CF2: Lessons Learned from CPM Perturbative Calculations

Andreas von Manteuffel (Michigan State University)

CPM session 128: From Amplitudes to Precision Theory for Future Colliders

Extending subtraction schemes to N3LO	*Fabrizio Caola*	
Zoom 6	13:30 - 13:48	
Discussion		
Zoom 6	13:48 - 13:53	
EW effects in parton showers at high energy	*Bryan Webber*	
Zoom 6	13:53 - 14:11	
Discussion		
Zoom 6	14:11 - 14:16	
Five-point QCD amplitudes at two loops	*Harald Ita*	
Zoom 6	14:16 - 14:34	
Discussion		
Zoom 6	14:34 - 14:39	
Higher-order corrections for HZ production and related techniques	*Zhao Li*	
Zoom 6	14:39 - 14:57	
Discussion		
Zoom 6	14:57 - 15:02	
General discussion and closing remarks	*Session organizers and participants*	
Zoom 6	15:02 - 15:15	

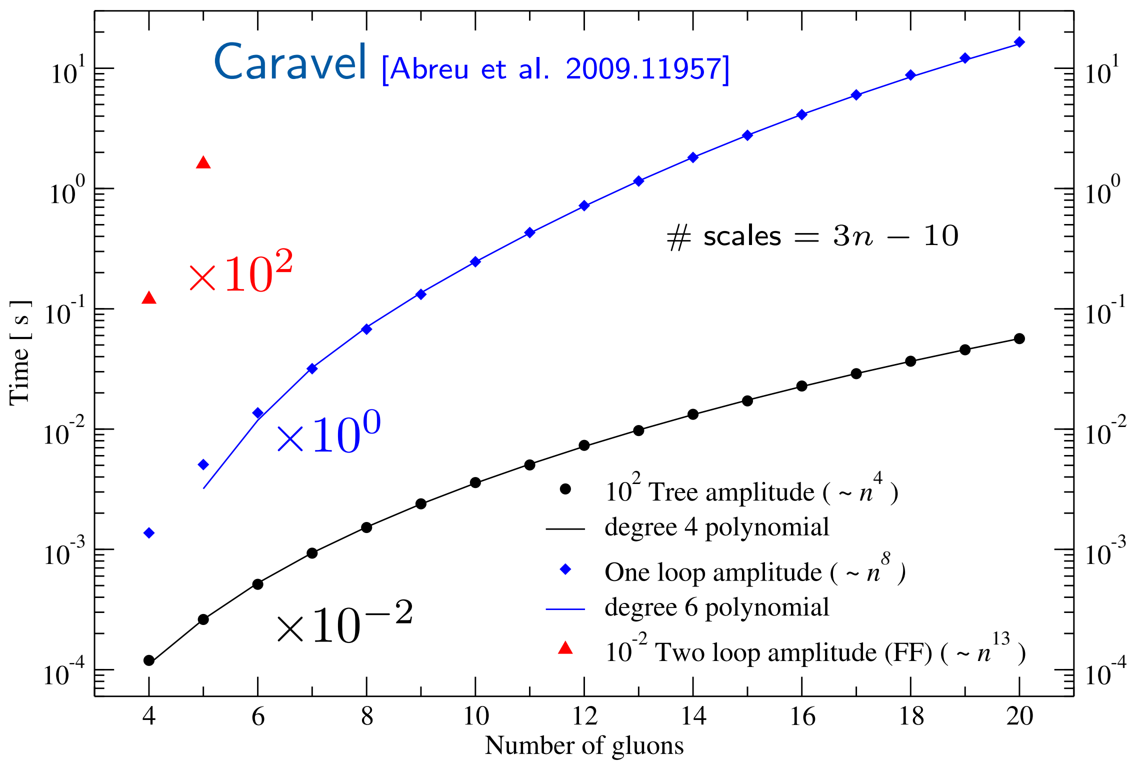
CPM session 84:
Computing Requirements & Opportunities in Theory

Computing roadmap overview	*Peter Boyle*
Zoom 10	12:45 - 12:55
Summary of needs in lattice	*Richard C. Brower*
Zoom 10	12:55 - 13:05
Summary of needs in Conformal Bootstrap	*David Simmons-Duffin*
Zoom 10	13:05 - 13:15
Summary of needs in Perturbative Amplitudes	*Fernando Febres Cordero*
Zoom 10	13:15 - 13:25
Discussion and Next Steps	
Zoom 10	13:25 - 13:45

- One LOI with dedicated CompF2 / Pert. Calc. component (Febres Cordero, Neumann, von Manteuffel)

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

[Slide from Fernando Febres Cordero's talk in session 84]

Computer needs (I)

- **NLO QCD corrections:** largely automated and efficient tools. Frontier includes high-multiplicity processes. Highlights:
  - $\sim$200k CPU hours for typical analysis
  - $\sim$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:
  - $\sim$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 $\sim$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

[Slide from Fernando Febres Cordero's talk in session 84]

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 $\sim$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!)

[Slide from Fernando Febres Cordero's talk in session 84]

Computer needs (III)

- **N^3LO 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!)

[Slide from Fernando Febres Cordero's talk in session 84]

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

[Slide from Fernando Febres Cordero's talk in session 84]