Status and prospects of SHERPA

QCD@LHC 2023 in Durham, UK

September 5 2023

Enrico Bothmann (ITP, U Göttingen)
on behalf of the SHERPA collaboration

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

MCnet

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

The SHERPA event generator framework

v2.2 release series [SHERPA collab. 1905.09127]

- Two multi-purpose matrix element (ME) generators:
  AMEGIC, COMIX

- Two parton showers (PS) generators:
  CSSHOWER, DIRE

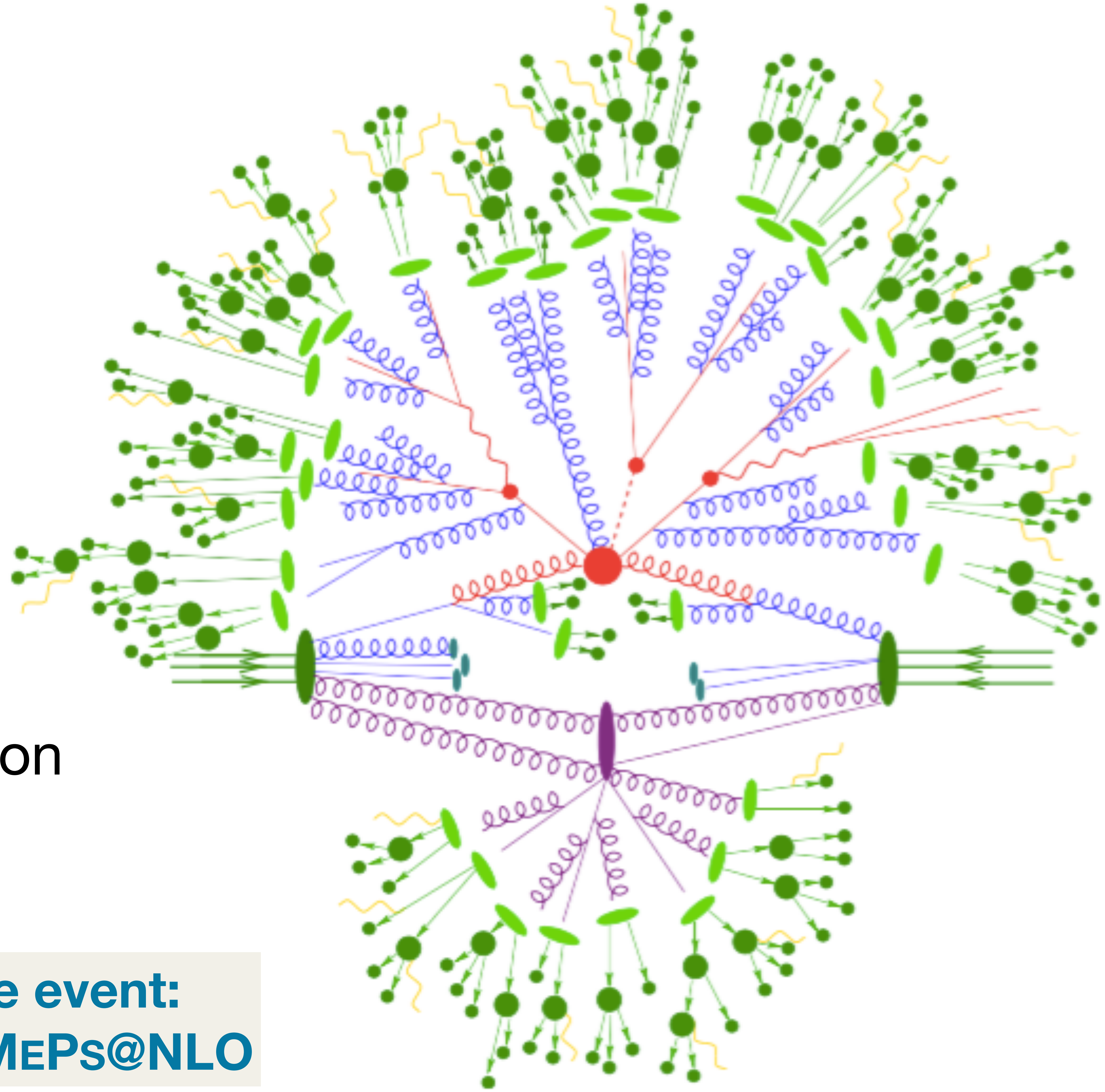
- A multiple interaction simulation à la PYTHIA

- A cluster fragmentation module

- A hadron and τ-lepton decay package

- A higher-order QED generator using YFS resummation

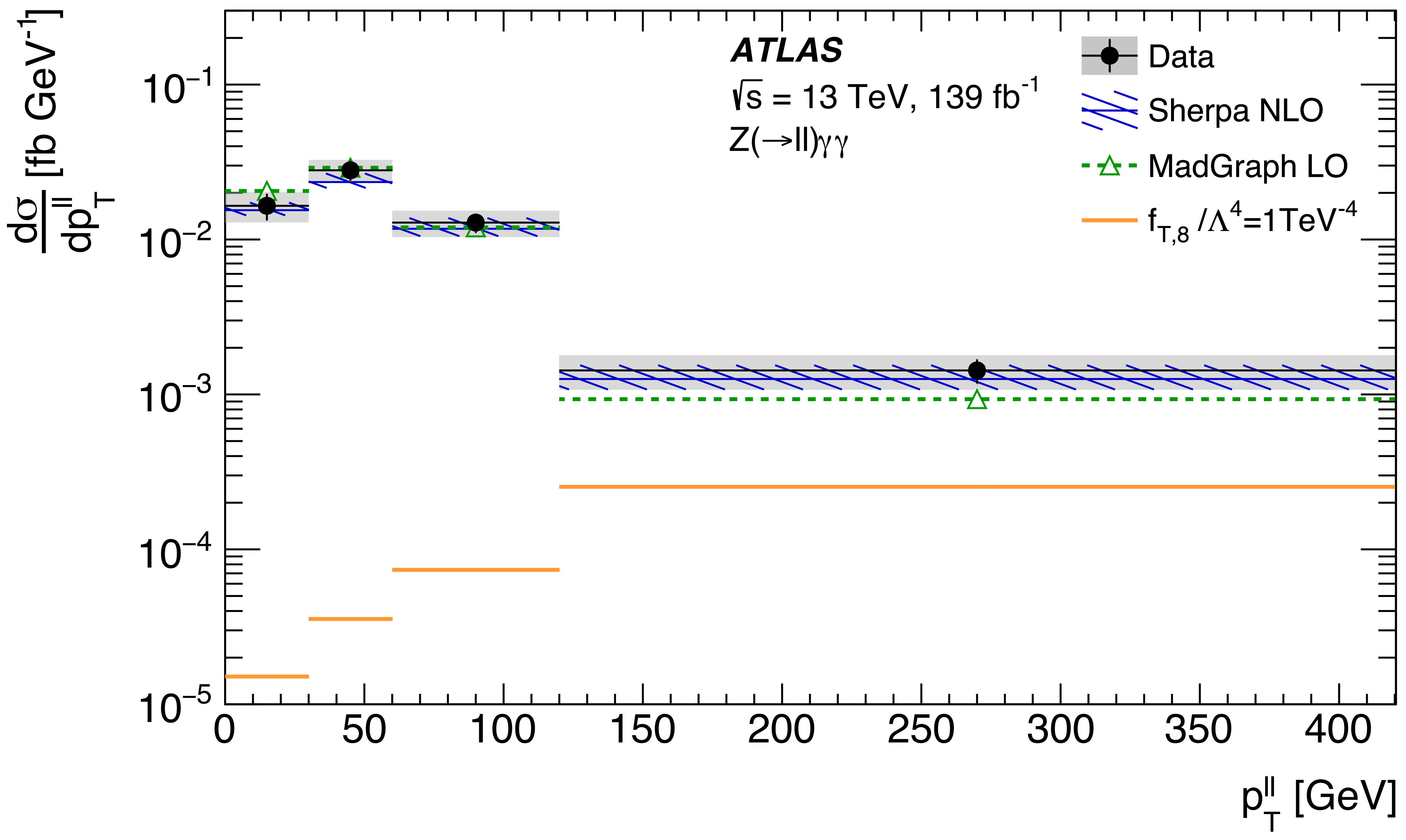
- Many add-ons

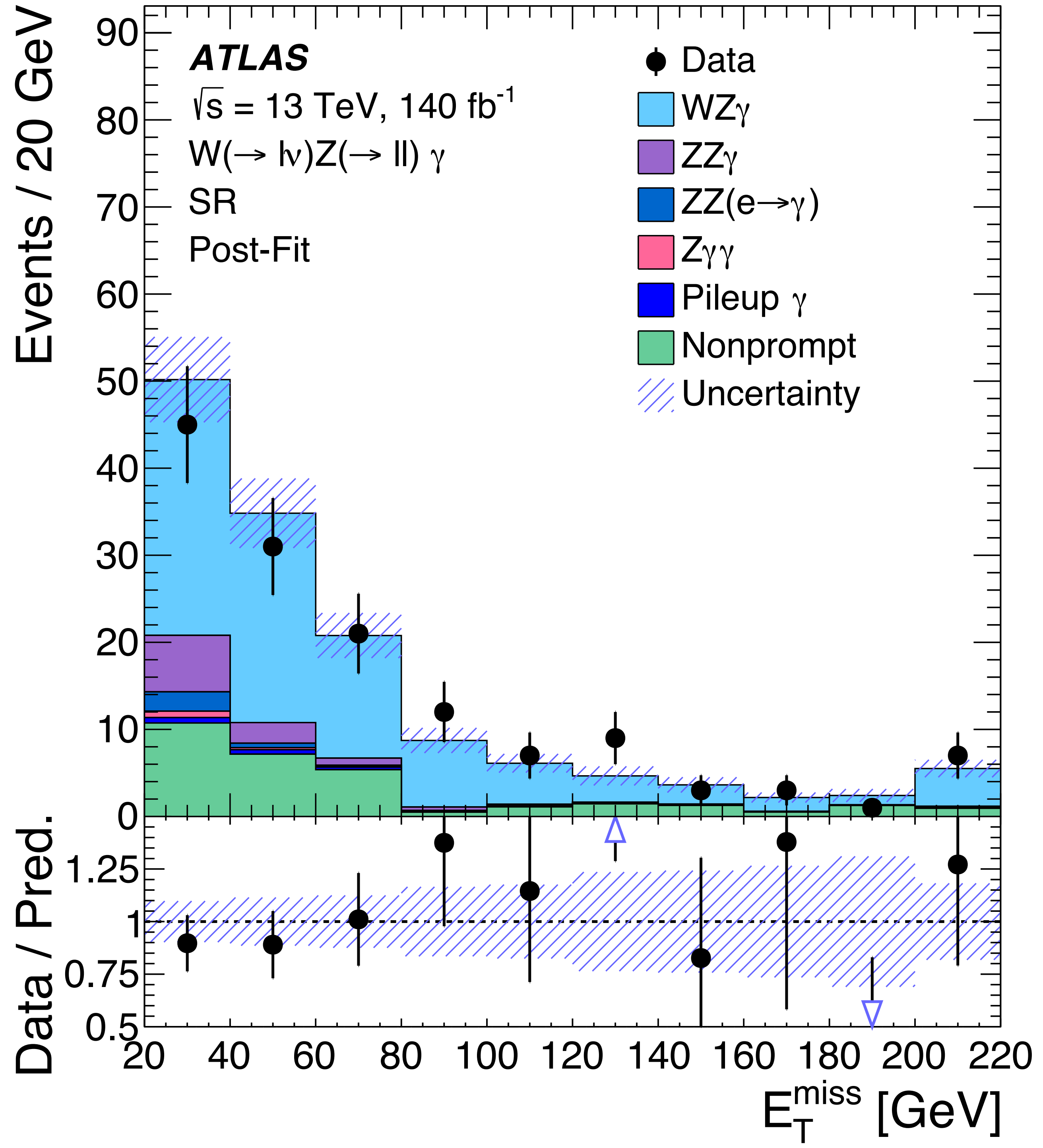


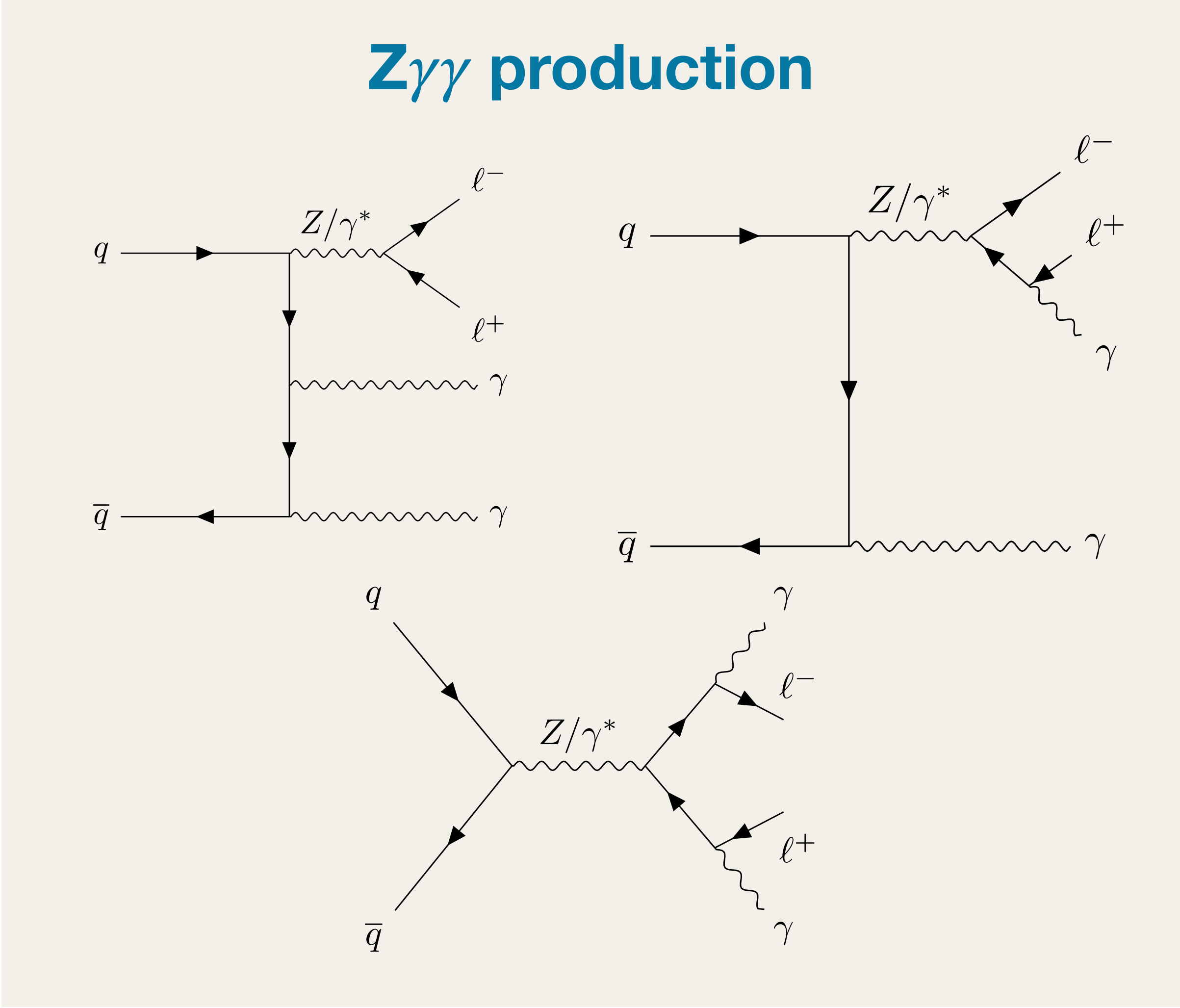
SHERPA's Traditional strength is the perturbative part of the event:
LO, NLO, NNLO, LOPS, NLOPS, NNLOPS, MEPS, MENLOPS, MEPS@NLO

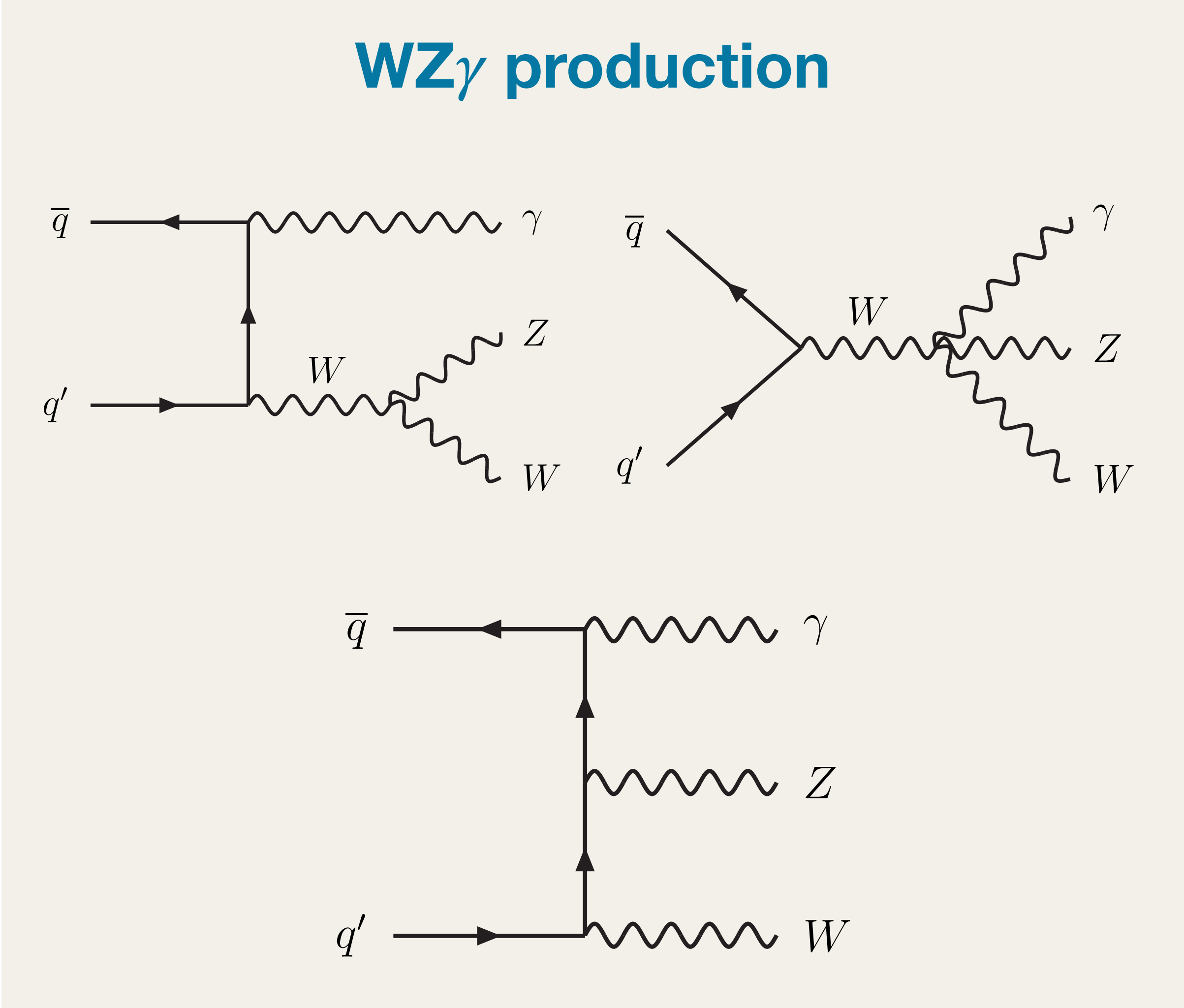
13 TeV Showcase I – rare triboson processes

Measure $Z\gamma\gamma$, observe $WZ\gamma$ and $W\gamma\gamma$ production [ATLAS 2211.14171, 2305.16994, 2308.03041]







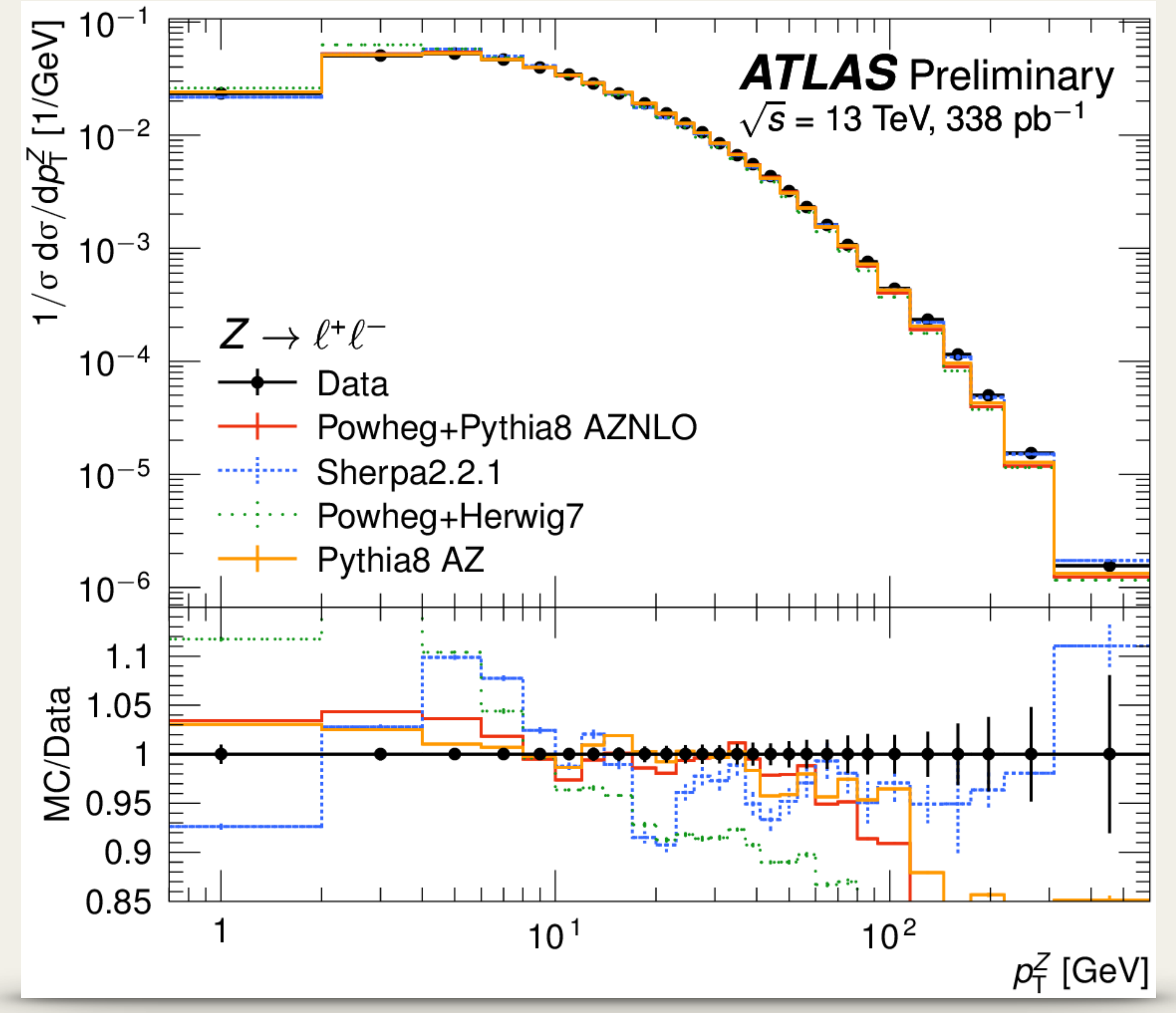


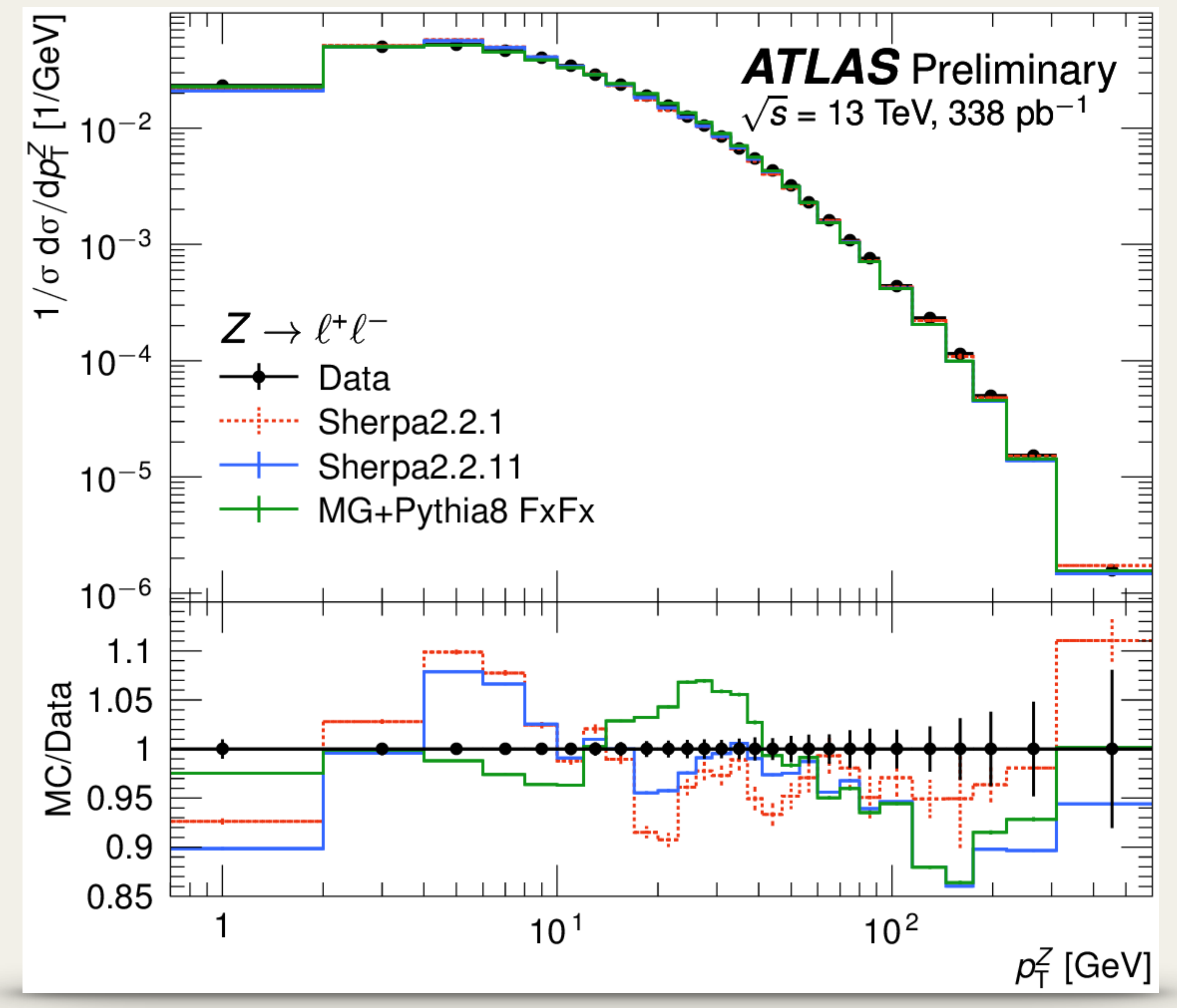
- $Z\gamma\gamma$ measurement: Sherpa 2.2.10 used for signal (0j@NLO+1,2j@LO), backgrounds $Z\gamma$+jets, ZZ, $WZ\gamma$, γ+jets and $\gamma\gamma$+jets
- tightened constraints on dimension-8 EFT operators

- $WZ\gamma$ observation: Sherpa 2.2.11 used for signal (0j@NLO+1,2j@LO), backgrounds $ZZ\gamma$, $Z\gamma\gamma$
- $W\gamma\gamma$ observation: Sherpa 2.2.10 used for signal (0j@NLO+1,2j@LO), backgrounds $WW\gamma$, $Z\gamma$ and $Z\gamma$

13 TeV Showcase II – precision measurements

V p_T [ATLAS-CONF-2023-028]





- Z p_T ultimate precision observable
- old Sherpa 2.2.1 sample shows sizable deviations at low p_T
  - deviation around 20 GeV reduced by improved splitting kernels in Sherpa 2.2.11 sample [2112.09588]
  - expect improvements at even lower p_T from upcoming Sherpa 3.0

(N)NLO + NLL′ accurate predictions for plain and groomed 1-jettiness in neutral current DIS → Daniel's talk (Wed)

Roadmap

Upcoming releases

2.3.0

- new HPC-ready HDF5 event pipeline;
  - ME-level unweighted events encoded in HDF5 event files; builds on [Höche 1905.05120]
  - later read in files for fast generation of merged+matched particle level events
    - Bonus: reuse of expensive ME-level events samples, e.g. for fast shower/hadronisation uncertainty studies etc.
  - https://zenodo.org/record/7754187 for sample H+jets files
  - publication+release very soon

3.0.0

- beta1 available for download and testing
- fully general & automated fixed-order NLO EW support [Schönherr 1712.07975]
- fully general & automated NLL EWSudakov corrections in all event generation modes (MEPS@NLO, …) [EB, Napoletano 2006.14635], [EB et al. 2111.13453]
- DY NNLO, EPA support, photon LHAPDFs, instantons, polarized cross sections for massive vector bosons …
- rewritten+retuned soft QCD, MPI, MinBias, Hadronisation, colour reconnection model [Chahal, Krauss 2203.11385]
- modernised: YAML input format, CMake build system, sphinx manual
- later in 3.x: Full NLL Alaric shower [Herren et al. 2208.06057] → Daniel's talk on Monday

Status and prospects of SHERPA

Efficiency

Precision

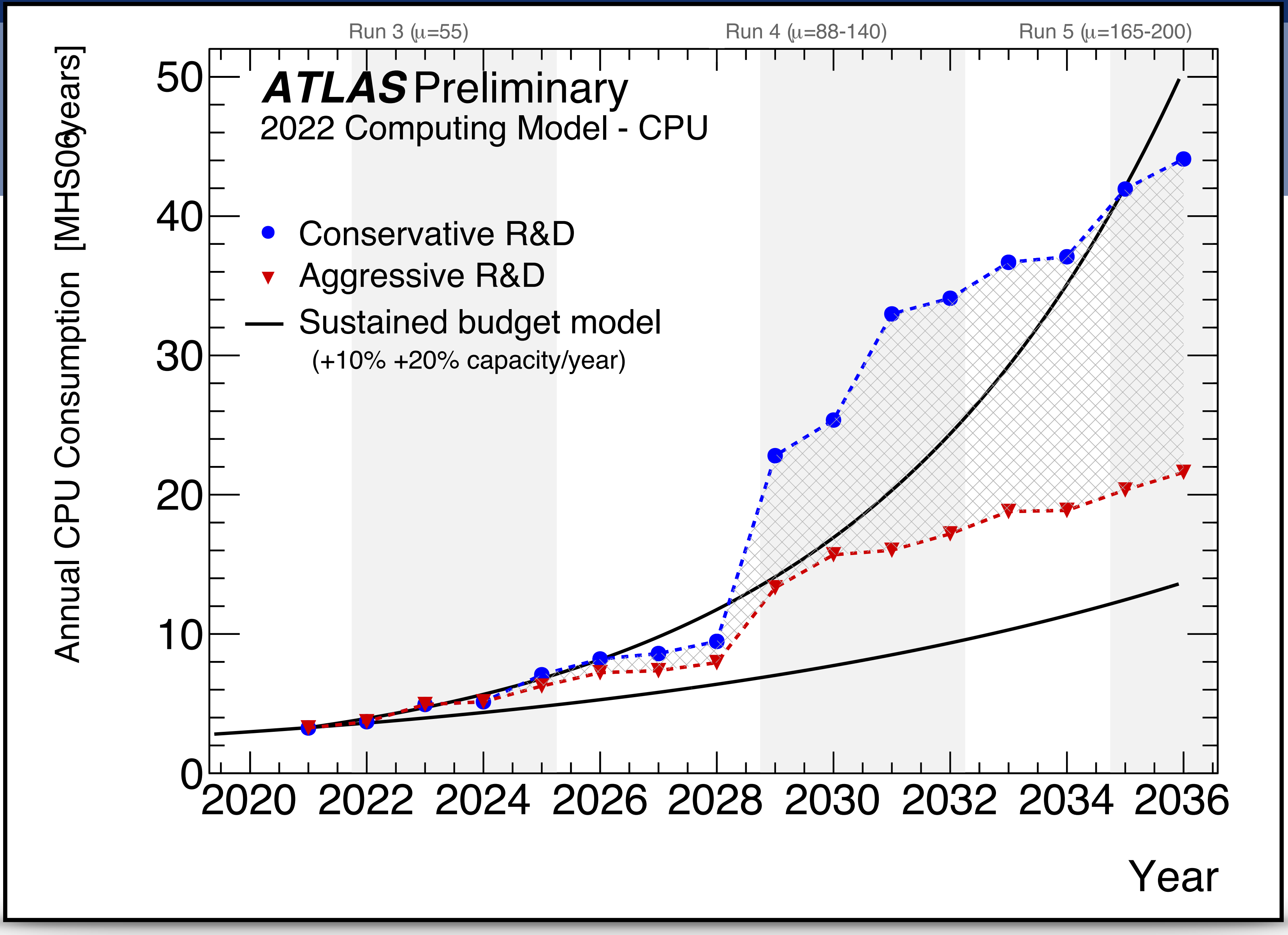
SHERPA+LHADPF Performance for (HL-)LHC

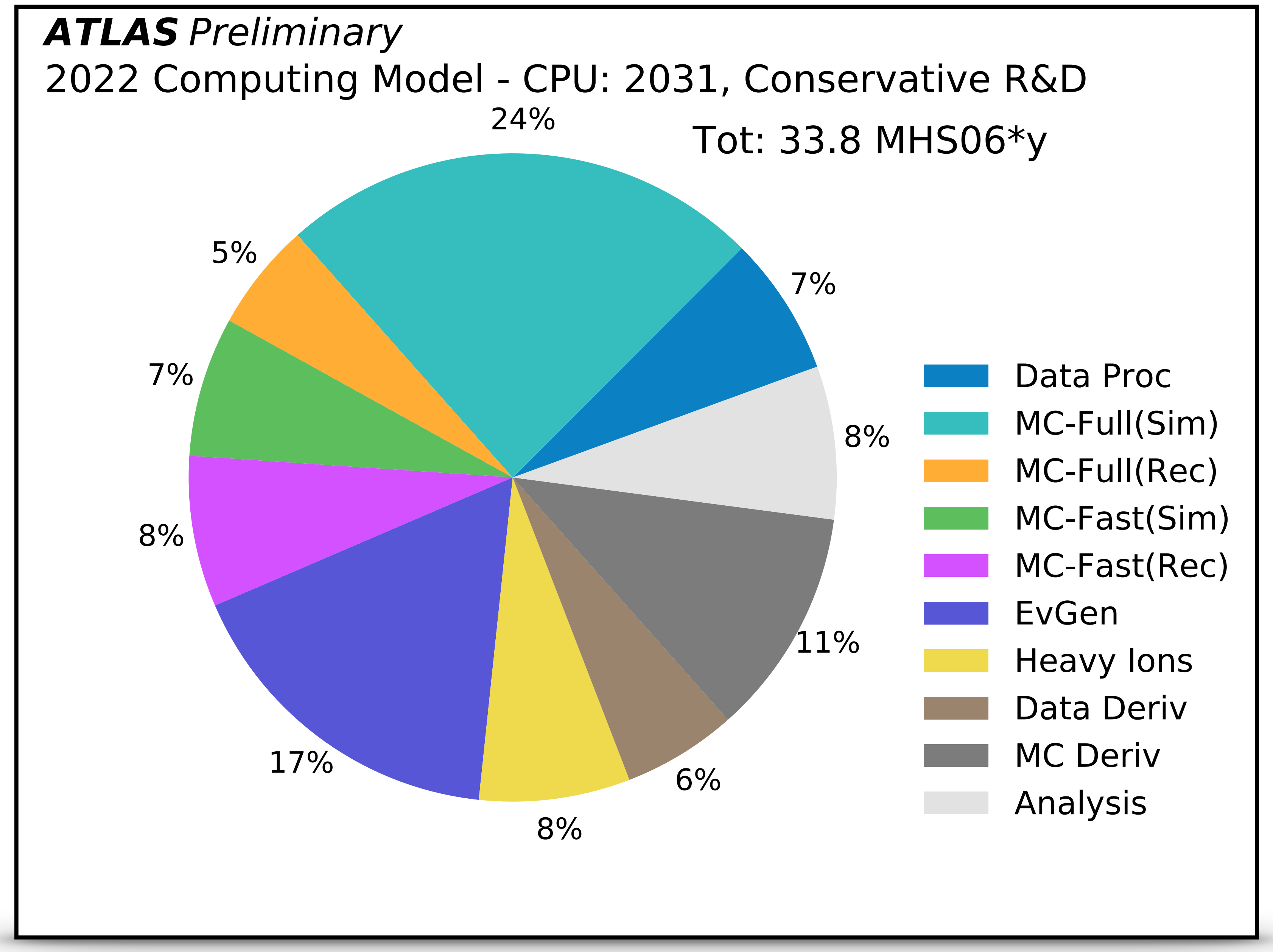
Overall profiling and tuning [EB et al. 2209.00843]

- MC event generation uses significant+increasing resources
- (HL-)LHC measurements in danger of being limited by MC statistics
- Explore reduction of CPU footprint for heaviest use cases, e.g. ATLAS default setup $Z + 0,1,2j$@NLO $+ 3,4,5j$@LO

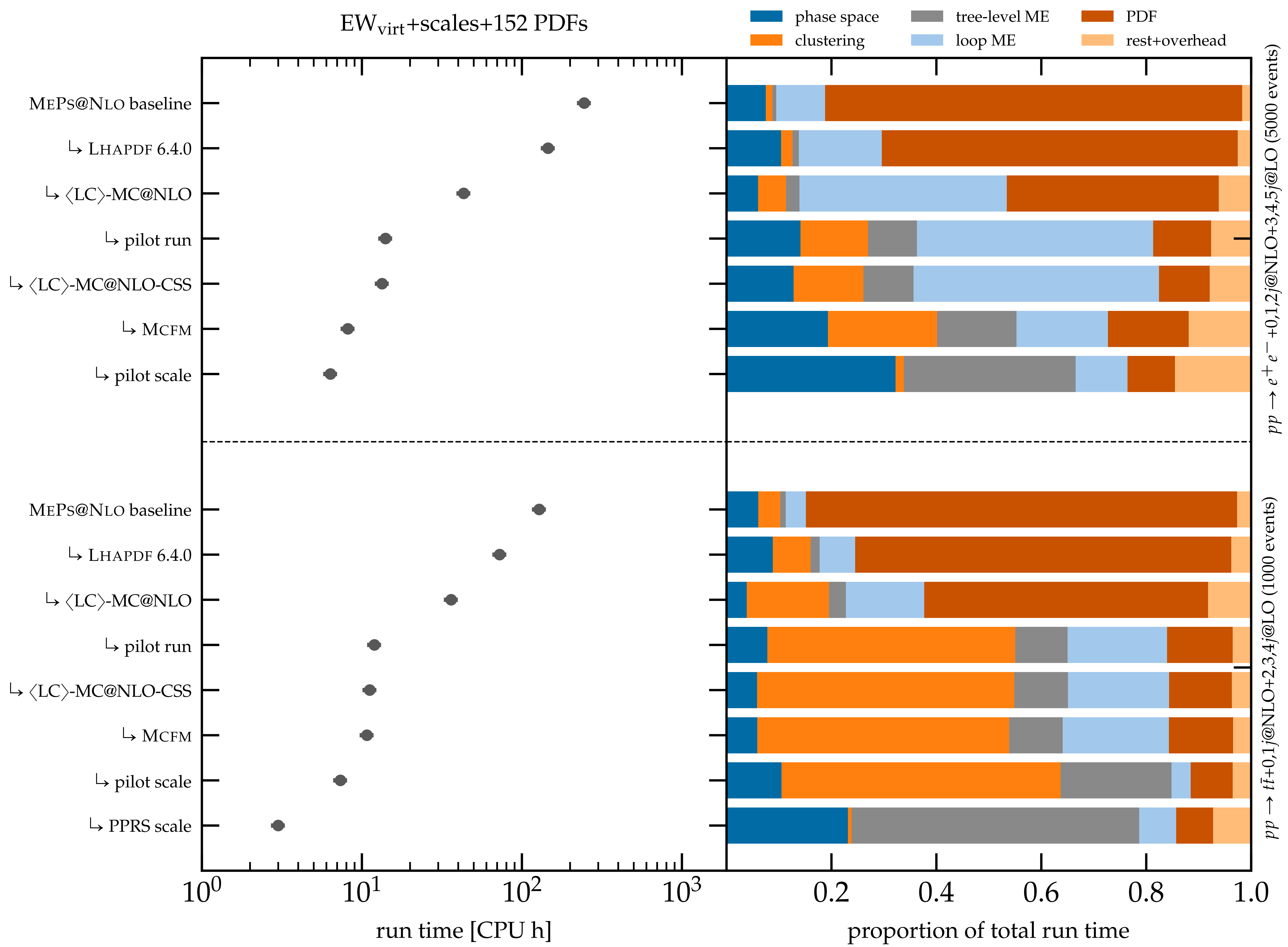
  1. LHAPDF improvement
  2. ⟨LC⟩-MC@NLO: reduce matching accuracy to leading colour, neglect spin correlations, i.e. S-MC@NLO → MC@NLO
     also useful to reduce negative event fractions [Danziger, Höche, Siegert 2110.15211]
  3. pilot run: minimal setup until PS point accepted, then rerun full setup
  4. ⟨LC⟩-MC@NLO-CSS: defer MC@NLO emission until after unweighting
  5. use analytical loop library where available
     here: OPENLOOPS → MCFM via interface [Campbell, Höche, Preuss 2107.04472]
  6. pilot scale definition in pilot run that requires no clustering
     small weight spread by correction to correct scale

- all new developments part of Sherpa 2.2.13 or later





SHERPA+LHAPDF Performance for (HL-)LHC [EB et al. 2209.00843] – Results



→ 39× speed-up for ATLAS $e^+e^- + \text{jets}$ setup

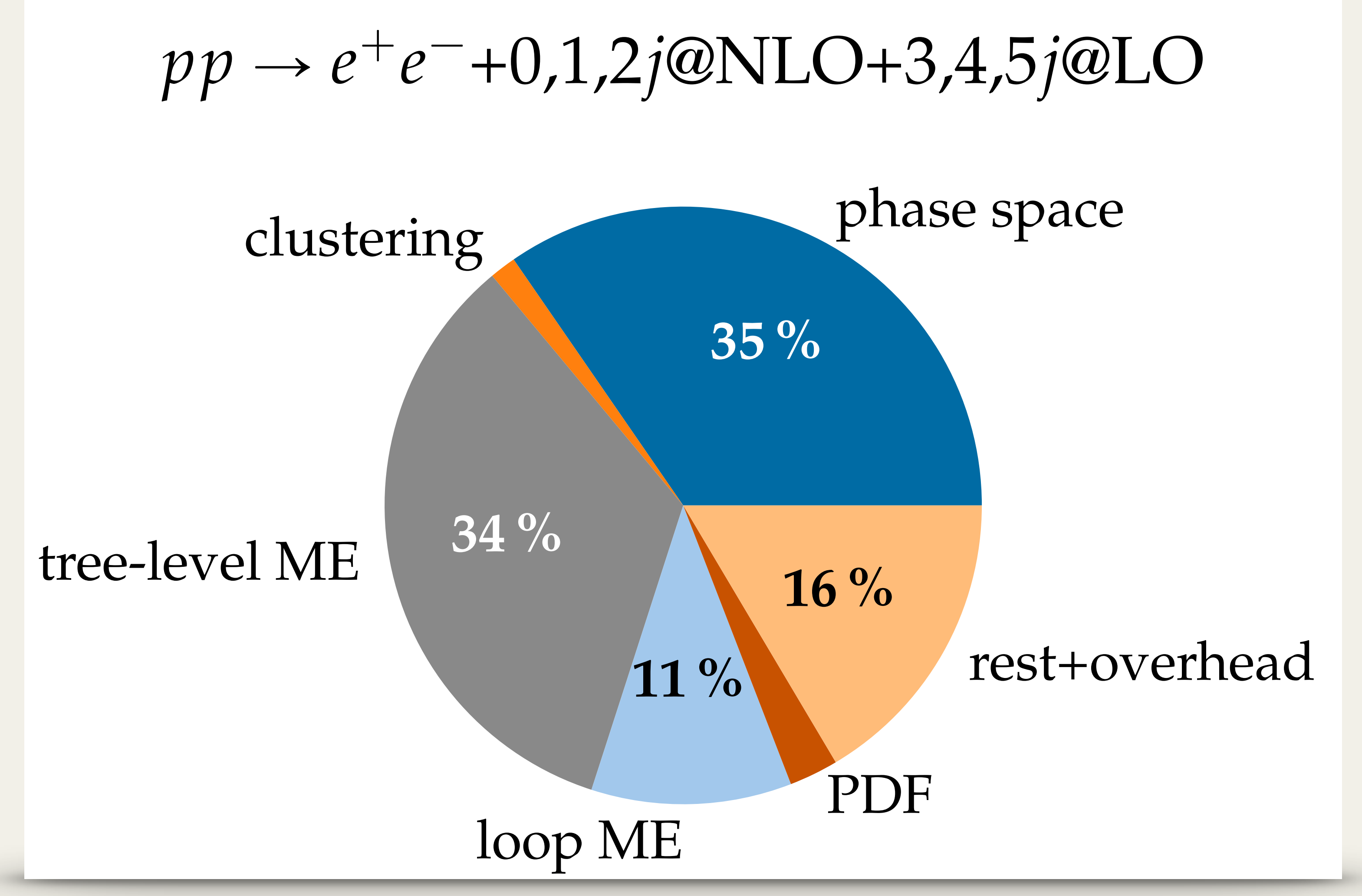
→ 43× speed-up for ATLAS $t\bar{t} + \text{jets}$ setup

Why stop here?

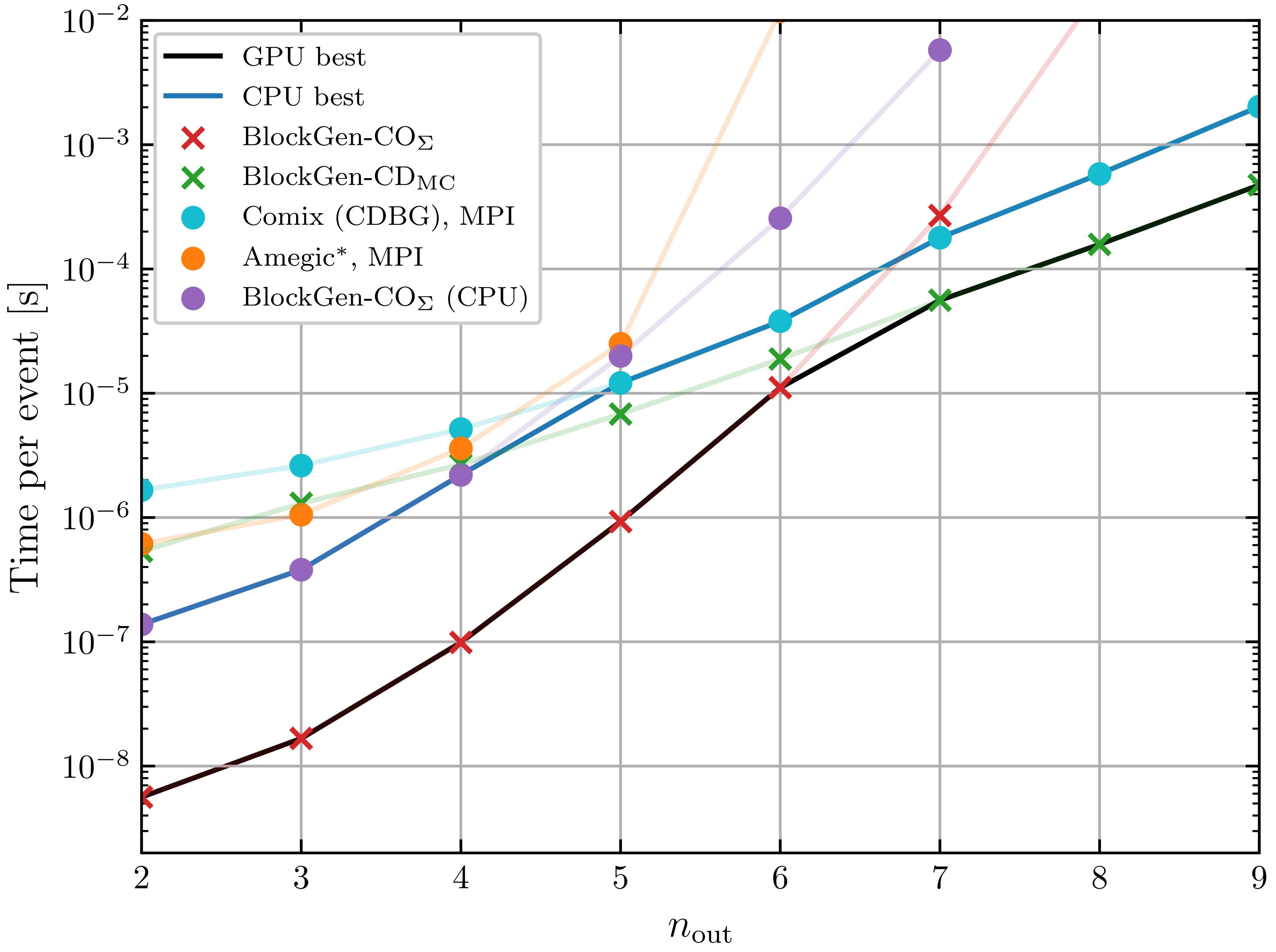
Port bottlenecks to GPU to increase physics range further

- HPC hardware increasingly heterogeneous
- other ongoing MC@GPU efforts: MadGraph5_aMC@NLO, pySecDec, MadFlow [Valassi et al. 2303.18244], [Heinrich et al. 2305.19768], [Carrazza et al. 2106.10279]; also see Zenny's talk (Wed)
- After performance improvements:
  tree-level ME and phase-space nearly 70 % of CPU usage
- Ongoing development from scratch of both components on CPU & GPU
  - concentrate on heavy hitters (V+jets, tt+jets, pure jets)
  - pick & adapt algorithms for GPU architecture
  - new ME generator Pepper (previously BlockGen)
    [EB, Giele, Höche, Isaacson, Knobbe 2106.06507]
  - new phase-space generator Chili
    [EB et al. 2302.10449]
  - use new HDF5 read-in of Sherpa 2.3.0 to integrate into existing pipeline for particle-level production for free!
  - bonus: very useful for parton-level Machine Learning studies, since training can happen exclusively on GPU

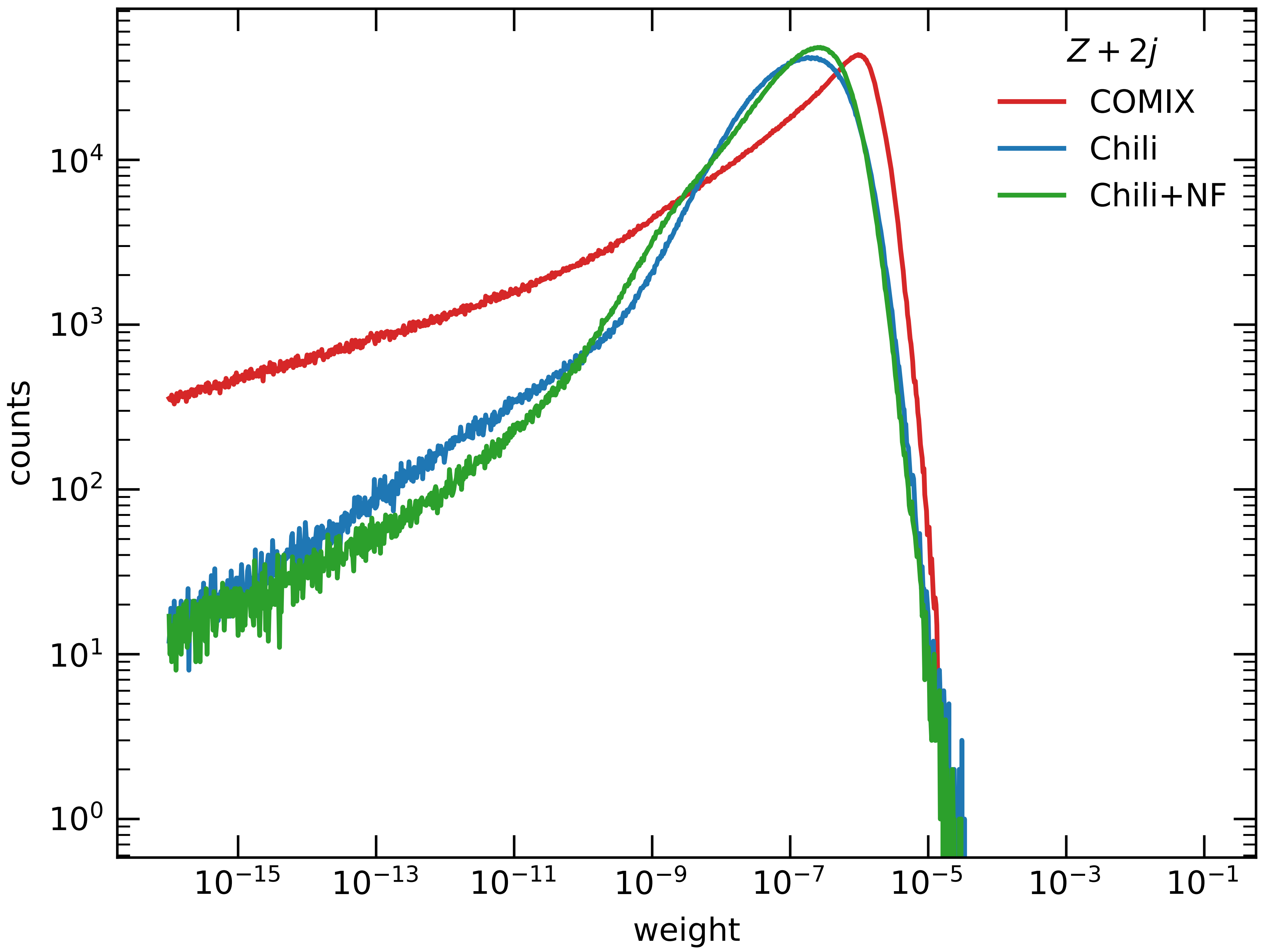
Relative CPU timing after tuning [EB et al. 2209.00843]:



Port bottlenecks to GPU to increase physics range further



→ Color-summed Berends-Giele recursion on GPU gives best performance in relevant multiplicity range, up to 150× speed-up



→ traditional phase-space parametrisation contains many channels that are not relevant for standard LHC event samples; CHILI uses much simpler (MCFM inspired) structure while achieving comparable sampling efficiency

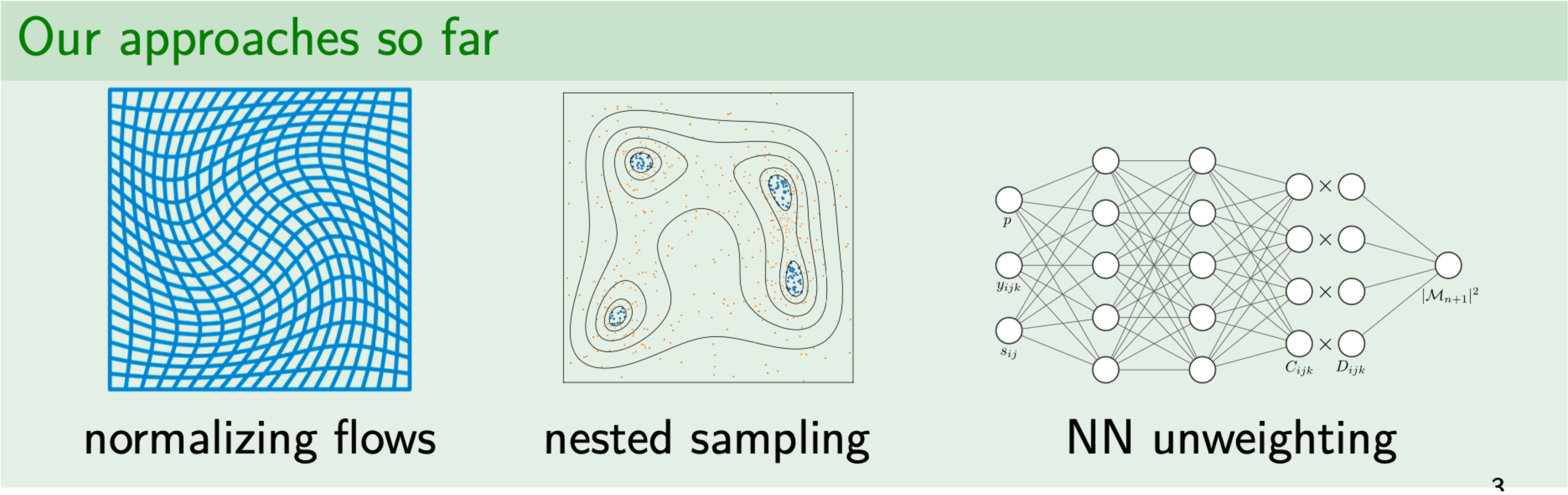
Upcoming publication of fully GPU-accelerated and HPC-ready partonic event generator

Machine Learning assisted event generation

paradigm: improve efficiency, but don't compromise on accuracy

Focus on same bottlenecks: partonic matrix elements, phase-space sampling.

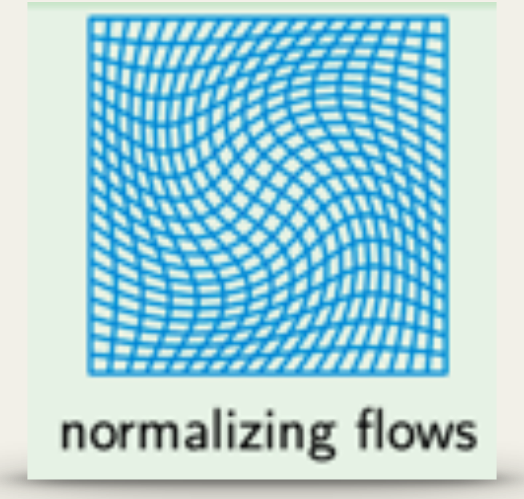
Make use of our Physics understanding.

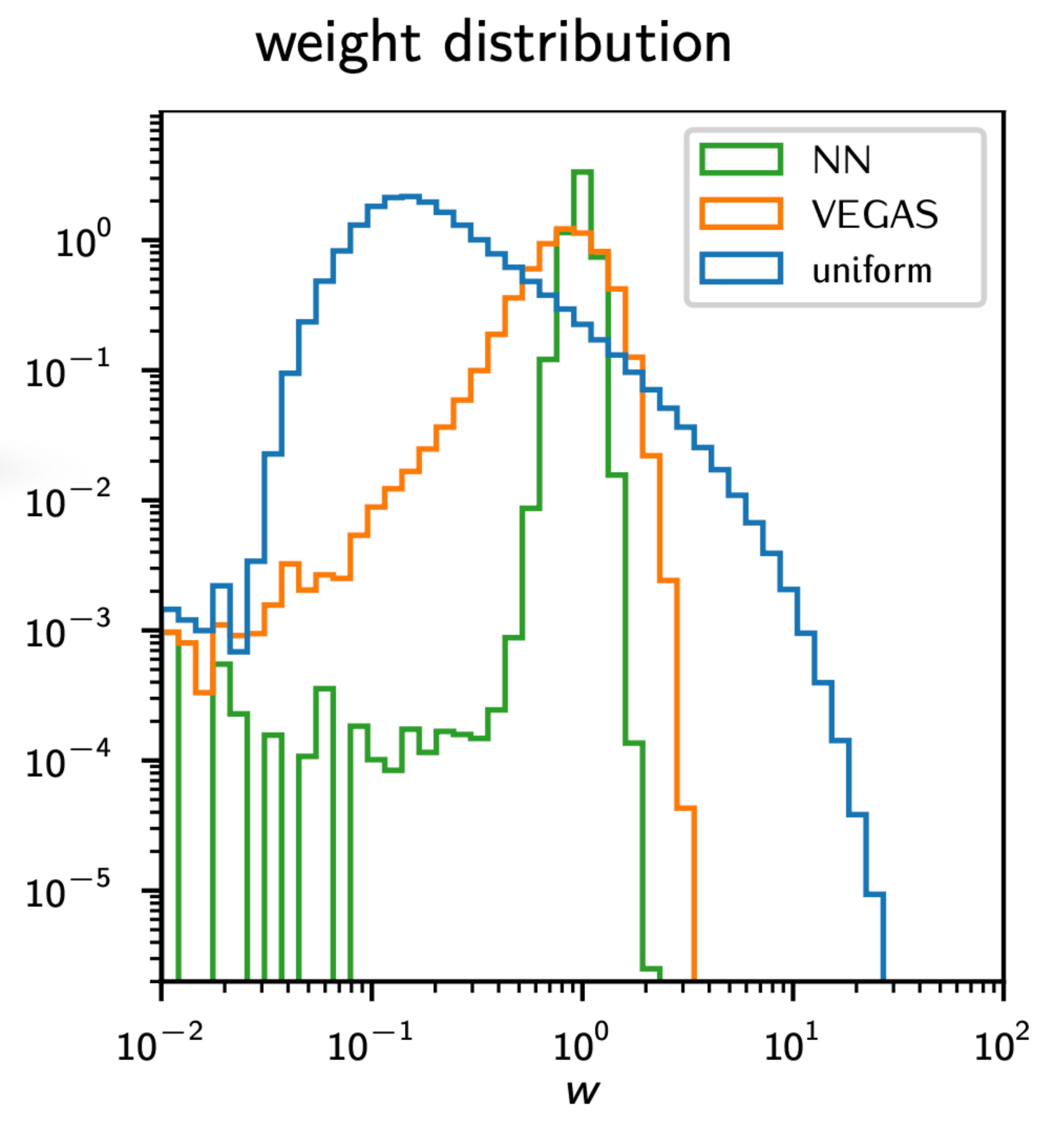


- Diffeomorphism parametrised by NN
- Drop-in replacement for VEGAS to optimise sampling

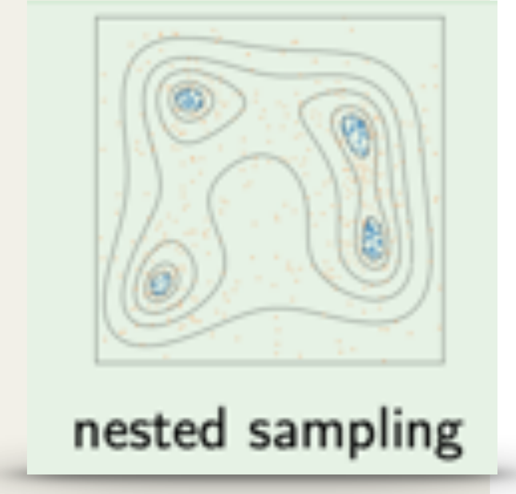
- Bayesian inference to optimise sampling
- Rich tooling available
- Short Markov chains: non-zero but low auto-correlation

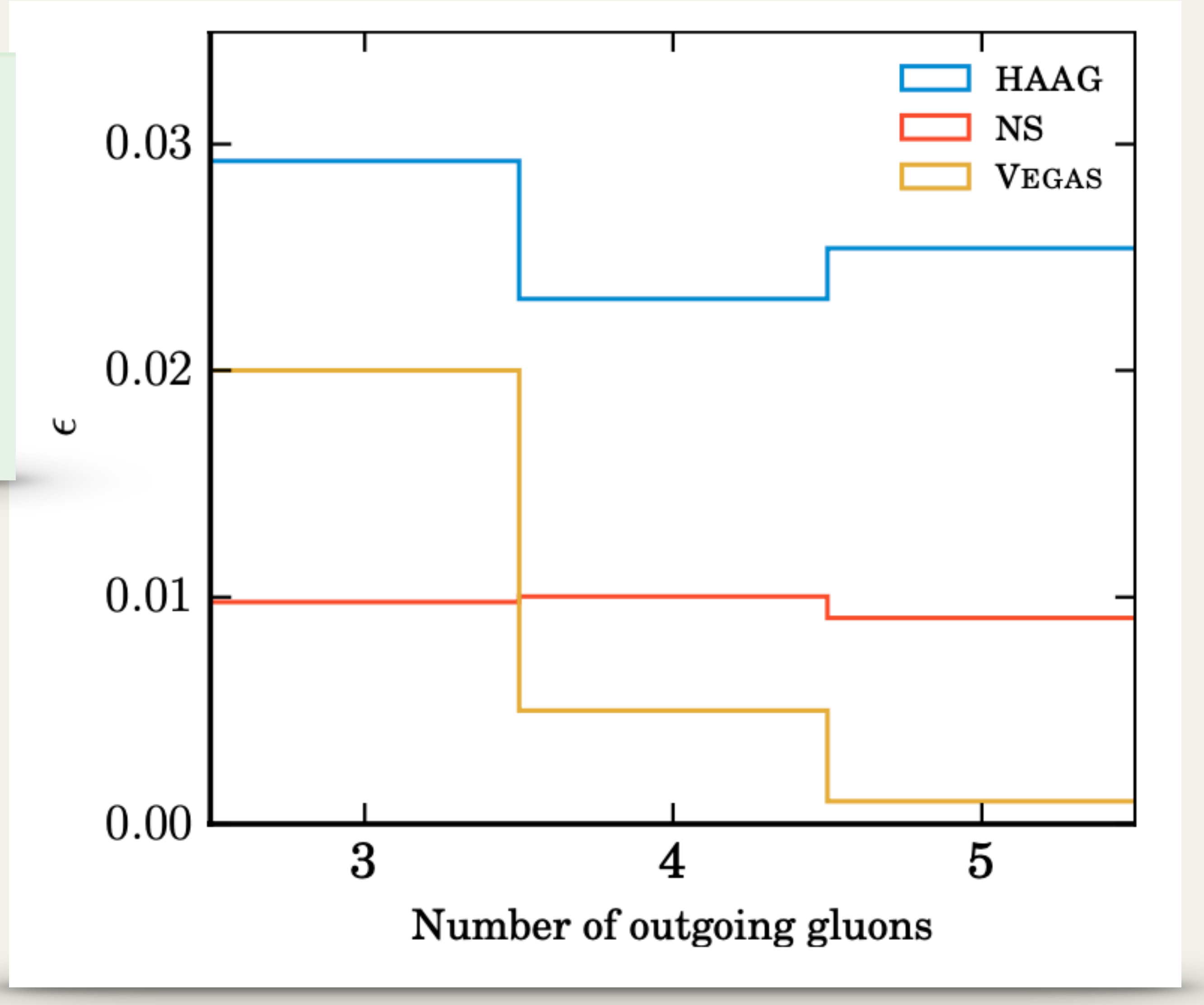
- use fast NN-based surrogate to reduce expensive ME evaluations
- recover true distribution by second unweighting step with exact ME



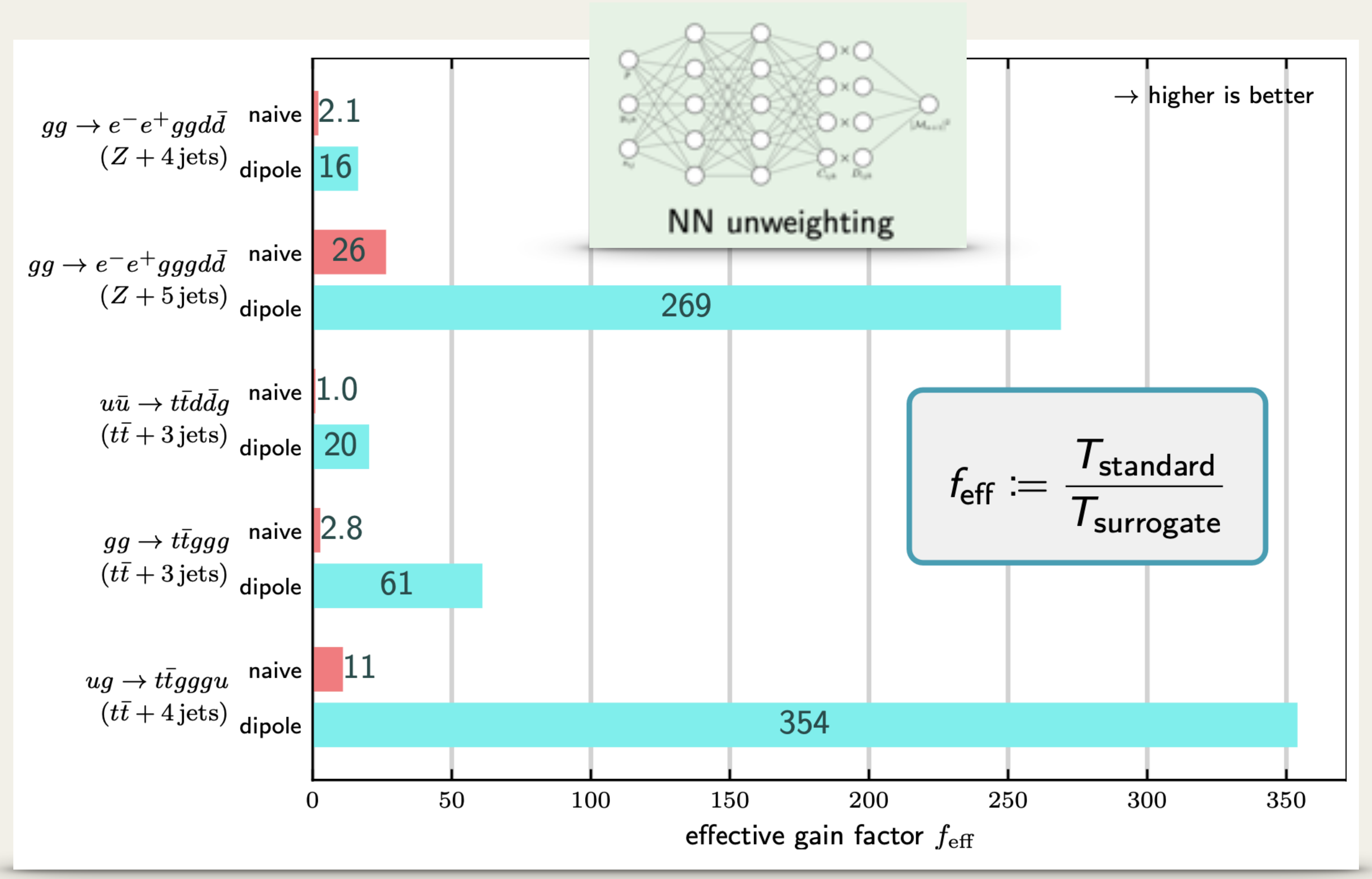


[EB, Janßen, Knobbe, Schmale, Schumann 2001.05478],
[Gao, Höche, Isaacson, Krause, Schulz 2001.10028]





[Handley, Janßen, Schumann, Yallup 2205.02030]



[Danziger, Janßen, Schumann, Siegert 2109.11964]
[Janßen, Maître, Schumann, Siegert, Truong 2301.13562]

- **Significant improvements for simple cases**
- **good scaling for sampling not yet solved/proven …**
- **… but NN unweighting offers working solution**

Status and prospects of SHERPA

Efficiency

Precision

New NLL-accurate shower algorithm

ALARIC → see D. Reichelt's talk (Mon) for details

- Framework to quantify log accuracy of parton showers established in
  [Dasgupta et al. 1805.09327, 2002.11114], also see Gavin's talk (Mon) & more refs. therein

- NLL accuracy requires that kinematics mapping of $n \to n+1$ phase space should not distort effects of pre-existing emissions on observables

  - extract NLL relevant effects by taking limit $\alpha_s \to 0$ at fixed $\lambda = \alpha_s \log v$, where v = resummed observable

  - PANSCALES developed & proven to fulfill requirements
    See Gavin's talks (Mon) & refs. therein

  - pre-existing showers in SHERPA do not meet this requirement

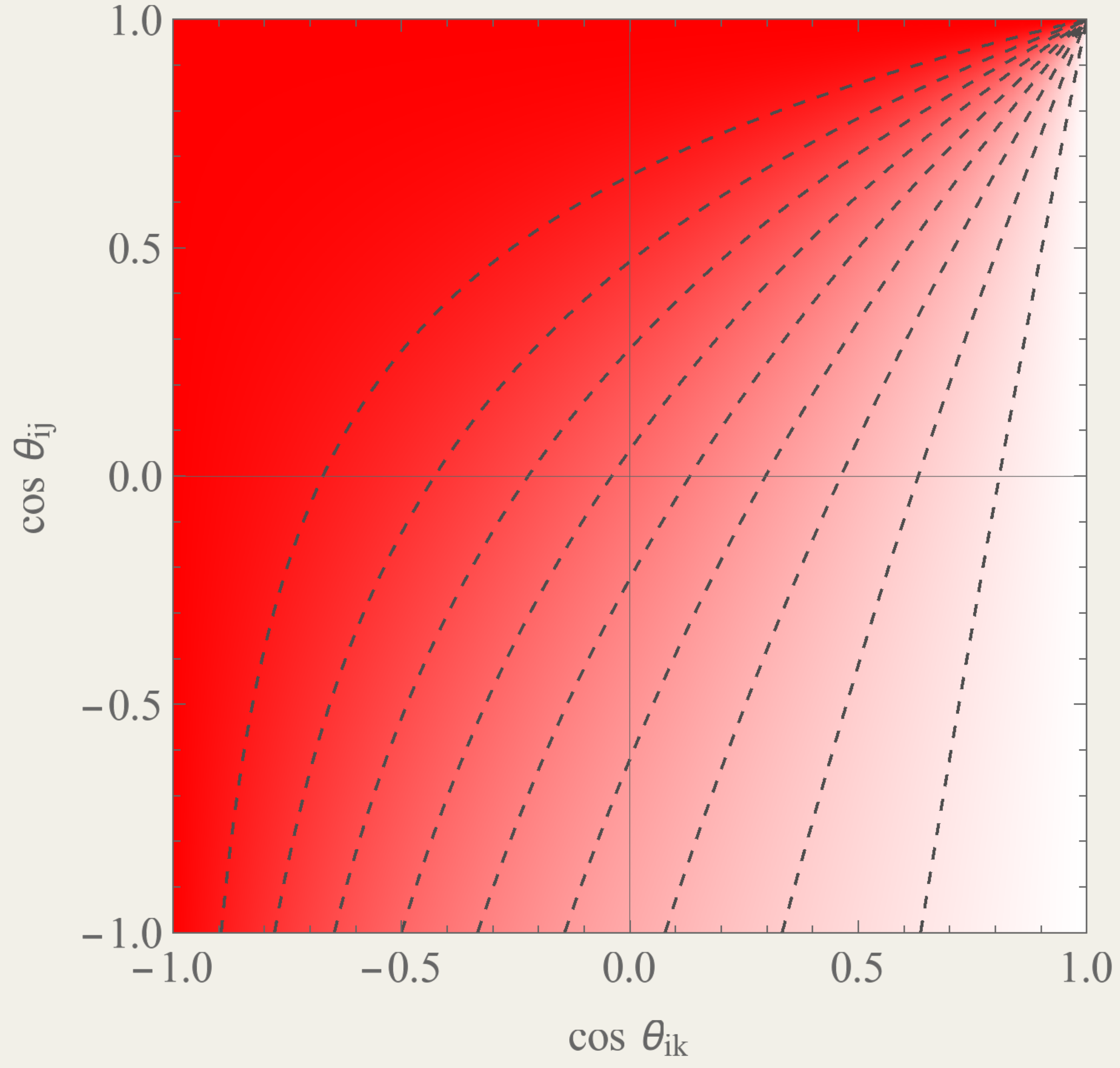
- new shower ALARIC
  [Herren Höche Krauss Reichelt Schönherr 2208.06057]

  - partial fractioning of eikonal → positive definite splitting function with full phase space coverage
    inspired by Catani & Seymour's treatment of identified hadrons

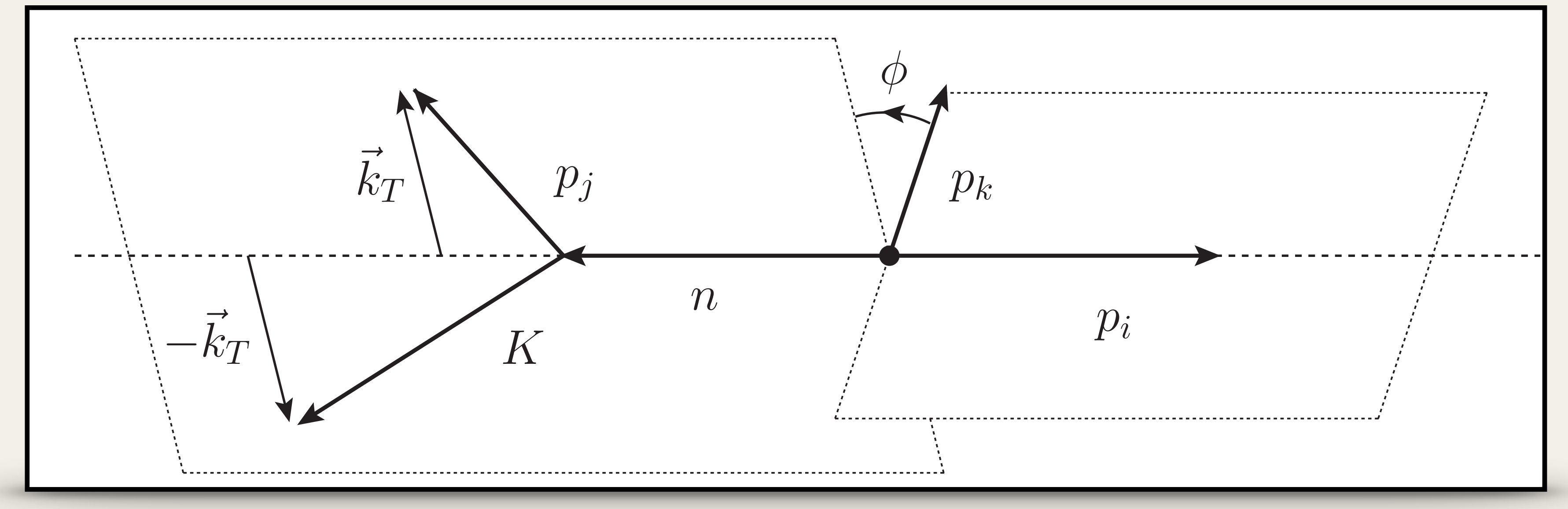
    - price: dependence of splitting functions on azimuthal angle

  - global kinematics scheme enables analytic proof of NLL accuracy + numerical validation

ALARIC's positive-definite radiator function

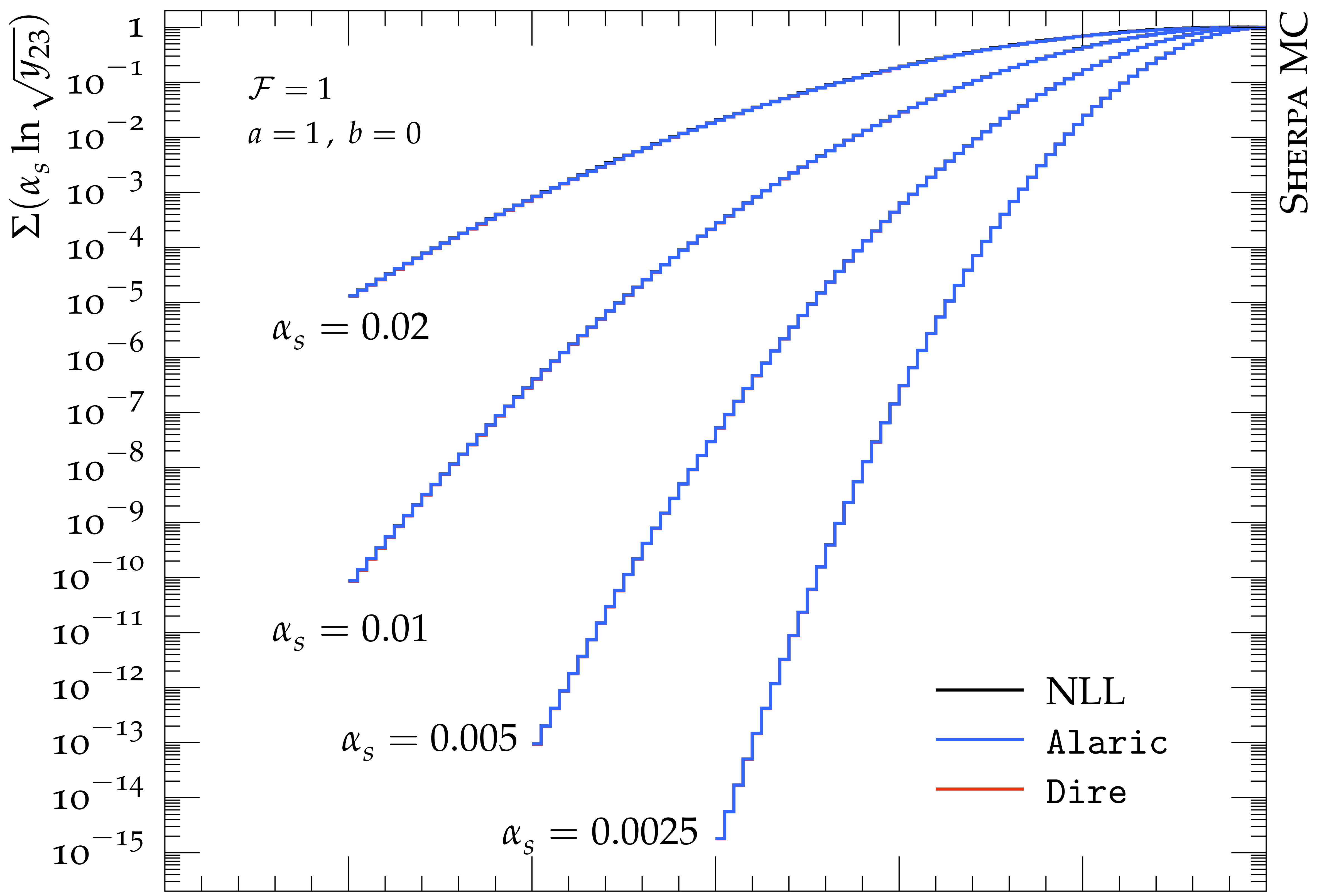


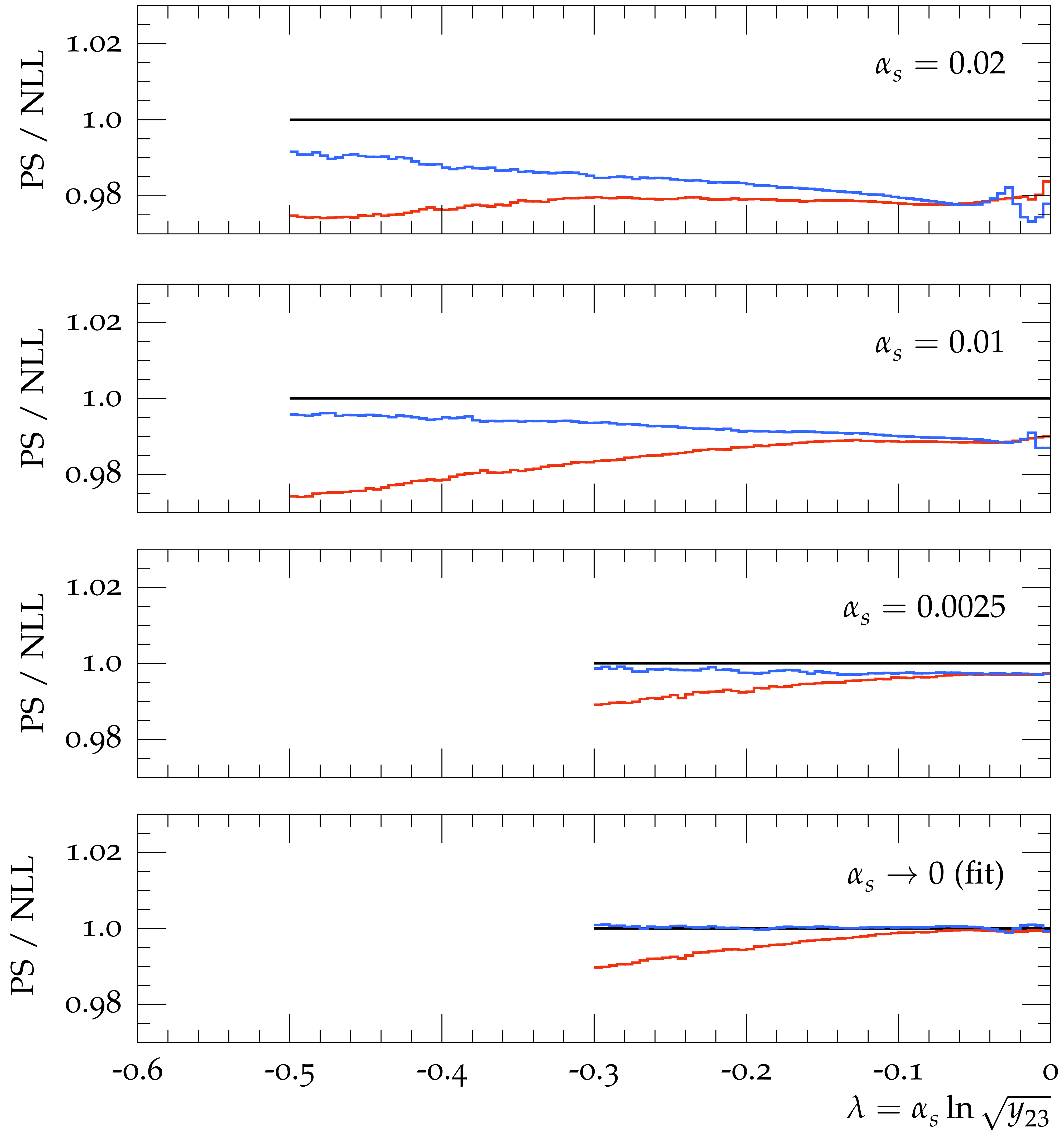
ALARIC's emission kinematics with ϕ dependence



New NLL-accurate shower algorithm

ALARIC → see D. Reichelt's talk (Mon) for details

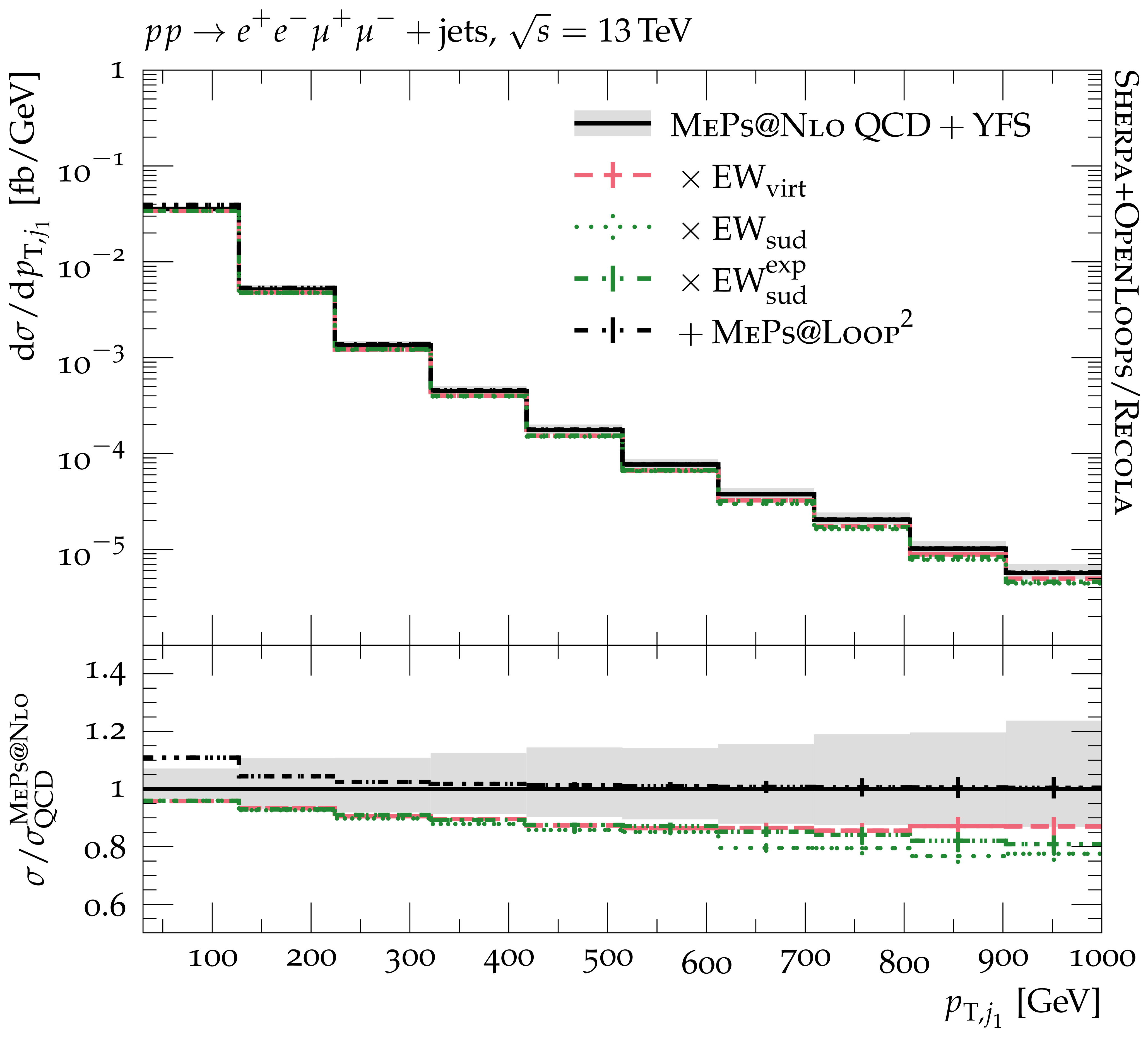




- Numerical Study taking limit $\alpha_s \to 0$ at fixed $\lambda = \alpha_s \log v$
- v = $2 \to 3$ clustering scale (Cambridge algorithm)
- DIRE has large deviations, but ALARIC flat wrt. NLL ✓

EW Sudakov logarithms

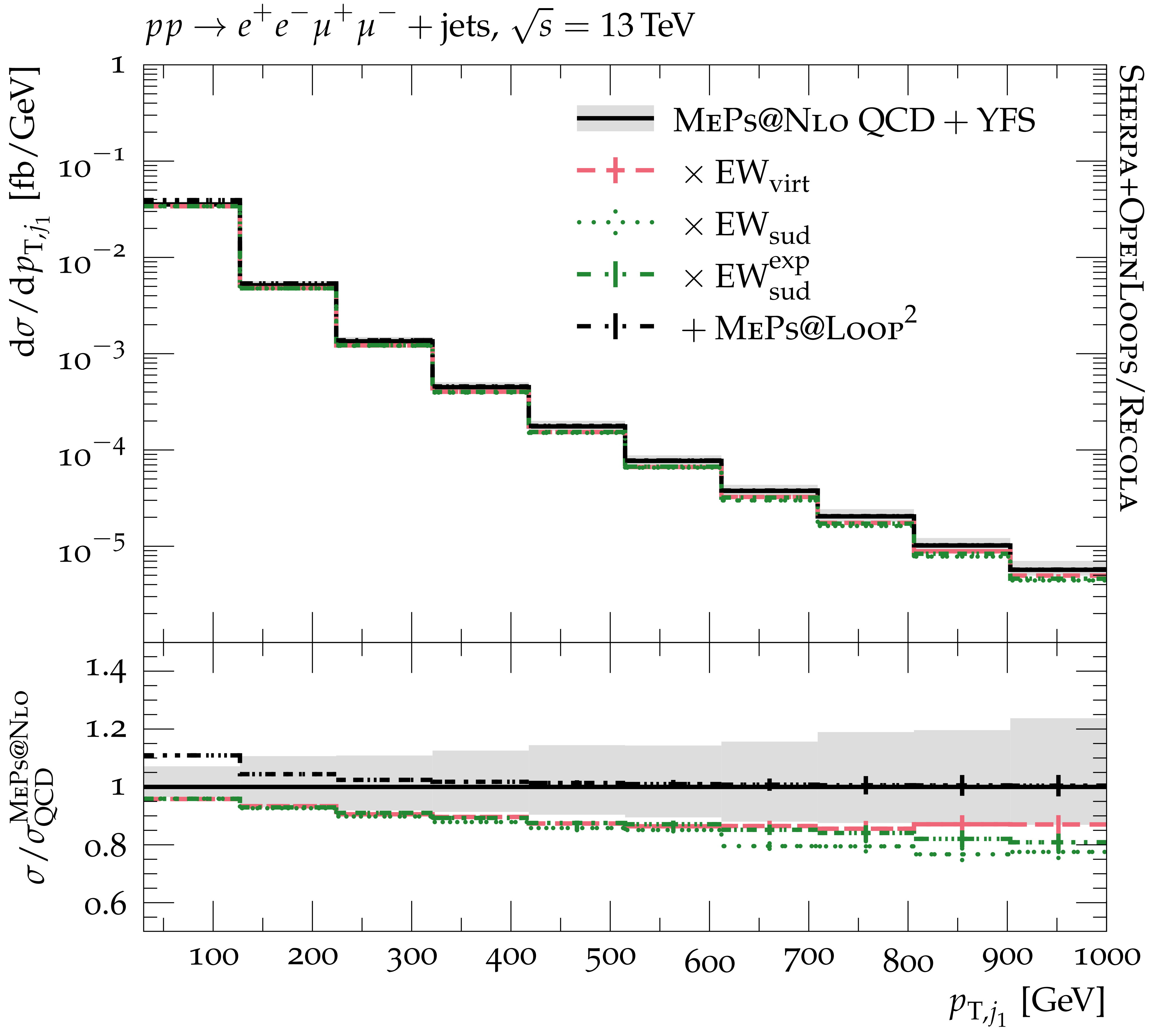
Automated implementation for all processes

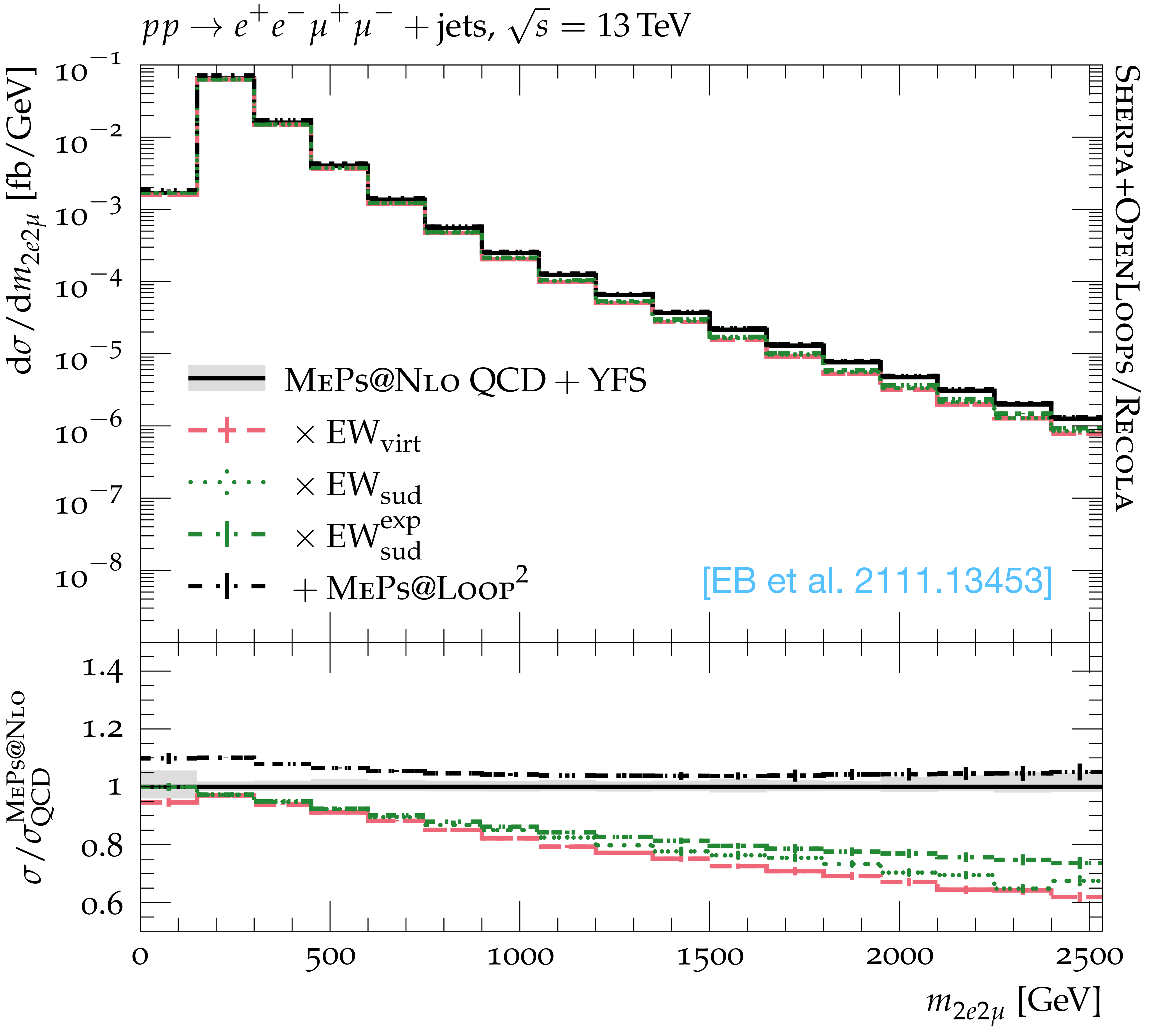


- Corrections due to soft/coll. EW gauge bosons coupled to external legs in high-energy limit
  (e.g. $p_T \gtrsim 1\,\text{TeV} \rightarrow \mathcal{O}(10\,\%)$ corrections)
- Corrections worked out in full generality
  [Denner, Pozzorini (2001) hep-ph/0010201]
- partial implementation in ALPGEN
  [Chiesa et al 1305.6837]
- In SHERPA fully automated as universal ME-level corrections applicable in all setups for any process, including MEPS@NLO predictions
  [EB, Napoletano 2006.14635], [EB et al. 2111.13453]
  - EW$_{\text{virt}}$ for $\mathcal{S}$ events, EW$_{\text{sud}}$ for $\mathcal{H}$ and LO events
  - YFS resummation for QED FSR
- Example: application to MEPS@NLO diboson production
  $pp \rightarrow 0{,}1j\,@\,\text{NLO} + 2{,}3j\,@\,\text{LO}$ [EB et al. 2111.13453]
- similar implementations in development for MadGraph5_aMC@NLO and OpenLoops [Pagani, Vitos, Zaro 2309.00452], [Recent talks by OpenLoops]

EW Sudakov logarithms

Automated implementation for all processes



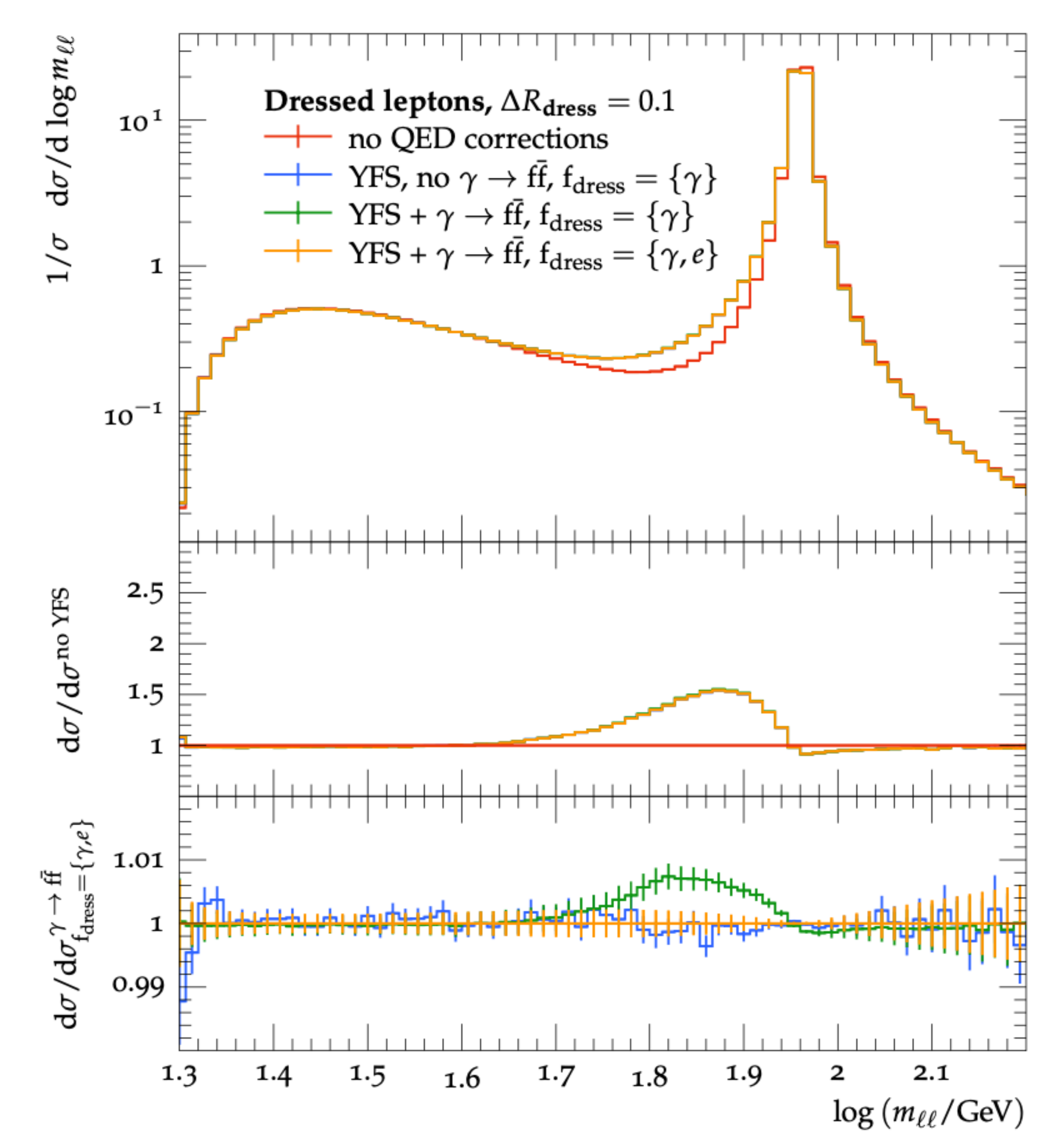


(Approximate) EW corrections outside of MEPS@NLO QCD uncertainty band

Resumming soft photons with YFS

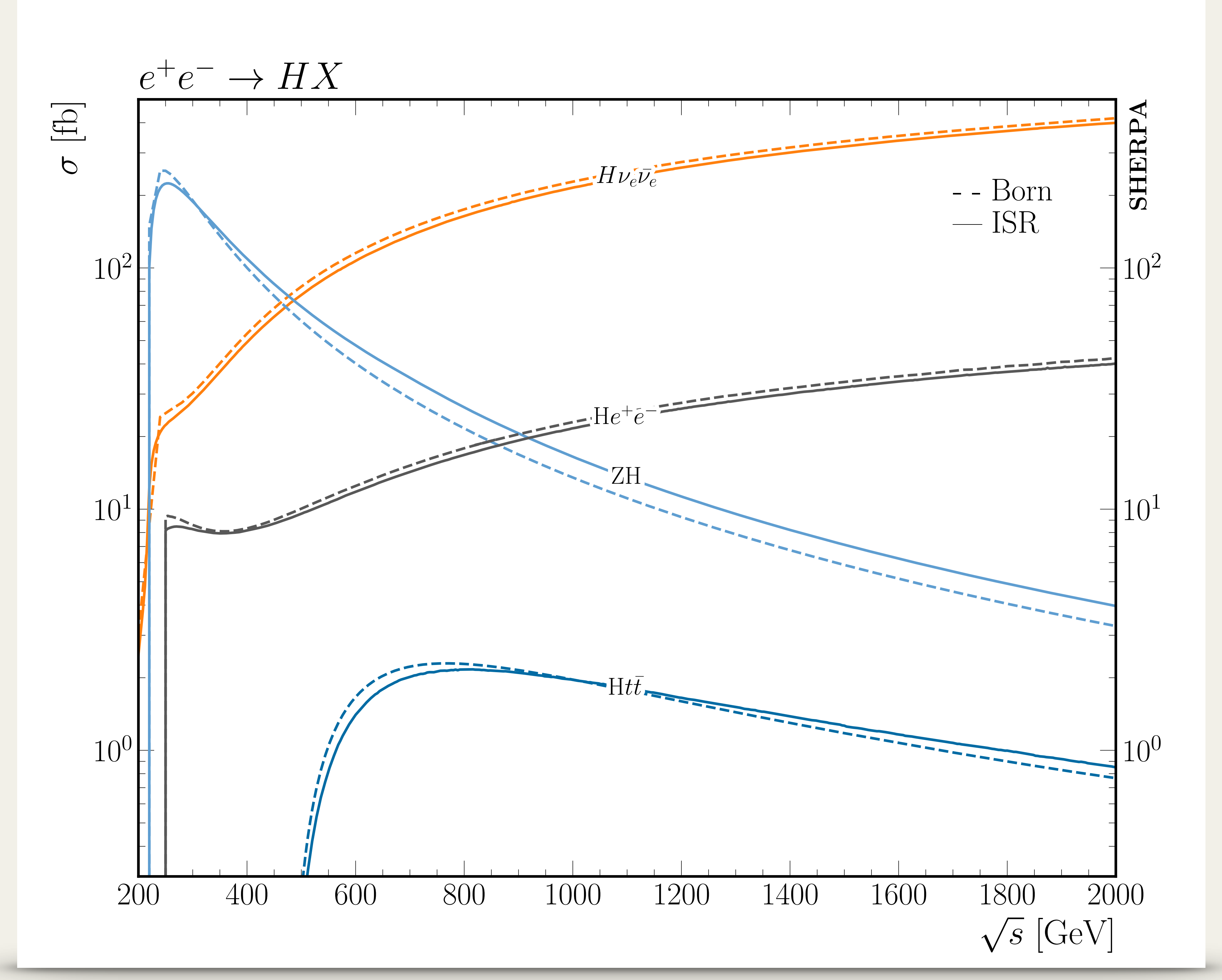
Recent developments in SHERPA

- photon splitting $\gamma \to e^+e^-$ [Flower, Schönherr 2210.07007]
  → Lois Flower's talk (Wed)
- Example: Dilepton invariant mass for $pp \to e^+e^-$:



Corrections up to 1 %, can be reigned in by refined dressing algorithm

- YFS in ISR for future lepton colliders
  [Krauss, Price, Schönherr 2203.10948]
- Application to Higgsstrahlung processes at lepton collider:



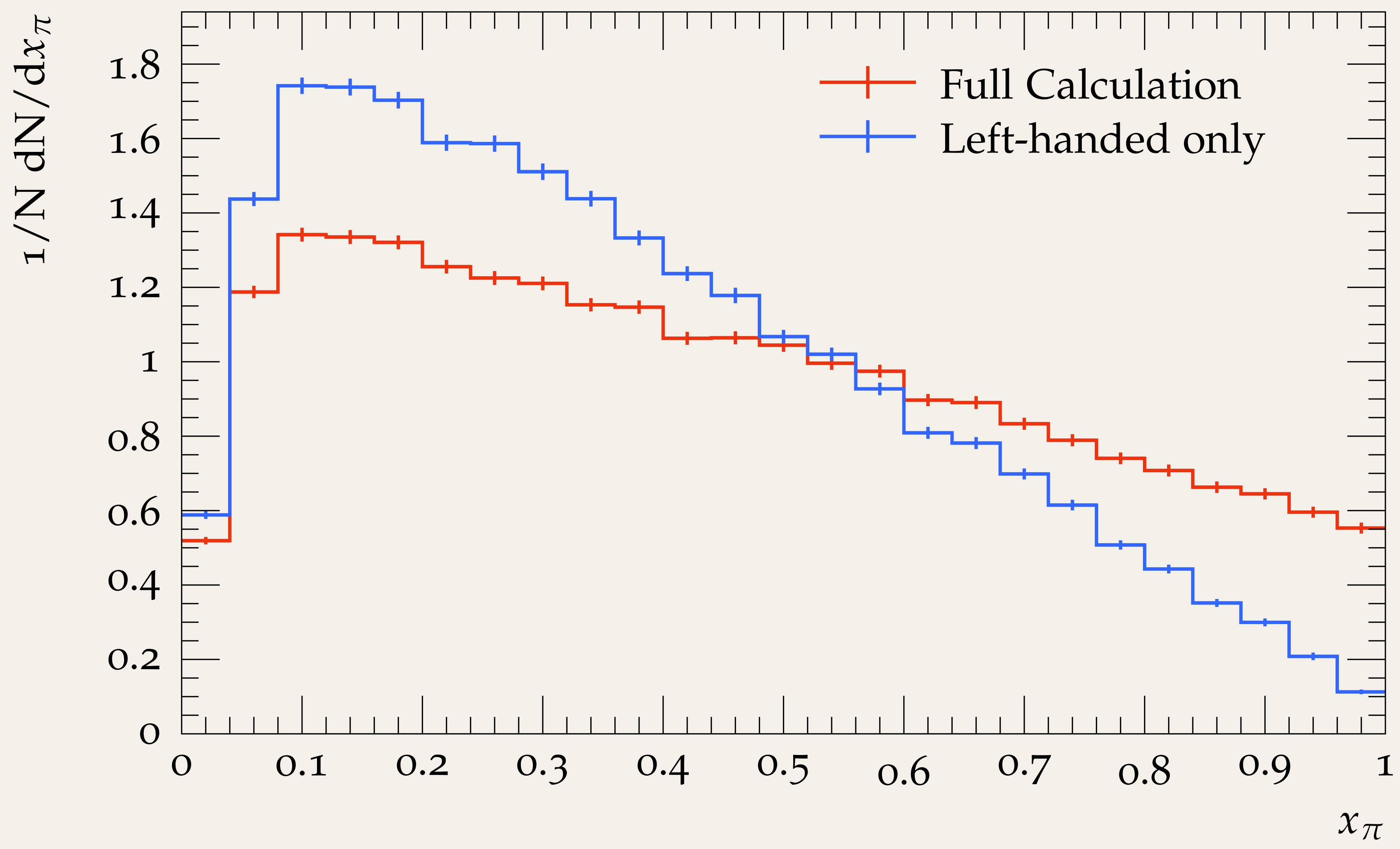
Process-independent implementation of YSF for ISR

Neutrino physics

Achilles + Sherpa

- Achilles is a newly developed neutrino event generator [Höche, Isaacson, Lopez Gutierrez, Rocco 2110.15319]
- paradigm: transfer LHC expertise+tooling in neutrino physics, developed in close collaboration with Sherpa
  - Achilles for nuclear physics effects
  - Sherpa's Comix for calculating leptonic currents, incl. BSM effects via Comix' UFO interface
  - study of ν_τ needs control over angular distribution of τ-lepton decay products
    - use interface to Sherpa for decays, incl. spin correlations across production and decays, QED showers [Isaacson, Höche, Siegert, Wang 2303.08104]

Momentum fraction for tau decay to single pion folded over DUNE's far-detector flux



→ improved description over conventional fully left-handed assumption of other neutrino generators

BSM physics

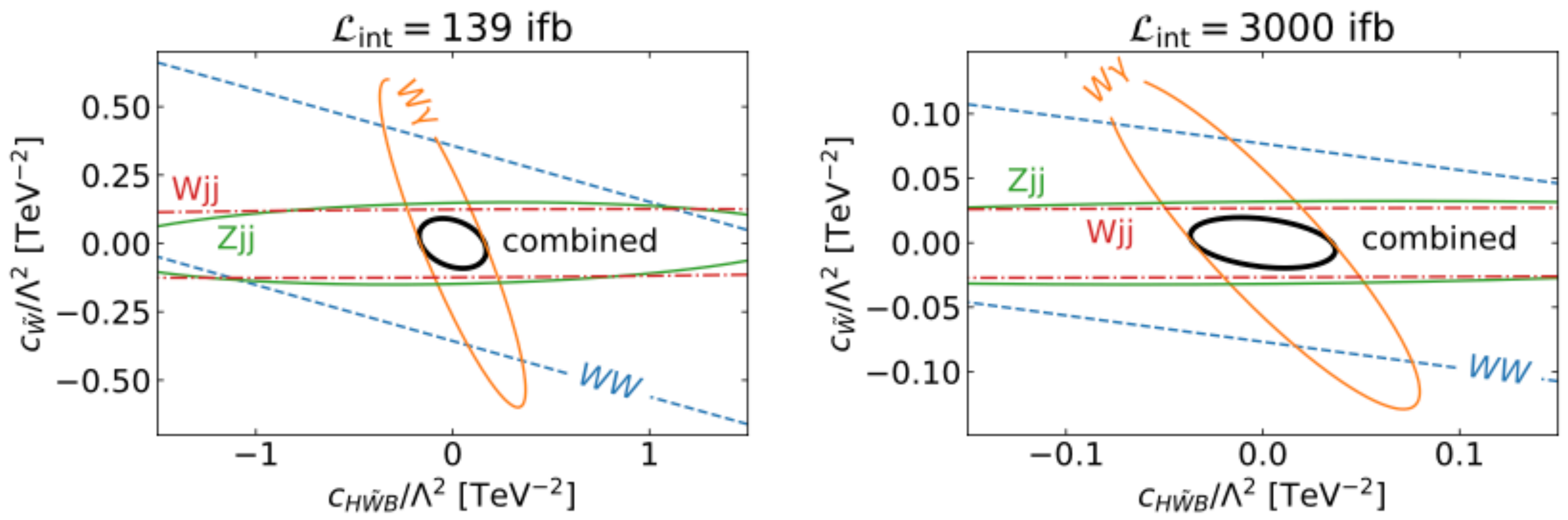
via UFO interface [Höche, Kuttimalai, Schumann, Siegert 1412.6478]

- full support for UFO model [Degrande et al. CPC183(2012)1201]
- UFO2 ongoing [Darmé et al. 2304.09883]
- Lorentz and colour structures built fully automatically
- automatic inclusion in hard decay module
  - identification of all $1 \to 2$ and $1 \to 3$ decay channels and calculation of LO widths
  - can select individual channels
  - spin correlations using spin density matrices [Richardson JHEP11(2201)029, Knowles CPC58(1990)271]

BSM physics

Calculating AGC limits using Sherpa+UFO [Biekötter, Gregg, Krauss, Schönherr 2102.01115]

- LO multi-leg with SMEFT model defined via UFO
- use public ATLAS and CMS SM measurements to constrain SMEFT parameters



Status and prospects of SHERPA

Conclusions

Conclusions

- **Efficiency improvements** (= increase physics range)
  - tuning exercise → factor-40 speed-up for heavy hitter ATLAS setups
  - porting bottlenecks to GPU → PEPPER+CHILI, integrated with Sherpa v2.3 via HDF5 event files
  - ML assisted event generation → NF, Nested Sampling, NN unweighting
- **Precision physics**
  - new NLL-accurate shower ALARIC
  - Fully automated EW_{sud} logarithms: application to MEPS@NLO ZZ production, but fully general
  - YFS developments: $\gamma \to e^+e^-$ splittings and ISR
  - Neutrino physics via ACHILLES+SHERPA
  - BSM physics via UFO

Backup

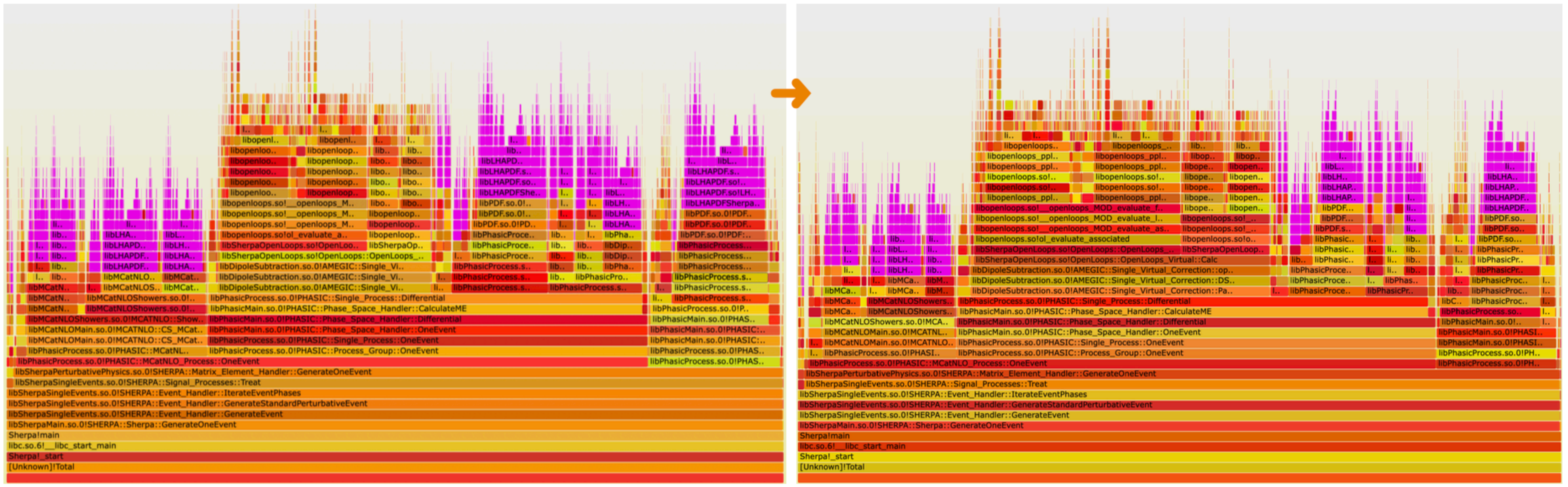
SHERPA+LHADPF Performance for (HL-)LHC

Overall profiling and tuning [EB et al. 2209.00843]

LHAPDF 6.2.3 → 6.4.0

- PDF grid caching for given (x, Q^2) point
  - repeated calls for different flavours / replicas benefit
  - caller side might need to reorder calls to benefit
- Use same interpolation grid structure across flavours
- Cache universal terms of polynomial interpolation
- up to 3x faster for single flavour, ~10x for all flavours

Sherpa flamegraph for ATLAS V+jets setup before/after LHAPDF 6.4.0

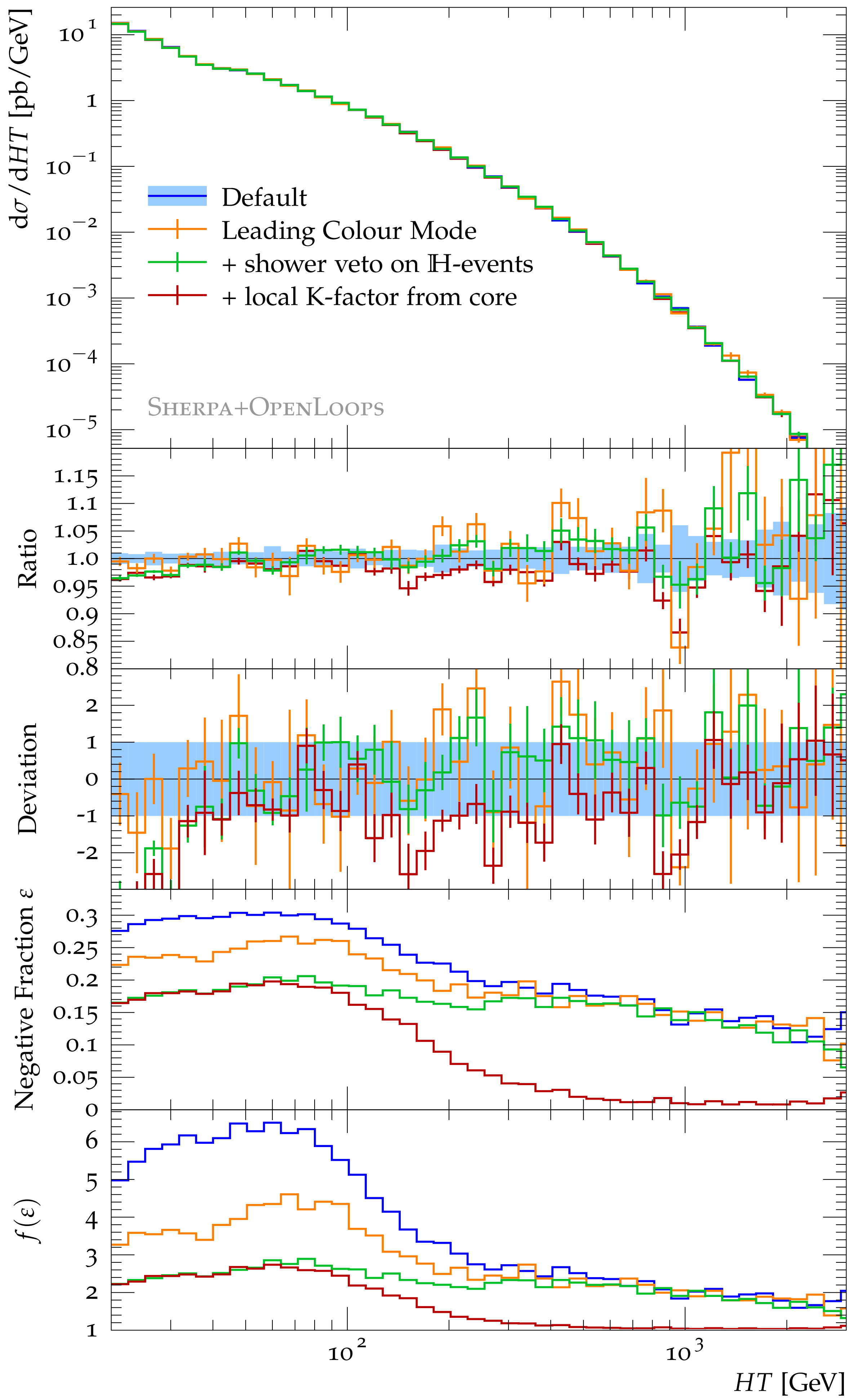


[Andy Buckley, Max Knobbe]

Negative weight fractions

Danziger, Höche, Siegert, arXiv:2110.15211, ATLAS arXiv:2112.09588

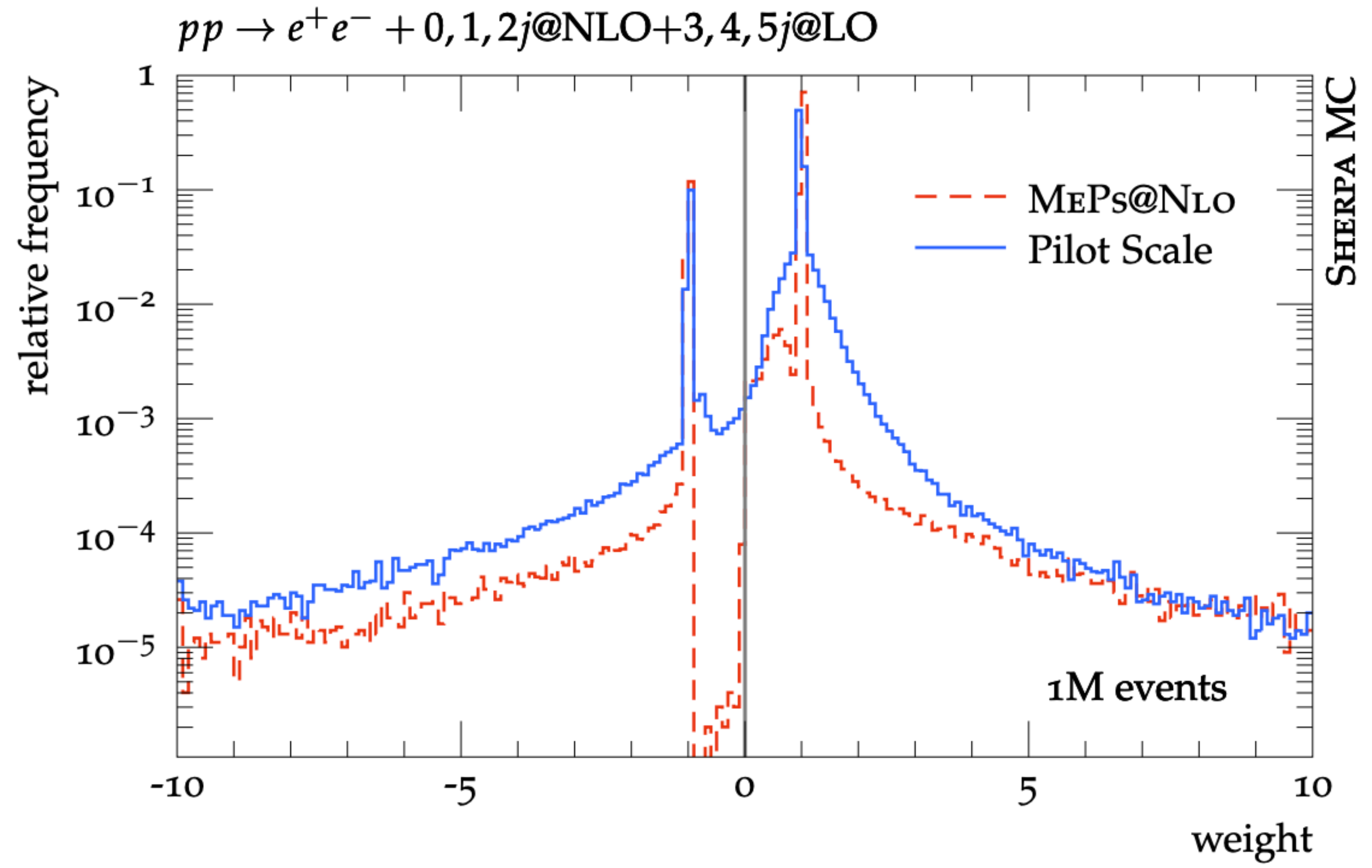
- explored three methods to improve the neg. weight fraction in SHERPA
  1) reduce matching accuracy to leading colour, neglect spin-correlations
  2) include jet veto on $\mathbb{H}$-events, as originally formulated in arXiv:2012.5030
  3) use local K-factor in NLO→LO merging from core configuration instead of highest multiplicity
- public since SHERPA-2.2.8 (Sep '19)

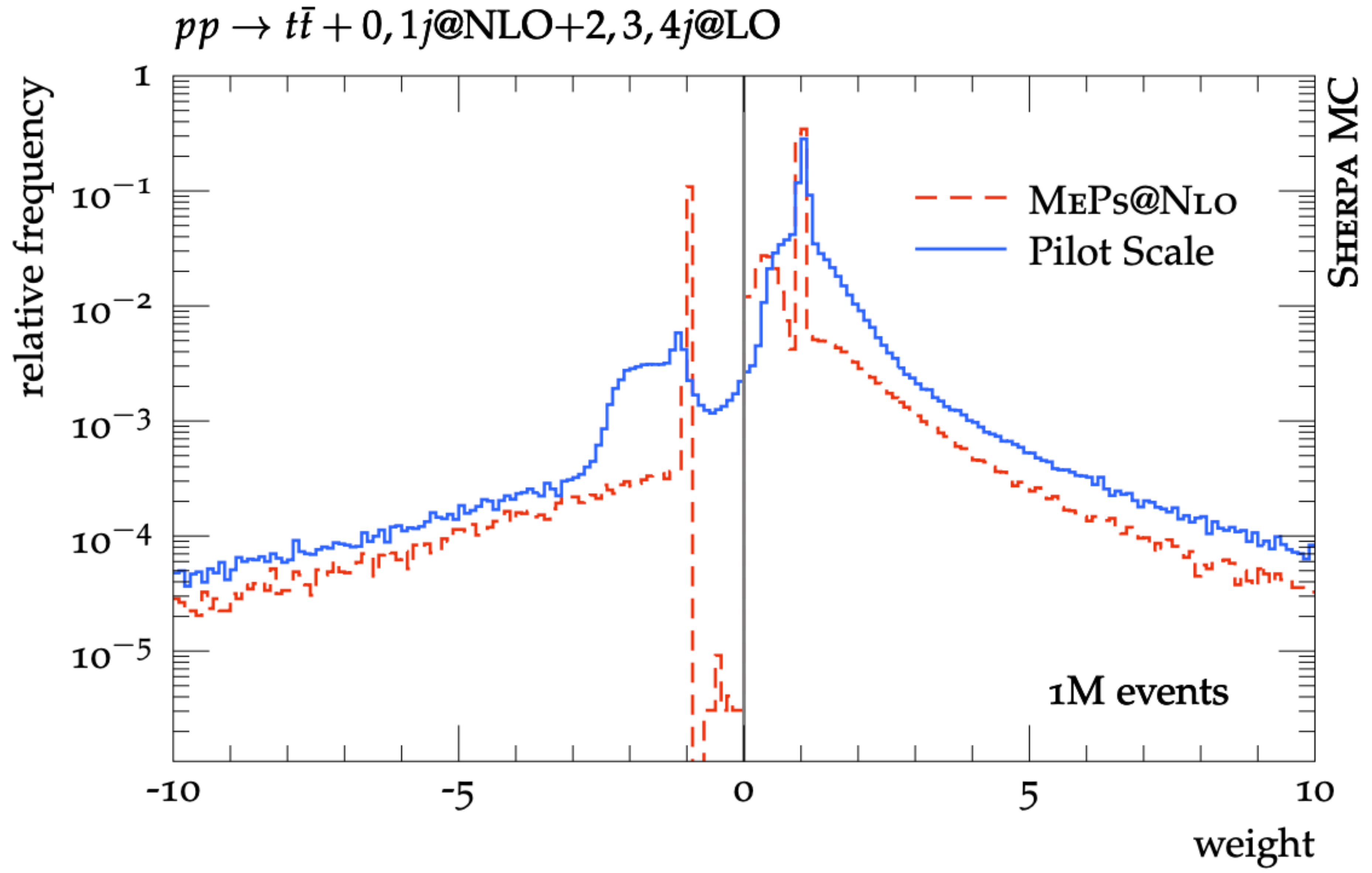


SHERPA+LHADPF Performance for (HL-)LHC

Overall profiling and tuning [EB et al. 2209.00843]

weight distribution broadening due to the use of a Pilot scale

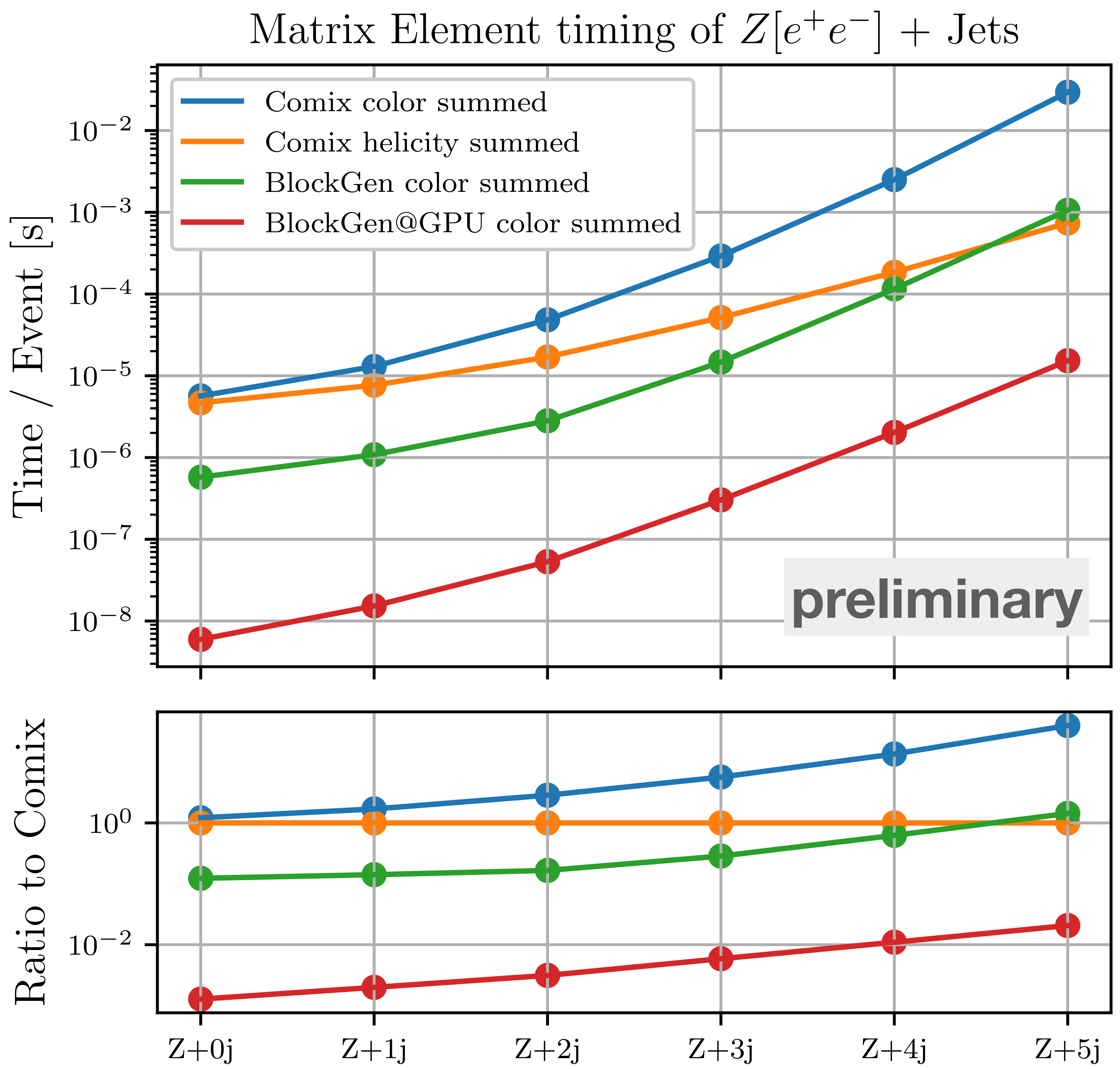




- effective reduction in efficiency from using the pilot scale typically $\lesssim 2$
- computing time reduction reduced by this, but in most cases still beneficial

Port bottlenecks to GPU

PEPPER vs. COMIX runtime per partonic event



Port bottlenecks to GPU

CHILI vs. COMIX runtime for given accuracy target

Process / MC accu	Default PS Time	Default PS # pts	New PS Time	New PS # pts
W+1j / 1‰	4m 52s	10.3M	2m 32s	3.10M
W+2j / 3‰	17m 12s	5.52M	13m 52s	2.53M
W+3j / 1%	46m 24s	7.48M	20m 16s	1.15M
H+1j / 1‰	2m 20s	1.83M	1m 36s	1.50M
H+2j / 3‰	4m 36s	2.32M	4m 4s	0.71M
H+3j / 1%	18m 12s	2.32M	12m 56s	0.63M

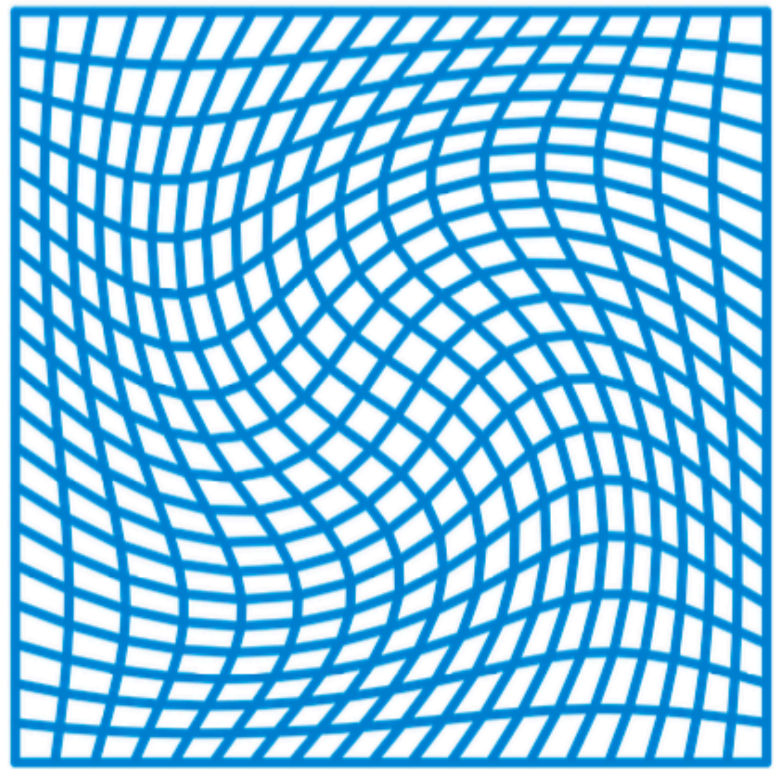
Process / MC accu	Default PS Time	Default PS # pts	New PS Time	New PS # pts
$t\bar{t}$+0j / 1‰	4m 38s	3.15M	4m 0s	3.59M
$t\bar{t}$+1j / 3‰	3m 12s	1.38M	3m 4s	1.47M
$t\bar{t}$+2j / 1%	11m 58s	1.47M	11m 20s	0.89M
2j / 1‰	12m 48s	2.98M	7m 44s	1.80M
3j / 3‰	22m 48s	6.80M	23m 12s	2.39M
4j / 1%	1h 25m	6.95M	50m 24s	0.91M

[EB et al. 2302.10449]

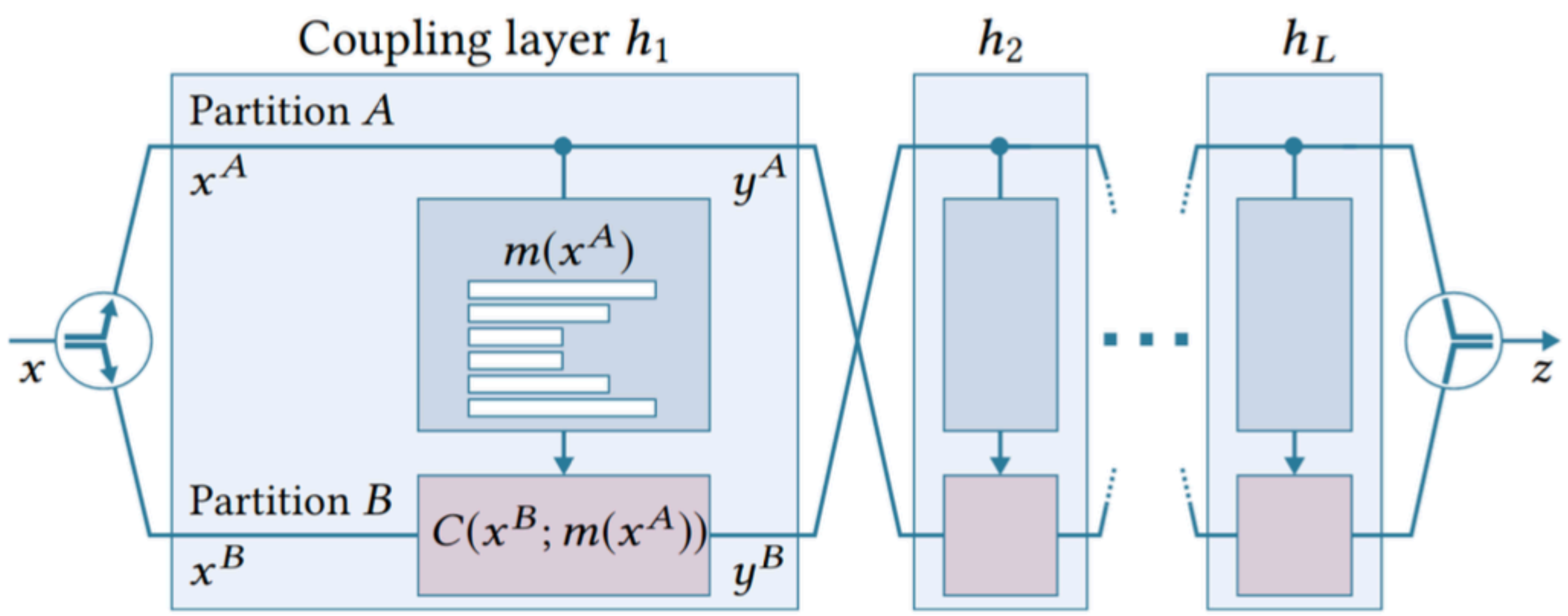
Normalizing Flows

Slide by Timo Janßen

- diffeomorphism parameterized by NNs
- layered mapping: $h = h_L \circ \cdots \circ h_2 \circ h_1$
- each coupling layer transforms part of input
- triangular Jacobian $\rightsquigarrow$ determinant costs $\mathcal{O}(d)$
- replacement for VEGAS



©wikimedia.org File:Diffeomorphism of a square.svg



Müller et al.: SIGGRAPH 2019

Normalizing Flows

Gain factors for V + n jets [Gao et al. 2001.10028]

unweighting efficiency $\langle w \rangle / w_{\max}$		LO QCD $n=0$	$n=1$	$n=2$	$n=3$	$n=4$	NLO QCD (RS) $n=0$	$n=1$
$W^+ + n$ jets	Sherpa	$2.8 \cdot 10^{-1}$	$3.8 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$8.3 \cdot 10^{-4}$	$9.5 \cdot 10^{-2}$	$4.5 \cdot 10^{-3}$
	NN+NF	$6.1 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$1.0 \cdot 10^{-2}$	$1.8 \cdot 10^{-3}$	$8.9 \cdot 10^{-4}$	$1.6 \cdot 10^{-1}$	$4.1 \cdot 10^{-3}$
	Gain	2.2	3.3	1.4	1.2	1.1	1.6	0.91
$W^- + n$ jets	Sherpa	$2.9 \cdot 10^{-1}$	$4.0 \cdot 10^{-2}$	$7.7 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$9.7 \cdot 10^{-4}$	$1.0 \cdot 10^{-1}$	$4.5 \cdot 10^{-3}$
	NN+NF	$7.0 \cdot 10^{-1}$	$1.5 \cdot 10^{-1}$	$1.1 \cdot 10^{-2}$	$2.2 \cdot 10^{-3}$	$7.9 \cdot 10^{-4}$	$1.5 \cdot 10^{-1}$	$4.2 \cdot 10^{-3}$
	Gain	2.4	3.3	1.4	1.1	0.82	1.5	0.91
$Z + n$ jets	Sherpa	$3.1 \cdot 10^{-1}$	$3.6 \cdot 10^{-2}$	$1.5 \cdot 10^{-2}$	$4.7 \cdot 10^{-3}$		$1.2 \cdot 10^{-1}$	$5.3 \cdot 10^{-3}$
	NN+NF	$3.8 \cdot 10^{-1}$	$1.0 \cdot 10^{-1}$	$1.4 \cdot 10^{-2}$	$2.4 \cdot 10^{-3}$		$1.8 \cdot 10^{-3}$	$5.7 \cdot 10^{-3}$
	Gain	1.2	2.9	0.91	0.51		1.5	1.1

Nested Sampling

Slide by Timo Janßen

Nested Sampling

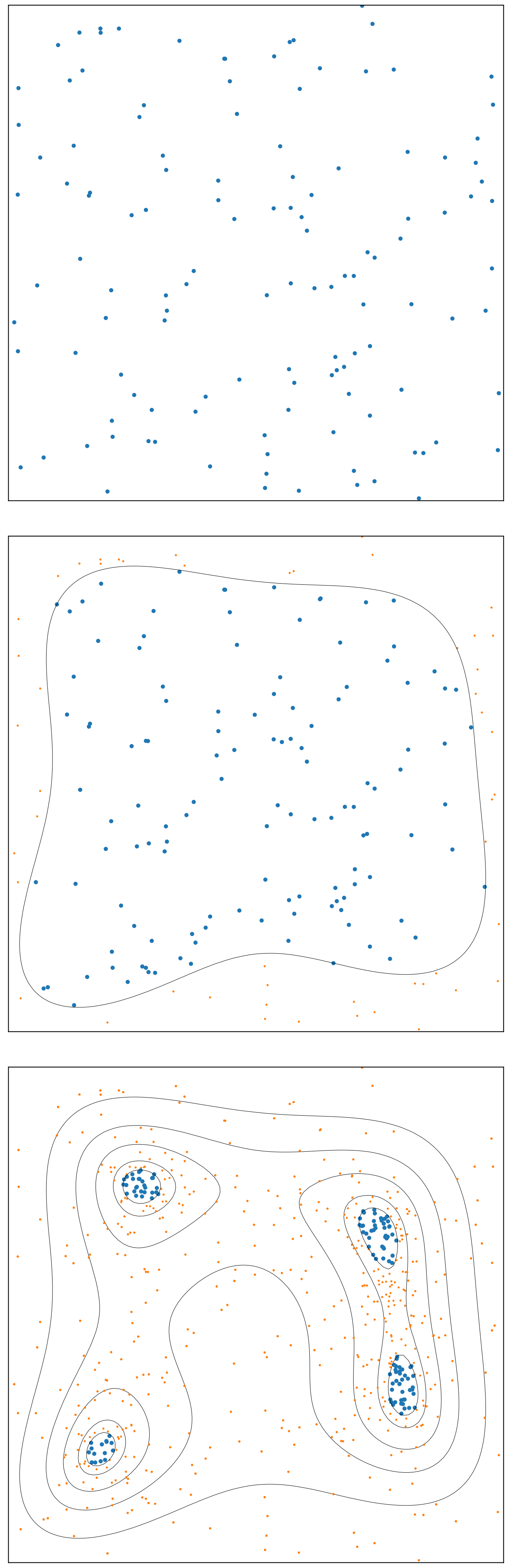
Meta algorithm

- draw ensemble of live points (uniformly)
- sort in order of likelihood, $\mathcal{L}$
- replace $\mathcal{L}_{\text{min}}$ by sampling uniformly, requiring $\mathcal{L} > \mathcal{L}_{\text{min}}$
- repeat until termination criterion reached
- dead points form representative sample of target distribution

Implementation

- PolyChord (Handley et al., 2015)
- use slice sampling (R. Neal, 2003) to evolve live points

$\rightarrow$ many short Markov chains $\rightsquigarrow$ low autocorrelation

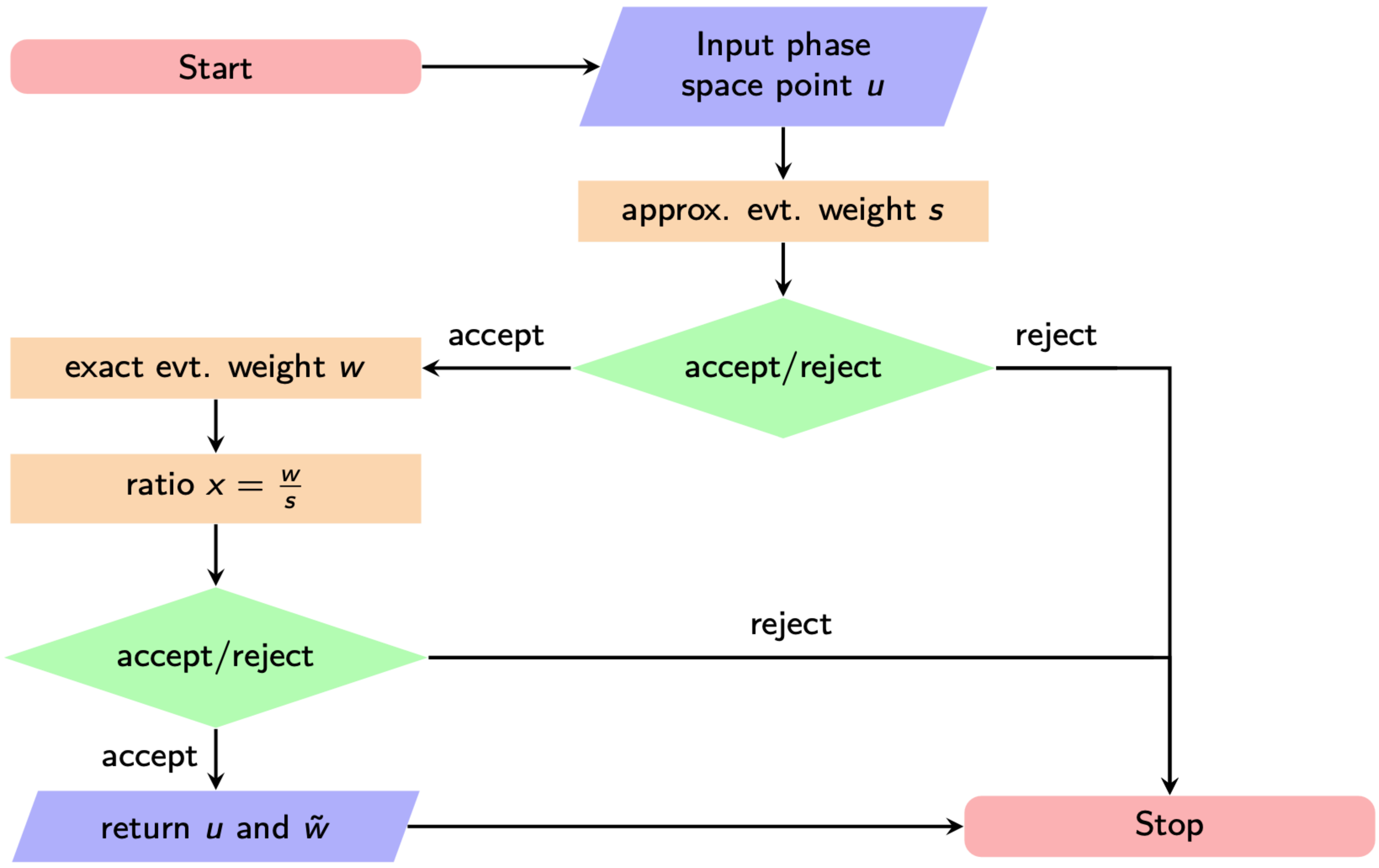


J. Skilling: AIP Conference Proceedings **735**, 395 (2004)

13

Surrogate unweighting

Algorithm [K. Danziger, TJ, S. Schumann, F. Siegert: SciPost Phys. 12, 164 (2022)]



Surrogate unweighting

Slide by Timo Janßen

Factorisation-aware matrix element emulation

soft/collinear factorisation properties

$$|\mathcal{M}_{n+1}|^2 \to |\mathcal{M}_n|^2 \otimes \mathbf{V}_{ijk}$$

[Catani, Seymour Nucl.Phys. B485 (1997) 291-419]

Ansatz

$$\langle|\mathcal{M}|^2\rangle = \sum_{\{ijk\}} C_{ijk} D_{ijk}$$

- $D_{ijk} = \langle V_{ijk} \rangle / s_{ij}$: spin-averaged Catani-Seymour dipoles divided by kinematic invariant
- C_{ijk}: coefficients fit by neural network

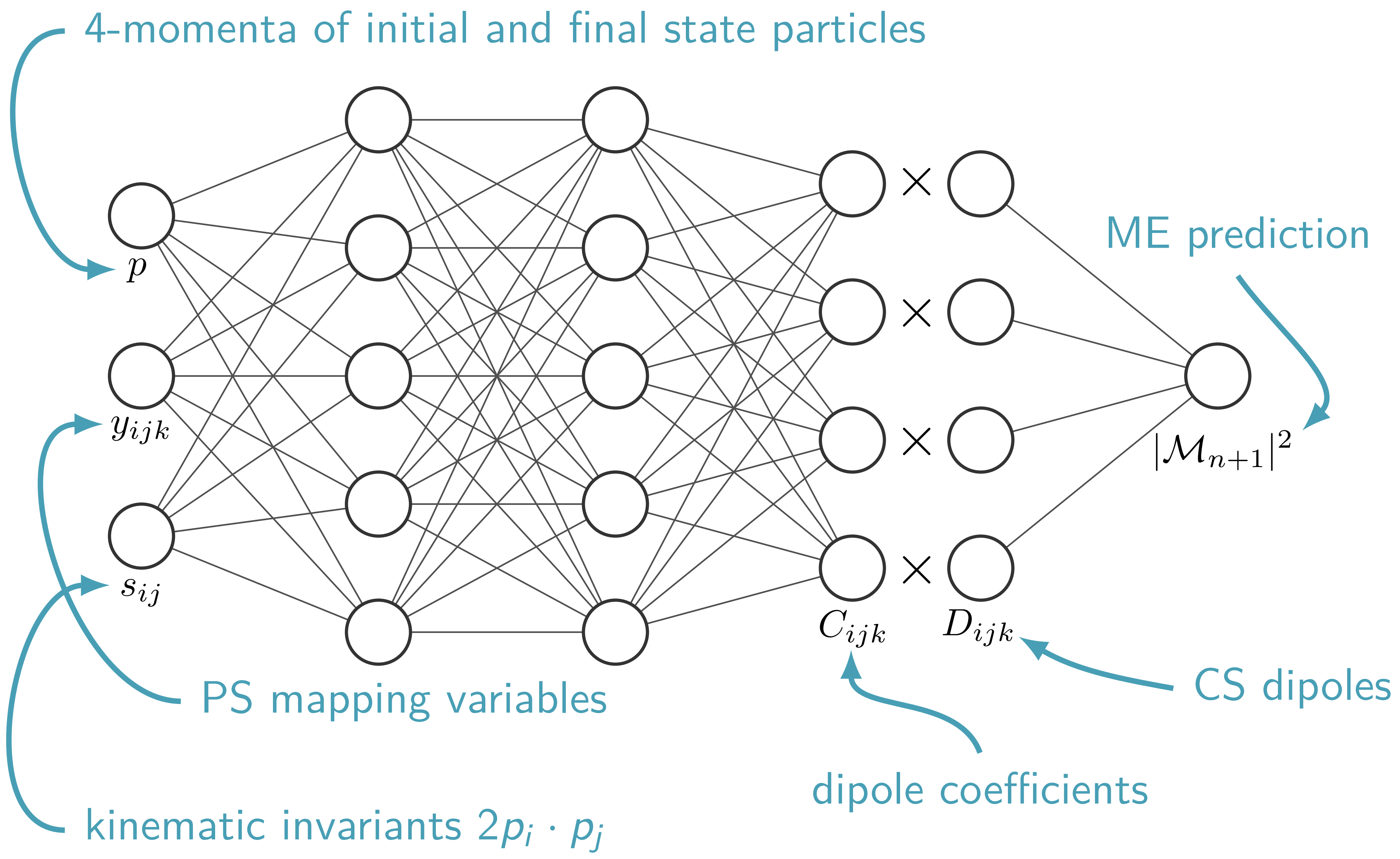
D. Maître, H. Truong: JHEP 11 (2021) 066

20

Surrogate unweighting

Slide by Timo Janßen

Factorisation-aware matrix element emulation



D. Maître, H. Truong: JHEP 11 (2021) 066

21

EWvirt & EWsud

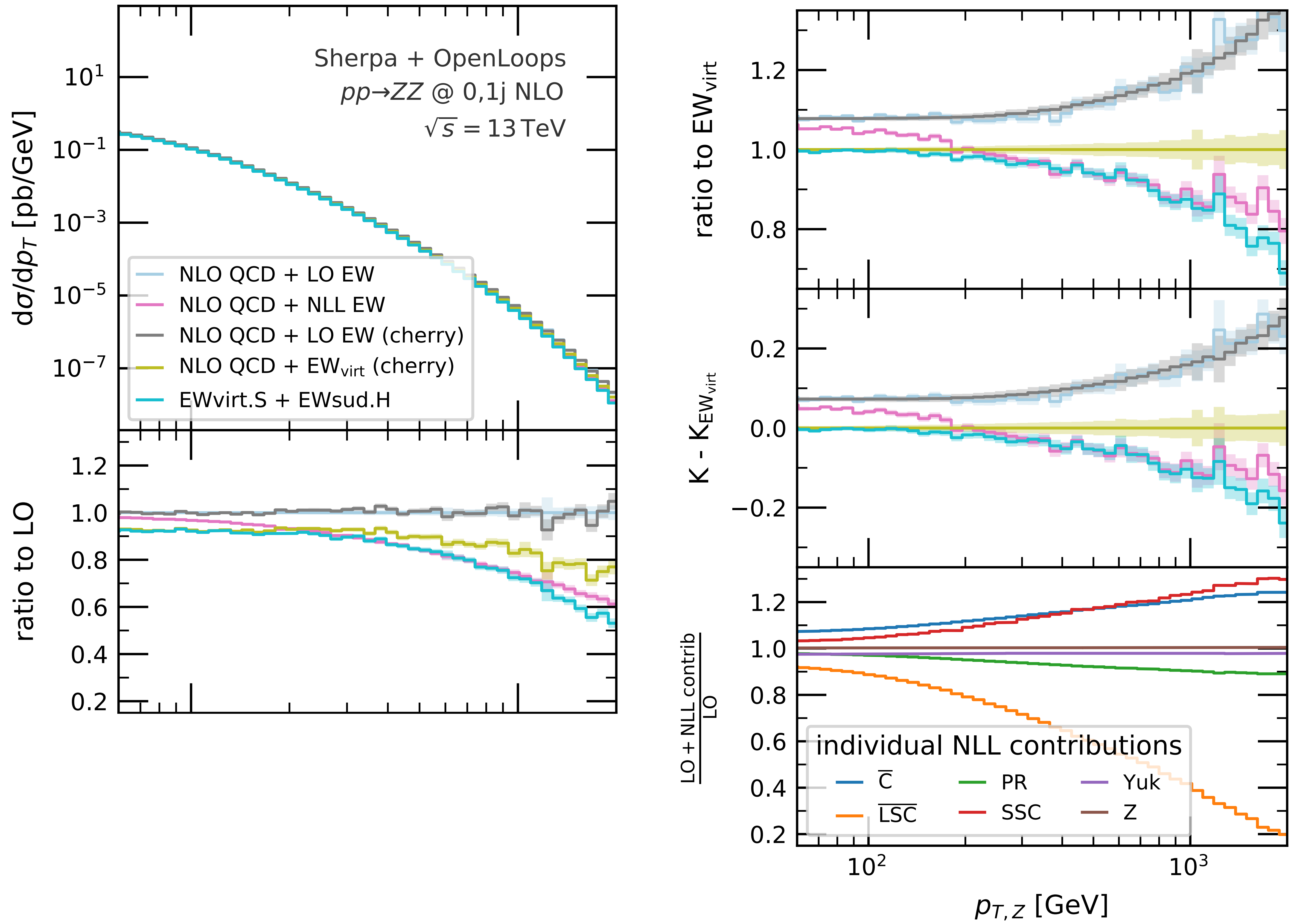
Comparative study in ZZ production

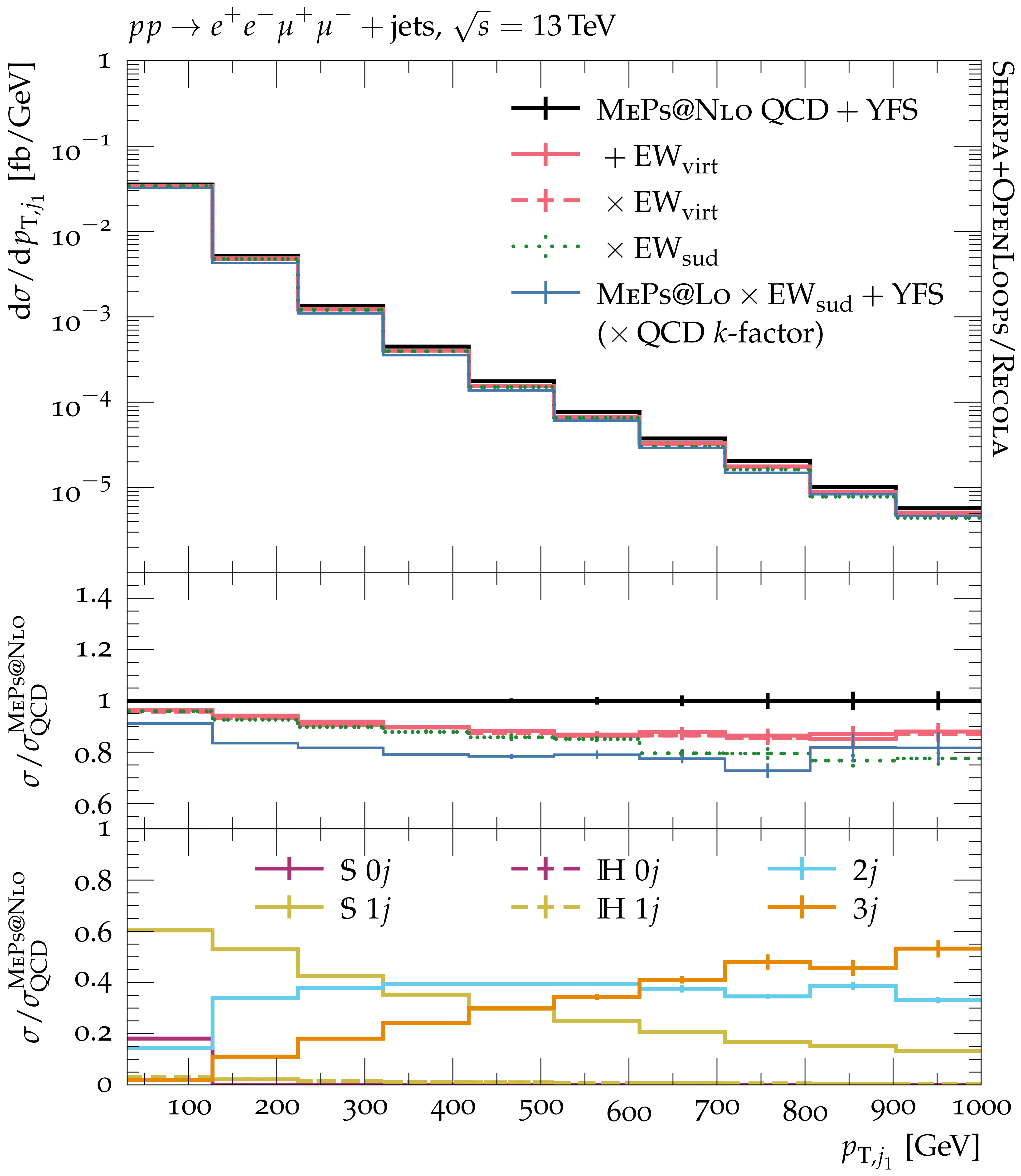
[EB et al. 2111.13453]

- Both schemes capture dominant logs in Sudakov region
- EWvirt:
  - subleading Born (can be sizable, e.g. in 3-jet production) [Reyer Schönherr Schumann 1902.01763]
  - approx. integrated real emission
  - finite terms in virtual loop
  - not applied to real-emission events
  - no subleading logs from RG
  - requires virtual loop ME
- don't expect perfect agreement, but so far we see K factors consistent within couple percent
- **proposal**: apply EWvirt to lower multis and EWsud to real-emission terms and higher multis, in a single merged sample („Hybrid“)

preliminary, MEPS@NLO ZZ production, 0,1j@NLO, 3,4j@LO

Hybrid: EWvirt.S + EWsud.H + EWsud.j4

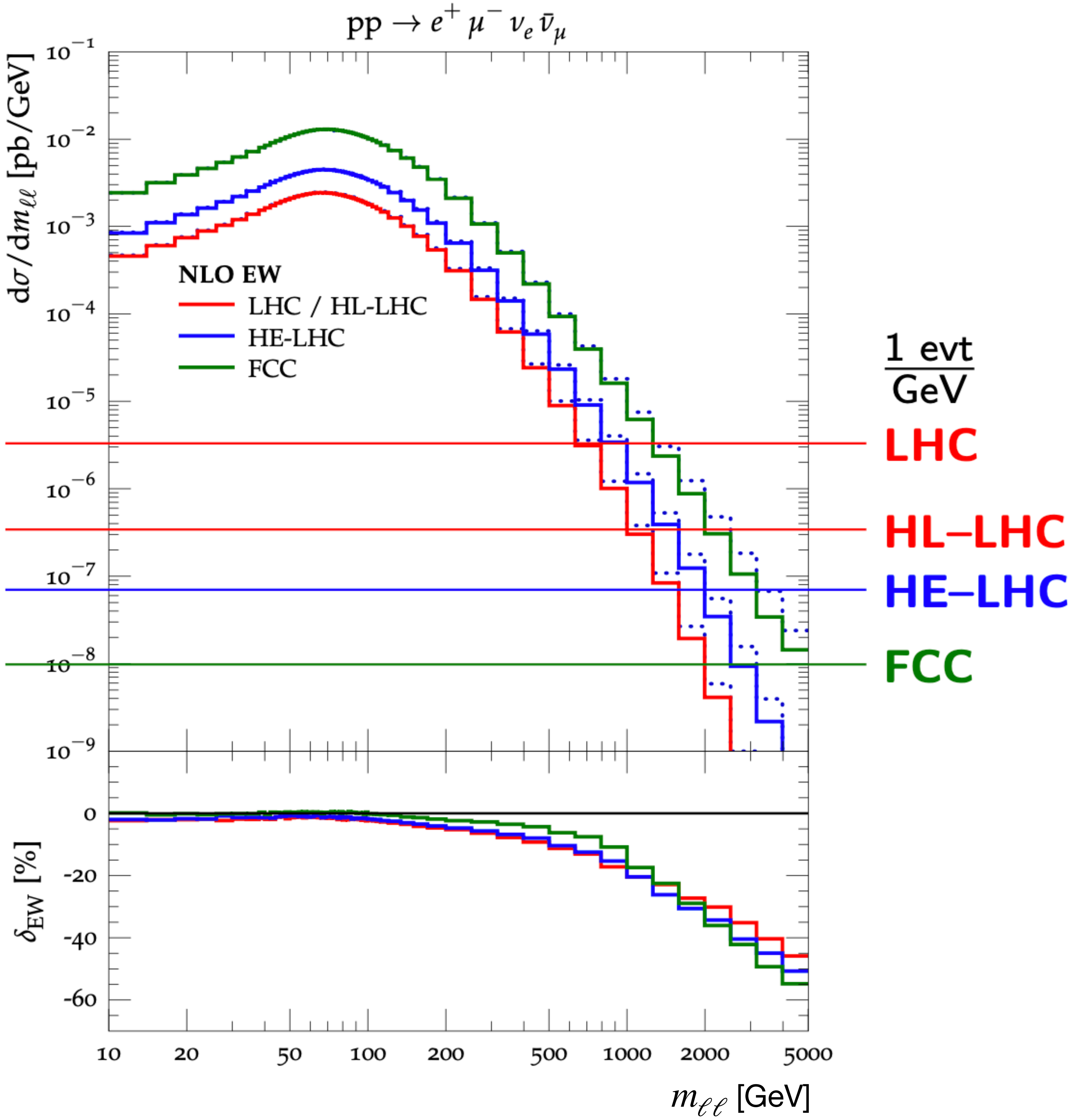




$pp \to e^+e^-\mu^+\mu^- + \text{jets}, \sqrt{s} = 13\,\text{TeV}$
MEPS@NLO QCD + YFS
$+\,\text{EW}_\text{virt}$
$\times\,\text{EW}_\text{virt}$
$\times\,\text{EW}_\text{sud}$
MEPS@LO $\times\,\text{EW}_\text{sud}$ + YFS
($\times$ QCD k-factor)
SHERPA+OPENLOOPS/RECOLA
$d\sigma/dp_{T,j_1}$ [fb/GeV]
$\sigma/\sigma^\text{MEPS@NLO}_\text{QCD}$
S 0j
S 1j
H 0j
H 1j
2j
3j
p_{T,j_1} [GeV]

Collider reach

- Plot taken from a talk by Marek Schönherr
- How far the integrated luminosity takes us into the Sudakov region



improved Z pT modelling by improved shower defaults and better intrinsic kT tune (PRELIMINARY)