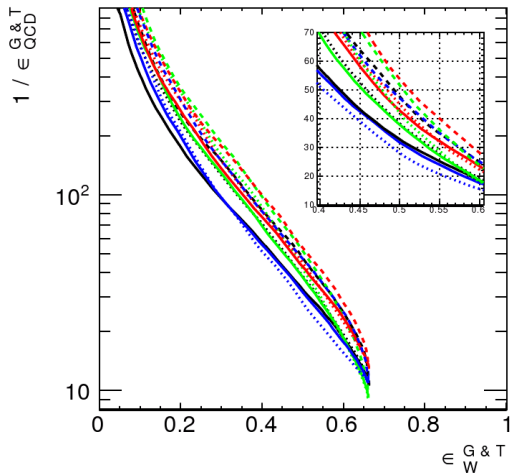


$200 < p_T^J < 250 \text{ GeV}, |y| < 1.2$

Эффект обрезаний по  $M_J, x_2, \tau_2^{wta} / \tau_1^{wta}$



$\epsilon_{\text{signal}}$  VS.  $\epsilon_{\text{QCD bkg}}$ ,  $350 < p_T(J) < 500$  GeV  
(CERN-PH-EP-2015-204)



**ATLAS** Simulation Internal

$\sqrt{s} = 8$  TeV Jet 4-momentum not calibrated

$|\eta^{\text{Truth}}| < 1.2$ ,  $350 < p_T^{\text{Truth}} < 500$  GeV, M Cut

- $C_2^{(\beta=1)}$  with anti- $k_t$   $R=1.0$  jets
- $D_2^{(\beta=1)}$  Trimmed ( $f_{\text{cut}}=5\%$ ,  $R_{\text{sub}}=0.2$ )
- $\tau_{21}^{\text{wta}}$
  
- $C_2^{(\beta=1)}$  with anti- $k_t$   $R=1.0$  jets
- $D_2^{(\beta=1)}$  Trimmed ( $f_{\text{cut}}=5\%$ ,  $R_{\text{sub}}=0.3$ )
- $\tau_{21}^{\text{wta}}$
  
- $C_2^{(\beta=1)}$  with C/A  $R=1.0$  jets
- $D_2^{(\beta=1)}$  Pruned ( $R_{\text{cut}}=0.5$ ,  $Z_{\text{cut}}=0.15$ )
- $\tau_{21}^{\text{wta}}$
  
- $C_2^{(\beta=1)}$  with C/A  $R=1.2$  jets
- $D_2^{(\beta=1)}$  Split-Filtered ( $\mu=1$ ,  $R_{\text{sub}}=0.3$ ,  $y_{\text{filt}}=15\%$ )
- $\tau_{21}^{\text{wta}}$

$\epsilon_{\text{signal}}$  VS.  $\epsilon_{\text{QCD bkg}}$ ,  $200 < p_T(J) < 250$  GeV  
 (CERN-PH-EP-2015-204)

