NASA/TM–20205003520

Marshall Space Flight Center Faculty Fellowship Program

N.F. Six, Program Director
Marshall Space Flight Center, Huntsville, Alabama

G. Karr, Compiler
The University of Alabama in Huntsville, Huntsville, Alabama

August 2020

The NASA STI Program…in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results…even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>

- E-mail your question via the Internet to <help@sti.nasa.gov>

- Phone the NASA STI Help Desk at 757–864–9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681–2199, USA

NASA/TM–20205003520

Marshall Space Flight Center Faculty Fellowship Program

N.F. Six, Program Director
Marshall Space Flight Center, Huntsville, Alabama

G. Karr, Compiler
The University of Alabama in Huntsville, Huntsville, Alabama

National Aeronautics and
Space Administration

Marshall Space Flight Center • Huntsville, Alabama 35812

August 2020

Acknowledgments

All are grateful to those who, through their diligence, brought the 2019 NASA Marshall Space Flight Center Faculty Fellowship program to completion. These professionals include Jody Singer, Larry Leopard, John Honeycutt, Steve Cash, Frank Six, Gerald Karr, Brooke Graham, Debora Nielson, Judy Drinnon, and Tammy Rowan.

TRADEMARKS

Available from:

NASA STI Information Desk
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681–2199, USA
757–864–9658

This report is also available in electronic form at
<http://www.sti.nasa.gov>

EXECUTIVE SUMMARY

The 2019 Marshall Faculty Fellowship Program involved 14 faculty in the laboratories and departments at Marshall Space Flight Center and one faculty researcher working from Colorado. These faculty engineers and scientists worked with NASA collaborators on NASA projects, bringing new perspectives and solutions to bear. This Technical Memorandum is a compilation of the research reports of the 2019 Marshall Faculty Fellowship program, along with the Program Announcement (Appendix A) and the Program Description (Appendix B). The research affected the following five areas:

(1) Materials
(2) Propulsion
(3) Spacecraft systems
(4) Vehicle systems
(5) Space science.

The materials investigations include Lunar Regolith for habitats, friction stir welding, and composite joints. Propulsion studies include cryogenic tank pressurization, transmitted torque in a cryogenic environment, and condensation in presence of noncondensables, Europa Lander Deorbit Stage, and catalyst development for a hybrid rocket. Spacecraft systems include wireless sensor networks and printed electronic inks. Vehicle systems studies were performed on Mars ascent vehicle analysis, architecture models, and Space Launch System manual steering. Space science studies include planetary lava flow. Our goal is to continue the Marshall Faculty Fellowship Program funded by Center internal project offices. Faculty Fellows in this 2019 program represented the following minority-serving institutions: Alabama A&M University, Southern University, Delgado Community College, and Dillard University.

2019 Marshall Space Flight Center Faculty Fellowship

From Left to Right:

Front Row—Frank Six, Tammy Winner, Charles Yang, Murphy Stratton, Tomekia Simeon, Bhaba Sarker, Alak Bandyopadhyay

Middle Row—Debora Nielson, Todd Lillian, Maria D. Cortes-Delgado, Stephen Whitmore, Juan Lorenzo, Md Abdus Salam, Brooke Graham

Back Row—Eugeniy Mikhailov, Sivaguru Ravindran, Gerald Karr

Not pictured: Jack Van Natta

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Multi-node Modeling of Autogenous Pressurization of a Cryogenic Propellant Tank with Simultaneous Injection and Venting

Alak Bandyopadhyay[1]
Alabama A & M University, Normal, AL 35762

Alok Majumdar[2]
and
Mark Rogers[2]

I. ABSTRACT

This paper presents a multi-node model of autogenous pressurization of liquid nitrogen in a flight tank using the Generalized Fluid System Simulation Program (GFSSP), a general purpose flow network code developed at NASA/Marshall Space Flight Center. This study considers two different models: (a) pressurization of the tank by nitrogen gas from a supply tank and (b) pressurization of the tank by using the boil-off propellant through an IVF (Integrated Vehicle Fluids) loop. The tank pressurization model considers two different liquid fill-levels of the tank (a) 75% and (b) 45%. Heat and mass transfer between the liquid and vapor has been modeled at the liquid vapor interface. Heat transfer between wall and vapor at the ullage has been accounted for by assuming heat transfer occurs by natural convection. The model also accounts for heat leak to the tank through the insulation and metal wall by heat conduction. The predicted pressures and temperatures are compared with the measured data and good agreement has been observed. In the second case, the tank is pressurized by using the boil-off fluid that recirculates through an IVF loop. Preliminary results from this model are also reported.

II. Introduction

Cryogenic Tanks are pressurized by inert gas such as Helium or Nitrogen to maintain the required pressure of the propellant delivered to the turbo-pump of a liquid rocket engine. Thermo-fluid system simulation tools are used to analyze the pressurization process of a cryogenic tank. Most system level codes (GFSSP and ROCETS) use a single node[1] to represent ullage which is the gaseous space in the tank. Ullage space in a cryogenic tank is highly stratified because the entering inert gas is at ambient temperature whereas the liquid propellant is at a cryogenic temperature. A single node model does not account for the effect of temperature gradient in the ullage. High fidelity Navier-Stokes based CFD model of Tank Pressurization is not practical for running a long duration transient model with thousands or millions of nodes. A possible recourse is to construct a multi-node model with system level code that can account for ullage stratification with conjugate heat transfer.

For the past few years, United Launch Alliance has been developing a propulsion system called Integrated Vehicle Fluids (IVF)[2] to improve the functional and reliability limits of upper stages for long-duration space missions. IVF uses boil-off propellants to drive thrusters for the reaction control system as well as to run small internal combustion engines (ICEs). The produced thrust is used for maneuvering the vehicle and to settle propellants during coast flight. Figure 1 shows a simplified schematic of the IVF system including the propellant tank and a fluid loop consisting of a compressor and heat exchanger instead of a helium tank in a conventional propulsion system. The compressor intakes propellant vapor from the tank ullage and drives it through a heat exchanger to heat it before it sends it back to the tank for pressurization. The heat exchanger receives heat from coolant of the ICE. The ICE provides power to the compressor and battery. The network flow solver program GFSSP[3] has been used to model the heat exchanger component and the complete IVF system by using one dimensional model (changing only in the tank axial direction) for temperature and pressure by Leclair et.al.[5] and Majumdar et.al.[4]. However both these models are unable to see any two dimensional effect within the tank.

The objective of the current study is to develop a multi-node computational model to simulate the pressurization of the tank due to a) the propellant injection from the top of the tank and simultaneous venting of ullage gas and (b) pressurization using a pressurization loop consisting of a blower and heat exchanger. In the loop model, the compressor is substituted with a blower and heat generated by the motor that runs the blower is added into the fluid. The test data are available with liquid Nitrogen. The model also considers the conjugate heat transfer to estimate the heat leak.

---

[1] Associate Professor, Electrical Engineering and Computer Science, Alabama A & M University.
[2] MSFC Summer Mentor and Collaborator

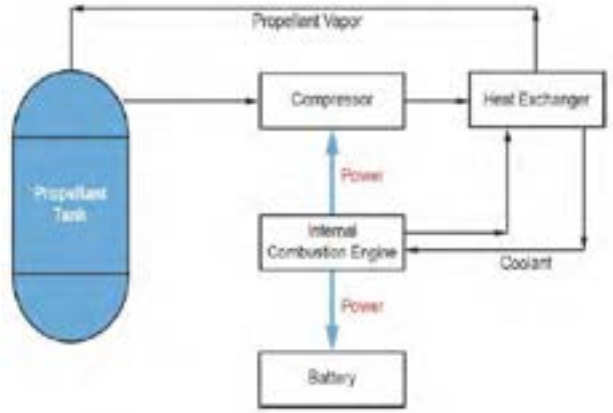


Figure 1. **Simplified Schematic of IVF System**

The loop model developed by Leclair et.al[5] has been integrated with the tank model developed in the first part of this work and the simulation has been carried out in the time range the test data is available.

III. Mathematical and Computational Model

In the present study, the tank model is discussed first as that is common to both problems discussed above. Once the tank model is established and compared with the test data, the IVF loop model is integrated for the second phase of this work.

1. Tank Pressurization due to propellant injection from supply line

For the tank model, the flow and heat transfer within the ullage space is considered along with the conjugate heat transfer between tank wall and the ullage. The interaction between ullage to liquid is modeled through the heat and mass transfer equations at the interface as described later in this section. Two different fill-levels of the tank are considered for the model: (a) tank is initially filled with 75% liquid (by volume) and (b) with 45% liquid (by volume). The ullage space is assumed to be filled with vapor. All the operating conditions including initial state of the ullage, the injector pressure and temperature conditions and vent valve operating conditions are taken from the experimental data. In this section, the computational model developed using GFSSP is described, followed by the heat and mass transfer model at the liquid-ullage interface and heat transfer between tank wall to the ullage space.

Computational Model Using GFSSP:

The entire ullage space is uniformly divided into 5 segments along the axial direction and uniformly divided into 5 segments in the radial direction. Nodes are placed at the center of each cell formed by this division. Figure 2 shows below the multi-node model of GFSSP with a total of 25 fluid nodes, 20 solid nodes and 44 branches. The mass and energy conservation equations in conjunction with the equation of state for a real fluid are solved in fluid nodes. The momentum equations of the fluid are solved in the branches. The energy conservation equations are solved in the solid nodes. The system of equations are solved by a hybrid numerical method[2] which is a combination of simultaneous Newton-Raphson method and successive substitution method.

Nodes 1 through 25 are representing the fluid nodes in the ullage space and nodes 28 through 46 (as shown with solid border line) represent the tank wall with two different layers for the metal and insulation as indicated. Nodes 1 through 5 are the ullage nodes closest to the liquid surface, nodes 5, 10, 15, 20 and 25 are close to tank walls, nodes 1, 6, 11, 16 and 21 are along the center line of the tank. The model is assumed to be axisymmetric. Node 26 represents the liquid node that contains the entire liquid mass of the tank. Node 27 is a boundary node representing the tank drain outlet. The injector flow comes in to the tank from the top of the tank and node 30 represents the inlet boundary to the tank, and similarly node 51 represent the exit boundary from the tank through which the fluid is vented out. The vent valve is modeled using a restriction option in GFSSP and the valve open-close area history with time is according to the test setup. Figure 3 shows the vent valve opening profile as a function of time for the (a) 75%-fill test study and (b) 45%-fill test study.

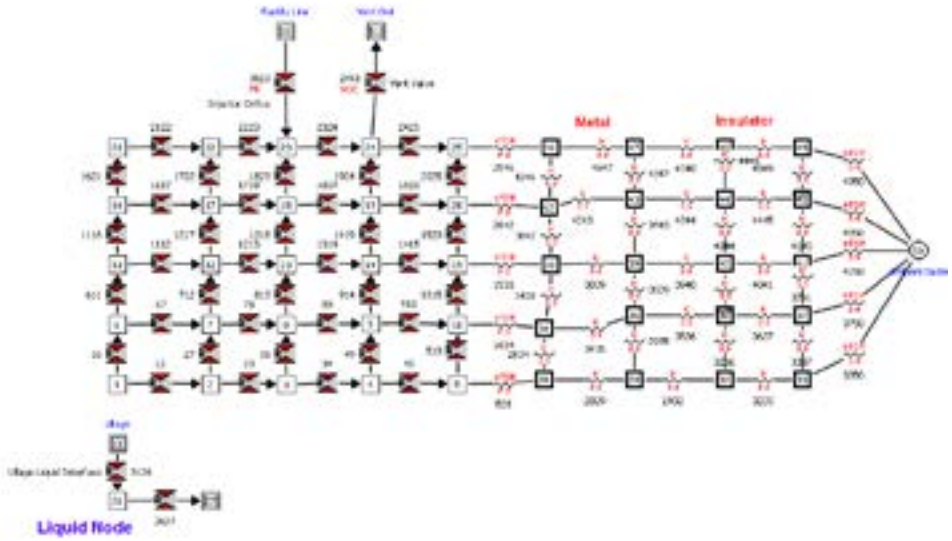
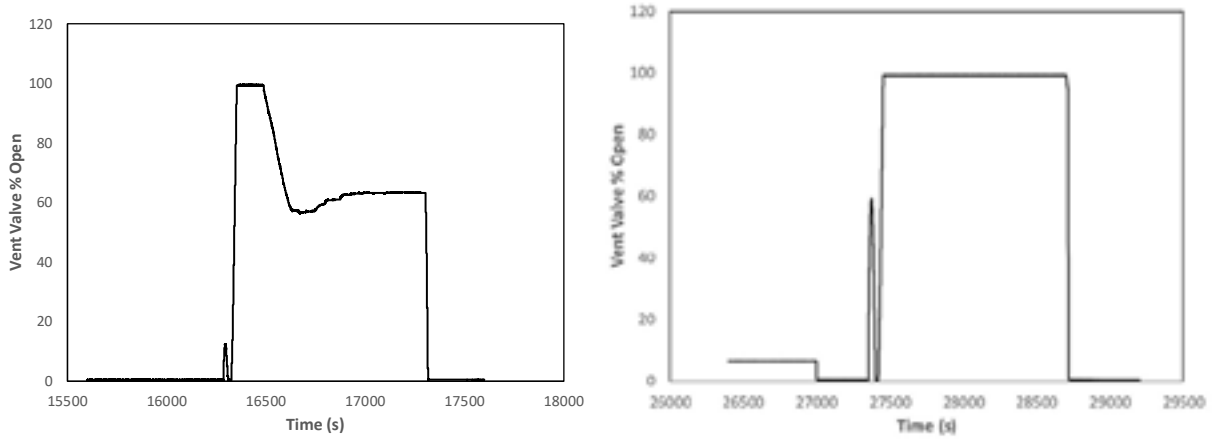


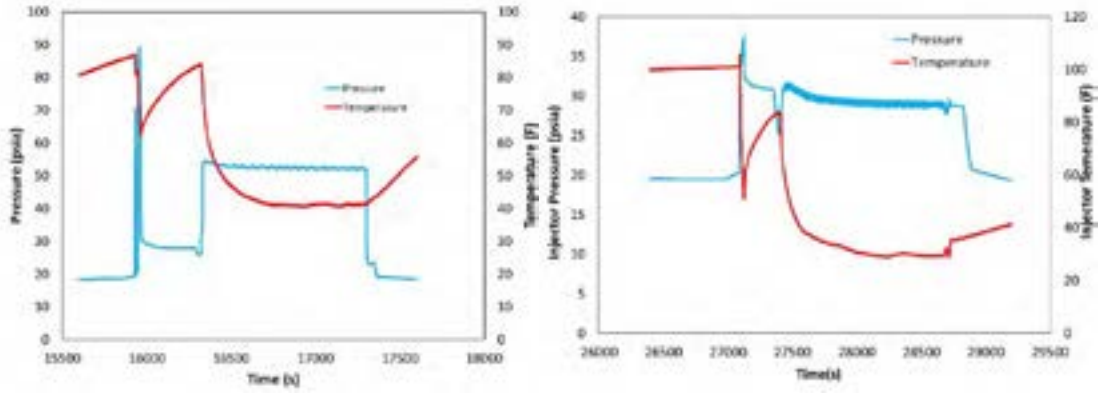
Figure 2. Multi-node Model of the Propellant Tank.



(a) (b)

Figure 3. Vent Valve percentage Open with time.

Node 30 in the GFSSP model (figure 2) represents the injector inlet to the tank through a valve with pressure and temperature conditions taken from the test data. The pressure and temperature inlet conditions for the injector flow are shown in figures 4(a) and 4(b) for 75% fill case and 45% fill case respectively.



(a) (b)

Figure 4: Propellant Injector Pressure and Temperature Test Data for (a) 75% and (b) 45% fill cases

Heat Transfer between Tank-wall to ullage

The heat transfer between the tank-wall and the ullage space is modeled in the current study by considering (a) heat transfer due to natural convection at the wall to ullage interface and (b) heat conduction in the tank wall with convective boundary condition at the external wall.

The heat transfer coefficient between the wall and ullage was computed from a natural convection correlation for a vertical plate[6]. The following set of equations was used for this correlation:

$$\mathrm{Nu} = [(\mathrm{Nu}_l)^m + (\mathrm{Nu}_t)^m]^{1/m} \quad m = 6 \tag{1}$$

$$\mathrm{Nu}_t = C_t^{\prime}\mathrm{Ra}^{1/3}/(1 + 1.4 \times 10^9 \mathrm{Pr}/\mathrm{Ra}) \tag{2}$$

$$\mathrm{Nu}_l = 2/\ln(1 + 2/\mathrm{Nu}^T) \tag{3}$$

$$\mathrm{Nu}^T = \bar{C}_l \mathrm{Ra}^{1/4}, \tag{4}$$

$$C_t^{\prime} = \frac{0.13\mathrm{Pr}^{0.22}}{\left(1 + 0.61\mathrm{Pr}^{0.81}\right)^{0.42}}, \tag{5}$$

and

$$\bar{C}_l = \frac{0.671}{\left[1 + \left(\frac{0.492}{\mathrm{Pr}}\right)^{9/16}\right]^{4/9}}, \tag{6}$$

Where Gr = Grashof number = $L^2\rho^2 g\beta\Delta T/\mu^2$

Pr = Prandtl number = $\mu C_P/k$

Ra = GrPr

Nu = hL/k. Subscripts t and l refer to turbulent and laminar, respectively.

Liquid-Ullage Heat and Mass Transfer Model for Self-Pressurization[7]

Figure 5 shows the schematic of ullage and liquid propellant where there is heat transfer between the ullage and the liquid propellant that also results in evaporative mass transfer.

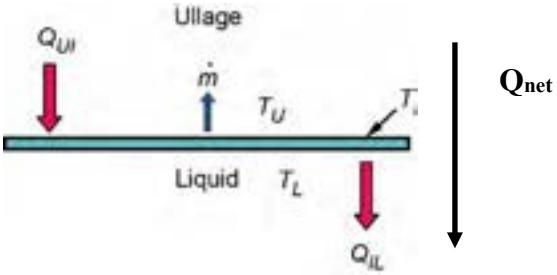


Figure 5. Evaporative Heat and Mass Transfer at liquid-vapor interface.

In this evaporative mass transfer model, a saturated layer is assumed at the interface between liquid and vapor so that $T_I = T_{sat}(P_v)$, where P_v is propellant vapor pressure in the ullage. The saturated layer receives heat from the ullage (Q_{UI}) and also rejects heat to the liquid (Q_{IL}). The difference in this heat rate contributes to the mass transfer in accordance with the law of energy conservation. The equations governing this process are as follows:
a) Heat transfer from ullage to interface layer:

$$Q_{UI}=h_{UI}A(T_U\ T_I) \tag{7}$$

b)Heat transfer from interface to liquid:

$$Q_{IL} = h_{IL}A(T_I - T_L)\ . \tag{8}$$

The evaporative mass transfer is expressed as

$$\dot{m} = \frac{Q_{UI} - Q_{IL}}{h_{fg}}\ . \tag{9}$$

h_{fg} is the enthalpy of evaporation, and the heat transfer coefficients h_{UI} and h_{IL} are computed from natural convection correlations given by:

$$h_{UI} = K_H C \frac{k_f}{L_c} \mathrm{Ra}^n = h_{IL}\ . \tag{10}$$

Where $C = 0.27$, and $n = 0.25$, K_H is a correction factor and was set to 0.5 to match the measured boil-off rate.

The heat transfer coefficient is assumed to be the same on both sides of the liquid vapor interface.

The net heat transfer rate from ullage to the liquid is given as below, and this has been used as a heat sink in the energy equation for the fluid in the ullage adjacent to the liquid node.

$$\dot{Q}_{net} = \dot{m}\left[C_{P,I}\left(T_I - T_L\right) + h_{fg}\right] \tag{11}$$

2. Integrated Model of Tank Pressurization by using closed IVF Loop Model

The second model consists of the tank model described in previous section integrated with a closed loop IVF model representing the blower and the heat exchanger. The IVF loop model of Leclair et. al.[5] has been shown in figure 6 below.

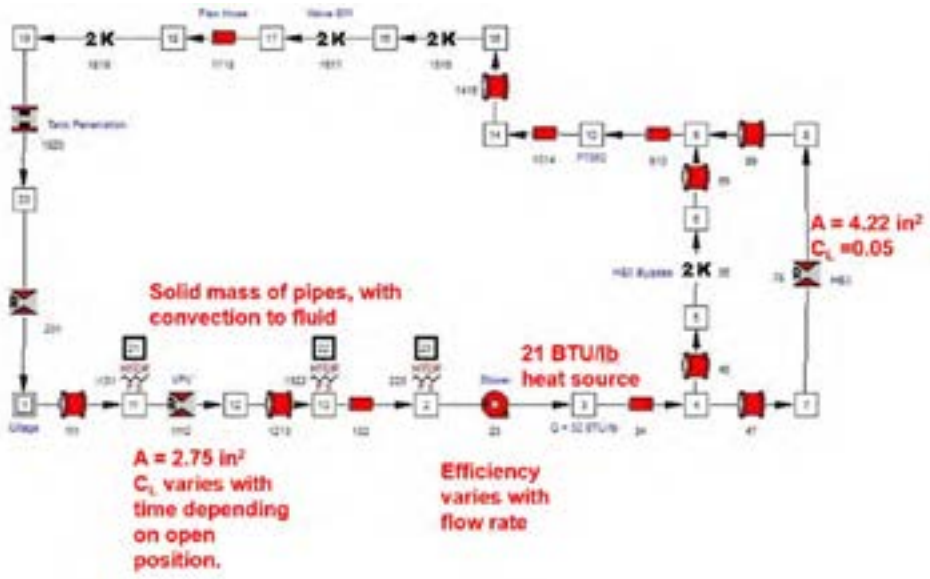


Figure 6. IVF Loop Model (Courtesy: Leclair et. al. [5])

This model shows two parallel paths for the flow to go: (a) through the heat exchanger (HEX) and (b) Heat Exchanger by-pass. The IVF test at NASA Marshall Propulsion Test Facility show three different testing: (a)Astros when Tank is pressurized by Facility Supply; open loop with vent gas exit to ambient, (b) Braves, when the Tank Pressurization is by closed loop bypassing the Heat Exchanger and (c) Cubs, when the Tank Pressurization is by closed loop with heat exchangers. For the current study, it is assumed that the flow is completely through the heat exchanger.

In this study, the integrated model consists of the tank and the IVF loop as shown in figure 7 below.

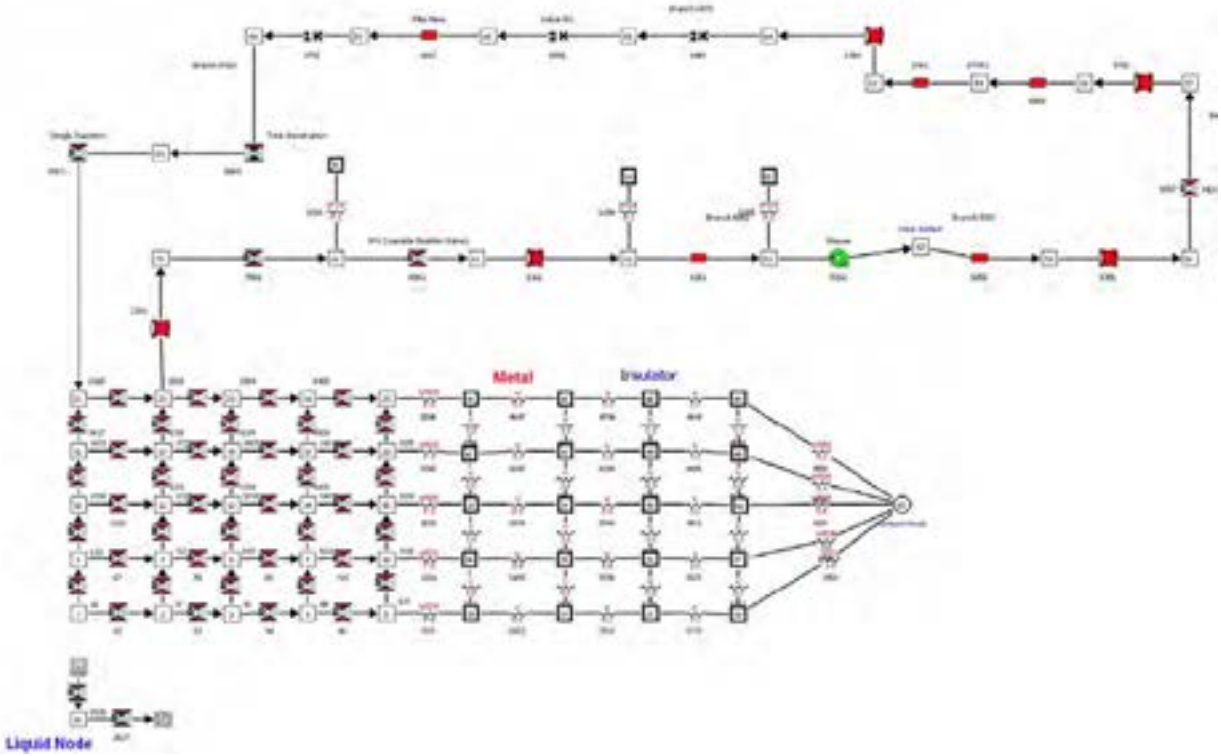


Figure 7. Integrated Tank-IVF Loop Model

The tank is pressurized due to continuous injection of hot gaseous Nitrogen that picks up heat in the blower and also in the heat exchanger. The motor is run with 25% power to drive the blower, and the blower efficiency is calibrated from steady state model to match the flowrate under steady state condition. It has been found that blower efficiency does not vary much as shown from the table given below. This tabulated data are generated from the simulation of a steady state model by providing the pressure and temperature at the blower input, and the pressure at the loop outlet (node 20), and the efficiency is varied to match the flow rate. For this purpose at three different times of the test data, the steady state loop model is run using GFSSP by adjusting the blower efficiency so as to match the flow rate and the results are as illustrated in the table below.

Time (s)	Pressure (blower Inlet, psia)	Pressure (loop exit)	Temp (blower Inlet, F)	Adjusted Efficiency (%)	Measured Flow Rate (lb/s)	Predicted Flowrate (lb/s)	Heat Added (Btu/lb)	Temp (measured at blower exit, F)	Temp (predicted At blower Exit, F)
20309	22.8	23.8	-129.2	5.5	0.054	0.055	20.4	-40.2	-40.4
20709	30.2	31.54	-135.4	7.2	0.071	0.072	23.6	-32.9	-32.8
21200	29.2	33.1	-112.4	6.5	0.066	0.066	20.45	-22.7	-22.5

Table 1. Calibration of Blower Using Steady State Model

As the efficiency does not vary much with flow rate, for the integrated model simulation, the blower efficiency is assumed to be constant at 7%. The heat rates are adjusted in the integrated model so as to match the flow rate through the loop as observed in test. The VPV (Variable Position Valve) is open during a short period as shown in figure 8 below.

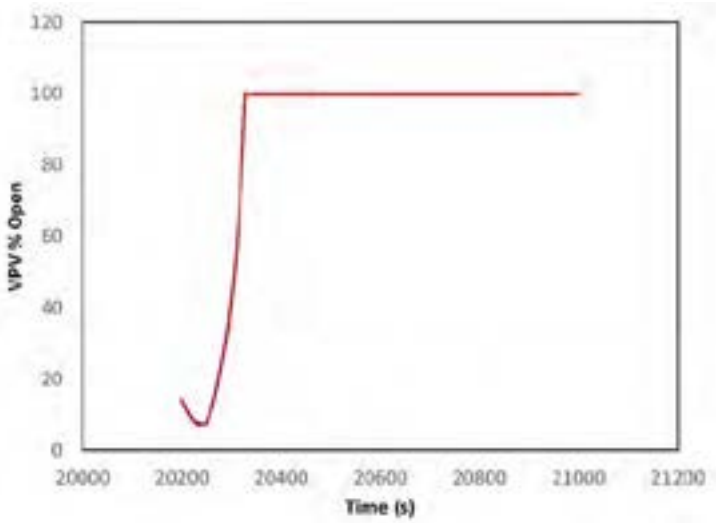


Figure 8. Variable Position Valve Percent Open as a function of time,

All the other operating conditions and initial conditions are taken from available test data.

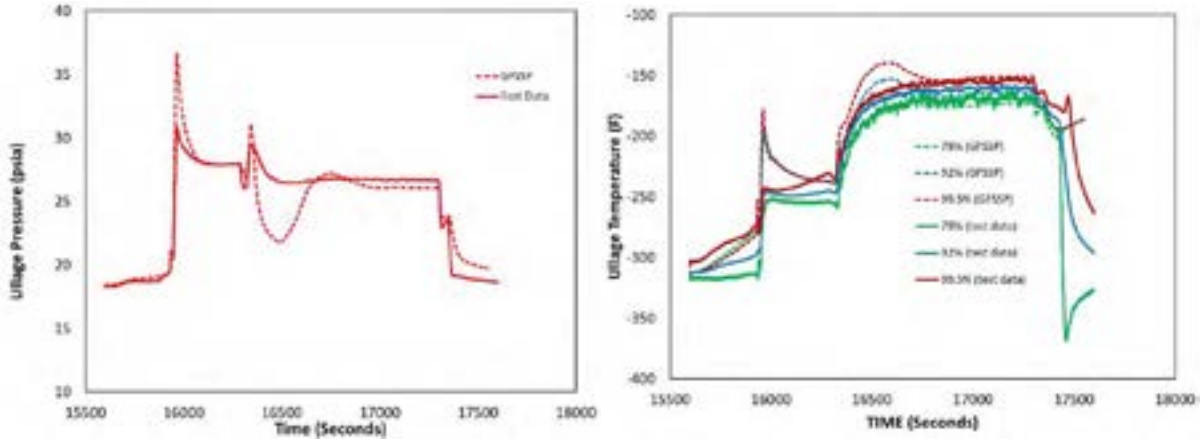
IV. Results and Discussion

The computed results for the tank pressurization due to propellant injection are presented first. Two different fill levels are considered for the simulation: (a) 75% Fill-level – 75% of the tank volume is initially filled with liquid propellant (in the current study liquid Nitrogen is used) and (b) 45% Fill-level – 45% of the total volume is filled with liquid propellant. The inputs for the simulations such as the vent flow valve open-close time variations and the injector pressure and temperature variations with time, are identical to test conditions and have been shown in figures 3 and 4 respectively in the previous section. The simulation is run for about 2000 seconds for the 75% fill-level case and about 2800 seconds for the 45% case. The test times show that the 75% case is tested first and then the 45% fill-level. Hence the modeling results are presented in similar way. The governing equations of mass, momentum and energy are solved

by marching in time with a step size of 0.01 second. The results are converged with an accuracy of 1 x 10^{-4}. The computational time in an Intel core i7-7820HQ processor with 16 GB RAM was 1.32 CPU second for 1 second of simulation time.

Case 1. Pressurization due to Propellant Injection, 75% Fill-Level

The computational results of ullage pressure and temperature from GFSSP simulations are compared with the test data and are as shown in figures 9(a) and 9(b) below.



(a) (b)

Figure 9. Comparison of computed data with test data for (a) ullage pressure and (b) ullage temperature.

The axial locations (measured from bottom of tank) are selected as per the node locations in the GFSSP model and the percentage indicates the volume of the tank as compared to total volume. Hence 79% means an axial position (from bottom of tank) that covers 79% of the total tank volume. The solid lines indicate the test data and the dotted lines indicate the numerical results from GFSSP. Good agreement is shown in ullage pressure prediction while comparing with the test data except the time zone when the vent valve opens 100% at a very fast rate as shown in vent-valve open history (figure 3a). The ullage temperature plot (figure 8b) shows the comparison of ullage temperature at three different locations in the ullage space: close to the liquid interface (79%), at the midpoint (92%) and at the top of the tank (99.5%). The GFSSP modeling shows higher ullage temperature corresponding to the test data in the time zone where the injector pressure and temperature were varying abruptly (see figure 6). When the injector pressure is relatively steady, agreement between the predictions and measurements is better. It shows about 20 degrees F stratification in the ullage space, which agrees with the test data. The wall temperature at various locations are also compared with the test data in figure 10. The test data were available only at the top of the tank and therefore the results are shown at 97% and 99.5% (top two nodal points in the model). The dotted lines correspond to modeling results and solid lines correspond to test data.

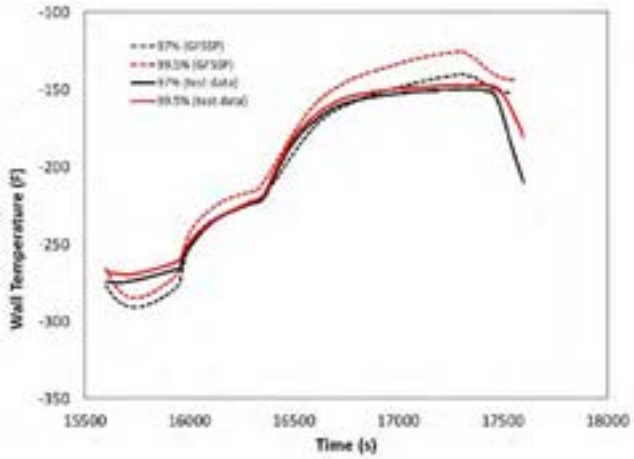


Figure 10. Internal wall temperature as a function of time (modeling vs test data)

The total heat transfer from the tank wall to the ullage space is computed and plotted as a function of time as shown in figure 11. In the steady state region, the computed heat transfer rate is about 2000 Watts. Negative Q implies that wall is heated by the ullage gas which happens during rapid pressurization.

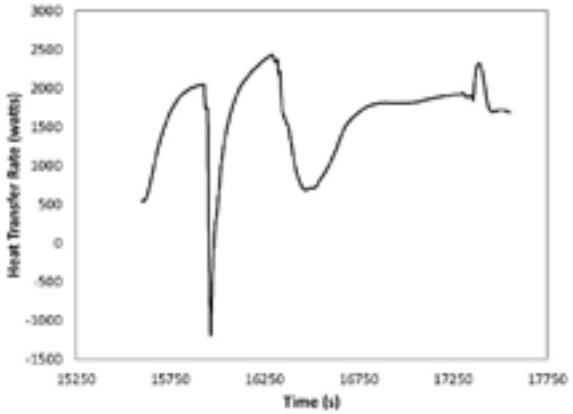


Figure 11. Heat Transfer rate from tank wall to ullage.

The heat transfer coefficient at the inner surface of the tank wall and ullage at three different axial locations (along the tank height) are shown in the figure 12.

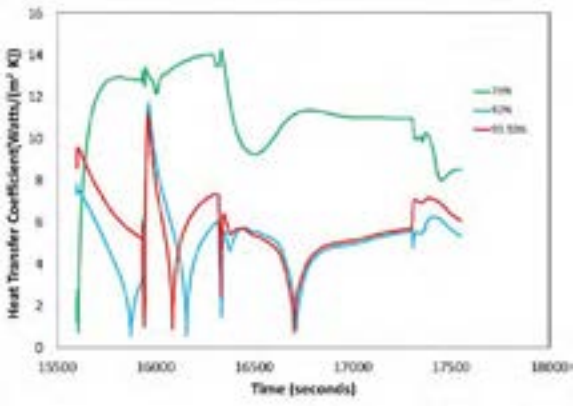


Figure 12. Computed heat transfer coefficient between the ullage and tank wall

Figure 13 shows the streamline trace and the temperature contours (at time = 17000 seconds, steady state region). The domain shown is the representation of the ullage space (tank top domain). The radial direction is along the X axis and tank height is along the Y axis. As expected the warmer region is at the top of the tank. The velocity vector near the top domain is due to the complex interaction of injector flow coming in and the vent flow going out. Increasing more number of grids near the top could have shown a better representation of the velocity vector plot, as well as temperature contours.

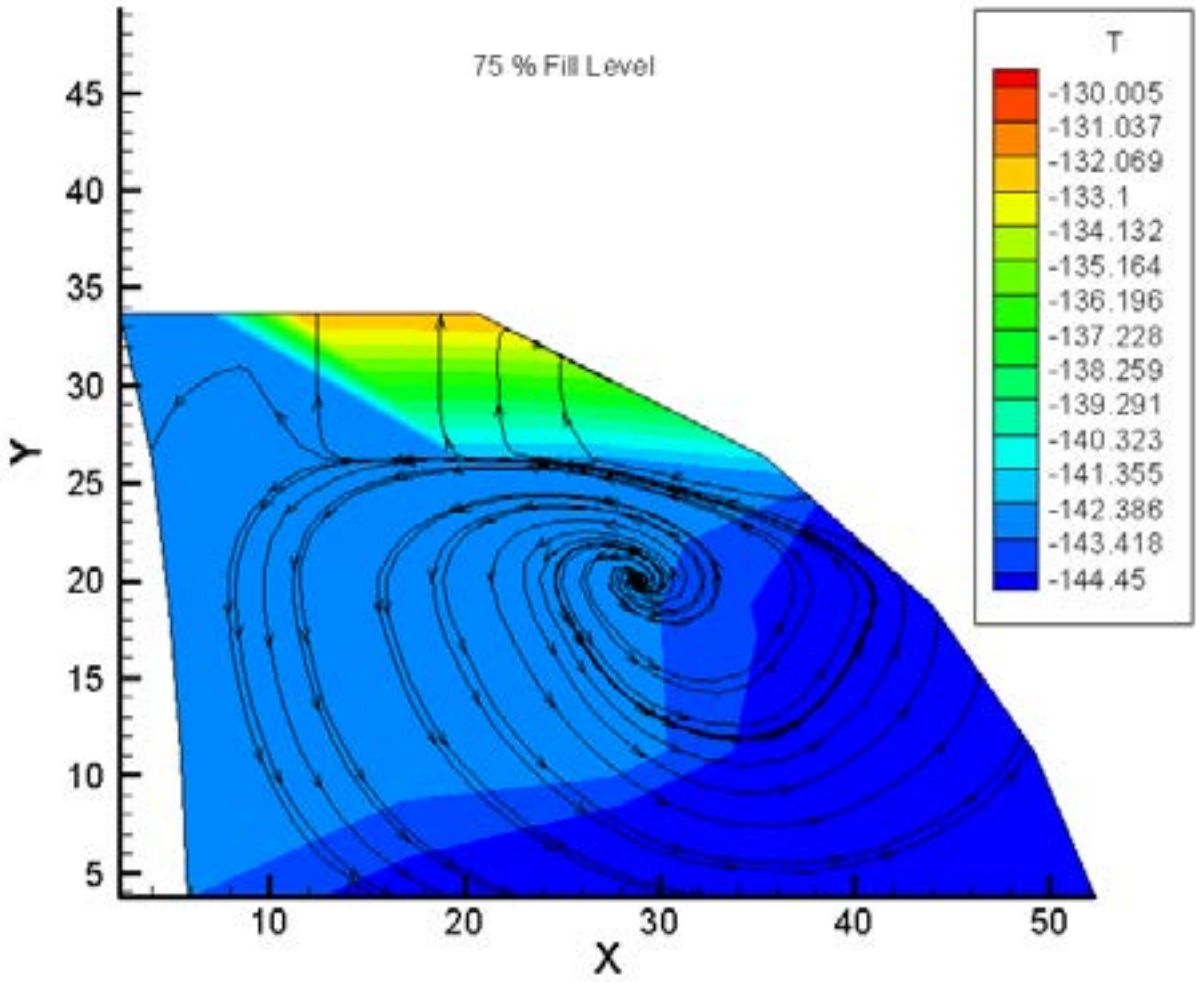


Figure 13. Temperature contour and Stream Traces at time = 17000 s.

Case 2. . Pressurization due to Propellant Injection, 45% Fill-Level

In the second test case, with 45% initially filled tank, the computed results from GFSSP are compared with test data for ullage pressure and ullage temperature as shown in figure 14 below.

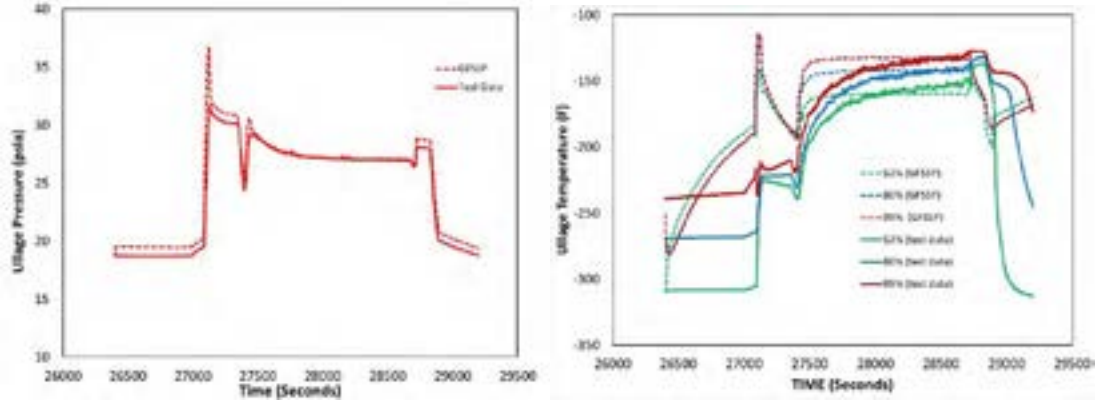


Figure 14. Comparison of computed data with test data for (a) ullage pressure and (b) ullage temperature.

Good agreement of the ullage pressure between model data and test data has been observed in figure 14. The ullage temperature also shows similar behavior in the steady state region.

The boil off of the propellant is computed as the amount of mass evaporated from the liquid surface and this has been plotted as a function of time in the figure 15 given below. For the test data, this value is computed as the difference between flow vented out from the ullage space and the injector flow coming in during steady state operation.

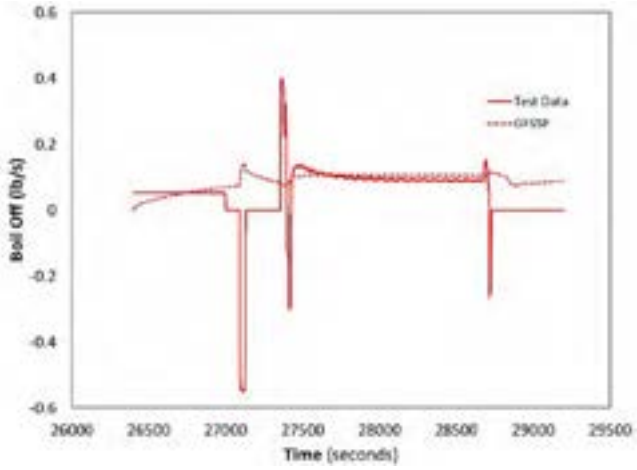


Figure 15: Computed and test data for boil-off.

The computed results compare well with the test data in the steady-state region when the vent valve is fully open. However, during the sudden opening and closing of the valve there are discrepancies.

The temperature contour and streamline traces for the 45%-fill case have been shown in figure 16.

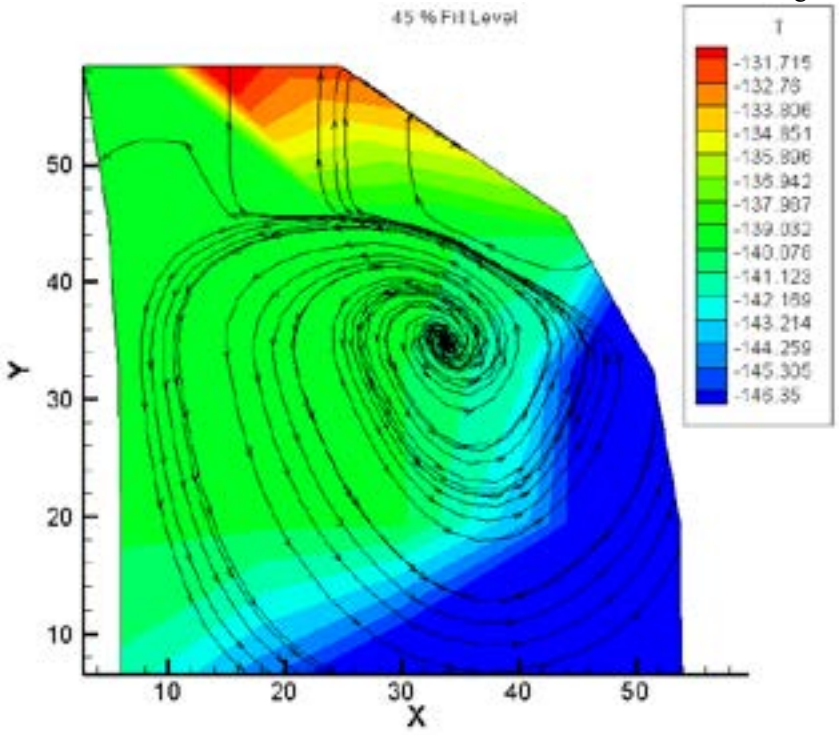


Figure 16. Temperature contour and Stream Traces at time = 28000 s.

Case 3: Results from the Integrated Model

The integrated model as illustrated in figure 8 earlier, has been used to compute the ullage pressure, temperature and flow rates with the operating conditions (initial conditions and VPV valve open according to figure 8). The computed values of ullage pressure are compared with the test data are as shown in figure 17 below.

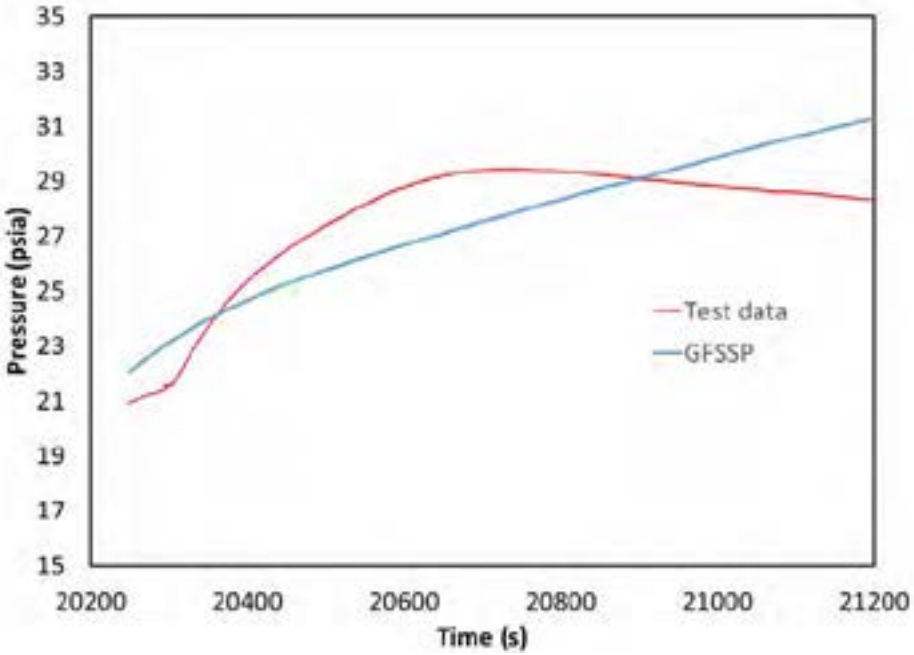


Figure 17. Ullage Pressure (Intgrated Model Computed Data and Test Data)

The corresponding flow rate through the blower is shown in figure 18 below.

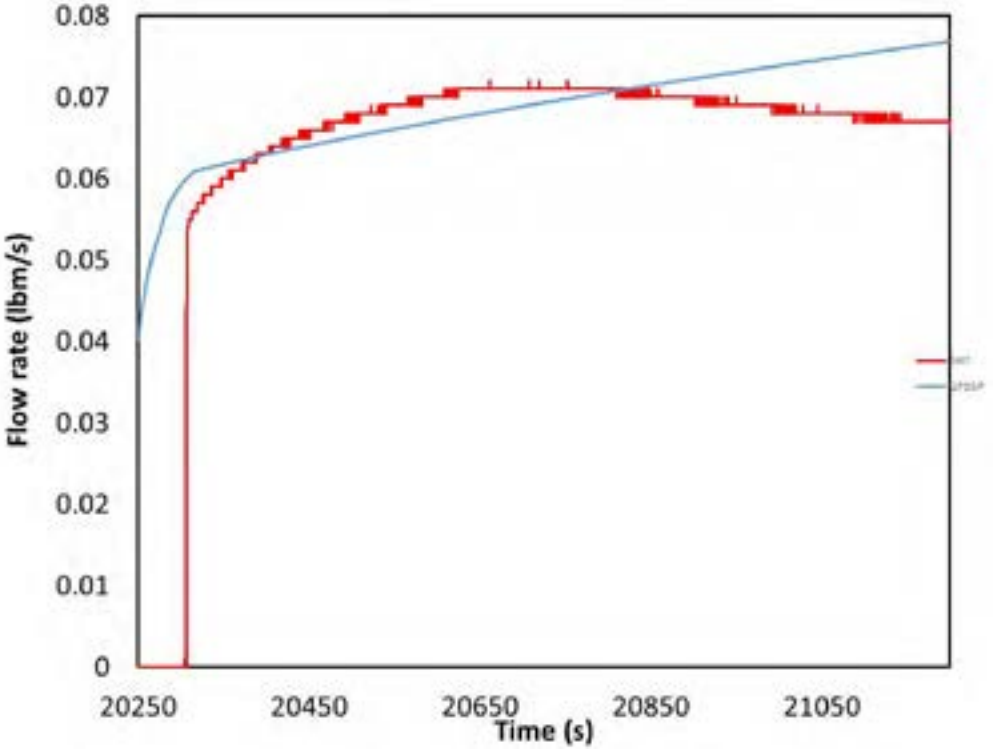


Figure 18. Flow Rate (Integrated Model Computed Data and Test Data)

The difference between the model (GFSSP) data and test data can be due to uncertainty of heat transfer correlation between wall and ullage. In the current model, natural convection correlations have been used. However, due to strong recirculation, heat transfer due to convection and/or mixed convection correlations might improve the computed solution. This work will be attempted in near future.

V. Conclusions

This paper demonstrates the feasibility of system level modeling of tank pressurization using multiple nodes. The ullage of a flight tank has been modeled using 25 nodes and 40 branches where mass and energy conservation equations were solved at the nodes and momentum equations are solved at the branches. Gravity, heat and mass transfer at the liquid vapor interface, and heat transfer between solid and fluid are accounted for in the governing equations. The model results have been verified by comparing with test data. The advantage of using multiple nodes in a system level code is that it allows prediction of recirculation and stratification with a fraction of the computational cost of a high fidelity Navier-Stokes code. The integrated model of the tank and closed loop needs further investigation.

Acknowledgments

The authors would like to acknowledge Dr. Robert J. Kenny/ER43 for supporting the fellowship awarded to Dr. Alak Bandyopadhyay under the NASA Summer Faculty program, Mr. Arthur H. Werkheiser/ST23 for making available the test data for model verification, Dr. Andre C. LeClair/ER43 and Mr. Juan G. Valenzuela/ER24 for many helpful discussions during the course of this work. The work was conducted at ER43/Thermal Analysis Branch in NASA/Marshall Space Flight Center.

References

[1]Majumdar, Alok and Steadman, Todd, "Numerical Modeling of Pressurization of a Propellant Tank", Journal of Propulsion and Power, Vol 17, No.2. March – April, 2001.

[2]Frank Zegler, "An Integrated Vehicle Propulsion and Power System for Long Duration Cryogen Spaceflight" , AIAA 2011-7355, AIAA SPACE 2011

[3]Majumdar, A. K., LeClair, A. C., Moore, R., and Schallhorn, P. A., "Generalized Fluid System Simulation Program, Version 6.0," , *NASA/TP—2016–218218*, March 2016.

[4]Majumdar, A.K, LeClair, A.C. and Hedayat, A., "Numerical Modeling of Pressurization of Cryogenic Propellant Tank for Integrated Vehicle Fluid System", AIAA Propulsion and Engineering Forum, 52nd *AIAA/SAE/ASEE Joint Propulsion Conference*, July 25-27, 2016, Salt Lake City, UT, pp 1-18

[5]LeClair, A.C., Hedayat, A. and Majumdar A.K., "Numerical Modeling of an Integrated Vehicle Fluids System Loop for Pressurizing a Cryogenic Tank", AIAA Propulsion and Engineering Forum, 53rd *AIAA/SAE/ASEE Joint Propulsion Conference*, 10-12 July 2017, Atlanta, GA

[6]Rohsenow WM, Hartnett JP, Cho YI., *Handbook of heat transfer*. 3rd edition,, McGraw-Hill; 1998. p. 4.13.

[7]Majumdar, Alok, Valenzuela Juan, LeClair, Andre and Moder, Jeff," Numerical modeling of self-pressurization and pressure control by a thermodynamic vent system in a cryogenic tank, Cryogenics 74 (2016) 113–122

NASA'S BUILDING BLOCK APPROACH: REAL-TIME HYBRID SIMULATION FOR DYNAMIC MODEL CALIBRATION AND VALIDATION

Maria D. Cortes-Delgado[1] and
University of Puerto Rico, Mayagüez Campus, Mayagüez, Puerto Rico, 00681

Jeffrey Peck, Jr.[2], Daniel Lazor[3]
And

Eric Stewart[4]

EV31 Structural Dynamics and System Integration Branch, NASA MSFC, Huntsville, AL, 35812 US

ABSTRACT

NASA programs implement the Building Block Approach for dynamic model calibration and validation. The Building Block Approach entails partitioning large space vehicles into substructures, where each substructure contractor is responsible for providing a test correlated dynamic model of their substructure. This procedure offers cost and schedule advantages, but it also faces some fundamental challenges. The main challenge when implementing the Building Block Approach is that the substructures are generally not tested in a configuration that replicates their behavior in the integrated vehicle. Hybrid Simulation methodology could be utilized with the Building Block Approach to help model the system-level dynamics when trying to conduct dynamic model calibration and validation of NASA's large space structures. Hybrid Simulation combines both analytical and experimental methodologies in a single simulation or test. The addition of Hybrid Simulation to the Building Block Approach would enable the determination of the effect of each substructure modal response to the system response as the system dynamics are modeled real-time during a substructure test. The main objective of the author's Summer Faculty Fellowship research project was to test the feasibility of applying Hybrid Simulation and Building Block Approach for the calibration of NASA's large space structures by means of a proof of concept test. The tasks designed to achieve the main objective were to: design a test setup to provide data on utilizing Hybrid Simulation and Building Block Approach for modal test calibration and validation, compare the results with conventional modal test calibration, to build a test setup and modeling platform for further parametric studies, and provide recommendations for future work. This paper discusses the completed tasks during the Summer Faculty Fellowship which included: the design and construction of the test setup to conduct the proof of concept test and recommendations for future work.

1. Introduction

In past programs, the National Aeronautics and Space Administration (NASA) has relied on an Integrated Vehicle Ground Vibration Test (IVGVT) for calibration of vehicle dynamic models prior to first flight. Some examples of launch vehicles where this approach has been used are the Saturn 1B, Saturn V, and the Space Transportation System (STS, i.e. the space shuttle). While this approach works well and is accepted by the technical community, it is very costly in terms of both money and schedule. For the Space Launch System (SLS), in order to

[1] Associate Professor, Department of Engineering Sciences and Materials, University of Puerto Rico, Mayaguez Campus.
[2] AST, NASA MSFC EV31 Structural Dynamics, Bldg. 4600, 2nd floor.
[3] AST, NASA MSFC EV31 Structural Dynamics, Bldg. 4600, 2nd floor.
[4] AST, NASA MSFC EV31 Structural Dynamics, Bldg. 4600, 2nd floor.

cut costs and schedule, NASA decided to follow the *Building Block Approach (BBA)* for dynamic model calibration and validation. The *BBA* consists in *partitioning* or dividing the main structure in multiple *substructures* and modeling and testing each individual substructure. In the *BBA,* for NASA programs, each contractor developing primary structural components was responsible for providing a test correlated dynamic model of their substructure. These correlated substructure models would then be integrated into a vehicle dynamic model for various verification analyses prior to flight. Although, the *BBA* does offer cost and schedule advantages when dealing with large space structures it also faces some fundamental challenges.

One of the main challenge inherent in the *BBA* is that the substructures are generally not tested in a configuration that replicates their behavior in the integrated vehicle. This results in these substructures being correlated to modes that are not guaranteed to participate in the vehicle modes of interest. Specifically, the substructure interfaces may not be sufficiently exercised. Thus, there will remain some uncertainty that the integrated vehicle model will not accurately represent the complete system dynamics in flight. In response to this uncertainty, the SLS program instituted several integrated vehicle tests prior to launch in an attempt to buy down risk. Related to this challenge is the question of how to design the substructure tests. If the substructure test is not configured to replicate the substructure's behavior while integrated in the vehicle, then it becomes difficult to identify modes that directly relate to the important flight modes, which results in less confidence that the substructure is adequately tested.

Hybrid Simulation (HS) and *Real-Time Hybrid Simulation (RTHS)* offer a solution to the challenges facing NASA's Building Block Approach for dynamic model calibration and validation. The creation of analytical models and experimental tests are commonly utilized methodologies to study the response of structures to dynamic loading. *HS* combines both analytical and experimental methodologies in a single simulation or test taking advantages of each research method. Fig. 1 shows the basic hybrid simulation concept where a structure is partition in analytical and experimental substructures where they could have overlapping or nonoverlapping boundaries. Figure 2 gives an example of how *HS* is utilized to test a wind turbine blade at the University of Denmark where they partition a structure (blade) into *analytical/numerical and experimental/physical substructures*.

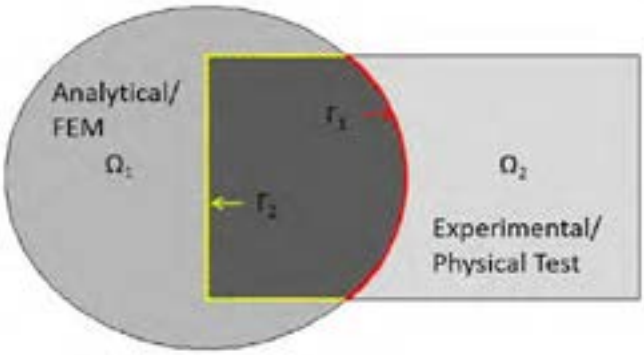


Figure 1. Basic Hybrid Simulation Concept.

Figure 2. Hybrid Simulation of a blade at Technique University of Denmark [1, 2]

Substructuring, allows to physically test, at full-scale, parts of the original structure while modeling the rest of the prototype (analytical substructure). Parts of the original structure that should be tested are elements or components that cannot be fully modeled/understood because its response may be nonlinear or a component to which the original structure is sensitive to model uncertainty. This allows for less complex and more cost-effective experimental tests and setups. Hybrid Simulation has been tested in the civil infrastructure vibrations community for three decades and deemed as a reliable test and simulation methodology to obtain the dynamic response of structures. However, this methodology has only been utilized for validation of civil structures and not for model calibration and/or space structures.

Real Time Hybrid Simulation (RTHS) emerges from hybrid simulation due to the need to test rate dependent (displacements and velocities) experimental substructures using a one-to-one time scale loading protocol [3, 4, 5]. *RTHS* has been mostly used to test simple configurations (columns, dampers, base isolators) on civil structures [6], but have also been used to test large and complex configurations for earthquake type loads [7].

Hybrid Simulation integrates both experimental and analytical methods when solving the equations of motion, as described in Eq. (1),

$$(M \cdot \ddot{u})^A + (C \cdot \dot{u})^A + (K \cdot u)^A + r^E = (-M^t \cdot \ell \cdot \ddot{u}_g)^A \quad \text{Eq. (1)}$$

Where the *A* and *E* subscripts means analytical and experimental substructures, respectively. *HS* can be displacement or velocity controlled. Furthermore, how the target displacements are imposed and the experimental restoring forces obtained determines if it is quasi-static or real-time *HS*. For quasi-static displacement-controlled *HS*, the testing loading time scale protocol runs at slower rates than the actual excitation time scale. Quasi-static *HS* are conducted when the experimental substructures are nor rate or velocity dependent. In *RTHS*, the testing loading time scale protocol is run at the same rate as the excitation time scale. *Hybrid Simulation* methodology could be utilized with the *Building Block Approach* to help model the system-level dynamics when trying to conduct dynamic model calibration and validation of NASA's large space structures. The *BBA* already utilizes substructure testing in their approach, but *H*S would offer the ability to conduct modal testing and calibration of one substructure at a time while the remainder of substructure(s) could be represented with analytical finite element models (FEMs). Therefore, NASA could obtain the actual forces at the boundaries between the substructures of their large space structures that they may encounter in real flight conditions. Furthermore, the procedure allows to, as the testing and calibration progresses, treat an initial experimental substructure into a calibrated analytical substructure that will help in the calibration and validation process of a *"new"* experimental substructure as shown in Fig. 3.

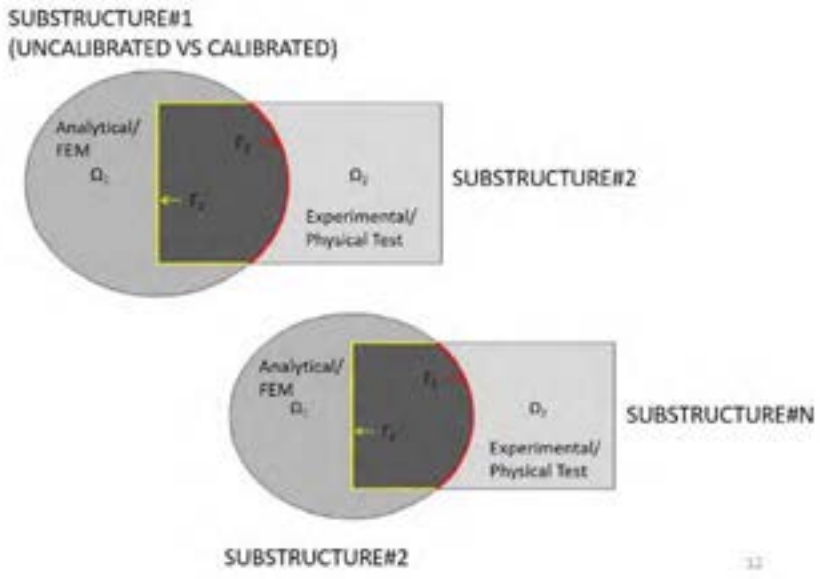


Figure 3. Proposed solution combining Hybrid Simulation and Building Block Approach.

Another benefit of combining *HS* and *BBA* is the determination of the target modes for the substructure modal tests. Furthermore, the addition of *HS* to the *BBA* would enable the determination of the effect of the substructure modal response to the system response as the system dynamics are modeled real-time during a substructure test.

2. Hybrid Simulation Fundamental Theory and Procedure

HS is a methodology that combines analytical and experimental substructures when simulating and obtaining the dynamic response of structures. In *HS,* the equations of motion are solved using a time-stepping integration procedure where at each time step the equations of motion are solved for the incremental deformations (displacements, velocities and accelerations), as presented in a very conceptual form in Eq. 1:

The time stepping procedure to run a typical hybrid simulation, based on Eq. (1), is presented in a flowchart on Fig. 4. Figure 5 illustrates how this procedure is applied from the numerical/analytical substructure to the physical/experimental substructure for a single degree of freedom (SDOF) structure. Following Fig. 4, the equation of motion is solved for the displacement at each step using responses at the previous steps. In the case of an explicit integration scheme, displacements are determined:

$$d_{i+1} = d_i + \Delta t v_i + \frac{1}{2}\Delta t^2 a_i \qquad \text{Eq. (3)}$$

Displacements are then sent and imposed on the experimental structure with actuators and to the numerical substructures as the target displacement.

The corresponding restoring forces, measured forces from the experimental substructure (from the Load Cell and LVDT on the actuator), and simulated forces from the numerical substructures are fed back to the equation of motion. The remaining responses at the time step including the velocity and acceleration are calculated based on the restoring and the external forces. The procedure is repeated until the end of the simulation time.

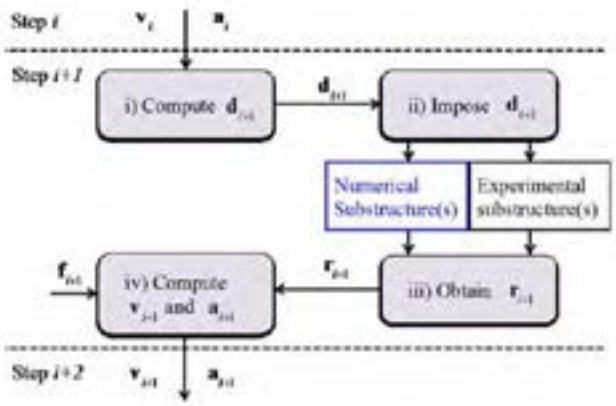


Figure 4. Time stepping procedure in a typical hybrid simulation [8].

3. Hybrid Simulation Configuration and Implementation

A basic Hybrid Simulation simulation/testing will require computational/analytical components and a physical/experimental component.

A. Physical/Experimental System Components

The physical system, as illustrated in Fig. 5, is comprised of: 1) servo-hydraulic system (servo-hydraulic actuators, servo-controllers and a hydraulic power system; 2) a test article with structural frames to be able to attach the actuators to the degrees of freedom needed to impose the commanded displacements and 3) a data acquisition system (sensors, A/D cards and data acquisition software) to measure the responses of the test article and send back to the computational/analytical component.

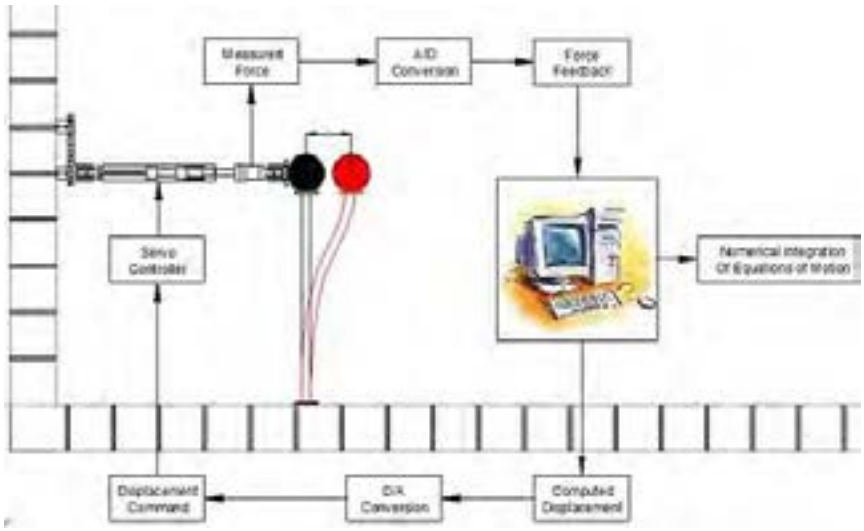


Figure 5. Hybrid Testing for a SDOF structure.

B. Computational/Analytical System Components

The computational system, as illustrated in Fig. 5, comprises: the computer hosting the numerical simulations (FEA, servo-controller, and data acquisition (DAQ) software) and a *master/coordinator* simulator component. The computer with the numerical simulations needs to be linked to the laboratory equipment in order to get the feedback from the sensors and receive next target displacements to be sent as command to the actuators. This data transfer between laboratory and computational system requires a custom coding program or available standard middleware packages such as UI-SIMCOR and OpenFresco. This would ensure a precise and accurate data transfer running at the necessary simulation time.

The master/coordinator simulator controls the time-stepping integration algorithm and coordinates the communication between substructures.

4. Hybrid Simulation at NASA MSFC – EV31

The main objective of the author's Summer Faculty Fellowship research project was to test the feasibility of applying Hybrid Simulation and Building Block Approach for the calibration of NASA's large space structures by means of a proof of concept test. The objectives of the proof of concept test were to: (1) design a test setup to provide data on utilizing Hybrid Simulation and Building Block Approach for modal test calibration and validation, (2) compare the results with conventional modal test calibration and validation, (3) build a test setup and modeling platform for further parametric studies, and (4) provide recommendations for future work.

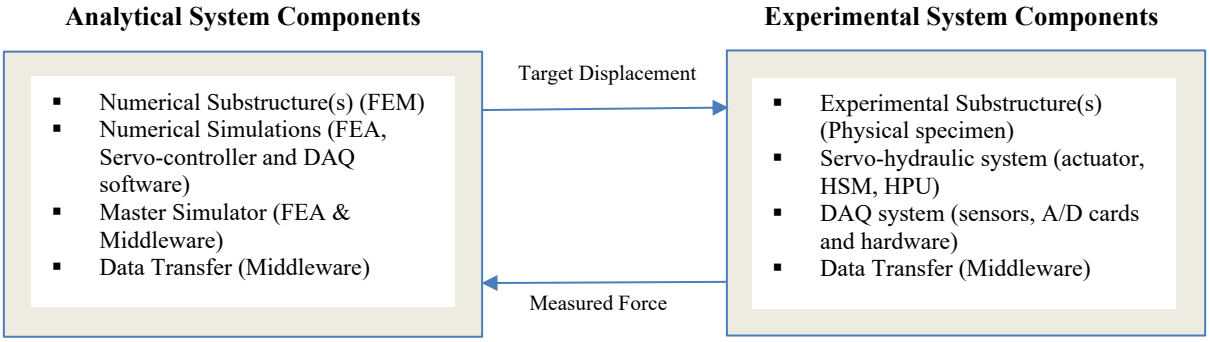


Figure 6. Basic Hybrid Testing Components.

Figure 6 lists the necessary analytical and experimental system components to run a simple Hybrid Simulation. As stated earlier, *HS* was created for the validation of civil structures under dynamic loading (usually earthquake excitation). Figure 6 was modified, in Fig. 7, to determine the basic configuration of the hybrid system components necessary for the determination of the modal properties of space structures. Furthermore, Fig. 7 lists the final analytical and experimental system component configuration for the implementation of the summer's proof of concept test.

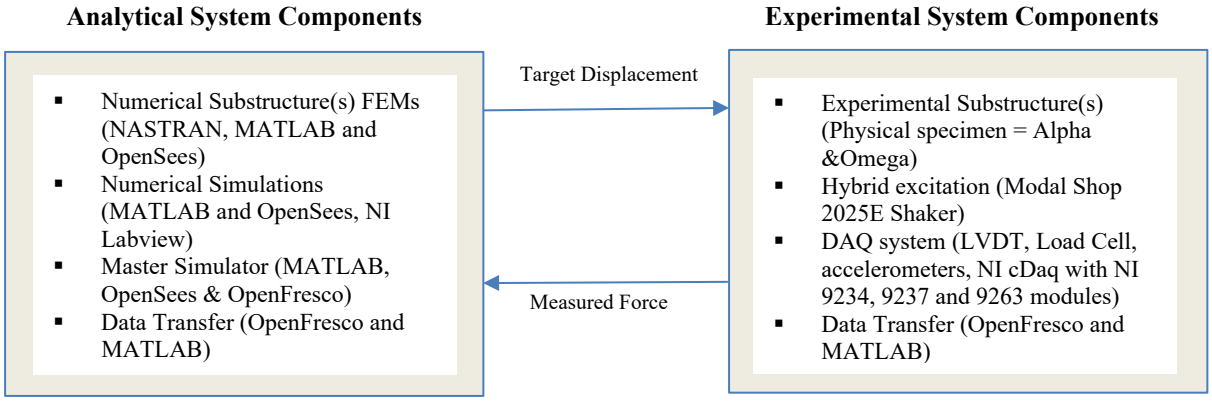


Figure 7. Proof of Concept Hybrid Testing Components.

5. Proof of Concept Test

This section discusses the design and setup of the summer's Proof of Concept Test. First, the prototype selected for the test will be described. Then, the prototype is analyzed to select the adequate analytical and experimental substructures. Finally, the final analytical and experimental system component configuration mentioned in Fig. 7, will be described in more detail manner.

A. Aerospace Prototype Structure

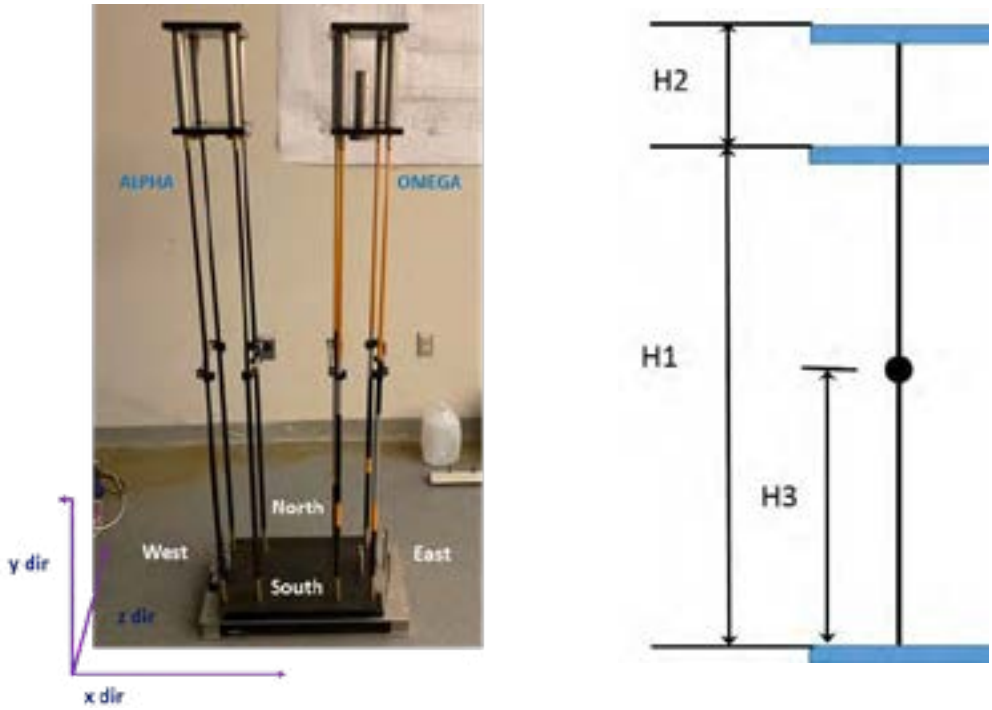
The aerospace prototype structure chosen for the proof of concept test is shown in Fig. 8. The prototype consists of two three-dimensional flexible towers with masses on the top. The two towers, *Alpha* and *Omega* are assumed to be fixed at the bottom to a steel plate. The dimensions of the tower's elements were measured, and the basic properties were determined as described in Table 1. The main tower structure is 50 inches high (*H1*) and has plan dimensions of 7.0 inches wide (*x-dir.*) and 4.0 inches in depth (*z-dir.*). The vertical elements are aluminum hollow sections that vary from bigger diameter at the bottom to smaller diameter at the top. The weight of the masses added on top of Alpha and Omega is 12.49 lbs. and 12.29 lbs., respectively. The Modulus of elasticity of the rods was calculated to be 1.224×10^7, with Poisson ratio of 0.3 and 0.000254 mass density.

Table 1. Vertical element properties along its height.

Height (in) (from base)	Diameter $_{outside}$ (in)	Diameter $_{inside}$ (in)	Area (in^2)	Inertia (in^4)	Polar Moment of Inertia (in^4)
13.75	0.655	0.608	0.04662	0.00233	0.00465
27.00	0.578	0.531	0.04094	0.00158	0.00315
41.875	0.389	0.342	0.02698	0.00045	0.00090

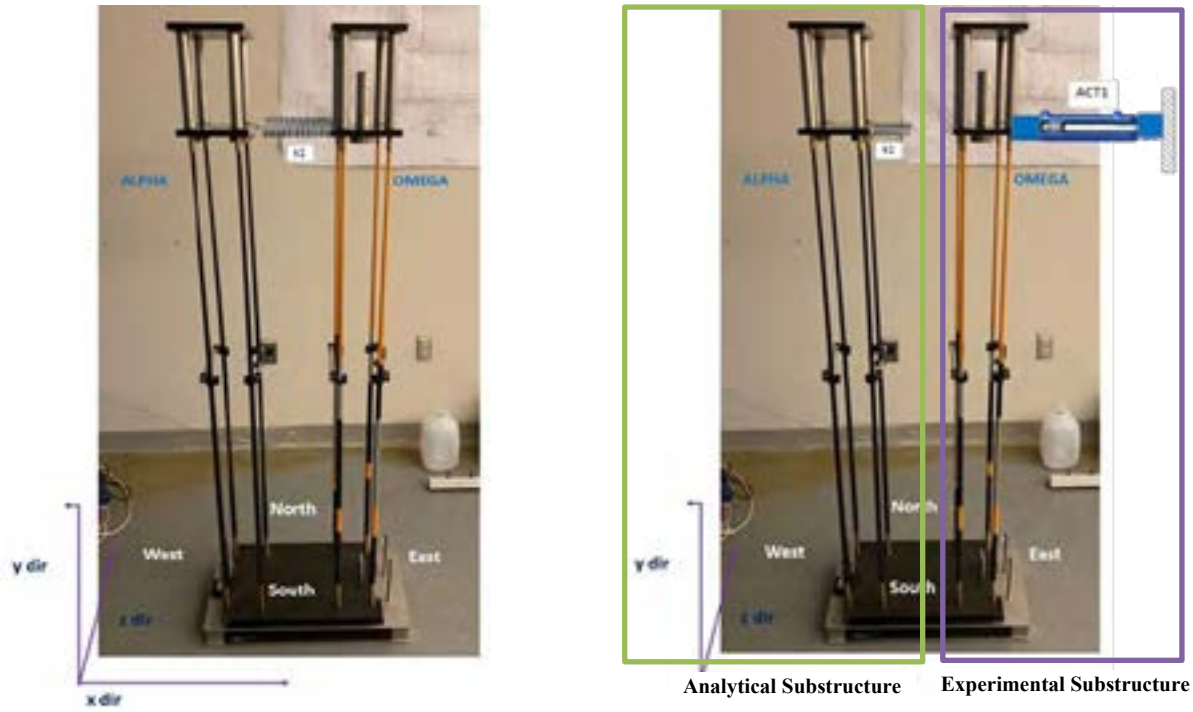
B. Coupled Prototype and Hybrid Coupled Substructures

The Alpha and Omega towers were coupled by means of a mechanical spring with measured coil stiffness, k_{spring} = 14lb/in, as shown in Fig. 9(a). In Figure 9(b), the coupled towers are partitioned in analytical and experimental substructures. The analytical substructure comprises the Alpha tower with the coupling spring finite element model (FEM) and the Omega tower will be the physical test specimen.



(a). Prototype structure. (b). Elevation profile.

Figure 8. Proof of Concept prototype structure.



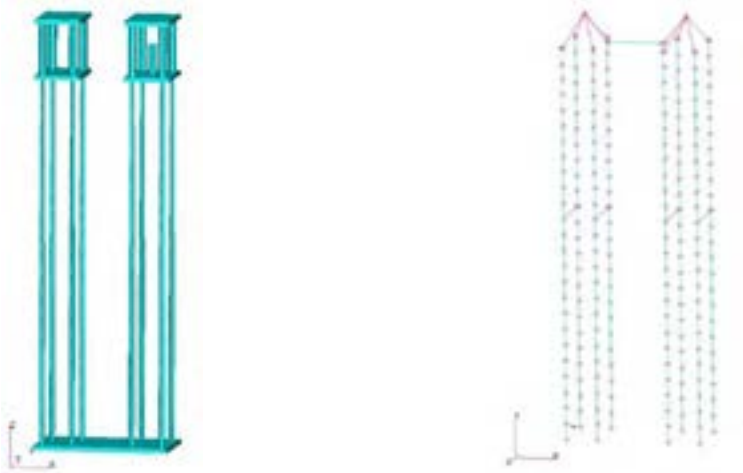
(a.) Coupled towers. (b). Hybrid Coupled structures.

Figure 9. Coupled and Hybrid Coupled structure.

C. Analytical System Components

1. Numerical Substructure FEMs

Finite Element Models (FEMs) with NASTRAN, OpenSees and MATLAB software packages were developed for the coupled towers to determine their fundamental modal properties. Two FEM of the coupled towers were developed: full geometry model and a reduced DOF model. Figure 10 shows the NASTRAN FEMs of the coupled towers.



(a). Full Geometry FEM (b). Reduced DOF model

Figure 10. NASTRAN FEM of the two towers.

The full geometry FEM was developed to get an early cut at the modal frequencies and shapes that the conventional modal test should target. The two top masses were developed using bar and plate elements with basic steel properties and measured dimensions. Rigid body mass matrices were calculated for the top masses and used for the early version of the reduced FEM. Figure 11 shows a rendering of the FEM top masses and their associated mass properties.

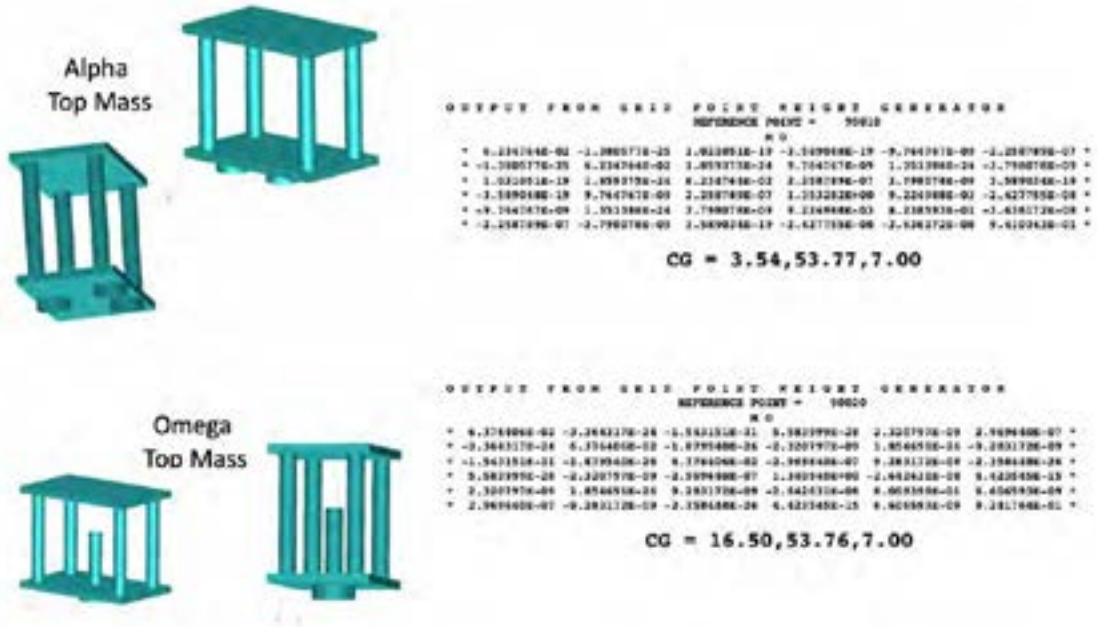


Figure 11. Top mass FEM representations and calculated mass properties.

Using the rigid body mass properties, the top masses were replaced with a lumped mass, connected to the flexible rods with a rigid element (RBE2). Table 2 shows the NASTRAN eigenvalue table for the first 10 modes.

Table 2. Reduced Model Eigenvalues and Frequencies.

R E A L E I G E N V A L U E S

MODE NO.	EIGENVALUE	RADIANS	CYCLES
1	5.737290E+01	7.574490E+00	1.205518E+00
2	2.271455E+02	1.507135E+01	2.398679E+00
3	2.299951E+02	1.516559E+01	2.413678E+00
4	5.230711E+02	2.287075E+01	3.639992E+00
5	1.279705E+03	3.577296E+01	5.693444E+00
6	1.297936E+03	3.602687E+01	5.733855E+00
7	8.976087E+04	2.996012E+02	4.768301E+01
8	9.281502E+04	3.046556E+02	4.848744E+01
9	1.237718E+05	3.518122E+02	5.599265E+01
10	1.237768E+05	3.518192E+02	5.599377E+01

The mode shapes associated with the first 8 modes are shown in Fig. 12.

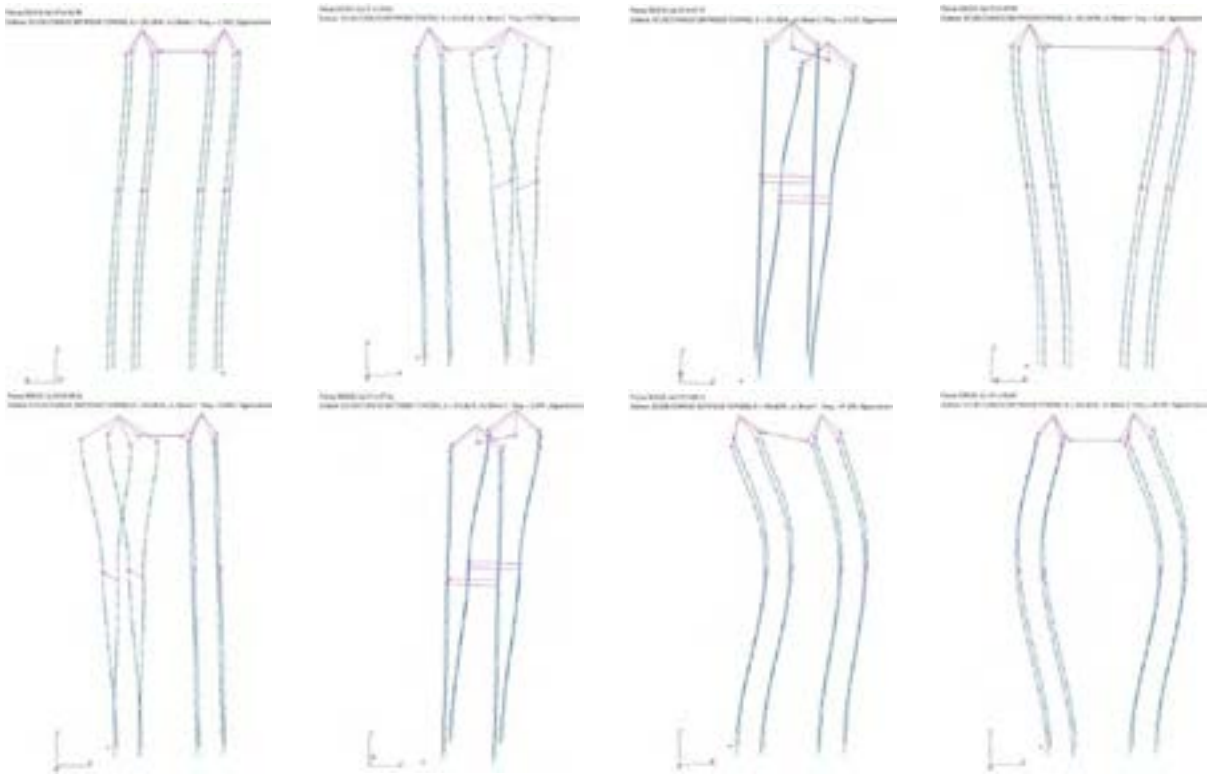


Figure 12. Mode Shapes for 1st - 8th Modes of the Reduced FEM

Looking at the results of the modal analysis of the reduced model, it was noted that the only low frequency mode with appreciable strain in the spring between the two towers is mode 4. This was the mode chosen for replication for the Proof of Concept Test.

2. Numerical Simulations

Analytical coupled simulations determine the expected responses of the coupled hybrid tests by simulating the experimental substructure with a FEM. The FEM software packages chosen to run the numerical coupled simulations for the Proof of Concept Test were MATLAB and OpenSees. These numerical simulations are being processed and are part of future work.

3. Master Simulator

OpenSees and MATLAB are the master/coordinator simulators chosen to control the time-stepping integration algorithm and the coordination of the communication between substructures

4. Data Transfer

The data transfer between laboratory and OpenSees and MATLAB will be carried out with the open source code OpenFresco.

D. Experimental System Components

Figure 13 shows the experimental system components for the proof of concept test consisting of the physical specimen, hybrid excitation, DAQ and data transfer system.

1. Physical Specimen

Omega was chosen as our physical specimen for the proof of concept hybrid test, as shown in Fig. 13. Target displacements calculated by the numerical simulation on the analytical substructure are going to be imposed on Omega and the restoring forces measured from a load cell attached to Omega, are going to be sent back to the analytical substructure for the next round of hybrid simulation.

(a). Experimental Substructure Concept

(b). Experimental Substructure Test Setup at Bldg. 4000 Room 1103.

Figure 13. Proof of Concept Test Setup.

1. Hybrid excitation

A modal shaker is going to apply the calculated target displacements on Omega. The modal shaker, from the Modal Shop type 2025E, is rigidly attached to a reaction frame and aligned at the middle of the top mass of Omega, as shown in Fig. 14.

2. DAQ system

The data acquisition consists of the necessary sensors and hardware to measure displacements, accelerations and force while conducting hybrid testing. An LVDT is attached at the top mass of Omega to control the commanded versus measured displacement, as shown in Fig. 14. Furthermore, a load cell is also attached to the top mass to measure forces needed to feed back the analytical substructure, as shown in Fig. 15. Accelerometers are mounted at the top mass of Omega and at different locations in order to determine the modal response of Omega during the hybrid test, as shown in Fig. 15. The hardware configured to acquire data and convert form digital to analog/ from analog to digital signal is National Instruments cDAQ with NI 9234, 9237 and 9263 modules, as shown in Fig. 16.

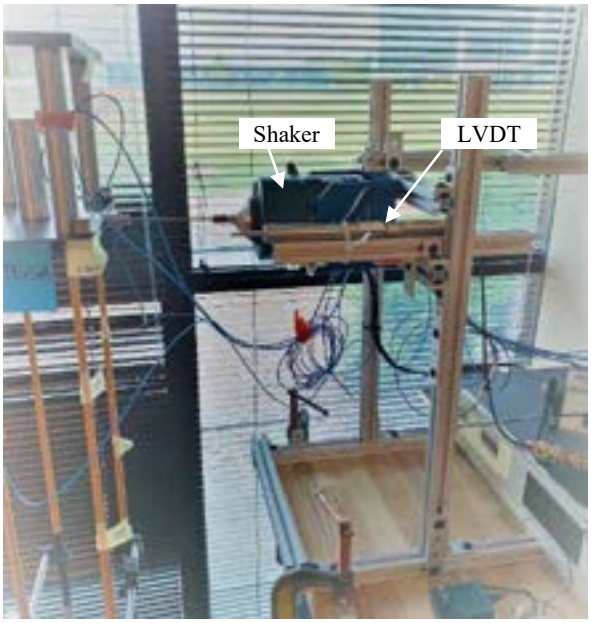


Figure 14. Modal Shaker and LVDT setup.

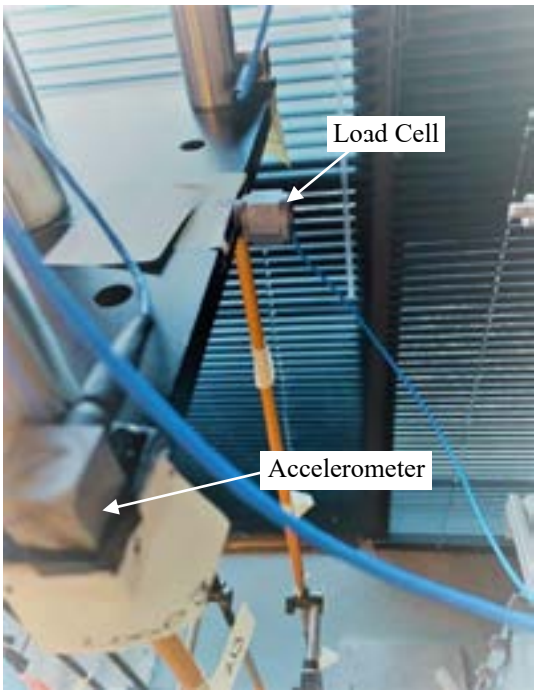


Figure 15. Load Cell and Accelerometer Setup for Omega.

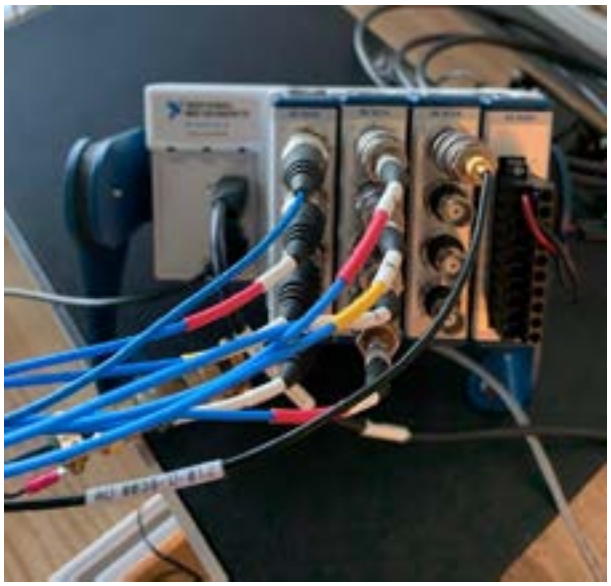


Figure 16. National Instrument DAQ system.

3. Data Transfer

The data transfer between laboratory and OpenSees and MATLAB will be carried out with the open source code OpenFresco.

6. Future Work: Proof of Concept Test

This section discusses and list recommendations for future work on the designed and built Proof of Concept Hybrid Test Setup. The author provides short- and long-term recommendations for this test setup.

A. Short term tasks

The hybrid test setup needs to close the loop of the targeted and measure shaker displacements. Next, it is necessary to run numerical coupled simulations to determine response of the system. (the author is currently conducting this runs). Finally, we must be able to post process data and compared with conventional modal test data, developed at EV31. Then, compare conventional modal test approaches with *HS* with *BBA* modal test data.

B. Long term tasks

To conduct more complex hybrid setups and fully explore its capabilities on large space structures the author recommends: to acquire servo-hydraulic system, SCRAMNET CARD, DSP-SpeedGoat and run Real-Time Hybrid Simulation with parametric studies on frequency, configurations and rotational and torsion degrees of freedom.

7. Acknowledgments

The author, M. Cortes-Delgado, thanks NASA MSFC EV31 collaborators for the summer research work. Furthermore, M. Cortes-Delgado acknowledges the IT and testing support from EV31 and ET40 personnel. Finally, the author would like to thank the NASA MSFC Summer Faculty Fellowship program for sponsoring this research and for their financial support.

8. References

[1] Chae, Y., Ricles, J.M. and Sause, R., Large-scale real-time hybrid simulation of a three-story steel frame building with magneto-rheological dampers, Earthquake Engng Struct. Dyn., 43:1915–1933

[2] Zhang, Z., Staino, A., Basu, B, Nielsen, Søren R.K., Performance evaluation of full-scale tuned liquid dampers (TLDs) for vibration control of large wind turbines using real-time hybrid testing, Engineering Structures Journal, Vol. 126, 417-431, 2016.

[3] K, Udagawa K, Seki M, Okada T, Tanaka H, Non-linear earthquake response analysis of structures by a computer-actuator on-line system, Bul earthquake eng resist struct res cent, 1, Institute of Industrial Science, University of Tokyo, 1980.

[4] Mahin SA, Shing PSB, Pseudodynamic method for seismic testing, J Struct Eng, Vol. 111(7), 1482–503, 1985.
[5] Nakashima M, Kato H, Takaoka E., Development of real-time pseudo dynamic testing. Earthq Eng Struct Dyn, Vol. 21(1), 79–92, 1992.
[6] Darby AP, Blakeborough A, Williams MS., Real-time substructure tests using hydraulic actuator. J Eng Mech, Vol. 125(10), 1133–9, 1999.
[7] Shao, X. and Griffith, C., An overview of hybrid simulation implementations in NEES projects, Engineering Structures Journal, Vol. 56, 1439-1451, 2013.
8 Nakata, N, Dyke, S, Zhang, J, Mosqueda, G, Shao, X, Hussam, M, Head, M , Erwin Bletzinger, M, Marshall, G A, Ou, G and Song , C, "Hybrid Simulation Primer and Dictionary", Hybrid Simulation Task Force, NEES, West Lafayette, Indiana, March 2014.

Dynamics of Multi-tether Electric Sails

Todd D. Lillian*

John Rakoczy†

The electric sail (E-sail) has been proposed as a novel propulsion system that, if successful, promises to facilitate missions that would be inconvenient or impractical to complete using traditional chemical propellants. The most common E-sail concept includes a central hub with several long positively charged tethers extending radially outward that harvest momentum from protons in the solar wind. Two primary obstacles remain before the E-sail technology is ready for implementation in scientific missions: (i) proving the electric sail effect works in space (and beyond earth's magnetosphere) as predicted by theory and laboratory experiments, and (ii) proving that the complicated dynamics of a multitether E-sail can be controlled throughout all phases of a scientific mission. The objective of this paper is to address the second obstacle by modeling and characterizing the dynamics of multitether E-sails in preparation for the development of necessary control systems. Here we derive a linear vibration model and implement it to find natural frequencies and mode shapes for a multitether Esail. We show that many of these frequencies can be estimated using simple models. Finally, we develop a full nonlinear dynamics model for multitether Esails. With this nonlinear model we identify a possible instability associated with tether extension and demonstrate a steering maneuver which successfully changes the axis of rotation of the E-sail.

I. Introduction

Traditional chemical propulsion systems have served the scientific community well for decades. However, these propulsion systems have proven inconvenient or impractical for some missions. For example, it took over 30 years (and multiple gravity assists [1]) for Voyager 1 and 2 to reach and return data from interstellar space. As another example, the Ulysses spacecraft required a gravity assist at Jupiter to escape the ecliptic plane [2] and complete 3 solar polar orbits over the course of about 17 years. At a minimum, the long duration and required gravity assists for such missions are inconveniences. Novel propulsion systems are worth considering as a means of overcoming these inconveniences on future missions.

One such propulsion system, the electric sail (E-sail), was concieved by Janhunen [3]. In general an E-sail consists of a long length of conducting wire that is positively charged and thereby repels protons in the solar wind to provide thrust. A more detailed and patented concept consists of a spinning spacecraft with many tethers extending radially outward [4]. Rotation of the system is used to develop tension on each tether. Depending on the voltage applied to an individual tether, a force can develop on the tether due to interaction with the solar wind. Disadvantages of the E-sail include its low thrust force and the need to be outside the earth's magnetosphere. However, a major advantage for an E-sail over the similar solar sail is that the thrust force decays approximately with the inverse of the distance from the sun rather than the inverse of the distance squared [5]. Two primary obstacles remain before E-sail technology is ready for implementation in scientific missions: (i) proving the electric sail effect works in space (and beyond earth's magnetosphere) as predicted by theory and laboratory experiments, and (ii) proving that the complicated dynamics of a multitether E-sail can be controlled throughout all phases of a scientific mission. The objective of this paper is to model and characterize the dynamics of multitether E-sails in preparation for future development of necessary control systems.

*Assistant Professor, Department of Mechanical Engineering, University of South Alabama, 150 Jaguar Drive, Shelby Hall 3128, Mobile, AL, 36688.

†Branch Chief, Controls Systems Design & Analysis Branch, Marshall Space Flight Center, Mail Code EV41, Huntsville, AL, 35812.

II. Methods

A. Continuous Time Equations of Motion

Here we derive the equations of motion for an e-sail with a hub and spoke configuration; see Figure 1. We define R_0 as an inertial reference frame for the system. The hub has a position x_A relative to the inertia frame, a direction cosine matrix R_A (which transforms from the body fixed reference frame to the inertial frame), a mass of m_A, and a diagonal inertia matrix J_A relative to the body fixed frame. Each of the N spokes/tethers supports a pod. The i^{th} pod has a position x_{Bi} relative to the inertia frame, a direction cosine matrix R_{Bi} (which transforms from the body fixed reference frame to the inertial frame), a mass of m_B and a diagonal inertia matrix J_B relative to its body fixed frame.

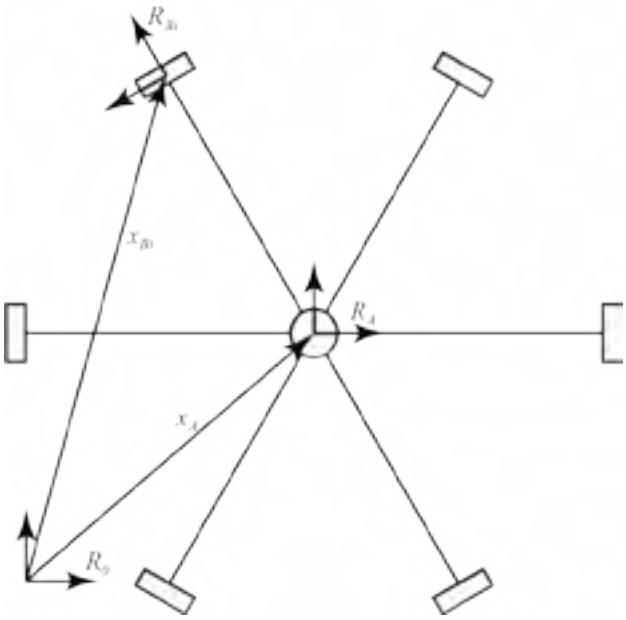


Fig. 1 Schematic of an E-sail with a hub and spoke configuration.

Here we represent the tethers as massless springs with modulus of elasticity E, cross sectional area A, and free length l_0. The i^{th} tether is attached to the hub at a point ρ_{Ai} relative to the mass center of the hub and $\rho_B = [-||\rho_B||,\ 0,\ 0]^T$ relative to the mass center of the i^{th} pod, each expressed in their respective body fixed frames. In our model, elasticity of the tethers is the only contribution to the potential energy of the system, which we write as

$$U = \sum_{i=0}^{N-1} \frac{EA}{2l_0} \left(||l_i|| - l_0\right)^2 . \tag{1}$$

Here l_i denotes the vector along the length of the i^{th} tether from hub to pod and can be written relative to the inertial frame as

$$l_i = (x_{Bi} + R_{Bi}\rho_B) - (x_A + R_A\rho_{Ai}) . \tag{2}$$

We write the kinetic energy of the system as

$$T = \frac{1}{2}m_A\dot{x}_A^T\dot{x}_A + \sum_{i=0}^{N-1}\frac{1}{2}m_B\dot{x}_{Bi}^T\dot{x}_{Bi} + \frac{1}{2}\Omega_A^T J_A\Omega_A + \sum_{i=0}^{N-1}\frac{1}{2}\Omega_{Bi}^T J_B\Omega_{Bi} . \tag{3}$$

Here dot represents the time derivative. Ω_A and Ω_{Bi} represent the angular velocitis of the hub and i^{th} pod respectively, each relative to the inertial frame and expressed in their body fixed frame. It follows that

$$\dot{R}_A = R_A\hat{\Omega}_A \tag{4}$$

$$\dot{R}_{Bi} = R_{Bi}\hat{\Omega}_{Bi}; \tag{5}$$

see, for example, [6]. Here the hat denotes a skew symmetric matrix formed by the elements of the associated vector such that the cross product can be written as a matrix multiplication, $a \times b = \hat{a}b = -\hat{b}a$.

In our model, we represent the net force due to solar wind interactions with the i^{th} tether as F_i. We approximate the tether to be a straight massless element with the net force acting at its midpoint. For these forces, we express the non conservative virtual work as

$$\delta W_{\text{nc,wind}} = \sum_{i=0}^{N-1} \frac{1}{2} F_i^T \delta \left(x_A + R_A \rho_{Ai} \right) + \sum_{i=0}^{N-1} \frac{1}{2} F_i^T \delta \left(x_{Bi} + R_{Bi} \rho_B \right). \tag{6}$$

Applying the extended Hamilton's principle we write

$$\int_{t_1}^{t_2} \left(\delta T - \delta U + \delta W_{\text{nc,wind}} \right) = 0. \tag{7}$$

We perform the variations following the approach of Lee et al. [6]. Specifically, we write the variations of the direction cosine matrices and angular velocities as follows

$$\delta R_A = R_A \hat{\eta}_A \tag{8}$$
$$\delta R_{Bi} = R_{Bi} \hat{\eta}_{Bi} \tag{9}$$
$$\delta \Omega_A = \dot{\eta}_A + \hat{\Omega}_A \eta_A \tag{10}$$
$$\delta \Omega_{Bi} = \dot{\eta}_{Bi} + \hat{\Omega}_{Bi} \eta_{Bi}. \tag{11}$$

Here η_A and η_{Bi} are Lie algabra elements representing the variation of the direction cosine matrices; see, for example, [6]. Applying these definitions, the integrand of Equation 7 can be expressed in terms of the following independent set of variations (i.e., virtual displacements): δx_A, δ_{Bi}, η_A, and η_{Bi}. The coefficients of these variations yield the following equations of motion

$$m_A \ddot{x}_A = \sum_{i=0}^{N-1} \frac{EA}{l_0 ||l_i||} (||l_i|| - l_0) l_i + \sum_{i=0}^{N-1} \frac{1}{2} F_i \tag{12}$$
$$m_B \ddot{x}_{Bi} = -\frac{EA}{l_0 ||l_i||} (||l_i|| - l_0) l_i + \frac{1}{2} F_i \tag{13}$$
$$J_A \dot{\Omega}_A = -\hat{\Omega}_A J_A \Omega_A + \sum_{i=0}^{N-1} \left(\frac{EA}{l_0 ||l_i||} (||l_i|| - l_0) \hat{\rho}_{Ai} R_A^T l_i + \frac{1}{2} \hat{\rho}_{Ai} R_A^T F_i \right) \tag{14}$$
$$J_B \dot{\Omega}_{Bi} = -\hat{\Omega}_{Bi} J_B \Omega_{Bi} - \frac{EA}{l_0 ||l_i||} (||l_i|| - l_0) \hat{\rho}_B R_{Bi}^T l_i + \frac{1}{2} \hat{\rho}_B R_{Bi}^T F_i \tag{15}$$
$$\dot{R}_A = R_A \hat{\Omega}_A \tag{16}$$
$$\dot{R}_{Bi} = R_{Bi} \hat{\Omega}_{Bi}. \tag{17}$$

B. Small Vibrations About a Reference State

In this section we consider a reference state in which the system rotates as a rigid body about its axis of symmetry, z-axis, with a constant angular velocity of magnitude $\bar{\Omega}$, accelerates with magnitude $\bar{a}$ along this axis of symmetry, and experiences the same constant solar wind load on each tether (in the direction of the solar wind). We can write

expressions for the variables in the reference state as follows

$$\bar{\ddot{x}}_A = \begin{bmatrix} 0 & 0 & \bar{a} \end{bmatrix}^T \tag{18}$$

$$\bar{x}_{Bi} = \bar{x}_A + \bar{R}_A \rho_{Ai} + \bar{l}_i - \bar{R}_{Bi} \rho_B \tag{19}$$

$$\bar{\Omega}_A = \begin{bmatrix} 0 & 0 & \bar{\Omega} \end{bmatrix}^T \tag{20}$$

$$\bar{\Omega}_{Bi} = \bar{R}_{Bi}^T \bar{R}_A \bar{\Omega}_A \tag{21}$$

$$\bar{\dot{\Omega}}_A = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \tag{22}$$

$$\bar{\dot{\Omega}}_{Bi} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \tag{23}$$

$$\bar{F}_i = \begin{bmatrix} 0 & 0 & ||\bar{F}|| \end{bmatrix}^T \tag{24}$$

$$\bar{R}_A = \begin{bmatrix} \cos(\bar{\Omega}t) & -\sin(\bar{\Omega}t) & 0 \\ \sin(\bar{\Omega}t) & \cos(\bar{\Omega}t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{25}$$

$$\bar{R}_{Bi} = \bar{R}_A \begin{bmatrix} \cos\left(\frac{2\pi i}{N}\right) & -\sin\left(\frac{2\pi i}{N}\right) & 0 \\ \sin\left(\frac{2\pi i}{N}\right) & \cos\left(\frac{2\pi i}{N}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\bar{\theta}_B) & 0 & \sin(\bar{\theta}_B) \\ 0 & 1 & 0 \\ -\sin(\bar{\theta}_B) & 0 & \cos(\bar{\theta}_B) \end{bmatrix} \tag{26}$$

$$\bar{l}_i = \bar{R}_A \begin{bmatrix} \cos\left(\frac{2\pi i}{N}\right)\cos(\bar{\theta}) \\ \sin\left(\frac{2\pi i}{N}\right)\cos(\bar{\theta}) \\ -\sin(\bar{\theta}) \end{bmatrix} ||\bar{l}_i|| \tag{27}$$

Here a bar is used to denote variables in the reference state. $\bar{\theta}_B$ and $\bar{\theta}$ are the angles at which the pods and tethers angle downward from the plane of the sail, respectively. Incidentally, to determine values of the variables in the reference state, one must substitute these expressions into the equations of motion for the system. If we denote small displacements from the reference state with ϵ_A and ϵ_{Bi} relative to the hub's body fixed frame and small changes in attitude with η_A and η_{Bi} we can write expressions for the variables as

$$\Omega_A = \bar{\Omega}_A + \dot{\eta}_A + \hat{\bar{\Omega}}_A \eta_A \tag{28}$$

$$\Omega_{Bi} = \bar{\Omega}_{Bi} + \dot{\eta}_{Bi} + \hat{\bar{\Omega}}_{Bi} \eta_{Bi} \tag{29}$$

$$R_A = \bar{R}_A + \bar{R}_A \hat{\eta}_A \tag{30}$$

$$R_{Bi} = \bar{R}_{Bi} + \bar{R}_{Bi} \hat{\eta}_{Bi} \tag{31}$$

$$x_A = \bar{x}_A + \bar{R}_A \epsilon_A \tag{32}$$

$$x_{Bi} = \bar{x}_{Bi} + \bar{R}_A \epsilon_{Bi} \tag{33}$$

$$l_i = \bar{l}_i + \bar{R}_A \epsilon_{Bi} - \bar{R}_{Bi} \hat{\rho}_B \eta_{Bi} - \bar{R}_A \epsilon_A + \bar{R}_A \hat{\rho}_{Ai} \eta_A. \tag{34}$$

Substitution of the above into Equations 12-15 and simplification yields the following linearized equations of motion

$$\begin{aligned}
\sum_{i=0}^{N-1} \frac{1}{2}\bar{R}_A^T f_i &= m_A\left(\hat{\bar{\Omega}}_A^2\epsilon_A + 2\hat{\bar{\Omega}}_A\dot{\epsilon}_A + \ddot{\epsilon}_A\right) \\
&\quad - \sum_{i=0}^{N-1}\frac{EA}{||\bar{l}||}\left(\frac{1}{||\bar{l}||^2}\{\bar{l}_i\}_A\{\bar{l}_i\}_A^T + \frac{||\bar{l}||}{l_0}I - I\right)(\epsilon_{Bi} - \bar{R}_A^T\bar{R}_{Bi}\hat{\rho}_B\eta_{Bi} - \epsilon_A + \hat{\rho}_{Ai}\eta_A) \qquad (35)\\
\frac{1}{2}\bar{R}_A^T f_i &= m_B\left(\hat{\bar{\Omega}}_A^2\epsilon_{Bi} + 2\hat{\bar{\Omega}}_A\dot{\epsilon}_{Bi} + \ddot{\epsilon}_{Bi}\right) \\
&\quad + \frac{EA}{||\bar{l}||}\left(\frac{1}{||\bar{l}||^2}\{\bar{l}_i\}_A\{\bar{l}_i\}_A^T + \frac{||\bar{l}||}{l_0}I - I\right)(\epsilon_{Bi} - \bar{R}_A^T\bar{R}_{Bi}\hat{\rho}_B\eta_{Bi} - \epsilon_A + \hat{\rho}_{Ai}\eta_A) \qquad (36)\\
\sum_{i=0}^{N-1}\frac{1}{2}\hat{\rho}_{Ai}\bar{R}_A^T f_i &= J_A\ddot{\eta}_A + \left(-\widehat{J_A\bar{\Omega}_A} + \hat{\bar{\Omega}}_A J_A\right)\dot{\eta}_A + \left(-\widehat{J_A\bar{\Omega}_A}\hat{\bar{\Omega}}_A + \hat{\bar{\Omega}}_A J_A\hat{\bar{\Omega}}_A\right)\eta_A \\
&\quad - \sum_{i=0}^{N-1}\frac{EA}{||\bar{l}||}\hat{\rho}_{Ai}\left(\frac{1}{||\bar{l}||^2}\{\bar{l}_i\}_A\{\bar{l}_i\}_A^T + \frac{||\bar{l}||}{l_0}I - I\right)(\epsilon_{Bi} - \bar{R}_A^T\bar{R}_{Bi}\hat{\rho}_B\eta_{Bi} - \epsilon_A + \hat{\rho}_{Ai}\eta_A) \\
&\quad - \sum_{i=0}^{N-1}\frac{EA}{l_0||\bar{l}||}(||\bar{l}|| - l_0)\hat{\rho}_{Ai}\widehat{\{\bar{l}_i\}_A}\eta_A - \sum_{i=0}^{N-1}\frac{1}{2}\hat{\rho}_{Ai}\widehat{\bar{R}_A^T\bar{F}_i}\eta_A \qquad (37)\\
\frac{1}{2}\hat{\rho}_B\bar{R}_{Bi}^T f_i &= J_B\ddot{\eta}_{Bi} + \left(-\widehat{J_B\bar{\Omega}_{Bi}} + \hat{\bar{\Omega}}_{Bi}J_B\right)\dot{\eta}_{Bi} + \left(-\widehat{J_B\bar{\Omega}_{Bi}}\hat{\bar{\Omega}}_{Bi} + \hat{\bar{\Omega}}_{Bi}J_B\hat{\bar{\Omega}}_{Bi}\right)\eta_{Bi} \\
&\quad + \frac{EA}{||\bar{l}||}\hat{\rho}_B\bar{R}_{Bi}^T\bar{R}_A\left(\frac{1}{||\bar{l}||^2}\{\bar{l}_i\}_A\{\bar{l}_i\}_A^T + \frac{||\bar{l}||}{l_0}I - I\right)(\epsilon_{Bi} - \bar{R}_A^T\bar{R}_{Bi}\hat{\rho}_B\eta_{Bi} - \epsilon_A + \hat{\rho}_{Ai}\eta_A) \\
&\quad + \frac{EA}{l_0||\bar{l}||}(||\bar{l}|| - l_0)\hat{\rho}_B\widehat{(\bar{R}_{Bi}^T\bar{R}_A\{\bar{l}_i\}_A)}\eta_{Bi} - \frac{1}{2}\hat{\rho}_B\widehat{\bar{R}_{Bi}^T\bar{F}_i}\eta_{Bi}. \qquad (38)
\end{aligned}$$

Here we denote the length vector of the i^{th} tether, expressed in the hub's body fixed frame, as

$$\{\bar{l}_i\}_A = \bar{R}_A^T\bar{l}_i. \qquad (39)$$

Note that $\{\bar{l}_i\}_A$ and the product $\bar{R}_{Bi}^T\bar{R}_A$ are constant in the reference state. Because $\bar{F}_i$ is along the axis of symmetry, $\bar{R}_A^T\bar{F}_i$ and $\bar{R}_{Bi}^T\bar{F}_i$ are constant too.

This system of equations can be written in a more compact format if we define

$$\begin{aligned}
X &= \begin{bmatrix} \epsilon_A^T & \eta_A^T & \epsilon_{B0}^T & \eta_{B0}^T & \epsilon_{B1}^T & \eta_{B1}^T & \dots \end{bmatrix}^T \qquad (40)\\
u &= \begin{bmatrix} f_0^T & f_1^T & \dots \end{bmatrix}^T \qquad (41)
\end{aligned}$$

such that

$$M\ddot{X} + C\dot{X} + KX = Du. \qquad (42)$$

M, C, and K are constants in our reference state with M diagonal, C nonsysmmetric, and K in general nonsymmetric. In contrast, D is time dependent. (Incidentally, D could be made constant if f_i were to be a vector expressed in the hub's body fixed frame.) If we define

$$Y = \begin{bmatrix} \dot{X} \\ X \end{bmatrix}, \qquad (43)$$

we can write this in first order form as follows

$$\begin{bmatrix} 0 & M \\ M & C \end{bmatrix}\dot{Y} + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}Y = \begin{bmatrix} 0 \\ D \end{bmatrix}u; \qquad (44)$$

see, for example, [7, 8]. Here the zeros denote the appropriately sized zero matrices.

Free vibrations of this system ($u = 0$) leads to the generalized eigenvalue problem

$$\begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} v_i = -\lambda_i \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} v_i \tag{45}$$

with eigenvalues λ_i and eigenvectors v_i. For this system, we expect all non-zero eigenvalues to be pure imaginary and come in complex conjugate pairs. The magnitude of the imaginary component of the eigenvalues represents natural frequencies and the eigenvectors represent complex modes of vibration. Because the system allows for rigid body modes, the system is defective (some eigenvectors will be repeated).

C. Discrete Time Equations of Motion

In this section we obtain discrete time equations of motion using a Lie group variational integrator similar to that outlined in [6]. Advantages of this approach include: energy is conserved (in the absence of external forces like F_i) and the structure of the direction cosine matrices R_A and R_{Bi} are preserved directly [6]. We begin by writing an expression for the discrete potential energy

$$U^k = \sum_{i=0}^{N-1} \frac{EA}{2l_0} \left(||l_i^k|| - l_0 \right)^2. \tag{46}$$

Here the superscript k is used to index through discrete points in time. We write a discrete approximation of the kinetic energy as

$$\begin{aligned} T^k = & \frac{1}{2\Delta t^2} m_A ||x_A^{k+1} - x_A^k||^2 + \sum_{i=0}^{N-1} \frac{1}{2\Delta t^2} m_B ||x_{Bi}^{k+1} - x_{Bi}^k||^2 \\ & + \frac{1}{\Delta t^2} \text{tr} \left[(I - \mathcal{F}_A^k) \tilde{J}_A \right] + \sum_{i=0}^{N-1} \frac{1}{\Delta t^2} \text{tr} \left[(I - \mathcal{F}_{Bi}^k) \tilde{J}_{Bi} \right]. \end{aligned} \tag{47}$$

Here the tilde denotes nonstandard inertia matrices defined as

$$\tilde{J}_A = \frac{1}{2} \text{tr}[J_A] I - J_A \tag{48}$$

$$\tilde{J}_{Bi} = \frac{1}{2} \text{tr}[J_{Bi}] I - J_{Bi}. \tag{49}$$

As in [6], $\mathcal{F}_A^k$ and $\mathcal{F}_{Bi}^k$ are rotations matrices that map changes from R_A^k to R_A^{k+1} and R_{Bi}^k to R_{Bi}^{k+1} as follows

$$R_A^{k+1} = R_A^k \mathcal{F}_A^k \tag{50}$$

$$R_{Bi}^{k+1} = R_{Bi}^k \mathcal{F}_{Bi}^k. \tag{51}$$

Furthermore, $\mathcal{F}_A^k$ and $\mathcal{F}_{Bi}^k$ are used to approximate angular velocities Ω_A^k and Ω_{Bi}^k as follows

$$\hat{\Omega}_A^k = \frac{1}{\Delta t} \left(\mathcal{F}^k - I \right) \tag{52}$$

$$\hat{\Omega}_{Bi}^k = \frac{1}{\Delta t} \left(\mathcal{F}_{Bi}^k - I \right). \tag{53}$$

The discrete approximation for the non conservative virtual work due to interactions with the solar wind is written as

$$\delta W_{\text{nc,wind}}^k = \sum_{i=0}^{N-1} \frac{1}{2} (F_i^k)^T \left(\delta x_A^k + R_A^k \hat{\eta}_A^k \rho_{Ai} + \delta x_{Bi}^k + R_{Bi}^k \hat{\eta}_{Bi}^k \rho_B \right). \tag{54}$$

We write a discrete version of the extended Hamilton's principle as follows

$$0 = \sum_{k=0}^{N_{\text{time}}-1} \left(\delta T^k + \frac{1}{2} \left(-\delta U^k - \delta U^{k+1} + \delta W_{\text{nc,wind}}^k + \delta W_{\text{nc,wind}}^{k+1} \right) \right) \Delta t, \tag{55}$$

where N_{time} is the number of discrete time points. Analogous to the continuous time case, the addends of Equation 7 can be expressed in terms of the following independent set of variations: δx_A, δ_{Bi}, η_A, and η_{Bi}. The coefficients of these variations yield the discrete equations of motion with which we include Equations 50 and 51 to obtain

$$x_A^{k+1} = 2x_A^k - x_A^{k-1} + \frac{\Delta t^2}{m_A}\sum_{i=0}^{N-1}\left(\frac{EA}{l_0}(||l_i^k|| - l_0)\frac{1}{||l_i^k||}l_i^k + \frac{1}{2}F_i^k\right) \tag{56}$$

$$x_{Bi}^{k+1} = 2x_{Bi}^k - x_{Bi}^{k-1} - \frac{\Delta t^2}{m_B}\left(\frac{EA}{l_0}(||l_i^k|| - l_0)\frac{1}{||l_i^k||}l_i^k - \frac{1}{2}F_i^k\right) \tag{57}$$

$$R_A^{k+1} = R_A^k \mathcal{F}_A^k \tag{58}$$

$$R_{Bi}^{k+1} = R_{Bi}^k \mathcal{F}_{Bi}^k \tag{59}$$

$$\begin{aligned}\mathcal{F}_A^{k+1}\tilde{J}_A - \tilde{J}_A(\mathcal{F}_A^{k+1})^T = &-\left(-\tilde{J}_A\mathcal{F}_A^k + (\mathcal{F}_A^k)^T\tilde{J}_A\right) - \Delta t^2\sum_{i=0}^{N-1}\frac{1}{2}\left(\rho_{Ai}(F_i^{k+1})^T R_A^{k+1} - (R_A^{k+1})^T F_i^{k+1}\rho_{Ai}^T\right)\\ &-\Delta t^2\sum_{i=0}^{N-1}\frac{EA}{l_0}(||l_i^{k+1}|| - l_0)\frac{1}{||l_i^{k+1}||}\left(\rho_{Ai}(l_i^{k+1})^T R_A^{k+1} - (R_A^{k+1})^T l_i^{k+1}\rho_{Ai}^T\right)\end{aligned} \tag{60}$$

$$\begin{aligned}\mathcal{F}_{Bi}^{k+1}\tilde{J}_B - \tilde{J}_B(\mathcal{F}_{Bi}^{k+1})^T = &-\left(-\tilde{J}_B\mathcal{F}_{Bi}^k + (\mathcal{F}_{Bi}^k)^T\tilde{J}_B\right) - \Delta t^2\frac{1}{2}\left(\rho_B(F_i^{k+1})^T R_{Bi}^{k+1} - (R_{Bi}^{k+1})^T F_i^{k+1}\rho_B^T\right)\\ &+\Delta t^2\frac{EA}{l_0}(||l_i^{k+1}|| - l_0)\frac{1}{||l_i^{k+1}||}\left(\rho_{Bi}(l_i^{k+1})^T R_B^{k+1} - (R_B^{k+1})^T l_i^{k+1}\rho_{Bi}^T\right).\end{aligned} \tag{61}$$

After using initial conditions to prescribe values of the variables at time steps 0 and 1, Equations 56-61 are solved in order for $k = 1$ and then repeated for $k = \{2,\ 3,\ ...\}$. Of these equations, only Equations 61 and 60 are implicit. Fortunately, however, each represents a nonlinear system of three scalar equations in three scalar unknowns which can be solved using the strategy outlined in [6].

As written, this discrete time model nonphysically allows tethers to support compressive loads when shortened. However, in our computational implementation we incorporate a Heaviside function to cancel the elastic terms of individual tethers when their length is less than l_0.

III. Results and Discussion

In preparation for designing a scalable multitether E-sail for a technical demonstration mission, here we present results for a system consisting of $N = 3$ aluminum tethers (E =70 GPa) with cross sectional area A=11,160 μm^2 and length l_0 =8 km. (Following [9], we assume the tether safely supports stesses up to about 10 MPa.) The hub is represented by a sphere of radius 0.113 m, uniform density, and mass m_A =12 kg. The pods are 3U cubesats with uniform density and mass m_B = 6 kg. The inertia matrices are

$$J_A = \frac{2}{5}(12\text{ kg})(0.113\text{ m})^2 I \tag{62}$$

$$J_B = \frac{1}{3}(6\text{ kg})(0.05\text{ m})^2\begin{bmatrix}10 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 10\end{bmatrix}. \tag{63}$$

Tether attachment points are evenly distributed around the equator of the hub and at the center of a long side of each pod such that

$$\rho_{Ai} = (0.113\text{ m})\begin{bmatrix}\cos\left(\frac{2\pi i}{N}\right) & \cos\left(\frac{2\pi i}{N}\right) & 0\end{bmatrix}^T\text{ m} \tag{64}$$

$$\rho_B = (0.05\text{ m})\begin{bmatrix}-1 & 0 & 0\end{bmatrix}^T\text{ m}. \tag{65}$$

The reference state angular velocity of the system is $\bar{\Omega}$ = 0.00088 rad/s. The parameter values presented in this section are rough estimates for what might fly on a technical demonstration mission rather than the results of a detailed design process.

$\lvert\text{Imag}\{\lambda\}\rvert$	Dominant motions
$0.000\bar{\Omega}$	rigid body translation along sail axis
$0.000\bar{\Omega}$	rigid body rotation about sail axis
$0.354\bar{\Omega}$	hub follows a circular path and pods follow an elliptical motion, nearly planar (Figure 2(a))
$0.888\bar{\Omega}$	3 independent modes in which the pods rotate about their body fixed x-axes
$0.999\bar{\Omega}$	out of plane motion in which the pods follow a pendulum-like motion (Figure 3(a))
$1.000\bar{\Omega}$	2 rigid body modes corresponding to translation within the plane
$1.001\bar{\Omega}$	out of plane motion in which the pods follow a pendulum-like motion (Figure 3(b))
$1.212\bar{\Omega}$	hub follows a circular path and pods follow an elliptical motion, nearly planar (Figure 2(b))
$1.582\bar{\Omega}$	Hub and pods translate along the axis of rotation out of phase with one another (Figure 4)
$144.9\bar{\Omega}$	In phase extensions of all tethers (Figure 5(a))
$191.2\bar{\Omega}$	Multi-phase extension of the tethers and the hub follows a circular path (Figure 5(b))
$192.0\bar{\Omega}$	Multi-phase extension of the tethers and the hub follows a circular path (Figure 5(c))
$219.9\bar{\Omega}$	3 modes in which the pods rotate about their body fixed z axes
$364.3\bar{\Omega}$	2 modes in which the hub rotates about axes in it's body fixed $x-y$ plane
$492.0\bar{\Omega}$	3 modes in which the pods rotate about their body fixed y axes
$515.2\bar{\Omega}$	rotation about the hub's body fixed z axis

Table 1 Summary of E-sail modes of vibration for $||\bar{F}_i|| = 0.00744$ N.

A. Linear Modes of Vibration

Following the approach developed in Section II.B and using parameter values described above and $||\bar{F}_i|| = 0.00744$ N, we obtain the modes of free vibration outlined in Table 1. (Incidentally, results qualitatively similar to those presented in this section can be obtained for $||\bar{F}_i|| = 0$ N; data not shown.) Although the eigenvalues/frequencies presented in the table are given in terms of the angular velocity $\bar{\Omega}$, it does not necessarily mean that the frequencies scale linearly with the angular velocity. Illustrations of some of these modes appear in Figures 2-5. In these figures, modal amplitudes are exaggerated to show detail; the analysis is valid for small deflections only. In fact, the linear model dictates that tethers support compression upon shortening. Because this is nonphysical, modal amplitudes must not change the length of the tethers by more than they are stretched in the reference state; here that is 0.3807 m.

Although rigid body modes often correspond to a frequency of 0 rad/s, here the frequency associated with rigid body translation in the plane is $\bar{\Omega}$. This is an artifact of our choice to use displacement coordinates, ϵ_A and ϵ_{Bi}, relative to a body fixed frame attached to the hub in the reference state. Also the modes illustrated in Figures 2-5 are drawn in this body fixed frame. In the inertial frame, a rigid body rotation about the spin axis and translation along the spin axis would be superimposed with these modes drawn relative to the body fixed frame.

The lowest frequency for a non rigid body mode, $0.354\bar{\Omega}$, corresponds to the motion illustrated in Figure 2(a). For this mode, out of plane motion is negligible. Incidentally, if amplitudes of this mode were to become large (well beyond the assumed small deflections), the pods could collide and result in a catastrophic failure. For systems with a larger number of tethers $N > 3$, the amplitude required for collision, in an analogous mode, would be significantly reduced. Consequently, auxiliary tethers [9], which connect neighboring pods, have been proposed to prevent collisions. Further work is needed to determine how these tethers might change the vibrations of the system. A similar mode depicted in Figure 2(b), might pose less of a concern because its frequency is larger, $1.212\bar{\Omega}$, and perhaps less likely to experience large amplitudes.

Table 1 identifies three modes with frequency $0.888\bar{\Omega}$ corresponding primarily to rotations about the body fixed x-axes of the pods. If the principal moments of inertia of the pods were identical (e.g., spherical pods), these modes would become rigid body modes with frequency $0\bar{\Omega}$. Because the torsional stiffness of the tethers is neglected, the stiffness of these modes comes from gyroscopic effects. If we were to incorporate torsional stiffness into our model, the small cross section and large length of the tethers suggests the stiffness would be very low and therefore have little impact on vibrations. Interestingly, if the principal moments of inertia of the pods about the y and z axes are swapped

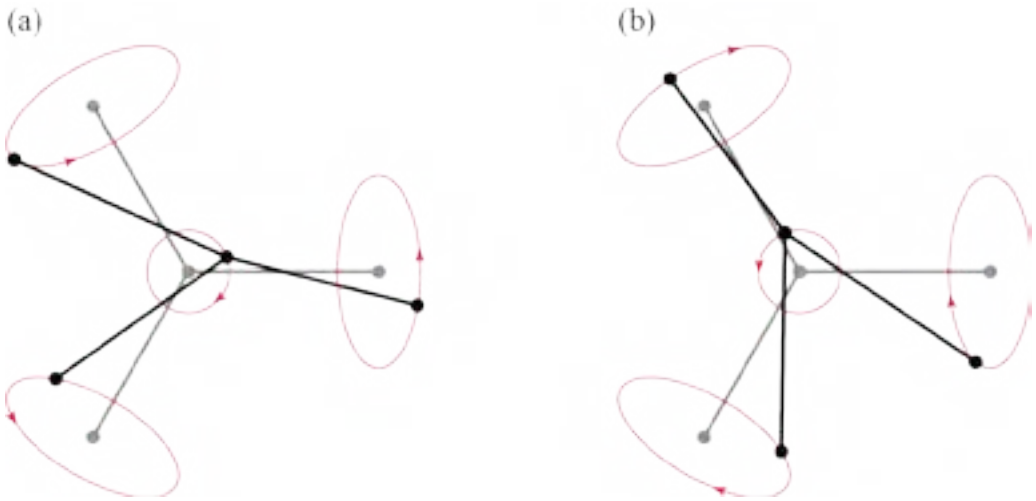


Fig. 2 Two nearly planar modes in which the hub and pods follow circular and elliptical paths (red) respectively, each relative to the reference state (gray).

the system becomes infinitesimally unstable. To justify this claim, consider a composite rigid body consisting of a mass m with principal moments of inertia (j_1, j_2, j_3) welded to an extremely long, massless, and *rigid* tether of length l_0. The components are welded such that the length of the rigid tether is colinear with the axis associated with j_1. The tether is supported at the end opposite the mass by a spherical joint. The inertia matrix for the composite body about the end of the tether is

$$J = \begin{bmatrix} j_1 & 0 & 0 \\ 0 & j_2 + ml_0^2 & 0 \\ 0 & 0 & j_3 + ml_0^2 \end{bmatrix}. \tag{66}$$

We can linearize the Euler equations for free motion about a reference state in which the system rotates about its third principal axis with angular velocity $\bar{\Omega}_3$ and obtain

$$J\dot{\bar{\omega}} + \begin{bmatrix} 0 & \bar{\Omega}_3(j_3 - j_2) & 0 \\ \bar{\Omega}_3(j_1 - j_3 - ml_0^2) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \omega = 0. \tag{67}$$

Here ω is a vector of small changes in angular velocity. The nonzero eigenvalues of the matrix are

$$\lambda^2 = \bar{\Omega}_3^2(j_3 - j_2)(j_1 - j_3 - ml_0^2). \tag{68}$$

Because l_0 is assumed to be large, $\lambda^2 < 0$ and the system is stable when $j_3 > j_2$. Aside from the potential for instability when $j_3 \leq j_2$ and rigid body modes (which have repeated eigenvectors), the system appears to be stable for small vibrations about the reference state. That is, the real part of each eigenvalue is effectively 0 (on the scale of computer precision) and eigenvectors for nonrigid body modes are linearly independent.

There are two modes in which the pods oscillate out of plane with distinct phases; see Figure 3. The frequencies of oscillation are slightly above and below the angular velocity of the reference state. Interestingly, we can estimate the frequency by modeling a single pod on a tether as a pendulum under gravitational acceleration which has a natural frequency of $\omega_n = \sqrt{\frac{g}{l_0}}$ for small oscillations. Specifically, substituting centripetal acceleration, corresponding to an angular velocity of $\bar{\Omega}$, for gravitational acceleration yields the frequency estimate

$$\omega_n \approx \sqrt{\frac{\bar{\Omega}^2 l_0}{l_0}} = \bar{\Omega}. \tag{69}$$

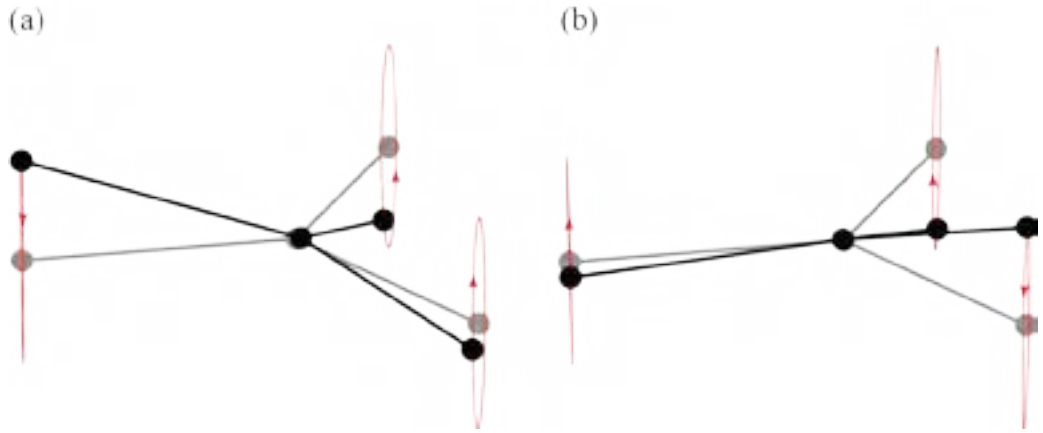


Fig. 3 Two out of plane modes in which pods oscillate like pendulums relative to the reference state (gray).

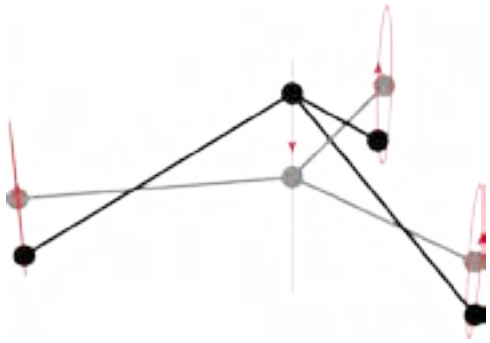


Fig. 4 Out of plane mode in which hub and pods oscillate out of phase with eachother along the axis of the e-sail relative to the reference state (gray).

Figure 4 shows a mode in which the hub and pods move out of plane with the pods out of phase with the hub. (This mode seems to be analogous to the 'flapping mode' identified in [10] and the oscillations in acceleration observed in [11].) Applying the Rayleigh-Ritz method to out of plane vibrations in which all tethers experience the same motion (symmetric), provides insight. To do so, we represent the hub and pod as point masses and neglect the stretch on the tether as well as the thrust force. Rotation of the sail with angular velocity $\bar{\Omega}$ results in the following expression for the tension in the tether as a function of position along its length s,

$$F_{\text{tension}}(s) = \bar{\Omega}^2 l_0 \left(\frac{1}{2} m_L l_0 + m_B\right) - \frac{1}{2} m_L \bar{\Omega}^2 s^2. \tag{70}$$

Here m_L is the mass per unit length of the tether. An arbitrary trial function, $\phi(s)$, representing a possible displacement on one of the tethers yields the Rayleigh quotient

$$R(s) = \frac{\int_0^{l_0} F_{\text{tension}}(s) \left(\phi'(s)\right)^2 ds}{\int_0^{l_0} m_L \left(\phi(s)\right)^2 ds + \frac{m_A}{N} \left(\phi(0)\right)^2 + m_B \left(\phi(l_0)\right)^2}, \tag{71}$$

which can be used to estimate the natural frequencies and mode shapes of a system; see, for example, [8]. Taking $m_L = 1.512$ mg/m and representing $\phi(s)$ with the set of all polynomials of order 9 yields an estimate for the lowest nonzero natural frequency of $1.584\bar{\Omega}$, closely matching the computed frequency for mode in Figure 4, $1.582\bar{\Omega}$. The corresponding mode is well represented by a first order polynomial. The next highest mode in this Rayleigh-Ritz approach is an order of magnitude higher, $15.90\bar{\Omega}$. Incidentally, the speed of a transverse wave traveling along a taught string, c, can be written as

$$c = \sqrt{\frac{F_{\text{tension}}}{m_L}}. \tag{72}$$

Assuming constant tension, $F_{\text{tension}} = \bar{\Omega}^2 l_0 m_B$, and using $m_L = 30.24$ mg/m yields $c = 35.06$ m/s. Although it would take nearly 4 minutes for a disturbance to travel from one end of the 8 km tether to the other, this is fast compared to how long it takes for the system to rotate, about 2 hr. In summary, if lower frequency modes are of primary interest, it appears reasonable to neglect the distribution of mass along the length of the tethers as we did with the models presented in Section II.

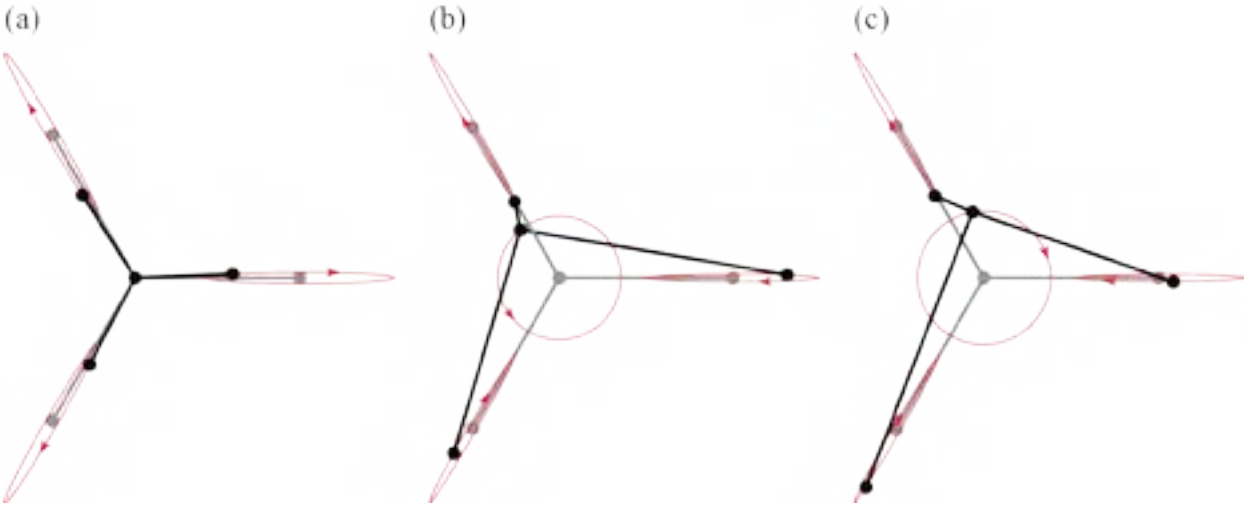


Fig. 5 Three nearly planar modes in which the tethers change length, each relative to the reference state (gray).

A simple model can also be used to predict the frequency of vibration of a mode in which all the tethers change length in phase with each other; see Figure 5(a). Specifically, we can think of each tether and pod as a mass spring system for which the natural frequency is

$$\omega_n \approx \sqrt{\frac{EA}{l_0 m_B}} \tag{73}$$

$$\approx 145.0\bar{\Omega}, \tag{74}$$

which is very close to the computed value, $144.9\bar{\Omega}$. Again, the modal amplitude shown in Figure 5(a) is beyond the assumed small displacements. In fact, if the amplitude were to approach that depicted in the Figure, the tether would become slack rather than support a compressive force as dictated by the potential energy function used to represent the system. We would expect the three extensional modes to be accurate as long as changes in length of the tether from the reference state remain smaller than the extension in the tethers in the reference state, 0.3807 m. However, we would expect that the other nonextentional modes could accurately represent displacements up to about 5-10% of l_0. Interestingly, the modes in Figure 5 show the pods following an elliptical path rather than a straight in and out motion. Conservation of angular moment explains this observation; when the tethers shorten the mass moment of inertia of the system is reduced and the angular velocity necessarily increases. Two modes in which the tethers extend out of phase with one another are depicted in Figures 5(b) and (c).

Another simple model can provide insight into modes in which the pods rotate with negligible translational motion. Specifically, we consider a disk shaped mass of mass m, radius ρ, mass moment of inertia J, and angle of rotation θ at the end of a very long tether of length l_0 ($l_0 \gg \rho$) subject to the centripetal acceleration along the length of the tether corresponding to an angular velocity of $\bar{\Omega}$. This leads to the following governing equation (for small rotations θ) and natural frequency

$$0 = J\ddot{\theta} + m l_0 \bar{\Omega}^2 \rho \theta \tag{75}$$

$$\omega_n = \sqrt{\frac{m l_0 \bar{\Omega}^2 \rho}{J}}. \tag{76}$$

Substituting $m = m_B$ and $\rho = ||\rho_B||$ and $J = [J_B]_{33}$ for rotation about a pod's z-axis or $J = [J_B]_{22}$ for rotation about a pod's y axis yields frequencies of $219.1\bar{\Omega}$ and $489.9\bar{\Omega}$ respectively. These values closely match the computed values of

$219.9\bar{\Omega}$ and $492.0\bar{\Omega}$; see Table 1. In addition, using $m = Nm_B$, $\rho = ||\rho_{Ai}||$, and $J = [J_A]_{33}$ for rotation about the hub's z-axis accurately predicts the frequency of vibration for hub rotation about z, $515.3\bar{\Omega}$ vs. $515.2\bar{\Omega}$; see Table 1.

B. Possible Instability Identified in Discrete Time Simulation

Here we employ our discrete time equations of motion (see Section II.C) to simulate dynamics of a multitether system. In this section, $F_i = 0$ N, $\Delta t = 0.01$ s, the tethers are initially slackened from their free length by 0.04 m (stretched by -0.04 m), and the initial angular velocity of the system is 0.00088 rad/s. In Figure 6 we plot the stretch in one tether and the angle about the z-axis of the corresponding pod relative to the hub as functions of time. As expected given the mode described by Figure 5(a), the stretch in the tether oscillates with what appears to be constant amplitude about the reference state value. Surprisingly, the angle of rotation of the pod oscillates with increasing amplitude for the duration of this simulation. Although the amplitude is small and the simulation doesn't prove that it will continue to grow, this behavior suggests the existence of a possible instability in the system. If this response is due to computer round-off, we argue that a similar response might occur for physical disturbances or manufacturing imperfections. Reducing Δt by an order of magnitude results in negligible changes to data presented in Figure 6, giving further evidence for an instability. Interestingly, each time the tether length reaches a minimum (corresponding to low tension), the amplitude of oscillations of the relative angle appears to make a step increase; see arrows in Figure 6(below). (Similar behavior results when the cross sectional area of the tethers is reduced about 20-fold and the initial stretch is 0.1 m such that the tethers remain taught.) This suggests that the possible instability is not a result of the inability of the tether to support compressive loads. Instead, there appears to be some nonlinear coupling between hub/pod rotational modes with tether extension modes (Figure 5). Inherent structural damping, which is expected to be small and therefore neglected in the current model, could be sufficient to inhibit growth of the rotations. That said, we believe maintaining tension will prove to be an important requirement for E-sail flight control systems. We expect increasing voltage to be an effective means of eliminating slack on a tether; studying this will require the model to be extended to account for transverse deformations of the tethers. Finally, increasing the distances $||\rho_A||$ and/or $||\rho_B||$, with the addition of rigid booms on hub and/or pods to which the tethers are attached, could alter the frequency and amplitude of rotational modes and possibly alleviate the instability.

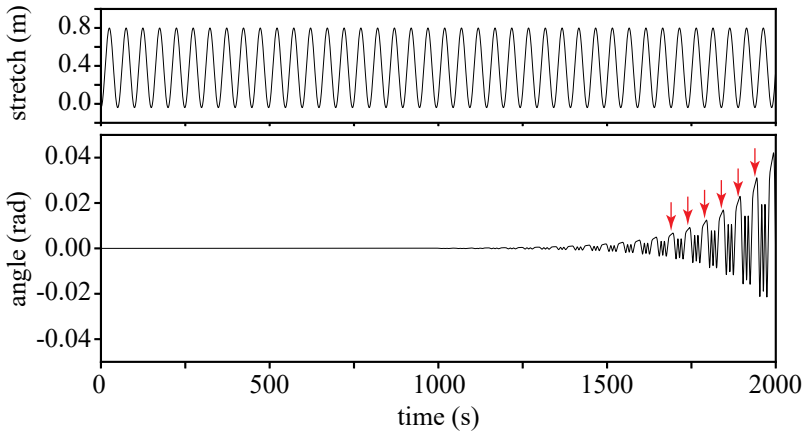


Fig. 6 (above) Stretch of a single tether as a function of time. (below) Relative angle of the corresponding pod with respect to the hub. Arrows mark times at which the stretch is negative (i.e., slack in the tether).

C. Plane of Rotation (and Thrust Vector) Can Be Directed

To simulate a steering maneuver resulting from varying the voltage on the tethers, we vary the force on the i^{th} tether in proportion to the x component of distance of its pod from the hub, $(x_{Bi} - x_A)_x$. Physically, the force on a tether from the solar wind acts perpendicular to its length; see, for example [9]. Therefore, we define n_i as a unit vector in the direction of the component of the solar wind perpendicular to the length of the tether such that we can write the force as

$$F_i = n_i n_i^T \bar{F} \frac{(x_{Bi} - x_A)_x}{||x_{Bi} - x_A||}. \tag{77}$$

Here $\bar{F}$ is a vector with magnitude 0.00744 N in the positive z direction of the inertial frame. Here we allow the solar wind to pull as well as push on the tethers. Although this is not physically possible, it facilitates plotting the steering maneuver because the mass center is relatively stationary. In reality a constant offset could be superimposed on the force which would eliminate the pulling and accelerate the center of mass of the system.

Figure 7(a) displays the trajectory of the three pods with this force. Clearly it is possible to change the plane of the sail. However, here changing the plane of the sail is accomplished by applying a torque, and thereby introducing angular momentum, along the negative y-direction. With the addition of this component of angular moment, which is perpendicular to the initial angular momentum about the z-axis, the net angular momentum of the system increases with time. As a result, in this simulation the stretch, and therefore the tension, in the tethers increases with time; see Figure 7(b). Tethers must be designed to withstand changes in the system's spin rate. This simulation also highlights the fact that adjusting the voltage on the tethers cannot change the component of angular momentum along the direction of the solar wind. However, if the E-sail were to be in orbit around the sun, the relative direction of the solar wind rotates through 360° suggesting that the angular momentum vector is controllable on the time scale of an orbit; see, for example [12]. It has been estimated that an E-sail would be capable of directing its thrust vector by about 30° [9].

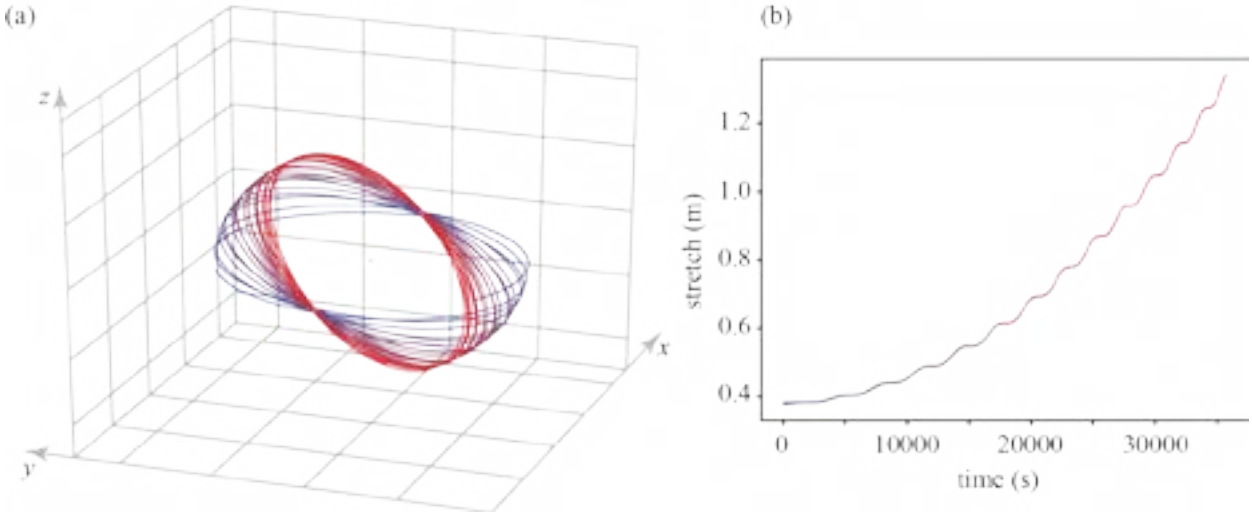


Fig. 7 (a) Trajectory of pods during a steering maneuver. Time is shown by the color change from blue to red. (b) Stretch of a single tether as a function of time.

IV. Future Work

This paper makes a significant step forward towards understanding the dynamics of E-sails, however there remains a lot of work to do before we can prove E-sail dynamics can be controlled throughout a mission. As a first step, we propose adding point masses along the length of the tethers to better account for their continuous distribution of mass and stiffness. Another straightforward step would be to address controllability and observability of each mode for the linearized model. The effects of variable solar wind could also be studied with the linearized model. Other, more challenging steps include: characterizing stability due to finite disturbances, model deployment dynamics, and develop control strategies.

Acknowledgments

This work is the direct result of a NASA Marshal Faculty Fellowship awarded to T. D. L. for the summer of 2019. T. D. L. thanks everyone on the E-sail team at Marshall Space Flight Center, especially John Rakoczy who served as his host. In addition, T. D. L. thanks Carlos Montalvo, a colleague at the University of South Alabama, for introducing E-sails dynamics to him and freely sharing his expertise and code. Finally, T. D. L. thanks his wife and 4 children for relocating to Huntsville with him for the summer.

References

[1] Kohlhase, C. E., and Penzo, P. A., "Voyager mission description," *Space Science Reviews*, Vol. 21, No. 2, 1977, pp. 77–101. doi:10.1007/BF00200846, URL https://doi.org/10.1007/BF00200846.

[2] Wenzel, K.-P., Marsden, R. G., Page, D. E., and Smith, E. J., "The Ulysses mission," *Astronomy and Astrophysics Supplement Series*, 1992.

[3] Janhunen, P., "Electric Sail for Spacecraft Propulsion," *Journal of Propulsion and Power*, Vol. 20, No. 4, 2004, pp. 763–764. doi:10.2514/1.8580, URL https://doi.org/10.2514/1.8580.

[4] Janhunen, P., "Electric sail for producting spacecraft propulsion," , 2010.

[5] Janhunen, P., and Sandroos, A., "Simulation study of solar wind push on a charged wire: basis of solar wind electric sail propulsion," *ANGEO*, Vol. 25, No. 3, 2007, pp. 755–767. URL https://www.ann-geophys.net/25/755/2007/.

[6] Lee, T., Leok, M., and McClamroch, N. H., "Lie group variational integrators for the full body problem in orbital mechanics," *Celestial Mechanics and Dynamical Astronomy*, Vol. 98, No. 2, 2007, pp. 121–144. doi:10.1007/s10569-007-9073-x, URL https://doi.org/10.1007/s10569-007-9073-x.

[7] Gutierrez-Wing, E. S., "Modal analysis of Rotating Machinery Structures," Ph.D. thesis, University of London, 2003.

[8] Meirovitch, L., *Fundamentals of Vibrations*, Waveland Press, Inc., 2001.

[9] Janhunen, P., Toivanen, P. K., Polkko, J., Merikallio, S., Salminen, P., Haeggström, E., Seppänen, H., Kurppa, R., Ukkonen, J., Kiprich, S., Thornell, G., Kratz, H., Richter, L., Krömer, O., Rosta, R., Noorma, M., Envall, J., Lätt, S., Mengali, G., Quarta, A. A., Koivisto, H., Tarvainen, O., Kalvas, T., Kauppinen, J., Nuottajärvi, A., and Obraztsov, A., "Invited Article: Electric solar wind sail: Toward test missions," *Review of Scientific Instruments*, Vol. 81, No. 11, 2010, p. 111301. doi:10.1063/1.3514548, URL https://doi.org/10.1063/1.3514548.

[10] Montalvo, C., and Wiegmann, B., "Electric sail space flight dynamics and controls," *Acta Astronautica*, Vol. 148, 2018, pp. 268 – 275. doi:https://doi.org/10.1016/j.actaastro.2018.05.009, URL http://www.sciencedirect.com/science/article/pii/S0094576517318313.

[11] Liu, F., Hu, Q., and Liu, Y., "Attitude Dynamics of Electric Sail from Multibody Perspective," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 12, 2018, pp. 2633–2646. doi:10.2514/1.G003625, URL https://doi.org/10.2514/1.G003625.

[12] Toivanen, P. K., and Janhunen, P., "Spin Plane Control and Thrust Vectoring of Electric Solar Wind Sail," *Journal of Propulsion and Power*, Vol. 29, No. 1, 2013, pp. 178–185. doi:10.2514/1.B34330, URL https://doi.org/10.2514/1.B34330.

Evaluation of a Piezo-ceramic Sensor

Juan M. Lorenzo[1]
Louisiana State University, Baton Rouge, Louisiana, 70803, U.S.A.

Donald A. Patterson [2]
Ario Labs LLC, Baton Rouge, Louisiana, 7081, U.S.A.

Renee Weber [3]
NASA Marshall Space Flight Center, Huntsville, 35805, U.S.A.

I. Abstract

Piezo-technology is widely used in the defense, aerospace and structural engineering fields. Current, off-the-shelf piezo-ceramic-type sensors can be used to develop miniature seismological instruments that permit non-invasive, shallow (< 1 m) high-resolution (10 cm) characterization of the regolith/soil profile on planetary bodies. We compare the signal-to-noise performance of a piezo-ceramic-type accelerometer to a proven, piezo-polymer-type sensor.

II. Introduction

A. Significance

Water is key for supporting future human missions on the Moon as well as Mars, and because buried H_2O ice can stiffen near-surface materials on these inner planetary bodies, seismic estimates of their strength have the potential to characterize the volume and distribution of buried H_2O ice. Specifically, high-frequency (kHz) seismic, piezo-electric sounding systems have the potential to reduce the ambiguity of in-situ resource mapping of the upper few m of regolith and soils and can complement other non-invasive methods which may not be able to penetrate as deeply (e. g., gamma, neutron, & radar spectroscopy). For this purpose, we envisage small arrays of piezo-sensors and piezo-sources integrated structurally into the landing pads of static landers (Figure 1) or rover wheels (Figure 2). A low-power and low-footprint seismic sub-system may characterize key aspects of the H_2O(s)-ice reservoirs foremost via their seismic velocities, and seismic attenuation calibrated to geomechanical and geotechnical properties.

Buried H_2O-ice reservoir characteristics such as depth, lateral continuity, distribution and degree of purity may enhance our understanding of volatile transport processes and growth within the lunar regolith and crust. Regolith in permanently shadowed regions of the Moon is key to understanding the genesis of the sequestered ice portion of the Lunar 'water' cycle[1] and to improving its potential as a sustainable resource for use by humans. Our focus lies within first few meters of the surface where potential ice-mining will be most feasible. In-situ characterization of H_2O-ice using landers or rovers, prior to excavation, helps preserve the geologic record of volatile deposition, and map sites suitable for mining.

---

[1] Professor, Department of Geology and Geophysics, Louisiana State University.
[2] Ario Labs LLC, 4055 Overlook Point Ct. Baton Rouge, LA 70817-1620, U.S.A.
[3] Chief Scientist, NSSTC 2047, 320 Sparkman Dr., Huntsville, AL 35805, U.S.A

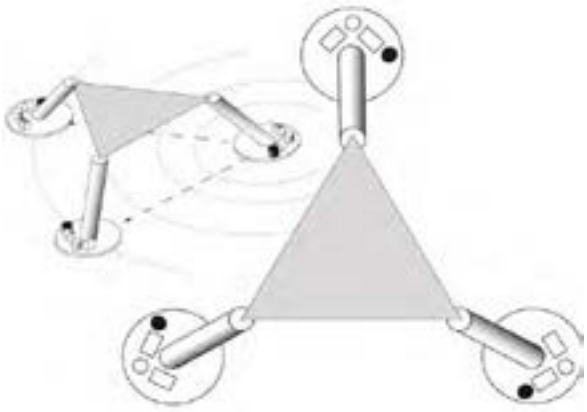


Figure 1. LEFT: Perspective view of conceptual lunar lander with payload area (triangle). Estimated ~ 2 m between landing pads. Piezo-actuators (black) create seismic pulses that travel between pads and sample the upper few meters of the lunar regolith. RIGHT: Three equally distributed piezo-sensors and one vertical piezo-actuator (green) are installed on each of 3 pads.

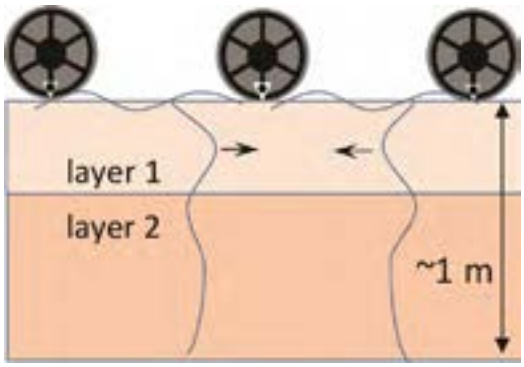


Figure 2. In concept, sensors (white triangles) and actuators (black circles) integrated into rover wheels can both generate and measure surface waves (blue curves). Multiple devices can be incorporated into each wheel. Seismic data are collected when the rover is stopped. Several seconds is sufficient to record data. Calibration against laboratory standards[2] may constrain estimates of ice-soil concentration.

Small piezo-electric sensors (1 cm x 1 cm footprint) can be used to extract soil properties under controlled laboratory conditions. However, redesign of these systems is needed for deployment under a much wider and more extreme range of physical conditions (e.g., temperature, radiation, accelerations). In a step toward space readiness, herein we evaluate and compare the performance of more suitable types of ceramic piezo-sensors against a standard high-performance polymer-based sensor.

B. Background to piezoelectric sensors and seismic instrumentation

Piezoelectric materials will produce an electric charge when stressed and conversely will change shape when subjected to an electric field. Piezoelectric materials have been widely used over the past 100 years, e.g., in radios and sonar. They have seen regular use since World War II in measuring and detecting shock waves from explosions (including atomic detonations) in environments up to 700°C (e.g.,[3,4,5]).

In general, polymer-based piezoelectric transducers (polyvinylidene fluoride-PVDF) are about an order of magnitude more sensitive than piezo-ceramic and natural crystals, but they lose sensitivity and are expected to become brittle as they transition to a glassy-brittle state (< -35 ºC), under space-temperature conditions. However, because they have been in use for decades in geotechnical soil studies[6,7] they do provide convenient and low-cost laboratory analogs (Figure 3).

Nevertheless, compared to PZT sensors, PVDF types have a limited operational temperature range[8]. Typically, they operate from -40°C to 85°C. Above 120°C the polymer starts to approach its Curie temperature and melts around 170°C. Below 0°C the d_{3n} parameters drop off sharply. In liquid Nitrogen (77°K), PVDF loses more than 99% of its sensitivity compared to PZT type sensors. At 77°K, the PZT also loses about 58% of its sensitivity but overall is still approximately 600 times more sensitive than the PVDF film. In addition, PZT and other types of ceramic sensors have Curie temperatures in excess of 200°C. Newer ceramic materials have even higher Curie temperatures, and for these reasons, ceramic type accelerometers are more appropriate for external mounting on a lunar lander.

Moreover, synthetic, piezo-transducers made of ceramic lead zirconate titanate materials (PZT, e.g., Figure 3) have long been under consideration by NASA[9]. Piezoelectric ceramics such as those that comprised the penetrometer (PZT-5A) on the Huygens lander were used to determine grain size of the surface of Titan. This material was chosen because of its sensitivity and durability to space temperatures and prolonged exposure to radiation levels over its 7-year mission. As well, for piezo-actuators which act as seismic source generators, lead magnesium niobate (PMN) and

PMN-PT (titanate) are examples of well-known electrostrictive materials (e. g., [10]) currently available commercially (e. g., TRS technologies) for use especially in cryogenic (NIST-below 180° C) conditions[11].

Previous successful use of piezo-ceramics in space on the Cassini-Huygens probe on Titan[12] their current technological maturity, low cost, low power consumption (μA), low mass (Table 1: e.g., 4.5 g) make them novel candidates for a technological leap that was not available during the Apollo era. Although piezo-sensors have 3-4 orders of magnitude less sensitivity when compared to the nearest equivalent MEMS sensor technology of the shortest-period seismometers (SEIS-P[13] : 0. 5 ng/√Hz @ 10Hz) of the InSight mission to Mars, SEIS-P has lower bandwidth (~40 Hz) was not designed to image the soils of Mars but primarily, to remove environmental noise and aid the longer-period seismometer to detect mars-global seismic events. Piezo-sensors have several advantages because they have a broader bandwidth (kHz) and finer resolution than Apollo missions that allows them to detect properties of the shallow regolith structure (< 3 m) plus they can be coupled with nearby, conveniently small (1 cm^3) piezo-actuator seismic sources.

In order to determine the shallow structure (~ 300 m) of the lunar regolith, Apollo 14, 16 and17 missions all employed active thumping (by astronauts) or explosions in their experiments[14]. But an envisaged static lunar lander will not have the ability to deploy an array of seismometers as did Apollo and a Huygen's-type probe will not be capable of determining geomechanical properties down to a few meters. Although outside the limit of this report, piezo-electric actuators can also fill that role with a small size (Figure 3).

C. Brief Outline

A PZT accelerometer is expected to be far more sensitive than a PVDF-polymer-based accelerometer under cold space conditions (liquid Nitrogen). Nevertheless, because these cold temperatures can reduce the PZT by about half, our goal herein is to evaluate whether we can increase the signal-to-noise ratio sensitivity of a PZT accelerometer, with reference to a PVDF polymer at room temperature, by introducing a newly designed differential charge amplifier. If so, then we expect that in a future stage, the same new electronics will also be tested at liquid Nitrogen conditions.

III. Tools and Methods

A. Background to Seismic Piezo-electric sensors

In previous upper soil seismic measurements[2,7] the electronics are designed around the TE Connectivity ACH-01 PVDF type accelerometer with an integrated JFET (Figure 3E, Table 1). The PVDF-type sensor has a room temperature sensitivity of 10 mV/g and a low acoustic impedance for better coupling to loose soil types. The ACH-01 has three lines from the sensor: +12 VDC, GND, and a signal output from the sensor's internal low noise JFET. The sensor output is passed through a high pass filter with -3dB low frequency roll-off of 20 Hz. It is then passed through a low-noise instrument amplifier with two gain settings (x100 and x1000, Figure 3, Table 1). Since piezoelectric sensors all exhibit a pyroelectric effect, the instrument amplifier is auto-zeroed to eliminate low frequency drift due to temperature changes of the environment on the sensor. The amplified signal is input to a differential output amplifier to drive a twisted pair cable connected to the final differential input, data acquisition system.

Although the instrument described above works well in the terrestrial environment, we must increase the range of physical conditions under which these sensor systems must perform in space, such as on the landing pads of a lunar lander. Herein we address two changes in our system in order to extend the range of working temperatures toward cryogenic conditions, for example, as is expected in the permanently shadowed areas of the Moon. First, we use a new electronic design and incorporate a PZT-based commercial, off-the-shelf sensor (est. US$ 400).

Commercially available PZT-type accelerometers usually contain an internal JFET to convert the charge generated into a voltage output. Most commonly, the other electrode of the sensor is connected to a metal casing and is grounded. This type of circuit produces a single-ended output which is subject to external noise coupling. A better method is to take advantage of the differential nature of the piezoelectric element. Instead of producing a single-ended charge-to-voltage, two low-noise and matched JFETs can produce a floating differential current output. Since no commercially-available ceramic piezo-sensor comes with such a front-end circuit configuration, we choose a DJB Instruments A/23/TS (A23TS) charge-output sensor (Figures 3, 4), with no internal JFET, to prototype our customized differential JFET circuit, which is installed directly at its output. Gate bias resistors are selected to produce a -3dB low-frequency roll-off of 8 Hz. These two differential JFETs are biased with a 100 μA current source and converted the sensor charge through a differential current output that is then sent through a twisted-pair cable to two, matched, transresistance amplifiers that convert the current to a voltage. These amplifiers are also filtered to produce an upper -3 dB frequency cutoff of 5 kHz. The differential voltage outputs of the transresistance amplifiers are then input to a gain-selectable, low-noise, differential instrument amplifier. Similar to the ACH-01 conditioning circuit (Figure 3, Table 1), the instrument amplifier contains an offset-zeroing circuit to compensate for temperature effects and other component offsets. The amplified and filtered output also goes to a fully differential output amplifier that drives a twisted pair line connected to the same data acquisition input as for the ACH-01.

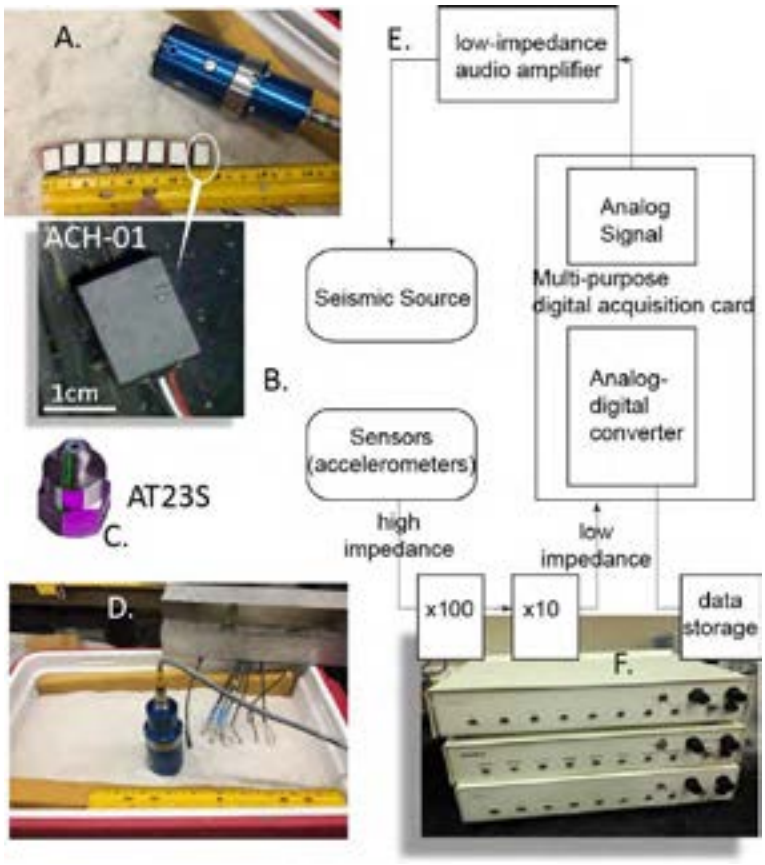


Figure 3. Sensor tests use (A) a magnetostrictive seismic source (blue cylinder) to compare polymer-based piezo-accelerometers (A & B) against a piezo-ceramic sensor (C) in a small sand tank (D) (E) Flow graph charts data acquisition steps[15]. Sensor voltage output is amplified x 100 (bottom right - F) and output as a differential-ended analog stream for digitization and data storage.

Seismic Sensors, Test and Acquisition System (Figure 3)	
	Sensors
For ACH-01	Piezo-electric accelerometer of polyvinylidene fluoride film composition (ACH-01 from Tyco); onboard charge amplifier, nominally flat response of ~9mV/g +/- 1mV, in 20 Hz to 20 kHz frequency range, ~ 8 g.
AT23S	PZT, 4.5 g.
Signal conditioning	100 and 1000-fold operational amplifier with differential output (Figure 3).
Sensor array dimensions	8 sensors, 0.03-.87 m source-receiver offsets, ~0.017 m sensor spacing.
	Digital Recording
Multi-purpose, digital acquisition card	(a) Onboard, PCI-based analog-to-digital acquisition (AD) card with an 8 differential-channel mode input (Model PCI-6251 from Nat. Instr.) software triggering, and low impedance analog output for source wavelet.
Instrument control software	Modified version of Multi-Function-Synch AI-AO.vi written in "G", a commercial virtual instrument software programming language (from National Instruments).
Sample rate	78.125 kS/s, per analog-input differential channel (8)
Nyquist frequency	~39 kHz
Input and output resolution	1 in 16 bits; 305 mV in 16 bits; 305 m*V* range.
Acquisition format	LabView© (Natl. Instr) ASCII format converted to SEGY[16]
Source wavelets	(a) Ricker wavelet, central frequency at 2 kHz, 23 samples at 50 kS/s, 50 micro-s wide side-lobes; synthesized digitally by PCI-6251 AD card. (b) Step-impulse, 20 kHz bandwidth
Seismic source generators	(a) Magnetostrictive ultrasonic transducer (Model CU- 18 from Etrema Products Inc.). Low-impedance audio amplifier (Model RMX 2450 from QSC Audio Products LLC) amplifies input Ricker source wavelet to drive this transducer at +150V (max)
Seismic software	Seismic Unix Processing System[17], for filtering, manipulation and display.
	Oscilloscope-function generator with automatic Bode Plotter (Velleman PCSGU250)
	Mechanical shaker/vibrator (PASCO Scientific Model ST-9324)
	Digital function generator-amplifier (PASCO Scientific Model PI-9587A)

Table 1- Nominal field, source and sensor equipment and software, and seismic acquisition parameters for the laboratory experiments (adapted from [15])

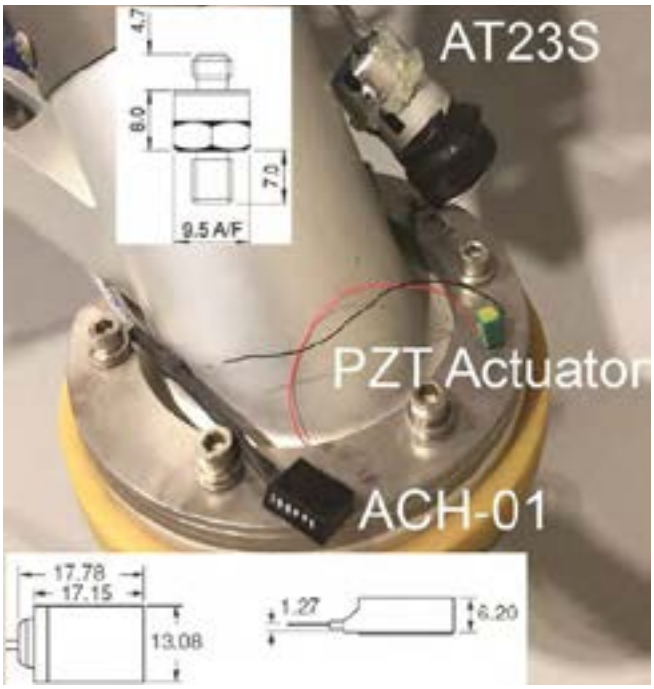


Figure 4. Mighty Eagle Lander pad and leg (NASA Marshall Space Flight Center) and two piezo-electric accelerometers (ACH-01 and AT23S) tested herein. The green piezo-stack actuator will be used in future development of a complete source plus sensor array. Both accelerometers are of comparable weight (Table 1) and size (all measurements are shown in mm). An aluminum cylinder is *only* a temporary, oversized housing to stabilize air moisture content for the front-end electronics.

B. Bode Plots

Generally, piezo-polymer sensors are an order of magnitude more sensitive than piezo-ceramic sensors. However, the major advantage of piezo-ceramics is their proven reliability and performance in space. Polymer-based sensors become brittle at low temperatures and show a marked reduction in their performance.

In order to compare the frequency and phase response characteristics of the ACH-01 against the AT23S, (Table 1) we use proprietary, automated, Velleman software (PCSGU250 V.114). We average all responses at each frequency for 2 s, over 0-10 kHz and normalize the output to the input Voltage (Figure 5). Low-frequency thresholds are different and conditioned by electronics.

Both the ACH-01 and AT23S display a similar phase shift response, in the most useful frequency ranges: < 10^3 Hz. The amplitude gain of the AT23S exceeds that of the ACH-01 and shows a linear increase with frequency (Figure 5). The equivalent response for the ACH-01 stays flat and nearly constant as per manufacturer's specification.

Overall, Bode plots show that the AT23S appears to be more sensitive. However, because the amplifiers for the ACH-01 and AT23S are not currently matched we prefer to conduct an additional signal-to-noise evaluation under common experimental conditions. We do expect the ACH-01 to be more sensitive at room temperatures, but we want to evaluate the relative benefit of our redesigned circuitry for the AT23S. Two, matched, onboard JFET charge amplifiers should improve the signal-to-noise which is a more useful indicator of usability of these tools in space conditions.

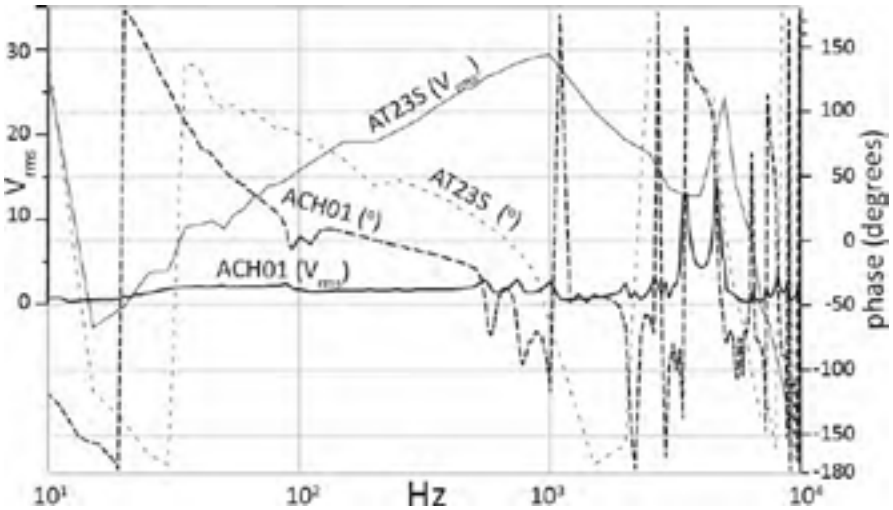


Figure 5. Gain (Vrms) and phase response of a polymer-based piezo-sensor (ACH-01, $30) compared to a piezo-ceramic sensor (AT23S-$400). Although typically less sensitive, the AT23S response is improved with onboard differential charge amplifiers and additional amplifiers prior to data acquisition. (Figure 2, Table 1). We note that pre-amplifiers with increase gain. For reference, off-the shelf cryogenic-rated sensors cost ~US $1000.

C. Signal-to-Noise Evaluation

Figure 6. Spatial arrangement of sensor and source for AT23S sensor. Separation between repeated locations of the magnetostrictive shaker/vibrator = 5 cm. The first offset between the source and the nearest sensor is 3.8 cm. Grain diameter: < 2^{-4} mm

Figure 7. Spatial arrangement of sensor and source for case of ACH-01 sensors. Fixed location of magnetostrictive shaker/vibrator, but ACH-01 sensors are buried 1 cm, and separated 5 cm apart. The first offset between the source and the nearest sensor is 3.8 cm. Grain diameter: < 2^{-4} mm.

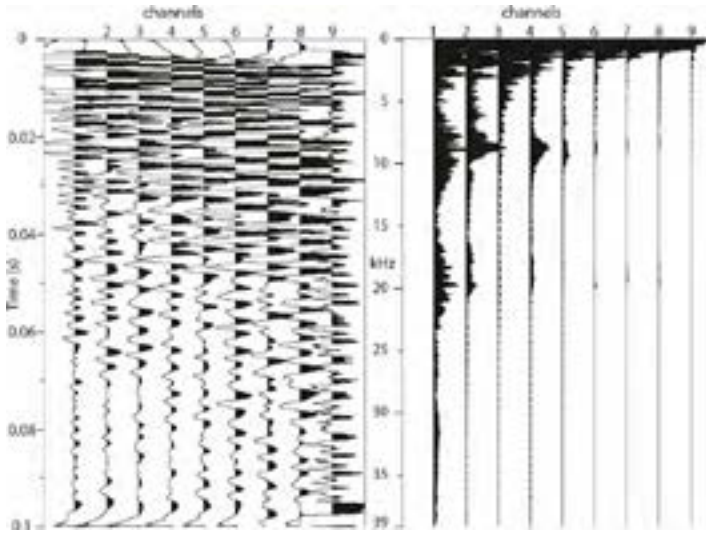


Figure 8. Representative data collected for the AS23S sensor. (LEFT) Channels 1 through 8 indicate variations of recorded voltage output versus time and offset (5 to 40 cm) between the sensor and a moving magnetostrictive shaker/vibrator.-- and channel 9 contains background noise. (RIGHT) Equivalent amplitude spectra for each respective channel, including that containing background noise (channel 9).

We derive a useful comparison of signal-noise-ratio (SNR) between sensors for a common experimental setup in a sand box containing very fine-to-silt sized, angular quartz sand (Figures 6 and 7). We analyze the change in energy among the different frequency components of the data (Figure 8) as a function of fixed distances (offsets) between a mechanical vibrator/shaker and sensors. We do not consider the energy contribution from the phase component. In order to estimate the signal-to-noise (in the frequency domain – Figure 9) we total the ratios of the data energy at each frequency to that found in the common background noise which is collected when the shaker is turned off. The original seismic vibration that enters the sand at the shaker/vibrator is designed to be a Ricker wavelet with a dominant

frequency of 2 kHz. But, by seismic attenuation through the sand, dominant frequency is reduced to < 1 kHz when it is received at the sensors.

Thanks to careful mechanical sieving, we can assume that sand body is sufficiently homogeneous so that only the distance between the sensor and the shaker will affect signal quality. That is, the nature of the seismic signal is dominated only by relative distance between the source and the sensor because the medium properties are laterally constant. As in the case of our single ATS23 sensor, we record both seismic data and background noise while the sensor remains fixed at one end of the sand box, but the source-to-sensor offset varies from 5-to-40 cm (Figure 6). For each recording, the shaker is placed at eight different locations across the sand box 5 cm apart (Figure 6). In the case of the ACH-01, 8 sensors placed in a line and separated by 5 cm (Figure 7) also record seismic data for the full range of sensor-to-shaker offset, as well as the background noise but with the shaker fixed at one end of the sand box. Experiments for both sensor types share a common wavefield and any possible edge effects are similarly shared, so that in practice we are able compare the SNR under identical conditions[2].

IV. Results and Recommendations

Newly designed electronics for a PZT-type sensor (AT23S) show that, although a PVDF-type sensor (ACH-01) displays a better signal-to-noise ratio (SNR) at room temperature, the PZT-type sensor can achieve a similar-order SNR as the ACH-01 (Figures 9 and 10). Specifically, SNR tests of the ACH-01 and its amplifier show a maximum value of ~ 160 centered at ~ 800 – 900 Hz, whereas the SNR of the AT23S is ~ 140 at ~700 Hz. Future experiments will focus on improvement of new electronics (1) to flatten a linear spectral response (Figure 5) that is observed for the AT23S and (2) to temperature-harden the electronic components so that we can continue (3) to test PZT sensor performance as well as piezo-ceramic type shakers/vibrators under liquid nitrogen temperatures.

In future work, causes in the secondary differences of the SNR between sensor types attributable to variable experimental conditions should be addressed as well. Variations can arise from differences in the mechanical coupling between different sensor enclosures (plastic versus Aluminum) to the sand body, enclosure resonance, cabling distribution, edge effects and the presence of various tools in the sand tank during the experiments (cf. Figures 6 and

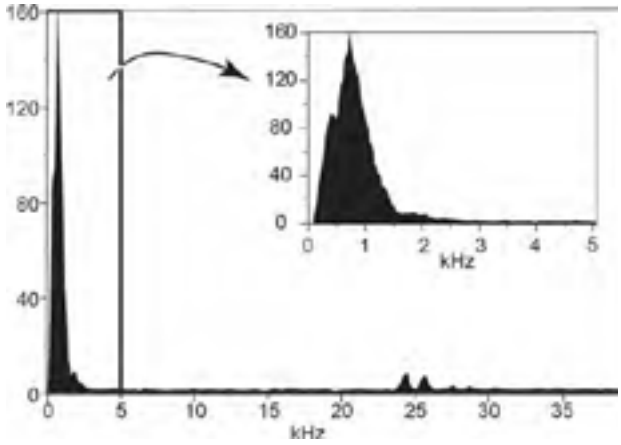


Figure 9. Signal-to-noise-variation ratio for a polymer-based ACH-01 sensor, up to the Nyquist frequency (~39 kHz), in a laboratory experiment. Usable frequencies reach ~5 kHz at most.

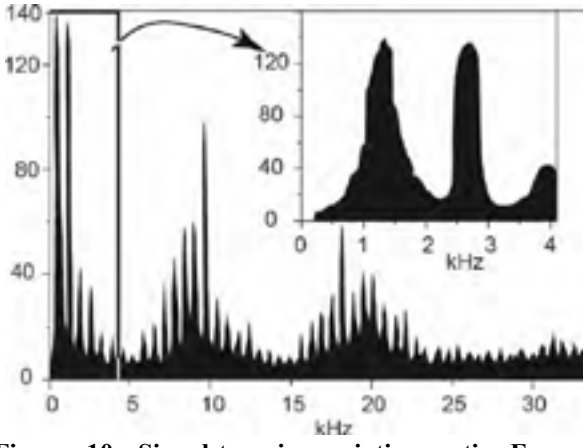


Figure 10. Signal-to-noise-variation ratio For a PZT-based AT23S sensor, up to the Nyquist frequency (~39 kHz) data in a laboratory experiment. Usable frequencies reach ~5 kHz.

7). Although similar experimental variations have not apparently influenced seismic analyses in past experiments[2] future investigations should remain vigilant. We envisage future use of miniature piezo-electric sensors and shakers/vibrators on landers and rovers to help interrogate for shallow (few meters) buried in-situ resources.

V. Acknowledgments

The first author thanks F. Dix of the Marshall Faculty Fellowship program for providing the opportunity to collaborate with members of the Lunar Planetary Science group at the National Space Science and Technology Center during my stay at the Marshall Space Flight Center, Huntsville, Alabama. Renee Weber was pivotal in providing space and a welcoming research environment together with H. Haviland-Fuqua, C. Fassett and M. Zanetti.

VI. References

[1]Jawin, E. R., Valencia, S. N., Watkins, R. N., Crowell, J. M., Neal, C. R., and Schmidt, G., "Lunar Science for Landed Missions Workshop Findings Report: Earth and Space Science:, 2019,, Vol. 6, No. 1, pp. 2-40.

[2]Lorenzo, J. M., Patterson, D. A., Karunatillake, S., Weber, R., Haviland, H., and Fassett, C., "Seismic Characteristics of the Shallow (0–1 m) Soils on the Moon and Mars: Ice in Soils", *50th Lunar and Planetary Science Conference*, 2019, Houston, Lunar and Planetary Institute, Abstract #3246.

[3]Dietrich, R. V., *The Tourmaline Group*, New York, Van Nostrand Reinhold Company, 1985, pp.1-300.

[4]Tressler, J., Alkoy S, and Newham, R. E., "Piezoelectric sensors and sensor materials": *J Electroceram*, 1998, Vol. 2, pp. 257-272.

[5]Zu, H., Wu, H., and Wang, Q.-M., "High-Temperature Piezoelectric Crystals for Acoustic Wave Sensor Applications*": IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 2016, Vol. 63, No. 3, pp. 486-505.

[6]Santamarina, J., Wakim, T., Tallin, A., Rab, F., and Wong, J., "Piezo film technology and applications in geotechnical testing",*Geotechnical Testing Journal*, 1991, Vol. 14, No. 4, pp. 363-370.

[7]Lorenzo, J.M., Hicks, J., Vera, E.E., "Integrated seismic and Cone Penetration Testing observations at a distressed earthen levee: Marrero, Louisiana, U.S.A.", Engineering Geology. *Journal of Engineering Geology* 2014, Vol.168C, pp. 59-68.

[8]Martinez, C., Zheng, Y., Daniel Easton, D., Kevin Farinholt, K., and Gyuhae Park, G., "Strain Sensors for High Field pulse Magnets", *27th Conference and Exposition on Structural Dynamic* 2009: (IMAC XXVII); Orlando, Florida, USA,200 9-12 February 2009, Vol. 1

[9]Hooker, M. W., "Properties of PZT-Based Piezoelectric Ceramics Between -150 and 250 C", 1998 NASA Langley Research Center; Hampton, VA United States.

[10]Blackwood, G. H., and Ealey, M. A., "Electrostrictive behavior in lead magnesium niobate (PMN) actuators. I. Materials perspective", *Smart Materials and Structures*, 1993, Vol. 2, No. 2, pp. 124-133.

[11]Xu, T.-B., Tolliver, L., Jiang, X., and Su, J., "A single crystal lead magnesium niobate-lead titanate multilayer-stacked cryogenic flextensional actuator", *Applied Physics Letters*, 2013, Vol. 102, No. 4, p. 042906.

[12]Lorenz, R .D. (1994) *Measurement Science Technology* 5, 1003-1041.

[13]Bowles, N. E., Pike, W. T., Teanby, N., Robert, G., S. B. Calcutt, J. Hurley, P. Coe1, J., Wookey, Dunton, P., Standley, I., Temple, J., Irsha, R., Taylor, J , Warre, T., and Charalambous, C., "Performance and Noise modelling of the Short Period Seismometer SEIS-SP, part of the SEIS instrument for NASA's 2016 InSight Mission", *LPSC* 2015, Houston.

[14]Cooper, M. R., Kovach, R. L., and Watkins, J. S., "Lunar near-surface structure", *Reviews of Geophysics*, 1974, Vol. 12, No. 3, pp. 291-308.

[15]Lorenzo, J. M., Smolkin, D. E., White, C., Chollett, S. R., Sun, T. "Benchmark hydrogeophysical data from a physical seismic model", *Computers & Geosciences* 2013 Vol. 50, pp. 44-5. Data sets downloadable from (http://github. com/cageo/Lorenzo-2012)

[16]Barry, K. M., Cavers, D. A., and Kneale, C. W., "Report on recommended standards for digital tape formats", *Geophysics*, 1975 Vol. 40, No. 2, pp. 344-352.

[17]Stockwell, J.W., "The CWP/SU: Seismic Unix package", *Computers & Geosciences*, 1999, Vol. 25, pp. 415-419.

Improving optical gyroscopes by using coupled cavities

Eugeniy E. Mikhailov[1,*] and David D. Smith[2]

[1] *Department of Physics, William & Mary, Williamsburg, Virginia 23187, USA*
[2] *NASA Marshall Space Flight Center, Huntsville, Alabama, USA*

(Dated: February 17, 2020)

We study coupled cavities laser system as a possible candidate for improving optical gyroscopes. Our theoretical study shows that this system can be set to the enhanced regime by proper choice of parameters. Our experiment shows that the system behaves in accordance to our model. However, due to the limited equipment available at the time of our experiment, we have not yet achieved high enough enhancement to be of interest for a practical device.

I. INTRODUCTION

"NASA's future missions show a diverse set of navigational challenges that cannot be supported with current methods. Onboard autonomous navigation and maneuvering techniques are critical"[1]. Optical gyroscopes, which are able to autonomously track the direction of a space craft, are addressing this navigation challenge. However, the sensitivity of a classical laser gyroscope scales proportionally to its area footprint and already reaches close to fundamental limit. A coupled cavities laser gyroscope is envisioned to give higher precision with the same footprint.

Optical gyroscopes utilize the Sagnac effect, where the frequency of the lasing changes due to relativistic effects in a rotating medium. This frequency shift can be thought to be due to a change in the optical path length of the laser, proportional to the rotation velocity and the area of the gyroscope. This connection allows testing of an optical gyroscope prototype without actually spinning the setup. Instead, it is sufficient to tune the cavity length.

The enhancement in the gyroscope response (pulling factor) is a ratio of the lasing frequency change in the system under the test (the coupled cavities laser, see Fig. 1) to an analogous setup without coupling under the same excitation.

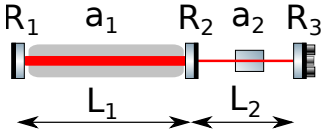


FIG. 1. Coupled cavity setup

The coupled-cavity laser is depicted in Fig. 1. It consists of the lasing medium (a_1) and two mirrors (R_1 and R_2), just as in a conventional laser. However, an additional mirror (R_3), which bounces/couples light back and creates a second cavity from mirrors (R_2 and R_3), allows tuning of the laser response. In this system the important measure is the dependence of the detuning of the lasing frequency Δ on the relative cavities detuning change δ.

II. THEORY

1. *Master equations for the coupled cavity laser*

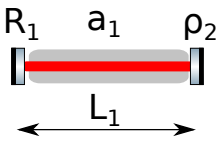


FIG. 2. Simplified coupled cavities setup

The reflectivity of a compound coupled cavities setup (ρ_{123}), shown in Fig. 1, can be expressed as a reflectivity of two mirror Fabri-Perot cavity shown in Fig. 2. Where we replace the R_2 mirror with a mirror which has a phase dependent reflectively ρ_{23} of a Fabri-Perot cavity consisting of mirrors R_2 and R_3. If we neglect losses inside of the mirrors, we have the following equation for the reflectivity (from the mirror R_1 side) at the probe light frequency (ω) of the coupled cavities

$$\rho_{123}(\omega) = -r_1 + \frac{a_1\rho_{23}(1-r_1^2)e^{i\phi_1}}{1-a_1\rho_{23}r_1e^{i\phi_1}} \tag{1}$$

where

$$\rho_{23}(\omega) = -r_2 + \frac{(a_2r_3)(1-r_2^2)e^{i\phi_2}}{1-(a_2r_3)r_2e^{i\phi_2}} \tag{2}$$

Here the r_i stands for the field reflectivity coefficient of the i_{th} mirror, a_i is amplification inside of the given cavity, and phase factors ϕ_1 and ϕ_2 are due to round trip travel time $t_i = 2L_i/c$ in either cavity 1 or 2. These factors are

$$\phi_1 = (\omega-\omega_1)t_2 = \Delta t_1 \tag{3}$$

$$\phi_2 = (\omega-\omega_2)t_2 = (\Delta-\delta)t_2 \tag{4}$$

* eemikh@wm.edu

where ω_1 and ω_2 are the resonances of the individual cavities consisting of mirrors R_1 and R_2 separated by distance L_1 or mirrors R_2 and R_3 with mirrors' separation L_2, respectively. Δ is the probe frequency detuning from the first cavity, and δ is the detuning of cavity 2 from cavity 1.

Lasing occurs when the reflectivity of the system (ρ_{123}) goes to infinity. This happens when denominator of the Eq. 1 goes to zero

$$1 - a_1 \rho_{23} r_1 e^{i\phi_1} = 0 \tag{5}$$

Substituting ρ_{23} [2] from Eq.2, we obtain

$$r_1 a_1 e^{i\phi_1} = r_2 + \frac{1 - r_2^2}{r_2 - r_3 e^{i\phi_2}} = G \tag{6}$$

here we made a substitution $r_3 a_2 \to r_3$, i.e. we absorb the second cavity amplification/losses into its mirror R_3 reflectivity. The frequency (ω) and the required saturated gain (a_1) values that satisfy Eq. 6 depend only on δ for a given r_1, r_2, and r_3.

Separating real and imaginary parts of Eq. 6, we obtain the expression for the required gain to achieve lasing

$$\begin{aligned}(a_1 r_1)^2 =& \Re(G)^2 + \Im(G)^2 \\ =& \left(\frac{\left(1 - r_2^2\right)\left(r_2 - \cos\varphi_2\, r_3\right)}{\left(r_2 - \cos\varphi_2\, r_3\right)^2 + \sin^2\varphi_2\, r_3^2} + r_2 \right)^2 \\ &+ \frac{\sin^2\varphi_2 \left(1 - r_2^2\right)^2 r_3^2}{\left(\left(r_2 - \cos\varphi_2\, r_3\right)^2 + \sin^2\varphi_2\, r_3^2\right)^2}\end{aligned} \tag{7}$$

Similarly, we can express

$$\phi_1 = \arctan(\Im(G)/\Re(G)) \tag{8}$$

Substituting the parameters we obtain the following expression for ϕ_1 as the function of ϕ_2

$$\phi_1(\phi_2) = \arctan\left(\frac{\sin\varphi_2 \left(1 - r_2^2\right) r_3}{r_2\left(\left(r_2 - \cos\varphi_2\, r_3\right)^2 + \sin^2\varphi_2\, r_3^2\right) + \left(1 - r_2^2\right)\left(r_2 - \cos\varphi_2\, r_3\right)} \right) = \Delta t_1 \tag{9}$$

2. *Derivation of the pulling factor*

Taking partial derivative of the left and right hand sides of Eq. 9, we can obtain a general expression for the pulling factor (PF)

$$PF = \frac{d\Delta}{d\delta} = \frac{1}{1 - \alpha/\beta(\phi_2)} \tag{10}$$

where $\alpha = t_1/t_2 = L_1/L_2$ is a ratio of cavities round trip times or lengths; and $\beta(\phi_2) = \frac{d\phi_1}{d\phi_2}$.

There are special points where $\beta(\phi_2) = \alpha$ and $PF \to \infty$. At these points our compound system has the highest response to the tiny variation of relative cavities detuning δ. To achieve maximum gyroscope response, one should keep the system at these conditions.

Unfortunately, the general expression for β is very cumbersome, but we can state that $\beta(\phi_2) < \beta_0$, where

$$\beta_0 = \beta(\phi_2 = 0) = \frac{r_3(1 - r_2^2)}{(1 - r_2 r_3)(r_2 - r_3)} \tag{11}$$

3. *Finding lasing frequency*

The parameters ϕ_1 and ϕ_2 cannot be measured directly in experiments, so we recast the solution in terms of the laser frequency shift with respect to the first cavity (Δ) and the relative cavity detuning (δ). Using Eqs. 3 and 4, we see that $\phi_1 = (\delta t_2 + \phi_2)\alpha$. Consequently, we can rewrite Eq. 9 only in term of ϕ_2 as

$$\phi_1(\phi_2) = (\delta t_2 + \phi_2)\alpha \tag{12}$$

Equation 12 is graphically depicted in Fig.3 (upper row). There are up to 3 possible roots for ϕ_2, which we find with numerical methods. Once we find them, we can reconstruct the dependence of the lasing frequency detuning (Fig.3 2nd from the top row), pulling factor (Fig.3 3rd from the top row), and required gain $a_1 r_1$ (Fig.3 the bottom row). Examination of the Δ vs δ dependency shows that it is possible to have multiple lasing frequencies for the same δ. When two branches merge, the solutions are degenerate, and this is also where PF diverges. At this point the $(\delta t_2 + \phi_2)\alpha$ line is tangential to the $\phi_1(\phi_2)$ curve. From the experimental point of view, this condition is very hard to achieve, since the range of acceptable δ goes to zero around these special points.

Analysis of Eq. 12 and Fig. 3, shows that a much more promising case with $|PF| > 1$ exists (see the left column of Fig. 3). Here, we can see rather large $|PF|$ around zero detunings (both δ and Δ) [3]. So $PF < -1$ is obtained when $\alpha/2 < \beta_0 \leq \alpha$. One can also see that the requirement for the gain grows as $\delta \to 0$, so not every lasing medium can reach such high PF values. We can obtain simplified expressions for the required gain to sustain lasing at $\delta \approx 0$.

$$(a_1(\delta = 0) r_1)^2 = \left(\frac{1 - r_2 r_3}{r_2 - r_3} \right)^2 \tag{13}$$

III. EXPERIMENT

We built the coupled-cavities laser utilizing a He-Ne tube of 25 cm length, capped with a high reflector mirror (R1) with curvature of 60 cm, and several different mirrors R2 and R3. Inside the R1-R2 cavity we placed a small pick off mirror (an anti-reflection coated waveplate with an estimated reflection of less than 1%) to send a small portion of intracavity lasing light to an optical spectrum analyzer with a free spectrum range 1.5 GHz. The other portion of the pick off beam was sent to a fast photodiode where a beat frequency between longitudinal laser modes was observed. Depending on the choice of L_1, the beat note was in the range of 200 — 500 MHz. This beat note allowed us to track the laser frequency shift, under the assumption that R2-R3 cavity was in close resonance with only one longitudinal mode. We mounted the $R3$ mirror on a piezo-electric transducer to tune the relative frequency shift δ. The optical spectrum analyzer (OSA) in conjunction with the beat note observation allowed us to track the lasing frequency shift Δ.

A. Observation of chaotic regime

Generally, lasers go into a chaotic regime when a part of the output beam is directed back into the laser with sufficient intensity. Our laser was no exception because the R3 mirror sent the output back. An example of such behavior is shown in Fig. 4, one can see that at certain δ there is no correlation of the laser output with its previous value.

To mitigate the chaotic behavior, we reduce a_2 by inserting neutral density filters into the second cavity. The resulting laser behavior is shown in Fig. 5. As the amplification factor a_2 decreased, the laser output became more and more stable. Unfortunately, at very low amplification factors the cavities became decoupled and the response to the second cavity detuning became very small.

B. Demonstration of the pulling factor

To resolve the small laser frequency shifts at small PF values, we switched to monitoring the beat note with a radio-frequency spectrum analyzer (Tektronix MCO4000). The motion of the beat note with respect to the beat frequency corresponding to the laser free spectral range (500.3 MHz) is shown in Fig. 6.

Notably, the beat note disappeared around high laser frequency change detunings (see the right most Fig. 6). This corresponded with a disappearance of one of the longitudinal modes on the OSA. This is predicted by our theory because the required gain is the highest at *delta* near zero (see Fig. 3). Our lasing medium is not able to provide the required gain , and, consequently, the lasing ceases.

The estimate of the highest pulling factor (the slope in the right most Fig. 6) is about 4×10^{-4}. While this number is very modest, it is in qualitative agreement with our theoretical model, which predicts that in the regime of low amplification factor in the second cavity the PF will be small.

IV. SUMMARY

Our theoretical model demonstrates availability of the high pulling regime, suitable for enhancing response of optical gyroscopes and thus improving space navigation. Our theory is in agreement with observed data. We have demonstrated response of the coupled cavity He-Ne laser to the small motion of the mirror external to the main cavity.

V. OUTLOOK

There were several experimental difficulties which impeded our progress: we did not have a way to fine tune the amplification factor, the He-Ne tube had limited gain which blocked us from the regime where the PF is the highest, our lasing tube produced laser frequency jitter which could limit our means for detection at higher PF values. None of these limitations are fundamental, and with better equipment we should be able to demonstrate much higher performance.

[1] NASA Technology Roadmap, TA-05.

[2] A careful reader would notice that we use $-\rho_{23}$ expression for this substitution. This is permitted, since equivalent result can be obtained by introducing the π phase shift to ϕ_1.

[3] The negative sign of PF is irrelevant for our discussion.

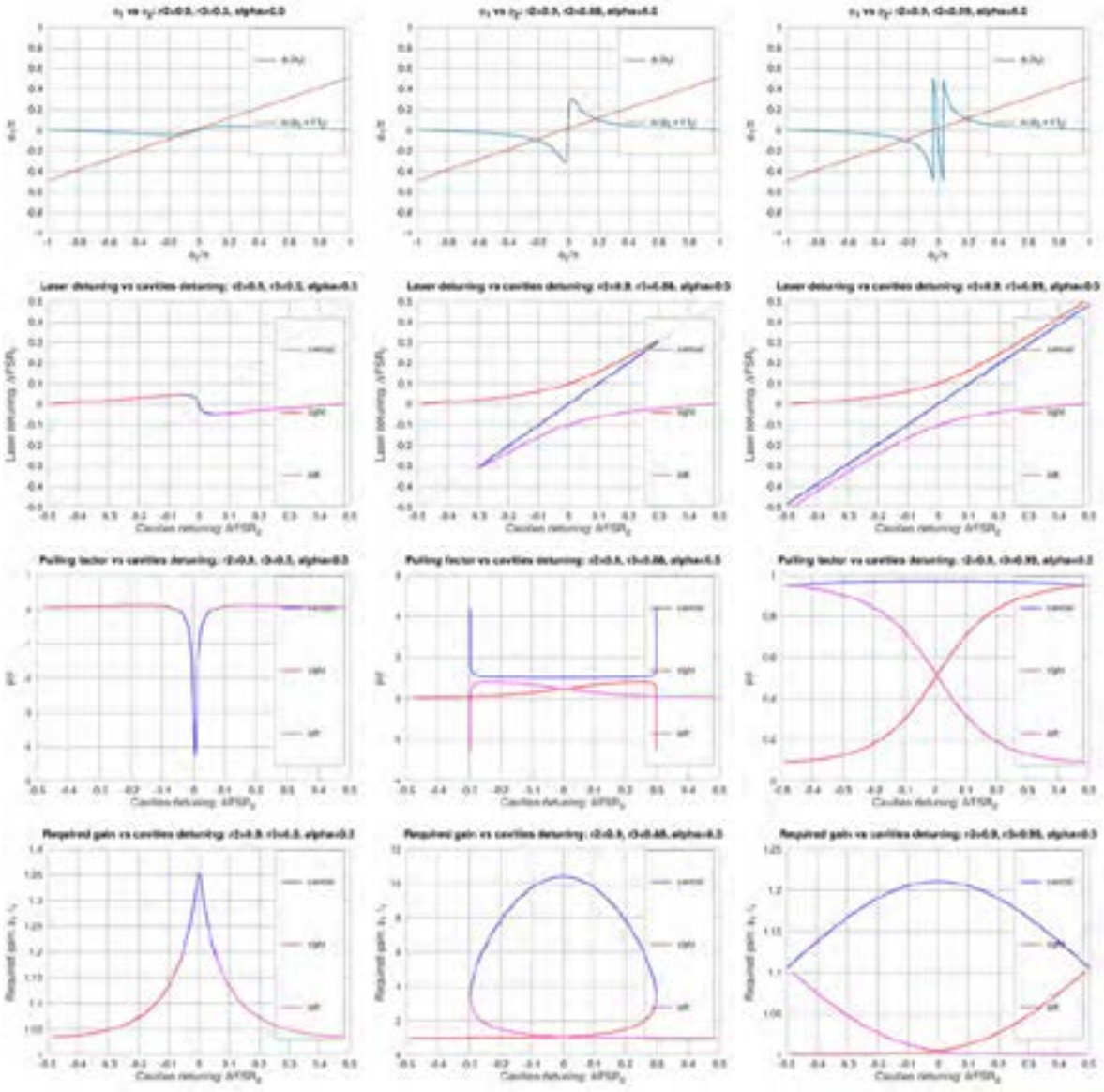


FIG. 3. Behavior of the coupled cavities laser for different parameters. Top raw is representation of Eq. 12. Circles depict the cases when Eq. 12 is satisfied. The second from the top raw is the laser frequency detuning (Δ) vs the relative cavities detuning (δ). The third from the top raw is pulling factor vs δ. The bottom raw is the required gain vs δ. For all plots $r_2 = 0.9$ and r_3 varies from left to right as 0.5, 0.88, and 0.99. Labels 'left', 'central, and 'right' correspond to the branches corresponding to different roots (circle positions) of the Eq. 12 depicted in the upper row. FSR_2 is the free spectral range of the second cavity $\pi c/L_2$.

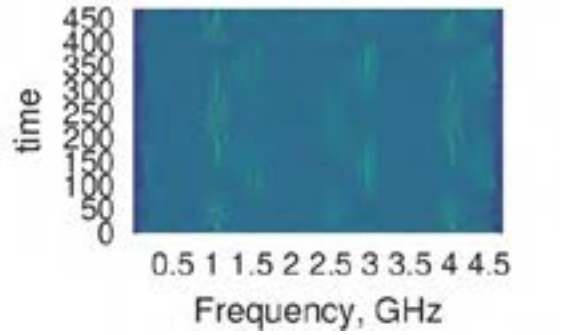


FIG. 4. Example of chaotic behavior. The frequency of lasing mode measured by (OSA) as function of time (δ sweep). $L_1 =$ 30 cm, $L_2 = 8$ cm. R2 is planar output coupler mirror, and R3 is output coupler with cuvature of 30 cm.

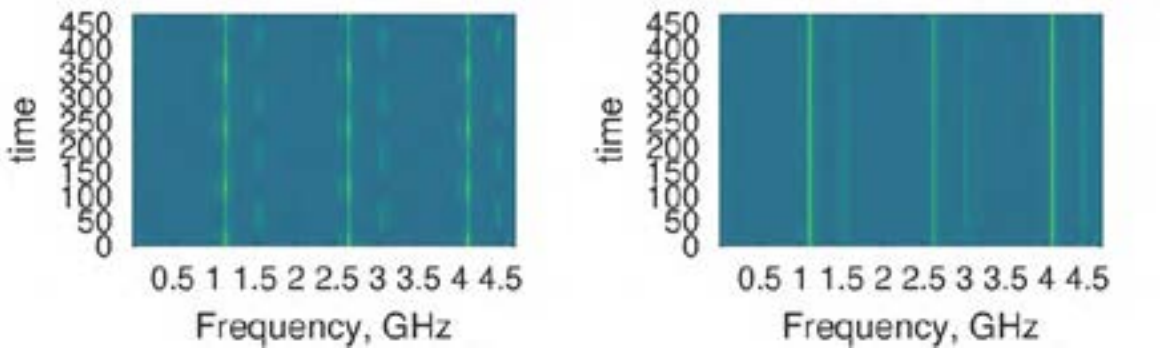


FIG. 5. Example of mitigated chaotic behavior. The frequency of lasing mode measured by (OSA) as function of time (δ sweep). Cavities' parameters are the same as in Fig. 4. Left $a_1 = 0.084$; Right $a_1 = 0$ set by opaque block inside of the second cavity.

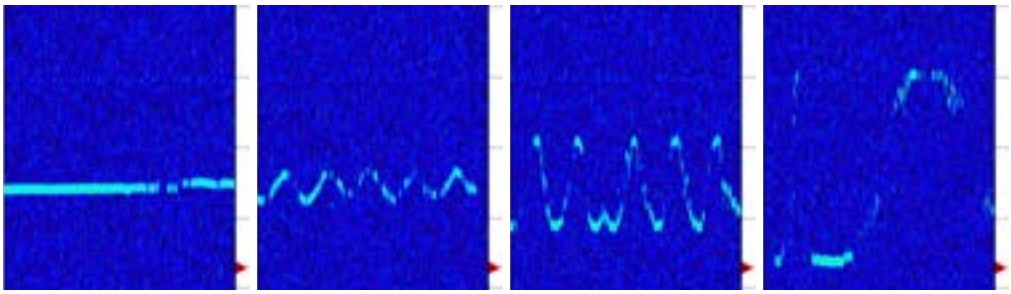


FIG. 6. The laser frequency shift by the change relative cavities detuning δ. The horizontal axis corresponds to sweep of δ by more than one FSR. The vertical axis span is 800 kHz, tick marks spacing is 200 kHz. Cavities' parameters are the same as in Fig. 4. Left to right change of the survival factor a_1: 0.0 (opaque block), 0.023, 0.086, 0.19.

A Preconditioned Quasi-Minimal Residual Algorithm for Solving Large Scale Network Flow Problems with GFSSP

S.S. Ravindran,
Professor, Propulsion Research Center, The University of Alabama in Huntsville, Huntsville, AL 35899.
and
Alok K. Majumdar,
Combustion and Thermal Analysis Branch, ER43, NASA MSFC, Huntsville, AL 35812.

We propose an Incomplete LU (ILU) factorization preconditioning technique and show the effectiveness of the preconditioned iterative linear solvers to network flow problems. Two Krylov subspace methods, namely, preconditioned quasi-minimal residual (QMR) method and preconditioned stabilized biconjugate gradient method (Bi-CGSTAB) have been investigated. Proposed solvers have been integrated into a network flow simulation software (Generalized Fluid System Simulation Program) based on the unstructured finite volume method. In order to access the performance and convergence of the algorithms, they are used to predict pressure surges in a pipeline that has entrapped air at one end of the pipe. Numerical predictions are compared with available experimental data and are found to be in good agreement. The results show that the ILU preconditioned iterative solvers significantly reduce computational time. Moreover, the results show that QMR method tends to provide better performance when coupled with ILU preconditioner.

I. Introduction

Preconditioned Krylov subspace methods are widely used for iteratively solving large-scale linear systems, especially those arising from computational fluid dynamics. However, their usage in network fluid flow computation is limited. Generalized Fluid System Simulation Program (GFSSP) [6] is a network simulation software based on unstructured upwind finite volume spatial discretization of conservation of mass, energy and species, and the momentum equation. The time discretization employs a backward-difference formula (BDF) based implicit time stepping which in general leads to a linear system of equations including a large sparse, badly conditioned non-symmetric matrix due to complex nature of physics and boundary conditions. Therefore, the performance of the implicit time stepping scheme decisively depends on the properties of the incorporated solver for the linear system of equations. Direct methods such as Gaussian elimination with partial pivoting for solving linear system can become prohibitive in terms of computer time and storage for large scale network flow problems.

For symmetric positive definite systems, conjugate gradient (CG) [4] and minimal residual (MINRES) are well accepted Krylov subspace methods. However, for nonsymmetric matrices there is no clear acceptable method. Generalized minimal residual (GMRES), quasi-minimal residual (QMR) and stabilized bi-conjugate gradient (Bi-CGSTAB) are some of the most popular Krylov subspace methods proposed in the literature. Among those, the last two enjoy three term recurrence. GMRES uses orthogonal basis for the construction of Krylov subspace but typically requires restarts undermining convergence. Bi-Conjugate gradient (Bi-CG) [2] method is the unsymmetric variant of the CG method based on the three term recurrence. In Bi-CG the approximations are constructed in such a way that residual $\mathbf{r}_j$ is orthogonal to another set of so-called "shadow residuals". Bi-CG, however, suffers from breakdown and irregular convergence behavior. Conjugate gradient squared (CGS) [11] was derived from Bi-CG with the goal to improve the convergence rate and to avoid adjoint vector multiplication. However, CGS tends to cause oscillations in the convergence behaviour. This lead to the development of a further variant of the BiCG algorithm called Bi-CG stabilized (Bi-CGSTAB) [12] which improves the convergence behavior and the expected convergence rate is about twice as that of Bi-CG algorithm. However, Bi-CGSTAB still suffers from breakdowns just like Bi-CG. In order to avoid probable breakdowns situations, a Bi-CG like approach called QMR was proposed in [3]. QMR combines desireable properties of the GMRES like minimization property and short recurrence like Bi-CG. Thus the scheme ensures smooth convergence behavior and requires low storage per iteration. A potential issue with all these methods is that they may suffer from breakdowns or near breakdowns. However, these apparent disadvantages of Krylov subspace methods can be overcome by effective preconditioners.

The goal of this paper is to investigate an ILU preconditioned QMR method for solving large sparse, badly conditioned nonsymmetric linear systems arising from network flow simulations. In order to carry out a systematic comparative study, we use a sudden valve opening of a pipe system with entrapped air as the benchmark problem.

II. Network Flow Modeling

A network fluid flow simulation software (GFSSP) has been used to model the benchmark problem. GFSSP is an unstructured finite volume-based network flow analysis program for analyzing thermofluid systems. A fluid network consists of boundary nodes, internal nodes, and branches to represent a fluid system. Boundary and internal nodes are connected through branches in series or parallel arrangements. At boundary nodes, pressures and temperatures are specified. Mass and energy conservation equations are solved in internal nodes. Flowrates are calculated in branches. Athermal systemconsists of solid andambient nodes connectedwith conductors. A fluid and solid node are connected with a solid to fluid conductor to model conjugate heat transfer. GFSSP uses a pressure-based scheme as pressure is computed from the mass conservation equation. The mass and momentum conservation equations and thermodynamic equation of state are solved simultaneously by the Newton–Raphson method while energy conservation equations of fluid and solid are solved separately but implicitly coupled with the other equations stated above. The conservation equations are solved in conjunction with the thermodynamic equation of state. From the computed pressure and enthalpy at the nodes, all other thermodynamic properties including density, viscosity, and thermal conductivity are evaluated from builtin thermodynamic property programs. For the saturated condition, vapor quality is calculated from liquid and vapor enthalpies at the node pressure. Density and other thermophysical properties of the liquid–vapor mixture are calculated as a function of vapor quality. The study of dynamics is carried out by using implicit time stepping schemes that allow users to use large time steps. While the implicit scheme has robust stability properties, it also requires solving a nonlinear system of equations by iterative Newton's method in which each iteration involves a nonsymmetric, badly conditioned sparse matrix solve.

III. Krylov Subspace Iterative Methods

Given a matrix $A \in \mathbb{R}^{n\times n}$ and vector $\mathbf{v} \in \mathbb{R}^n$, the k^{th} Krylov subspace generated by them, denoted by

$$\mathcal{K}_k(A, \mathbf{v}) := \text{span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{k-1}\mathbf{v}\}\,.$$

In order to solve $A\mathbf{x} = \mathbf{b}$, let $\mathbf{x}_0$ be some initial guess to the solution, and $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$ be the initial residual vector. A Krylov subspace method incrementally finds approximate solution $\mathbf{x}_k$ from the Krylov subspace $\mathcal{K}_k(A, \mathbf{v})$. Two commonly used methods for the construction of basis of the subspace $\mathcal{K}_k(A, \mathbf{v})$ are Arnoldi iteration and bi-Lanczos iteration. The Arnoldi iteration is a procedure for constructing the orthogonal basis for $\mathcal{K}_k(A, \mathbf{v})$. Starting from the a unit vector $\mathbf{q}_1 = \mathbf{v}/\|\mathbf{v}\|_2$, it iteratively constructs $Q_{k+1} = [\mathbf{q}_1|\mathbf{q}_2|\ldots|\mathbf{q}_k|\mathbf{q}_{k+1}]$ with orthonormal columns by solving

$$h_{k+1,k}\mathbf{q}_{k+1} = A\mathbf{q}_k - h_{1k}\mathbf{q}_1 - \ldots - h_{kk}\mathbf{q}_k$$

where $h_{ij} = \mathbf{q}_i^T A\mathbf{q}_j$ for $j \le i$ and $h_{k+1,k} = \|A\mathbf{q}_k - h_{1k}\mathbf{q}_1 - \ldots - h_{kk}\mathbf{q}_k\|$. The Arnoldi iteration has a k-term recurrence and thus its computational cost increases as k increases. For this reason, it uses restarts but the restart may undermine its convergence. The GMRES is a Krylov subspace method that computes at the k^{th} step the best least squares solution $\mathbf{x}_k$ from the Krylov subspace $\mathcal{K}_k(A, \mathbf{b})$. The Arnoldi iteration is used to find this vector. The bi-Lanczos iteration also known as Lanczos bi-orthogonalization is an alternative way to construct the basis of the Krylov subspace $\mathcal{K}_k(A, \mathbf{v})$. Starting with the vector $\mathbf{v}_1 = \mathbf{v}/\|\mathbf{v}\|_2$, we iteratively construct $\mathbf{v}_{k+1} = [\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_k|\mathbf{v}_{k+1}]$ by solving $\beta_k\mathbf{v}_{k+1} = A\mathbf{v}_k - \gamma_{k-1}\mathbf{v}_{k-1} - \alpha_k\mathbf{v}_k$. If $\mathcal{K}_k \neq \mathcal{K}_{k-1}$, then the columns of V_k form a basis of $\mathcal{K}_k(A, \mathbf{v})$ and

$$A\mathbf{v}_k = \mathbf{v}_{k+1}\widetilde{T}_k\,, \tag{1}$$

where

$$\widetilde{T}_k = \begin{bmatrix} \alpha_1 & \gamma_1 & & & & & \\ \beta_1 & \alpha_2 & \gamma_2 & & & & \\ & \beta_2 & \alpha_3 & . & & & \\ & & . & . & . & & \\ & & & . & . & . & \\ & & & & . & . & \gamma_{k-1} \\ & & & & & \beta_{k-1} & \alpha_k \\ & & & & & & \beta_k \end{bmatrix}$$

is a $(k+1) \times k$ tri-diagonal matrix. To determine α_i and γ_i, we construct another Krylov subspace $\mathcal{K}(A^T, \mathbf{w})$ whose basis is given by the columns of $\mathbf{w}_{k+1} = [\mathbf{w}_1 | \mathbf{w}_2 | \ldots | \mathbf{w}_{k+1}]$ subject to the bi-orthogonality condition

$$\mathbf{w}_{k+1}^T \mathbf{v}_{k+1} = \mathbf{v}_{k+1}^T \mathbf{w}_{k+1} = I_{k+1} \,. \tag{2}$$

It follows from (1) and (2) that $\alpha_k = \mathbf{w}_k^T A \mathbf{v}_k$. Let $\mathbf{v} = \mathbf{v}_k$ and $\mathbf{w} = \mathbf{w}_k = \mathbf{v}^{-T}$ be complete basis vectors of $\mathcal{K}_k(A, \mathbf{v})$ and $\mathcal{K}_k(A^T, \mathbf{w})$, respectively. Then, we have $\mathbf{w}^{-1} A^T \mathbf{w} = S$, where $S = \mathbf{v}^T A^T \mathbf{v}^{-T}$ and $A^T \mathbf{w}_k = \mathbf{w}_{k+1} \widetilde{S}_k$, where $\widetilde{S}_k$ is the leading $(k+1) \times k$ submatrix of S. Therefore

$$\gamma_k \mathbf{w}_{k+1} = A^T \mathbf{w}_k - \beta_{k-1} \mathbf{w}_{k-1} - \alpha_k \mathbf{w}_k \,.$$

The Krylov subspace methods, Bi-conjugate gradient (Bi-CG) and quasiminimal residual methods are based on bi-Lanczos iterations. In Bi-CG, the residual is constructed to be orthogonal to a Krylov subspace generated with A^T and some vector $\mathbf{s}_0$. That is, the two sets of residuals $\{\mathbf{r}_i\}$ and $\{\widetilde{\mathbf{r}}_i\}$ (the set obtained by using A^T called "shadow residuals") are produced which are bi-orthogonal. Similarly, the two sets of direction vectors $\{\mathbf{p}_i\}$ and $\{\widetilde{\mathbf{r}}_i\}$ are produced from the residuals, which are mutually A-conjugate. However, Bi-CG suffers from breakdown situation ($\mathbf{v}_k^T \mathbf{w}_k = 0$) or near breakdown ($\mathbf{v}_k^T \mathbf{w}_k \approx 0$) and exhibits irregular convergence behaviour. This motivated Freund [3] to propose the QMR algorithm in which the projected overdetermined tri-diagonal system is solved in a least squares sense. Since the basis vectors for the Krylov subspace, generated by the two-sided Lanczos process, are bi-orthogonal and in general not orthogonal, this approach does not lead to a minimum residual approximate solution and thus the name quasi-minimum residual (QMR).

In [11] Sonneveld observed that the computational effort to produce the shadow residuals $\{\widetilde{\mathbf{r}}_i\}$ in Bi-CG could as well be used obtain an additional reduction of the Bi-CG residuals $\{\mathbf{r}_i\}$. This lead him to propose an algorithm known as conjugate gradient squared (CGS) that computes approximations $\mathbf{x}_k$ with a residual of the form $\mathbf{r}_k = q_k(A)\mathbf{r}_k$, where $\mathbf{q}_k$ is some appropriate polynomial of degree k. In case $q_k(A)$ gives an additional reduction, CGS is an attractive method. Unfortunately, in many situations, the CGS choice for q_k leads to amplifications of $\mathbf{r}_k$ instead of reduction. Van der Vorst [12] proposes to take for $\mathbf{q}_k$ a product of appropriate 1-step minimal residual polynomials, i.e., degree one polynomials of the form $(1 - \omega_k t)$ for some optimal ω_k. This choice leads to what is known as the Bi-CGSTAB algorithm.

The Krylov subspace methods demonstrate excellent convergence for many types of problems. However, they may stagnate or do not converge in some cases. Usually these situations are alleviated by selecting a good pre-conditioner. Roughly speaking, a preconditioner is a matrix M, whose inverse approximates A^{-1} and $M^{-1}\mathbf{v}$ can be computed efficiently. Qualitatively, M is a good preconditioner if $M^{-1}A$ is well-conditioned and has tightly clustered eigenvalues than those of A. With a preconditioner, instead of solving $A\mathbf{x} = \mathbf{b}$, one solves $M^{-1}A\mathbf{x} = M^{-1}\mathbf{b}$ by using the Krylov subspace $\mathcal{K}(M^{-1}A, M^{-1}\mathbf{b})$ instead of $\mathcal{K}(A, \mathbf{b})$. In order to precisely define preconditioner, $A = D + L + U$, where D is the diagonal of A, L is the strict lower triangular part, and U is the strict upper triangular part. Some of the simplest preconditioners are the Gauss-Seidel with $M = (D + L)$, Jacobi with $M = D$ and SOR with $M = \frac{1}{\omega}(D + \omega L)$, where $\omega > 0$. However, these preconditioners fail to be invertible if the diagonal entries of A are zero. We propose to investigate a preconditioner based on incomplete LU factorization as follows: Set $M_1 = L + I$, $M_2 = U$ and define $M = M_1 M_1$. The preconditioned QMR algorithm follows the two term recurrence version without look-ahead. The pseudo-code for the QMR Algorithm can be summarized as follows:

Algorithm 1: Preconditioned QMR
Input: $A, M = M_1 M_2 \in \mathbb{R}^{n \times n}$ invertible, $\mathbf{b}, \mathbf{x}_0 \in \mathbb{R}^n$.
output: Approximate solution of $A\mathbf{x} = \mathbf{b}$.

(i) $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$; $\widetilde{\mathbf{v}}_1 = \mathbf{r}_0$; $M_1\mathbf{y}_1 = \widetilde{\mathbf{v}}_1$, $\rho_1 = \|\mathbf{y}_1\|_2$.
(ii) Choose $\widetilde{\mathbf{w}}_1$, for e.g., $\widetilde{\mathbf{w}}_1 = \mathbf{r}_0$; $M_2^T\mathbf{z}_1 = \widetilde{\mathbf{w}}_1$, $\xi_1 = \|\widetilde{\mathbf{z}}_1\|_2$
(iii) $\gamma_0 = 1$, $\eta_0 = -1$
(iv) For $i = 1, 2, \ldots$ do
  (a) if $\rho_i = 0$ or $\xi_i = 0$, method fails
  (b) $\mathbf{v}_i = \widetilde{\mathbf{v}}_i/\rho_i$; $\mathbf{y} = \mathbf{y}/\rho_i$
  (c) $\mathbf{w}_i = \widetilde{\mathbf{w}}_i/\xi_i$; $\mathbf{z} = \mathbf{z}/\xi_i$
  (d) $\delta_i = \mathbf{z}^T\mathbf{y}$; if $\delta_i = 0$ method fails
  (e) $\widetilde{\mathbf{y}} = (M_2)^{-1}\mathbf{y}$
  (f) $\widetilde{\mathbf{z}} = (M_1^T)^{-1}\mathbf{z}$
  (g) 1. If $i = 1$, $\mathbf{p}_1 = \mathbf{y}$; $\mathbf{q}_1 = \mathbf{z}$
      2. else $\mathbf{p}_i = \mathbf{y} - (\xi_i\delta_i/\epsilon_{i-1})\mathbf{p}_{i-1}$;
         $\mathbf{q}_i = \mathbf{z} - (\rho_i\delta_i/\epsilon_{i-1})\mathbf{q}_{i-1}$
      3. end if
  (h) $\widetilde{\mathbf{p}} = A\mathbf{p}_i$
  (i) $\epsilon_i = \mathbf{q}_i^T\widetilde{\mathbf{p}}$; if $\epsilon_i = 0$ method fails
  (j) $\widetilde{\mathbf{v}}_{i+1} = \widetilde{\mathbf{p}} - \beta_i\mathbf{v}_i$
  (k) $\mathbf{y} = M_1^{-1}\widetilde{\mathbf{v}}_{i+1}$
  (l) $\rho_{i+1} = \|\mathbf{y}\|_2$
  (m) $\widetilde{\mathbf{w}}_{i+1} = A^T\mathbf{q}_i - \beta_i\mathbf{w}_i$
  (n) $\mathbf{z} = (M_2^T)^{-1}\widetilde{\mathbf{w}}_{i+1}$
  (o) $\xi_{i+1} = \|\mathbf{z}\|_2$; $\theta_i = \rho_i/(\gamma_{i-1}\beta_i|)$; $\gamma_i = 1/\sqrt{1+\theta_i^2}$; if $\gamma_i = 0$ method fails.
  (p) $\eta_i = -\eta_{i-1}\rho_i\gamma_i^2/(\beta_i\gamma_{i-1}^2)$
  (q) 1. if $i = 1$, $\mathbf{d}_1 = \eta_1\mathbf{p}_1$; $\mathbf{s}_1 = \eta_1\widetilde{\mathbf{p}}$
      2. else $\mathbf{d}_i = \eta_i\mathbf{p}_i + (\theta_{i-1}\gamma_i)^2\mathbf{d}_{i-1}$; $\mathbf{s}_i = \eta_i\widetilde{\mathbf{p}} + (\theta_{i-1}\gamma_i)^2\mathbf{s}_{i-1}$
      3. end if
  (r) $\mathbf{x}_i = \mathbf{x}_{i-1} + \mathbf{d}_i$; $\mathbf{r}_i = \mathbf{r}_{i-1} - \mathbf{s}_i$
  (s) Check for convergence
  (t) end

The pseudo-code for the Bi-CGSTAB algorithm can be summarized as follows:

Algorithm 2: Preconditioned BiCGSTAB
Input: $A, M = M_1M_2 \in \mathbb{R}^{n\times n}$ invertible, $\mathbf{b}, \mathbf{x}_0 \in \mathbb{R}^n$.
output: Approximate solution of $A\mathbf{x} = \mathbf{b}$.

(i) $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$.
(ii) Choose $\widetilde{\mathbf{r}}_0$ such that $\mathbf{r}^T\mathbf{r} \neq 0$ (for e.g., $\widetilde{\mathbf{r}}_0 = \mathbf{r}_0$)
(iii) $\rho_0 = \alpha = \omega_0 = 1$
(iv) $\mathbf{v}_0 = \mathbf{p}_0 = 0$
(v) For $i = 1, 2, \ldots$ do
  (a) $\rho_i = \widetilde{\mathbf{r}}_0^T\mathbf{r}_{i-1}$
  (b) $\beta = (\rho_i/\rho_{i-1})(\alpha/\omega_{i-1})$
  (c) $\mathbf{p}_i = \mathbf{r}_{i-1} + \beta(\mathbf{p}_{i-1} - \omega_{i-1}\mathbf{v}_{i-1})$
  (d) $\widehat{\mathbf{p}} = M^{-1}\mathbf{p}$
  (e) $\mathbf{v}_i = A\widehat{\mathbf{p}}_i$
  (f) $\alpha = \rho_i/\widetilde{\mathbf{r}}_0^T\mathbf{v}_i$
  (g) $\mathbf{h} = \mathbf{x}_{i-1} + \alpha\widehat{\mathbf{p}}_i$
  (h) If $\mathbf{h}$ is accurate enough, then set $\mathbf{x}_i = \mathbf{h}$ and quit.
  (i) $\widehat{\mathbf{s}} = M^{-1}\mathbf{s}$
  (j) $\mathbf{s} = \mathbf{r}_{i-1} - \alpha\mathbf{v}_i$
  (k) $\mathbf{t} = A\widehat{\mathbf{s}}$
  (l) $\omega_i = \mathbf{t}^T\mathbf{s}/\mathbf{t}^T\mathbf{t}$
  (m) $\mathbf{x}_i = \mathbf{h} + \omega_i\widehat{\mathbf{s}}$
  (n) If $\mathbf{x}_i$ is accurate enough, quit

(o) $\mathbf{r}_i = \mathbf{s} - \omega\mathbf{t}$

IV. Computational Results

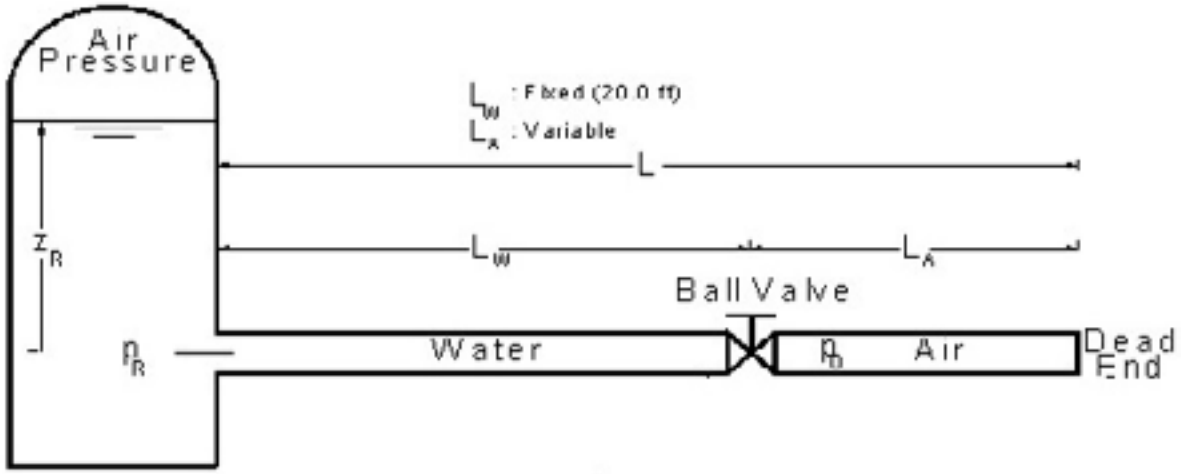


Fig. 1 Schematic of water hammer experimental setup [10].

In Figure 1, a long pipe is attached to a reservoir containing liquid water at one end, and it is closed at the other end, as shown in Fig. 1. The liquid water and entrapped air regions in the pipe are separated by a ball valve. The dimension of the pipe and other controlling parameters, such as reservoir-to-air pressure ratio, length of air column, etc., are taken from [10] so that the numerical results can be compared to the experimental data. The ball valve is closed until about 0.15s, and then gradually opens to 100% in about 0.4s. The two most important controlling parameters for this problem are the reservoir pressure (p_R) and the fractional air length present in the pipe as compared to the total pipe length ($\alpha_g = L_A/L$). The initial length for the water volume in the pipe (L_w) is fixed to 20 ft, and initial length of air column in the pipe (L_A) varies from a low of 1.23 ft to 16.23 ft, the value of α ranging from 0.0579 to 0.448, respectively. The ratio of reservoir pressure to the initial pressure of the entrapped air ($P_R = p_R/p_0$) varies from 2 to 7, i.e., the reservoir pressure (p_R) range being 29.4 to 102.9 psi. The pipe diameter is 1.025 in. The entrapped air and water are initially at 14.7 psia and 60^0F, respectively. The objective of this study is to predict the transient pressure at different points along the length of the pipe. A methodology has been developed to model the dynamics of the liquid – air interface by coupling the mass and momentum-conservation equation of liquid system and the thermodynamic equation of state for the air [1]. The methodology has been implemented in the finite volume procedure [6] of the GFSSP to model the experimental setup of Lee [10].

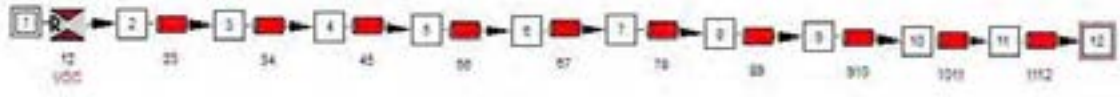
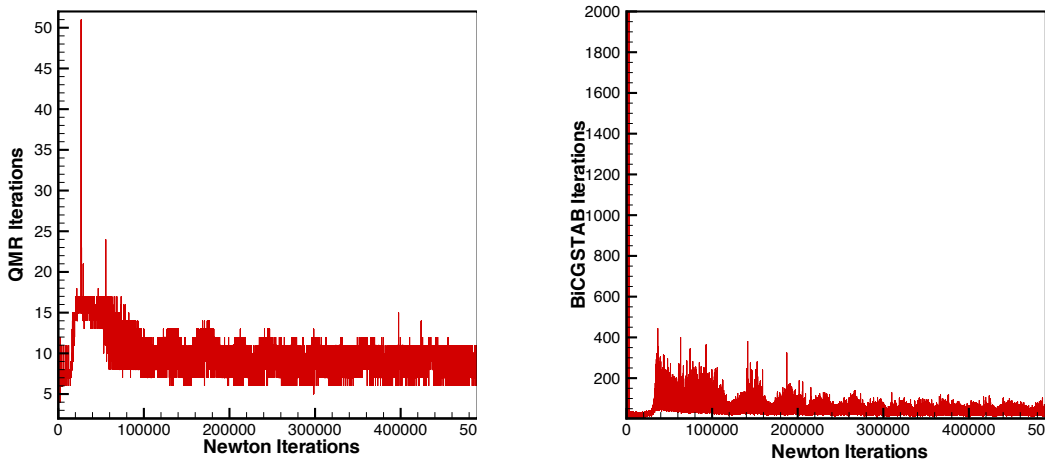


Fig. 2 A ten-branch GFSSP model.

The physical domain is split into a set of finite volume with a number of segments, as shown in Fig. 2. Node 1 is the boundary node that represents the reservoir. Node 12 has an interface with an imaginary control volume containing air



(a) QMR **(b) BiCGSTAB**

Fig. 3 Number of linear iterations against number of Newton iterations

only. The imaginary control volume has a fixed amount of air, but the volume changes as it is pressurized due to the fluctuation of pressure at node 12. Thereby, the volume of node 12 changes as the volume of the imaginary control volume changes. The entire liquid column is divided into 10 equal length pipe segments. The pressure, temperature, and mass flow rate are computed in each of the internal nodes and the velocity is computed in each branch. The ILU preconditioned QMR and BiCGSTAB linear solvers are incorporated into the solution process for the large scale network flow problems simulated by GFSSP. The effectiveness of the iterative linear solvers is proved by solving the water-hammer (pressure surge) problem described above. All the numerical experiments reported in this paper were carried out for about 45% initial air volume ($\alpha = 0.4491$). For this model problem, we compare the numerical results obtained from the preconditioned QMR and BiCGSTAB with the experimental data. The initial guess was the zero vector, and the iteration was stopped when a reduction of the initial residual by six orders of magnitude had been achieved. Our numerical experiments show that the approximate solution converged even when a reduction of the initial residual by three orders of magnitude. For nonlinear problems the accuracy of the linear solution can impact the convergence of the Newton nonlinear solver. In some cases it may be necessary to manually specify the iterative solver relative tolerance to improve the convergence of the Newton solver or to improve performance. The maximum number of iterations for the linear solver is set at 2000. If the residual does not converge to tolerance within the maximum number of linear iterations, the linear iterative solver is said to have encountered a non-convergence. Although non-convergence was never observed with preconditioned QMR solver, it did occur with BiCGSTAB solver during the early stages of time accurate simulations. However, the analysis will continue running and in this case the Newton-Raphson iterations continue to converge, see Figure 3(b).

Figure 4 shows the transient pressure at the end of the pipe with $\alpha = 0.449$. As observed from this figure, numerical results using the Krylov subspace methods (QMR and BiCGSTAB) matches quite well with that of the experimental data. Table 4.1 compares the runtimes of the two Krylov subspace methods. It also compares the number of Krylov iterations of the two iterative methods. It can be seen that in each case ILU preconditioned QMR performs better than BiCGSTAB.

In addition to the robustness issues (relating mainly to the rate of convergence or stagnation), the iterative solver is expected to outperform the direct sparse solver only for large-scale network flow models (even when the model

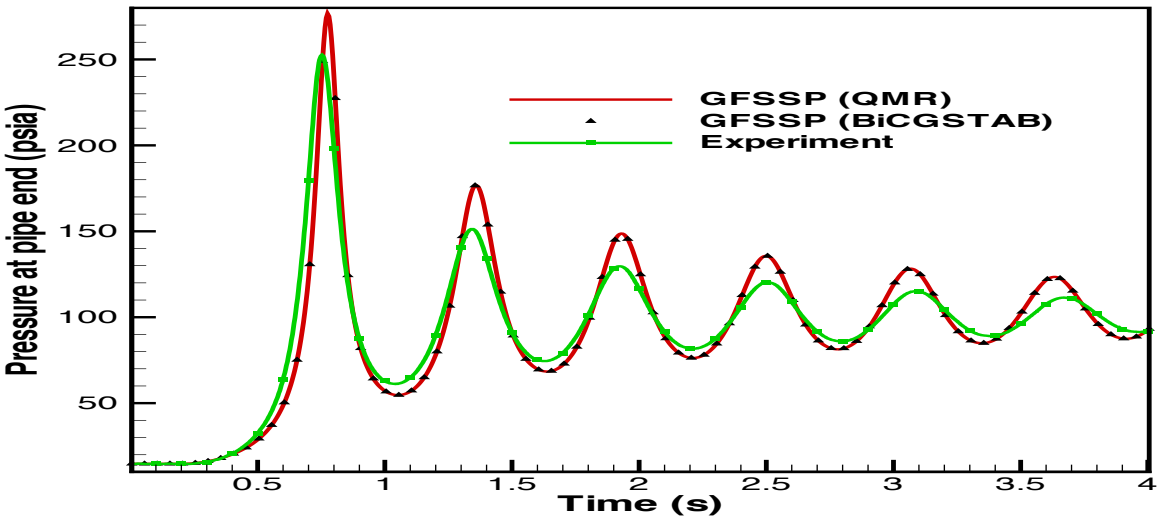


Fig. 4 Predicted air pressure using preconditioned QMR algorithm for $P_R = 7$ at about 45% initial air volume ($\alpha = 0.4491$). Also shown are the experimental data [10].

is well conditioned) that require a very large number of floating point operations for factorization. Typically, for a well-conditioned network flow model, the number of degrees of freedom must be greater than one thousand before the iterative solver will be comparable to the direct solver (Gaussian elimination with partial pivoting) in terms of run time.

Table 4.1. Comparison of computational time, number of linear and Newton iterations for QMR and BiCGSTAB solvers.

Linear solver	Number of linear iterations	Number of Newton iterations	CPU (seconds)
QMR	2.0827×10^7	516507	1067
BiCGSTAB	2.6689×10^7	539307	1211.5

V. Conclusion

In this paper, we have presented an ILU preconditioned QMR and compared its performance with another Krylov subspace method, namely Bi-CGSTAB. We compared these methods in terms of convergence and run times. Overall, our results show that QMR with ILU preconditioner tends to be the most effective for network flow problems in propulsion systems. This is because many of these problems lead to ill-conditioned matrices. Our investigations also draw attention to Bi-CGSTAB. Note that with effective preconditioner, Bi-CGSTAB converges almost as smoothly as preconditioned QMR. However, QMR is still faster than Bi-CGSTAB. For future work, there are several topics which could be considered to further improve the linear solver's performance in GFSSP. It could be advantageous to consider use of more advanced preconditioners (e.g., ILU(0)) that balances the overhead of its computation against the stability that it adds. The QMR algorithm presented here follows the two term recurrence version without look-ahead, presented by [3]. This version of QMR is simpler to implement than the full QMR method with look-ahead, but it is susceptible to breakdown of the underlying Lanczos process. Other implementational variations are whether to scale Lanczos vectors or not, or to use three-term recurrences instead of coupled two-term re-currences. These decisions we believe will have implications for the stability and the efficiency of the algorithm.

Acknowledgments

The author Dr. S.S. Ravindran would like to thank the support of NASA Marshall Faculty Fellowship program for this research. Dr. Ravindran also would like to thank the guidance and support of Dr. Alok Majumdar, Dr. Robert J

Kenny of ER43 at NASA MSFC, Dr. Frank Six, the director of NASA MSFC Faculty Fellowship Program, and Dr. Gerald Karr of the Alabama Space Grant Consortium.

References

[1] A. Bandyopadhyay and Alok K. Majumdar, *Network Flow Simulation of Fluid Transients in Rocket Propulsion System*, Journal of Propulsion and Power, **30**, No. 6 (2014), pp. 1646-1653.

[2] R. Fletcher, *Conjugate Gradient Methods for Indefinite Systems*, in: Lecture Notes in Mathematics, vol. 506, Springer-Verlag, Berlin, Heidelberg, New York, 1976, pp. 73–89.

[3] R.W. Freund and N.M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numerische Mathematik, **60** (1991), pp. 315–339.

[4] M.R. Hestenes, E.L. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards Section, B **49** (1952), pp. 409-436.

[5] C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. Res. Nat. Bur. Standards, **49** (1952), pp. 33-53.

[6] Alok K. Majumdar, A.C. LeClair, R. Moore and P.A. Schallhorn, *Generalized Fluid System Simulation Program*, Version 6.0, NASA TM-2013-217492, Oct. 2013.

[7] Alok K. Majumdar and S.S. Ravindran, *Numerical Prediction of Conjugate Heat Transfer in Fluid Network*, AIAA Journal of Propulsion and Power, **27** (3), pp.620-630, 2011.

[8] Alok K. Majumdar and S.S. Ravindran, *Fast, Nonlinear Network Flow Solvers for Fluid and Thermal Transient Analysis*, International Journal of Numerical Methods for Heat and Fluid Flow, **20**(6), pp. 617-637, 2010

[9] N.H. Lee and C.S. Martin, *Experimental and Analytical Investigation of Entrapped Air in a Horizontal Pipe*, Proceedings of the 3rd ASME/JSME Joint Fluids Engineering Conference, American Soc. Of Mechanical Engineers, Fairfield, NJ, July 1999, pp. 1–8.

[10] N.H. Lee, *Effect of Pressurization and Expulsion of Entrapped Air in Pipelines*, Ph.D. Thesis, Georgia Inst. of Technology, Atlanta, Aug. 2005.

[11] P. Sonneveld, *GS, a fast Lanczos-type solver for nonsymetric linear systems*, SIAM J.Sci. Statist. Comput., **10** (1989), pp. 36–52.

[12] H.A. Van der Vorst, *Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems*, SIAM Journal on Scientific and Statistical Computing, **13**, pp. 631-644.

Promoting Routing Protocol and Data Visualization for Wireless Sensor Network

Md A. Salam[1]
Southern University, Baton Rouge, Louisiana, 70813

And

Kosta Varnavas[2]
NASA Marshall Space Flight Center, Huntsville, Alabama, 35812

Abstract

The application of wireless sensors in space mission control is tremendous. Wireless sensors drive down the spacecraft complexity and cost by reducing mass and power requirements. Wireless sensor nodes are operated by batteries. It is essential that we maintain the battery life for a longer lite. There are many proposals on how to preserve the battery life. In this research, we focus on the routing protocols to prolong the network lifetime. Our proposed protocol is based on two basic protocols, namely, LEACH (Low-Energy Adaptive Clustering Hierarchy) and PEGASIS (Power Efficient Gathering in Sensor Information System). The proposed method compared favorably with these two basic protocols and our previously proposed cluster-based protocol in terms of network lifetime. Moreover, we have conducted research on sensor data visualization. Sensor data from the environment is collected and simulated using Python programming and visualized on the web for viewing it from anywhere at any time.

Nomenclature

WSN = Wireless Sensor Network
LEACH = Low Energy Adaptive Clustering Hierarchy
PEGASIS = Power Efficient Gathering in Sensor Information Systems
BS = Base Station
CH = Cluster Head
MN = Member Nodes
TDMA = Time Division Multiple Access
RF = Radio Frequency
NASA = National Aeronautics and Space Administration
MSFC = Marshall Space Flight Center
EEPROM = Electrically Erasable Programmable Read-Only Memory

I. Introduction

This report integrates the advancement of the routing protocol and data visualization techniques for wireless sensor networks. The application of wireless sensor network is enormous. In our daily life, we interact with sensors on many occasions. These sensors are deployed in security and surveillance systems, environmental monitoring, industries, precision agriculture, disaster response, automotive vehicle, military, spacecraft, underwater, and many more places of importance [1], [2]. There are three main components of a wireless sensor network, namely, monitored environments, sensor nodes, and sink. The monitored environment or sensed event can be static (for example, measuring temperature) or dynamic (for example, monitoring a mobile vehicle). Sensor nodes have sensing, computing, and communicating capabilities [3]. These nodes collaborate among themselves to sense and collect crucial data such as audio, video, seismic, or others as necessary. After collecting the data, these nodes perform computation and finally transmit data to the neighboring nodes or directly to the base station (i.e. sink) depending upon the routing protocol. The number of sensor nodes for each application varies ranging from tens to hundreds or even thousands depending upon the application and the size of the network. These wireless sensor nodes are usually powered by batteries and they are energy-constrained. Once the battery of a sensor node is dead or below a certain threshold value, it is

[1] Professor, Department of Computer Science, Southern University
[2] Space Systems Department, Electronic Design Branch/ES36

hard to replace it, and therefore, the node is considered as a dead node. There are many research proposals that focus on how to prolong the battery life, i.e. the network lifetime by using the routing protocols [4]. An energy-efficient routing protocol can improve the lifetime of a network and therefore the degree of network performance [5]. It is necessary to design a wireless communication protocol that will maximize the node's lifetime and minimize the node failure by collaborating with neighboring nodes [6], [7].

There is a great demand for sensor network research in space applications. We can highlight the following necessities for space research.

- In order to accomplish precision landing and hazard avoidance on the Moon, Mars and other planetary surfaces where Earth GPS is not available, a suite of Guidance, Navigation, and Control (GNC) sensors are required [https://explornet.nasa.gov/events/3232]
- Lower mass and power requirements drive down spacecraft complexity and costs, and thus are key factors for sensors [https://explornet.nasa.gov/docs/DOC-49594]
- Developing sensors technology to replicate current capabilities to view the Sun, stars, and infrared signals is very important [3D Printing in Space, National Academy of Sciences, 2014]
- The appropriate characterization, analysis, and quality control sensors have to be designed and built as part of the manufacturing equipment [3D Printing in Space, National Academy of Sciences, 2014]

In this technical report, we have presented a new wireless communication protocol which is based on two basic protocols, namely, LEACH and PEGASIS. Earlier [8], we have proposed another protocol which was also based on these two basic protocols. This newly proposed protocol outperforms our previously proposed protocol as well as these two basic protocols. In the following sections, we will briefly explain the two basic protocols, previously proposed protocol, and provide detail explanation of the newly proposed protocol. At the end, we will compare the performances of these protocols. Finally, we discussed the data visualization techniques of wireless sensor networks.

II. Basic Protocols

In this section, we will provide a brief description of three protocols: LEACH, PEGASIS, and previously proposed cluster-based protocol.

A. LEACH Protocol

The LEACH (Low-Energy Adaptive Clustering Hierarchy) protocol was developed at the MIT Lab by Heinzelman and et al. [9]. It is a self-organizing and adaptive clustering hierarchy protocol. The operation of LEACH is done into two phases: the set-up phase and the steady-state phase. The cluster formation is done during the set-up phase. Depending upon the node's energy level and the number of times it serves as a cluster head (CH), a node is chosen as a cluster head. Once a node is decided as a cluster head, it broadcasts joining request messages to its neighboring nodes. If the neighboring nodes receive multiple joining requests, it will join the one who has a stronger signal. The member nodes must inform the cluster head that they have joined in that cluster. Once a cluster is formed, the cluster head node creates a TDMA (time division multiple access) schedules among the member nodes in that cluster to receives data. In the steady-state phase, member nodes within a cluster transmit sensed data to the cluster head node of the same cluster. Once the cluster head receives all the sensed data from its own cluster member nodes, it conducts signal processing to compress the received data and transmits to the base station (BS).

Fig. 1 shows how the LEACH protocol works. The basic protocol was simulated with 100 sensor nodes with five clusters. The nodes marked with the same symbol belong to the same cluster and the cluster heads are marked with a filled dark circle. These clusters are dynamic, i.e. the cluster and cluster head change after each round. The number of cluster in a network depends upon the environments, size of the network, and other network parameters. The base station is fixed and located far from sensor nodes. These nodes are considered homogeneous and energy-constrained.

The main drawback of LEACH is that the cluster head node transmits data directly to the base station and it causes a lot of energy consumption if the base station is far from that cluster head. There are several modified version of LEACH protocols available such as LEACH-C (centralized) [10], Energy-LEACH [11], multihop-LEACH [11], LEACH-B (balanced) [12] and other [13].

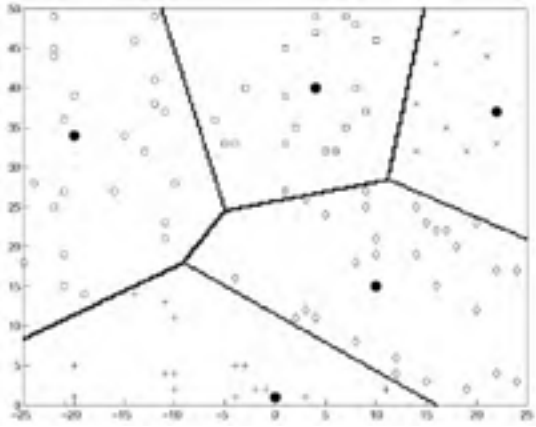


Fig. 1. LEACH protocol operation [9]

B. PEGASIS Protocol

The PEGASIS (Power Efficient GAthering in Sensor Information Systems) protocol was proposed right after the LEACH protocol at the Aerospace Corporation Lab by Lindsey and Raghavendra [14].

In PEGASIS, a greedy algorithm is used to form the chain among the sensor nodes assuming that the base station has the global knowledge of the entire network. The chain formation starts from the furthest node from the base station. This initial node (i.e. the furthest node from the base station) will connect to the nearest neighbor node and that nearest neighbor node will connect to another nearest neighbor node and this will continue until all nodes are connected in the chain based on the greedy algorithm. The node closest to the base station (BS) will be a leader node who will be responsible to transmit data to the base station. Once a node dies, the chain is reconstruction in the same manner excluding the dead node. Each node performs data fusion and transmits data to the nearest neighbor node to whom it is connected. Eventually, all the sensed information come to the leader node and it transmit to the base station (Fig. 2).

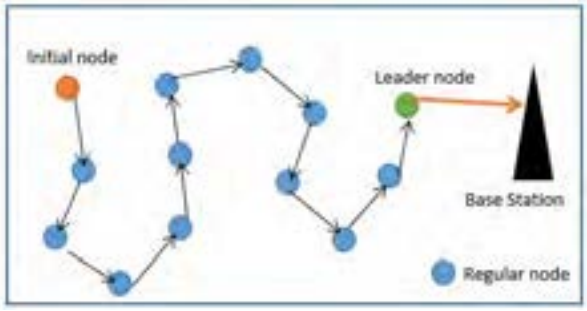


Fig. 2. PEGASIS protocol architecture

The PEGASIS protocol outperforms the LEACH protocol by approximately 2x the number of rounds when 1%, 20%, 50%, and 100% of nodes die for a 50m x 50m network [12]. Since the publication of PEGASIS, there have been many modified version of it available by many scientists around the globe. Li et al. [15] have proposed an ant colony based algorithm instead of the greedy algorithm to form the chain. Feng et al. have proposed another version of PEGASIS [16] which assigns each node weight and uses a weighting mechanism to select the transmitter node.

C. Cluster-Based Proposed Protocol

Our previously proposed protocol was based on LEACH and PEGASIS protocols [8]. In that protocol, we assume that the base station (BS) has the knowledge of the sensor node's physical location and it calculates the distance of each node from the base station. It forms clusters among the sensor nodes based upon the principle of LEACH protocol and in each cluster, the nodes nearest to the base station serve as a cluster head (CH) or leader node. Since the base station has the knowledge of each sensor node, the furthest node from the base station in each cluster will be the initial node to start the transmission to nearest node in the chain of a specific cluster as per the principle of PEGASIS protocol. Once the chain formation is done in each cluster, the base station will calculate the distances of each cluster head and the cluster head furthest from the base station will be the initial node to start transmission to the nearest cluster head node in the chain.

Fig. 3 explains the chain formation and data transmission of the previously proposed cluster-based protocol. Here we divided the nodes into five clusters with cluster heads CH1, CH2, CH3, CH4, and CH5. Each cluster has an initial node and a cluster head node. Chain formation starts at the initial node and ends at the cluster head node. Among the cluster heads (CH1, CH2, CH3, CH4, and CH5), the furthest cluster head (CH1) node from the base station becomes the initial node and nearest cluster head (CH5) node from the base station becomes the final or leader node to transmits data to the base station.

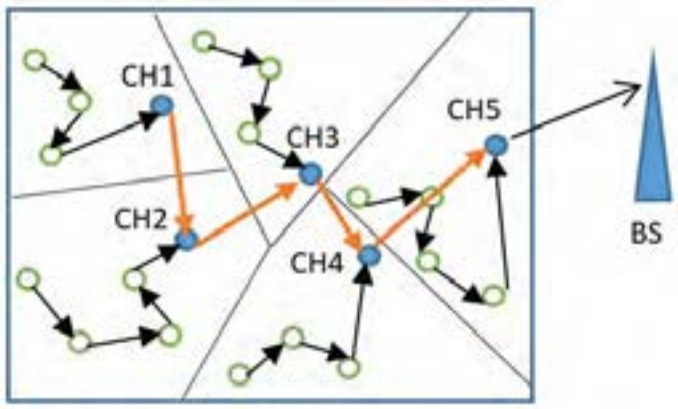


Fig. 3. Cluster-based proposed protocol architecture

III. Hierarchical Based Proposed Protocol

Our proposed protocols are based on two basic protocols: LEACH and PEGASIS. Previously proposed protocol was based on cluster and the present proposed protocol is based on hierarchy. In the following sections, we will focus on the proposed hierarchy based protocol.

A. Architecture of the Proposed Protocol

The proposed hierarchy based protocol architecture is shown in Fig. 4. The white circles represent member nodes and the black circles represent cluster head (CH) of a cluster. In this proposed algorithm, it is assumed that the base station (BS) has the knowledge of the physical location of each sensor node and BS calculates the distances of all the sensor nodes. The nodes are categorized into different levels based upon their distances from the BS. In this proposed architecture, the entire network is divided into three hierarchical levels and sensor nodes in each level form clusters among themselves. The cluster formation is performed based on the LEACH algorithm. As in LEACH, in each cluster, there is a cluster head and the cluster head is chosen depending on the distance from the base station (BS) and energy level. The chain formation among the nodes in a cluster and among the cluster heads of various clusters is done based on the PEGASIS algorithm.

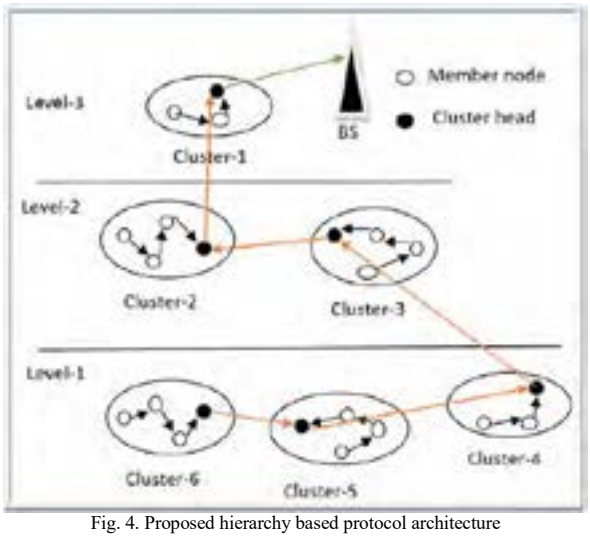


Fig. 4. Proposed hierarchy based protocol architecture

The working principle of the proposed architecture is described in the following steps.

Step-1: The BS calculates the distances of all the nodes of the entire network.
Step-2: The various levels of the nodes are categorized based on the distances.
Step-3: In each level, clusters are formed based on the working principle of LEACH algorithm.
Step-4: In each cluster, the node nearest to the base station is chosen as a cluster head node.
Step-5: In each cluster, the chain formation is done based on the PEGASIS algorithm. The furthest node from the BS will be the initial node and the cluster head will be the final or leader node.
Step-6: Among all the cluster heads in the same level, the furthest distance CH from the base station will be the initial data transmission node and the cluster head nearest to the base station will be the final node. This final node will transmit data to the next initial node of the next level which is also chosen in the same manner.
Step-7: Finally, the last cluster head in the last level will send all the data from the entire network to the base station.

In Fig. 4, there are six clusters, three in level-1, two in level-2, and one in level-1. After the formation of chains in each level and each cluster, i.e. among the six clusters, the cluster head in cluster-6 initiates the data transmission (since it is the furthest CH from the base station) and it sends data to cluster head 4 through cluster head 5. The cluster head 4 transmits data to the next level to cluster head 3. After receiving the data from cluster head 3, the cluster head 2 transmits it to the cluster head 1. Finally, cluster head 1 transmits the entire network data to the base station.

B. Simulation of the Proposed Protocol

In our simulation, we have considered 60 sensor nodes to analyze the network performance. Java program is coded according to the proposed algorithm. The base station is located at (100, 100). Initially the all the nodes in the network will have same energy of 1 Joule per node. In each round of the transmission, each node will transmit 2000-bit data packet to the base station. When the energy dissipation of a particular node goes below the threshold value, it is considered as a dead node. The following figures (Figs. 5, 6, and 7) show the status of the nodes at their different rounds. The green color represents alive node and the red color represents a dead node. The star at the upper right corner at the location of x=100 and y=100 represents the base station.

Fig. 5 displays the initial set up and (x, y) location of all the 60 sensor nodes. It is assumed that all the sensor nodes have the same amount of energy to start with. They are all alive at this round of the network.

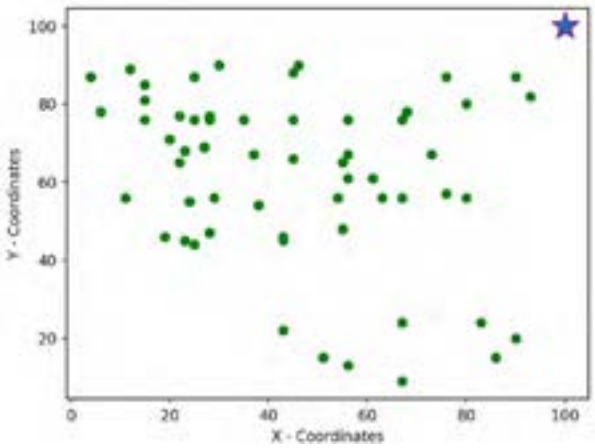


Fig. 5. Initial setup of sensor nodes for the proposed protocol.

After the 10th round, we see only six nodes are dead and most of them are alive (Fig. 6). Unlike the minimum transmission energy (MTE) protocol or direct transmission protocol as described in [9], [10], where the initial energy dissipation is concentrated either closest areas to the base station or furthest areas from the base station, respectively. In MTE protocol, the closer nodes are being used by the further nodes as a router to transmit data to the base station in the chain which causes the nearest nodes to die earlier. Whereas, in direct transmission protocol, each node transmit directly to the base station which causes the further nodes to die faster compared to the nearer nodes. In either of these cases, areas where sensors are dead become unmonitored. In the proposed method, we have noticed that dead nodes are not concentrated in one single area rather it is distributed among various areas and we still are able to monitor the field of deployment without any disruption.

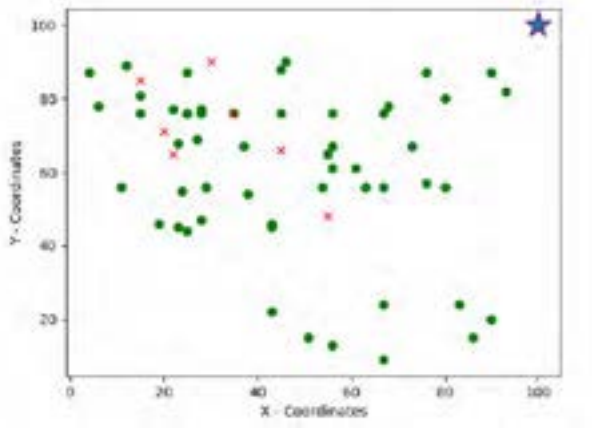


Fig. 6. Nodes status after 10th round for the proposed protocol.

The Fig.7 shows the various nodes status after the 30th round. Here we notice that the number of survival nodes are more compared to other protocols. Moreover, the alive nodes are not concentrated in one single area rather it is distributed almost all over the deployed areas which makes this protocol more promising compare to other protocols.

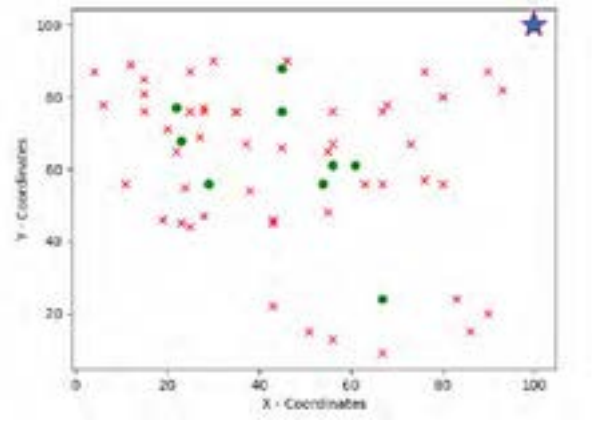


Fig. 7. Nodes status after 30th round for the proposed protocol.

The survival status of different nodes are summarized in Fig. 8. We can deduct from this figure which node is disappearing at what round. For example, node number 60 dies at round 10 and node number 2 dies in round 15.

The complete list of all the 60 sensor nodes are provided in Table 1. The first column shows the 60 sensor nodes in sequence starting from node number 1 through node number 60. The second column displays the x and y coordinate of a sensor node. For example, the (x, y) coordinate of node 20 is (56, 61) and the (x, y) coordinate of node 40 is (20, 71).

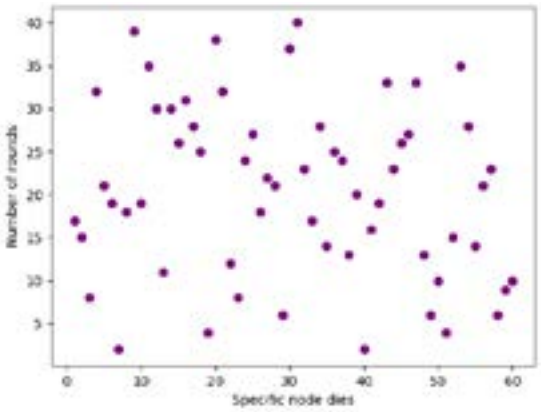


Fig. 8. Nodes' survival status.

The third column represents the distances of each node from the base station. The base station is located at (100, 100). For example, the distance from node 10 to the base station is 42.6 and the distance from node 50 to the base station is 70.7. The distances are calculated as follows. For example, node number 31 is located at (x_{31}, y_{31}) = (45, 56) and base station which is located at (x, y) = (100,100) can be calculates as:

$$d = \sqrt{[(x-x_3)^2 + (y-y_3)^2]} = \sqrt{[(100-54)^2 + (100-56)^2]} = 63.7$$

In all of these three protocols (PEGASIS, proposed cluster and proposed hierarchy), we kept the sensor nodes at the same physical location, i.e. these sensor nodes are static.

Table 1. Node's location, distance, and the round at which a specific node dies for three protocols

Sensor nodes	Location	Distance from BS	PEGASIS	Proposed Cluster	Proposed Hierarchy
1	76,14	89.3	12	26	17
2	12,89	88.7	15	25	15
3	15,85	86.3	20	24	8
4	67,24	82.9	3	20	32
5	80,56	48.3	2	5	21
6	25,44	93.6	3	28	19
7	45,66	64.7	8	13	2
8	28,47	89.4	25	27	18
9	45,76	60	7	12	39
10	73,67	42.6	5	3	19
11	29,56	83.5	25	21	35
12	27,69	79.3	25	17	30
13	37,67	71.1	3	14	11
14	56,67	55	4	9	30
15	46,90	54.9	2	7	26
16	22,77	81.3	5	18	31
17	11,56	99.3	9	18	28
18	25,76	78.7	20	16	25
19	22,65	85.5	15	23	4
20	56,61	58.8	7	11	38
21	61,61	55.2	20	14	32
22	24,55	88.3	4	26	12
23	35,76	69.3	4	15	8
24	4,87	96.9	25	29	24
25	51,15	98.1	12	31	27
26	67,76	40.8	5	8	18
27	43,22	96.6	25	28	22
28	83,24	77.9	10	20	21

29	90,87	16.4	1	6	6
30	63,56	57.5	13	15	37
31	54,56	63.7	3	16	40
32	43,46	78.5	12	15	23
33	25,87	76.1	20	18	17
34	67,56	55	25	12	28
35	15,76	88.3	7	24	14
36	19,46	97.3	15	30	25
37	56,76	50.1	9	10	24
38	76,87	27.3	2	7	13
39	23,45	94.6	8	16	20
40	20,71	85.1	6	22	2
41	68,78	38.8	1	9	16
42	38,54	77.2	8	22	19
43	23,68	83.4	3	23	33
44	67,9	96.8	20	31	23
45	56,13	97.5	20	32	26
46	25,76	78.7	4	24	27
47	45,88	56.3	8	12	33
48	28,77	75.6	2	20	13
49	35,76	69.3	13	17	6
50	30,90	70.7	6	18	10
51	55,48	68.8	3	15	4
52	28,76	75.9	17	21	15
53	55,65	57	10	9	35
54	43,45	79.2	13	26	28
55	80,80	28.3	8	8	14
56	6,78	96.5	5	11	21
57	76,57	49.2	9	10	23
58	86,15	86.1	12	28	6
59	93,82	19.3	4	6	9
60	15,81	87.1	10	29	10

The fourth column demonstrates the lifetime of each sensor when using the PEGASIS algorithm. For example, node 49 dies in round 13 and node 59 dies in round 4.

The fifth column displays the lifetime of sensor nodes when we apply the proposed cluster-based algorithm. For example, node 7 dies in round 13 and node 57 dies in round 10.

The sixth column shows the nodes' lifetime status for the proposed hierarchy based protocol. For example, 46 dies in round 27 whereas node 54 dies in round 28.

If we look carefully at Table 1, we find that the survival status of the proposed hierarchy based protocol is better compared to the PEGASIS and proposed cluster-based protocols. For example, node 31 dies in round 40 in case of proposed hierarchy based protocol whereas the same node dies earlier in rounds 3 and 16 for PEGASIS and proposed cluster-based protocol, respectively. Only a few cases some of our nodes are dying earlier than PEGASIS and proposed cluster-based protocol due to their location and functionality. For example, node 40 dies in round 2 for the proposed hierarchy based protocol whereas it dies in round 6 and 22 for PEGASIS and proposed cluster-based protocol, respectively. In overall, the nodes live longer for the proposed hierarchy based protocol compared to other two protocols. This instance is shown in Fig. 9. Here, it is showing the various percentage of nodes death at different round. For example, the survival rate of first 10% nodes is better in the proposed cluster-based protocol compared to the proposed hierarchy based protocol and PEGASIS. The proposed hierarchy based protocol compared favorable after the death of 50% nodes. The 100% disappearance of all the nodes occurs after approximately in round 26, 32, and 40 for PEGASIS, proposed cluster, and proposed hierarchy protocol, respectively.

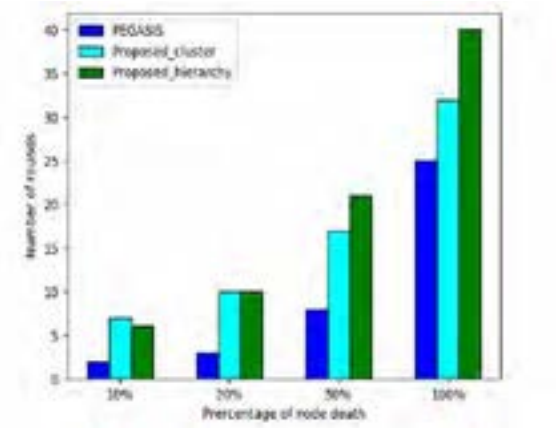


Fig 9. Performance results of PEGASIS, cluster-based proposed protocol, and hierarchy based proposed protocol

Here we consider that the network lifetime of the wireless sensor network exists until the last surviving sensor node dies or the energy of the last sensor node goes below the threshold value to be considered as a dead node. Fig. 10 compares the lifetime of these three protocols.

Nodes are started to die earlier in PEGASIS protocol and the entire network became dead after the 26th round of data transmission. Whereas, in the proposed cluster-based protocol, the entire network is alive until their 32nd round. Among these three protocols, the proposed hierarchy based protocol outperforms other protocols, it survives till the 40th round of data transmission. Both of the proposed cluster-based and hierarchy based algorithms extend the lifetime of the network compared to LEACH and PEGASIS. Since PEGASIS outperforms LEACH, we have not compared here with LEACH protocol.

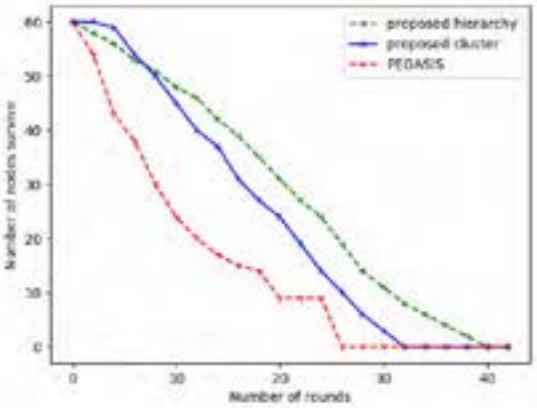


Fig. 10. Comparison of network lifetime among the PEGASIS, proposed cluster and proposed hierarchy based protocols

IV. Data Visualization

One of the most important tasks of a wireless sensor network is to monitor the environment remotely. There are two main parts of the WSN applications: deployment area and remote monitoring & controlling area. The deployment area is the environment where we deploy the sensors to monitor the activities of interest such as temperature, motion, humidity, and others. In the monitoring area, we remotely collect the data from the sensors and serially feed to the computer as shown in Fig. 9. The first block showing the sensors that we deployed in the field. The second block is an RF receiver and receiving data wirelessly from the sensor field and feeding to a computer using a serial port. The third block is a computer or a decision support unit that is used to process the data and monitor the field. The final bock is a data visualization. The authorized users can visualize the data over the cloud through a web server from anywhere and anytime.

Figure 9. Schematic diagram for sensor network and data analyzer.

V. Hardware Components

In our experiments, we used sensors manufactured by the Synapse Wireless Inc. [18] which is located in Huntsville, Alabama. We used sensor with Model SM200, PART RF200P81. The RF200P91 is a reliable through-hole module that integrates with wireless networks. It has a line of sight communicate range up to 1500 feet and it can tolerate temperature ranging from -40° C to +85° C. It has a 4-kilobyte internal EEPROM (Electrically Erasable Programmable Read Only Memory). We deployed five of these sensors in the field to monitor temperature from the surrounding area. The left side of Fig. 10 shows these sensor nodes that we have used and the right side of that figure shows the wireless RF receiver. This receiver is feeding data to a computer through a serial port.

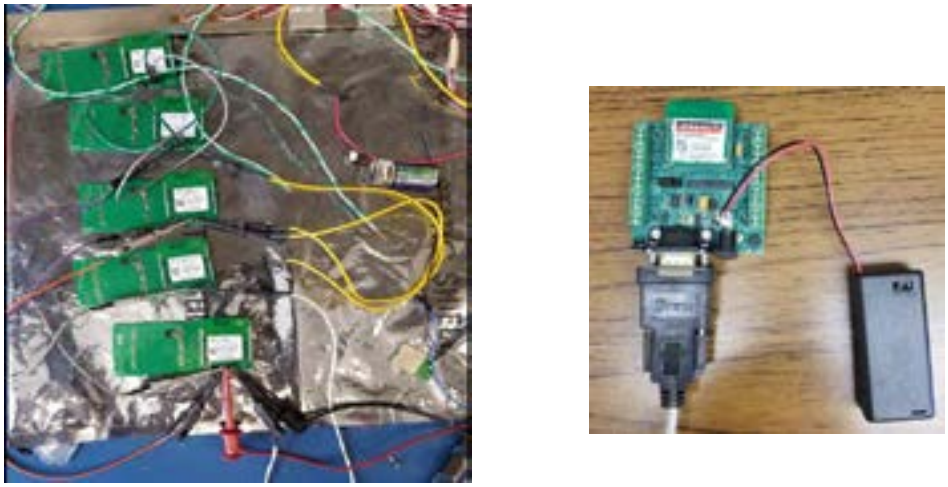


Figure 10. Sensor array (left) and RF data receiver (right).

VI. Data Analysis

Sensor data visualization is achieved using Python programming. We have used *matplolib*, which is a Python 2D plotting library for producing publication-quality figures [19]. Figure 11 displays the data from five individual sensors, namely, sensor 1, sensor 2, sensor 3, sensor 4, and sensor 5 Due to their placement location, different sensors are providing different temperatures values in their raw format. For example, the temperature of sensor 1 varies from 467 to 469 in their raw format whereas, the temperature of sensor 2 varies from 463 to 465. In the X-axis, we are keeping only fresh 50 samples and removing the older values. We can change the display duration using our Python program.

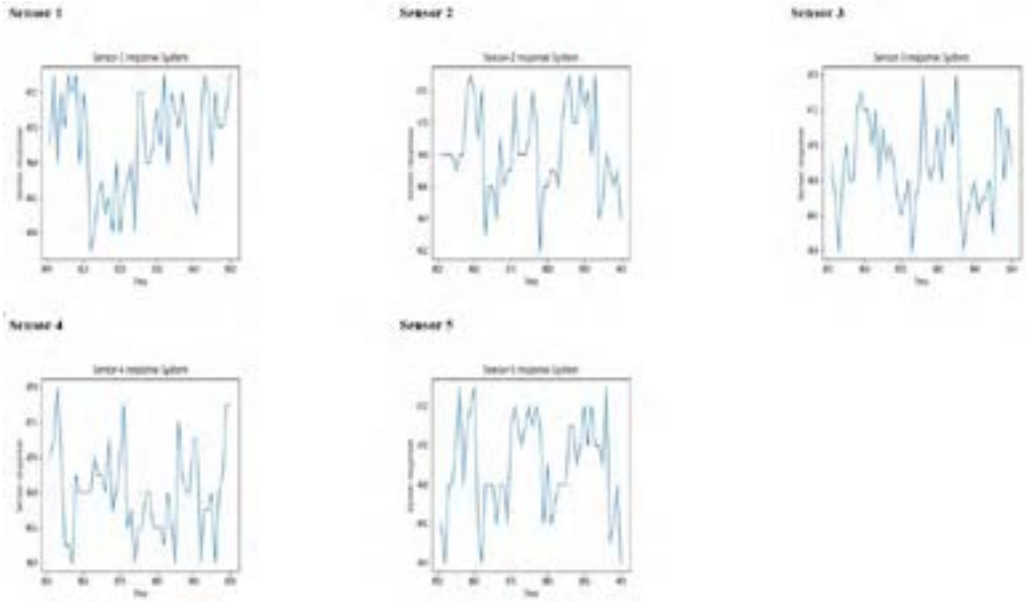


Figure 11. Sensor data display in line graph

We can display the graph as per our demand in different format and shape. Figure 12 shows the five sensors in bar graph format. Here the temperature is ranging from 464 to 469 in their raw format in a room environment. We kept only fresh 50 samples of the temperature. Based on the application, we can easily change the number of samples to be displayed.

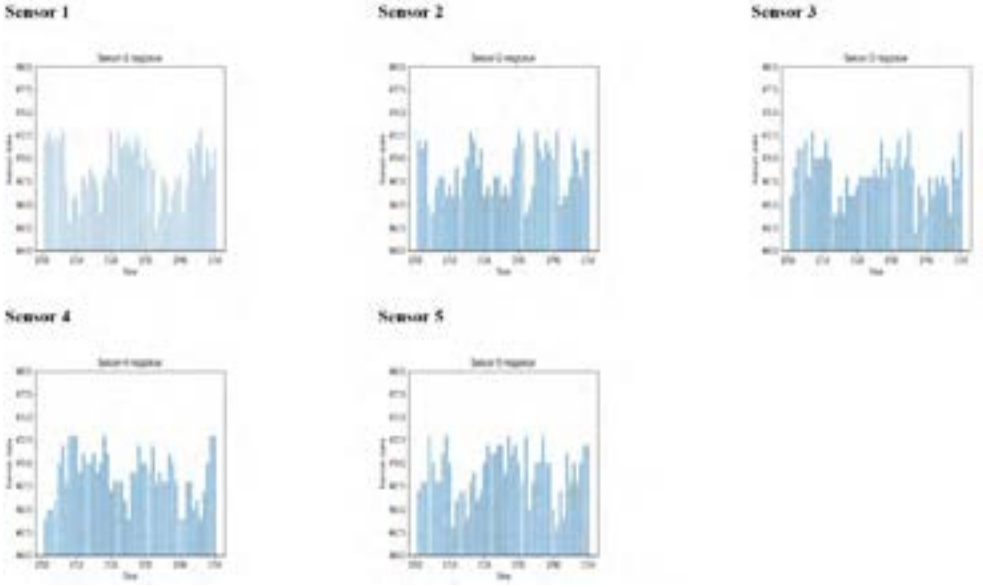


Figure 12. Sensor data display in bar graph

VII. Conclusion

In this paper, we have proposed a new routing protocol for wireless sensor network that is based on the two basic protocols, namely, LEACH and PEGASIS. LEACH is a clustering-based adaptive hierarchy based protocol and PEGASIS is a chain based protocol. There have been tremendous investigations to explore these two protocols since they are published. Many modified versions of these protocols are available in the research domain as well as in the field of application. Initially, we proposed the cluster-based algorithm [8] and then we proposed here the hierarchy based algorithm. This hierarchy based algorithm outperforms our previously proposed cluster-based protocol and it is more energy-efficient.

In our simulation, we have considered that the sensor nodes are stationary for the duration of their lifetime and all of them are homogeneous, i.e. of the same type. We have only considered the lifetime of the network in our simulation since it is a very crucial parameter for a sensor network.

In the future, we plan to consider a dynamic network with mobile sensors. Moreover, we plan to include other network factors such as network delay, bit error rate, data compression, and propagation delay of network. We also would like to explore on sensors that could survive at cryogenic to very high in aerospace application [17]. Our future research direction could include batteries that may operate adequately in these extreme environments.

In data visualization, we have considered only five sensors. In practice, we may have hundreds of sensors. Our programs are easily upgradable to accommodate any large number of sensors.

Acknowledgments

The authors would like to thank NASA MSFC Faculty Fellowship program and their staff members for providing the opportunity to conduct this research. We also would like to thanks engineers and scientists at NASA MSFC ES36 group who took time and shared their thoughts and ideas on using wireless sensor network technologies in space applications.

References

[1] M. Othmana and K. Shazalib, "Wireless sensor network applications: a study in environment monitoring system," *International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012),* Procedia Engineering 41 (2012), pp. 1204–1210.

[2] K. Bouabdellaha , H. Noureddine, and S. Larbi, "Using wireless sensor networks for reliable forest fires detection," *The 3rd International Conference on Sustainable Energy Information Technology (SEIT 2013),* Procedia Computer Science 19 (2013), pp. 794–801.

[3] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," Computer Networks 38 (2002) 393–422.

[4] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," Ad Hoc Networks 3 (2005) 325–349.

[5] K. Romer and F. Mattern, "The Design Space of Wireless Sensor Networks," *IEEE Wireless Communications*, Vol.11, Issue 6, Dec. 2004.

[6] S. Sin, M. Woo, and C.Raghavendra, "Power-aware routing in mobile ad hoc networks," *In proceedings ACM/lEEE Mobicom'98*,1998.

[7] M. Zorzi and R. Rao, "Energy Management in Wireless Communications," *In Proceedings 6th WINLAB Workshop on Third Generation Wireless Information Network*, Mar. 1997.

[8] M. Salam, K. Varnavas, and B. Maddireddy, "Prolonging wireless sensor network lifetime using routing protocol," *in Proceedings of the 6th IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE),* Dec. 11-13, Huntsville, Alabama, USA.

[9] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in the *Proceedings of the Hawaii International Conference on System Sciences*, Jan. 4-7, 2000, Maui, Hawaii.

[10] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," IEEE Trans. On Wireless Communications, vol. 1, no. 4, Oct. 2002.

[11] F. Xiangning and S. Yulin, "Improvement on LEACH protocol of wireless sensor network," *in the Proceeding of the 2007 International* Conference on *Sensor Technologies and Applications (SENSORCOMM 2007)*, Valencia, Spain.

[12] M. Tong and M. Tang, "LEACH-B: An improved LEACH protocol for wireless sensor network," *in the Proceeding of the 6th IEEE International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, Chengdu, China, Sep. 2010.

[13] H. Shwe, A. Kumar, and P. Chong, "Building efficient multi-level wireless sensor networks with cluster-based routing protocol," *KSII Transaction on Internet ad Information Systems,* vol. 10, no. 9, pp. 4272-4286, Sep. 2016.

[14] S. Lindsey and C. Raghavendra, "PEGASIS: power efficient gathering in sensor information systems," *in the Proceedings of the IEEE Aerospace Conference*, Big Sky, Montana, Mar. 2002.

[15] T. Li, F. Ruan, Z. Fan, J. Wang, and J. Kim, "An Improved PEGASIS Protocol for Wireless Sensor Network," *in the IEEE Proceeding of the 2015 3rd International Conference on Computer and Computing* Science (COMCOMS), Hanoi, Vietnam.

[16] S. Feng, B. Qi, and L. Tang, "An improved energy-efficient PEGASIS-based protocol in wireless sensor networks,", *in the Proceeding of the IEEE 8th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Shanghai, China, July 2011.

[17] W. Willson and G. Atkinson, "Wireless sensors for space applications," *Sensors & Transducers Journal*, vol. 13, Special Issue, Dec. 2011, pp. 1-9.

[18] Synapse Wireless, Inc. www.synapsewireless.com.

[19] Matplotlib, www.matplotlib.org.

2019 NASA-MSFF Final Report:

Some Logistics and Maintenance Concerns in the Gateway: Issues and Optimal Solution Methodologies

Bhaba R. Sarker[1]
Louisiana State University, Baton Rouge, LA 70803

and

H. Charles Dischinger, Jr.[2]
NASA Marshall Space Flight Center, Huntsville, AL 35808

Abstract

In the Gateway, logistical and maintenance activities that include spare parts, tools and inspection and repair will be needed frequently, as the launching spacecraft will be adding new modules and equipment to outfit Habitat with life supporting facilities for longer time as a service station to land man on the Moon's surface and to explore deep space. In this endeavor, new logistics and maintenance problems will need to be solved within specified time limitation. This paper addresses some important issues relating to suppliers' compliances, lateness in assembly and disassembly or maintenance operations, reliability of the operational system, availability of service facilities, spare parts location problem and general logistics problems, and proposes the respective solution methodologies that may be applied to Gateway operations. Finally, some general findings and recommendations are made as to the improvement of design and operational concept of the Gateway.

Nomenclature

C = $\{c_p;\ p = 1, \ldots, P\}$, penalty vector (dollars/day)
c_p = Penalty cost in phase p (dollars/day), p = 1, …, P
d_j^p = Due date of the element $j \epsilon G_p$ in phase p
G_p = Group of task or elements/components being processed or worked on in phase p
L_j^p = Lateness (in days) of element $j \epsilon G_p$ in phase p (p = 1, …, P).
r_c = Ready/current time of review
σ_j = Standard deviation of processing times of task $j \epsilon G_p$
Q_j = unreliability/unavailability of a component j
R_j = Reliability/Availability of a component j
t_j^p = Expected processing time of the component $j \epsilon G_p$
τ_j^p = Remaining time of the element $j \epsilon G_p$
Z = Total penalty cost (equivalently, minimizing lateness of maximizing system reliability)

I. Introduction

The deep-space outpost, known as the *Gateway,* is now a key element of NASA's exploration plans beyond Low Earth Orbit, This effort intends to put astronauts down near the Moon's south pole by 2024. The Gateway whose estimated service life in the orbit is about 15 years will provide an orbiting base around the Moon from which astronauts can descend to the lunar surface or go farther into space. This outpost will help humanity extend its footprint out into deep space and enable a variety of interesting scientific and commercial activities on and around the Moon.

[1] Elton G. Yates Distinguished Professor, Department of Mechanical and Industrial Engineering, 3290T Patrick F. Taylor Hall, Louisiana State University, Baton Rouge, LA 70803 (bsarker@lsu.edu).

[2] Team Lead, Human Factor and Integrated Logistics Systems (MSFC-EV74), Building-4203/Room-6100, NASA Marshall Space Flight Center, Huntsville, AL 35808 (charles.dischinger@nasa.gov).

The Gateway operations are primarily categorized administratively into two phases: Phase 1 is to place a man and the first woman on the lunar surface by 2024 and Phase 2 is focused on advancing the technologies that will foster a sustainable presence on and around the Moon. Phase 2 will be a lasting and productive presence enabled by reusable systems, accessible to contributing partners for repeatable trips to multiple destinations across the lunar surface.

In the process of the Gateway Phase 1 operations, first Power and Propulsion Element (PPE) is transported to lunar orbit, followed by several other elements, including Minimal Habitation Capability (MHC), Logistics Module (LM) Ascent Element (AE), Descent Element (DE), Transfer Element (TE), and Orion. All these elements are assembled together to form the full Gateway when crewed Orion is docked. The Gateway's additional habitation volume, environmental control and life support (ECLS) systems, airlock, external robotics, and science capabilities will be provided by NASA, industry, and international partners. The docking ports provided by additional Gateway modules will allow landing system components and commercial cargo ships to dock with the Gateway for refueling and aggregation.

As additional modules are added, the Gateway's function as a reusable command module will increase. During Phase 1 operations, without a U.S. or international habitat module at the Gateway, the crew is limited to surface expeditions of no more than four days. This assumes a 21-day mission that includes a five-day crew transit to Gateway in Orion, 2 days transit from Gateway to the Moon, 3-4 days on the surface, 2 days transit back to Gateway and 5 days transit from the Gateway back to the Earth. During Phase 2, with a U.S. or international habitat module at the Gateway that includes its own substantial ECLS systems, total mission durations in the lunar vicinity can be extended from 21 to beyond 30 days, allowing for longer surface expeditions of up to two weeks (assuming supporting assets on the lunar surface are available). Longer stays at the Gateway in Phase 2 also will enhance human and robotic missions on the surface and support deep space science activities.

Despite many favorable and unfavorable criticism by aerospace and NASA experts [*cf.,* Berger 2018, Black 2019], NAC HEO (NASA Advisory Council Human Exploration and Operations) updated the situation on Human Exploration and Operations Committee Status (Bowersox 2018, Sloss 2018) and Crusan (2018) reported on the future human exploration planning for lunar orbital platform-Gateway and science. Two major documents on Gateway Concept of Operations [DSG-CONOP-001 (2018) and DSG-CONOP-002 (2019)] basically summarized the general concept of this undertaking. Gateway Phase 1 Pre-Decision Report [June 21, 2019] also provides some highlights on it. Recently, Neeley (2019) provided a phased approach for the proposed Gateway and Maintenance process. The most updated conceptual design for maintenance and logistics systems are given by Price (2019a, b). In all these documents, the researchers provide the conceptual design architecture of the Gateway and general policy the contractors should follow to streamline the Boots on the Mission by 2024, but none indicated how the internal maintenance operations will be performed by the crews or by the automation.

When the Gateway will be operational, it will have various types of manned and unmanned activities in the station such as handling logistics problems (cargo handling, stowing, stacking, disposal, etc.) and maintenance problems (needed replacement, adjustment, or repair of components etc.). Since the details of these are not explicitly mentioned, here, we intend to address some important issues relating to suppliers' compliances, lateness in assembly and disassembly or maintenance scheduling operations, reliability of the operational system, availability of service facilities, spare parts location problem and general logistics problems. In this research, we investigate those problems that may creep into in the Gateway operations and prescribe some solution approaches to address these problems for more efficient and effective operations.

It may be noted that, in this research, parts, tools, elements, components, jobs, and tasks have been interchangeably used to mean the same point of discussion depending on the problem classifications such as a system, configuration, structure, component assembly/disassembly, maintenance, and/or logistics operation. For example, when we say 'job' or 'task', it could be a cleaning task or turning a screw in maintenance, a part in assembly/disassembly, or a subsystem in the Gateway operation. Therefore, the interpretation of it is left with the reader's perception in the context of discussion.

II. Problem Description

While preparing for the Gateway project, many unknowns to complete the Artemis mission are to be investigated as to the PPE (Power and Propulsion Equipment) in 2022, MHAB (Manned Habitat) in 2023, LM (Lunar Module) in 2024, and Orion and human landing systems (AE, DE and TE). In completing all these Gateway elements, spare parts or tools will be frequently needed, as the logistics spacecraft will be adding new equipment. MHAB needs to be

outfitted with all life supporting facilities and elements for longer time as a service station to land humans on the lunar surface and for further exploration of the deep space (for example, including Mars missions).

Each logistics mission will occur when Gateway is unoccupied; each would be required to deliver at least 3,400 kilograms of pressurized cargo and 1,000 kilograms of unpressurized cargo to the Gateway. The logistics spacecraft must be able to remain docked to the Gateway for up to three years and depart with at least as much cargo as it delivered for disposal "to a government-approved safe and stable end-of-mission disposal orbit or other government-approved disposal location." The company providing the logistics services would be responsible for launching the cargo spacecraft and will have performed at least one successful launch prior to the first cargo mission.

The crews' productive time in the Gateway is very limited and valuable as a maintenance resource, since they do not recommend, facilitating science experiments, no more than 3 hours of maintenance work per day. Maintaining all operational activities to sustain the MHAB system operable for human living in all adverse and constrained conditions is a prime concern of the successful operations of the Gateway mission. As the astronaut's time is valuable, the following problems are pertinent with respect to the total mission time and launch deadline.

(a) Compliance and Lateness:

(i) In order to meet the deadline on 'Boots on the Moon' objective, late logistics delivery puts subsequent activity at risk. A 3-phase penalty system imposed by the NASA for late delivery of components and services influences the companies to comply with the contract. The problem is how to assess such penalties and/or warning system for them to comply with the contract/deadline.

(ii) How to determine the lateness of the ongoing manufacturing operation or assembly/disassembly jobs or the maintenance tasks while the project is ongoing.

(b)Maintenance Scheduling

(i) Given the planned and unknown (unpredictable) maintenance activities on board the Gateway, appearing in either serial or parallel operational relationship, the problem is how to best estimate the times the crews may take such that they can complete the task within the given deadline.

(ii) The maintenance sequence of the tasks will have impact on the completion time, waiting time, maximum lateness, number of tardy tasks, etc. So, there exists a problem how to schedule such assembly operations or maintenance/logistics activities on board.

(c) Reliability Issues

(i) While configuring a subsystem or replacing the parts/components in a serial or parallel configuration, the reliability issues prevail tremendously in increasing the system reliability. A study with respect to the number of parts being used, types of system configuration and reliabilities of individual parts needs to be pursued to see how such a design or a part's characteristic affects the system reliability for the Gateway.

(ii) Further, failure characteristics also affect the availability of the system; so, a repair policy is to be prescribed for a system with different component failure distribution.

(d) Spare parts/Components Locations

For EVM (external vehicular maintenance) operations, autonomous robotic maintenance will require the best location for spare parts/tools at the Gateway (Price 2019a). This study needs to investigate such problems from an optimal or best perspective given the circumstantial conditions around the robot.

(e) *Logistics Operations*

Finally, the logistical issues of supplying the payload, handling the cargo, waste disposal, and the unloading and stowing them efficiently need to be studied from different perspectives of technical, human factor issues, scheduling and maintenance.

While each of the issues is a complex problem in itself, we tried to integrate them both theoretically and numerically with detailed computations. The implication of such a problem situation in the Gateway operations is also discussed so that the NASA design and maintenance operation analysts can foresee similar issues and their ramifications, for future endeavors.

III. Controlling Gateway Suppliers' Compliance: An Optimization Problem

The Gateway operation starts from the ground level (i.e., at the manufacturers and supplier's level). Supplying the contracted materials to the Gateway system is of prime concern in the sense that its nonconformance and lateness affect the mission operations and deadlines. Here we, thus, study the impact of noncompliance issues, its lateness in completing a task and the impact in maintenance.

A. Noncompliance Penalty

At levels, 2-3, as per the announcement (cf. Gebhardt, June 17, 2019), a contractor agrees their contracted tasks (or Gateway element) must meet the launch mission deadline failure to which a penalty will be assessed by NASA. Assuming a 3-phase task completion milestone, the phases are defined here along with penalty per day as follows [see Fig 1]:

Phase-1:	Contract effective time to 12 months prior to launch time, L:	$ 1,000/day
Phase-2:	12 months prior to 6 months prior to launch time, L:	$10,000/day
Phase-3:	6 months prior to launch time, L:	$20,000/day

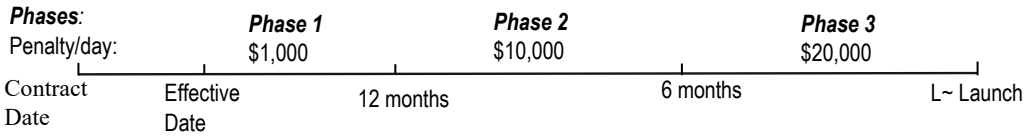


Fig. 1. Penalty cost for manufacturers or suppliers' lateness in different phases of NASA contracts

The contractor/manufacturer/supplier's contract is effective from certain date from which time NASA checks on their compliance of the completion of individual vital elements/components.

(a) The supplier is penalized for their lateness of these individual components as per the data provided above (Gebhardt, 2019). As the time approaches to launch time for a mission, the penalty is assessed at higher rate to discourage the supplier to be late, which eventually induces the higher reliability to meet the mission completion.
(b) The penalty assessment problem is further restricted or excused depending on the delay grace period and types of irrecoverable problems and other factors. For example, compliance of 90-day time window in Phase 1, and 7-day window in Phase 3, etc.

An optimization problem can be formulated to find which elements needs to be attended well to minimize the penalty (which is good for the company) resulting in reducing the risk of lateness to the launch mission deadline. For example, let G_p be the group of elements (components or tasks) being worked at a point in time on in time-phase p = 1, 2, …, 3 as defined above. Now, if L_j^p is the lateness of an element $j \epsilon G_p$ and $C = \{c_p\} = (c_1, c_2, \dots, c_p)$ is the penalty cost vector for $c_1 < c_2 < \dots < c_p$ dollars/day for phase p = 1, 2, 3, respectively, then the total penalty Z is to:

$$\text{Minimize } Z = \sum_{p=1}^{3} c_p \sum_{j \epsilon G_p} L_j^p = c_1 \sum_{j \epsilon G_1} L_j^1 + c_2 \sum_{j \epsilon G_2} L_j^2 + c_3 \sum_{j \epsilon G_3} L_j^3$$
$$= \sum_{p=1}^{3} c_p \sum_{j \epsilon G_p} max(0, r_c + \tau_j^p - d_j^p), \quad (1)$$

where $L_{j \epsilon G_p}^p = max\{0, r_c + \tau_j^p - d_j^p\}$, r_c is the ready/review (current) time, τ_j^p is the remaining time of the element and d_j^p is the due date of element $j \epsilon G_p$ in phase p. Note that the remaining time $\tau_j^p = t_j^p$, the expected processing time of the component $j \epsilon G_p$ if the component is evaluated at the beginning of its processing. It is stated in some contract information that the monitory penalty to a company may be subjected to the grace period and other time and system constraints as prescribed by the NASA contracts.

Obviously, for example, with a 3-phase lateness penalty cost, $C = (c_1, c_2, c_3)$ = (1000, 10000, 20000) dollars/day, the objective function in eq. (1) is simply written as

$$Z = \sum_{p=1}^{3} c_p \sum_{j \epsilon G_p} L_j^p = 1000 \sum_{j \epsilon G_1} L_j^1 + 10000 \sum_{j \epsilon G_2} L_j^2 + 20000 \sum_{j \epsilon G_3} L_j^3. \quad (2)$$

where $L^p_{j\epsilon\{1,2,3\}} = max\{0, r_c + \tau^p_j - d^p_j\}$. So, as a component that becomes late in any later operational/manufacturing phase, will be assessed higher penalty depending on how late the component is as it approaches to the deadline either within the company itself or to the NASA agreement. The contractor will inherently feel or be ardent to complete the project/component in time in fear of incurring high penalty cost to be assessed by NASA.

The objective function eq. (1) can also equivalently (indirectly) maximize the system reliability to meet the launching deadline because of relatively higher lateness penalty closer to the components' due dates.

B. Evaluation of Lateness

An assessment of lateness of a component either at a Gateway supplier level (ground level) or in inspection and maintenance check of different functions/activities at the Habitat in the Lunar orbit may be performed effectively if the required data are measured and collected properly. It may seem funny to penalize the crews for lateness but assessing such a measure will indirectly alert them to be prompt in completing their tasks.

The lateness penalty C basically serves as the weight factor on different stages of completion/maintenance. For example, at the supplier level which could be a contractor, a manufacturer or a transportation agency. The current time/date of review, r_c, is known at any point in time, the remaining time of completion of an activity/component, τ^p_j, is best estimated by the assessor (in this case an inspector at ground level or a crew at the HAB), and finally the completion deadline or the due date d^p_j is already scheduled in the operations chart in any phase *p*. So, the lateness function $L^p_j = max\{0, r_c + \tau^p_j - d^p_j\}$ is easily obtainable for all components. The weight or penalty function C is prescribed by the operations controller or contract document.

As an example, let's assume a single component in a single phase of operations, wherein the review (current) time r_c = August 14, 2019, the remaining time of completion τ_{k_3}= 7 days and the due date of it is $d^3_{k_3}$ = August 19, 2019 in phase p =3. So, the lateness $L^3_1 = max\{0, 14 + 7 - 19\}$ = 2 days in phase 3, which will incur a penalty cost of ($20000/day) × (2 days) = $40,000. This is just for a single component in a single phase, so this process of computations must be completed for all components in a project for all phases. In this case, if the due date would have been August 26 instead of August 19, the lateness would have been $L^3_1 = max\{0, 14 + 7 - 26\}$ = 0 day, which means that the component is completed 5 days earlier than the schedules deadline, and hence, there is no penalty.

For Gateway maintenance operations in the orbit, this time units could be in days, hours or minutes and the process of evaluation follows the similar approach. Maintenance time is very critical and important for making the Habitat time more effective and efficient to prepare for descending/ascending operation successfully. So, to illustrate it further, let the maintenance time data for some arbitrary tasks in the Gateway be given as in Table 1.

Table 1. Time data for some tasks G_p = {2, 3, 5, 7, 8} in the Gateway [Assume r_c= 8:30AM]

Task $j\epsilon G_p$	Task Description	Processing Time, t^p_j (min)	Start time	τ^p_j (min)	d^p_j (clock time)	$L^p_j = max\{0, r_c + \tau^p_j - d^p_j\}$
2	Lubrication	30	8:10	10	8:50	0
3	Servicing	55	8:15	40	9:00	10
5	Tightening	10	8:00	0	8:20	0
7	Preserving	15	8:20	5	8:30	5
8	Adjustment	35	8:25	30	8:55	5
Sum (in min):	-	145	-	85	-	20

Given the total lateness $\sum_{j\epsilon G_p} L^p_j$ is 20 minutes (see Table 1) in the Gateway maintenance operation, the crews have to decide whether they can afford to be late based on other schedules of the subsequent activities, or multiple crews have to be engaged in completing the tasks as early as possible. Since the Gateway does not have phase penalty due to lateness, the penalty function will have no direct impact on the crews but on the orbiter to meet the astronomical schedule for landing to complete all tasks within specified days in the lunar orbit. So, it will impose a burden on the crews to complete the task faster at a higher performance rate both technically and ergonomically.

As indicated earlier, t^p_j is the expected processing time of a component $j\epsilon G_p$. For non-measurable task times, practitioners usually consider *optimistic* (fast time), *most likely* (normal time), and *pessimistic* (slow) times to conduct such tests. While processing the maintenance work on board not fully autonomously, the task completion time could be variant, in which case, for non-repeating tasks, either pessimistic time (i.e., higher value) or $t^p_j \leftarrow (t^p_j + 3\sigma_j)$ should be considered to be 99.95% reliable in serial or sequential tasks (i.e., in string operations that follow one after another). On a further analysis of these arbitrary operations now calls for situation on operations sequences of the tasks that have impact on the completion of the assignment as discussed now.

IV. Assessment of the Lateness in Task Completion or Operation Maintenance

The Gateway will have numerous activities relating to logistics refurbishing/stowing, and maintenance of different regular and unscheduled tasks. Given the limited time assigned to the crews, these tasks must be completed efficiently and effectively in timely fashion to operate the whole Gateway successfully for lunar landing. So, scheduling the operational or maintenance tasks efficiently and objectively should be a prime concern of the crews and the NASA control system. Here, we thus show some methodological approaches to perform such activities.

A. Time Estimation of Serial or Parallel Maintenance Operations:

A set of maintenance tasks G_p at the Gateway could be independent of one other (e.g., knobs checking, lamp replacement, filter changing, door fixing, etc.), and hence, such a task may not affect the operations of others, or could be dependent on one another affecting the operations/maintenance sequence of this whole set of tasks (e.g., docking of logistics module, opening the door, stowing the supplies to right stacks, etc.). The first case is known as *parallel* maintenance operations while the latter case is serial/sequential maintenance.

If all independent maintenances are performed in parallel, the most optimistic upper estimated of completion time of all task is

$$T_p = \max_{j \epsilon G_p} \{t_j^p + 3\sigma_j\} \tag{3}$$

while if they are dependent (if all are in serial sequence), the upper estimated time of completion is

$$T_s = \sum_{j \epsilon G_p} t_j^p + 3\sqrt{\sum_{j \epsilon G_p} \sigma_j^2}, \tag{4}$$

where σ_j^2 is the variance of processing time of task, $j \epsilon G_p$.

To illustrate it, let's assume the data in Table 2 from which the parallel processing time, T_p = *max* {36,67,13,21,44} = 67 minutes and serial processing time, T_s =145+3$\sqrt{34}$ ≈ 162.5 minutes. Clearly, sequential maintenance operations should be avoided as much as possible to minimize the maintenance time, preferring the parallel processing of all maintenance activities as much as possible.

Table 2. Time data for some tasks G_p = {2, 3, 5, 7, 8} in the Gateway

Task $j \epsilon G_p$	t_j^p (min)	σ_j^2 (min^2)	$t_j^p + 3\sigma_j$ (min)
2	30	4	36
3	55	16	67
5	10	1	13
7	15	4	21
8	35	9	44
Computation:	$\sum_{j \epsilon G_p} t_j^p$ = 145	$\sum_{j \epsilon G_p} \sigma_j^2$ = 34	$\max_{j \epsilon G_p} \{t_j^p + 3\sigma_j\}$ = 67

Project Management Issues:

(a) Because of the precedence constraints of some of the jobs, the project may form a network. *Critical Path Method* (CPM) can be used to find the critical jobs for deterministic time parameters and some of these tasks will have slack time (i.e., flexible schedule) while other tasks will have tight (critical) schedule. Any lateness in the critical activities will incur lateness in the whole project.

(b) For variabilities in task completion times, each task may have optimistic (fastest), most likely (average) or pessimistic (slowest) time estimates while all tasks will have different durations. In such a case, *Program Evaluation and Review Technique* (PERT), originally developed by the U.S. Navy in 1952, can be applied to determine the critical activities on which the crews will have special attention to complete the project without lateness, in failure of which the availability of the system will be disturbed.

The problem is to determine the critical task element, time and the reliability of completion of task elements within certain specified time to determine the crew's maintenance time as well as to schedule the tasks on board the Gateway or on the ground to refurbish the logistics materials in time.

B. Maintenance Task Sequencing: Minimizing Lateness

The following few techniques may be valuable for some commonly encountered maintenance and logistics problems for the Gateway operation.

(a) *SPT first*: It is established from the general theory of scheduling (French 1982) that if the jobs (here tasks) are independent, then the total or mean waiting time and mean processing time of jobs can be minimized if the tasks are performed in *shortest processing time* (SPT) order; that is, the task with SPT should be performed first followed by the task with next higher processing time and so on. In other words, writing mathematically, schedule n jobs in the group G_p as

$$S_{SPT}: < j: j \epsilon G_p > \text{ such that } t_{[1]} \leq t_{[2]} \leq t_{[3]} \leq \cdots \leq t_{[n-1]} \leq t_{[n]}, \quad (5)$$

where $t_{[k]}$ (k = 1,2, …, n) indicates the processing time of job j, t_j, in the kth position of all processing times arranged in ascending order. Let's take an example in Table 3 where a 7-job processing sequence is given the last row.

Table 3: Shortest Processing Time (SPT) Rule for G_p = {1,2,3,4,5,6,7}

Task j:	1	2	3	4	5	6	7
Processing time, t_j	8	5	2	9	4	3	10
Processing order, $t_{[k]}$:	$t_{[5]}$	$t_{[4]}$	$t_{[1]}$	$t_{[6]}$	$t_{[3]}$	$t_{[2]}$	$t_{[7]}$

Here, $t_{[1]} = t_3 = 2$, $t_{[3]} = t_5 = 4$, $t_{[5]} = t_1 = 8$ and so on. Hence, the sequence of performing all tasks following the SPT order is given as S_{SPT}: <3 - 6 - 5 – 2 – 1 – 4 - 7> which minimizes the mean waiting time as well as the mean processing time of all tasks. Such a policy in maintenance of different tasks in the Gateway or Habitat will reduce the average waiting time of sub-systems requiring maintenance.

(b) *EDD first:* Similarly, the Earliest Due Date (EDD rule) policy will minimize the maximum tardiness (lateness) of the tasks with due dates (which could be in days, hours or minutes) if they are performed according to

$$S_{EDD}: < j: j \epsilon G_p > \text{ such that } d_{[1]} \leq d_{[2]} \leq d_{[3]} \leq \cdots \leq d_{[n-1]} \leq d_{[n]}, \quad (6)$$

which means that the job which has the earliest due date should be scheduled/performed first followed by the job with corresponding higher due date and so on.

Table 4: Data for EDD Rule for critical tasks for G_p = {1,2,3,4,5,6,7}

Task j:	1	2	3	4	5	6	7
Processing time (hours), t_j	8	5	2	9	4	3	10
Due date, d_j	10	23	25	35	4	5	30
Processing order, $d_{[k]}$:	$d_{[3]}$	$d_{[4]}$	$d_{[5]}$	$d_{[7]}$	$d_{[1]}$	$d_{[2]}$	$d_{[6]}$
Completion time, C_j	15	20	22	41	4	7	32
Lateness, $L_j = max\{0, C_j - d_j^p\}$:	5	0	0	6	0	2	2

Using the due dates in Table 4, S_{EDD}: <5 - 6 - 1 – 2 – 3 – 7 - 4> which minimizes the maximum lateness (or tardiness) of all tasks to $L_{max} = T_{max}$ = max{5, 0, 0, 6, 0, 2, 2} = 6 minutes and the mean lateness of all tasks. Here the mean tardiness is $\bar{T}_{EDD} = \sum_{j \epsilon G_p} L_j \,/n$ = 15/7 ≈ 2.143 hours.

(c) Other more complex issues:

(i) *Jobs with due dates and precedence constraints:* If a set or subset of maintenance tasks are dependent (i.e., if there exists a precedence relationship among some tasks), then Lawler's algorithm may be applied to solve such a problem to minimize the maximum lateness in performing the maintenance tasks in strings of jobs. In this algorithm, the job with the latest due date without violating the precedence relationship should be scheduled last as $S_{Lawler}: < j: j \epsilon G_p >$.

(ii) *Multiple strings with due dates:* Other important aspects of maintenance activities in the Gateway operations are to minimize the maximum lateness in performing the maintenance work for multiple operations with due dates either in sequence or in parallel onboard or on the ground.

(iii) *Number of tardy jobs:* Another aspect is also to minimize the number of tardy maintenance jobs with due dates, for lack of which many other subsequent complicated issues may ensue, resulting in other regenerative problems that may be costly and/or unwanted.

(iv) *Lateness for string operations:* For minimizing the maximum lateness of maintenance, an optimum solution may be found by applying Moore-Hodgson rule for independent maintenance activities and Smith Algorithm for *string* operations (dependent tasks) [see French 1982].

Only one level of maintenance is considered for Gateway because no down-mass is available for return. Gateway maintenance can only be performed by automation (robotics or onboard software), ground intervention by the flight control team, or the on-orbit crew. As the system configuration and design mature in future, down-mass may be considered for maintenance as well for repetitive use of certain component or equipment.

V. Gateway Subsystem Reliability

The Gateway will have a host of activities involving different tasks (components/elements) to meet the mission deadline or internal maintenance or logistics schedules. There are many tasks which may be completed in parallel to others while some of these activities will have precedence requirement meaning one or more jobs have to be done before/after another set of job(s), or a group of jobs may be performed in parallel to other set(s). The complete structure of these relationships will result effectively in an activity relationship network of tasks/components. The most preceded job/element triggers the beginning of the project and completion of the last set of job(s) will be the end of the project time. The term 'project' is used to mean an assigned set of jobs to be completed. The difference of the last task completion time and the first task start time is the total time required to complete the project (which is basically the makespan of the project). The following system constraints may prevail in the project:

(a) The manpower resources are always a limitation, especially 2-4 crews in the Gateway limiting no more than 4 jobs/elements can be completed at any point in time (depending on how many crews are on board and how they perform) whereas at the ground level manpower restriction may be relaxed, but human factor and ergonomic issues may prevail as well. Multiple, parallel activities (simultaneous) may also cause human errors/mistakes affecting time to complete the task on time and eventually lowering the reliability of the system.

(b) The NASA, from ergonomic point of view, also restricts crews' total maintenance time to minimal. Given a crewed mission for 30 days a year, the crews must complete the maintenance in one month while non-crewed maintenance works at the Gateway are to be completed autonomously during the remaining time.

The reliability of the system may be evaluated using the following basic concept of different configurations of the systems to be evaluated. Depending on the complexity of the maintenance or logistics operations, whether they are in complex network typed relationship, many advanced techniques could be applied to enhance the system reliability.

A. Reliability of Serial or Parallel Systems:

In the Gateway Habitat module or, in general, any unit/subsystem of the space station is composed of many physical elements/components. The reliability of the Gateway operations depends on the reliability of the individual components or subsystems. A few basic configurations of a system and the reliability of its components (serial or parallel) are now analyzed to evaluate such a system [see Fig 2]. Let R_j be the reliability and Q_j be the unreliability of a component j $(j = 1,2, \dots, n)$ to perform its desired assigned task/function.

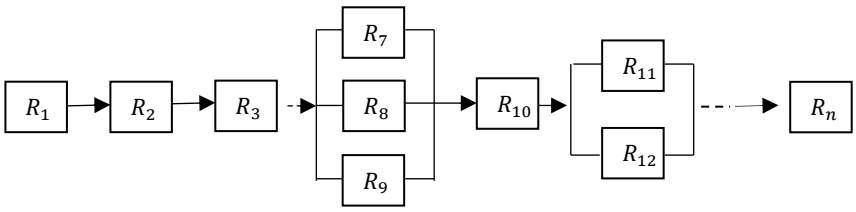


Fig. 2. Serial and parallel operations for maintenance logistics problems

(a) *Serial or chain systems:* If n components are in series (or chain) with each component reliability $R_j(j = 1,2, \dots, n)$, then the reliability of the series system, R_s, is evaluated as

$$R_s = \prod_{j=1}^{n} R_j \tag{7}$$

and the system unreliability,

$$Q_s = 1 - R_s = 1 - \prod_{j=1}^{n} R_j = 1 - \prod_{j=1}^{n}(1 - Q_j). \tag{8}$$

(b) *Parallel systems:* If n components are in parallel with each component reliability $R_j(j = 1,2, \dots, n)$, then the reliability of the parallel system is computed as

$$R_p = 1 - \prod_{j=1}^{n} Q_j = 1 - \prod_{j=1}^{n}(1 - R_j) \quad (9)$$

and the system unreliability,

$$Q_p = \prod_{j=1}^{n} Q_j = \prod_{j=1}^{n}(1 - R_j). \quad (10)$$

For example, let's take a 3-componet series sub-system G_s= (1,2,3) with respective reliabilities, R_1 = 0.9, R_2 = 0.8 and R_3 = 0.95 from which serial sub-system, we get the reliability $R_s = \prod_{j \in G_s} R_j$ = (0.9)(0.8)(0.95) = 0.684 which is very low, whereas for a parallel sub-system G_p = (7,8,9) with the same respective reliabilities R_7 = 0.9, R_8 = 0.8 and R_9 = 0.95, the parallel system reliability $R_p = 1 - \prod_{j \in G_p} Q_j$ = 1- (0.1)(0.2)(0.05) = 0.999 which is very high.

Some observations:

These results immediately reflect several scenarios that the Gateway or logistics system should follow. There are three main perspectives the manufacturers/suppliers or the crews need to follow:

(i) *Number of parts in serial/parallel configuration:* A system should have more redundant or alternative parts in parallel to increase the reliability of the system and a minimal number of sequential parts/components to decrease the unreliability in a serial structure of the system, that is, *minimize the sequential length of maintenance activities and maximize the parallelism of the components.*

(ii) *System configuration:* Since the parallel/redundant parts increase the system reliability, more redundant parts should be used in each stage of a serial system so that the stage reliability is increased effectively increasing the system reliability as well, that is*, multiple parts in parallel at each stage of a serial system will help increase its reliability*.

(iii) *Reliable/unreliable parts:* For less reliable parts, more redundant/parallel parts should be used whereas for high reliable parts, parallel configuration may be avoided to minimize the total cost, that is, *less reliable parts in parallel operations and high reliable parts in serial maintenance system in increase the system reliability*.

B. Preventive versus Corrective Maintenance: Improving the Availability

In any operation, the system could be operational (up or working) or nonoperational (down or not working) due to component or whole system failure. The mean time between two consecutive failures is known as mean time between failures (MTBF) and the mean time to failure (MTTF) is the average uptime or operational time whereas mean time to repair (MTTR) is the average of down time or nonoperational time [see Fig. 3].

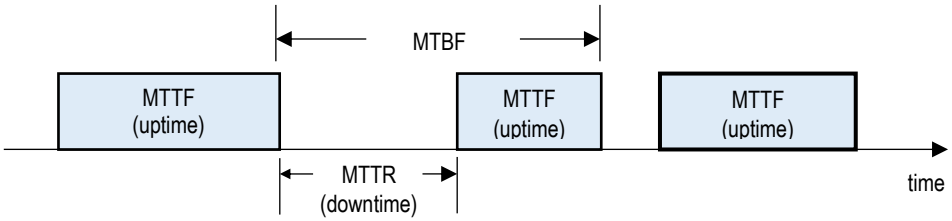


Fig. 3. Mean uptime, repair time and system availability

Mathematically MTBF = MTTF + MTTR. Hence, the reliability (R) of a system is simply given by

$$R = \frac{MTTF}{MTTF + MTTR} \quad (11)$$

If the repair time is kept at minimum, the reliability of system, R, increases; so, by employing an automatic replacement of the components can be achieved in relatively short time compared to the uptime (i.e., MTTF), the system reliability/availability can be improved to almost 100%. Thus, a subsystem with either redundant parallel components or an efficient automation process to perform the maintenance job will yield high performance.

Scheduled or preventive maintenance is performed at constant intervals of time, even if the system is still working satisfactorily. Such a process prolongs the life of components, decreases the number of failures and increases the mean time to failure (MTTF) of the system. Corrective (forced) maintenance follows in-service failures. In other words, nothing is done until the system fails. As soon as this occurs, needed replacement, adjustment, or repair of components is done to restore the system to normal operation. In a way, corrective maintenance may be thought of as repair.

In the case of parallel redundant systems, all redundant paths must fail for the system to fail. Failure of one or more redundant components will not be detected and rectified, unless there is periodic inspection and preventive

maintenance. It is obvious that the MTTF will increase with increasing (i.e., decreasing time between) frequency of inspections, the MTTF is constant and is equal to the value computed for a parallel redundant system. With proper inspection and preventive maintenance, the MTTF can be increased dramatically. By contract, if maintenance is not done properly, it is possible for the MTTF to be less than the value computed for the parallel redundant system.

The Gateway should be designed to be highly reliable. For Gateway, the operational availability (A_0) must be at the highest possible level for manned operations at the time lunar surface operation and perhaps to a lesser degree for unmanned operations. We noted that operational availability can be increased by maximizing the MTBF and minimizing the MTTR. The MTBF can be maximized by designing highly reliable systems and conducting preventive maintenance tasks that prevent unexpected failures. Price (2019a,b) mentioned that the MTTR can be minimized by incorporating common, modular, and accessible hardware designs that consider ease of maintenance; incorporating diagnostic capabilities that allow failures to be quickly identified at the lowest level; and by ensuring that necessary spare parts are available. Increasing operational availability will allow the crew to focus on mission objectives, rather than spend crew hours conducting maintenance.

VI. Spare Parts/tools Location at Gateway

Each part/component in the Gateway adds to the total weight of it that has an impact on the mass and volume restriction. Therefore, a good tool planning and multi-functional use of it will render an efficient use of these resources for the Gateway. A framework of the tool location and allocation may be proposed for the tool refurbishing process. Time will be wasted to go and fetch the tool(s) to the point of need and returning them to the storage location. This unnecessary travel of the crews is counted against the allowable maintenance time. Also, some tools may be used for some specific equipment for which locating such tools away from the needed location will also affect the crews productive time.

An optimization problem can be formulated and solved to determine the optimal combination and locations of these tools at the minimal cost, subject to other weight and volume constraints that the Spacecraft and/or Gateway will allow, but for practical purposes with limited use of them, some general thumb rule will help the crews improve the system performance and availability. Further, these can be classified in different categories using different metrics based on the characteristics of tools:

- Types, functions of tools, and flexibility
- Outer dimension (convex/polyhedral hull) to configure the space requirement;
- Weight of the tools or parts
- Tool functions needed for the Gateway operations and their requirement specifications

Tool kits can be formed in different ways depending on the functional need and location(s) of requirements. Arranging tool kits is a typical knapsack problem and the complexity and sizes of the kits will depend on those metrics mentioned above. Such tool kits or multi-functional tools or parts may be then stored accordingly. The following general principles for selecting and locating the tools may be followed to improve the crews productive time and system availability.

(a) If the special tools uniquely needed for some specific instrument, they should be located at that (dedicated) place and general multi-functional tools are to be located in general area, then the travel time of astronauts can be minimized.

(b) If tools are RFID-tagged, a computer database can also identify the location of tools when the Gateways elements are added for Phase 2 operations. For Phase 1 this storing process should be simple and easily accessible to the crews.

(c) Multiple tools and/or multi-functional tools are to be used to minimize the volume requirement and weight of the tooling system.

VII. Logistics Operations: Payload Integration

Many reports indicate that request for proposal (RFP) includes options for additional logistics services beyond the baseline transport of pressurized and unpressurized cargo. Other options include refueling, additional payload power, long-term habitation and the ability to be co-manifested on a Space Launch Services mission. Some reports indicate that NASA will provide a firm fixed-price contract for logistics services to avoid eventually costly open-end contracts when the latter is cheaper at the beginning than the first one.

The Gateway hardware is usually shipped to the payload integration site in a ready state to be integrated as a payload in a just-in-time approach. The element modules and components should not require any maintenance at the payload integration site—all tests and inspections should be done before it reaches the site. This will minimize the on-board testing, maintenance and disposal of parts/components.

The payload could of different types such as electrical and mechanical hardware, parts and components, subsystem, tools, liquid and gas, and perishable items like food and drinkable items, to mention a few. Requirement of all these items varies depending on the time in mission and the need arising at the Gateway. So each logistics module the payload should refurbished accordingly in ready-to-use form so that crews time is minimized. Each item tagged with proper identification will help sort and store them correctly in right place.

VIII. General Finding and Recommendations

While it is relatively closer to the Earth, operations in lunar orbit and on the surface will help us hone our skills for Mars mission which is very far. Of utmost importance is astronaut safety. Aboard the Gateway and on the Moon, astronauts will validate the capability to execute both nominal and contingency physical and behavioral health operations autonomously and with time delay. The Gateway also has the potential to conduct complex orbital maneuvers in the Earth-Moon system to maximize science return, and to practice the operational procedures required for transits deeper into the solar system.

On and around the Moon, we will learn how to efficiently and safely investigate extraterrestrial bodies, synergistically using automation and robotics to maximize science return, employing dust mitigation strategies, advancing space suit designs, and using in-space, on-demand manufacturing and *in-situ resource utilization* (ISRU), to reduce dependence on tools and supplies sourced from Earth.

A. On-orbit Maintenance:

(a) The element module/component projects should consider the limitations for crew time, limited tool availability, accessibility, limited stowage space, and spare availability when determining organizational maintenance.

(b) Priority of tasks and parallel operations should be identified to perform maintenance activities (for example, visual inspections, lubrication, tightening and minor adjustment, servicing, preserving, electrical work, etc.)

(c) Major assemblies/modular components causing equipment malfunction can be easily traced by using easy-to-interpret Built-in Test Equipment, caution and warning lights, simple go/no-go indicators, installed instrumentation, or easy-to-use-and-interpret external diagnostic or fault isolation devices, such as volt ohm meters [Price 2019].

(d) Replacement of major assemblies or modular components that are identified as worn out, damaged, or otherwise defective or condition monitored basis should be changed on a timely fashion change, preferably under scheduled maintenance policy to avoid catastrophic failure.

(e) Rerouting of electric connections is better and more efficient than replacement since it cost less money and time. Light structural repairs (such as straightening, fluid line repair and connector repair) should be done on board. A five-year maintenance cycle should be followed to limited preventive maintenance due to limited crew availability.

B. Loose-end Architecture

In both logistics modules and Gateway operations, many conceptual design architectures are proposed for both component configuration and day-to-day maintenance operation. It is noted that most of the reports, without mentioning any document, working paper or presentations, general suggestions or a potential prescriptions for design and operations are proposed but no alternative action plans is given when those proposed prescriptions fail or are not workable in not far-fetched future. Figure 4 illustrates such an example of having an action plan. If the action fulfils (YES) there is no problem—the proposition will follow as perceived usually, but if it does not (NO), then what to do (loose-end)? The planners should provide the complete and exhaustive guidelines for all alternatives so that no ambiguity or unsatisfied/non-complied part of it falls in the crack. A check and balance guideline and measure should be in place as a liability at level 3 downward. Otherwise, this type of incomplete information may be an expensive mistake causing abandonment of the subsystem or increase in maintenance or repair time.

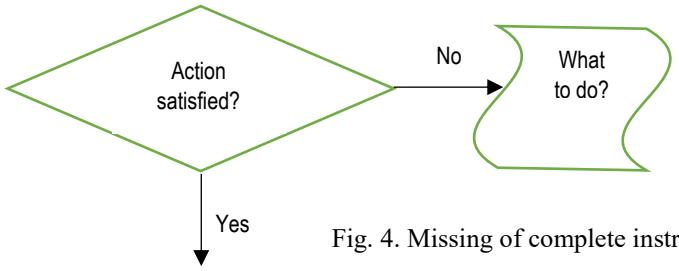


Fig. 4. Missing of complete instructional or methodological step(s)

C. Resources for Maintenance

(a) *Down-mass effect*: The Gateway maintenance concept must consider an important factor, the limited available resources throughout the life cycle of the program, which needs to consider, specifically the down-mass, up-mass, onboard storage and capacity, crew time and the Habitat environment. Since down-mass is treated almost as disposal, return of some of its items to the depot for repair should not be considered because of technical and capacity problem, and expenses involved; thus, both preventive and corrective maintenance must be done onboard. Further, since the down-mass is disposal (i.e., one-time use), it should be designed at the minimum cost for a single-time use. In advanced stage of Gateway project for longer duration, the descent unit (which is now designed to be disposal) may be build reusable to same cost and time.

(b) Modularity: Obviously, both up-mass and onboard stowage are constrained resources for Gateway. Price (2019) indicated that modular hardware designs must be applied that reduce hardware to the smallest replaceable unit and utilize common component/system interfaces that allow items to be interchangeable between modules/elements. Modularity early in design will pay off in life cycle by reducing resupply mass and enabling commonality across modules at the lowest level practical in each system (see Price 2019a). From an engineering perspective, interchangeability of parts or elements will minimize the number of uniqueness and hence minimize the number of parts/components or tools. Further, commonality of interfaces, interchangeable parts, common tools, and interoperability should be adopted as much possible so that such items can be used in multiple modules.

(c) *Crew time and automation:* Since the crew time availability for maintenance should be minimal throughout the life of *Gateway* to facilitate the astronauts to other productive function such as experiments, testing, and data collections, to mention a few. The following issues and approaches are pertinent to the Gateway manned and unmanned operations:

- Automation to identify the need for maintenance or failure of a component to the lowest practical level will minimize the amount of time required for troubleshooting by the onboard crew and/or flight operations team such that the crew time is focused on scientific experiments and the maintenance activity. Automation via onboard software and/or ground control can help in preventive maintenance and recovery from failure.
- While the early phases of the Gateway operation are dependent on crews for maintenance, robotics or automation may help do the internal maintenance to save the crews time in later phases of the program,
- External robotic capabilities should be preferred to perform the external maintenance outside the Gateway and EVA capability will help backup for contingency maintenance activities if external robotic capabilities are unable to resolve a failure. Interfacing capability of equipment for preventive maintenance should be incorporated to allow robotic assets to perform the necessary activities (such as removing and replacing components).

D. Phase 1 Maintenance Operations

As PPE and Logistics module are supplied to the Gateway with sufficient reliability to avoid immediate inspection and maintenance during Phase I, maintenance need to focus on the Mini HAB only. When the modules will be added to the Gateway, the Mini-HAB should not require corrective maintenance within first five years of its commissioning to service. Physical maintenance activities, performed by crew, can only be performed during the 30 days a year when a crew is on-board the Gateway. NASA needs to be cautious that, sending PPE and Logistics module to the Gateway does not mean that those elements might not have been thoroughly checked at the manufacturer level—the NASA must ensure the certification through various inspection and proper documentation or preferably through test or site inspection.

It is anticipated that minimal up-mass will require maintenance resources; therefore, the designs that require a minimal common set of tools to perform the maintenance should be considered on board. Only organizational maintenance is to dominate in the orbit, so no intermediate or depot maintenance is to be considered and internal maintenance should be kept at minimal. To minimize the crew time required for maintenance, items (cables, wire harnesses, and fluid lines) should be clearly labeled and easily accessible to the crew. A robust diagnostics capability should be built in to the system to identify failed items to the lowest level without relying on external test equipment. This will enable both crew and flight control teams to quickly diagnose and return failed items to an operational state, enabling increased focus on the mission [see Price July 2019a]. Further, the robotic arm will not be delivered to Gateway until in later assembly phases, so EVA capability is not available, and hence, no external maintenance is to be performed during the first phase. Extensibility to such activities will ensue when other modules are added to the Gateway after Phase 1.

E. Optimal methodologies:

This research investigated some current planning and strategic policy for the Gateway project. Knowing different facts of the situation, the authors looked at those problems from different methodological perspectives such as lateness in suppliers' compliances, scheduling of onboard activities, enhancement of reliability of systems needing maintenance allocation and location of spare parts/components, and the logistics operation. In this section, we recommend some results to improve the system performance.

(a) Sub-system Configuration and Reliability Measurement:

The parallel configuration of parts/components in a subsystem element should be considered preferably to increase the reliability. The following results are evident from the study:

(i) *Number of parts:* More redundant or alternative parts in parallel system and minimal number of sequential parts/components in serial system increases the reliability (So, minimize the sequential length of maintenance activities and maximize the parallelism of the components).
(ii) *System configuration:* Multiple parts in parallel at each stage of a serial system will increase its reliability.
(iii) *Reliable/unreliable parts:* Less reliable parts in parallel operations and high reliable parts in serial maintenance system in increase the system reliability.
(iv) *Maintenance Policy:* Preventive/scheduled maintenance is preferable to corrective maintenance for most replaceable items.

(b) Lateness and project management:

When a contract fails to meet the target deadline as spelled in the contract, subsequent operations may affect meeting their target deadlines as well, possibly resulting failure of other milestones of the mission. Therefore, a penalty assessment policy will help NASA to minimize this unwanted occurrence. The scope of this aspect can be extended to meet the deadline in maintenance activities on board the Gateway, in which case, the weight, instead of penalty cost, will be more appropriate to assess the degree of being late.

(d) Scheduling Maintenance Task:

Three-sigma upper estimation of times in evaluating the total completion time will be appropriate to keep the crews under less stress. It is proven that applications of SPT and EDD rules in scheduling jobs at ground level or on board the Gateway will certainly minimize the waiting time, the number of lateness and the maximum lateness, which are crucial measures of performance in the Gateway operation. CPM/PERT will help schedule all tasks in the project.

(e) Spare parts and Logistics

(i) More multifunctional tools and spare parts/components should be used in the Gateway to minimize both volume and weight. General-purpose parts/tools should be located in a common place while special parts/tools should be stored close to the place of need.
(ii) The items shipped through logistics module should be stacked and stowed in ready-to-use form with minimal packaging but properly labeled.

The main challenge of the crews is to identify the problems and match the solution methods correctly as described earlier. Given the detailed methodologies and the brief recommendations above, NASA as well as the associated companies will be the main stakeholders of ripping the benefits of these applications such as minimum maintenance and waiting times, minimum lateness, minimal maximum tardiness, increased reliability and availability of systems.

IX. Conclusions

As understood, Gateway is the staging station in a cislunar orbit for landing a man and a first woman on the Moon with a hope to extend other missions to deep space including the Mars. The size (volume and mass) of the Gateway define the orbit and other Mini-HAB environment. So the primary objective of this lunar mission is to refurbish the Gateway and sustain it for a longer period of at least 30 days such that crews can prepare well for landing on the lunar surface and/or prepare for launching into the deep space to hone the technology and science better for future endeavors.

Sustaining the Gateway properly will require a productive and efficient operation of it. This research highlights some existing propositions so far thought by many NASA and outside experts and proposes some new ideas as to the maintenance and reliability of system. A number of optimal strategies for maintenance activities have been emphasized to improve the system performance (minimizing maintenance time and increasing system reliability). While many of these methodologies have elucidated from the Gateway related problems through numerical examples with some general theories, the several applications are also cited from concurrent and contemporary perspectives on logistics and maintenance issues to be applied to the Gateway operation.

Angularity or the tangential thoughts from the traditional, straight designs and maintenance operations may entail some new advantages to the existing system. While the NASA projects are, in general, not a mass production system, testing the individual parts, components or elements at the ground level (manufacturers or suppliers) will help identify the mistakes or human factor issues in early stage better than discovering it at the Gateway in lunar orbit. Unfortunately, the non-electric components/elements being tested at ground level are usually not sent to space. Therefore, a serious and careful attention must be given to select flawless components when they are manufactured.

People involved in the actual operation may not know many of these techniques and applications yet the general knowledge of this information, if not even applied fully, will help mitigate the maintenance problem to an extent the reward for which will eventually be tremendous to improve the system reliability. For future spacecraft designs, modularity, repairability, upgradability and flexibility should be an important proposition in the on-orbit operation to minimize time, money and dependability on other resources.

Acknowledgements

This work was supported by *2019 NASA Summer Faculty Fellowship Program* and conducted at the Human Factors and Integrated Logistics Systems Branch (EV74), Marshall Space Flight Center, Huntsville, AL 35808. The authors are thankful to Mr. James R. Neeley, Ms. Sally Richardson and Mr. Kent Kimmons for providing some documents and valuable suggestions to complete the project. The authors express gratitude to Mr. Tommy R. Smith, EV74 Branch Chief and other Division authorities for providing the opportunities and resources. Professor Bhaba Sarker is also thankful to Dr. N. Frank Six, the Director of Summer Faculty Fellowship Program, Dr. Gerald Karr and Ms. Brooke Graham for providing the opportunity and timely assistance to work at MSFC in summer 2019.

References

Berger, E. (2018), "Former NASA administrator says Lunar Gateway is 'a stupid architecture,'" https://arstechnica.com/science/2018/11/former-nasa-administrator-says-lunar-gateway-is-a-stupid-architecture/ , November 15, 2018.

Berger, E. (2019), "Facing 2024 deadline, NASA issues a report defending the Lunar Gateway: Gateway supports the acceleration of landing on the surface of the Moon in 2024," https://arstechnica.com/science/ 2019/05/facing-2024-deadline-nasa-issues-a-report-defending-the-lunar-gateway/ , May 2, 2019.

Black, G. (2018), "The Lunar Orbital Platform – Gateway: an unneeded and costly diversion," *The Space Review, SSI 50: The Space Settlement Enterprise,* http://www.thespacereview.com/article/3676/1, *May 14, 2018.*

Black, G. (2019), "NASA's flawed plan to return humans to the Moon," *The Space Review, SSI 50: The Space Settlement Enterprise,* http://www.thespacereview.com/article/3676/1 *Monday, March 18, 2019.*

Bowersox, K. (2018), "Human Exploration and Operations Committee Status," NAC HEO Committee Meeting, Thursday, December 8, 2018. https://www.nasa.gov/sites/default/files/atoms/files/ human_exploration_ operations_committee_report_082918_tagged.pdf.

Crusan, J. (2018), "Future Human Exploration Planning: Lunar Orbital Platform-Gateway and Science Workshop Findings," Advanced Exploration Systems, NASA Headquarters, Washington, DC. https://www.nasa.gov/sites/default/files/atoms/files/20180327-crusan-nac-heoc-v8.pdf March 27, 2018.

Davis, D., Bhatta, S., Howell, Kathleen, Jang, J.-W., Whitley, R., Clark, F., Guzzetti, D., Zimovan, E., and Barton, G. (2017), "Orbit Maintenance and Navigation of Human Spacecraft at Cislunar Near Rectilinear Halo Orbits,"

Report #JSC-CN-38626 (Preprint AAS-17-269), Presented at the *27th AAS/AIAA Space Flight Mechanics Meeting*; February 05, 2017 - February 09, 2017; San Antonio, TX; United States. February 5, 2017.

Davis, J.P., Mayberry, J.P. and Penn, J.P. (2019), "On-orbit servicing: Inspection, repair, refuel, upgrade and assembly of Satellines in space," *Aerospace,* Center for Space Policy and Strategy, April 2019

DSG-CONOP-001 (2018), *Gateway Concept of Operations*, NASA, Washington, DC, August 31, 2018.

DSG-CONOP-002 (2019), *Utilization Concept of Operations (CONOPS) for the Gateway*, NASA, Washington, DC, May 7, 2019.

Foust, J. (2019), "NASA to seek proposals for Gateway logistics," *Spacenews, June 17, 2019,* https://spacenews.com/nasa-to-seek-proposals-for-gateway-logistics/, July 17, 2019.

French, S. (1982), *Sequencing and Scheduling: An Introduction to the Mathematics of the Job-Shop*, Ellis Horwood, Oxford, UK.

GAO (US Government Accountability Office), (2018), *NASA Assessments of Major Projects: Report to Congressional Committee*, Report #GAO-18-280SP, Contact: Christina Chaplain at Chaplainc@gao.gov, 441 G. Street, Washington, DC 20548; May 2019, 120 pages.

Gateway Phase 1 Gateway Pre-Decision Report, June 21, 2019.

Gebhardt, C. (2019), "NASA determines best course for commercial lunar Gateway Resupply," *NASA Spacelight.com,* June 17, 2019. [Jim Bridenstine @ JimBridenstine, June 14, 2019 and CBS News]. https://www.nasaspaceflight.com/2019/06/nasa-commercial-lunar-gateway-resupply/

Gerstenmaier, W. H. (2018), "Human Exploration & Operations Overview," Lecture slides, August 2018.

Gill, T. (2018), "NASA's Lunar Orbital Platform – Gateway," The Space Congress Proceedings, 17. February 28, 2018. https://commons.erau.edu/cgi/viewcontent.cgi?article=3711&context=space-congress-proceedings, July 17, 2019.

Hahn, J. (2016), "Looks aside, NASA's Orion is "Light years ahead of what they had in Apollo," *ARS Technica*, https://arstechnica.com/science/2016/03/looks-aside-nasas-orion-is-lightyears-ahead-of-what-they-had-in-apollo/, March 27, 2016.

HRE-XSI (2019), *I-HAB System Requirements Document (Draft)*, European Space Agency, ESA-DSG-HAB-RS-0001_1.0_IHAB_SRD, first released on July 16, 2019, pp. 1-107 [Provided by J. R. Neeley, July 17, 2019].

Kline, R. C. and Sherbrooke, C. C. (1993), *Estimating Spares Requirement for Space Station Location Freedom using the M-SPARE model,* Report NS101R2 (July 1993), Logistics Management Institute, 6400 Goldsboro Road, Bethesda, MD 20817.

Neeley, J. R. (2019), "Proposed Gateway and Maintenance Process-Phased Approached," Gateway Pre-Decision Report, B-4203/6108A. NASA Marshal Space Flight Center, Huntsville, AL; June 27, 2019.

Patterson, L.P. (2001), "On-orbit maintenance operations strategy for the international space station – Concept and implementation," NASA Johnson Space Center (source of Acquisition), https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100042525.pdf, July 18, 2019.

Price, C. (2019a), "Maintenance Concept," A whitepaper prepared for Gateway Operations, July 2019.

Price, C. (2019b), "Gateway Maintenance Philosophy," A presentation made in July 2019.

Ramakumar, R. (1993), *Engineering Reliability: Fundamental and Applications*, Prentice Hall, Englewood Cliffs, New Jersey 07632, USA.

Sloss, P. (2018), "NASA updates Lunar Gateway plans," *NASA Spaceflight.com,* https://www.nasaspaceflight.com/2018/09/nasa-lunar-gateway-plans/ September 11, 2018.

Trujillo, D.J. and Scharmer, J.B. (2012), "Reliability, Availability, and Maintainability Consideration in the Design and Evaluation of Physical Security Systems," Report #SAND2012-4740C, Sandia National Laboratories, P.O. Box 5800, MS-0785, Albuquerque, NM 87185-0785.

Computational and Experimental Approaches to Understanding the Shape Memory of Ionic Polyimides for Additive Manufacturing

Tomekia M. Simeon[1]
Dillard University, Department of Chemistry, New Orleans, LA 70122

Enrique M. Jackson[2]
NASA, Marshall Space Flight Center (MSFC), AL 35812

Jason Bara[3]
The University of Alabama, Department of Chemical and Biological Engineering, Tuscaloosa, AL 35487

Abstract

The accuracy of the Fragment Molecular Orbital method (FMO) was investigated using the Pair Interaction Energy Decomposition Analysis (PIEDA). First principles *ab initio* and energy decomposition analysis based on large-scale FMO calculations were used to obtain the interaction energies of two supramolecular complexes of polyimides broken down into fifteen fragments. Calculated interaction energies qualitatively explained the binding affinity shown in prior thermal analysis experimental results from previous summers. A comprehensive set of basis sets: STO-3G, 6-31G*, 6-311G*, 6-31++G**, aug-cc-pVDZ, and aug-cc-pVTZ was employed; along with HF and DFT methods. In this report, we present only results using HF and the DFT LPBE0AC functional with the aug-cc-pVDZ basis set. Our results indicate the HF calculated complexation energies agree qualitatively with the energetic ordering from LPBE0AC calculations with an aug-cc-pVDZ basis, both for structures dominated by hydrogen-bonding and π-π stacking interactions. When the PIEDA energies are decomposed into components, and we find that the electrostatic interactions dominate while induction and dispersion makes a significant contribution to the overall energy.

[1] Assistant Professor, Dept. of Chemistry, Dillard University, New Orleans, LA 70122.
[2] Aerospace Polymeric Engineer, Materials and Process Lab, EM22, MSFC, AL 35812.
[3] Associate Professor, Dept. of Chemical & Biological Engineering, The University of Alabama, Tuscaloosa, AL 35487.

I. Introduction

Ionic polyamides (Fig. 1) are a new, hybrid polymer that builds upon the well-known chemistries underlying Kevlar, Nomex and Nylon – some of the most important commercial polymers of the 20th and 21st centuries. However, none of these polymers exhibits shape memory and only Nylon polymers have been shown to be directly useable for 3-D printing (additive manufacturing). Recently our collaborators have developed new monomers from asymmetric imidazole-amine species and the same diacid chloride building blocks from which Kevlar, Nomex and Nylon polymers are built. When these diiimide-dimidazole monomers are reacted with a suitable co-monomer such p-dichloroxylene an ionic polyamide is formed. From these ionic polyamides, the chloride anions are then anion-exchanged with a thermally stable, non-nucleophilic anion such as bistriflimide to form melt-processable ionic polyamides which are suitable for use in fused deposition modeling (FDM) based 3-D printing as well as conventional manufacturing processes (e.g., extrusion, injection molding, fiber spinning, etc.).

Much to our surprise, the ionic polyamide materials synthesized exhibited shape memory properties. Without any special thermal or chemical processing, the ionic polyamides were observed to behave as thermoplastic elastomers. This behavior was discovered through recovery of ionic polyamide during drying in a vacuum oven.

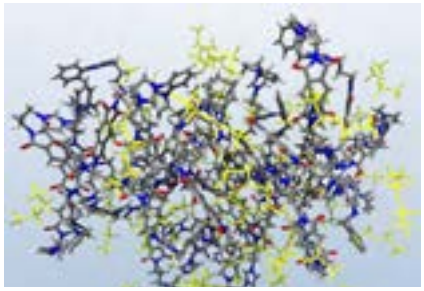


Fig. 1. Structure of ionic polyamide polymer.

While attempting to pull the ionic polyamide from a glass dish at 150°C, it was discovered that the polymer could be processed into a coil shape by wrapping it around a 1/4" diameter stainless steel rod. Upon cooling to room temperature, the coil was observed to be deformable through stretching with rapid recovery to the original size within 1-2 minutes. Further experiments were performed to produce discs of the ionic polyamide. These discs (~1/8" thick and 2" diameter) could be folded in half but would rapidly spring back to their original fully extended state.

As mentioned, these materials were not expected to have shape memory behavior and they were originally produced as a comparative study to similar ionic polyimide materials which also show promise for 3-D printing applications. However, although they are also highly thermally stable and amenable to a variety of processing techniques, the ionic polyimides do not exhibit such shape memory behavior. It is hypothesized that the additional H-bonding sites on the ionic polyamides are responsible for the shape-memory behaviors as there is very little difference in the chemical compositions of ionic polyamides and ionic polyamides other than the H-bonding donor and acceptor sties. When sliced/cut/shredded, the particles of the ionic polyamides also exhibit self-healing behavior and will fuse back together over short time scales (seconds to minutes). The physical and mechanical properties of ionic polyamides have been of interest to modelers for decades[1-5], but past work has not provided the level of understanding needed to describe all the intermolecular forces and shape-memory behaviors involved in polymer folding.

This project seeks to advance the use of ionic polyamides for 3-D printing by first providing a systematic protocol for computational modeling to further scaling up synthesis experimentally. At this present time, developing a robust thermal and mechanical property database, creating expertise in thermal processing techniques, producing filaments for FDM 3-D printing, is ideal for advancing aerospace science.

The goal of this project is twofold: 1) to use electronic structure results to determine H-bonding and π-π stacking interaction energies that can be used to calibrate for MD simulations and 2) calibrate the FMO-PIEDA which is a method that breaks the intermolecular interactions into its various components, such as, electrostatics and repulsion. Therefore, calibrating the DFT results, and providing a decomposition of the energy into components that can provide insight into the description of larger polymers. Details of the structural models used to study H-bonding and π-π stacking interactions and the calculated geometric parameters are presented in Methods. The results of our structural modeling, while the FMO-PIEDA results are presented and compared with HF results. Finally in conclusions, the main findings are summarized.

II. Computation Model

Structural Parameters: A total of two polyimides and poly(ILs) was optimized and the folded versus unfolded conformations were studied. For the purpose of this paper we focus on PMDA API ortho xylene polymer complex (Figure 1) and use it as a benchmark for future FMO-MD and FMO-PIEDA calculations. The details involving the solvation of the complex and then fragmentation is shown in Figure 2 and the short chain of the oligomer is shown in Figure 3. A fragmentation scheme to understand π-π stacking and H-bonding in the systems, is shown in Figure 4 and 5, wherein the parent molecule is broken into two PMDA API ortho xylene-I (**PMDA-API-I**) and PMDA API ortho xylene-II (**PMDA-API-II**); these geometries represent two configurations taken at various temperatures during MD modeling, 300° and 600°K.

The relative contributions of π-π stacking interactions and the H-bonding interactions to the overall complex binding is established in both complexes. The scheme for performing the single point calculations for the FMO-PIEDA, is the same as for our previous studies concerning, DFT-SAPT.[6-8] Whereas, the constrained geometry optimizations at the HF/3-21G(d) level of theory have been performed to account for the addition of hydrogen atoms to take care of dangling bonds. The final geometries are then subjected to single-point calculations and their interaction energy is then broken down to the various components.

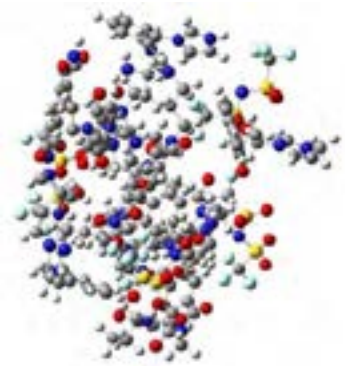


Figure 1. PMDA API ortho xylene polymer complex.

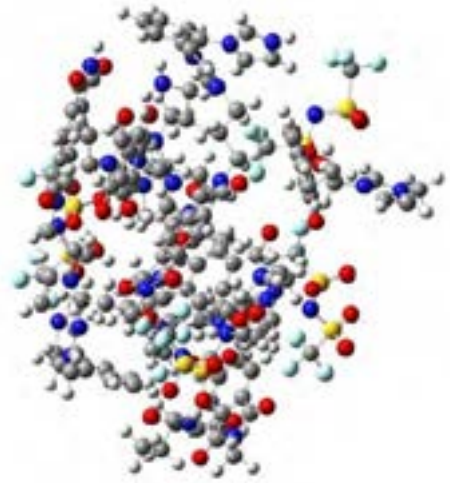


Figure 2a. PMDA-API-I/II

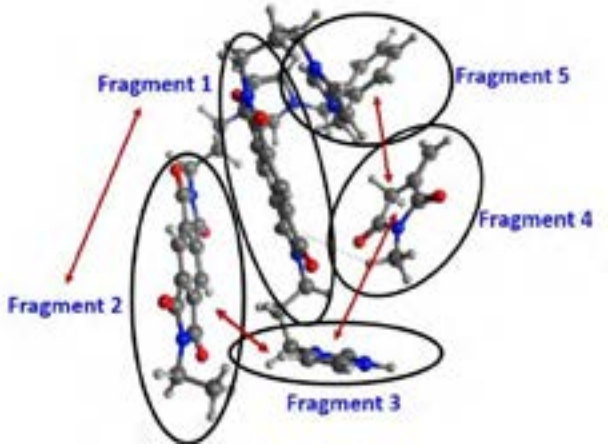


Figure 2b. Fragments in Polymer

Figure 2. Optimized PMDA API ortho xylene polymer calculated demonstrating regions of π-π stacking and hydrogen-bonding interactions. The lowest-energy conformers were then geometry optimized using the PM6 semi-empirical method – including SCRF (SMD) implicit solvation using Gaussian '16.[18,1 9] The PM6 method was employed for monomer geometry optimization using the SMD solvation model, with acetonitrile as the solvent (ε = 35.688). This model includes a parameterized function to reproduce dispersion corrections. The structure (**Polymer**) was then broken down to fragmented systems as labeled **Fragments in Polymer**.

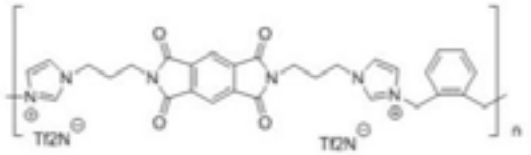


Figure 3. PMDA API ortho xylene polymer short oligomer.

Figure 4. The fragmentation scheme for π-π stacking in PMDA API ortho xylene polymer.

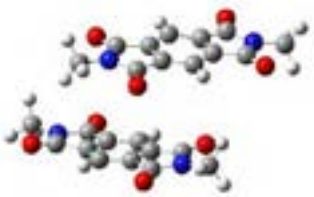


Figure 5. The fragmentation scheme for H-bonding in PMDA API ortho xylene polymer labeled.

III. Theoretical Methods

The calculations were performed using the fragment molecular orbital (FMO) method.[9] Although the classical MD calculations involve the use of standard methods, the FMO application is both new and potentially quite important given that the unusual (highly bent) structures of ionic polyamides might not be accurately described by the standard force fields.

FMO enables the use of electronic structure methods to describe large macromolecules that can be broken into discrete units. The FMO method is a practical approach for using electronic structure methods to study challenging problems for supramolecular complexes, like the polymer in Fig. 1, as all the calculations are done on fragments. Not only does this reduce computational effect but it also enables the development of calculations that scale linearly with system size, as demonstrated already for similar systems.[10-12]

For studied models the Facio software package[13] (a graphics program for molecular modeling and visualization) was used, which is a considerable improvement over the manual input needed in the past. This has a FMO gui control panel that fragments and generates a file for supramolecular complexes. The recent version is now specifically, capable of automatically defining and fragmenting individual segments, so preparing an input for over ~100 fragments (over 11,000 atoms) is straightforward.

For the 15 fragments for our test models, the next step is to compute the fragments electron density distributions. In the full quantum-mechanical calculation, each fragment is immersed in a Coulomb field due to the remaining part of the system (the environment), to which the exchange interaction with other fragments is added and the electron density is fully relaxed. In FMO, one adds the environmental Coulomb field and neglects the corresponding exchange and charge-transfer interactions. Consequently, the fragment (monomer) densities are converged self-consistently giving fully polarized fragment densities and their respective energies. In the second step, one obtains the quantum-mechanical interaction between pairs of fragments (dimers), which are accompanied by performing dimer calculations in the Coulomb field due to the remaining fragments.[14]

The FMO expression for the total energy is

$$\sum_{FMO} = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} + E_I + E_J) \quad (1)$$

Where monomer (E_I) and dimer (E_{IJ}) energies are obtained from the corresponding calculations on N fragments (monomers) and their pairs (dimers) in the external Coulomb field due to the remaining monomers. Overall, the (E_{IJ}) term is the SCF energy of dimer I + J in the external Coulomb field of the other N-2 fragments. In this regard, the task of obtaining a quantum mechanical energy is greatly simplified for large supramolecular complexes.

The FMO computations also provide Pair Interaction Energies (PIEs), also known as inter-fragment interaction energies (IFIE), between fragments.[15-17] The PIEs can be used during the analysis of interaction between polymers residues and surrounding molecules and is derived from the following

$$\Delta E_{IJ} = (E'_{IJ} - E'_I - E'_J) + Tr(\Delta D_{IJ} V_{IJ}) \quad (2)$$

where ΔD_{IJ} and V_{IJ} are the difference density matrix and the environmental electrostatic potential for dimer *IJ* from other fragments, E'_I and E'_{IJ} are the monomer energies and the dimer energies without the environmental electrostatic potential. The IFIE analysis is useful to our proposed research in that the results generated can be plotted in two-dimensions to highlight "*hot-spots*" of ionic interactions.

Similarly, the Pair Interaction Energy Decomposition Analysis (PIEDA[18]) in FMO is advantageous in elucidating the interactions associated with polymer bending, binding and intermolecular interactions.

In FMO, the PIEDA is given by

$$\Delta E_{IJ}^{int} = \Delta E_{IJ}^{ES} + \Delta E_{IJ}^{EX} + \Delta E_{IJ}^{CT+mix} + \Delta E_{IJ}^{DI} \quad (3)$$

where the total interaction value ΔE_{IJ}^{int} is decomposed into electrostatic $= \Delta E_{IJ}^{ES}$, exchange-repulsion ΔE_{IJ}^{EX}, charge-transfer plus higher order mixed terms ΔE_{IJ}^{CT+mix} and dispersion ΔE_{IJ}^{DI} contributions.

FMO was selected over other methods because it can easily be combined with a number of current electronic structure techniques including wave function-based theories that include for dispersion effects and do not suffer from self-interaction errors. The consequences of dispersion (along with electrostatic) forces in ILs are enormous, so it is essential to include them in our studies. For these model complexes, constrained geometry optimizations at the HF/3-21G(d) level of theory have been performed to account for the addition of hydrogen atoms to take care of dangling bonds. The final geometries are then subjected to single-point calculations at the HF and the LPBE0AC functional using the aug-cc-pVDZ basis set. Energies in kcal/mol.

IV. Results

We examined the errors obtained between two fragmentations of PMDA API ortho xylene polymer labeled PMDA API-I and PMDA API-II (Fig. 2). The errors of total energy of the entire complex; the sum of the all PIEDAs between each fragment in the polymers were examined for these two fragmentations at the FMO-HF/STO-3G, 6-31G*, 6-311G*, 6-31++G**, aug-cc-pVDZ, and aug-cc-pVTZ levels and with the DFT-LPBE0AC functional. Results for both methods are reported for the HF and the LPBE0AC functional and aug-cc-pVDZ basis set. The discrepancy of total energy between the two fragmentations was less than ~2.1 kcal/mol and the errors of the PIEDA interaction energies ranged from ~2.0 to 4.5 kcal/mol (Tables 1-4). Although the larger fragmentation led to more reliable estimations of energies, between each complex the errors accounted for just a fraction, the errors were negligibly small for qualitative understanding.

Figure 6 show the total interaction energies between dimer fragments within fragmented polymer **I** and **II**. As shown in Figure 6, Fragments 1-3 and notably fragments 13-15 are strongly destabilized, whereas energies indicate that the center of the complex is relatively more energetically stable, or its globular configuration is preferred. Although, HF and DFT-LPE0AC are vastly different in terms of computational efficiency, the overall energetic trend is the same in both polymers. Hence, for future calculations to save computational time and effort supramolecular complexes can be optimized, for example with HF/3-21G* basis set and calibrated with various force fields (such as, Amber and Charmm) with larger basis sets coupled single point calculations with a DFT functional.

Moreover, our previous work considered a computationally expensive effort into fragmenting the interaction energies using DFT-SAPT. [19-21] While DFT-SAPT is not scalable to significantly larger systems, this comparison in this study with SAPT and PIEDA provides an important benchmark for further improvements in DFT functionals that can be scaled. In addition, the comparison between the V*DZ* and AV*DZ* basis sets demonstrated it is much better for a balanced description to increase the basis to AVDZ size, hence the reason we used it in this study. On average, our energies overall using FMO scaled quite well with results reported in ref 19.

Since FMO scaled fairly well and is computationally less expensive, for this reason other polymer models will follow a similar protocol. It is explained now. Briefly, the minimum energy studies of the PMDA API ortho xylene polymers were isolated by means of simulated annealing (SA) molecular dynamics (MD) simulations. For them, the system is first thermalized at a given temperature and subsequently slowly cooled down towards 0 K. The calculations were performed with explicit solvation in DMSO using NAMD and the AMBER12 force field with partial charges calculated using the Extended Huckel Theory. Both force fields were developed for modeling proteins, with some recent additions for "small molecules". The aim for using "small molecules" or fragments is to observe the change in structure as the temperature is increased. One question that is of interest here is there a structural change that corresponds with an observed endotherm using experimental techniques incorporating instrumentation, such as DSC and TGA?

To continue, our systems was first allowed to equilibrate at a selected temperature for 1200 ps and subsequently cooled down to 0 K in 100 ps, 800 ps, or 1200 ps using a linear temperature ramp. A broad sampling of the spacial geometries was further enforced by employing two different initial temperatures of 300 K and 600 K, and two largely different initial geometries: fully extended and folded (Figure 7c and Figure 2a).

Note that while the interaction between the fragments are largely repulsive, this is compensated by the attractive interactions between the π-π stacking between the rings and the counter ions. This leads to an effective counter ion-driven net attractive among the various fragmented rings which accounts for almost 90% of the folding energy stabilization between fragment dimers 7-12. When considering the total energies, we see electrostatic interaction accounted for over 55% of the total stabilization energy, respectively, followed by induction (which accounts for charge) then dispersion. In other words, almost all stabilizations for these fragments were derived from the interaction with the globular configuration. These results suggest that the shape specificity was driven by the interactions between these fragments, in which electrostatic interactions including hydrogen bonding interactions are important.

Summarizing, the results indicate that the stability of the H-bonding and π-π stacked pleated conformation observed in PMDA API ortho xylene polymer is due to interactions within the complex. The organized ionic and solvent environment that the counter ions contribute provides great starting structures once they are extracted for computational modeling in vacuum.

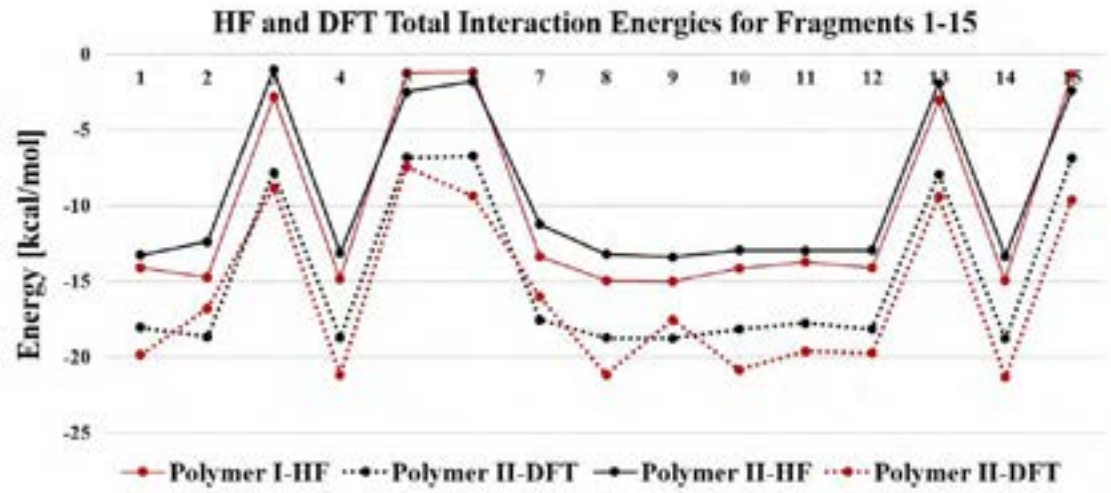


Figure 6. The total interaction energies between the pair fragments for Polymer I and Polymer II.

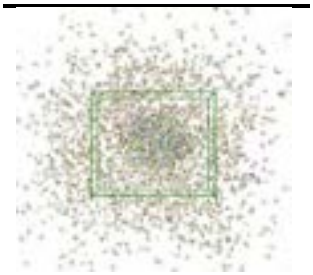


Figure 7a.
Molecular Dynamics Simulations simulation with solvent outside the PBC

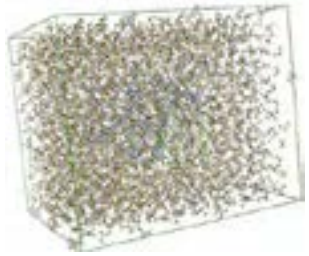


Figure 7b.
Molecular Dynamic simulation of calculation for *ab initio* calculations.

Figure 7c.

PMDA API ortho xylene structure globular structure extended.

Table 1: Stacking energies for PMDA-API-I structure from FMO Theory Approach. Interaction energy contributions obtained with using the HF functional with the aug-cc-pVDZ basis set. Energies in kcal/mol.

	Fragments	$E_{el}^{(1)}$	$E_{exch}^{(1)}$	$E_{ind}^{(2)}$	$E_{ind-exch}^{(2)}$	$E^{(1)}$	$E^{(2)}$	$E^{(1)}+E^{(2)}$
1	Frg 1-Frg.2	-45.518	43.055	-22.303	10.654	-2.462	-11.648	-14.111
2	Frg 2-Frg.3	-45.820	42.858	-22.311	10.555	-2.962	-11.756	-14.718
3	Frg 3-Frg.4	-32.037	37.216	-17.758	9.749	5.179	-8.009	-2.830
4	Frg 4-Frg.5	-45.170	41.790	-21.689	10.243	-3.379	-11.446	-14.826
5	Frg 5-Frg.6	-36.220	44.868	-21.579	11.690	8.648	-9.889	-1.242
6	Frg 6-Frg.7	-35.939	44.408	-21.301	11.666	8.469	-9.635	-1.166
7	Frg 7-Frg.8	-44.897	42.862	-21.908	10.579	-2.035	-11.330	-13.365
8	Frg 8-Frg.9	-43.613	39.633	-20.604	9.669	-3.980	-10.934	-14.915
9	Frg 9-Frg.10	-43.657	39.640	-20.660	9.678	-4.018	-10.982	-15.000
10	Frg 10-Frg.11	-45.458	42.882	-22.113	10.561	-2.577	-11.552	-14.129
11	Frg 11-Frg.12	-45.640	43.566	-22.347	10.729	-2.073	-11.618	-13.691
12	Frg 12-Frg.13	-45.491	43.004	-22.187	10.573	-2.487	-11.615	-14.102
13	Frg 13-Frg.14	-32.195	37.244	-17.835	9.763	5.050	-8.0716	-3.022
14	Frg 14-Frg.15	-45.193	41.685	-21.665	10.220	-3.508	-11.445	-14.954
15	Frg 15-Frg.16	-36.140	44.644	-21.463	11.637	8.504	-9.826	-1.322

Table 2: Inter/intra energies for PMDA-API-I structure from FMO Theory Approach. Interaction energy contributions obtained with using the LPBE0AC functional with the aug-cc-pVDZ basis set. Energies in kcal/mol.

	Fragments	$E_{el}^{(1)}$	$E_{exch}^{(1)}$	$E_{ind}^{(2)}$	$E_{ind-exch}^{(2)}$	$E_{disp}^{(2)}$	$E_{disp-exch}^{(2)}$	$E^{(1)}$	$E^{(2)}$	$E^{(1)}+E^{(2)}$
1	Frg 1-Frg.2	-43.106	48.868	-26.836	15.583	-16.010	3.474	5.763	-23.789	-18.027
2	Frg 2-Frg.3	-43.431	48.716	-26.845	15.475	-16.017	3.460	5.285	-23.926	-18.641
3	Frg 3-Frg.4	-31.049	41.708	-21.828	13.961	-13.439	2.861	10.658	-18.446	-7.787
4	Frg 4-Frg.5	-42.801	47.601	-26.145	15.082	-15.844	3.416	4.800	-23.492	-18.692
5	Frg 5-Frg.6	-35.055	49.668	-26.253	16.510	-14.931	3.250	14.613	-21.425	-6.812
6	Frg 6-Frg.7	-34.791	49.217	-25.966	16.481	-14.896	3.245	14.427	-21.136	-6.709
7	Frg 7-Frg.8	-42.579	48.591	-26.428	15.504	-16.122	3.498	6.011	-23.549	-17.538
8	Frg 8-Frg.9	-41.288	45.210	-24.825	14.269	-15.363	3.276	3.922	-22.644	-18.723
9	Frg 9-Frg.10	-41.357	45.283	-24.913	14.294	-15.314	3.264	3.926	-22.669	-18.743
10	Frg 10-Frg.11	-43.084	48.681	-26.656	15.503	-16.061	3.488	5.597	-23.726	-18.129
11	Frg 11-Frg.12	-43.266	49.392	-26.943	15.727	-16.194	3.531	6.125	-23.879	-17.753
12	Frg 12-Frg.13	-43.157	48.859	-26.741	15.504	-16.071	3.474	5.702	-23.834	-18.132
13	Frg 13-Frg.14	-31.174	41.739	-21.902	13.978	-13.432	2.859	10.564	-18.497	-7.932
14	Frg 14-Frg.15	-42.808	47.491	-26.111	15.050	-15.824	3.409	4.684	-23.475	-18.792
15	Frg 15-Frg.16	-34.976	49.445	-26.119	16.441	-14.901	3.243	14.468	-21.336	-6.867

Table 3: Stacking energies for PMDA-API-II structure from FMO Theory Approach. Interaction energy contributions obtained with using the HF functional with the aug-cc-pVDZ basis set. Energies in kcal/mol.

	Fragments	$E_{el}^{(1)}$	$E_{exch}^{(1)}$	$E_{ind}^{(2)}$	$E_{ind-exch}^{(2)}$	$E^{(1)}$	$E^{(2)}$	$E^{(1)}+E^{(2)}$
1	Frg 1-Frg.2	-46.353	44.985	-23.050	11.159	-1.369	-11.891	-13.259
2	Frg 2-Frg.3	-49.614	50.291	-25.523	12.465	0.677	-13.058	-12.381
3	Frg 3-Frg.4	-35.356	43.604	-20.697	11.433	8.248	-9.265	-1.017
4	Frg 4-Frg.5	-50.160	50.262	-25.685	12.473	0.102	-13.213	-13.111
5	Frg 5-Frg.6	-33.293	39.429	-18.853	10.227	6.137	-8.627	-2.490
6	Frg 6-Frg.7	-34.592	41.853	-20.039	10.995	7.261	-9.044	-1.783
7	Frg 7-Frg.8	-50.849	53.226	-26.831	13.218	2.379	-13.614	-11.235
8	Frg 8-Frg.9	-48.636	48.113	-24.561	11.891	-0.523	-12.670	-13.193
9	Frg 9-Frg.10	-48.044	47.162	-24.129	11.624	-0.882	-12.504	-13.386
10	Frg 10-Frg.11	-47.023	46.194	-23.468	11.356	-0.828	-12.112	-12.940
11	Frg 11-Frg.12	-46.991	46.153	-23.497	11.386	-0.838	-12.111	-12.949
12	Frg 12-Frg.13	-49.110	49.049	-24.962	12.102	-0.060	-12.860	-12.920
13	Frg 13-Frg.14	-34.454	41.435	-19.793	10.879	6.981	-8.914	-1.934
14	Frg 14-Frg.15	-49.654	49.370	-25.286	12.246	-0.284	-13.041	-13.325
15	Frg 15-Frg.16	-33.196	39.380	-18.804	10.219	6.185	-8.586	-2.401

Table 4: Inter/intra energies for PMDA-API-II structure from FMO Theory Approach. Interaction energy contributions obtained with using the LPBE0AC functional with the aug-cc-pVDZ basis set. Energies in kcal/mol.

	Fragments	$E_{el}^{(1)}$	$E_{exch}^{(1)}$	$E_{ind}^{(2)}$	$E_{ind-exch}^{(2)}$	$E_{disp}^{(2)}$	$E_{disp-exch}^{(2)}$	$E^{(1)}$	$E^{(2)}$	$E^{(1)}+E^{(2)}$
1	Frg 1-Frg.2	-42.617	46.829	-24.329	12.873	-15.827	3.220	4.212	-24.063	-19.851
2	Frg 2-Frg.3	-47.132	56.594	-30.690	18.038	-17.511	3.907	9.462	-26.256	-16.794
3	Frg 3-Frg.4	-33.038	44.455	-22.096	13.044	-14.005	2.835	11.417	-20.223	-8.806
4	Frg 4-Frg.5	-46.518	51.828	-26.468	13.416	-16.810	3.413	5.310	-26.449	-21.139
5	Frg 5-Frg.6	-32.254	44.049	-23.108	14.628	-13.681	2.953	11.795	-19.208	-7.413
6	Frg 6-Frg.7	-32.170	42.560	-21.308	12.508	-13.693	2.739	10.390	-19.754	-9.364
7	Frg 7-Frg.8	-48.319	59.590	-32.233	19.061	-18.182	4.078	11.272	-27.277	-16.005
8	Frg 8-Frg.9	-45.087	49.535	-25.257	12.732	-16.357	3.290	4.448	-25.592	-21.144
9	Frg 9-Frg.10	-45.556	53.286	-29.100	16.913	-16.873	3.721	7.729	-25.248	-17.519
10	Frg 10-Frg.11	-43.595	47.560	-24.101	12.114	-16.007	3.199	3.966	-24.795	-20.829
11	Frg 11-Frg.12	-43.145	48.004	-24.687	13.050	-16.101	3.287	4.859	-24.452	-19.592
12	Frg 12-Frg.13	-45.159	51.032	-26.255	13.869	-16.646	3.425	5.873	-25.608	-19.735
13	Frg 13-Frg.14	-32.145	42.271	-21.157	12.461	-13.582	2.724	10.127	-19.554	-9.427
14	Frg 14-Frg.15	-46.060	50.913	-26.050	13.164	-16.650	3.363	4.853	-26.173	-21.320
15	Frg 15-Frg.16	-30.917	40.121	-20.037	11.671	-13.068	2.595	9.2033	-18.839	-9.635

V. Conclusions

We have isolated minimum energy structures for two PMDA API ortho xylene polymer orientations by means of simulated annealing molecular dynamics simulations in both vacuum and a high dielectric medium. We then use the FMO functional to study the total interaction energies, which are later broken into the electrostatic, dispersion, and induction contributions. The stacking and inter/intra-bonded interactions were investigated and reported along with a comparison of HF and the DFT functional.

We learn that the folding in vacuum is primarily driven by Coulombic interactions among the charged species, naming ionic+rings and its counter ions. In fact, the presence of countering leads to an effective attraction between the stacking cyclic rings that accounts for roughly half the folding energy. In conclusion, we learn FMO-PIEDA interaction energies scale well with high level *ab intio* calculations. Therefore, the functionalization of these materials chemical, conductive and mechanical properties may answer to the problem of additive manufacturing polymers.

VI. Acknowledgements

This work was supported by the Louisiana Board of Regents Summer Assisted Research (SAR) Faculty Fellowship. This work was conducted at the Marshall Space Flight Center. Acknowledgements are owed to Dr. Enrique Jackson, Dr. Brittany Brown, Dr. Jason Bara of the University of Alabama Tuscaloosa and Dr. Kendall Byler of the University of Alabama Huntsville.

VII. Bibliography

1. Thar J, Brehm M, Seitsonen AP, Kirchner B *J Phys Chem B* **2009** *113(46)*, 15129-15132.
2. Cremer T, Kolbeck C, Lovelock KRJ, Paape N, Wo¨lfel R, Schulz PS, Wasserscheid P, Weber H, Thar J, Kirchner B, Maier F, Steinru¨ck H-P *Chem Eur J* **2010** *16(30)*, 9018-9033.
3. Maginn EJ *J Phys Condens Matter* **2009** *21(37)*, 373101-373118.
4. Watanabe H, Doi H, Saito S, Matsugami M, Fujii K, Kanzaki R, Kameda Y, Umebayashi Y *J Mol Liq* **2016** *217*, 35-42.
5. Izgorodina EI, Seeger ZL, Scarborough DLA, Tan SYS *Chem Rev*. **2017** *117(10)*, 6696-6754.
6. Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation-Theory Approach to Intermolecular Potential-Energy Surfaces of Van-Der-Waals Complexes, *Chem. Rev.* **1994**, *94*, 1887-930.
7. Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular Potentials Based on Symmetry-Adapted Perturbation Theory with Dispersion Energies from Time-Dependent Density-Functional Calculations, *J. Chem. Phys.* **2005**, *123*.
8. Hesselmann, A.; Jansen, G. The Helium Dimer Potential from a Combined Density Functional Theory and Symmetry-Adapted Perturbation Theory Approach Using an Exact Exchange-Correlation Potential, *PCCP* **2003**, *5*, 5010-14.
9. Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M., *Chem. Phys. Lett*. **1999** *313*,701-706.
10. Simeon,T., Ratner, M.A., Schatz G.C. *J. Phys. Chem. A*, **2013** *117 (33)*, 7918-7927.
11. Szczepaniak M and Moc J. *J Phys Chem. A* **2014** *118(36)*, 7925-7938.
12. Bazargan, G. and Sohlberg, K. *Int. Rev. Phys. Chem*. **2018** *37(1)*, 1-82.
13. Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; Iwasawa, Y.; Nakata, K.; Chuman, H.; Nakano, T. *J Chem. Inf. Model*, **2006** *46*, 221-230.
14. Fedorov, D.G.; Ishida, T.; Uebayasi, M.; Kitaura, K. *J. Phys. Chem. A* **2007** *111*, 2722-2732.
15. Amari, S.; Aizawa, M.; Zhang, J.; Fukuzawa, K.; Mochizuki, Y.; Iwasawa, Y.; Nakata, K.; Chuman, H.; Nakano, T. *J Chem. Inf. Model*, **2006** *46*, 221-230.
16. Nemoto,T.; Fedorov, D.G.; Uebayasi,M.; Kanazawa,K.; Kitaura,K.; Komeiji, Y. *Comput. Biol. Chem*. **2005** *29*, 434-439.
17. Sugiki, S.-I.; Matsuoka, M.; Usuki, R.; Sengoku,Y.; Kurita,N.; Sekino, H.; Tanaka, S. *J. Theor. Comput. Chem.* **2005** *4*, 183-195.
18. Fedorov, D. G. and Kitaura, K. *J. Comput. Chem*. **2007** *28*, 222-237.
19. Simeon, T.M. NASA Technical Reports **2018**.

20. Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M., *Chem. Phys. Lett*. **1999** *313*,701-706.
21. Simeon,T., Ratner, M.A., Schatz G.C. *J. Phys. Chem. A*, **2013** *117 (33)*, 7918-7927.

Reliable Expandable Satellite Testbed (REST) System Development and Implementation

Murphy C. Stratton[1]

The University of Illinois at Urbana-Champaign, IL, 61801

This paper reports on the implementation and early development of the REST system in the Small Projects Rapid Integration and Testing Environment (SPRITE) Laboratory. This paper discusses the way this system can be used in the future to enhance the reliability and success of missions. The installation, development, problems, and future work of the REST framework are detailed in this report. The Advanced Real Time Environment for Modeling, Integration, and Simulation (ARTEMIS) simulation environment and components will be reviewed as well. Future users of the REST system will benefit by avoiding the errors reported in this paper, which concludes with a discussion about the importance of decreasing spacecraft failure rates and how this system will benefit Marshall Space Flight Center (MSFC).

I. Introduction

The REST framework was developed to reduce the high failure rates of satellite missions. Specifically, CubeSat missions have ~60% failure rate among University projects. REST is used for initial prototyping of the simulation environment, flight software development, Input/ Output (I/O) bus communication development, qualification testing, acceptance testing, and mission planning (Hunt, 2019). The installation of the system in the SPRITE lab will aid in the testing, design, and development of spacecraft hardware and avionics communication by using hardware in the loop testing. Real or simulated hardware models can be used in the simulation to allow for more robust testing of the satellite environment. This testbed incorporates the ARTEMIS and Managed Automation Environment for Simulation, Test and Real Time Operations (MAESTRO) environments which were developed for the NASA Space Launch System (SLS) program (Hunt, 2019)[4].

---

[1] Laboratory Manager of the CubeSat Lab, Department of Aerospace Engineering

The U.S. Army Space and Missile Defense Command (SMDC) needed a way to test CubeSats and SmallSats using a hardware in the loop (HWIL) system. Simulated orbital operational models created to test the hardware and software of these SmallSats provides a rigorous testing platform. Increasing the availability and standard of functional systems level testing and integration would substantially decrease the mission failure rates among CubeSats/SmallSats. This type of HWIL testing allows the user to detect problems early in the spacecraft and software development lifecycle. Checking and testing the orbital operational modes of the satellite prior to launch would significantly increase mission success (Dynamic Concepts, 2016)[7].

The REST system that was delivered to the SPRITE lab required configuration and installation of ARTEMIS to be able to do basic simulations. It was unnecessary to install the MAESTRO user interface at this time, as it is not needed to run a simulation and only provides a visual interface to the operator. This project consists of getting the system up and running and to demonstrate that the system can simulate a piece of hardware, then replace the simulated model with an actual piece of component hardware. For this project, the Star Tracker emulator was chosen as the hardware component that would be tested. The hardware would be programmed using the Core Flight Executive (cFE) software framework and would be successfully integrated into the ARTEMIS simulation.

The result would be to demonstrate a test incorporating each hardware component in a satellite. Prior to the REST system being delivered to NASA MSFC, a demonstration test was performed using the Iodine Satellite (iSAT) hardware. Initial test scripts that ran on the iSAT flight were developed to test the communication with each hardware component and verified the interface with a basic set of commands. The iSAT I/O board was used for I/O communication and power. They performed component tests, where the REST flight computer software model was used to communicate with each piece of hardware one at a time. Lastly, system level tests were performed using the iSAT flight computer hardware with the REST software component models.

The hardware architecture for the REST system consists of several components, including a model node, recorder node, uninterruptible power supply (UPS), power distribution unit (PDU), Ethernet switch, computer monitor, and a keyboard, video, and monitor (KVM) switch. The lab test configuration for the REST system is shown in Figure 1. All components are assembled in a 19” server rack. The model and recorder node components are Concurrent ImaGen Rackmount Real-Time Servers running a RedHawk Linux Real-Time OS, each box also has six peripheral component interconnect express (PCIe) slots. The Concurrent boxes perform the real-time computing necessary to run the ARTEMIS simulation environment.

The models require complex calculations and often the simulations require the computing power from multiple machines at once to meet the required computation time, which is made possible by transferring data between them. The model and recorder nodes also utilize I/O boards to interface with the satellite avionics hardware. The I/O interface allows for serial bus communication, the sending/receiving of analog/digital signals, and the transfer of data between simulation computers for data synchronization and recording. A SeaLevel 7804e RS-

232/422/485 Serial Interface board was used for communication between the star tracker emulator and the REST system.

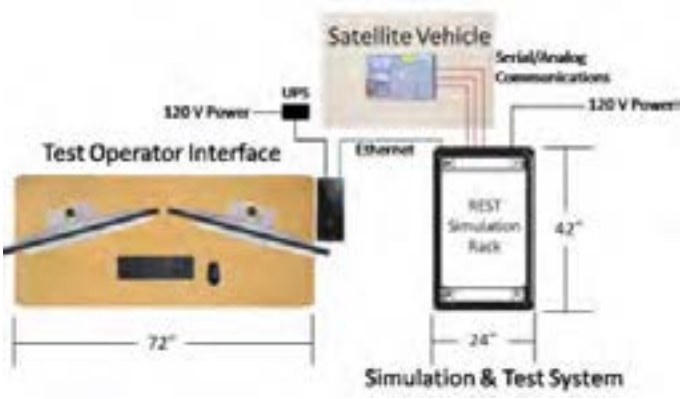


Figure 1: REST Lab Configuration (Hunt, 2019)

II. Software Architecture

A. ARTEMIS Framework

ARTEMIS simulates any physical system that can be modeled and is made of building blocks that are used to model avionics hardware and interface with real communication buses, which means the simulation must run in real time. ARTEMIS operates in the Redhawk Linux environment, meaning first time users should be familiar with Linux and the C language. Simple functions that are called in the model use the shared memory and infiniband buses to send and receive messages between executables (Hunt, 2019)[4]. ARTEMIS was originally created to test the avionics hardware but has since supported a "plug and play' method with each model being a different executable (Hunt, 2019)[4]. Individual executables can be interchanged for actual avionics hardware, and not disrupt the simulation. Due to the data transfer between executables, the synchronization module sets up a 'Blackboard' and writes the output at the end of each frame.

The model architecture in ARTEMIS is comprised of subsystem, component, and core simulation (CoreSim) models. The subsystem models send forces and environmental behavior to the CoreSim model, and they include the engines. The avionics boxes onboard a satellite are simulated using the component models. Some component models include the power distribution unit and the flight computer (Hunt, 2019)[4]. Figure 2 shows the interaction between the different models.

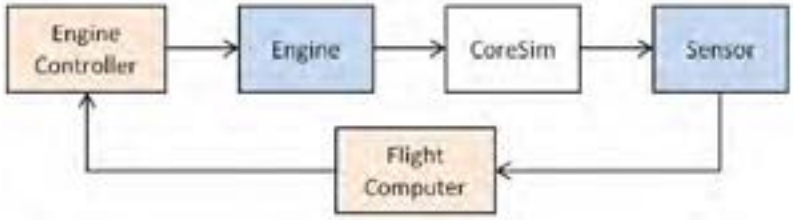


Figure 2: The interaction between the subsystems, the components, and the CoreSim models (Hunt, 2019)

The ARTEMIS framework is able to operate in both real-time and non-real-time environments and is the basis to perform the execution and synchronization for the simulation (Hunt, 2019)[4]. Several modules make up the ARTEMIS framework, they include:

- Simulation Executive
- Synchronization
- Input/ Output (I/O) Layer
- Models
- Data Recorder

The executive module is the basis of the ARTEMIS framework and provides the main interface for each model executable and handles processing tasks. The synchronization module keeps each model executable in sync with the simulation schedule and manages the blackboard shared memory. The I/O layer module allows for the interface between bus components communicating between each other, which represents the communication between real hardware components of a spacecraft. The models module implements each component running in a simulation. The data recorder module records the hardware bus data (Hunt, 2019)[4]. Figure 3 demonstrates the ARTEMIS modules interface.

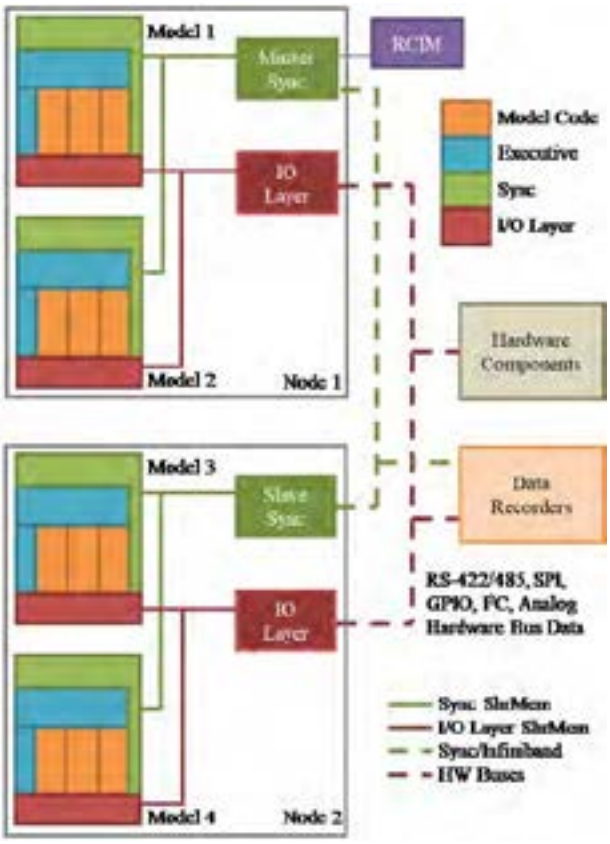


Figure 3: ARTEMIS Module Interface Layout (Hunt, 2019)

ARTEMIS contains a CoreSim model that is utilized by the REST framework. The CoreSim model simulates system dynamics and environmental effects, reacts to forces from effector models, and provides state data to sensor models (Hunt, 2019)[4]. Simulated satellites are represented by the CoreSim and it is made of four main components: vehicle forces, configuration manager, vehicle dynamics, and environment (Hunt, 2019)[4]. Figure 4 shows the CoreSim components and how they interact with each other.

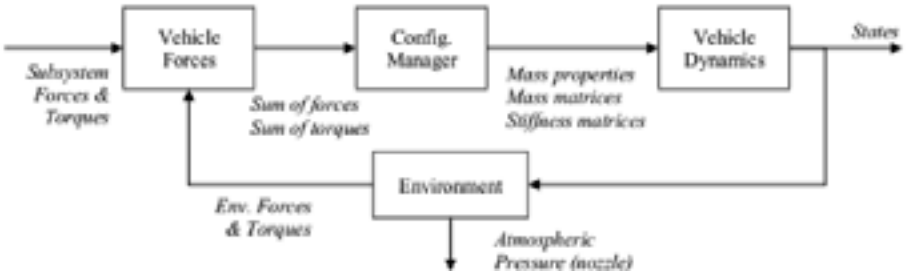


Figure 4: CoreSim Simulation Components Interface (Hunt, 2019)

B. Core Flight System

The core Flight System (cFS) is a flight software environment, created at NASA Goddard, which is both platform and mission independent. Inside of the cFS environment is a reusable core Flight Executive (cFE) with flexible, compliant applications. The ability to reuse existing code decreases the amount of time it takes to create robust flight software. cFS also allows for easier collaboration between groups like universities and organizations. The scalability of cFS is also a plus, as it can be used to program small CubeSat assemblies to larger spacecraft. In the scope of this project, cFS was used as a way to interface with a piece of tangible hardware through the ARTEMIS framework. An application was written in cFS to interface with the star tracker for the Mars Ascent Vehicle (MAV). The cFS code was integrated into the ARTEMIS code, specifically into the flight computer code. This way it can be demonstrated that it is possible to replace a simulated model in ARTEMIS with a piece of real hardware, and run a successful test. The biggest difficulty was integrating cFS into the ARTEMIS code (Medina, 2019)[6].

C. ARTEMIS Beginner's Tutorials

The tutorials were a way to introduce the beginner ARTEMIS user to the system and get them familiar with its basic functions. Tutorials one and four were attempted and they were both run successfully, although not without having to debug numerous compile errors. The ARTEMIS system was installed onto the REST system in the SPRITE lab. The user can connect to the compnode or recorder node on the rack through an Ethernet connection. To login to the to the REST system the user must secure shell (i.e. SSH) into either the recorder node or the computer node. Using the command line in a Linux environment, the user enters the IP address for either the computer or recorder node.

In ARTEMIS, any system can be simulated through models. The operating environment consists of a command line, and the simulation is run from the terminal window. ARTEMIS uses

scripts, which help developers design and test their models. Some important scripts include the 'configure_artemis' and 'look_at' scripts, which are needed when compiling and running a simulation. The 'configure_artemis' script configures the environment for the simulation and sets the environment variables. The 'look_at' script prints the executable list to the screen. Green executables means they were built correctly, while red are not. When simulations are being compiled and run, environment variables are needed and are assigned by the 'configure_artemis' script (Hunt, 2019)[3].

To compile ARTEMIS it is necessary to use makefiles due to the large number of executables. Using 'GNU Make', the user can compile multiple executables with a single command. ARTEMIS uses several makefiles which are used to compile certain types of executables. Certain directories, like the Executive and component directories, have their own makefiles, which helps decrease the number of make commands. However, users do not need to go to each directory with a makefile to run a 'make' command. Both the ARTEMIS and REST frameworks can be compiled from the root/REST/ARTEMIS/Scripts directory using the makefile there. This makefile calls the rest of the makefiles, which allows the user to compile ARTEMIS in a fast, simple way. Figure 5 shows the executable list after configuring ARTEMIS and running a simple 'make' command. The 'Metedata_REST' and 'A_data_recorder' are shown in red because they are needed for running the MAESTRO interface, which was not used in the scope of this project, therefore they were not compiled. There are several 'make' options available to the user, including **make all**, **make clean**, and **make spotless**. Make clean and make spotless refer to the deletion of object files and executables (Hunt, 2019)[3].

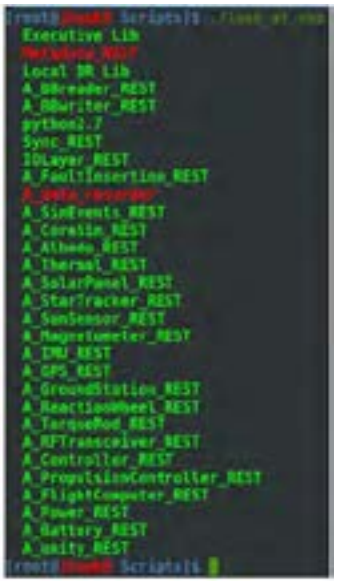


Figure 5: The compile and make commands entered into the terminal and the resulting executable list (Hunt, 2019)

To run an ARTEMIS simulation, the user must navigate to the ARTEMIS/Scripts directory and enter the command into the terminal. This will show the user the Environment variables for both ARTEMIS and Sim, as well as the Blackboard file. The user then enters a 'make' command (depending on the desired output) to compile the simulation. To run the actual simulation the run command must be entered. Run scripts are created to simplify running the simulation by decreasing the number of arguments on the command line by handling the different parameters. The ARTEMIS run script is located in the Scripts directory. An example of a run script is shown in figure 6 (Hunt, 2019)[2].

Figure 6: Run script example (Hunt, 2019)

The **SYNC** parameter refers to '-noX' flag which is important if the user is not running as a root user, for this project the simulations were always run by a root user. The **GROUP** parameter defines groups and enables executables for them. The **ON** and **OFF** parameters enable and disable the model executables for the startARTEMIS command. The **TIMER** and **TICK** parameters apply the timing conditions for the simulation. The **RECORD** parameter creates a directory to store the run-time data from the simulation and will create a 'Default' directory if one is not specified. The **FAULT** parameter refers to the fault insertion that is enabled when the **'-f'** flag is used (Hunt, 2019)[2].

D) Tutorial 1: Spring Mass Damper

The first tutorial went through how to create a new model and wrapper function for a vehicle model, how to read input from an XML file, how to use the data recorder and integrator, and how to run a single model executable. In this tutorial, the user creates a model of a spring mass damper (SMD) system and integrates it into the system. The user learns how to create a standard wrapper function for the new model. Wrapper functions are provided by models and are executed during the simulation phases. There are several simulation phases including **Set Sim Name**, **Init Read**, and **Sync Register.** These phases have important functions such as loading input files for the models and registering the model through the Sync module (Hunt, 2019)[2].

The next part of the tutorial teaches the user about the directory structure for the SMD model. The model files are located under the 'Core Sim' directory, which includes the data, include, object (obj), source (src), and bin folders. This section shows the user how to set up the

blackboard and make data available, how to read the input from an XML file that describes the system properties, and how to use the data recorder and integrator. The last section describes how to run a single model executable using the ARTEMIS framework. This section describes how to edit the configure scripts, makefiles, and XML data files. After running the command line inputs the model executable list will show whether or not the '**A_SMD_REST'** executable model was successfully built depending on if the model title is red or green. The majority of data files in ARTEMIS use the XML format and the tutorial describes how a few of them need to be edited before running the simulation (Hunt, 2019)[2].

The executable configuration file is extremely important and editing this file is the last step before running any simulation. This file is in comma separated value (CSV) format and is used by the startARTEMIS script to know which executables are to be used. The names of the models are listed as the top row in the file and the user must manually add the SMD model as a new column in the list.

If the user is running the command as 'root', then it is not necessary to include the 'sudo' command. If the simulation runs successfully, the user will be able to check their results using the Bresenham plotting tools (see ARTEMIS Developer's Guide). The user may experience compile errors, including formatting/ syntax errors and file path errors in the makefiles. Errors found in this tutorial were reported to the original author and should be updated soon.

E) Tutorial 4: Using the I/O Layer for Digital Communication

This tutorial introduces the user to the input/output (I/O) layer for asynchronous communications. The I/O layer is an important component in the ARTEMIS system, as it allows the user to interface with both real and simulated hardware. Non-real time processes are performed by the I/O layer and receives commands used by the models in the IOLayerLibrary. A simulated asynchronous bus allows the user to communicate between the two model executables created for this tutorial, which are the sun sensor and the flight computer. The I/O layer supports several interface devices including RS-422, RS-485, and RS-232 serial devices (Hunt, 2019)[2]. Some of the basic I/O layer functions include the:

- *vIoAsyncInitWrapper*: this wrapper function establishes the I/O layer and calls the IOLayerInit function.
- *IOLayerInit:* this function communicates with the I/O layer to enable the device that is called by the model.
- *IOLayerRead:* this function returns the most recent data stored in the memory buffer.
- *IOLayerWrite:* this function moves the user buffer and stores it in the shared memory buffer and is then processed by the I/O layer.
- *IOLayerClose:* this function disables the connection with the I/O device.

The tutorial then prompts the user to create the sun sensor and flight computer component models. Directories are created for both models under the Models/Component directory. Each model directory contains the object, source, and include files. The flight computer model

contains many functions including vInitRead, vSyncRegsiter, and vWriteIO. These functions have very important tasks: creating the standard wrapper for the model, registering with the Blackboard for memory sharing between models, and sending messages allowed by the I/O layer (Hunt, 2019)[2].

The user then creates the shared object, which is the shared library used by component models. According to the author of the *ARTEMIS Developer's Guide*, the "component model shared library implements vendor specific communication protocols, process timing, signal noise, signal accuracy, and fault behavior". The user can then create the makefile for the Sun sensor which creates the shared library. The user can then enter commands into the terminal in order to build the model component executables.

The tutorial then prompts the user to create the I/O layer configuration XML file. This file configures the hardware interfaces needed for the simulation and includes the device name with its necessary buses. The user must also edit the existing REST configuration XML file by adding the flight computer and sun sensor components and parameters. The user must also add these components/parameters to other XML files such as the SWDelivery_Orig.xml and DataRecord.xml files. Once the user has made the necessary modifications to the XML files and the exe_config.csv file, they can then enter commands to compile and run the ARTEMIS simulation (Hunt, 2019)[2].

When trying to initially run this model, several compile errors occurred. Many of them were simple, overlooked syntax errors caused by the user copying and pasting code from the tutorial into the ARTEMIS scripts. This caused numerous formatting errors, so future users should be very cautious of this. Users trying to debug their runtime errors should use the Meld file comparison tool in Linux to see what changes were made to their original code. Other errors occurred, including build errors from the A_data_recorder and metadata_REST executables. A lot of time was spent trying to debug these errors, but it was for naught as it was realized that the data recorder and metadata are not being used for this simulation, as they are used by the MAESTRO interface. Future users should keep this in mind and not panic when they see these two executables show up as red in the model list as they actually do not exist. Another error encountered was the **"what(): Failed to open shared memory: No such file or directory."** error. According to the author of the *ARTEMIS Developer's Guide*, this error could actually be ignored and said that it was caused by the shared object file not being found. The author also stated that future users should check that it is named correctly in the XML file, it is located in the correct place, and is called correctly when trying to open the file. Users should also make sure that anything listed under the flight computer matches what is configured and to make the appropriate changes to the exe_config.csv file. Once these tutorials are successfully completed, the user should feel comfortable in experimenting with the REST system.

III. **Experimentation**

The main purpose of this project was to have the REST system up and running in the SPRITE lab as soon as possible. The ES53 Avionics and Software Ground Systems Test Branch constructed the rack with OS, and DCI assisted with the ARTEMIS installation. DCI also

produced the main bulk of documentation for the REST/ARTEMIS systems. They suggested working through tutorials one and four in the ARTEMIS Developers Guide in order get a better understanding of the system. Once the tutorials were successfully debugged and running correctly, it was time to begin trying to talk to the system with actual hardware. The delivered system came without any interface boards, therefore, the system could not interface with the star tracker hardware. A Sealevel 7804e I/O board was acquired along with an interface cable with connectors corresponding to the ports on the board. The board was installed and verified to be recognized by the REST system.

The Sealevel 7804e I/O serial board has eight (8) ports that can be configured for RS-232, RS-422, or RS-485 serial communication. The goal was to get the star tracker emulator hardware to communicate with the REST system through the I/O board, using the RS-485 2-wire communication protocol. The RS-485 protocol allows for multiple listening and commanding devices on the same lines as long as the devices have unique unit addresses. It is important to enable and disable the transmit functionality at the correct time due to the two (2) lines for transmit and receive. The Sealevel board came with the eight (8) ports already configured to the different serial protocols.

A loopback test was performed using port 2 on the I/O board which was configured for the RS-232 protocol. This first test was performed to make sure the ports on the I/O board could be opened, and send and receive simple telemetry data. Code written for the star tracker in cFS was installed onto REST compnode 1. The idea was to run the loopback test using the modified cFS star tracker code on the REST system. The python modules “libzmq” and “PyQt4” installed on REST were different versions than the ones installed to run the ground system in cFS, causing undefined symbol errors. The ground system tool allows the user to send commands to their app in cFS using a graphical user interface (GUI). Updating the zmq module on REST fixed one of the errors, but then another error occurred due to the ‘SIP.so’ file installed on the REST system being built for a different version of python than the one on the REST system. The solution was to use virtual environments to install the necessary versions of the python modules on the REST system and use pointers in the cFS code point to the correct libraries. The modules could not be installed globally because it would adversely affect the MAESTRO configuration. Changing the baud rate for the RS-232 loopback test also fixed some errors and resulted in a successful loopback test.

The same test was then performed using a special cable that connects the star tracker to a 5V power supply and to the REST system through the I/O board. No telemetry was seen from the star tracker emulator meaning it could have been working incorrectly. The star tracker emulator does not have any lights indicating it was powered on however, it drew consistent current which meant it was probably on. This meant the problem could be with the RS-485 ports on the board or the code. A simple program called ‘serial.c’ was written to test the serial connection between the REST rack and the star tracker emulator. No errors occurred when opening the port and the write instructions returned without errors, but there was still no telemetry data returned from the star tracker emulator. The response generated from the emulator is a read command that is 5 bytes (0xEC, Address MSB, Address LSB, Len MSB, Len LSB), so 60 bytes total are read from

the emulator's telemetry buffer. The function writes the star tracker read command to the serial port, then reads the serial port file descriptor. The next step tested was whether the emulator would send data while not being connected through the I/O board.

The next test performed was to see if the star tracker emulator was working correctly. The emulator used was different than the actual star tracker flight model in that it did not contain the optics or the detector. The commands and interfaces used were still the same as the flight model, so the same interface control document (ICD) could be used. To test the star tracker emulator, a simple USB to 485 serial cable and some MATLAB scripts provided by the ES35 Avionics Subsystem Branch – GNC Hardware and Systems were used. To get the star tracker emulator to successfully send telemetry data. After sending a read command to the star tracker emulator, accurate voltages and telemetry data were received. This meant that the emulator was working correctly, meaning the fault with not talking to the emulator through the I/O board resided with the code.

The next step was to talk to the star tracker using the cFS code written by the ES52 Software Development Branch. Here, the cFS code was being executed on the REST system, but not using the ARTEMIS framework. Data was being transmitted from the emulator, but it was not in a readable form. This was because data from the star tracker was in big endian byte order, while the REST system used little endian byte order. The fix to this problem came from swapping the byte order in the cFS code. Data from the star tracker was also being cut off due to the line length being set to 172 characters and was fixed by setting the line length to 512. Another issue was the data was only printing once, but this was fixed by putting the emulator into idle mode before sending the test telemetry command. The next step was publishing the data into another app that received the published telemetry.

Once the cFS code was running successfully on the REST system, it was time to try running the code inside of the ARTEMIS framework. The power supplied to the emulator was correct, drawing ~150 mA at 5 Volts. The cable was good and was tested for continuity issues prior to being used. The code was also good, since it was running successfully on the REST system. This meant that the error was the RS-485 ports on the I/O board. After contacting Sealevel, it was discovered that since the RS-485 two wire protocol is more complex, it was necessary to enable the transmit for those specific ports in the serial.c script. The I/O board uses 16C954 universal asynchronous receive/transmit (UART) for its ports. There is an 'ioctl' command that enables the auto RS-485 mode, which is not supported by the serial driver. These ports have an auto RS-485 mode that use the data terminal ready (DTR) pin to control the RS-485 line in half duplex. This is enabled by writing to multiple registers of the UART. Code was taken from an ARTEMIS script that does this and put into the 'serial.c' application. Fixing this allowed the I/O board to interface with the star tracker, but not using cFS.

When trying to use the 'ioctl' to set the RS-485, a system error occurred due to the serial drivers not supporting that command. This meant registers would need to be manipulated. Another error in trying to talk to the emulator through cFS on the REST system was that the code from cFS was in 32-bit, while the REST system uses 64-bit. REST's 64-bit could not be changed, so the next step was to try and run cFS as 64-bit. However, an error occurred when

trying to compile cFS as 64-bit that referred to the elf2tbl utility needing 32-bit elf files to run. A possible solution was to build the elf tables as 64-bit using a branch of elf2cfetbl from GitHub. This was successful but caused a segmentation fault in the cFS house-keeping application. This is a problem if the user needs to use the housekeeping application. This was the furthest point reached in on this project.

IV. Future Work and Conclusion

Future work on this project will be getting the cFS code to run inside of the ARTEMIS framework successfully. Once this is accomplished, more simulated models can be replaced by actual hardware. Eventually, full tests can be performed using a complete spacecraft bus. This would allow CubeSat projects from outside universities and organizations to test their systems using rigorous standards. The long-term effects would be outstanding, causing more CubeSat projects to be successful. This would enhance the scientific value of these projects and provide invaluable experience to young engineers.

The completion of a REST system at the MSFC would open the door to many future partnerships and collaborations. Universities would greatly benefit from using the REST system because they would have a way to test their flight software and hardware early in the mission development cycle while also utilizing the expertise provided by NASA MSFC. Targeting system errors early gives the mission a much better chance of success. This is crucial for increasing the success rate of CubeSat missions among universities and organizations. NASA MSFC would also benefit highly from this by gaining experience in testing the system with different types of simulated and real hardware.

The value of the REST system to on-going and future projects completely justifies any necessary work being put into this project. The industry greatly needs a system that can test out the kinks in small CubeSat assemblies as well as much larger, more complex systems. By next summer, the SPRITE lab should have its own complete REST system which will be available to outside organizations. Users who work on this project in the future should heed the error warnings in this report in order to cut down on the amount of time spent identifying and resolving such issues. Users should also read the ARTEMIS Developers guide, ARTEMIS Installation Guide, ARTEMIS User's Guide, and REST User's in their entirety and work through the tutorials before attempting to modify, design, or test anything on the REST system.

V. Acknowledgments

A special thanks to the DCI, ES52, ES53, and ES35 teams for their assistance and great contributions to this project. Also, thanks to the Education Office of MSFC for the opportunity to participate in the 2019 MSFC Summer Faculty Fellow program.

VI. References

[1]Hunt, M. (2019). *ARTEMIS Installation Guide* (1st ed., Tech.). Huntsville, AL: Dynamic Concepts.

[2]Matras, A., Tobbe, P., Tolbert, D., Ipnar, J., Smith, J., Molina, J., Hunt, M. (2019). *ARTEMIS Developer's Guide* (1st ed., Tech.). Huntsville, AL: Dynamic Concepts.

[3]Hunt, M. A. (2019). *ARTEMIS User's Guide* (1st ed., Tech.). Huntsville, AL: Dynamic Concepts.

[4]Williams, P., Young, J., Pearce, A., & Hunt, M. A. (2019). *REST User's Guide* (1st ed., Tech.). Huntsville, AL: Dynamic Concepts.

[5]*SeaLevel UltraCOMM 8.PCIe* [User Manual]. (2017, February).

[6]Medina, B. (n.d.). Core Flight Software System (CFS) [Digital image]. Retrieved July/August, 2019.

[7]Dynamic Concepts, Inc. (2016). *RELIABLE EXPANDABLE SATELLITE TESTBED (REST) SYSTEM REQUIREMENTS DOCUMENT (SRD)* (1st ed., p. 6) (USA, United States Army Space and Missile Defense Command). Huntsville, AL. Retrieved July 31, 2019.

Baseline "Scout" Lander Mission Analysis for in-Situ Lunar Lava Tube Exploration

Stephen A. Whitmore,* Thomas P. Lampton,† and Jonathan Jones‡
Utah State University 1, Logan, UT, 84321

Low-budget lunar "scout" mission concepts for in-situ exploration of lava tube "skylight" surface openings are proposed. The primary objective is inexpensive and repeatable missions to investigate potential lava tube openings that may provide structures that are suitable for development as human habitats. If such missions were enabled, the collected in-situ database can be used to scrub the list of potential skylight targets, substantially enhancing success probabilities of more comprehensive and costly follow-on exploration missions. This study investigates capabilities of 4 emerging non-defense, commercial launch vendors for delivery of the in-situ landing payloads to lunar orbit. A typical landing system would weigh approximately 10kg, and land only the sensors and optical systems necessary to "vet" the surface feature for future human habitat development. Because the "scout" missions allow cheap and repeatable access to promising surface features, targets can be repeatedly investigated. All 4 vendors examined possess the capability of delivering sufficient mass to lunar orbit to allow multiple 10 kg spacecraft to access the lunar surface. The option of having redundant landers available significantly reduces mission risk; and should the first landing attempt be successful, the second lander provides the option to explore another surface target that lies along the orbital track. Starting from a nominal 200 km altitude lunar orbit, this study concludes that approximately 50% of the on-orbit mass is deliverable to the lunar surface; however, it appears that solid propulsion systems provide insufficient impulse precision for reliable mission outcomes. Use of a throttleable or restartable final descent stage, coupled with closed-loop energy management, is likely mandatory for surface payloads approximating 10 kg mass.

Nomenclature

Symbols

A^* = Nozzle Throat Area, cm^2
a = Orbit Semi Major Axis, km
C_F = Thrust Coefficient
E = Actual Spacecraft Total Energy Level, MJ
E_{opt} = Optimal Spacecraft Total Energy Level, MJ
ΔE = "Dead-Band" Around Optimal Energy Level , MJ
e = Orbit Eccentricity
h = Mean Altitude, km
F_{axial} = Axial Thrust Component, N

* Professor, Mechanical and Aerospace Engineering, Utah State University.
† Technical Assistant for Solid Propulsion, ER-62, NASA Marshall Space Flight Center
‡ Lead Artemis Landing Systems, ER-62, NASA Marshall Space Flight Center

F_{thrust} = Instantaneous Spacecraft Thrust Level, *N*

g_o = Standard Earth Acceleration of gravity, *9.8067 m/sec²*

I_{sp} = Specific Impulse, *s*

I_{tot} = Total Required Impulse for Lunar Capture Stage, *s*

L_f = Stage Propellant Load Mass Fraction, $M_p/M_{initial}$

L^* = Ratio of Combustor Chamber Volume to Nozzle Throat Area, *m*

M_{GTO} = Payload Mass Delivered to Geostationary Transfer Orbit (*GTO), kg*

$M_{initial}$ = Initial Stage or Spacecraft Mass, *kg*

M_{LEO} = Payload Mass Delivered to Low Earth Orbit (*LEO), kg*

$M_{P,LLO}$ = Propellant Mass Burned, Lunar Transfer Orbit (*LTO*) to Low Lunar Orbit (*LLO*), *kg*

M_{p1} = Optimal Propellant Mass Burned without Loiter, *kg*

M_{p2} = Propellant Mass Burned with Loiter, *kg*

$M_{stage,\ dry}$ = Residual Mass of Expended Kick Stage, *kg*

$M_{T,LLO}$ = Total Mass Delivered to *LLO*, *kg*

$M_{u,LTO}$ = Usable Mass Delivered to *LTO*, *kg*

$M_{u,LLO}$ = Usable Mass Delivered to LTO, *kg*

m = Instantaneous Mass of Lander Spacecraft, *kg*

O/F = Oxidizer to Fuel Ratio

$R_{GTO\,a}$ = *GTO* Radius at Apogee, *km*

$R_{LEO\,a}$ = *LEO* Radius at Apogee, *km*

$R_{LEO\,p}$ = *LEO* Radius at Perigee, *km*

$R_{\mu L}$ = Mean Radius of Lunar Orbit Relative to Earth, *km*

r = Radial Distance, *km*

r_{moon} = Mean Lunar Radius Based on Volume, *1736.60 km*

$V_{GTO\,p}$ = Perigee Velocity of *GTO* Orbit, *km/s*

$V_{LEO\,p}$ = Perigee Velocity of *LEO* Orbit, *km/s*

$V_{LCO\,p}$ = Perigee Velocity of *LCO* Orbit, *km/s*

$V_{LTO\,p}$ = Perigee Velocity of *LTO* Orbit, *km/s*

V_r = Vertical Velocity Component, *km/sec*

V_v = Horizontal Velocity Component, *km/sec*

$\Delta\Omega$ = Change in Right Ascension Angle, *deg.*

ΔV = Velocity Impulse, *m/s*

ΔV_{GTO} = Velocity Impulse Required for *GTO* Insertion, *m/s*

ΔV_{LLO} = Velocity Impulse Required for *LLO*, *m/s*

ΔV_{LTO} = Velocity Impulse Required for *LTO* Insertion, *m/s*

ε_{LLO} = Normalized Final *LLO* Comparison Parameter

γ = Flight Path Angle, $tan^{-1}(V_r/V_v)$, *deg.*

$\mu_{\oplus}$ = Earth's Planetary Gravitational Constant, *3.9860044 x 10⁵ km²/s²*

μ_{moon} = Moon's Planetary Gravitational Constant, *4.903226 x 13^5 km^2/s^2*
ν = True Anomaly Angle, *deg*
Ω = Right Ascension of the Ascending Node, *deg.*
ω = Argument of Perigee, *deg.*

Acronyms

ABS = Acrylonitrile Butadiene Styrene
ECI = Earth-Centered Inertial Coordinate System
GTO = Geostationary Transfer Orbit
HTPB = Hydroxyl-Terminated Poly Butadiene
INS = Inertial Navigation System
IPR = Injector Pressure Ratio
KSC = NASA Kennedy Space Center Launch Complex
LCROSS = Lunar Crater Observation and Sensing Satellite
LEO = Low Earth Orbit
LCI = Lunar Centered Inertial Coordinate System
LCO = Lunar Capture Orbit
LCS = Lunar Capture Stage
LLO = Low Lunar Orbit
LQT = Linear Quadratic Tracking
LRO = Lunar Reconnaissance Orbiter
LTO = Lunar Transfer Orbit
PID = Proportional Integral derivative
NGIS = Northrop Grumman Innovation Systems
RAAN = Right Ascension of the Ascending Node
RMS = Root-Mean Square
SLS = Space Launch System
ScMC = Scout Mission Concept
SSO = Sun Synchronous Orbit
STT = Surface Transfer Trajectory
TRL = Technology Readiness Level
WLA5 = Australian Woomera, Area 5 Launch Complex
WFF = NASA Wallops Flight Facility Launch Complex

I. Introduction

Since its arrival in lunar orbit in 2009, the Lunar Reconnaissance orbiter (LRO) has imaged over 200 lunar surface features that exhibit the characteristics of "skylights" into extensive subsurface lava tubes and caverns.[1] Based on properties of similar terrestrial features, some of these caverns may be structurally stable and extend for miles underground. In fact, the Indian

Chandrayaan-1[2] orbiter has imaged a lunar surface rille near the lunar equator with an intact lava tube section that extends for more than 2 km with a width of 350 meters.

Such extensive underground caverns have the potential to serve as starting points for human habitats; providing natural protection against solar flares, cosmic rays, and micrometeorite impacts. The walls of these "skylight" structures can also provide supports for airtight structures where a fabric dome seals off the surface inlet. Finally, some of these underground features are now believed to contain ancient water frozen as ice. In fact, multiple lunar probes including Clementine[3] and the impact sensor which accompanied the LRO orbiter during the *Lunar Crater Observation and Sensing Satellite* (LCROSS) mission,[4] have identified substantial amounts subsurface ice near the moon's south pole.

A. On the Need for in-Situ "Scout" Data

Several authors, including Whittaker,[5] Ximenes et al,[6] and Kerber,[7] have proposed large-scale missions to explore these underground caverns. Typically, these sub-surface exploration mission concepts are highly complex, and require that substantial mass and technical support infrastructure be delivered to the lunar surface. Such undertakings are bound to be very expensive. Although many potential "skylight" features have been identified, it must be noted that many of these features observed from orbit may, in fact, be the result of moraines at the tail end of a volcanic surface flow rather than volcanic skylights. As such, this lack of definitive knowledge with regard to surface feature morphology presents considerable mission risk associated with undertaking complex and costly sub-surface exploration missions.

During planning for such missions, a database of preliminary in-situ surface "scout" data would be incredibly valuable. Because the "scout" mission would allow cheap and repeatable access to promising surface features, multiple targets could be investigated and the resulting in-situ data used in conjunction with other on-orbit images to winnow down the list of potential targets to a significantly smaller, but more promising set. At the end of this process only the most vetted and verified sites would remain. This "winnowing" process would substantially enhance the potential for mission success associated with comprehensive follow-on exploration missions.

B. Low-Cost "Skylight" Surface Scout Missions

This proposal will investigate low-budget mission concepts that will allow in-situ investigation of candidate skylight surface features using small spacecraft that leverage a variety of available option to reach lunar orbit. Potential lunar-delivery options will be described later in the next section. A typical lander proposed for this scout-mission would have and on-orbit mass of approximately 10 kg, and would deliver only necessary sensor options including a spotlight, wide-view camera, and magnetometer to look for the presence of valuable metallic ores that may be minable to set up a permanent habitat. Some minimal computational power, telemetry, and accelerometers for guidance and control filters, must also be included on the landing vehicle.

Images provided by the previously described lunar orbiting spacecraft have identified that a typical "skylight" opening is quite small, with diameters ranging from slightly greater than 5 meters to less than 900 meters.[8] Large skylights are indicative of partial collapse of the subsurface features, and such unstable tubes would not be ideal for habitat development. Thus, to be useful the scout vehicle must be able to "hit" a very small surface opening with a rather unsophisticated small lander. This lander must reach the target site without the aid of surface landing beacons and expensive/heavy onboard Inertial Navigation Systems (INS).

II.Analysis of Low-Cost Launch Options

A variety of launch options exist for reaching lunar orbit with the lander payloads. Although "rideshare" options using NASA's Space Launch System (*SLS*) remain open, the low rate of proposed launches makes SLS rideshare a non-attractive option for the types of in-situ missions described in the previous section. Instead, this study will investigate the possibility of using non-defense, commercial launch services. The following sub-section investigates this option in greater detail.

C. Non-defense Commercial Launch Options and System Capabilities

Over the past half-decade, a variety of non-defense, commercial vendors specifically targeting the nano-launch market have emerged. These vendors include Rocket Labs,[9] Vector Launch,[10] Relativity Space,[11] and Gilmour Space Technologies.[12] Each of these companies has made significant progress since 2018, and all are expected to launch low earth orbit (*LEO*) payloads within the next 3-4 years. Rocket Labs has successfully orbit 6 of the 7 payloads that it has attempted using its *Electron* rocket. Recently Rocket Labs teamed with NASA to deliver a small spacecraft payload to low earth orbit (*LEO*). [13] DARPA has selected the Vector as a finalist for its SmallSat Launch competition. Relativity Space has multiple 2021 launch contracts for its *Terran-1* rocket. Recently, Australian venture capital firm Blackbird Ventures has teamed up with the Silicon Valley Mike-Cannon Brookes firm to fund development of Gilmour's launch vehicle.[14] The *ERIS*-4 launch vehicle is expected to deliver it first *LEO* payloads near the end of 2021.

Table 1. Low Earth Orbit (*LEO*) Capabilities of 4 Emerging Commercial Space-Launch Operations

Launch Provider, Vehicle	***LEO* Payload**	***LEO* Orbit Parameters**	**Launch Cost**	**Launch Site**	**Maturity/Company Notes**
Rocket Labs, Electron	221 kg	180 x 700 km @ 37.95º Inclination	$6M, $27,000/kg	NASA Wallops (Va. USA)	6/7 Successful *LEO* Launches
Vector Launch, Vector-H	263 kg	200 x 700 km @ 37.95º Inclination	$5-7M, 19,000.kg-27,000kg	NASA Wallops (Va. USA)	DARPA Launch Prize Finalist
Relativity Space, Terran-1	1000 kg	200 x 700 km @ 28.5° Inclination	$10M, $10,000.kg	NASA KSC (Fl. USA)	Multiple Launch Contracts for 2021.
Gilmour Space, ERIS-5	450 kg	200 x 500 km @ 31º Inclination	$10 - $15.2M, $22,000-$34,000/kg[§]	Woomera LA5, (S.A. Australia)	SubOrbital demo Launch from LA5, Orbital Expected 2021.

Table 1 lists the launch capabilities and associated payload costs for each of these vendors, and notes the current status of the primary launch technologies. Clearly, there are other potential commercial options; however, these 4 launch vendors are the most highly developed for the nano-launch market, and are considered as typical of additional options that will emerge over the next 5 years. Currently, these vendors are targeting the low-earth orbit small-spacecraft and CubeSat

[§] Cost Listed in Australian Dollars.

market; however, each has set sights on heavier lift-options with *GTO* (Geostationary-Transfer Orbit) capabilities with the proposed systems.

Table 2. Geostationary-Transfer Orbit (*GTO*) Payload Capability Extrapolated from the Data of Table 1.

Launch Provider, Vehicle	***LEO* Payload**	***LEO* to *GTO* Burn** ΔV	**Required Change in Orbital Energy**	**Extrapolated *GTO* Payload**	**Consumed Propellant for *GTO***	***GTO* Launch Cost**
Rocket Labs, Electron	221 kg	2.315 km/sec	21.17 MJ/kg	97.9 kg	123.1 kg	\$62,300/kg
Vector Launch, Vector-H	263 kg	2.333 km/sec	21.12 MJ/kg	115.8 kg	147.2 kg	\$43,200/kg-\$60,450/kg
Relativity Space, Terran-1	1000 kg	2.333 km/sec	21.12 MJ/kg	440.3 kg	559.7 kg	\$22,700/kg
Gilmour Space, ERIS-5	450 kg	2.388 km/sec	21.52 MJ/kg	194.3 kg	255.7 kg	\$51,500/kg-\$78,200/kg

Table 2 presents the results of calculations performed using the uses data of Table 1 to approximate the total payload that each of the above systems can deliver to *GTO* orbit. Presented are the required ΔV, assuming orbit transfer at apogee, the required change in energy per unit mass, and the delivered *GTO* payload. The consumed propellant listed in column 7 of Table 2 is calculated as the difference between the *LEO* and *GTO* delivered payloads. Although the analysis presented by Table 2 calculates the delivered *GTO* mass assuming a *LEO*-to-*GTO* burn, it is likely that the actual trajectory will be such that the launch vehicle burns directly into the *GTO* orbit from the launch conditions. Thus, the calculated *GTO* mass estimates are conservative, and the delivered *GTO* payloads are likely marginally higher than the presented values.

The calculations of Table 2 were performed assuming a *GTO* Apogee radius $R_{GTO\,a}$ of 42,164.2 km (altitude 35,793.2) km, and an upper stage specific impulse (I_{sp}) of 290 sec. The *GTO* transfer burn occurs at *LEO* orbit perigee, and the associated velocity impulse is the difference between *GTO* and *LEO* orbit perigee velocities. As derived from the Vis Viva equation,[15]

$$\Delta V_{GTO} = V_{GTO_p} - V_{LEO_p} = \sqrt{\frac{2\cdot\mu_{\oplus}}{R_{LEO_p}}}\cdot\left(\sqrt{\frac{R_{GTO_a}}{R_{LEO_p}+R_{GTO_a}}} - \sqrt{\frac{R_{LEO_a}}{R_{LEO_p}+R_{LEO_a}}}\right). \quad (1)$$

The end-of burn payload mass M_{GTO} is calculated from the required ΔV_{GTO} and Vendors' stated LEO payload mass M_{LEO},

$$M_{GTO} = \frac{M_{LEO}}{e^{\left(\frac{\Delta V_{GTO}}{g_0 \cdot I_{sp}}\right)}}. \quad (2)$$

D. The Lunar Orbit Around the Earth

Figure 1 shows the relative geometry of the earth's rotational axis, the plane of the ecliptic, the lunar orbit, and the Moon's axis of rotation. The Earth's equatorial axis is inclined at approximately 23.44° with regard to the plane of the ecliptic. The angle between the ecliptic and the lunar equator is 1.54°. The angle between the Earth's equatorial plane and the orbital plane of the Moon varies within a range from 23.44° ± 5.145°. This angle changes according to the precessional period of the Moon's orbit, approximately once every 18.6 years.

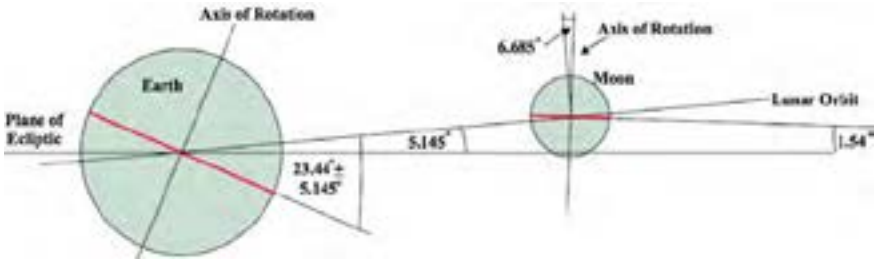


Figure 1. Orientation of the Lunar Orbit Relative to Earth's Equator and the Ecliptic Plane.

Figure 2 plots the inclination of the lunar orbit with regard to Earth's equator beginning on January 1 2020 and running through December 2030. The inclination angle lies at 22.9º in 2020, peaks at approximately 28.6º in early 2025, and drops back below 23º in late 2030. The mean inclination angle over this ten-year period is approximately 26.5º. Thus, depending on the launch date and site, the inclination angle change required in order to reach lunar orbit varies within a *10*° range, and can make a significant difference with regard to the ΔV required to reach Low Lunar Orbit. (*LLO*). Launching during the period starting January 1, 2024 through January 1, 2026 where the moon is at an inclination angle above 28º, presents the most advantageous lunch-period for the proposed mission.

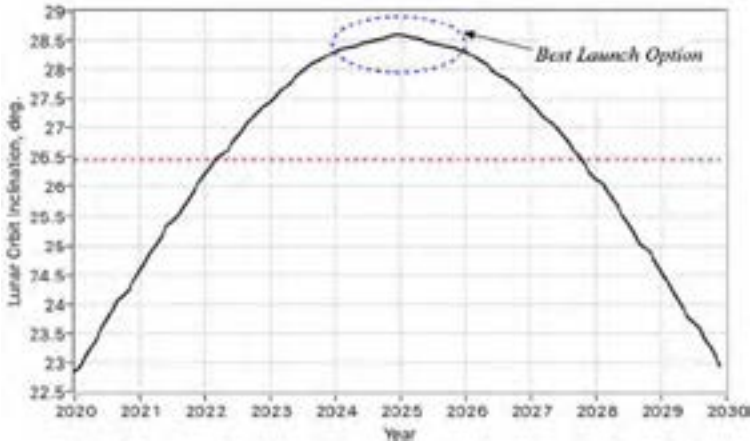


Figure 2. Lunar Orbit Inclination, 2020 through 2030.

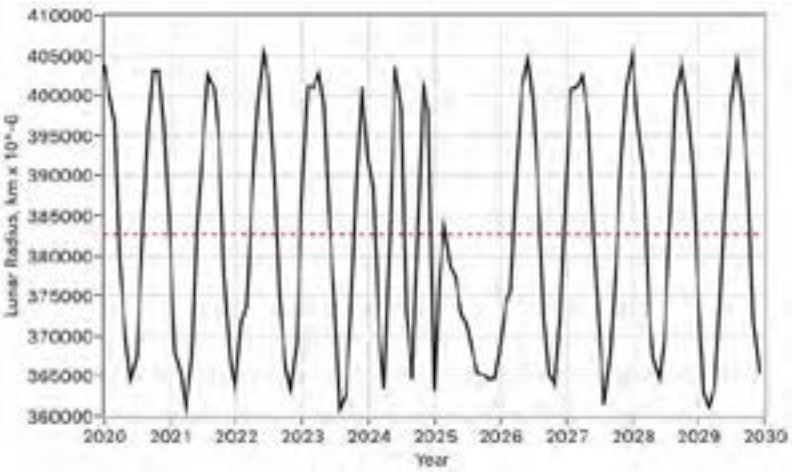


Figure 3. Earth-Moon Distance on January 1, 2020 through 2030.

The lunar orbit radius varies from a perigee of approximately *363,000 km* to and apogee of slightly greater than *405,000 km,* with a typical synodic period of period of approximately 29.5 days. Figure 3 plots the distance from the earth to the Moon on January 1, beginning in 2020 and running through December 2030.[16] From this data set, the mean is calculated to be approximately *382,600 km*. That radius value will be used for all follow-on *LTO* calculations to be presented in the next sub-section.

E. Lunar Transfer Orbit (LTO) Analysis

The data presented by Table 2 show promise for delivering at least a reasonable payload size to Lunar Transfer Orbit. As with the previous section, the required ΔV and delivered payload masses are calculated assuming a transfer burn from the *LEO* perigee. The methods of Table 2 are now repeated to calculate the required ΔV and delivered payload masses for insertion into *LTO* transfer from the original LEO parking obit. Only direct lunar transfers are considered in this analysis.

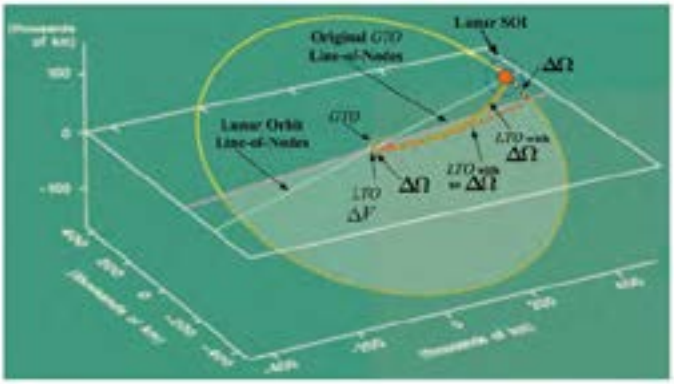


Figure 4. Schematic of LEO-TO-LTO Maneuver with and without $\Delta\Omega$ (RAAN) ΔV Maneuver.

Lunar Orbit can be reached by direct lunar transfer without a plane change, only if the line-of-nodes of the *LTO* and those of the mean lunar orbit coincide. The launch window for a direct transfer trajectory results from the angular difference between the launch site latitude and the lunar declination angle.** As long as the launch latitude is greater than the lunar orbit inclination angle, the launch time can be set so that the nodal line of the initial *LEO* orbit points in the direction of the lunar right ascension of the ascending node (*RAAN*), Ω.

Thus, direct lunar rendezvous, where the lunar line-of-nodes coincides with the launch *RAAN,* is possible only twice per month. Assuming that the launch window timing is such that the original *LEO RAAN* orbit does not intercept lunar orbit precisely along the line-of-nodes, then a $\Delta\Omega$ correction is required during the transfer burn. This launch analysis allows up to a $\pm 5^o$ $\Delta\Omega$ correction; a value that corresponds to approximately 18.2 hours of right ascension. Thus, the

** Declination angle is defined as the orbit inclination angle with respect to the Earth's equator, and is the equivalent of terrestrial geocentric latitude projected onto the celestial sphere.

launch window is expanded by an equivalent number of hours. Figure 3 illustrates this *RAAN* change during the *LTO* insertion.

If the plane change is performed as a combined maneuver initiated at the *LEO* orbit apogee, the required total ΔV is calculated using Equation (3). In Eq. 3 $V_{LEO\,p}$ is the orbital velocity at the original *LEO* orbit perigee, $V_{LTO\,p}$ is the transfer orbit perigee velocity calculated assuming an apogee altitude corresponding to the mean lunar orbit radius $R_{\mu L}$, and $\Delta\Omega$ is the budgeted *RAAN* correction [Ref. 15], 23

$$\Delta V_{\substack{LEO \\ LTO}} = \sqrt{V^2_{LEO_P} + V^2_{LTO_P} - 2 \cdot V_{LEO_P} \cdot V_{LTO_P} \cdot \cos\Delta\Omega} \quad . \tag{3}$$

Table 3 presents the results of the LTO insertion calculations. The calculations of Table 3 assume a lunar arrival time of *12:00 hours GMT* on Jan 1, 2024. A full 5º of *RAAN* change at *LTO* insertion is also assumed. The precise ΔV and delivered payload values will need to be slightly adjusted for the actual flight date and arrival time, and the required RAAN change. Note that the ΔV required to enter *LTO* with the maximum 5º *RAAN* change is about *1/3rd* higher than the value required for *GTO* insertion. This moderate cost result is expected since the escape ΔV from a 200 km *LEO* orbit is only slightly higher than this *LTO* value, approximately *3.23km/s*.

Table 3. Lunar-Transfer Orbit (*LTO*) Payload Capability Extrapolated from the Data of Table 1.

Launch Provider, Vehicle	***LEO* Payload**	***LEO* to *LTO* Burn ΔV *(with $\pm 5^o$ $\Delta\Omega$)***	***LTO* Payload**	**Consumed Propellant for *LTO***	**Usable Payload @ *LTO***	***GTO* Launch Cost (Usable Payload)**
Rocket Labs, Electron	221 kg	3.100 km/sec	74.29 kg	146.71 kg	48.40 kg	$123,964/kg
Vector Launch, Vector-H	263 kg	3.117 km/sec	87.88 kg	175.12 kg	56.98 kg	$ 87,754/kg-$122,855/kg
Relativity Space, Terran-1	1000 kg	3.117 km/sec	334.15 kg	665.85 kg	216.65 kg	$ 46,159/kg
Gilmour Space Technologies, ERIS-5	450 kg	3.170 km/sec	147.632 kg	302.38 kg	94.26 kg	$106,090/kg-$159,136/kg

The resulting mass to *LTO* is reduced by an additional 24-25% compared to *GTO*. The consumed propellant listed in column 6 of Table 3 is calculated as the difference between the *LEO* and *GTO* delivered payloads. The "usable" payload listed in column 6 of Table 3, assumes that the depleted transfer stage dry-mass is approximately 15% of the initial propellant mass loading, and decrements that value from the extrapolated *LTO* total payload mass. This assumption is equivalent to 17.65% of the consumed propellant for the propulsion stage.

Of the 4 launch systems, the Rocket Labs Electron vehicle delivers the smallest "useable" payload of slightly less than 50 kg to *LTO*; while the Relativity Space Terran-1 delivers the largest, approximately 217 kg. The Terran-1 also has the lowest projected launch cost; with the Gilmour ERIS-5 projected to have the highest costs to reach Lunar transfer orbit. The Terran-1 and ERIS-5 Launch Vehicles are the only options that will allow more than 90 kg to be inserted into *LTO*.

F. Reaching Lunar Orbit

As described previously Lunar Orbit can be reached by direct lunar transfer without a plane change, only if the line-of-nodes of the *LTO* and those of the mean lunar orbit coincide. Also, the moon must be at the target point as the spacecraft arrives. With a properly phased argument of perigee ω, as set by the *initial LEO* insertion burn time; and *RANN,* adjusted during the insertion burn, to match the moon's line-of-nodes; then as the spacecraft approaches the lunar *Sphere of Influence* (*SOI*), it becomes possible to drop into lunar orbit with a highly-elliptical, but non-hyperbolic Lunar Capture Orbit (*LCO*).

Figure 5 shows a representative direct launch *LTO* trajectory from the NASA Kennedy Space Center (*KSC*) Launch Complex, emulating a possible launch trajectory that would be available Terran-1 launch vehicle. The *LTO* trajectory of Figure 5 is designed so that the spacecraft arrives at its closest approach to the moon at 200 km lunar altitude, on or near January 1, 2024. Here, the spacecraft departs with a due east launch, setting the resulting *LTO* inclination at 28.5^o -- the latitude of the *KSC* launch complex -- and the perigee altitude at approximately 200 km. Figures 5(a) and 5(b) show the spacecraft and lunar position as it approaches the moon's *SOI* interface at approximately 66,000 km distance. Here the spacecraft is captured by the lunar gravity field, and falls toward the lunar surface. Figures 5(c) and 5(d) show the spacecraft and lunar position at the closest approach to the lunar surface.

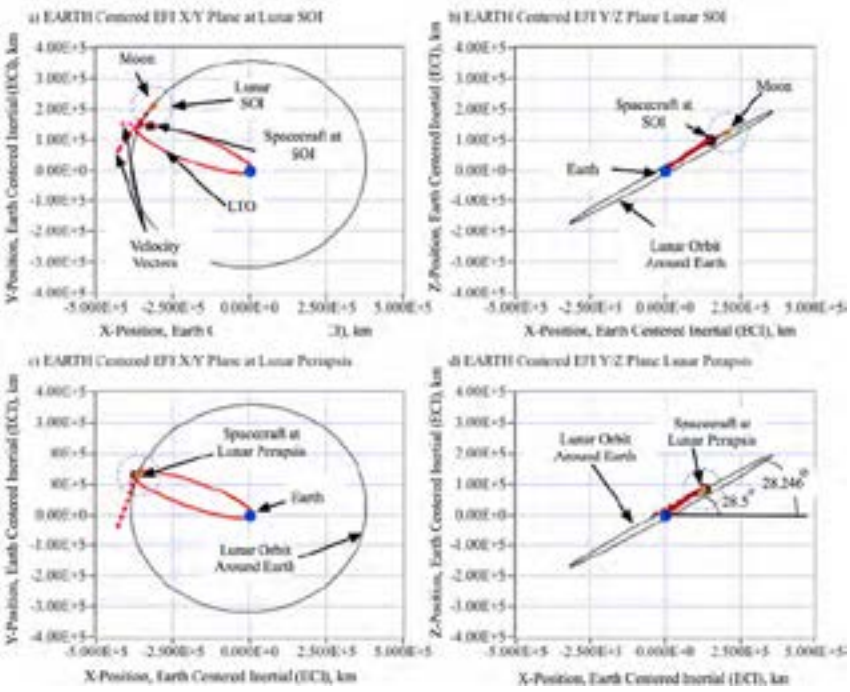


Figure 5. Schematic of Representative Direct LTO Transfer from 200 km Perigee Altitude, and Lunar Arrival January 1, 2024.

G. Patched-Conic Analysis of Lunar Orbit Insertion

Assuming that the appropriate $\Delta\Omega$ *LTO* burn adjustment in is performed, then the spacecraft approaches the moon along the line-of-nodes, and only a single retrograde ΔV burn at the orbit periapsis is required to enter the final circular *LLO*. Figure 6 shows this final maneuver. Figure 6(a) shows the initial elliptical highly *LCO* and the final circular *LLO* looking down on the lunar equatorial plane. Figure 6(b) shows the same data, but viewed from the horizontal plane, facing in the direction of the Equinox of Ares.

The initial highly elliptical *LCO* is calculated by differencing the *LTO* orbit position and velocity vectors from the lunar plane orbit position and velocity vectors in the original *ECI* coordinates. The resulting differenced-vectors are then rotated to plane of the lunar orbit, and then again the plane of the lunar equator, which is inclined to the lunar orbit plane by 6.685°. The

position and velocity vectors in Lunar Centered Inertial (*LCI*) are then used to calculate the moon-relative orbital elements using the analytical methods as detailed by Sellers, et al. [15]

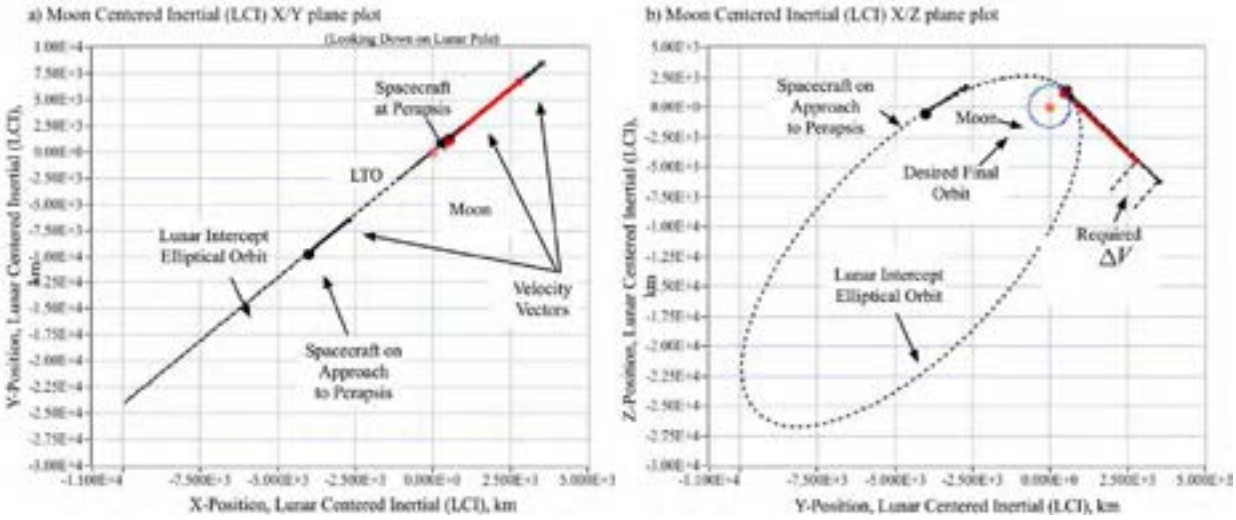


Figure 6. Schematic Showing Initial Lunar Capture Orbit (*LCO*), and Resulting Low Lunar Orbit (*LLO*) after Orbit Insertion Burn.

For each of the launch service providers previously listed by Tables 1-3, Table 4 summarizes the results of the patched-conic analysis for lunar orbit insertion. In all cases the final lunar orbit (*LLO*) is polar, inclined at 90º to the lunar equator, and has a mean orbital altitude of 200 m. The orbit right ascension varies depending upon the precise arrival time in lunar orbit. The target arrival date for the spacecraft to lunar orbit is, on or near, January 1, 2024. Listed, are the previously estimated *LTO* usable mass $M_{u,\ LTO}$; the orbital elements of the transfer *LTO* and initial lunar capture *LCO*; the ΔV requirements of the final burn to circularize the orbit; the extrapolated masses to *LLO*, the required total impulse I_{tot} for the Lunar Capture Stage (*LCS*), and the estimated cost per kg of usable payload delivered to the *LLO*. The assumed vacuum I_{sp} for the *LCS kick motor* is 280 seconds.

The "*usable*" *LLO* payload listed by the 6th column of Table 4, decrements the mass of depleted LCS motor. As with the LTO analysis this calculation assumes that the depleted capture stage dry-mass is approximately 15% of the initial propellant mass loading, and decrements that value from the extrapolated *LTO* total payload mass. This assumption is equivalent to 17.65% of the consumed propellant for the propulsion stage. Equation (4) shows an example pf this calculation for the *Rocket Labs*, *Electron* launch system. Here, the calculation sequence is derived from the rocket equation (Seller's [15], Chapt 6.)

Total Mass to LLO,

$$M_{T,LLO} = \frac{M_{u,LTO}}{e^{\left(\frac{\Delta V_{LLO}}{g_0 \cdot Isp}\right)}} = \frac{48.40_{kg}}{e^{\left(\frac{539_{m/\sec}}{9.8067_{m/\sec^2} \cdot 290_{\sec}}\right)}} = 38.37_{kg}$$

Propellant Consumed from LTO to LLO,

$$M_{p,LLO} = M_{u,LTO} - M_{T,LLO} = M_{u,LTO} \cdot \left(1 - \frac{1}{e^{\left(\frac{\Delta V_{LLO}}{g_0 \cdot Isp}\right)}}\right) = 48.40_{kg} - 39.37_{kg} = 8.63_{kg}$$

Usable Mass to LLO,

$$M_{u,LLO} = M_{T,LLO} - \varepsilon \cdot M_{p,LLO} = 39.77kg - 0.1765 \cdot 8.65_{kg} = 38.25 \quad . \qquad (4)$$

The conclusions of Table 4 are similar to those of the previous *LTO* analysis. Of the 4 launch systems, the *Rocket Labs Electron* vehicle delivers the smallest "useable" payload of slightly greater than 37 kg to *LLO*; while the *Relativity Space Terran-1* delivers the largest, slightly less than 170 kg. The Terran-1 also has the lowest projected launch cost; with the Gilmour ERIS-5 projected to have the highest costs to reach Lunar Orbit, but can deliver approximately 72 kg to *LLO*. The Terran-1 and ERIS-5 Launch Vehicles are the only options that will allow more than 70 kg to be inserted into *Low Lunar Orbit*.

Table 4. Results of Patch-Conic Analysis: Required ΔV and Low Lunar Orbit (LLO) Payload Capability.

Launch Provider, Vehicle	**Usable Payload @ *LTO***	***LTO* Orbit Elements (*ECI*)**		***LCO* Orbit Elements (*LCI*)**	***LTO*-to-*LLO* ΔV**	**Estimated Mass to *LLO***		**Required Total Impulse, I_{tot}**	***LLO* Launch Cost**
Rocket Labs, Electron	48.40 kg	*a*	204,594 km	19,207.9 km	0.6016	Total	38.88 kg	26,148.5 N-sec (5876.72 lbf-sec)	$161,305
		e	0.9680	0.8992		Used Propellant	9.523 kg		
		i	37.95°	89.99 °		Usable Payload	37.20 kg		
		Ω	-3.376°	248.14°					
		ω	-19.764°	135.73 °					
Vector Launch, Vector-H	56.98 kg	*a*	204,621 km	19,395.3 km	0.6022 km/sec	Total	45.76 kg	30811.3 N-sec (6924.67 lbf-sec)	$ 114,210 - $159,894
		e	0.9679	0.9002		Used Propellant	11.22 kg		
		i	37.95°	90.00°		Usable Payload	43.78 kg		
		Ω	-3.379°	248.14°					
		ω	-19.762°	135.67°					
Relativity Space, Terran-1	216.65 kg	*a*	204,621 km	18,748.8 km	0.6002 km/sec	Total	174.11 kg	116,803 N-sec (26,250.8 lf-sec)	$ 60,022
		e	0.9679	0.8967		Used Propellant	42.54 kg		
		i	28.50°	89.99°		Usable Payload	166.61 kg		
		Ω	3.733°	247.53°					
		ω	-25.734°	122.93°					
Gilmour Space, ERIS-5	94.26 kg	*a*	204,969 km	23,756.72 km	0.6128 km/sec	Total	75.41 kg	51,770.9 N-sec (11,625.2 lbf-sec)	$138,737 - $210,880
		e	0.9679	0.9185		Used Propellant	18.85 kg		
		i	31.00°	90.00°		Usable Payload	72.08 kg		
		Ω	1.350°	247.58°					
		ω	-23.753°	132.89°					

H. Sizing the Lunar Capture Stage (LCS)

The total impulse I_{tot} value listed by the 6^{th} column of Table 4, gives guidance that allows a top-level sizing estimate of the *Lunar Capture Stage (LCS)* to be performed. As mentioned previously, the launch vehicle is assumed to place the spacecraft onto the lunar transfer orbit phased with the proper argument of perigee and RAAN, and the *LCS* is tasked with performing only a single retrograde orbit insertion burn. This analysis will assume that a solid motor can be sized appropriately to deliver the required impulse value.

For simplicity, and as cost saving measure to avoid non-recurrent engineering development costs for this mission, this analysis selects the *LCS* stage as the "nearest match" from the most current Northrop Grumman Innovation Systems (*NGIS*) Solid Propulsion Products Catalog.[17] Since these "off the shelf" motors do not precisely match the impulse values as shown by Table 4, the precise DV that is delivered to the payload, the final usable mass delivered to *LLO*, and the resulting semi-major axis and eccentricity (or apsis and periapsis) of the resulting *LLO* must be re-calculated.

Based on the actual delivered total impulse I_{tot} and I_{sp} of the kick motor, the recalculated "delivered" ΔV is

$$\Delta V_{delivered} = g_0 \cdot I_{sp} \cdot \ln\left(\frac{I_{sp} \cdot g_o \cdot M_{initial}}{I_{sp} \cdot g_o \cdot M_{initial} - I_{tot}} \right)_{deliverd} . \tag{5}$$

The recalculated usable mass $M_{u,\,LLO}$ is now

$$M_{U,\,LLO} = M_{initial} \cdot e^{\left(-\frac{\Delta V}{g_0 \cdot I_{sp}} \right)_{delivered}} - M_{stage,dry} , \tag{6}$$

where, $M_{stage,\,dry}$ is the residual mass of the expended kick stage. The semi-major axis and eccentricity of the resulting *LLO* is calculated by applying the Vis-Viva equation to the desired original 200 km altitude orbit, to the orbit achieved using the actual delivered impulse value. Here,

$$\left[V_{LLO}^2 \right]_{desired} - \left[V_{LLO}^2 \right]_{delivered} = \mu \cdot \left[\left(\frac{1}{a_{LLO}} \right)_{delivered} - \left(\frac{1}{a_{LLO}} \right)_{desired} \right] . \tag{7}$$

In Eq. (7) the subscript $()_{desired}$ denotes the desired quantity as presented by Table (4); the subscript $()_{delivered}$ denotes the actual value that is delivered by the selected lunar perigee kick motor. Eq. (7) can be re-written in terms of *"desired"* and *"delivered"* ΔV, and when solved for the actual delivered *LLO* semi-major axis the result is

$$\left(a_{LLO} \right)_{delivered} = \frac{\left(a_{LLO} \right)_{desired}}{1 + \frac{2}{\mu} \cdot \left(a_{LLO} \right)_{desired} V_{LCO_P} \cdot \left(\Delta V_{desired} - \Delta V_{delivered} \right) \cdot \left(1 - \frac{\left(\Delta V_{desired} + \Delta V_{delivered} \right)}{2 \cdot V_{LCO_P}} \right)} . \tag{8}$$

In Eq. (8) the burn is assumed to occur impulsively at the *LCO* periapsis radius where *the* velocity is $V_{LCO\,p.}$

Table 5 presents the recalculated *LLO* values using the closest match kick-motors values selected from Ref. [17]. For each of the launch vehicles, generally there exists an NGIS star-motor that accomplishes the *LLO* insertion task successfully. For the Electron launch vehicle, it is

required to bundle two small Star 6 motors together is order to generate sufficient impulse to insert the spacecraft into *LLO*. Depending upon whether the actual delivered impulse of the *NGIS* stage is smaller or larger than the "desired" impulse of Table 4, the orbit is either slightly larger, or smaller than the ideal 200 km altitude circular orbit that was previously baselined.

Table 5. Modified Low Lunar Orbit (LLO) Orbit Parameters and Payload Capability using NGIS Star Motor as Lunar Capture Stage (*LCS*).

Launch Provider, Vehicle	**Usable Payload @ *LTO***	**NGIS Motor Designation**		***Δ V* Delivered**	***LLO* Orbit Parameters *(Revised)***		**Delivered Mass to *LLO* *(Revised)***		***Revised LLO* Cost**
Rocket Labs, Electron	48.40 *kg*	**Star 6 *(2 x)***		0.6314 *m/sec*	*a*	1867.46 *km*	Total:	38.70 *kg*	$165,400
		I_{tot}	27,382.2 *N-sec*		*e*	0.03702	Used Propellant	9.70 *kg*	
		I_{sp}	287.9 *sec*		h_a	200.00 *km*	**Usable Payload**	**36.28 *kg***	
		L_f	0.8		h_p	61.72 *km*			
Vector Launch, Vector-H	56.98 *kg*	**Star 10**		0.5783 *m/sec*	*a*	1997.01 *km*	Total:	45.05 *kg*	$126,800 - $177,520
		I_{tot}	29,366.7 *N-sec*		*e*	*0.03025*	Used Propellant	11.93 *kg*	
		I_{sp}	250.95 *sec*		h_a	*320.83 km*	**Usable Payload**	**39.43 *kg***	
		L_f	0.68		h_p	*200.0 km*			
Relativity Space, Terran-1	216.65 kg	**Star 13B**		0.6033 *m/sec*	*a*	1929.23 *km*	Total:	169.55 *kg*	$ 61,300
		I_{tot}	115,909 *N-sec*		*e*	0.0038204	Used Propellant	47.10 *kg*	
		I_{sp}	286.58 *sec*		h_a	200.00 *km*	**Usable Payload**	**163.13 *kg***	
		L_f	0.88		h_p	185.26 *km*			
Gilmour Space, ERIS-5	94.26 kg	**Star 12**		0.5431 *km/sec*	*a*	2126.77 *km*	Total:	75.98 *kg*	$150,250 - $228,280
		I_{tot}	46,052.3 *N-sec*		e	0.08942	Used Propellant	18.28 *kg*	
		I_{sp}	56.82 *sec*		h_a	580.33 *km*	**Usable Payload**	**66.56 *kg***	
		L_f	0.66		h_p	200.00 *km*			

Only the Terran-1's required impulse level is a close match to the available NGIS Star motor. The resulting orbit has a low eccentricity, with a semi-major axis that less than 1% different from the initial baseline *LLO* orbit. The NGIS Star 13B motor (Ref [17], p. 79) is ideally sized to deliver the Terran-1 payload into *LLO*. The STAR 13B incorporates a titanium case developed for the STAR 13 with the propellant and nozzle design of an earlier STAR 13 apogee motor. The motor design was qualified in 1983 and was previously used twice for space missions; 1) in 1984 to adjust orbit inclinations of the Active Magnetosphere Particle Tracer Experiment satellite launched from a Delta 180 stage and 2) in 1988 as a kick motor for a classified missile defense experiment.

I. Down Select for "Scout" Mission Launch Vehicle

The data of Table 5 were used to rank the 4 launch vendors according to capability and cost. Four Factors will be considered in the final Pugh-decision matrix. These are 1) Payload to *LLO*, 2) Closeness of final LLO to original base, 3) Cost-per-kilogram of payload to *LLO*, and 4) TRL maturity of the System. Each of these factors are weighted equally in the decision matrix. For each factor the score varies from best (1) to worst (4), low score wins. Table 5 presents the results of

the Pugh analysis. For Launch Vendors with High and Low cost estimates, the median value is used for this analysis. The orbit deviation is calculated by the root sum square of the periapsis and apsis radius differences between the achieved orbit and the baseline, divided by the baseline circular orbit radius, i.e.

$$\varepsilon_{LLO} = \frac{1}{a_{LLO_{baseline}}} \sqrt{\left(R_{p_{achieved}} - a_{LLO_{baseline}}\right)^2 + \left(R_{a_{achieved}} - a_{LLO_{baseline}}\right)^2} \quad . \tag{9}$$

Table 6 shows the Pugh decision matrix results. The Terran-1 is the clear winner based up on these criteria, with the Vector-H coming in a distant second. the Gilmore space, *ERIS-5*, due to its lower *TRL* development stage, and higher reported launch costs ranks last of the 4 launch system considered by this study. The Electron vehicle offers the highest *TRL* maturity; but delivers the smallest payload to *LLO,* and a cost that is more than twice the reported cost for the Terran-1.

Due to the comparative immaturity of the Terran-1 launch system and the unproven business track-record of Relativity Space; down-selecting to the Terran-1 substantially increases mission risk. Considering this increased risk; however, the significantly increased payload capacity and reduced cost make this increased risk more than acceptable. The Terran-1 payload to *LLO* and the corresponding *LLO* orbit parameters listed by Table 5 will be used to complete the final lunar descent study phase of this project.

Table 6. Pugh Decision Matrix for Down-Selection of Launch Services Provider.

Launch Provider, Vehicle	**Usable Mass to *LLO***	**Orbit Deviation from Baseline**	***LLO* Cost**	**Maturity, TRL**	**Total Score, Rank**
Rocket Labs, Electron	Mass: 36.28 *kg* Rank: 4	Deviation: 7.14% Rank: 3	Cost: $165,400 Rank: 3	TRL: 9 (operational) Rank: 1	Total: 11 **Rank: 3**
Vector Launch, Vector-H	Mass: 39.43 *kg* Rank: 3	Deviation: 6.23% Rank: 2	Cost: $152,160 Rank: 2	TRL: 8 (Actual System Qualified) Rank: 2	Total: 9 **Rank: 2**
Relativity Space, Terran-1	Mass: 163.13 *kg* Rank: 1	Deviation: 0.77% Rank: 1	Cost: $ 61,300 Rank: 1	TRL: 7 (System Undergoing Qualification) Rank: 3	Total: 6 **Rank: 1**
Gilmour Space, ERIS-5	Mass: 66.56 *kg* Rank: 2	Deviation: 19.64% Rank: 4	Cost: $189,235 Rank: 4	TRL: 6 (Scaled Model Suborbital launch) Rank: 4	Total: 14 **Rank: 4**

III.Descending to the Lunar Surface: Scout Mission Concept (*ScMC*) of Operations

Images provided by the previously described lunar orbiting spacecraft have identified that a typical "skylight" opening is quite small, with diameters ranging from slightly greater than 5 meters to less than 900 meters.[18] Large skylights are indicative of partial collapse of the subsurface features, and such unstable tubes would not be ideal for habitat development. Thus, to be useful the scout vehicle must be able to "hit" a very small surface opening with a rather unsophisticated small lander. This lander must reach the target site without the aid of surface landing beacons and expensive/heavy onboard Inertial Navigation Systems (*INS*).

This study will conclude by proposing Scout Mission Concept (*ScMC*) of operations that sends multiple small spacecraft as opposed to one large spacecraft to the lunar surface. Due to the rather large *LLO* payload capacity of the Terran-1 lunch vehicle, more than 160 kg; there exists the option to send multiple landers to the lunar surface to investigate the target site. The option of having redundant landers available significantly improves the odds of hitting the desired skylight opening, and would clearly reduce mission risk.

Landers, hereon referred to as "MeatBalls," that do not hit the target opening can be used as communication relays, increasing signal bandwidth, and allowing high-definition video and other data to be relayed back to the orbit platform. The orbiting ScMC platform serves to relay surface telemetry data to the Artemis gateway or earth-based deep-space receiving networks (DSN). Should early landing attempt be successful, then having backup landers provides the option to explore another surface target that lies along the orbital track. Figure 7 shows a "cartoon" schematic of the proposed *ScMC* architecture.

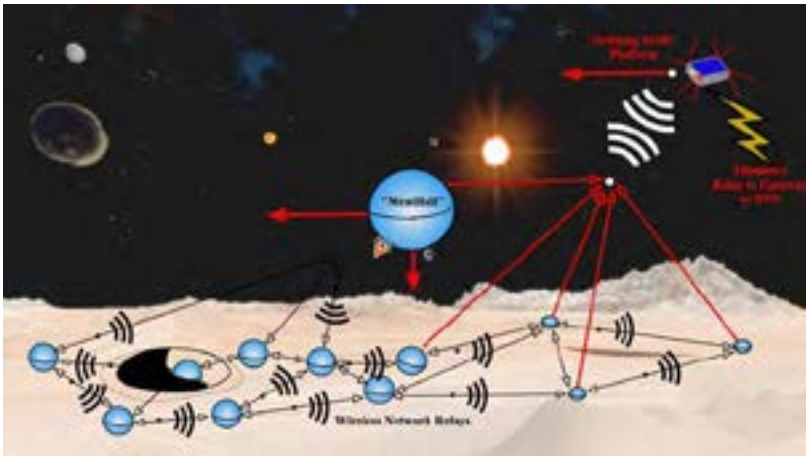


Figure 7. *ScMC* "MeatBall" Lander System Architecture.

A. Top-Level Concept for ScMC LLO Platform

As will be shown in the next section, the lander "MeatBalls" will be sized to weigh 10-kg each, including the de-orbit propulsion module and the propulsion system required for landing. Because the Terran-1 can deliver up to 163 kg to *LLO*, then there exists the possibility for the orbiting Scout platform to carry up to 10, bowling-ball sized, 10-kg MeatBall landers while still leaving up to 63 kg of residual mass that is available for the *ScMC Orbiting Platform*. This mass-budget does not include the mass of the previously-described Star 13B *LCS* motor, which was budgeted in the usable mass calculation of Table 6.

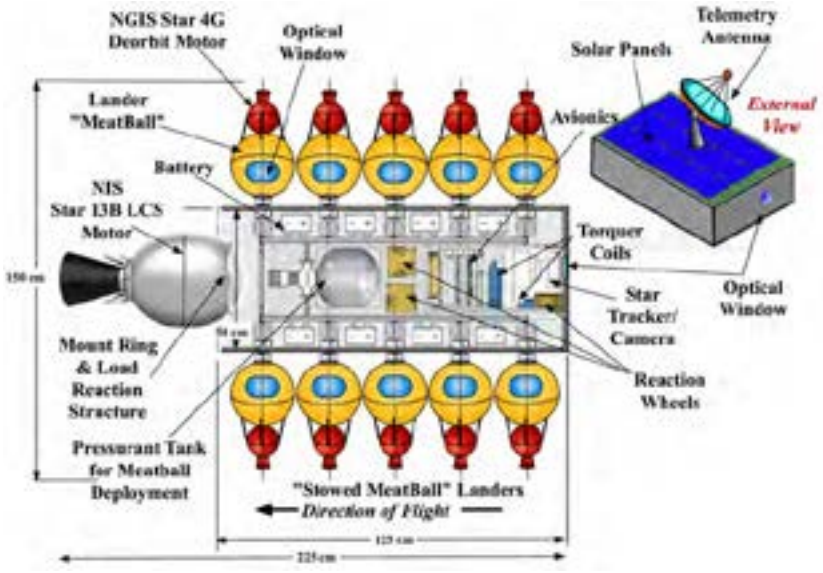


Figure 8. Top-Level Schematic of the *ScMC* Orbiting Platform Design.

Figure 8 shows a top-level layout of the *ScMC Orbiting Platform*. The platform is drawn showing 10 Stowed MeatBall Landers awaiting deployment. Sufficient avionics including the flight computer, a strap down Inertial Navigation Sensor, and a Star-Tracking Camera System are allowed

in order to provide the detailed orbital ephemeris data required to phase the lander deployment times properly. Attitude is assumed to be controlled using reaction wheels, and at least two magnetic torque coils are included as a contingency in case the reaction wheels need desaturation. Although the moon's magnetic fields are considerably weaker and more irregular than Earth's dipole magnetic field, and generally point away from the lunar surface;[19] studies have estimated that there exits sufficient strength to desaturate reaction wheels, given sufficient time of up to 50 days.[20] The external surface of the platform is paneled with photovoltaics to supply power to the subsystem.

The "MeatBall" deployment system assumes a pneumatic separation system, with MeatBalls on opposite sides of the system plumbed to the same pressure source, so separation occurs simultaneously. Simultaneous deployment avoids generating spin torques are the 10-kg masses are ejected. The stored MeatBall deployment gases also provide the option for using a cold-gas propulsion system to desaturate reaction wheels, in the event that the magnetic coils cannot provide sufficient torque. Mechanisms to spin the MeatBalls during deployment could easily be built into the system, these options have not, however, been investigated at this point.

Not including the Star 13B kick motor and the stowed MeatBall landers, the orbiting platform is roughly 1.25 meters long, 50 cm wide, and 25 cm deep. If the kick motor length is included, then the platform length grows to approximately 225 cm. When the sizes of the stowed MeatBall landers are accounted for, then the width of the deployed system grows to 1.5 meters. Methods for packaging this system into the upper stage of the Terran-1 launch vehicle have not been investigated at this point.

B. "MeatBall" Landing Vehicle Form Factor

For this study, the target volume for the landing system, is equivalent to a 6-U CubeSat. A spherical form factor for the lander is preferred due to its ability right itself after landing using a simple gyroscope or reaction wheel set and a vertical center of gravity offset. A spherical design also offers the possibility of bouncing the landing craft along the surface at a survivable speed, allowing a "hole in one" without the need for precise landing control, navigation aids, or hover guidance systems. Attitude control of the lander is assumed to be controlled using reaction wheels

According to international standards,[21] each "U" of the spacecraft would have dimensions of *10 x 10 x 10 cm*, and weigh no more than 1.333 kg per "U". Thus, a 6-U lander would weigh approximately *8 kg*, and have *6000 cm^3* of available payload volume. This size limit would then allow for a lunar deorbit propulsion module weight of up to 2-kg, with the resulting on-orbit mass of approximately 10 kg for each "MeatBall." A 6-U volume has an equivalent spherical diameter of *22.54 cm*. The internal layout and avionics of the MeatBall systems has not been investigated at this point.

J. Selected de-Orit Thruster for the MeatBall Landers

For this study, it is assumed that the de-orbit function for each MeatBall will be facilitated using an off-the shelf star motor selected from the NGIS product catalog. A trade study identified one motor, the NGIS Star 4G that is ideally suited for the MeatBall deorbit function. The STAR 4G motor was developed as a low-cost, high mass fraction orbit adjustment motor deploying constellations of nanosatellites. The motor incorporates a non-eroding throat insert to maximize specific impulse, and to minimize impulse variability. The Star 4G has never been demonstrated during a spaceflight.

This solid-propellant thruster is only 11.3 cm in diameter, yet it delivers 2627.1 N-sec of total impulse at a specific impulse of 269.4 sec. The mean thrust level is approximately 245 N. The total burn time for this motor is slightly longer than 10 seconds. Based on the 10kg Mass of the MeatBall, the deceleration rate during the burn is rather benign, approximately 2.5 earth g's. A total of 995 grams of propellant are consumed during a nominal burn, and the approximate retrograde ΔV is 279 m/sec for the 10 kg MeatBall lander. Figure 9 shows the Star 4G motor during test firing.

Figure 9. Star 4G Motor During Static Test Firing.

Figure 9b shows the thrust, chamber pressure, impulse, and massflow time history profiles during a typical 10-second vacuum burn. (Ref. [17], p. 69) The massflow profile was calculated from chamber pressure data using a one-dimensional De Laval Flow analysis,[22] with the combustion properties set to match the measured thrust profile. Figure 10(a) shows the measured (solid black lone) and calculated (dashed red line) thrust profiles. The plotted thrust and massflow profiles will be used later in this section for calculating the lunar surface deorbit trajectory. The inert mass of the motor after burnout is approximately *508 grams*.

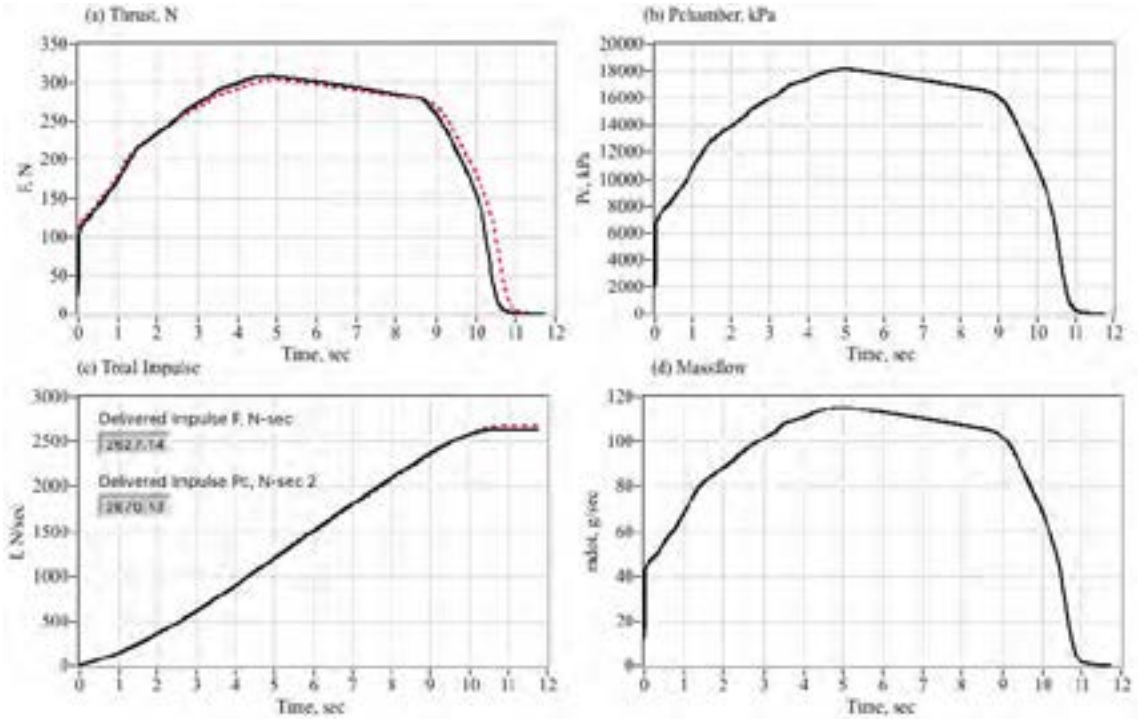


Figure 10. Vacuum Burn Time History Profiles for Star 4G Motor.

K. Modeling the Lunar Descent Trajectory

Because the motor burns for the descent to the lunar surface must distributed over a significant portion of the time of flight to the surface, simple impulsive Keplerian ΔV calculation will not be sufficiently accurate for mission analysis and planning. Instead the point mass equations of motor are written in two dimensions, and then integrated as a boundary value calculation with the spacecraft terminal velocity being a critical constrain parameter. The final downrange of the lander

is not considered to be a problem constraint as that value can be adjusted by the de-orbit time, once the total downrange of the landing trajectory is calculated. As derived by Whitmore (2019),[23] Equation (10) presents the 2-Dimensional state-space model describing the planar, ballistic flight of the spacecraft from de-orbit to the lunar surface,

$$\begin{bmatrix} \dot{V}_r \\ \dot{V}_\nu \\ \dot{r} \\ \dot{\nu} \\ \dot{m} \end{bmatrix} = \begin{bmatrix} \frac{V_\nu^2}{r} + \frac{F_{thrust}\sin(\gamma)}{m} - \frac{\mu_{moon}}{r^2} \\ -\frac{V_r V_\nu}{r} + \frac{F_{thrust}\cos(\gamma)}{m} \\ V_r \\ \frac{V_\nu}{r} \\ -\frac{F_{thrust}}{g_0 \cdot I_{sp}} \end{bmatrix}. \quad (10)$$

In this analysis the instantaneous thrust vector F_{thrust} is always assumed to point along the direction of the flight path angle γ, calculated as

$$\gamma = \tan^{-1}\left(\frac{V_r}{V_\nu}\right), \quad (11)$$

The parameters (V_r, V_ν) are the local vertical and horizontal velocity components, r is the instantaneous radial distance to the moon's center, and m is the current instantaneous spacecraft mass. Terms on the left-hand side of Eq. (10) are the derivatives of the state vector components. Starting with the initial *LLO* position and velocity at the time of deorbit, the descent trajectory is calculated by numerically integrating Eq. (10) over time.

Nominal Descent Trajectory of 10 kg MeatBall from Terran-1 LLO Apsis

A traditional Apollo-style propulsive-hover landing presents a significant mass penalty for this problem. During the Apollo final approach, the lander would zero-out the vertical velocity and hover above the target; gradually removing kinetic energy while lowering altitude. As the vehicle hovers each second costs a velocity increment ΔV that is proportional to the acceleration of lunar gravity, or approximately 1.65 m/sec. Thus, a 3-minute hover period is equivalent to a wasted ΔV, of approximately 297 m/sec, or more than 100% of ΔV input required to deorbit from the nominal 200 km lunar parking orbit. These associated gravity losses would severely restrict the available payload deliverable to the lunar surface.

Instead the descent must be significantly more deliberate with the spacecraft "slamming on the brakes" shortly before impacting the lunar surface. The timing of the braking maneuver is critical for success of the mission. Burning too soon along the descent trajectory has the effect of slowing

the spacecraft too fast to the point where the vehicle slows to zero velocity at a point above the lunar surface. At this point because the spacecraft weighs less than the braking thrust level in lunar g's, the vehicle will rapidly gain altitude and accelerate. In some of the cases, depending on mass allowed for propellant consumption, the lander will actually re-orbit or depart lunar orbit altogether. Burning too late along the descent trajectory does not allow sufficient time for the vehicle to decelerate to a safe landing velocity. The vehicle impacts at a non-survivable velocity.

Figure 11 shows the results of a descent and landing trajectory that is well optimized and allows a safe-landing using the prescribed motor parameters. Starting with the initial conditions of Table 1, the state equations (1) are integrated over time as an initial value problem. The initial de-orbit burn uses the prescribed Star 4G thrust profile of Figure 10. As calculated earlier, the initial Star 4G de-orbit burn reduces the orbital speed by 276.6 *m/sec,* and starts the freefall to the lunar surface along the shown track. Following the de-orbit burn, the *508 grams* of residual mass from the expended stage is discarded, and the lander mass is *8,497 grams* during the descent.

Following the de-Orbit burn, the problem becomes a boundary value calculation where the burn altitude and timing are adjusted by the boundary value solver in order to drive the vertical and horizontal landing velocities less than *1 m/sec*, and within 1 meter of the lunar surface. From the previous discussion, recall that the ballistic simulation assumes the thrust vector is always directed against the direct line of travel in order to achieve optimal performance. Burn loiter, where the thrust vector is directed off axis from the flight path, were not allowed for this simulation.

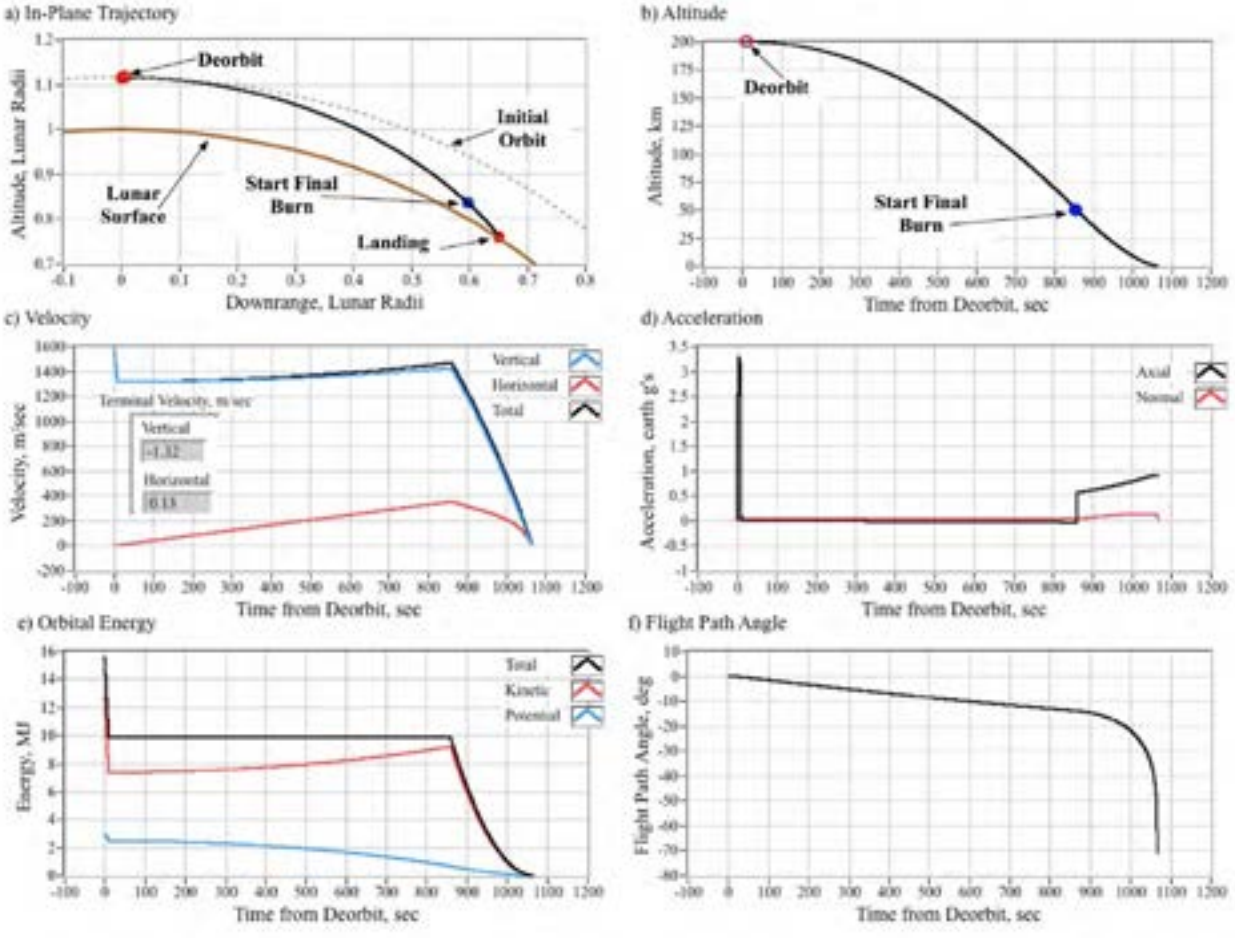


Figure 11. Well-optimized MeatBall Lander Descent, Approach, and Landing Trajectory.

Table 7 summarizes the key events and the corresponding trajectory parameters for the de-orbit, descent, and landing. Recall from Table 5 that the Star 13B motor delivers the spacecraft that into a slightly elliptical *LLO* whose perigee is slightly lower than the desired 200 km altitude. the initial *LLO* orbit parameters, and also shows the calculated state parameters corresponding to the apsis position within the *LLO*. Note that de-orbit is most efficient at the orbit apsis where the horizontal velocity is lowest.

Key events including the de-orbit burn, initiation of the final landing burn, and impact are indicated. Peak velocity during the free fall occurs at an altitude of approximately *50* km above the lunar surface. Figure 11(f) shows the flight path angle time history plot. The initial approach trajectory is rather shallow; but as the vehicle slows right before landing, the flight path angle approaches vertical rather quickly. Even at a height of 10 km above the lunar surface, a little more than 2 minutes from landing, the flight path angle is still only -20º. At this flight path angle ΔV gravity losses are accumulating at a rate of only 0.56 m/sec.

For this analysis the baseline the final descent motor is assumed to deliver a constant *50 N* of thrust over the entire burn profile, with a specific impulse of *290 seconds* During the descent an additional *3,632 grams* of propellant are consumed. The final landing mass is *4,865 grams*. The time from initiation of the braking burn to lunar impact is burn duration is precisely *206.56 seconds*. The second burn delivers 1585.9 m/sec of ΔV. *The total mission delta V is approximately 1,862.5 m/sec.* The entire mission time from de-orbit to impact is *1,067.89 seconds,* or just slightly less than *18 minutes*. The down range from the de-orbit burn is slightly greater than 1,200 km or approximately 11% of a full revolution.

The burnout altitude is *0.53* meters above the lunar surface. The vertical and horizontal impact velocities are -1.32 and 0.13 m/sec, respectively. The total impact energy is approximately *4.33 Joules*, or the equivalent to dropping the lander from a height of 9 cm (3.5 in.) on Earth. This impact would be easily survivable by properly-mounted lander instrumentation systems.

Table 7. Summary of Key Events and Spacecraft Parameters During de-Orbit, Descent, and Landing

Event	Time	Orbital Elements		Position		Velocity		Mass	γ
Initial Orbit	0 sec	*a*	*1929.23 km*	*h*	*200 km*	V_r	*0.0 km/sec*	*10 kg*	*0.0º*
		e	*0.00382044*	*ν*	*180.00º*	V_ν	*1,591.25 m/sec*		
De-Orbit	10.91 sec	*a*	*1469.72*	*h*	*199.833 km*	V_r	*-2.20 m/sec*	*9.005 kg*	*-0.10º*
		e	*0.31756*	*ν*	*180.469º*	V_ν	*1,314.5 4m/sec*		
Final Burn Initiation	860.44 sec	*a*	*1469.72*	*h*	*48.674 km*	V_r	*-351.83 m/sec*	*8.497 kg*	*-13.86º*
		e	*0.31756*	*ν*	*180.469º*	V_ν	*1,425.84 m/sec*		
Landing Motor Burnout	1,067.75 sec	*a*	*868.3*	*h*	*0.534 meters*	V_r	*-0.01 m/sec*	*4.865 kg*	*-70.92º*
		e	*1.0*	*ν*	*220.5º*	V_ν	*0.13 m/sec*		
Impact		*a*	*868.3*	*h*	*0 meters*	V_r	*-1.32 m/sec*	*4.865 kg*	*-84.38º*

	1,067.89 sec	*e*	*1.0*	*v*	*220.479°*	V_v	*0.13 m/sec*		

L. Discussion of Landing Trajectory Results

Based upon the results of Figure 11, it appears that if the initial trajectory is carefully prescribed and the de-orbit burn timelines are carefully controlled, that it would be theoretically possible for the MeatBall lander system to be configured with a solid-motor descent stage. However, it is doubtful that the required precision levels can be achieved; especially for such a small spacecraft.

For example, with a 3 second decrease (1%) in I_{sp} of the lander descent propulsion system, the spacecraft mass now depletes at a rate that is 3% faster than nominal. This faster burn rate results in a motor burn time that is 2.14 seconds shorter. Assuming a propellant mass load of *3,632 grams (*from previous the optimal solution*),* the *motor now burns out at* an *altitude of 126 meters above the lunar surface with a vertical velocity of -7.3 m/sec*. The impact velocity resulting from this fall is now *21.5 m/sec*. The resulting impact energy is now approximately *1,126 Joules*. This impact energy is now greater by a factor of more than 250 times. It is doubtful that the lander would survive this impact.

Similarly, should the I_{sp} increase by 3 seconds, the mass of the spacecraft now depletes at a rate that is 3% slower than nominal. Because the space craft now weighs more during the entire burn period, a larger total impulse is required to give the same ΔV; and the impact velocity is now *-36.8 m/sec* in the vertical direction. The vehicle also weighs slightly more at *5.03 kg*. The resulting impact energy is now *3,406 Joules*, or the equivalent of dropping the lander from an altitude of 69 meters on Earth. This fall would clearly not be survivable unless the lander systems were highly ruggedized.

On the Need for Closed Loop Energy Management During Final Approach.

Similar arguments can also be made with regard to initial de-orbit burn and final burn initiation altitudes and initiation times. Other factors that must be considered are variations of the local lunar gravity field, and altitude of the landing site terrain. Clearly, some sort of loop closure is necessary during landing in order to ensure a reasonable degree of mission success. Potential solutions include loiter, pulsed burning, and throttle. These methods can be used to track the optimal-trajectory energy curve associated as shown by Figure 11(c).

Clearly, calculating the energy level along the descent trajectory requires accurate tracking of the of the spacecraft position, velocity, and altitude. Thus, a suite of inertial sensors with sufficient accuracy levels will be required as a part of the MeatBall instrumentation package. Because the flight to the surface after de-orbit is rather short (18 minutes) and the acceleration levels are low (less than 3.5 g's), sufficient accuracy can likely be achieved by an integration-loop using strap down accelerometers and rate gyros, de-biased shortly before launch using higher-fidelity instrumentation available on the orbiting platform.

Because the mass varies by more than 50% during the descent, an accurate estimate of the total spacecraft mass must also be generated. Because the mass cannot be directly measured during flight, it must be calculated using a dynamic model of the vehicle, and some measure of the thruster chamber pressure in order to calculate the nozzle exit massflow. Because there are no drag forces acting, the measured accelerations can also be used to estimate the spacecraft thrust, and the massflow can be inferred through the motor thrust coefficient, C_F. For both estimates an accurate model of the thermodynamic properties of the combustion gasses and the rate of erosion of the nozzle throat area A^* will be required. Throat erosion is especially problematic for the final descent

motor where the burn exceeds 200 seconds. Any errors associated with the throat erosion rate will directly translate to an error in the associated massflow measurement.

Calculating the total energy curve in real time is a rather complex undertaking, and beyond the scope of the study. However, it must be understood that estimating this feedback parameter is an essential part of the closed-loop landing strategy, and filtering methods for performing this calculation from available data sources must be developed. A comprehensive error analysis and sensitivity study will be required in order to identify the measurement requirements.

Closed-Loop Loiter Control

Of these methods, Loiter, where the vehicle pitch angle, and hence the thrust vector, oscillates at some amplitude about the flight path angle, has the highest technology readiness Level (*TRL*) maturity. It has been previously demonstrated on launch vehicles using solid rocket motors.[24] Loiter works by providing sufficiently more impulse in your propulsion system to ensure that the mission objectives can be accomplished under all reasonable flight conditions, 95 or 99% confidence levels are typically used. By monitoring the current energy level, the loiter wastes some of this extra energy depending upon the current energy state.

For example, assume that the vehicle control system prescribes a loiter angle where the vehicle precesses at some prescribed cone angle about the direction of flight. The coning motion keeps the mean flight path and heading angles unchanged while reducing the total impulse delivered along the axial direction. At a maximum offset angle of θ_{max} the ratio of the axial "working" thrust to the total thrust is

$$\frac{F_{axial}}{F_{thrust}} = \cos\theta_{\max} \quad . \tag{12}$$

The loiter control system manages the coning amplitude to keep the vehicle energy near or just below the optimal value as shown by Figure 11(e). Should the vehicle energy level rise above the scheduled value of Fig. 11(e), then the cone angle is nulled and the thrust is directed along the flight path.

Thus, a 15º loiter angle would then deliver approximately 96.6% of the total available impulse along the axial direction. A 25º loiter angle reduces the delivered axial impulse to 90.6%. This loss of "working" impulse is equivalent to a drop in specific impulse, and more propellant mass must be consumed to deliver the same amount of effective total impulse. The required propellant M_{p2} can be calculated based on the original optimal propellant mass M_{p1}, and the initial mass at the beginning of the loiter burn, $M_{initial}$.

$$M_{p_2} = M_{initial}\left(1-\left(1-\frac{M_{p_1}}{M_{initial}}\right)^{\frac{Isp_1}{Isp_2}}\right) \quad . \tag{14}$$

For the previous example, where the initial mass for the final burn is 8.497 kg and the optimal propellant mass consumed was 3,632 grams, and the original Isp_1 was *290 seconds*, then a 25º loiter angle results in an effective RMS Isp_2 of *262.83 seconds.* The resulting increase in the propellant mass required to achieve the same "working" impulse is now 3,905 grams. This change is an increase of 273 grams, or an increase of 7.5% compared to the original propellant mass of 3,632 grams. Thus, the landing payload from the optimal example would be reduced to 4,592

grams, compared to the optimal landing mass of *4,865 grams*. This mass loss does not include any propellant costs that would be associated with the loiter attitude control system and the propellant that would be required to maintain the commanded coning angle. The frequency of the loiter coning oscillation would be determined by the time response of the attitude control system, and the acceptable drift from the mean line of travel during the oscillation period.

Pulse-Burn Control Strategy

Pulse burn energy management and control can be adapted from traditional method used for propulsive spacecraft attitude control. This control would be actuated only during the final descent phase of the mission. Sufficient impulse margin must be allowed in order for the spacecraft to correct a higher-than-nominal energy level up on reaching the final burn initiation point or "high gate." The energy management algorithm uses a balance of potential and kinetic energy to predict the vehicle apogee altitude based on the previously-described estimation filter calculations of velocity, (*kinetic energy per unit mass*), altitude (*potential energy per unit mass*), and spacecraft mass. As the motor burns against the direction of flight, the vehicle energy dissipation is given by

$$\Delta E = m \cdot \int_0^t F_{thrust} \cdot V \cdot dt \quad . \tag{15}$$

If the vehicle energy level currently lies above the Optimal curve as shown by Figure 11(e), the motor is actuated and burns until the energy levels drops below the prescribed energy curve. Once the required energy state is achieved, then thrust is terminated. In order to minimize the number of required propulsion system actuation a "dead-band" ΔE around the desired energy state is allowed. Continuously, along the trajectory the energy state is calculated from the inertial and mass data

$$E = m \cdot \left(\frac{V_r^2 + V_\nu^2}{2} + \frac{\mu_{moon} h}{\left(r_{moon} + h\right) \cdot r_{moon}} \right) , \tag{16}$$

and this value is compared to the optimal energy level, E_{opt}. This dead-band control strategy dead-band strategy is depicted by Figure 12. The dead-band is scheduled as a percentage of the nominal energy level for a given altitude or time of flight. If the current vehicle energy is above the upper actuation limit, shown by the upper dashed on Fig. 12, then thrust is actuated. The thrust continues until the motor drops below the lower actuation limit. Below this limit, shown by the lower dashed line on Figure 12, thrust is terminated. Thrust will not re-activate until the energy level climbs back above the upper limit again. Table 8 shows the dead-band control logic. As the vehicle approaches the landing point the dead-band size drops proportionately in order to gain

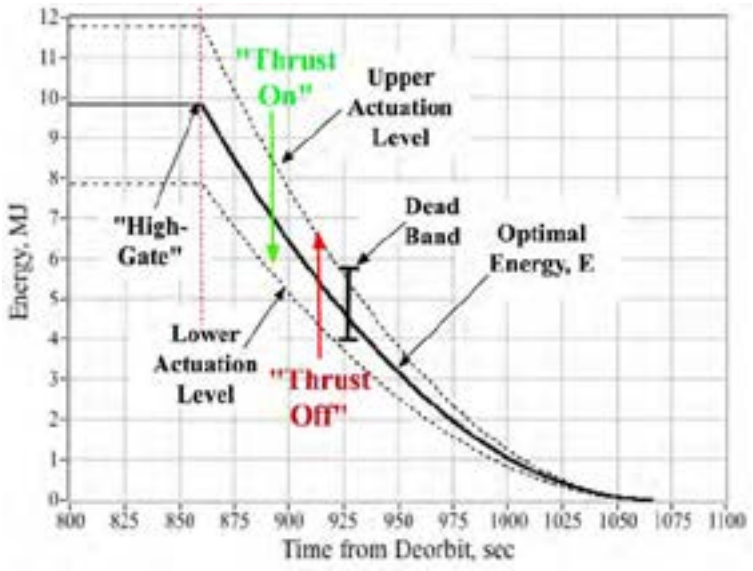


Figure 12. Dead-Band Energy Management Strategy

more precise energy management control. The size of this dead-band will be determined by the startup time, and number of available ignitions for the propulsion system.

Table 8. Energy-Management Dead-Band Control Logic.

```
if (h > h_hi-gate)
→ F_thrust = "off"
else
{
   if (E > E_opt + ΔE)
      → F_thrust = "on"
   else
   {
      if (E > E_opt − ΔE && F_thrust = "on")
         → F_thrust = "on"
      else
      {
         if (E > E_opt − ΔE && F_thrust = "off")
            → F_thrust = "off"
         {
            if (E < E_opt − ΔE)
               → F_thrust = "off"
          }
   }
}
```

Throttle Control

The final option for closed-loop energy management is throttle control. Throttle control is an available option only for liquid bi-propellant, mono-propellant, or hybrid rocket engines. Solid rocket motors cannot be throttled in any reasonable sense, and certainly not in a closed-loop. Casiano, et al[25] present a comprehensive survey of the throttling options and the associated issues for bi-propellant liquid rocket systems.

Operational use of digital electronic engine control for closed-loop throttle jet engines has been in existence for three decades now; however, this technology is still very immature for rocket systems. The deep throttling capability required for this problem, with turn-down ratios greater than 50%, is difficult to achieve for bi-propellant rocket systems. For bi-prop rockets, the upstream injector feed pressures and area ratios directly control the total mass flow, chamber pressure, and fuel oxidizer-to-fuel (*O/F*) ratio. Proper fuel and oxidizer atomization are critical for stable combustion.[26] Maintaining a sufficiently high-pressure drop across the injector for satisfactory atomization sets a practical lower limit to the depth of throttling that can be achieved by pressure modulation alone.

Theoretically, a liquid rocket system can be throttled to any level by lowering the oxidizer and fuel feed pressures upstream of the injector. However, reducing propellant flow rates causes the upstream injector pressure to drop faster than the chamber pressure. At some point the injector pressure ratio (IPR) becomes so low that coupling occurs between the chamber and propellant feed system. As a rule of thumb, a pressure ratio of 1.25 or greater across the injector is required to insure proper combustion stability. This requirement limits the ability to throttle liquid rockets by using feed pressure only, typically 60–70% of the nominal operating thrust level. Because of this pressure ratio limit, simple pressure-fed bi-propellant rockets are almost impossible to throttle deeply.

Other factors also contribute to the complexity of deep-throttle liquid rocket systems. Both the fuel and oxidizer valves are required to precisely maintain near-optimal oxidizer-to-fuel (*O/F*)

ratios over a wide range of propellant mass flow rates. The combustor L^* (the ratio of chamber volume to nozzle throat area) is typically configured for a near-optimal *O/F* ratio. In a deeply throttled engine, small variations in either propellant flow rate can result in a significantly skewed *O/F* ratio, and this off-design condition can interact with the chamber L^* to produce either incomplete combustion or combustion instability. Either case will produce a suboptimal combustor performance compared with the full-throttle motor.

Because of the aforementioned difficulties, deep-throttle liquid rocket engines nearly always require variable geometry injection systems. Injectors for deep-throttled liquid systems are generally based on the TRW pintle injector design.[27] The unique geometry and injection characteristics of the pintle injector distinguish it from injectors typically used on liquid-bipropellant rocket engines. While quite effective in allowing precise engine throttle control, pintle mechanisms are complex and heavy. Bi-propellant pintle engines would require two separate injectors including the logic controller, drive mechanisms, injection ports, and feed systems. The associated pintle-drive power requirements would likely be incompatible with the very small (6-U) spacecraft size being considered here.

In contrast to bi-prop engine, hybrid rocket systems are significantly easier to throttle deeply. When a liquid rocket is throttled, both the oxidizer and fuel mass flow rates can be directly controlled using both the upstream feed pressure and injector flow area. In contrast, when a hybrid motor is throttled, only oxidizer mass flow rate is directly controlled and the fuel mass flow rate is an indirect response to the change in oxidizer mass flow. The fuel mass flow rate is driven primarily by the rate of fuel pyrolysis and the continuously variable fuel port surface burn area. As the oxidizer flux to the motor is reduced, the fuel pyrolysis rate also drops, and the motor naturally adjusts the O/F ratio to stay well within an operable range.[28]

Whitmore et al (2014, 1)[29] have demonstrated the successful turn down of a nitrous oxide, Hydroxyl-Terminated Poly Butadiene (HTPB) hybrid rocket motor from 800 N to less than 12 N, a 67:1 turndown ratio. At throttle levels approaching 20% of nominal, the nitrous oxide exiting the throttle valve is entirely in a vapor state. The vapor chokes the injector and eliminates feed system coupling. In addition to the self-compensating *O/F* ratio described in the previous paragraph, this two-phase effect is another likely reason for the unexpected combustion stability observed at very deep-throttle levels.

Closed-Loop throttle-control allows energy to be managed through a continuous tracking process. Multiple control options including commonly-used Linear Quadratic Tracking (*LQT*),[30] and Proportional Integral Derivative (*PID*)[31] control techniques are available. Previously, Whitmore et al (2014, 2),[32] and Whitmore and Spurrier,[33] have demonstrated close-loop control hybrid systems using PID control. Ref. [32] throttled an 800-N Nitrous Oxide /HTPB hybrid using PID control with chamber pressure feedback to track prescribed chamber pressure profiles. Ref. [33] achieved closed loop-throttle on a 900 N hybrid rocket using Gaseous oxygen and additively manufactured acrylonitrile-butadiene-styrene (ABS) as the propellants. For both motors closed-loop throttling was achieved using an expensive commercial off-the-shelf ball-valve and actuator hardware. The ball-valve provided a fast response variable area flow-path that would choke at higher turn-down levels. The choked ball valve directly modulated the oxidizer flow through the variable flow area, and flow choking aided in decoupling any combustor pressure feedback to the upstream oxidizer flow path.

Multiple and open- and closed-loop tests were performed to demonstrate that closed-loop control can significantly reduce the run-to-run burn variability typical of hybrid rocket motors. Closed-loop proportional/integral control algorithms featuring thrust or pressure feedback are used

to track prescribed step and linear ramp profiles. Control law gains were tuned a priori using a numerical model and then adjusted using the actual test hardware. Response profiles were optimized according the integral absolute error criterion. Test results indicate that, to a 95% confidence level, closed-loop throttling significantly reduces the mean run-to-run thrust variability from 9.1 to less than 1%.

Throttling control offers an elegant solution for the previously-described energy tracking problem. Throttling offers the ability to optimize the descent trajectory, and maximize the payload delivery to the lunar surface. There are, however, packaging and reliability issues associated with applying approach for bi-propellant systems. Hybrids offer a good packaging solution, and closed-loop throttle has been previously achieved with good success during ground tests. Unfortunately, to date no operational rocket system with closed-loop throttle has ever been flown, and the TRL is relatively low. Small throttled hybrid offers a promising "green" solution, but significant development will be required before a throttled system can be deployed for the proposed Lunar Scout Mission.

M. Proposed Propulsion Solution

As a follow-up to this preliminary investigation it is proposed to design, develop, and test a novel, 3-D printed hybrid propulsion solution, optimized for SmallSat Lunar Landing missions. The key enabling technology for this mission, the High-Performance "Green" Hybrid Propulsion (*HPGHP*) system, has been under development at the Propulsion Research Laboratory at Utah State University for nearly decade. In March 2018 a flight experiment containing a 10-N *HPGHP* prototype was launched aboard a Terrier-Malemute sounding rocket from NASA Wallops Flight facility. During the mission the thruster was successfully fired 5 times at apogee of 172 km. Thus, the *HPGHP* system has an established spaceflight heritage.

HPGHP leverages the unique electrical breakdown characteristics of additively-manufactured plastics like ABS, polystyrene, and polyamide. Additive manufacturing changes the electrical breakdown properties of these materials and is an essential element of the *HPGHP* technology. When FDM printed materials are subjected to an inductive charge, an electrical-arc is produced along the layered surface, pyrolyzing material and seeding combustion with the introduction of an oxidizing flow. The system has been engineering to a high level of reliability and the number of possible ignitions is limited only by the amount of fuel present. Typical startup sequences consume less than 5 watts of power for less than 1/4 second, consuming less than 25 joules of total energy. No preheat is required. Once started, the system can be rapidly fired multiple times with additional energy inputs required.

Previous tests of a prototype SmallSat system have demonstrated ignition latencies less than 100 ms. Repeatable minimum impulse bits of less than 0.25 N have been demonstrated by prototype systems during both ambient and vacuum ground tests. The ability to perform essentially an unlimited number of restarts allows this system to precisely manage the spacecraft energy state during descent to the lunar surface; a capability not available to solid-propulsion systems, or systems using ionic liquid propellant that exhibit considerably higher latencies.

The proposed system uses 3-D printed acrylonitrile butadiene styrene (ABS) or polyamide as the fuel material; with "nytrox," a solution of nitrous-oxide and gaseous oxygen, as the oxidizer. The solution of nitrous oxide into the nitrous oxide essentially eliminates the potential for an uncontrolled decomposition reaction; and makes the system inherently safe, with a low storage pressure. The propellants are environmentally friendly, and ABS is 100% recyclable. The

propellants are space storable, and actually demonstrate better performance under near-cryogenic conditions. Published results from round tests have demonstrated that specific impulse values greater than 300 seconds achievable under vacuum conditions. Propellants based on ionic liquids, due to their very large preheat requirement; are unsuitable for deep-space missions where near-cryogenic temperatures are encountered. Ionic liquid propellants can achieve specific impulse values only as high as 235 seconds. Closed-Loop throttling of Nitrous-Oxide based hybrid motors has been previously demonstrated (Ref [32]). Throttling has never been successfully demonstrated using ionic liquid propellants.

It is proposed to design, package, and ground-test a 50 N HPGHP system into a 6-U "MeatBall" form factor. On-demand ignition reliability will be demonstrated under both ambient and vacuum conditions. This project will demonstrate the capability to fire this unit in pulsed mode so as to track a prescribed energy profiles necessary for both lunar landing and sample return missions. Throttle and loiter control options will also be investigated. Hardware in the loop (HIL) simulations with the actual propulsive hardware are proposed. Early, ambient tests of the system will be performed at Utah State, and following successful demonstration of HIL tests; the system will be move to NASA MSFC for both ambient and vacuum chamber verification testing. Ground tests proposed for this project will use lab-weight motor components; however, trade studies are proposed to identify properly sized and compatible flight weight, space qualified components that world be used to build the actual spaceflight hardware.

IV.Conclusion

This study has assessed mission concepts low-budget lunar “scout” missions that allow for in-situ exploration of lava tube "skylight" surface openings. The primary objective is inexpensive and repeatable missions to investigate potential lava tube openings that may provide structures that are suitable for development as human habitats. If such missions were enabled, the collected in-situ database can be used to scrub the list of potential skylight targets, substantially enhancing success probabilities of more comprehensive and costly follow-on exploration missions.

This study investigated the capabilities of 4 emerging non-defense, commercial launch vendors for sending small payloads into Lunar Transfer Orbits, and has sized the final kick stage based on available apogee kick motors available from a prime defense contractor. A "cataloged" motor was selected for the Lunar Capture Stage in order to eliminate the associated non-recurrent engineering and development costs for this mission. Following a trade study where 4 primary factors 1) usable mass to low lunar orbit, 2) final orbit deviation from a baseline 200 km circular orbit, 3) payload cost per kilogram to lunar orbit, and 4) launch system development and technology maturity; Vector Space and its Terran-1 launch vehicle was down-selected as the primary option.

The Terran-1 is capable of delivering up to 163 kg of usable mass into low lunar orbit. This significant mass to lunar orbit offer the possibility of an orbiting Scout platform that can carry up to 10, bowling-ball sized, 10-kg "MeatBall" landers while still leaving over 60 kg of residual mass for the platform structure and avionics. The option of having redundant landers available significantly reduces mission risk; and should the first landing attempt be successful, the second lander provides the option to explore another surface target that lies along the orbital track. These small spacecraft would carry only the necessary avionics and optical systems necessary to “vet” the surface feature for future human habitat development.

Starting from a nominal 200 km altitude lunar orbit, this study concludes that approximately 50% of the on-orbit mass is deliverable to the lunar surface. The initial de-orbit stage is based on a small *cataloged* solid-motor that was previously developed for SmallSat orbit insertion. This

motor provides sufficient impulse to send the MeatBall payload onto a shallow lunar intercept that tracks approximately 12km downrange. This study concludes however, that an unregulated solid propulsion system with a set a-priori total propellant mass provides insufficient impulse precision for reliable mission outcomes. Clearly, some sort of loop closure is necessary during landing in order to ensure a reasonable degree of mission success. Use of "loiter" where the vehicle cones about the velocity vector, or a throttleable/restartable final descent stage operating in pulsed mode, coupled with closed-loop energy management, is likely mandatory for this mission. Top-level control solutions are evaluated for each of these strategies. solutions include loiter, pulsed burning, and throttle. Of the proposed solutions, loiter has the highest technology readiness level, but also costs the most in terms of propellant usage.

Finally, a potential propulsion solution is proposed based on High-Performance "Green" Hybrid Propulsion system, has been under development at the Propulsion Research Laboratory at Utah State University for nearly decade. The proposed system uses 3-D printed acrylonitrile butadiene styrene or polyamide as the fuel material; with "nytrox," a solution of nitrous-oxide and gaseous oxygen, as the oxidizer. The solution of nitrous oxide into the nitrous oxide essentially eliminates the potential for an uncontrolled decomposition reaction; and makes the system inherently safe, with a low storage pressure. The propellants are space storable and deliver up to 300 seconds of vacuum specific impulse.

It is proposed to design, package, and ground-test a 50 N system into a 6-U "MeatBall" form factor. On-demand ignition reliability will be demonstrated under both ambient and vacuum conditions. This project will demonstrate the capability to fire this unit in pulsed mode so as to track a prescribed energy profiles necessary for both lunar landing and sample return missions. Throttle and loiter control options will also be investigated. Hardware in the loop simulations with the actual propulsive hardware are proposed.

Acknowledgments

The author wishes to Thank the NASA Marshall Space Flight Center Office of Stem Engagement and Dr. Frank Six for Sponsoring this work through the Summer Faculty Fellowship program. I also want to thank Dr. Gerald Karr and Brooke Graham of the University of Alabama, Huntsville for facilitating the no-Cost extension that allowed the completion of this report.

References

[1] Wagner, R. V., and Robinson, M. S., "Distribution, formation mechanisms, and significance of lunar pits", Icarus, *No. 237, July 15, 2014, pp. 52–60. http://dx.doi.org/10.1016/j.icarus.2014.04.002.*

[2] Arya, A. S., Rajesekhar, R. P., Thangjam, Ajai, K., and Kiran, K. S., "Detection of potential site for future human hab*itability on the Moon using Chandrayaan-1 data", Current Science, Vol. 100 No. 4*, Feb. 25, 2011, pp. 524-529. https://www.currentscience.ac.in/php/toc.php?vol=100&issue=04.

[3] Nozette, S., Lichtenberg, C. L., Spudis, P, Bonner, R., Ort, W., Malaret, E., Robinson, M., and Shoemaker, E. M., "The Clementine Bistatic Radar Experiment," *Science*, Vo. 274, No. 5292, November 29, 1996, pp. 1495-1498. *https://science.sciencemag.org/content/274/5292/1495.full.pdf.*

[4] Lakdawalla, E. S., "LCROSS Lunar Impactor Mission, "Yes, We Found Water," *J. Planetary Society*, Nov. 13, 2009, *http://www.planetary.org/news/2009/1113_LCROSS_Lunar_Impactor_Mission_Yes_We.html,* [Retrieved 29 August, 2019]

[5] Whittaker, W., "Technologies Enabling Exploration of Skylights, Lava Tubes, and Caves, Final Report, NIAC Phase I, Contract No. NNX11AR42G, September 2012. *https://www.nasa.gov/pdf/718393main_Whittaker_2011_PhI_Cave_Exploration.pdf,* [Retrieved 29 August, 2019]

[6] Ximenes, S., Elliot, J. O., and Bannova, O., "Defining a Mission Architecture and Technologies for Lunar Lava Tube Reconnaissance," *Thirteenth ASCE Aerospace Division Conference on Engineering, Science, Construction, and Operations in Challenging Environments, and the 5th NASA/ASCE Workshop On Granular Materials in Space Exploration*, April 2012. *https://doi.org/10.1061/9780784412190.038*.

[7] *Kerbarm L., "Moon Diver Mission Concept - Descending into a Moon Cave to Better Understand the Solar System's Largest Volcanic Eruptions," Online Journal, Keck Institute for Space Studies, http://www.kiss.caltech.edu/lectures/2018_Moon_Diver.html, [Retrieved 29 August, 2019]*

[8] Dvorsky, G. "Could This Lunar Cave Provide Shelter for a Future Moon Colony?," *GizModo*, October 18, 2014, *https://io9.gizmodo.com/these-intriguing-lunar-caves-could-provide-shelter-for-1607153182*, [Retrieved 29 August, 2019]

[9] Anon., *Space Launch Report:* Rocket Lab Electron Data Sheet," *http://www.spacelaunchreport.com/electron.html,* [Retrieved 29 August, 2019]

[10] Anon., "Vector-H, Go Where You Want, When You Want," *https://vector-launch.com/vector-h*/, [Retrieved 15 September, 2019].

[11] Anon., "From Raw Material to Flight in Less than 60 days," *https://www.relativityspace.com/Terran*, [Retrieved 15 September, 2019].

[12] Anon., "Gilmour Space, "*https://www.gspacetech.com/launch,"* [Retrieved 29 August, 2019].

[13] Clark, S., " "NASA, Rocket Lab partner on successful satellite launch from New Zealand," *Spaceflight Now*, December 17, 2018, *https://spaceflightnow.com/2018/12/17/nasa-rocket-lab-partner-on-successful-satellite-launch-from-new-zealand/*, [Retrieved 29 August, 2019]

[14] McDuling, J., "Blackbird shoots for the moon and Mars with Gilmour Space investment," *Financial Review*, May 29, 2017, *https://www.afr.com/technology/blackbird-shoots-for-the-moon-and-mars-with-gilmour-space-investment-20170526-gwe9rq*, [Retrieved 29 August, 2019].

[15] Sellers, J. J., Astore, W. J., Giffen R. B., and Larson W. J., "Understanding Space: An Introduction to Astronautics," 3rd edition, McGraw-Hill, 2003, pp. 143, 168-172, *ISBN-13: 978-0073407753, ISBN-10: 0073407755.*

[16] Spurmann, J., "Lunar Transfer Trajectories," DLR Document TN 10-02, V1.0, Feb. 03, 2010, *Space Flight Technology, German Space Operations Center (GSOC), ftp://dim13.org/pub/doc/TN_1002.pdf.* [Retrieved 15 September, 2019].

[17] Anon., "The Value of Performance, Northrop Grumman Propulsion Products Catalog," OSR No. 16-S-1432, June 2018. *https://www.northropgrumman.com/Capabilities/PropulsionSystems/Documents/NGIS_MotorCatalog.pdf.* [Retrieved 15 September, 2019].

[18] Dvorsky, G. "Could This Lunar Cave Provide Shelter for a Future Moon Colony?," GizModo, October 18, 2014, https://io9.gizmodo.com/these-intriguing-lunar-caves-could-provide-shelter-for-1607153182, [Retrieved 29 August, 2019]

[19] Purucker, M. E., "E. and Nicholas, J. B., "Global spherical harmonic models of the internal magnetic field of the Moon based on sequential and coestimation approaches," *Journal of Geophysical Research*, Vol. 115, 2010, pp. 401–409. https://doi.org/10.1029/2010JE003650

[20] Penrod, J. A., "Feasibility of Using Lunar Magnetic Fields to Control CubeSat Attitude," B.S. Thesis, *University of Colorado, Department of Aerospace Engineering*, 2015. https://hanspeterschaub.info/Papers/grads/JamesPenrod.pdf. [Retrieved 29 August, 2019]

[21] Mehrparvar, A., "CubeSat Design Specification, Rev. 13, *The CubeSat Program, Cal Poly SLO*, https://static1.squarespace.com/static/5418c831e4b0fa4ecac1bacd/t/56e9b62337013b6c063a655a/1458157095454/cds_rev13_final2.pdf, [Retrieved 15 September, 2019].

[22] Anderson, J. D., *Modern Compressible Flow*, 3rd Edition, New York: The McGraw Hill Companies, Inc., 2003, Chapter 4, pp. 127-187. ISBN-13: 978-0072424430. https://libcat.lib.usu.edu/search/i0070016542

[23] Whitmore, S. A., "MAE 5540 Lecture Notes, Chapter 3," Published Online, http://mae-nas.eng.usu.edu/MAE_5540_Web/propulsion_systems/section3/section3.2.pdf, Slide 57. [Retrieved 15 September, 2019].

[24] Betts, J. T., "Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, vol. 21, no. 2, pp. 193–207, 1998. https://doi.org/10.2514/2.4231

[25] Casiano, M. J., Hulka, J. R., and Yang, V., "Liquid-Propellant Rocket Engine Throttling" A Comprehensive Review," *Journal of Propulsion and Power*, Vol. 26, No. 5, Sept-Oct 2010, pp. 897-912. https://doi.org/10.2514/1.49791

[26] Betts, E. M., and Frederick, R., "Historical Systems Study of Liquid Rocket Engine Throttling Capabilities," AIAA Paper 2010-0863, *46th AIAA/ASME/SAW/ASEE Joint Propulsion Conference*, Nashville TN, July 2010. https://doi.org/10.2514/6.2010-6541

[27] Dressler, G. A., and Bauer, J. M., "TRW Pintle Engine Heritage and Performance Characteristics," AIAA Paper 2000-3871, *36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, Las Vegas, NV, July 2000. https://doi.org/10.2514/6.2000-3871

[28] Waidmann, W., "Thrust Modulation in Hybrid Rocket Engines," *Journal of Propulsion and Power*, Vol. 4, No. 5, 1988, pp. 421–427. https://doi:10.2514/3.23083

[29] Whitmore, S. A., Peterson, Z. W., and Eilers, S. D., "Deep Throttle of a Nitrous Oxide Hydroxyl-Terminated Poly Butadiene Hybrid Rocket Motor, *Journal of Propulsion and Power*, Vol. 30, No. 1, Jan-Feb. 2014, pp. 78-86. https://doi:10.2514/3.34967

[30] Catlin, D. E., "The Linear Quadratic Tracking Problem," *Applied Mathematical Sciences Estimation, Control, and the Discrete Kalman Filter*, Vol. 71, Springer Link Publishing, 1989, pp. 164-187. https://doi.org/10.1007/978-1-4612-4528-5_8

[31] Araki, M., "PID Control," *Encyclopedia of Life Support Systems (EOLSS), Control Systems, Robotics, and Automation,* Vol. II, 2009, pp. 58-79. http://www.eolss.net/ebooks/Sample%20Chapters/C18/E6-43-03-03.pdf.

[32] Whitmore, S. A., Peterson, Z. W., and Eilers, S. D., "Closed-Loop Throttling of a Hybrid Rocket Motor," *Journal of Propulsion and Power*, Vol. 30, No. 2, March-April, 2014, pp. 78-86. https://doi:10.2514/1.B34294

[33] Whitmore, S. A., Spurrier, Z. S., "Throttled Launch-Assist Hybrid Rocket Motor for an Airborne NanoSat Launch Platform", *51st AIAA/SAE/ASEE Joint Propulsion Conference, AIAA Propulsion and Energy Forum*, (AIAA 2015-3940) Orlando FL USA, July 2015. http://dx.doi.org/10.2514/6.2015-3940

Technology Transfer and Technical Writing at NASA/MSFC

[1]Tammy S. Winner, Ph.D.

University of North Alabama, Florence, AL, 35630

[2]Terry L. Taylor, Manager, Technology Transfer Office

NASA, Marshall Space Flight Center, Huntsville AL

I. Introduction and Background

The following white paper will report the findings of a 10 week Faculty Fellowship during the summer of 2019 at the NASA Marshall Space Flight Center (MSFC) in Huntsville, Alabama. I was accepted to participate in this program by Dr. Frank Six and placed under the supervision of Terry Taylor, Chief of the Technology Transfer Team at MSFC. The title of this ongoing research is "Technology Transfer Technical Writing." The goal of the Faculty Fellowship program at MSFC is threefold, as stated in the handbook, "…to provide research experience in current NASA projects, to learn about Marshall's role in space exploration, and to understand the partnership between NASA, the private sector and academia." The following report will articulate how I attempted to achieve the stated goals of the program during my time as a Faculty Fellow and suggest future research opportunities.

Within the National Aeronautics and Space Act of 1958, a directive was issued to "provide for the widest practicable and appropriate dissemination of the results of its missions." (NASA, 203). And, in 1962 the Technology Transfer Program was established. It is the longest-running mission to date at NASA. Numerous examples can be cited to support the missions' efforts to make the American tax payer aware of the ways in which space exploration contributes to a better life here on Earth. While operating as a Faculty Fellow within the Technology Transfer Team I attended meetings and seminars, wrote awards nominations, organized materials for a 12-month trade show calendar in collaboration with shipping and receiving, interviewed eight of the 15 team members on their roles and contributions to the mission, and learned about the yearly publication generated by the Technology Transfer Team, *Spinoff*. I was also introduced to the various Feld center websites, software catalog, licensing systems, and patent processes. But why would a Professor of Writing want to spend a summer at NASA's MSFC?

Good question. I earned my Ph.D. in Rhetoric and Composition with a concentration in Technology and Writing in 1999 from Indiana University of Pennsylvania. For the past twenty

[1] Associate Professor, Dept. of English/Professional Writing, University of North Alabama.
[2] Manager, Technology Transfer Office, NASA/MSFC, Huntsville, AL

years I have been studying the ways in which technology is a writing tool for all writers, from all places. In 1997, as a graduate student, I attended a workshop at NASA's Classroom of the Future in Wheeling West Virginia. Two decades later, NASA is still teaching me how to be a better teacher—using technology. My ten-week Faculty Fellowship at MSFC during the summer of 2019 enabled me to conduct research using the following qualitative research methods: floating observations, interviews and focus groups all in an unaltered, naturalistic setting. I kept copious notes, collected numerous artifacts from various departments within the agency and attended seminars, group meetings, tours, and special events in which I was encouraged to take part. It was especially timely to take part in this research during the 50th Anniversary of Apollo. Much excitement surrounded this event and many "good" stories were told this summer that, perhaps, would not have been mentioned had it not been the 50th anniversary of the Lunar Landing. The goal of my naturalistic inquiry was to identify the occupational gaps, if any, that were believed to exist between newly hired NASA employees who were within 2 years of graduating from college.

II. Floating Observations Identify Occupational Gaps and Create New Assignments

One of the many outcomes of my experience as a Faculty Fellow at MSFC was the ability to designing Technical Communication Assignments for both undergraduate and graduate students studying Technical Writing at University of North Alabama. The goals of these various writing intensive tasks are to: tighten the occupational gap some students may face after graduating from UNA and entering the Technical Communication workforce in Huntsville, AL, and strengthen existing partnerships between MSFC and UNA. The Technical Communication assignments and lectures I was able to create in collaboration with MSFC staff are as follows:

- Rhetorical Analysis of a NASA registered US Patent
- Document Design of a TOPS (Technology Opportunity Sheet) based on a US patented technology out of MSFC
- Preparation of White Papers using AIAA Style Guide
- Understanding the Nomenclature of NASA Technical Writing Tasks
- Proposal Writing (undergraduate level)
- Advanced Proposal Writing (graduate level)
- How to research and compose an award nomination for an employee you don't know
- Collaborative Writing Groups using the Fostering Resilient Systems: Anticipate, Monitor, Respond, Learn
- Dense to Condense: Research Poster Design of a Complex Problem
- Top 10 Soft Skills Technical Communicators Must Have in Today's Workforce

These writing assignments and lectures will be piloted during the 2019-2020 academic calendar at UNA by me. Once they are student tested, they will be revised accordingly and shared with other lecturers. Additionally, the results of the floating observations, i.e., "teachable

moments" at NASA/MSFC will be presented at the annual Rocky Mountain Modern Language Association in El Paso, Texas, October, 2019.

III. Strengthening Existing Partnerships between MSFC and UNA via TC Assignments at the Graduate and Undergraduate Level

The 10-week Faculty Fellowship at NASA/MSFC during the summer of 2019 allowed this seasoned Professor of Technical Communication the opportunity to build many educational bridges with various teams throughout the Agency. In addition to attending various lectures, I knocked on many office doors, shook hands, collected business cards and asked the question that served to be the driving force behind my research, "What do new employees entering MSFC right out of college need to know…that they don't know." The answers to this question, It is my hope, will serve to strengthen existing partnerships between MSFC and UNA.

IV. Recommendations and Conclusion

After attending a lecture by Scott Hutchins at MSFC, I am also interested in pursuing a Cooperative Agreement Notice (CAN) with one or more partners at MSFC/NASA. The tentative idea for the CAN is a Technology Transfer Writing Program. The idea would be to work with the office of Technology Transfer to design a series of writing assignments that can be used to educate students within the Alabama Public School System on the ways in which NASA space technology has improved life on Earth. The duality of this assignment would be clearly supported with not only the topics, but the format, style, language and tone of each document/writing sample. It would also continue to support existing relationships between Alabama and NASA.

While the original foci of this research was to improve collaboration and strengthen the existing partnership between UNA and NASA/MSFC, the outcomes proved to be far vaster. As an Associate Professor of Writing at UNA, my time as a Faculty Fellow at NASA/MSFC energized me to renew and rethink my current pedagogy. I found myself questioning the content of the current Technical Writing texts, rethinking my student learning objectives within my syllabi, and giving more credence to the soft skills my students need to possess in order to not just survive, but thrive as a newly hired employee at NASA. Perhaps the hardest question I have to answer, based on my time at NASA/MSFC this summer is – "do we need to require Professional Writing majors at UNA to take a coding class?" The answer, based on my qualitative research is yes. But the questions surrounding English majors being required to take coding classes still remain with regard to "when" and "how much."

Future studies are needed to address, but are not limited to: continuing to identify the lack of certain soft skills needed to transition from the college classroom to the workforce at NASA, understanding the communication preferences among the various age groups and genders working at MSFC, and studying communication gaps between the various generations working at NASA/MSFC. I would like to reapply to the Faculty Fellowship next summer, but I would like

to conduct my research at Headquarters or Goddard Space Flight Center (GSFC). It is my belief that taking these observations to another geographic location within the Agency will further add to the existing data.

Forward Joint of Payload Adaptor for the Space Launch System

2019 NASA Marshall Space Flight Center Summer Faculty Fellowship Final Report

Charles Yang[1]
Wichita State University, Wichita, KS 67260

Majid Babai[2], William E. Guin[3], *and* Justin R. Jackson[4]
NASA Marshall Space Flight Center, Huntsville, AL 35811

A composite Payload Adaptor (PLA) is being designed and fabricated at NASA Marshall Space Flight Center for the Block 1B heavy-lift Space Launch System (SLS). Functions as the primary structural interface between the payload and the body of the lunch vehicle, the PLA is shaped as a frustum of a cone and is connected to the payload at the forward end and the launch vehicle at the aft end. The objective of this study is to design the pinned joints at the forward end by estimating the bearing capacity of each pinned joint and the load transfer at the forward end of the PLA. Finite element models with progressive failure analysis as well as experiments were conducted to predict the load carrying capacity of the pinned joints. Three-dimensional finite element models were developed in ABAQUS/Standard. User subroutines of three sets of failure criteria were implemented to capture the material nonlinearity and property degradations due to damage. Different composite facesheet layups were included in this investigation.

Nomenclature

d = index of lamina shear damage/nonlinearity
e_m = index of matrix failure
e_f = index of matrix failure
e_{fm} = index of fiber/matrix shear failure
E_x = lamina Young's modulus in the fiber direction
E_y = lamina Young's modulus in the resin direction
E_z = lamina Young's modulus in the out-of-plane direction
h = shell element thickness
G_{xy}, G_{xy}, G_{xy}, = lamina shear moduli
X_c = lamina compressive strength in the fiber direction
X_t = lamina tensile strength in the fiber direction
Y_c = lamina compressive strength in the resin direction
Y_t = lamina tensile strength in the resin direction
α = coefficient of lamina shear nonlinearity
γ_{xy}, γ_{xy}, γ_{xy}, = lamina shear strains
ν_{xy}, ν_{xy}, ν_{xy}, = Poisson's ratios
σ_x = lamina stress in the fiber direction
σ_y = lamina stress in the resin direction
σ_z = lamina stress in the out-of-plane direction

[1] Professor, Department of Aerospace Engineering, MSFC Faculty Fellow
[2] Advanced Manufacturing Chief, EM 42, MSFC Collaborator
[3] Materials Engineer, EM 42
[4] Materials Engineer, EM 42

I. Introduction

Advanced composites have been the "material of choice" in aerospace industries in the past decades due their high strength-to-weight ratio compared to traditional metallic materials. Research and development efforts as well as experience have gradually pushed the applications of advanced composite materials from secondary, non-essential parts such as cosmetic aircraft interior panels to primary load-carrying structures such as the all-composite fuselage of Boeing 787 and Airbus A350, just to name a few. NASA's intension to take advantage of the benefits from advanced composite materials can be seen from the Composite Crew Module (CCM) constructed from 2006 for the Constellation Program Crew Exploration Vehicle. Even though the CCM was not actually adopted for any space exploration projects, the technologies industry gained have assisted in future composites applications.

NASA's heavy-lift Space Launch System (SLS) program is aiming at sending astronauts and/or payloads to cis-lunar space as well as to Mars and other deep-space destinations in the 2020s. The SLS is currently tasked to land astronauts on the moon by 2024 and maintain US presence on the moon by 2028. Part of the new flight hardware development of the SLS program is the composite sandwich Payload Attach Fitting (PAF). As shown in Fig. 1, the PAF is a cone-shaped structure used to connect the payload to the launch vehicle. The main function of the PAF is to support the weight and inertia force during launch and deployment of the payload. The PAF is joined to the payload at the forward end and to the launch vehicle body at the aft end via metal fittings. Figure 2 shows the schematic of a metal clevis fitting at the forward end. The purpose of this study is twofold: (1) investigate the load carrying capacity of pinned joints of composite sandwich panels, as shown in Fig. 3 and (2) design the PAF panel and the joints by determining the section force and moment distributions on the PAF under various load cases.

II. Pinned Joint Failure Load of Composite Sandwich Panels

While there has not been much research work conducted on pin-bearing or bolt-bearing damage of composite honeycomb sandwich panels, pin- and bolt-bearing damage on solid composite laminates have been studied in depth [1-14]. Assume the foam core in the PAF panel does not take significant bearing load, the two composite facesheets of the sandwich panel are bearing all the load transferred in a pinned joint. Therefore, the pin-bearing capacity of the facesheets which are two solid composite laminates is used to determine the load carrying capacity of the pinned joints of composite honeycomb sandwich panels. Two methods were used in the current study: (A) Tsai-Hill first-ply failure theory was applied directly to the bearing stress and (B) three-dimensional progressive failure analysis using finite element methods.

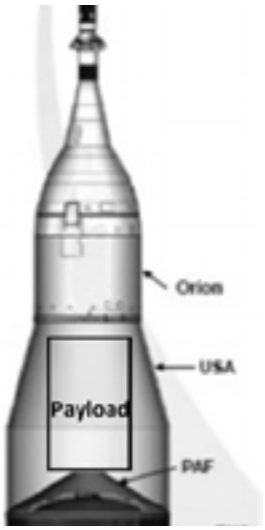


Figure 1. Location of PAF.

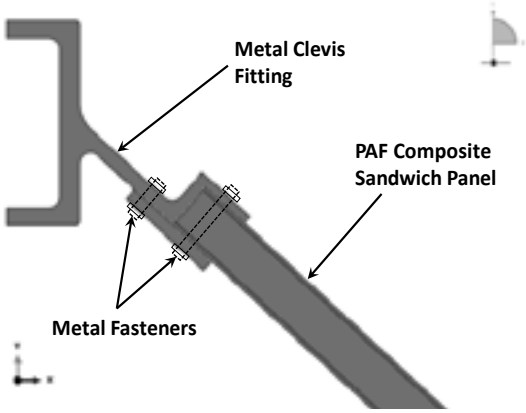


Figure 2. Metal clevis fitting.

A. Tsai-Hill First-Ply Failure Criterion

The application of first-ply theory was published and suggested by Chamis in a NASA Technical Memorandum

[14] in 1988. He suggested a simplified equation for estimating the load carry capacity of a pinned or bolted joint against bearing failure as

$$F_{failure} = \sigma_{c,allowable} dt \tag{1}$$

where F_{failure} is the applied compressive force at fastener hole at bearing failure, $\sigma_{\text{c,allowable}}$ is the allowable compressive stress of the composite laminate, d is the fastener diameter, and t is the laminate thickness. In order to determine $\sigma_{c,allowable}$ of the two composite laminates, finite element models of a square composite panel were constructed using the commercial finite element software MSC Patran/Nastran which features Tsai-Hill first-ply failure criterion. The finite element model of 10×10 two-dimensional composite shell elements, as shown in Fig. 4, with x-directional constraint on the left edge and a line load on the right edge was used for this purpose. $\sigma_{\text{c,allowable}}$ is determined once the Tsai-Hill criterion is satisfied. The joint bearing load F_{failure} was then obtained by Equation (1) with the determined $\sigma_{\text{c,allowable}}$.

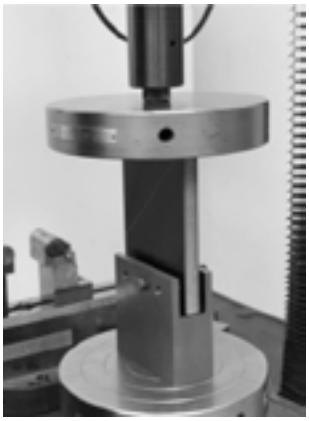


Figure 3. Pinned joint test.

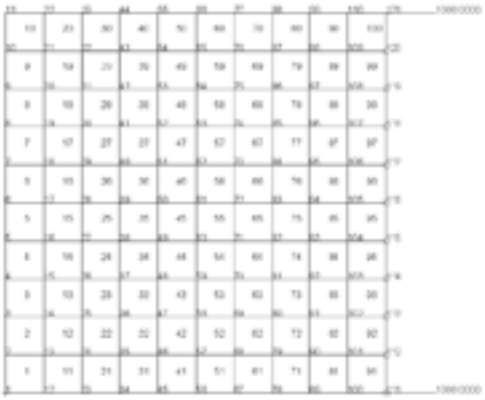


Figure 4. MSC Patran/Nastran model of a square laminate.

B. Three-Dimensional Finite Element Models with Progressive Failure Analysis

Hashin failure criteria [15] were expanded to cover all three directions [16] and was used to predict the failure load of pinned composite joints. Shear nonlinearity in all three directions were included in the approach. Besides the shear nonlinearity in three directions, the other three failure modes are matrix failure, fiber failure, and fiber/matrix shearing failure.

(1) Shear Nonlinearity: The shear nonlinearity is described as

$$\gamma = \frac{\sigma}{G} + \alpha\sigma^3 \tag{2}$$

for each of the x-, y- and z-directions, where α is the nonlinear coefficient. Chang and Lessard [17] determined α to be $0.8\text{x}10^{-14}$ for T300/976 composite lamina for the shear stress/strain relationship in the x-y plane. If the damage parameter d's are used, the shear moduli in the degraded form become $G_{xy} = (1\text{-}d_{xy})\ G_{xy,\text{original}}$, $G_{xz} = (1\text{-}d_{xz})\ G_{xz,\text{original}}$, and $G_{yz} = (1\text{-}d_{yz})\ G_{yz,\text{original}}$ and d_{xy}, d_{xz} and d_{yz} can be derived as

$$d_{xy} = \frac{3\alpha G_{xy}\sigma_{xy}^2 - 2\alpha\sigma_{xy}^3/\gamma_{xy}}{1 + 3\alpha G_{xy}\sigma_{xy}^2} \tag{3}$$

$$d_{xz} = \frac{3\alpha G_{xz}\sigma_{xz}^2 - 2\alpha\sigma_{xz}^3/\gamma_{xz}}{1 + 3\alpha G_{xz}\sigma_{xz}^2} \tag{4}$$

$$d_{yz} = \frac{3\alpha G_{yz}\sigma_{yz}^2 - 2\alpha\sigma_{yz}^3/\gamma_{yz}}{1 + 3\alpha G_{yz}\sigma_{yz}^2} \tag{5}$$

(2) Matrix Failure: With the failure index e_m defined for matrix failure if greater than 1, when $\sigma_y > 0$ or $\sigma_z > 0$,

$$e_m^2 = \left(\frac{\sigma_y}{Y_t}\right)^2 + \frac{2\sigma_{xy}^2/G_{xy} + 3\alpha\sigma_{xy}^4}{2S_{xy}^2/G_{xy} + 3\alpha S_{xy}^4} + \frac{2\sigma_{yz}^2/G_{yz} + 3\alpha\sigma_{yz}^4}{2S_{yz}^2/G_{yz} + 3\alpha S_{yz}^4} \tag{6}$$

or

$$e_m^2 = \left(\frac{\sigma_z}{Y_t}\right)^2 + \frac{2\sigma_{xz}^2/G_{xz} + 3\alpha\sigma_{xz}^4}{2S_{xz}^2/G_{xz} + 3\alpha S_{xz}^4} + \frac{2\sigma_{yz}^2/G_{yz} + 3\alpha\sigma_{yz}^4}{2S_{yz}^2/G_{yz} + 3\alpha S_{yz}^4} \tag{7}$$

when $\sigma_y < 0$ or $\sigma_z < 0$,

$$e_m^2 = \left(\frac{\sigma_y}{Y_c}\right)^2 + \frac{2\sigma_{xy}^2/G_{xy} + 3\alpha\sigma_{xy}^4}{2S_{xy}^2/G_{xy} + 3\alpha S_{xy}^4} + \frac{2\sigma_{yz}^2/G_{yz} + 3\alpha\sigma_{yz}^4}{2S_{yz}^2/G_{yz} + 3\alpha S_{yz}^4} \tag{8}$$

or

$$e_m^2 = \left(\frac{\sigma_z}{Y_c}\right)^2 + \frac{2\sigma_{xz}^2/G_{xz} + 3\alpha\sigma_{xz}^4}{2S_{xz}^2/G_{xz} + 3\alpha S_{xz}^4} + \frac{2\sigma_{yz}^2/G_{yz} + 3\alpha\sigma_{yz}^4}{2S_{yz}^2/G_{yz} + 3\alpha S_{yz}^4} \tag{10}$$

where S_{xy} and S_{xz} are the shear strengths in the x-y and x-z planes, respectively.

(3) Fiber Failure: With the failure index e_f defined for fiber failure if greater than 1, when $\sigma_x > 0$,

$$e_f^2 = \left(\frac{\sigma_x}{X_t}\right)^2 + \frac{2\sigma_{xy}^2/G_{xy} + 3\alpha\sigma_{xy}^4}{2S_{xy}^2/G_{xy} + 3\alpha S_{xy}^4} + \frac{2\sigma_{xz}^2/G_{xz} + 3\alpha\sigma_{xz}^4}{2S_{xz}^2/G_{xz} + 3\alpha S_{xz}^4} \tag{11}$$

when $\sigma_x < 0$,

$$e_f^2 = \left(\frac{\sigma_x}{X_c}\right)^2 \tag{12}$$

(4) Fiber/Matrix Shearing Failure: With the failure index e_{fm} defined for fiber/matrix shearing failure if greater than 1, when $\sigma_x < 0$,

$$e_{fm}^2 = \left(\frac{\sigma_x}{X_c}\right)^2 + \frac{2\sigma_{xy}^2/G_{xy} + 3\alpha\sigma_{xy}^4}{2S_{xy}^2/G_{xy} + 3\alpha S_{xy}^4} + \frac{2\sigma_{xz}^2/G_{xz} + 3\alpha\sigma_{xz}^4}{2S_{xz}^2/G_{xz} + 3\alpha S_{xz}^4} \tag{13}$$

Table 1 shows the degraded material properties after each of the failure criteria is met according to Olmendo and Santiuste's study [16]. As can be seen, constant knock-down factors are used to the moduli once the related failure modes are activated.

Table 1. Material degradation due to failure [16].

Failure Mode	E_x	E_y	E_z	ν_{xy}	ν_{xz}	ν_{yz}	G_{xy}	G_{xz}	G_{yz}
G_{xy} Nonlinearity	E_x	E_y	E_z	ν_{xy}	ν_{xz}	ν_{yz}	1	G_{xz}	G_{yz}
G_{xz} Nonlinearity	E_x	E_y	E_z	ν_{xy}	ν_{xz}	ν_{yz}	G_{xy}	1	G_{yz}
G_{yz} Nonlinearity	E_x	E_y	E_z	ν_{xy}	ν_{xz}	ν_{yz}	G_{xy}	G_{xz}	1
Matrix	E_x	$0.4E_y$	$0.4E_z$	0	0	0	G_{xy}	G_{xz}	$0.2G_{yz}$
Fiber	$0.14E_x$	$0.4E_y$	$0.4E_z$	0	0	0	$0.25G_{xy}$	$0.25G_{xz}$	$0.2G_{yz}$
Fiber/Matrix Shearing	E_x	E_y	E_z	0	0	ν_{yz}	$0.25G_{xy}$	$0.25G_{xz}$	G_{yz}

The commercial finite element software ABAQUS/Standard was used for the progressive failure analysis of this study. Eight-noded, three-dimensional reduced integration solid elasticity elements (C3D8R) were used for the models with the progressive damage criteria. Because the geometry is symmetric in the width direction but the stacking sequences of the laminates are asymmetric, finite element models with half of the width and the entire thickness of the laminate were constructed. Figures 5 and 6 show the finite element mesh of the half-model of a 4"x6" composite plate pinned with a 0.375"-diameter steel pin. One element was used in the thickness direction for each ply. Due to symmetry, the nodes on the left edge were applied with symmetry boundary conditions which constrains the displacements in the x-direction. The surface on the back of the laminate was applied with elastic foundation constraint to simulate the response from the core. Constraint equations were applied to the nodes on the upper end of the laminate so that they have the same y-directional displacement to simulate the loading condition of the test setup. Surface contact was applied between the pin hole of the laminate and the steel pin. The nodes within a 0.1786"-diameter circle of the center of the steel pin were fixed for both x- and y-directional displacements. Elastic material property option "Engineering Constants" was selected for the composite plies where values of E_x, E_y, E_z, ν_{xy}, ν_{xz}, ν_{yz}, G_{xy}, G_{xz}, and G_{yz} were supplied. Material degradation models were applied only to the elements around the pin hole on the laminate, as shown in Fig. 7, in order to reduce the computation time assuming failure modes occur only within this region. Six field variables were added to the material properties of these elements to specify the six failure modes. Options of "User Defined Field" and "Depvar" with six dependent variables were selected.

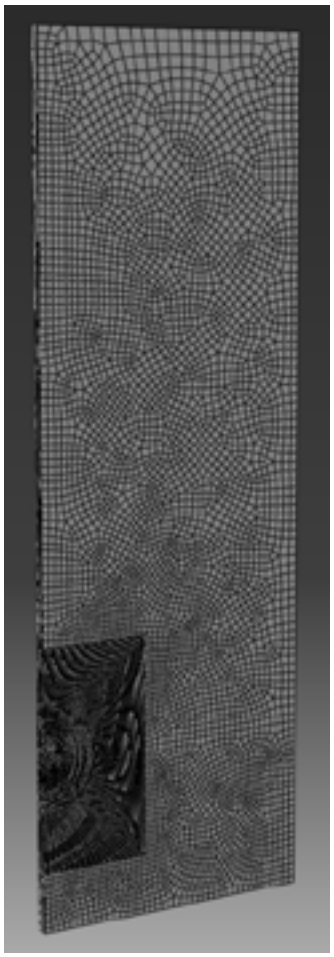


Figure 5. 3D Finite element mesh.

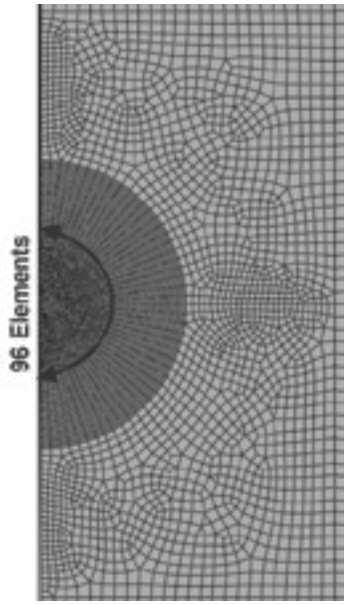


Figure 6. Mesh density around fastener.

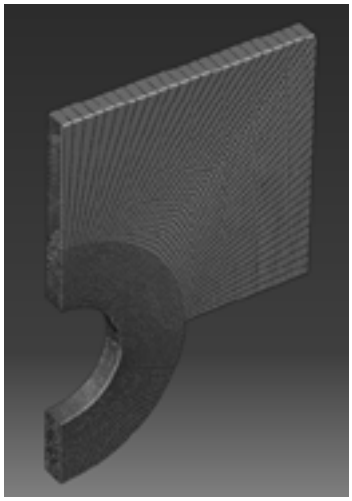


Figure 7. Elements with damage features.

C. Results from Pin Bearing Analyses

All material properties of the composite lamina used in the finite element models were obtained from the literature. Figure 8 shows the typical force history of the models. The program terminated when it failed to converge which means the laminate can no longer take a higher compressive load. The maximum force which the model reached was considered as the bearing load capacity of the laminate.

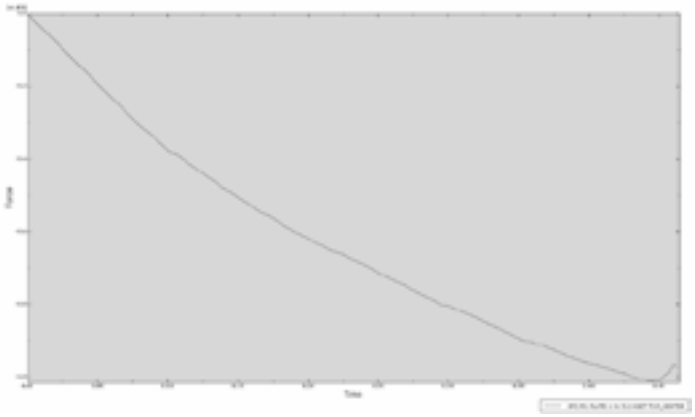


Figure 8. Typical force history.

The deviations of the calculated bearing failure loads from the test data are listed in Table 2 for five different laminate configurations. As can be seen, the first-ply failure theory underestimated the bearing failure load because considerable numbers of 45-deg plies were included in the laminate specimens. The deviations have a large spread which is due to the difference in plies constitution. Under uniaxial compressive loading 45-deg plies do not contribute to the laminate strength as much as 0-deg plies based on the first-ply failure theory. However, they do substantially contribute to pin bearing strength of the laminate. On the other hand, the progressive damage model resulted in much more consistent deviations from the test results, although a more sophisticated process to determine the moduli degradations as functions of extent of damage is necessary to yield more accurate predictions.

Table 2. Comparison between predicted bearing failure load and test data.

Laminate Configuration	Deviation from Test Data	
	First-Ply Failure Theory	Progressive Damage Model
1	-12%	27%
2	-25%	26%
3	-14%	13%
4	-41%	11%
5	-24%	11%

III. Section Forces of PAF under Various Load Cases

In order to design the fastened joints at the forward end of the PAF, the section forces and bending moments of the PAF were calculated using the finite element method. The commercial finite element software ABAQUS was used for this purpose. Figure 9 shows the finite element model of the PAF with 3D 4-noded, layered shell reduced integration (S4R) elements. The boundary condition applied was hinged nodes at the bottom of the PAF. Several designed load cases of combinations of axial force, lateral force, lateral bending moment, and pressure were applied to the finite element models via a hat-shaped load introduction structure as shown in Figures 10 and 11. The load introduction structure actually simulates the stiffness of the payload and its sub-system and it determines how the loads will be transferred into the PAF. It should be noted that the hat-shaped load introduction structure is nowhere near the real design but only serves the purpose of demonstrating how its stiffness affects the section forces and bending moments in the PAF.

*SF*1, *SF*2, *SF*3, *SF*4, and *SF*5 are the section forces and *SM*1, *SM*2, and *SM*3, are the bending/twisting moments of the PAF per unit width, where the 1-, 2-, and 3(*n*)-directions (meridional, hoop, and thickness directions) are defined in Figure 12. These section forces and moments are related to the in-plane and out-of-plane stresses as

$$(SF1, SF2, SF3, SF4, SF5) = \int_{-h/2}^{h/2} (\sigma_{11}, \sigma_{22}, \sigma_{12}, \sigma_{13}, \sigma_{23}) dz \tag{14}$$

$$(SM1, SM2, SM3) = \int_{-h/2}^{h/2} (\sigma_{11}, \sigma_{22}, \sigma_{12}) z dz \tag{15}$$

where σ_{11} and σ_{22} are the membrane stresses in the meridional (1-direction) and hoop directions (2-direction), respectively, z (3, or n-direction) is the coordinate in the thickness direction of the PAF sandwich panel and has its origin at the mid-plane, σ_{12}, σ_{13}, and σ_{23} are the shear stresses, and h is the thickness of the PAF sandwich panel.

Another finite element model constructed was one with the simulated payload separation system added between the PAF and the load introduction structure as shown in Figure 13. The payload separation system was simulated as a metal ringabove the forward end of the PAF. Same as the hat-shaped load introduction structure, the added metal ring only serves the purpose of demonstrating the effects of added stiffness of the payload system on the section forces and moments of the PAF but does not reflect any concept or the design of the payload separation.

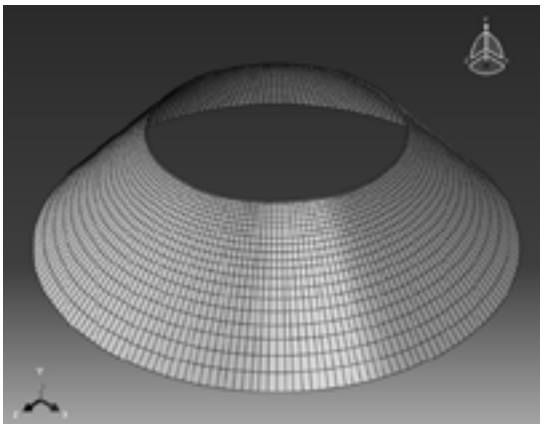


Figure 9. Geometry and FE model of PAF.

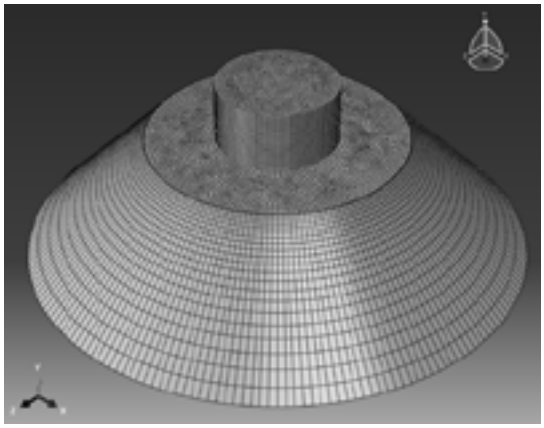


Figure 10. PAF with load introduction structure.

Because the purpose of the finite element models was to calculate the panel force at the forward end which will translate to the bearing load of the fastened joints, only section force in the 1-direction is considered for the joint design. The desired PAF panel needs to be able to withstand all forces in other directions. Once the PAF panel's section forces and moments are calculated via the finite element models, the critical facesheet force per unit width in the 1-direction (SFf1) was calculated by adding the bending force to the membrane force as

$$Sff1 = \frac{SF1}{2} \pm \frac{SM1}{h_f} \tag{16}$$

where h_f is the distance between the centers of the two facesheets of the PAF panel. This is under the assumption that the core carries insignificant force and the facesheets are thin compared to the panel thickness. Assume that two symmetric facesheets are used for the PAF panel, the equivalent panel section forces SF1eq was then calculated by multiplying SFf1 by 2.

$$SF1eq = 2Sff1 \tag{17}$$

The same procedures were used to calculate SF2eq and SF3eq for the panel design. To demonstrate the influence of the stiffness of the load introduction structure on the section forces, two additional thicknesses of the load introduction structure were applied and section forces were calculated under load case B.

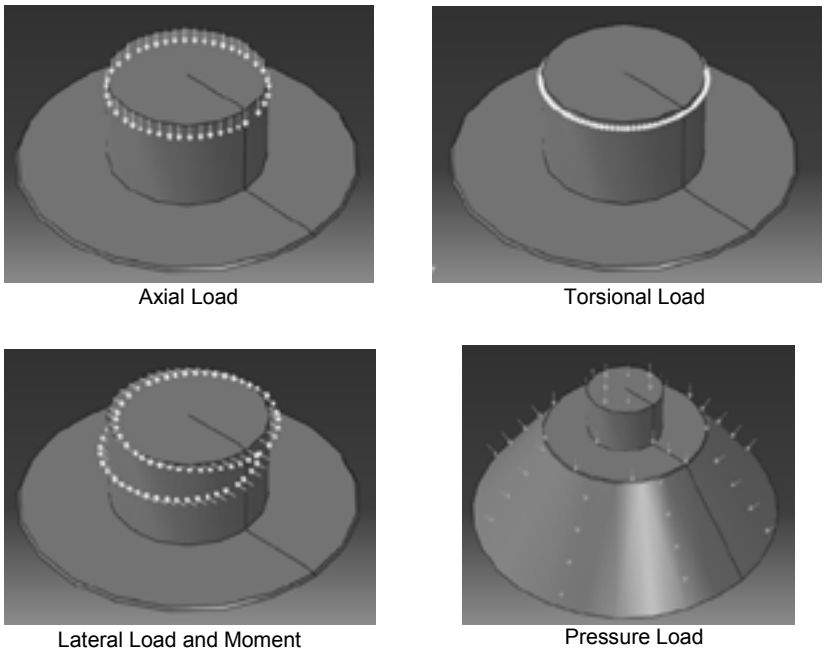


Figure 11. Applied loads via the load introduction structure.

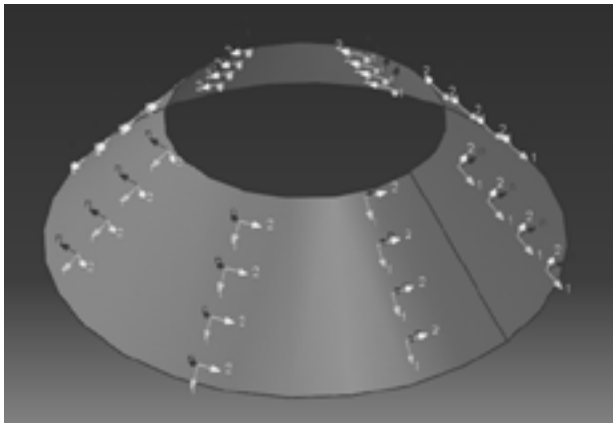


Figure 12. PAF coordinates.

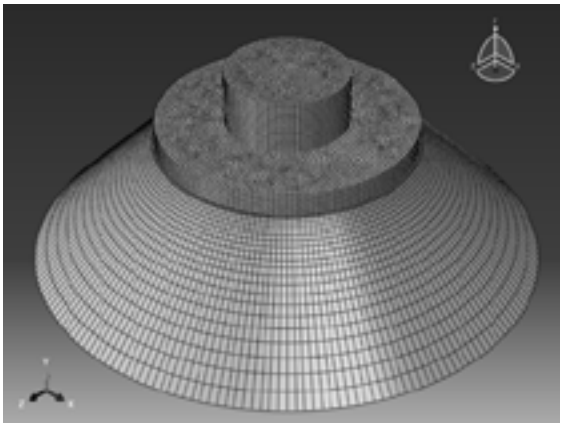


Figure 13. PAF with load introduction system and payload separation system.

After the finite element results were obtained, as expected, both the most critical section force SF1 and section bending moment SM1 occurred at the forward end of the PAF where the joints are located. Tables 3-5 demonstrate the scaled section forces and bending moments based on the minimum value under load case A. Table 3 lists all scaled critical section forces, including SF1eq, under different load cases. Also can be seen in Table 3, the stiffness of the load introduction structure greatly affects the section bending moments, especially on SM1 at the forward end. Tables 4 and 5 list the scaled critical section forces of the PAF with the simulated payload separation system of different stiffnesses added to the models. As can be seen, the stiffness of the payload separation system also has significant influence on the section forces.

IV. Conclusion

Three-dimensional finite element models with progressive failure analysis have shown a good potential to be an effective method to predict the ultimate failure loads of composite structures such as pinned composite joints. Such method has become attractive because these models can be executed on a personal computer with a reasonable runtime nowadays thanks to the increased computation power of these machines. The current study demonstrated better accuracy and less scatter of such models compared with the Tsai-Hill first-ply failure criterion on the predicted bearing capacity of composite laminates. A more sophisticated procedure to determine the degraded material moduli will further improve the accuracy of the models. Once the bearing capacity of the pinned joint is determined, the section forces at the forward end of the PAF need to be accurately determined in order to design the joints. Based on the finite element results from this study, the geometry, stiffness, and design of the payload and its subsystems have significant influence on the PAF's section forces for any given combination of loads. The design of the PAF and its joints at the forward end without given design criteria/detailed design of the payload and its subsystems above the PAF will be very difficult and unreliable.

Table 3. Section forces and moments of PAF without payload separation system.

Load Case	SF1 (Max Min)	SF2 (Max Min)	SF3 (Max Min)	SF4 (Max Min)	SF5 (Max Min)	SM1 (Max Min)	SM2 (Max Min)	SM3 (Max Min)	SF1eq (Max Min)	SF2eq (Max Min)	SF3eq (Max Min)
A	1	1	1	1	1	-168	-72.22	-1.28	1	1	1
	3.93	-3.46	-1.06	-11.20	-1.00	1	1	1	18.7	-2.18	-1.06
B	-2.48	3.77	80.0	1.80	3.00	-168	-66.7	-3.48	-8.58	1.32	2.36
	5.54	-5.46	-88.0	-10.0	-3.00	18.6	9.44	3.48	24.1	-3.00	-2.61
C	-3.37	3.23	48.0	4.80	2.00	-35.0	-13.9	-2.39	-14.2	1.68	1.44
	2.26	-3.15	-54.0	-2.00	-2.00	86.4	35.0	2.39	8.75	-1.32	-1.61
B_{thin}	-2.68	28.2	119	36.8	18.4	-1732	-767	-11.5	-35.0	19.1	3.72
	6.06	-6.46	-128	-156	-19.2	368	161	10.2	76.0	-6.53	-4.03
B_{thick}	-2.47	4.62	75.0	15.6	1.40	-36.4	-20.0	-2.83	-9.50	1.92	2.22
	5.49	-9.46	-83.0	-10.8	-1.40	17.27	11.7	2.83	18.9	-3.42	-2.44

Table 4. Section forces and moments of PAF with payload separation system of required shear stiffness.

Load Case	SF1 (Max Min)	SF2 (Max Min)	SF3 (Max Min)	SF4 (Max Min)	SF5 (Max Min)	SM1 (Max Min)	SM2 (Max Min)	SM3 (Max Min)	SF1eq (Max Min)	SF2eq (Max Min)	SF3eq (Max Min)
A	1	1	1	-16.0	1	1	1	1	1	1	1
	3.73	121	-1.07	1	-1.00	-2.31	-3.09	-1.07	6.56	-78.1	-1.06
B	-2.38	-74.9	2.30	-24.0	3.00	1.44	2.00	3.20	-4.14	48.2	2.35
	5.27	171	-2.57	9.70	-3.00	-3.44	-4.55	-3.27	9.38	-111	-2.59
C	-3.20	-101	1.37	-8.90	2.00	1.88	2.55	2.13	-5.48	64.7	1.41
	2.15	70.0	-1.57	13.00	-2.00	-1.31	-1.82	-2.13	3.74	-45.0	-1.59
B_{thin}	-2.37	-78.6	2.27	-28.0	2.00	1.94	2.55	3.13	-4.55	52.1	2.29
	5.22	185	-2.53	11.0	-2.00	-5.50	-6.73	-3.13	11.0	-125	-2.50
B_{thick}	-2.40	-74.6	2.33	-23.0	3.00	1.50	2.00	3.20	-4.11	48.1	2.38
	5.27	169	-2.57	9.60	-3.00	-3.13	-4.18	-3.27	9.14	-108	-2.59

Table 5. Section forces and moments of PAF with payload separation system of twice the required shear stiffness.

Load Case	SF1 (Max Min)	SF2 (Max Min)	SF3 (Max Min)	SF4 (Max Min)	SF5 (Max Min)	SM1 (Max Min)	SM2 (Max Min)	SM3 (Max Min)	SF1eq (Max Min)	SF2eq (Max Min)	SF3eq (Max Min)
A	1	1	1	-22.9	1	1	1	1	1	1	1
	3.73	54.6	-1.07	1	-1.00	-4.27	-5.41	-1.00	8.81	646	-1.06
B	-2.37	-33.9	2.27	-32.9	2.00	2.64	3.51	3.08	-5.53	-401	2.33
	5.28	77.0	-2.53	14.3	-2.00	-6.18	-7.84	-3.08	12.5	916	-2.58
C	-3.20	-45.4	1.37	-12.4	1.00	3.45	4.32	2.08	-7.40	-537	1.39
	2.15	31.6	-1.53	18.6	-1.00	-2.36	-2.97	-2.08	5.00	371	-1.58
B_{thin}	-2.37	-35.0	2.23	-38.6	1.00	3.64	4.59	2.69	-6.23	-434	2.24
	5.22	80.6	-2.47	15.7	-1.00	-9.55	-11.4	-3.08	14.91	1017	-2.45
B_{thick}	-2.38	-33.7	2.30	-31.4	2.00	2.55	3.24	3.08	-5.47	-399	2.36
	5.28	76.0	-2.53	12.9	-2.00	-5.55	-7.30	-3.08	12.11	891	-2.58

Acknowledgments

The first author of this report would like to express his appreciation to the Marshall Space Flight Center Faculty Fellowship Program Director, Dr. Frank Six, and all the coordinators including Dr. Gerald Karr, Brooke Graham, and Katie Hayden, for providing this unique opportunity to work with NASA researchers in a NASA environment. A special gratitude he gives to the MSFC Collaborators, Dr. William Guin and Justin Jackson. They have spent time and efforts to make this Fellowship meaningful and fruitful. Many thanks also go to Melina Zaharias, Dawn Phillips, and Robert Wingate for conducting the tests and providing critical PAF information.

References

[1]Thoppul, S. D., Finegan, J., and Gibson, R. F., "Mechanics of Mechanically Fastened Joints in Polymer–Matrix Composite Structures – A review," *Composites Science and Technology,* Vol. 69, 2009, pp. 301-329.

[2]Xiao, Y., and Ishikawa, T., "Bearing Strength and Failure Behavior of Bolted Composite Joints (part I: Experimental Investigation)," *Composites Science and Technology,* Vol. 65, 2005, pp. 1022-1031.

[3]Xiao, Y., and Ishikawa, T., "Bearing Strength and Failure Behavior of Bolted Composite Joints (part II: Modeling and Simulation)," *Composites Science and Technology,* Vol. 65, 2005, pp. 1032-1043.

[4]Camanho, P. P., and Lambert, M., "A Design Methodology for Mechanically Fastened Joints in Laminated Composite Materials," *Composites Science and Technology,* Vol. 66, 2006, pp. 3004-3020.

[5]Broughton, W. R., Crocker, L. E., and Gower, M. R. L., "Design Requirements for Bonded and Bolted Composite Structures," NPL Report MATC(A)65, January 2002.

[6]İçten, B. M., and Sayman, O., "Failure Analysis of Pin-Loaded Aluminum–Glass–Epoxy Sandwich Composite Plates," *Composites Science and Technology,* Vol. 63, 2003, pp. 727-737.

[7]Shokrieh, M. M., "Failure of Laminated Composite Pinned Connections," Master's Thesis, Dept. of Mechanical Engineering, McGill Univ., Montreal, Canada, 1991.

[8]Zhou Y., Fei, Q., and Tao, J., "Profile design of loaded pins in composite single lap joints: From circular to non-circular," *Results in Physics*, Vol. 6, 2016, pp. 471-480.

[9]Okutan, B., and Karakuzu, R., "The Failure Strength For Pin-loaded Multi-directional Fiber-glass Reinforced Epoxy Laminate," *Journal of Composite Materials*, Vol. 36, No. 24, 2002, pp. 2695-2712.

[10]Okutan, B., and Karakuzu, R., "The strength of pinned joints in laminated composites," *Composites Science and Technology,* Vol. 63, 2003, pp. 893-905.

[11]Lessard, L. B., and Shokrieh, M. M., "Two-Dimensional Modeling of Composite Pinned-Joint failure ," *Journal of Composite Materials*, Vol. 29, No. 5, 1995, pp. 671-697.

[12]Wang, H. S., Hung, C. L., and Chang, F. K., "Bearing Failure of Bolted Composite Joints. Part I: Experimental Characterization," *Journal of Composite Materials*, Vol. 30, No. 12, 1996, pp. 1284-1313.

[13]Hung, C. L. and Chang, F. K., "Bearing Failure of Bolted Composite Joints. Part II: Model and Verification," *Journal of Composite Materials*, Vol. 30, No. 12, 1996, pp. 1359-1400.

[14]Chamis, C. C., "Simplified Procedures for Designing Composite Bolted Joints," NASA Technical Memorandum 100281, 1988.

[15]Hashin, Z. "Failure Criteria for Unidirectional Fibre Composites," *J. of Applied Mechanics*, Vol. 47, 1980, pp. 329–334.

[16]Olmedo, Á. and Santiuste, C., "On the prediction of bolted single-lap composite joints," *Composite Structures*, Vol. 94, No. 6, 2012, pp. 2110–2117.

[17]Chang, F-K., and Lessard, L. B., "Damage Tolerance of Laminated Composites Containing an Open Hole and Subjected to Compressive Loadings: Part I—Analysis," *J. of Composite Materials*, Vol. 25, No. 1, 1991, pp. 2–43.

APPENDIX A—NASA MARSHALL SPACE FLIGHT CENTER FACULTY FELLOWSHIP PROGRAM ANNOUNCEMENT

MARSHALL FACULTY FELLOWSHIP PROGRAM

JUNE 3, 2019 – AUGUST 9, 2019

Application Deadline February 8, 2019

- The Marshall Space Flight Center is offering Faculty Fellowships for qualified STEM faculty at U.S. colleges and universities to conduct research with NASA colleagues during a ten-week residential program in Huntsville, Alabama.
- Faculty Fellows will receive stipends of $15,000 (Assistant Professor, Research Faculty), $17,000 (Associate Professor), or $19,000 (Professor).
- A relocation allowance of $1,500 will be provided to those fellows who live more than fifty miles from MSFC and a $500 travel supplement for one round-trip.
- Applicants must be U.S. citizens who hold full-time teaching or research appointments at accredited U.S. universities or colleges.
- During the ten-week program, fellows are required to conduct their research on-site at the Marshall Space Flight Center. A written final report is required at the end of the Fellowship.

Women and under-represented minorities, and persons with disabilities are encouraged to apply.

Website : https://www.uah.edu/asgc/applications/marshall-faculty-fellowship

APPENDIX B—NASA MARSHALL SPACE FLIGHT CENTER FACULTY FELLOWSHIP PROGRAM DESCRIPTION

NASA Marshall Faculty Fellowship Program

Program Description

- The Marshall Faculty Fellowship program is a residential research experience. Fellows are required to conduct their research, during the ten-week program, on-site at the Marshall Space Flight Center.
- Participants cannot receive remuneration from other entities or other programs or other university or government sources during the Faculty Fellowship 10-week period.
- An oral presentation by the Fellow to the Marshall group with which s/he has been affiliated is required, near the end of the fellowship period.
- A written final report is required at the end of the Fellowship.
- A written evaluation of the program by the Fellow is expected at the end of the Fellowship.

Eligibility

- US citizen
- Full time teaching or research appointment at accredited US university or college.
- Fellowship is awarded for one summer period, but Fellow may apply again for a second year.
- Women, under-represented minorities, and persons, with disabilities are encouraged to apply.

Selection

The applications selected to be Faculty Fellows will be chosen by the Marshall group which has been assigned the area of investigation (concentration) chosen by the applicant.

Marshall Collaborator

A Marshall Collaborator will be identified to serve as the co-investigator and day-to-day contact. At the end of the ten-week period, the Faculty Fellow and the Marshall Collaborator will prepare a white paper summarizing the summer effort, including results and recommending follow-up work.

Compensation

Stipends for Faculty Fellows are set as follows for the 10-week period:

Assistant Professors and Research Faculty	$15,000
Associate Professors	$17,000
Professors	$19,000

A relocation allowance of $1,500 will be provided to fellows who live more than fifty miles from the Marshall Center.

A travel supplement of $500 will be provided to those fellows receiving the relocation allowance.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01–08–2020	Technical Memorandum	

4. TITLE AND SUBTITLE

Marshall Space Flight Center Faculty Fellowship Program

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

N.F. Six, Program Director and G. Karr, Compiler*

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

George C. Marshall Space Flight Center
Huntsville, AL 35812

8. PERFORMING ORGANIZATION REPORT NUMBER

M–1512

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546–0001

10. SPONSORING/MONITOR'S ACRONYM(S)

NASA

11. SPONSORING/MONITORING REPORT NUMBER

NASA/TM–20205003520

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified-Unlimited
Subject Category 31
Availability: NASA STI Information Desk (757–864–9658)

13. SUPPLEMENTARY NOTES

Prepared by the Academic Affairs Office, Office of Human Capital
*The University of Alabama in Huntsville

14. ABSTRACT

The 2019 Marshall Faculty Fellowship Program involved 14 faculty in the laboratories and departments at Marshall Space Flight Center and one faculty researcher working from Colorado. These faculty engineers and scientists worked with NASA collaborators on NASA projects, bringing new perspectives and solutions to bear. This Technical Memorandum is a compilation of the research reports of the 2019 Marshall Faculty Fellowship program, along with the Program Announcement (Appendix A) and the Program Description (Appendix B). The research affected the following five areas: Materials, Propulsion, Spacecraft Systems, Vehicle Systems, and Space Science.

15. SUBJECT TERMS

wireless sensors, propulsion, green propellant, materials, 3D printing, friction stir welds, instruments, solar neutrinos, gamma-ray polarimeter, lightning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk at email: help@sti.nasa.gov
U	U	U	UU	174	19b. TELEPHONE NUMBER *(Include area code)* STI Help Desk at: 757–864–9658

National Aeronautics and
Space Administration
IS02
George C. Marshall Space Flight Center
Huntsville, Alabama 35812