Dominant Suborbital Space Tourism Architectures

Markus Guerster and Edward F. Crawley

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

DOI: 10.2514/1.A34385

In the early stages of maturity of a system built for a specific function, it is common for the solutions to lie in a broad architectural space, in which numerous concepts are being developed, built, and tested. As the product matures, certain concepts become more dominant. This pattern can currently be observed in the suborbital tourism industry, in which the obvious question is what system architecture will provide the best combination of cost and safety and in the long run become the dominant architecture. This paper addresses this question by defining a broad architectural space of thousands of possibilities and exploring it comprehensively. We identified 33 feasible architectures, 26 of which had not been proposed earlier. A genetic algorithm optimizes each architecture with respect to the launch mass (a proxy for cost) and operational safety. The launch mass has been calculated in a design analysis framework consisting of four modules: weight/size, propulsion, aerodynamics, and trajectory. A validation of this framework shows relative differences below 7%. A quantitative safety analysis is developed and validated with a survey of 11 leading experts. We identify six dominant architectures. These form a set from which optimal variants are likely to come.

I. Introduction

G
enerally, in the early stages of maturity, systems built for a specific function lie in a broad architectural space with numerous concepts being developed, built, and tested. As the product matures, certain concepts become more dominant, and the variety of concepts in use decreases [1]. Consider the wide range of flying machines in the decades before and after the Wrights. History teaches us that the original architectural decisions made by the Wrights (e.g., biplane, pusher propeller, and canard) do not always survive as the dominant architecture [1]. This pattern of a wide variety of concepts can currently be observed in the suborbital tourism industry.

The birth of the suborbital space tourism dates to May 1996, when the XPrize was created by Peter Diamandis, the founder of the XPrize Foundation. In 2004, it was renamed to the Ansari XPrize after a large donation from the Ansari family [2]. This competition challenged teams around the world to build a reusable, private-funded, and manned spaceship. The first team to demonstrate a system capable of carrying three people to 100 km above the Earth’s surface twice within two weeks received the $10 million prize. Twenty-six teams from seven nations proposed their concepts. Finally, the Mojave Aerospace Ventures team, which was led by Burt Rutan and his company Scaled Composites and financed from Paul Allen, won the competition on 4 October 2004. Richard Branson licensed Burt Rutan’s technology and created Virgin Galactic. The competition launched a billion-dollar market for suborbital space travel [2]. Besides Virgin Galactic, Blue Origin and XCOR have also spent significant effort in developing suborbital vehicles. However, adverse financial conditions forced XCOR to lay off most of its employees in June 2017 [3,4]. This delayed the completion of their Mark 1 prototype for an indefinite period, leaving Blue Origin and Virgin Galactic as the major current commercial players in the United States. Their concepts are highly different, raising the obvious question: what system architectures of suborbital space tourism vehicles will provide the best combination of performance, cost, and safety?

System Architecture provides us with the methodology to approach this question. We will develop approaches and design analysis frameworks that can be used to evaluate suborbital tourism architectures with respect to their benefit (performance and safety) and cost (launch mass being used as a proxy for cost).

System Architecture emerged as a discipline in the late 1980s. The methodology has proven its success in early phases of complex aerospace engineering system designs [5,6]. System Architecture is a process used in the very early stages of system development, in which the focus lies on carefully identifying and making the decisions that define the highest-level design trades. These architectural decisions determine much of the system’s performance, cost, and safety. We use architecture to mean the entities of a system as well as the relationship between those entities [7,8]. This idea of architecture as decisions is used later to describe suborbital tourism vehicles [6]. A good architecture is then defined by providing the most benefit for a given cost, or equivalent performance with least cost [9]. Solution-specific metrics must be defined for each problem statement [10,11].

As each architecture can be represented by a set of decisions [12], we can encode these decisions into mathematical design variables. If the benefit and cost of the system become the objective functions, we can view the system architecting process as an optimization problem. Recent reductions in computational costs allow us to explore larger architectural spaces in greater detail [13].

A design analysis framework is used to compute the benefit and cost metrics, generating output parameters for given input design variables that are defined by the architecture. Our literature search on existing design frameworks parallels the one executed by Frank et al. [14–16] and Burgaud et al. [17]. We found 12 sizing and synthesis codes [14–16,18–29]. They were compared in terms of computational speed, availability, ease of use, ability to explore the architectural space, and their level of abstraction suitable for conceptual studies. Our comparison of the design analysis framework evaluation codes shows that Frank’s code has the best overall performance. Fortunately, Frank provided us with his design analysis framework.

The specific objective of this work is to identify a small set of dominant architectures for suborbital space tourism. Using system architecture approaches, we will fix the system performance to the XPrize standard (isoperformance [30]), search through the architectural space for suborbital space tourism vehicles, and evaluate optimized architectures for safety and launch mass using a parametric model. We chose to optimize toward launch mass instead of cost because direct estimation of costs has a higher uncertainty and therefore would lead to results with higher inherent uncertainty. We made this a two-step process, in which we first optimize toward mass (this paper) and then perform a commercial discounted cash flow analysis based on the optimized results (see [31,32]). Besides the minimum altitude of 100 km, the isoperformance is characterized by a defined number of participants (four throughout our analysis). Finally, we will assess the undominated architectures and identify the small handful of variants that merit more refined design analysis.
In Sec. II, we define the plausible architectural space. Then, we build a parametric model for each viable architecture and optimize them in Sec. III. Finally, we visualize and make sense of the result to support the decision-making process in Sec. IV.

II. Defining Plausible Architectural Space

A common axiom is that a larger architectural space yields better architectures and variants. Based on this, the objective of this section is to define the architectural space as broadly as is feasible. As used later, we use the word \textit{variant} to describe an instance of an architecture with a specific set of design variables. Thus, every architecture has its own variant space, and the combination of all variant spaces results in the architectural space (see Sec. III.A). We identify four main steps to define the architectural space:

1. Build a database of existing concepts in the architectural space of interest.
2. Define a morphological decision matrix based on the database.
3. Define logical, reasonable, and model-limitation constraints.
4. Enumerate the feasible architectures that represent the plausible architectural space.

These four steps are described for our case of suborbital tourism vehicles in the following four subsections, Secs. II.A–II.D. In reality, the definition of the morphological decision matrix, the reasoning about the constraints, and the enumeration of the feasible architectures are all coupled. Consequently, this is an iterative process.

A. Building Database of Existing Concepts

The first step in defining the architectural space is to look at existing concepts that have similar requirements. These can be concepts from the documents submitted in competitions, current competitors who have already decided on an architecture, or existing studies in the market. The literature research turned up team summaries of almost all the companies that competed in the Ansari XPrize, previous studies of the market, texts about space tourism, and data extracted from companies’ websites. The sources reviewed include 1) a 1998 report from the Associate Administrator for Commercial Space Transportation [33]; 2) the 2002 report “Suborbital Industry at the Edge of Space” [37]; and 7) 2016 current websites from ARCA [38], Blue Origin [39], Virgin Galactic, XCOR, Copenhagen Suborbital [40], Dassault-Aviation, EADS [41], CosmoCourse [42], and Scaled Composites.

We found 37 companies with seven distinguishable architectures based on our morphological matrix, which is described in the following section.

B. Defining Morphological Decision Matrix

The concept of using a morphological matrix to study the architectural space of different configurations dates back to Zwicky in 1969 [43]. Since then, it has often been used as a decision support tool [44]. In the form of the morphological matrix that we use, rows represent architectural decisions, and columns are the possible options (see Fig. 1). The options for a certain decision must be mutually exclusive. An architecture is then defined by choosing one option for each decision. They do not have to be from the same column. Our final decision matrix is shown in Fig. 1. Together with the constraints from Sec. III.C, these decisions and options define our architectural space. The matrix consists of one vehicle-level decision about the number of modules, followed by five decisions for the first module and three for any possible second module. We define a module as a system that separates during flight and lands independently. The module with the higher number has the higher energy state. The vehicle can consist of either one or two modules. If there is only one module, the first module is crewed, and no second module exists (see Fig. 2 and architectures 1–5 in Fig. 3). If the vehicle has two modules, the second module is crewed, and the first module is uncrewed (architectures 6–33 in Fig. 3). We assume that during the ascent the crew in module 2 can control module 1, but module 1 needs to land autonomously.

In our morphological decision matrix shown in Fig. 1, there is one decision on vehicle level about the number of modules (nModules). This influences which of the following decisions need to be considered. If the vehicle only consists of one module (nModules = 1), then the decisions about the second module are the default (Wing 2 = no, Rocket Engine 2 = no, and Landing mode 2 = None). On module level, the decisions cover all three operational aspects: launching, flying, and landing. The launching or takeoff mode can be

\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{Decision} & \textbf{Option 1} & \textbf{Option 2} & \textbf{Option 3} & \textbf{Option 4} & \textbf{# of options} \\
\hline
\textbf{Take-off mode} & \textbf{Horizontal} & \textbf{Vertical} &  &  & 2 \\
\hline
\textbf{Wing} & \textbf{No} & \textbf{Yes} &  &  & 2 \\
\hline
\textbf{Jet Engine} & \textbf{No} & \textbf{Yes} &  &  & 2 \\
\hline
\textbf{Rocket Engine} & \textbf{No} & \textbf{Yes} &  &  & 2 \\
\hline
\textbf{Landing mode} & \textbf{Gliding} & \textbf{Horizontal Powered} & \textbf{Parachute} & \textbf{Rocket} & 4 \\
\hline
\textbf{Wing} & \textbf{No} & \textbf{Yes} &  &  & 2 \\
\hline
\textbf{Rocket Engine} & \textbf{No} & \textbf{Yes} &  &  & 2 \\
\hline
\textbf{Landing mode} & \textbf{Gliding} & \textbf{Parachute} & \textbf{Rocket} & \textbf{None} & 4 \\
\hline
\end{tabular}

Fig. 1 Morphological matrix with eight architectural decisions. Blue shaded options represent Blue Origin’s New Shepard vehicle. Full enumeration would lead to 2048 unconstrained architectures.
horizontal or vertical. An independent launching decision does not exist for the second module because the second module has the same takeoff mode as the first module and then separates during the ascent phase. Whether the modules are to be oriented on top of each other or in piggyback style is not modeled.

The flying decision covers three issues: if the module has wings or not, if it has jet engines or not, and if it has a rocket engine or not. By excluding the decision about the jet engine for the second module, we implicitly integrate the constraints that the final ascent must be made by a rocket engine. If a jet engine is part of the architecture, then it must be on the first module. A commercially available jet engine (Mach under 2) does not impart enough excess energy in the atmosphere to ballistically ascend to 100 km.

Finally, the last decision for each module is the landing mode. The options cover unpowered and powered landing modes. Unpowered modes include gliding and parachute landing, the orientation of the 14 first one of which is horizontal and of the latter of which is vertical. Powered modes are horizontal powered landing (Hor. Powered) and vertical powered landing with a rocket engine (Rocket). We assume a horizontal powered landing with a rocket engine is not feasible due to challenges associated with its throttling over a greater range. As the second module cannot have a jet engine, it has no horizontal powered landing option. A full combinatorial enumeration (product of the numbers in the last column) leads to 2048 unconstrained architectures identified.

We apply this morphological matrix to Blue Origin’s New Shepard to provide an example of how this matrix can be used to represent an architecture (indicated by the blue marked options in Fig. 1). By our definition, this vehicle is a two-module vehicle. The first module is a vertical launching rocket with a rocket engine. It has no wings and no jet engine. The landing mode is vertically powered (rocket). The crew-carrying second module (the capsule) separates during flight, has no propulsion nor wings, and lands with a parachute. Note that the decisions in our matrix do not include scaling attributes like the number of participants, passengers, or crew. We treat the number of participants as a preset parameter (four for all examples in this paper) and the number of pilots as a design variable.

C. Defining of Logical, Reasonable, and Model-Limitation Constraints

The morphological matrix of Fig. 1 defines the architectural space for all 2048 unconstrained architectures. We define three kinds of constraints, which, when applied to the architectural space of Fig. 1, will significantly reduce the number to the feasible architectures. Logical constraints exclude some combinations of options through simple logic. Reasonable constraints exclude combinations that a knowledgeable and well-informed observer would conclude are almost certainly dominated [45]. Model-limitation constraints flag options that cannot be analyzed by the general model and are treated as special cases or eliminated by reasonable constraints.

Model-limitation constraints identify options that the modeling capability cannot analyze. For example, our main model was able to analyze all but three of the published concepts for suborbital tourism. Those that cannot be characterized include systems that launch from underneath the water, systems that launch from balloons, and systems with more than two modules. These would have to be independently analyzed to test their comparative risk and cost, and this was beyond the scope of the study. But our reasoning is that these three concepts are unlikely to be dominant architectures. For example, 36 out of 37 companies in the data set proposed a concept with one or two modules, and only one uses a three-module rocket. Since one or two propulsion modules can easily supply enough energy to reach the Karman line at 100 km altitude, it is likely this architecture will be dominated. One company proposes a launch from under the water. One can imagine that this type of launch adds complexity, operational challenges, and cost without any obvious benefit; therefore, it was excluded. Finally, we excluded the balloon option as an ascent method as our design framework does not have a model to optimize and evaluate the metrics for a vehicle with this feature.

There are additional constraints that exclude infeasible architectures. Some of them are logically not possible; for example, with reference to Fig. 1, if the number of modules is set to 1 (nModules == 1), then logically there is no second module, and therefore Wing2 and Rocket Engine2 must be No, and Landing Mode 2 must be None. Other constraints are more subjective or reasonable. They cover aspects in which domain experts can reasonably argue that the implementation of this combination of options results in higher development challenges or operational complexity without providing benefit in the metrics of interest. An example is the exclusion of horizontal takeoff without wings. One can think of systems in which a horizontal takeoff without a lifting wing may be achievable (e.g., by providing lift with vertically oriented engines). However, we can argue with engineering judgment that this increases complexity without providing a benefit. For an exhaustive matrix of the 23 constraints used in the analysis and their justification, refer to the work by Guerster [46].

D. Enumeration of Feasible Architectures

At this point, we have created the morphological matrix and identified all the constraints among the decision’s options. With this information, we can now define our architectural space by enumerating all unconstrained architectures and comparing each with the constraints matrix. If the architecture contains a set of options that is restricted by one of the 23 constraints, it is excluded from the list of viable architectures. The resulting list includes 33 distinct feasible architectures that satisfy all constraints out of the 2048 unconstrained solutions. This list of 33 architectures therefore defines our architectural space. We have developed a matrix to visually represent these 33 feasible architectures. The matrix, shown in Fig. 3, has the concept of the first module as columns and for the second module as rows. There are six possible concepts for the first module and, by coincidence, the same number for the second module (including the choice that there is no second module). Three of the 36 combinations in the matrix of Fig. 3 are excluded by the constraint that there must be a rocket engine on either the first or second module [47]. The pictograms in the top row for the first module and the left column for the second module help to visualize the concepts of the modules and their attributes. A description of the meaning of these pictograms is provided in Fig. 4.

The 33 architectures are numbered, starting from the top left to the bottom right. The three infeasible ones are marked by an asterisk. The architectures indicated with an asterisk were investigated by Frank et al. [14–16]. If we screen our database of 37 proposed companies and match them to one of the 33 architectures, we find that there are just seven distinct architectures contained in the database of 37 companies. For each of these, an example company is indicated. This correspondence does not include an exact match for design variables as our design optimization may have arrived at a different number of pilots, participants, or propellant type. For example, Virgin Galactic’s WhiteKnightTwo and SpaceShipTwo are matched to architecture 14.

Virgin Galactic uses this architecture with a specific set of design variables, resulting in what we call in this paper a variant. In other words, Virgin Galactic developed a variant within the group of architecture 14. There is a huge number of other variants with

\[ n \approx 2^{13} \times 2^{13} = 2^{26} \]

As indicated in Fig. 2, we do not allow pilots on the uncrewed module 1. Hence, the WhiteKnightTwo carrier aircraft is slightly outside of our architectural space since it has two pilots. Nevertheless, it is an example of design for architecture 14.
various combination of design variables. The design variables are defined by Frank’s sizing code. They span, for example, different rocket engine propellant types, the number of participants, the wing’s aspect ratio, and the turbine inlet temperature. For a full description of the 27 design variables with their limits, refer to the work by Guerster [46]. The optimization procedure described in Sec. III varies these design variables to find a set of variants with the lowest mass and risk within each architecture of the 33 architectures.

In summary, we identified 33 viable architectures. Of these, only seven appear in the database of company proposals. We have laid the foundation to discover new designs for safer and affordable suborbital space tourism by identifying 26 new architectures. In the design parameters and identify dominant architectures.

III. Architectural Space Exploration

Having defined the architecture space in Sec. II, in this section, we will describe the methodology of how we will evaluate this architectural space (Sec. III.A). We will then discuss the design framework for performance and safety assessment and outline their validation in Secs. III.B and III.C.

A. Methodology

The methodology of the architectural space exploration is shown in Fig. 5. An architecture definition function takes the decisions-encoded architectures as an input. As each architecture has a different list of design variables, this architecture definition function extracts the active variables.
33 viable architectures

Architecture definition

Generic design and requirement variables list

Design variables

Launch mass

Safety Metric

Overall Residual Safety-Risk Metric

Genetic algorithm

Launch mass

Design variables

Launch mass

Overall Residual Safety-Risk Metric

Fig. 5 Schematic overview of the architectural space exploration methodology. A Genetic Algorithm optimizes each architecture within its variant space in the discrete risk and continuous mass dimensions.

design variables with their lower and upper boundaries as well as their design variable type (continuous, discrete, or categorical) from a generic list. For example, one design variable is the number of pilots in the final module, which can be 0, 1, or 2. A second example is the type of rocket engine and its propellant combination. The rocket engine can be solid, liquid (LOX/LH₂ or LOX/RP₁), or hybrid (hypergolic, LOX/HTPB, LOX/Paraffin, N₂O/HTPB, or N₂O/paraffin). In addition, there are preset parameters such as the maximum altitude, acceptable load factor, seat pitch, and number of participants. The design variables are then populated with initial values by a non-dominated sorting genetic algorithm II Genetic Algorithm (GA) [48]. To explore a broad space within each architecture, the population is set to 2000. This increases the likelihood that the converged variants are global optimums. The design framework takes the design variables from the GA as input and outputs the total launch mass. In doing so, the design framework 1) sizes the module, 2) calculates the propulsion, 3) calculates the aerodynamic properties, and 4) plots the trajectory. This is an iterative optimization process, as all four calculations depend on each other. Depending on the vehicle’s number of modules, the design framework must be executed once or twice within the GA loop described previously. As a next step, the safety assessment calculates the discrete Overall Residual Safety-Risk Metric (ORSRM) based on the design variables. This safety assessment and the total launch mass are the objective functions of the GA, and the new generation is generated by crossover and mutation. The new population is then evaluated again until the maximum number of generations is reached. Fifty generations generally assure convergence [46]. Since the optimization of the various architectures is not coupled, the problem can be broken down into 33 independent subproblems. Therefore, the optimization process is done in parallel for all architectures, resulting in a Pareto front with the most promising variants for each architecture (Fig. 5). Overlying the individual Pareto fronts from each architecture’s variant space results in the architectural space, which is shown in Fig. 5. The actual calculated architectural space is discussed in Sec. IV.

B. Design Framework

We base our design framework on Frank et al.’s contributions [14–16] and make two significant modifications:

1) We extend their code to be capable of exploring our larger architectural space.
2) We improve the computational performance by two orders of magnitude.

Without the latter adjustment, the optimization would be computationally intractable. In the following, we first give a conceptual overview of the design framework developed by Frank et al. [14–16], followed by a description of our two contributions and a validation.

1. Schematic Overview

Frank et al.’s [14–16] design framework evaluates the mass of the vehicle based on the current state of the design variables. It consists of four submodules: the weight/size, propulsion, aerodynamics, and trajectory submodules (outlined in blue in Fig. 6). The propulsion and trajectory calculations are further specialized to those for jet and rocket engines. The optimizer for the design framework iterates the propellant masses and the burn time for the jet and rocket engines until the relative differences with the previous estimates are below a defined limit. For each iteration, the weight and size are recalculated, the propulsion system is sized, the aerodynamic properties are extracted, and finally

![Diagram showing the design framework process]

Fig. 6 Overview of the design framework consisting of the weight/size, propulsion, aerodynamic, and trajectory submodules. An internal loop updates the propellant estimation until convergence is achieved.
the separation is not part of the optimization. If the module has only a rocket engine, the transition altitude is not considered. If the module has both a rocket and a jet engine, there is an additional transition velocity, and altitude must be optimized. The definition of the corresponding variables is shown in Fig. 7. It is assumed that if the module has a jet engine and a rocket engine it launches with just the jet engines until the transition altitude, shuts down the jet engines, and uses the rocket engine for the remaining ascent. We do not consider the simultaneous operation of both engine types. If the first module only has a jet engine or a rocket engine, neither transition altitude nor velocity needs to be considered.

3. Increasing Computational Efficiency by Factor of 95
A run-time investigation of Frank’s initial code shows that around 95% of the evaluation is taken up by optimizing the trajectory. With this, it is impractical to explore our large architectural space. Hence, we have developed a new approach that estimates the propellant mass with a fraction of the computational effort while retaining accuracy. Instead of using a gradient-based solver, we make use of the already existing feedback loop within the framework (see Fig. 6) to achieve convergence within the trajectory module (further details are in [46]). The computer time needed for the design framework evaluation is reduced by 95%. This allows us to search through a much larger architectural space.

4. Validation of Design Framework
Data from Virgin Galactic, Blue Origin, and Rocketplane XP are used to validate the design framework’s weight, sizing, jet, and rocket trajectory optimization features. We compare the takeoff masses of these vehicles with the results from the design framework (see Table 1). Because of the lack of testing, no prototype of the Rocketplane XP exists, and therefore no measured takeoff mass is available. However, detailed design study data are available on its official website [49] and in other sources. Both New Shepard and the SpaceShipTwo are in the test phase. We have cited several sources for each project in Table 1.

The table shows relative differences between 1 and 7% with respect to the closest reported value. Given the range and uncertainty in the sources, this is a reasonable result, especially with this variety of features in the architectural space.

C. Safety Assessment
The architectural decisions have a significant and lasting impact on how safe the system can be built. However, practical methods for quantitative analysis of the safety, like the top-down Fault Tree Analysis or the bottom-up Failure Mode and Effect Analysis, are difficult to apply in conceptual scenario tradeoffs [50–52]. Our approach is an adaption of Dulac and Leveson’s risk analysis for early system architecture trade studies, leading to a quantitative measure, the ORSRM. We provide in the following a conceptual description of the assessment and its main conclusions. A summary of the input data for the assessment is shown in the Appendix. For a detailed step-by-step description, refer to [53]. We use three factors for the safety assessment:

Severity is the worst possible consequence of a hazard [51]. Leveson [51] developed a custom scale for an early manned space exploration analysis, which we will use here. Its four categories range from “less than minor injury” to “loss of life.”

Likelihood is the probability that a hazard occurs [51]. We adapt the terminology with six categories from the standard risk matrix [51]. They range from impossible to frequent. We anchor the scale with the category occasional, meaning it is likely that the hazard occurs once over the vehicle’s lifespan.

Mitigation is the potential of the design to mitigate a hazard [51]. We use the mitigation impact scale developed by Leveson [51]. The five-category scale ranges from “no mitigation potential” to “complete elimination of the hazard from the design.”

An overview of the assessment is depicted in Fig. 8.

The process consists of four steps:
1) In step 1, a top-level hazard list is created based on NASA’s generic hazards lists for the Space Shuttle [54] and the Constellation Program. It covers in total 11 hazards at different operational points of the mission. They cover five main areas: 1) fire and explosion, 2) incorrect propulsion/trajectory/control during ascent, 3) loss of structural integrity, 4) loss of life support, and 5) incorrect stage separation if the vehicle has two modules. The first three areas apply to the ascent phase as well as the descent phase.
2) In step 2, severity and likelihood factors are assigned to each of the 11 hazards defined in step 1. The product of these factors defines the resulting risk from this hazard.
3) In step 3, the mitigation factor assesses the potential of mitigating each hazard during the design phase. This factor depends on the variant of the vehicle and therefore is a function of the architectural decisions and the two design variables: the number of pilots and propellant type. The dependency on the design variables generates the horizontal bands seen in Sec. IV. An example of applying the highest mitigation factor complete elimination of the hazard from the design is an architecture with only a single module. The rationale is that, since there is no second module, there is no stage separation and therefore the incorrect stage separation hazard can never occur.
4) In Step 4, the ORSRM is calculated by following the procedure defined by Dulac and Leveson [52]. The procedure first calculates the maximum possible mitigation potential for each hazard and then

<table>
<thead>
<tr>
<th>Vehicle name</th>
<th>Reference weight, kg</th>
<th>Design framework calculated weight, kg</th>
<th>Relative difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue Origin’s New Shepard</td>
<td>35,000</td>
<td>32,710</td>
<td>6.5%</td>
</tr>
<tr>
<td>Virgin Galactic’s SpaceShipTwo</td>
<td>9740–13,608</td>
<td>13,495</td>
<td>0.8% with respect to 13,608 kg</td>
</tr>
<tr>
<td>Rocketplane XP</td>
<td>8840–9072</td>
<td>9,641</td>
<td>6.3% with respect to 9072 kg</td>
</tr>
</tbody>
</table>

*aData available online at http://www.insideadventures.com/rocketplane-xp.html [retrieved 15 March 2017].
*bData available online at http://www.astronautix.com/r/rocketplanexp.html [retrieved 15 March 2017].
comparing it to the achieved mitigation potential of the current variant. This relative measure is then multiplied by the severity and likelihood factors defined in step 2. Averaging over those values results in the ORSRM. A higher value corresponds to a higher risk.

To validate our assessment of safety risk, we conducted a survey among 11 informed experts to compare their collective opinion of the relative safety of architectures with those calculated by us using the ORSRM approach. The experts included three former astronauts, three industry experts (one from each of the three major suborbital tourism companies), two former high-level technical leaders of government aerospace organizations, and three academic experts with broad knowledge of operations in aerospace. We asked each of these 11 experts to compare the safety of eight pairwise combinations of architectures, denoted as A and B. The questionnaire consisted of eight questions with three choices each. The choices were A) architecture A is safer, B) architecture B is safer, or C) no significant difference is perceived. The questions were chosen to cover the main safety decision in the architectural space. A summary of the questionnaire is shown in Table 2, in which architectures A and B refer to the architecture numbers in Fig. 3. Our ranking of the relative safety of the safety was calculated through the ORSRM and shown in Table 2. These results were hidden for the experts during the survey. If the same propellant and number of pilots is assumed for the compared set of architectures, a single ORSRM value can be calculated for each architecture. Our results that can be inferred from Table 2 are that in comparisons 1, 2, 3, 6, and 8 choice B is safer, for comparisons 5 and 7 choice A is safer, and for comparison 4 there is no difference.

By cross-tabulating our assessment with the 11 responses, we were able to derive the contingency table shown by Table 3. The diagonals show a strong correlation between the ranking by ORSRM values with the responses from the 11 experts (81.82% for choice A, 60.00% for choice B, and 81.82% for choice C). For the cases in which the responses do not match our assessment, no clear trend for either of the remaining two options can be identified.

This survey among 11 experts shows a good correlation between their perception of safety risk and our ranking based on the ORSRM.

### Table 2: Questionnaire for validation of the safety assessment with the eight questions we asked 11 experts (or each question, the ORSRM result is given for comparison)

<table>
<thead>
<tr>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture A</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>18</td>
<td>3</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>ORSRM</td>
<td>8.82</td>
<td>8.82</td>
<td>8.82</td>
<td>5.76</td>
<td>5.76</td>
<td>5.76</td>
<td>2.40</td>
<td>2.40</td>
</tr>
<tr>
<td>Architecture B</td>
<td>1</td>
<td>3</td>
<td>18</td>
<td>17</td>
<td>29</td>
<td>4</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>ORSRM</td>
<td>7.22</td>
<td>5.17</td>
<td>5.76</td>
<td>5.76</td>
<td>8.03</td>
<td>2.83</td>
<td>4.40</td>
<td>1.66</td>
</tr>
</tbody>
</table>

No stronger correlation with somewhat subjective judgments of architectural safety would be expected at this early point in design, so we will proceed with our ORSRM ranking as the basis of architectural screening.

### IV. Results and Discussion

In Sec. II, we defined the 33 architectures as our architectural space, and in Sec. III, we described the architectural space exploration methodology, including performance and safety assessment. This section presents and discusses the optimized results for four participant vehicles.

We first show an example objective space for one architecture with its variants. Figure 9 shows the final population of architecture 1 with the ORSRM score on the vertical axis and the total launch mass on the horizontal axis. Each black dot represents one variant of architecture 1. As conceptually introduced in Fig. 5, the launch mass is continuous, while the risk assessment shows discrete bands. As shown later in Table 4, the different number of pilots causes the three different bands. All variants in the top band have no pilot (higher risk, lower mass), the middle band has one pilot, and the bottom band has two pilots (lower risk, higher mass). The smaller fluctuations in the ORSRM within one band are caused by the different propellant types (solid, liquid, and hybrid).

When we combine all architectures into one plot, to reduce clutter, we do not show the variants. Figure 10 shows the overall Pareto front containing all 33 architectures. Each Pareto front is labeled with the architecture number defined in Fig. 3. The number is plotted close to the optimized variant of the architecture with the highest ORSRM, i.e., on the top left of the line. The dashed black line combines the dominant solution of all architectures and displays the overall Pareto front. Furthermore, the architectures are colored depending on their wing attribute. If both modules have a wing, they are colored blue; if only the first module (the lower-energy module) has wings but not the second, they are colored red; if only the second module has wings but not the first, they are colored orange; and if neither of the modules has wings, they are colored gray. In addition, the four architectures closest to Blue Origin, Virgin Galactic, XCOR, and Rocketplane XP are encoded with markers. The Blue Origin architecture 18 has larger dots (♦), the Virgin Galactic architecture 14 has a downward-pointing triangle (▾), the XCOR architecture 4 has crosses (✗), and the Rocketplane XP architecture 5 has filled diamonds (♦). Our results represent point designs, and do not take the modeling uncertainty into account. However, we can roughly estimate these modeling uncertainties by considering the validation results of both metrics. On average, the relative deviation is less than 5% for the mass. The contingency analysis would suggest that about 60–80% of the ORSRM indicates a ranking that correlates with expert opinion. Using these values, we can see from Fig. 10 that the Pareto architectures and variants are still relatively distinct and properly positioned.

Examining the results, we see that the XCOR architecture 4 is on the Pareto front for medium risk and mass. The Rocketplane XP architecture 5 is around 1 t heavier but slightly safer. The Blue Origin architecture 18 has an additional 1 t launch mass compared to the Rocketplane XP and is riskier. The Virgin Galactic architecture 14 has a slightly larger mass than the Rocketplane XP but is slightly
Architectures 17 and 18 differ in the landing mode of their first module: parachute and rocket powered, respectively. As can be seen from the figure, the rocket powered landing increases the mass and slightly increases the risk for the vehicle. The mass increase is due to the additional systems and propellant needed. The same reasoning explains the difference between architectures 1 and 2. The minimum total launch mass for a vehicle with four participants is 4.7 t.

It is worth noting here that three out of the four (Virgin Galactic, XCOR, and Rocketplane XP) more developed architectures are on the Pareto front. Only Blue Origin’s New Shepard is dominated. We explain this by the higher goal of Blue Origin to enter the orbital launch market and using suborbital tourism as a stepping stone to develop the necessary technology. This results in a suboptimal architecture for the tourism market, even though it might be safer when considering the Blue Origin’s multiple development programs.

In general, decreasing mass increases the risk. This is an expected behavior and true for most technical systems [6]. There is a tradeoff between risk and mass (which is often a proxy for cost). The data show a clear trend from the heavier and safer options on the bottom right to the lighter and riskier architectures on the top left.

There are some trends revealed by the aggregated data. The centroids shown in Fig. 10 indicate the mean value of the attribute for all similar architectures, which in Fig. 10 is the wing attribute (the centroids are calculated by averaging over all variants with the same attribute). As already introduced in the safety assessment, Sec. III.C, wings reduce risk, but the results here show that they also increase mass. The architectures with two wings are the heaviest but the safest. If the vehicles have no wings (gray lines) they are riskier but lighter. The architectures with a wing on either the first or second module are in between. And the red and orange centroids show having a wing on the second (higher energy) module is safer but heavier. This makes sense, since the weight of the wings on the second module has to be carried by the first module, increasing its size and hence increasing the overall mass of the vehicle. A wing on the second module, which would provide more abort options, would also be slightly safer.

To further investigate the influence of the number of engines on the mass and risk, we recolored the Pareto fronts according to their number of rocket engines on the entire vehicle (see Fig. 11). Red indicates vehicles with two rocket engines, meaning that each module has one rocket engine. The blue architectures have one rocket engine (either module 1 or module 2). None of the architectures with two rocket engines is on the Pareto front; these architectures are completely dominated by those with only one rocket engine. This leads us to the conclusion that the vehicles with a single rocket engine provide the best combinations of lower mass and risk. For orbital ascents, splitting the required energy between stages results in a performance benefit [55]. Because of the lower total required ΔV, the reverse is true for suborbital vehicles, as the results show.

As can be seen from Fig. 3, architectures 16, 27, and 33 have three engines, a rocket and jet engine on the first module and an additional rocket engine on the second module. Our engineering intuition is that these architectures should be heavier and will be dominated by others. And, indeed, the model confirms this intuition. They are the three heaviest architectures in Fig. 10 with total launch masses around 13–14 t.

We can similarly investigate the influence of the architectural decisions on the number of modules and the takeoff mode. Figure 12a shows that on average the one-module vehicles are considerably lighter but slightly riskier than the two-module architectures. The separation capability of a two-module vehicle makes them slightly

---

**Fig. 9** Architecture 1 used for an example plot of the final population. The black dots represent the variants, and the gray line is the resulting Pareto front for architecture 1.

**Fig. 10** Pareto front of the 33 architectures with respect to the ORSRM and total launch mass; the number of participants is set to 4. The black dashed line indicates the combined Pareto front (nPAX, number of participants).

**nPAX: 4**

---
The mass difference is significant; the average two-module vehicle is around twice as heavy as the average one-module vehicle. In Fig. 12b, the influence of the takeoff mode on the two objective functions can be observed. On average, a vertical takeoff results in a slightly lighter vehicle but increases the risk by almost 50%. The horizontal takeoff needs wings, which need to be sized for the larger takeoff mass compared to wings used for landing only. These heavier wings cause the higher mass of the horizontal takeoff option.

To identify the architectures, which merit further design development, we look at the architectures on the Pareto front in Fig. 11, which are 1, 3, 4, 5, 14, and 9, running from lighter and riskier to heavier and safer. They are shown in Fig. 13 using the pictograms defined in Fig. 4. All but one of these architectures on the Pareto front have wings. Since the architectures are ordered from the lighter and riskier to the heavier and safer, the previously discussed influence of the wing on the metrics can be seen. Again, the safer vehicles have wings, and the two safest architectures have two modules with wings. The second module adds safety due to separation capabilities, which must be traded off against its higher mass. By comparing architectures 3 and 4, we see the vertical takeoff mode for a winged module leads to a lighter vehicle since the wings can be sized for the lighter landing mass. However, the vertical launch method is riskier.

Four out of six recommended architectures have already been proposed; our discovery is that architecture 9 is the safest while still having a competitive launch mass.

Since the ORSRM depends not only on the architectural decisions but also on two design variables (discussed in Sec. III.C), there can be more than one variant per architecture on the Pareto front (see Fig. 9). All undominated variants for the six architectures identified on the Pareto front are listed in Table 4 with their ORSRM value, launch mass, number of pilots, and rocket propellant type for modules 1 and 2. The variants for each architecture are numbered, with 1 being the riskiest within an architecture. Since we investigate the tradeoff between risk and mass, this riskiest design is also the lightest.

The three columns on the right show the design variables that influence the ORSRM within an architecture. With an increasing number of pilots, the variants become safer. The propellant for all vertical launchers with a one-module vehicle is solid (architectures 1 and 3), whereas, for the horizontal takeoff in architecture 5, the lighter two variants have a liquid LOX/LH2 rocket engine. The safer but heavier two variants are powered by a solid rocket engine.

In the following, we compare variants 2 and 3 of architecture 5 to understand some detailed tradeoffs (see Table 5 for a summary of further design variables). The system optimization of these two variants shows that the transition velocity is equal, but the transition altitude is different (see Fig. 7 for the definition of transition velocity).
and altitude). The rocket engine ignites at around 6 km for a liquid rocket compared to 8 km for the solid rocket case. The higher specific impulse ($I_{sp}$) of LOX/LH2 (450 s compared to 296 s) makes the use of the rocket engine preferable for a greater energy increment ($\Delta v$). Despite the lower thrust of the LOX/LH2 engine (61 kN compared to 137 kN), the longer burn time of 150 s (compared to 58 s) leads to a greater energy increment. In addition, due to the higher $I_{sp}$ of LOX/LH2, less propellant is burned. Even though the liquid engine has a lower thrust with similar chamber pressure and nozzle expansion ratio, the overall complexity is higher than the solid rocket engine, and therefore its dry mass is more than twice as much. The sum of the propellant mass and rocket engine weight is 3182 kg for the liquid case and 3399 kg for the solid engine. This is the major contribution to the different launch masses in Table 4. Other minor effects are associated with different jet engine sizing, jet fuel consumptions, and structural differences. This example illustrates how the model and the resulting data can be used to not only understand architectural decisions but also the technical tradeoffs on a system level.

and altitude). The rocket engine ignites at around 6 km for a liquid rocket compared to 8 km for the solid rocket case. The higher specific impulse ($I_{sp}$) of LOX/LH2 (450 s compared to 296 s) makes the use of the rocket engine preferable for a greater energy increment ($\Delta v$). Despite the lower thrust of the LOX/LH2 engine (61 kN compared to 137 kN), the longer burn time of 150 s (compared to 58 s) leads to a greater energy increment. In addition, due to the higher $I_{sp}$ of LOX/LH2, less propellant is burned. Even though the liquid engine has a lower thrust with similar chamber pressure and nozzle expansion ratio, the overall complexity is higher than the solid rocket engine, and therefore its dry mass is more than twice as much. The sum of the propellant mass and rocket engine weight is 3182 kg for the liquid case and 3399 kg for the solid engine. This is the major contribution to the different launch masses in Table 4. Other minor effects are associated with different jet engine sizing, jet fuel consumptions, and structural differences. This example illustrates how the model and the resulting data can be used to not only understand architectural decisions but also the technical tradeoffs on a system level.

V. Conclusions

This paper aimed to address the question of which system architecture of a suborbital space tourism vehicle will provide the best combination of launch mass and safety. The paper approached this problem by first defining the plausible architectural space. With this approach, a set of 33 feasible architectures, which can be described in a single matrix, could be found. Of these 33, 26 had not been proposed earlier. Based on the architectural descriptions, this paper optimized each of the 33 architectures with respect to two objective functions: launch mass and safety. The launch mass was calculated by a design framework based on Frank [14] with two significant modifications. To quantitatively evaluate the safety, Dulac and Leveson's [51,52] approach was followed. Severity and likelihood factors were used to characterize the risk of each hazard, and a mitigation factor quantified the potential of the individual architectural features to mitigate the hazard. The paper found some main safety trends in designing a suborbital vehicle. A validation with 11 experts showed the validity of this paper’s assessment. The depth of modeling allowed performing informed tradeoffs on the system level, like the type of rocket engine. Finally, on architectural level, the comparison of the Pareto fronts of the architectures was possible. This led to the following conclusions for four-participant suborbital space tourism vehicles:

1) Wings are safer but heavier; wings on the second module are safer but heavier compared to wings on first module.
2) Architectures with one rocket engine completely dominate architectures with two rocket engines, and they dominate those with two rocket engines and an additional jet engine.
3) Single-module architectures weigh on average half of two-module architectures but are slightly riskier.
4) A vertical takeoff design results in a slightly lighter vehicle but increases risk by 50% on average.
5) The six dominant architectures (1, 3, 4, 5, 14, and 9) merit a more refined design analysis.

The value of this paper’s analysis lies in the comparison of the vast architectural space. This paper evaluated and optimized 26 novel architectures with thousands of variants and identified major tradeoff criteria. The limitation of this paper’s analysis was that the launch mass was used as a proxy for cost. Even though there is a strong correlation between mass and cost, a life-cycle cost assessment would be better. The extension of the described approach and the developed code to orbital launch vehicle can be an area of future work.

Appendix:

A.1 Steps 1–2: Hazard List with Assigned Severity and Likelihood Factors

The first step is to identify the hazards that might occur during system operation. For some systems, government agencies have mandated attention to specific hazards. For example, the U.S. Department of Defense prescribes a minimum baseline of hazards that must be considered when constructing a nuclear weapon system [51]. However, in most cases, the hazards must be determined by the analyst. There are a few structured ways to enable individuals or a group of people to apply their knowledge (e.g., with what-if questions). For an extensive list of possible activities, refer to [50].

Especially in the space sector, publicly available generic hazards lists for the space shuttle [54] and the Constellation Program, we derived a list of system-level hazards for each module. Table A1 shows them organized by mission phase, starting with the generally applicable hazards and followed by
hazards in the ascent phase as well as the descent phase. The descent phase of the uncrewed module is not considered, since no hazards threaten participants. For example, a fire or explosion of the uncrewed module during descent does not directly harm the participants. However, it is not desirable, as this hazard jeopardizes the mission success and destroys the equipment. Compromises to

<table>
<thead>
<tr>
<th>Module</th>
<th>Phase</th>
<th>ID</th>
<th>Hazard</th>
<th>severity factor $f_S$</th>
<th>likelihood factor $f_L$</th>
<th>resulting risk factor $f_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crewed</td>
<td>General</td>
<td>C-G1</td>
<td>Loss of life support of crewed module (including power, temperature, oxygen, air pressure, CO$_2$, food, water, etc.)</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Ascent</td>
<td>C-A1</td>
<td></td>
<td>Flammable substance in presence of ignition source during ascent of crewed module (fire + explosion)</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Ascent</td>
<td>C-A2</td>
<td></td>
<td>Incorrect propulsion/trajectory/control during ascent of crewed module</td>
<td>3</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Ascent</td>
<td>C-A3</td>
<td></td>
<td>Loss of structural integrity during ascent of crewed module (due to aerodynamic loads, vibrations, etc.)</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Descent</td>
<td>C-D1</td>
<td></td>
<td>Flammable substance in presence of ignition source during descent of crewed module (fire + explosion)</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Descent</td>
<td>C-D2</td>
<td></td>
<td>Incorrect propulsion/trajectory/control during descent of crewed module</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Descent</td>
<td>C-D3</td>
<td></td>
<td>Loss of structural integrity during descent of crewed module (due to aerodynamic loads, vibrations, etc.)</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Uncrewed</td>
<td>General</td>
<td>U-G1</td>
<td>Incorrect stage separation between uncrewed and crewed module</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Ascent</td>
<td>U-A1</td>
<td></td>
<td>Flammable substance in the presence of ignition source during ascent of uncrewed module (fire + explosion)</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Ascent</td>
<td>U-A2</td>
<td></td>
<td>Incorrect propulsion/trajectory/control during ascent of uncrewed module</td>
<td>3</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Ascent</td>
<td>U-A3</td>
<td></td>
<td>Loss of structural integrity during ascent of uncrewed module (due to aerodynamic loads, vibrations, etc.)</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Table A1 Hazard list with associated severity factor $f_S$, likelihood factor $f_L$, and resulting risk factor $f_R$ [hazards are grouped by phases and crewed vs uncrewed module designs and described, together with their identifications (ID)].

Fig. A1 Hazard mitigation database (description is in the text) (prop, propellant; traj, trajectory; TM, take-off mode; LAm, landing mode of the first module; and HP, horizontal powered).
these other categories are not taken into account in our safety-risk assessment.

The loss of life support hazard C-G1 has the lowest value caused by a relatively moderate severity 3 combined with a low likelihood 2. Some of the hazards have equal risk values but different individual entries for the severity and likelihood. If the module explodes, the participants will lose their lives. Compared to this, the severity of an incorrect propulsion/trajectory/control is lower, as there are possibilities to reestablish a secure state. However, an engine loss or incorrect trajectory (U-A2 or C-A2) is more likely to occur than an explosion. The higher likelihood combined with the lower severity leads to the same risk amplitude. Excluding the likelihood factor at this stage would lead to unreasonably higher values for the highly severe but less likely hazards, compared to the less severe but more likely ones. The risk factor varies from 6 to 12. Except for C-G1, the risk factors vary around ±20% from 10. Because of the linear weighting, this also means that the ORSRM is affected by around 20%. As seen in the result Sec. IV, the ORSRM score varies approximately from 1 to 9. This means that the mitigation scale has significantly more influence on the safety assessment than the severity and likelihood factors (which result in the risk factor). In standard space projects, a mitigation strategy is developed during the design to reduce these risks. We account for this effect by estimating this mitigation potential using the mitigation factor described in the following section.

A.2 Step 3: Mitigation Potential

After building the hazard list and calculating the risk factors, we determine the mitigation potential depending on the architectural decisions. Leveson developed the mitigation impact scale shown in Table A2. It ranges from 1 to 4, with higher numbers meaning more impact. The highest factor means that the design choice can completely mitigate the hazard; e.g., the choice of a single module vehicle mitigates the incorrect stage separation U-G1 (see Fig. A1). The dash means that the design choice has no mitigation potential. With this scale and the hazards, we build a matrix called the hazard mitigation database. It is shown in Fig. A1. The columns show the hazards identified in Table A2. The rows are the architectural decision with their options as well as two important design variables, which impact the safety of the vehicle. These are the number of pilots varying from 0 to 2 and the type of propellant: solid, liquid, or hybrid. Including these design variables results in up to 27 different values within an architecture (three options for the number of pilots, three options for module-1 propellant type, and three options for module-2 propellant type). However, not every combination leads to a variant on the Pareto front. If we exclude the design variables from the risk assessment, each architecture has its own independent ORSRM value.

To populate Fig. A1, we first decide on whether the decision has a mitigation potential on the hazard or not. For example, the number of modules has no influence on the fire and explosion hazard of the crewed module C-G1. Every noninfluenced combination is marked by a dash. For the other fields, we discussed the factors and compared the options inside a decision. The values shown in Fig. A1 are called mitigation factors.

Since the hazards differ for the uncrewed and crewed modules, the decisions of modules 1 and 2 have to be matched depending on the number of modules. If the architecture is a single-module vehicle, module 1 is the crewed module, and the uncrewed module does not exist (see the discussion in Fig. 2). And therefore, hazards U-A3 until U-A3 are not applicable (red shaded box). For a two-module vehicle, the uncrewed module is module 1, and module 2 matches the crewed module. In this case, all hazards are applicable (blue shaded boxes).

The vehicle decision is independent of the number of modules, and the mitigation factors always apply to the hazards (gray box). However, if the vehicle is a single-module architecture, the right four columns with the uncrewed module hazards are not considered in the assessment.

Acknowledgments

This work was partially funded by the Heinrich und Lotte Mühlfenzl Stiftung. The authors would like to thank Christopher P. Frank and Dimitri Mavris for sharing their code with us and we gratefully acknowledge their contributions.

References


doi:10.1115/1.2893288


doi:10.1002/sys.1529-6858


doi:10.2514/1.A32562


doi:10.1109/6.591665


