PHYSICAL REVIEW D **90,** 015010 (2014)

CMSSM, naturalness and the "fine-tuning price" of the Very Large Hadron Collider

Andrew Fowlie*

National Institute of Chemical Physics and Biophysics, Ravala 10, Tallinn 10143, Estonia

(Received 25 March 2014; published 10 July 2014)

The absence of supersymmetry or other new physics at the Large Hadron Collider (LHC) has lead many to question naturalness arguments. With Bayesian statistics, we argue that natural models are most probable and that naturalness is not merely an aesthetic principle. We calculate a probabilistic measure of naturalness, the Bayesian evidence, for the Standard Model (SM) with and without quadratic divergences, confirming that the SM with quadratic divergences is improbable. We calculate the Bayesian evidence for the constrained minimal supersymmetric Standard Model (CMSSM) with naturalness priors in three cases: with only the M_Z measurement; with the M_Z measurement and LHC measurements; and with the M_Z measurement, m_h measurement and a hypothetical null result from a $\sqrt{s} = 100$ TeV Very Large Hadron Collider (VLHC) with 3000/fb. The "fine-tuning price" of the VLHC given LHC results would be ~400, which is slightly less than that of the LHC results given the electroweak scale (~500).

DOI: 10.1103/PhysRevD.90.015010 PACS numbers: 12.60.Jv

I. INTRODUCTION

Weak-scale supersymmetry (SUSY) [1–4] was supposed to solve the naturalness problem of the Standard Model (SM) [5,6], but it was absent in the ATLAS [7] and CMS [8] searches at the Large Hadron Collider (LHC) in 20/fb with center-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. Although ATLAS and CMS will continue their searches for SUSY at $\sqrt{s} = 13$ TeV, a new $\sqrt{s} = 100$ TeV Very Large Hadron Collider (VLHC) might be built [9].

There are numerous motivations for SUSY. The theoretical motivations for SUSY (see, e.g., Ref. [10]) are, *inter alia*, that it completes the maximal symmetries of the S matrix and connects with gravity and superstrings. The phenomenological and experimental motivations for SUSY (see, e.g., Ref. [11]) are that it unifies the gauge couplings at the anticipated scale, that the lightest SUSY particle could explain the measured abundance of dark matter in the Universe and that it predicts that the mass of the lightest Higgs boson is $m_h \lesssim 135$ GeV. Perhaps the strongest motivation for SUSY, however, is that it solves the technical naturalness problem of the SM, if SUSY particles are sufficiently light. The LHC results, however, suggest that SUSY particles might not be sufficiently light [12,13] and have led many to question naturalness arguments [14–18].

We argue in Sec. II that the best measure of naturalness is Bayesian evidence. We measure naturalness in the SM in Sec. III and in the constrained minimal supersymmetric SM (CMSSM) [19–21] in Sec. IV by calculating their Bayesian evidences with "honest" or "naturalness" priors. We evaluate the consequences for naturalness of hypothetical null results from a $\sqrt{s} = 100$ TeV VLHC with Bayesian statistics, i.e., the "fine-tuning price" of the VLHC [22–24], by calculating the Bayesian evidence in this scenario. Learning this price could motivate building the VLHC [25]. We argue that our comparison between the SM and the CMSSM was fair in Sec. V. We discuss the μ problem of the MSSM [26] in the context of Bayesian statistics in Sec. VI, and conclude in Sec. VII. For similar analyses, see, e.g., Refs. [27–33].

*Andrew.Fowlie@KBFI.ee

II. BAYESIAN EVIDENCE

For a pedagogical introduction to Bayesian statistics, see, e.g., Ref. [34]. In Bayesian statistics, probability is a numerical measure of belief in a proposition. With Bayes' theorem, our belief in a model given experimental data is given by

$$p(\text{model}|\text{data}) = \frac{p(\text{data}|\text{model}) \times p(\text{model})}{p(\text{data})}, \tag{1}$$

where $\mathcal{Z} \equiv p(\text{data}|\text{model})$ is the *Bayesian evidence*, $p(\text{model})$ is our prior belief in the model, and $p(\text{data})$ is a normalization constant. We can eliminate the normalization constant if we consider a ratio of probabilities for model_a and model_b:

$$\underbrace{\frac{p(\text{model}_a|\text{data})}{p(\text{model}_b|\text{data})}}_{\text{Posterior odds},\theta'} = \underbrace{\frac{p(\text{data}|\text{model}_a)}{p(\text{data}|\text{model}_b)}}_{\text{Bayes factor},B} \times \underbrace{\frac{p(\text{model}_a)}{p(\text{model}_b)}}_{\text{Prior odds},\theta}. \tag{2}$$

Our *prior odds*, θ, is a numerical measure of our relative belief in model_a over model_b, *before* considering experimental data. The Bayes factor, B, updates our *prior odds*, θ, with the experimental data, resulting in our *posterior odds*, θ'. Our *posterior odds* is a numerical measure of our relative belief in model$_a$ over model$_b$, *after* considering experimental data. The Bayes factor is the ratio of the models' evidences.

Let us make our discussion more concrete. From an experiment, one can construct a "likelihood function" giving the frequentist probability of obtaining the data, given a particular point, $\vec{x}$, in a model's parameter space,

$$\mathcal{L}(\vec{x}) = p(\text{data}|\vec{x}, \text{model}). \tag{3}$$

The likelihood function for a measurement is typically a Gaussian function (by the central limit theorem). It could be, e.g., the probability of measuring a Higgs mass $m_h = 125$ GeV given a particular parameter point $\vec{x}$ in a SUSY model. With Bayes' theorem, it can be readily shown that the evidence is an integral over the likelihood,

$$\mathcal{Z} = \int \mathcal{L}(\vec{x})\pi(\vec{x}) \prod \mathrm{d}x, \tag{4}$$

where $\pi(\vec{x}) \equiv p(\vec{x}|\text{model})$ is our *prior*; our prior belief in the model's parameter space. Priors are somewhat subjective and there might exist a spectrum of assigned priors amongst investigators. All investigators, however, will make identical conclusions from the evidence, if the likelihood is sufficiently informative.

Because individual evidences are somewhat meaningless (e.g., the evidence has dimension $[1/\text{data}]$), it is necessary to compare the evidence against that of a reference model with a Bayes factor. If the Bayes factor is greater than (less than) 1, the model in the numerator (denominator) is favored. The interpretation of Bayes factors is somewhat subjective, though we have chosen the Jeffreys' scale, Table I, to ascribe qualitative meanings to Bayes factors. If a Bayes factor is sufficiently large, all investigators will conclude that a particular model is favorable, regardless of their prior odds for the models. The Jeffreys' scale is, however, only a guide for interpreting a Bayes factor; the full result is the posterior odds found by multiplying the Bayes factor by the prior odds in Eq. (2).

The Bayes factor quantitatively incorporates a principle of economy widely known as Occam's razor and in physics as "fine-tuning" or "naturalness" [36–38]. It is insightful to consider the evidence $\mathcal{Z} = p(\text{data}|\text{model})$ a function of the data normalized to unity, i.e., as a sampling distribution [39]. Natural models "spend" their probability mass near the obtained data, i.e., a large fraction of their parameter space agrees with the data. Complicated models squander their probability mass away from the obtained data. This is illustrated in Fig. 1. Bayesian statistics formalizes Occam's razor, fine-tuning and naturalness arguments. Naturalness is no longer a nebulous, aesthetic criterion; it is formalized and justified by Bayesian statistics.

We measure the "fine-tuning price" of new experimental data with a partial Bayes factor. A partial Bayes factor, P, updates our relative belief in model$_a$ over model$_b$ with new experimental data,

$$P \cdot \frac{p(\text{model}_a|\text{data})}{p(\text{model}_b|\text{data})} = \frac{p(\text{model}_a|\text{data} + \text{newdata})}{p(\text{model}_b|\text{data} + \text{newdata})}. \tag{5}$$

It can be readily shown that a partial Bayes factor is a ratio of Bayes factors,

$$P = \frac{p(\text{data} + \text{newdata}|\text{model}_a)}{p(\text{data} + \text{newdata}|\text{model}_b)} \frac{p(\text{data}|\text{model}_b)}{p(\text{data}|\text{model}_a)}. \tag{6}$$

See, e.g., Ref. [33] for a comprehensive discussion of partial Bayes factors. Having introduced our formalism, we are ready to calculate evidences in the SM and CMSSM.

TABLE I. The Jeffreys' scale for interpreting Bayes factors [35], which are ratios of evidences. We assume that the favored model is in the numerator, though this could be readily inverted.

Grade	Bayes factor, B	Preference for model in numerator
0	$B \leq 1$	Negative
1	$1 < B \leq 3$	Barely worth mentioning
2	$3 < B \leq 20$	Positive
3	$20 < B \leq 150$	Strong
4	$B > 150$	Very strong

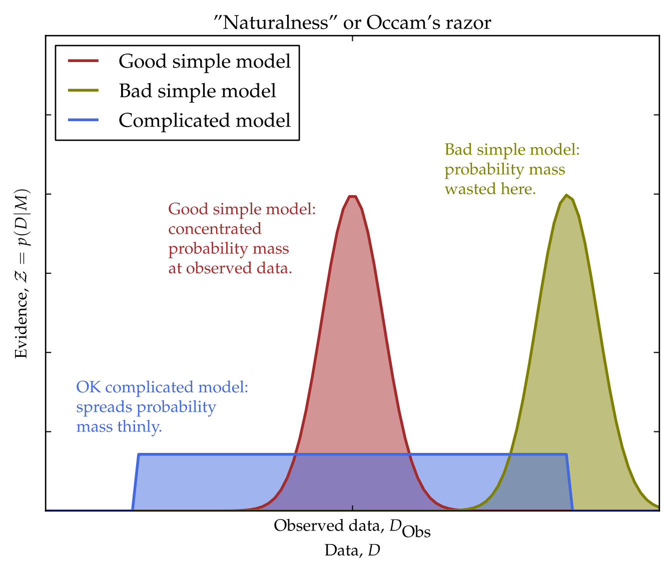

FIG. 1 (color online). Illustration of the evidence, interpreted as a sampling distribution, originally from Ref. [39]. The observed evidence is the evidence evaluated at the observed data. The red line shows a model that concentrates its probability mass at the observed data: it is a good, simple model. The green line shows a model that concentrates its probability mass away from the observed data: it is a bad, simple model. The blue line shows a model that thinly spreads its probability mass around the observed data: it is an OK, complicated model.

III. BAYESIAN EVIDENCE FOR THE STANDARD MODEL

If the SM is coupled to the Planck scale, it suffers from a well-known fine-tuning problem, the "hierarchy problem" [5,6]. The dimension-2 coupling, μ^2, in the Higgs potential,

$$V = \mu^2\phi^2 + \lambda\phi^4, \tag{7}$$

must be incredibly fine-tuned. The dressed coupling must be $\sim -(100\ \mathrm{GeV})^2$, but the bare coupling receives a positive quadratic correction $\sim M_\mathrm{P}^2$. Let us calculate the evidence for the SM, given the electroweak scale and that the Higgs mass is $\sim$125 GeV. While naively our Higgs potential is described by μ^2 and λ, let us instead write the dressed dimension-2 coupling as the sum of a bare coupling and a quadratic correction,

$$\mu^2 = \mu_0^2 + \Delta\mu^2, \tag{8}$$

and treat μ_0^2, $\Delta\mu^2$ and λ as separate parameters. *A priori*, if the SM is coupled to the Planck scale, M_p, we expect that $\Delta\mu^2 \sim M_\mathrm{P}^2$, and that $\lambda \sim 1$, whereas we have no idea about the scale of μ_0^2. Let us formalize these thoughts with logarithmic, scale invariant priors $\pi(x) \propto 1/x$:

$$\Delta\mu^2 \text{ between } 10^{36} \quad \text{and} \quad 10^{40}\ \mathrm{GeV}^2, \tag{9}$$

$$\mu_0^2 \text{ between } 10^{0} \quad \text{and} \quad 10^{40}\ \mathrm{GeV}^2, \tag{10}$$

$$\lambda \text{ between } 10^{-3} \quad \text{and} \quad 10^{1}. \tag{11}$$

We also note that *a priori* μ_0^2 could be positive or negative.

We calculate the evidence for the SM given the M_Z measurement [40] and the LHC $m_h \sim 125$ GeV measurement [40–42]. We approximate the likelihood functions for the measurements of M_Z and m_h as Dirac delta functions:

$$\mathcal{Z}_{\text{only } M_Z} = \frac{\int \delta(M_Z - 91.1876\ \mathrm{GeV}) \frac{\mathrm{d}\mu_0^2}{\mu_0^2} \frac{\mathrm{d}\lambda}{\lambda} \frac{\mathrm{d}\Delta\mu^2}{\Delta\mu^2}}{\int \frac{\mathrm{d}\mu_0^2}{\mu_0^2} \frac{\mathrm{d}\lambda}{\lambda} \frac{\mathrm{d}\Delta\mu^2}{\Delta\mu^2}},$$
$$\mathcal{Z}_{m_h \text{ and } M_Z} = \frac{\int \delta(M_Z - 91.1876\ \mathrm{GeV}) \delta(m_h - 125.9\ \mathrm{GeV}) \frac{\mathrm{d}\mu_0^2}{\mu_0^2} \frac{\mathrm{d}\lambda}{\lambda} \frac{\mathrm{d}\Delta\mu^2}{\Delta\mu^2}}{\int \frac{\mathrm{d}\mu_0^2}{\mu_0^2} \frac{\mathrm{d}\lambda}{\lambda} \frac{\mathrm{d}\Delta\mu^2}{\Delta\mu^2}}. \tag{12}$$

The denominators normalize our logarithmic priors. We integrate the Dirac delta functions with tree-level formulas for the Higgs and Z-boson masses (see, e.g., Ref. [43]),

$$m_h = \sqrt{-2\mu^2}, \tag{13}$$

$$M_Z = g\sqrt{\frac{-\mu^2}{2\lambda}}. \tag{14}$$

We calculate evidences by performing the integrals in Eq. (12) for two models:

(1) The SM with quadratic divergences, $\Delta\mu^2 \sim M_\mathrm{P}^2$, and
(2) The SM without quadratic divergences, $\Delta\mu^2 = 0$. References [44,45] argue that quadratic divergences vanish in theories with classical scale invariance without modifications to the Z-boson or Higgs boson masses.

The resulting evidences are in Table II. Unsurprisingly, the evidence for the SM with quadratic divergences is

TABLE II. Bayesian evidences and Bayes factors for the SM with quadratic divergences, SM without quadratic divergences and CMSSM. The headings indicate which experimental results were included. The final column is the fine-tuning price, as measured by partial Bayes factors.

	M_Z	M_Z, m_h and LHC	M_Z, m_h and VLHC
Evidences, $\mathcal{Z}$	GeV^{-1}	GeV^{-2}	GeV^{-2}
SM with quadratic divergences	9×10^{-37}	2×10^{-40}	2×10^{-40}
SM no quadratic divergences	1×10^{-4}	2×10^{-7}	2×10^{-7}
CMSSM	8×10^{-5}	3×10^{-10}	7×10^{-13}
Bayes factors, $B = \mathcal{Z}_a/\mathcal{Z}_b$			
CMSSM/SM with quadratic divergences	9×10^{31}	2×10^{30}	4×10^{27}
SM no quadratic divergences/CMSSM	2×10^{0}	7×10^{2}	3×10^{5}
Partial Bayes factors, $P = B_{i+1}/B_i$			
SM no quadratic divergences/CMSSM	$\sim$2	$\sim$500	$\sim$400

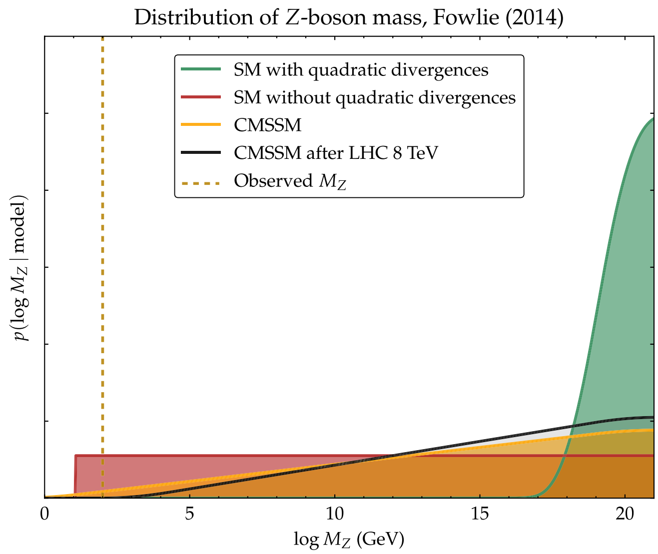

FIG. 2 (color online). The probability distribution of the Z-boson mass in the various models. The area under each plot is equal to 1.

minuscule compared to that for the SM without quadratic divergences. The Bayes factors in Table II are more than 10^{30} against the SM with quadratic divergences (150 is considered "very strong" on the Jeffreys' scale).

Let us interpret the evidence as a sampling distribution for the expected Z-boson mass, i.e., plot the evidence as a function of M_Z (Fig. 2). As expected, the SM with quadratic divergences squanders its prediction for the Z-boson mass near M_{p}, far away from the measured M_Z. The SM with quadratic divergences is unnatural. Because without quadratic divergences one can make no prediction for the magnitude of M_Z, the SM without quadratic divergences is somewhat unnatural and complicated.

Now that we have completed the somewhat trivial exercise of calculating the evidences for the SM, let us calculate the evidences for the CMSSM.

IV. BAYESIAN EVIDENCES FOR THE CMSSM

The Z-boson mass, or, equivalently, the scale of electroweak symmetry breaking, is predicted in the MSSM via radiative electroweak symmetry breaking. At tree level [4],

$$\frac{1}{2}M_Z^2 = -\mu^2 + \frac{m_{H_d}^2 - m_{H_u}^2 \tan^2\beta}{\tan^2\beta - 1}. \tag{15}$$

This expression is problematic; it contains the "little-hierarchy problem" [46] and the related "μ problem" [26]. From experiments, we know that M_Z is $\sim$100 GeV. The MSSM predicts M_Z via a cancellation between the SUSY breaking parameters, $m_{H_u}^2$ and $m_{H_d}^2$, and a SUSY preserving parameter in the superpotential, μ. If the SUSY breaking scale is greater than the measured value of M_Z, a cancellation between such large numbers is somewhat miraculous. This is the little-hierarchy problem. This problem is statistical in nature; we are concerned that the MSSM is unlikely because its parameters must be fine-tuned; i.e., it might only agree with experiments in a small fraction of its parameter space.

With a simplified Eq. (15),

$$\frac{1}{2}M_Z^2 \simeq -\mu^2 - m_{H_u}^2, \tag{16}$$

we found an analytic expression for the evidence from Eq. (4) as a function of M_Z in the CMSSM with logarithmic priors. We plot this expression as a function of the Z-boson mass in Fig. 2. While the CMSSM is somewhat fine-tuned, the fine-tuning of the SM with quadratic divergences is far worse. The SM dimension-2 coupling is quadratically sensitive to the UV; the highest scales must enter our expression for M_Z. In the SM with quadratic divergences, the cancellation resulting in M_Z must involve quantities $\sim M_{\text{P}}$. In the CMSSM, we require a cancellation, but the cancellation could be at any scale up to M_{p}.

Fine-tuning is typically measured with a sensitivity, for example, that was originally proposed in Refs. [37,38], the Barbieri-Giudice measure,

$$\Delta_i = \frac{x_i}{M_Z^2}\frac{\partial M_Z^2}{\partial x_i}, \tag{17}$$

where x_i are the model's parameters. The reciprocal of this measure is, indeed, similar to our Bayesian evidence, in that a small Barbieri-Giudice measure indicates that the model might spend its probability mass around the measured value of M_Z. This is illuminated by rewriting Eq. (17),

$$\Delta_i^{-1} = \left[\frac{\Delta M_Z^2}{M_Z^2}\frac{x_i}{\Delta x_i}\right]^{-1} \propto \frac{\Delta x_i}{x_i}. \tag{18}$$

The reciprocal of the Barbieri-Giudice measure is proportional to the local fraction of the model's parameter space in which M_Z varies by ΔM_Z; in similarity, the evidence is a measure of the fraction of the model's prior volume in which the model agrees with experiments [47–49]. The Barbieri-Giudice measure lacks, however, a formal interpretation and is, furthermore, a property of a point in the model's parameter space, rather than of the model itself (cf. the evidence).

Bayesian evidence automatically penalizes fine-tuning. Focusing mechanisms (e.g., the focus point [50–52]) are automatically incorporated. We must, however, choose "honest" priors. In the CMSSM, we ought to formulate our prior beliefs in μ and b, the fundamental parameters, defined as

$$W \supset \mu H_u H_d, \tag{19}$$

$$\mathcal{L}_{\text{Soft}} \supset -b^2 H_u H_d + \text{c.c.} \tag{20}$$

For pragmatism, however, we exchange μ and b for $M_Z$and $\tan\beta$ via e.g., Eq. (15). We ought to transform our priors with the appropriate Jacobian, resulting in *effective priors* for M_Z and $\tan\beta$ [28,29,31,47–49]. With logarithmic priors for μ and b, our effective priors are

$$\pi(M_Z) = \frac{\partial\mu}{\partial M_Z}\pi(\mu) = \frac{2\mu}{M_Z}\Delta_\mu^{-1}\pi(\mu) = \mathrm{const}\Delta_\mu^{-1}, \quad (21)$$

$$\pi(\tan\beta) = \frac{\partial b}{\partial\tan\beta}\pi(b) = \frac{\mathrm{const}}{b}\frac{\partial b}{\partial\tan\beta}. \quad (22)$$

The effective prior for M_Z reveals the formal relationship between Bayesian statistics and the Barbieri-Giudice measure [47,48]. With the Barbieri-Giudice measure, the statistical nature of the problem is latent [27,53]; it is now manifest.

We calculated the evidence exactly in the CMSSM with honest priors. Let us make our prior choices clear, because it is a potential source of confusion. For the fundamental CMSSM parameters and priors we pick[1]

$$\begin{aligned}
&m_0 \text{ log prior between } 1\ \text{GeV} \quad \text{and} \quad M_\mathrm{P},\\
&m_{1/2}/m_0 \text{ log prior between } 10^{-3} \quad \text{and} \quad 10^3,\\
&A_0/m_0 \text{ linear prior between } -5 \quad \text{and} \quad 5,\\
&b/m_0 \text{ log prior between } 10^{-3} \quad \text{and} \quad 10^3,\\
&\mu \text{ log prior between } 1\ \text{GeV} \quad \text{and} \quad M_\mathrm{P}. \qquad (23)
\end{aligned}$$

We anticipate that a breaking mechanism might distribute the SUSY breaking masses about a common SUSY breaking scale [48], which we pick as m_0. We do not consider mechanisms in which SUSY breaking parameters are split into distinct groups separated by many orders of magnitude [54,55]. We call this choice of priors and parametrization our *de jure* priors.

Were we to numerically calculate the evidence for the CMSSM with our *de jure* priors, we would waste CPU time considering parameter space with incorrect M_Z. For the purpose of our numerical calculation, we transform our *de jure* priors into our equivalent *de facto* priors,

$$\begin{aligned}
&m_0 \text{ log prior between } 1\ \text{GeV} \quad \text{and} \quad 20\ \text{TeV},\\
&m_{1/2}/m_0 \text{ log prior between } 10^{-3} \quad \text{and} \quad 10^3,\\
&A_0/m_0 \text{ linear prior between } -5 \quad \text{and} \quad 5,\\
&\tan\beta \text{ effective prior between } 1 \quad \text{and} \quad 60,\\
&M_Z \text{ effective prior, fixed } 91.1876\ \text{GeV}, \qquad (24)
\end{aligned}$$

[1]One might wonder whether we should pick, e.g., m_0^2 rather than m_0 as a fundamental parameter, since it is the square which appears in the soft-breaking Lagrangian. Because we pick logarithmic priors, however, the choice is irrelevant.

where the effective priors are in Eq. (21). The "missing" parameter space in our *de facto* priors at $M_\mathrm{SUSY} \gg 20$ TeV is irrelevant in our calculation, because it contains negligible evidence. The "missing" parameter space, however, results in differences in normalization between our *de jure* and *de facto* priors, which we correct by hand. Fortunately, because the sign of μ is identical at the electroweak and M_P scales, we require no Jacobian to transform our prior for $\operatorname{sign}\mu$ from M_P to the electroweak scale.

We pick informative, Gaussian priors for the SM nuisance parameters m_t, m_b, $1/\alpha_\mathrm{em}$ and α_s [40]. Reference [47] stresses that the top and bottom masses are derived parameters; the input parameters are the Yukawa couplings, y_t and y_b. In the CMSSM, the relationship between fermion masses and the Yukawa couplings includes factors of $\sin\beta$ and $\cos\beta$. We should pick priors for the Yukawa couplings rather than for the fermion masses; however, at leading order with logarithmic priors for the Yukawa couplings, there is no effective prior associated with $(y_t, y_b) \to (m_t, m_b)$. At leading order, our treatment of the SM nuisance parameters is equivalent to picking logarithmic priors for the Yukawa couplings.

We calculated the CMSSM's mass spectrum and effective priors with SOFTSUSY [56]. We used MultiNest [57] with PyMultiNest [58] to perform the integral in Eq. (4). We found the evidence for three cases:

(1) $M_Z = 91.1876$ GeV [40] only in our likelihood (fixed by our *de facto* priors),
(2) M_Z, $m_h = 125.9 \pm 0.4 \pm 2.0$ GeV [40–42,59] and the null result from the LHC in 20/fb [7] in our likelihood, and
(3) M_Z, m_h and a hypothetical null result from the VLHC in 3000/fb [60] in our likelihood.

In the first case, our likelihood for M_Z is a Dirac delta function. In the second case, our likelihood for m_h is a Gaussian with theoretical and experimental errors added in quadrature, and we veto points that are excluded by an ATLAS search for jets and missing energy [7]. In the last case, we consider the potential consequences of the $\sqrt{s} = 100$ TeV VLHC, by vetoing points that would be excluded by a null result in 3000/fb [60], i.e., points with $m_{\tilde{g}} \lesssim 16$ TeV and $m_{\tilde{q}} \lesssim 16$ TeV or points with $m_{\tilde{g}} \lesssim 13.5$ TeV.

The evidences for the CMSSM in our three cases are shown are shown in Table II. Let us discuss the results case by case:

(1) *M_Z only in our likelihood*. The Bayes factor favors the CMSSM over the SM with quadratic divergences by $\sim 10^{32}$; as anticipated, the CMSSM is favored by naturalness. The Bayes factor favors the SM without quadratic divergences over the CMSSM by only ~ 2, which is "barely worth mentioning" on the Jeffreys' scale in Table I. Prior to LHC experiments, the CMSSM was not unnatural.
(2) M_Z, m_h and LHC 20/fb *in our likelihood*. The Bayes factor favors the CMSSM over the SM with

quadratic divergences by $\sim 10^{30}$; the little-hierarchy problem in the CMSSM is minuscule compared with the hierarchy problem in the SM with quadratic divergences. The Bayes factor favors the SM without quadratic divergences over the CMSSM by ~ 700 (150 is very strong on the Jeffreys' scale). Relative to the SM without quadratic divergences, the evidence for the CMSSM diminishes by a factor of ~ 500; this is the fine-tuning price of the LHC.

(3) *M_Z, m_h and a hypothetical null result from VLHC* 3000/fb *in our likelihood*. The Bayes factor favors the SM without quadratic divergences over the CMSSM by $\sim 10^5$. Relative to the SM without quadratic divergences, the evidence for the CMSSM diminishes by a further factor of ~ 400. The fine-tuning price of null results from the VLHC (~ 400) would be similar to, though slightly less than, that of the LHC (~ 500).

Note that in all cases, however, the Bayes factors favor the CMSSM over the SM with quadratic divergences by $\gtrsim 10^{27}$. The fine-tuning prices for the experiments in the CMSSM are illustrated in Fig. 3 by the logarithm of the Bayes factor for the SM without quadratic divergences against the CMSSM.

The posterior probability density (see, e.g., Ref. [61] for an introduction) is a by-product of the MultiNest evidence calculation. With M_Z, m_h and null results from the LHC in our likelihood, the posterior probability density for $(m_0,\ m_{1/2})$ confirms that the focus point [50–52] at $m_0 \sim 8$ TeV and $m_{1/2} \lesssim 2$ TeV is favored. With only M_Z in our likelihood, unsurprisingly, we find that $M_{\mathrm{SUSY}} \sim M_Z$ is favored by M_Z, i.e., by naturalness.

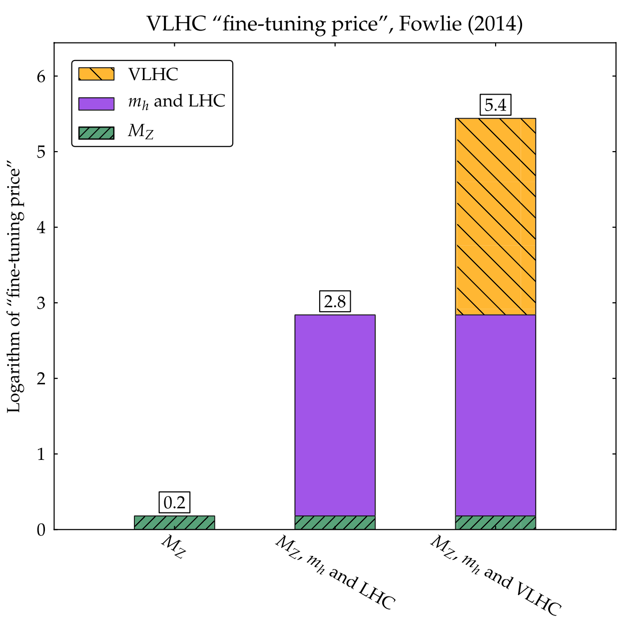

FIG. 3 (color online). The fine-tuning prices of the M_Z measurement, LHC experiments and hypothetical null results from the VLHC. Our fine-tuning prices are the Bayes factors for the SM without quadratic divergences against the CMSSM broken down by experiment. M_Z indicates the measurement of the Z-boson mass, m_h; LHC indicates the LHC Higgs mass measurement and null results from LHC; and VLHC indicates hypothetical null results in 3000/fb at $\sqrt{s} = 100$ TeV. The logarithm is base 10.

V. EFFECTIVE VERSUS UV-COMPLETE THEORIES

We interpreted the SM as an effective theory valid only below a cutoff scale, $\Lambda = M_{\mathrm{P}}$, above which, we presumed, quantum field theory (QFT) is significantly modified by new physics (NP) related to gravity. The SM has a single relevant operator, $\mu^2\phi^2$. We considered the finite bare mass, μ^2, to be a physical parameter originating from NP. We parametrized our ignorance of μ^2 with a logarithmic prior. We assumed that quadratic corrections to μ^2 are unaffected by NP below the Planck scale, hinted at by, e.g., neutrino masses, inflation and dark matter.

If one attempts to remove the cutoff from the SM, $\Lambda \to \infty$, the bare mass diverges. The renormalized mass, μ_{R}^2, might be considered to be fundamental. The renormalized mass differs from the bare mass by a quadratic correction and scheme-dependent terms. Because there are no quadratic corrections to the renormalized mass, it runs logarithmically from the Planck scale to the electroweak (EW) scale. The hierarchy problem is hidden in counterterms. There are numerous problems with such an approach, e.g., triviality.

We, however, interpreted the CMSSM as an ultraviolet (UV) complete theory valid at all scales. The fundamental parameters were renormalized SUSY breaking masses defined at the renormalization scale $\mu = M_{\mathrm{GUT}}$ in the minimal subtraction scheme, rather than bare masses. Was it fair to compare the SM as an effective theory with the CMSSM as a UV-complete theory? While with Bayesian evidence one can compare any models that make predictions for the experimental data, comparisons are interesting only if the models are realistically interpreted.

Suppose we instead interpreted the CMSSM as an effective theory valid only below a cutoff scale, $\Lambda = M_{\mathrm{GUT}}$, at which new GUT physics is important, or the Planck scale, at which gravitational interactions mediate SUSY breaking. Divergences are no worse than logarithmic in supersymmetric models. The bare SUSY breaking masses at the cutoff scale would be similar to renormalized SUSY breaking masses at the GUT or Planck scales; the bare and renormalized masses would differ by logarithmic corrections. No fine-tuning of the EW scale is hidden by parametrizing the CMSSM as a UV-complete theory with renormalized masses.[2]

[2] It is possible, however, that focusing mechanisms are disfavored if SUSY breaking masses are unified at the Planck scale rather than at the GUT scale [62].

The comparison was fair, though its outcome was perhaps inevitable. Although the SM was vastly disfavored, it was important in the analysis; it was a reference model against which we judged the severity of the change in the "fine-tuning price" in the CMSSM.

VI. THE μ PROBLEM

A problem emerges from our honest choice of prior for μ, which aggravates the fine-tuning problem. The μ parameter is a symmetry conserving parameter in the superpotential. *A priori*, it is unrelated to a symmetry breaking scale. This is problematic; phenomenologically it must be that $\mu \sim M_{\mathrm{SUSY}}$. The evidence for a model in which we expect $100\ \mathrm{GeV} \lesssim \mu \lesssim M_{\mathrm{P}}$ and observe $\mu \sim M_{\mathrm{SUSY}}$ could be smaller than that for a model in which we expect $\mu \sim M_{\mathrm{SUSY}}$ and observe $\mu \sim M_{\mathrm{SUSY}}$. This is the μ problem; in our formulation, its statistical nature is manifest. Equation (21) reveals the μ problem and the fine-tuning problem; the μ problem is that $\pi(\mu \approx M_{\mathrm{SUSY}})$ is small and the fine-tuning problem is that $\partial\mu/\partial M_Z$ is small, resulting in a small prior belief in the observed electroweak scale, $\pi(M_Z)$. The ratio of evidences for a model that predicts $M_Z \lesssim \mu \lesssim M_{\mathrm{P}}$ and an "almost-so" model that predicts, e.g., $10^{-1} M_{\mathrm{SUSY}} \lesssim \mu \lesssim 10^3 M_{\mathrm{SUSY}}$, is approximately

$$\frac{\ln\left(\frac{10^3 M_{\mathrm{SUSY}}}{10^{-1} M_{\mathrm{SUSY}}}\right)}{\ln\left(\frac{M_{\mathrm{P}}}{M_Z}\right)} \approx \frac{1}{5}. \tag{25}$$

A similar result applies to SUSY models with a Giudice-Masiero mechanism [63].

The little-hierarchy problem in the CMSSM is ~30 times worse than the μ problem. The μ problem contributes a factor of only ~5 to a Bayes factor for an "almost-so" model against the CMSSM. The Bayes factor with M_Z, m_h and null results from the LHC favors the SM without quadratic divergences over the CMSSM by ~700; the little-hierarchy problem contributes a factor of ~150 and the μ problem contributes a factor of ~5.

VII. CONCLUSIONS

The absence of SUSY or other new physics at the LHC has led many to question naturalness arguments. Drawing upon the literature, we clarified the relationship between Bayesian statistics and naturalness, concluding that natural models are most probable and that naturalness is not merely an aesthetic principle. We calculated the Bayesian, probabilistic measure of naturalness, the evidence, for the SM with and without quadratic divergences, demonstrating that the SM with quadratic divergences is improbable. We calculated the evidence for the CMSSM in three cases: with only the M_Z measurement; with the M_Z measurement and LHC measurements; and with the M_Z measurement and a hypothetical null result from the VLHC with 3000/fb. The latter allowed us to quantitatively understand the potential fine-tuning price of the VLHC. We found that the fine-tuning price of null results from the VLHC (~400) would be slightly less than that of the LHC (~500). We hope this result might help to inform preliminary discussions and plans for the VLHC.

ACKNOWLEDGMENTS

I thank M. Raidal and A. Strumia for helpful criticisms of my manuscript. This work was supported in part by Grant No. IUT23-6, CERN+, and by the European Union through the European Regional Development Fund and by ERDF project 3.2.0304.11-0313, Estonian Scientific Computing Infrastructure (ETAIS).

[1] A. Salam and J. A. Strathdee, Nucl. Phys. **B76**, 477 (1974).
[2] H. E. Haber and G. L. Kane, Phys. Rep. **117**, 75 (1985).
[3] H. P. Nilles, Phys. Rep. **110**, 1 (1984).
[4] S. P. Martin, arXiv:hep-ph/9709356.
[5] E. Gildener, Phys. Rev. D **14**, 1667 (1976).
[6] L. Susskind, Phys. Rev. D **20**, 2619 (1979).
[7] ATLAS Collaboration, Tech. Rep. ATLAS-CONF-2013-047, CERN, Geneva, 2013.
[8] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 06 (2014) 055.
[9] Future Circular Collider Kickoff Meeting (University of Geneva, Geneva, Switzerland, 2014).
[10] M. Dine, *Supersymmetry and String Theory: Beyond the Standard Model* (Cambridge University Press, Cambridge, England, 2007).
[11] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events* (Cambridge University Press, Cambridge, England, 2006).
[12] G. F. Giudice, arXiv:1307.7879.
[13] J. L. Feng, Annu. Rev. Nucl. Part. Sci. **63**, 351 (2013).
[14] R. Foot, A. Kobakhidze, K. L. McDonald, and R. R. Volkas, Phys. Rev. D **89**, 115018 (2014).
[15] S. Dubovsky, V. Gorbenko, and M. Mirbabayi, J. High Energy Phys. 09 (2013) 045.
[16] M. Heikinheimo, A. Racioppi, M. Raidal, C. Spethmann, and K. Tuominen, Mod. Phys. Lett. A **29**, 1450077 (2014).
[17] M. Farina, D. Pappadopulo, and A. Strumia, J. High Energy Phys. 08 (2013) 022.
[18] A. de Gouvea, D. Hernandez, and T. M. P. Tait, Phys. Rev. D **89**, 115005 (2014).

[19] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, Phys. Rev. Lett. **49**, 970 (1982).

[20] R. L. Arnowitt and P. Nath, Phys. Rev. Lett. **69**, 725 (1992).

[21] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, Phys. Rev. D **49**, 6173 (1994).

[22] P. H. Chankowski, J. R. Ellis, and S. Pokorski, Phys. Lett. B **423**, 327 (1998).

[23] R. Barbieri and A. Strumia, Phys. Lett. B **433**, 63 (1998).

[24] A. Strumia, J. High Energy Phys. 04 (2011) 073.

[25] N. Arkani-Hamed, in [9].

[26] J. E. Kim and H. P. Nilles, Phys. Lett. **138B**, 150 (1984).

[27] L. Giusti, A. Romanino, and A. Strumia, Nucl. Phys. **B550**, 3 (1999).

[28] B. C. Allanach, Phys. Lett. B **635**, 123 (2006).

[29] B. C. Allanach, K. Cranmer, C. G. Lester, and A. M. Weber, J. High Energy Phys. 08 (2007) 023.

[30] F. Feroz, B. C. Allanach, M. Hobson, S. S. AbdusSalam, R. Trotta, and A. M. Weber, J. High Energy Phys. 10 (2008) 064.

[31] M. E. Cabrera, J. A. Casas, and R. R. de Austri, J. High Energy Phys. 07 (2013) 182.

[32] D. Kim, P. Athron, C. Balázs, B. Farmer, and E. Hutchison, arXiv:1312.4150.

[33] C. Balazs, A. Buckley, D. Carter, B. Farmer, and M. White, Eur. Phys. J. C **73**, 2563 (2013).

[34] R. Trotta, Contemp. Phys. **49**, 71 (2008).

[35] H. Jeffreys, *Theory of Probability* (Clarendon Press, Oxford, 1961), 3rd ed.

[36] A. Grinbaum, Found. Phys. **42**, 615 (2012).

[37] R. Barbieri and G. F. Giudice, Nucl. Phys. **B306**, 63 (1988).

[38] J. R. Ellis, K. Enqvist, D. V. Nanopoulos, and F. Zwirner, Mod. Phys. Lett. A **01**, 57 (1986).

[39] D. J. C. MacKay, Ph.D. thesis, California Institute of Technology, 1991.

[40] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[41] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).

[42] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).

[43] T. P. Cheng and L. F. Li, *Gauge Theory of Elementary Particle Physics* (Oxford University Press, New York, 1984).

[44] W. A. Bardeen, Proceedings of the Ontake Summer Institute on Particle Physics, Ontake Mountain, Japan, 1995, http://inspirehep.net/record/404517.

[45] K. A. Meissner and H. Nicolai, Phys. Lett. B **648**, 312 (2007).

[46] R. Barbieri and A. Strumia, arXiv:hep-ph/0007265.

[47] M. E. Cabrera, J. A. Casas, and R. Ruiz de Austri, J. High Energy Phys. 03 (2009) 075.

[48] M. E. Cabrera, J. A. Casas, and R. Ruiz de Austri, J. High Energy Phys. 05 (2010) 043.

[49] S. Fichet, Phys. Rev. D **86**, 125029 (2012).

[50] K. L. Chan, U. Chattopadhyay, and P. Nath, Phys. Rev. D **58**, 096004 (1998).

[51] J. L. Feng, K. T. Matchev, and T. Moroi, Phys. Rev. D **61**, 075005 (2000).

[52] J. L. Feng, K. T. Matchev, and D. Sanford, Phys. Rev. D **85**, 075007 (2012).

[53] A. Strumia, arXiv:hep-ph/9904247.

[54] G. F. Giudice and A. Romanino, Nucl. Phys. **B699**, 65 (2004).

[55] N. Arkani-Hamed and S. Dimopoulos, J. High Energy Phys. 06 (2005) 073.

[56] B. C. Allanach, Comput. Phys. Commun. **143**, 305 (2002).

[57] F. Feroz, M. P. Hobson, and M. Bridges, Mon. Not. R. Astron. Soc. **398**, 1601 (2009).

[58] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel *et al.*, arXiv:1402.0004.

[59] B. C. Allanach, A. Djouadi, J. L. Kneur, W. Porod, and P. Slavich, J. High Energy Phys. 09 (2004) 044.

[60] T. Cohen, T. Golling, M. Hance, A. Henrichs, K. Howe, J. Loyal, S. Padhi, and J. G. Wacker, J. High Energy Phys. 04 (2014) 117.

[61] A. Fowlie, A. Kalinowski, M. Kazana, L. Roszkowski, and Y. L. S. Tsai, Phys. Rev. D **85**, 075012 (2012).

[62] N. Polonsky and A. Pomarol, Phys. Rev. Lett. **73**, 2292 (1994).

[63] G. F. Giudice and A. Masiero, Phys. Lett. B **206**, 480 (1988).