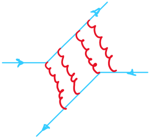
Perturbative Amplitudes

Present Computational Needs
(and hints about the Future)



Fernando Febres Cordero
Department of Physics, Florida State University

Snowmass Community Planning Meeting,
84. Computing Requirements & Opportunities in Theory,
October 2020

Collecting Information

Gathering initial data on computing requirements by several state-of-the-art precision calculations for the LHC:

1. CPU hours, GPU usage/testing
2. Disk space requirements, I/O needs
3. RAM usage
4. HPC readiness, batch needs

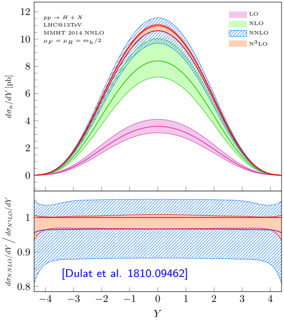
As reference we survey several very recent studies:

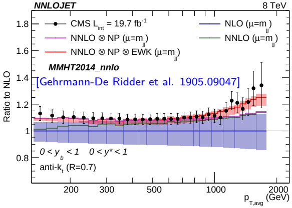
- High-multiplicity NLO QCD: fully off-shell studies of [$t\bar{t}j$], [$t\bar{t}\gamma$] and [$t\bar{t}W$] production, [6-particle 1-loop analytic amps] · · ·
- Loop-induced NLO QCD with many scales: [H+jet], [HH], [EFT HH] production, · · ·
- NNLO QCD: [H+jet with MCFM], [H+jet with NNLOJet], triple-photon production by [Chawdry et al.] and by [Kallweit et al.] (to appear), · · ·
- N^3LO QCD corrections: [inclusive W] production
- Multi-loop, multi-scale amplitudes and form factors: [four-loop QCD form factors], [three-loop massive self energies], [2-loop five-parton amplitudes], [full-color 2-loop five-gluon amplitude], [Ntuples @ NNLO] · · ·
- Multi-loop, multi-scale integrals: [2-loop five-point massless integrals], [2-loop five-point 1-mass integrals], · · ·

Disclaimer: more systematic review needed!

Fixed order calculations for collider phenomenology

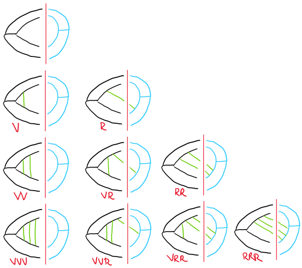
- First principle calculations, relating fundamental theory parameters to measurements
- Systematically improvable through higher-order perturbative effects
- Though limited to inclusive observables, can be interfaced to Monte Carlo programs to simulate realistic collider events
- A necessity in the search of new physics at hadron colliders





Anatomy of N^{k}LO calculations

Take dijet production at hadron colliders, where :



		typical th uncertainties
LO:	α_s^2	$\sim 50\%$
NLO:	α_s^3	$\sim 17\%$
NNLO:	α_s^4	$\sim 7\%$
N^{3}LO:	α_s^5	few $\%$

State-of-the-art in QCD corrections for dijet production is NNLO

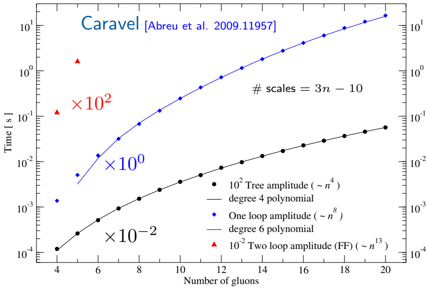
[Gehrmann-De Ridder, Gehrmann, Glover, Huss, Pires], [Czakon, van Hameren, Mitov, Poncelet]

Key ingredients:

- Strategy to handle and cancel IR divergences [See Caola's talk at CPM # 128]
- Multi-loop multi-leg scattering amplitudes [See Ita's talk at CPM # 128]

Multi-scale amplitudes: computational complexity

Take as ex. the numerical computation of n-gluon color-ordered amps



- Polynomial complexity to compute color-ordered amplitudes
- Dramatic computational increase in loop order
- Asymptotic behavior characterizes algorithm, but minimal impact in pheno
- Combinatorial growth in amplitudes needed for summed MEs

Analytic computations for processes with not many scales (say 5 - 8) can considerably improve efficiency $\longrightarrow$ tame/handle the typical analytical exponential complexity growth!

Amplitude analytic expressions found for:

- selected 6-particle 1-loop amps
- essentially all 4-particle 2-loop amps
- selected 5-particle 2-loop amps
- selected 4-particle 3-loop amps

Computer needs (I)

- **NLO QCD corrections:** largely automated and efficient tools. Frontier includes high-multiplicity processes. Highlights:
  - ∼200k CPU hours for typical analysis
  - ∼100GB output files for flexibility
  - few GBs of RAM per running core
  - HPC ready, MPI for preparation, MPI/array jobs for event sampling. Common cluster setup $100 - 1000$ cores, jobs lasting order 1 day
  - **Recent** algorithms developed to extract multi-scale (6-particle) 1-loop analytic expressions from high-precision floating-point evaluations
- **Loop-induced NLO QCD corrections:** exploits numerical integration of multi-scale two-loop amplitudes. Frontier heavy-particle production with loop-mass effects (where analytic control is limited). Highlights:
  - ∼100k CPU hours typical analysis, preparation stages might require building blocks (phase-space point evaluations) with a median of about 2 hours (but a large variance)
  - Mid disk needs, object files can reach hundreds of MB, compile time hours and up-to few days, and executables in total reach ∼100GB
  - HPC employed including clusters with CPUs and with GPUs
  - **Recent** improvements exploit quasi-random integration sampling, multi-variate interpolation grid, matching to analytic series expansions, etc

Computer needs (II)

- **NNLO QCD corrections:** huge progress over the last decade. $2 \to 2$ processes available. Frontier lies in the recent triphoton studies. Some programs available publicly (Matrix, MCFM, $\cdots$). Highlights:
    - Few 10k's and upto 100k CPU hours for typical runs, commonly needing several iterations to obtain final results for an analysis
    - Initial exploration of ntuple-format files show large sets, in the order of ∼1TB of disk space
    - HPC exploited for Monte Carlo integration over phase space in clusters of order 1000 cores, jobs lasting few days
    - **Large amount of developments** including building multi-variate interpolation grids with machine-learning algorithms for optimization and error estimates, aiding numerical algorithms with analytic building blocks, optimizing special function evaluations, etc.
- **Multi-loop integrals and amplitudes:** state-of-the-art analytic computations deal with huge computational challenges, normally associated with a swell of *intermediate* terms. Frontier depends on loop order, associated to the number of scales in the problem. Highlights:
    - Recent analytic computations have reach order 100k CPU hours, with an important level of uncertainty of requirements
    - Often few TBs of RAM required, potentially covered by swap partitions on SSDs. On array jobs often *main* thread is the only high-memory consuming
    - Processing large amounts of temporary data, up to 100TB in over a day of running. *Checkpointing* files can easily grow to 1TB
    - Recently HPC employed for computer algebra procedures. Peculiar challenges on batch policies given complexity of computer algebra algorithms, and on the usage of proprietary software (Mathematica, Maple, etc)
    - **Huge progress in recent years** including introduction of finite field arithmetic in QFT calculations, functional reconstruction algorithms, algebraic geometry techniques, differential equations for Feynman integrals and other numerical approaches, etc. Often after a *heavy* computation results are relatively compact and highly efficient (non-trivial simplification procedures!)

Computer needs (III)

- **N^{3}LO QCD corrections:** first-ever computations at this level for hadron colliders recently achieved: Higgs production (differential on rapidity) and W (Drell-Yan) production. They are expressed as single-variable expressions and, though advanced analytic tools are required to produce them, computational requirements are relatively small. This will dramatically change for more complicated processes or for fully differential calculations. A rule of thumb for example for ggH fully differential is that complexity will be at least that of $gg \to gH$ production at NNLO (stable over unresolved regions!)

Outlook

- Current computational needs for state-of-the-art calculations mostly covered by mid-size clusters
- While moving from current capabilities (see e.g. [Les Houches Wishlist]) to new challenges for the HL-LHC era computer-time usage will surge (typical over 1M CPU hours per analysis ??)
- HPC systems will be required, most notably for computer algebra:
  - Dedicated node pools (high-memory, large-disk resources) within large HPC systems?
  - Hybrid schemes for time limits in queueing systems?
  - Checkpointing through storage of virtual machines?
  - Licensing schemes for proprietary software in large HPC systems?
  - Several attempts to use GPUs, flexibility limited
- As complexity of computer libraries grows, need for dedicated computer science support for development and maintenance of our software frameworks