GenEvA

A New Framework for Event Generation

Jesse Thaler (Berkeley)

with Christian Bauer and Frank Tackmann arXiv:0801.4026 [Physics] arXiv:0801.4028 [Techniques]

Monte Carlo in LHC Era

All experimental searches and measurements are (in one way or another) Monte Carlo sensitive.

How will we understand BSM backgrounds? $pp \rightarrow W + \text{jets} \quad pp \rightarrow Z + \text{jets}$ $pp \rightarrow t\bar{t} + \text{jets}$

Heavy resonances + QCD radiation. Multiple scales and potentially large logarithms.

TeVatron Example

(conversations with Beate Heinemann)

$$p\bar{p} \rightarrow Z + b / p\bar{p} \rightarrow Z$$

0.0023 ("NLO")

0.0035 ("LO")

This is important calibration for heavy flavor.

TeVatron Example

(conversations with Beate Heinemann)

$$p\bar{p} \rightarrow Z + b / p\bar{p} \rightarrow Z$$

 $0.0037 \pm 0.0006 (CDF)$

0.0023 ("NLO")

0.0035 ("LO")

This is important calibration for heavy flavor.

Scorecard

"NLO" = MCFM w/ Pythia UE + Had.

- + Order α_s^2
- Some Leading Logarithms
- + Proper Bottom Mass Treatment
- No PS/ME merging
- + All Angular Correlations

"LO" = Pythia Out-of-the-Box

- Order αs
- + All Leading Logarithms
- Ad Hoc Bottom Mass Treatment
- + "Normalized" PS/ME merging
- Some Angular Correlations

Two fundamentally different approaches, each with benefits and drawbacks.

Parton Showers

Perturbative								
α_s Expansion								

Fixed n-body Phase Space

Parton Showers

Pe	rturbative
αs	Expansion

Fixed n-body Phase Space

Soft Collinear Limit

Parton Showers Recursive Phase Space



Fixed n-body Phase Space

Soft Collinear Limit

Parton Showers Recursive Phase Space



Existing Tools

Merge successes of fixed-order calculations with successes of parton showers?

PS/ME Merging

Supplement Tree-Level Matrix Elements with Sudakov Information (CKKW, MLM, Lönnblad, ...)

MC@NLO

Combine Loop-Level Matrix Elements with Sudakov Information (FW, POWHEG, ...)

Traditional Approach

$d\sigma = \mathrm{MC}\left(|\mathcal{M}|^2 \, d\Phi\right)$

Dead zones? Double counting? Negative weights? Ambiguities?

Traditional Approach

Vetoed Showers, Modified Scale Choices

 $d\sigma = \mathrm{MC}\left(|\mathcal{M}|^2 \, d\Phi\right)$

Dead zones? Double counting? Negative weights? Ambiguities?

Traditional Approach

Vetoed Showers, Modified Scale Choices

$d\sigma = \mathrm{MC}\left(|\mathcal{M}|^2 \, d\Phi\right)$

Subtractions, Sudakovs, Multiple Samples

Dead zones? Double counting? Negative weights? Ambiguities?



Perturbative α_s Expansion



Soft Collinear Limit

Recursive Phase Space

Perturbative α_s Expansion

Fixed n-body Phase Space

Soft Collinear Limit

Recursive Phase Space



Calculational Merging

Soft Collinear Limit

Fixed n-body Phase Space





Calculational Merging

Soft Collinear Limit

Fixed n-body Phase Space

Algorithmic Merging

Recursive Phase Space

The GenEvA Framework

$d\sigma = |\mathcal{M}(\mu)|^2 dMC(\mu)$

No dead zones, no double counting, no negative weights, no incalculable ambiguities.

The GenEvA Framework

Calculations Algorithms $d\sigma = |\mathcal{M}(\mu)|^2 d\mathrm{MC}(\mu)$

No dead zones, no double counting, no negative weights, no incalculable ambiguities.

The GenEvA Framework



No dead zones, no double counting, no negative weights, no incalculable ambiguities.

GENerate EVents Analytically

- Algorithmic Side
 - "Deriving" the Master Formula
 - A New Approach to Phase Space
- Calculational Side
 - Proof-of-Concept Amplitudes
 - LO/LL Merging (Analog of PS/ME Merging)
 - NLO/LL Merging (Analog of MC@NLO)
 - NLO/LO/LL Merging (New!)
 - Technical Details
- GenEvA at the LHC



Ultimate Goal: Hadronic Collisions with Heavy Resonances

Current Status: Leptonic Collisions with Massless Partons

$$e^+e^- \rightarrow n \text{ jets}$$



There is real code....

```
+---- Command Line
| GenEvA --cms 1000 --cut 10 --numStat 10000 --best 6 50
+-----
```

```
+---- Event Generation Information

Process: e- e+ -> j j

Center-of-Mass Energy: 1000 GeV

Matching Scale: 50 GeV with maximum multiplicity 6

Shower Cutoff: 10 GeV

Generation: Events are matched to NLO/LO matrix element.
```

+.	Run S	Statistics							
	Process:	NumGen	NumKept	NumStat	StatEff	NumUnw	UnwEff	Sigma +/- dS (pb)	(error%)
	Global:	19771	18674	10000.3	0.536	6485.0	0.347	0.253007 +/- 0.001779	(0.70%)
	2j:	2303	2303	2303.0	1.000	2303.0	1.000	0.089849 +/- 0.001760	(1.96%)
	3j :	8480	7383	6406.3	0.868	3539.7	0.479	0.129731 +/- 0.001333	(1.03%)
	4j:	5629	5629	3351.1	0.595	905.4	0.161	0.029322 +/- 0.000462	(1.57%)
	5j :	2492	2492	1187.3	0.476	254.1	0.102	0.003693 +/- 0.000104	(2.81%)
	6j :	867	867	326.1	0.376	82.2	0.095	0.000412 +/- 0.000023	(5.49%)
+.									

+---- Thank you for running GenEvA

....and it's reasonably user-friendly.

GenEvA Master Formula

Generic Solution to Merging Fixed-Order Calculations with Parton Showers

$$d\sigma = \left|\mathcal{M}(\mu)\right|^2 d\mathrm{MC}(\mu)$$



Traditional Approach $d\sigma = \mathrm{MC}\left(|\mathcal{M}|^2 \, d\Phi\right)$ $E_{\rm CM}$ Partonic μ -----Showering $\Lambda_{
m QCD}$ Hadronization



Traditional Approach $d\sigma = \mathrm{MC}\left(|\mathcal{M}|^2 \, d\Phi\right)$

Three Technical Problems

Infrared Divergences
 Scale Dependence
 Double Counting

GenEvA Framework $d\sigma = |\mathcal{M}(\mu)|^2 dMC(\mu)$

Three Conceptual Solutions

- I. Infrared Divergences? → Merge QCD Approximations
- 2. Scale Dependence? → Merge Calc. with Pheno. Models
- 3. Double Counting? → Merge Phase Space Algorithms

GenEvA Framework $d\sigma = |\mathcal{M}(\mu)|^2 d\mathrm{MC}(\mu)$














I. Different QCD Approx.

We have Fixed Order Expansion in α_s .

We have (Sub-)Leading Logarithms in Soft-Collinear Limit.

Infrared Divergences should Cancel between Trees and Loops.

$$|\mathcal{M}|^2 \to |\mathcal{M}(\tilde{\mu})|^2$$

Infrared divergences cancelled in definition of "amplitude". Infrared scale $\tilde{\mu}$ needed to resum $\alpha_s \log^2 r$ terms.

Partonic Calculations

$$|\mathcal{M}^{\mathrm{CKKW}}(\tilde{\mu})|^2 \simeq |\mathcal{M}^{\mathrm{tree}}|^2 \Delta(E_{\mathrm{CM}},\tilde{\mu})$$

$$|\mathcal{M}_n^{\mathrm{NLO}}(\tilde{\mu})|^2 \simeq |\mathcal{M}_n^{\mathrm{tree}+\mathrm{loop}}|^2 + \int_{\tilde{\mu}} |\mathcal{M}_{n+1}^{\mathrm{tree}}|^2$$

$$|\mathcal{M}^{\mathrm{MC}\otimes\mathrm{NLO}}(\tilde{\mu})|^2 \neq |\mathcal{M}^{\mathrm{NLO}}(\tilde{\mu})|^2 \Delta(E_{\mathrm{CM}},\tilde{\mu})$$

2. Calculations vs. Models

Calculations Available for Finite Number of Particles.

Need Parton Shower to fill out Phase Space.

Need Hadronization Model for Detector Simulation.

 $\mu \to \mu$

If "amplitude" has correct leading logarithms, interface with parton shower will be smooth if μ scale is the same.

3. Phase Space Algorithms

Field Theory Calculations need Fixed Number of Final States.

Parton Showers need Variable Number of Final States.

Want Every Phase Space Point Covered Once and Only Once.

$MC(d\Phi) \rightarrow dMC(\mu)$

Replace two event generation frameworks with one master framework that solves double counting by construction.

$$d\sigma = \mathrm{MC}\left(\left|\mathcal{M}
ight|^2 d\Phi
ight)$$

Infrared Divergences $\left|\mathcal{M}
ight|^2
ightarrow \left|\mathcal{M}(ilde{\mu})
ight|^2$ Merge QCD Approx.
Scale Dependence $ilde{\mu}
ightarrow \mu$ Merge Calc. w/ Model
Double Counting $\mathrm{MC}\left(d\Phi
ight)
ightarrow d\mathrm{MC}(\mu)$ Merge Algorithms
 $d\sigma = \left|\mathcal{M}(\mu)
ight|^2 d\mathrm{MC}(\mu)$





GenEvA Phase Space

Understanding the Effect of the Parton Shower



Partonic Phase Space



The Parton Shower



Additional Emissions



How to avoid double counting between 2-body showered and 3-body unshowered?



dMC is dΦ organized in terms of showered areas. Double-counting solved by construction. Simple to say, technically challenging to implement.

Complete Phase Space



$$d\sigma = \sum_{n=2}^{n_{\max}} |\mathcal{M}_n(\mu)|^2 d\mathrm{MC}_n(\mu)$$

The amplitude is a function of n-body phase space, but influences $(\geq n)$ -body phase space through shower.

What is the Shower?

Parton shower fills out phase space starting from hard scattering matrix element.

$$d\sigma = \left|\mathcal{M}_2^{\text{hard}}\right|^2 d\mathrm{MC}_2(E_{\mathrm{CM}})$$

What is the Shower?

Parton shower fills out phase space starting from hard scattering matrix element.

$$d\sigma = \left|\mathcal{M}_2^{\text{hard}}\right|^2 d\mathrm{MC}_2(E_{\mathrm{CM}})$$

There must be an equivalent description of same physics with no shower!

$$d\sigma = \sum_{n=2}^{\infty} \left| \mathcal{M}_n^{\text{shower}} \right|^2 d\Phi_n$$

What is the Shower?

There is also an equivalent description of the same physics with part shower, part "matrix element"!

$$d\sigma = \sum_{n=2}^{n_{\max}} \left| \mathcal{M}_n^{\text{shower}}(\mu) \right|^2 d\text{MC}_n(\mu)$$

The scale μ gives this interpolation meaning, by capturing correct leading-logarithmic dependence.



Improving Monte Carlo

$$d\sigma = \sum_{n=2}^{n_{\max}} |\mathcal{M}_n(\mu)|^2 d\mathrm{MC}_n(\mu)$$

Choose the best possible expression for $\left|\mathcal{M}_n(\mu)\right|^2$

and lower μ and raise n_{max} as far as possible.

GenEvA Amplitudes

Comparing Different Expansions of QCD

 $|\mathcal{M}(\mu)|^2$

Terminology

LL: Leading Logarithms Correct Sudakov Factors in Soft/Collinear Limit

LO: Tree-Level Matrix Elements Correct Quantum Interference in Large Angle Limit

NLO: Next-to-Leading Order Everything Correct to Order α_s

LO Tree-Level Generators (ALPGEN, MadEvent, CompHep, Apacic, Whizard, Helac, ...)









LO/LL Analog of PS/ME Merging (CKKW, MLM, Lönnblad, ...)











NLO Loop-Level Generators (MCFM, NLOJet, PHOX, ...)





NLO/LL Analog of MC@NLO (FW, POWHEG, ...)





NLO/LO/LL GenEvA Best











NLO/LO/LL GenEvA Best (New!)



Figure of Merit?

How would you know whether we have actually achieved an NLO/LO/LL sample?

Normalization

The μ -dependence should scale likeNo LL: $\alpha_s \log^2 \mu$ LO/LL: $\alpha_s \log \mu$ NLO/LL: $\alpha_s^2 \log^2 \mu$

Shape

A merged sample should interpolate between the two underlying differential distributions.

Cross Section Scaling



Baseline Shower



LO/LL Calculation



LO/LL answer is smaller than either approximation.

LO/LL Calculation



LO/LL answer is smaller than either approximation.

NLO/LL Calculation



A "Goldilocks" Interpolation

NLO/LO/LL Calculation



Interpolates between PS/ME Merging and MC@NLO!

NLO/LO/LL Calculation



Interpolates between PS/ME Merging and MC@NLO!
Isolated Components



Non-trivial combination of five different samples.

Isolated Components



Only single-logarithmic change in total distribution.

"Data" Comparison



GenEvA Details

Strategy to Merge Different Approximation Schemes

$$\left|\mathcal{M}^{\mathrm{A}}(\mu)\right|^{2}$$
 vs. $\left|\mathcal{M}^{\mathrm{B}}(\mu)\right|^{2}$



$$\left|\mathcal{M}^{\text{Best}}(\mu)\right|^2 = \left|\mathcal{M}^{\text{A}}(\mu)\right|^2$$







NLO/LO/LL



C: NLO/LL B: LO/LL A: Shower (MC@NLO) (PS/ME Merging)



C: NLO/LL B: LO/LL A: Shower (MC@NLO) (PS/ME Merging)



Same four-vectors are determined by multiple shower histories. Dominant history is the most singular one.

LO/LL Merging



In singular regions of phase space:

$$\left|\mathcal{M}_n^{\text{tree}}\right|^2 \to \sum_j Q_j$$

Interference terms in tree-level matrix element with Sudakovs from shower "matrix element"?

$$\frac{\text{LO/LL Merging}}{\left|\mathcal{M}_{n}^{\text{LO/LL}}(\mu)\right|^{2}} = \left|\mathcal{M}_{n}^{\text{tree}}\right|^{2} \sum_{i} \frac{Q_{i}}{\sum_{j} Q_{j}} \Delta_{i}(\mu)$$

Shower doesn't factorize, but in singular regions:

$$\frac{Q_{\text{dom}}}{\sum_{j} Q_{j}} \to 1 \qquad \frac{Q_{\text{other}}}{\sum_{j} Q_{j}} \to 0$$

$$\left|\mathcal{M}_{n}^{\mathrm{LO/LL}}(\mu)\right|^{2} \simeq \left|\mathcal{M}_{n}^{\mathrm{tree}}\right|^{2} \Delta_{\mathrm{dom}}(\mu)$$

Equivalent to CKKW in singular regions.

NLO/LL Merging

$$\sigma_2(\mu) = \sigma_{\rm NLO} \Delta_R(\mu)$$

$$\frac{d\sigma_3(t)}{dt} = \sigma_{\rm NLO} R(t) \Delta_R(t)$$
$$= \frac{d\sigma_3^{\rm tree}(t)}{dt} + \mathcal{O}(\alpha_s^2)$$

As shown by POWHEG, turn NLO calculation into "shower" with novel "splitting function". By construction, cross section is correct to NLO.

GenEvA Outlook

Hadronic Collisions, Heavy Resonances, Advanced Matrix Elements

The GenEvA Framework



No dead zones, no double counting, no negative weights, no incalculable ambiguities.

GenEvA IL or CH To be relevant for the LHC, we need... Calculations Algorithms $d\sigma = |\mathcal{M}(\mu)|^2 |dMC(\mu)|$ Proper Fact./Renorm. Scale Treatment **Proper Mass Treatment**

Parton Distribution Functions ISR/FSR Interference Proper Mass Treatment Interface with p_⊥ Showers ISR/FSR Double Counting Resonance/Showerer

These are technical issues, not conceptual ones. Consequence of μ appearing in both calculations and algorithms.

Theory Challenge $\left|\mathcal{M}^{\mathrm{Best}}(\mu)\right|^2$

SCET Matrix Elements

Subleading-logarithmic treatment of multiple scales?

NNLO/NLO/LO/NLL/LL

Describe NⁱLO observables accurate to NⁱLO and N^jLL observables accurate to N^jLL, simultaneously?

Preliminary SCET Work

(Matrix Elements from Matthew Schwartz)





Backup Slides

In Case You Were Wondering...

Reweighting (Simple)



Reweighting (Jacobian)



Reweighting (Not I-I)



Reweighting (GenEvA)



MadEvent Comparison

process	MadEvent	GenEvA	process	MadEvent	GenEvA
LO 3 (fb)	216.71 ± 0.21	216.77 ± 0.22	LO 5 (ab)	2542 ± 3	2543 ± 3
$u\bar{u}g$	86.62 ± 0.13	86.60 ± 0.18	$u\bar{u}ggg$	912 ± 2	912 ± 2
$d ar{d} g$	21.75 ± 0.07	21.55 ± 0.10	$d\bar{d}ggg$	227.5 ± 0.9	228.3 ± 0.8
$s \overline{s} g$	21.63 ± 0.06	21.73 ± 0.10	$u\bar{u}d\bar{d}g$	33.8 ± 0.2	34.3 ± 0.4
$c\overline{c}g$	86.71 ± 0.13	86.70 ± 0.18	$u \bar{u} u \bar{u} g$	25.6 ± 0.2	25.7 ± 0.3
LO 4 (fb)	36.44 ± 0.04	36.49 ± 0.04	LO 6 (ab)	67.9 ± 0.3	68.0 ± 0.2
$u\bar{u}gg$	14.00 ± 0.03	14.00 ± 0.02	$u\bar{u}gggg$	22.41 ± 0.09	22.29 ± 0.12
$d ar{d} g g$	3.504 ± 0.013	3.511 ± 0.011	$u \bar{u} u \bar{u} g g$	1.117 ± 0.006	1.14 ± 0.03
$u \bar{u} d \bar{d}$	0.175 ± 0.001	0.180 ± 0.003	$u \bar{u} u \bar{u} u \bar{u} u \bar{u}$	0.005 ± 0.001^{-1}	0.005 ± 0.001
$u \bar{u} u \bar{u}$	0.132 ± 0.001	0.132 ± 0.002	$u\bar{u}d\bar{d}s\bar{s}$	0.019 ± 0.001^{-1}	0.020 ± 0.005

MadEvent Comparison



MadEvent Comparison

	η_{eff}	$T_{\rm eff} \ ({\rm msec})$	$T_{0.9} (\text{msec})$
GenEvA LO 3	0.789	0.57	0.62
GenEvA LO/LL inc. 3	0.965	0.47	< 0.47
MadEvent 3	0.982	2.6	< 2.6
MadEvent $u \bar{u} g$	0.994	3.0	< 3.0
GenEvA LO 4	0.525	1.7	2.2
GenEvA LO/LL inc. 4	0.713	1.3	1.5
MadEvent 4	0.809	11.1	11.4
MadEvent $u \bar{u} g g$	0.752	5.4	5.7
GenEvA LO 5	0.390	10.0	15
GenEvA LO/LL inc. 5 $$	0.557	8.6	10.8
MadEvent 5	0.843	62	64
MadEvent $u ar{u} g g g$	0.833	27	27
GenEvA LO 6	0.298	160	250
GenEvA LO/LL inc. 6	0.396	150	230
MadEvent 6	0.809	1900	2300
MadEvent $u \bar{u} g g g g$	0.784	330	350

GenEvA Efficiency



More Interpolations



More Components



Differential Scaling



Differential Scaling

