Optimising loop amplitude evaluations

Simon Badger (University of Turin)

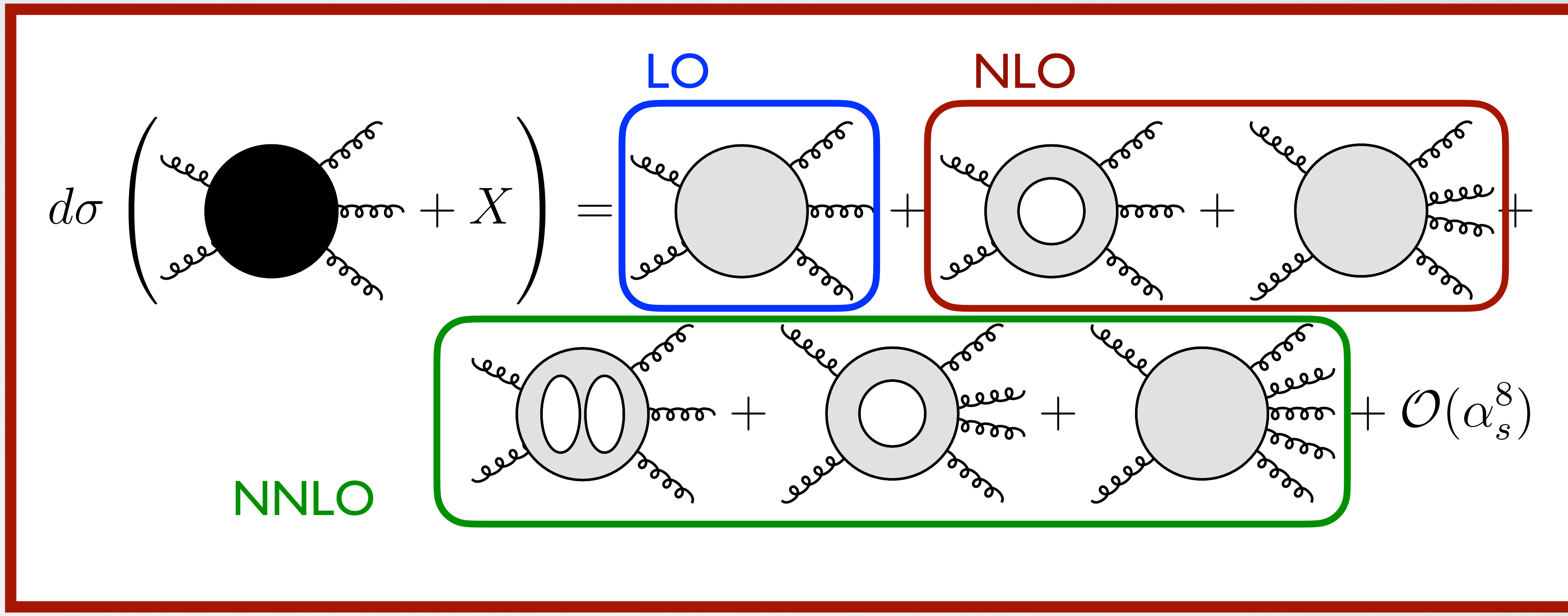
Event generators' and N(n)LO codes' acceleration
CERN
14th November 2023

well known problem:

multi-scale amplitudes difficult to integrate

CPU cost scales badly with higher orders

(especially when IR singular)



some solutions already exsist

automated and optimised MCs

multi-core processing

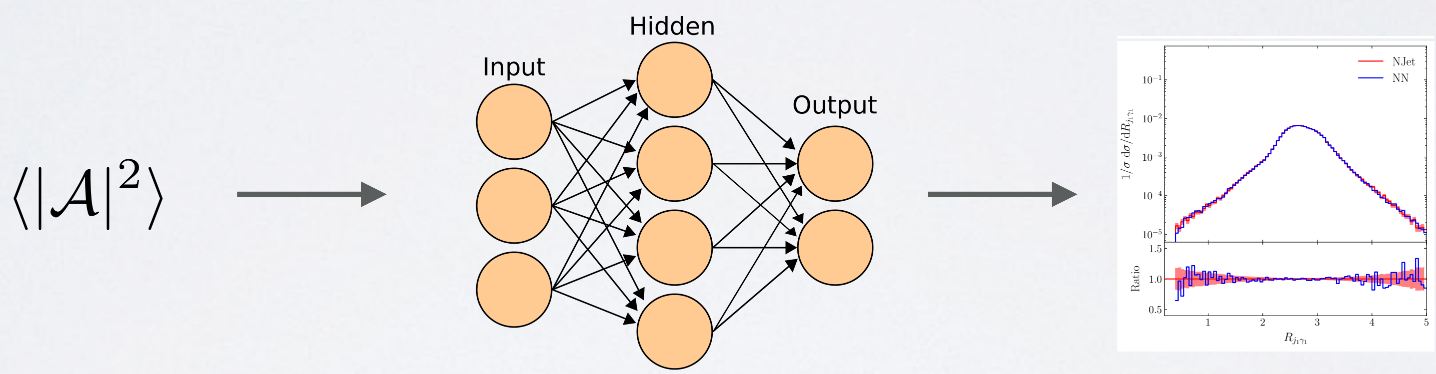
re-weighting lower orders

…as well as other methods discussed at this meeting

this talk: testing amplitude neural network for loop induced processes

gg →γγg and gg →γγgg

- Realistic hadron collider setup with SHERPA
- Precision evaluations: determining NN errors

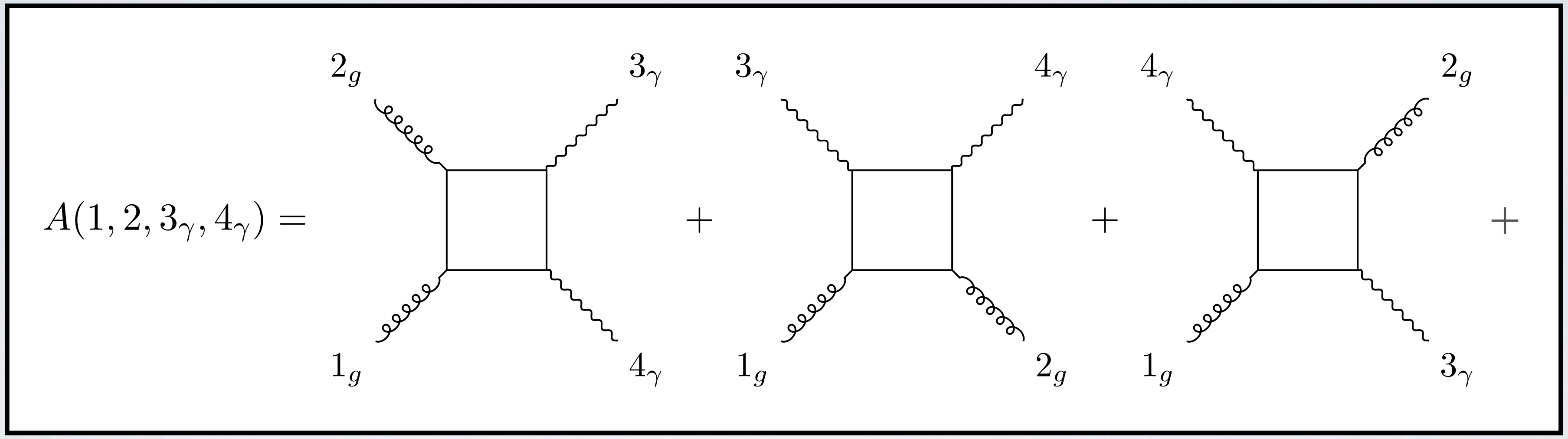


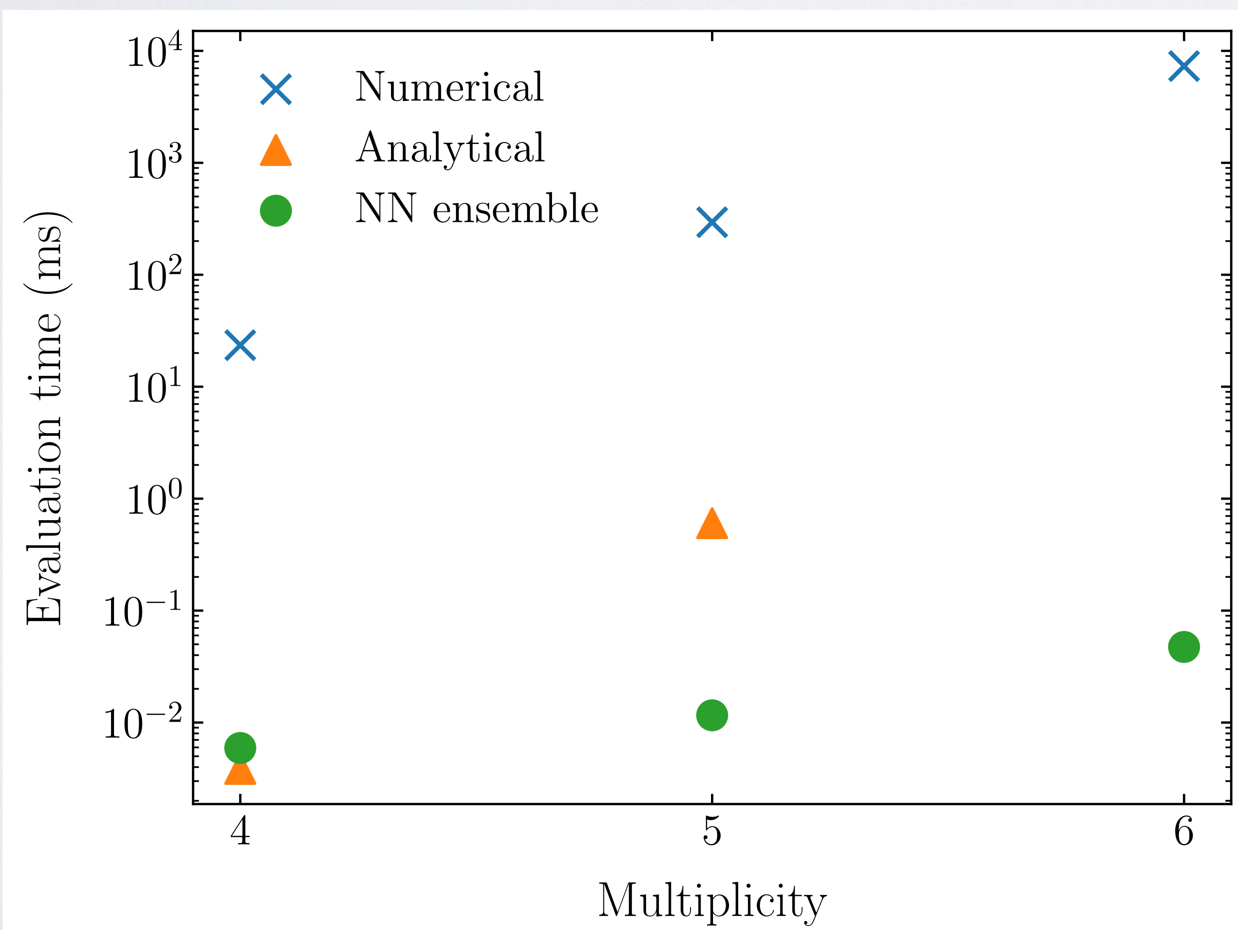
Aylett-Bullock, SB, Moodie JHEP 08 (2021) 066

SB, Butter, Luchmann, Pitz, Plehn SciPost Phys.Core 6 (2023) 034

- Reflections on experience at NNLO

How expensive are the loop amplitudes…





numerical evaluation with **NJET (V3)**
(SB, Biedermann, Moodie, Uwer, Yundin)
https://bitbucket.org/njet/njet/wiki/Home

(also includes some 2L 5pt amplitudes)

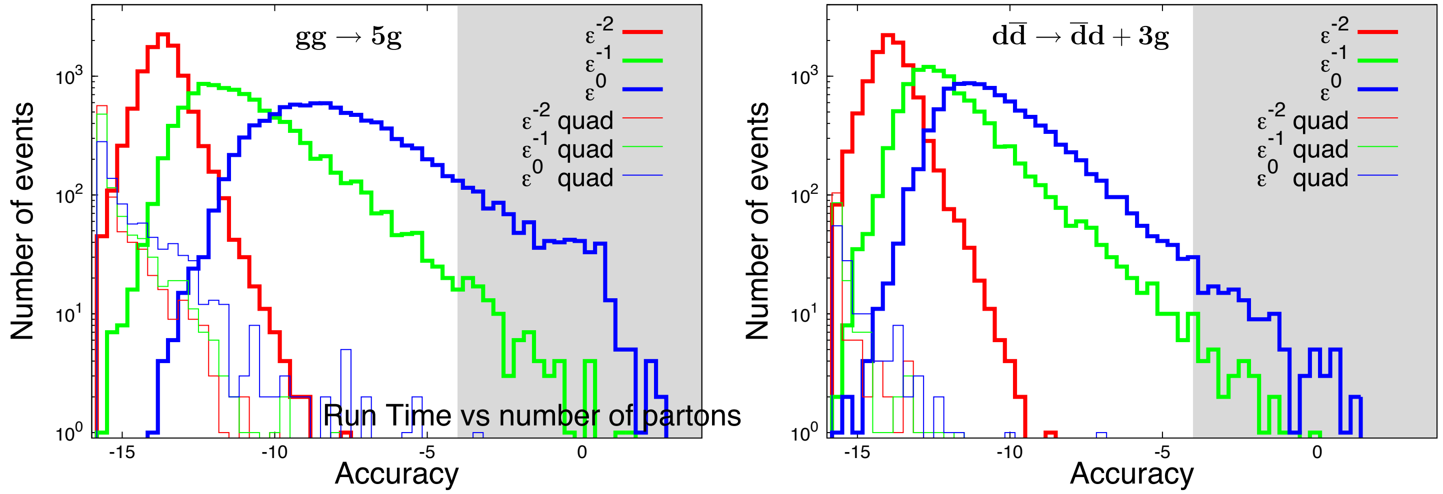
evaluation time including
error estimates

Precision is Important

standard one-loop matrix elements provide high precision evaluations:

OpenLoops, GoSam, Madloop, BlackHat, Recola, Helac-NLO…

one-loop pp → 5 partons with NJET



(a) (b)

Figure 5. Accuracy for 5-jet amplitudes: (a) shows the seven gluon process and (b) the $d\bar{d} \to \bar{d}d + 3g$ process. The thicker histograms show computations in double precision whereas the thinner curves show the distribution in quadruple precision for points which did not pass the relative accuracy of 10^{-4} when calculated in double precision. Red histograms show the $\frac{1}{\varepsilon^2}$ poles, green histograms the $\frac{1}{\varepsilon}$ and blue histograms the finite part of the amplitudes.

use the ability to switch numerical precision (e.g. qd for 32 or 64 digits)

dimension scaling tests, gauge invariance checks,…

Wishlist

an amplitude approximation which is:

- simpler to train/fit than generating events using traditional methods
- simplest model which fits a generic process
- reliable error estimates
- robust against changes in phase-space (cuts, jet algorithms, scale variations, etc.)
- simple to distribute

first attempt: e+e- → jets

SB, (Aylett-)Bullock [2002.07516]

- Single NN does badly
- Splitting IR sectors via FKS sectors improves reliablity
- **Error estimates** by varying model initialisation (ensemble of networks)
- K-factors work better than tree-level (no 1/s poles)
- Various tests suggest single run speed improvements at least x10

$$\mathcal{S}_{i,j} = \frac{1}{D_1 s_{ij}}, \quad D_1 = \sum_{i,j \in \mathcal{P}_{\mathrm{FKS}}} \frac{1}{s_{ij}},$$

$$\mathrm{d}\sigma^{(X)} = \sum_{i,j} S_{i,j}\,\mathrm{d}\sigma^{(X)},$$

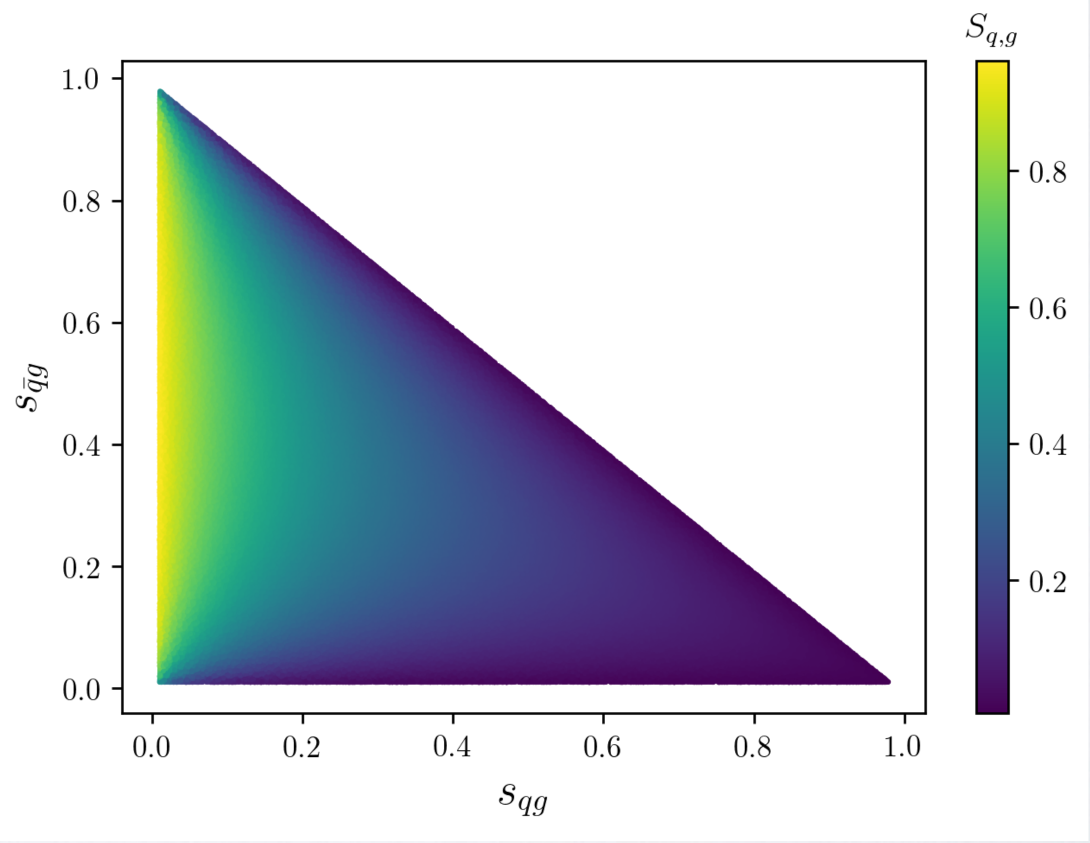
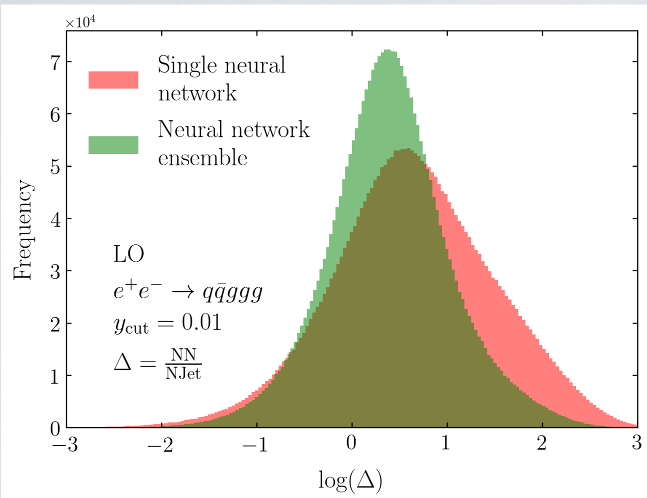
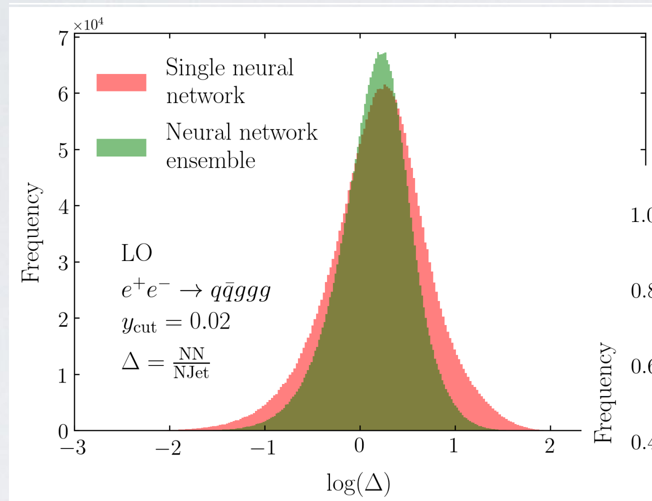


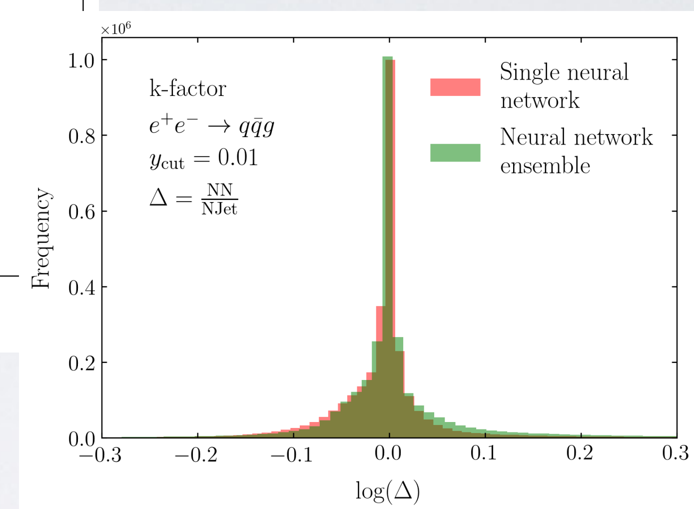
Figure 2: Behaviour of the $S_{q,g}$ FKS partition function

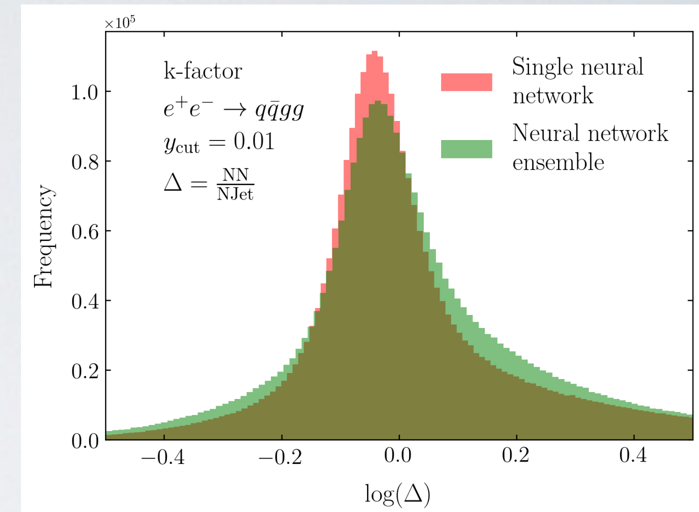
first attempt: e+e- → jets

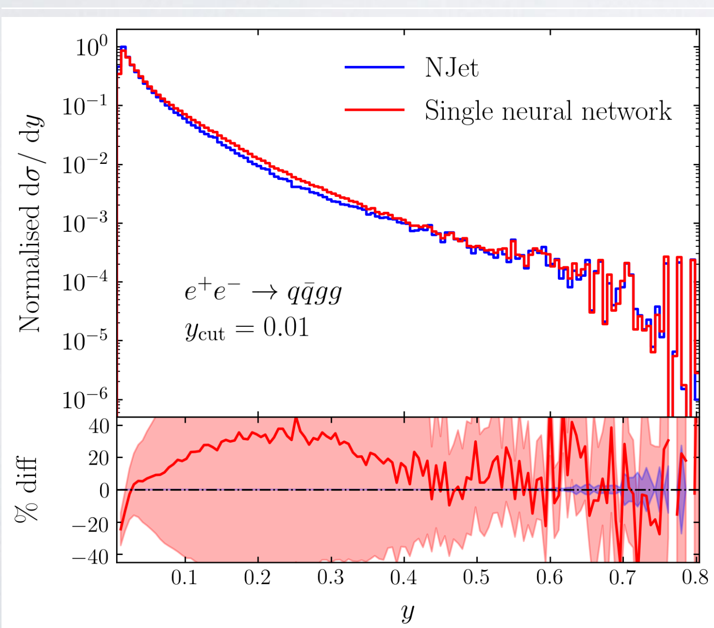
SB, (Aylett-)Bullock [2002.07516]

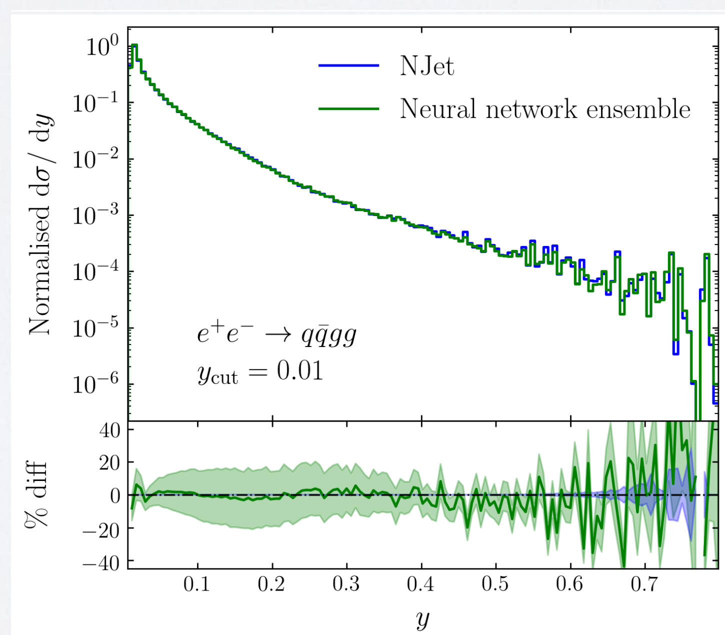




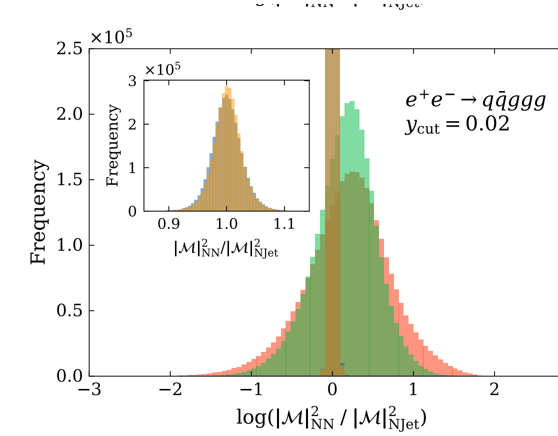








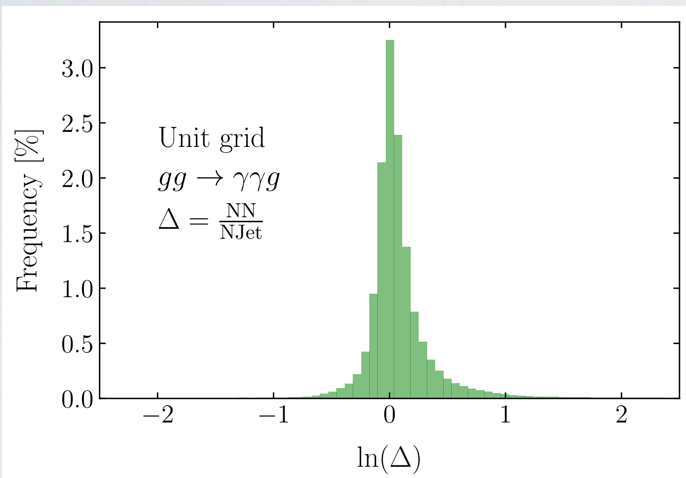
tree-level factorisation aware approach performs better



Maitre, Truong [2107.06625]

second attempt: gg → γγ+gluons

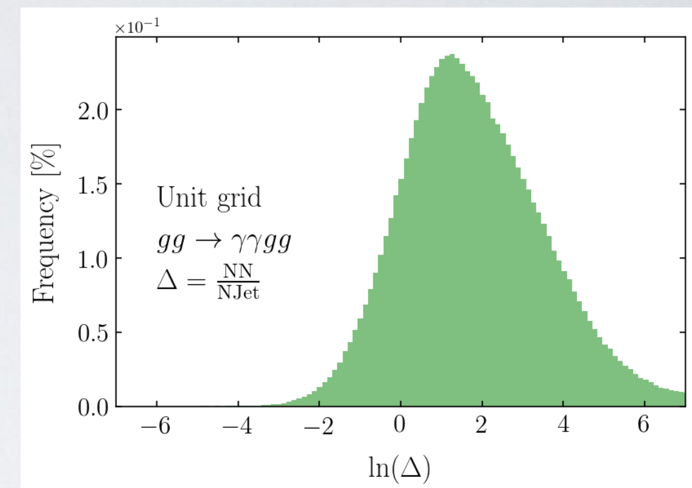
Aylett-Bullock, SB, Moodie [2106.09474]

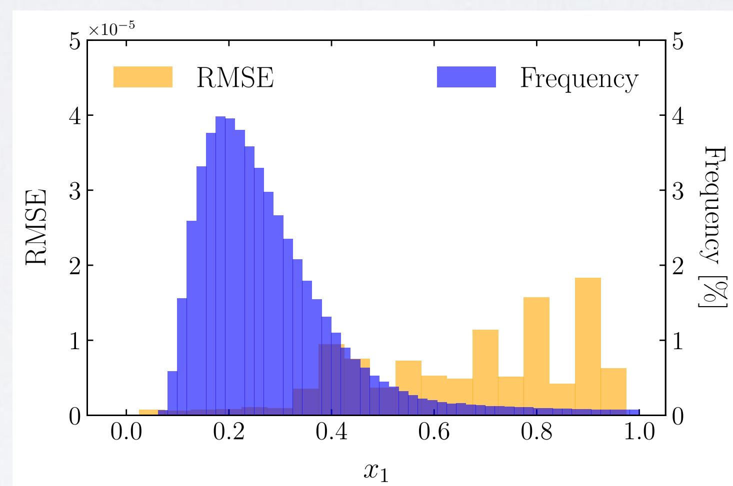


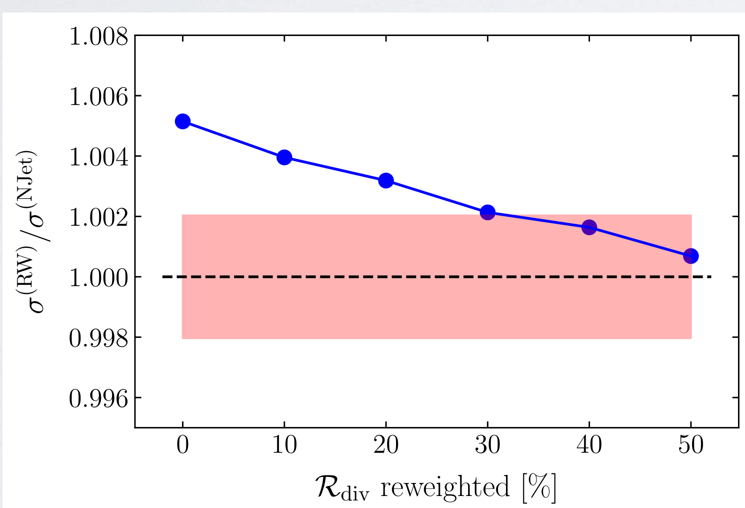
$p_{T,j} > 20\,\text{GeV}$ $R_{\gamma,j} > 0.4$ $|\eta_j| < 5$

$p_{T,\gamma_1} > 40\,\text{GeV}$ $R_{\gamma,\gamma} > 0.4$ $|\eta_\gamma| < 2.37$

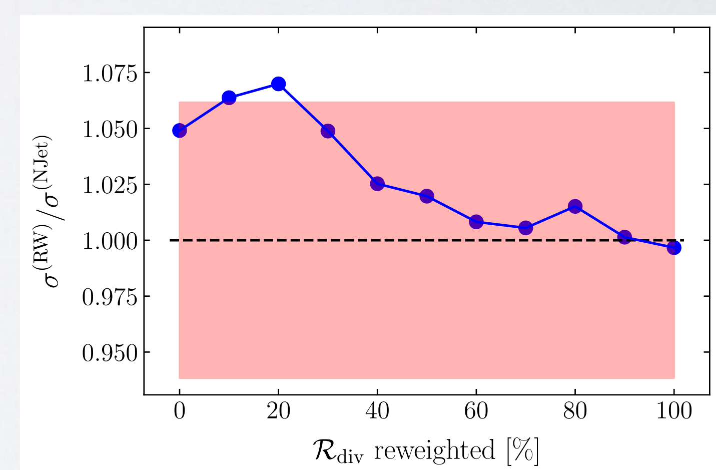
$p_{T,\gamma_2} > 30\,\text{GeV}$







7–9% of the total phase space



speed-up ~ $N_{\text{infer}}/N_{\text{train}}$

third attempt: gg → γγ+gluons with **Bayesian networks**

SB, Butter, Luchmann, Pitz, Plehn [2206.14831]

- weights and biases are replaced with (Gaussian) distributions
- optimised training times vs. pure ensemble approach

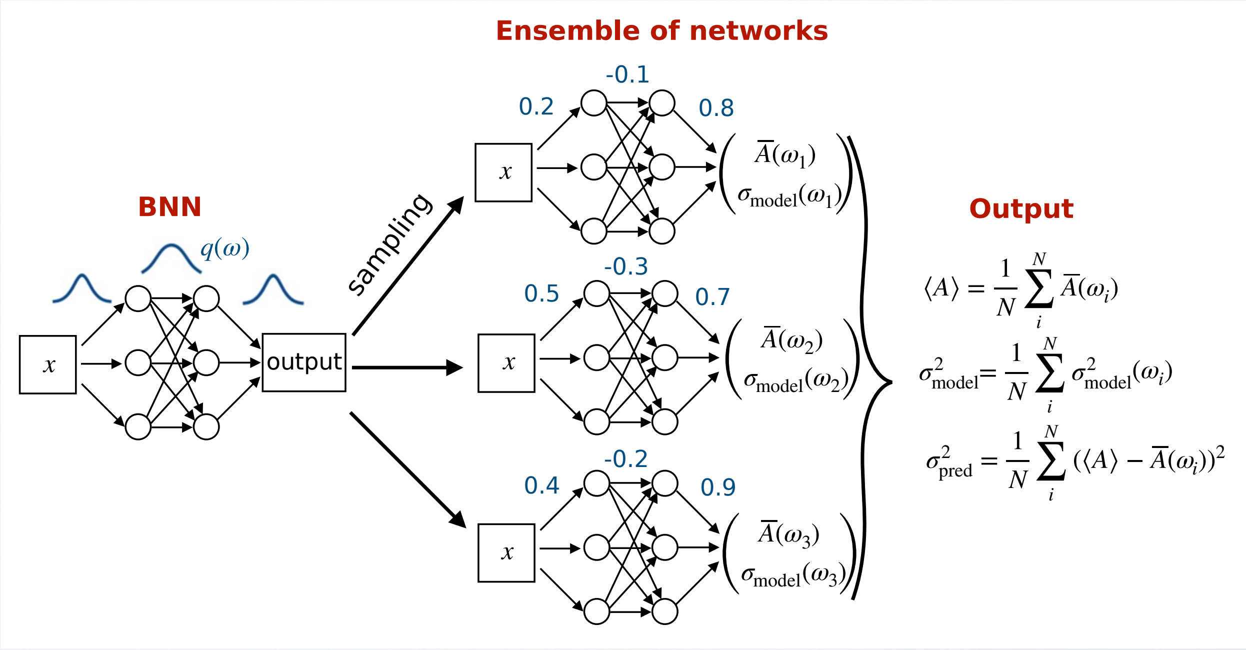
- better defined error estimates
- improved training via loss and performance boosting

single BNN

$$A_j \to \log\left(1 + \frac{A_j}{\sigma_A}\right)$$

~6k params for 2 → 3

~600k params for 2 → 4

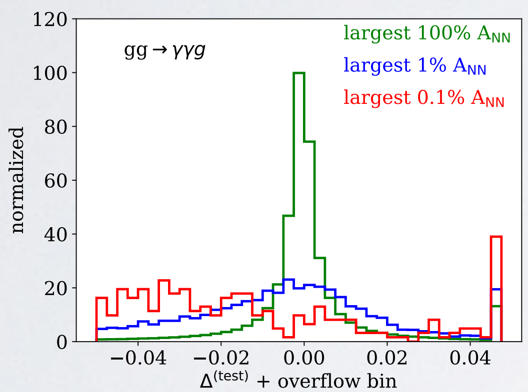


third attempt: gg → γγ+gluons with Bayesian networks

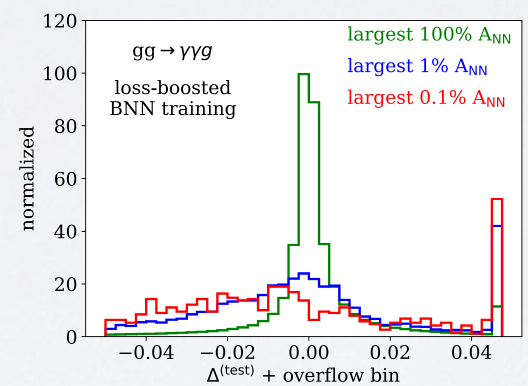
SB, Butter, Luchmann, Pitz, Plehn [2206.14831]

- weights and biases are replaced with (Gaussian) distributions
- optimised training times vs. pure ensemble approach

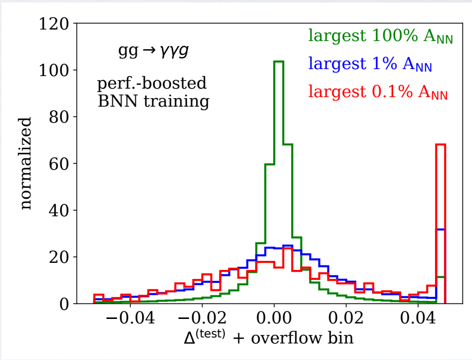
- better defined error estimates
- improved training via loss and performance boosting

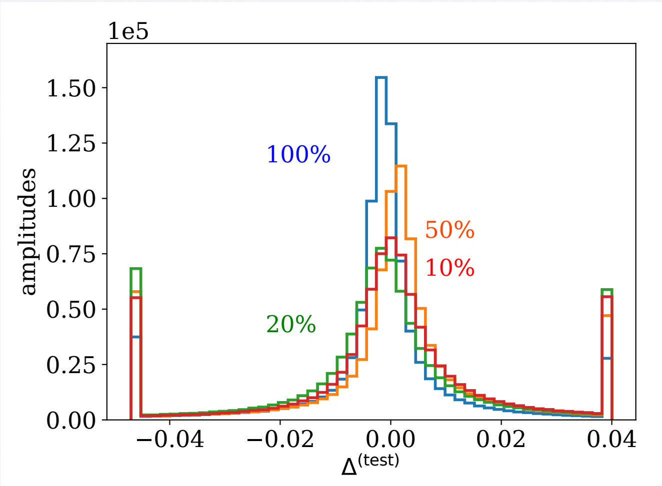


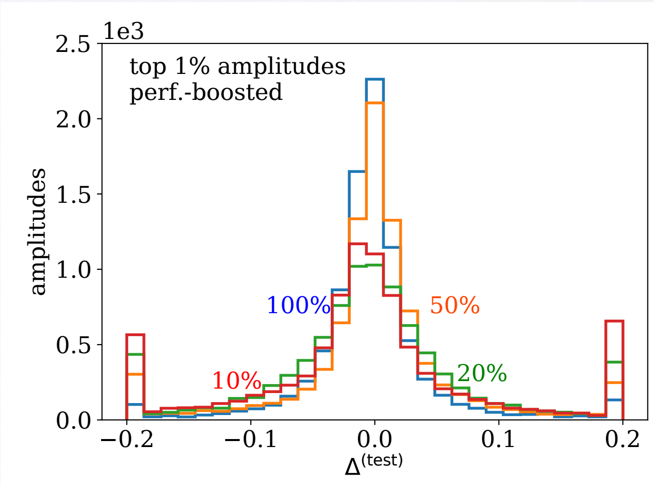
loss boosting



performance boosting



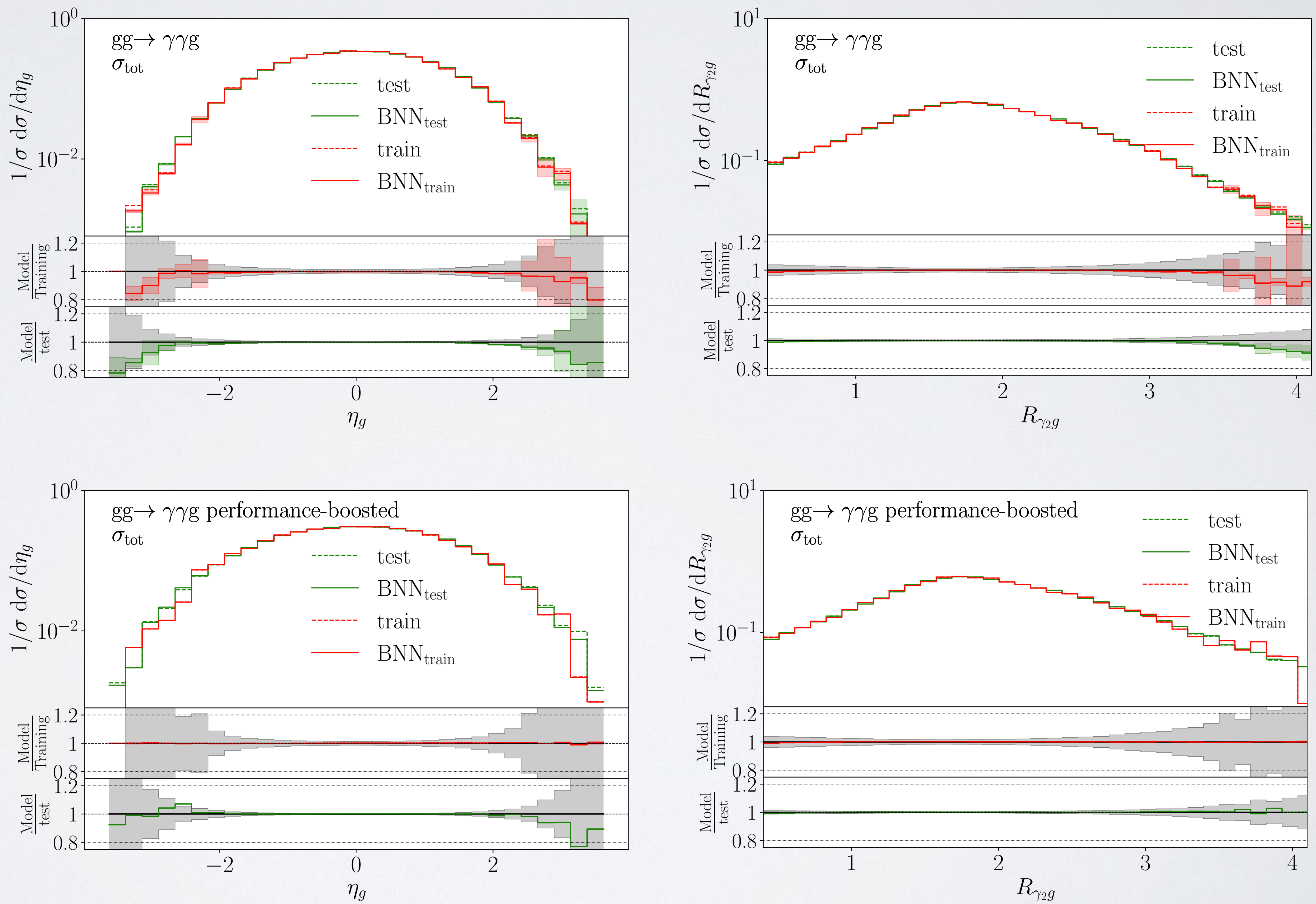




third attempt: gg → γγ+gluons with Bayesian networks

SB, Butter, Luchmann, Pitz, Plehn [2206.14831]

performance boosting reduces uncertainties in tails

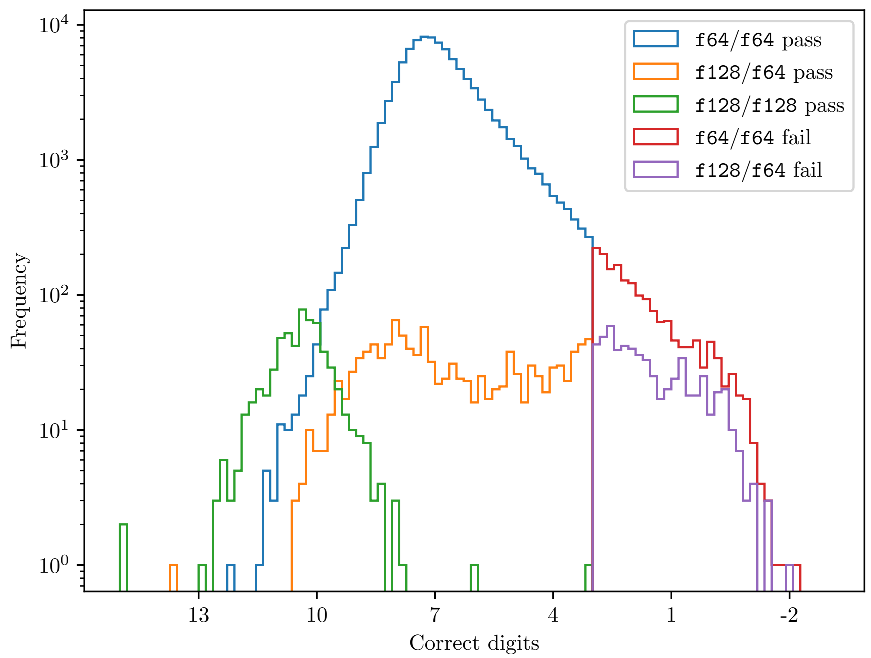


Higher Order Applications

precision frontier has moved to NNLO (QCD) and beyond

$2 \to 2$ @3-loop,
$2 \to 3$@ 2-loop

e.g. 2L $gg \to \gamma\gamma g$



SB, Gehrmann, Marcoli, Moodie 2109.12003

reasonable performance for 2→3 with **analytic** finite remainders (extracted using Finite Field sampling)

$$A^{(L),k}_{i;j} = \sum_{s=-2L}^{o(L)} \sum_r \epsilon^s \, c_{r,s}(\vec{x}) \, \mathrm{mon}_r(f, \mathfrak{c})$$

pentagon functions [Chicherin, Sotnikov, Zoia]

precision requirements more subtle

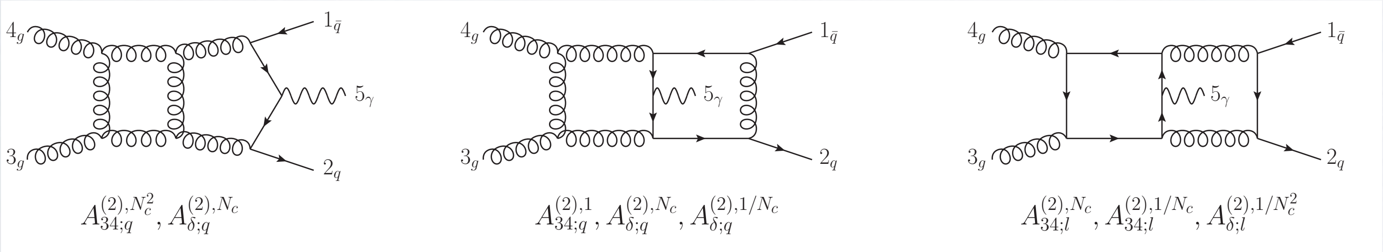
instabilities can be in both rational coefficients and special functions

All Two-Loop Feynman Integrals for Five-Point One-Mass Scattering, Abreu et al. [2306.15431]

Single Photon plus Two-Jets

SB, Czakon, Hartanto, Moodie, Peraro, Poncelet, Zoia [2304.06682]

first full colour 2 → 3 differential cross section



IR subtraction with STRIPPER approach

analytic 2L finite remainders

comparison with ATLAS [1912.09866]

1. We require at least two jets defined with the anti-k_T algorithm [106] for jet radius $R = 0.4$ that have minimal transverse momentum of $p_T(j) > 100$ GeV and maximal rapidity $|\eta(j)| < 2.5$.

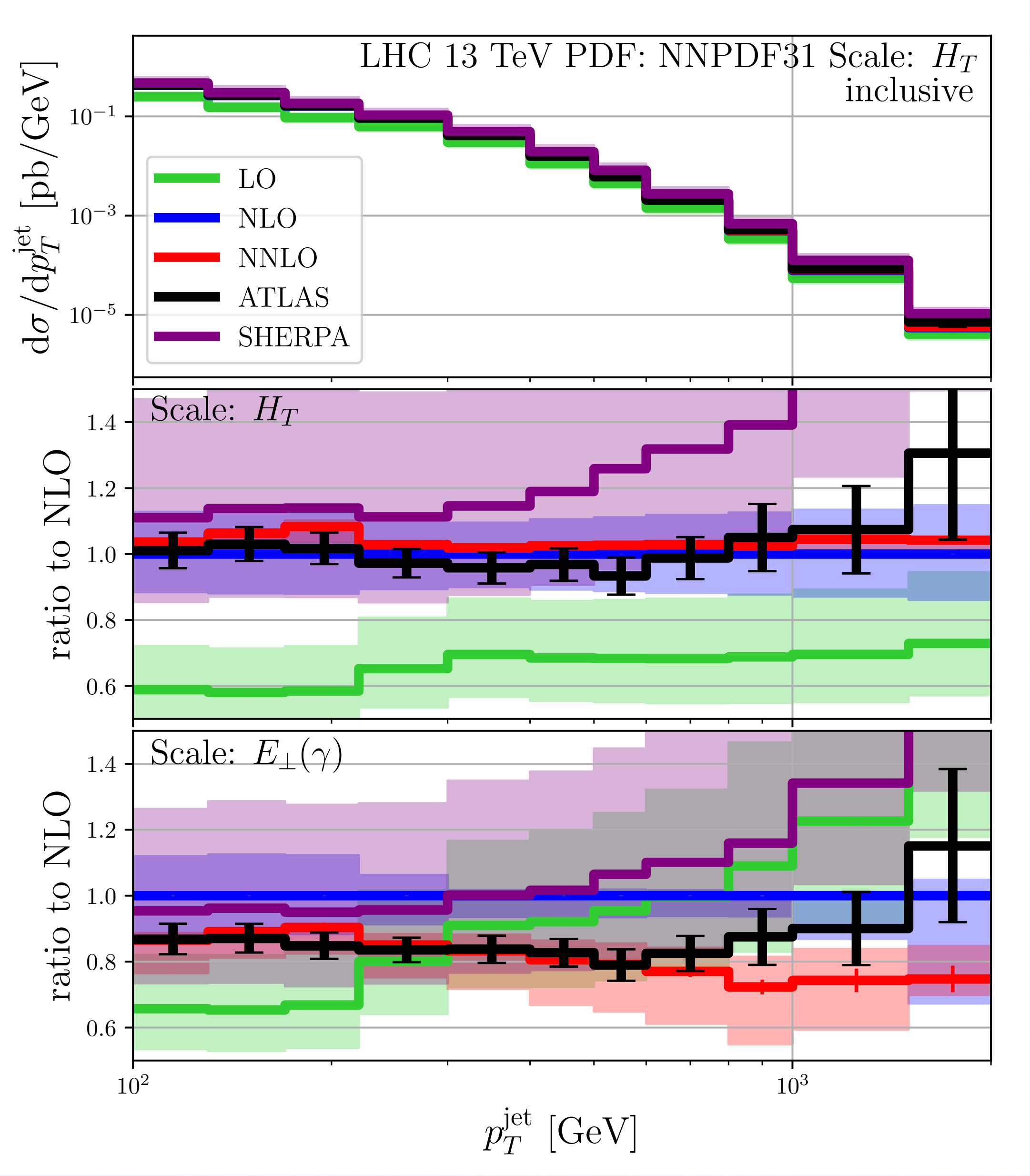
2. The identified jets must be separated from the photon by $\Delta R(\gamma, j) > 0.8$.

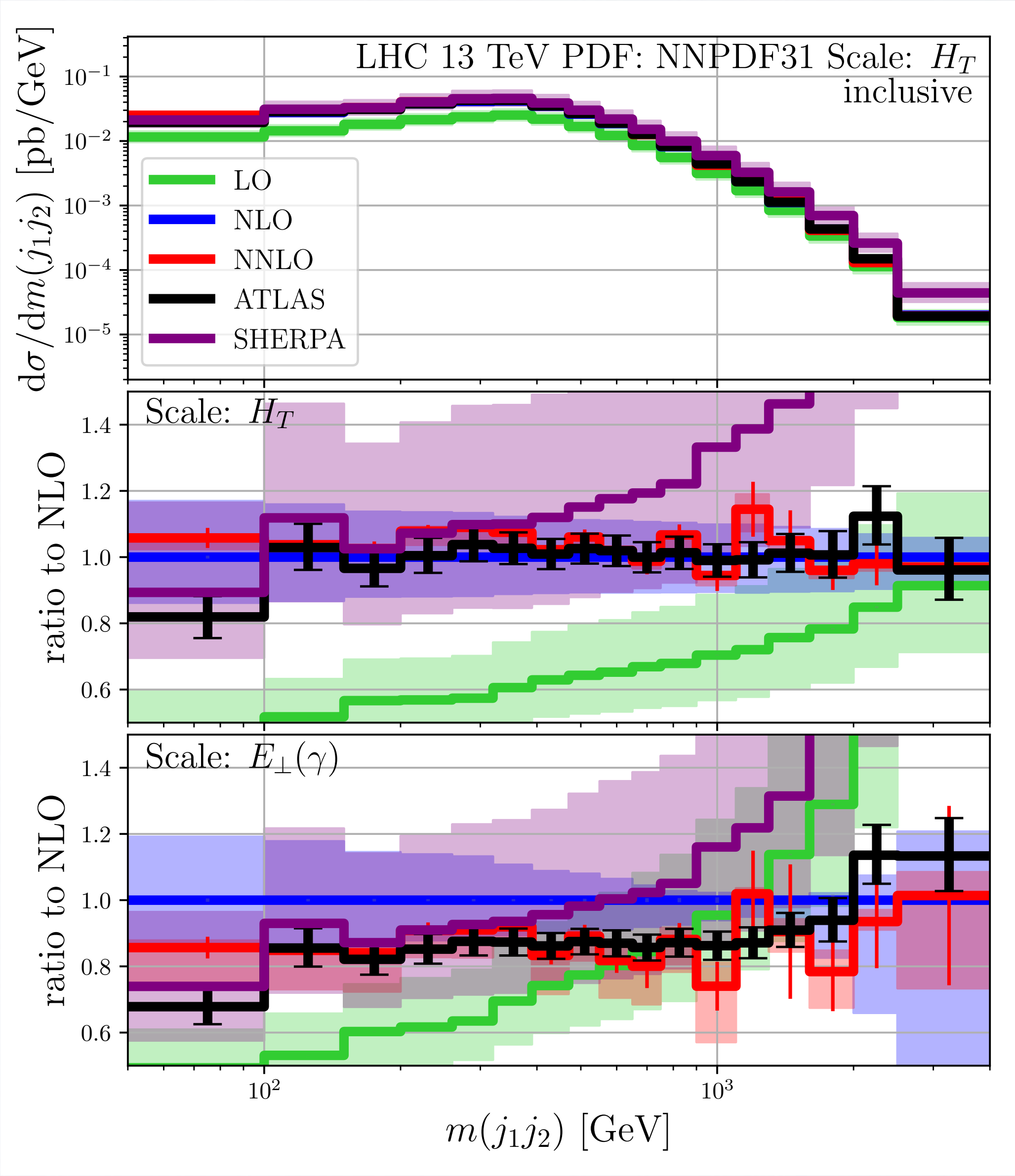
3. One isolated photon must be present in the final state with $E_\perp(\gamma) \geq 150$ GeV, $|\eta(\gamma)| \leq 2.37$ excluding $1.37 \leq |\eta(\gamma)| \leq 1.56$.

RV and RR using ME from OpenLoops and AvH libraries

Single Photon plus Two-Jets

SB, Czakon, Hartanto, Moodie, Peraro, Poncelet, Zoia [2304.06682]

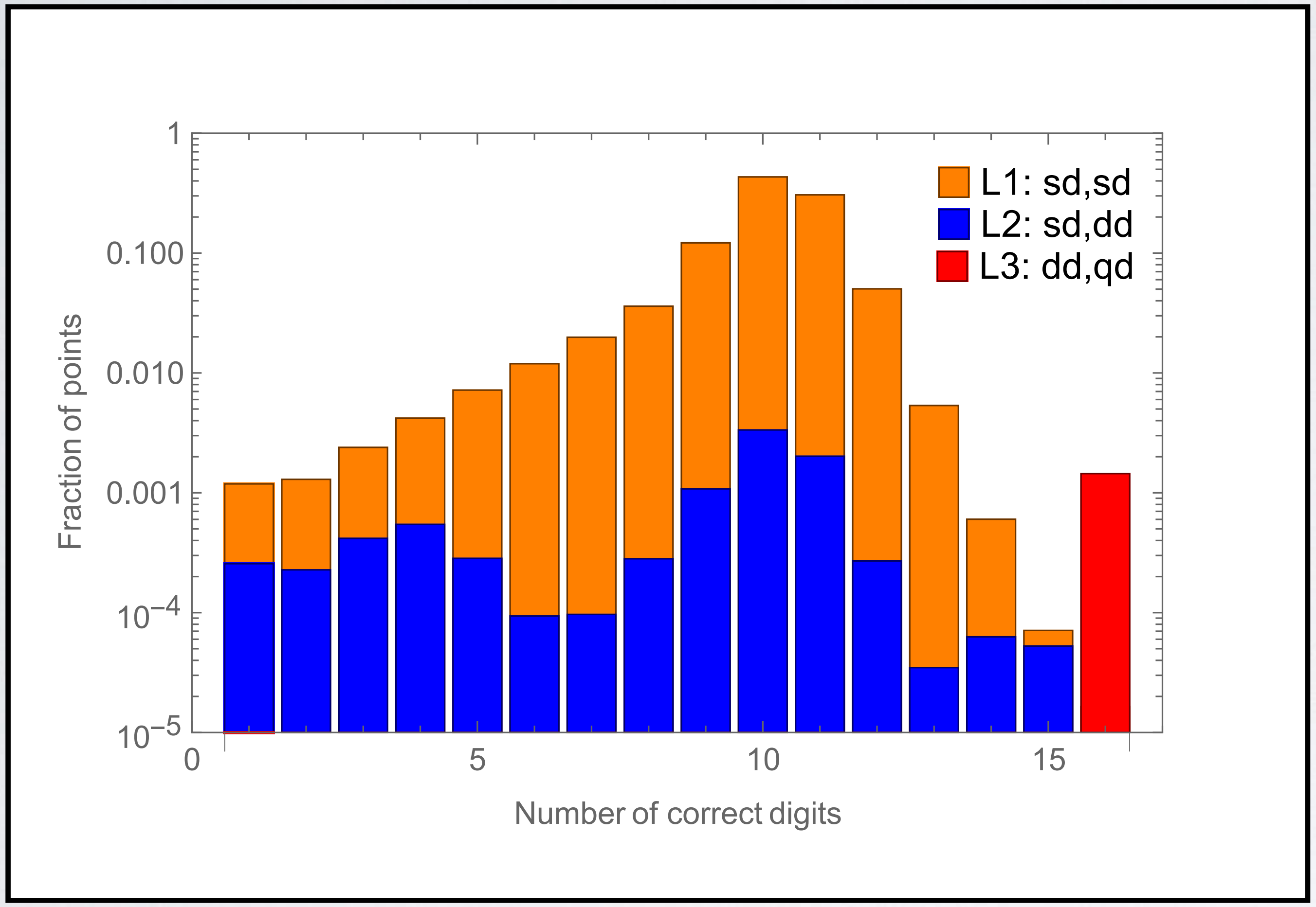




excellent overall agreement with data

Single Photon plus Two-Jets

SB, Czakon, Hartanto, Moodie, Peraro, Poncelet, Zoia [2304.06682]



total cost ~550 kCPUh

VV:RV:RR ~ 1: 10: 40

Note: the 1st time is always a challenge - many improvements possible

double virtual call average ~16s per point
[NB: 1 double precision call O(1s)]

Future Applications

speeding up amplitude calls with NN looks like a viable option for leading order codes

I would still be interested in better control for the number of correct digits

speeding up amplitude virtual amplitudes can help, but need to make impact on real radiation

two-loop 2 → 3 not dominated by virtuals, largely thanks to well studied special function basis

This is (probably) not going to continue: pp → ttj, pp→ ttH, → WWj

at some stage analytic formula unfeasible - must take numerical route

special function basis in 'pentagon function' form will not be possible (eg. if there are elliptic structures)

expect dramatic change in evaluation time

Conclusions

I still have some questions…

phase-space : how can we use amplitude values to minimize the required number of training points

squared amplitudes vs (ordered) helicity amplitudes

analytic vs. numeric: can we improve the architecture to better satisfy amplitude properties? (c.f. bootstrap techniques)