PHYSICAL REVIEW D **89,** 095007 (2014)

SUSY naturalness without prejudice

D. M. Ghilencea*

Theoretical Physics Department, National Institute of Physics and Nuclear Engineering (IFIN-HH) Bucharest, P.O. Box MG-6 077125, Romania
CERN Theory Division, CH-1211 Geneva 23, Switzerland

(Received 12 March 2014; published 8 May 2014)

Unlike the Standard Model (SM), supersymmetric models stabilize the electroweak (EW) scale v at the quantum level and predict that v is a function of the TeV-valued SUSY parameters (γ_α) of the UV Lagrangian. We show that the (inverse of the) covariance matrix of the model in the basis of these parameters and the usual deviation $\delta\chi^2$ (from $\chi^2_{\rm min}$ of a model) automatically encode information about the "traditional" EW fine-tuning measuring this stability, provided that the EW scale $v \sim m_Z$ is indeed regarded as a function $v = v(\gamma)$. It is known that large EW fine-tuning may signal an incomplete theory of soft terms and can be reduced when relations among γ_α exist (due to GUT symmetries, etc.). The global correlation coefficient of this matrix can help one investigate if such relations are present. An upper bound on the usual EW fine-tuning measure ("in quadrature") emerges from the analysis of the $\delta\chi^2$ and the s-standard deviation confidence interval by using $v = v(\gamma)$ and the theoretical approximation (loop order) considered for the calculation of the observables. This upper bound avoids subjective criteria for the "acceptable" level of EW fine-tuning for which the model is still "natural."

DOI: 10.1103/PhysRevD.89.095007 PACS numbers: 12.60.Jv, 12.60.-i

I. A NATURAL TEST FOR SUSY MODELS

A. Introduction

Unlike the Standard Model (SM), supersymmetric models (MSSM, NMSSM, etc.) stabilize the electroweak (EW) scale v at the quantum level and make a prediction for it. The combined Higgses EW VEV v, or the Z boson mass $m_Z \propto v$, is a derived quantity that is a function $v = v(\gamma)$, where γ_α denote the following Lagrangian UV parameters: the TeV-valued soft masses, soft couplings, and μ. The function $v = v(\gamma)$ is obtained from the minimization of the Higgs potential and from the fact that the Higgs couplings are fixed, to lowest order, by gauge interactions (unlike the SM case where the Higgs self-coupling is arbitrary). Whether this prediction successfully recovers the measured value $m_Z^0 \approx 91.187$ GeV of the Z boson is a natural test of SUSY. This regards m_Z as an observable to be fitted. This view, adopted here, remains true to the original motivation of SUSY.

This naturalness test received much attention from theorists who long ago introduced fine-tuning [1] measures [2,3] for it. However, precision data fits [4] often prefer to keep m_Z as a fixed input (constant) equal to m_Z^0 rather than as an observable as well that depends on the Lagrangian parameters and then no χ^2 "cost" to fit m_Z is usually reported.[1] At the same time, some data fits still report the EW fine-tuning in which $v \sim m_Z$ is indeed a function $v = v(\gamma)$, often giving a large variation of m_Z (about its fixed input value). It is not clear to us how results relying on two different assumptions [v fixed constant or a function $v = v(\gamma)$] can be combined consistently to draw a clear conclusion. This is due to the following questions (particularly Q_3 below):

Q_1: Is the likelihood to fit the observable m_Z related to the EW fine-tuning "cost"?

Q_2: What is the link of the total likelihood to fit a set of observables to EW fine-tuning?

Q_3: How do we compare a model with a good fit (of m_Z and other data) and "large" EW fine-tuning to one with nearly as good a fit but less fine- tuning (for the same data)?

Question Q_1 is even more compelling given that we do not know what an "acceptable" value of the EW fine-tuning is. One can address $Q_{1,2,3}$ by regarding m_Z as an observable and by using standard tools to test models, as discussed in the likelihood approach [5,6] or earlier in the Bayesian case [7–9]. These works suggested that EW fine-tuning is related to the likelihood (χ^2) or the posterior probability to fit the data that includes the observable m_Z. In this letter we explore this relation further, using a different approach, and a χ^2 (frequentist) analysis. As detailed below, we study a possible relation of the covariance matrix of the model to the EW fine-tuning. This connection is not examined in the literature even though each of these aspects were studied in the past separately. This is the main purpose of this work.

As an example, consider the MSSM case with the Higgs potential minimum conditions in a standard notation, fixing the EW scale ($v \propto m_Z$) and tan β (or B),

*dumitru.ghilencea@cern.ch

[1] This is partly due to technical and historical reasons from pre-SUSY fits where m_Z is an input; with m_Z output also the parameter scans would be very ineffective, as many points are ruled out by Z mass.

$$\frac{m_Z^2}{2} = -\mu^2 + \frac{m_1^2 - m_2^2 \tan^2\beta}{\tan^2\beta - 1} + \ldots$$
$$2m_3^2 = (m_1^2 + m_2^2 + 2\mu^2)\sin 2\beta + \cdots. \qquad (1)$$

The dots stand for quantum corrections to the quartic Higgs couplings and $m_{1,2,3}$ are one-loop soft masses. In precision data fits, one traditionally replaces m_Z "by hand" by its measured mass m_Z^0 to fix μ instead (as a function of remaining γ_α), or "fine-tune" the independent γ_α to reproduce m_Z^0. Ultimately, this amounts to using a Dirac δ distribution for the observable m_Z. Further [5,6],

$$\delta(1 - m_Z/m_Z^0) = \frac{1}{\Delta}\delta[n^\alpha(1 - \gamma_\alpha/\gamma_\alpha^0)],$$
$$\gamma = \{m_0, \mu, A_0, B_0, M_1, M_2, M_3, \ldots\}. \qquad (2)$$

Here $m_Z \sim v$ is a function of parameters γ_α as shown in Eq. (1), with γ_α^0 components of the set γ that respect the condition $m_Z(\gamma_\alpha^0) = m_Z^0$ and n^α are the components of the normal to the surface defined by this equation; finally,

$$\Delta \equiv \left\{\sum_\alpha \left(\frac{\partial \ln m_Z(\gamma)}{\partial \ln \gamma_\alpha}\right)^2_{\gamma=\gamma^0}\right\}^{1/2}. \qquad (3)$$

Since Δ emerged from fixing the EW scale condition ($m_Z = m_Z^0$) associated with fine-tuning, we can only interpret it as a derived, unique measure of fine-tuning (not chosen).

The message is that the distribution in the lhs of Eq. (2) chosen as a likelihood (for observable m_Z) somehow "knows" about the EW fine-tuning Δ. This may not be too surprising, but it hints to a deeper connection. The EW scale $v = v(\gamma)$ enters in many observables and also correlations among these can be present. Therefore, they can also have significant individual fine-tunings associated. We usually refer to fine-tuning of m_Z, but one can similarly discuss, for example, the fine-tuning of the Higgs mass, etc, since this is also closely related to the hierarchy problem. Then a relation similar to Eq. (2) can be present between each observable and the set of parameters γ. This seems to suggest an underlying connection of the likelihood (or usual χ^2) associated with a set of EW observables to their fine-tuning and to the distribution of the parameters of the model about their central values (of maximal likelihood or min χ^2); these could be connected as in the example above, by some general form of Δ. These comments indicate an affirmative answer to question Q_1. For a more detailed discussion see [5,6].

In the following we explore some of these issues further. We consider that

(1) in practice one does not have Dirac distributions for m_Z or other observables,
(2) correlations can exist between m_Z and other observables: m_h, m_H, etc.,
(3) and other observables can also depend on the EW scale $v = v(\gamma)$ (and/or on γ).

With these in mind we study the link of the likelihood to fit the data and its deviation from the maximal value, to the EW fine-tuning (Δ above); equivalently, in a chi-square language, we study the link of the deviation $\delta\chi^2$ from the minimal value $\chi^2_{\min}$, to the EW fine-tuning, Δ. We find that, just as the likelihood to fit m_Z^0 contains Δ [Eq. (2)], in the general case [with (1), (2), (3)] of the total likelihood, EW fine-tuning is automatically present in the covariance matrix $\tilde{M}$ in the basis of the fundamental parameters (γ_α) and thus also in the deviation $\delta\chi^2$, provided that we regard v as a function $v = v(\gamma)$, (as predicted by SUSY). Thus the matrix $\tilde{M}$ has a more fundamental role than the EW fine-tuning and contains information about the stability of the EW scale under UV variation of SUSY parameters. This view also ends a long-held distinction between EW fine-tuning (to fit m_Z) and that to fit other observables (m_h, m_H, etc) that also depend on v and that are thus ultimately linked to the hierarchy problem (just as $m_{Z,W}$ are).

B. The link of the likelihood (χ^2) to EW fine-tuning

Consider a model with a number of observables $\mathcal{O}_i$ ($i = 1, 2, \ldots, n$) of central experimental values $\mathcal{O}_i^0$ fitted using a set of SUSY parameters[2] γ_α, ($\alpha = 1, 2.., s$) that enter in the Lagrangian with $s < n$ and $n_{df} = n - s$ (n_{df}: number of degrees of freedom). The general form of Eq. (1) of the two minimum conditions of the scalar potential is

$$v = v(\gamma; \beta), \qquad \tan\beta = \tan\beta(\gamma, v)$$
$$\gamma_\alpha : m_0, \mu, A_0, B_0, M_1, M_2, M_3, \ldots \qquad (4)$$

leading to[3] $v = v(\gamma, \beta(\gamma))$; to simplify the notation, hereafter we refer to this dependence as $v(\gamma)$. From a Taylor series for $m_Z \sim v$ about a particular point γ^0,

$$m_Z = m_Z(\gamma^0) + \left(\frac{\partial m_Z}{\partial \gamma_\alpha}\right)_{\gamma=\gamma^0}(\gamma_\alpha - \gamma_\alpha^0) + \cdots. \qquad (5)$$

Assume for a moment that γ^0 is a solution[4] to $m_Z(\gamma) = m_Z^0$, with $m_Z^0 \approx 91.187$ GeV. Eq. (5) can be rewritten as

$$\frac{m_Z - m_Z(\gamma^0)}{m_Z(\gamma^0)} = \Delta n^\alpha \frac{\gamma_\alpha - \gamma_\alpha^0}{\gamma_\alpha^0} + \mathcal{O}((\gamma_\alpha - \gamma_\alpha^0)^2),$$
$$n^\alpha \equiv \frac{1}{\Delta}\left(\frac{\partial \ln m_Z}{\partial \ln \gamma_\alpha}\right)_{\gamma=\gamma^0}, \qquad (6)$$

[2] γ can include nuisance variables (Yukawa couplings, etc.) eliminated later by integration/profiling.

[3] One often replaces β by some other parameter, like B_0, with no implications below.

[4] Note that with an overconstrained set of parameters γ_α by the set of observables ($n_{df} > 0$), γ^0 above should actually denote the set that minimizes the global χ^2 of all observables, including m_Z, rather than the solution to $m_Z(\gamma) = m_Z^0$ (see Sec. I. D). Then $m_Z(\gamma^0)$ does not reproduce the central measured value, but a value that should be within few standard deviations from it (say $2\sigma_z$); then the difference in the lhs of (6) should again be understood as $2\sigma_z$.

where n^α denote the components of the normal to the surface $m_Z(\gamma^0) = m_Z^0$, with $n_\alpha n^\alpha = 1$. Here Δ is given in Eq. (3) with Eq. (4).

For the recently measured $m_h \approx 126$ GeV [10], we know that minimal Δ, upon varying all allowed parameters γ_α and $\tan\beta$, is $\Delta \sim \mathcal{O}(1000)$ [9] in the MSSM-like models with different boundary conditions for the soft terms. Note however that in the general version of the NMSSM (GNMSSM) a value of $\mathcal{O}(20)$ is still possible [11]. Then to keep m_Z of Eq. (6) within $2\sigma_z$ of its central measured value[5] m_Z^0 or equivalently $\delta m_Z/m_Z^0 = 4.6 \times 10^{-5}$, one must keep each parameter γ_α within an order of $\delta\gamma_\alpha/\gamma_\alpha^0 \approx 4.6 \times 10^{-8}$ of its value γ_α^0. For simplicity, assume that all parameters other than one of them, say μ, are fixed and take for example $\mu_0 \approx 1$ TeV, so it would mean $\delta\mu = 46$ keV. Such accuracy $\delta\mu$ or $\delta\gamma_\alpha$ needed to compensate the large Δ in the rhs of (6), can be reached by a fine scan of the parameter space; but a deviation by few keV deviates m_Z by more than $2\sigma_z$ from its measured value, if $\Delta \approx 1000$. So a good stability of the χ^2 fit of m_Z and a large Δ, are not easily compatible (also recall that γ_α^0 are (over)constrained to keep under control χ^2 due to other observables). This problem reflects a relation of the "traditional" fine-tuning of Eq. (3) and the associated χ^2 "cost" to fit m_Z or other observables that depend on $v(\gamma)$. If one insists on keeping m_Z a fixed input number (equal to m_Z^0), the problem remains because in the above discussion one replaces m_Z by any another observable that depends on $v = v(\gamma)$. Therefore, the relation of Δ to χ^2 is important and usually overlooked in the literature.

If relations among initial γ_α exist, dictated for example by UV symmetries (SU(5), etc), they can reduce[6] Δ [9,13]. So a large Δ can simply be a sign of our ignorance of the UV physics, telling us that our theory of soft terms is inappropriate or incomplete. Aside from this possibility, dramatic fine-tuning of γ_α could be "natural" if γ_α are related to a fundamental constant of nature, whose accurate determination is crucial for the theory.

There are, however, limits to how much one can "fine-tune" γ_α in a given loop order. Indeed, γ_α^0 are determined from the condition of minimizing total χ^2 computed using a theoretical calculation of the observables in a fixed loop-order. This calculation is affected by an error from ignored higher loops (an example is the 2–3 GeV theoretical error of the Higgs mass at two-loop [14]). So the perturbation theory alone inevitably introduces a theoretical error $\sigma_{\rm th}$ to each γ_α^0. Then only points with $\sigma_{\gamma_\alpha} > \sigma_{\rm th}$ are actually relevant[7]; with this bound, from Eq. (6) Δ then has an upper bound if one insists to keep m_Z within say $2\sigma_z$ from its central value. So $\delta\chi^2$ and Δ are related.

The above discussion can be extended to all observables that depend on the EW scale $v(\gamma)$; then a relation between each observable and the amount of tuning of γ_α is present, like in Eq. (6), with a similar connection to their χ^2 contribution.[8] It is then natural to expect a more general connection between the total χ^2 (or more generally, the likelihood) of all observables including m_Z and their fine-tuning with respect to γ_α. The generalisation of σ_γ and σ_z discussed above is the covariance matrix, therefore the latter could be the missing link in this connection (see later, Sec. I. D).

C. Fixing the EW scale and the relation to Δ

Let us denote by $L(\mathcal{O}|\gamma)$ the total likelihood to fit some observables $\mathcal{O}_i$ other than m_Z. We impose on this likelihood the condition of fixing the EW scale that motivated SUSY, that we regard just as a condition to fit $m_Z \sim v$ of Eq. (1) to its central measured value. We first take a Dirac delta distribution for m_Z. Assuming that we can factorize this distribution[9] from $L(\mathcal{O}|\gamma)$ of other data, the total likelihood that accounts also for fixing the EW scale is $L(\mathcal{O}, m_Z^0|\gamma) = L(\mathcal{O}|\gamma) L(m_Z^0|\gamma)$ with [5,6]

$$L(m_Z^0|\gamma) = \delta\left(1 - \frac{m_Z(\gamma)}{m_Z^0}\right) = \frac{1}{\Delta}\delta(n^\alpha(1 - \gamma_\alpha/\gamma_\alpha^0)). \quad (7)$$

In the last step we used Eq. (6), with Δ as in Eq. (3). The argument of the Dirac δ function ensures that one must be on the surface predicted in SUSY by the minimum condition $m_Z(\gamma^0) = m_Z^0$ giving $n^\alpha(1 - \gamma_\alpha/\gamma_\alpha^0) = 0$, or $\gamma_\alpha = \gamma_\alpha^0$. For a detailed discussion and interpretation of Eq. (7) see [5,6].[10,11] Further, it is illustrative to go beyond the Dirac δ used in Eq. (7), so one can take

[5] as usually done in the data fits for any observable.

[6] Δ can also be reduced by "new physics" in the Higgs sector that can increase the Higgs mass [11,12].

[7] A naive estimate of $\sigma_{\rm th}$ can be, at three-loop order, $1/(16\pi^2)^3$, which is larger than 4.6×10^{-8} mentioned. More correctly, $\sigma_{\rm th}$ of each parameter is found numerically from the error of ignored loops in the theoretical value of the observables (such as that of m_h mentioned) that depend on that parameter.

[8] In this case γ_α^0 will correspond to the maximal likelihood point.

[9] We relax this assumption in Sec. I. D.

[10] As a side remark, Eq. (7) can be formally integrated in one general direction (combination of γ_α) and rewritten as [5]: $\tilde{L}(m_Z^0|\gamma) = (1/\Delta)|_{\gamma_\alpha = \gamma_\alpha^0}$. So $1/\Delta$ that emerged is the likelihood "cost" of respecting the SUSY condition of fixing the EW scale and is part of total $L(\mathcal{D}, m_Z^0|\gamma)$ to fit the data that includes m_Z^0. A similar interpretation was noticed on phenomenological grounds in [15]. With this, one can provide a very simple estimate of an upper bound on Δ. The contribution to total χ^2 due to m_z alone (say $\delta\chi^2$) equals, according to the last equation $\delta\chi^2 = -2\ln\tilde{L} = -2\ln(1/\Delta)$ (under some assumptions). By demanding a "good fit" i.e. that total $\chi^2/n_{df} \approx 1$, (χ^2 includes contributions from other observables) one has $\Delta < \exp(n_{df}/2)$ [5,6], which with usual $n_{df} \sim 10$ gives an upper value $\Delta \sim 100$. One objection to this approach is that $\tilde{L}$ is not normal which affects the goodness of the fit criterion $\chi^2/n_{df} \sim 1$.

[11] The minimal value of Δ (with all γ_α allowed to vary), grows with the Higgs mass $\Delta \sim \exp(m_h/\mathrm{GeV})$ (see Figs. 1–8 and 13–16 in [9], also [16]). As a result an error $\delta m_h = 2 - 3$ GeV that is the theoretical uncertainty of m_h prediction [14] brings an uncertainty factor $\approx \exp(3) \approx 20$ for Δ and accordingly for its $\delta\chi^2$ effect. This uncertainty means that $\Delta \approx 20$ and $\Delta \approx 400$ can be seen as equally "acceptable." Thus the results [5,6] should be regarded as a general estimate rather than a strict criteria of viability.

$$L(m_Z^0|\gamma) = \frac{m_Z^0}{\sqrt{2\pi}\sigma_z}\exp\left(-\frac{m_Z^{02}}{2\sigma_z^2}(m_Z/m_Z^0-1)^2\right), \quad (8)$$

which recovers the lhs of Eq. (7) when $\sigma_z \to 0$. From the Taylor expansion in Eq. (5) about γ_α^0 one finds, to a first approximation, with Δ as in Eq. (3)

$$L(m_Z^0|\gamma) = \frac{1}{\Delta}L(\gamma) + \cdots$$
$$\Rightarrow L(\mathcal{O}, m_Z^0|\gamma) = \frac{1}{\Delta}L(\gamma)L(\mathcal{O}|\gamma) + \cdots. \quad (9)$$

The first equation is similar to Eq. (7). $L(\gamma)$ is the associated normalized distribution of the output values of parameters γ; if all γ are fixed to γ^0 except one of them (γ_α), then

$$L(\gamma) = \frac{1}{\sqrt{2\pi}\sigma_{\gamma_\alpha}}\exp\left(-\frac{1}{2\sigma_{\gamma_\alpha}^2}(\gamma_\alpha/\gamma_\alpha^0-1)^2\right) \quad (10)$$

with estimated

$$\sigma_{\gamma_\alpha} = \frac{1}{\Delta}\frac{\sigma_z}{m_Z^0}\gamma_\alpha^0; \quad \text{if } \sigma_{\gamma_\alpha} \geq \sigma_{\rm th} \Rightarrow \Delta \leq \frac{\sigma_z}{\sigma_{\rm th}}\frac{\gamma_\alpha^0}{m_Z^0}. \quad (11)$$

Δ that emerged in (9), (11) is the sole consequence of the condition of fixing the EW scale ($m_Z = m_Z^0$) that is usually associated with fine-tuning, so it is a unique, derived measure (of fine-tuning) from this constraint. Δ also relates the normalized likelihood (of m_Z) to the normalized Gaussian distribution of γ_α, about the central value γ_α^0; such relation is more generic (see later). Since Δ enters in the expression of σ_{γ_α}, this suggests again that in more complex cases the generalization of σ_{γ_α}, the error matrix, could be related to the EW fine-tuning. Finally, if one demands $\sigma_{\gamma_\alpha} \geq \sigma_{\rm th}$ (from the loop-order accuracy $\sigma_{\rm th}$ that affects γ_α^0) an upper bound on Δ emerges in Eq. (11). This is actually a strong bound, even assuming $\sigma_{\rm th} \sim \sigma_z$, then $\Delta \leq \mathcal{O}(10)$ for TeV valued γ's.

D. The general case: More observables, correlations, and fine-tuning

So far we ignored the correlations of m_Z with other observables (for example with the loop-corrected Higgs masses m_h, m_H) or the fact that other observables also depend on the EW scale. We include these effects to examine the relations of the total likelihood (χ^2) of these observables, its deviation from its maximal (min χ^2) value, and of the covariance matrix, to the EW fine-tuning. The study is restricted to Gaussian distributions for observables so is equivalent to a simple χ^2 analysis with $\chi^2 = -2\ln L$; the analysis can be extended to general likelihoods.

Consider the observables $\mathcal{O}_j$, $(j = 1, 2 \ldots n)$, of experimental central values $\mathcal{O}_j^0$, and to simplify the notation we now assume that m_Z is also one of them, $\mathcal{O}_n = m_Z$; they are functions of γ_α, $(\alpha = 1, 2, \ldots, s)$, so $\mathcal{O}_i = \mathcal{O}_i(\gamma, v(\gamma))$, with $v(\gamma)$ as in Eq. (4). The total likelihood $L(\mathcal{O}|\gamma)$ due to all $\mathcal{O}_i$ must then be maximized with respect to the SUSY parameters γ_α. It is convenient to work with a dimensionless form of this likelihood,[12]

$$L(\mathcal{O}|\gamma) = (2\pi)^{-\frac{n}{2}}(\det M)^{-\frac{1}{2}}\exp[-1/2 u_i (M^{-1})_{ij} u_j],$$
$$u_i \equiv \mathcal{O}_i/\mathcal{O}_i^0 - 1, \quad (12)$$

with a (dimensionless) covariance matrix $M_{ij} = \rho_{ij}\sigma_i\sigma_j/(\mathcal{O}_i^0\mathcal{O}_j^0)$, $\rho_{ij} = \rho_{ji}$ denote the correlation coefficients, with $\rho_{ii} = 1$. Let γ^0 denote the solution of the condition to maximize $L(\mathcal{O}|\gamma)$. Although not appropriate for high precision numerical studies, to illustrate the main idea a Taylor expansion can be used

$$\mathcal{O}_i(\gamma) = \mathcal{O}_i(\gamma^0) + (\gamma_\alpha - \gamma_\alpha^0)\left(\frac{d\mathcal{O}_i}{d\gamma_\alpha}\right)_{\gamma=\gamma^0} + \cdots, \quad (13)$$

then

$$L(\mathcal{O}|\gamma) = \frac{\kappa}{\Delta}L(\gamma) + \cdots, \quad (14)$$

where κ is a constant and[13]

$$\Delta \equiv [\det M \det \tilde{M}^{-1}]^{\frac{1}{2}}, \qquad \tilde{M}^{-1} \equiv \mathcal{J}^T M^{-1}\mathcal{J},$$
$$\mathcal{J}_{i\alpha} \equiv \frac{1}{\mathcal{O}_i^0}\left[\frac{d\mathcal{O}_i}{d\ln\gamma_\alpha}\right]_{\gamma=\gamma^0} \quad (15)$$

$\tilde{M}_{ij}$ is a $s \times s$ matrix, $\mathcal{J}_{i\alpha}$ is a $n \times s$ matrix, and

$$L(\gamma) \equiv (2\pi)^{-\frac{s}{2}}(\det\tilde{M})^{-\frac{1}{2}}\exp[-1/2\tilde{\gamma}_\alpha\tilde{M}_{\alpha\beta}^{-1}\tilde{\gamma}_\beta],$$
$$\tilde{\gamma}_\alpha \equiv \gamma_\alpha/\gamma_\alpha^0 - 1. \quad (16)$$

$L(\gamma)$ is the normalized distribution of γ_α about central γ_α^0 that maximizes it and contains correlations. Equation (14) has similarities to Eqs. (7), (9).

Let us examine the matrix $\tilde{M}$ and assume for simplicity that M_{ij} is diagonal, then

$$\tilde{M}_{\alpha\beta}^{-1} = \sum_{i=1}^n\left\{\left(\frac{d(\mathcal{O}_i/\sigma_i)}{d\ln\gamma_\alpha}\right)\left(\frac{d(\mathcal{O}_i/\sigma_i)}{d\ln\gamma_\beta}\right)\right\}_{\gamma=\gamma^0},$$
$$\alpha, \beta = 1, 2, \ldots s. \quad (17)$$

[12]The "dimensionful" form of the total likelihood is $\mathcal{L}(\mathcal{O}|\gamma) = (2\pi)^{-n/2}(\det K)^{-1/2}\exp[-1/2(\mathcal{O}_i - \mathcal{O}_i^0)(K^{-1})_{ij}(\mathcal{O}_j - \mathcal{O}_j^0)]$; K is the dimensionful covariance matrix, $K_{ij} = \sigma_i\sigma_j\rho_{ij}$; $\rho_{ij} = \rho_{ji}$ account for correlations; $\rho_{ii} = 1$. $\mathcal{L}$ is equal to L used in the text up to a constant, $L(\mathcal{O}|\gamma) = |\mathcal{O}_1^0\mathcal{O}_2^0 \ldots \mathcal{O}_n^0| \times \mathcal{L}(\mathcal{O}|\gamma)$.

[13]In a χ^2 language: $\kappa = (2\pi)^{-n_{df}/2}\exp(-\chi^2_{\rm min}/2)$ with $\chi^2_{\rm min} = u_i(\gamma^0)M_{ij}^{-1}u_j(\gamma^0)$ and $\delta\chi^2 = \chi^2 - \chi^2_{\rm min}$ with $\delta\chi^2 = -2\ln[L(\mathcal{O}|\gamma)/L(\mathcal{O}|\gamma^0)] = -2\ln[L(\gamma)/L(\gamma^0)]$.

This expression shows the relevance of the variations of $\mathcal{O}_i$ "normalized" to their σ_i, which is somewhat expected on physical grounds. Further, $\mathcal{O}_i$ are functions of $v(\gamma)$, $\mathcal{O}_i = \mathcal{O}_i(\gamma, v(\gamma))$, which is relevant in establishing the relation of this matrix to the traditional fine-tuning. As a result of this dependence, the matrix $\tilde{M}^{-1}$ contains new terms,

$$\tilde{M}^{-1}_{\alpha\beta} = \tilde{M}^{-1}_{\alpha\beta}|_{v=\mathrm{const}} + \sum_{i=1}^{s}\left\{\left(\frac{\partial\mathcal{O}_i/\sigma_i}{\partial\ln v}\right)^2\left(\frac{\partial\ln v}{\partial\ln\gamma_\alpha}\right)\left(\frac{\partial\ln v}{\partial\ln\gamma_\beta}\right)\right\}_{\gamma=\gamma^0} + \sum_{i=1}^{s}\left\{\left(\frac{\partial\mathcal{O}_i/\sigma_i}{\partial\ln v}\right)\left(\frac{\partial\ln v}{\partial\ln\gamma_\alpha}\right)\left(\frac{\partial\mathcal{O}_i/\sigma_i}{\partial\ln\gamma_\beta}\right) + (\alpha\leftrightarrow\beta)\right\}_{\gamma=\gamma^0}, \tag{18}$$

which are not present in the traditional approach of numerical data fits in which v is actually a constant (fixed input) [4]. So each entry $\tilde{M}^{-1}_{\alpha\beta}$ automatically contains the EW fine-tuning represented by the partial derivatives of v with respect to $\gamma_\alpha, \gamma_\beta$, from all observables that depend on v.

This is an interesting result and suggests it is worth studying other properties of $\tilde{M}$. First, the trace

$$\mathrm{Tr}\tilde{M}^{-1} = \sum_{i=1}^{n}\sum_{\alpha=1}^{s}\left(\frac{d\mathcal{O}_i/\sigma_i}{d\ln\gamma_\alpha}\right)^2_{\gamma=\gamma^0} = \sum_{i=1}^{n}\left(\frac{\partial\mathcal{O}_i/\sigma_i}{\partial\ln v}\right)^2_{\gamma=\gamma^0}\times\underbrace{\sum_{\alpha=1}^{s}\left(\frac{\partial\ln v}{\partial\ln\gamma_\alpha}\right)^2_{\gamma=\gamma^0}}_{\Delta^2} + \cdots, \tag{19}$$

contains terms proportional to the traditional EW fine-tuning (second sum above), with contributions from all $\mathcal{O}_i$ that depend on $v(\gamma)$. This correction is also missed if v is a constant while retaining only the explicit dependence of $\mathcal{O}_i$ on γ. These results can also be examined for the single observable case, such as m_Z, m_h, m_H. Being invariant under the choice of basis (of parameters), the trace has some physical meaning and then so does the EW fine-tuning that emerges from it. So $\tilde{M}$ with $v = v(\gamma)$ seems more fundamental than the fine-tuning that was introduced in the past on physical grounds. These observations are easily extended if the initial M_{ij} is not diagonal.

The conclusion is that the "usual" EW fine-tuning is automatically present in the analysis of precision data fits provided that one includes the EW scale as $v = v(\gamma)$ predicted by SUSY. This result has not been investigated numerically[14]; one can reevaluate the precision data fits to treat v as a function $v = v(\gamma)$ and include observables that depend on it (m_Z, m_h, etc.) together with the additional likelihood "cost" it could bring. In this picture, the traditional EW fine-tuning *per se* and its numerical value may be less relevant since $\tilde{M}$ contains the information related to these.

Further, one can also consider the determinant of the (inverse) covariance matrix $\tilde{M}^{-1}$ in the basis of the fundamental parameters. It is actually more relevant to consider this determinant relative to that of the initial matrix M, and this gives exactly Δ of Eq. (15). This factor related the normalized $L(\mathcal{O}|\gamma)$ and $L(\gamma)$ in Eq. (14) and it is a measure of their relative width.[15] For simplicity, take the case when the number of observables $\mathcal{O}_i$ equals that of the parameters γ_α ($n = s$). Then,

$$\Delta = (\det \mathcal{J}^T\mathcal{J})^{1/2}. \tag{20}$$

In particular, for two observables (say m_h and m_Z) and two parameters,

$$\Delta = \Delta_1\Delta_2[1-\xi_{12}]^{1/2}, \tag{21}$$

where

$$\xi_{12} \equiv \frac{1}{\Delta_1^2\Delta_2^2}[\mathcal{J}_{1\alpha}\mathcal{J}_{2\alpha}]^2\Delta_k$$
$$= \left\{\sum_\alpha\left(\frac{\mathcal{O}_k(\gamma^0)}{\mathcal{O}_k^0}\frac{d\ln\mathcal{O}_k}{d\ln\gamma_\alpha}\right)^2_{\gamma=\gamma^0}\right\}^{1/2}, \quad k=1,2. \tag{22}$$

With $\mathcal{O}_i$ functions of $v(\gamma)$, the individual fine-tunings $\Delta_{1,2}$ of $\mathcal{O}_{1,2}$ include the EW fine-tuning and are part of this more general Δ. So we see again that fine-tuning of the observables and of the EW scale is included by the covariance matrix.[16] The above results answered question Q_2 in the Introduction.

[14] It may be possible that numerical studies account for this effect in a different way. With v (m_Z^0) a fixed input, the EW minimum condition brings instead a dependence say $\mu = \mu(\gamma_\alpha)$ where γ_α denote parameters other than μ. With this dependence it is possible to account for the above effect, but the presence in the covariance matrix of the EW fine-tuning seen above is not manifest and is overlooked.

[15] In information theory [17] $\ln\Delta$ is interpreted as the change of the differential entropy when going from a multivariate Gaussian distribution (of observables $\mathcal{O}_i$) to another one (here of parameters γ_α).

[16] Δ of Eq. (21) is smaller then when the variations of $\mathcal{O}_i$ are orthogonal ($\xi_{12} = 0$ i.e. independent $\mathcal{O}_i$).

From our result above it is clear that the usual criterion of a good fit in a model $\chi^2/n_{df} \approx 1$ ($\chi^2 \equiv -2\ln L$) imposed on this matrix in a numerical analysis of the EW data with $v = v(\gamma)$ should then automatically take fine- tuning into account; this can then bring bounds on EW fine-tuning. In a simplified setup and under additional assumptions, this procedure was used in [5,6] to set bounds on Δ. We do not pursue this method here.

In the following let us be more general and analyze instead the s-standard deviation confidence interval, defined by the surface,[17]

$$-2\ln L(\gamma') \le -2\ln L(\gamma^0) + s^2, \tag{23}$$

with $L(\gamma^0) = L_{\max}$ and $\gamma' = \gamma'(s)$. In χ^2 language this becomes $\delta\chi^2 = \chi^2 - \chi^2_{\min} \le s^2$. In our approximation this condition becomes, from Eqs. (16), (17),

$$\delta\chi^2 = \tilde{\gamma}_\alpha \tilde{M}^{-1}_{\alpha\beta}\tilde{\gamma}_\beta = \sum_{i=1}^{n}\left\{\left(\frac{d\mathcal{O}_i/\sigma_i}{d\ln\gamma_\alpha}\right)_{\gamma=\gamma^0}(\gamma'_\alpha/\gamma^0_\alpha - 1)\right\}^2 \le s^2 \tag{24}$$

(implicit sum over α). With $\mathcal{O}_i$ a function of $v(\gamma)$, for fixed s and assuming that $|\gamma'_\alpha - \gamma^0_\alpha| \ge \sigma_{\mathrm{th},\alpha} > 0$ this condition brings a bound on the EW fine-tuning. We introduced *ad hoc* $\sigma_{\mathrm{th},\alpha}$ as a theoretical error of computing γ^0_α, from the maximal likelihood (min χ^2) condition in which theoretical values of observables are affected by ignored higher loops errors[18] (including effects of the RG flow for γ_α these observables depend on).

From inequality (24) for one observable only (m_Z),

$$\Delta \le \frac{s\sigma_z}{m_Z(\gamma^0)}\left|\frac{n^\alpha(\gamma'_\alpha - \gamma^0_\alpha)}{\gamma^0_\alpha}\right|^{-1}, \tag{25}$$

with Δ as in Eq. (3). Equation (25) also gives, by varying each γ_α separately with the remaining ones fixed, then adding these results in quadrature:

$$\Delta \le \frac{s\sigma_z}{m_Z(\gamma^0)}\left\{\sum_{\alpha\ge 1}^{s}\left(\frac{\gamma^0_\alpha}{\sigma_{\mathrm{th},\alpha}}\right)^2\right\}^{1/2}. \tag{26}$$

This bound is similar to that discussed in Sec. I. B and Eq. (11) and depends on the experimental and theoretical errors. The strength of this bound depends on the values of $\sigma_{\mathrm{th},\alpha}$ and σ_z (more generally σ_i of all $\mathcal{O}_i$) and can be enhanced by the presence of more observables, see Eq. (24). Finally, using this approach in precision data fits of the EW data, a plot of the lhs of Eq. (24) giving $\delta\chi^2$, as a function of Δ, for current value of the Higgs mass could illustrate the role that fine-tuning plays in deciding if a model is realistic. Bounds (25), (26) can be generalized when original M_{ij} is not diagonal (i.e. when Eq. (17) is not valid anymore).

The matrix $\tilde{M}$ has another interesting feature. A large traditional EW fine-tuning, which is more a problem of supersymmetry breaking than of supersymmetry itself, can signal that our theory of soft terms (γ_α) is incomplete. As mentioned, relations among soft masses (such as GUT relations among the gaugino masses, etc) can reduce its value. There is then the possibility that some parameters γ_α could be related. Such relations can be captured by the matrix $\tilde{M}$ as off-diagonal entries. This means that properties of this matrix can help us identify the fundamental γ_α under the constraints of the model. Indeed, there exists the so-called "global" correlation coefficient of one such parameter (γ_α) with the rest, defined as

$$\rho_\alpha = \sqrt{1 - [\tilde{M}_{\alpha\alpha}(\tilde{M}^{-1})_{\alpha\alpha}]^{-1}}, \quad 0 \le \rho_\alpha \le 1. \tag{27}$$

ρ_α measures the total amount of correlation between γ_α and all other parameters γ_β ($\beta \ne \alpha$). If $\rho_\alpha = 0$ then γ_α is an independent variable while if $\rho_\alpha \to 1$ there is full correlation of γ_α with one linear combination of the other parameters; this is captured by the off-diagonal terms when inverting the matrix. ρ_α could help a better understating of the SUSY breaking soft terms. In this sense ρ_α, $\alpha = 1, 2..s$ could also be used as a new measure of EW fine-tuning defined as $\tilde{\Delta} = \max|\rho_\alpha|$.

Interestingly, another coefficient was discussed previously, for the correlation between pairs of parameters, $\rho_{\alpha\beta}$ [4]; this was used to define a new measure of EW fine-tuning as $\max|\rho_{\alpha\beta}|$, where $\rho_{\alpha\beta} \sim \tilde{M}_{\alpha\beta}/(\sigma_\alpha\sigma_\beta)$. This measure was reached on physical grounds and again supports the connection of the matrix $\tilde{M}$ to fine-tuning, emphasized here. We insist that when one computes the coefficient ρ_α as well as $\rho_{\alpha\beta}$, the EW scale v can be regarded as a function of γ's, to reflect this original prediction of SUSY.

To conclude, a traditional frequentist analysis of the EW observables (including m_Z) with the constraint that v is a function $v = v(\gamma)$ is a test that remains true to the original motivation of SUSY. The EW fine-tuning due to all observables is automatically captured by the covariance matrix if $v = v(\gamma)$, and in this case there may be no need to discuss Δ separately. Upper bounds on the EW fine-tuning emerge, as shown in Eq. (25), from the standard deviation interval constraint discussed (imposed on this matrix), see Eq. (23) and the ignored higher-loops error ($\sigma_{\mathrm{th},\alpha}$) affecting the theoretical calculation of the UV parameters γ_α of the Lagrangian.[19]

This discussion relied on a Taylor expansion of $\mathcal{O}_i$ to linear order which, although illustrative for our purpose, is

[17]The value of s depends on the number of degrees of freedom n_{df}.

[18]An example is that of the 2–3 GeV higher loop error mentioned in Sec. I. B for m_h.

[19]Upper bounds on Δ also emerge from the criterion $\chi^2/n_{df} \approx 1$ (good fit) [5,6], not discussed here.

not acceptable for precision studies. A numerical approach, with $v = v(\gamma)$, can avoid this approximation.

II. CONCLUSIONS

Unlike the SM, its supersymmetric versions stabilize the EW scale $v \sim m_Z$ at the quantum level and predict that v is a derived quantity, function of the SUSY UV parameters γ_α (soft masses, couplings and μ), so $v = v(\gamma)$. Whether this SUSY prediction successfully recovers its experimental value is the natural test of this theory. This view remains true to the original motivation of SUSY. Past estimates showed that fixing the EW scale to its measured value affects the likelihood to fit the data by a factor related to the EW fine-tuning. Here we examined this problem in a different, more general approach.

The result is that the covariance matrix in the basis of the parameters γ_α automatically encodes information about the EW fine-tuning provided that the EW scale is regarded as a function $v = v(\gamma)$ (rather than a constant). Note that such connection between this matrix and the EW fine-tuning was not previously examined in the literature, even though each of these aspects were studied separately. Further, the trace of the inverse of the covariance matrix and its determinant also contain the EW fine-tuning due to all EW observables that depend on v. This indicates that the EW fine-tuning is somewhat less fundamental since this matrix includes its effects through the variations of the EW scale v with respect to γ_α, (closely related to Δ).

A consequence of the above result is that the evaluation of the traditional EW fine-tuning *per se* is then less relevant for the viability of a model as long as with $v = v(\gamma)$ a good $\delta\chi^2$ of the observables (including m_Z) is still possible in that model, within the theoretical approximation (loop order) considered. From this condition and approximation one can subsequently infer numerical bounds on the EW fine-tuning. More explicitly, the deviation $\delta\chi^2$ from the minimal value $\chi^2_{\rm min}$ is affected by the EW fine-tuning; so for a s-standard deviation confidence interval (region) and a given theoretical error (loop order), a bound on the traditional measure of the EW fine-tuning ("in quadrature") was obtained [Eqs. (25), (26)]. This eliminates subjective criteria for "acceptable" numerical values for Δ.

At present the above effect seems to be overlooked in the precision data fits in the frequentist approach where v is actually a fixed, input constant, so the fine-tuning-related corrections to the covariance matrix shown in the text [Eq. (18)] seem to be ignored; or they are indeed included but in a way which does not make manifest the role of the EW fine-tuning that we showed. This effect needs further numerical investigation. Our result also answers how to compare a model with a good fit of the data but significant EW fine-tuning against a model with nearly as good a fit but less EW fine- tuning: since fine-tuning effects are included in the covariance matrix, one simply chooses the model with the best fit obtained with $v = v(\gamma)$. This answers our remaining question (Q_3) in the Introduction.

A large EW fine-tuning can be an indication of our ignorance of the details of the SUSY breaking mechanism and of the lack of a theory of soft terms. It is known that symmetries that relate the soft terms can reduce its value. The global correlation coefficient of the covariance matrix can show if a particular parameter γ_α is correlated with a combination of the rest. This could help one trace the more fundamental SUSY parameters and better understand the relation of fine-tuning to supersymmetry breaking.

ACKNOWLEDGMENTS

The author thanks Stefan Pokorski for many interesting discussions. He also thanks Philip Bechtle, Werner Porod, and Tim Stefaniak for a related discussion. This work was supported by a grant from the Romanian National Authority for Scientific Research, CNCS-UEFIS-CDI, Project No. PN-II-ID-PCE-2011-3-0607, and in part by Programme "Nucleu" PN 09 37 01 02.

[1] L. Susskind, Phys. Rev. D **20**, 2619 (1979); K. Wilson, as acknowledged in this paper.

[2] J. R. Ellis, K. Enqvist, D. V. Nanopoulos, and F. Zwirner, Mod. Phys. Lett. A **01**, 57 (1986); R. Barbieri and G. F. Giudice, Nucl. Phys. **B306**, 63 (1988).

[3] G. W. Anderson and D. J. Castano, Phys. Lett. B **347**, 300 (1995).

[4] P. Bechtle, T. Bringmann, K. Desch, H. Dreiner, M. Hamer, C. Hensel, M. Kramer, and N. Nguyen, J. High Energy Phys. 06 (2012) 098; K. Kowalska, S. Munir, L. Roszkowski, E. M. Sessolo, S. Trojanowski, and Y.-L. S. Tsai, Phys. Rev. D **87**, 115010 (2013); A. Fowlie, M. Kazana, K. Kowalska, S. Munir, L. Roszkowski, E. M. Sessolo, S. Trojanowski, and Y.-L. S. Tsai, Phys. Rev. D **86**, 075010 (2012); C. Strege, G. Bertone, F. Feroz, M. Fornasa, R. R. de Austri, and R. Trotta, J. Cosmol. Astropart. Phys. 04 (2013) 013; G. Bertone, D. G. Cerdeno, M. Fornasa, R. Ruiz de Austri, C. Strege, and R. Trotta, J. Cosmol. Astropart. Phys. 01 (2012) 015; C. Strege, G. Bertone, D. G. Cerdeno, M. Fornasa, R. Ruiz de Austri, and R. Trotta, J. Cosmol. Astropart. Phys. 03 (2012) 030; O. Buchmueller, R. Cavanaugh, M. Citron, A. De Roeck, M. J. Dolan, J. R. Ellis, H. Flacher and S. Heinemeyer *et al.*, Eur. Phys. J. C **72**, 2243 (2012).

[5] D. Ghilencea, Nucl. Phys. **B876**, 16 (2013); D. Ghilencea and G. G. Ross, Nucl. Phys. **B868**, 65 (2013).

[6] D. Ghilencea, *Proc. Sci.*, Corfu2013 (**2013**) 034.
[7] B. C. Allanach, K. Cranmer, C. G. Lester and A. M. Weber, J. High Energy Phys. 08 (2007) 023; M. E. Cabrera, J. A. Casas, and R. Ruiz de Austri, J. High Energy Phys. 03 (2009) 075; M. E. Cabrera, J. A. Casas, and R. Ruiz d Austri, J. High Energy Phys. 05 (2010) 043; S. S. AbdusSalam, B. C. Allanach, F. Quevedo, F. Feroz, and M. Hobson, Phys. Rev. D **81**, 095012 (2010).
[8] S. Fichet, Phys. Rev. D **86**, 125029 (2012).
[9] D. Ghilencea, H. M. Lee, and M. Park, J. High Energy Phys. 07 (2012) 046.
[10] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012); S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012); S. Chatrchyan *et al.*, Report No. ATLAS-CONF-2012-162, CMS-PAS-HIG-12-045.
[11] G. G. Ross, K. Schmidt-Hoberg, and F. Staub, J. High Energy Phys. 08 (2012) 074; G. G. Ross and K. Schmidt-Hoberg, Nucl. Phys. **B862**, 710 (2012), and a similar result is found in the effective approach to GNMSSM: S. Cassel, D. M. Ghilencea, and G. G. Ross, Nucl. Phys. **B825**, 203 (2010).
[12] M. Carena, K. Kong, E. Ponton, and J. Zurita, Phys. Rev. D **81**, 015001 (2010); I. Antoniadis, E. Dudas, D. M. Ghilencea, and P. Tziveloglou, Nucl. Phys. **B831**, 133 (2010); **B848**1 (2011); **B841**157 (2010); M. Carena, E. Ponton, and J. Zurita, Phys. Rev. D **82**, 055025 (2010); **85**035007 (2012); A. Brignole, J. A. Casas, J. R. Espinosa, and I. Navarro, Nucl. Phys. **B666**, 105 (2003); S. Cassel and D. M. Ghilencea, Mod. Phys. Lett. A **27**, 1230003 (2012).
[13] D. Horton and G. G. Ross, Nucl. Phys. **B830**, 221 (2010).
[14] B. C. Allanach, Comput. Phys. Commun. **143**, 305 (2002); G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein, Eur. Phys. J. C **28**, 133 (2003); S. Heinemeyer, Int. J. Mod. Phys. A **21**, 2659 (2006).
[15] P. Ciafaloni and A. Strumia, Nucl. Phys. **B494**, 41 (1997); A. Strumia, arXiv:hep-ph/9904247.
[16] S. Cassel, D. Ghilencea, and G. G. Ross, Phys. Lett. B **687**, 214 (2010); S. Cassel, D. Ghilencea, and G. G. Ross, Nucl. Phys. **B835**, 110 (2010).
[17] T. Cover and J. Thomas, *Elements of Information Theory* (Wiley, New York, 2006).