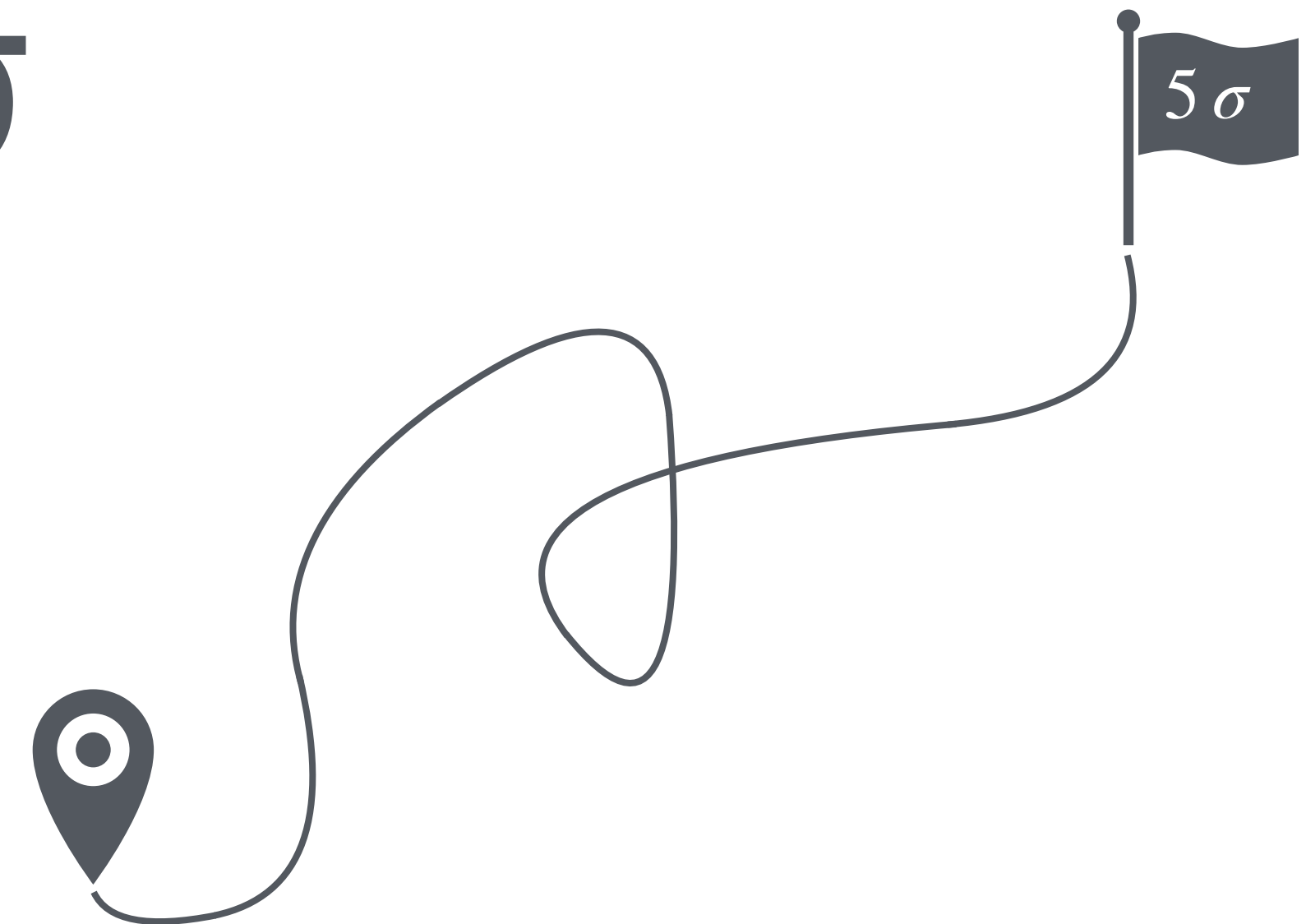


The journey to 5σ

Anke Ackermann
Kirchhoff Institute for Physics
Heidelberg University

Annweiler am Trifels, July 2024



How to make an observation



Pick an appropriate final state

Apply an event selection to maximise your S/B

Estimate the irreducible backgrounds

Check if you are above 5σ

Pick an appropriate final state

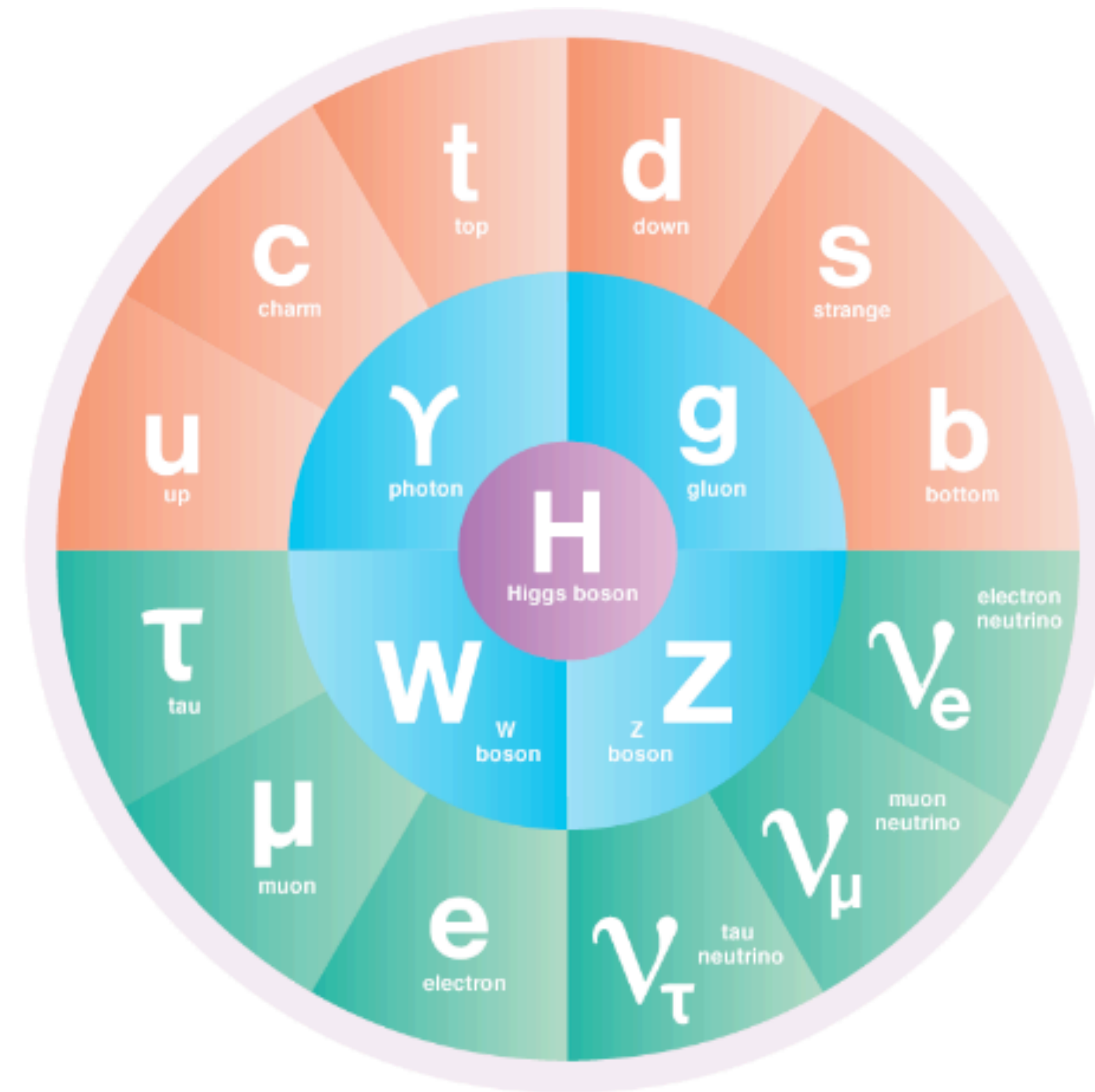


pp

q \bar{q}

gg

...



photons

electrons

muons

jets

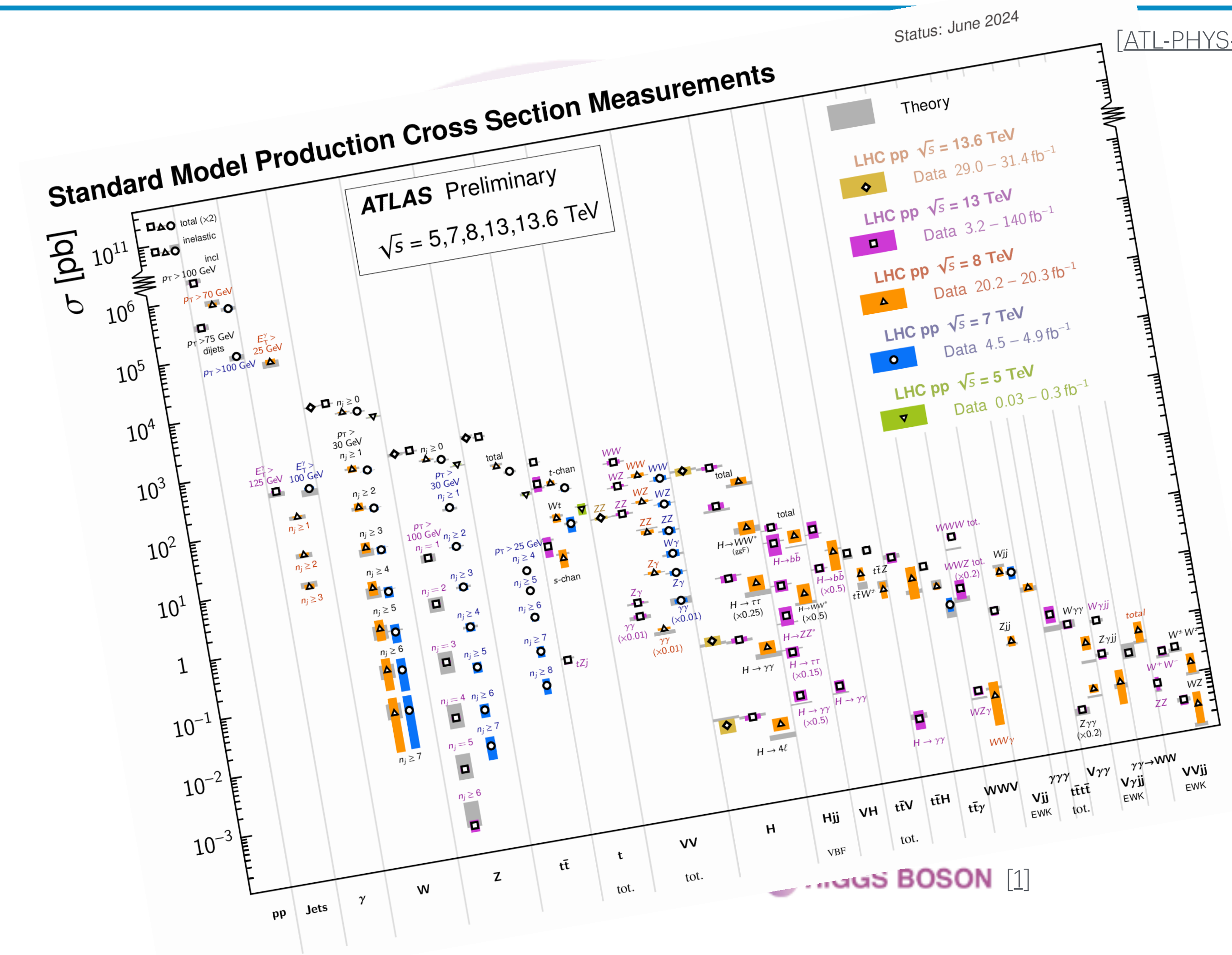
neutrinos/
MET

● QUARKS ● LEPTONS ● BOSONS ● HIGGS BOSON [1]

Pick an appropriate final state, but ...

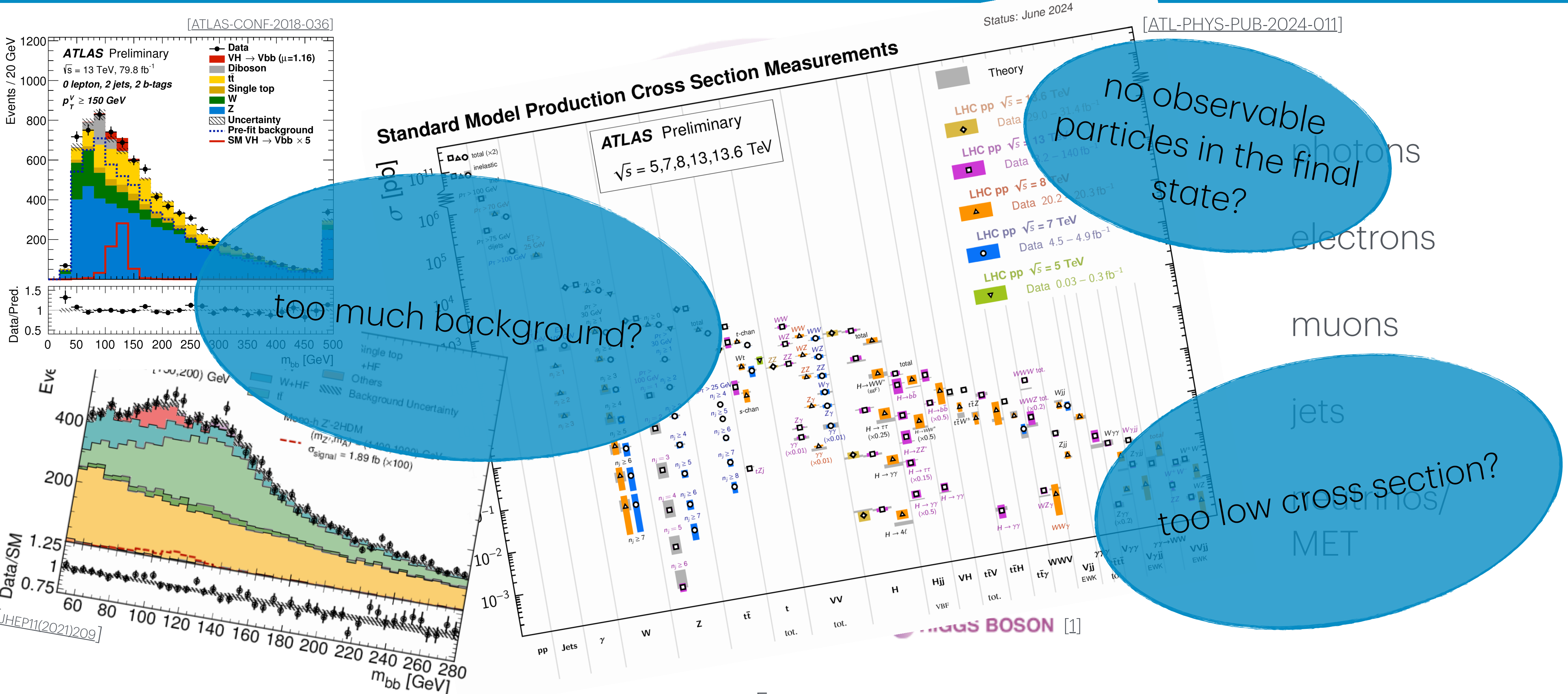


pp
q \bar{q}
gg
...



photons
electrons
muons
jets
neutrinos/
MET

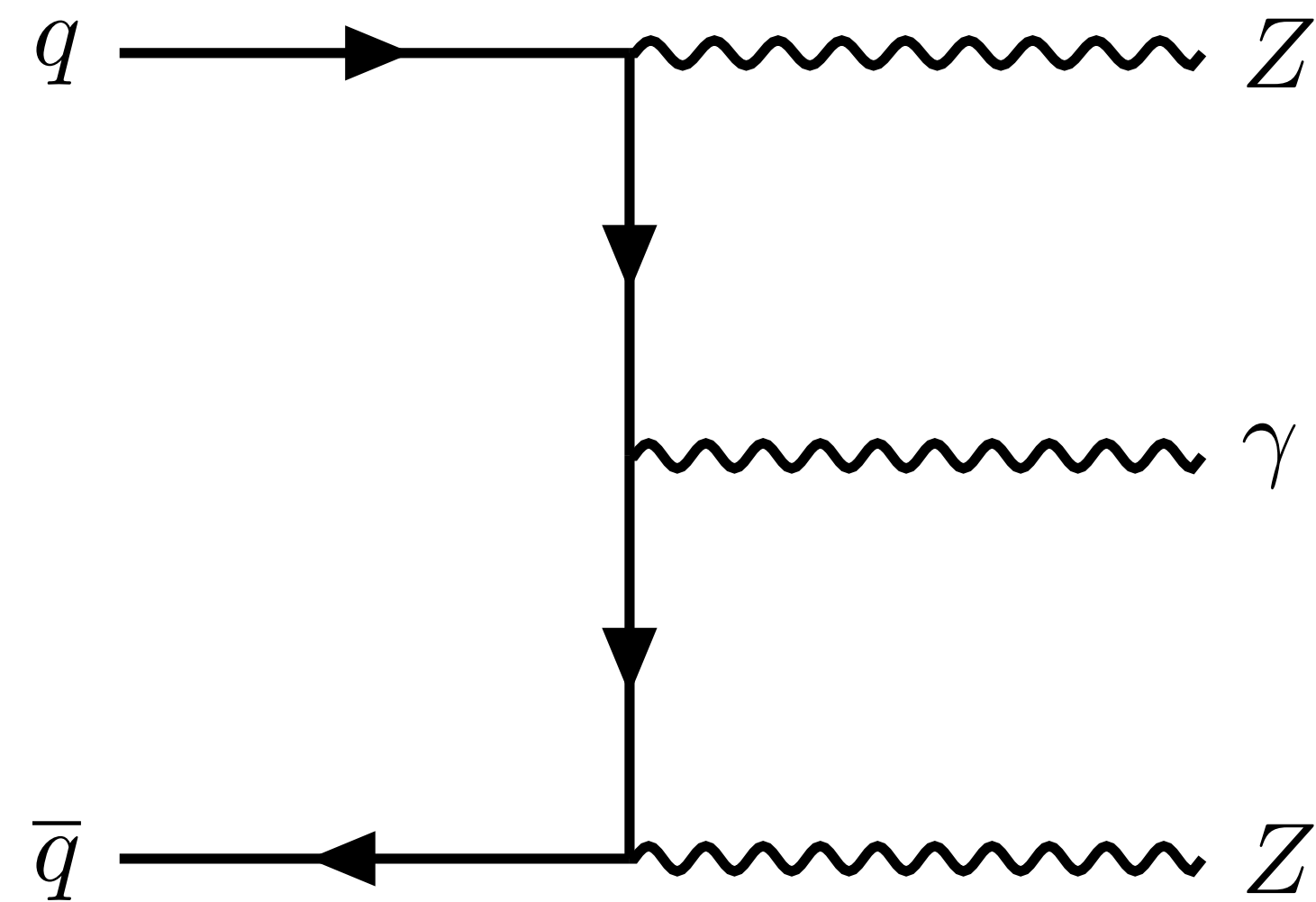
Pick an appropriate final state, but ...



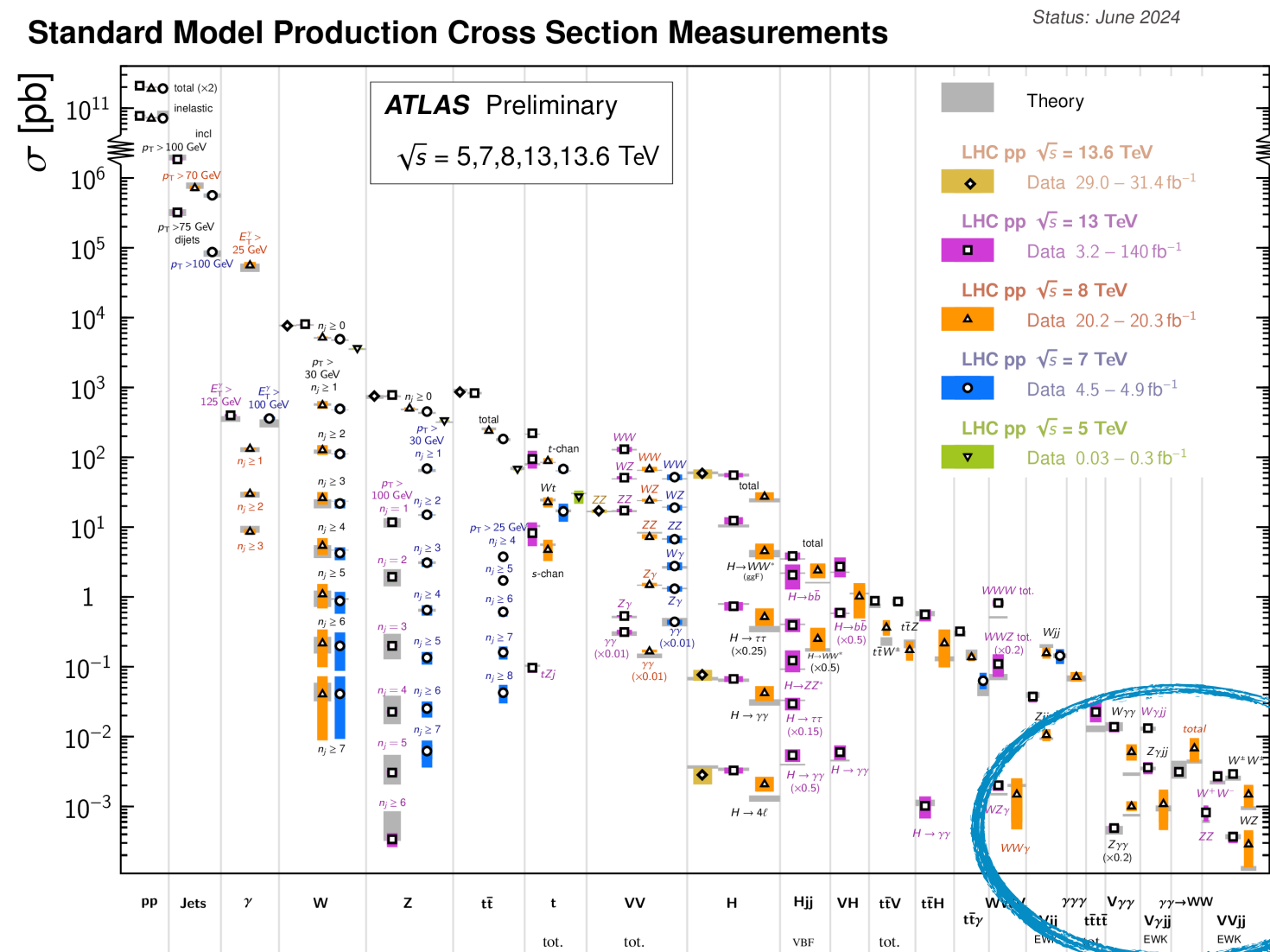
An appropriate final state



$$q\bar{q} \rightarrow ZZ\gamma$$

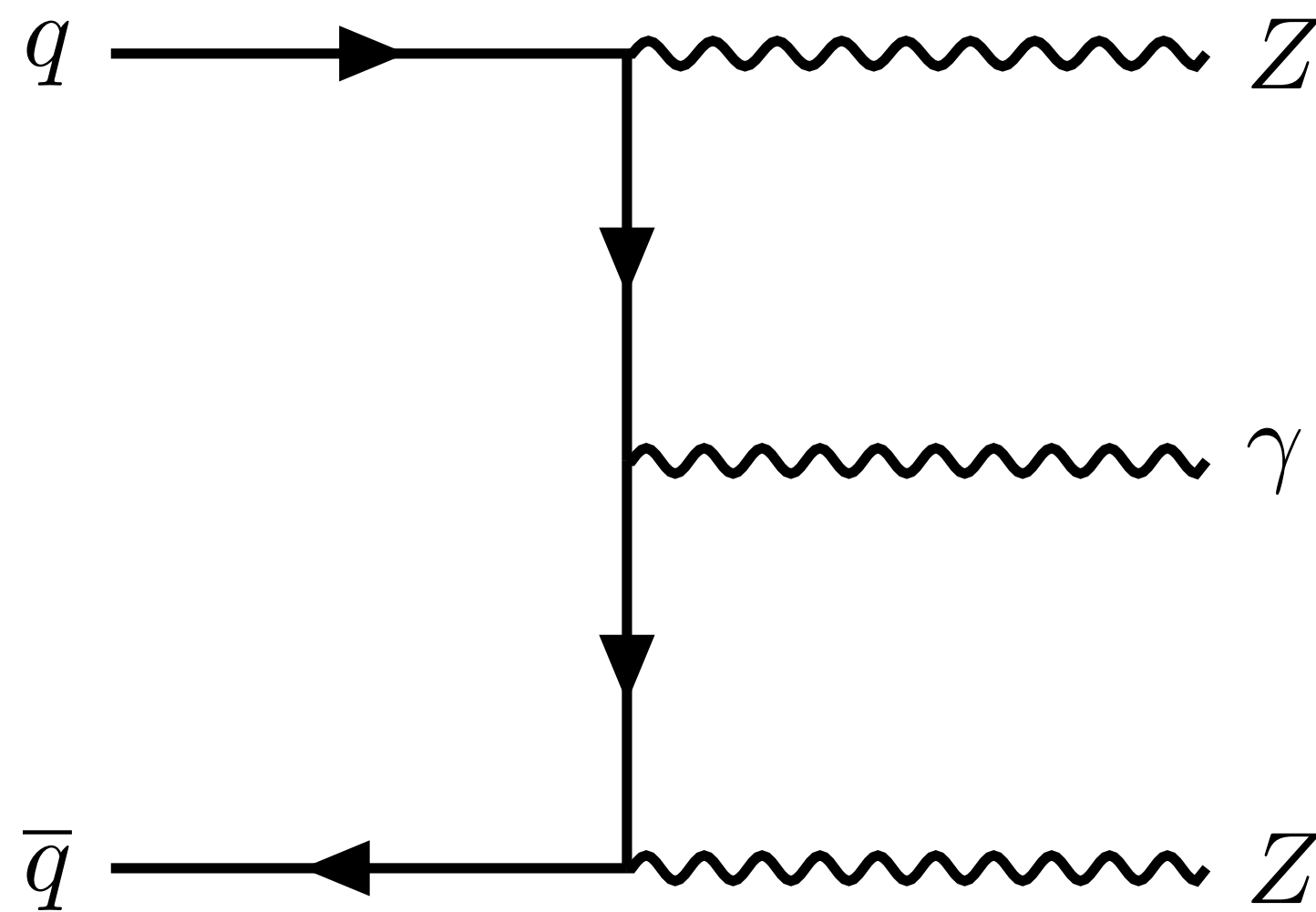


An appropriate final state

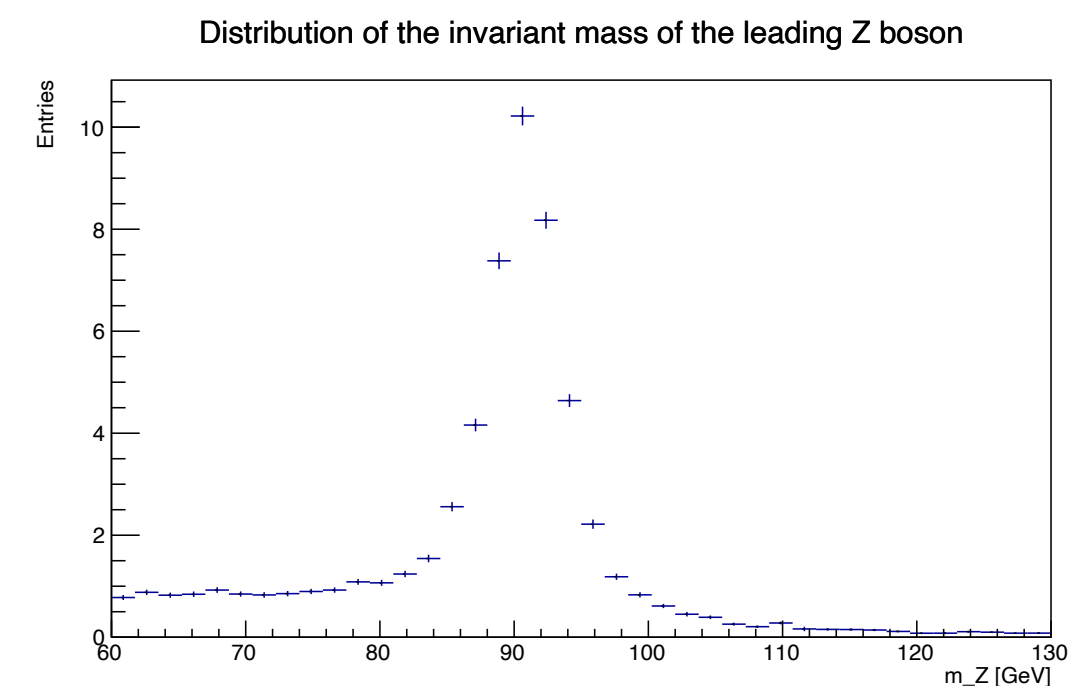
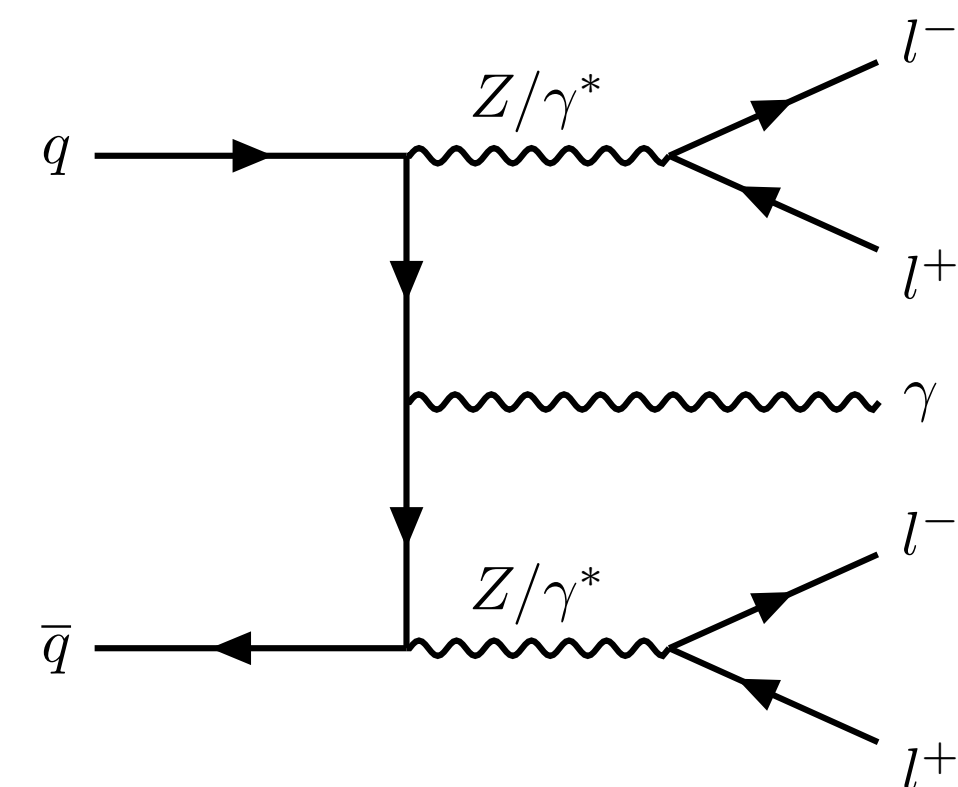


background? ✓

$$q\bar{q} \rightarrow ZZ\gamma$$



Final state particles ✓



How to make an observation



Pick an appropriate final state



Apply an event selection to maximise your S/B

Estimate the irreducible backgrounds

Check if you are above 5σ

Apply a good event selection



1.

Select the final
state particles

2.

Use properties of
the production
channel

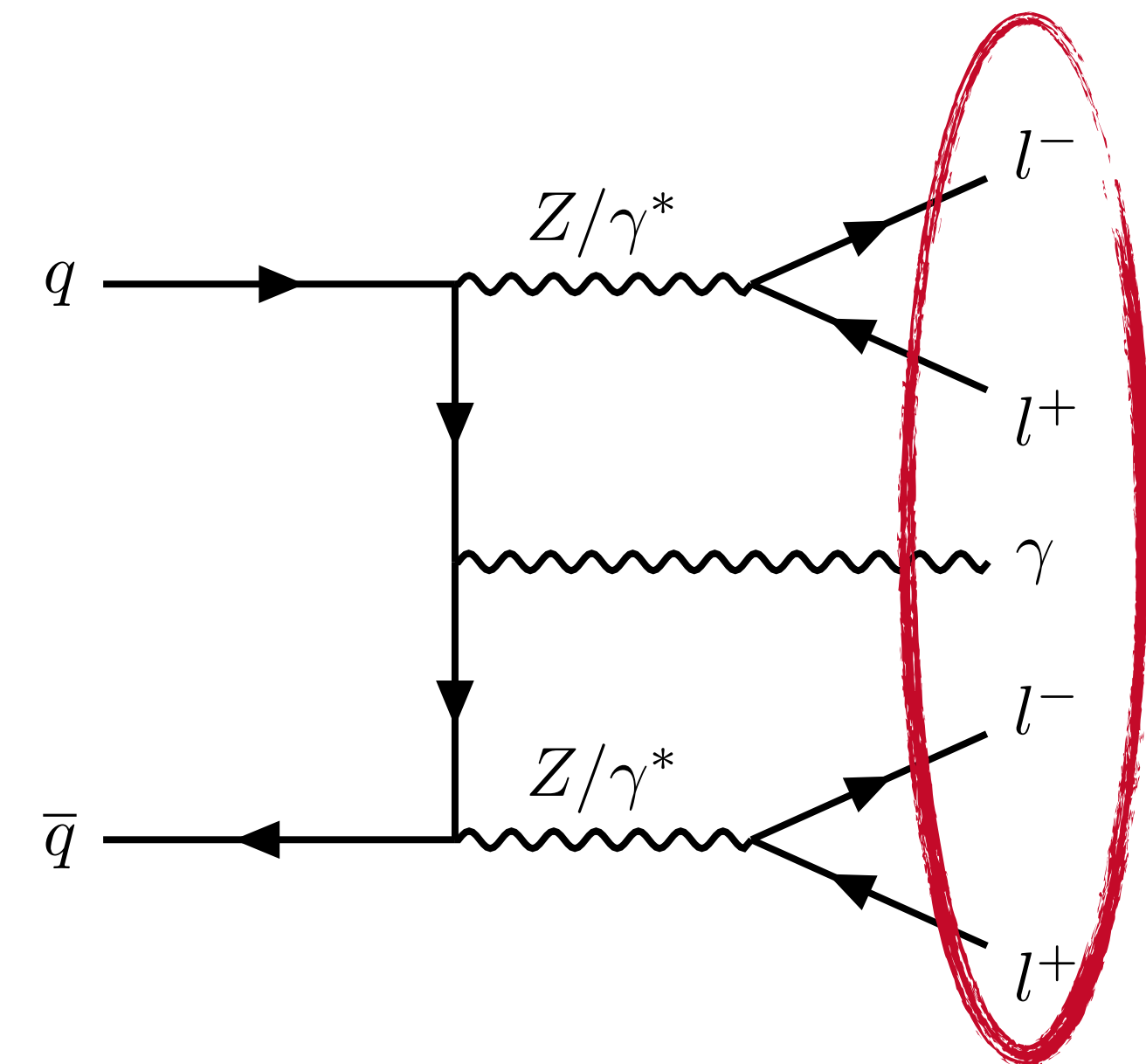
3.

Remove as much
background as
possible

Select the final state particles



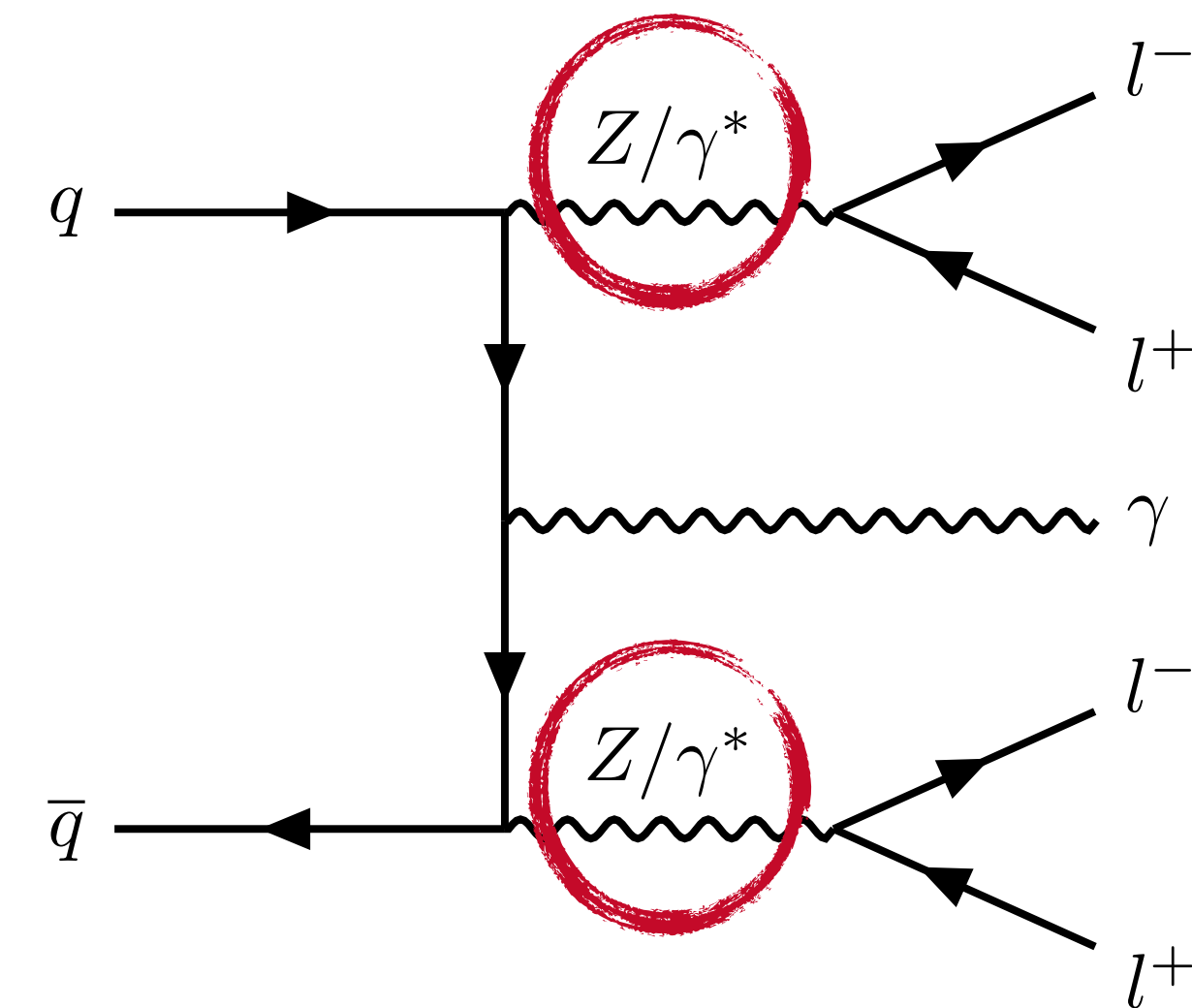
- 2 pairs of 2 opposite-sign same-flavour (OSSF) leptons $\rightarrow 4e, 4\mu, 2e2\mu$
- 1 photon, $p_T > 20 \text{ GeV}$
- invariant mass $m_{\ell\ell} > 40 \text{ GeV}$
- $\min(|m_{\ell\ell,1} - m_Z| + |m_{\ell\ell,2} - m_Z|)$
- FSR rejection



Use properties of the Z boson



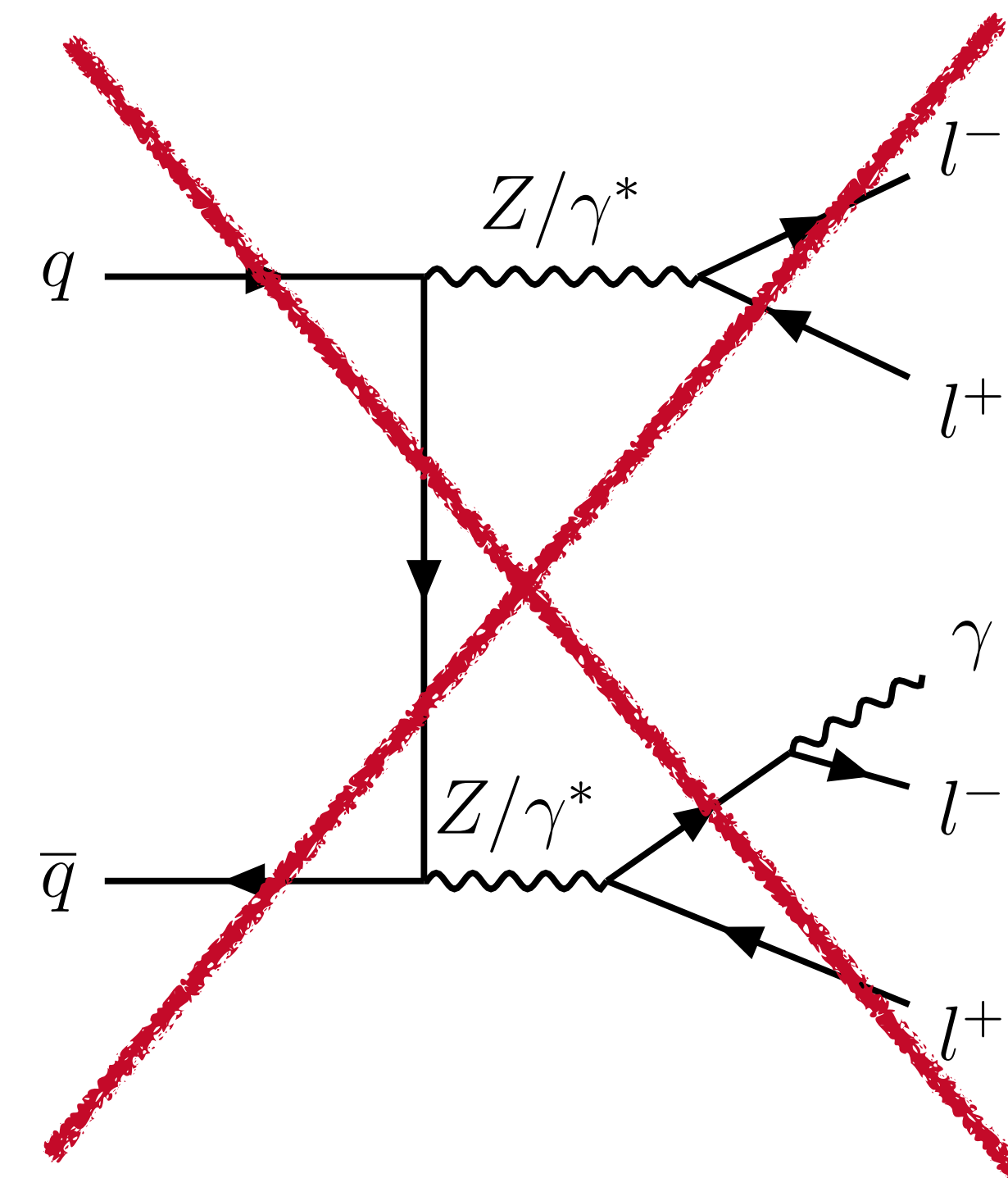
- 2 pairs of 2 opposite-sign same-flavour (OSSF) leptons $\rightarrow 4e, 4\mu, 2e2\mu$
- 1 photon, $p_T > 20 \text{ GeV}$
- invariant mass $m_{\ell\ell} > 40 \text{ GeV}$
- $\min(|m_{\ell\ell,1} - m_Z| + |m_{\ell\ell,2} - m_Z|)$
- FSR rejection



Remove background



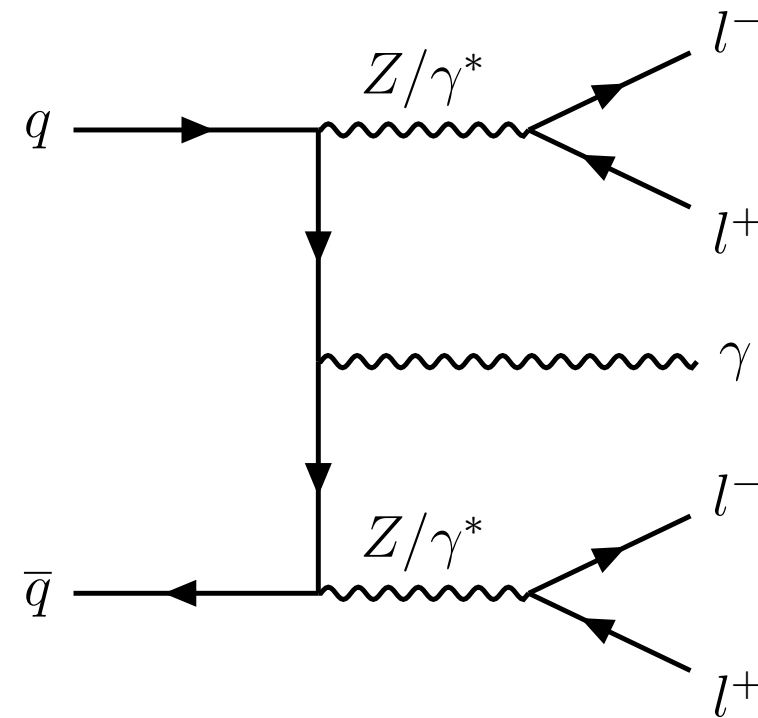
- 2 pairs of 2 opposite-sign same-flavour (OSSF) leptons $\rightarrow 4e, 4\mu, 2e2\mu$
- 1 photon, $p_T > 20$ GeV
- invariant mass $m_{\ell\ell} > 40$ GeV
- $\min(|m_{\ell\ell,1} - m_Z| + |m_{\ell\ell,2} - m_Z|)$
- FSR rejection



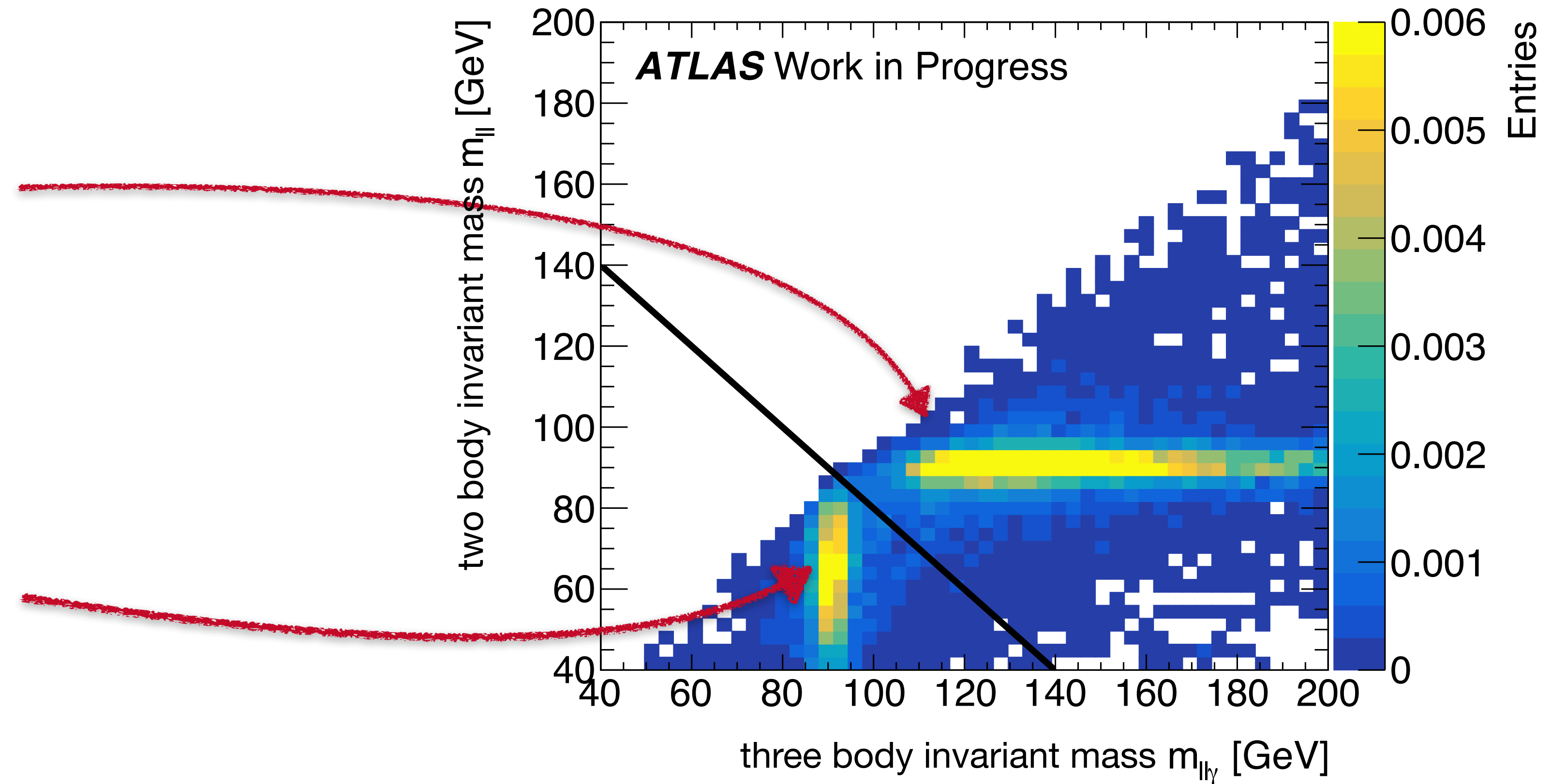
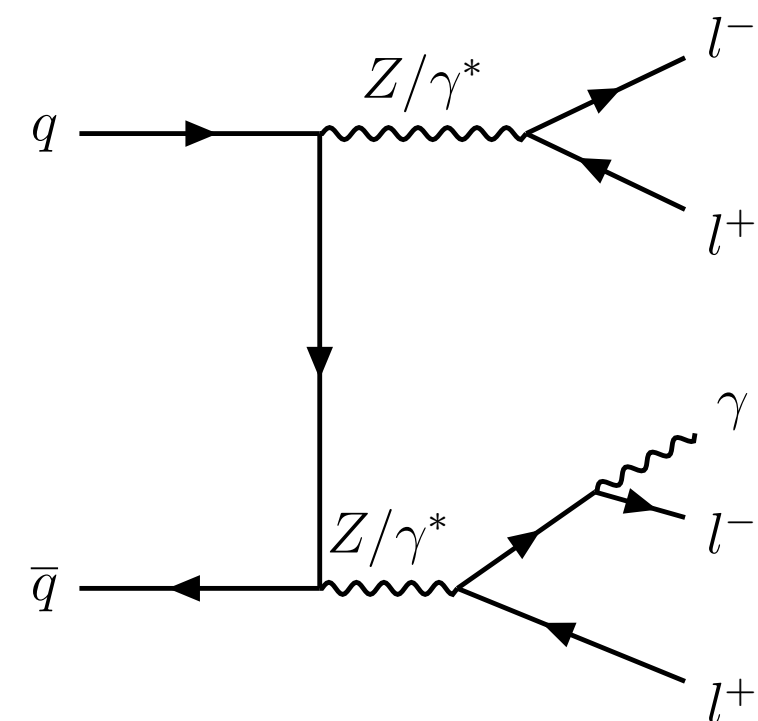
Remove background



Signal



FSR



$$\text{FSR rejection: } (m_{ll\gamma} + m_{ee}) < 2 m_Z$$

How to make an observation



Pick an appropriate final state



Apply an event selection to maximise your S/B



Estimate the irreducible backgrounds

Check if you are above 5σ

Estimate the irreducible background



1.

Determine
potential
backgrounds

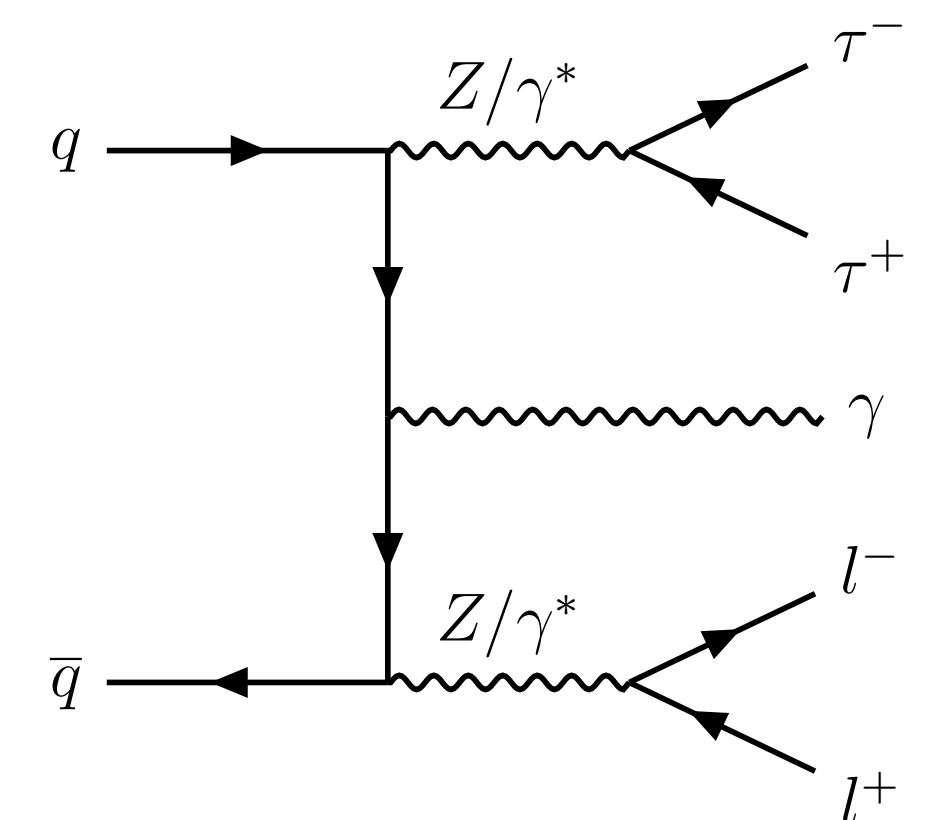
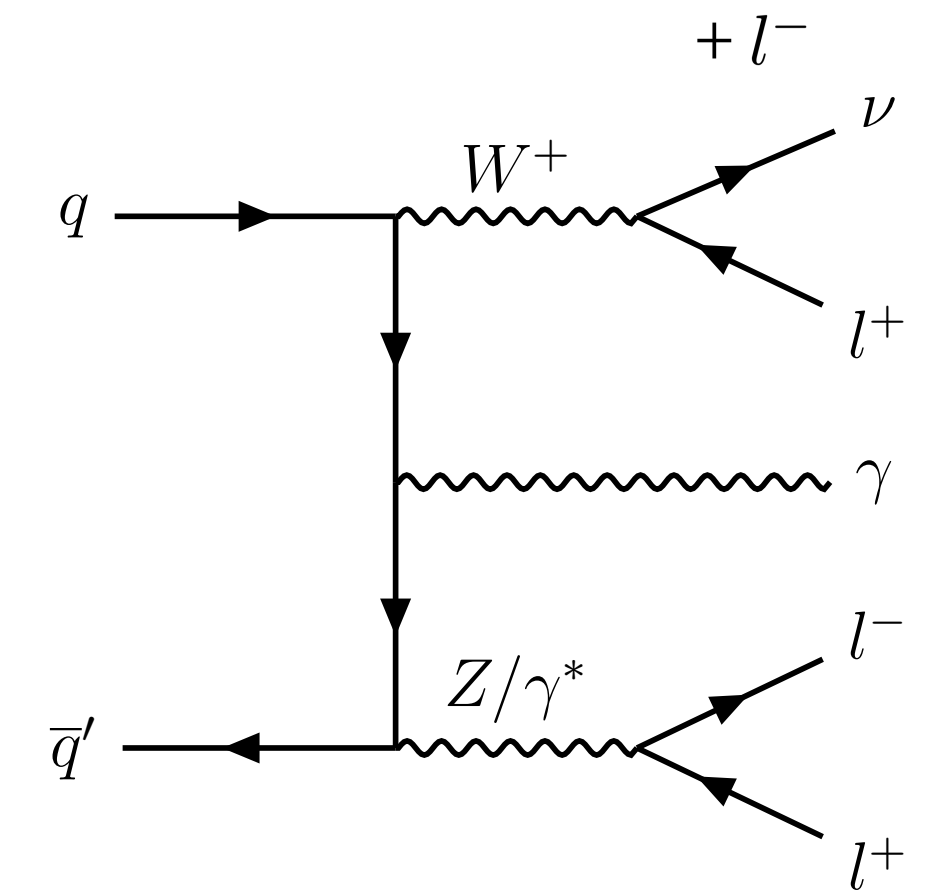
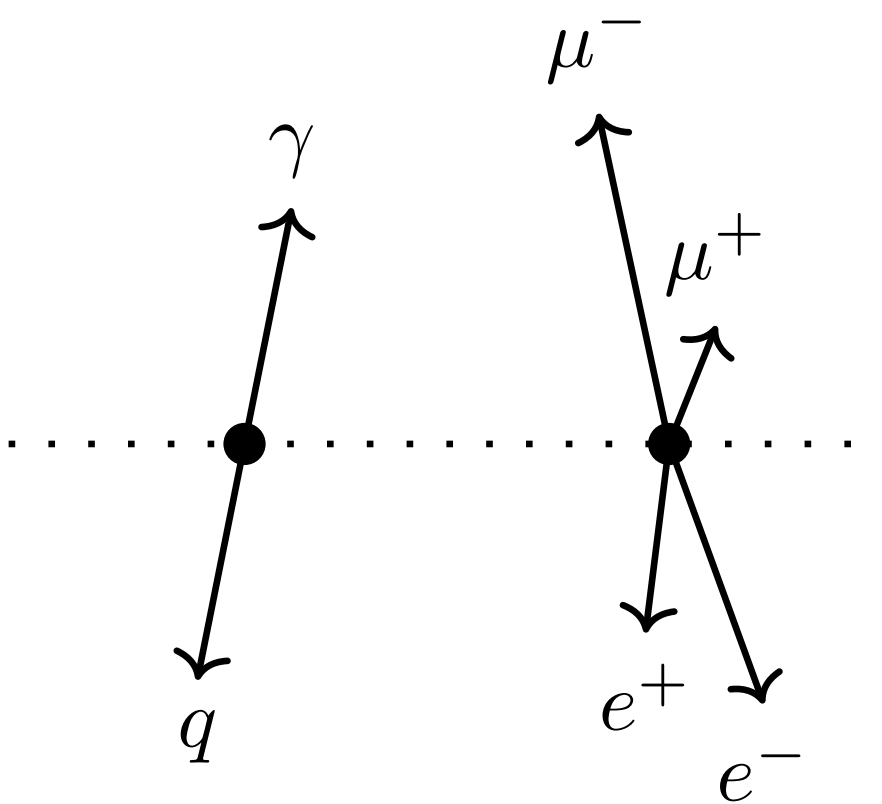
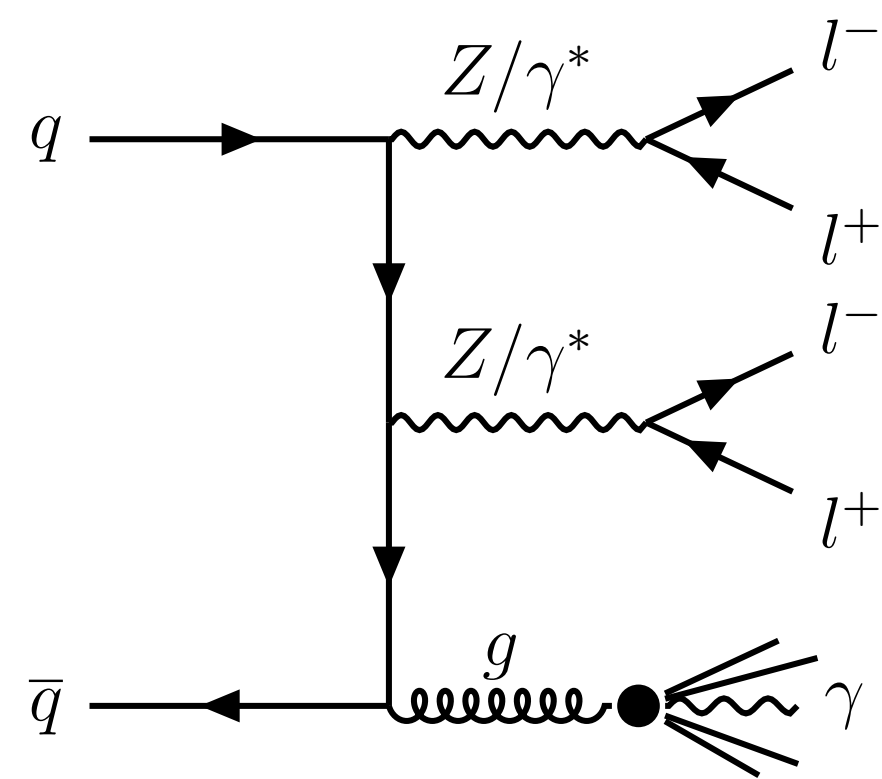
2.

Estimate the
potential
backgrounds

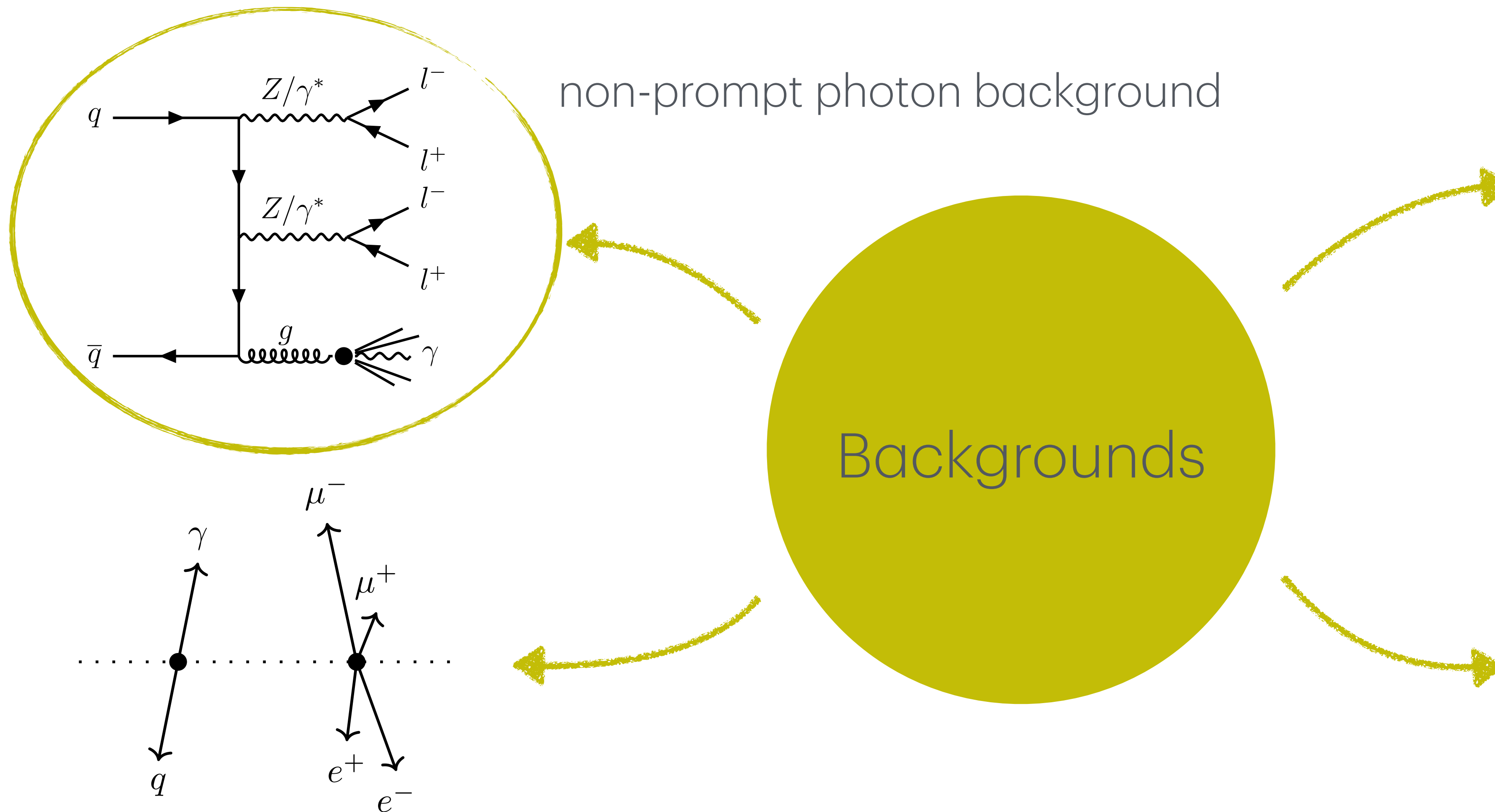
3.

Keep the
uncertainties
low

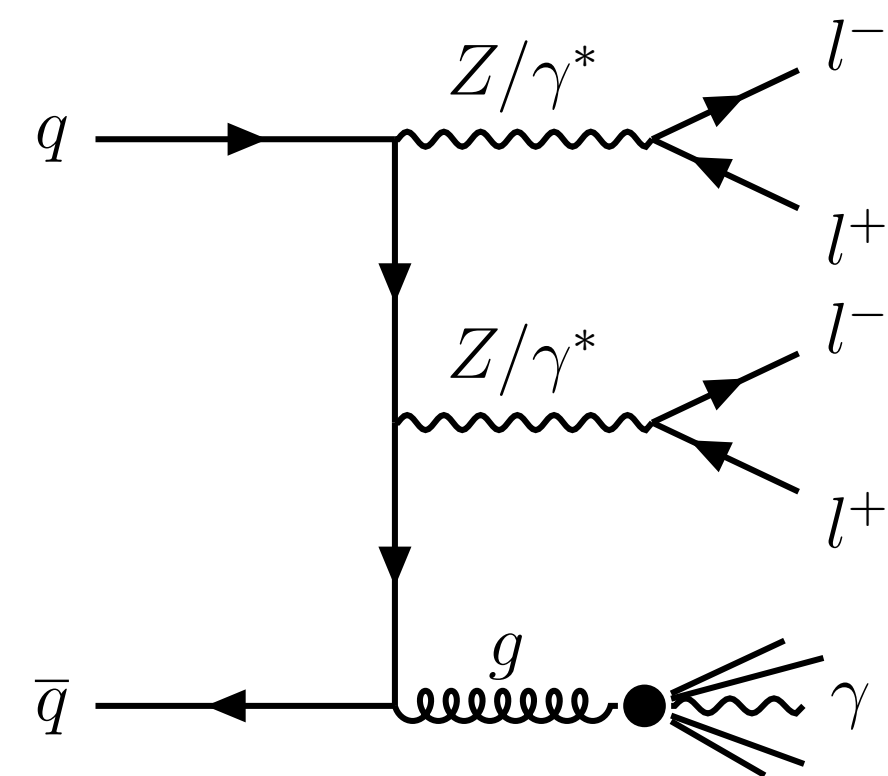
Potential backgrounds



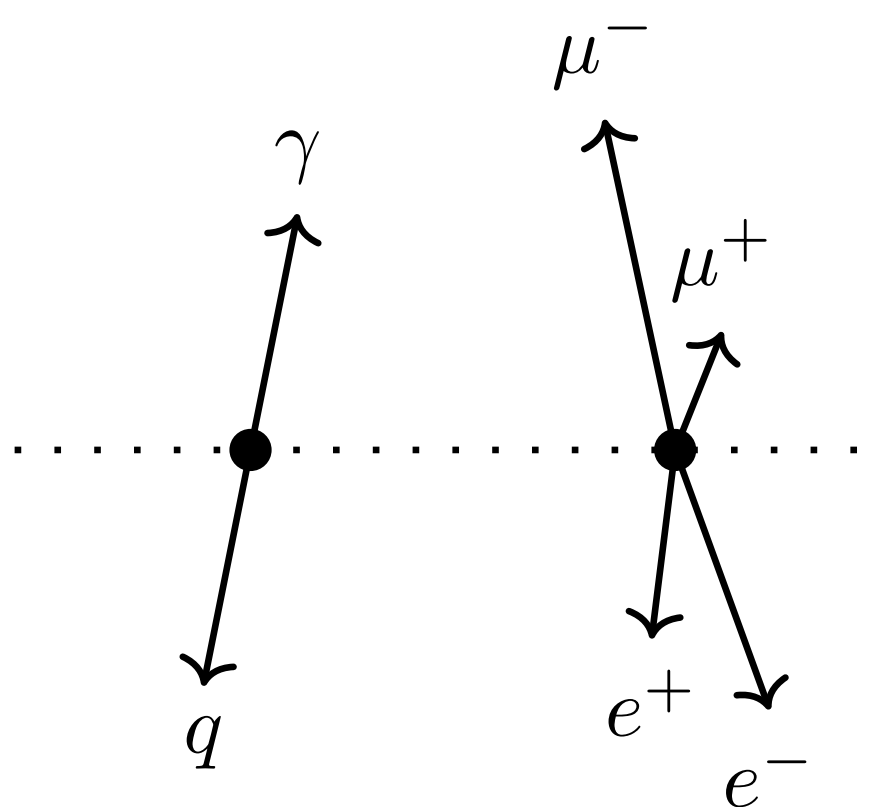
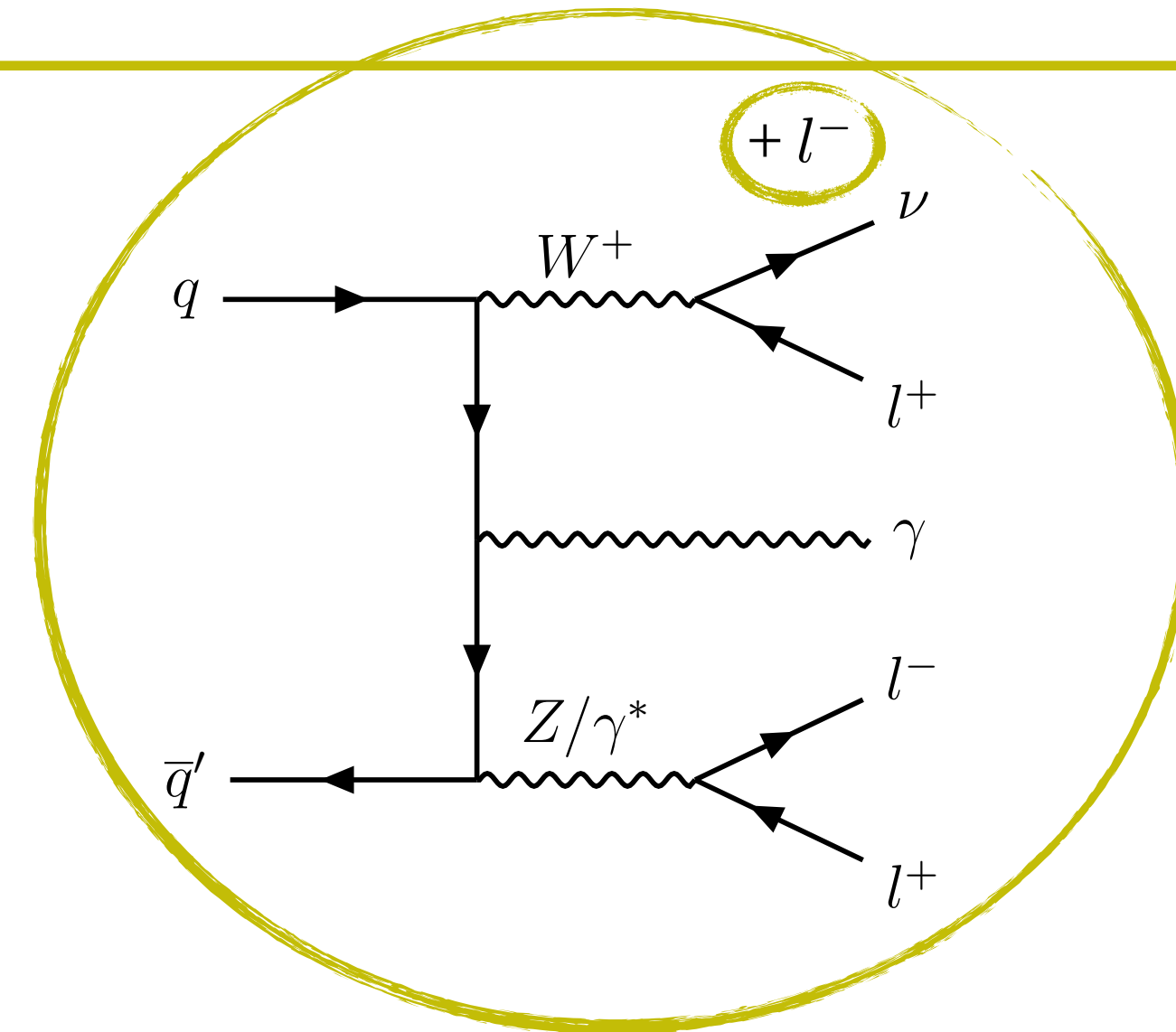
Potential backgrounds



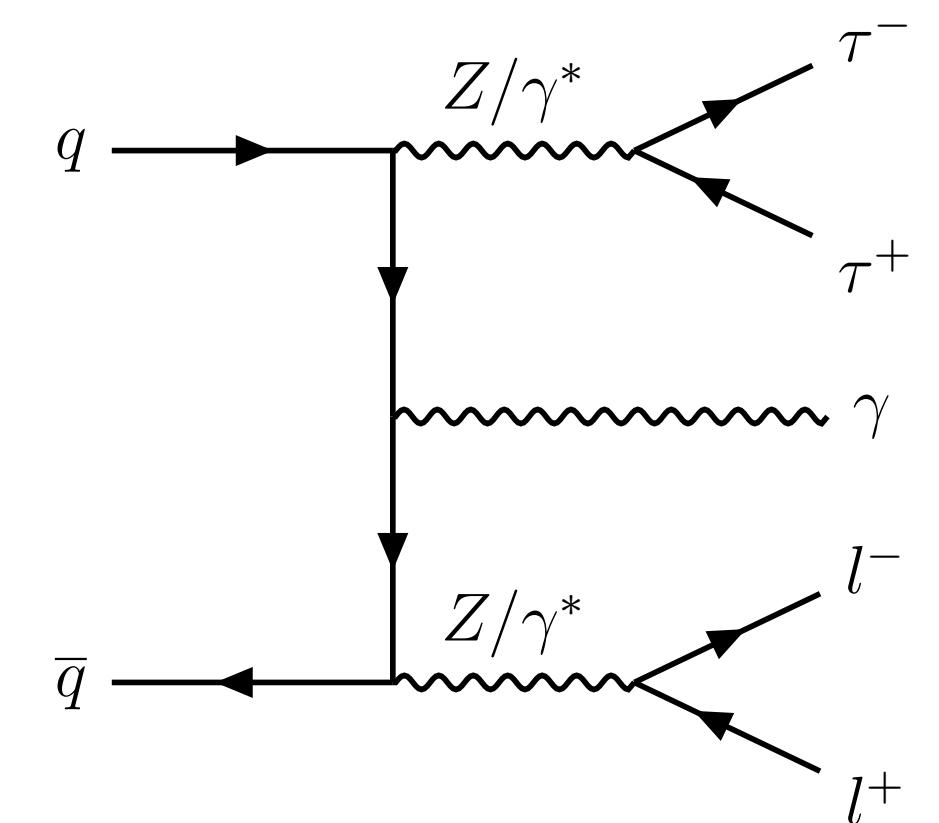
Potential backgrounds



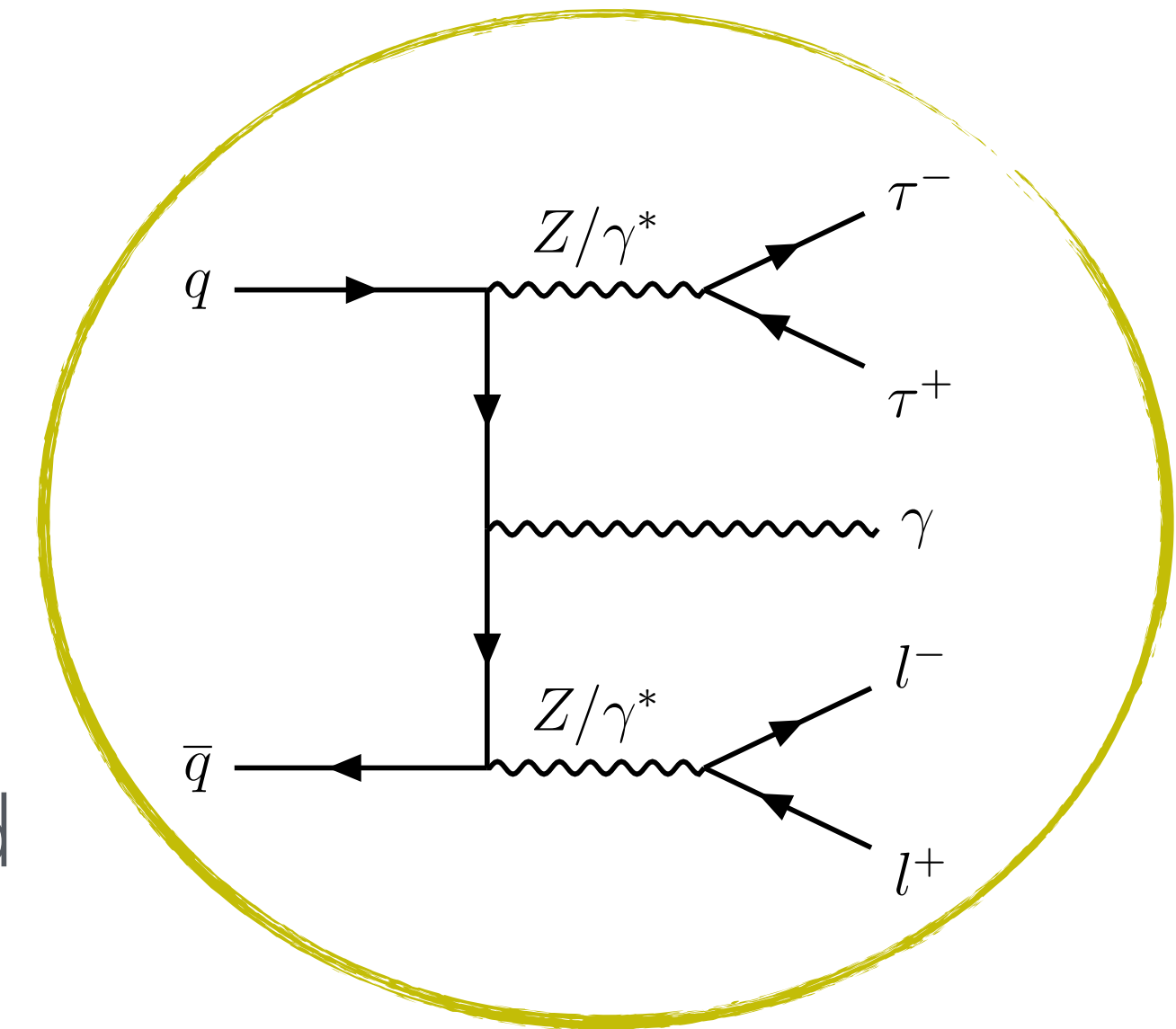
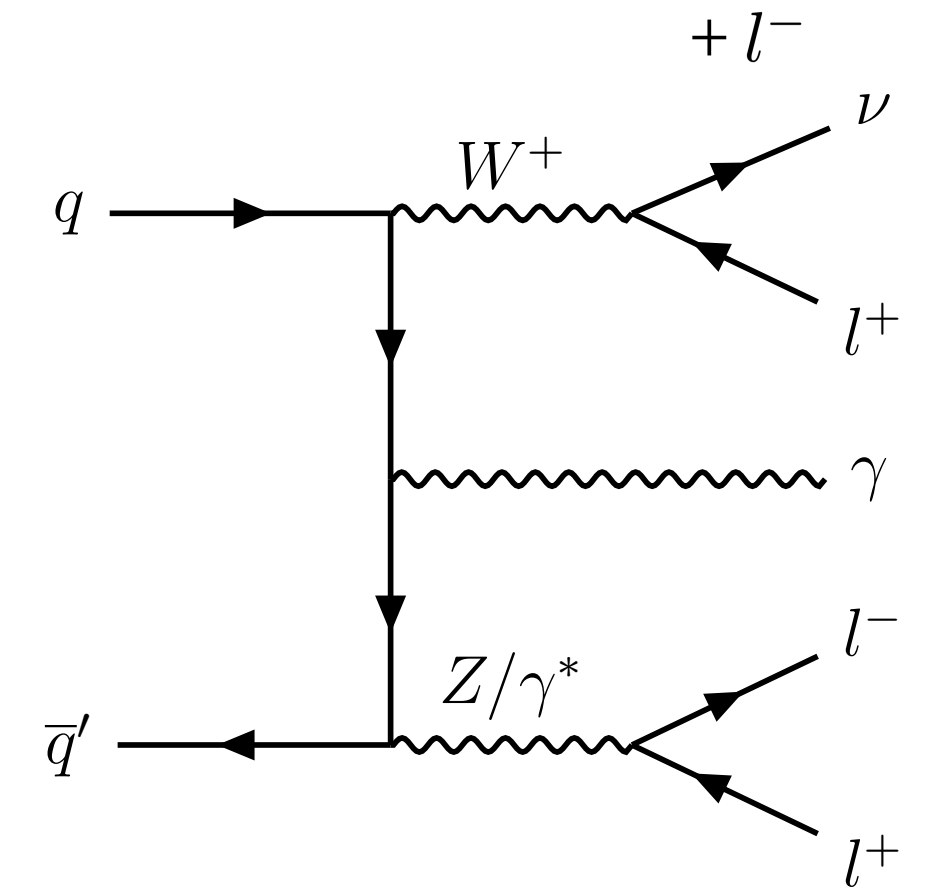
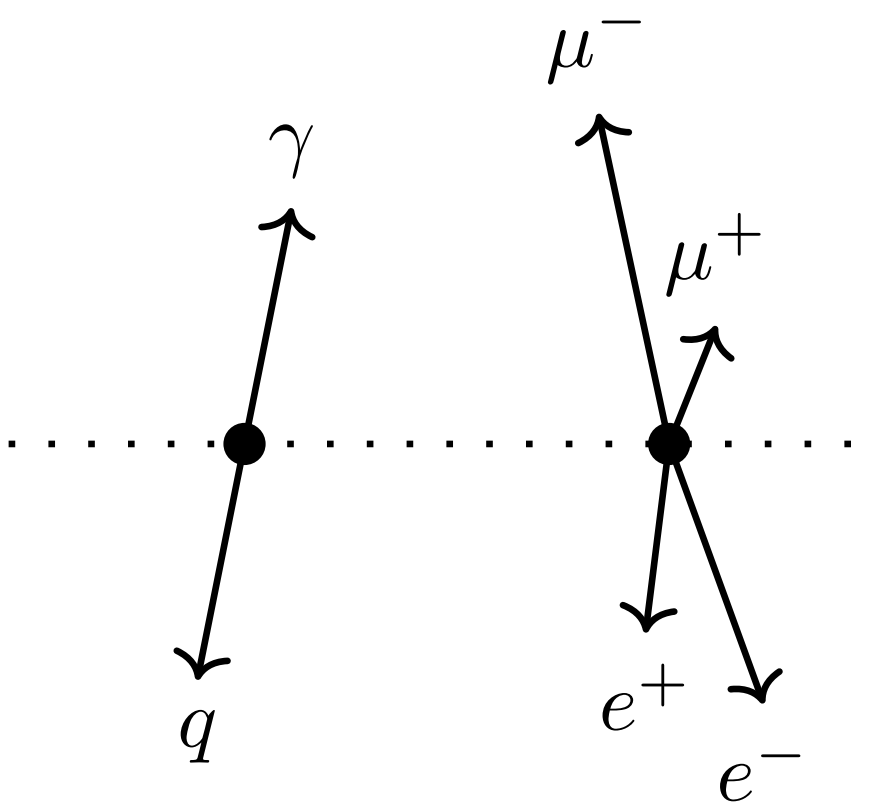
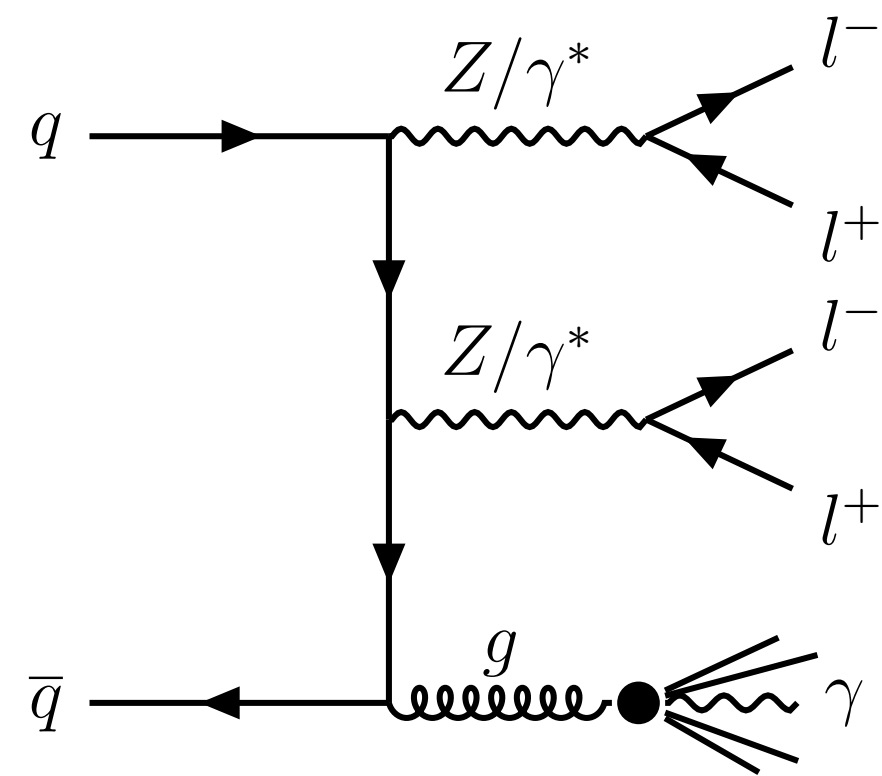
misidentified lepton background



Backgrounds

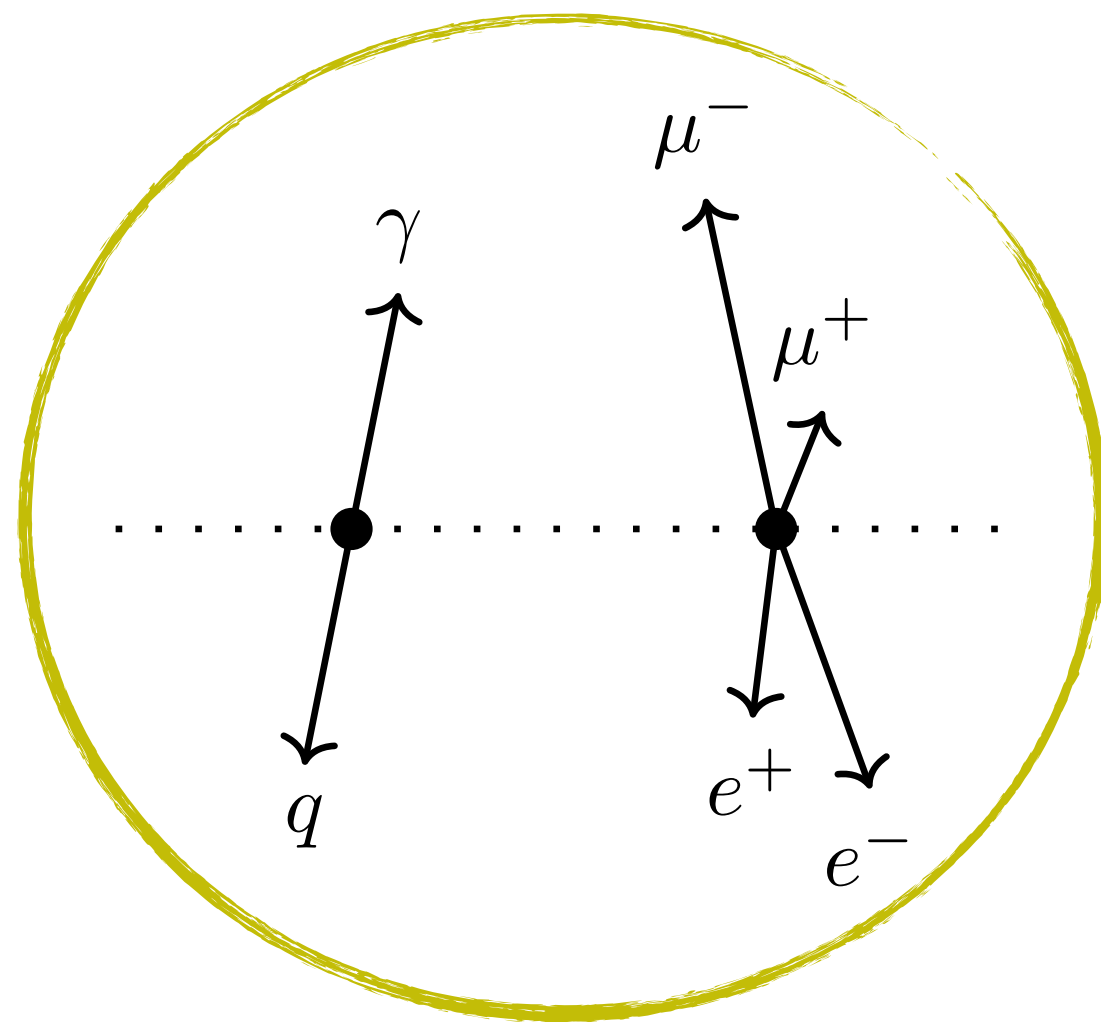
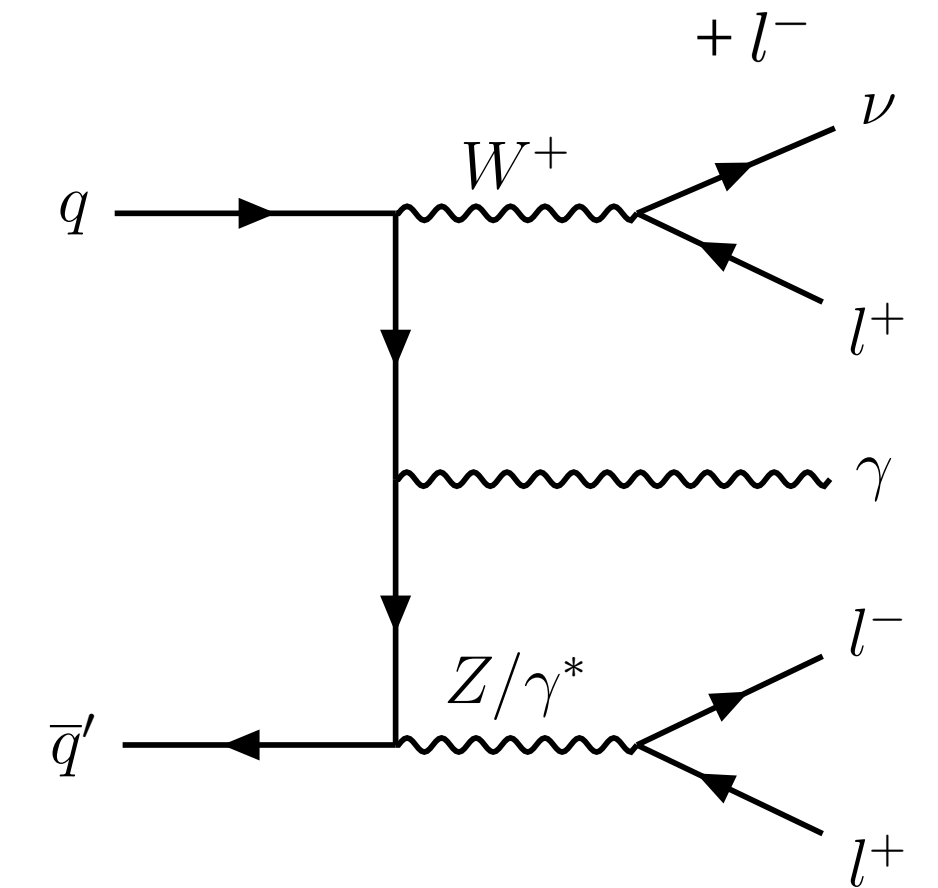
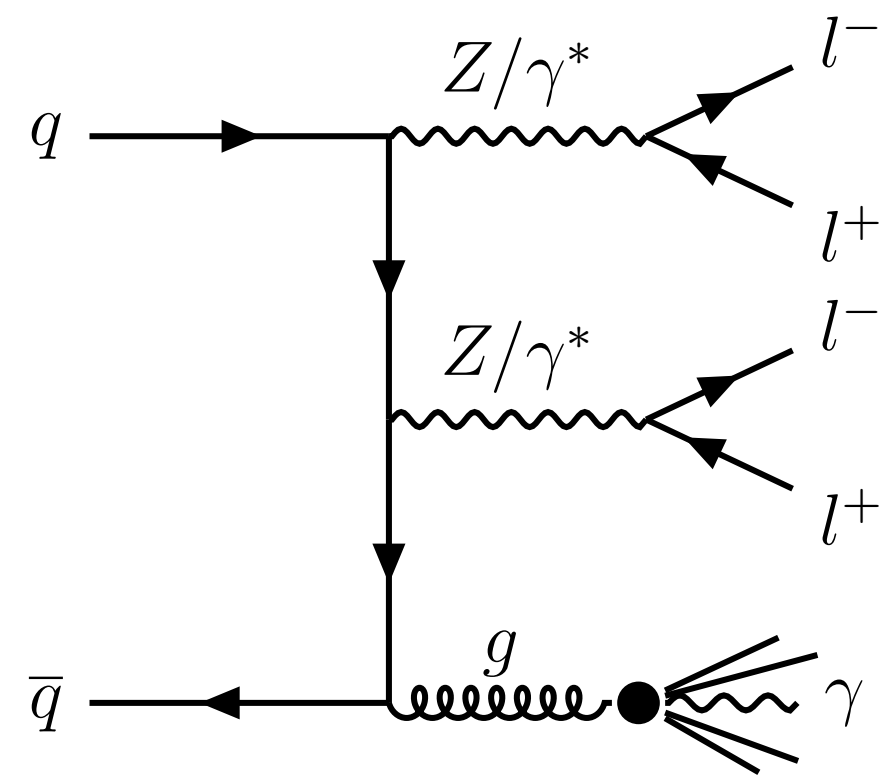


Potential backgrounds

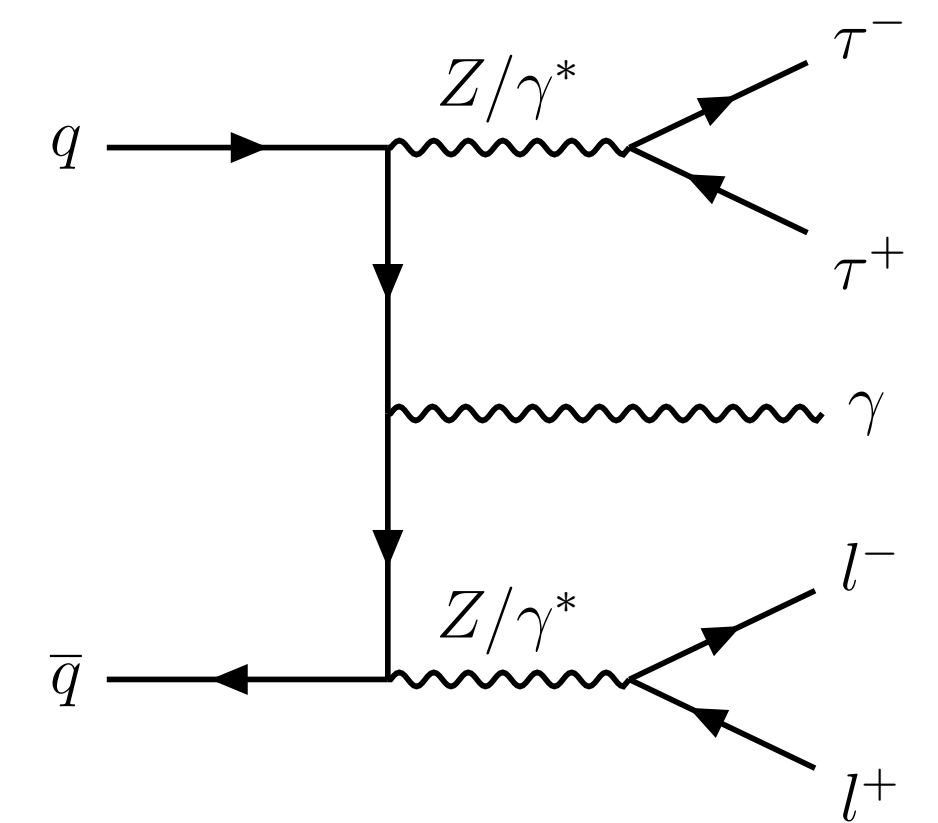


$Z \rightarrow \tau^+\tau^-$ background

Potential backgrounds

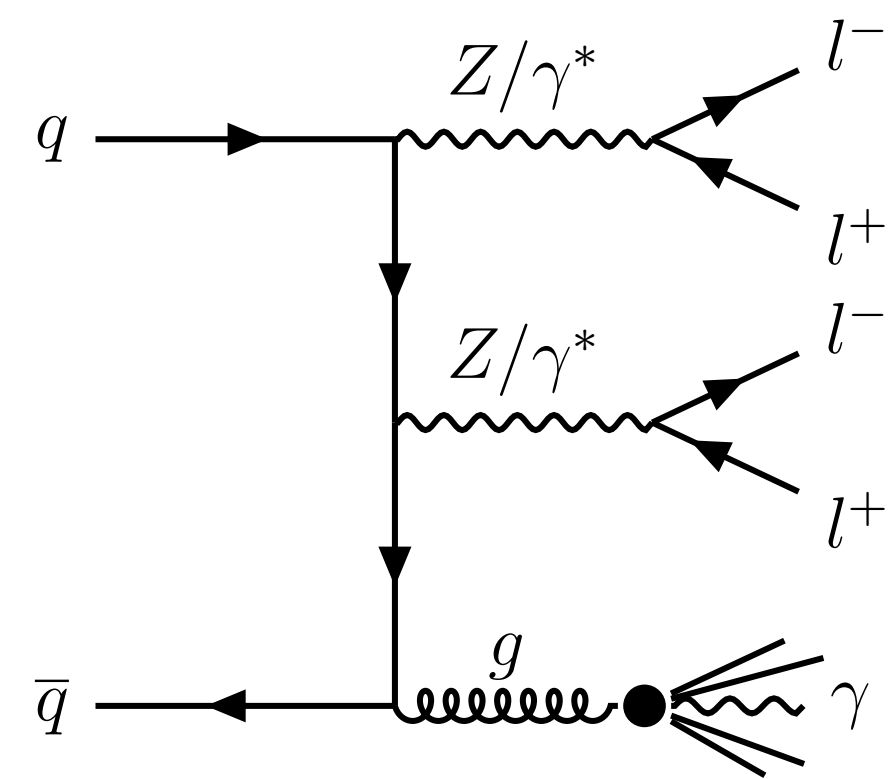


pileup background

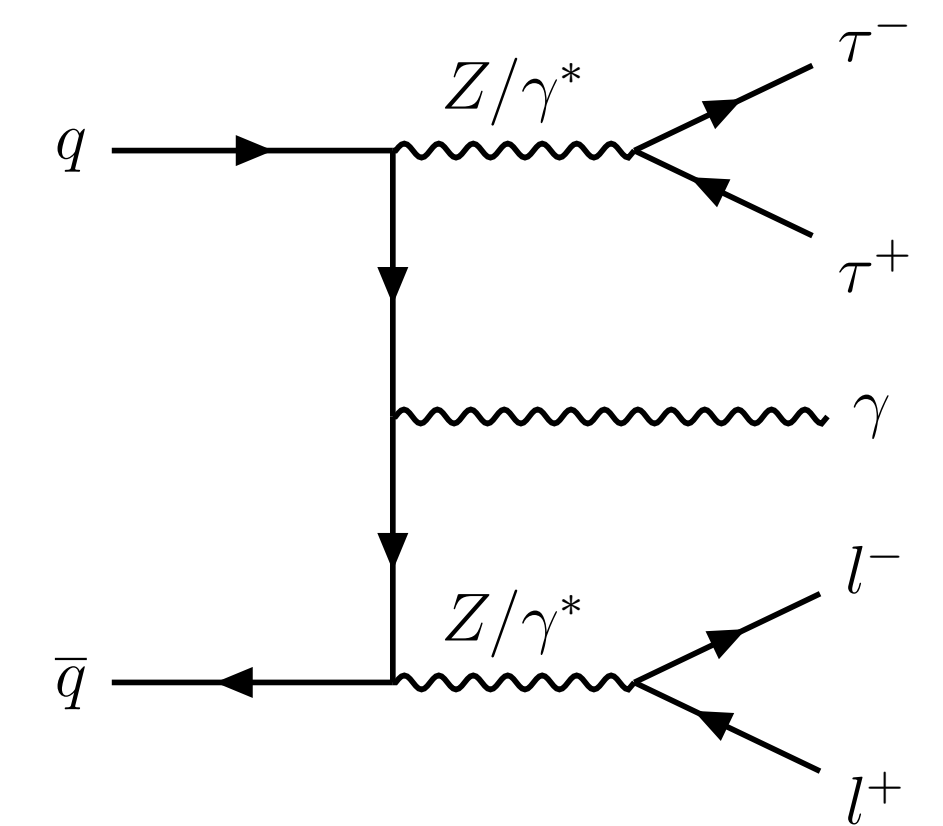
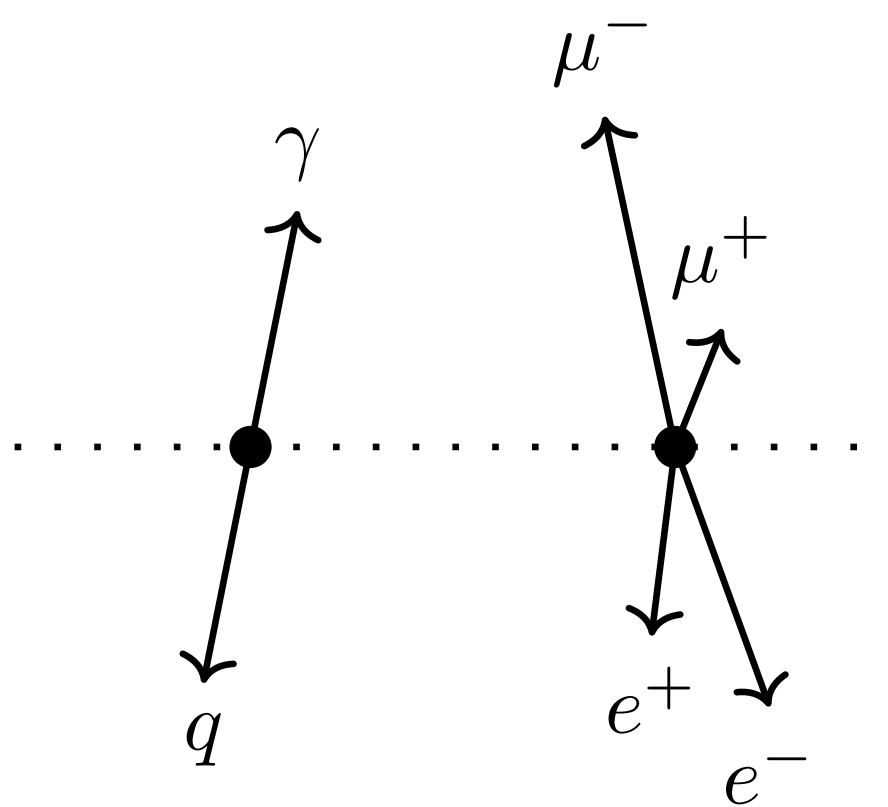
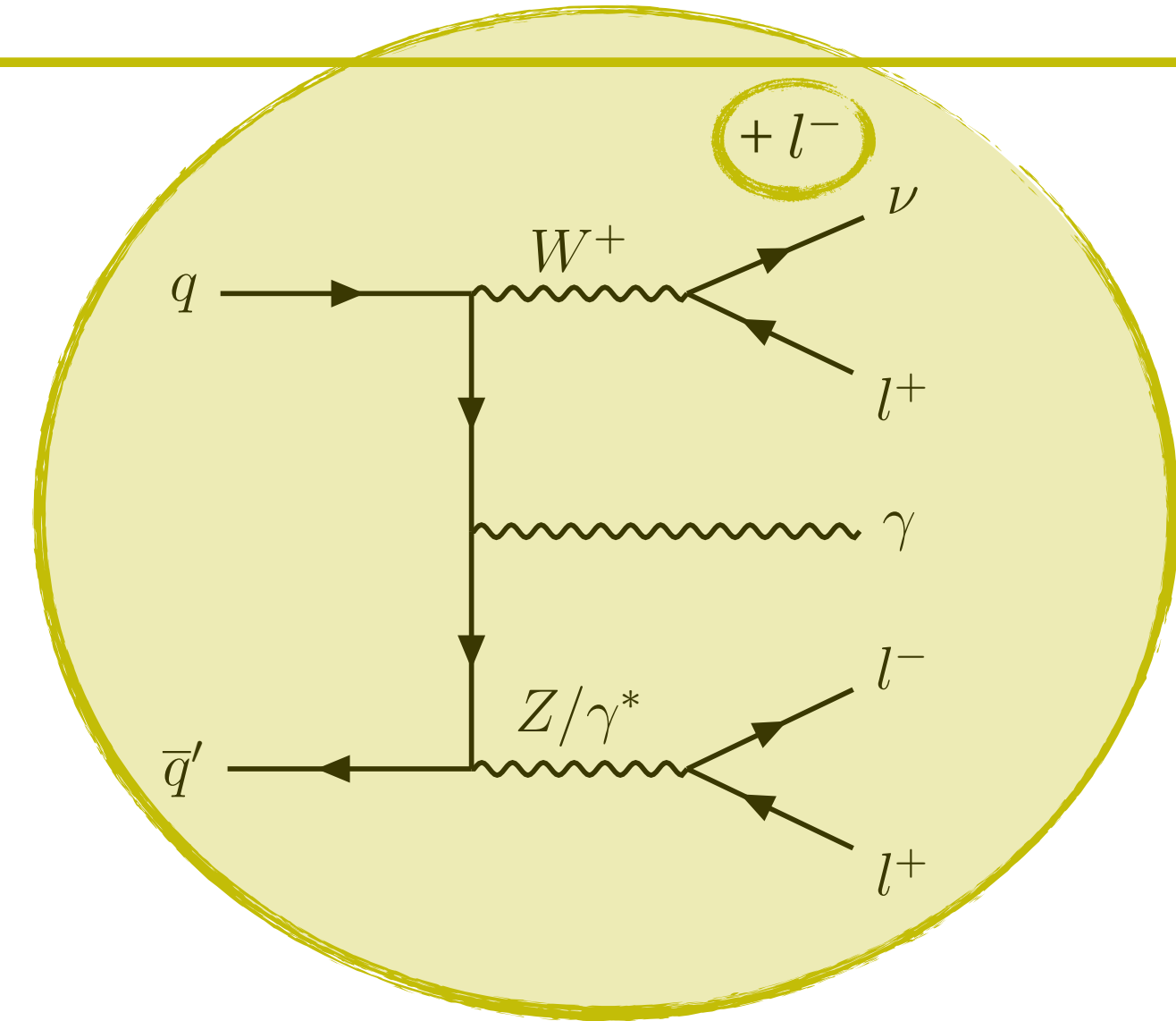


Annweiler am Trifels, ZZ γ

Potential backgrounds



misidentified lepton background

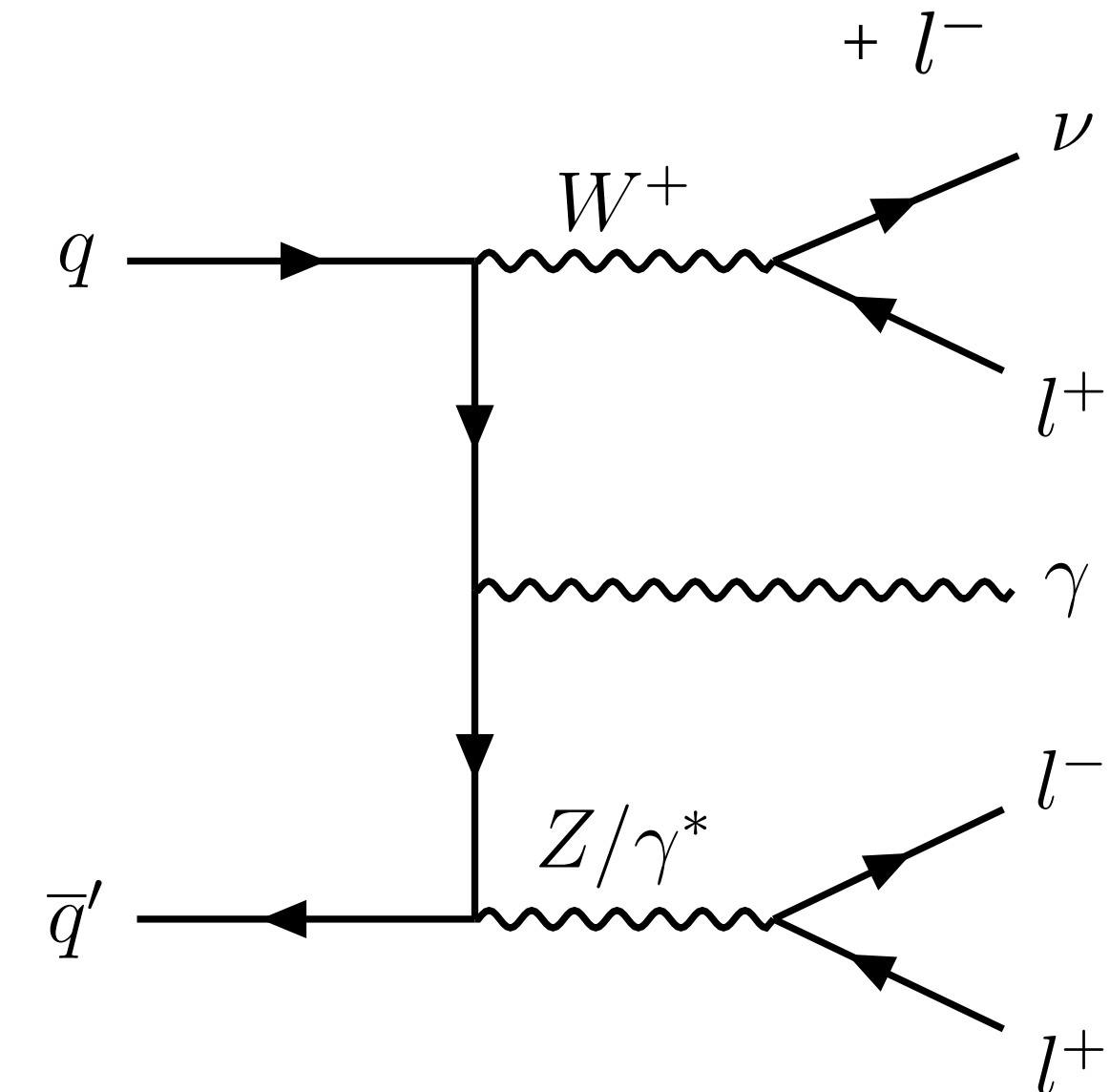


Estimate the background



Overview - misidentified leptons:

- non-prompt leptons mainly from jets
- main backgrounds: $WZ\gamma$, $t\bar{t}\gamma$
- method: matrix method + fake factor
 - real efficiencies from MC
 - fake efficiencies from data (Z CR)



Estimate the background



measurable
quantities

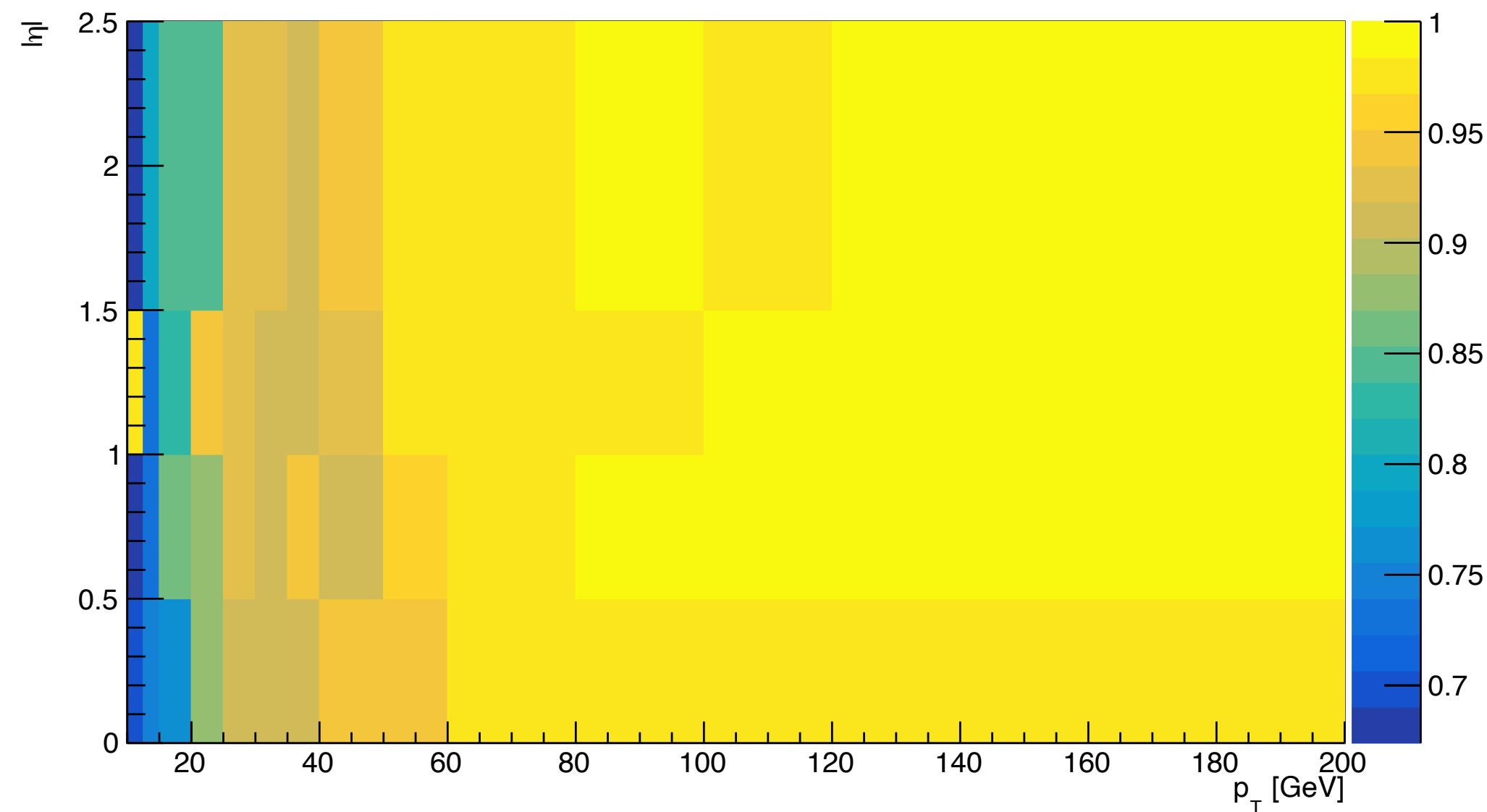
tight leptons
loose leptons

$$\begin{pmatrix} N^t \\ N^l \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_r \\ N_f \end{pmatrix}$$

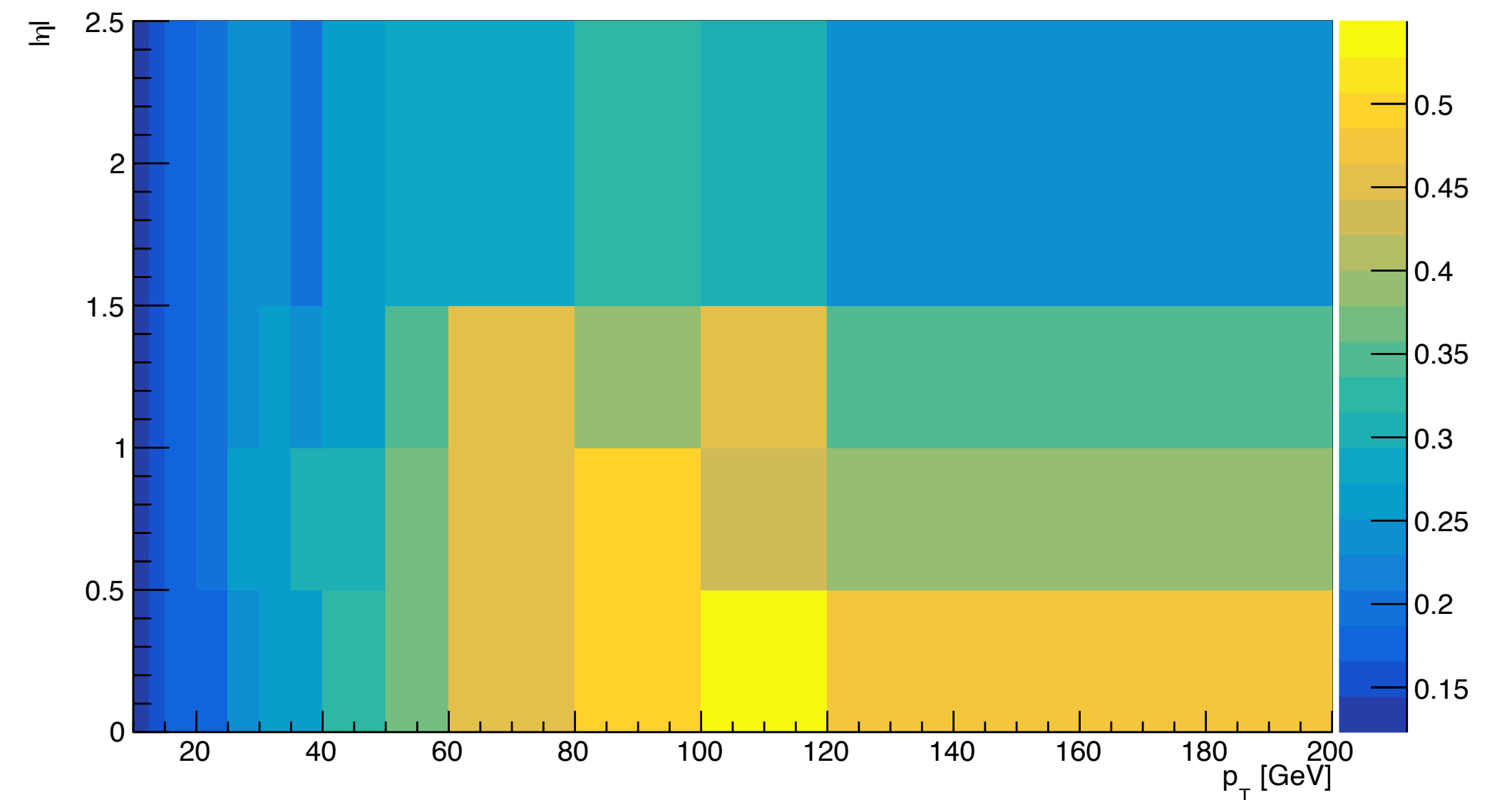
properties

real leptons
fake leptons

Real efficiencies of the electrons



Fake efficiencies of the electrons



Estimate the background



$$\begin{array}{l} \text{CR1} \\ \text{CR2} \end{array} \begin{pmatrix} N^{tt} \\ N^{tl} \\ N^{lt} \\ N^{ll} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1 \bar{r}_2 & r_1 \bar{f}_2 & f_1 \bar{r}_2 & f_1 \bar{f}_2 \\ \bar{r}_1 r_2 & \bar{r}_1 f_2 & \bar{f}_1 r_2 & \bar{f}_1 f_2 \\ \bar{r}_1 \bar{r}_2 & \bar{r}_1 \bar{f}_2 & \bar{f}_1 \bar{r}_2 & \bar{f}_1 \bar{f}_2 \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}$$

$$N_{\text{SR}}^{\text{fake lepton}} = r_1 f_2 N^{rf} + f_1 r_2 N^{fr} + f_1 f_2 N^{ff}$$

Estimate the background



$$\begin{array}{l} \text{CR1} \\ \text{CR2} \end{array} \begin{pmatrix} N^{tt} \\ N^{tl} \\ N^{lt} \\ N^{ll} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1 \bar{r}_2 & r_1 \bar{f}_2 & f_1 \bar{r}_2 & f_1 \bar{f}_2 \\ \bar{r}_1 r_2 & \bar{r}_1 f_2 & \bar{f}_1 r_2 & \bar{f}_1 f_2 \\ \bar{r}_1 \bar{r}_2 & \bar{r}_1 \bar{f}_2 & \bar{f}_1 \bar{r}_2 & \bar{f}_1 \bar{f}_2 \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}$$

$$N_{\text{SR}}^{\text{fake lepton}} = \sum_{i,j} \frac{r_i f_i}{r_i - f_i} \frac{r_j (1 - f_j)}{r_j - f_j} N_{\text{CR1},i,j} - \sum_{i,j} \frac{r_i f_i}{r_i - f_i} \frac{r_j f_j}{r_j - f_j} N_{\text{CR2},i,j}$$

$$N_{\text{CR1},i,j}^{\text{data}} - N_{\text{CR1},i,j}^{\text{ZZ}\gamma}$$

Estimate the background



- CR1: SR with 1 lepton loose & !tight
 - data: 1 ± 1
 - $4\ell+\gamma$ signal sample: 0.436 ± 0.024
- CR2: SR with 2 lepton loose & !tight
 - data: 1 ± 1
 - $4\ell+\gamma$ signal sample: 0.0121 ± 0.00



electrons	matrix method
p_T	-0.012 ± 0.002
η	-0.009 ± 0.001
p_T, η	-0.012 ± 0.002
muons	
p_T	0.17 ± 0.18
η	0.05 ± 0.07
p_T, η	0.15 ± 0.16

Estimate the background



- CR1: SR with 1 lepton loose & !tight
 - data: 1 ± 1
 - $4\ell+\gamma$ signal sample: 0.436 ± 0.024
- CR2: SR with 2 lepton loose & !tight
 - data: 1 ± 1
 - $4\ell+\gamma$ signal sample: 0.0121 ± 0.00



electrons	matrix method
p_T	-0.012 ± 0.00
η	-0.00
p_T, η	± 0.002
muons	
p_T	17 ± 0.18
η	0.07
p_T, η	0.15

unphysical results

too large uncertainties

Estimate the background



- CR1: SR with 1 lepton loose & !tight
 - data: 1 ± 1
 - $4\ell+\gamma$ signal sample: 0.436 ± 0.024
- CR1': SR with 1 lepton loose & !tight, no photon
 - data: 553 ± 24
 - 4ℓ MC sample: 233.6 ± 2.0

$$\frac{N_{ZZ\gamma}^{\ell \text{ fake}}}{N_{ZZ\gamma}} = \frac{N_{ZZ}^{\ell \text{ fake}}}{N_{ZZ}}$$

$$\left. \begin{aligned} \frac{N_{ZZ\gamma}^{CR_1}}{N_{ZZ\gamma}} &= \frac{0.436 \pm 0.004}{5.83 \pm 0.08} = 0.0748 \pm 0.0042 \\ \frac{N_{ZZ}^{CR_1'}}{N_{ZZ}} &= \frac{233.6 \pm 2.0}{3112.6 \pm 7.2} = 0.0750 \pm 0.0007 \end{aligned} \right\} 0.06 \sigma$$

Estimate the background



Misidentified leptons in ZZ:

electrons	fake estimate
p_T	$27.2 \pm 6.2 \pm 2.5$
η	$18.0 \pm 4.5 \pm 1.5$
p_T, η	$26.5 \pm 6.6 \pm 2.6$
muons	
p_T	$6.9 \pm 1.3 \pm 1.1$
η	$6.6 \pm 1.1 \pm 1.2$
p_T, η	$6.8 \pm 1.5 \pm 1.3$

Systematic uncertainties

- stat. uncertainty from efficiencies propagated
- syst. uncertainty through binning (one bin η , 3 bins p_T)
- variations of prompt MC background $\pm 10\%$
- difference to fake factor method

Estimate the background



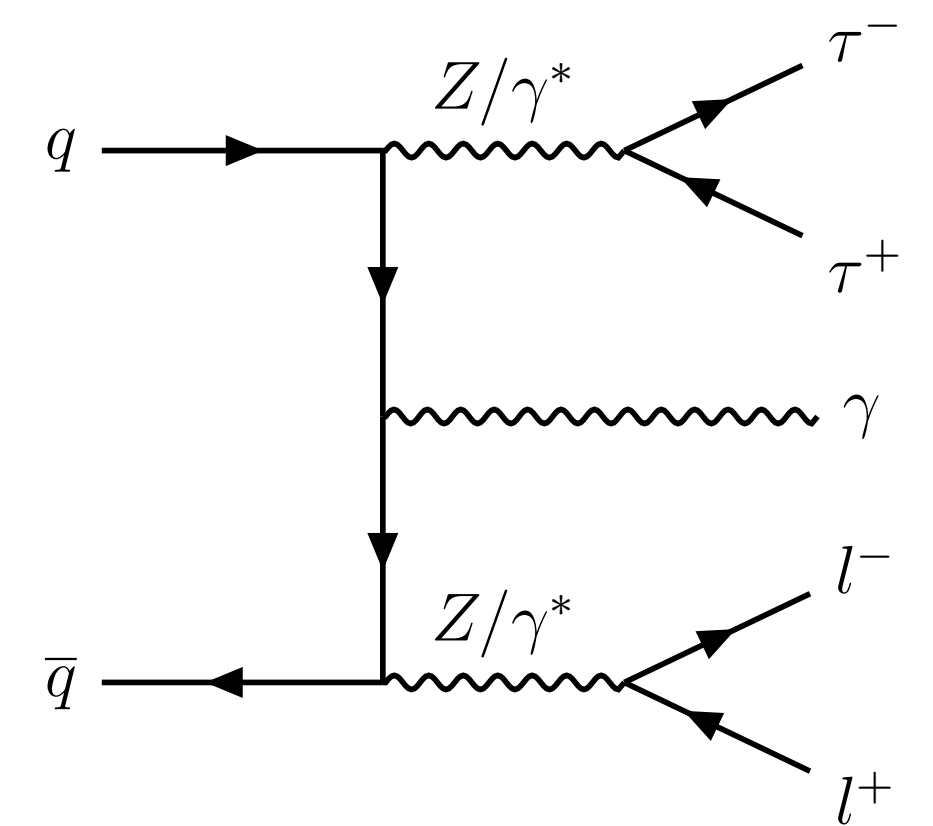
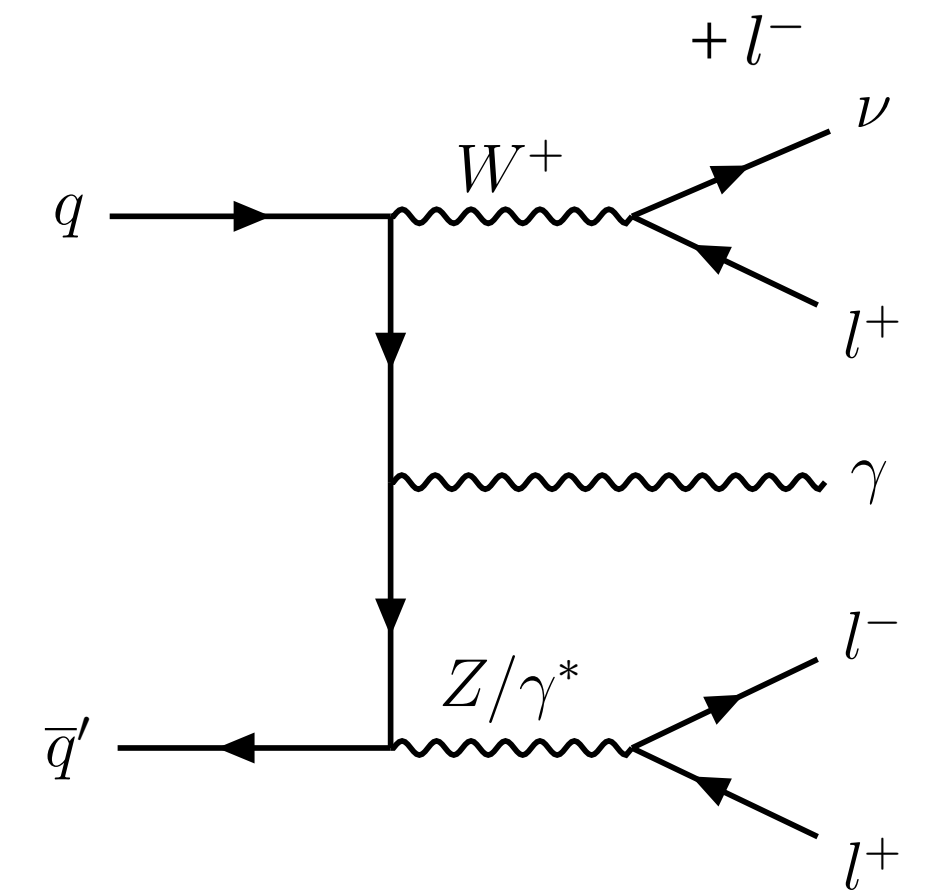
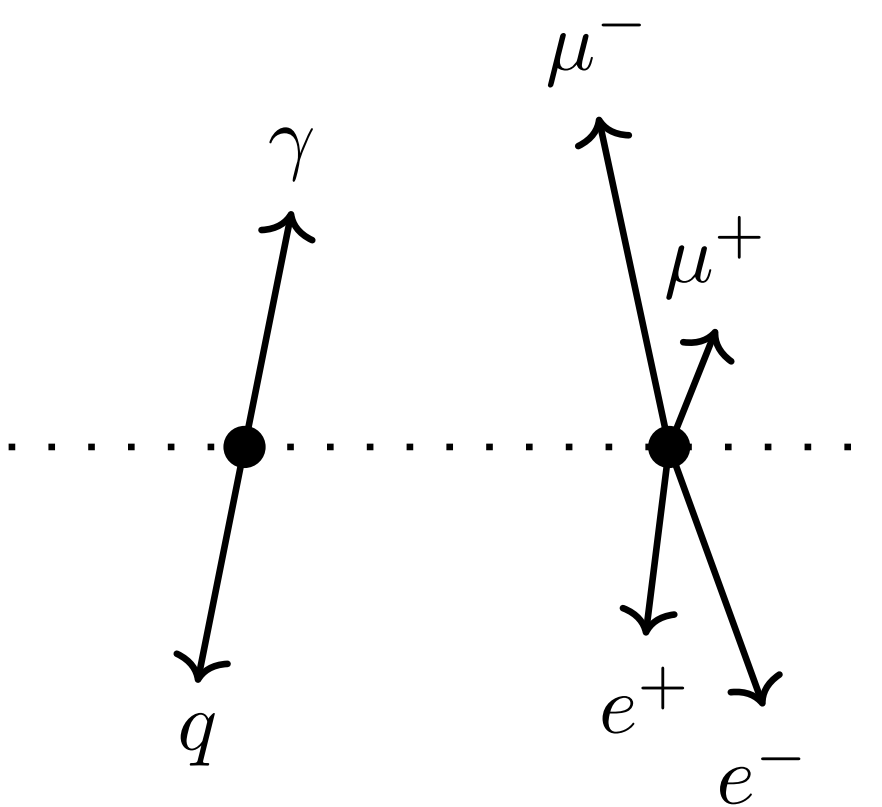
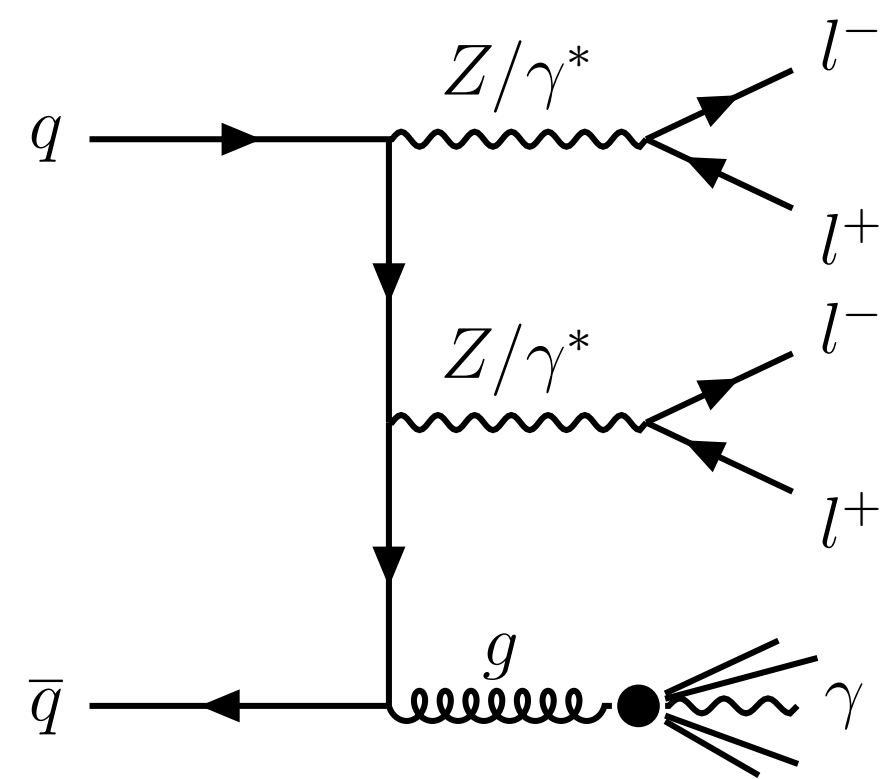
$$N_{ZZ\gamma}^{\ell \text{ fake}} = \frac{N_{ZZ}^{\ell \text{ fake}}}{N_{ZZ}} N_{ZZ\gamma}$$

$$N_{ZZ\gamma}^{\ell \text{ fake}} = 0.062 \pm 0.014$$

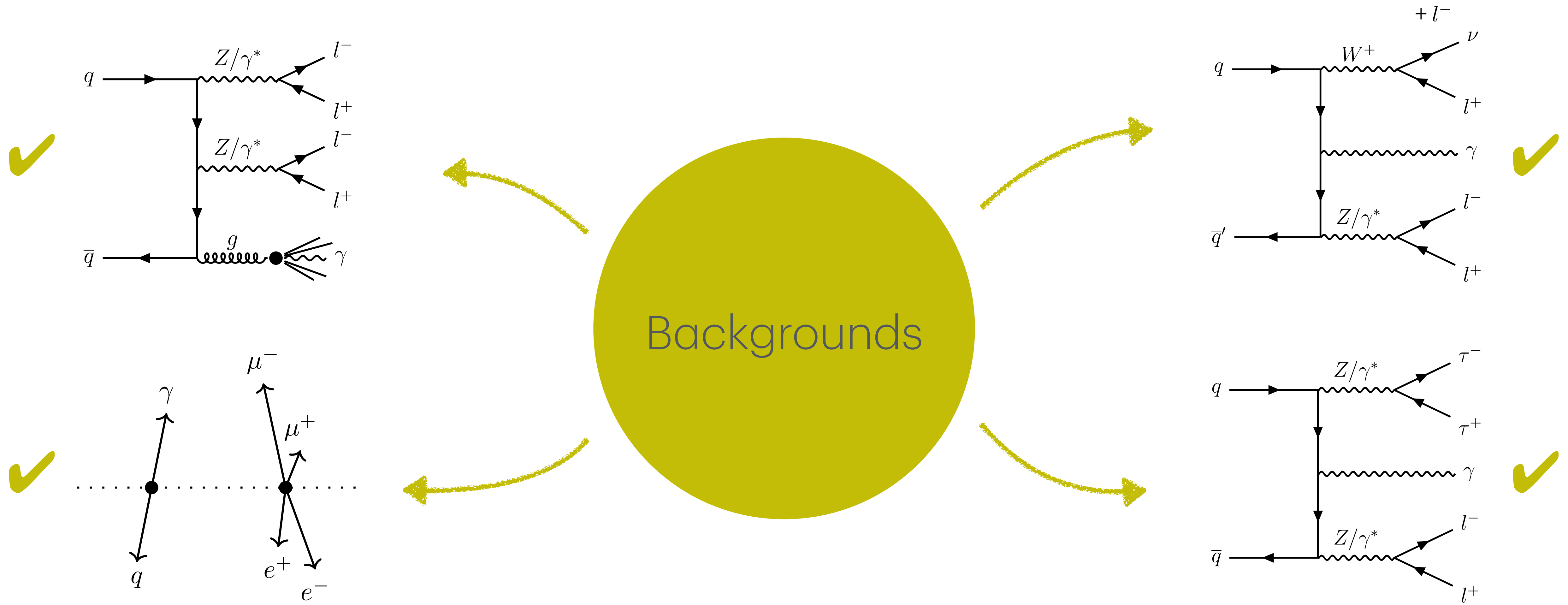
Misidentified leptons in $ZZ\gamma$:

electrons	fake estimate
p_T	0.051 ± 0.013
η	0.034 ± 0.009
p_T, η	0.050 ± 0.013
muons	
p_T	0.013 ± 0.003
η	0.012 ± 0.003
p_T, η	0.013 ± 0.004

Potential backgrounds



Potential backgrounds



How to make an observation



Pick an appropriate final state



Apply an event selection to maximise your S/B



Estimate the irreducible backgrounds



Check if you are above 5σ

Derive the signal significance



1.

Combine all
backgrounds

2.

Get the right
formula

3.

Above 5σ ?

Overview - expected events



	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 \pm 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 \pm 0.03
Signal	6.57 \pm 0.21
Non-prompt Photons	0.49 \pm 0.09
Misidentified Leptons	0.062 \pm 0.014
Pile up	0.056 \pm 0.014
$Z \rightarrow \tau\tau$ contribution	0.035 \pm 0.006
Total Background	0.64 \pm 0.09
N_{tot}	7.21 \pm 0.23

Overview - expected events



	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 \pm 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 \pm 0.03
Signal	6.57 \pm 0.21
Non-prompt Photons	0.49 \pm 0.09
Misidentified Leptons	0.062 \pm 0.014
Pile up	0.056 \pm 0.014
$Z \rightarrow \tau\tau$ contribution	0.035 \pm 0.006
Total Background	0.64 \pm 0.09
N_{tot}	7.21 \pm 0.23

What formula?



$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$

$$Z = \frac{s}{b}$$

Z : significance

n : number of events

s : number of signal events

b : number of background events

σ : total background uncertainty

$$Z = \frac{s}{\sqrt{s + b}}$$

$$Z = \frac{s}{\sqrt{b + \sigma^2}}$$

$$Z = \frac{s}{\sqrt{b}}$$

$$Z = \sqrt{2 \left(n \ln \left[\frac{n}{b} \right] - (n - b) \right)}$$

What formula?



Poisson-Poisson model with asymptotic formulae

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$

$$Z = \frac{s}{b}$$

Signal to noise/
background ratio

Z: significance
n: number of events
s: number of signal events
b: number of background events
 σ : total background uncertainty

Gaussian approximation w/ uncertainty

$$Z = \frac{s}{\sqrt{b + \sigma^2}}$$

$$Z = \frac{s}{\sqrt{s + b}}$$

Signal significance
used by TMVA

$$Z = \frac{s}{\sqrt{b}}$$

Gaussian approximation w/o uncertainty

Poisson-Poisson model with asymptotic
formulae w/o uncertainty

$$Z = \sqrt{2 \left(n \ln \left[\frac{n}{b} \right] - (n - b) \right)}$$

[ATL-COM-GEN-2018-026]

What formula?



Poisson-Poisson model with asymptotic formulae

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$

$$Z = \frac{s}{b}$$

Signal to noise/
background ratio

Z : significance

n : number of events

s : number of signal events

b : number of background events

σ : total background uncertainty

Gaussian approximation w/ uncertainty

$$Z = \frac{s}{\sqrt{b + \sigma^2}}$$

$$Z = \frac{s}{\sqrt{s + b}}$$

Signal significance
used by TMVA

$$Z = \frac{s}{\sqrt{b}}$$

Gaussian approximation w/o uncertainty

Poisson-Poisson model with asymptotic
formulae w/o uncertainty

$$Z = \sqrt{2 \left(n \ln \left[\frac{n}{b} \right] - (n - b) \right)}$$

[ATL-COM-GEN-2018-026]

Above 5σ ?



	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 ± 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 ± 0.03
Signal	6.57 ± 0.21
Non-prompt Photons	0.49 ± 0.09
Misidentified Leptons	0.062 ± 0.014
Pile up	0.056 ± 0.014
$Z \rightarrow \tau\tau$ contribution	0.035 ± 0.006
Total Background	0.64 ± 0.09
N_{tot}	7.21 ± 0.23

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$

$$Z = 4.58 \sigma$$

How to make an observation



Pick an
appropriate
final state



Optimise the
event selection

Estimate the
irreducible
backgrounds

Check if you
are above 5σ

Repeat 

Above 5σ ?



	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 ± 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 ± 0.03
Signal	6.57 ± 0.21
Non-prompt Photons	0.49 ± 0.09
Misidentified Leptons	0.062 ± 0.014
Pile up	0.056 ± 0.014
$Z \rightarrow \tau\tau$ contribution	0.035 ± 0.006
Total Background	0.64 ± 0.09
N_{tot}	7.21 ± 0.23

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$

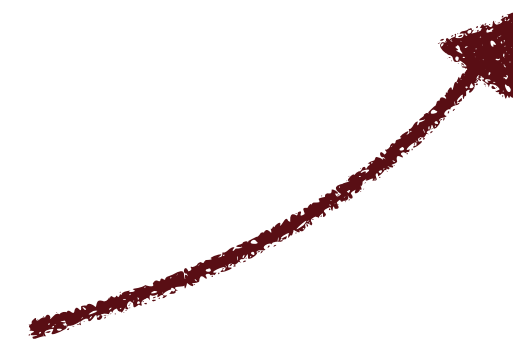
$$Z = 4.58 \sigma$$

Above 5σ ?



	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 ± 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 ± 0.03
Signal	6.57 ± 0.21
Non-prompt Photons	0.49 ± 0.09
Misidentified Leptons	0.062 ± 0.014
Pile up	0.056 ± 0.014
$Z \rightarrow \tau\tau$ contribution	0.035 ± 0.006
Total Background	0.64 ± 0.09
N_{tot}	7.21 ± 0.23

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}$$



$$Z = 5\sigma$$





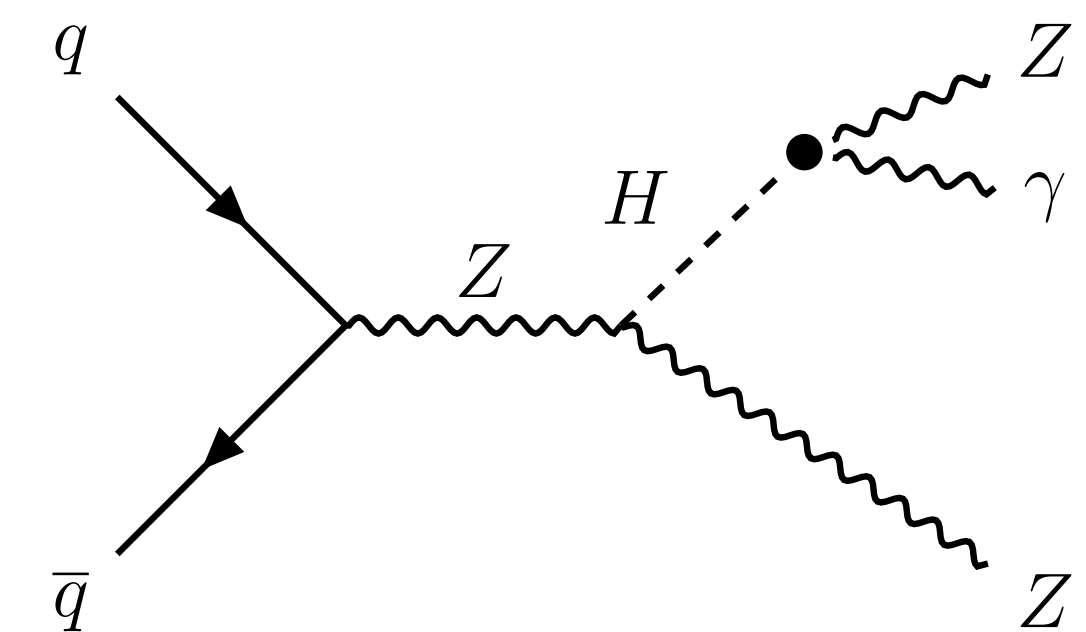
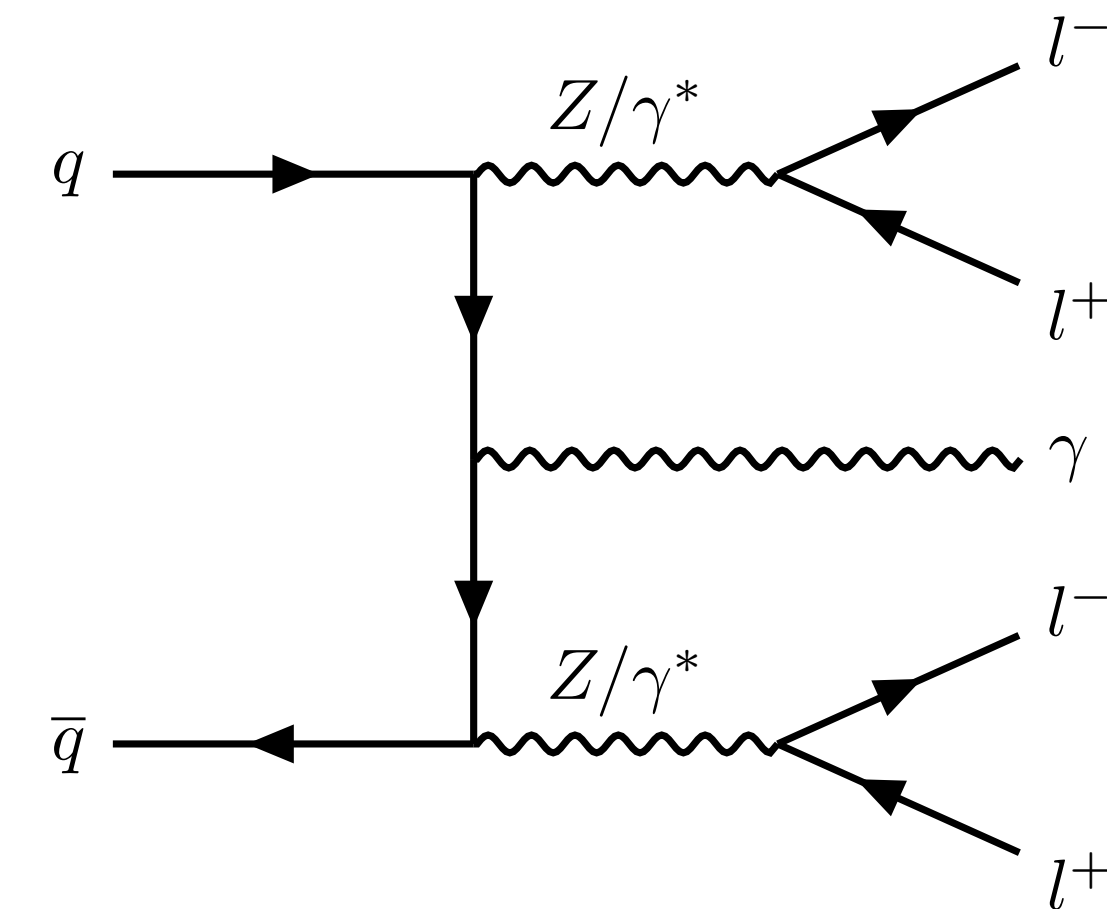
Backup

Analysis strategy



- fully leptonic decay: $Z \rightarrow ee, \mu\mu$
- signal: $4\ell + \gamma$
- $pp \rightarrow ZZ\gamma \rightarrow 4\ell + \gamma$ (incl. $pp \rightarrow ZH(\rightarrow Z\gamma) \rightarrow 4\ell + \gamma$)
- analysis of full Run 2 data (140 fb^{-1})

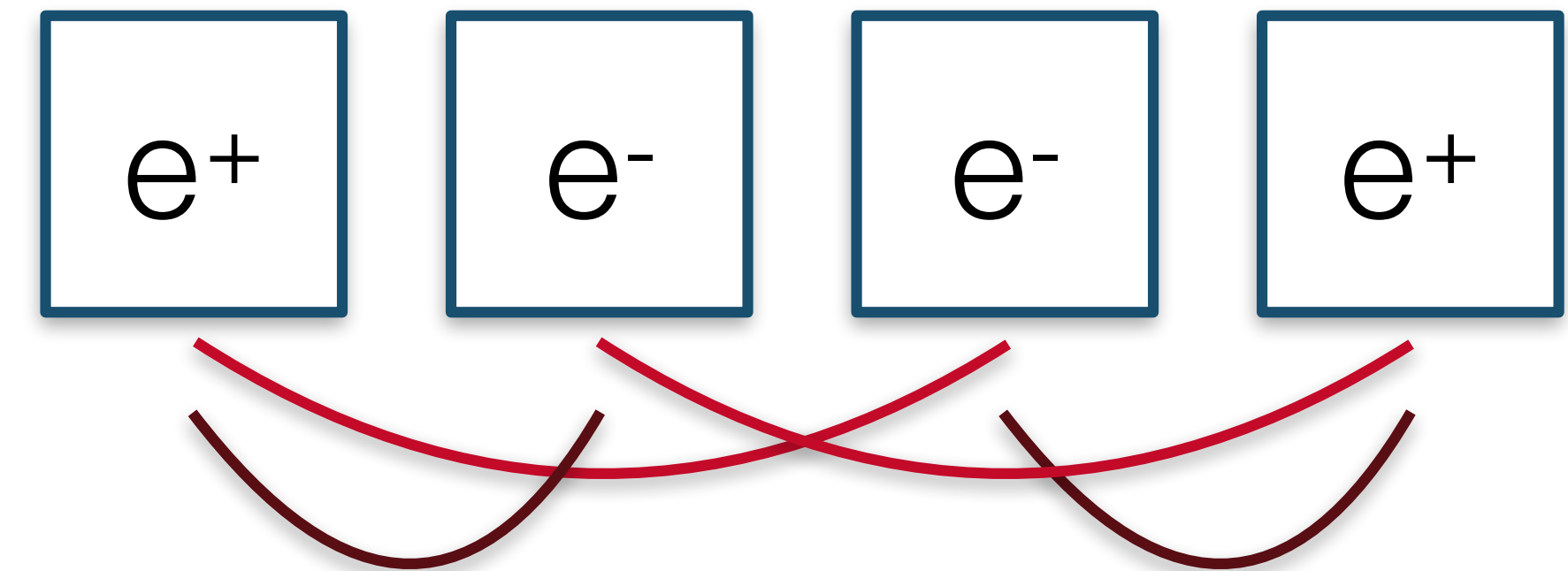
ATLAS Work in Progress	Number of Events
SHERPA 2.2.11 $llll\gamma$	5.83 ± 0.21
POWEGPYTHIA8 $ZH(\rightarrow Z\gamma)$	0.74 ± 0.03
Signal	6.57 ± 0.21



Event selection



- 2 pairs of 2 opposite-sign same-flavour (OSSF) leptons $\rightarrow 4e, 4\mu, 2e2\mu$
- invariant mass $m_{\ell\ell} > 40 \text{ GeV}$
- $\min(|m_{\ell\ell,1} - m_Z| + |m_{\ell\ell,2} - m_Z|)$
- 1 photon, $p_T > 20 \text{ GeV}$
- FSR rejection



Event selection



	Electrons	Muons	Photon
Kinematics	$p_T > 30, 25, 10, 10\text{GeV}$ $ \eta < 2.47$ excl. $1.37 < \eta < 1.52$	$p_T > 30, 25, 10, 10\text{GeV}$ $ \eta < 2.5$	$p_T > 20\text{GeV}$ $ \eta < 2.37$ excl. $1.37 < \eta < 1.52$
Identification	tight (l_1), medium	medium	tight
Isolation	Loose	PflowTightFixedRad (l_1), PflowLooseFixedRad	FixedCutLoose
Multiplicity		≥ 2 OSSF pairs	$\Delta R(l, \gamma) > 0.4$ ≥ 1 photon
Mass	$m_{ll} > 40 \text{ GeV}, (m_{ll\gamma} + m_{ll}) > 2 \cdot m_Z$		
Trigger	single lepton triggers		

Signal Systematics



Source	Relative uncertainty [%]
Photon identification efficiency	1.2
Photon isolation efficiency	1.0
Electron-Photon energy resolution	-0.19/ + 0.14
Electron-Photon energy scale	0.4
Electron identification efficiency	2.1
Electron isolation efficiency	0.11
Electron reconstruction efficiency	0.5
Muon isolation efficiency	0.6
Muon reconstruction efficiency	0.5
Muon trigger efficiency	0.15
Pile-up reweighting	-1.4/+1.6
Monte Carlo signal statistics	1.3
Theoretical uncertainties	x.xx
Integrated luminosity	0.83
Total systematic uncertainty signal	3.6

Table 1: Systematic uncertainty sources that contribute less than 0.1% are not shown. Theoretical uncertainties are not yet included.

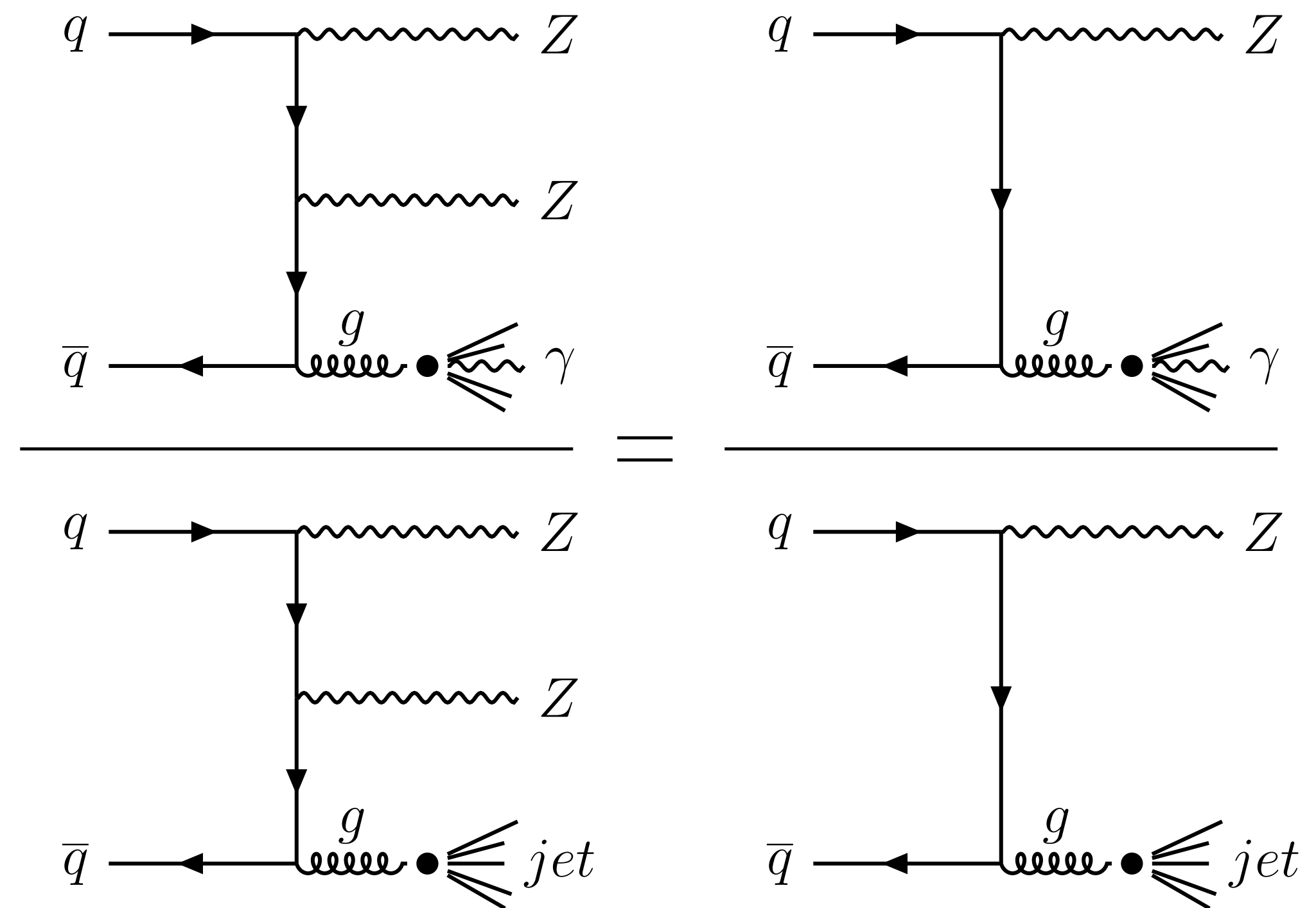
Non-prompt photon estimation



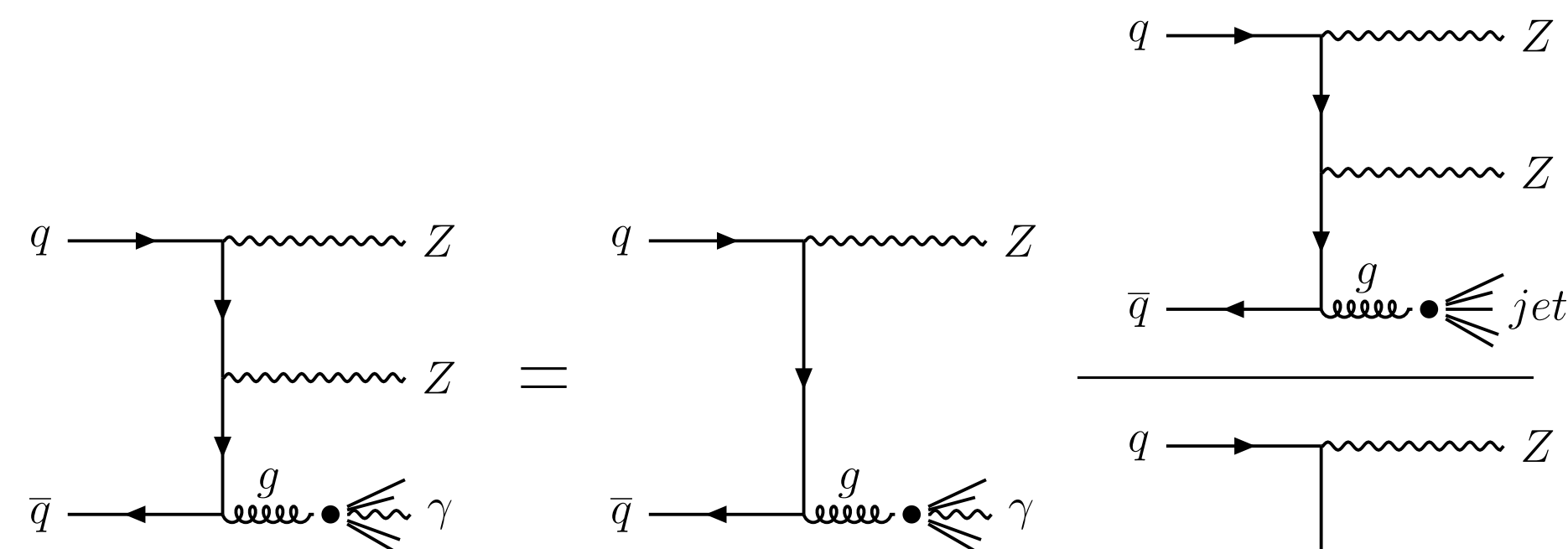
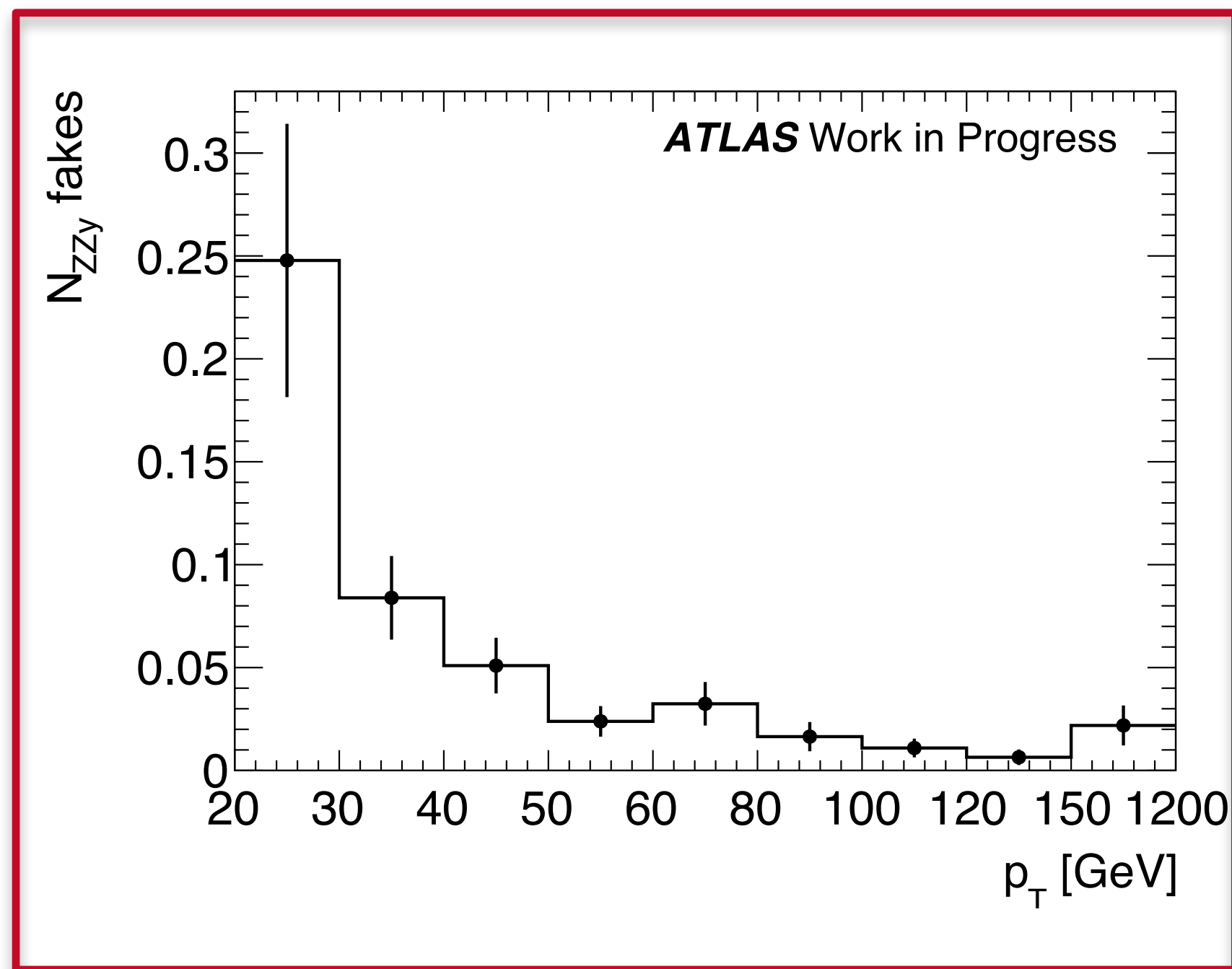
Ansatz:

The ratio of non-prompt photons produced in jets is independent of the simultaneous produced particles.

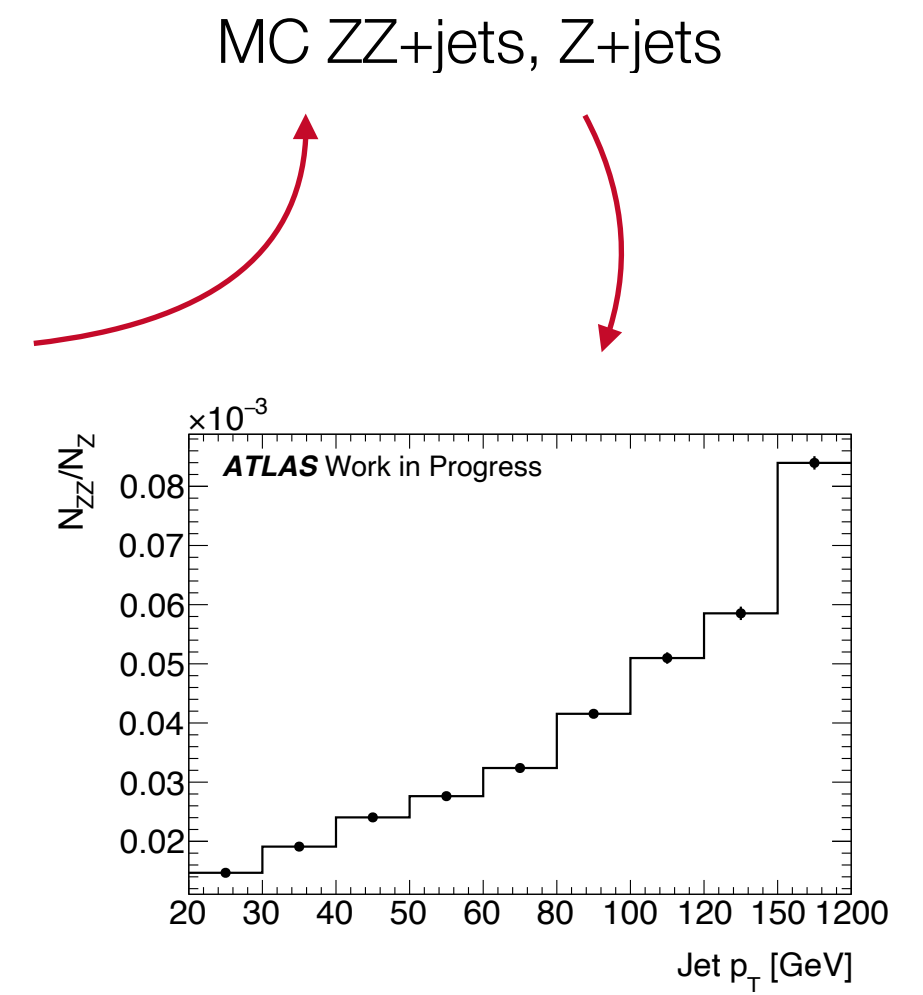
$$\frac{N_{ZZ\gamma}^{\gamma \text{ non-prompt}}}{N_{ZZ}^{\text{jets}}} = \frac{N_{Z\gamma}^{\gamma \text{ non-prompt}}}{N_Z^{\text{jets}}}$$



Non-prompt photon estimation



Estimation from Zγ Analysis



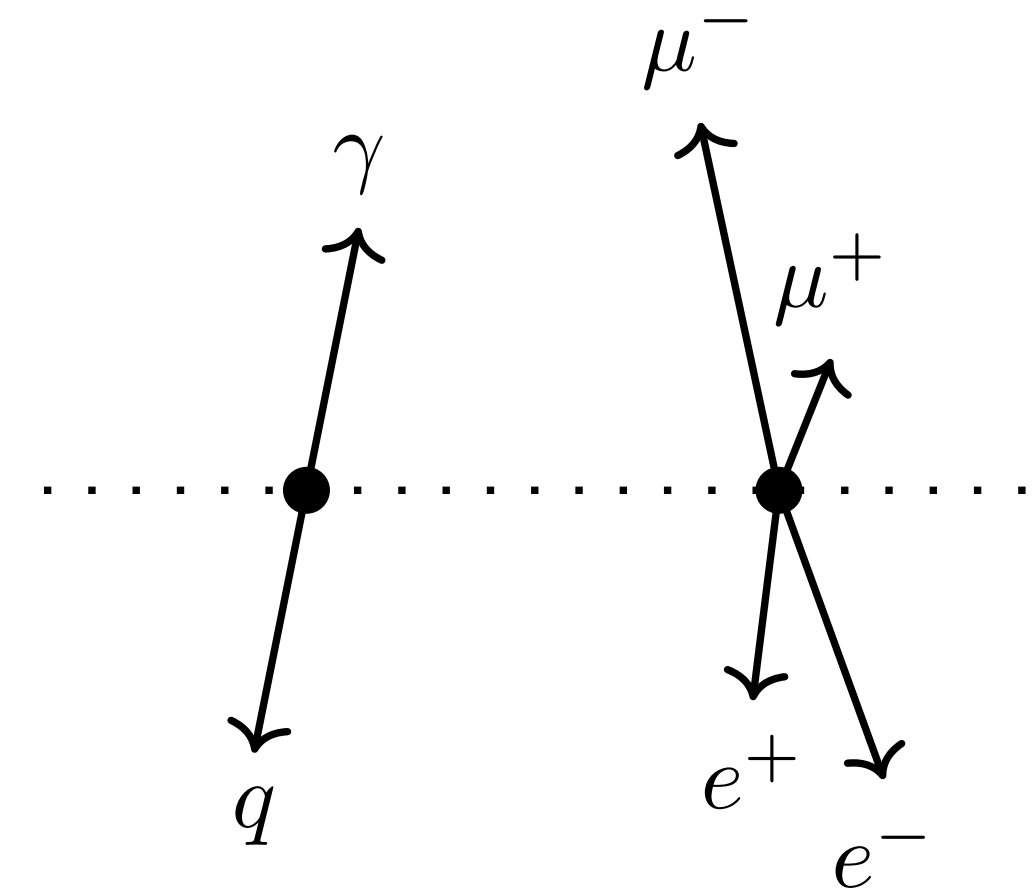
$$N_{ZZ\gamma}^{\gamma \text{ fake}} = 0.49 \pm 0.003 \text{ (stat.)} \pm 0.09 \text{ (syst.)}$$

Pile-up



Two independent interactions in the same bunch crossing

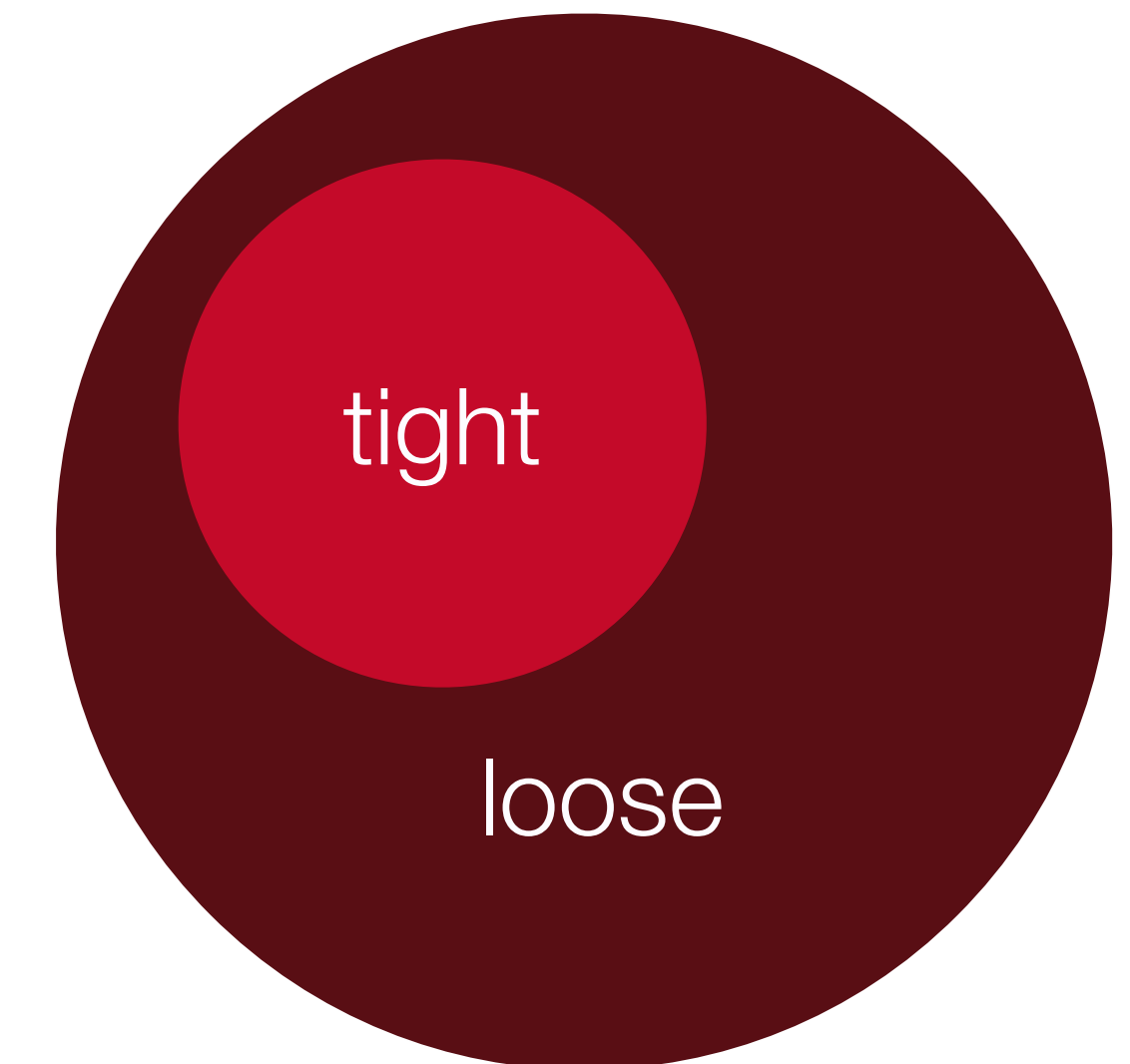
- Pile up contribution: $ZZ + \gamma$
 - ↳ photon cannot be associated to any vertex
- Ansatz: $N_{ZZ+\gamma} \sim P_{ZZ} P_{\gamma}$
- Determine pile up on particle level with MC overlay ($ZZ+\gamma$)
- Apply correction to derive estimate on reconstruction level



Matrix method



- baseline selection (**loose**) > signal selection (**tight**)
- real efficiency: $\frac{\text{prompt tight leptons}}{\text{prompt loose leptons}}$
 - ↳ derive with MC
- fake efficiency: $\frac{\text{fake tight leptons}}{\text{fake loose leptons}}$
 - ↳ derive in fake enriched CR with data (Z CR)

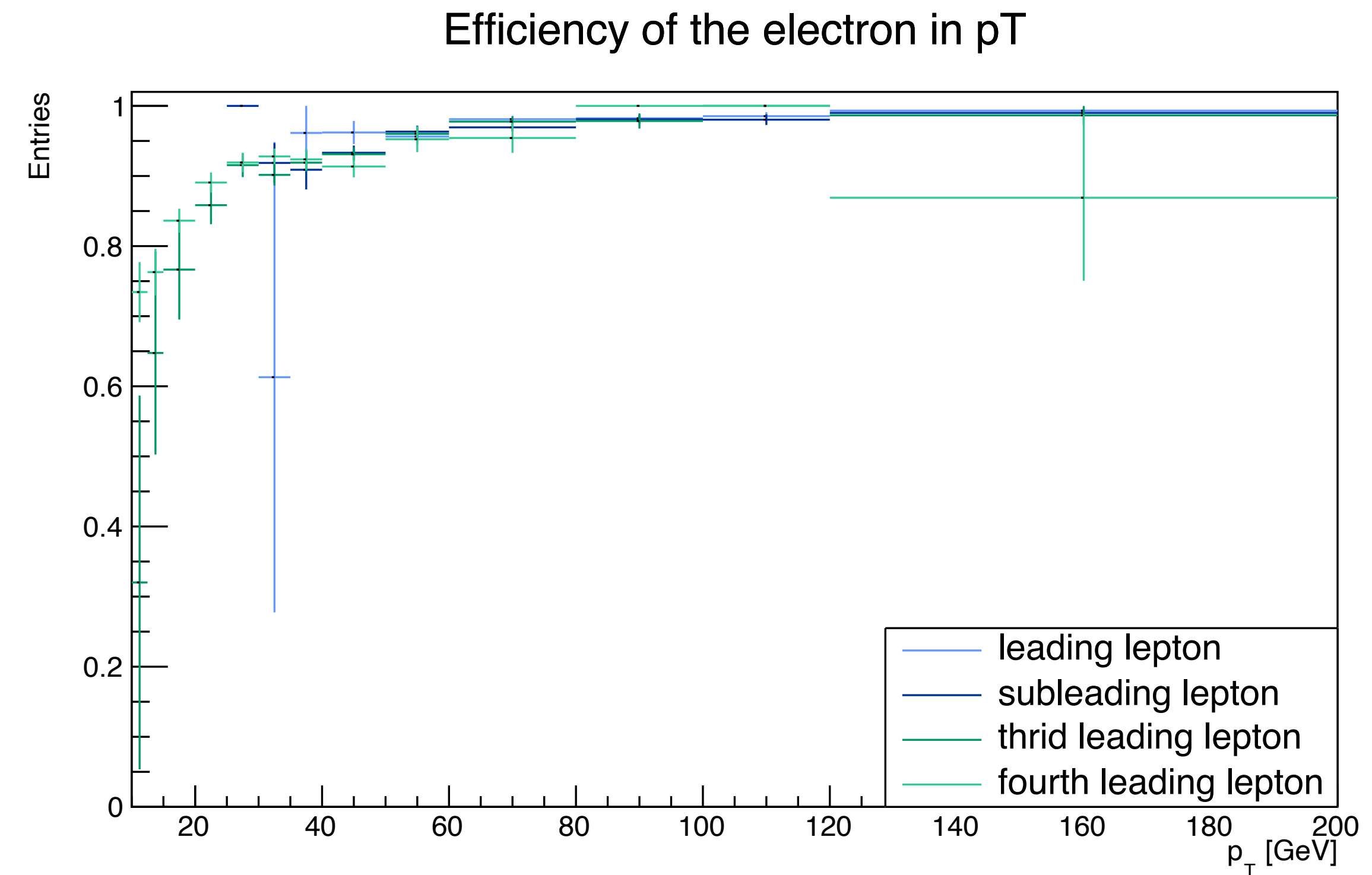


- loose leptons:
 - electrons: loose ID, no iso
 - muons: fail tight d0 significance, no iso
- tight leptons:
 - electrons: medium ID, loose iso
 - muons: tight d0 significance, loose iso

Real efficiencies



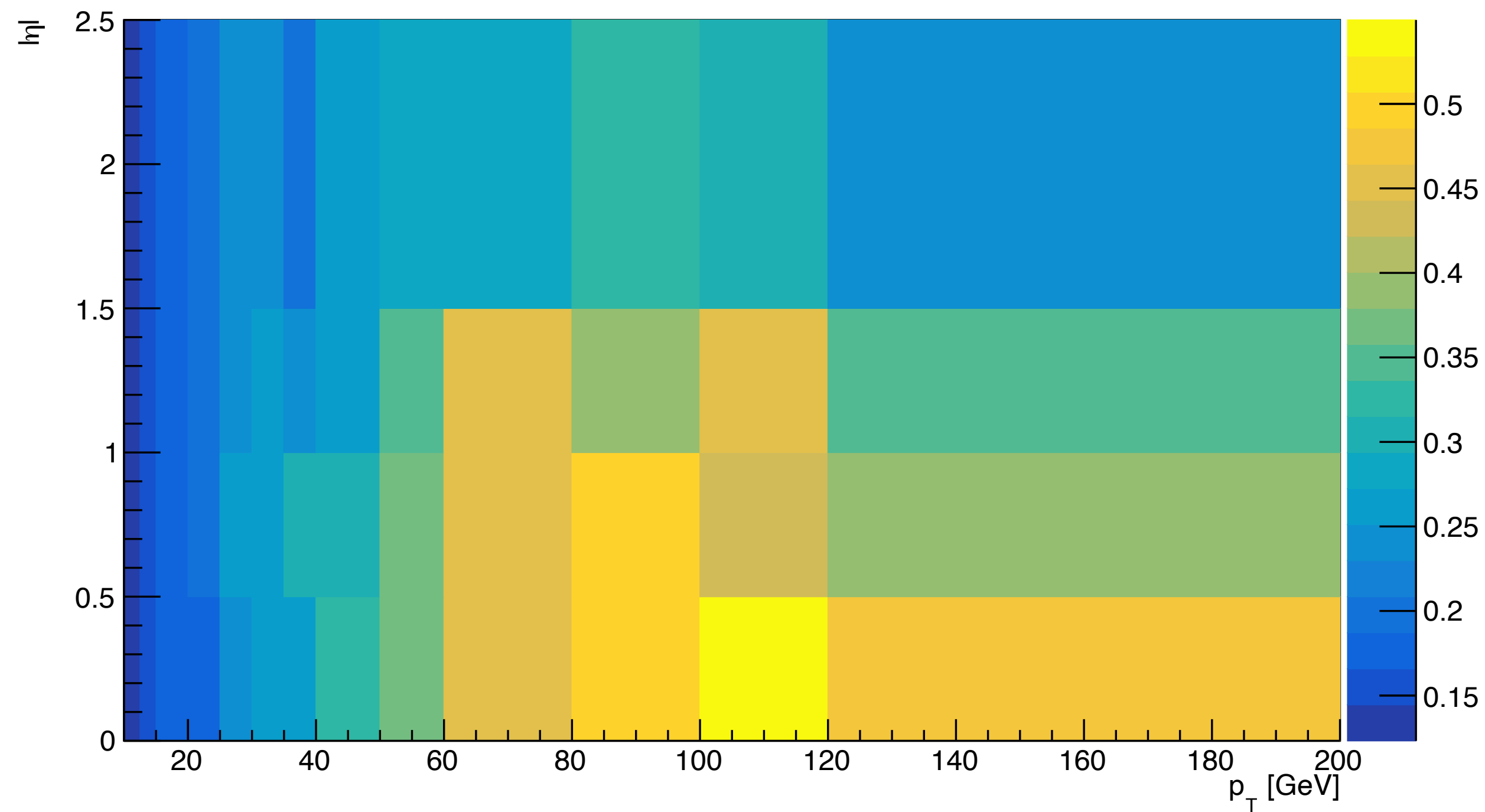
- $ZZ\gamma$ SR \rightarrow MC
- ansatz: efficiency is independent of lepton (leading, subleading, etc.)
- comparison of lepton efficiencies in MC
- small differences \rightarrow features of lepton p_T distribution



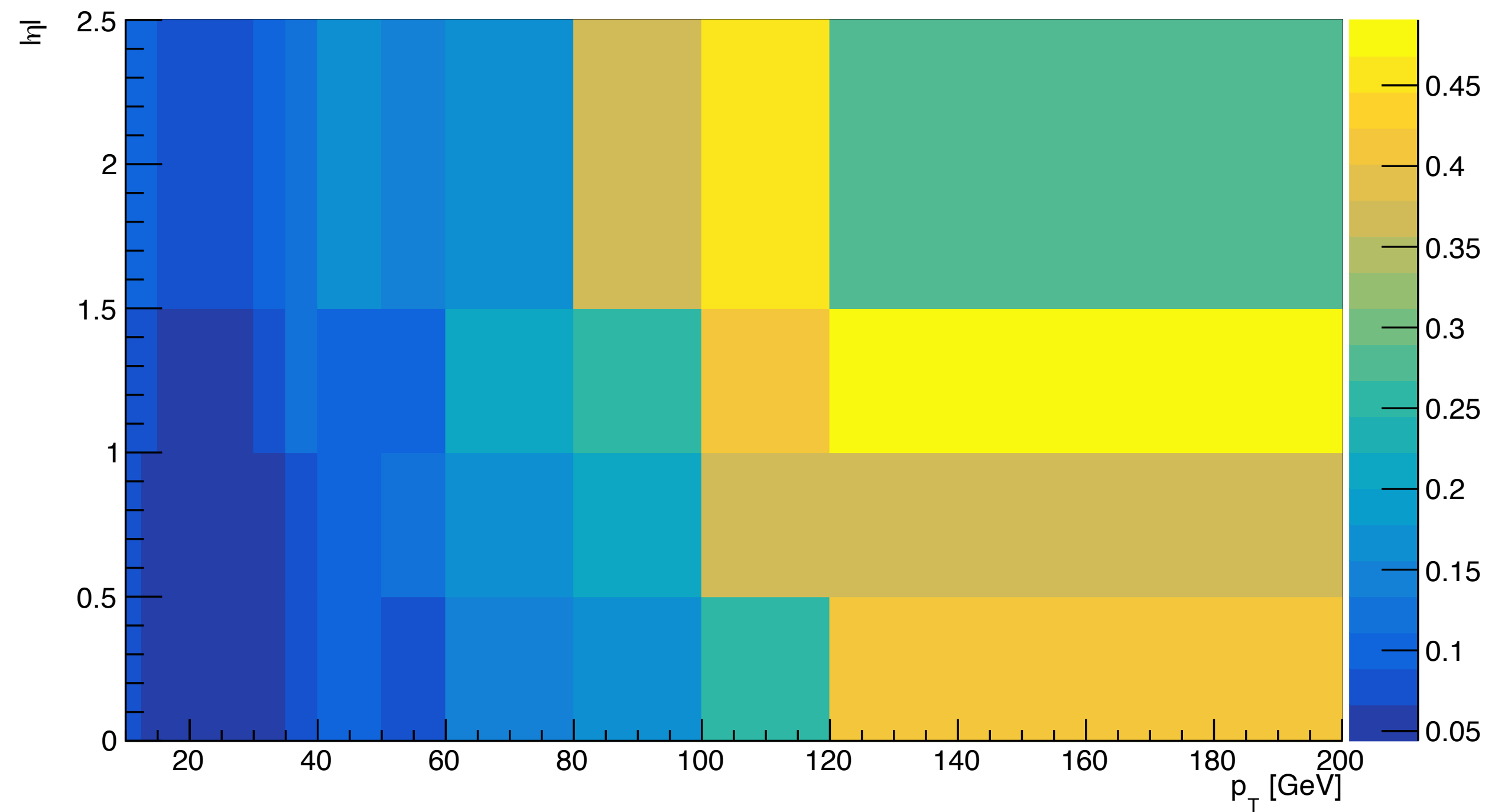
Fake efficiencies



Efficiency of the electrons



Efficiency of the muons



Misidentified leptons - validation



- VR1: SR with 1 OSSF pair & 1SSSF pair
 - ZZy SR - data: 0 ± 0 ✓
 - ZZ SR - data: 29 ± 5 ✓
- VR2: SR with 1 OSSF pair & 1OSOF pair
 - ZZy SR - data: 0 ± 0 ✓ (w/o mass cut: 1 ± 1)
 - ZZ SR - data: 32 ± 6 ✓

$$N_{ZZ\gamma}^{\ell \text{ fake}} = 0.062 \pm 0.014$$

$$N_{ZZ}^{\ell \text{ fake}} = 33 \pm 7$$