The pMSSM likelihood approach to CMS data and implications thereof

J. Gunion (UC Davis) for the CMS Collaboration
with S Bein (FSU), S. Kraml (LPSC Grenoble), H.B. Prosper (FSU),
S. Sekmen (CERN), L. Vanelderen (UHH), W. Waltenberger (HEPHY)

SUSY at the Near Energy Frontier, November 10, 2013

The pMSSM approach:

- Addresses limitations of particular analyses, e.g. SMS scenarios.
- Tests SUSY at the Electroweak scale, avoids GUT-scale prejudices re signatures.
  $\Rightarrow$ Who knows how and at what scale Beyond-MSSM Physics enters.

Talk based on SUS-12-030 and already public results to be found in future SUS-13-020.
Special thanks to S. Sekmen for rapid response to my 'theory' plot requests.

The p(henomenological)MSSM

The pMSSM is a 19-dimensional parameterization of the 124 parameter MSSM defined at the SUSY scale $m_{\rm SUSY} = \sqrt{m_{\tilde{t}_1} m_{\tilde{t}_2}}$.

- Basic ingredients:
  - Assumes R-parity conservation
  - no new sources of CP violation
  - no flavor changing neutral currents
  - 1st and 2nd generation sfermion are assumed degenerate in mass and their A terms are assumed to be of negligible importance
  - sfermion mass matrices and trilinear couplings are assumed flavor-diagonal
  - the lightests neutralino is assumed to be the LSP

- This leaves 19 SUSY parameters
  - M_1, M_2, M_3 — no GUT relations assumed.
  - $\tan\beta$, μ (higgsino mass parameter), m_A (mass of CP-odd Higgs).
  - 10 sfermon mass parameters, $m_{\tilde{Q}_1} = m_{\tilde{Q}_2}$, $m_{\tilde{U}_1} = m_{\tilde{U}_2}$, $m_{\tilde{D}_1} = m_{\tilde{D}_2}$, $m_{\tilde{L}_1} = m_{\tilde{L}_2}$, $m_{\tilde{E}_1} = m_{\tilde{E}_2}$, $m_{\tilde{Q}_3}$, $m_{\tilde{U}_3}$, $m_{\tilde{D}_3}$, $m_{\tilde{L}_3}$, $m_{\tilde{E}_3}$
  - A_t, A_b, A_τ

Construction of the pMSSM prior

The purpose of this study is to assess what current data tell us, and do not tell us, about the MSSM using the more tractable pMSSM as a proxy. Using these data we perform a global Bayesian analysis that yields posterior probability densities of model parameters, masses and observables. We work within the pMSSM sub-space,

$$
\begin{gathered}
-3\,\mathrm{TeV} \leq M_1, M_2 \leq 3\,\mathrm{TeV} \\
0 \leq M_3 \leq 3\,\mathrm{TeV} \\
-3\,\mathrm{TeV} \leq \mu \leq 3\,\mathrm{TeV} \\
0 \leq m_A \leq 3\,\mathrm{TeV} \\
2 \leq \tan\beta \leq 60 \\
0 \leq \tilde{Q}_{1,2}, \tilde{U}_{1,2}, \tilde{D}_{1,2}, \tilde{L}_{1,2}, \tilde{E}_{1,2}, \tilde{Q}_3, \tilde{U}_3, \tilde{D}_3, \tilde{L}_3, \tilde{E}_3 \leq 3\,\mathrm{TeV} \\
-7\,\mathrm{TeV} \leq A_t, A_b, A_\tau \leq 7\,\mathrm{TeV}, \qquad (1)
\end{gathered}
$$

A point in this space will be denoted by θ.

The posterior density of θ given data D is given by

$$p(\theta|D) \sim L(D|\theta)\, p(\theta), \tag{2}$$

where $L(D|\theta)$ is the likelihood and $p(\theta)$ is the prior probability density, or prior for short. Since we consider multiple independent measurements D_i, the combined likelihood is given by $L(D|\theta) = \prod_i L(D_i|\theta)$.

The prior encodes any knowledge we have about θ independent of the measurements D_i. It may for example encode information from measurements other than D_i or theoretical assumptions. The latter include the following.

1. The sparticle spectrum must be free of tachyons and cannot lead to color or charge breaking minima in the scalar potential.

2. We require that EWSB be consistent.

3. The Higgs potential must be bounded from below.

4. Finally, in this study, we also require that the LSP be the lightest neutralino, $\widetilde{\chi}_1^0$.

These requirements already 'distort' the otherwise initially flat distributions in the θ parameters.

We partition the data into two parts: i) the preCMS-measurements (preCMS) listed in Table 1 and ii) a limited set of the CMS SUSY and Exotica search results (CMS). With this partitioning, the posterior density becomes

$$p(\theta|D) \sim L(D^{\mathrm{CMS}}|\theta)\, L(D^{\mathrm{preCMS}}|\theta)\, p_0(\theta) = L(D^{\mathrm{CMS}}|\theta)\, p^{\mathrm{preCMS}}(\theta), \qquad (3)$$

where $p_0(\theta)$ is the prior at the start of the inference chain and $p^{\mathrm{preCMS}}(\theta) \sim L(D^{\mathrm{preCMS}}|\theta)\, p_0(\theta)$ can be viewed as a prior that encodes the information from the preCMS-measurements. This partitioning allows us to assess the impact of the CMS results on the pMSSM parameter space while being consistent with constraints from the previous measurements.

Table 1: The measurements that are the basis of our pMSSM prior $p^{\rm preCMS}(\theta)$. All except m_h at the LHC were used to sample points from the pMSSM parameter space via Markov Chain Monte Carlo (MCMC).

i	Observable $\mu_j(\theta)$
1	$BR(b \to s\gamma)$
2	$BR(B_s \to \mu\mu)$
3	$R(B_u \to \tau\nu)$
4	Δa_μ
5	m_t
6	$m_b(m_b)$
7	$\alpha_s(M_Z)$
8a	$m_h \Rightarrow$ pre-LHC: $m_h^{low} = 112$
8b	$m_h \Rightarrow$ LHC: $m_h^{low} = 120$, $m_h^{up} = 130$
9	sparticles $\Rightarrow$ LEP limits
10	prompt $\tilde{\chi}_1^\pm \Rightarrow c\tau(\tilde{\chi}_1^\pm) < 10$ mm

We take the prior $p_0(\theta) = \text{constant}$ at the start of the inference chain.

The experimental results and bounds 1 – 8a and 9 in Table 1 were included in the preCMS likelihood $p^{\rm preCMS}(\theta)$ used in the MCMC scan.

Since we have chosen the prior $p_0(\theta) = \text{constant}$, the sampled points also

constitute a discrete representation of the preCMS likelihood as a function of the pMSSM parameters θ.

Approximately 20 million points were sampled from the pMSSM sub-space, using multiple MCMC chains, from which a random sub-sample of 7205 points were selected subject to the flat prior $120 \leq m_h \leq 130$ GeV.

This prior: (a) is wide enough to permit a study of the influence of the constraints on the mass of the neutral Higgs boson, h, on the posterior densities; (b) allows for the significant uncertainty in the theoretical prediction for m_h.

For each point in the sub-sample of 7205 points, we calculated the predictions $\mu_l(\theta)$ for the observables measured in the included CMS analyses and the associated likelihoods $L(D_l^{\rm CMS}|\mu_l(\theta))$, where $D_l^{\rm CMS}$ denotes the CMS data associated with the prediction $\mu_l(\theta)$.

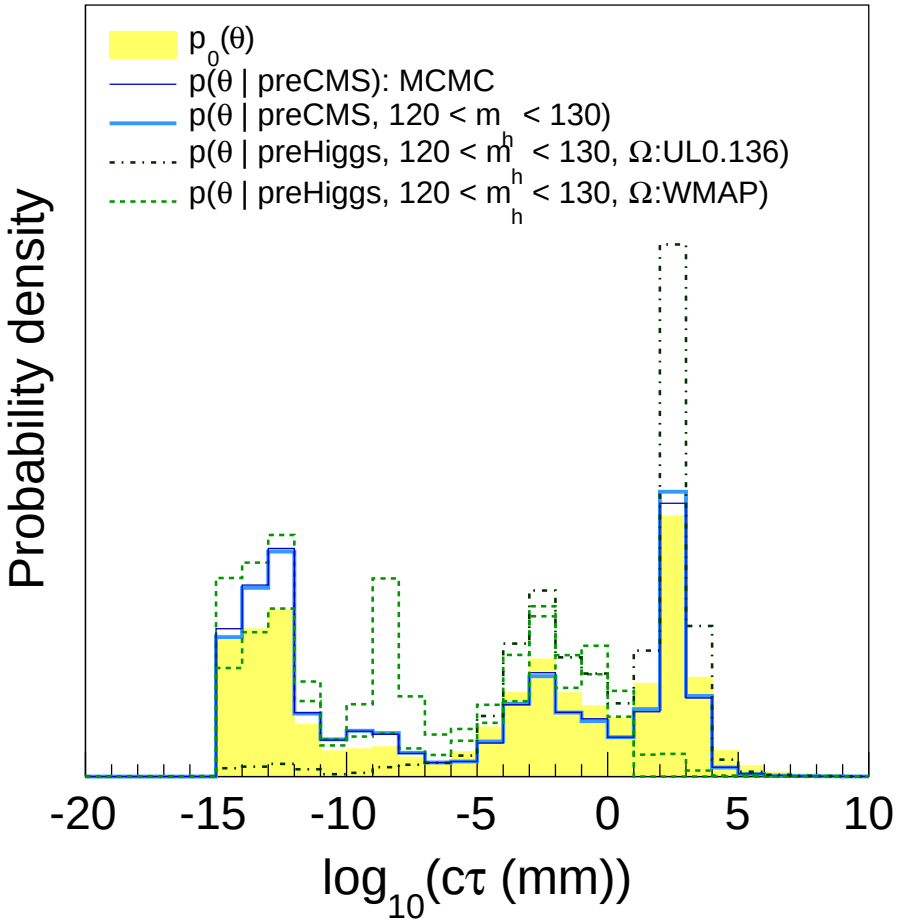
The chargino lifetime

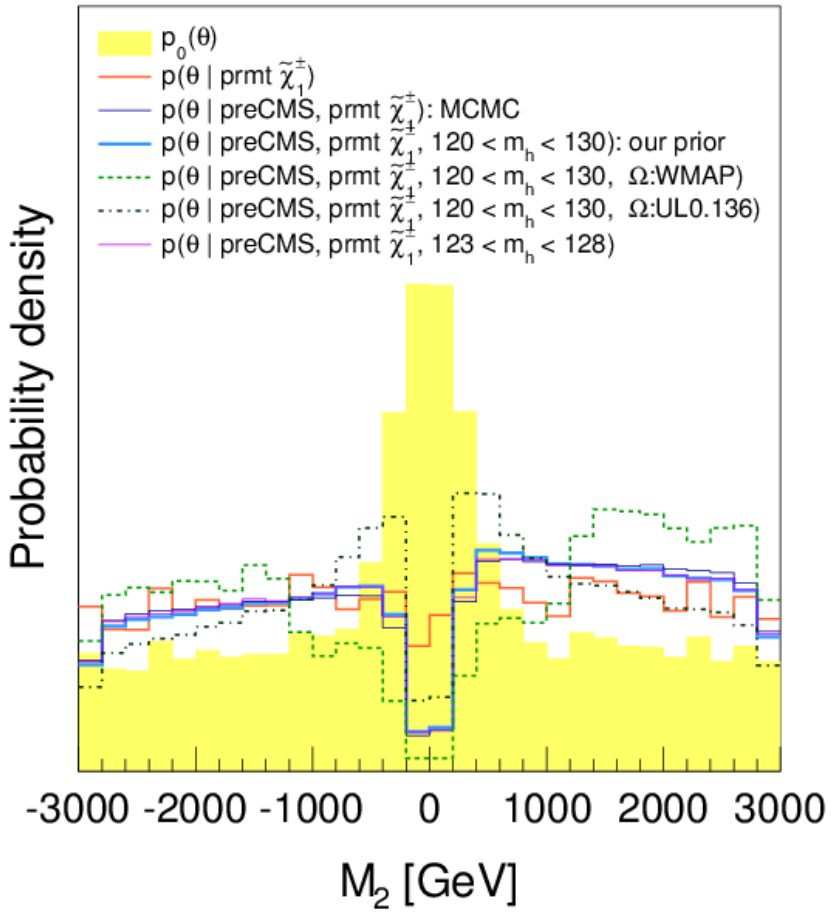
- Letting M_1, M_2 and μ, vary freely over the same range implies that about $2/3$ of the time M_2 or μ will be the smallest mass parameter in the neutralino mass matrix.

  This implies that in a considerable portion of the pMSSM parameter space the $\tilde{\chi}_1^{\pm}$ will be close in mass to the $\tilde{\chi}_1^0$: $\Delta m \sim$ few GeV, ($|\mu| < M_1, M_2$ — higgsino $\tilde{\chi}_1^0$) or almost degenerate, $\Delta m \sim$ few hundred MeV, ($M_2 < |\mu|, M_1$ — wino $\tilde{\chi}_1^0$).

  When the $\tilde{\chi}_1^{\pm}$–$\tilde{\chi}_1^0$ mass difference becomes very small, the charginos are long-lived and can traverse the detector before they decay. Investigating such scenarios thoroughly requires dedicated searches involving lost tracks, dE/dx or time of flight measurements, which was outside the scope of our CMS study.

  Despite the $p_0(\theta)$ being flat in θ, the $c\tau$ distribution is not, in particular having a big peak at large $c\tau$.

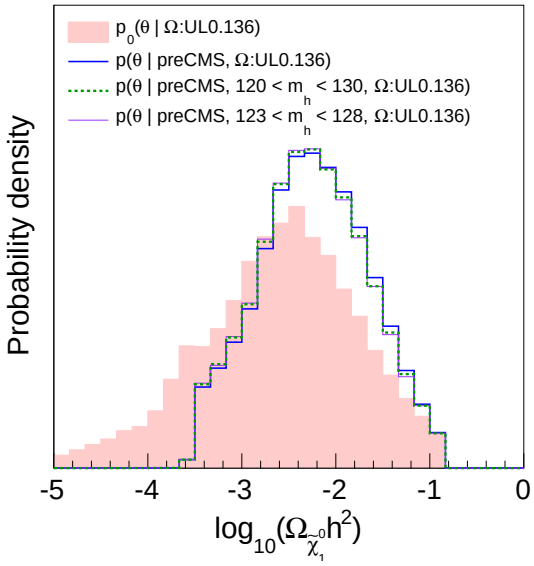


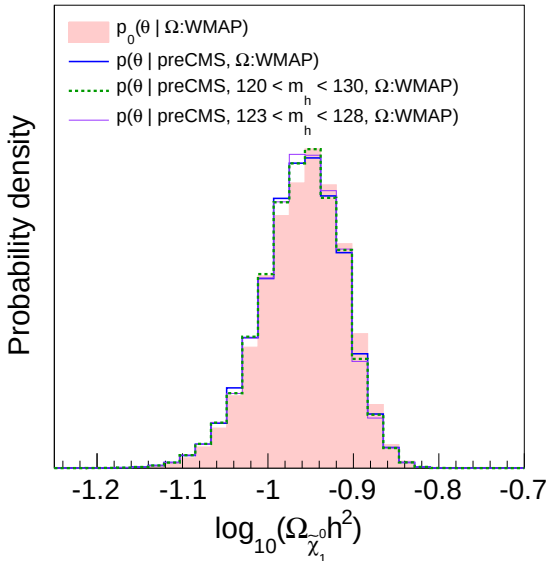


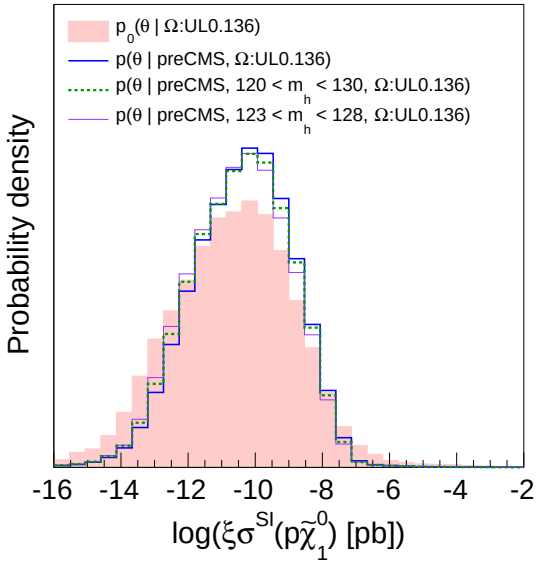
Requiring $\Omega h^2 < 0.136$ (WMAP upper bound) suggests large preference for $10 \text{ mm} < c\tau < 1$ m, a region that is very difficult experimentally.

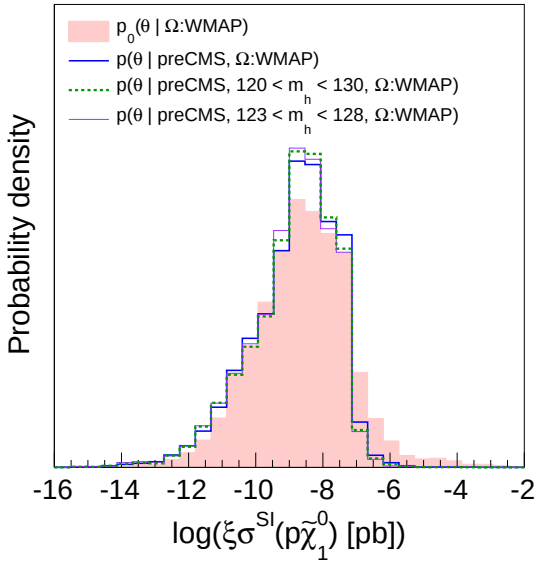
Requiring Ωh^2 inside the WMAP window kills this peak completely, and is generally speaking a somewhat improbable part of parameter space.

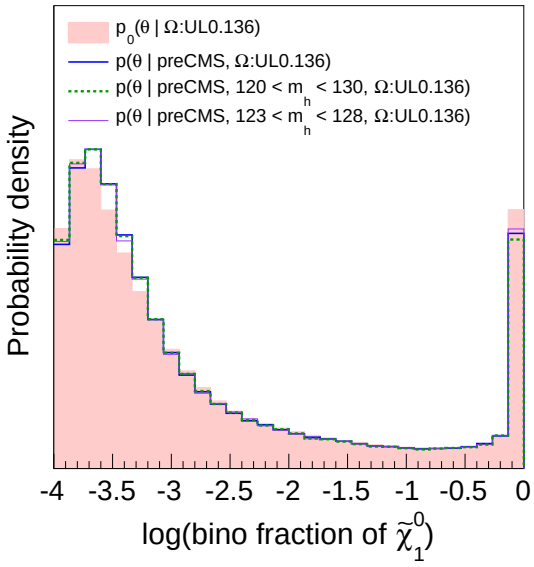
- A few comparisons between ULO.136 ($\Omega h^2 < 0.136$) and WMAP window.

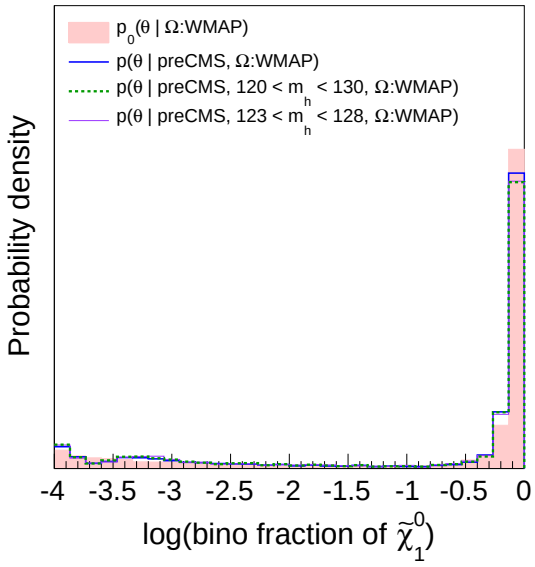


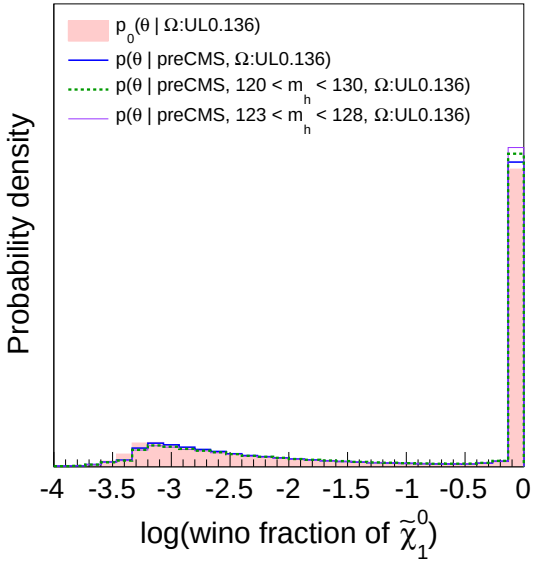


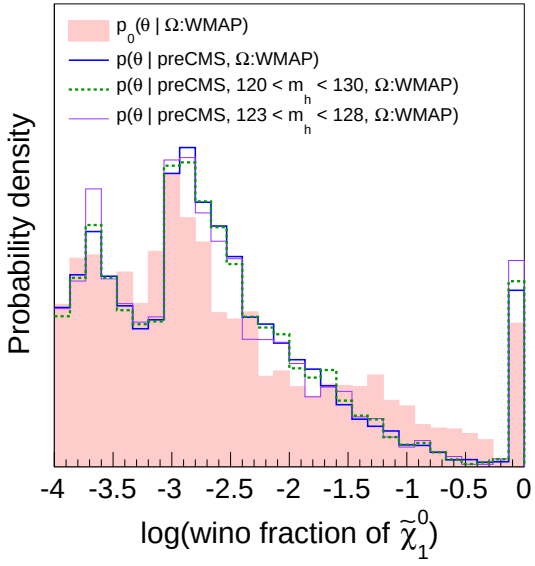


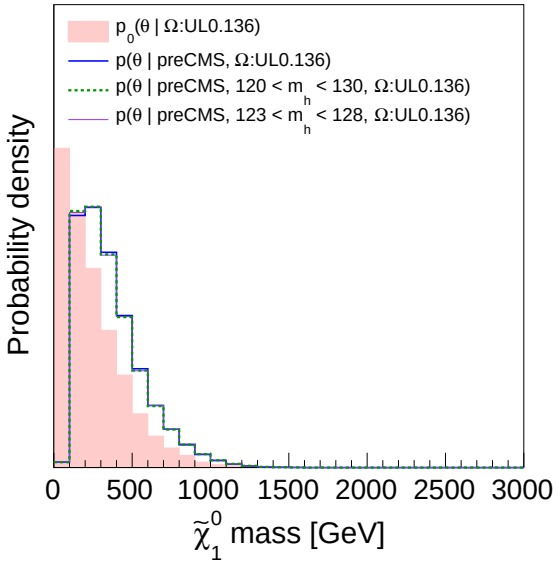


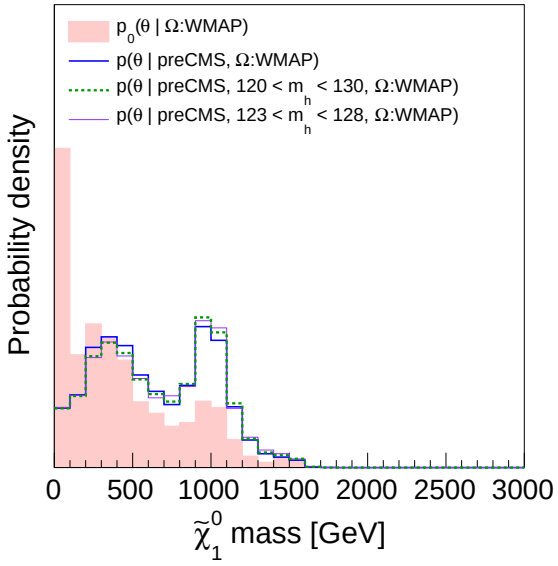

$p_0(\theta \mid \Omega{:}UL0.136)$
$p(\theta \mid preCMS, \Omega{:}UL0.136)$
$p(\theta \mid preCMS, 120 < m_h < 130, \Omega{:}UL0.136)$
$p(\theta \mid preCMS, 123 < m_h < 128, \Omega{:}UL0.136)$
Probability density
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0
log(bino fraction of $\tilde{\chi}_1^0$)

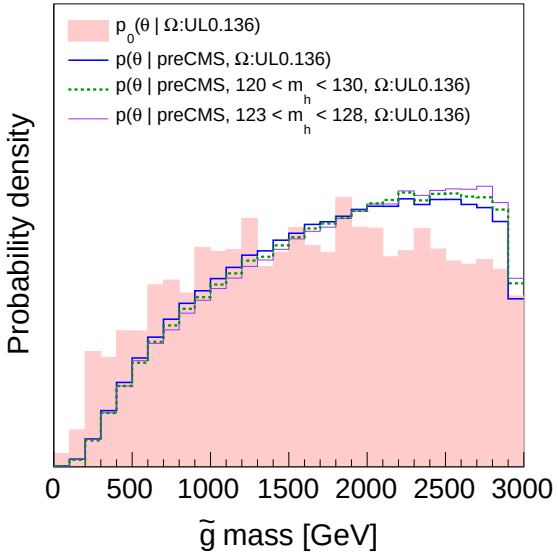

$p_0(\theta \mid \Omega{:}WMAP)$
$p(\theta \mid preCMS, \Omega{:}WMAP)$
$p(\theta \mid preCMS, 120 < m_h < 130, \Omega{:}WMAP)$
$p(\theta \mid preCMS, 123 < m_h < 128, \Omega{:}WMAP)$
Probability density
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0
log(bino fraction of $\tilde{\chi}_1^0$)

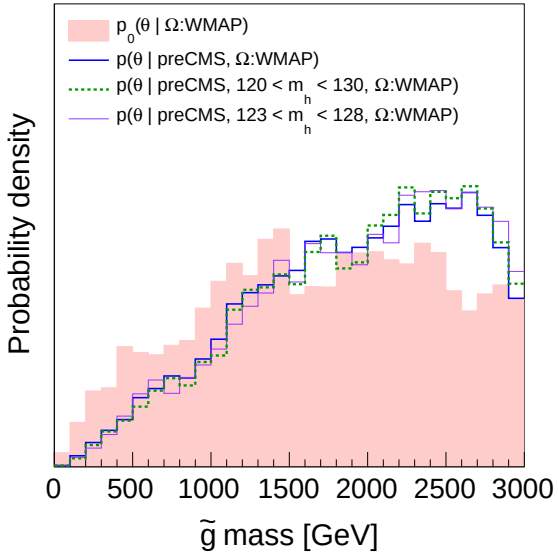

$p_0(\theta \mid \Omega{:}UL0.136)$
$p(\theta \mid preCMS, \Omega{:}UL0.136)$
$p(\theta \mid preCMS, 120 < m_h < 130, \Omega{:}UL0.136)$
$p(\theta \mid preCMS, 123 < m_h < 128, \Omega{:}UL0.136)$
Probability density
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0
log(wino fraction of $\tilde{\chi}_1^0$)


$p_0(\theta \mid \Omega{:}WMAP)$
$p(\theta \mid preCMS, \Omega{:}WMAP)$
$p(\theta \mid preCMS, 120 < m_h < 130, \Omega{:}WMAP)$
$p(\theta \mid preCMS, 123 < m_h < 128, \Omega{:}WMAP)$
Probability density
-4 -3.5 -3 -2.5 -2 -1.5 -1 -0.5 0
log(wino fraction of $\tilde{\chi}_1^0$)

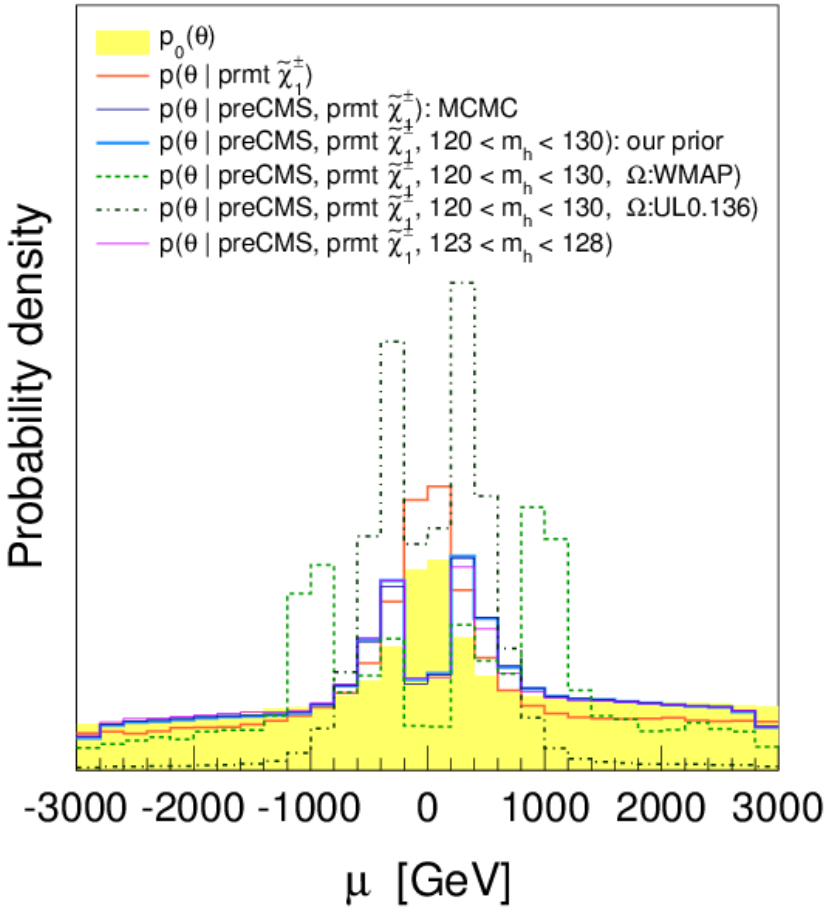

p0(θ | Ω:UL0.136)
p(θ | preCMS, Ω:UL0.136)
p(θ | preCMS, 120 < mh < 130, Ω:UL0.136)
p(θ | preCMS, 123 < mh < 128, Ω:UL0.136)
Probability density
0 500 1000 1500 2000 2500 3000
χ̃10 mass [GeV]

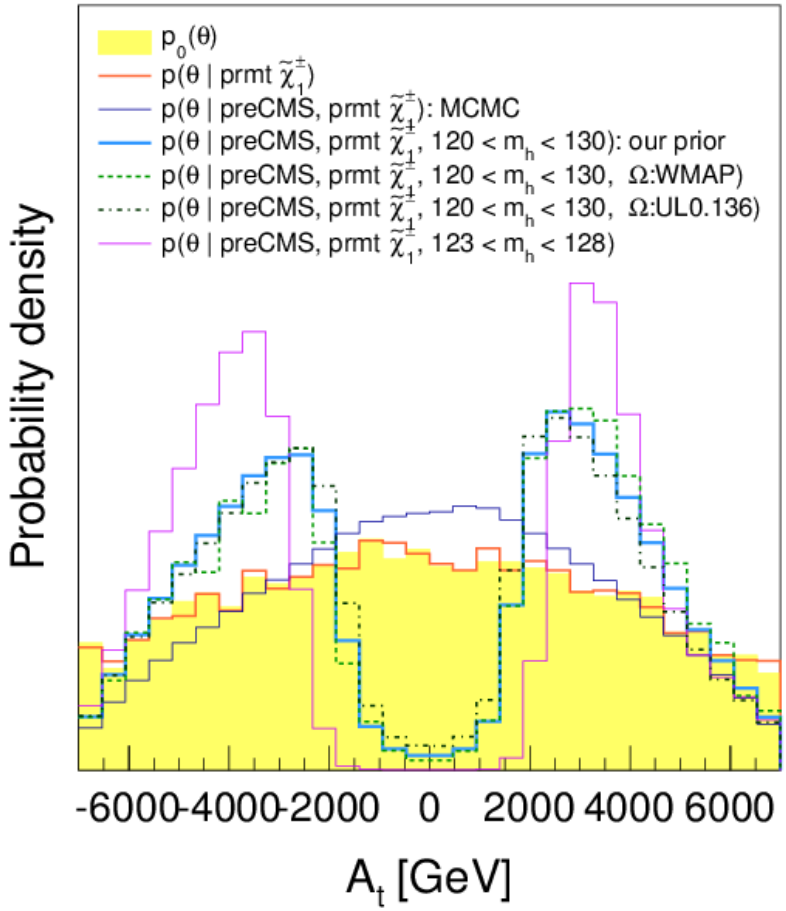

p0(θ | Ω:WMAP)
p(θ | preCMS, Ω:WMAP)
p(θ | preCMS, 120 < mh < 130, Ω:WMAP)
p(θ | preCMS, 123 < mh < 128, Ω:WMAP)
Probability density
0 500 1000 1500 2000 2500 3000
χ̃10 mass [GeV]


p0(θ | Ω:UL0.136)
p(θ | preCMS, Ω:UL0.136)
p(θ | preCMS, 120 < mh < 130, Ω:UL0.136)
p(θ | preCMS, 123 < mh < 128, Ω:UL0.136)
Probability density
0 500 1000 1500 2000 2500 3000
g̃ mass [GeV]

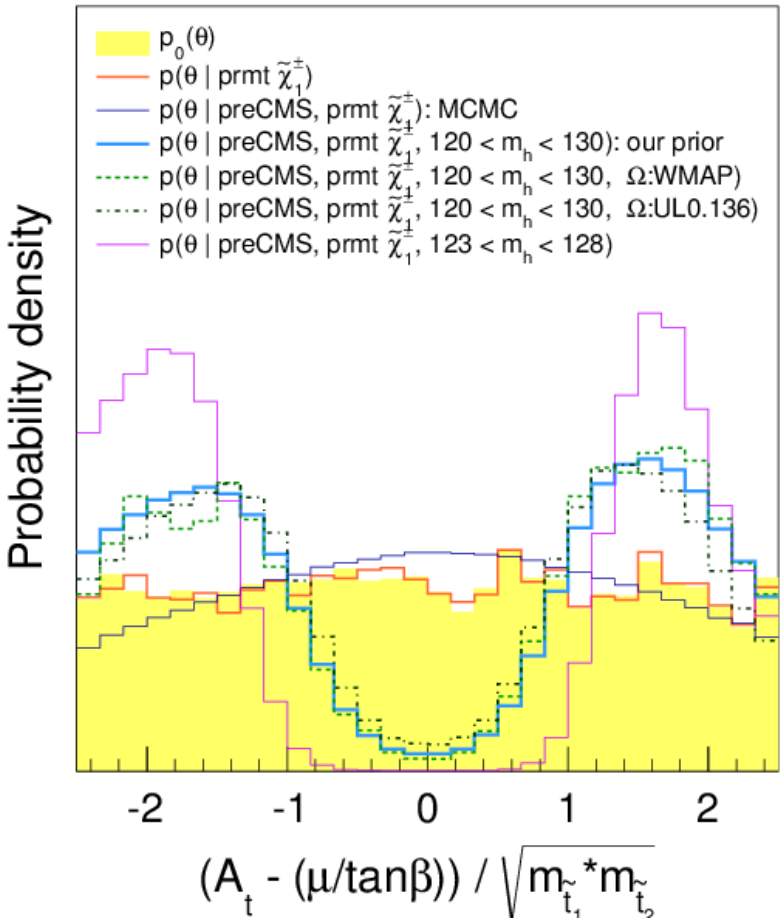

p0(θ | Ω:WMAP)
p(θ | preCMS, Ω:WMAP)
p(θ | preCMS, 120 < mh < 130, Ω:WMAP)
p(θ | preCMS, 123 < mh < 128, Ω:WMAP)
Probability density
0 500 1000 1500 2000 2500 3000
g̃ mass [GeV]

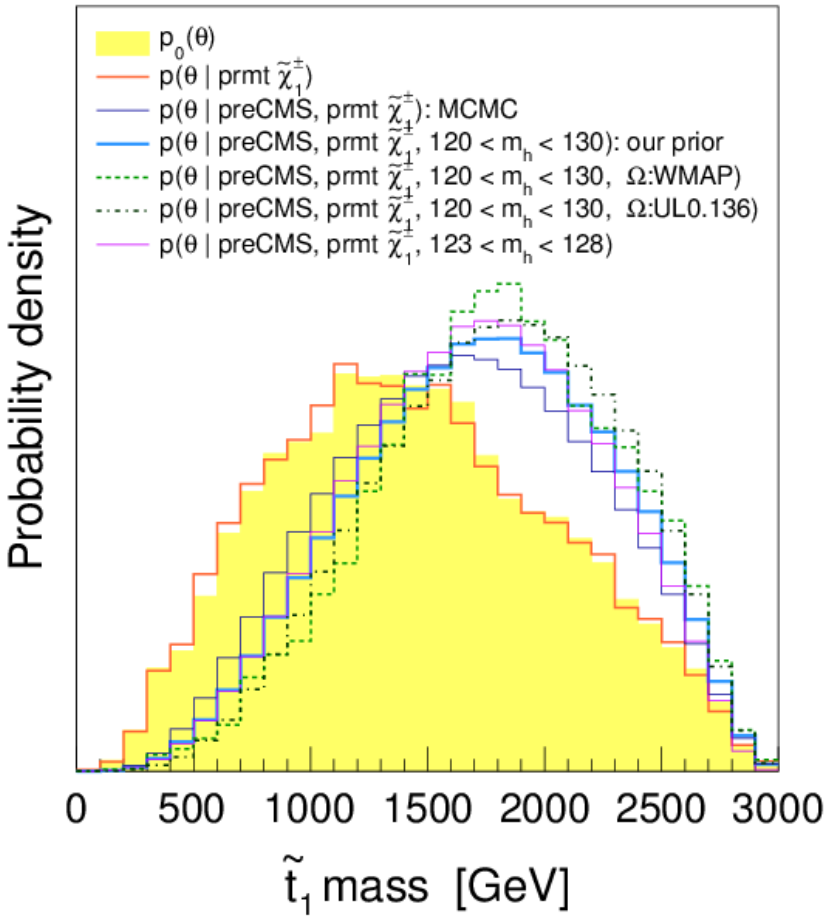
- The analyses we employed assumed prompt chargino decay. Thus, we imposed a model prior of $c\tau(\widetilde{\chi}_1^{\pm}) < 10$ mm $\Rightarrow$ dramatically suppresses "natural" peak at small M_2. $\Rightarrow$ preCMS distributions below.



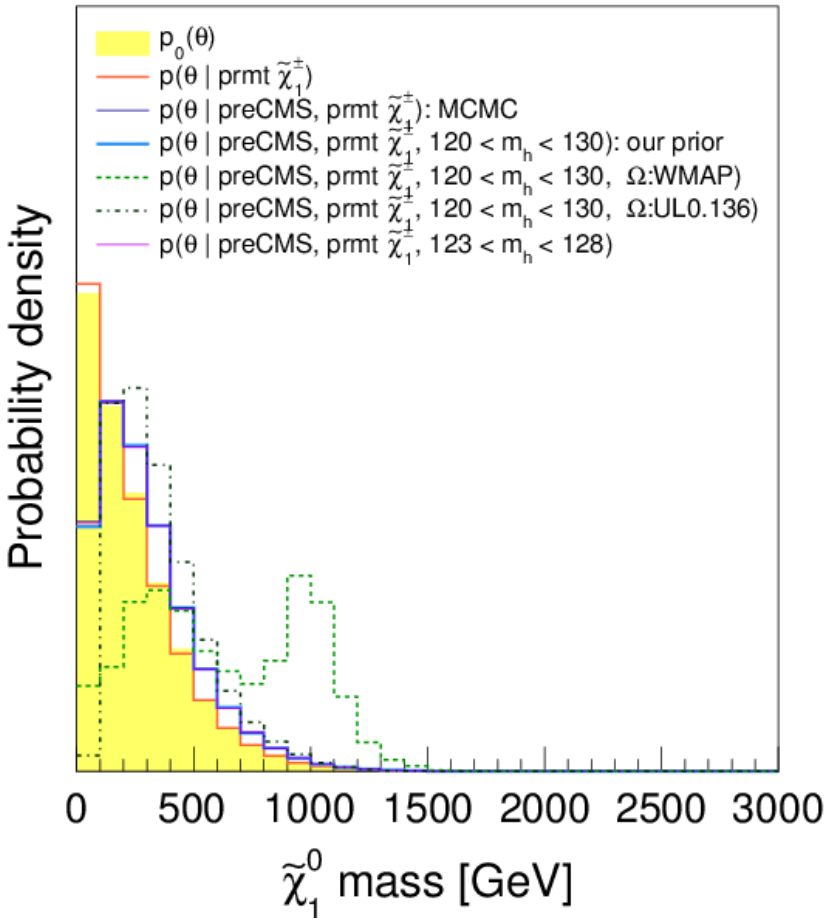


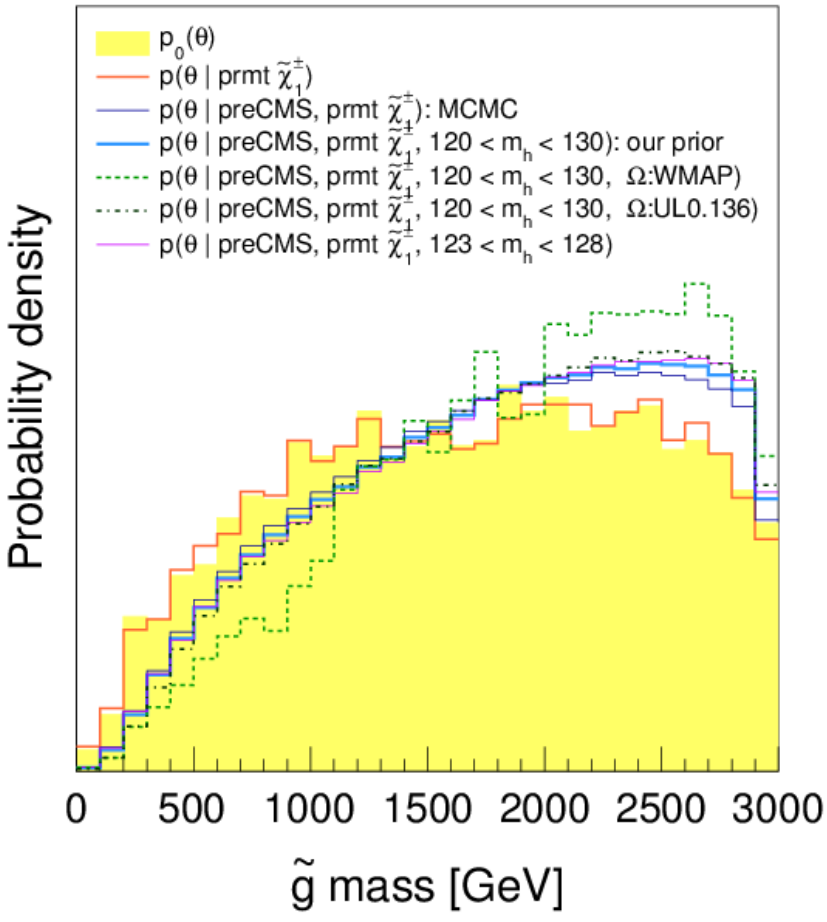
Note: Higgs mass + WMAP $\Rightarrow$ preference for large μ (less "natural"), A_t and $X_t/m_{\rm SUSY}$.





Note: even at the preCMS, $m_{\tilde{t}_1} > 500$ GeV is preferred. This is because a light stop contributes too much to $b \to s\gamma$. $\Rightarrow$ stop-focused analyses are only just able to probe this region.





Note: at the preCMS level, the $\widetilde{\chi}_1^0$ mass is most probably low, but the $\widetilde{g}$ mass has a large spread, including values that are not that large.

Note: WMAP window suggests large probability for large $m_{\widetilde{\chi}_1^0}$ and pushes $\widetilde{g}$ mass to higher values.

Implementation of CMS analyses — 7 TeV

Our first study (SUS-12-030) was for 7 TeV data (~ 5 fb^{-1}). We list in Table 2 the CMS analyses we have implemented in this study. These analyses, which are a subset of the 7 TeV CMS SUSY analyses, cover a variety of final states.

Table 2: List of implemented CMS analyses, which are used for building the CMS likelihood $L(D^{\rm CMS}|\theta)$.

Analysis	CERN doc. no & reference
Hadronic $H_T + H_T^{miss}$ search	CMS-SUS-12-011
Hadronic $H_T + E_T^{miss} + b$-jets search	CMS-SUS-12-003
Hadronic $H_T + E_T^{miss} + \tau$s search	CMS-SUS-12-004
Hadronic monojet $+E_T^{miss}$ search	CMS-EXO-11-059
Leptonic same sign (SS) 2ℓ search	CMS-SUS-11-010
Leptonic opposite sign (OS) 2ℓ search	CMS-SUS-11-011
Leptonic electroweakino (EWKino) search	CMS-SUS-12-006

Each of the analyses of Table 2 provides results for a number of search regions. When these search regions are disjoint, we combine the results by taking the product of the likelihoods calculated for each search region. In the end we provide the

posterior distributions for each analysis separately, presenting the combined result for the analysis when the search regions are disjoint, or presenting the results for individual search regions when the search regions overlap.

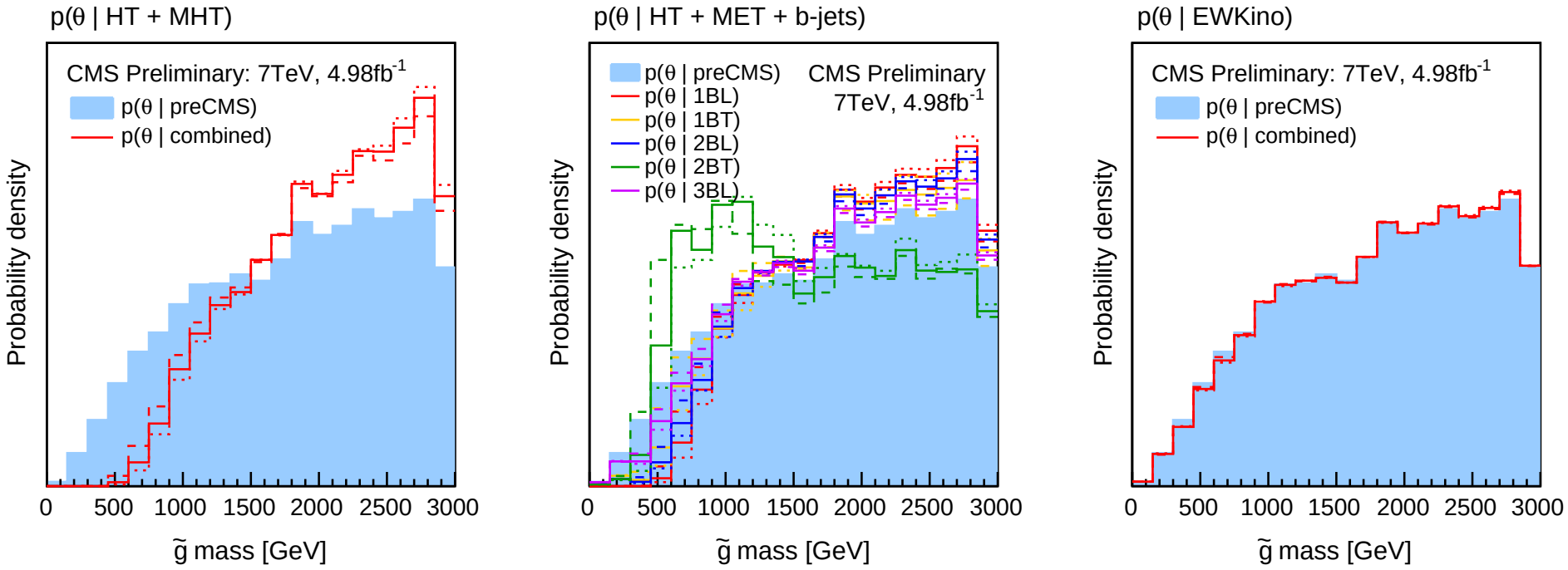


Figure 1: Marginalized 1D posterior probability distributions for $\tilde{g}$ mass. The line histograms in the three plots show posterior densities after including three of the seven implemented CMS analyses: $H_T + H_T^{miss}$, $H_T + E_T^{miss} + b$-jets and EWKino.

The hadronic searches favor higher $\tilde{g}$ masses, except for the 2BT ($2b$-jets + tight H_T) search region defined by $H_T > 600$ GeV, $E_T^{miss} > 300$ GeV and $N_{b\text{-jets}} \geq 2$

of the $H_T + E_T^{miss} + b$-jets search, which prefers lower gluino masses. The latter is due to a 2.2σ excess observed in that channel (SUS-12-003). For the $\tilde{u}_R$, $\tilde{u}_L$ and $\tilde{d}_{R,L}$ squarks higher mass values are favored except for the 2BT search region of the $H_T + E_T^{miss} + b$-jets analysis.

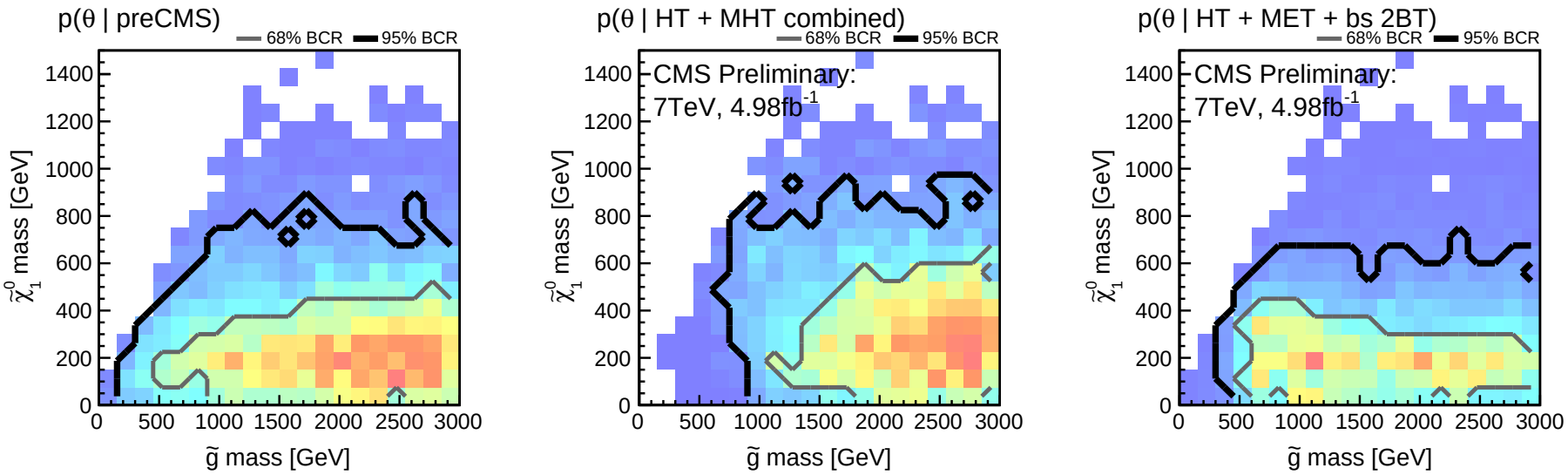


Figure 2: Marginalized 2D posterior probability distributions for $\tilde{\chi}_1^0$ mass versus $\tilde{g}$ mass. In each row, the 1st plot shows the preCMS posterior density. The 2nd and 3rd plots show posterior densities after applying the $H_T + H_T^{miss}$ combined and $H_T + E_T^{miss} + b$-jets 2BT results respectively. The grey and black contours enclose the 68% and 95% Bayesian credible regions respectively.

Our results show that there are still many low-mass scenarios in the pMSSM parameter space to which current CMS searches are insensitive. There are several

reasons.

- cross sections are simply too low for available integrated luminosities.
- mass splittings between some SUSY particles are small, yielding jets and leptons of lower p_T on average than those that can be studied with current searches.
- the E_T^{miss} is lower on average than the E_T^{miss} thresholds in current analyses.

A quantiative measure of insensitivity: A search is insensitive to new physics if the analysis cannot distinguish between the background plus signal hypothesis, denoted by H_1, and the background-only hypothesis, denoted by H_0. To measure the relative probabilities of the two hypotheses H_1 at θ and H_0, we use the local Bayes factor

$$B_{10}(\theta) = L(D^{\rm CMS}|\theta, H_1)/L(D^{\rm CMS}|H_0), \tag{4}$$

to distinguish it from the global Bayes factor $B_{10} = L(D^{CMS}|H_1)/L(D^{CMS}|H_0)$, in which the likelihood $L(D^{\rm CMS}|\theta, H_1)$ times the prior $p(\theta|H_1) = p(\theta)$ is marginalized with respect to θ. This we map to

$$Z = {\rm sign}(\ln {\rm B}_{10})\sqrt{2|\ln {\rm B}_{10}|}, \tag{5}$$

which is a signed Bayesian analog of the frequentist "n-sigma". In conventional language, the case $B_{10}(\theta) \gg 1$ would indicate a signal at "Z-sigma significance", while the case $1/B_{10}(\theta) \gg 1$ would indicate a signal exclusion at "Z-sigma significance". Note that in our definition of Z, negative values correspond to exclusions while positive values are associated with potential observations. We define a search to be **insensitive** if

$$|Z| \leq 2. \tag{6}$$

A point with $Z > 5$ would signify a discovery while $Z < -1.64$ would mean that the point is excluded at 95% confidence level (CL).

When the various analyses, and or the search regions within a single analysis, considered are not exclusive, we employ Z with the largest absolute value, that is,

$$\begin{aligned} l_{best} &\equiv \arg\max_{l}(|Z_l|), \\ Z_{best} &= Z_{l_{best}}, \end{aligned} \tag{7}$$

where l_{best} is the index of the analysis with the largest absolute significance. Whenever there is a combined likelihood available for an analysis, as in $H_T + H_T^{miss}$, OS 2ℓ and EWKino analyses, we include the single Z value obtained from that likelihood into

the Z_{best} calculation. Whenever a combined likelihood cannot be obtained from the search regions in an analysis, as in $H_T + E_T^{miss} + b$-jets, $H_T + E_T^{miss} + \tau$s, monojet $+E_T^{miss}$ or SS 2ℓ analyses, we include the Z value from each search region in the Z_{best} calculation.

Exploring the unexplored

In order to quantify what regions of the pMSSM have been missed by the searches considered in this study, a set of pMSSM points has been defined for which $|Z| \leq 2$. This set of points are referred to as the "unexplored" or non-excluded points. (Note that, evidently, the points are "unexplored" only with respect to the analyses we have included in this study.)

A total of 4504 of all 7205 points are found in this set. It is especially interesting to focus on those "unexplored" points that have a large production cross section (i.e., with $\sigma > 10$ fb).

Of the 4504 points, 2198 fulfill this criterion. We refer to the latter as "unexplored high-σ points". Fig. 3 shows the total production cross sections of the unexplored pMSSM points.

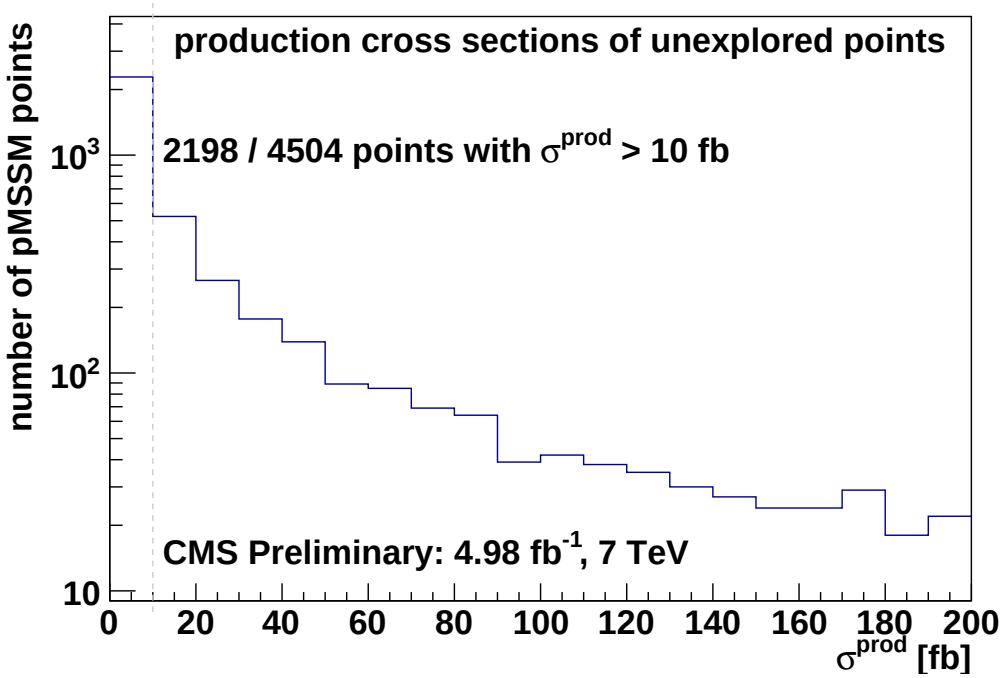


Figure 3: Number of unexplored pMSSM points, as a function of the production cross section.

Unexplored high-σ points

Why do the 2198 unexplored pMSSM points with $\sigma^{\rm prod} > 10$ fb escape detection? The dominant production mechanisms associated with such points are: (1) EWK production of charginos and neutralinos; (2) and, with lower frequency, squark-pair production.

SMS decomposition reveals that mass degeneracy issues are often the issue.

The $|Z_{best}|$ value can be used to construct a binary likelihood of the combination of all analyses where $p(|Z_{best}| < 2|\theta) = 1$ and $p(|Z_{best}| \geq 2|\theta) = 0$. The distributions of points with $|Z_{best}| < 2$ can be viewed as the probability densities of the so-called $non\text{-}excluded\ points$, that is, points that lie in the region we have called unexplored.

Comparing $|Z_{best}| < 2$ with the preCMS distributions for selected sparticle masses and for the total sparticle production cross section, we find:

- The CMS analyses have the most impact on $\tilde{g}$ and light-flavor squark masses as well as on the cross section.

- This is followed by the 3rd generation squarks and EW gauginos.

- In terms of the analyses, the $H_T + H_T^{miss}$ results are more decisive compared to jets $+ E_T^{miss} + b$-jets. This will also apply to the currently-public $8\ \mathrm{TeV}$ results of the next section.

In Fig. 4, we plot the marginalized 1D posterior densities of the best significance Z_{best}. The left histogram depicts the preCMS distribution of this quantity, that is, the best significance distribution incorporating preCMS likelihoods only.

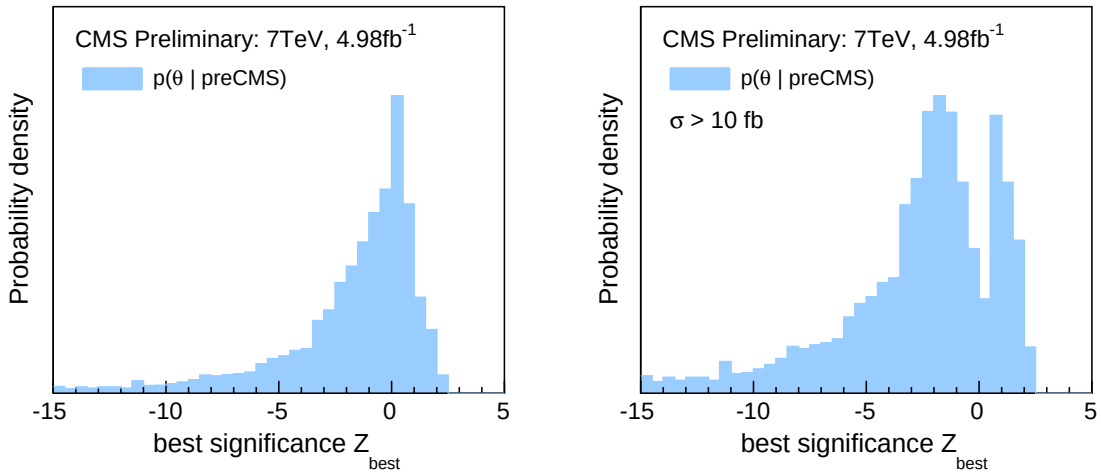


Figure 4: Unexplored pMSSM points.

About 32% of the pMSSM points we analyze have $\sigma < 10$ fb. About 63% of the points in the left plot have $|Z_{best}| < 2$, and hence are inaccessible with 5 fb^{-1} of 7 TeV LHC data.

The right plot of Fig. 4 shows the best significance distribution for points with $\sigma > 10$ fb, which are much more likely to be accessible with 5fb^{-1} of 7 TeV LHC data. The probability of having $|Z_{best}| < 2$ in this case is 45%. As expected, the best significance distribution is wider, and the impact of the CMS analyses in this accessible subset of the pMSSM sub-space is much greater.

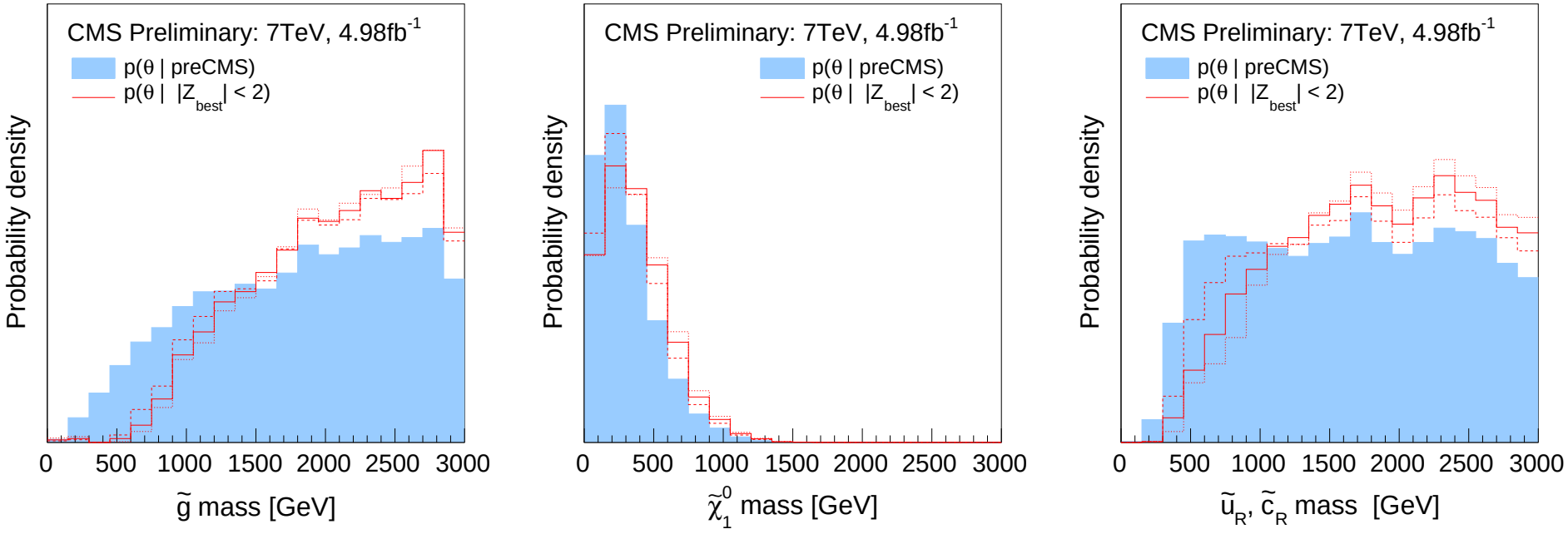


Figure 5: Unexplored pMSSM points as a function of various sparticle masses.

Public results for 8 TeV (19.5 fb^{-1})

The 8 TeV data is quite effective at whittling away more of the pMSSM parameter space associated with $c\tau < 10$ mm. Consider, *e.g.* the $m_{\tilde{\chi}_1^0}$ vs. $m_{\tilde{g}}$ plane. Of the currently public results, the HT+MHT (inclusive) analysis (SUS-13-012) is more effective than the HT+MHT+b-jets analysis (SUS-12-024). Of course, we are working to combine these with other analyses, in particular the EWkino analyses.

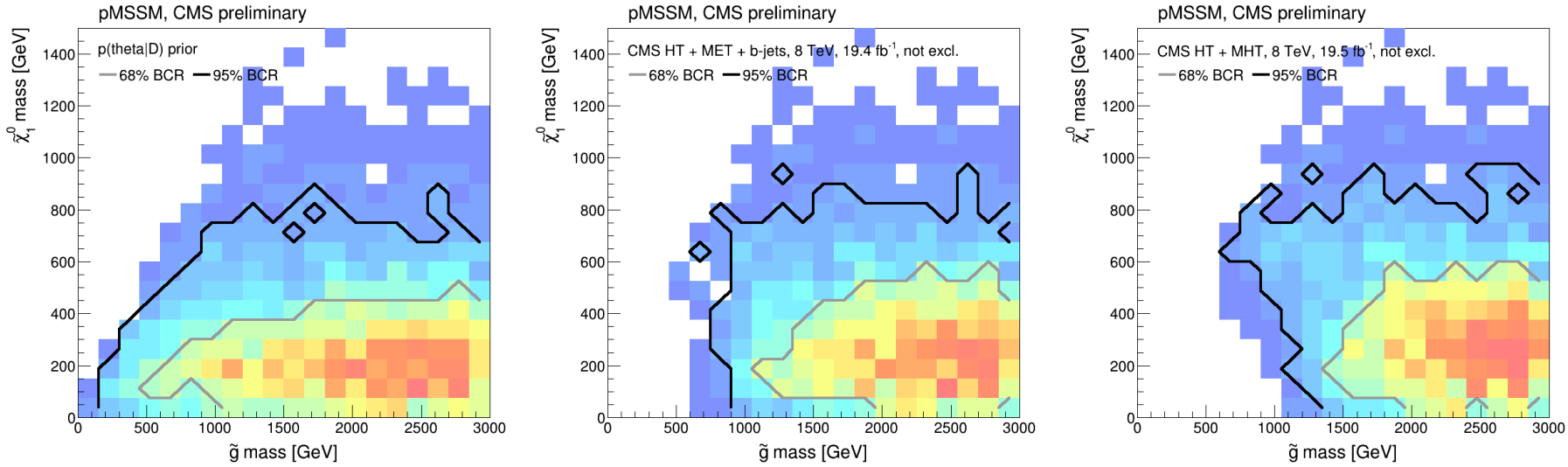


Figure 6: 2D distributions before 8 TeV data, after HT+MHT+b-jets, and after HT+MHT analyses.

Another perspective is provided by examining the 1D distributions. Below we show the distributions for 95% excluded ($|Z| > 1.64$) vs. non-excluded ($|Z| < 1.64$) in two cases: (1) the mass of the lightest colored particle; and (2) the signal cross section

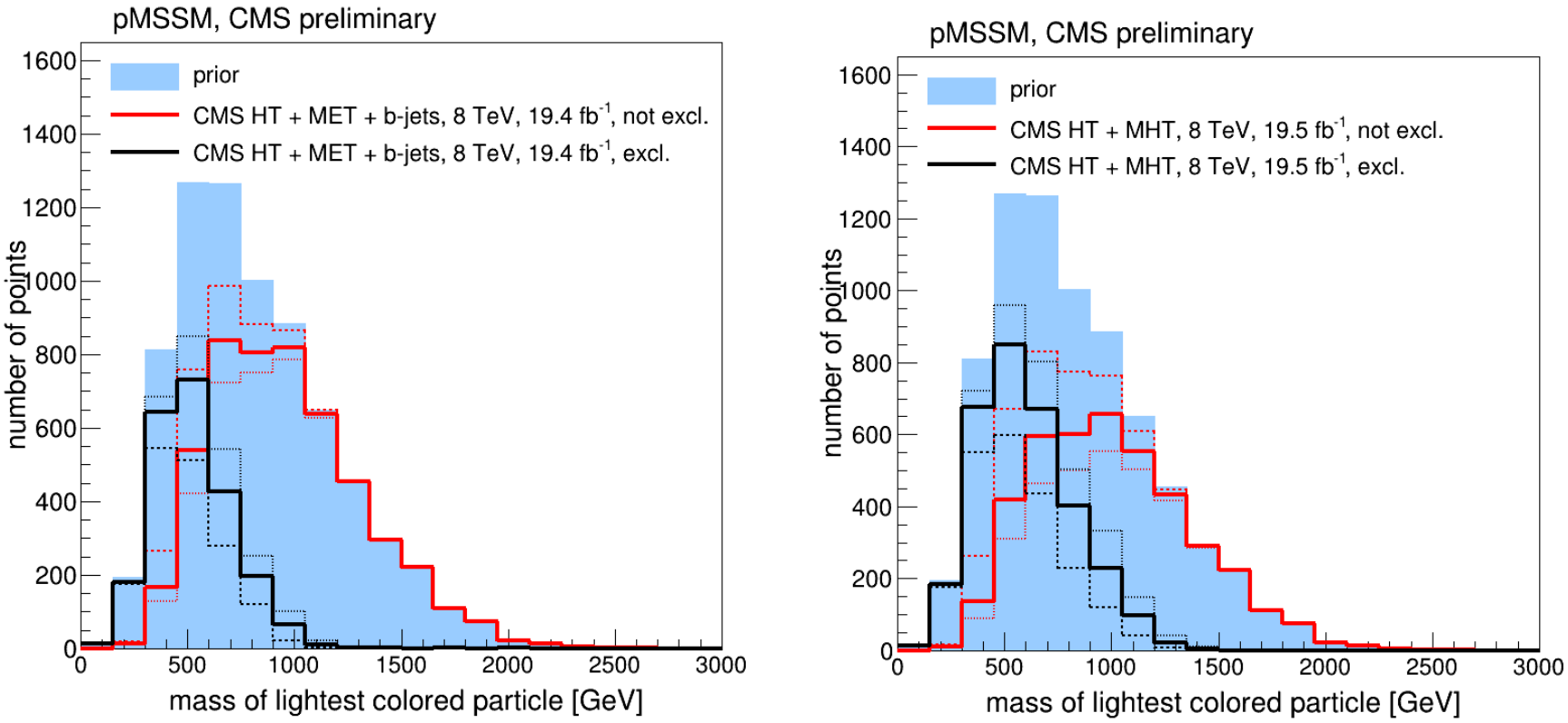


Figure 7: Mass of the lightest colored particle after 8 TeV data for HT+MHT+b-jets and HT+MHT, compared to preCMS. Excluded ($|Z| > 2$) vs. non-excluded ($|Z| < 2$).

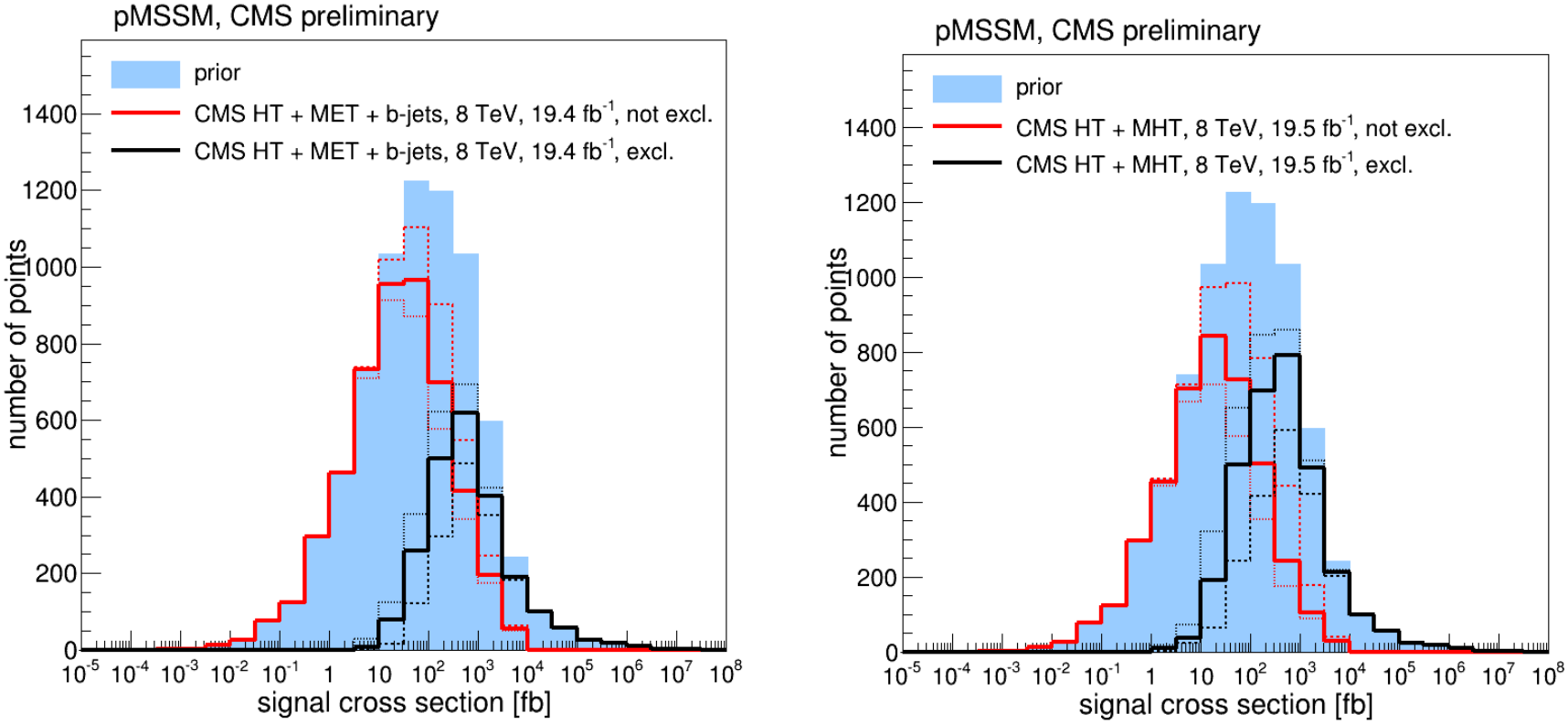


Figure 8: Cross section after 8 TeV data for HT+MHT+b-jets and HT+MHT, compared to preCMS. Excluded ($|Z| > 2$) vs. non-excluded ($|Z| < 2$).

Conclusions

- The pMSSM provides an interesting perspective on low-energy supersymmetry.

  Its relevance depends upon your point of view regarding how simple the extrapolation to the GUT scale should be.

  My opinion: we should not give up on models that are fine at the EW scale just because they don't evolve benignly up to the GUT scale.

  The pMSSM is an excellent way to categorize and explore the full MSSM parameter space.

- The Bayesian method for doing the interpretation makes it feasible to interpret multi-parameter models.

- It is interesting to note that WMAP-window models typically have small $c\tau$ for the $\widetilde{\chi}_1^{\pm}$ and will therefore be constrained by 'prompt' ($c\tau < 10$ mm) decay analyses as presented here.

In contrast, if we allow for other sources of dark matter, we might wish to take seriously the much more easily realized $\Omega h^2 < 0.01$ models which predict a large probability for 10 mm $< c\tau < 1$ m. Such models will be a real challenge at the LHC since the typical event will have 'stubs', 'DITS' (short curly tracks) and so forth rather than simple missing energy. (see the early papers by JG and S. Mrenna, for example).

- CMS data (and ATLAS also) is significantly impacting the pMSSM parameter space, excluding most, but certainly not all, of the high σ models.

- In the case of unexcluded high-σ models, small mass splittings are primarily to blame for lack of sensitivity. $\Rightarrow$ might gain sensitivity using more refined analyses of current data.

  But, there are many low-σ models that can only be explored with more energy and luminosity at the LHC. $\Rightarrow$ both are coming!

- We are working on the combination of $7 + 8$ TeV results.