Fluorescence Confocal Laser Scanning for in vivo Cochlear Imaging

BRIAN FROST

Electrical Engineering PhD Candidate, Columbia University
bf2458@columbia.edu

Abstract: Confocal microscopy is a well-established imaging modality, wherein an image of a sample is reconstructed pointwise through strong focus of light both at the sample and the detector. Confocal systems allow for imaging of volumes as well, but do not generally offer imaging at depths as deep as optical coherence tomography or ultrasound imaging. In cochlear mechanics research, confocal microscopy has been used to image structures such as the hair cells and the tectorial membrane *in situ* or *ex vivo*. Recent developments in cochlear mechanics have tended towards the use of optical coherence tomography for *in vivo* imaging and vibrometry. However, optical coherence tomography does not provide an image of the tectorial membrane in most cases. We present here a design for a confocal microscope system which can be used alongside an optical coherence tomography system to image the tectorial membrane *in vivo*. We simulate a number of these systems in Zemax OpticStudio to determine the optimal parts required to build this system alongside a currently operating optical coherence tomography system.

1. Research Report

1.1. Motivation

At the Fowler Laboratory at Columbia University Irving Medical Center, we study mammalian cochlear mechanics *in vivo*. Optical coherence tomography (OCT) has become popular in the cochlear mechanics field, as it facilitates imaging and nanometer-scale vibrometry at a depth non-invasively. The cochlea is a very small and delicate structure, and most invasive imaging modalities will compromise the cochlea's functionality.

OCT works based on reflected light, and thus does not require fluorescent labeling of tissues to properly function. This is generally a benefit of OCT, however there exist structures within the ear which have very low reflectance. For example, the tectorial membrane (TM) is a gelatinous structure that lies above the hearing organ within the cochlea which we have been unable to resolve in OCT images of gerbil cochleae. For the sake of brevity, we will not discuss the role the TM plays in cochlear mechanics, but it should be noted that its function is somewhat understudied due to its being quite difficult to image *in vivo*.

We would like to perform experiments at the Fowler Laboratory in which we change the properties of the TM through profusion or drug injection, and then measure the effects such a change has on the vibrations within the cochlea. The *in vivo* vibrometry will, as usual, be performed using OCT, but it is important that we can see how the structure of the TM has actually changed as a result of our profusion or injection. While this need not occur *in vivo*, it is important that it is performed without having to break the bone encasing the cochlea so as to ensure that the TM is not torn, moved or dried out in any way. Thus, an imaging modality that can perform imaging at a depth and that can actually produce an image of the TM is needed.

Strimbu et al showed in 2018 that the TM has a higher Ca2+ concentration than the fluid surrounding it, making it a good candidate for Calcium indicator fluorescence imaging. In Strimbu's work, confocal microscopy was used to produce images of the TM after staining it with Asante Calcium Red-1 (ACR-1). These images were taken using a Zeiss LSM 780 confocal

laser scanning microscope (CLSM). [1]

Motivated by the work of Strimbu and his colleagues, we would like to use CLSM and ACR-1 to image the TM. Given that we already have a ThorLabs Telesto OCT system [2], it would be best if the CLSM could be built as an add-on to the OCT system. As both system's use 2-D scanning mirrors, and require relatively expensive scanning objective lenses, a lot of money could be saved by simply using the same scanning hardware and objective lens for both systems.

1.2. Theory of Confocal Laser Scanning Microscopy

CLSM systems can acquire light from only one "point" at a time, and reconstruct three-dimensional volumes by scanning the sample. Planar (or XY) scanning is usually achieved through the use of XY mirror galvonometers, while the depth (or Z) dimension is usually scanned through motion of the sample towards or away from the objective lens. The meaning of "point" is important here – the resolution of the CLSM in all three dimensions is determined by the spot size in all three dimensions for the system when it is acquiring light from a single XYZ coordinate in the sample.

CLSM systems usually operate based on fluorescence, in which a sample is excited by light at one wavelength and it, in turn, emit light at a longer wavelength. Fluorescence imaging with CLSM can be understood through the consideration of excitation and fluorescence independently. That is, we can consider each wavelength of light in the system separately.

Figure 1 shows the excitation direction of a CLSM system. With collimated light as the assumed input, scanning mirrors direct the angle at which light enters the objective lens. Light focuses at a single point determined by this angle, and that point will act as a fluorescent source.

The optical system in Figure 2 shows the configuration of a CLSM collecting light from a single coordinate in the sample – this is the fluorescent point of view. The size of the pinhole at the point detector determines both the resolution and the light intensity detected. For example, in an ideal system, a 0-diameter detector would receive light only from the exact focus of the objective. This is, of course, infeasible – not only are points non-existant, but the point spread function (PSF) of a lens is not simply an impulse but instead has volume in the three-dimensional sample. As a result of the non-ideal PSF, light from other points will necessarily enter the pinhole as well. Beyond this, a smaller pinhole will necessarily collect less light, and thus can suffer from significant signal-to-noise ratio (SNR) decreases.

It is important that we characterize the resolution of this system in all three dimensions, and also determine the effect of pinhole size on the system. We know that the lateral resolution (in XY) is determined for classical microscopes by the numerical aperture (NA) of the objective and is a result of the half-maximum width of the Airy's disk: $r = \frac{1.22\lambda}{2\text{NA}}$. In the case of fluorescent confocal microscopy, the detected light from the sample is also focused by the same optical elements to this same point. That is to say that there is a PSF for the excitation direction and a second for the detection direction, each for a lens with the same numerical aperture. The outcome is a narrower PSF which is the product of the excitation and detection PSFs. From this, we can derive

$$r_{XY} \approx \frac{0.51\lambda}{\text{NA}}.$$

Importantly, this derivation assumes the excitation and fluoresced wavelengths are near. The important takeaway is that fluorescent CLSM has a smaller XY resolution than conventional microscopes. Notable also is that in scanning, the PSF will change as a result of spherical aberrations, so the true XY resolution will be determined by the largest XY resolution observed in the field-of-view (FOV) of the device.

Axial resolution (that is, resolution in the depth or Z dimension) is slightly more complicated, as the behavior of light traveling deeper into a material is determined largely by the refractive index n. [3] The following formula is often used for the axial full-width half-maximum (FWHM)

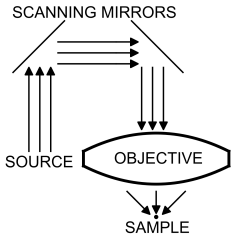

Fig. 1. The excitation direction of a CLSM system.

of the CLSM system:

$$r_Z \approx \frac{.88\lambda}{n - \sqrt{n^2 - \mathrm{NA}^2}}.$$

From these equations we see that n, λ and NA are the parameters which determine resolution entirely. Visible light is often used, mostly due to the availability of fluorescent dyes in the visible spectrum, but it is also notable that visible light CLSM will have a better resolution than infrared CLSM. It is also clear that a large NA lens is preferable, although wide cones of rays can introduce more error due to spherical aberrations. Lastly, n is a property of our sample, and cannot be changed.

The pinhole, in simulation, can be made any size without having to worry about light intensity issues. However, in real samples, it is not advised that the pinhole is made any smaller than the diameter of the Airy disk of the collector lens. A larger pinhole allows in light from a broader range of points and thus increases the XYZ resolution from that determined by the equations shown above. Any smaller pinhole will allow in less light, but at very small pinholes, the “ifinitely small pinhole diameter" approximation does yield better resolution. The derivation of these equations is left out (can be found in [3]), but they are given by:

$$r_{0,XY} \approx \frac{0.37\lambda}{\mathrm{NA}},$$

$$r_{0,Z} \approx \frac{0.64\lambda}{n - \sqrt{n^2 - \mathrm{NA}^2}}.$$

1.3. Sources of Non-Ideality and Metrics of Interest

There are a few important sources of non-ideality in CLSM systems, and these non-idealities will guide how we characterize our simulated system. The equations above for XY and Z resolutions provide an important starting point – the PSF for the excitation and fluorescent direction of the objective lens are both important.

The first important aspect to note here is that a lens does not necessarily behave the same way in both forward and backward directions – spherical aberrations can play varying roles depending on the side of the lens at which light is incident. Furthermore, the excitation and fluoresced light are at different wavelengths, so they are impacted differently by the chromatic properties of the lens. For example, we know that lenses have different foci for different wavelength light.

This suggests that the PSF of the objective lens in its forward configuration is of interest, as is the PSF of fluoresced light in the backwards direction. We may also be interested in measuring the effect on focal length by changing wavelength within the visible range in both directions of operation. For our particular case, a lateral (XY) resolution of anything below 5 μm is acceptable, as this is would allow us to make out structural changes along the tectorial membrane’s width.

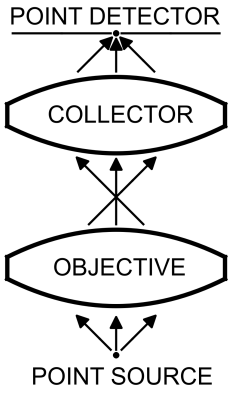

Fig. 2. A simplified diagram of how a confocal microscope performs point-wise detection.

The TM is about 200-300 μm thick, so axial (Z) resolution much greater than 50 μm or so would not be acceptable.

Clear also from the resolution equations, we know that the NA of the objective is very important. In the fluorescent direction, the cone of rays which enter the lens is determined only by the geometry of the lens, so the NA is a fixed value. On the other hand, the beam diameter in the excitation direction can be varied independent of the lens geometry. While larger NA theoretically offers better resolution, spherical aberrations and astigmatism impact a lens' operation more as the NA increases. Thus, it will be important to determine an optimal beam diameter for the system, which will depend largely on the objective lens.

All of these metrics are dependent, of course, on the angle at which light enters the objective. This is importantly linked to the FOV of the device. We want to maximize the FOV within a given resolution constraint. For our purposes, we use the OCT to image the basilar membrane (BM) in the basal turn of the cochlea in gerbils. The TM rests atop the BM, and at basal turns the BM is no more than 200 μm wide. Thus, a 200μm or larger field of view is necessary for our purposes.

We must also recall the role of the collector lens. The collector's role is to focus quasi-monochromatic fluoresced light, so chromatic aberrations are of little impact. However, we do not want the collector to be an aperture stop for the system (or else we will lose signal power), so the collector must be at least somewhat "filled" with light. Thus, spherical aberrations play a large role in the collector's operation. The spot size of the collector for large input angle is not actually important for the resolution of the system, but it is important for maximal light collection.

Looking at the Zeiss LSM series of CLSMs, we can see that our constraints are well within the realm of possibility. These systems give an 11.3mm $\times$ 11.3 mm field of view with 1.25$\times$ objectives, so even with a 10$\times$ objective, a 200 μm range should be feasible. We know also that 1 μm resolution is achievable with high NA objectives and small pinholes. [3]

2. Simulation Report

2.1. Zemax Simulation Strategy

We wish to determine the best choices for an objective lens and a collector lens so as to meet the constraints described in the previous section. To do so, we consider the two operation directions of the CLSM independently.

2.1.1. Excitation Direction

In the excitation direction, as shown in Figure 1, we have only the objective lens to consider. In Zemax, we create a collimated light source of varying beam diameter, and place two scanning mirrors in the light path. Lastly, we have the objective and the object plane at which the light focuses. This architecture is shown in the Zemax environment in Figure 3.

For ACR-1, the excitation wavelength is 500 nm and the fluorescence wavelength is 640 nm. [4] We test objective lenses from the scanning lens catalog of Thorlabs within this wavelength range, all of which have Zemax models provided through the Thorlabs website.

From the excitation direction alone, we can obtain information about the field of view of the system. We desire a field of view of at least 200 μm $\times$ 200 μm, with a lateral resolution of 5μm. Thus, we are interested in the range of scanning mirror angles in which the FWHM at the object plane is less than 5 μm.

Given the symmetries of the objective lenses tested, the x- and y-FWHM values will be the same if the x- and y- mirrors are scanned by the same amount – that is, if we vary the position being imaged in the object plane along a line segment at a 45 degree angle from the horizontal, the FWHM values for x and y will be the same. Given this symmetry, we can vary the scanning mirror angles in this way to test the field of view through a single dimension of measurement.

We care about FOV for both wavelengths, as the fluorescent wavelength's PSF also determines the resolution of the system (as discussed in the research report). So long as both wavelengths provide a sufficiently large FOV, the lateral resolution and FOV constraints can be said to be met by the given lens.

For each lens we must also test the maximum beam diameter before which the resolution becomes worse as a result of spherical aberrations. We would like to achieve a higher beam diameter as according to the reslution equations given in the research report, a higher NA yields lower resolution (both lateral and axial). However, a weakening of the focus as a result of spherical aberrations has the opposite effect.

2.1.2. Fluorescence Direction

While the fluorescence direction, as discussed in the research report, also has some tightening effects on the lateral resolution, our criteria in the excitation direction are enough too ensure our lateral resolution specification is met. We choose to do this partly because the fluorescent direction is far harder to simulate.

In Zemax, we consider the object plane as containing spaced sources at the fluorescent wavelength emitting cones of light. These cones enter the objective, and then the collector. They are lastly focused at the image plane. This setup is shown in Figure 4.

Within sequential Zemax, a pinhole cannot be simulated. Instead, we can simply observe the intensity of light radiated from different points in the object within some distance from the center of the image plane. However, this does not account for the heightened resolution that can be achieved with very small pinholes, as described in the research report.

As a result, we do not so much observe this entire fluorescence direction architecture, but rather the collector alone. The collector is meant to collect quasi-monochromatic light in a relatively large cone and bring it to a sharp focus at the pinhole. This means that chromatic aberrations do not affect the operation of this device, whereas spherical aberrations greatly impact the collector's performance.

To choose the best possible collector, we consider the Thorlabs catalog of aspheric lenses in the relevant wavelength range, and consider their NA and spot size. We do not want the collector to act as an aperture stop for the system, so the collector NA must be larger than that of the objective. We choose to consider only aspheric lenses, as they are designed to be minimally affected by spherical aberrations.

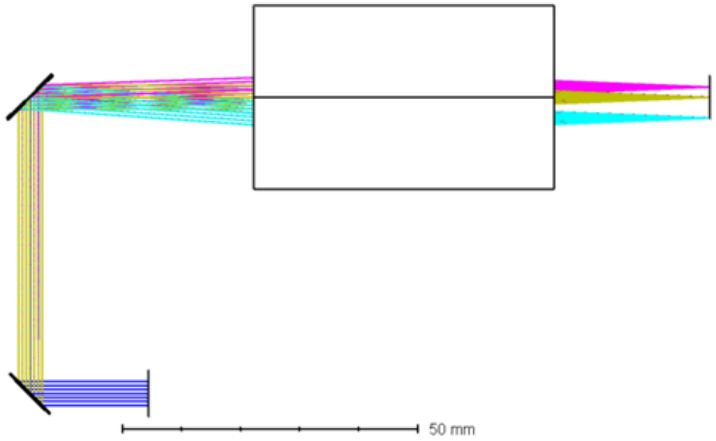

Fig. 3. Screenshot of the Zemax architecture for the excitation direction. Different color beams correspond to different configurations of the scanning mirrors, i.e. to different points of focus along the object plane. The objective lens used in this specific image is the SL50.

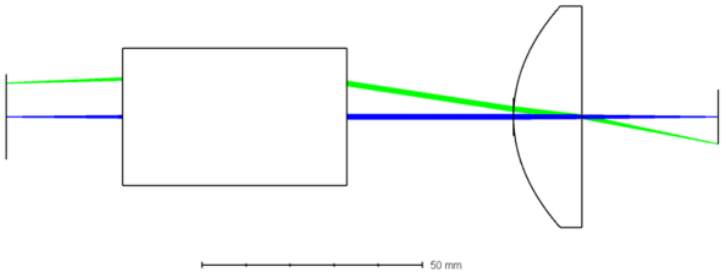

Fig. 4. Screenshot of the Zemax architecture for the fluorescence direction. Different color beams correspond to different object locations. The objective lens used in this specific image is the SL50.

2.1.3. Axial Resolution

The FOV measurements and beam diameter measurements in the excitation direction give us sufficient information on the lateral resolution of the system, as well as the FOV, the impacts of chromatic aberrations and the impacts of spherical aberrations in the objective. Our study of collector spherical aberrations gives us an understanding of all remaining aberrations in the system. However, we have still failed to characterize axial resolution.

This is quite difficult to do directly using the Zemax sequential software, not only due to the fact that pinholes cannot be simulated, but also due to the fact that three-dimensional point-spread functions are not available. Thus, we must compute the axial resolution indirectly. We have the objective lens NA from the beam diameter determined, and we know that the cochlea is mostly filled with water, so $n = 1.33$. From this, we can determine axial resolution at moderate and small pinhole sizes.

Although it is imprecise, we then test these axial resolution values by changing the distance between the objective and the object plane. We can get an understanding of the axial resolution by seeing how the maximum height of the PSF varies. We can do this at both the center and extreme locations.

2.2. Lenses Tested

The Thorlabs catalog offers a number of objective lenses which they call “scanning lenses", optimized for imaging modalities which require scanning light sources (such as OCT or confocal). Within this category, exactly three lenses fall within our wavelength range: the LSM03-VIS (hereafter LSM03), the CLS-SL and the SL50-CLS2 (hereafter SL50). The specifications for these lenses can be found on the Thorlabs website.

The LSM03 and SL50 offer very similar working distances – 25.1 mm and 26.4 mm respectively. The CLS-SL on the other hand, offers a 54 mm working distance. While a longer working distance is a wonderful convenience, it unfortunately also leads to a lower NA, and thus a higher lateral resolution.

For collectors, we consider one CNC-polished aspheric lens and one MRF-polished aspheric lens. MRF-polished lenses are designed to have diffraction limited spot sizes, while CNC-polished lenses can have larger NAs.

Of the CNC-polished lenses, we choose a “high-precision" lens in our wavelength range: the ASL10142-A (hereafter ASL). This lens has an NA of 0.143, and a focal length of 79mm. For the MRF-polished lens, we choose the AL2550H-B (hereafter AL). All MRF-polished lenses available have the same NA (0.2), so for a fair comparison to the CNC-polished lens, we choose one with a similar focal length (in this case 50 mm). Again, the specifications for these lenses are available on the Thorlabs website.

2.3. Results

2.3.1. Excitation Direction

We consider the LSM03, SL50 and CLS-SL in the configuration shown in Figure 3. We first hold the scanning mirrors fixed to determine the max beam diameter after which the resolution increases, using light at the excitation wavelength. We define this point either by the point after which the PSF begins to spread as the input beam diameter is raised or the point after which the height of the PSF at the object plane decreases below 0.9 (90% max). The results are shown in Table 1.

The CLS-SL allows a far larger beam diameter than the other two tested lenses. The CLS-SL also has a higher working distance, however, so its numerical aperture is actually lower than those of the other two tested lenses. This table also gives the estimated axial resolution for moderate and small pinholes according to the formulas in the research report, with λ being the average of the excitation and fluorescence wavelengths (570 nm) and $n = 1.33$. We can see that the SL50 shows an axial resolution around the desired value of 50 microns, although it is a bit larger. The CSL-SL and LSM03, unfortunately, give an axial resolution far higher than is desired. As a result, we do not test with them in our other experiments.

Objective Lens	Beam Diameter	NA	Axial Resolution	Small-Pinhole Axial Resolution
SL50	7.0 mm	0.133	75.2 μm	54.7 μm
LSM03	5.0 mm	0.100	133.2 μm	96.90 μm
CLS-SL	11 mm	0.101	128.4 μm	93.4 μm

Table 1. Optimal beam diameters for the viable objective lenses, as well as the resultant NAs, and estimated axial resolution values for small and moderate pinhole sizes.

Next, we test the FOV. We move the scanning mirrors in tandem, and observe the angle at which either the FWHM is greater than 5 microns or the maximum of the PSF is less than 0.5.

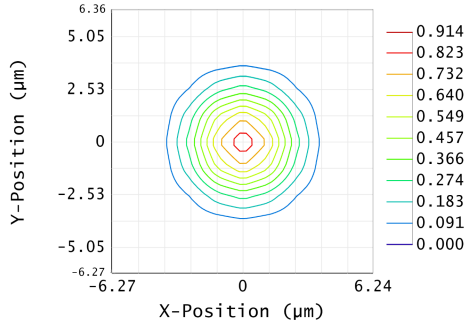

Fig. 5. Contour plot of the PSF of the excitation direction using the SL50 objective at the center of the FOV.

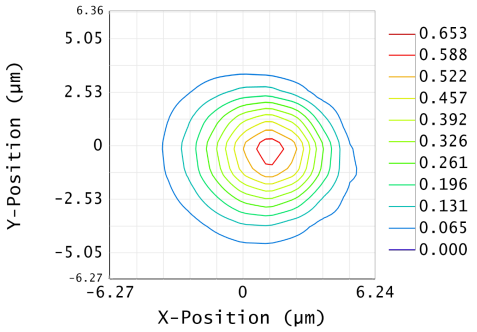

Fig. 6. Contour plot of the PSF of the excitation direction using the SL50 objective at the corner of the FOV.

We do this for both the excitation and fluorescence wavelength. The results are given in Table 2.

The SL50 has an FOV far larger than our 200 μm $\times$ 200 μm requirement, both for the excitation wavelength and the fluorescent wavelength. The PSF at the center of the FOV is shown in Figure 5, and the PSF at the corner of the field of view is shown in Figure 6.

Objective Lens	Working Distance (mm)	FOV at 500nm	FOV at 640nm
SL50	26.4	3.5mm×3.5mm	2.6mm×2.6mm

Table 2. FOV results for the tested lens.

To test the axial resolution values in Table 1, we move the object closer to and further from the objective lens. We do this both for the center of the FOV and the corner. The results are summarized in Table 3.

This shows that the PSF decrease to about half of the focus value (or less) 75 microns away from the focus. This is consistent with our axial resolution of about 75 microns. There are two reasons this measurement is not entirely valid, however. For one, Zemax sequential mode does not allow for the inclusion of a material with varying index of refraction. Thus, this measurement assumes an index of 1, artificially *lowering* the measured resolution.

On the other hand, like the lateral resolution, the measured axial resolution in the excitation

direction is only one component of the axial resolution. The total axial resolution depends also on the fluorescent direction. This artificial *raises* the measured resolution. The effects of these two inconsistencies aside, this is the best we can do with sequential Zemax to estimate axial resolution.

Distance from Focus	0	25 μm	50 μm	75 μm
SL50 Center PSF Peak	0.914	0.785	0.616	0.437
SL50 Corner PSF Peak	0.653	0.526	0.386	0.280

Table 3. PSF peak values within the field of view of the SL50 lens, at 500 nm, as the object is moved closer to and further from the lens. This result is consistent with the computed axial resolution.

In sum, we must pick the SL50 based on its axial resolution. While the other lenses prove to have sufficient FOV and lateral resolutions, their axial resolution is so poor that they may not be considered. The FOV and working distance for the SL50 are better than necessary, but unfortunately the axial resolution is just barely sufficient. The SL50 lens costs $3,257.18.

2.3.2. Collector Lens

Finally we consider a comparison between our two collector lenses – the CNC-polished ASL, and the MRF-polished AL. We want to pick the lens which gives the smallest spot size when sufficiently filled with light.

Each lens has a much larger diameter than the beam diameter used for the SL50 lens. We know that ideally, the MRF-polished lens should have a diffraction-limited spot size. We consider the lenses completely filled with light, as well as the lenses with 9mm input beam diameter. The corresponding spot diagrams are shown in Figure 7 and 8.

It is clear that the MRF-polished lens is far less affected by spherical aberrations, as at both the larger and smaller input beam diameters, its spot size is incredibly small. The CNC-polished lens behaves far more ideally for the 9 mm input beam, showing it is affected by spherical aberrations significantly. Still, it never behaves as ideally as the MRF-polished lens.

The AL lens costs $318.27, while the ASL lens costs $364.67. Thus, we choose the AL lens as it performs better for our purposes and costs less money.

2.4. Conclusion

Using the Thorlabs lens catalog and the Zemax OpticStudio simulation software, we designed and simulated a fluorescence confocal laser scanning microscope for the purpose of *in vivo* cochlear imaging. We were able to achieve a sufficient lateral resolution over a much larger FOV than we require, however our axial resolution was made unfortunately high by our necessity for a large working distance. However, the axial resolution of the system designed, especially in the small pinhole case, is certainly good enough to make out the boundaries of the tectorial membrane as desired.

To integrate this system with the currently operating SD-OCT system, we would need to purchase the objective and collector, as well as a dichroic mirror, a point detector and a laser light source. The laser light source will likely dominate this cost. Unfortunately, no scanning lens in the Thorlabs catalog is rated for a wavelength range large enough to be used for both our confocal and OCT system. However, the changing of an objective is a quick process, and as the scanning mirrors need not be changed between the OCT and confocal system, the modalities can be rapidly swapped.

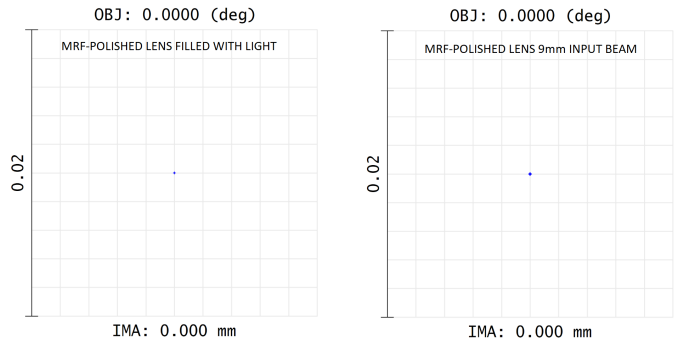

Fig. 7. Spot diagram for the AL MRF-polished lens, axis labels in microns. On the left, the lens is filled with light, i.e. we make use of the entire diameter of the lens. On the right, we use a 9 mm input beam diameter.

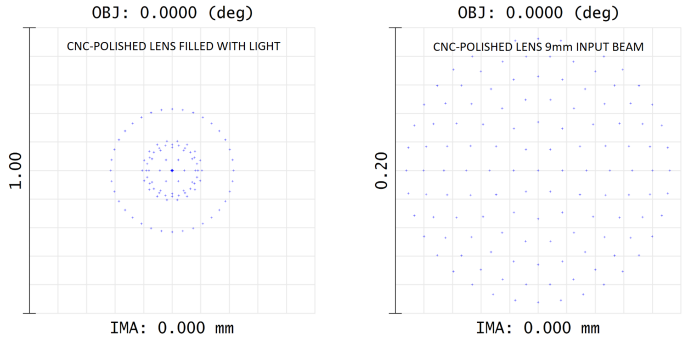

Fig. 8. Spot diagram for the ASL CNC-polished lens, axis labels in microns. On the left, the lens is filled with light, i.e. we make use of the entire diameter of the lens. On the right, we use a 9 mm input beam diameter.

To solve the problem of needing to swap, there exist lenses with greater wavelength ranges such as the Mitutoyo 378-823-5. Unfortunately, Mitutoyo lenses do not have Zemax models, and cannot be used for our simulations. However, it is possible that a Mitutoyo lens could perform as well or better than the SL50 for confocal imaging – this could be tested once the system is built, as we have a set of Mitutoyo objectives at Fowler lab.

References

1. C. E. Strimbu, S. Prasad, P. Hakizimana, and A. Fridberger, "Control of hearing sensitivity by tectorial membrane calcium," Proc. Natl. Acad. Sci. **116**, 5756–5764 (2019).
2. Thorlabs, Inc., 56 Sparta Avenue Newton, NJ 07860, *SD-OCT Base Unit Spectral Domain OCT System Base Units: Telesto, Ganymede, and Callisto Series User Manual* (2018).
3. Carl Zeiss MicroImaging GmbH, 07740 Jena, Germany, *LSM 710, LSM 780, LSM 710 NLO, LSM 780 NLO and ConfoCor 3 Operating Manual* (2010).
4. K. L. Hyrc, A. Minta, P. R. Escamilla, P. P. Chan, X. A. Meshik, and M. P. Goldberg, "Synthesis and properties of asante calcium red–a novel family of long excitation wavelength calcium indicators," Cell Calcium **54**, 320–333 (2013).