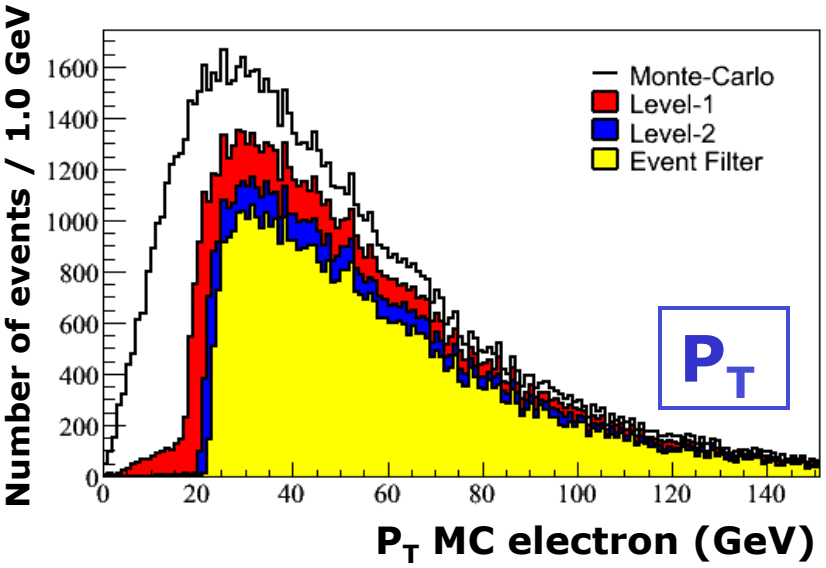
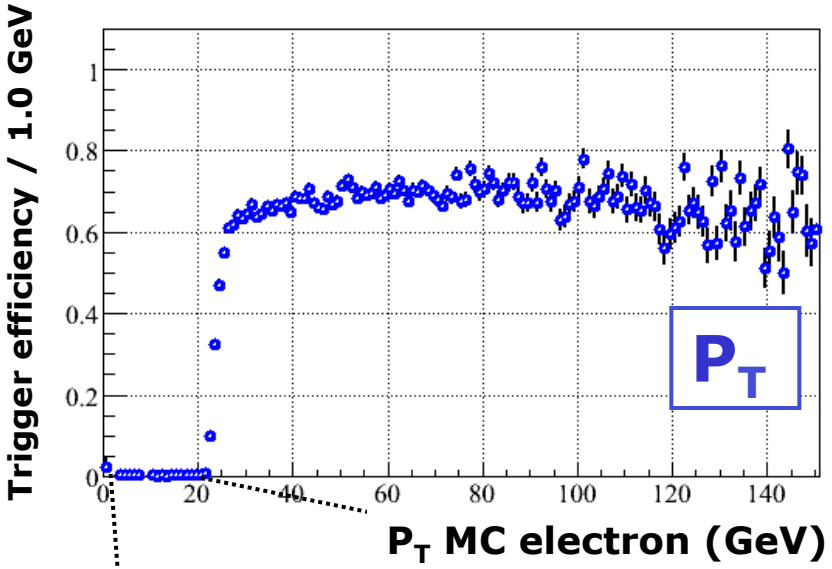
Binomial Error intervals
&
Extra leptons in ttbar events

Ivo van Vulpen
(Nikhef)

Binomial Error intervals

Trigger efficiencies: Example: EM25i in tt(electron) events



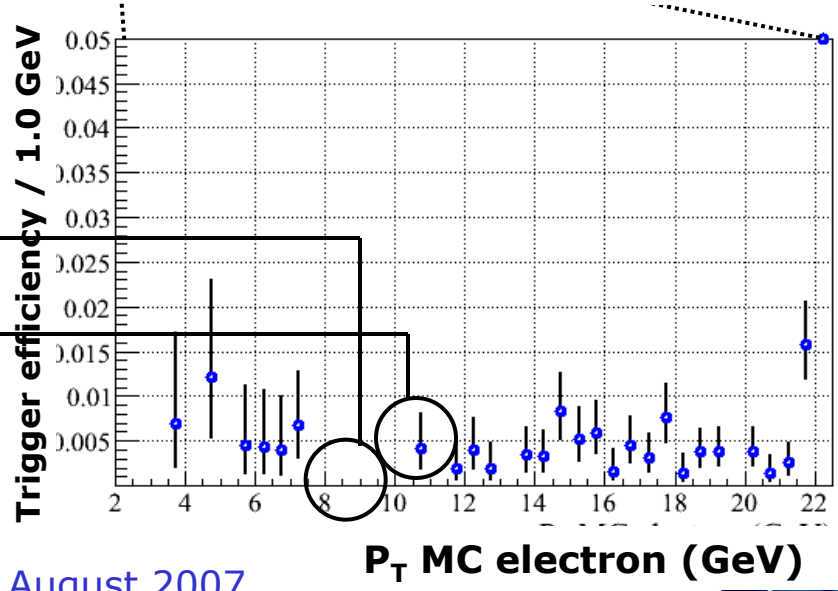


Issues when making the plot:

A How to treat weighted events ?
- $N_{tot}<0$, $N_{sel}<0$ or $N_{sel}>N_{tot}$

B Correct Binomial errors
i.e. not simply sqrt(ε(1-ε)/N)

→ Discuss in (T5) Trigger meeting



Binomial distribution

$$f(N_{sel} \mid N_{tot}, p) = \frac{N_{tot}!}{N_{sel}!(N_{tot} - N_{sel})!} p^{N_{sel}} (1-p)^{N_{tot} - N_{sel}}$$

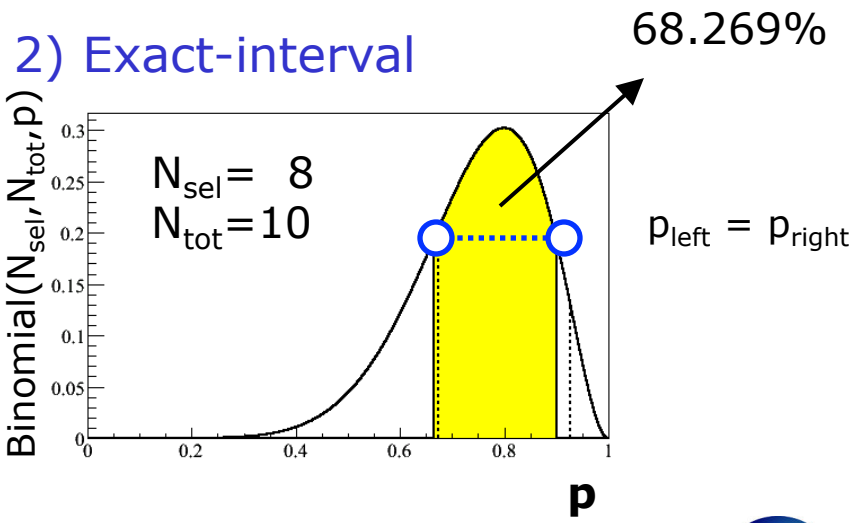
- Select N_{sel} events out of a sample of N_{tot} events
  Throw a dice 10x → probability to find 4x1 = $(^{10}_{4})(0.16)^4(0.84)^6$ =4.8%

- Given N_{tot} and N_{sel} → information on p (efficiency) and it's uncertainty

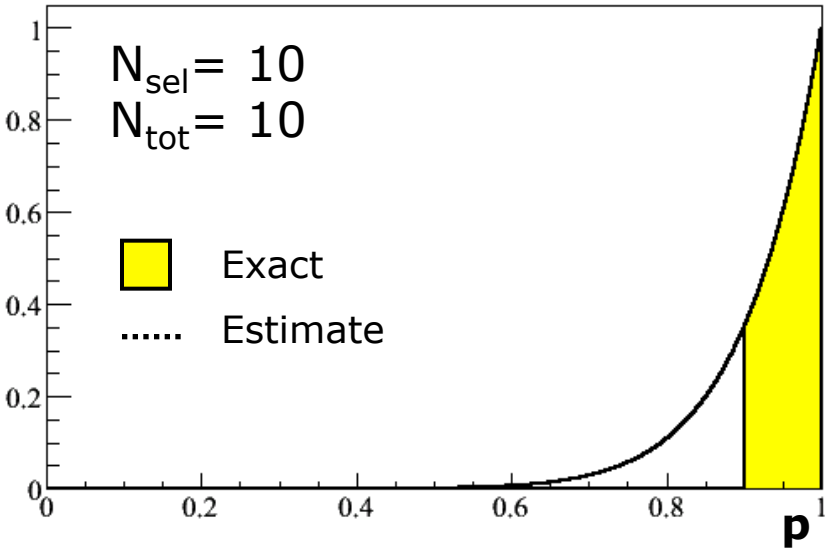
1) Estimated-interval

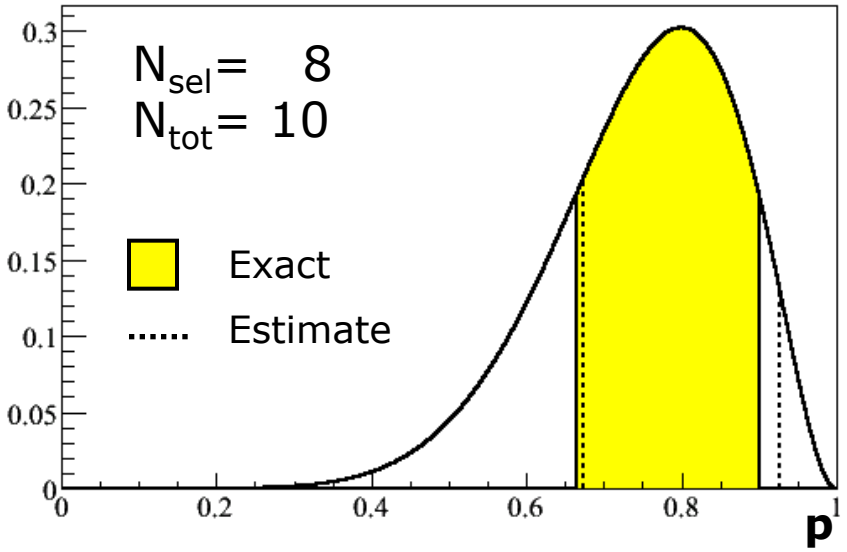
$$\Delta p = \sqrt{\frac{p(1-p)}{N_{tot}}}$$

2) Exact-interval

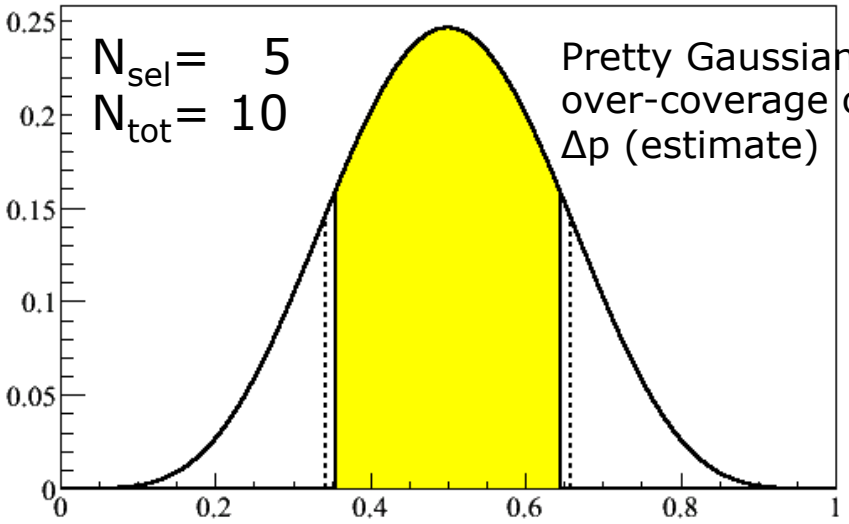


Difference in uncertainty region [estimate/exact]



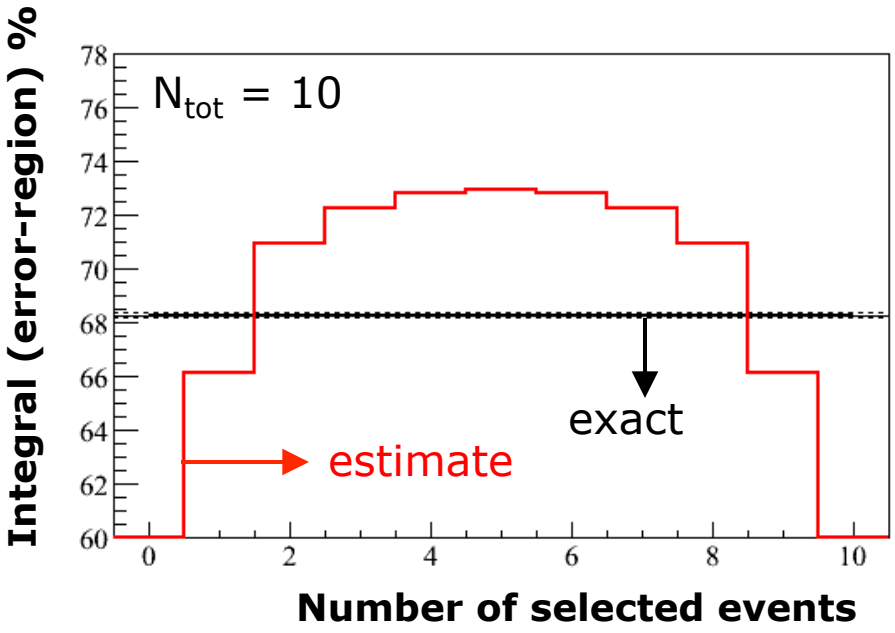


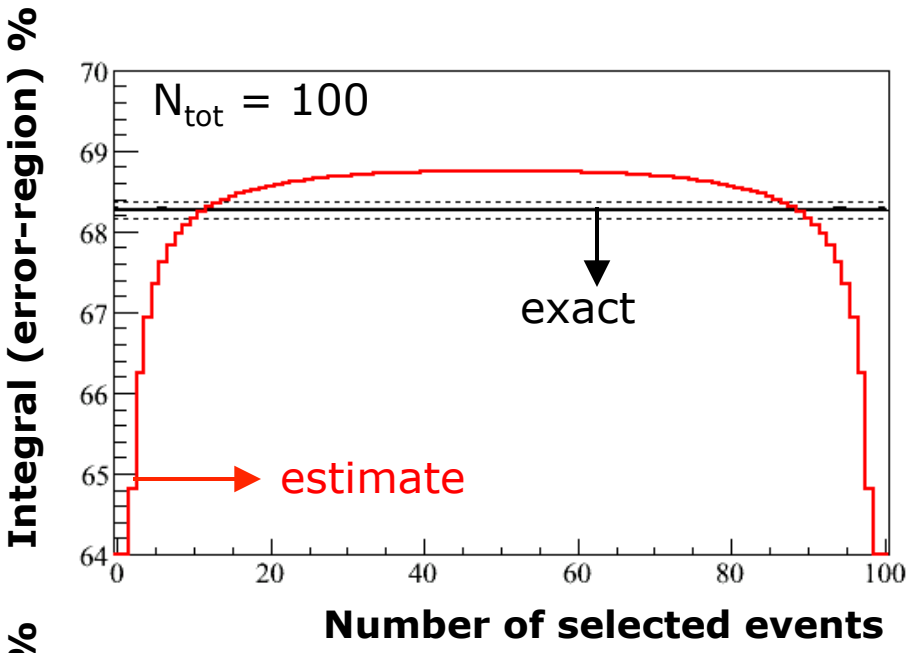
N_{sel},N_{tot}	Δp(estimate)	Δp(exact)
10,10	± 0.000	+0.000 −0.100
8,10	± 0.126	+0.101 −0.136
5,10	± 0.158	+0.145 −0.144



Pretty Gaussian, but over-coverage of Δp (estimate)

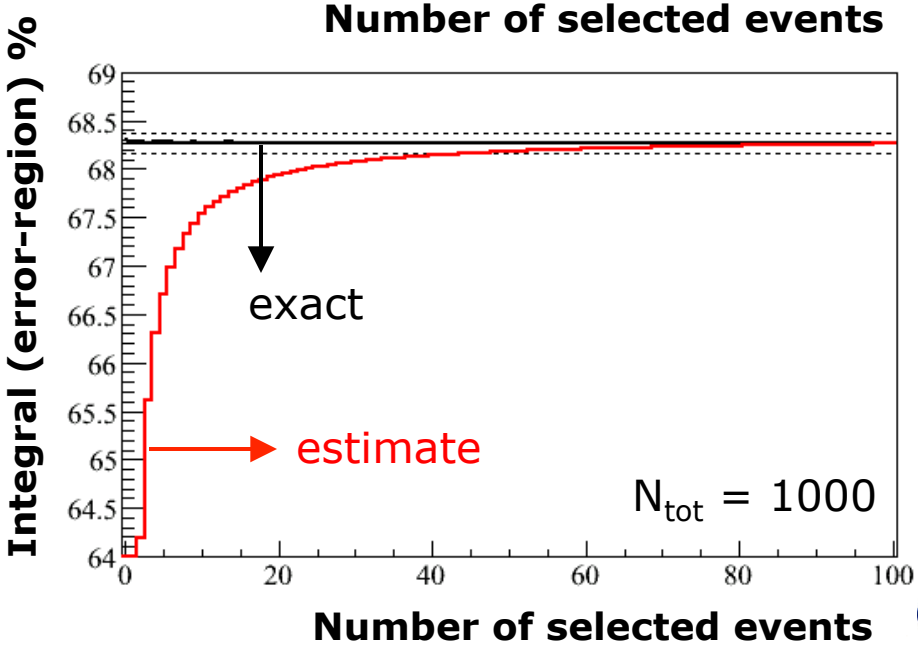
Difference in coverage [estimate/exact]





Coverage: integral Δp region

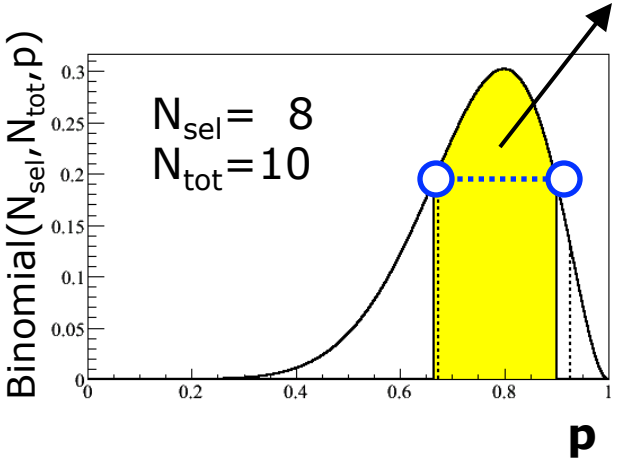
- Under/over-coverage when using estimated-interval



68.269% B

Computing exact interval

Step left/right from p_{max} until:
a) integral is gaussian one sigma equiv.
b) $p_{left} = p_{right}$



At each step need:

Probability: $$f(N_{sel} \mid N_{tot}, p) = \frac{N_{tot}!}{N_{sel}!(N_{tot} - N_{sel})!} p^{N_{sel}} (1-p)^{N_{tot} - N_{sel}}$$

→ In root `TMath::Binomial`

Integral: $$\int_0^x x^k (1-x)^{N-k} dx = B(p, k+1, N-k+1)$$

→ Incomplete Beta function
In root `TMath::BetaIncomplete`

Numerical problems when computing exact interval:

$$f(N_{sel} \mid N_{tot}, p) = \frac{N_{tot}!}{N_{sel}!(N_{tot} - N_{sel})!} p^{N_{sel}} (1-p)^{N_{tot} - N_{sel}}$$

N_{sel}=200, N_{tot}=1000 at p=0.5: $6.6\ 10^{+215}$ x $6.2\ 10^{-61}$ x $1.5\ 10^{-241}$

Solution: just remove Binomial coefficient and use logarithms.

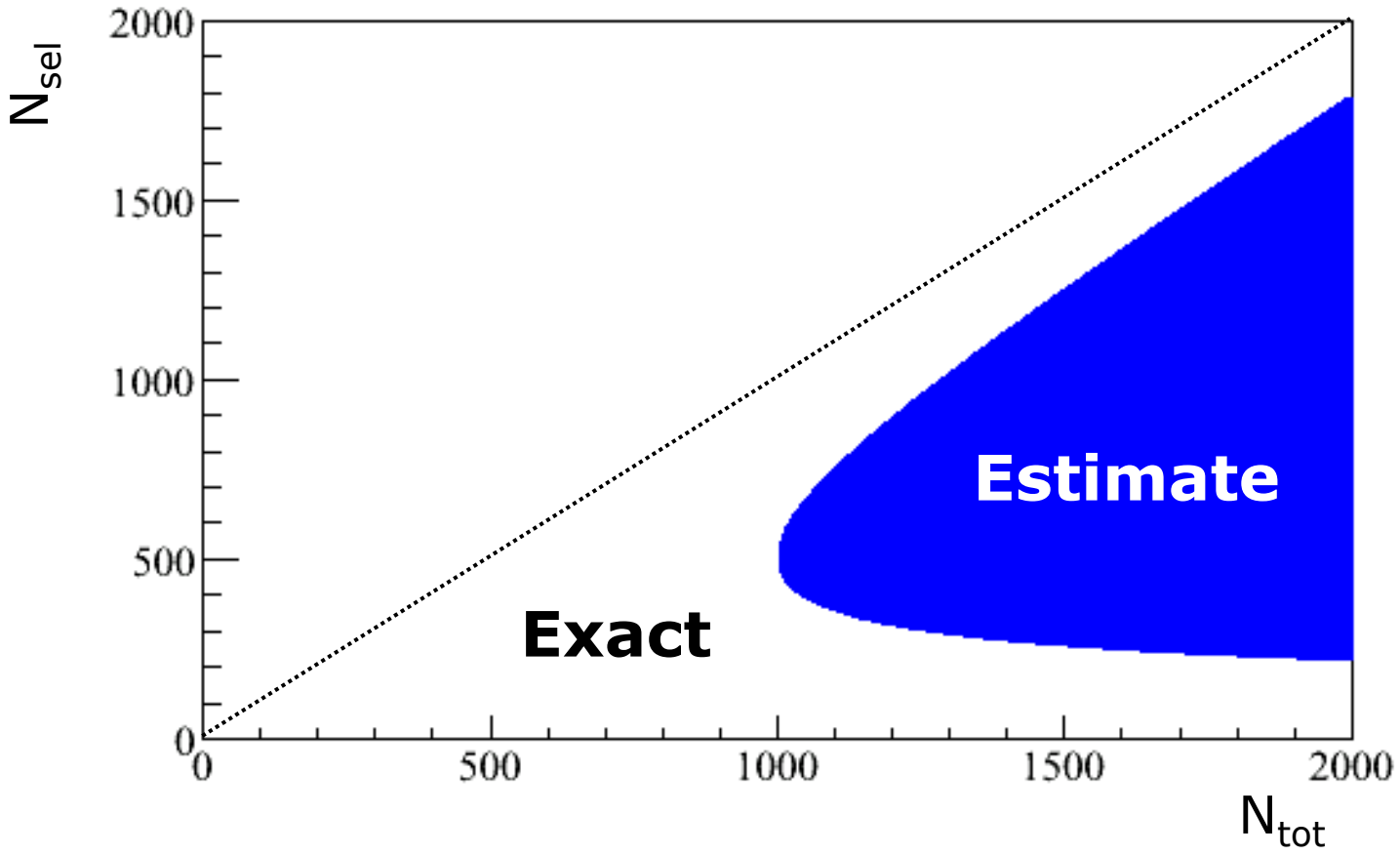
$$\int_0^x x^k (1-x)^{N-k} dx = B(p, k+1, N-k+1)$$

$$\text{Beta}(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = e^{\ln(\Gamma(x)) + \ln(\Gamma(y)) - \ln(\Gamma(x+y))}$$

= TMath::BetaIncomplete(p,k+1,N-k+1)
= something / TMath::Beta(k+1,N-k+1)
= something / TMath::Exp(TMath::LnGamma(p)+TMath::LnGamma(q)-TMath::LnGamma(p+q))
Problem if argument < -700 since 10-300 = 0 → Integral = inf or NaN

Solution: switch to estimate if this happens

Where do you use estimate:



In principle could do exact in larger region by using Stirling's approximation for Beta(x,y) and work with logarithms.

Existing code in Root

→ Exact Binomial interval can already be computed in Root:

`TGraphAsymmErrors::BayesDivide(const TH1* pass, const TH1* total)`
(code by Andy Haas and Marc Paterno)

`TMath::BinomialI(n.p.k)` → should be updated
(coded by Anna Kreshuk)

Contains already the computation of the correct interval

→ Short routine presented in previous slides:

http://www.nikhef.nl/~ivov/BinomialErrorInterval/**BinomialErrorInterval.C**

Please check/verify:
BinomialErrorInterval(Int_t N_{sel}, Int_t N_{tot} , Double_t *ErrorRegion)

To get printout: Example(N_{tot},N_{sel})

Weighted events

- MC@NLO weighted events +1 and -1 (~13.4% has w=-1)
- What to do with bins where: $N_{tot} < 0$, $N_{sel} < 0$ or $N_{sel} > N_{tot}$?

Several private discussions, but no consensus, no clear strategy.

1) Use weight = 1 in all plots
   + 'An electron is an electron'
   - Trigger could depend on weight (event topology, phase space, …)

2) Combine the two trigger curves for weight=+1 and weight=-1
   Not sure how-to, analysis dependent ?

What do you do with weighted events ?

Should have discussion in ATLAS (T5-group) and arrive at common approach to create turn-on curves.

Extra leptons in ttbar events

- **Number of selected events for 100 pb^{-1} (no trigger)**

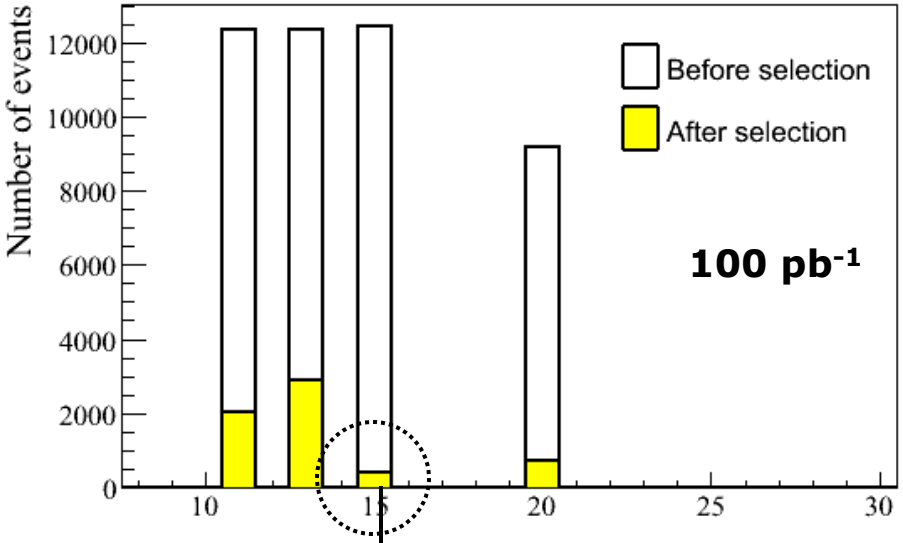
	Before selection	After selection	Selection efficiency	Mass window 141<Mt<189 +M_W-cut
TTbar(5200)	**46277.8**	**6172.0**	**13.3%**	**1420.9**
electron	12329.2	2168.6	17.6%	527.3
muon	12350.2	2871.4	23.2%	719.9
tau	12415.0	427.1	3.4%	105.2
fully-leptonic	9181.7	704.8	7.7%	68.7
W+jets	**79266.7**	**1601.0**	**2.0%**	**129.8**

→ Similar numbers to Udine-group/Akira

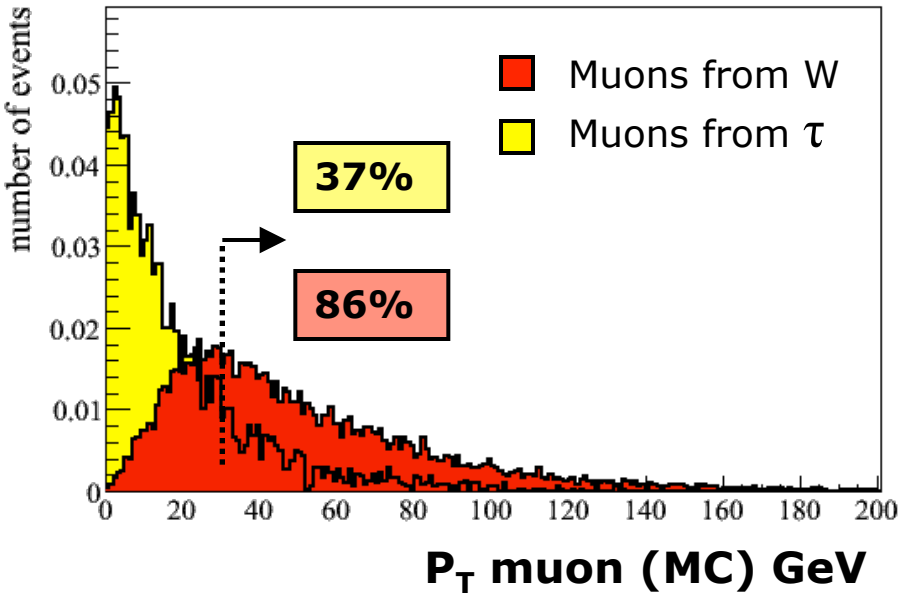
Issues: Electron definition (isolation/medium/cracks)

Leptons from top events: example: muons from tau's

11/13/15 = semi-leptonic tt (elec/muon/tau)
20 = fully-leptonic tt



Alexander Doxiadis



Efficiency for tau's (in a muon analysis):
~ ε(muon) * 0.17 * (0.37/0.86) * ..

QCD background in commissioning analysis

→ QCD events passing event selection:

- Jet requirements — No problem
- E_T-miss > 20 GeV — ~ 9% — (rel. 11, FastSim estim.)
- Single isolated lepton — ? — (electron or muon)

→ Origin extra leptons:

Semi-leptonic B(D)-decays

Photon conversions

Non-prompt

π faking electron

Fake

→ Determine rate:

QCD FastSim (Release 11)

TTbar FullSim (next slide)

QCD fullsim ??

10^{-5} (egama) → 10^{-4} SUSY/TOP ?

QCD FastSim estimate:

+ Correct jet topologies

- How well do you model the relevant parameters ?

TTbar FullSim:

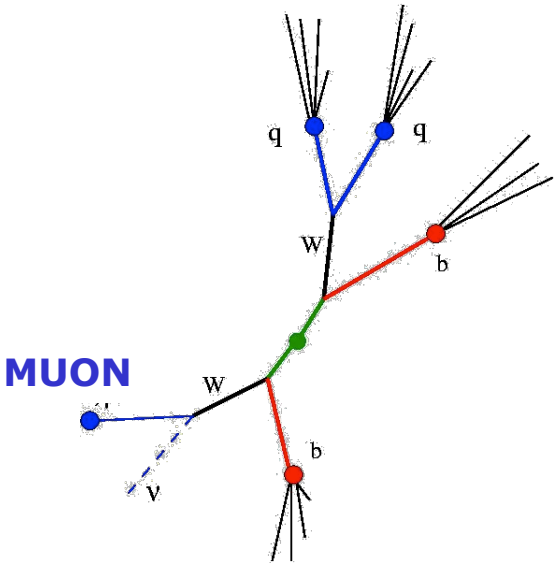
+ Best modeling of detector, photon conversions etc.

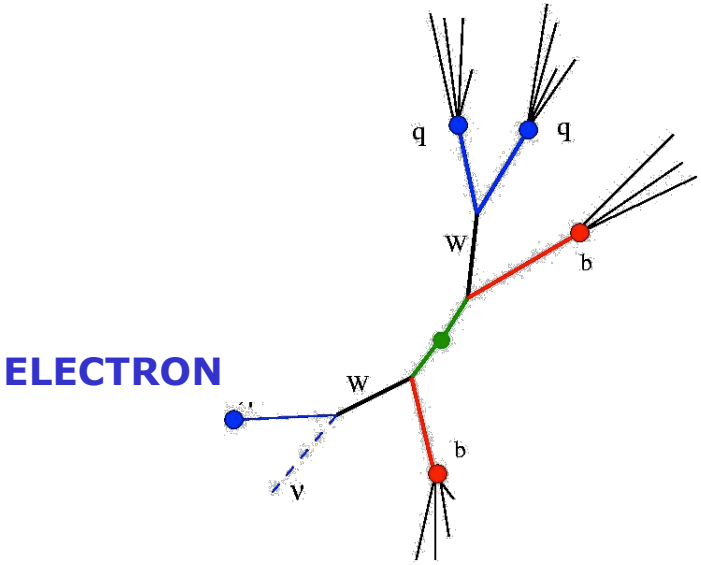
- Different jets:
  * B/Q-jets in ttbar, Q and gluon-jets in QCD
  * higher-Pt jets in ttbar
  * Different topologies

→ Can we translate it to QCD events ?

QCD Fullsim:

Not available, only SUSY sample with hard filter cuts
Issue: should we produce it ?





Do we see electrons ?

Do we see muons ?

Then what … → # 'wrong' leptons per jet
→ Find origin of these extra leptons
→ Get # QCD jets and compute QCD efficiency

TTbar sample (5200): no weights → ~520k events with ~2.1M jets (P_T>20 GeV)

- **Monte-Carlo statistics (no weights, *before* event selection)**

# tt(muon) events	=	139,105	# tt(electron) events	=	138,652
# jets (P_T> 20 GeV)	=	565,646	# jets (P_T> 20 GeV)	=	606,897
# tight electrons	=	343	# good muons	=	207
↓			↓		
'Extra' electrons/jet	=	60.6 *10^{-5}	'Extra' muons/jet	=	34.1 *10^{-5}

- **Expected # events (*after* event selection, *100 pb^{-1}*, no trigger) for the electron and muon analysis separately**

Analysis	tt(electron)	tt(muon)	tt(tau)	tt(leptonic)	TOTAL
1 electron	2163.1	2.5	177.4	299.7	2642.7
1 muon	5.4	2868.9	249.7	405.1	3529.2

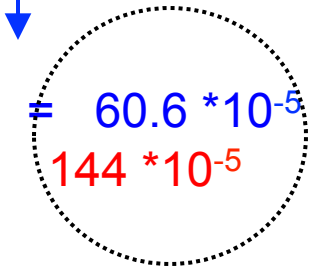
- **Monte-Carlo statistics (no weights, *before* event selection)**

tt(muon) events = 139,105
jets (P_T> 20 GeV) = 565,646

tt(electron) events = 138,652
jets (P_T> 20 GeV) = 606,897

tight electrons = 343
medium electrons = 814

good muons = 207



'Extra' electrons/jet = 60.6 *10^{-5}
144 *10^{-5}

'Extra' muons/jet = 34.1 *10^{-5}

TIGHT

- **Expected # events (*after* event selection, *100 pb^{-1}*, no trigger) for the electron and muon analysis separately**

Analysis	tt(electron)	tt(muon)	tt(tau)	tt(leptonic)	TOTAL
1 electron	2163.1	2.5	177.4	299.7	2642.7
1 muon	5.4	2868.9	249.7	405.1	3529.2

MEDIUM

- **Expected # events (*after* event selection, *100 pb^{-1}*, no trigger) for the electron and muon analysis separately**

Analysis	tt(electron)	tt(muon)	tt(tau)	tt(leptonic)	TOTAL
1 electron	2646.7	12.6	240.3	338.6	3238.2
1 muon	4.4	2860.2	249.1	361.6	3475.3

More electrons … but you loose some muons, probably because of Nleptons ==1 cut

Coming 2 weeks:

Manuel

1) Search for origin extra electrons: B(D) decays, photon conv., **fakes**

   Seach for origin extra muons: B-decays ?

Alexander

→ Study dependence on Pt (objects) and jet-type
   Disentangle parameters

→ Check FUllSim Ttbar estimate (Pt, eta) versus Egamma estimate
→ Redo our study with medium electrons
→ Run over di-jet sample from Egamma group

3) Study difference TTbar Fast/Full Simulation

   Some things CAN be modelled correctly others CAN NOT

4) Decide if we need separate FullSIm QCD sample ~ 300k ?

Martijn