Rapid Design Space Exploration for Conceptual Design of Hypersonic Missions

Michael J. Grant*, Ian G. Clark†, and Robert D. Braun‡
Georgia Institute of Technology, Atlanta, GA, 30332

During conceptual design, multidisciplinary optimization is often performed using computationally intensive direct methods. Prior work has shown that rapid design studies can be performed using fast indirect methods, but several optimization techniques including discrete dynamic programming, nonlinear inversion, and pseudospectral methods are required to construct a suitable initial guess within the design space. In this investigation, a simplified methodology is developed to eliminate these optimization techniques, enabling rapid design space exploration using continuation of indirect methods alone. This is made possible by initially converging to a simple solution that is outside of the design space of interest, and solutions within the design space of interest are quickly accessed through continuation from this initial solution. As an initial step to automate this continuation process, state transition tensors are used to predict optimal solutions throughout the design space. A methodology is developed to provide accurate predictions of trajectories with varying flight times, and the error of these predictions is controlled to regulate the continuation process. This approach provides flexibility to adapt to future computational capabilities and serves as an initial step to bridge the gap between conceptual trajectory design and onboard trajectory planning.

Nomenclature

DSM	Design Structure Matrix
GPU	Graphics Processing Unit
LGL	Legendre-Gauss-Lobatto
MDO	Multidisciplinary Design Optimization
NLP	Nonlinear Programming
RK	Runge-Kutta
STT	State Transition Tensor
A	vehicle reference area, m^2
C_D	drag coefficient
C_L	lift coefficient
D	drag force magnitude, N
H	scale height, m
L	lift force magnitude, N
L/D	lift to drag ratio
m	vehicle mass, kg
p	order of Taylor series solution
r	radial magnitude, m
r_e	Earth radius, m
t	time, s

*Graduate Research Assistant, Guggenheim School of Aerospace Engineering, AIAA Student Member.
†Visiting Assistant Professor, Guggenheim School of Aerospace Engineering, AIAA Member.
‡David and Andrew Lewis Associate Professor of Space Technology, Guggenheim School of Aerospace Engineering, AIAA Fellow.

v	relative velocity magnitude, m/s
$\mathbf{x}'$	augmented state vector
α	angle of attack, deg
β	ballistic coefficient, kg/m^2
μ	gravitational parameter, m^3/s^2
γ	relative flight path angle, rad
θ	downrange subtended angle, rad
λ	costate
ρ_0	atmospheric density at the surface, kg/m^3
$\mathbf{\Phi}$	state transition tensor
ϕ	bank angle, deg
*	reference solution value

I. Introduction

TRADITIONALLY, conceptual design is performed using various tools provided by disciplinary experts. The interaction among these tools during hypersonic mission design is often illustrated using a design structure matrix (DSM) as shown in Figure 1. In this methodology, design variables are often chosen within each discipline, and a multidisciplinary design optimization (MDO) scheme is used to optimize the overall system.[1,2,3,4,5,6,7] These schemes are often classified as direct optimization methods which operate on a finite number of design variables. As a result, continuous design variables such as hypersonic trajectories must be discretized for incorporation into these methods. Collocation, pseudospectral, and explicit shooting methods are three common trajectory optimization methods used in system design studies.[8,9,10,11] A common collocation method arranges nodes along a path, and the trajectory is optimized assuming a cubic interpolation between the nodes. In this method, no assumption is made *a priori* regarding the form of the trajectory. Alternatively, pseudospectral methods implement a more efficient quadrature scheme such as Legendre-Gauss-Lobatto with the assumption that the trajectory can be approximated as a polynomial. Both of these approaches are implicit methods that require use of nonlinear programming (NLP) solvers such as SNOPT.[12] Rather than solving a large NLP problem, explicit shooting methods discretize the control history and propagate the equations of motion to a terminal point. While shooting methods require less memory, satisfying path constraints such as maximum heat rate and g-loading can be challenging during the shooting process. However, shooting methods allow global searching algorithms widely used in systems studies, such as genetic algorithms and particle swarm optimizers, to be extended to include trajectory optimization.[11,13] Although shooting methods are easy to implement with existing tools, prior work has shown that continuation of indirect methods can improve the efficiency of entry corridor design studies.[14]

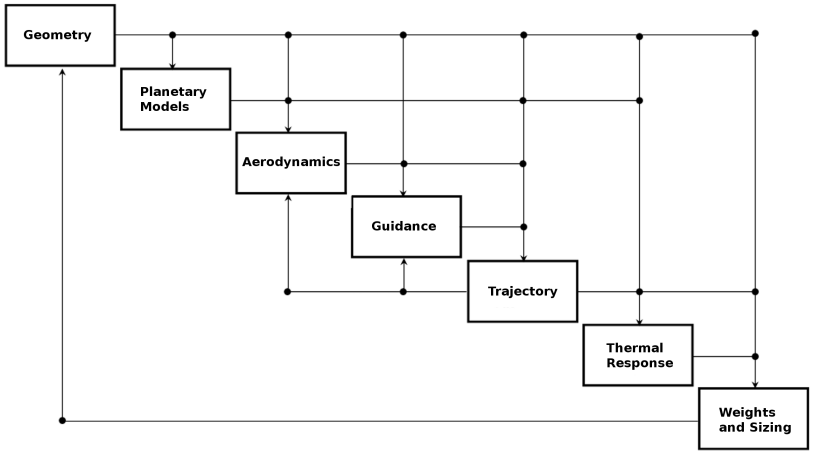

Figure 1. Example design structure matrix for entry systems.[15]

II. Rapid Trajectory Optimization Background

In this prior work, a rapid trajectory design methodology is formed through the combination and advancement of disparate trajectory optimization techniques that had been developed over the previous century into a unified framework that is capable of solving a wide range of design problems. Specifically, this framework implements discrete dynamic programming, nonlinear inversion, pseudospectral methods, indirect methods, and continuation. This methodology balances the advantages and disadvantages of both direct and indirect methods as described in Table 1. While direct methods are generally slower than indirect methods when converging to an optimal solution, direct methods have a much larger region of attraction that enables this convergence.[16,17] As a result, direct methods, including those previously discussed, have been widely adopted over indirect methods for design studies. However, many solutions are desired from design studies, and, in many cases, trades in designs are performed in succession. Rather than executing successive direct methods, this methodology capitalizes on the computational efficiency of indirect methods to perform design studies using continuation.[14] This process has been shown to be successful for a range of vehicle shapes, constraints, environment parameters, initial conditions, and terminal conditions. This trajectory optimization process has also been extended to include vehicle shape, allowing rapid simultaneous hypersonic aerodynamic and trajectory optimization for conceptual design.[18] While this methodology enables the rapid conceptual design of hypersonic missions, several limitations remain.

Table 1. Comparison between direct and indirect methods.

	Advantages	*Disadvantages*
Direct Methods	Large region of attraction	Computationally intensive
	Widespread NLP solvers exist	Optimality not guaranteed
Indirect Methods	Rapid convergence	Small region of attraction
	Necessary conditions satisfied	Costates introduced

II.A. Prior Methodology Limitations

II.A.1. Framework Complexity

In order to successfully execute the design methodology, several optimization techniques are required. These techniques include discrete dynamic programming, nonlinear inversion, pseudospectral methods, indirect methods, and continuation. The integration of these algorithms greatly increases the complexity of the process and would likely limit its adoption by the design community.

II.A.2. Path Cost Limitation and Sensitivity of Discrete Dynamic Programming

In the prior methodology, path constraints, such as maximum heat rate and g-loading, as well as design constraints, such as minimum control authority, are shown to reduce the number of trajectory options that define an entry corridor. As such, discrete dynamic programming is used to construct a good initial guess within this corridor. While this approach is very efficient to minimize path cost throughout the corridor, incorporating terminal cost into the objective would be problematic. Furthermore, the dynamic programming solution provides a good initial guess within altitude-velocity space. As such, the remaining states must be estimated using nonlinear inversion of the altitude-velocity solution. Along certain portions of the the trajectory, the estimation of flight path angle is highly sensitive to the slope of the altitude-velocity solution.[14] For certain problems, these numerical challenges may reduce the quality of the initial guess provided by discrete dynamic programming. These challenges could be addressed through various grid densities or adaptive mesh techniques, but these solutions would continue to add complexity into the process.

II.A.3. Pseudospectral Complexities

In the previous work, the solution from a pseudospectral method is used to provide a good initial guess required for the continuation of indirect methods. Pseudospectral methods approximate the trajectory as a polynomial by collocating nodes at efficient locations such as the Legendre-Gauss-Lobatto (LGL) points shown in Figure 2. Note that for a given number of nodes, the location of these nodes is fixed and problem independent. Additionally, these nodes are highly clustered near the endpoints. While this reduces the onset of the Runge phenomenon when interpolating high order polynomials near the endpoints, high dynamic regions of many hypersonic missions, such as entry, occur in the middle of the trajectory where the dynamic pressure peaks. This is observed when comparing the LGL node locations with the node locations of an error-controlled, adaptive Runge-Kutta (RK) method for a representative Earth entry problem shown in Figure 2. Without assuming a form of the solution *a priori*, the adaptive RK method clusters nodes during high dynamic periods which occur near the center of the trajectory.

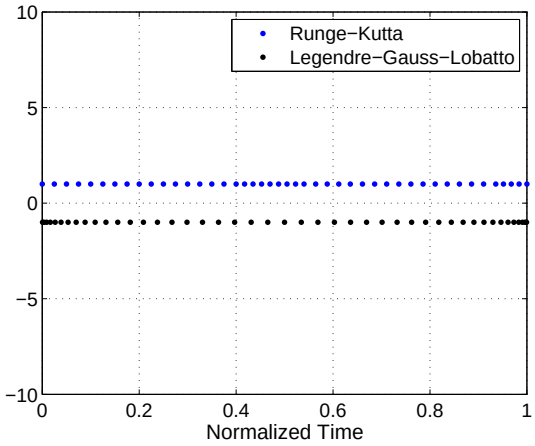

Figure 2. Comparison of node locations for a representative Earth entry problem (45 nodes for both methods).

Collocating nodes at the LGL points requires use of an NLP solver in which the equations of motion are enforced as constraints at the nodes. Due to the clustering of nodes at the endpoints, a sufficient number of nodes must be chosen to ensure accuracy of the interpolated dynamics during the high dynamic region near the center of the trajectory. While adaptive mesh refinement methods have been developed to address this issue, successive trajectory optimizations with varying node counts and arrangements are required.[19,20] Furthermore, NLP solvers such as SNOPT employ penalty functions that may result in additional complexity during convergence to the optimal solution. Finally, the converged pseudospectral solution is mapped to an indirect solution using the Covector Mapping Theorem.[21,22] The mapping requires identification of the discontinuous changes in costates that occur at the entrance of path constraints. This identification can be challenging to accomplish in an automated manner when this discontinuity resides between the collocated nodes.[23,24,25]

II.A.4. The Disparity Between Conceptual Design and Onboard Operations

In general, the open-loop optimal trajectories simulated during conceptual design are not consistent with the closed-loop trajectories flown by fielded systems. When developing a flight vehicle, substantial effort is placed on the selection and development of a suitable guidance algorithm. Once a guidance algorithm has been selected, modifications are often made to tailor the algorithm to a specific mission.[26] Efforts such as these often complicate the design process and are generally omitted from conceptual design.

During conceptual design, optimal solutions are used to analyze various vehicle and trajectory combinations over a range of environment and mission requirements uncertainties that form the design space as notionally shown in Figure 3. Analogously, real-time trajectory planning would enable optimal trajectory solutions to be constructed during flight to identify the best path within a range of expected flight conditions that form a subset of this design space. As a result, the need for traditional guidance algorithms would be eliminated, and the conceptual design process would better reflect onboard operations. If the prior methodology is used, then an appropriate balance between the use of pseudospectral methods and continuation of indirect methods is required to perform efficient exploration of this design space. For example, if the pseudospectral solution creates an anchor point in this design space to seed the continuation of indirect solutions, then the designer must decide where to create these anchor points. Furthermore, many options likely exist when continuing from the anchor solution to another solution of interest as shown in Figure 4. The optimal continuation policy may influence the choice of anchor solution locations, but this policy is problem dependent and would likely require substantial computational resources. If the continuation of indirect methods is used to perform real-time trajectory planning, then this analysis would likely improve

the onboard performance of the continuation process. However, this analysis would likely not be worthwhile during conceptual design due to the required increase in computational resources. Finally, while the prior methodology could provide a consistent rapid optimization framework for both conceptual design and real-time trajectory planning, the complexity resulting from the use of many disparate optimization techniques would likely lead to additional challenges during certification for onboard operations.

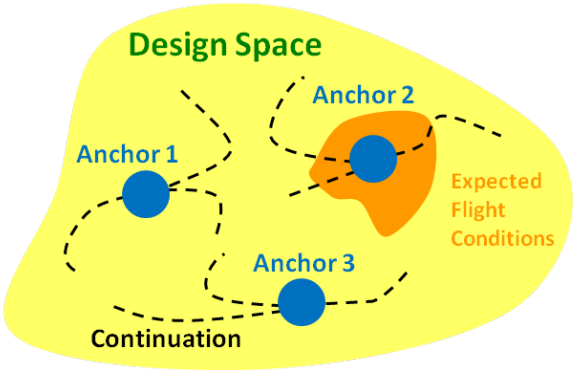

Figure 3. Example of design space exploration using continuation from anchor solution.

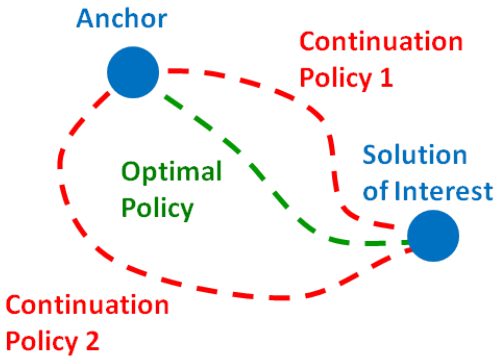

Figure 4. Example of optimal continuation policy.

This prior methodology is analogous to many MDO algorithms in which relevant design solutions are immediately sought. As such, pseudospectral methods are required to seed the continuation process due to the difficulty of providing a sufficient initial guess to converge to an indirect solution, especially in the presence of path constraints such as maximum heat rate and g-loading. Consequently, these processes are generally confined to the plane of the design space represented in Figure 3. If solutions of minimal immediate value to the designer are also considered, then optimal solutions throughout the design space can be efficiently constructed using continuation of indirect methods alone. This is accomplished by seeding the indirect method with a simple, short, unconstrained trajectory solution outside of the design space as opposed to a constrained pseudospectral solution within the design space. This greatly simplifies the process by eliminating the dynamic programming, nonlinear inversion, and pseudospectral steps of the prior methodology. As a result, a simplified methodology that relies only on the continuation of indirect methods is developed to address the aforementioned limitations and perform rapid design space exploration.

III. Rapid Design Space Exploration Using Continuation

For this study, a planar entry trajectory is assumed with equations of motion shown in Eq. (1)-(4), where t is the time, r is the radial magnitude, θ is the downrange subtended angle, v is the relative velocity magnitude, γ is the relative flight path angle, D is the drag force magnitude, m is the mass of the vehicle, μ is the gravitational parameter, L is the lift force magnitude, and ϕ is the bank angle. For comparison with the prior methodology, trajectories are again optimized to minimize total heat load.[14] To minimize heat load, the heat rate must be maximized along every portion of the trajectory, and this result is evident from the optimal solutions presented in this report. Nominal environment parameters are chosen as shown in Table 2, and a high performance blunted conic is chosen with geometric and aerodynamic parameters shown in Table 3. An entry mass of 4080 kg is assumed, resulting in a large ballistic coefficient, β, of 6,070 kg/m^2 for illustrative purposes in the following examples.

Table 2. Environment parameters.

Parameter	*Value*
Scale Height, H	7200 m
Surface Density, ρ_o	1.217 kg/m^3
Gravitational Parameter, μ	3.986e14 m^3/s^2
Earth Radius, r_e	6378000 m

$$\frac{dr}{dt} = v \sin\gamma \tag{1}$$

$$\frac{d\theta}{dt} = \frac{v\cos\gamma}{r} \tag{2}$$

$$\frac{dv}{dt} = -\frac{D}{m} - \frac{\mu \sin\gamma}{r^2} \tag{3}$$

$$\frac{d\gamma}{dt} = \frac{L\cos\phi}{mv} + \left(\frac{v}{r} - \frac{\mu}{vr^2}\right)\cos\gamma \tag{4}$$

Table 3. Blunted conic parameters.

Geometric Parameter	*Value*	*Aerodynamic Parameter*	*Value*
Length	6.92 m	α	10 deg
Nose Radius	0.025 m	C_D	0.144
Base Radius	1.22 m	C_L	0.311
Cone Half-Angle	10 deg	L/D	2.16

III.A. Construction of the Initial Indirect Solution Outside of the Design Space

As previously mentioned, it is generally impractical for the designer to provide a sufficient initial guess within the constrained design space without the use of pseudospectral methods. Alternatively, the designer can easily provide a sufficient initial guess to converge to a short, unconstrained, optimal trajectory that is outside of the design space. For the following example, a terminal altitude of 6 km and velocity of 1.4 km/s is enforced. After choosing a reasonable terminal flight path angle of -49^o and terminal costate values of zero, an initial guess can be constructed through reverse propagation of the states and costates for a short period of time which was chosen to be one second for this example. Before the trajectory is propagated, the lift of the vehicle is artificially reduced by a factor of 100. This prevents switches in control that could result along the unconstrained arc as described by the switching structure depicted in Table 4.[23,24,25] Finally, the initial altitude and velocity resulting from the reverse propagation are chosen as fixed initial conditions.

Table 4. Control switching structure.

$\lambda_\gamma < 0$	Bank = 0 deg
$\lambda_\gamma > 0$	Bank = 180 deg
$\lambda_\gamma = 0$	Bank is indeterminate

With this initial guess as well as fixed initial and terminal conditions, an indirect method is able to quickly converge to the short, unconstrained, minimum heat load solution. The resulting trajectory and costates are shown in Figures 5 and 6, respectively. As expected, the costates are smooth throughout the entire unconstrained trajectory. As with many common indirect method solvers, only the necessary conditions of optimality are guaranteed to be satisfied. The optimality of this short, unconstrained solution can be easily verified for this relatively simple trajectory optimization problem. As expected to maximize heat rate throughout the trajectory, the optimal solution consists of a full lift-up trajectory as described by the switching structure. The optimality of this solution could also be verified through analysis of the sufficient conditions of optimality. However, this is left as future work.

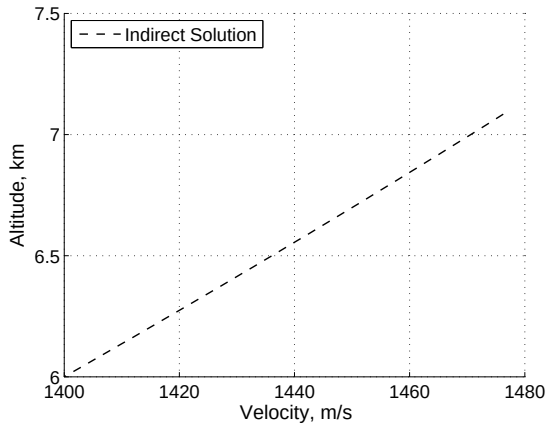

Figure 5. Initial indirect solution.

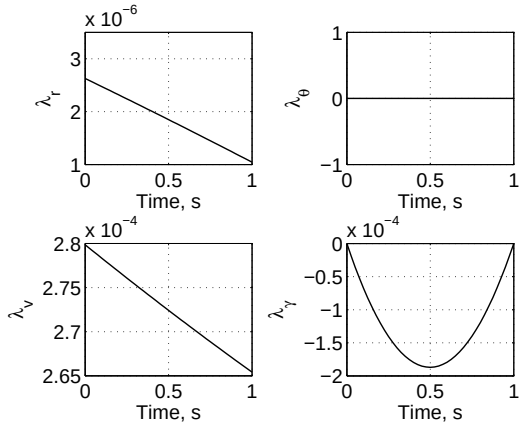

Figure 6. Costates from initial indirect solution.

III.B. Unconstrained Continuation to the Desired Initial State

The short, unconstrained trajectory is outside of the design space and is not of immediate interest to the designer. However, this solution serves as the starting point from which the trajectory is evolved to a solution of interest within the design space. With this converged indirect solution, the unconstrained trajectory must be extended through continuation to the set of desired entry conditions. The evolution of the optimal trajectories and corresponding costates is shown in Figures 7 and 8, respectively. During the continuation process, the prior solution serves as an initial guess to converge to the subsequent nearby solution. As the trajectory is lengthened, the switching structure is monitored to ensure no switches occur in the bank angle. During this process, λ_γ is always negative, implying that the artificial decrease in lift was sufficient to prevent switches in control. Additionally, each successive solution is monitored during the continuation process to determine if large changes occur between solutions. For example, the initial flight path angle is monitored to determine if the vehicle is initially traveling towards the target as desired. If the distance between solutions is too large, then the indirect method solver may converge to a solution in which the vehicle is initially traveling away from the target. If this occurs, then smaller distances are chosen between solutions during the continuation process, ensuring convergence to a proper initial flight path angle.

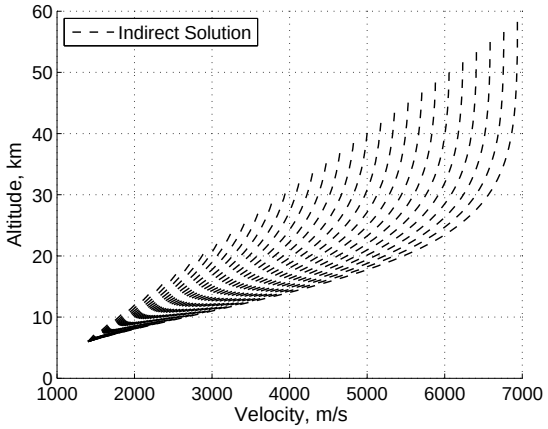

Figure 7. Trajectories from unconstrained continuation.

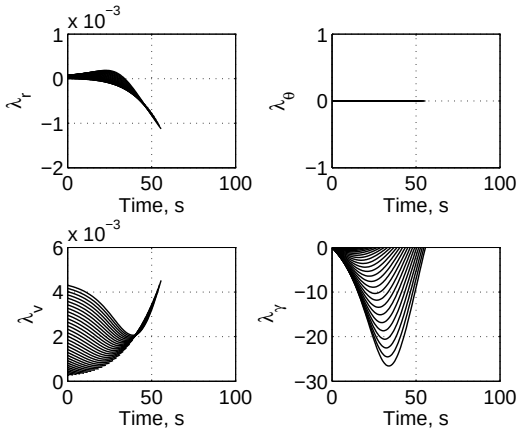

Figure 8. Costates from unconstrained continuation.

III.C. Introduction of Path Constraints

With the trajectory fully extended to include the desired initial conditions, path constraints such as maximum heat rate and g-loading can be incrementally introduced. In this process, path constraints that define a flight corridor are introduced at a location of tangency to the unconstrained trajectory as depicted in Figure 9. This approach serves two purposes. First, indirect methods require the number and order of unconstrained and constrained trajectory arcs to be specified prior to optimization. This is difficult to estimate *apriori*, and, in the prior methodology, this sequence is provided from the pseudospectral solution. After identifying the magnitude of the constraint that provides this tangency condition, a determination can be made as to whether or not the constraint is initially active and should be included in the optimization. Second, by introducing a constraint at the point of tangency, the states and time of the optimal trajectory remain unchanged. Only the costates change with the introduction of a discontinuity as a result of the corner conditions.[23,24,25] This provides a gradual change in solution necessary for convergence of the indirect method.

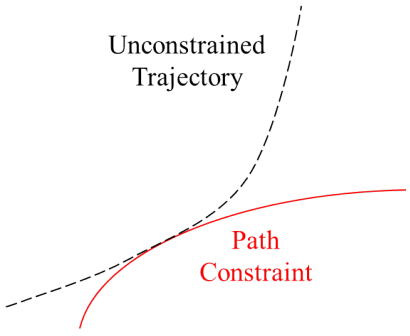

Figure 9. Introduction of path constraint.

In the prior unconstrained continuation process, lift is artificially decreased to prevent switches in bank angle. As path constraints are introduced, the lift of the vehicle must be incrementally restored to the actual value to enable the vehicle to maneuver along these constraints. As an example, a maximum heat rate

constraint is introduced at a location of tangency along the final trajectory from the previous unconstrained continuation process. Once the heat rate constraint is introduced in this manner, it is incrementally decreased to the desired value. As the heat rate constraint is decreased, the lift of the vehicle is restored to prevent saturation in bank angle while traveling along the constraint. At the end of this process, the lift of the vehicle is restored to 50% of the actual value. The resulting trajectories and costates from this process are shown in Figures 10 and 11, respectively. As shown, the initial flight path angle must be shallowed as the heat rate constraint is decreased to prevent the vehicle from violating the constraint. As expected, discontinuities occur in the costates at the time the vehicle enters the heat rate constraint. Note that as the magnitude of the heat rate constraint is reduced and the lift of the vehicle is increased, both the length of the trajectory and time of the entrance to the heat rate constraint increase. It is also important to note that the switching structure only applies to the unconstrained portions of the trajectory. While λ_γ is shown to be positive during the middle portion of each trajectory, this occurs while the vehicle is traveling along the heat rate constraint and, as a result, does not conflict with the switching structure. As expected, λ_γ is verified to be negative along the unconstrained arcs.

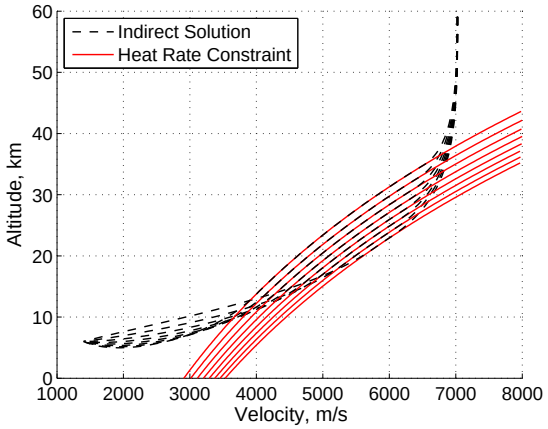

Figure 10. Trajectories from continuation of heat rate constraint and L/D.

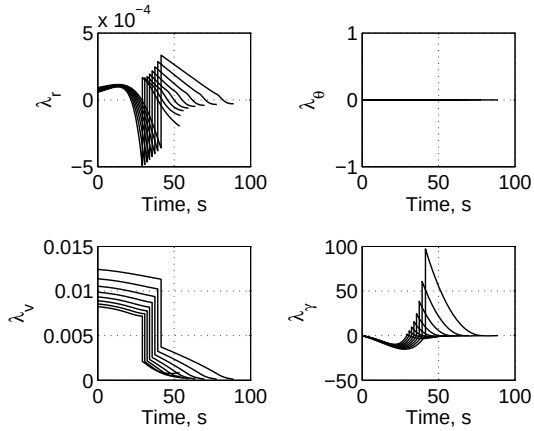

Figure 11. Costates from continuation of heat rate constraint and L/D.

After the continuation of the heat rate constraint is completed, additional path constraints such as maximum g-loading can also be introduced and modified in a similar manner. The trajectories and costates that result from continuation of the g-loading constraint are shown in Figures 12 and 13, respectively. During this process, the lift of the vehicle is fully restored. As expected, a second discontinuity occurs where the vehicle enters the g-loading constraint at around 80 s. At the end of this process, a constrained indirect solution is obtained within the design space of interest. It is important to note that this process is executed quickly through continuation of fast indirect methods, eliminating the need for the slower processes of discrete dynamic programming, nonlinear inversion, and pseudospectral methods associated with the previous methodology. Consequently, this eliminates the need to construct anchor solutions at strategic locations throughout the design space. Instead, any point within the design space can be quickly accessed through continuation from a simple solution outside of the design space. As a result, rapid design space exploration can be performed using continuation of fast indirect methods alone.

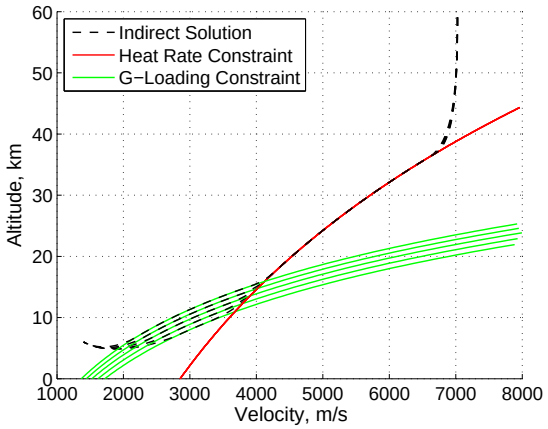

Figure 12. Trajectories from continuation of g-loading constraint and L/D.

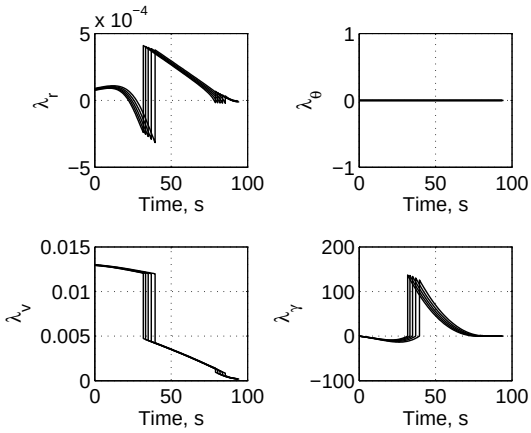

Figure 13. Costates from continuation of g-loading constraint and L/D.

In the prior example, algorithm monitors are used to ensure the initial flight path angle and lift restoration provides feasible solutions during the continuation process. However, parameters are changed manually during this continuation process. Without prior knowledge of optimal solutions during the continuation process, small steps must be conservatively chosen after trial and error to ensure that each prior solution serves as a good initial guess to converge to the next indirect solution. If indirect solutions could be predicted during the continuation process, then improved initial guesses could be constructed. Additionally, if the error of this prediction could also be estimated, then the choice in step size can be automated to perform efficient design space exploration.

IV. Unconstrained Solution Prediction Using State Transition Tensors

In the prior example, the construction of optimal solutions within the design space begins with the continuation of unconstrained solutions to the desired initial condition. As such, the prediction of unconstrained solutions must be addressed. While traditional MDO methods rely on direct optimization techniques that perform a constrained search for the optimum, indirect methods convert the optimization problem into a root-solving problem of boundary conditions. If a relationship between these boundary conditions can be constructed, then optimal solutions can be predicted during the continuation process. This relationship is constructed using State Transition Tensors (STTs).

IV.A. State Transition Tensors

In this work, STTs are used to relate changes in initial states and costates from a reference solution (denoted by *) to changes in states and costates at a future time as shown in Eq. (5), where $\mathbf{x}'_i$ is the i^{th} component of an augmented state vector that also includes costates and $\mathbf{\Phi}_{i,k_1\cdots k_p}$ is the corresponding p^{th} order STT. As a result, deviations in the augmented state with n elements can be expressed as a Taylor series of STTs to order m as shown in Eq. (6), and deviations in the rate of the augmented state can expressed in the form of Eq. (7). The equations of motion can be combined with the costate rate equations obtained from the necessary conditions of optimality to construct a relationship for the rate of the augmented state as shown in Eq. (8). The corresponding deviations in the rate of the augmented state can also be calculated by expressing Eq. (8) as a Taylor series as shown in Eq. (9), where the Jacobian of the augmented state is shown in Eq. (10). After substitution of Eq. (6) into Eq. (9), the STTs can be calculated by matching coefficients with Eq. (7).[27] An automated process has been developed to perform this coefficient matching for any order m chosen by the designer as well as to construct the appropriate subroutines for computing the STTs associated with the trajectory optimization of interest. After computing the STTs, predictions can be made during the continuation process in both the states and costates for nearby optimal solutions.

$$\mathbf{\Phi}_{i,k_1\cdots k_p} = \frac{\partial^p \mathbf{x}'_i}{\partial \mathbf{x}'^o_{k_1} \cdots \partial \mathbf{x}'^o_{k_p}} \tag{5}$$

$$\delta \mathbf{x}'_i(t) = \sum_{p=1}^{m} \frac{1}{p!} \left(\sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_p=1}^{n} \mathbf{\Phi}_{i,k_1\cdots k_p} \delta \mathbf{x}'^o_{k_1} \cdots \delta \mathbf{x}'^o_{k_p} \right) \tag{6}$$

$$\delta \dot{\mathbf{x}}'_i(t) = \sum_{p=1}^{m} \frac{1}{p!} \left(\sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_p=1}^{n} \dot{\mathbf{\Phi}}_{i,k_1\cdots k_p} \delta \mathbf{x}'^o_{k_1} \cdots \delta \mathbf{x}'^o_{k_p} \right) \tag{7}$$

$$\dot{\mathbf{x}}'_i(t) = \mathbf{f}'_i[t, \mathbf{x}'(t)] \tag{8}$$

$$\delta \dot{\mathbf{x}}'_i(t) = \sum_{p=1}^{m} \frac{1}{p!} \left(\sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_p=1}^{n} \mathbf{f}'^*_{i,k_1\cdots k_p} \delta \mathbf{x}'_{k_1} \cdots \delta \mathbf{x}'_{k_p} \right) \tag{9}$$

$$\mathbf{f}'^*_{i,k_1\cdots k_p} = \left. \frac{\partial^p \mathbf{f}'_i}{\partial \mathbf{x}'_{k_1} \cdots \partial \mathbf{x}'_{k_p}} \right|_{\mathbf{x}'=\mathbf{x}'^*} \tag{10}$$

Higher order STTs are often used in astrodynamics applications to maintain accuracy for the long propagation times required.[28,29,30,31] For relatively short hypersonic missions, higher order STTs could potentially be used to accurately predict optimal solutions at greater distances from the reference solution. If the optimal solutions throughout the design space can be expressed as a convergent STT series, then the accuracy of the predictions will improve as the order of the STT approximation is increased. However, not all functions can be approximated by a convergent Taylor series. A scalar example of each scenario will illustrate the considerations made when predicting optimal solutions using STTs.

IV.B. Taylor Series Examples

A scalar Taylor series approximation about the reference point a can be expressed in the form of Eq. (11), where $a = 0$ in the following examples. As shown in Figure 14, the Taylor series approximation of the function e^x appears to improve for all x as the order of the approximation increases. Alternatively, Figure 15 illustrates that the error of the Taylor series increases as the order of the approximation increases for the function $\log(1+x)$ where $x > 1$. During the continuation process, the exact solution of the predicted optimal trajectories is not known *a priori*, and as a result, this comparison cannot be made to directly determine if the STT prediction improves as the order of the approximation increases. Instead, convergence of the STT series can be indirectly observed from convergence of the approximate STT solutions as the order of the approximation increases. For a convergent series, the discrepancy between the approximate solution of order p, denoted as f_p, and of order $p+1$, denoted as f_{p+1}, should decay as p increases. For the convergent Taylor series example, this decay is observed at various values of x as shown in Figure 16. For the divergent Taylor series example, this decay is observed for $x < 1$ as shown in Figure 17. However, for $x > 1$, this decay only occurs for small orders of p. As p increases, the discrepancy also eventually increases. As shown in both examples, the discrepancy is likely to decrease as the order p increases for x near the reference point $a = 0$. The discrepancy between an approximation of order p and an approximation of order $p+1$ is shown in Eq. (12). For $|x-a| << 1$, $(x-a)^{p+1}$ will likely dominate this expression, resulting in the observed decay in discrepancy for both examples. However, as x is chosen farther from the reference point, the discrepancy will only decrease with order p if $\frac{f^{(p+1)}(a)}{(p+1)!}$ decreases in magnitude faster than $(x-a)^{p+1}$ increases in magnitude.

$$f(x+a) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(p)}(a)}{p!}(x-a)^p \tag{11}$$

$$|f_{p+1} - f_p| = \left| \frac{f^{(p+1)}(a)}{(p+1)!}(x-a)^{p+1} \right| \tag{12}$$

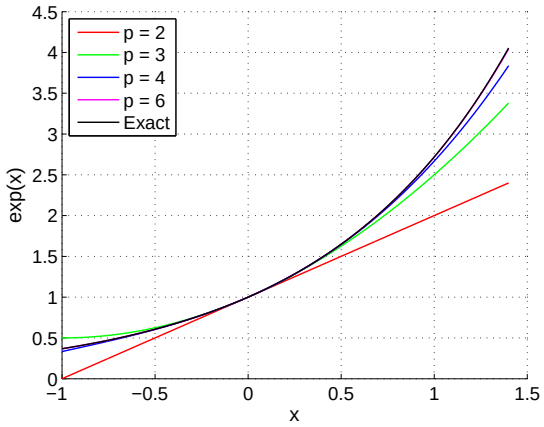

Figure 14. Example convergent Taylor series approximations for $\mathbf{f(x) = e^x}$.

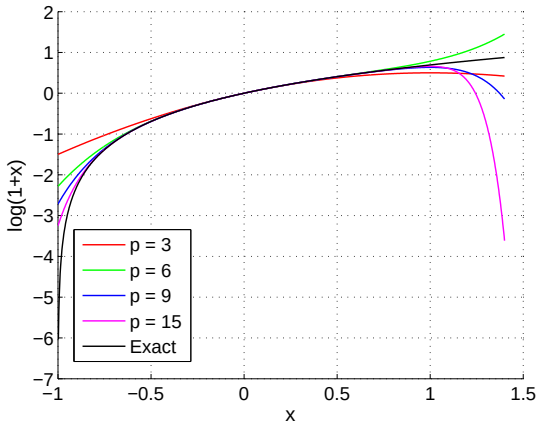

Figure 15. Example divergent Taylor series approximations for $\mathbf{f(x) = \log(1 + x)}$.

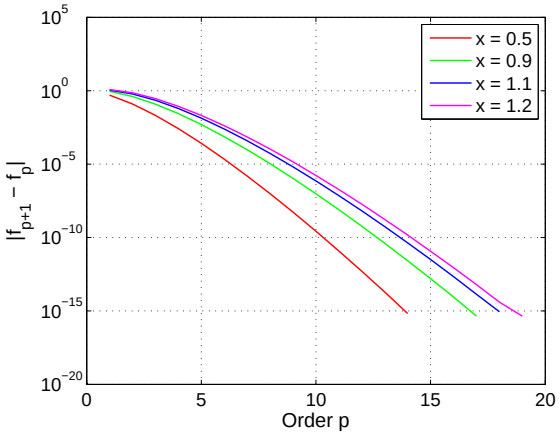

Figure 16. Discrepancy in convergent Taylor series estimates at various x locations for $\mathbf{f(x) = e^x}$.

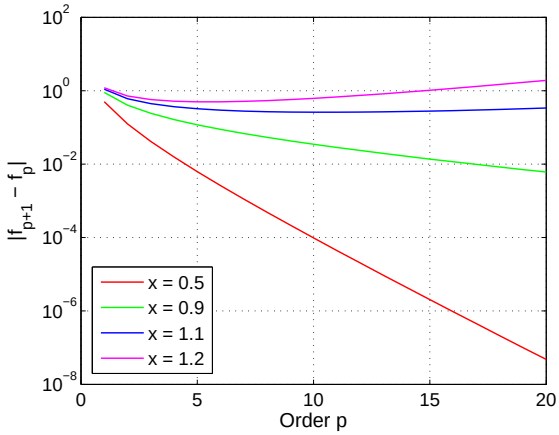

Figure 17. Discrepancy in divergent Taylor series estimates at various x locations for $\mathbf{f(x) = \log(1 + x)}$.

IV.C. Optimal Solution Prediction Using State Transition Tensors

The prior Taylor series examples illustrate two important considerations when predicting optimal solutions during the continuation process. First, the accuracy of predicted solutions is not guaranteed to increase as the order of the STT approximation increases. As such, higher order STT approximations should only be used to improve accuracy if the observed discrepancy decreases. Second, for well-posed hypersonic problems, optimal solutions can likely be predicted to any degree of accuracy using STTs of sufficient order as long as these solutions are sufficiently close to the reference solution. As predictions are made farther away from the reference solution, the accuracy of predictions with increasing order p is governed by the optimal solution space. To highlight the advantages of using higher order STTs, optimal trajectories from a continuation process are also predicted using both first and second order STTs as shown in Figure 18. During this process, the prior converged indirect solution serves as the reference trajectory when calculating the STTs. The errors in predicted altitude and velocity with the corresponding optimal solutions are shown in Figure 19. As desired, the predicted errors from the second order approximation are smaller throughout each trajectory than the errors from the first order approximation. Additionally, the accuracy of these predictions suggests that larger steps can be taken during the continuation process and that the error in predicted solutions can be controlled using higher order STT approximations.

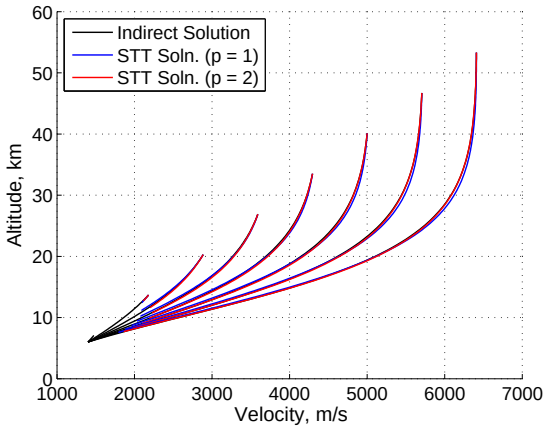

Figure 18. Improved STT-predicted trajectories with order of approximation.

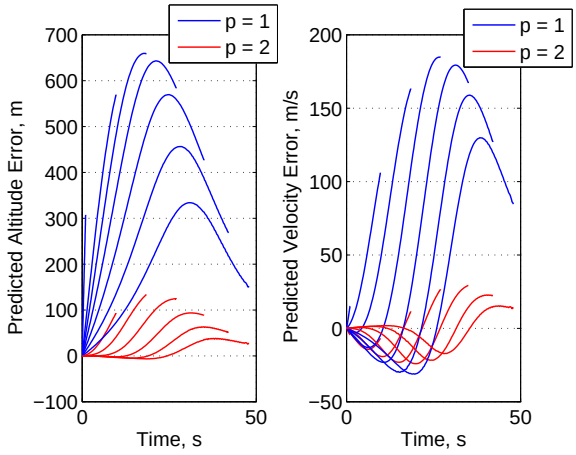

Figure 19. Error in predicted altitude and velocity for various orders of approximation.

As an initial step to automate the continuation process, a methodology has been developed to predict optimal solutions throughout the unconstrained trajectory space using STTs. By controlling the error of this prediction, the methodology dynamically determines the appropriate change in initial conditions originally chosen through trial and error as shown in Figure 7. Furthermore, these predicted solutions are used to provide better initial guesses that were originally formulated using the prior indirect solution. To predict optimal, unconstrained trajectory solutions, boundary conditions must be satisfied at both the initial and terminal points of the trajectory. In general, the STTs provide an analytic mapping between these boundary conditions, enabling an analytic root-solving process to be used to quickly construct predicted optimal solutions. During the unconstrained continuation process, the initial altitude and velocity are incrementally increased to eventually match the desired initial conditions while the terminal conditions remain fixed. As a result, the time of each optimal trajectory increases during this process. However, the STTs are only able to relate changes in initial conditions to changes in terminal conditions with respect to the indirect solution that serves as the reference trajectory during computation of the STTs. This is evident in Figure 18 in which the predicted STT solutions terminate before reaching the desired terminal point. To predict optimal trajectories with longer flight times, the indirect solution used to construct the STTs must be propagated beyond the original terminal conditions by a time dt_f as shown in Figure 20. As such, STTs can then be used to predict the change in initial conditions, dx_i, that yields an optimal trajectory with this increased flight time to the desired terminal state.

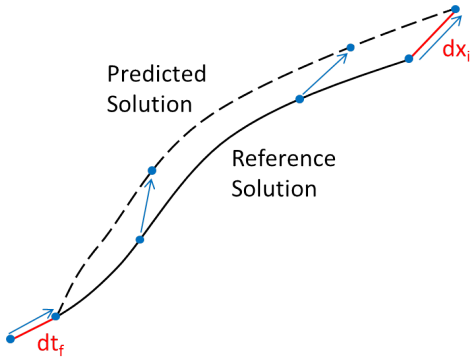

Figure 20. Optimal solution prediction of trajectories with varying initial conditions and fixed terminal conditions.

While this process enables the prediction of optimal trajectories with greater flight times, many solutions in dx_i likely exist for a given choice in dt_f. To enforce a unique solution for dx_i, all changes in the initial state are restricted to the analytic shooting line that connects the initial state of the indirect solution to the desired initial condition as shown in Figure 21. Additionally, as the magnitude of dx_i increases as dt_f increases, the STT-predicted solutions will reside at greater distances from the initial indirect solution. Consequently, the error associated with these predictions will also likely increase. Without knowledge of the exact optimal solutions during this process, the error at the boundary conditions is approximated as the discrepancy between STT-predicted solutions of order p and of order $p+1$. To limit the number of steps during the continuation process, the magnitude of dx_i should be chosen to be as large as possible without violating the maximum approximate error chosen by the designer. However, the proper choice in dt_f that provides this desired magnitude of dx_i is not known *a priori*. As such, an intermediate rapid continuation process is employed to predict optimal solutions along the analytic shooting line using STTs.

During this intermediate process, dt_f is incrementally increased, and the analytic mapping between the initial and final conditions of the STT-predicted solutions is used to converge to the appropriate dx_i along the analytic shooting line. During this process, the STTs provide the opportunity to perform an analytic shooting in which dx_i is iterated to converge to the appropriate terminal conditions. Once the proper change in initial conditions is found, the final time is incrementally increased again from $dt_{f,1}$ to $dt_{f,2}$, and a new initial condition, $dx_{i,2}$, is identified through analytic shooting of the STTs using $dx_{i,1}$ as an initial guess. This process is repeated until dt_f can no longer be increased without exceeding the limit in approximate error chosen by the designer. Note that the intermediate continuation process is naturally regulating. As the indirect solution is propagated beyond the terminal conditions using an error-controlled, adaptive RK scheme, the sequence of dt_f values used by the intermediate continuation process adapts to the dynamics of the optimization problem. If the dynamics of the problem are relatively strong, then smaller steps are taken in dt_f, resulting in more steps during the intermediate continuation process.

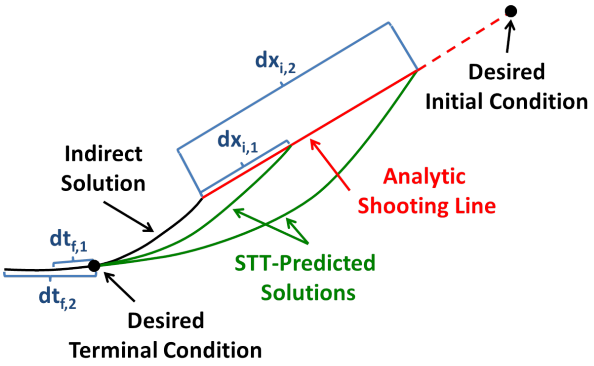

Figure 21. Example of unconstrained analytic shooting using STT predictions.

To maximize computational efficiency, the analytic shooting process is initially performed using an STT of order one, and the approximate error of each solution is calculated using an STT of order two. As the limit in approximate error is reached, the order of the STT used for each calculation can be increased as long as the discrepancy is decreased. If no higher order STT solution is available due to the increase in computational requirements, then the STT-predicted solution is used as an initial guess to converge to a new indirect solution, and the process is repeated. As an example, this process is executed using a maximum STT order of two shown in Figure 22 and a maximum order of three shown in Figure 23. As expected, as the maximum order of the STT increases, larger changes in dx_i can be predicted with the same limit in approximate error. Note that this error includes predicted flight times, and as a result, the STT-predicted solutions to not reach the terminal point. However, these predictions do provide greatly improved initial guesses that enable rapid convergence to the nearby indirect solution also shown in Figures 22 and 23.

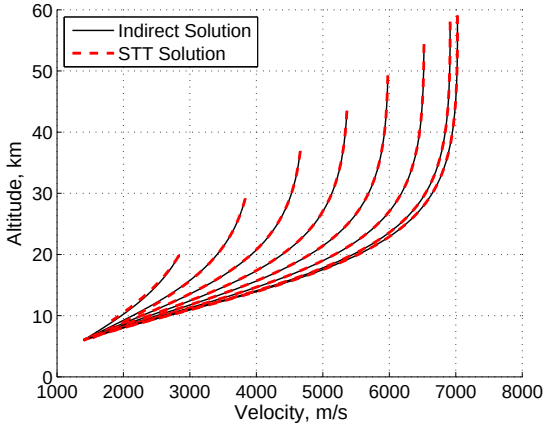

Figure 22. Unconstrained continuation using 2nd order STT predictions.

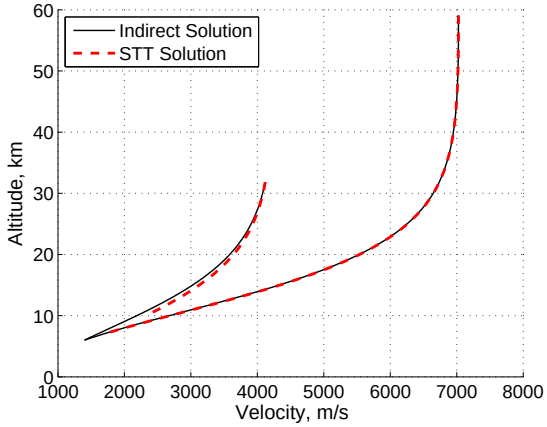

Figure 23. Unconstrained continuation using 3rd order STT predictions.

V. Addressing Prior Methodology Limitations and Future Work

As shown, the complexities of the prior methodology that relies on the combination of many trajectory optimization techniques including discrete dynamic programming, nonlinear inversion, pseudospectral methods, and continuation of indirect methods are eliminated. Instead, this framework can be collapsed into a design methodology that only relies on the continuation of indirect methods. As a result, the path

cost limitation and sensitivity of the discrete dynamic programming process are eliminated. Additionally, the complexities of pseudospectral methods associated with predetermined node locations, mesh refinement techniques, use of penalty functions in NLP solvers, and identification of discontinuous changes in costates are also eliminated. Alternatively, indirect methods enable the use of single and multiple shooting methods that employ an error-controlled, adaptive RK scheme. As a result, path constraints are directly enforced, eliminating the need for penalty functions, and discontinuous changes in costates that result from these path constraints can be directly observed during the continuation process. The investigation of efficient multiple shooting tools, such as BNDSCO, and symplectic integration techniques is left as future work.[32] Furthermore, the prediction of optimal solutions should be expanded to include various vehicle parameters, environment parameters, and path constraints. During the manual continuation process, solutions are monitored to ensure convergence. For example, the lift of the vehicle is increased in a manner that enables the vehicle to follow path constraints while preventing any switches in control. The automation of this monitoring would also enable automation of the continuation process as optimal solutions are predicted using STTs. Furthermore, the continuation of indirect methods may eliminate the disparity between conceptual design and onboard operations.

The dynamic node arrangement provided by adaptive RK schemes ensures that the dynamics of the trajectory are satisfied to an accuracy specified by the designer. This eliminates the need for mesh refinement techniques that may be required when node locations are predetermined as shown in Figure 2. Additionally, the dynamic node arrangement efficiently clusters nodes in regions of most interest that have strong dynamics, whereas pseudospectral methods cluster nodes near the endpoints with generally weak dynamics for hypersonic missions. The efficient clustering of nodes by adaptive RK schemes could assist onboard real-time trajectory planning operations in which optimal trajectory solutions are only accepted if the dynamics at many locations throughout the entire solution are validated to a certain level of accuracy. Furthermore, the use of penalty functions, internal subproblems, and slack variables by NLP solvers greatly increase the complexity of identifying optimal solutions. As a result, the certification of these algorithms for flight operations would also increase in complexity. Alternatively, indirect methods only require a root-solver and a numerical integration scheme. Onboard calculations such as these have already been planned to support the entry monitor system of the Orion command module.[33] As such, the continuation of indirect methods could potentially bridge the gap between conceptual design and onboard trajectory operations. Further advancements in this approach should be evaluated during conceptual design as well as in a flight-relevant hardware environment to address additional gaps that may remain. For example, additional real-time considerations including estimation of vehicle and atmospheric properties could also be incorporated into the conceptual design process.

As methods are evaluated for both conceptual design and onboard applications, considerations must be made for future advances in ground and onboard computing. In both environments, limited computational resources demand rapid calculation of optimal solutions. For example, the emergence of the massive parallelization of scientific computing provided by graphics processing units (GPU) will likely alter the manner in which conceptual design is performed. While NLP solvers can be parallelized, STTs can be computed with relatively small amounts of memory.[34] This enables the relatively limited but fast memory on a GPU to be utilized for rapid parallel computation of STTs.[35] As the capability of GPUs increases, higher order STTs can be computed in parallel to reduce the number of steps required during the continuation process. Additionally, the analytic shooting line should be evaluated to determine if a more efficient path for the continuation policy can be constructed. Finally, as the balance between serial speed and parallelization matures in future computing, indirect methods can also be adapted through proper division of the problem and use of multiple shooting techniques.

VI. Conclusions

In this investigation, a simplified methodology is developed to perform rapid design space exploration that only relies on the continuation of fast indirect methods. This approach eliminates the complexities associated with the integration of discrete dynamic programming, nonlinear inversion, and pseudospectral methods required by a prior design methodology before the continuation of indirect methods. These optimization techniques are required in the prior methodology due to the challenges associated with constructing a good initial guess to converge to an indirect solution within a constrained design space. In this work, a good initial guess is shown to be easily constructed for a short, unconstrained, optimal trajectory that resides outside of

the design space of interest. As such, solutions within the design space of interest are shown to be efficiently constructed through continuation of indirect methods from this unconstrained solution. During this process, path constraints are introduced at points of tangency with the trajectory, and examples with maximum heat rate and g-loading constraints illustrate that the indirect method is able to converge to optimal solutions as the magnitude of these constraints is modified. For these examples, the continuation process is performed manually. As an initial step toward the automation of the continuation process, state transition tensors are shown to provide accurate predictions of optimal solutions throughout the unconstrained trajectory space. A methodology is developed to perform accurate predictions of trajectories with varying flight times, and the error of these predictions is controlled to identify the required steps during the continuation process. Finally, the use of state transition tensors and continuation of indirect methods could be used for both conceptual design studies and onboard real-time trajectory planning. This approach is easily adaptable to future computing considerations and would serve as an initial attempt to bridge the gap between conceptual design and onboard operations.

References

[1]Acton, D. E. and Olds, J. R., "Computational Frameworks for Collaborative Multidisciplinary Design of Complex Systems," AIAA 98-4942, *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis, MO, 2-4 Sep. 1998.

[2]Olds, J. R., "The Suitability of Selected Multidisciplinary Design and Optimization Techniques to Conceptual Aerospace Vehicle Design," AIAA 92-4791, *4th AIAA/USAF/NASA/OAI Symposium on Multidisciplinary Analysis and Optimization*, Cleveland, OH, 21-23 Sep. 1992.

[3]Braun, R. D., Powell, R. W., Lepsch, R. A., Stanley, D. O., and Kroo, I. M., "Comparison of Two Multidisciplinary Optimization Strategies for Launch-Vehicle Design," *Journal of Spacecraft and Rockets*, Vol. 32, No. 3, 1995.

[4]Brown, N. F. and Olds, J. R., "Evaluation of Multidisciplinary Optimization Techniques Applied to a Reusable Launch Vehicle," *Journal of Spacecraft and Rockets*, Vol. 43, No. 6, 2006.

[5]Perez, R. E., Liu, H. H. T., and Behdinan, K., "Evaluation of Multidisciplinary Optimization Approaches for Aircraft Conceptual Design," AIAA 2004-4537, *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Albany, NY, 30 Aug. - 1 Sep. 2004.

[6]Steinfeldt, B., Theisinger, J., Korzun, A., Clark, I., Grant, M., and Braun, R., "High Mass Mars Entry Descent and Landing Architecture Assessment," AIAA 2009-6684, *AIAA Space 2009*, Pasadena, CA, 14 - 17 Sept. 2009.

[7]Garcia, J. A., Brown, J. L., Kinney, D. J., Bowles, J. V., Huynh, L. C., Jiang, X. J., Lau, E., and Dupzyk, I. C., "Co-Optimization of Mid Lift to Drag Vehicle Concepts for Mars Atmospheric Entry," AIAA 2010-5052, *10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, Chicago, IL, 28 Jun. - 1 Jul. 2010.

[8]Bibeau, R. and Rubenstein, D., "Trajectory Optimization for a Fixed-Term Reentry Vehicle Using Direct Collocation and Nonlinear Programming," AIAA-2000-4262, *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Denver, CO, 14-17 Aug. 2000.

[9]Herman, A. and Conway, B., "Direct Optimization Using Collocation Based on High-Order Gauss-Lobatto Quadrature Rules," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 3, 1996.

[10]Josselyn, S. and Ross, I. M., "Rapid Verification Method for the Trajectory Optimization of Reentry Vehicles," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 3, 2003.

[11]Grant, M. J. and Mendeck, G. F., "Mars Science Laboratory Entry Optimization Using Particle Swarm Methodology," AIAA 2007-6393, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Hilton Head, SC, 20-23 Aug. 2007.

[12]User's Guide for SNOPT Version 7: Software for Large-Scale Nonlinear Programming, University of California, San Diego, CA, Stanford University, Stanford, CA, 16 Jun. 2008.

[13]Lafleur, J. and Cerimele, C., "Mars Entry Bank Profile Design for Terminal State Optimization," AIAA 2008-6213, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Honolulu, HI, 18-21 Aug. 2008.

[14]Grant, M. J., Clark, I. G., and Braun, R. D., "Rapid Entry Corridor Trajectory Optimization for Conceptual Design," AIAA 2010-7810, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Toronto, Ontario, Canada, 2-5 Aug. 2010.

[15]Otero, R. E. and Braun, R. D., "The Planetary Entry Systems Synthesis Tool: A Conceptual Design and Analysis Tool for EDL Systems," IEEEAC 1331, *2010 IEEE Aerospace Conference*, Big Sky, MT, Mar. 2010.

[16]Kirk, D. E., *Optimal Control Theory: An Introduction*, Prentice-Hall, Inc., 1970.

[17]McCausland, I., *Introduction to Optimal Control Theory*, John Wiley and Sons, Inc., 1969.

[18]Grant, M. J., Clark, I. G., and Braun, R. D., "Rapid Simultaneous Hypersonic Aerodynamic and Trajectory Optimization Using Variational Methods," *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Portland, OR, 8-11 Aug. 2011 (to be published).

[19]Darby, C. L., Hager, W. W., and Rao, A. V., "An hp-Adaptive Pseudospectral Method for Solving Optimal Control Problems," *Optimal Control Applications and Methods*, Vol. 32, 2010.

[20]Zhao, Y. and Tsiotras, P., "Density Functions for Mesh Refinement in Numerical Optimal Control," *Journal of Guidance, Control, and Dynamics*, Vol. 34, No. 1, 2011.

[21]Gong, Q., Ross, I. M., Kang, W., and Fahroo, F., "Connections Between the Covector Mapping Theorem and Convergence of Pseudospectral Methods for Optimal Control," *Journal of Computational Optimization and Applications*, Vol. 41, No. 3, 2008.

[22]Gong, Q., Ross, I. M., Kang, W., and Fahroo, F., "On the Pseudospectral Covector Mapping Theorem for Nonlinear Optimal Control," *45th IEEE Conference on Decision and Control*, San Diego, CA, 13-15 Dec. 2006.

[23]Bryson, A. E. and Ho, Y.-C., *Applied Optimal Control*, Taylor and Francis, 1975.

[24]Elsgloc, L. D., *Calculus of Variations*, Dover Publications, Inc., 2007.

[25]Petrov, I. P., *Variational Methods in Optimum Control Theory*, Academic Press Inc., 1968.

[26]Mendeck, G. F. and Carman, G. L., "Guidance Design for Mars Smart Landers Using The Entry Terminal Point Controller," AIAA-2002-4502, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Monterey, CA, 5-8 Aug. 2002.

[27]Park, R. S. and Scheeres, D. J., "Nonlinear Mapping of Gaussian Statistics: Theory and Applications to Spacecraft Trajectory Design," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 6, 2006.

[28]Turner, J. D., Majji, M., and Junkins, J. L., "High-Order State and Parameter Transition Tensor Calculations," AIAA-2008-6453, *AIAA/AAS Astrodynamics Specialist Conference*, Honolulu, HI, 18-21 Aug. 2008.

[29]Sengupta, P., Vadali, S. R., and Alfriend, K. T., "Second-Order State Transition for Relative Motion Near Perturbed, Elliptic Orbits," *Celestial Mechanics and Dynamical Astronomy*, Vol. 97, No. 2, 2007.

[30]Russell, R. P. and Lantoine, G., "A Hybrid Differential Dynamic Programming Algorithm for Robust Low-Thrust Optimization," AIAA-2008-6615, *AIAA/AAS Astrodynamics Specialist Conference*, Honolulu, HI, 18-21 Aug. 2008.

[31]Russell, R. P. and Lantoine, G., "Optimal Control of Relative Motion in Arbitrary Fields: Application at Deimos," AAS 10-313, *Kyle T. Alfriend Astrodynamics Symposium*, Monterey, CA, 17-19 May 2010.

[32]Oberle, H. J. and Grimm, W., BNDSCO: A Program for the Numerical Solution of Optimal Control Problems, University of Hamburg, Hamburg, Germany, Oct. 2001.

[33]Baird, D., Grant, M., Kadwa, B., Gillespie, E., Matthews, D., Penny, W., Zak, T., and Bihari, B., "Orion Entry Display Feeder and Interactions with the Entry Monitor System," AIAA 2010-8062, *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Canada, 2-5 Aug. 2010.

[34]Dayde, M., "Parallel Algorithms for Nonlinear Programming Problems," *Journal of Optimization Theory and Applications*, Vol. 61, No. 1, 1989.

[35]Arora, N., Russell, R. P., and Vuduc, R. W., "Fast Sensitivity Computations for Trajectory Optimization," AAS 09-337, *AIAA/AAS Astrodynamics Specialist Conference*, Pittsburg, PA, 9-13 Aug. 2009.