Giant K factors

Gavin Salam

CERN, Princeton & LPTHE/CNRS (Paris)

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

ITP, Bern 11 May 2011



dσ/dm [log scale] Signal QCD prediction

mass

mass peak

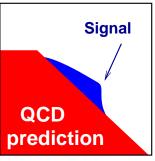
New resonance (e.g. Z') where you see all decay products and reconstruct an invariant mass

QCD may:

- swamp signal
- smear signal

leptonic case easy; hadronic case harder

mass edge



mass

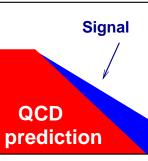
New resonance (e.g. R-parity conserving SUSY), where undetected new stable particle escapes detection.

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

QCD may:

- swamp signal
- smear signal

high-mass excess



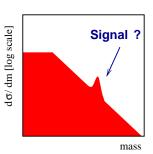
mass

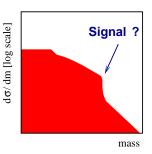
Unreconstructed SUSY cascade. Study effective mass (sum of all transverse momenta).

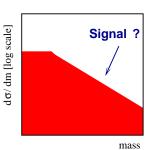
Broad excess at high mass scales.

Knowledge of backgrounds is crucial is declaring discovery.

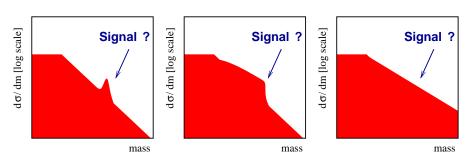
QCD is one way of getting handle on background.





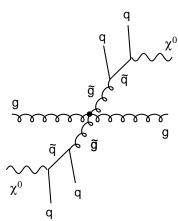


THIS TALK

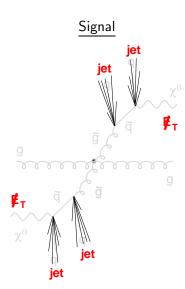


THIS TALK

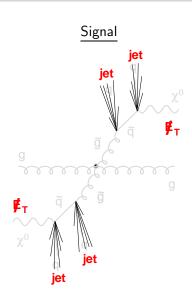
Signal



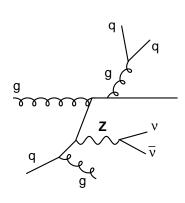
SUSY example: gluino pair production



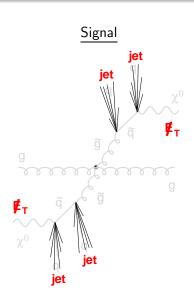
SUSY example: gluino pair production



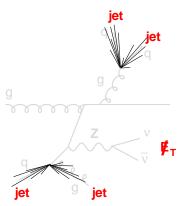
Background



SUSY example: gluino pair production



Background

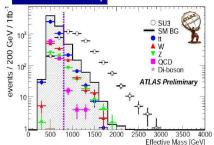


SUSY searches: what excesses?

Atlas selection [all hadronic]

- no lepton
- MET > 100 GeV
- 1^{st,2nd} jet > 100 GeV
- 3rd,4th jet > 50 GeV
- MET / m_{eff} > 20%

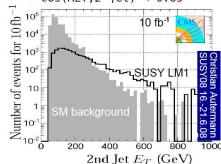
Christian Autermann SUSY08 16.-21.6.08



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
- $-0.95 < \cos(\text{MET}, 1^{\text{st}} \text{jet}) < 0.3$
- $cos(MET, 2^{nd}jet) < 0.85$

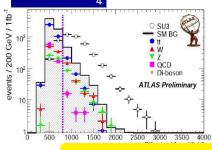


SUSY searches: what excesses?

Atlas selection [all hadronic]

- no lepton
- MET > 100 GeV
- 1^{st,2nd} jet > 100 GeV
- 3rd,4th jet > 50 GeV
- MET / m_{eff} > 20%

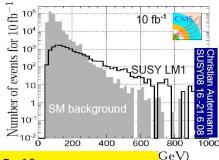
Christian Autermann SUSY08 16.-21.6.08



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
- $-0.95 < \cos(\text{MET}, 1^{\text{st}} \text{jet}) < 0.3$
- $cos(MET, 2^{nd}jet) < 0.85$



SUSY \simeq factor 5–10 excess

$$\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$$
$$\alpha_s \simeq 0.1$$

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

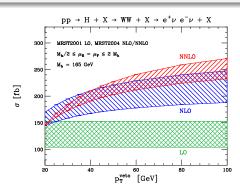
It isn't

Rules of thumb:

LO good to within factor of 2

NLO good to within scale

uncertainty



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%

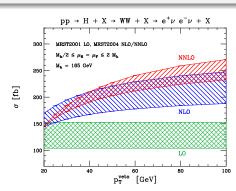
It isn't

Rules of thumb:

LO good to within factor of 2

NLO good to within scale

uncertainty



Anastasiou, Melnikov & Petriello '04 Anastasiou, Dissertori & Stöckli '07

$$\sigma = c_0 + c_1 \alpha_s + c_2 \alpha_s^2 + \dots$$
$$\alpha_s \simeq 0.1$$

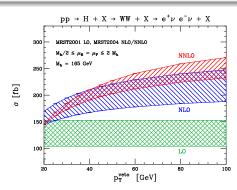
That implies LO QCD (just c_0) should be accurate to within 10%

It isn't

Rules of thumb:

LO good to within factor of 2

NLO good to within scale
uncertainty



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%

It isn't

Rules of thumb:

LO good to within factor of 2

NLO good to within scale
uncertainty

We don't always have NLO for the background (e.g. Z+4 jets, a 2 \rightarrow 5 process). Though amazing recent progress

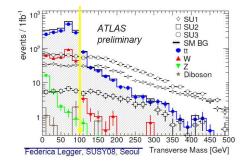
 $2 \rightarrow 4$: Blackhat, Rocket, Helac-NLO, BDDP $2 \rightarrow 5$ (W+4j): Blackhat

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



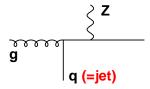
A conservative QCD theory point of view:

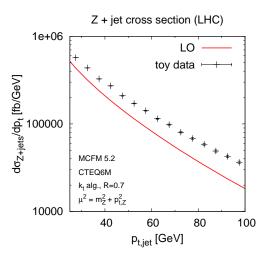
It's hard to be sure: since we can't (yet) calculate Z+4 jets beyond LO.

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" $2 \to 2$ NNLO $\sim e^+e^- \to \gamma^*/Z \to 3$ jets, NNLO: Gehrman et al '08, Weinzierl '08
- ▶ But let's pretend we only know it to LO, and look at the p_t distribution of the hardest jet (no other cuts keep it simple)





stage 1: get control sample

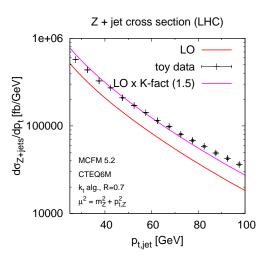
Check LO v. data at low p_t

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

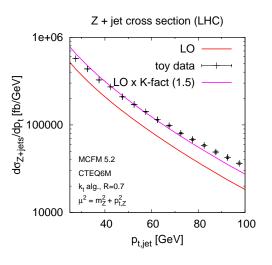


stage 1: get control sample

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

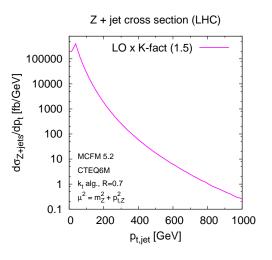


stage 1: get control sample

Check LO v. data at low p_t

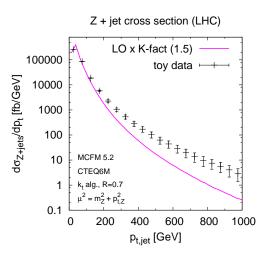
- normalisation off by factor 1.5 (consistent with expectations)
 So renormalise LO by K-fact
- ▶ shape OKish

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



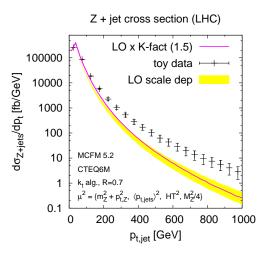
stage 2: look at high p_t

- ▶ good agreement at low p_t, 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



stage 2: look at high p_t

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



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

Is 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 5.7σ 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. . .

Is it:

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

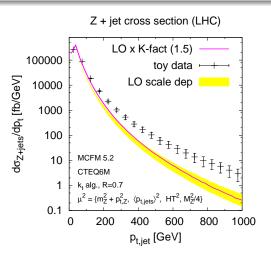
Here it's just a toy illustration. Later this year it may be for real:

▶ Do Nature / Science / PRL accept the paper?

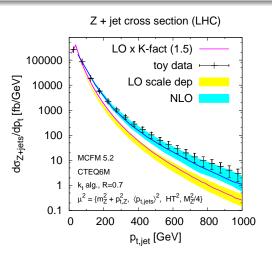
Discovery of New Physics at the TeV scale

We report a 5.7σ 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. . .



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



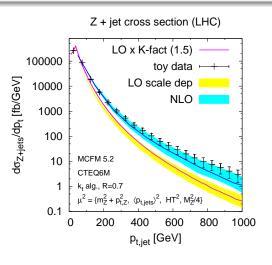
Unlike for SUSY multi-jet searches, in the Z+jet case we do have NLO.

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



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

Once NLO is included the excess disappears

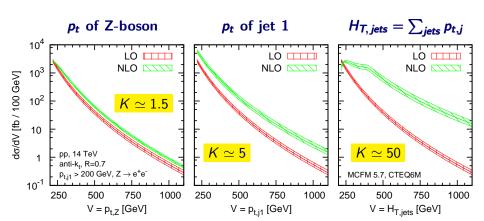
The "toy data" were just the upper edge of the NLO band

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

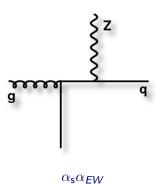
Related observations also by Bauer & Lange '09; Denner, Dittmaier, Kasprzik & Muck '09

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

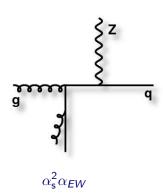
NB: DYRAD, MCFM consistent

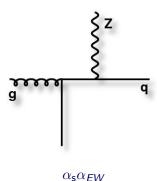


"Giant *K*-factors"

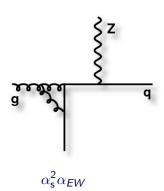


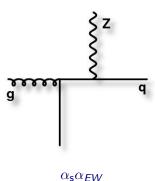
Next-to-Leading Order



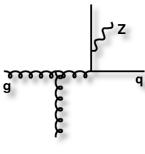


Next-to-Leading Order

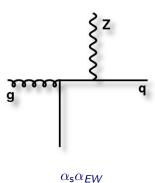




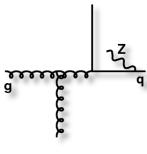
Next-to-Leading Order



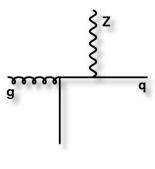
$$\alpha_s^2 \alpha_{EW} \ln^2 \frac{p_t}{M_Z}$$



Next-to-Leading Order

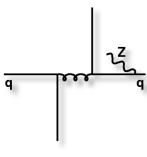


$$\alpha_s^2 \alpha_{EW} \ln^2 \frac{p_t}{M_Z}$$



 $\alpha_{\mathsf{S}}\alpha_{\mathsf{EW}}$

Next-to-Leading Order



$$\alpha_{\rm s}^2 \alpha_{EW} \ln^2 rac{p_t}{M_Z}$$

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

NLO driven in part by qq parton luminosity: large at pp colliders

Is this example not a little contrived?

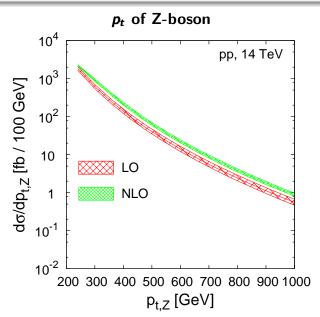
After all, experiments would surely notice unexpected event topology such as that here.

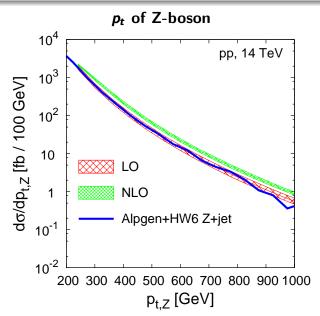
We actually first saw the problem in a more complex process: $Wb\bar{b}$ as a background to boosted Higgs searches (with "wrong" cuts). The more complicated the process, the trickier the diagnosis of the problem.

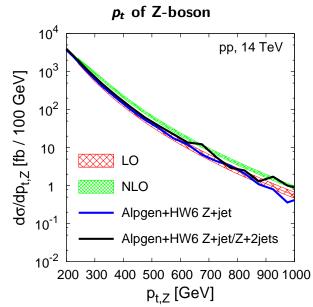
It's enough to get this wrong once, leading to "unwarranted" press-releases and major subsequent embarassment In day-to-day work the experiments don't just use LO.

Instead, they

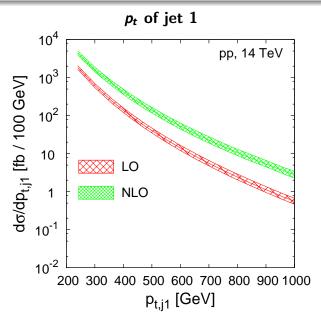
- ► Take Z + jet, Z + 2 jet samples, etc.,
- Attach a parton shower to each of them
- and combine with procedures such as MLM matching or CKKW matching.

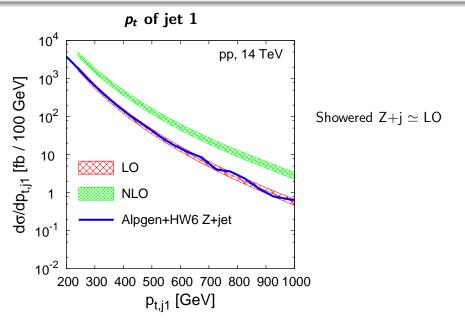


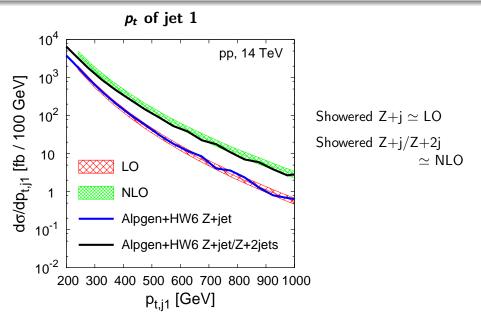


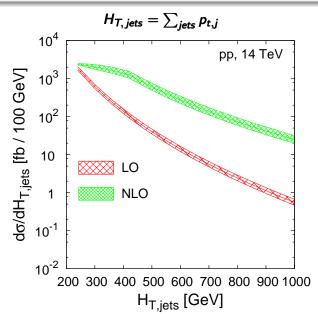


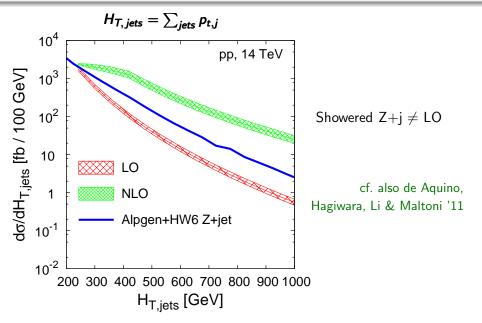
All predictions similar and stable

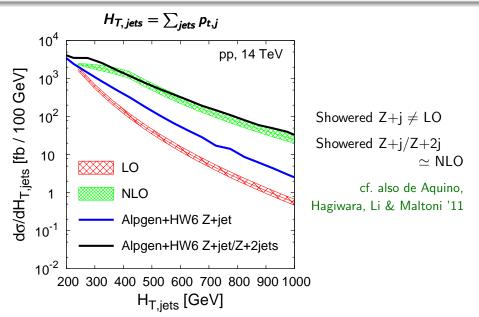












It's great that (the widely-used) MLM/CKKW matching procedure correctly approximates the NLO giant K-factor.

But what happens at NNLO?

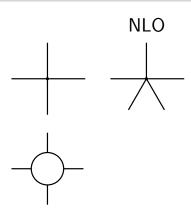
a natural question when LO \rightarrow NLO convergence is poor Despite 10 years' calculation, the answer is not yet known

Our strategy: get an approximation to NNLO

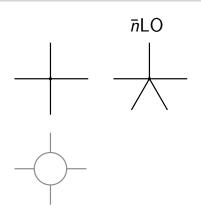
specifically, we will approximate a subset of the loop contributions, with a method dubbed "LoopSim"

LO





NNLO

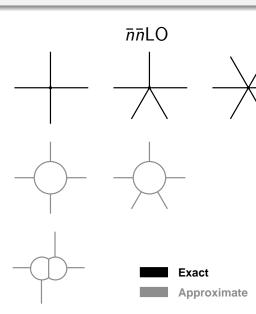


Our naming scheme:

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

▶ \bar{n} LO: approx 1-loop diagrams

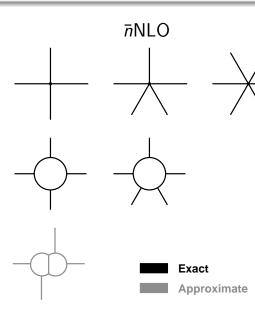




Our naming scheme:

For each loop that we approximate, replace N $ightarrow ar{n}$

- ▶ \bar{n} LO: approx 1-loop diagrams
- ▶ $\bar{n}\bar{n}$ LO: approx 1- and 2-loops



Our naming scheme:

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

- ▶ \bar{n} LO: approx 1-loop diagrams
- ▶ $\bar{n}\bar{n}$ LO: approx 1- and 2-loops
- $ightharpoonup \bar{n}NLO$: approx 2-loop only

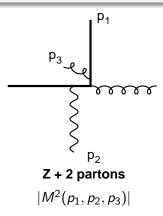
First try $Z + jet @ \bar{n}LO$:

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

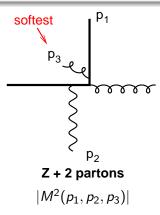
and add in process with one extra jet.

[i.e. include Z + 2 jets @ LO]

approximate the 1-loop Z+jet term, by requiring cancellation of all divergences [those from singly unresolved limit of Z+2 jets]



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

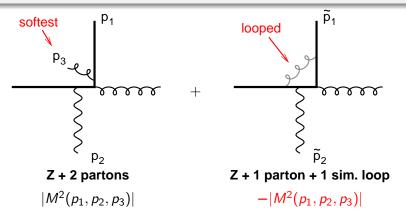


Identify softest or most collinear parton

[with help of a jet algorithm]

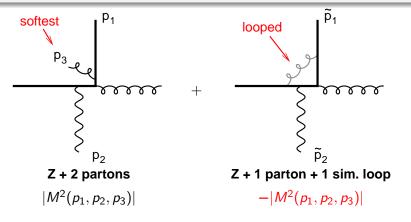
• "Loop" it \equiv remove it from event, reshuffle other momenta; weight of looped event is $(-1)\times$ weight of tree-level event

The LoopSim idea: \bar{n} LO



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

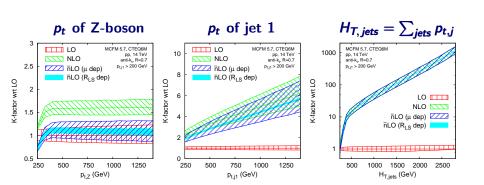
The LoopSim idea: \bar{n} LO



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

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

\bar{n} LO results (*K*-factors, normalised to LO)



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

Works with similar "quality" to MLM matching

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



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

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

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

$$|M^2(p_1,p_2,p_3,p_4)| \qquad -|M^2(p_1,p_2,p_3,p_4)| \qquad -|M^2(p_1,p_2,p_3,p_4)| \qquad +|M^2(p_1,p_2,p_3,p_4)|$$

Simultaneously loop each of the 2 softest emissions: provides approx of 2-loop

Total of tree plus approx 1- and 2-loop pieces gives zero

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

Simultaneously loop each of the 2 softest emissions: provides approx of 2-loop

Total of tree plus approx 1- and 2-loop pieces gives zero

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 $\bar{n}NLO$

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

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

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

$$K_{\mathsf{NNLO}} = \frac{\sigma_{\mathsf{NNLO}}}{\sigma_{\mathsf{LO}}} \sim K_{\mathsf{NLO}} \gg 1$$

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 $Z+2j@NLO \rightarrow Z+j@\overline{n}NLO$

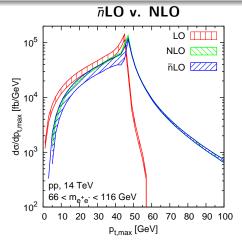
with MCFM

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

with NLOjet++



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

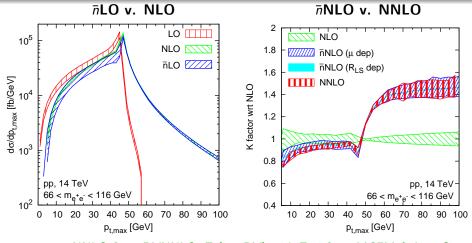


Z (i.e. DY) with $Z{+}j$ from MCFM & LoopSim

For $p_{tell} \gtrsim \frac{1}{2}M_Z + \Gamma_Z$ (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

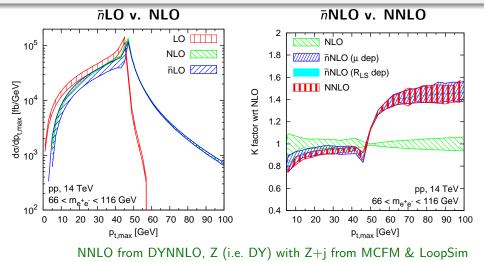


NNLO from DYNNLO, Z (i.e. DY) with Z+j from MCFM & LoopSim

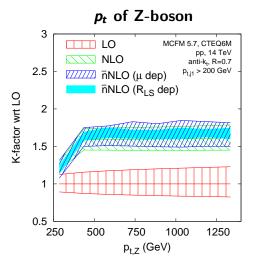
For $p_{t\,ell}\gtrsim \frac{1}{2}M_Z+\Gamma_Z$ (giant K-factor!) it had to work For $p_{t,\ell}\lesssim \frac{1}{2}M_Z+\Gamma_Z$ it's remarkable that it still works



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

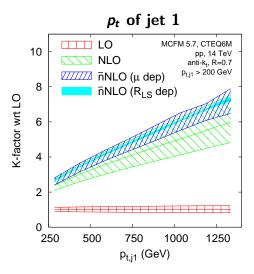


For $p_{t\,ell}\gtrsim \frac{1}{2}M_Z+\Gamma_Z$ (giant K-factor!) it had to work For $p_{t,\ell}\lesssim \frac{1}{2}M_Z+\Gamma_Z$ it's remarkable that it still works

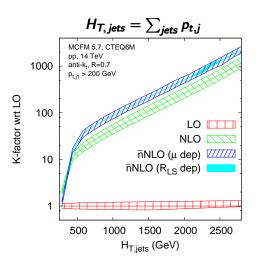


- ► *p*_{tZ} distribution didn't have giant *K*-factors.
- ▶ *n*NLO brings no benefit

To get improvement you would need exact 2-loop terms



- ▶ p_{tj} distribution seems to converge at \bar{n} NLO
- ▶ scale uncertainties reduced by ~ factor 2



- ► Signficiant further enhancement for *H*_{T,iets}
- ▶ \bar{n} NLO brings clear message:

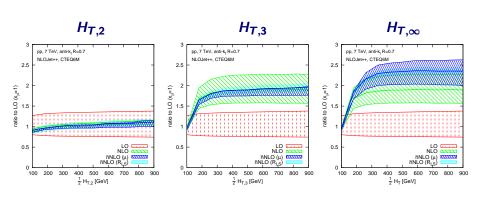
 $H_{T,jets}$ is not a good observable!

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,

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



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?

Be aware that giant K-factors exist

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

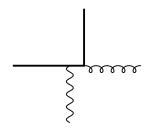
New tool to get good predictions in such cases: LoopSim

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

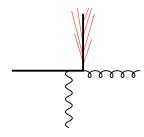
Watch out for H_T

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

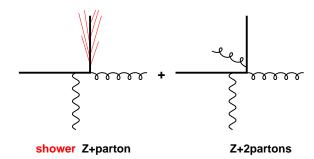
EXTRAS

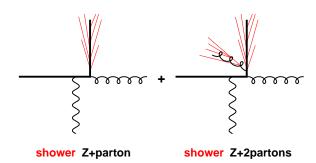


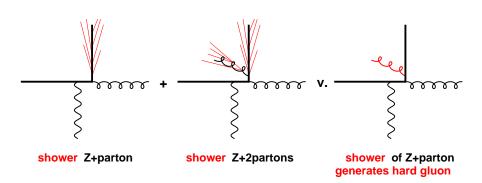
Z+parton

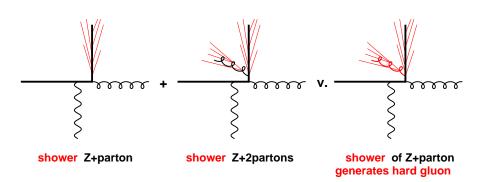


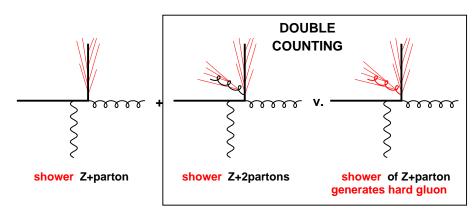
shower Z+parton





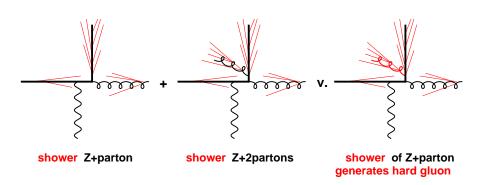






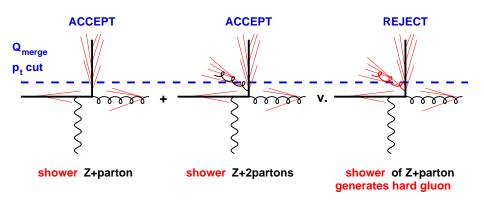
 $\mathsf{Z}+\mathsf{parton}$ implicitly includes part of $\mathsf{Z}+2$ partons It's just that the 2nd parton isn't always explicitly "visible"

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



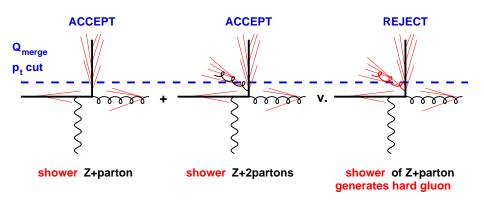
- ► MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.
- ▶ In a few slides, we will try to do that without the parton shower

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



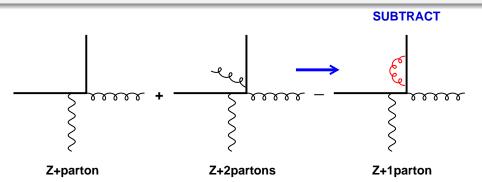
- ► MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.
- ▶ In a few slides, we will try to do that without the parton shower

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



- ► MLM merging relies on parton shower to help figure out what fraction of Z + parton is really Z + 2 partons.
- ▶ In a few slides, we will try to do that without the parton shower

cartoon of the LoopSim idea

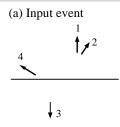


softest particle of Z+2 is "looped" = removed from event (kinematics reshuffled)

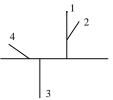
- ▶ For every Z + 2 parton $(2 \rightarrow 3)$ event, figure out what what $2 \rightarrow 2$ event it would really have come from "Loop" the softest parton [Don't actually explicitly calculate any loop diagrams: simulate the loops]
- ▶ Subtract that $2 \rightarrow 2$ event

Unlike MLM, no cutoffs on $2 \rightarrow 3$ events If done properly, divergences will cancel

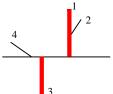
cartoon of LoopSim "looping" procedure



(b) Attributed emission seq.



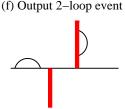
(c) Born particle ID



(d) Output 1–loop event



(e) 2nd output 1–loop event



Use jet algorithm to assign a branching structure to event

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

à la CKKW

Define operators:

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

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

"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

$$E_{n,\ell} \equiv$$
 event with n partons and ℓ exact loops $U_{orall,\ell} \equiv$ operator to apply when ℓ exact loops known $U_{orall,1}(E_{n,0}) = U_{orall}(E_{n,0}) - U_{orall}(U_1(E_{n,0}))$ $U_{orall,1}(E_{n,1}) = U_{orall}(E_{n,1})$

$$Z+j@\bar{n}NLO = Z+j@NLO + U_{\forall,1}(Z+2j@NLO_{only})$$

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

Not the case in methods that merge with parton showers too