

## ATOMIC PHYSICS

# A measurement of the atomic hydrogen Lamb shift and the proton charge radius

N. Bezginov<sup>1</sup>, T. Valdez<sup>1</sup>, M. Horbatsch<sup>1</sup>, A. Marsman<sup>1</sup>, A. C. Vutha<sup>2</sup>, E. A. Hessels<sup>1\*</sup>

The surprising discrepancy between results from different methods for measuring the proton charge radius is referred to as the proton radius puzzle. In particular, measurements using electrons seem to lead to a different radius compared with those using muons. Here, a direct measurement of the  $n = 2$  Lamb shift of atomic hydrogen is presented. Our measurement determines the proton radius to be  $r_p = 0.833$  femtometers, with an uncertainty of  $\pm 0.010$  femtometers. This electron-based measurement of  $r_p$  agrees with that obtained from the analogous muon-based Lamb shift measurement but is not consistent with the larger radius that was obtained from the averaging of previous electron-based measurements.

The Lamb shift—the difference in energy between the two most tightly bound excited states (the  $2S_{1/2}$  and  $2P_{1/2}$  states) in the hydrogen atom—has played a pivotal role in explaining the fundamental interactions between charged particles since the advent of quantum mechanics. The Dirac theory of relativistic quantum mechanics predicts that the  $2S_{1/2}$  and  $2P_{1/2}$  states have the same energy (1). This energy degeneracy is accidental, in that it occurs only if the force between the electron and the proton is exactly proportional to the reciprocal of their separation squared ( $1/r^2$ ), as predicted by Coulomb's law. The very existence of the Lamb shift indicates that Coulomb's law fails for short distance scales. Three reasons for the failure were proposed in the early days of quantum mechanics. First, and most central to this work, the electron can penetrate inside the proton and experiences a smaller force while inside (2). Second, the Heisenberg uncertainty principle allows an electron and its antiparticle (the positron) to appear and then disappear (i.e., to be created and annihilated), so long as the pair is in existence for a sufficiently short time span. When such a pair exists inside of a hydrogen atom, the field inside the atom separates the positive and negative charges, and this vacuum polarization modifies Coulomb's law (3). Third, the electron can interact with itself; however, calculations of this self energy led to the unfortunate conclusion that the effect is infinite. In 1947, however, only months after Lamb and Retherford definitively showed that the  $2S_{1/2}$  and  $2P_{1/2}$  states are not degenerate (4), Bethe proposed that there would be a finite residual effect if the infinite effect of a free electron is (carefully) subtracted from the infinite effect for an electron within a hydrogen atom (5).

All three effects contribute to the Lamb shift. In each case, it is the energy of the  $2S_{1/2}$  state that is affected, because it is in this state that the electron and proton can overlap. The  $2P_{1/2}$  state is almost entirely unaffected owing to the centrifugal force that stems from its angular momentum and keeps the electron away from the proton. The predictions of vacuum polarization and self energy, and their confirmation by Lamb, were foundational for the development of the theory of quantum electrodynamics (QED), which is still believed to properly describe electromagnetic interactions.

In the decades after Lamb's measurement, increasingly precise measurements of the Lamb shift were performed, culminating in a measurement with a precision of 9 kHz made by Lundeen and Pipkin in 1981 (6). A decades-long concerted effort by many theoretical physicists has allowed the QED prediction for the Lamb shift interval to also become increasingly precise [(7) and references therein], allowing for increasingly stringent tests of QED theory. Here, we present a high-precision measurement of the Lamb shift in which we use the frequency-offset separated oscillatory field (FOSOF) technique (8), which was developed for this measurement. We measure an energy difference between the  $2S_{1/2}$  ( $F = 0$ ) and  $2P_{1/2}$  ( $F = 1$ ) states of 909.8717 MHz (multiplied by Planck's constant,  $h$ ). The uncertainty on our measurement is  $\pm 3.2$  kHz. Here,  $F$  refers to the hyperfine state, as shown in Fig. 1. This measurement has direct consequences for the size of the proton [the root mean square (RMS) charge radius,  $r_p$ ] and for tests of the theory of QED.

## The proton size

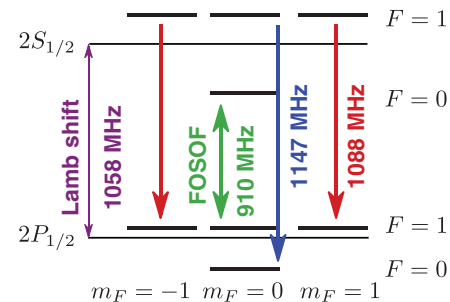
Because the electron can penetrate inside of the proton, the size of the proton affects the Lamb shift, but this contribution is small: only  $\sim 0.01\%$  of the shift. However, increasingly accurate Lamb shift experiment and theory became sensitive to this small contribution, and the uncertainty in the size of the proton became a limiting factor

in allowing for tests of QED using the Lamb shift. Measurements of the Lamb shift (along with the assumption that QED calculations were correct) became a way to determine  $r_p$ . The Committee on Data for Science and Technology (CODATA) 2014 value of the radius (7) includes both this determination and determinations using elastic scattering of electrons:  $r_p$ [CODATA 2014] = 0.8751(61) fm. However, the most precise determination of the proton radius comes from measuring the  $2S \rightarrow 2P$  Lamb shift in muonic hydrogen (9, 10):  $r_p$ [muonic] = 0.84087(39) fm. The large discrepancy [0.0342(61) fm, or  $\sim 4\%$ ] between these two values (the first determined entirely using electrons; the second determined entirely using muons) is referred to as the proton radius puzzle (11, 12) and has led to speculation about whether muons and electrons interact differently with the proton. In this work, we measure the hydrogen Lamb shift, the direct analog of the muonic measurement, in an attempt to help resolve the puzzle.

## Measurement technique

Our measurement (Fig. 2) uses a fast beam of hydrogen atoms created by passing protons (accelerated to a kinetic energy of 55 keV) through a molecular hydrogen gas target. Collisions with the  $H_2$  molecules cause about half of the protons to neutralize into hydrogen atoms. Approximately 4% are created in the  $2S_{1/2}$  state, which is metastable, with a lifetime of one-eighth of a second.

The neutral atoms travel with a speed  $v$  of  $\sim 3$  mm/ns, or 1% of the speed of light. They pass between 70-cm-long deflector plates, where an electric field of 20 V/cm deflects the remaining protons out of the beam. All four  $2S_{1/2}$  states with  $F = 0$  and  $F = 1$  (Fig. 1) are equally populated at the start, but only the  $F = 0$  state survives the passage through two radio frequency (rf) cavities (blue and red in Figs. 1 and 2) that have their rf intensity and frequencies tuned to



**Fig. 1. Energy levels of hydrogen relevant to our experiment.** Shown are the  $2S_{1/2}$  and  $2P_{1/2}$  energy levels, indicating the Lamb shift, as well as the hyperfine ( $|Fm_F\rangle$ ) states of atomic hydrogen. The green arrow indicates the transition measured in this work; the transitions marked with red and blue arrows are used to remove populations from the  $2S_{1/2}$  ( $F = 1$ ) states. Here,  $F$  and  $m_F$  are the total angular momentum and its projection along the direction of the rf fields.

<sup>1</sup>Department of Physics and Astronomy, York University, Toronto, Ontario M3J 1P3, Canada. <sup>2</sup>Department of Physics, University of Toronto, Toronto, Ontario M5S 1A7, Canada. \*Corresponding author. Email: hessels@yorku.ca

transfer more than 99.9% of the  $F = 1$  atoms to the  $2P_{1/2}$  state. The  $2P_{1/2}$  state has a lifetime of 1.6 ns and decays to the  $1S_{1/2}$  ground state over a distance scale of 0.5 cm.

The  $2S_{1/2} (F = 0) \rightarrow 2P_{1/2} (F = 1)$  transition measured in this work is driven (green in Figs. 1 and 2) as the atoms pass through a pair of waveguides. The waveguides are electrically shorted at the top end so that the rf fields are reflected back on themselves to form a standing wave and the atoms pass through at the anti-node of this standing wave. The  $2S_{1/2} (F = 0)$  atoms that survive these fields (after passing through two more cavities to once again remove any unwanted  $F = 1$  population) are detected by applying an electric field that mixes the  $2S_{1/2}$  and  $2P_{1/2}$  states. The mixture quickly decays to the  $1S_{1/2}$  ground state by emitting a 121.6-nm Lyman- $\alpha$  photon, and the photon is efficiently detected after passing out of our vacuum system through a  $MgF_2$  window and photoionizing an acetone molecule.

### FOSOF

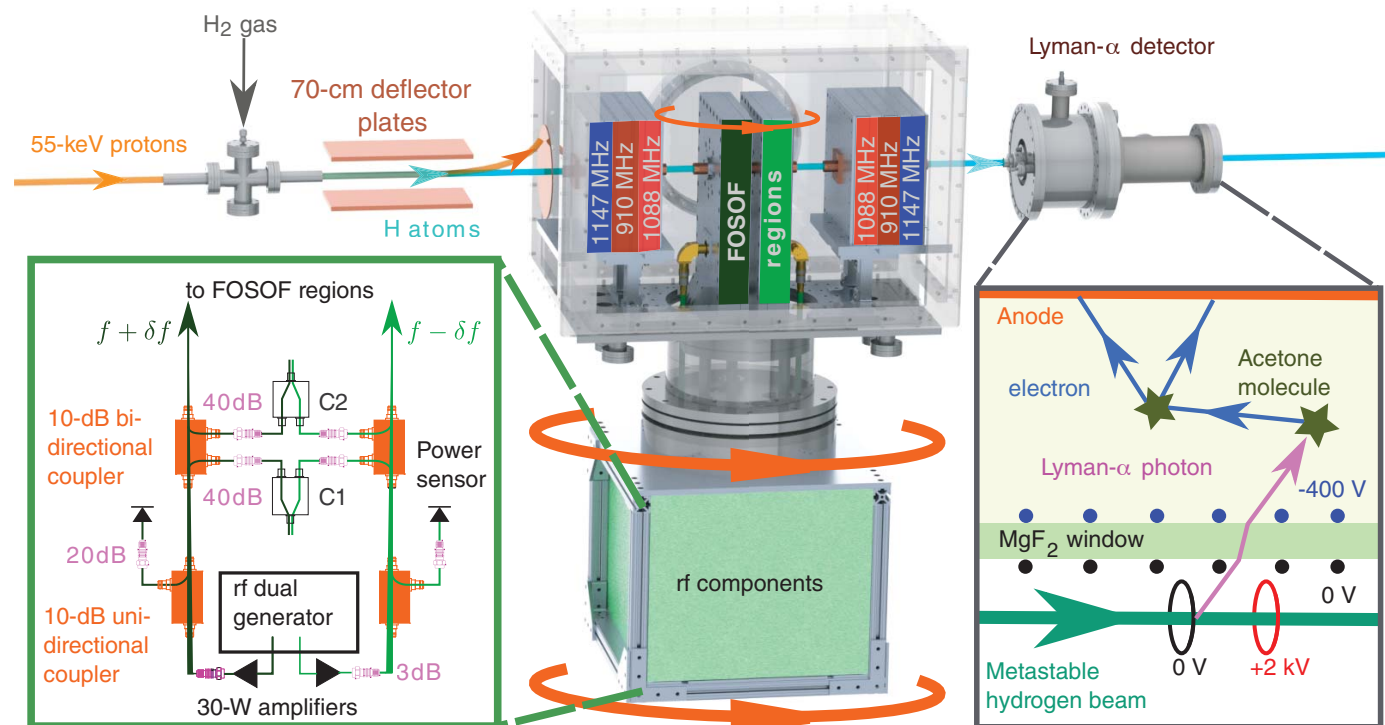
The measurement is performed using the recently developed FOSOF technique (8, 13), which is a modification of the Ramsey technique (14)

of separated oscillatory fields. For FOSOF, the frequencies of the two separated fields are offset from each other ( $f - \delta f$  and  $f + \delta f$ , with a frequency difference  $2\delta f$  set to 625 Hz for this work), so that the relative phase of the two fields varies continuously in time. The combined effect of the two FOSOF regions (green in Fig. 2) for driving the  $2S_{1/2} (F = 0) \rightarrow 2P_{1/2} (F = 1)$  transition depends on this phase, and the number of Lyman- $\alpha$  photons observed varies in time, as shown (in red) in Fig. 3A. This signal consists of a large constant component stemming from  $2S_{1/2}$  atoms that survive all of the rf fields and a small sinusoidal component caused by a progression between constructive and destructive interference from the two FOSOF regions as their relative phase varies. The sinusoidal signal is small because of the short lifetime of the  $2P_{1/2}$  state (1.6 ns), but it still shows a signal-to-noise ratio of 30:1, despite only 6 ms of averaging time represented by each of the data points in Fig. 3A.

Also shown in Fig. 3A is a 625-Hz reference signal obtained by beating the two rf frequencies. The key to the FOSOF technique is that the Lyman- $\alpha$  signal and the reference signal will be in phase when  $f$  is set to the atomic resonant frequency  $f_0$ , and the phase difference  $\Delta\theta$  between

the two is proportional to  $f - f_0$ . The phase difference occurs because the rf and the atoms accumulate phase at rates determined by  $f$  and  $f_0$ , respectively, during the time it takes the atoms to traverse the distance between the two FOSOF regions.

To obtain a precision measurement, the phase difference  $\Delta\theta$  has to be measured to an accuracy of better than 1 mrad. Given that filtering and time delays can also cause phase shifts, we employ three techniques to ensure that unintended phase shifts do not affect our measurement. First, we take data with the two FOSOF regions set to  $f - \delta f$  and  $f + \delta f$  (Fig. 3A) and change the frequencies to  $f + \delta f$  and  $f - \delta f$  (Fig. 3B). As seen in the figure,  $\Delta\theta$  has opposite signs in these two cases, so that an average  $[\Delta\theta^{(AB)}]$  of  $\Delta\theta^{(A)}$  and  $-\Delta\theta^{(B)}$  cancels any unintended phase shifts related to the limited bandwidth of the detection system. This frequency change is performed every few seconds. Second, we physically rotate by  $180^\circ$  the entire FOSOF system [both the out-of-vacuum parts (the generator, amplifiers, cables, and rf monitoring system) and the in-vacuum parts (the green waveguides in Fig. 2)]. The whole system is rotated as a single unit by using a 32-cm-diameter rotational feedthrough for the



**Fig. 2. The measurement apparatus.** Metastable  $2S_{1/2}$  atoms are created by colliding a beam of protons with  $H_2$  molecules. Deflector plates remove the protons, and rf cavities (red and blue) remove  $2S_{1/2} (F = 1)$  atoms. The  $2S_{1/2} (F = 0)$  atoms are driven to the  $2P_{1/2} (F = 1)$  state in a pair of FOSOF regions (green), which have rf frequencies that are offset from each other. The number of surviving  $2S_{1/2} (F = 0)$  atoms is measured by mixing them in an electric field and observing the resulting Lyman- $\alpha$  photons

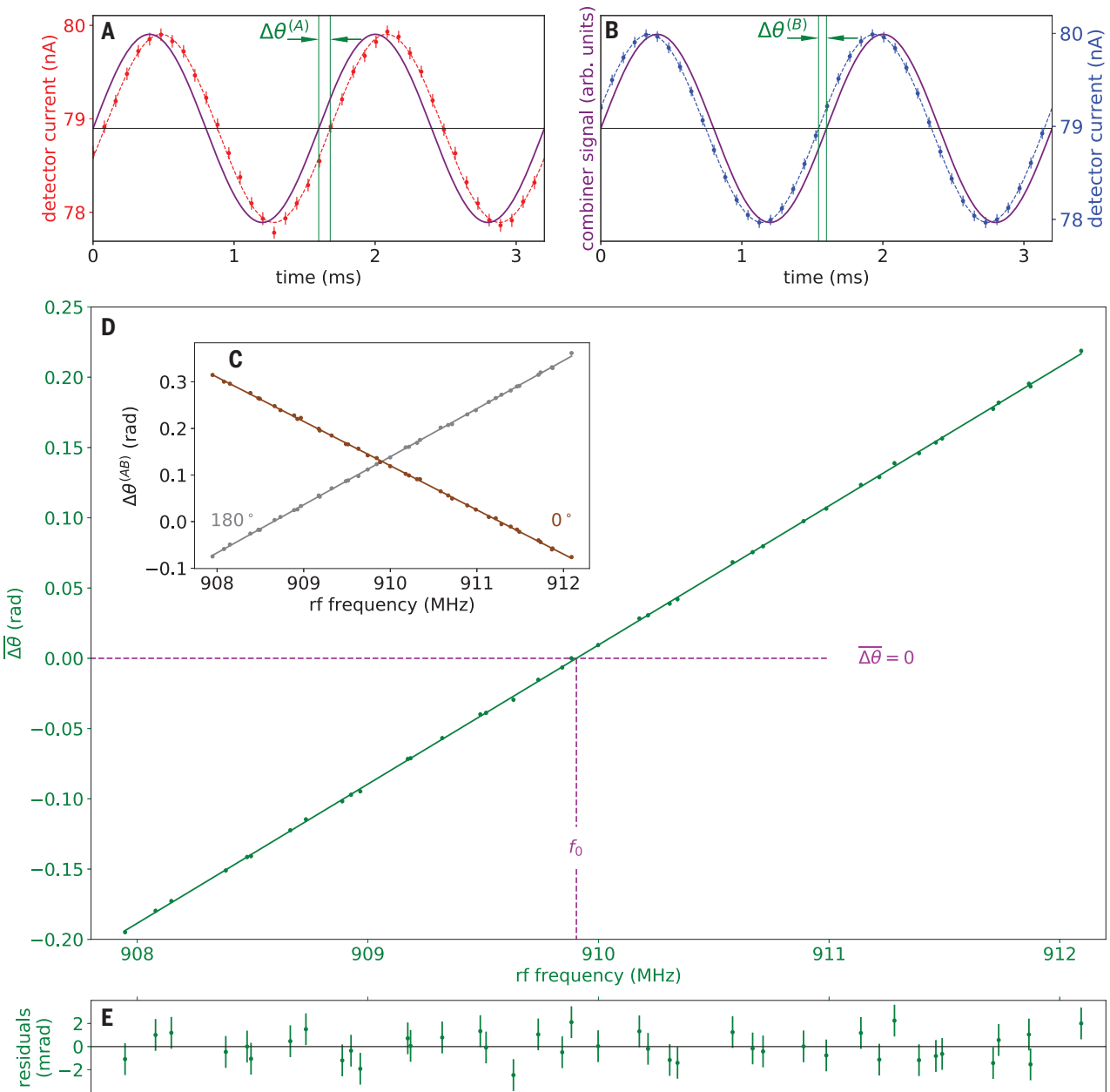
via an efficient gas-ionization detector. Key to the success of the measurement is the fact that the entire FOSOF system (generator, amplifiers, monitors, and in-vacuum FOSOF waveguides) can be rotated by  $180^\circ$ , so that the atoms can encounter the two fields in the reverse order. The additional 910-MHz cavities shown (brown) are used to test for systematic effects. The relative phase of the rf going to and reflecting back from the FOSOF regions is measured by rf combiners C1 and C2.

vacuum connection, with all critical components constructed rigidly and temperature stabilized to ensure that the rf system is unaffected by the rotation. This physical rotation is performed approximately every hour, and, similar to the results shown in Fig. 3, A and B, it flips the sign of  $\Delta\theta$ . The average  $\langle\Delta\theta\rangle$  of  $\Delta\theta^{(AB)}[180^\circ]$  and  $-\Delta\theta^{(AB)}[0^\circ]$  cancels any imperfections in the rf system and our measurement of the reference signal. The

third technique we use to ensure that we are correctly determining  $\Delta\theta$  is to measure the beat signal twice: once by combining the rf fields before they enter the FOSOF regions (C1 in Fig. 2) and once after they return from these regions (C2). The important parameter is the relative phase of the rf fields at the position of the atoms, but the consistency of the C1 and C2 phases provides evidence that both are

accurate measures of the phase at the atomic position.

Figure 3C shows 90 min of FOSOF data. The brown and gray lines show the measured  $\Delta\theta$  as a function of rf frequency  $f$  for the  $0^\circ$  and  $180^\circ$  rotation, respectively, of the FOSOF system. The intersection of the two curves provides a measure of the atomic resonant frequency  $f_0$ . A better way to obtain  $f_0$  is to take the average of



**Fig. 3. The FOSOF signal.** (A and B) The FOSOF technique measures the phase difference  $\Delta\theta$  between the atomic signal [red and blue in (A) and (B), respectively] and the reference signal (purple). The sign of  $\Delta\theta$  depends on whether the atoms first encounter the  $f + \delta f$  or  $f - \delta f$  rf fields. In particular, in (A) the atoms first travel through the rf field region of frequency  $f - \delta f$  and then through the field region of frequency  $f + \delta f$ . For the plot in (B), the order of the encountered

frequencies is reversed.  $|\Delta\theta^{(A)}|$  is larger than  $|\Delta\theta^{(B)}|$  because of a phase offset caused by the limited bandwidth of the detection system. (C) The average of  $\Delta\theta^{(A)}$  and  $-\Delta\theta^{(B)}$  cancels this phase offset and is shown versus  $f$  for the two orientations of the FOSOF regions. (D) Average of  $\Delta\theta(0^\circ)$  (brown) and  $-\Delta\theta(180^\circ)$  (gray). The straight-line fit determines the  $f_0$  at which  $\Delta\theta = 0$ . (E) The residuals from the fit in (D) show that the data are fit well [ $\chi^2(39) = 29.1$ ] by a simple straight line.

$\Delta\theta(180^\circ)$  and  $-\Delta\theta(0^\circ)$  (8), because this average cancels any (possibly frequency-dependent) phase lags. This average is shown in Fig. 3D, where the intercept ( $\Delta\theta = 0$ ) determines  $f_0$ . A straight-line fit determines this intercept to a statistical accuracy of better than 2 kHz, and the residuals (E) show that the data are well fit by the straight line. The data are taken with a randomized ordering of the frequencies  $f$ , and all of the data for this work are taken blind, with an offset (unknown to the experimenters until the end) added to all  $f$  values.

The three techniques discussed above are intended to cancel any effects due to relative phase errors. Given this measurement's sensitivity to submilliradian phase shifts, as well as the need to be conservative for such a high-precision measurement, we consider the possibility of a remaining relative phase error  $\phi_0$  (which would indicate that the use of C1 and C2 does not, on average, represent the relative phase that the atoms experience, even after employing the three techniques). To explore such an effect, we measure the relative phase of the fields in the two FOSOF regions by inserting probes into the tubes through which the atoms would otherwise travel and using a third rf combiner to determine the relative phase. This test shows that the relative phase in the cavities agrees with that measured by C1 and C2 to within  $\pm 0.18$  mrad. This phase uncertainty leads to an uncertainty of 1.5 kHz in our measured  $f_0$ .

The data in Fig. 3 are taken with a beam speed of  $v = 3.22$  mm/ns, a spacing of  $d = 4$  cm between the two FOSOF regions, and an rf electric field amplitude of  $E_{\text{rf}} = 18$  V/cm. In all, 116 data sets

similar to that depicted in Fig. 3 are taken, with 18 different combinations of parameters ( $v$ ,  $d$ ,  $E_{\text{rf}}$ ). The parameter ranges used are: 1.99 mm/ns  $\leq v \leq 3.22$  mm/ns, 4 cm  $\leq d \leq 7$  cm, and 5 V/cm  $\leq E_{\text{rf}} \leq 24$  V/cm. Smaller  $v$  and  $E_{\text{rf}}$  and larger  $d$  are preferred to control systematic effects, but these values are more difficult to achieve as the signal-to-noise ratio scales as  $(E_{\text{rf}})^2$  multiplied by  $\exp[-(d/v)/(3.2 \text{ ns})]$ . The statistical uncertainty obtained from combining all data sets is  $< 0.5$  kHz, which allows for a careful study of systematic effects.

### Effect of other states

One concern for the present measurement is the effect of other hyperfine transitions and transitions involving states with  $n \geq 3$ . This concern is heightened by our analysis of unexpectedly large shifts from  $n = 3$  states for Lamb shift measurements (15). It is also heightened because we still see 0.3% of our Lyman- $\alpha$  signal, even when we employ all six of the rf cavities shown in Fig. 2, despite the fact that our modeling indicates that we should be able to reduce our  $2S_{1/2}$  population to  $< 0.1\%$ . We believe that this 0.3% comes from collisional repopulation of the  $2P_{1/2}$  state in our detector, as the percentage increases with increasing pressure, but we cannot exclude the possibility that it might come from cascades from higher- $n$  states. However, in the following paragraphs, we discuss four experimental tests that show that there are no substantial effects from higher- $n$  states.

First, a third rf cavity (the one labeled 910 MHz in Fig. 2) is employed to transfer  $2S_{1/2}$  ( $F = 0$ ) atoms to the  $2P_{1/2}$  ( $F = 1$ ) state. This third cavity

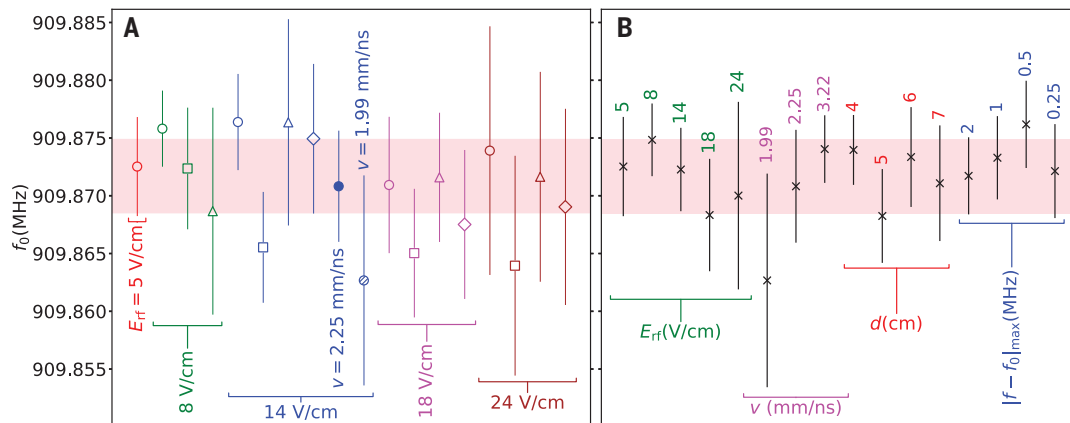
reduces the intended FOSOF signal size, but, because the rf cavity is tuned to the  $2S_{1/2}$  ( $F = 0$ )  $\rightarrow$   $2P_{1/2}$  ( $F = 1$ ) transition frequency, it has relatively little effect on other states. Consequently, the relative population of other states [relative to the population of the  $2S_{1/2}$  ( $F = 0$ ) state] is larger; thus, any shifts that could result from these other states would also be larger. No substantial shifts are seen, as shown in table S3. Second, the deflection plates in Fig. 2 remove almost all of the  $n = 3$  population because they mix the longer-lived  $3S_{1/2}$  states with the short-lived  $3P$  states. Some data (table S4) are taken with no deflection field, and the absence of this field also does not reveal a shift. This test shows that  $n = 3$  shifts similar to those discussed in (15) do not play a major role here. Third, different experimental parameters ( $v$ ,  $d$ ,  $E_{\text{rf}}$ ) would lead to different contributions from other states. The consistency of our results for these different parameters provides further evidence that there are no contributions from transitions involving other states. Fourth, both the vacuum pressure near the FOSOF regions (table S4) and the pressure of the  $H_2$  gas target (table S3) were varied, the first to check for collisional repopulation of other states by the background gas and the second because different  $H_2$  pressures are expected to lead to different initial population distributions of  $n$  states. Again, no notable shifts were found.

### AC Stark shifts and modeling

A substantial systematic effect in the present measurement is the ac Stark shift (caused by  $E_{\text{rf}}$ ) due to the effect of the  $2S_{1/2}$  and  $2P_{1/2}$  states on

**Table 1. Systematic corrections.** Shown are systematic corrections and corrected centers for the 18 parameter sets ( $d$ ,  $v$ ,  $E_{\text{rf}}$ ) used in our measurement. Systematic effects for Doppler shift ( $\Delta_{\text{Dop}}$ ), ac Stark shift ( $\Delta_{\text{ac}}$ ), and phase error ( $\Delta_\phi$ ) are listed with their uncertainties. The rightmost column provides the corrected line center along with the total statistical and systematic uncertainties. The bottom row indicates weighted averages.

$d$ (cm)	$v$ (mm/ns)	$E_{\text{rf}}$ (V/cm)	weight (%)	$\Delta_{\text{Dop}}$ (kHz)	$\Delta_{\text{ac}}$ (kHz)	$\Delta_\phi$ (kHz)	$f_0$ [corrected] (kHz)
4	3.22	5	6.3	-52.6(1.1)	5.5(0.8)	0.0(2.0)	909,872.5(3.5)(2.4)
4	3.22	8	15.3	-52.6(1.1)	13.5(1.3)	0.0(2.0)	909,875.8(2.0)(2.6)
4	3.22	14	0.0	-52.6(1.1)	42.9(3.3)	0.0(1.9)	909,876.4(1.1)(4.0)
4	2.25	14	3.0	-25.7(0.5)	32.4(2.2)	0.0(1.8)	909,870.8(3.9)(2.9)
4	1.99	14	0.5	-20.0(0.4)	28.6(2.3)	0.0(1.6)	909,862.7(8.6)(2.9)
4	3.22	18	0.0	-52.6(1.1)	75.1(5.4)	0.0(1.7)	909,870.9(0.9)(5.8)
4	3.22	24	0.0	-52.6(1.1)	152.9(10.5)	0.0(1.7)	909,873.9(1.4)(10.6)
5	3.22	8	10.5	-52.6(1.1)	10.0(1.4)	0.0(1.6)	909,872.4(4.6)(2.4)
5	3.22	14	11.1	-52.6(1.1)	34.4(2.9)	0.0(1.6)	909,865.5(3.3)(3.4)
5	3.22	18	3.4	-52.6(1.1)	61.6(4.6)	0.0(1.5)	909,865.0(2.5)(4.9)
5	3.22	24	0.0	-52.6(1.1)	128.8(9.0)	0.0(1.4)	909,863.9(2.2)(9.2)
6	3.22	8	4.2	-52.6(1.1)	8.3(1.1)	0.0(1.4)	909,868.7(8.7)(2.0)
6	3.22	14	11.7	-52.6(1.1)	29.1(2.4)	0.0(1.3)	909,876.4(4.3)(2.9)
6	3.22	18	9.1	-52.6(1.1)	52.4(3.8)	0.0(1.3)	909,871.6(3.7)(4.2)
6	3.22	24	0.0	-52.6(1.1)	110.9(7.7)	0.0(1.2)	909,871.6(4.5)(7.8)
7	3.22	14	11.6	-52.6(1.1)	26.3(2.0)	0.0(1.2)	909,874.9(5.9)(2.5)
7	3.22	18	11.1	-52.6(1.1)	46.8(3.2)	0.0(1.2)	909,867.5(5.3)(3.6)
7	3.22	24	2.4	-52.6(1.1)	98.6(6.8)	0.0(1.1)	909,869.0(4.8)(6.9)
weighted average:				-51.6(1.0)	29.5(2.3)	0.0(1.5)	909,871.7(1.4)(2.9)



**Fig. 4. Observed values for the atomic resonant frequency,  $f_0$ .**

(A) Consistent centers are found for the 18 ( $v$ ,  $d$ ,  $E_{rf}$ ) parameter sets used. Circles, squares, triangles, and diamonds represent  $d = 4, 5, 6,$  and  $7$  cm, respectively. (B) Averaged  $f_0$  values for each  $v$ ,  $d$ , and  $E_{rf}$  also agree, as do

$f_0$  values obtained with the use of different frequency ranges. The pink band shows the  $1\sigma$  uncertainty range for the current measurement. Numbers above the data points in (B) give the value of the parameters listed below the data points.

each other (the Bloch-Siegert shift) and the effect of the  $2P_{3/2}$  state on the  $2S_{1/2}$  state. For a constant  $E_{rf}$ , the shift is straightforward to calculate, but for separated fields, a density-matrix numerical modeling of the entire experiment, similar to that in (16), is required. All 400 density-matrix elements for the  $n = 1$  and  $n = 2$  states are numerically integrated through the 700-ns trajectory of the experiment, including all of the electric, magnetic, and rf fields of the experiment and including phase averaging over these fields. The equations used for this modeling are described by Marsman *et al.* (16). The numerical calculations are intensive and are carried out on a cluster of computers. The rf and dc fields themselves are also calculated numerically. The shifts that result from this modeling are found using the method shown in fig. S2, and these shifts vary between 5 and 155 kHz (depending on  $E_{rf}$ ,  $d$ , and  $v$ , as shown in Table 1). The modeled shifts are proportional to  $E_{rf}^2$  for small  $E_{rf}$  (with some additional  $E_{rf}^4$  dependence at larger  $E_{rf}$ ) and show a small quadratic dependency on the distance of the atomic trajectory from the axis, defined by the 20-mm-diameter tubes through which the atoms enter and exit the rf regions.

We deduce the value of  $E_{rf}$  for the rf power used in our FOSOF regions by comparing modeling to experiment for the effectiveness at reducing the  $2S_{1/2}$  ( $F = 0$ ) population (fig. S3). This process is complicated by the fact that the shift is dependent on the RMS distance of the atoms from the axis and by our incomplete understanding of the 0.3% residual signal present when all quench cavities are employed. We use an RMS off-axis distance of 1.75 mm (as would be expected from the experimental geometry and collimation) and include the 0.3% residual in our modeling, but we also include 50% uncertainties in both of these quantities. Because we do not want our final result to be substantially dependent on modeling, we con-

firm from the consistency of data (Fig. 4) with different modeled ac Stark shifts that the calculated shifts are correct to 5%, and we include an additional 5% uncertainty to all calculated shifts. The inclusion of the 5% uncertainty causes us to give a low weight to the measurements with larger ac Stark shift. In the weighted average of our measurements, the average ac Stark shift is 29.5(2.3) kHz.

### Time dilation

The speed  $v$  of our atomic beam can be accurately measured by using rf pulses to remove  $2S_{1/2}$  atoms in the first FOSOF region and seeing the effect of a second set of time-delayed pulses in the second FOSOF region. This method is illustrated in fig. S1. Any unanticipated time delays can be canceled by rotating the FOSOF system by  $180^\circ$ . The results of these speed measurements are shown in table S1. A second measure of  $v$  comes from comparing (table S2) the measured FOSOF slope, as in Fig. 3D, to that of the modeled line shape. The two methods of determining the speed produce consistent results. The determined speed is  $v = 3.22(3)$  mm/ns for the 55-keV protons (table S1). The resulting time-dilation correction is 52.6(1.0) kHz. Additionally, data are taken with 21- and 27-keV protons [with speeds  $v = 1.99(2)$  and  $2.25(2)$  mm/ns and time-dilation corrections of 20.0(0.4) and 25.7(0.5) kHz] to confirm this correction (table S1). The first-order Doppler shift is negligible because the fields in the FOSOF regions are almost-perfect standing waves (owing to an almost-perfect reflection from the waveguide short) and because the atomic velocity and the waveguide propagation axis are perpendicular.

### DC Stark shift

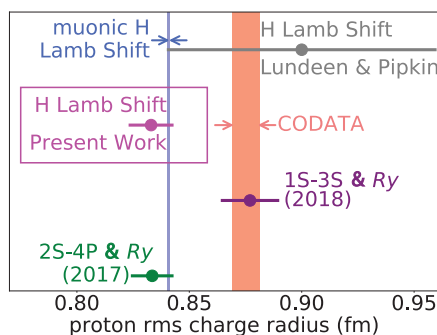
Magnetic fields transverse to the beam velocity are canceled using large coils (to better than 20 mG over the measurement volume) to avoid

a  $10 \text{ kHz}/(\text{V}/\text{cm})^2$  quadratic dc Stark shift for the atoms that the magnetic field would cause when relativistically transformed into the frame of reference of the atoms. One could also be concerned about possible charging of the surfaces of the FOSOF regions by the energetic atoms. However, no shifts are found even when the deflection field is turned off (allowing all of the protons to pass through the FOSOF regions) and when poorer collimation is used, in which case there are direct trajectories for protons and fast atoms to hit the surfaces of the FOSOF regions. As a result, a shift of much less than 1 kHz is expected when the beam is well collimated and free of protons.

### Summary of measurements

A summary of measurements for all 18 parameter sets ( $v$ ,  $d$ ,  $E_{rf}$ ) is shown in Fig. 4A and Table 1. The 18 values show excellent agreement. The error bars in the plot include both statistical uncertainties and the systematic uncertainties discussed in the preceding paragraphs. Averages of measurements for each value of  $d$ ,  $v$ , and  $E_{rf}$  (Fig. 4B) also show excellent consistency. The averages shown are weighted averages with the statistical uncertainties added in quadrature and each type of systematic uncertainty added linearly. The  $|f - f_0|_{\text{max}}$  points in Fig. 4B show that there is no dependence on the range of frequencies used. Ideally, a larger range of frequencies would be used, and we have taken limited confirming data over a larger frequency range. Another FOSOF measurement (13) has tested for possible FOSOF systematics with a better signal-to-noise ratio and shows no dependence on frequency range or any other FOSOF parameter.

Because the separation  $d$  is the most important parameter that we vary in the measurement (in that it has a strong effect on both signal size and line shape), we equally weight the results at the four separations. The individual weights for



**Fig. 5. Summary of proton radius data.**

Shown are values for the proton RMS charge radius from our measurement, muonic hydrogen, CODATA 2014, and the measurements of Beyer *et al.* (18) and Fleurbaey *et al.* (19) combined with that of Parthey *et al.* (20). Also shown in gray is the value from Lundeen and Pipkin (6, 16).

the 18 measurements are shown in Table 1 and minimize the final uncertainty (combined statistical and systematic), subject to the condition that the total weight for each separation is 25%. The last row of Table 1 shows our final measured result

$$f_0^{\text{avg}} = 909.8717(32) \text{ MHz}$$

Here, the uncertainty of 3.2 kHz comes from a combination of a 1.4-kHz statistical uncertainty, a 2.3-kHz uncertainty in the ac Stark shift, a 1.0-kHz uncertainty in the time-dilation correction, and a 1.5-kHz phase measurement uncertainty. The contribution from hyperfine structure to this interval is 147.9581 MHz (17), and correcting for this contribution leads to a Lamb shift of 1057.8298(32) MHz.

### Comparison to other work

Our measurement is lower than the measurement of Lundeen and Pipkin (6)— $f_0^{\text{L\&P}}$  [original] = 909.887(9) MHz—by 1.5 standard deviations. However, our recent reanalysis (16) of their work (using the modeling developed for this work) led to a small shift and larger uncertainties:  $f_0^{\text{L\&P}}$  [reanalyzed] = 909.894(20) MHz, which agrees with the present work.

A value of the proton radius can be deduced from the current measurement (8, 17)

$$r_p[\text{this work}] = 0.833(10) \text{ fm}$$

which is in excellent agreement (Fig. 5) with the muonic hydrogen Lamb shift value but disagrees with the CODATA 2014 value (7).

Two additional measurements in hydrogen that have been published within the past year can also be used to determine the proton radius: a measurement of the 2S→4P interval (18) and a measurement of the 1S→3S interval (19). Both of these measurements require a precise value of the Rydberg constant to determine  $r_p$ . When combined with an existing very precise measurement of the 1S→2S interval (20), they predict the values of  $r_p$  shown in Fig. 5. The values from (18) and (19) disagree.

A combination of our work and the measurement of the muonic hydrogen Lamb shift (9, 10) allows for a direct comparison of measurements of the proton charge radius using the analogous measurements for the muon-based and electron-based determinations. Consistent charge radii are found from the two measurements.

### REFERENCES AND NOTES

1. P. A. M. Dirac, *Proc. R. Soc. London Ser. A* **117**, 610–624 (1928).
2. E. C. Kemble, R. D. Present, *Phys. Rev.* **44**, 1031–1032 (1933).
3. E. A. Uehling, *Phys. Rev.* **48**, 55–63 (1935).

4. W. E. Lamb, R. C. Retherford, *Phys. Rev.* **72**, 241–243 (1947).
5. H. A. Bethe, *Phys. Rev.* **72**, 339–341 (1947).
6. S. R. Lundeen, F. M. Pipkin, *Phys. Rev. Lett.* **46**, 232–235 (1981).
7. P. J. Mohr, D. B. Newell, B. N. Taylor, *J. Phys. Chem. Ref. Data* **45**, 043102 (2016).
8. A. C. Vutha, E. A. Hessels, *Phys. Rev. A* **92**, 052504 (2015).
9. R. Pohl *et al.*, *Nature* **466**, 213–216 (2010).
10. A. Antognini *et al.*, *Science* **339**, 417–420 (2013).
11. R. Pohl, R. Gilman, G. A. Miller, K. Pachucki, *Annu. Rev. Nucl. Part. Sci.* **63**, 175–204 (2013).
12. J. C. Bernauer, R. Pohl, *Sci. Am.* **310**, 32–39 (2014).
13. K. Kato, T. D. G. Skinner, E. A. Hessels, *Phys. Rev. Lett.* **121**, 143002 (2018).
14. N. F. Ramsey, *Phys. Rev.* **76**, 996 (1949).
15. A. Marsman, M. Horbatsch, E. A. Hessels, *Phys. Rev. A* **96**, 062111 (2017).
16. A. Marsman, M. Horbatsch, Z. A. Coriveau, E. A. Hessels, *Phys. Rev. A* **98**, 012509 (2018).
17. M. Horbatsch, E. A. Hessels, *Phys. Rev. A* **93**, 022513 (2016).
18. A. Beyer *et al.*, *Science* **358**, 79–85 (2017).
19. H. Fleurbaey *et al.*, *Phys. Rev. Lett.* **120**, 183001 (2018).
20. C. G. Parthey *et al.*, *Phys. Rev. Lett.* **107**, 203001 (2011).
21. N. Bezginov *et al.*, Data for: A measurement of the atomic hydrogen Lamb shift and the proton charge radius, Harvard Dataverse (2019); doi:10.7910/DVN/MJ2R0C.

### ACKNOWLEDGMENTS

We thank I. Ferchichi, V. Isaac, C. Storry, S. Lundeen, Z. Coriveau, K. Kato, and T. Skinner for contributions to this work. **Funding:** Funding for this work was provided by the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, the Ontario Research Fund, the U.S. National Institute for Standards and Technology, the Canada Research Chair program, Compute Canada, Compute Ontario, and York University. **Author contributions:** N.B., T.V., A.C.V., and E.A.H. contributed to all aspects of the experimental work, including design of the experiment, collection of data, and analysis. M.H., A.M., and E.A.H. contributed equally to computational modeling. **Competing interests:** The authors have no competing interests. **Data and materials availability:** The data collected and the program used for modeling are made publicly available at dataverse.org (21).

### SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/365/6457/1007/suppl/DC1  
Figs. S1 to S3  
Tables S1 to S4

24 July 2018; resubmitted 30 August 2018  
Accepted 30 July 2019  
10.1126/science.aau7807

## A measurement of the atomic hydrogen Lamb shift and the proton charge radius

N. Bezginov, T. Valdez, M. Horbatsch, A. Marsman, A. C. Vutha and E. A. Hessels

*Science* **365** (6457), 1007-1012.  
DOI: 10.1126/science.aau7807

### Unraveling the proton puzzle

The discrepancy between the proton size deduced from the Lamb shift in muonic hydrogen and the average, textbook value based on regular (electronic) hydrogen has puzzled physicists for nearly a decade. One possible resolution could be that electrons interact with protons in a different way than muons do, which would require "new physics." Bezginov *et al.* measured the Lamb shift in electronic hydrogen, which allowed for a direct comparison to the Lamb shift measured in muonic hydrogen. The two results agreed, but the discrepancy with the averaged value remains.

*Science*, this issue p. 1007

#### ARTICLE TOOLS

<http://science.sciencemag.org/content/365/6457/1007>

#### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2019/09/04/365.6457.1007.DC1>

#### REFERENCES

This article cites 20 articles, 2 of which you can access for free  
<http://science.sciencemag.org/content/365/6457/1007#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2019 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works