

b-tagging at ~~LHC~~
ATLAS

花垣和則 / 大阪大学

Overview

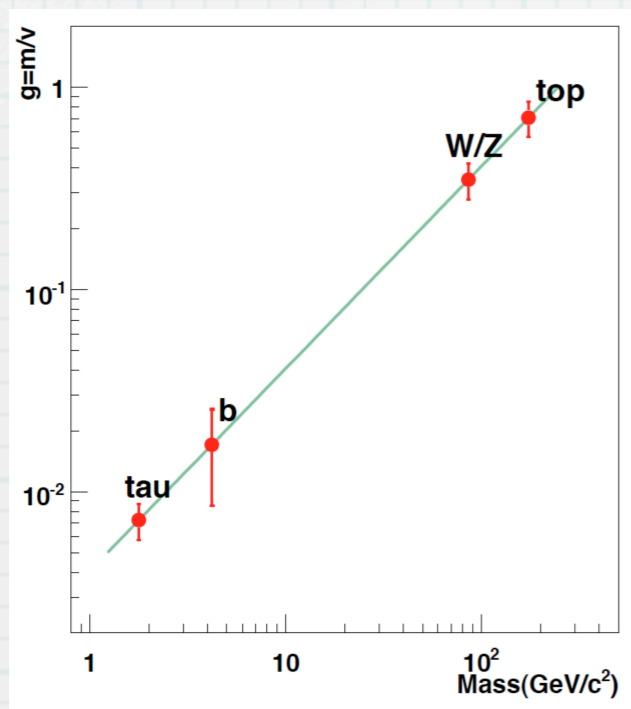
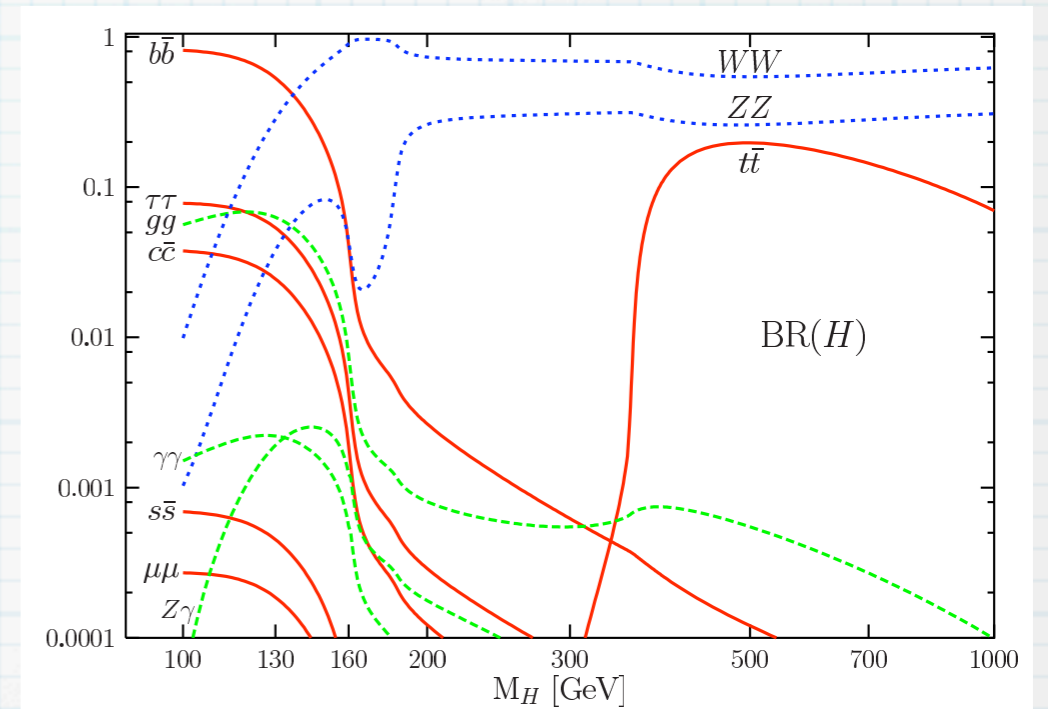
- ◆ b-quark起源のジェット (b-jet)を同定すること
 - ❖ ジェットの起源となるパートの識別
 - ◆ b-jet, c-jet, tau-jet, light-jet (u,d,s,gluon)

- ◆ このトークの流れ
 - ❖ b-taggingがなぜ必要か
 - ❖ Overview -- b-taggingのためのアイデア --
 - ❖ アルゴリズムの説明とMCでのperformance
 - ❖ データ取得に向けて

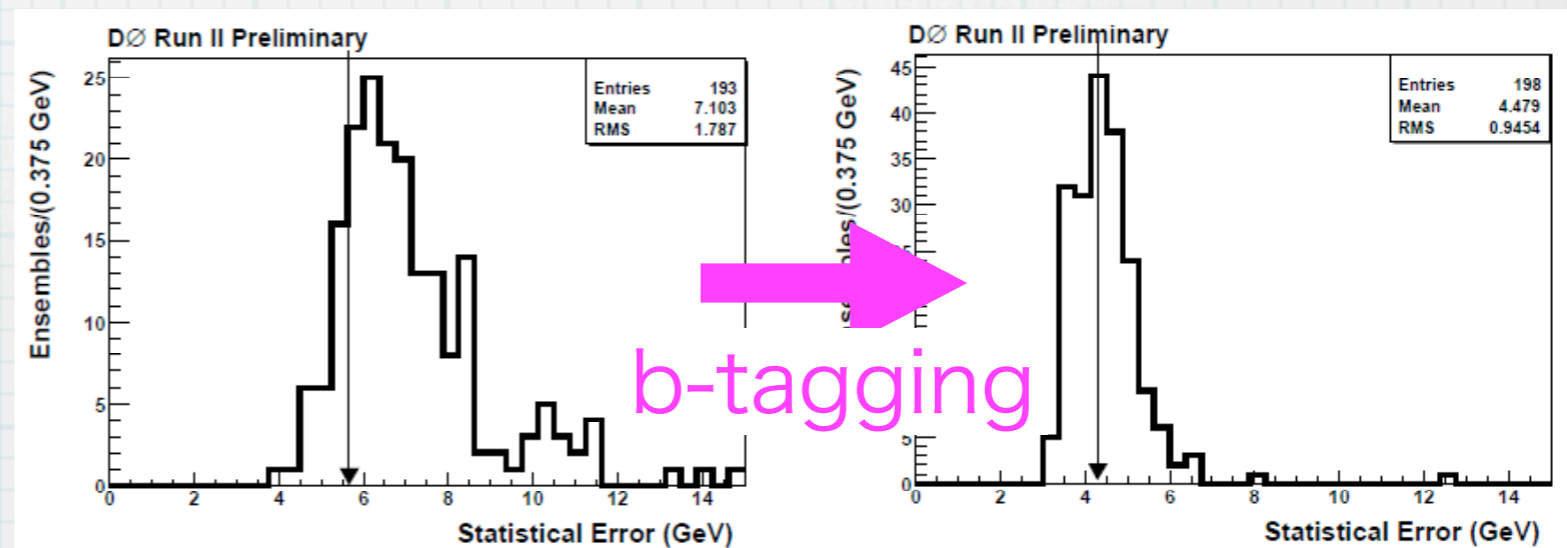
- ◆ (注) このトークは私の偏見で構成されています

Motivation

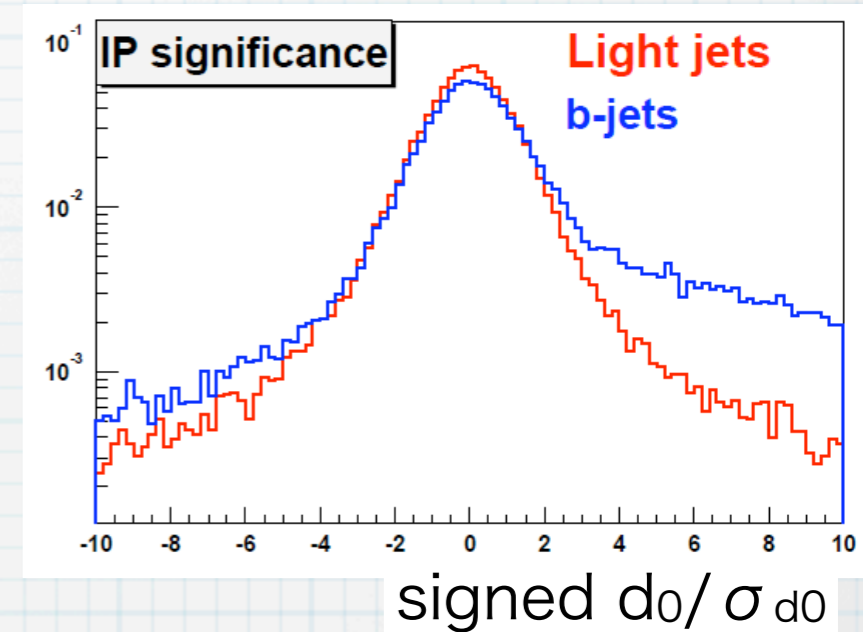
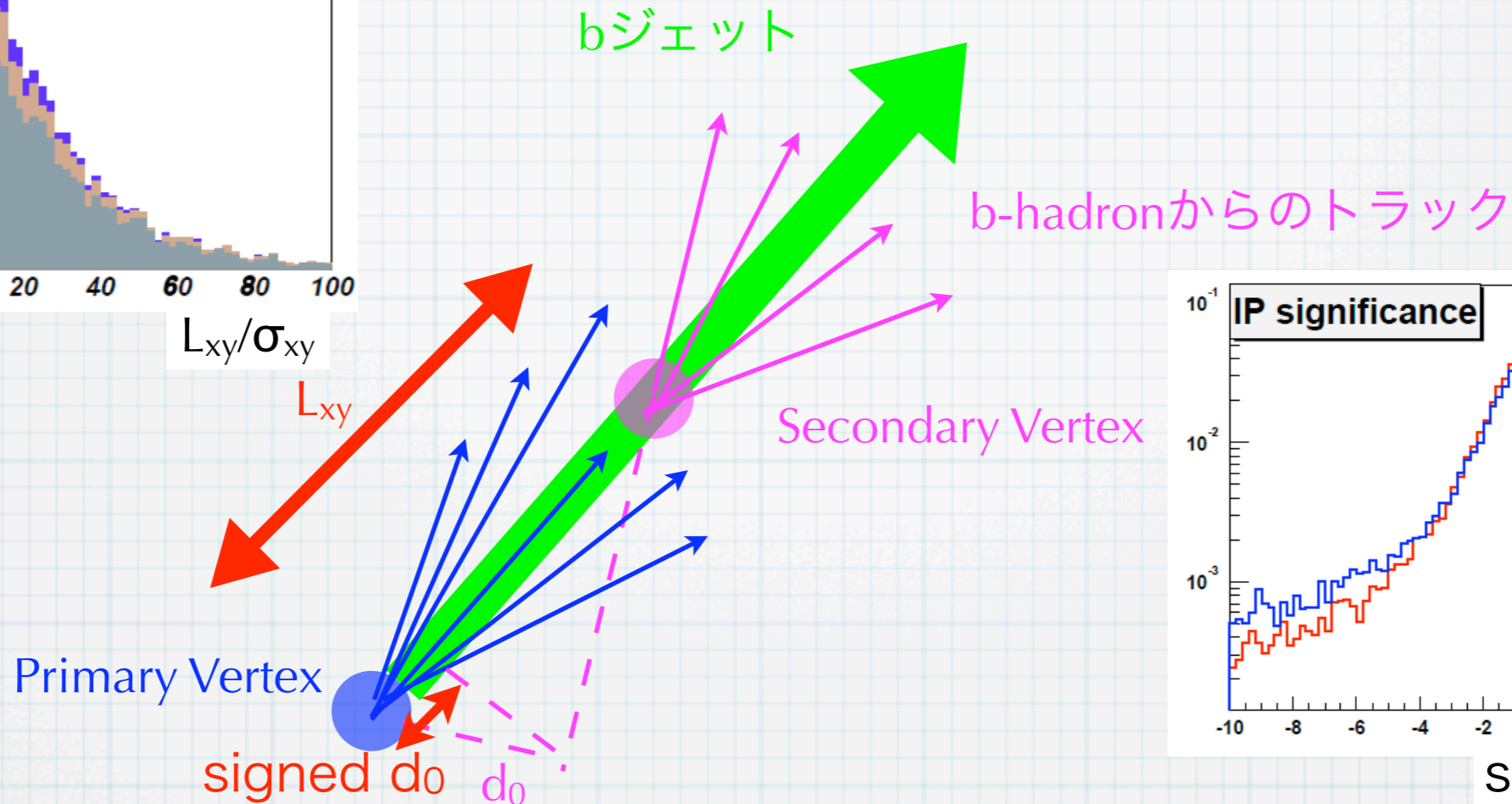
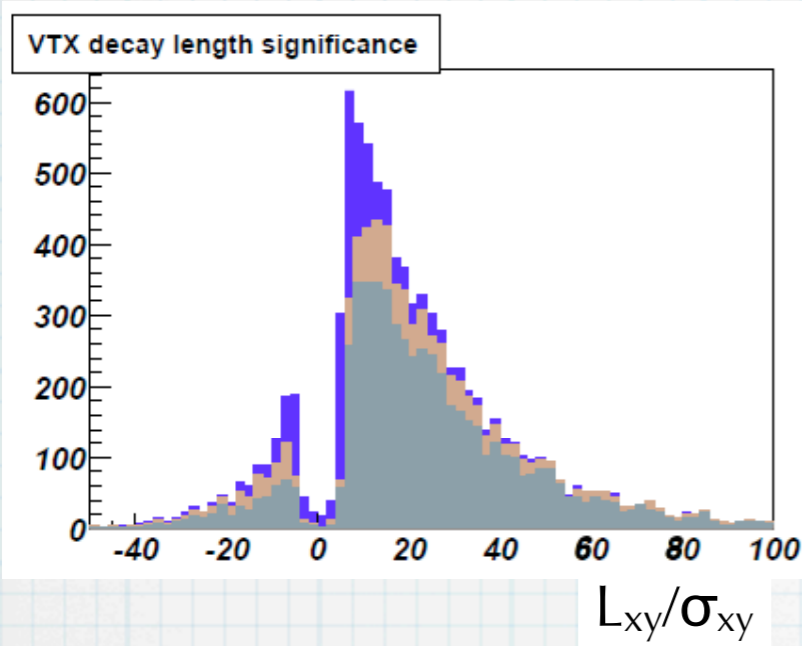
- ◆ 終状態にb-quarkを含む物理
 - ❖ Higgsが軽いとき $H \rightarrow b\bar{b}$
 - ◆ $t\bar{t}H$, $b\bar{b}H/A$, $q\bar{q}H$ etc.
 - ◆ もしHiggsが発見されたら湯川結合の測定



- ◆ Some SUSY signature
- ◆ top : $\text{Br}(t \rightarrow bW) \sim 100\%$
 - ❖ 例として質量測定

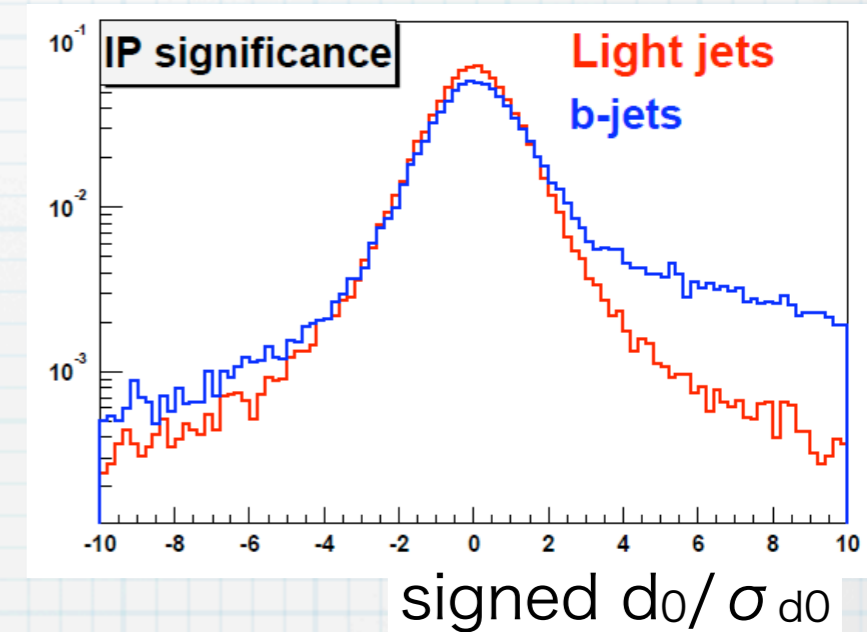
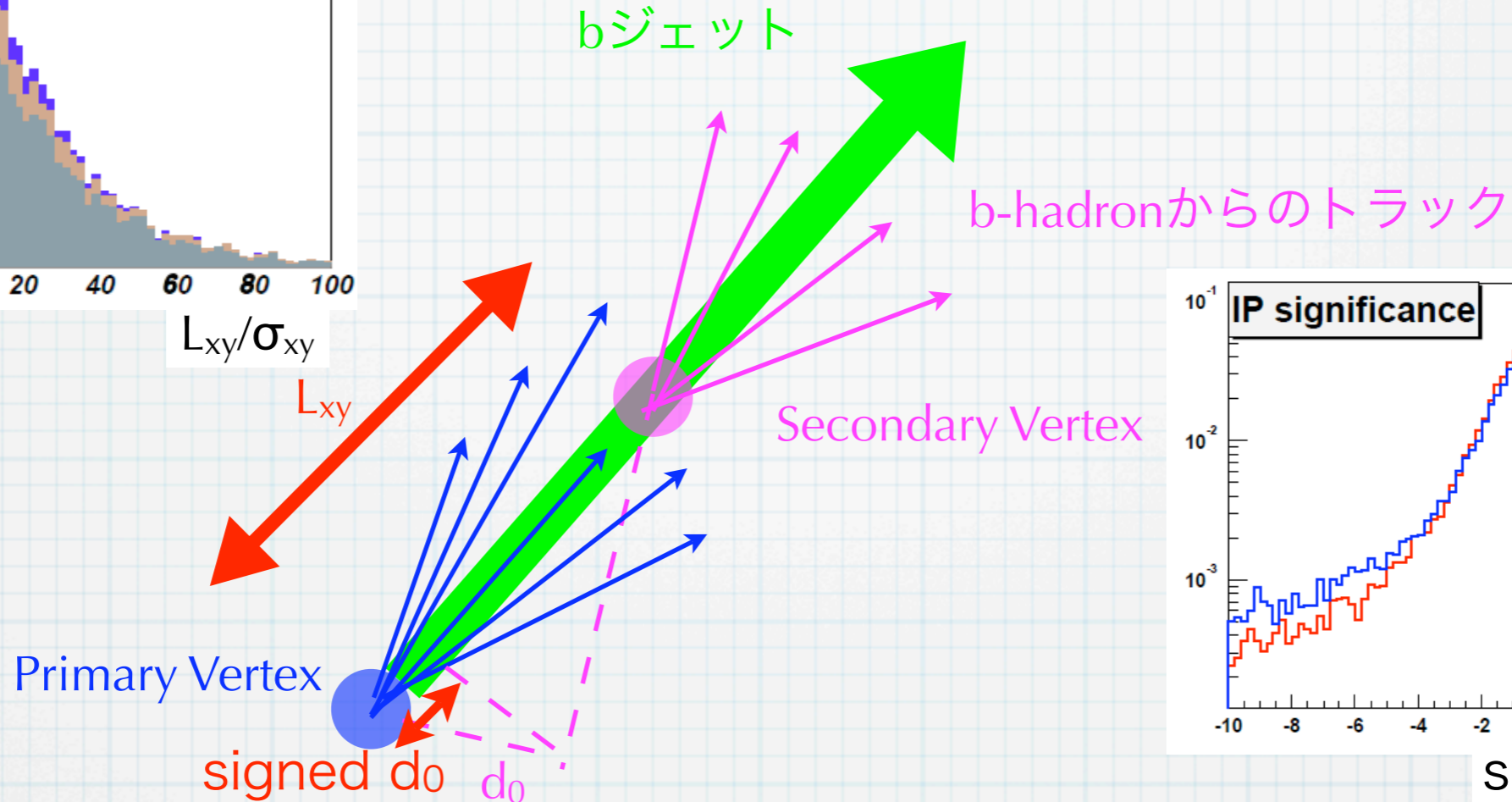
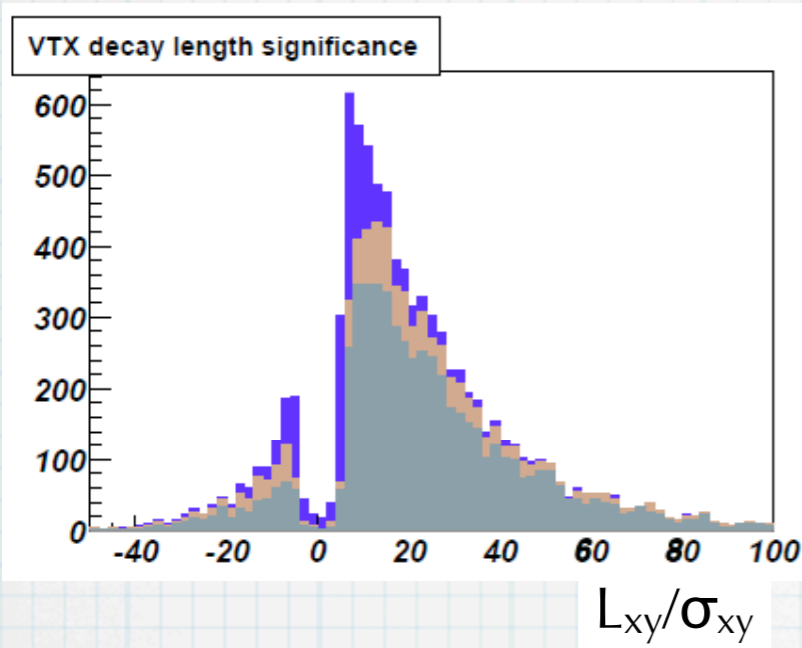


b-taggingの概略



- ◆ b-hadronは崩壊するまでに数ミリ飛べる ($c\tau \sim 400-500 \mu\text{m}$)
 - ❖ 大きなインパクトパラメータdを持つトラック
 - ❖ secondary vertex の存在 (大きな L_{xy})
 - ← si-pixel, si-strip で荷電粒子の飛跡を精度よく測定
- ◆ $\text{Br}(b \rightarrow lX) \sim 11\% + \text{Br}(b \rightarrow c \rightarrow lX) \sim 11\%$ ($l=e$ or μ)
 - ← ジェット近傍のミュオン (あるいは電子) を見つける

b-taggingの概略

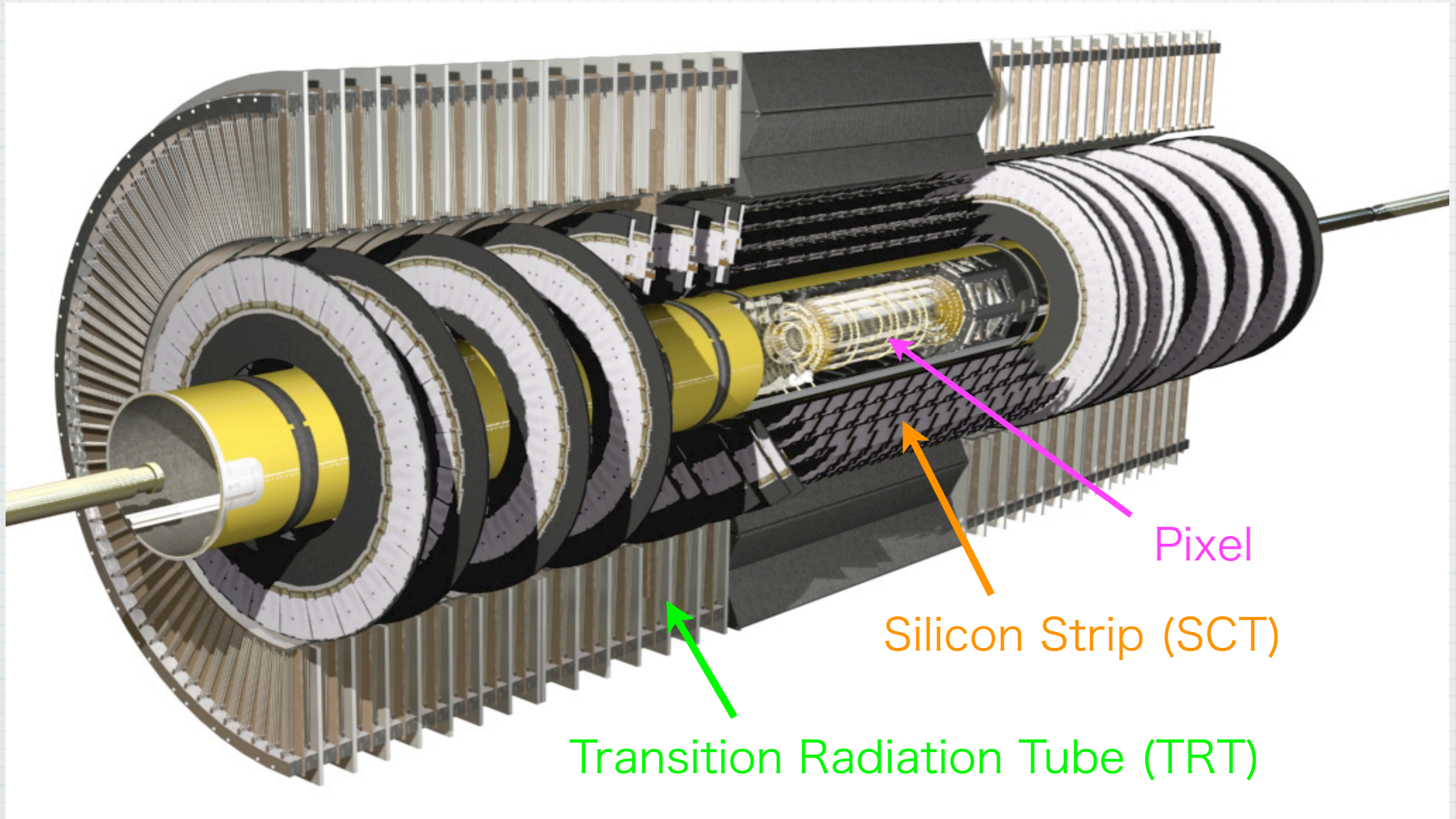


- ◆ b-hadronは崩壊するまでに数ミリ飛べる ($c\tau \sim 400-500 \mu\text{m}$)
 - ❖ 大きなインパクトパラメータdを持つトラック
 - ❖ secondary vertex の存在 (大きな L_{xy})



今日の話

ATLAS 内部飛跡検出器群



アルゴリズムの説明と

MCでのperformance

インパクトパラメータによるtagging

- ◆ Signed Impact Parameter significance ($\equiv S_{IP}$)を使う

- ❖ Counting method

- ◆ ある閾値を設けて大きな S_{IP} を持つトラックの数を数える

- ❖ Jet Probability

- ◆ $\mathcal{R}(s)$ をresolution function として

$$\mathcal{P}_{trk} = \frac{\int_{-a}^{|S_{IP}|} \mathcal{R}(s) ds}{\int_{-a}^0 \mathcal{R}(s) ds}$$

← light jet から来る
トラックらしさ

$$\mathcal{P}_{jet}^{\pm} = \Pi^{\pm} \times \sum_{j=0}^{N_{trk}^{\pm}-1} \frac{(-\log \Pi^{\pm})^j}{j!} \quad \text{with} \quad \Pi^{\pm} = \prod_{i=1}^{N_{trk}^{\pm}} \mathcal{P}_{trk}(S_{IP}^{\pm} > 0)$$

- ❖ Likelihood ratio

- ◆ b-, c-, light-jet それぞれについて S_{IP} のlikelihoodを用意する

$$W_{Jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(S_i)}{u(S_i)}$$

simple
robust

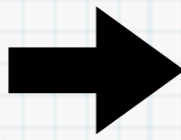
Tevatron
でも使用

sophisticated
not robust

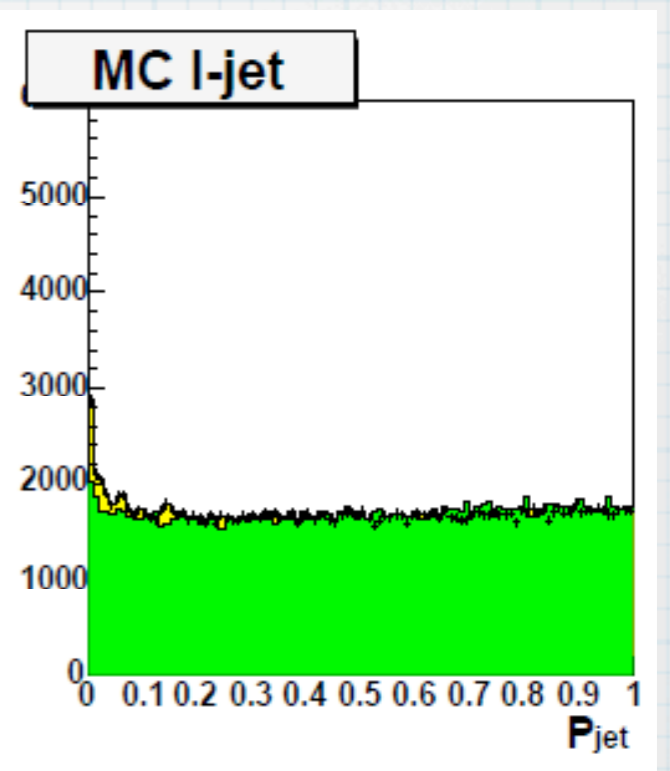
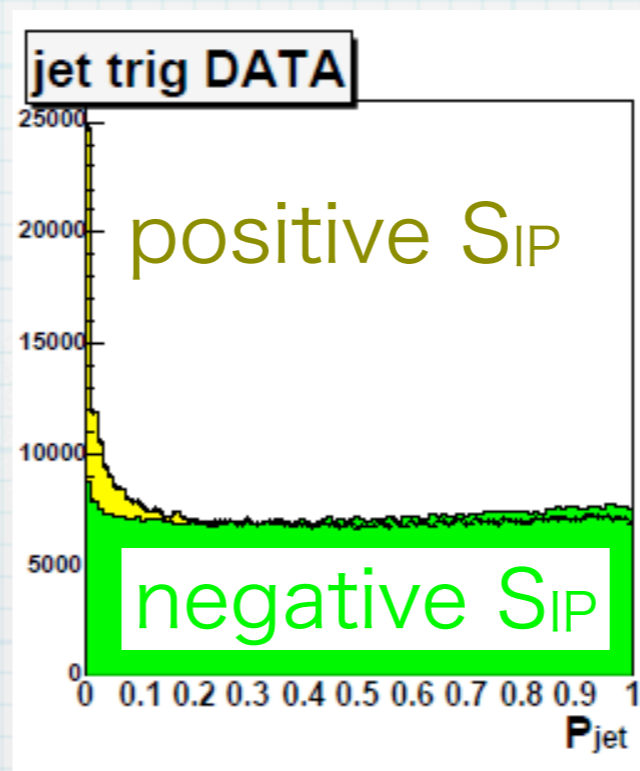
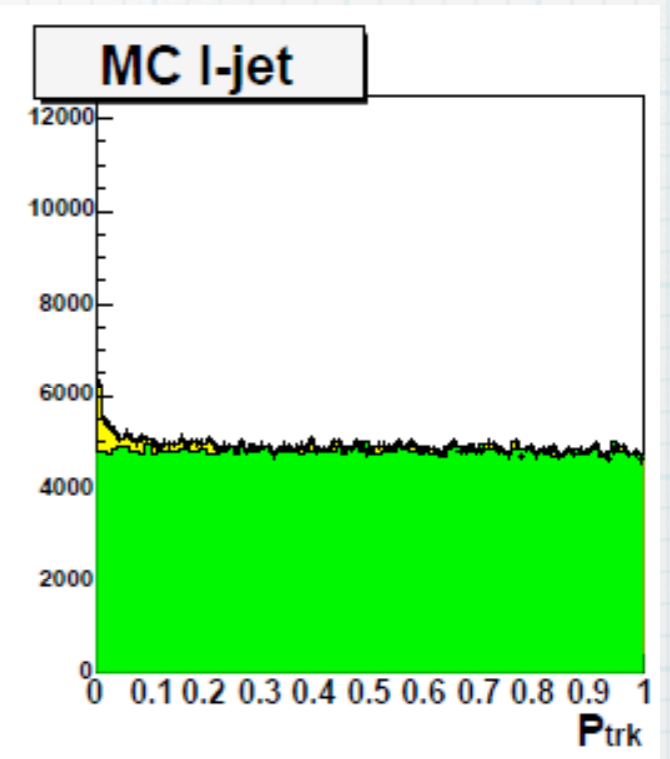
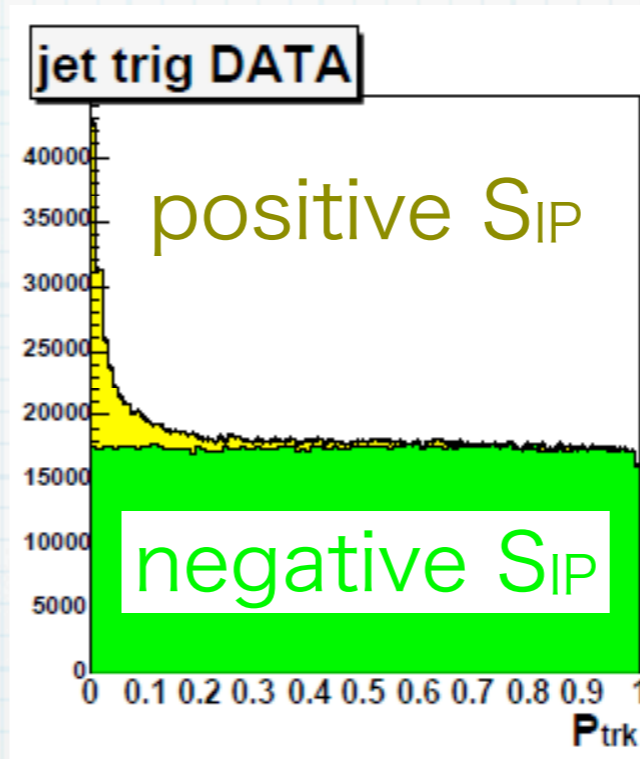
Jet Probability

- ◆ 例としてDØでの分布

$$\mathcal{P}_{trk} = \frac{\int_{-a}^{|S_{IP}|} \mathcal{R}(s) ds}{\int_{-a}^0 \mathcal{R}(s) ds}$$



negative S_{IP} を使うので
 (=light jetだけでよい)
 calibration sample
 に困らない
 →実験初期でもreliable



$$\mathcal{P}_{jet}^{\pm} = \Pi^{\pm} \times \sum_{j=0}^{N_{trk}^{\pm}-1} \frac{(-\log \Pi^{\pm})^j}{j!}$$



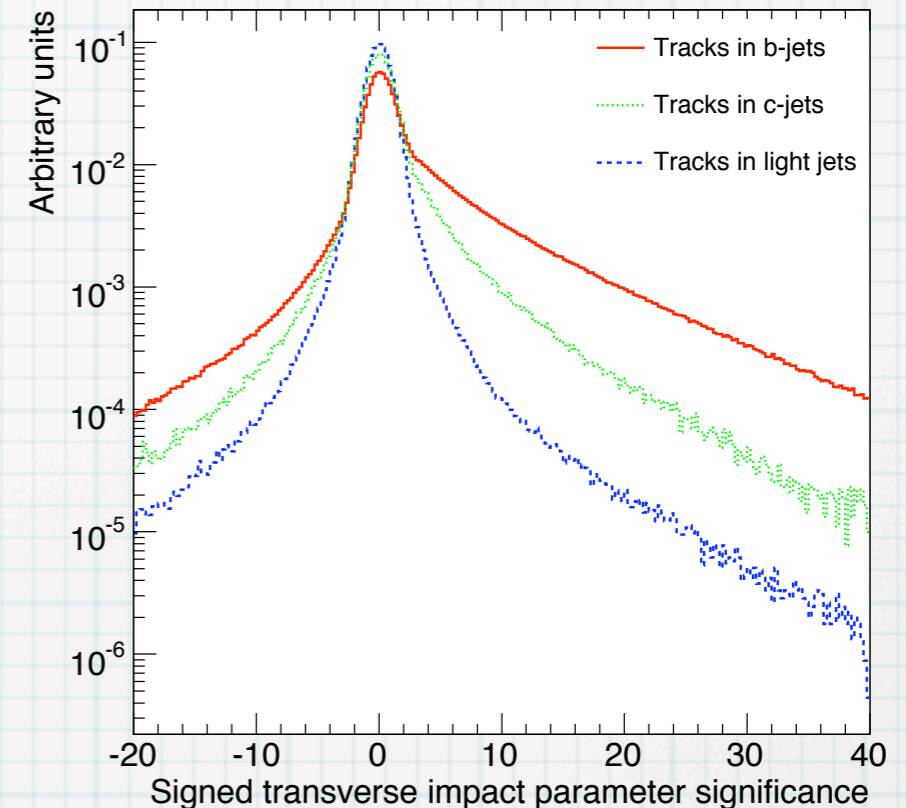
$$\Pi^{\pm} = \prod_{i=1}^{N_{trk}^{\pm}} \mathcal{P}_{trk}(S_{IP}^{\pm} > 0)$$

Likelihood Ratio

- ◆ S_{IP} の確率分布を b-, c-, light-jet それぞれに対して準備 (reference histogram)

$$W_{Jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(S_i)}{u(S_i)}$$

- ❖ likelihood ratio なので別の discriminant を簡単に足せる
- ❖ Jet Probability method では light-jet の S_{IP} だけを使っている
 - ◆ naïvè に W_{jet} のほうが discriminant power が大きいと予想できる
 - ◆ b(c)-jetに対するreferenceの作成が問題

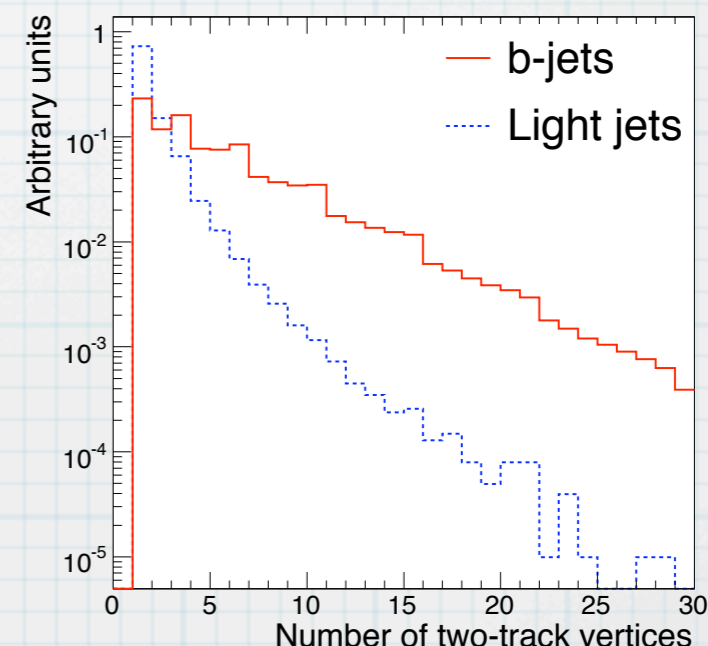
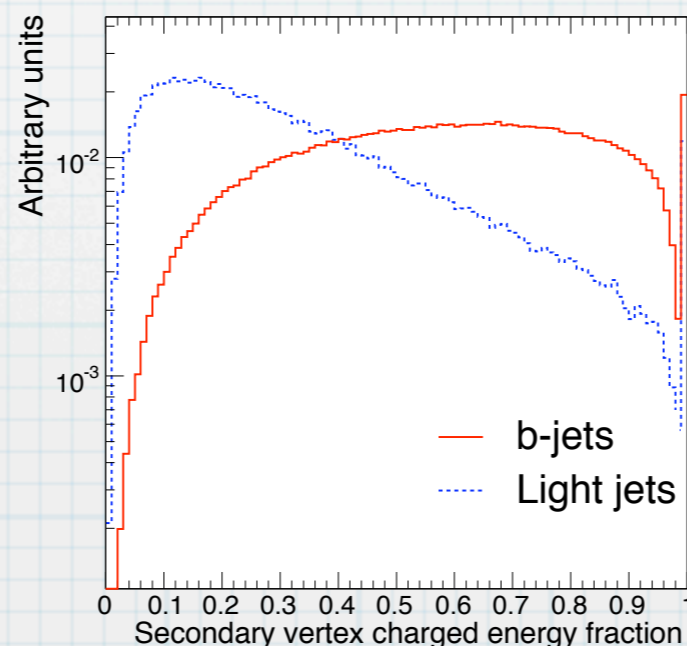
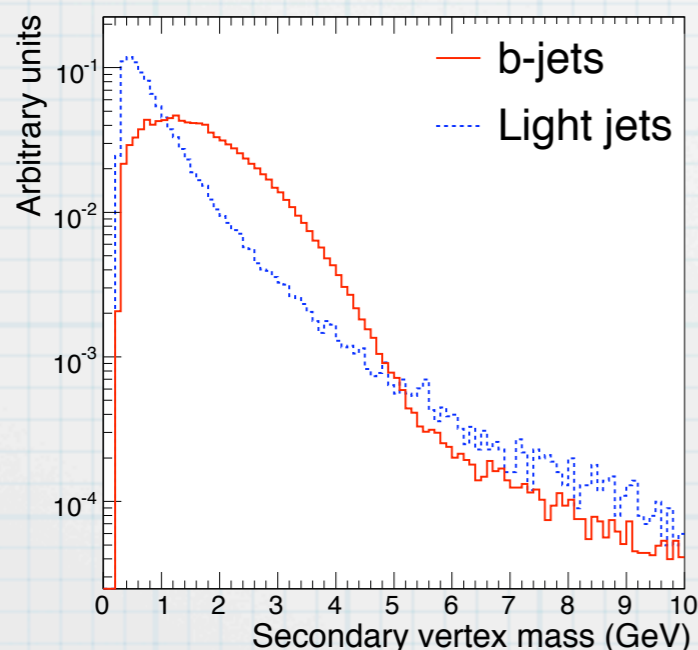


Secondary Vertex reconstruction

- ◆ p-p衝突地点から有意に離れた b-hadronの崩壊地点 (secondary vertex) を探す

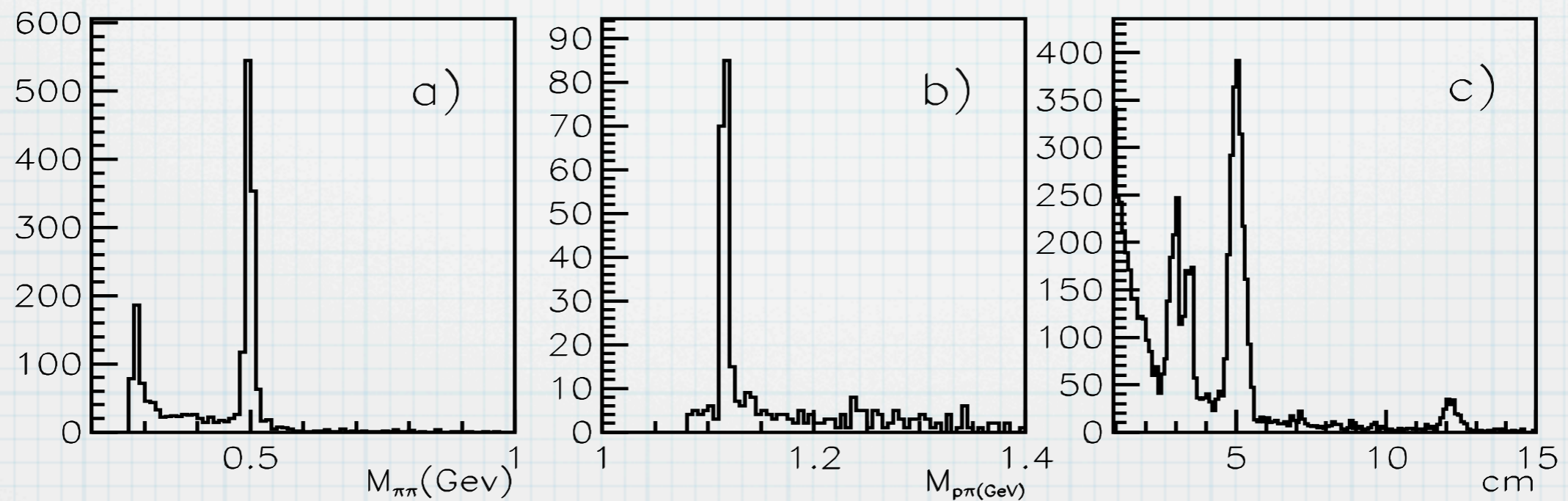
sec. vtx rec. efficiency	
Process	ϵ_b
WH(100 GeV) 3 layers	57.0%
WH(100 GeV) 2 layers	52.2%
WH(400 GeV) 3 layers	77.2%
WH(400 GeV) 2 layers	70.3%

- ❖ b-hadronの崩壊地点を測定してるわけだから
 - ◆ secondary vertex の再構築に使われたトラックを使って不変質量
 - ◆ 他にもdiscriminantとして使える変数あり

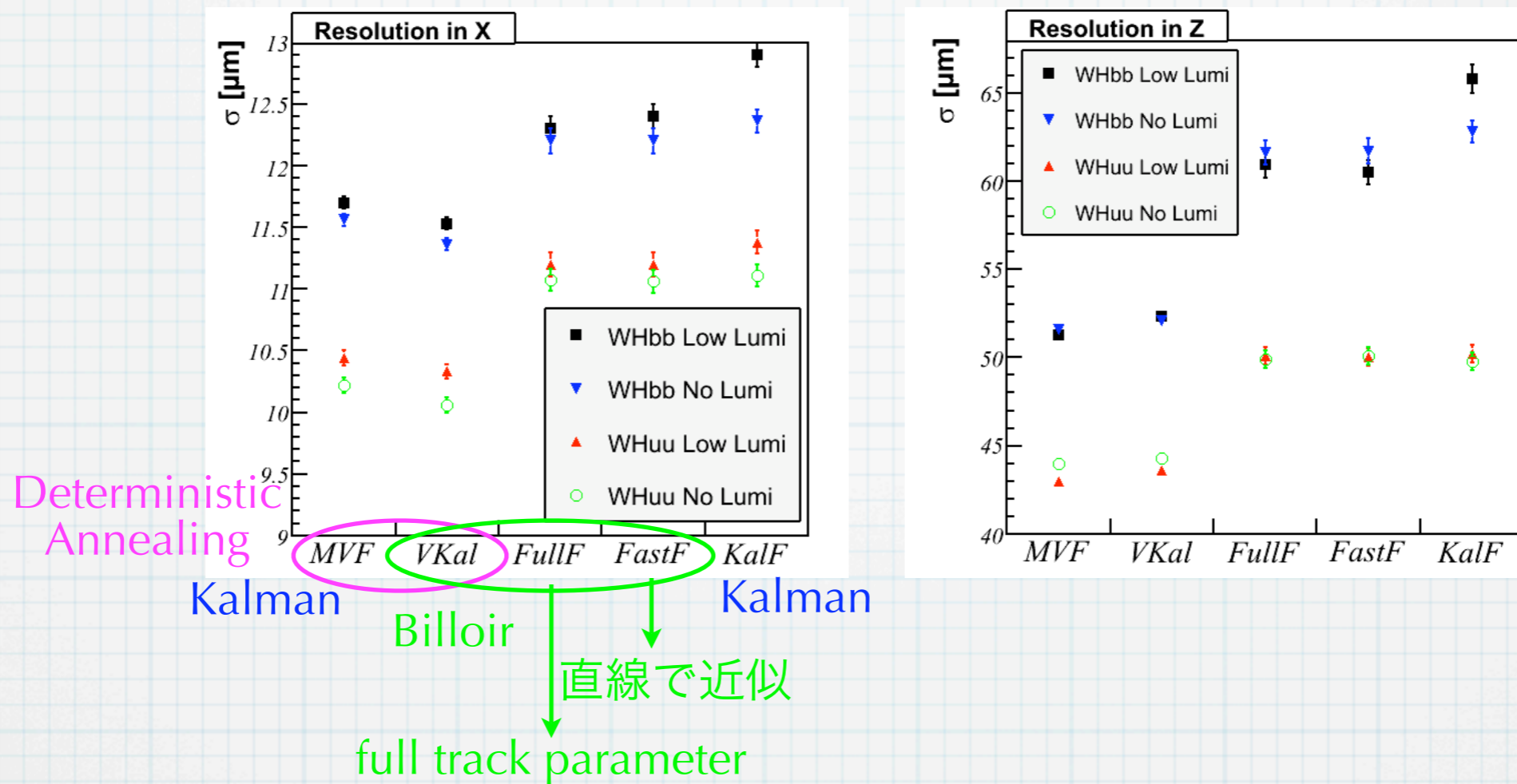


V^0 粒子のカット

- ◆ IP based, secondary vertex based どちらも light quark 起源の長寿命粒子 (Ks, Λ , photon conversion) は、2トラックの不変質量 (とvertexの位置) で除く



Vertexingについて

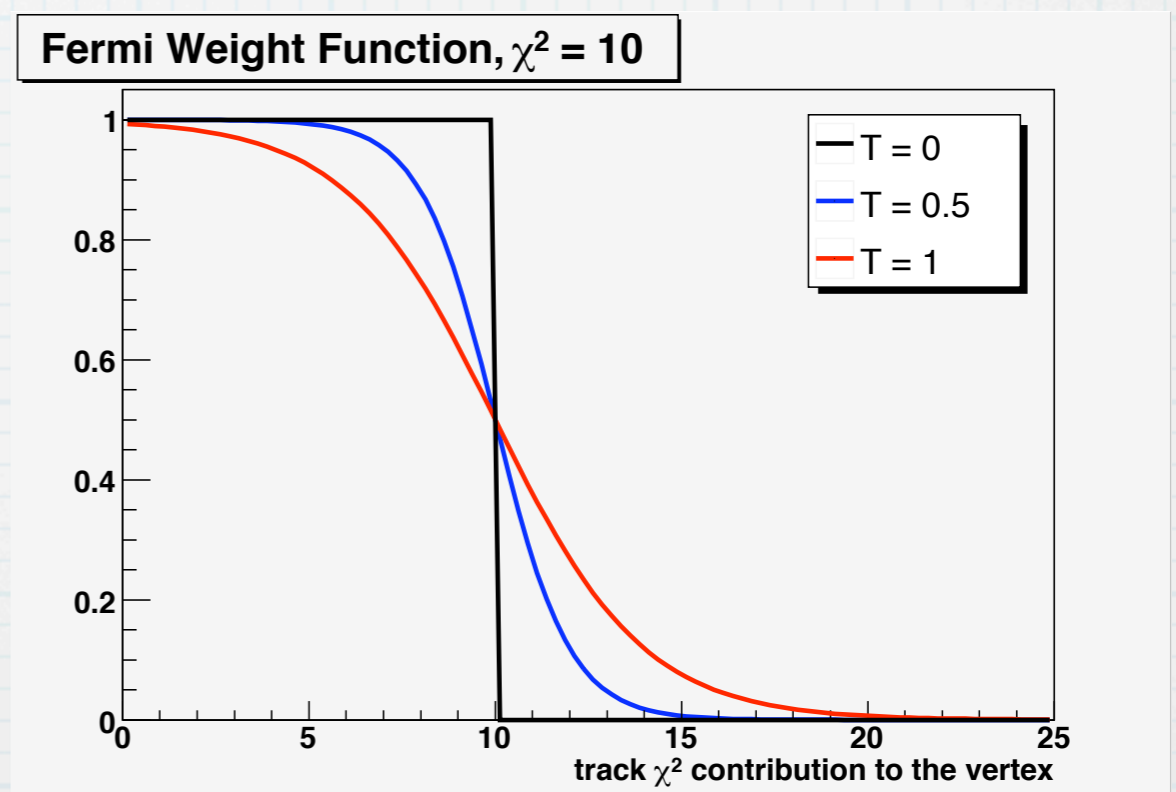


- ◆ ごちゃごちゃありますが、
 - ❖ 基本的には全部 Kalman Filter
 - ◆ Billoir method も formalism が違うだけ
 - ❖ Deterministic Annealing (adaptive fitter) だけは やっておいたほうがよい

Adaptive Vertex Fitter

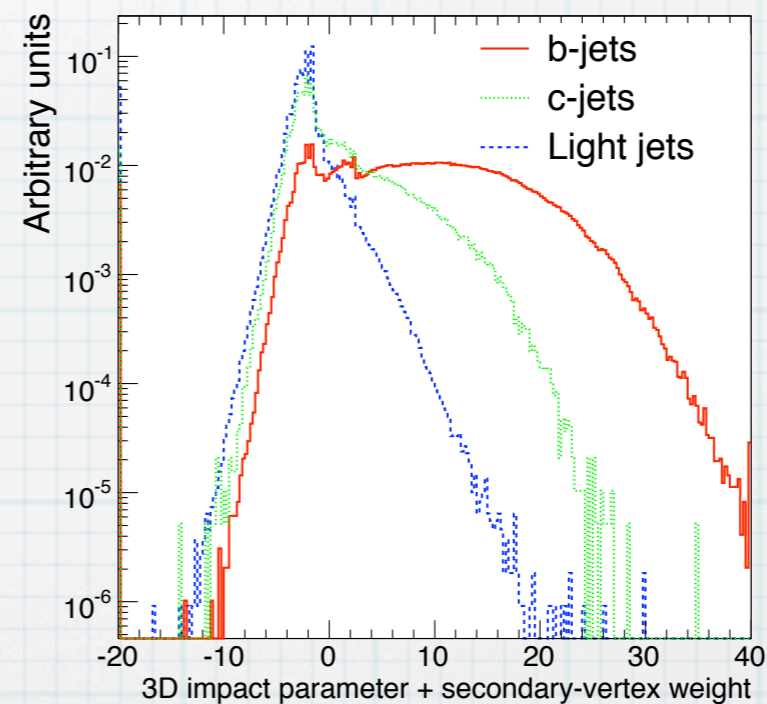
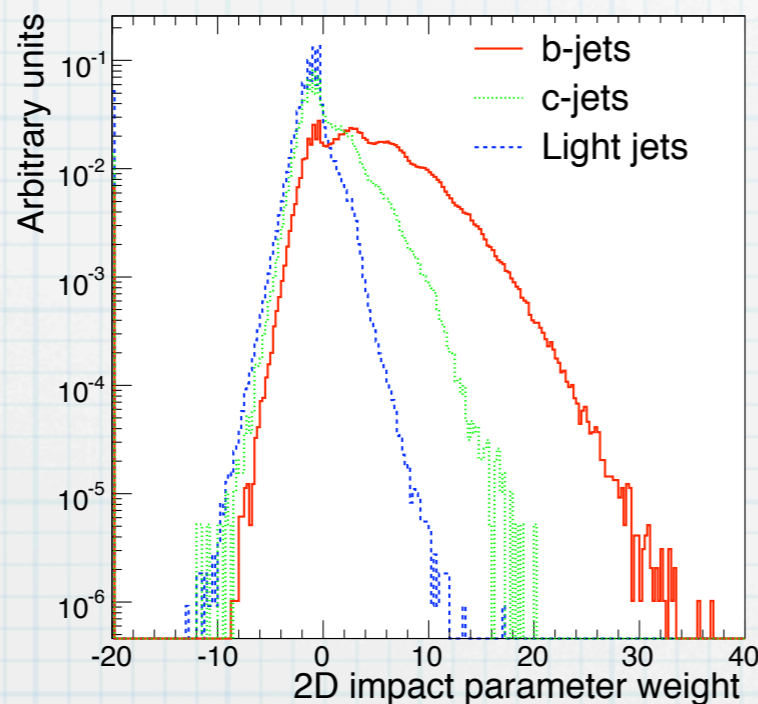
- ◆ ほとんど全てのトラックをエラーにweightをかけることで使う
⇔ 今までの方法ではエラー（など）に閾値を設けてクリーンなトラックだけ使っていた

$$w_i = \frac{1}{1 + e^{(\chi_i^2 - \chi_{cutoff}^2)/2T}}$$



実際に使われてるアルゴリズム

- ◆ ATLASのデフォルトはlikelihood ratio
 - ❖ 2D SIP + 1D SIP + 10ページの変数3つ
 - ❖ 他にstudyされてるのは
 - ◆ 2D SIP 単体、2D+1D SIP, JetProb., charm decay
まで考慮してのvertexing + default

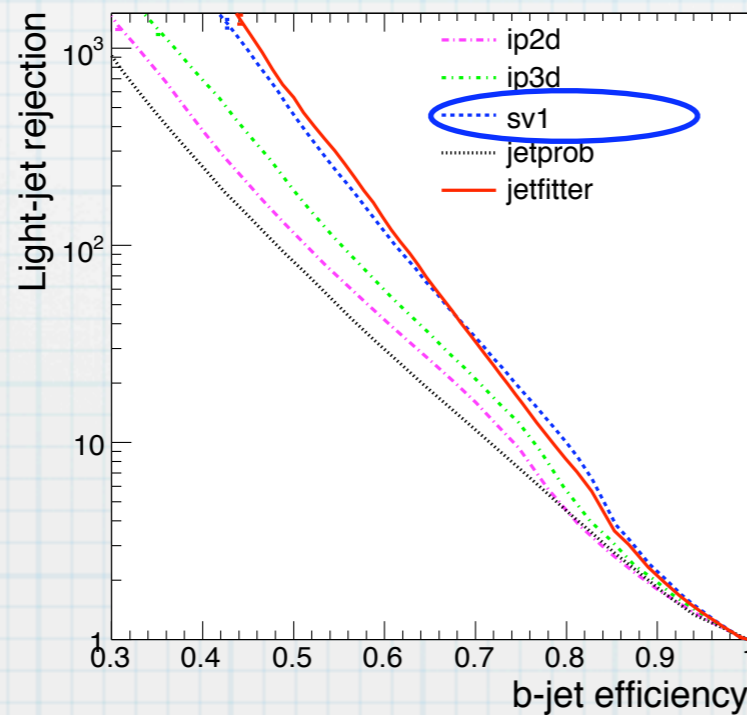


- ◆ CMSは JetProb. をデフォルトとして薦めてるらしい
 - ❖ 他は counting SIP, secondary vtx. reconstruction

Efficiency and Fake Rejection Factor in MC

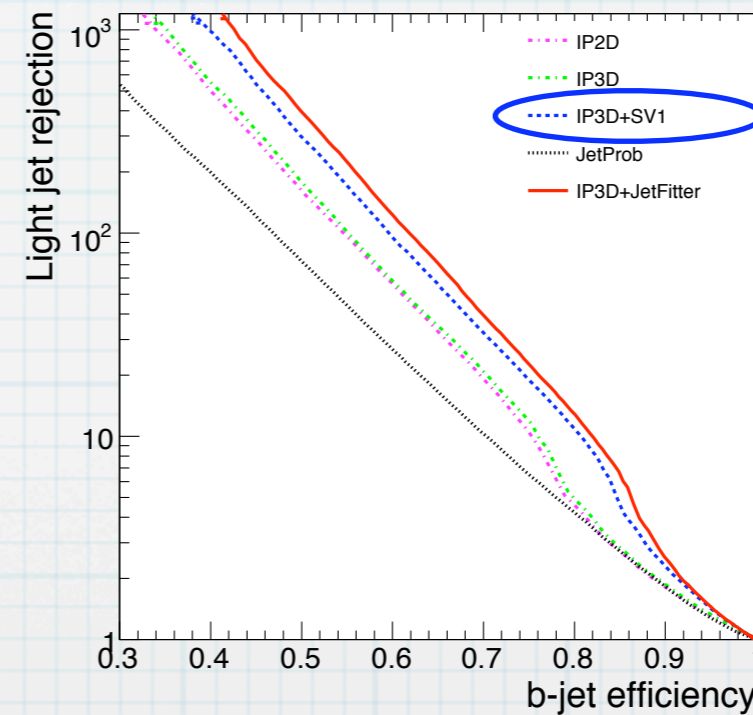
$$\varepsilon_q = \frac{\text{Number of jets of real flavour } q \text{ tagged as } b}{\text{Number of jets of real flavour } q}$$

- ◆ likelihood を作るための reference histogram は large tt, ttjj sample
 - ❖ 注意：jet p_T , η などでbin切りされていないので performance には sample dependence あり



(a) WH events ($m_H=120$ GeV)

比較的ソフトなprocess

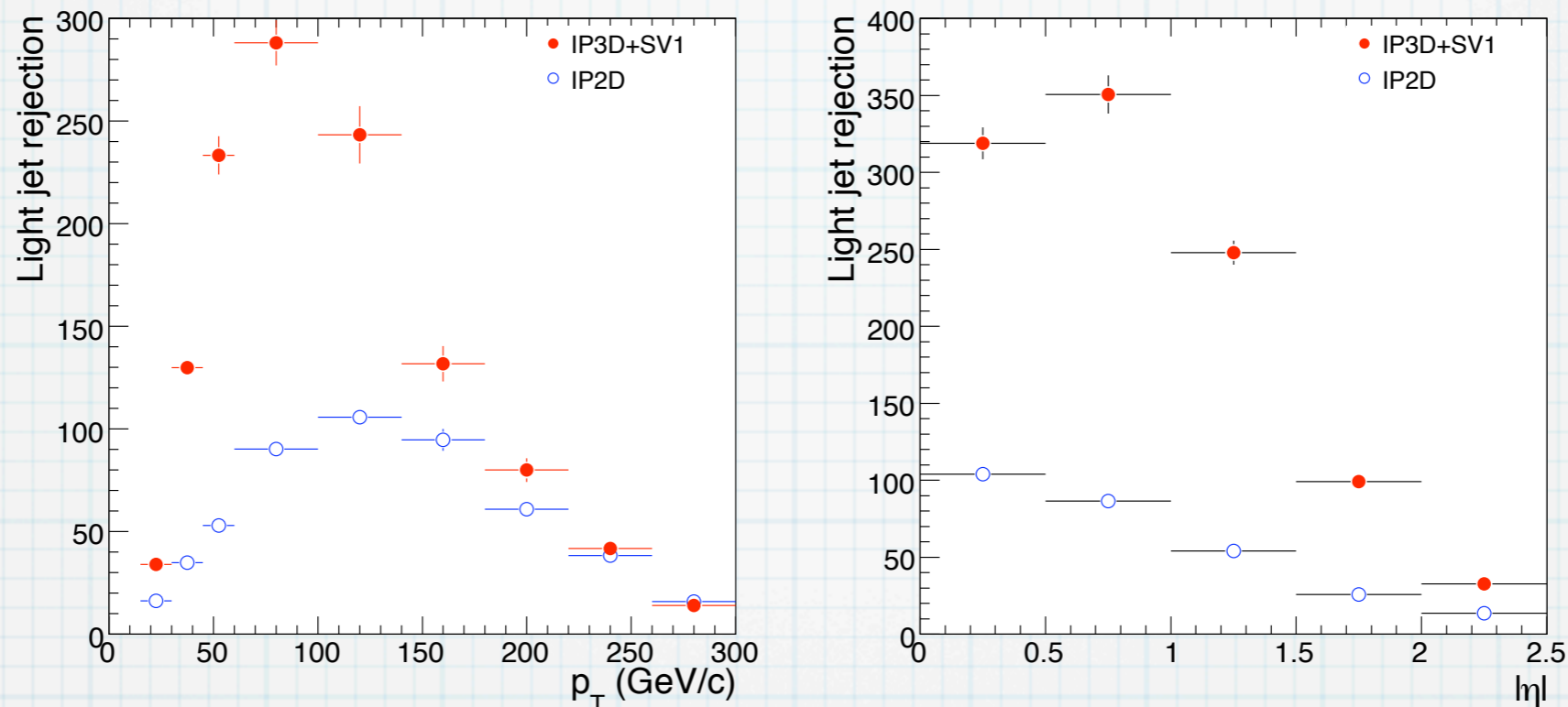


(b) WH events ($m_H=400$ GeV)

ハードなprocess

デフォルト

Jet p_T and η dependence



- ◆ p_T と η 分布自体に相関があることに注意
 - ❖ 本当は $\eta - p_T$ 平面上で見るべき
- ◆ Low p_T , high η での degradation
 - ❖ 粒子の多重散乱 ← 物質量が多い (後述)
- ◆ High p_T での degradation
 - ❖ 粒子がよりcollimateされて粒子密度が上がる
 - ❖ そもそもfragmentationからの粒子が増える
 - ❖ ビームパイプの外まで粒子が飛んでしまう

Summary of Performance at ATLAS

WH ($M_H=120\text{GeV}$)

	JetProb	IP2D	IP3D	IP3D+SV1	JetFitterCOMB
Purified light jet rejection, 50%	83 ± 1	116 ± 2	190 ± 3	457 ± 13	554 ± 17
Purified light jet rejection, 60%	30 ± 0	42 ± 0	59 ± 1	117 ± 2	133 ± 2

WH ($M_H=400\text{GeV}$)

	JetProb	IP2D	IP3D	IP3D+SV1	JetFitterCOMB
Purified light jet rejection, 50%	71 ± 1	163 ± 3	179 ± 3	297 ± 7	395 ± 11
Purified light jet rejection, 60%	27 ± 0	56 ± 1	58 ± 1	95 ± 1	123 ± 2
Purified c-jet rejection, 50%	7.9 ± 0.1	9.7 ± 0.1	10.7 ± 0.2	12.5 ± 0.2	12.7 ± 0.2
Purified c-jet rejection, 60%	4.7 ± 0.0	5.7 ± 0.1	6.1 ± 0.1	6.8 ± 0.1	7.3 ± 0.1

tt and ttj

	JetProb	IP2D	IP3D	IP3D+SV1	JetFitterCOMB
Raw light jet rejection, 50%	92 ± 0	142 ± 1	219 ± 1	423 ± 4	593 ± 6
Purified light jet rejection, 50%	99 ± 0	181 ± 1	293 ± 2	732 ± 10	863 ± 12
Raw light jet rejection, 60%	31 ± 0	49 ± 0	67 ± 0	144 ± 1	180 ± 1
Purified light jet rejection, 60%	33 ± 0	56 ± 0	76 ± 0	194 ± 1	213 ± 2

- ◆ 数字を現段階で真剣に受け取らないで下さい
 - ❖ performanceを落とす様々な効果は次頁以降

実験家が気になること

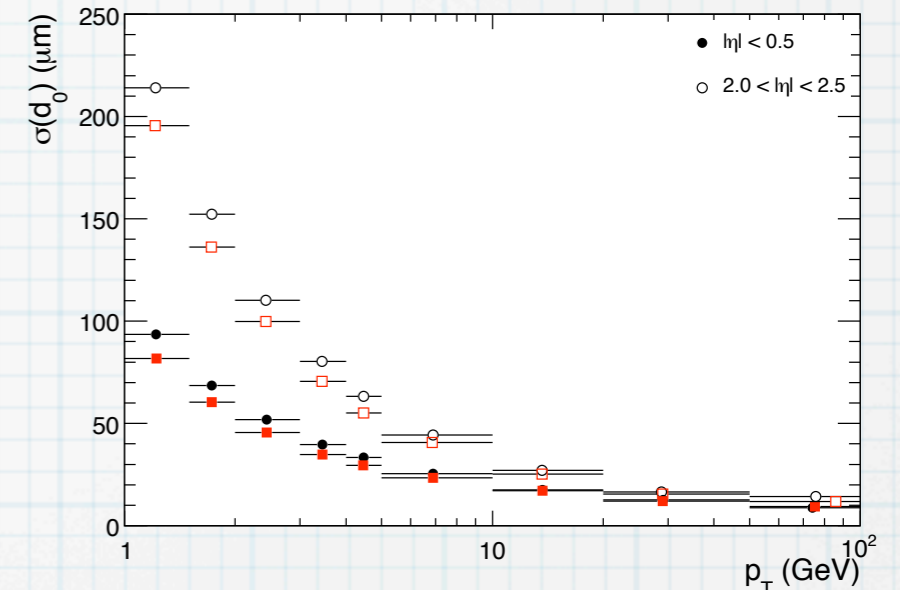
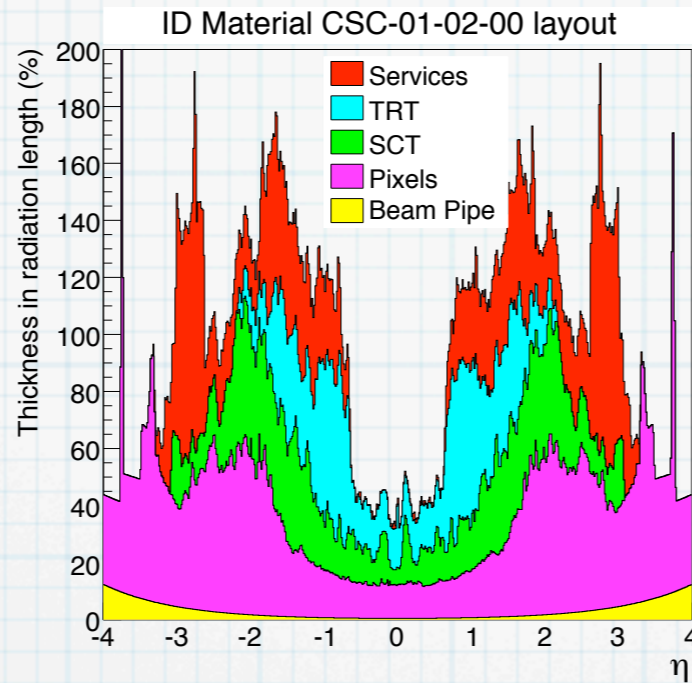
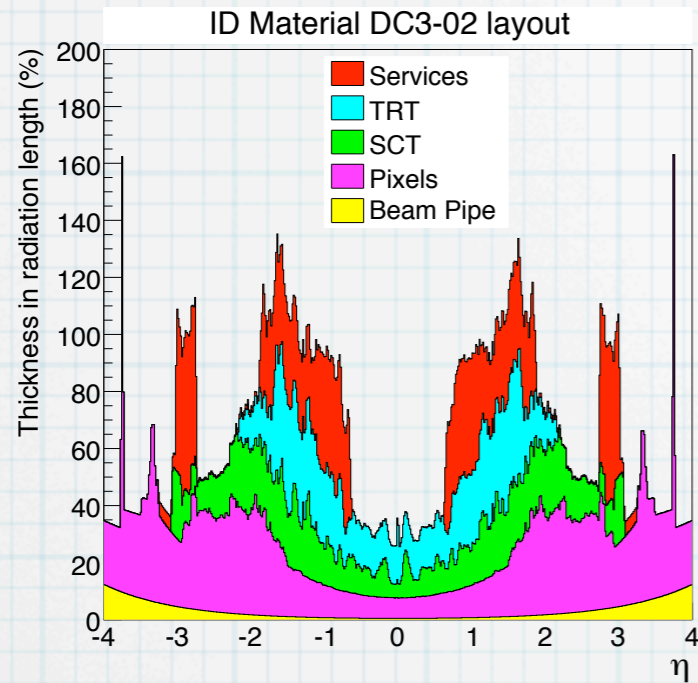
検出器の Performance

- ◆ Baselineの解析ではPixelの5%が死んでると仮定
 - ❖ 1%、あるいは3%のdeadだと…
 - ◆ tracking efficiency は最大2.5%程度向上

	IP2D	IP3D	IP3D+SV1
$\varepsilon_b = 60\%$			
Baseline rejection (5% of dead pixels)	48 ± 1	90 ± 2	201 ± 8
Relative change with 1% of dead pixels	1.08	1.09	1.20
Relative change with 3% of dead pixels	1.00	1.00	1.09

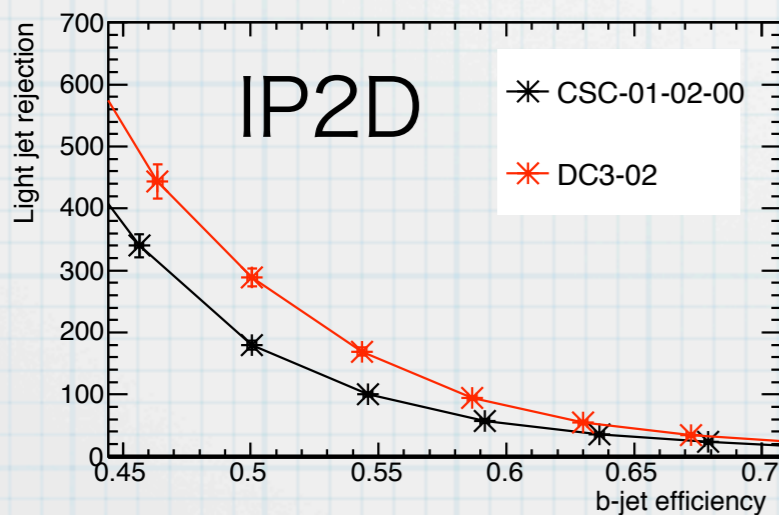
- ◆ Secondary vertex reconstruction のほうが sensitive
 - ❖ もっと死んでるときの study はされていない (?)

物質量

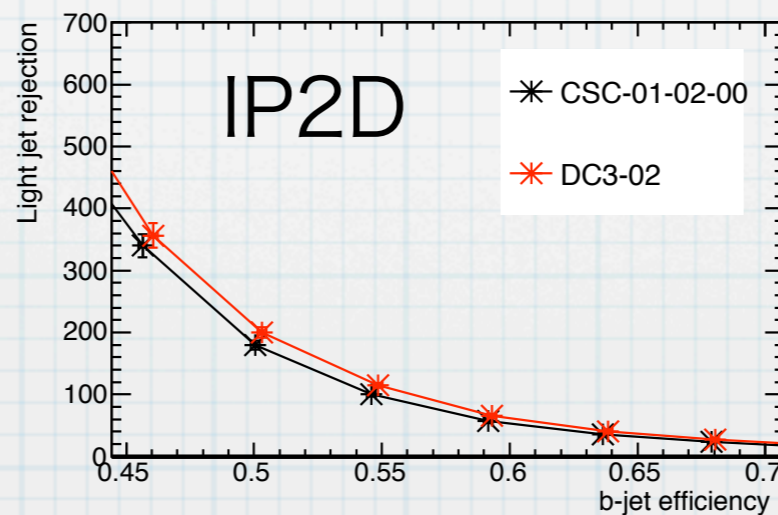


- ◆ Degradation comes from
 - ❖ worse resolution
 - ❖ more secondary particles

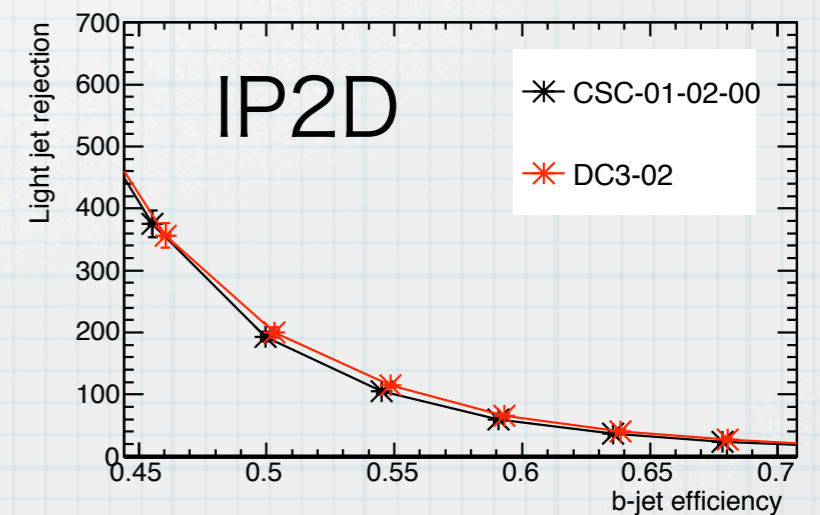
Performanceはこの geometryを使って評価



(a) Baseline



(b) DC3 impact parameter rescaled



(c) DC3 impact parameter rescaled and CSC extra secondaries removed

Misalignment of Pixel detector

random 10

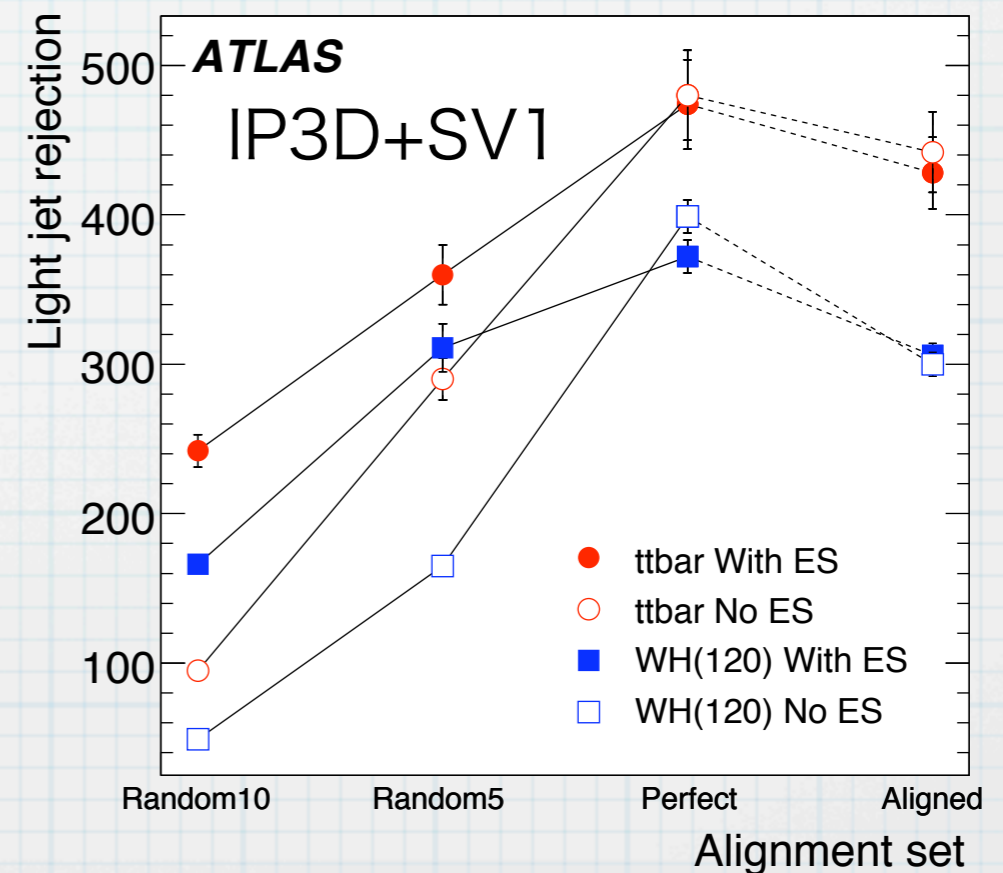
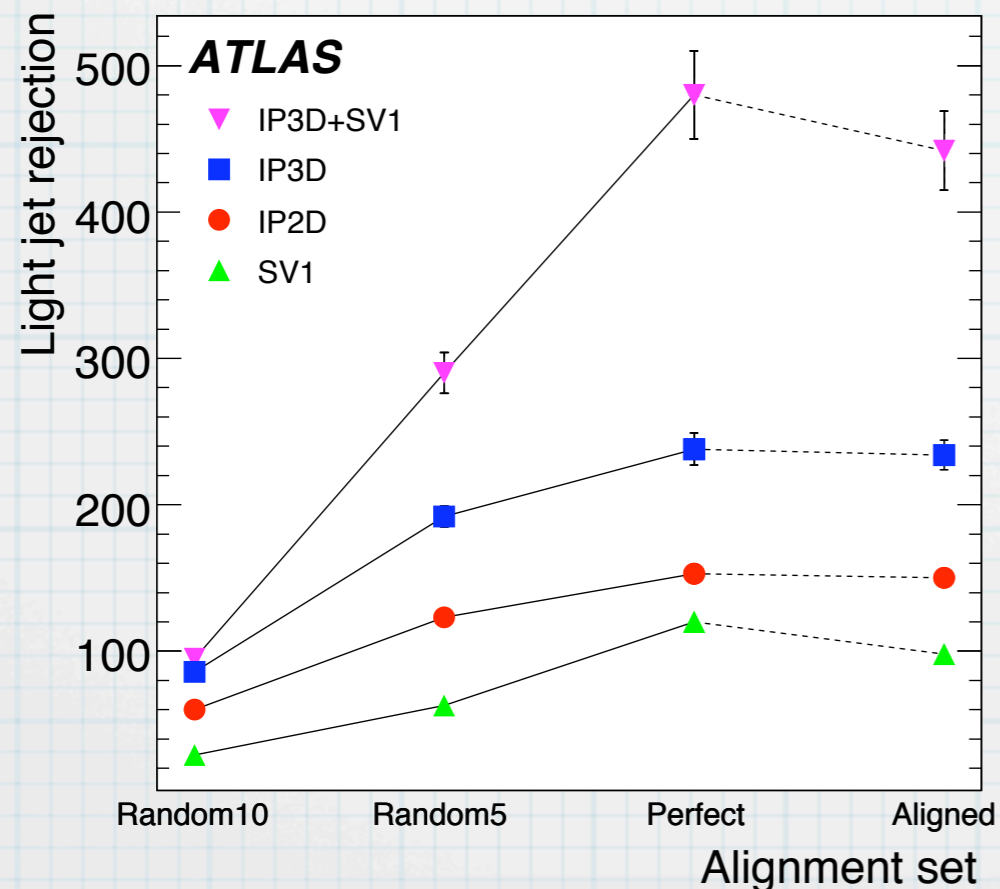
Level	x	y	z	RotX	RotY	RotZ
Module	10	30	30	0.3	0.5	0.2
layer	10	10	15	0.05	0.05	0.1
disk	10	10	30	0.2	0.2	0.1
Whole Pixel	10	10	15	0.1	0.1	0.1

random 5

Level	x	y	z	RotX	RotY	RotZ
Module	5	15	15	0.15	0.3	0.1
layer	7	7	10	0.02	0.02	0.05
disk	7	7	20	0.1	0.1	0.05
Whole Pixel	7	7	10	0.05	0.05	0.05

$$\sigma'^2 = a^2 \cdot \sigma^2 + c^2$$

	All <i>a</i>	Perfect <i>c</i> (μm)	Random10 <i>c</i> (μm)	Random5 <i>c</i> (μm)	Aligned <i>c</i> (μm)
Pixel Barrel Phi	1.03	0	31	13	3
Pixel Barrel Eta	0.97	0	71	34	13
Pixel Endcap Phi	1.05	0	30	14	3
Pixel Endcap Eta	1.08	0	43	11	15
SCT Barrel	0.78	0	0	2	7
SCT Endcap	0.86	0	6	5	8
TRT Barrel	0.82	0	11	3	37
TRT Endcap	0.77	0	11	10	19

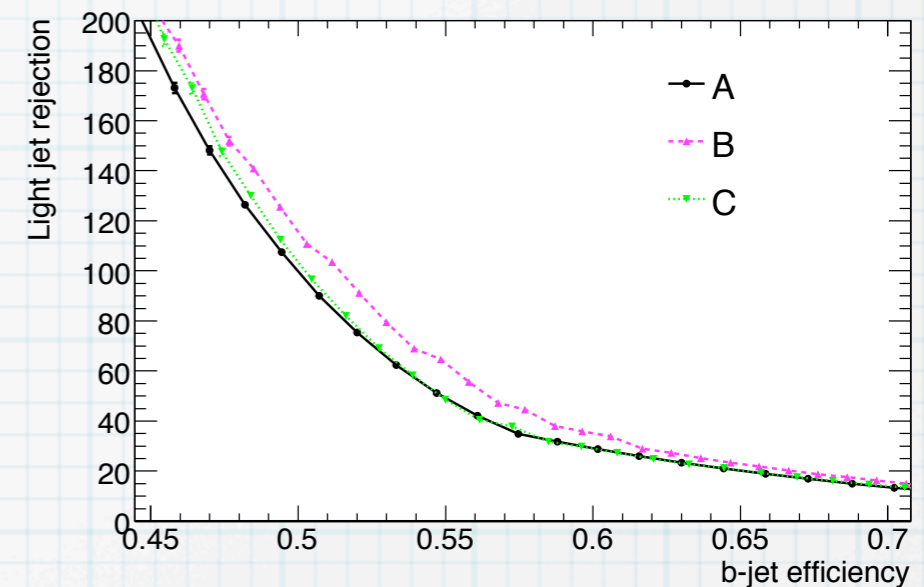


Calibration

◆ Jet Probability

A: ジェットに含まれる全てのトラックを使ってcalibration

B: 真にIPから来てるトラックのみを使った場合



❖ トラックのcategory分けがされていない

◆ $p(\sin\theta)^{2/3}$, pixel(SCT)の数、などで分けるべき

❖ σ_{d_0} のチューニング ← d_0 の分布自体でやれる

◆ Likelihood ratio

❖ t-tbar を使う (?)

◆ efficiencyの測定は出来るかもしれないが、multi-dimensional likelihood に耐え得る統計を集めるのは難しい

Efficiency と Fake Rejection の 測定

◆ Efficiency

- ❖ t-tbar で b-tag を使わずに b-jetサンプル
- ❖ dijet事象を用いて Sytem 8 と呼ばれる方法

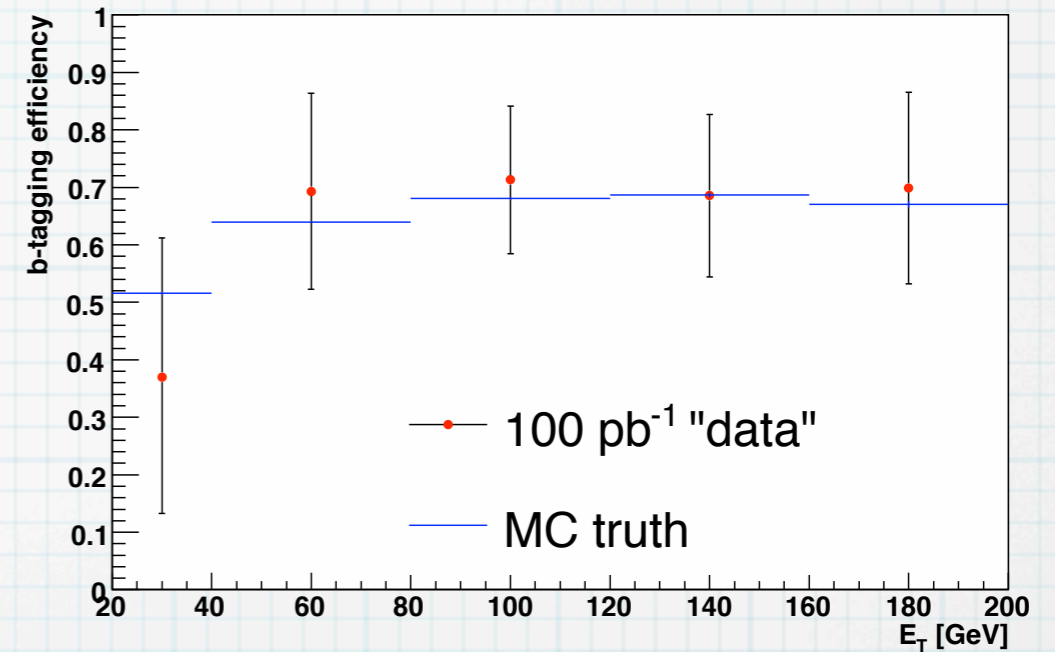
◆ Fake Rejection Factor

- ❖ S_{IP} のマイナスの分布 (=検出器のresolutionによって生じる) を使って評価

◆ Sample dependence をなくすためには、上記2つを以下のn次元空間で測定する必要がある

- ❖ jet p_T , η ← 最低この2つ
- ❖ jet multiplicity, luminosity
- ❖ などなど

from t-tbar events

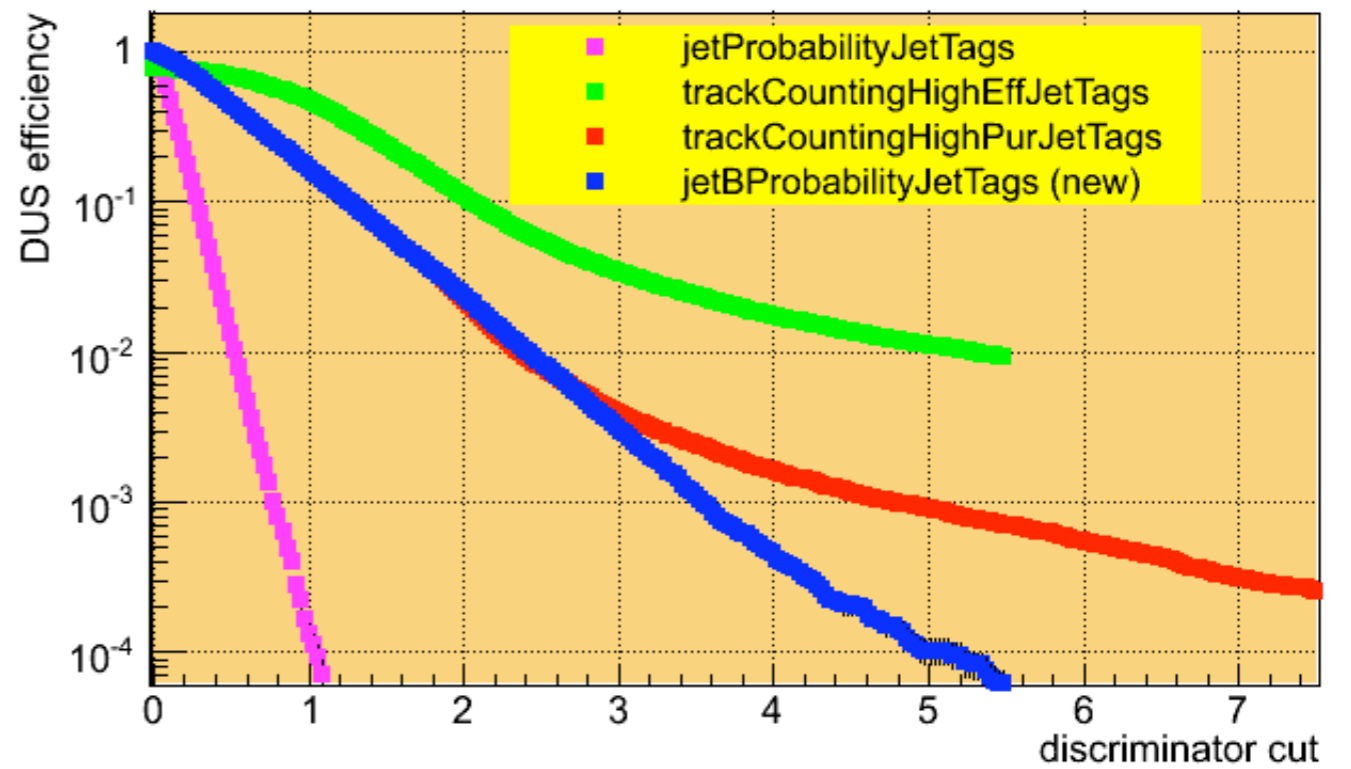
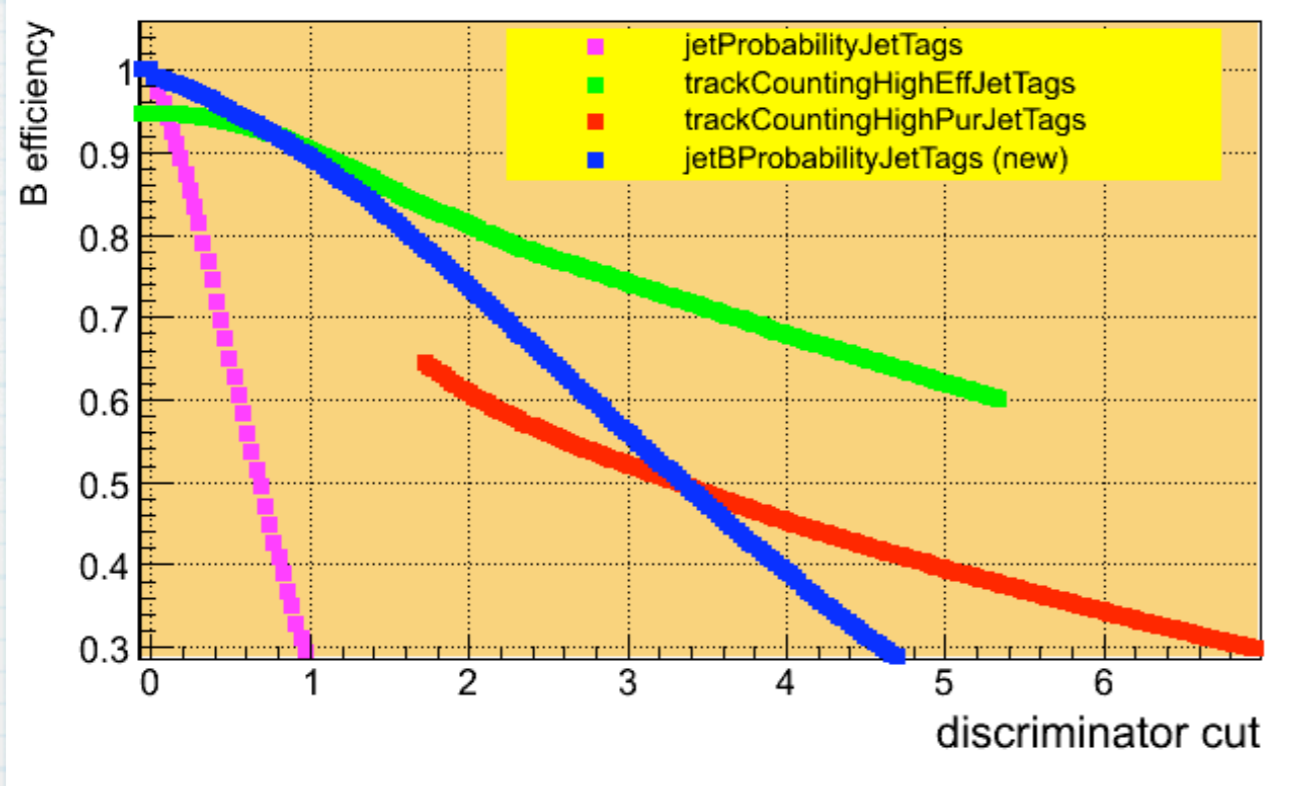


Conclusions

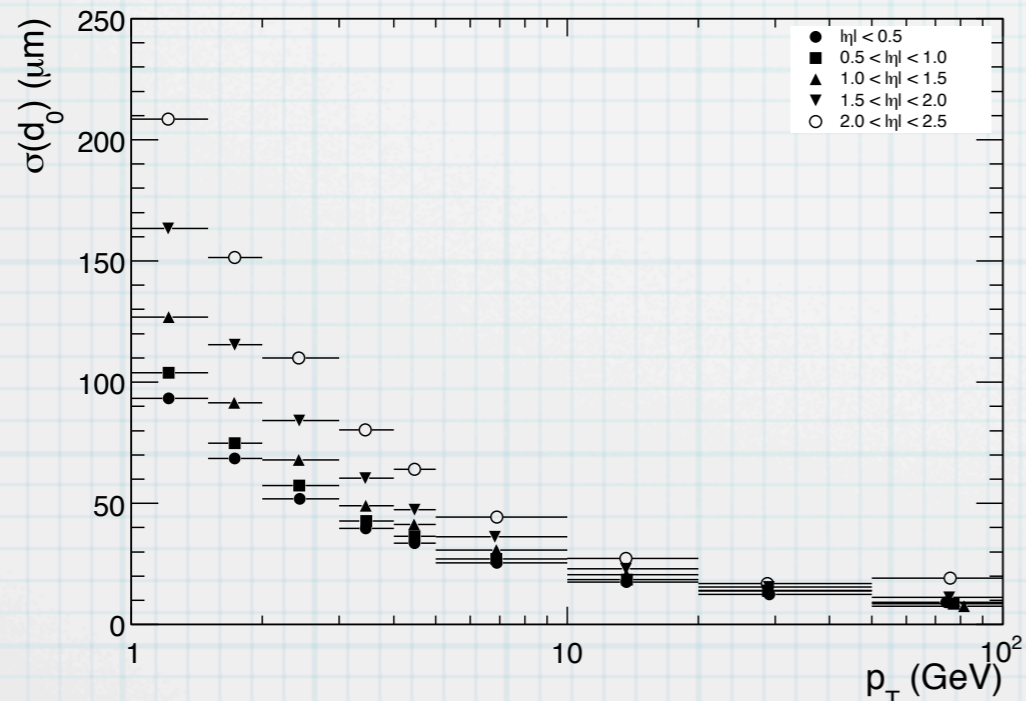
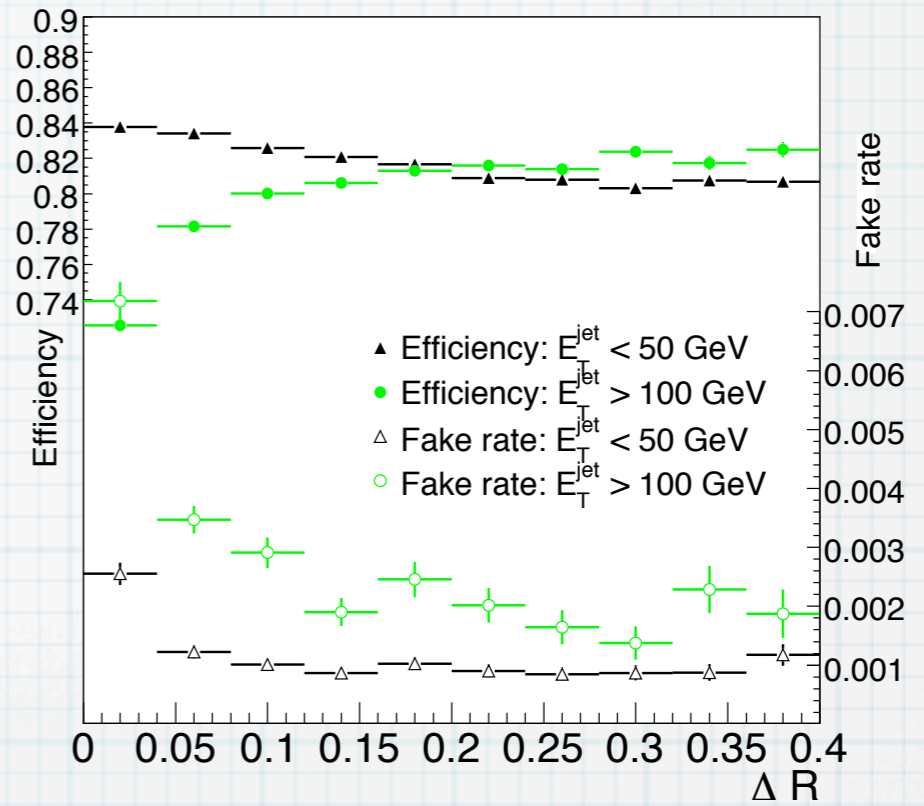
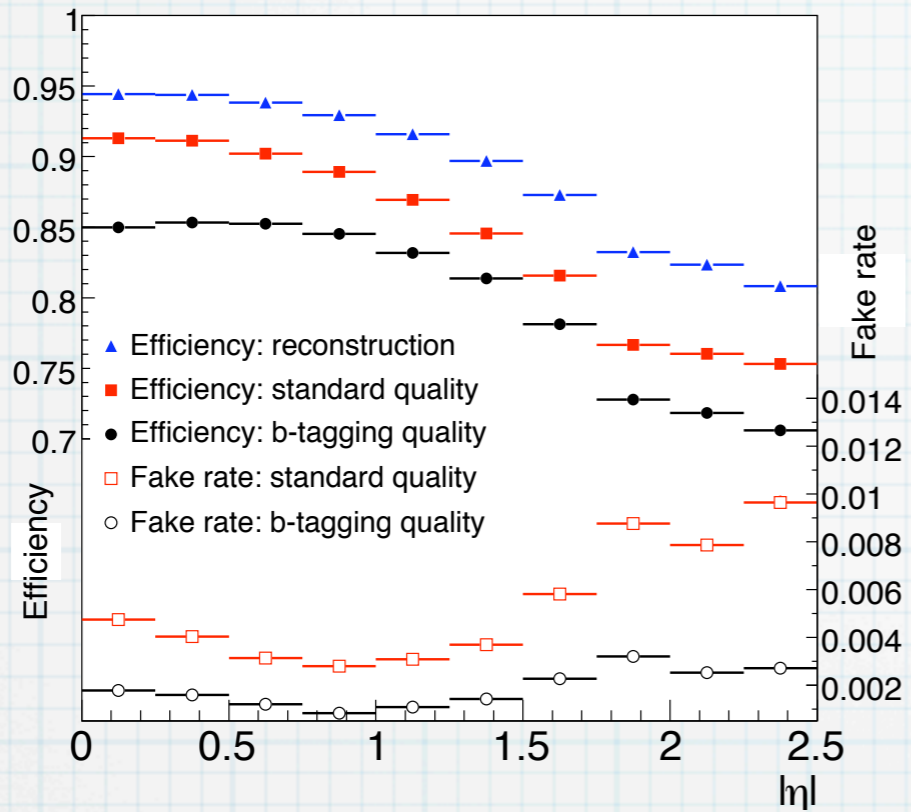
- ◆ ATLAS (LHC) で使われる b-jet tagging の方法について紹介した
- ◆ シミュレーションから期待されるPerformanceは様々な効果によってdegradeする可能性
 - ❖ Calibrationするのが難しい場合は n次元空間で efficiency と Fake Rejection を測定すべし
- ◆ 実験初期にreliableなのは Jet Probability
 - ❖ 豊富な calibration sample
- ◆ もっときめ細かなcalibrationが不可欠
 - ❖ Sample dependence
 - ❖ 少なくともトラックのqualityに応じたcalibration
←劇的な改善が期待できる (現在進行中)

Backup

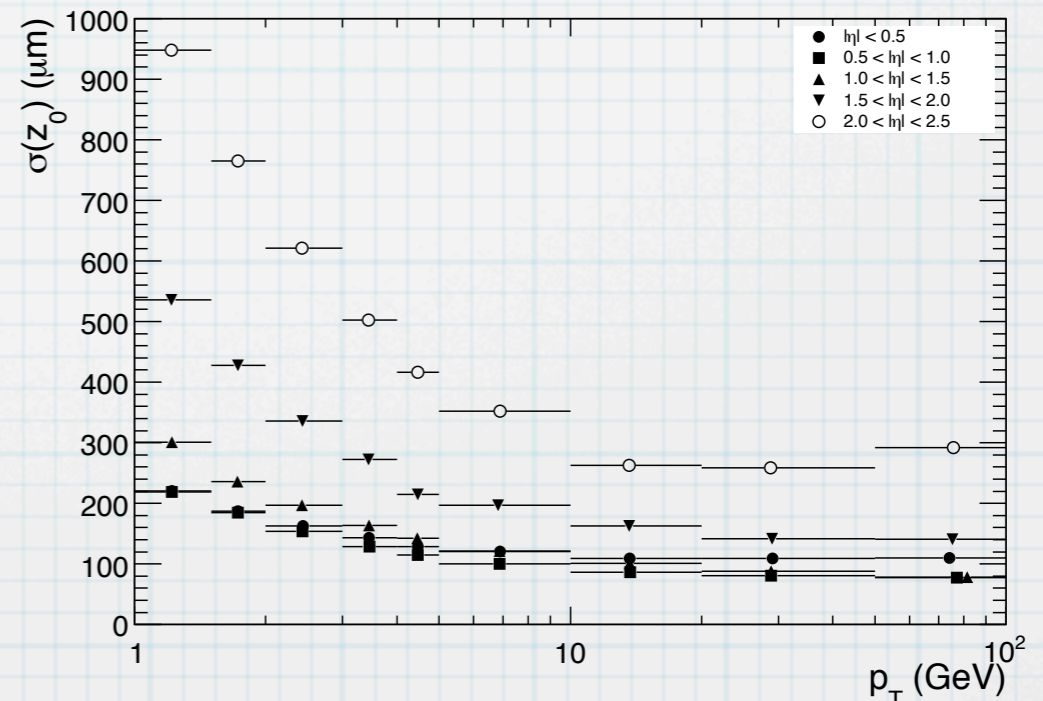
CMS



Tracking Performance



(a) Transverse impact parameter



(b) Longitudinal impact parameter

Effect of misalignment

Table 4: Efficiency of track reconstruction (statistical errors only)

Setup	efficiency [%]
Perfect	97.1 ± 0.2
Random10	95.3 ± 0.2
Random10 + error scaling	97.1 ± 0.2
Aligned	96.9 ± 0.2
Aligned + error scaling	96.9 ± 0.2

Table 5: Ratio of fake tracks from the total accepted tracks (statistical errors only)

Setup	fake tracks [%]
Perfect	2.33 ± 0.02
Random10	2.46 ± 0.02
Random10 + error scaling	2.29 ± 0.02
Aligned	2.34 ± 0.02
Aligned + error scaling	2.20 ± 0.02

Table 6: Ratio of primary vertex outliers (statistical errors only)

Setup	outliers in x [%]	outliers in z [%]
Perfect	1.7 ± 0.1	4.1 ± 0.1
Random10	5.5 ± 0.2	8.3 ± 0.2
Random10 + error scaling	2.8 ± 0.1	6.1 ± 0.2
Aligned	3.2 ± 0.1	8.2 ± 0.2
Aligned + error scaling	3.2 ± 0.1	8.0 ± 0.2

Table 7: Primary vertex resolution (statistical errors only)

Setup	res. in x [μm]	shift in x [μm]	res. in z [μm]	shift in z [μm]
Perfect	11.4 ± 0.1	-0.13 ± 0.07	51.1 ± 0.4	-8.2 ± 0.3
Random10	15.1 ± 0.1	4.2 ± 0.1	63.0 ± 0.4	1.4 ± 0.4
Random10 + error scaling	13.2 ± 0.1	2.6 ± 0.1	56.6 ± 0.4	2.3 ± 0.4
Aligned	13.9 ± 0.1	-0.18 ± 0.09	53.7 ± 0.4	-91.5 ± 0.4
Aligned + error scaling	13.8 ± 0.1	-0.15 ± 0.09	55.4 ± 0.4	-91.6 ± 0.4

System 8

$$n = n_b + n_{cl}$$

$$p = p_b + p_{cl}$$

$$n^{LT} = n_b \varepsilon_b^{LT} + n_{cl} \varepsilon_{cl}^{LT}$$

$$p^{LT} = p_b \varepsilon_b^{LT} + p_{cl} \varepsilon_{cl}^{LT}$$

$$n^{SMT} = n_b \varepsilon_b^{SMT} + n_{cl} \varepsilon_{cl}^{SMT}$$

$$p^{SMT} = p_b \varepsilon_b^{SMT} + p_{cl} \varepsilon_{cl}^{SMT}$$

$$n^{both} = n_b \varepsilon_b^{LT} \varepsilon_b^{SMT} + n_{cl} \varepsilon_{cl}^{LT} \varepsilon_{cl}^{SMT}$$

$$p^{both} = p_b \varepsilon_b^{LT} \varepsilon_b^{SMT} + p_{cl} \varepsilon_{cl}^{LT} \varepsilon_{cl}^{SMT}$$