MC Generators for Parallel & Distributed Computing

Stefan Höche

SLAC National Accelerator Laboratory

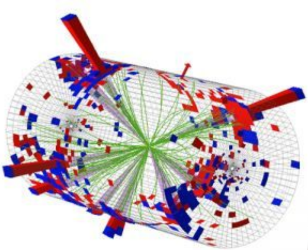
Future Trends in NP Computing

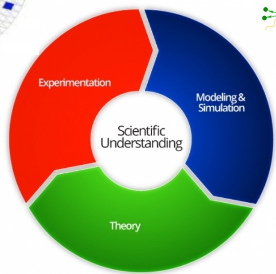
JLab, March 16, 2016

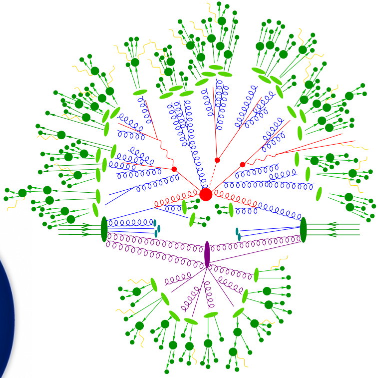
U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

MC Generators – Linking theory to experiment







$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i \bar{\Psi} \not{D} \Psi + h.c.$$

Toolkit inventory

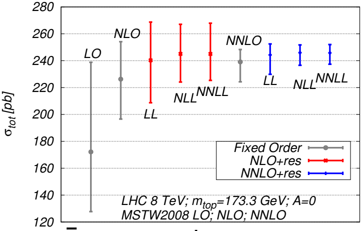
SLAC

Current state of development

- Parton shower Monte Carlo (Herwig,Pythia,Sherpa,...)
- Automated NLO calculations (BlackHat,GoSam,Helac,MadLoop,MadGolem,NJet,OpenLoops,...)
- Matching to parton shower (aMC@NLO,Herwig,POWHEG Box,Sherpa,...)
- Merging of NLO calculations (aMC@NLO,Helac,Pythia,Sherpa,...)

Cutting edge technology & future directions

- Inclusive NNNLO (gg→H)
- Differential NNLO (V/H(+jet),$\gamma\gamma$,VV,...)
- NNLO+N^{x}LL resummation (gg→H,$t\bar{t}$,...)
- NNLO+parton shower (W,Z,gg→H)



$t\bar{t}$ cross section arXiv:1305.3892

Parallel and distributed computing – Why?

SLAC

- Final states in collider experiments typically very complex
  interesting signals small and in tricky corners of phase space
  $\rightarrow$ Need many events to simulate backgrounds accurately $\mathcal{O}(10M)$
- Precise theoretical predictions needed to extract physics parameters
  $\rightarrow$ Simulation of one event can take anywhere from 10ms to 100s
- Theory uncertainty estimates mandatory to control systematics
  $\rightarrow$ Multiplies computational effort by factor of $\mathcal{O}(1-10)$

A single high-precision MC prediction costs $\mathcal{O}(250k)$ CPU hours

Parallel and distributed computing – How?

MC simulations / NLO pQCD calculations can be split into

- Integration step
  - Determine total cross section and maximum for MC simulation
  - Use adaptive MC integrators to reduce variance
  - Store results in form of weight factors / grids
- Event generation step
  - Use weight factors / grids to increase efficiency
  - Produce full events rather than cross sections only (Parton shower, Hadronization, Hadron decays, ...)

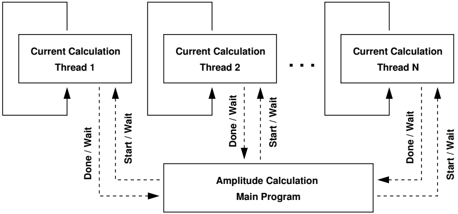
Natural separation into HPC and HTC domain

Optimizing resource usage could mean

- Integration performed in parallel
- Event generation distributed

Current parallelization strategy

- Amplitude calculation in Dyson-Schwinger method and phase-space recursion easily thread-parallelized

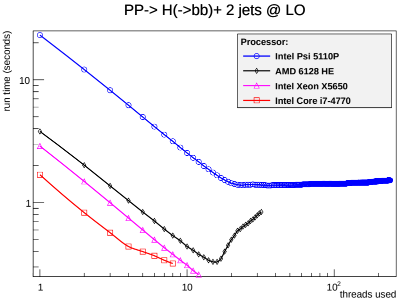


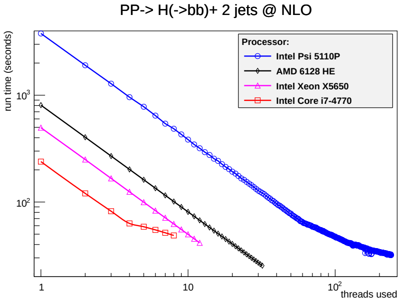
- Low efficiency as number of threads limited by propagator structure
- Proof-of-concept in event generator Comix arXiv:0808.3674

$gg \to ng$	Cross section [pb]				
n $\sqrt{s}$ [GeV]	8 1500	9 2000	10 2500	11 3500	12 5000
Comix	0.755(3)	0.305(2)	0.101(7)	0.057(5)	0.026(1)

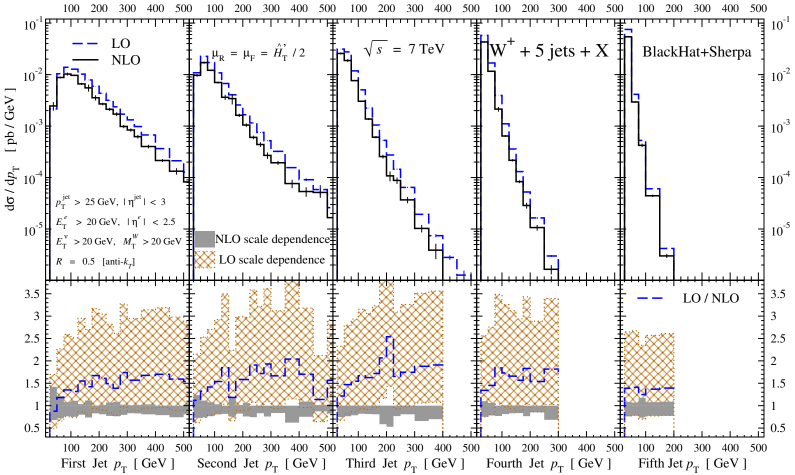
Current parallelization strategy

- MC integration / optimization of adaptive integrator easily MPI parallelizable, often thread-parallelizable
- MPI used in MC event generator Sherpa arXiv:1304.1253
- MPI and OpenMP used in MCFM arXiv:1503.06182



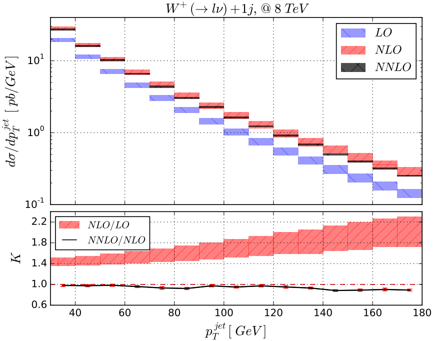


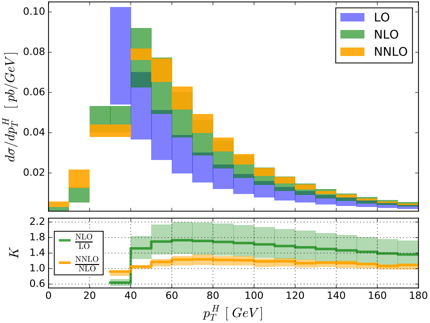
Example applications - Using small clusters



- W+5jet NLO calculation from BlackHat+Sherpa arXiv:1304.1253
- Combination of generalized unitarity approach (virtual corrections) and Dyson-Schwinger method (Born and real-emission corrections)

Example applications - Using NERSC / LCFs





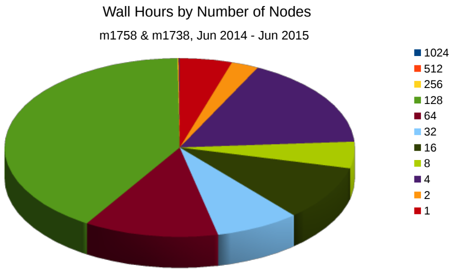
- W/H+1jet NNLO calculation using MCFM arXiv:1504.02131, arXiv:1505.03893
- Jettiness subtraction technique to regularize IR divergences at NNLO
  Soft function computed numerically arXiv:1504.02540

HPC usage projections from Snowmass 2013

Type of calculation	CPU hours per project	projects per year
NLO parton level	300,000	10-12
Matrix Element Method	200,000	3-5
NNLO parton level	250,000	5-6
Precision event generation	200,000	3
Exclusive jet cross sections	300,000	1-2
Parton Distributions	50,000	5-6
MSSM phenomenology	500,000	10
BSM constraints	150,000	2
Model building	100,000	1-2

- Projected total of $\geq$6M CPUh for pQCD and $\geq$5.45M CPUh for Pheno
- Prone to rapid changes depending on theory and technology developments

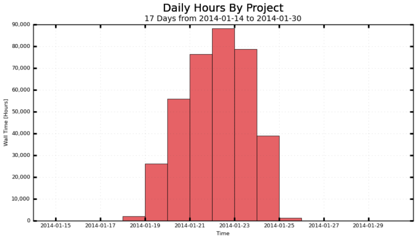
Actual HPC usage by 2015



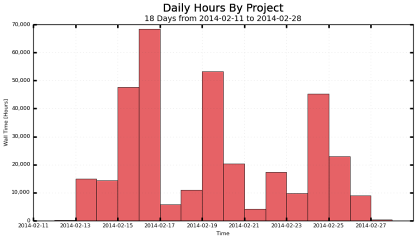
- NERSC usage June 2014 - June 15 $\sim$6.07M CPUh (pQCD only)
- Still at small-scale, but potential for growth ($\rightarrow$ NNLO applications)

MC event generators on the Open Science Grid

- H+≤2jets at NLO
  ∼481k CPU-h (6d real time)
  opportunistic usage



- $t\bar{t}+$ ≤2jets at NLO
  ∼345k CPU-h (14d real time)
  opportunistic usage



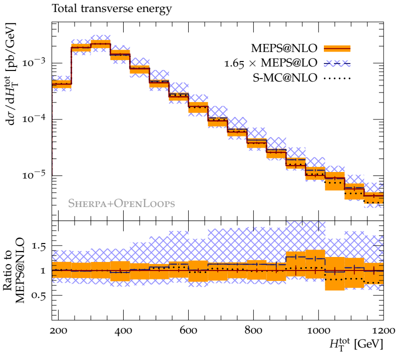
SLAC

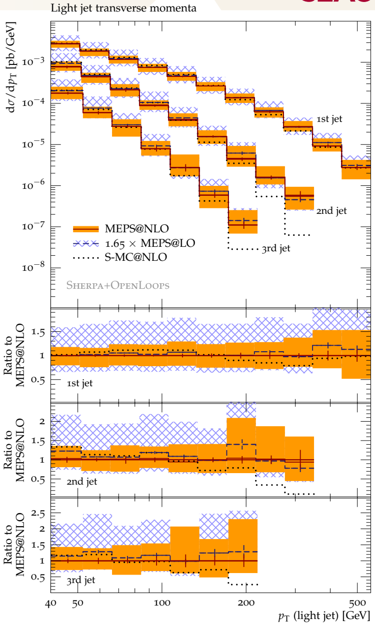
User Experience

- Tried pre-staging but code small enough to be sent with job input
  → allows rapid turnaround in development!
- HTPC working (using custom-compiled MPICH2, sent with input)
  → allows larger simulations and mid-scale integration jobs
- Some simulations currently memory-bound
  → not enough (known) high-mem nodes for high-multiplicity NLO

Example applications

- First matched/merged sim for $t\bar{t}$+2j including $t\bar{t}$+0,1,2j@NLO arXiv:1402.6293
- Largely reduced theory uncertainty for both for measurement (p_T, N_{jet}) and BSM search (H_T) observables



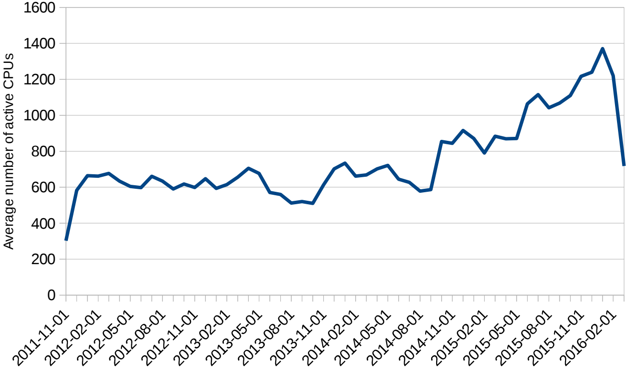


MC generators and volunteer computing

SLAC

Test4Theory / MCplots project arXiv:1306.3436

- 2.68 trillion events simulated on LHC@Home platform
- Useful for time-non-critical validation & tuning



Summary / Outlook

So far

- MC simulations catching on to mid-scale parallel computing
- Batch & distributed computing still main production mode
- Ideally both use cases combined to exploit new precision MC

To do

- Make output of MC simulations useful beyond a single analysis
  $\rightarrow$ storage and management of results ($\sim$ 2-20TB per event sample)
- Better exploitation of HTPC: 16 core parallelism often sufficient