System level design trades for the Submillimeter Probe of the Evolution of Cosmic Structure (SPECS)

Raymond J. Sedwick* and David W. Miller
MIT Space Systems Laboratory

ABSTRACT

The Submillimeter Probe of the Evolution of Cosmic Structure (SPECS) is a proposed mission being considered by NASA to take high resolution sky images at the submillimeter wavelengths. At these wavelengths, we can observe highly red-shifted light that has been traveling towards the Earth over a great deal of time, allowing observation of the developing universe in the distant past. To achieve Hubble class resolution, the system must synthesize an aperture that is 1000 meters in diameter, making interferometry the only currently reasonable approach to design. This paper addresses a high level system trade of a general class of rotating tethered architectures that reel apertures in and out radially to fill the u-v plane. A metric of cost per image is used to identify how aperture area and number of apertures should be selected, and considerations of how spectral information is acquired is used to generate design ideas for reducing the overall system mass.

Keywords: System Trades, Space Telescopes, SPECS, Interferometry

1. INTRODUCTION

At its highest level, SPECS is a Michelson Imaging Fourier Transform Spectrometer intended to achieve Hubble class spatial resolution and high spectral resolution over a broad spectrum in the submillimeter wavelengths. The most immediately derivable requirement is that to achieve this resolution at such long wavelengths, the telescope must be able to synthesize an aperture that is roughly 1000 meters in diameter. Following the lead of Farley and Quinn¹, we will assume that a spinning tethered architecture is most suitable for this particular mission. Our entry point into this study is then to assume an architecture which performs interferometric imaging using some number of tethered apertures, pairwise interfered to fill in the Fourier u-v imaging plane.

As a performance metric, we will assert that a good architecture has a low cost per image taken over its operational lifetime. In addition, we will assume that there is some minimum requirement on the number of images that must be taken, and some maximum budget cap constraint on the project. If we are lucky, there will be at least one architecture in the trade space that can meet both of these constraints. Otherwise, whatever trade space was defined must be broadened to include other technologies, or ultimately the constraints must be relaxed. In our analysis, we will attempt to stay as general as possible, for as long as possible, to avoid introducing design decisions on an ad hoc basis.

2. THE PERFORMANCE METRIC

The first step is then to formulate this performance metric in terms of system parameters that can be quantified. First, the cost may be expressed in many different ways, depending on what information is available, and what parameters are being considered. Again in trying to stay general, we want to start with a simple but comprehensive formulation, and we will assume that there are three primary terms in the cost. The first is a constant term, which would capture elements of the design that do not scale strongly with the size, number or lifetime of the spacecraft. The second term is components that do indeed scale with the number of spacecraft, which would include launch, fabrication and possibly some elements of the design and operations. The last term is primarily operations, and we will assume that it scales with the lifetime of the spacecraft, but is not strongly related to either number of spacecraft or other design details. For a large number \( n \) of spacecraft, this assumption would probably not be valid, since the complexity of operations or cost of autonomy development could begin to be affected. As we will show later, few spacecraft are actually preferred, so that this assumption is more valid.

* sedwick@mit.edu; Ph: 617-253-7831; Fx: 617-258-5940; http://ssl.mit.edu/; MIT Space Systems Laboratory, 77 Massachusetts Ave, Rm 37-431, Cambridge, MA 02139
The first cut at the lifecycle cost of the system is then

\[ \text{Cost} = C_0 + C_n n + C_T T_{\text{life}} \]  \hspace{1cm} (2.1)

We then take the cost that scales with the number of spacecraft and further subdivide it into three parts. The first is a constant, which would account for launch cost of the bus, as well as possibly aspects of the design costs that were driven by the number of apertures. The second is proportional to the aperture area, which would include launch cost of both the aperture and other components that would scale accordingly. The last is proportional to \( A^{1.58} \), which is a curve fit that has been shown by Meinel to accurately capture the development cost of very large optical telescopes.\(^2\) The scaling has also been found to apply to small amateur and professional telescopes less than a meter in diameter, although the coefficient is not the same. It is not obvious that this relationship can be applied to submillimeter wavelength systems, but it is also not likely that the cost of these apertures would scale at a higher power than optical telescopes, and for our purposes this represents a useful upper boundary. The final form of the cost is then given by

\[ \text{Cost} = C_0 + (C_1 + C_2 A + C_3 A^{1.58}) n + C_4 T_{\text{life}} \]  \hspace{1cm} (2.2)

Clearly there can be components of the system that do not fall into one of these scaling laws. For instance, re-targeting requires propellant, the amount of which scales with both the number of times a maneuver is performed, as well as the number of spacecraft. To capture this would require adding a term that scaled both with \( n \) and \( T_{\text{life}} \). However, this term does not affect the results that follow. Contributions of other components as we get into a higher level of detail will be less significant as compared to the terms already captured, and would also not have a large effect. As another example, there may be a component that can be identified as scaling with the square of the aperture area, but the assumption is that the contribution of this term is negligible as compared to those already present, for reasonable aperture sizes. This contribution of the cost can always be estimated by comparing it with any terms above that can be easily evaluated under reasonable assumptions. Also, with the presence of the zero\(^{\text{th}}\) and first order dependence on \( A \), we can always add a linear approximation to any power law around some nominal design point. Given the different contributions for the system, determining the best values for the \( C \) parameters is still a difficult task, and would most likely be met with opposition regardless of the choice. The hope is instead that one can be convinced that given a particular set of technologies, the cost scaling can be reasonably captured with the correct constants inserted in the above equation.

The other component of the performance metric is the number of images that can be acquired over the lifetime of the system. For this estimate, we will assume that the system is always in a fully operational state, operating continuously, and either imaging or re-targeting. The amount of time that the system spends imaging is normalized by the total time to image and re-target, and this ratio will be called the duty cycle (\( DC \) ) of the system. The total number of images is then given by

\[ \text{Number of images} = \frac{T_{\text{life}}}{T_{\text{image}} + T_{\text{target}}} = DC \left( \frac{T_{\text{life}}}{T_{\text{image}}} \right) \]  \hspace{1cm} (2.3)

The operational lifetime \( (T_{\text{life}}) \) of the system is a specified requirement, and the imaging time \( (T_{\text{image}}) \) is determined by the signal to noise (\( SNR \)) requirements of the system. Before evaluating the imaging time, it is worth noting that a more rigorous calculation of the utility (in this case number of images) can be made in light of component failures through the process of Markov modeling of the system states\(^3\). However, the accuracy of this approach is dependent on the ability to assign realistic values to component failure rates, and unless the component is mass-produced and tested, a statistically significant value of this rate cannot be obtained. Because many of the components are very cutting edge technology, an accurate description of their statistical behavior would be difficult to determine. At best, we can try to make qualitative arguments of design decisions that will lead to a more robust system.

Returning then to evaluating the time to take an image, the background limited \( SNR \) for a system like SPECS imaging a point source is given by

\[ SNR \sim A \sqrt{n T_{\text{image}}} \left( 1 + \frac{FOV}{2 R_e \Delta \theta} \right)^{-1/2} \]  \hspace{1cm} (2.4)
where $n$ is the number of apertures, $\Delta \theta$ is the spatial resolution, $B$ is the maximum baseline, $R_d$ is the desired spectral resolution and $FOV$ is the desired field of view. The source of these terms will be explained in section 3. All of these parameters except for aperture size and number of apertures are assumed specified by requirements. The proportionality depends on the signal strength being measured, the background intensity and the efficiency of the instruments, and overall is not a parameter that can be controlled. We therefore simplify the relationship to

\[
SNR = \beta A \sqrt{n T_{\text{image}}}
\]  

(2.5)

where $\beta$ is just a constant of proportionality. We can now rewrite the performance metric as

\[
\text{Cost} = DC \left( \frac{SNR}{\beta} \right)^3 \left[ \frac{C_0 + (C_1 + C_2 A + C_3 A^{1.35}) n}{n A^2 T_{\text{life}}} + \frac{C_4}{n A^2} \right]
\]

(2.6)

Because the only free parameters are $n$ and $A$, we can deduce some interesting design issues from this expression. For one, because the $A^2$ in the denominator dominates all powers of $A$ in the numerator, the minimum cost per image design will have the largest possible aperture. Clearly if there was a component that scaled with aperture area to an exponent greater than two (as discussed earlier), then there would be some aperture size above which the cost per image again begins to increase. However, because the cost coefficient of such a component is assumed to be small, the area at which that would happen would be unrealistically large. The other point where this scaling can break down is if the aperture goes beyond a size where it can be packed the next larger launch fairing. At this point, some on-orbit assembly or the development of a new launch vehicle would cause a huge penalty to the laws given above and invalidate their use. There is therefore a maximum area that can be assumed if the above relationship is to be used.

Since the performance metric improves as both the aperture size and number are increased, we need to look at the problem another way. Instead, let’s consider the cost variation of the system under the assumption of a fixed number of images. Fixing the number of images is essentially fixing the image time, or setting $A^2 n$ equal to a constant we will define as $k$. Solving for $n$, and substituting into the expression for cost results in

\[
\text{Cost} = C_0 + \left( \frac{C_1 + C_2 A + C_3 A^{1.35}}{A^2} \right) k + C_4 T_{\text{life}}
\]

(2.7)

So, similarly to the performance metric, the aperture size dominates the cost scaling and indicates that it’s better to go with the largest aperture. Under the assumption of a constant number of images, this also says that we should use the fewest number of apertures possible. To perform the large baseline interferometry, the minimum number of apertures needed is two. However, to combine the light at equal path lengths from each aperture, a third spacecraft is needed halfway between. We will discuss later the benefits of making this a third aperture and changing the geometry to a triangle, as is currently planned. Given a particular choice for $A$ and $n$, one can always guarantee a minimum number of images, if enough lifetime can be achieved. Under this condition, cost per image decreases, asymptotically reaching a minimum value, whereas the actual cost of the system continues to climb.

One point worth noting is the importance of the scaling law for $SNR$. If photon shot noise was instead identified as the dominant source of noise, then the aperture area that appears in the $SNR$ would move under the radical. The result of this seemingly innocuous change is that the $A^2$ that appears in the denominator of equations 2.6 and 2.7 is reduced to simply $A$. Now the cost per image, and the cost assuming constant number of images both have aperture area appearing in the numerator and denominator, implying that an optimum aperture size exists. The optimum for each is different, and as an example aperture size that minimizes the lifecycle cost for a fixed number of images is given by

\[
A = \left( \frac{C_1}{0.35 C_3} \right)^{0.34}
\]

(2.8)
As one would expect, the more expensive the aperture development cost is in relation to the rest of the bus cost, the smaller the optimum value of the aperture would become. Interestingly, the aperture launch cost does not affect the result, since it too scales with aperture area.

3. IMAGING

Having looked at the highest level design issues, more detail on the actual imaging process is necessary to draw further conclusions. As the name implies, the Michelson Fourier Transform Imaging Spectrometer combines several technologies to achieve its science. A Michelson interferometer takes the same light over two paths, and interferes it at a point on a focal plane. As the relative path difference is varied, the intensity on the detector will fluctuate, as different wavelengths of light pass in and out of phase. If this intensity is sampled as a function of path difference, the resulting data can be Fourier transformed to arrive at the spectral content.

There are then two aspects to its imaging nature. First is that this light is brought in from two apertures, separated by some distance, and looking along the same line of sight. Given a particular wavelength of light, the separation distance causes these apertures to be tuned to a specific spatial variation of this light across the sky. Sampling over many of these spatial frequencies, a Fourier transform can again be used to produce a brightness map of the sky at this particular wavelength of light. Because changing the path difference produced spectral information at each spatial frequency, a brightness map for each wavelength of light sampled can be generated. The other imaging aspect is that instead of having a single detector, with a field of view of $\lambda/D$, the science light is focused onto an $N \times N$ array of detector pixels, extending the FOV out to $N\lambda/D$.

Each interferogram produced while sampling would therefore appear as an image on the array (if it were visible) at a resolution $\lambda/D$, as if looking through one of the apertures.

To invert the spatial frequency content into a brightness map, a sufficient number of samples in the u-v plane must be made, where sufficient is not a very well defined parameter. Algorithms such as CLEAN and Maximum Entropy have been developed to fill in the missing data in sparsely sampled images, with good success. For our analysis, we will assume complete u-v filling, with the understanding that the results can be scaled to more sparse sampling through modification of the duty cycle defined earlier.

To determine how the u-v plane gets filled, we use the fact that the u-v coverage is the autocorrelation of the aperture coverage. The autocorrelation is given mathematically as

$$\gamma(u,v) = \int_{-\infty}^{\infty} f \left( \frac{x}{\lambda} \right) \cdot f \left( \frac{u-x}{\lambda}, \frac{v-y}{\lambda} \right) \ du \ dv$$

but there is also a very useful physical analogy. In Figure 1, we consider a distribution of apertures in a plane, centered at the origin, and make a copy of the distribution which can be offset at any point (x, y) in the plane. In this case we have selected three apertures at the corners of an equilateral triangle. Placing the copy at each point, we integrate over the total amount of overlapping area between the two copies, and record the total at the corresponding point (u, v) in the u-v plane. When all the points have been recorded, the resulting map is the u-v coverage for the aperture distribution. The movable copy of the distribution is shown slightly reduced in size (just for clarity), and in the six regions of the plane where non-zero overlap will occur. Since non-zero overlap occurs any time two of the apertures (one original, one copy) are within one diameter of each other, the resulting coverage regions are of diameter 2D. In this case, we see that the three apertures result in seven spots in the u-v plane. In general, n apertures in physical space will map into $n(n-1)$ spots away from the origin, and one at the origin, corresponding to the dc component of spatial frequency. A configuration of only two apertures at (x, 0) and (-x, 0) would result in u-v coverage resembling spots 3, 4 and 5 in the figure.

SPECS fills the u-v plane by rotating the configuration around the line of sight, and reeling the apertures in and out. As the configuration rotates, each spot in the u-v plane traces out a spiral path, and with every spot diameter (2D) that it travels, a new u-v point is covered. Full u-v coverage is obtained when the u-v plane spots move in or out just far enough during rotation that no overlap occurs. For the three aperture case shown above, the apertures must be reeled out one diameter every 60 degrees. This is not obvious in the physical domain, which appears to require that an aperture be reeled out one diameter every 120 degrees, resulting in double coverage. Since it is the baseline vector between apertures that matters, it can be seen
in the physical domain after rotating 60 degrees that the baseline appearing horizontally at the top of the triangle is the same as the baseline that originally appeared horizontally at the bottom of the triangle.

For configurations where the \( n \) apertures are evenly distributed around a perimeter, the reel rate that achieves complete u-v filling is found from

\[
\frac{\dot{r}}{\theta} = \frac{D}{\left[ T / n(n-1) \right]} = \frac{n(n-1)D}{2\pi} \theta = \frac{n(n-1)D}{2\pi} \left( \frac{v}{r} \right)
\]

where \( T \) is the period of rotation at a given radius. Given a velocity profile as a function of radius, this equation can be integrated to determine the time required to fill the u-v plane out to the desired baseline. Instead we recognize that the spot area of the filled aperture in the u-v plane is \( A_{aperture} = \pi(2R)^2 = \pi B^2 \) and the spot area filled by a given configuration of apertures is \( A_{\text{conf}} = n(n-1)\pi D^2 \), not counting the central spot, since it is common to all configurations. Assuming the same time is spent at each configuration, the time available to spend is then given as

\[
T_{\text{conf}} = n(n-1) \left( \frac{D}{B} \right)^2 \frac{T_{\text{image}}}{\text{SNR}} = \left( \frac{4}{\pi} \right) \left( \frac{\text{SNR}}{\beta B} \right)^2 \frac{(n-1)}{A}
\]

where the expression for \( \text{SNR} \) given earlier has been substituted in for \( T_{\text{image}} \). Depending on physical limitations of the delay line design, it is possible that the time available based on \( \text{SNR} \) is insufficient for the delay line to fully stroke. On the other hand, perhaps the time needed for a single configuration is short enough that the resulting \( \text{SNR} \) would not be high enough. Ideally the delay line could be engineered not to be the limiting factor, but this remains to be seen.

Although we speak of the time spent at a given configuration, this is a bit misleading because the apertures are always in motion. In fact, at any given moment, the spot in the u-v plane representing a baseline measurement will most likely be spanning two distinct u-v points. This is shown graphically in Figure 2.

As the spot in the u-v plane moves through a region representing a u-v point to be measured, the amount of overlap with that point changes. At half of a configuration time before it is centered on the u-v point, the spot first contacts the region. It has just passed through the position of being completely centered in the previous u-v point. One fourth of the configuration time later, it is halfway between two u-v points, and another one fourth later it is centered. The process then repeats for the following u-v point. The total amount of time that a spot is in contact with a designated u-v point is actually two configuration times, although the time over which at least half of the information is coming from one single u-v point is just
one configuration time. A weighting factor (represented graphically in Figure 2) must be applied to the u-v data to try to separate out the information that is coming from the different u-v points. It has been suggested that only measurements to one side of the zero path difference point of the delay line need to be taken, since a real sky brightness will produce a Hermitian spectrum, half of which is therefore redundant. This would reduce the stroke distance by half. However, the need to separate out data from two adjacent u-v points might require taking a two sided measurement, since the Hermitian nature of the data may provide the additional information necessary to effectively separate the u-v point data.

Figure 2. Pass Through Sampling of u-v Points

Over a period of $T_{conf}$, the delay line must stroke far enough to provide the necessary spectral resolution. Not all of the distance needs to be provided by actual linear motion, since multiple reflections can be used as a multiplier to length, but this comes at a cost of sensitivity. The first requirement on the delay line is the resolution that is desired by the science community, which has been set to a value of $R = \lambda/\Delta\lambda = 10,000$. The path difference necessary to achieve a given resolution for the one sided measurements is $R\lambda$. However, to perform the desired interferometry over all of the pixels in the imaging array, the delay line must pass through the zero path difference (ZPD) points of each pixel in the array, which are not collocated. Given an $N \times N$ array of pixels, the path length difference necessary to reach the ZPD of pixels on both sides of the array is $(NB/D)\lambda$. The ratio of what is required to achieve the science requirement in all pixels to achieving the science for a single pixel is then

$$\frac{NB\lambda}{D/R\lambda} = \frac{NB}{R\lambdaD} \quad \text{or} \quad \frac{NB\lambda + 2R\lambda}{2R\lambda} + 1 = \frac{NB}{2RD} + 1$$

(3.4)

where the first case is for a one sided stroke, and the other is for a two sided stroke. These results are independent of the wavelength or the number of reflections. The reason for the difference is shown in Figure 3.

Figure 3. Difference Between Double (left) and Single (right) Sided Measurements

For the double sided case, $R\lambda/2$ data points for science must be taken to each side of the points needed for the reach the ZPD points over the entire FOV. However, in the single sided case, simply stroking over the ZPD points will automatically collect all the necessary data, since the one sided transformation can be taken from either side of the ZPD point.
We will re-write the second form in terms of parameters that have requirements as follows

\[
\frac{NB}{2R_d D} + 1 = \left( \frac{N\lambda}{D} \right) \left( \frac{B}{\lambda} \right) \frac{1}{2R_d} + 1 = \frac{FOV}{2R_d \Delta \theta} + 1 \tag{3.5}
\]

This term appeared in the definition of \( SNR \) given in equation 2.4. For a fixed imaging time, if you are forced to take more data points than you desire, this requires a faster sampling rate. This provides frequencies over a larger band than is necessary, which basically get thrown away. This effectively reduces the efficiency of the process, and introduces the factor in equation 3.5 into the \( SNR \) calculation. This is distinct from the case when more data is collected at the same data rate, but over a longer period of time. When this is the case, the additional data falls within the frequency band of interest, and averaging can be used to actually increase the \( SNR \). For the additional requirements of \( FOV = 0.24 \) deg and \( \Delta \theta = 50 \) mas, this ratio is about 1.7 for the one sided measurements, and 1.8 for the two sided case.

### 4. CONFIGURATION ISSUES

The assumption that was made about having the same time for each sample has significant impact on the system configuration design. Equation 3.2 assumed that each u-v point was passed over exactly one time during the imaging process, and therefore the time allotted to each point is just \( D/v \). This then requires that the apertures be moving at constant velocity during the entire imaging process. Unfortunately, a spinning collection of apertures changing radius will not maintain a constant velocity profile due to angular momentum conservation. In fact, the velocity scales as \( 1/r \), meaning that if the apertures are spinning at a radius of 50 meters (suggested as the minimum measurement distance), they will be moving 10 times faster than when they are at the maximum radius of 578 meters (the radius assuming 3 apertures in a triangle with 1000 meter baseline).

To address this problem, Farley and Quinn\(^1\) produced two designs, the SPECS HEX and SPECS Tetra-Star that use counter masses as governors to control the speed of the apertures (see Figure 4). The HEX is a hub and spoke configuration that reeels out the three counter masses as the apertures are reeled in, thus giving a degree of freedom to control the moment of inertia of the system. Because of Coriolis acceleration as the apertures and masses are reeled in and out (accelerating each in a different direction), a 17 meter radius central truss is used to keep the pendula from getting too close to one another. This size was based on the time set by the \( SNR \) requirement, and what was considered a safe approach distance. There were two main drawbacks to this design. The first was the presence of the central truss, since it is not contributing directly to the imaging process, but adds mass to the system. The second was the speed limit that was imposed by the Coriolis acceleration.

![Figure 4. SPECS HEX and SPECS Tetra-Star Configurations by Farley and Quinn](image-url)
During imaging, this is not a problem, since the speed limit was designed to match the desired imaging time. However, to minimize propellant usage for re-targeting, it is desirable to have the system spinning as slowly as possible, i.e. in the largest moment of inertia configuration. Therefore the system should be returned to the same configuration each time before re-targeting, requiring twice the effective time for an image, or setting the duty cycle defined earlier to just under one half. In reality, a trade would have to be done to determine whether simply putting more propellant on board would be the better way to go, but nonetheless it introduces an inefficiency into the system.

Both of these issues were addressed in the Tetra-Star concept. Here, the counter masses were tethered externally, with each attached to two of the apertures. Now, the central truss has been eliminated, and the only speed limit is that the force of the counter masses on the apertures does not cause the inner tethers to go slack. To further simplify the design, they made the external tethers of fixed length. The velocity profile of the apertures with radius is no longer constant, however the variation is much less than the factor of ten that originally existed.

Unfortunately, even with this innovation in the design, the counter masses represent more or less dead mass that is not contributing to the imaging process. To see if there is any way around the constant velocity requirement we return to the details of the imaging process. At each u-v point, the delay line must stroke through all the necessary path difference to acquire complete science data for each pixel in the imaging array. Because the fringe intensity is simply a function of path difference (and not time), there is no reason that if the apertures were stationary a subset of these measurements couldn't be made at non-contiguous times. Half of the delays could be recorded, and then some arbitrary length of time could pass (provided it is short on the cosmological time scale) before the remaining data is taken. Likewise, even with the apertures moving, a different portion of the delays could be recorded each time one of the spots in the u-v plane passed a particular point. This is shown graphically in Figure 5.

![Figure 5. Combining Delay Line Data from Different Times](image)

Here, the shaded square represents the u-v point to be recorded. At velocity $v$, the time to pass through this point is exactly the desired configuration time ($T_{conf}$). If instead the velocity is $3v$, then one third of the delay offsets can be recorded during the first pass, the second third during the second, and so on. After three passes, all the data has been recorded. Because the same total amount of time has been spent in both cases, the total imaging time is no different than if a constant velocity had been maintained all along. To maintain a smooth radial reel rate, each pass now needs to have some overlap with adjacent u-v points in the radial direction, just as was considered for the tangential. Again, a weighting function can be applied to distribute the data to the proper u-v point.

The superposition of non-contiguous data sets also has some other potential benefits. Referring back to the discussion of re-targeting when the system is rotating at its slowest rate, we would want to perform this maneuver when the apertures are extended to the maximum baseline. This would still mean imaging during a reel-in and then back tracking to the maximum
baseline for re-targeting. Although the reel-out portion could be done more quickly than the reel-in, it will still significantly affect the duty cycle of the system. With a non-overlapping coverage approach such as was originally derived, the approach of imaging both on the way in and the way out does not provide full coverage, since the gaps left by reeling in at twice the rate will not be filled when reeling out at twice the rate. However, since overlap would increase linearly from zero at the maximum radius to 90% at the shortest radius, imaging on both the reel-in and reel-out phases becomes possible. Granted, the complexity of combining the data increases significantly, but there is fundamentally no reason that it cannot be done.

5. OTHER DESIGN CONSIDERATIONS

By recording the data in the way described in the last section, the need for regulating the velocity of the configuration is lifted, and with it the need for the counter masses and extra tethers. This represents a significant savings in mass and hardware complexity. There are still a few details brought up earlier in the paper that must be addressed. It was mentioned in section 2 that although the minimum number of apertures needed for interferometry is two, the minimum number of spacecraft is three. This is because an equal path length for the light must be provided when the two different sources are combined. There are several possible reasons why this third spacecraft should itself also be a collector, and in fact all three spacecraft should be identical collector/combiners. First is that a significant amount of overhead is already associated with designing and building a third entirely different spacecraft. There are certainly gains in only having to design a single spacecraft, and even learning curve savings in building them. Unfortunately, to properly answer the question of which architecture would actually cost less, much more detailed estimates would have to be made.

Up until now, we have focused entirely on the system performance under the assumption that everything is working. However, the possibility always exists that some part of the system may fail over time. For the two collector and one combiner case, if either collector fails or the combiner fails, then the system as a whole has failed and the productivity goes immediately to zero. This is because the probability that the system is functioning properly can be represented by a product of the probabilities that each components is functioning, since all components must be working for the system to work. If instead the system is designed so that each spacecraft is both a collector and a combiner, then even if the system suffers some failures it will still operate, although in a degraded state. For instance, if one of the combiners fails, then two of the three baselines can still collect data, and if one of the collectors fails, then there will still be one baseline that can collect data. For these two cases, the duty cycle defined earlier would simply reduce from nearly one, to roughly 2/3, and then to 1/3. In fact, if we were lucky enough that the failed collect was not on either of the spacecraft with a failed combiner, then all three components could fail and the system would still operate. This is because for the three identical collector/combiner case, we essentially have three separate systems, all functioning in parallel. A complete system failure would require that all three systems fail, the probability of which is the cube of the probability that any one fails. If the probability of one of the three failing is 0.1 (a huge number), there is only a 0.001 chance the system will fail as a whole.

A way to more quantitatively examine the robustness of the system would be to recast the number of images into a utility function where the productivity at any given time is weighted by the probability of being in a particular operational state. As mentioned before, applying a Markov model to the system will then allow the determination of what the most probable states of the system will be over time, and evaluation of an expectation value for the total number of images collected. It does appear however that the redundancy, design savings and learning curve afforded by going to three identical spacecraft would more than offset what would appear to be a marginal cost of additional launch mass.

6. CONCLUSIONS

It has been assumed that aperture area and number of apertures are two of the primary drivers of cost and performance for SPECS. Constructing a performance metric of cost per function using these as free parameters, it was shown that better performance comes from increasing the aperture area and decreasing the number of apertures. Although two apertures are the hard minimum for interferometry, three spacecraft must be launched in total, and it has been argued that the performance gain in going to three identical collector/combiner spacecraft outweighs the marginal cost increase. It also allows the system to continue taking measurements even if one of the apertures or two of the combiners fail, greatly increasing the robustness of the system.
For a fixed speed delay line, it is necessary to provide the same imaging time at each u-v point, to allow all path length differences to be achieved. Existing designs have cleverly addressed this requirement by using counter masses to control the rotational speed of the system. As a result of the imaging process, however, it appears possible to simulate a constant time spent at each point by concatenating $M$ noncontiguous (in time) measurements at any given u-v point, each of duration $1/M$ of the desired measurement time, without loss of information. This would allow the elimination of the counter masses to control the rotation rate, and would add the possibility of increasing the duty cycle by allowing imaging to occur in both the reel-in and reel-out phases.

There are other issues to be addressed that will affect the final system configuration, one of which is the thermal emission from the tethers or even the other spacecraft. With the triangular configuration, and no central combiner, the line of sight for the light path between spacecraft runs parallel to the tether. It must be determined whether emissions from the tether will significantly affect the system performance. If not, then most of the other design details will be common problems with any configuration (tether dynamics, etc.), and the architecture outlined here appears to offer many advantages.

7. ACKNOWLEDGMENTS

This work was funded under the Cross Enterprise Technology Development Program through NASA Goddard Space Flight Center, under the technical supervision of David A. Quinn. The authors would like to thank John C. Mather for his invaluable assistance in understanding the signal to noise scaling for SPECS.

8. REFERENCES

1 R.E. Farley and D.A. Quinn, Tethered Formation Configurations: Meeting the Scientific Objectives of Large Aperture and Interferometric Science, AIAA 2001-4770.

2 Meinel, Cost Scaling Laws Applicable to Very Large Optical Telescopes, SPIE, 1979.

