Dear Editor and Referee,

Thank you both for taking the time and effort to review our manuscript again. We have taken your comments to the manuscript, and below you can find detailed responses to the issues you raised. Once again, the paper has been significantly improved through our dialogue, and we hope this now meets your approval.

Reviewer's Comments:

In this revision, the authors added new content that substantially improved the quality of the manuscript. However, there are still some major problems and concerns that need to be addressed. In addition, the authors made many imprecise statements and mistakes. A further revision is needed.

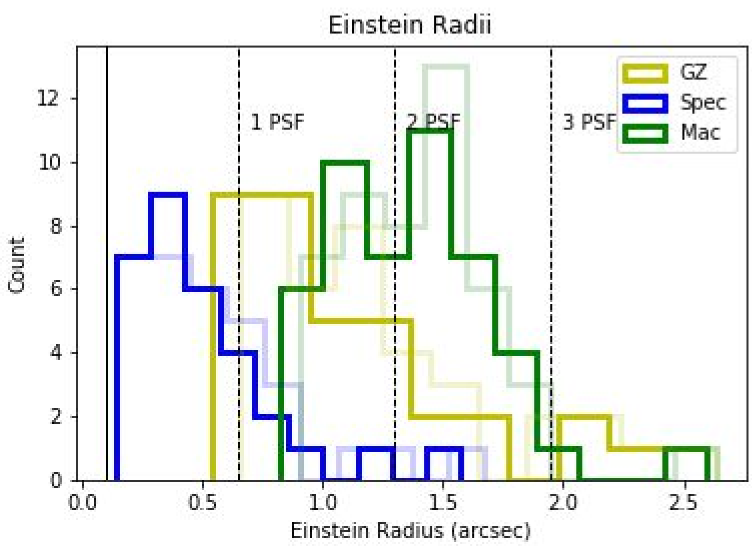
Major comments:
>>> 1. The formula for theta_E seems wrong to me. I think it should be 10**(8.09) in the denominator. That being said, it seems like the authors are not using that formula for theta_E, because it will otherwise be much smaller than what's shown in the Figures. Nevertheless, I'm questionable about all the results relevant to theta_E.

<<< Thank you for pointing this out. It was a typo in the manuscript; as written the calculation uses Gpc, which accounts for the 10^3 difference. As we present H_0 in units of km s^{-1} Mpc^{-1}, it does make more sense to write this calculation in terms of Mpc. This change in wording has been completed, though it does not change any of the calculated values.

2. Directly adopting the f_dm of 61% from Barnabe+ 2011 for lens candidates in this manuscript is inappropriate for two reasons. First, the 16 lenses in Barnabe+ 2011, with a mean stellar mass of 10^11.3, are more massive than most of the lens candidates in the GAMA spectroscopy and GalaxyZoo samples. Considering the f_dm - mass correlation, lower f_dm values are expected for those lens candidates. Second, the f_dm in Barnabe+ 2011 corresponds to the dark-matter fraction within the effective radius, while what is needed for estimating theta_E is the dark-matter fraction within the Einstein radius, which is generally much smaller than the effective radius according to the SLACS lenses. So instead of underestimating the Einstein radius as the authors claimed, I think using f_dm=61% will overestimate it for all three samples.

<<< We thank the referee for pointing out the fact that the Einstein radius will be overestimated for lens candidates with stellar masses much smaller than 10^{11} solar masses. A new paragraph in Section 3.3 “Einstein Radius Estimation” has been added to

address this. Barnabe-2011 selected their candidates and utilized stellar mass measurements from Auger-2009 in their calculations. Auger-2009 reported a mean stellar mass fraction within the Einstein radius of 40%, implying f_{DM} = 60% within the Einstein radius. Auger-2010 analyzed 73 SLACS candidates and presented a linear relation in logarithmic space between dark matter fraction within half the effective radius (which they claim to be comparable to the Einstein radius) and the stellar mass. The average f_{DM} of the data presented there is ~ 55%. We applied the empirical fit to our data, which does indeed account for lower dark matter fraction of smaller galaxies. However, this fit (like all assumptions taken from SLACS candidates) is taken from high-mass galaxies, and therefore does not in itself solve the relationship of f_{DM} to stellar mass in lower-mass galaxies. This change to a sliding f_{DM} based on stellar mass results in small shifts of total-mass estimates and Einstein radius estimates for all candidates. This is shown in the figure below, compared with the Einstein radius estimated from the average f_{DM} of 61% (the lighter faded lines). Note, this plot is binned slightly differently from the one in the paper. We have shown it this way because it shows the differences between the two calculations more clearly.



This difference does not meaningfully impact our results. However, we chose to recalculate everything throughout the paper using this more appropriate sliding dark matter fraction, and everything in the paper now includes this adjustment. As the referee will certainly note, this led to a review of GalaxyZoo selection. This reevaluation of the material was a healthy and instructive review of our work, and our conclusions remain almost entirely unchanged by these improvements to candidate sample and Einstein radius estimates.

3. I'd be very cautious about using "maximum total mass" despite the leading "a soft constraint". To me, "maximum" indicates a limit, while in fact there is no hard limit. Some SLACS or many BELLS lenses clearly have Einstein radii larger than the radius of the SDSS or BOSS fiber.

Also, this "mass" is simply not a good property of the lens because it changes with the source redshift. It seems not meaningful to introduce this "maximum mass". If the authors really want to, I suggest using a term like characteristic mass.

<<< We have considered the language and understand the concern. We have clearly stated in the paper that there is no hard limit to identification, and that it is possible for features of lenses with Einstein radii larger than the fiber to be recovered in a mixed spectroscopy approach. We have changed the wording to reflect this, introducing a “fiber radius mass” defined as follows: the precise mass of a lensing system (of which the lens and source redshift are known) for which the estimated Einstein radius equals the radius of the spectroscopic fiber, assuming the fiber has been positioned precisely at the center of the lensing object. The geometry and surface brightness of the features, the placement of the aperture being not exactly centered on the center of the object, and chance alignment can and do allow the detection of lensing features of objects with Einstein radii larger than the fiber. However, as has been documented, the probability of detection goes down rapidly when the Einstein radius becomes larger than the aperture. This is the area where the imprecision lies, due to the complexities of observation and the fact that these lensing features are most frequently not close to being exact Einstein rings. That does not change what the Einstein radius is, for which there is a specific enclosed mass (for a given measured source and lens redshift) that corresponds to the radius of the spectroscopic fiber. Since we do have source and lens redshifts for the GAMA spectroscopy candidates, this fiber radius mass is a significant property. If the enclosed mass is higher than the fiber radius mass, the Einstein radius will extend beyond the fiber radius and diminish background source flux, corresponding to a diminished probability of detection by mixed spectroscopy. In this way, the fiber radius mass represents a meaningful upper likelihood constraint to the mass of GAMA spectroscopy candidates and is therefore an essential element to understanding the disparity between GAMA spectroscopy candidates and higher mass candidates of the other two methods. In addition, we have addressed the issue of unknown source distances for LinKS and GalaxyZoo candidates by calculating curves that show the values for which the combination of lens redshift and stellar mass result in Einstein radius estimates that equal the fiber radius of 1 arcsecond, parameterized by the ratio of the source distance to lens distance. These calculations are presented in the body of the manuscript and in a new figure.

4.It's worth emphasising that the comparisons in Figure 15 need to be taken with caution because different selections have been applied. For example, the SLACS, S4TM, and BELLS candidates all require certain S/N thresholds in the detected emission lines. S4TM and BELLS candidates also require the expected Einstein radii to be larger than certain thresholds. These requirements are not applied to the GAMA spectroscopy sample. In addition, only a subset of all the spectroscopically identified candidates in the SLACS survey that have HST observations are reported in Bolton+ 2008.

<<< SDSS S/N is by definition different from GAMA but a S/N requirement was used for both GAMA as a whole and GAMA lens selection. See Holwerda+ (2015). The numbers presented for the SLACS survey are based on our best read of the published SLACS literature. We assumed the 131 SLACS targets reported in Bolton+ (2008) were representative. If we misread the reported SLACS numbers, please correct us. We note the difference in emphasis in both the text already and now in the caption of Fig 15 as well.

SLACS and S4TM do not report the number of candidate objects that were the start for their HST Snapshot programs, just the number of objects observed, which is intrinsically lower for a snapshot program. Snapshot efficiencies are typically around 35-40% but varies greatly. So the S4TM and SLACS numbers are lower limits. The BELLS survey reported a number for a GO program and all their candidates were observed with HST. Our read of the SLACS and S4TM survey papers could not reveal their snapshot target list length or HST observing efficiency.

Other comments:
Abstract
>>> Please make it clear that the 0.72 per sq. deg refers to the density of candidates.

<<< This has been addressed.

Introduction
>>> The statement of "Strong gravitational lenses appear in every respect to be just like other elliptical galaxies" is imprecise. Obviously there are also spiral lenses and cluster lenses.

<<< This is true and has been addressed.

>>> It's necessary to clarify that strong lensing only provides a robust total-mass measurement in the inner region of the lens.

<<< That is true. We re-emphasized this in section 3.3 and in the introduction.

>>> Lensing doesn't help to constrain the Fundamental plane.

<<< It does! As soon as one has the measures for the mass, the m/l is constrained, which a fundamental ingredient in this fundamental plane. That is for further studies of these lenses however.

>>> Please add at least Vegetti+ 2012 for dark-matter substructure, Suyu+ 2017 for H0LiCOW.

<<< This has been addressed.

>>> I don't quite understand "volume-weighted to greater mass". Can the author explain on "volume-weighted"?

This refers to the Malmquist bias that results in higher mass galaxies being over-represented in apparent luminosity limited surveys.

Section 2
>>> Please add Shu+ 2016 for the BELLS survey.

<<< This has been addressed.

>>> I don't understand the statement "LRG sample leads to effective constraints on redshift to mr < 19.5".

<<< Spectroscopic redshifts can be obtained for SDSS using the 4000A break. For the LRG subsample that means a fainter limiting depth to which spectroscopy can be done. We have included this point.

>>> I'm sorry about my incorrect statement of "the SLACS survey imposed a redshift cut of z > 0.15" in my previous reply. After checking the relevant references, it's unclear whether the SLACS survey made such an explicit cut. In fact, ~20 SLACS lenses are at z < 0.15. Also, the lens redshift range of the SLACS survey is 0.05-0.5 instead of 0.15-1.

<<< We thank you for your input in every regard! Thank you for clarifying your response.

>>> Move "In the following calculations we adopt ..." to the end of Introduction. Other relevant cosmological parameters, i.e. Omega_m, Omega_Lambda, need to be defined as well.

<<< This has been addressed.

Section 2.1
>>> "verified" is misleading. Chan+ 2016 suggested 10 of them are probable lens systems, and only one of them has a valid lens model.

<<< This has been addressed.

>>> "structurally misses lower redshift and higher mass lens candidates" as compared to what samples?

<<< As compared to the candidates it does find. An additional statement referring to later sections has been added for clarity.

Section 2.2.2
>>> Please refer to Marshall+ 2016 when introducing citizen science method.

<<< This has been included.

Section 3.1
>>> It's worth explicitly mentioning that you only consider PG+ELG blends from Holwerda+ 2015 in this manuscript, because the other types of blends could also be lenses as discussed in Chan+ 2016.

<<< Thank you for pointing this out. We have added a clarifying to section 2.1 where we introduced the GAMA sample.

Figure 3
>>> For the sake of clarity, please use passive galaxy redshift and emission-line galaxy redshift for the labels.

<<< The figure has been reassessed and presented now explicitly as the relationship between ELG and PG redshift. This is much more clear, and we appreciate the referee's suggestion.

Figure 4
>>>It'd be better to show two more higher-score candidates here, because a lower limit of 30% is applied.

<<< Two candidates, with scores ~ 35 and 40% have been added to the figure.

Section 3.2
>>> I suggest moving "The GalaxyZoo team imposed a redshift restriction of ..." to Section 2.2.2 where the GalaxyZoo sample is first introduced.

<<< This has been addressed.

Section 3.3
It'd be better to use the velocity dispersion to estimate the Einstein radius. That way one doesn't need to fold in the dark-matter fraction, which brings another source of uncertainty. If spectra are available for all the lensing galaxies, it should be straightforward to compute the velocity dispersions, say using pPXF. If that doesn't work, a fiducial velocity dispersion value may be used. One just needs to make a sensible choice of the fiducial value.

<<< This is indeed the best way to estimate Einstein radii. The GAMA spectroscopic signal-to-noise and spectroscopic resolution is not of such high quality that the velocity dispersion is something we can reliably measure. We discussed this within the GAMA team as a potential application but it proved unfeasible (Baldry private communication). The pPXF idea is something we absolutely want to pursue and we have asked for time on GEMINI and Keck for high s/n spectroscopy which would allow for a kinematic mass estimate, bypassing our need to assume a M/L (stellar pop model) and a dark matter fraction (kinematic measure). It is the next step in this project but unfortunately not feasible with current data.

>>> Does the stellar mass from MAGPHYS catalog also assume a Chabrier IMF? If yes, please explicitly say it.

<<< Yes! This has been addressed.

>>> "distance" should be angular diameter distance.

<<< This has been addressed.

Section 4.1
>>> The current Einstein radius estimation using f_dm for the three GAMA/KiDS samples is quite uncertain, so comparisons with measured Einstein radii of the SLACS/S4TM/BELLS samples seem inappropriate/meaningless. I also suggest using expected Einstein radii instead of "Einstein radii" when referring to the three samples in this work.

<<< The term "estimated" has been added to all references to the estimations conducted on the three samples. We thank the referee for this suggestion.

Section 4.2
>>> It would be useful to quote the mean/median redshift and stellar mass of the three samples.

<<< This has been addressed.

Section 4.3
>>> It may be better to move this section to the beginning of Section 4.

<<< We believe the overlaps are best introduced following the description of sample characteristics.

Section 5

>>> More details on the K-S tests are needed, for example, how large are the parent samples that are compared to the three lens candidate samples.

<<< Parent sample sizes have been included for all three K-S tests, and the machine learning parent sample has been revised. Section 5 has been rewritten to include these changes.

>>> As far as I understand, the LinKS sample does not have the r < 20 cut.

<<< Petrillo-2017 and Petrillo-2019, which detail the selection process in more detail, do include the r < 20 cut in the description of the LRG selection. Petrillo-2018, where the LinKS candidates are presented, presents all other criteria exactly the same as the other two papers but moves the statement of the r < 20 cut to Section 2.1 referring to the selection of the “full” sample, from which the LRG sample is selected.

>>> The statement of "Lensing searches of any kind prefer higher mass" seems too strong. I also don't agree that there is an upper mass limit for the spectroscopy method. Please see my other comment at the beginning.

<<< See above for our position on the upper mass constraint and a fiber radius mass parameter. The statement in quotations above has been expanded to clarify and qualify the statement in the context of this study and in the existing strong lens literature.

Figure 13
>>> What does "the minimum (1") utilized for the training set" mean? I don't see any 1" cut in Petrillo+ 2019. Please clarify.

<<< Petrillo-2019 lists the parameters used for the construction of the “mock lenses” that constitute the positives for their training set. They describe the Einstein radius parameters to be selected from a logarithmic distribution ranging 1.0-5.0 arcsec. Petrillo-2018 refers to this in the first paragraph of Section 4.2. This has also been added to the paper in the introduction of the LinKS survey (Section 2.2.1).

Section 6.2 (and Figure 10)
>>> First of all, the f_dm of 61% can not be directly used. From that f_dm, one can only get the stellar mass within the effective radius. The resulting "upper limit to stellar mass" can not be directly compared to the total stellar masses of the other two samples. Even if the stellar mass within the same radius for the other two samples can be derived, them being higher than the "upper limit to stellar mass" does not necessarily lead to what the authors concluded about "This ensures ..." because 1) their source redshifts could be different; 2) their dark-matter fractions could be different.

<<< As discussed above, a first estimate using a dark matter fraction of 61% is informative for the purposes of this paper, but we chose to review these estimates utilizing a sliding dark matter fraction as a function of stellar mass. While the source redshift is unknown for the other two samples, we have shown that there does exist parameter space where the stellar mass does far exceed the *fiber radius mass* for GAMA spectroscopy candidate configurations and for a range of possible source-to-lens distance ratios. The word “ensure” does imply a certainty that is not present here and has been changed to reflect the uncertainty. We thank the referee for the input.

Section 6.4
>>> It's confusing to see that only 7 of the LinKS candidates pass the LRG selection. From my understanding, all LinKS candidates from Petrillo+ 2019 passed the LRG selection in the first place. Please clarify.

<<< Thank you for pointing this out. A mistake was made in the application of the criteria to the data, and a review of that work has resulted in new numbers and new conclusions from this section. Section 6.4 has been rewritten.

Figure 15
>>> The density for the BELLS sample is wrong. The BELLS survey contains 45 candidates selected from the first partial year of BOSS data, which is roughly 10% of the final BOSS survey. So the density will be more like 45/(10000 sq. deg * 10%) = 0.045.

<<< Thank you for clarifying for us. The plot has been updated with the correct number.

>>> For the sake of clarify, please include how exactly those densities are computed with necessary references in the manuscript, or at least in the reply.

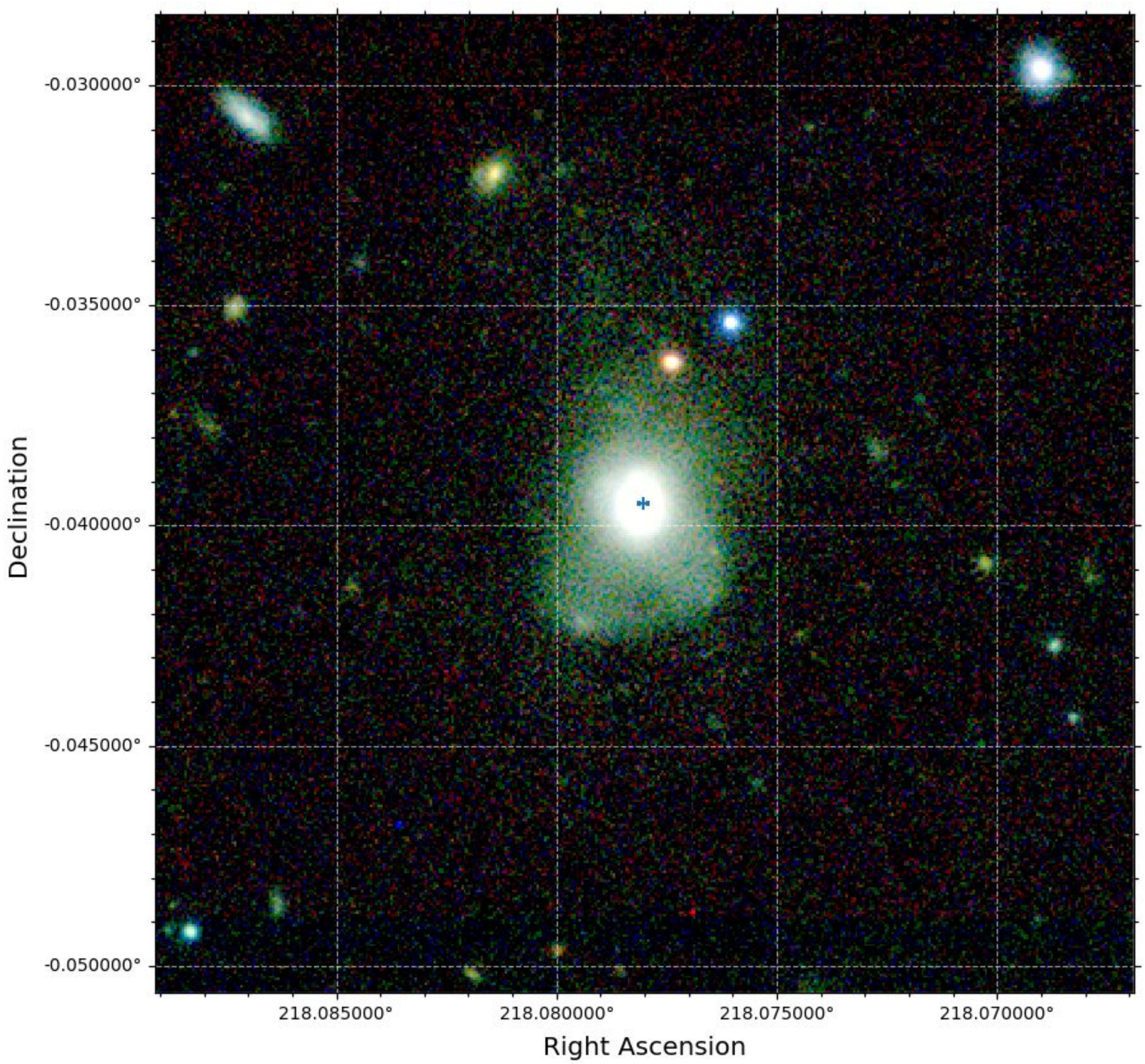
<<< Here are the numbers used to compute the candidate density for each survey.
SLACS: 8000 deg^2; r < 19.5; 131 candidates; *0.016 deg^{-2}*
(SLACS V - Bolton et al 2008; SDSS I/II from sdss.org)
S4TM: 9380 deg^2; r < 19.5; 118 candidates; *0.013 deg^{-2}*
(S4TM - Shu et al 2017; SDSS-DR7 - Abazajian et al 2009) - SDSS-DR7
BELLS: 1000 deg^2; r < 19.5; 45 candidates; *0.045 deg^{-2}*
(BELLS - Brownstein et al 2011)
GAMA spectroscopy (equatorial): 180 deg^2, r < 19.7; 47 candidates; *0.2611 deg^{-2}*
LinKS: 904 deg^2; r < 25.0; FWHM (r) ~ 0.65; 308 candidates; *0.34 deg^{-2}*
(Petrillo et al 2018)
LinKS (in GAMA equatorial): 180 deg^2; 47 candidates; *0.2611 deg^{-2}*
GalazyZoo (in GAMA equatorial): 180 deg^{-2}; 16 candidates; *0.089 deg^{-2}*
DES: 5000 deg^2, r < 24.9, FWHM ~ 1.06 (g); Candidates: 511; *0.10 deg^{-2}*
(Jacobs et al 2019)
DECaLS: 9,000 deg2; r < 22.5; FWHM ~ 1.0; Candidates: 335; *0.037 deg^{-2}*

(Huang et al 2019)

Section 6.5

>>> I'm surprised that there are features seen in the KiDS data for G593852 but not in the HST data. Can the authors show the KiDS data for this system? What's the exposure time in F625W for G593852?

<<< Below is the KiDS image of G593852. It has been added to the figure. The F625W is blue but shallow exposure and the ring-like structure is not particularly blue compared to the lens.



Section 7
>>> I'd be cautious in calling the three methods "ineffective means". I think it's fine to say each of them has its own limitation.

<<< This phrase refers to their being ineffective means to vet the other two methods. Their limitations in the identification of lens candidates is discussed separately.

>>> "extend candidate selection to a lower mass range" as compared to what?

<<< This lower mass range is in reference to the mass range occupied by the majority of previously identified lens candidates, including those identified in SDSS. A phrase has been added to the sentence for clarity.

The authors wish to express once more our sincere appreciation for the time and consideration of this manuscript. We wish you the best in these unprecedented times.

With highest regards,

Shawn Knabel