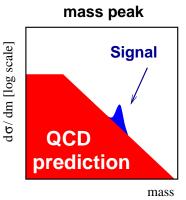
Giant K factors

Gavin Salam

CERN, Princeton & LPTHE/CNRS (Paris)

Work performed with Mathieu Rubin and Sebastian Sapeta, arXiv:1006.2144

IPN Lyon 25 March 2011

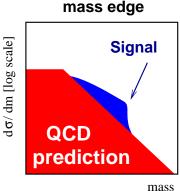


New resonance (e.g. Z') where you see all decay products and reconstruct an invariant mass

QCD may:

- swamp signal
- smear signal

leptonic case easy; hadronic case harder



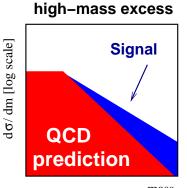
De New resonance (e.g. R-parity conserving SUSY), where undetected new stable par-

SUSY), where undetected new stable particle escapes detection.

Reconstruct only *part* of an invariant mass \rightarrow kinematic edge.

QCD may:

- swamp signal
- smear signal



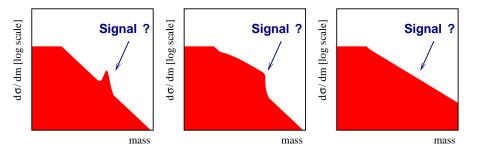
mass

Unreconstructed SUSY cascade. Study *effective* mass (sum of all transverse momenta).

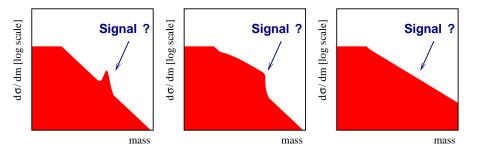
Broad excess at high mass scales.

Knowledge of backgrounds is crucial is declaring discovery.

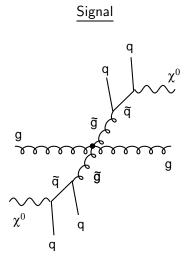
QCD is *one way* of getting handle on back-ground.

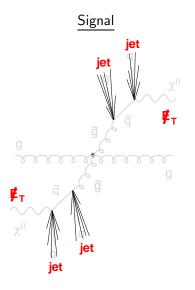


THIS TALK

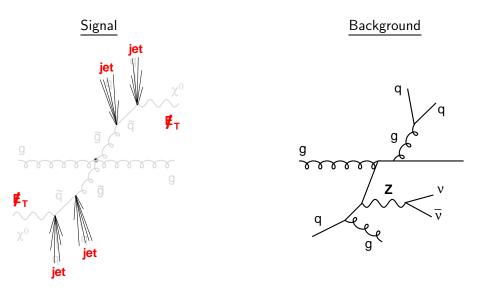


THIS TALK

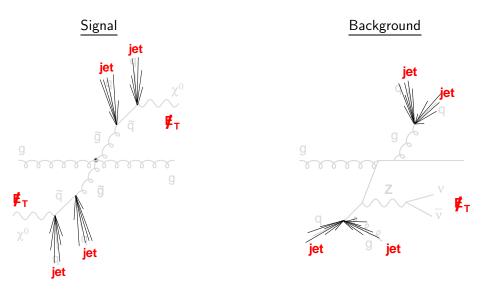




[Introduction]



[Introduction]

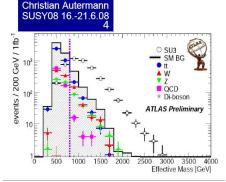


SUSY searches: what excesses?

Atlas selection [all hadronic]

- no lepton
- MET > 100 GeV
- 1^{st,}2nd jet > 100 GeV
- 3rd,4th jet > 50 GeV

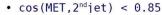
• MET / m _ > 20%

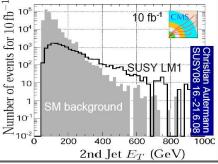


CMS selection [leptonic incl.]

(optimized for 10fb⁻¹, using genetic algorithm)

- 1 muon pT > 30 GeV
- MET > 130 GeV
- 1st, 2nd jet > 440 GeV
- 3rd jet > 50 GeV
- -0.95 < cos(MET,1stjet)<0.3

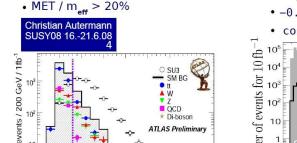




SUSY searches: what excesses?

Atlas selection [all hadronic]

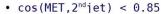
- no lepton
- MET > 100 GeV
- 1^{st,2nd} jet > 100 GeV
- 3rd,4th jet > 50 GeV

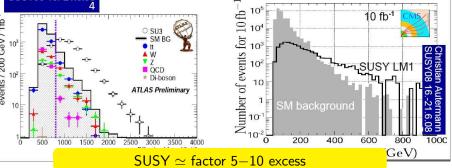


CMS selection [leptonic incl.]

(optimized for 10fb⁻¹, using genetic algorithm)

- 1 muon pT > 30 GeV
- MET > 130 GeV
- 1st, 2nd jet > 440 GeV
- 3rd jet > 50 GeV
- -0.95 < cos(MET,1stjet)<0.3



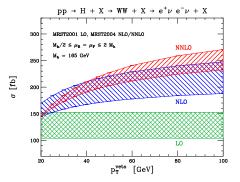


 $\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$ $\alpha_s \simeq 0.1$ That implies LO QCD (just c_0) should be accurate to within 10%

lt isn't

Rules of thumb: LO good to within factor of 2 NLO good to within scale uncertainty

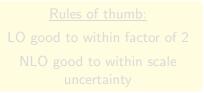
This talk is about an example where these rules fail spectacularly, the lessons we learn, and the solutions we can apply.



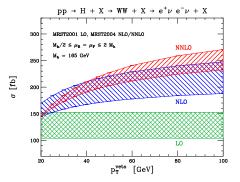
Anastasiou, Melnikov & Petriello '04 Anastasiou, Dissertori & Stöckli '07 $\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$ $\alpha_s \simeq 0.1$

That implies LO QCD (just c_0) should be accurate to within 10%

lt isn't



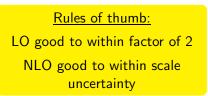
This talk is about an example where these rules fail spectacularly, the lessons we learn, and the solutions we can apply.



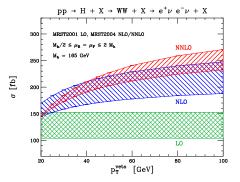
Anastasiou, Melnikov & Petriello '04 Anastasiou, Dissertori & Stöckli '07 $\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$ $\alpha_s \simeq 0.1$

That implies LO QCD (just c_0) should be accurate to within 10%

lt isn't



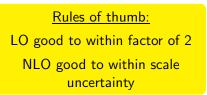
This talk is about an example where these rules fail spectacularly, the lessons we learn, and the solutions we can apply.



Anastasiou, Melnikov & Petriello '04 Anastasiou, Dissertori & Stöckli '07 $\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$ $\alpha_s \simeq 0.1$

That implies LO QCD (just c_0) should be accurate to within 10%

lt isn't



This talk is about an example where these rules fail spectacularly, the lessons we learn, and the solutions we can apply.

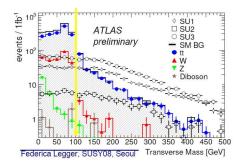
Control samples

We don't always have NLO for the background (e.g. Z+4 jets, a $2 \rightarrow 5$ process). Though amazing recent progress $2 \rightarrow 4$: Blackhat, Rocket, Helac-NLO, BDDP $2 \rightarrow 5$ (W+4j): Blackhat

Must then rely on LO (matched with parton showers). How does one verify it?

Common "data-driven" procedure: [roughly]

- Get control sample at low p_t
- SUSY should be small(er) contamination there
- Once validated, trust LO prediction at high-p_t



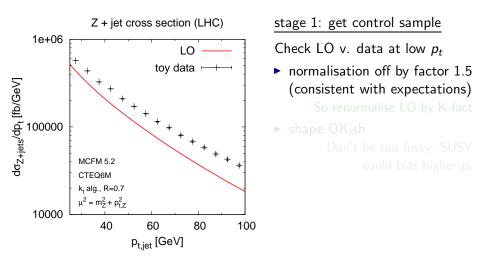
A conservative QCD theory point of view:

It's hard to be sure: since we can't (yet) calculate Z+4 jets beyond LO. But we would tend to think it is safe, as long as control data are within usual factor of two of LO prediction

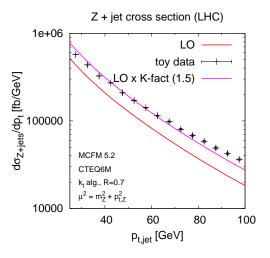
Illustrate issues with toy example: Z+jet production

- ► It's known to NLO and a candidate for "first" 2 → 2 NNLO $\sim e^+e^- \rightarrow \gamma^*/Z \rightarrow 3$ jets, NNLO: Gehrman et al '08, Weinzierl '08
- But let's pretend we only know it to LO, and look at the pt distribution of the hardest jet (no other cuts — keep it simple)

Toy data, control sample



Toy data, control sample



stage 1: get control sample

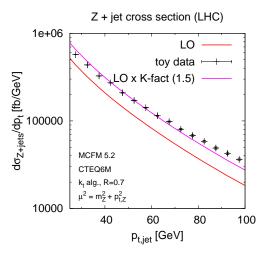
Check LO v. data at low p_t

 normalisation off by factor 1.5 (consistent with expectations) So renormalise LO by K-fact

shape OKish

Don't be too fussy: SUSY could bias higher p_t

Toy data, control sample

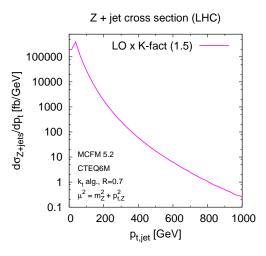


stage 1: get control sample

Check LO v. data at low p_t

- normalisation off by factor 1.5 (consistent with expectations) So renormalise LO by K-fact
- shape OKish

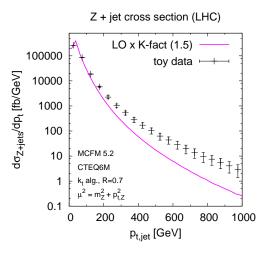
Don't be too fussy: SUSY could bias higher p_t



stage 2: look at high p_t

- good agreement at low pt, by construction
- excess of factor \sim 10 at high p_t
- check scale dependence of LO [NB: not always done except e.g. Alwall et al. 0706.2569] still big excess

Toy data, high p_t

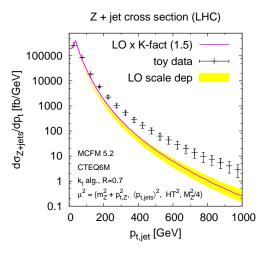


stage 2: look at high p_t

- ▶ good agreement at low pt, by construction
- excess of factor ~ 10 at high p_t

 check scale dependence of LO [NB: not always done except e.g. Alwall et al. 0706.2569] still big excess

Toy data, high p_t



stage 2: look at high p_t

- ▶ good agreement at low p_t, by construction
- excess of factor ~ 10 at high p_t
- check scale dependence of LO [NB: not always done except e.g. Alwall et al. 0706.2569] still big excess

ls it:

- QCD + extra signal?
- ▶ just QCD? But then where does a *K*-factor of 10 come from?

Here it's just a toy illustration. Later this year it may be for real:

▶ Do Nature / Science / PRL accept the paper?

Discovery of New Physics at the TeV scale We report a 5.7 σ excess in MET + jets production that is consistent with a signal of new physics ...

Do we proceed immediately with a linear collider? It'll take 10–15 years to build; the sooner we start the better

▶ At what energy? It would be a shame to be locked in to the wrong energy...

ls it:

- QCD + extra signal?
- ▶ just QCD? But then where does a *K*-factor of 10 come from?

Here it's just a toy illustration. Later this year it may be for real:

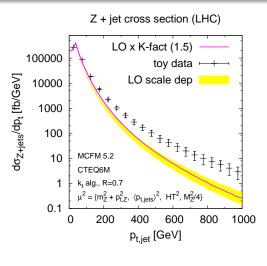
Do Nature / Science / PRL accept the paper?

Discovery of New Physics at the TeV scale We report a 5.7 σ excess in MET + jets production that is consistent with a signal of new physics ...

Do we proceed immediately with a linear collider? It'll take 10–15 years to build; the sooner we start the better

• At what energy? It would be a shame to be locked in to the wrong energy...

Open the box...



Unlike for SUSY multi-jet searches, in the Z+jet case we do have NLO.

Once NLO is included the excess disappears

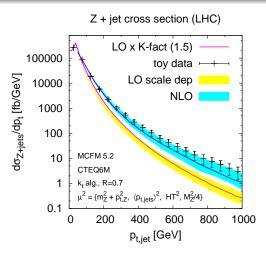
The "toy data" were just the upper edge of the NLO band

Example based on background work for Butterworth, Davison, Rubin & GPS '08

Related observations also by Bauer & Lange '09; Denner, Dittmaier, Kasprzik & Muck '09

Hold on a second: how does QCD give a K-factor $\mathcal{O}(5-10)$? NB: DYRAD, MCFM consistent

Open the box...



Unlike for SUSY multi-jet searches, in the Z+jet case we do have NLO.

Once NLO is included the excess disappears

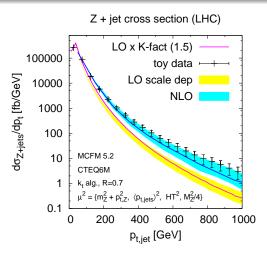
The "toy data" were just the upper edge of the NLO band

Example based on background work for Butterworth, Davison, Rubin & GPS '08

Related observations also by Bauer & Lange '09; Denner, Dittmaier, Kasprzik & Muck '09

Hold on a second: how does QCD give a K-factor $\mathcal{O}(5-10)$? NB: DYRAD, MCFM consistent

Open the box...



Unlike for SUSY multi-jet searches, in the Z+jet case we do have NLO.

Once NLO is included the excess disappears

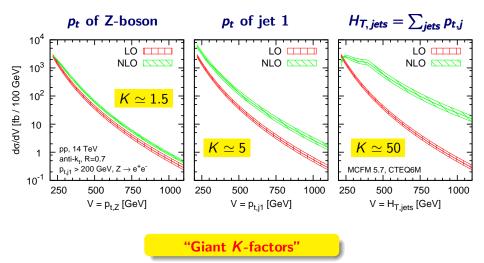
The "toy data" were just the upper edge of the NLO band

Example based on background work for Butterworth, Davison, Rubin & GPS '08

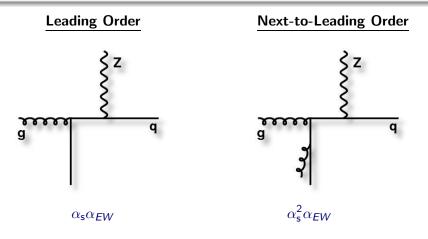
Related observations also by Bauer & Lange '09; Denner, Dittmaier, Kasprzik & Muck '09

Hold on a second: how does QCD give a K-factor O(5-10)? NB: DYRAD, MCFM consistent

What about other observables?

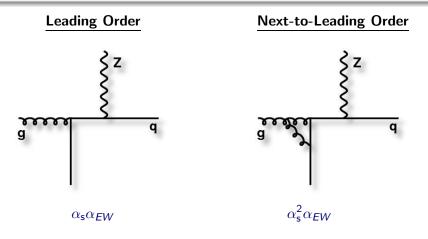


Why the large K-factor?



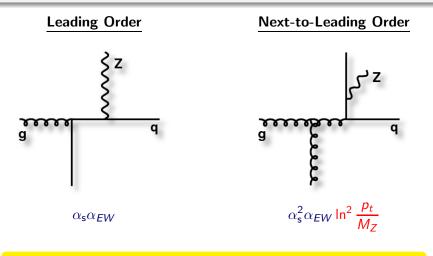
LHC will probe scales well above EW scale, $\sqrt{s} \gg M_Z$. EW bosons are light. New logarithmically enhanced topologies appear.

Why the large K-factor?



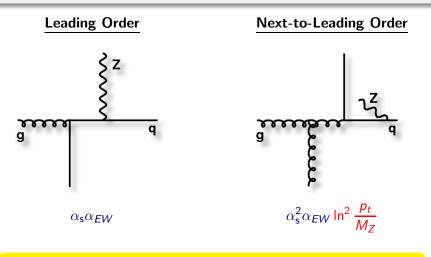
LHC will probe scales well above EW scale, $\sqrt{s} \gg M_Z$. EW bosons are light. New logarithmically enhanced topologies appear.

Why the large K-factor?



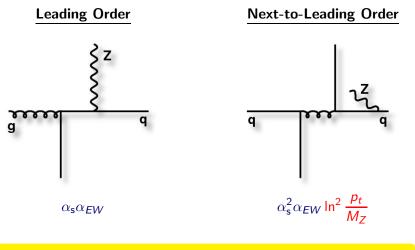
LHC will probe scales well above EW scale, $\sqrt{s} \gg M_Z$. EW bosons are **light**. New logarithmically enhanced topologies appear.

Why the large K-factor?



LHC will probe scales well above EW scale, $\sqrt{s} \gg M_Z$. EW bosons are **light**. New logarithmically enhanced topologies appear.

Why the large K-factor?



LHC will probe scales well above EW scale, $\sqrt{s} \gg M_Z$. EW bosons are **light**. New logarithmically enhanced topologies appear.

Z+jet @ LO \rightarrow NLO shows huge correction

What happens at NNLO? Despite 10 years' calculation, the answer is not yet known

Our strategy: get an approximation to NNLO

specifically, approximate the elements of the calculation that contribute only a small part of the giant K-factor, i.e. the loops.

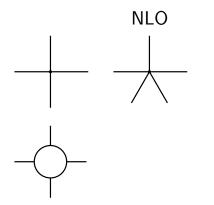
[An example]

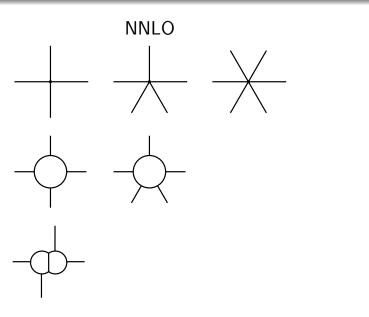
Ingredients of (N)NLO

LO

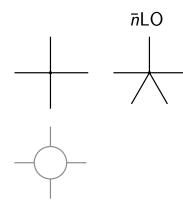
[An example]

Ingredients of (N)NLO





Approximations to (N)NLO



Our naming scheme:

For each loop that we approximate, replace N $\rightarrow \bar{n}$

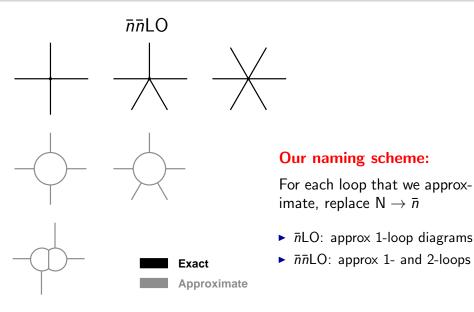
▶ *n*LO: approx 1-loop diagrams



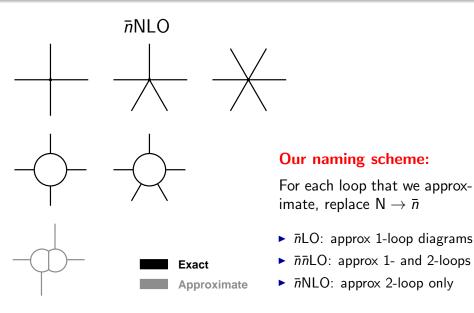


Approximate

Approximations to (N)NLO



Approximations to (N)NLO

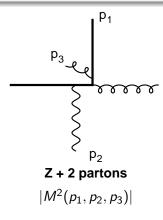


First try $Z + jet @ \bar{n}LO$:

Take the "leading" process [Z + jet @ LO]

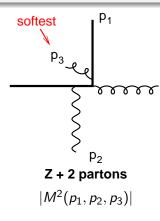
and add in process with one extra jet. [i.e. include Z + 2 jets @ LO]

The LoopSim idea: $\bar{n}LO$



- Identify softest or most collinear parton [with help of a jet algorithm]

The LoopSim idea: nLO

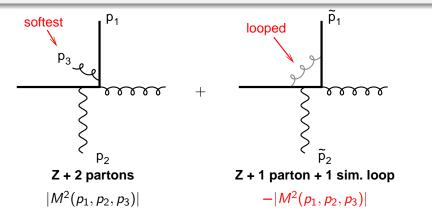


Identify softest or most collinear parton [with

[with help of a jet algorithm]

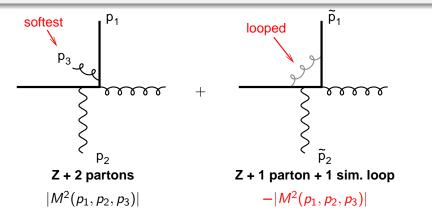
"Loop" it ≡ remove it from event, reshuffle other momenta; weight of looped event is (−1)× weight of tree-level event

The LoopSim idea: $\bar{n}LO$



- Identify softest or most collinear parton [with help of a jet algorithm]
- "Loop" it ≡ remove it from event, reshuffle other momenta; weight of looped event is (-1)× weight of tree-level event

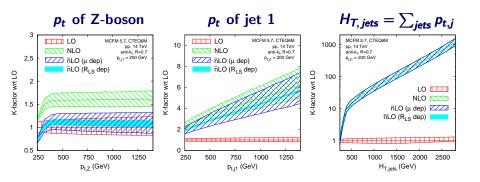
The LoopSim idea: nLO



- Identify softest or most collinear parton [with help of a jet algorithm]
- Loop" it ≡ remove it from event, reshuffle other momenta; weight of looped event is (-1)× weight of tree-level event

This cancels all the "single-unresolved" divergences in the Z+2 events

Giant K-factors (@ IPN Lyon)



When the K-factors are large, $\bar{n}LO$ agrees well with NLO

MLM matching also does a similar job cf. de Aquino, Hagiwara, Li & Maltoni '11

MLM/CKKW matching also effectively provide \bar{n} LO type accuracy

How dooes LoopSim compare to MLM/CKKW?

- 1. Does not rely on shower (✓: simplicity; ✗: not easily integrated with shower MCs)
- 2. Does not need arbitrary separation of Z+1/Z+2/etc. samples with (hard-to-choose) momentum cutoff

3. Can easily be extended beyond LO matching



LoopSim at $\bar{n}\bar{n}LO$ and $\bar{n}NLO$

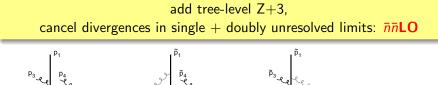
add tree-level Z+3, cancel divergences in single + doubly unresolved limits: $\overline{n}\overline{n}LO$



 $|M^2(p_1, p_2, p_3, p_4)|$



LoopSim at $\bar{n}\bar{n}$ LO and \bar{n} NLO

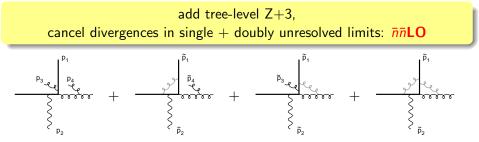


 $|M^2(p_1, p_2, p_3, p_4)| = -|M^2(p_1, p_2, p_3, p_4)| = -|M^2(p_1, p_2, p_3, p_4)|$

Separately loop either of the 2 softest emissions: provides approx of 1-loop

p2

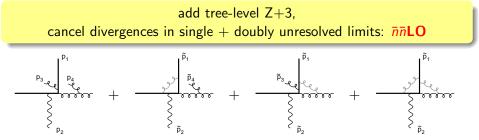




$$\begin{split} |M^2(p_1,p_2,p_3,p_4)| & -|M^2(p_1,p_2,p_3,p_4)| & -|M^2(p_1,p_2,p_3,p_4)| & +|M^2(p_1,p_2,p_3,p_4)| \\ \text{Simultaneously loop each of the 2 softest emissions: provides approx of 2-loop} \\ & \text{Total of tree plus approx 1- and 2-loop pieces gives zero} \end{split}$$



LoopSim at $\bar{n}\bar{n}$ LO and \bar{n} NLO



$$\begin{split} |M^2(p_1,p_2,p_3,p_4)| & -|M^2(p_1,p_2,p_3,p_4)| & -|M^2(p_1,p_2,p_3,p_4)| & +|M^2(p_1,p_2,p_3,p_4)| \\ \text{Simultaneously loop each of the 2 softest emissions: provides approx of 2-loop} \\ & \text{Total of tree plus approx 1- and 2-loop pieces gives zero} \end{split}$$

add in (exact Z+2 @ 1-loop) – (approximate Z+2 @ 1-loop) + extra simulated 2-loop piece to cancel new Z+2@1-loop divergences This is $\bar{n}NLO$

[LoopSim] L[nNLO]

The 2-loop piece has the topology of the LO diagram.

The "mistake" we make in approximating it should therefore be a "pure" $\mathcal{O}(\alpha_s^2)$ correction, without any large enhancements from new NLO type topologies.

$$\sigma_{\bar{n}\mathsf{NLO}} = \sigma_{\mathsf{NNLO}} + \mathcal{O}\left(\alpha_{\mathsf{s}}^{2}\sigma_{\mathsf{LO}}\right)$$
$$= \sigma_{\mathsf{NNLO}}\left(1 + \mathcal{O}\left(\frac{\alpha_{\mathsf{s}}^{2}}{\mathsf{K}_{\mathsf{NNLO}}}\right)\right)$$

 $K_{
m NNLO} = rac{\sigma_{
m NNLO}}{\sigma_{
m LO}} \sim K_{
m NLO} \gg 1$

The *relative* contribution of the neglected piece is suppressed by the large K-factor.

 \bar{n} NLO should be a good approximation to NNLO when the K-factor is large and due to new higher-order topologies.

Testing \bar{n} **NLO, in 3 processes** [making use of existing NLO calculations]

1. Z@NLO and Z+j@NLO \rightarrow Z@ \bar{n} NLO

with MCFM; compare to exact NNLO from DYNNLO

2. Z+j@NLO and Z+2j@NLO \rightarrow Z+j@ \bar{n} NLO

with MCFM

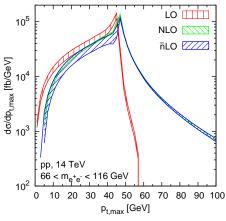
3. 2j@NLO and 3j@NLO \rightarrow 2j@ \bar{n} NLO

with NLOjet++

[*n*NLO results] └[Drell-Yan]

Validation: Drell-Yan lepton p_t , \bar{n} NLO v. NNLO

nLO v. NLO



Z (i.e. DY) with Z+j from MCFM & LoopSim

For $p_{t\,ell} \gtrsim \frac{1}{2}M_Z + \Gamma_Z$ (giant *K*-factor!) it had to work For $p_{t,\ell} \lesssim \frac{1}{2}M_Z + \Gamma_Z$ it's remarkable that it still works

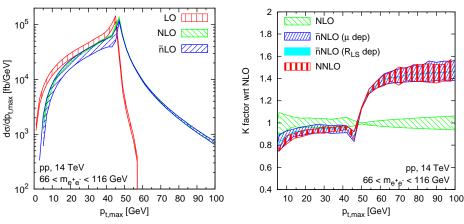
Validation: Drell-Yan lepton p_t , \bar{n} NLO v. NNLO

*n***NLO v. NNLO**

nLO v. NLO

[*n*NLO results]

[Drell-Yan]



NNLO from DYNNLO, Z (i.e. DY) with Z+j from MCFM & LoopSim

For $p_{t\,ell} \gtrsim \frac{1}{2}M_Z + \Gamma_Z$ (giant *K*-factor!) it had to work For $p_{t,\ell} \lesssim \frac{1}{2}M_Z + \Gamma_Z$ it's remarkable that it still works

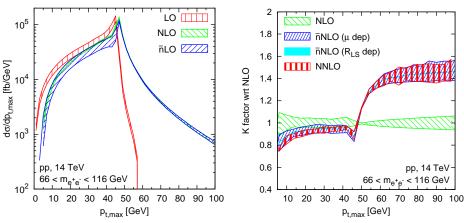
Validation: Drell-Yan lepton p_t , \bar{n} NLO v. NNLO

*n***NLO v. NNLO**

nLO v. NLO

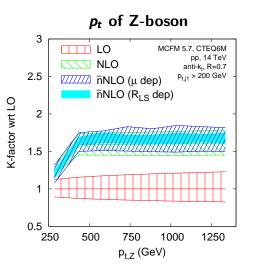
[*n*NLO results]

[Drell-Yan]

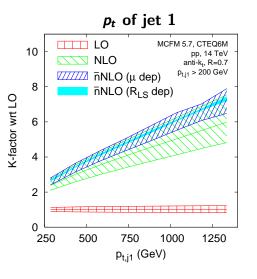


NNLO from DYNNLO, Z (i.e. DY) with Z+j from MCFM & LoopSim

For $p_{t\,ell} \gtrsim \frac{1}{2}M_Z + \Gamma_Z$ (giant K-factor!) it had to work For $p_{t,\ell} \lesssim \frac{1}{2}M_Z + \Gamma_Z$ it's remarkable that it still works

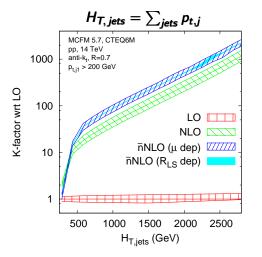


- *p_{tZ}* distribution didn't have giant *K*-factors.
- *n*NLO brings no benefit
 To get improvement you would need exact 2-loop terms



- *p_{tj}* distribution seems to converge at *n*NLO
- \blacktriangleright scale uncertainties reduced by \sim factor 2

 $[\bar{n}NLO \text{ results}]$ [Z+jet]



- ► Significant further enhancement for H_{T,jets}
- ► *n*NLO brings clear message:

H_{T,jets} is not a good observable!

[*n*NLO results] └[Z+jet]

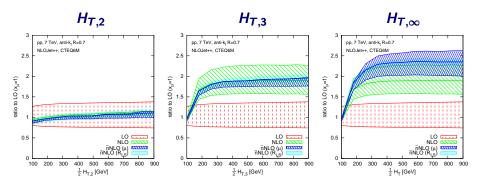
H_T (effective mass) type observables are widely used in searches

- H_T has a steeply falling distribution (like p_{tj} , p_{tZ})
- ► At each order (NLO, NNLO), an extra (soft) jet contributes to the H_T sum
 e.g. from ISR
- ► That shifts *H*_T up, which translates to a substantial increase in the cross section

We can test this hypothesis for plain jet events, using a truncated sum,

$$H_{T,n} = \sum_{i=1}^{n} p_{t, \text{jet } i}$$

 $H_{T,n}$ in (di)jet events



A clear message: for a process with *n* objects at lowest order, use $H_{T,n}$

Do you know what gets used in your experiment's searches? Many writers of LHC SUSY proceedings didn't...

Be aware that giant K-factors exist

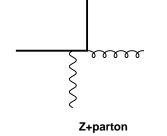
Always look one order beyond the leading order, for example with $$\rm MLM/CKKW$$ matching

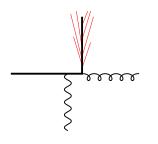
New tool to get good predictions in such cases: LoopSim Basically an "operator" to generate approximations to unknown loops Combine Z+j@NLO, Z+2j@NLO to get " \bar{n} NLO" Z+jet It sometimes works even beyond "giant-K-factor" regions

Watch out for H_T

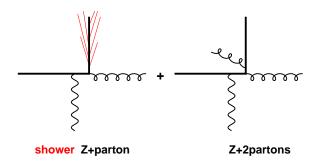
Even for simple processes, it converges very poorly unless you define it carefully (limit number of objects in sum)

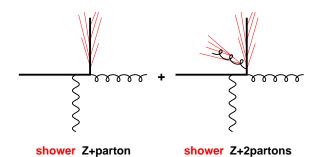
EXTRAS

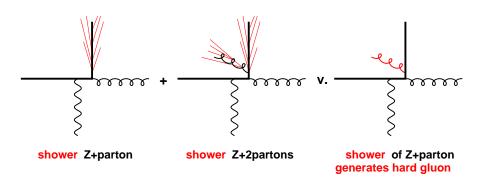


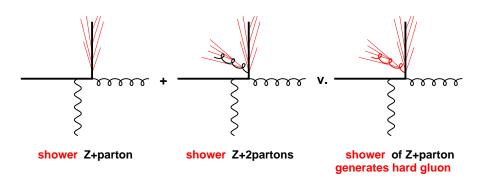


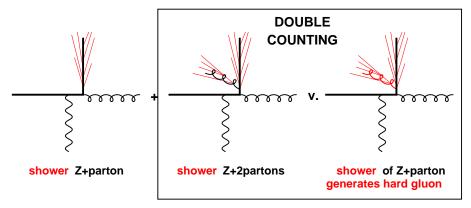
shower Z+parton







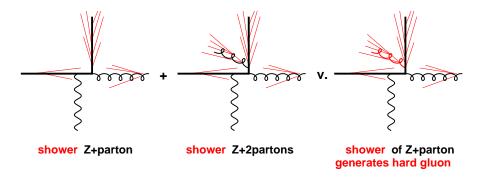




 $Z + {\rm parton\ implicitly\ includes\ part\ of\ } Z + 2 \ {\rm partons\ }$ It's just that the 2nd parton isn't always explicitly "visible"

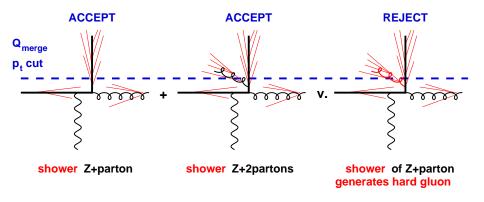
cartoon of MLM merging of Z+j and Z+2j

[Extras] └[How MLM works]



 MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.

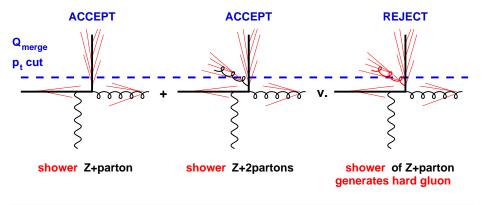
Our aim is to do that without the parton shower



 MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.

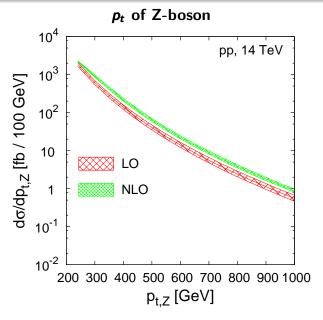
Our aim is to do that without the parton shower



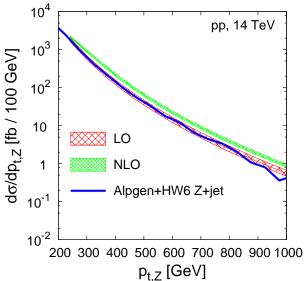


 MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.

Our aim is to do that without the parton shower



 p_t of Z-boson



pt of Z-boson 10⁴ pp, 14 TeV 10³ do/dp_{t,Z} [fb / 100 GeV] 10² 10 LO NLO 1 Alpgen+HW6 Z+jet 10⁻¹ Alpgen+HW6 Z+jet/Z+2jets 10⁻² 200 300 400 500 600 700 800 900 1000 p_{t.Z} [GeV]

All predictions similar and stable

[Extras] └[MLM results]

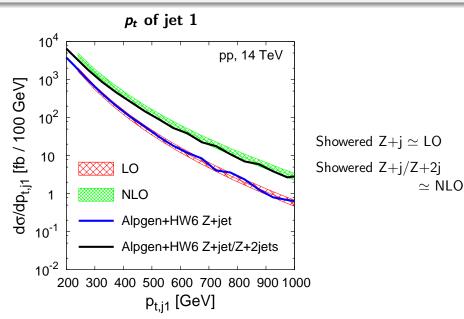
 p_t of jet 1 10⁴ pp, 14 TeV 10³ dơ/dp_{t,j1} [fb / 100 GeV] 10² 10 LO \boxtimes NLO 1 10⁻¹ 10⁻² 200 300 400 500 600 700 800 900 1000 p_{t,j1} [GeV]



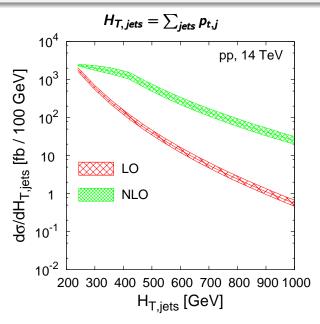
 p_t of jet 1 10⁴ pp, 14 TeV 10³ dơ/dp_{t,j1} [fb / 100 GeV] 10² 10 LO NLO 1 Alpgen+HW6 Z+jet 10⁻¹ 10⁻² 200 300 400 500 600 700 800 900 1000 p_{t,i1} [GeV]

Showered Z+j \simeq LO

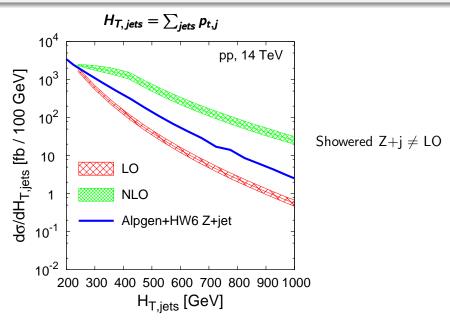




[Extras] └[MLM results]

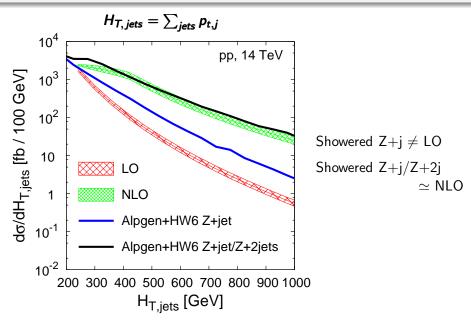


[E×tras] └[MLM results]



[E×tras] └[MLM results]

Testing Alpgen + Herwig + MLM Matching

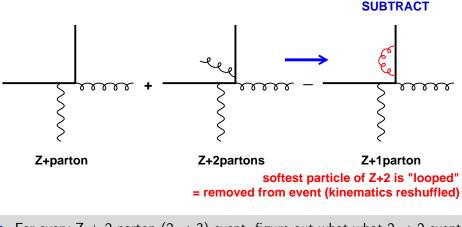


cartoon of the LoopSim idea

2011-02-02

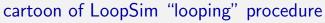
33 / 28

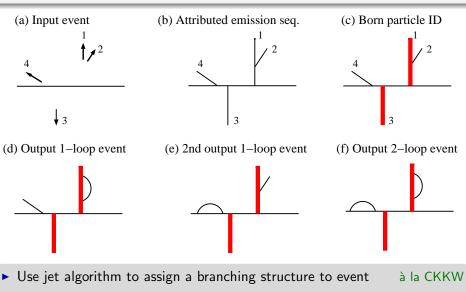




For every Z + 2 parton (2 → 3) event, figure out what what 2 → 2 event it would really have come from "Loop" the softest parton [Don't actually explicitly calculate any loop diagrams: simulate the loops]
 Subtract that 2 → 2 event Unlike MLM, no cutoffs on 2 → 3 events If done properly, divergences will cancel

Giant K-factors (@ IPN Lyon)





The particles that are softest are the ones that will be "looped"

[Extras]

└ [LoopSim details]

[Extras] └[LoopSim details] *n*LO accuracy

Define operators:

 $U_{\ell}(\text{event E}) \equiv \text{all simulated } \ell\text{-loop events from E}$

$$U_{orall}(ext{event}) \equiv \sum_{\ell=0} U_\ell(ext{event})$$

"U" stands for unitarisation (cancellation of all divergences) sum of all diagrams (essentially) adds up to zero

To combine Z+j with Z+2j take

 $Z+j@\bar{n}LO \equiv Z+j@LO + U_{\forall}(Z+2j@LO)$

we use " $\bar{n}LO$ " to emphasize that this is a crude approximation to an actual NLO calculation — the exact loops are missing NB: U_{\forall} here includes $\ell = 0, 1$

\bar{n} NLO: merging Z+j and Z+2j, both @NLO

Just replace simulated loops with exact loops Apply LoopSim to exact 1-loop to get (e.g.) simulated 2-loop terms

$$\begin{split} E_{n,\ell} &\equiv \text{event with } n \text{ partons and } \ell \text{ exact loops} \\ U_{\forall,\ell} &\equiv \text{operator to apply when } \ell \text{ exact loops known} \\ U_{\forall,1}(E_{n,0}) &= U_{\forall}(E_{n,0}) - U_{\forall}(U_1(E_{n,0})) \\ U_{\forall,1}(E_{n,1}) &= U_{\forall}(E_{n,1}) \end{split}$$

 $\mathsf{Z}+\mathsf{j}@\bar{n}\mathsf{NLO}=\mathsf{Z}+\mathsf{j}@\mathsf{NLO}+\mathit{U}_{\forall,1}(\mathsf{Z}+2\mathsf{j}@\mathsf{NLO}_{\mathsf{only}})$

Extension to NLO, NNLO, multi-leg, etc. is almost trivial in LoopSim

Not the case in methods that merge with parton showers too

└ [LoopSim details]