Limits in HEP experiments for pedestrians

Stefan Schmitt, DESY

SFB 676 Block meeting

Hamburg February 23/24, 2012

Outline

- Frequentist and Bayesian probability
- Probability in High energy physics
- Confidence intervals, limits
- Data analysis
- Limit calculation
- Downward fluctuations
- Expected limits
- Some Examples

Frequentist/Bayesian probability

- Frequentist view: probabilities describe the outcomes of experiments

  Models have unknown parameters. Probabilities to make an observation are given as a function of the model parameters
- Bayesian extension: probabilities are also used to describe the "degree of belief" in the model parameters.

  → The model parameters themselves can have probabilities assigned.

- Remark: throughout this talk a "model" is a cross-section prediction with certain free parameters

Probabilities in high energy physics

- Probabilty theory: predict number of events given the model parameters and the experimental setup

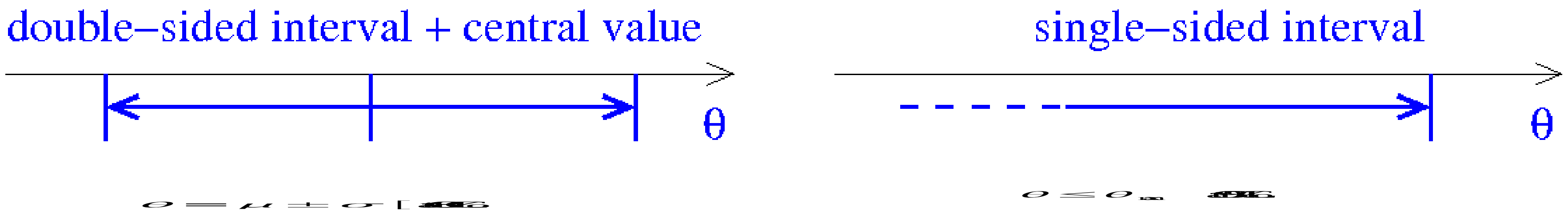
- **Frequentist analysis:** give for each model the probability of the observation (there is no probability of the model itself)
- **Bayesian analysis:** assign a probability (degree of belief) to each model
- High energy physics: preference for **Frequentist** methods, in particular for discoveries

  → Bayesian methods are not discussed further in this talk

Confidence intervals, Limits

- Confidence intervals tell about parameters of a model
- Confidence level (CL): associated probability

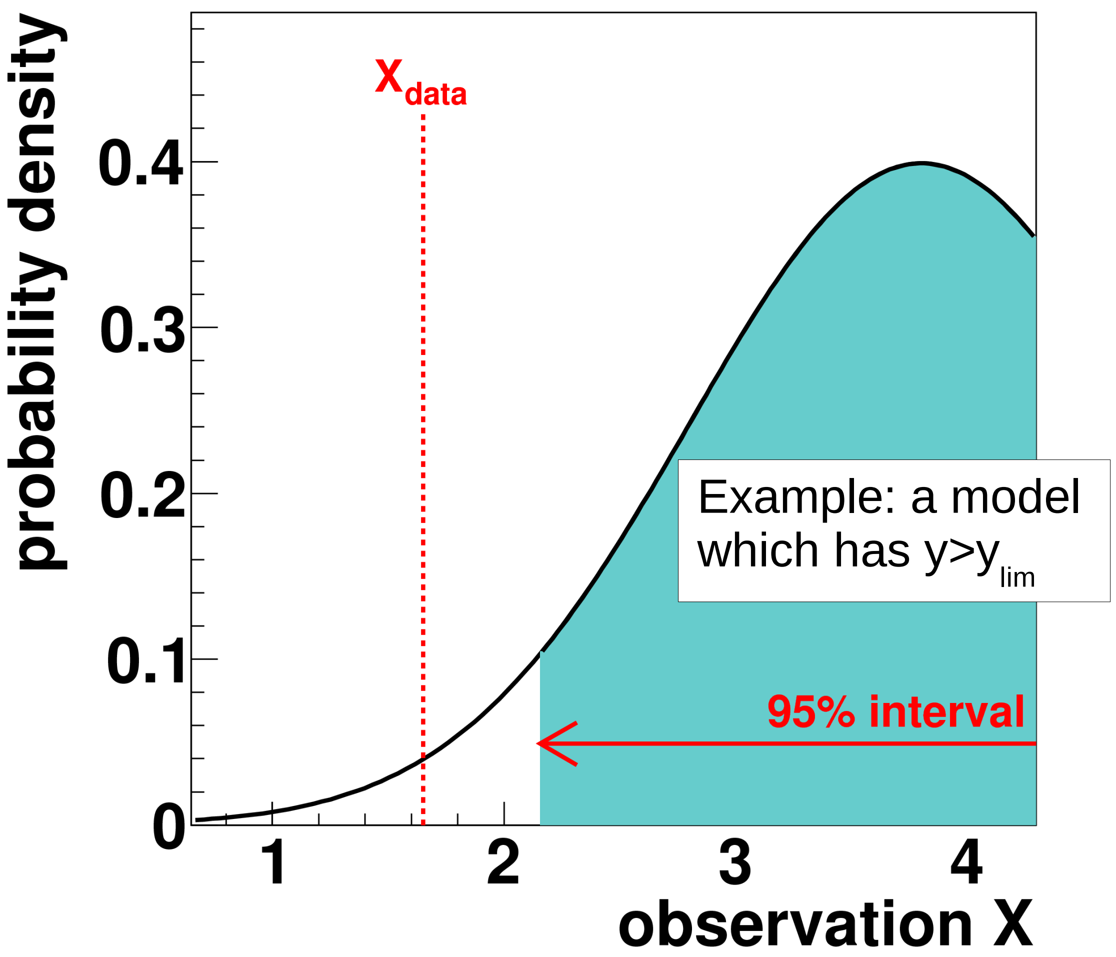
- Double-sided interval: measurement (usually CL=68%)
- Single-sided interval: limit (often CL=95%)



Frequentist exclusion limits

- What the statement "$y>y_{lim}$ is excluded at 95% CL" really means:

  Models having parameter $y>y_{lim}$ predict the observation at a probability of less than 5%

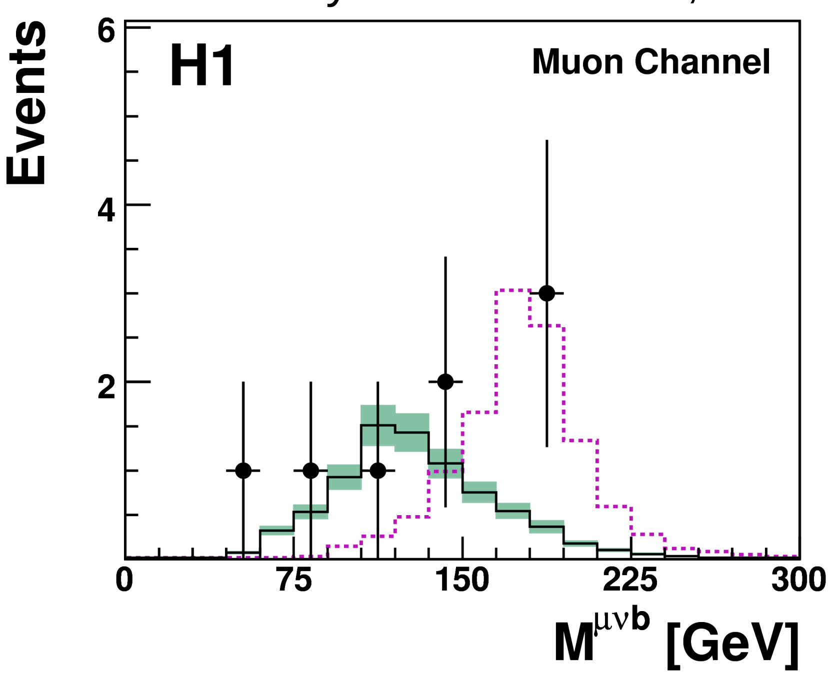


- The model together with the experimental conditions defines a probability density
- Excluded models are those where the measurement is outside their respective 95% region

Data analysis on one slide

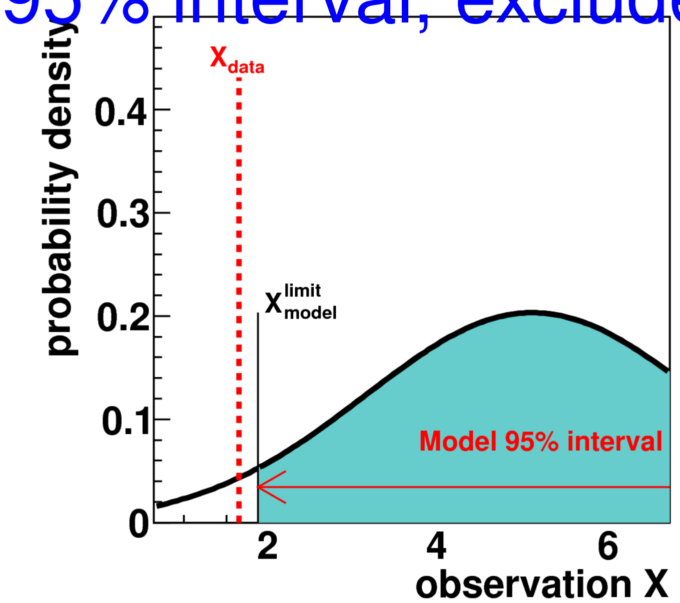
- Count Data events in bins (distinct phase-space regions)
- Predictions are based on event simulations “Monte Carlo generators”, plus detector simulation
- Predictions have systematic uncertainties
  - “theory”: scale uncertainties, parton densities, ...
  - “experimental”: calibration, luminosity, ...
- For each bin: Poisson's law to predict N events and Poisson mean has systematic uncertainties

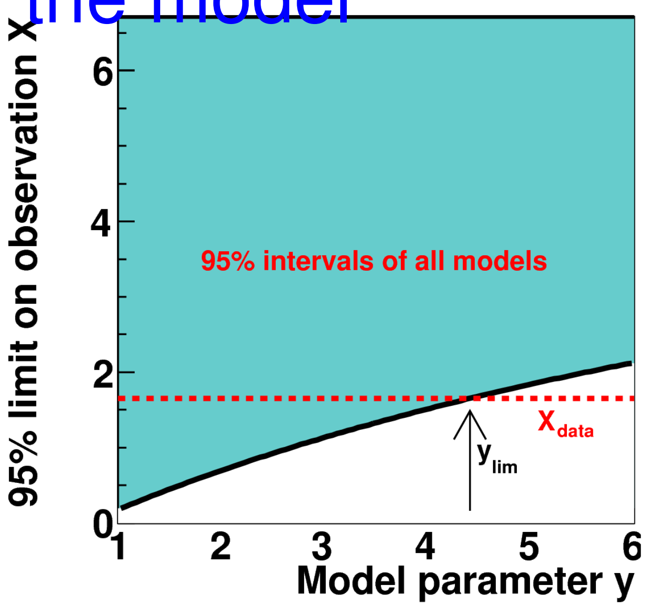
Search for Single Top Quark Production at HERA, Phys.Lett.B678:450, 2009



Limit calculation on one slide

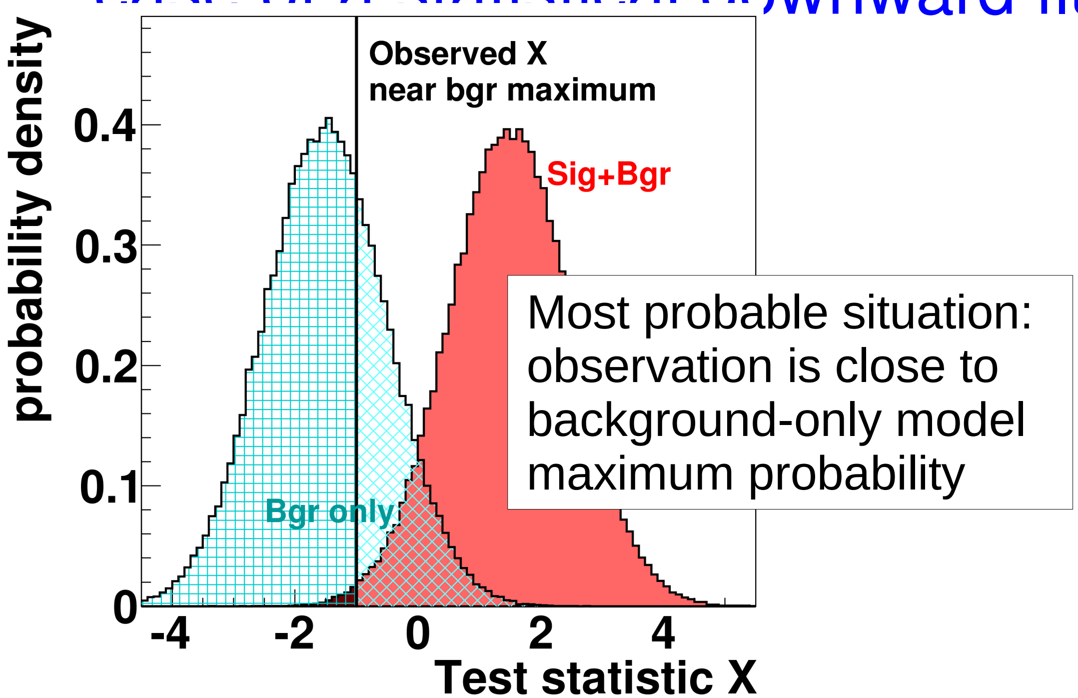
- Combine information from all bins in one variable X, robust against systematic uncertainties (e.g. profile likelihood)
- Sample the probability density of X for a given model using toy experiments
- If X(data) is outside the 95% interval, exclude the model
- Simple case: one model parameter, monotonic relation to X

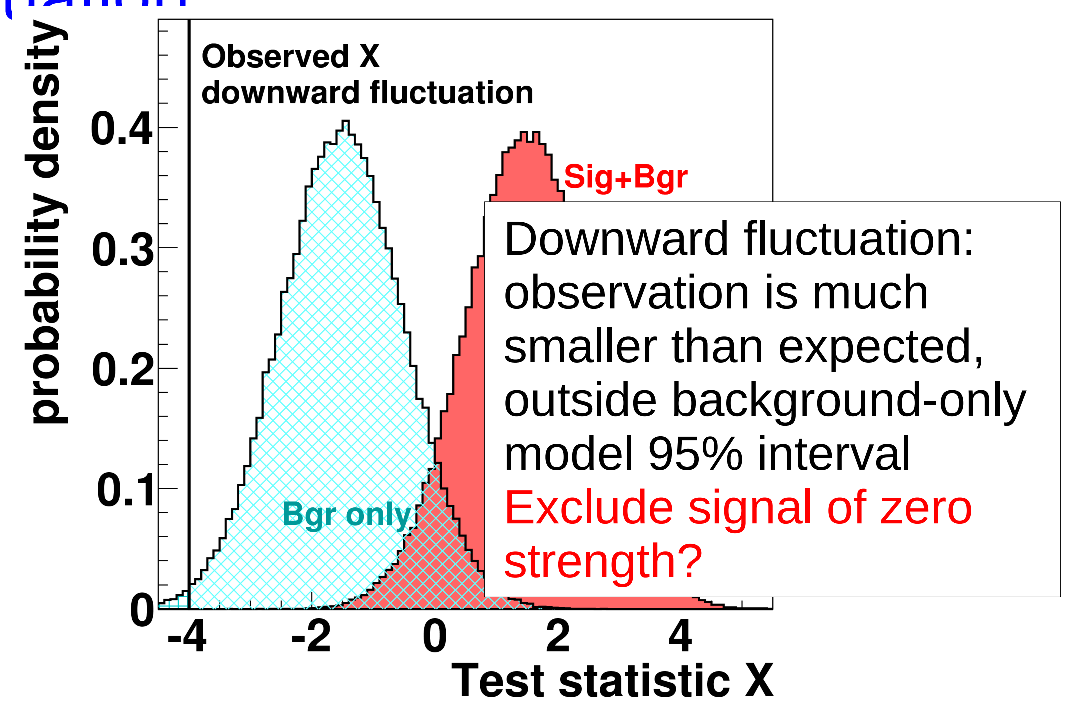




Problems with downward fluctuations

- “Standard” search: signal of unknown strength on top of background
- One model parameter is the signal cross section (or a coupling)
- Observable X rises with increasing number of signal events
- Problem: possible exclusion of background only model (signal=0) in case of a statistical downward fluctuation



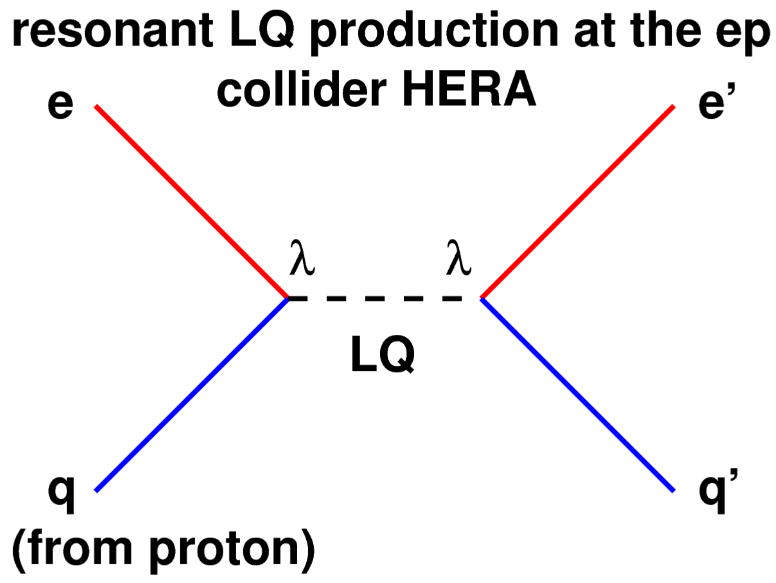


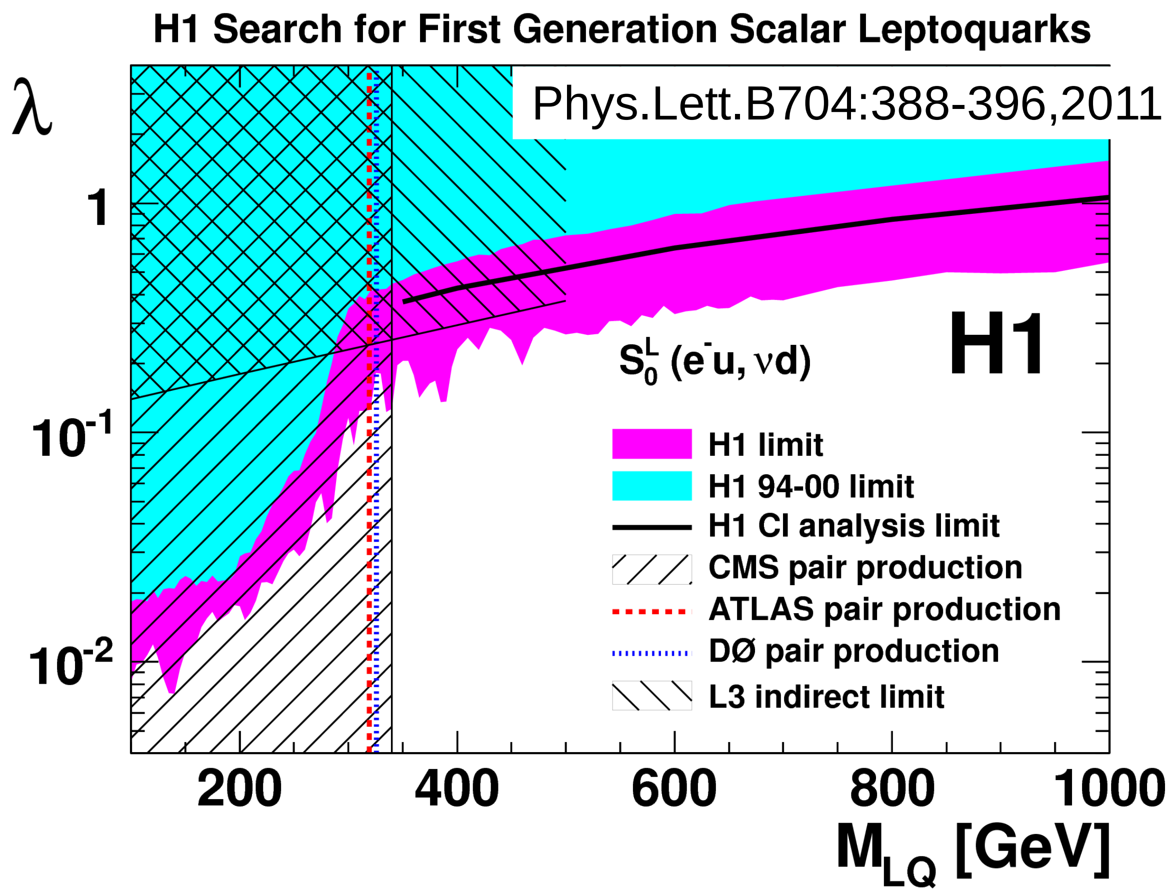
Treatment of downward fluctuations

- Problem: statistical downward fluctuations may result in a limit much better than the experimental sensitivity
- Two basic solutions
  - Power constraints: quote expected limit in case of downward fluctuations
  - CL_S method: normalize probability to background-only probability
- In either case, the limits are somewhat more conservative as compared to the textbook Frequentist method
- Status: CL_S method is used for most HEP limits

Limit example 1: leptoquark search

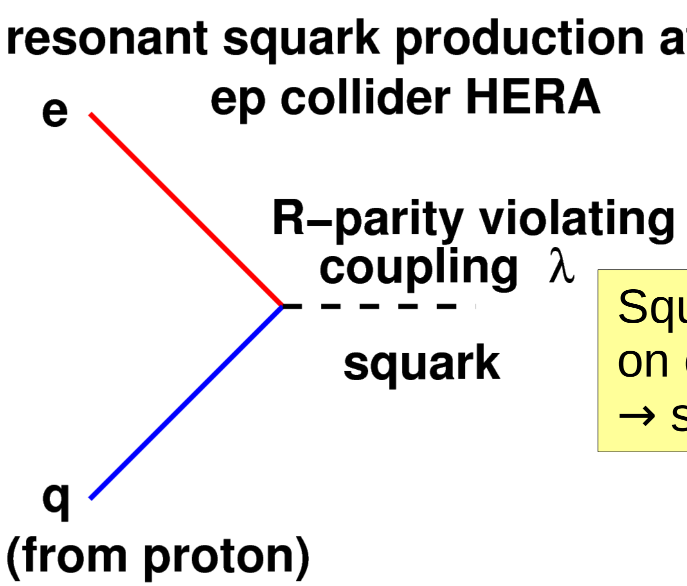
- Limits on one parameter are often shown as a function of another model parameter
- Here: leptoquark coupling (y-axis) and mass (x-axis)
- Colored regions are excluded



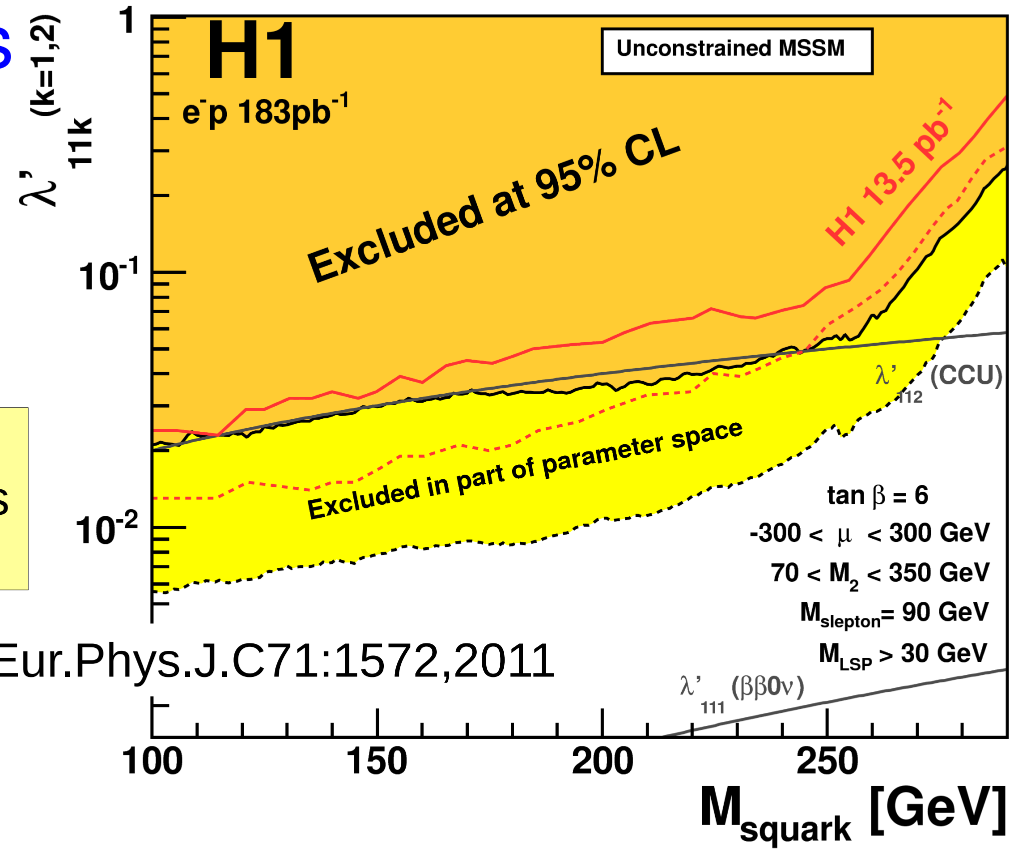


Limit example 2: R-parity violating SUSY

- Similar to Leptoquark analysis, coupling wrt Squark mass
- Scan of other (hidden) model parameters for each point
- Exclude a set of models if the (coupling,mass) is excluded for any setting of the other parameters



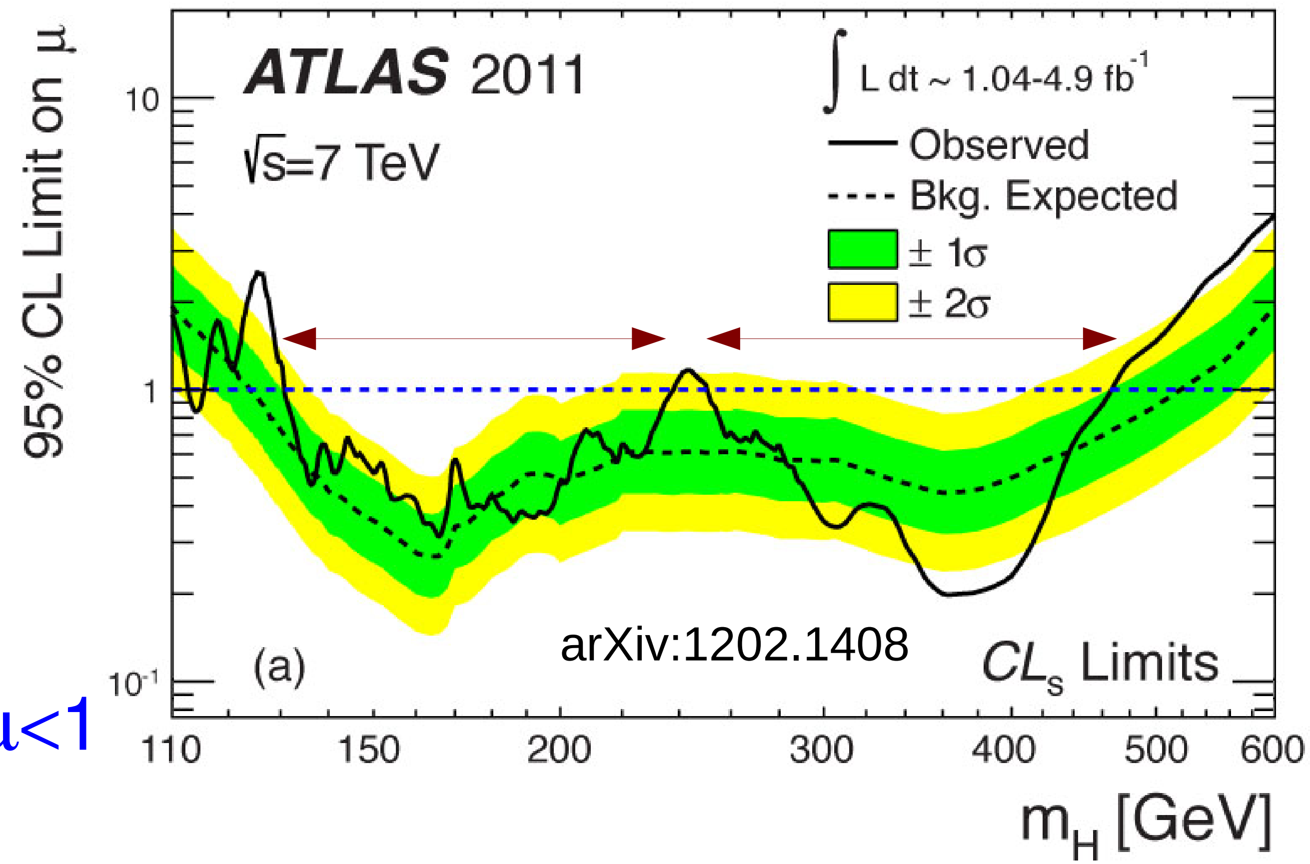
Squark decay depends on other SUSY parameters → scan



Eur.Phys.J.C71:1572,2011

Limit example 3: LHC SM Higgs search

- The model only has one unknown parameter, the Higgs mass
- A scaling factor μ on the Higgs cross section is used as a second parameter
- Standard model: $\mu=1$
- Solid line: observed limit, Region above the curve is excluded
  → SM Higgs is excluded where the solid line has $\mu<1$



Observed and expected limit

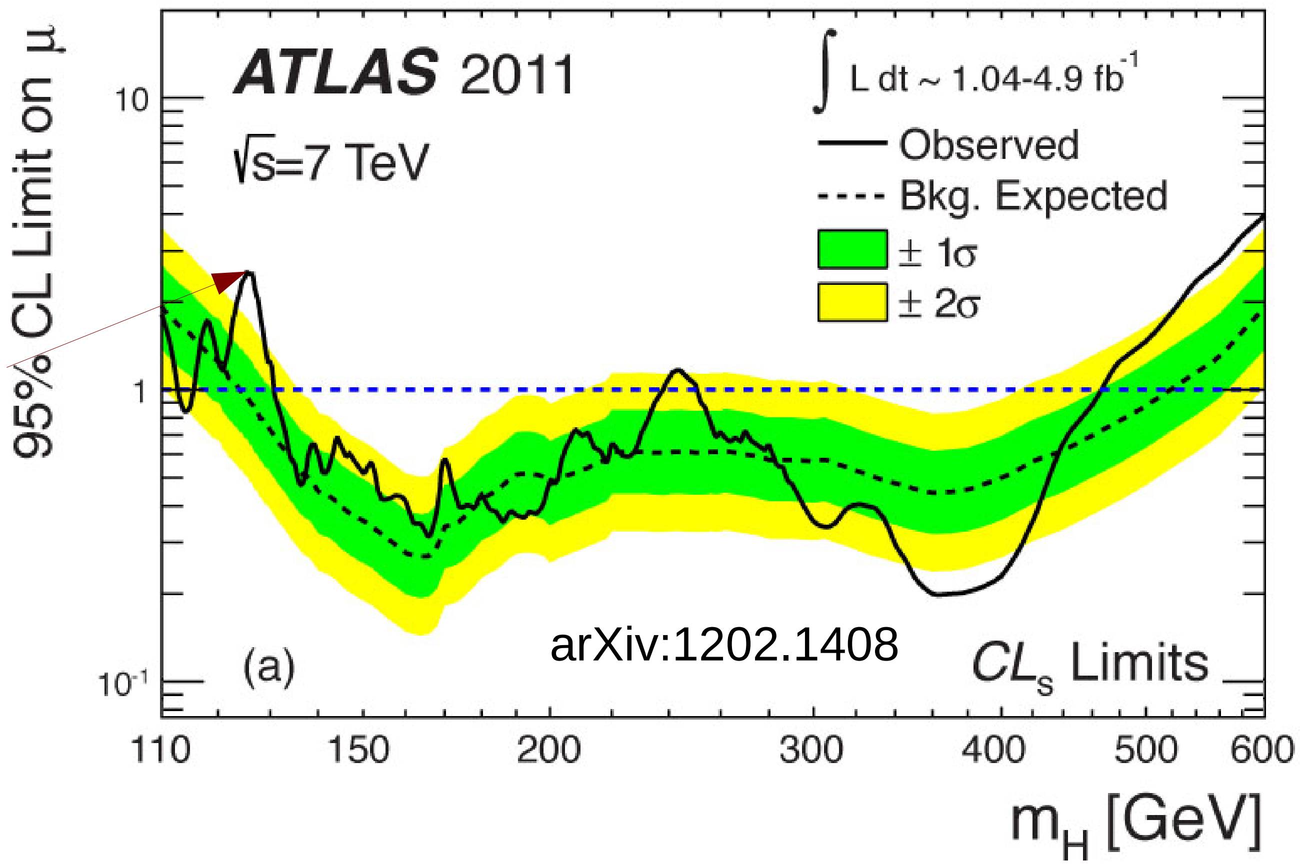
- Observed limit: compare data to models
- Expected limit: compare background-only model to other models
- Quantifies the sensitivity of the experiment
- Also useful: show 68% or 95% region of expected limits
- What if the observed limit deviates from the expected limit

  (a) the experiment or the background model are not understood well

  (b) there are candidate events in favour of the signal

Limit example 3: LHC SM Higgs search

- The model only has one unknown parameter, the Higgs mass
- A scaling factor μ on the Higgs cross section is used as a second parameter
- Green, yellow: 1,2 σ around expected limit
- Solid line above yellow line → Higgs candidates



Summary/Outlook

- Limits in experimental particle physics have been discussed
  - Frequentist definition is used in most cases
  - Calculation using toy experiments
  - Limits are usually set on model parameters, 2D plot of two parameters
  - For more than two parameters, the other parameters are either fixed or scanned
  - Expected limit shows the experiment's sensitivity
- Not covered in this talk
  - Technical details (profile likelihood, power constraints, CL_S, etc)
  - Bayesian methods
  - ...