QED with McMule for Belle II

Yannick Ulrich

(he/him)

for the McMule team

University of Liverpool

2 March 2026

what is McMule?

fixed-order NNLO QED framework Monte Carlo for MUons and other LEptons

- provided: matrix elements by us or others
- output: physical cross section for any physical observable
- McMule: phase space generation, subtraction, stabilisation, integration, event generation, etc.

- all leptonic $2 \to 2$ processes in QED at NNLO (+ a few others)
- stable public version is an integrator
- generator on development branch

Get the code here: `https://mule-tools.gitlab.io`
Read the docs here: `https://mcmule.readthedocs.io`

McMule

processes in McMule

process	experiment	physics motivation	order
$e\mu \to e\mu$	MUonE	HVP to $(g-2)_\mu$	NNLO+
$\ell p \to \ell p$	P2, Muse, Prad, QWeak, ...	proton radius and weak charge	NNLO
$eN \to eN$	PRad, ULQ2	background	+
$e^-e^- \to e^-e^-$	Prad 2	normalisation	NNLO
	MOLLER, ...	$\sin^2\theta_W$ at low Q^2	
$e^+e^- \to e^+e^-$	any e^+e^- collider	luminosity measurement	NNLO
$ee \to \ell\ell$	VEPP, BES, Daphne, ...	R-ratio	NNLO±
	Belle	τ properties	
$ee \to \gamma\gamma$	Daphne	dark searches	NNLO−
	any e^+e^- collider	luminosity measurement	
$e\nu \to e\nu$	DUNE	flux & $\sin^2\theta_W$	NNLO−
$\mu \to \nu\bar\nu e$	MEG	ALP searches	NNLO+
	DUNE	beam-line profiling	
$\mu \to \nu\bar\nu e\gamma$	MEG, Mu3e, Pioneer	background	NLO
$\mu \to \nu\bar\nu eee$	MEG, Mu3e	background	NLO
$ee \to \pi\pi$	VEPP, BES, Daphne, ...	R-ratio	+
$ee \to \pi\pi\gamma$	VEPP, BES, Daphne, ...	R-ratio	+
$ee \to \ell\ell\gamma$	VEPP, BES, Daphne, ...	R-ratio	+

process	experiment	physics motivation	order
$e\mu \to e\mu$	MUonE	HVP to $(g-2)_\mu$	NNLO$+$
$\ell p \to \ell p$	P2, Muse, Prad, QWeak, ...	proton radius and weak charge	NNLO
$eN \to eN$	PRad, ULQ2	l	$+$
$e^-e^- \to e^-e^-$	Prad 2	on	NNLO
	MOLLER, ...	low Q^2	
$e^+e^- \to e^+e^-$	any e^+e^- col	measurement	NNLO
$ee \to \ell\ell$	VEPP, BES,		NNLO$\pm$
	Belle	s	
$ee \to \gamma\gamma$	Daphne	es	NNLO$-$
	any e^+e^- col	measurement	
$e\nu \to e\nu$	DUNE	θ_W	NNLO$-$
$\mu \to \nu\bar\nu e$	MEG		NNLO$+$
	DUNE	filing	
$\mu \to \nu\bar\nu e\gamma$	MEG, Mu3e, Pioneer	background	NLO
$\mu \to \nu\bar\nu eee$	MEG, Mu3e	background	NLO
$ee \to \pi\pi$	VEPP, BES, Daphne, ...	R-ratio	$+$
$ee \to \pi\pi\gamma$	VEPP, BES, Daphne, ...	R-ratio	$+$
$ee \to \ell\ell\gamma$	VEPP, BES, Daphne, ...	R-ratio	$+$

goal: world domination

measuring $(g-2)_\tau$

BSM effects in $(g-2)_\ell$ scale as $m_\ell^2 \to$ heavier leptons are better probes

- 'classical' methods (stored / trapped) are not possible for τ
- current way: $\gamma\gamma \to \tau\tau$ leads to $\delta(g-2)_\tau \sim 10^{-3}$
  - in $ee \to ee\tau\tau$ [DELPHI 2004]
  - $NN \to NN\tau\tau$ [del Aguila, Cornet, Illana 91, ATLAS 23, CMD 22]
- a sensible probe of BSM would need to be 10^{-6}, based on $(g-2)_\mu$ and $m_\tau^2/m_\mu^2 \sim 300$ scaling
- only process with good enough statistical reach: $ee \to \tau\tau$ on Υ resonance
- since $\sigma \sim |F_1(s) + F_2(s)|^2$ and $(g-2) = F_2(0)$, this would be possible
- however, $F_1 \gg F_2$, would need 10^{-6} measurement of cross section

BSM effects in $(g-2)_\ell$ scale as $m_\ell^2 \to$ heavier leptons are better probes

- ‘classical’ methods (store … e for τ
- current way: $\gamma\gamma \to \tau\tau$ le…
  - in $ee \to ee\tau\tau$ [DELI…
  - $NN \to NN\tau\tau$ [del… 23, CMD 22]
- a sensible probe of BSM … ed on $(g-2)_\mu$ and $m_\tau^2/m_\mu^2 \sim 300$ scaling
- only process with good e… $\to \tau\tau$ on Υ resonance
- since $\sigma \sim |F_1(s) + F_2(s)$… would be possible
- however, $F_1 \gg F_2$, would need 10^{-6} measurement of cross section

asymmetry for $(g-2)_\tau$

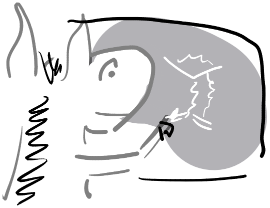
asymmetries to the rescue [Crivellin, Hoferichter, Roney 21]

- we need to cancel the $|F_1|^2$ term which can be done using spin asymmetries
- τ spin can be reconstructed from semileptonic tau decay (similar to muon spin in FNAL experiment)
- start with polarised electron beam (helicity λ), then

$$\sigma_\lambda \sim A\ s_y \Im(F_2) + B\ \lambda s_x \Big(|F_1|^2 + \Re(F_2)\Big) + C\ \lambda s_z |F_1 + F_2|^2$$

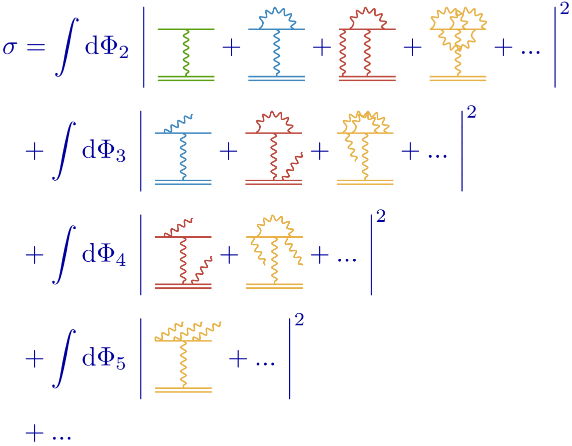
- define helicity difference $\sigma_{\text{pol}} = \sigma_{+1} - \sigma_{-1}$ to isolate B and C term
- integrate over polarisation angles to isolate $\Re(F_2)$
- only problem: need polarised beam to measure $\Re(F_2)$
  (EDM can be measured without this)
- very ambitious measurement (one-loop QED corrections are about 260×10^{-6})

⇒ need theory input



theory background

anatomy of an NNNLO calculation



challenges to overcome

- divergent phase space integration

$\Rightarrow$ FKS$^\ell$

- numerical instabilities

$\Rightarrow$ next-to-soft stabilisation

- virtual amplitudes with $m \neq 0$

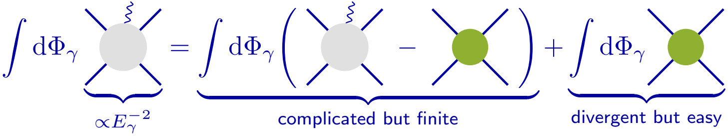
$\Rightarrow$ OpenLoops (one-loop) massification (two-loop)

- negative weights

$\Rightarrow$ cell resampling

FKS$^\ell$ subtraction scheme

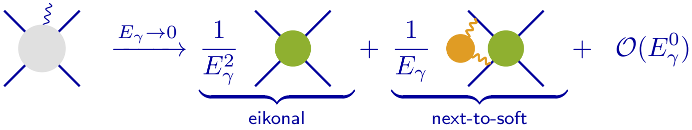
subtract universal counter term from divergent real correction



- works to all order in QED [Engel, Signer, YU 19]
- no resolution parameter ω_c
- unphysical & arbitrary $0 < \xi_c \lesssim 1$
- singularities are treated locally $\to$ stable numerical integration

next-to-soft stabilisation

real-virtual corrections trivial in principle, delicate in practise

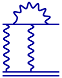


- based on LBK theorem [Low 58; Burnett, Kroll 67] and extensions [Engel, Signer, YU 21; Engel 23]
- if $E_\gamma < E_{\text{NTS}} \approx 10^{-3}\sqrt{s}/2$, switch to NTS expansion rather than full expression
- introduces small theory error $\mathcal{O}(10^{-3}) \times \sigma^{(2)} = \mathcal{O}(10^{-6})$
  $\Rightarrow$ well below the N^{3}LO
- significant speed-up: 7 days vs. 3 months

massification

two-loop integrals with masses are really difficult

- but $m_e^2 \ll m_\tau^2 \sim s \sim Q^2$
- expand in m_e^2/Q^2



$$\sim A \log^2 \frac{m_e^2}{Q^2} + B \log \frac{m_e^2}{Q^2} + C + \mathcal{O}\Big(\frac{m_e^2}{Q^2}\Big)$$

- can be done easily by using $m_e = 0$ result up to three-loop
  [Penin 06; Becher, Melnikov 07; Engel, Gnendiger, Signer, YU 18; YU 23]
- introduces small theory error $\mathcal{O}(10^{-2}) \times \sigma^{(2)} = \mathcal{O}(10^{-5})$
  $\Rightarrow$ well below statistical error

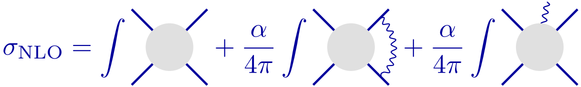
all of this was for an integrator

- calculate arbitrary differential distributions
- event generation by just dumping momenta to file ("garden hose approach")
- if $r \times N$ of N weights are negative, we need $\propto 1/(1-2r)^2$ events

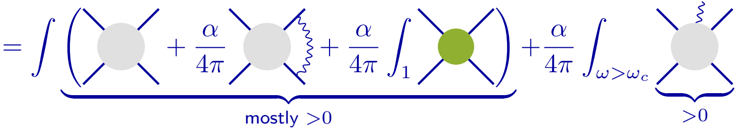
$\Rightarrow$ reduce r as much as possible by cancelling negative weights as early as possible

- optimisations from splitting integrand goes away

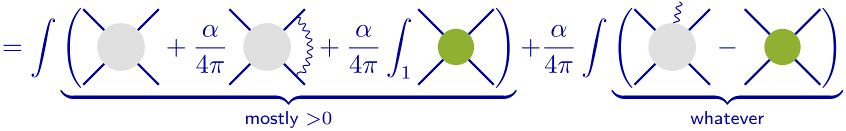
... at NLO for simplicity



- slicing: fairly few negative weights but numerically construct $\log\omega_c$



- subtraction: easier integration but lots and lots of negative weights ($\mathcal{O}(5\%)$ at NLO, more at NNLO)



can we get both?

two observations

1. cross section $\sigma = \int_\mathcal{C} \mathrm{d}\sigma > 0$, irregardless of the size of integration region $\mathcal{C}$
2. experiments have a finite resolution

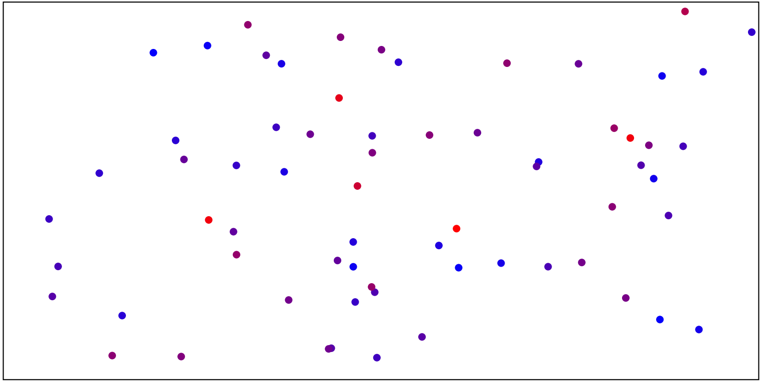
   (we already knew that because we can't see soft photons)

algorithm to remove negative weights [Andersen, Maier 21]

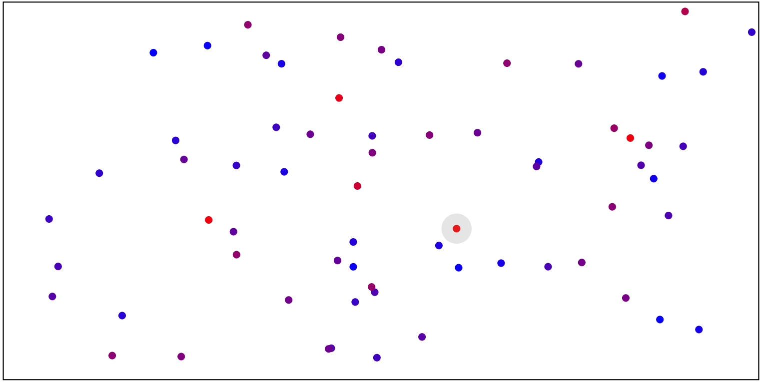
- pick an event with $w_i < 0$
- find nearby events until $\sum_{i\in\mathcal{C}} w_i > 0$
- if $\mathcal{C}$ gets too big (events become resolvable), abort (or add more events)
- else $w_i \to \frac{\sum_{j\in\mathcal{C}} w_j}{\sum_{j\in\mathcal{C}} |w_j|} w_i$

we can remove negative weights without biasing physical observables!

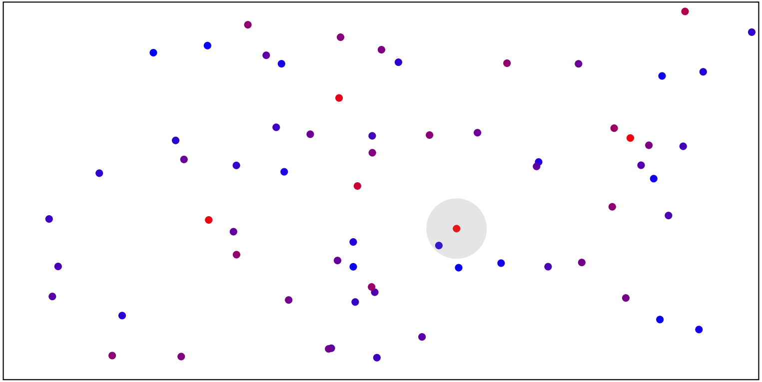
example



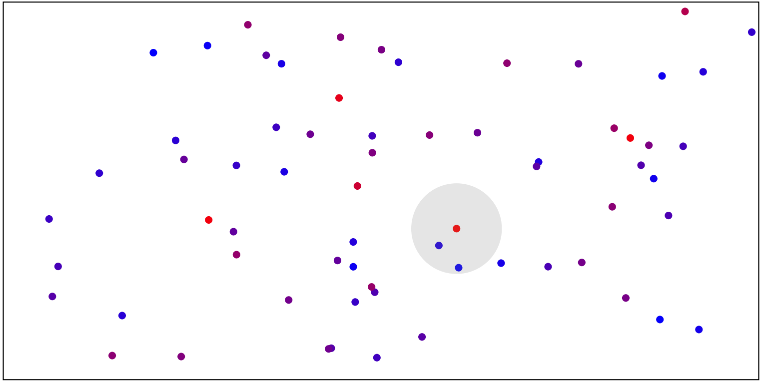
example



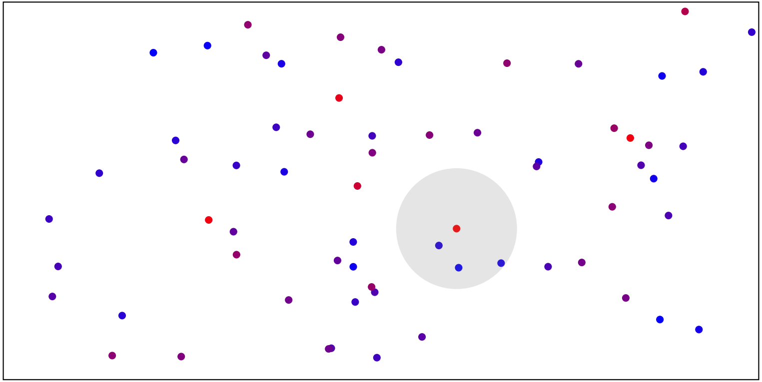
example



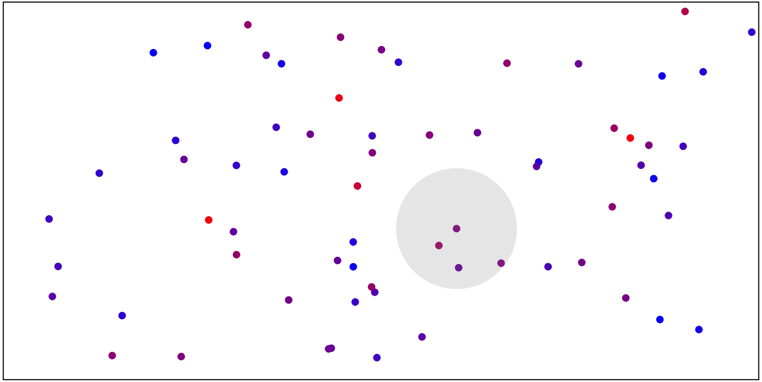
example



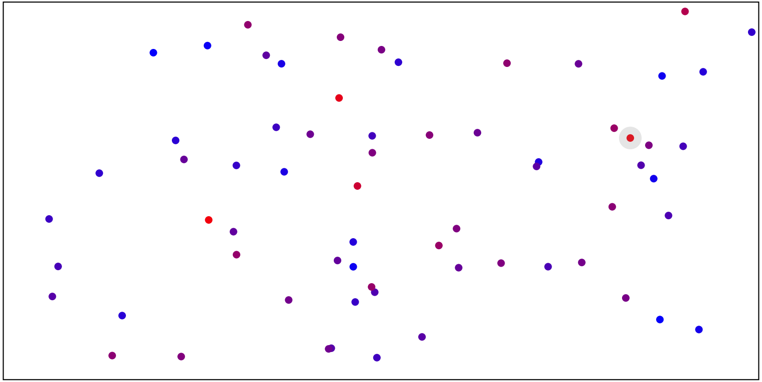
example



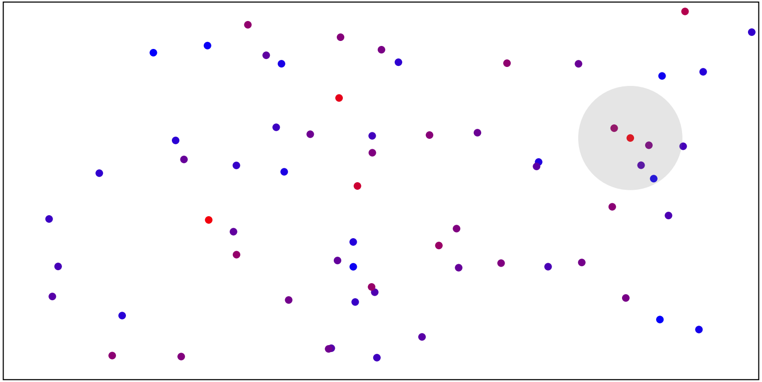
example



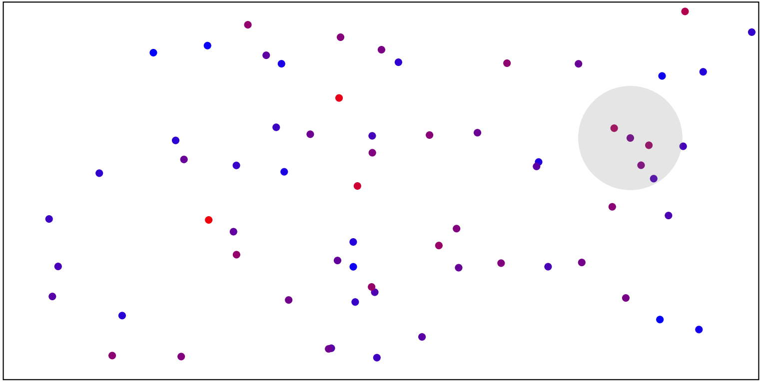
example

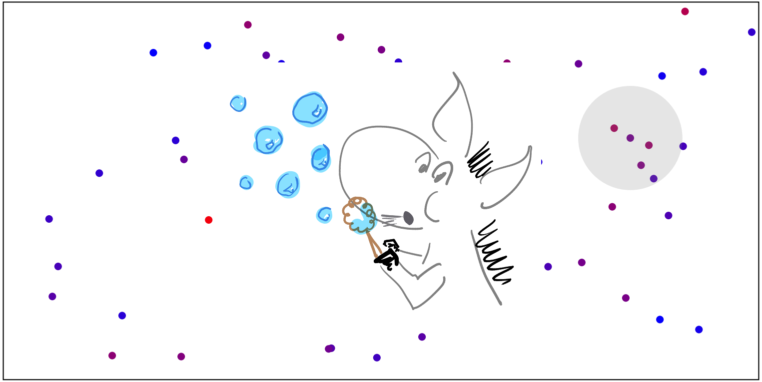


example



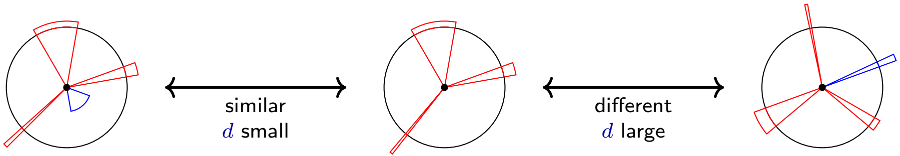
example





we need to define a metric in event space $d(e_1, e_2) \geq 0$

- doesn't really matter how we do this as long as IR safe
  (events with soft photons are near each other)
- ideally: events that look similar are closer to each other than those that don't



results for Belle II

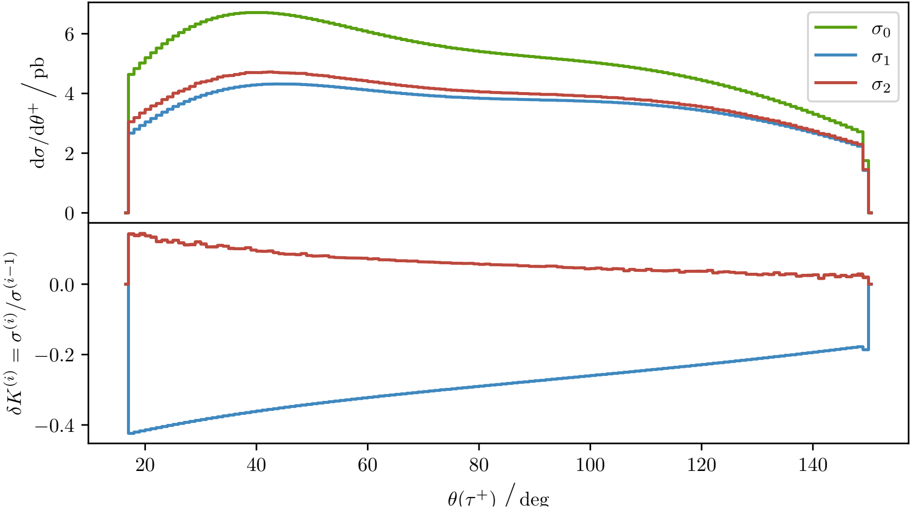
setup for Belle II

the following is an example and would need to be completed / tweaked!

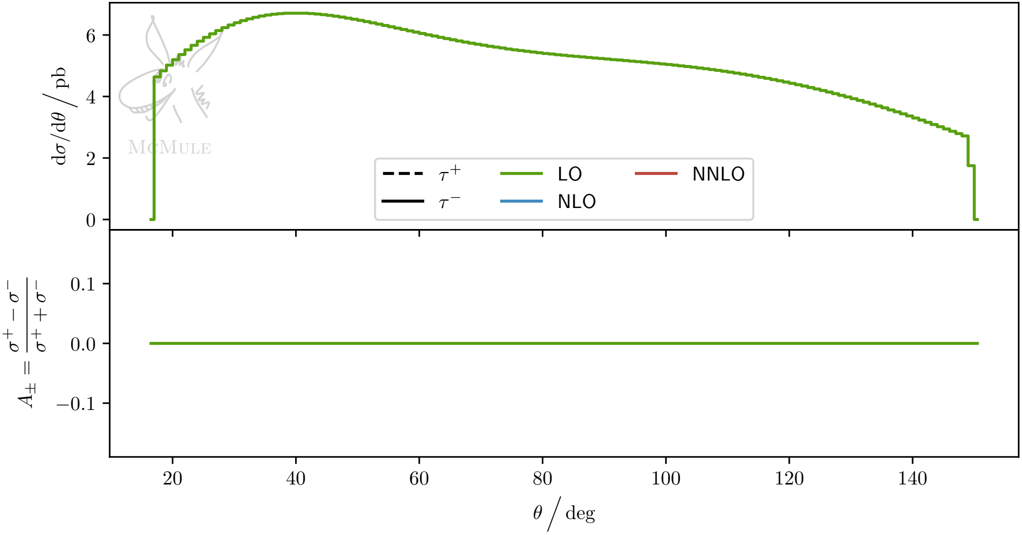
- two beams: $4\,\mathrm{GeV}$ (e^-) and $4\,\mathrm{GeV}$ (e^+)
- cuts: $17^\circ < \theta < 150^\circ$, no photon with $E_\gamma > 50\,\mathrm{MeV}$
- for today: no polarisation but NNLO from [Broggio, Engel, Ferroglia, Mandal, Mastrolia, Rocco, Ronca, Signer, Torres Bobadilla, Zoller, YU 22]
- for polarisation at NLO, see [Gogniat, Hoferichter, YU 25], for EW see [Kollatzsch, YU 22]
- split into initial-state corrections (easy) and rest (hard)
- about 9.5 CPU months ($\lesssim$ 2 weeks wall clock)

find runcards and analysis scripts here

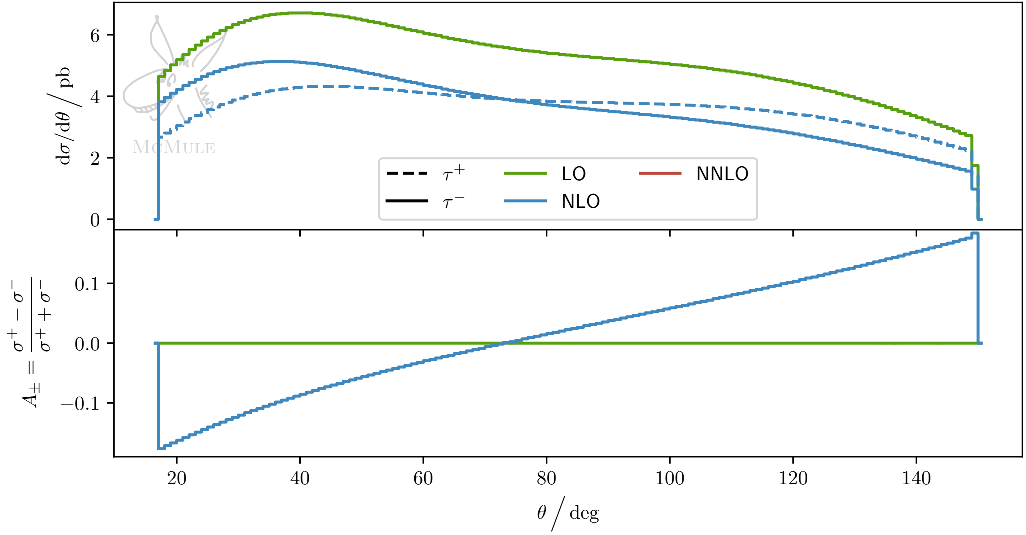
size of corrections



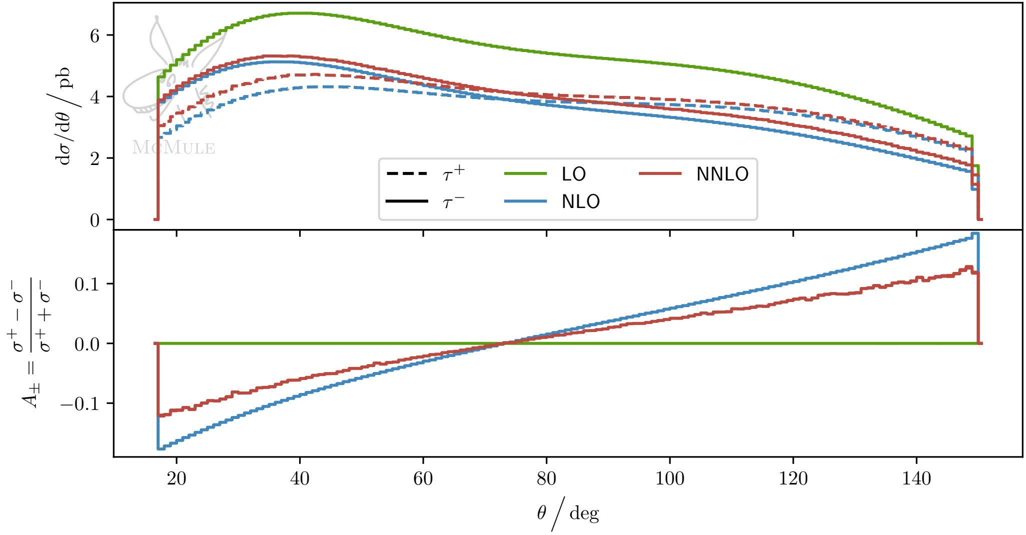
τ^+-τ^- asymmetry



τ^+-τ^- asymmetry

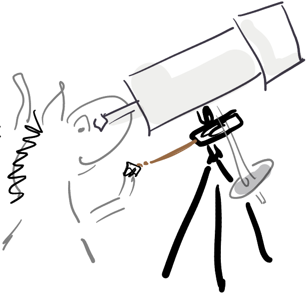


τ^+-τ^- asymmetry



specific to $ee \to \tau\tau$

- fully polarised NNLO using form factors by [Ronca, Torres Bobadilla 2?]
- approximate N^{3}LO [Badger, Czakon, Hartanto, Moodie, Peraro, Poncelet, Zoia 23; Crisanti, Dave, Mastrolia, Ronca, Smith, Torres Bobadilla 26]
- full mass dependence at NNLO plausible but very difficult



general

- merging event generator
- resummation of soft photon emission
- miscellaneous performance improvements

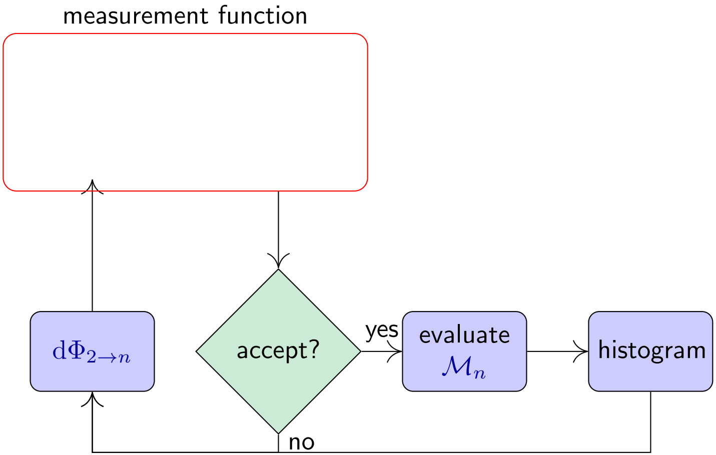
please get in touch so we can help you use McMule!
yannick.ulrich@cern.ch

f.l.t.r.: S.Kollatzsch (Zurich & PSI), A.Signer (Zurich & PSI), V.Sharkovska (Zurich & PSI), S.Gündogdu (Zurich & PSI), D. Moreno (PSI), A.Coutinho (IFIC), Y.Ulrich (Liverpool), D. Radic (Zurich & PSI), L.Naterop (Zurich & PSI), M.Rocco (Turin)
not shown: not shown: F.Hagelstein (Mainz), N.Schalch (Oxford), P.Banerjee (Cosenza), M.Ronchi (Mainz), Y.Fang (PSI), G.Billis (PSI), J.Wilson (Liverpool)

McMule

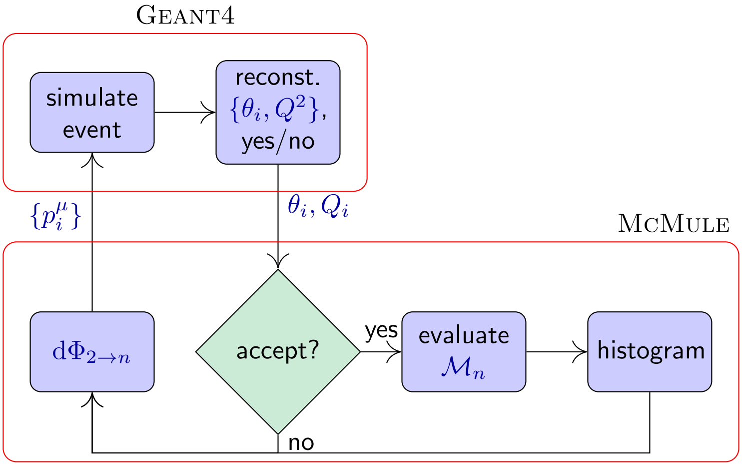
`mule-tools.gitlab.io`

- PRad II is adopting McMule into their analysis flow



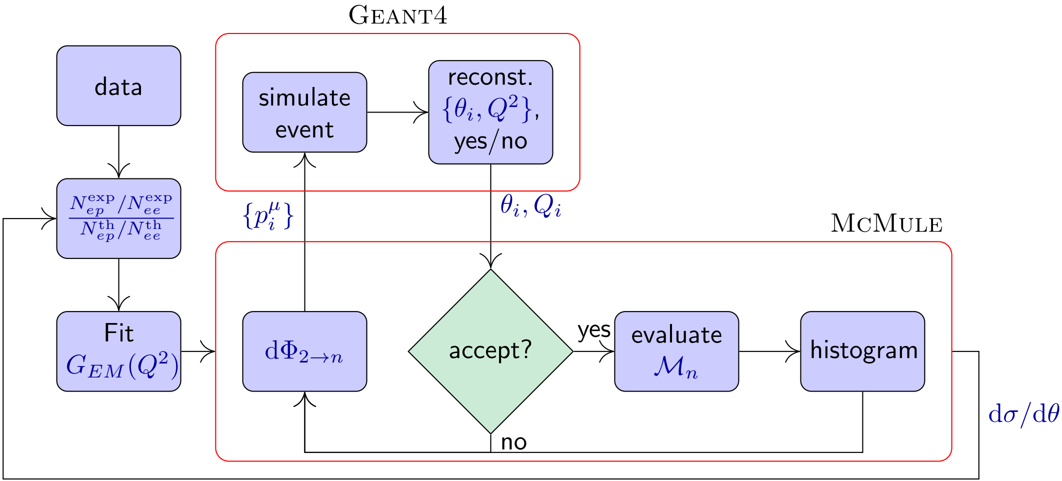
gitlab.comyannickulrich/simple_geant4_package

- PRad II is adopting McMule into their analysis flow



gitlab.comyannickulrich/simple_geant4_package

- PRad II is adopting McMule into their analysis flow



gitlab.comyannickulrich/simple_geant4_package

- buidling & validating generator
- generator efficiency $\sim (w_{\rm max} - w_{\rm min})/\langle w\rangle$

Order	ξ_c	N	r	$[w_{\rm min}, w_{\rm max}]/\langle w\rangle$	feature
LO	n/a	99.7M	0.0%	$[0.03, 7.7]$	
NLO	0.1	220M	0.5%	$[-3.6, 3.6]\times 10^7$	
	1.0	216M	3.5%	$[-2.1, 2.1]\times 10^6$	
NNLO	0.1	21.1G	0.7%	$[-8.6, 8.6]\times 10^7$	
	1.0	19.8G	9.2%	$[-2.3, 2.3]\times 10^8$	

- buidling & validating ger
- generator efficiency $\sim$ $(\imath$

Order	ξ_c	N	]/$\langle w\rangle$	feature
LO	n/a	9		
NLO	0.1	22	10^7	
	1.0	2	10^6	
NNLO	0.1	2	10^7	
	1.0	1	10^8	

- buidling & validating generator
- generator efficiency $\sim (w_{\text{max}} - w_{\text{min}})/\langle w \rangle$

Order	ξ_c	N	r	$[w_{\text{min}}, w_{\text{max}}]/\langle w \rangle$	feature
LO	n/a	99.7M	0.0%	$[0.03, 7.7]$	
NLO	0.1	220M	0.5%	$[-3.6, 3.6] \times 10^7$	
			0.1%	$[-1.3, 19] \times 10^3$	cres
	1.0	216M	3.5%	$[-2.1, 2.1] \times 10^6$	
			0.3%	$[-7.0, 26] \times 10^3$	cres
NNLO	0.1	21.1G	0.7%	$[-8.6, 8.6] \times 10^7$	
		3.6G	0.1%	$[-6.9, 6.9] \times 10^5$	cres
	1.0	19.8G	9.2%	$[-2.3, 2.3] \times 10^8$	
		1.2G	0.5%	$[-3.3, 2.4] \times 10^4$	cres

- buidling & validating generator
- generator efficiency $\sim (w_{\rm max} - w_{\rm min})/\langle w\rangle$

Order	ξ_c	N	r	$[w_{\rm min}, w_{\rm max}]/\langle w\rangle$	feature
LO	n/a	99.7M	0.0%	$[0.03, 7.7]$	
NLO	0.1	220M	0.5%	$[-3.6, 3.6]\times 10^7$	
			0.1%	$[-1.3, 19]\times 10^3$	cres
		51M	0	$[+0.0, 1.6]\times 10^4$	subs
	1.0	216M	3.5%	$[-2.1, 2.1]\times 10^6$	
			0.3%	$[-7.0, 26]\times 10^3$	cres
		61M	0	$[+0.0, 2.5]\times 10^4$	subs
NNLO	0.1	21.1G	0.7%	$[-8.6, 8.6]\times 10^7$	
		3.6G	0.1%	$[-6.9, 6.9]\times 10^5$	cres
	1.0	19.8G	9.2%	$[-2.3, 2.3]\times 10^8$	
		1.2G	0.5%	$[-3.3, 2.4]\times 10^4$	cres

- buidling & validating generator
- generator efficiency $\sim (w$ … $- w$ … $)/\langle w\rangle$

Order	ξ_c	N…		…$]/\langle w\rangle$	feature
LO	n/a	9…			
NLO	0.1			10^7	
		22…		…0^3	cres
		51…		10^4	subs
	1.0			10^6	
		21…		…0^3	cres
		61…		10^4	subs
NNLO	0.1	21…		10^7	
		3.6G	0.1%	$[-6.9, 6.9] \times 10^5$	cres
	1.0	19.8G	9.2%	$[-2.3, 2.3] \times 10^8$	
		1.2G	0.5%	$[-3.3, 2.4] \times 10^4$	cres