The MEM for Analyses at the LHC

A Practitioner's View

2nd MEM Workhop | Zurich | Switzerland · January 9, 2014

Patrick Rieck · Oliver Maria Kind · Thomas Lohse · Sören Stamm
rieck@physik.hu-berlin.de

Institut für Physik
Humboldt-Universität zu Berlin
Germany

Importance of the MEM

Role in HEP Analyses

Maximum Significance

- Likelihood-ratio as most powerful test statistic for simple hypothesis tests:

$$\frac{f(x|\mathrm{H}_1)}{f(x|\mathrm{H}_0)} > k_\alpha$$

- Practically impossible to compute a complete p.d.f. $f(x|\mathrm{H})$ from MC

Analysis Approaches

Type	Discriminant	Usage
Integration	$\int_\Omega \mathrm{d}\Phi\ f(\Phi)$	Choose an appropriate region Ω, **cut and count**
Marginalization	$\frac{\mathrm{d}f}{\mathrm{d}y} = \int \mathrm{d}\Phi f(\Phi)\delta(y - g(\Phi))$	Choose an appropriate function g, e.g. from **multivariate analysis**
Approximation	$\int \mathrm{d}\Phi \frac{1}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\Phi} W(x|\Phi)$	Simplified computation of the p.d.f., known as **matrix element method**

Importance of the MEM

Advantages and Disadvantages

Advantages

- It makes sense
  - Using complete event information
  - Not based on a particular MC simulation
- It's flexible
  - Searches
  - Parameter estimations
  - Decay reconstruction
- It's powerful
  - High discriminating power
  - Sensitivity to signal *and* background
  - Impact of systematics at least as good as for MVAs

Disadvantages

- Most existing packages analysis specific
- It's slow ⇒ Need to speed up computations

Importance of the MEM

Advantages and Disadvantages

Advantages

- It makes sense
  - Using complete event information
  - Not based on a particular MC simulation
- It's flexible
  - Searches
  - Parameter estimations
  - Decay reconstruction
- It's powerful
  - High discriminating power
  - Sensitivity to signal *and* background
  - Impact of systematics at least as good as for MVAs

Disadvantages

- Most existing packages analysis specific
- It's slow ⇒ Need to speed up computations

- So far only few MEM-based analyses at the LHC
- MEM needed in order to reach quality of certain Tevatron analyses

Importance of the MEM

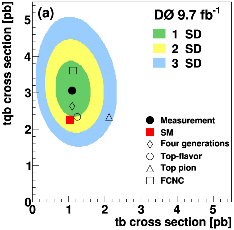
Advantages and Disadvantages

Advantages

- It makes sense
  - Using complete event information
  - Not based on a particular MC simulation
- It's flexible
  - Searches
  - Parameter estimations
  - Decay reconstruction
- It's powerful
  - High discriminating power
  - Sensitivity to signal *and* background
  - Impact of systematics at least as good as for MVAs

Disadvantages

- Most existing packages analysis specific
- It's slow ⇒ Need to speed up computations



[Phys. Lett. B 726, 656 (2013)]

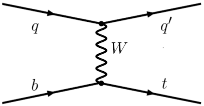
- So far only few MEM-based analyses at the LHC
- MEM needed in order to reach quality of certain Tevatron analyses
- Recent example: evidence for single-top production in the *s*-channel

MEM for Analyses at the LHC

Single-Top Physics

- ▸ Production of single top-quarks by EW interactions, access to $|V_{tb}|$, polarized top-quarks
- ▸ Dealing with large backgrounds - $t\bar{t}$, $W+\text{jets}$

- Achieved cross section measurements:
  - Tevatron : s- and t-channel simultaneously
  - LHC : t- and Wt-channel
- MEM as vital part of Tevatron analyses

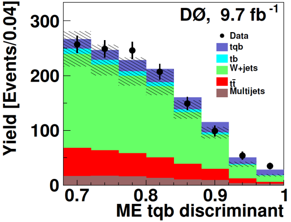


Discriminant:

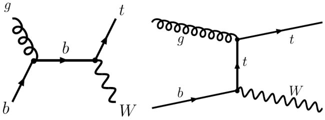
$$\mathcal{D}(X) = \frac{L_{\text{sig}}(X)}{L_{\text{sig}}(X) + L_{\text{bkg}}(X)}$$

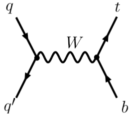
$$L_{\text{H}}(X) = \sum_{\text{proc } i \,\in\, \text{H}} f_i^{\text{H}}\, \mathcal{P}(X|i)$$

X = measured event,
f_i^{H} = event fraction of process i given hypothesis H



[Tsai, Yun-Tse, FERMILAB-THESIS-2013-09]





MEM for Analyses at the LHC

Implementation

- Development from scratch
  - Could use MadWeight in principle, however version 4 was not fast enough
  - Need flexiblity for adjustments (see below)

$$\mathcal{P}(X) = \frac{1}{\varepsilon\sigma} \sum_{p\,\in\,\text{permutations}} \int \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi \sum_{i,j} \frac{f_i(x_1) f_j(x_2)}{2 x_1 x_2 s_{\text{had}}} |\mathcal{M}_{i,j}|^2 W_p(X|\Phi)$$

- Integrate and sum up each jet–parton assignment
- Sum up initial states for each integrand
- PDFs f_i, f_j from LHAPDF (version 5)
- Amplitudes $|\mathcal{M}_{i,j}|^2$ from MCFM or MadGraph
- Transfer functions W_p involve efficiencies
- MC integration using VEGAS (CUBA implementation)
- Effective treatment of QCD radiation in progress (transverse boosts)

MEM for Analyses at the LHC

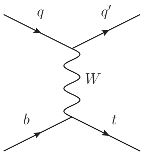
Implementation

- C++ code using ROOT
- Several processes available:
  single top s, t and Wt-channel, $t\bar{t}$, $W+2/3$ jets, $W+b\bar{b}$(+jet)
- Specific phase space generation for these processes,
  allowing for different jet multiplicities
- Simple I/O using customizable ROOT trees
- Easy to extend using inheritance from base classes for:
  - Processes
  - Phase space generators
  - Transfer functions
- Easy to use
  - Likelihood computation and I/O setup
    using simple ROOT C-scripts (see next slide)

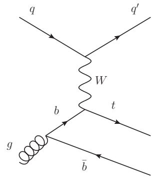
Run Script:

```
MemMgr *mgr = new MemMgr;
mgr->SetCollider(MemMgr::kPP, 8000.);
mgr->SetPdfMgr("cteq66");

MemTFcnSet *tfcnATLAS = new MemTFcnSet(MemTFcnAtlasBase::kMC12);

MemProcSgTop_tChannel_2jets *procSgTop =
    new MemProcSgTop_tChannel_2jets("tChannel", "SgTop t-channel", 172.5);
procSgTop->GetMCMgr()->SetEpsRel(0.05);
procSgTop->SetTFcnSet(tfcnATLAS);
mgr->AddProcess(procSgTop_tChannel);

mgr->SetEvtReader(new MemEvtReaderGeneric);
mgr->SetInputTreeName("t_mem");
mgr->AddInputFile("myMemInput.root");

mgr->SetEvtWriter(new MemEvtWriterGeneric);
mgr->SetOutputFile("MyMemOutput.root");
mgr->SetOutputTree("t_llh", "MEM Likelihodd Tree");

mgr->Run();
```

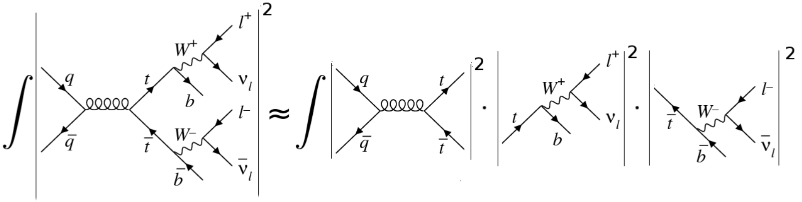
- Using MCFM amplitudes
  1. Fast computations, e.g.
     single-top *s*-channel:
     gain factor of 2 compared to MadGraph
  2. Computing all initial states simultaneously
  3. Flexibility due to different types of amplitudes, e.g.
     - 4- or 5-flavour scheme for single-top
     - Neglecting spin correlations for $t\bar{t}$



5 flavour scheme
($bq \rightarrow tq'$)



4 flavour scheme
($gq \rightarrow tbq'$)



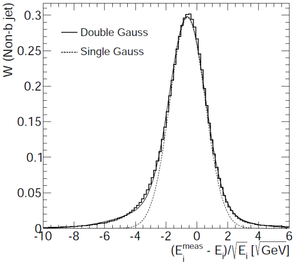
MEM for Analyses at the LHC

Transfer Functions

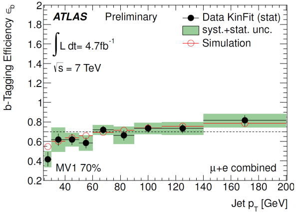
- **Double Gaussian** TFcns for e, μ and jet-energies, matching to LO parton level
- **Fixing all angles,** fixed e and μ-energies optionally
- Use $\not{E}_\mathrm{t}$ and $\varphi_{\not{E}_\mathrm{t}}$ TFcns, matched to neutrino transverse momentum
- Incorporating **reconstruction efficiencies**
  - Single object TFcn:

$$W(\mathbf{p}^{\mathrm{reco}}|\,\mathbf{p}^{\mathrm{true}}) = \varepsilon(\mathbf{p}^{\mathrm{true}}) \cdot \mathrm{DoubleGaussian}(\mathbf{p}^{\mathrm{reco}}|\,\mathbf{p}^{\mathrm{true}})$$

  - **Reconstruction** and **b-tag** efficiencies
  - $1 - \varepsilon(\mathbf{p}^{\mathrm{true}})$ for objects not reconstructed



[J. Erdmann et al., submitted to NIM A, 1312.5595 [hep-ex]]

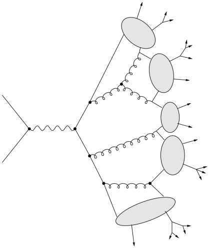


[ATLAS Collaboration, ATLAS-CONF-2012-097 (2012)]

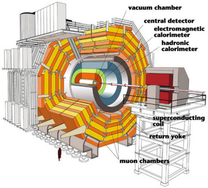
- MEM: simple approximation of highly non-trivial p.d.f.
- Estimating goodness of approximation e.g. by

1. Checking marginalizations of MEM likelihood:
   - Analytic integration of a subset of final state objects
   - Numerical evaluation of marginalizations
   - Comparison to full MC
2. Comparing MEM likelihood from full simulation with toy MC (next slide)

$$\int \mathrm{d}\Phi \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi} W(X|\Phi) \;\approx$$



+



?

e.g. lepton p_T marginalization:

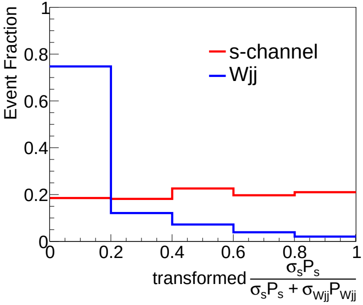
$$\frac{\mathrm{d}N}{\mathrm{d}p_\mathrm{T}^\mathrm{lep}} = \int\int \mathrm{d}\Phi \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi} W(X|\Phi) \prod_i^{i\neq\mathrm{lep}} \mathrm{d}^3 p_i \mathrm{d}\vartheta_\mathrm{lep} \mathrm{d}\varphi_\mathrm{lep}$$

$$= \int \mathrm{d}\Phi \frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Phi} \varepsilon(X|\Phi) W_\mathrm{lep}(p_\mathrm{T}^\mathrm{lep,\ reco} |\, p_\mathrm{T}^\mathrm{lep,\ true})$$

Results on parton level

- Analyse single-top *s*-channel vs. $W + 2$ jets
- Compare results with full MC simulation

1. LO event generation using MadEvent, application of transfer functions
   - Smearing of momenta
   - Event selection according to efficiencies
2. Computation of *s*-channel vs. $W + 2$ jets likelihood ratio



Likelihood ratio single-top *s*-Channel vs. $W + 2$ jets

"transformed": transformation making the signal distribution flat

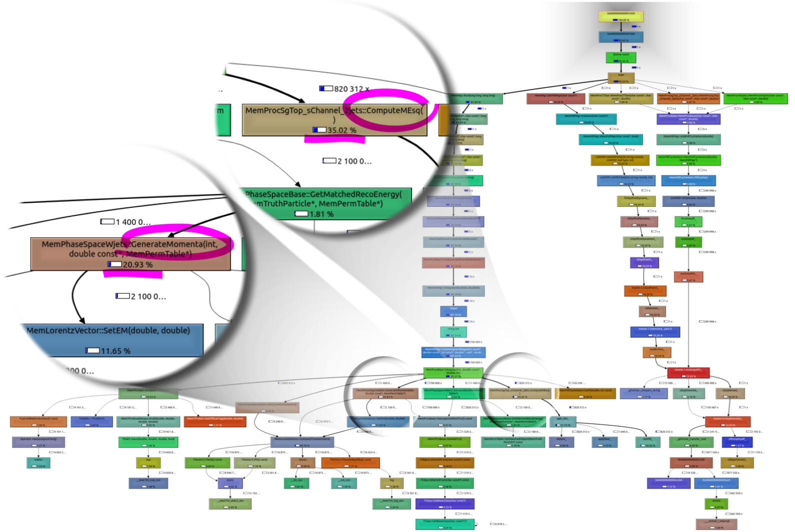
MEM Speed-Up

Overview

Means	Explanation
Compilation	Compiler options and version
Profiling	Examine program execution
Reduction of permutations	Use b-tag information, drop permutations with small contribution
Symmetrization of $\lvert\mathcal{M}\rvert^2$	For decays like $W \to qq$, $H \to b\bar{b}$ instead of permutations
Fast amplitude code	Fast MCFM amplitudes, no spin correlations (whenever reasonable), narrow width approximation
PDF caching	Faster than LHAPDF
δ Functional TFcns	Simplifies integral (whenever reasonable)
Phase space generation	Appropriate parameterization
MC integration setup	Iteration frequency of importance sampling, quasi-random vs. pseudo-random numbers
Systematics estimation	During nominal computation (see below)

MEM Speed-Up

Profiling (Valgrind)



MEM Speed-Up
Systematics Computations

Practically spending most time computing likelihoods for systematic variations:

Type	Abundance	Example	Re-Run MEM?
Scale factor	many	b-tagging efficiency	no
Modeling	few	QCD radiation	yes
4-Vector variations	many	jet energy scale	possibly not

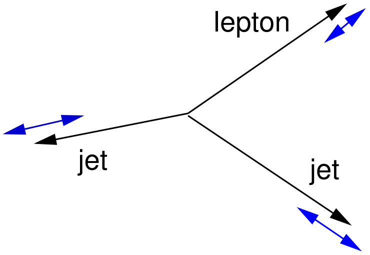
- 4-vector variations: rescaling only magnitudes of momenta

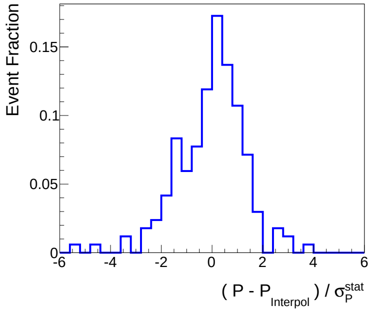
  $\Rightarrow$ Can compute systematic likelihoods approximately from nominal computation, two ways:

  1. Taylor expansion of transfer function
  2. Computation of additional likelihoods during nominal computation (see next slide)

MEM Speed-Up
Systematics Computations (ctd)

- Variation of nominal event and computation of likelihoods
- Interpolation of likelihoods
  - Using polynomial, 2^{nd} order in energies of jets and leptons
- Reduction of computation time in case of
  - Low multiplicities (⇒ few variations)
  - Many systematics





Interpolation vs. exact result

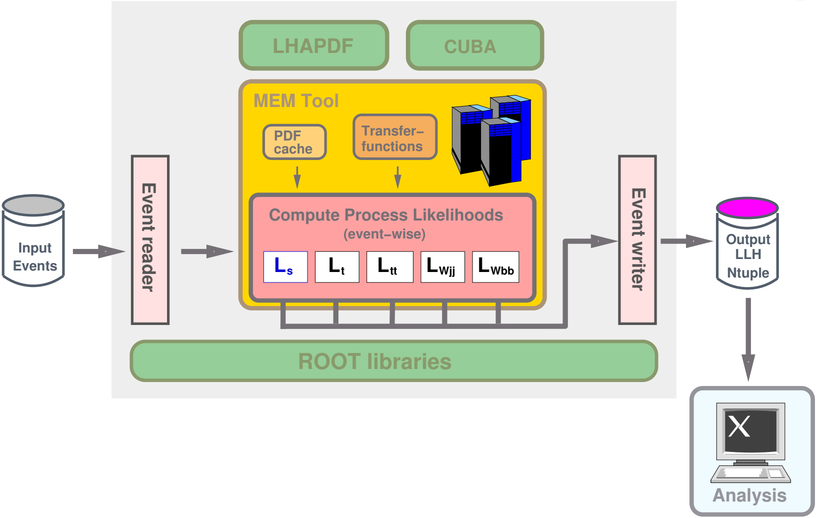
Summary

- High potential for improving analyses at the LHC
- Developing fast and flexible code
  - Easy to use
  - Easy to extend
- Significant improvement by incorporating efficiencies in transfer functions
- Ideas to test goodness of p.d.f. approximation
- Several means to speed up the MEM
- Open project, collaboration is welcome

Backup

MEM for Analyses at the LHC

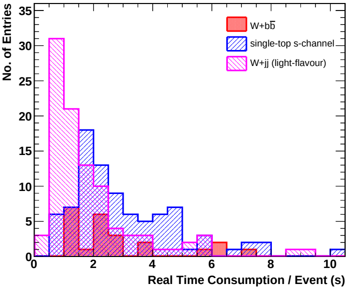
Workflow



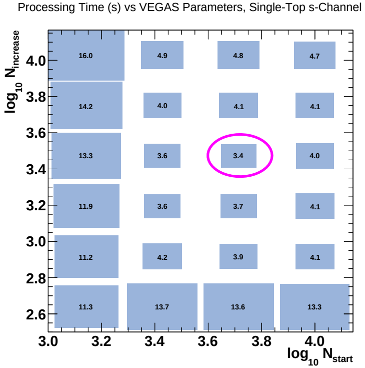
MEM for Analyses at the LHC
MC Integration Setup

Further means to reduce computation time

- Different computation time for different processes
  ⇒ Focus efforts on major ones
- Investigation of computation time vs. importance sampling frequency



Computation time for different processes



Practically spending most time computing likelihoods for systematic variations:

- 4-vector variations: rescaling only magnitudes of momenta
  $\Rightarrow$ Can compute systematic likelihoods approximately from nominal computation, two ways:
  1. Taylor expansion of transfer function

$$\mathcal{P}(X_{\text{sys}}) \approx \mathcal{P}(X_{\text{nom}}) + \mathcal{P}'(X_{\text{nom}}) \cdot (X_{\text{sys}} - X_{\text{nom}})$$

$$\begin{pmatrix} \mathcal{P}(X_{\text{sys}}) \\ \mathcal{P}'(X_{\text{nom}}) \end{pmatrix} = \int \mathrm{d}\Phi \frac{1}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\Phi} \begin{pmatrix} W(X|\,\Phi) \\ W'(X|\,\Phi) \end{pmatrix}$$

  2. Computation of additional likelihoods during nominal computation, interpolation using a set of varied events:

$$\mathcal{P}(X) \approx \sum_{n_{\text{lep}},\ n_{\text{jet1}},\ n_{\text{jet2}}\ \in \mathbb{N}_0}^{n_{\text{lep}}+n_{\text{jet1}}+n_{\text{jet2}}\ \leq 2} c_{n_{\text{lep}},n_{\text{jet1}},n_{\text{jet2}}} E_{\text{lep}}^{n_{\text{lep}}} E_{\text{jet1}}^{n_{\text{jet1}}} E_{\text{jet2}}^{n_{\text{jet2}}}$$