## Track 3 Summary

Computations in Theoretical Physics: Techniques and Methods

## Anke Biekötter - Leonardo Cosmai Joshua Davies - Latifa Elouadrhiri

ACAT 2022 -
24-28 October 2022
Villa Romanazzi Carducci, Bari, Italy


## ACAT 2022



## ACAT 2022



Thank you for your contributions!

## Track 3 Highlights

- Monte Carlo generation
- Precision frontier
- Beyond Standard Model physics
- Towards Quantum Computing


## This is a biased selection

## Simon Badger

from theory to experiment


## Speeding up Monte Carlo event generators



- Performance analysis
- Pilot runs (what do we need when?)
- New architectures - GPUs, vector CPUs
- Portability (Kokkos, Alpaka, ...)
- Physics ideas and analytic results


## Breakdown of CPU budget in V+jets

Christian Gütschow



Neural Importance Sampling - Results

## Enrico <br> Bothmann






Smaller impact for more complicated (multi-channel) processes, similar in [Gao et al., Phys. Rev. D 101 (2020) no.7, 076002] GPU evaluation of MEs desirable for efficient training of. talks by M. Knobbe, R. Wang and A. Valassi

- Alternative to ML-assisted phase space sampling: directly learn target distribution using autoregressive flows, GANs, VAEs

if no suriectivity guarantee $\rightarrow$ might miss tails of distributions and get small bias in overall integration result


## Enrico <br> Bothmann


$p p \rightarrow e^{+} e^{-}+0,1,2 j @ \mathrm{NLO}+3,4,5 j @ \mathrm{LO}$

Max Knobbe

Smaller impact for more complicated (multi-channel) processes, similar in [Gao et al., Phys. Rev. D 101 (2020) no.7, 076002] GPU evaluation of MEs desirable for efficient training of. talks by M. Knobbe, R. Wang and A. Valassi

- Alternative to ML-assisted phase space sampling: directly learn target distribution using autoregressive flows, GANs, VAEs

if no surjectivity guarantee $\rightarrow$ might miss tails of distributions and get small bias in overall integration result
- Benchmark performance for gluon-only process
- Relevant test, since as-many-gluon-as-possible amplitudes make up largest portion of computing time for jet-processes
- Compare different color treatments:
color-dressing/summing/sampling
- Color-sampled algorithms scale similar to color-summed approaches
Color-summing scales worse than color-dressing, bu faster up to roughly $5-6$ outgoing jets
Caveat: Color-sampling comes with penalty factor from slower convergence
$\Rightarrow$ Algorithmic choice: Sum colors



## ME on a GPU

## Andrea Valassi

## MadEvent/CUDA for gg $\rightarrow$ tt̄gg (improved at ACAT2022)

| CUDA grid size |  | ICHEP2022 | madevent |  | standalone |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 8192 |  | 524288 |
| $g g \rightarrow t \overline{t g}$ | $\begin{gathered} \text { MEs } \\ \text { precision } \end{gathered}$ |  | $\begin{gathered} t_{\mathrm{TOT}}=t_{\mathrm{Mad}}+t_{\mathrm{MEs}} \\ {[\mathrm{sec}]} \end{gathered}$ | $N_{\text {events }} / t_{\text {TOT }}$ [events/sec] | $N_{\text {events }} / t_{\text {MEs }}$ [MEs/sec] |  |  |
| Fortran | double | $58.3=5.2+53.1$ | 1.55 E 3 ( $=1.0$ ) | 1.70 E 3 ( $=1.0$ ) | - | - |
| CUDA | double | $6.1=5.7+0.36$ | 1.49E4 (x9.6) | 2.54 E 5 ( x149) | 2.51 E | 4.20E5 (x247) |
| CUDA | float | $5.7=5.4+0.24$ | 1.59 E 4 (x10.3) | 3.82E5 (x224) | 3.98E5 | 8.75E5 (x515) |

## Enrico

Bothmann


Smaller impact for more complicated (multi-channel) processes, similar in [Gao et al., Phys. Rev. D 101 (2020) no.7, 076002 GPU evaluation of MEs desirable for efficient training cf. takks by M. Knobbe, R. Wang and A. Valassi Alternative to ML-assisted phase space sampling: directly learn target distribution using autoregressive flows, GANs, VAEs

if no surjectivity guarantee $\rightarrow$ might miss tails of distributions and get small bias in overall integration result

## Max Knobbe

- Benchmark performance for gluon-only process
- Relevant test, since as-many-gluon-as-possible amplitudes make up largest portion of computing time fo jet-processes
- Compare different color treatments:
color-dressing/summing/sampling
- Color-sampled algorithms scale similar to color-summed approaches
Color-summing scales worse than color-dressing, bu
faster up to roughly $5-6$ outgoing jets
Caveat: Color-sampling comes with penalty factor from slower convergence
$\Rightarrow$ Algorithmic choice: Sum colors


$$
p p \rightarrow e^{+} e^{-}+0,1,2 j @ \mathrm{NLO}+3,4,5 j @ \mathrm{LO}
$$

## ME on a GPU

## Andrea Valassi

## MadEvent/CUDA for gg $\rightarrow$ tt̄gg (improved at ACAT2022)

| CUDA grid size |  | ICHEP2022 | madevent |  | standalone |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 8192 |  | 524288 |
| $g g \rightarrow t \overline{t g} g$ | $\begin{gathered} \text { MEs } \\ \text { precision } \end{gathered}$ |  | $\begin{gathered} t_{\text {TOT }}=t_{\text {Mad }}+t_{\text {MEs }} \\ {[\text { [sec] }} \end{gathered}$ | $N_{\text {events }} / t_{\text {TOT }}$ [events/sec] | $N_{\text {events }} / t_{\text {MEs }}$ [MEs/sec] |  |  |
| Fortran | double | $58.3=5.2+53.1$ | 1.55 E 3 ( $=1.0$ ) | 1.70E3 ( $=1.0$ ) | - | - |
| CUDA | double | $6.1=5.7+0.36$ | 1.49 E 4 (x9.6) | $2.54 \mathrm{E5}$ (x149) | 2.51 E5 | 4.20E5 (x247) |
| CUDA | float | $5.7=5.4+0.24$ | 1.59E4 (x10.3) | 3.82E5 (x224) | 3.98E5 | 8.75 E5 (x515) |

Reduced the overhead from scalar Fortran MadEvent overhead from $10 \%$ to $5 \%$ of initial Fortran (improved handling of MLM merging) Maximum allowed overall speedup from Amdah's law is now increased from $\times 10$ to $\times 20-$ which we do achieve

Enrico

| CUDA grid size |  | ACAT2022 | $\begin{array}{\|c\|} \hline \text { madevent } \\ \hline 8192 \end{array}$ |  | standalone |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | 524288 |  |
| $g g \rightarrow t \overline{t g} g$ | $\begin{gathered} \text { MEs } \\ \text { precision } \end{gathered}$ |  | $\begin{gathered} t_{\mathrm{TOT}}=t_{\mathrm{Mad}}+t_{\mathrm{MEs}} \\ {[\mathrm{sec}]} \end{gathered}$ | $N_{\text {events }} / t_{\text {Tot }}$ [events/sec] | $N_{\text {events }} / t_{\text {MEs }}$ [MEs/sec] |  |  |
| Fortran | double | $55.4=2.4+53.0$ | 1.63 E 3 ( $=1.0$ ) | 1.70E3 ( $=1.0$ ) | - | - |  | $p p \rightarrow e^{+} e^{-}+0,1,2 j @ \mathrm{NLO}+3,4,5 \mathrm{j} \mathrm{LO}$ |
| CUDA | double | $2.9=2.6+0.35$ | 3.06E4 (x18.8) | 2.60 E 5 ( $\times 152$ ) | 2.62E5 | 4.21E5 (x) |  |  |



- Benchmark performance for gluon-only process
- Relevant test, since as-many-gluon-as-possible amplitudes make up largest portion of computing time for jet-processes
- Compare different color treatments:
color-dressing/summing/sampling
- Color-sampled algorithms scale similar to color-summed
approaches approaches
Color-summing scales worse than color-dressing, bu
faster up to roughly $5-6$ outgoing jets
Caveat: Color-sampling comes with penalty factor from slower convergence
$\Rightarrow$ Algorithmic choice: Sum colors

remember: aim for $w=f / g \approx 1$, i.e. peaked distribution of $w$


Smaller impact for more complicated (multi-channel) processes, similar in [Gao et al., Phys. Rev. D 101 (2020) no.7, 076002] GPU evaluation of MEs desirable for efficient training cf. talks by M. Knobbe, R. Wang and A. Valassi
Alternative to ML-assisted phase space sampling: directly learn target distribution using autoregressive flows, GANs, VAEs

if no surjectivity guarantee $\rightarrow$ might miss tails of distributions and get small bias in overall integration resu

## Taylor Childers

## Performance of Kokkos

- So, does Kokkos provide equivalent performance?
- Plot shows early versions of BlockGen ${ }_{20}^{10}$ calculating the process: $\mathrm{gg} \rightarrow$ njets
- Time per Event on $y$-axis, number of outgoing partons on $x$-axis
- Compare CPU with C++, GPU with CUDA, and GPU with Kokkos
- Can see the CUDA is $100 x$ faster than the CPU for this example
- 
- Kokkos is slightly less performant than CUDA at low multiplicity (low computational complexity), but reaches comparable performance as multiplicity increases

Simon

## Precision Frontier

## Johann Usovitsch

Precision test of the Standard Model Future prospects
Overview of future experiments as of 2022

## Experiment uncertainty Theory uncertainty

|  | Experiment |  |  |  |
| :--- | :--- | :--- | :--- | :--- |
|  | ILC | CEPC | FCC-ee | Current |
|  | $3-4$ | 3 | 10.3 | 4 |
| $M_{W}[\mathrm{MeV}]$ |  | Theory uncertainty |  |  |
| $\sin ^{2} \theta_{\text {eff }}^{1}\left[10^{-5}\right]$ | 1 | 2.3 | $? 0.6$ | 4.5 |
| $\Gamma_{Z}[\mathrm{MeV}]$ | 0.8 | 0.5 | 0.10 .025 | 0.4 |
| $R_{f}\left[10^{-5}\right]$ | 14 | 17 | $\not 1$ | 15 |

- Recent update from [Alain Blondel, Patrick Janot, Eur.Phys.J.Plus 137 (2022) 1]
- To match the precision of the experiment we compute 3-loop and 4-loop Standard Model predictions

Badger



IR frontier AIID


N3LO splititng functions, analytic

- Numerical methods
- Avoiding algebraic complexity
- Physics informed
- Exploiting known structures

Simon
Badger


- Numerical methods
- Avoiding algebraic complexity
- Physics informed
- Exploiting known structures

Simon
Badger


- Numerical methods
- Avoiding algebraic complexity
- Physics informed
- Exploiting known structures


## Numerical approaches



Uncertainties

Result

- 100 nodes
- 4 hidden layers
- $4 \mathrm{M} \times 800=3.2 \mathrm{~B}$ PS points


Machine learning the primitive

## Numerical approaches - physics informed

$$
K_{n+1}=C_{0}+\sum_{\{i j k\}} C_{i j k} \frac{X_{i j k}^{1}}{X_{i j k}^{0}}
$$

Comparison to naive model



$15 / 22$

Learning K factors/matrix elements

- coefficients of antenna functions

Results: effective gain factors for LHC multi-jet processes


## Analytic approaches

Finite fields

## Local unitarity

$N_{f}$ part @ ${ }^{3} \mathrm{LO} e^{+} e^{-} \rightarrow j j$

$N_{f}^{2}$ part $@ N^{3} \mathrm{LO} \quad e^{+} e^{-} \rightarrow j j$

| Channel | $\mathrm{f} 64 / \mathrm{f} 64$ |  | Evaluation strategy |  |
| :--- | ---: | ---: | ---: | ---: |
|  | Time (s) | $f(\%)$ | Time (s) | $f(\%)$ |
| $g g \rightarrow g g g$ | 1.39 | 69 | 1.89 | 77 |
| $g g \rightarrow \bar{q} q g$ | 1.35 | 91 | 1.37 | 91 |
| $q g \rightarrow q g g$ | 1.34 | 92 | 1.57 | 93 |
| $q \bar{q} \rightarrow g g g$ | 1.34 | 93 | 1.38 | 93 |
| $\bar{q} Q \rightarrow Q \bar{q} g$ | 1.14 | 99 | 1.16 | 99 |
| $\bar{q} \bar{Q} \rightarrow \bar{Q} g$ | 1.36 | 99 | 1.39 | 99 |
| $\bar{q} g \rightarrow \bar{q} Q \bar{Q}$ | 1.36 | 99 | 1.39 | 99 |
| $\bar{q} q \rightarrow Q \bar{Q} g$ | 1.14 | 99 | 1.14 | 99 |
| $\bar{q} g \rightarrow \bar{q} q \bar{q}$ | 1.84 | 99 | 1.90 | 99 |
| $\bar{q} \bar{q} \rightarrow \bar{q} \bar{q} g$ | 1.82 | 99 | 1.94 | 99 |
| $\bar{q} q \rightarrow q \bar{q} g$ | 1.71 | 99 | 1.77 | 99 |
| $g g \rightarrow \gamma \gamma g *$ | 9 | 99 | 26 | 99 |

Ryan Moodie (Turin) Two-loop five-point amplitudes in massless QCD with finite fields


## Lorentz Invariant Phase Space



## Charles R. Harris et al. "Array programming with NumPy". In: Nature 585 (2020), pp. 357-362. Dol: <br> 10.1038/s41586-020-2649-2

${ }^{2}$ Fredrik Johansson et al. mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18),
$\qquad$ ${ }^{3}$ Aaron N
${ }^{4}$ Wolfram Decker et al. SIngular 4-3-0 - A computer algebra system for polynomial computations http://www.singular.uni-kl.de. Giuseppe De Laurentis

## Non-perturbative physics

Simulation at multiple $\lambda$ values
For a lattice action :
$S\left(\phi, m_{\text {fixed }}, \lambda\right) \longrightarrow p\left(\phi \mid \lambda_{i}\right)$

## Ankur Singha



New Theory Prediction Pipeline
Produce FastKernel (FK) tables! Felix Hekhorn


Alberto Martini

Proton-proton @ 0.9-13 TeV, Predictions Gabor Biro


- So far: everything at $\sqrt{s}=7 \mathrm{TeV} \rightarrow$ the ONLY energy, where the models were trained
- Good agreement for all observable quantities as predictions for other LHC energies
- Multiplicity scaling?



## Preparation of detector configuration



## Beyond Standard Model

- Looking for new physics
- LHC and beyond
- Model independence
- Recasting
- Symmetries


## Sven Krippendorf



No direct optimisation available: embedding in deep layer

## We need: group input with the same meaning together



## The LHC g.o.f. challenge

By analysing the LHC data, we would like to find evidence of failure of the SM theory, suggesting need of BSM.

This is a tremendously hard gof problem!
BSM is tiny departure from SM, or large in tiny prob. region
Affecting few (unknown) observables over $\infty$ many we can measure
Model-dependent $\mathrm{H}_{1}$
BSM searches

- Optimise sensitivity to one
specific BSM model
- Fail to discover other models.
What if the right theoretical
model is not yet formulated?


## Model-independent

 searches- Could reveal truly unexpected new physical laws.
- No hopes to find Optimal strategy. For a Good strategy, we need a
good choice of $\mathrm{H}_{\mathrm{w}}$.

vity to one What if the right theoretical model is not yet formulated?


## Beyond Standard Model

## The Normalised AutoEncoder <br> <br> Alo

 <br> <br> Alo}Theo Heimel

No complexity bias!
More rubust and reliable anomaly detection

Visualisation of the latent space Looks like a mess, but very useful for interpreting the results and diagnosing problems with the training!


Matrix Element Method

- Process with theory parameter $\alpha$, hard-scattering momenta $x_{\text {hard }}$
- Likelihood at hard-scattering level given by differential cross section

$$
p\left(x_{\text {hard }} \mid \alpha\right)=\frac{1}{\sigma(\alpha)} \frac{\mathrm{d} \sigma(\alpha)}{\mathrm{d} x_{\text {hard }}}
$$

- Neyman-Pearson lemma $\Longrightarrow$ optimal use of information
- Differential cross section only known analytically at hard-scattering level



## Barry Dillon

Dilion - Universität Heidelberg - Anomaly searches for new physics at the LHC
Anomaly detection

## Beyond Standard Model

\section*{The Normalised AutoEncoder

Theo Heimel

## orrenson, Krämer

## orrenson, Krämer

No complexity bias!
More rubust and reliable anomaly detection

Visualisation of the latent space
Looks like a mess, but very useful for interpreting the results and diagnosing problems with the training!

Matrix Element Method

- Process with theory parameter $\alpha$, hard-scattering momenta $x_{\text {hard }}$
- Likelihood at hard-scattering level given by differential cross section

$$
p\left(x_{\text {hard }} \mid \alpha\right)=\frac{1}{\sigma(\alpha)} \frac{\mathrm{d} \sigma(\alpha)}{\mathrm{d} x_{\text {hard }}}
$$

- Neyman-Pearson lemma $\Longrightarrow$ optimal use of information



## Beyond Standard Model



Jamie Yellen

Dark Matter MC event generation


Henri Sieber

## Useful Near-Term Quantum Algorithms

- Quantum Machine Learning (QML): Quantum advantages proven for

Learning complex patterns w/ quąntum feature maps (arXiv:2010.02174)
Exponential gain in predicting certain worst-case error (arXiv:2101.02464)
Quantum correlations used in generative modeling (arXiv:2101.08354)

- Quantum Chemistry and Materials Studies

Variational quantum eigensolvers (VQE) for energy estimation
Quantum simulation of dynamics of excitation
Study of quantum many-body phenomena

- Optimization Problems: Quantum Approximate Optimization Algorithm

IONQ

## Quantum Machine Learning

## Top Tagging through MPS



Jack
Araz

A hard decision making process
Our best classical (not quantum!) classifier with diameter, grading, histologic type, multifocality, in situ component, PgR :

| 70.8 (70.3-71.1) | $69.8(69.3-70.2)$ | $74.8(72.9-75.1)$ | $61.0(60.3-61.7)$ |
| :--- | :--- | :--- | :--- | reporting the $1^{\text {st_}} 3^{\text {rd }}$ interquartile range after 10 ten-fold cross-validations.



Domenico
Pomarico

## Quantum Circuit Born machine (QCBM)

. Sample from a variational pure state $|\psi(\theta)\rangle$ by projective measurement with probability given by the Born rule: $\boldsymbol{p}_{\boldsymbol{\theta}}(\boldsymbol{x})=|\langle\boldsymbol{x} \mid \boldsymbol{\psi}(\boldsymbol{\theta})\rangle|^{2}$


Michele
2. Training (Hybrid loop):
KL divergence
Adversarial
Delgado and Hamilton, arXiv:2203.03578 (2022) Zoufal, et al., npj Quantum Inf 5 , 103 (2019). In the phase space. Kyrienko, et al., arXiv: 2202.08253 (2022).

3. Why the Maximum Mean Discrepancy MMD ?

```
Resource efficient for NISQ devices
```


## Quantum chemistry and fluids

Variational quantum eigensolver


## NEXT STEPS

- Build Quantum Circuit for the Collision and Streaming of qLBM
- Implement the Quantum Circuit in the Intel Quantum SDK
- Finally, Solve a simple Fluid Dynamics problem using this Circuit.
- Validate the results



## Optimisation

Quantum annealers

## Frameworks

$$
\begin{gathered}
\chi^{2}=\sum_{i j} V_{a} C_{a b}^{-1} V_{b}, \quad V_{a}=O_{a}^{(\exp )}-O_{a}^{(\mathrm{th})}(c) \\
O_{a}^{(\mathrm{th)}}(c)=A_{a}+\sum_{i} B_{a i} c_{i}+\sum_{i j} C_{a i j} c_{i} c_{j}
\end{gathered}
$$




Qibo
Qibo is an open-source full stack API for quantum simulation and quantum hardware control and calibration.


Andrea
Pasquale

## The future

- Faster, more precise calculations and event generation
- New physics: model independent searches
- Towards Quantum Computing

There are a lot of exciting ideas for us. Happy coding!

Anke Biekötter - Leonardo Cosmai Joshua Davies - Latifa Elouadrhiri

