

Multi-Fidelity Trajectory Optimization with Response Surface-Based Aerodynamic Prediction

Colonno, M. R.¹

*Stanford University, Palo Alto, California, Zip 94305 and
Space Exploration Technologies, Hawthorne, CA, 90250*

Reddy, S.²

Space Exploration Technologies, Hawthorne, California, 90250

and

Alonso, J. J.³

Stanford University, Palo Alto, California, Zip 94305

A methodology for the optimization of general atmospheric trajectories is presented. The trajectory optimization problem is formulated as a constrained minimization problem then discretized for use with a Gauss pseudospectral method. Assuming a robust optimizer, the fidelity of an optimum solution is limited by the accurate prediction of input physics. These include aerodynamic properties, atmospheric properties, wind velocity and direction, and mass properties. In addition, the free stream conditions at which aerodynamic simulations are performed are, in turn, a function of the trajectory and vary with mission. Here, a Kriging surface method is proposed which allows for the efficient adaptive sampling of subsequent physical data for increased levels of fidelity. This adaptive approach to high-fidelity physical data decreases the cost and cycle time required for mission planning or day-of-launch criteria. Results from an ascent and reentry example are presented and discussed in the context of current methods.

Nomenclature

a	=	semimajor axis
C_D	=	drag coefficient
C_L	=	lift coefficient
d	=	unit vector in drag direction
e	=	eccentricity
EIF	=	expected improvement function
f	=	generic function
F	=	thrust
g	=	gravitational acceleration vector
g_0	=	gravitational acceleration at planet's surface
h	=	altitude
i	=	inclination
I_{sp}	=	specific impulse
J	=	optimization objective function
l	=	length scale
l	=	unit vector in lift direction
M	=	Mach number
m	=	mass

¹ Chief Aerodynamic Engineer, Space Exploration Technologies, Member AIAA

² Dynamics Engineer, Space Exploration Technologies, Member AIAA

³ Associate Professor, Aeronautics & Astronautics, Stanford University, Member AIAA

p	=	pressure
\mathbf{p}	=	vector of orbital parameters
q	=	dynamic pressure ($\rho_\infty V_\infty^2/2$)
R	=	planetary radius
Re	=	Reynolds number
Re'	=	Reynolds number per unit length
\mathbf{r}	=	unit vector in the radial (vertical) direction
S	=	reference area for aerodynamic coefficients
s	=	interpolation error
T	=	temperature
$[\mathbf{T}]$	=	coordinate transformation matrix
t	=	time
t_b	=	burn time
\mathbf{u}	=	thrust direction unit vector
V	=	velocity
W	=	control stability constraint
x	=	axial direction
\mathbf{x}	=	position vector
\mathbf{y}	=	vector in response surface variable space
α	=	angle of attack
β	=	yaw angle
γ	=	ratio of specific heats
\mathbf{w}	=	rotation rate vector
μ	=	planetary gravitational constant = Gm_{planet}
ρ	=	density
Ω	=	longitude of the ascending node
ω	=	argument of periapsis
ν	=	true anomaly

Subscripts

0	=	initial state
CP	=	center of pressure
G	=	gimbal location (thrust vector control)
f	=	final state
∞	=	ambient (free stream)
E	=	Earth
ECI	=	Earth-centered inertial coordinates
o	=	orbital value
P	=	propellant
PL	=	payload

Superscripts

T	=	transpose
\bullet	=	rate ($\partial/\partial t$)
$\hat{\cdot}$	=	predicted by response surface

I. Introduction

THE high cost of access to space places great emphasis on the efficient design of launch vehicles and the subsequent optimization of trajectories for maximum performance. With a fixed launch vehicle design, ascent trajectories are optimized for a given payload and desired set of orbital parameters. This optimization process necessarily depends on a robust method for dynamic optimization and a multidisciplinary set of physical input data. The latter generally includes aerodynamic properties of the vehicle over a wide Mach range, detailed mass property

data which is constantly changing as propellant mass is expelled, and atmospheric data (including wind conditions) which may be taken just before launch or reentry. All of these sources of data have associated uncertainties which must be accounted for in establishing the mission margin and day-of-launch liftoff criteria. Since margin must be established to cover all reasonable uncertainties, increased uncertainty results in smaller mass delivered to orbit.

Algorithms to solve general dynamic optimization (or applied optimal control) problems have evolved over time from the original work of Byrson^{1,2} to many more recent solution methods^{3,4,5,7,14}. Here, we use a Gauss pseudospectral method (GPM) of Ref. 3 through the TOMLAB GPOCS software¹⁹. GPOCS is an example of a *direct* solution method in which the continuous-time optimal control problem is transcribed into a nonlinear programming problem (NLP) which is subsequently solved using any well-established algorithms (SNOPT⁶ in the case of GPOCS). This is contrasted with an *indirect* solution method in which the first-order necessary conditions for optimality are analytically derived from the optimal control problem itself and the subsequent boundary-value problem solved numerically for extremal paths. Current numerical methods for trajectory optimization are discussed in greater detail in Refs. 4,5,14 among others.

With a robust dynamic optimization tool, the fidelity-limiting aspect of trajectory optimization shifts to the physical input data. In this paper we are primarily concerned with: 1) the quantification and minimization of uncertainties in physical data through adaptive sampling, 2) maximizing flight performance over a range of mission profiles, and 3) providing a rapid-response means by which to establish liftoff criteria with manageable computational cost. Items (1) and (2) are addressed during mission planning while (3) is dependent upon atmospheric conditions evaluated before launch. The goal of this work is to provide a means by which a trajectory can be optimized to a specified level of uncertainty with minimal computational cost or (in many cases equivalently) the minimum cycle time.

II. Problem Formulation: Ascent

Dynamic optimization or applied optimal control problems are generally formulated as constrained minimizations in which the equations of motion are treated as constraint functions. A vector of continuous control parameters are sought which extremize a function of the state variables, often at the end of the time period of interest. This general problem statement is then discretized for use with an appropriate numerical method to perform the actual optimization. The sections below detail the objective and constraint functions for a general three degree of freedom (3-DOF) launch vehicle ascent problem.

A. Equations of Motion

Here, we begin with the 3-DOF equations of motion in Earth-center-inertial (ECI) coordinates (named by convention, this formulation is general to any central gravitational field) for a general multi-stage vehicle, Eqs. (1). When properly scaled, Eqs. (1) form a non-stiff set of ten coupled ordinary differential equations (ODEs). These were chosen for convenience and numerical performance, but any reference frame can be used.

$$\begin{aligned}
 \frac{\partial \mathbf{x}_{ECI}}{\partial t} &= \mathbf{V}_{ECI} \\
 \frac{\partial \mathbf{V}_{ECI}}{\partial t} &= \frac{1}{m} (F_i \mathbf{u} + C_D q_\infty S \mathbf{d} + C_L q_\infty S \mathbf{l} + \mathbf{g}(\mathbf{x})) \\
 \frac{\partial m}{\partial t} &= -\dot{m}_i \\
 \frac{\partial \mathbf{u}}{\partial t} &= \mathbf{w}
 \end{aligned} \tag{1}$$

The vector \mathbf{u} is the thrust direction which, in the 3-DOF approximation, we assume is also the vehicle orientation. Here, we use the rate of change of \mathbf{u} , or \mathbf{w} , as the control parameter making \mathbf{u} a “pseudostate”. This formulation allows constraints to easily be placed on \mathbf{w} continuously or at the endpoints of each stage’s flight. The thrust and mass flow vary for each stage, denoted by the subscript i . For a multi-stage vehicle, discontinuities occur in the vehicle mass between stages as the previous stage’s empty structural mass is released which require a multi-phase formulation of the optimal control problem. These are accounted for numerically with link conditions on each of the state variables, Eqs. (2). The lift and drag forces act parallel to and perpendicular to the total relative air velocity (\mathbf{V}_{tot}), respectively. These directions are denoted by the unit vectors \mathbf{d} and \mathbf{l} in Eqs. (1). The gravitational

acceleration, $\mathbf{g}(\mathbf{x})$, can be as detailed as required, but the spherical approximation (\mathbf{g} directed along $-\mathbf{r}$) and J2 oblateness correction of the WGS84¹⁷ model are commonly used.

$$\begin{aligned}
\Delta \mathbf{x}_{\text{ECI}}^{i \rightarrow j} &= 0 \\
\Delta \mathbf{V}_{\text{ECI}}^{i \rightarrow j} &= 0 \\
\Delta m^{i \rightarrow j} &= -m_{S_i} \\
\Delta \mathbf{u}^{i \rightarrow j} &= 0
\end{aligned} \tag{2}$$

To complete the problem formulation the objective, boundary conditions (constraints which are evaluated at the initial and final times of each stage), and path constraints (constraints which are evaluated continuously in time) must be defined. These are discussed in the sections below. We seek the vector $\mathbf{u}(t)$ which minimizes J subject to the equations of motion (EOM) and the constraints discussed below.

B. Objective

For ascent trajectories, the objective is simple: minimize the propellant mass required to deliver the payload mass, m_{PL} , to the desired orbit. This is equivalent to maximizing the mass at the end of flight (burnout mass), stated as a minimization problem in Eq. (3). Note that the final time t_f is, therefore, not known a priori. For trajectories which do not include a coast phase this is equivalent to a minimum time problem. It is not known a priori if $m(t_f)$ is positive; the mission may not be possible and the optimization procedure must account for cases in which the vehicle has insufficient performance.

$$J = -m(t_f) \tag{3}$$

C. Boundary Conditions

The initial conditions $\mathbf{x}_{\text{ECI}}(t_0)$ and $\mathbf{V}_{\text{ECI}}(t_0)$ are set by the location of the launch site, with initial velocity equal to the local rotational speed of the planet (if any). $m(t_0)$ is simply the gross liftoff mass. The initial orientation vector is assumed to be vertical, $\mathbf{u}(t_0) = \mathbf{r}$. The boundary conditions at t_f are more complex, requiring insertion into an orbit with the desired set of orbital elements $\mathbf{p} = \{a, i, e, \Omega, \omega\}^T$, summarized below. These are transformed into ECI coordinates (and vice versa) at t_f to establish $\mathbf{x}_{\text{ECI}}(t_f)$ and $\mathbf{V}_{\text{ECI}}(t_f)$, while $m(t_f)$ is unconstrained as the target of the optimization. Note that the orbital elements listed above lack the true anomaly, ν , as we assume the insertion can take place anywhere on the desired orbit. Hence, the boundary conditions $\mathbf{x}_{\text{ECI}}(t_f)$ and $\mathbf{V}_{\text{ECI}}(t_f)$ are not unique.

During a stage separation event a brief coast phase is typically required to fully shutdown the previous stage's propulsion system, engage the separation system, and activate the current stage's propulsion system. Depending on the stage sizing and desired orbit, burnout may occur at altitudes where aerodynamic forces are still significant. Without any thrust to counter aerodynamic moments, the new stage can rotate off course or collide with the previous stage during separation. Hence, limits on α at burnout may be placed on each stage. No constraints are placed on $\mathbf{u}(t_f)$, though $\mathbf{w} = 0$ may be desirable at t_f and at stage burnouts.

D. Path Constraints

Any number of additional constraints can be placed on the flight depending on the vehicle used, the requirements of the payload, and the mission. These constraints may be placed directly on the state or control variables or on functions thereof. Typical constraints include maximum axial and lateral accelerations, maximum $q_\infty \alpha$ product, maximum stagnation point heat transfer, and rotation rates during flight and at stage burnouts. The normalization of the orientation vector \mathbf{u} is also treated as a path constraint. Refer to the launch vehicle example below for some specific constraint examples.

A simplified stability criterion is also included to account for limitations on the vehicle's control system. When neglecting the rotational DOFs, no limits on the moments about the vehicle center of gravity (CG). When the transient behavior of the control system and the vehicle structural dynamics are included, the rate of change of the moment is often the stability-limiting criterion. For ascent trajectories, this is primarily a function of the wind shear or rate of change of the aerodynamic moment. A single instantaneous criterion is insufficient because stability is a function of both the wind shear's duration and amplitude. Changes in wind velocity which have high amplitude but

short apparent duration (product of the vehicle's vertical speed and rate of change with altitude) or low amplitude and long duration are generally not a stability concern.

Here, we use the stability criterion $W_{ij}(t, \tau_j)$, Eq. (4), an integration over a pulse duration τ which must be defined through an independent control system analysis. The integrated lift force rate is normalized with respect to the maximum side force which can be exerted by the control system. No specification is made about the type of control system, which may involve trust vector control (TVC), aerodynamic surfaces, or both. The length scales l_{CP} and l_G are representative of the moment arm between the center of pressure (CP) and the CG and the point at which control forces are applied (for most launch vehicle profiles, the total vehicle or stage length is applicable) and the CG, respectively. In general, the CP location itself is a function of M_∞ and α but is presently assumed to be constant over the relatively short duration τ . The integration duration may be a function of flight time to account for changes in the vehicle mass properties. Alternatively, several different values of τ could be used to form several different constraints. (The subscript ij refers to the i^{th} stage and j^{th} pulse duration.) It is worth noting that the $|\partial C_L / \partial t|$ term is a function of both the vehicles aerodynamic properties and wind conditions. Application of the stability constraint W_{ij} assumes a passively unstable vehicle with CP forward of the CG. For passively stable vehicles or stages this constraint is neglected.

$$\frac{S_i}{F_i \sin(\delta_{\max_i})} \left(\frac{l_{CP}}{l_G} \right)^{t+\tau/2} \int_{t-\tau/2}^{t+\tau/2} q_\infty(t) \left| \frac{\partial C_L}{\partial t} \right| dt \leq W_{ij}(t, \tau_j) \quad (4)$$

III. Physical Data

Evaluation of Eqs. (1) requires a set of multidisciplinary physical data, including aerodynamic properties, wind data and a gravitational model. We will assume that gravitational acceleration as a function of position, $\mathbf{g}(\mathbf{x})$, is known a priori. In addition, we will assume that the vehicle liftoff and burnout masses are measured with associated uncertainties. (It is noted that a 6-DOF simulations would additionally require the moment of inertia matrix as a function of vehicle mass.) Aerodynamic data and atmospheric conditions, including the wind velocity, are required. The sections below detail the data required in each category and the limitations on currently available tools.

A. Aerodynamic Properties

The aerodynamic properties of the launch vehicle are typically the most expensive to determine and cannot, in general, be refined adaptively close to the time of launch. For a 3-DOF simulation the vehicle is treated as a point mass having a lift and drag coefficients which are functions, in general, of M_∞ , α , and Re . For vehicles which are not axisymmetric, the sideslip angle β would additionally be required and some assumption would have to be made about the vehicle's roll orientation during flight. Here, the Reynolds number per unit length, Re' , is used for convenience since it is a function of the free stream coefficients only (equivalent to taking the length scale in $Re = 1.0$ m). In a 6-DOF analysis, the aerodynamic side (yawing) force and moments or CP location would be required to evaluate the rotational degrees of freedom. Summarizing, the 3-DOF optimization requires $C_L(M_\infty, \alpha, Re')$ and $C_D(M_\infty, \alpha, Re')$ data throughout the flight.

One of the greatest challenges of launch vehicle aerodynamics is the large range of flow conditions over which data is required. This includes values of M_∞ from zero to 25 or greater, widely varying flow physics, and possibly large values of α early in flight due to wind. Fortunately from a trajectory optimization standpoint, the density (and hence dynamic pressure) typically falls to very small values before the physics of the free stream gases changes from a continuum fluid. However, the trajectory is not known a priori and a continuous, differentiable evaluation of the aerodynamic forces is required for an optimization algorithm such as that of GPOCS. Whether aerodynamic data comes from wind tunnel testing or computational fluid dynamics (CFD), it is unlikely to span the required range of flow conditions.

Typically, data is available in the subsonic and lower supersonic M_∞ range while data at very high M_∞ has to extrapolated. Care must be taken to examine the $M_\infty \rightarrow 0$ and $M_\infty \rightarrow \infty$ limits for all aerodynamic coefficients to avoid infinite or negative values. Though $q_\infty \rightarrow 0$ at these limits, numerical instabilities or inaccurate solutions may result. Refer to the example in the following section for a specific discussion. Ascent trajectories with typical launch vehicle properties tend to result in maximum q_∞ conditions in the low supersonic range while reentry trajectories tend to experience maximum q_∞ conditions at much high M_∞ , placing additional emphasis on high-fidelity aerodynamic data in the hypersonic range.

B. Atmospheric Data

Atmospheric data includes physical properties density, pressure, and temperature as well as wind velocity. All of these data are generally assumed to be functions of altitude only, but in general may be functions of spherical position (latitude and longitude) as well. These data are typically sampled as close to the liftoff (or reentry) time as possible and often carry the greatest uncertainties. While the wind conditions have a minimal effect on the performance of the trajectory (ΔV_D), they have a profound effect on control stability. A full 6-DOF simulation is required to accurately determine response to a given wind profile, but a maximum wind shear criterion is typically included as a day-of-launch constraint and can be determined through a 3-DOF simulation (discussed above).

It is worth noting that q_∞ , the drag direction vector \mathbf{d} , and lift direction vector \mathbf{l} are all functions of the wind velocity. The total relative air velocity, \mathbf{V}_{tot} , is the vector sum of $-\mathbf{V}_{ECI}$ and \mathbf{V}_w while q_∞ is based on the magnitude of \mathbf{V}_{tot} , Eq. (5). Wind data is specified relative to the planet's surface (spherical coordinates) and has to be transformed into ECI coordinates, taking into account the planet's rotation. The vector \mathbf{d} is given by Eq. (7) and the vector \mathbf{l} by Eq. (8). The angle of attack, α , is given by Eq. (9).

$$\mathbf{V}_{tot} = [\mathbf{T}]_{E \rightarrow ECI} \mathbf{V}_w - \mathbf{V}_{ECI} \quad (5)$$

$$q_\infty = \frac{1}{2} \rho_\infty |\mathbf{V}_{tot}|^2 \quad (6)$$

$$\mathbf{d} = \frac{\mathbf{V}_{tot}}{|\mathbf{V}_{tot}|} \quad (7)$$

$$\mathbf{l} = -(\mathbf{u} \times \mathbf{d}) \times \mathbf{d} \quad (8)$$

$$\alpha = \cos^{-1}(-\mathbf{u} \cdot \mathbf{d}) \quad (9)$$

Evaluation of M_∞ number requires the local speed of sound and the thrust is a function of the local ambient pressure. Using a one-dimensional gasdynamic approximation, the thrust at altitude is given by Eq. (10), where p_d is the ambient pressure at which the design thrust F_d .

$$F(p_\infty) = F_d + A_e(p_d - p_\infty) \quad (10)$$

IV. Response Surface Approach

Due to the critical dependence of the solution to the physical data discussed above, a methodology which bounds uncertainties throughout the flight and provides a mechanism for adaptive sampling of physical data is advantageous. It is worth emphasizing that the physical data and optimal trajectory are *codependent*. The conditions at which the aerodynamic coefficients are analyzed, for example, depend upon the trajectory and vice versa. This codependence inevitably involves iterating to a solution which is generally expensive and may be required on a mission-by-mission basis. In addition, the traditional methodology provides no mechanism for adaptively refining the solution and supporting physical data where the greatest improvement in the optimal trajectory (performance or accuracy) may be had. Kriging surfaces, frequently used in global optimization applications, are ideally suited to this task. Here we use the Matlab¹⁶ Design and Analysis of Computer Experiments (DACE) Kriging Toolbox¹³, based on the development discussed in Refs. 11,15.

In order to utilize Kriging surfaces effectively, data samples much be chosen in a sufficiently random manner. Ordered sets of data, such as Mach or α sweeps, tend to perform poorly due to the use of radial basis functions. A wide variety of methods are available for sampling in design of experiments (DOE) applications. Here we choose Latin Hypercube sampling, frequently used in response surface modeling. This method is convenient in (but not limited to) applications which contain only simple bounds on the independent variables.

The number of data samples is constrained by the available budget (computational or otherwise). The "10- n rule", where n is the independent variable dimensionality, is often used as a general guideline but more data should be used if available. With the data evaluated at the sample points, response surface creation is generally inexpensive. Correlation and regression functions must be chosen in addition to bounds other parameters discussed

in Refs. 11,13,15. Evaluating the resulting surface quality can be done in a number of ways; here we choose minimization of the standardized cross-validated residual (SCVR), based on the “one-out error”. The error associated with the value (and gradient) at a given point in the surface, \mathbf{y}_i , can be estimated from the Kriging surface. The one-out error, denoted s_{-i} , is defined in Eq. (11) as the difference between the actual value of the function of interest at point \mathbf{y}_i , $f(\mathbf{y}_i)$, and the value predicted by a response surface which does not include the data point, $f_{-i}'(\mathbf{y}_i)$, normalized by the estimated error at that location, $s'(\mathbf{y}_i)$. The parameters mentioned above are tuned to minimize the maximum of the SCVR with the initial sample population via any appropriate optimization method.

$$\text{SCVR} = \frac{f(\mathbf{y}_i) - f_{-i}'(\mathbf{y}_i)}{s_{-i}(\mathbf{y}_i)} \quad (11)$$

Finally, the concept of the expected improvement function (EIF), Eq. (12), is introduced for use in adaptive refinement. The EIF, designed for use with global optimization problems, is used to identify subsequent sample points that have the greatest potential to resolve a minimum in f . The EIF takes into account both the value of the function and the estimated error, yielding new sample points which have small f (minimization is assumed) and large s' . Since numerous local maxima of the EIF will occur in regions between the current samples, a genetic algorithm (GA) is generally best suited to locate the global maximum of the EIF. Here, we use a standard GA for engineering applications developed in Refs. 9,10 for adaptive sampling. It is noted that in this application the EIF will be evaluated along a path through the response surface resulting in a one-dimensional function for the EIF. Hence, a simpler minimization method (relative to a GA) would likely also be effective.

$$\text{EIF} = \frac{s'}{\sqrt{2\pi}} \left(z \int_{-\infty}^z e^{-\frac{x^2}{2}} dx + e^{-\frac{z^2}{2}} \right) \quad (12)$$

$$z \equiv \frac{f_{\min} - f'}{s'}$$

While the EIF is used with objective function in an optimization application, here we apply it to identify regions of physical data where subsequent analyses yield the greatest reduction in uncertainties. Using aerodynamics as an example, the EIF based on q_∞ , EIF($-q_\infty, s'$) is an obvious choice. After generating an initial set of aerodynamic response surfaces and an initial optimum trajectory for a given mission, adaptive sampling based on subsequent maxima of EIF($-q_\infty, s'$) will improve regions of aerodynamic data which have both large values of q_∞ and s' . This allows expensive analyses to be clustered around point where increased accuracy has the greatest benefit to the simulation of the trajectory.

The EIF provides a flexible and robust tool for refinement across any of the physical data or a combination thereof. It may be of benefit to refine analyses around the maximum stability factor (W), wind velocity, stage separation events, or where any other limiting atmospheric flight load conditions take place. The response surface approach serves to decouple, to some extent, trajectory optimization from the analyses which provide the physical data.

V. Multi-Fidelity Approach

The fidelity of the physics used in the optimization process has a profound effect on the computational cost. Physical models must be evaluated at every control point in every iteration for use in the equations of motion. In addition convergence rate of Gauss pseudospectral method used in this work, while faster and more forgiving than other methods previously used, is strongly affected by the quality of the initial guess. Hence, it is advantageous to “tier” the fidelity upward sequentially, using lower-fidelity physical models to initialize more expensive, higher-fidelity optimizations. It should be noted that this process does not guarantee convergence to a global optimum. While it is intuitive to expect that a higher-fidelity optimum path lies close in the variable space to a slightly lower-fidelity path, the author is not aware of any general proof to this effect. This is an acknowledged limitation of the method presented here, though experimentation with numerous examples did not produce any optimum paths which were different when fidelity was tiered versus initialized at the highest level.

VI. Launch Vehicle Example

For the purposes of demonstrating the method, consider a two-stage launch vehicle with the properties summarized below. These values are typical of a medium-lift launch (several thousand kg to low Earth orbit) vehicle using RP-1 (purified kerosene) and liquid oxygen (LOX). We'll consider a launch from the site of this conference (numerical demonstration purposes only) to a 185 km circular polar orbit. The various parameters and constraints listed below are intended to be fictional but rational.

Mass Properties:

$$\begin{aligned}
 m_{S,1} &= 13,000 \text{ kg} \\
 m_{S,2} &= 2,900 \text{ kg} \\
 m_{P,1} &= 240,000 \text{ kg} \\
 m_{P,2} &= 5,000 \text{ kg} \\
 m_{PL} &= 2,000 \text{ kg}
 \end{aligned}$$

Propulsion:

$$\begin{aligned}
 F_1 &= 3,800 \text{ kN} \\
 I_{sp,1} &= 270 \text{ s (sea level)} \\
 F_2 &= 483.0 \text{ kN} \\
 I_{sp,2} &= 325 \text{ s (vacuum)}
 \end{aligned}$$

Launch Site (Reno, NV):

$$\begin{aligned}
 \text{Latitude} &= 39^\circ 31' 20.86'' \text{ N} \\
 \text{Longitude} &= 119^\circ 46' 45.58'' \text{ W} \\
 \text{Elevation} &= 1350 \text{ m (4430 ft)}
 \end{aligned}$$

Orbit:

$$\begin{aligned}
 a &= 6,563.1 \text{ km (185 km altitude)} \\
 i &= 90^\circ \text{ (polar)} \\
 e &= 0.0 \text{ (circular)} \\
 \Omega &= \text{unconstrained} \\
 \omega &= \text{unconstrained}
 \end{aligned}$$

Constraints:

$$\begin{aligned}
 |\dot{w}| &\leq 10 \text{ deg/s} \\
 q_\infty \alpha &\leq 15.0 \text{ atm-deg}
 \end{aligned}$$

Stability Criterion:

$$\begin{aligned}
 W_{\max} &= 5.0 \\
 \tau &= 3.0 \text{ s}
 \end{aligned}$$

Five levels of fidelity are chosen, summarized below, including adaptive sampling of input physical data. The sections below detail the steps of a multi-fidelity trajectory optimization and the adaptively sampling process. Finally, selected results are presented and discussed.

Fidelity Levels:

Level	Atmosphere	Wind	Aerodynamics	Adaptive Sampling
1	none	None	None	No
2	exponential	None	constant $C_D = 1.0$, No Lift	No

3	exponential	Response Surfaces	constant $C_D = 1.0$, No Lift	Yes
4	MSISE-90 splines	Response Surfaces	Response Surfaces	Yes
5	MSISE-90 splines	Response Surface	Response Surfaces	N/A

A. Aerodynamic Response Surface Creation

Here, we will assume aerodynamic coefficients C_D and C_L are given by analytical expressions Eqs. (13) and (14) respectively. This allows the response surface estimation to be compared directly to the function, but it is worth emphasizing that this would not be possible in general. Note that Eqs. (13-14) has finite, positive limits as $M_\infty \rightarrow 0$ and $M_\infty \rightarrow \infty$. Note that we have implicitly assumed an axisymmetric vehicle and neglected the dependence of C_D and C_L on Re .

$$C_D = \cos(\alpha) \left(\frac{0.4750M_\infty^2 - 0.7127M_\infty + 0.3049}{M_\infty^2 - 1.914M_\infty + 1.042} \right) \quad (13)$$

$$C_L = 80 \sin(\alpha) \left(\frac{49.50M_\infty^2 + 0.1317M_\infty + 46.08}{M_\infty^3 + 593.9M_\infty^2 - 528.6M_\infty + 962.1} \right) \quad (14)$$

The aerodynamic response surfaces must be generated as if the Eqs. (19-20) were unknown functions. With no a priori knowledge of the path taken through $\{M_\infty, \alpha\}$ space, an initial sample of 50 points (CFD solutions or wind tunnel tests) is used and the function is sampled via Latin Hypercube over the range $0 < \alpha < 10^\circ$ and $0 < M_\infty < 3.0$. The resulting initial response surfaces, resulting from a minimization of the SCVR as discussed above, are plotted in Figs. (1-2) against the exact functional forms. Details of the response surfaces are summarized below. If CFD data were used, these response functions could correspond to inviscid or viscous solutions depending on the level of fidelity desired for the initial estimate. Since the fidelity levels listed above do not make any distinction between the fidelity of the aerodynamic data itself, we'll assume these are evaluated at the highest level of fidelity.

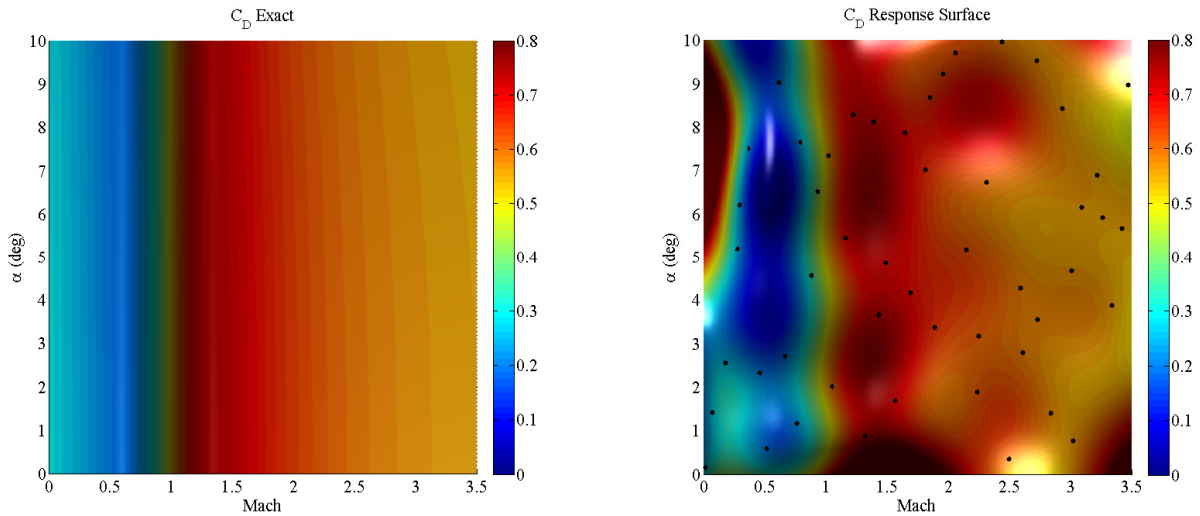


Figure 1. Comparison of C_D response surface (right) with Eq. (13), (left). Note significant inaccuracies in C_D throughout the domain shown due to the strong dependence on M_∞ .

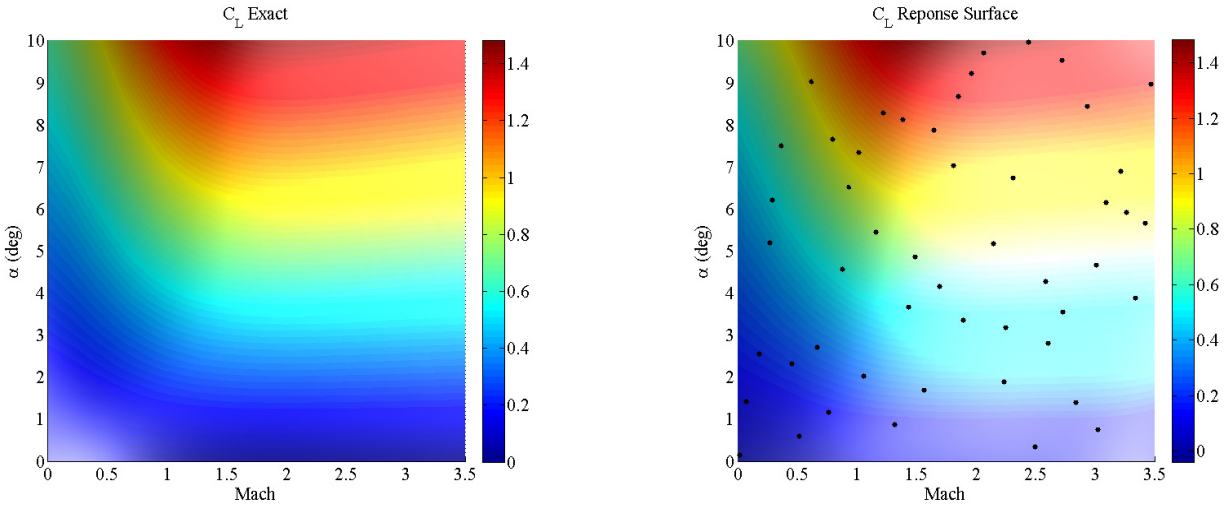


Figure 2. Comparison of C_L response surface (right) with Eq. (14), (left). C_L is very smooth and nearly linear in α for small α resulting in a close comparison with the initial sample.

B. Wind Data Response Surface Creation

A fictional but representative set of balloon-sampled wind data was used for the purposes of this case study, plotted below in Figs. (3). It was assumed that wind is horizontal and its magnitude and direction are a function only of altitude. All data were sampled in altitude increments of approximately 0.25 km up to a maximum altitude of approximately 22.0 km. Above this altitude, the wind velocity held constant in magnitude at the maximum altitude sample value while the direction was held constant at the average value. A one-dimensional response “surface” was created using the Kriging method discussed above for $V_w(h)$ and the wind direction. It is worth noting that balloon-based wind measurement typically must include some accounting for *persistence*, or the extent to which the measured profile changes with time, but this is presently neglected.

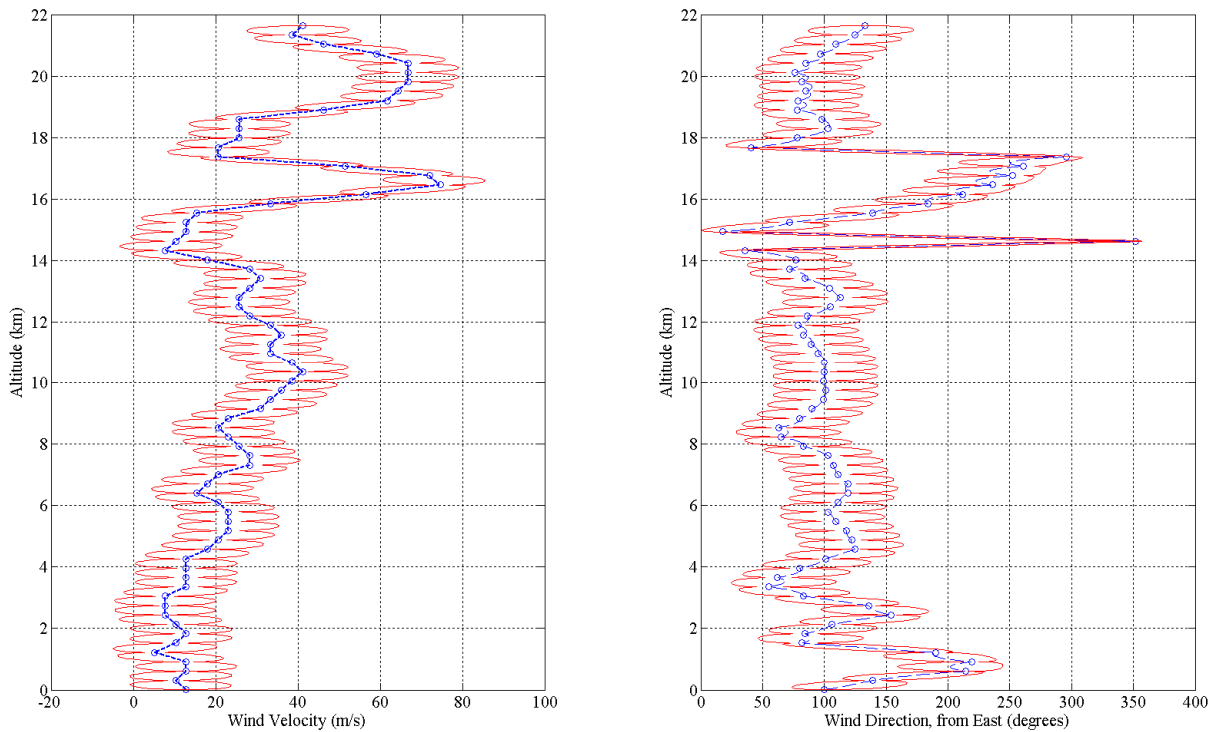


Figure 3. Wind speed (left) and direction (right) for the representative profile with \pm three s' range (red) from one-dimensional response surface. In this simple case, local maxima occur between each data point.

C. Atmospheric Data

Typically, $\rho_\infty(h)$, $p_\infty(h)$, and $T_\infty(h)$ would be sampled from balloon data in addition to the wind data discussed above. Here, for the purposes of simplicity, we assume the MSISE-90 model^{8,12,18}. Many models are available for various regions of the atmosphere; the MSISE-90 model is well-suited for applications spanning a large range in altitude. These we implemented by fitting and subsequently evaluating cubic splines over the range $0 \text{ km} < h < 200 \text{ km}$. Low-fidelity $\rho_\infty(h)$ and $p_\infty(h)$ distributions were taken as exponential, Eqs. (15-16). The scale height was taken as a constant 7.3 km based on sea-level data in Ref. 20, but various with altitude in general. It is noted that T_∞ can vary substantially in the upper atmosphere depending on solar and seasonal conditions but here is only used for the computation of the speed of sound. Above approximately 40 km, the density and hence q_∞ are low enough that errors in the speed of sound (and hence M_∞) have a small impact on the resulting optimum trajectory but a continuous evaluation of M_∞ is helpful for numerical performance. The ideal gas equations of state was used to compute the speed of sound from T_∞ based on a constant $\gamma = 1.4$.

$$\rho_\infty = \rho_{\text{SL}} e^{-h/h_S} \quad (15)$$

$$p_\infty = p_{\text{SL}} e^{-h/h_S} \quad (16)$$

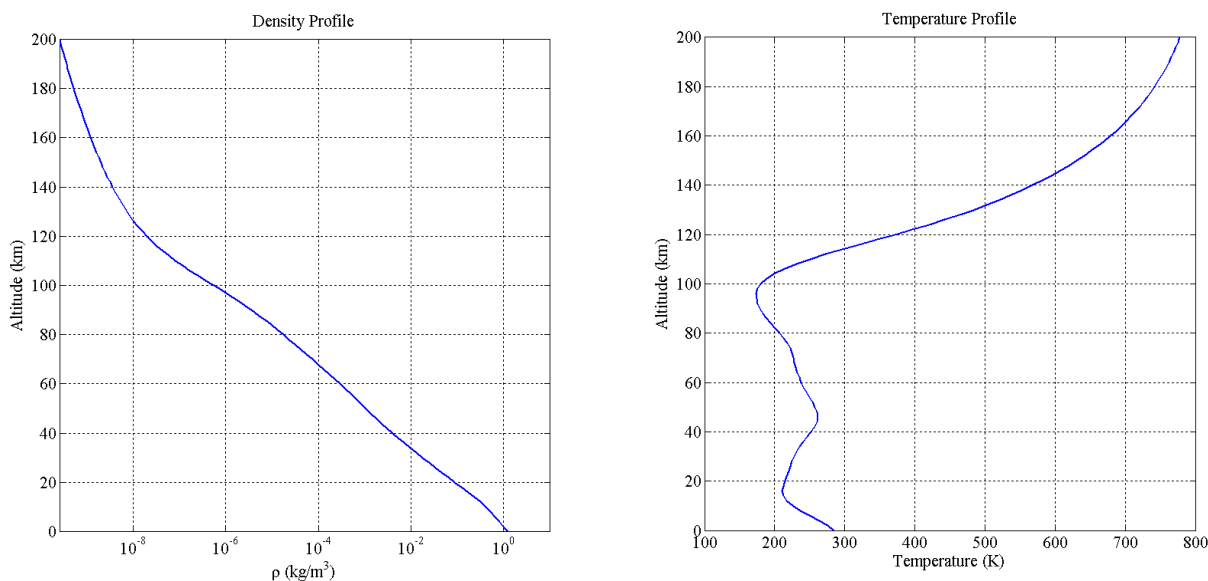


Figure 4. Density (left) and temperature (right) profiles from the MSIS-E-90^{8,12,18} model up to 200 km.

D. Results

Optimizations were performed at each level of fidelity, initializing each new level of fidelity from the previous solution. Additional data points were adaptively sampled based on EIF($-q_\infty, s'$) until the adaptive sampling budget of 20 points on the aerodynamic response surfaces was exhausted. Selected paths through $\{M_\infty, \alpha\}$ space are shown in Fig. (5) and EIF results, based on the original C_D and C_L response surfaces for consistency, are shown in Fig. (6). Results of the high-fidelity optimum are shown in Fig. (7).

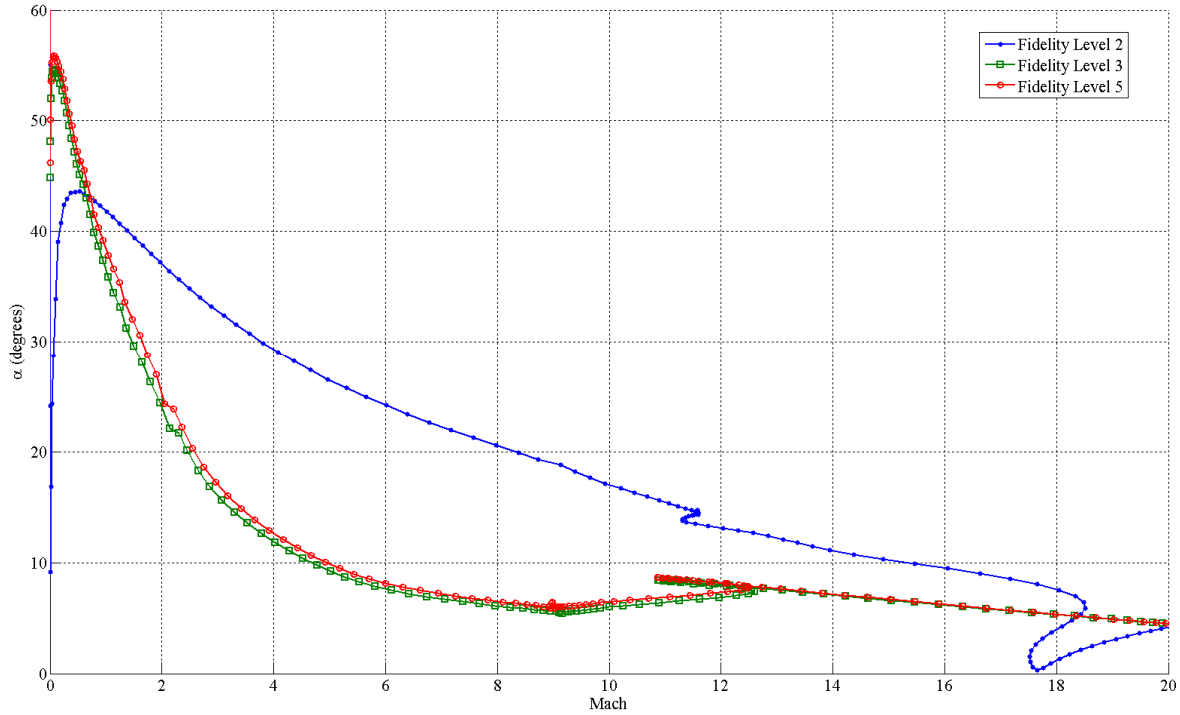


Figure 5. Mach- α flight profiles for various levels of fidelity. Note that M_∞ actually decreases in certain flight regimes due to an increasing speed of sound. Fidelity levels which utilize wind data have $\alpha(t_0) = 90^\circ$ due to the vertical liftoff constraint. Note that much of the flight is spent at $\alpha > 10^\circ$, the maximum used in the initial response surface data range.

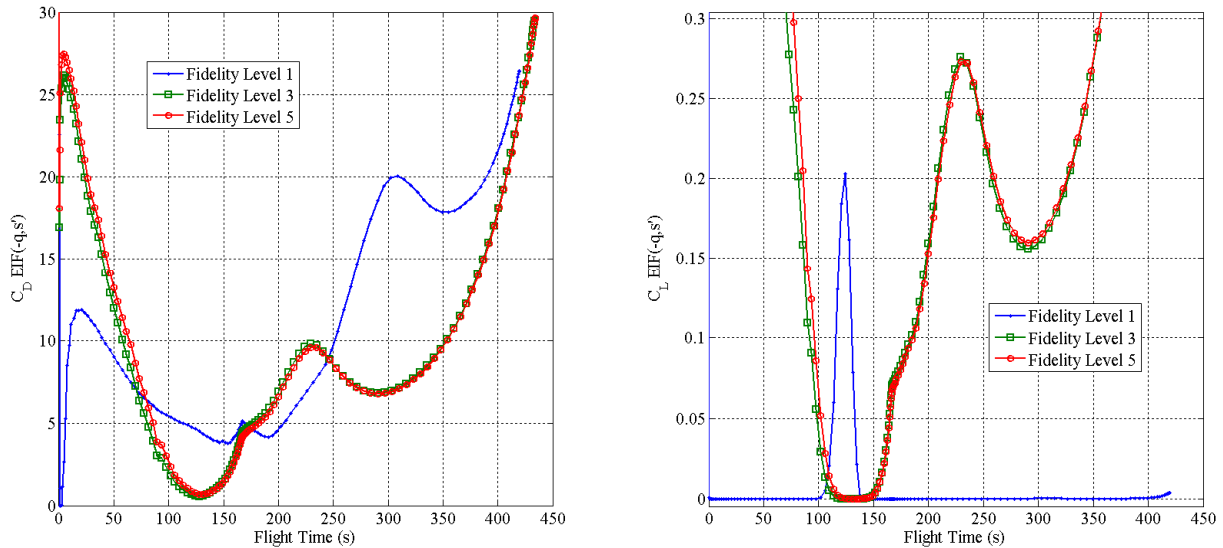


Figure 6. EIF(- q_∞, s^2) results for C_D and C_L during flight, based on the original (unadapted) response surfaces shown in Figs. (1-2) above. Note that EIF values are lower during the portion of the flight covered by the M_∞ range of the initial response surfaces and increases in regions of high α . Fidelity level 1, which does not include wind, dramatically alters the α profile and hence the EIF values. Note that C_D EIF values are much larger due to the relatively poor fit representation provided by the C_D response surface.

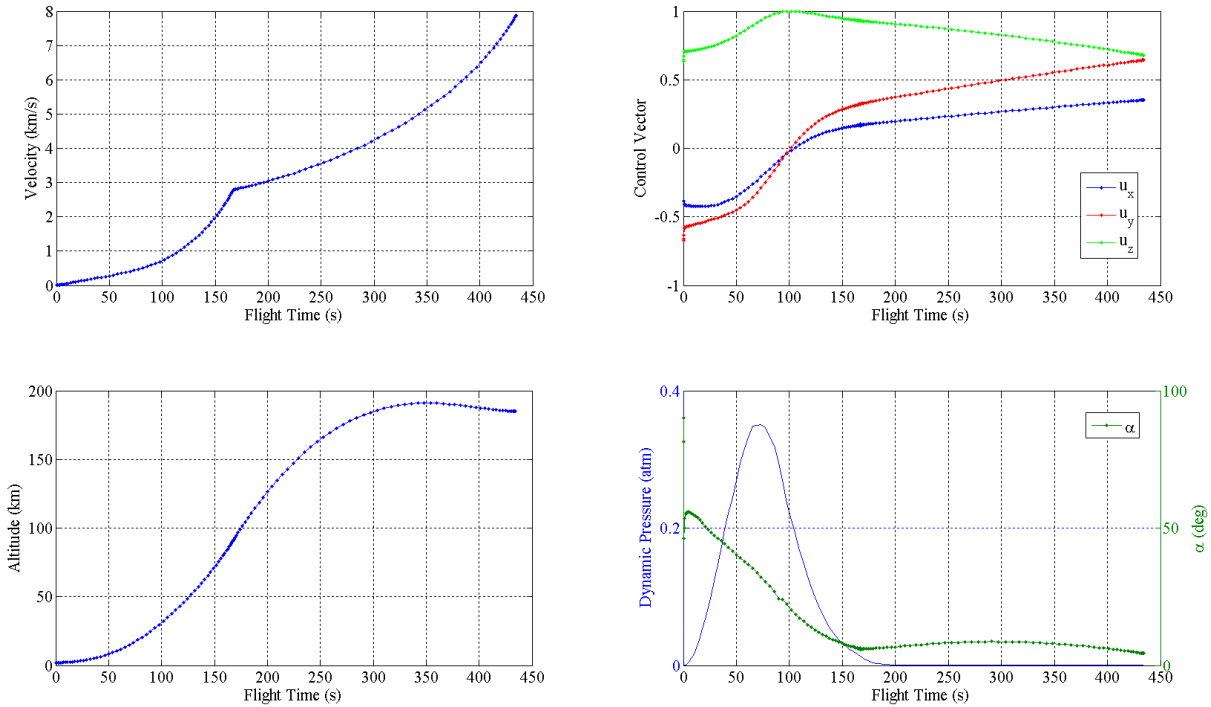


Figure 7. Optimum trajectory results from highest fidelity level simulation: total airspeed velocity (including wind), upper left, control vector components in ECI coordinates, upper right, altitude (lower left), and dynamic pressure with α , lower right. Stage division is visible in the velocity plot where a discontinuity exists in acceleration. Note that the optimum path in this particular case rises above the final altitude before adding speed via thrust with a downward component.

These results are driven primarily by $\alpha(t)$, a function primarily of wind conditions. Note that the initial range, $0 < \alpha < 10^\circ$, was too small for the lower- M_∞ portion of the flight. This observation emphasizes the usefulness of the EIF as a tool for adaptively sampling. Simply adding points in the region of maximum q_∞ would, in this example, be of little value in increasing the fidelity of the solution overall. By taking into account the estimated error as well, additional data can be tuned to regions of interest for a particular mission. As M_∞ becomes large late in the flight, decreasing q_∞ competes with rapidly increasing error (recall the original upper limit of M_∞ was 3.5), resulting in large values of the EIF.

VII. Conclusions & Future Directions

A multi-fidelity, response surface based trajectory optimization tool was developed and tested with an ascent trajectory example. By initializing higher levels of fidelity with the solutions of lower fidelity optimization and adaptively sampling physical data in independent variable regions of greatest impact, high-fidelity optima can be determined with minimal computational cost or solution time. Since atmospheric conditions are typically sampled shortly before liftoff and continuously change, rapid determination of launch availability can be critical to the mission success. Here, accuracy very close to that of a solution utilizing exact aerodynamic data was achieved with a small number of samples relative to increasing the sampling resolution of the entire variables space.

There are numerous future directions to improve the capability studied here. A key improvement is the generalization to 6-DOF, requiring a control system model. The rotational degrees of freedom have physical time scales \ll the orbital time scale used in a 3-DOF simulation resulting in a stiff set of ODEs. This presents numerous problems to numerical optimization methods, the least of which is the number of control points required for adequate resolution of the independent rotational states. Practically, the rotational states must be driven by guidance which is a function of state. Since GPOCS does not rely on shooting (integrating the ODEs forward in time), feedback control can be implemented with appropriate transient behavior.

Accounting for the rotational DOFs also required a complete set of aerodynamic data, including pitch, yaw, and roll moments. This increases the dimensionality of the response surface needed for aerodynamic performance, placing additional emphasis on adaptive sampling for efficient high-fidelity results. The simplified stability criterion

used here would be replaced with more accurate measures of controllability for day-of-launch criteria evaluation. Depending on the control system used, this may result in more EIF objectives for adaptive sampling.

Finally, extending the physical disciplines to include other results of interest to ascent and reentry trajectories would increase the usefulness of an EIF-based approach to adaptive refinement. These would include flight environments such as heat transfer and structural dynamics (natural modes). Path or endpoint constraints on stagnation point heat transfer or dynamic amplification, for example, could be included with the supporting physics and subsequently adaptively refined for a given mission based on any combination of criteria. Structural dynamic response can be particularly sensitive to payload mass properties, making adaptive refinement in this area particularly useful in mission planning especially if a source for unsteady aerodynamic prediction were available.

Acknowledgments

The author would like to thank Marcus Edvall and Anil Rao of Tomlab Optimization, Inc. for their support and advice with the GPOCS package used for dynamic optimization in this work.

References

- ¹Bryson, A. E., *Applied Linear Optimal Control*, Cambridge University Press, 2002.
- ²Bryson, A. E., *Dynamic Optimization*, Addison-Wesley, 1999.
- ³Bensen, D. A., Huntington, G. T., Thorvaldsen, T. P., and Rao, A. V., "Direct Trajectory Optimization and Costate Estimation via an Orthogonal Collocation Method", *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 6, Nov-Dec. 2006.
- ⁴Betts, J. T., "Survey of Numerical Methods for Trajectory Optimization", *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 2, 1998, pp. 193-207.
- ⁵Betts, J. T., *Practical Methods for Optimal Control Using Nonlinear Programming*, SIAM Press, Philadelphia, 2001.
- ⁶Gill, P. E., Murray, W., and Saunders, M. A., "SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization," *SIAM Review*, Vol. 47, No. 1, 2005, pp. 99-131.
- ⁷Hargraves, C. R., and Paris, S. W., "Direct Trajectory Optimization Using Nonlinear Programming and Collocation", *Journal of Guidance, Control, and Dynamics*, Vol. 10, No. 4, 1987, pp. 338-342.
- ⁸Hedin, A. E., "Extension of the MSIS Thermospheric Model into the Middle and Lower Atmosphere", *Journal of Geophysical Research*, Vol. 96, No. A2, pp. 1159-1172, 1991.
- ⁹Holst, T. L. and Pulliam, T.H., "Evaluation of Genetic Algorithm Concepts Using Model Problems, Part I: Single Objective Optimization", NASA/TM-2003-212812, 2003.
- ¹⁰Holst, T. L. and Pulliam, T.H., "Evaluation of Genetic Algorithm Concepts Using Model Problems, Part II: Multi-Objective Optimization", NASA/TM-2003-212813, 2003.
- ¹¹Jones, D. R., Schonlau, M., and Welch, W.J., "Efficient Global Optimization of Expensive Black-Box Functions", *Journal of Global Optimization*, Vol. 33, 2005.
- ¹²Labitzke, K., Barnett, J. J., and Edwards, B. (editors), *Handbook MAP 16, SCOSTEP*, University of Illinois, Urbana, 1985.
- ¹³Lophaven, S. N., Nielsen, H.B., and Søndergaard, J., "DACE: A Matlab Kriging Toolbox", Version 2.0, Technical Report IMM-TR-2002-12, Technical University of Denmark, 2002.
- ¹⁴Ross, I. M., Fahroo, F., "A Perspective on Methods for Trajectory Optimization", *AIAA/AIS Astrodynamics Specialist Conference and Exhibit*, Aug. 2002.
- ¹⁵Sóbester, A., Leary, S.J., and Keane, A. J., "On the Design of Optimization Strategies Based on Global Response Surface Approximate Models", *Journal of Global Optimization*, Vol. 13, 1998.
- ¹⁶Matlab, Software Package, Ver. 7.5.0, MathWorks, Inc., 2007.

¹⁷NIMA TR8350.2: "Department of Defense World Geodetic System 1984, Its Definition and Relationship with Local Geodetic Systems." 3rd ed., July, 1997.

¹⁸NSSDC (Goddard Space Flight Center) ModelWeb MSISE-90 Model:
<http://nssdc.gsfc.nasa.gov/space/model/atmos/msise.html>

¹⁹TOMLAB, Software Package, Ver. 5.9, Tomlab Optimization, 2007.

²⁰"U.S. Standard Atmosphere", U.S. Government Printing Office, Washington, D.C., 1976.