Authors' response to referees comments

We thank both referees for their detailed comments. We have modified our paper significantly in response, including the addition of a significant amount of new experimental data on the role of heating and loss during imaging. A description of these changes and their location in the revised manuscript is provided below (shown as underlined).

A detailed response to each of the points raised by the referees is also provided. Together we believe this provides a comprehensive response to the objections raised by the referees.

Reference numbers refer to the bibliography in the paper.

Response to Referee 1:

the reader wonders if instead their experimental system was just not compatible with a proper optical tweezer set-up.

What is a "proper optical tweezer set up"? Perhaps it is a setup that provides micron scale localisation of a single atom. Our work clearly meets this definition. However, optical tweezers containing many atoms are also widely used, for example in Rydberg quantum optics experiments. If by "proper" one means "conventional" (ie using lenses of NA>0.5), then we could have easily implemented such a set-up, as we have already done 8 years ago for Rb [Maxwell et al., Phys. Rev. Lett. **110**, 103001 (2010)]. The result would have been a system very similar to the recent work on Sr and Yb tweezer arrays from other groups [15,16,17,21].

Instead our design started from optical lattice clocks, where control over systematic errors at the 10^{-19} level is required. We sought to preserve as many features of conventional clocks as possible, including an interrogation region free from dielectric surfaces and optical access for lattice beams if required. This paper describes the solution that we arrived at, which is different from the ground-breaking work in [15,16,17,21] (and, as we point out, from most alkali tweezer experiments), but which nevertheless functions as a "proper" tweezer set-up.

We have added new experimental data that shows that high-fidelity detection of N = 0, 1, 2 atoms in our tweezer can be achieved (section 6). As a result, a new member of the team who contributed to this data has been added as an author (F. Leroux).

They claim that their avoidance of parity projection is due to the larger size of their tweezer; however, this claim was not fully substantiated. A carefully study of references [15,16] shows that parity projection can be avoided during imaging and cooling in tightly-focused tweezers if desired.

We do not claim that parity projection cannot be avoided in [15,16], merely that it *is* avoided in our setup. Parity projection in Norcia et al. [16] was obtained by the application of alternating pulses of 461nm light and sideband cooling light with a 10% duty cycle, for a total duration of 20 ms. This is the same protocol that is subsequently used for imaging. The authors state that their observed parity projection is due to light-assisted collisions induced by the resulting total of 2 ms of 461 nm light. In our experiment, atoms are exposed to the 461 nm imaging light for a comparable duration and at much higher intensity, and yet we do not observe parity projection.

The major difference between our experiment and [16] is that the volume of our tweezer $V \propto \omega_x \omega_y z_0$ is 15 times larger. Since the light assisted collision rate is proportional to the square of

the density, this leads to a negligible collision rate during our imaging duration. This is the basis of our claim – an imaging duration that led to parity projection in [16] did not do so in our experiment.

We have added this quantitative estimate to the paper, along with a discussion contrasting the different approaches to parity projection in [15] and [16] (section 6 paragraph 3).

The authors acknowledge that heating of the atoms during imaging leads to a reduction in the fluorescence signal for longer exposures, but they do not discuss the possible connection between this observation and their claim of observing atom-resolved loss.

As the referee correctly states, we have a large differential light shift, which combined with heating can lead to a decrease in scattering rate. We have examined the data in more detail and one sees some data that exhibits sudden jumps like 5(c) (old version) and 5(d) (old version) as well as some that exhibit the gradual reduction in count rate one would expect for heating. Therefore, it is not in general possible to achieve atom-resolved loss using the data presented in Fig. 5(c) (old version).

To respond properly to this key point, we have taken a substantial amount of additional data, and performed a quantitative analysis of the imaging fidelity. The results appear in Figs 5, 6 and 7 in the new version. As a result, section 6 has been completely rewritten, and the abstract, introduction and conclusions have also been modified to take into account the new data.

To summarize, we added Sisyphus cooling during imaging as was done in previous works from the Caltech group. We are able to show that this removes the loss of fluorescence signal due to heating, with the lifetime now limited only by depumping to the 5s4d $^1D_2$ state (new Fig 5).

A clear contribution from N = 2 atoms in the tweezer is visible in the data. We have quantified the fidelity with which we can determine the atom number, and we find that two atoms can be discriminated from one with good fidelity. This data substantiates our claim to number-resolved imaging. Time-resolved measurements showing atom-resolved loss are also presented; the cumulative effect of imaging errors means that at present this can't be used for deterministic loading as stated in the text. In future, this limitation can be overcome by switching the dipole trap to 813nm, where the 5s4d $^1D_2$ state is trapped and using repumping, as was shown by the Caltech group [21].

We have also applied the same quantitative analysis to a single 200 us frame of the original data shown in Fig. 5(a), and we find that the performance of this imaging method is surprisingly good in terms of both fidelity and loss for determining the presence of a single atom. However, as expected heating means it does not work well enough to resolve two atoms from one atom. This data is shown in the new Fig. 6.

Lastly, we have alternated imaging and cooling, along the lines of Norcia et al. [16], but additionally exploiting the fast readout of our camera to only readout frames where the probe light is on. This essentially eliminates limitations due to the dark count of our camera.

We believe this dataset and the accompanying text provides a much clearer analysis of the performance of our imaging system and the role of heating.

For example, the tweezer depth is often described as "shallow", or "deep", or "very deep". I understand that the maximum depth used is 7.5 mK, presumably for the 1S0 state. The trap depths of the 1P1 state are only mentioned in reference to a differential Stark shift.

Trap depths have been added throughout. The $^1P_1$ trap depth is not mentioned elsewhere as this state is anti-trapped for a tweezer wavelength of 532nm.

Also, Fig 4 should have a horizontal axis of trap depth in mK or MHz … Perhaps a double horizontal axis (top and bottom) could be used.

The conversion from power to trap depth requires knowledge of the trap waist. This value is obtained from the gradient of the fit in Fig 4. Therefore, it would be misleading to simply convert the axis. However a double axis is a good idea, and we have made that change.

The Rayleigh length and an estimate of the axial trap frequency should be stated. Also, it would be helpful if an estimate of the trap frequency for the radial direction could be given earlier, even before it is carefully measured. It is helpful for understand "shallow" vs "deep".

These values have been added to the paper.

Finally, I believe the overall story of this paper is unclear. The "new" features are: 1. Longer working distance lens, 2. SPAD detection, and 3. Claimed novel ability to operate without parity projection and the claimed atom-resolved loss during imaging. The advantages of 1 are far oversold, 2 is interesting but not utilized since there is only one tweezer, and 3 is partially wrong and not entirely convincing.

Essentially the author correctly identifies the story of our paper – points 1-3 are the main points. However, we do not agree with the comments. Concerning (1) we stress that our approach is different rather better, and motivate it clearly by the desire to keep stray field systematics low in our experiment (see response to point 2 below). Concerning (2), there are now several techniques that are well established (over 10 years for some) for scaling from one tweezer to many (eg DMMs, SLMs, AOMs, microlens arrays etc). As refs [15,16] show, their extension to Sr does not present any particular issues. Instead we concentrate on how the detector (which is new) scales, which is the motivation for the data shown in Fig. 2. As we show in Fig. 6 and 7, using the SPAD camera enables new imaging modalities not possible with detectors of lower frame rate. Concerning (3) we have added substantial additional data to back up our claim to be able to resolve the atom number in the trap, and quantify the performance of our imaging system. We clearly do avoid parity projection, and though this *ability* is not unique to our experiment, number-resolved detection has not been shown in previous experiments with Sr (though it has in the alkalis).

Overall, we conclude (in section 7) that a different approach to tweezer arrays in Sr is possible, with potential advantages and some drawbacks compared to the ground-breaking work of the JILA and Caltech groups. Given the current interest in tweezer arrays and in AEA arrays in particular, we believe these results will be of interest to other groups.

Other points:

1. *The authors choose to cite references [15] and [16] individually in several cases…*
   We thank the referee for drawing our attention to the photos. The paper has been altered accordingly. The other two examples have been modified to cite both papers.

2. *…it seems that the authors oversell their potential advantage regarding dielectric surfaces and charging problems. Such problems are not significant in recent experiments with Rb Rydberg atoms in tweezers at Harvard, for instance. Also, why does the use of a longer working distance preclude a large NA?*
   For a stray charge or small patch potential on a dielectric surface, the associated DC Stark shift will reduce as $(1/d^4)$ where d is the distance to the surface. Comparing our setup (d = 37mm) to that of e.g. Norcia et al (d = 3 mm) this leads to a reduction in systematic error of 10^{-4}. Even comparing to other experiments where d = 10 mm, the reduction is still significant. It is not true that charging problems are insignificant in recent experiments in Harvard. In the Supplementary Material to [10] the authors explain that nearly continuous exposure of the glass cell to UV light is essential to stabilise the electric field environment. This presumably removes surface Rb atoms by light induced atomic desorption (LIAD) – a technique that has not been shown to work with Sr or Yb. It remains to be seen whether the cell approach in Sr [15,16] will cause problems when pushed to the limits of precision. The risk associated with these issues is reduced by our choice of a longer working distance, which is well motivated by our goal of achieving an environment capable of precision measurement. We have strengthened this argument in the paper (section 2 paragraph 2).
   A longer working distance does not preclude a higher NA but it gets harder to source the required large-diameter aspheres. Multi-element lenses in UHV are possible but challenging. Optical access was not a primary motivation, though it is certainly an advantage of our approach. Following the referee we have add a sentence to this effect to the paper (section 2 paragraph 2)

3. *The authors should address the implied scalability of their system...*
   At 532 nm and 813 nm, a power of 8mW (36mW) is required to reach a 0.5 mK trap depth. However before ramping for imaging we require only 0.16 mW (0.72mW) . Arrays of >100 traps are therefore feasible with reasonable laser power. Axial confinement is inherently weaker, but with ground state cooling the spatial extent would already be small (0.4 microns). Given our high degree of optical access, a lattice could easily be implemented if required. For a typical spacing of 3um and assuming a conventional 2D square array, we are limited by the number of pixels on our detector (eg 32x32 = 1024 sites) rather than the transverse field of the lens. The size and shape of the required addressing beams then depends on the chosen array size and geometry.

Response to Referee 2

This would be a nice feature to have, if NA were not also severely sacrificed. Currently, there are major demonstrated advantages to a high-NA setup (reduced optical power requirements, lower *effects of atomic motion, possibility of sideband cooling, higher collection efficiency), but no demonstration that setups with ~5mm separation between atoms and glass surfaces are incompatible with Rydberg physics.*

See our response to referee 1 above. Current experiments rely on methods like LIAD that have not been proven to work for Sr. We have shown that we can implement Sisyphus cooling to control atomic motion, and image with a fidelity comparable to setups with a higher NA.

No quantitative demonstration of number resolved imaging is presented. This claim cannot be made without providing the fidelity with which different number states can be distinguished, and the measured rate of atom loss associated with imaging at this fidelity. Further, it is claimed that this imaging technique can be used to to deterministically prepare loaded tweezers. In estimating the fidelity of this approach, no mention is made of the imaging fidelity, which would limit one's ability to load single atoms.

As described above, we have comprehensively addressed this point by providing new data and a quantitative analysis of the fidelity and loss. We find that we can substantiate the claim to be able to resolve the number of atoms in the tweezer, but that at the present time we cannot use this technique to deterministically prepare loaded tweezers.

No quantitative measurement of parity projection avoidance is presented. The authors claim that parity projection is avoided by using a larger trap, but no quantitative information is presented to substantiate this claim. Further, the authors argue that the lack of parity projection in their system contrasts with previous strontium tweezer experiments. However, as indicated in the appendices of references 15, 16, substantially longer parity projection steps are used to isolate single atoms, presumably because parity projection actually takes a while in those systems as well.

As stated in response to referee (1) above, we have substantially modified our discussion of parity projection to better reflect the work of [15,16] and to provide a quantitative argument for the link to the trap volume.

(4) The authors claim that the use of a SPAD array provides advantages over CCD or CMOS cameras, and specifically that unlike CCD or CMOS cameras, SPAD arrays are limited by photon shot noise instead of dark counts or readout noise. However, the cooled EMCCD cameras used in tweezer and quantum gas microscope experiments are typically dominated by shot noise (with a contribution from the excess noise factor associated with gain, though this can also be avoided), and have much lower dark count rates than those of the SPAD array, much higher quantum efficiency, and negligible readout noise. There may be a cost advantage to the SPAD array, but if this is the reason to use them then it should be stated.

The main advantage of the SPAD array is speed. For short exposure times and a small number of incident photons, the SPAD camera has better signal-to-noise (SNR), as the dark count is negligible, and the EMCCD is dominated by the excess noise referred to by the referee. For longer imaging times, the EMCCD camera has better SNR, as the excess noise doesn't increase with exposure time, whereas the SPAD dark count does. In this regime, the much higher quantum efficiency of an EMCCD

is also an advantage too. We have added some text discussing this trade-off to section 3 paragraph 2.

The new data in Fig. 5, 6 and 7 shows this nicely – for a short 200us exposure we are able to image with high fidelity and loss due to the absence of readout noise (new Fig. 6). This sub-ms regime cannot easily be accessed by EMCCD cameras. Using cooling and a long exposure we clearly see the limitation of the high dark count mentioned by the referee (see Fig. 5). Here an EMCCD camera would be a better choice. However, we show that we can largely overcome this limitation by alternating cooling and imaging pulses, and by exploiting the fast time response of the camera to only read out when the probe light is on (Fig. 7). In this way we largely eliminate the effect of dark noise on our imaging fidelity.

Other issues:

• *figure 3a: I think the X axis here is supposed to be microseconds, not milliseconds. After 50ms, gravity would have caused the atoms to drop by over 1 cm.*

• *figure 3b: Perhaps there is an issue with the axes here as well. The line implies a velocity of around 50m/s, which is not what one wold expect for the quoted temperatures.*

There were indeed errors on the axes of Figs 3(a) and (b) which have been corrected.

Are the temperatures quoted in the axial or radial directions, or are the two assumed to be equal?

Both measurement methods measure the radial temperature. We have added a sentence to make this clear.