Computing Frontier: CpF T4 Perturbative QCD

Stefan Höche and Laura Reina

Snowmass Energy Frontier Workshop

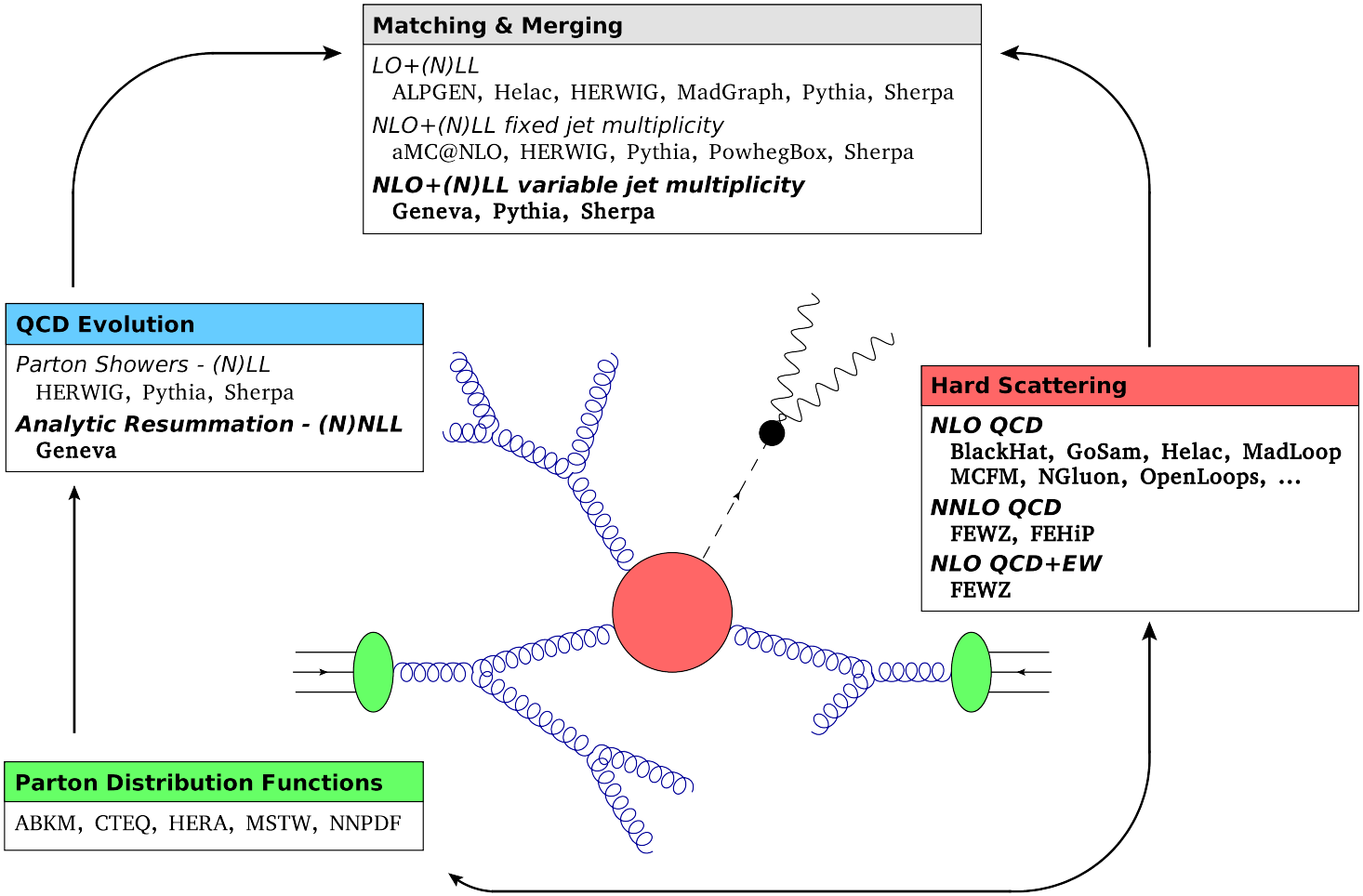
BNL, April 3-6 2013

Theoretical HEP and computing

Main areas in which HEP theory relies on HTC/HPC

- Fixed-order (N)NLO calculations
  - ▷ Test new theoretical concepts in high-multiplicity processes
  - ▷ Generate parton-level event samples for theory & experiment
- QCD/EW resummation
  - ▷ Validate algorithms for high fixed-order & logarithmic accuracy
  - ▷ Compute predictions/uncertainty estimates for many observables
- MC event generators
  - ▷ Validate algorithms to simulate high-multiplicity events at NLO
  - ▷ Tuning, release preparation, tests of applicability
- Interpretation of LHC data in terms of models for new physics
  - ▷ Generate large events samples, test variety of BSM scenarios
  - ▷ Archive data and/or plots for usage by experiments
- Lattice QCD

Broad impact of Perturbative QCD on collider physics



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

We expect to receive input and interact with several 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

and assess present and future needs, learn from research in computer science and technology, set the agenda

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 not 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

Computational Tools: Distributed Computing

Local Computing Clusters

- Batch processing
- Large-scale parallel computing (SC centers, e.g. NERSC)

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`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

HPC for Higher-Order Perturbative QCD

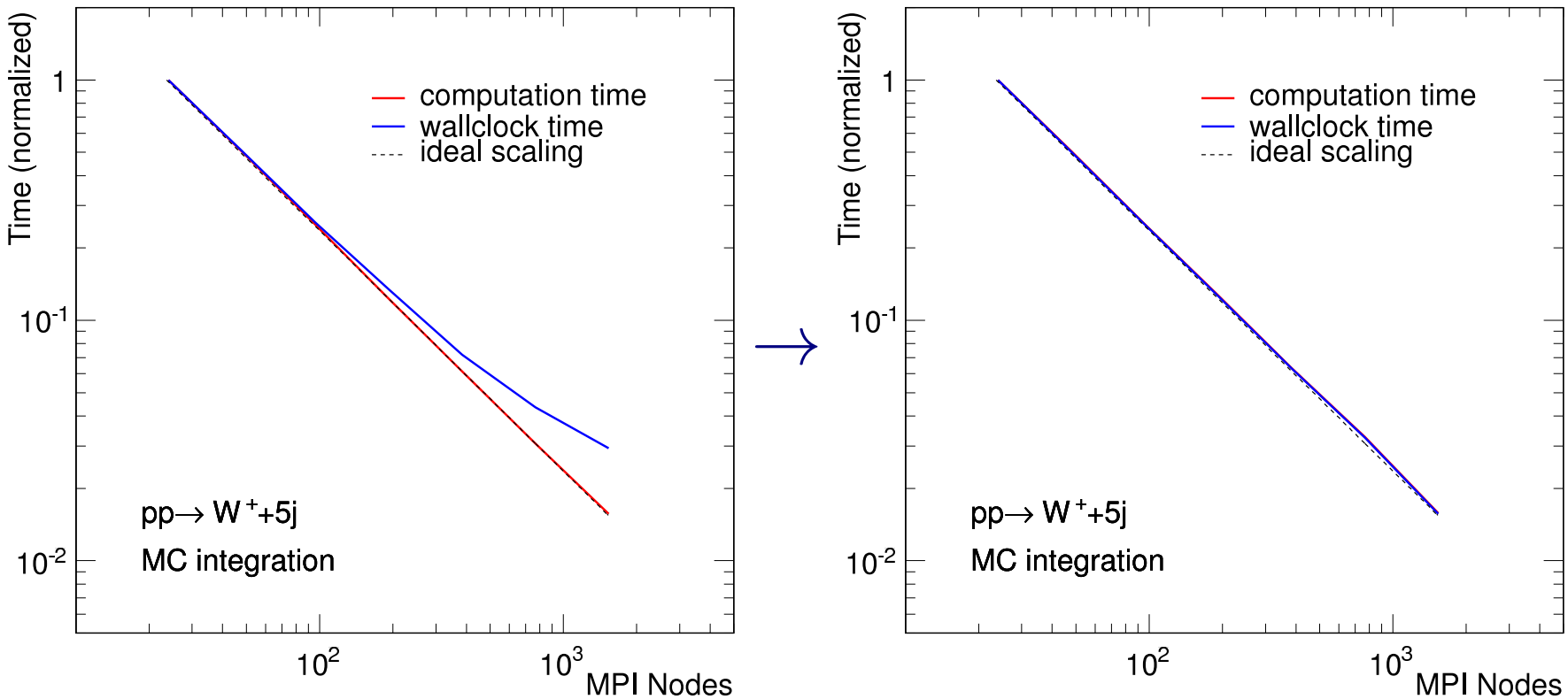
Current status:

- No conceptual problems in breaking down NLO calculations into standard modular steps, automate them, and make them public
- Dedicated effort started in the context of Snowmass 2013 to address this project and have a realistic estimate of necessary resources
- First exploratory studies at NERSC show enormous potential to utilize codes which are readily available

Expected benefits of HPC:

- Various existing calculations/software can be made public within a common well-tested framework → see tutorial, S. Höche, Thursday 12:30 pm, Room C
- Allows experimental studies involving extensive multiple runs without depending on the computer- and manpower available to code authors
- New cutting-edge calculations can fully exploit parallel architectures and the availability of large resources

Experience gained at NERSC so far



- ▷ Initially large MPI overhead in pQCD MC when run on many nodes
- ▷ Discussion with S. Habib (ANL) → removed bottlenecks
- ▷ Much to gain from HPC experience in Cosmology / Lattice QCD
- ▷ Existing NLO codes succesfully interfaced with Sherpa

→ see tutorial, S. Höche, Thursday 12:30 pm, Room C

Summary

- HPC at NERSC tested for cutting edge pQCD / MC
  $\sim$ scaling shown, ready for production
- Small-scale MPI runs also tested on OSG
  Thanks to G. Garzoglio, T. Levshina, M. Slyz!
- White paper on HPC/HTC for HEP circulated
  The computing needs of theoretical high energy physics at the Energy Frontier
  [Boughezal, Dixon, Hewett, Hoeche, Petriello, Reina, Rizzo]
- Collecting responses from community

Ongoing Activities

- HPC tutorial at East Coast meeting of the Energy Frontier (S. Höche, Thursday 12:30 pm, Room C)
- Signing up people with interest in HPC for Snowmass study
- First follow-up planned after Loopfest (May 16, 2013)
- Build knowledge base "HPC for THEP"