Available online at www.sciencedirect.com

ScienceDirect

Nuclear Physics B 884 (2014) 379–395

NUCLEAR PHYSICS B

www.elsevier.com/locate/nuclphysb

Probing the scale of New Physics at the LHC: The example of Higgs data

Sylvain Fichet

International Institute of Physics, UFRN, Av. Odilon Gomes de Lima, 1722, Capim Macio, 59078-400 Natal, RN, Brazil

Received 12 February 2014; received in revised form 10 April 2014; accepted 29 April 2014

Available online 9 May 2014

Editor: Hong-Jian He

Abstract

We present a technique to determine the scale of New Physics (NP) compatible with any set of data, relying on well-defined credibility intervals. Our approach relies on the statistical view of the effective field theory capturing New Physics at low energy. We introduce formally the notion of testable NP and show that it ensures integrability of the posterior distribution. We apply our method to the Standard Model Higgs sector in light of recent LHC data, considering two generic scenarios. In the scenario of democratic higher-dimensional operators generated at one-loop, we find the testable NP scale to lie within $[10, 260]$ TeV at 95% Bayesian credibility level. In the scenario of loop-suppressed field strength-Higgs operators, the testable NP scale is within $[28, 1200]$ TeV at 95% Bayesian credibility level. More specific UV models are necessary to allow lower values of the NP scale.

1. Introduction

Several major experimental and theoretical facts like the measurement of neutrino masses, proofs of the existence of dark matter, as well as the hierarchy problem or the striking hints for Grand Unification all point towards the existence of physics beyond the Standard Model (SM). Although there are strong expectations that such New Physics (NP) should show up at an energy scale close to the electroweak scale, direct searches for new states have so far turned out to be

E-mail address: sylvain.fichet@lpsc.in2p3.fr.

http://dx.doi.org/10.1016/j.nuclphysb.2014.04.025

unsuccessful. Indirect constraints from electroweak precision measurements at LEP also push the NP scale Λ above the electroweak scale.

Overall, it seems that Λ should be substantially higher than the electroweak scale, $\Lambda \gg m_Z$. This paradigm is adopted in a large amount of propositions of new physics. We adopt this fairly general hypothesis in the present work. It implies that the NP involved in physical processes at an energy scale $E \ll \Lambda$ can be integrated out. This results in a low-energy effective theory, consisting of the Standard Model supplemented by infinite series of local, higher-dimensional operators (HDOs) involving negative powers of the NP scale Λ,

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_{i,m} \frac{\alpha_i}{\Lambda^{m_i}} \mathcal{O}_i . \tag{1}$$

Considering a set of experimental observations through this effective description of new physics, we can wonder what information can be obtained about Λ. For a dataset perfectly compatible with the SM, it is common to derive a lower bound on Λ, barring some fine-tuned cancellations among HDO-induced contributions. On the other hand, if data show a deviation with respect to the SM, arbitrary high values of Λ should be also disfavoured, as the effective theory reduces to the SM in the decoupling limit $\Lambda \to \infty$ and cannot explain the discrepancy. Finding a general method to consistently infer the range of Λ compatible with some data – whether they deviate or not from the SM – is the subject of the present work.

We are going to use the effective theory approach within the framework of Bayesian statistics. An important feature of the Bayesian framework is that any irrelevant parameter can be consistently eliminated in a well-defined way through integration. Here we will be mainly interested in the probability distribution of Λ, $p(\Lambda|\text{data})$, which will be obtained through integration over all the α_i coefficients. Adopting a Bayesian view is appropriate to account for the generic character of the scenario we will consider (i.e. it ensures that no fine-tuning is present in the scenario).[1]

The outline of this note is as follows. In Section 2 we shortly review the basics of Bayesian inference and discuss its application to effective theories. In Section 3 we show that one has to require NP to be testable to obtain an integrable posterior. The basic MCMC setup and conceptual subtleties inherent to our approach are discussed in Section 4. Although inference on Λ applies to any kind of data, it is particularly motivated by current LHC results. In Section 5 we apply our method to the Standard Model Higgs sector, using the latest pieces of information available from CMS, ATLAS and Tevatron. We discuss the leading constraints and the necessary conditions favouring lower values of the NP scale.

2. Effective theory and Bayesian inference

Let us briefly review necessary notions of Bayesian statistics (see [2] for an introduction). In this approach, the notion of probability p is defined as the degree of belief about a proposition. Our study lies in the domain of Bayesian inference, which is based on the relation

$$p(\theta|d, \mathcal{M}) \propto p(d|\theta, \mathcal{M})\, p(\theta|\mathcal{M}). \tag{2}$$

[1] Adopting such viewpoint already provided useful tools to treat anarchical models of the SM flavour sector, see [1].

In our case $\theta \equiv \{\Lambda, \alpha_{1\ldots n}\}$ are the parameters of the higher-dimensional operators (HDOs) defined in Eq. (1). The parameter space will be denoted by $\mathcal{D}$. [2] $\mathcal{M}$ is the Standard Model extended with HDOs, and d represents the experimental data. The distribution $p(\theta|d, \mathcal{M})$ is the so-called posterior distribution, $p(d|\theta, \mathcal{M}) \equiv L(\theta)$ is the likelihood function encoding experimental data, and $p(\theta|\mathcal{M})$ is the prior distribution, which represents our a priori degree of belief in the parameters.

The posterior distribution is the core of our results. Being interested in the new physics scale, we focus on the marginal posterior $p(\Lambda|d, \mathcal{M})$, obtained by integrating all HDO coefficients α's,

$$p(\Lambda|d, \mathcal{M}) \propto \int d^n\alpha_i \; p(\Lambda, \alpha_i|\mathcal{M}) L(\Lambda, \alpha_i). \tag{3}$$

The prior and posterior distributions do not need to be normalized to unity to carry out the inference process in its broader meaning. For example, assuming some significant deviation from the SM is present in the data, it is sufficient to look at the bump in Λ's improper posterior to have a good idea about the values of Λ favored by the data. However, to go further and determine intervals associated with an actual probability (Bayesian Credible intervals), the posterior does need to be normalizable. More precisely, the posterior needs to be "proper". It should be integrable on a unbounded domain like $\mathbb{R}$. Over a bounded domain, the integral should be independent of the bound of the domain, unless the bound is well justified.

In the rest of this section we will observe that the Λ's posterior is improper. We will find the conceptual subtlety at the origin of this improperness, then propose a slight conceptual change leading to a proper Λ's posterior. In this work we consider as valuable the ability to determine Bayesian Credible intervals, and thus to have proper posteriors. However, even without paying particular interest to properness and Bayesian Credible intervals, the conceptual observations and their consequences that we will present below are in any case relevant for anyone interested in inference on Λ.

For concreteness, we give to the NP scale a logarithmically uniform distribution,

$$p(\Lambda) \propto \frac{1}{\Lambda}. \tag{4}$$

By doing so, all the orders of magnitude are given the same probability weight. This is arguably the most objective choice, justified by the "principle of indifference" [3,4]. [3] There is no sensible argument to fix the upper bound on Λ. The prior of Λ is therefore improper.

Similarly, we give uniform priors to the α's. Contrary to the domain of Λ, there are well justified bounds on α's because of perturbativity of the EFT approach. Indeed for $\Lambda > 4\pi v$, perturbativity implies $\alpha_i \in [-16\pi^2, 16\pi^2]$. We refer to [5] for more details about the bounds on HDOs coefficients. Although the priors adopted above are well motivated, the whole approach including the upcoming statements remains valid for any kind of priors, as long as the domain of α's remains bounded.

Determining the posterior distribution of Λ is a standard Bayesian procedure. However a peculiarity of the Λ posterior is that in the decoupling limit $\Lambda \to \infty$, the likelihood tends to its

[2] Notice in general $\mathcal{D}$ should also enclose the SM parameters. However this is not relevant for the present work. In the Higgs sector study we will perform, modifications of the SM parameters do appear, but they can always be expressed in terms of the HDO parameters.

[3] The "principle of indifference" maximizes the objectiveness of the priors. Once a transformation law $\gamma = f(\theta)$ irrelevant for a given problem is identified, this principle provides the most objective prior by identifying $p_\Theta \equiv p_\Gamma$ in the relation $p_\Theta(\theta)d\theta = p_\Gamma(\gamma)d\gamma$.

SM value $L \to L^{\rm SM}$. As the logarithmic prior of Λ is also improper, it turns out that the posterior distribution is improper in the Λ direction,

$$p(\Lambda \to \infty|d) \propto \frac{L_{\rm SM}}{\Lambda}. \tag{5}$$

To understand the origin of this improperness, let us introduce the notion of "testability", carrying the usual meaning as given e.g. in philosophy of science (see e.g. [6]). Considering the effective Lagrangian Eq. (1), we observe that, for $\Lambda \to \infty$, the new physics cannot manifest itself in the data. It is therefore not testable at $\Lambda \to \infty$. However, the behaviour of $p(\Lambda|d)$ in the decoupling limit does not seem to reflect this fact, as it remains constant up to the $1/\Lambda$ factor coming only from the prior.

Let us be more precise by translating the notion of testability in a formal way. We adopt the following definition as a Bayesian translation of testability. "A model is testable with respect to the SM for a given dataset d whenever $L \neq L^{\rm SM}$". We would like to know what happens to our posterior when we require testability. For a continuous parameter space, requiring testability corresponds to excising a slice $\Omega_{\Lambda, L^{\rm SM}}$ of the parameter space, defined as $\Omega_{\Lambda, L^{\rm SM}} = \{\alpha_i | L(\Lambda, \alpha_i) = L^{\rm SM}\}$. Therefore, by requiring testability of the HDOs-extended SM, inference is made on the possibilities of new physics which are actually testable by the data. Stated differently, to the initial question "What can we learn about Λ from d?", we already know that the answer is "Nothing" whenever $L = L^{\rm SM}$. We therefore discard this particular possibility, to investigate the NP which can be actually probed by d.

The fact that the requirement of testability leads to a proper posterior will be demonstrated in Section 3 and in Appendix. We admit it for the rest of this section. Requiring testability, the marginal posterior of the NP scale Λ is then expressed as

$$p^*(\Lambda|d, \mathcal{M}) = p\big(\Lambda|L \neq L^{\rm SM}, d, \mathcal{M}\big). \tag{6}$$

In our approach this distribution is the relevant object to inform us on the NP scale and will be therefore at the center of our interest for subsequent applications. We refer the reader to Section 3 for a formal discussion.

Notice this subtlety about testability usually does not matter in cases where the posterior is proper. Typically, the likelihood is continuous and bounded, such that the subdomain $\Omega_{\Lambda, L^{\rm SM}}$ has measure zero. Excluding this subdomain therefore does not change integrals of the posterior, and leaves the results of inference unchanged. The requirement of testability becomes important in our case because the posterior is improper. More generally this problem is susceptible to appear whenever the NP scale is a free parameter of the model.

Some qualitative comments can be made about the different effects driving the shape of the $p^*(\Lambda|d, \mathcal{M})$ posterior. Both tails will drop to zero, fast enough to let the distribution be integrable. Let us consider the low-Λ tail of the distribution. Even though experimental constraints push Λ to high values, it often happens that some precise cancellations between various HDOs contributions allow Λ to go to low values. However, the regions of parameter space in which precise cancellations occur have weak statistical weight by construction, such that their unnatural character is built-in the Bayesian approach (see [7] for more considerations on naturalness). We conclude that the low-Λ tail is set by the trade-off between goodness-of-fit and possible fine-tuning. Considering the high-Λ tail, if the data d are compatible with the SM, the shape is asymptotically independent of d, and is only dictated by the probability of $\mathcal{M}$ to be testable. In contrast, when d shows some deviation with respect to the SM, a good fit of the deviation favours low values of Λ. The high-Λ tail is thus shaped by the two effects. It is set by default to

a profile depending only on $\mathcal{M}$, which is overwhelmed by the shape dictated by goodness-of-fit once an excess appears in the data. The high-Λ tail behaviour can be observed in the toy model of Appendix D.

3. Inference on testable new physics

In this section we scrutinize the posterior to better understand how its integral diverges. We then show that the requirement of having testable NP leads to a proper posterior. It is necessary to use the framework of Lebesgue integration to treat rigorously the following questions. In doing so, we will introduce the Lebesgue measure μ.[4] In what follows we let Λ go to infinity, such that the designation "proper" is equivalent to "integrable". Various proofs are reported in Appendix, as well as a useful example of explicit computation within a toy model.

Let us first show that the integral of the posterior distribution diverges, i.e.

$$\int\limits_{\Lambda,\alpha_i\in\mathcal{D}} p(\Lambda,\alpha_i|d,\mathcal{M})d\mu = \infty. \tag{7}$$

To do so, let us rewrite Eq. (7) to make appear the manifold $\Omega_{\Lambda,\hat{L}}$ defined by fixed values of the likelihood,[5]

$$\int\limits_{(\Lambda,L)} d\mu \int\limits_{\Omega_{\Lambda,L}} p(\Lambda,\alpha_i|d,\mathcal{M})\frac{1}{J}d\mu(\Omega_{\Lambda,L}), \tag{8}$$

where

$$\Omega_{\Lambda,\hat{L}} = \{\alpha_i\,|L(\Lambda,\alpha_i) = \hat{L}\}. \tag{9}$$

The Jacobian is $J = (\sum_i(\partial L/\partial\alpha_i)^2)^{1/2}$. The marginal posterior in the (Λ, L) plane

$$p(\Lambda,L|d,\mathcal{M}) = \int\limits_{\Omega_{\Lambda,L}} p(\Lambda,\alpha_i|d,\mathcal{M})\frac{1}{J}d\mu \tag{10}$$

generates a measure ν over (Λ, L) such that

$$d\nu = p(\Lambda,L|d,\mathcal{M})d\mu \tag{11}$$

where μ is the Lebesgue measure. It is shown in Appendix A that $p(\Lambda, L|d, \mathcal{M})$ tends to a Dirac peak (i.e. ν tends to the Dirac measure) in the decoupling limit,

$$p(\infty, L^{\rm SM}|d,\mathcal{M}) \propto \delta(L - L^{\rm SM}). \tag{12}$$

A schematic picture of the $p(\Lambda, L|d, \mathcal{M})$ distribution is shown in Fig. 1.

Employing Radon–Nikodim (RN) decomposition, the measure ν can be decomposed as

$$\nu = \nu_c + \nu_d, \tag{13}$$

[4] This is the appropriate measure for continuous probability spaces. For brevity we will omit the argument of the integrand when no ambiguity is possible. We are going to use the extended real set $\bar{\mathbb{R}} = \mathbb{R}\cup\{-\infty,\infty\}$.

[5] We apply the coarea formula to Eq. (7), in which we identify the surjective mapping with the likelihood function $L : \mathcal{D} \to \mathbb{R}$.

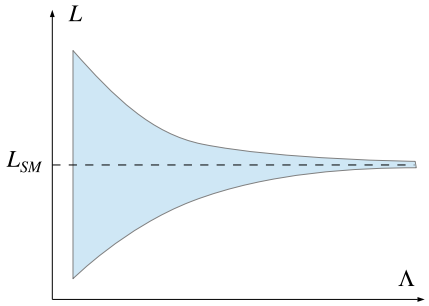

Fig. 1. A picture of the support of the posterior distribution in the (Λ, L) plane, $p(\Lambda, L|d, \mathcal{M})$. The support tends to the point $L = L_{\rm SM}$ for $\Lambda \to \infty$.

where ν_c is absolutely continuous with respect to Lebesgue measure while ν_d is discrete. The discrete measure satisfies

$$p\big(\Lambda, L|L = L^{\rm SM}, d, \mathcal{M}\big) = \frac{d\nu_d}{d\mu}, \tag{14}$$

and we can then identify our "excised" marginal posterior as

$$p^*(\Lambda|d, \mathcal{M}) = p\big(\Lambda, L|L \neq L^{\rm SM}, d, \mathcal{M}\big) = \left|\frac{d\nu_c}{d\mu}\right|. \tag{15}$$

The presence of the absolute value is related to a non-trivial subtlety in the definition of the excised probability space, that is discussed in Appendix C. In the decomposition of ν defined by Eqs. (13)–(15), it turns out that the contribution from the discrete measure ν_d is infinite,

$$\int\limits_{(\Lambda,L)} d\nu_d = \int p\big(\Lambda, L\big|L = L^{\rm SM}, d, \mathcal{M}\big)d\mu \sim \int d\Lambda\, dL\, \Lambda^{-1}\delta\big(L - L^{\rm SM}\big) = \infty. \tag{16}$$

In contrast, one can show that the ν_c measure leads to a contribution

$$\int\limits_{(\Lambda,L)} d\nu_c \propto \int p\big(\Lambda, L\big|L \neq L^{\rm SM}, d, \mathcal{M}\big)d\mu \propto \int d\Lambda\, \Lambda^{-m-1}, \tag{17}$$

which is finite, for any HDO with dimension $m + 4$. The proofs of Eqs. (16), (17) are given in Appendix B.

From this point of view, it appears that the divergent part of the posterior is localized on the subspace $\Omega_{\Lambda, L^{\rm SM}}$. It is precisely the domain where the new physics cannot be tested by the data. Requiring testability, we reduce the parameter space to $\mathcal{D}\backslash\Omega_{\Lambda, L^{\rm SM}}$, such that only the contribution Eq. (17) remains in the posterior integral. This contribution being finite, the posterior of testable NP is well proper.

We can check that the requirement of testability is harmless regarding the experimental information. Let us recall that the likelihood function comes initially from an experimental probability density function (PDF) $p_X(x)$ associated with some observable X. We assume that p_X has no discrete component. The repartition function of the observable X is

$$p(X < x) = \int\limits_{[-\infty,x]} p_X\, d\mu. \tag{18}$$

Expressing x as a function of (Λ, α_i), the likelihood function is then $L(\Lambda, \alpha_i) = p_X(x(\Lambda, \alpha_i))$. The domain $\Omega_{\Lambda, L^{\rm SM}}$ is mapped onto the SM value of the observable $x^{\rm SM}$. Excluding this domain amounts to excluding the point $x^{\rm SM}$ from the experimental density. A single point having measure zero, this leaves the repartition function unchanged. We conclude that the restriction from $\mathcal{D}$ to $\mathcal{D}\backslash\Omega_{\Lambda, L^{\rm SM}}$ leaves the experimental information invariant.

4. The MCMC setup

In the present work we are going to evaluate posterior distributions by means of a Markov Chain Monte Carlo (MCMC) method. The basic idea of a MCMC is setting a random walk in the parameter space such that the density of points asymptotically reproduces a target function, in our case the posterior distribution. Any marginalization is then performed through a simple binning of the points of the Markov chain along the appropriate dimension. We refer to [8,2] for details on MCMCs and Bayesian inference. Our MCMC method uses the Metropolis–Hastings algorithm with a symmetric, Gaussian proposal function. We check the convergence of our chains using an improved Gelman and Rubin test with multiple chains [9]. The first 10^4 iterations are discarded (burn-in).

Some precautions about the MCMC method are necessary regarding the subtleties about improper posteriors discussed in Sections 2, 3. Indeed, using the MCMC method, we are not working with the exact continuous posterior distributions, as the one discussed in Section 3. Instead, we are manipulating histograms which are estimators of the exact posteriors. These estimators are discrete distributions

$$\hat{p}_{N,\Delta^{(n)}}(\Lambda, \alpha_i | d, \mathcal{M}), \tag{19}$$

where N is the number of points and $\Delta^{(n)}$ is the bin size along the various dimensions. The estimator tends to its estimand $p(\Lambda, \alpha_i|d, \mathcal{M})$ when $N \to \infty$, $\Delta \to 0$, i.e. in the continuum limit with infinite sampling.

Notice the bin size can be optimized for a given N. Too large bins give a poor estimation of the distribution, while too thin bins suffer from large binomial noise. It exists therefore an optimal bin size to minimize estimators uncertainty. As far as we know it is commonly determined in a ad-hoc way. We proceed similarly in this note.

In the continuous case, we found in Section 3 that the $L = L^{\rm SM}$ subdomain (i.e. $\Omega_{\Lambda, L^{\rm SM}}$) shall be excluded to obtain a proper posterior. This feature is translated into the discrete estimator case as follows. Let us evaluate $\hat{p}$ without the $L \neq L^{\rm SM}$ restriction. Considering $\hat{p}$ in the (Λ, L) plane, for $\Lambda \to \infty$, the only non-zero bin of $\hat{p}$ is the bin containing the value $L^{\rm SM}$. This is the discrete equivalent of the Dirac peak obtained in Eq. (12). To obtain the estimator of $p(\Lambda|L \neq L^{\rm SM}, d, \mathcal{M})$, we have therefore to excise this bin. This is the discrete equivalent of the $L \neq L^{\rm SM}$ restriction. The fact that we exclude a seemingly finite slice of the parameter space should not be surprising, as for the estimator $\hat{p}$, space is not continuous but discrete. Finally, the upper bound $\Lambda < \Lambda_{max}$ also has to be finite in practice. For a given finite N and a given bin size, there exists a finite value $\Lambda = \tilde{\Lambda}$ above which all points of $\hat{p}$ are in the $L^{\rm SM}$ bin. In practice one has therefore to make sure that Λ_{max} is large enough such that $\tilde{\Lambda} < \Lambda_{max}$.

5. Probing Λ in the Higgs sector

In this section we apply the inference process defined through Sections 2–4 to the Standard Model Higgs sector extended with higher-dimensional operators. The theoretical treatment of

HDOs and the analysis of data we used are the same as realized in the recent work [5]. Here we briefly review the main points of the analysis, and refer to this work for any further theoretical and experimental details.

The Higgs sector is supplemented by a set of CP-even dimension-6 operators, whose basis is chosen to be[6]

$$\mathcal{O}_6 = |H|^6, \qquad \mathcal{O}_{D^2} = |H|^2 |D_\mu H|^2, \qquad \mathcal{O}'_{D^2} = \left|H^\dagger D_\mu H\right|^2, \tag{20}$$

$$\mathcal{O}_{WW} = H^\dagger H \left(W^a_{\mu\nu}\right)^2, \qquad \mathcal{O}_{BB} = H^\dagger H \left(B_{\mu\nu}\right)^2, \qquad \mathcal{O}_{WB} = H^\dagger W_{\mu\nu} H B_{\mu\nu}, \tag{21}$$

$$\mathcal{O}_{GG} = H^\dagger H \left(G^a_{\mu\nu}\right)^2, \tag{22}$$

$$\mathcal{O}_D = J^a_{H\,\mu} J^a_\mu, \qquad \mathcal{O}'_D = J^Y_{H\,\mu} J^Y_\mu, \tag{23}$$

$$\mathcal{O}_f = 2 y_f |H|^2 H \bar{f}_L f_R. \tag{24}$$

Here J_H and J_f are $SU(2)$ or $U(1)_Y$ currents involving the Higgs field and the fermion f respectively, and $J = \sum_f J_f$ are the SM fermion currents coupled to B_μ and W_μ.

This choice of basis is such that the field strength–Higgs operators $\mathcal{O}_{FF}$'s cannot be generated at tree-level in a perturbative UV theory. We therefore consider two general cases, "democratic HDOs" and "loop-suppressed $\mathcal{O}_{FF}$'s", depending on whether or not the $\mathcal{O}_{FF}$'s are loop-suppressed with respect to the other HDOs. Moreover, in important classes of models like for the R-parity conserving MSSM, the HDOs can only be generated at one-loop. We will therefore consider two cases within the democratic HDOs scenario, one with tree-level HDOs, $\alpha_i \in [-16\pi^2, 16\pi^2]$, and one with loop-level HDOs, $\alpha_i \in [-1, 1]$. For the case of loop-suppressed $\mathcal{O}_{FF}$'s, we assume that the unsuppressed HDOs are generated at tree-level. We therefore investigate three scenarios whose features are summarized in Table 1. In case of tree-level HDOs, perturbativity of the HDO expansion $|\alpha|/\Lambda^2 < 1/v^2$ imposes an additional constraint for $\Lambda < 4\pi v$. We take custodial symmetry to be an exact symmetry of the theory, such that the operators $\mathcal{O}'_D$, $\mathcal{O}'_{D^2}$ are set to zero. Finally, we emphasize that these scenarios are generic, in the sense that they encompass all known UV models in addition to the realizations not yet thought of. This implies that features predicted only by specific UV models, like suppression of HDOs or precise cancellations between HDOs, will get a small statistical weight, as we consider the whole set of UV realizations.

Concerning data, we take into account the results from Higgs searches at the LHC and at Tevatron as well as electroweak precision observables and trilinear gauge couplings. Higgs results [28, 10–27] have to be exploited with care as HDOs modify both Higgs decays and production. We use results (partly) accounting for correlations between the subchannels when they are available. When estimated decomposition into production channels are unavailable, we take the relative ratios of production cross sections for a SM Higgs [29,28] as a reasonable approximation. The Higgs mass is set to $m_h = 125.5$ GeV, close to the combined mass measurement from the two experiments, since it is not yet possible to take it as a nuisance parameter without losing the correlations between production channels. We take into account the electroweak precision observables using the Peskin–Takeuchi S and T parameters [30,31]. Beyond S and T, the W and Y parameters [32] should be used in the HDO framework. However the constraints arising from these parameters are by far negligible with respect to our other constraints. Experimental values

[6] The operator $\mathcal{O}_6$ plays no role in what follows and is listed here only for completeness.

Table 1
Summary of the setup of the scan in the three scenarios we consider. The α_{FF} coefficients (where $FF = WW, WB, BB, GG$) correspond to the field-strength–Higgs operators. In both cases we take custodial symmetry to be an unbroken symmetry.

	Democratic HDOs		Loop-suppressed $\mathcal{O}_{FF}$'s
	Tree-level	One-loop	
Λ	$[v, \infty[$	$[v, \infty[$	$[v, \infty[$
$\lvert\alpha_{FF}\rvert$	$[0, \Lambda^2/v^2]$ if $\Lambda < 4\pi v$ $[0, 16\pi^2]$ else,	$[0, 1]$	$[0, 1]$
Other $\lvert\alpha\rvert$	$[0, \Lambda^2/v^2]$ if $\Lambda < 4\pi v$ $[0, 16\pi^2]$ else,	$[0, 1]$	$[0, \Lambda^2/v^2]$ if $\Lambda < 4\pi v$ $[0, 16\pi^2]$ else.

of S and T are taken from the latest SM fit [33], $S = 0.05 \pm 0.09$ and $T = 0.08 \pm 0.07$ with a correlation coefficient of 0.91. Regarding constraints on TGV, we take into account the LEP measurements [34].

Applying the method described in Sections 2–4, we obtain the normalizable posterior distributions $p^*(\Lambda|d)$. One can always normalize them to unity such that we will designate them as probability density functions (PDFs). It turns out that the posterior PDF of the NP scale for tree-level and one-loop democratic HDO is essentially the same under a shift $\log_{10}\Lambda \to \log_{10}\Lambda - \log_{10}(4\pi) \approx \log_{10}\Lambda - 1.10$, i.e. a rescaling $\Lambda \to \Lambda/4\pi$. This happens because the region $|\alpha| \in [0, \Lambda^2/v^2]$ for tree-level HDOs has a negligible impact on the posterior, such that the tree-level and one-loop scenarios can be identified through a rescaling of Λ. The posterior PDFs of the NP scale for the various scenarios are shown in Fig. 2. The 68% and 95% Bayesian credible intervals (BCIs) of Λ for democratic HDOs are respectively $[200, 1400]$, $[123, 3300]$ TeV for tree-level HDOs and $[16, 110]$, $[9.8, 260]$ TeV for loop-level HDOs. We find 68% and 95% BCIs of $[62, 533]$, $[28, 1200]$ TeV for the scenario of loop-suppressed $\mathcal{O}_{FF}$'s.

We find the leading constraint on Λ to be the Higgs data for democratic HDOs, while these are the electroweak observables for loop-suppressed $\mathcal{O}_{FF}$'s. This can be understood as follows. The $\mathcal{O}_{FF}$ operators are mapped onto field strength–Higgs anomalous couplings, among which the $\zeta_g h(G_{\mu\nu})^2$ and $\zeta_\gamma h(F_{\mu\nu})^2$ couplings. Given that the corresponding SM couplings are generated at one-loop, $\zeta_{g,\gamma}$ need to be sensibly suppressed to not induce large deviations in the predictions of gluon fusion and $h \to \gamma\gamma$ processes (see [5] for details). For democratic HDOs, this need of small $\zeta_{g,\gamma}$ pushes Λ to high values in order to suppress the $\mathcal{O}_{FF}$'s. In contrast, for the scenario of loop-suppressed $\mathcal{O}_{FF}$'s, the $\zeta_{g,\gamma}$'s are already loop-suppressed with respect to other anomalous couplings by assumption. This alleviates the aforementioned constraint, leaving the S, T measurements as leading constraints.

Having identified the leading constraints, we may comment about the necessary conditions allowing more specific UV models to reach lower values of Λ. For models having democratic HDOs, a suppressed $\mathcal{O}_{GG}$ is required to reduce the ζ_g coupling. The ζ_γ coupling being proportional to $s_w^2\alpha_{WW} + c_w^2\alpha_{BB} - \frac{1}{2}c_w s_w \alpha_{WB}$, precise cancellations among these various terms may occur within an appropriate UV model, while they are unprobable (i.e. fine-tuned) in the generic scenario. Note both conditions on ζ_g and ζ_γ need to be fulfilled in order to lower the values of Λ. If only one of the ζ's is suppressed, the outcome will still remain similar to the left plot of Fig. 2. This occurs in particular when these ζ's are generated perturbatively. In that case one has $\zeta_g/\zeta_\gamma \approx g_s^2/g_Y^2 \gg 1$, such that ζ_γ is naturally suppressed with respect to ζ_g, which then becomes the leading constraint. Concerning models with loop-suppressed $\mathcal{O}_{FF}$'s, the main condition to

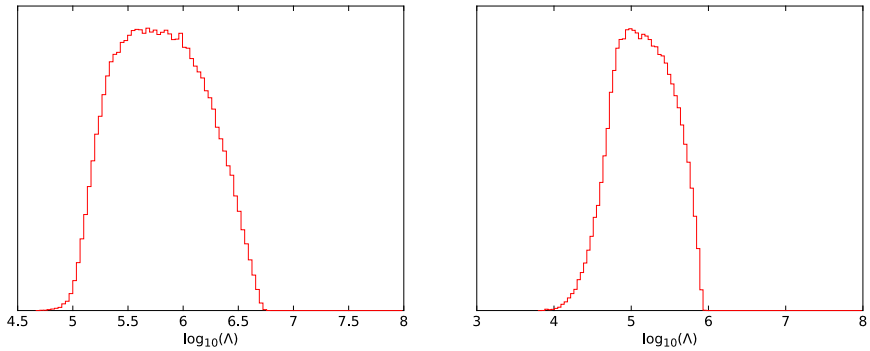

Fig. 2. Posterior PDFs of Λ ($p^*(\Lambda|d,\mathcal{M})$) in GeV units for tree-level democratic HDOs (right) and loop-suppressed $\mathcal{O}_{FF}$'s (left). The PDF of Λ for one-loop democratic HDOs has the same shape as for tree-level democratic HDOs and is shifted by $\log_{10} 4\pi$ towards low values of Λ.

reach a lower Λ is to have a suppressed $\mathcal{O}_D$. This operator induces the main contribution to the S parameter, $\alpha S \approx s_w^2 \alpha_D v^2/\Lambda^2$, other contributions to S, T being loop-suppressed (see [5]).

The PDFs presented above are given for an optimal size of the bins.[7] To exemplify the uncertainty associated with the MCMC estimation of the PDFs, we compute the 95% BCIs obtained with twice more bins and twice less bins. We find the variations over $\log_{10}\Lambda$ to be $O(2\%)$. The origin of these variations lies in the uncertainty inherent to the concrete estimation method presented in Section 4, and is not related to the formal inference process described in Sections 2, 3.

6. Conclusion

Whenever one considers a set of data – showing or not a significant deviation from the Standard Model, it is interesting to ask what information can be obtained about the energy scale of a possible underlying new physics. We present a method to consistently infer the distribution of Λ from any dataset. In doing so we use a statistical view of the unknown NP parametrized by higher-dimensional operators. To obtain a proper posterior, necessary to create Bayesian credible intervals, we point out the requirement that NP has to be testable by the data.

We formally demonstrate using Lebesgue integration that this requirement implies proper posteriors. In doing so we introduce a subspace where the likelihood itself is taken as a random variable. Some conceptual subtleties related to this trick are discussed, and a helpful toy model is introduced in the appendix. Given that Monte Carlo Markov Chains methods are commonly used to realize statistical inference, we describe the concrete implementation of this inference process in MCMCs.

As an illustration, we apply our approach to the SM Higgs sector, in light of recent data. Building on the recent work [5], we consider the scenarios of democratic HDOs and loop-suppressed $\mathcal{O}_{FF}$'s. For democratic HDOs, we obtain 95% Bayesian credible intervals of $[123, 3300]$ TeV and $[9.8, 260]$ TeV, for respectively tree-level and loop-generated HDOs. For loop-suppressed $\mathcal{O}_{FF}$'s, we find the 95% BCI $[28, 1200]$ TeV, assuming that unsuppressed HDOs are generated at tree-level. More specific UV models suppressing some particular HDOs or predicting otherwise fine-tuned relations are necessary to favour lower values of the NP scale.

[7] Namely we use 400 bins for a sampling $N = O(2 \times 10^6)$.

Acknowledgements

The author would like to thank Gero von Gersdorff and Béranger Dumont for intense discussions relative to this study. The author is grateful to Ivan Gordeli and Gero von Gersdorff for reading the manuscript. The author acknowledges the Brazilian Ministry of Science, Technology and Innovation for financial support, and the LPSC and the Ecole Polytechnique for hospitality during part of this work.

Appendix

The effective Lagrangian equation (1) contains in general HDOs of arbitrary dimensions. Here, for simplicity we will consider HDOs of a single dimension labelled by m. The generalization to the case of HDOs with arbitrary dimension is straightforward.

Appendix A. Asymptotics of $p(\Lambda, L|d, \mathcal{M})$

Let us show that the marginal posterior in the (Λ, L) plane

$$p(\Lambda, L|d, \mathcal{M}) = \int\limits_{\Omega_{\Lambda,L}} p(\Lambda, \alpha_i|d, \mathcal{M}) \frac{1}{J}\, d\mu \tag{A.1}$$

tends to be proportional to the Dirac peak in the decoupling limit,

$$p(\Lambda = \infty, L) \propto \delta\big(L - L^{\mathrm{SM}}\big). \tag{A.2}$$

Proof. In the decoupling limit $\Lambda \to \infty$, L tends to L^{SM}. Thus, for any arbitrary small $\delta L > 0$, it exists a finite $\Lambda = \tilde{\Lambda}$ such that $\Omega_{\Lambda,L} = \emptyset$ for any $\Lambda > \tilde{\Lambda}$ and $|L - L^{\mathrm{SM}}| > \delta L$. In the decoupling limit with $L \neq L^{\mathrm{SM}}$, the integration domain $\Omega_{\Lambda > \tilde{\Lambda}, L}$ therefore reduces to the null set. This implies

$$p\big(\Lambda \to \infty, L \big| L \neq L^{\mathrm{SM}} d, \mathcal{M}\big) = 0. \tag{A.3}$$

Let us now study the behaviour for $L = L^{\mathrm{SM}}$. Defining $\partial_i L = \partial L(x_i)/\partial x_i$, a Λ^{-m} factor out from the Jacobian $J = (\sum_i (\partial L/\partial \alpha_i)^2)^{1/2}$,

$$J = \Lambda^{-m} \left(\sum_i (\partial_i L)^2 \right)^{1/2}. \tag{A.4}$$

The $\partial_i L$ are finite by hypothesis, such that $J = O(\Lambda^{-m})$. The posterior $p(\Lambda, \alpha_i|d, \mathcal{M})$ is therefore $O(\Lambda^{m-1})$ once one takes into account the log prior. For any $m \geqslant 2$, the distribution gets therefore infinite,

$$p\big(\Lambda = \infty, L | L = L^{\mathrm{SM}}, d, \mathcal{M}\big) = \infty. \tag{A.5}$$

We deduce from Eqs. (A.3), (A.5) that $p(\Lambda = \infty, L|L = L^{\mathrm{SM}}, d, \mathcal{M})$ is proportional to a Dirac peak centred on $L = L^{\mathrm{SM}}$. □

Appendix B. Integration of the posterior

Starting from

$$dv_d = p\big(\Lambda, L\big|L \neq L^{\rm SM}, d, \mathcal{M}\big)d\mu, \tag{B.1}$$

$$d|v_c| = p\big(\Lambda, L\big|L = L^{\rm SM}, d, \mathcal{M}\big)d\mu, \tag{B.2}$$

we want to show that $\int_{(\Lambda,L)} dv_d = \infty$ and that $\int_{(\Lambda,L)} dv_c$ is finite. One assumes $v_c > 0$.

Proof. Let us denote the marginal posterior along Λ as

$$p(\Lambda, d, \mathcal{M}) \equiv f(\Lambda) \tag{B.3}$$

for simplicity. We write $f = f^* + f^{\rm SM}$ with

$$f^*(\Lambda) = \int p\big(\Lambda, L|L \neq L^{\rm SM}, d, \mathcal{M}\big)\, d\mu(L) \tag{B.4}$$

and

$$f^{\rm SM}(\Lambda) = \int p\big(\Lambda, L|L = L^{\rm SM}, d, \mathcal{M}\big)\, d\mu(L), \tag{B.5}$$

such that $\int dv_d = \int f^{\rm SM}(\Lambda)\, d\mu$ and $\int dv_c = \int f^*(\Lambda)\, d\mu$.

We define the simple function

$$f_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\mathbb{I}(E_k) + n\mathbb{I}(n,\infty), \tag{B.6}$$

with $E_k = [f^{-1}(\frac{k}{2^n}), f^{-1}(\frac{k+1}{2^n})]$. f_n converges pointwise to f and we have $f_n(x) \leqslant f(x)$, such that $\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$ by the Monotone Convergence Theorem (MCT). We define $k^{\rm SM}$ such that $k^{\rm SM}/n < f^{\rm SM} < (k^{\rm SM}+1)/n$. We then have $f_n = f_n^{\rm SM} + f_n^*$ with

$$f_n^{\rm SM} = \frac{k^{\rm SM}}{2^n}\mathbb{I}\big(E_k^{\rm SM}\big), \tag{B.7}$$

$$f_n^* = \sum_{k=0,k\neq k^{\rm SM}}^{n2^n-1} \frac{k}{2^n}\mathbb{I}(E_k) + n\mathbb{I}(n,\infty), \tag{B.8}$$

and $\lim_{n\to\infty} f_n^{\rm SM} = f^{\rm SM}$, $\lim_{n\to\infty} f_n^* = f^*$.

Let us compute $\int f_n^{\rm SM} d\mu$ where μ is the Lebesgue measure. Given that $L \to L^{\rm SM}$ for $\Lambda \to \infty$, for any arbitrary small $\delta L = k^{\rm SM}/2^n - L^{\rm SM}$, it exists a finite $\tilde{\Lambda}$ such that $f \in E_{k^{\rm SM}}$ for any $\Lambda \in [\tilde{\Lambda}, \infty]$. Therefore $\mu(E_{k^{\rm SM}}) = \infty$. This implies $\int f_n^{\rm SM}(\Lambda)\, d\mu = \infty$, then $\int f^{\rm SM}(\Lambda)\, d\mu = \infty$ by the MCT, and thus $\int dv_d = \infty$.

Let us now compute $\int f_n^* d\mu$. $\mu(E_{k\neq k^{\rm SM}})$ is finite. We have to show that the sum over n converges. To do so we first simplify f^* using the $\Lambda \to \infty$ limit. The Limit Comparison Test (LCT) will ensure that the simpler function has the same integrability features as f^*. We will denote the successive simplified functions by $\hat{f}^*$.

We factorize the Λ prior and factorize the likelihood function out from the first integral such that

$$f^*(\Lambda) = \frac{1}{\Lambda}\int d\mu(L)\, L \int\limits_{\Omega_{\Lambda,L}} \frac{1}{J}\, d\mu(\Omega_{\Lambda,L}). \tag{B.9}$$

We replace the derivatives $\partial_i L$ in the Jacobian of Eq. (B.9) by their values at $\Lambda \to \infty$. The LCT ensures that this simplified function as the same integrability properties as Eq. (B.9) as their limits for $\Lambda \to \infty$ are the same. We can then integrate over $d\mu(\Omega_{\Lambda,L})$ and obtain

$$\hat{f}^* \propto \Lambda^{m-1} \int d\mu(L)\, L\, \mu(\Omega_{\Lambda,L}). \tag{B.10}$$

For any finite $\tilde{\Lambda}$, one can expand the likelihood with respect to $\tilde{\Lambda}/\Lambda$,

$$L = L^{\rm SM} + \frac{\partial L}{\partial \Lambda^{-1}}\bigg|_{\Lambda\to\infty} \cdot \frac{1}{\Lambda} + O\left(\frac{\tilde{\Lambda}^2}{\Lambda^2}\right). \tag{B.11}$$

L can be reexpressed as

$$L = L^{\rm SM} + \partial_i L|_\infty \cdot \alpha_i \frac{\tilde{\Lambda}}{\Lambda} + O\left(\frac{\tilde{\Lambda}^2}{\Lambda^2}\right). \tag{B.12}$$

The LCT ensures that one can replace L by its truncated expansion to study the integrability of f^*. In this limit, $\mu(\Omega_{\Lambda,L})$ is the volume of a hyperplane in the $\{\alpha_i\}$ space defined by Eq. (B.12). We can write it as $\mu(\Omega_{\Lambda,L}) = \mathcal{V}\{(L - L^{\rm SM})^2 \Lambda^{2m}\}$ such that the squared likelihood and the Λ dependence appear explicitly. We are left with studying the integrability of

$$\hat{f}^* = \Lambda^{m-1} \int d\mu(L)\, L\, \mathcal{V}\big\{\big(L - L^{\rm SM}\big)^2 \Lambda^{2m}\big\}. \tag{B.13}$$

Changing variable $(L - L^{\rm SM})^2\Lambda^{2m} \to (L - L^{\rm SM})^2$ factors out a Λ^{-2m} factor. The remaining integral $\frac{1}{2}\int d(L - L^{\rm SM})^2\, \mathcal{V}\{(L - L^{\rm SM})^2\}$ gives $\frac{1}{2}\mu(\{\alpha_i\})$ which is bounded by hypothesis.[8] For example in the tree-level democratic HDOs case we have $\mu(\{\alpha_i\}) = (32\pi^2)^m$. We have therefore $\hat{f}^* = \frac{1}{2}\Lambda^{-m-1}\mu(\{\alpha_i\})$. $\hat{f}^*$ being Riemann integrable over $[\Lambda_{min}, \infty]$ and absolutely convergent, it is therefore Lebesgue integrable. We deduce that $\int \hat{f}_n^* d\mu$ converges for $n \to \infty$, thus $\int f_n^* d\mu$ converges as well by the LCT, the integral $\int f^* d\mu$ is therefore finite and so is $\int d\nu_c$. □

Appendix C. Probability definition in the excised space

Here we discuss the subtlety that leads to the apparition of the absolute value on $|d\nu_c/d\mu|$ in Eq. (15). We stress that this discussion mainly matters at the formal level. In practice, for example when computing the posterior $p^*(\Lambda|d)$ using the MCMC method of Section 4, this question will not appear.

First, notice that we expressed our posterior distribution as a function of the likelihood L. This is perfectly allowed, as the likelihood can be just seen as a random variable as another. However the likelihood is also a conditional probability. Our "excised" space $\mathcal{D}\backslash\Omega_{\Lambda,L^{\rm SM}}$ is thus rather particular.

Second, let us note that in the Kolmogorov axioms of probability, the positivity axiom can be seen as a simple sign convention. For any sample space Ω, requiring $p(\Omega) = -1$ and $p(E) \leqslant 0$, $\forall E \in \Omega$, the subsequent results just change by a sign flip. Let us denote by $p^{(-)}$ the probabilities defined in this way, and by $p^{(+)}$ the usual positive probabilities. One of the consequence of using the $p^{(-)}$ system is that the expectation of a random variable X is given by

$$\langle X\rangle = -\int dx\, x\, p_X^{(-)}(x). \tag{C.1}$$

[8] Recall that this is imposed by perturbativity of couplings of the UV theory.

When using such convention, a crucial point is that the conditional probabilities must still be taken positive, contrary to the actual probabilities – inconsistencies would appear otherwise due to the probability multiplications. The freedom to switch between the $p^{(+)}$ and $p^{(-)}$ system of probabilities is just a symmetry of the classical probability theory.

Keeping these points in mind, let us now compute the naive expectation $\langle L^*\rangle'_\Lambda$ of the likelihood L over the excised parameter space $\mathcal{D}\backslash\Omega_{\Lambda,L^{\rm SM}}$. To do so we use the RN decomposition of Eq. (13), and obtain

$$\langle L^*\rangle'_\Lambda = \int L\, d\nu_c = \langle L\rangle_\Lambda - L^{\rm SM}. \tag{C.2}$$

It is clear that $\langle L\rangle_\Lambda$ is not necessarily larger than $L^{\rm SM}$. This typically happens when data disfavor the model with respect to the SM. We deduce that $\langle L^*\rangle'_\Lambda$ can take both signs. But in the two paragraphs above, we emphasized that L is a conditional probability, and as such must be positive whatsoever. We conclude that ν_c has to be taken as a probability measure of the $p^{(-)}$ kind, whenever $\langle L\rangle_\Lambda - L^{\rm SM} < 0$. The actual expectation is then $\langle L^*\rangle_\Lambda = -\int L\, d\nu_c$, which is positive as it should. We thus end up with the prescription that the measure ν_c is taken as a probability $p^{(+)}$ or $p^{(-)}$ when $\langle L\rangle_\Lambda - L^{\rm SM}$ is respectively positive or negative. Finally, as soon as we restrict ourselves to the excised space, we always have the freedom to switch between $p^{(-)}$ and $p^{(+)}$. Choosing to deal only with $p^{(+)}$, the probability density over the excised space is expressed as

$$p^*(\Lambda|d,\mathcal{M}) = \begin{cases} d(\nu-\nu_s)/d\mu & \text{if } \langle L\rangle_\Lambda - L^{\rm SM} > 0, \\ d(\nu_s-\nu)/d\mu & \text{if } \langle L\rangle_\Lambda - L^{\rm SM} < 0, \end{cases} \tag{C.3}$$

hence the absolute value in Eq. (15).

Appendix D. The BSM coin

To exemplify our approach, let us adopt a simple NP model. Suppose that the SM predicts that a certain coin is fair. It comes Heads or Tails with probability $1/2$. Suppose that a HDO modifies the probability such that the coin is not fair anymore,[9]

$$p(H|\Lambda,\alpha) = 1/2 + \alpha/\Lambda, \qquad p(T|\Lambda) = 1/2 - \alpha/\Lambda. \tag{D.1}$$

The SM is recovered for $\Lambda \to \infty$, or if $\alpha = 0$.

Λ is given a logarithmic prior over $[2,\infty[$, α is given a flat prior over $[-1,1]$. Let us assume that the coin is tossed twice and comes down "H, T". We toss the coin only twice for simplicity of the subsequent expressions. A more complicated likelihood would unnecessarily complicate the formulas. In doing so, data favor the SM hypothesis. We can thus expect a likelihood $\langle L\rangle_\Lambda < L_{\rm SM}$.

The SM likelihood $L_{\rm SM}$ is

$$p(\text{HT},\text{SM}) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}. \tag{D.2}$$

We now work out the SM+HDO likelihood,

$$p(\text{HT}|\Lambda,\alpha) = (1/2+\alpha/\Lambda)(1/2-\alpha/\Lambda) = \frac{1}{4} - \frac{\alpha^2}{\Lambda^2}. \tag{D.3}$$

[9] We are grateful to the referee for pointing out to us this simple example.

Let us first compute the posterior PDF of Λ without any "excision". It is given by

$$p(\Lambda|\mathrm{HT}) \propto \int d\alpha\, p(\mathrm{HT}|\Lambda,\alpha)\, p(\Lambda)\, p(\alpha) \tag{D.4}$$

$$p(\Lambda|\mathrm{HT}) \propto \int_{-1}^{1} d\alpha \left(\frac{1}{4} - \frac{\alpha^2}{\Lambda^2}\right) \frac{1}{\Lambda}\,\frac{1}{2} \tag{D.5}$$

$$p(\Lambda|\mathrm{HT}) \propto \left(\frac{1}{4\Lambda} - \frac{1}{3\Lambda^3}\right). \tag{D.6}$$

As expected the posterior $p(\Lambda|\mathrm{HT})$ is not integrable over $[2,\infty[$, i.e. it is improper.

Let us now proceed to the excision. What we want to compute is the distribution $p^*(\Lambda|\mathrm{HT})$,

$$p^*(\Lambda|\mathrm{HT}) = \int_{L\neq L_{\mathrm{SM}}} p(\Lambda, L|\mathrm{HT})\, dL. \tag{D.7}$$

From a one-to-one variable change using $\alpha = \Lambda\sqrt{1/4 - L}$, we compute

$$p(\Lambda, L|\mathrm{HT}) \propto \frac{L}{\sqrt{1-4L}}. \tag{D.8}$$

The measure

$$d\nu = p(\Lambda, L|\mathrm{HT})dL \tag{D.9}$$

is singular in $L = 1/4 = L_{\mathrm{SM}}$, such that one can decompose it as $d\nu = d\nu_d + d\nu_c$ where

$$d\nu_d = p(\Lambda, L|L = L_{\mathrm{SM}}, \mathrm{HT})\delta(L - L_{\mathrm{SM}})dL. \tag{D.10}$$

$$d|\nu_c| = p(\Lambda, L|L \neq L_{\mathrm{SM}}, \mathrm{HT})|dL. \tag{D.11}$$

Plugging the decomposition into the integral of Eq. (D.7), we have

$$p^*(\Lambda|\mathrm{HT}) = \left|\int p(\Lambda, L|\mathrm{HT})dL - p(\Lambda, L_{\mathrm{SM}}|\mathrm{HT})\right|. \tag{D.12}$$

Let us work out the two terms on the right-hand side of the equation above. The first one is just Eq. (D.6) written differently,

$$\begin{aligned}\int p(\Lambda, L|\mathrm{HT})dL \propto \int_{\frac{1}{4}-\frac{1}{\Lambda^2}}^{\frac{1}{4}} dL\, \frac{L}{\sqrt{1-4L}} &= -\frac{1}{12}\Big[\sqrt{1-4L}(2L+1)\Big]_{\frac{1}{4}-\frac{1}{\Lambda^2}}^{\frac{1}{4}} \\ &= \left(\frac{1}{4\Lambda} - \frac{1}{3\Lambda^3}\right).\end{aligned} \tag{D.13}$$

The second one is

$$p(\Lambda, L = L_{\mathrm{SM}}|\mathrm{HT}) \propto \frac{L_{\mathrm{SM}}}{\Lambda} = \frac{1}{4\Lambda}. \tag{D.14}$$

The proportionality constant is the same for both terms. The divergent piece cancels between both terms, leaving

$$p^*(\Lambda|\mathrm{HT}) \propto \frac{1}{3\Lambda^3}. \tag{D.15}$$

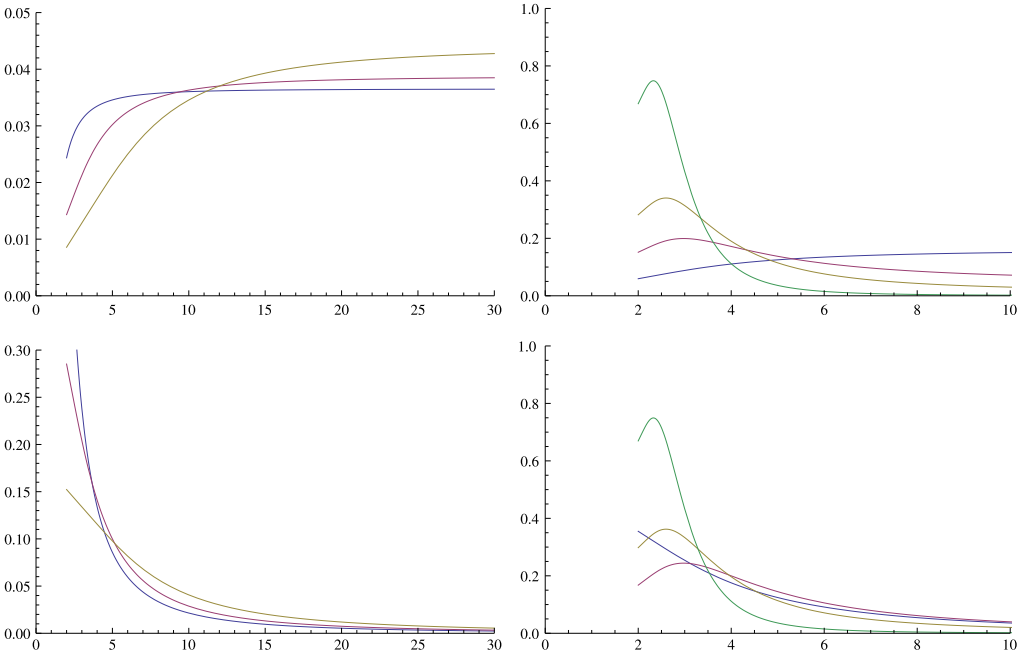

Fig. 3. Examples of $p(\Lambda|d)$ and $p^*(\Lambda|d)$ distributions for the BSM coin, assuming various data. Left panel: $p(\Lambda|d) \times \Lambda$ (top) and $p^*(\Lambda|d) \times \Lambda$ (bottom) distributions for $(H, T) = (1, 1), (5, 5), (20, 20)$ in respectively blue, purple, yellow. Right panel: $p(\Lambda|d) \times \Lambda$ (top) and $p^*(\Lambda|d) \times \Lambda$ (bottom) distributions for $(H, T) = (5, 5), (5, 15), (5, 20), (5, 30)$ in respectively blue, purple, yellow, green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

As a final illustration, $p(\Lambda|d)$ and $p^*(\Lambda|d)$ are shown on Fig. 3 for various outcomes of the BSM coin tossing. As discussed in Section 2, the shapes remain roughly identical when data are compatible with the SM. In contrast, a bump appears in $p(\Lambda|d)$ when the data favor the BSM hypothesis. The high-Λ tail of $p^*(\Lambda|d)$ drops increasingly quick with the increase of BSM evidence.

References

[1] F. Brummer, S. Fichet, S. Kraml, The supersymmetric flavour problem in 5D GUTs and its consequences for LHC phenomenology, J. High Energy Phys. 1112 (2011) 061, arXiv:1109.1226 [hep-ph].
[2] R. Trotta, Bayes in the sky: Bayesian inference and model selection in cosmology, Contemp. Phys. 49 (2008) 71, arXiv:0803.4089 [astro-ph].
[3] J. Press, Subjective and Objective Bayesian Statistics: Principles, Models, and Applications, 2nd edition, Wiley Series in Probability and Statistics, 1996.
[4] E.T. Jaynes, Prior probabilities, IEEE Trans. Syst. Sci. Cybern. 4 (3) (1968) 227–241.
[5] B. Dumont, S. Fichet, G. von Gersdorff, A Bayesian view of the Higgs sector with higher dimensional operators, arXiv:1304.3369 [hep-ph].
[6] T.S. Kuhn, The Structure of Scientific Revolutions, University of Chicago Press, 1996.
[7] S. Fichet, Quantified naturalness from Bayesian statistics, Phys. Rev. D 86 (2012) 125029, arXiv:1204.4940 [hep-ph].
[8] B.C. Allanach, C.G. Lester, Multi-dimensional mSUGRA likelihood maps, Phys. Rev. D 73 (2006) 015013, arXiv:hep-ph/0507283.
[9] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Stat. Sci. 7 (1992) 457–511.

[10] ATLAS-CONF-2013-012, Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb^{-1} of proton–proton collision data.
[11] ATLAS-CONF-2013-014, Combined measurements of the mass and signal strength of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton–proton collision data.
[12] ATLAS-CONF-2013-013, Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton–proton collision data.
[13] ATLAS-CONF-2013-030, Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton–proton collision data.
[14] ATLAS-CONF-2013-034, Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton–proton collision data.
[15] ATLAS-CONF-2012-161, Search for the Standard Model Higgs boson in produced in association with a vector boson and decaying to bottom quarks with the ATLAS detector.
[16] CMS-PAS-HIG-13-001, Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel.
[17] CMS-PAS-HIG-13-002, Properties of the Higgs-like boson in the decay $H \to ZZ \to 4\ell$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV.
[18] CMS-PAS-HIG-13-003, Evidence for a particle decaying to W^+W^- in the fully leptonic final state in a standard model Higgs boson search in pp collisions at the LHC.
[19] CMS-PAS-HIG-12-039, Search for SM Higgs in $WH \to WWW \to 3\ell 3\nu$.
[20] CMS-PAS-HIG-12-042, Evidence for a particle decaying to W^+W^- in the fully leptonic final state in a standard model Higgs boson search in pp collisions at the LHC.
[21] CMS-PAS-HIG-12-045, Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV.
[22] CMS-PAS-HIG-12-044, Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks for HCP 2012.
[23] CMS-PAS-HIG-12-025, Search for Higgs boson production in association with top quark pairs in pp collisions.
[24] CMS-PAS-HIG-13-004, Search for the Standard-Model Higgs boson decaying to tau pairs in proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV.
[25] CMS-PAS-HIG-13-006, Search for the standard model Higgs boson in the Z boson plus a photon channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV.
[26] Aurelio Juste, Standard Model Higgs boson searches at the Tevatron, talk at HCP2012, 15 Nov. 2012, Kyoto, Japan, http://kds.kek.jp/conferenceDisplay.py?confId=9237.
[27] Yuji Enari, $H \to b\bar{b}$ from Tevatron, talk at HCP2012, 14 Nov. 2012, Kyoto, Japan, http://kds.kek.jp/conferenceDisplay.py?confId=10808.
[28] Tevatron New Physics Higgs Working Group, CDF Collaboration, D0 Collaboration, Updated combination of CDF and D0 searches for Standard Model Higgs boson production with up to 10.0 fb^{-1} of data, arXiv:1207.0449 [hep-ex].
[29] LHC Higgs Cross Section Working Group, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.
[30] M.E. Peskin, T. Takeuchi, A new constraint on a strongly interacting Higgs sector, Phys. Rev. Lett. 65 (1990) 964.
[31] M.E. Peskin, T. Takeuchi, Estimation of oblique electroweak corrections, Phys. Rev. D 46 (1992) 381.
[32] R. Barbieri, A. Pomarol, R. Rattazzi, A. Strumia, Electroweak symmetry breaking after LEP-1 and LEP-2, Nucl. Phys. B 703 (2004) 127, arXiv:hep-ph/0405040.
[33] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Moenig, M. Schott, et al., The electroweak fit of the Standard Model after the discovery of a new boson at the LHC, Eur. Phys. J. C 72 (2012) 2205, arXiv:1209.2716 [hep-ph].
[34] LEPEWWG/TGC/2005-01, A combination of results in charged triple gauge boson couplings measured by the LEP experiments.