

MSSM Forecast for the LHC

Maria Eugenia Cabrera

Instituto de Física Teórica, IFT-UAM/CSIC
U.A.M., Cantoblanco,
28049 Madrid, Spain
E-mail: maria.cabrera@uam.es

Alberto Casas

Instituto de Física Teórica, IFT-UAM/CSIC
U.A.M., Cantoblanco,
28049 Madrid, Spain
E-mail: alberto.casas@uam.es

Roberto Ruiz de Austri

Instituto de Física Corpuscular, IFIC-UV/CSIC
Valencia, Spain
E-mail: rruiz@ific.uv.es

ABSTRACT: We perform a forecast of the MSSM with universal soft terms (CMSSM) for the LHC, based on an improved Bayesian analysis. We do not incorporate ad hoc measures of the fine-tuning to penalize unnatural possibilities: such penalization arises from the Bayesian analysis itself when the experimental value of M_Z is considered. This allows to scan the whole parameter space, allowing arbitrarily large soft terms. Still the low-energy region is statistically favoured (even before including dark matter or g-2 constraints). Contrary to other studies, the results are almost unaffected by changing the upper limits taken for the soft terms. The results are also remarkable stable when using flat or logarithmic priors, a fact that arises from the larger statistical weight of the low-energy region in both cases. Then we incorporate all the important experimental constraints to the analysis, obtaining a map of the probability density of the MSSM parameter space, i.e. the forecast of the MSSM. Since not all the experimental information is equally robust, we perform separate analyses depending on the group of observables used. When only the most robust ones are used, the favoured region of the parameter space contains a significant portion outside the LHC reach. This effect gets reinforced if the Higgs mass is not close to its present experimental limit and persists when dark matter constraints are included. Only when the g-2 constraint (based on e^+e^- data) is considered, the preferred region (for $\mu > 0$) is well inside the LHC scope. We also perform a Bayesian comparison of the positive- and negative- μ possibilities.

KEYWORDS: Supersymmetric Effective Theories, Beyond Standard Model, Supersymmetry Phenomenology.

Contents

1. Introduction	1
2. The set up for the scan of the MSSM	3
2.1 Electroweak breaking	3
2.2 Treatment of nuisance variables and Yukawa couplings	4
2.3 Variables for the MSSM scan	6
2.4 Priors in the initial parameters	8
3. High-energy vs Low-energy regions and the EW breaking	12
4. Experimental Constraints	16
4.1 EW and B-physics observables, and limits on particle masses	17
4.2 Constraints from $(g - 2)_\mu$	22
4.3 Constraints from Dark Matter	27
5. Negative sign of μ	29
5.1 Results	29
5.2 Positive versus negative μ	31
6. Summary and conclusions	35

1. Introduction

The idea of an LHC forecast for the Minimal Supersymmetric Standard Model (MSSM) is to use all the present (theoretical and experimental) information available to determine the relative probability of the different regions of the MSSM parameter space. This includes theoretical constraints (and perhaps prejudices) and experimental constraints, such as electroweak precision tests. For recent work on this subject see refs. [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. An appropriate framework to perform such forecast is the Bayesian approach (for a review see ref. [11]), which allows a sensible statistical analysis and to separate in a neat way the objective and subjective pieces of information.

The probability density of a particular point in the parameter space, say $\{\theta_i^0\}$, given a certain set of *data*, is the so-called posterior probability density function (pdf), $p(\theta_i^0|\text{data})$, which is given by the fundamental Bayesian relation

$$p(\theta_i^0|\text{data}) = p(\text{data}|\theta_i^0) p(\theta_i^0) \frac{1}{p(\text{data})} . \quad (1.1)$$

Here $p(\text{data}|\theta_i^0)$ is the likelihood (sometimes denoted by \mathcal{L}), i.e. the probability density of measuring the given data for the chosen point in the parameter space¹. $p(\theta_i^0)$ is the prior, i.e. the “theoretical” probability density that we assign a priori to the point in the parameter space. Finally, $p(\text{data})$ is a normalization factor which plays no role unless one wishes to compare different classes of models, so for the moment it can be dropped from the previous formula. One can say that in eq. (1.1) the first factor (the likelihood) is objective, while the second (the prior) contains our prejudices about how the probability is distributed a priori in the parameter space, given all our previous knowledge about the model.

Ignoring the prior factor is not necessarily the most reasonable or “free of prejudices” attitude. Such procedure amounts to an implicit choice for the prior, namely a completely flat prior in the parameters. However, choosing e.g. θ_i^2 as initial parameters instead of θ_i , the previous flat prior becomes non-flat. So one needs some theoretical basis to establish, at least, the parameters whose prior can be reasonably taken as flat.

Besides, note that a choice for the allowed ranges of the various parameters is necessary in order to make statistical statements. Often one is interested in showing the probability density of one (or several) of the initial parameters, say θ_i , $i = 1, \dots, N_1$, but not in the others, θ_i , $i = N_1 + 1, \dots, N$. Then, one has to *marginalize* the latter, i.e. integrate in the parameter space:

$$p(\theta_i, i = 1, \dots, N_1|\text{data}) = \int d\theta_{N_1+1}, \dots, d\theta_N p(\theta_i, i = 1, \dots, N|\text{data}) . \quad (1.2)$$

This procedure is very useful and common to make predictions about the values of particularly interesting parameters. Now, in order to perform the marginalization, we need an input for the prior functions *and* for the range of allowed values of the parameters, which determines the range of the definite integration (1.2). A choice for these ingredients is therefore inescapable in trying to make LHC forecasts.

In the present paper we apply these concepts to the study of the MSSM [12]. More precisely, we will consider a standard framework, often called CMSSM or MSUGRA, in which the soft parameters are assumed universal at a high scale (M_X), where the supersymmetry (SUSY) breaking is transmitted to the observable sector, as happens e.g. in the gravity-mediated SUSY breaking scenario. Hence, our parameter-space is defined by the following parameters:

$$\{\theta_i\} = \{m, M, A, B, \mu, s\} . \quad (1.3)$$

¹Frequentist approaches, which are an alternative to Bayesian ones, are based on the analysis of the likelihood function in the parameter space; see ref. [10] for a recent frequentist analysis of the MSSM.

Here m , M and A are the universal scalar mass, gaugino mass and trilinear scalar coupling; B is the bilinear scalar coupling; μ is the usual Higgs mass term in the superpotential; and s stands for the SM-like parameters of the MSSM. The latter include the $SU(3) \times SU(2) \times U(1)_Y$ gauge couplings, g_3, g, g' , and the Yukawa couplings, which in turn determine the fermion masses and mixing angles.

In sect. 2 we explain the set up for this study. In our opinion we have improved previous analyses in several aspects. Namely, we have not made ad hoc assumptions to penalize fine-tuned regions. A nice consequence is the absence of dependences on the initial ranges for the MSSM parameters. Besides, we have done a rigorous treatment of the nuisance variables (in particular Yukawa couplings) and we have made a satisfactory choice of priors for the initial parameters (actually two different choices to evaluate the dependence on the prior). In sect. 3 we compare the relative probability of the high- and low-energy (i.e. accessible to LHC) regions of the MSSM parameter space. We show that, for any reasonable prior, the low-energy region is statistically favoured only after properly incorporating the information about the scale of electroweak breaking. Sections 4 and 5 are devoted to include all the important experimental constraints into the analysis, for positive and negative μ -parameter respectively. In this way, we obtain a map of the probability density of the MSSM parameter space, i.e. the MSSM forecast for the LHC. We distinguish between the most robust experimental data (EW observables, limits on masses of supersymmetric particles, etc.) and more controversial data ($g_\mu - 2$) or model-dependent constraints (Dark Matter), performing separate analyses depending on the group of observables used. The comparison between the positive- and negative- μ cases is done in sect. 5. In sect. 6 we present a summary of the analysis and the main conclusions.

2. The set up for the scan of the MSSM

2.1 Electroweak breaking

The main motivation of low-energy SUSY is the nice implementation of the electroweak (EW) breaking, where the EW scale (or, equivalently, the Z mass) does not suffer from enormous (quadratic) radiative corrections. Actually, in the MSSM the EW breaking occurs naturally in a substantial part of the parameter space. This success is greatly due to the SUSY radiative contributions to the Higgs potential. Of course, in our analysis, the points of the parameter space that do not have a correct EW breaking are to be discarded, as usual.

It is common lore that the parameters of the MSSM, $\{m, M, A, B, \mu\}$, should not be far from the experimental EW scale in order to avoid unnatural fine-tunings to obtain the correct size of the E.W breaking. There is a rich literature [13, 14, 15, 16, 17] about the best way to quantify this fine-tuning. It is normally understood that regions of the parameter space with large fine-tuning (typically at large values of the soft parameters) are to be considered unnatural and thus uninteresting. The exception to this rule are landscape-like scenarios, which we do not consider in detail in this paper. Previous Bayesian

studies of the MSSM have attempted to incorporate this criterion by implementing some penalization of the fine-tuned regions, e.g. using a conveniently modified prior for the MSSM parameters [2, 16, 4]. Another (more usual) practice has been to restrict the range of the soft parameters to \gtrsim few TeV, but this makes the results dependent on the actual ranges considered.

However, since the naturalness arguments are deep down statistical arguments, one might expect that an effective penalization of fine-tunings should arise from the Bayesian analysis itself, with no need of introducing “naturalness priors” or restricting the soft terms to the low-energy scale. It was shown in ref. [18] that this is indeed the case (see also [17] for a previous observation in this sense). Let us briefly recall the argument. The key point is to consider M_Z^{exp} as experimental data on a similar foot to the others, entering the total likelihood, \mathcal{L} . For the sake of simplicity let us approximate the likelihood associated to the Z mass as a Dirac delta, so

$$p(\text{data}|s, m, M, A, B, \mu) \simeq \delta(M_Z - M_Z^{\text{exp}}) \mathcal{L}_{\text{rest}} , \quad (2.1)$$

where $\mathcal{L}_{\text{rest}}$ is the likelihood associated to all the physical observables, except M_Z . Now, we can take advantage of this Dirac delta to marginalize the pdf in one of the initial parameters, e.g. μ , performing a change of variable $\mu \rightarrow M_Z$:

$$\begin{aligned} p(s, m, M, A, B | \text{data}) &= \int d\mu p(s, m, M, A, B, \mu | \text{data}) \\ &= \int dM_Z \left[\frac{d\mu}{dM_Z} \right] p(\text{data}|s, m, M, A, B, \mu) \\ &\simeq \mathcal{L}_{\text{rest}} \left[\frac{d\mu}{dM_Z} \right]_{\mu_Z} p(s, m, M, A, B, \mu_Z) . \end{aligned} \quad (2.2)$$

where μ_Z is the value of μ that reproduces the experimental value of M_Z for the given values of $\{s, m, M, A, B\}$, and $p(s, m, M, A, B, \mu)$ is the prior in the initial parameters (still undefined). This marginalized pdf can be written as

$$p(s, m, M, A, B | \text{data}) = 2 \mathcal{L}_{\text{rest}} \frac{\mu_0}{M_Z} \frac{1}{c_\mu} p(s, m, M, A, B, \mu_0) . \quad (2.3)$$

where $c_\mu = \left| \frac{\partial \ln M_Z^2}{\partial \ln \mu} \right|$ is the conventional Barbieri-Giudice measure [13, 14] of the degree of fine-tuning. Thus, the presence of this fine-tuning parameter in the denominator penalizes the regions of the parameter space with large fine-tuning, as desired.

As we will see in full detail in sect. 3, this is enough to make the high-scale region of the parameter space, say soft terms \gtrsim few TeV, statistically insignificant; which allows in turn to consider a wide range for the soft parameters (up to the very M_X). In consequence, the results of our analysis are essentially independent on the upper limits of the MSSM parameters, in contrast with previous studies.

2.2 Treatment of nuisance variables and Yukawa couplings

It is common in statistical problems that not all the parameters that define the system are of the same interest. In the problem at hand we are interested in determining the probabil-

ity maps for the MSSM parameters that describe the new physics, i.e. $\{m, M, A, B, \mu\}$, but not (or not at the same level) for the SM-like parameters, denoted by $\{s\}$. However, the *nuisance parameters* $\{s\}$ play an important role in extracting experimental consequences from the MSSM. The usual technique to eliminate nuisance parameters is simply marginalizing them, i.e. integrating the pdf (2.3) in the $\{s\}$ variables (for a review see ref. [19]). When the value of a nuisance parameter is in one-to-one correspondence to a high-quality experimental piece of information (included in $\mathcal{L}_{\text{rest}}$), this integration simply selects the “experimental” value of the nuisance parameter, which thus becomes (basically) a constant with no further statistical significance in the analysis. Note that in that case the prior on such nuisance parameter becomes irrelevant. In the MSSM, nuisance parameters of this class are the gauge couplings, $\{g_3, g, g'\}$, which thus can be extracted from the analysis (for an extended discussion see [18]).

In the pure SM a similar argument can be used to eliminate the Yukawa couplings, since they are in one-to-one correspondence to the quark and lepton masses. However, in the MSSM these masses depend also on the relative value of the two expectation values of the two Higgses, i.e. $\tan\beta \equiv v_2/v_1$. At tree-level, the up-type-quark, down-type-quark and charged lepton masses go like $m_u \sim y_u v \sin\beta$, $m_d \sim y_d v \cos\beta$ and $m_e \sim y_e v \cos\beta$ respectively; where $v^2 = 2(v_1^2 + v_2^2) = (246 \text{ GeV})^2$ is proportional to the Z mass squared. Note that $\tan\beta$ is a derived quantity, which is obtained upon minimization of the scalar potential $V(H_1, H_2)$, and thus takes different values at different points of the MSSM parameter space. This means that two viable MSSM models (with the same fermion masses) will have in general very different values of the Yukawa couplings, and thus the theoretical prior, $p(y)$, will play a relevant and non-ignorable role in their relative probability. Any Bayesian analysis of the MSSM amounts to an explicit or implicit assumption about the prior in the Yukawa couplings.

In previous Bayesian analyses of the MSSM the role of the Yukawa couplings was basically ignored: their values were just taken as needed to reproduce the experimental fermion masses, within uncertainties. Even if this procedure can be seen as “sensible”, it is worth wondering which kind of prior $p(y)$ corresponds to. As shown in ref. [18] this can be worked out by marginalizing the Yukawa couplings, using the experimental information about the fermion masses. Let us discuss briefly see how this works.

Just for the sake of the argument, let us approximate the associated likelihood as a product of Dirac deltas

$$\mathcal{L}_{\text{fermion masses}} = \delta(m_t - m_t^{\text{exp}}) \delta(m_b - m_b^{\text{exp}}) \dots \quad (2.4)$$

and the fermion masses by the tree-level expressions

$$m_t = \frac{1}{\sqrt{2}} y_t^{\text{low}} v s_\beta, \quad m_b = \frac{1}{\sqrt{2}} y_b^{\text{low}} v c_\beta, \quad \text{etc.} \quad (2.5)$$

where $s_\beta \equiv \sin\beta$, $c_\beta \equiv \cos\beta$ and y_i^{low} are the low-energy Yukawa couplings. Suppose further that $y_i^{\text{low}} = R_i y_i$, where y_i are the high-energy Yukawa couplings (and thus the input parameters) and the renormalization-group (RG) factor R_i does not depend on y_i

itself. Now, it is easy to work out the factor introduced in the pdf when the y_i variables are marginalized:

$$\begin{aligned} \int [dy_t dy_b \dots] p(y, m, M, A, B | \text{data}) &= \int [dy_t dy_b \dots] p(y) \delta(m_t - m_t^{\text{exp}}) \delta(m_b - m_b^{\text{exp}}) \dots \\ &\sim p(y) \left| \frac{dy_t}{dm_t} \right| \left| \frac{dy_b}{dm_b} \right| \dots = p(y) s_\beta^{-1} c_\beta^{-1} \dots \end{aligned} \quad (2.6)$$

where $p(y)$ denotes the prior in the Yukawa couplings (which we assume that factorizes from the other priors). Eq. (2.6) represents the footprint of the Yukawa couplings in the pdf. Now, taking logarithmically flat priors in the Yukawas, i.e. $p(y_i) \propto 1/y_i$, then the $s_\beta^{-1} c_\beta^{-1} \dots$ factors get cancelled. This is therefore the prior implicitly assumed in the previous Bayesian analyses, and the one we will adopt in this paper. Remarkably, for independent reasons, we find the logarithmically flat prior for Yukawa couplings a most sensible choice. Certainly there is no convincing origin for the experimental pattern of fermion masses, and thus of Yukawa couplings. However it is a fact that these come in very assorted orders of magnitude (from $\mathcal{O}(10^{-6})$ for the electron to $\mathcal{O}(1)$ for the top), suggesting that the underlying mechanism may produce Yukawa couplings of different orders with similar efficiency; and this is the meaning of a logarithmic prior.

Of course the above discussion is oversimplified. The physical (pole) masses include radiative corrections. Besides, the RG factor R_i for the top Yukawa coupling has an important dependence on the Yukawa itself. These subtleties have been incorporated to the full analysis.

2.3 Variables for the MSSM scan

Although our initial set of variables is $\{m, M, A, B, \mu, s\}$, and this is the one on which we have to set our theoretical prior, for the purposes of scanning the MSSM parameter space it is much more convenient to trade some of them by other parameters with more direct phenomenological significance. We have already seen that it is worth to trade μ by M_Z , which is automatically integrated out. Similarly, we have seen that the Yukawa couplings are nuisance variables that are profitably traded by the physical fermion masses and easily integrated out. Besides all this, it is highly advantageous to trade the initial B -parameter by the derived $\tan \beta$ parameter. The main reason is that for a given viable choice of $\{m, M, A, \tan \beta\}$, there are exactly two values of μ (with opposite sign and the same absolute value at low energy) leading to the correct value of M_Z . Thus working in one of the two (positive and negative) branches of μ , each point in the $\{m, M, A, \tan \beta\}$ space corresponds exactly to one model, whereas a point in the $\{m, M, A, B\}$ space may correspond to several models, introducing a conceptual and technical complication in the analysis.

Consequently we should compute the whole Jacobian, J , of the transformation

$$\{\mu, y_t, B\} \rightarrow \{M_Z, m_t, t\}, \quad t \equiv \tan \beta. \quad (2.7)$$

Then the *effective prior* in the new variables becomes

$$p_{\text{eff}}(g_i, m_t, m, M, A, \tan \beta) \equiv J|_{\mu=\mu_Z} p(g_i, y_t, m, M, A, B, \mu = \mu_Z) \quad (2.8)$$

where we have already marginalized M_Z using the associated likelihood $\sim \delta(M_Z - M_Z^{\text{exp}})$ (recall that μ_Z is the value of μ that reproduces the experimental M_Z .) In eqs.(2.7, 2.8) we have made explicit the dependence just on the top Yukawa coupling and mass, but for other fermions goes the same.

So, to prepare the scan in the new variables we need an explicit evaluation of the Jacobian factor. For that we must know the dependence of the old variables on the new ones. This dependence is extracted from the minimization equations of the Higgs scalar potential, $V(H_1, H_2)$ (which connect $\{\mu, B\}$ with $\{M_Z, \tan \beta\}$), and from the relation between the yukawa couplings and fermion masses. This dependences, even when radiative corrections are included, have the form

$$\mu = f(M_Z, y, t), \quad y = g(M_Z, m_t, t), \quad B = h(\mu, y, t), \quad (2.9)$$

where f, g, h are well defined functions, for which we give approximate analytical expressions below. Here we have made explicit only the dependence on the variables involved in the change of variables (2.7). Note that y depends on M_Z since $v \propto M_Z$. In consequence

$$J = \frac{\partial f}{\partial M_Z} \frac{\partial g}{\partial m_t} \frac{\partial h}{\partial t} . \quad (2.10)$$

where the factor $\partial f/\partial M_Z$ carries essentially the fine-tuning penalization discussed in subsect. 2.1.

For the numerical analysis we have evaluated J using the `SoftSusy` code [20] which implements the full one-loop contributions and leading two-loop terms to the tadpoles for the electroweak breaking conditions with parameters running at two-loops. This essentially corresponds to the next-to-leading log approximation.

However, it is possible to give an analytical and quite accurate expression of J by working with the tree-level potential with parameters running at one-loop (i.e. essentially the leading log approximation). Then

$$\mu_{\text{low}}^2 = \frac{m_{H_1}^2 - m_{H_2}^2 t^2}{t^2 - 1} - \frac{M_Z^2}{2} \quad (2.11)$$

$$B_{\text{low}} = \frac{s_{2\beta}}{2\mu_{\text{low}}} (m_{H_1}^2 + m_{H_2}^2 + 2\mu_{\text{low}}^2) \quad (2.12)$$

$$y_{\text{low}} = \frac{m_t}{v s_\beta} . \quad (2.13)$$

Here the “low” subscript indicates that the quantity is evaluated at low scale (more precisely, at a representative supersymmetric mass, such as the geometric average of the stop

masses). The soft masses $m_{H_i}^2$ are also understood at low scale. For notational simplicity, we have dropped the subscript t from the Yukawa coupling. Note that all these low-energy quantities contain an implicit dependence on the top Yukawa coupling through the corresponding RG equations. Besides,

$$\mu_{\text{low}} = R_\mu(y)\mu, \quad B_{\text{low}} = B + \Delta_{RG}B(y), \quad y_{\text{low}} \simeq \frac{yE(Q_{\text{low}})}{1 + 6yF(Q_{\text{low}})}, \quad (2.14)$$

where $R_\mu(y), \Delta_{RG}B(y)$ are definite functions of y ; Q is the renormalization scale, $F = \int_{Q_{\text{high}}}^{Q_{\text{low}}} E \ln Q$, and $E(Q)$ is a definite function that depends just on the gauge couplings [21]. Plugging eqs.(2.11–2.14) into eqs.(2.9, 2.10) it is straightforward to get an explicit expression for J . Plugging the latter back into eq. (2.8) we get an approximate form for the effective prior

$$p_{\text{eff}}(m_t, m, M, A, \tan \beta) \propto \left[\frac{E}{R_\mu^2} \right] \frac{y}{y_{\text{low}}} \frac{t^2 - 1}{t(1 + t^2)} \frac{B_{\text{low}}}{\mu_Z} p(m, M, A, B, \mu = \mu_Z), \quad (2.15)$$

where we have taken a logarithmically flat prior for the Yukawa couplings (i.e. $p(y_i) \propto y_i^{-1}$), as discussed above. This is the prior to be used when the MSSM parameter space is scanned in the usual variables $\{m, M, A, \tan \beta\}$ and μ is taken as required to reproduce the correct EW breaking (the information about M_Z^{exp} is thus automatically incorporated). Let us stress that its form stems just from the relation between the initial variables and the phenomenological ones, indicated in eq.(2.7), and it is not “subjective” at all. Besides, the prefactor in the r.h.s. of eq.(2.15) (which is essentially the Jacobian) is valid for any MSSM, not just the CMSSM. The subjectivity lies in the $p(m, M, A, B, \mu)$ piece, i.e. the prior in the initial parameters, for which we have still to make a choice. Furthermore, the prefactor in eq.(2.15) contains the above-discussed penalization of fine-tuned regions, something that may be not so obvious, but that will become clear in sect. 3. Finally, the form of the prefactor implies an effective penalization of large $\tan \beta$, reflecting the smaller statistical weight of this possibility [it occurs only for small B_{low} , as is clear from eq.(2.12)].

2.4 Priors in the initial parameters

The choice of the prior in the initial parameters, $\{m, M, A, B, \mu\}$ must reflect our knowledge about them, before consideration of the experimental data (to be included in the likelihood piece). In our case, we have already made some non-trivial, though quite reasonable, assumptions about them, namely the hypothesis of universality of the soft terms (which is supported by the strong constraints from FCNC processes) at a very high scale (this restricts the analysis to scenarios where the transfer of SUSY breaking is suppressed by a high scale, as happens e.g. in models with gravity-mediated SUSY breaking).

To go further we must consider the dynamical origin of the parameters. Four of them, $\{m, M, A, B\}$, are soft SUSY-breaking parameters. They typically go like $\sim F/\Lambda$, where F is the SUSY breaking scale, which corresponds to the dominant VEV among the auxiliary fields in the SUSY breaking sector (it can be an F -term or a D -term) and Λ is the messenger scale, associated to the interactions that transmit the breaking to the observable

sector. Since the soft-breaking terms share a common origin it is logical to assume that their sizes are also similar. Of course, there are several contributions to a particular soft term, which depend on the details of the superpotential, the Kähler potential and the gauge kinetic function of the complete theory (see e.g. ref. [22]). So, it is reasonable to assume that a particular soft term can get any value (with essentially flat probability) of the order of the typical size of the soft terms or below it. There are special cases, like split SUSY scenarios, where the soft terms can be classified in two groups that feel differently the breaking of SUSY. In those instances, the priors should also be considered in two separate groups. But those cases are out of the scope of the present analysis, which is focussed on the simplest, most conventional and less baroque framework, which consists of a common SUSY breaking origin and transmission for all the soft terms. The μ -parameter is not a soft term, but a parameter of the superpotential. However, it is desirable that its size is related (e.g. through the Giudice-Masiero mechanism [23]) to the SUSY breaking scale. Otherwise, one has to face the so-called μ -problem, i.e. why should be the size of μ similar to the soft terms', as is required for a correct electroweak breaking (see eq.(2.11)). Thus, concerning the prior, we can consider μ on a similar foot to the other soft terms.

Now, we are going to make the previous discussion more quantitative. Let us call M_S the typical size of the soft terms in the observable sector, $M_S \sim F/\Lambda$. Then, we define the ranges of variation of the initial parameters as

$$\begin{aligned}
-qM_S &\leq B \leq qM_S \\
-qM_S &\leq A \leq qM_S \\
0 &\leq m \leq qM_S \\
0 &\leq M \leq qM_S \\
0 &\leq \mu \leq qM_S
\end{aligned} \tag{2.16}$$

where q is an $\mathcal{O}(1)$ factor. We have considered here the branch of positive μ . For the negative one we simply replace $\mu \rightarrow -\mu$. We have taken the same q for all the parameters, since we find no reason to make distinctions among them. Note that we can take $q = 1$ with no loss of generality, provided M_S is allowed to vary in the range $0 \leq M_S \leq \infty$. In practice, to avoid divergences in the priors, we have to take a finite range for M_S , say

$$M_S^0 \leq M_S \leq M_X, \quad M_S^0 \sim 10 \text{ GeV} \tag{2.17}$$

Nevertheless, the values of the upper and lower limits of the M_S range are going to be irrelevant, as it will become clear soon. Consequently, we can still take $q = 1$.

We have discussed the ranges of the parameters, but not the shape of the priors. As already stated, we find reasonable to assume (conveniently normalized) flat priors for the soft parameters inside the ranges (2.16), i.e.

$$p(m) = p(M) = p(\mu) = \frac{1}{M_S}, \quad p(A) = p(B) = \frac{1}{2M_S} \tag{2.18}$$

Still we have to decide what is the prior in M_S , and it is at this point where we have to take the decision of assuming a flat or logarithmic prior in the scale of SUSY breaking. We

have considered the two possibilities throughout the paper. The comparison of the results from both choices will give us a measure of the prior-dependence of the analysis.

Logarithmic prior

Let us start assuming a logarithmic prior in M_S , which we consider the most reasonable option, since it amounts to consider all the possible orders of magnitude of the SUSY breaking in the observable sector on the same foot (this occurs e.g. in conventional SUSY breaking by gaugino condensation in a hidden sector). Then,

$$p(M_S) = N_{M_S} \frac{1}{M_S}, \quad (2.19)$$

where N_{M_S} is a normalization constant, which turns out to be completely irrelevant. Now, we can marginalize M_S , which thus disappears completely from the subsequent analysis, leaving a prior which depends just on the $\{m, M, A, B, \mu\}$ parameters²:

$$\begin{aligned} p(m, M, A, B, \mu) &= \frac{N_{M_S}}{4} \int_{\max\{m, M, |A|, |B|, \mu, M_S^0\}}^{M_X} \frac{1}{M_S^6} dM_S \\ &= \frac{N_{M_S}}{20} \left[\frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^5} - \frac{1}{M_X^5} \right] \\ &\simeq \frac{N_{M_S}}{20} \frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^5} \end{aligned} \quad (2.20)$$

Of course, the prefactor is just an irrelevant normalization constant. Note that we have neglected the $1/M_X^5$ term, which simply forced the prior to strictly vanish in the M_X limit. This effect is only appreciable when one of the parameters is close to M_X , otherwise it is completely negligible (note the fifth power in the denominators). On the other hand, as mentioned in subsect. 2.1 and will become clear soon, once the EW breaking is incorporated to the analysis, regions of very large initial parameters become irrelevant. In consequence, eq.(2.20) is an excellent approximation. Note that the value of M_X disappears from the analysis. The value of the lower limit on M_S , i.e. M_S^0 , is also irrelevant. Note that its presence in the denominator of eq.(2.20) avoids the prior to diverge when the parameters are very small. This “regulating” effect is only felt when *all* the parameters are below M_S^0 . However, we know that this region will be killed by the experimental data once they are taken into account (through the likelihood piece in the pdf). E.g. the upper bounds on chargino masses require $|\mu| \gtrsim 100$ GeV. Hence, the value of M_S^0 plays no relevant role, apart from the formal regularization of the prior. Let us recall that the above prior (2.20) is the one to be plugged in eq.(2.8) (or in the approximated expression (2.15)) to get the effective prior in the scan parameters.

It is funny to compare the prior of eq.(2.20) with a “more conventional” logarithmic prior, i.e. $p(m, M, A, B, \mu) \propto 1/(m, M, A, B, \mu)$. First of all, the “more conventional” prior is not regulated unless one imposes that the parameters should not go below some low-scale

²This procedure is a “hierarchical Bayesian technique”, first used in ref. [4], but using complicated functions that were not possible to integrate analytically.

(or that the prior does not behave logarithmically flat in that region). But then the results are sensitive to the cut-off scale chosen. Note that the prior for phenomenologically viable points, with e.g. very small A and large μ (thus avoiding the constraints from chargino masses), will depend on the precise treatment of this region. Apart from this annoyance, the conventional logarithmic prior treats the parameters as uncorrelated objects. This produces non-realistic distortions. E.g. a point of the parameter space where some parameters are very large, but the others are very small, can have a value of the prior (i.e. an assigned probability) larger than another point where all the parameters are $\mathcal{O}(\text{TeV})$. However, this goes against the expectative that all the initial parameters are likely to have similar sizes, as they share a common dynamical origin. In other words, it is not sensible to increase the prior probability (in a very significant amount) just because one of the parameters is abnormally small, compared to the others. These problems are nicely avoided by the simple prior (2.20), reflecting the way it has been constructed.

To finish this discussion, let us note that the prior (2.20) does not have the form of a product of individual priors defined for each parameter. Still, we can get the form of the prior for just one parameter, marginalizing the others before including any experimental information. For instance, the prior in the gaugino mass, M , is obtained by marginalizing in m, A, B, μ , leading to

$$\mathcal{P}(M) \propto \frac{1}{\max\{M, M_S^0\}} \quad (2.21)$$

where we have neglected $\sim 1/M_X$ contributions. This has indeed the form of a logarithmically flat prior. Of course, similar individual priors are obtained for the other parameters.

Flat prior

We can now repeat the previous analysis, assuming a flat prior for M_S , which amounts to consider all the values of the SUSY breaking in the observable sector on the same foot. Hence we maintain the ranges for the parameters, eqs. (2.16), (2.17), and the flat priors inside those ranges, eq. (2.18). We just replace the logarithmically flat prior in M_S , eq. (2.19), by a flat prior

$$p(M_S) = N_{M_S} \quad (2.22)$$

where N_{M_S} is an irrelevant normalization constant $\sim 1/M_X$. Again, to obtain the prior in the $\{m, M, A, B, \mu\}$ variables, we marginalize in M_S . The previous result (2.20) becomes now

$$p(m, M, A, B, \mu) \propto \frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^4} \quad (2.23)$$

where, once more (and for the same reasons) we have neglected a $1/M_X^4$ contribution. The difference with eq.(2.20) is that now we have one power less in the denominator. Again,

the prior (2.23) is the one to be plugged back into eq.(2.8) in order to get the effective prior in the scan parameters.

We can also repeat the exercise of obtaining the prior for an individual parameter, say M , by marginalizing the others. In this case, the previous equation (2.21) becomes

$$\mathcal{P}(M) \sim \ln \frac{M_X}{\max\{M, M_S^0\}} , \quad (2.24)$$

In essence this is a flat prior in M , as it does not change much along orders of magnitude. E.g. in the $100 \text{ GeV} \leq M \leq 4 \text{ TeV}$ range it just changes a 13%. Again, the other parameters go in a similar way.

3. High-energy vs Low-energy regions and the EW breaking

At first sight it may seem that the assumption of a logarithmic prior, see eqs.(2.20, 2.21), amounts to a strong preference for the low-energy region of the parameter space, i.e. for $\{m, M, A, B, \mu\}$ not far from the EW scale. However, this is not true. We may ask the following question: What is a priori the relative probability that a parameter, say M , lies in the low-energy (accessible to the LHC) region, $100 \text{ GeV} \lesssim M \lesssim 2 \text{ TeV}$ versus the chance that it lies at a higher scale, $2 \text{ TeV} \lesssim M \lesssim M_X$. Using eq.(2.21), it is clear that this relative probability is

$$\frac{\mathcal{P}(100 \text{ GeV} \leq M \leq 2 \text{ TeV})}{\mathcal{P}(2 \text{ TeV} \leq M \leq M_X)} \simeq \frac{1}{12} \quad (3.1)$$

(in an obvious notation). I.e. in the initial set-up the most probable situation is that SUSY escapes LHC detection, even with logarithmic prior. Note that this is not so if one cuts-off the ranges of the parameters at a few TeV, as is very usually done, but we allow them to vary all the way up to M_X . Of course, the situation is much more dramatic for a flat prior. Using eq.(2.24) we see that in that case

$$\frac{\mathcal{P}(100 \text{ GeV} \leq M \leq 2 \text{ TeV})}{\mathcal{P}(2 \text{ TeV} \leq M \leq M_X)} \simeq 3 \times 10^{-12} . \quad (3.2)$$

Hence, the flat-prior set-up assigns a negligible initial probability to LHC detection.

Fortunately things change for better as soon as we incorporate the experimental information about the size of the EW breaking, i.e. M_Z^{exp} . We have already discussed in subsect. 2.1 how M_Z^{exp} can be used to marginalize μ , leaving a footprint of fine-tuning penalization. Now we are going to be more precise. We will take the effective prior in the scan variables, given by eq.(2.8) or by the approximate expression (2.15). Recall that these expressions already incorporate the experimental information on M_Z and the marginalization of μ .

Now, we can evaluate once more the relative probability between the low- and high-energy regions (say for the M parameter again), but with this effective prior, i.e. incorporating the information about the EW breaking scale. For the sake of clarity we present now an analytical discussion, with some approximations, that gives correctly the essential

results and allows to show the physical reasons behind them. At the end we will present the numerical results.

Hence, we have to marginalize the $\{m, A, \tan \beta\}$ parameters, since the μ -parameter has already been marginalized. Let us first perform the integration in $\{m, A\}$. Note that for a given value $M = M_1$, and $\tan \beta$ fixed, only a portion of the $\{m, A\}$ plane will be able to accommodate, by adjusting the μ -parameter, the required EW breaking. Let us call this region R_1 . Therefore the integration is only extended to R_1

$$\mathcal{P}(M_1, \tan \beta) \sim \int_{R_1} dm dA p_{\text{eff}}(m_t, m, M, A, \tan \beta) \quad (3.3)$$

where $p_{\text{eff}}(m_t, m, M, A, \tan \beta)$ is given by eq.(2.8) or eq.(2.15). We have to compare this probability with the one for a different gaugino mass, say $M_2 > M_1$ (we keep $\tan \beta$ fixed). The expression for $\mathcal{P}(M_2, \tan \beta)$ is completely analogous to eq.(3.3). The only subtle point is what is the new allowed region, R_2 . An approximate way to determine it is the following. In almost all the parameter space the squared SUSY-parameters, $\{m^2, M^2, A^2, B^2, \mu^2\}$ are much larger than M_Z^2 . Therefore, the combination of them producing the correct value of M_Z is almost identical to the one producing $M_Z = 0$. So we can approximate R_1 and R_2 as the regions giving $M_Z = 0$. Now, if we neglect for a moment RG effects, it is clear that for each point $\{m_1, A_1\} \in R_1$ there is another point $\{m_2, A_2\} \in R_2$, producing the same breaking, given by $m_2 = \frac{M_2}{M_1} m_1$, $A_2 = \frac{M_2}{M_1} A_1$, $\mu_2 = \frac{M_2}{M_1} \mu_1$, $B_2 = \frac{M_2}{M_1} B_1$. In other words, $R_2 \sim \frac{M_2}{M_1} R_1$. RG effects do not in principle modify this relation since they are proportional to the very soft terms. However, there is a residual effect: since the running goes from M_X (where the SUSY parameters are defined) until the scale where the EW breaking is evaluated (\sim stop masses), there is logarithmic correction, $\propto \log(M_2/M_1)$, which would slightly modify the shape of R_2 . On the other hand, there will be points in R_2 that will go out of the allowed ranges of the parameters. I.e., R_2 will be slightly smaller than $\frac{M_2}{M_1} R_1$. This means that we are slightly overestimating the weight of the high-scale parameter space, which is a conservative attitude for this discussion. Now we can express the integration in the R_2 region in terms of the R_1 one,

$$\begin{aligned} \mathcal{P}(M_2, \tan \beta) &\sim \int_{R_2} dm_2 dA_2 p_{\text{eff}}(m_2, M_2, A_2, \mu_2, \tan \beta) \\ &= \int_{R_1} dm_1 dA_1 \left(\frac{M_1}{M_2}\right)^3 p_{\text{eff}}(m_1, M_1, A_1, \mu_1, \tan \beta) \\ &= \left(\frac{M_1}{M_2}\right)^3 \mathcal{P}(M_1, \tan \beta) , \end{aligned} \quad (3.4)$$

Here have used the fact that, when we assume a logarithmic initial prior, the effective prior, p_{eff} , scales as

$$p_{\text{eff}}(m_2, M_2, A_2, \mu_2, \tan \beta) = \left(\frac{M_1}{M_2}\right)^5 p_{\text{eff}}(m_1, M_1, A_1, \mu_1, \tan \beta) . \quad (3.5)$$

This can be easily noticed from the approximate expression of p_{eff} in eq.(2.15). This relation is exact, essentially, up to small RG-effects in the $\frac{B_{\text{low}}}{\mu_Z}$ factor involved in (2.15).

The last step is to marginalize $\tan\beta$. But this will not affect the relative probability we are interested in, since the factor obtained by integrating $\tan\beta$ is identical for M_1 and for M_2 . Alternatively, we can leave $\tan\beta$ fixed at some arbitrary value. The important point is that the relative probability goes like

$$\frac{\mathcal{P}(M_2, \tan\beta)}{\mathcal{P}(M_1, \tan\beta)} \sim \frac{M_1^3}{M_2^3}. \quad (3.6)$$

In other words,

$$\mathcal{P}(M, \tan\beta) \sim \frac{1}{M^3} \quad (3.7)$$

This should be compared with the $\mathcal{P}(M) \sim \frac{1}{M}$ behaviour obtained in eq.(2.21), when the experimental M_Z was not taken into account. We see that the pdf in M has gained two powers of M in the denominator. This is the fine-tuning penalization that arises on its own from the Bayesian analysis. Now the relative probability of the low-energy (accessible to the LHC) region versus the probability of a higher scale becomes

$$\frac{\mathcal{P}(100 \text{ GeV} \leq M \leq 2 \text{ TeV})}{\mathcal{P}(2\text{TeV} \leq M \leq M_X)} \simeq \left(\frac{2 \text{ TeV}}{100 \text{ GeV}} \right)^2 \sim 10^2 \quad (3.8)$$

to be compared with eq.(3.1) before including the EW breaking in the analysis. In conclusion, once the EW breaking is correctly incorporated in the Bayesian analysis (but not before!), 99% of the probability lives in the low-energy (LHC-relevant) region of the parameter space. Note that this is achieved without invoking other kinds of constraints (like Dark Matter or $g-2$ constraints) that are often used to set the scale of the soft terms not far from the EW scale. Hence, the main reason to believe that SUSY should be accessible at LHC scales comes from the EW breaking itself, i.e. the original motivation for phenomenological SUSY. We find this result very satisfactory. Needless to say that for the other parameters the results go in a completely analogous way.

We have discussed the impact of the EW breaking scale when a logarithmic prior is used. For a flat prior, it is straightforward to repeat the above discussion, taking into account that the prior $p(m, M, A, B, \mu)$ [and thus $p_{\text{eff}}(m, M, A, \mu, \tan\beta)$] has one power of mass less in the denominator. Therefore at the end of the day we arrive at

$$\mathcal{P}(M, \tan\beta) \sim \frac{1}{M^2} \quad (3.9)$$

to be compared with the almost flat behaviour before including the experimental EW information, see eq.(2.23). Consequently, the relative probability of the low-energy and high-energy regions of the parameter space, for a flat prior, becomes

$$\frac{\mathcal{P}(100 \text{ GeV} \leq M \leq 2 \text{ TeV})}{\mathcal{P}(2\text{TeV} \leq M \leq M_X)} \simeq \frac{2 \text{ TeV}}{100 \text{ GeV}} \simeq 10 \quad (3.10)$$

So, even assuming a flat prior for the typical size of the soft breaking terms, up to the M_X scale, we see that the EW breaking is sufficient to put 90% of the total probability in

the LHC-interesting region. This contrasts strongly with previous analysis, and, again, we consider it very satisfactory.

We have checked the previous arguments by performing the analysis in a numerical way. For the posterior samples we adopt the MultiNest [24] algorithm as implemented in the `SuperBayeS` code [25]. It is based on the framework of Nested Sampling, recently invented by Skilling [26, 27]. MultiNest has been developed in such a way as to be an extremely efficient sampler even for likelihood functions defined over a parameter space of large dimensionality with a very complex structure as it is the case of the CMSSM. The main purpose of the Multinest is the computation of the Bayesian evidence and its uncertainty but it produces posterior inferences as a by-product. For the marginalization procedure we have used the above-discussed ranges for our priors, i.e. from 0 to M_X for m , M and $|A|$. Besides, we have used $2 < \tan\beta < 62$. The lower limit comes from present bound on the Higgs mass. The upper one is to avoid Yukawa couplings beyond the perturbative regime.

In Fig. 1 the red line shows the prior in M (the other parameters are marginalized) using logarithmic (left panel) or flat (right panel) initial prior for the scale of SUSY breaking in the observable sector (M_S). The blue bar distributions show the pdf once the EW breaking is incorporated in the analysis, i.e. the effective prior in the scan variables, p_{eff} , see eq.(2.8) and the approximate form (2.15). The logarithmic scale in the horizontal axes allows to see that most of the probability, which initially lies in the high-energy region (M above the TeV scale), flows dramatically into the low-energy region once the EW breaking is considered. Actually, most of the probability falls inside the LHC discovery reach (even with with just 1 fb^{-1} [28, 29]). Quantitatively, the results are in good agreement with the previous discussion.

Besides making the high-energy parameter space quite irrelevant, the EW breaking has another dramatic effect, which is visible in Fig. 1. Namely, the probability distributions (pdfs) based on a logarithmic or on a flat prior are quite similar, *after* the incorporation of the EW scale. That is, the favoured regions of the parameter space are quite independent of the choice of the prior. Normally, a behaviour of this kind is attributed to the fact that the data are powerful enough to select a region of the parameter space, so that the general expression of the pdf, eq.(1.1), is dominated by the likelihood piece. However this is not the case here. As a matter of fact, concerning the likelihood, there are points with arbitrary large parameters that are as good as the low-energy ones, since they correctly reproduce M_Z^{exp} , the only data so far considered. The low-energy region is preferred because it is statistically much more significant, as we have discussed above. But this is a *Bayesian* effect, non-existent in a frequentist analysis. Therefore the situation is very good from the Bayesian point of view: the results are quite independent from the type of prior, but to see the preferred regions we need the Bayesian procedures.

To finish this section, let us note that the previous statistical argument supports low-energy supersymmetry breaking (in the observable sector), even in a landscape scenario. In other words, even if there were many more vacua with supersymmetry breaking at large

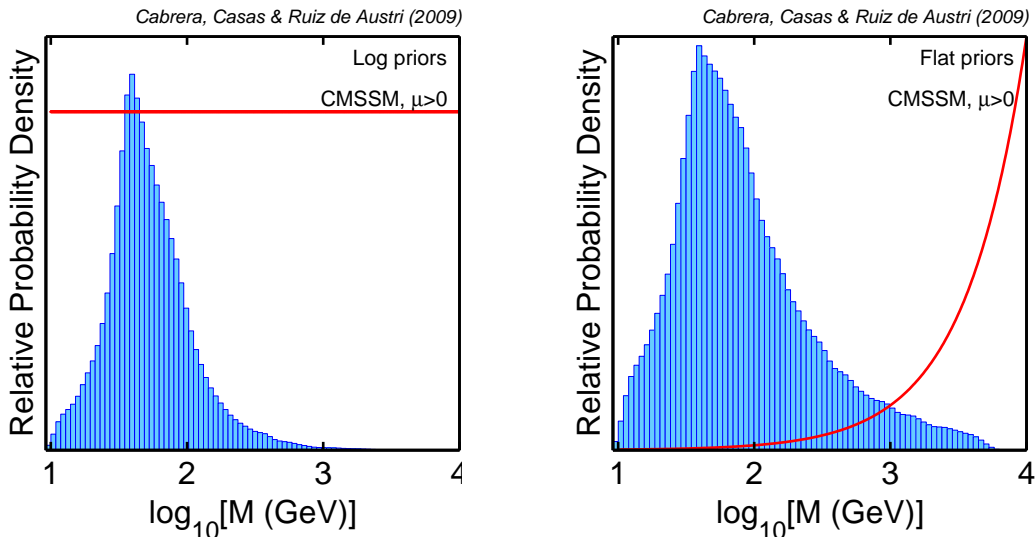


Figure 1: 1D marginalized posterior probability distribution of the M parameter for logarithmic (left panel) and flat (right panel) priors in the $\mu > 0$ case for a scan including the information about the EW breaking (M_Z^{exp}). The red lines represent the marginalized prior. All given in arbitrary units.

scale, most of realistic vacua would correspond to low-energy supersymmetry breaking, for rather generic a-priori distributions of all possible vacua (for related work in this line see [30]).

4. Experimental Constraints

In this section we will incorporate all the relevant experimental information to the likelihood piece of the probability distribution (all but M_Z^{exp} , which has already been taken into account). This amounts to include many experimental observables and bounds, with their error bars, and to calculate the predictions for them in the MSSM.

As originally demonstrated in [1, 3], the values of the relevant SM-like parameters (*nuisance parameters*) can strongly influence some of the CMSSM predictions. For our analysis we take the set

$$\left\{ M_t, m_b(m_b)^{\overline{MS}}, \alpha_{\text{em}}(M_Z)^{\overline{MS}}, \alpha_s(M_Z)^{\overline{MS}} \right\}, \quad (4.1)$$

where M_t is the pole top quark mass, while the other three parameters (the bottom mass, the electromagnetic and the strong coupling constants) are all evaluated in the \overline{MS} scheme at the indicated scales. The constraints on the SM nuisance parameters are given in Table 1.

On the other hand, there are the experimental values of accelerator and cosmological observables, which are listed in Table 2. Instead of including all this information at once and show the results, we find more illustrative to do it in several steps. This will allow to show the effect of the various types of data on the probability distributions (which are

SM (nuisance) parameter	Mean value μ	Uncertainty σ (exper.)	Ref.
M_t	172.6 GeV	1.4 GeV	[31]
$m_b(m_b)^{\overline{MS}}$	4.20 GeV	0.07 GeV	[32]
$\alpha_s(M_Z)^{\overline{MS}}$	0.1176	0.002	[32]
$1/\alpha_{\text{em}}(M_Z)^{\overline{MS}}$	127.955	0.03	[33]

Table 1: Experimental mean μ and standard deviation σ adopted for the likelihood function for SM (nuisance) parameters, assumed to be described by a Gaussian distribution.

sometimes opposite). On the other hand, not all the data are on the same foot of quality and reliability and it is convenient not to mix them from the beginning.

In order to avoid a proliferation of plots we examine first the positive μ branch. In the next section we will show the relevant plots and results for negative μ and perform a comparison of the relative probability of the two possibilities.

4.1 EW and B-physics observables, and limits on particle masses

We start by considering the most reliable and robust pieces of experimental information: EW and B(D)-physics observables and lower bounds on the masses of supersymmetric particles and the Higgs mass. The complete list of the observables of this kind used in our analysis is given in Table 2 (all the entries except those concerning a_μ and dark matter constraints).

To calculate the MSSM spectrum we use `SoftSusy` [20]. We discard points suffering from unphysicalities: no self-consistent solutions to the RGEs, no EW breaking and tachyonic states. Furthermore, we require the neutralino to be the lightest supersymmetric particle (LSP) in order to be an acceptable dark matter candidate. The latter condition might be relaxed, as discussed in subsect. 4.3 below. For details on how the EW observables are computed the reader is directed to [3].

Roughly speaking, the MSSM parameter space is quite unconstrained by EW (LEP) observables, except for quite small values of the SUSY soft-terms (i.e. when the SUSY corrections are sizeable). This is logical. As it is well known, the MSSM is free from the Little Hierarchy problem, understood as the tension between LEP observables and the need of new physics at $\mathcal{O}(\text{TeV})$ scales to avoid the hierarchy problem [34]. This is because R-parity prevents from tree-level SUSY contributions to higher order SM operators. In consequence, unless supersymmetric masses are quite small, the effect of SUSY on LEP observables is not important.

Concerning B-physics observables, the branching ratio for the $B \rightarrow X_s \gamma$ decay (the most important one) has been computed with the numerical code `SusyBSG` [35] using the full NLO QCD contributions, including the two-loop calculation of the gluino contributions presented in [36] and the results of [37] for the remaining non-QCD $\tan \beta$ -enhanced contributions. The supersymmetric contributions to $b \rightarrow s \gamma$ grow with decreasing masses

Observable	Mean value	Uncertainties		ref.
	μ	σ (exper.)	τ (theor.)	
M_W	80.398 GeV	27 MeV	15 MeV	[39]
$\sin^2 \theta_{\text{eff}}$	0.23149	17×10^{-5}	15×10^{-5}	[39]
$\delta a_\mu \times 10^{10}$ (e^+e^-)	29.5	8.8	2.0	[40]
$\delta a_\mu \times 10^{10}$ (τ)	14.8	8.2	2.0	[41]
ΔM_{B_s}	17.77 ps $^{-1}$	0.12 ps $^{-1}$	2.4 ps $^{-1}$	[42]
$BR(\overline{B} \rightarrow X_s \gamma) \times 10^4$	3.52	0.33	0.3	[43]
$\frac{BR(B_u \rightarrow \tau \nu)}{BR(B_u \rightarrow \tau \nu)_{SM}}$	1.28	0.38	-	[43]
$\Delta_{0-} \times 10^2$	3.6	2.65	-	[44]
$\frac{BR(B \rightarrow D \tau \nu)}{BR(B \rightarrow D e \nu)} \times 10^2$	41.6	12.8	3.5	[45]
R_{l23}	1.004	0.007	-	[46]
$BR(D_s \rightarrow \tau \nu) \times 10^2$	5.7	0.4	0.2	[47]
$BR(D_s \rightarrow \mu \nu) \times 10^3$	5.8	0.4	0.2	[47]
$\Omega_\chi h^2$	0.1099	0.0062	$0.1 \Omega_\chi h^2$	[48]
	Limit (95% CL)		τ (theor.)	ref.
$BR(\overline{B}_s \rightarrow \mu^+ \mu^-)$	$< 5.8 \times 10^{-8}$		14%	[49]
m_h	> 114.4 GeV (SM-like Higgs)		3 GeV	[50]
ζ_h^2	$f(m_h)$ (see text)		negligible	[50]
$m_{\tilde{q}}$	> 375 GeV		5%	[32]
$m_{\tilde{g}}$	> 289 GeV		5%	[32]
other sparticle masses	As in table 4 of ref. [3].			

Table 2: Summary of the observables used in the analysis to constrain the CMSSM parameter space. Upper part: Observables for which a positive measurement has been made. $\delta a_\mu = a_\mu^{\text{exp}} - a_\mu^{\text{SM}}$ denotes the discrepancy between the experimental value and the SM prediction of the anomalous magnetic moment of the muon $(g - 2)_\mu$. Lower part: Observables for which only limits currently exist.

of the supersymmetric particles *and* with increasing $\tan \beta$. However the effect is never dramatic for masses around 1 TeV or larger. The other B(D)-physics observables summarized in Table 2 have been computed with the code **SuperIso** (for details on the computation of the observables see [38] and references therein). Both codes have been integrated into **SuperBayes**.

Experimental bounds used in the analysis are indicated in the second part of Table 2. These include bounds on supersymmetric masses (squarks, sleptons, gluinos, charginos, neutralinos) and the Higgs mass. In general, the constraints on supersymmetric masses tend obviously to cut off the region of the parameter space with too small values of m, M . On top of this, the bound on the Higgs mass is most relevant, and deserves special attention, as we are about to see. For details on how the likelihood is computed we refer to ref. [3].

Fig. 2 (upper panels) show the pdf for the gaugino mass parameter, M , once all this experimental information is incorporated. Again, the left (right) panels correspond

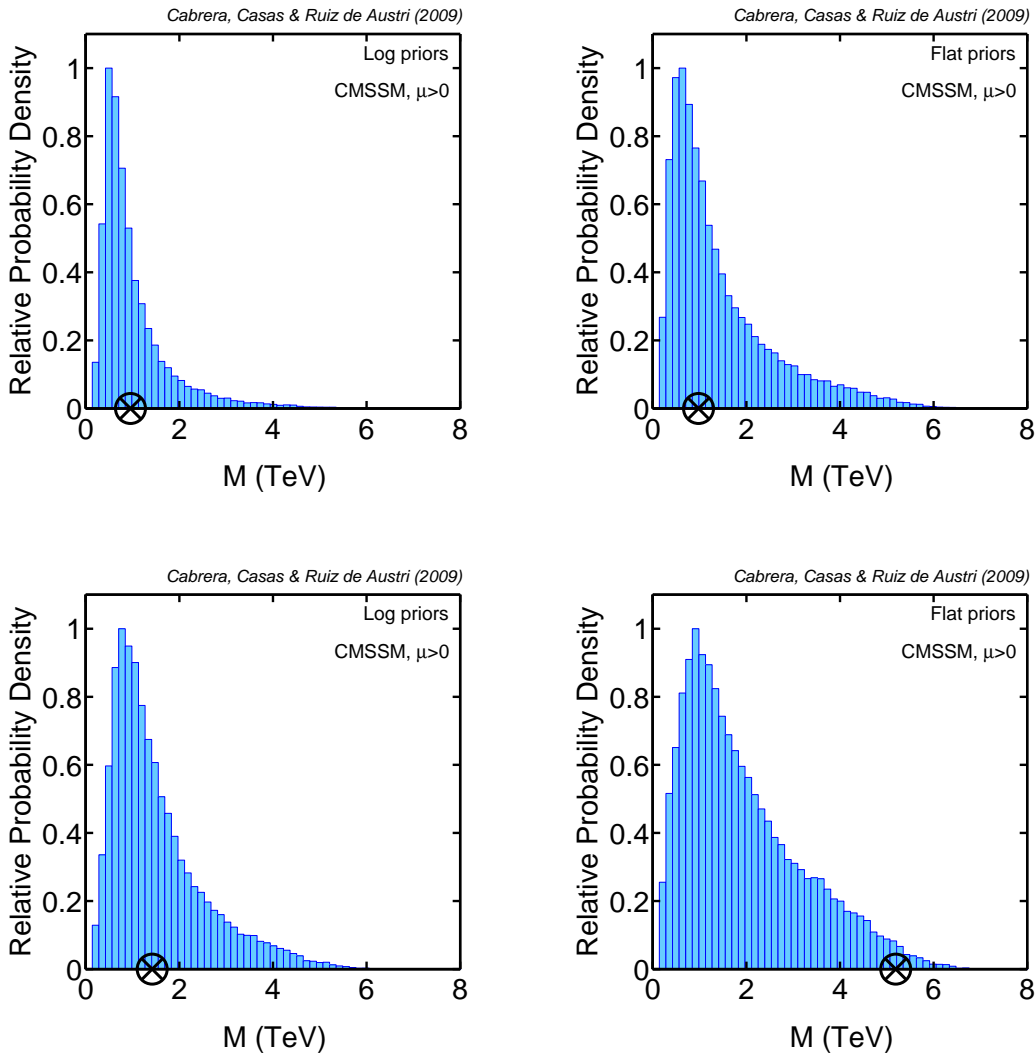


Figure 2: Upper panels show the 1D marginalized posterior probability distribution of the M parameter for logarithmic (left panel) and flat (right panel) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking (M_Z^{exp}), collider limits on Higgs and superpartner masses, and EW and B(D)-physics observables. Lower panels show the same but imposing a bound for the Higgs mass of $m_h \geq 120$ GeV. The cross gives the best-fit point.

to a logarithmic (flat) initial prior for the scale of SUSY breaking in the observable sector (M_S). The reason to show the pdf of M is to facilitate the comparison with the analogous probability distribution before the inclusion of the new pieces of experimental information (Fig. 1). Clearly, the bulk of the probability is now pushed into the high-energy region. This effect is basically due to the Higgs mass bound. As discussed above, concerning the other observables, everything works fine, as long as SUSY is not at too low scale. On the other hand, it is well known that in the MSSM the tree-level Higgs mass is bounded from above by M_Z , so radiative corrections (which grow logarithmically with the stop masses) are needed.

prior	Mean value	Best fit	68% (95%) range
log	116.7	117.9	[112.9 : 120.6] ([110 : 124.8])
flat	118.6	117.6	[114.4 : 123] ([111 : 126.8])

Table 3: Higgs mass mean value and best fit for logarithmic and flat priors, and the 68% and 95% Bayesian equal-tails credibility intervals. All numbers are given in GeV units.

It is possible to be more quantitative by considering the dominant 1-loop correction [51] to the theoretical upper bound on m_h in the MSSM:

$$m_h^2 \leq M_Z^2 \cos^2 2\beta + \frac{3m_t^4}{2\pi^2 v^2} \log \frac{M_{\tilde{t}}^2}{m_t^2} + \dots \quad (4.2)$$

where m_t is the (running) top mass and $M_{\tilde{t}}$ is an average of stop masses. Hence, for a given lower bound on the Higgs mass, m_h^{\min} , one needs

$$M_{\tilde{t}} \gtrsim e^{-2.1 \cos^2 2\beta} e^{(m_h^{\min}/62 \text{ GeV})^2} m_t. \quad (4.3)$$

Thus, an increase Δm_h^2 on the lower bound of the Higgs mass squared approximately translates into a multiplicative factor for $M_{\tilde{t}}$:

$$M_{\tilde{t}} \rightarrow M_{\tilde{t}} e^{\Delta m_h^2 / (62 \text{ GeV})^2}, \quad (4.4)$$

and a similar increase can be expected in the initial parameters m , M .

To illustrate these facts, we have re-done the pdfs assuming a different value of the Higgs mass bound, say $m_h \geq 120$ GeV. Of course this would correspond to the real situation if the Higgs mass turns out finally to lie in this range. According to the previous argument, we can expect now a longer push of the probability distribution into the high-energy region. And this is what happens, as it is shown in Fig.2 (lower panels). The effect is very important, given the modest increase in the Higgs mass bound. Larger shifts in m_h have an exponentially larger effect, as discussed above. So, if the MSSM is true and we wish to detect it at LHC, let us hope that m_h is close to the present experimental limit.

Fig.3 shows some representative probability distributions for individual (initial and derived) parameters, i.e. once all the rest are marginalized. The dimension-full parameters (m , A) follow a trend similar to that of the gaugino mass, M (which was already shown in Fig. 2). On the other hand, large values of $\tan \beta$ are penalized, mainly due to the Jacobian factor in the probability distribution, see eq.(2.15). Fig. 4 shows the probability distribution for the Higgs mass. The corresponding Bayesian credibility intervals, representing the 68% and 95% of the total probability, are given in Table 3. The central value for the Higgs mass is at 117–118 GeV. From that table one can see the robustness of the results under changes of the prior, which suggests that the posterior is dominated by the likelihood. This is confirmed by the little discrepancy among the mean value of the posterior pdf and the best fit.

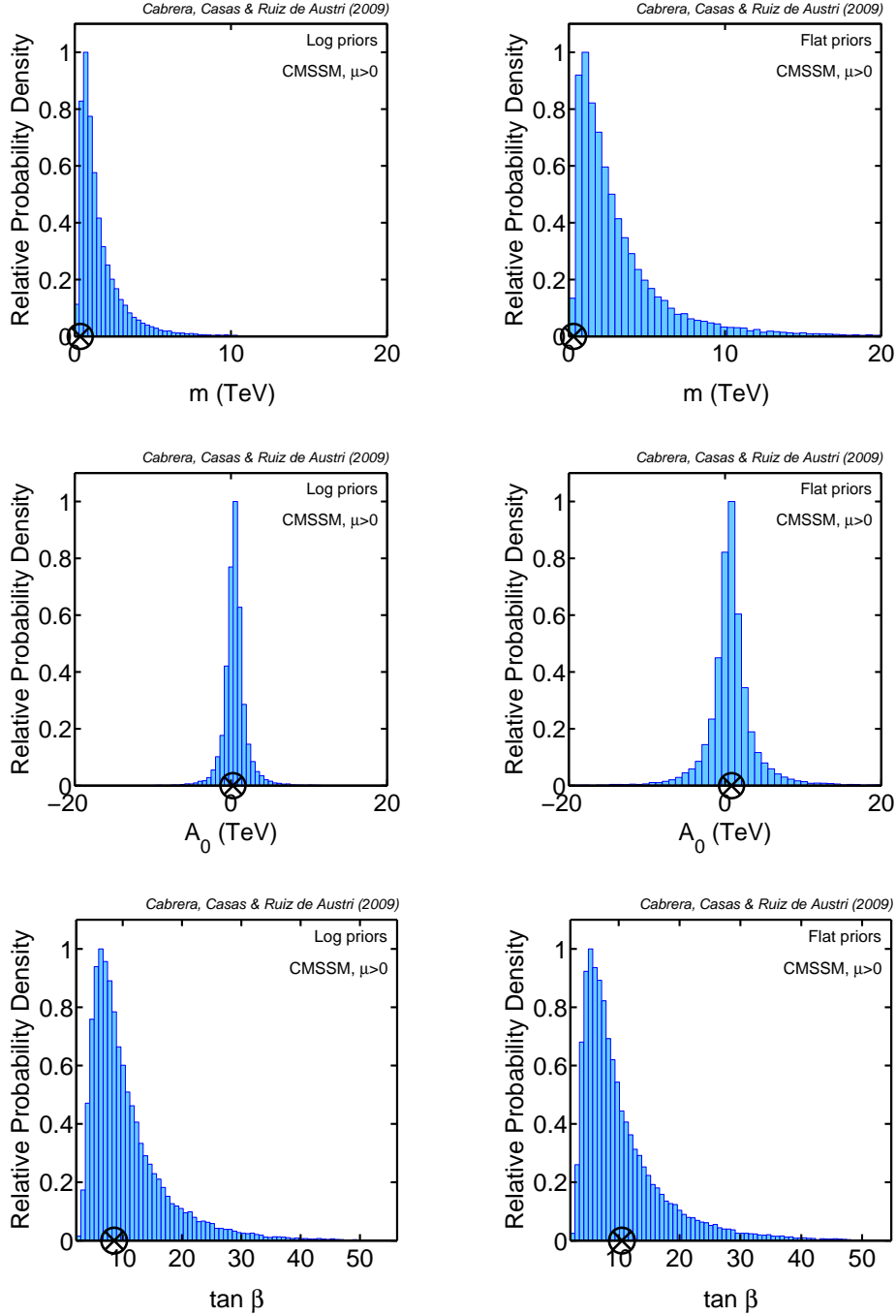


Figure 3: 1D marginalized posterior probability distribution of the CMSSM parameters for logarithmic (left panels) and flat (right panels) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking (M_Z^{exp}), collider limits on Higgs and superpartner masses, and EW and B(D)-physics observables. The cross gives the best-fit point.

Fig.5 shows the probability distribution in the $\{M, m\}$ and $\{\tan \beta, M\}$ planes (i.e. when all the parameters but two are marginalized). The results of Figs. 3–5 are shown all

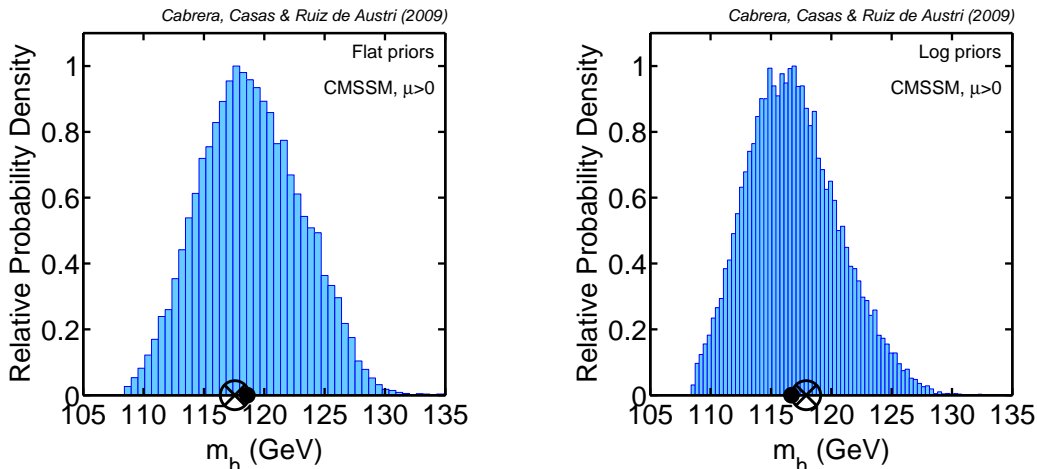


Figure 4: As Fig. 3, for the Higgs mass. The small filled circle represents the mean value of the posterior pdf.

for logarithmic (left panels) and flat (right panels) priors, exhibiting a remarkable stability, which has already been discussed. In the $\{M, m\}$ plots we have shown also the discovery reach of LHC for 1 fb^{-1} and 100 fb^{-1} (with a center-of-mass energy of 14 TeV). These lines have taken from ref.[29] and, strictly, correspond to $A = 0$, $\tan \beta = 45$, but they provide a good indication of the LHC discovery potential in the short and medium term (for similar analyses see [28]). Now, it is clear that a substantial (though still non-dominant) part of the probability falls out of the LHC reach, an effect that it is more important for flat prior. This means that if we are unlucky, supersymmetry could evade LHC detection in the short, or even the long, term. On top of this, let us recall that if the Higgs mass is not close to its present experimental value, the preferred regions of the parameter space are quickly pushed to high-energy (see discussion about Fig.2), thus jeopardizing the discovery of supersymmetry.

4.2 Constraints from $(g - 2)_\mu$

The magnetic anomaly of the muon, $a_\mu = \frac{1}{2}(g-2)_\mu$ has been a classical and powerful test for new physics. At present, the experimental uncertainties in the experimental and theoretical determinations are on the verge of strongly constraining, or even giving a positive signal, of new physics. However, the situation is still somewhat uncertain, due essentially to inconsistencies between alternative determinations of the SM hadronic contribution, more precisely the contribution coming from the hadronic vacuum polarization diagram, say $\delta_{\text{had}}^{\text{SM}} a_\mu$.

This contribution can be expressed in terms of the total hadronic cross section $e^+e^- \rightarrow$ hadrons. Using direct experimental data for this cross section, one obtains a final result for a_μ , which is at more than 3σ from the current experimental determination, namely $\delta a_\mu = a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = 29.5 \pm 8.8 \times 10^{-10}$. This has been often claimed as a signal of new

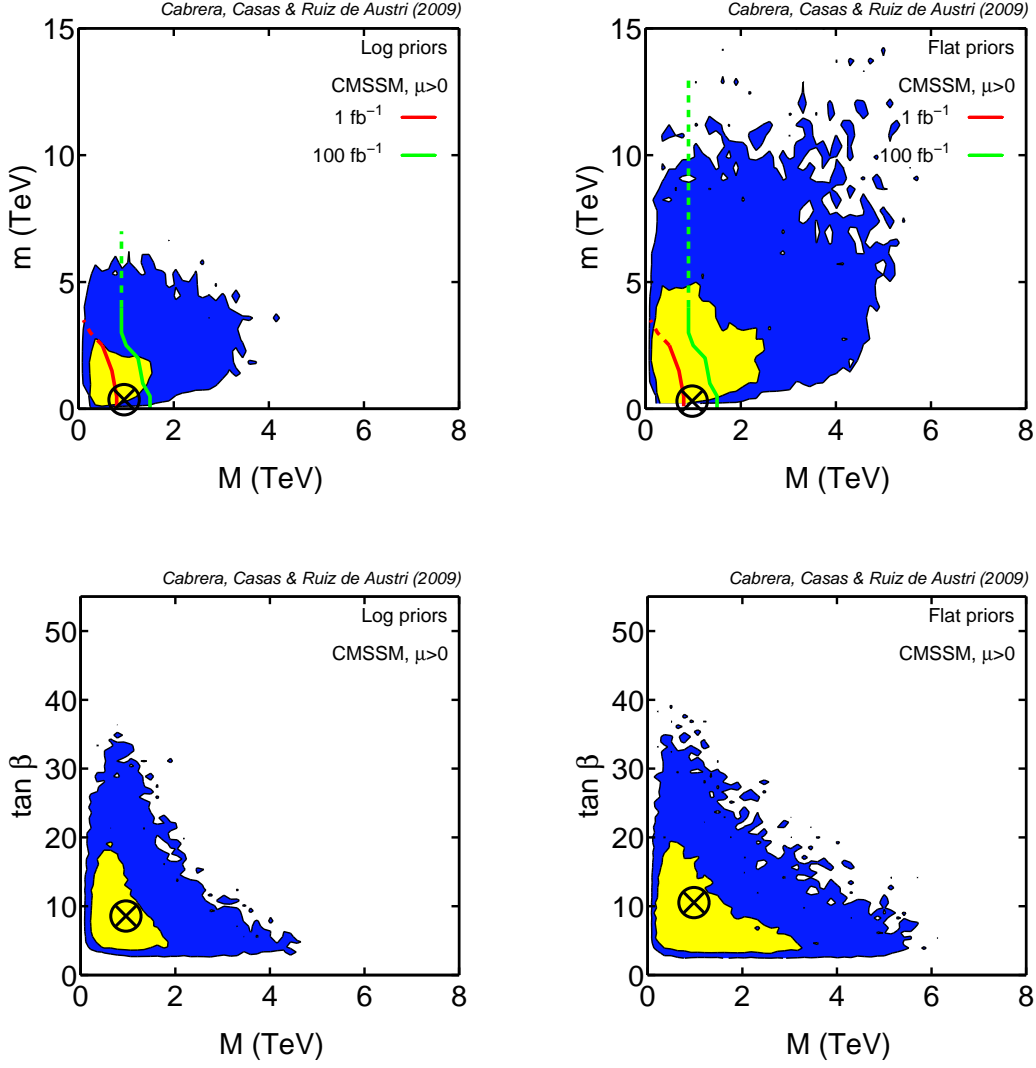


Figure 5: 2D marginalized posterior probability distribution for logarithmic (left panels) and flat (right panels) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking (M_Z^{exp}), collider limits on Higgs and superpartner masses, and EW and B(D)-physics observables. The inner and outer contours enclose respective 68% and 95% joint regions. The red(green) lines show discovery reach of LHC with 1(100) fb^{-1} . The cross gives the best-fit point.

physics. Obviously, if one accepts this point of view, the discrepancy should be cured by contributions of new physics, in our case MSSM contributions. The immediate implication is that supersymmetric masses should be brought to quite small values, in order to produce a large enough contribution, $\delta^{\text{MSSM}} a_\mu$, to reconcile theory and experiment. Hence, SUSY should live at low-energy (accessible to LHC), mainly because of a_μ . This is an independent argument from the the one based on the size of the EW scale, which has been discussed in sect.3.

The previous statement is quite strong. History has taught us that many experimental

observables, in apparent disagreement with the SM prediction, have eventually converged with it. This occurred due to both experimental and theoretical subtleties and difficulties, that sometimes had not been fully understood or taken into account. Although, obviously, a_μ is a most relevant test for the SM, and hopefully will be a first indication of physics beyond the SM, it is perhaps prudent not taking for granted that this is so indeed. As a matter of fact, the experimental $e^+e^- \rightarrow$ hadrons cross section shows some inconsistencies between different groups of experimental data (see [52] for a recent account). This is especially notorious if one considers the hadronic τ -decay data, which are theoretically related to the e^+e^- hadronic cross section. Using the τ -data, the 3.3σ disagreement becomes 1.8σ , i.e. one comes back to the SM realm. Although the more direct e^+e^- data are usually preferred to evaluate a_μ^{SM} , this discrepancy is warning us to be cautious about this procedure.

To illustrate this situation, we have performed two alternative analyses. In the first one we use the evaluation of $\delta_{\text{had}}^{\text{SM}} a_\mu$ based on e^+e^- data. In the second, we use the one based on τ -data. We compute $\delta_{\text{had}}^{\text{SM}} a_\mu$ at full one-loop level adding the logarithmic piece of the quantum electro-dynamics two-loop calculation plus two-loop contributions from both stop-Higgs and chargino-stop/sbottom [53]. The effective two-loop effect due to a shift in the muon Yukawa coupling proportional to $\tan^2 \beta$ has been added as well [54].

Using $\delta_{\text{had}}^{\text{SM}} a_\mu$ from e^+e^- data

In this case, the inclusion of the a_μ constraint has a dramatic effect, as mentioned above. The preferred values of the soft terms are pushed into the low-energy region. Actually, the push is so strong that the predictions for other observables, in particular $b \rightarrow s \gamma$, start to be too large. This tension has been pointed out in ref.[55], and we would like to illustrate it here presenting some representative plots. Fig.6 shows the (non-normalized) pdf for the m -parameter in three different cases (taking always a logarithmic prior): a) using EW + Bounds + B-physics, as in subsect 4.1 (blue solid line); b) using EW + Bounds + a_μ (red dashed line); c) using EW + Bounds + B-physics + a_μ (green dashed-dotted line).

Clearly, the effect of just a_μ is to bring the preferred region for the soft terms from ~ 1 TeV to ~ 300 GeV. This effect is remarkably stable against variations of the type of prior, indicating that the data are now powerful enough to essentially select a region of the parameter space. Let us also mention that large values of $\tan \beta$ become now much more likely, being normally associated to the region of larger soft masses (recall that $\delta^{\text{MSSM}} a_\mu$ grows with decreasing masses and increasing $\tan \beta$). When both $b \rightarrow s, \gamma$ and a_μ are taken into account, there is almost no region of the parameter space able to reproduce both experimental results within 2σ . Therefore the likelihood factor gets suppressed, and the “preferred” region of the parameter space (illustrated here by the green line) is somehow an average of the two previous cases. This tension between $b \rightarrow s, \gamma$ and a_μ can also be noticed by looking at Fig.7, where the left and right panels show the pdfs of $BR(B \rightarrow X_s \gamma)$ and $\delta^{\text{MSSM}} a_\mu$ respectively, with the same code for the lines as in Fig. 6. Comparing the position

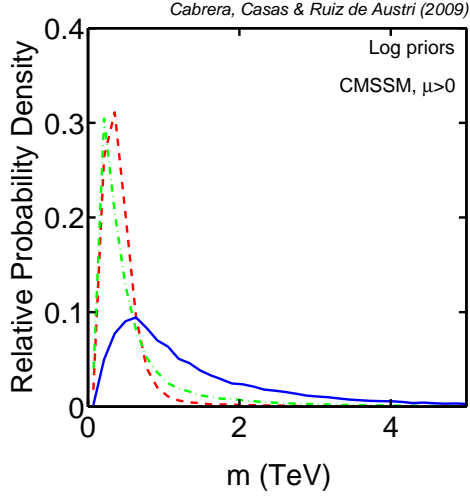


Figure 6: Non-normalized 1D marginalized posterior probability distribution of the m parameter for logarithmic prior and $\mu > 0$ including SM nuisance parameters constraints, EW breaking (M_Z^{exp}), EW observables, collider limits on Higgs and superpartner masses: + B(D)-physics observables (blue line); + a_μ (red line); + B(D)-physics and a_μ (green line).

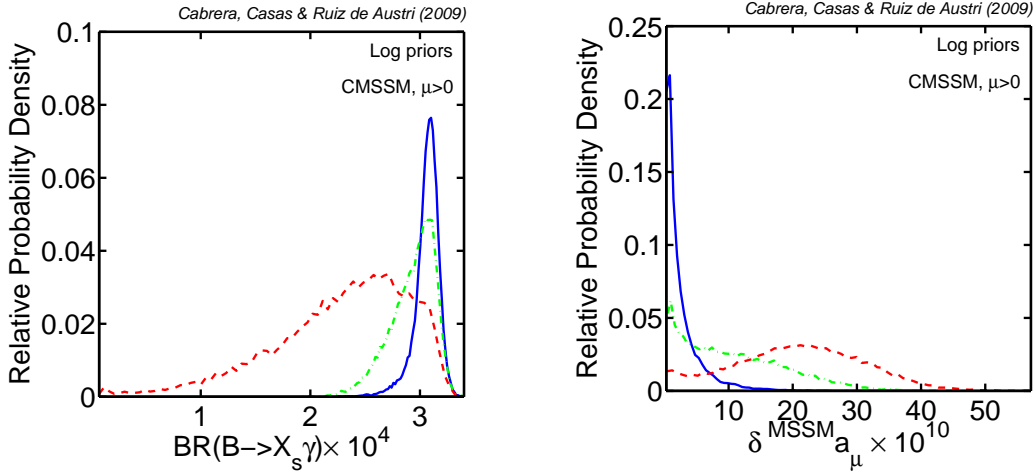


Figure 7: Non-normalized 1D marginalized posterior probability distribution for $BR(B \rightarrow X_s \gamma)$ (left panel) and for $\delta^{\text{MSSM}} a_\mu$ (right panel). The code for the lines is as in Fig. 6.

of the bulk of the probability distribution with the central values and uncertainties quoted in Table 2, it is clear that the most favourable cases are not really satisfactory reproducing the two measurements.

Let us also remark that, if the Higgs mass turns out to be $\mathcal{O}(10)$ GeV above the present experimental limit, the tension between the Higgs mass and a_μ^{exp} would be dramatic and could not be reconciled: m_h (a_μ^{exp}) would require too large (small) soft masses, see the discussion in subsect. 4.1.

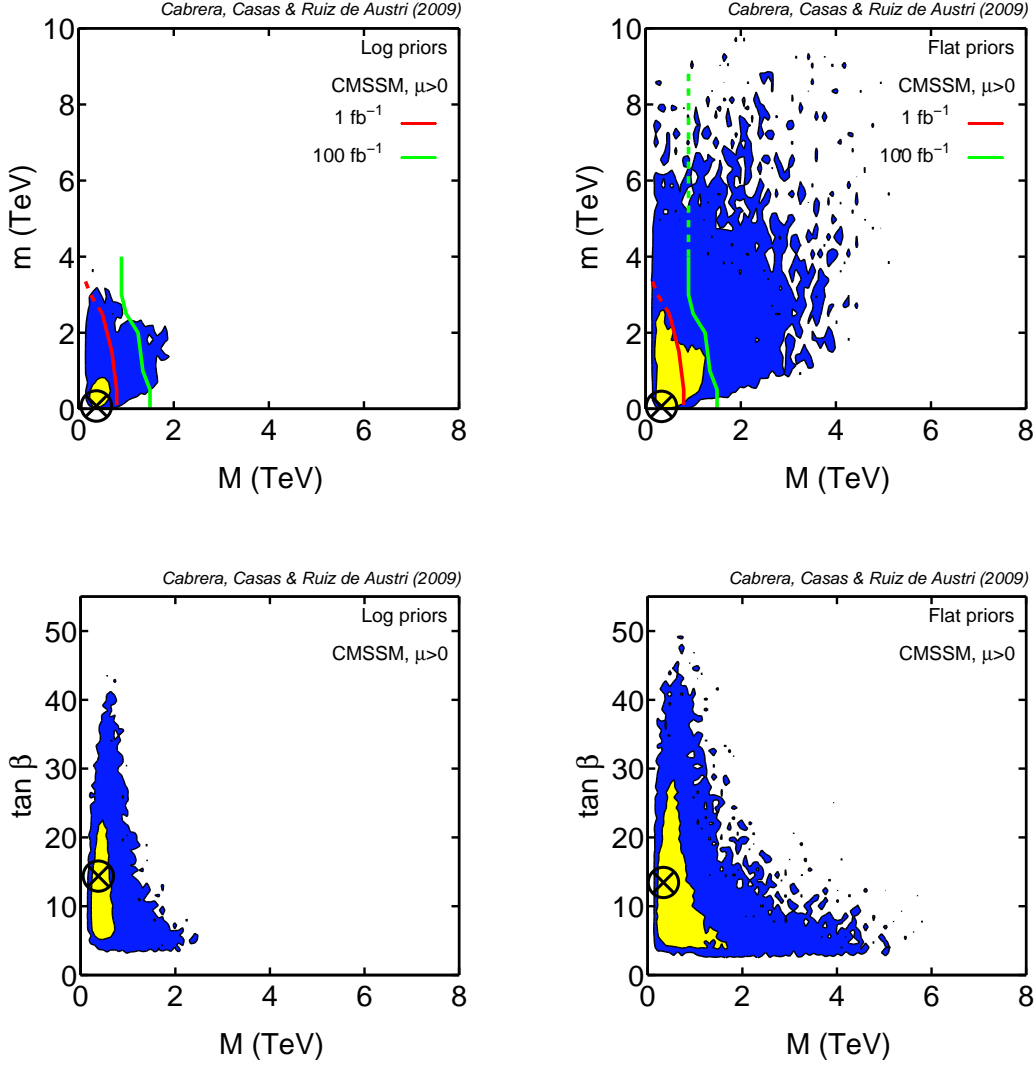


Figure 8: As in Fig. 5 but with the additional constraint from a_μ , based on e^+e^- data.

Fig.8 shows the probability distribution in the $\{M, m\}$ and $\{\tan \beta, M\}$ planes, as in Fig.5, once the a_μ constraint (based on e^+e^- data) is included. Comparison with Fig. 5 clearly shows the big push of the soft terms into the low-energy region. Actually, most of the probability falls now within the LHC reach (even in the short term), which is great news for the potential discovery of SUSY (if the a_μ discrepancy is really there).

Using $\delta_{\text{had}}^{\text{SM}} a_\mu$ from τ data

In this case, there is no big discrepancy between a_μ^{SM} and a_μ^{exp} , so $\delta^{\text{MSSM}} a_\mu$ does not need to be large. Consequently, the probability distributions are essentially unchanged by the inclusion of the a_μ constraint, and are very similar to those shown in subsect. 4.1 (Figs. 2–5).

Consequently, if a_μ is not a signal of new physics, the size of EW breaking continues to be the only piece of data that brings SUSY to scales accessible to LHC (apart from Dark Matter considerations, which we examine next).

4.3 Constraints from Dark Matter

There are different astrophysical and cosmological observations that offer impressive evidence of the existence of Dark Matter (DM) in the universe (see Table 2 for a recent determination of Ω_{DM}). On the other hand, the consistency with the observed large structure of the universe favours cold dark matter (CDM), i.e. non-relativistic matter at the beginning of galaxy formation. This leads to the hypothesis of a weakly interacting massive particle (WIMP) as the component of CDM.

Supersymmetry offers a good candidate for such a WIMP, namely the LSP, which is stable in the standard (R-parity conserving) SUSY formulations (for a review see [56]). Although, depending on the models, there are several possibilities for the SUSY WIMP, the most popular and natural candidate is the lightest neutralino, χ^0 , which is the LSP in most of the CMSSM parameter space. However the calculations show that typically too many neutralinos are produced after inflation. Therefore some efficient annihilation mechanism is required in order to bring Ω_{DM} down to the allowed range. In the context of CMSSM there are four such mechanisms known, which take place in four different regions of the parameter space:

Bulk region: Neutralinos can be annihilated (into leptons) via sleptons if the masses of the latter are not high. This requires rather small m and M soft parameters, in potential conflict with the Higgs mass bound.

Focus Point region: For moderate or large values of $\tan\beta$ the electroweak scale is quite insensitive to the variation of m . For large enough values of m , the μ parameter decreases, which drives the LSP to get a significant Higgsino-component, making its annihilations (into vector bosons) more efficient.

Co-annihilation region: If the mass of the second lightest supersymmetric particle (NLSP) is close to that of the LSP, the annihilation of the latter is enhanced through co-annihilation processes. In the CMSSM this mechanism takes place typically with an stau NLSP. In the parameter space this corresponds to a rather narrow region with $M > m$.

Higgs funnel region: When the mass of the pseudoscalar A^0 -boson becomes close to twice the neutralino LSP and $\tan\beta$ is large, the annihilation occurs quite efficiently through the A^0 resonance.

In order to evaluate the viability of supersymmetric CDM in each point of the CMSSM parameter space, we use the `MicrOMEGAs` code [57] integrated into `SuperBayes`. The corresponding likelihood, assuming that all the CDM is made up of neutralinos, is then incorporated to the pdf in the Bayesian scan. Fig.9 shows the resulting probability distribution in

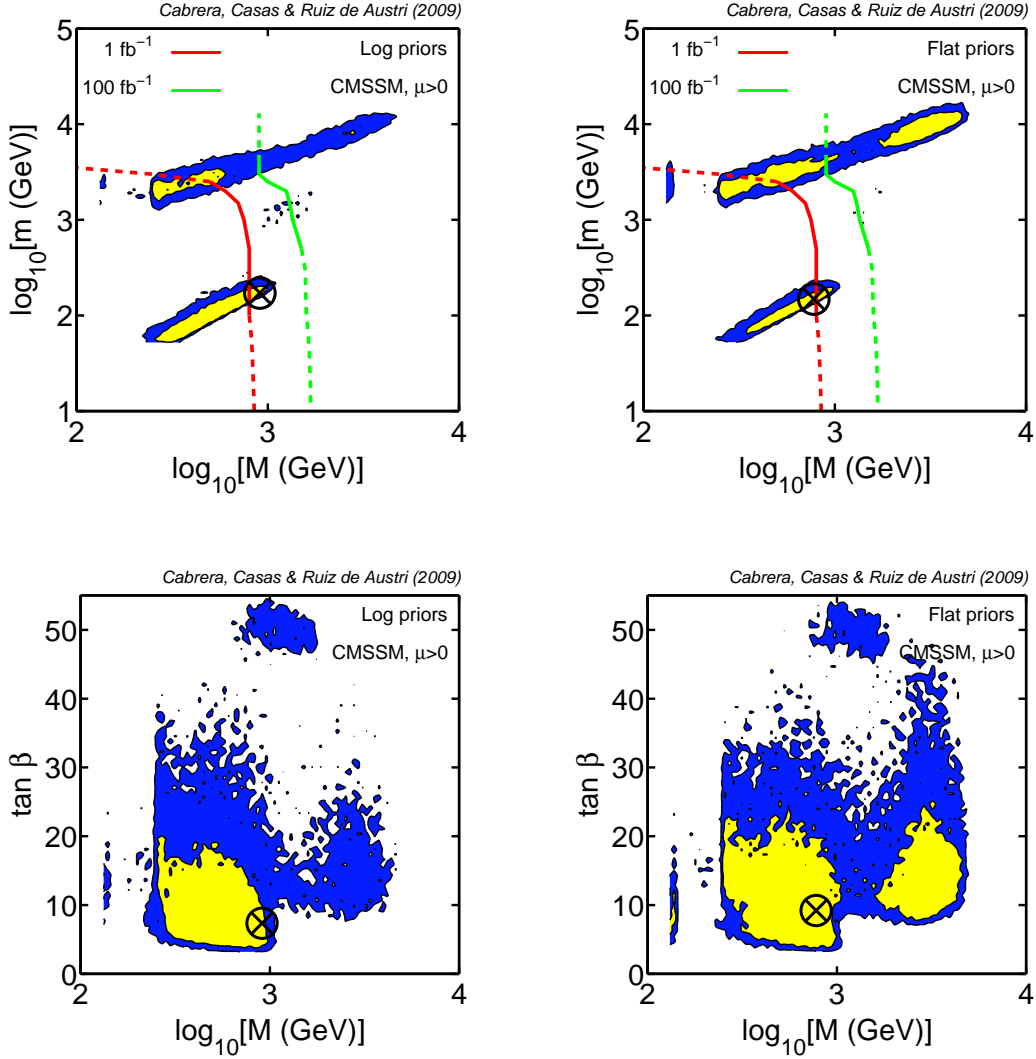


Figure 9: As in Fig. 5 but with an additional constraint from the WMAP CDM abundance.

the $\{M, m\}$ and $\{\tan \beta, M\}$ planes (i.e. when all the parameters but two are marginalized) for logarithmic and flat priors. In these figures we have *not* included the information about a_μ . The $\{M, m\}$ -plane plots show a kind of blurring with respect to usual plots in the literature, due to the integration in the variables $A, \tan \beta$. Still, the above-mentioned four viable regions are visible in Fig.9. The $\{\tan \beta, M\}$ plots show that some of the regions require large $\tan \beta$ (in particular the Higgs-funnel region), which thus becomes more probable than before-including CDM constraints. Although the favoured regions are qualitatively similar for logarithmic and flat priors, quantitatively the area of highest probability is extended into larger (even inaccessible to LHC) soft masses in the case of flat prior. This is because the DM constraints, though quite severe, do not select a unique region of the parameter space but several ones, located in different zones of the CMSSM parameter space, as discussed above. Consequently, the prior assumed for the parameter space plays

a relevant role when comparing the relative probability of these regions.

Regarding the impact on the LHC potential of discovery, roughly speaking, including DM constraints the low-energy gets favoured and therefore the detection of SUSY at the LHC, as can be seen by comparing Figs. 5 and 9. However, there survive large (though less probable, especially for log prior) high-energy areas out of the LHC reach. Consequently, again, if we are unlucky, even if DM is supersymmetric, it could escape LHC detection (especially if the Higgs mass is not close to its experimental limit).

In any case, again, one should be cautious at interpreting these results as a robust constraint on the CMSSM. Certainly, they are so with an “standard” cosmology. However, it could happen that other regions of the MSSM parameter space are cosmologically viable if, e.g. the overproduction of CDM is diluted by electroweak baryogenesis. Admittedly, the latter is not a most natural or popular scenario of inflation, but mechanisms for it have been explored [58]. Alternatively, the LSP could be unstable assuming tiny violations of R-parity, see e.g. [59]. In these cases the observed dark matter should be provided by other candidate, e.g. an axion. But this is not a drawback for the model. Of course, CDM constraints are extremely interesting and they have to be taken into account. But it seems sensible not to put them at the same level as e.g. electroweak observables.

Finally, Fig. 10 shows the $\{M, m\}$ -plane plots when the a_μ constraint (based on e^+e^- data) is incorporated to the analysis as well. Clearly, the regions with “too large” soft masses (to reproduce the a_μ^{exp}) are now suppressed, leaving a quite definite region at low-energy. More precisely, the bulk and co-annihilation regions are now clearly selected amongst the various possibilities to obtain Ω_{DM} . We stress, however, that in this case one should be cautious about both the Ω_{DM} and the a_μ constraints. Note in particular that, if the a_μ constraint is based on τ data, it does not produce relevant restrictions and, consequently, the corresponding plots are quite similar to those of Fig. 9.

5. Negative sign of μ

So far all the results and plots presented correspond to $\mu > 0$. The analysis for $\mu < 0$ is completely similar. The most worth-mentioning difference is that with $\mu < 0$ the MSSM contributions to a_μ have negative sign and thus become useless to reconcile theory and experiment (a discrepancy that is only present *if* $\delta_{\text{had}}^{\text{SM}} a_\mu$ is evaluated using $e^+e^- \rightarrow \text{had}$ data). On the other hand, the contributions to $b \rightarrow s, \gamma$ have now positive sign, which is the “right” sign to push the theoretical result closer to the experimental value (see Table 1). This effect, however, has less impact than a_μ in the distribution of probability.

5.1 Results

The results for $\mu < 0$ are summarized in Figs. 11, 12, 13 and 14, which are as previous Figs. 5,8,9 and 10, but with opposite sign of μ .

Fig. 11 shows the posterior distribution function when only the most robust set of data (EW and B(D)-physics observables, and limits on particle masses) are taken into account.

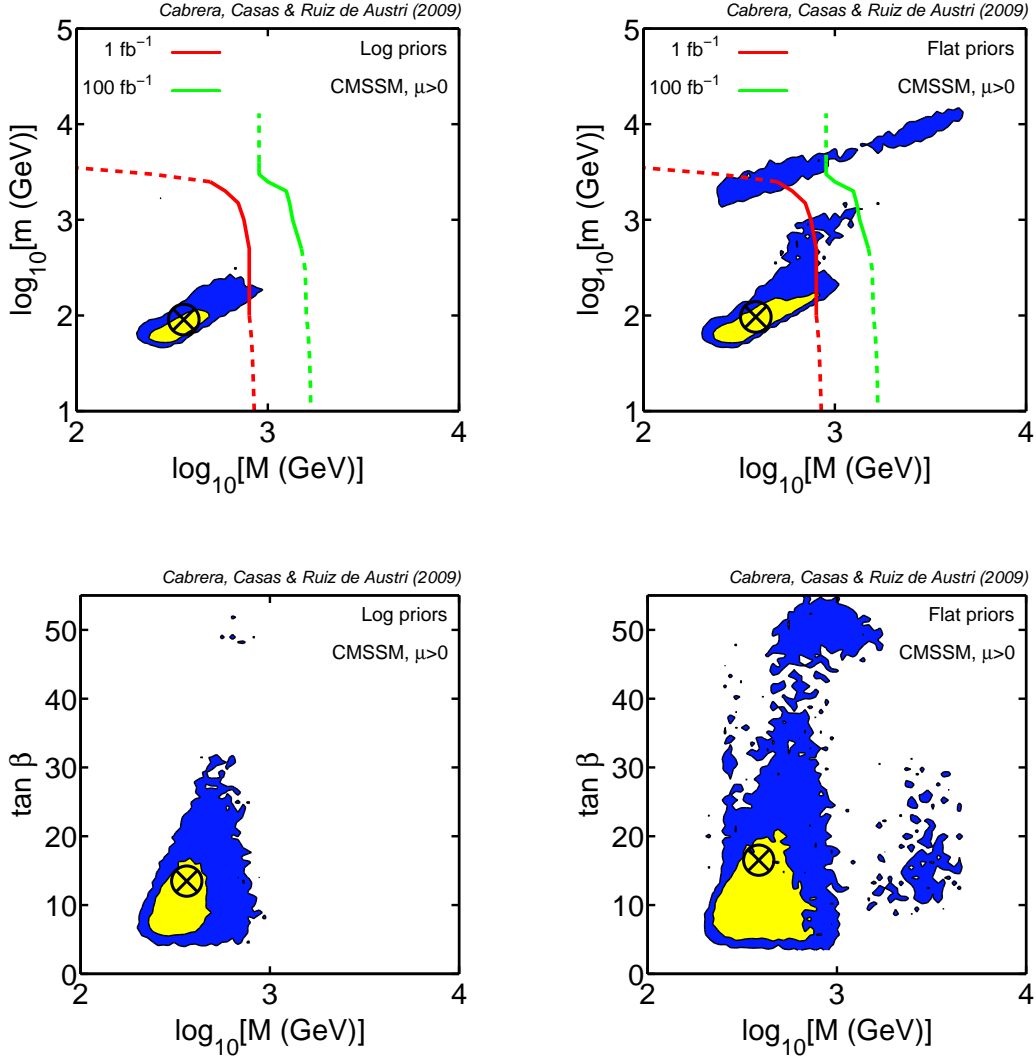


Figure 10: As in Fig. 5 but with an additional constraint from the WMAP CDM abundance.

Because of the above-mentioned $b \rightarrow s, \gamma$ observable, the distribution is now slightly shifted to smaller soft masses (now “it pays” to have a moderately sizeable SUSY contribution to this process), as it is clear from comparison with Fig. 5. The effect is welcome, as it pushes SUSY towards regions of the parameter space more accessible to LHC. However the impact is far from dramatic.

Fig. 12 shows the posterior when a_μ (evaluated using e^+e^- data) is included in the analysis. Now the difference with the analogue for positive μ (Fig. 8) is really dramatic. Recall that now the SUSY contributions to a_μ have the wrong sign, so it does not pay to have smaller soft masses. Consequently, Fig. 12 is similar to Fig. 11 (i.e. before including a_μ constraints) and, actually, the soft masses are pushed to slightly higher values.

Fig. 13 shows the posterior when one considers the previous robust set of data (not including a_μ) plus the constraints from Dark Matter. Since Dark Matter has a great

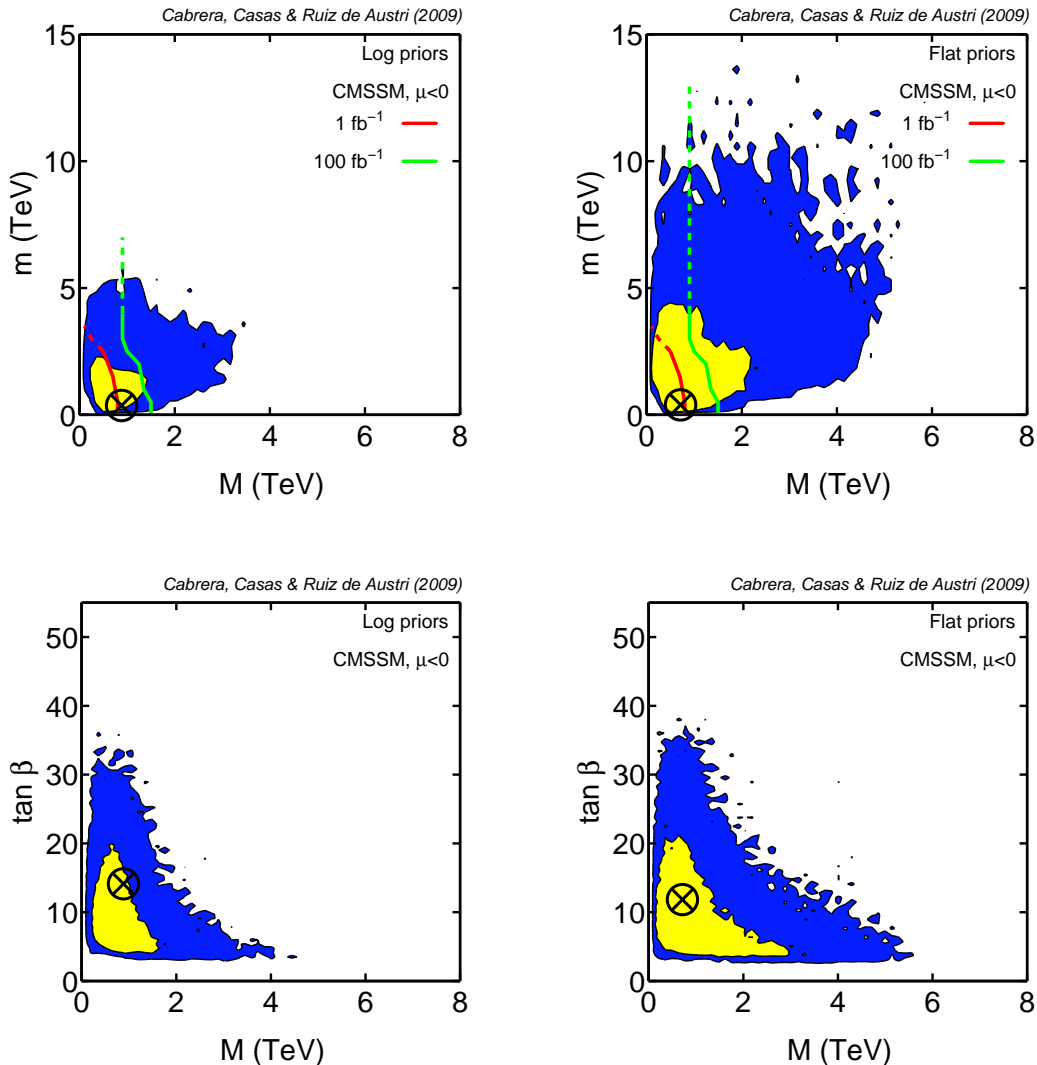


Figure 11: As in Fig. 5 but with $\mu < 0$.

potential to select preferred regions in the parameter space, the results are quite similar to those for $\mu > 0$, Fig. 9; and the same comments hold here.

Finally, Fig. 14 shows the posterior when all the experimental information, including a_μ (evaluated using e^+e^- data) is taken into account. Similarly to our above discussion of Fig. 12, the results do not change much after the inclusion of the a_μ constraint. In consequence, Fig. 14 is quite similar to Fig. 13, with a certain penalization of too small soft masses. Again, this is in strong contrast with the $\mu > 0$ case, where the low-energy (bulk and co-annihilation) regions were preferred (see Fig. 10).

5.2 Positive versus negative μ

In order to compare the relative probability of the $\mu > 0$ and $\mu < 0$ branches, one has

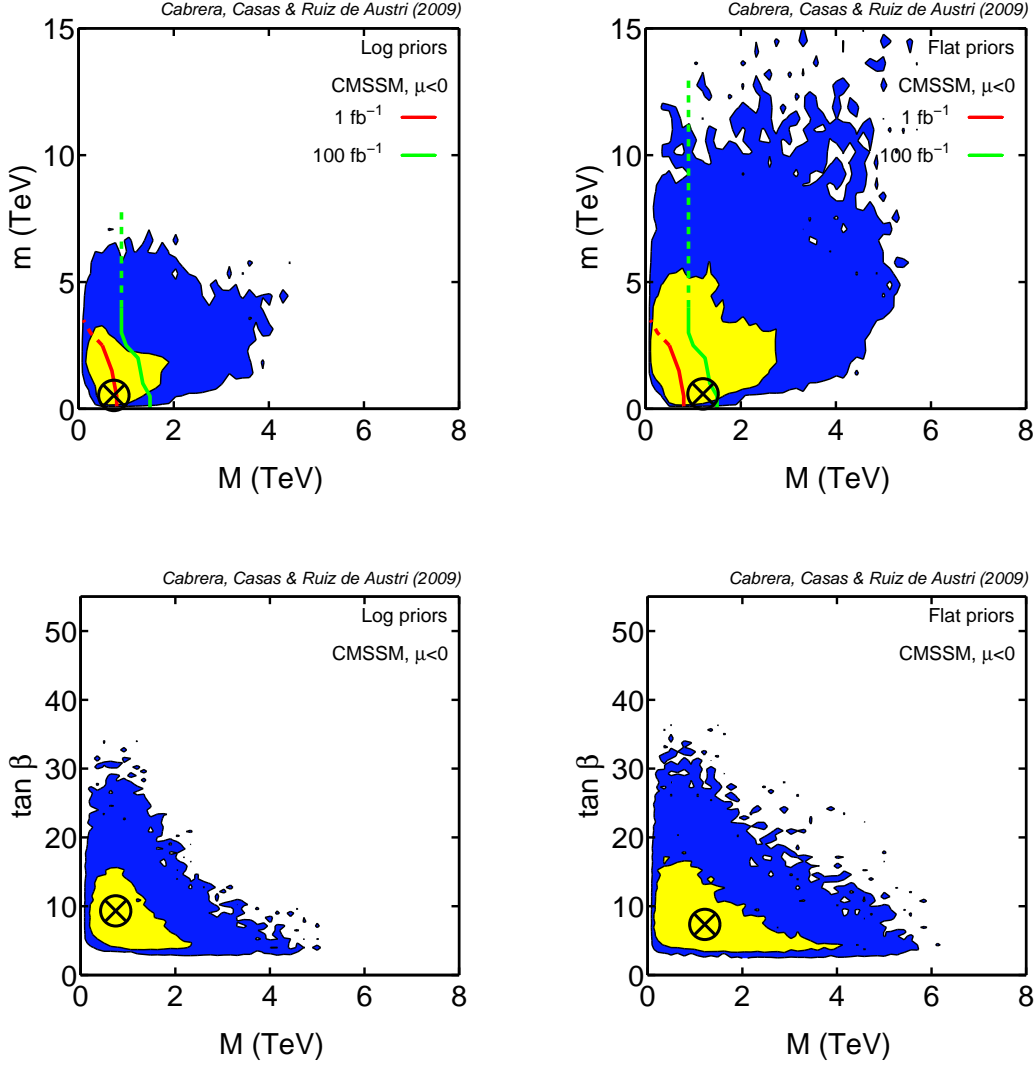


Figure 12: As in Fig. 8 but with $\mu < 0$.

to evaluate the Bayesian evidences of both cases. We recall that the Bayesian evidence is the piece in the denominator of eq.(1.1), i.e. $p(\text{data})$, sometimes called \mathcal{Z} . Note that, integrating both sides of eq.(1.1), and using the fact that $p(\theta_i|\text{data})$ must be correctly normalized, one simply obtains

$$\mathcal{Z} \equiv p(\text{data}) = \int d\theta_1 \cdots d\theta_N p(\text{data}|\theta_i) p(\theta_i), \quad (5.1)$$

i.e. the evidence is the integral of the likelihood times the prior, and therefore it is a measure of the global probability of the model. Note that in doing parameter inference within a given model, \mathcal{Z} plays the role of a normalization factor and can be (and it is) usually ignored. However, it is the key quantity to compare the relative probability of two different models. Let $\mathcal{M}_1, \mathcal{M}_2$ be two models with prior probabilities $p(\mathcal{M}_1), p(\mathcal{M}_2)$.

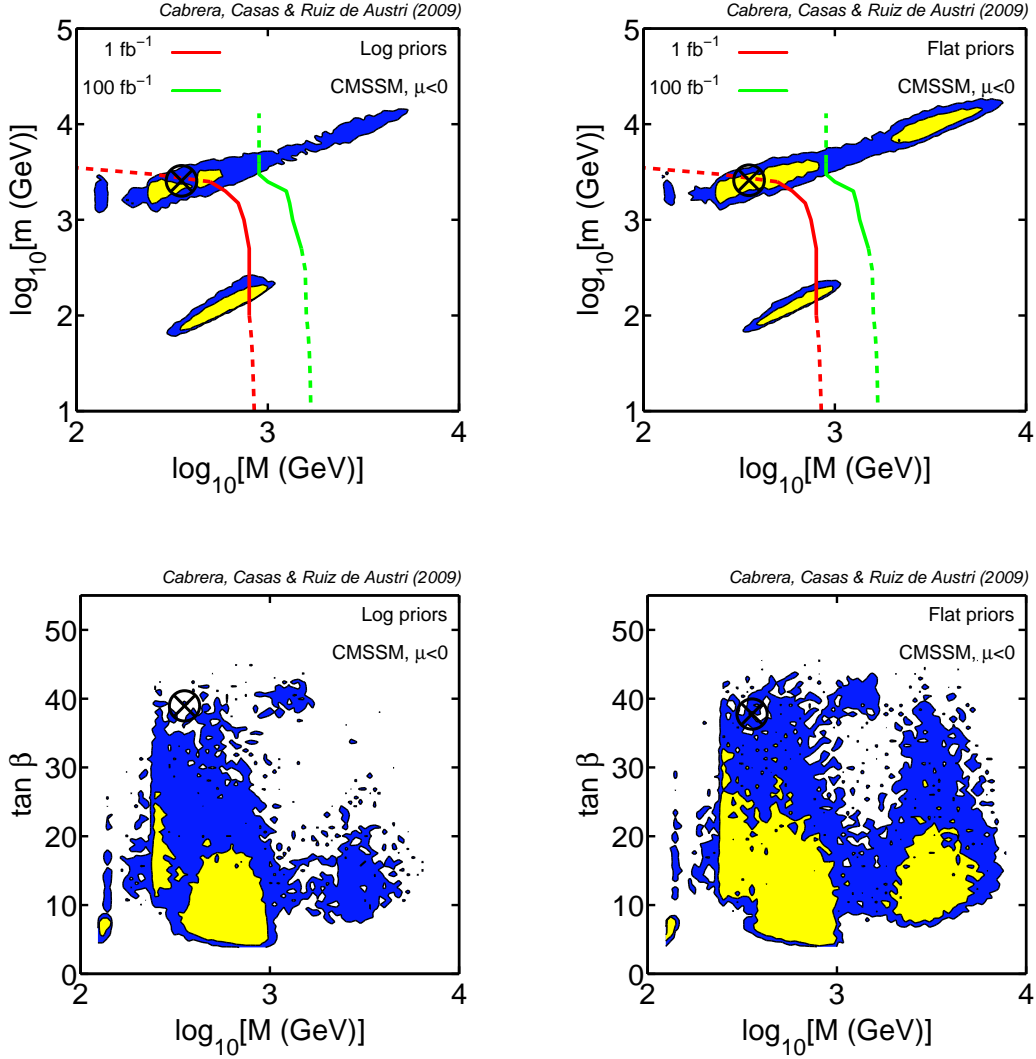


Figure 13: As in Fig. 9 but with $\mu < 0$.

Then the relative posterior probability of the two models, given a set of data, is simply

$$\frac{p(\mathcal{M}_1|data)}{p(\mathcal{M}_2|data)} = \frac{\mathcal{Z}_1 p(\mathcal{M}_1)}{\mathcal{Z}_2 p(\mathcal{M}_2)} = B_{12} \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}, \quad (5.2)$$

where $B_{12} \equiv \mathcal{Z}_1/\mathcal{Z}_2$ is called the Bayes factor and $p(\mathcal{M}_1)/p(\mathcal{M}_2)$ is the prior factor, often set to unity.

The natural logarithm of the Bayes factor provides a useful indication of the different performance of two models. In Table 4, we summarize the translation of the Bayes factor to relative probabilities and a conventional interpretation of them [60], which we follow in this paper.

The evaluation of the Bayesian evidence is in general a numerically challenging task, as it involves a multidimensional integral over the whole parameter space. In addition

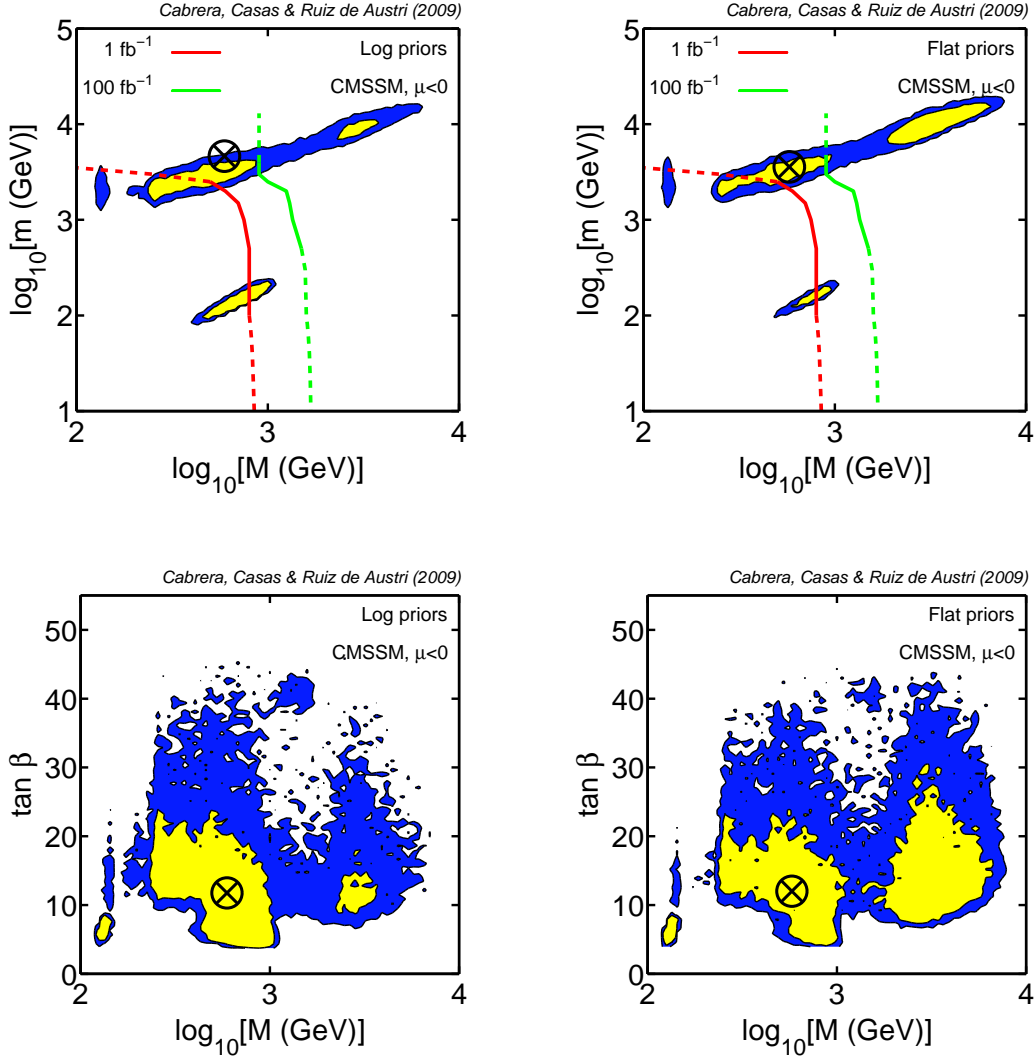


Figure 14: As in Fig. 10 but with $\mu < 0$.

$ \ln B_{10} $	Odds	Probability	Strength of evidence
< 1.0	$\lesssim 3 : 1$	< 0.750	Inconclusive
1.0	$\sim 3 : 1$	0.750	Weak Evidence
2.5	$\sim 12 : 1$	0.923	Moderate Evidence
5.0	$\sim 150 : 1$	0.993	Strong Evidence

Table 4: The scale we use for the interpretation of model probabilities.

the likelihood is often multi-modal, or it has strong degeneracies that confine the posterior to thin sheets in parameter space. Standard techniques as thermodynamic integration [61] have been proposed for a reliably estimation of the evidence, but they are extremely computationally-expensive. Perhaps the most elegant algorithm proposed up to now is the

Observables	$\ln B_{+-}$ (flat)	$\ln B_{+-}$ (log)
EW + Bounds + B-physics	-0.32 ± 0.048	-0.48 ± 0.049
EW + Bounds + B-physics + $(g - 2)_\mu$	0.81 ± 0.043	1.73 ± 0.052
EW + Bounds + B-physics + Ω_{DM}	-0.31 ± 0.068	-0.66 ± 0.066
EW + Bounds + B-physics + $(g - 2)_\mu + \Omega_{DM}$	1.9 ± 0.065	3.71 ± 0.068

Table 5: The natural log of the Bayes factor ($\ln B_{+-}$) for $\mu > 0$ and $\mu < 0$. A positive (negative) value indicates a preference for $\mu > 0$ ($\mu < 0$).

nested sampling as referred to in Sec. 3. This technique has greatly reduced the computational cost of model selection.

Let us now come back to our task, namely comparing the evidences for the positive and negative μ branches, considered here as different models with equal prior probabilities. We have applied the MultiNest algorithm to obtain the Bayes factor of these two models, $B_{+-} = \mathcal{Z}_+/\mathcal{Z}_-$. Of course the results depend on the experimental information considered, which enters the likelihood piece in eq.(5.1). The results are given in Table 5.

The first column of Table 5 indicates the set of experimental data taken into account, the notation is self-explanatory and corresponds to the different cases previously defined. The discussion of subsect. 5.1 allows to understand the numbers of the table. When only the most robust pieces of experimental information are used (first row), the performance of both models is similar. The $\mu < 0$ branch is slightly favoured, due to its capability to reproduce the central value of $b \rightarrow s, \gamma$, but the effect is not really significant, as is shown by a value of $|\ln B_{+-}|$ well below 0.75, see Table 4. This holds when Ω_{DM} constraints are incorporated into the analysis (third row of Table 5). On the other hand, $(g - 2)_\mu$ constraints (when evaluated using $e^+e^- \rightarrow \text{had}$ data) clearly favour the $\mu > 0$ branch, as discussed above, which is reflected in the numbers of the second and fourth rows of Table 5. Using the conventions of Table 4, we see that the global evidence in favour of positive μ is weak-to-moderate (not strong but already significant). Note that this effect is stronger for log prior, since in that case the high-energy region (the preferred one for $\mu < 0$) gets an additional penalization. Likewise, when Ω_{DM} constraints are included (at the same time as $(g - 2)_\mu$), the preference for positive μ gets even stronger. This is because, Ω_{DM} constraints favours (in terms of statistical weight) the low-energy region of the parameter space, and this is the region strongly preferred (penalized) by a_μ constraints for positive (negative) μ .

6. Summary and conclusions

The idea of an MSSM forecast for the LHC is to use all the present (theoretical and experimental) information available to determine the relative probability of the different regions of the MSSM parameter space. This includes theoretical constraints (and perhaps prejudices) and experimental constraints. An appropriate framework to perform such forecast is

the Bayesian approach, which allows a sensible statistical analysis, clearly identifying the objective and subjective pieces of information. The latter are incorporated in the *prior*, which represents the “theoretical” probability density that is assigned a priori to each point in the parameter space. Ignoring the prior factor is not necessarily the most reasonable or “free of prejudices” attitude. Such procedure is equivalent to a completely flat prior in the parameters. But one needs some theoretical basis to establish, at least, the parameters whose prior can be reasonably taken as flat. Besides, a choice for the allowed ranges of the various parameters is necessary in order to make statistical statements. The Bayesian approach allows to keep track of the influence of the prior (whether explicit or implicit) upon the results.

In this paper we have performed a Bayesian analysis of the MSSM with universal soft terms at high energy (a scenario sometimes denoted CMSSM or MSUGRA). We have improved previous studies by means of a careful handling of the various pieces of information:

- First, we do not incorporate ad hoc measures of the fine-tuning to penalize regions of the parameter space with large soft terms. Such penalization arises from the Bayesian analysis itself when the experimental value of M_Z is considered on the same foot as the rest of experimental information (and not as a constraint of the model). Nicely, this permits to scan the whole parameter space, allowing arbitrarily large soft terms. Still, the low-energy region is statistically favoured (even before including dark matter or a_μ constraints). Incidentally, this statistical argument supports low-energy supersymmetry breaking (in the observable sector), even in a landscape scenario. On the other hand, in a frequentist analysis (thus ignoring the prior factor), the high-energy region is essentially non-disfavoured at all (before including dark matter or a_μ constraints), since it works as well as the ordinary SM. Using fine-tuning arguments to penalize it, would take us back to the choice of an implicit (and non-trivial) prior.
- We have done a rigorous treatment of the nuisance variables, in particular Yukawa couplings, showing that the usual practice of taking the Yukawas as required to reproduce the fermion masses, approximately corresponds to taking logarithmically flat priors in the Yukawa couplings. We argue that this is a most reasonable choice.
- Although we start with the usual MSSM initial parameters, $\{m, M, A, B, \mu\}$ (plus Yukawa couplings and other nuisance variables) we use an efficient (and actually quite common) set of variables to scan the MSSM parameter space. Besides trading μ by M_Z and the Yukawa couplings by the fermion masses, it is extremely convenient to trade B by $\tan\beta$. These changes introduce a global Jacobian factor in the density probability when working in the new (and more suitable) parameters for the scan. Once the information about M_Z^{exp} is incorporated (by marginalizing M_Z) the *effective prior* in the new variables inherits the Jacobian factor, as is explicit in eq.(2.8). A quite accurate analytical expression for it is given in eq.(2.15), which is valid for any MSSM (not just the CMSSM). This effective prior contains inside the above-mentioned penalization of fine-tuned regions, but we stress that the latter has not

been introduced by hand. Actually, these expressions for the effective prior contain no ad hoc constraints or prejudices, since the prior in the initial variables is still undefined.

- We have developed a sensible prior in the initial variables. Our basic assumption has been that the soft-breaking terms share a common origin and, hence, it is logical to assume that their sizes are also similar. So, it is reasonable to assume that a particular soft term can get any value (with essentially flat probability) of the order of the typical size of the soft terms in the observable sector, $M_S \sim F/\Lambda$, or below it. Then, concerning the prior in M_S itself, we have taken the two basic choices: a flat prior and a logarithmic prior (i.e. flat in the magnitude of M_S). We perform the analysis for the two priors, even though we think the logarithmic one is more realistic, taking into account what we know about mechanisms of SUSY breaking. This allows us to quantify the dependence of the results on the choice of the prior.

The second part of the paper (sections 3-5) is devoted to incorporate all the important experimental constraints to the analysis, obtaining a map of the probability density of the MSSM parameter space, i.e. the MSSM forecast for the LHC. Since not all the experimental information is equally robust, we perform separate analyses depending on the group of observables used. The main results are the following:

- First we include only the most robust experimental data: E.W. and B(D)-physics observables, and lower bounds on the masses of supersymmetric particles and the Higgs mass. Then, the favoured region of the MSSM parameter space lies at low-energy, but there is a significant portion out of the LHC reach. This effect is more prominent in the case of flat prior, but it is visible both for flat and logarithmic prior. The main responsible for this situation is the lower bound on the Higgs mass: to become consistent with experiment the Higgs mass needs sizeable radiative corrections, which require larger soft terms. We show that increasing the Higgs mass in a few GeV affects substantially the amount of parameter space within the LHC reach. In consequence, if we wish to detect SUSY at LHC, let us hope that the Higgs mass is close to the present experimental limit. We also show the present preferred credibility interval for the Higgs mass (Table 3).
- Then we add the information about a_μ . As is well-known, the impact of this observable depends dramatically on the way one computes the SM hadronic contribution. Using $e^+e^- \rightarrow \text{had}$ data (the most common choice in the literature), the soft terms are dramatically pushed into the low-energy region (for $\mu > 0$), well inside the LHC reach. Actually, the push is so strong that the predictions for other observables, in particular $b \rightarrow s \gamma$, start to be too large. Furthermore, if the Higgs mass turns out to be $\mathcal{O}(10)$ GeV above the present experimental limit, the tension between the Higgs mass and a_μ would be dramatic and could not be reconciled: $m_h(a_\mu)$ would require too large (small) soft masses.

Using τ -decay data, instead $e^+e^- \rightarrow \text{had}$, there is no big discrepancy between a_μ^{SM} and a_μ^{exp} , so the SUSY contribution does not need to be large. Consequently, the probability distributions are essentially unchanged by the inclusion of the a_μ constraint. Although the more direct e^+e^- data are usually preferred to evaluate a_μ^{SM} , this discrepancy is warning us to be cautious about this procedure.

- We then consider the impact of Dark Matter (DM) constraints (unplugging a_μ constraints), namely we require that the WIMP responsible for Ω_{DM} is the supersymmetric LSP (typically the lightest neutralino). As is known, this selects four regions in the parameter space: Bulk, Focus point, Co-annihilation and Higgs-funnel. When all the SUSY parameters, but the universal scalar and gaugino mass (m and M respectively), are marginalized, these regions appear clearly in the $m - M$ plots. One can observe a certain blurring with respect to usual plots in the literature, due to the integration in the other variables.

Roughly speaking, including DM constraints the low-energy gets favoured and therefore the detection of SUSY at the LHC. However, there survive large high-energy areas out of the LHC reach. Consequently, again, if we are unlucky, even if DM is supersymmetric, it could escape LHC detection (especially if the Higgs mass is not close to its present experimental limit). In any case, we have stressed that one should be cautious at interpreting these results as a robust constraint on the CMSSM, discussing possible ways-out for the (in principle) disfavoured regions.

- Finally we consider all constraints at the same time (including a_μ and Ω_{DM}). The bulk and co-annihilation regions are now clearly selected (for $\mu > 0$) amongst the various possibilities to obtain Ω_{DM} . Again, these results have to be taken with caution.
- We perform a similar analysis for $\mu < 0$. The most important change in the results is that when a_μ (using e^+e^- data) is taken into account, this scenario becomes unfavoured, as it cannot reproduce the experimental measure. Besides, there is no push of the MSSM parameters into the low energy region (actually, the opposite is true, but in a mild way).

We also perform a Bayesian comparison of both scenarios, showing quantitatively how better the $\mu > 0$ case performs than the $\mu < 0$ one. Actually, the advantage of the positive- μ case only occurs when a_μ constraints (using e^+e^- data) are taken into account.

In summary, LHC offers an exciting horizon for SUSY discovery, but there is still a non-negligible possibility that it escapes detection, especially if the Higgs mass is not close to its present experimental value. So we should cross fingers.

Acknowledgements

We thank P. Slavich for interesting discussions and suggestions. M. Eugenia Cabrera thanks

Ben Allanach for useful discussions and the hospitality of DAMTP (Cambridge) during a two months visit. J.A. Casas thanks A. Ibarra for useful discussions and the support and hospitality of TUM (Munich) during a one month visit. This work has been partially supported by the MICINN, Spain, under contract FPA 2007–60252, the Comunidad de Madrid through Proyecto HEPHACOS S-0505/ESP–0346, and by the European Union through the UniverseNet (MRTN–CT–2006–035863). M. E. Cabrera acknowledges the financial support of the CSIC through a predoctoral research grant (JAEPre 07 00020). The work of R. Ruiz de Austri has been supported in part by MEC (Spain) under grant FPA2007-60323, by Generalitat Valenciana under grant PROMETEO/2008/069 and by the Spanish Consolider-Ingenio 2010 Programme CPAN (CSD2007-00042). The use of the ciclope cluster of the IFT-UAM/CSIC is also acknowledged.

References

- [1] B. C. Allanach and C. G. Lester, *Phys. Rev. D* **73** (2006) 015013 [arXiv:hep-ph/0507283].
- [2] B. C. Allanach, *Phys. Lett. B* **635** (2006) 123 [arXiv:hep-ph/0601089].
- [3] R. R. de Austri, R. Trotta and L. Roszkowski, *JHEP* **0605** (2006) 002 [arXiv:hep-ph/0602028].
- [4] B. C. Allanach, K. Cranmer, C. G. Lester and A. M. Weber, *JHEP* **0708** (2007) 023 [arXiv:0705.0487 [hep-ph]].
- [5] L. Roszkowski, R. Ruiz de Austri and R. Trotta, *JHEP* **0707**, 075 (2007) [arXiv:0705.2012 [hep-ph]].
- [6] O. Buchmueller *et al.*, *JHEP* **0809** (2008) 117 [arXiv:0808.4128 [hep-ph]].
- [7] R. Trotta, F. Feroz, M. P. Hobson, L. Roszkowski and R. Ruiz de Austri, arXiv:0809.3792 [hep-ph].
- [8] J. Ellis, *Eur. Phys. J. C* **59** (2009) 335 [arXiv:0810.1178 [hep-ph]].
- [9] S. S. AbdusSalam, B. C. Allanach, F. Quevedo, F. Feroz and M. Hobson, arXiv:0904.2548 [hep-ph].
- [10] O. Buchmueller *et al.*, *Eur. Phys. J. C* **64** (2009) 391 [arXiv:0907.5568 [hep-ph]].
- [11] G. D’Agostini, arXiv:hep-ph/9512295; R. Trotta, *Contemp. Phys.* **49** (2008) 71 [arXiv:0803.4089 [astro-ph]].
- [12] For a review see S. P. Martin, arXiv:hep-ph/9709356.
- [13] J. R. Ellis, K. Enqvist, D. V. Nanopoulos and F. Zwirner, *Mod. Phys. Lett. A* **1** (1986) 57.
- [14] R. Barbieri and G. F. Giudice, *Nucl. Phys. B* **306** (1988) 63.
- [15] B. de Carlos and J. A. Casas, *Phys. Lett. B* **309** (1993) 320 [arXiv:hep-ph/9303291]; G. W. Anderson and D. J. Castano, *Phys. Lett. B* **347** (1995) 300 [arXiv:hep-ph/9409419]; P. Athron and D. J. Miller, *Phys. Rev. D* **76** (2007) 075010 [arXiv:0705.2241 [hep-ph]]; P. Ciafaloni and A. Strumia, *Nucl. Phys. B* **494** (1997) 41 [arXiv:hep-ph/9611204]; J. A. Casas, J. R. Espinosa and I. Hidalgo, *JHEP* **0411** (2004) 057 [arXiv:hep-ph/0410298]; [ibid], *JHEP* **0503** (2005) 038 [arXiv:hep-ph/0502066].

- [16] L. Giusti, A. Romanino and A. Strumia, Nucl. Phys. B **550** (1999) 3 [arXiv:hep-ph/9811386].
- [17] A. Strumia, arXiv:hep-ph/9904247; F. Feruglio, A. Strumia and F. Vissani, Nucl. Phys. B **637** (2002) 345 [Addendum-ibid. B **659** (2003) 359] [arXiv:hep-ph/0201291].
- [18] M. E. Cabrera, J. A. Casas and R. Ruiz de Austri, JHEP **0903**, 075 (2009) [arXiv:0812.0536 [hep-ph]].
- [19] J. O. Berger, B. Liseo and R. L. Wolpert, Statistical Science 1999, Vol. 14, No. 1, 1-28
- [20] B. C. Allanach, Comput. Phys. Commun. **143** (2002) 305 [arXiv:hep-ph/0104145].
- [21] L. E. Ibanez and C. Lopez, Nucl. Phys. B **233** (1984) 511.
- [22] V. S. Kaplunovsky and J. Louis, Phys. Lett. B **306** (1993) 269 [arXiv:hep-th/9303040].
- [23] G. F. Giudice and A. Masiero, Phys. Lett. B **206** (1988) 480.
- [24] F. Feroz and M. P. Hobson Mon. Not. Roy. Astron. Soc. **384** 449 (2008).
- [25] Available from: <http://superbayes.org>
- [26] J. Skilling, Nested sampling, in R. Fischer, R. Preuss and U. von Toussaint (Eds) , 735 (Amer. Inst. Phys. conf. proc. 2004), pp. 395–405.
- [27] J. Skilling, Bayesian Analysis **1** 833–861 (2006).
- [28] G. L. Bayatian *et al.* [CMS Collaboration], J. Phys. G **34** (2007) 995; G. Aad *et al.* [The ATLAS Collaboration], arXiv:0901.0512 [hep-ex].
- [29] H. Baer, V. Barger, A. Lessa and X. Tata, JHEP **0909** (2009) 063 [arXiv:0907.1922 [hep-ph]].
- [30] L. Susskind, arXiv:hep-th/0405189.
- [31] By CDF Collaboration and D0 Collaboration, arXiv:0803.1683 [hep-ex].
- [32] W.-M. Yao *et al.*, *J. Phys.* **G33** (2006) 1 and 2007 partial update for the 2008 edition.
- [33] K. Hagiwara, A. D. Martin, D. Nomura and T. Teubner, *Phys. Lett.* **B 649** (2007) 173.
- [34] R. Barbieri and A. Strumia, arXiv:hep-ph/0007265. R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, Nucl. Phys. B **703** (2004) 127 [arXiv:hep-ph/0405040].
- [35] G. Degrossi, P. Gambino and P. Slavich, Comput. Phys. Commun. **179** (2008) 759 [arXiv:0712.3265 [hep-ph]].
- [36] G. Degrossi, P. Gambino and P. Slavich, Phys. Lett. B **635** (2006) 335 [arXiv:hep-ph/0601135].
- [37] G. D’Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, Nucl. Phys. B **645**, 155 (2002) [arXiv:hep-ph/0207036].
- [38] F. Mahmoudi, Comput. Phys. Commun. **180**, 1579 (2009) [arXiv:0808.3144 [hep-ph]].
- [39] See <http://lepewwg.web.cern.ch/LEPEWWG>.
- [40] J. P. Miller, E. de Rafael and B. L. Roberts, Rept. Prog. Phys. **70**, 795 (2007) [arXiv:hep-ph/0703049].
- [41] M. Davier *et al.*, arXiv:0906.5443 [hep-ph].
- [42] The CDF Collaboration, *Phys. Rev. Lett.* **97** (2006) 062003. and *Phys. Rev. Lett.* **97** (2006) 242003.

- [43] Heavy Flavor Averaging Group (HFAG) (E. Barberio *et al.*), arXiv:0808.1297 [hep-ex].
- [44] B. Aubert *et al.* [BABAR Collaboration], arXiv:0808.1915 [hep-ex].
- [45] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **100** (2008) 021801 [arXiv:0709.1698 [hep-ex]].
- [46] M. Antonelli *et al.* [FlaviaNet Working Group on Kaon Decays], arXiv:0801.1817 [hep-ph].
- [47] A. G. Akeroyd and F. Mahmoudi, JHEP **0904** (2009) 121 [arXiv:0902.2393 [hep-ph]].
- [48] J. Dunkley *et al.* [The WMAP Collaboration], arXiv:0803.0586 [astro-ph].
- [49] The CDF Collaboration, Phys. Rev. Lett. **97** (2006) 242003.
- [50] The LEP Higgs Working Group, <http://lephiggs.web.cern.ch/LEPHIGGS>; G. Abbiendi *et al.* [the ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration and the OPAL Collaboration, The LEP Working Group for Higgs Boson Searches], Phys. Lett. **B 565** (2003) 61.
- [51] J. R. Ellis, G. Ridolfi and F. Zwirner, Phys. Lett. B **262** (1991) 477; Y. Okada, M. Yamaguchi and T. Yanagida, Prog. Theor. Phys. **85** (1991) 1; H. E. Haber and R. Hempfling, Phys. Rev. Lett. **66** (1991) 1815; R. Barbieri, M. Frigeni and F. Caravaglios, Phys. Lett. B **258** (1991) 167.
- [52] M. Davier, A. Hoecker, B. Malaescu, C. Z. Yuan and Z. Zhang, arXiv:0908.4300 [hep-ph].
- [53] S. Heinemeyer, D. Stockinger and G. Weiglein, Nucl. Phys. B **699**, 103 (2004) [arXiv:hep-ph/0405255].
- [54] S. Marchetti, S. Mertens, U. Nierste and D. Stockinger, Phys. Rev. D **79** (2009) 013010 [arXiv:0808.1530 [hep-ph]].
- [55] F. Feroz, M. P. Hobson, L. Roszkowski, R. Ruiz de Austri and R. Trotta, arXiv:0903.2487 [hep-ph].
- [56] G. Jungman, M. Kamionkowski and K. Griest, Phys. Rept. **267** (1996) 195 [arXiv:hep-ph/9506380].
- [57] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, Comput. Phys. Commun. **149** (2002) 103 [hep-ph/0112278]; Comput. Phys. Commun. **174**, 577 (2006) [hep-ph/0405253].
- [58] L. Knox and M. S. Turner, Phys. Rev. Lett. **70** (1993) 371 [arXiv:astro-ph/9209006]; J. Garcia-Bellido, D. Y. Grigoriev, A. Kusenko and M. E. Shaposhnikov, Phys. Rev. D **60** (1999) 123504 [arXiv:hep-ph/9902449].
- [59] A. Ibarra and D. Tran, JCAP **0902** (2009) 021 [arXiv:0811.1555 [hep-ph]].
- [60] H. Jeffreys, Theory of probability, 3rd edn, Oxford Classics series (reprinted 1998) (Oxford University Press, Oxford, UK, 1961).
- [61] J. O. Ruanaidh and W. Fitzgerald, Numerical Bayesian Methods Applied to Signal Processing. Springer Verlag: New York, 1996.