A new approach to Naturalness in SUSY models.

D. M. Ghilencea

CERN Theory Division, CH-1211 Geneva 23, Switzerland and
Theoretical Physics Department, National Institute of Physics and Nuclear Engineering
IFIN-HH Bucharest MG-6, Romania.
E-mail: dumitru.ghilencea@cern.ch

We review recent results that provide a new approach to the old problem of naturalness in supersymmetric models, without relying on subjective definitions for the fine-tuning associated with fixing the EW scale (to its measured value) in the presence of quantum corrections. The approach can address in a model-independent way many questions related to this problem. The results show that naturalness and its measure (fine-tuning) are an intrinsic part of the likelihood to fit the data that includes the EW scale. One important consequence is that the additional constraint of fixing the EW scale, usually not imposed in the data fits of the models, impacts on their overall likelihood to fit the data (or \(\chi^2/n_{df}, n_{df}^f\): number of degrees of freedom). This result has negative implications for the viability of currently popular supersymmetric extensions of the Standard Model.

CERN-PH-TH-2013-065
1. Naturalness: many questions, few answers.

The purpose of this talk is to provide a new, different perspective to the old problem of naturalness in particle theory, based on recent results [1, 2] that we review here. As pointed out long ago [3], in order to fix the electroweak (EW) scale in the Standard Model (SM) to its current value, in the presence of quantum corrections, a tremendous amount of “fine” tuning of parameters is required (roughly 1 part in $10^{33}$), and this is thought to be unnatural. This tuning is just another face of the large hierarchy between the EW and the Planck scales and indicates when a model is unnatural, assuming its validity up to the Planck scale. However, “new physics” can enter at some lower scale. For example, additional symmetries can naturally explain such hierarchies [4]. One such possibility is to build models based on scale (conformal) symmetry [5], not presented here. Another possibility is to use low-energy supersymmetry, which is the case discussed below. With supersymmetry broken at a scale $m_{\text{susy}}$ above few TeV, there is however a “remnant” of the fine-tuning mentioned in the SM that worsens as $m_{\text{susy}}$ is increased, due to negative SUSY searches. So the level of fine tuning became a “measure” of the success of SUSY as a natural solution to the hierarchy problem i.e. a measure of the naturalness of a model.

The problem is that we do not have a widely accepted definition for fine tuning $\Delta$ and criteria to decide when a model is natural. This brings many open questions. The first definition was [6]

$$\Delta = \max_{\vec{\gamma}} \left| \frac{\partial \ln v}{\partial \ln \vec{\gamma}} \right|, \quad \vec{\gamma} = \{m_0, m_{1/2}, A_0, B_0, \mu_0 \cdots \}. \quad (1.1)$$

where $\vec{\gamma}$ are components of the set $\vec{\gamma}$ of SUSY-related parameters that define the model and $v$ is the EW scale (higgs vev). This definition is motivated from the physics point of view, and measures the stability of the EW scale at the quantum level under small variations of $\vec{\gamma}$, after SUSY is broken. It was used in many tests for the naturalness of the models. But an immediate question arises: $\Delta$ depends on parameters $\vec{\gamma}$, so should these parameters include only those that SUSY introduces or also some non-SUSY ones as well, like nuisance variables (e.g. Yukawa couplings)? For example top Yukawa coupling could bring the dominant contribution to $\Delta$, so should we include $y_t$ on the list of parameters in (1.1)? Further, one can also use other definitions for $\Delta$, see [7], or more recently [8], or for example:

$$\Delta = \left\{ \sum_{\vec{\gamma}} \left( \frac{\partial \ln v}{\partial \ln \vec{\gamma}} \right)^2 \right\}^{1/2}, \quad \vec{\gamma} = \{m_0, m_{1/2}, A_0, B_0, \mu_0 \cdots \}. \quad (1.2)$$

Do all these different definitions for $\Delta$ lead to similar conclusions for the viable regions of the parameter space of the model? what is the right definition of fine tuning? Another drawback is that $\Delta$ provides a local measure in the $\vec{\gamma}$ space of the quantum cancellations that fix the EW scale, while to compare models, one could say that a more global measure is desirable. How do we compare models of different $\Delta$, over the entire parameter space?

Even if all theorists agree on a particular definition for $\Delta$, then what $\Delta$ is acceptable for a model to be considered natural? In principle we want $\Delta$ be small, so the EW scale is stable when varying $\vec{\gamma}$, in the presence of the radiative corrections; but is $\Delta = 100$ (i.e. tuning of 1 part in 100)
acceptable? or $\Delta = 10000$? If no sign of TeV-scale supersymmetry is found at the LHC, is there a value of $\Delta$ for which we should give up a model based on low-energy supersymmetry?

Further questions arise when comparing two models, as discussed in the following. Firstly, for a supersymmetric model that can fit the experimental data represented by a set of observables $O_{\text{ex}}^i$ of theoretical values $O_{\text{th}}^i$, one defines

$$\chi^2 = \sum_i \frac{(O_{\text{th}}^i - O_{\text{ex}}^i)^2}{\sigma_i^2}$$

(1.3)

A good $\chi^2$ fit requires that ($n_{df}$ denotes the number of degrees of freedom)

$$\frac{\chi^2}{n_{df}} \approx 1,$$  (1.4)

where $n_{df} = n_O - n_p$.

Suppose now that we have two models A and B: model A has a very good fit $\chi^2/n_{df} \approx 1$ but fine-tuning $\Delta$ (according to some definition) of order $\mathcal{O}(1000)$. So the model is likely but rather ... unnatural! Consider that model B has a good $\chi^2/n_{df}$ but slightly worse compared to model A, but its fine tuning much improved $\mathcal{O}(10)$ (according to the same definition). How do we decide which model of the two is better? This situation is confusing, more so since current data fits indeed report separately the value of $\chi^2/n_{df}$ and the value of $\Delta$. Also the possibility of having a realistic model (i.e. good likelihood) that is unnatural, or a natural one but unrealistic reflects a puzzling situation. We would like to answer these questions, but since we cannot even agree on a definition for $\Delta$, how to address them?

It is then highly desirable not to rely on any definition for fine-tuning and to use more fundamental arguments (and eventually derive a measure for naturalness from these). The first step to understand what happens is this: one tunes the parameters $\gamma_i$ to fit the observables. This brings a $\chi^2/n_{df}$ “cost”. One must also tune the same parameters $\gamma_i$ to fix the EW scale, which is central to the issue of fine tuning $\Delta$. Clearly, there is no technical difference between the two tunings. That means that total $\chi^2$ and $\Delta$ must be technically related. But what is their relation?

Let us think from a different perspective and provide the central idea. While we cannot agree on a definition for fine-tuning (naturalness) in a model, we do know if a model is realistic or not: we test it against experimental data. We do so by computing the likelihood to fit the data (frequentist approach) or the Bayesian probability of the model. These are well defined mathematical (probabilistic) tools used to test a model. Next, recall the original goal of SUSY, central to the issue of naturalness: fixing the EW scale ($m_Z$) to its measured value, in the presence of quantum corrections. If we regard this as a constraint and impose it on the current likelihood to fit the data, the associated physical problem of naturalness (fine-tuning) should be captured by the mathematics that describes this constraint! So one should compute the likelihood to fit the data with this constraint imposed! Surprisingly, this is something not yet done by the current methods that compute the likelihood (or $\chi^2$) to fit the data in SUSY models. This is the central idea that we follow below.

Without assuming any definition for fine-tuning or reference to this concept, we show how the constraint of fixing the EW scale automatically leads to a correction factor that impacts on (worsens) the current likelihood to fit the data. Once this mathematical relation is established one can answer, in a model-independent way and on probabilistic grounds, the questions listed above.

Let us then compute the likelihood to fit the data in a SUSY model under the mathematical constraint of fixing the electroweak scale to its measured value, in the presence of quantum corrections. The constraint encodes the physics behind the naturalness problem, and leads to a new perspective on this issue. Firstly note that in accurate data fits one usually performs a fit with the following observables (in a standard notation):

\[ \sin^2 \theta_{e}^{\text{lep}}, \Gamma_Z, \delta a_\mu, m_h, m_W, \Omega_{DM}, BR(B \rightarrow X_s \gamma), BR(B_s \rightarrow \mu^+ \mu^-), BR(B_d \rightarrow \tau \nu), \Delta M_B. \] (2.1)

However, fixing the EW scale to its accurately measured value \( m_H^0 \), is not on this list, so the current likelihood tests of supersymmetric models do not account for this effect (constraint).

For each observable \( O_i \) denote the corresponding probability \( P(O_i | \gamma, y) \), which is usually taken a Gaussian. \( P(O_i | \gamma, y) \) depends on parameters \( \gamma = \{ \gamma_1, \gamma_2, \ldots \} \) introduced by SUSY and on other parameters too, like Yukawa couplings \( y = \{ y_1, y_2, \ldots \} \) that denote nuisance variables. These are variables that can be removed from the total likelihood by either maximizing it wrt to them (for \( \gamma \) fixed) or by integrating over them wrt some measure. Since observables \( O_i \) are independent, one multiplies their probability distributions to obtain a total distribution \( \mathcal{L}(\text{data} | \gamma, y) \) where “data” stands for all observables. With “data” fixed, \( \mathcal{L} \) is a function of the sets \( \gamma, y \) only, so we denote it as \( \mathcal{L}(\text{data} | \gamma, y) \equiv L(\gamma, y) \), with

\[ L(\gamma, y) = \prod_i P(O_i | \gamma, y); \quad \gamma = \{ m_0, m_{1/2}, \mu_0, m_0, A_0, B_0, \ldots \}; \quad y = \{ y_1, y_2, \ldots \} \] (2.2)

One then computes the corresponding \( \chi^2 \equiv -2 \ln L \), by seeking a region of the parameter space where \( \chi^2 / n_{\text{dof}} \approx 1 \) at the minimum, by tuning \( \gamma_j \) and \( y_k \) to fit observables \( O_i \). If one remains true to the original goal of SUSY, one should impose on \( L \) the additional constraint of fixing the EW scale to the measured value \( m_H^0 \) and analyze its impact on \( L \). This is what we shall do below. For technical details see the original papers [1].

We start with the scalar potential of supersymmetric models (\( H_{1,2} \) higgs doublets):

\[
V = m_1^2 |H_1|^2 + m_2^2 |H_2|^2 -(m_3^2 H_1 H_2 + \text{h.c.}) + (\lambda_4/2) |H_1|^4 + (\lambda_5/2) |H_2|^4 + \lambda_3 |H_1|^2 |H_2|^2
+ \lambda_4 |H_1 H_2|^2 + [ (\lambda_5/2) (H_1 H_2)^2 + \lambda_6 |H_1|^2 (H_1 H_2) + \lambda_7 |H_2|^2 (H_1 H_2) + \text{h.c.} ]
\] (2.3)

Some \( \lambda_j \) are non-zero only at quantum level (considered here). Introduce effective couplings \( \lambda, m \)

\[
\lambda = \lambda_1 / 2 \cos^4 \beta + \lambda_2 / 2 \sin^4 \beta + (\lambda_3 + \lambda_4 + \lambda_5) / 4 \sin^2 2 \beta + (\lambda_6 \cos^2 \beta + \lambda_7 \sin^2 \beta);
\]

\[
m^2 = m_1^2 \cos^2 \beta + m_2^2 \sin^2 \beta - m_3^2 \sin 2 \beta,
\] (2.4)

We also denote \( v^2 \equiv \langle h_1^0 \rangle^2 + \langle h_2^0 \rangle^2 \). Then the EW minimum conditions determine \( v \) and \( \tan \beta \):

\[
v = (-m^2/\lambda)^{1/2} = 0, \quad \text{and} \quad \tan \beta - \tan \beta_0 (\gamma, y, v) = 0.
\] (2.5)
where $\beta_0$ is the root of the second minimum condition. Let us denote
\begin{equation}
\tilde{v}(\gamma, y, v, \beta) \equiv (-m^2/\lambda)^{1/2}
\end{equation}
which at the quantum level depends on the arguments shown.

The “constrained” likelihood, hereafter denoted $L_\text{w}$, to fit the data and also to fix the electroweak scale to its value ($m_Z^0$) can be written in terms of the usual likelihood $L$ to fit observables other than the EW scale itself, as follows:
\begin{equation}
L_\text{w}(\gamma, y) = m_Z^0 \int d\beta \delta\left[ v - (-m^2/\lambda)^{1/2} \right] \delta\left[ \tan \beta - \tan \beta_0(\gamma, y, v) \right] \delta(m_Z - m_Z^0) L(\gamma, y, v, \beta) \\
= v_0 \delta\left[ v_0 - \tilde{v}(\gamma, y, v_0, \beta_0(\gamma, y, v_0)) \right] L(\gamma, y, v_0, \beta_0(\gamma, y, v_0)) \\
= \frac{v_0}{|\nabla v|_0} \delta\left[ n_i (\ln z_i - \ln z_i^0) \right] L(\gamma, y, v_0, \beta_0(\gamma, y, v_0)), \quad \text{with} \quad z_i \equiv \{ \gamma_j, y_k \}.
\end{equation}
with a sum over repeated index $i$; $m_Z^0$ in front of the integral is a normalization factor and
\begin{equation}
m_Z = (g_1^2 + g_2^2)^{1/2} v_2/2, \quad m_Z^0 = (g_1^2 + g_2^2)^{1/2} v_0/2, \quad v_0 = 246 \text{ GeV}
\end{equation}
$m_Z$ is the theoretical mass of Z boson, $m_Z^0$ is its measured value ($\approx 91.2$ GeV) and $g_1, g_2$ are gauge couplings of U(1), SU(2). Under integral (2.7) two Dirac delta functions impose EW minimum conditions (2.5). In the presence of these functions, the integrals over $v$ and $\tan \beta$ are just a formal way to say that we solved the constraints of the EW minimum. Actually, the integral over $v$ can be regarded as an integral over $m_Z \propto v$ which is an observable of probability distribution $\delta(m_Z - m_Z^0)$; the role of this delta function is to fix the EW scale to the measured value. Using $\delta(m_Z - m_Z^0)$ is justified since $m_Z$ is accurately measured. In the last step in (2.7) a Taylor expansion of $\tilde{v}$ near $v_0$ was done, and $n_i$ are components of the normal to the surface: $v_0 - \tilde{v}(\gamma, y, v_0, \beta_0(\gamma, y, v_0)) = 0$ which has the solution $\gamma_j = \gamma_j^0, y_k = y_k^0$. So $n_i = (\partial \tilde{v}/|\nabla \tilde{v}|)_0$ where subscript “o” indicates evaluation at $\gamma_j^0, y_k^0$. $|\nabla (\ldots)|_0$ denotes the gradient (in parameter space $\ln \gamma_j, \ln y_k$), evaluated again at $\gamma_j^0, y_k^0$ [1].

From eq.(2.7), we conclude that the “constrained” likelihood $L_\text{w}$ is nonzero if $\gamma_j = \gamma_j^0, y_k = y_k^0$. These parameters are correlated by $v_0 - \tilde{v}(\gamma_j^0, y_k^0, v_0, \beta_0(\gamma_j^0, y_k^0, v_0)) = 0$ so one could eliminate one of them (usually chosen to be $\mu_0$). With this in mind, the result of the last equation becomes:
\begin{equation}
L_\text{w}(\gamma_j^0, y_k^0) = \frac{L(\gamma_j^0, y_k^0, v_0, \beta_0(\gamma_j^0, y_k^0, v_0))}{\Delta_\gamma(\gamma_j^0, y_k^0)}.
\end{equation}
where $\Delta_\gamma \equiv |\nabla \ln \tilde{v}|_0$, giving
\begin{equation}
\Delta_\gamma(\gamma_j^0, y_k^0) = \left\{ \sum_{z_i = \{ \gamma_j, y_k \}} \left( \frac{\partial \ln \tilde{v}}{\partial z_i} \right)^2 \right\}^{1/2}
\end{equation}
where notice that the sum over $z_i$ extends over both $\gamma_j, \gamma_k, \ldots$ and $y_j, y_k, \ldots$ parameters\footnote{In the last step in eq.(2.7) we used $\ln z_i$ as fundamental variables instead of $z_i$, where $z_i = \gamma_j, \gamma_k, \ldots; y_j, y_k, \ldots$; this ensures dimensionless arguments under the last Dirac delta function. This equation can however be written in terms of $\gamma_j$ and $y_k$ after prior normalization of $\gamma_j$ to some scale (like $v_0$, etc), to ensure they are dimensionless. In this case the expression of $\Delta_\gamma$ in eq. (2.10) is changed by replacing $\ln \gamma_j \to \gamma_j/v_0$ and $\ln y_k \to y_k$.}.
The result in (2.9) shows that to maximize $L_w$ one should maximize not the traditional likelihood $L$ (as done in present data fits), but its ratio to the quantity $\Delta_q$ that emerged above. Points in the parameter space with large $\Delta_q$ can reduce the overall $L_w$. The presence of $\Delta_q$ in the constrained $L_w$ is entirely due to imposing on the traditional likelihood $L$, of the constraint of fixing the EW scale in the presence of quantum corrections. Since this constraint encodes the physics of the naturalness problem, the quantity $\Delta_q$ in (2.9) can only be interpreted as a model-independent measure of naturalness i.e. fine-tuning. It is for this reason that in the following we often refer to $\Delta_q$ as “fine-tuning”. We stress that $\Delta_q$ is a quantity that was derived when computing $L_w$, and not an ad-hoc definition! $\Delta_q$ of (2.9) depends on SUSY parameters $\gamma_j$ and on nuisance variables like Yukawa couplings $y_k$. This answers the question on the parameters with respect to which one should actually compute fine-tuning.

Further, using common assumptions such as Gaussian distributions for the observables, then $\chi^2 = -\ln L$, in which case the relation in (2.9) can be rewritten as

$$\chi^2_w = \chi^2 + 2\ln \Delta_q.$$  \hspace{1cm} (2.11)

A SUSY model that fits well the data and also fixes the EW scale must have total $\chi^2_w/n_{df} \approx 1$. Until now one would only try to satisfy the weaker condition $\chi^2/n_{df} \approx 1$ associated with fixing observables other than $m_Z$, and which is not easily satisfied in many SUSY models.

Relation (2.11) for $\chi^2_w$ makes no distinction between tuning parameters to fit the EW scale or some other observable. Indeed, fixing the EW scale brings a $\chi^2$ “cost” ($2\ln \Delta_q$) for the model when $\Delta_q > 1$, just like any other observable. Eq.(2.11) also shows that we should actually minimize $\chi^2_w$, i.e. the sum of two contributions. A small $\Delta_q$ is preferable, in agreement with the traditional physical intuition (although it does not have to be minimized on its own). Finally, eq.(2.11) answers the old question of how to compare two models: one with “good” $\chi^2/n_{df}$ but “bad” fine-tuning and another with “not as good” $\chi^2/n_{df}$ but “not as bad” fine-tuning: one should simply minimize and compare their total $\chi^2_w/n_{df}$ which encodes the information of their likelihood to fit the data that includes the EW scale. In brief, naturalness is an intrinsic part of the likelihood to fit the data.

We can also derive a criterion for what is acceptable fine tuning in a SUSY model: we demand a good fit for the constrained likelihood, i.e. $\chi^2_w/n_{df} \approx 1$ or so. This requires that

$$\Delta_q < \exp(n_{df}/2)$$ \hspace{1cm} (2.12)

which should be comfortably respected (given the extra $\chi^2$-“cost” of other observables). We thus have an answer to the question regarding an upper bound for fine tuning, beyond which the model may not be regarded as “natural” anymore. This bound is general and independent of the model details other than its number of degrees of freedom $n_{df}$. In many popular SUSY models $n_{df}$ is of order 10, meaning that an acceptable upper bound is somewhere in the region $\Delta_q \ll \exp(5) \approx 150$. What happens if $\Delta_q \geq \exp(n_{df}/2)$? In that case $\chi^2_w/n_{df}$ becomes larger than unity even before fixing observables other than the EW scale itself! Obviously, for a given value of $\chi^2_w/n_{df}$ such as 2 or larger we have a poor fit and one is unable to fit the data and (fix) the EW scale at the same time. Then the model can be ruled out probabilistically.
The relation of $\chi^2_w$ to $\Delta_q$ provided a probabilistic interpretation for $\Delta_q$. This helps us appreciate better the significance of a given value of $\Delta_q$. Suppose we have $\Delta_q = 10^6$ so $\chi^2_w$ receives a correction from fixing the EW scale equal to $2 \ln 10^6 \approx 27.6$. What does this mean? From eq.(1.3), assuming that one particular observable is $\approx 5.25 \sigma$ deviations from its central value would bring a correction to total $\chi^2_w$ that is similar to that of $\Delta_q = 10^6$. By this argument we have the correspondence $\Delta_q \approx 10^6 \leftrightarrow 5.25 \sigma$, and similarly $\Delta_q \approx 1000 \leftrightarrow 3.5 \sigma$ and $\Delta_q \approx 100 \leftrightarrow 3 \sigma$, and so on. To see this in a broader context, remember that currently the disagreement between the muon anomalous magnetic moment in theory versus experiment is about $3 \sigma$ deviations [10]. According to the above correspondence this is similar to having $\Delta_q \approx 100$. We shall discuss shortly the numerical values of $\Delta_q$ in popular SUSY models.

3. How the Bayesian approach accounts for naturalness.

Here we show how to extend our previous results to the Bayesian approach to fit the data, based on [2] (section 2) reviewed here. In this approach, one computes the Bayesian probability (“evidence”) $p(\text{data})$ of a model, given the data. Here $p(\text{data})$ is an integral over the parameter space $\gamma$ (and eventually $\gamma$ too) of the product between the likelihood $L$ and the initial probability distributions $p(\gamma)$ (“priors”) for the parameters

$$ p(\text{data}) = \int d\gamma_1...d\gamma_n p(\gamma) L(\gamma,y,v,\beta) \tag{3.1} $$

This result for $p(\text{data})$ is used to compare the relative probability (viability) of different models.

To extend our results we simply use (3.1) with the replacement $L \rightarrow L_w$. Eq.(2.7) that imposes the constraint of fixing the EW scale, is multiplied by $p(\gamma)$ and integrated over $\gamma$. The result is

$$ p(\text{data}) = m^0_Z \int d\gamma_1...d\gamma_n p(\gamma) d\nu d(\tan \beta) $$
$$ \times \delta[v - \tilde{v}(\gamma,y,v,\beta)] \delta[\tan \beta - \tan \beta_0(\gamma,y,v)] \delta(m_Z - m^0_Z) L(\gamma,y,v,\beta) \tag{3.3} $$
$$ = v_0 \int d\gamma_1...d\gamma_n p(\gamma) \delta[v_0 - \tilde{v}(\gamma,y,v_0,\beta_0(\gamma,y,v_0))] L(\gamma,y,v_0,\beta_0(\gamma,y,v_0)) $$
$$ = \int_{\Sigma} dS_{\ln \gamma} \frac{1}{|\nabla_{\ln \gamma} \ln \tilde{v}|} L(\gamma,y,v_0,\beta_0(\gamma,y)) \gamma_1...\gamma_n p(\gamma), \quad \Sigma : v_0 = \tilde{v}(\gamma,y,v_0,\beta_0(\gamma,y,v_0))$$

This is just a global version in parameter space of the local result in (2.7), (2.9); in deriving this result, the priors are spectators. The integral above is over the surface $\Sigma$ defined by $v_0 - \tilde{v}(\gamma,y,v_0,\beta_0(\gamma,y,v_0)) = 0$. Note the emergence of the suppression factor $|\nabla_{\ln \gamma} \ln \tilde{v}| = \Delta_q$ where $\Delta_q$ is that given by eq.(2.10) but with the sum restricted to $\gamma$ parameters ($\nabla_{\ln \gamma}$ denotes the gradient computed wrt $\ln \gamma_j$ variables). If one integrated (3.3) over Yukawa (nuisance) parameters too,

---

2 A probabilistic interpretation for the inverse of $\Delta$ was suggested, on physical grounds, in [9].

3 One can also use that (the surface integral is over $S$ defined by $g = 0$):

$$ \int_{|z|} f(z_1,\ldots,z_n) dz_1\ldots dz_n = \int_{S_{g=0}} dS_{n-1} f(z_1,\ldots,z_n) \frac{1}{|\nabla_{\delta} g|} \tag{3.2} $$
then $\Delta_q$ is exactly that of eq.(2.10) [2]. The $1/\Delta_q$ factor under the integral of $p(\text{data})$ in the last line of (3.3) acts as an extra prior and is solely due to the constraint of fixing the EW scale via $\delta(m_Z - m'_Z)$, as also seen in the local case. By comparing this result for $p(\text{data})$, evaluated numerically for different models, one decides which model is more likely, while taking into account the effect of $1/\Delta_q$ associated with naturalness. Finally, eq.(3.3) simplifies further for logarithmic priors, $p(\gamma) = 1/(\gamma_1\gamma_2 \ldots \gamma_n)$.

Let us mention an equivalent form of the result in (3.3) which is

$$p(\text{data}) = \int dS_\gamma \frac{1}{|\nabla_\gamma \ln \bar{v}|} L[\gamma, y, v_0, \beta_0(\gamma, y, v_0)] p(\gamma)$$

(3.4)

where the gradient $\nabla_\gamma$ and $dS_\gamma$ are defined with respect to parameters $\gamma$ instead of $\ln \gamma$ used in (3.3). In this case $\Delta_q$ changes accordingly, $\Delta_q = |\nabla_\gamma \ln \bar{v}|$.

Unlike the “frequentist” approach discussed in the previous section, current data fits in the Bayesian approach already include the “cost” of naturalness in their results. Early attempts were based on a choice of priors that go like $1/\Delta$ where $\Delta$ was an ad-hoc definition for fine-tuning wrt a given parameter [11]. That meant that points that were badly fine-tuned wrt that particular choice of fine-tuning, were giving a smaller contribution to global $p(\text{data})$. More recently, a so-called “Jacobian factor” was identified to account for a similar effect [12, 13], starting from assumptions similar to those used here (see [13] for more details). It would be interesting to establish the exact relation of this Jacobian factor to the factor $1/\Delta_q$ that we identified above to account for naturalness.

This ends our discussion on the Bayesian approach and on how $\Delta_q$ emerges in this case.

4. The minimal value of $\Delta_q$ in supersymmetric models and its impact on $\chi^2_{\nu}/n_{df}$.

The next step in this analysis is to answer what values $\Delta_q$ of eq.(2.10) takes in current SUSY models. We then examine the impact it has on total $\chi^2_{\nu}/n_{df}$, see eq.(2.11). Can one satisfy condition (2.12)? To this purposes we use the results in [2]. For related studies see [15, 16, 17, 19, 20, 21, 22] and for recent data fits in SUSY models see [17, 18]. We use eq.(2.10) evaluated wrt to the SUSY parameters $\gamma$ only, listed below for each model considered; this restriction underestimates $\Delta_q$, so from this point of view our analysis is conservative. The models considered here are [23]:

- the constrained MSSM (CMSSM), of parameters $\gamma = \{m_0, m_{1/2}, \mu, A_0, B_0\}$, in a standard notation for soft masses of squarks and sleptons ($m_0$), gaugino ($m_{1/2}$), usual $\mu$ term ($\mu$), trilinear ($A_0$) and bilinear ($B_0$) SUSY-breaking terms.
- the NUHM1 model: this is a CMSSM-like model but with Higgs masses in the ultraviolet (uv) different from $m_0$, $m_{h_1}^{\text{uv}}$, $m_{h_2}^{\text{uv}} \neq m_0$, with parameters $\gamma = \{m_0, m_{1/2}, \mu, A_0, B_0, m_{h_1}^{\text{uv}}, m_{h_2}^{\text{uv}}\}$.
- the NUHM2 model: this is a CMSSM-like model with non-universal Higgs mass, $m_{h_1}^{\text{uv}} \neq m_{h_2}^{\text{uv}} \neq m_0$, with independent parameters $\gamma = \{m_0, m_{1/2}, \mu, A_0, B_0, m_{h_1}^{\text{uv}}, m_{h_2}^{\text{uv}}\}$.
- the NUGM model: this is a CMSSM-like model with non-universal gaugino masses $m_{\lambda_i}, i = 1, 2, 3$, with $\gamma = \{m_0, \mu, A_0, B_0, m_{\lambda_1}, m_{\lambda_2}, m_{\lambda_3}\}$.
- the NMSSM model: compared to the CMSSM, in this case there is an additional gauge singlet, so $\gamma = \{m_0, m_{1/2}, \mu, A_0, B_0, m_{1/2}, m_3\}$ where $m_3$ is the singlet soft mass.
A new approach to Naturalness.

Figure 1: $\Delta_q$ versus $M_{higgs}$ in various models, from ref.[2]; lightest grey (0) area: excluded by SUSY mass bounds; darker grey (1): excluded by $b \to s \gamma, B \to \mu^+ \mu^-$, $\delta \rho$; dark grey (2): excluded by condition $\delta a_\mu \geq 0$. Area inside the closed contour: allowed by data and with 2$\sigma$ deviation of $g - 2$: $\delta a_\mu \leq (25.5 + 2 \times 8) \times 10^{-10}$; $\delta a_\mu^{max}$ is shown colour encoded. Area outside closed contour: $\delta a_\mu^{max} \leq (25.5 - 2 \times 8) \times 10^{-10}$ ($2\sigma$). Only in the NUGM is one close to satisfying the $g - 2$ constraint within $2\sigma$, not too surprising given its non-universal gaugino masses. In all plots the dark matter relic density was computed [14] and can be fitted within $3\sigma$ [2]. For similar plots in NMSSM or GNMSM see [15] while for the GNMSM in the limit of a massive gauge singlet (integrated out) see plots in [16].

- the general NMSSM model (GNMSSM): this is an extension of the NMSSM with a bilinear superpotential term, $M S^2$, where $S$ is the singlet superfield. So we have one more parameter ($M$) in addition to the parameters of the NMSSM.

In all models a complete scan over the entire parameter space $\gamma$ is performed [2] including $\tan \beta$ with standard experimental constraints$^4$. The result for $\Delta_q$ (computed at two-loop using SOFTSUSY [24]), is plotted in Figure 1 as a function of the SM-like higgs mass $m_h$ (computed at two-loop leading log), for each model, except the NMSSM, GNMSSM cases. Notice the log scale in the plots (on OY axis), indicating an approximately exponential increase of $\Delta_q$ wrt $m_h$ [19]. We find it interesting that if one also computed $\Delta$ defined by (1.1), the plots have an identical form to those for $\Delta_q$, but only slightly shifted along OY axis, to a lower $\Delta$ by a factor between 1 and 2.

From Figure 1, after fixing the higgs mass to its measured value $m_h \approx 126$ GeV at the LHC [25] one immediately identifies in each model the minimal value of $\Delta_q$ obtained after scanning the whole parameter space. The results are summarized in Table 1 for minimal $\Delta_q$ and its corresponding $\delta\chi^2 = 2\ln\Delta_q$ contribution, for $m_h$ near the pre-LHC bound (115 GeV) and for the current value (126 GeV). The results in this table include additional numerical results for the NMSSM and also GNMSSM models for which we did not display the plots of $\Delta_q$ as a function of $m_h$, but used instead

$^4$For the experimental data used see table 1 of [2]. The muon $\delta a_\mu$ is not imposed but shown separately in Figures 1.
A new approach to Naturalness.

<table>
<thead>
<tr>
<th>Model</th>
<th>$n_p$</th>
<th>Approx</th>
<th>$\Delta q(115)$</th>
<th>$\delta \chi^2(115)$</th>
<th>$\Delta q(126)$</th>
<th>$\delta \chi^2(126)$</th>
<th>$n_{df}$</th>
<th>$\chi^2_w/n_{df}(126)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMSSM</td>
<td>5</td>
<td>2-loop</td>
<td>15</td>
<td>5.42</td>
<td>1800</td>
<td>14.99</td>
<td>9</td>
<td>2.66</td>
</tr>
<tr>
<td>NUHM1</td>
<td>6</td>
<td>2-loop</td>
<td>100</td>
<td>9.21</td>
<td>1500</td>
<td>14.63</td>
<td>8</td>
<td>2.83</td>
</tr>
<tr>
<td>NUHM2</td>
<td>7</td>
<td>2-loop</td>
<td>85</td>
<td>8.89</td>
<td>1300</td>
<td>14.34</td>
<td>7</td>
<td>3.05</td>
</tr>
<tr>
<td>NUGM</td>
<td>7</td>
<td>2-loop</td>
<td>15</td>
<td>5.42</td>
<td>1000</td>
<td>13.82</td>
<td>7</td>
<td>2.97</td>
</tr>
<tr>
<td>NMSSM</td>
<td>6</td>
<td>1-loop</td>
<td>12</td>
<td>4.97</td>
<td>&gt;200</td>
<td>10.59</td>
<td>8</td>
<td>2.32</td>
</tr>
<tr>
<td>GNMSSM</td>
<td>7</td>
<td>1-loop</td>
<td>12</td>
<td>4.97</td>
<td>27</td>
<td>6.59</td>
<td>7</td>
<td>1.94</td>
</tr>
</tbody>
</table>

Table 1: The correction $\delta \chi^2 \equiv 2\ln \Delta q(\gamma)$, the number of parameters $n_p$ and degrees of freedom $n_{df}$ in SUSY models, for $m_h \approx 115$ GeV corresponding to the pre-LHC bound and for $m_h \approx 126$ GeV. For 115 GeV, one had $\delta \chi^2/n_{df} < 1$, thus leaving the possibility of having $\chi^2_w/n_{df} \approx 1$. For the 126 GeV value, $\delta \chi^2/n_{df} > 1$ in all models above, except the GNMSSM where this is close to unity. This means that the requirement of a good fit ($\chi^2_w/n_{df} \approx 1$) is already violated by the fixing of the EW scale alone, before considering the additional $\chi^2$ cost of other observables. Assuming these give the usual $\chi^2/n_{df} \approx 1$ i.e. a good fit, adding the impact of $\delta \chi^2$ leads to the total $\chi^2_w/n_{df}$ displayed in the last column. $n_{df}$ may vary, depending on the exact number of observables fitted. The numerical values of $\Delta q$ are from [2] (first 4 models) and correspond to the plots in Figure 1, while for GNMSSM and NMSSM we used the results in [15] and [17].

The results of [15], [17]. Although not shown in the table, for all models $\delta \chi^2$ shown varies by $\approx 1$ for a similar sign variation of the higgs mass by 1 GeV.

The results in Table 1 show that $\Delta q$ has a significant impact on the overall quality of the fit (i.e. the value of $\chi^2_w/n_{df}$) since it brings corrections $\delta \chi^2/n_{df} > 1$ for models other than the GNMSSM. This means that the requirement of a good fit ($\chi^2_w/n_{df} \approx 1$) is already violated by the fixing of the EW scale alone, before considering the additional $\chi^2$ cost of other observables. So naturalness, encoded by the constraint of fixing the EW scale, has a very significant $\chi^2_w$ cost. Assuming (ideally) a good fit under current experimental constraints, i.e. $\chi^2/n_{df} \approx 1$, one finds that overall $\chi^2_w/n_{df} > 2.3$ in the models discussed. An exception is the GNMSSM model, where this result is improved mildly ($\approx 1.94$).

There is another interesting aspect in Table 1. While in some models $\Delta q$ at 126 GeV decreases relative to the CMSSM case, the correction $\delta \chi^2/n_{df}$ and total $\chi^2_w/n_{df}$ can actually increase (!), because of the stronger effect of decreasing $n_{df}$ (larger $n_p$). Thus, unless it is very significant, a reduction of $\Delta q$ due to new physics (more parameters) does not necessarily lead to a reduction of $\chi^2_w/n_{df}$ and a better fit. In GNMSSM the reduction of $\Delta q$ is significant enough and $\chi^2_w/n_{df}$ is improved relative to the CMSSM (for a recent GNMSSM study see [26]).

5. Conclusions

Our conclusions are best summarized by listing the old, unanswered questions in the introduction on the problem of fine-tuning in SUSY models, together with the answers that we found.

- Is there a relation between naturalness and the likelihood to fit the data, and if so, can we derive, on mathematical grounds, an expression for fine-tuning?
To answer this, one should recall the original goal of SUSY, central to the physical problem of naturalness (fine-tuning): to fix the EW scale to its measured value, in the presence of quantum corrections (while assuming the model is valid up to the Planck scale). If we regard this as a constraint and impose it on the current likelihood to fit the data the associated physical problem of naturalness (fine-tuning) should be captured by the mathematics that describes the constraint. We computed the “constrained” likelihood \( L_w \) by imposing this constraint on the likelihood \( L \) to fit the data (other than the EW scale) and found that \( L_w = L/\Delta q \). So one should maximize \( L_w \), not \( L \). The factor \( \Delta q \) is due to this constraint alone, so it encodes the mathematical effect of the naturalness; thus, it can only be regarded as a derived fine-tuning expression. Rather surprisingly, current data fits (in the frequentist approach) do no account for this (suppressing) effect on the likelihood, thus leading to more optimistic results.

- What are the parameters with respect to which \( \Delta q \) should be computed?

In the literature it is often unclear whether to include Yukawa couplings among these parameters. If one computes the constrained likelihood \( L_w \) as explained, there is no choice: these parameters include both the SUSY parameters and the nuisance variables (Yukawa couplings, etc) as well. One can maximize \( L_w \) wrt Yukawa (SUSY parameters fixed) and thus eliminate them. Alternatively, one can marginalize \( L_w \) over nuisance variables in which case \( \Delta q \) is computed with respect to the SUSY parameters only.

- What is an “acceptable” value for \( \Delta q \) identified above?

The relation between the likelihoods \( L_w = L/\Delta q \), has a correspondent in \( \chi^2_w = \chi^2 + 2 \ln \Delta q \) (assuming \( \chi^2 = -2 \ln L \)). So total \( \chi^2_w \) accounts for tuning the parameters to fix both the usual observables and the EW scale, and it makes no distinction between these tunings which are treated on equal footing, as it should be the case. In particular, electroweak scale (“fine”) tuning is nothing but \( \chi^2 \) “cost”! In other words, naturalness is an intrinsic part of the likelihood to fit the data that includes the EW scale. For a good fit one should then have a total \( \chi^2_w/n_{df} \approx 1 \) where \( n_{df} \) is the number of degrees of freedom. This implies a necessary bound \( \Delta q < \exp(n_{df}/2) \), which should be comfortably respected (i.e. \( \Delta q \ll \exp(n_{df}/2) \), since often \( \chi^2/n_{df} \) is itself close to unity). For the popular models discussed this means \( \Delta \ll 90 \) (\( n_{df} \leq 9 \)).

- Can we rule out supersymmetric models based on the size of the quantity \( \Delta q \) alone?

The answer is yes; when \( \Delta q \gg \exp(n_{df}/2) \) then the total \( \chi^2_w/n_{df} \) is so poor that the model cannot fit the data anymore and is ruled out. Note however that the addition of new physics, that can actually be beyond the LHC reach, could reduce \( \Delta q \), as seen for the “general NMSSM” (GNMSSM) model.

- Does \( \Delta q \) have a physical meaning?

The relation \( \chi^2_w = \chi^2 + 2 \ln \Delta q \) gives \( \Delta q \) a probabilistic interpretation. A value of say \( \Delta q = 10^6 \) has an effect on \( \chi^2_w \) similar to one observable that contributes to \( \chi^2_w \) being \( \approx 5.25\sigma \) deviations from the central value. By the same rule, \( \Delta q = 100 \) is equivalent to a deviation of \( \approx 3\sigma \), which is the order of deviation of the muon magnetic moment.

- How can we compare two models and decide which is better: one that has a very good \( \chi^2 \) fit in the traditional sense, but significant fine tuning and one that has a nearly-as-good \( \chi^2 \) fit but less fine-tuning?
A new approach to Naturalness.

The question arises since accurate data fits usually report separately $\chi^2/n_{df}$ and fine tuning $\Delta$ computed according to some definition. We showed that to compare such models, one should actually compare their total $\chi^2_w/n_{df}$, and this answers the question above. This is the ‘frequentist’ case. In the Bayesian approach one compares instead the “constrained” probabilities of the two models $p(\text{data})$ that contain an additional $1/\Delta_q$ factor under their integral. This factor again emerges (as an extra prior) from the naturalness constraint of fixing the EW scale at quantum level.

- How can one reduce $\Delta_q$ and also $\chi^2_w/n_{df}$ to improve the data fits of SUSY models?
  The easiest way to reduce $\Delta_q$ is to increase the effective quartic higgs coupling $\lambda$, whose small value is at the origin of large $\Delta_q$ in most SUSY models\(^5\). One can increase $\lambda$ by new physics (supersymmetric!) beyond the MSSM higgs sector\(^{[27]}\) and possibly beyond the LHC reach too! The new physics can be represented by additional massive states that couple to the higgs sector: singlets, SU(2) doublets, additional massive $U(1)$ boson, etc. However, adding new physics increases the number of parameters (reduces $n_{df}$). Therefore, unless the reduction in $\Delta_q$ is very significant (to $\Delta_q < 10$ or so), the total $\chi^2_w/n_{df}$ can actually increase due to the reduced $n_{df}$ and to the mild (log) dependence of $\chi^2_w$ on $\Delta_q$. Of all models discussed (MSSM or NMSSM-like) only the GNMSSM model had $\Delta_q < 10$; this model emerges as the one with the best chance of a good fit ($\chi^2_w/n_{df}$) in the sense discussed above.

- Can one compare the relative viability of two models based on the values of $\Delta_q$ alone?
  The answer is negative. It is the total $\chi^2_w/n_{df}$ that matters when comparing models, not $\Delta_q$ alone. This is because one model can have a smaller $\Delta_q$ (by adding more parameters) but a larger $\chi^2_w/n_{df}$ than in the second model (see previous point).

To conclude, naturalness in SUSY models is an intrinsic part of the likelihood to fit the experimental data that includes the EW scale. This helps address many open questions on this issue, as shown above. The results above indicate that for some popular SUSY models the value obtained for the corrected likelihood (i.e. $\chi^2_w/n_{df}$) already questions their viability, except perhaps the case of GNMSSM. This can be investigated further by performing accurate data fits that evaluate the new $\chi^2_w/n_{df}$ that we identified.

Acknowledgements: The author thanks Hyun Min Lee, Myeonghun Park and Graham Ross, for their collaboration on the original papers reviewed here (ref. \([1,2]\)). This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS - UEFISCDI, project number PN-II-ID-PCE-2011-3-0607.

References


\(^5\)For details on this issue, see first reference in \([19]\).
A new approach to Naturalness.


A new approach to Naturalness.


A new approach to Naturalness.


