OASIS:
Optimal Analysis-Specific Importance Sampling for event generation

Prasanth Shyamsundar

University of Florida

New Perspectives 2020, Fermilab

July 20, 2020

Who I am, professionally

- Graduate student at the University of Florida.

- Defending my dissertation tomorrow.

- Joining the Fermilab Quantum Institute as a postdoc at the end of September.

OASIS motivation

- ▶ Simulations in HEP are computationally expensive.
  - — Detector simulation is the most resource intensive part of the pipeline.
  - — Projected HL-LHC computational requirements may not be met. "Billion dollar problem"
  - — Need to speed up the simulation pipeline.
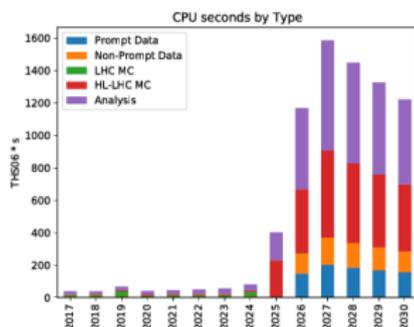
CMS
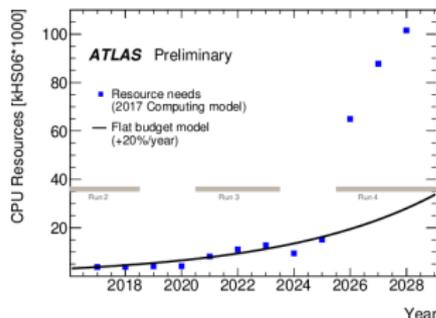
ATLAS

J. Albrecht *et al.* [HEP Software Foundation], "A Roadmap for HEP Software and Computing R&D for the 2020s," Comput. Softw. Big Sci. **3**, no.1, 7 (2019) [arXiv:1712.06982 [physics.comp-ph]].

**OASIS reduces the number of simulated events needed by experiments.**

- ► Why do we need simulated events in HEP?
  - — To estimate the distribution of our analysis variables under different theory models.

- ► Why do we need **more** simulated events in HEP?
  - — To estimate the distribution of our analysis variables **better**.
  - — Reduce the error bars on the estimates.

Do we need to reduce error bars everywhere*?
  - — Not really... It is better to focus on the sensitive regions of the phase-space.

*maintaining the same ratios between error bars

An example: Top mass measurement
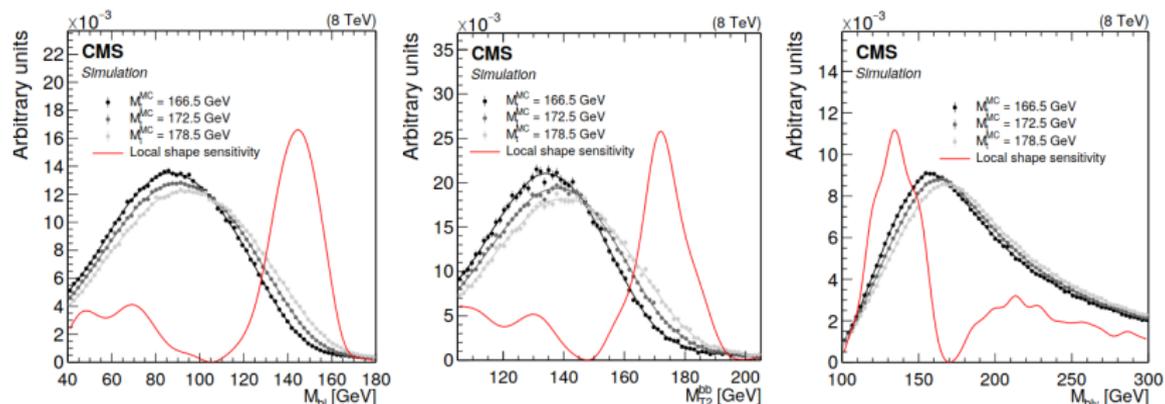
A. M. Sirunyan *et al.* [CMS], "Measurement of the top quark mass in the dileptonic $t\bar{t}$ decay channel using the mass observables $M_{b\ell}$, M_{T2}, and $M_{b\ell\nu}$ in pp collisions at $\sqrt{s} = 8$ TeV," Phys. Rev. D **96**, no.3, 032002 (2017) [arXiv:1704.06142 [hep-ex]].

▶ Different regions of the phase-space are sensitive to the value of a parameter (or presence of a signal) to different extents.

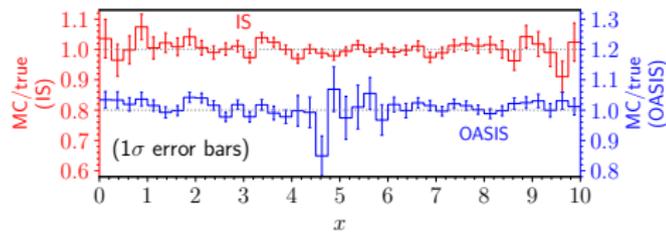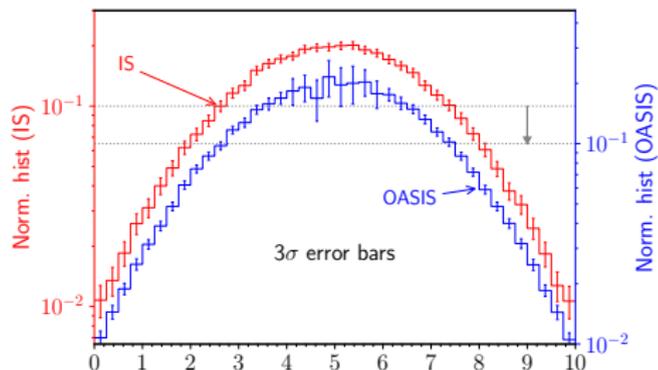▶ Different regions of the phase-space have different utilities to the analysis.

OASIS: Optimal Analysis Specific Importance Sampling

IDEA: Choose the event sampling distribution based on the utility to the analysis at hand. (Remember to weight them appropriately)

local sensitivity
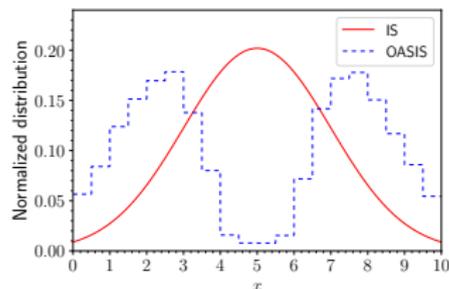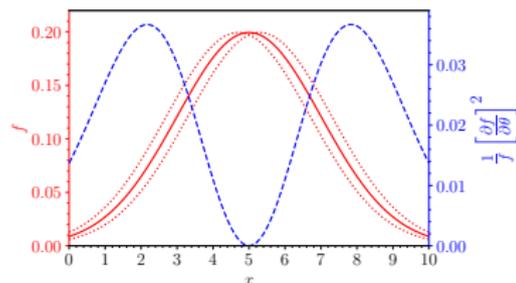
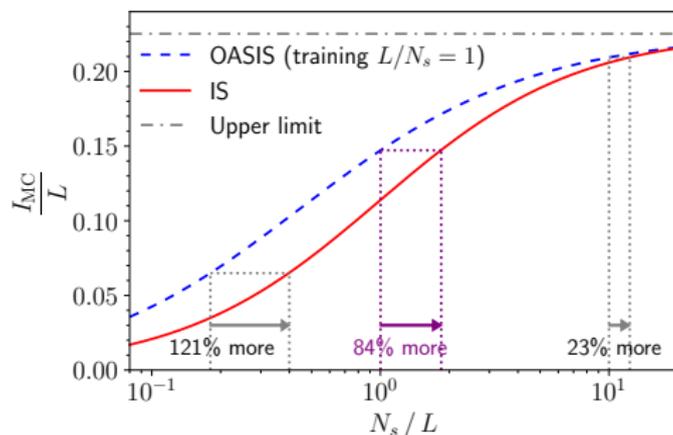sampling distributions

weighted histograms

Why implement OASIS?

I_{MC}/L captures the sensitivity of the analysis to the parameter value.

OASIS can potentially (will probably) save hundreds of millions of dollars in computing budget.

The two aspects of OASIS

- ▶ The math and stats of figuring out the optimal sampling distribution in the analysis-level variables.
- ▶ The technology for achieving this optimal sampling distribution — using parton level importance sampling.

The two aspects of OASIS

- ▶ The math and stats of figuring out the optimal sampling distribution in the analysis-level variables.
- ▶ The technology for achieving this optimal sampling distribution — using parton level importance sampling.

Check out:

K. T. Matchev and P. Shyamsundar, *"OASIS: Optimal Analysis-Specific Importance Sampling for event generation,"* [arXiv:2006.16972 [hep-ph]].

Tune into:

LPC Physics Forum
Thursday, Jul 30 at 1:30pm CT

OASIS: "Better" simulated events to allow for fewer simulated events — Prasanth Shyamsundar

https://indico.cern.ch/event/939512/

The two aspects of OASIS

- ▶ The math and stats of figuring out the optimal sampling distribution in the analysis-level variables.
- ▶ The technology for achieving this optimal sampling distribution — using parton level importance sampling.

Check out:

K. T. Matchev and P. Shyamsundar, *"OASIS: Optimal Analysis-Specific Importance Sampling for event generation,"* [arXiv:2006.16972 [hep-ph]].

Tune into:

LPC Physics Forum
Thursday, Jul 30 at 1:30pm CT

OASIS: "Better" simulated events to allow for fewer simulated events — Prasanth Shyamsundar

https://indico.cern.ch/event/939512/

Thank you!