Characterizing Dust Distributions in Virgo Cleanrooms

University of Padova, University of Florida IREU 2021

Zach Yarbrough

12 August 2021

1 Introduction

Located near Pisa, Italy and hosted by the European Gravitational Observatory (EGO), Advanced Virgo is a gravitational wave detector that is actively contributing to the study of our universe. Advanced Virgo is a Michelson interferometer, as displayed in Figure 1, comprised of two perpendicular arms that are three kilometers long and home to laser beams that measure miniscule fluctuations in space time.

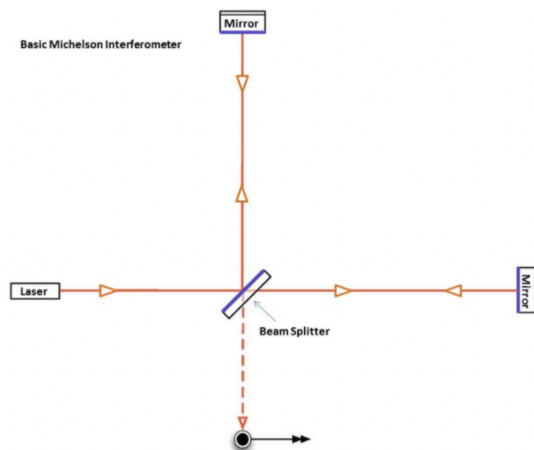

Figure 1: A simplified visualization of a Michelson interferometer, the same type that is used by Virgo [5].

These distortions caused by gravitational waves cause the length of Virgo's arms, and the light within, to change, thereby creating a measurable difference in phase. These differences, however, are minuscule; so small that footsteps several stories above the detection bench can negatively impact the data. This kind of contamination that can detrimentally affect measurements is commonly referred to as noise. There are several ongoing projects investigating noise reduction from sources like seismic shifts, changes in thermal properties that shift optical elements, and stray light, which we discuss in this paper.

1.1 Stray Light

Stray light is defined as unwanted light in an optical system. On some scales stray light can be nothing more than a minor inconvenience, while in other, more sensitive systems, a small amount of unwanted illumination can alter the outcome of many measurements. In the case of the Virgo interferometer, gravitational waves are measured with a sensitivity on the order of 10^{-12} photon wavelengths, so precise that even a few scattered photons can impact resulting measurements.

The phase of stray light differs from that of the light in the expected optical pathway and causes noise when recombining with the nominal beam. This phase shift is a result of scattering, a process by which photons may be removed ("scattered") from the beam before reentering with a different phase. Scattering occurs due to a number of factors, including motion of optical or mechanical elements due to seismic or human activity, intrinsic properties of elements within the nominal beam path, and surface contamination due to particulates. Of these sources of scattering, two of the most common are surface roughness of optical elements and particle surface contamination [1]. Reducing noise within the Virgo interferometer is of a very high priority, and although dust contamination is low in the cleanroom environments of Virgo, the sensitivity of detection makes the characterization of dust critically important.

1.2 Mie Theory

To understand and model the scattering of incident rays — and therefore analyze the effect dust has on scattered light — we utilize Mie scattering theory. Mie theory considers particles to be spherical and utilizes both forward and back scattering, resulting in modeling capabilities of particles approximately one photon wavelength in diameter. This particular modeling theory uses the particle density function $f(D)$ to predict the bidirectional scattering distribution function (BDSF). The particle density function describes the cleanliness of an optical surface and is given by

$$f(S, CL, D) = -\frac{d}{dD} N_p(S, CL, D) \tag{1}$$

where N_p is the number of particles per unit area whose diameters are bigger than or equal to D [2]. IEST-STD-CC1246D describes N_p as:

$$N_p(S, CL, D) = 10^{|S|[\log_{10}^2(CL) - \log_{10}^2(D)]} \tag{2}$$

where S is the slope of N_p versus D on a $\log - \log^2$ scale, CL is the surface cleanliness level, and D is particle diameter in μm. Lower values of S, also called the particle distribution slope, correspond to cleaner surfaces. IEST-STD-CC1246D defines an idealized slope value as -0.926, a value that indicates a "recently cleaned surface" [4]. This value, however, has not been experimentally observed, thus necessitating calculations of S for each specific experimental environment.

While Mie theory possesses great potential for raytracing simulations, it requires a wide knowledge of particle size distribution on the optical element in question, a measurement that can be difficult to ascertain with acceptable accuracy in an experimental setting. Thus, it was imperative that our imaging process be optimized and repeatable, capable of producing accurate values of particle properties that could be used in the powerful Mie models.

2 Imaging Process

As discussed in Duden [3], high quality images of each optical element must be obtained to accurately identify and characterize particles and their distribution. In order to do so, images were taken using a 10-megapixel camera with a 2/3” sensor, 35mm focal length, and accompanying lens at a distance of 110mm [6]. Samples were placed on a tilting table several centimeters high and illuminated by an LED light ring approximately 50cm above the optical element.

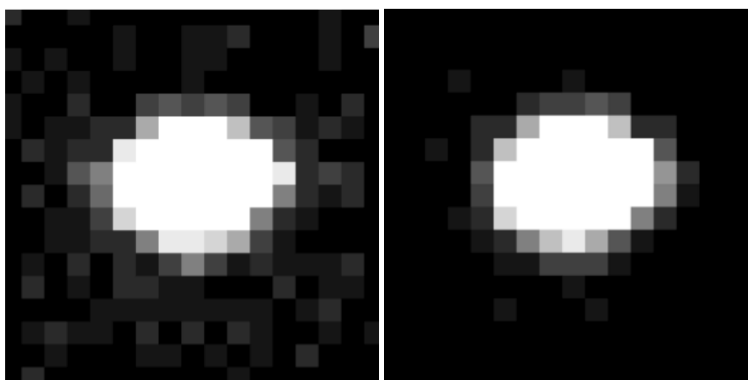

Figure 2: A single image of a particle (Left), and the same particle after combining 4 nominally identical images (Right)[3].

3 M4 O3 Analysis

The nature of the imaging setup made it impossible to capture the entirety of the wafer in one image, so each wafer was divided into three sections: top, middle, and bottom. Each section was then subsequently imaged. 10 images were taken of the same field, allowing us to import them as a stack into ImageJ and combine them using a z-projection. Combining multiple nominally identical images improves the analysis by reducing contributions due to background particles, as shown in Figure 2.

Random fluctuations in pixel luminosity may cause false counts; thus, combining multiple images reduces the impact of these variations and allows for increased accuracy in particle identification.

3.1 Processing

Once the optical elements, or wafers, had been imaged and nominally identical images combined using ImageJ, they were processed using MATLAB.
Following the code written by Waller [7] and referencing the list made by Duden

[3], the process for identifying particles and calculating their various characteristics is as follows:

1. Import image
2. Subtract background illumination
3. Convert image to grayscale
4. Convert grayscale image to binary
5. Remove particles above or below a chosen size from binary image
6. Print binary image
7. Outline, number, and label particles on binary image
8. Measure area of particles
9. Determine location of particles
10. Measure major and minor axes of ellipses drawn around particles
11. Output particle labels, areas, locations, and ellipse aspect ratios to csv file

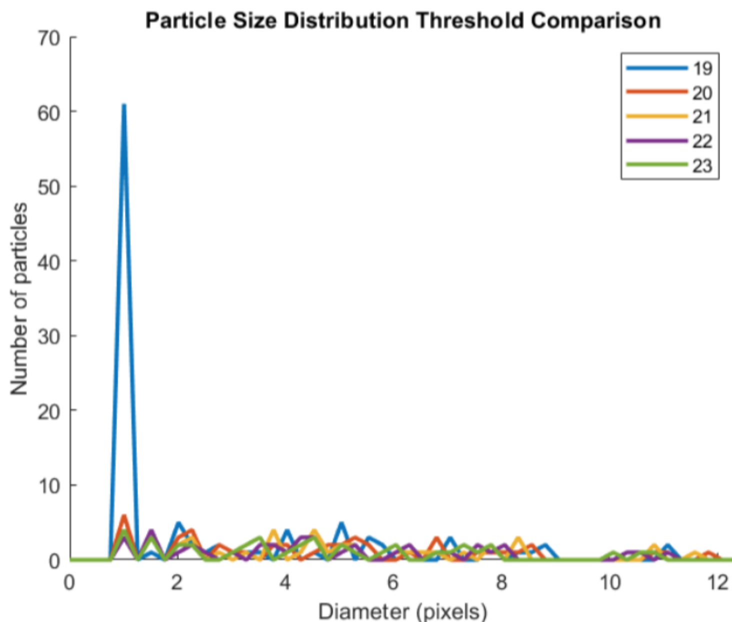

Figure 3: Threshold values 19-23, demonstrating a threshold value that is too low and therefore skewing data [1].

When binarizing an image, one must choose a threshold. Because the images were taken with a variety of apertures and exposure times, we analyzed the combined stacks with a wide range of thresholds to observe the different results and make observations that would eventually allow an "optimal" threshold to be chosen. Threshold values that are too low can allow background contamination to be counted as particles — therefore skewing the number of small particles to higher

values that are not representative of the actual distribution — and threshold values that are too high eliminate the less intense pixels around the edges of actual particles, possibly resulting in inaccurate diameter and area measurements. This is demonstrated in Figure 3, where threshold 19 is obviously too low as the data is skwewd towards particles with smaller diameters.

Ten nominally identical images were taken of each section of the the M4 O3 mirrors. Images were taken with both aperture 8 and 16 so as to determine the nominal aperture.

Initially, optimal thresholds were selected by collaborative observation and discussion. Once the threshold had been determined, our processing technique allowed us to export the diameters of the identified particles in micrometers. These diameters were then plotted against a linear fit according to Mie theory to determine several characteristic parameters: slope, cleanliness level, and R2 score.

4 M4 O3 Results

Figure 4 demonstrates the $\log(Particles/0.1\text{m}^2$ with Diameter $\geq D)$ as a function of $\log^2$(Particle Diameter D, μm). In order to determine optimal camera settings, the same portion of the M4 wafer was imaged with different settings. In this case, we analyzed the differences in apertures. The plots on the left were taken with a camera aperture of 16, and the plots on the right used an aperture of 8. As demonstrated by the increased number of points, aperture 8 allows in more light than aperture 16 — smaller aperture labels, like 8 in this case, correspond to larger shutter openings and therefore more light.

The advantage of a larger aperture and more light is that the particles can be resolved more accurately. Larger apertures yield higher resolution which is always beneficial when analyzing particles with diameters on the order of a few microns. The disadvantage, however, is that larger apertures make it easier for background contamination to impact the results. More light means more random fluctuations in pixel intensity and possible incorrect identification. As noted previously, background contamination appears in the analysis as disproportionately high amounts of small particles.

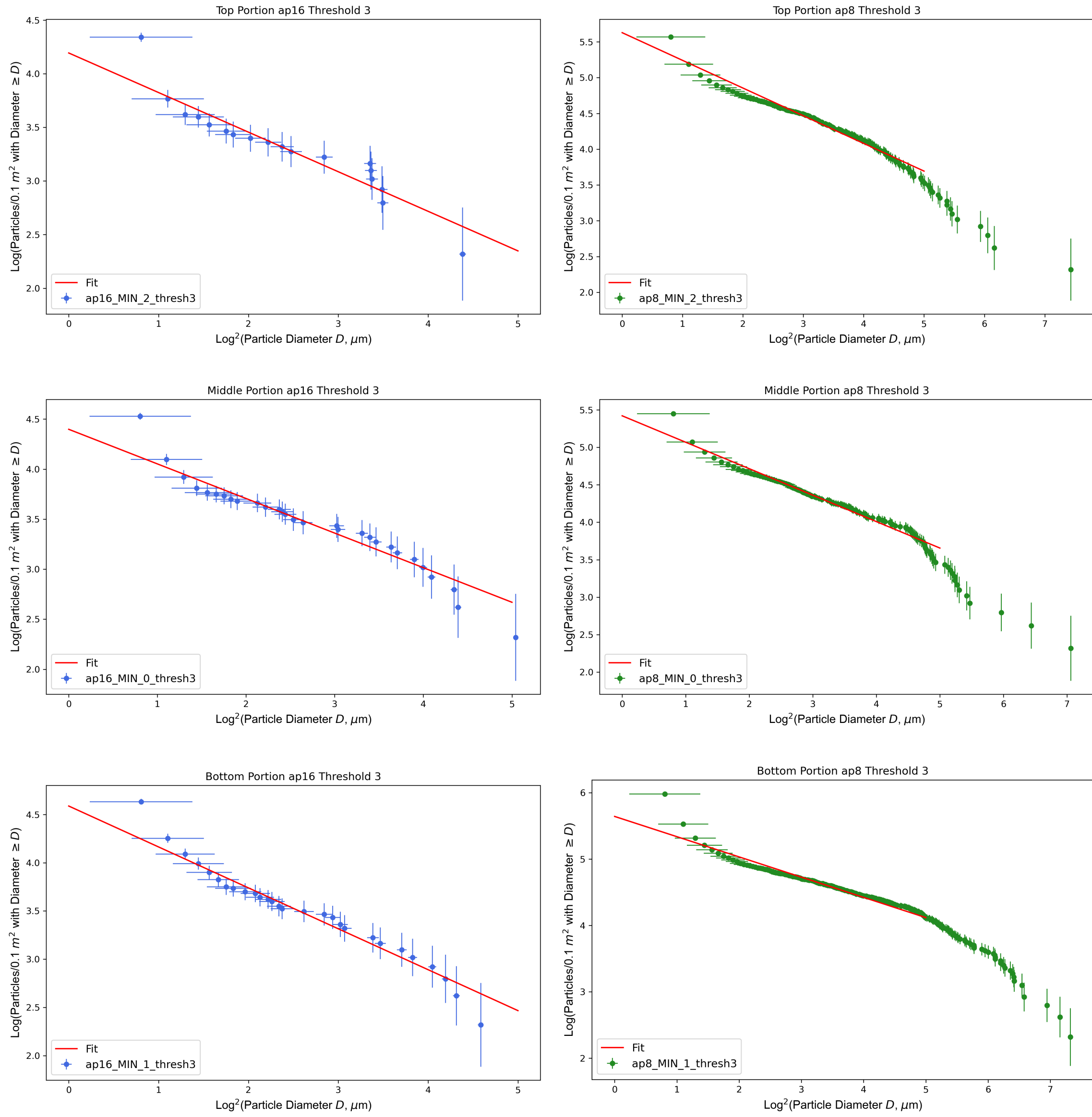

Figure 4: Particle distributions and respective fits for the top, middle, and bottom portions of the M4 optical element, imaged with aperture 16 (Left) and aperture 8(Right).

The table below gives the slopes and uncertainties for the fit functions corresponding to aperture 8 and 16. As mentioned previously, an ideal slope is described as approximately -0.926, a value which is not approached by any of the figures shown here.

Wafer	Slope	Uncertainty
Top portion aperture 16	-0.36899	±0.03458
Middle portion aperture 16	-0.34574	±0.01844
Bottom portion aperture 16	-0.42428	±0.018917
Top portion aperture 8	-0.38683	±0.00669
Middle portion aperture 8	-0.35288	±0.00636
Bottom portion aperture 8	-0.30472	±0.00443

Additionally, computed CL values differed from expected, if not reasonable, values by an order of magnitude. Interestingly, CL depends heavily on the largest particles captured in the distributions. Many of the plots reveal several large particles that are large enough to be seen by the naked eye, outliers. Because no such particle is evident in the visual analysis of the wafer, it is reasonable to consider the possibility that the analysis process has grouped together several particles of an average size and combined them in to one or more exceptionally large particles, heavily skewing the CL values. The difference in both S and CL values of the aperture 8 and 16 analysis reflects poorly on the imaging procedure to this point.

Another important factor to consider when determining optimal aperture is diffraction limiting. While the greatest resolution (approximately $5 \mu m$) is obtained using the greatest numerical aperture (smallest shutter opening), some apertures, like aperture 16, greatly impact the pixel resolution limit. Thus, larger apertures including both aperture 16 and 8 could be considered a poor choice for the imaging process in question.

4.1 Conclusions

Having arrived at the conclusion that apertures 8 and 16 were unsuitable for our purposes, it became clear that there were many changes to be made to the imaging process. Specifically, experimental results demonstrated that pursuing high resolution, and therefore using the greatest numerical aperture (smallest shutter opening), quickly becomes unsustainable as resolution deteriorates due to diffraction limitation.

In order to develop an imaging procedure with the most accuracy — which would consequently yield the most meaningful parameters for simulations — aperture selection and dependency needed to be better understood.

4.2 Revisions

Following our conclusions, it became clear that new apertures were needed for the imaging process. To compare aperture candidates, we used specialized objects to observe differences between apertures. Produced in the Virgo machine shop and displayed in Figure 5, our specialized objects were staircases that had step intervals comparable to the diameter of particles.

We imaged these steps using all of the apertures possible on our camera — including apertures 8 and 16 to confirm previous conclusions. We then displayed the images in large format and used digital zoom to investigate further, aiming to count the number of steps that each aperture captured in focus.

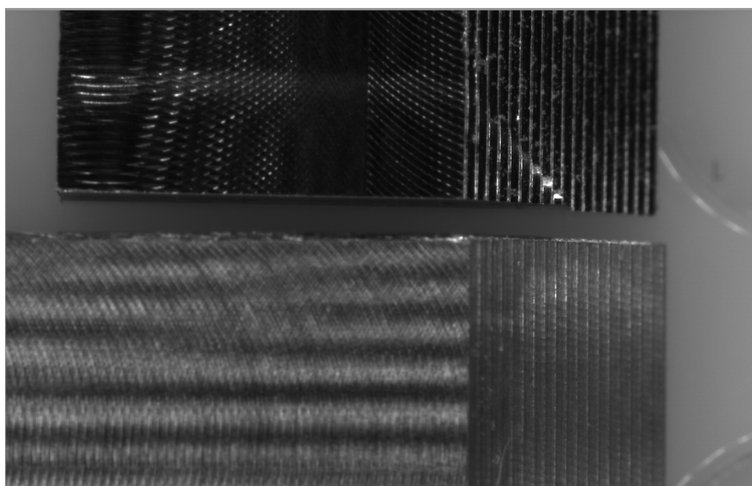

Figure 5: Custom staircase imaging elements used to determine best aperture candidates.

This process was difficult and resulted in a great deal of discussion, as our only way to identify one aperture as being better than another was by visual analysis. Knowing that the physical properties, and therefore quantifiable metrics, of the particles captured by our chosen apertures would soon be evaluated, the choice itself was not critically important. Eventually, apertures 2,8 and 5.6 were chosen for further analysis because they produced the most steps in focus.

5 Calibrated Dust

After selecting the two best apertures for the imaging process, we proceeded to the next portion of the process: analyzing calibrated dust particles. To this point, there has been no way to check the accuracy or repeatability of our imaging setup and analysis. To explore this further, we imaged three different kinds of particles, Molybdenum, Titanium, and Aluminum, all of which were made up of a known elemental

combination and had well documented statistics. The known values for these powders include mean particle diameter, mean particle diameter of the 90th percentile of particles, and the mean diameter of the 10th percentile, all in micrometers.

Knowing these characteristic measurements provides a metric for the evaluation of our camera setup and processing procedure. If the analysis can identify and measure the particles effectively, analyzing a distribution of particles with a known mean diameter should return that same value.

5.1 Calibrated Dust Analysis

Particles of Molybdenum, Titanium, and Aluminum were deliberately scattered on pristine mirrors. These wafers were then imaged using apertures 2.8 and 5.6. Similar to the analysis carried out on the M4 O3 wafer, ten images of the same field were taken and then combined using a z-projection. In this case it was determined that imaging the entirety of the wafer was not necessary, so a discreet portion of each wafer was carefully selected.

Combined stacks were then processed using the same MATLAB code and procedure used on the M4 O3 wafer. Notably, even though the detection of particles and calculation of properties remained consistent, one change was made to the resulting plots. Instead of plotting number of particles versus diameter, number of particles was plotted against area of particles in pixels as shown in Figure 6. Because pixels are discretized, the fainter pixels on the fringes of particle boundaries have a chance to be misidentified. Plotting as a function of area reduced ambiguity by eliminating contributions from possible misattributions.

Once the changes had been made to avoid using diffraction-limited apertures, our focus shifted to optimal threshold values. Specifically, if we could establish a method for determine optimal threshold values for each set of parameters, we could then compare the threshold values to determine how much the resultant distributions depended on camera settings. In other words, changes in optimal threshold values will allow us to understand if, and by how much, camera settings impact the statistics and results of identical wafer regions and their distributions.

5.2 Optimal Threshold Determination

Before optimal thresholds could be compared, a metric had to be established to describe what makes a threshold optimal. To accomplish this, we used the known statistics of each particle type. As shown in Figure 7, the processing code easily outputs a number of values, including mean particle diameter, diameter of the 90th percentile, and the number of particles. Using this information, we developed the following process to determine the optimal threshold:

Find a range of thresholds that result in a mean diameter value ± 5 the known value. For example, the mean diameter of an Aluminum particle in micrometers

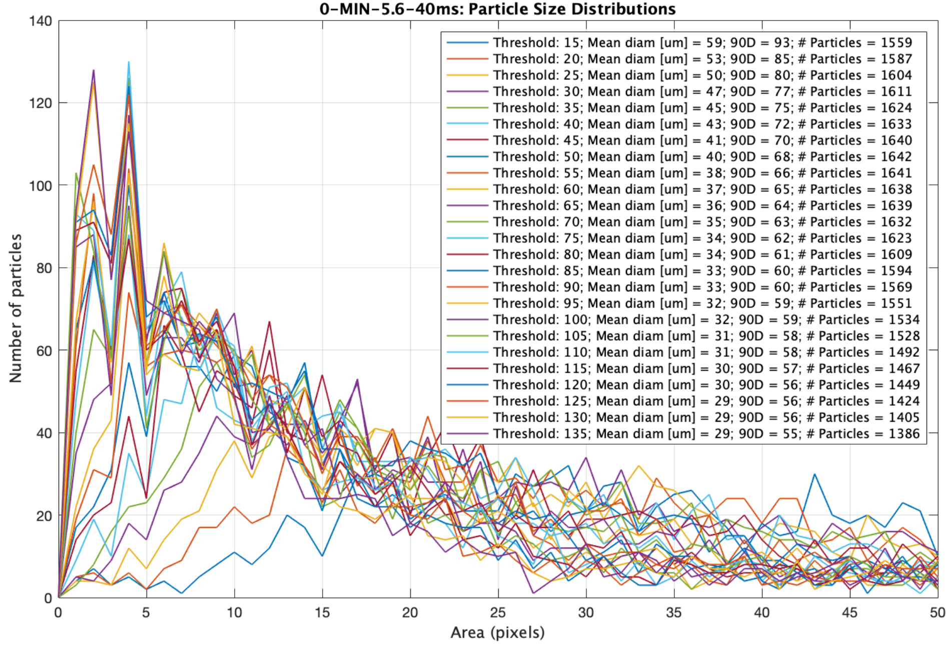

Figure 6: A plot demonstrating number of particles as a function of particle area (pixels). Representing aperture 5.6, exposure time 40 milliseconds.

is known to be $44.7\mu m$, so we selected a range of thresholds that resulted in mean diameter values from 40 to 50 after rounding. This is demonstrated in Figure 7, the range of $45\mu m \pm 5$ is made up of approximately 20 threshold values, ranging from 24 to 48. Creating this window of threshold values allows us to confidently say that the most optimal threshold value lies somewhere within these bounds.

Next, we repeated these steps using the known value for the 90th percentile of particle diameters. This characteristic tells us that 90% of threshold diameters fall below this value, and in the case of Aluminum that value is $69\mu m$. Rounding to $70\mu m$, we searched for values that produce a range of $\pm 5\mu m$ so as to confidently say the optimal threshold for the 90th percentile diameter value is within those bounds. As demonstrated in Figure 8, that threshold range happens to be 36-62.

Now in possession of threshold ranges that we know to contain the optimal threshold for producing mean diameter values and 90th percentile diameter values respectively, we can calculate the residual differences between observed and expected values. Equations 3 and 4 give us the residual value for each threshold when computed.

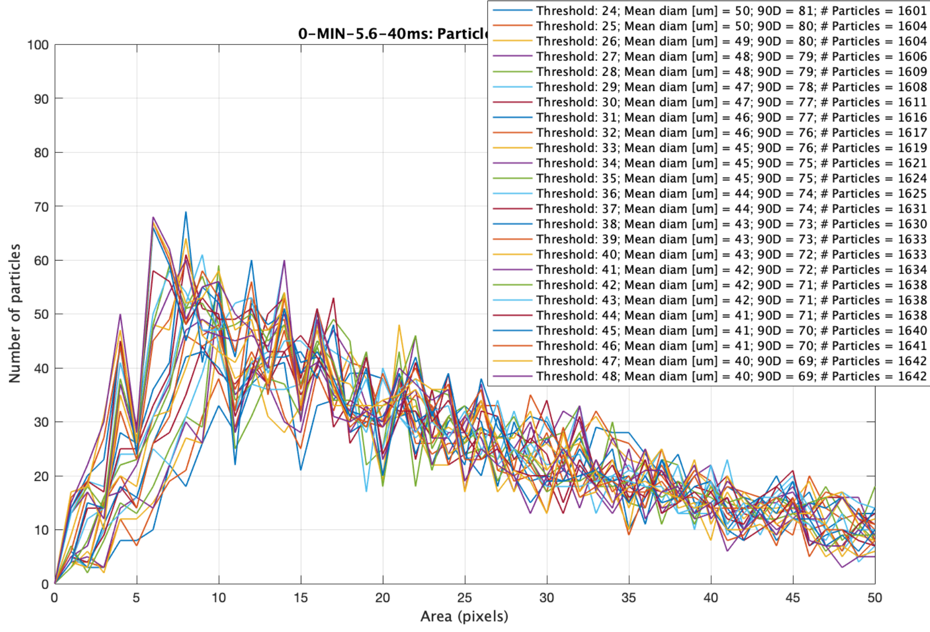

Figure 7: Threshold values 24-48 applied to aperture 5.6 exposure time 40 milliseconds, resulting in a mean threshold range of 65 $\pm 5 \mu m$.

$$\frac{(\text{Measured mean } D \text{ value - expected mean } D \text{ value})}{\text{expected mean } D \text{ value}} \tag{3}$$

$$\frac{(\text{Measured mean } 90D \text{ value - expected mean } 90D \text{ value})}{\text{expected mean } 90D \text{ value}} \tag{4}$$

We can then plot equations 3 and 4 against one another to visualize the effectiveness of each threshold. We know that the "ideal" threshold lies within our range of values — one of these thresholds results in the most accurate statistics for the calibrated particles. Equations 3 and 4 also tell us that the best thresholds will return values close to zero — if the difference between the measured and expected value is zero then the measured value is accurate and the threshold that was used to produce that value is optimal. Thus, we can intuitively say that, once the results of equations 3 and 4 are plotted against one another, the best threshold value will be closest to the origin. As Figure 9 demonstrates, there are clearly certain thresholds that offer better statistical return than others. In this case, the threshold closest to the origin, and therefore the optimal threshold for aperture 5.6 exposure time 40 milliseconds is 30.

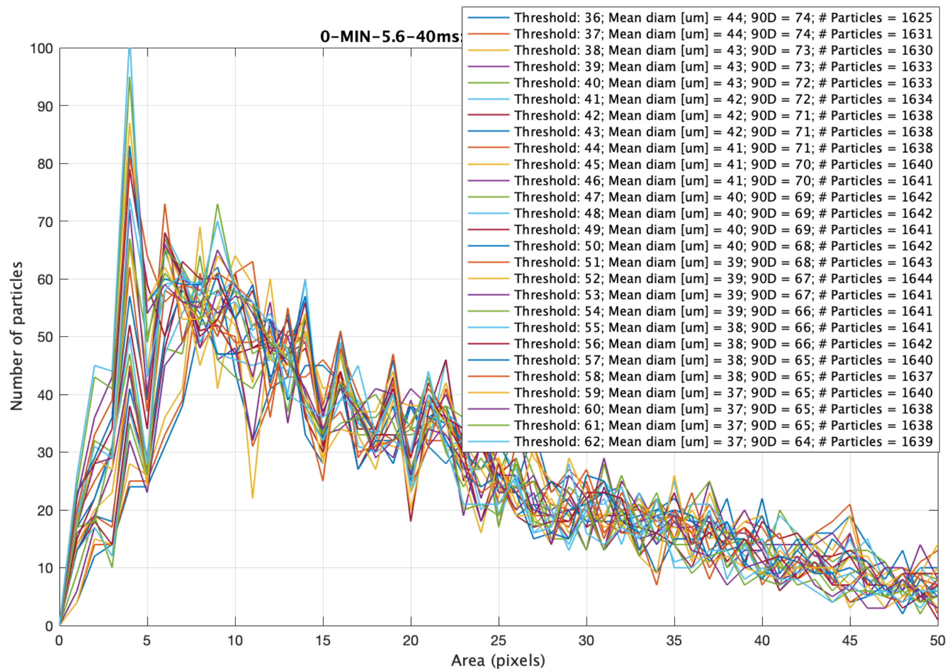

Figure 8: Threshold values 36 to 62 applied to aperture 5.6 exposure time 40 milliseconds, resulting in a 90th percentile threshold range of 70 $\pm 5 \mu m$.

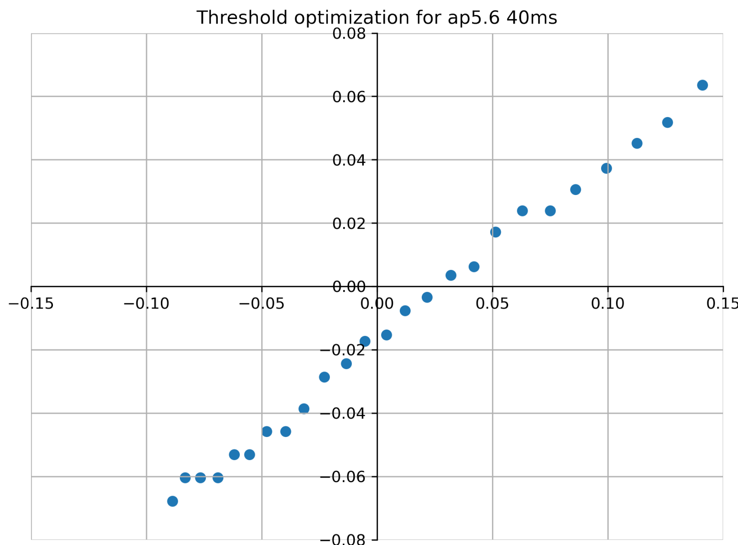

Figure 9: Calculated residuals for all threshold values. The optimal threshold for this particular image is the value closest to the origin.

5.3 Calibrated Thresholds Results and Conclusions

This same analysis was carried out for images that had different camera settings, including aperture 5.6, exposure time 12ms, aperture 2.8, exposure time 3ms, and aperture 2.8, exposure time 10ms. Figure 10 displays the optimal threshold values for each group of settings discussed above.

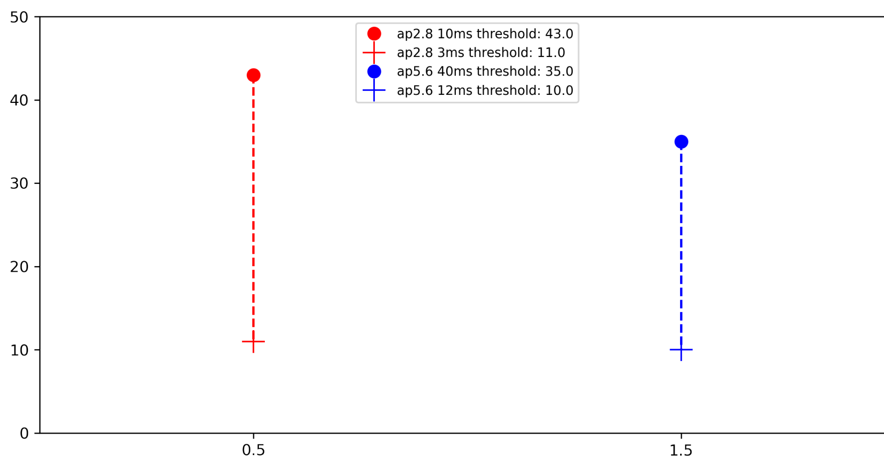

Figure 10: Optimal thresholds for all imaging parameters (x axis values are arbitrary).

This figure is interesting for a number of reasons. More significant than the threshold values themselves are the differences between optimal thresholds for each group of settings. Clearly, there is a greater difference between the optimal threshold values of the images taken with aperture 2.8 (depicted in red) and those taken with aperture 5.6 (depicted in blue). The difference is small but potentially significant, indicating possible volatility inherent to aperture 2.8. The only change between the two sets of images taken with aperture 2.8 was the exposure time, a controlled variable that did not involve touching the imaging setup or the wafer. Bearing this in mind, the possibility for an even larger difference between optimal threshold values increases when other, unforeseen factors are taken into consideration. In other words, there is a larger difference in observed optimal thresholds between apertures 2.8 and 5.6 under optimal conditions. While this is insignificant in lab conditions, developing a repeatable procedure to be applied anywhere within the Virgo optical scheme requires minimization of uncertainty, and, in the case of aperture 2.8, volatility. Thus, these results point to aperture 5.8 being a more stable, viable choice for continued analysis.

An additional interesting point is that the threshold values of the lower exposure times are very similar. Recalling that images taken with an aperture of 2.8 and exposure time of 3 milliseconds have been exposed to the same amount of light as images

taken with aperture 5.8 and exposure time 12 milliseconds, these values for optimal thresholds indicate similarities between the two images which are not observed at higher exposure time (and therefore larger amounts of light). This observation presents the option to expose particle fields for less time, therefore subjecting them to less light, in order to reduce ambiguity within the imaging procedure. That is to say that

To explore this question and optimal threshold determination technique further, these steps were repeated for the same groups of settings, aperture 2.8 exposure times 3 milliseconds and 10 milliseconds, and aperture 5.6 exposure times 12 milliseconds and 40 milliseconds.

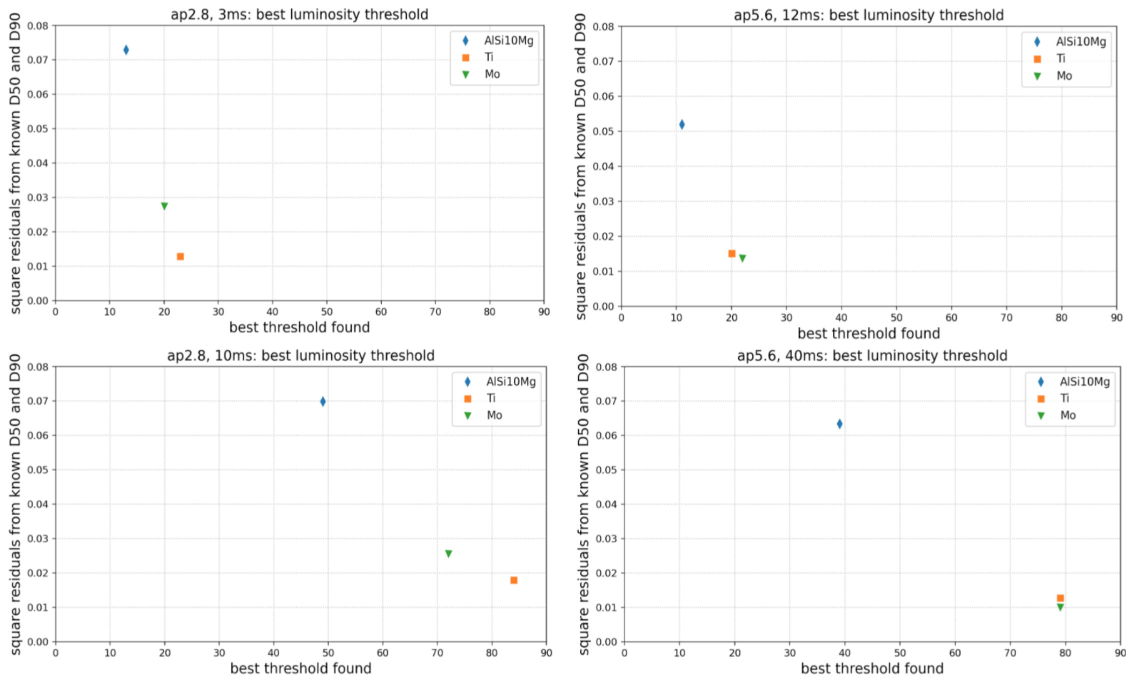

Figure 11: Optimal threshold values for each element captured with each set of parameters.

Figure 11 shows the best threshold value for each element taken with each group of settings. The instability of aperture 2.8 is demonstrated once again by comparing the best threshold values for Titanium and Molybdenum. In this case, threshold values for Aluminum are not as significant as we do not expect particles of that size to appear in Virgo cleanrooms. The images taken with aperture 5.6 show very similar threshold values for the two different particles, whereas images taken with aperture 2.8 are much farther apart. This is especially interesting when we consider the differences in diameter between Titanium and Molybdenum. As previously mentioned, Molybdenum particles are much smaller than Titanium, so it is important to note the ability of aperture 5.6 to produce very similar — if not identical — opti-

mal threshold values for two particles with much different sizes; this is a significant observation. In the future, we hope to implement this imaging procedure on a large scale with the goal of understanding dust distributions in any given environment. Within said environments, there is no way to know the average diameter of particles as we currently do with calibrated ones. Thus, the similarities between threshold values even when the particles are vastly different in size indicate that these specific settings, aperture 5.6 exposure times 12 and 40 milliseconds, are best suited to producing accurate results.

6 Results and Conclusions

As a whole, this project helped us to understand our imaging process on a deeper level. Work by previous investigators encouraged us to image and analyze the M4 O3 mirrors. The models produced thereafter resulted in values for slope and cleanliness level that were not representative of expected values, indicating a possible weakness in the imaging process. Moreover, diffraction limitations eliminated the expected benefits of apertures 8 and 16. Considering these two facts prompted a deeper investigation into the imaging process, specifically aperture optimization.

After choosing apertures 2.8 and 5.6 for further consideration, we began working with calibrated particles in order to determine the effectiveness of our procedure.

Developing our own metric for identifying optimal thresholds created a common metric shared by all images that we could then use to compare the newly selected apertures. Visually representing these results in unique ways provided insight into the differences between apertures 2.8 and 5.6, specifically a possible volatility and uncertainty possessed by aperture 2.8 that made 5.6 the best candidate for obtaining accurate results. Future work will inevitably investigate apertures 2.8 and 5.6 in more detail, observing the selected thresholds' reaction to wafer tilt, variable particle diameter, and eventually genuine dust from Virgo optical elements.

In summation, the development of a repeatable, accurate imaging and analysis procedure has proved to be no small task, and the investigations carried out in this paper are invaluable steps towards that eventual conclusion.

7 Acknowledgements

I would like to extend many thanks to my mentors Dr. Livia Conti and Dr. Giacomo Ciani for their relentless support and investment in my full potendial. I also owe a great deal of thanks to my talented colleague and friend Beatrice D'Angelo for being an invaluable collaborator as well as a constant source of encouragement. Thank you to the entire team at the University of Florida for extending this opportunity for collaboration and scientific exploration. Finally, thank you to the National

Science Foundation for supporting this program and numerous projects through grants PHY-1950830 and PHY-1460803.

References

[1] B. D'Angelo *Dust Analyses. 2021.*

[2] E. Fest *Stray light analysis and control. PM229:228, 2013.*

[3] E. Duden *Final Report. 2020.*

[4] IEST-STD-CC1246D *Product cleanliness levels and contamination control program, Institute of Environmental Sciences and Technology (2002).*

[5] LIGO Laboratory, Caltech. n.d. *LIGO's Interferometer.*

[6] F. Sorrentino *"Optical system for dust contamination monitoring via Si wafer inspection (CRQ 2017/008)," Virgo internal document VIR-0463A-17 (2017).*

[7] S. Waller *Process for aLIGO Particle Imaging Using a Nikon D7100 and Matlab. Jul 2015. LIGO-T1400515-v9.*