natureresearch

https://doi.org/10.1038/s41586-020-2964-7

Supplementary information

Determination of the fine-structure constant with an accuracy of 81 parts per trillion

In the format provided by the
authors and unedited

Peer Review File

Manuscript Title: Determination of the fine-structure constant with 81 parts-per-trillion accuracy

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

This is a very important paper that deserves publication in Nature. My compliments to the authors.

A new measurement of the fine structure constant is reported. This fundamental constant is critically important in physics -- perhaps most because it (with the electron magnetic moment measurement) provides the most precise test of the most accurate prediction of the standard model of particle physics.

They are the only research group to have attained the reported accuracy. Their methods are novel and interesting. Their method and report quality seems beyond reproach. Uncertainties seem to be carefully evaluated.

This new result differs significantly from the a result derived from Cs. This is carefully discussed.

Less carefully discussed is that this new result also disagrees with the previous Rb-derived measurement made by the authors of this report. The authors would discuss and hopefully explain their disagreement with themselves. They mention later in one sentence that they had missed a systematic effect in their previous measurement. If possible they should correct their previous result or explicitly say that they retract it, or revise the error bars for the previous measurement. They say that the missed systematic explains the discrepancy but they need to justify this claim more concretely.

The authors describe their "groundbreaking experimental setup." It would be better if they they used a less congratulatory label and left it to readers to decide if it was groundbreaking. Instead they should summarize what features of their apparatus make it better, more reliable and more robust compared to their earlier measurement and those of their competitors. Readers need to understand why this is the measurement that should be trusted.

Mueller used Bragg diffraction (n=5) and Block oscillation with N=125. This paper uses Raman (n=1) and a large Bloch oscillation (N=500). Interference time T is almost same. Sensitivity on h/m is proportional to (n+N). Probably this is the major gain. Is it?

Center of the two interferometer arms are located at the nearly same position. This cancels gravity gradient systematics. This could be a big advantage. Is it?

It is difficult to understand what "We recorded data with different experimental parameters" means. What parameters? Why are these the crucial parameters? How much were they changed? How much did the result vary?

A simulation code "models precisely the underlying physics ..." On what basis is this claim made? Why should a reader believe the claim?

In the paragraph that starts "Table 1 presents our error budget" it would be good to replace vague claims with real numbers. Some examples:

-- gravity gradient cancels out -- to what extent? How is this known.

-- earth's rotation is suppressed -- to what extent? How is this known.

-- beam alignment is corrected in real time -- to what extent, how is this known?

-- ultra-stable cavity -- how stable? how is this known.

-- regularly measured -- how often?

-- utilizing a larger beam size -- larger than what? how large?

This paragraph does not need to be a lot longer, but it needs concrete numbers, and specifics -- rather then vague claims.

p4, after eq (1) "m/m_e is known to better than 0.1ppb…"
• m(Rb)/m_e's uncertainty is 69 ppt. This paper's new uncertainty on h/m is 140ppt. It would be better to write the uncertainty of m(Rb)/m_e explicitly instead of just saying "known to be better".

p5, last paragraph
• How cold is the sample? Maybe some reference about preparation.
• Two Raman pi pulse are used, but their function is not discussed.

P6, above eq (3)
• "inverting Raman beams" is not very clear to me. \delta\omega_R detuning is imposed to which pulse when inverted? Maybe the inverted trajectory can be shown in fig 2 as well.

P7, second paragraph
• Some brief mentioning about the magnetic shield should be included. That was 300 ppt uncertainty in the previous Rb measurement.

P9, forth line from the bottom "Both uncertainties are dominated by a_e(Exp)" This seems to be a major overstatement. The magnetic moment uncertainty is 0.28 ppt and the Cs uncertainty is 0.23 ppt. These are roughly the same. One is not dominated by the other. The new Rb uncertainty is smaller at 0.09ppt.

The authors are to be commended for a wonderful measurement. With clarification of the points raised, this work clearly merits publication in Nature.

Referee #2 (Remarks to the Author):

I strongly recommend publication of the article entitled "Determination of the fine- structure constant below 0.1 part-per-billion and Standard Model prediction of the electron g-factor below 0.1 part-per-trillion". In this paper the authors provide a reliable value of the fine-structure constant with the world-best precision. This one simple fact is sufficient for me why this article should be published in Nature.

The fine-structure constant, denoted by α, is a dimensionless number that governs interactions among charged objects through electro-magnetic force. This means that α concerns almost all phenomena in the universe from elementary-particle physics to condensed-matter physics, chemistry, biology, engineering and many others. Especially, after the new SI for measurement launched in 2019, the fine-structure constant occupies a central role in determination of other fundamental physical constants.

The work done by the authors is measurement of Rb-atom recoil h/m_{Rb} using atomic interferometry. Basics of the experiment was invented by a Nobel laureate S. Chu using Cs atom and a new method precisely determining α was then established. The Stanford Cs-atom measurement was continued by H. Müller and their latest value of α was published in Science 2018 as Ref.[4] . This is currently the world-best value of α.

A French group begun the atomic recoil measurement by using Rb atom. In 2009 and in 2011, their results on α were published as Ref. [21] and as Ref. [13], respectively. The authors of the submitted article to Nature are successors of these experiments. The most remarkable achievement of the new experiment is discovery of a new source of the systematic error. In the error budget table Table.1, this new effect is stated as "Phase shifts in Raman phase lock loop" and entirely explains the 2.4 σ shift of the value from the 2011 result. Unfortunately, the article does not contain any detail of evaluation of this shift. I wish the authors to publish a paper describing more details of the experiment after publication of Nature.

The new result of α obtained by the authors are very interesting. First of all, because of precision achieved by this experiment, many restrictions to beyond-the-Standard- Model of elementary particles physics can be obtained. Together with the Cs-atom result by Ref. [4], these two articles should be recognized as the starting points of new error of elementary-particle physics through low-energy phenomena.

The α value of this paper is consistent with the one obtained from the electron g- factor, but 5 σ different from the one from the Cs-atom experiment. This situation must stimulate not only experimentalists working AMO-physics but also particle theorists pursuing physics beyond the Standard Model.

From the above reasons, this article should be promptly published in Nature.

Below is a list of suggested modifications to be considered by the authors:

Reconsider the title of the paper. I prefer that the title include some words related to the Rb-atomic interferometry and/or Rb recoil h/m in order to make clear distinction from the recent Cs result Ref. [4]. As tests of the Standard Model, three topics, electron g-factor, compositeness of electron, and Be anomaly, are discussed in the article. So, please remove the electron g from the title.

PDF page2, 1st paragraph.

- QED perturbation series are known up to the eighth-order (or four-loop level) and the tenth-order (or five-loop level).

Same place.

- The latest tenth-order result of QED is given in Atoms, 2019, 7(1), 28 by T. Aoyama et al. This gives a slightly different value for ae(Rb2020) in Page 9. This is an insignificant change, but people working in precision physics tend to care about it.

PDF page 4, below Eq. (1).

- Recent two determination of the Rydberg constant has 3 or 4 sigma discrepancy. Please explain why the discrepancy does not affect determination of *α* in Eq. (1).

Same place.

- The measured quantity h/m_{Rb} is immediately converted to the mass of Rb in kg. This is not clear for readers who are not familiar with the new SI. Please let readers remind that *h* is a determined constant.

Same place.

- The statement about (m/m_e) might be misread that the ratio (m/m_e) itself is measured. In fact, the relative atomic masses of $A(m)$ and $A(m_e)$ are measured. This is why AME2016 is quoted in Figure 4 and "Relative mass" is found in the error budget Table 1. Please rephrase the text.

Figure 4, Left.

- The vertical axis must be "Contributions to δa_e"? To avoid misunderstanding, change $\delta a_\mu(m_e/m_\mu)^2$ to $(m_e/m_\mu)^2\delta a_\mu$ as in the text.

Referee #3 (Remarks to the Author):

The paper describes very significant progress in physics, a measurement of the fine-structure constant with an error bar of 81 ppt. This is the most precise determination of this important constant of nature. The paper describes the impact of this measurement on fundamental physics. Notably, it disagrees with previous direct measurements of alpha, including one performed by the same group. The authors also describe constraints on possible dark-sector candidates, including a new analysis for a candidate that has been suggested by the anomalous decay of excited states of beryllium-8.

Unfortunately, it seems that the paper has been submitted without supplementary material. Since this paper will undoubtedly be read and scrutinized widely, and because of the high standards of the precision measurement community, the paper should be exceptionally clear on the systematic error budget. Recent precision measurements in atomic physics of the fine-structure constant and other quantities that have comparable implications for fundamental physics have all rigorously supported their analyses of systematic effects with detailed discussions, see, e.g., [1] followed by [2], and [3-7]. Every entry in the systematic error budget must be justified with formulas and data that allow an independent assessment of the plausibility of the result. This information can, e.g., be included as supplementary material. For example:

1. Relevant quantities specifying the atom interferometer geometry should be given, such as the pulse separation time, the number and timing of Bloch oscillations, and the details on how these parameters were varied during data taking.

2. The error associated with the Monte-Carlo simulation used to evaluate systematic effects should be specified. The authors should describe how they have verified that the simulation accurately reproduces the behavior of the experiment when parameters are varied.

3. The authors rightfully point out that they have identified the problem of small-scale variations of the laser wavevector. This important point should be explained in detail. How have they characterized the laser beam, and how has the resulting correction and its error been calculated?

The manuscript should show appropriate data (perhaps beam images as well as derived quantities such as the root-mean-square intensity variation and its typical length scale, or the autocorrelation function) that allow the reader to reproduce the calculation.

4. How have the authors made sure that the vacuum viewports do not introduce additional small-scale variations, especially since the atoms are located quite closely to the viewports in the experiment, giving the intensity variations little space to disperse?

5. The largest correction in the error budget, 108.2 +/- 5.4 parts in 10^11, is from the Gouy phase. Have the authors assumed a Gaussian beam, and how have they verified the shape of the beam? Neither optical fibers nor apodizing filters deliver particularly good Gaussian beams. How has the beam waist radius and its location been measured?

6. The authors should show equations for the interferometer phase arising from systematic effects such as the Coriolis acceleration and/or the Sagnac effect as well as gravity gradients, and explain how they measured the parameters entering these equations to arrive at the numbers stated in the error budget.

8. How has the quasi-thermal motion of the atoms been characterized, e.g., the distribution of the initial positions and velocities, and how do these numbers enter the error budget?

9. Can they specify the nature of the "phase shift in the Raman phase lock loop" by describing this loop in sufficient detail to allow an independent assessment of the induced error bar?
Can the authors specify the nature of the magnetic shield they used, and show data how the effect of residual fields was characterized?

The above list may not be complete - the authors should make every effort to support every entry in the error budget at a level where it can be independently scrutinized.

In conclusion, this work shows every sign of a landmark result, but the above information is required for an informed assessment of the paper. I recommend asking the authors to respond before the paper can be reviewed in detail.

A few minor remarks:

The paper is very well-written.

Plotting an Alan variance to integration times more than 1/10 or 1/3 of the total data taking time is not very meaningful. This should be corrected.

There should be a citation for the Rb and electron mass, so that readers in 10 years (that will undoubtedly exist!) will know the source for these values

[1] D. Hanneke, S. Fogwell, and G. Gabrielse, PRL 100, 120801 (2008)
[2] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse, Phys. Rev. A 83, 052122 (2011)
[3] ACME Collaboration, Nature 562, 18 (2018)
[4] ACME Collaboration, Science 343, 17 (2014)
[5] N. Bezginov, T. Valdez, M. Horbatsch, A. Marsman, A. C. Vutha, and E. A. Hessels, Science 365, 1007–1012 (2019)
[6] A. Beyer, L. Maisenbacher, A. Matveev, R. Pohl, K. Khabarova, A. Grinin, T. Lamour, D. C. Yost, T. W. Hänsch, N. Kolachevsky, T. Udem, Science 358, 79–85 (2017)
[7] W. B. Cairncross, D.l N. Gresh, M. Grau, K. C. Cossel, T. S. Roussy, Y. Ni, Y. Zhou, J. Ye, and E. A. Cornell, PRL 119, 153001 (2017)

Author Rebuttals to Initial Comments:

Responses to the Referees:

We warmly thank the three referees for their very pertinent remarks and suggestions which were useful to us in improving the understanding and clarity of our work.
All their remarks converge towards the need to add a section detailing the experimental design and protocol as well as the methods we used to reduce and evaluate systematic effects. To this end, we have added a Methods section where we hope to have answered all their questions. We have also taken their comments into account and added factual information in the main text.
In the following, we answer to the referees remarks.
Questions from the referees are in italic. Our answers are in blue.

Referee 1 (Remarks to the Author):
Less carefully discussed is that this new result also disagrees with the previous Rb-derived measurement made by the authors of this report. The authors would discuss and hopefully explain their disagreement with themselves. They mention later in one sentence that they had missed a systematic effect in their previous measurement. If possible they should correct their previous result or explicitly say that they retract it, or revise the error bars for the previous measurement. They say that the missed systematic explains the discrepancy but they need to justify this claim more concretely.
The problem we have identified on the phase lock loop is a possibility to explain the difference between the two measurements. It is not possible to evaluate this effect on the ten years old experiment since the PPL that has been changed.
One of the new systematic concerns the correlation between local intensity and recoil. This effect could not explain the discrepency as the shift would be the other way.
To be honest we have no clear explanation for this discrepancy. Our new measurements makes the 2011 results de facto obsolete, because of the dif- ference of uncertainties and it is not worth to reevaluate the systematics on the 2011 measurement.

The authors describe their "groundbreaking experimental setup." It would be better if they they used a less congratulatory label and left it to readers to decide if it was groundbreaking. Instead they should summarize what features of their apparatus make it better, more reliable and more robust compared to their earlier measurement and those of their competitors. Readers need to understand why this is the measurement that should be trusted.
We agree with your remark. We changed groundbreaking for "new". New features of the setup are the robustness and reliability of the design which we describe in detail in the Methods. The main difference with the previous set-up is that the atomic interferometer is realized in a separate and longer area where the magnetic field is perfectly controlled.

Mueller used Bragg diffraction (n=5) and Block oscillation with N=125. This paper uses Raman (n=1) and a large Bloch oscillation (N=500). Interference time T is almost same. Sensitivity on h/m is proportional to (n+N). Probably this is the major gain. Is it?
The sensitivity to h/m in the Berkeley experiment scales as n(n+N), so their theoretical sensitivity is actually higher than ours. We think the major gain in our experiment arises from a better signal to noise provided by the Raman transition. Furthermore Bragg diffraction requires a precise calibration of a phase offset that limits the statistics.

Center of the two interferometer arms are located at the nearly same position. This cancels gravity gradient systematics. This could be a big advantage. Is it?
Yes.
Gravity gradient is a relatively large systematics on the Berkeley experiment. While the theoretical

calculation of the correction is simple, a precise mea- surement of the gradient is tricky in a lab environment. Our method is a clear advantage.

It is difficult to understand what "We recorded data with different experi- mental parameters" means. What parameters? Why are these the crucial parameters? How much were they changed? How much did the result vary?
We have added details in the Methods section. In figure 3C of the main text, data with different parameters are shown. They are corrected for systematic effect and are consistent.

A simulation code "models precisely the underlying physics ..." On what basis is this claim made? Why should a reader believe the claim?
In the paragraph that starts "Table 1 presents our error budget" it would be good to replace vague claims with real numbers. Some examples:

- *gravity gradient cancels out – to what extent? How is this known.*
- *earth's rotation is suppressed – to what extent? How is this known.*
- *beam alignment is corrected in real time – to what extent, how is this known?*
- *ultra-stable cavity – how stable? how is this known.*
- *regularly measured – how often?*
- *utilizing a larger beam size – larger than what? how large?*

This paragraph does not need to be a lot longer, but it needs concrete numbers, and specifics – rather then vague claims.
We have added in the main text key factual informations on the issues raised above by the referee 1. We have also detailed in the Methods the evaluation of the different systematic effects. Details and references for the model used by the Monte Carlo simulation are also detailed in the Methods.

p4, after eq (1) m/m_e is known to better than 0.1ppb. . . " $m(Rb)/m_e$'s uncer- tainty is 69 ppt. This paper's new uncertainty on h/m is 140ppt. It would be better to write the uncertainty of $m(Rb)/m_e$ explicitly instead of just saying "known to be better".
We have taken into account your remark and wrote explicitly the uncertainty of both relative masses.

p5, last paragraph
How cold is the sample? Maybe some reference about preparation.
In the section "Experimental setup" of the Methods we describe the prepa- ration of atomic sample as well as its features.

Two Raman pi pulse are used, but their function is not discussed.
In the section "Experimental sequence" of the Methods we explain the function of the two π Raman pulses in between the Bloch pulses of the atomic elevator.

P6, above eq (3)
"inverting Raman beams" is not very clear to me. $\delta\omega_R$ detuning is imposed to which pulse when inverted? Maybe the inverted trajectory can be shown in fig 2 as well.
By inverting Raman beams, we mean inverting the Raman beam direction. This now clear in the paper. For a Raman transition (by opposition to a Bragg diffraction), the two lasers used to drive the transition have a large frequency difference and originate from two independent lasers. The inversion is obtained by changing the laser which is retroreflected. It is not possible to change the direction of the recoil by changing the Raman frequency.
We decided to not modify Fig 2C : superposing the four spectra on the same plot would reduce the clarity of the picture.
Explanation on the inversion of the Raman beam are given in the method and on the extended data

figure 1 a.

P7, second paragraph
Some brief mentioning about the magnetic shield should be included. That was 300 ppt uncertainty in the previous Rb measurement..
It is done in the Methods. Yes in the previous Rb measurement the uncer- tainty due to the 2nd order Zeeman effect was 300 ppt,

P9, forth line from the bottom "Both uncertainties are dominated by a_e(Exp)
This seems to be a major overstatement. The magnetic moment uncertainty is 0.28 ppt and the Cs uncertainty is 0.23 ppt. These are roughly the same. One is not dominated by the other. The new Rb uncertainty is smaller at 0.09ppt.
We agree with your remark. We change the sentence by : The uncertainty on δa_e is now dominated by a_e(Exp).

Referee 2 (Remarks to the Author):
Below is a list of suggested modifications to be considered by the authors:

1. *Reconsider the title of the paper. I prefer that the title include some words related to the Rb-atomic interferometry and/or Rb recoil h/m in order to make clear distinction from the recent Cs result Ref. [4]. As tests of the Standard Model, three topics, electron g-factor, compos- iteness of electron, and Be anomaly, are discussed in the article. So, please remove the electron g from the title.*

Thank you ! We have changed the title in accordance with your recommendations.

2. *PDF page2, 1st paragraph. QED perturbation series are known up to the eighth-order (or four-loop level) and the tenth-order (or five-loop level).*

Thank you for you comment.

3. *Same place. The latest tenth-order result of QED is given in Atoms, 2019, 7(1), 28 by T.Aoyama et al. This gives a slightly different value for a_e(Rb2020) in Page 9. This is an insignificant change, but people working in precision physics tend to care about it.*

All the calculations were actually performed using values from this reference. A mistake occurred in the citation. Thank you for your remark.
To simplify, we are now citing the first paper on the 10th order calcula- tion (Aoyama Phys. Rev.Lett.109, 111807 (2012)) and the latter that you mentioned.

4. *PDF page 4, below Eq. (1). Recent two determination of the Ryd- berg constant has 3 or 4 sigma discrepancy. Please explain why the discrepancy does not affect determination of α in Eq. (1).*

The value of the Rydberg constant used to evaluate our value of α is from the web site: https://physics.nist.gov/cuu/Constants/index.html. It is the recent adjustment of the Rydberg constant made by CODATA 2018 which considers the discrepancy mentioned above and takes into account recent results from regular and muonic hydrogen.
Moreover, the precision with which the Rydberg constant is known is one to two order of magnitude better than for the relative masses or h/m. The impact of this discrepancy would be negligible. We therefore preferred to use the new adjustment of CODATA 2018 which will be published soon.

5. *Same place. The measured quantity h/m_{Rb} is immediately converted to the mass of Rb in kg.*

This is not clear for readers who are not famil- iar with the new SI. Please let readers remind that h is a determined constant.

Now we wrote : In the international system f units adopted in 2019 where h has a fixed value, we obtain $m87_{Rb}$ = ...

6. *Same place. The statement about (m/m_e) might be misread that the ratio (m/m_e) itself is measured. In fact, the relative atomic masses of A(m) and A(me) are measured. This is why AME2016 is quoted in Figure 4 and "Relative mass" is found in the error budget Table 1. Please rephrase the text.*

Thanks to the referee for this remark. We rephrased the text as follow: The atom-to-electron mass ratio m/m_e is obtained from the ratio the relative atomic masses $A_r(m)$ of the atom (known at 69 ppt for rubid- ium and the relative atomic masse $A_r(m_e)$ of the electron (known at 30 ppt).

7. *Figure 4, Left. The vertical axis must be "Contributions to δa_e"? To avoid misunderstanding, change $\delta a_\mu(me/m_\mu)^2$ to $(m_e/m_\mu)^2\delta a_\mu$ as in the text.*

Of course it is more clear, thank you, its done

Referee 3 (Remarks to the Author):

Unfortunately, it seems that the paper has been submitted without supplementary material. Since this paper will undoubtedly be read and scrutinized widely, and because of the high standards of the precision measurement com- munity, the paper should be exceptionally clear on the systematic error budget. Recent precision measurements in atomic physics of the fine-structure con- stant and other quantities that have comparable implications for fundamental physics have all rigorously supported their analyses of systematic effects with detailed discussions, see, e.g., [1] followed by [2], and [3-7]. Every entry in the systematic error budget must be justified with formulas and data that allow an independent assessment of the plausibility of the result. This information can, e.g., be included as supplementary material. For example:

1. *Relevant quantities specifying the atom interferometer geometry should be given, such as the pulse separation time, the number and timing of Bloch oscillations, and the details on how these parameters were varied during data taking.*
2. *The error associated with the Monte-Carlo simulation used to evaluate systematic effects should be specified. The authors should describe how they have verified that the simulation accurately reproduces the behavior of the experiment when parameters are varied.*
3. *The authors rightfully point out that they have identified the problem of small-scale variations of the laser wavevector. This important point should be explained in detail. How have they characterized the laser beam, and how has the resulting correction and its error been calculated? The manuscript should show appropriate data (perhaps beam images as well as derived quantities such as the root-mean-square intensity variation and its typical length scale, or the autocorrelation function) that allow the reader to reproduce the calculation.*
4. *How have the authors made sure that the vacuum viewports do not introduce additional small-scale variations, especially since the atoms are located quite closely to the viewports in the experiment, giving the intensity variations little space to disperse?*

We use fused silica viewports mounted in non-magnetic flange. We also measured the laser beam profile using a camera CCD after the viewport. We even tested different viewports. We did not observe any noticeable differences with and without the viewport when analyzing the images.

5. *The largest correction in the error budget, 108.2 +/- 5.4 parts in 10^{11}, is from the Gouy phase.*

Have the authors assumed a Gaussian beam, and how have they verified the shape of the beam? Neither optical fibers nor apodizing filters deliver particularly good Gaussian beams. How has the beam waist radius and its location been measured?

6. *The authors should show equations for the interferometer phase aris- ing from systematic effects such as the Coriolis acceleration and/or the Sagnac effect as well as gravity gradients, and explain how they mea- sured the parameters entering these equations to arrive at the numbers stated in the error budget.*
7. *How has the quasi-thermal motion of the atoms been characterized, e.g., the distribution of the initial positions and velocities, and how do these numbers enter the error budget?*
8. *Can they specify the nature of the "phase shift in the Raman phase lock loop" by describing this loop in sufficient detail to allow an indepen- dent assessment of the induced error bar? Can the authors specify the nature of the magnetic shield they used, and show data how the effect of residual fields was characterized?*
9. *Can they specify the nature of the "phase shift in the Raman phase lock loop" by describing this loop in sufficient detail to allow an indepen- dent assessment of the induced error bar? Can the authors specify the nature of the magnetic shield they used, and show data how the effect of residual fields was characterized?*

The above list may not be complete - the authors should make every effort to support every entry in the error budget at a level where it can be independently scrutinized.
We have added a Methods section where we answer to the nine issues raised by the referee: 1/ Detailed description of the experimental setup, 2/ A de- tailed description of interferometer sequence, 3/ Details on the Monte Carlo simulation, 4/ Detailed description of evaluation of our systematic effects.

Plotting an Alan variance to integration times more than 1/10 or 1/3 of the total data taking time is not very meaningful. This should be corrected.
Thank you for the remark. We have removed these points from the plot.

There should be a citation for the Rb and electron mass, so that readers in 10 years (that will undoubtedly exist!) will know the source for these values
This is done. Moreover, we have added the value in the table of the error budget.

D. Hanneke, S. Fogwell, and G. Gabrielse, PRL 100, 120801 (2008)
D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse, Phys. Rev. A 83, 052122 (2011)
ACME Collaboration, Nature 562, 18 (2018)
ACME Collaboration, Science 343, 17 (2014)
N. Bezginov, T. Valdez, M. Horbatsch, A. Marsman, A. C. Vutha, and E. A. Hessels, Science 365, 1007–1012 (2019)
A. Beyer, L. Maisenbacher, A. Matveev, R. Pohl, K. Khabarova, A. Grinin, T. Lamour, D. C. Yost, T. W. Ha¨nsch, N. Kolachevsky, T. Udem, Science 358, 79–85 (2017)
W. B. Cairncross, D.l N. Gresh, M. Grau, K. C. Cossel, T. S. Roussy, Y. Ni, Y. Zhou, J. Ye, and E. A. Cornell, PRL 119, 153001 (2017)

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

I commend the authors for their thoughtful responses and their improved manuscript. The added "Methods" is a very useful addition.

With what is now provided, readers will have two major questions:

1. Given that there are now two "atom recoil" measurements that differ by much more than the uncertainties assigned, why should this measurement be believed rather than the other?

2. Is the major message of this work to demonstrate that "atom recoil" determination of alpha are unreliable, or is the major message about less discrepancy between the SM prediction and measurement.

I hope that the authors feel that they have done all that they can to address these two issues.

I do wish that several of the thoughtful comments to my referee suggestions had made it into the manuscript in a paragraph that readers would find helpful. For example,

---"The main difference with the previous setup is that the atom interferometer is realized in a separate and long area where the magnetic field is" very carefully controlled. ("perfectly" would be too strong.)

---Although other experiments have a potentially higher sensitivity, "the major gain in our experiment arises from a better signal-to-nois provide by the Raman transition. Furthermore, Bragg defraction required a precise calibration in a phase offset that limits the statistics."

---One nice feature of our method is that there is no need to measure a gravity gradient.

It is true that if this result and the uncertainty are believed, then the uncertainty in the comparison of the SM prediction and measurement is due to the electron measurement uncertainty. This appropriate point is made strongly.

However, the point that may trouble readers is that there are now three different "atom recoil" determination of the fine structure constant that disagree a great deal. One interpretation is that there is one or more uncertainties in these types of measurements that have not yet been recognized. If this is so, what limits the comparison of the SM prediction and the electron measurement is in fact is systematic uncertainties in the atom recoil measurements -- not the magnetic moment measurements.

I note that the change in the manuscript from "This explains" to "This could explain" is very significant, even a bit troubling at this point. From the initial manuscript I thought the the authors understood why their new result differed so much from their old one. Based upon this I suggested that they be a bit clearer and more specific to help readers like me understand how to weight the two results.

Now the manuscript says "This could explain" which is very different. I thank the authors for their clear statement in their response, namely "To be honest we have no clear explanation for this discrepancy. Our new measurements makes the 2011 results de facto obsolete, because of the difference
of uncertainties and it is not worth to reevaluate the systematics on the 2011 measurement." I wish that some form of this had made it into the revised manuscript.

IN summary

The revised manuscript is much more worthy of this wonderful result. It certainly should be published. If the authors choose to address the points raised above, the manuscript would stand the test of time better. I hope that they believe that they have done all that they can to address these points.

Referee #2 (Remarks to the Author):

The revised version of the paper with a new title “New determination of the fine-structure constant from the recoil of a Rb atom” must be published in Nature without any delay. In this version, the authors describe any details of their experiments. They are clearly explained with more extended data.
One request is that the reference [12] must be cited in the second line from the bottom of page 10 for the sentence “our previous measurement”. Because the 2.4 sigma sift from the previous measurement is one of the most curious part of this work.
Congratulations on the authors’ achievement.

Referee #3 (Remarks to the Author):

I applaud the authors for a mostly well-done measurement. The new materials that have been submitted show the great effort that went into the error budget. Nevertheless, I have a number questions and concerns that should be addressed:

1. Beam-related systematics should be determined from measurements and not just guessed, especially because they constitute the largest systematic shift by far.

a. For example, the authors did "not precisely control the overlap between the laser beams” and instead just assumed an uncertainty of +/-500 microns. In absence of data, I doubt that the shift can be estimated that well; for example, translation of a Gaussian beam of 4.9 mm waist parallel to itself by 500 microns reduces the overlap integral with another beam, or the fiber-coupling efficiency, by only 0.5%, which is noticeable only in a careful measurement. The authors should present their measurement of this shift and describe how much it affects the error budget.

b. Many systematic effects depend on the position or velocity of atoms, but the authors "don’t control the precise initial position of the atomic cloud". The authors have performed simulations where the overlap of the beams between themselves and with the atomic cloud are varied. The authors should present their measurement of the magnitude of these variations that underlies these simulations.

c. The Raman and the second Bloch beams are coming from two separate fibers. The authors should describe how they have established that the beam waists and their locations are matched longitudinally and transversely and how this affects the systematic errors.

In absence of such data, some contributions to the beam-related systematics seem to have come out of thin air. It is particularly important to address this since they could potentially affect many terms in the error budget.

2. The authors increase the Raman intensity with time to compensate for the motion of atoms out of the beam. However, atoms in the middle of the beam are now illuminated with a higher intensity, and thus experience a different AC Stark shift than atoms further away. If the atoms are not all detected with the same probability, how has this been taken into account for the error estimate?

3. According to extended data table 1, the authors have used only two values of TR and NB. How much data has been taken at each value, and in which sequence? Is this sequence suitable for detecting and eliminating variations of the light-shift phase over the period in which the data is taken, or similar phase offsets?

4. While I agree that the measurement in Fig. 4 conclusively estimates the effect of small-scale intensity fluctuations for Bloch oscillations, would the Raman beam splitters suffer from a similar effect? What might be the cause for the extra 2% intensity modulations at 50 micron length scale that they have to add to make the model agree with the measurement? Would one get a different result if one added a different combination of amplitude fluctuation magnitude and length scale?

5. In detection, the authors don't select atoms in the center of the atom cloud, they image all of them. Also, their detection scheme pushes the atoms around before detecting the total population. Could this result in a systematic effect? They didn't mention any Monte Carlo simulation looking at these effects.

6. Based on their phase equation, they have a maximum of 2 Mrad of phase. The quoted error in the fine structure constant corresponds to only 0.3 mrad, so the method by which the fluorescence signals are analyzed needs to be checked at that level of precision. Can the authors describe how the fringes were fitted?

7. Similar to beam-induced errors, the microwave phase errors add a large offset that has to be subtracted. The authors offer little explanation, other than saying that it may be a "nonlinearity" in the phase detector. Can the authors check if this could be due to the group delay dispersion of the low-pass filter? The authors characterize such phase errors by an oscilloscope and note that changing the oscilloscope bandwidth to 20 MHz does not introduce additional shifts larger than 0.01 ppb in h/m. In absence of a quantitative understanding of the effect, should this be added as an error bar to the measurement?

8. Can the authors amend Fig. 3 or the description in the text to show where the ramp of 25 MHz/s is applied, and can they state the frequencies (in MHz) of the signals after the mixers? I am asking because these quantities enter the phase shift caused by group delay dispersion.

Minor comments

Apodizing filters are known to result in non-Gaussian beams. From the text in the supplemental material, I think this has been taken into account as the Monte-Carlo simulation used actual images of the beam. Is this true, and if so, can this be explicitly stated?

Author Rebuttals to First Revision:

Responses to the Referees:
In the following, we answer to the referees remarks.
Questions from the referees are in italic. Our answers are in blue.

Referee 1 (Remarks to the Author):

1. *Given that there are now two "atom recoil" measurements that differ by much more than the uncertainties assigned, why should this measurement be believed rather than the other?*
2. *Is the major message of this work to demonstrate that "atom recoil" determination of alpha are unreliable, or is the major message about less discrepancy between the SM prediction and measurement.*

I hope that the authors feel that they have done all that they can to address these two issues.
After our measurement in 2011, we have identified several key points. We built a new experimental setup and improved the modelling of several system- atics. We believe that we have properly studied all the systematic effects and did a detailed description of their evaluation in the paper. We are confident in this new result.
No reason has been identified to explain the difference with Berkeley caesium measurement. Nevertheless, all that can be said is that we have much higher sensitivity (about 3 times), which has allowed us to experimentally study systematic effects.
The message of this work is that we did a new determination of the fine structure constant with an improved uncertainty.

I do wish that several of the thoughtful comments to my referee suggestions had made it into the manuscript in a paragraph that readers would find help- ful. For example,
—"The main difference with the previous setup is that the atom interferom- eter is realized in a separate and long area where the magnetic field is" very carefully controlled. ("perfectly" would be too strong.)
You are right, we have added on page 16: "One main difference with the previous setup is that the atom interferometer is realized in a separate and long tube where the magnetic field is precisely controlled thanks to a uniformly wound solenoid shielded by two layers of µ-metal."

—Although other experiments have a potentially higher sensitivity, "the ma- jor gain in our experiment arises from a better signal-to-noise provide by the Raman transition. Furthermore, Bragg diffraction required a precise calibra- tion in a phase offset that limits the statistics."
We did not include this remark in the manuscript because we used the same method as in 2011. The major gain in sensitivity is not due to the Raman technique, which certainly gives a better signal to noise than Bragg diffrac- tion, but to several improvements in the experimental set-up.

—One nice feature of our method is that there is no need to measure a gravity gradient.
We added in page 23:
"We use the atomic elevator to position the atoms in the interferometer zone. The nice feature of this method is that we can choose v_0 and z_0 so that the average position for the four spectra (E_R, E_B) is the same, thus mitigating the effect of the gravity gradient. "

It is true that if this result and the uncertainty are believed, then the uncer- tainty in the comparison of the SM prediction and measurement is due to the electron measurement uncertainty. This appropriate point is made strongly.
However, the point that may trouble readers is that there are now three differ- ent "atom recoil" determination of the fine structure constant that disagree a great deal. One interpretation is that there is one or more uncertainties in these types of measurements that have not yet been recognized. If this is so, what limits the comparison of the SM prediction and the electron measure- ment is in fact is systematic uncertainties in the atom recoil measurements
– not the magnetic moment measurements.
I note that the change in the manuscript from "This explains" to "This could explain" is very significant, even a bit troubling at this point. From the initial manuscript I thought the the authors understood why their new result differed so much from their old one. Based upon this I suggested that they be a bit clearer and more specific to help readers like me understand how to weight the two results.
Now the manuscript says "This could explain" which is very different. I thank the authors for their clear statement in their response, namely "To be honest we have no clear explanation for this discrepancy. Our new measurements makes the 2011 results de facto obsolete, because of the difference of uncertainties and it is not worth to reevaluate the systematics on the 2011

measurement." I wish that some form of this had made it into the revised manuscript.
We affirm that the new systematic effects identified in this work have not been taken into account in the uncertainty budget of 2011 measurement. Unfortunately, today we do not have any data available, or even the possi- bility of re-evaluating these effects retrospectively, so we cannot state with certitude that these effects explain the discrepancy between these two mea- sures. To emphasize this point and for the sake of honesty, we have replaced "Explain" with "Could explain". To take into consideration your comment, we have added explicitly in the manuscript (page 11): This could explain the 2.4σ discrepancy between our new measurement and $\alpha_{LKB2011}$. Unfortunately we do not have available data to evaluate retrospec- tively the contributions of the phase shift in the Raman phase lock loop and of short scale fluctuations of laser intensity to the 2011 measurement. Thus, we cannot firmly state that these two effects are at the cause of the 2.4 σ discrepancy between our two measurements.

In summary
The revised manuscript is much more worthy of this wonderful result. It certainly should be published. If the authors choose to address the points raised above, the manuscript would stand the test of time better. I hope that they believe that they have done all that they can to address these points.

Referee 2 (Remarks to the Author):
The revised version of the paper with a new title "New determination of the fine-structure constant from the recoil of a Rb atom" must be published in Nature without any delay. In this version, the authors describe any details of their experiments. They are clearly explained with more extended data. One request is that the reference [12] must be cited in the second line from the bottom of page 10 for the sentence "our previous measurement". Because the sigma sift from the previous measurement is one of the most curious part of this work. Congratulations on the authors' achievement.
Thank you, it's done.

Referee 3 (Remarks to the Author):

1. *Beam-related systematics should be determined from measurements and not just guessed, especially because they constitute the largest systematic shift by far.*
   We strongly refute and object the use of the term "guessed". Our team has been working for more than 30 years in the field of high-precision measurements. We have never, but never "guessed" a systematic effect. For example, we have worked for three years to understand how to take into account the effect of wave-front distortions on recoil measurement. When we experimentally observed a systematic effect that depends on the efficiency of Bloch oscillations (see S. Bade et al. Phys. Rev. Lett. 121, 073603 (2018).), we did not guess the correction, but we took the time to understand and properly evaluate the correction and the uncertainty of this effect. The model we developed was used by Holger Muller's team to estimate the correction on the measure of cesium recoil.

   - *For example, the authors did "not precisely control the overlap between the laser beams" and instead just assumed an uncertainty of +/-500 microns. In absence of data, I doubt that the shift can be estimated that well; for example, translation of a Gaussian beam of 4.9 mm waist parallel to itself by 500 microns reduces the overlap integral with another beam, or the fiber-coupling efficiency, by only 0.5, which is noticeable only in a careful measurement. The authors should present their measurement of this shift and describe how much it affects the error budget.*

It is obvious that the overlap between the Raman and Bloch beams from the top fiber and the Bloch beam from the bottom fiber has been carefully controlled. This is also the case concerning the overlap of the atomic cloud and the laser beams. By writing "In the experiment we do not precisely control the overlap between the laser beams (only the direction is well controlled), so we took an uncertainty of ±500 µm for this overlap. We also don't control the precise initial position of the atomic cloud with respect to the axis of the laser beam. Simulations were performed with different sets of parameters of the overlap of the beams between themselves and with the atomic cloud.", we wanted to say that the precision of the overlap, of the order of 10% of the waist, is not as precise as the beam alignment (less than 10^{-5}) and we wanted also to say that we have not implemented an automatic procedure as is the case for beam alignment (checked every 45 minutes). We made the overlap adjustments at the beginning of the measurement and checked them whenever we observed a drop in fringe contrast or in the number of atoms. We have modified the text (see pages 25-26) to be more explicit; we also provide more details on our alignment procedures.
We hope this answers the questions 2a, b and c.

- *Many systematic effects depend on the position or velocity of atoms, but the authors "don't control the precise initial position of the atomic cloud". The authors have performed simulations where the overlap of the beams between themselves and with the atomic cloud are varied. The authors should present their measurement of the magnitude of these variations that underlies these simulations.*
  We agree that many systematics depend on the atomic trajectory. However, in the experiment the cloud expands and at the begin- ning of the interferometer the width of the cloud is larger than the waist of the laser beams. The distribution of the position of atoms that contributes to the signal is determined mainly by the laser beam and not the initial position of the cloud. Indeed, the survival probability of the atoms at the edge of the Bloch beam is too low and these atoms are lost during the atomic elevator step. So the Bloch oscillations act as a spatial filter that shapes the transverse distribution of the laser beam.
  Therefore, the average correction of those systematics is not very sensitive to the initial position. This was confirmed by the Monte Carlo simulation.
  We did not measure the magnitude of these variations but evaluated the uncertainty of the alignment procedure (see below) and use it in the simulation.

- *The Raman and the second Bloch beams are coming from two sep- arate fibers. The authors should describe how they have established that the beam waists and their locations are matched longitudinally and transversely and how this affects the systematic errors.*
  We added in page 25-26 of manuscript: In the experiment, the precision on the overlap of the laser beams is less than 10% of the waist. This is not as precise as the beam alignment (less than 10^{-5}). We have not implemented an automatic procedure for this less sensitive alignment.
  The adjustments of the overlap between the upward and down- ward beams and the overlap between the atomic cloud and the laser beams were carefully made at the beginning of the measure- ment and re-checked if we observe a drop in the contrast of atomic fringes or in the number of atoms. Our method for controlling beam overlap consists in using, during the alignment procedure described above, a razor blade mounted on a micro-metric trans- lation. This technique usually used to estimate the waist of laser beam allows us to superpose the centers of laser beams at better than 10%, so we took an uncertainty of ±500 µm for this overlap. To align the laser beam

overlap with the atomic cloud, we use Rabi oscillations with the two co-propagating Raman beams (downward Raman beam). The procedure consists in minimizing the pulse duration that satisfies the π condition by laterally displacing the laser beams (the fiber is placed on a micro-metric translation). Of course, at each lateral position, the injection of the beams into the fibers is re-optimized.

In absence of such data, some contributions to the beam-related systematics seem to have come out of thin air. It is particularly important to address this since they could potentially affect many terms in the error budget.

2. *The authors increase the Raman intensity with time to compensate for the motion of atoms out of the beam. However, atoms in the middle of the beam are now illuminated with a higher intensity, and thus experience a different AC Stark shift than atoms further away. If the atoms are not all detected with the same probability, how has this been taken into account for the error estimate?*
   This intensity ramp was determined experimentally so that the velocity dependence phase shift observed on each spectra is cancelled. This optimisation was performed with the same experimental condition as in the measurement and therefore takes into account the efficiency of the detection.
   Note also that the goal of the intensity ramp is not only to compensate the AC Stark shift (which is also compensated by the inversion of the Raman direction and the Bloch direction) but to remove the velocity dependence on the shift at the level of individual spectra.

   In response to your relevant remark, in the manuscript page 28, we replaced ”This number has been calibrated through the identification of a velocity-dependent phase shift during Raman transitions induced by the imbalance of average intensity between the Raman pulses.”
   By
   ”This number corresponds to the laser intensity ramp that reduces the velocity-dependent phase shift during Raman transitions, induced by the unbalance of the average internal intensity between Raman pulses. Calibration of the slope of the intensity ramp was performed under the same experimental conditions as those of the measurement of the ratio h/m in order to ensure the same detection efficiency.”

3. *According to extended data table 1, the authors have used only two val- ues of TR and NB. How much data has been taken at each value, and in which sequence? Is this sequence suitable for detecting and eliminating variations of the light-shift phase over the period in which the data is taken, or similar phase offsets?*
   As shown in figure 3C, the value of h/m is determined from a set of data obtained with four sets of (T_R, N_B, τ_B). The more sensitive configuration (20ms, 500 BOs, 12 µs) was used as a reference. Other configurations were always taken together with the reference one by alternating randomly the configurations after each spectra (typically every 5 minutes). Long term variations of the values over the measurement period have not been observed (cf. Allan variance). The different values of T_R and N_B were not used to eliminate the light-shift but to check the consistency of the error budget.

4. *While I agree that the measurement in Fig. 4 conclusively estimates the effect of small-scale intensity fluctuations for Bloch oscillations, would the Raman beam splitters suffer from a similar effect? What might be the cause for the extra 2% intensity modulations at 50 micron length scale that they have to add to make the model agree with the measurement? Would one get a different result if one added a different combination of amplitude fluctuation*

magnitude and length scale?
The simulation account for the small scale fluctuations for both Raman transitions and Bloch oscillations.
We have no a confirmed reason to explain the additional noise that should be added to fit the experiment. We believe that it comes from the viewport of the vacuum chamber.
In Bade et al., Phys. Rev. Lett.121, 073603 (2018), we show that there is a universal behaviour of the shift that depends on the efficiency of Bloch oscillation and on σ/l^2, where σ is the RMS relative intensity noise and l is the correlation length. Therefore the interpolation at the working intensity of the shift measured at low intensity is robust and independent of the choice of σ and l.

5. *In detection, the authors don't select atoms in the center of the atom cloud, they image all of them. Also, their detection scheme pushes the atoms around before detecting the total population. Could this result in a systematic effect? They didn't mention any Monte Carlo simulation looking at these effects.*
   No. With our experiment based on Raman transitions, there is no large bias related to the detection scheme at the contrary of experiment based on Bragg diffraction. We did not perform any Monte Carlo simulation on the detection scheme.

6. *Based on their phase equation, they have a maximum of 2 Mrad of phase. The quoted error in the fine structure constant corresponds to only 0.3 mrad, so the method by which the fluorescence signals are analyzed needs to be checked at that level of precision. Can the authors describe how the fringes were fitted?*
   We do not have 2 Mrad of phase because the phase shift induced by acceleration of atoms is compensated by the frequency shift $\delta\omega_R$ of the Raman beams.
   We use a regular least square method to find the frequency for which the phase is zero. We checked that the value is unchanged by changing the set of points used for the fit procedure.

7. *Similar to beam-induced errors, the microwave phase errors add a large offset that has to be subtracted. The authors offer little explanation, other than saying that it may be a "nonlinearity" in the phase detector. Can the authors check if this could be due to the group delay dispersion of the low-pass filter? The authors characterize such phase errors by an oscilloscope and note that changing the oscilloscope bandwidth to 20 MHz does not introduce additional shifts larger than 0.01 ppb in h/m. In absence of a quantitative understanding of the effect, should this be added as an error bar to the measurement?* Thank you for this relevant question.
   We have checked that groups delay induced a phase shift by inserting different filter in the PLL and we also checked that the measurement of the phase shift with the scope compensate the measured phase shift. We added in manuscript page 30: "Experimentally, we have also checked that this method can compensate the phase shift observed when filters with different cutoff frequencies are inserted in the PLL.
   Following your recommendation we have added 0.01 ppb to the error bar. We added at the end of the method.
   "We then find a correction to h/m of (-82.1 ±1.2) 10^{-11}. The uncertainty includes a conservative uncertainty of 0.01 ppb on the phase measurement with the oscilloscope."
   Please note that including 0.01 ppb does not change the uncertainty on our value of the fine structure constant.

8. *Can the authors amend Fig. 3 or the description in the text to show where the ramp of 25 MHz/s is applied, and can they state the frequen- cies (in MHz) of the signals after the mixers? I am asking because these quantities enter the phase shift caused by group delay*

dispersion.

The ramp of 25 MHz/s appear in $\delta\omega_{RF}$. This variable appear explicitly in the Fig 3 (note that we have corrected a typo : $\delta\omega_{R}$ in Fig 3.b being $\delta\omega_{RF}$). Figure 3.b was also modified in order to display the position of the $\pi/2$ pulses of the interferometer.

9. *Apodizing filters are known to result in non-Gaussian beams. From the text in the supplemental material, I think this has been taken into account as the Monte-Carlo simulation used actual images of the beam. Is this true, and if so, can this be explicitly stated?*
This is true and explicitly stated in the sentence "The term in Laplacian (associated with the Gouy phase) has been calculated from pictures taken with a CCD" (§ Wave front corrections).

Reviewer Reports on the Second Revision:

Referee #1 (Remarks to the Author):

I compliment the authors on a very impressive and important measurement. I have no further comments. As in my first review, I strongly recommend publication.

Referee #3 (Remarks to the Author):

The authors' replies to the previous round of reviews are good and have further improved the manuscript. This work will be welcomed by the community as a landmark precision measurement. Accordingly, I have only minor additional suggestions. Addressing them should, in my opinion, not be a condition for publication. They are merely suggestions to add information that, I feel, the community would be interested to know.

I commend the authors for carefully addressing the issue of the tension between the three atomic measurements of the fine structure constant, avoiding speculation when no data is available.

The authors' work determining the effect of small-scale intensity fluctuations by measuring the correlation between alpha and the beam intensity is excellent. I agree with their interpretation that the additional noise is likely arising from the viewports of the vacuum chamber. If so, the small-scale fluctuations would be a function of the distance between the viewports and the atoms, and this distance is slightly different for the various beam splitters in the upward and downward-accelerating interferometers shown in Fig. 2. Some readers of the paper might wonder whether this difference can cause a systematic shift in the result. Can the authors add a comment on how (and how well) this can be ruled out?

As a minor observation, the accuracy of the best previous measurement has been 0.20 ppb, and the one of the current measurement is 81 ppt. It is thus a bit optimistic to talk about a "threefold improvement" in alpha, and I suggest to remove that statement. (This observation in no way diminishes the enormous value of the measurement.)

I hope this paper can be published in Nature soon. Let me take this opportunity to congratulate the authors for adding another impressive accomplishment to their series of measurements of the fine structure constant.

Author Rebuttals to Second Revision:

Responses to Referee #3:

Questions from the referees are in italic. Our answers are in blue. We thank the referee for relevant remarks:

1/ The authors' work determining the effect of small-scale intensity fluctuations by measuring the correlation between alpha and the beam intensity is excellent. I agree with their interpretation that the additional noise is likely arising from the viewports of the vacuum chamber. If so, the small-scale fluctuations would be a function of the distance between the viewports and the atoms, and this distance is slightly different for the various beam splitters in the upward and downward-accelerating interferometers shown in Fig. 2. Some readers of the paper might wonder whether this difference can cause a systematic shift in the result. Can the authors add a comment on how (and how well) this can be ruled out?

As explained in the method, to evaluate the systematic effect due to wavefront distortions we take into account the propagation of the laser beams during the sequence of the atomic interferometer. The noise that we add to fit experimental data is also propagated during both upward and downward-accelerating interferometers. Thus if this noise arises from the viewports of the vacuum, the evaluation of the systematic effect takes into account the position of the atoms with respect to the viewports.

2/ As a minor observation, the accuracy of the best previous measurement has been 0.20 ppb, and the one of the current measurement is 81 ppt. It is thus a bit optimistic to talk about a "threefold improvement" in alpha, and I suggest to remove that statement. (This observation in no way diminishes the enormous value of the measurement.)

You are right, we replace the sentence: "our result is a three-fold improvement in \alpha accuracy over the previous cesium measurement" by "Our result improves the accuracy on α by a factor of 2.5 over the previous cesium recoil measurement".