Status of the fourth fermion generation before ICHEP2012: Higgs data and electroweak precision observables

Otto Eberhardt, 1,* Alexander Lenz, 2,† Andreas Menzel, 3,‡ Ulrich Nierste, 1,§ and Martin Wiebusch 1,‖

1Institut für Theoretische Teilchenphysik, Karlsruhe Institute of Technology, D-76128 Karlsruhe, Germany
2CERN-Theory Division, PH-TH, Case C01600, CH-1211 Geneva 23, Switzerland
3Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, D-12489 Berlin, Germany

(Received 1 July 2012; published 9 October 2012)

We perform a global fit of the parameters of the Standard Model with a sequential fourth generation (SM4) to LHC and Tevatron Higgs data and electroweak precision data. Using several likelihood ratio tests, we compare the performance of the SM4 and the Standard Model (SM3) at describing the measured data. Since the SM3 and SM4 are not nested (i.e., the SM3 cannot be considered as a special case of the SM4 with some parameters fixed), the usual analytical formulae for p-values in likelihood ratio tests do not hold. We thus apply a new method to compute these p-values. For a Higgs mass of 126.5 GeV and fourth-generation quark masses above 600 GeV, we find that the SM4 is excluded at 3.1σ.

DOI: 10.1103/PhysRevD.86.074014 PACS numbers: 14.65.Jk, 12.15.Ff, 14.60.St, 14.80.Ec

I. METHOD AND INPUTS

In this paper, we study the Standard Model with a sequential fourth generation (SM4), which differs from the established Standard Model (denoted by SM3) by an additional fermion generation. We treat the masses of the extra fermions as free parameters and allow for arbitrary flavor mixings among the quarks of the four generations in our fits. Large mixings of the fourth-generation lepton doublet with those of the first three generations are ruled out [1] from data on lepton-flavor violating decays and lepton-flavor universality [2]. Recent NA62 data constrain these mixing angles even further [3]. Including lepton mixing within the allowed range has a negligible impact on the electroweak precision observables (EWPOs). In the absence of lepton mixing, the decay of the Higgs boson into neutrinos is invisible as long as the fourth-generation charged lepton is heavier than the corresponding neutrino. This invisible Higgs decay mode increases the total Higgs width and potentially counterbalances the effect of the enhanced gg → H production mechanism [4,5], because the branching fractions into the observed final states are reduced [6,7]. Allowing for (even small) mixing of the fourth with the other lepton doublets can render the neutrino decay mode visible. Since we want to quantify the level at which the SM4 is ruled out, we may confine ourselves to the most conservative scenario with an unmixed fourth-generation lepton doublet. Like the SM3, the SM4 can be studied with Dirac or Majorana neutrinos. In the fits presented in this paper, we use Dirac neutrinos. In our conclusions, we briefly discuss the (marginal) changes in the results expected for Majorana neutrinos.

From a model-building point of view, the hierarchy between three almost massless neutrinos and a fourth neutrino with mass of order of the electroweak scale can be motivated by a symmetry enforcing massless neutrinos in the exact symmetry limit: e.g., three right-handed neutrino fields might carry some U(1) charge while the fourth neutrino field and the left-handed lepton doublets are uncharged under this new symmetry. The Yukawa couplings are small spurions breaking this symmetry, leading to three tiny neutrino masses and tiny mixings between the fourth and the other generations.

A sequential fourth generation of fermions decouples neither from the production cross section σ(gg → H) nor from the Higgs decay rate into photons. Consequently, current LHC Higgs data put the SM4 under serious pressure [7–10]. In a recent publication [7], we presented a global fit of the SM4 parameters to EWPOs and Higgs signal strengths measured at Tevatron and the LHC. The signal strength is defined as

\[ \mu(X → H → Y) = \frac{\sigma(X → H)B(H → Y)}{\sigma(X → H)B(H → Y)_{SM3}} \]  

Here, we update our results with all available data and analyze the status of the SM4 prior to the ICHEP2012 conference. We also compute the statistical significance (p-value) at which the SM4 is excluded. As explained in Ref. [7], the computation of the p-value is nontrivial: due to the nondecoupling nature of the fourth-generation fermions, the SM3 cannot be regarded as a special case of the SM4, i.e., the two models are not nested. Analytical formulae for p-values only hold for nested models, and thus the p-value of the SM4 has to be computed numerically. To this end, a new C++ framework for maximum likelihood fits and likelihood ratio tests called my Fitter [11] was written. The implementation is discussed in Ref. [12].

In total, the following aspects of our previous analysis have been improved:
(1) The masses of all four fourth-generation fermions are now consistently treated as free parameters. To avoid nonperturbative Yukawa couplings and constraints from direct searches of fourth-generation quarks, we require $600 \text{ GeV} \leq m_{F^c}, \ m_t \leq 800 \text{ GeV}$. We are aware that for fermion masses of 800 GeV, the validity of perturbation theory is questionable at best. However, reducing the upper limit for the fermion masses can only lead to larger $\chi^2$ values in the SM4 and thus to smaller $p$-values. In this sense, the upper limit of 800 GeV is a conservative estimate.

(2) The signal strength for $pp \to H \to \tau\tau$ measured at the LHC [13] is included in the analysis.

(3) In the global fit, the Higgs mass is no longer fixed at 125 GeV, but is allowed to float in the range where experimental data on the Higgs signal strengths is available, i.e., $115 \text{ GeV} \leq m_H \leq 150 \text{ GeV}$ [14].

(4) Since, for a variable Higgs mass, no separate $H \to \gamma\gamma$ signal strengths for the gluon fusion and vector boson fusion production modes are available, we only use the combined signal strength for $pp \to H \to \gamma\gamma$ as input.

(5) For the two cases $m_H = 126.5 \text{ GeV}$ (the preferred Higgs mass of the SM3) and $m_H = 147 \text{ GeV}$ (the preferred Higgs mass of the SM4), we perform likelihood ratio tests to compare the performance of the SM3 and SM4 at describing the measured data.

Regarding the last point, a few more comments are in order. In likelihood ratio tests, the difference $\Delta \chi^2$ of minimal $\chi^2$ values obtained in the SM3 and the SM4 is used as a test statistic. One then assumes that the measured observables are random variables distributed around the prediction of one model (e.g., the SM4) with a spread determined by their errors and computes the probability ($p$-value) that a random set of “toy observables” leads to a $\Delta \chi^2$ which is more extreme (e.g., more SM3-like) than the $\Delta \chi^2$ value obtained from the real data. Note that this is different from the goodness-of-fit analysis presented in Ref. [9], which used the $\chi^2$ value of the SM4 as a test statistic and therefore did not compare the performance of the SM3 and the SM4. Also, the $H \to \tau\tau$ signal strengths were not included in their analysis.

Unfortunately, the likelihood ratio tests cannot be done (by us) if the Higgs mass is treated as a free parameter. In that case, the signal strengths measured in each invariant mass bin of each Higgs decay mode would have to be treated as separate observables, and we do not have any information on statistical correlations between adjacent bins. Thus, we only perform likelihood ratio tests for specializations of the SM3 and SM4, where the Higgs mass is fixed to $m_H = 126.5 \text{ GeV}$ (the value preferred by the global SM3 fit) or $m_H = 147 \text{ GeV}$ (the value preferred by the global SM4 fit). Then, only the signal strengths at $m_H = 126.5 \text{ GeV}$ and $m_H = 147 \text{ GeV}$ have to be treated as independent observables and correlations between these observables can safely be neglected.

Note, however, that the information from all invariant mass bins is encoded in our $\chi^2$ function. So, for example, the $\chi^2$ value at $m_H = 147 \text{ GeV}$ has a contribution due to the fact that there is a signal at $m_H = 126.5 \text{ GeV}$. If the model under consideration had a Higgs boson outside the discovery reach of LHC (or no Higgs boson at all), the theory prediction for all signal strengths in all invariant mass bins would be zero. This leads to a constant contribution to the $\chi^2$, which we are allowed to drop. Now, assume that the model has a Higgs boson with some mass $m_H$ and a predicted signal strength $\mu_{\text{ex}}(m_H)$. Let $\Delta \mu(m_H)$ and $\Delta \hat{\mu}(m_H)$ be the measured signal strength and experimental error for the corresponding invariant mass bin. After dropping the constant, the $\chi^2$ function is

$$\chi^2(m_H) = \frac{[\hat{\mu}_{\text{ex}}(m_H) - \mu_{\text{ex}}(m_H)]^2 - [\Delta \mu_{\text{ex}}(m_H)]^2}{[\Delta \hat{\mu}(m_H)]^2}.$$  (2)

If there is a clear signal at the Higgs mass $m_H$, the second term gives a large negative contribution to the $\chi^2$ function. This contribution is not present if $m_H$ is in a region without a signal, so the minimum of the $\chi^2$ function will usually be at a Higgs mass close to the signal.

In the present analysis, the following experimental inputs are used:

(i) $\hat{\mu}(pp \to H \to WW')$ measured by ATLAS [16],
(ii) $\hat{\mu}(pp \to H \to \gamma\gamma)$ measured by ATLAS [17],
(iii) $\hat{\mu}(pp \to HV \to Vbb)$ measured by CDF and D0 [18],
(iv) $\hat{\mu}(pp \to H \to ZZ')$ and $\hat{\mu}(pp \to H \to \tau\tau')$ measured by ATLAS [13],
(v) the EWPOs $M_Z, \Gamma_Z, \sigma_{\text{had}}, A_{\ell\ell}^U, A_{\ell\ell}^V, A_{\ell\ell}^B, A_{\ell\ell}^H, A_{\ell\ell}, A_{\ell\ell}, R_{\ell} = \Gamma_{\ell\ell}/\Gamma_{\text{had}}, \ R_A, \ R_B, \ \sin^2\theta_W^{\text{eff}}$ measured at LEP and Stanford Linear Collider [19] as well as $m_t, M_W, \Gamma_W$ and $\Delta m_{\text{had}}^{(5)}$ [20].
(vi) the lower bounds $m_{VP} \geq 600 \text{ GeV}$ (from the LHC) [21–24] and $m_{t_q} > 101 \text{ GeV}$ (from LEP2) [20].

Unfortunately, there is no data for signal strengths as a function of the Higgs mass from CMS.

On the theory side, the global fits with a variable Higgs mass were done with the CKMfitter software [25]. The EWPOs in the SM4 were calculated with the method described in Ref. [26], using FeynArts, FormCalc and LoopTools [27–29] to compute the SM4 corrections to the EWPOs. The EWPOs in the SM3 were calculated with the ZFitter software [30–32]. The Higgs width and branching ratios in the SM4 and SM3 were calculated with HDECAY v. 4.45 [33], which implements results of Refs. [34–37]. The SM3 Higgs production cross sections were taken from Ref. [38] (LHC) and Refs. [39,40] (Tevatron). For the numerical integration required to compute the $p$-values, we use the Degas code [41] which was developed in the context of Refs. [42,43].
II. RESULTS

To show the impact of the $H \to \tau \tau$ signal strength, we plot the minimal $\chi^2$ value with and without the $H \to \tau \tau$ input as a function of the mass $m_{\nu_4}$ of the fourth-generation neutrino in Fig. 1. We see that for $m_{\nu_4} \lesssim 60$ GeV, the minimum $\chi^2$ values are almost the same with and without the $H \to \tau \tau$ input. For $m_{\nu_4} \gtrsim 65$ GeV, the $H \to \tau \tau$ input increases the minimum $\chi^2$ by more than 20. We also see that without the $H \to \tau \tau$ input, the SM4 favors large values of $m_{\nu_4}$. With the $H \to \tau \tau$ signal strengths included, the smallest $\chi^2$ values are obtained for $m_{\nu_4}$ between 50 and 60 GeV.

This can be understood as follows: the production rate of Higgs bosons in gluon fusion is enhanced by a factor of 9 in the SM4 due to the contributions from additional heavy quark loops. On the other hand, the effective $HWW$, $HZZ$ and $H\gamma\gamma$ couplings are suppressed by the higher order corrections discussed in Ref. [37]. No such suppression is possible for $H \to \tau \tau$, so we would expect a $H \to \tau \tau$ signal strength of 9. The only way to reduce this signal strength is to open the invisible $H \to \nu_4 \bar{\nu}_4$ decay mode, which then suppresses all branching ratios by a common factor. Thus, for large values of $m_{\nu_4}$, the fit gets considerably worse if the $H \to \tau \tau$ channel is included.

Figures 2 and 3 show the minimum $\chi^2$ value as a function of the Higgs mass $m_H$ in the SM3 and SM4, respectively. The solid lines show the results of the combined analysis of signal strengths and EWPOs, while for the dashed lines only, the Higgs signal strengths (including $H \to \tau \tau$) were used as inputs. We see that the SM3 clearly prefers a Higgs mass near 126.5 GeV. This is in agreement with a similar analysis presented in Ref. [44]. There is another local minimum at $m_H = 145$ GeV, but with a considerably larger $\chi^2$ value. The $\chi^2$ function of the SM4 in the combined analysis of signal strengths and EWPOs also has one minimum at $m_H = 126.5$ GeV and another one at $m_H = 147$ GeV. Here, the $\chi^2$ values are almost the same, but still larger than the minimal $\chi^2$ value of the SM3, $\chi^2 = 23.3$, by about 8 units. Note that for non-nested models or models with bounded parameters, the relation between $\chi^2$ values and $p$-values is no longer given by Wilks’s theorem. Thus, in the case of the SM4, the number of degrees of freedom is an ill-defined concept, and the $p$-values have to be calculated by numerical simulation. For the simulations, we used the $my$ Fitter package [11]. Further details on the statistical issues and the $my$ Fitter simulation method can be found in Ref. [12]. For $m_H = 147$ GeV, the signals at invariant masses near 126.5 GeV would be interpreted as statistical fluctuations. Then, the data would be better described by the SM4 because it has more mechanisms for suppressing its Higgs signals. These mechanisms were discussed in Ref. [7].

Figure 4 shows the pulls of the Higgs signal strengths for the SM3 with a Higgs mass of 126.5 GeV and the SM4 with
a Higgs mass of 126.5 GeV or 147 GeV. We see that in the SM4 with $m_H = 126.5$ GeV, the measured $H \rightarrow \tau\tau$ signal strength deviates by more than 4σ from its predicted value. This is due to the effect mentioned in the discussion of Fig. 1. For the SM4 with $m_H = 147$ GeV, the measured signal strengths for the invariant mass bin at 147 GeV are in good agreement with their theory predictions. However, in that case, the $\chi^2$ receives a large contribution due to the fact that the measured values of the signal strengths in the invariant mass bin at 126.5 GeV deviate from their predicted values of zero.

Table I shows the $p$-values obtained from the likelihood ratio tests for the two SM4 Higgs masses. We see that, based on the Higgs signal strengths alone, the SM4 scenario with $m_H = 126.5$ GeV is ruled out at almost 4σ while the scenario with $m_H = 147$ GeV is only excluded at 3σ. At a fixed Higgs mass of 126.5 GeV, the electroweak fit is actually better in the SM4 than in the SM3. Thus, if the EWPOs are included in the fit, the $p$-value increases to 2 permille, which corresponds to 3.1σ. The lower bound $m_{t',b'} \geq 600$ GeV is not essential for this result; relaxing this bound to $m_{t',b'} \geq 400$ GeV decreases the minimum-$\chi^2$ by 0.6. For the SM4 scenario with $m_H = 147$ GeV, the $p$-value drops to 0.74 permille ($3\sigma$). In any case, the SM4 is excluded at more than 3σ.

**III. CONCLUSIONS**

We presented a combined analysis of Higgs signal strengths and EWPOs in the context of the Standard Model with three or four fermion generations. The SM3 is in good agreement with the experimental data and the best-fit Higgs mass is 126.5 GeV. The SM4, on the other hand, struggles to describe the Higgs signal strengths measured at Tevatron and the LHC. The $\chi^2$ function of the SM4 has two minima at $m_H = 126.5$ GeV and $m_H = 147$ GeV with essentially the same $\chi^2$ value, which is larger than the minimal $\chi^2$ value of the SM3 by 8 units. The second minimum of the SM4 $\chi^2$ function occurs because the SM4 cannot reproduce the signal strengths measured at 126.5 GeV very well, so that a SM4 with a Higgs mass nowhere near the observed signals describes the data equally well as a SM4 with $m_H = 126.5$ GeV. To quantitatively compare the performance of the SM3 and SM4 at describing the data, we performed likelihood ratio tests for fixed Higgs masses $m_H$ of 126.5 GeV in the SM3 and $m_H = 126.5$ GeV, 147 GeV in the SM4. The $p$-values were computed with a new numerical method [12] for likelihood ratio tests of non-nested models. If EWPOs and signal strengths are included in the fit, we find $p$-values of $2.0 \times 10^{-3}$ and $0.74 \times 10^{-3}$, respectively, which means that the SM4 is excluded at the 3σ level. While this result is obtained for Dirac neutrinos, it will change only marginally for the case of Majorana neutrinos with two fourth-generation mass eigenstates $\nu_A$, $\nu_S$: the fit to the signal strengths will return the same invisible Higgs width, now corresponding to the sum of the four decay rates $\Gamma(H \rightarrow \nu_A \nu_S, \nu_A \nu_A)$. A marginal difference occurs once the EWPOs are included: choosing the $\nu_A, \nu_S$ mass splitting such that the eigenstate with the larger SU(2) doublet component becomes heavier, one can slightly improve the quality of the electroweak fit. The improvement is negligible, as indicated by the shallowness of the minimum of the SM4 $\chi^2$ function in Fig. 1. While the SM4 is under severe pressure, a sequential fourth generation may still be viable in conjunction with an extended Higgs sector [45–48].

**ACKNOWLEDGMENTS**

We thank Geoffrey Herbert and Heiko Lacker for important contributions to the SM4 part of the CKMfitter code. We also thank Jérôme Charles for CKMfitter software support and very useful inputs regarding likelihood ratio tests for non-nested models. We acknowledge support by the DFG through grants NI1105/2-1, LA2541/1-1, LE1246/9-1, and Le1246/10-1.
[14] A lattice study has found the lower bound $m_H \geq 500\text{GeV}$ for $m_{\ell} = m_Z = 700\text{GeV}$ [15]. We interpret this result such that the perturbative vacuum state is metastable for $m_H \approx 125\text{GeV}$ and the heavy quark masses used by us. Therefore, Ref. [15] per se does not invalidate our analysis.
[38] S. Dittmaier et al. (LHC Higgs Cross Section Working Group), arXiv:1101.0593.