Diffraction at the LHC

at the boundary of experiment and theory

Mikael Mieskolainen, Risto Orava

Helsinki Institute of Physics (HIP), University of Helsinki

Diffractive and Electromagnetic processes - WE-Heraeus Summer School 2015
Bad Honnef

mikael.mieskolainen@cern.ch

19.8.2015

Hmm?

Let us first quickly go through the current Monte Carlo models of inclusive *pp*-diffraction (& soft QCD / minimum bias). 15 min talk, so we basically skip these. [FYI, CERN Forward Physics Yellow Report has a review of these]

Then introduce a new analysis technique, a technique which generalizes large rapidity gaps (LRG) and differential gap distribution measurements

Finally discuss the interplay between Monte Carlo and the new analysis technique! How to get via probabilistic way:

$$\sigma_{inel} \triangleq \sigma_{SDL} + \sigma_{SDR} + \sigma_{DD} + \sigma_{ND} + (\sigma_{CD}) \tag{1}$$

Traditional large rapidity gap (LRG) analysis

The de-facto kinematical signature of diffraction (coherence)

Search for a gap of $\Delta\eta \geq 3$ units (same as $\xi = 1 - p_z^f/p_z^i = M_X^2/s \leq 0.05$) by requiring no tracks, hits or energy deposit over some experimental threshold in the given η-interval. **NB! No gap definition without corresponding p_T "threshold" definition..**

So LRG $\simeq$ Diffraction, but not with $=$. With fixed (or floating) gaps, one can certainly select a subset of diffractive events, no doubt. And reject most of the non-diffractive events where gaps are mostly coming from hadronization *fluctuations* (presumably exponentially supressed). However, the "full picture" requires different approaches.

A few words about (soft QCD) Monte Carlo models

Mostly Regge theory based

1. Classic triple Regge $\propto \frac{1}{M_X^2} e^{-bt}$ parametrization

`PYTHIA 6,8`, with MPI ($2 \rightarrow 2$ QCD with $p_T \rightarrow 0$ regular.) for ND. P8 includes 5 different parametrizations for "Pomeron flux", including Min-Bias Rockefeller (MBR), and p_T spectrum equivalent with PHOJET, P6 with softer p_T spectrum.

`PHOJET`, last official update in $\sim$ 2001

2. Cosmic Ray Shower generators

`QGSJet-II-04`, Gribov's Reggeon Field Theory (RFT) based

`SIBYLL`, Dual Parton Model with minijet production

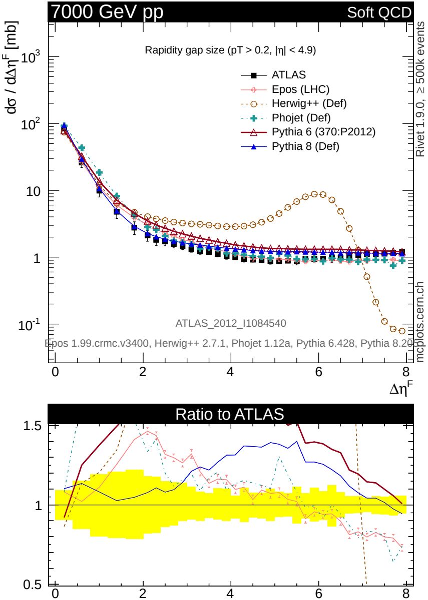
`EPOS LHC`, simultaneous parton ladders, the ridge structure in $(\Delta\eta, \Delta\phi)$

3. Interesting new models

`SHRiMPS` (Sherpa), KMR ladder evolution, Good-Walker for low mass N^*

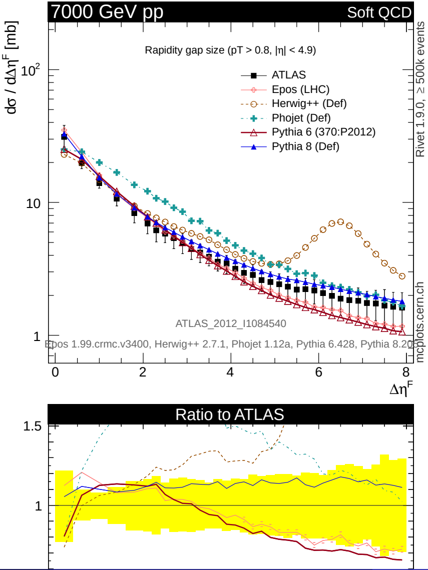
`DIPSY` (Lund), Dipole evolution (inspired by LL BFKL model by Mueller)

Floating gaps $d\sigma/d\Delta\eta^F$, $p_T > 0.2$ GeV, $|\eta| < 4.9$

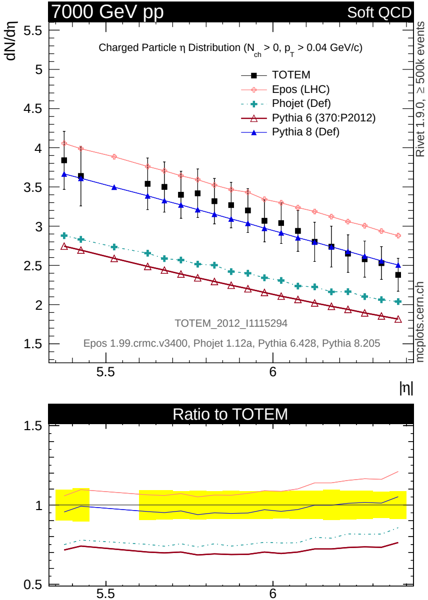


Floating gaps $d\sigma/d\Delta\eta^F$, $p_T > 0.8$ GeV, $|\eta| < 4.9$

Trivial observation: higher p_T threshold produces artificial gaps.



Forward $dN_{ch}/d\eta$; mcplots.cern.ch



QCD analyses...

- As an example: jet algorithms. Nowadays a full machinery of claimed technology for jet substructure, color flow inversion, radiation patterns, probabilistic jets (QJets), special techniques for highly boosted...
- **What is needed: A fully coherent QCD analysis framework, analysing the event topology and particle/energy flow with all Q^2 scales...** Transition from non-diffractive to diffractive, from low-Q^2 to multijet events with underlying event understood. Able to resolve perturbative phenomena from non-perturbative, beyond the usual IRC safe criteria. One day maybe...

Probabilistic multivariate analysis in space of $\mathbb{R}^N$

No large rapidity gaps explicitly required!

Basic idea: Vectorize tracking, hits, (& calorimetry) over experimentally available pseudorapidity η into N dimensional vector $\mathbf{x} \in \mathbb{R}^N$. This approach uses optimally the final state topology and particle/energy flow.

Google out previous work:

M.M., *The Existence and Uniqueness of Diffraction at the LHC*, Talk given at `Diffraction 2014`.

M.M., *Bayesian Classification of Hadronic Diffraction in the Collider Detector at Fermilab*, MSc Thesis, 2013.

M. Kuusela, J.W. Lämsä, E. Malmi, P. Mehtälä, and R. Orava., *Multivariate techniques for identifying diffractive interactions at the LHC*, International Journal of Modern Physics A, 2010.

Previous work: Pairwise posteriori distributions (CDF)

Distributions below demonstrate the non-unique signature of real events and continuum transitions between the experimental signatures. How do you do show this with LRG selection..? Can't do it.

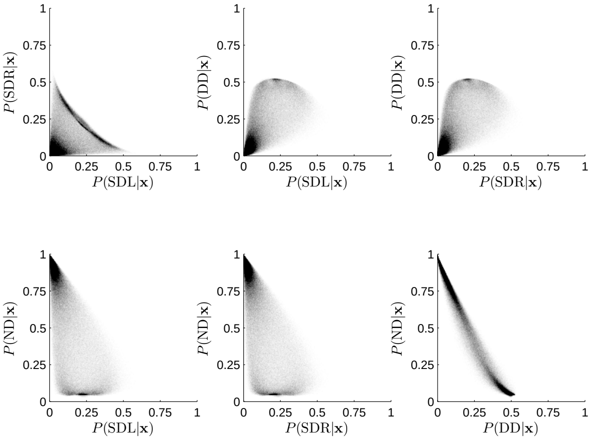


Figure : CDF $\sqrt{s} = 1.96$ TeV 0-bias data, MLR-ℓ_1 algorithm + PYTHIA 6.x MC.

Previous work: Regularization paths with PYTHIA 6

ℓ_1-regularization induces rapidity gaps as a limit when $\lambda \to \infty$

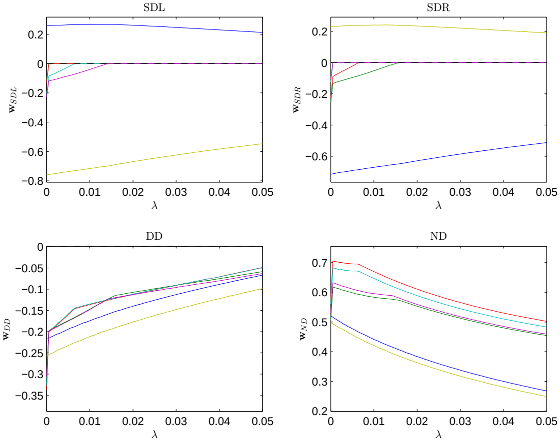


Figure : On y-axis the coefficients of $\mathbf{w}_j$ in order: $w_i :=$ (blue, green, red, light blue, purple, yellow), with binning $\mathbf{d}_\eta = (-3.6, -1.8, -0.9, 0, 0.9, 1.8, 3.6)$, such that $\eta_{\min,\max}(w_i) \in [d_i, d_{i+1}]$. Variables are calorimeter deposits integrated over ϕ.

New method: Topological combinatorics

Take simplifying binary limit of real valued multivariate analysis $\mathbb{R}^N \to \mathbb{B}^N$, $\mathbb{B} = \{0, 1\}$

Figure : x-axis $\sim \eta$ with 6 possible divisions of pseudorapidity span, with ϕ integrated over. Number of unique binary vectors per space: $2^2, 2^4, 2^6, 2^{12}, 2^{24}, 2^{48}$ (top to bottom).

These different combinations $\sim$ partial cross sections σ_i capture a huge range of different final state topologies and particle flows!

How to choose the division of η-space?

Naturally there are experimental boundary conditions

Formally, we have function spaces V_j ordered by (gap) *resolution* in rapidity

$$V_{-\infty} \subset \cdots \subset V_{-1} \subset V_0 \subset V_1 \cdots \subset V_\infty \quad (2)$$

where $V_{-\infty} = \{f(\eta) = 0\}$ ($\sim$ empty detector) and $V_\infty = L^2(R)$ (any-square integrable functions). In practise, one must go with the

experimental limitations and physical interpretation in mind.

If one builds all resolution combinations, the analysis procedure has a mathematical interpretation as a multiresolution wavelet basis tree expansion with *Haar wavelets*. (Intuitively a bit like renormalization group flow.)

Bayesianism: Posterior $\propto$ Density $\times$ Prior

Golden rule idea behind Bayesian inference: Update your prior knowledge with the new measurement

$$P(C = j | \mathbf{X} = \mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}|j)P(j)}{f_{\mathbf{X}}(\mathbf{x})} = \frac{f_j(\mathbf{x})P_j}{\sum_{j'=1}^{|\mathcal{C}|} f_{j'}(\mathbf{x})P_{j'}} \tag{3}$$

Densities (likelihoods) f_j
with $j = 1, \ldots, |\mathcal{C}|$, ($\mathcal{C}$ is a discrete set of scattering processes) encapsulate the **theoretical input** about differential cross sections (e.g. triple Pomeron $1/M_X^2$) + hadronization phase (e.g. Lund string) and detector/reconstruction response (GEANT)

Priors P_j
encapsulate the theoretical integrated cross sections, e.g. single diffraction $P_{SD} \propto \int\int dM_X^2 dt \frac{d^2\sigma_{SD}}{dM_X^2 dt}$ (MC) $\times$ efficiency $\times$ acceptance (GEANT)

Example of fixed resolution binary combinatorial analysis

So we have multivariate classification with discrete class densities f_j

Generator level detector combinatorics 2^N (here $N = 4$) simulation with PYTHIA 8:

Table : First five signatures out of $2^4 = 16$ possible, class fractions are $x_j \sim f_j \times P_j$, partial cross sections per given combination denoted with σ_i. $\eta_{--}, \eta_{-}, \eta_{+}, \eta_{++}$ denote 4 regions in pseudorapidity. **Numbers just for an illustration**.

ID	η_{--}	η_{-}	η_{+}	η_{++}	x_{ND}	x_{SDL}	x_{SDR}	x_{DD}	x_{CD}	σ_i (mb)
0	0	0	0	0	0.00	0.38	0.39	0.17	0.06	3.4417
1	0	0	0	1	0.00	0.00	0.65	0.31	0.03	1.2377
2	0	0	1	0	0.03	0.00	0.46	0.27	0.24	0.4832
3	0	0	1	1	0.04	0.00	0.57	0.36	0.03	3.9924
4	0	1	0	0	0.03	0.46	0.00	0.27	0.24	0.4797

⋮

Cross-sections via probabilities

"Soft classification", is basically a mixture estimation/inversion problem

It is well-known that conditional expectation values obey the so-called *iterated expectation* relation

$$\mathbb{E}[h(\mathbf{X},\mathbf{Y})] = \mathbb{E}[\mathbb{E}[h(\mathbf{X},\mathbf{Y})|\mathbf{Y}]] = \mathbb{E}[\mathbb{E}[h(\mathbf{X},\mathbf{Y})|\mathbf{X}]], \quad (4)$$

where $\mathbf{X},\mathbf{Y}$ are random vectors and $h(\mathbf{X},\mathbf{Y})$ some arbitrary function of those.

Using this, one can show easily that integrating (summing) posteriori probabilities over an event sample size of n results in relative cross section for the k-th scattering process class

$$\frac{\sigma_k}{\sigma_{inel}^{vis}} \cong \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{|\mathcal{C}|}\delta(j-k)P(j|\mathbf{x}_i) \quad (5)$$

Different *philosophies* to choose priors

Bayesian vs Frequentist inference...

1. **Fully Bayesian**, induce distributions for the priors P_j using domain knowledge, and maybe some physical constraints as unitarity, Regge factorization, some symmetry etc. and finally integrate over posteriors $\rightarrow$ a simple sampling algorithm needed $\rightsquigarrow$ Bayesian credibility intervals a natural side-effect

2. **Point priors**, semi-Bayesian where one uses e.g. priors from the given MC model $\rightarrow$ Just use the Bayes' formula shown earlier, no computational complications

3. **Maximum Marginal Likelihood**, $\arg\max_{\{P_j\}} \prod_{i=1}^{n} \sum_{j=1}^{|\mathcal{C}|} f_j(\mathbf{x}_i) P_j$ i.e. maximize the denominator (evidence) in the Bayes' formula over the number of n events. No closed form solution, but iterative $\circlearrowleft$ Expectation Maximization (EM) algorithm can be derived (classic frequentist mixture density problem)

Analysis output

So, *model independent* measurement is the measurement of partial cross sections σ_i, where $i = 1, \ldots, 2^N$, i.e., for each i-th combination.

In principle, unfolding mapping $U : \{\sigma_i\}_{\text{detector}} \rightarrow \{\sigma_i\}_{\text{corrected}}$ must be done for comparison with theory / MC models (as with every measurement, in principle). Or other way around, MC must be folded.

What is *model dependent* is the **probabilistic mapping**: $P : \{\sigma_i\} \rightarrow \{\sigma_{SD}, \sigma_{DD}, \sigma_{ND}\}$. For this, we use **MC input** in terms of class likelihoods f_j. However, the class priors P_j can be estimated from data $\Rightarrow$ semi-model dependent.

RIVET[1] analysis for future proof comparison and MC tuning!

[1] The Rivet project (Robust Independent Validation of Experiment and Theory) is a toolkit for validation of Monte Carlo event generators: https://rivet.hepforge.org/

Conclusions

Details of the analysis techniques will be discussed in future...

We introduced a generalized soft diffraction analysis: Multiresolution topological combinatorics $\supset$ gap distributions $\supset$ large rapidity gaps

Probabilistic multivariate approach can naturally handle the **non-unique** experimental signature between diffraction / non-diffraction and deals *naturally* with experimental limitations such as p_T-thresholds.

What is really needed, is better interplay between theoretical and experimental (+algorithmic) definitions of diffraction!