

OPTIMIZATION OF EARLY-PHASE CISLUNAR NAVIGATION CONSTELLATIONS FOR USERS NEAR THE LUNAR SOUTH POLE

Mark Hartigan*, Dillon Smith*, and E. Glenn Lightsey†

To meet the needs of burgeoning global scientific and strategic interest in cislunar space, organizations such as NASA and the ESA have expressed interest in establishing a cislunar-based position, navigation, and timing (PNT) system. Minimal system implementations (defined as between 4 and 8 satellites) will provide critical PNT services to near-term cislunar missions and set the stage for constellation expansion and the establishment of a global lunar navigation service. This work proposes a set of fixed metrics through which cislunar PNT constellations can be evaluated, including: coverage, gap time, dilution of precision, user equivalent range error, and stationkeeping costs. Several minimal implementations that are proposed in literature are first examined, then optimized for certain design traits to improve system performance. Finally, simulations are conducted to compare constellation performance for end users – specifically ground stations and rovers.

INTRODUCTION

National and international interest is rapidly turning towards building a sustained human presence at the moon and in cislunar space. The White House Office of Science and Technology Policy has released a National Cislunar Science and Technology Strategy that identifies the need for scalable and interoperable cislunar positioning, navigation, timing, and communications (PNTC) capabilities to support NASA’s Artemis program and beyond.¹ NASA’s Moon-to-Mars architecture identifies this same need and is seeking comments and feedback on providing PNTC infrastructure for cislunar space.² NASA is currently in the early stages of developing such a cislunar constellation – called LunaNet – and has released details outlining the need for interoperability,³ while the European Space Agency has targeted the same need within their Moonlight Initiative.⁴ Current literature is broadly focused on either optimizing late-stage, full-scale (greater than 18 satellites) constellations or analyzing discrete examples of minimal systems to achieve targeted objectives.^{4,5} However, minimal system implementations (defined as consisting of between 4 and 8 satellites) will provide critical PNT services to near-term cislunar missions and set the stage for constellation expansion and the establishment of a global lunar navigation service. This work aims to bridge this lunar PNT design gap by first analyzing several minimal implementations that are proposed in literature, and then optimizing certain design parameters to improve system performance.

The focus of this analysis is on minimal systems capable of providing navigation solutions to users near the lunar south pole using trilateration of pseudorange measurements derived from signals broadcast by orbiting satellites. Such a service should strive to minimize navigation errors,

*Graduate Research Assistant, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA.

†David Lewis Professor of Space Systems Technology, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA.

maximize service availability to users, and operate for as long as possible within the available resources. Orbital perturbations around the moon – largely from lunar non-spherical gravity, the Earth, and Sun – greatly limit the stability of lunar orbits and therefore exacerbate stationkeeping costs, reducing mission life. However, select orbit types, including Elliptical Lunar Frozen Orbits (ELFOs) and Near Rectilinear Halo Orbits (NRHOs), possess long term stability and thus prove to be advantageous for a PNT system.^{6,7} Within this subset of orbits, those providing the lowest Positional Dilution of Precision (PDOP) to the user’s location are desirable. Using orbital parameters bounded by stability and a cost function minimizing PDOP, orbit geometries and satellite spacing are optimized in order to provide the lowest position error to users while utilizing a minimal number of satellites.

The method used to derive these results employs the NASA-developed General Mission Analysis Tool (GMAT) through the Python API in order to propagate lunar orbits. Use of the API allows for GMAT to be integrated into a Python-based optimization wrapper in which stable optimal lunar orbits are obtained. User navigation error (UNE) is then derived from PDOP and modeled error sources such as clock drift, receiver noise, and multipath error. Metrics are produced for comparison between the different constellations found by this study and those examined in literature. Other metrics analyzed are coverage, defined as when a target can obtain a position fix within an operable threshold, and gap time, which is when coverage is not available. A developed set of standard metrics allows for the comparison between proposed implementations in an objective design trade analysis.

SIGNAL DESIGN

For a fully deployed PNT solution with global lunar coverage, the strongest candidate for signal design currently is to employ a GNSS-like service that utilizes radionavigation signals consisting of a carrier frequency modulated with a pseudo-random noise (PRN) code and pertinent navigation data. These PRN codes are unique to each satellite and repeat at fixed, well-defined intervals so by measuring the part of the PRN code received, the user can derive time-of-flight from multiple nodes to triangulate their position. Modulating a PRN code on the carrier frequency creates a spread-spectrum signal, which is beneficial for many reasons:⁸ while the base signal is below the ambient RF noise floor, if the PRN code is known, the signal-to-noise ratio is very high; the autocorrelation properties of PRN codes mean that any multipath errors greater than one chip width are eliminated; and it is robust against radio frequency interference.

Another consideration is operating spectrum, or the carrier frequency of the navigation signal. Currently, LunaNet has specified a carrier frequency of 2491.75 MHz⁹ which conforms to the Space Frequency Coordination Group recommendation given in SFCG 32-2R4*. Analysis in the following sections will be focused on the currently-provided LunaNet specifications; considerations must be made for other frequency choices to not interfere with planned radioastronomy on the far side of the moon and remain within recommendations of the SFCG. The far side of the moon is permanently hidden from the Earth, and therefore devoid of RF interference; this makes the far side invaluable for radio astronomy and is of high scientific interest to keep it interference-free. Current SFCG recommendations are using the 2483.5-2500 MHz band for dedicated PNT signals, while the 2025-2110 MHz and 2200-2290 MHz bands may be used for communications containing PNT ranging signals (all of the above fall into the IEEE S band of radio frequencies).

*[https://www.sfcgonline.org/Recommendations/REC_SFCG_32-2R4_\(Freqs_for_Lunar_Region\).pdf](https://www.sfcgonline.org/Recommendations/REC_SFCG_32-2R4_(Freqs_for_Lunar_Region).pdf)

METRICS OF INTEREST

In order to compare proposed constellations of an early stage cislunar PNT system, standard metrics must be introduced. Determining these metrics will allow for an equitable comparison to be carried out. Some metrics are explored in the current work, and others need to be implemented in the future to allow for a comprehensive study.

System Coverage

Coverage refers to the amount of time that a user has access to the PNT system. For this study, a user is said to have access to the system when it is in view of at least four PNT satellites to enable pseudorange measurements. Furthermore, a satellite is in view of a target when there is a clear line of sight between the satellite and the user, and if the satellite has an angle with the local horizon of at least 5 degrees. The inclusion of the 5 degree constraint takes into account the varying geography of the lunar surface to ensure general satellite visibility. Along with coverage, the average number of satellites in view for locations of interest will be kept as a comparison metric. Even though only four satellites are required to be in view to obtain a PNT solution, a greater number of satellites allows for a better range estimate.

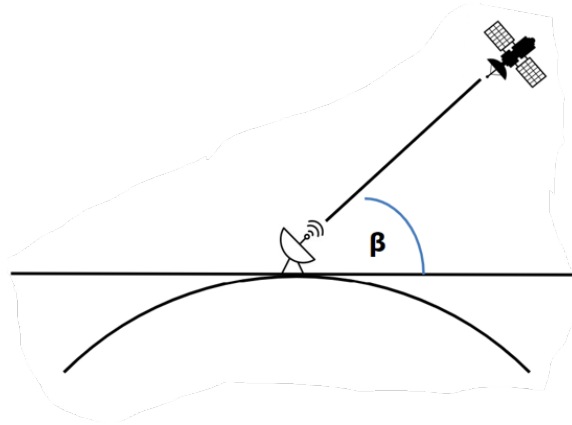


Figure 1: Visibility requirements with $\beta \geq 5^\circ$.

Gap Time

Gap time refers to the amount of time that a user does not have access to the PNT system. Therefore, it is defined as when there are fewer than four satellites in view of a target user, and it will be calculated as the difference between the total propagation time and the total amount of time that coverage is available.

$$T_{gap} = T_{prop} - T_{cov} \quad (1)$$

A variant of gap time that may prove to be more insightful is instead calculating the maximum continuous gap time that a user will experience. This is because many users will possess inertial navigation systems independent of the PNT system that they will rely on during the absence of coverage. In turn, users will maintain a position estimate during gap periods that will become increasingly worse the longer the gap time. Therefore, frequent periods of smaller gaps may be preferable to fewer periods of larger gaps.

Positional Dilution of Precision

Positional Dilution of Precision (PDOP) refers to the uncertainty introduced into a position measurement due to the geometry of the system. Ideally, PDOP for a given constellation should be as small as possible to maintain higher accuracy PNT measurements, and this occurs when the PNT satellite geometry is the most diverse – illustrated in Figure 2.

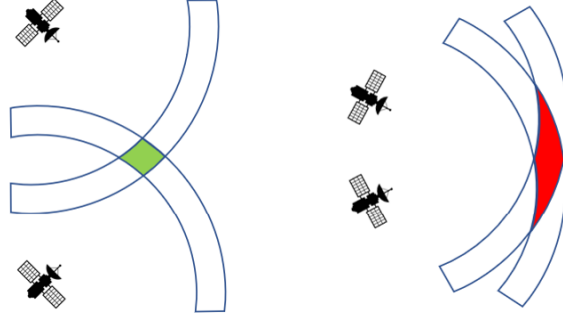


Figure 2: The effect of PNT satellite geometry on measurement uncertainty.

PDOP is only defined for a target if the user can receive at least four signals from different PNT satellites at one time, since at least four range measurements are required in order to determine the position of a user in 3D space using trilateration. To calculate PDOP, the matrix H must first be created using position knowledge of the system.

$$H = \begin{bmatrix} \frac{x_1 - x}{||r_1||} & \frac{y_1 - y}{||r_1||} & \frac{z_1 - z}{||r_1||} & -1 \\ \frac{x_2 - x}{||r_2||} & \frac{y_2 - y}{||r_2||} & \frac{z_2 - z}{||r_2||} & -1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_n - x}{||r_n||} & \frac{y_n - y}{||r_n||} & \frac{z_n - z}{||r_n||} & -1 \end{bmatrix} \quad (2)$$

Here, x , y , and z denote the Cartesian coordinates of a user, while the x_1 , y_1 , and z_1 denote the Cartesian coordinates of one PNT satellite. The $||r||$ terms represent the distance between the PNT satellite and the user. It is noted that to meet the requirement of utilizing at least four range measurements, n must be ≥ 4 . Using the H matrix, the Geometric Dilution of Precision (GDOP) covariance matrix can be calculated.

$$Q = (H^T H)^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{xt} & \sigma_{yt} & \sigma_{zt} & \sigma_t^2 \end{bmatrix} \quad (3)$$

Within the covariance matrix, Q , the diagonal elements represent the variances of the x , y , and z positional uncertainties along with the time uncertainty, t . The off-diagonal elements represent the covariances between terms. To extract the PDOP, the root sum is taken of the positional variances.

$$PDOP = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \quad (4)$$

User Equivalent Range Error

Besides dilution of precision, the other variable required to calculate user navigation error is user equivalent range error (UERE). Several factors contribute to UERE; only a few key ones will be covered.

Clock Errors Navigation solution accuracy relies on the accuracy of both the broadcast node ephemerides and clock drift models. No clock is perfectly accurate with respect to a reference time: oscillators are subject to bias, drift, temperature effects, and random walk- and random jitter-type noise. For this reason, the GPS control segment regularly adjusts the onboard clocks in GPS satellites back to GPS Time; they also provide a model of clock drift in the form of polynomial coefficients. Clock drift at time t_1 can be characterized by Equation (5), where t_0 is the starting epoch, f_0 is the nominal oscillator frequency, Δf is the frequency bias, $\dot{f}$ is the frequency drift, and $\ddot{f}$ is the oscillator noise.

$$\Delta t(t_1) = \Delta t(t_0) + \frac{\Delta f}{f_0}(t_1 - t_0) + \frac{\dot{f}}{2f_0}(t_1 - t_0)^2 + \int_{t_0}^{t_1} \frac{\ddot{f}(t)}{f_0} dt \quad (5)$$

The U.S. Air Force's Global Positioning Systems Directorate regularly updates clock correction parameters in each satellite's navigation message providing the clock bias (a_0), clock drift (a_1), and clock drift rate (a_2) coefficients in Equation (6) below.¹⁰

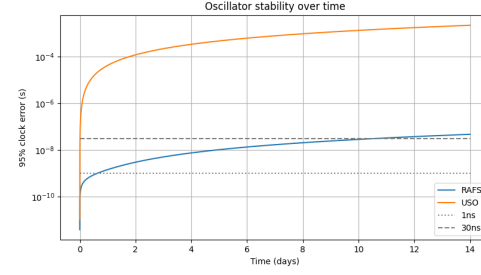
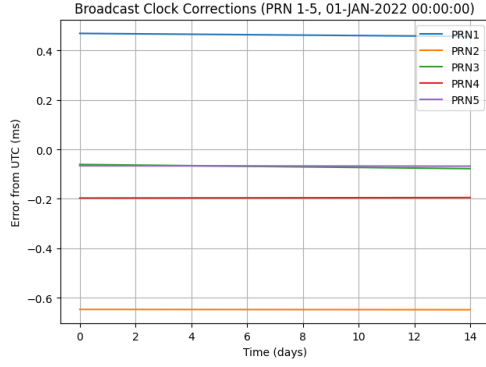
$$\Delta t(t_1) = a_0 + a_1(t_1 - t_0) + a_2(t_1 - t_0)^2 \quad (6)$$

Additionally, Figure 3a plots these corrections for a few sample satellites over two weeks. It is evident that the clock drift is largely driven by clock bias on the order of milliseconds, while the drift is less than a nanosecond per second and drift rate is roughly zero. According to the GPS Standard Positioning Service standards,¹¹ the 95% statistical error from Coordinated Universal Time (UTC) shall never exceed 30 nanoseconds – and in practice is usually less than 7 [8, Chapter 5.6].

However, the GPS ephemerides (as well as Galileo and GLONASS) do not attempt to model random noise, the final term in Equation (5). One way to account for random walk and random jitter noise is including in the error budget Hadamard variance $\sigma_H^2(\tau)$,¹² which estimates the relative frequency deviation of an oscillator due to random fluctuations over N intervals of length τ . The equation for Hadamard variance is

$$\sigma_H^2(\tau) = q_1\tau^{-1} + \frac{1}{6}q_2\tau + \frac{11}{120}q_3\tau^3 \quad (7)$$

From this equation, the RMS time error of the oscillator can be written as $\tau * \sigma(\tau)$, where τ is the time since the oscillator was last synchronized [8, Chapter 4.2.2]. Using the fitted models for a Rubidium Atomic Frequency Standard (RAFS) and Ultra-Stable Oscillator (USO) from,¹³ the 95% error is plotted in Figure 3b. As shown, the 30-nanosecond standard set by GPS is quickly violated by a USO but takes around 11 days uncorrected from a RAFS for its uncertainty to rise above that threshold; to meet the 1-nanosecond performance GPS typically achieves, it would require accurate calibrations roughly twice daily.



(a) Broadcast clock corrections from GPS satellites (PRN codes 1-5) on Jan 1, 2022.

(b) Estimated 95% oscillator deviation over time.

Figure 3: Clock properties.

For this analysis, it will be assumed that constellations adopt an update schedule that maintains the 30-nanosecond standard. How this is achieved in practice will vary greatly based on system implementation.

Orbit Determination Errors Directly related to clock corrections are the node ephemerides, which in GPS are broadcast to the user to compute node positions for pseudorange measurements. In GNSS, these ephemerides are distributed by using ground station tracking to get a precise orbit determination (OD) solution, which is then uploaded into the satellite navigation message. In the lunar environment before any lunar stations are established, this OD solution will have to be computed onboard the nodes as there likely will not be enough resources available to track every node from the ground and update ephemerides daily. Therefore any satellite OD covariance will be passed along to the user – unless lunar ground stations for tracking are constructed in later phases. Small et. al.¹³ have demonstrated that a lunar satellite is capable of achieving a mean 1σ RSS position error of approximately 3.7 meters using weak GNSS with outages, optical navigation ranging to the moon using center-finding, and a RAFS clock.

Receiver Noise and Multipath Receiver noise and multipath errors are difficult to distinguish as they appear as very similar types of noise. Multipath errors arise from the radio signal arriving through an indirect path to the receiver. Luckily, due to the strong autocorrelation properties of PRN codes, any multipath errors greater than one chip width (around 59 meters based on current LunaNet specifications⁹) are eliminated. Multipath errors in GPS may range from 1 to 5 meters depending on if the environment is benign or highly reflective [8, Chapter 5.4.2].

Receiver noise is characterized by the following equation [8, Chapter 10.5]:

$$\sigma_{\Delta\tau} = cT_c \sqrt{\frac{d}{4TC/N_0}} \quad (8)$$

where c is the speed of light, T_c is the chipping time of the navigation signal, d is the receiver correlation spacing, T is the receiver averaging time, and C/N_0 is the signal power to noise spectral density ratio. According to LunaNet specifications, the planned chip width is 19.55 ns, which corresponds with a chipping rate of 5.115 Mcps.⁹ The correlation spacing $d \in [0.1, 1]$ is the space

between samples in the delay lock loop, and is commonly set to 0.1 to minimize noise – lower values do not yield greater benefit because of filters in the delay lock loop that round over the correlation peak. It is worth noting that in order to implement these values, the receiver needs a processor a speed of at least $\frac{1}{dT_c} \text{ Hz}$. T is the time over which measurements are averaged to yield a single solution; longer times typically yield better results, but if the receiver is moving quickly it might begin conflating measurements. Typical values range from 10 seconds in low-speed applications to 0.01-0.02 seconds for near-instantaneous measurements. Finally, the signal-to-noise ratio is comprised of many factors, but in the Earth environment C/N_0 is nominally in the region of 40 – 46 dB-Hz.¹⁴

Since one chip width is the minimum position resolution obtainable by users, improving the clock rate onboard each node and subsequently reducing chip width can improve navigation accuracy from code measurements. Unfortunately, limitations may then begin to arise in receivers as required processing speed is at least $\frac{1}{dT_c} \text{ Hz}$.

Lunar Regolith One large unknown on the lunar surface is the lunar regolith. Not much research has been done into its effects in radio signals; however, airborne particles from the regolith kicked up during takeoff or landing can attenuate signals if they are on a similar scale to the wavelength of the signal. Figure 4 plots the grain size distribution of the lunar regolith. Based on this data, most

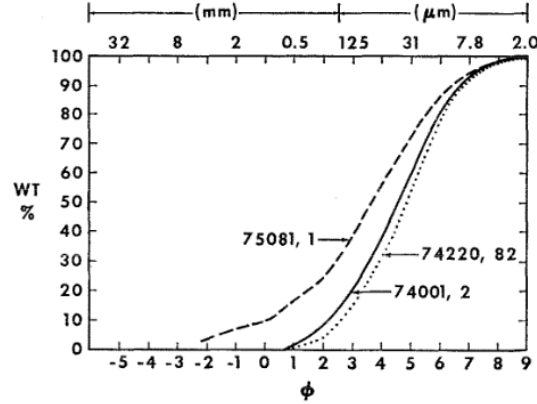


Figure 4: Grain size distribution of mature lunar soil (75081) and orange soil (74220).¹⁵

(> 99%) of regolith particles are under 5mm in grain size, which would then only attenuate signals on the order of 60 GHz or greater. Current LunaNet specifications target the 2483.5-2500MHz band for navigation signals, which is well beneath the attenuation threshold. A final consideration for the lunar regolith is its refractive index and potentially dispersive properties, similar to that of the ionosphere or troposphere. Literature available in this area is limited, but the error that arises is formulated as a path integral from the satellite to receiver [8, Chapter 5.3.1] as in Equation (9):

$$\Delta\tau = \frac{1}{c} \int_{\text{Satellite}}^{\text{Receiver}} [n(l) - 1] dl. \quad (9)$$

Here, c is the speed of light and $n(l)$ is a position-dependent refractive index of the media the wave is passing through. The refractive index also varies with the frequency of the signal passing through it – since the carrier and code are transmitted at different frequencies, this causes a phenomenon called

code-carrier divergence. Around Earth, the ionosphere and troposphere are primary contributors; however, these stretch tens to hundreds of miles above the Earth's surface, so waves transmitted by GNSS satellites travel through them often for hundreds of miles. If a similar phenomenon occurs in lunar regolith, it may only be for tens of meters and therefore impact the signal much less. In any case, this merits further research.

Table 1 gives some nominal values for 95% error bounds from the sources discussed above at a maximum Age-of-Data (AOD). If the clock model uncertainty is kept to the 30 ns bound mentioned previously, this would yield a maximum uncertainty of approximately 9 meters. Based on the orbit determination (OD) uncertainty computed through simulations by Small et al.¹³ which use weak GNSS (accounting for outages at perilune), optical center-finding of the Moon, and a Rubidium standard, this would transfer an average uncertainty of 9 m to the user. Assuming a relatively benign multipath environment when compared to the Earth, a multipath error of roughly 2 m is reasonable.

Table 1: Collected error sources at maximum AOD.

Source	95% RSS Position Error
Clock Model	8.994 <i>m</i>
OD Uncertainty	9.081 <i>m</i>
Receiver Noise	19.818 <i>m</i>
Multipath	1.960 <i>m</i>
Lunar Regolith	0 <i>m</i>
Signal-In-Space Errors	12.931 <i>m</i>
Total	23.663 <i>m</i>

If the lunar regolith remains undisturbed and does not attenuate any of the signal, that results in a total signal-in-space error of 13 meters for the nominal LunaNet configuration based on the assumptions outlined above. Using a receiver with GPS-like C/N_0 and a low averaging time ($T = 0.02$ s) results in receiver noise of 20 m totalling a final UERE of 23.6 meters. This value for receiver noise is relatively high and can likely be driven lower with more modern receiver designs.

User Navigation Error

User navigation error (UNE) refers to the compounded uncertainty a positional measurement carries due to geometric, systematic, and environmental sources. UNE is defined as the product of the PDOP and user equivalent range error.

$$\text{UNE} = \text{PDOP} \cdot \text{UERE} \quad (10)$$

PDOP is defined as it was previously, and UERE is the combination of environmental and system induced uncertainties. For this study, UERE will remain as a constant value that is calculated as the root sum square of the contributors shown in Table 1.

Stationkeeping Costs

A goal for designing a cislunar PNT architecture should be that it has longevity. Similar to GNSS architectures, a cislunar system should aim to be of service for decades. In the cislunar

environment, stationkeeping costs can rise considerably depending on the type of orbit. Two large categories of stable lunar orbits that are being explored are Elliptical Lunar Frozen Orbits (ELFOs) and Near Rectilinear Halo Orbits (NRHOs). ELFOs are constructed such that the drift in their orbital parameters are minimized, and analytical solutions to calculating families of ELFOs exist.¹⁶ NRHOs are a different group of orbits that utilize the Lagrangian points within the Earth-Moon system. These unique areas of space with balancing gravitational forces allow for stable orbits to be achieved.¹⁷ Specifically, NRHOs exhibit highly eccentric orbits that exist at the edge of Lagrangian points which allow for near lunar passes.

The calculation of stationkeeping costs is not performed in this study. In the future, it can be done by utilizing Lambert’s problem to calculate the Δv cost of transferring from a drifted orbit to the initial orbit at the point of lowest cost. This study utilizes constellations containing ELFOs and NRHOs in order to keep stationkeeping costs low without directly using it as a comparison metric. However, it is noted that not all ELFOs and NRHOs required the same minimum level of station keeping. Therefore, this metric should be incorporated into a future study.

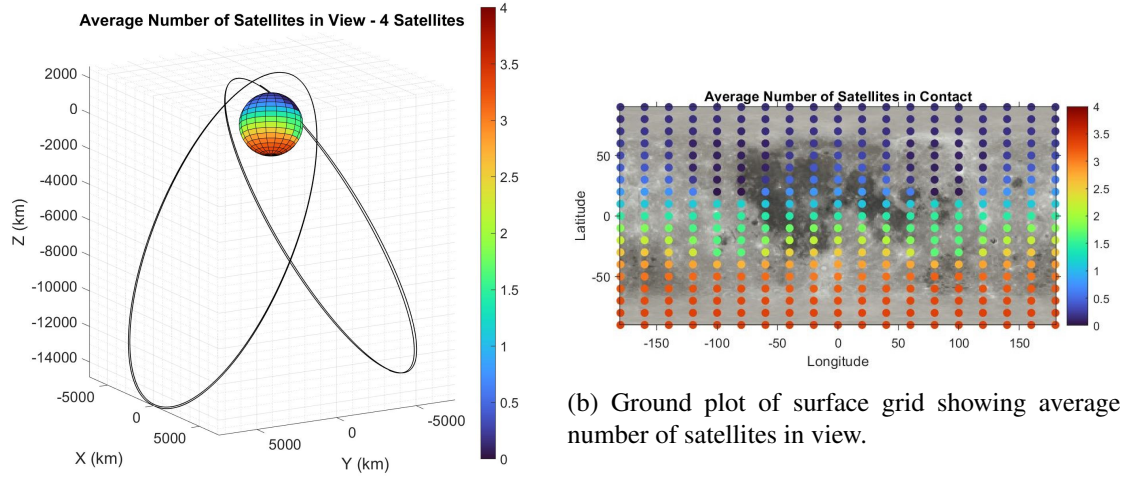
CONSTELLATION EVALUATION

The purpose of this software utility is to provide the necessary tools to analyze constellations on the basis of navigation performance. To this end, this paper will explore several early-phase minimal constellations for navigation at the lunar south pole. The lunar south pole has been identified by NASA, the ESA, and others as a location of primary interest on the moon – mostly due to the existence of permanently sunlit regions and several craters that potentially contain water ice. Therefore, initial satellites deployed into this constellation would best be utilized by providing coverage to the lunar south pole. As this has been a recent topic of interest to the cislunar community, this paper will be analyzing two proposed constellations from literature and two novel designs.

Each constellation is first propagated in GMAT; next, the metrics of interest that define performance are calculated using MATLAB. All plots produced during this process maintain a standard set of configurations in order to facilitate the comparisons between constellations. Key outputs include orbit visualizations with the average number of satellites in view displayed as a color map on the surface of the moon, plots indicating the number of satellites in view to a ground station as a function of time, and plots of the UNE over time for the south pole. All of these outputs will be explored as they are applied to constellations. To assist with determining metrics, a grid of ground stations was created on the lunar surface, spread equally apart by 10 degrees latitudinally and 20 degrees longitudinally. This results in a grid of 325 ground stations.

Constellations from Literature

The first constellation explored is adapted from Bhamidipati et al.,¹⁸ depicted in Figure 5a. This four-satellite constellation consists of two mirrored orbital planes each containing two satellites. The planes are separated by a right ascension of 180° , within which each satellite is separated by 180° . These orbits have an altitude of 9750.5 km, an eccentricity of 0.7, an inclination of 63.5° , and an argument of perilune of 90° .



(a) Orbit visualization with 3D surface plot.

Figure 5: Global performance visualization for the four-satellite constellation.

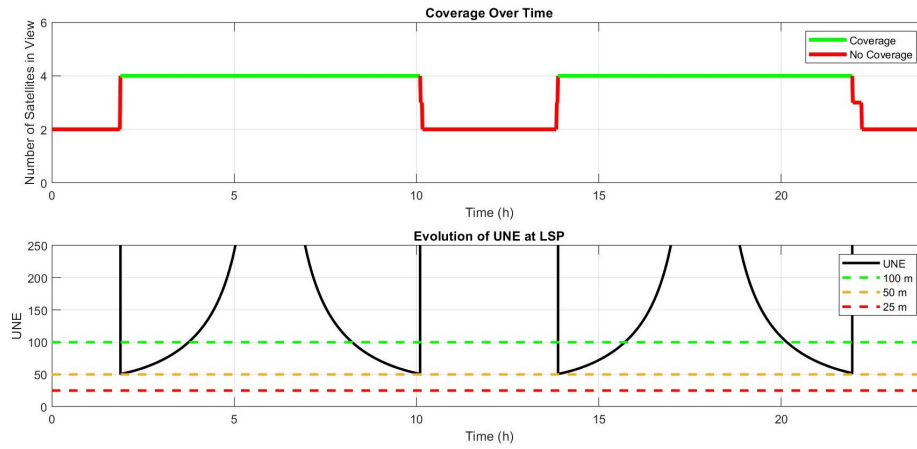


Figure 6: Four satellite constellation; number of satellites in view of the lunar south pole over time (top) and how the 3σ UNE changes over time (bottom).

This constellation exhibits a common characteristic that will be seen in most other constellations, which is that minimal designs of a cislunar PNT system utilize two orbits that are mirrored across the lunar poles. The color map denotes the average number of satellites in view at all ground locations over the course of the constellation propagation. Each lunar surface panel corresponds to one of the ground stations discussed previously, in which each panel is centered around one point grid point. By looking at Figure 5a, it can immediately be seen that the lunar south pole receives the highest coverage as it exhibits the highest average number of satellites in view. It is also noted that when examining future constellations with more than four satellites, the color map will maintain an upper limit of four satellites in order to allow for a direct comparison. Figure 5b unwraps the 3D surface plot shown in 5a into a 2D latitude-longitude scatter plot. Similarly, each node represents one of

the ground stations defined previously. A trend is present where the average number of satellites in view drops off as the latitude increases. It is worth noting that as the average decreases from four satellites in view, it does not mean that there are never four satellites in view simultaneously. Many higher latitude locations are still able to obtain a position fix for at least some period of the constellation propagation despite their low averages. It is clear, however, that lower latitudes are favored (the goal of the design).

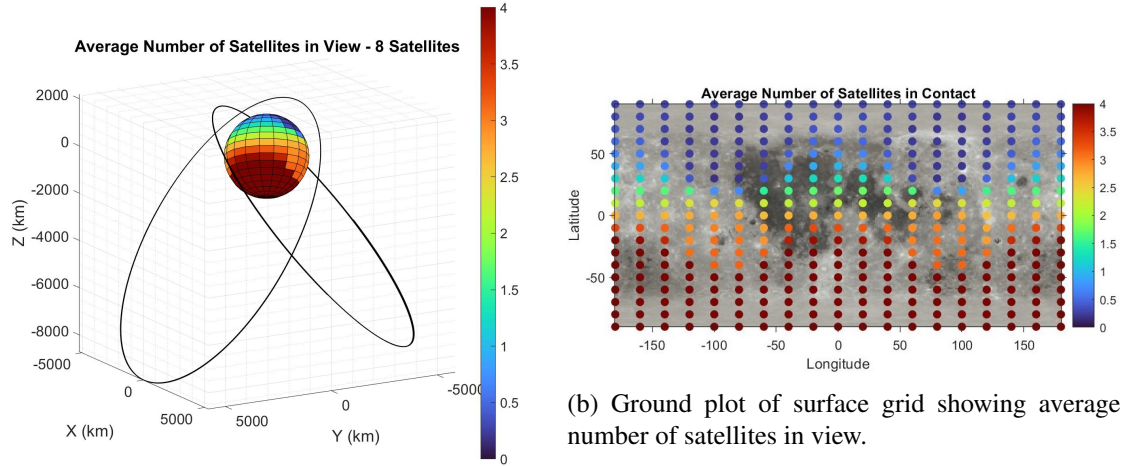
A useful output from post-processing is the coverage over time plot shown in the top of Figure 6. For the four satellite constellation, it is shown that the lunar south pole remains in coverage of the cislunar PNT system for the majority of the 24 hour propagation. The geometry of the constellation, having two satellites in mirrored orbits with a true anomaly separation of 180 degrees, results in a regular switch between four to two satellites in view. To better understand the performance implications, the PDOP for each time step of the propagation is multiplied by the total UERE displayed in Table 1. The resulting UNE over time plot for the lunar south pole is displayed as the bottom plot in Figure 6. This plot also features thresholds for desirable UNE values. This constellation results in a minimum UNE of 50 m, which occurs just as four satellites come into and leave from view. This makes sense as when two satellites become visible or no longer visible, the geometric layout of all satellites will be the most diverse, thus reaching the lowest PDOP value. Furthermore, the absence of a UNE measurement at a time step represents a period where there are not enough satellites in view to make a measurement, which aligns with the coverage over time plot. It is noted that UNE can only achieve a lower bound that matches the value of the UERE as PDOP can only be reduced to a minimum value of 1. Therefore, achieving a low PDOP is an attainable goal for designing a constellation, but overall performance will be limited by how much the UERE can be minimized.

Table 2: Periods of coverage and gap for the four-satellite constellation.

Metric	Hours
Total Coverage Time	16.31
Max. Continuous Coverage	8.23
Total Gap Time	7.69
Max. Continuous Gap	3.76

As seen in Table 2, the constellation spends a majority of the propagation period providing coverage to the lunar south pole. However, there are large gap periods that occur, which might place a constraint on any user that does not have a high fidelity inertial measurement unit (IMU) to use for navigation in the absence of coverage. These results are expected for an early stage constellation with the minimum number of satellites possible, and it is expected that gap times will decrease as additional satellites are added.

The second constellation is comprised of eight satellites and was conceived by Murata et al.,¹⁹ depicted in Figure 7a. Their work used two elliptical lunar frozen orbits (ELFOs) separated by a right ascension of 180° ; in one, each satellite is at a true anomaly of $90i^\circ$, $i = 0, 1, 2, 3$ and in the other $(45 + 90i)^\circ$, $i = 0, 1, 2, 3$. The selected orbits have a semimajor axis of 6541.4 km, eccentricity of 0.6, inclination of 56.2° , and argument of perilune of 90° .



(a) Orbit visualization with 3D surface plot.

Figure 7: Global performance visualization for the eight-satellite constellation.

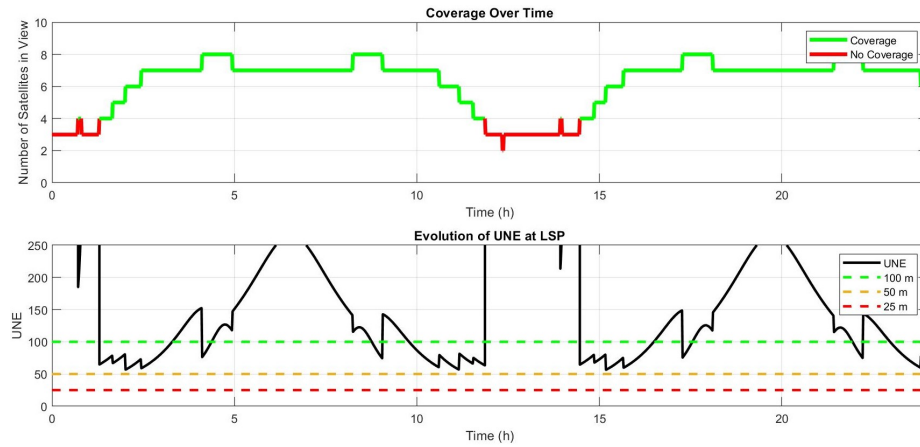


Figure 8: Eight satellite constellation; number of satellites in view of the lunar south pole over time (top) and how the 3σ UNE changes over time (bottom).

Table 3: Periods of coverage and gap for the eight-satellite constellation.

Metric	Hours
Total Coverage Time	20.27
Max. Continuous Coverage	10.58
Total Gap Time	3.73
Max. Continuous Gap	2.05

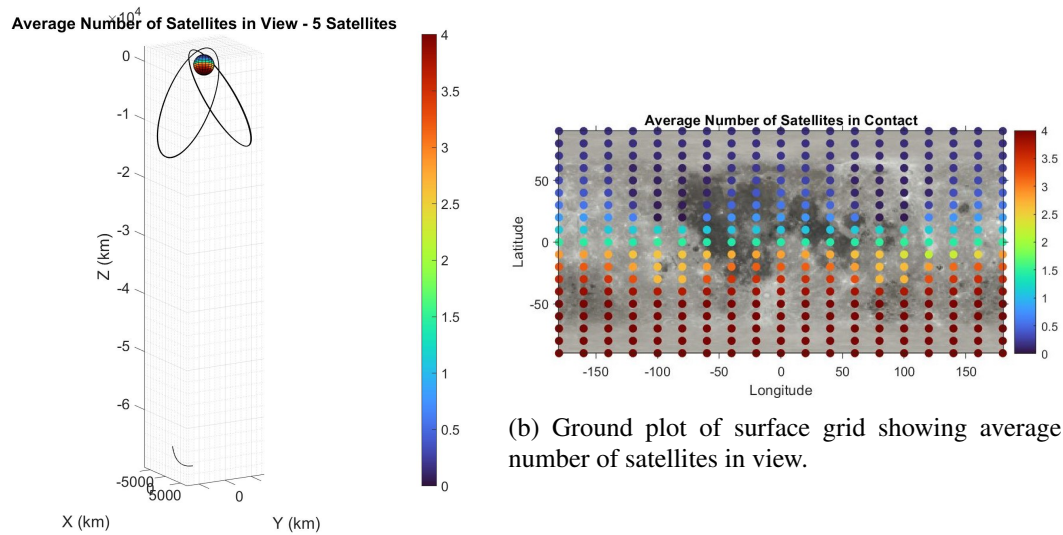
The increase in satellites from the previous constellation results in a higher average coverage – as

shown in Figure 7b – with an average coverage of ≥ 4 expanding to more latitudes. True anomalies between the two orbital planes are staggered, which manifests in the gradual changes from three to eight visible satellites in Figure 8. However, the equispaced nature means that very rarely are all eight satellites ever in view of the south pole. Overall there is a better performance w.r.t. UNE, but the minimum is higher than that of the four-satellite constellation, indicating that the constellation geometry has room for improvement. As expected, the total coverage time has increased over the previous constellation; it would be preferable to reduce this to zero, which should be possible with an eight-satellite constellation.

Novel Constellations

The third constellation to be analyzed is an adaption of the first four-satellite constellation. It uses the same geometry but adds an additional satellite in a near-rectilinear halo orbit (NRHO) similar to that planned for use by the Lunar Gateway (Figure 9a). This new node will increase the number of visible satellites for long periods at a time due to its highly elliptical orbit, improving geometry and reducing dilution of precision. Currently, NRHOs are being studied by NASA and other partners to in order to verify their stability. One such mission, CAPSTONE, is currently utilizing the NRHO selected to be used for the Lunar Gateway mission to gather more data on its characteristics.²⁰

For this study, the NRHO satellite remains in view of the south pole for the entire constellation propagation. This results in a direct enhanced performance over the four satellite case, as seen in Figures 9a and 9b. Importantly, there will be periods of time in which this satellite is not in view of the south pole, resulting in performance similar to the four-satellite case. There is not a substantial increase in coverage time; the main gains occur in UNE, shown in Figure 10 – a majority of measurements sit under the 100-meter threshold. Having a satellite directly overhead greatly improves PDOP calculations, explaining the increase in performance.



(a) Orbit visualization with 3D surface plot.

(b) Ground plot of surface grid showing average number of satellites in view.

Figure 9: Global performance visualization for the five-satellite constellation.

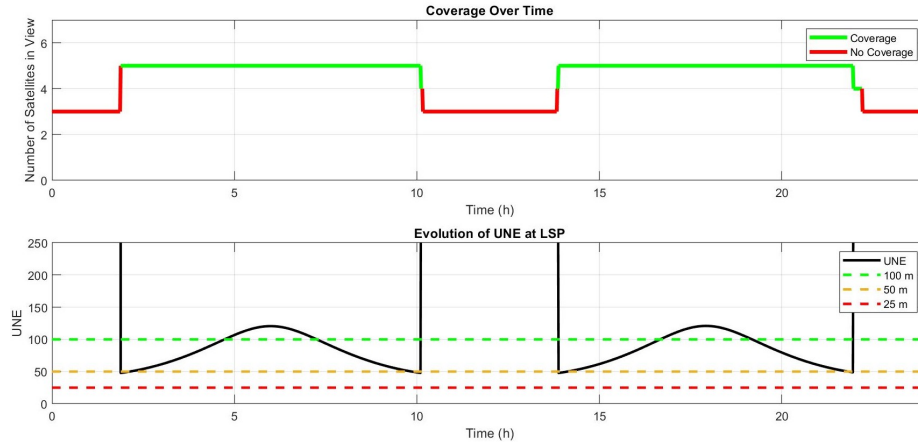


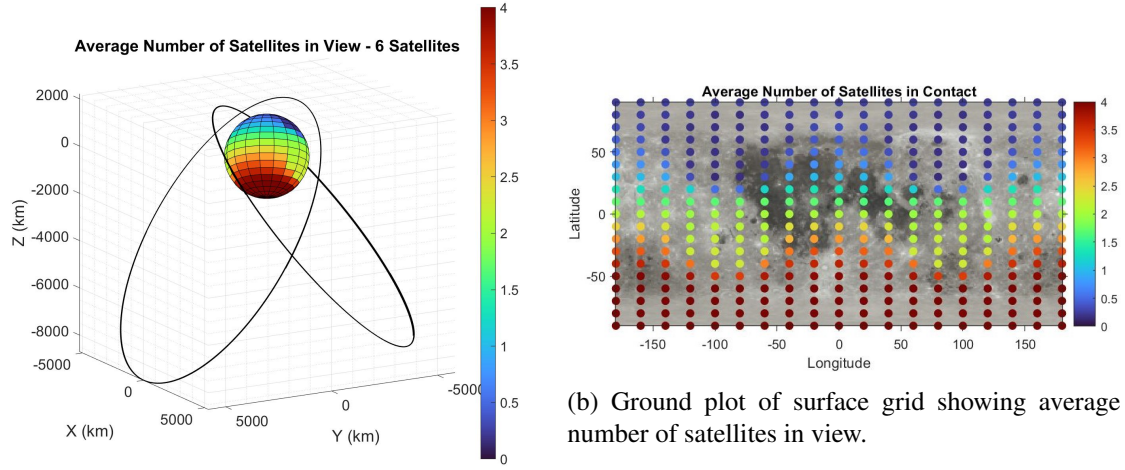
Figure 10: Five satellite constellation; number of satellites in view of the lunar south pole over time (top) and how the 3σ UNE changes over time (bottom).

Table 4: Periods of coverage and gap for the five-satellite constellation.

Metric	Hours
Total Coverage Time	24.00
Max. Continuous Coverage	24.00
Total Gap Time	0
Max. Continuous Gap	0

Finally, a six-satellite constellation is considered as an adaptation of the one proposed by Murata et al (Figure 11a). It maintains the same two mirrored orbital planes, but replaces the equispaced satellites with one string-of-pearls formation in each orbit, with the satellites following each other at true anomalies of 0° , 160° , and 200° . This spacing was adapted from a single-ELFO three-satellite constellation explored by Ely and Lieb.²¹

Coverage shown in Figure 11 appears similar to that seen in Figure 7. More interestingly in Figure 12, this constellation has no drop in coverage – despite having fewer satellites than the eight-satellite constellation. While it performs better in terms of coverage time, the UNE performs more poorly; this exemplifies the tradeoff between coverage and low UNE, showing one does not imply the other. Table 5 shows the lack of coverage gaps, though some missions may consider the periods of large UNE as effectively gaps in coverage depending on navigation requirements.



(a) Orbit visualization with 3D surface plot.

Figure 11: Global performance visualization for the six-satellite constellation.

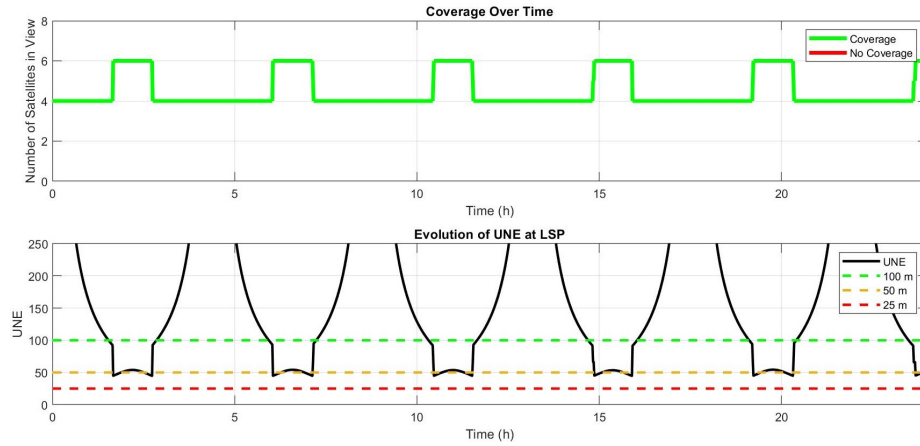


Figure 12: Six satellite constellation; number of satellites in view of the lunar south pole over time (top) and how the 3σ UNE changes over time (bottom).

Table 5: Periods of coverage and gap for the six-satellite constellation.

Metric	Hours
Total Coverage Time	16.64
Max. Continuous Coverage	8.36
Total Gap Time	7.36
Max. Continuous Gap	3.68

CONSTELLATION OPTIMIZATION

As shown in the previous section, improvements can quickly be made to constellations by making architectural changes. This begets the question, how much more performance can be obtained from a minimal constellation? To address this, the starting true anomalies of satellites can be iterated upon to find the optimal spacing to maximize performance. An objective function can be defined s.t. an optimizer seeks to find

$$\min_{\bar{f}_0} F(\bar{f}_0) = \frac{1}{N * P_{avail}^2} \sum_{i=1}^N \text{UNE}_i, \quad (11)$$

where $\bar{f}_0$ represents a vector of starting true anomalies for each satellite in the orbit. In the objective function, UNE is assigned a penalty value of 2000 when coverage is not available; N is the total number of time steps in the iteration; and P_{avail} is the proportion of time in coverage.

This optimization was performed in MATLAB using the Nelder-Mead simplex method; such a direct search method was chosen due to the nonlinear and difficult-to-compute nature of the function gradient. Similar first-order methods, such as a conjugate gradient method computed using finite differences, fail to converge to a stationary point due to errors in the gradient approximation.

Six-Satellite Optimization

It was found that small changes to the initial guess affect optimization results greatly; to counteract this, a series of optimizations were conducted using a variety of starting locations. Table 6 displays the best-performing cases. It can be seen that Case 1 performs best when optimizing the function value, but Case 4 results in the best average PDOP, where gap time is not accounted for.

Table 6: Optimal starting true anomalies for the six-satellite constellation.

Case	$f_{0,1}$	$f_{0,2}$	$f_{0,3}$	$f_{0,4}$	$f_{0,5}$	$f_{0,6}$	$F(f_0)$	Avg. PDOP
1	0.00	178.64	223.41	79.99	186.32	211.72	195.84	10.86
2	170.14	148.04	257.22	0.00	128.47	184.67	199.18	10.78
3	96.88	186.96	271.75	134.49	201.04	226.31	199.27	9.19
4	111.94	190.23	215.71	44.83	180.89	231.62	200.28	8.65
Original	0.00	160.15	199.85	0.00	160.15	199.85	937.28	28.42

A comparison plot between the performance of the original constellation and Case 1 can be seen in Figure 13, with associated statistics in Table 7. Overall, the optimized case performs much better than the original. An interesting note is that the optimized case has a larger minimum UNE than the original, which may be attributed to the fact that the objective function targets average UNE rather than minimum.

Table 7: Statistical metrics for 3σ UNE of the six-satellite constellation.

Case	Coverage	Active UNE Mean	Active UNE Var.	Min. UNE	Max. UNE
1	99.9%	146.61 <i>m</i>	1630.30 <i>m</i> ²	53.84 <i>m</i>	960.31 <i>m</i>
Original	100%	223.89 <i>m</i>	5863.72 <i>m</i> ²	44.66 <i>m</i>	1487.97 <i>m</i>

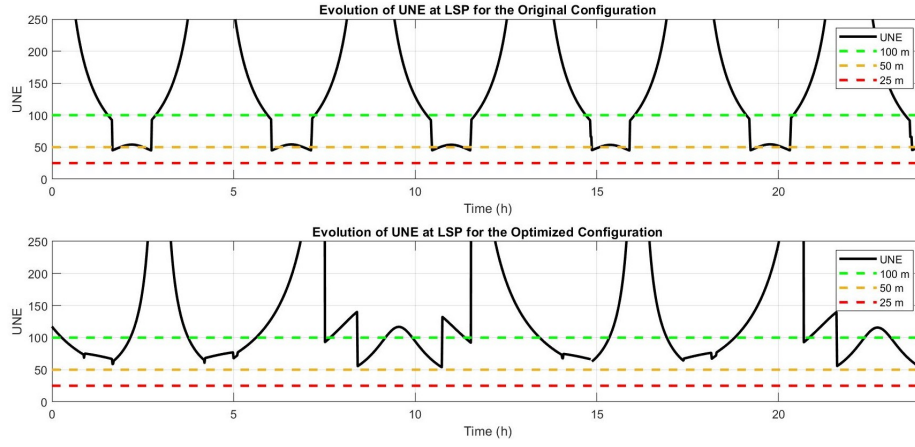


Figure 13: 3σ UNE at the lunar south pole for the original six-satellite case (top) and optimization Case 1 (bottom).

Eight-Satellite Optimization

This optimization was performed using the same process mentioned previously; notably, it takes many more iterations than the six-satellite case due to the increase in design variables. Table 8 displays the best-performing cases. All optimized cases outperform the original model, while Case 1 performs the best both in minimizing the objective function and having the lowest average PDOP; a marked improvement made by this case is achieving full coverage, which the original constellation failed to do.

Table 8: Optimal starting true anomalies for the eight-satellite constellation.

Case	$f_{0,1}$	$f_{0,2}$	$f_{0,3}$	$f_{0,4}$	$f_{0,5}$	$f_{0,6}$	$f_{0,7}$	$f_{0,8}$	$F(f_0)$	PDOP
1	43.7	183.34	152.58	216.87	42.9	152.22	182.82	217.22	76.11	3.53
2	39.5	151.30	183.05	217.49	43.7	216.50	182.73	152.23	76.15	3.54
3	0.00	148.61	179.93	212.49	0.00	147.51	212.42	179.71	76.17	3.64
4	43.9	147.29	181.94	212.14	0.01	151.47	179.74	216.24	77.03	3.61
Orig.	0	90	180	270	45	135	225	315	198.4	4.52

A comparison plot between the performance of the original constellation and Case 1 can be seen in Figure 14, with associated statistics in Table 9. Having complete coverage is a major improvement, allowing for an average UNE below the 100-meter threshold. The optimized case also reaches a minimum UNE of 38.5 meters, well below the intermediate threshold of 50 meters.

Table 9: Statistical metrics for 3σ UNE of eight six-satellite constellation.

Case	Coverage	Active UNE Mean	Active UNE Var.	Min. UNE	Max. UNE
1	100%	87.74 m	914.51 m ²	38.46 m	167.45 m
Original	84.5%	132.9 m	4098.89 m ²	56.49 m	287.85 m

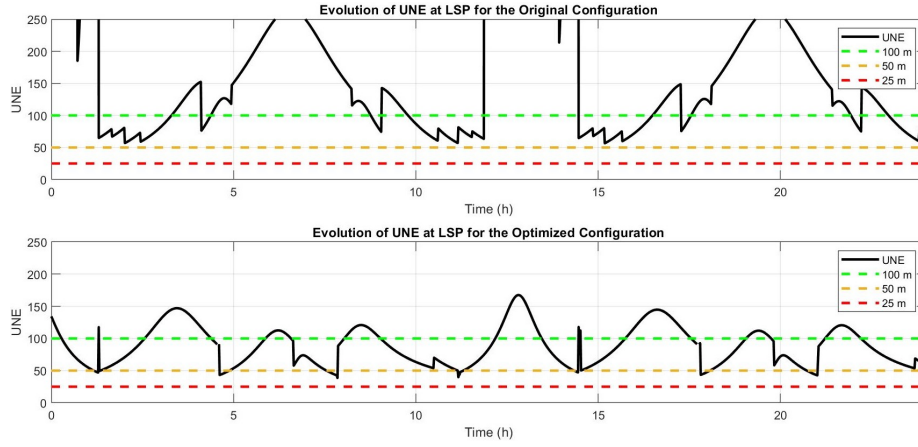


Figure 14: 3σ UNE at the lunar south pole for the original eight-satellite case (top) and optimization Case 1 (bottom).

NAVIGATION PERFORMANCE

For users navigating using these constellations, there will be a couple options depending on requirements for timeliness of solution acquisition. If navigation solutions are not required in real-time, users can employ delayed estimation methods like batch estimation. Batch estimation is most useful for stationary users or with well-characterized dynamical trajectories; in this case, real-time navigation solutions would not be required. The benefits of a batch estimator like a linear unbiased minimum variance estimator are the ability to achieve more precise solutions and a resilience to service outages. However, many users such as lunar orbiters or rovers have uncertain dynamics and a need for real-time navigation solutions. Batch estimators may be infeasible as precise solutions are not available immediately and it becomes inaccurate to compute state transition matrices to align all measurements with a single time. In these cases, users are best served using real-time kinematic fixes or implementing a form of Kalman filter.

This section will analyze two lunar use cases: first, that of a ground station or other stationary user; second, a lunar rover that is equipped with an accelerometer and driving around with a somewhat uncertain trajectory – the uncertainty being characterized by an acceleration standard deviation of $\sigma_{\text{Rover}} = 2 * 10^{-7} \text{ m/s}^2$. Perfect knowledge of the motion of the moon is assumed, and measurements are collected every 60 seconds for a period of 24 hours. Both scenarios will be gathering position measurements using each of the six navigation constellations discussed previously, comparing the results between each.

Ground Station

Figure 15 depicts 24-hour navigation simulations of a ground station placed at -75° latitude near the lunar south pole. This and Figure 16 show 3σ RMS uncertainty thresholds, derived from the linear covariance; these thresholds were verified with Monte Carlo runs but the runs were omitted for brevity. This stationary ground station is attempting to estimate its position on the lunar surface with no *a priori* information using a linear batch filter. Results are shown for each of the six constellations analyzed and are summarized in Table 10.

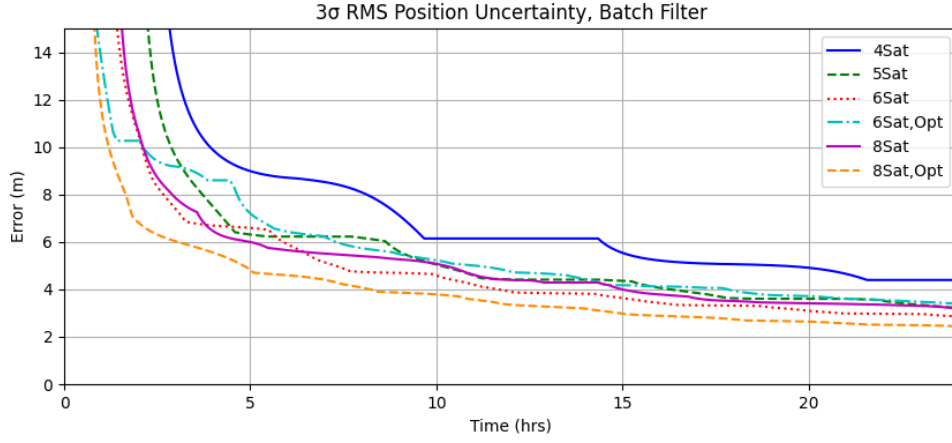


Figure 15: Ground station obtaining its position with a batch filter using various constellations.

Table 10: 3σ RSS position error for the ground station after 24 hours.

Constellation	3σ error, Batch filter
4-satellite	4.394 <i>m</i>
8-satellite	3.185 <i>m</i>
5-satellite	3.215 <i>m</i>
6-satellite	2.854 <i>m</i>
6-satellite (Optimized)	3.403 <i>m</i>
8-satellite (Optimized)	2.430 <i>m</i>

Generally, the constellations from literature are characterized by having few service outages accompanied by long periods of coverage. However, the navigation error during these periods is typically much higher than the novel constellations. The result is that even with more frequent service outages in the 5- and 6-satellite constellations, it results in generally faster convergence to a solution when the system dynamics are well-known. A large contributing factor to high-accuracy solutions is the user's access to low-UNE measurements – once a user converges to a solution within approximately 10 m, measurement accuracy becomes more important than quantity. This effect can be seen when comparing the six-satellite constellation's non-optimized and optimized configuration, where the original constellation achieves higher accuracy in a batch filter after 24 hours due to its lower minimum UNE (seen in Table 7). Achieving better performance would be a matter of reformulating the objective function in Equation (11) to prioritize periods of low UNE.

Ultimately, the optimized eight-satellite constellation achieves the best results because of its low minimum UNE (Table 9) and general availability of higher-quality measurements than the other constellations.

Lunar Rover

Figure 16 depicts 24-hour navigation simulations of a lunar surface rover following the same trajectory. This rover is driving around near the south pole and navigating using both the satellite constellation and an accelerometer in an extended Kalman filter.

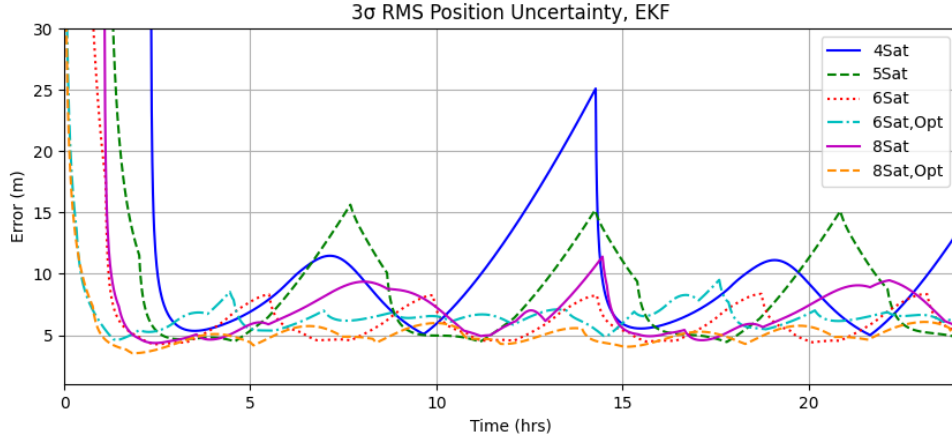


Figure 16: Lunar surface rover navigating with an EKF using various constellations.

In these plots, error spikes occur due to uncertainty in the dynamics as the rover navigates using its IMU while constellation measurements are not available. As such, the constellations from literature – characterized by more prolonged periods of outages and higher average dilution of precision during periods of coverage – perform worse than the novel constellations. This is because the gap time between coverage is longer than in the novel constellations and DOP is greater on average; while there may be more gaps in the five- and six-satellite constellations, their shorter nature allows the EKF to not spiral out of control as much and recover sooner once measurements are reacquired.

Comparing the original and optimized six-satellite constellation, higher peak uncertainty but fewer high peaks can be seen in the optimized version. This is due to the optimizer shrinking the gaps of very high UNE; however, it still suffers from having a higher minimum UNE than the original constellation. Looking at the eight-satellite constellations tells a different story though. The optimized version has few – if any – large peaks in uncertainty and the average error is 3-5 meters lower. This can be attributed to the optimized solution both eliminating gap time and reducing the average and minimum UNE over the interval.

CONCLUSIONS

This paper establishes a methodology for evaluating different minimal implementations of cislunar navigation constellations and provides a comparison of proposed designs on concrete evaluation metrics. The ability to optimize for certain design traits is successfully demonstrated, and navigation performance for representative users near the lunar south pole is simulated and discussed. This work provides a foundation for deployment of early-stage navigation constellations in the cislunar domain that aligns with strategic objectives such as LunaNet design recommendations. Subjects of future research are the inclusion of carrier-phase and range-rate measurements from the constellation for improved performance; reformulation of optimization problems to better target user needs, incorporating more design parameters than just true anomaly; considering the navigation performance of orbital users; and the evolution of this minimal architecture into a constellation with global PNT coverage and local service augmentations.

REFERENCES

- [1] M. Daniels and K. Bretl, “National Cislunar Science & Technology Strategy,” Nov. 2022.
- [2] K. Goodliff, N. Merancy, S. Bhakta, M. Rucker, P. R.-P. Chai, T. Ashurst, P. Troutman, and C. Stromgren, “Moon-to-Mars Architecture Definition Document,” Tech. Rep. ESDMD-001, NASA, Apr. 2023.
- [3] D. J. Israel, K. D. Mauldin, C. J. Roberts, J. W. Mitchell, A. A. Pulkkinen, L. V. D. Cooper, M. A. Johnson, S. D. Christe, and C. J. Gramling, “LunaNet: a Flexible and Extensible Lunar Exploration Communications and Navigation Infrastructure,” *2020 IEEE Aerospace Conference*, Big Sky, MT, USA, IEEE, Mar. 2020, pp. 1–14, 10.1109/AERO47225.2020.9172509.
- [4] P. Giordano, F. Malman, R. Swinden, P. Zoccarato, and J. Ventura-Traveset, “The Lunar Pathfinder PNT Experiment and Moonlight Navigation Service: The Future of Lunar Position, Navigation and Timing,” Long Beach, California, Feb. 2022, pp. 632–642, 10.33012/2022.18225.
- [5] F. Pereira and D. Selva, “Analysis of Navigation Performance with Lunar GNSS Evolution,” Long Beach, California, Feb. 2022, pp. 514–529, 10.33012/2022.18210.
- [6] T. A. Ely, “Stable Constellations of Frozen Elliptical Inclined Lunar Orbits,” *The Journal of the Astronautical Sciences*, Vol. 53, Sept. 2005, pp. 301–316, 10.1007/BF03546355.
- [7] D. Davis, B. Sagar, K. Howell, J.-W. Jang, R. Whitley, F. Clark, D. Guzzetti, E. Zimovan, and G. Barton, “Orbit Maintenance and Navigation of Human Spacecraft at Cislunar Near Rectilinear Halo Orbits,” *Proceedings of the 27th AAS/AIAA Space Flight Mechanics Meeting*, San Antonio, TX, Univelt, Feb. 2017.
- [8] P. Misra and P. Enge, *Global positioning system: signals, measurements, and performance*. Lincoln, Mass: Ganga-Jamuna Press, 2. ed ed., 2006.
- [9] D. Israel and N. Babu, “Draft LunaNet Interoperability Specification,” July 2022.
- [10] A. Flores and D. Nicholas, “NAVSTAR GPS Space Segment/Navigation User Interfaces,” Apr. 2021.
- [11] F. Moorefield, “Global Positioning System Standard Positioning Service Performance Standard,” Apr. 2020.
- [12] S. T. Hustell, W. G. Reid, L. J. D. Crum, L. H. S. Mobbs, and J. A. Buisson, “Operational Use of the Hadamard variance in GPS,” *Proceedings of the 28th Annual Precise Time and Time Interval Systems and Applications Meeting*, Reston, Virginia, Dec. 1996, pp. 201–214.
- [13] J. L. Small, L. M. Mann, J. M. Crenshaw, C. J. Gramling, J. J. Rosales, L. B. Winternitz, M. A. Hassounneh, D. A. Baker, S. Hur-Diaz, and A. J. Liounis, “Lunar Relay Onboard Navigation Performance and Effects on Lander Descent to Surface,” Long Beach, California, Feb. 2022, pp. 587–601, 10.33012/2022.18221.
- [14] N. Kubo, K. Kobayashi, and R. Furukawa, “GNSS Multipath Detection Using Continuous Time-Series C/N0,” *Sensors*, Vol. 20, July 2020, p. 4059, 10.3390/s20144059.
- [15] G. H. Heiken, D. S. McKay, and R. Brown, “Lunar deposits of possible pyroclastic origin,” *Geochimica et Cosmochimica Acta*, Vol. 38, Nov. 1974, pp. 1703–1718, 10.1016/0016-7037(74)90187-2.
- [16] T. Nie and P. Gurfil, “Lunar frozen orbits revisited,” *Celestial Mechanics and Dynamical Astronomy*, Vol. 130, Oct. 2018, p. 61, 10.1007/s10569-018-9858-0.
- [17] E. Zimovan, K. Howell, and D. C. Davis, “Near Rectilinear Halo Orbits and Their Application in Cislunar Space,” *3rd IAA Conference on Dynamics and Controls of Space Systems*, Vol. 20, Moscow, Russia, May 2017, p. 40.
- [18] S. Bhamidipati, T. Mina, and G. Gao, “A Case Study Analysis for Designing a Lunar Navigation Satellite System with Time-Transfer from Earth-GPS,” Long Beach, California, Feb. 2022, pp. 407–419, 10.33012/2022.18202.
- [19] M. Murata, I. Kawano, and S. Kogure, “Lunar Navigation Satellite System and Positioning Accuracy Evaluation,” Long Beach, California, Feb. 2022, pp. 582–586, 10.33012/2022.18220.
- [20] M. R. Thompson, A. Forsman, S. Chikine, B. C. Peters, T. Ely, D. Sorensen, J. Parker, and B. Cheetham, “Cislunar Navigation Technology Demonstrations on the CAPSTONE Mission,” Long Beach, California, Feb. 2022, pp. 471–484, 10.33012/2022.18208.
- [21] T. A. Ely and E. Lieb, “Constellations of elliptical inclined lunar orbits providing polar and global coverage,” *The Journal of the Astronautical Sciences*, Vol. 54, Mar. 2006, pp. 53–67, 10.1007/BF03256476.