Universidad Autónoma de Madrid

Facultad de Ciencias
Departamento de Física Teórica

Consejo Superior de Investigaciones Científicas

Instituto de Física Teórica IFT
UAM-CSIC

A Bayesian trek in the search of Supersymmetry and other New Physics at the LHC

María Eugenia Cabrera Catalán

Madrid, Octubre 2011.

Universidad Autónoma de Madrid

Facultad de Ciencias
Departamento de Física Teórica

Consejo Superior de Investigaciones Científicas

Instituto de Física Teórica IFT
UAM-CSIC

A Bayesian trek in the search of Supersymmetry and other New Physics at the LHC

Memoria de Tesis doctoral realizada por

María Eugenia Cabrera Catalán

presentada ante el Departamento de Física Teórica
de la Universidad Autónoma de Madrid
para la obtención del Doctorado

Proyecto dirigido por

J. Alberto Casas
Científico Titular del Instituto de Física Teórica UAM - CSIC

Madrid, Octubre 2011.

Contents

Introduction: Motivations, Goals and Main Results

There is a broad consensus in the scientific community about the validity of the Standard Model (SM), which describes with impressive precision the phenomenology of particle physics until energies around 1 TeV. However the SM presents deficiencies which suggest that it is an effective theory coming from a more fundamental one. From the phenomenological point of view, the most important reason to expect new physics is perhaps the overwhelming evidence of dark matter in the universe, which cannot be explained within the SM framework. From the theoretical point of view, the SM, as any other quantum field theory, is inconsistent with gravitation at the quantum level, though this is not problematic from the practical point of view until energies comparable with the Planck mass ($\sim 10^{19}$ GeV), completely inaccesible for present experiments. Besides, the process of electroweak symmetry breaking is a source of theoretical trouble.

As a matter of fact, a crucial prediction of the mechanism of electroweak breaking in the SM, namely the Higgs boson, has not yet been discovered. Furthermore, the SM predicts huge radiative contributions to the Higgs mass, which would make the electroweak scale many orders of magnitude larger than its actual value, unless some miraculous cancellation takes place or there is new physics beyond the SM at $\mathcal{O}$(TeV), i.e. very likely within the reach of LHC. Actually, the two main physical goals of LHC are the discovery of the Higgs boson and the detection of new physics beyond the SM. The LHC is already working at impressive performance and, although it has not discovered yet the Higgs boson nor new physics, it is already puting important constraints on them. Undoubtedly these are very changing and exciting times for particle physics, and the work of the present thesis has been realized in this context.

More precisely, most of our work has been done in relation with the possible existence and detection of Supersymmetry (SUSY) at LHC. SUSY is one of the few candidates for physics beyond the SM that is really well motivated from the theoretical point of view and allows to perform detailed caculations and thus realize precise predictions. Indeed SUSY has been the most extensively studied candidate for new physics in the last decades and the first LHC analyses are using SUSY as a paradigmatic scenario of new physics to present their constraints on physics beyond the SM. Of course, this does not mean that SUSY is really there, but clearly is a most serious scenario to be considered in the light of the LHC.

One unpleasant aspect of SUSY, and indeed of most of the scenarios of new physics, is that the low-energy theory is defined by a large number of unknown parameters. In the case of SUSY, most of them are soft-breaking terms that have to do with the unknown mechanism of SUSY breaking. This makes dificult to forsee which values of those parameters are most probable and to identify which kind of SUSY is really there, even if there is a signal of new physics and SUSY is the actual theory behind.

Our primary goal has been precisely to establish, in the most rigorous possible way, the distribution of probability in the supersymmetric parameter space. Such analysis can only be done using Bayesian techniques and with the help of modern algorithms of sampling, and requires to take into account all the experimental constraints and theoretical information related to SUSY. The analysis involves also some subtleties in the statistical treatment (e.g. concerning Jacobian factors) that have not taken into account in previous literature. The set up for a Bayesian analysis of SUSY is described in chapter 1 of the present memory, and it is valid, not only for the work presented here but for any future exploration of this kind.

A particularly interesting result found has to do with the existence of a quantification of fine-tuning. It is common lore that the higher the supersymmetric masses, the more fine-tuned is the process of electroweak breaking, in order to produce the correct electroweak scale. Such fine-tuning has been quantified in a rather arbitrary (though reasonable) way and incorporated in supersymmetric analyses as an *ad hoc* penalization in the parameter space, so that high-energy values of the soft parameters become disfavoured. We have found that such penalization arises in an automatic way when performing a correct Bayesian analysis. Besides, the resulting penalization of the fine-tuning is very similar to (though not exactly the same as) the standard measures of fine-tuning.

We have done most of our work in the context of the so-called Constrained Minimal Supersymmetric Standard Model (CMSSM). This is a particular class of minimal supersymmetric models (MSSM), characterized by the universality of soft terms at high energy. The CMSSM is both reasonable and representative of most MSSMs, since universality is to a great extent required by experimental constraints on flavour violation and, besides, usual realistic scenarios of SUSY do not differ much from the CMSSM. At the time of starting our work LHC had not yet begun to work and produce the first significant results. In consequence, we considered all the available experimental data that constrain SUSY in any way: direct limits on supersymmetric masses, Higgs boson mass, electroweak observables, B and D observables, etc. Experimental data concerning $g_\mu - 2$ and dark matter are also incorporated, but in a separate way (switching them on and off), since they are not as robust as the other constraints or even controversial. The details and results of our analysis are presented in chapter 2, together with a comparison with the previous literature, including frequentist approaches, which are an alternative and complementary procedure to Bayesian ones. Our results indicate that there is a considerable probability that, even if SUSY is there, it is not discovered in the first stages of LHC.

Bayesian techniques are also required to quantify present or potential tensions between observables within a specific model, in our case the MSSM. In this way we have pointed out that a "large" Higgs mass (i.e. not close to the present lower bound) is rather incompatible

with the $g-2$ observable, if the latter is evaluated using $e^+e^- \to$ had data. We have quantified this tension as a function of the (yet unknown) Higgs mass. These results are presented in chapter 3.

Our next motivation has been the incorporation of direct LHC data in the searches of new physics using Bayesian techniques. Due to the complexity of the LHC experiment, much of the comparison between LHC data and theoretical predictions has to be made by confronting experimental histograms (in different variables) and theoretical histograms produced by simulations. Thus our goal has been to develope a rigorous and effective way to compare experimental and theoretical histograms, and calculate the corresponding likelihood, taking into account the different sources of statitical and systematic uncertainties. This is very useful to extract as much information as possible from the comparison between experimental data with theoretical simulations, optimizing the chances of identifying new physics at the LHC. We have illustrated the latter point by showing how a search in the CMSSM parameter space, using Bayesian techniques, can effectively find the correct values of the CMSSM parameters by comparing histograms of events with multijets + missing energy displayed in the effective-mass variable. The procedure is in fact very efficient to identify the true supersymmetric model, in the case supersymmetry is really there and accesible to the LHC. But, of course, the technique can be applied to any scenario of new physics. These results are presented in chapter 4.

Finally, in the last months we have seen the first relevant LHC results concerning new physics and Higgs. An already substantial portion of the CMSSM parameter space has been discarded, though there is still room for SUSY. On the other hand, the window for the Higgs mass has shrunk dramatically. It is funny that the latter leads to upper bounds on the supersymmetric masses, which are complementary to the direct lower bounds from SUSY searches. We have found out which are the present bounds in the MSSM parameter space as a result of the present and future upper bounds on the Higgs mass, both in an ordinary and a split-SUSY context. It is remarkable that, thanks to the LHC data, one is not free any more to say that supersymmetric masses can be safely pushed arbitrarily to a large scale (e.g. of order the Planck mass), even paying a fine-tuning price. These results are expounded in chapter 5.

Introducción: Motivaciones, Objetivos y Resultados Principales

Existe un amplio consenso en la comunidad científica acerca de la valides del Modelo Estándar (ME), que describe con impresionante precisión la fenomenología de física de partículas hasta energías de alrededor de 1 TeV. Sin embargo el ME presenta deficiencias que sugieren que es una teoría efectiva proveniente de una teoría mas fundamental. Desde el punto de vista fenomenológico, la razón mas importante para esperar nueva física es quizás la abrumadora evidencia de la existencia de materia oscura en el universo, que no puede ser explicada por el ME. Desde el punto de vista teórico, el ME, como cualquier otra teoría cuántica de campo, es inconsistente con gravedad a nivel cuántico, aunque esto no es problemático, desde un punto de vista práctico, hasta energías comparables con la masa de Plank ($\sim 10^{19}$ GeV), completamente inaccesible para los experimentos actuales. Además, el proceso de ruptura de la simetría electrodébil es una fuente de problemas teóricos.

De hecho, una predicción crucial del mecanismo de ruptura electrodébil en el ME, el bosón de Higgs, aún no ha sido descubierta. Además, el ME predice inmensas contribuciones radiativas a la masa de Higgs, que podrían hacer que la escala electrodébil sea muchos ordenes de magnitud mayor que su valor actual, a menos que se de una "milagrosa" cancelación o exista nueva física mas allá del ME a $\mathcal{O}$(TeV), muy probablemente dentro del alcance del LHC. Actualmente, los dos objetivos físicos principales del LHC son el descubrimiento del bosón de Higgs y la detección de nueva física mas allá del ME. El LHC ya esta trabajando con un desempeño impresionante y, aunque aún no se ha descubierto el ni bosón de Higgs ni nueva física, ya esta poniendo nuevos limites de exclusión. Indudablemente estos son tiempo muy cambiantes y emocionantes para la física de partículas, y el trabajo de la presente tesis ha sido realizado en este contexto.

Mas precisamente, la mayoría de nuestro trabajo ha sido realizado en relación con la posible existencia y detección de Supersimetría (SUSY) en el LHC. SUSY es una de las pocas candidatas a física mas allá del ME que está muy bien motivada desde el punto de vista teórico y permite realizar cálculos detallados y con ello hacer predicciones precisas. En realidad SUSY ha sido la teoría mas estudiada como candidata a nueva física en las últimas décadas. Los primeros análisis de LHC están utilizando SUSY como un escenario paradigmático de nueva física para presentar las restricciones sobre física mas allá del ME. Por supuesto, esto no significa que SUSY existe, pero claramente es el escenario mas serio para ser considerado

a la luz del LHC.

Un aspecto desagradable de SUSY, y de hecho de la mayoría de los escenarios de nueva física, es que la teoría a bajas energías esta definida por un número muy grande de parámetros desconocidos. En el caso de SUSY, la mayoría de estos parámetros son términos de ruptura *soft* que tienen que ver con el desconocido mecanismo de ruptura de SUSY. Esto hace difícil prever cual es el valor mas probable que estos parámetros pueden tener e identificar que tipo de SUSY está detrás, aún si realmente existe esa señal de nueva física y SUSY es la teoría correcta detrás de dicha señal.

Nuestro principal objetivo ha sido precisamente establecer, de la forma mas rigurosa posible, la distribución de probabilidad del espacio de parámetros Supersimétrico. Dicho análisis puede realizarse solamente utilizando técnicas Bayesianas y con la ayuda de modernos algoritmos de muestreo, se requiere tener en cuenta todos las restricciones experimentales y la información teórica relacionada con SUSY. El análisis también involucra sutilezas en el tratamiento estadístico (por ejemplo, con respecto al factor jacobiano) que no han sido tomadas en cuenta en previos anáisis. La puesta a punto de un análisis Bayesiano para SUSY está descrita en el capítulo 1 de la presente memoria, y es válido, no solo para este trabajo sino para cualquier exploración futura de este tipo.

Uno de los resultados que hemos encontrado es particularmente interesante, tiene que ver con la existencia de una cuantificación del *fine-tuning*. Es bien sabido que cuanto mas grande sean las masas supersimétricas, mas *fine-tuning* tendrá el proceso de ruptura electrodébil para poder producir la escala electrodébil correcta. Dicho *fine-tuning* ha sido cuantificado de una manera mas bien arbitraria (aunque razonable) e incorporada en análisis supersimétricos como una penalización *ad hoc* en el espacio de parámetros, de modo que valores de alta energía para los parámetros *soft* están desfavorecidos. Hemos encontrado que dicha penalización surge de forma automática cuando realizamos un análisis Bayesiano correctamente. Además, la penalización de *fine-tuning* resultante es muy parecida (aunque no exactamente la misma) a la medición estándar de *fine-tuning*.

Hemos hecho la mayor parte de nuestro trabajo en el contexto del llamado Modelo Supersimétrico Minimal Constreñido (CMSSM). Este es una clase particular de Modelo Supersimétrico Minimal (MSSM), caracterizado por la universalidad de los términos *soft* a altas energías. El CMSSM es razonable y representativo de la mayoría de los MSSMs, ya que la universalidad es en gran parte requerida por los resultados experimentales de violación de sabor, además, escenarios realistas de SUSY usualmente no difieren mucho del CMSSM. En el tiempo en que empezamos este trabajo el LHC no había empezado a producir los primeros resultados significativos. En consecuencia, consideramos todos los datos experimentales disponibles que de alguna manera constriñen SUSY: limites directos sobre masas supersimétricas, la masa del bosón de Higgs, observables electrodébiles, física del B y D, etc. También fueron incorporados los datos experimentales concernientes a $g_\mu - 2$ y materia oscura, pero de forma separada (quitándolos e incluyéndolos), ya que no son tan robustos como los otros datos o mas bien son controversiales. Los detalles y resultados de nuestro análisis se presentan en el capítulo 2, junto con una comparación de nuestro trabajo con trabajos previos, incluyendo enfoques frecuentistas, que son procedimientos alternativos y

complementarios al enfoque Bayesiano. Nuestro resultado indica que existe una probabilidad considerable de que aun si SUSY esta allí, no sea descubierta en la primera etapa del LHC.

Las técnicas Bayesianas son también requeridas para cuantificar presentes o potenciales tensiones entre observables para un modelo dado, en nuestro caso el MSSM. De este modo, hemos señalado que una masa "grande" del Higgs (lejos del presente límite experimental) es mas bien incompatible con el observable $g-2$, si es evaluado usando los datos de $e^+e^- \to$ had. Hemos cuantificado esta tensión como función de la masa del Higgs (aún desconocida). Este resultado es presentado en el capítulo 3.

Nuestra siguiente motivación ha sido la incorporación de datos del LHC en la búsqueda de nueva física usando técnicas Bayesianas. Debido a la complejidad del LHC, la mayor parte de la comparación de los datos de LHC y las predicciones teóricas han tenido que hacerse comparando histogramas experimentales (en distintas variables) con histogramas teóricos obtenidos con simulaciones. Nuestro objetivo ha sido desarrollar un método riguroso y efectivo para comparar histogramas teóricos y experimentales, calcular su correspondiente likelihood, teniendo en cuenta las distintas fuentes de errores estadísticos y sistemáticos. Esto es muy útil para extraer tanta información como sea posible de la comparación de datos experimentales con simulaciones teóricas, optimizando las posibilidades de identificar nueva física en el LHC. Hemos ilustrado este último punto mostrando cómo una búsqueda en el espacio de parámetros del CMSSM, usando técnicas Bayesianas, puede efectivamente encontrar el punto correcto del espacio de parámetros del CMSSM comparando histogramas con eventos de multijets + energía perdida mostrados en la variable de la Masa Efectiva. El procedimiento es de hecho muy eficiente para identificar el verdadero modelo supersimétrico, en el caso en el que supersimetría exista y sea accesible al LHC. Pero, por supuesto, la técnica puede ser aplicada a muchos escenarios de nueva física. Este resultado está presente en el capítulo 4.

Finalmente, en los últimos meses hemos visto los primero resultados relevantes de LHC con respecto al Higgs y nueva física. Una región sustancial del espacio de parámetros del CMSSM ha sido descartada, aunque hay todavía espacio para SUSY. Por otro lado, la ventana para la masa del Higgs se ha reducido dramáticamente. Es divertido que esto último conduzca a límites superiores en las masas supersimétricas, que son complementarios con el presente límite inferiores provenientes de las búsquedas de SUSY. Hemos calculado cuales son los límites actuales en el espacio de parámetros del MSSM como resultado del presente y futuros límites superiores en la masa del Higgs, ambos en el contexto del MSSM y de split-SUSY. Es de destacar que, gracias a los datos del LHC, ya no se es libre de decir que las masas supersimétricas pueden estar a arbitrariamente alta escala (por ejemplo a la escala de Plank), aún pagando un precio por el *fine-tuning*. Estos resultados están expuestos en el capítulo 5.

Chapter 1

Bayesian statistics, Supersymmetry and Naturalness

As mentioned in the Introduction, supersymmetry is one of the few well motivated candidates for physics beyond the SM. *If* supersymmetry is responsible for the hierarchy problem of the electroweak scale, there are big chances that could be discovered at the LHC. This has motivated an important effort to anticipate which kind of supersymmetric model is more likely to be there (in the case it is) or, in more precise words, which region of the parameter space of the minimal supersymmetric standard model (MSSM) is more probable, taking into account the present (theoretical and experimental) wisdom about the model. This wisdom includes theoretical constraints (and perhaps prejudices) and experimental constraints. The appropriate framework to evaluate this probability is the Bayesian approach, which allows to separate in a neat way the objective and subjective pieces of information. In this chapter we construct the set up for any Bayesian analysis of supersymmetry, which will be applied in later chapters.

In sect. 1.1 we address some basic aspects of the Bayesian approach. In sect. 1.2 we briefly introduce the MSSM framework and some notation. In sect. 1.3 we discuss the potential fine-tuning associated with the process of electroweak breaking in the supersymmetric context, and show that a penalization of this fine-tuning arises from the Bayesian analysis itself (with no ad hoc assumptions as in analyses of previous literature), upon the marginalization of the μ parameter. We also present a rigorous treatment of the Yukawa couplings. In sect. 1.4 we use an efficient set of variables to scan the MSSM parameter space, trading in particular B by $\tan\beta$, giving the effective prior in the new parameters. In sect. 1.5 we present a satisfactory choice of priors for the initial parameters (actually two different choices to evaluate the dependence on the prior). In sect. 1.6 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.

1.1 Bayesian Statistics

Let us start by recalling some basic notions of Bayesian inference. We refer the reader to [1, 2] for further details. In the Bayesian analysis one tries to make inferences about the relative probability of different "states of nature" (corresponding to different values of the parameters defining the model) upon the observation of different data. Normally this determination takes the form of a probability distribution since the theoretical computations and the experimental data are affected by different kinds of errors and uncertainties.

For a model defined by a set of parameters $\{\theta\}$, the probability of $\{\theta\}$ given a set of *data*, the posterior probability density function (pdf), is denoted by $p(\theta|\text{data})$ and it is obtained via Bayes theorem as

$$p(\theta|\text{data}) \;=\; \frac{p(\text{data}|\theta)\; p(\theta)}{p(\text{data})}\;. \tag{1.1}$$

Here $p(\text{data}|\theta)$ is the likelihood function [1]. $p(\theta)$ is the prior, i.e. the probability density that we assign the points in the parameter space before seeing the data. Finally, $p(\text{data})$ is the *evidence*, sometimes called $\mathcal{Z}$, a normalization factor unless one wishes to compare different classes of models as we will see below.

One can say that in eq. (1.1) the the likelihood is objective, while the prior is subjective, since it contains our prejudices about which regions of the parameter space are more "natural" or "expectable". It is desirable that the results of the analysis are as independent as possible of the chosen prior. This happens if the data are powerful enough to select a very small region of the parameter space, so that eq. (1.1) is dominated by the likelihood, i.e. essentially the pdf is non-zero just in the narrow region of non-vanishing $p(\text{data}|\theta)$. However, in many instances this is not the case.

The somewhat subjective character of the prior, $p(\theta)$, has often motivated to ignore its presence, identifying in practice $p(\theta|\text{data})$ with $p(\text{data}|\theta)$. However, it must be noticed that this procedure implicitly implies a choice for the prior, namely a completely flat prior in the parameters. This is not necessarily the most reasonable or "free of prejudices" attitude. Note for example that using 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.

If one is interested in the most probable value of one (or several) of the initial parameters, say $\theta_i,\ i = 1, ..., n_1$, but not in the others, $\theta_i,\ i = n_1 + 1, ..., N$, we have to *marginalize* the latter, called "Nuisance parameters", i.e. integrate in the parameter space:

$$p(\theta_i,\ i = 1, ..., n_1|\text{data}) \;=\; \int \mathrm{d}\theta_{n_1+1}, ..., d\theta_N\; p(\theta_i,\ i = 1, ..., N|\text{data})\;. \tag{1.2}$$

This procedure is very useful and common to make predictions about the values of particularly interesting parameters. It must be noticed that, in order to perform the marginalisation,

[1] Frequentist approaches, which are an alternative to the Bayesian framework, are based on the analysis of the likelihood function in the parameter space.

$\lvert \ln B_{10} \rvert$	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 1.1: The scale we use for the interpretation of model probabilities.

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).

The evidence, integrating both sides of eq.(1.1), and using the fact that $p(\theta|\text{data})$ must be correctly normalized, one simply obtains

$$\mathcal{Z} \equiv p(\text{data}) = \int d\theta_1 \cdots d\theta_N \; p(\text{data}|\theta_i) \; p(\theta_i) \; , \tag{1.3}$$

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.

If one is interested in constraining the parameters of a model, the evidence is merely a normalization constant, and can therefore be dropped. However, the evidence is very useful in the context of Bayesian model comparison. When two different models (or hypotheses) are used to fit the data, the ratio of their evidences gives the relative probability of the two models in the light of the data. Let $\mathcal{M}_1$, $\mathcal{M}_2$ be two models with prior probabilities $p(\mathcal{M}_1)$, $p(\mathcal{M}_2)$. Then the relative posterior probability of the two models, given a set of data, is simply

$$\frac{p(\mathcal{M}_1|data)}{p(\mathcal{M}_2|data)} = \frac{\mathcal{Z}_1 \; p(\mathcal{M}_1)}{\mathcal{Z}_2 \; p(\mathcal{M}_2)} = B_{12} \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}, \tag{1.4}$$

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 1.1, we summarize the translation of the Bayes factor to relative probabilities and a conventional interpretation of them [3].

By the other hand, if one wants to quantify the tension between two set of data,

$$\{\text{data}\} = \{\mathscr{D}, D\}. \tag{1.5}$$

Here $\mathscr{D}$ represents the subset of observations, whose compatibity with the rest of the observations, D, (which are assumed to be correct) we want to test. One can construct the quantity $p(\mathscr{D}|D)$, i.e. the probability of measuring a certain value for $\mathscr{D}$, given the known values of the remaining observables, D,

$$p(\mathscr{D}|D) = \frac{p(\mathscr{D}, D)}{p(D)}. \tag{1.6}$$

Here $p(\mathscr{D}, D) = p(\text{data})$ is the joint evidence, as given by Eq. (1.3), i.e., the global probability of measuring both sets of data at the same time, and $p(D)$ is its equivalent but just for the D subset. The latter is a normalization factor which will soon cancel out.

Now, the consistency of $\mathscr{D}^{\rm obs}$ (the measured $\mathscr{D}$) with the rest of data, D, (given an specific model), can be tested by comparing $p(\mathscr{D}^{\rm obs}|D)$ with the value obtained using different values of $\mathscr{D}$, in particular the one that maximizes such probability, say $\mathscr{D}^{\rm max}$ (assuming the same reported error at the new central value). This gives a measure of the likelihood of the actual data, $\mathscr{D}^{\rm obs}$, under the assumption that the model is correct:

$$\frac{p(\mathscr{D}^{\rm obs}|D)}{p(\mathscr{D}^{\rm max}|D)} \;=\; \frac{p(\mathscr{D}^{\rm obs}, D)}{p(\mathscr{D}^{\rm max}, D)} \;\equiv\; \mathscr{L}(\mathscr{D}^{\rm obs}|D). \tag{1.7}$$

$\mathscr{L}(\mathscr{D}^{\rm obs}|D)$ is analogous to a likelihood ratio in data space, but integrated over all possible values of the parameters of the model. Therefore, it can be used as a test statistics for the likelihood of the data being tested, $\mathscr{D}^{\rm obs}$, in the context of the model used. Such test was called $\mathscr{L}-$test in Ref. [4]. Note that, as mentioned above, the $p(D)$ factor cancels out in the expression of $\mathscr{L}(\mathscr{D}^{\rm obs}|D)$, which is simply given by the ratio of the joint evidences.

1.2 The MSSM framework

Let us now particularize some of the previous general statements to the MSSM (for a review see [5]). 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:

$$\{\theta_i\} \;=\; \{m, M, A, B, \mu, s\}\ . \tag{1.8}$$

Here m, M and A are the universal scalar mass, gaugino mass and trilinear scalar coupling, while B is the bilinear scalar coupling. These are soft SUSY-breaking terms. The universality assumption for them is in part justified by the need of keeping the FCNC processes under control and it does come out naturally in several schemes of SUSY breaking mediation, e.g. minimal SUGRA or gauge-mediated models (for a review see [6] and [7] respectively). The μ parameter is the usual Higgs mass term in the superpotential, presumably with a magnitude similar to the soft breaking terms, as it is demanded by a successful electroweak breaking (see below). Finally, 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.

The universality assumption for the soft breaking termsis in part justified by the need of keeping the FCNC processes under control and it does come out naturally in several schemes of SUSY breaking mediation, e.g. minimal SUGRA or gauge-mediated models (for a review

see [6] and [7] respectively). Beside these four parameters one has to include the μ-parameter (i.e. the Higgs mass term in the superpotential) as an additional independent parameter, presumably with a magnitude similar to the soft breaking terms, as it is demanded by a successful electroweak breaking (see below). The notation used here is consistent with refs. [5, 8].

An important difference from the SM is that the MSSM contains two Higgs doublets, H_1, H_2, with expectation values $v_i = \langle H_i^0 \rangle$ determined by the parameters of the model upon minimization of the scalar potential, $V(H_1, H_2)$. They have to fulfill $2(v_1^2+v_2^2) = v^2 = (246\,\mathrm{GeV})^2$. The down-type-quark masses go like $m_d \sim y_d v_1 = y_d v \cos\beta$, where $\tan\beta \equiv v_2/v_1$. Similarly for the up-type-quarks $m_u \sim y_u v_2 = y_u v \sin\beta$, and for the charged leptons, $m_e \sim y_e v_1 = y_e v \cos\beta$. Hence the values of the Yukawa couplings which give the observed fermion masses depend on the derived parameter $\tan\beta$, a fact that will be relevant later in our discussion.

1.3 Connection between the Bayesian approach and the fine-tuning measure

It is common lore that the parameters of the MSSM, $\{m, M, A, B, \mu\}$, should not be far from the electroweak scale in order to avoid unnatural fine-tunings to obtain the correct scale of the electroweak breaking. This can be easily appreciated from the minimization of the tree-level form of the scalar potential, $V(H_1, H_2)$, which gives the expectation values of the Higgses, and thus the value of $M_Z^2 = \frac{1}{2}(g^2 + g'^2)(v_1^2 + v_2^2)$; namely

$$M_Z^2 = 2\,\frac{m_{H_1}^2 - m_{H_2}^2 \tan^2\beta}{\tan^2\beta - 1} - 2\mu^2 \quad . \tag{1.9}$$

Unless the $\mu-$term and the soft masses m_{H_i} (which upon the renormalization running depend also on the other soft terms) are close to the electroweak scale, a funny cancellation among the various terms in the right hand side of (1.9) is necessary to get the experimental M_Z.

A conventional measure of the degree of fine-tuning is given by the Barbieri-Giudice fine-tuning parameters [9]:

$$c_i = \left| \frac{\partial \ln M_Z^2}{\partial \ln p_i} \right|, \tag{1.10}$$

which weigh up the sensitivity of M_Z with respect to the parameters of the model, p_i. The global measure of the fine-tuning is taken as $c \equiv \max\{c_i\}$ or $c \equiv \sqrt{\sum c_i^2}$ [9, 10, 11, 12].

Previous studies have attempted to incorporate this fine-tuning measure to the Bayesian approach through the prior $p(p_i)$. In particular, in refs. [13, 14] a prior $p(p_i) \propto 1/c$ was proposed[2]. In principle this is not unreasonable since $1/c$ approximately indicates the probability of a cancellation among the various terms contributing to M_Z^2 to give a result $\lesssim (M_Z^{\mathrm{exp}})^2$.

[2]Another prior designed to catch the naturalness criterion has been proposed in ref. [15].

This can be intuitively seen as follows. Expanding $M_Z^2(p_i)$ around a point in parameter space that gives the desired cancellation, say $\mathcal{P}^0 \equiv \{p_i^0\}$, up to the linear term in the parameters, one finds that only a small neighborhood $\delta\mathcal{P} \sim \mathcal{P}^0/c$ around this point gives a value of M_Z^2 smaller or equal to the experimental value [10]. Hence, if one assumes that $\mathcal{P}$ could reasonably have taken any value of the order of magnitude of $\mathcal{P}^0$, then only for a small fraction $\sim 1/c$ of this region one gets $M_Z^2 \lesssim (M_Z^{\rm exp})^2$, thus the rough probabilistic meaning of c.

However, though reasonable, the above-mentioned proposals for priors are rather arbitrary, as the very measure of the fine-tuning is. On the other hand, 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" *ad hoc*[3]. This is in fact the case, as we are about to see.

Let us consider M_Z as an experimental data, on a similar foot to the rest of physical observables. Then the total likelihood reads

$$p(\text{data}|s, m, M, A, B, \mu) \;=\; N_Z \; e^{-\frac{1}{2}\chi_Z^2} \; \mathcal{L}_{\rm rest} \;, \tag{1.11}$$

where s represents the SM-like parameters, $\mathcal{L}_{\rm rest}$ is the likelihood associated to all the physical observables, except M_Z, and

$$\chi_Z^2 = \left(\frac{M_Z - M_Z^{\rm exp}}{\sigma_Z} \right)^2 \;, \tag{1.12}$$

where $\sigma_Z \ll M_Z^{\rm exp}$ is the experimental uncertainty in the Z mass; finally $N_Z = 1/\sqrt{2\pi\sigma_Z}$ is a normalization constant. Let us now use this sharp dependence on M_Z to marginalize the pdf in the $\mu-$parameter, performing a change of variable $\mu \to M_Z$:

$$\begin{aligned} p(s, m, M, A, B|\text{ data}) \;&=\; \int d\mu \; p(s, m, M, A, B, \mu|\text{data}) \\ &=\; N_Z \int dM_Z \left[\frac{d\mu}{dM_Z} \right] e^{-\chi_Z^2} \; \mathcal{L}_{\rm rest} \; p(s, m, M, A, B, \mu) \\ &\simeq\; \mathcal{L}_{\rm rest} \left[\frac{d\mu}{dM_Z} \right]_{\mu_0} p(s, m, M, A, B, \mu_0) \;. \end{aligned} \tag{1.13}$$

where μ_0 is the value of μ that reproduces the experimental value of M_Z for the given values of $\{s, m, M, A, B\}$. In the last line of (1.13) we have approximated $N_Z \; e^{-\frac{1}{2}\chi_Z^2} \simeq \delta(M_Z - M_Z^{\rm exp})$. Essentially the same result is obtained by performing the $\mu-$integration in the stationary point approximation. Now, comparing (1.13) to the definition of fine-tuning parameters (1.10), we can write

$$p(s, m, M, A, B|\text{ data}) \;=\; 2 \; \mathcal{L}_{\rm rest} \; \frac{\mu_0}{M_Z} \; \frac{1}{c_\mu} \; p(s, m, M, A, B, \mu_0) \;. \tag{1.14}$$

Several comments are in order here. First, the presence of the fine-tuning parameter, $1/c_\mu$, penalizes the regions of the parameter space with large fine-tuning, as desired. Actually

[3]This has also been noted and taken into account in [16, 17]

eq. (1.14) is very similar to multiply by hand the initial prior in the parameters by a factor $1/c$, as in ref. [14]. The difference is that here the factor $1/c_\mu$ has not been put by hand: it comes out from the marginalization in μ. Moreover the prior $p(s, m, M, A, B, \mu_0)$ is still undefined. If one takes it as flat, then one gets the same as in ref. [14], but with one factor μ in the numerator (still the regions of large fine-tuning are penalized since c_μ goes parametrically as $\sim \mu^2$). If one takes logarithmically flat priors, i.e. $p(\mu) \propto 1/\mu$, then eq. (1.14) would formally coincide with the procedure of multiplying the theoretical prior $p(s, m, M, A, B)$ by a factor $1/c$. This is reasonable: the usual naturalness criteria implicitly assume that for a given value of one parameter, say $\mu = \mu_0$, the prior probability is distributed around μ_0 [10, 12] with a width $\sim \mu_0$ [see the brief discussion in the paragraph after eq. (1.10)]. This is equivalent to assume that the value $\mu = \mu_0$ has a prior probability $\propto 1/\mu_0$. Actually this is the reason why, according to usual fine-tuning arguments, large soft parameters are more unlikely than small ones: for the former the region of the parameter space that produces the observed electroweak scale is much narrower than for the latter, *not* in absolute value, but *compared* to the size of the soft parameters in each case. Assuming flat priors there would be no reason to prefer soft parameters of the electroweak size instead of e.g. order $M_{\rm GUT}$. The fact that even for flat priors we still get a penalty factor μ/c_μ comes from the assumption of a prior flat in μ instead of μ^2, which is the quantity that appears in the cancellation [see e.g. eq. (1.9)].

We find very satisfactory that the usual parameter to quantify the degree of fine-tuning emerges from the Bayesian approach "spontaneously", not upon subjective assumptions, especially taking into account that there has been much discussion in the literature about its significance and suitability, see e.g. refs. [18, 19, 20, 10, 11, 12, 13]. Actually, one gets simply c_μ instead c, as defined in eq. (1.10). Of course there is nothing special with the $\mu-$parameter, except the fact that we have chosen to marginalize it using the experimental information about M_Z, which is the usual practice. Had we chosen to marginalize another parameter, say M, we would have got c_M, but of course at the end the results would be the same. A convenient way to view eq. (1.14) is to imagine that we start with an MSSM parameter space $\{s, m, M, A, B\}$ where μ has been eliminated using the experimental value of M_Z. Then the pdf appears as the likelihood associated to the experimental information (except $M_Z^{\rm exp}$) times an effective prior

$$p_{\rm eff}(s, m, M, A, B) \;=\; 2\,\frac{\mu_0}{M_Z}\,\frac{p(\mu_0)}{c_\mu}\,p(s, m, M, A, B)\;, \tag{1.15}$$

where for simplicity we have assumed that the prior in μ factorizes from the rest. This means that the initial prior gets multiplied by a factor $2\frac{\mu_0}{M_Z}\,\frac{p(\mu_0)}{c_\mu}$ that carries the fine-tuning penalty. In Fig. 1.1 we have plotted this factor in representative slices of the $\{s, m, M, A, B\}$ parameter space (using the two basic choices $p(\mu) \propto$ const., $p(\mu) \propto 1/\mu$) for some illustrative and physically relevant cases. In all of them large soft parameters get penalized (except partially for focus-point regions [21, 22]). There are no ad hoc assumptions for this result, it just comes out from the value of $M_Z^{\rm exp}$ and the marginalization of μ.

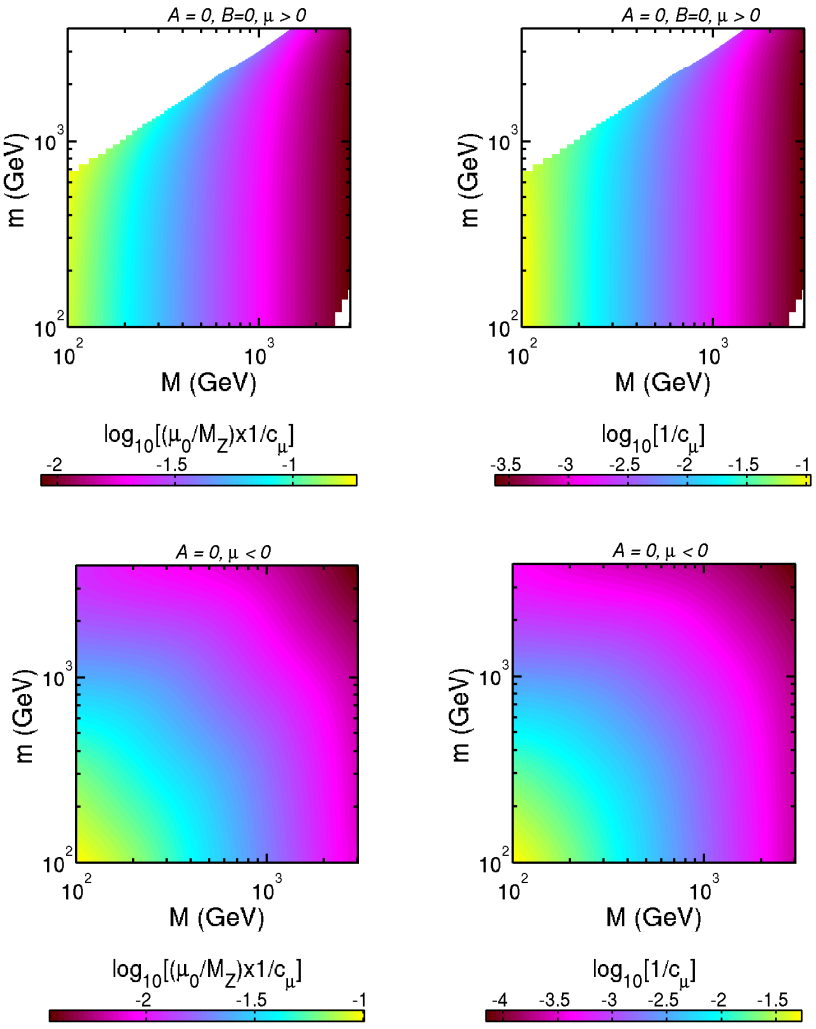

Figure 1.1: Values of the factor $\mu\, p(\mu)/(M_Z c_\mu)$ (in logarithmic units and up to a convenient proportionality constant) in the $\{m, M\}$ plane for $\mu > 0, A = 0,\ B = 0$ (upper plots), and for $\mu < 0,\ A = 0$ and the minimal SUGRA relation $B = A - m$ (lower plots), using the two basic initial priors, $p(\mu) \propto$ const. (left plots), $p(\mu) \propto 1/\mu$ (right plots). The plotted factor appears in the effective prior given in eq. (1.15).

1.3.1 Nuisance variables and the role of the Yukawa couplings

It is common in statistical problems that not all the parameters that define the system are of interest. In the problem at hand we are interested in determining the probability regions for the MSSM parameters that describe the new physics, i.e. $\{m, M, A, B, \mu\}$, but not (or not at the same level) in 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 (1.14) in the $\{s\}$ variables (for a review see ref. [23]). When the value of a nuisance parameter is in one-to-one correspondence to a high-quality experimental piece of information (included in $\mathcal{L}_{\rm rest}$), 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. In particular, the prior on such nuisance parameter becomes irrelevant. In the MSSM, nuisance parameters of this class are the gauge couplings, $\{g_3, g, g'\}$[4], which thus can be extracted from the analysis.

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 value of $\tan\beta \equiv v_2/v_1$, which is a derived quantity that 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 order to make these points more explicit, let us temporarily simplify the discussion approximating the experimental likelihood related to the fermion masses as

$$\mathcal{L}_{\text{fermion masses}} \;=\; \delta(m_t - m_t^{\rm exp})\;\delta(m_b - m_b^{\rm exp})\;.... \tag{1.16}$$

(which is a fair approximation). This is a factor of the global likelihood, $\mathcal{L}_{\rm rest}$. Likewise, let us approximate the theoretical values of the fermion masses as

$$m_t = \frac{1}{\sqrt{2}} y_t^{\rm low} v s_\beta, \quad m_b = \frac{1}{\sqrt{2}} y_b^{\rm low} v c_\beta, \qquad \text{etc.} \tag{1.17}$$

where $s_\beta \equiv \sin\beta$, $c_\beta \equiv \cos\beta$ and $y_i^{\rm low}$ are the low-energy Yukawa couplings. As it is well-known these expressions correspond to the running masses. The physical (pole) masses include a radiative correction that we have ignored here, but not in our full analysis. A further simplification is to assume $y_i^{\rm low} = R_i y_i$, where y_i are the high-energy Yukawa couplings (and thus the input parameters) and the renormalization-group factor R_i does not depend on y_i itself (this is not a good approximation for the top Yukawa coupling, but we will assume it momentarily for the sake of clarity). Now, the marginalization in the Yukawa couplings can be readily done, integrating the pdf given by eq. (1.14) in the y_i variables. Writing just the

[4]Strictly speaking, the initial theoretical inputs are the gauge couplings at high energy, which are related to the experimental (low-energy) ones by the renormalization-group running. This running depends on the other MSSM parameters through the position of thresholds associated with different particles. Hence, two viable MSSM models have slightly different values of the gauge couplings at high energy, and thus the theoretical prior on the couplings would play an (almost insignificant) role in the statistical comparison of the two models.

relevant terms we get

$$\int [dy_t\ dy_b \cdots]\ p(y, m, M, A, B|\ \text{data}) \ = \ \int [dy_t\ dy_b \cdots]\ p(y)\delta(m_t - m_t^{\rm exp})\ \delta(m_b - m_b^{\rm exp}) \cdots$$
$$\sim p(y)\ \left|\frac{dy_t}{dm_t}\right| \left|\frac{dy_b}{dm_b}\right| \cdots \ = \ p(y)\ s_\beta^{-1}\ c_\beta^{-1} \cdots \qquad (1.18)$$

where $p(y)$ denotes the prior in the Yukawa couplings (which we assume that factorizes from the other priors). Eq. (1.18) represents the footprint of the Yukawa couplings in the pdf. Note that the factors $s_\beta^{-1}\ c_\beta^{-1} \cdots$ arise from the change of variables $y_i \to m_i$, even if the likelihood is not approximated by deltas. There are as many such factors as quarks and leptons. This amounts to a dramatic modulation of the relative probability of MSSM regions with different $\tan\beta$ if one chooses a flat prior, $p(y) = \text{const}$. If, instead, one takes logarithmically flat priors, i.e. $p(y_i) \propto 1/y_i$, then the $s_\beta^{-1}\ c_\beta^{-1} \cdots$ factors get cancelled, so that the elimination of the Yukawa couplings does not leave a footprint in the probability density of the (non-nuisance) MSSM parameter space, $\{m, M, A, B, \mu\}$.

In previous Bayesian analyses of the MSSM the role of the Yukawa couplings was not considered to this extent. Essentially, their values were taken as needed to reproduce the experimental fermion masses, within uncertainties. As we have seen, this practice approximately corresponds to assuming logarithmically flat priors in the Yukawa couplings[5].

The above discussion is however oversimplified. As already mentioned, the marginalization in the top Yukawa coupling (and sometimes the bottom one) produces extra factors due to the dependence of R_t on y_t. Actually, since one is marginalizing simultaneously in the Yukawa couplings and the $\mu-$parameter one has to evaluate the full Jacobian of the transformation $\{\mu, y_t\} \to \{M_Z, m_t\}$, which introduces additional contributions. Furthermore, the picture gets more complicated due to the fact that, for a given choice of $\{m, M, A, B\}$, there may be several values of μ leading to the correct value of M_Z with different values of $\tan\beta$ and thus of the Yukawa couplings. This means that in the marginalization one has to sum over all these possibilities. This is technically annoying and reduces the clarity of the approach. These drawbacks can be eliminated by trading in the statistical analysis the initial $B-$parameter by the derived $\tan\beta$ parameter, as we discuss in the next section.

Let us finally mention that in the analysis of ref. [14] the fermion masses themselves, rather than the Yukawa couplings, were taken as SM-like variables. The advantage of such procedure is that these nuisance variables are in obvious one-to-one correspondence to the experimental data. Then the priors on the masses become almost irrelevant, and they can be integrated out, almost without leaving any footprint. However, this has two problems. First, the fermion masses are obviously derived quantities and should not be taken as initial input variables, even if this makes life easier. Second, such procedure introduces completely artificial factors, as it will become clear at the end of the next section.

[5]Actually, 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.

1.4 Efficient variables to scan the MSSM parameter space

In MSSM analyses it is normally very advantageous, both for theoretical and phenomenological reasons, to trade the initial $B-$parameter by the derived $\tan\beta$ parameter. On the phenomenological side, $\tan\beta$ is a parameter that appears explicitly in the predictions for many physical processes, such as cross sections, branching ratios, etc. (this is unlike B, that enters only in a very indirect way). Thus it is convenient to get the probability density of the MSSM parameter space as a function of $\tan\beta$. On the theoretical side, 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, as mentioned in the previous section.

Changing variables $B \to \tan\beta$ amounts to a factor $dB/d\tan\beta$ in the pdf. On the other hand, we have seen in sect. 1.3 that it is convenient to trade μ and y_t by M_Z and m_t, as this makes the marginalization of these variables easier and more transparent. Thus we should compute the whole Jacobian, J, of the transformation

$$\{\mu, y_t, B\} \;\to\; \{M_Z, m_t, t\}, \qquad t \equiv \tan\beta \;, \tag{1.19}$$

so that, in the new variables, the pdf reads

$$p(g_i, m_t, m, M, A, \tan\beta|\text{ data}) \;=\; \mathcal{L}_{\rm rest}\; J|_{\mu=\mu_0}\; p(g_i, y_t, m, M, A, B, \mu=\mu_0) \;. \tag{1.20}$$

Here we have made explicit the dependence on the gauge couplings, and the top Yukawa coupling and mass, but not on the other fermions'. In this equation we have already marginalized M_Z using the associated likelihood $\sim \delta(M_Z - M_Z^{\rm exp})$ (recall that μ_0 is the value of μ that reproduces the experimental M_Z.) The combination

$$p_{\rm eff}(g_i, m_t, m, M, A, \tan\beta) \;\equiv\; J|_{\mu=\mu_0}\; p(g_i, y_t, m, M, A, B, \mu=\mu_0) \tag{1.21}$$

can be viewed as the effective prior in the new, more convenient, variables to scan the MSSM. Note that, as discussed in subsect. 1.3.1, the gauge couplings are fairly irrelevant for the statistical analysis, so we will drop them in what follows. In order to work out J we need the dependence of the old variables on the new ones, which can be derived from the minimization equations of the scalar potential, $V(H_1, H_2)$, and from the expression of the top pole mass. For the numerical analysis we have used the SOFTSUSY code [8] which implements the full one-loop contributions and leading two-loop terms to the tadpoles for the electroweak symmetry breaking conditions with parameters running at two-loops. This essentially corresponds to the next-to-leading log approximation. However, in order to highlight the most relevant facts it is useful to write down the expressions arising from the

minimization of the tree-level potential with parameters running at one-loop (i.e. essentially the leading log approximation):

$$\mu^2_{\rm low} = \frac{m^2_{H_1} - m^2_{H_2} t^2}{t^2-1} - \frac{M_Z^2}{2} \tag{1.22}$$

$$B_{\rm low} = \frac{s_{2\beta}}{2\mu_{\rm low}}(m^2_{H_1} + m^2_{H_2} + 2\mu^2_{\rm low}) \tag{1.23}$$

$$y_{\rm low} = \frac{m_t}{v\ s_\beta}\ . \tag{1.24}$$

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 $m^2_{H_i}$ are also understood at low scale. For notational simplicity, we have dropped the subscript t from the Yukawa coupling. We are not making explicit the role of the bottom Yukawa coupling, which is treated in a similar foot to the top one. Note that all these low-energy quantities contain an implicit dependence on the top Yukawa coupling through the corresponding renormalization-group equations (RGEs). The effect of the one-loop corrections on the effective potential to the previous expressions is incorporated by correcting the soft masses $m^2_{H_i}$ with one-loop tadpole effects along the lines of ref. [24]. Similarly the pole top mass is given by the running top mass, appearing in eq. (1.24), plus a radiative correction $\Delta_{rad} m_t$. Eqs. (1.22–1.24), even when corrected with the mentioned radiative effects, have the structure

$$\mu = f(M_Z, y, t), \quad y = g(M_Z, m_t, t), \quad B = h(\mu, y, t)\ , \tag{1.25}$$

where we only make explicit the dependence on the variables involved in the change of variables (1.19). Note that y depends on M_Z since $v \propto M_Z$. Notice also that, unlike eqs. (1.22–1.24), eqs. (1.25) are defined in terms of the the high-energy parameters.

From eqs. (1.25) it is straightforward to evaluate the Jacobian J of the transformation (1.19), and thus the effective prior (1.21). J gets simply

$$J = \left| \begin{array}{ccc} \frac{\partial \mu}{\partial M_Z} & \frac{\partial \mu}{\partial t} & \frac{\partial \mu}{\partial m_t} \\ \frac{\partial B}{\partial M_Z} & \frac{\partial B}{\partial t} & \frac{\partial B}{\partial m_t} \\ \frac{\partial y}{\partial M_Z} & \frac{\partial y}{\partial t} & \frac{\partial y}{\partial m_t} \end{array} \right| = \frac{\partial f}{\partial M_Z}\ \frac{\partial g}{\partial m_t}\ \frac{\partial h}{\partial t}\ , \tag{1.26}$$

where the factor $\partial f/\partial M_Z$ carries essentially the fine-tuning penalization discussed in sect. 1.3

We can give an analytical and quite accurate expression of J by using the approximate equations (1.22–1.24), and expressing the low-energy values of μ, B, y in terms of the high-energy ones through the integrated 1-loop RGEs. Schematically,

$$\mu_{\rm low} = R_\mu(y)\mu, \quad B_{\rm low} = B + \Delta_{RG} B(y)\ , \tag{1.27}$$

where $R_\mu(y), \Delta_{RG}B(y)$ are definite functions of y (and other parameters, but not μ and B), see e.g. ref. [5] above. Similarly,

$$y_{\rm low} \simeq \frac{yE(Q_{\rm low})}{1+6yF(Q_{\rm low})} \ , \tag{1.28}$$

where Q is the renormalization scale, $F = \int_{Q_{\rm high}}^{Q_{\rm low}} E \ln Q$, and $E(Q)$ is a definite function that depends just on the gauge couplings [25]. Plugging (1.27) and (1.28) into eqs. (1.22–1.24) we get explicit expressions for the f, g, h functions. The relevant derivatives, to be plugged in (1.26), read

$$\frac{\partial f}{\partial M_Z} = -\frac{M_Z}{\mu}\frac{1}{2R_\mu^2} = -\frac{M_Z}{\mu_{\rm low}}\frac{1}{2R_\mu} \tag{1.29}$$

$$\frac{\partial h}{\partial t} = B_{\rm low}\ \frac{1-t^2}{t(1+t^2)} \tag{1.30}$$

$$\frac{\partial g}{\partial m_t} = \frac{E}{v\ s_\beta}\left(\frac{y}{y_{\rm low}}\right)^2 \ . \tag{1.31}$$

Let us comment briefly on these expressions. As mentioned above, eq. (1.29) is essentially the fine-tuning factor $2\mu/(M_Z c_\mu)$ obtained in sect. 1.3 [eq. (1.14)]. It penalizes large scales for μ. Eq. (1.30) counts the volume conversion from dB to dt and it is proportional to a soft mass just for dimensional reasons. Note that this factor penalizes low scales. This is easy to understand looking at eq. (1.23): for a given interval in $\tan\beta$, the larger the values of the soft masses and μ, the larger the corresponding interval in B is. So larger B is favoured. Note, however, that the size of the interval of B relative to the value of B itself (which is statistically meaningful) is essentially constant. Indeed, the B-factor in eq. (1.30) will be cancelled in the pdf if one uses logarithmic flat priors for the soft terms, $p(B) \propto 1/B$. This reasoning is similar to that after eq. (1.14). Finally, eq. (1.31) corresponds to eq. (1.18) of our preliminar discussion. In particular, the $1/s_\beta$ factor corresponds to the same factor in (1.18).

The Jacobian of the transformation (1.19) is given by the product of the three factors of eqs. (1.29–1.31),

$$J = \frac{1}{4}(g^2+g'^2)^{1/2}\left[\frac{E}{R_\mu^2}\right]\ \frac{B_{\rm low}}{\mu}\frac{t^2-1}{t(1+t^2)}\left(\frac{y}{y_{\rm low}}\right)^2 s_\beta^{-1} \ . \tag{1.32}$$

In the previous derivation we have considered just the top Yukawa coupling in the change of variables (1.19). Once the others fermions are taken into account, the Jacobian gets a s_β^{-1} factor for each $u-$type quark and a c_β^{-1} factor for each $d-$type quark and charged lepton, as discussed in sect. 1.3. Now, recall that the effective prior in the new variables is the product of J by the initial prior, as expressed in eqs. (1.20, 1.21); so taking a logarithmically flat prior for the Yukawa couplings (i.e. $p(y_i) \propto y_i^{-1}$) the $s_\beta^{-1}, c_\beta^{-1}$ factors get cancelled in the effective

prior and the pdf. For the top Yukawa coupling (and sometimes for the bottom one) this cancellation still leaves a residual dependence on $\tan\beta$ since $\left(\frac{y}{y_{\rm low}}\right)^2 s_\beta^{-1} \times \frac{1}{y} \propto \frac{y}{y_{\rm low}}$, which through (1.28) depends on y itself and thus on $\tan\beta$.

Therefore, the effective prior defined by eq. (1.21) takes the approximate form

$$p_{\rm eff}(m_t, m, M, A, \tan\beta) \;\; \propto \;\; \left[\frac{E}{R_\mu^2}\right] \; \frac{y}{y_{\rm low}} \frac{t^2-1}{t(1+t^2)} \frac{B_{\rm low}}{\mu_0} p(m, M, A, B, \mu = \mu_0) \; . \qquad (1.33)$$

Let us stress that its form stems just from the relation between the initial variables and the phenomenological ones, indicated in eq.(1.19), and it is not "subjective" at all. Besides, the prefactor in the r.h.s. of eq.(1.33) (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.(1.33) contains the above-discussed penalization of fine-tuned regions, something that may be not so obvious, but that will become clear in sect. 1.6. Finally, the form of the prefactor implies an effective penalization of large $\tan\beta$, reflecting the smaller statistical weight of this possibility. Actually, the implicit fine-tuning associated to a large $\tan\beta$ was already noted in ref. [26, 27], where it was estimated to be of order $1/\tan\beta$, in agreement with eq.(1.33). This is logical. From eq.(1.23) we see that

$$\frac{1}{\tan\beta} = \frac{\mu_{\rm low} B_{\rm low}}{m_{H_1}^2 + m_{H_2}^2 + 2\mu_{\rm low}^2} \qquad (1.34)$$

The denominator of this expression has the size of the typical soft terms (which we will call M_S). Therefore a large $\tan\beta$ requires abnormally small $\mu_{\rm low} B_{\rm low}$. As a matter of fact, μ cannot be very small, otherwise the mass of the lightest chargino would be below the experimental limit. Therefore $\tan\beta$ requires very small $B_{\rm low}$. But this cannot be naturally arranged since the radiative contributions to B (i.e. its RG evolution from high to low scale) are sizeable (of order M_S) [5]. Thus small $B_{\rm low}$ requires a tuning between its initial (high scale) value and the radiative corrections.

In this section we have argued so far that the sensible initial choice of independent parameters of the MSSM is $\{g_i, y_t, m, M, A, B, \mu\}$, while for practical reasons it is most convenient to work with the set $\{g_i, m_t, m, M, A, \tan\beta, M_Z\}$ (and signμ). M_Z is eliminated from the analysis using its extremely sharp likelihood. The effective prior in the new variables is then given by eqs. (1.21, 1.26), for which we gave explicit approximate expressions in eqs. (1.32, 1.33).

It is interesting to wonder what would have been the result, if one had insisted in taking directly m_t as an initial (nuisance) variable, so that the transformation (1.19) would have just involved $\{\mu, B\} \;\; \to \;\; \{M_Z, t\}$, as has been done e.g. in ref. [15]. This approach is explained in Appendix A.

1.5 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. [28]). 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 $\mu-$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 [29]) to the SUSY breaking scale. Otherwise, one has to face the so-called $\mu-$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.(1.22)). 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

$$\begin{aligned} -qM_S \leq \quad & B \quad \leq qM_S \\ -qM_S \leq \quad & A \quad \leq qM_S \\ 0 \leq \quad & m \quad \leq qM_S \\ 0 \leq \quad & M \quad \leq qM_S \\ 0 \leq \quad & \mu \quad \leq qM_S \end{aligned} \tag{1.35}$$

where q is an $\mathcal{O}(1)$ factor. We have considered here the branch of positive μ. For the negative one we simply replace $\mu \to -\mu$. 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

$$M_S^0 \leq M_S \leq M_X \,, \quad M_S^0 \sim 10 \text{ GeV} \tag{1.36}$$

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 (1.35), i.e.

$$p(m) = p(M) = p(\mu) = \frac{1}{M_S} \,, \quad p(A) = p(B) = \frac{1}{2M_S} \tag{1.37}$$

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. 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,

$$p(M_S) = N_{M_S} \frac{1}{M_S} \,, \tag{1.38}$$

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 $\{m, M, A, B, \mu\}$ parameters[6]:

$$\begin{aligned} p(m, M, A, B, \mu) &= \frac{N_{M_S}}{4} \int_{\max\{m,M,|A|,|B|,\mu,M_S^0\}}^{M_X} \frac{1}{M_S^6} \, dM_S \\ &= \frac{N_{M_S}}{20} \left[\frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^5} - \frac{1}{M_X^5} \right] \\ &\simeq \frac{N_{M_S}}{20} \frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^5} \end{aligned} \tag{1.39}$$

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.

[6]This procedure is a "hierarchical Bayesian technique", first used in ref. [15], but using complicated functions that were not possible to integrate analytically.

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 sect 1.3 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.(1.39) 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.(1.39) 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 (1.39) is the one to be plugged in eq.(1.21) (or in the approximated expression (1.33)) to get the effective prior in the scan parameters.

It is funny to compare the prior of eq.(1.39) 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 (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 $\mathcal{O}$(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 (1.39), reflecting the way it has been constructed.

To finish this discussion, let us note that the prior (1.39) 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

$$\mathcal{P}(M) \propto \frac{1}{\max\{M, M_S^0\}} \tag{1.40}$$

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. (1.35), (1.36), and the flat priors inside those ranges, eq. (1.37). We just replace the logarithmically flat prior in M_S, eq. (1.38), by a flat prior

$$p(M_S) = N_{M_S} \tag{1.41}$$

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 (1.39) becomes now

$$p(m, M, A, B, \mu) \propto \frac{1}{[\max\{m, M, |A|, |B|, \mu, M_S^0\}]^4} \tag{1.42}$$

where, once more (and for the same reasons) we have neglected a $1/M_X^4$ contribution. The difference with eq.(1.39) is that now we have one power less in the denominator. Again, the prior (1.42) is the one to be plugged back into eq.(1.21) 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 (1.40) becomes

$$\mathcal{P}(M) \sim \ln \frac{M_X}{\max\{M, M_S^0\}} \ , \tag{1.43}$$

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

1.6 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.(1.39, 1.40), 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 Gev $\lesssim M \lesssim$ 2 TeV versus the chance that it lies at a higher scale, 2 TeV $\lesssim M \lesssim M_X$. Using eq.(1.40), it is clear that this relative probability is

$$\frac{\mathcal{P}(100 \text{ GeV} \leq M \leq 2 \text{ TeV})}{\mathcal{P}(2\text{TeV} \leq M \leq M_X)} \simeq \frac{1}{12} \tag{1.44}$$

(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.(1.43) we see that in that case

$$\frac{\mathcal{P}(100\ \text{GeV} \leq M \leq 2\ \text{TeV})}{\mathcal{P}(2\text{TeV} \leq M \leq M_X)} \simeq 3 \times 10^{-12}\ . \tag{1.45}$$

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 sect. 1.3 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.(1.21) or by the approximate expression (1.33). 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

$$\mathcal{P}(M_1, \tan\beta) \sim \int_{R_1} dm\ dA\ p_{\text{eff}}(m_t, m, M, A, \tan\beta) \tag{1.46}$$

where $p_{\text{eff}}(m_t, m, M, A, \tan\beta)$ is given by eq.(1.21) or eq.(1.33). 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.(1.46). 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,

$$\begin{aligned}\mathcal{P}(M_2,\tan\beta) &\sim \int_{R_2} dm_2\ dA_2\ p_{\rm eff}(m_2,M_2,A_2,\mu_2,\tan\beta)\\ &= \int_{R_1} dm_1\ dA_1\ \left(\frac{M_1}{M_2}\right)^3\ p_{\rm eff}(m_1,M_1,A_1,\mu_1,\tan\beta)\\ &= \left(\frac{M_1}{M_2}\right)^3\ \mathcal{P}(M_1,\tan\beta)\ ,\end{aligned} \tag{1.47}$$

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

$$p_{\rm eff}(m_2,M_2,A_2,\mu_2,\tan\beta) \ = \ \left(\frac{M_1}{M_2}\right)^5\ p_{\rm eff}(m_1,M_1,A_1,\mu_1,\tan\beta)\ . \tag{1.48}$$

This can be easily noticed from the approximate expression of $p_{\rm eff}$ in eq.(1.33). This relation is exact, essentially, up to small RG-effects in the $\frac{B_{\rm low}}{\mu_Z}$ factor involved in (1.33).

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

$$\frac{\mathcal{P}(M_2,\tan\beta)}{\mathcal{P}(M_1,\tan\beta)} \ \sim \ \frac{M_1^3}{M_2^3}\ . \tag{1.49}$$

In other words,

$$\mathcal{P}(M,\tan\beta) \ \sim \ \frac{1}{M^3} \tag{1.50}$$

This should be compared with the $\mathcal{P}(M) \ \sim \ \frac{1}{M}$ behaviour obtained in eq.(1.40), 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

$$\frac{\mathcal{P}(100\ {\rm GeV}\le M\le 2\ {\rm TeV})}{\mathcal{P}(2{\rm TeV}\le M\le M_X)} \simeq \left(\frac{2\ {\rm TeV}}{100\ {\rm GeV}}\right)^2 \sim 10^2 \tag{1.51}$$

to be compared with eq.(1.44) 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 $p_{\rm eff}(m, M, A, \mu, \tan\beta)$] has one power of mass less in the denominator. Therefore at the end of the day we arrive at

$$\mathcal{P}(M, \tan\beta) \;\sim\; \frac{1}{M^2} \tag{1.52}$$

to be compared with the almost flat behaviour before including the experimental EW information, see eq.(1.42). Consequently, the relative probability of the low-energy and high-energy regions of the parameter space, for a flat prior, becomes

$$\frac{\mathcal{P}(100\ \text{GeV} \leq M \leq 2\ \text{TeV})}{\mathcal{P}(2\text{TeV} \leq M \leq M_X)} \simeq \frac{2\ \text{TeV}}{100\ \text{GeV}} \simeq 10 \tag{1.53}$$

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 [30] algorithm as implemented in the `SuperBayeS` code [31]. It is based on the framework of Nested Sampling, recently invented by Skilling [32, 33]. 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 [34]. The upper one comes from imposing Yukawa couplings in the perturbative regime [35, 36, 37]. The precise value $\tan\beta < 62$ has been chosen to allow an strict comparison with previous analyses in refs. [38, 15] However, the precise value of this upper bound turns out to be irrelevant, as the region of very large $\tan\beta$ is strongly suppressed (see the discussion after eq. (1.33) and ref. [39]).

In Fig. 1.2 the red line shows the prior in M (upper panels) and m (lower panels), when the other parameters are marginalized, using logarithmic (left panels) or flat (right panels) initial prior for the scale of SUSY breaking in the observable sector (M_S). The blue bar

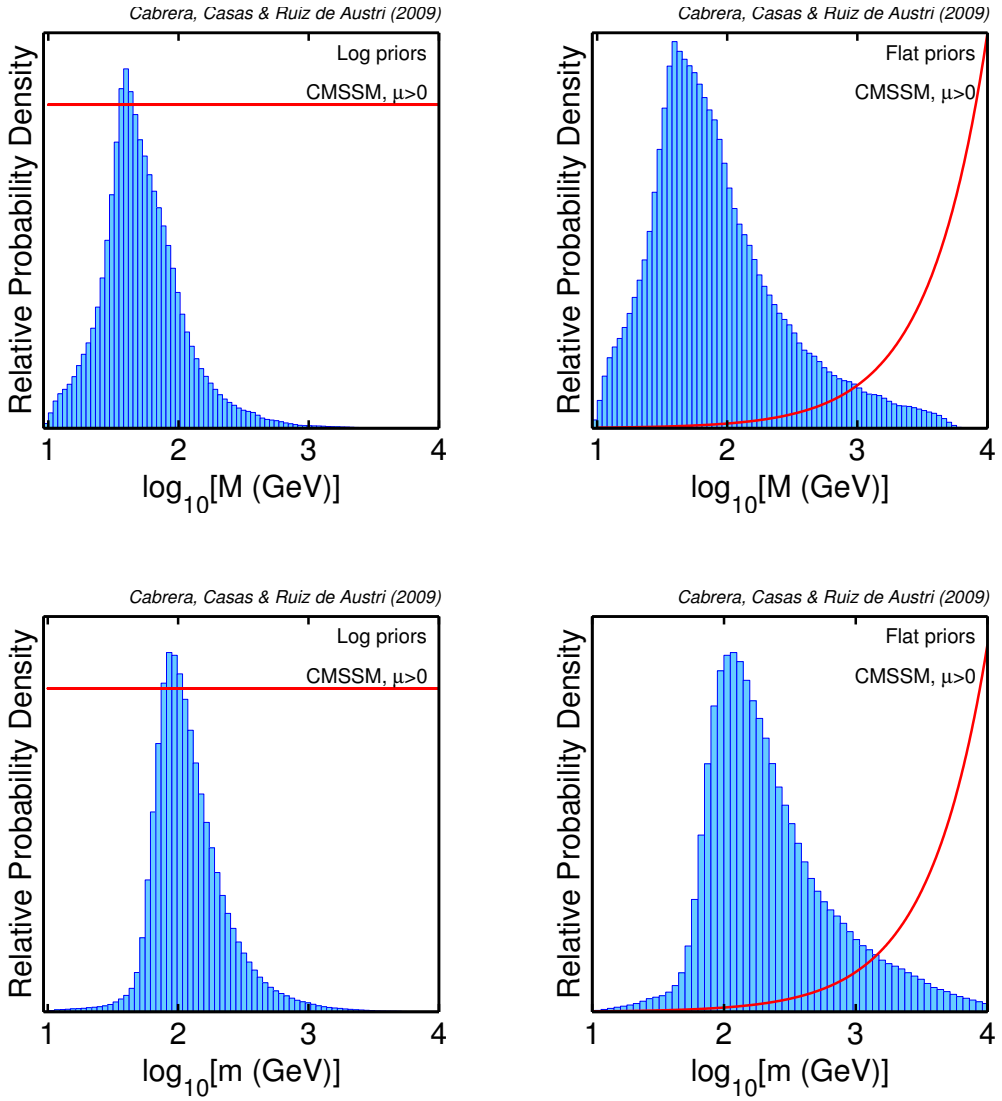

Figure 1.2: 1D marginalized posterior probability distribution of the M and m parameters (upper and lower panels respectively) for logarithmic (left panels) and flat (right panels) priors in the $\mu > 0$ case, for a scan including the information about the EW breaking ($M_Z^{\rm exp}$). The red lines represent the marginalized prior. All given in arbitrary units.

distributions show the pdf once the EW breaking is incorporated in the analysis, i.e. the effective prior in the scan variables, $p_{\rm eff}$, see eq.(1.21) and the approximate form (1.33). The logarithmic scale in the horizontal axes allows to see that most of the probability, which initially lies in the high-energy region (M, 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 just 1 fb^{-1} [40, 41, 42]). Quantitatively, the results are in good agreement with the previous discussion. Although the distribution of probability above 1 TeV is almost invisible in the plots (especially for log priors), it is

actually different from zero and follows from the approximate law of eq. (1.50), (1.52). Notice also that, at this stage, all the points are equally "best-fit points" (even at extremely large M, m), since they are equally in reproducing M_Z, the only experimental information so far considered.

Besides making the high-energy parameter space quite irrelevant, the EW breaking has another dramatic effect, which is visible in Fig.1.2. 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 $M_Z^{\rm exp}$, 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 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 [43]).

Chapter 2

CMSSM forecast for the LHC

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 experimental constraints. For previous work on this subject see refs. [44, 14, 38, 15, 45, 46, 47, 48, 49, 50].

In this chapter we use the Bayesian set up described in chapter 1 to map the probability density of the CMSSM parameter space. We include all the important experimental constrains into the analysis, for positive and negative μ-parameter respectively. 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. We also make a comparison between the positive- and negative-μ cases. Finally, we compare our results with those of previous literature, including frequentist approaches, which are an alternative and complementary procedure to Bayesian ones.

2.1 Experimental Constraints

In this section we will perform a scan of the CMSSM parameter space using the Bayesian framework described in the previous chapter. We incorporate all the relevant experimental information to the likelihood piece of the probability distribution (all but $M_Z^{\rm exp}$, 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 [44, 38], 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

$$\left\{M_t,\ m_b(m_b)^{\overline{MS}},\ \alpha_{\rm em}(M_Z)^{\overline{MS}},\ \alpha_s(M_Z)^{\overline{MS}}\right\}, \tag{2.1}$$

where M_t is the pole top quark mass, while the other three parameters (the bottom mass, the

SM (nuisance) parameter	Mean value μ	Uncertainty σ (exper.)	Ref.
M_t	172.6 GeV	1.4 GeV	[51]
$m_b(m_b)^{\overline{MS}}$	4.20 GeV	0.07 GeV	[52]
$\alpha_s(M_Z)^{\overline{MS}}$	0.1176	0.002	[52]
$1/\alpha_{\rm em}(M_Z)^{\overline{MS}}$	127.955	0.03	[53]

Table 2.1: Experimental mean μ and standard deviation σ adopted for the likelihood function for SM (nuisance) parameters, assumed to be described by a Gaussian distribution.

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 2.1.

On the other hand, there are the experimental values of accelerator and cosmological observables, which are listed in Table 2.2. 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 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.

2.1.1 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 2.2 (all the entries except those concerning a_μ and dark matter constraints).

To calculate the MSSM spectrum we use `SoftSusy` [8], where SUSY masses are computed at full one-loop level and the Higgs sector includes two-loop leading corrections [54]. 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. 2.1.3 below. In our treatment of the radiative corrections to the electroweak observables M_W and $\sin^2\theta_{\rm eff}$ we include full two-loop and known higher order SM corrections as computed in ref. [55, 56], as well as gluonic two-loop MSSM corrections obtained in [57].

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) [58, 59]. 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 $\mathcal{O}$(TeV) scales to avoid the hierarchy problem [60, 61]. 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 $B \to X_s\gamma$ decay (the most important one) has been computed with the numerical code `SusyBSG` [62] using the full NLO QCD contributions, including the two-loop calculation of the gluino contributions presented in [63] and the results of [64] for the remaining non-QCD $\tan\beta$-enhanced contributions. The supersymmetric contributions to $b \to s\gamma$ grow with decreasing masses of the supersymmetric particles *and* with increasing $\tan\beta$. For $\mu > 0$ they have the "wrong sign", so larger supersymmetric masses are preferred. However the SUSY contribution is never dramatic for masses around 1 TeV or larger. For the determination of ΔM_{B_s} we use expressions from ref. [65] which include dominant large $tan\beta$-enhanced beyond-LO SUSY contributions from Higgs penguin diagrams. 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 [66] and references therein). Both codes have been integrated into `SuperBayes`.

Experimental bounds used in the analysis are indicated in the second part of Table 2.2. 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. [38].

For the quantities for which positive measurements have been made (as listed in the upper part of Table 2.2), we assume a Gaussian likelihood function with a variance given by the sum of the theoretical and experimental variances, as motivated by eq. (3.3) in ref. [38]. For the observables for which only lower or upper limits are available (as listed in the bottom part of Table 2.2) we use a smoothed-out version of the likelihood function that accounts for the theoretical error in the computation of the observable, see eq. (3.5) and fig. 1.2 in ref. [38]. In particular, in applying a lower mass bound from LEP-II on the Higgs boson h we take into account its dependence on its coupling to the Z boson pairs ζ_h^2, as described in detail in ref. [79]. When $\zeta_h^2 \simeq 1$, the LEP-II lower bound of 114.4 GeV (95% CL) [78] applies. For arbitrary values of ζ_h, we apply the LEP-II 95% CL bounds on m_h, which we translate into the corresponding 95% CL bound in the (m_h, ζ_h^2) plane. We then add a conservative theoretical uncertainty $\tau(m_h) = 3$ GeV, following eq. (3.5) in ref. [38]. The best fit is then defined as the maximum value of the joint likelihood function.

Fig. 2.1 (upper panels) show the pdf for the gaugino mass parameter, M, once all this experimental information is incorporated. Again, the left (right) panels correspond 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. 1.2). 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,

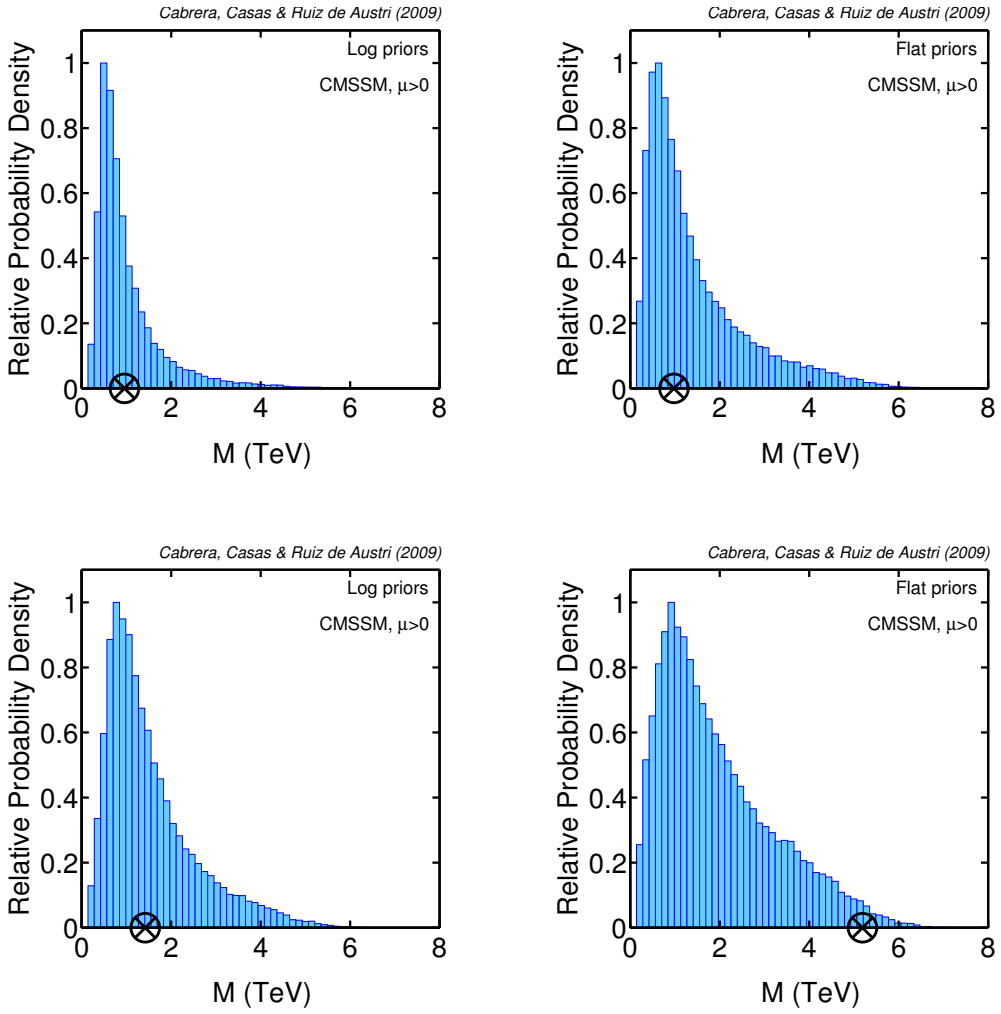

Figure 2.1: Upper panels show the 1D marginalized posterior probability distribution of the M parameter for logarithmic (left panel) and flat (right panel) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking ($M_Z^{\rm exp}$), collider limits on Higgs and 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 corresponds to the best-fit point, defined as the one with highest likelihood.

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.

It is possible to be more quantitative by considering the dominant 1-loop correction [80] to the theoretical upper bound on m_h in the MSSM:

$$m_h^2 \leq M_Z^2 \cos^2 2\beta + \frac{3m_t^4}{2\pi^2 v^2} \log \frac{M_{\tilde{t}}^2}{m_t^2} + ... \tag{2.2}$$

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, $m_h^{\rm min}$, one needs

$$M_{\tilde{t}} \;\gtrsim\; e^{-2.1\cos^2 2\beta} e^{\left(m_h^{\rm min}/62\ {\rm GeV}\right)^2} m_t \ . \tag{2.3}$$

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

$$M_{\tilde{t}} \;\to\; M_{\tilde{t}}\, e^{\Delta m_h^2/(62\ {\rm GeV})^2} \ , \tag{2.4}$$

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. 2.1 (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[1].

Fig. 2.2 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. 2.1). On the other hand, large values of $\tan\beta$ are penalized, mainly due to the Jacobian factor in the probability distribution, see eq.(1.33) and the subsequent discussion. It is worth to remark that this penalization of $\tan\beta$ contrasts with other Bayesian analyses, where the prior for $\tan\beta$ was taken as flat. Here it arises from the above-mentioned Jacobian factor and therefore has nothing to do with a particular choice of priors. Fig. 2.3 shows the probability distribution for the Higgs mass. One can see that there are a significant number of points which evade the LEP-II 114.4 GeV lower bound for the SM Higgs. This reflects the fact that we have employed the full likelihood function in the (m_h, ζ_h^2) plane as described above and which allows points with low Higgs masses where $\zeta_h^2 = sin^2(\beta - \alpha) \ll 1$. The corresponding Bayesian credibility intervals, representing the 68% and 95% of the total probability, are given in Table 2.3. 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. Notice also the little discrepancy among the mean value of the posterior pdf and the best fit.

Fig. 2.4 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. 2.2–2.4 are shown all 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.[42]. They arise from a study of events with $\rm N_{jets} \geq 2$ and an

[1]Certainly, it is well-known that a Higgs above 125 GeV is not easy to arrange in the MSSM, and that is at the origin of the difficulties. What the present analysis shows, in a more direct way, is how improbable is to arrange a large m_h (see also Fig. 2.3 below) and the implications for the discovery of SUSY at the LHC.

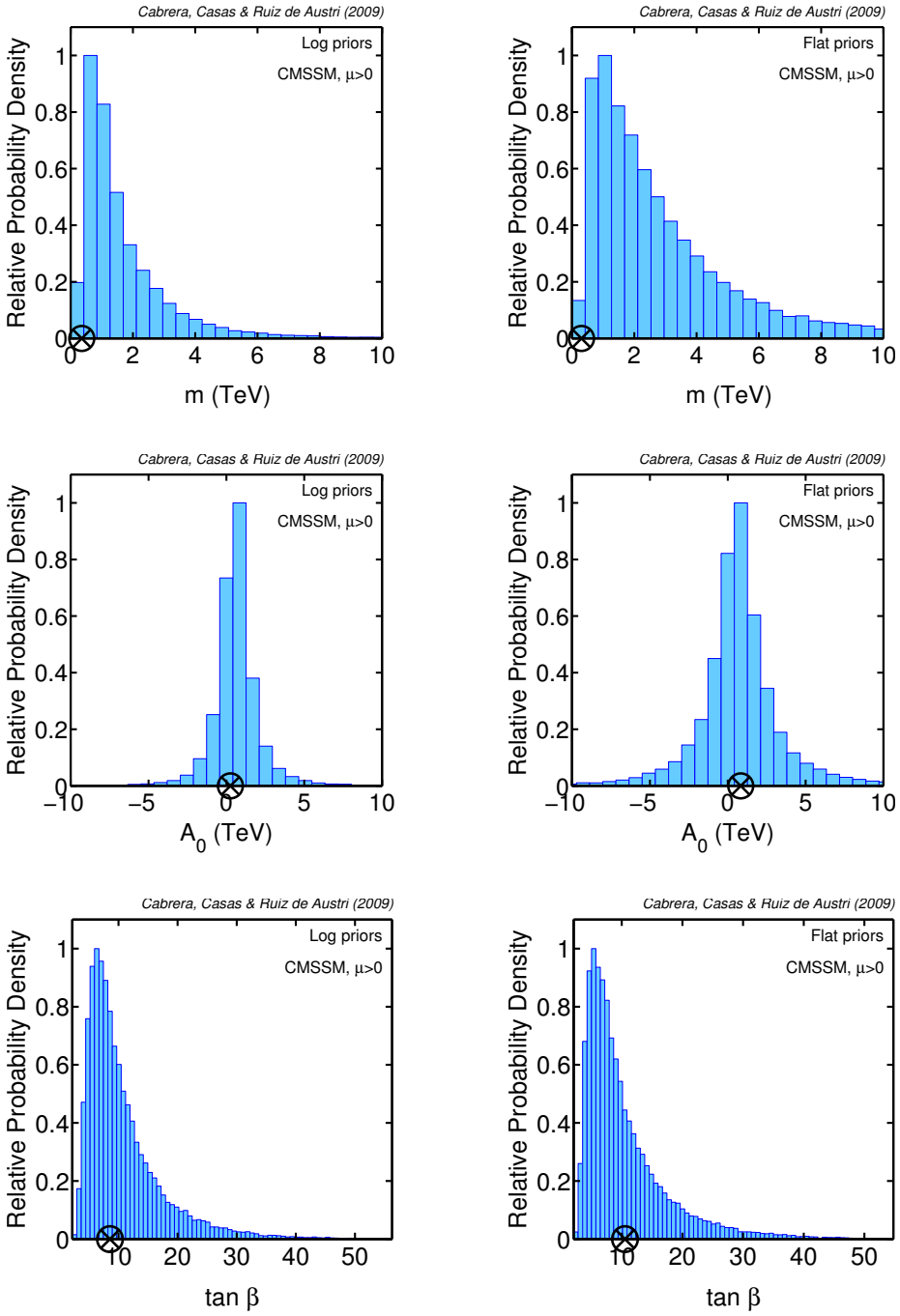

Figure 2.2: 1D marginalized posterior probability distribution of the CMSSM parameters for logarithmic (left panels) and flat (right panels) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking ($M_Z^{\rm exp}$), collider limits on Higgs and superpartner masses, and EW and B(D)-physics observables. The cross corresponds to the best-fit point.

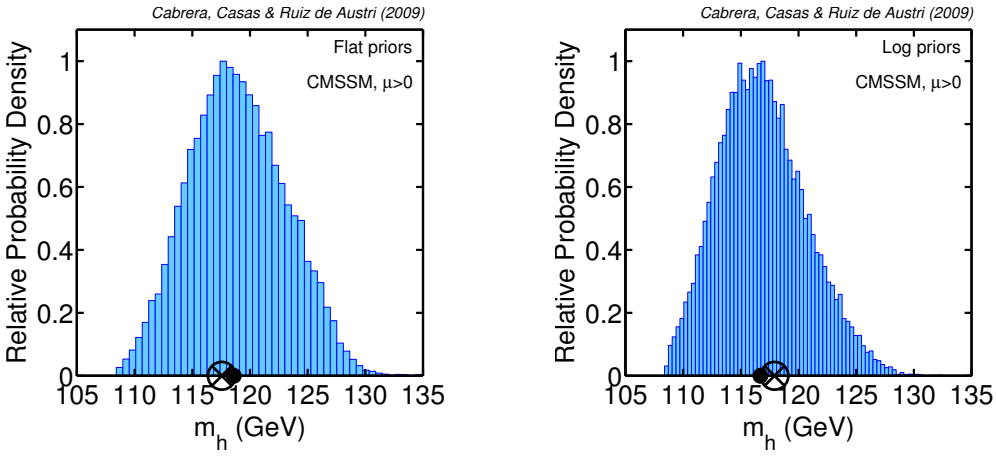

Figure 2.3: As Fig. 2.2, for the Higgs mass. The small filled circle represents the mean value of the posterior pdf and the cross corresponds to the best-fit point.

optimization of the cuts on $E_T^{\rm missing}$. (For a more detailed explanation of the procedure used see [29]). Strictly speaking, the lines 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 [40]). 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. 2.1), thus jeopardizing the discovery of supersymmetry.

2.1.2 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 $\delta^{\rm SM}_{\rm had} a_\mu$.

This contribution can be expressed in terms of the total hadronic cross section $e^+e^- \to$ 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 $\delta a_\mu = a_\mu^{\rm exp} - a_\mu^{\rm SM} = 29.5 \pm 8.8 \times 10^{-10}$. This has been often claimed as a signal of new 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

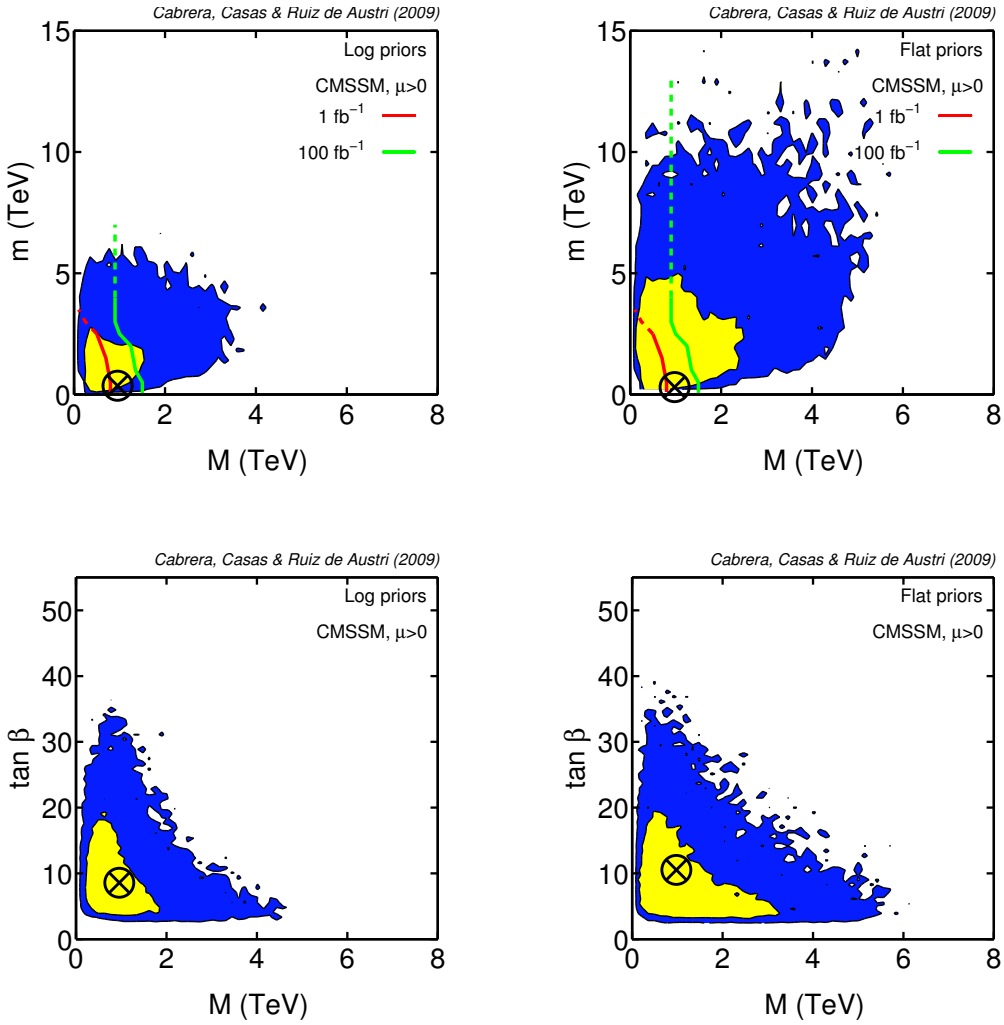

Figure 2.4: 2D marginalized posterior probability distribution for logarithmic (left panels) and flat (right panels) priors in the $\mu > 0$ case for a scan including SM nuisance parameters constraints, EW breaking ($M_Z^{\rm exp}$), collider limits on Higgs and superpartner masses, and EW and B(D)-physics 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 corresponds to the best-fit point.

a large enough contribution, $\delta^{\rm MSSM} a_\mu$, 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.1.6.

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 $e^+e^- \to$ hadrons cross section shows some inconsistencies between different groups of experimental data (see [81] for a recent account). This is especially notorious if one considers the hadronic $\tau-$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 $a_\mu^{\rm SM}$, 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 $\delta_{\rm had}^{\rm SM} a_\mu$ based on e^+e^- data. In the second, we use the one based on τ-data. We compute $\delta_{\rm had}^{\rm SM} a_\mu$ 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 [82]. The effective two-loop effect due to a shift in the muon Yukawa coupling proportional to $\tan^2\beta$ has been added as well [83].

Using $\delta_{\rm had}^{\rm SM} a_\mu$ 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 $b \to s\ \gamma$, start to be too large. This tension has been pointed out in ref.[4], and we would like to illustrate it here presenting some representative plots. Fig. 2.5 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 2.1.1 (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 $\delta^{\rm MSSM} a_\mu$ grows with decreasing masses and increasing $\tan\beta$). When both $b \to s, \gamma$ 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 $b \to s, \gamma$ and a_μ can also be noticed by looking at Fig. 2.6, where the left and right panels show the pdfs of $BR(B \to X_s\gamma)$ and $\delta^{\rm MSSM} a_\mu$ respectively, with the same code for the lines as in Fig. 2.5. Besides, we have include a gaussian in each panel (solid black line), proportional to the likelihood, and thus centered at the experimental value with the experimental uncertainty. Comparing the position of the bulk of the probability distribution with the likelihood, it is clear that the most favourable cases are not really satisfactory reproducing the two measurements simultaneously, even though we have not attempted to quantify this tension in a rigorous way.

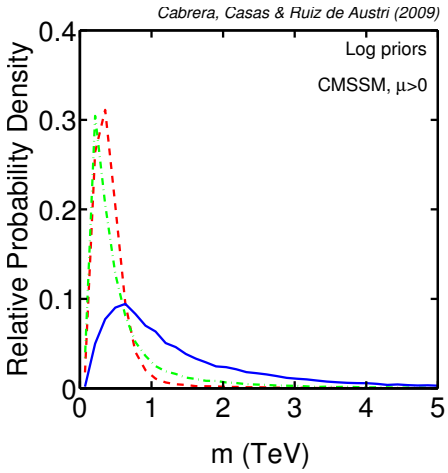

Figure 2.5: Non-normalized 1D marginalized posterior probability distribution of the m parameter for logarithmic prior and $\mu > 0$ including SM nuisance parameters constraints, EW breaking ($M_Z^{\rm exp}$), EW observables, collider limits on Higgs and superpartner masses: + B(D)-physics observables (blue solid line); + a_μ (red dashed line); + B(D)-physics and a_μ (green dashed-dotted line).

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 $a_\mu^{\rm exp}$ would be dramatic and could not be reconciled: m_h ($a_\mu^{\rm exp}$) would require too large (small) soft masses, see the discussion in subsect. 2.1.1.

Fig. 2.7 shows the probability distribution in the $\{M, m\}$ and $\{\tan\beta, M\}$ planes, as in Fig. 2.4, once the a_μ constraint (based on e^+e^- data) is included. Comparison with Fig. 2.4 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 $\delta_{\rm had}^{\rm SM} a_\mu$ from τ data

In this case, there is no big discrepancy between $a_\mu^{\rm SM}$ and $a_\mu^{\rm exp}$, so $\delta^{\rm MSSM} a_\mu$ 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. 2.1.1 (Figs. 2.1–2.4).

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).

2.1.3 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

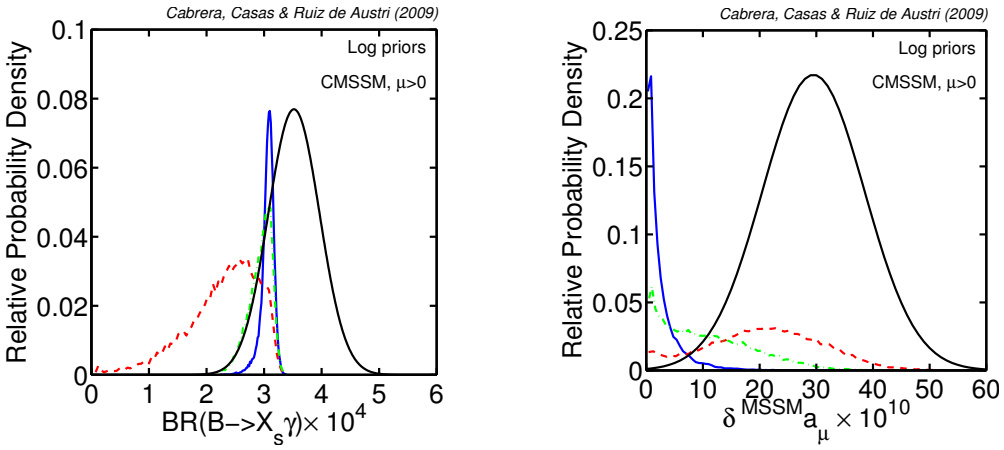

Figure 2.6: Non-normalized 1D marginalized posterior probability distribution for $BR(B \to X_s\gamma)$ (left panel) and for $\delta^{\rm MSSM}a_\mu$ (right panel).The code for the lines is as in Fig. 2.5. Besides, the black (solid) gaussians represent the experimental likelihood.

of $\Omega_{\rm DM}$). 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 [84]). 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 $\Omega_{\rm DM}$ 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 $M > m$.

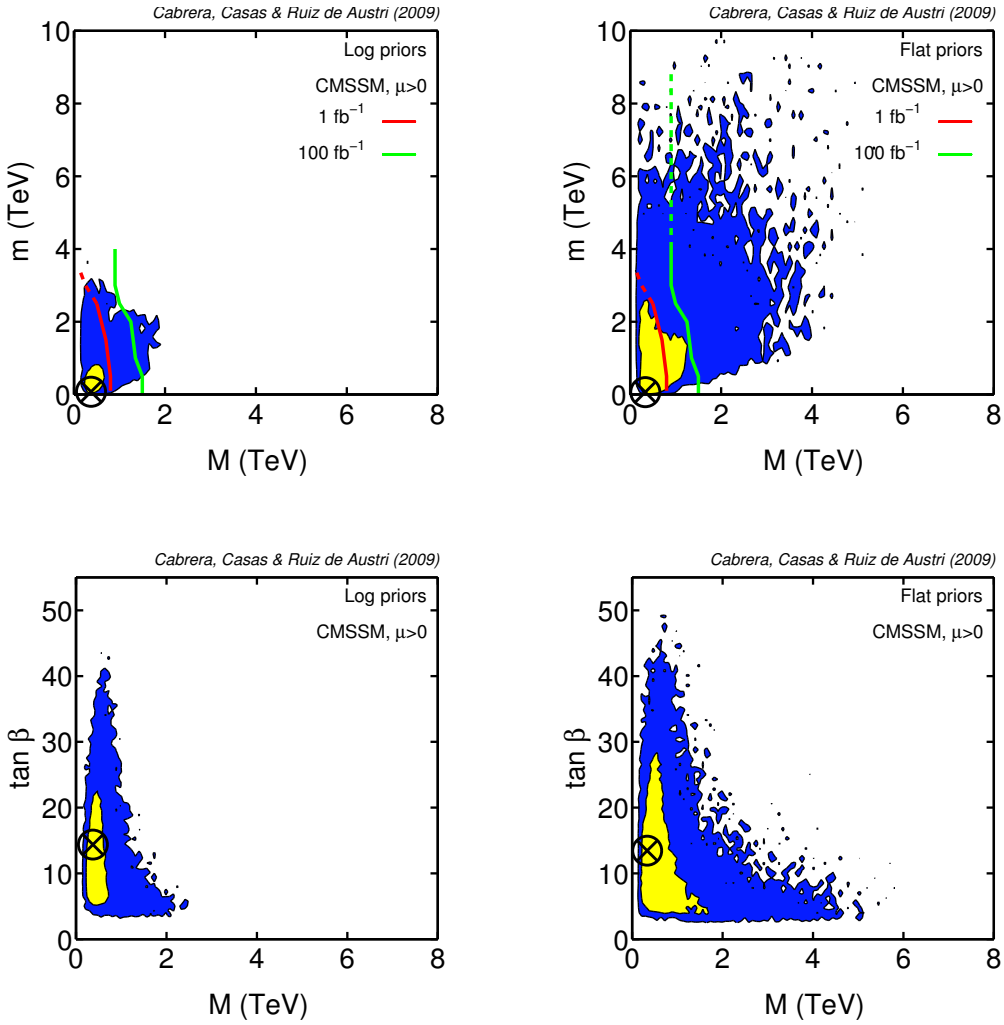

Figure 2.7: As in Fig. 2.4 but with the additional constraint from a_μ, based on e^+e^- data.

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 [85, 86] 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. 2.8 shows the resulting probability distribution in the $\{M, m\}$ and $\{M, \tan\beta\}$ 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. 2.8. On the other hand, the $\{M, \tan\beta\}$ plots show two big preferred regions. The largest one ocurrs at $M < 1$ TeV and contains (mixed) the

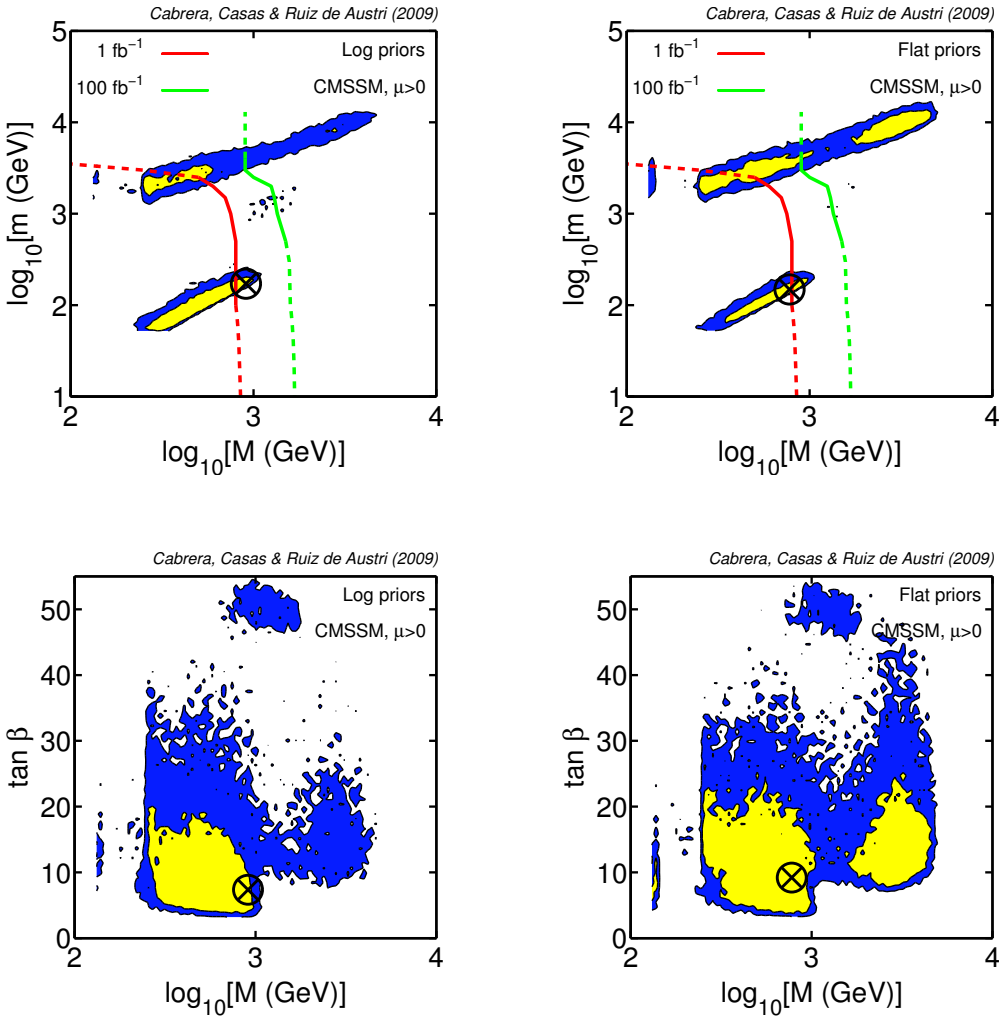

Figure 2.8: As in Fig. 2.4 but with an additional constraint from the WMAP CDM abundance.

Co-annihilation, Bulk and part of the Focus Point regions. The second one occurs at $M > 1$ TeV and corresponds to the part of the Focus Point region that needs moderate to large values of $\tan\beta$. Besides, the very small island around $M = 200$ GeV (also visible in the $\{M, m\}$ plots) corresponds also to the Focus Point region. Finally, the Higgs funnel region, which becomes significant for very large values of $\tan\beta$, is located around $\tan\beta = 50$. Let us remark that, since some of the previous regions require large $\tan\beta$, the latter 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. 2.4 and 2.8. 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 [87, 88]. Alternatively, the LSP could be unstable assuming tiny violations of R-parity, see e.g. [89]. 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. 2.9 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 $a_\mu^{\rm exp}$) 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. 2.8.

2.2 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* $\delta_{\rm had}^{\rm SM} a_\mu$ is evaluated using $e^+e^- \to$ had data). On the other hand, the contributions to $b \to s, \gamma$ 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.

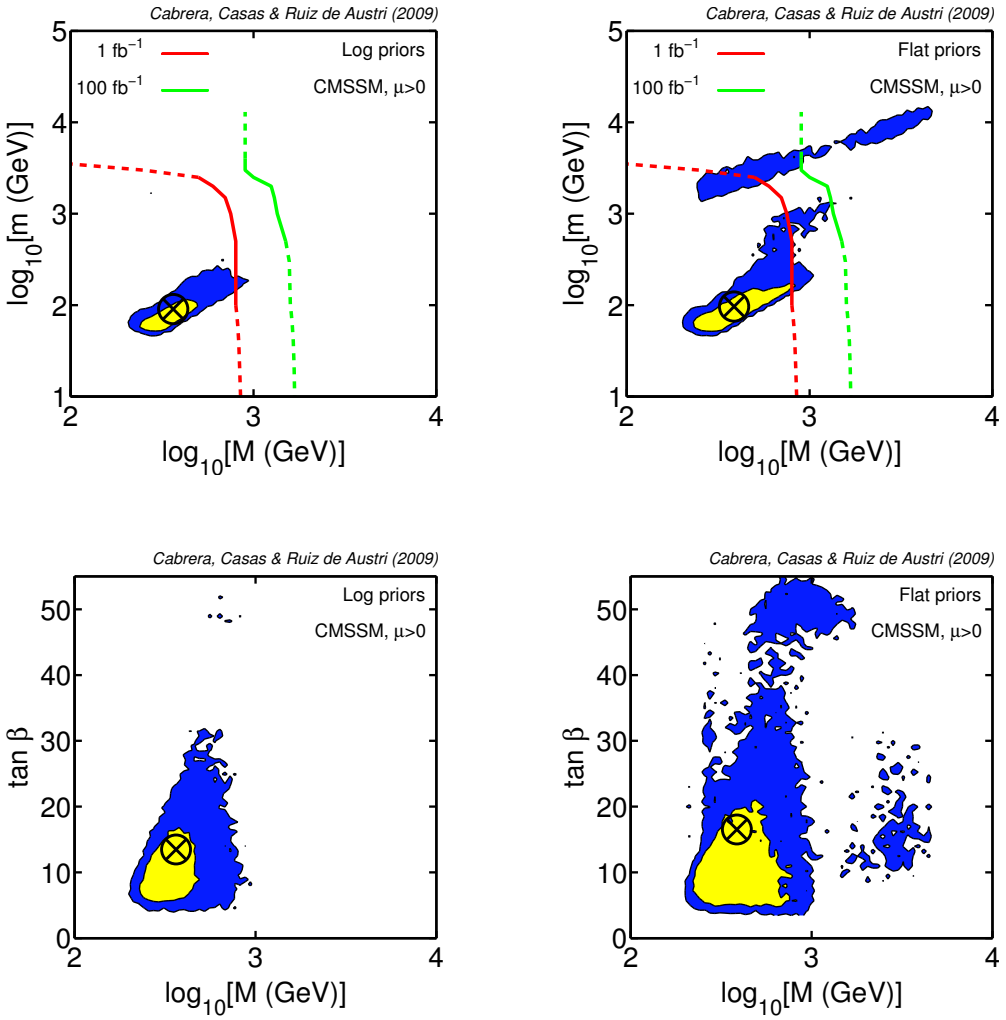

Figure 2.9: As in Fig. 2.4 but with an additional constraint from the WMAP CDM abundance.

2.2.1 Results

The results for $\mu < 0$ are summarized in Figs. 2.10, 2.11, 2.12 and 2.13, which are as previous Figs. 2.4, 2.7, 2.8 and 2.9, 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. Because of the above-mentioned $b \to s, \gamma$ 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.

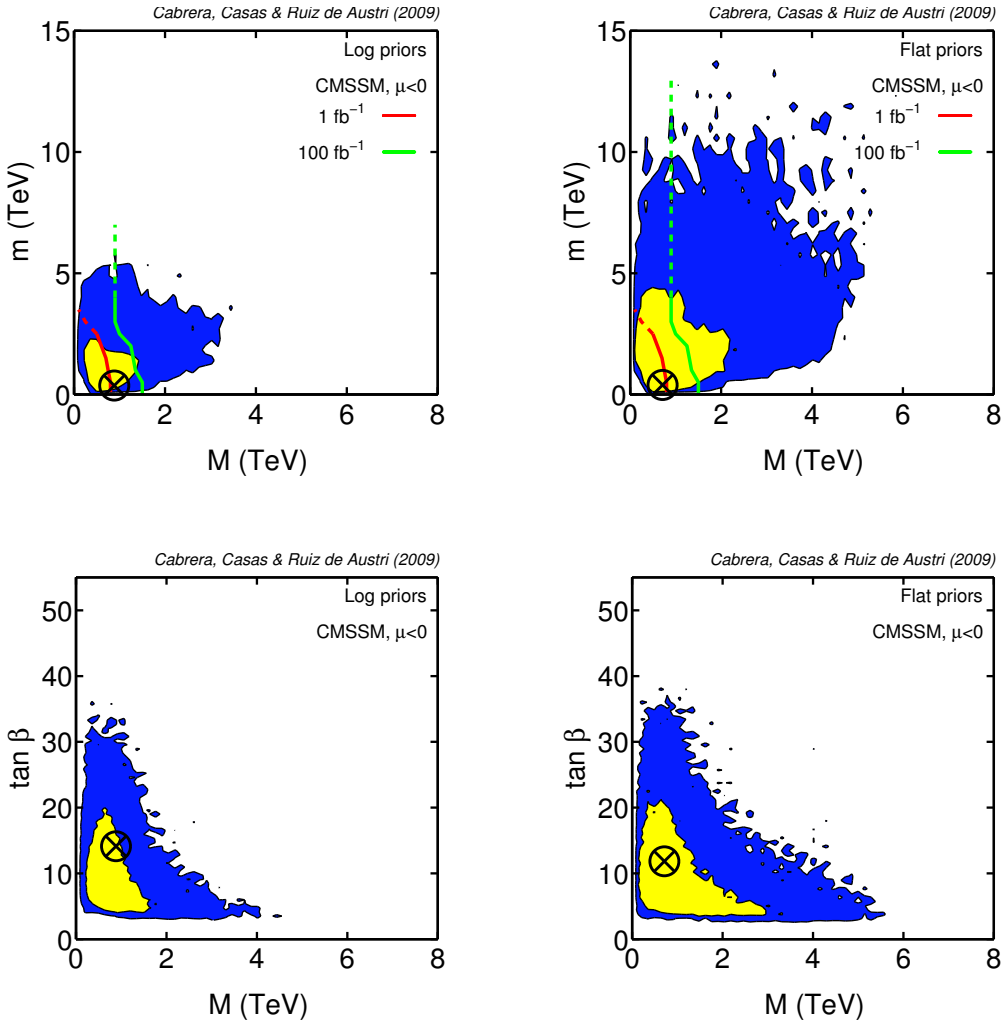

Figure 2.10: As in Fig. 2.4 but with $\mu < 0$.

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 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

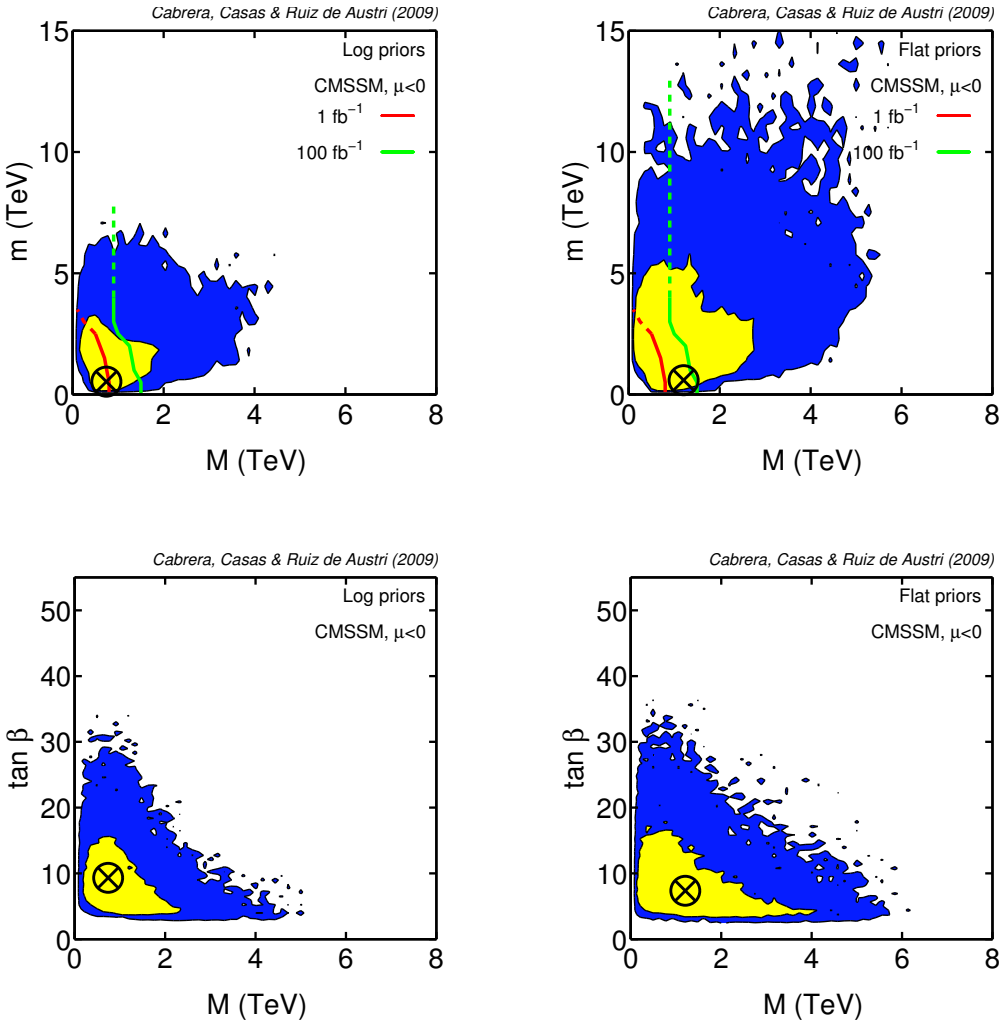

Figure 2.11: As in Fig. 2.7 but with $\mu < 0$.

co-annihilation) regions were preferred (see Fig. 10).

2.2.2 Positive versus negative μ

In order to compare the relative probability of the $\mu > 0$ and $\mu < 0$ branches, one has to evaluate the Bayesian evidences of both cases, 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.(1.3). The results are given in Table 2.4.

The first column of Table 2.4 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 sect. 2.2.1 allows to understand the numbers of the table. When only the

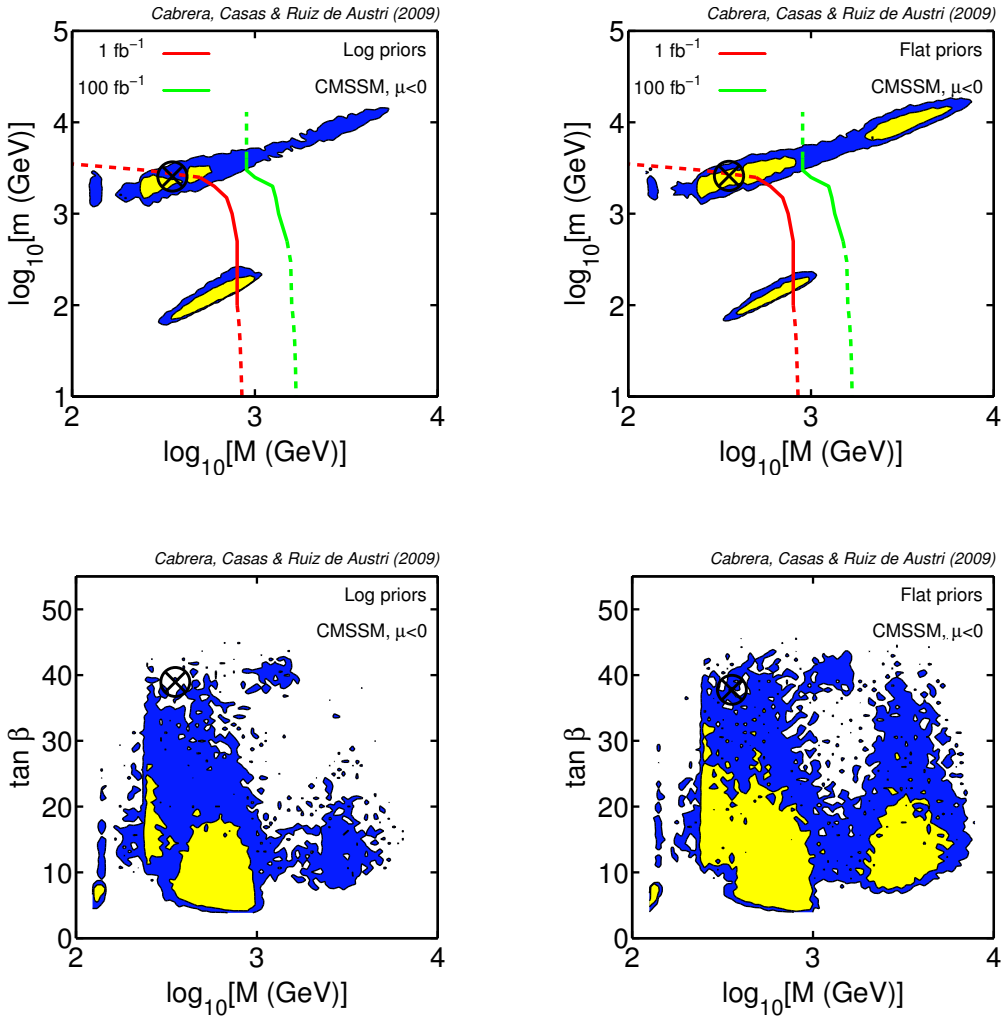

Figure 2.12: As in Fig. 2.8 but with $\mu < 0$.

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 $b \to s, \gamma$, but the effect is not really significant, as is shown by a value of $|\ln B_{+-}|$ well below 0.75, see Table 1.1. This holds when Ω_{DM} constraints are incorporated into the analysis (third row of Table 2.4). On the other hand, $(g-2)_\mu$ constraints (when evaluated using e^+e^- →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 2.4. Using the conventions of Table 1.1, 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.

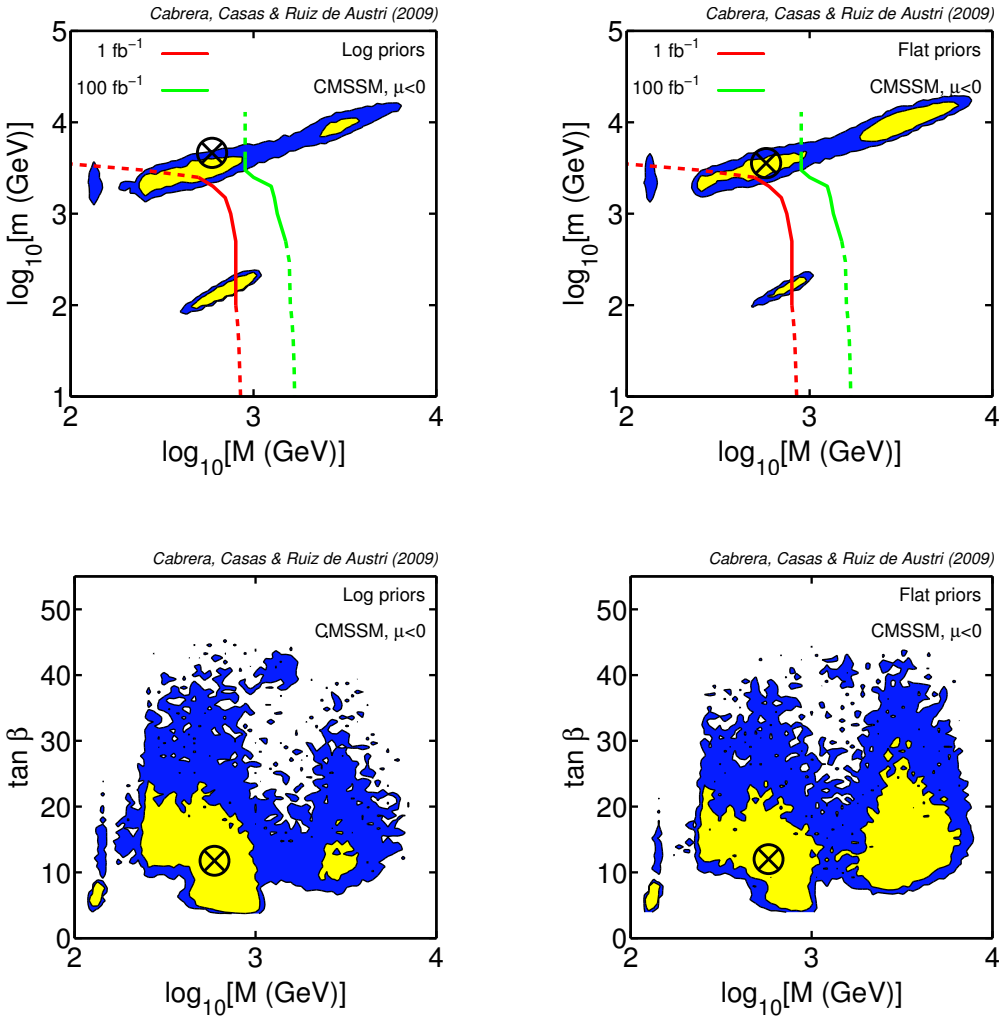

Figure 2.13: As in Fig.2.9 but with $\mu < 0$.

2.3 Comparison to previous work

Some of the previous work in this subject has been collected in refs. [44, 14, 38, 15, 45, 46, 47, 48, 49, 50] All of them are Bayesian analyses, except refs.[46, 50]. However, a fair comparison with our work is tricky, since these articles often make assumptions very different from us about the priors and ranges of the initial parameters (and even about which are the initial parameters). Also they may include different pieces of experimental information. The last point is dramatic regarding $(g-2)_\mu$, as is clear from subsect. 2.1.2. Nevertheless it is interesting to compare our work with this previous literature, to make clearer how all these differences do affect the results and conclusions. For the sake of concreteness, we have considered five previous representative works, corresponding to refs.[44, 38, 15, 49, 50].

In ref.[44], which was pioneering in MSSM Bayesian analyses, $\tan\beta$ was considered as an initial parameter, with flat prior. As a result, there is no penalization of the large $\tan\beta$

region, which thus becomes even favoured by experimental data (probably because of Dark Matter constraints, see below). Besides, the authors include always the experimental data concerning $(g-2)_\mu$ (based on e^+e^- data) and Dark Matter constraints. Finally, the priors for the soft terms are taken as flat, with ranges bound by 2 TeV. Hence their Fig. 2.1 would correspond to our Fig. 2.10 (they are based on essentially the same experimental data). Actually, the $\{M, m\}$ plots of the two figures are not very different, although ours favour more clearly the low-energy region (due to the incorporation of the electroweak scale, as discussed in sect. 1.3). This effect would have been more impressive if in ref.[44] they had unplugged the $(g-2)_\mu$ and Dark Matter data. And much more if, besides, they had widened the allowed range of the soft terms. On the other hand, their $\{M, \tan\beta\}$ plots favour more clearly the region of very large $\tan\beta$ (Higgs Funnel region). In our opinion this effect is not realistic, since $\tan\beta$ is clearly a derived parameter, and this fact introduces a Jacobian factor in the associated probability distribution, penalizing large $\tan\beta$.

In ref.[38] the initial assuptions were similar to those of ref.[44] (and the results are consistent with each other). Therefore the comparison with our results is also similar. In this case, however, the authors tried two different classes of ranges for the soft parameters (up to 2 TeV and 4 TeV) and, also, they probed to disconnect $(g-2)_\mu$. From their Fig. 16, it is clear that, by unplugging $(g-2)_\mu$, the preferred region for M goes from 0.5–1 TeV to 1–1.5 TeV. Comparing to our Fig. 2.8 (which is now the corresponding one), we see that in our analysis the high-energy region is more penalized, which is not surprising.

In ref.[15], a refined version of the analysis of ref.[44] was presented. In this case, $\tan\beta$ was considered a derived parameter (which introduces a Jacobian factor). Also, M_Z was marginalized, as in our case (for a detailed comparison between the two procedures see ref.[90]). Therefore, the initial set up of ref.[15] is the most similar one to ours. Their priors, however, are quite different and somewhat arbitrary (though reasonable). They would correspond more or less to our logarithmic priors, allowing very large ranges for the parameters. In their results the authors observed indeed a penalization of the high-energy region, which they attributed to the choice of the priors. We think, however, that it is mainly a consequence of the marginalization of M_Z, and the effective penalization of fine-tuning that it entails (something that is far from obvious at first sight). In their Fig. 2.2 they compare their results with those of ref.[44]. There one can clearly see the extra penalization of the high-energy region. The $\{M, m\}$ and $\{M, \tan\beta\}$ plots of that figure correspond to the (log prior) plots of our Fig. 2.10. Indeed, both figures are quite consistent (theirs are even more tilted towards low energy, probably due to the additional effect of their choice of priors). Unfortunately, they do not explore unplugging $(g-2)_\mu$ and Dark Matter data, so a comparison with other results and plots of our work is not possible.

In ref.[49] a Bayesian analysis of the so-called pMSSM ("phenomenological MSSM") was presented. This model has many initial parameters (~ 20), all of them defined at low-energy. Apart from that, the set up of the analysis was similar to that of ref.[44]. In particular they took $\tan\beta$ as an initial parameter, and considered flat priors and finite ranges for the soft parameters (< 4TeV), including $(g-2)_\mu$ and Dark Matter experimental data in all instances. In order to make any comparison with our work, one has to focus on particular quantities.

A good example is the gluino mass, $M_{\tilde{g}}$, which for mSUGRA is $\sim 2.5M$. From their Fig. 2.2 the peak of the probability distribution of $M_{\tilde{g}}$ is around 2–3 TeV, which would correspond to $M \sim 1$ TeV. This should be compared with our Fig. 2.10 (the one based on a similar set of experimental data). In our case the peak of the distribution is around 400 GeV, showing that we get an extra penalization of the high-energy region, as explained in this work. Unplugging $(g-2)_\mu$ and Dark Matter, the differences would have been more dramatic, especially if the allowed ranges of the parameters were stretched.

Finally, in ref.[50] a frequentist analysis of the MSSM was presented. This is a point of view complementary to the Bayesian approach, followed here. The authors of ref.[50] perform a scan of the parameter space of the CMSSM (and also of the so-called NUHM1 model), evaluating the likelihood (based on the χ^2). This leads to zones of estimated probability (inside contours of constant χ^2) around the best fit points in the parameter space. Their Fig. 1.2 (($\{M, m\}$ plane) corresponds to our Fig. 2.10. However, notice that, in their case, the unplotted variables are optimized to obtain the best χ^2, whereas in our case they are marginalized. Nevertheless, it is remarkable that the two figures are quite similar (especially comparing with our log-prior plot). This is an encouraging result. Indeed, the frequentist and Bayesian approaches must converge when the quality of data increases. This means that the bulk of the probability is centered around the best-fit points. This coincidence is also observed when the authors probe to unplug $(g-2)_\mu$ (compare their Fig. 2.1 to our Fig. 2.8). Since they do not explore to unplug Dark Matter, it is not possible to make further comparisons. It is likely, that in that case their 68% and 95% c.l. regions become much more extended in the parameter space, thus taking up a large portion of the high-energy (non-accesible to LHC) region. Notice that a frequentist approach cannot penalize those regions from fine-tuning arguments. Fine-tuning has to do with statistical weight (see sect. 1.3) and a frequentist analysis is based in likelihood, i.e. the ability to reproduce the experiment. Without $(g-2)_\mu$ and Dark Matter data, the experimental reasons to stick to low-energy are much less powerful. In other words, without $(g-2)_\mu$ and Dark Matter it is likely that the convergence between frequentist and Bayesian approaches is still weak.

Observable	Mean value	Uncertainties		ref.
	μ	σ (exper.)	τ (theor.)	(exper.)
M_W	80.398 GeV	27 MeV	15 MeV	[67]
$\sin^2\theta_{\rm eff}$	0.23149	17×10^{-5}	15×10^{-5}	[67]
$a_\mu^{\rm exp}\times10^{10}$	11659208.9	6.33	-	[52]
$\delta a_\mu\times10^{10}$ (e^+e^-)	29.5	8.8	2.0	[68]
$\delta a_\mu\times10^{10}$ (τ)	14.8	8.2	2.0	[69]
ΔM_{B_s}	17.77 ps^{-1}	0.12 ps^{-1}	2.4 ps^{-1}	[70, 71]
$BR(\overline{B}\to X_s\gamma)\times10^4$	3.52	0.33	0.3	[72]
$\frac{BR(B_u\to\tau\nu)}{BR(B_u\to\tau\nu)_{SM}}$	1.28	0.38	-	[72]
$\Delta_{0-}\times10^2$	3.6	2.65	-	[73]
$\frac{BR(B\to D\tau\nu)}{BR(B\to De\nu)}\times10^2$	41.6	12.8	3.5	[74]
R_{l23}	1.004	0.007	-	[75]
$BR(D_s\to\tau\nu)\times10^2$	5.7	0.4	0.2	[76]
$BR(D_s\to\mu\nu)\times10^3$	5.8	0.4	0.2	[76]
$\Omega_\chi h^2$	0.1099	0.0062	$0.1\,\Omega_\chi h^2$	[77]
	Limit (95% CL)	τ (theor.)		ref.
$BR(\overline{B}_s\to\mu^+\mu^-)$	$<5.8\times10^{-8}$	14%		[71]
m_h	>114.4 GeV (SM-like Higgs)	3 GeV		[78]
ζ_h^2	$f(m_h)$	negligible		[78]
$m_{\tilde{q}}$	>375 GeV	5%		[52]
$m_{\tilde{g}}$	>289 GeV	5%		[52]
other sparticle masses	As in table 4 of ref. [38].			

Table 2.2: 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. $\delta a_\mu = a_\mu^{\rm exp} - a_\mu^{\rm SM}$ denotes the discrepancy between the experimental value and the SM prediction of the anomalous magnetic moment of the muon $(g-2)_\mu$. As explained in the text, for each quantity we use a likelihood function with mean μ and standard deviation $s=\sqrt{\sigma^2+\tau^2}$, where σ is the experimental uncertainty and τ represents our estimate of the theoretical uncertainty. Lower part: Observables for which only limits currently exist. The likelihood function is given in ref. [38], including in particular a smearing out of experimental errors and limits to include an appropriate theoretical uncertainty in the observables. m_h stands for the light Higgs mass while $\zeta_h^2 \equiv g^2(hZZ)_{\rm MSSM}/g^2(hZZ)_{\rm SM}$, where g stands for the Higgs coupling to the Z and W gauge boson pairs. The references for the theoretical calculations are given in the text.

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 2.3: 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.

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 + $\Omega_{\rm DM}$	-0.31 ± 0.068	-0.66 ± 0.066
EW + Bounds + B-physics + $(g-2)_\mu$ + $\Omega_{\rm DM}$	1.9 ± 0.065	3.71 ± 0.068

Table 2.4: The natural log of the Bayes factor (ln B_{+-}) for $\mu > 0$ and $\mu < 0$. A positive (negative) value indicates a preference for $\mu > 0$ ($\mu < 0$).

Chapter 3

Higgs mass vrs $(g-2)_\mu$, a test of consistency for the CMSSM

Supersymmetry has been often invoqued as the new physics that might reconcile the experimental muon magnetic anomaly, a_μ, with the theoretical prediction (basing the computation of the hadronic contribution on e^+e^- data). However, in the context of the CMSSM, the required supersymetric contributions (which grow with decreasing supersymmetric masses) are in potential tension with a possibly large Higgs mass (which requires large stop masses). In the limit of very large m_h supersymmetry gets decoupled, and the CMSSM must show the same discrepancy as the SM with a_μ. But it is much less clear for which size of m_h does the tension start to be unbearable. In this chapter we show how tension can be quantified with the help of Bayesian techniques.

In section 3.1 we review the status of $(g-2)_\mu$ and the supersymmetric prediction for it. In section 3.2 we show the potential tension between the requirement of suitable SUSY contributions to the muon anomaly and a possibly large Higgs mass. In section 3.3 we quantify such tension as a function of m_h, with the help of Bayesian techniques. In section 3.4 we show how the probability distributions of the most relevant parameters (universal scalar and gaugino masses, and $\tan\beta$) change with increasing m_h.

3.1 $(g-2)_\mu$ and supersymmetry

The magnetic anomaly of the muon, $a_\mu = \frac{1}{2}(g-2)_\mu$ has been a classical and powerful test for new physics. As it is known, the present experimental value and some of the theoretical determinations of a_μ show a remarkable discrepancy, suggesting physics beyond the Standard Model (SM) to account it. However, the situation is still uncertain, due essentially to inconsistencies between alternative determinations of the contribution coming from the hadronic vacuum-polarization diagram, say $\delta^{\rm SM}_{\rm had}a_\mu$.

This contribution can be expressed in terms of the total hadronic cross section $e^+e^- \to$ had.

Using direct experimental data for the latter, one obtains a final result for a_μ, which is at more than 3 σ from the current experimental determination [**81**], namely

$$\delta a_\mu \;\; = \;\; a_\mu^{\rm exp} - a_\mu^{\rm SM} = 25.5 \pm 8.0 \;\; \times 10^{-10} \tag{3.1}$$

(the quoted error bars are 1 σ). This discrepancy has been often claimed as a signal of new physics. Obviously, if one accepts this point of view, the discrepancy should be cured by contributions from physics beyond the SM.

Admittedly, such claims are too strong. We are quite aware of past experimental observables in apparent disagreement with the SM prediction, which have eventually converged with it. This has occured due to both experimental and theoretical subtleties that sometimes had not been fully understood or taken into account. As a matter of fact, the experimental $e^+e^- \to$ *had* cross section exhibits some inconsistencies between different groups of experimental data. Using only BABAR data the discrepancy reduces to 2.4 σ, while without it the discrepancy becomes 3.7 σ, [81]. The inconsistency is specially notorious if one considers hadronic τ decay data, which are theoretically related to the $e^+e^- \to$ *had* cross section. Using just τ-data the disagreement becomes 1.9 σ, [81], [69]. Although the more direct e^+e^- data are usually preferred to evaluate $a_\mu^{\rm SM}$, these inconsistencies are warning us to be cautious about the actual uncertainties involved in the determination of $a_\mu^{\rm SM}$.

If one takes the discrepancy between theory and experiment shown in eq.(3.1) as a working hypothesis, one has to consider possible candidates of new physics able to provide the missing contribution to reproduce $a_\mu^{\rm exp}$. The Minimal Supersymmetric Standard Model (MSSM) is then a natural option. We will consider here the simplest and most extensively analyzed version of the MSSM, namely the so-called constrained MSSM (CMSSM).

The main supersymmetric (CMSSM) contributions to a_μ come from 1-loop diagrams with chargino-sneutrino and neutralino-smuon exchange [91]. In general, these contributions, say $\delta^{\rm MSSM} a_\mu$, are larger for smaller supersymmetric masses and can be just of the right magnitude to reconcile theory and experiment (thus constraining the CMSSM parameter space).

3.2 Higgs mass vs. $(g-2)_\mu$

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 to reconcile the theoretical predictions with the present experimental lower bound, $m_h > 114.4$ GeV (SM-like Higgs). Roughly speaking, a Higgs mass above 130 GeV requires supersymmetric masses above 1 TeV. In this regime one can expect SUSY to be decoupled, so that the prediction for a_μ becomes close to $a_\mu^{\rm SM}$. Hence, a large Higgs mass in the MSSM would necessarily amounts to a > 3 σ discrepancy between the experimental and the theoretical values of a_μ (evaluated via $e^+e^- \to$ had).

The main goal of this chapter is to quantify the tension between m_h and a_μ in the context of the CMSSM. This is useful since it allows to put an educated upper bound on the Higgs

mass, which will depend on the discrepancy one is ready to tolerate. Conversely, it tells us from which minimum value of $m_h^{\rm exp}$ we will have to give up either the CMSSM assumption or the theoretical evaluation of a_μ via $e^+e^- \to$ had (with the quoted uncertainties).

For the sake of the discussion, we will give now some approximate analytical expressions for m_h and $\delta a_\mu^{\rm MSSM}$. In the MSSM the tree-level squared Higgs mass plus the one-loop leading logarithmic contribution is given by

$$m_h^2 \;\simeq\; M_Z^2 \cos^2 2\beta \;+\; \frac{3 m_t^4}{2\pi^2 v^2}\left[\log\frac{m_{\tilde t}^2}{m_t^2} + \frac{X_t^2}{M_S^2}\left(1 - \frac{X_t^2}{12 M_S^2}\right)\right] \;+\; \cdots$$

Here $\tan\beta$ is the ratio of the expectation values of the two MSSM Higgs fields, $\tan\beta \equiv \langle H_u\rangle/\langle H_d\rangle$; m_t is the (running) top mass and $m_{\tilde t}$ is the geometrical average of the stop masses. Besides,

$$X_t \equiv A_t + \mu \cot\beta, \tag{3.2}$$

where A_t is the top trilinear scalar coupling, and M_S^2 is the arithmetical average of the squared stop masses. All the quantities in eqs.(3.2), (3.2) are understood at low energy (for more details see e.g. ref.[92, 80, 93, 94, 95, 96, 97]). Subdominant terms not written in eq.(3.2) can be important for a precise determination of m_h, and we have included them in the numerical analysis. The previous equations tell us how m_h grows with increasing supersymmetric masses and also with increasing $\tan\beta$. Besides, the contribution associated to the stop mixing (second term within the square brackets in eq.(3.2)) is maximal at $X_t = \sqrt{6} M_S$.

In the MSSM the tree-level squared Higgs mass plus the one-loop leading logarithmic contribution is given by (3.2).

On the other hand, as mentioned above, the supersymmetric contribution to the muon anomaly, $\delta^{\rm SUSY} a_\mu$, arises mainly from 1-loop diagrams with chargino-sneutrino and neutralino-smuon exchange. This contribution increases with increasing $\tan\beta$ and decreasing supersymmetric masses. See refs.[98, 99, 82, 83, 100].

Although the analytical expressions are complicated, one can get an intuitive idea of the parametric dependence by considering the extreme case where the masses of all supersymmetric particles are degenerate at low energy[1]: $M_1 = M_2 = \mu = m_{\tilde\mu L} = m_{\tilde\mu R} = m_{\tilde\nu} \equiv M_{\rm SUSY}$. Then [101],

$$\delta^{\rm SUSY} a_\mu \simeq \frac{1}{32\pi^2}\frac{m_\mu^2}{M_{\rm SUSY}^2} g_2^2 \tan\beta\, {\rm sign}(M_2\mu). \tag{3.3}$$

Examining the approximate expressions (3.2) and (3.3), it is clear that a large m_h and a large $\delta^{\rm SUSY} a_\mu$ will be more easily obtainable (and thus compatible) for larger $\tan\beta$. On the contrary, the larger the supersymmetric masses the larger m_h but the smaller $\delta^{\rm SUSY} a_\mu$, and this is the origin of the potential tension.

[1] This limit is often used because of the simplification of the formulae it implies. However, it is unachievable in the CMSSM.

However, it is difficult from the previous expressions (or the more sophisticated ones) to conclude for which size of m_h does the tension start to be unbearable. The reason is that a particular value of the Higgs mass, say $m_h = 120$ GeV, can be achieved through eq.(3.2) with different combinations of $\tan\beta$, stop masses and X_t. Besides, there are many ways, i.e. very different regions in the MSSM parameter space, in which these quantities can have similar *low-energy* values. Still, the corresponding contribution $\delta^{\rm SUSY} a_\mu$ can change significatively from one region to another. Unless one performs a complete scan of the parameter space one cannot conclude that the required value of $\delta^{\rm SUSY} a_\mu$ is unattainable for $m_h = 120$ GeV. On the other hand, if it is attainable, but only in an extremely tiny portion of the parameter space, this implies a tension between the two observables since the consistency between m_h and a_μ requires a severe fine-tuning. And it is possible, in principle, to quantify such tension.

In the analysis we have included two-loop leading corrections for the Higgs sector [102, 103, 104, 105, 54]. $\delta^{\rm SUSY} a_\mu$ was computed 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 [82]. The effective two-loop effect due to a shift in the muon Yukawa coupling proportional to $\tan^2\beta$ has been added as well [83].

Next we expound how a systematic analysis of this kind can be done with the help of Bayesian techniques. This will allow us to quantify the tension between m_h and a_μ as a function of m_h.

3.3 Quantifying the tension between m_h and a_μ

As mention in sec. 1.1, when two different models (or hypotheses) are used to fit the data, the ratio of their evidences gives the relative probability of the two models in the light of the data. For an application to model selection in the context of the CMSSM, see [106].

In order to quantify the tension between m_h and a_μ, following Ref. [4] we separate the complete set of data in two subsets:$\mathscr{D}$ is the experimental value of a_μ, whereas D is given by all the standard electroweak observables, B- and D-physics observables, limits on supersymmetric masses, etc (for the complete list of experimental data used, see Table 2.2). D includes also the value of m_h that we are probing, and thus provisionaly assumed to be the actual one. Hence, we will not consider any experimental error in the value of m_h, just the uncertainty associated to the theoretical calculation (estimated as ± 2 GeV).

The consistency of the measured muon anomaly, $\mathscr{D}^{\rm obs}$, with the rest of data, D, in the context of the model (CMSSM), can be evaluated using the $\mathscr{L}-$test, (1.7). In our case, the value of $\mathscr{D}^{\rm max}$ depends on the value of m_h probed. For very large m_h, say $m_h \geq 135$ GeV, SUSY must decouple, so $\mathscr{D}^{\rm max}$ should approach the SM prediction. Hence, in this limit one expects $\mathscr{L}(\mathscr{D}^{\rm obs}|D)$ to show a $3.2\,\sigma$ discrepancy; in other words, $-2\ln\mathscr{L}(\mathscr{D}^{\rm obs}|D) \to 3.2^2$. However, the expression (1.7) allows us to evaluate this likelihood for any intermediate value of m_h, and so we can evaluate how quickly this limit is reached as a function of the assumed value for m_h.

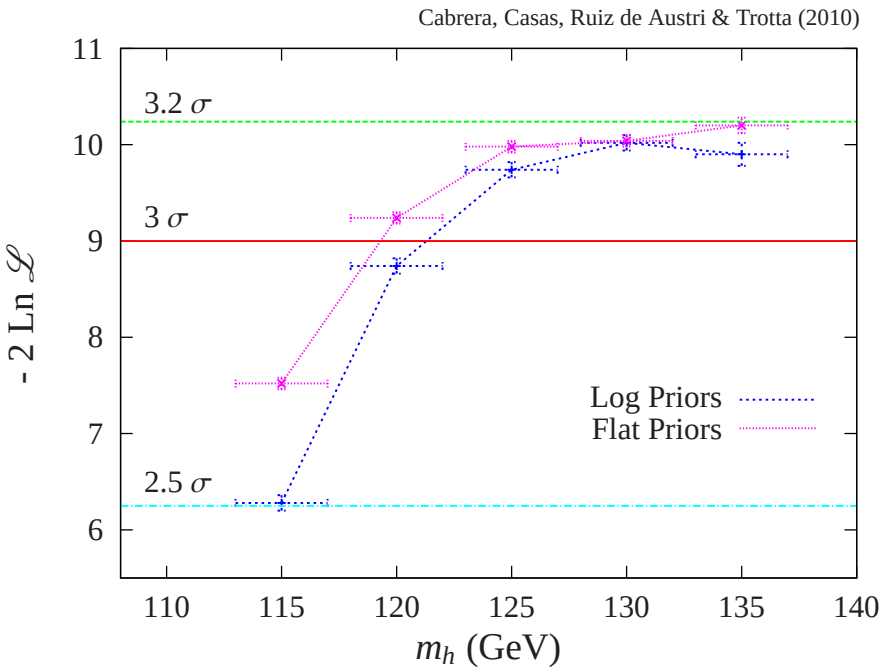

Figure 3.1: The $-2\ln\mathscr{L}$ test statistics, as defined in Eq. (1.7), as a function of the assumed value for m_h in the CMSSM framework with logarithmic (blue, lower dotted line) and flat (violet, upper dotted line) priors. Horizontal lines denote thresholds of 2.5 σ, 3 σ and 3.2 σ discrepancy.

For the numerical calculation we have used the MultiNest [30, 107, 47] algorithm as implemented in the `SuperBayeS` code [31, 38, 45]. It is based on the framework of Nested Sampling, recently invented by Skilling [32, 33]. 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 at no extra computational cost.

Fig. 3.1 shows the value of $-2\ln\mathscr{L}$ (the analogous of the usual χ^2) for different values of the Higgs mass, $m_h(\text{GeV}) = 115, 120, 125, 130, 135$, and for two different choices of initial priors for the CMSSM parameters, namely log prior (red line) and flat prior (blue line). The precise shape of the log and flat priors used here is the one derived in ref.[108], to which the reader is referred, which take into account the likelihood associated to the electroweak breaking process. The horizontal error bars reflect the uncertainty in the theoretical computation of m_h in the MSSM, while the vertical error bars come from sources of error in the computation of $\mathscr{L}$, mainly the numerical accuracy of the evidence returned by MultiNest. Lines of conventional confidence levels thresholds in terms of number of σ are shown as well for comparison.

From the figure we see that the likelihood of the experimental value of a_μ approaches asymp-

totically the expected 3.2 σ discrepancy for large values of m_h, for both types of priors. As mentioned above, this is logical and it represents a nice cross-check of the reliability of the whole procedure. Besides, Fig. 3.1 tells us how fast this convergence is reached as m_h increases. And, as a matter of fact, the convergence is very fast. At $m_h = 125$ GeV the maximum level of discrepancy is already achieved, indicating that SUSY has decoupled, and thus the prediction for a_μ coincides with the SM one. If we require less than 3 σ discrepancy, we need $m_h \lesssim 120$ GeV. This is a prediction of the CMSSM provided we accept the calculation of a_μ based on e^+e^- data. For a larger Higgs mass we should give up either the CMSSM model or the computation of a_μ based on e^+e^-; or accept living with such inconsistency. Let us also note that, even assuming a Higgs mass as low as it can be, the minimum level of discrepancy is about 2.5 σ. However, most of this tension with a_μ comes from $b \to s, \gamma$ data [4], rather from the value of the Higgs mass. This can be checked by repeating the analysis excluding all the experimental information (except M_Z and the assumed Higgs mass). The resulting plot is similar to that of Fig. 3.1, except the $m_h = 115$ GeV point, which shows a ~ 1.5 σ discrepancy.

It is an interesting exercise to compute how our conclusions would change if a_μ became more precisely measured in the future (keeping the same central value). If one continued to assume the theoretical evaluation of a_μ based on e^+e^- data, the signal for new physics would obviously become stronger. In this case, the tension between a large Higgs mass and the experimental a_μ would get more unbearable. We have done this excercise, by changing (artificially) the experimental uncertainty of $a_\mu^{\rm exp}$, so that the discrepancy with the SM result be 5 σ, something that could happen in the next years. Now, in the context of the CMSSM, the value of $-2\ln \mathscr{L}(\mathscr{D}^{\rm obs}|D)$ must approach asymptotically such 5 σ discrepancy, and this is indeed what we observe, as shown in Fig. 3.2. In this hypothetical situation, a Higgs mass above 120 Gev would imply a discrepancy larger than 4 σ with the muon anomaly in the context of the CMSSM. Actually, the present lower bound, $m_h \geq 114.4$ GeV, would already be inconsistent with the muon anomaly at the 3 σ level.

This gives a fair idea of the tensions within the CMSSM to accommodate a value of a_μ as the measured one (basing the theoretical calculation on present e^+e^- data).

3.4 Probability distributions for supersymmetric parameters

It is also interesting to investigate the probability distributions of the CMSSM parameters for various assumed values of the Higgs mass. Figure 3 (upper panels) shows the marginalized probability distribution functions (pdfs) of m, M assuming a value of $m_h = 115, 120, 125$ (GeV), as well as adding in all present-day constraints mentioned above. The location of the peak in the posterior pdf increases with the assumed Higss mass since, as mentioned in section 3.2, in the MSSM a large m_h requires large radiative contributions, which grow logarithmically with the stops masses. This happens even though large values of m and M are penalized both for a natural electroweak breaking (see ref. [90, 108]) and by the need of

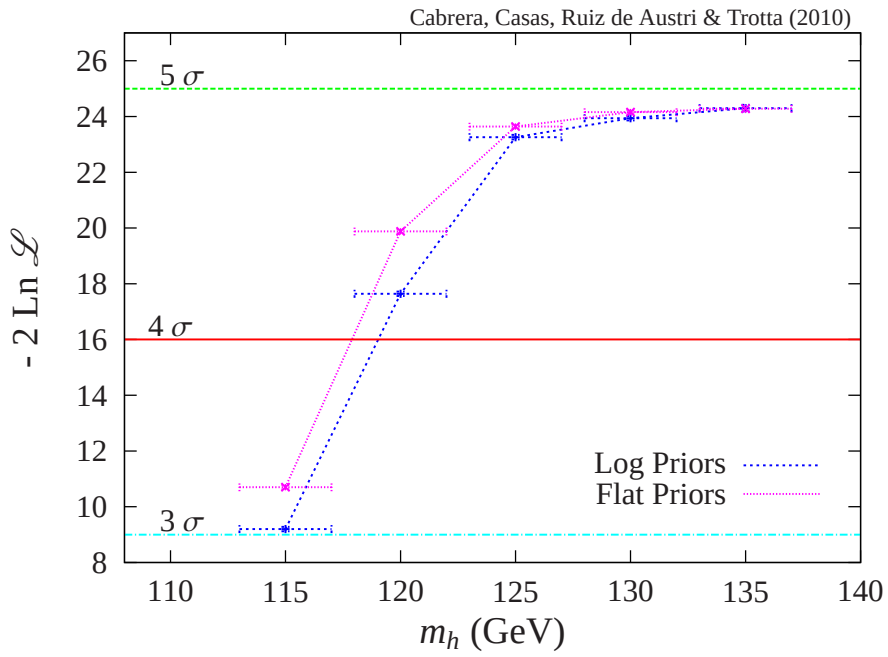

Figure 3.2: As in Fig. 3.1, but (artificially) assuming an improved experimental determination of a_μ, so that the SM discrepancy becomes 5 σ. Now the horizontal lines denote thresholds of 3 σ, 4 σ and 5 σ discrepancy.

a sizeable $\delta^{\rm SUSY} a_\mu$. The model "prefers" to reproduce m_h at the cost of not reproducing a_μ rather than viceversa. Note here that for increasing soft masses the discrepancy of a_μ with the experimental value approaches 3.2 σ, but if the soft masses are not large enough, the discrepancy associated to m_h would be much more severe.

Fig. 3.3 (lower panel) shows the pdf of $\tan\beta$ for $m_h = 115, 120, 125$ (GeV). Its shape is the result two competing effects. On the one hand, large values of $\tan\beta$ are severely penalized for the electroweak breaking, see chapters 1 and 2. On the other hand, the need of a sizeable $\delta^{\rm SUSY} a_\mu$ favours large $\tan\beta$ (see the approximate expression (3.3)). Fig. 3.3 shows the balance between these two effects. [The Higgs mass increases also with $\tan\beta$, but the effect is only important for small values of $\tan\beta$, see eq.(3.2)]. Now, since for larger m_h the soft masses are larger, with the side-effect of suppressing $\delta^{\rm SUSY} a_\mu$, one might expect that the preferred value of $\tan\beta$ increases with m_h, to compensate this in eq. (3.3). However, this effect is not very important, as it is apparent in Fig. 3. To understand this, let us approximate (for the sake of the argument) $M_{susy} \sim m_{\tilde{e}_{L,R}}$ in eq. (3.3) and use [5], [109]

$$\begin{aligned} m^2_{\tilde{e}_L} &\simeq m^2 + 0.54 M^2, \\ m^2_{\tilde{e}_R} &\simeq m^2 + 0.15 M^2, \\ m^2_{\tilde{t}} &\simeq 3.36 M^2 + 0.49 m^2 - 0.05 A^2 - 0.19 AM + m_t^2. \end{aligned} \tag{3.4}$$

Since m_h increases (logarithmically) with $m^2_{\tilde{t}}$, while $\delta^{\rm SUSY} a_\mu$ is suppressed by $m^2_{\tilde{e}_{L,R}}$, it might seem that the most efficient way to reproduce both is to increase M rather than m (note the

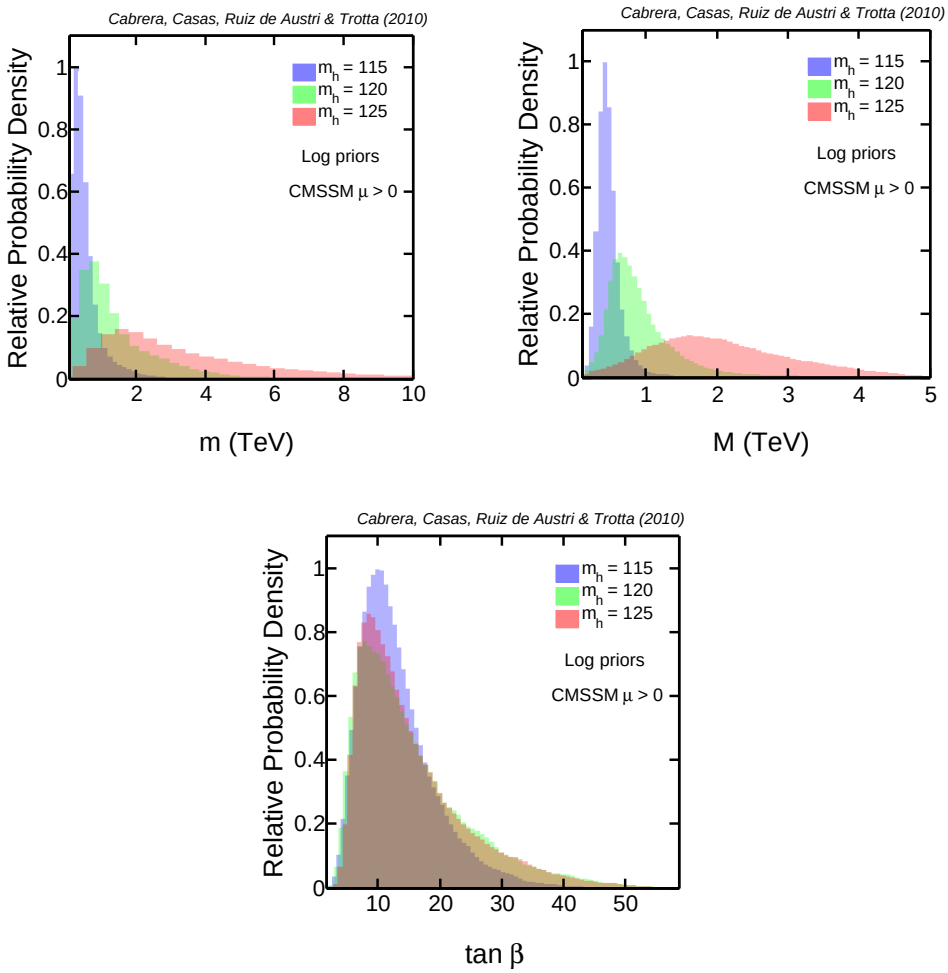

Figure 3.3: Probability distribution functions of m, M (upper panels) and $\tan\beta$ (lower panel) for $m_h = 115$ GeV (blue), 120 GeV (green) and 125 GeV (red).

different dependences on M in eqs. (3.4)). The problem is that the fine-tuning grows very fast with M; in other words, the number of points in the parameter space with correct EW breaking decresases very quickly. In consequence this possibility is statistically penalized. On the contrary, for small M and large m, if $\tan\beta > 8$, there is a focus-point region, with small fine-tuning. This region is statistically favoured, though this is counteracted by the penalization arising from the suppression in $\delta^{\rm SUSY} a_\mu$. This cannot be compensated by larger values of $\tan\beta$, since in this regime very big values of $\tan\beta$ (as would be needed for such compensation) start to be forbidden as we increase m. In consequence, a very large $\tan\beta$ is hardly favoured by an increasing m_h.

Finally, let us mention that a lot of effort has been done in the literature to determine the most probable region of the parameter space of the CMSSM [50, 49, 48, 47, 46, 15, 38, 14, 44, 110, 111]. This includes both Bayesian approaches (as the one followed here) and frequentist

ones. The latter (which can be considered as complementary to the Bayesian ones) are based on the analysis of the likelihood function in the parameter space. Thus they do not penalize regions from fine-tuning arguments, as we mention in sec 2.3. In consequence, following a frequentist approach it would be much more hard to show up the tension between m_h and $g-2$. On the other hand, the present analysis differs from the previous ones in the fact that several hypothetic future scenarios, depending on the value of the Higgs mass, are considered and compared.

Chapter 4

Histogram comparison as a powerful tool for the search of new physics at LHC. Application to CMSSM

In this chapter we deal with the incorporation of direct LHC data in the searches of new physics, which is the logical next step after performing the previous forecasts.

The LHC is already probing new physics beyond the reach of past experiments. At any stage of this enterprise, i.e. with the available data at any time, there are two main questions to address: 1) Is there any signal of New Physics (NP)? and 2) In the positive case, which NP is it? In order to optimize the answer to these questions there is an intense activity to explore assorted strategies for the search of NP. The task is challenging, due in part to the fact that LHC data, though very rich, are not as clean as those from an e^+e^- collider. Besides, the theoretical calculations are also subject to great uncertainties and rely to some extent on Monte Carlo simulations.

Most of the LHC data can be organized in form of histograms with number of events of a certain kind (e.g. those presenting multijets + missing energy) displayed in different variables [40, 41, 112]: M_{eff}, $\not{p}_T$, α_T, etc. In many cases the comparison with the simulations is done just by comparing the total number of events after performing different cuts in the variables involved. In this way, both ATLAS and CMS have already posed meaningful bounds [113, 114] on NP scenarios, in particular on the already discussed in previous chapters CMSSM [115]. More precisely, the most powerful bounds on the CMSSM have been obtained by considering events with several jets and missing transverse momentum. Somehow, the study of the total number events, choosing different cuts on the sets of data, amounts to partially compare the shapes of the experimental and the theoretical (simulated) histograms; although in a way which is not optimal.

As mentioned above, even if we are quite sure to have a signal of NP, we face the problem of identifying the model producing such signal. Of course the variety of scenarios of NP is enormous, which makes the job very complex. Even playing in the framework of a given

scenario, such as the CMSSM, the sole study of the number of events of a certain kind is not enough to determine the parameters of the model, due to the existence of big degeneracies in such determination. Again, this situation can be improved by probing different cuts in the sets of data. But, once more, this is not an optimized way of comparing theory and experiment, since the richness of the data is not completely exploited.

In this chapter we present an effective and rigorous way to compare experimental and theoretical histograms, incorporating the different sources of uncertainty involved in the task. In our opinion, in an experiment with the characteristics of the LHC this is a useful tool to extract as much information as possible from the comparison between experimental data with theoretical simulations. We illustrate this usefulness by showing how a search in the CMSSM parameter space, using Bayesian techniques, can effectively find the correct values of the CMSSM parameters by comparing histograms of events with multijets + missing energy displayed in the M_{eff} variable. This procedure could be very efficient to identify the true supersymmetric model, in the case supersymmetry is really there and accessible to the LHC.

In section 4.1 we establish the notation and the statistical basis for the rigorous comparison between the experimental and the theoretical histograms. Section 4.2 is devoted to the incorporation of extra sources of uncertainty, in particular systematic ones. At the end of this section we give our final formula for the complete likelihood of a theoretical model by histogram comparison. In section 4.3 we illustrate the proposed technique by a showing how a search in the constrained-MSSM parameter space, using Bayesian techniques, can effectively find the correct values of the MSSM parameters by comparing histograms of events with multijets + missing energy displayed in the effective-mass variable. But, of course, the technique can be applied to any scenario of new physics.

4.1 Comparison of histograms. Statistical uncertainties

4.1.1 Basic Ingredients and Notation

Suppose we have experimental data, e.g. multijet + missing energy events at LHC, organized in an histogram upon some variable M, e.g. the effective mass of the events, as defined in ref. [41]. Let us call K the number of bins of the histogram. Each bin corresponds to a (central) value of the effective mass, M_i. We will denote the bin contents (number of events for each M_i) by v_i. The total number of events is $v = \sum_i v_i$.

Leaving apart for the moment all sources of systematic and accidental uncertainties, except the purely statistical ones, the probability that the experiment produces the actual data, v_i,

is given by a Poisson distribution

$$\mathcal{P}(v_i) = \prod_{i=1}^{K} \frac{\nu_i^{v_i}}{v_i!} e^{-\nu_i}, \tag{4.1}$$

where ν_i are the expected values (or "means") of the distribution. The values of ν_i are in principle calculable (at some degree of precision) provided we knew the theory responsible for them, e.g. the Standard Model. But we are precisely trying to uncover unknown NP, therefore ν_i are unknown.

On the other hand, working within a scenario of NP defined by some parameters, θ_a (in our example the parameters of the CMSSM), we can in principle calculate the means under, supposedly, the same conditions of energy and luminosity as the experiment. We will denote μ_i these theoretical means. Of course, μ_i depend on the point in the parameter space, i.e. the precise model under consideration. If the model is the true one, then $\nu_i = \mu_i$. This is the so-called "null-hypothesis". The likelihood of a point of the parameter space is the corresponding probability of producing the observed data, v_i, under the null-hypothesis, i.e.

$$\mathcal{P}(v_i) = \prod_{i=1}^{K} \frac{\mu_i^{v_i}}{v_i!} e^{-\mu_i}. \tag{4.2}$$

The likelihood is a crucial quantity to compare the viability of the different regions of the parameter space, both in frequentist and Bayesian analyses (see, e.g., [2]). In particular, in Bayesian analyses one is interested in determining the probability density of a point of the parameter space, θ_a, given an experimental set of data (in our case, v_i). This is the so-called posterior probability density function, $p(\theta_a|\text{data})$, which is given by the fundamental Bayesian relation

$$p(\theta_a|\text{data}) \;=\; p(\text{data}|\theta_a)\; p(\theta_a)\; \frac{1}{p(\text{data})}. \tag{4.3}$$

Here $p(\text{data}|\theta_a)$ is the above-mentioned likelihood, i.e. the probability of obtaining the observed data if the model defined by the θ_a parameters is the true one; while $p(\theta_a)$ is the prior, i.e. the "theoretical" probability density that we assign a priori to the point in the parameter space; and $p(\text{data})$ is a normalization factor that ensures that the total probability is one.

In order to compute the likelihood (4.2) we need the theoretical means, μ_i. However, in practice one does not have at disposal a complete evaluation of μ_i, but rather a simulation of the process using diverse computation codes. The results of the simulation can also be organized in an histogram with K bins, associated with the same values of the effective mass, M_i. The bin contents of the simulation are denoted by u_i, with total number of events $u = \sum_i u_i$. Of course, the values of u_i obey also a Poisson statistics

$$\mathcal{P}(u_i) = \prod_{i=1}^{K} \frac{\mu_i^{u_i}}{u_i!} e^{-\mu_i}. \tag{4.4}$$

Here we have again left aside for the moment all sources of systematic and accidental uncertainties associated with the theoretical simulation, except the purely statistical ones.

4.1.2 Computation of the likelihood

As mentioned, usually the codes provide values for u_i, but not for μ_i. If we had enough computation time we could obtain a good evaluation of the theoretical means, μ_i, since, increasing the statistics, the bin contents would approach the mean values with decreasing relative uncertainty. This would be the case if we knew from the beginning which specific model we want to test, but this procedure is not efficient if we want to scan the parameter space, testing thousands or millions of models (points in that space). So, identifying u_i with μ_i is not justified unless u_i is large. The relation between them is given by eq.(4.4). Since we are not sure about the values of μ_i, we cannot directly calculate the likelihood $\mathcal{P}(v_i)$ from eq.(4.2). The best we can do is calculate $\mathcal{P}(v_i|u_i)$, i.e. the probability of getting the experimental data, v_i, under the assumption that the model is the true one (null-hypothesis), given that t he simulation has produced u_i,

$$P(v_i|u_i) = \int \prod_{i=1}^{K} d\mu_i \mathcal{P}(v_i|\mu_i)\mathcal{P}(\mu_i|u_i). \tag{4.5}$$

Here $\mathcal{P}(v_i|\mu_i)$ is given by the Poisson distribution (4.2) and $\mathcal{P}(\mu_i|u_i)$ denotes the probability that the theoretical means are μ_i, given that the simulation has produced the u_i–histogram. $\mathcal{P}(\mu_i|u_i)$ is not known, we must infer it using the Bayes theorem,

$$P(\mu_i|u_i) = \frac{\mathcal{P}(u_i|\mu_i)\mathcal{P}(\mu_i)}{\int d\mu_i \mathcal{P}(u_i|\mu_i)\mathcal{P}(\mu_i)}, \tag{4.6}$$

where $\mathcal{P}(u_i|\mu_i)$ is the probability for each individual bin, given by the Poisson distribution (4.4), and $\mathcal{P}(\mu_i)$ is the prior for μ_i. Since $\mathcal{P}(u_i|\mu_i)$ is peaked around $u_i = \mu_i$, the dependence on the prior, $\mathcal{P}(\mu_i)$, is small, but nevertheless it is there. The simplest procedure here is to take a flat prior for $\mathcal{P}(\mu_i)$. Then the $\mathcal{P}(\mu_i)$ cancels in the numerator and the denominator of eq.(4.6) (the latter becomes simply 1), and we can identify

$$\mathcal{P}(\mu_i|u_i) \equiv \mathcal{P}(u_i|\mu_i). \tag{4.7}$$

Now eq.(4.5) reads

$$P(v_i|u_i) = \int \prod_{i=1}^{K} d\mu_i \frac{\mu_i^{v_i}}{v_i!} e^{-\mu_i} \frac{\mu_i^{u_i}}{u_i!} e^{-\mu_i} = \prod_{i=1}^{K} \frac{(u_i+v_i)!}{u_i!\ v_i!}\ \ 2^{-1-u_i-v_i}. \tag{4.8}$$

This formula represents our best estimate the likelihood, although is only valid when the non-statistical sources of uncertainty, both in the experimental and in the theoretical side, are ignored (they are incorporated in the next section). Note that expression (4.8) avoids the problem of the empty bins in the theoretical simulation. I.e. if one simply identified $\mu_i = u_i$, then the presence of an empty bin ($u_i = 0$) would make the whole likelihood - eq.(4.2)- vanishing. Therefore the $\mathcal{P}(\mu_i|u_i)$ piece in the calculation of the likelihood, eq.(4.5), is important, at least for bins with low statistics.

4.1.3 Separation of normalization and shape tests

Suppose for a moment we could calculate all μ_i with great accuracy and that we keep ignoring other sources of uncertainties different from the statistical ones. Then, the likelihood is simply given by the Poisson distribution $\mathcal{P}(v_i)$, as given by eq.(4.2).

Now, it is interesting that that expression can be separated in a test for the global normalization (the total number of events) and a test for the shape. Namely

$$P(v_i) = \prod_{i=1}^{K} \frac{\mu_i^{v_i}}{v_i!} e^{-\mu_i} = \mathcal{P}(\text{norm}) \times \mathcal{P}(\text{shape}), \tag{4.9}$$

where

$$\begin{aligned} \mathcal{P}(\text{norm}) &= \frac{\mu^v}{v!} e^{-\mu} \qquad \text{with } \mu = \sum_i \mu_i, \\ \mathcal{P}(\text{shape}) &= \prod_{i=1}^{K} \frac{\mu_i^{v_i}}{\mu^{v_i}} \frac{(v!)^{1/K}}{v_i!} = V \prod_{i=1}^{K} (\mu_i \frac{v}{\mu})^{v_i}, \end{aligned} \tag{4.10}$$

with

$$V = \frac{v!}{v^v} \prod_{i=1}^{K} \frac{1}{v_i!} = \text{const.} \tag{4.11}$$

Notice that both $\mathcal{P}(\text{norm})$, $\mathcal{P}(\text{shape})$ are proportional given by Poisson distributions. In particular, $\mathcal{P}(\text{shape})$ is proportional to a Poisson distribution, where the means μ_i are re-normalized so that they would fit perfectly the total number of events:

$$\mathcal{P}(\text{shape}) = V \prod_{i=1}^{K} (\mu_i \frac{v}{\mu})^{v_i} = (e^v \prod_{i=1}^{K} v_i!) V \prod_{i=1}^{K} \frac{(\mu_i \frac{v}{\mu})^{v_i}}{v_i!} e^{-\mu_i \frac{v}{\mu}} = \frac{v! e^v}{v^v} \prod_{i=1}^{K} \frac{(\mu_i \frac{v}{\mu})^{v_i}}{v_i!} e^{-\mu_i \frac{v}{\mu}}. \tag{4.12}$$

This expression is really independent of the global normalization. i.e. if we make $\mu_i \to a\mu_i$, then $\mathcal{P}(\text{shape})$ remains the same. This also tells us that if we fix the shape of a simulated histogram and allow to change its global normalization (i.e. we allow $\mu_i \to a\mu_i$), the total probability (4.9) is always maximal when the global mean, μ, coincides with the total number of events, v.

The interesting thing about separating normalization and shape tests is that one can treat the extra sources of uncertainty for both (mainly systematic errors) in a separate way, as will become clear in the next section. E.g. one may consider that the amount of systematic uncertainty in the global normalization is larger than in the shape, and hence it is useful to separate the two tests.

4.2 Incorporating other sources of uncertainty. Systematic errors

4.2.1 General strategy

There are several sources of uncertainty in the comparison of the experimental data with the theoretical predictions. First, there is the statistical uncertainty, associated to the Poisson distributions, which has been the subject of the previous section. Besides there are additional sources of -mainly systematic- uncertainty, both in the experimental side (calibration, ..., etc.) and in the theoretical one (K-factors, pdfs, etc.). However, for practical purposes, we can treat the experimental data as if they were free from systematic errors and "absorb" *all* the experimental systematic uncertainty in the theoretical side.

We will call μ_i^{th} the means that, in the simulation process, have produced the theoretical (u_i) histogram. Now, due to the systematic uncertainty, we *cannot* identify them directly with the "true means", μ_i, which are the real ones associated with the model under consideration, and thus the ones that, supposedly, have "produced" the experimental histogram (v_i) under the null-hypothesis. The relation between them can be expressed as

$$\mu_i(M) \;=\; F(M_i)\;\mu_i^{th}, \tag{4.13}$$

where $F(M)$ is some "transfer function" on the effective mass (M) that encodes all (experimental and theoretical) systematic uncertainties. This function can depend on a number of unknown parameters, though we know it cannot be completely arbitrary (below we give an ansatz for $F(M)$).

Now, in analogy with eq.(4.5), the best estimate for the likelihood is

$$P(v_i|u_i) = \int DF \int D\mu_i^{th}\; \mathcal{P}(v_i|\mu_i)\; \mathcal{P}(\mu_i^{th}|u_i)\; \mathcal{P}(\mu_i|\mu_i^{th}), \tag{4.14}$$

where $\mathcal{P}(\mu_i|\mu_i^{th}) \equiv \mathcal{P}(F)$ is still to be guessed, and the integration measure DF is written in a symbolic form. The first two factors in the integrand are statistical probabilities, as in (4.5). The third factor contains the systematic uncertainty (if we decided to ignore it, then we would simply take $\mathcal{P}(\mu_i|\mu_i^{th}) \equiv \delta(F-1)$).

We can write explicit expressions for the three factors in eq.(4.14). The first factor,$\mathcal{P}(v_i|\mu_i)$, is given by the Poisson distribution (4.2). Regarding the second factor, recall that (taking a flat prior for μ_i^{th}) we can identify

$$\mathcal{P}(\mu_i^{th}|u_i) \;\sim \mathcal{P}(u_i|\mu_i^{th}) \;=\; \prod_{i=1}^{K} \frac{(\mu_i^{th})^{u_i}}{u_i!} e^{-\mu_i^{th}}. \tag{4.15}$$

Finally, we have to make ansatz for the F function and its probability, $\mathcal{P}(\mu_i|\mu_i^{th}) \equiv \mathcal{P}(F)$. Since it is convenient to separate the uncertainties associated to the global normalization and to the shape, we express eq.(4.13) as

$$\mu_i \;=\; F(M_i)\;\mu_i^{th} \;=\; f\;g_i\;\mu_i^{th}. \tag{4.16}$$

Here f and g_i carry the uncertainty in the global normalization and in the shape, respectively. With this definition, g_i obey the relation

$$\sum_i g_i \; \mu_i^{th} \;=\; \sum_i \mu_i^{th} \;\equiv\; \mu^{th}, \tag{4.17}$$

i.e. the g_i parametrize systematic errors that modify the shape of the histogram without changing the total number of events. The situation $f = g_i = 1$ corresponds to the absence of systematic errors, but we have to assign a non-vanishing probability to the possibility that f, g_i depart from that ideal situation. Thus we write

$$\mathcal{P}(\mu_i|\mu_i^{th}) \equiv \mathcal{P}(f, g_i) = \mathcal{P}(f)\; \mathcal{P}(g). \tag{4.18}$$

For the moment we do not write a concrete ansatz for $\mathcal{P}(f)$, $\mathcal{P}(g)$ (this is postponed to the next subsection). So, the likelihood (4.14) is given by

$$\mathcal{P}(v_i|u_i) = \int D\mu_i^{th} \int Df Dg \;\left(\prod_{i=1}^{K} \frac{\mu_i^{v_i}}{v_i!} e^{-\mu_i}\right)\left(\prod_{i=1}^{K} \frac{(\mu_i^{th})^{u_i}}{u_i!} e^{-\mu_i^{th}}\right) \mathcal{P}(f)\mathcal{P}(g), \tag{4.19}$$

with $\mu_i = f g_i \mu_i^{th}$. In this expression Df, Dg are symbolic ways to express integration over all the possibilities for f, g_i.

4.2.2 Ansätze for the transfer functions

In eq.(4.16) we have written the "transfer" function, F, that encodes the systematic uncertainty, as

$$F(M_i) \;=\; f\; g_i, \tag{4.20}$$

but so far we have not established on which parameters the F-function –and thus the quantities f, g_i– depend. A simple and handy choice for practical purposes is to take the very values of $\{f, g_i\}$ as those independent parameters. Alternatively, since systematic errors must depend on M in a smooth way, we could parametrize $F(M)$ as a smooth function, e.g. $F \sim f\; \sum_\alpha P_\alpha$, where P_α are $\sim$ Legendre Polynomials and the summation contains just a few terms. Then, the F function would be defined by the a_α coefficients (together with the global normalization factor, f). This would be sensible, but it leads to very cumbersome expressions, difficult to handle. On the other hand, since in practice v_i and u_i are both quite smooth (apart from statistical noise), only sets of values of $F(M_i)$ that vary smoothly with M can lead to a simultaneous fit of both histograms. In other words, chaotic values of $F(M_i)$ (or, equivalently, g_i) will be strongly penalised by the $\mathcal{P}(v_i|\mu_i)$ piece (first factor in eq.(4.19)). So, even if those eccentric choices for g_i are not specially penalised by $\mathcal{P}(g)$, they are by other factors in the likelihood and become irrelevant. In consequence, choosing $\{f, g_i\}$ as independent parameters is a reasonable option

Concerning the integration measures, we could simply take $Df = df$, $Dg = \prod_i dg_i$. However, since $\{f, g_i\}$ are defined as multiplicative factors in eq.(4.16), it seems much more sensible

to use their magnitudes as the actual unknowns. This is equivalent to choose $\{\ln f, \ln g_i\}$ as the independent parameters. Then,

$$Df \equiv \frac{1}{f}\, df, \quad Dg \equiv \prod_{i=1}^{K} \frac{1}{g_i}\, dg_i. \tag{4.21}$$

Of course, since $\{f, g_i\}$ are never far from 1, it does not make a big difference to use $\{f, g_i\}$ or $\{\ln f, \ln g_i\}$, but it can be checked that the second option leads to a more stable and satisfactory test. Note that, in principle, the g_i variables are subject to condition (4.17), so there are in fact $K-1$ independent g_i variables. However, for the moment we have ignored such complication in writing (4.21).

Finally, concerning the probabilities $\mathcal{P}(f)$, $\mathcal{P}(g)$, we can take them as gaussians centered around $f = g_i = 1$. The argument of these gaussians must be essentially the "squared-distance" of $\{f, g_i\}$ to their central values, i.e. $\mathcal{P}(f) \sim \exp\{-\frac{1}{2}(f-1)^2\}$ and $\mathcal{P}(g) \sim \exp\{-\frac{1}{2}\int dM (g(M)-1)^2\} \sim \exp\{-\frac{1}{2}\sum_i (g_i-1)^2\}$. A nice fact here is that $\mathcal{P}(g)$ appears naturally factorized as $\prod_i \mathcal{P}(g_i)$, which is very convenient for analytical manipulations.

A suitable (and equivalent at first order), way to express these ansätze is by using the logarithmic variables, $\{\ln f, \ln g_i\}$:

$$\mathcal{P}(f) = \frac{1}{\sqrt{2\pi}\Delta_f} e^{-\frac{1}{2}\left(\frac{\ln f}{\Delta_f}\right)^2}, \tag{4.22}$$

$$\mathcal{P}(g) \propto \frac{1}{\Delta_g^K} e^{-\frac{1}{2}\sum_i \left(\frac{\ln\ g_i}{\Delta_g}\right)^2}, \tag{4.23}$$

where the widths Δ_f, Δ_g measure our degree of ignorance about the magnitude of f, g_i. Note that the use of logarithmic variables allows to maintain the whole range of integration of the gaussians, $[-\infty, \infty]$, without artificial cuts to keep $\{f, g_i\}$ positive.

In any case, we will go as far as possible in the analysis without specifying the precise ansätze for $\mathcal{P}(f)$, $\mathcal{P}(g)$.

4.2.3 Separation of normalization and shape tests

Coming back to our expression (4.19) for the likelihood, we note that the first factor of (4.19) may be decomposed, as in eqs.(4.9)-(4.10), into a factor for the global-normalization and another for the shape:

$$\left(\prod_{i=1}^{K} \frac{\mu_i^{v_i}}{v_i!} e^{-\mu_i}\right) \;=\; \frac{\mu^v}{v!} e^{-\mu}\; V \;\prod_{i=1}^{K} (\mu_i \frac{v}{\mu})^{v_i}, \tag{4.24}$$

where V is given in eq.(4.11). Since the *total* number of events is normally large the global-normalization factor can be approximated by a Dirac delta,

$$\frac{\mu^v}{v!} e^{-\mu} \;\simeq \delta(\mu - v) = \delta(f\mu^{th} - v) = \frac{1}{\mu^{th}} \delta(f - v/\mu^{th}). \tag{4.25}$$

Analogously, the second factor of (4.19) can be written as

$$\prod_{i=1}^{K} \frac{(\mu_i^{th})^{u_i}}{u_i!} e^{-\mu_i^{th}} \;\simeq\; \delta(\mu^{th} - u) \times U \prod_{i=1}^{K} (\mu_i^{th} \frac{u}{\mu^{th}})^{u_i}, \tag{4.26}$$

with

$$U = \frac{u!}{u^u} \prod_{i=1}^{K} \frac{1}{u_i!} \;=\; const. \tag{4.27}$$

We can use the presence of these deltas to extract pieces of the integrand of eq.(4.19) outside the sign of integration. Hence,

$$\mathcal{P}(v_i|u_i) \propto \mathcal{P}(f = \frac{v}{u}) \int D\mu_i^{th} Dg \left(\prod_{i=1}^{K} \frac{(\frac{v}{u} g_i \mu_i^{th})^{v_i}}{v_i!} e^{-\frac{v}{u} g_i \mu_i^{th}} \right) \left(\prod_{i=1}^{K} \frac{(\mu_i^{th})^{u_i}}{u_i!} e^{-\mu_i^{th}} \right) \mathcal{P}(g). \tag{4.28}$$

Note that we have made explicitely the integration in $\int Df = \int (1/f) df$, but not in $\int d\mu^{th}$. However the implicit presence of the $\delta(\mu^{th} - u)$ in the integrand, as expressed in eq.(4.26), has allowed us to replace $\mu^{th} \to u$ in a consistent way.

Assuming in the previous expression that Dg and $\mathcal{P}(g)$ are factorizable as products of K factors, like in eqs.(4.21), (4.23), makes much easier the integration in practice. We know that the g_i are subject to condition (4.17), so strictly speaking we only have $K - 1$ independent g_i variables and this factorization is not complete. In spite of this, assuming a complete factorization is a sensible and good approximation. The reason is the following. In eq.(4.28) the Poisson distribution in the first factor of the integrand only departs appreciably from zero when $\sum g_i \mu_i^{th} \simeq u \simeq \mu^{th}$. This can be checked by doing again a decomposition of such distribution as in (4.9, 4.10) and noting that the global-normalization piece of the decomposition is essentially a Dirac delta, $\delta(\sum \frac{v}{u} g_i \mu_i^{th} - v)$. So, even assuming that Dg and $\mathcal{P}(g)$ are factorizable, and thus integrating over sets of $\{g_1, g_2, ...g_K\}$ which do not respect (4.17), the Poisson distribution in the first factor of the integrand causes that only those sets that obey condition (4.17) will contribute appreciably to the integral. Alternatively one could understand this procedure considering that in the expression (4.16), the g_i variables that encode M_i-dependent systematic errors can also distort the total number of events. This is in fact a quite realistic situation. Then, the relation (4.17) is not to be imposed and Dg and $\mathcal{P}(g)$ become factorizable as products of K factors (strictly speaking). The trouble is that the previous separation between normalization and shape cannot be done exactly. But, if the g_i only amount to slight distorsions of the total normalization (in other words, $\mathcal{P}(g)$ penalizes much more severely the variation in the normalization than $\mathcal{P}(f)$) the separation (4.26) is a good approximation and (4.28) is valid.

Now, taking profit of the factorization of Dg and $\mathcal{P}(g)$ we can make explicitely the integration

in the μ_i^{th} variables, with no need of specifying the ansätze for $\mathcal{P}(f)$, $\mathcal{P}(g)$:

$$\begin{aligned}\mathcal{P}(v_i|u_i) \quad &\propto \quad \mathcal{P}(f=\frac{v}{u}) \\ &\times \quad \prod_{i=1}^{K}\left(\frac{(u_i+v_i)!}{u_i!v_i!}\int dg_i \frac{1}{g_i}\left(\frac{v}{u}g_i\right)^{v_i}\left(1+\frac{v}{u}g_i\right)^{-1-u_i-v_i}\mathcal{P}(g_i)\right). \qquad (4.29)\end{aligned}$$

In this expression, the factor of the first line, $\mathcal{P}(f=\frac{v}{u})$, carries the test for the global normalization: it is only sensitive to the mismatch between the experimental total number of events, v, and the theoretical one, u. The remaining factor (second line) corresponds to the test of the shape. It is funny to check that indeed, for given u, v, this expression has a maximum at $u_i = (u/v)v_i$.

Eq.(4.29) represents our final expression to evaluate the likelihood of a simulated histogram, u_i, confronted to the experimental one, v_i. (A modified version is given in eq.(5.11) of Appendix B to incorporate the fact that the luminosity of the simulated histogram may be different from that of the experimental one.) This expression amounts to realize K integrals, which can be done numerically at low cost in computing time, even if one needs to probe thousands or millions of histograms, corresponding to points in the parameter space of a theoretical scenario. All this is illustrated in the next section.

4.3 Application to the CMSSM

4.3.1 Set up

To show the power of the histogram-comparison technique, we will simulate LHC data assuming that nature lives in a standard benchmark SUSY model. This simulation will be considered as our (mock) experimental data. Then we will scan the CMSSM parameter space using Bayesian techniques to find out the most probable region of parameters, showing to which extent the histogram-comparison between the mock data and the theoretical prediction is capable to determine the "true" model.

To scan the CMSSM we use the set up described in chapter 1. In the likelihood we consider (besides the value of $M_Z^{\rm exp}$) the experimental bounds on the masses of supersymmetric particles and the lightest Higgs boson (see ref.[38] for details), and the mock LHC data of multijet events plus missing energy, which is the main focus of this section[1]. More precisely, for the sake of the simulated LHC data we work under the hypothesis that natu re lies in the so-called SU9 benchmark point, defined in ref. [41]. This is specified by the following values of the CMSSM parameters:

$$m = 300 \text{ GeV}, \quad M_{1/2} = 425 \text{ GeV}, \quad A = 20\ , \quad \tan\beta = 20, \quad \mu > 0. \qquad (4.30)$$

[1]Most of the electroweak precision tests are currently being surpassed by LHC data. Other observables, like $b \to s, \gamma$ would have a moderate impact in the analysis, but we want to focus on the impact of the LHC, which is already the dominant part of the likelihood. For the same reason we have not included the somewhat controversial $g-2$ data or Dark Matter constraints.

The corresponding values of the squark mass (first two generations) and the gluino mass are $m_{\tilde{q}} = 920$ GeV, $M_{\tilde{g}} = 994$ GeV. This point of the CMSSM parameter space is on the verge of being excluded by the last analyses by ATLAS and CMS [116, 117]. Of course, assuming the SU9 point is just an artifact to show the histogram-comparison technique at work, combined with Bayesian analysis. The LHC simulation has been performed using Pythia version 6.419 [118] with events generated at $E_{CM} = 14$ TeV, and selecting those satisfying the following cuts[2]:

- Three or more jets with $p^T > 30$ GeV and $|\eta| > 3.0$. The hardest with $p^T > 180$ GeV and $|\eta| > 1.7$, the second with $p^T > 110$ GeV .
- $p^T_{\rm miss} > 200$ GeV .
- $\Delta\phi_1 > 0.3,\ \ \Delta\phi_2 > 0.3,\ \ \Delta\phi_3 > 0.3$.
- $\sqrt{\Delta\phi_2^2 + (\pi - \Delta\phi_1)^2} > 0.5$, $\sqrt{\Delta\phi_1^2 + (\pi - \Delta\phi_2)^2} > 0.5$ and $\Delta\phi_2 > \pi/9$.
- $H_T = \sum_{i=2} p_i^T + p^T_{\rm miss} > 500$ GeV .

where $\Delta\phi_i := \Delta\phi({\rm jet}_i - p^T_{\rm miss})$. Concerning the luminosity we have considered 10^4 supersymmetric events, upon which we impose the previous cuts. Since the total cross section for SUSY production in the SU9 model is 2.4 pb, this corresponds to a luminosity of about 4.2 fb^{-1}. The histogram of number of events as a function of the effective mass, $M_{\rm eff}$, is displayed in Fig. 4.1, where the SM background has been extracted. For each event, $M_{\rm eff}$ is defined as [41]

$$M_{\rm eff} \ = \ \sum_j |p_j^T| + \not{p}^T . \tag{4.31}$$

where j runs over all jets satisfying the previous cuts. Note that this effectively implies a lower bound $M_{\rm eff} \geq 680$ GeV for the events considered.

Using the notation of sects. 4.1, 4.2, the bin contents of the "experimental" histogram of Fig. 4.1 are the v_i quantities. For each point scanned in the parameter space we compute a simulated histogram (the u_i quantities in sects. 4.1, 4.2), and then evaluate the likelihood through eq. (4.28). In order to use that expression, we have to specify the $\mathcal{P}(f)$, $\mathcal{P}(g)$ functions, which encode the systematic uncertainty assigned to the total number of events and the shape of the histograms respectively. In our case, we have used for them the gaussian profiles (4.22), (4.23). The width of the first, Δ_f, reflects the uncertainty in the total number of events due to (mainly) theoretical uncertainties associated with the K-factors and the parton distribution factions. We have been pretty conservative, assuming $\Delta_f = 0.5$; in other words we accept that a factor 2 or 1/2 in the total number of events is plausible[3]. On

[2]We have followed the strategy given in sect. 13.5 of ref. [40]

[3]In the present scan we are using a tree-level Pythia simulation of the parton events; which implies an important uncertainty about the K-factors. In a 1-loop-refined simulation this uncertainty could be assumed smaller.

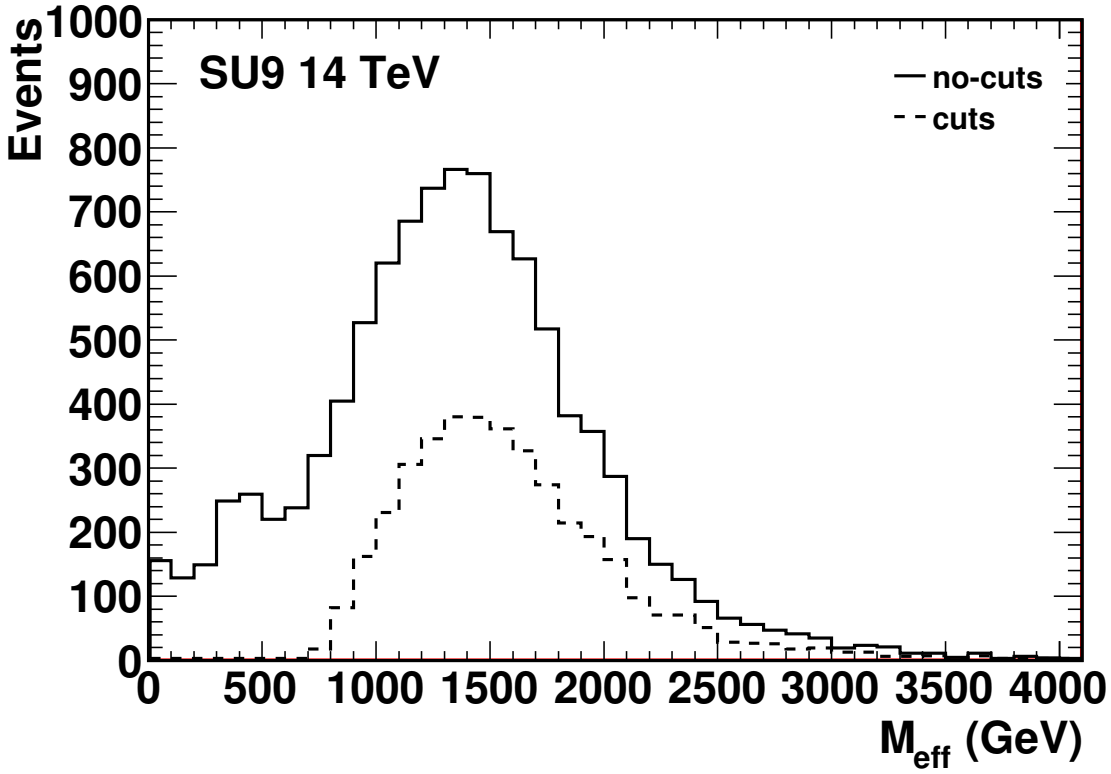

Figure 4.1: Effective mass distribution expected for the SUSY SU9 model without and with cuts applied as given in sect. 4.3

the other hand, from eq.(4.23), Δ_g goes as the typical systematic uncertainty in the shape times $\sqrt{K}$. E.g. if, *once* the uncertainty affecting the global normalization is extracted, one estimates that there remains a systematic uncertainty in the shape which is of order 10% for every bin, then o ne has to use $\Delta_g \simeq 0.1\sqrt{K}$ to reproduce that uncertainty at 1σ. In our case the total number of bins is $K = 10$, so we estimate $\Delta_g = 0.2$ as a reasonable choice.

To conclude our discussion of the likelihood, let us note that the total cross section varies from point to point in the CMSSM parameter space. This implies that considering 10^4 initial supersymmetric events does not correspond to the same luminosity for each point scanned in the parameter space. But of course, the comparison of histograms must be realized under the same conditions of luminosity. This feature can be easily incorporated into the histogram-comparison technique discussed in sect. 4.2, leading to a slight and straightforward correction in the expression (4.28) for the likelihood, which now becomes eq.(5.11) of Appendix B (see that appendix for more details). And that is the expression that we have finally used to compute the likelihood when scanning the CMSSM parameter space.

4.3.2 Results

We have computed the distribution of the posterior in the CMSSM parameter space using a modified version of the public `SuperBayeS` package [31][4] adopting MultiNest v2.8 [30, 107] as a scanning algorithm. We use as running parameters, namely the number of live points $n_{\rm live} = 2000$ and a tolerance parameter tol = 1. Our final inferences for each of the log and flat priors are obtained from chains generated with approximately 10^5 likelihood evaluations.

We have also included in our likelihood the limits on the lightest Higgs and SUSY masses provided by LEP and Tevatron. For details on the implementation see ref. [38].

For the marginalization procedure we have used $[0, M_X]$ as the range for m, $M_{1/2}$ and $|A|$. Besides, we have used $2 < \tan\beta < 62$. See a detailed discussion about this in sect. 1.5.

In order to show the potential of the histogram-comparison technique, we have performed the analysis twice: switching off and on the shape test.

Test for the total number of events

First we compute the (LHC-part of the) likelihood associated with a particular point in the CMSSM parameter space by comparing the prediction for the total number of supersymmetric events, satisfying the cuts specified in the previous subsection, with the "experimental" result for that number (i.e. after extracting the SM background).

This means that for each point examined we compute an histogram of events, u_i, but we just compare the total number, $u = \sum u_i$, with the experimental one, $v = \sum v_i$, using the first factor of eq.(4.28). More precisely, in order to incorporate the fact that the effective luminosity used in the simulation may change from point to point, we have actually used the –slightly modified– first factor of eq.(5.11),

$$\text{LHC} - \text{likelihood} \ \propto \mathcal{P}(f = \frac{v}{Lu}) \ , \tag{4.32}$$

where L is the quotient of the experimental luminosity and the luminosity of the simulation. We recall that the $\mathcal{P}(f)$ function carries all the uncertainties affecting the total number of events, except the purely statistical ones (which are subdominant when that number is large), and is given by the gaussian (4.22) with $\Delta_f = 0.5$; as explained in the previous subsection.

Fig. 4.2 (upper panels) shows the posterior pdf in the $M_{1/2} - m$ plane, after marginalizing the rest of the parameters: A, $\tan\beta$, together with the previosly marginalized μ and the nuisance SM parameters, $\{s\}$. As discussed below, the cross section for the kind of events considered (multijets + missing energy) is actually fairly insensitive to the values of A, $\tan\beta$, so the marginalization in these parameters does not change appreciably the probability distribution in the $M_{1/2} - m$ plane. The left (right) panel corresponds to log (flat) priors for the soft terms. The shape of these plots can be easily understood. Since we are fitting a unique quantity, namely the total number of events, and we have two parameters, $\{M_{1/2}, m\}$, we

[4]For this work, the public `SuperBayeS` code has been modified to interface with Pythia 6.419 [118].

can expect a degeneracy in the parameter space, which is in fact the case. The elongated shape of the allowed region, especially visible in the flat prior case, is in fact a widening –due to the uncertainties– of the l ine where the degeneracy is exact, which includes of course the "true model", i.e. SU9. This is marked with a red diamond in the plots. Note that for small gaugino mass, the squark masses become irrelevant, provided they are large enough, since in that case the dominant production are gluino pairs, whose masses do not depend on m. This is reflected in the vertical form of the region for small $M_{1/2}$.

Now, the shape of the line of degeneracy, somehow visible in the upper plots of Fig. 4.2, depends on the cuts used to select events. Note that, as the values of the soft terms get smaller, the cuts used to bound the energy of the first and second jets, and the missing transverse momentum (see previous subsection) become more and more inappropriate: many events with three or more jets plus missing momentum do not pass the cuts. As a consequence the counted total number of events of this kind is dramatically cut out and can become equal to the experimental one. This enhances artificially the statistical weight of the low energy region. As a result the maximum value of the pdf, and its averaged central value (marked by a green dot), are shifted from the "true model" (marked by a red diamond).

There are ways to counteract these nasty effects. Playing with different cuts, the degeneracy gets partially broken and it is possible to discard larger regions of the parameter space. For instance, one can compare the total number of events using several choices for the lower bound on $M_{\rm eff}$. Somehow, this equivales to test the shape of the experimental and theoretical histograms, but not in the most efficient way. This is improved using the histogram-comparison technique explained in sections 4.1, 4.2, which we will apply shortly to this analysis.

Fig. 4.2 (lower panels) shows the posterior in the $\tan\beta - A$ plane, after marginalizing the rest of the parameters. As mentioned above, the cross section of the type of events considered does not depend appreciably on A and $\tan\beta$, and this is reflected in the plots. The preference for rather small values of both A and $\tan\beta$ is essentially a consequence of the Jacobian factor (1.32) in the posterior. The Jacobian automatically penalizes regions of the parameter space where fine-tuning is needed to reproduce the electroweak scale. This disfavors large values for both A and $\tan\beta$.[5] The remarkable insensitivity to A and $\tan\beta$ is physically due to the fact that the CMSSM spectrum is not much dependent on the the values of A and $\tan\beta$, except for the mixing effects in the mass matrices of stops (and sbottoms and stau for large $\tan\beta$), charginos and neutralinos. Even for these matrices the effect is normally quite small. Thus the production rates of squarks and gluinos are quite independent of A and $\tan\beta$. Once the supersymmetric particles are created, their decay rates are not very relevant for the cross section of the process considered (multijets + missing momentum), and, in any case, they are quite independent of these parameters too. This insensitivity to A and $\tan\beta$ could be partially cured by complementing the present analysis by a separated study of those events involving leptons [119], but that discussion is outside the scope of this work.

[5]This is an statistical effect which is not visible in frequentist approaches, where the basic quantity is the likelihood and fine-tuning is not penalized, unless such penalization is artificially incorporated.

Incorporation of the shape test

Now we repeat the analysis, but computing the likelihood associated with the LHC data with the use of the whole expression (4.28), which takes into account not only the total number of events, but also the comparison of the histogram shapes. Again, in order to incorporate the fact that the luminosity of the simulation changes from point to point in the parameter space we use the modified formula (5.11):

$$\begin{aligned} \text{LHC}-\text{likelihood} \quad &\propto \quad \mathcal{P}(f=\frac{v}{Lu}) \\ &\times \quad \prod_{i=1}^{K}\left(\frac{(u_i+v_i)!}{u_i!v_i!}\int dg_i \frac{1}{g_i}\left(\frac{v}{u}g_i\right)^{v_i}\left(1+\frac{v}{u}g_i\right)^{-1-u_i-v_i}\mathcal{P}(g_i)\right) \end{aligned} \tag{4.33}$$

We recall that $\mathcal{P}(g)$ carries all the systematic uncertainties affecting the shape of the histograms, and is given by the gaussian (4.23) with $\Delta_g = 0.2$; as explained in the previous subsection. Note that, as could be expected, the correction due to the difference in luminosity does not affect the shape-part of the likelihood.

Fig. 4.3 is as Fig. 4.2, but after including the likelihood associated to the shape in the analysis. The upper panels show, for log and flat priors, the posterior pdf in the $M_{1/2} - m$ plane, after marginalizing the rest of the parameters. As expected, the test of the theory is now much more efficient and the previous degeneracies dissapear (note the different ranges of the axes of the two figures). This illustrates the potential of making use of all the information contained in the theoretical and experimental histograms when computing the likelihood of a model, provided the various sources of uncertainty are properly taken into account.

The lower panels of Fig. 4.3 show the posterior in the $\tan\beta - A$ plane. Again, the cross section of the type of events considered does not depend appreciably on these parameters, which is reflected in the plots. Still, introducing the test for the shape slightly improves the sensitivity of the search to the values of A and $\tan\beta$, but that sensitivity is anyway very small.

Figs. 4.4 and 4.5 show, for logarithmic and flat priors, the unidimensional posteriors for m, $M_{1/2}$, A and $\tan\beta$, after marginalization of all the parameters, except the one plotted in each graph. The shape of these functions reflects the previous discussion. It is worth-noticing the great precision in the determination of the gaugino mass, which comes from the fact that, due to the renormalization group running, $M_{1/2}$ is the parameter that dominantly determines the low-energy spectrum of the CMSSM.

Finally, we note that the posteriors have in all cases a very slightly dependence on the type of prior used, reflecting the robustness of the approach.

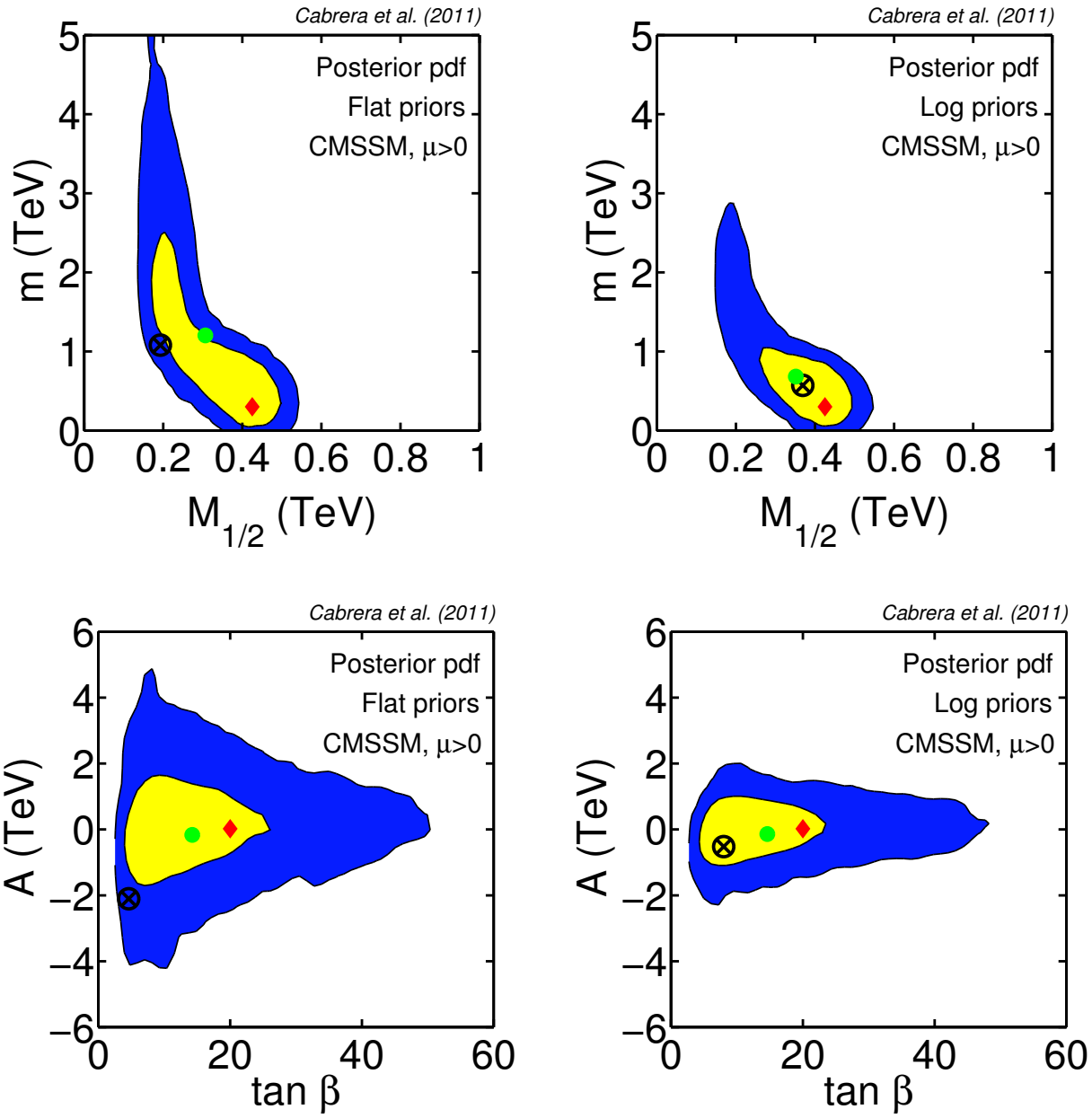

Figure 4.2: 2D marginalized posterior probability distribution for flat (left panels) and logarithmic (right panels) priors using the normalization test. The inner and outer contours enclose respective 68% and 95% joint regions. The small filled circle represents the mean value of the posterior pdf, the cross corresponds to the best-fit point and the diamond to the SU9 model.

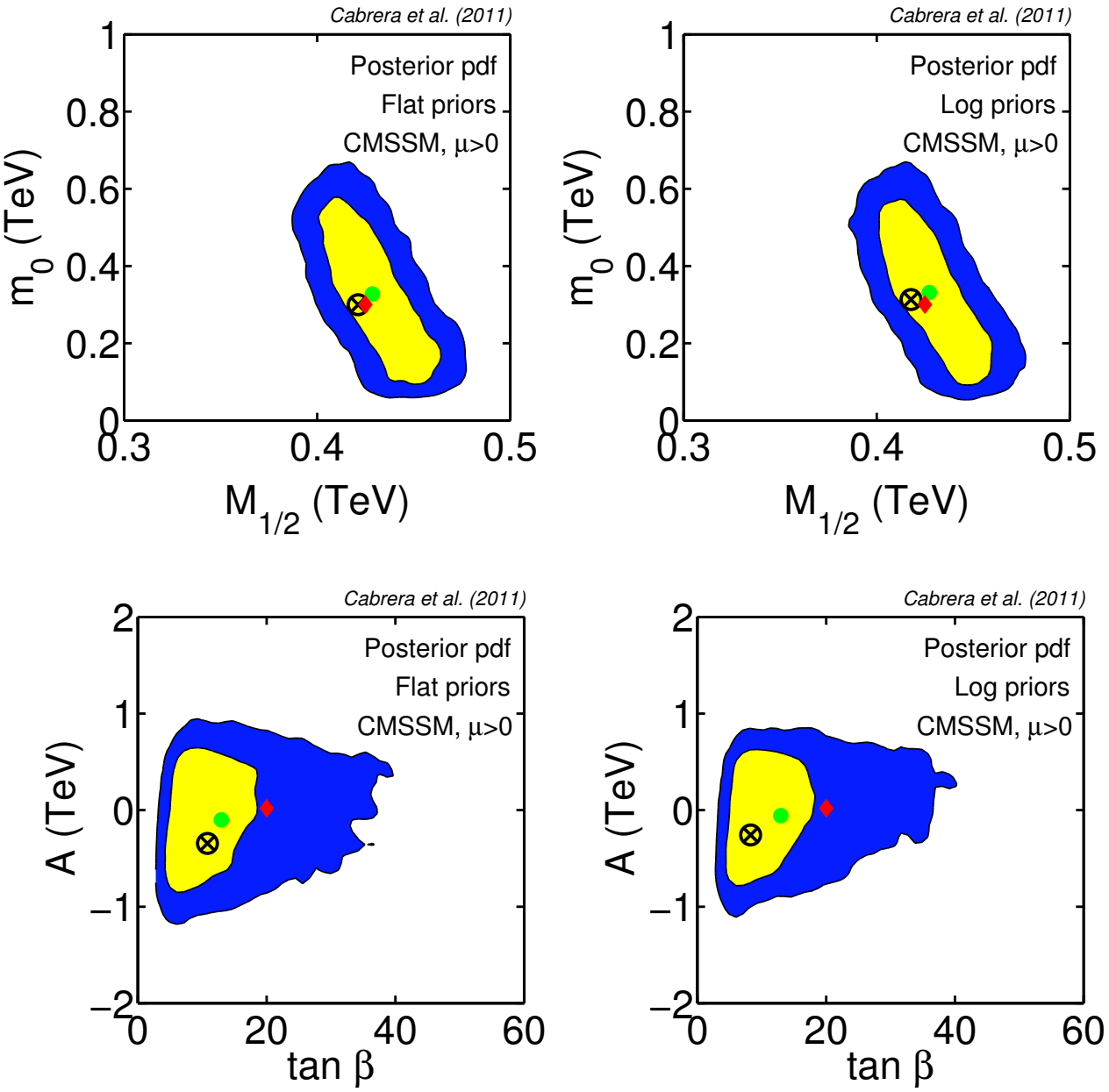

Figure 4.3: 2D marginalized posterior probability distribution for flat (left panels) and logarithmic (right panels) priors using both normalization and shape tests. The inner and outer contours enclose respective 68% and 95% joint regions. The small filled circle represents the mean value of the posterior pdf, the cross corresponds to the best-fit point and the diamond to the SU9 model.

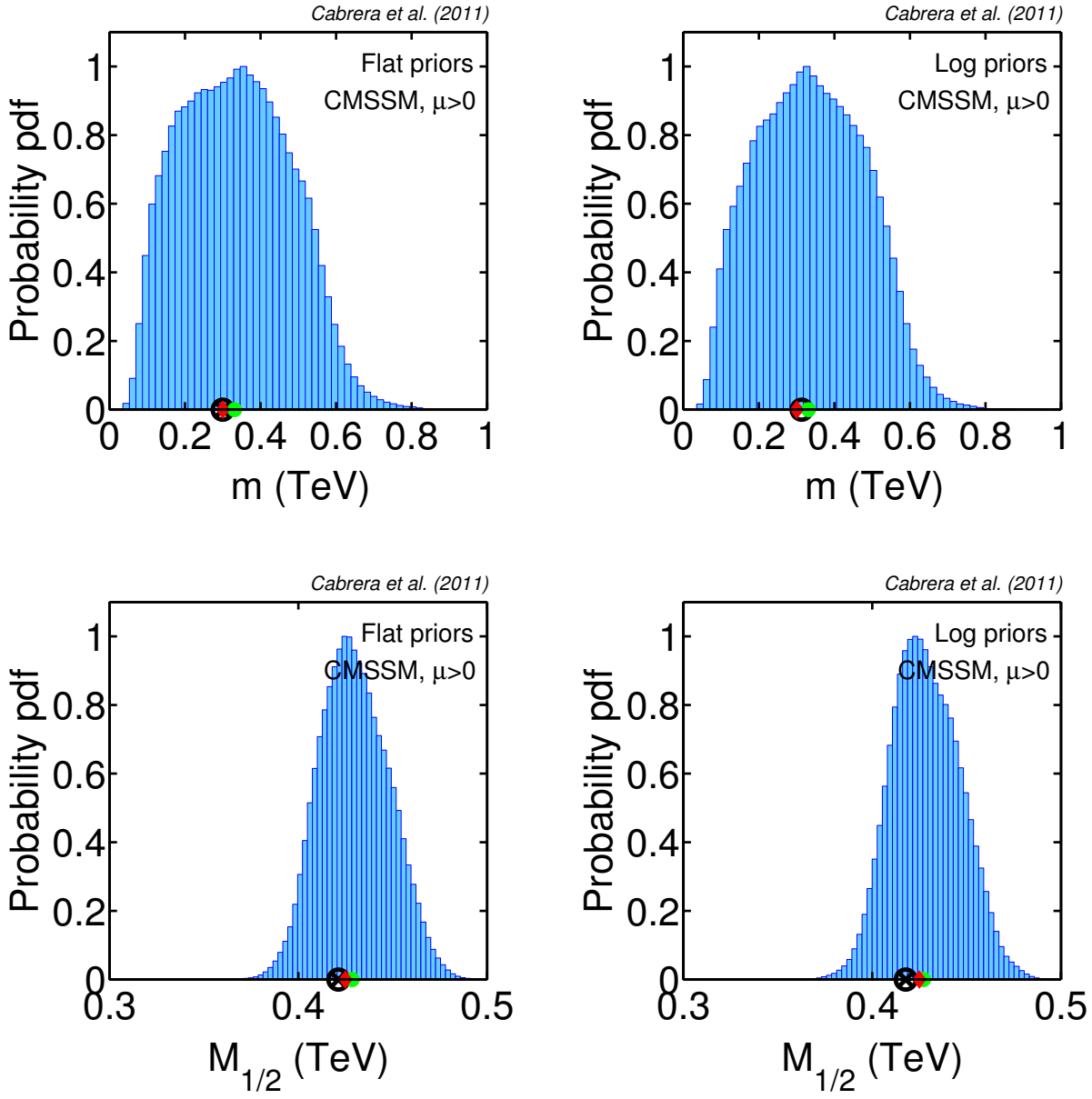

Figure 4.4: 1D marginalized posterior probability distribution of the m and $M_{1/2}$ parameters (upper and lower panels respectively) for flat (left panels) and logarithmic (right panels). The small filled circle represents the mean value of the posterior pdf, the cross corresponds to the best-fit point and the diamond to the SU9 model.

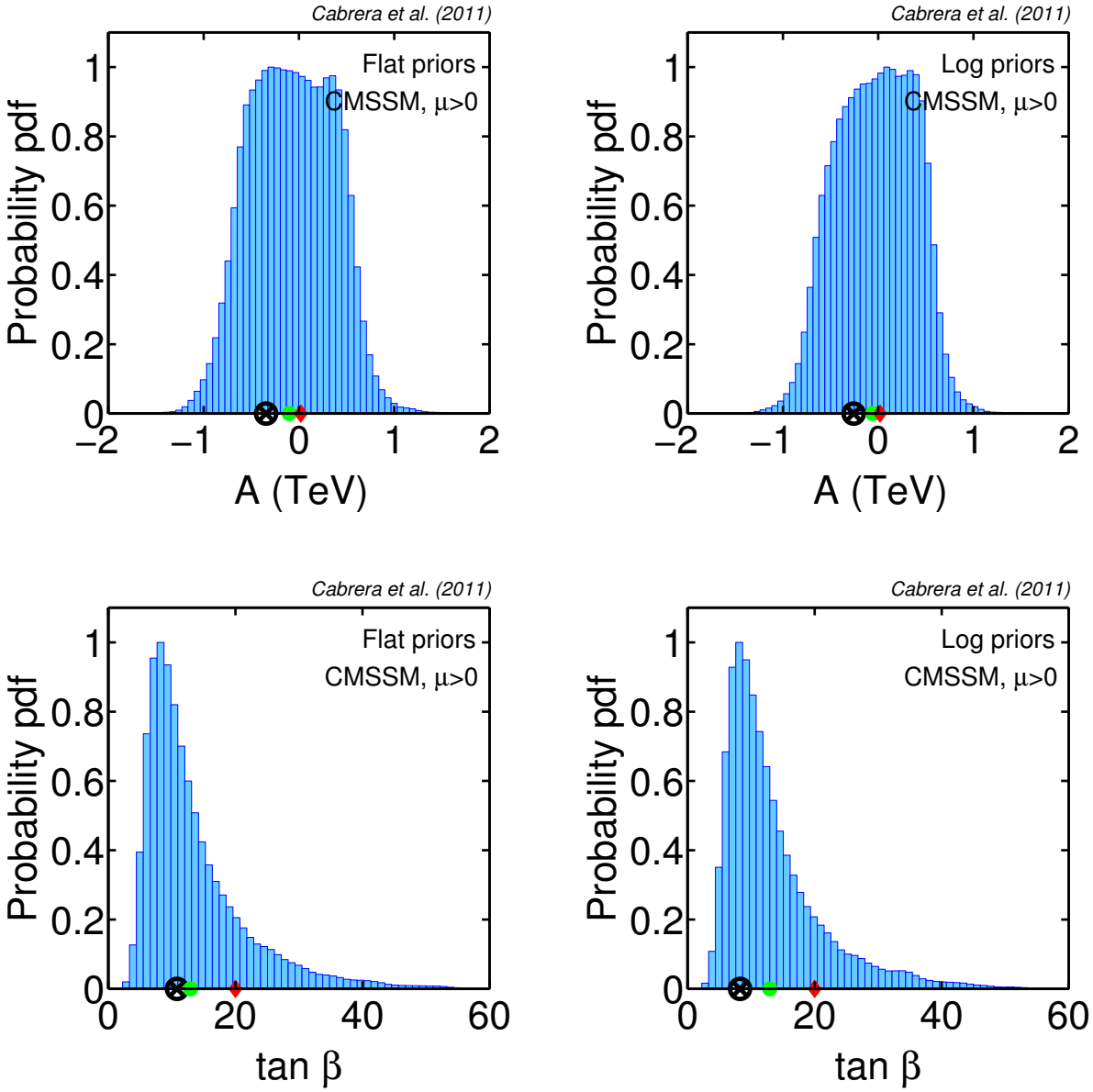

Figure 4.5: 1D marginalized posterior probability distribution of the A and $tan\beta$ parameters (upper and lower panels respectively) for flat (left panels) and logarithmic (right panels). The small filled circle represents the mean value of the posterior pdf, the cross corresponds to the best-fit point and the diamond to the SU9 model.

Chapter 5

Upper bounds on SUSY masses from the LHC

In the previous chapter we showed how powerful could be the Bayesian analysis to decode future new physics signals from LHC. In this chapter we depart from the use of Bayesian tools and explore an indirect, but powerful, way to put *upper* bounds on supersymmetric particles using the recent LHC bounds on the Higgs mass.

The experimental bounds on the Higgs mass are rapidly changing. Besides the LEP lower bound, $m_h > 114.4$ GeV [78], LHC has recently extended the 95% CL *excluded* region around $2M_W$ to 149 GeV $< m_h <$ 190 GeV, and has excluded a new range at 295 GeV $< m_h <$ 450 GeV [120]. For sure we are likely to see stronger limits (and hopefully a discovery) as the LHC luminosity keeps growing.

These bounds put constrains on the parameter space of the Standard Model: they directly translate into bounds on the self-coupling of the Higgs. At tree-level the relation reads $m_h^2 = 2\lambda_{\rm tree} v^2$, where λ is the SM Higgs quartic-self-coupling and $v = 174.1$ GeV is the Higgs expectation value. This is not especially exciting per-se since there is no particular prediction for that coupling in the pure SM. On the other hand, there are models where the value of the Higgs self-coupling is related to other parameters of the theory; hence a bound on m_h can put constrains on those parameters. One classic example is supersymmetry (SUSY) where, in the minimal model (MSSM), $\lambda_{\rm tree} = \frac{1}{4}(g^2 + g'^2)\cos^2 2\beta$. Here g, g' are the $SU(2) \times U(1)$ gauge couplings and $\tan\beta = \langle H_u\rangle/\langle H_d\rangle$, i.e. the ratio of expectation values of the two MSSM Higgs fields. This relation means, in particular, that at tree-level the mass of the Higgs in the MSSM is bounded by the mass of the Z-boson (91.1 GeV). As it is well-known, radiative corrections increase m_h, which can then get compatible with its experimental (LEP) lower bound, at the expense of requiring a relatively heavy spectrum ($\gtrsim 1$ TeV) of superpartners, which in turn introduces some fine tuning. Much work has been devoted to this important feature of the MSSM [92, 80, 93, 94, 95, 96, 97, 121, 122].

Our approach here is *opposite* and complementary: using the *upper* bound on the Higgs

mass to put an upper bound on the masses of supersymmetric particles.

It is common to hear that "SUSY cannot be ruled out", meaning that one can always increase the masses of superpartners to avoid its discovery at the LHC or any conceivable experiment. However, for the above-mentioned reasons, in the MSSM the Higgs mass cannot be arbitrarily large. Actually, the radiative correction to m_h^2 depends logarithmically on the SUSY masses (principally on stop masses). This is easy to understand. The MSSM tree-level relation, $\lambda_{\rm tree} = \frac{1}{4}(g^2+g'^2)\cos^2 2\beta$, breaks down at the threshold scale where supersymmetric particles become decoupled. Below that SUSY-threshold, λ runs down to the electroweak scale following the SM renormalization group (RG) equation. The SUSY-threshold scale is essentially given by (an average of the two) stop masses, since, in the one-loop effective potential they are responsible for the largest contribution to the Higgs quartic-coupling to be matched with the SM-effective-theory (for details see e.g. [97]). Hence, the enhancement of λ, and thus of m_h^2, goes logarithmically with the ratio of the SUSY-threshold scale to the electroweak scale.

Since the value of λ at the SUSY-threshold scale is always perturbative, the Higgs mass in the MSSM necessarily obeys the SM perturbativity upper bound. For the extreme case when the supersymmetric masses are as large as M_P, and so is the threshold of new physics, this bound reads $m_h \lesssim 180$ GeV [123], which is already overtaken by the recent LHC exclusion ranges on m_h quoted above. In other words, for the MSSM the only relevant experimentally allowed range for m_h is

$$114.4 \text{ GeV} < m_h < 149 \text{ GeV} \ . \tag{5.1}$$

These upper and lower bounds on m_h translate, respectively, into upper and lower bounds on the masses of the supersymmetric particles. The latter can be complemented with the recent direct LHC bounds on the size of the supersymmetric masses [124, 125], giving the window of scales where the MSSM can live. In this chapter we are going to show explicitly this window, describing how it will evolve as the LHC improves the limits on the mass of the Higgs or discovers its existence. We will see that the upper bounds on the MSSM scale are not yet of physical significance, but they will get much stronger in the near future.

We have evaluated the Higgs mass, starting with the tree-level value of λ at SUSY-threshold scale, corrected with 1-loop threshold corrections

$$\lambda(M_{\rm SUSY}) = \frac{1}{4}(g^2+g'^2)\cos^2 2\beta + \frac{3}{16\pi^2}\frac{m_t^4}{v^4}X_t \tag{5.2}$$

with

$$X_t = \frac{2(A_t - \mu\cot\beta)^2}{M_{\rm SUSY}^2}\left(1 - \frac{(A_t-\mu\cot\beta)^2}{12M_{\rm SUSY}^2}\right) \tag{5.3}$$

in where m_t is the running top mass corresponding to a pole mass $M_t = 173.1 \pm 1.253$ GeV [126]; A_t is the trilinear scalar coupling, μ is the mass parameter for the Higgses in the superpotential; and $M_{\rm SUSY}$ represents a certain average of the stop masses [97, 121, 122].

Note that, unless the combination $|A_t - \mu \cot\beta|$ becomes larger than $\sqrt{12} M_{\rm SUSY}$, which is very odd (and almost in coflict with charge and color breaking bounds), the value of X_t is in the range $0 \leq X_t \leq 6$.

The above value of λ has to be run down to the electroweak scale, say $Q_{\rm ew}$. Then the Higgs mass is given by

$$m_h \simeq 2\lambda(Q_{\rm ew}) v^2 \tag{5.4}$$

This relation gets one-loop radiative corrections (in the effective SM theory), which are rendered negligible by choosing appropriately $Q_{\rm ew}$. A nearly optimal choice is $Q_{\rm ew} = M_t$ [97]. Finally, to get the pole mass, M_h, one has to add (pretty small) radiative corrections. We have performed the previously-sketched calculation of m_h using the 2-loop SM RG equation of λ, which is coupled to the RG equations of the other SM parameters, in especial the top Yukawa coupling, y_t, and the strong coupling, α_3. The complete set of RG equations can be found e.g. in ref.[121, 127].

Let us mention that in the literature there are approximate analytic formulae for the Higgs mass, obtained by approximating the running of λ at a certain order in the leading log expansion [121, 122]. The first terms of those formulae read

$$m_h^2 \simeq M_Z^2 \cos^2 2\beta + \frac{3}{4\pi^2}\frac{m_t^4}{v^2}\left[\frac{1}{2}X_t + \log\frac{M_{\rm SUSY}^2}{M_t^2}\right] + \cdots \tag{5.5}$$

These approximations are only valid for $M_{\rm SUSY} \lesssim 1$ TeV, so they are *not* applicable to our problem. However, eq.(5.5) is useful to qualitatively understand the numerical results, so we have shown it explicitly.

Now we are ready to obtain upper bounds on $M_{\rm SUSY}$ as a function of $\tan\beta$ and the upper bound on m_h. It should be kept in mind that $M_{\rm SUSY}$ essentially stands for "stop masses". Indeed, in the usual MSSM scenarios the masses of all supersymmetric particles are of the same order (say, within a factor of 10 or less). Therefore, with this caveat, the following results are valid for essentially any MSSM model. A notable exception are split-SUSY models [128, 129], where the masses of scalar superpartners are very high but the gauginos and higgsinos are still relatively light. Actually, the results for split-SUSY are analogous to those of the ordinary MSSM, but we will discuss them separately.

Fig. 5.1 shows the bands of constant Higgs mass in the $\tan\beta - M_{\rm SUSY}$ plane. The width of the bands comes from the various sources of uncertainty, to be discussed shortly. The behaviour shown in Fig. 5.1 can be *qualitatively* understood from the approximate expression (5.5). In particular, the larger $\tan\beta$ the bigger the first (tree-level) term in Eq.(5.5) becomes, and thus a smaller value of $M_{\rm SUSY}$ is required to reproduce m_h. The width of each band (darker part) has been obtained by varying X_t within its range, $0 \leq X_t \leq 6$. Note that this uncertainty arises from our ignorance about the values of the remaining MSSM parameters. On top of that uncertainty, we have added the error coming from experimental uncertainties in the theoretical computation of m_h, resulting in the wider lighter bands. The experimental uncertainty is dominated by the one in the top mass $M_t = 173.1 \pm 1.25$ GeV. Additional

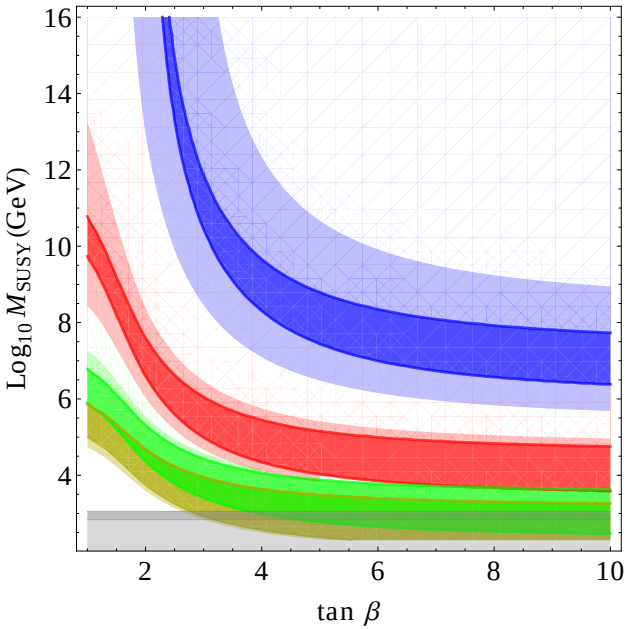

Figure 5.1: Bands of constant Higgs mass in the $\tan\beta - M_{SUSY}$ plane for the MSSM. From top to bottom $m_h = 140, 130, 120, 115$ GeV. The low (gray) horizontal band stands for the direct LHC lower bounds on M_{SUSY} (see text).

sources of experimental error, such as the one in α_3, are negligible when added in quadrature. Let us remark that we have *not* added in quadrature the uncertainties coming from the ignorance about X_t and M_t, but in a direct way (to avoid statistical inconsistencies); thus the light band represents an overestimate of the total error. On the other hand, there is an intrinsic theoretical error coming from the higher-loop effects not considered in the computation of m_h, which we estimate in ~ 2 GeV [97, 121, 122]. This is negligible when added in quadrature to the other sources of error.

Now, each band of Fig. 5.1 represents the future upper bound on M_{SUSY}, as soon as the upper bound on m_h reaches the corresponding value. The present relevant bound, $m_h < 149$ GeV, does not constrain the MSSM parameter space in a significant way. Actually, the corresponding band is outside Fig. 5.1. But it is clear from the figure that as soon as new LHC bounds on m_h are reported, the MSSM parameter space will become significantly cornered from above. Note also that the $m_h = 115$ GeV band corresponds to the present *lower* bound on M_{SUSY}. So in Fig. 5.1 we see the future evolution of the MSSM window. On the other hand, if LHC discovers the Higgs, say at $m_h = 130$ GeV, the associated band in Fig. 5.1 gives the allowed region of the MSSM parameter space.

We have complemented the latter lower bound on M_{SUSY} with the direct lower bounds that LHC has already put on the MSSM parameter space [124, 125]. They translate into the grey band at the bottom of Fig. 5.1, which has been obtained as follows. LHC bounds on the MSSM are presented by ATLAS and CMS as exclusion regions in the constrained

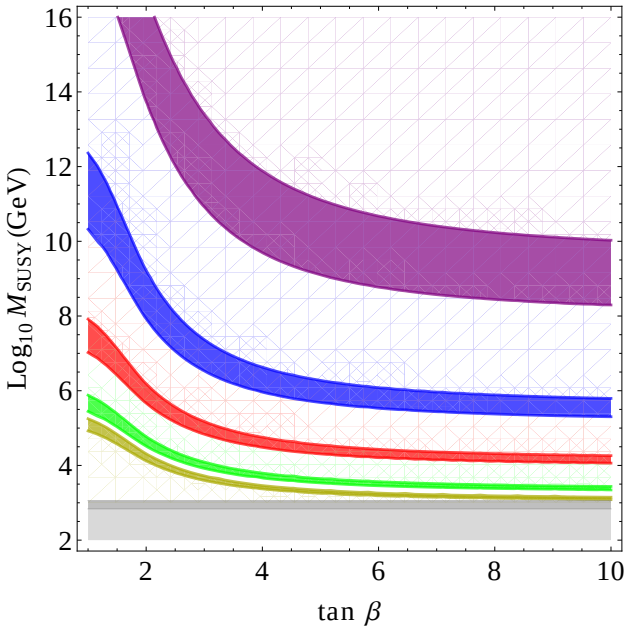

Figure 5.2: The same as Fig. 5.1 but for split-SUSY. From top to bottom $m_h = 150, 140, 130, 120, 115$ GeV.

MSSM (CMSSM) parameter space, As explained in sect. 1.2, the CMSSM is characterised by universality of the soft terms at M_X. We have extracted the data from the ATLAS exclusion plot with `EasyNData` [130] and then calculated the exclusion contour for the gluino mass, $M_{\tilde{g}}$, versus M_{SUSY} with the `SoftSUSY` package [8]. It turns out that M_{SUSY} must be larger than 750-1000 GeV (the precise value depends on the value of $M_{\tilde{g}}$). This uncertainty in the bound is reflected in the narrow darker grey strip on top of the light one. Note that, strictly, the grey band corresponds to a lower bound for the CMSSM case; but one does not expects big changes in ordinary MSSM models.

One could take the attitude of only considering low M_{SUSY}, say $M_{\mathrm{SUSY}} \lesssim \mathcal{O}(\mathrm{TeV})$, as reasonable, in order to avoid fine-tuning to get the correct electroweak scale. However, it has been suggested that in a landscape scenario such fine-tuning can be largely compensated by the overabundance of vacua with SUSY broken at a high scale, in which the anthropic principle would operate, see e.g. ref. [43]. As we have seen, this kind of scenario is going to be tested by LHC very soon. The split-SUSY framework [128, 129], which we are going to discuss next, is in fact a popular variant of the above-mentioned landscape scenario.

In split-SUSY one supposes that the spectrum of superpartners is split, having scalars at higher scales whereas gauginos and higgsinos remain light, usually thanks to an approximate R-symmetry. In this way, split-SUSY is kept consistent with the perturbative unification of gauge couplings at M_X and still contains Dark Matter candidates. The approximate R-symmetry is also responsible to keep $A-$ and $\mu-$parameters small, something necessary for the radiative stability of the scenario. So in split-SUSY there are two well-separated

SUSY thresholds (thus the name).The prediction for m_h in split-SUSY is done following a similar approach to the one previously discussed and applied for the MSSM. This prediction was already considered in the second seminal split-SUSY paper of ref. [128, 129]. We have updated that calculation (using e.g. upgraded information about the top mass) and studied in a systematic way the bounds on $M_{\rm SUSY}$ depending the evolution of the upper bound on m_h. The main difference with respect to the MSSM case is that in split-SUSY the gluinos remain active and contribute significantly to the RG equations between the upper and the lower SUSY-thresholds.

The results are shown in Fig. 5.2, which is analogous to the previous Fig. 5.1 for the MSSM. Due to the smallness of A_t and μ (and thus of X_t), the threshold correction for λ at the high SUSY threshold is negligible, see Eq. (5.2). In consequence, the width of the bands in Fig. 5.2 is only due to the experimental error on M_t. Note also that $M_{\rm SUSY}$ corresponds to the upper SUSY-threshold ($\sim$ stop masses), and therefore the plotted upper bounds are absolute upper bounds for SUSY in the split-SUSY scenario. In contrast to the MSSM case, there is a region of the parameter space, large $\tan\beta$ and very heavy masses, that is already excluded with today's bound on the Higgs mass; yielding $M_{\rm SUSY} < 10^{11}$ GeV. As soon as the upper bound on m_h reaches 140 GeV, the exclusion will hold for any $\tan\beta$. This is most relevant for split-SUSY, since in this scenario we *do expect* the upper SUSY-threshold to lie at very high energy; typically as large as M_X or M_p, something which is going to be probed very soon by LHC.

Summary and Conclusions

The work presented here has been realized in the context of the search of new physics at the LHC. Most of it is related with the possible existence and detection of Supersymmetry (SUSY), but some results are applicable to any scenario of new physics. Likewise, most of this work, but not all, has been performed with the help of Bayesian techniques. One of our key ideas has been to exploit all the available (theoretical and experimental) wisdom about the theory examined in order to establish the corresponding distribution of probability in the associated parameter space. For this goal Bayesian inference is the appropriate tool, specially when the experimental data are not powerful enough to select a small region in that parameter space.

We have first proposed an improved Bayesian analysis framework to explore the MSSM, careful handling of the various pieces of information. The main results and conclusions here are the following:

- We have found that a penalization of fine-tuning arises in an automatic way when performing a correct Bayesian analysis. Besides, the resulting penalization is very similar to (though not exactly the same as) the standard measures of fine-tuning, in particular the Barbieri-Giudice one. This result is in fact completely general and goes beyond the use of supersymmetry as the possible new physics beyond the Standard Model.

- We have done a rigorous treatment of the nuisance variables, in particular Yukawa couplings. We have shown that the easiest and usual practice of taking the Yukawas "as required", approximately corresponds to taking logarithmically flat priors in the Yukawa couplings, which on the other hand is not an unreasonable choice.

- The use of an efficient (and actually quite common) set of variables to scan the MSSM parameter space, $\{m_t, m_b; m, M, A, \tan\beta, \text{sgn }\mu\}$, implies to inherit a Jacobian factor (1.21) which contains inside the above-mentioned penalization of fine-tuned regions. A quite accurate analytical expression for it is given (1.33), which is valid for any MSSM. 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 sensible (flat and logarithmic) priors in the initial variables, using the 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.

These improvements make the approach much more rigorous and, besides, allow 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).

We have the applied this framework to the realization of a forecast of the MSSM for the LHC. We have focussed in its constrained version (CMSSM). Our main results here are the following:

- Including 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, the favoured region of the MSSM parameter space lies at low-energy, but there is a significant portion out of the LHC reach.
- Adding the information about a_μ, using $e^+e^- \to$ had data, the soft terms are dramatically pushed into the low-energy region (for $\mu > 0$), well inside the LHC reach. This does not happen using τ-decay data instead $e^+e^- \to$ had, warning us to be cautious about strong conclusions from g-2.
- Roughly speaking, including dark matter 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.

Bayesian techniques are also useful to quantify present or potential tensions between observables within a specific model. In this way we have pointed out that a "large" Higgs mass (i.e. not close to the present lower bound) is rather incompatible with the $g-2$ observable, if the latter is evaluated using $e^+e^- \to$ had data. We have quantified this tension as a function of the (yet unknown) Higgs mass, showing that

- For $m_h \geq 125$ GeV the maximum level of discrepancy ($\sim 3.2\ \sigma$) is already achieved, indicating that SUSY has decoupled, and thus the prediction for a_μ coincides with the SM one.
- Given present-day data, requiring less than a 3 σ discrepancy, implies $m_h \lesssim 120$ GeV.

Due to the complexity of the LHC experiment, much of the comparison between LHC data and theoretical predictions has to be made by confronting experimental histograms (in different variables) and theoretical histograms produced by simulations. In this sense, we have presented a rigorous and effective method to compare experimental and theoretical histograms, evaluating the total likelihood, and applied it to a physically relevant case. In doing this we have taken into account that, besides the statistical uncertainties inherent to the histograms, there are additional sources of systematic error.

The procedure can be easily incorporated to both frequentist and Bayesian analyses, since both are based on the likelihood of the theoretical models. In the two approaches, incorporating the total likelihood optimizes the chances of picking up a signal of new physics and, once the signal is found, identifying which new physics is behind. E.g. if the new physics is supersymmetry, it allows to find in an optimal way the parameters of the supersymmetric model.

We have illustrated the latter point by showing how a search in the CMSSM parameter space, using Bayesian techniques, can effectively find the correct values of the CMSSM parameters by comparing histograms of events with multijets + missing energy displayed in the effective-mass variable. The procedure is in fact very efficient to identify the true supersymmetric model, in the case supersymmetry is really there and accessible to the LHC. But, of course, the technique can be applied to any scenario of new physics.

Finally, we have taken into account the relevant experimental results that the LHC is already giving. In particular, we have used the current and forthcoming LHC upper bounds on the Higgs mass to put upper bounds on supersymmetric masses, $M_{\rm SUSY}$; using the fact that in the MSSM the quartic Higgs coupling, and therefore the Higgs mass, is a function of the SUSY masses (in particular stop masses). Right now there is no significant constraint on the parameter space of the MSSM but very soon there will be, as can be seen in Fig. 5.1. On the other hand, for split-SUSY a significant part of the parameter space can already be excluded on these grounds, as can be seen in Fig. 5.2. As the LHC continues to improve the bounds on m_h the allowed region will be more and more shrunk, therefore showing that, even if SUSY is not found, it can not be hidden way. Eventually (and hopefully) the Higgs will be discovered, and one can use these results to establish the region of the $M_{\rm SUSY} - \tan\beta$ plane consistent with the actual value of m_h.

Resumen y Conclusiones

El trabajo presentado ha sido realizado en el contexto de la bsqueda de nueva fsica en el LHC. La mayor parte del est realacionada con la posible deteccin de Supersimetra, pero algunos resultados son aplicables a cualquier escenario de nueva fsica. Asimismo, la mayor parte de este trabajo, pero no toda, se ha realizado con la ayuda de tcnicas Bayesianas. Una de nuestras ideas claves ha sido explotar toda la informacin disponible (tanto erica como experimental) acerca de la teora para establecer la correspondiente distribucin de probabilidad en el espacio de parmetros asociado. Para este objetivo, la inferencia Bayesiana es la herramienta apropiada, especialmente cuando los datos experimentales nos son suficientes para seleccionar una regin pequea del espacio de parmetros.

En primer lugar hemos propuesto un marco de anlisis Bayesiano mejorado para explorar el MSSM, con un manejo cuidadoso de las distintas piezas de informacin. Los principales resultados y conclusiones son los siguientes:

- Hemos demostrado que realizando un anlisis Bayesiano correcto, aparece una penalizacin automtica del fine-tuning. Adems, dicha penalizacin es muy similar, aunque no exactamente igual, a las medidas "estndar" de fine-tuning, concretamente la de Barbieri-Giudice. Este resultado es completamente general, y de hecho va ms all de la consideracin de la supersimetra como la posible nueva fsica ms all del Modelo Estndar.

- Hemos realizado un tratamiento riguroso de las variables "nuisance", en particular los acoplos de Yukawa, demostrando que la prctica usual de tomar estos acoplos "tal como se requieren" para reproducir las masas corresponde aproximadamente a tomar priors logartmicos en los acoplos de Yukawa; lo que es una eleccin razonable.

- El uso de un conjunto de variables eficiente (y de hecho muy comn) para escanear el espacio de parmetros del MSSM, a saber $\{m_t, m_b; m, M, A, \tan\beta, \text{sgn } \mu\}$, implica heredar el factor Jacobiano (1.21), que contien en su interior la mencioanda penalizacin de regiones "fine-tuneadas". Hemos dado una expresin analtica aproximada para ese Jacobiano en eq.(1.33), vlida para cualquier MSSM. Estas expresiones no contienen asunciones ad hoc o prejuicios.

- Hemos desarrollado priors sensatos (planos y logartmicos) para las variables iniciales del MSSM, usando la asuncin bsica de que los trminos de ruptura soft tienen un origen comn; por lo que es lgico suponer que sus magnitudes son tambin similares.

Todas estas mejoras hacen el procedimiento mucho ms riguroso y, adems, permiten escanear el espacio de parmetros completo, incluyendo trminos soft arbitrariamente grandes. An as, la regin de baja energa est estadsticamente favorecida, incluso sin incluir datos de materia oscura o $g-2$.

Hemos aplicado este esquema a la realizacin de un "forecast" del MSSM para el LHC, centrndones en su versin constreida (SMSSM). Nuestros resultados aqu son los siguientes:

We have the applied this framework to the realization of a forecast of the MSSM for the LHC. We have focussed in its constrained version (CMSSM). Our main results here are the following:

- Incluyendo slo los datos experimentales ms robustos: observables E.W. and fsica del B(D), y cotas inferiores para las masas supersimtricas y del Higgs, la regin ms favorable del MSSM est a baja energa, pero hay una porcin significativa fuera del alcance del LHC.
- Aadiendo la informacin relativa a a_μ, usando datos de $e^+e^- \to$ had, los trminos soft son empjados espectacularmente dentro del alcance del LHC. Esto no pasa usando datos de τ-decay en vez de $e^+e^- \to$ had, lo que aconseja ser cautos a la hora de sacar conclusiones fuertes a partir de g-2.
- Grosso modo, la inclusin de datos de materia oscura empuja la regin favorecida del MSSM hacia baja energa, favoreciendo su deteccin en el LHC. Sin embargo, sobreviven grandes reas del espacio de parmetros fuera del alcance del LHC.

Las tcnicas Bayesianas son tambin tiles para cuantificar tensiones entre observables dentro del contexto de un modelo especfico. De esta forma, hemos seanalado que una masa "grande" del Higgs (i.e. no cercana a su cota inferior actual) es bastante incompatible con el observable $g-2$, si ste ltimo se evalua usando datos de $e^+e^- \to$ had data. Hemos cuantificado esta tensin como una funcin de la (todava desconocida) masa del Higgs, demostrando que:

- Para $m_h \geq 125$ GeV se alcanza ya el mximo nivel de discrepancia ($\sim 3.2\ \sigma$), incdcando que la supersimetra se desacopla, por lo que la prediccin para a_μ coincide con la del Modelo Estndar.
- Con los datos actuales, exigir una discrepancia menor que 3 σ discrepancy, implica $m_h \lesssim 120$ GeV.

Dada la complejidad del LHC, la comparacin econ las predicciones tericas ha de realizarse confrontando histogranas experimentales (en diferentes variables) con los producidos por simulaciones tericas. En este sentido, hemos presentado un mtodo riguroso y efectivo para comparar histogramas tericos y experimentales, evaluado la likelihood completa, y lo hemos aplicado a un caso fsico relevante. AL hacer esto hemos tenido en cuenta, adems de las incertidumbres estadsticas, la existencia de fuentes de error sistemtico adicional.

El procedimiento puede ser fcilmente incorporado tanto a anlisis Bayesianos como frecuentistas, ya que ambos se basan en la likelihood de los modelos tericos. En ambos casos, la incorporacin de la likelihood completa mejora las posibilidades de encontrar una seal de nueva fsica y, una vez es encontrada, identificar qu nueva fsica est detrs. Por ejemplo, si la nueva fsica es la supersimetra, el procedimiento permite encontrar de forma ptima los parmetros del modelo.

Hemos ilustrado el ltimo punto mostrando como una bsqueda en el espacio de parmetros del CMSSM, usando tcnicas Bayesianas, puede localizar de forma eficaz los valores correctos de los parmetros gracias a la comparacin de histogramas de eventos con multijets + missing energy, distribuidos en la variable de "masa efectiva". El procedimiento es realmente muy eficiente para encontrar el modelo supersimtrico correcto, en el caso de que la supersimetra est ah, y accesible al LHC. Pero naturalmente, la tcnica puede aplicarse a cualquier escenario de nueva fsica.

Finalmente, hemos tenido en cuenta los resultados relevantes que el LHC est ya arrojando. Concretamente, hemos usado las cotas superiore actuales y futuras sobre la masa del Higgs para establecer cotas superiores en las masas supersimtricas, $M_{\rm SUSY}$; usando el hecho de que en el MSSM el acoplo curtico del Higgs (ypor tanto su masa) es una funcin de las masas supersimtricas, en larticular las de los stops. Ahora mismo la restriccin en el espacio de parmetros del MSSM no es importante, pero lo ser muy pronto, como puede deducirse de la Fig. 5.1. Por otro lado, una parte importante del espacio de parmetros resulta ya excluido para split-SUSY, ver Fig. 5.2. A medida que el LHC mejore sus cotas en m_h, la regin permitida se har cada vez menor, demosrando que, incluso si no se encuentra la supersimetra, sta no puede ocultarse en escalas altas de forma arbitraria. Al final (esperamos) el Higgs ser encontrado, y uno puede usar estos resultados para establecer la regin del plano $M_{\rm SUSY} - \tan\beta$ consistente con el valor real de m_h.

Agradecimientos

A mis padres, Dominga Eugenia Catalán Gómez y Edgar Fidencio Cabrera Tello, por el gran esfuerzo que hicieron para que yo tuviera un abanico abierto de oportunidades. Gracias por incentivarme y apoyarme incondicionalmente para que siguiera adelante.

A mi asesor, Alberto Casas, por guiarme y enseñarme a dar los primeros pasos en la investigación, por trasmitirme sus conocimientos siempre con entusiasmo. Gracias por charlas informales de física, llenas de analogía, ejemplos muy sencillos e historia.

A Roberto Ruiz de Austri por haber compartido sus conocimientos conmigo, por haberme acompañado en estos mis primeros pasos en la investigación.

A Fernando Quevedo por su confianza y su apoyo.

A Mis compañeros de doctorado, con los que tuve el gusto de compartir estos cuatro años.

Appendix A: Using Fermion masses as initial parameters of the SM.

As is mentioned at the end in sect. 1.4, ref. [15] take fermion masses as initial parameters of the SM. As argued in sect. 1.3, it is theoretically bizarre to take m_t as a fundamental variable, instead of y_t. However, one may gain the bonus of almost no sensitivity to the prior in m_t, since this is essentially fixed by the experiment. This is true, but this procedure introduces extremely counter-intuitive contributions to the Jacobian, as we will see briefly. The new 2-variable Jacobian is given by

$$J_2 = \begin{vmatrix} \frac{\partial \mu}{\partial M_Z}\Big|_{t,m_t} & \frac{\partial \mu}{\partial t}\Big|_{M_Z,m_t} \\ \\ \frac{\partial B}{\partial M_Z}\Big|_{t,m_t} & \frac{\partial B}{\partial t}\Big|_{M_Z,m_t} \end{vmatrix} , \tag{5.6}$$

where the subscripts emphasize which variables have to be kept frozen in the partial derivations. Now, using the definitions (1.25), it is straightforward to obtain

$$J_2 = \frac{\partial f}{\partial M_Z}\frac{\partial h}{\partial t} + \frac{\partial g}{\partial M_Z}\left(\frac{\partial f}{\partial y}\frac{\partial h}{\partial t} - \frac{\partial f}{\partial t}\frac{\partial h}{\partial y}\right) + \frac{\partial f}{\partial M_Z}\frac{\partial h}{\partial y}\frac{\partial g}{\partial t} . \tag{5.7}$$

It is amusing that this expression for the Jacobian is more complicated than in the 3-variable case, eq. (1.26). This comes from the fact that the derivatives in (5.6) contain contributions coming from the dependence of μ and B on y, which is in turn a function of t and M_Z, eq. (1.24). These contributions were cancelled inside the 3-variable Jacobian thanks to the third row in the matrix of eq. (1.26), but they are not cancelled here and give rise to the second and third terms in eq. (5.7). Of course, if one ignores that μ and B in eqs.(1.22, 1.23) have an implicit dependence in the top Yukawa, as was done in ref. [15], these contributions do not appear; but this is not a good approximation, as we are about to see. Note that the first term in (5.7) is similar to the 3-variable Jacobian given by eq. (1.26)[1], whose physical significance (including the information about fine-tuning) was discussed after eq. (1.31). This term goes parametrically as B/μ. However the second term goes parametrically as $Bm^2/\mu M_Z^2$, and thus is much more important for large soft terms, which then become

[1]Thus the resemblance of the result of ref. [15] to our approximate expression (1.32), except for the RG and s_β^{-1} factors.

strongly favoured (contrary to the intuitive expectatives). Therefore there is no reason to ignore such term. In consequence, the expressions used in ref. [15] are much closer to using y_t as a fundamental variable with logarithmically flat prior than to using m_t.

Appendix B: Histogram comparison when experiment and the simulation have different luminosities

Some expressions of sects. 4.2 and 4.3 have to be modified when the effective luminosity of the simulation is not the same as the experimental one. In practice, the former can change from point to point when scanning the parameter space since typically one simulates a fixed number of supersymmetric events (say 10^4 events), but obviously the cross section changes throughout the parameter space. Of course one could adjust at every point the luminosity so that it coincides with the experiment, but normally this is costly in running time, and it is unnecessary, since the comparison can still be made as described next.

Let us call L^{th}, L^{exp} the luminosities of the theoretical simulation and the experiment, respectively, and suppose for a moment there are no systematic errors. Then the means that, under the null-hypothesis, are responsible for the experimental data, v_i, are *not* the ones of the simulation, μ_i, but

$$\hat{\mu}_i = \frac{L^{exp}}{L^{th}}\ \mu_i \ \equiv\ L\ \mu_i\ . \tag{5.8}$$

Hence eq.(4.8) becomes

$$P(v_i|u_i) = \int \prod_{i=1}^{K}\ d\mu_i \frac{(L\mu_i)^{v_i}}{v_i!} e^{-L\mu_i}\ \frac{\mu_i^{u_i}}{u_i!} e^{-\mu_i} = \prod_{i=1}^{K} \frac{(u_i+v_i)!}{u_i!\ v_i!}\ \ L^v\ (1+L)^{-1-u_i-v_i}\ . \tag{5.9}$$

Once systematic uncentainty is taken into account (4.2), everything is actually easy to handle since the luminosity factor L plays the role of a systematic and universal factor affecting the means of the simulation. More precisely, the equation (4.16), that relates the true means to be compared with the experiment, μ_i, with those of the simulation, μ_i^{th}, becomes

$$\mu_i\ =\ L\ f\ g_i\ \mu_i^{th}\ . \tag{5.10}$$

Therefore the subsequent equations remain the same with the simple change $f \to Lf$. In particular, the likelihood given by eq.(4.29) becomes now

$$\mathcal{P}(v_i|u_i) \propto \mathcal{P}(f=\frac{v}{Lu})\ \prod_{i=1}^{K}\left(\frac{(u_i+v_i)!}{u_i!v_i!}\ \int dg_i \frac{1}{g_i}\left(\frac{v}{u}g_i\right)^{v_i}\left(1+\frac{v}{u}g_i\right)^{-1-u_i-v_i}\mathcal{P}(g_i)\right)\ . \tag{5.11}$$

This is the formula we have used in our scan of the CMSSM parameter space.

Bibliography

[1] R. Trotta, *Applications of Bayesian model selection to cosmological parameters*, *Mon. Not. Roy. Astron. Soc.* **378** (2007) 72–82 [astro-ph/0504022].

[2] R. Trotta, *Bayes in the sky: Bayesian inference and model selection in cosmology*, *Contemp. Phys.* **49** (2008) 71–104 [0803.4089].

[3] H. Jeffreys, *Theory of probability*. Oxford Classics series, Oxford University Press, Oxford, UK, 3rd ed., 1998.

[4] F. Feroz, M. P. Hobson, L. Roszkowski, R. Ruiz de Austri and R. Trotta, *Are $BR(b \to s\gamma)$ and $(g-2)_\mu$ consistent within the Constrained MSSM?*, 0903.2487.

[5] S. P. Martin, *A Supersymmetry Primer*, hep-ph/9709356.

[6] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, *Phys.Rept.* **110** (1984) 1–162.

[7] C. F. Kolda, *Gauge mediated supersymmetry breaking: Introduction, review and update*, *Nucl.Phys.Proc.Suppl.* **62** (1998) 266–275 [hep-ph/9707450].

[8] B. Allanach, *SOFTSUSY: a program for calculating supersymmetric spectra*, *Comput.Phys.Commun.* **143** (2002) 305–331 [hep-ph/0104145].

[9] R. Barbieri and G. Giudice, *Upper Bounds on Supersymmetric Particle Masses*, *Nucl.Phys.* **B306** (1988) 63.

[10] P. Ciafaloni and A. Strumia, *Naturalness upper bounds on gauge mediated soft terms*, *Nucl.Phys.* **B494** (1997) 41–53 [hep-ph/9611204].

[11] J. Casas, J. Espinosa and I. Hidalgo, *Implications for new physics from fine-tuning arguments. 1. Application to SUSY and seesaw cases*, *JHEP* **0411** (2004) 057 [hep-ph/0410298].

[12] J. Casas, J. R. Espinosa and I. Hidalgo, *Implications for new physics from fine-tuning arguments. II. Little Higgs models*, *JHEP* **0503** (2005) 038 [hep-ph/0502066].

[13] L. Giusti, A. Romanino and A. Strumia, *Natural ranges of supersymmetric signals*, *Nucl.Phys.* **B550** (1999) 3–31 [hep-ph/9811386].

[14] B. Allanach, *Naturalness priors and fits to the constrained minimal supersymmetric standard model*, *Phys.Lett.* **B635** (2006) 123–130 [hep-ph/0601089].

[15] B. C. Allanach, K. Cranmer, C. G. Lester and A. M. Weber, *Natural priors, CMSSM fits and LHC weather forecasts*, *JHEP* **0708** (2007) 023 [0705.0487].

[16] A. Strumia, *Naturalness of supersymmetric models*, hep-ph/9904247. This paper is mainly a brief discussion of hep-ph 9911386, with only few original points.

[17] F. Feruglio, A. Strumia and F. Vissani, *Neutrino oscillations and signals in beta and 0nu2beta experiments*, *Nucl.Phys.* **B637** (2002) 345–377 [hep-ph/0201291].

[18] B. de Carlos and J. Casas, *One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem*, *Phys.Lett.* **B309** (1993) 320–328 [hep-ph/9303291].

[19] G. W. Anderson and D. J. Castano, *Measures of fine tuning*, *Phys.Lett.* **B347** (1995) 300–308 [hep-ph/9409419].

[20] P. Athron and . Miller, D.J., *A New Measure of Fine Tuning*, *Phys.Rev.* **D76** (2007) 075010 [0705.2241].

[21] J. L. Feng, K. T. Matchev and T. Moroi, *Multi - TeV scalars are natural in minimal supergravity*, *Phys.Rev.Lett.* **84** (2000) 2322–2325 [hep-ph/9908309].

[22] J. L. Feng, K. T. Matchev and T. Moroi, *Focus points and naturalness in supersymmetry*, *Phys.Rev.* **D61** (2000) 075005 [hep-ph/9909334].

[23] B. L. J. O. Berger and R. L. Wolpert, *Integrated likelihood methods for eliminating nuisance variables*, *Statistical Science* **14, 1** (1999) 1–28.

[24] D. M. Pierce, J. A. Bagger, K. T. Matchev and R.-j. Zhang, *Precision corrections in the minimal supersymmetric standard model*, *Nucl.Phys.* **B491** (1997) 3–67 [hep-ph/9606211].

[25] L. E. Ibanez and C. Lopez, *N=1 Supergravity, the Weak Scale and the Low-Energy Particle Spectrum*, *Nucl.Phys.* **B233** (1984) 511.

[26] L. J. Hall, R. Rattazzi and U. Sarid, *The Top quark mass in supersymmetric SO(10) unification*, *Phys.Rev.* **D50** (1994) 7048–7065 [hep-ph/9306309].

[27] A. E. Nelson and L. Randall, *Naturally large tan BETA*, *Phys.Lett.* **B316** (1993) 516–520 [hep-ph/9308277].

[28] V. S. Kaplunovsky and J. Louis, *Model independent analysis of soft terms in effective supergravity and in string theory*, *Phys.Lett.* **B306** (1993) 269–275 [hep-th/9303040].

[29] G. Giudice and A. Masiero, *A Natural Solution to the mu Problem in Supergravity Theories*, *Phys.Lett.* **B206** (1988) 480–484.

[30] F. Feroz and M. P. Hobson, *Multimodal nested sampling: an efficient and robust alternative to MCMC methods for astronomical data analysis*, *Mon. Not. Roy. Astron. Soc.* **384** (2008) 449–463 [0704.3704].

[31] *http://superbayes.org*.

[32] J. Skilling, *Nested sampling*, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering* (R. P. R. Fischer and U. von Toussaint, eds.), vol. 735, pp. 395–405, 2004.

[33] J. Skilling, *Nested sampling for general bayesian computation*, *Bayesian Analysis* **1** (2006), no. 4 833–860.

[34] **ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaborations, LEP Working Group for Higgs Boson Searches** Collaboration, S. Schael *et. al.*, *Search for neutral MSSM Higgs bosons at LEP*, *Eur.Phys.J.* **C47** (2006) 547–587 [hep-ex/0602042].

[35] J. Casas, J. Espinosa and H. Haber, *The Higgs mass in the MSSM infrared fixed point scenario*, *Nucl.Phys.* **B526** (1998) 3–20 [hep-ph/9801365].

[36] H. Baer, C.-h. Chen, M. Drees, F. Paige and X. Tata, *Supersymmetry reach of Tevatron upgrades: The Large tan Beta case*, *Phys.Rev.* **D58** (1998) 075008 [hep-ph/9802441].

[37] M. Jurcisin and D. Kazakov, *Infrared quasifixed points and mass predictions in the MSSM. 2. Large tan(beta) scenario*, *Mod.Phys.Lett.* **A14** (1999) 671–688 [hep-ph/9902290].

[38] R. R. de Austri, R. Trotta and L. Roszkowski, *A Markov chain Monte Carlo analysis of the CMSSM*, *JHEP* **0605** (2006) 002 [hep-ph/0602028].

[39] J. R. Ellis, T. Falk, G. Ganis, K. A. Olive and M. Srednicki, *The CMSSM parameter space at large tan beta*, *Phys.Lett.* **B510** (2001) 236–246 [hep-ph/0102098].

[40] **CMS** Collaboration, G. L. Bayatian *et. al.*, *CMS technical design report, volume II: Physics performance*, *J. Phys.* **G34** (2007) 995–1579.

[41] **The ATLAS** Collaboration, G. Aad *et. al.*, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, 0901.0512.

[42] H. Baer, V. Barger, A. Lessa and X. Tata, *Supersymmetry discovery potential of the LHC at s**(1/2) = 10-TeV and 14-TeV without and with missing E(T)*, *JHEP* **0909** (2009) 063 [0907.1922].

[43] L. Susskind, *Supersymmetry breaking in the anthropic landscape*, hep-th/0405189.

[44] B. Allanach and C. Lester, *Multi-dimensional mSUGRA likelihood maps*, *Phys.Rev.* **D73** (2006) 015013 [hep-ph/0507283].

[45] L. Roszkowski, R. Ruiz de Austri and R. Trotta, *Implications for the Constrained MSSM from a new prediction for $b \to s\gamma$*, *JHEP* **07** (2007) 075 [0705.2012].

[46] O. Buchmueller, R. Cavanaugh, A. De Roeck, J. Ellis, H. Flacher *et. al.*, *Predictions for Supersymmetric Particle Masses using Indirect Experimental and Cosmological Constraints*, *JHEP* **0809** (2008) 117 [0808.4128].

[47] R. Trotta, F. Feroz, M. P. Hobson, L. Roszkowski and R. Ruiz de Austri, *The Impact of priors and observables on parameter inferences in the Constrained MSSM*, *JHEP* **12** (2008) 024 [0809.3792].

[48] J. Ellis, *Prospects for Discovering Supersymmetry at the LHC*, *Eur.Phys.J.* **C59** (2009) 335–343 [0810.1178].

[49] S. S. AbdusSalam, B. C. Allanach, F. Quevedo, F. Feroz and M. Hobson, *Fitting the Phenomenological MSSM*, *Phys.Rev.* **D81** (2010) 095012 [0904.2548].

[50] O. Buchmueller, R. Cavanaugh, A. De Roeck, J. Ellis, H. Flacher *et. al.*, *Likelihood Functions for Supersymmetric Observables in Frequentist Analyses of the CMSSM and NUHM1*, *Eur.Phys.J.* **C64** (2009) 391–415 [0907.5568].

[51] **Tevatron Electroweak Working Group and CDF and D0 Collaboration** Collaboration, *A Combination of CDF and D0 Results on the Mass of the Top Quark*, 0803.1683.

[52] **Particle Data Group** Collaboration, W. Yao *et. al.*, *Review of Particle Physics*, *J.Phys.G* **G33** (2006) 1–1232.

[53] K. Hagiwara, A. Martin, D. Nomura and T. Teubner, *Improved predictions for g-2 of the muon and alpha(QED) (M**2(Z))*, *Phys.Lett.* **B649** (2007) 173–179 [hep-ph/0611102].

[54] A. Dedes, G. Degrassi and P. Slavich, *On the two loop Yukawa corrections to the MSSM Higgs boson masses at large tan beta*, *Nucl.Phys.* **B672** (2003) 144–162 [hep-ph/0305127].

[55] M. Awramik, M. Czakon, A. Freitas and G. Weiglein, *Precise prediction for the W boson mass in the standard model*, *Phys.Rev.* **D69** (2004) 053006 [hep-ph/0311148].

[56] M. Awramik, M. Czakon, A. Freitas and G. Weiglein, *Complete two-loop electroweak fermionic corrections to $\sin^2\theta^{\rm lept}_{\rm eff}$ and indirect determination of the Higgs boson mass*, *Phys.Rev.Lett.* **93** (2004) 201805 [hep-ph/0407317].

[57] A. Djouadi, P. Gambino, S. Heinemeyer, W. Hollik, C. Junger *et. al.*, *Leading QCD corrections to scalar quark contributions to electroweak precision observables*, *Phys.Rev.* **D57** (1998) 4179–4196 [hep-ph/9710438].

[58] W. de Boer and C. Sander, *Global electroweak fits and gauge coupling unification*, *Phys.Lett.* **B585** (2004) 276–286 [hep-ph/0307049].

[59] S. Heinemeyer, W. Hollik and G. Weiglein, *Electroweak precision observables in the minimal supersymmetric standard model*, *Phys.Rept.* **425** (2006) 265–368 [hep-ph/0412214].

[60] R. Barbieri and A. Strumia, *The 'LEP paradox'*, hep-ph/0007265.

[61] R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, *Electroweak symmetry breaking after LEP-1 and LEP-2*, *Nucl.Phys.* **B703** (2004) 127–146 [hep-ph/0405040].

[62] G. Degrassi, P. Gambino and P. Slavich, *SusyBSG: A Fortran code for $BR[B \to X(s)\gamma]$ in the MSSM with Minimal Flavor Violation*, *Comput.Phys.Commun.* **179** (2008) 759–771 [0712.3265].

[63] G. Degrassi, P. Gambino and P. Slavich, *QCD corrections to radiative B decays in the MSSM with minimal flavor violation*, *Phys.Lett.* **B635** (2006) 335–342 [hep-ph/0601135].

[64] G. D'Ambrosio, G. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl.Phys.* **B645** (2002) 155–187 [hep-ph/0207036].

[65] J. Foster, K.-i. Okumura and L. Roszkowski, *New Higgs effects in B-physics in supersymmetry with general flavour mixing*, *Phys.Lett.* **B609** (2005) 102–110 [hep-ph/0410323].

[66] F. Mahmoudi, *SuperIso v2.3: A Program for calculating flavor physics observables in Supersymmetry*, *CPHCB,180,1579-1613.2009* **180** (2009) 1579–1613 [0808.3144].

[67] *http://lepewwg.web.cern.ch/LEPEWWG*.

[68] J. P. Miller, E. de Rafael and B. L. Roberts, *Muon g-2: Review of Theory and Experiment*, *Rept. Prog. Phys.* **70** (2007) 795 [hep-ph/0703049].

[69] M. Davier *et. al.*, *The Discrepancy Between tau and e+e- Spectral Functions Revisited and the Consequences for the Muon Magnetic Anomaly*, *Eur. Phys. J.* **C66** (2010) 127–136 [0906.5443].

[70] **CDF - Run II Collaboration** Collaboration, A. Abulencia *et. al.*, *Measurement of the $B_s^0 - \overline{B}_s^0$ Oscillation Frequency*, *Phys.Rev.Lett.* **97** (2006) 062003 [hep-ex/0606027].

[71] **CDF Collaboration** Collaboration, A. Abulencia *et. al.*, *Observation of B0(s) - anti-B0(s) Oscillations*, *Phys.Rev.Lett.* **97** (2006) 242003 [hep-ex/0609040].

[72] **Heavy Flavor Averaging Group** Collaboration, E. Barberio *et. al.*, *Averages of b−hadron and c−hadron Properties at the End of 2007*, 0808.1297.

[73] **BABAR Collaboration** Collaboration, B. Aubert *et. al.*, *Measurement of Branching Fractions and CP and Isospin Asymmetries in $B \to K^*\gamma$*, 0808.1915.

[74] **BABAR Collaboration** Collaboration, B. Aubert *et. al.*, *Observation of the semileptonic decays $B \to D^*\tau^-\overline{\nu}(\ \tau^)$ and evidence for $B \to D\tau^-\overline{\nu}(\ \tau^)$*, *Phys.Rev.Lett.* **100** (2008) 021801 [0709.1698].

[75] **FlaviaNet Working Group on Kaon Decays** Collaboration, M. Antonelli *et. al.*, *Precision tests of the Standard Model with leptonic and semileptonic kaon decays*, 0801.1817.

[76] A. Akeroyd and F. Mahmoudi, *Constraints on charged Higgs bosons from $D(s) + - \to \mu + -\nu$ and $D(s) + - \to \tau + -\nu$*, *JHEP* **0904** (2009) 121 [0902.2393].

[77] **WMAP Collaboration** Collaboration, J. Dunkley *et. al.*, *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data*, *Astrophys.J.Suppl.* **180** (2009) 306–329 [0803.0586].

[78] **LEP Working Group for Higgs boson searches, ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration** Collaboration, R. Barate *et. al.*, *Search for the standard model Higgs boson at LEP*, *Phys.Lett.* **B565** (2003) 61–75 [hep-ex/0306033].

[79] L. Roszkowski, R. R. de Austri and R. Trotta, *On the detectability of the CMSSM light Higgs boson at the Tevatron*, *JHEP* **0704** (2007) 084 [hep-ph/0611173].

[80] J. R. Ellis, G. Ridolfi and F. Zwirner, *On radiative corrections to supersymmetric Higgs boson masses and their implications for LEP searches*, *Phys. Lett.* **B262** (1991) 477–484.

[81] M. Davier, A. Hoecker, B. Malaescu, C. Z. Yuan and Z. Zhang, *Reevaluation of the hadronic contribution to the muon magnetic anomaly using new e+e- → pi+pi- cross section data from BABAR*, *Eur. Phys. J.* **C66** (2010) 1–9 [0908.4300].

[82] S. Heinemeyer, D. Stockinger and G. Weiglein, *Electroweak and supersymmetric two-loop corrections to (g-2)(mu)*, *Nucl.Phys.* **B699** (2004) 103–123 [hep-ph/0405255].

[83] S. Marchetti, S. Mertens, U. Nierste and D. Stockinger, *Tan(beta)-enhanced supersymmetric corrections to the anomalous magnetic moment of the muon*, *Phys.Rev.* **D79** (2009) 013010 [0808.1530].

[84] G. Jungman, M. Kamionkowski and K. Griest, *Supersymmetric dark matter*, *Phys.Rept.* **267** (1996) 195–373 [hep-ph/9506380].

[85] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, *MicrOMEGAs: A Program for calculating the relic density in the MSSM*, *Comput.Phys.Commun.* **149** (2002) 103–120 [hep-ph/0112278].

[86] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, *micrOMEGAs: Version 1.3*, *Comput.Phys.Commun.* **174** (2006) 577–604 [hep-ph/0405253].

[87] L. Knox and M. S. Turner, *Inflation at the electroweak scale*, *Phys.Rev.Lett.* **70** (1993) 371–374 [astro-ph/9209006].

[88] J. Garcia-Bellido, D. Y. Grigoriev, A. Kusenko and M. E. Shaposhnikov, *Nonequilibrium electroweak baryogenesis from preheating after inflation*, *Phys.Rev.* **D60** (1999) 123504 [hep-ph/9902449].

[89] A. Ibarra and D. Tran, *Decaying Dark Matter and the PAMELA Anomaly*, *JCAP* **0902** (2009) 021 [0811.1555].

[90] M. E. Cabrera, J. A. Casas and R. Ruiz de Austri, *Bayesian approach and Naturalness in MSSM analyses for the LHC*, *JHEP* **03** (2009) 075 [0812.0536].

[91] D. Stockinger, *The muon magnetic moment and supersymmetry*, *J. Phys.* **G34** (2007) R45–R92 [hep-ph/0609168].

[92] J. R. Ellis, G. Ridolfi and F. Zwirner, *Radiative corrections to the masses of supersymmetric Higgs bosons*, *Phys. Lett.* **B257** (1991) 83–91.

[93] Y. Okada, M. Yamaguchi and T. Yanagida, *Upper bound of the lightest Higgs boson mass in the minimal supersymmetric standard model*, *Prog. Theor. Phys.* **85** (1991) 1–6.

[94] Y. Okada, M. Yamaguchi and T. Yanagida, *Renormalization group analysis on the Higgs mass in the softly broken supersymmetric standard model*, *Phys. Lett.* **B262** (1991) 54–58.

[95] H. E. Haber and R. Hempfling, *Can the mass of the lightest Higgs boson of the minimal supersymmetric model be larger than m(Z)?*, *Phys. Rev. Lett.* **66** (1991) 1815–1818.

[96] R. Barbieri, M. Frigeni and F. Caravaglios, *The Supersymmetric Higgs for heavy superpartners*, *Phys. Lett.* **B258** (1991) 167–170.

[97] J. A. Casas, J. R. Espinosa, M. Quiros and A. Riotto, *The Lightest Higgs boson mass in the minimal supersymmetric standard model*, *Nucl. Phys.* **B436** (1995) 3–29 [hep-ph/9407389].

[98] G. Degrassi and G. Giudice, *QED logarithms in the electroweak corrections to the muon anomalous magnetic moment*, *Phys.Rev.* **D58** (1998) 053007 [hep-ph/9803384].

[99] S. Heinemeyer, D. Stockinger and G. Weiglein, *Two loop SUSY corrections to the anomalous magnetic moment of the muon*, *Nucl.Phys.* **B690** (2004) 62–80 [hep-ph/0312264].

[100] P. von Weitershausen, M. Schafer, H. Stockinger-Kim and D. Stockinger, *Photonic SUSY Two-Loop Corrections to the Muon Magnetic Moment*, *Phys.Rev.* **D81** (2010) 093004 [1003.5820].

[101] T. Moroi, *The Muon Anomalous Magnetic Dipole Moment in the Minimal Supersymmetric Standard Model*, *Phys. Rev.* **D53** (1996) 6565–6575 [hep-ph/9512396].

[102] S. Heinemeyer, W. Hollik and G. Weiglein, *The Masses of the neutral CP - even Higgs bosons in the MSSM: Accurate analysis at the two loop level*, *Eur.Phys.J.* **C9** (1999) 343–366 [hep-ph/9812472].

[103] G. Degrassi, P. Slavich and F. Zwirner, *On the neutral Higgs boson masses in the MSSM for arbitrary stop mixing*, *Nucl.Phys.* **B611** (2001) 403–422 [hep-ph/0105096].

[104] A. Brignole, G. Degrassi, P. Slavich and F. Zwirner, *On the O(alpha(t)**2) two loop corrections to the neutral Higgs boson masses in the MSSM*, *Nucl.Phys.* **B631** (2002) 195–218 [hep-ph/0112177].

[105] A. Brignole, G. Degrassi, P. Slavich and F. Zwirner, *On the two loop sbottom corrections to the neutral Higgs boson masses in the MSSM*, *Nucl.Phys.* **B643** (2002) 79–92 [hep-ph/0206101].

[106] F. Feroz *et. al.*, *Bayesian Selection of sign(mu) within mSUGRA in Global Fits Including WMAP5 Results*, *JHEP* **10** (2008) 064 [0807.4512].

[107] F. Feroz, M. P. Hobson and M. Bridges, *MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics*, *Mon. Not. Roy. Astron. Soc.* **398** (2009) 1601–1614 [0809.3437].

[108] M. E. Cabrera, J. A. Casas and R. Ruiz d Austri, *MSSM Forecast for the LHC*, *JHEP* **05** (2010) 043 [0911.4686].

[109] J. A. Casas, J. R. Espinosa and I. Hidalgo, *The MSSM fine tuning problem: A Way out*, *JHEP* **01** (2004) 008 [hep-ph/0310137].

[110] C. Beskidt, W. de Boer, T. Hanisch, E. Ziebarth, V. Zhukov *et. al.*, *Constraints on Supersymmetry from Relic Density compared with future Higgs Searches at the LHC*, *Phys.Lett.* **B695** (2011) 143–148 [1008.2150].

[111] O. Buchmueller, R. Cavanaugh, D. Colling, A. De Roeck, M. Dolan *et. al.*, *Frequentist Analysis of the Parameter Space of Minimal Supergravity*, *Eur.Phys.J.* **C71** (2011) 1583 [1011.6118].

[112] **CMS** Collaboration, *Search strategy for exclusive multi-jet events from supersymmetry at cms*, tech. rep., 2009. PAS-SUS-09-001.

[113] *https://twiki.cern.ch/twiki/bin/view/AtlasPublic*.

[114] *https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResults*.

[115] G. L. Kane, C. F. Kolda, L. Roszkowski and J. D. Wells, *Study of constrained minimal supersymmetry*, *Phys.Rev.* **D49** (1994) 6173–6210 [hep-ph/9312272].

[116] *Search for supersymmetry with jets, missing transverse momentum and one lepton at sqrt(s) = 7 tev*, Tech. Rep. ATLAS-CONF-2011-090, CERN, Geneva, Jun, 2011.

[117] **CMS** Collaboration, *Search for supersymmetry in all-hadronic events with missing energy*, Tech. Rep. PAS-SUS-11-004, 2011.

[118] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026 [hep-ph/0603175].

[119] H. Baer, C.-h. Chen, M. Drees, F. Paige and X. Tata, *Probing minimal supergravity at the CERN LHC for large tan Beta*, *Phys.Rev.* **D59** (1999) 055014 [hep-ph/9809223].

[120] **CMS** Collaboration, *Sm higgs combination*, tech. rep., 2011.

[121] M. S. Carena, J. Espinosa, M. Quiros and C. Wagner, *Analytical expressions for radiatively corrected Higgs masses and couplings in the MSSM*, *Phys.Lett.* **B355** (1995) 209–221 [hep-ph/9504316].

[122] H. E. Haber, R. Hempfling and A. H. Hoang, *Approximating the radiatively corrected Higgs mass in the minimal supersymmetric model*, *Z.Phys.* **C75** (1997) 539–554 [hep-ph/9609331].

[123] B. Grzadkowski and M. Lindner, *STABILITY OF TRIVIALITY MASS BOUNDS IN THE STANDARD MODEL*, *Phys.Lett.* **178B** (1986) 81.

[124] *Search for squarks and gluinos using final states with jets and missing transverse momentum with the atlas detector in sqrts = 7tev proton-proton collisions*, Tech. Rep. ATLAS-CONF-2011-086, CERN, Geneva, Jun, 2011.

[125] **CMS** Collaboration, *Search for supersymmetry in all-hadronic events with α_t*, Tech. Rep. CMS-PAS-SUS-11-003, 2011.

[126] **CDF and D0 Collaboration** Collaboration, *Combination of CDF and D0 Results on the Mass of the Top Quark using up to 5.6 fb^{-1} of data*, 1007.3178. Long author list - awaiting processing.

[127] J. Espinosa and M. Quiros, *Two loop radiative corrections to the mass of the lightest Higgs boson in supersymmetric standard models*, *Phys.Lett.* **B266** (1991) 389–396.

[128] N. Arkani-Hamed and S. Dimopoulos, *Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC*, *JHEP* **0506** (2005) 073 [hep-th/0405159].

[129] G. Giudice and A. Romanino, *Split supersymmetry*, *Nucl.Phys.* **B699** (2004) 65–89 [hep-ph/0406088].

[130] P. Uwer, *EasyNData: A Simple tool to extract numerical values from published plots*, 0710.2896.