# From LHC results to BSM fits

## Gfitter SUSY Results

## Experience of working with RooFit/RooStats "Workspaces"

Max Baak (CERN)
on behalf of the Gfitter group (*)

LHCC: Global BSM fits and LHC data
11 February 2011, CERN


http://cern.ch/Gfitter

(*) M.B. (CERN), H. Flächer (CERN), M. Goebel (Univ. Hamburg, DESY), J. Haller (Goettingen), A. Höcker (CERN), D. Ludwig (Univ. Hamburg, DESY), K. Mönig (DESY), M. Schott (CERN), J. Stelzer (DESY)

# Overview: From LHC data to BSM fits

Yesterday

- Gfitter results obtained sofar with global electroweak fit (Johannes Haller)
  - SM fit, BSM constraints through STU formalism

Today

- Minimal SuperGravity results by Gfitter

- Focus of this presentation:
  Inclusion of *current* LHC data via RooStats workspaces
  - Step-by-step practicalities of using workspaces from/in SUSY environment

# Context

- As member of ATLAS:
  - Promotion of the RooWorkspace concept in ATLAS SUSY group.
  - Actively involved in the search for SUSY of 1-lepton + jets + MET
  - Combination of different search channels
  - Setting of exclusion / discovery limits in SUSY mass plane
- As member of Gfitter:
  - Ensure that results are easily useable when made public to outside groups
  - Use of these results in combined BSM fit

# Some quotes from yesterday

- "Fit machinery is ready to perform combined fits based on actually observed LHC data. But we need the data in appropriate form from the LHC collaborations."
  – Peter Wienemann

- "Include LHC results on SUSY searches in the likelihood f 'n:
  - depends on what CMS and ATLAS will publish
  - employ workspace of RooFit/RooStats?

  " – Leszek Roszkowski

# The Gfitter Project


## A Generic Fitter Project for HEP Model Testing

- Gfitter = state-of-the-art model testing tool for LHC era
- Gfitter software
  - Modular, object-oriented C++, relying on ROOT, XML, python
  - Core package with data-handling, fitting, and statistics tools
  - Independent "plug-in" physics libraries: SM, 2HDM, oblique parameters, SUSY, ...
- Gfitter features
  - Various fitting tools: Minuit (1/2), Genetic Algorithms, Simulated Annealing, etc.
  - Full statistics analysis: goodness-of-fit, p-values, parameter scans, MC analyses.
  - Consistent treatment of statistical, systematic, theoretical uncertainties (Rfit prescription), correlations, and inter-parameter dependencies.
    - Theoretical uncertainties included in $\chi^2$ with flat likelihood in allowed ranges
- Main publication: EPJ C60, 543-583,2009 [arXiv:0811.0009]
- Also available at: www.cern.ch/Gfitter
- Today: updates and new results

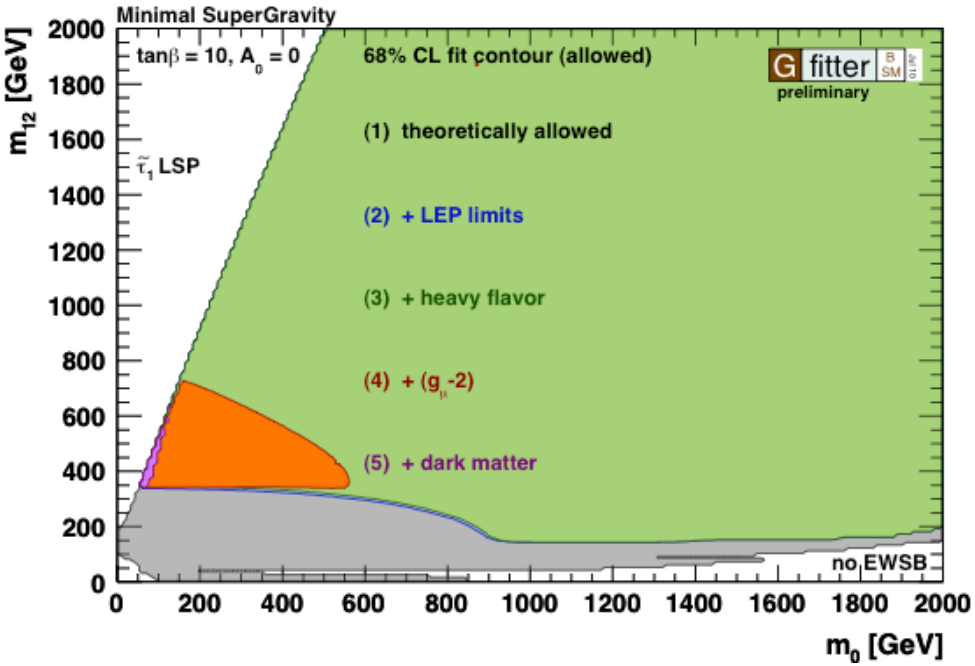
# Minimal SuperGravity

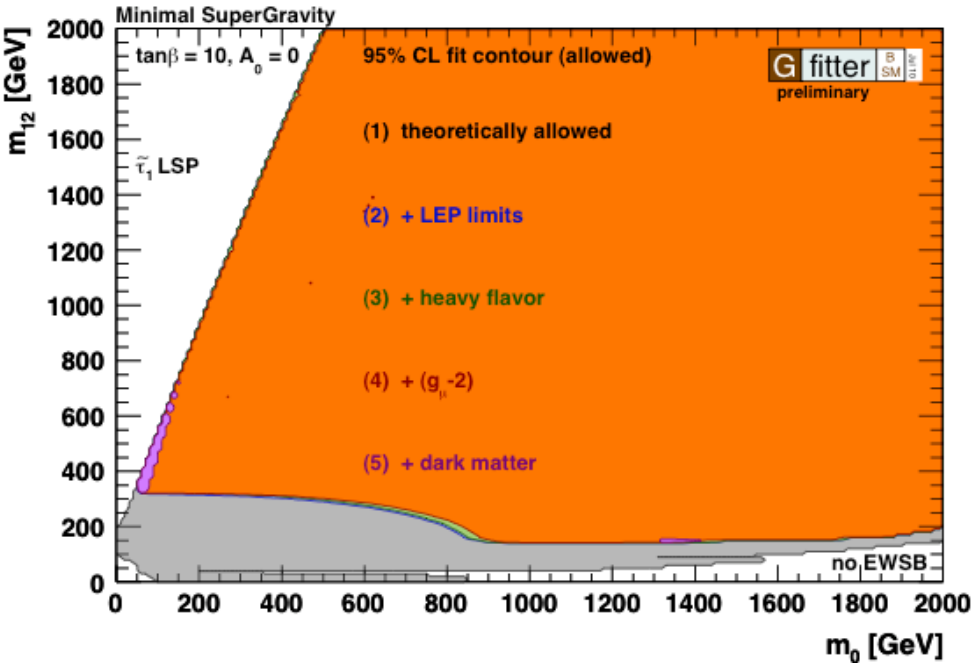
- Supersymmetry may solve many shortcomings of the SM (hierarchy problem, unification of coupling constants, DM candidate)
- mSUGRA: highly constraining breaking mechanism at GUT scale
- Breaking mediated by gravitational interaction

- Simplified model sofar tested in Gfitter: mSUGRA
- Determined by 5 parameters
  - $m_{1/2}$: mass of fermions at GUT scale
  - $m_0$ : mass of scalar particles at GUT scale
  - $\tan\beta$ : ratio of two Higgs VEVs
  - $A_0$: tri-linear coupling of the Higgs
  - sgn( $\mu$ ): sign of Higgsino mass term

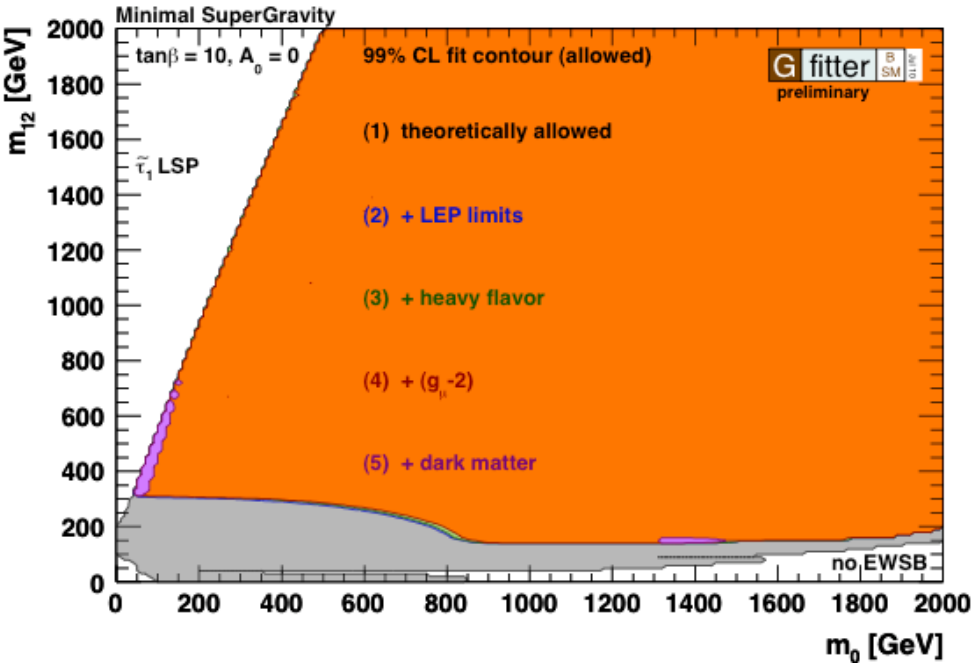
SoftSUSY – B.C. Allanach, Comp. Phys. Com. 143 (2002) 305-331
FeynHiggs – M. Frank et al JHEP0702, 047 (2007)
SuperIso – F. Mahmoudi, JHEP12, 026 (2007)
migroMegas – G. Belanger et al, IRFU-10-24, LAPTH-012-10

# Constraints on mSUGRA 1

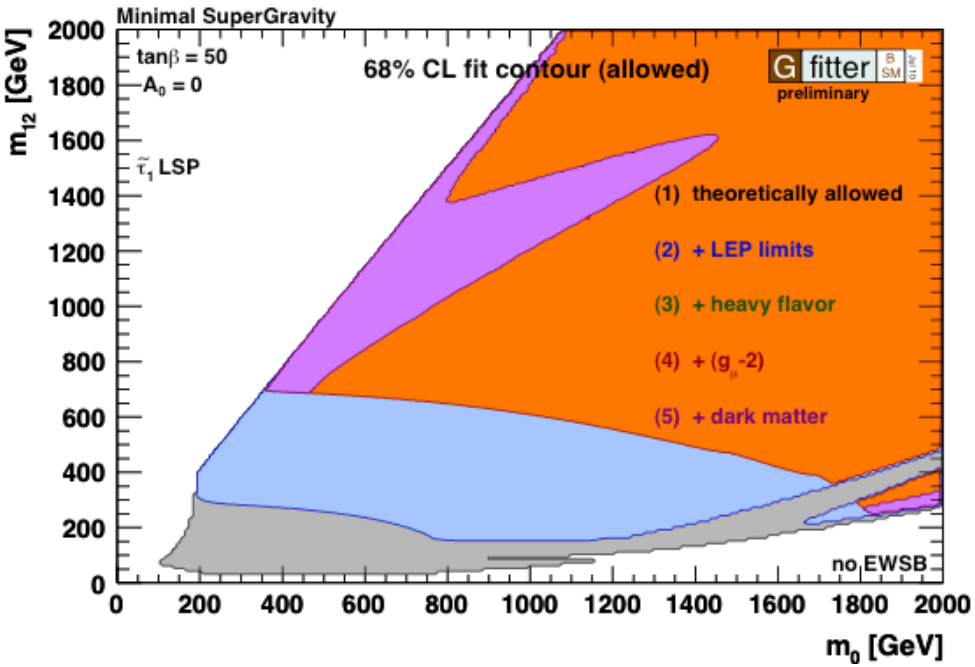


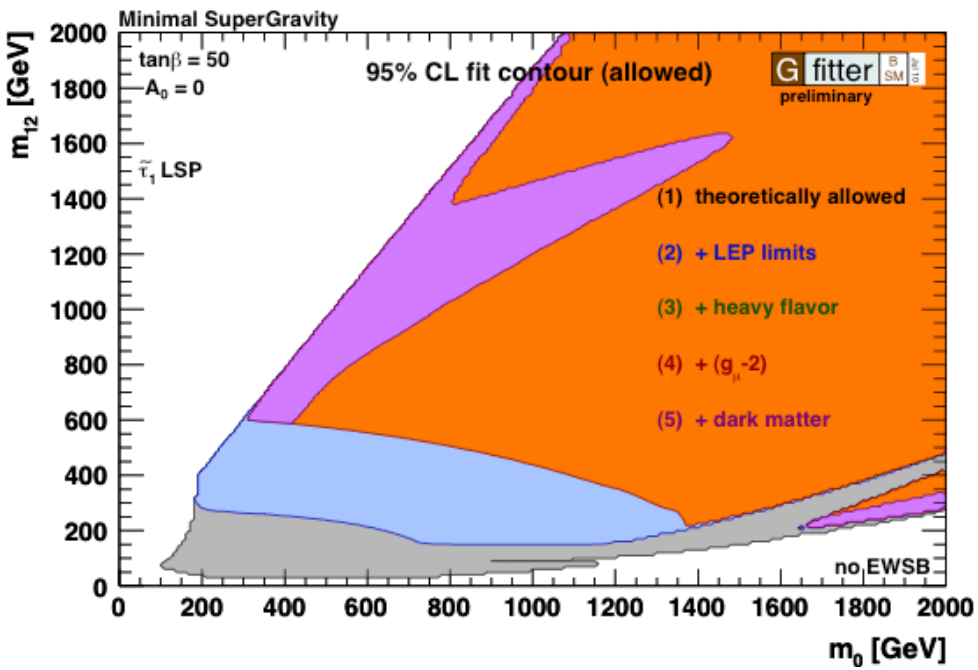




- Limits on m0, m12 by including in fit
  - Theoretical limits, e.g. by requiring a non-charged LSP
  - LEP limits on Higgs, neutralino, slepton masses
  - Constraints from heavy flavor physics
  - Constraint from $g_\mu$-2
  - From the relic density

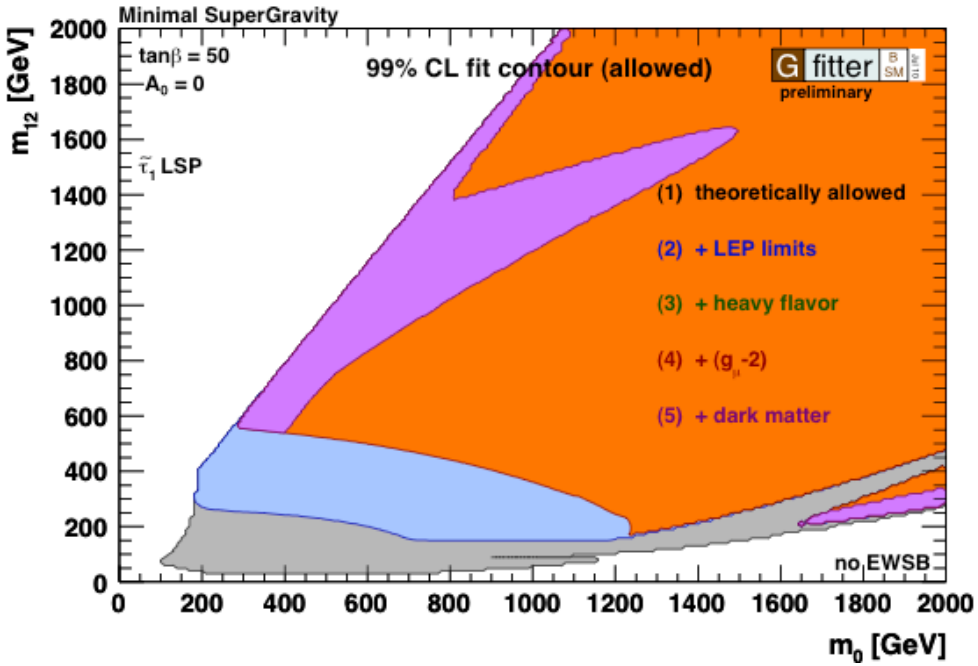


# Constraints on mSUGRA 2

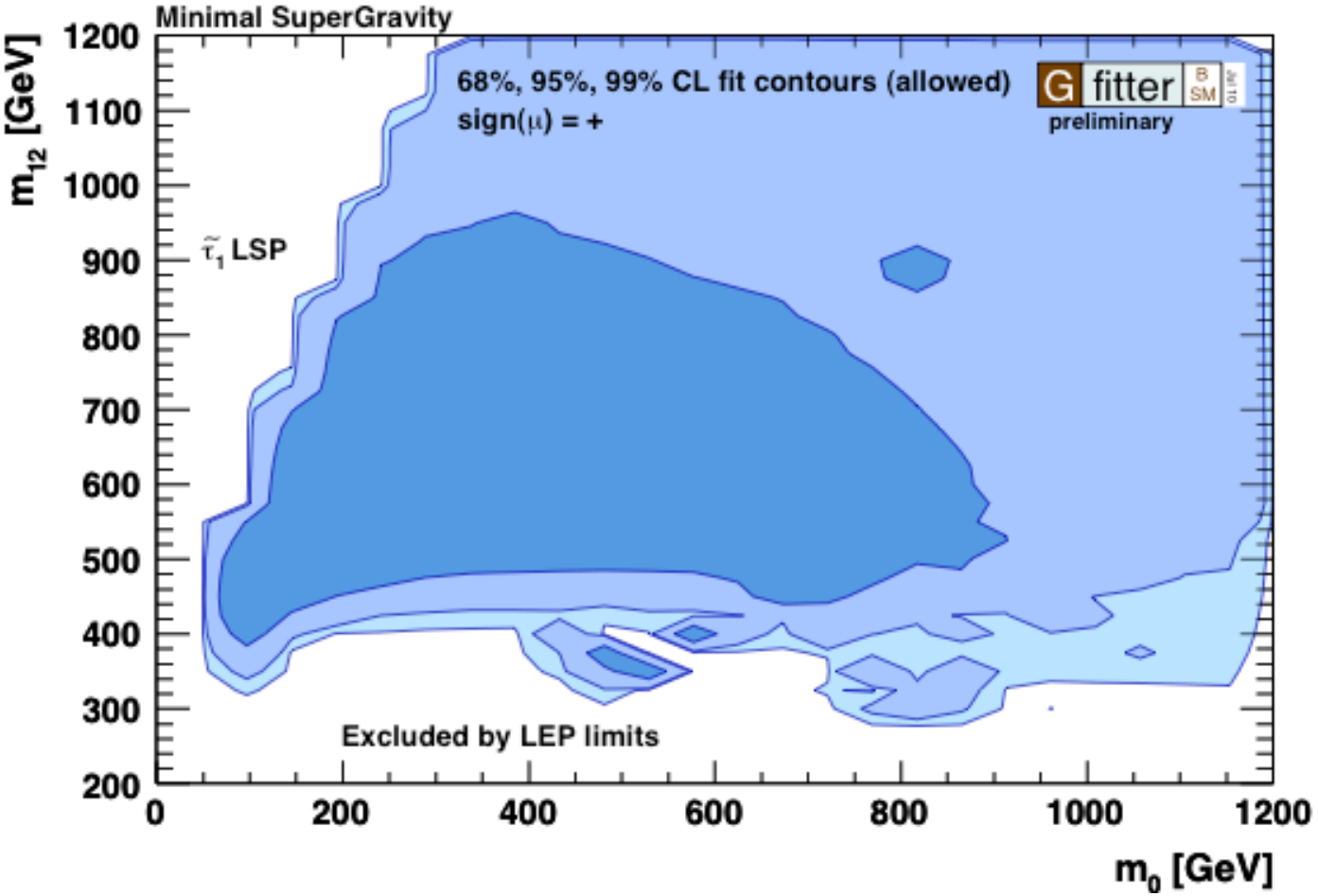



- Limits on m0, m12 by including in fit
  - Theoretical limits, e.g. by requiring a non-charged LSP
  - LEP limits on Higgs, neutralino, slepton masses
  - Constraints from heavy flavor physics
  - Constraint from $g_\mu$-2
  - From the relic density

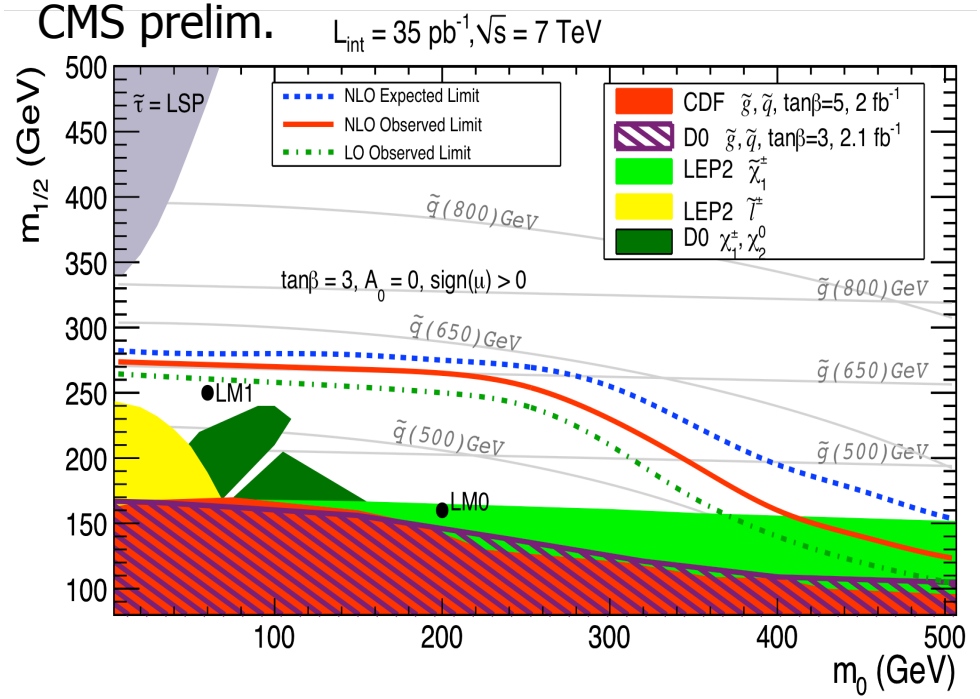


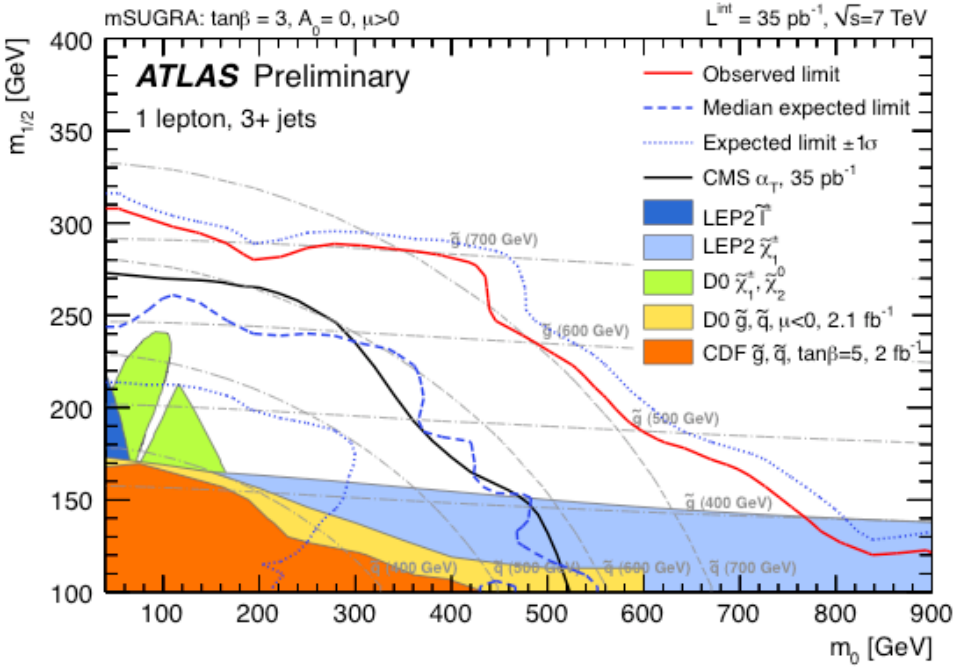
# Constraints on mSUGRA 3




- tan $\beta$ and $A_0$ are floating parameters
- Constraints: LEP limits, heavy flavor, $g_\mu$-2, dark matter density
  - No EW precision observables included
- 68%, 95%, 99% CL fit contours of allowed region are shown

# LHC preliminary results

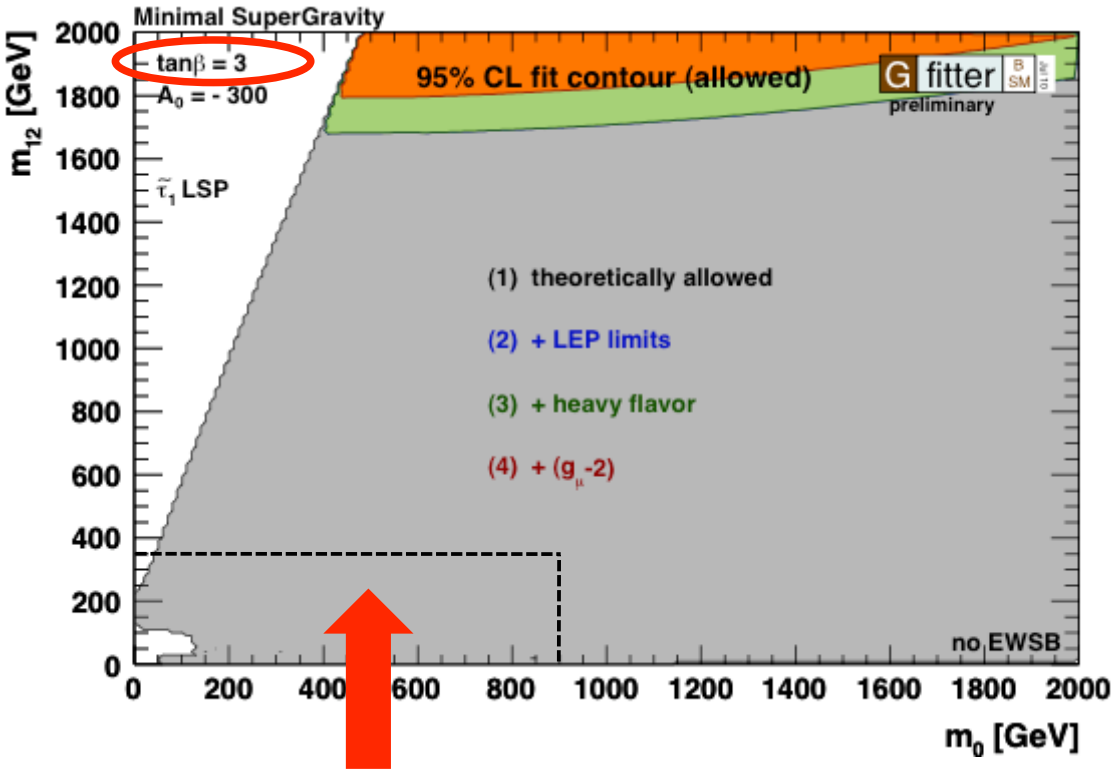



- ATLAS and CMS gentlemen's agreement: tan $\beta$ =3 and $A_0$ =0
- Wish to use the results from LHC that are coming out now!
  - Not 35/fb, but 35/pb.
  - Combine these results
  - Apply (combined) constraint in (mSUGRA) BSM fit
- How to do this exactly? (topic of this presentation)

# LHC Results (in perspective)

- LHC results presented sofar have only moderate dependence on tan $\beta$
  - Very small dependence in production cross-section, some in decay topology
- (Good news, because in mSUGRA tan $\beta$ =3 is already excluded by LEP Higgs limit.)




Region covered by ATLAS & CMS

- To use the current LHC results in mSUGRA fit, need to:
  - Either, assume no dependence on tan $\beta$ and $A_0$
  - Or, calculate dependence yourself
  - Or, experiments present several results for several tan $\beta$ values.

# Combination of analyses

- ATLAS
  - All SUSY analyses done with RooFit / RooStats workspaces
- CMS
  - (I know that) 0-lepton + jets ($\alpha_T$) based on workspaces
- Combination of ATLAS analyses is straight-forward technically, including all correlated systematics.
- Confident this is also relatively simple between ATLAS and CMS.

# SUSY Combination Example

Workspace for 0 lepton channel

Individual constraints on x-sec, lumi, JES

Workspace for 1 muon channel

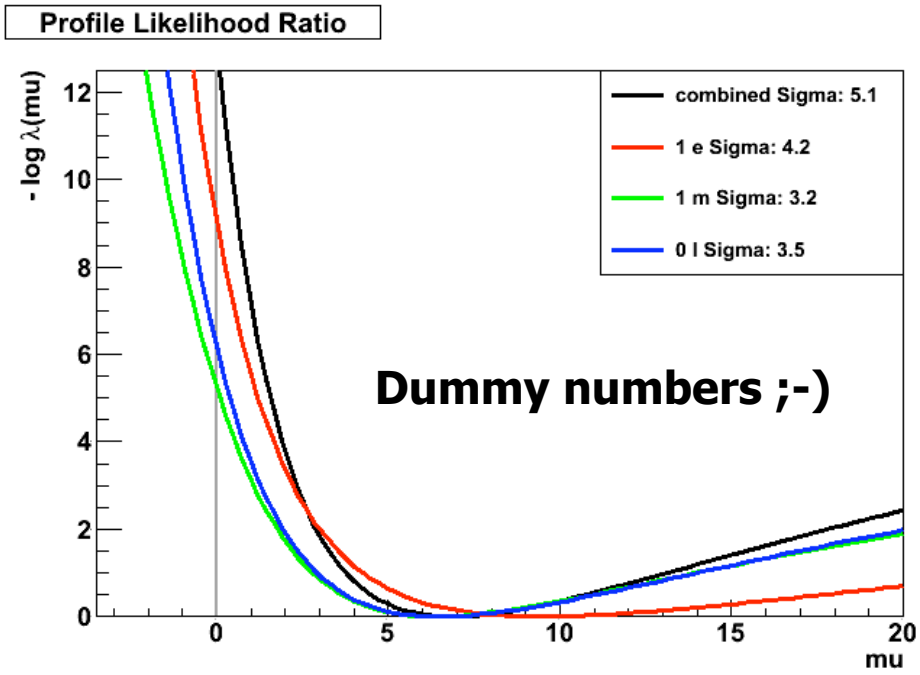
Individual constraints on x-sec, lumi, JES

Workspace for 1 electron channel

Individual constraints on x-sec, lumi, JES

- Identify common model parameters (x-section uncertainty, JES, lumi, etc) between channels

- Remove and replace common/ correlated constraints in final pdf
  - Eg. don't apply luminosity constraint twice

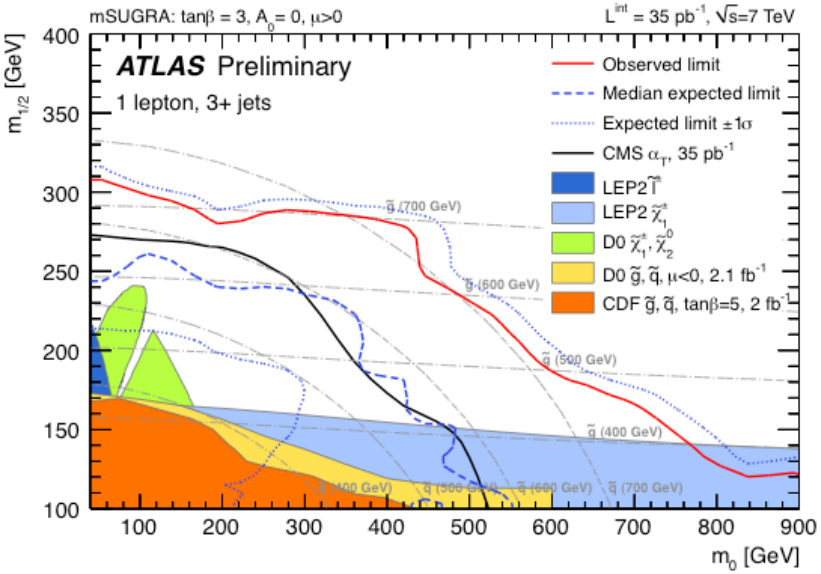


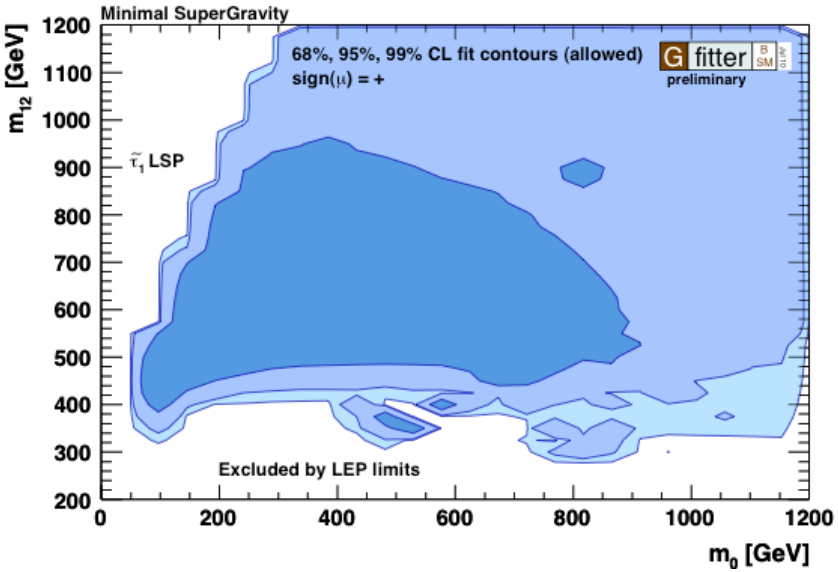

Create final pdf in combined workspace, before passing on to RooStats

- Straight-forward exercise with RooFit/RooStats !

# Let's connect these two ;-)

- Left: ATLAS / CMS result of counting experiment
- Right: combined constraints on mSugra




- Use full information: central result + full error
  - Ie. Full likelihood needed, not just 95% CL exclusion bounds
- Obtained by using workspaces!

# Example counting experiment

- Workspace: Toy model, "inspired" by ATLAS 1-lepton analysis results
  - "Cut-and-count" analysis.

- Enriched signal region
  - Observed number of events : 2 events
  - Expected number of background events (from control samples) : 2 ± 1 events
    - ➢ Consistent with no signal. This gives constraint in NP
  - Signal region described by Poisson distribution, uncertainties by Gaussians

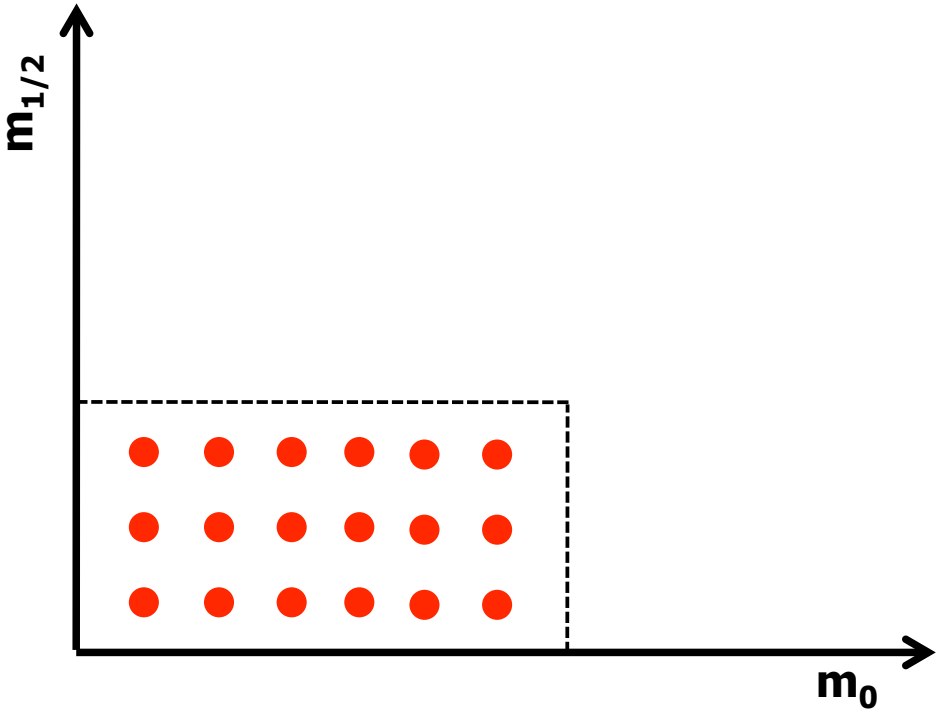
- Signal expectation from fully simulated MC
- … according to a simplified model
  - Currently mSUGRA for both ATLAS and CMS. But can be any model!

- 3 nuisance parameters: describe (un-)correlated uncertainties on background and signal.
- 3 simplified-model parameters, to describe signal:
  - In this toy example: $m_0$, $m_{1/2}$, $\tan\beta$

- Realistic workspace: setup bit simpler, but very similar to that used in ATLAS

# Simulation

- Only limited number of mass points generated by simulation
  - Full simulation = expensive
- Usual: 2d or 3d grid planes.
- Generated parameter values of chosen simplified model.
  - mSugra: $m_0$, $m_{1/2}$, tan beta
- Eg. Alternative models
  - m(squark) vs m(gluino)
  - m(gluino) vs m(LSP)
- In range of expected experimental sensitivity
- Possibly irregular grid spacing
  - ATLAS has combination of fine grid and coarse grid.

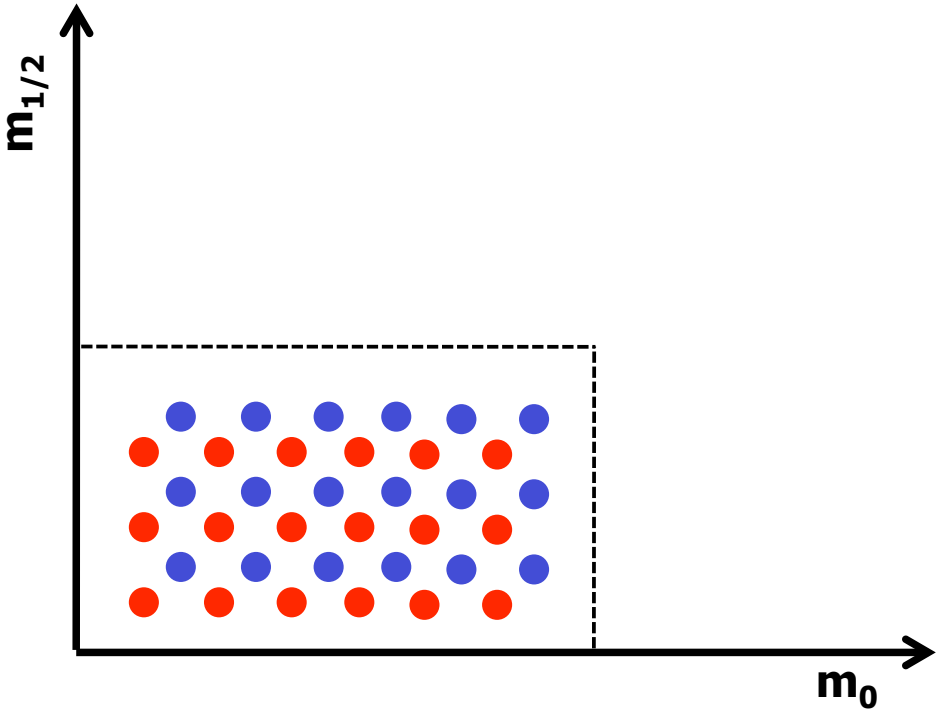


# Interpolation

- Grid only has discrete parameter values
- Whereas: continuous parameter used in simplified model

- Experiments may have non-overlapping grids
  - Eg. ATLAS and CMS both have sgn( $\mu$ )=+, tan $\beta$ = 3 and $A_0$ = 0
  - But many non-overlapping points
  - Often no direct averaging possible!

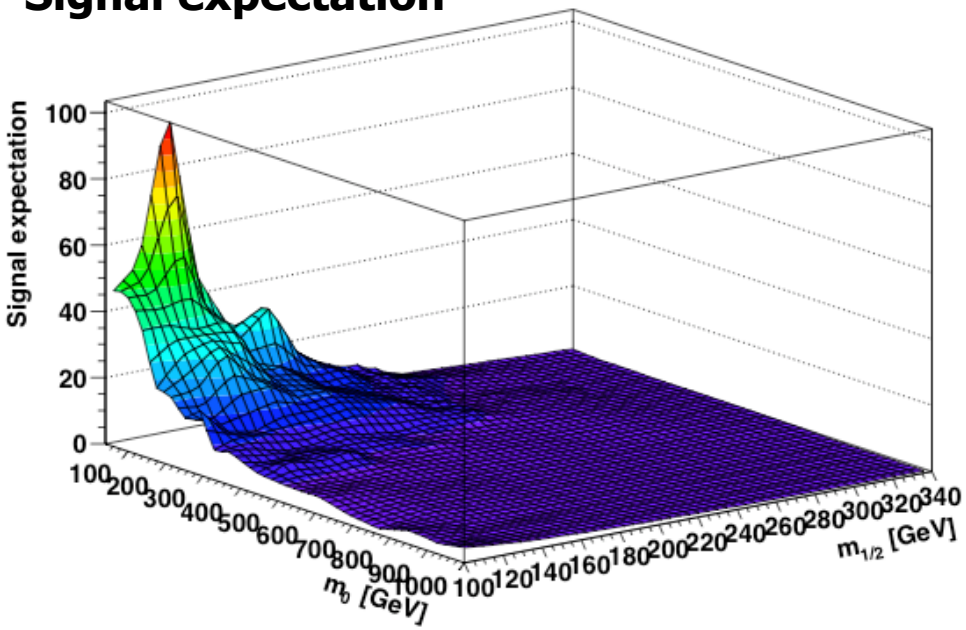
- Interpolation is needed



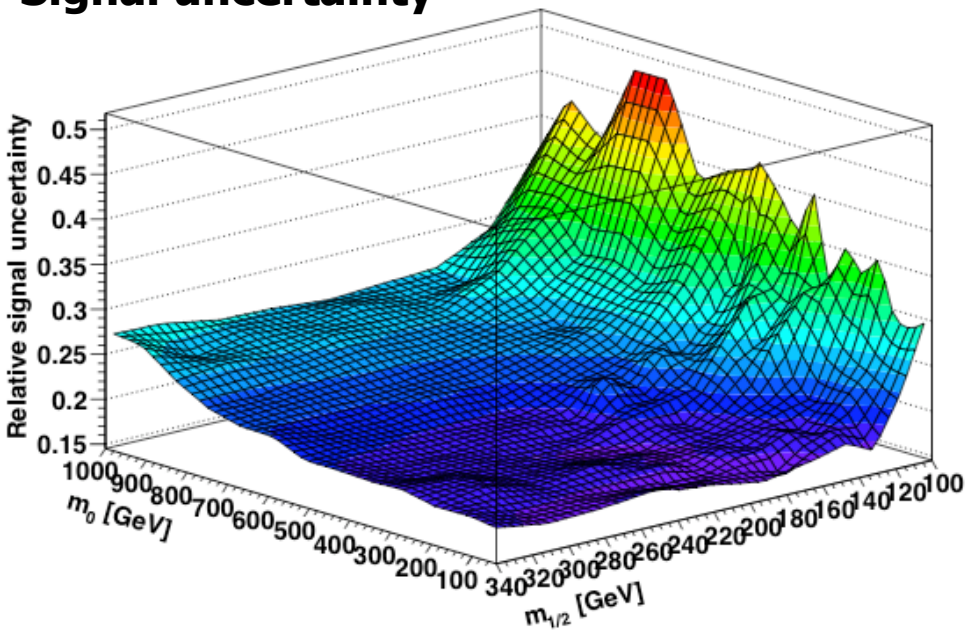
# Interpolation examples

- Linear, non-linear, smoothed interpolation?
  - Non-trivial with >1 interpolation parameters
  - This will make a difference!
- Interpolation in what variable?
  - Signal expectation? Only works if background is constant over the grid.
    - ➢ Selections may change over grid to favor certain decay topologies
  - Likelihood?
- RooFit / RooStats allows for 3-dimensional interpolation
  - Can be extended for N-dim.
  - But: probably don't want simplified models with more than 3 pars.

**Signal expectation**



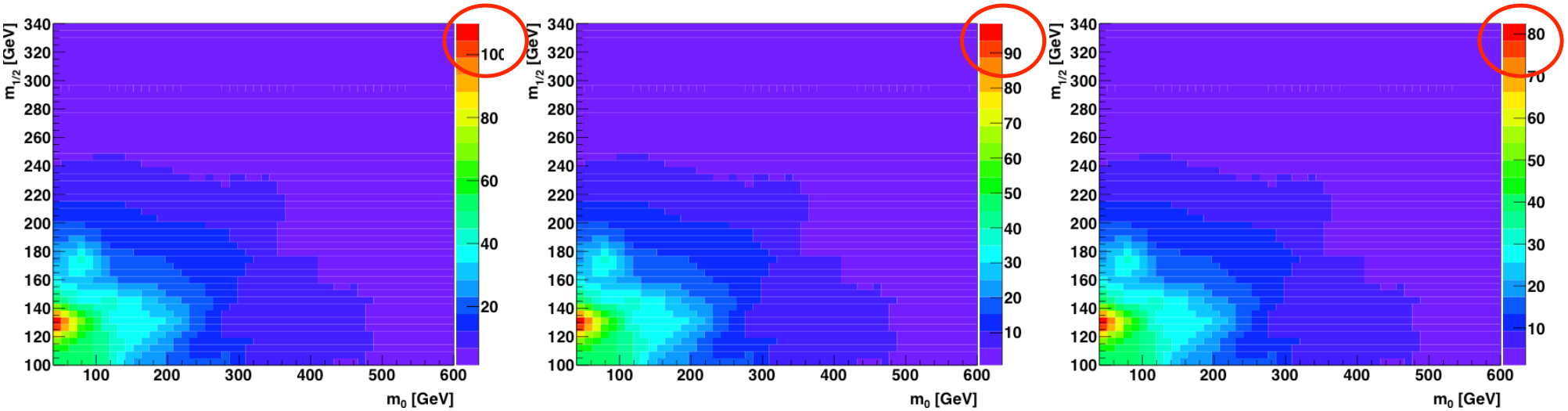

**Signal uncertainty**



# Interpolation in 3rd parameter: tanβ

- Signal expectation for different tan $\beta$ values
- Ie. Interpolation in three variables

**tanβ=3** **tanβ=10** **tanβ=40**

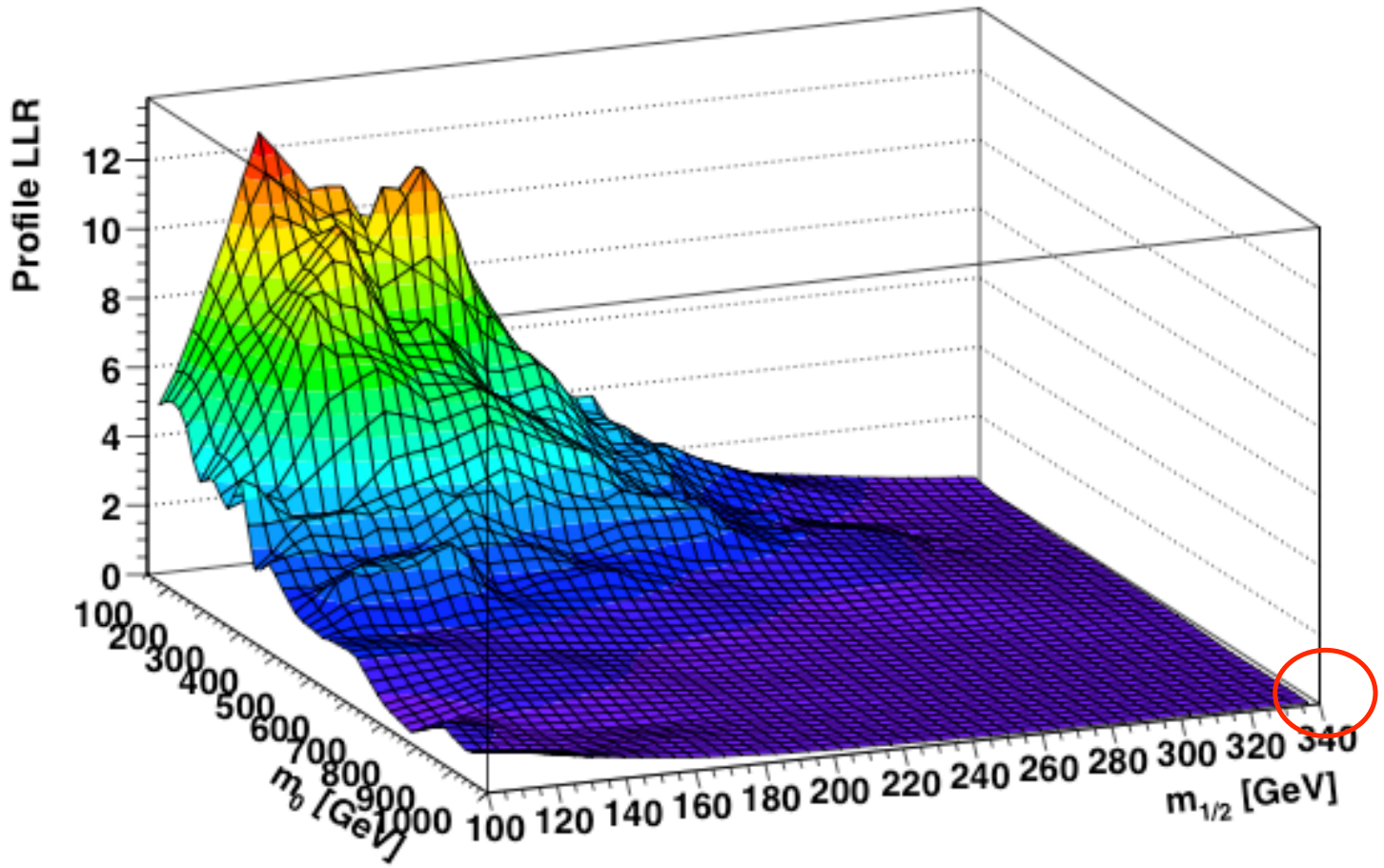



- Note that normalization changes for different tan $\beta$ values.
- Remember, these are just toy values …

# Fit example

- With interpolation, can now directly fit for simplified-model parameters! Here: $m_0$, $m_{12}$, $\tan\beta$
- Eg. global fit minimum at $m_0$=1000, $m_{1/2}$ = 340 GeV, as expected

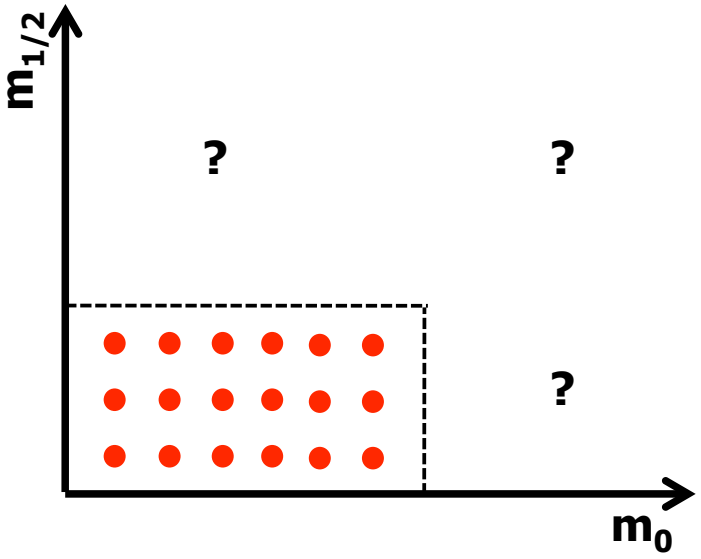
**Profile likelihood curve from workspace**

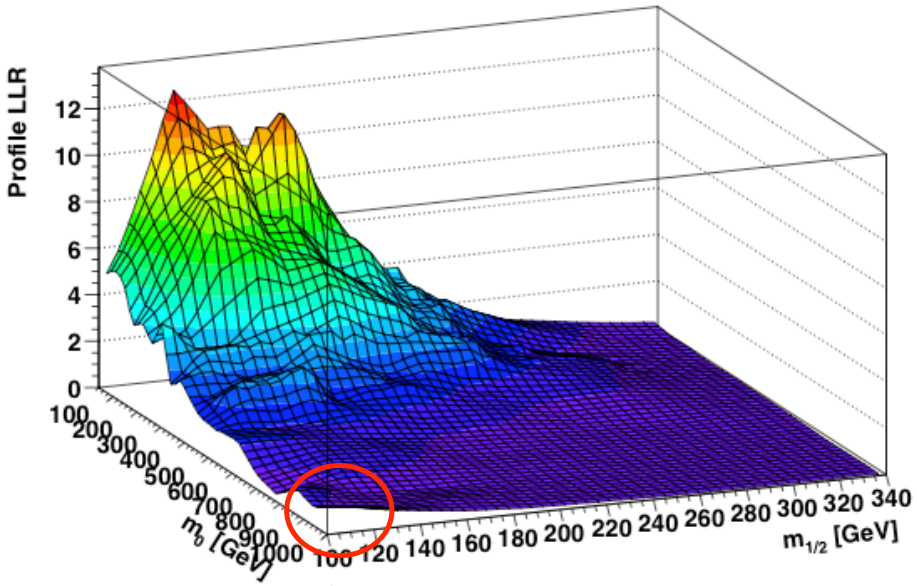



**Preferred region in fit has expectation of 0 signal events**

# Extrapolation out of the box?

- Experiments only generate MC grid in limited sensitivity region.
- Likely smaller than full range of interest for model parameters.
- What about extrapolation to values outside the grid?
- Extrapolation = unreliable. Should be prevented!
  - Workspace prediction should be cut-off outside parameter range.
    - ➢ But how exactly?




**Sensitivity not yet zero at boundary.**

# Remarks on statistics

- Published workspace gives a lot of analysis information, but workspace alone not is not sufficient.

For example: Published exclusion limits

- Extensive statistical coverage tests normally done by experiments
  - Missing from published log likelihood ratio.
- Experiments also publish confidence levels! In this case, it is unclear how to combine them with the indirect constraints.
- For example: Use of published log likelihood ratio directly in fit may well give undercoverage or overcoverage.
  - Eg. LLR known to provide overcoverage in case of low background expectation.

- Good news is: full information to do this coverage study available from workspace

# Treatment of nuisance parameters

- Three nuisance parameters present in RooWorkspace example
  - Describe correlated and uncorrelated uncertainties on signal and background

- Two possible choices:

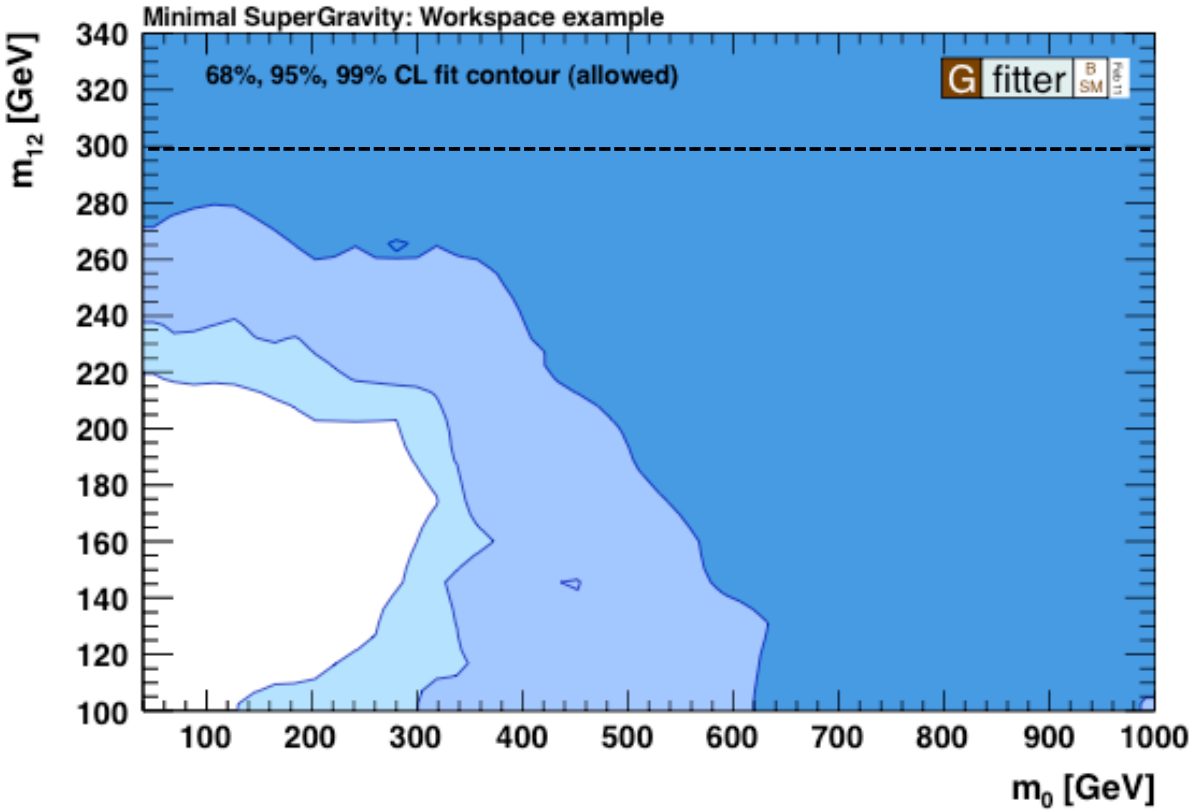
  1. (Profile) log likelihood ratio: Nuisance parameters profiled out by experiments
     - ➢ BSM fit focuses fully on simplified-model parameters
  2. Use likelihood directly: Nuisance parameters treated by BSM fit

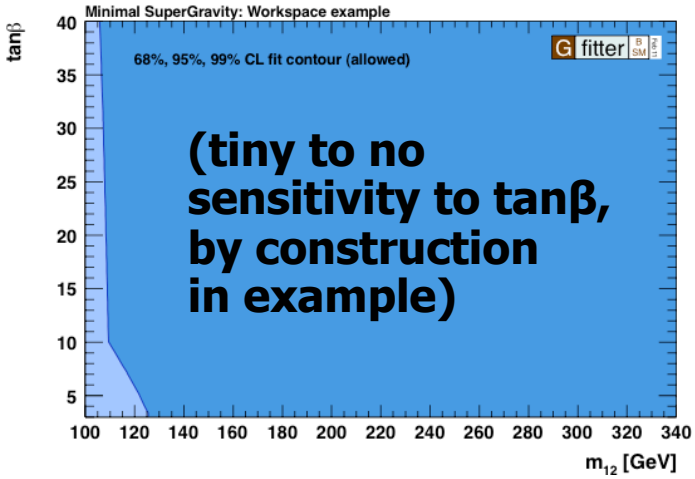
- Gfitter: R-fit approach
  - Uniform constraints on nuis. Parameters, free within ±1 sigma.

- In this example: profile log likelihood ratio from workspace added directly to $\chi^2$ minimized by Gfitter.
  - Only parameters considered are simplified-model parameters.

# Implementation in Gfitter

- Implementation in Gfitter relatively straight-forward
  - (Possible after browsing through a few instructional examples.)




- (Sorry, combined constraint not finished in time …)
  - At this stage straight-forward to complete.
  - Example does not yet affect current favored mSUGRA region in $m_0$-$m_{12}$ plane.

# Conclusions 1

- Workspaces are proven, great tool for combination of different analyses and different experiments.
  - Already (much) used in ATLAS and CMS experiments
- Will most certainly be used for SUSY combination between experiments
- Now that SUSY are starting to come out, (ATLAS) workspaces will hopefully be public soon.
  - Discussion on digital publication of workspaces has started. Decision to be taken soon (needs to be taken by management).
- Though necessary tool, workspace is not sufficient.
  - Statistical interpretation is missing, though should be reproducable.
  - A published analysis (= workspace) remains model-dependent.

# Conclusions 2

- All ingredients available in RooFit / RooStats that allow the presentation of full results of *simplified models* by ATLAS or CMS to outside fitter groups.
- Straight-forward to integrate into existing Gfitter mSUGRA fit.

- Some practical details to be considered
  - Interpolation needed for combination of channels, and turning discrete grids into continuous distributions.
    - ➢ Solved technically.
  - "Model-parameters out of boundary" effects.
    - ➢ Set likelihood ratio to zero outside box?

- Test workspace with ATLAS-like result is available at:
  - https://twiki.cern.ch/twiki/bin/view/RooStats/BenchmarkWorkspaces
  - Also available: Correlated edge-point measurements in workspace form
    - ➢ by Kyle
- Please try out and provide feedback!