Towards an accurate and efficient event generation for multi-jet processes

Timea Vitos[1]

Uppsala University,
ELTE Eötvös Loránd Budapest

High Precision for Hard Processes
Torino, September 10-13, 2024

[1] *In collaboration with Rikkert Frederix, Lund University*

Disclaimer

Back to tree level in the next 20 minutes!

Focus on **high-multiplicity**!

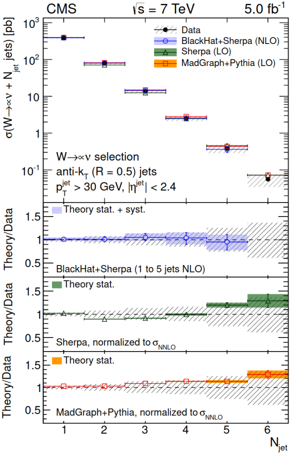
Today's talk

Today's talk

High-multiplicity processes: why bother?

- Hadronic decays of Standard Model background: leads to **multi-jet signatures**
  $\rightarrow \geq 5$ well-separated final-state jets in the detector
- (Almost) any new physics search includes decays leading to such signatures
- These are the processes

  $pp \rightarrow$ gluons + quarks (+ colour singlets)
- Colour singlets contribute to phase space but not to colour interaction



High-multiplicity processes in matrix element generators

- Current colour treatment: **colour decomposition** of amplitudes:

$$\langle|\mathcal{M}|^2\rangle \propto \sum_{\sigma_k,\sigma_l} \underbrace{C(\sigma_k,\sigma_l)}_{\text{colour matrix}} \underbrace{\mathcal{A}(\sigma_k)(\mathcal{A}(\sigma_l))^*}_{\text{dual amplitudes}} \tag{2}$$

  where $\sigma_{k,l}$ are some permutations of the external particles (each dual amplitude $\mathcal{A}(\sigma_k)$ corresponds to a **colour order**)

- Colour matrix size for n final state particles: $\sim n! \times n!$
- This factorial bottleneck is present already at **tree-level**: place for a first-step improvement

Today's talk

Colour decomposition of amplitudes

$$\langle|\mathcal{M}|^2\rangle \propto \sum_{\sigma_k,\sigma_l} C(\sigma_k,\sigma_l)\mathcal{A}(\sigma_k)(\mathcal{A}(\sigma_l))^*$$

- Expand in N_C^{-2} (large-N_C limit)

$$C(\sigma_k,\sigma_l) = \underbrace{a_0 N_C^x}_{\text{Leading colour (LC)}} + \underbrace{a_1 N_C^{x-2}}_{\text{Next-to-leading colour (NLC)}} + \mathcal{O}(N_C^{x-4}) \quad \forall k,l \quad (3)$$

  keeping all terms yields full colour (FC)
- Schematically, we obtain a sparse colour matrix in this NLC approximation:

$$C(\sigma_k,\sigma_l) \to \begin{pmatrix} \mathrm{LC} & 0 & 0 & 0 & 0 & \mathrm{NLC} \\ 0 & \mathrm{LC} & 0 & \mathrm{NLC} & 0 & 0 \\ 0 & 0 & \mathrm{LC} & 0 & 0 & 0 \\ 0 & \mathrm{NLC} & 0 & \mathrm{LC} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{LC} & 0 \\ \mathrm{NLC} & 0 & 0 & 0 & 0 & \mathrm{LC} \end{pmatrix}$$

- Truncation at NLC greatly improves the number of terms in the colour sum[2]

[2] R. Frederix, T. Vitos, arXiv:2109.10377

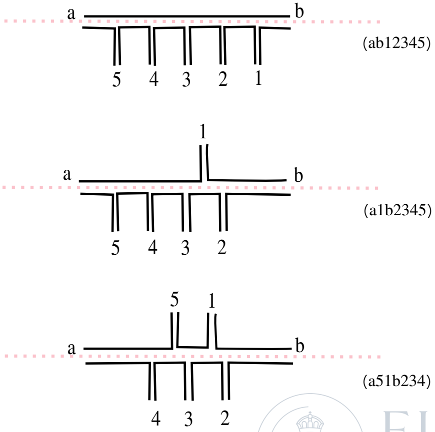
Some conventions for colour ordering

- Colour order classified by the string of particles in the process

$$ab \to 1 \dots n \quad (4)$$

- Same ordering when quarks around, with each quark line being independent with respect to the overall order

$$(\underbrace{q_1 g_1 g_2 g_3 \overline{q}_1}_{\text{quark line 1}} \underbrace{q_2 g_4 g_5 \overline{q}_2}_{\text{quark line 2}}) \quad (5)$$



Today's talk

Proposal for an efficient event generator

- The colour-summed and averaged matrix element using the colour decomposition

$$\langle|\mathcal{M}|^2\rangle \propto \sum_{\{c\},\{h\}} \sum_{\sigma_k,\sigma_l} C(\sigma_k,\sigma_l)\mathcal{A}(\sigma_k,\{h\})(\mathcal{A}(\sigma_l,\{h\}))^* \quad (6)$$

- Expansion in N_C^{-2} (for all-gluon process):

$$\langle|\mathcal{M}|^2\rangle = \frac{N_C^n(N_C^2-1)}{2^{(n-2)}} \sum_{\sigma_k} \left(\underbrace{|\mathcal{A}(\sigma_k,\{h\})}_{\text{LC}}|^2 + \mathcal{O}(N_C^{-2}) \right) \quad (7)$$

- LC-amplitude interferences of MHV-type have simple and well-known peak structure

$$\mathcal{A}(\sigma_k,h)(\mathcal{A}(\sigma_k,h))^* \sim \frac{1}{s_{\sigma_k(1)\sigma_k(2)}s_{\sigma_k(2)\sigma_k(3)}\cdots s_{\sigma_k(n)\sigma_k(1)}} \quad (8)$$

with $s_{\sigma(i)\sigma(j)} = (p_{\sigma(i)} + p_{\sigma(i+1)})^2$

Proposal for an efficient event generator

- Re-write the integrated cross section as:

$$\int \sum_{\{c\}} |\mathcal{M}|^2 d\Phi_n = \sum_{\sigma_k} \int |\mathcal{A}_{\sigma_k}|^2 \underbrace{\frac{\sum_{\{c\}} \sum_{\sigma_k, \sigma_l} \mathcal{A}_{\sigma_k} C_{\sigma_k \sigma_l} \mathcal{A}^*_{\sigma_l}}{\sum_{\sigma_k} |\mathcal{A}_{\sigma_k}|^2}}_{\text{reweight factor}} d\Phi_n \qquad (9)$$

  one obtains $(n-1)! \to \lfloor \frac{n-2}{2} \rfloor$ different channels for LC generation
- **Idea**: generate events with approximated integrand which is close to the full integrand → LC approximation of integrand
- Generate FC-accurate events in two steps:

  1. **Integration**: approximate peak structure of integrand with LC accuracy
  2. **Reweighting**: reweight LC events to FC with reweight factor

Today's talk

Phase-space integration

- Integration over the Lorentz-invariant phase space

$$d\Phi_n \sim \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}$$

- With the overall momentum conservation $\delta^4(P_{\rm in} - P_{\rm out})$, this is a $3n - 4$ dimensional phase space
- For hadron collisions, we have an additional integral

$$\int dx_1 dx_2 \qquad (10)$$

- We investigate two setups:

LHC setup

- Collision with protons at $\sqrt{s} = 14$ TeV
- Cuts: $p_T > 30$ GeV, $|\eta| < 6.0$, $\Delta R > 0.4$

Fixed-energy setup

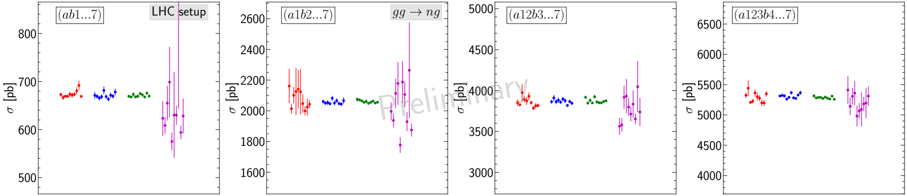
- Partonic collision at $\sqrt{\hat{s}} = 1$ TeV
- Cuts: $\sqrt{|2p_i \cdot p_j|} \geq 30$ GeV

Phase-space integration

- Any variable change from d^3p works: but some work better than others!
- We make a comparison of four choices of integration variables:

  HAAG[3] t-channel[4] $2 \to 3$[5] p_T-based

- Compare the Monte Carlo error for the cross section and spread of result for 10 random seeds, with fixed number of grid points

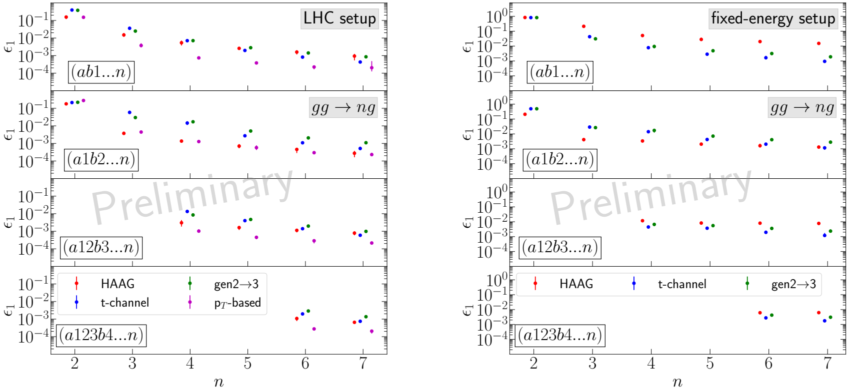


[3] A. van Hameren, C.G. Papadopoulos, arXiv:hep-ph/0204055
[4] E. Byckling, K. Kajantie, 10.1016/0550-3213(69)90271-5
[5] E. Byckling, K. Kajantie, 10.1103/PhysRev.187.2008

Phase-space integration

- Efficiency measure of LC events:

$$\epsilon_1 = \frac{N_{\text{events}}}{N_{\text{total}}} \tag{11}$$



Today's talk

Reweighting the LC events to FC and NLC

- From the set of LC-accurate events, generating FC accuracy, one reweights each event with $r_{\mathrm{LC}\to\mathrm{FC}}$ according to

$$\sum_{\sigma_k} \int |\mathcal{A}_{\sigma_k}|^2 \underbrace{\frac{\sum_{\{c\}} \sum_{\sigma_k,\sigma_l} \mathcal{A}_{\sigma_k} C_{\sigma_k \sigma_l} \mathcal{A}^*_{\sigma_l}}{\sum_{\sigma_k} |\mathcal{A}_{\sigma_k}|^2}}_{\text{reweight factor } r_{\mathrm{LC}\to\mathrm{FC}}} \mathrm{d}\Phi_n$$

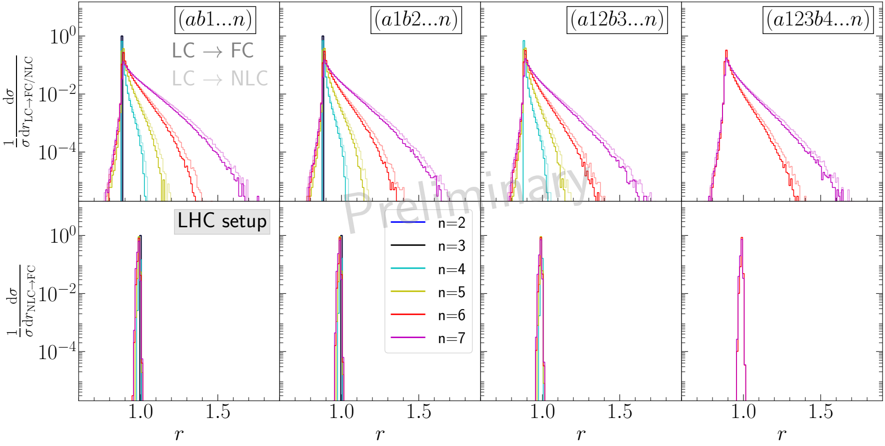
- An alternative is to reweight to NLC accuracy, reducing significantly the amount of terms needed in the colour sum

$$\frac{\sum_{\{c\}} \sum_{\sigma_k,\sigma_l} \mathcal{A}_{\sigma_k} C_{\sigma_k \sigma_l} \mathcal{A}^*_{\sigma_l}}{\sum_{\sigma_k} |\mathcal{A}_{\sigma_k}|^2} \to \underbrace{\frac{1}{\sum_{\sigma_k} |\mathcal{A}_{\sigma_k}|^2} \left(\sum_{\{c\}} \sum_{\sigma_k} \mathcal{A}_{\sigma_k} \times \left[\sum_{w} C_w \mathcal{A}^*_w \right] \right)}_{r_{\mathrm{LC}\to\mathrm{NLC}}} \tag{12}$$

where all $C_w \sim \mathrm{NLC}$

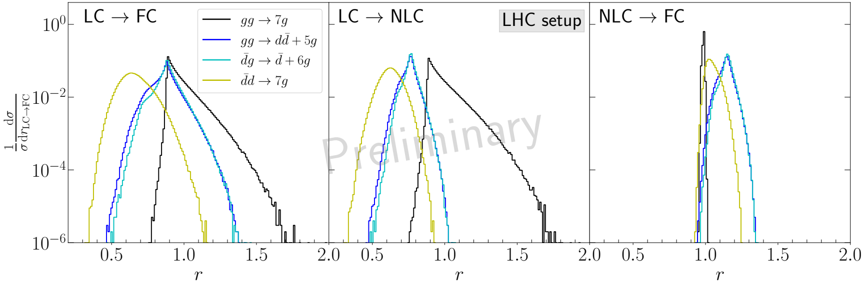
Weight spread of reweighting procedure

- For a secondary unweighting step, we need events which are kept with high efficiency and no very high weights $\rightarrow$ narrow weight spread
- Multiplicity dependence of this weight spread for all-gluon process shows peaked distribution for the reweight factors



Weight spread of reweighting procedure

- For processes with one quark line the reweighting efficiency is similar:

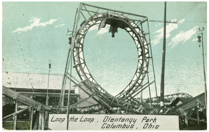


Summary and outlook

- Comparison of integration variable choice:
  → overall, the 2→3 method is the best-performing
  → for higher multiplicities, for partonic processes and for simple colour orders, HAAG performs better
  → a p_T-based integration is less effieinct, even with physical cuts
- Reweighting technique is promising:
  → relatively small weight spread for all process types

Future prospects

- Extend to physical processes and LHC physics
- Extend to NLO computations (colour treatment in loops)



Thank you for listening!

Appendix

Colour matrix: fundamental decomposition

For n-gluon amplitudes

- Matrix-element

$$\mathcal{M} = g^{n-2} \sum_{\sigma \in S_{n-1}} \mathrm{Tr}[T^{a_1} T^{a_{\sigma(1)}} ... T^{a_{\sigma(n-1)}}] \mathcal{A}(1, \sigma(1), ..., \sigma(n-1)) \tag{13}$$

Not a minimal set

Dual amplitudes are related by the Kleiss-Kuijf relation (dual Ward identity)

$$\mathcal{A}(1, 2, 3, 4...., n) + \mathcal{A}(2, 1, 3, 4, ..., n) + ... + \mathcal{A}(2, 3, 4, ..., 1, n) = 0 \tag{14}$$

- Squared matrix-element

$$|\mathcal{M}|^2 = (g^2)^{n-2} \sum_{k,l=1}^{(n-1)!} C_{kl} \mathcal{A}(1, \sigma_k(1), \ldots, \sigma_k(n-1)) \left(\mathcal{A}(1, \sigma_l(1), \ldots, \sigma_l(n-1))\right)^*$$

- Colour matrix (size $(n-1)! \times (n-1)!$):

$$C_{kl} = \sum_{\text{col.}} \mathrm{Tr}[T^{a_1} T^{a_{\sigma_k(1)}} \ldots T^{a_{\sigma_k(n-1)}}] \left(\mathrm{Tr}[T^{a_1} T^{a_{\sigma_l(1)}} \ldots T^{a_{\sigma_l(n-1)}}]\right)^* \tag{15}$$

Colour matrix: results

For n-gluon amplitudes

Including the adjoint decomposition: matrix size $(n-2)! \times (n-2)!$

all-gluon			
n	Fundamental	Colour-flow	Adjoint
4	6 (6)	6 (6)	2 (2)
5	11 (24)	16 (24)	5 (6)
6	24 (120)	36 (120)	18 (24)
7	50 (720)	71 (720)	93 (120)
8	95 (5040)	127 (5040)	583 (720)
9	166 (40320)	211 (40320)	4162 (5040)
10	271 (362880)	331 (362880)	31649 (40320)
11	419 (3628800)	496 (3628800)	-
12	620 (39916800)	716 (39916800)	-
13	885 (479001600)	1002 (479001600)	-
14	1226 (6227020800)	1366 (6227020800)	-

Adjoint decomposition

For n-gluon amplitudes

- The amplitude is now

$$\mathcal{M} = \sum_{\sigma \in S_{n-2}} \left(F^{a_{\sigma(2)}} \dots F^{a_{\sigma(n-1)}} \right)_{a_1 a_n} \mathcal{A}(1, \sigma(1), \dots, \sigma(n), n), \tag{16}$$

  with $(F^a)_{bc} = if^{abc}$
- Minimal basis: $(n-2)!$ independent dual amplitudes
- Smaller colour matrix: but **LC not only on diagonal!**
- No found algorithm (yet) to get NLC elements

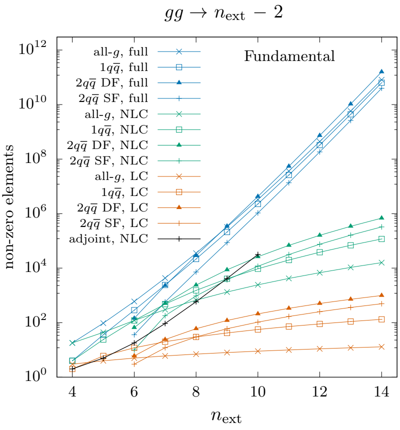
Colour-flow decompositions

For one quark line plus n-gluon amplitudes: the full projection of U(1) gluons

$$
\begin{aligned}
\mathcal{M}_{1qq} = &\sum_{\sigma\in S_n} \delta^{i_q}_{j_{\sigma(1)}} \delta^{i_{\sigma(1)}}_{j_{\sigma(2)}} \dots \delta^{i_{\sigma(n-1)}}_{j_{\sigma(n)}} \delta^{i_{\sigma(n)}}_{j_q} \mathcal{A}_{1qq}(q,\sigma(1),\dots,\sigma(n),\bar{q}) \\
&+ \left(\frac{-1}{N}\right) \sum_{\sigma\in S_n} \delta^{i_q}_{j_{\sigma(1)}} \delta^{i_{\sigma(1)}}_{j_{\sigma(2)}} \dots \delta^{i_{\sigma(n-1)}}_{j_q} \delta^{i_{\sigma(n)}}_{j_{\sigma(n)}} \mathcal{A}_{1qq}(q,\sigma(1),\dots,\sigma(n-1),\bar{q},\sigma(n)) \\
&+ \left(\frac{-1}{N}\right)^2 \frac{1}{2!} \sum_{\sigma\in S_n} \delta^{i_q}_{j_{\sigma(1)}} \delta^{i_{\sigma(1)}}_{j_{\sigma(2)}} \dots \delta^{i_{\sigma(n-2)}}_{j_q} \delta^{i_{\sigma(n-1)}}_{j_{\sigma(n-1)}} \delta^{i_{\sigma(n)}}_{j_{\sigma(n)}} \\
&\qquad\qquad \mathcal{A}_{1qq}(q,\sigma(1),\dots,\sigma(n-2),\bar{q},\sigma(n-1),\sigma(n)) \\
&+ \dots \\
&+ \left(\frac{-1}{N}\right)^n \frac{1}{n!} \sum_{\sigma\in S_n} \delta^{i_q}_{j_q} \delta^{i_{\sigma(1)}}_{j_{\sigma(1)}} \dots \delta^{i_{\sigma(n-1)}}_{j_{\sigma(n-1)}} \delta^{i_{\sigma(n)}}_{j_{\sigma(n)}} \mathcal{A}_{1qq}(q,\bar{q},\sigma(1),\dots,\sigma(n)).
\end{aligned}
$$

Results for gg initiated processes[6]

- For external particles $n_{\text{ext}} \in [4, 14]$
- Blue: full colour   Green: NLC   Red: LC



[6] R. Frederix, T. Vitos arXiv:2109.10377

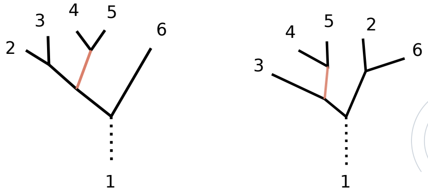
Integration

Amplitude evaluation

- Each of the $\mathcal{A}_{\sigma_k}$ computed in a recursive fashion
  $\rightarrow$ recycle parts of the amplitudes
- For a fixed colour-ordering (example for $n = 6$):

$$\begin{gathered} \{1,2\} \quad \{2,3\} \quad \{3,4\} \quad \{4,5\} \\ \{(1,2),(2,3)\} \quad \{(2,3),(3,4)\} \quad \{(3,4),(4,5)\} \\ \{((1,2),(2,3)),((2,3),(3,4))\} \quad \{((2,3),(3,4)),((3,4),(4,5))\} \\ \{(((1,2),(2,3)),((2,3),(3,4))),(((2,3),(3,4)),((3,4),(4,5)))\} \end{gathered} \tag{17}$$

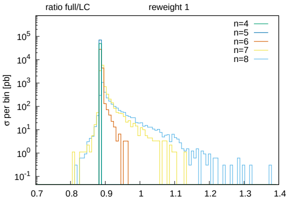
- Example: same current (4,5) appears in the two colour-orderings
  $\sigma = (1,2,3,4,5,6)$ and $\sigma = (1,3,4,5,2,6)$

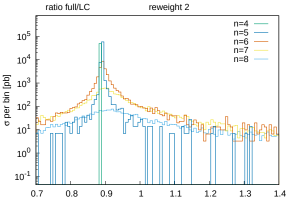


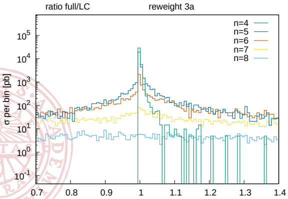
Reweighting the LC events to FC

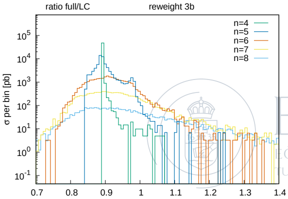
Obtaining NLC and FC accurate predictions from the LC events can be done in several ways:

1. Perform FC sum (sum over colour configurations)
2. Pick subset of permutations (using phase space symmetry)
3. a) Pick specific colour configuration randomly (consistent with colour ordering) or
   b) use an average of all colour configurations





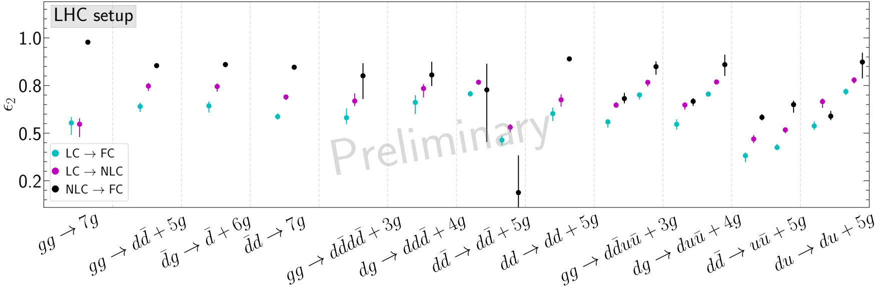




Weight spread of reweighting procedure

- Use a secondary unweighting efficiency to measure the spread of weights:

$$\epsilon_2 = \frac{1}{N} \sum_i \frac{r_i}{r_{\max}} \quad (18)$$



Weight spread of reweighting procedure

- Reweight factor for varying colour ordering for various one quark line processes: