

Theoretical prediction for the muonium hyperfine-structure interval and its accuracySavely G. Karshenboim ^{*}*Ludwig-Maximilians-Universität, Fakultät für Physik, 80799 Munich, Germany;**Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany;**and Pulkovo Observatory, St. Petersburg, 196140, Russia*

Evgeny Yu. Korzinin

D. I. Mendeleev Institute for Metrology, St. Petersburg, 190005, Russia

(Received 19 May 2020; revised 23 December 2020; accepted 19 January 2021; published 8 February 2021)

Because there are new projects on the determination of the hyperfine interval in the ground state of muonium, we revisit different approaches to constructing its theoretical prediction and comparison with experimental data. We discuss a controversy in the accuracy of two recent predictions [Mohr *et al.*, *Rev. Mod. Phys.* **88**, 035009 (2016), and Eides, *Phys. Lett. B* **795**, 113 (2019)] and produce a constraint on possible systematic errors in the determination of the muonium hyperfine-structure interval theoretically or experimentally. In particular, a possible additional theoretical term, that may be either a missing “regular” theoretical contribution or a contribution due to new physics, is estimated as $-0.14(52)$ kHz. The constraint is based on all available data through a least-squares-adjustment procedure, including all their correlations. The result is close to the one of Eides [*Phys. Lett. B* **795**, 113 (2019)].

DOI: [10.1103/PhysRevA.103.022805](https://doi.org/10.1103/PhysRevA.103.022805)**I. INTRODUCTION**

Theoretical predictions play important roles in physics, serving different purposes. Sometimes it is important to predict a possible value of a certain quantity prior its measurement which may simplify a related experiment, e.g., by reducing a possible region of scanning. Sometimes it serves as an overall consistency check of the data and theory. It is also important to compare theory and experiment in order to look for a possible systematic error in both of them or to constrain a possible new physics term.

Depending on the purpose, one has to perform different procedures. The central value and the accuracy of a prediction depends on its purpose. For example, if for a new experiment one needs to estimate a value of the muonium hyperfine-structure (HFS) interval as accurately as possible, the best estimate is its previous experimental value. No theoretical calculation can produce a better prediction.

For muonium, the most interesting purposes of a comparison of theory and experiment include a constraint on a parameter space available for new physics and to check the overall consistency of the related theory and data.

Two theoretical predictions of the $1s$ muonium HFS interval were published recently:

$$\Delta\nu_{\text{th}} = 4463\,302.868(271) \text{ kHz } [6.1 \times 10^{-8}] [1], \quad (1)$$

$$\Delta\nu_{\text{th}} = 4463\,302.872(516) \text{ kHz } [1.2 \times 10^{-7}] [2], \quad (2)$$

which have quite different uncertainties (by approximately a factor of 2). In addition to the uncertainty given in parentheses, we also present a fractional uncertainty in square brackets. The prediction (1) (from Ref. [1]) was recently criticized in Ref. [2]. The prediction in Ref. [2] is perfectly consistent with predictions published in theoretical reviews (see Refs. [3,4] for recent ones; see also Ref. [5] for an older one) before the publication of Ref. [1].

Often a comparison of a theory and an experiment presents a “clean” case when the experimental input required for the theoretical expression comes from other fields. The specifics of the muonium HFS interval is that the input for the theory comes from the same experiment that the theoretical prediction is supposed to be compared with. That concerns both predictions mentioned above, which may create certain complications in their interpretation.

In this paper, we examine the legitimacy and accuracy of these two predictions. We demonstrate that both are legitimate but are most suitable for different purposes. We demonstrate that although they have uncertainties different by approximately a factor of 2, they constrain a possible new physics term at roughly the same level [as $-0.17(52)$ and $-0.13(52)$ kHz, respectively]. [To make the results comparable, we apply the same theoretical expression, which marginally shifts the central value of prediction (2).] Both

^{*}savely.karshenboim@mpq.mpg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

predictions also constrain a possible systematic error in experimental determination of the muonium HFS interval as 0.09(27) and 0.08(27) kHz, respectively. Notably, that level is somewhat different from the level of the theoretical constraint. We also consider other evaluations of the data for consistency tests and comparison of theory and experiment for muonium.

II. DATA AND RELATIONS

Prior a discussion on the numerical values of the predictions, let us review the structure of the related theoretical expression and the scope of the available data.

The leading term for the $1s$ HFS interval, the so-called Fermi energy, is defined here as

$$\nu_F = \frac{16}{3} R_\infty c \alpha^2 \frac{\mu_\mu}{\mu_p} \frac{\mu_p}{\mu_B} \left(1 + \frac{m_e}{m_\mu}\right)^{-3}, \quad (3)$$

where we use a standard notation for masses and magnetic moments of the involved particles, the fine structure constant, the Bohr magneton, the speed of light, and the Rydberg constant. The presence of the proton magnetic moment is dictated by the method of the calibration of the magnetic field at the intermediate stages. Following notation in Ref. [1], we consider μ_μ and μ_p as positive values and μ_e as a negative one. Elementary charge e and Bohr magneton μ_B are positive. The choice of the signs for the magnetic moments in, e.g., original (experimental) papers [6,7] is all positive, which is more consistent with a relativistic consideration. However, publication [1] presents a comprehensive summary on theory and experiment, which we rely on, while the rival publication [2] has also chosen the same convention about the signs, so we follow it here.

There are different definitions applied in the literature. The Fermi energy is a splitting related to a nonrelativistic energy of the interaction of the magnetic moments of an electron and a nucleus. One may use Dirac's value of the electron magnetic moment or a complete one. The nucleus in muonium, a bound system of an electron and a (positively charged) muon, is a lepton and therefore its anomalous magnetic moment is also small $a_\mu \approx a_e \approx \alpha/2\pi$. One may define the Fermi energy with the Dirac's value of its magnetic moment $e\hbar/2m_\mu$ or a complete one,

$$\mu_\mu = (1 + a_\mu) \frac{e\hbar}{2m_\mu}.$$

The definition of the Fermi energy in Ref. [1] uses the Dirac's value for the magnetic moments of both particles, while the effects due to the anomalous magnetic moments of the constituents are to be considered as corrections to the leading term. Here, following Ref. [2], we use Dirac's value of the electron magnetic moment and a complete one for the muon magnetic moment. That makes the quantum electrodynamics (QED) part of the expression applicable for the hydrogen atom. We present the Fermi energy in (3) as a combination of μ_μ/μ_p , which may be directly extracted from experiment (see below), and several fundamental constants, which are known more accurately than any determinations of μ_μ/μ_p . For all the values in (3), except μ_μ/μ_p , we apply values of the constants following Ref. [1]. Their uncertainty is very small and has been neglected by us.

We consider various procedures which may lead to very close results. We are to give various numerical values with a somewhat superficial accuracy in order to show how they are affected by modifications of the procedure. Neglecting uncertainties of various fundamental constants, involved in ν_F in (3), in particular we neglect the uncertainty in the reduced mass factor there and the uncertainty in g_p and a_μ , required to transfer μ_μ/μ_p to m_e/m_μ (see below). Those effects could affect the uncertainty of values considered here at the level of a percent of their uncertainties. Such an accuracy in the uncertainty is of no interest by itself, but it may be interesting to see how the uncertainty changes once we alternate our approach.

All the values of the fundamental constants in ν_F are known more accurately than μ_μ/μ_p and m_e/m_μ , which are directly related through the identity

$$\frac{\mu_\mu}{\mu_p} = \frac{2(1 + a_\mu) m_p m_e}{g_p m_e m_\mu},$$

with all the factors but μ_μ/μ_p and m_e/m_μ , known with a negligible uncertainty. Either of those two mentioned constants, μ_μ/μ_p or m_e/m_μ , could be chosen as a base constant to be found by an evaluation of the data, while the other is to be derived from it. The choice is a matter of taste and has no practical consequences since, at the current level of accuracy in determination of μ_μ/μ_p and m_e/m_μ , they are related through an (almost) exact relation.

The theoretical expression for the $1s$ Mu HFS interval may be presented as

$$\Delta\nu_{\text{Mu}} = \nu_F(1 + Q_{\text{QED}} + Q_{\text{hadr}} + Q_{\text{weak}}), \quad (4)$$

where Q_{QED} is the QED correction to the Fermi energy, Q_{hadr} stands for various hadronic contributions (through intermediate states), and Q_{weak} presents a small contribution due to the weak interactions.

As for the theoretical contributions, we follow here a consideration in Ref. [1]. In principle, we do not completely agree with it; however, the purpose of this paper is to consider different approaches to a calculation of a theoretical prediction, while various rather minor corrections will be considered elsewhere. The theoretical expression contains α , m_e/m_μ , a_e , a_μ , etc., in various small corrections, beside ν_F . For their calculations, we apply the values of constants following Ref. [1]. The uncertainty of the calculation of the corrections due to that is negligible. In particular, we neglect the uncertainty of m_e/m_μ in the reduced-mass factor $(1 + m_e/m_\mu)^{-3}$ in (3).

Theory by itself contains an uncertainty due to uncalculated terms. It is estimated in Ref. [1] as

$$\delta_{\text{Mu}} = \pm 85 \text{ Hz}. \quad (5)$$

All that leaves the theoretical expression with two free parameters, namely, μ_μ/μ_p (or m_e/m_μ) and δ_{Mu} , which is convenient to consider as a free parameter, for which we have a pseudodatum (5). We treat the pseudodatum as a result of an estimation on the same ground as we treat results of measurements (cf. Ref. [1]). Following Ref. [1], we use m_e/m_μ as a fit parameter.

The experimental data include results of two papers, which present two runs of an experiment at Los Alamos Meson Physics Facility (LAMPF) performed by an international collaboration. We refer to those runs as to experiment (exp.) 1 [6]

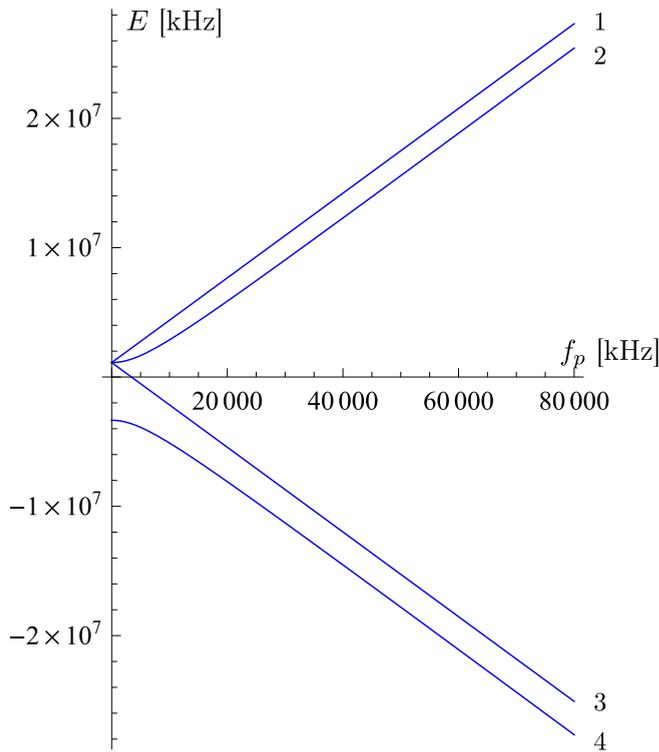


FIG. 1. The Breit-Rabi sublevels of the $1s$ hyperfine manifold in muonium in the presence of a magnetic field [8]. The magnetic field B is parameterized with the related free-proton spin-precession frequency $f_p = 2\mu_p B/h$, as done in Refs. [6,7]. The zero energy is defined as the center of gravity of the HFS multiplet (which does not depend on the value of the applied magnetic field). The measurements of small splitting ν_{12} and ν_{34} were performed at $f_p^{(1)} = 57\,972.993$ kHz [6] and $f_p^{(2)} = 72\,320.000$ kHz [7].

and exp. 2 [7]. The experiments measured transitions between the Breit-Rabi sublevels of the HFS manifold (see Ref. [8] for details as well as the experimental papers [6,7] and references therein) of the ground state in muonium in the presence of a magnetic field (see Fig. 1). Two small intervals, ν_{12} and ν_{34} , were measured in Refs. [6,7] by detecting the muon's spin flip. The values of two measured frequencies were correlated within each experiment.

The magnetic field was calibrated by measuring the proton spin precession in water. The absolute probe with water utilized in Ref. [7] had fractional uncertainty of 3.4×10^{-8} [9]. (In the earlier experiment [6], it was less accurate.)

To arrive at a free-proton frequency from a water one, one has to introduce corrections due to molecular shielding [10,11], temperature dependence [12], bulk diamagnetic shape correction [13], and other effects [9]. Some of those corrections were improved later on. In particular, the water shielding factor was measured in Refs. [10,11] in comparison to atomic hydrogen shielding factor, which was calculated. The theory of it in comparison with the one used in Refs. [9–11] was marginally improved (see, e.g., Ref. [14]) and the corrections have been used in the other parts of the evaluation in Ref. [1]. For example, the shielding factor for the magnetic moment of the proton in water at 25° C applied in the probe is $2.5692(14) \times 10^{-5}$, while it is $2.5687(15) \times$

10^{-5} according Ref. [15] and $2.5691(11) \times 10^{-5}$ following Ref. [1]. The variations in the uncertainty and central value are tiny. Such marginally important details should be either explicitly neglected (for a reason) at the data-selection stage, or included into the whole least-squares-adjustment (LSA) procedure in such a way that then the procedure should determine whether any particular correction is to be ignored (say, being smaller than the rounding errors). That is important because from measurement to measurement the situation may change, and the algorithm should be prepared for such a change.

The restored would-be free-proton spin-precession frequency f_p has a different uncertainty in exp. 1 and exp. 2. The one due to the calibration of the probe [9] (for Ref. [7]) is roughly two in the units in the last presented digit (i.e., approximately 2 Hz). However, since the dependence of $\nu_{12/34}$ on f_p is known (see below), one may transfer corrections and uncertainties in f_p to the values of measured $\nu_{12/34}$. All the involved parameters are known with an accuracy sufficient to calculate a shift and/or uncertainty of ν 's due to a correction and/or uncertainty of f_p . That allows ones to operate with exactly known free-proton spin-precession frequencies. (Each of two big measurements [6,7] consisted of a large number of runs. The effects of the magnetic field were considered on a run-by-run basis.)

The measured quantities, ν_{12} and ν_{34} (see Fig. 1), were evaluated in the original experimental papers [6,7] and their results are cited in Ref. [2]. Meanwhile, in publications of the CODATA task group on fundamental constants, starting with Ref. [15], the data have been rearranged as

$$\begin{aligned} \nu(f_p) &= \nu_{34} - \nu_{12}, \\ \Delta\nu_{\text{Mu}} &= \nu_{12} + \nu_{34}, \end{aligned} \quad (6)$$

and the values of $\nu(f_p)$ and $\Delta\nu_{\text{Mu}}$ have further been treated as (correlated) experimental data. Their correlation coefficients have been calculated in a proper way and such a rearrangement is an identity transformation (except of rounding errors).

The result on the rearranged data are given in Table I, following Ref. [1]. The input in the last line is the uncertainty of the theoretical expression which is presented as a pseudo-datum constraint by an estimation.

The data evaluation is based on an LSA procedure, i.e., on a multivariate least-squares evaluation of correlated data, which is described, e.g., in publications of the Particle Data Group (see, e.g., Ref. [16]) as well as in a CODATA task group publication [15]. The data are compared with a fit, i.e., with an expression that contains the fit parameters. Following [1], those parameters are m_e/m_μ and δ_{Mu} . The method is a standard generalization of the LSA procedure for correlated data, which we have to use once we intend to present a theoretical constraint as a single number, after accumulating information from five measured or estimated input, two pairs of which are correlated as seen from Table I.

Now we have to discuss observational equations that express measured quantities in terms of the fit parameters. They are summarized in Table II. We note that for this paper we consider all the fundamental constants that are known with the accuracy essentially better than the one of determination of m_e/m_μ as exactly known ones.

TABLE I. Experimental and theoretical input data for the evaluation of m_μ/m_e . Here, $r(a, b)$ is the correlation coefficient for quantities a and b .

Quantity	Value	Comment
$\nu(f_1) = \nu(f_p^{(1)})$	627 994.77(14) kHz	Exp. 1 at $f_p^{(1)} = f_1 = 57 972.993$ kHz, [6]
$\Delta\nu_{\text{Mu}}^{(1)}$	4463 302.88(16) kHz	$r(\Delta\nu_{\text{Mu}}^{(1)}, \nu(f_1)) = 0.227$
$\nu(f_2) = \nu(f_p^{(2)})$	668 223.166(57) kHz	Exp. 2 at $f_p^{(2)} = f_2 = 72 320.000$ kHz, [7]
$\Delta\nu_{\text{Mu}}^{(2)}$	4463 302.765(53) kHz	$r(\Delta\nu_{\text{Mu}}^{(2)}, \nu(f_2)) = 0.195$
$\delta_{\text{Mu}}^{(3)}$	0 ± 0.085 kHz	Theoretical estimation, following Ref. [1]

We also have to note that the observational equations are not standard mathematical identities but a presentation of details of a related LSA procedure, i.e., have a meaning similar to some lines of a code. Their left-hand part is presented by the observable quantities. It is expected that within the least-square procedure we put there the related experimental or theoretical results, i.e., the data. The right-hand part of the observational equations presents functions of the fit parameters. Together, the central values of the observable quantities, their uncertainty and correlations, and the fit functions are used to construct the value of χ^2 to be minimized.

We have already discussed $\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$, the theoretical expression for the muonium HFS interval [see (4)]. The other expression, mentioned in Table II, describes $\nu(f_p)$ in observational equation I. It is of the form (see, e.g., Ref. [8])

$$\nu(f_p) \doteq \left(\frac{g'_\mu(\text{Mu})}{g_\mu} \frac{\mu_\mu}{\mu_p} + s_e \right) f_p + \sqrt{(\Delta\nu_{\text{Mu}})^2 + f_p^2 \left(s_e - \frac{g'_\mu(\text{Mu})}{g_\mu} \frac{\mu_\mu}{\mu_p} \right)^2}, \quad (7)$$

where

$$s_e = \frac{g'_e(\text{Mu})}{g_e} \frac{\mu_e}{\mu_p}.$$

Here, $g'_e(\text{Mu})$, $g'_\mu(\text{Mu})$ are the g factors of the bound electron and muon in the $1s$ state of muonium. They are considered

TABLE II. Observational equations. The equations are not standard mathematical identities but a presentation of details of a related LSA procedure. Their left-hand part is presented by the observable quantities and suggests that within an LSA fitting we put there the related experimental or theoretical results, i.e., the data. We number the equations as I, II, and III and the data pieces as 1 [related to exp. 1 [6]], 2 (related to exp. 2 [7]), and 3 (an estimation of theoretical uncertainty in (5)). The right-hand part presents functions of the fit parameters. We use here the output values from Ref. [1] for all the involved fundamental constants except of m_e/m_μ , which acts as one of two fit parameters of our evaluation. The other fit parameter of our consideration is, following Ref. [1], a value of δ_{Mu} .

No.	Equation
II, I2	$\nu(f_p) \doteq \nu(f_p, \Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}))$
III, II2	$\Delta\nu_{\text{Mu}} \doteq \Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$
III3	$\delta_{\text{Mu}} \doteq \delta_{\text{Mu}}$

with sufficient details in Ref. [1]. As we mentioned μ_e and g_e (following [1]) are negative, which makes s_e also negative.

To apply (7) as an observational equation in Table II, we have to express it in terms of the fit parameters, i.e., to substitute $\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$ for $\Delta\nu_{\text{Mu}}$, which is an identity transformation as far as we consider also observational equation II.

The last observational equation (III) also needs a comment. We mentioned referring to (5) that δ_{Mu} is a theoretical uncertainty for the muonium-HFS-interval theory. That is not entirely correct. The quantity stands for a deviation of a truncated (with unknown terms ignored) theoretical expression with only central values of the coefficients for the related known terms from the actual (true) value of $\Delta\nu_{\text{Mu}}$. Such a truncated expression is incomplete and should deviate from the actual value of muonium HFS. If the theory is well understood, the deviation can be estimated theoretically and the input datum for the deviation is a theoretical estimation of the uncertainty of the truncated expression with central values, denoted as $\delta_{\text{Mu}}^{(3)}$. In Table I, we give such an input value, while the observational equation in Table II contains δ_{Mu} both in the left- and right-hand parts. The one on the left indicates that the quantity δ_{Mu} is measurable and a datum for it is available. The one on the right is a fit parameter and indicates that with performing a (standard) LSA we are to find an output value of δ_{Mu} as a deviation, which depending on relation between the accuracy of the overall data may be dominated by theoretical estimation (as it has happened) or by a certain comparison of the truncate theory and experiment [which could be if the experiment would be more accurate and produce an accurate value of m_e/m_μ through (7)].

If we would suggest that new physics is possible or that a certain theoretical term has been missing, the deviation of the truncated result for the HFS interval from its actual value should consist of two terms $\varepsilon + \delta_{\text{Mu}}$, where the second term presents an estimated uncertainty of the (standard) theory, while the first term is supposed to describe a missing theoretical piece to be constrained.

III. TWO THEORETICAL PREDICTIONS

We have already given the numerical values of the predictions in (1) and (2). Now we first describe them below and then analyze them.

As we see from the consideration above, the muonium case is a very specific one. A theoretical prediction speaks in the terms of the Fermi energy (3), a calculation of which requires a value of m_e/m_μ (or μ_μ/μ_p). To determine the latter,

we have to use $\nu(f_p)$ (measured in the same experiment as $\Delta\nu_{\text{Mu}}$) and a certain presentation of $\Delta\nu_{\text{Mu}}$ [cf. (7)], which is not an ordinary situation for a theoretical prediction. That means that while calculating a theoretical prediction it is unavoidable to use at least one datum [on $\nu(f_p)$] from the same experiment as determines $\Delta\nu_{\text{Mu}}$. That also means that one has to deal somehow with the value of $\Delta\nu_{\text{Mu}}$ on an early stage of calculating its theoretical prediction. The possibilities are either to use its experimental value (to which the theoretical prediction is to be eventually compared) (cf. Ref. [2]) or the theoretical expression (4) (cf. Ref. [1]), which would make the dependence of the theoretical prediction on the details of the theoretical expression quite complicated. The problem is solved differently in the discussed predictions.

The prediction (2) from Ref. [2] was obtained by directly using original evaluation in the experimental papers [6,7] to determine m_e/m_μ . The expression (7) converted to

$$\frac{\mu_\mu}{\mu_p} = \frac{(\Delta\nu_{\text{Mu}})^2 - [\nu(f_p)]^2 + 2s_e f_p \nu(f_p)}{4s_e f_p^2 - 2f_p \nu(f_p)} \frac{g_\mu}{g'_\mu(\text{Mu})} \quad (8)$$

allows one to determine m_e/m_μ (through μ_μ/μ_p) directly from the experimental data as has been done in original experimental papers [6,7]. To do that in each paper, they used the measured value of $\Delta\nu_{\text{Mu}}$ instead of its presentation in terms of the fit parameters.

The next step was a calculation of the value of the theoretical Q coefficients in (4). The theoretical state of the art could be found in both cited papers [1,2]. Theory is presented there somewhat differently but the difference in the central value and the uncertainty is not of big importance. There is also a number of small additional corrections to the original experimental results which should be introduced for a proper procedure, but all of them are marginal.

A list of marginal effects, which should be revisited while applying the results from experimental papers [6,7], includes an update of the corrections to the shielding in muonium, i.e., the value of $g'(\text{Mu})/g$ [17–20], an update on a_μ [21], and an update on various values of involved fundamental constants, the most important of which for HFS in muonium are updates on g_p [14] (see also a calculation of the shielding factor in hydrogen [17,22] to correct the related experiment [23]) and m_e/m_p [24,25].

In addition to the updates, one has to mention that the result of each paper applied a value of $\Delta\nu_{\text{Mu}}$ measured within the same experiment. The physical value of $\Delta\nu_{\text{Mu}}$ does not depend on an experiment and a proper procedure should somehow combine the results of both experiments for $\Delta\nu_{\text{Mu}}$. However, the accuracy of evaluation of m_e/m_μ from each experiment is limited by the value of $\nu(f_p)$, essentially less accurate than $\Delta\nu_{\text{Mu}}$. By combining two results on $\Delta\nu_{\text{Mu}}$, we improve the accuracy of a quantity which is already known with a superficial accuracy and therefore the changes would be minimal.

A more rigorous procedure (following the same logics) would be first to dismiss the pseudodatum $\delta_{\text{Mu}}^{(3)}$ and find m_e/m_μ . Technically that means the use of the experimental results on $\Delta\nu_{\text{Mu}}$ in observational equation I. With that value of m_e/m_μ in hand and now with the estimation from (5), the next stage is to obtain a value of a theoretical prediction. Such

a value is marginally different from (2). Here, referring to results on base of the prediction from Ref. [2], we mean that rigorous procedure.

The rival prediction [see (1)] [1] was obtained in a different way. An evaluation was performed with the experimental data on $\Delta\nu_{\text{Mu}}$ excluded. Technically, that means that to extract a value of m_e/m_μ from $\nu(f_p)$ one has to apply a theoretical expression for $\Delta\nu_{\text{Mu}}$. Afterward, with a value of m_e/m_μ (and all necessary correlation coefficients) one may find theoretical prediction (1). The value of m_e/m_μ and a theoretical value on $\Delta\nu_{\text{Mu}}$ calculated with this mass ratio have been obtained without any use of the experimental data on $\Delta\nu_{\text{Mu}}$ and therefore the outcome may be compared with those experimental values, which makes the comparison legitimate.

The uncertainty of two predictions is different by approximately a factor of 2. As we explain below, both are in a certain sense legitimate and both deal with certain elements of circling. The latter, if properly understood, does not make the predictions incorrect or useless, but makes their use not a straightforward one.

One of the most important applications of a theoretical prediction is a comparison of theory and experiment in order to constrain a new physics or to look for a doubtful theoretical term of a regular theory within standard physics. If we add an additional term to a theoretical expression, the value of a theoretical prediction will be shifted. In the case of the prediction (2) [2], the shift will be by the same value as we add to the expression. In the case of (1) [1], a value of the shift of the theoretical prediction is a subject of a certain calculation because there are two contributions to the sensitivity coefficient. One is a correction to the theoretical expression by itself, which is shifted by the new possible term, while the other is a shift in m_e/m_μ , since the theoretical expression (with a possible new term) is to be applied to find m_e/m_μ . If such a sensitivity is not given, those who use prediction (1) could be confused and produce an incorrect constraint.

Another issue one has to remember that either theoretical prediction is correlated with experimental values of $\Delta\nu_{\text{Mu}}$. The result [2] is correlated because it is based on the use of experimental values of $\Delta\nu_{\text{Mu}}$ to which the comparison is to be made. The other issue is that the values of $\nu(f_p)$ are correlated with values for $\Delta\nu_{\text{Mu}}$ from the same experiment anyway (see Table I). It would be more useful together with a theoretical prediction for the muonium HFS interval to present its correlation coefficient with the experimental value(s) of $\Delta\nu_{\text{Mu}}$ in order to correctly calculate the uncertainty of the comparison of theory and experiment.

It is useful to neglect a less accurate experiment [6] and derive the related predictions from a single more accurate measurement [7]. Both predictions are based in this case on explicit equations which follow from (8). For the theoretical prediction ν_{th} , which is of the same functional form [see (4)], they use different values of m_e/m_μ , namely

$$\frac{m_e}{m_\mu} = \mathcal{F}(\nu(f_2), \Delta\nu_{\text{Mu}}^{(2)}), \quad (9)$$

for a prediction similar to the one from Ref. [2] and

$$\frac{m_e}{m_\mu} = \mathcal{F}(\nu(f_2), \Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}^{(3)})) \quad (10)$$

TABLE III. Results of different evaluations following the logic of Ref. [2]. The results of exp. 1 and exp. 2 on m_e/m_μ are obtained by using (8) with the experimental values of $\Delta\nu_{\text{Mu}}$ as the input. That is, m_e/m_μ is an output datum of certain evaluations, which was subsequently applied to obtain a value of the muonium HFS interval. The theoretical estimation $\delta_{\text{Mu}}^{(3)}$ was omitted while extracting m_e/m_μ , but included while subsequently calculating $\Delta\nu_{\text{th}}$. LSA' is an LSA with pseudodatum on δ_{Mu} dismissed. Compared to the original results, the evaluation of the data from exp. 1 [6] and exp. 2 [7] is slightly updated, following Ref. [1].

Evaluation	m_e/m_μ	$\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}^{(3)})$ [kHz]	Comment
Exp. 1	206.768 220(77) [3.7×10^{-7}]	4463 304.1(17) $\times 10^9$ [3.7×10^{-7}]	
Exp. 2	206.768 283(25) [1.2×10^{-7}]	4463 302.78(56) $\times 10^9$ [1.2×10^{-7}]	
Average	206.768 276(24) [1.2×10^{-7}]	4463 302.91(52) $\times 10^9$ [1.2×10^{-7}]	
LSA'	206.768 276(24) [1.2×10^{-7}]	4463 302.91(52) $\times 10^9$ [1.2×10^{-7}]	LSA w/o $\delta_{\text{Mu}}^{(3)}$

for the one similar to the prediction from Ref. [1]. Here m_e/m_μ in the right-hand part of the second equation is the same as in the left-hand part, which makes this determination legitimate but indirect.

Function \mathcal{F} , which follows from (8), is the same for the both determinations, but its arguments are different. Both identities perfectly determine the value of m_e/m_μ , but the values which result from those two equations have different central values, uncertainties, correlations with experimental value $\nu_{\text{Mu}}^{(2)}$, and (crucially for a comparison of theory and experiment) sensitivities to modification of the theory. Consequently, similar differences appear when one calculates $\Delta\nu_{\text{th}}$ using the related values of m_e/m_μ .

Below we consider two predictions in more detail and suggest an appropriate procedure. The purpose of the following sections is to find the sensitivity coefficients (for the additional terms) and to check the strength of these constraints which rely on nontrivial sensitivity coefficients. It is clear that a comparison of the theoretical predictions above and the experimental data relies on the same set of theoretical and experimental information but uses different compositions of it. We also discuss which is the best procedure to combine all the available information and produce the strongest constraint on additional terms and which is the best way to compare the theory and the experiment.

IV. ACCURACY OF DIFFERENT EVALUATIONS AS A FUNCTION OF THEORETICAL UNCERTAINTY

Let us consider now the predictions more accurately. We start with the prediction in (2) [2]. It utilizes a value of m_e/m_μ which may be obtained from an LSA procedure with theoretical data piece in (5) removed from a consideration. A subsequent calculation of $\Delta\nu_{\text{th}}$ with such a value of m_e/m_μ contains the estimation of (5). Since only one source of the uncertainty dominates, namely the uncertainty of $\nu(f_p)$, one may perform a simplified procedure which has outcome marginally different from the mentioned LSA. One may evaluate both experiments separately without any use of theory and derive related values of m_e/m_μ (see Table III) and $\Delta\nu_{\text{Mu}}$. To not use the theory for the extraction of m_e/m_μ means that one has to apply relation (7) while using an experimental value of $\Delta\nu_{\text{Mu}}$ in it. The weighted average (over two experiments) of m_e/m_μ allows one to reproduce prediction (2), while the average of $\Delta\nu_{\text{Mu}}$ gives the experimental value to compare to.

An element missing within such a consideration is a correlation between the values of the prediction and of the experimental results on $\Delta\nu_{\text{Mu}}$, described by correlation coefficient $r(\Delta\nu_{\text{Mu}}, \Delta\nu_{\text{th}})$ for $\Delta\nu_{\text{th}}$ from Ref. [2] that requires comparing the theory and the experiment. Within a marginal correction, it coincides with

$$r(\Delta\nu_{\text{Mu}}, m_e/m_\mu) = 0.109. \quad (11)$$

The correlator in (11) basically follows the value related to the most accurate experiment [7]. If we include only its data, the result is $r = 0.110$ instead of $r = 0.109$. In the case of a single experiment, we may use explicit formulas for m_e/m_μ [cf. (8)] and find contributions to the correlator due to the presence of $\Delta\nu_{\text{Mu}}$ in (8) and due to a (known) correlation between $\Delta\nu_{\text{Mu}}$ and $\nu(f_p)$ within the same experiment. The result is $r = -0.088 + 0.198 = 0.110$, which demonstrates that both sources of the correlation are important. The correlation coefficient is required to find the uncertainty of the eventual comparison of the theoretical prediction and the experimental result. Since the predicted value of $\Delta\nu_{\text{Mu}}$ is essentially less accurate than the experimental one, the correlation plays a marginal role.

If the predicted and experimental results are consistent (and they are), then any constraint for new physics or any additional theoretical term of any other nature is to be consistent with zero. If the results would not agree, the nonzero correlation is not as important as the fact that both numerical values in (11) are extracted from the same experiment(s) and any discrepancy would equally compromise both of them. The interpretation depends on how a possible systematic error, theoretical or experimental, is involved. We will return to this question later on, but first we consider the other prediction [1], which *seems* more accurate.

Prediction (1) is derived in Ref. [1] by using an LSA procedure from which the data pieces, related to a direct experimental determination $\Delta\nu_{\text{Mu}}$, are removed. We reproduce it below by using two procedures with a rearrangement of the data and a subsequent calculation of averages, which allows for a direct transparent interpretation. Those procedures are helpful to explain various properties of prediction (1). (The average of a few individual values for the same quantity in case of their correlations is understood here as their linear combination which has the minimal possible uncertainty. Such a value can be in particular found through a certain least-square procedure.)

TABLE IV. Results of different evaluations following logics of [1]. The results marked with Exp. 1 and 2 are obtained by using (8) with a theoretical expression of $\Delta\nu_{\text{th}}$ for $\Delta\nu_{\text{Mu}}$. LSA'' is an LSA with the experimental data on $\Delta\nu_{\text{Mu}}$ dismissed. With those data dismissed, δ_{Mu} , being a fit parameter, has its output value equal to $\delta_{\text{Mu}}^{(3)}$. Evaluation of the data from Exp. 1 [6] and Exp. 2 [7] is consistent with the procedure described in Ref. [1].

Evaluation	m_e/m_μ	$\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}^{(3)})$ [kHz]	Comment
Exp. 1	206.768254(33) [1.6×10^{-7}]	4463 303.4(7) [1.6×10^{-7}]	
Exp. 2	206.768283(14) [6.6×10^{-8}]	4463 302.77(30) [6.6×10^{-8}]	
Average	206.768279(13) [6.1×10^{-8}]	4463 302.863(272) [6.1×10^{-8}]	
LSA''	206.768279(13) [6.0×10^{-8}]	4463 302.864(271) [6.1×10^{-8}]	LSA w/o $\Delta\nu_{\text{Mu}}^{(1)}, \Delta\nu_{\text{Mu}}^{(2)}$

A. Procedure 1

Similarly to the reproduction of (2) above, let us consider two experiments separately. We again rely on (8). Following Ref. [1], we decline to use the experimental data on $\Delta\nu_{\text{Mu}}$ at this stage. Therefore, the only way to make relation (8) useful (after dismissing the experimental data on $\Delta\nu_{\text{Mu}}$) is to apply a theoretical description of $\Delta\nu_{\text{Mu}}$ in terms of the fit parameters, i.e., $\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$. The results of such an evaluation are given in Table IV.

The individual results in Table IV derived from two experiments separately are correlated through δ_{Mu} . That only marginally affects their average value. It is interesting to see how the fractional uncertainty of m_e/m_μ and correlation coefficient $r(m_e/m_\mu, \delta_{\text{Mu}}^{(3)})$ evolve while changing δ_{Mu} . (We denote the running input value as $\delta_{\text{Mu}}^{\text{inp}}$, while $\delta_{\text{Mu}}^{(3)}$ stands for its actual value.)

As one might anticipate, if we increase the (estimated) uncertainty of theory, the uncertainty of each experimental determination of m_e/m_μ increases and the correlation coefficient $r(m_e/m_\mu, \delta_{\text{Mu}})$ is not negligible (see Figs. 2 and 3 for detail). That definitely proves that the value of m_e/m_μ obtained in such a way is theory dependent, which means that prediction (1) [1] has a nontrivial dependence on details of the

theoretical expression and its sensitivity to a modification of theory should be studied additionally.

One can note from the plots that once the uncertainty magnification factor $\delta_{\text{Mu}}^{\text{inp}}/\delta_{\text{Mu}}^{(3)}$ leads to a theoretical uncertainty $\delta_{\text{Mu}}^{\text{inp}}$ which exceeds the level of uncertainty of m_e/m_μ extracted from the both experiments, the uncertainty of each separate extraction increases with $\delta_{\text{Mu}}^{\text{inp}}$, while their average value (with all the correlations taken into account) reaches a limit with a finite accuracy

$$\left(\frac{m_e}{m_\mu}\right)_{\delta_{\text{Mu}} \rightarrow \infty}^{\text{average}} = 206.768\,41(16) [7.7 \times 10^{-7}]. \quad (12)$$

The presence of such a limit means that the individual results of two experiments become strongly correlated. Apparently, the correlation comes from the fact that with a magnified theoretical uncertainty, in some moment it starts to dominate in the uncertainty of each individual values of m_e/m_μ . The correlation is a positive one. Since the individual results are positively correlated, an outcome for the average, more accurate than the individual results by themselves, is possible if the linear combination for the average has coefficients with opposite signs. Counterintuitive negative weights for a weighed average are possible for correlated data.

To explain such a behavior, we are to perform an equivalent procedure, which has different intermediate steps.

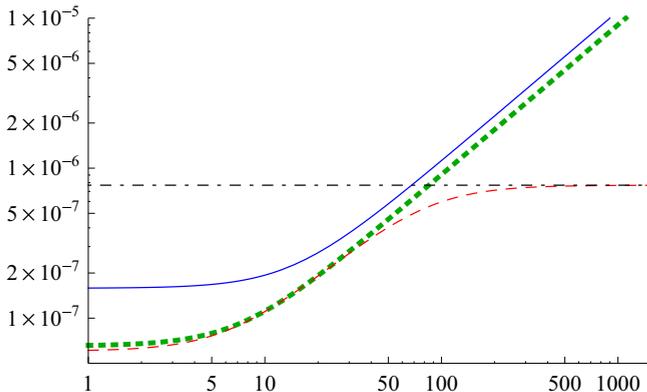


FIG. 2. Fractional uncertainty of determination of m_e/m_μ following procedure 1 for Exp. 1 (green dotted), Exp. 2 (blue solid), and their combined value (through LSA'') (with their correlation taken into account) as a function of the uncertainty magnification factor $\delta_{\text{Mu}}^{\text{inp}}/\delta_{\text{Mu}}^{(3)}$, where $\delta_{\text{Mu}}^{\text{inp}}$ a magnified uncertainty of the theoretical input estimation, while $\delta_{\text{Mu}}^{(3)}$ is its actual value. The black dot-dashed line relates to an asymptotic limit (12).

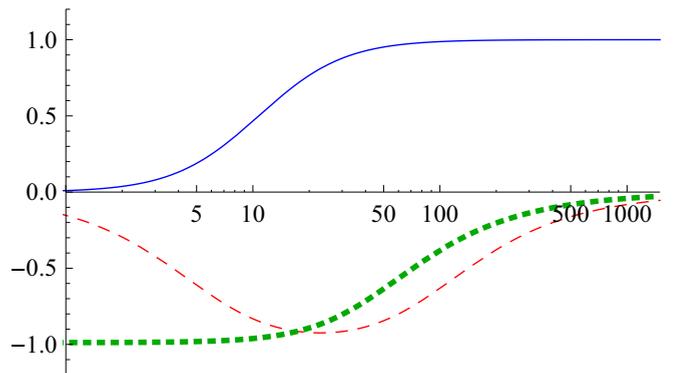


FIG. 3. Correlations coefficients as a function of the uncertainty magnification factor $\delta_{\text{Mu}}^{\text{inp}}/\delta_{\text{Mu}}^{(3)}$: $r[(m_e/m_\mu)^{\text{Exp.1}}, (m_e/m_\mu)^{\text{Exp.2}}]$ (blue solid), $r(m_e/m_\mu, \delta_{\text{Mu}})$ (red dashed), and $r[(m_e/m_\mu)_{\text{dep}}, (m_e/m_\mu)_{\text{ind}}]$ (green dotted), where m_e/m_μ without any superscripts is a combine value from two experiments, obtained within LSA''.

B. Procedure 2

The key data used for determination of m_e/m_μ are the results on $\nu(f_p)$ which can be evaluated by using (8). Above, we did the evaluation for each experiment separately and averaged the outcome. Apparently in the case of evaluation of an experiment in order to apply the identity (8) to determination to m_e/m_μ , we have to use a certain input for $\Delta\nu_{\text{Mu}}$, theoretical or experimental. However, once we evaluate the data of two measurements simultaneously, we have another option. Identity (8), applied to two different magnetic fields (at different f_p 's), allows ones without any other input (for $\Delta\nu_{\text{Mu}}$) to derive a certain value of μ_μ/μ_p (and consequently of m_e/m_μ),

$$\frac{\mu'_\mu}{\mu_p} = \frac{[-v^2(f_1) + 2s_e f_1 \nu(f_1)] - [-v^2(f_2) + 2s_e f_2 \nu(f_2)]}{[4s_e f_1^2 - 2f_1 \nu(f_1)] - [4s_e f_2^2 - 2f_2 \nu(f_2)]}, \quad (13)$$

and of $\Delta\nu_{\text{Mu}}$,

$$(\Delta\nu_{\text{Mu}})^2 = -\frac{\frac{-v^2(f_1)+2s_e f_1 \nu(f_1)}{4s_e f_1^2-2f_1 \nu(f_1)} - \frac{-v^2(f_2)+2s_e f_2 \nu(f_2)}{4s_e f_2^2-2f_2 \nu(f_2)}}{\frac{1}{4s_e f_1^2-2f_1 \nu(f_1)} - \frac{1}{4s_e f_2^2-2f_2 \nu(f_2)}}. \quad (14)$$

The former of these two equations, after use of numerical values of the shielding factor $g'_\mu(\text{Mu})/g_\mu$ and fundamental constants g_p and a_μ , gives a theory-independent value of the mass ratio which we denote here as $(m_e/m_\mu)_{\text{ind}}$. The latter of the identities produces a value of the Mu HFS interval, which together with a theoretical expression gives us a theory-dependent value of the mass ratio, which we denote here as $(m_e/m_\mu)_{\text{dep}}$. The values in the left-hand part of the identities are explicitly defined through the measurable quantities and we can easily find their correlation:

$$r\left(\frac{\mu'_\mu}{\mu_p}, \Delta\nu_{\text{Mu}}\right) = -0.988. \quad (15)$$

With the actual sources of the uncertainties [with the uncertainty due to $\nu(f_p)$ strongly dominating], one can also write (within the rounding error)

$$r[(m_e/m_\mu)_{\text{ind}}, (m_e/m_\mu)_{\text{dep}}] = r\left(\frac{\mu'_\mu}{\mu_p}, \Delta\nu_{\text{Mu}}\right), \quad (16)$$

where the values in the right-hand part are defined through (13) and (14), respectively.

The value of μ'_μ/μ_p in (13) depends neither on a theoretical expression nor on experimental data for $\Delta\nu_{\text{Mu}}$. One can use it to immediately find a theory-independent value of the mass ratio

$$\left(\frac{m_e}{m_\mu}\right)_{\text{ind}} = 206.768\,409(159). \quad (17)$$

In the meantime, the value of $\Delta\nu_{\text{Mu}}$ in (14) is a kind of an indirect experimental result on the muonium HFS interval

$$(\Delta\nu_{\text{Mu}})_{\text{ind}} = 4463\,305.8(3.6)\text{ kHz}. \quad (18)$$

Being combined with a known theoretical expression, $\Delta\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$, it allows one to find a theory-dependent value:

$$\left(\frac{m_e}{m_\mu}\right)_{\text{dep}} = 206.768\,143(165). \quad (19)$$

Those two indirect values of m_e/m_μ and $\Delta\nu_{\text{Mu}}$ are complementary to the results on $\Delta\nu_{\text{Mu}}$ in Table I and to the result on m_e/m_μ in Table III. They are independent of the results obtained in a standard evaluation in original experimental papers [6,7] and prediction paper [2] (and given in the mentioned tables) in a sense that cannot be expressed in their terms, but they are correlated with the results in the mentioned tables. While the results in experimental papers and in Ref. [2] present the results of each experiment in terms of $\Delta\nu_{\text{Mu}}$ and m_e/m_μ , the indirect values use a part of the results of each of two experiments [namely, their $\nu(f_p)$] and take advantage of the known shape of their dependence of f_p . Importance of the indirect results depends on the composition of the data (the uncertainties and their correlations, the overall consistency, and the values of the applied magnetic field, parameterized with f_p).

It is crucial that all the input is the same as for LSA with omitted data on $\Delta\nu_{\text{Mu}}$ in Ref. [1] and in procedure 1 above, while all the rearrangements are identities. The average of two results obtained here should coincide with prediction (1) [1], which is counterintuitive, because the intermediate predictions of procedure 2 are very inaccurate. Their fractional uncertainty is about 8×10^{-7} for both values [cf. (17) and 19]. However, the value of their correlation coefficient in (16) indicates that those two results are strongly anticorrelated and their combination should be essentially more accurate than both of them. Since the uncertainties are roughly the same, their half sum should be a good estimation of their average. The correlation coefficient $r[(m_e/m_\mu)_{\text{ind}}, (m_e/m_\mu)_{\text{dep}}]$ by approximately 1% deviates from minus unity [cf. (15)] and thus we expect that the uncertainty of our estimation with the half sum should be an order of magnitude smaller than the individual uncertainties of the considered values, which is in a reasonable agreement with our expectation. (The correlation coefficient enters an expression with squares of the uncertainties and therefore the uncertainty of the half sum squared should be two orders of magnitude below the squared individual uncertainties.)

We note that the very presence of a theory-independent value of m_e/m_μ in procedure 2 delivers an explanation of what happens if we increase the uncertainty magnification factor within procedure 1. Within the latter, both individual results (the results per experiment) become less accurate, being strongly correlated. Since two procedures are identical in terms of their input and output, both the theory-independent value and the theory-dependent one are certain linear combinations of the results per experiment. With the increase of the uncertainty magnification factor, the theory-dependent contribution $(m_e/m_\mu)_{\text{dep}}$ becomes very uncertain, while the theory-independent one $(m_e/m_\mu)_{\text{ind}}$ becomes dominant in the determination of the average, and, in particular, determines the limit in (12). In other words, the uncertainty of the average of the individual results cannot be less accurate than the uncertainty of the theory-independent value in (18).

Now, let us consider in detail how two inaccurate results of procedure 2 could produce an accurate average in the case of the current theoretical uncertainty $\delta_{\text{Mu}}^{(3)}$ as an input. It is helpful to consider the dependence of their fractional accuracy and their correlation coefficient as a function of the uncertainty magnification factor (see Figs. 4 and 3 for details).

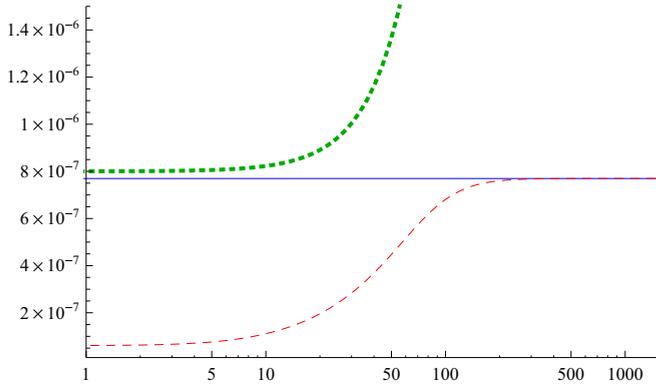


FIG. 4. Fractional uncertainties of various values of m_e/m_μ as function of the uncertainty magnification factor: for theory-dependent (green dotted), theory-independent (blue solid), and combined (red dashed) values.

While the magnification factor is relatively close to unity, the anticorrelation produces an accurate average of two inaccurate results, while with increase of the magnification factor the theory-independent value becomes dominant. The behavior of the correlator $r(m_e/m_\mu, \delta_{\text{Mu}})$ changes as well. The correlation with theory $r(m_e/m_\mu, \delta_{\text{Mu}})$ is important in the transition area but unimportant for large and low values of the magnification factor.

Procedure 2 allows one easily to find the sensitivity of outcome for m_e/m_μ on an additional theoretical term. We can adapt the individual m_e/m_μ results to a calculation of theoretical prediction of the HFS interval and its sensitivity to a theoretical addition. First, we note that in procedure 2 we have to combine two correlated values of m_e/m_μ . The correlator is known [see (15)]. In such a case, we have two alternative procedures with the same result. One of them is a mini-LSA when we have to find an average with two data points. The other is to present the outcome as a linear combination of two correlated input values with arbitrary coefficients and to choose them to minimize the uncertainty of the combination. The results should be the same. Minimizing the uncertainty, we find

$$\begin{aligned} \left(\frac{m_e}{m_\mu}\right)_{\text{aver}} &\simeq 0.510 \left(\frac{m_e}{m_\mu}\right)_{\text{ind}} + 0.490 \left(\frac{m_e}{m_\mu}\right)_{\text{dep}} \\ &= 206.768\,279(13) \quad [6.1 \times 10^{-8}]. \end{aligned} \quad (20)$$

The theoretical prediction linearly depends on m_e/m_μ in a sense that the numerical value of all the corrections beyond the linear term are small and indistinguishable to use of any m_e/m_μ which appears through the paper and, in particular, for $(m_e/m_\mu)_{\text{ind}}$ and $(m_e/m_\mu)_{\text{dep}}$. Applying the combination in (20), we find

$$\begin{aligned} \Delta v_{\text{th}}^{\text{aver}} &\simeq \Delta v_{\text{th}} \left[0.510 \left(\frac{m_e}{m_\mu}\right)_{\text{ind}} + 0.490 \left(\frac{m_e}{m_\mu}\right)_{\text{dep}} \right] \\ &= 0.51 \Delta v_{\text{th}} \left[\left(\frac{m_e}{m_\mu}\right)_{\text{ind}} \right] + 0.49 \Delta v_{\text{th}} \left[\left(\frac{m_e}{m_\mu}\right)_{\text{dep}} \right] \\ &\simeq 4463\,302.863(272) \text{ kHz} \quad [6.1 \times 10^{-8}]. \end{aligned} \quad (21)$$

TABLE V. The output data of a simplified version of LSA from Ref. [1]. Here, $\nu = 3$ is the number of degrees of freedom of the data evaluation.

Quantity	Value
m_μ/m_e	206.768 2826(46) [2.2×10^{-8}]
δ_{Mu}	-4(83) Hz
χ^2	1.24
χ^2/ν	0.41
Δv_{Mu}	4463 302.776(50) kHz [1.1×10^{-8}]

Here, the first term in the middle line is a calculation of a theoretical prediction with a value m_e/m_μ which does not depend on theory, while the other one depends on it. If we add an additional term to theory

$$\Delta v_{\text{th}} \rightarrow \Delta v'_{\text{th}} = \Delta v_{\text{th}} + \varepsilon, \quad (22)$$

we shift the prediction in the first term in the middle line by ε . In the meantime, the second value is a result of a calculation of the muonium HFS interval with a value of m_e/m_μ extracted from an indirect measurement of Δv_{Mu} with a use of a theory [see (14)]. Indeed, such a calculation should reproduce the value of Δv_{Mu} we started with. If we modify the theory, the value of m_e/m_μ would be shifted in such a way that ensures that we reproduce Δv_{Mu} from (14). That means that only the first term can be affected by a modification of the theory. Therefore, we find

$$\frac{d \Delta v_{\text{th}}^{\text{aver}}}{d\varepsilon} = 0.510. \quad (23)$$

The result has been also confirmed by a direct calculation.

Adding an additional term to the theoretical expression does not change its uncertainty and the shift affects only the central value of the outcome. While presenting the theoretical predictions in (1) and (2), we noted that they have uncertainties different by approximately a factor of 2. Now we see that their sensitivity to an additional theoretical term is also different by approximately the same factor, which makes the constraints on the additional theoretical term of the same strength.

V. MODEL-INDEPENDENT CONSISTENCY TEST

As already mentioned, there are two basic reasons to compare theory and experiment. One is to make a consistency check and the other is to constrain a possible missing term in the theoretical expression which may be either a result of an error in current calculations or a result of new physics. We start here with the former problem. One of the possibilities to make a consistency check is to perform an LSA, which is to be a simplified version of the one in Ref. [1]. Since the muonium sector does not interfere too much with an accurate determination of the involved constants (such as α , R_∞), except for m_e/m_μ , we may adopt their numerical value from Ref. [1] and perform a fit with two parameters only, m_e/m_μ and δ_{Mu} . The result is summarized in Table V and the value of χ^2/ν confirms perfect overall consistency.

Performing an LSA does not mean a direct comparison of theory and experiment, but it produces a good consistency

check. An inconsistency could show up in different ways and we should consider them one by one.

First of all, we have to be sure that the experimental results (without involvement of any theory) are consistent, and only next we have to look for consistency of theory and experiment. The direct experimental results on $\Delta\nu_{\text{Mu}}$ contain two values (one from each experiment [6,7]) obtained by rearrangement of their results in terms of $\nu(f_p)$ and $\Delta\nu_{\text{Mu}}$ (see Table I). There is one more value, given in (18), obtained by the probing the field dependence by combining $\nu(f_p)$ from two experiments. All three values are perfectly consistent.

Next, consider a determination of m_e/m_μ without any use of the theoretical expression for $\Delta\nu_{\text{Mu}}$. We again have three values. Two of them are obtained through (8) from each experiment separately. The third value is an indirect one from (17). All three determinations are in a perfect agreement.

The indirect values in both triads are independent of two others in a sense discussed above.

After concluding that the experimental results have a perfect internal consistency, we turn to a model-independent comparison of theory and experiment. The theoretical expression for the value of the HFS interval expresses the value of the HFS interval in muonium, $\Delta\nu_{\text{Mu}}$, in terms of certain fundamental constants. A prediction therefore contains a value of μ_μ/μ_p (or, which is equivalent, a value of m_e/m_μ), the uncertainty of which dominates in the numerical value of any prediction. The value of m_e/m_μ is in one or the other way obtained from the same experiment as the experimental value of $\Delta\nu_{\text{Mu}}$ to be compared with. (A determination of m_e/m_μ from an experiment of a different type is possible (see, e.g., Refs. [26,27]), but such a method is less accurate than the method applied in Refs. [6,7].

The most clean model-independent way to compare theory and experiment is to perform a comparison of a pure experimental value (or values) and a pure theoretical one, where the uncertainty of the latter does not contain any experimental contribution, while the uncertainty of the former is determined by the experimental uncertainties only. The value the best suited such a comparison is $\Delta\nu_{\text{Mu}}/\nu_F$.

A theoretical expression for the muonium HFS interval anyway expresses a theoretical prediction in units of the Fermi energy ν_F [cf. (4)]. On the experimental side, the value of ν_F by definition [see, e.g., (3)] is just a symbol for a certain combination of values of several fundamental constants. It is not a theoretical construct in a sense that it is not based on theoretical assumptions. A theoretical assumption is that it is the leading contribution to the muonium HFS interval and it is useful to express theory in terms of ν_F , but that does not relate to the numerical value of ν_F by itself. To obtain an experimental value of ν_F , one has to evaluate the experimental data without any use of the theory which is close to the evaluation in Ref. [2].

In principle, $\Delta\nu_{\text{Mu}}$ and ν_F , determined from an experiment without any use of the theory for the muonium HFS interval, have correlated uncertainties which are not of any practical importance here, since the different individual fractional uncertainties are not comparable and one (for ν_F) is essentially larger than the other (for $\Delta\nu_{\text{Mu}}$). A summary of the experimental and theoretical determination of $\Delta\nu_{\text{Mu}}$ and ν_F is given in Fig. 5. The theoretical value here follows [1].

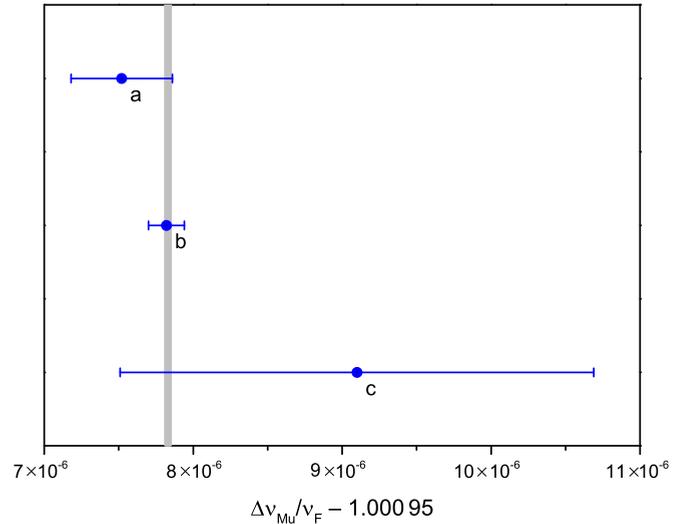


FIG. 5. Consistency test: determination of $\Delta\nu_{\text{Mu}}/\nu_F$. Here superscript ^a is the experimental result from Ref. [6], superscript ^b is from Ref. [7]; and superscript ^c is the combined result of two experiments with using indirect values (18) and (17). The band presents the value obtained by the theoretical evaluation following Ref. [1].

In addition to the correlations parameterized with correlation coefficient $r(\nu(f_p), \Delta\nu_{\text{Mu}})$, there may be an overlooked systematic effect that affects the result of all the measurements within the same experiment. Therefore, possible systematic errors in $\Delta\nu_{\text{Mu}}$ and ν_F , a search of which is one of the reasons for a comparison of theory and experiment, are also correlated. We summarize individual experimental results and their average in Table VI. The theoretical result is also presented there.

The plot and Table VI contain a relatively inaccurate experimental value of $\Delta\nu_{\text{Mu}}/\nu_F$, obtained from a comparison of two values of $\nu(f_p)$ obtained at different f_p . The numerator, $\Delta\nu_{\text{Mu}}$, is from (18), while the denominator, ν_F , is obtained with help of (17). To find their ratio, one has to take their strong anticorrelation [see (15)] into account.

The ratio $\Delta\nu_{\text{Mu}}/\nu_F$ discussed above is also suitable to constrain a new physics or any other additional theoretical term of interest, since a modification of theory [as parameterized in (22)] simply shifts the theoretical value of $\Delta\nu_{\text{Mu}}/\nu_F$ by

TABLE VI. Experimental and theoretical values of $\Delta\nu_{\text{Mu}}/\nu_F$. The average value is an average over two values from different experiments. The “Exp. c” value is related to data point c in Fig. 5. The theoretical value here is found following Ref. [1]. The ϵ LSA value is a result of a modified LSA evaluation, considered below in Sec. VI.

Source	$\Delta\nu_{\text{Mu}}/\nu_F$	u_r
Exp. 1	1.000 957 54(37)	3.7×10^{-7}
Exp. 2	1.000 957 82(12)	1.2×10^{-7}
Exp. c	1.000 9591(16)	16×10^{-7}
Average	1.000 957 792(116)	1.16×10^{-7}
ϵ LSA	1.000 957 791(116)	1.16×10^{-7}
Theory	1.000 957 82(2)	1.9×10^{-8}

ε/v_F . Actually, ε/v_F is the fractional value of the would-be new-physics term.

The ratio $\Delta v_{\text{Mu}}/v_F$ may be also used to constrain possibly missed experimental systematic effects. As mentioned, the correlation between the uncertainties of the numerator and denominator (for the results from individual experiments) has no practical importance for a calculation of the uncertainty of their ratio. In the meantime, the correlations are closely related to the problem of sensitivity to various possibly missed systematic effects. The interpretation of the comparison of the experimental values of $\Delta v_{\text{Mu}}/v_F$ to the theory is very close to the approach in Ref. [2]. We consider the sensitivity of its prediction to the presence of an overseen experimental systematic error in Sec. VI.

There is another rigorous way to constrain new physics or possible missing systematic effects that is considered in the next section.

VI. MODIFIED LSAS

As far as our concern is muonium physics, the LSA performed by CODATA [1] could be reduced for muonium data to a smaller LSA, described in Sec. II. We refer to it as a standard LSA. It allows us to perform an overall consistency test of muonium data and theory, as explained in Sec. V. If we intend to compare two values, say, theory and experiment, we may remove a part of data from the LSA procedure, to find an adjusted value and compare it with the remaining data (cf. LSA' in Table III and LSA'' in Table IV). The labels "theory" and "experiment" are not unambiguous, because a theoretical prediction usually needs an experimental input. In particular, to obtain theoretical prediction Δv_{th} in Ref. [1], it was suggested to remove direct experimental values on Δv_{Mu} but to keep all theory available. The outcome is to be compared to the direct experimental values of Δv_{Mu} . An alternative procedure, close to the one applied in Ref. [2], was to remove a theoretical datum on δ_{Mu} , which disables any application of theory, in order to find a value of m_e/m_μ and to compare the theoretical expression with such a value of m_e/m_μ to the direct experimental values on Δv_{Mu} . In Ref. [1], such a value of m_e/m_μ is also found through a rigorous LSA procedure [see Eq. (227) there] and it nearly coincides with the related value from Ref. [2].

Instead of splitting the data into two parts in order to compare one to the other, we use another approach for a comparison. We note, that because of a variety of the data, their correlations, and their different sensitivities to various effects, many meaningful constraints are rather model dependent. Therefore, we choose to directly introduce a model-dependent correction, say, an additional theoretical term ε [see (22)] and to perform an LSA with one more fit parameter. The LSA procedure should take into account all the correlations and the sensitivities to ε and produce a direct result on the additional term. Note that by observing the outcome from the constraints based on two different theoretical predictions we have found that they are rather comparable since the twice smaller uncertainty in prediction (1) is compensated by a halved sensitivity in (23) to modifications of theory. The uncertainties of two predictions are strongly correlated because the uncertainty of $v(f_p)$ from the best experiment of two [6,7] strongly

dominates and first determines the uncertainty of the extracted m_e/m_μ and next of any reasonable present-day theoretical prediction. Combining two strongly correlated predictions cannot improve their accuracy but somewhat shifts the central value. For a correct procedure to combine them, it is necessary to consider a modified LSA with the introduction of ε .

Below we consider two modified LSA procedures. They deal with the standard data, but their observational equations are modified. Different theoretical predictions have different sensitivities to various hypotheses on a new term. Technically, such a hypothesis means an introduction of a shift to certain values, which, in a practical sense, models a certain overseen systematic effect. If more than one value is shifted, the correlation of the shifts should be presented explicitly. In the simplest case, either only one value is a subject of a shift or all the shifts are expressed in the terms of a single parameter. The resulting LSA has one additional parameter and its outcome includes a constraint on it. Here we first consider a modification of the theory according to (22). That would have a sense under a suggestion of new physics or of a large overseen contribution from a regular QED theory. Since the parameter is denoted as ε , we refer to such an evaluation as to ε LSA. The related observational equations are given in Table VII.

Note that ε is not a model-independent difference of the theory and the experiment but a modification of the theory while the experiment remains intact. That is set by (7). Both experimental and theoretical values of Δv_{Mu} may contain an error we intend to constrain, but Eq. (7) contains the actual value of Δv_{Mu} and therefore we have to introduce the theoretical correction ε , e.g., due to a possible contribution of new physics, explicitly. If we are to constrain a possible experimental systematic error, we should introduce another substitution,

$$\Delta v_{\text{Mu}} \rightarrow \Delta v'_{\text{Mu}} = \Delta v_{\text{Mu}} + \varepsilon', \quad (24)$$

where we have data available on Δv_{Mu} , while the true value of the Mu HFS interval, which should be compared with the theory and be introduced in observational equation II in Table VII, is $\Delta v'_{\text{Mu}}$. We refer to such an evaluation as to ε' LSA. The observational equations are given in Table VII.

The ε LSA differs from a standard LSA because of the observational equation II in Table VII. Appearance of ε breaks the relation between the standard theoretical expression $\Delta v_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$ and the observable value Δv_{Mu} for which direct experimental data are available. In the meantime, a theoretical value, related to the actual HFS interval, which enters observational equation I, contains ε . ε LSA should bring all these things together.

One can understand how the evaluation is done without running any least-squares evaluation. $\Delta v_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}) + \varepsilon$ presents the actual value of the HFS interval, for which experimental data are available. They are not really available for $\Delta v_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$ and ε separately. Therefore the LSA procedure should basically put the experimental values of Δv_{Mu} in equation I. Equation III determines the output value for δ_{Mu} which should coincide with the input value, while equation II serves to find ε by comparing $\Delta v_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}) + \varepsilon$ with Δv_{Mu} . This logical line agrees with the way of a determination of m_e/m_μ and a subsequent theoretical prediction as done

TABLE VII. Observational equations for ε LSA and ε' LSA. ε LSA is defined as an LSA with an additional parameter ε introduced by a substitution in (22). ε' LSA is an LSA with an additional parameter ε' introduced by a substitution in (24).

No.	Equations for ε LSA	Equations for ε' LSA
II, I2	$\nu(f_p) \doteq \nu(f_p; \nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}) + \varepsilon)$	$\nu(f_p) \doteq \nu(f_p; \nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}))$
III, II2	$\Delta \nu_{\text{Mu}} \doteq \nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}) + \varepsilon$	$\Delta \nu_{\text{Mu}} \doteq \nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}}) - \varepsilon'$
III3		$\delta_{\text{Mu}} \doteq \delta_{\text{Mu}}$

in Ref. [2]. ε LSA involves some correlations and updates ignored in Ref. [2], but they are of reduced numerical importance.

The most important issue is that, as discussed above, there is another logical line, related to the CODATA's prediction [1], and if the sensitivity to ε is introduced correctly (which is unavoidable within ε LSA where it is done automatically—see equation II), then the constraint on ε is comparable with one from Ref. [2] (see explanations in Sec. IV). A comparison of the constraints is given in Table VIII. Since the prediction from Refs. [2] and [1] have different sensitivities and therefore are differently affected by the substitution (22), their comparison allows for one more constraint, which is given in the same table.

Two constraints obtained from the original predictions of Refs. [1,2] have almost the same uncertainty, which is dominated by the uncertainty of $\nu(f_p)$, with the difference invisible because of their rounding. Combining all the constraints within ε LSA cannot really improve the strength of the constraint.

A naïve constraint on new-physics term ε , which would be derived following the prediction from Ref. [1] and ignoring the sensitivity (23), would be

$$\varepsilon = -0.09(27) \text{ kHz}, \tag{25}$$

which apparently has the uncertainty incorrect by a factor of 2.

The observational equations for ε' LSA are in certain respect similar to those for ε LSA, but in some respects they are quite different and suggest different logics of the evaluation. The physical meaning of ε' introduced according (24) is an additional term for experimental values, such as a missing systematic effect which affects the results on $\Delta \nu_{\text{Mu}}$, but not on $\nu(f_p)$. Technically, that means that the measured value of $\Delta \nu_{\text{Mu}}$ deviates by ε' from the actual value of the interval, as indicated in observational equation II. Meanwhile, expression $\nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})$ correctly describes the actual value of the

TABLE VIII. Constraint on an additional theoretical term ε defined in (22). The constraint following the theoretical prediction from Ref. [1] is obtained with the sensitivity coefficient (23) taken into account.

Constraint on ε [kHz]	Ref.
-0.134(523)	Following Ref. [2]
-0.172(523)	Following Ref. [1]
-0.094(523)	Comparing Eqs. (1) and (2)
-0.135(523)	From ε LSA

HFS interval. Therefore equation I is the same as for the standard LSA.

While observational equation II breaks the relation between the measured value $\Delta \nu_{\text{Mu}}$ and the theoretical expression (with appropriated values of the parameters) in the same way for ε LSA and ε' LSA, observational equation I requires the actual value of the HFS interval as a quantity. In the case of ε LSA, the theory is uncertain because of a possible additional term, but we could use the data from the experiments for the Mu HFS interval. In the case of ε' LSA, in contrast, the uncertain part is with the measurements of $\Delta \nu_{\text{Mu}}$, while the theory is reliable. That is what produces a technical asymmetry in a comparison of the theory and the experiment.

The results are summarized in Table IX. To find the results arising from the original constraints in Refs. [1,2], one has to calculate the related sensitivities, first of the mass ratio $\partial(m_e/m_\mu)/\partial\varepsilon'$ and next of the theoretical prediction. In the case of the CODATA's constraint (1) it is zero. For (2), we obtain

$$\frac{\partial \ln \Delta \nu_{\text{th}}(m_e/m_\mu, \delta_{\text{Mu}})}{\partial \varepsilon'} = \frac{\partial \ln(m_e/m_\mu)}{\partial \varepsilon'} = -2.189 \times 10^{-10} \text{ Hz}^{-1}. \tag{26}$$

As well as in the case of the constraint on ε , we find that all the constraints have the same accuracy (after rounding), but slightly different central values. One of the constraints (based on the prediction from Ref. [1]) is easy to obtain while the prediction from Ref. [2] requires additional calculations [of the related sensitivity]. In particular, a naïve constraint using the prediction (2) while ignoring the sensitivity (26) produces the result

$$\varepsilon' = -0.13(52) \text{ kHz}, \tag{27}$$

with a twofold larger uncertainty than a correct constraint shown in Table IX.

We note that the constraint on missing experimental error significantly differs from the one on an additional theoretical term. That is not a surprise. A comparison of theory and experiment is symmetric when those two values can be

TABLE IX. Constraint on a missing systematic error ε' introduced in (24). The constraint following the theoretical prediction from Ref. [2] is obtained with the sensitivity coefficient (26).

Constraint on ε' [kHz]	Ref.
0.084(267)	Following [2]
0.088(267)	Following [1]
0.084(267)	From ε' LSA

found completely independently of each other. In the case of muonium, in addition to their comparison in observational equation II of Table VII, the actual value of $\Delta\nu_{\text{Mu}}$ enters also observational equation I, without which a theoretical prediction is not possible. That means that a calculation of a prediction requires a value of m_e/m_μ , for which either a theoretical value of $\Delta\nu_{\text{Mu}}$ or an experimental one could be applied, depending on our purpose. As we have learned above from a comparison of the theoretical predictions from Refs. [2] and [1], that makes a difference. A quantitatively meaningful model-independent comparison of the theory and the experiment is not possible with the current composition of muonium data. The uncertainty of constraints on overseen experimental and theoretical systematic errors is different, and so is the sensitivity of the predictions [1,2] to different effects.

VII. CONCLUSIONS

Theoretical predictions may serve different purposes, on the base of which we form a model of modification of standard equations by introducing certain additional terms to be constraints. The outcome is model dependent and it is important to state the purpose of a particular theoretical prediction. A constraint on an additional theoretical term (such as a new-physics contribution or a disputed term of a regular theory) and on an additional experimental term (such as a possible source of a systematic error) may be at a somewhat different level of accuracy (cf., e.g., constraints in Tables VIII and IX). Different predictions can be optimized for different purposes.

A prediction as a numerical value and its uncertainty may be not sufficient by itself for an efficient comparison to the experiment. In the meantime, the muonium experimental data has demonstrated a perfect overall consistency and once it is established a comparison of the theory and the experiment serves as a final consistency check. In particular, both recent theoretical predictions (from Refs. [2] and [1]), discussed in this paper, are equally suitable for such a test and they confirm the consistency at the same level of significance. The related deviation of the theory from the experiment is at the level of $0.25\text{--}0.32\sigma$.

However, for further applications of the predictions, it is required to present the sensitivity of the predictions to the targeted phenomena. For instance, to constrain an additional theoretical term it is necessary to explain what happens with the prediction once we add such a term to the theoretical expression. Such an effect can be not obvious (see the interpretation of the CODATA's prediction in (1) of the muonium HFS interval in Sec. IV). Naïve comparisons ignoring the sensitivity coefficients lead in certain cases to incorrect results for both recent theoretical predictions [1,2] [cf. (25) and (27)]. One of the naïve evaluations [starting from (1)] produces an incorrect constraint for an additional theoretical term, and the other [starting from (2)] an incorrect constraint for an overseen experimental error.

The other important issue is that a prediction is usually supposed to be compared with experimental values. In the case of the already existing experimental values, a correlation coefficient relating the uncertainties of the prediction and of the experiment is required. The optimized prediction for an

additional theoretical term is close to that of Eides [2] [see (2)]. A correlation between the theoretical prediction and the experiment is present in our evaluations, but fortunately it is only marginally important.

Speaking about the existing results on muonium, we note that none of the mentioned values of $\Delta\nu_{\text{th}}$ is perfectly suited for a comparison of theory and experiment. Use of such values is limited by the accuracy of the input parameter m_e/m_μ or, alternatively, μ_μ/μ_p . A value which can be found from theory and applied for a comparison with a future experiment is a value of $\Delta\nu_{\text{Mu}}/\nu_F$. This quantity can be determined experimentally in a measurement of the HFS interval at magnetic field. (Dealing with a value of ν_F determined without any theoretical input is close to the logics of Ref. [2]; however, to make a comparison for future more accurate experiments we will apparently use their ν_F , i.e., we are to produce a new theoretical prediction of Mu HFS interval in Hz.) Until an accurate independent determination of m_e/m_μ appears (through, e.g., a measurement of the $1s\text{--}2s$ transition in muonium), it is better to operate with $\Delta\nu_{\text{Mu}}/\nu_F$. One can compare results on such a value, extracted from different experiments, and compare experimental and theoretical results (see Fig. 5 and Table VI).

Above, we have described a rigorous approach to constrain a parameter space left for a new-physics term. Theory can be almost continuously updated, and sometimes it is. New experimental results can appear as well. For this reason, it is important to deal with estimations that should be simple for understanding and would allow for modifications and updates. A value of $\Delta\nu_{\text{Mu}}/\nu_F$, calculated rigorously from all available experimental data, has a fractional uncertainty of 1.16×10^{-7} . If we ignore the correlations between $\Delta\nu_{\text{Mu}}$ and ν_F , the uncertainty is increased to 1.17×10^{-7} . Excluding correlation between $\nu(f_p)$ and $\Delta\nu_{\text{Mu}}$ within the same experiments would change the uncertainty to 1.19×10^{-7} . If we consider only the most accurate experiment [7] and ignore the related correlations the uncertainty of $\Delta\nu_{\text{Mu}}/\nu_F$, it is 1.25×10^{-7} . In other words, a rigorous consideration is capable of providing a marginal improvement of the uncertainty compared with a straightforward evaluation of the data of the best experiment only. A comparison of two experiments [6,7] is necessary to check their consistency, but once the consistency is established, one may use the data only from the more accurate experiment.

We have to emphasize that a short answer to a controversy of two theoretical predictions (from Refs. [2] and [1]) is that to constrain a new-physics term or a possible missing term of QED theory, the prediction (2) [2] can be directly used and the required corrections are rather marginal. Prediction (1) [1] is not incorrect, but it is confusing and cannot be used in a straightforward way. Any straightforward use of it leads to incorrect results with the central value and the uncertainty underestimated by approximately a factor of 2 [cf. (25) and Table VIII]. All the constraints on new physics derived following the theoretical prediction from Ref. [1] and ignoring the sensitivity coefficient of (23) should be corrected.

The central values of all the individual predictions in Table VIII are somewhat different but still roughly the same because of a good overall consistency of the muonium data (see Table V). The latter means that all the possible additional terms, whatever they are, should be consistent with zero.

As explained, the strongest and the most rigorous constraint can be obtained with a complete use of the available information through a relatively complicated procedure. With an experimental progress in muonium physics (see, e.g., Refs. [28,29]), a difference of its outcome and a simplistic approach of Ref. [2] may be important.

To conclude, we have also to mention that the theoretical uncertainty in Refs. [2] and [1] is somewhat different, which is not that important as far as it is essentially below the uncertainty due to determination of ν_F . In this paper, we deliberately avoided any discussion of its value. The paper is focused on a comparison of various procedures to produce theoretical predictions on the muonium HFS interval and on their subsequent use for a comparison of theory and experiment. We have applied a theoretical expression from Ref. [1], where all the related contributions are fully described. Having in mind a possible experimental progress, the theoretical uncertainty should be revisited, which is a completely separate issue.

Personally, we believe that a real theoretical uncertainty should be somewhat larger than that in Refs. [2] and [1]. However, as one can see from plots in Figs. 2 and 4, even a magnification of the uncertainty by a factor of 5 would

weakly affect the accuracy of the average value of m_e/m_μ and therefore of the theoretical prediction for the muonium hyperfine interval.

The choice of an appropriate approach to produce a theoretical prediction, whatever it is, and to constrain a space for an additional theoretical term, becomes a key question for the interpretation of the muonium HFS data because of differences in existing predictions. This question is revisited in this paper. Our constraint for the new theoretical term (and in particular a new-physics one) is $-0.14(52)$ kHz. The result is consistent with a result based on Ref. [2]. It is also consistent with a result based on the prediction from Ref. [1] once its sensitivity factor is taken into account. Such a result is also derived in this paper.

ACKNOWLEDGMENTS

The authors are grateful to M. I. Eides, V. G. Ivanov, K. Jungmann, D. Newell, B. Ohayon, V. A. Shelyuto, K. Shimomura, and R. Szafron for stimulating conversations. The work was supported in part by the DFG (Grant No. KA 4645/1-1) and by ERC (under H2020 Grant No. 742247).

-
- [1] P. J. Mohr, D. B. Newell, and B. N. Taylor, *Rev. Mod. Phys.* **88**, 035009 (2016).
 - [2] M. I. Eides, *Phys. Lett. B* **795**, 113 (2019).
 - [3] M. I. Eides, H. Grotch, and V. A. Shelyuto, *Phys. Rep.* **342**, 63 (2001).
 - [4] S. G. Karshenboim, *Phys. Rep.* **422**, 1 (2005).
 - [5] J. Sapirstein and D. R. Yennie, in *Quantum Electrodynamics*, edited by T. Kinoshita (World Scientific, Singapore, 1990), pp. 560–672.
 - [6] F. G. Mariam, W. Beer, P. R. Bolton, P. O. Egan, C. J. Gardner, V. W. Hughes, D. C. Lu, P. A. Souder, H. Orth, J. Vetter, U. Moser, and G. zu Putlitz, *Phys. Rev. Lett.* **49**, 993 (1982).
 - [7] W. Liu, M. G. Boshier, S. Dhawan, O. van Dyck, P. Egan, X. Fei, M. Grosse Perdekamp, V. W. Hughes, M. Janousch, K. Jungmann, D. Kawall, F. G. Mariam, C. Pillai, R. Prigl, G. zu Putlitz, I. Reinhard, W. Schwarz, P. A. Thompson, and K. A. Woodle, *Phys. Rev. Lett.* **82**, 711 (1999).
 - [8] V. W. Hughes and G. zu Putlitz, in *Quantum Electrodynamics*, edited by T. Kinoshita (World Scientific, Singapore, 1990), pp. 822–904. Note: the classical book by H. A. Bethe and E. E. Salpeter (*Quantum Mechanics of One and Two Electron Atoms*, Springer-Verlag, Berlin, 1957) contains errors in the related formula (47.9).
 - [9] X. Fei, V. W. Hughes, and R. Prigl, *Nucl. Instrum. Methods Phys. Res., Sect. A* **394**, 349 (1997).
 - [10] W. D. Phillips, W. E. Cooke, and D. Kleppner, *Phys. Rev. Lett.* **35**, 1619 (1975).
 - [11] W. D. Phillips, W. E. Cooke, and D. Kleppner, *Metrologia* **13**, 179 (1977).
 - [12] B. W. Petley and R. W. Donaldson, *Metrologia* **20**, 81 (1984).
 - [13] Y. I. Neronov and A. E. Barzakh, *Sov. Phys. JETP* **48**, 769 (1978) [*Zh. Eksp. Teor. Fiz.* **75**, 1521 (1978)].
 - [14] G. Schneider, A. Mooser, M. Bohman, N. Schön, J. Harrington, T. Higuchi, H. Nagahama, S. Sellner, C. Smorra, K. Blaum, Y. Matsuda, W. Quint, J. Walz, and S. Ulmer, *Science* **358**, 1081 (2017); A. Mooser, S. Ulmer, K. Blaum, K. Franke, H. Kracke, C. Leuteritz, W. Quint, C. C. Rodegheri, C. Smorra, and J. Walz, *Nature (London)* **509**, 596 (2014).
 - [15] P. J. Mohr and B. N. Taylor, *Rev. Mod. Phys.* **72**, 351 (2000).
 - [16] M. Tanabashi *et al.* (Particle Data Group), *Phys. Rev. D* **98**, 030001 (2018).
 - [17] S. G. Karshenboim and V. G. Ivanov, *Can. J. Phys.* **80**, 1305 (2002).
 - [18] E. A. Moore, *Mol. Phys.* **97**, 375 (1999).
 - [19] N. C. Pyper, *Mol. Phys.* **97**, 381 (1999); N. C. Pyper and Z. C. Zhang, *ibid.* **97**, 391 (1999).
 - [20] M. I. Eides and H. Grotch, *Ann. Phys.* **260**, 191 (1997).
 - [21] G. W. Bennett, B. Bousquet, H. N. Brown, G. Bunce, R. M. Carey, P. Cushman, G. T. Danby, P. T. Debevec, M. Deile, H. Deng, W. Deninger, S. K. Dhawan, V. P. Druzhinin, L. Duong, E. Efstathiadis, F. J. M. Farley, G. V. Fedotovitch, S. Giron, F. E. Gray, D. Grigoriev, M. Grosse-Perdekamp, A. Grossmann, M. F. Hare, D. W. Hertzog, X. Huang, V. W. Hughes, M. Iwasaki, K. Jungmann, D. Kawall, M. Kawamura, B. I. Khazin, J. Kindem, F. Krienen, I. Kronkvist, A. Lam, R. Larsen, Y. Y. Lee, I. Logashenko, R. McNabb, W. Meng, J. Mi, J. P. Miller, Y. Mizumachi, W. M. Morse, D. Nikas, C. J. G. Onderwater, Y. Orlov, C. S. Özben, J. M. Paley, Q. Peng, C. C. Polly, J. Pretz, R. Prigl, G. zu Putlitz, T. Qian, S. I. Redin, O. Rind, B. L. Roberts, N. Ryskulov, S. Sedykh, Y. K. Semertzidis, P. Shagin, Yu. M. Shatunov, E. P. Sichtermann, E. Solodov, M. Sossong, A. Steinmetz, L. R. Sulak, C. Timmermans, A. Trofimov, D. Urner, P. von Walter, D. Warburton, D. Winn, A. Yamamoto, and D. Zimmerman (Muon (g-2) Collaboration), *Phys. Rev. D* **73**, 072003 (2006).

- [22] S. G. Karshenboim and V. G. Ivanov, *Phys. Lett. B* **566**, 27 (2003).
- [23] P. F. Winkler, D. Kleppner, T. Myint, and F. G. Walther, *Phys. Rev. A* **5**, 83 (1972).
- [24] T. Beier, H. Häffner, N. Hermanspahn, S. G. Karshenboim, H.-J. Kluge, W. Quint, S. Stahl, J. Verdú, and G. Werth, *Phys. Rev. Lett.* **88**, 011603 (2001).
- [25] S. Sturm, F. Köhler, J. Zatorski, A. Wagner, Z. Harman, G. Werth, W. Quint, C. H. Keitel, and K. Blaum, *Nature (London)* **506**, 467 (2014).
- [26] V. Meyer, S. N. Bagayev, P. E. G. Baird, P. Bakule, M. G. Boshier, A. Breitrück, S. L. Cornish, S. Dychkov, G. H. Eaton, A. Grossmann, D. Hübl, V. W. Hughes, K. Jungmann, I. C. Lane, Y.-W. Liu, D. Lucas, Y. Matyugin, J. Merkel, G. zu Putlitz, I. Reinhard, P. G. H. Sandars, R. Santra, P. V. Schmidt, C. A. Scott, W. T. Toner, M. Towrie, K. Träger, L. Willmann, and V. Yakhontov, *Phys. Rev. Lett.* **84**, 1136 (2000).
- [27] F. Maas, B. Braun, H. Geerds, K. Jungmann, B. E. Matthias, G. zu Putlitz, I. Reinhard, W. Schwarz, L. Williams, L. Zhang, P. E. G. Baird, P. G. H. Sandars, G. Woodman, G. H. Eaton, P. Matousek, T. Toner, M. Towrie, J. R. M. Barr, A. I. Ferguson, M. A. Persaud, E. Riis, D. Berkeland, M. Boshier, and V. W. Hughes, *Phys. Lett. A* **187**, 247 (1994).
- [28] P. Strasser, M. Abe, M. Aoki, S. Choi, Y. Fukao, Y. Higashi, T. Higuchi, H. Iinuma, Y. Ikedo, K. Ishida, T. Ito, T. U. Ito, M. Iwasaki, R. Kadono, O. Kamigaito, S. Kanda, K. Kawagoe, D. Kawall, N. Kawamura, M. Kitaguchi, A. Koda, K. M. Kojima, K. Kubo, M. Matama, Y. Matsuda, Y. Matsudate, T. Mibe, Y. Miyake, T. Mizutani, K. Nagamine, S. Nishimura, T. Ogitsu, N. Saito, K. Sasaki, S. Seo, H. M. Shimizu, K. Shimomura, T. Suehara, M. Tajima, K. S. Tanaka, T. Tanaka, J. Tojo, D. Tomono, H. A. Torii, E. Torikai, A. Toyoda, Y. Tsutsumi, K. Ueno, Y. Ueno, D. Yagi, A. Yamamoto, T. Yamanaka, T. Yamazaki, H. Yasuda, M. Yoshida, and T. Yoshioka (MuSEUM Collaboration), *EPJ Web Conf.* **198**, 00003 (2019).
- [29] I. Belosevic, A. Antognini, Y. Bao, A. Eggenberger, M. Hildebrandt, R. Iwai, D. M. Kaplan, K. S. Khaw, K. Kirch, A. Knecht, A. Papa, C. Petitjean, T. J. Phillips, F. M. Piegsa, N. Ritjoho, A. Stoykov, D. Taqq, and G. Wichmann, *Hyp. Inter.* **240**, 41 (2019).