Update on RooStatsCms

Grégory Schott
Danilo Piparo
Günter Quast

Universität Karlsruhe

Atlas-CMS statistics forum, September 25th, CERN

RooStatsCms status → *current global status is "production"*

RooStatsCms only

- Intuitive model "factory"
  - Build the analysis model from a text configuration file "datacard"
  - Datacard also describe nuisance parameters (and correlations)
  - Building a combined model for the combined analysis
- Outputs a standard RooFit PDF describing the analysis
  Remember what parts are signal and background contributions

RooStatsCms and RooStatsKarlsruhe

- Statistical methods distributed in RooStatsKarlsruhe (public since March'08)
- Implementation of nuisance parameters and correlations *completed**
  - *Can be marginalized or profiled*
- Statistical methods:
  - LimitCalculator (CL_B, CL_{SB}, CL_S) *completed**
  - PLScan (profile likelihood) *completed**
  - FCCalculator (fully frequentist approach) *validation to complete*
  - Bayesian approach and Markov chains *being investigated*

  ** strong implementation, tested and used by CMS analyses*
- Batch friendly: decomposition in sub-jobs; results stored in ROOT files
  - Results can be merged and exploited by results classes

You will find in the backups applications to CMS analyses

- Reproducing CMS physics TDR significances — ***VALIDATIONS***
  - one of the validation and X-check we are doing
  - other ones being:
    - comparison to other statistical codes used in CMS
    - comparison to classes in ROOT (TLimit, TFeldmanCousins, ...)
    - CDF single-top statistical package "Bill"

- CL_B, CL_{SB} and CL_S in the VBF $H \rightarrow \tau\tau$ analysis — ***PRODUCTION USAGE***
  - significance using CL_B
  - upper-limits on σ_H / σ_H^{SM} using CL_S
  - production of LEP- and Tevatron- like plots showing the results
  - those results have been approved yesterday by the CMS collaboration

- Comparison to the profile likelihood approach in this analysis
  - over-coverage of upper-limits with low signal
  - proper coverage with larger signal

- Application to other CMS analyses
  - not in backups (non-public CMS results)
  - regular reports to CMS Higgs WG: come hear us to keep up-to-date!
  - strong interrest by other CMS working groups

RooStatsKarlsruhe Class Index

C | F | L | M | N | P | S

Main Page | Namespaces | Classes | Files

Alphabetical List | Class List | Class Hierarchy | Class Members

RooStatsKarlsruhe by doxygen 1.5.1

RooStatsKarlsruhe Class Index

C | F | L | M | N | P | S

C
Constraint
ConstrBlock2
ConstrBlock3
ConstrBlockArray

F
FCCalculator
FCResults

L
LEPBandPlot
LikelihoodCalculator
LimitCalculator
LimitPlot
LimitResults

M
Minus2LnQCalculator

N
NLLPenalty

P
PdfCombiner
PLScan
PLScanPlot
PLScanResults

S
StatisticalMethod
StatisticalPlot

Statistical Methods – Mother: StatisticalMethod

LimitCalculator | PLScan | FCCalculator

Constraints
Mother:
NLLPenalty.cc

Constraint.cc
ConstrBlock2.cc
ConstrBlock3.cc

ConstrBlockArray.cc

Job results collection from batch system

Statistical Results – Mother: StatisticalResult

LimitResults | PLScanResults | FCResults

Statistical Plots – Mother: StatisticalPlot

LimitPlot | PLScanPlot (add also FC curves) + LEPBandPlot | ExclusionBandPlot

Conclusion

- RooStatsCms is in production phase: used for CMS analysis results
  - used by the VBF H $\rightarrow \tau\tau$ analysis: approved by CMS yesterday
  - some other CMS analyses using the tool as well
  - we provide user support within CMS to use the tool
    - documentation (installation, methods, usage), lots of examples, doxygen-style comments of classes methods and members
      http://www-ekp.physik.uni-karlsruhe.de/~RooStatsKarlsruhe
    - workshop with hands-on-exercises in June'08 *(another one planned)*
- Limited manpower (responsabilities for T1-FZK, Higgs, QCD analyses)
  - Need to set priorities:
  - is now on doing a combination of the CMS Higgs analyses (in October)
- Worries with RooStats related to manpower and organizational structures
  - *technical maintenance of such a package*
  - *how to make decisions on contributions to go in*
  - *test and validation + release management*

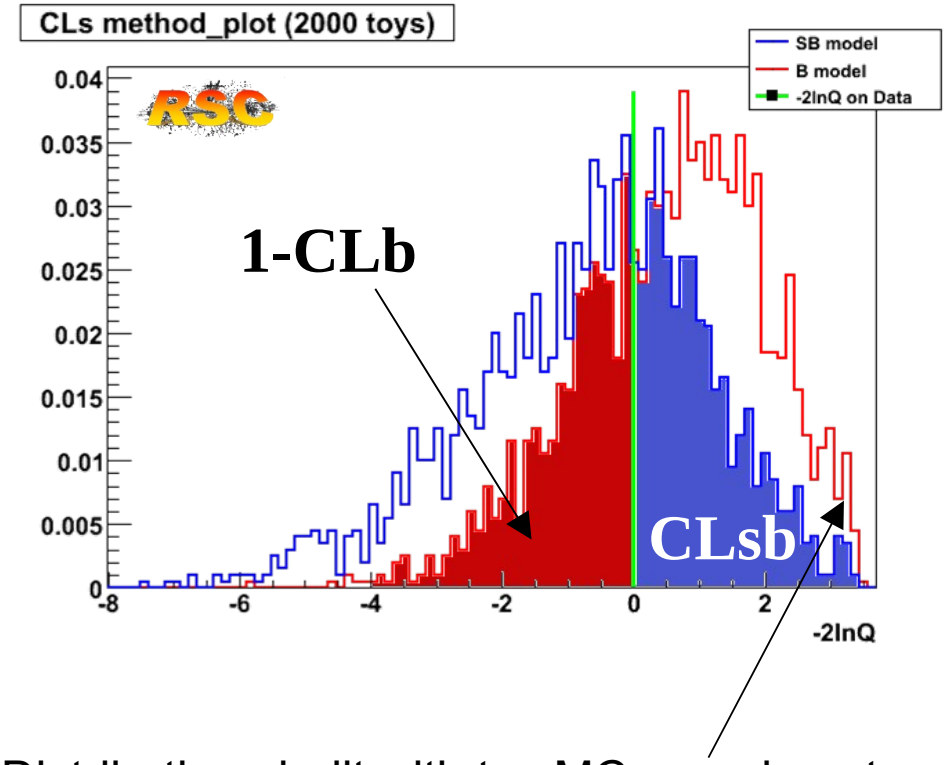
BACKUPS

All statistics methods based on the likelihood function build using RooFit

- Build the analysis model(s) using RooStatsCms
- Number counting analysis $$L(N;n_s,n_b)=\frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}$$
- Analysis using PDFs
$$L(x;n_s,n_b,\theta)=\frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!}\prod_{i=1}^{N} n_s f_s(x,\theta)+n_b f_b(x,\theta)$$
- Combination of analyses:
  - Simultaneously apply to the data of each analysis its likelihood function
- Vary nuisance parameters in toy-MC experiments
- Generate toy data samples
- Apply fit constraints (if needed), add to the *- log L* a term:
  for correlated gaussians: $\log L_P \sim 0.5\cdot(\underline{m}-\underline{m}_0)^T \cdot V^{-1} \cdot (\underline{m}-\underline{m}_0)$, V is correlation matrix
  other types of nuisance parameter shapes possible
- Compute the likelihood over that data sample (with or without fit)

The “CLs” technique

- The name “CL_s”: just a part of the method
- Used at LEP and Tevatron
- Idea: *separation of hypotheses analysing distributions of likelihood functions ratios*
- Variable $Q=L_{s+b}/L_b$ with L_{s+b} ,L_b likelihoods in the sig+bkg and bkg only hypotheses



Distributions built with toy MC experiments

Treatment of systematics:

For every toy MC experiment, before the generation of the toy dataset, parameters affected by systematics are properly fluctuated.

The “CLs” technique – Significance

- CL_B : background CL, measure the compatibility of the experiment with the B-only hypothesis

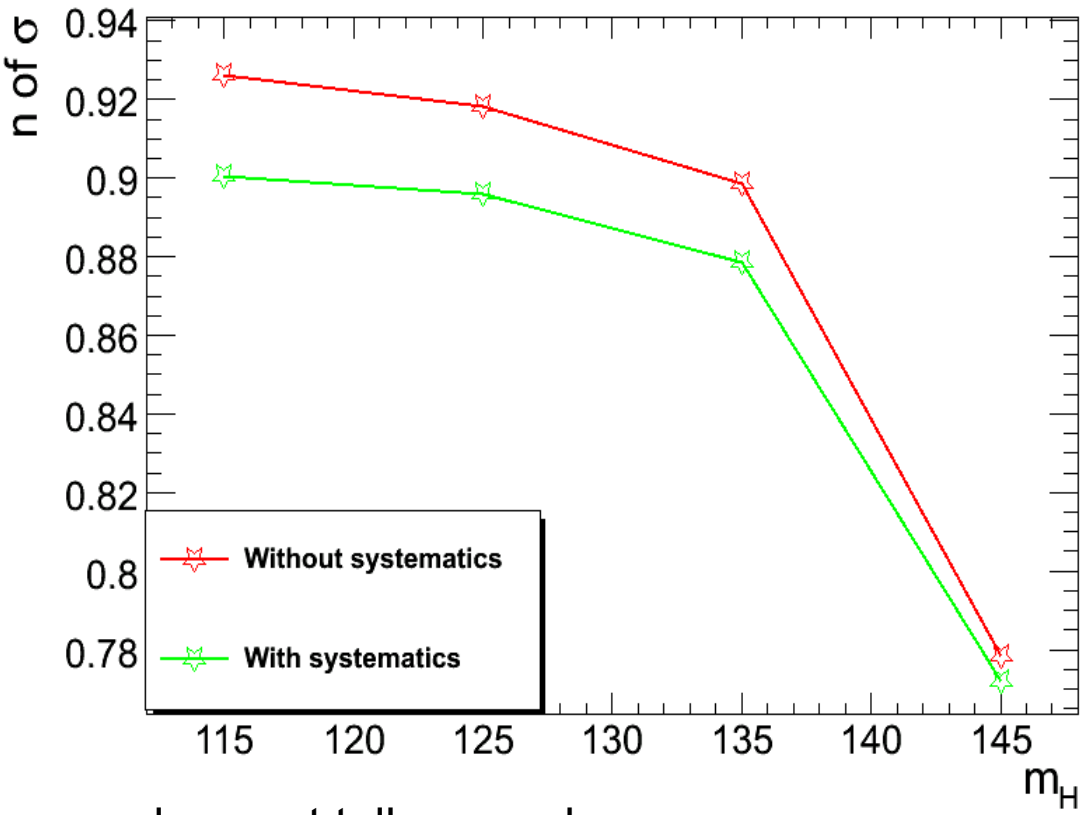
- $1 - CL_B$: probability for a B-only experiment to give a more S+B-like likelihood ratio than the observed one

- CL_{S+B} : measure the compatibility of the experiment with the S+B hypothesis
  - if CL is small (< 5%) the S+B hypothesis can be excluded at more than 95% CL but it does not mean that the signal hypothesis is excluded at that level

- CL_S : the signal significance is a-priori defined to be: $CL_S = CL_{S+B} / CL_B$

H→ττ: Significance

- Significance calculated for the H→ττ analysis using CLb



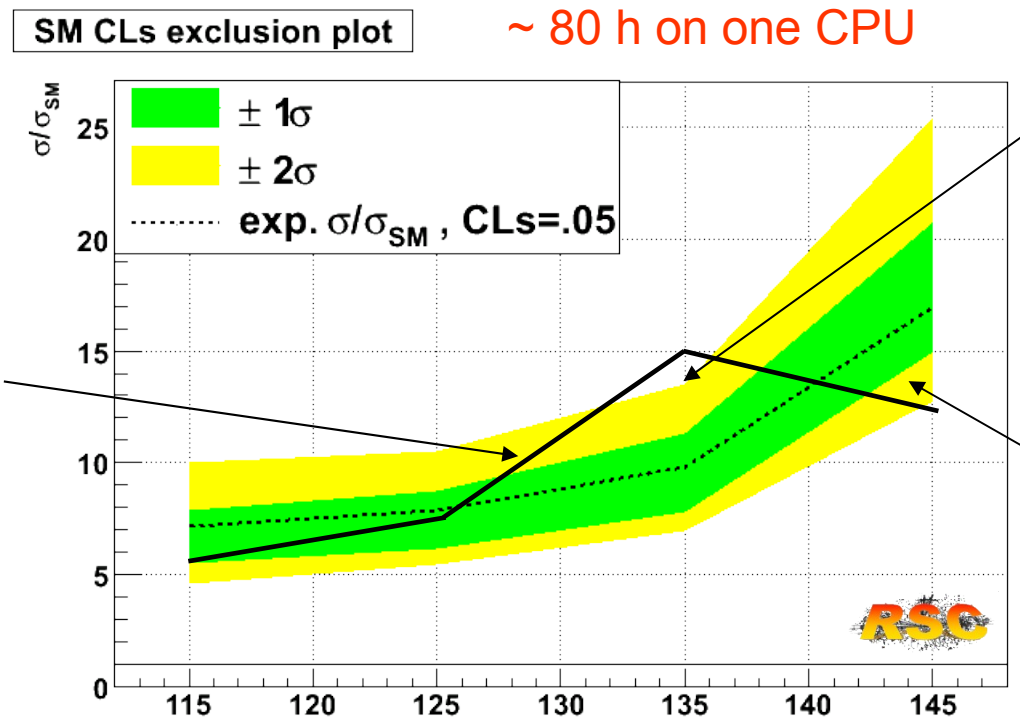
- In this case significance does not tell us much.
- The question becomes:

"Which production cross section can I exclude with the data I have?"

The “CLs” technique – Exclusion

- Method used upon review committee request
- Assume to observe only background
- Amplify the SM production cross section by a factor necessary to obtain CL_s=0.05

  → “95% exclusion”



Obtained with real data

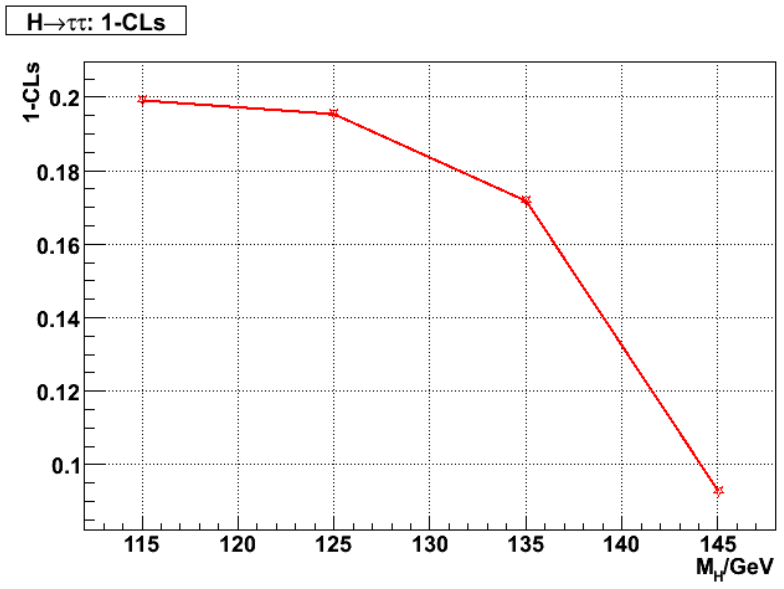
Less exclusion power than expected: “bad luck”

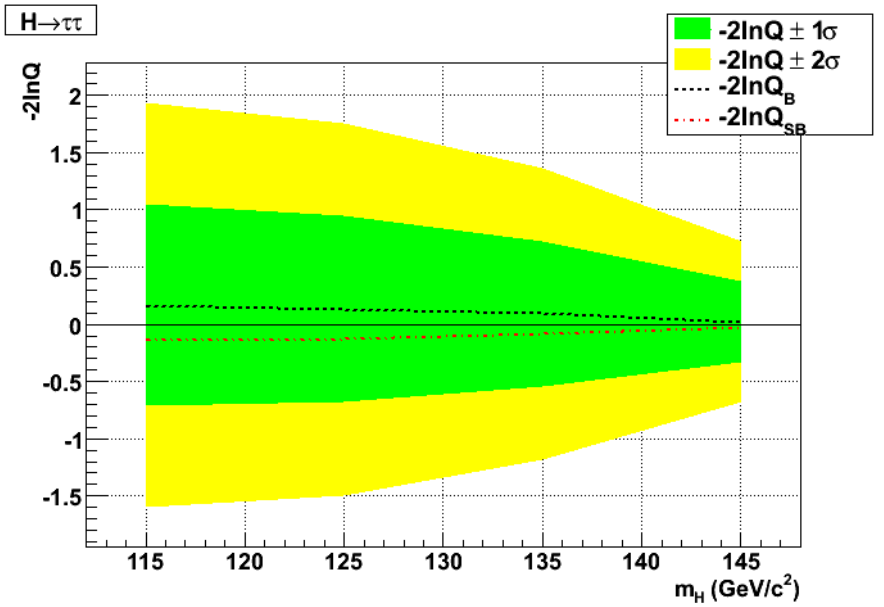
More exclusion power than expected: “good luck”

Bands:

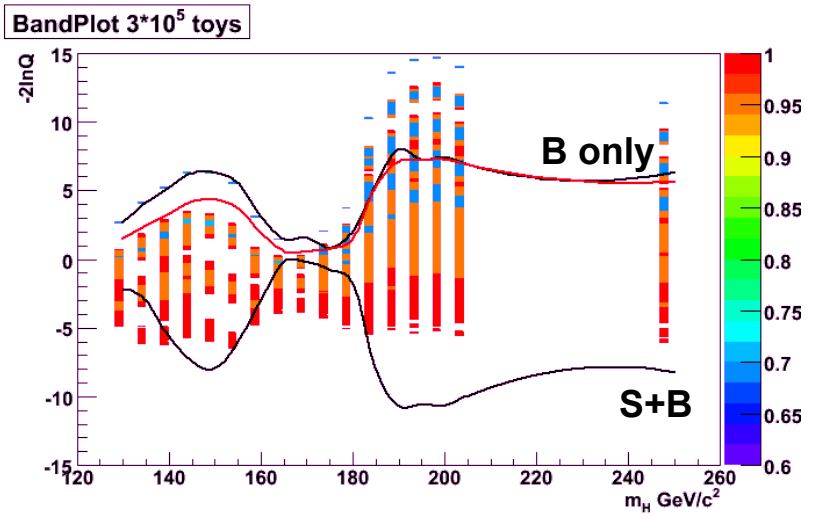
- Assume to observe $N_b + n \cdot \mathrm{sqrt}(N_b)$, where n=2,1,-1,-2 for the -2,-1,1,2 sigma band border respectively
- Systematics taken into account in distributions of -2lnQ

Other plots





In "another" analysis :



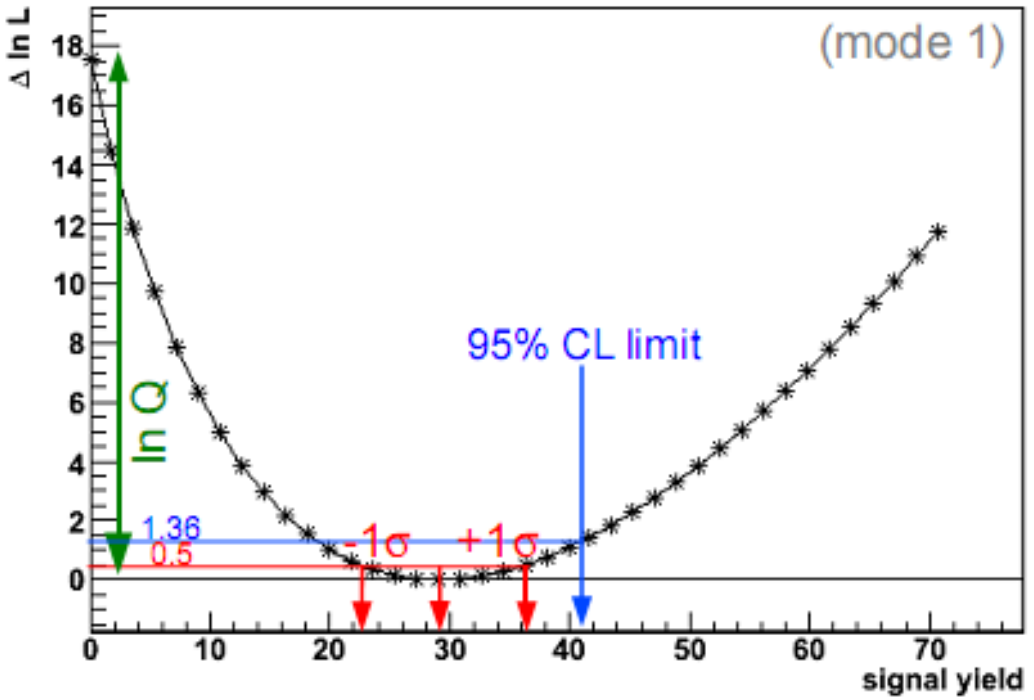
The “profile likelihood” technique

1) Each nuisance parameter x_i becomes a fit parameter

2) Add to the combined -log(likelihood) a term $\frac{1}{2} (\mathbf{x}-\mathbf{x}_m)^T \mathrm{cov}^{-1} (\mathbf{x}-\mathbf{x}_m)$ to take into account constraints
(in gaussian case)
($\mathbf{x}_m$ and cov represents the vector of external constraints and their correlated errors)

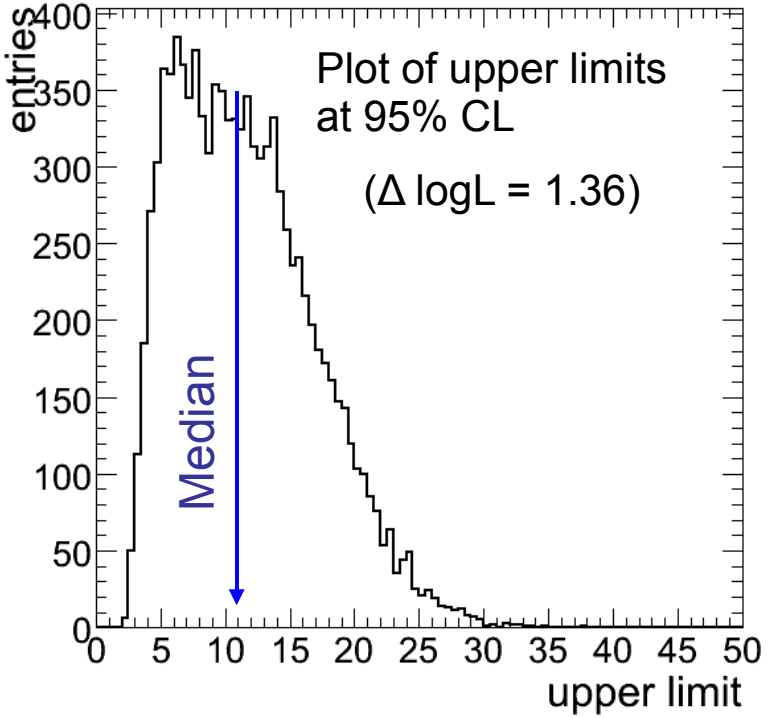
3) Vary the signal yield and minimise w.r.t. all other parameters
=> thus obtaining the "profile likelihood"

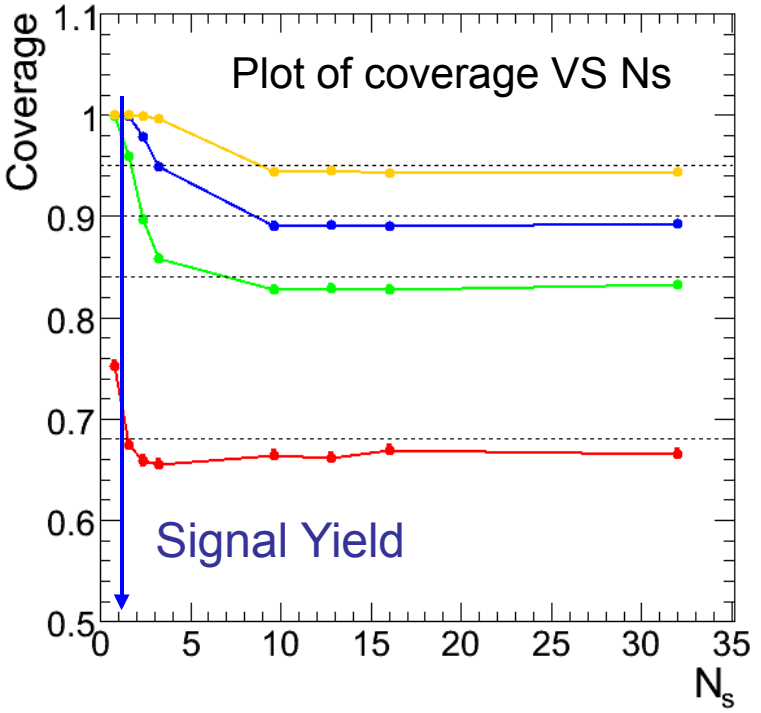
Signal significance:
$S = \mathrm{sqrt}\,(2\,\mathrm{ln}Q)$
$= \mathrm{sqrt}\,(2\,\mathrm{ln}\,L_B - 2\,\mathrm{ln}\,L_{S+B})$



Profile likelihood – limits and coverage

- To compare with the $H \rightarrow \tau\tau$ (no systematics here)
- Profile the likelihood function and search for the upper-limit using $\Delta \log L$
- Much faster (1 single fit, i.e. a minute or two)
- With profile likelihood the 95% CL UL is 10.71 events = 6.7 SM cross section
  - to compare to ~5.5 with CL_s
- Test of coverage:
  - For low signal yields, the profile likelihood method largely over-covers
  - The method works well for large signal (and luminosity)





Combination of analyses

- Significance - sqrt(2lnQ) - curves for various analyses.
- The CMS PTDR studies are compared with the one obtained with RSC