

Self-consistent extraction of spectroscopic bounds on light new physics

Cédric Delaunay, Jean-Philippe Karr, Teppei Kitahara, Jeroen C.J. Koelemeij, Yotam Soreq, Jure Zupan

▶ **To cite this version:**

Cédric Delaunay, Jean-Philippe Karr, Teppei Kitahara, Jeroen C.J. Koelemeij, Yotam Soreq, et al.. Self-consistent extraction of spectroscopic bounds on light new physics. Physical Review Letters, 2023, 130 (12), pp.121801. 10.1103/PhysRevLett.130.121801 . hal-03839767

HAL Id: hal-03839767
https://hal.science/hal-03839767

Submitted on 25 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LAPTH-063/22, CERN-TH-2022-158, KEK-TH-2454

Self-consistent extraction of spectroscopic bounds on light new physics

Cédric Delaunay,[1,2,*] Jean-Philippe Karr,[3,4,†] Teppei Kitahara,[5,6,7,‡] Jeroen C. J. Koelemeij,[8,§] Yotam Soreq,[9,¶] and Jure Zupan[10,**]

[1] *Laboratoire d'Annecy-le-Vieux de Physique Théorique, CNRS – USMB, BP 110 Annecy-le-Vieux, F-74941 Annecy, France*
[2] *Theoretical Physics Department, CERN, Esplanade des Particules 1, Geneva CH-1211, Switzerland*
[3] *Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-Université PSL, Collège de France, 4 place Jussieu, F-75005 Paris, France*
[4] *Université d'Evry-Val d'Essonne, Université Paris-Saclay, Boulevard François Mitterrand, F-91000 Evry, France*
[5] *Institute for Advanced Research & Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Nagoya 464–8602, Japan*
[6] *KEK Theory Center, IPNS, KEK, Tsukuba 305–0801, Japan*
[7] *CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*
[8] *LaserLaB, Department of Physics and Astronomy, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands*
[9] *Physics Department, Technion – Israel Institute of Technology, Haifa 3200003, Israel*
[10] *Department of Physics, University of Cincinnati, Cincinnati, Ohio 45221,USA*

arXiv:2210.10056v1 [hep-ph] 18 Oct 2022

Fundamental physical constants are determined from a collection of precision measurements of elementary particles, atoms and molecules. This is usually done under the assumption of the Standard Model (SM) of particle physics. Allowing for light new physics (NP) beyond the SM modifies the extraction of fundamental physical constants. Consequently, setting NP bounds using these data, and at the same time assuming the CODATA recommended values for the fundamental physical constants, is not reliable. As we show in this Letter, both SM and NP parameters can be simultaneously determined in a consistent way from a global fit. For light vectors with QED-like couplings, such as the dark photon, we provide a prescription that recovers the degeneracy with the photon in the massless limit, and requires calculations only at leading order in the small new physics couplings. At present, the data show tensions partially related to the proton charge radius determination. We show that these can be alleviated by including contributions from a light scalar with flavor non-universal couplings.

I. INTRODUCTION

Precision measurements of atomic and molecular properties play a dual role in fundamental physics. On the one hand, assuming the Standard Model (SM) of particle physics, these are used to determine two of the SM parameters, the fine-structure constant, α, and the electron mass, m_e (through the Rydberg constant $R_\infty \equiv \alpha^2 m_e c/(2h)$), along with a number of other observables such as the charge radii and relative atomic masses of the proton and deuteron. An example is the determination of fundamental physical constants by the Committee on Data of the International Science Council (CODATA) [1].

On the other hand, precision measurements can be used to search for new physics (NP) beyond the SM. Such searches have been conducted using measurements of single particle observables [2–4], atomic systems [5–10], and molecular systems [11–14], see [15] for a review. The presence of NP would manifest itself as a discrepancy between measurements and theoretical SM predictions. The difficulty here is that in many cases the SM predictions depend on the fundamental physics parameters, which in turn were extracted from data by CODATA *under the assumption that the SM is correct, and no NP exists*. In general, the presence of NP would affect the extraction of fundamental constants, possibly reducing the claimed sensitivity of NP searches. This subtlety is more often than not ignored in the literature.

In this Letter we propose and carry out a self-consistent determination of constraints on light NP models by performing a global fit, simultaneously extracting the SM and NP parameters. We go well beyond the previous studies [5, 6, 16], which were performed only on subsets of data. We pay special attention to the potentially problematic limit of massless NP. The challenge is that the SM predictions are calculated to a higher perturbative order than the leading order (LO) NP contributions, which can then lead to incorrect limiting behaviour for very light NP. Below, we provide a prescription, valid to LO in NP parameters, that corrects for such mismatches in the theoretical predictions, and leads to the proper massless NP limit.

The global fit shows several $3\,\sigma$ (3 standard deviations) discrepancies between observables and predictions, as-

* cedric.delaunay@lapth.cnrs.fr
† karr@lkb.upmc.fr
‡ teppeik@kmi.nagoya-u.ac.jp
§ j.c.j.koelemeij@vu.nl
¶ soreqy@physics.technion.ac.il
** zupanje@ucmail.uc.edu

suming the SM. These anomalies are well known: they correspond to the measurements constituting the proton charge radius puzzle [17–19], with the addition of new measurements of hydrogen transitions [20, 21]. Reference [20] showed the tension of their $2S_{1/2}-8D_{5/2}$ measurement with other hydrogen data is relaxed in the presence of an additional Yukawa-like interaction. Our global analysis, which determines simultaneously both the SM and NP parameters, shows for the first time that all these deviations can be largely accounted for in a single NP model – a light scalar that couples to gluons, electrons and muons.

II. NEW PHYSICS BENCHMARK MODELS

We focus on minimal extensions of the SM, where either a light scalar boson, ϕ, or a light vector boson, ϕ_μ, is added to the spectrum of SM particles. The new light particle is assumed to have parity conserving interactions with the SM electrons and muons, as well as with light quarks, resulting in couplings to neutrons and protons[1]. The interaction Lagrangian is therefore given by

$$\mathcal{L}_{\rm int} = \sum_{i=e,\mu,n,p} g_i \overline{\psi}_i (\Gamma\cdot\phi)\psi_i \,, \tag{1}$$

where $\Gamma\cdot\phi \equiv \phi, \gamma^\mu\phi_\mu$ for spin $s=0,1$ bosons, respectively. Taking the nonrelativistic limit for ψ_i, and working at LO in g_i, the tree level exchange of ϕ or ϕ_μ induces a Yukawa-like nonrelativistic potential,

$$V^{ij}_{\rm NP}(r) = (-1)^{s+1}\alpha_\phi q_i q_j \frac{e^{-m_\phi r}}{r} \,, \tag{2}$$

between particles ψ_i and ψ_j, separated by a distance r. The NP coupling constant, $\alpha_\phi \equiv |g_e g_p|/(4\pi) > 0$, gives the strength of the NP induced potential between electrons and protons. The strength of NP interactions between fermions ψ_i and ψ_j, relative to the electron–proton one, is given by the product of effective NP couplings, $q_i q_j$, where $q_i \equiv g_i/\sqrt{|g_e g_p|}$. In particular, for the electron–proton system the product of effective NP couplings can take the values, $q_e q_p = \pm 1$. For $q_i q_j > 0$ the potential (2) is attractive (repulsive) for spin 0 (1) mediator ϕ, and vice versa for $q_i q_j < 0$.

In the numerical analysis, we consider the following benchmark NP models:

a. Dark photon. The light NP mediator is a vector boson with couplings to the SM fermions proportional to their electric charges. A UV complete realization is an additional abelian gauge boson with field strength $F'_{\mu\nu}$, that couples to the SM through the renormalizable kinetic mixing interaction, $-(\epsilon/2)F'_{\mu\nu}F^{\mu\nu}$ [22], where $F_{\mu\nu}$ is the electromagnetic field strength. To LO in ϵ this yields $\alpha_\phi = \alpha\epsilon^2$ and $q_e = q_\mu = -q_p = -1$, $q_n = 0$.

b. $B-L$ gauge boson. The difference of baryon (B) and lepton (L) numbers is non-anomalous, and can be gauged without introducing new fermions [23, 24]. Light $B-L$ gauge boson with gauge coupling g_{B-L} gives rise to the NP potential in (2) with $\alpha_\phi = g^2_{B-L}/(4\pi)$. The charges $q_e = q_\mu = -q_p = -q_n = -1$ coincide with the dark photon ones, except for neutron. Comparison of $B-L$ and dark photon bounds illustrates the importance of performing spectroscopy of different isotopes of the same species, such as hydrogen and deuterium.

c. Scalar Higgs portal. A light scalar mixing with the Higgs boson [25, 26] inherits the SM Yukawa structure, giving $\alpha_\phi = \sin^2\theta\, m_e\kappa_p m_p/(4\pi v^2) \lesssim 1.8\times10^{-10}$ where $v\simeq 246\,$GeV is the SM Higgs vacuum expectation value, and θ the scalar mixing angle. The effective leptonic ($\ell = e,\mu$) charges are $q_\ell = m_\ell/\sqrt{m_e\kappa_p m_p}$, while the effective nucleon charges ($N = p,n$) are given by $q_N = \kappa_N m_N/\sqrt{m_e\kappa_p m_p}$ with $\kappa_p \simeq 0.306(14)$ and $\kappa_n \simeq 0.308(14)$ [27–32] (see also Sec. S4). Since couplings to muons and nucleons are enhanced by $q_\mu/q_e = m_\mu/m_e \simeq 200$ and $g_N/q_e = m_N/m_e \simeq 2\times10^3$, respectively, this NP benchmark highlights the relevance of muonic atom and molecular spectroscopy.

d. Hadrophilic scalar. A scalar with $q_\ell = 0$ and $q_N = \kappa_N m_N/\sqrt{m_e\kappa_p m_p}$, *i.e.*, with vanishing couplings to leptons, highlights the importance of molecular hydrogen ion spectroscopy as a probe of internuclear interactions [11, 14]. For expedience we take g_N to be the same as for the Higgs portal, but this could be relaxed in general.

e. Up-lepto-darko-philic (ULD) scalar. In order to evade strong bounds from $K^+\to\pi^+ + X_{\rm inv}$ searches, where $X_{\rm inv}$ are invisible particles that escape the detector, see Sec. V, we adopt a particular version of a light scalar benchmark. The ULD scalar has enhanced couplings to leptons, $q_\ell = m_\ell/\sqrt{m_e\kappa'_p m_p}$, and reduced couplings to nucleons (due to couplings to only the up quark), $q_N = \kappa'_N m_N/\sqrt{m_e\kappa'_p m_p}$, with $\kappa'_p \simeq 0.018(5)$ and $\kappa'_n \simeq 0.016(5)$, and $\alpha_\phi = k^2 m_e\kappa'_p m_p/(4\pi v^2)$, with k a dimensionless parameter controlling the overall strength of interactions, which is varied in the fit. The ϕ is assumed to predominantly decay to invisible states, possibly related to the dark matter, which evades constraints from beam dump experiments. See Sec. S4 in the supplemental material for further details, including results for an additional NP benchmark model– the scalar photon.

[1] Extension to parity non-conserving couplings and additional particles is straightforward.

III. DATASETS

The adjustment of parameters, *i.e.*, the fitting procedure, presented in this work has been carried out using two different datasets, CODATA18 and DATA22. The CODATA18 dataset consists of data that was used in the latest CODATA adjustment in Ref. [1], but restricted only to the subset most relevant for constraining NP. This subset contains observables related to the determination of the Rydberg constant R_∞, the proton and deuteron radii, r_p and r_d respectively, the fine-structure constant α, and the relative atomic masses of the electron, proton, and deuteron: $A_{\rm r}(e)$, $A_{\rm r}(p)$ and $A_{\rm r}(d)$, respectively. The inputs are listed in Tables S1, S3, and include theory uncertainties in Table S2. The other observables and parameters included in the CODATA 2018 adjustment are very weakly correlated with the selected data, and can be neglected for our purpose.

The DATA22 dataset combines the updated CODATA18 inputs with the additional data that improve the overall sensitivity to NP (see Table S6 and S9). In particular, we include the measurements of transition frequencies in simple molecular or molecule-like systems, the hydrogen deuteride molecular ion (HD^+) [33–35], and the antiprotonic helium atom ($\bar{p}^3$He and $\bar{p}^4$He) [36, 37]. These have an enhanced sensitivity to the NP models with mediators that have large couplings to quarks (and thus nuclei). The three benchmark models of this type are the Higgs portal, hadrophilic and ULD scalars, cf. Sec II.

The CODATA18 dataset is used as a reference point to verify the implementation of the inputs and the adjustment procedure, while DATA22 is used to obtain our nominal results. The full list of data in the two datasets, as well as further discussion of the importance of including certain observables when constraining NP, is given in Supplemental Material, Sections. S2 and S3.

IV. LEAST-SQUARES ADJUSTMENT WITH NEW PHYSICS

The experimental data are compared to the theoretical predictions with NP following the linearized least-squares procedure [38]. The theoretical prediction for an observable $\mathcal{O}$ takes the form,

$$\mathcal{O} = \mathcal{O}_{\rm SM}(g_{\rm SM}) + \mathcal{O}_{\rm NP}(g_{\rm SM}, \alpha_\phi, m_\phi) + \delta\mathcal{O}_{\rm th}\,, \qquad (3)$$

where $\mathcal{O}_{\rm SM}$ is the state of the art SM prediction, and depends on the SM parameters $g_{\rm SM} = \{R_\infty, r_p, r_d, \alpha, A_{\rm r}(e), A_{\rm r}(p), A_{\rm r}(d)\}$, while the NP contribution $\mathcal{O}_{\rm NP}$ depends in addition on α_ϕ and m_ϕ. The theoretical uncertainties are included as in Ref. [1], by adding a normally distributed variable $\delta\mathcal{O}_{\rm th}$ with zero mean and standard deviation equal to the estimated uncertainty of the theoretical expression. The $\delta\mathcal{O}_{\rm th}$'s are treated as yet another set of input data and varied in the fit, along with $g_{\rm SM}$, α_ϕ, and m_ϕ, in order to minimize the χ^2 function constructed from the input data and theory predictions (see also Sec. S1).

The SM theoretical predictions for atomic transition frequencies, the electron anomalous magnetic moment, and bound-electron g-factors are from Ref. [1] (see references therein). The predictions for the $HD^+$and $\bar{p}$He transition frequencies are from Ref. [39, 40] and [41–43], respectively, and are updated with the latest CODATA recommended values, see Sec. S3 for details.

The NP contributions to atomic and molecular ion transition frequencies are obtained using (time-independent) first-order perturbation theory [44, 45]. We use exact nonrelativistic wavefunctions for hydrogen-like atoms and very precise nonrelativistic numerical ones from a variational method of Ref. [46] for $HD^+$and $\bar{p}$He. Expectation values of the Yukawa potentials in Eq. (2) are calculated for a grid of m_ϕ values, taking advantage of the fact that their matrix elements in the chosen basis can be obtained in an analytical form. The precision is limited to $\mathcal{O}(\alpha^2)$ because of the neglected relativistic corrections to the wavefunction. The NP contribution to the free-electron $(g-2)_e$ arises at one-loop [47, 48], while for bound electrons we include an additional tree-level contribution from electron-nucleus interaction [49]. Finally, we assume NP to have negligible effects in atom recoil measurements as well as relative atomic mass measurements from cyclotron frequency measurements in Penning traps.

We pay particular attention to the possible degeneracy between the determination of SM and NP parameters. In the $m_\phi \to 0$ limit, the dark photon is completely degenerate with the QED photon, since couplings of the two are aligned, $q_i = Q_i$, and thus only the combination $\alpha + \alpha_\phi$ can be determined from data. This degeneracy should be retained in the theoretical predictions (3), which in principle requires calculating NP effects to the same very high order as the SM. We propose an alternative procedure, which uses the state-of-the-art SM calculations but requires NP contribution only at LO in α_ϕ, and reproduces the correct $q_i \to Q_i$, $m_\phi a_0 \ll 1$ limit, where $a_0 \equiv \alpha/(4\pi R_\infty) = (\alpha m_e)^{-1}$ is the Bohr radius.

For *light vectors* we rewrite the NP potential in Eq. (2) as the sum of the Coulomb-like potential with QED coupling Q_i plus the remainder,

$$V_{\rm NP}^{ij}(r) = \alpha_\phi \frac{Q_i Q_j}{r} + \widetilde{V}_{\rm NP}^{ij}(r)\,, \qquad (4)$$

where $\widetilde{V}_{\rm NP}^{ij}(r) \equiv \alpha_\phi \left(q_i q_j e^{-m_\phi r} - Q_i Q_j\right)/r$. The theory predictions are evaluated at LO in $\widetilde{V}_{\rm NP}(r)$, while the NP Coulomb term and the related relativistic corrections are evaluated to the same order as the SM, which amounts to replacing $\alpha \to \alpha + \alpha_\phi$ in the SM predictions. For any observable $\mathcal{O}$ the theoretical prediction is then

$$\mathcal{O} = \mathcal{O}_{\rm SM}\left(\alpha + \alpha_\phi\right) + \widetilde{\mathcal{O}}_{\rm NP}\left(\alpha + \alpha_\phi, \alpha_\phi, m_\phi\right)\,, \qquad (5)$$

where $\mathcal{O}_{\rm SM}$ is the SM contribution now expressed as a function of $\alpha + \alpha_\phi$ and $\widetilde{\mathcal{O}}_{\rm NP}$ is the NP contribution from

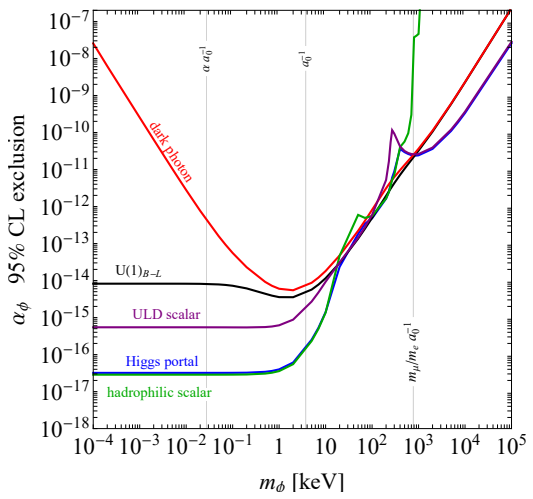


FIG. 1. The 95% CL bounds on the NP coupling constant α_ϕ as a function of the new boson's mass m_ϕ for the benchmark NP models of Sec. II, as indicated. Other model-dependent constraints may apply (see text).

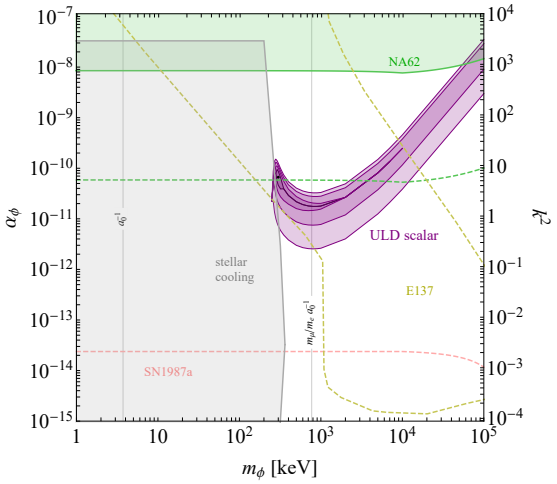


FIG. 2. The constraints on ULD scalar in the α_ϕ, m_ϕ plane, with purple-shaded 1, 2, 3, 4 σ CL regions favored by the DATA22 dataset (black dot is the best-fit point). Exclusions are by SN1987a [50, 51] (below the pink line, absent if ϕ invisible decay dominates), NA62 $K^+ \to \pi^+ X_{\rm inv}$ search [52] (green, the dashed line is a naive NNLO estimate), stellar cooling [53] (gray), and E137 [54, 55] (between yellow dashed lines, absent if ϕ invisible decay dominates).

$\widetilde{V}_{\rm NP}$. In the $m_\phi \to 0, q_i \to Q_i$ limit the potential $\widetilde{V}_{\rm NP}$ vanishes, and all theory predictions are the SM ones, but shifted by $\alpha \to \alpha + \alpha_\phi$. For massive dark photon with $m_\phi a_0 \ll 1$ the leading effect of $\widetilde{V}_{\rm NP}$ is parametrically $\widetilde{\mathcal{O}}_{\rm NP} \propto m_\phi^2$. Note that for massless $B-L$ the potential $\widetilde{V}_{\rm NP}$ vanishes in hydrogen but not in deuterium where $\widetilde{\mathcal{O}}_{\rm NP} \propto q_n$, thus breaking the degeneracy between the SM and NP contributions when $m_\phi \to 0$.

For *light scalars* there is no degeneracy with QED in the massless mediator limit; it is lifted by relativistic corrections. We can use directly the state-of-the-art SM predictions, and simply add to them the NP contribution due to the potential (2) at LO, without any special treatments, see also Sec. S7.

V. RESULTS

First, we perform the control fit, *i.e.*, the least-squares adjustment assuming SM, based on the CODATA18 dataset with inflated experimental uncertainties when there are tensions in the data [1], see also Sec. S2. Resulting χ^2 per degree-of-freedom (dof) is $\chi^2_{\rm SM}/\nu_{\rm dof} \simeq 0.95$ ($\nu_{\rm dof} = 78-44 = 34$), indicating an overall good description by the SM, and the use of correct expansion factors. The output $g_{\rm SM}$ values and relative uncertainties (see Table S5) are in excellent agreement ($\lesssim 0.2\,\sigma$) with the latest CODATA recommended values [1], validating our procedure.

Next, we perform adjustments based on the DATA22 dataset, assuming either the SM or one of the NP benchmark models in Sec. II. We do not inflate experimental errors, since mild tensions in the data could be a hint of NP. The SM-only hypothesis still describes the data relatively well, with $\chi^2_{\rm SM}/\nu_{\rm dof} \simeq 1.4$ ($\nu_{\rm dof} = 102-62 = 40$), despite known tensions in the proton charge radius puzzle data and the recent hydrogen $2S_{1/2} - 8D_{5/2}$ transition [20].

Figure 1 shows the 95 % confidence level (CL) upper bounds on α_ϕ as function of m_ϕ for the NP benchmark models, Sec. II. The strongest exclusion is always reached around $m_\phi \sim a_0^{-1} \sim 4\,$keV, and stays roughly constant for lighter m_ϕ (except for dark photon due to degeneracy with QED in the $m_\phi \to 0$ limit, see Sec. IV). Deuterium observables translate to a $\sim 2\times$ stronger bound on $B-L$ at $m_\phi \sim a_0^{-1}$, compared to dark photon. The significantly stronger bounds on the Higgs portal and hadrophilic scalar for $m_\phi \lesssim 10\,$keV are due to the $\sim \kappa_p m_p/m_e \simeq 500$ enhancement in internucleon interactions (compared to electron–nucleon potential), affecting the HD^+ observables. For heavier NP, $m_\phi a_0 \gtrsim 1$ (m_μ/m_e) in hydrogen (muonic hydrogen), the interaction is point-like, with suppressed electron (muon) wave function overlap, and the bounds decouple as $\propto 1/m_\phi^2$ (and more quickly for hadrophilic scalar). The bounds are stronger for Higgs portal and ULD scalar due to $\sim m_\mu/m_e \simeq 200$ enhanced effects in muonic hydrogen.

For $m_\phi a_0 \gtrsim m_\mu/m_e$ the Higgs portal and ULD scalar are statistically preferred over the SM at the $\sim 4\,\sigma$ and $\sim 5\,\sigma$ level, respectively. Figure 2 shows the preferred region for the ULD scalar, around the best-fit point, $m_\phi = 300\,$keV and $\alpha_\phi = 6.7 \times 10^{-11}$. This NP hint is supported mostly from the recent measurements of the hydrogen $2S_{1/2} - 8D_{5/2}$ and $1S_{1/2} - 3S_{1/2}$ transitions [20, 21], as well as muonic deuterium, *cf.* Sec. S5. While these tensions between data and the SM predic-

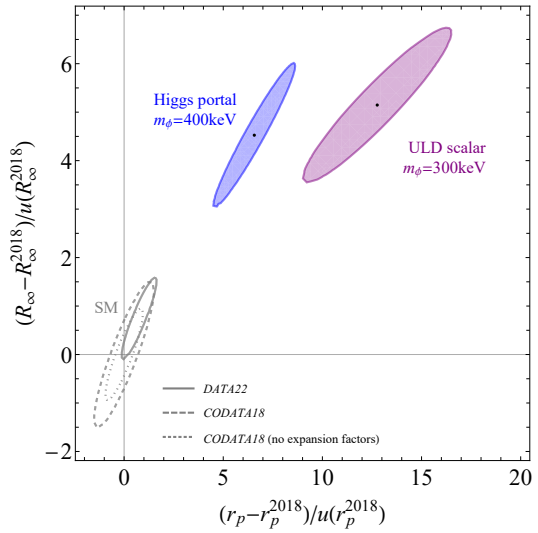


FIG. 3. The 68% CL regions for simultaneous determinations of the Rydberg constant R_∞ and the proton radius r_p assuming either the SM-only hypothesis (gray) or including putative NP contributions from a 400 keV Higgs portal scalar (blue) or 300 keV ULD scalar (purple). The solid lines use DATA22 dataset, the dashed (dotted) lines the CODATA18 dataset with (without) errors inflated by expansion factors. Both R_∞ and r_p are shown in terms of normalized deviations from the central values of the CODATA 2018 analysis, Ref. [1].

tion are not new, our analysis shows that all tensions can be significantly ameliorated when including NP interactions due to a single light scalar. The favored NP mass is close to the (inverse) Bohr radius of muonic atoms, $a_0^{-1} \times m_\mu/m_e \sim$ MeV, due to the large muon-electron coupling ratio in these models, contrasting with scalars having weaker or vanishing coupling to muons (see Sec. S4 and [20]). However, other constraints require the scalar to have rather a nontrivial pattern of couplings, see Sec S4 A. For ULD scalar the E137 [54, 55] bounds are evaded since ϕ decays predominantly to an invisible dark sector. Since ϕ couples to up quarks and not directly to heavy quarks and gluons, the bound from NA62 search for $K^+ \to \pi^+\phi$ [52] is weakened [51, 56]. Finally, the minimal ULD model induces a too large contribution to $(g-2)_\mu$, however this can be suppressed in less minimal versions with a custodial symmetry [57].

The presence of NP also impacts the determination of the fundamental constants in the SM. Figure 3 shows the 68 % CL determination of r_p and R_∞, subtracting the CODATA 2018 recommended values and normalizing to respective errors. The SM-parameter uncertainties increase in the presence of NP and the central values shift outside the nominal SM ellipse, shown explicitly in Fig. 3 for the Higgs portal and ULD scalar model. Because of the degeneracy with the photon the uncertainty on α in the dark photon model increases as $1/m_\phi^2$ for masses below 10 eV (see Fig. S7) and eventually becomes comparable to α itself for $m_\phi \sim 0.1$ meV, while $\alpha+\alpha_\phi$ remains well constrained.

VI. CONCLUSIONS

Extracting bounds on light NP from a global fit to spectroscopic and other precision data requires both SM and NP parameters to be determined simultaneously. The possibility of NP contributions changes the extracted allowed ranges of SM parameters, a change that can be quite substantial, see Fig. 3. Furthermore, we provided a prescription to consistently include NP corrections from light vectors. It requires calculations of NP contribution only at leading order, and recovers the expected degeneracy between dark photon and QED in the massless mediator limit.

At present, spectroscopic data show tensions that could either be due to unknown or under-appreciated systematics, or to light NP. We showed that the $\sim 4\sigma$ anomaly in data can be explained by a flavor non-universal light scalar model.

ACKNOWLEDGMENTS

We would like to thank Dmitry Budker and Gilad Perez for useful discussions and comments on the manuscript. The work of CD is supported by the CNRS IRP NewSpec. The work of TK is supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Early-Career Scientists (Grant No. 19K14706) and the JSPS Core-to-Core Program (Grant No. JPJSCCA20200002). The work of YS is supported by grants from the NSF-BSF (No. 2018683), the ISF (No. 482/20), the BSF (No. 2020300) and by the Azrieli foundation. JZ acknowledges support in part by the DOE grant de-sc0011784. This work was also supported in part by the European Union's Horizon 2020 research and innovation programme, project STRONG2020, under grant agreement No. 824093.

[1] E. Tiesinga, P. J. Mohr, D. B. Newell, and B. N. Taylor, "CODATA recommended values of the fundamental physical constants: 2018*," Rev. Mod. Phys. **93** (2021) 025010.

[2] D. Hanneke, S. Fogwell, and G. Gabrielse, "New Measurement of the Electron Magnetic Moment and the Fine Structure Constant," Phys. Rev. Lett. **100** (2008) 120801 [arXiv:0801.1134].

[3] **Muon g-2** Collaboration, "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm," Phys. Rev. Lett. **126** (2021) 141801 [arXiv:2104.03281].

[4] **ACME** Collaboration, "Improved limit on the electric dipole moment of the electron," Nature **562** (2018) 355–360.

[5] J. Jaeckel and S. Roy, "Spectroscopy as a test of Coulomb's law: A Probe of the hidden sector," Phys. Rev. D **82** (2010) 125020 [arXiv:1008.3536].

[6] S. G. Karshenboim, "Constraints on a long-range spin-independent interaction from precision atomic physics," Phys. Rev. D **82** (2010) 073003 [arXiv:1005.4872].

[7] C. Delaunay, R. Ozeri, G. Perez, and Y. Soreq, "Probing Atomic Higgs-like Forces at the Precision Frontier," Phys. Rev. D **96** (2017) 093001 [arXiv:1601.05087].

[8] J. C. Berengut *et al.*, "Probing New Long-Range Interactions by Isotope Shift Spectroscopy," Phys. Rev. Lett. **120** (2018) 091801 [arXiv:1704.05068].

[9] C. Frugiuele, J. Pérez-Ríos, and C. Peset, "Current and future perspectives of positronium and muonium spectroscopy as dark sectors probe," Phys. Rev. D **100** (2019) 015010 [arXiv:1902.08585].

[10] C. Frugiuele and C. Peset, "Muonic vs electronic dark forces: a complete EFT treatment for atomic spectroscopy," JHEP **05** (2022) 002 [arXiv:2107.13512].

[11] E. J. Salumbides, *et al.*, "Bounds on fifth forces from precision measurements on molecules," Phys. Rev. D **87** (2013) 112008 [arXiv:1304.6560].

[12] M. Borkowski, *et al.*, "Weakly bound molecules as sensors of new gravitylike forces," Sci. Rep. **9** (2019) 14807 [arXiv:1612.03842].

[13] S. Alighanbari, G. S. Giri, F. L. Constantin, V. I. Korobov, and S. Schiller, "Precise test of quantum electrodynamics and determination of fundamental constants with HD+ ions," Nature **581** (2020) 152–158.

[14] M. Germann, *et al.*, "Three-body QED test and fifth-force constraint from vibrations and rotations of HD^+," Phys. Rev. Res. **3** (2021) L022028.

[15] M. S. Safronova, *et al.*, "Search for New Physics with Atoms and Molecules," Rev. Mod. Phys. **90** (2018) 025008 [arXiv:1710.01833].

[16] M. P. A. Jones, R. M. Potvliege, and M. Spannowsky, "Probing new physics using Rydberg states of atomic hydrogen," Phys. Rev. Res. **2** (2020) 013244 [arXiv:1909.09194].

[17] R. Pohl *et al.*, "The size of the proton," Nature **466** (2010) 213–216.

[18] A. Antognini *et al.*, "Proton Structure from the Measurement of $2S-2P$ Transition Frequencies of Muonic Hydrogen," Science **339** (2013) 417–420.

[19] J.-P. Karr, D. Marchand, and E. Voutier, "The proton size," Nat. Rev. Phys. **2** (2020) 601–614.

[20] A. D. Brandt, *et al.*, "Measurement of the 2S1/2-8D5/2 Transition in Hydrogen," Phys. Rev. Lett. **128** (2022) 023001 [arXiv:2111.08554].

[21] A. Grinin, *et al.*, "Two-photon frequency comb spectroscopy of atomic hydrogen," Science **370** (2020) 1061.

[22] B. Holdom, "Two U(1)'s and Epsilon Charge Shifts," Phys. Lett. B **166** (1986) 196–198.

[23] A. Davidson, "$B-L$ as the fourth color within an $SU(2)_L \times U(1)_R \times U(1)$ model," Phys. Rev. D **20** (1979) 776.

[24] R. E. Marshak and R. N. Mohapatra, "Quark - Lepton Symmetry and $B-L$ as the U(1) Generator of the Electroweak Symmetry Group," Phys. Lett. B **91** (1980) 222–224.

[25] B. Patt and F. Wilczek, "Higgs-field portal into hidden sectors." hep-ph/0605188.

[26] D. O'Connell, M. J. Ramsey-Musolf, and M. B. Wise, "Minimal Extension of the Standard Model Scalar Sector," Phys. Rev. D **75** (2007) 037701 [hep-ph/0611014].

[27] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, "Remarks on Higgs Boson Interactions with Nucleons," Phys. Lett. B **78** (1978) 443–446.

[28] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, "Dark matter direct detection rate in a generic model with micrOMEGAs 2.2," Comput. Phys. Commun. **180** (2009) 747–767 [arXiv:0803.2360].

[29] P. Junnarkar and A. Walker-Loud, "Scalar strange content of the nucleon from lattice QCD," Phys. Rev. D **87** (2013) 114510 [arXiv:1301.1114].

[30] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, "micrOMEGAs_3: A program for calculating dark matter observables," Comput. Phys. Commun. **185** (2014) 960–985 [arXiv:1305.0237].

[31] F. Bishara, J. Brod, B. Grinstein, and J. Zupan, "From quarks to nucleons in dark matter direct detection," JHEP **11** (2017) 059 [arXiv:1707.06998].

[32] F. Bishara, J. Brod, B. Grinstein, and J. Zupan, "DirectDM: a tool for dark matter direct detection." arXiv:1708.02678.

[33] S. Alighanbari, G. S. Giri, F. L. Constantin, V. I. Korobov, and S. Schiller, "Precise test of quantum electrodynamics and determination of fundamental constants with HD^+ ions," Nature **581** (2020) 152–158.

[34] S. Patra, *et al.*, "Proton-electron mass ratio from laser spectroscopy of HD^+ at the part-per-trillion level," Science **369** (2020) 1238.

[35] I. V. Kortunov, *et al.*, "Proton-electron mass ratio by high-resolution optical spectroscopy of ion ensembles in the resolved-carrier regime," Nature Phys. **17** (2021) 569 [arXiv:2103.11741].

[36] M. Hori, *et al.*, "Two-photon laser spectroscopy of antiprotonic helium and the antiproton-to-electron mass ratio," Nature **475** (2011) 484.

[37] M. Hori, *et al.*, "Buffer-gas cooling of antiprotonic helium to 1.5 to 1.7 K, and antiproton-to-electron mass ratio," Science **354** (2016) 610.

[38] P. J. Mohr and B. N. Taylor, "CODATA Recommended Values of the Fundamental Physical Constants: 1998," Rev. Mod. Phys. **72** (2000) 351–495.

[39] V. I. Korobov, L. Hilico, and J. P. Karr, "Fundamental Transitions and Ionization Energies of the Hydrogen Molecular Ions with Few ppt Uncertainty," Phys. Rev. Lett. **118** (2017) 233001 [arXiv:1703.07972].

[40] V. I. Korobov and J. P. Karr, "Rovibrational spin-averaged transitions in the hydrogen molecular ions," Phys. Rev. A **104** (2021) 032806.

[41] V. I. Korobov, "Calculation of transitions between metastable states of antiprotonic helium including

relativistic and radiative corrections of order $R_\infty\alpha^4$," Phys. Rev. A **77** (2008) 042506.
[42] V. I. Korobov, "Bethe logarithm for resonant states: Antiprotonic helium," Phys. Rev. A **89** (2014) 014501 [arXiv:1312.4299].
[43] V. I. Korobov, L. Hilico, and J.-P. Karr, "mα^7 -Order Corrections in the Hydrogen Molecular Ions and Antiprotonic Helium," Phys. Rev. Lett. **112** (2014) 103003.
[44] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum mechanics; 1st ed.* Wiley, New York, NY, 1977. Trans. of : Mécanique quantique. Paris : Hermann, 1973.
[45] A. Messiah, *Quantum mechanics*. Dover Publications, 2014.
[46] V. I. Korobov, "Coulomb three-body bound-state problem: Variational calculations of nonrelativistic energies," Phys. Rev. A **61** (2000) 064503.
[47] R. Jackiw and S. Weinberg, "Weak interaction corrections to the muon magnetic moment and to muonic atom energy levels," Phys. Rev. D **5** (1972) 2396–2398.
[48] F. Jegerlehner and A. Nyffeler, "The Muon g-2," Phys. Rept. **477** (2009) 1–110 [arXiv:0902.3360].
[49] V. Debierre, C. H. Keitel, and Z. Harman, "Fifth-force search with the bound-electron g factor," Phys. Lett. B **807** (2020) 135527.
[50] G. Raffelt, "Limits on a CP-violating scalar axion-nucleon interaction," Phys. Rev. D **86** (2012) 015001 [arXiv:1205.1776].
[51] B. Batell, A. Freitas, A. Ismail, and D. Mckeen, "Probing Light Dark Matter with a Hadrophilic Scalar Mediator," Phys. Rev. D **100** (2019) 095020 [arXiv:1812.05103].
[52] **NA62** Collaboration, "Search for a feebly interacting particle X in the decay $K^+ \to \pi^+ X$," JHEP **03** (2021) 058 [arXiv:2011.11329].
[53] E. Hardy and R. Lasenby, "Stellar cooling bounds on new light particles: plasma mixing effects," JHEP **02** (2017) 033 [arXiv:1611.05852].
[54] J. D. Bjorken, *et al.*, "Search for Neutral Metastable Penetrating Particles Produced in the SLAC Beam Dump," Phys. Rev. D **38** (1988) 3375.
[55] Y.-S. Liu, D. McKeen, and G. A. Miller, "Validity of the Weizsäcker-Williams approximation and the analysis of beam dump experiments: Production of a new scalar boson," Phys. Rev. D **95** (2017) 036010 [arXiv:1609.06781].
[56] C. Delaunay, *et al.*, "work in progress.".
[57] R. Balkin, *et al.*, "Custodial symmetry for muon g-2," Phys. Rev. D **104** (2021) 053009 [arXiv:2104.08289].
[58] R. Bouchendira, P. Cladé, S. Guellati-Khélifa, F. Nez, and F. Biraben, "New Determination of the Fine Structure Constant and Test of the Quantum Electrodynamics," Phys. Rev. Lett. **106** (2011) 080801.
[59] R. H. Parker, C. Yu, W. Zhong, B. Estey, and H. Müller, "Measurement of the fine-structure constant as a test of the Standard Model," Science **360** (2018) 191 [arXiv:1812.04130].
[60] W. J. Huang, *et al.*, "The AME2016 atomic mass evaluation (I). Evaluation of input data; and adjustment procedures," Chin. Phys. C **41** (2017) 030002.
[61] M. Wang, *et al.*, "The AME2016 atomic mass evaluation (II). Tables, graphs and references," Chin. Phys. C **41** (2017) 030003.
[62] A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team. "NIST Atomic Spectra Database (ver. 5.9)," Available: https://physics.nist.gov/asd. National Institute of Standards and Technology, Gaithersburg, MD., 2021.
[63] F. Heiße, *et al.*, "High-Precision Measurement of the Proton's Atomic Mass," Phys. Rev. Lett. **119** (2017) 033001.
[64] S. L. Zafonte and R. S. V. Dyck, "Ultra-precise single-ion atomic mass measurements on deuterium and helium-3," Metrologia **52** (2015) 280–290.
[65] V. A. Yerokhin, K. Pachucki, and V. Patkóš, "Theory of the Lamb Shift in Hydrogen and Light Hydrogen-Like Ions," Ann. Phys. **531** (2019) 1800324.
[66] S. G. Karshenboim, A. Ozawa, V. A. Shelyuto, R. Szafron, and V. G. Ivanov, "The Lamb shift of the 1s state in hydrogen: Two-loop and three-loop contributions," Phys. Lett. B **795** (2019) 432.
[67] R. Szafron, E. Y. Korzinin, V. A. Shelyuto, V. G. Ivanov, and S. G. Karshenboim, "Virtual Delbrück scattering and the Lamb shift in light hydrogenlike atoms," Phys. Rev. A **100** (2019) 032507.
[68] S. G. Karshenboim and V. A. Shelyuto, "Three-loop radiative corrections to the 1s Lamb shift in hydrogen," Phys. Rev. A **100** (2019) 032513.
[69] L. Morel, Z. Yao, P. Cladé, and S. Guellati-Khélifa, "Determination of the fine-structure constant with an accuracy of 81 parts per trillion," Nature **588** (2020) 61–65.
[70] X. Fan, T. G. Myers, B. A. D. Sukra, and G. Gabrielse, "Measurement of the Electron Magnetic Moment." arXiv:2209.13084.
[71] A. Czarnecki, J. Piclum, and R. Szafron, "Logarithmically enhanced Euler-Heisenberg Lagrangian contribution to the electron gyromagnetic factor," Phys. Rev. A **102** (2020) 050801.
[72] W. J. Huang, M. Wang, F. G. Kondev, G. Audi, and S. Naimi, "The AME 2020 atomic mass evaluation (I). Evaluation of input data, and adjustment procedures," Chin. Phys. C **45** (2021) 030002.
[73] M. Wang, W. J. Huang, F. G. Kondev, G. Audi, and S. Naimi, "The AME 2020 atomic mass evaluation (II). Tables, graphs and references. Available: https://www-nds.iaea.org/amdc/," Chin. Phys. C **45** (2021) 030003.
[74] F. Heiße, *et al.*, "High-precision mass spectrometer for light ions," Phys. Rev. A **100** (2019) 022518 [arXiv:2002.11389].
[75] S. Rau, *et al.*, "Penning-trap mass measurements of the deuteron and the HD^+ molecular ion," Nature **585** (2020) 43–47 [arXiv:2203.05971].
[76] D. J. Fink and E. G. Myers, "Deuteron-to-Proton Mass Ratio from the Cyclotron Frequency Ratio of H_2^+ to D^+ with H_2^+ in a Resolved Vibrational State," Phys. Rev. Lett. **124** (2020) 013001.
[77] D. J. Fink and E. G. Myers, "Deuteron-To-Proton Mass Ratio from Simultaneous Measurement of the Cyclotron Frequencies of H_2^+ and D^+," Phys. Rev. Lett. **127** (2021) 243001.
[78] J. J. Krauth, *et al.*, "Measuring the alpha-particle charge radius with muonic helium-4 ions," Nature **589** (2021) 527.
[79] J. Krauth, *The Lamb shift of the muonic helium-3 ion*

and the helion charge radius. PhD thesis, Ludwig-Maximilians University, Munich, 2017.
[80] W. Xiong *et al.*, "A small proton charge radius from an electron–proton scattering experiment," Nature **575** (2019) 147–150.
[81] Z.-F. Cui, D. Binosi, C. D. Roberts, and S. M. Schmidt, "Fresh Extraction of the Proton Charge Radius from Electron Scattering," Phys. Rev. Lett. **127** (2021) 092001 [arXiv:2102.01180].
[82] S. Schiller and V. Korobov, "Tests of time independence of the electron and nuclear masses with ultracold molecules," Phys. Rev. A **71** (2005) 032505.
[83] V. I. Korobov, "Bethe logarithm for resonant states: Antiprotonic helium," Phys. Rev. A **89** (2014) 014501.
[84] D. T. Aznabayev, A. K. Bekbaev, and V. I. Korobov, "Leading-order relativistic corrections to the rovibrational spectrum of $H_2{}^+$ and HD^+ molecular ions," Phys. Rev. A **99** (2019) 012501.
[85] V. I. Korobov, "Metastable states in the antiprotonic helium atom decaying via Auger transitions," Phys. Rev. A **67** (2003) 062501.
[86] V. I. Korobov and Z.-X. Zhong, "Bethe logarithm for the H_2^+ and HD^+ molecular ions," Phys. Rev. A **86** (2012) 044501.
[87] V. I. Korobov, "Leading-order relativistic and radiative corrections to the rovibrational spectrum of H2+ and HD+ molecular ions," Phys. Rev. A **74** (2006) 052506.
[88] V. I. Korobov, "Calculation of transitions between metastable states of antiprotonic helium including relativistic and radiative corrections of order $R_\infty\alpha^4$," Phys. Rev. A **77** (2008) 042506.
[89] V. I. Korobov and T. Tsogbayar, "Relativistic corrections of order $m\alpha^6$ to the two-centre problem," J. Phys. B: At. Mol. Opt. Phys. **40** (2007) 2661.
[90] K. Pachucki, "Recoil Effects in Positronium Energy Levels to Order α^6," Phys. Rev. Lett. **79** (1997) 4120.
[91] K. Pachucki, "Radiative recoil correction to the Lamb shift," Phys. Rev. A **52** (1995) 1079.
[92] A. Czarnecki and K. Melnikov, "Expansion of Bound-State Energies in Powers of m/M," Phys. Rev. Lett. **87** (2001) 013001.
[93] T. Tsogbayar and V. I. Korobov, "Relativistic correction to the $1s\sigma$ and $2p\sigma$ electronic states of the H_2^+ molecular ion and the moleculelike states of the antiprotonic helium $He^+\bar{p}$," J. Chem. Phys. **125** (2006) 024308.
[94] V. I. Korobov, L. Hilico, and J.-P. Karr, "Theoretical transition frequencies beyond 0.1 ppb accuracy in H_2^+, HD^+, and antiprotonic helium," Phys. Rev. A **89** (2014) 032511.
[95] V. I. Korobov, L. Hilico, and J.-P. Karr, "Bound-state QED calculations for antiprotonic helium," Hyperfine Interactions **233** (2015) 75.
[96] J. P. Karr, L. Hilico, and V. I. Korobov, "One-loop vacuum polarization at $m\alpha^7$ and higher orders for three-body molecular systems," Phys. Rev. A **95** (2017) 042514.
[97] A. Crivellin, M. Hoferichter, and M. Procura, "Accurate evaluation of hadronic uncertainties in spin-independent WIMP-nucleon scattering: Disentangling two- and three-flavor effects," Phys. Rev. D **89** (2014) 054021 [arXiv:1312.4951].
[98] **Flavour Lattice Averaging Group** Collaboration, "FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)," Eur. Phys. J. C **80** (2020) 113 [arXiv:1902.08191].
[99] **xQCD** Collaboration, "πN and strangeness sigma terms at the physical point with chiral fermions," Phys. Rev. D **94** (2016) 054503 [arXiv:1511.09089].
[100] **MILC** Collaboration, "Intrinsic strangeness and charm of the nucleon using improved staggered fermions," Phys. Rev. D **88** (2013) 054503 [arXiv:1204.3866].
[101] S. Durr *et al.*, "Lattice computation of the nucleon scalar quark contents at the physical point," Phys. Rev. Lett. **116** (2016) 172001 [arXiv:1510.08013].
[102] S. Durr *et al.*, "Sigma term and strangeness content of octet baryons," Phys. Rev. D **85** (2012) 014509 [arXiv:1109.4265]. [Erratum: Phys.Rev.D 93, 039905 (2016)].
[103] C. Alexandrou, *et al.*, "Nucleon axial, tensor, and scalar charges and σ-terms in lattice QCD," Phys. Rev. D **102** (2020) 054517 [arXiv:1909.00485].
[104] S. Borsanyi, *et al.*, "Ab-initio calculation of the proton and the neutron's scalar couplings for new physics searches." arXiv:2007.03319.
[105] W. Altmannshofer, S. Gori, A. L. Kagan, L. Silvestrini, and J. Zupan, "Uncovering Mass Generation Through Higgs Flavor Violation," Phys. Rev. D **93** (2016) 031301 [arXiv:1507.07927].
[106] F. J. Botella, G. C. Branco, M. Nebot, and M. N. Rebelo, "Flavour Changing Higgs Couplings in a Class of Two Higgs Doublet Models," Eur. Phys. J. C **76** (2016) 161 [arXiv:1508.05101].
[107] D. Ghosh, R. S. Gupta, and G. Perez, "Is the Higgs Mechanism of Fermion Mass Generation a Fact? A Yukawa-less First-Two-Generation Model," Phys. Lett. B **755** (2016) 504–508 [arXiv:1508.01501].
[108] **CMS** Collaboration, "Evidence for Higgs boson decay to a pair of muons," JHEP **01** (2021) 148 [arXiv:2009.04363].
[109] N. Bar, K. Blum, and G. D'Amico, "Is there a supernova bound on axions?" Phys. Rev. D **101** (2020) 123025 [arXiv:1907.05020].
[110] P. S. B. Dev, R. N. Mohapatra, and Y. Zhang, "Revisiting supernova constraints on a light CP-even scalar," JCAP **08** (2020) 003 [arXiv:2005.00490]. [Erratum: JCAP 11, E01 (2020)].
[111] H. Gao and M. Vanderhaeghen, "The proton charge radius," Rev. Mod. Phys. **94** (2022) 015002.

Self-consistent extraction of spectroscopic bounds on light new physics
Supplemental Material

Cédric Delaunay, Jean-Philippe Karr, Teppei Kitahara, Jeroen C. J. Koelemeij, Yotam Soreq, and Jure Zupan

In the supplemental material we give further details on the linearized least squares method (Sec. S1), the CODATA18 dataset (Sec. S2), the DATA22 dataset (Sec. S3), the new physics benchmarks models (Sec. S4), new physics improvements relative to the SM (Sec. S5), the extraction of fundamental constants in the presence of NP (Sec. S6), and the $m_\phi \to 0$ limit (Sec. S7).

S1. THE LINEARIZED LEAST SQUARES METHOD

We use the method of linearized least squares employed by the CODATA to determine both the fundamental and NP constants. For completeness we summarize below briefly the fitting procedure, while further details can be found in Appendix E of Ref. [38].

The N input data w_i are either measured quantities or estimated errors on theoretical calculations (for these the central values are taken to be zero), and are listed in Tables S1-S4 and S6-S7, S9, see also Sections S2 and S3 for a detailed discussion of the two datasets we use. The predictions for w_i are functions of M fitted parameters (adjusted constants) z_j,

$$w_i \doteq f_i(z) \equiv f_i(z_1, z_2, \ldots, z_M)\,. \tag{S1}$$

The dotted equality sign means that the left and right hand sides are not equal in general, since the set of equations is usually overdetermined ($N \geq M$), but should agree within experimental errors, if the assumed physics model and measurements are correct. Most of the f_i are nonlinear functions, which we linearize, *i.e.*, Taylor expand around some initial values s_j, chosen close enough to the expected values of z_j such that the higher-order terms can be neglected,

$$w_i \doteq f_i(s) + \sum_{j=1}^{M} \frac{\partial f_i(s)}{\partial z_j}(z_j - s_j) + \cdots\,. \tag{S2}$$

Linearization allows for more efficient determination of adjusted constants. Observational equations (S1) are now a set of overdetermined linear equations. In matrix notation these are

$$Y \doteq AX\,, \tag{S3}$$

where Y is an N-dimensional vector of $y_i \equiv w_i - f_i(s)$, A is an $N \times M$ matrix with elements $a_{ij} \equiv \partial f_i(s)/\partial z_j$, while X is an M-dimensional vector of $x_j \equiv z_j - s_j$. The best fit value of X is obtained by minimizing the χ^2 function,

$$\chi^2 = (Y - AX)^T V^{-1} (Y - AX)\,, \tag{S4}$$

where V is the $N \times N$ covariance matrix constructed from variances $u_{ii} = u_i^2$ and covariances $u_{ij} \equiv u(w_i, w_j)$ for inputs $w_{i,j}$ (here $u_i \equiv u(w_i)$ is the standard uncertainty). Updated correlation coefficients for the DATA22 dataset are listed in Tables S8, S10. The result of the fit, vector $\hat{X}$ with components $\hat{x}_j$ that minimizes χ^2, is given by

$$\hat{X} = G A^T V^{-1} Y\,, \tag{S5}$$

where $G \equiv (A^T V^{-1} A)^{-1}$ is the covariance matrix of $\hat{X}$, *i.e.*, an $M \times M$ matrix that captures the uncertainties and correlations of the adjusted constants z_j.

The solution $\hat{X}$ is only an approximation of the exact solution to the nonlinear problem. More precise values of the adjusted fundamental constants are obtained by iterating the above procedure using $s_j + \hat{x}_j$ as starting values, until the new $\hat{x}_j$'s are negligibly small compared to their estimated uncertainties. Similarly to the CODATA, we use the convergence condition

$$\sum_{j=1}^{M} \frac{\hat{x}_j^2}{u^2(\hat{x}_j)} < 10^{-20}\,. \tag{S6}$$

In practice, this condition is satisfied after a few iterations.

Label	Input datum	Value (kHz)	Rel. uncert.
A1	$\nu_{\rm H}(2S_{1/2}-4S_{1/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	4 797 338(10)	2.1×10^{-6}
A2	$\nu_{\rm H}(2S_{1/2}-4D_{5/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	6 490 144(24)	3.7×10^{-6}
A3	$\nu_{\rm D}(2S_{1/2}-4S_{1/2})-\frac{1}{4}\nu_{\rm D}(1S_{1/2}-2S_{1/2})$	4 801 693(20)	4.2×10^{-6}
A4	$\nu_{\rm D}(2S_{1/2}-4D_{5/2})-\frac{1}{4}\nu_{\rm D}(1S_{1/2}-2S_{1/2})$	6 494 841(41)	6.3×10^{-6}
A5	$\nu_{\rm D}(1S_{1/2}-2S_{1/2})-\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	670 994 334.606(15)	2.2×10^{-11}
A6	$\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	2 466 061 413 187.035(10)	4.2×10^{-15}
A7	$\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	2 466 061 413 187.018(11)	4.4×10^{-15}
A8	$\nu_{\rm H}(1S_{1/2}-3S_{1/2})$	2 922 743 278 659(17)	5.8×10^{-12}
A9	$\nu_{\rm H}(2S_{1/2}-4P)$	616 520 931 626.8(2.3)	3.7×10^{-12}
A10	$\nu_{\rm H}(2S_{1/2}-8S_{1/2})$	770 649 350 012.0(8.6)	1.1×10^{-11}
A11	$\nu_{\rm H}(2S_{1/2}-8D_{3/2})$	770 649 504 450.0(8.3)	1.1×10^{-11}
A12	$\nu_{\rm H}(2S_{1/2}-8D_{5/2})$	770 649 561 584.2(6.4)	8.3×10^{-12}
A13	$\nu_{\rm D}(2S_{1/2}-8S_{1/2})$	770 859 041 245.7(6.9)	8.9×10^{-12}
A14	$\nu_{\rm D}(2S_{1/2}-8D_{3/2})$	770 859 195 701.8(6.3)	8.2×10^{-12}
A15	$\nu_{\rm D}(2S_{1/2}-8D_{5/2})$	770 859 252 849.5(5.9)	7.7×10^{-12}
A16	$\nu_{\rm H}(2S_{1/2}-12D_{3/2})$	799 191 710 472.7(9.4)	1.2×10^{-11}
A17	$\nu_{\rm H}(2S_{1/2}-12D_{5/2})$	799 191 727 403.7(7.0)	8.7×10^{-12}
A18	$\nu_{\rm D}(2S_{1/2}-12D_{3/2})$	799 409 168 038.0(8.6)	1.1×10^{-11}
A19	$\nu_{\rm D}(2S_{1/2}-12D_{5/2})$	799 409 184 966.8(6.8)	8.5×10^{-12}
A20	$\nu_{\rm H}(2S_{1/2}-6S_{1/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-3S_{1/2})$	4 197 604(21)	4.9×10^{-6}
A21	$\nu_{\rm H}(2S_{1/2}-6D_{5/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-4S_{1/2})$	4 699 099(10)	2.2×10^{-6}
A22	$\nu_{\rm H}(1S_{1/2}-3S_{1/2})$	2 922 743 278 678(13)	4.4×10^{-12}
A23	$\nu_{\rm H}(1S_{1/2}-3S_{1/2})$	2 922 743 278 671.5(2.6)	8.9×10^{-13}
A24	$\nu_{\rm H}(2S_{1/2}-4P_{1/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	4 664 269(15)	3.2×10^{-6}
A25	$\nu_{\rm H}(2S_{1/2}-4P_{3/2})-\frac{1}{4}\nu_{\rm H}(1S_{1/2}-2S_{1/2})$	6 035 373(10)	1.7×10^{-6}
A26	$\nu_{\rm H}(2S_{1/2}-2P_{3/2})$	9 911 200(12)	1.2×10^{-6}
A27	$\nu_{\rm H}(2P_{1/2}-2S_{1/2})$	1 057 862(20)	1.9×10^{-5}
A28	$\nu_{\rm H}(2P_{1/2}-2S_{1/2})$	1 057 845.0(9.0)	8.5×10^{-6}
A29	$\nu_{\rm H}(2P_{1/2}-2S_{1/2})$	1 057 829.8(3.2)	3.0×10^{-6}

TABLE S1. Input data of the electronic hydrogen and deuterium measurements for the CODATA18 dataset, taken from Table X of Ref. [1].

Finally, we define for each input datum a normalized residual

$$R_i \equiv \frac{w_i - f_i(\hat{z})}{u_i}\,, \tag{S7}$$

where $\hat{z}$ denotes the final value of the adjusted constants. Ignoring correlations, the contribution of each input datum i to the minimal χ^2 is R_i^2.

S2. THE CODATA18 DATASET

The CODATA 2018 dataset contains all the inputs from Ref. [1] related to the determination of the Rydberg constant R_∞, the proton and deuteron radii, r_p and r_d, respectively, the fine-structure constant α, and the relative atomic masses of the electron, proton, and deuteron: $A_{\rm r}(e)$, $A_{\rm r}(p)$ and $A_{\rm r}(d)$, respectively. The other observables and parameters included in the CODATA 2018 adjustment are very weakly correlated with the selected data, and can be neglected for our purposes.

The selected data include atomic transitions and non-spectroscopic observables, which depend on the main fundamental constants as follows. The inputs primarily used to determine R_∞, r_p, and r_d are:

A) measurements of transition frequencies in electronic hydrogen and deuterium, labeled as Ai, $i = 1, \ldots, 29$, and listed in Table S1,

Label	Input datum	Value (kHz)	Rel. uncert.
B1	$\delta_{\rm H}(1S_{1/2})/h$	0.0(1.6)	4.9×10^{-13}
B2	$\delta_{\rm H}(2S_{1/2})/h$	0.00(20)	2.4×10^{-13}
B3	$\delta_{\rm H}(3S_{1/2})/h$	0.000(59)	1.6×10^{-13}
B4	$\delta_{\rm H}(4S_{1/2})/h$	0.000(25)	1.2×10^{-13}
B5	$\delta_{\rm H}(6S_{1/2})/h$	0.000(12)	1.3×10^{-13}
B6	$\delta_{\rm H}(8S_{1/2})/h$	0.0000(51)	9.9×10^{-14}
B7	$\delta_{\rm H}(2P_{1/2})/h$	0.0000(39)	4.8×10^{-15}
B8	$\delta_{\rm H}(4P_{1/2})/h$	0.0000(16)	7.6×10^{-15}
B9	$\delta_{\rm H}(2P_{3/2})/h$	0.0000(39)	4.8×10^{-15}
B10	$\delta_{\rm H}(4P_{3/2})/h$	0.0000(16)	7.6×10^{-15}
B11	$\delta_{\rm H}(8D_{3/2})/h$	0.000 000(13)	2.6×10^{-16}
B12	$\delta_{\rm H}(12D_{3/2})/h$	0.000 0000(40)	1.8×10^{-16}
B13	$\delta_{\rm H}(4D_{5/2})/h$	0.000 00(17)	8.2×10^{-16}
B14	$\delta_{\rm H}(6D_{5/2})/h$	0.000 000(58)	6.3×10^{-16}
B15	$\delta_{\rm H}(8D_{5/2})/h$	0.000 000(22)	4.2×10^{-16}
B16	$\delta_{\rm H}(12D_{5/2})/h$	0.000 0000(64)	2.8×10^{-16}
B17	$\delta_{\rm D}(1S_{1/2})/h$	0.0(1.5)	4.5×10^{-13}
B18	$\delta_{\rm D}(2S_{1/2})/h$	0.00(18)	2.2×10^{-13}
B19	$\delta_{\rm D}(4S_{1/2})/h$	0.000(23)	1.1×10^{-13}
B20	$\delta_{\rm D}(8S_{1/2})/h$	0.0000(49)	9.6×10^{-14}
B21	$\delta_{\rm D}(8D_{3/2})/h$	0.000 0000(95)	1.8×10^{-16}
B22	$\delta_{\rm D}(12D_{3/2})/h$	0.000 0000(28)	1.2×10^{-16}
B23	$\delta_{\rm D}(4D_{5/2})/h$	0.000 00(15)	7.5×10^{-16}
B24	$\delta_{\rm D}(8D_{5/2})/h$	0.000 000(19)	3.8×10^{-16}
B25	$\delta_{\rm D}(12D_{3/2})/h$	0.000 0000(58)	2.5×10^{-16}

TABLE S2. Input data for the additive energy corrections accounting for missing contributions to the theoretical description of H and D energy levels for the CODATA18 dataset, taken from Table VIII of Ref. [1]. Relative uncertainties are with respect to the binding energy.

Label	Input datum	Value	Rel. uncert.
C1	$E_{\rm LS}(\mu{\rm H})$	202.3706(23) meV	1.1×10^{-5}
C2	$E_{\rm LS}(\mu{\rm D})$	202.8785(34) meV	1.7×10^{-5}
C7	$\delta E_{\rm LS}(\mu{\rm H})$	0.0000(129) meV	6.4×10^{-5}
C8	$\delta E_{\rm LS}(\mu{\rm D})$	0.0000(210) meV	1.0×10^{-4}
C9	r_p	0.880(20) fm	2.3×10^{-2}
C10	r_d	2.111(19) fm	9.0×10^{-3}

TABLE S3. Input data of the muonic hydrogen and deuterium Lamb shifts (LS), as well as the proton and deuteron charge radii from the electron-proton and electron-deuteron scatterings for the CODATA18 dataset, taken from Table XVIII of Ref. [1]. For the additive corrections C7 and C8, the relative uncertainty is with respect to the value of the corresponding theoretical quantity.

B) the additive theory uncertainties on theory predictions for the relevant energy levels, labeled by Bi, $i = 1, \ldots, 25$, listed in Table S2, and

C) the inputs for muonic hydrogen and deuterium and electron scattering, labeled by Ci, with $i = 1, 2, 7, 8$ and $i = 9, 10$, respectively (see Table S3).

The dependence of an electronic transition frequency $\nu^i_{\rm SM}$ on physical constants is described, within the SM, by the

Label	Input datum	Value	Rel. uncert.
D1	$a_e \equiv \frac{1}{2}(g-2)_e$	$1.159\,652\,180\,73(28)\ \times 10^{-3}$	2.4×10^{-10}
D2	δ_e	$0.000(18) \times 10^{-12}$	1.5×10^{-11}
D3	$h/m_{\mathrm{Rb}}(^{87}\mathrm{Rb})$	$4.591\,359\,272\,9(57)\ \times 10^{-9}\ \mathrm{m^2 s^{-1}}$	1.2×10^{-9}
D4	$h/m_{\mathrm{Cs}}(^{133}\mathrm{Cs})$	$3.002\,369\,472\,1(12)\ \times 10^{-9}\ \mathrm{m^2 s^{-1}}$	4.0×10^{-10}
D5	$A_{\mathrm{r}}(^{87}\mathrm{Rb})$	$86.909\,180\,531\,2(65)$	7.4×10^{-11}
D6	$A_{\mathrm{r}}(^{133}\mathrm{Cs})$	$132.905\,451\,961\,0(86)$	6.5×10^{-11}
D7	$\omega_{\mathrm{s}}/\omega_{\mathrm{c}}(^{12}\mathrm{C}^{5+})$	$4376.210\,500\,87(12)$	2.8×10^{-11}
D8	$\Delta E_{\mathrm{B}}(^{12}\mathrm{C}^{5+})/hc$	$43.563\,233(25)\ \times 10^{7}\ \mathrm{m^{-1}}$	5.8×10^{-7}
D9	δ_{C}	$0.0(2.5)\ \times 10^{-11}$	1.3×10^{-11}
D10	$\omega_{\mathrm{s}}/\omega_{\mathrm{c}}(^{28}\mathrm{Si}^{13+})$	$3912.866\,064\,84(19)$	4.8×10^{-11}
D11	$A_{\mathrm{r}}(^{28}\mathrm{Si})$	$27.976\,926\,534\,99(52)$	1.9×10^{-11}
D12	$\Delta E_{\mathrm{B}}(^{28}\mathrm{Si}^{13+})/hc$	$420.6467(85)\ \times 10^{7}\ \mathrm{m^{-1}}$	2.0×10^{-5}
D13	δ_{Si}	$0.0(1.7)\ \times 10^{-9}$	8.3×10^{-10}
D14	$\omega_{\mathrm{c}}(\mathrm{d})/\omega_{\mathrm{c}}(^{12}\mathrm{C}^{6+})$	$0.992\,996\,654\,743(20)$	2.0×10^{-11}
D15	$\omega_{\mathrm{c}}(^{12}\mathrm{C}^{6+})/\omega_{\mathrm{c}}(p)$	$0.503\,776\,367\,662(17)$	3.3×10^{-11}
D19	$A_{\mathrm{r}}(^{1}\mathrm{H})$	$1.007\,825\,032\,241(94)$	9.3×10^{-11}
D21	$\Delta E_{\mathrm{B}}(^{1}\mathrm{H}^{+})/hc$	$1.096\,787\,717\,430\,7(10)\ \times 10^{7}\ \mathrm{m^{-1}}$	9.1×10^{-13}
D23	$\Delta E_{\mathrm{B}}(^{12}\mathrm{C}^{6+})/hc$	$83.083\,850(25)\ \times 10^{7}\ \mathrm{m^{-1}}$	3.0×10^{-7}

TABLE S4. Input data relevant for fundamental constants other than the Rydberg constant and nuclear charge radii for the CODATA18 dataset, taken from Table XXI of Ref. [1]. The D1-D13 inputs are relevant for the fine-structure constant and the electron mass, while the D14-D23 ones are relevant for the proton and deuteron masses. For the additive corrections D2, D9 and D13 of the a_e and bound g-factors in carbon (C) and silicium (Si), respectively, the relative uncertainty is relative to the value of the corresponding theoretical quantity.

dataset		CODATA18		DATA22		
constant	unit	value	rel. uncert.	value	rel. uncert.	shift
$R_\infty c$	Hz	$3.289\,841\,960\,2512(64) \times 10^{15}$	1.9×10^{-12}	$3.289\,841\,960\,2563(35) \times 10^{15}$	1.1×10^{-12}	$+0.9\sigma$
r_p	fm	$0.8414(19)$	2.2×10^{-3}	$0.8428(11)$	1.3×10^{-3}	$+0.8\sigma$
r_d	fm	$2.128\,04(74)$	3.5×10^{-4}	$2.128\,59(43)$	2.0×10^{-4}	$+0.8\sigma$
α		$7.297\,352\,569\,3(11) \times 10^{-3}$	1.5×10^{-10}	$7.297\,352\,564\,47(68) \times 10^{-3}$	9.3×10^{-11}	-4.3σ
$A_{\mathrm{r}}(e)$	u	$5.485\,799\,090\,67(15) \times 10^{-4}$	2.8×10^{-11}	$5.485\,799\,090\,397(94) \times 10^{-4}$	1.7×10^{-11}	-1.7σ
$A_{\mathrm{r}}(p)$	u	$1.007\,276\,466\,622(54)$	5.4×10^{-11}	$1.007\,276\,466\,596(14)$	1.3×10^{-11}	-0.5σ
$A_{\mathrm{r}}(d)$	u	$2.013\,553\,212\,744(41)$	2.0×10^{-11}	$2.013\,553\,212\,542(15)$	7.4×10^{-12}	-5.0σ

TABLE S5. The values and relative standard uncertainties of the main fundamental constants resulting from the least-squares adjustment based on the CODATA18 and DATA22 datasets without new physics. The symbol u denotes the unified atomic mass unit. The last column indicates the shift of the DATA22 value from the CODATA18 one in units of the CODATA18 uncertainty.

simplified expression

$$\nu^i_{\mathrm{SM}} = 2cR_\infty \left[\frac{a_i}{1 + m_e/m_N} + b_i(\alpha) + c_i(r_N) \right] , \tag{S8}$$

where $a_i \equiv 1/n_i'^2 - 1/n_i^2$, with n_i, n_i' the principal quantum numbers of the initial and the final state, respectively, while m_N is the mass of the nucleus, and i runs over all the transitions in Table S1. The coefficient $b_i(\alpha)$ denotes higher-order relativistic and QED corrections, which take the form of a power series in α and $\ln(\alpha)$, with the leading term starting at order $\mathcal{O}(\alpha^2)$. Finally, $c_i(r_N)$, where r_N is the nuclear charge radius, denotes the finite-nuclear-size and nuclear-polarizability corrections, where the leading term is of order $(r_N/a_0)^2$. The Bohr radius a_0 is given by

$$a_0^{-1} \equiv 4\pi R_\infty/\alpha \, . \tag{S9}$$

Since the hydrogen transition frequencies depend only weakly on m_e/m_N and α, these parameters have to be determined by other means, using inputs listed in Table S4, and labeled as Di, $i = 1, \ldots, 23$. The fine-structure

constant α is extracted from two different methods: (i) a comparison of theoretical and experimental results for the anomalous magnetic moment of the electron [2] and (ii) atom-recoil experiments [58, 59] that measure the mass m_A of atom A in units of the Planck constant h, combined with values of the Rydberg constant and masses of the electron and atom A in atomic mass units. Following Ref. [1], the CODATA 2018 dataset therefore also includes the values of relative atomic masses for the relevant atoms. These are taken from the 2016 Atomic Mass Evaluation (AME) [60, 61] (see Table S4).

The electron relative atomic mass is obtained from the measurements of spin-flip and cyclotron frequencies in a hydrogenic ion A using a theoretical calculation of the bound-electron g-factor $g_e(A)$. The atomic masses of the ions are deduced from the AME 2016 values of their neutral counterparts by subtracting the mass of the missing electrons and theoretically correcting for their binding energies [62].

The relative atomic mass of proton is extracted from the AME 2016 value of the hydrogen atom (using the theoretical binding energy) along with the more recent measurement [63]; for the relative atomic mass of deuteron, only the most recent measurement [64] is taken into account. The correlation coefficients for the A, B and D datasets are listed in Table IX and Table XXII of Ref. [1], respectively.

Finally, for the ease of comparison we follow Ref. [1], and add expansion factors for the CODATA18 analysis, in the case where no new physics is considered. The expansion factors reduce the tension in the data: we multiply by a factor of 1.6 the quoted errors for the proton radius data, *i.e.*, to all the A, B, and C inputs, and a factor of 1.7 for the proton mass data, *i.e.*, the D15 and D19 items. In the DATA22 analysis, on the other hand, we do not use the expansion factors.

Extracted values of $g_{\rm SM}$ using CODATA18 data and assuming no new physics contributions are listed in Table S5.

S3. THE DATA22 DATASET

The DATA22 dataset contains the CODATA18 inputs, but with updated values for both the experimental and theoretical inputs, Table S7, as well as the additional input data, Table S6. Together these then provide an improved sensitivity to NP effects.

To the list of hydrogen atom observables we added the two latest measurements of the $1S_{1/2} - 3S_{1/2}$ [21] and $2S_{1/2} - 8D_{5/2}$ [20] transition frequencies, and took into account recent theory improvements in the calculation of $nS_{1/2}$ energy levels [65–68]. The resulting slightly reduced theoretical uncertainties and their correlation coefficients are provided in Tables S7 and S8, respectively.

Among the data related to α determination, the recoil measurement of Ref. [69] was replaced by the improved measurement from the same group [58], and the measurement of the electron magnetic moment of Ref. [2] by the recently improved measurement [70]. Concerning the electron mass, the only change is the recently discovered long-distance contribution of order $\alpha^2(Z\alpha)^5 \ln(Z\alpha)$ [71] that is included in the theoretical expression for the bound-electron g-factor.

The input data for all of the utilized relative atomic masses were updated to their AME2020 values [72, 73]. In particular, these values take into account the recent high-precision measurements of the proton, deuteron and HD^+ masses [74, 75] and the deuteron–proton mass ratio [76]. Note that the latest measurement of m_p/m_d [77] is not included in the AME2020. We chose not to add it separately because of the possible (unknown to us) correlations with the AME2020 values of the hydrogen and deuterium atomic masses.

Moreover, we introduced measurements of transition frequencies in simple molecular or molecule-like systems: the hydrogen deuteride molecular ion (HD^+) [33–35], and the antiprotonic helium atom ($\bar{p}^3$He and $\bar{p}^4$He) [36, 37]. The main merit of adding these systems is their large sensitivity to the electron–nucleus mass ratios, enhanced by about three orders of magnitude relative to hydrogen. Due to this, the uncertainty on m_e/m_p is currently the largest contribution to the uncertainty of the SM prediction [39, 40]. On the other hand, this then allows for improved determinations of the m_e/m_p mass ratio from comparisons of theory and experiment [33–35, 40]. These determinations are in good agreement with the latest Penning trap measurements [72–76] and have similar uncertainties, in the 10^{-11} range. They constitute a test of the SM, which, despite being less precise than that obtained from hydrogen-like atoms, is much more sensitive to the NP models in which the mediators have increased couplings to nuclei. Examples of such models are the Higgs portal and the hadrophilic scalar, introduced in Section II in the main text. This feature has been exploited to constrain Yukawa-type forces between hadrons [11, 14], and is the major benefit of including the HD^+data in the DATA22 dataset. Similarly, the $\bar{p}$He spectroscopy leads to the determination of the antiproton-to-electron mass ratio (or, equivalently, the m_p/m_e ratio, assuming CPT symmetry) with an uncertainty slightly below 10^{-9} [37]. Although the achieved precision is lower, these data provide useful constraints on NP models for higher mediator masses relative to the HD^+ constraints, due to the smaller average distance between the nuclei [14]. The $\bar{p}$He energy levels depend on additional parameters, *i.e.*, the masses and charge radii of the α particle and the helion. Their masses are deduced from the AME 2020 values of helium-3 and helium-4 masses and their theoretical binding

Label	Input datum	Value	Rel. uncert.	Reference
A30	$\nu_{\rm H}(1S_{1/2}-3S_{1/2})$	2 922 743 278 665.79(72) kHz	2.5×10^{-13}	Grinin *et al.* [21]
A31	$\nu_{\rm H}(2S_{1/2}-8D_{5/2})$	770 649 561 570.9(2.0) kHz	2.6×10^{-12}	Brandt *et al.* [20]
D1	$a_e\equiv\frac{1}{2}(g-2)_e$	1.159 652 180 59(13) $\times10^{-3}$	1.1×10^{-10}	Fan *et al.* [70]
D3	$h/m_{\rm Rb}(^{87}{\rm Rb})$	4.591 359 258 90(65) $\times10^{-9}$ m^{2}s^{-1}	1.4×10^{-10}	Morel *et al.* [69]
D5	$A_{\rm r}(^{87}{\rm Rb})$	86.909 180 529(6)	6.9×10^{-11}	AME 2020 [73]
D6	$A_{\rm r}(^{133}{\rm Cs})$	132.905 451 958(8)	6.0×10^{-11}	AME 2020 [73]
D9	$\delta_{\rm C}$	0.0(9.4) $\times10^{-12}$	4.9×10^{-12}	Czarnecki *et al.* [71]
D13	$\delta_{\rm Si}$	0.0(5.8) $\times10^{-10}$	2.8×10^{-10}	Czarnecki *et al.* [71]
D11	$A_{\rm r}(^{28}{\rm Si})$	27.976 926 534 42(55)	2.0×10^{-11}	AME 2020 [73]
D14	$A_{\rm r}(^{2}{\rm H})$	2.014 101 777 844(15)	7.4×10^{-12}	AME 2020 [73]
D15	$\Delta E_{\rm B}(^2{\rm H}^+)/hc$	1.097 086 145 529 9(10) $\times10^{7}$ m^{-1}	9.1×10^{-13}	NIST ASD 2021 [62]
D19	$A_{\rm r}(^{1}{\rm H})$	1.007 825 031 898(14)	1.4×10^{-11}	AME 2020 [73]
D23	$\Delta E_{\rm B}(^{12}{\rm C}^{6+})/hc$	− − −	− − −	− − −
E1	$\nu_{\rm HD^+}((0,0)-(0,1))$	1 314 925 752.910(17) kHz	1.3×10^{-11}	Alighanbari *et al.* [33]
E2	$\nu_{\rm HD^+}((0,0)-(1,1))$	58 605 052 164.24(86) kHz	1.5×10^{-11}	Kortunov *et al.* [35]
E3	$\nu_{\rm HD^+}((0,3)-(9,3))$	415 264 925 501.8(1.3) kHz	3.1×10^{-12}	Patra *et al.* [34] + Germann *et al.* [14]
G1	$\nu_{\bar{\rm p}^4{\rm He}}((32,31)-(31,30))$	1 132 609 226.7(4.0) MHz	3.5×10^{-9}	Hori *et al.* [37]
G2	$\nu_{\bar{\rm p}^4{\rm He}}((33,32)-(31,30))$	2 145 054 858(7) MHz	3.4×10^{-9}	Hori *et al.* [36]
G3	$\nu_{\bar{\rm p}^3{\rm He}}((32,31)-(31,30))$	1 043 128 581(6) MHz	6.2×10^{-9}	Hori *et al.* [37]
G4	$\nu_{\bar{\rm p}^3{\rm He}}((35,33)-(33,31))$	1 553 643 100(10) MHz	6.7×10^{-9}	Hori *et al.* [36]
I1	$E_{\rm LS}(\mu^4{\rm He})$	1378.521(48) meV	3.5×10^{-5}	Krauth *et al.* [78]
I2	$E_{\rm LS}(\mu^3{\rm He})$	1258.586(49) meV	3.9×10^{-5}	Krauth [79]

TABLE S6. The new inputs for the DATA22 dataset. $\nu_{\rm HD^+}((v,L)-(v',L'))$ corresponds to the spin-averaged frequency of the transition between rovibrational states (v,L) and (v',L') of HD^+, where v (v') denote the initial (final) vibrational state, and L (L') the initial (final) orbital angular momentum quantum number. Similarly, $\nu_{\bar{\rm p}{\rm He}}((n,l)-(n',l'))$ is the frequency of the transition between states (n,l) and (n',l') of $\bar{p}$He , where n (n') and l (l') respectively denote the principal and orbital quantum numbers of the antiproton in the initial (final) state. The "− − −" symbol indicates that the input datum has been removed from the CODATA18 dataset.

Label	Input datum	Value (kHz)	Rel. uncert.
B1	$\delta_{\rm H}(1S_{1/2})/h$	0.0(1.3)	4.0×10^{-13}
B2	$\delta_{\rm H}(2S_{1/2})/h$	0.00(16)	2.0×10^{-13}
B3	$\delta_{\rm H}(3S_{1/2})/h$	0.000(49)	1.3×10^{-13}
B4	$\delta_{\rm H}(4S_{1/2})/h$	0.000(21)	1.0×10^{-13}
B5	$\delta_{\rm H}(6S_{1/2})/h$	0.000(11)	1.2×10^{-13}
B6	$\delta_{\rm H}(8S_{1/2})/h$	0.0000(47)	9.2×10^{-14}
B17	$\delta_{\rm D}(1S_{1/2})/h$	0.0(1.2)	3.5×10^{-13}
B18	$\delta_{\rm D}(2S_{1/2})/h$	0.00(15)	1.8×10^{-13}
B19	$\delta_{\rm D}(4S_{1/2})/h$	0.000(18)	9.0×10^{-14}
B20	$\delta_{\rm D}(8S_{1/2})/h$	0.0000(46)	8.9×10^{-14}

TABLE S7. Updated input data for the additive energy corrections accounting for missing contributions to the theoretical description of H and D energy levels. Uncertainties of non-S energy levels are not given here as they are unchanged with respect to Table VIII of [1].

energies. The charge radii are determined from the muonic helium spectroscopy data [78, 79], which we therefore added to the DATA22 dataset.

The new inputs for the DATA22 dataset are listed in Table S6. The updates of theoretical uncertainties for hydrogen and deuterium levels, based on Refs. [65–68], and the correlation matrix (calculated following the method of Ref. [1]) are collected in Tables S7 and S8, respectively. Moreover, an expansion factor of 2.4 is applied for the fine-structure constant data, *i.e.*, datapoints D3 and D4, because in this case the discrepancies cannot be due to NP, and are most

Updated correlation coefficients in hydrogen and deuterium				
$r(\mathrm{B1},\mathrm{B2}) = 0.9960$	$r(\mathrm{B1},\mathrm{B3}) = 0.9948$	$r(\mathrm{B1},\mathrm{B4}) = 0.9860$	$r(\mathrm{B1},\mathrm{B5}) = 0.5409$	$r(\mathrm{B1},\mathrm{B6}) = 0.5393$
$r(\mathrm{B1},\mathrm{B17}) = 0.9421$	$r(\mathrm{B1},\mathrm{B18}) = 0.9393$	$r(\mathrm{B1},\mathrm{B19}) = 0.9273$	$r(\mathrm{B1},\mathrm{B20}) = 0.4648$	$r(\mathrm{B2},\mathrm{B3}) = 0.9948$
$r(\mathrm{B2},\mathrm{B4}) = 0.9860$	$r(\mathrm{B2},\mathrm{B5}) = 0.5409$	$r(\mathrm{B2},\mathrm{B6}) = 0.5393$	$r(\mathrm{B2},\mathrm{B17}) = 0.9393$	$r(\mathrm{B2},\mathrm{B18}) = 0.9421$
$r(\mathrm{B2},\mathrm{B19}) = 0.9273$	$r(\mathrm{B2},\mathrm{B20}) = 0.4648$	$r(\mathrm{B3},\mathrm{B4}) = 0.9848$	$r(\mathrm{B3},\mathrm{B5}) = 0.5402$	$r(\mathrm{B3},\mathrm{B6}) = 0.5387$
$r(\mathrm{B3},\mathrm{B17}) = 0.9382$	$r(\mathrm{B3},\mathrm{B18}) = 0.9382$	$r(\mathrm{B3},\mathrm{B19}) = 0.9261$	$r(\mathrm{B3},\mathrm{B20}) = 0.4642$	$r(\mathrm{B4},\mathrm{B5}) = 0.5354$
$r(\mathrm{B4},\mathrm{B6}) = 0.5339$	$r(\mathrm{B4},\mathrm{B17}) = 0.9299$	$r(\mathrm{B4},\mathrm{B18}) = 0.9299$	$r(\mathrm{B4},\mathrm{B19}) = 0.9432$	$r(\mathrm{B4},\mathrm{B20}) = 0.4601$
$r(\mathrm{B5},\mathrm{B6}) = 0.2929$	$r(\mathrm{B5},\mathrm{B17}) = 0.5101$	$r(\mathrm{B5},\mathrm{B18}) = 0.5101$	$r(\mathrm{B5},\mathrm{B19}) = 0.5035$	$r(\mathrm{B5},\mathrm{B20}) = 0.2524$
$r(\mathrm{B6},\mathrm{B17}) = 0.5086$	$r(\mathrm{B6},\mathrm{B18}) = 0.5086$	$r(\mathrm{B6},\mathrm{B19}) = 0.5021$	$r(\mathrm{B6},\mathrm{B20}) = 0.9830$	$r(\mathrm{B7},\mathrm{B8}) = 0.0001$
$r(\mathrm{B9},\mathrm{B10}) = 0.0001$	$r(\mathrm{B11},\mathrm{B12}) = 0.6738$	$r(\mathrm{B11},\mathrm{B21}) = 0.9428$	$r(\mathrm{B11},\mathrm{B22}) = 0.4797$	$r(\mathrm{B12},\mathrm{B21}) = 0.4781$
$r(\mathrm{B12},\mathrm{B22}) = 0.9428$	$r(\mathrm{B13},\mathrm{B14}) = 0.2061$	$r(\mathrm{B13},\mathrm{B15}) = 0.2392$	$r(\mathrm{B13},\mathrm{B16}) = 0.2421$	$r(\mathrm{B13},\mathrm{B23}) = 0.9738$
$r(\mathrm{B13},\mathrm{B24}) = 0.1331$	$r(\mathrm{B13},\mathrm{B25}) = 0.1351$	$r(\mathrm{B14},\mathrm{B15}) = 0.2225$	$r(\mathrm{B14},\mathrm{B16}) = 0.2252$	$r(\mathrm{B14},\mathrm{B23}) = 0.1128$
$r(\mathrm{B14},\mathrm{B24}) = 0.1238$	$r(\mathrm{B14},\mathrm{B25}) = 0.1257$	$r(\mathrm{B15},\mathrm{B16}) = 0.2613$	$r(\mathrm{B15},\mathrm{B23}) = 0.1309$	$r(\mathrm{B15},\mathrm{B24}) = 0.9698$
$r(\mathrm{B15},\mathrm{B25}) = 0.1459$	$r(\mathrm{B16},\mathrm{B23}) = 0.1325$	$r(\mathrm{B16},\mathrm{B24}) = 0.1455$	$r(\mathrm{B16},\mathrm{B25}) = 0.9692$	$r(\mathrm{B17},\mathrm{B18}) = 0.9979$
$r(\mathrm{B17},\mathrm{B19}) = 0.9851$	$r(\mathrm{B17},\mathrm{B20}) = 0.4938$	$r(\mathrm{B18},\mathrm{B19}) = 0.9851$	$r(\mathrm{B18},\mathrm{B20}) = 0.4938$	$r(\mathrm{B19},\mathrm{B20}) = 0.4874$
$r(\mathrm{B21},\mathrm{B22}) = 0.3404$	$r(\mathrm{B23},\mathrm{B24}) = 0.0729$	$r(\mathrm{B23},\mathrm{B25}) = 0.0740$	$r(\mathrm{B24},\mathrm{B25}) = 0.0812$	

TABLE S8. Updated correlation coefficients $r(\mathrm{Bi},\mathrm{Bj}) \geq 0.0001$ between input data for the hydrogen and deuterium energy corrections.

Label	Input datum	Value (kHz)	Rel. uncert.
F1	$\delta_{\mathrm{HD}^+}(0,0)/h$	0(21.28)	5.4×10^{-12}
F2	$\delta_{\mathrm{HD}^+}(0,1)/h$	0(21.27)	5.4×10^{-12}
F3	$\delta_{\mathrm{HD}^+}(0,3)/h$	0(21.17)	5.4×10^{-12}
F4	$\delta_{\mathrm{HD}^+}(1,1)/h$	0(20.80)	5.4×10^{-12}
F5	$\delta_{\mathrm{HD}^+}(9,3)/h$	0(18.18)	5.2×10^{-12}
H1	$\delta_{\bar{\mathrm{p}}^4\mathrm{He}}(31,30)/h$	0(1454)	6.0×10^{-11}
H2	$\delta_{\bar{\mathrm{p}}^4\mathrm{He}}(32,31)/h$	0(1514)	6.6×10^{-11}
H3	$\delta_{\bar{\mathrm{p}}^4\mathrm{He}}(33,32)/h$	0(1595)	7.2×10^{-11}
H4	$\delta_{\bar{\mathrm{p}}^3\mathrm{He}}(31,30)/h$	0(1595)	6.9×10^{-11}
H5	$\delta_{\bar{\mathrm{p}}^3\mathrm{He}}(32,31)/h$	0(1625)	7.4×10^{-11}
H6	$\delta_{\bar{\mathrm{p}}^3\mathrm{He}}(33,31)/h$	0(1804)	8.5×10^{-11}
H7	$\delta_{\bar{\mathrm{p}}^3\mathrm{He}}(35,33)/h$	0(1998)	1.0×10^{-10}
I3	$\delta E_{\mathrm{LS}}(\mu^4\mathrm{He})$	0.000(293) meV	2.1×10^{-4}
I4	$\delta E_{\mathrm{LS}}(\mu^3\mathrm{He})$	0.000(521) meV	4.1×10^{-4}

TABLE S9. Input data for the additive energy corrections accounting for missing contributions to the theoretical description of energy levels of three-body systems HD^+ (F) and $\bar{p}\mathrm{He}$ (H), and the Lamb shift of muonic helium ions (I).

likely due to overlooked or underestimated systematics. Note that in the DATA22 inputs, we do not include newer input values for r_p from e-p scattering. These include a new result of the PRad collaboration, $r_p = 0.831(14)$ fm [80], and a reanalysis of modern data, giving $r_p = 0.847(8)$ fm [81]. In both cases, the determination of the proton charge radius from the electron-proton scattering data is not precise enough to make an appreciable difference in the fit.

The SM calculations of energy levels in the three-body systems HD^+ and $\bar{p}\mathrm{He}$ were carried out using the nonrelativistic QED (NRQED) approach. Energy levels are expanded in powers of α:

$$E = E^{(0)} + E^{(2)} + E^{(3)} + E^{(4)} + E^{(5)} + E^{(6+)} + E_{\mathrm{nuc}} + E_{\mathrm{other}}\,. \tag{S10}$$

Here $E^{(n)}$ is the total contribution of order $R_\infty \alpha^n$, E_{nuc} corresponds to nuclear structure (finite size and polarizability) corrections, and E_{other} to the muonic and hadronic vacuum polarization corrections. The first term, $E^{(0)}$, is the nonrelativistic energy. It was calculated by solving the three-body Schrödinger equation with high accuracy using a variational method [46, 82], in conjunction with the coordinate rotation (CCR) method for the case of resonant (quasibound) states in antiprotonic helium [83]. The leading relativistic correction, $E^{(2)}$, described by the Breit-Pauli

Correlation coefficients in $HD^+$and $\bar{p}$He				
$r(\mathrm{F1},\mathrm{F2}) = 1.0000$	$r(\mathrm{F1},\mathrm{F3}) = 1.0000$	$r(\mathrm{F1},\mathrm{F4}) = 1.0000$	$r(\mathrm{F1},\mathrm{F5}) = 0.9980$	$r(\mathrm{F2},\mathrm{F3}) = 1.0000$
$r(\mathrm{F2},\mathrm{F4}) = 1.0000$	$r(\mathrm{F2},\mathrm{F5}) = 0.9980$	$r(\mathrm{F3},\mathrm{F4}) = 1.0000$	$r(\mathrm{F3},\mathrm{F5}) = 0.9980$	$r(\mathrm{F4},\mathrm{F5}) = 0.9982$
$r(\mathrm{H1},\mathrm{H2}) = 0.9950$	$r(\mathrm{H1},\mathrm{H3}) = 0.9893$	$r(\mathrm{H1},\mathrm{H4}) = 0.9683$	$r(\mathrm{H1},\mathrm{H5}) = 0.9941$	$r(\mathrm{H1},\mathrm{H6}) = 0.9771$
$r(\mathrm{H1},\mathrm{H7}) = 0.9744$	$r(\mathrm{H2},\mathrm{H3}) = 0.9980$	$r(\mathrm{H2},\mathrm{H4}) = 0.9693$	$r(\mathrm{H2},\mathrm{H5}) = 0.9996$	$r(\mathrm{H2},\mathrm{H6}) = 0.9855$
$r(\mathrm{H2},\mathrm{H7}) = 0.9891$	$r(\mathrm{H3},\mathrm{H4}) = 0.9646$	$r(\mathrm{H3},\mathrm{H5}) = 0.9984$	$r(\mathrm{H3},\mathrm{H6}) = 0.9868$	$r(\mathrm{H3},\mathrm{H7}) = 0.9957$
$r(\mathrm{H4},\mathrm{H5}) = 0.9686$	$r(\mathrm{H4},\mathrm{H6}) = 0.9526$	$r(\mathrm{H4},\mathrm{H7}) = 0.9513$	$r(\mathrm{H5},\mathrm{H6}) = 0.9859$	$r(\mathrm{H5},\mathrm{H7}) = 0.9906$
$r(\mathrm{H6},\mathrm{H7}) = 0.9830$				

TABLE S10. Correlation coefficients $r(\mathrm{Fi},\mathrm{Fj}) \geq 0.0001$ and $r(\mathrm{Hi},\mathrm{Hj}) \geq 0.0001$ between input data for the $HD^+$and $\bar{p}$He energy corrections, respectively.

Hamiltonian, was calculated in [84] for HD^+rovibrational states and in [83, 85] for $\bar{p}$He. The $E^{(3)}$ term gives the leading radiative correction. It involves a numerically challenging quantity, the Bethe logarithm, which was obtained with high precision in [86] for $HD^+$and in [83] for $\bar{p}$He resonant states. Other relevant numerical data for this correction can be found in [83, 84]. The $R_\infty\alpha^4$-order correction, $E^{(4)}$, is made up of several contributions. One- and two-loop radiative corrections are given in [87, 88]. The relativistic correction was calculated in [89], and complete numerical data for HD^+can be found in the Supplemental material of [40]. The relativistic-recoil contribution was estimated from hydrogenlike atom theory [90], and the radiative-recoil term was taken from [91, 92].

It is worth noting that the relativistic correction, as well as several higher order corrections (the $R_\infty\alpha^5$ and $R_\infty\alpha^6$ terms), were calculated in the adiabatic approximation, where the wave function is written as a product of electronic and nuclear (rovibrational) wavefunctions. Electronic wavefunctions were obtained by solving the Schrödinger equation in the field of two clamped nuclei using a variational method [93]. In the first step, the $R_\infty\alpha^4$ relativistic corrections for the bound electron were calculated for a range of internuclear distances [89]. Then, the electronic curves were averaged over the nuclear wavefunction. The effect of corrections to the vibrational wavefunction induced by electronic corrections also needs to be taken into account [39, 40].

The radiative correction $E^{(5)}$ comprises the one-loop self-energy [94, 95] and the vacuum polarization (the Uehling potential) [96]. Complete numerical data for these two contributions are available in the Supplemental Material of Ref. [40] for HD^+. The Wichmann-Kroll vacuum polarization term, as well as two-loop and three-loop corrections are also included [94]. It should be noted that in the case of antiprotonic helium, the Uehling correction has only been estimated from the hydrogen atom theory [95].

One- and two-loop radiative corrections of order $R_\infty\alpha^6$ have been partially calculated in HD^+ in [39, 40] (complete numerical data are available in the Supplemental Material of [40]). In antiprotonic helium, they have only been estimated using the hydrogen atom theory [95]. In the nuclear finite-size and structure correction E_{nuc}, we include the leading-order finite-size correction [87, 88]. In HD^+, we take into account the higher-order nuclear corrections for the deuteron as described in [40]. In $\bar{p}$He , we also include these corrections for the helium-3 and helium-4 nuclei using the theoretical expressions presented in [65]. Similar corrections for the proton or antiproton are negligibly small at the present level of experimental and theoretical accuracy. The last term of Eq. (S10), denoted by E_{other}, corresponds to the muonic and hadronic vacuum polarization contributions, with the explicit expressions given in [40].

Theoretical uncertainties, Table S9, and their correlations, Table S10, were estimated following an approach similar to [1]. In HD^+, by far the largest sources of uncertainty are the yet uncalculated nonlogarithmic contributions of order $R_\infty\alpha^6$ in the one-loop self-energy and in the two-loop radiative correction [39, 40]. Smaller uncertainties come from the use of the adiabatic approximation in the calculation of some of the high-order correction terms ($R_\infty\alpha^4$ to $R_\infty\alpha^6$), from the relativistic-recoil correction of order $R_\infty\alpha^4 m_e/m_p$, and from the deuteron finite-size correction of order $R_\infty\alpha^4$. All these uncertainties are assumed to be fully correlated ("type u_0" uncertainties in the terminology of Ref. [1]), leading to correlation coefficients very close to 1. In $\bar{p}$He, additional sources of uncertainty come into play. Firstly, there is a larger uncertainty associated with one- and two-loop $R_\infty\alpha^6$-order corrections, since these terms have only been estimated from hydrogen atom results. Secondly, the one-loop vacuum polarization term of order $R_\infty\alpha^5$ has a significant uncertainty, for the same reason. Finally, some of the operator expectation values involved in the calculation of $E^{(2)}$ and $E^{(3)}$ have non-negligible numerical uncertainties. These numerical uncertainties are assumed to be uncorrelated ("type u_n" [1]), leading to smaller correlation coefficients with respect to HD^+.

The sensitivity of theoretical energy levels to the SM parameters (R_∞, α, and nuclear charge radii) is easily obtained by differentiating Eq. (S10), with the exception of particle masses. The dependence on particle masses mainly comes from the nonrelativistic energy (first term of Eq. (S10)), and is obtained numerically. The corresponding sensitivity coefficients are thus calculated numerically using the approach outlined in [82].

The values of g_{SM} determined from the fit to DATA22 data under the hypothesis of just SM, *i.e.*, no new physics

contributions, are listed in Table S5. We reiterate that in this adjustment the experimental errors were not increased by expansion factors, unlike in the CODATA18 dataset, except for the input data D3 and D4. The extraction of SM and NP parameters in the case of new physics is discussed in the main text and in Sections S5 and S6.

S4. FURTHER DETAILS ON NP BENCHMARK MODELS

In the main text we introduced in Section II five NP benchmark models that were then used to illustrate simultaneous extraction of SM and NP parameters, highlighting relevance of different spectroscopic data. Dark photon, $B-L$ gauge boson, and light Higgs-mixed scalar are all models that were already widely discussed in the literature. Hadrophilic scalar is a straight-forward modification of light Higgs-mixed scalar, taking the couplings in the leptonic sector to vanish. Below we give further details on the more involved modification, the ULD scalar, and also introduce another NP benchmark model, the scalar photon model, which was omitted from the discussion in the main text for brevity.

A. Up-Lepto-Darko-Philic (ULD) scalar

At low energies the ULD benchmark model comprises the SM Higgs and an additional light scalar singlet ϕ. The light scalar couples to up quarks, electrons and muons, and to a dark sector (SM singlet) fermion χ. For ease of comparison we use the notation for the couplings that is reminiscent of the Higgs-mixed scalar benchmark model, *i.e.*,

$$\mathcal{L}_\phi = k\frac{m_\ell}{v}\phi\bar{\ell}\ell + k\frac{m_u}{v}\phi\bar{u}u + y_\chi\phi\bar{\chi}\chi\,, \tag{S11}$$

with $\ell = e, \mu$ and k a model-dependent constant, however, the underlying high scale model is different. The couplings of ϕ to leptons and nucleons are then given by,

$$\mathcal{L}_{\rm eff} = g_\ell\phi\bar{\ell}\ell + g_N\phi\bar{N}N\,, \tag{S12}$$

where

$$g_\ell = k\frac{m_\ell}{v}\,, \qquad g_N = k\frac{\kappa'_N m_N}{v}\,, \tag{S13}$$

with $\kappa'_p \simeq 0.018(5)$ and $\kappa'_n \simeq 0.016(5)$. Here and for the scalar mixed portal couplings we use $\sigma^{p,n}_{u,d}$ values from [31] that were obtained following the procedure in [97], from $\sigma_{\pi N} = (50 \pm 15)$ MeV with conservative errors [31] which covers the spread between lattice QCD and pionic atom determinations. For the scalar operator of the strange quark we use the FLAG value $\sigma^p_s = \sigma^n_s = (52.9 \pm 7.0)$ MeV [98], obtained by averaging the $N_f = 2+1$ lattice QCD results [29, 99–102] (see also [99, 103, 104]). The NP parameters in the non-relativistic potential (2) are therefore given by

$$q_\ell = \frac{m_\ell}{\sqrt{m_e\kappa'_p m_p}}\,, \qquad q_N = \frac{\kappa'_N m_N}{\sqrt{m_e\kappa'_p m_p}}\,, \qquad \alpha_\phi = k^2\frac{m_e\kappa'_p m_p}{4\pi v^2}\,. \tag{S14}$$

Because the electron-proton coupling ratio is enhanced by $\kappa_p/\kappa'_p \sim \mathcal{O}(20)$ relative to the Higgs portal model, ϕ contributions to electron and muon $g-2$ and electron beam dump signals are significantly enhanced at fixed α_ϕ. Whenever necessary, bounds from beam dump experiments can be evaded assuming ϕ decays dominantly to additional invisible particles that could be related to dark matter. The ϕ contribution to muon $g-2$ can be reduced in less minimal models that include a custodial symmetry [57]. In such models the ϕ is accompanied by a light pseudoscalar particle which leads to a new spin-dependent force. Below we consider two distinct UV realizations of the ULD benchmark model.

The first possibility is a light scalar that couples to up quarks, electrons and muons through dimension five operators,

$$\mathcal{L}_\phi = \frac{y'_\ell}{\Lambda}\phi\,\bar{L}_\ell H\ell_R + \frac{y'_u}{\Lambda}\phi\,\bar{Q}_u\tilde{H}u_R + \text{h.c.}\,, \tag{S15}$$

where L_ℓ,Q_u and ℓ_R, u_R are the SM left-handed doublets and right-handed singlets, respectively, and $\tilde{H} = i\sigma_2 H^*$, with H the Higgs doublet. We assume that $y'_{\ell,u} = A \times m_{\ell,u}/v$, with A a universal dimensionless constant. The ϕ couplings to electrons, muons and up quarks are then the same as for the SM Higgs, but rescaled by $k = Av/\Lambda$. The

favored region of the ULD parameter space is for $|k| \simeq 2.4$, see Fig. 2. This requires the y' couplings to be larger than the corresponding SM Yukawa couplings by a factor of $A \simeq 100 \times (\Lambda/10\,\text{TeV})$. The largest is the coupling to muons, $y'_\mu \simeq 0.04 \times (\Lambda/10\,\text{TeV})$, which remains well perturbative even for a relatively high value of the cutoff scale Λ. The higher dimensional operators could arise from extra vector-like fermions at the scale Λ,

$$\mathcal{L}_{\rm UV} = \lambda_\ell \phi \bar{\Psi}^\ell_L \ell_R + x_\ell \bar{L}_\ell H \Psi^\ell_R + \lambda_u \phi \bar{\Psi}^u_L u_R + x_u \bar{Q}_u \tilde{H} \Psi^u_R + \text{h.c.}\,, \tag{S16}$$

where $\Psi^{\ell,u}$ are vector-like $\text{SU}(2)_L$ singlets of mass $M_{\ell,u}$. Integrating out the Ψ's yields the dimension-five operators in Eq. (S15) with $y'_\ell/\Lambda = x_\ell \lambda_\ell / M_\ell$ and $y'_u/\Lambda = x_u \lambda_u / M_u$.

In the second example of a UV model we supplement the SM field content by an extra Higgs doublet H' and a light singlet S that mixes with H', giving a light mass eigenstate $\phi \simeq S - \sin\theta' H'$, where $\theta' \ll 1$ is the mixing angle. The mixings of the two scalars with the SM Higgs is assumed to be small. The Yukawa couplings of H' to the SM fermions are assumed to be diagonal in the same basis as for the SM Higgs, with the only nonzero values the Yukawa couplings to the electron, the muon and the up quark. The SM Higgs H remains the dominant source of the electron, muon and up-quark masses and gives the entirety of the mass to the remaining SM fermions, while a subdominant parts of the electron, muon and up quark masses are due to the H' vacuum expectation value, $\langle H' \rangle = (0, v')/\sqrt{2}$ (for more general models of this type see, *e.g.*, [105–107]). That is, for electron, muon and up quark we have

$$m_{\ell,u} = m^H_{\ell,u} + m^{H'}_{\ell,u}\,, \tag{S17}$$

and $m_i = m^H_i$ for $i = \tau, d, s, c, b, t$. As a simple ansatz we take $m^{H'}_{\ell,u} = \delta \times m^H_{\ell,u}$, with $\delta \ll 1$ a universal factor. The ϕ couplings to electrons, muons and up quarks are then the same as for the SM Higgs, just rescaled by $k = \delta \sin\theta' v/v'$. The smallness of δ guarantees that the SM Higgs boson decays to muons, $h \to \mu^+\mu^-$, remain close to the SM predictions, in agreement with measurements, which require $\delta \lesssim 0.35$ at 90% CL [108]. Note that for $|k| \sim \mathcal{O}(1)$ the H' VEV must be smaller than the weak scale, $v'/v \simeq \delta \sin\theta' \ll 1$.

B. The scalar photon

Finally, we introduce an additional NP benchmark model, the "scalar photon", which will prove useful in the discussion of the $m_\phi \to 0$ limits in Section S7 below. The model consists of SM and a new light scalar that has the same pattern of couplings to the SM fermions as the photon, *i.e.*, $q_e = q_\mu = -q_p = -1$ and $q_n = 0$. That is, the only difference between the dark photon and the scalar photon model is the spin of the mediator. The comparison of bounds for the two models is therefore very informative, and highlights the importance of different subsets of data as well as the importance of implementing the $m_\phi \to 0$ behavior correctly. For instance, the right panel of Fig. S1 shows that the bounds on α_ϕ are rather different in the two models, despite the identical pattern of couplings. For $m_\phi \gtrsim m_e$, the bound is mostly set by the one-loop NP contribution to the electron $g-2$, which is larger for scalar photon. For $m_\phi a_0 \sim \mathcal{O}(1)$, the sensitivity of the fit to NP is dominated by hydrogen spectroscopy, which currently favors an addition of a repulsive NP interaction, thus yielding a stronger bound for dark photon. For very light masses $m_\phi \lesssim 1/a_0$, the bounds on α_ϕ weaken both for the scalar photon as well as for dark photon, due to the degeneracy with the QED photon. However, while for dark photon the degeneracy with the QED photon is exact in the $m_\phi \to 0$ limit, for massless scalar photon the degeneracy is lifted by $\mathcal{O}(\alpha^2)$ relativistic corrections. The bound on the scalar photon fine-structure constant therefore flattens below small enough m_ϕ, for which the NP contributions to the hydrogen energy levels become comparable to the relativistic corrections, *i.e.*, for $a_0 m_\phi \sim \mathcal{O}(\alpha)$.

The least-squares adjustment also reveals that the DATA22 dataset favors the scalar photon model over the SM-only hypothesis at the 4.8σ level, with the best-fit point at $m_\phi = 0.2\,\text{keV}$ and $\alpha_\phi = 2.7 \times 10^{-13}$. The 1σ, 2σ, 3σ, and 4σ CL contours are shown as orange shaded regions in Fig. S2 (right panel), along with the exclusions from stellar cooling (grey), SN1987a cooling (red), NA62 search for $K^+ \to \pi^+ + \text{inv}$ (green) and E137 (dark green). The overlap of all the bounds excludes the region favored by the adjustment. Note that even if the SN cooling does not apply [109] the favored region is still excluded by the combination of other constraints.

Figure S1 (left panel) shows the allowed range for α_ϕ at 99% CL assuming a given scalar photon mass (orange shaded region), with the solid line showing the best fit α_ϕ for each value of m_ϕ. Note that for very light ϕ, with mass below around 20 eV the value $\alpha_\phi = 0$ is allowed at 99% CL, *i.e.*, the very light scalar photon is not preferred over the SM. Similarly, the heavy scalar photon, with mass above about 100 keV, is not preferred over the SM, in agreement with Fig. S2 (right panel).

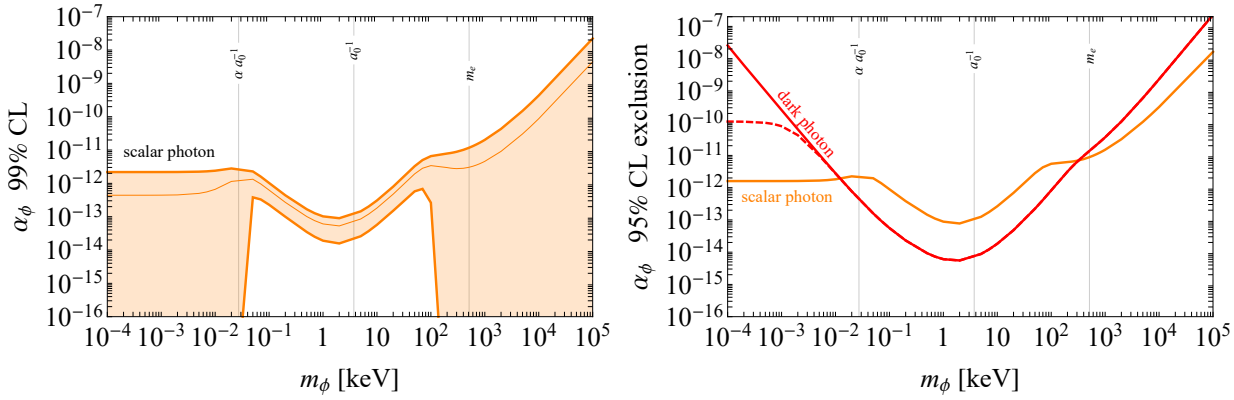


FIG. S1. (left) 99% CL interval on the NP coupling α_ϕ as function of m_ϕ for the scalar photon benchmark model. (right) Comparison of the 95%CL exclusion on α_ϕ in the dark photon (red) and scalar photon (orange) models. The dashed line denotes the unphysical bound in the dark photon case that results from the same fit but without redefining α in order to recover the correct behavior in the long-range ($m_\phi \to 0$) regime.

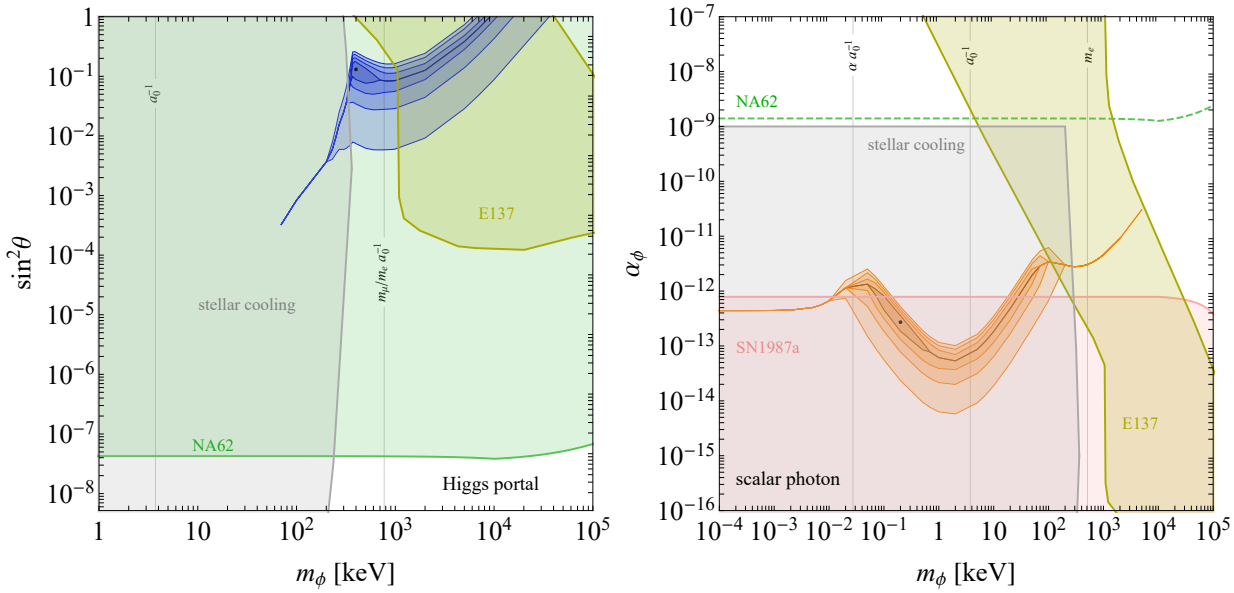


FIG. S2. NP parameters favored by atomic spectroscopy in the Higgs portal (left) and the scalar photon (right) benchmark models, profiling over the SM parameters. The black dot indicates the best-fit point and the blue-shaded (left) and orange-shaded (right) areas represent the $1, 2, 3, 4\sigma$ confidence regions around it. The other shaded areas denote excluded region by SN1987a [50, 51, 110] (pink), the $K^+ \to \pi^+ X$ search at NA62 [52] (green, the dashed line is an NNLO estimate), stellar cooling [53] (gray) and the E137 beam dump experiment [54, 55] (yellow).

S5. NEW PHYSICS IMPROVEMENTS RELATIVE TO THE SM

The fit to DATA22 dataset assuming only the SM, *i.e.*, no NP, gives the minimum χ^2 per degree of freedom $\chi^2_{\rm SM}/\nu_{\rm dof} \simeq 1.4$ ($\nu_{\rm dof} = 102 - 62 = 40$). Ignoring correlations, the largest contributions to the χ^2 are from the hydrogen observables A12–A15, A23 and A31 in Tables S1 and S6, see also Fig. S3.

The adjustments exhibit no significant ($> 2\sigma$) preference over the SM for the gauged $B - L$, dark photon and hadrophilic scalar models, see Fig. S4. The ULD scalar and the Higgs portal, on the other hand, show a preference over the SM for $m_\phi a_0 \gtrsim m_\mu/m_e$. This NP evidence can already be anticipated from the bounds on α_ϕ shown in Fig. 1. For masses heavier than $m_\phi a_0 \gtrsim m_\mu/m_e$ the main constraint is from muonic hydrogen, so that one would expect, due to enhanced couplings of ϕ to muons, the bounds on ULD scalar and Higgs portal to be stronger by a factor $\mathcal{O}(m_\mu/m_e)$ relative to all the other models we consider, in which $q_e = q_\mu$. The bounds in Fig. 1, however, are found

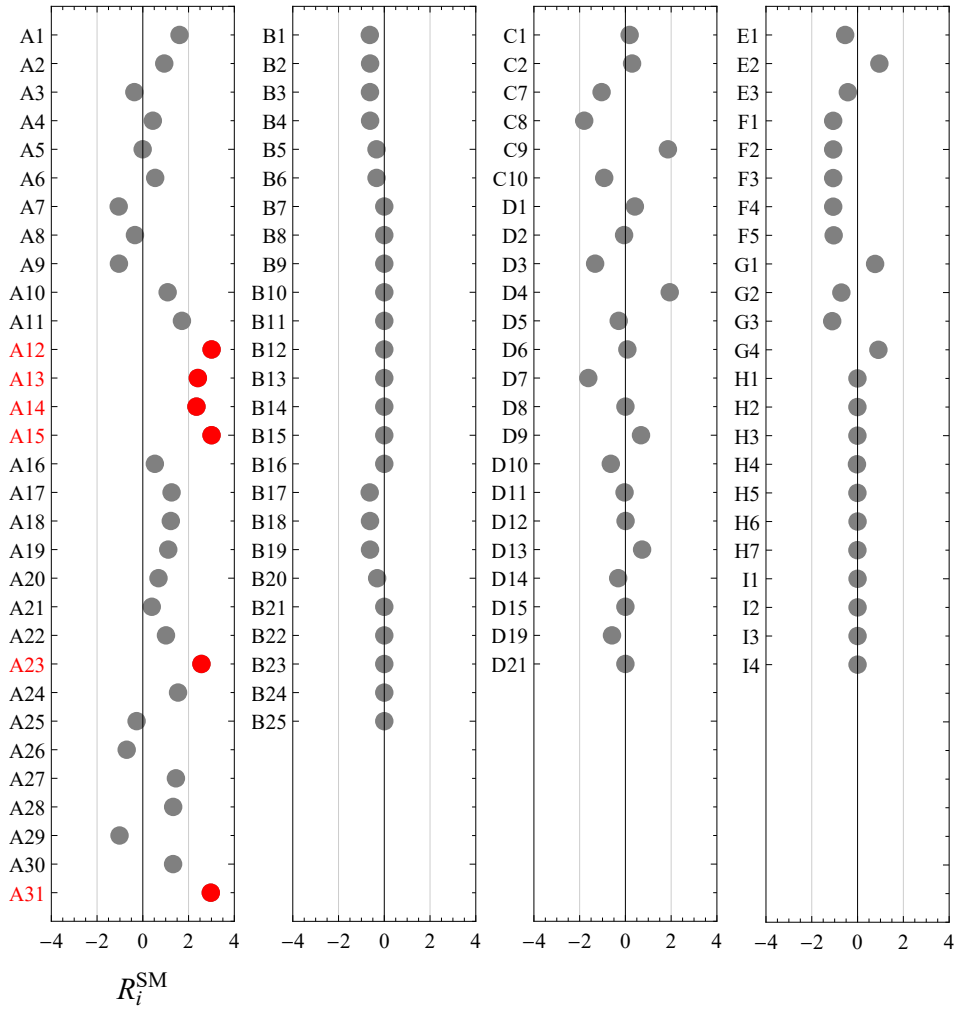


FIG. S3. Normalized residuals $R_i^{\rm SM}$, see Eq. (S7), of individual input data within the DATA22 dataset. Input data satisfying $|R_i^{\rm SM}| \geq 2$ are indicated in red.

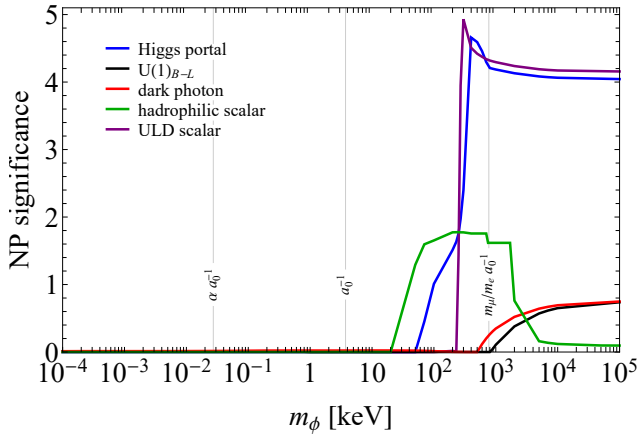


FIG. S4. NP significance relative to the SM, $Z \equiv \sqrt{\chi^2_{\rm SM} - \chi^2_{\rm NP}}$, where $\chi^2_{\rm NP}$ ($\chi^2_{\rm SM}$) is the minimum χ^2 with NP (without NP) for a fixed NP mass.

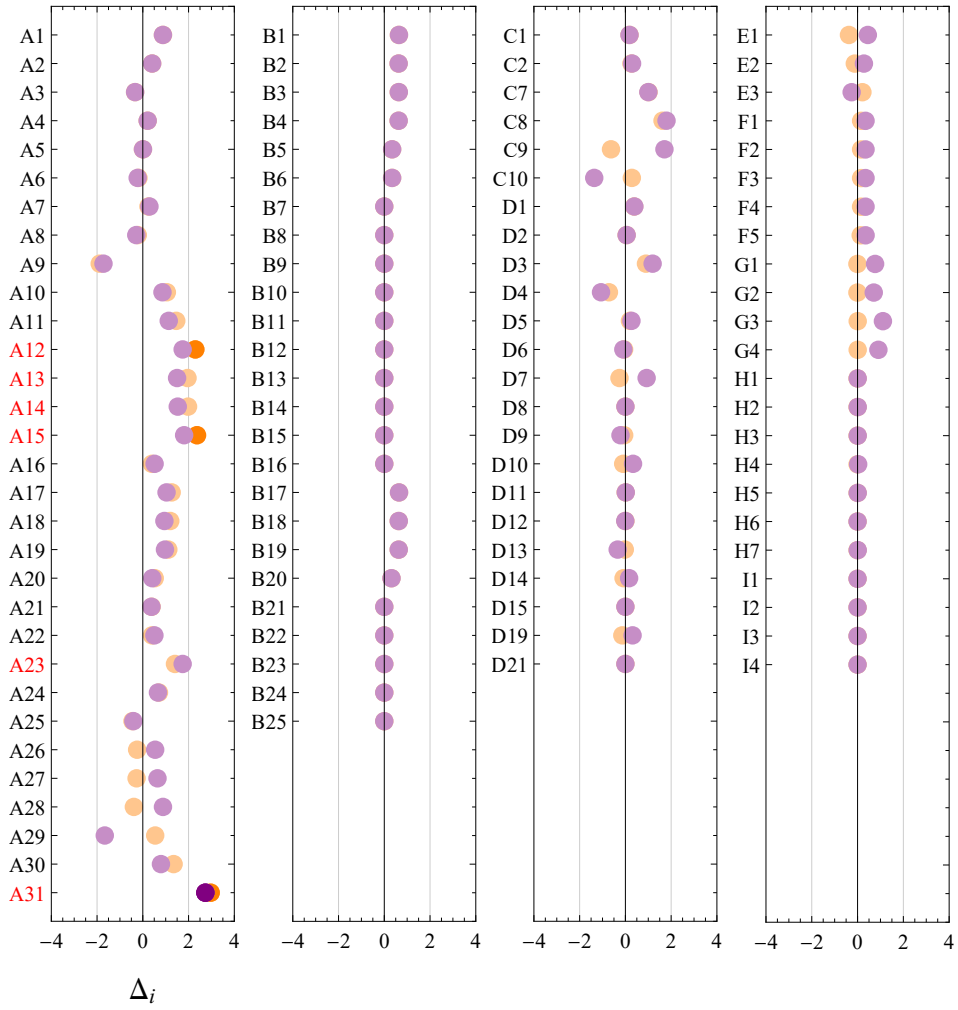


FIG. S5. Difference of squared normalized residuals, $\Delta_i \equiv \mathrm{sign}(\delta R_i^2)\sqrt{|\delta R_i^2|}$, of individual input data within the DATA22 dataset, where $\delta R_i^2 \equiv (R_i^{\rm SM})^2 - (R_i^{\rm NP})^2$ is the contribution of each input datum i to the χ^2 difference between SM and NP ignoring correlations with other input data. $R_i^{\rm NP}$ are the normalized residuals of the scalar photon (orange) and the ULD scalar (purple) models, evaluated at their respective best-fit points. Darker colors indicate input data with $\sqrt{|\delta R_i^2|} \geq 2$. Input data satisfying $|R_i^{\rm SM}| \geq 2$ are indicated in red.

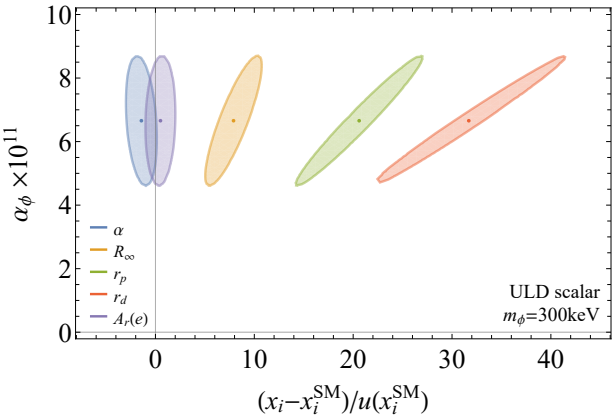


FIG. S6. The 68 % CL covariance ellipses of the fundamental constants $x_i = \alpha, R_\infty, r_p, r_d, A_{\rm r}(e)$ (color coding as indicated) and the NP coupling constant α_ϕ for a 300 keV ULD scalar, using the DATA22 dataset. The x-axis is shifted by $x_i^{\rm SM}$, the extracted value of x_i for the SM-only hypothesis, and normalized by its uncertainty.

for the ULD and Higgs portal models to be weaker by a factor of $\mathcal{O}(20)$ than the naive expectations, for heavy ϕ. This is an indication that the data favors NP models with large μ-to-e coupling ratio over the SM in this mass range.

The significance of the deviation is at the $\sim 4\sigma$ level for the Higgs portal and the $\sim 5\,\sigma$ level for the ULD scalar. Figure 2 in the main text shows the preferred region in the ULD model parameters, where the best-fit point is $m_\phi = 300\,\text{keV}$ and $\alpha_\phi = 6.7 \times 10^{-11}$. Similarly, the left panel in Fig. S2 shows the preferred region for the Higgs portal parameters, with the best-fit point given by $m_\phi = 400\,\text{keV}$ and $\alpha_\phi = 2.5 \times 10^{-11}$. In both models the NP evidence receives support mostly from the recent measurements of the hydrogen $2S_{1/2} - 8D_{5/2}$ and $1S_{1/2} - 3S_{1/2}$ transitions [20, 21], as well as muonic deuterium, see Fig. S5. These tensions between data and the SM prediction are not new. The authors of Ref. [20] already pointed out the inconsistency of their $2S_{1/2} - 8D_{5/2}$ measurement with hydrogen theory and discussed NP interpretations in the form of Yukawa potential as well as its impact on the determination of R_∞. The $1S_{1/2} - 3S_{1/2}$ hydrogen and muonic deuterium Lamb shift measurements are known pieces of the so-called proton-radius puzzle [19, 111]. Our analysis shows that both tensions can be significantly ameliorated by postulating the existence of a single light scalar mediator, with the pattern of couplings to the SM fermions such as in the Higgs portal or in the ULD model, with only the latter also avoiding other, non-spectroscopic constraints.

S6. EXTRACTION OF FUNDAMENTAL CONSTANTS IN THE PRESENCE OF NP

In this section we discuss the effect of NP on the uncertainty of extracted fundamental constants. Figure 3 in the main text shows the 68% CL region for simultaneous determinations of the proton charge radius r_p and the Rydberg constant R_∞, assuming the SM-only hypothesis, using either the CODATA18 or DATA22 datasets, as well as for the ULD and the Higgs portal models (DATA22 only). As expected, the extracted values of r_p and R_∞ are highly correlated regardless of the existence of NP. NP induces significant shifts in the extracted values of R_∞ and r_p, notably when the data shows evidence for nonzero α_ϕ. Furthermore, the larger the shift in the extracted value of the fundamental constant, the more important is the correlation with α_ϕ, see Fig. S6.

Figure S7 illustrates the impact of allowing in the fit the possibility of NP. The relative uncertainties on the extracted values of fundamental constants, g_{SM}, are plotted as a function of the NP mass, for all six benchmark models. In many cases the uncertainties on the extracted fundamental constants change significantly, by factors of $\mathcal{O}(1)$, even if the NP parameters are strongly disfavored by data. For dark photon, in particular, the uncertainty on α increases as $\sim m_\phi^{-2}$ for ϕ masses below $\sim 10^{-2}\,\text{keV}$. This is a result of a degeneracy between α and α_ϕ in the $m_\phi \to 0$ limit, see the inset in Fig. S7. The combination $\alpha + \alpha_\phi$, however, is well determined.

S7. FURTHER DETAILS ON THE $m_\phi \to 0$ LIMIT

The extraction of bounds on NP parameters in the $m_\phi \to 0$ limit requires special care, especially for new vector bosons that couple to the SM fermions in a similar way than the QED photon. Massless dark photon, in particular, is completely degenerate with the QED photon. In the massless limit all the effects due to the exchanges of a dark photon are therefore absorbed in the SM predictions by performing the shift $\alpha \to \alpha + \alpha_\phi$, *i.e.*, the massless dark photon is unobservable. The reason is that for a massless dark photon one can always choose a linear combination of photon and dark photon fields that does not couple to the SM currents, and redefine the orthogonal linear combination as the SM photon. This behavior should be reflected in theoretical predictions. That is, in the $m_\phi \to 0$ limit, $\mathcal{O}_{\text{SM}}$ only depends on $\alpha + \alpha_\phi$, so that this sum can be interpreted as the SM fine-structure constant. Note that the shift $\alpha \to \alpha + \alpha_\phi$ also implies a redefinition of the Rydberg constant extracted from hydrogen, $R_\infty \to (\alpha + \alpha_\phi)^2 m_e/(4\pi)$, and of the Bohr radius, $a_0^{-1} \to 4\pi R_\infty/(\alpha + \alpha_\phi)$.

The treatment of NP corrections in Eqs. (4), (5) is designed such that *a)* it is always correct to leading order in α_ϕ, and *b)* it reproduces the correct result for the massless dark photon, *i.e.*, for $m_\phi \to 0$ and $q_i = Q_i$. The degeneracy of dark photon with the QED photon is broken either by having $m_\phi \neq 0$ (while still $q_i = Q_i$ for all particles) or by having a coupling for at least one particle differ from the QED one, $q_i \neq Q_i$. For infinitesimal deformations from the massless dark photon limit the leading observable NP effect is captured by the potential $\widetilde{V}_{\text{NP}}$. For $m_\phi a_0 \ll 1$, keeping the NP couplings still aligned with QED, $q_i = Q_i$, the leading non-trivial term in $\tilde{V}_{\text{NP}}$ scales as $\mathcal{O}(m_\phi^2)$, indicating a parametric loss of sensitivity to NP for small m_ϕ, $\widetilde{\mathcal{O}}_{\text{NP}} \propto (m_\phi a_0)^2$. (The term linear in m_ϕ is independent of r and thus unobservable.) For massless vectors with couplings that deviate inifinitesimally from the QED couplings, *i.e.*, $q_i = Q_i + \delta q_i$ with $\delta q_i \ll Q_i$, the leading effect scales as $\tilde{V}_{\text{NP}}^{ij} \propto \delta q_i + \delta q_j$.

For massive vectors, $m_\phi a_0 \sim \mathcal{O}(1)$, with charges that differ significantly from the QED ones, $q_i \neq Q_i$, the prescription in Eqs. (4), (5) is equivalent to working to leading order in V_{NP} and not shifting the SM predictions at all. That is, for m_{NP} nonzero and $q_i \neq Q_i$ such as massive $B - L$ gauge boson, both prescriptions are strictly speaking

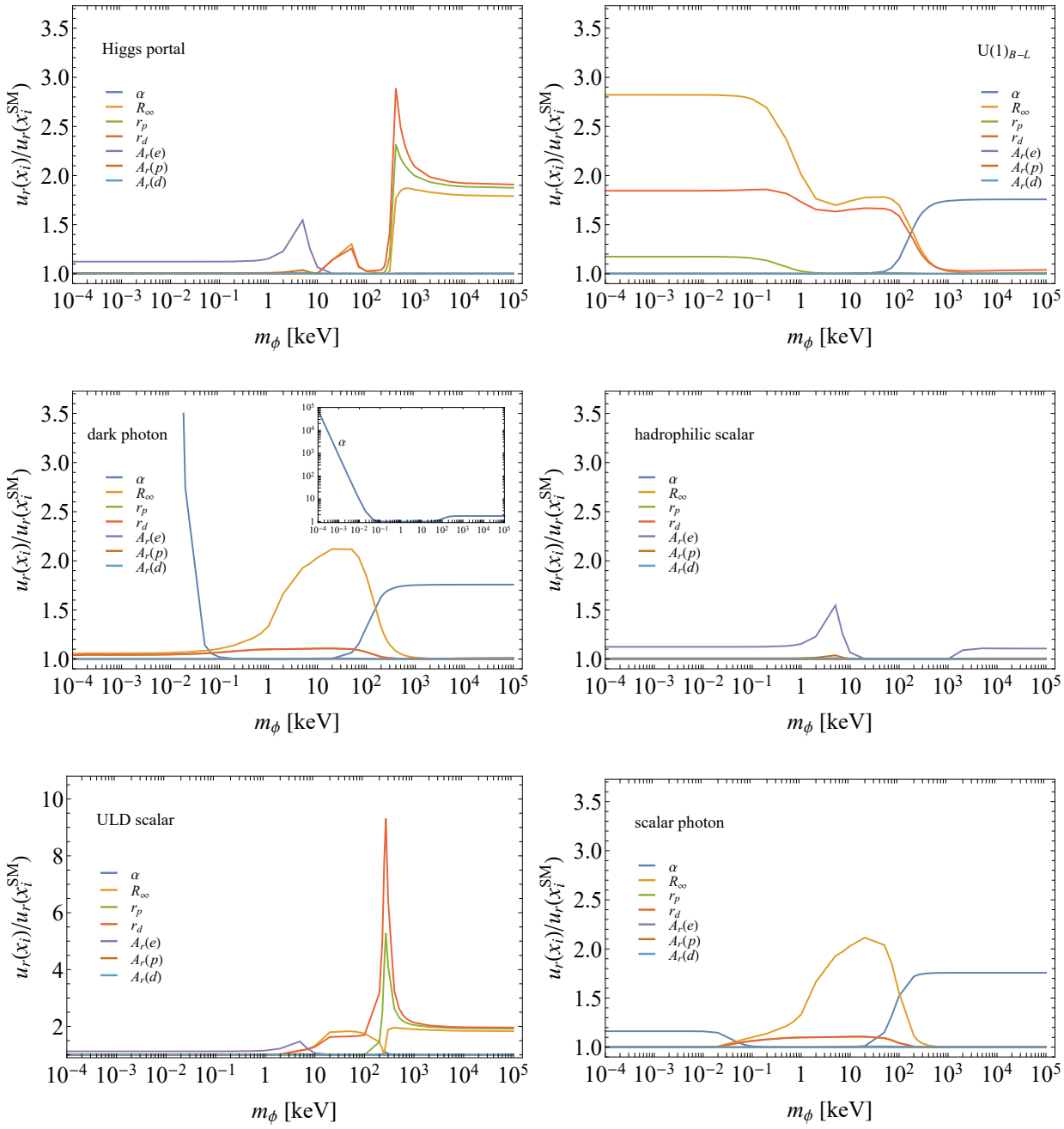


FIG. S7. The ratio of relative uncertainties $u_r(x_i) \equiv u(x_i)/x_i$ on fundamental constants $x_i = \alpha, R_\infty, r_p, r_d, A_{\rm r}(e)$ (color coding as indicated) when extracted assuming NP relative to the extraction in the SM, for all six NP benchmark models.

only correct to leading order in α_ϕ and one can in principle use either of the two. However, for massless $B-L$ the prescription in Eqs. (4), (5) will reproduce correctly many of the higher order corrections in hydrogen (though not all, for instance hadronic vacuum polarization type contributions are not correctly captured, as are higher order corrections for deuterium), and this may offer some benefit.

Finally, we point out that there is partial degeneracy also between QED photon and the massless scalar photon model with $q_i = Q_i$. However, the degeneracy of QED photon and massless scalar photon is only approximate. It is broken by relativistic corrections, which are different for vector and scalar degrees of freedom. This corrections are $\mathcal{O}(\alpha^2)$ in atomic and molecular transitions. This means that it is consistent to work with unshifted SM predictions and add the NP contributions at leading order using unsubtracted potential $V_{\rm NP}$ also in the case of scalar photon, as

we did for the other three scalar benchmarks in the main text.