Computing in Perturbative QCD - Draft Charge

Stefan Höche and Laura Reina[2]

Community Planning Meeting (CPM2012)

FNAL, 10/12/12

[2] for the Computing Frontier Working Group CpF T4 (perturbative QCD)

Scope of the study

Define computing needs to carry out calculations required to answer the physics questions posed by Energy Frontier groups

- HE1: The Higgs Boson
- HE2: Precision Study of Electroweak Interactions
- HE3: Fully Understanding the Top Quark
- HE5: Quantum Chromodynamics and the Strong Force

Project into the future whether needs can be met

If they can only be met through new research in computer science or technology set research agenda

Background of this study

- Broad impact of perturbative QCD on collider physics
- Large particle multiplicity needed in calculations to meet experimental needs (for both signal and background)
- Higher perturbative orders required for stability/accurateness of theoretical predictions
- Sophisticated computational means to face challenges and provide flexible/fast/reliable theoretical tools

Building on the enormous progress in NLO and NNLO calculations in recent years, in the context of Snowmass 2013 we would like to

- Share and document experiences
- Identify successful computational models
- Open the discussion for new ideas and further improvements

NLO

Status

- Conceptual/technical challenges largely met
- New & old techniques for one-loop QCD implemented or being implemented in several (public) codes
- Interface with Parton Shower Monte Carlo at NLO available

Issues to consider

- Availability of codes, grade of automation, expandability versatility (e.g. implementation of cuts, jet vetos), user friendliness
- Scaling: How do automated codes perform with increasing number of particles (massive particles)? (Example: processes like $t\bar{t}jj$ or $t\bar{t}b\bar{b}$ at NLO still far from being available in an efficient way).
- Can improved computing help to better exploit existing tools? (e.g. provide power to run w/ different parameters or cuts provide storage needed for large event files / ntuples)

NNLO

Status

- State of the art is $2 \to 2$ processes with massive particles (e.g. $t\bar{t}$ hadroproduction) or $2 \to 1$ processes fully differentially.
- Still big challenges to be met in computing both two-loop corrections and double-parton emission (still building tools).

Questions to be addressed

- How are NNLO calculations evolving: Are they going to be mainly analytical or mainly numerical in nature?
- Can computational issues and bottlenecks be identified already? How do we expect the need of computational power to scale?
- Are there intrinsically different computational issues at NNLO compared to NLO?

Computational Tools: Parallelization

Multi-Threading

- Communication across processor cores (CPU/GPU)
- Shared memory between all threads - implicit communication
- Not scalable $\rightarrow$ reduction of processing time by at most # of cores

Message Passing Interface (MPI)

- Communication across processor cores or computing nodes
- No shared memory between threads - communicate explicitly
- Scalable $\rightarrow$ “arbitrary” reduction of processing time
- On clusters limited by network bandwidth
- Insufficient handling of node failure in current MPI libs

Local Computing Clusters

- Batch processing
- Large-scale parallel computing

The Open Science Grid (OSG)

- Multi-disciplinary partnership to federate local, regional, community and national cyberinfrastructures to meet needs of research and academic communities (http://www.opensciencegrid.org)
- Capable of absorbing peak loads which are difficult or impossible to handle by single sites
- Details of resource allocation hidden from user
- Few overhead compared to local computing clusters
- MPI capable, but no inter-node communication yet

Draft charge

Perform survey on resource usage and needs of current projects

- CPU / GPU time
- RAM / Cache
- Disk storage
- Network
- Software

Identify limitations, e.g.

- Insufficient computing / storage resources
- Lack of or deficiencies in software frameworks

Draft Charge

Identify infrastructure needed to perform calculations required for precision SM and BSM measurements

- Role of parallel computing models
- Benefits of inter-thread vs. inter-process communication
- Prospects for using GPUs on a large scale
- Gains from consolidating resources
- Prospects for access to parallel computing on the Grid
- Role of the network
- Storage needs and allocation
- Dependences on local resources

Draft Charge

Identify software needed to perform calculations required for precision SM and BSM measurements

- Role of the software environment
- Importance of proprietary software
- Limitations in programming languages
- Limitations in libraries for parallel computing
- New common libraries