natureresearch

https://doi.org/10.1038/s41586-020-3006-1

Supplementary information

Entanglement on an optical atomic-clock transition

In the format provided by the authors and unedited natureresearch

Peer Review File

Manuscript Title: Entanglement on an Optical Atomic Clock Transition

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

This manuscript reports the first demonstration of spin-squeezing on an optical atomic clock transition and the implementation of an optical atomic clock sequence showing a performance beyond the standard quantum limit.

This is a significant result, representing an important step forward in the field of quantum metrology. State-of-the-art optical lattice clocks operate at or near the standard quantum limit, but techniques for precision enhacement by means of entanglement have up to now only been demonstrated on microwave or radio-frequency transitions. Transferring spin-squeezing and the associated quantum entanglement to an optical transition is technologically challenging and it paves the way to significant precision enhancements in future clocks and atom interferometers. The results reported in this manuscript will find many interested readers in the quantum science and precision measurement communities and beyond.

The manuscript is for the most part well-written. The introduction and motivation are clear, accessible to a broad audience, and include appropriate citations to earlier work. The experimental demonstration of spin-squeezing on the optical clock transition of ultracold Yb atoms (Figs 1-3) is very well described. The characterization of the squeezed state in terms of the Wineland parameter is solid.

My main concerns and questions relate to the data in Fig. 4, its presentation and the conclusions drawn from it. This figure shows data from an experiment where the full optical atomic clocks sequence is implemented, either with a squeezed spin state (SSS) or with a coherent spin state (CSS). The performance is characterized in terms of the Allan deviation, which is the appropriate figure of merit. However, I am concerned about the role of local oscillator (LO) phase noise in these measurements, how this noise is subtracted from some of the data, how this is presented in the manuscript, and the claims derived from it.

Fig. 4a reports the Allan deviation measured with a Ramsey sequence on the optical atomic clock transition, either with a CSS or a SSS and for two different Ramsey times. If I understood it right, no noise subtraction or other data processing is performed. The data for short Ramsey time (tauR1 = 0.17 ms) shows that the stability is improved if a SSS is used as compared to a CSS. Although this Ramsey time is much shorter than in state-of-the-art atomic clocks, the measurement is a proof-of-principle demonstration that spin-squeezing can enhance the precision in an optical atomic clock sequence. To show this enhancement the SQL should be indicated in Fig. 4a, for the parameters of the Ramsey sequence with tauR1 = 0.17 ms, where data is shown for both the CSS and SSS. In this way one could directly see the improvement with respect to the SQL in a conceptually straightforward way.

The data in Fig. 4b, on the other hand, involves subtraction of local oscillator noise, which is measured independently with a longer Ramsey time where this noise dominates. How exactly is the LO phase noise subtracted? What is the clock performance before and after subtraction? This is probably straightforward and I can guess how it works, but I think this should be explained explicitly in the methods.

Because the LO noise is subtracted, this data cannot be considered a demonstration of the actual clock operation. It is true that the LO noise is the signal one would like to measure (and feed back) when operating the system as a full atomic clock (as stated in the methods), but in the present setup the LO noise is way too large to actually do this. The data in Fig. 4b are rather a characterization measurement showing how the clock would perform if it was operated with a better local oscillator showing less noise. This is interesting, but the way it is presented with the improvement beyond SQL indicated, it could be misunderstood as showing the actual clock performance beyond the SQL. I think the discussion of this figure needs to be modified and extended to properly explain these points.

Two further points:

The Dick noise is mentioned in the beginning of the manuscript, but at the end it is not discussed further. Because of the small duty cycle it would be significant in the present work if a fully operational clock was implemented. This should be mentioned in the context of the results, otherwise the reader may wonder why this noise is no longer considered.

There is a problem with notation (throughout the entire paper and methods): I think the Ramsey contrast (length of atomic spin vector appearing in the Wineland parameter) is |<\vec{S}>| not <|\vec{S}|>, see Wineland et al, PRA 46, 6797 (1992) or also the earlier paper by the Vuletic group, PRL 104, 250801 (2010). Is this simply a notational error or is there more behind it? Perhaps the authors indeed want to make a statement about <|\vec{S}|> as the radius of the Bloch sphere, which is insensitive to phase noise, but then this is not the quantity that appears in the Wineland parameter. This should be clarified.

Overall I think it should be stated more clearly in the manuscript that this is not yet a fully operational optical atomic clock but rather the implementation of an optical atomic clock sequence with a spin-squeezed state and the demonstration that the atomic system provides a stability beyond the SQL. The main limitation in the present work is the short coherence time of the LO, which should be stated more clearly, including its implications. The title and abstract also require modifications to avoid misinterpretation of the results by a non-expert audience. This being said, I want to emphasize again that the results in this paper are significant even if they do not yet show a fully operational entanglement-enhanced optical atomic clock and that the results are an important step forward in the field of quantum metrology.

Referee #2 (Remarks to the Author):

The authors of the manuscript 'Entanglement-enhanced optical atomic clock' describe the generation and detection of a spin-squeezed state on an optical transition, namely the clock transition in Yb-171. They compare the clock instability between operation with a coherent and a squeezed state and observe a significant improvement of clock stability as shown in Fig. 4. I believe that these results are highly significant and I congratulate the authors on their beautiful experiment. I believe that the observation of spin-squeezing on an optical transition is a major scientific achievement that warrants publication in Nature. A very strong point of the method is its compatibility with the operation of a highly accurate lattice clock.
Nevertheless, I think that some sentences in the manuscript are phrased such that they imply incorrect or at least exaggerated statements and should be adapted. In my opinion this will not diminish the scientific value of the paper but put it in its proper context. The authors claim in the abstract (line 15) that lattice clocks operate at or near the SQL. Though this statement is not incorrect, in the context of the paper it implies that non-classical states will be necessary and valuable to improve the clock performance of state-of-the-art lattice clocks. However, the statement disregards that the cited systems show a substantial noise contribution from local

oscillator (LO) noise via the Dick effect, which impedes strong improvement of the clock stability by spin squeezed states. I agree that dead time-free interrogation can overcome this problem as the authors state later, but in the abstract the wording is misleading.
Secondly, the authors also write (line 17) that performance beyond the SQL has been achieved in microwave clocks. I think this will mislead many readers of Nature and does not do justice to the people operating microwave atomic clocks with orders of magnitude better accuracy and stability as in the referenced demonstrations. Again, I value these demonstrations but it is a far too strong statement to call them clocks without making a clear distinction to the real metrological systems.
A third general remark concerns the subtraction of LO noise contributions from the data (around line 100). In the actual clock demonstration in Fig 4 a, no operation beyond the SQL has been achieved, which contradicts the statement in lines 122 and 123. I agree that the data implies that beyond SQL operation would be possible with either a dead time-free clock or a perfect, noise-free LO. However, I do not think that LO noise is a mere technical problem – perfect lasers would e.g. allow for extremely long interrogation times (disregarding other technical issues, too) and thus very low SQL, in consequence making spin squeezed states practically unnecessary. Therefore, I am not convinced that removing LO noise from the data to demonstrate sub-SQL operation of the clock is valid.
More detailed comments concern:
Fig. 3 and Fig. 4 both use 'time' as x axis. It would be helpful to add 'Ramsey' and 'averaging' or something similar.
Fig. 4: For me, it is not directly clear what is shown in panel a: the instability of the self-comparison data (including noise from both clock cycles) or an inferred instability of a single clock cycle? Is T_c the time for a single interrogation or does it include the dead time of the second cycle used for self-comparison? The references 27 and 32 are not very helpful here, 32 does not use self-comparison, I think. The data reduction form the instability in a self-comparison to a single lock instability does involve some subtle issues with the LO noise as is discussed PRA 92, 063814.
Lines 110/111: What do the authors mean by 'the LO noise is the same in both sequences'? Why should it be different? Or do they want to say that the clock noise arising from LO noise is the same in both cycles?
Lines 130 – 132: With the best interrogation lasers available and in the described scenario, the instability would be about 3E-17 to 4E-17 at 1 s due to the Dick effect, slightly above the SQL. Why do the authors limit the interrogation time to 0.3 s when current lasers promise interrogation times of several seconds (and a corresponding SQL)?
LO noise (line 268 and following): I am confused by the LO noise the authors estimate (line 278). After a coarse estimate of the Dick effect-limited instability for the given parameters, I would expect an instability in Fig. 4a that is about two orders of magnitude worse. Can it be that LO noise and Dick noise have been summarized in the number given in line 280?

Referee #3 (Remarks to the Author):

The authors report the creation of a many-atom entangled state on a 171Yb optical transition with 350(40) atoms, and demonstrate an optical-lattice clocks with an Allan deviation below the standard quantum limit (SQL). As the systematic uncertainty of the optical clocks rapidly decreases, reaching 9x10-19, it became increasingly important to improve further the clock stability (that will allow to reach a given accuracy faster). For lattice clocks, further reduction of the systematic uncertainty to 10-19 level and below will require a detailed investigation of the higher-order light shifts. Ability to very quickly average down and reach 10-18 (and below) level of accuracy is essential for these studies as well as future operation of clocks and their applications requiring fast averaging. For example, some dark matter searches need excellent short-term stability. Clock performance beyond the SQL has been achieved for microwave transitions, and a generation of spin squeezing between Zeeman sublevels of the electronic ground state of 171Yb by the same group. It is more difficult to maintain phase coherence at higher (optical) frequencies, and this work represents an important achievement. A metrological gain of 4dB over the SQL and reaching a given stability 2.8(3) times faster than the same clock operated at the SQL was reported.

The spin squeezed state (SSS) on the optical-clock transition is generated by the first spin squeezing between the two nuclear sublevels of the electronic 1S0 ground state using the interaction between the atoms and the optical cavity and then transferring the population of the upper sublevel into the excited clock state. This is an elegant solution that builds on prior work. The experiments are very thorough, and the paper is clearly written. Literature references, abstract, introduction are appropriate (see my note below about the conclusion).

This is the first time that the operation of the optical clock beyond the SQL with spin-squeezed states was demonstrated, representing a major achievement in quantum metrology.
The paper will be of immediate interest to a broad range of researches involved in quantum information and precision measurement with quantum systems as well as a wide range of their applications spanning several fields, and I recommend its publication in Nature after the authors consider the comments below.

(1) In the sentence on page 3 "The SSS is nearly uncertainty limited, with a detected (inferred) area 1.9 (1.6) times larger than the limit set by the Heisenberg's uncertainty principle [15]." what detected and inferred areas refers too?

(2) The caption of Fig. 2 states that the error bar in this and the following figures represent the 68% confidence interval. Why not use a 95% confidence interval, which is more common?

(3) In light of (2), are all errors in the text, such as 2.8(3), represent one sigma?

(4) In the sentence on page 4, "The pi-pulse has a transfer efficiency of 0.95, thus setting a limit on the observable noise reduction of 13 dB", why is the efficiency 0.95? Is this the limitation of the present set up or is there some other issue?

(5) The issue of the time that entanglement can be maintained is very important. Its evaluation is complicated by the local oscillator (LO) noise. The paper states that SSS is no longer sufficiently entangled to overcome the SQL at 0.2 s, but SSS still can offer a metrological gain over the CSS for up to 0.5 s. On a minor note some clarification would be useful why SSS is still better than CSS if entanglement is not sufficient. More important question: even taking optimistically the SSS entanglement time of 0.5s, this is still below the probe times allowed by the best available local oscillators. The long-term goal of developing SSS with optical clocks is to be eventually able to beat the SQL when all other parameters defining stability in Eq. (1) are at their limits. While I understand that this is the first-time demonstration, it is essential to add some prospects for the future ability to maintain SSS state for much longer times. If LOs that are coherent for a 50-100 seconds become possible, one can use a long probe time to achieve superb stability. Other schemes such as correlation spectroscopy can be also used to overcome LO noise and improve stability without SSS. Do the authors see SSS still capable of providing a metrological gain in this scenario in the future, i.e., in a perfect experiment in the next 5-10 years, can you see maintaining SSS for 50 seconds or more? I understand that this may be hard to answer by at least some projection is needed to ensure that ultimately SSS with provide a metrological gain that cannot be achieved by any other (relatively simpler) means.

(6) In a question related to (5), the last paragraph promises four orders of magnitude improvement in stability with no discussion. Some further discussion of the challenges is needed here (perhaps at least in the supplement).

(7) The schemes used in this work will complicate present optical clock setups. How much of a problem is this, i.e., do authors foresee the appearance of new systematic effects at the 10-18 -

10-19 levels? I understand that any assessment of the future systematic is beyond the scope here, so I just what to see if there are any more or less obvious problems that will appear with the production and use of the SSS states which do not appear during the CSS clock operation.

Referee 1

This manuscript reports the first demonstration of spin-squeezing on an optical atomic clock transition and the implementation of an optical atomic clock sequence showing a performance beyond the standard quantum limit.

This is a significant result, representing an important step forward in the field of quantum metrology. State-of-the-art optical lattice clocks operate at or near the standard quantum limit, but techniques for precision enhancement by means of entanglement have up to now only been demonstrated on microwave or radio-frequency transitions. Transferring spin-squeezing and the associated quantum entanglement to an optical transition is technologically challenging and it paves the way to significant precision enhancements in future clocks and atom interferometers. The results reported in this manuscript will find many interested readers in the quantum science and precision measurement communities and beyond.

The manuscript is for the most part well-written. The introduction and motivation are clear, accessible to a broad audience, and include appropriate citations to earlier work. The experimental demonstration of spin-squeezing on the optical clock transition of ultracold Yb atoms (Figs 1-3) is very well described. The characterization of the squeezed state in terms of the Wineland parameter is solid.

My main concerns and questions relate to the data in Fig. 4, its presentation and the conclusions drawn from it. This figure shows data from an experiment where the full *optical atomic clocks sequence is implemented, either with a squeezed spin state (SSS) or with a coherent spin state (CSS). The performance is characterized in terms of the Allan deviation, which is the appropriate figure of merit. However, I am concerned about the role of local oscillator (LO) phase noise in these measurements, how this noise is subtracted from some of the data, how this is presented in the manuscript, and the claims derived from it.*

We would like to thank referee 1 for his/her careful and expert review of our work. We have followed all of Referee 1's suggestions/comments, and have implemented the following changes:

Comment 1.1

> *Fig. 4a reports the Allan deviation measured with a Ramsey sequence on the optical atomic clock transition, either with a CSS or a SSS and for two different Ramsey times. If I understood it right, no noise subtraction or other data processing is performed. The data for short Ramsey time (tauR1 = 0.17 ms) shows that the stability is improved if a SSS is used as compared to a CSS. Although this Ramsey time is much shorter than in state-of-the-art atomic clocks, the measurement is a proof-of-principle demonstration that spin-squeezing can enhance the precision in an optical atomic clock sequence. To show this enhancement the SQL should be indicated in Fig. 4a, for the parameters of the Ramsey sequence with tauR1 = 0.17 ms, where data is shown for both the CSS and SSS. In this way one could directly see the improvement with respect to the SQL in a conceptually straightforward way.*

This is a good suggestion by the referee. Since the Fig. 4a in our originally submitted version of the manuscript included two datasets with different Ramsey times, and correspondingly different levels of the SQL, it would have been confusing to add both to the same figure. Instead, we have now split Fig. 4a into two part (new Figs. 4a and b), and the previous Fig. 4b has now become Fig. 4c. In the new Figs. 4a and 4b we now indicate both the corresponding SQL and the contribution of LO phase noise to the (in-)stability of the clock, as suggested by the referee. We have adjusted the figure caption and the discussion in the text accordingly.

Comment 1.2

> *The data in Fig. 4b, on the other hand, involves subtraction of local oscillator noise, which is measured independently with a longer Ramsey time where this noise dominates. How exactly is the LO phase noise subtracted? What is the clock performance before and after subtraction? This is probably straightforward and I can guess how it works, but I think this should be explained explicitly in the methods.*

The new Figs. 4a and b now clarify the relative contributions of the quantum projection noise and the LO noise. The LO noise is subtracted in quadrature, and the clock performance before and after subtraction can now be seen by comparing Figs. 4a and 4c. This is now also discussed in more details in the Methods, section "LO noise subtraction".

Comment 1.3

> *Because the LO noise is subtracted, this data cannot be considered a demonstration of the actual clock operation. It is true that the LO noise is the signal one would like to measure (and feed back) when operating the system as a full atomic clock (as stated in the methods), but in the present setup the LO noise is way too large to actually do this. The data in Fig. 4b are rather a characterization measurement showing how the clock would perform if it was operated with a better local oscillator showing less noise. This is interesting, but the way it is presented with the improvement beyond SQL indicated, it could be misunderstood as showing the actual clock performance beyond the SQL. I think the discussion of this figure needs to be modified and extended to properly explain these points.*

We thank the referee for this suggestion and agree that the data in Fig. 4b (now Fig. 4c) are a demonstration measurement that shows how much improvement one would get with a better local-oscillator laser. We have accordingly modified and extended the discussion of figure 4 (lines 106 - 118). We now state "*In Fig. 4 we demonstrate the implementation of an optical atomic clock Ramsey sequence with an entangled state input and show that the atomic system provides a stability beyond the SQL.*" (This follows a wording suggested by the referee, see also our response to Comment 1.1).

Comment 1.4

> *The Dick noise is mentioned in the beginning of the manuscript, but at the end it is not discussed further. Because of the small duty cycle it would be significant in the present work if a fully operational clock was implemented. This should be mentioned in the context of the results, otherwise the reader may wonder why this noise is no longer considered.*

Following this suggestion, we have expanded the conclusion of our manuscript, and now discuss in more detail the impact of Dick noise, and methods to reduce it (lines 132 - 135).

Comment 1.5

> *There is a problem with notation (throughout the entire paper and methods): I think the Ramsey contrast (length of atomic spin vector appearing in the Wineland*

parameter) is |<\vec{S}>| not <|\vec{S}|>, see Wineland et al, PRA 46, 6797 (1992) or also the earlier paper by the Vuletic group, PRL 104, 250801 (2010). Is this simply a notational error or is there more behind it? Perhaps the authors indeed want to make a statement about <|\vec{S}|> as the radius of the Bloch sphere, which is insensitive to phase noise, but then this is not the quantity that appears in the Wineland parameter. This should be clarified.

We agree with the referee; This was a notational error, and we have corrected it. This quantity is now written as it should be, i.e., |<\vec{S}>|.

Comment 1.6

Overall I think it should be stated more clearly in the manuscript that this is not yet a fully operational optical atomic clock but rather the implementation of an optical atomic clock sequence with a spin-squeezed state and the demonstration that the atomic system provides a stability beyond the SQL. The main limitation in the present work is the short coherence time of the LO, which should be stated more clearly, including its implications. The title and abstract also require modifications to avoid misinterpretation of the results by a non-expert audience. This being said, I want to emphasize again that the results in this paper are significant even if they do not yet show a fully operational entanglement-enhanced optical atomic clock and that the results are an important step forward in the field of quantum metrology.

We thank the referee for this suggestion and have modified the title and abstract to clarify what we are demonstrating in this work, and what still needs to be done in the future. In particular, we now state in the abstract “Here we report creation of a many-atom entangled state on an OLC transition, and use it to demonstrate a Ramsey sequence with an Allan deviation below the SQL after subtraction of the LO noise.” We have also modified the discussion of Fig. 4 and the conclusion (see Comment 1.3). Furthermore, we have changed the title of the paper to “Entanglement on an Optical Atomic Clock Transition” in order to avoid misinterpretation of our results by non-experts.

Referee 2

The authors of the manuscript ‘Entanglement-enhanced optical atomic clock’ describe the generation and detection of a spin-squeezed state on an optical transition, namely the clock transition in Yb-171. They compare the clock instability between operation with a coherent and a squeezed state and observe a significant improvement of clock stability as shown in Fig. 4. I believe that these results are highly significant and I congratulate the authors on their beautiful experiment. I believe that the observation of spin-squeezing on an optical transition is a major scientific achievement that warrants *publication in Nature. A very strong point of the method is its compatibility with the operation of a highly accurate lattice clock. Nevertheless, I think that some sentences in the manuscript are phrased such that they imply incorrect or at least exaggerated statements and should be adapted. In my opinion this will not diminish the scientific value of the paper but put it in its proper context.*

We would like to thank referee 2 for the positive, very constructive and insightful comments about our manuscript. We have followed all of the referee's recommendations by making the changes described below:

Comment 2.1

> *The authors claim in the abstract (line 15) that lattice clocks operate at or near the SQL. Though this statement is not incorrect, in the context of the paper it implies that non-classical states will be necessary and valuable to improve the clock performance of state-of-the-art lattice clocks. However, the statement disregards that the cited systems show a substantial noise contribution from local oscillator (LO) noise via the Dick effect, which impedes strong improvement of the clock stability by spin squeezed states. I agree that dead time-free interrogation can overcome this problem as the authors state later, but in the abstract the wording is misleading.*

We thank the referee for pointing out this possible misunderstanding. We have rewritten the abstract, where we now state "The stability of optical-lattice clocks (OLCs) is limited both by the interrupted interrogation of the atomic system by the local-oscillator (LO) laser (Dick noise [5]), and by the standard quantum limit (SQL) that arises from the quantum noise associated with discrete measurement outcomes. While schemes for removing the Dick noise have been recently proposed and implemented *…*".

Comment 2.2

> *Secondly, the authors also write (line 17) that performance beyond the SQL has been achieved in microwave clocks. I think this will mislead many readers of Nature and does not do justice to the people operating microwave atomic clocks with orders of magnitude better accuracy and stability as in the referenced demonstrations. Again, I value these demonstrations but it is a far too strong statement to call them clocks without making a clear distinction to the real metrological systems.*

We thank the referee for this comment. We certainly did not mean to mislead readers about the beautiful and important work done with accurate microwave atomic clocks and have rephrased our statement in the abstract "… performance beyond the SQL by engineering quantum correlations (entanglement) between the atoms [9–20] has been demonstrated only in proof-of-principle experiments with microwave clocks of limited stability*.*"

Comment 2.3

> *A third general remark concerns the subtraction of LO noise contributions from the data (around line 100). In the actual clock demonstration in Fig 4 a, no operation beyond the SQL has been achieved, which contradicts the statement in lines 122 and 123. I agree that the data implies that beyond SQL operation would be possible with either a dead time-free clock or a perfect, noise-free LO. However, I do not think that LO noise is a mere technical problem – perfect lasers would e.g. allow for extremely long interrogation times (disregarding other technical issues, too) and thus very low SQL, in consequence making spin squeezed states practically unnecessary. Therefore, I am not convinced that removing LO noise from the data to demonstrate sub-SQL operation of the clock is valid.*

We agree with the referee that due to the large LO noise, we have not achieved full clock operation (with feedback to the LO) below the SQL, but rather demonstrated that the atomic system provides a stability beyond the SQL (see also Comment 1.6 of referee 1). To avoid confusion about this point, we have replaced the original statement by the statement that we "… demonstrate the implementation of an optical atomic clock Ramsey sequence with an entangled state input and show that the atomic system provides stability beyond the SQL" (lines 106 - 107). We have also modified the abstract accordingly, and altered the title of the paper.
While in the limit of a perfect LO indeed one can make the interrogation time very long and thus does not require spin squeezing, there is a range of LO instability, relevant in practice, where spin squeezing helps. This is the range where due to the higher-resolution readout $(\Delta\phi)_{SS}$ of the LO phase offered by spin squeezing, the frequency stability of the LO $\Delta\omega = (\Delta\phi)_{SS} /T$ is improved. Currently we cannot directly demonstrate this because by the time we can do a second measurement of the LO phase after having done the first measurement with improved resolution due to spin squeezing, the LO phase has drifted off by much more than that resolution $(\Delta\phi)_{SS}$. However, by improving the LO noise and making it similar to that achieved in other state-of-the-art clocks, and removing Dick noise, one can make use of the improved phase measurement. Other possible gains from spin squeezing are discussed in the conclusion that has been rewritten.

Comment 2.4

> *Fig. 3 and Fig. 4 both use 'time' as x axis. It would be helpful to add 'Ramsey' and 'averaging' or something similar.*

We thank the referee for this suggestion. We have renamed the axes of Fig. 3 and Fig. 4 to ‘Hold time’ and ‘Averaging time’, respectively.

Comment 2.5

> *Fig. 4: For me, it is not directly clear what is shown in panel a: the instability of the self-comparison data (including noise from both clock cycles) or an inferred instability of a single clock cycle? Is T_c the time for a single interrogation or does it include the dead time of the second cycle used for self-comparison? The references 27 and 32 are not very helpful here, 32 does not use self-comparison, I think. The data reduction form the instability in a self-comparison to a single lock instability does involve some subtle issues with the LO noise as is discussed PRA 92, 063814.*

The Allan deviation plot shows the inferred instability of a single clock cycle; Since in our system there is no correlation (see Methods, LO noise subtraction) between two consecutive cycles, we infer the single clock cycle instability by dividing the data that include noise from both clock cycles by \sqrt{2}. T_c is the time for a single interrogation. We also agree that PRA 92, 063814 is a much better and explanatory reference. We have expanded the caption of Fig. 4 to clarify these points, and we now cite PRA 92, 063814.
Please note that to comply also with the suggestion by Referee 1 (see Comment 1.1) to indicate the SQL in the figure, we have now split the original figure 4a into 4a and 4b.

Comment 2.6

> *Lines 110/111: What do the authors mean by ‘the LO noise is the same in both sequences’? Why should it be different? Or do they want to say that the clock noise arising from LO noise is the same in both cycles?*

We agree with the referee that our original phrasing was unclear. We have now rewritten the corresponding paragraph (lines 106 - 118), and discuss in more detail how the LO noise is determined (see also Comment 1.2). Indeed the clock noise from LO noise is the same in both cycles.

Comment 2.7

> *Lines 130 – 132: With the best interrogation lasers available and in the described scenario, the instability would be about 3E-17 to 4E-17 at 1 s due to the Dick effect, slightly above the SQL. Why do the authors limit the interrogation time to 0.3 s when current lasers promise interrogation times of several seconds (and a corresponding SQL)?*

We agree with the referee that this was perhaps a too conservative assumption. We have now significantly expanded the conclusion, using a longer proposed Ramsey time of 2s (set by the natural lifetime of the excited state in Yb in combination with the squeezing) to estimate the possible clock performance with spin squeezing (see lines 124 - 126). Sr or other atoms would allow even longer times.

Comment 2.8

> *LO noise (line 268 and following): I am confused by the LO noise the authors estimate (line 278). After a coarse estimate of the Dick effect-limited instability for the given parameters, I would expect an instability in Fig. 4a that is about two orders of magnitude worse. Can it be that LO noise and Dick noise have been summarized in the number given in line 280?*

We agree with the referee that the formulation we used may have been confusing. We have rewritten this section of the Methods (“LO Noise”), and we now explicitly state that the value $\tau_{LO}=6$~ms for the coherence time includes also the Dick noise contribution (lines 302 - 303). The LO frequency noise on the 4s-time cycle time scale dominates the clock instability. The expected Dick noise, considering an LO noise dominated by low frequency noise (flicker and Brown), agrees with our measurement.

Referee 3

The authors report the creation of a many-atom entangled state on a 171Yb optical transition with 350(40) atoms, and demonstrate an optical-lattice clocks with an Allan deviation below the standard quantum limit (SQL). As the systematic uncertainty of the optical clocks rapidly decreases, reaching 9x10-19, it became increasingly important to improve further the clock stability (that will allow to reach a given accuracy faster). For lattice clocks, further reduction of the systematic uncertainty to 10-19 level and below will require a detailed investigation of the higher-order light shifts. Ability to very quickly average down and reach 10-18 (and below) level of accuracy is essential for these studies as well as future operation of clocks and their applications requiring fast averaging. For example, some dark matter searches need excellent short-term stability. Clock performance beyond the SQL has been achieved for microwave transitions, and a generation of spin squeezing between Zeeman sublevels of the electronic ground state of 171Yb by the same group. It is more difficult to maintain phase coherence at higher (optical) frequencies, and this work represents an important achievement. A metrological gain of 4dB over the SQL and reaching a given stability 2.8(3) times faster than the same clock operated at the SQL was reported.

The spin squeezed state (SSS) on the optical-clock transition is generated by the first spin squeezing between the two nuclear sublevels of the electronic 1S0 ground state *using the interaction between the atoms and the optical cavity and then transferring the population of the upper sublevel into the excited clock state. This is an elegant solution that builds on prior work. The experiments are very thorough, and the paper is clearly written. Literature references, abstract, introduction are appropriate (see my note below about the conclusion).*

This is the first time that the operation of the optical clock beyond the SQL with spin-squeezed states was demonstrated, representing a major achievement in quantum metrology.

The paper will be of immediate interest to a broad range of researches involved in quantum information and precision measurement with quantum systems as well as a wide range of their applications spanning several fields, and I recommend its publication in Nature after the authors consider the comments below.

We would like to thank referee 3 for the very positive and constructive comments and suggestions that have helped further improve our paper. Below, we attach the changes we have made to our manuscript, attending to all the suggestions and comments the referee has made.

Comment 3.1

> *In the sentence on page 3 "The SSS is nearly uncertainty limited, with a detected (inferred) area 1.9 (1.6) times larger than the limit set by the Heisenberg's uncertainty principle [15]." what detected and inferred areas refers too?*

Detected area refers to the directly measured state area that includes all measurement noise. The “inferred area” is the area of the reconstructed state after we have removed the detection noise. Since it is not central to our present result, we have removed the inferred area values from that sentence, and now only quote the directly measured state area including all measurement noise.

Comment 3.2

> *The caption of Fig. 2 states that the error bar in this and the following figures represent the 68% confidence interval. Why not use a 95% confidence interval, which is more common?*

This question about the convention seems to strongly depend on the community, and while many publications in the metrological community use the 95% confidence interval, the vast majority of papers on optical atomic clocks, and spin-squeezing and entanglement, use the 68% (1 sigma) confidence interval. See, for example, neutral-

atom-clock papers [Oelker et al., Nature Phot. 13, 714 (2019), McGrew et al., Nature 564, 87 (2018), Schioppo et al., Nature Phot. 11, 48 (2017), Takano et al., Nature Phot. 10, 662 (2016), Ushijima et al., Nature Phot. 9, 185 (2015), Nicholson et al, Nature Comm. 6, 6896 (2015), Bloom et al., Nature 506, 71 (2014)], and spin-squeezing papers [Riedel et al. Nature, 464, 7292 (2010), Hosten et al. ", Nature 529, 7587 (2016), Leroux et al. , Physical Review Letters, 104, 7 (2010)].
Based on this, we think that we are in line with most of the work on optical atomic clocks (as well as spin squeezing) when we use the 68% confidence interval. However, if the referee insists, we can change it to 95%.

Comment 3.3

In light of (2), are all errors in the text, such as 2.8(3), represent one sigma?

This is correct, all the errors mentioned in the text correspond to one sigma confidence interval. We now specify this in the text and in the figure captions.

Comment 3.4

In the sentence on page 4, "The pi-pulse has a transfer efficiency of 0.95, thus setting a limit on the observable noise reduction of 13 dB", why is the efficiency 0.95? Is this the limitation of the present set up or is there some other issue?

It is a limitation of the present set-up, and in particular, the non-zero temperature of atoms. We expect the efficiency of the pi-pulse to increase with even better cooling. We have clarified this point by modifying the sentence as follows (line 83): “The pi-pulse has a transfer efficiency of 0.95, mostly limited by finite atomic temperature”.

Comment 3.5

The issue of the time that entanglement can be maintained is very important. Its evaluation is complicated by the local oscillator (LO) noise. The paper states that SSS is no longer sufficiently entangled to overcome the SQL at 0.2 s, but SSS still can offer a metrological gain over the CSS for up to 0.5 s. On a minor note some clarification would be useful why SSS is still better than CSS if entanglement is not sufficient. More important question: even taking optimistically the SSS entanglement time of 0.5s, this is still below the probe times allowed by the best available local oscillators. The long-term goal of developing SSS with optical clocks is to be eventually able to beat the SQL when all other parameters defining stability in Eq. (1) are at their limits. While I understand that this is the first-time *demonstration, it is essential to add some prospects for the future ability to maintain SSS state for much longer times. If LOs that are coherent for a 50-100 seconds become possible, one can use a long probe time to achieve superb stability. Other schemes such as correlation spectroscopy can be also used to overcome LO noise and improve stability without SSS. Do the authors see SSS still capable of providing a metrological gain in this scenario in the future, i.e., in a perfect experiment in the next 5-10 years, can you see maintaining SSS for 50 seconds or more? I understand that this may be hard to answer by at least some projection is needed to ensure that ultimately SSS with provide a metrological gain that cannot be achieved by any other (relatively simpler) means.*

We thank the referee for this suggestion and important question. First, the answer to the minor note: Both the CSS and the SSS have some of the same imperfections (finite contrast, imperfect state detection), and the SSS with its lower quantum uncertainty can still outperform the CSS used in the experiment even if it is no longer demonstrably entangled (i.e. no longer better than a perfect CSS which represents the SQL). A related explanation is now given in lines 115-117, "… below the CSS that is also subject to imperfect state detection and contrast loss".

Now to the important questions: The lifetime of the SSS is limited by i) dephasing of the collective atomic spin (relative to some perfect absolute frequency reference), ii) dephasing between the atoms, and iii) atom loss. Regarding problem i), the atomic transition needs to be controlled (and stable) no matter whether one uses a CSS or an SSS, and it seems likely that the atomic ensemble can be sufficiently controlled in the future to achieve the corresponding absolute stability, possibly using new transitions that are less sensitive to blackbody radiation, etc. Item ii) can be addressed by making the environment seen by the atoms as homogeneous as possible, cooling all the atoms to the ground state of a 3D lattice, etc; we see no fundamental or stringent technical limitations there. Problem iii) is perhaps the most mundane problem and will ultimately be limited by the vacuum; if one wants to maintain 10dB of spin squeezing (less than 10% of atom loss) for 60 seconds, then one needs a vacuum lifetime of 10 minutes. Some experiments with cold atoms are approaching those lifetimes. In short, we see no fundamental or medium-term technical obstacle to achieve the performance numbers considered by the referee. We have expanded the conclusion to add some discussion of this question posed by the referee, but due to length constraints of the manuscript we cannot elaborate in as much detail as this interesting question deserves.

We also would like to point out that some proposed fundamental tests require optical clocks with a finite bandwidth, and these applications can profit directly and near-term from spin squeezing. This is now mentioned in the concluding paragraph.

Comment 3.6

In a question related to (5), the last paragraph promises four orders of magnitude improvement in stability with no discussion. Some further discussion of the challenges is needed here (perhaps at least in the supplement).

The improvement in stability is mostly by employing a better stabilization cavity for the clock laser, and reproducing some of the demonstrated solutions from the literature to improve the frequency stabilization of the clock laser to the cavity. This future improvement is somewhat technical, and we have added a discussion in the Methods section, rather than in the main manuscript.

Comment 3.7

The schemes used in this work will complicate present optical clock setups. How much of a problem is this, i.e., do authors foresee the appearance of new systematic effects at the 10-18 - 10-19 levels? I understand that any assessment of the future systematic is beyond the scope here, so I just what to see if there are any more or less obvious problems that will appear with the production and use of the SSS states which do not appear during the CSS clock operation.

The main difference compared to standard setups is the proximity of the mirrors, piezo actuators, and the effect of the cavity mode on the clock transition. We have evaluated those and find that with proper shielding, appropriate distance from the mirrors, and control of the cavity mode frequency during the Ramsey time, there should be no strong obstacles to control and suppress additional systematic effects induced by the spin squeezing setup at the 10^{-19} scale.

We mention this now in the conclusion of the main paper (lines 136-137) and have added a section in Methods (“Potential systematic effects”) where we briefly discuss potential systematic effects when applying the presented scheme to state-of-the-art clocks.

In summary, we have complied with all suggestions/comments by the referees and are resubmitting the paper for publication in Nature.

With kind regards,

Vladan

(on behalf of all the authors)

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

The authors satisfactorily addressed all my concerns. In the resubmitted manuscript, the results are presented in an appropriate way that should allow also non-expert readers to easily understand what was achieved in this work. I enthusiastically recommend publication, this is an important step forward in quantum metrology with atomic clocks.

Minor point: I spotted a typo in the Methods section on LO noise subtraction. The references to Figs. 4a and 4b are interchanged (Fig. 4a shows the short Ramsey time and Fig. 4b the long one).

Referee #2 (Remarks to the Author):

I thank the authors for their detailed reply and adequate adaptations of the manuscript. The paper is in its present form an important contribution to the field of quantum enhanced measurements. The results are presented in proper relation to other activities. I congratulate the authors and recommend publication in the present form.

Referee #3 (Remarks to the Author):

The authors addressed all of my comments and I recommend publication of the paper in its present form.