Publishing non-interpreted (model-independent) experimental results

David Côté (CERN)

Characterization of New Physics

5th November 2010

Thanks to B. Petersen, A. Höcker, M. Baak, J. Boyd, T. Carli, D. Froidevaux, S. Caron, M. Pierini and K. Cranmer for their help to prepare this talk!! :-)

I will present personal ideas of the authors, not official policies of the ATLAS & CMS Collaborations.

Overview

- Total *vs.* fiducial cross-sections
- Generated *vs.* reconstructed information
- Model-(in)dependence of experimental results
- Examples of experimental strategies to reduce model-dependence

Total *vs.* fiducial cross-sections

- Fiducial cross-section: partial effective cross-section for given signal topology in a restricted kinematical range
  - largely model-independent: interpretable *a posteriori* with any theoretical calculation
    - first and foremost, experiments should always publish this
    - but knowing true kinematical information is not trivial (next slide)
  - data analyzed but not interpreted
    - provides all the experimental facts but no explicit information about fundamental physics
- Total cross-section
  - experiments first interpret their data for benchmark theoretical models agreed by the community at large
    - also very important but fully model-dependent
  - difficult to re-interpret with other theoretical calculations, including those to come in the future
    - mapping to simplified models aims to address this aspect
    - experiments may publish more total cross-sections on demand
- Both approaches should be pursued

Generated *vs.* reconstructed information

- Preferred way: provide results in terms of *true* information, unfolded from experimental effects
  - true η and p_T corrected for resolution effects
  - leptons before final-state radiation (FSR)
  - jets at hadron level
    - hadron-level implies that theorists use parton shower generators (Pythia, Herwig, *etc.*)
    - parton level also possible but not rigorously defined (not an experimental observable)
  - missing transverse energy of undetectable particles
    - corrected for mis-reconstruction and resolution effects
      - dominated by true missing energy at high values
  - unfolding of involved variables can become complicated
    - for instance: multivariate outputs
- Alternative: publish reconstructed information and provide a public simulation program that can be used to connect to true information
  - see next slide

Public toy simulation of ATLAS & CMS

- Demand for super fast simulators of ATLAS & CMS by non-members
  - only for approximate studies
  - useful when published unfolded data is not available
    - *e.g.* to explore phenomenology at future luminosities
  - better work without big ATLAS/CMS software frameworks
  - high-speed is a real advantage for multi-parameter scans
- Delphes[1] and PGS[2] seem to be popular programs
  - rumor is that Delphes provides decent precision
  - maybe some parametrizations could be officially stamped by the ATLAS & CMS Collaborations?
- Another alternative: ATLAS/CMS could provide a public web interface for generic simulation requests
  - this was done by the DØ[3] and Fermi Collaborations
  - recent proposal: RECAST[4] (see talk by I. Yavin)

1: http://arxiv.org/abs/0903.2225
2: http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs.html
3: "Quaero", Phys. Rev. Lett. 87, 231801 (2001)
4: http://recast.it/

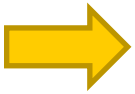
Model-(in)dependence of experimental information (I)

- Largely model-independent:
  - integrated luminosity
  - signal yield in a restricted kinematical region
    - identify and count signal events after some filtering
    - evaluate statistical and systematic uncertainties
  - signal reconstruction efficiency in the restricted kinematical region above
    - derived from control samples and detector simulation
    - constant for each decay topology at first order
      - divide data in binned sub-samples if necessary
    - residual model-dependence at 2^{nd} order:
      - lepton/jet isolation
      - η dependence
      - selection on number of particles (*e.g.* exactly one lepton)
      - multivariate analysis (?)

Sufficient to derive an effective fiducial cross-section

Model-(in)dependence of experimental information (II)

- Fully model-dependent:
  - absolute reconstruction efficiency
    - fraction of events outside the probed kinematical region must be derived from theoretical calculations and Monte-Carlo generators
  - total cross-section including all final states
    - B.F.(observed final-state) x cross-section is measured



Necessary to derive a total cross-section.

Examples of experimental strategies

The residual model-dependence of signal efficiency can be significantly reduced by using adequate analysis strategies.

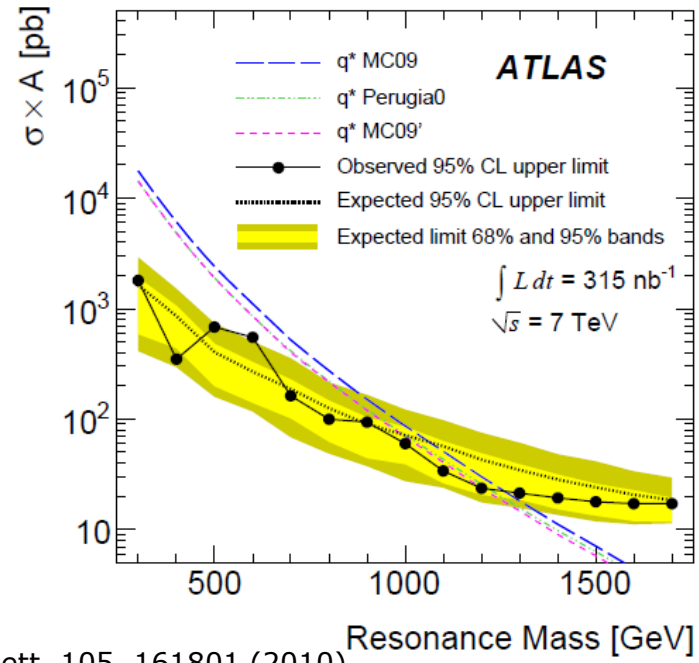
- reconstruction efficiency is not uniform in η
  - better detector performance in barrel than forward regions
  - problem: *generated* η distribution of signal is unknown
  - solutions:
    - apply η-dependent efficiency corrections
    - split dataset in η bins in which the generated distribution is effectively flat
- signal efficiency varies for different topologies
  - relax particle isolation requirements
  - use small jet cone size
  - combine several final states together, for example:
    - if only measuring yield of 1 lepton: $N_{1lep}=80$
      - generated signal can be a mixture of decays to 1,2,n leptons
    - but if measuring: $N_{1lep}=80$, $N_{2lep}=0$, $N_{3lep}=0$
      - then signal clearly comes only from decays to 1 lepton

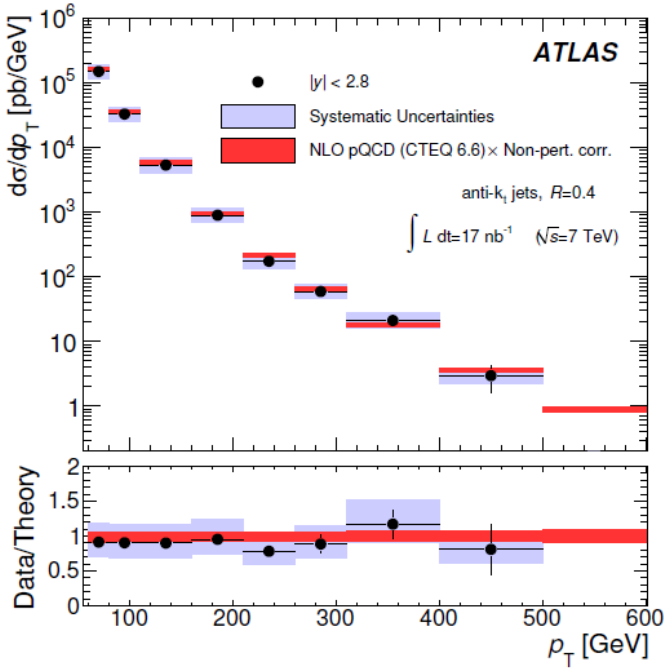
Examples of experimental strategies

- maximize the fraction of signal phase space where good S/B can be obtained
  - keeps measurement sensitive to as much signal models and topologies as possible
  - reduces uncertainty on absolute efficiency
  - use detector information as much as possible rather than cutting on phase space
    - *e.g.* "tight" particle identification, not "loose"
  - work hard to master the backgrounds instead of escaping with excessively tight kinematic cuts
    - even if difficult at background-dominated LHC experiments

Application so far in early measurements

- Selected ATLAS papers
  - model-independent search for dijets resonance[1]
    - published cross-section x signal acceptance
    - acceptance formulated in terms of reconstructed information
  - inclusive jet and dijets cross-section[2]
    - published fiducial cross-section for p_T>60 GeV and |y|<2.8
    - simple reco→true correction of jet p_T





1: Phys. Rev. Lett. 105, 161801 (2010)
2: http://arxiv.org/abs/1009.5908

Conclusions

- Factorize experimental results from their theoretical interpretation
  - publish new physics limits or observation in terms of model-independent fiducial cross-sections
    - interpretable with any calculation *a posteriori*
  - also publish total cross-sections for chosen models
- Factorize expertises/work of experimentalists and theorists
  - use unfolding algorithms to provide experimental results in terms of generated (*true*) information
  - alternatives (formulate theory in terms of reconstructed info):
    - ATLAS & CMS provide toy simulation programs that theorists use
    - ATLAS & CMS fulfill requests for simulating new models themselves