

Strategic risk dominance in collective systems design

Paul T. Grogan¹ and Ambrosio Valencia-Romero¹

¹ School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, USA

Abstract

Engineered system architectures leveraging collaboration among multiple actors across organizational boundaries are envisioned to be more flexible, robust, or efficient than independent alternatives but also carry significant downside risks from new interdependencies added between constituents. This paper transitions the concept of risk dominance from equilibrium selection in game theory to engineering design as a strategic measure of collective stability for system of systems. A proposed method characterizes system design as a bi-level problem with two or more asymmetric decision-makers. A measure of risk dominance assesses strategic dynamics with respect to the stability of joint or collaborative architectures relative to independent alternatives using a novel linearization technique to approximate linear incentives among actors. An illustrative example case for an asymmetric three-player design scenario shows how strategic risk dominance can identify and mitigate architectures with unstable risk-reward dynamics.

Key words: game theory, value-driven design, systems engineering, system of systems, collaborative design

1. Introduction

Collaboration across organizational boundaries in engineered systems presents an important tradeoff during conceptual design. Distributed architectures where multiple actors cooperate for mutual benefit seek superior performance compared to independent architectures. Potential improvement may derive from increased flexibility (de Weck *et al.* 2004), robustness (Brown and Eremenko 2006), or efficiency (Oates 2008). However, collaborative architectures also introduce new interdependencies between constituent systems which can lead to degraded performance or failure if not treated as a system of systems (Maier 1998).

Understanding and reasoning about the tradeoff between risk and reward in collaborative system architectures remains a critical area of research. It underlies ongoing challenges in managing inter-agency collaboration in joint projects (National Research Council 2011) as well as new laws requiring public agencies such as the National Oceanic and Atmospheric Administration (NOAA) to consider purchasing data from commercial providers to supplement government missions (United States of America 2017). Improved methods to understand fundamental collaborative dynamics early in the conceptual design and architecture selection phase could help avoid costly coordination failures.

Risk fundamentally deals with the interaction between probability and consequence of alternative scenarios (Kaplan and Garrick 1981). In engineering design, decision-makers routinely employ risk analysis methods to help choose

Received 10 January 2019
Revised 14 October 2019
Accepted 16 October 2019

Corresponding author

P. T. Grogan
pgrogan@stevens.edu

Published by Cambridge
University Press
© The Author(s) 2019
Distributed as Open Access under
a CC-BY 4.0 license
(<http://creativecommons.org/licenses/by/4.0/>)

Des. Sci., vol. 5, e24
journals.cambridge.org/dsj
DOI: 10.1017/dsj.2019.23

the **Design Society**
a worldwide community

 **CAMBRIDGE**
UNIVERSITY PRESS

among alternative design concepts (Lough *et al.* 2009). Traditional perspectives on risk for engineered systems consider potential impacts of external factors such as natural disasters or attacks and internal factors such as component fatigue, failure, or error. However, collaborative systems exhibit an additional source of uncertainty attributed to coordination failures among interacting decision-makers. This strategic source of risk is not addressed by methods that view engineering design as a centralized decision-making process.

This paper transitions the analytical concept of risk dominance from equilibrium selection in game theory to engineering design to measure the relative risk of coordination failures in collaborative systems at a level suitable for conceptual trade studies. Risk dominance recognizes the fragility of joint decisions and seeks to balance potential rewards with downside risk of coordination failure. This paper contributes a method to formulate collective design problems as a strategic design game and measure risk dominance. This paper provides a rigorous treatment of risk dominance in engineering design with two or more asymmetric players and overcomes barriers in transitioning fundamental theory to application. Results of this work can be used to inform conceptual phase architecture trades between collaborative and independent alternatives.

The remainder of this paper is organized as follows. Section 2 reviews applications of economics and game theory in engineering design literature and introduces the stag hunt game as the intellectual foundation for collaborative systems. Section 3 proposes a method to formulate collective systems design as strategic design games to assess strategic risk dominance with two or more players. Section 4 introduces an application case to show how risk dominance can identify and mitigate potential sources of coordination failure in an asymmetric three-player design scenario. Finally, a conclusion summarizes contributions, assumptions and limitations, and future work.

2. Background

2.1. Utility-based methods in systems design

This section builds on a line of literature dating to early works by Simon (1959) that develop and apply economic methods for decision-making to a broad class of problems including engineering design. From this perspective, engineering design selects the alternative concept with highest value (under expectation) measured using von Neumann–Morgenstern utility (Hazelrigg 1998). This approach is normative to organize and process individual preferences to support decision-making rather than descriptive to explain why certain decisions were made (Thurston 2001). Multi-criteria decision analysis methods including multi-attribute utility theory, analytic hierarchy process, and others help to formulate decisions for complex problems (Velasquez and Hester 2013).

Applying multi-criteria decision analysis to collective systems design problems characteristically transforms each participating actor's preferences into objective functions, defining an optimization problem. Although arguments have been made both against and in favor of this approach, most engineering design literature does not address essential uncertainty resulting from interactive effects between independent decision-makers. For instance, comparing engineering design to a social choice problem and grounded on Arrow's (1963) Impossibility Theorem, Hazelrigg (1997) argues that an optimal solution can only be reached

if all actors share the same utility function; otherwise, any attempt to maximize aggregated individual gains is bound to result in 'irrational' outcomes – assuming that a 'rational' design maximizes every designer's expected utility. In contrast, Scott and Antonsson (1999) state the purpose of engineering problems is to meet requirements, not individuals' wishes, and thus engineering design exists on a continuum between single- and multi-actor problems and any disparity between system requirements and decision-makers' preferences should be made explicit. However, in a collective systems design process with independent decision authority, actors strategically retain information and, as a result, do not have full knowledge about each other's preferences.

Recent design research emphasizes value-centric or value-driven design methods building on rational decision-making theory to maximize system value rather than meeting requirements at minimum cost (Collopy and Hollingsworth 2011). A related class of methods for tradespace exploration enumerate a design space and evaluate one or more design attributes to visualize a set of alternatives (Ross *et al.* 2004). These approaches generally treat risk as uncertainty or variation in value which can be analyzed with Monte Carlo sampling of a stochastic value function (Walton and Hastings 2004; O'Neill *et al.* 2010; Daniels and Paté-Cornell 2017). When applied to collective systems design, tradespace exploration results in computationally intense calculations due to combinatorial factors of a large design space coupled with interactive effects between actors. More importantly, this type of risk should not be considered as an explicit attribute to be traded during concept evaluation but only as uncertainty on other attributes (Abbas and Cadenbach 2018). More focused analysis of strategic dynamics is necessary to understand risk in collaborative systems.

2.2. Game-theoretic methods in systems design

In an attempt to reconcile multi-criteria decision analysis methods and limitations imposed from social choice applied to engineering design, Franssen and Bucciarelli (2004) demonstrate how a game-theoretic approach to collective systems design can help designers reach satisfactory outcomes without disregarding the implications of conflicting actors' preferences. Game theory analyzes strategic decision-making among multiple interacting actors or players. In the context of engineering design, the players are the design actors – cognizant individuals, computational agents, design organizations, or indirect stakeholders with undefined strategic interests (usually modeled as players from 'nature'). Individual utility functions, typically referred to as payoff functions, describe the players' preferences over the possible outcomes resulting from their actions. Each player also has a set of complete contingent plans or sequences of actions, i.e. strategies, from which to choose to maximize their expected gains depending on how much information they have about the other players' actions.

Game-theoretical approaches to engineering design frequently do not provide enough details about what aspects of the problem equate to 'strategies' and what other elements constitute the strategic setting of collective action or 'game' to be studied. An existing body of work applies game theory to engineering design by equating design alternatives to strategy spaces (Vincent 1983; Lewis and Mistree 1997; Briceño 2008; Wernz and Deshmukh 2010). In these works, the strategy set is composed of continuous functions linked to the functional attributes of the system and game-theoretical methods inform design decision-making at a

Table 1. Stag hunt game with $u_i = \frac{2}{3}$

Player 1 Strategy (s_1)	Player 2 Strategy (s_2)	
	Hare (ϕ_2)	Stag (ψ_2)
Hare (ϕ_1)	2	3
Stag (ψ_1)	0	4
	3	4

low level of abstraction. Normative methods evaluate and select design decisions based on idealistic scenarios of cooperative or non-cooperative strategic equilibria (Papageorgiou *et al.* 2016).

The majority of contributions to engineering design literature grounded in game-theoretic methods search for stable design sets under strong assumptions of rationality – namely Nash equilibria – as solutions to the multi-actor design problem. To improve outcomes designated by Nash equilibria, some works develop methods to further explore the strategy space beyond rational reaction strategy sets (Gurnani and Lewis 2008; Herrmann 2010). Other works explore subgame perfect equilibria as solution concepts to game-theoretical models of engineering design (Bhatia *et al.* 2016; Kang *et al.* 2016). Finding equilibria is computationally intense and sensitive to the strategy space definition, thus the number of design alternatives to be assessed should be kept at a minimum to allow for the best use of classical game-theoretical methods. Moreover, the existence of more than one equilibrium still leaves selection of a ‘best’ option in the air and rekindles the debate about what an optimal solution means from an rational/objective perspective.

2.3. Stag hunt game and risk dominance

In contrast to existing works applying game theory to engineering design which conflate design and strategy decisions for general problems, this paper adopts a simple strategic context to evaluate dynamics for a specific class of problems related to collective systems design. Analysis of risk dominance, a concept from equilibrium selection literature, provides insights about the relationships between interacting design actors and the relative stability of collaborative systems.

The stag hunt is a canonical game theory problem that models fundamental challenges in collective decision-making (Skryms 2004). It follows the narrative of two hunters deciding between two alternatives to either hunt *stag* or *hare*. A stag hunt provides a desirable reward but requires joint participation of both hunters (i.e. a single stag hunter goes home hungry). A hare hunt yields a modest reward and can either be performed alone or jointly. Depending on the particular game, an independent hare hunt may be more or less rewarding than a joint hare hunt; however, both cases must be preferred to a failed stag hunt and less desirable than a successful stag hunt.

Table 1 shows a normal form payoff matrix for an example symmetric two-player stag hunt game with payoffs of 2 for a joint hare hunt, 3 for an individual hare hunt, 4 for a successful stag hunt, and 0 for a failed stag hunt. Rather than

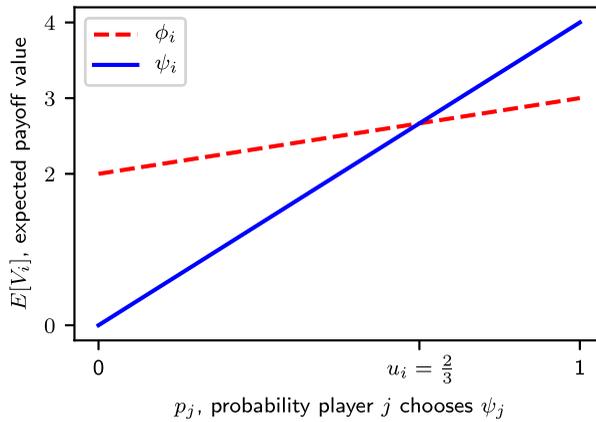


Figure 1. The expected value of strategies for player i as a function of p_j , the probability that player j chooses ψ_j for the example game in Table 1.

absolute wealth or resource quantities (e.g. amount of food), this paper treats payoff values as von Neumann–Morgenstern utilities that already account for behavioral factors such as loss aversion and diminishing sensitivity within each strategy context (Arrow 1971; Tversky and Kahneman 1992).

The strategy space $\mathcal{S}_i = \{\phi_i, \psi_i\}$ denotes the hare and stag strategies, respectively, for player i . Inspection of selected strategy sets for two players (s_1, s_2) shows both $\phi = (\phi_1, \phi_2)$ and $\psi = (\psi_1, \psi_2)$ are Nash equilibria because neither player has unilateral incentive to deviate away from these points. From an equilibrium perspective both stag and hare strategies are stable; however, there are clear differences in risk and reward.

Harsanyi and Selten (1988) develop theory for equilibrium selection in bipolar games based on the concept of risk dominance. Similar to how some equilibria exhibit payoff dominance (i.e. the stag equilibrium ψ yields higher payoffs), risk dominance is a desirable feature that captures resistance to losses.

In specific games such as the stag hunt with two strategies and two Nash equilibria (bipolar games), further analysis compares alternative strategies from a rational (expected value maximization) perspective. Figure 1 visualizes the expected value for player i as a function of the probability that player j chooses strategy ψ_j (i.e. $p_j = Pr\{s_j = \psi_j\}$). For low values of p_j the hare strategy ϕ_i provides the highest expected value. Similarly, for high values of p_j the stag strategy ψ_i provides the highest expected value. However, the two lines intersect at a point $0 < u_i < 1$ which measures the minimum probability of player j choosing strategy ψ_j for it to be rational for player i to choose strategy ψ_i .

A closed-form expression in Eq. (1) computes u_i for any two-player bipolar game for a given payoff function V .

$$u_i(\phi, \psi) = \frac{V_i(\phi_i, \phi_j) - V_i(\psi_i, \phi_j)}{(V_i(\phi_i, \phi_j) - V_i(\psi_i, \phi_j)) - (V_i(\psi_i, \psi_j) - V_i(\phi_i, \psi_j))}. \quad (1)$$

For the example in Table 1, $u_1 = u_2 = 2/3$. In other words, the expected value-maximizing decision for either player is to choose a stag hunt if and only if they estimate a better than two-in-three chance that their partner chooses a stag hunt.

Table 2. Stag hunt game with $u_i = 0.2$

Player 1 Strat. (s_1)	Player 2 Strategy (s_2)	
	Hare (ϕ_2)	Stag (ψ_2)
Hare (ϕ_1)	2	3
Stag (ψ_1)	2	0
	0	11
	3	11

Table 3. Stag hunt game with $u_i = 0.8$

Player 1 Strat. (s_1)	Player 2 Strategy (s_2)	
	Hare (ϕ_2)	Stag (ψ_2)
Hare (ϕ_1)	2	3
Stag (ψ_1)	2	-2
	-2	4
	3	4

Variations on the stag hunt game produce different values of u_i . For example, Table 2 shows an alternative game for a scenario with a ‘trophy’ stag which increases the upside payoff from 4 to 11, lowering the threshold for economic collaboration to $u_i = 0.2$. Alternatively, Table 3 shows a scenario with an ‘injury’ incurred from a solitary stag hunt which decreases the downside payoff from 0 to -2, raising the threshold for economic collaboration to $u_i = 0.8$. These scenarios show that the structure of the payoff function influences perception of collaboration and is central to the concept of risk dominance.

In the literature, u_i is called the *normalized deviation loss* because its mathematical expression resembles a loss associated with deviating away from the equilibrium ϕ in the numerator normalized by the total losses associated with deviating away both equilibria in the denominator. Counter-intuitively, large deviation losses insulate a decision. The example in Table 2 shows the deviation loss from 11 to 3 promotes the stag strategy while the example in Table 3 shows the large deviation loss from 2 to -2 promotes the hare strategy.

Selten (1995) proposes a quantitative metric to measure risk dominance for bipolar games with linear incentives called the weighted average log measure (WALM) of risk dominance (see Appendix A for details). Linear incentives assume each player’s payoff can be expressed as a linear combination of whether other players participate in the collective strategy and is further addressed in Appendix B. Equation (2) defines the WALM of risk dominance for an n -player game, where u_i are normalized deviation losses and w_i are influence weights based on an influence matrix A which measures player interdependence (Selten 1995).

$$R(\phi, \psi) \equiv \sum_{i=1}^n w_i(A) \ln \frac{u_i}{1 - u_i}. \tag{2}$$

This expression maps bipolar games to a real-number scale that measures the risk dominance of the collective strategy ψ relative to the independent strategy

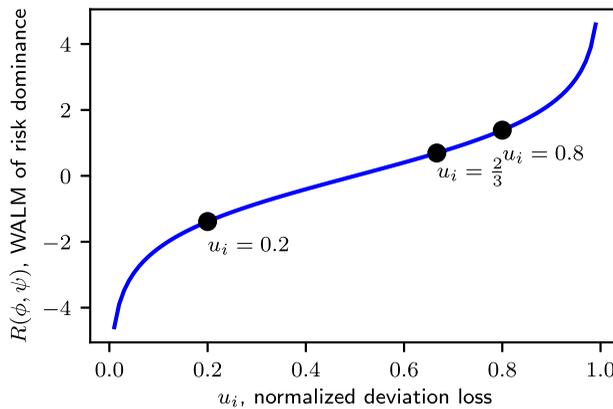


Figure 2. WALM of risk dominance is the logit function of the normalized deviation loss u_i for symmetric games with $n = 2$ players.

ϕ . For the objective case with no knowledge of other players’ actions (equivalent to $p_j = 0.5$), $R > 0$ indicates ϕ is risk dominant while $R < 0$ indicates ψ is risk dominant. In all other cases with partial information leading to a probability distribution $f(p_j)$, R provides a relative measure of risk dominance.

WALM risk dominance can be simplified for games with $n = 2$ players because weights are defined as $w_1 = w_2 = 1/2$. Furthermore, payoff symmetry with $u_1 = u_2 = u_i$ further reduces the expression to Eq. (3) which is simply the logit function of u_i visualized in Figure 2.

$$R(\phi, \psi) = \ln \frac{u_i}{1 - u_i} = \ln \left(\frac{V_i(\phi_i, \phi_j) - V_i(\psi_i, \phi_j)}{V_i(\psi_i, \psi_j) - V_i(\phi_i, \psi_j)} \right). \quad (3)$$

The stag hunt games in Tables 1–3 have $R_1 = \ln 2 \approx 0.69$, $R_2 = \ln 0.25 \approx -1.39$, and $R_3 = \ln 4 \approx 1.39$. In Table 2, $R_2 < 0$ indicates the collective strategy ψ is risk dominant. In Table 1 and Table 3, $R_3 > R_1 > 0$ indicates the independent strategy ϕ is risk dominant and *more strongly* so in Table 3 compared to Table 1. Risk dominance is normative for strategy selection in non-cooperative cases; however, in other cases it is only a relative measure of strategic dynamics across games.

While direct analysis of individual incentives in Figure 1 provides an intuitive explanation of strategic dynamics in two-player bipolar games, the more general formulation for WALM of risk dominance detailed in Appendix A handles asymmetric games with $n \geq 2$ players where players may express different or conflicting normalized deviation losses and pairwise interactions.

3. Risk dominance for collaborative systems

This section develops a method to convert a collective system design problem into a strategic design game to formulate and measure risk dominance.

3.1. Strategic design games

Game theory is a strategic analysis method that works at a high level of abstraction. In the context of engineering design, strategic decision-making best corresponds to architecture selection in conceptual design rather than more detailed design

Table 4. Stag hunt design utilities

Design	Strategic Context			
	ϕ_i, ϕ_j	ϕ_i, ψ_j	ψ_i, ϕ_j	ψ_i, ψ_j
Atlatl	0	0	0	4
Bow	1	1	1.5	3.5
Club	1	1	0	2
Dog	2	3	0	0

Table 5. Stag hunt game with $u_i = 0.5$

Player 1	Player 2 (s_2, d_2)	
(s_1, d_1)	ϕ_2, Dog	ψ_2, Bow
ϕ_1, Dog	2	3
	2	1.5
ψ_1, Bow	1.5	3.5
	3	3.5

decisions in preliminary design. This section defines the concept of a strategic design game as a multi-actor value model for engineering decision-making to permit strategic analysis (Grogan *et al.* 2018).

A strategic design game distinguishes between two levels of decisions: strategy decisions $s_i \in \mathcal{S}_i$ govern collective behavior among players and design decisions $d_i \in \mathcal{D}_i$ specify system configurations. A corresponding multi-actor value function $[V_i^{s_1, \dots, s_n}(d_1, \dots, d_n)]$ for n players in Eq. (4) maps design and strategy decisions to values (utilities) for each player.

$$V : \prod_{i=1}^n \mathcal{D}_i \times \mathcal{S}_i \rightarrow \mathbb{R}^n. \tag{4}$$

While the design spaces \mathcal{D}_i may be large or unbounded and unique to each player, the strategy space $\mathcal{S}_i = \{\phi_i, \psi_i\}$ is limited to two options: choosing between independent (ϕ_i) or collective (ψ_i) action. Within each strategy space, design decisions are evaluated based on the context of the governing strategy.

To illustrate this concept, reconsider the stag hunt game from Table 1 with design variables to choose the hunting weapon from the symmetric design space $\mathcal{D} = \{\text{Atlatl}, \text{Bow}, \text{Club}, \text{Dog}\}$. A multi-actor value model in Table 4 evaluates each design alternative in possible strategic contexts. Within fixed equilibrium contexts the best hare-hunting design is Dog with utility 2 and the best stag-hunting design is Atlatl with utility 4. However, as previously investigated in Table 1, this combination requires a probability greater than $u_i = 2/3$ to pursue a stag hunt. In cases with unreliable partners, the stag-hunting design Bow in Table 5 may be more desirable because it only requires a probability greater than $u_i = 0.5$ to pursue a stag hunt, although it only provides utility 3.5 if successful.

Table 6. Strategic design game for three players

	$s_2 = \phi$	$s_2 = \psi$	$s_2 = \phi$	$s_2 = \psi$
$s_1 = \phi$	\mathcal{V}_1^ϕ \mathcal{V}_2^ϕ \mathcal{V}_3^ϕ	\mathcal{V}_1^ϕ $V_2^\psi(d_2)$ \mathcal{V}_3^ϕ	\mathcal{V}_1^ϕ \mathcal{V}_2^ϕ $V_3^\psi(d_3)$	\mathcal{V}_1^ϕ $V_2^\psi(d_2, d_3)$ $V_3^\psi(d_2, d_3)$
$s_1 = \psi$	$V_1^\psi(d_1)$ \mathcal{V}_2^ϕ \mathcal{V}_3^ϕ	$V_1^\psi(d_1, d_2)$ $V_2^\psi(d_1, d_2)$ \mathcal{V}_3^ϕ	$V_1^\psi(d_1, d_3)$ \mathcal{V}_2^ϕ $V_3^\psi(d_1, d_3)$	$V_1^\psi(d_1, d_2, d_3)$ $V_2^\psi(d_1, d_2, d_3)$ $V_3^\psi(d_1, d_2, d_3)$
	$s_3 = \phi$		$s_3 = \psi$	

Applied to engineering design, the independent strategy ϕ_i corresponds to systems designed and operated with few external dependencies. The collective strategy ψ_i represents potential performance gains from collaborative systems having numerous interdependencies which operate at risk of degraded performance due to coordination failures. The resulting strategic design game resembles a binary game with zero, one, or two pure strategy Nash equilibria. Games with zero equilibria have no stable strategy sets, limiting the use of normative analysis. Games with one equilibrium exhibit a dominant strategy and do not benefit from further analysis. Therefore, only cases with two equilibria (i.e. bipolar games) benefit from and are valid to measure risk dominance.

More detailed analysis of a strategic design game benefits from two simplifying assumptions in Eq. (5) and shown in Table 6 for a normal form strategic design game with $n = 3$ players.

$$V_i^{s_1, \dots, s_n}(d_1, \dots, d_n) = \begin{cases} V_i^\phi(d_i) & \text{if } s_i = \phi_i \\ V_i^\psi(d_k : s_k = \psi_k) & \text{otherwise.} \end{cases} \quad (5)$$

These assumptions limit interaction effects between participants inside and outside a collective strategy and are not strictly required for analysis but help to communicate the method and results using simplified notation.

The first simplification approximates the multi-actor value function by a single-actor value function V_i^ϕ when player i chooses an independent strategy ϕ_i . This assumes no strong interaction effects between independent players and others and allows local design optimization in Eq. (6).

$$\mathcal{V}_i^\phi = \max_{d \in \mathcal{D}_i} V_i^\phi(d). \quad (6)$$

The second simplification aggregates all candidate designs for players participating in a collective strategy ψ including player i . For the special case where no other players select the collective strategy (i.e. $s_j = \phi_j \forall j \neq i$), a single-actor value function $V_i^\psi(d_i)$ replaces the multi-actor one. This formulation emphasizes interaction effects among participants in a collective strategy but assumes no strong interaction effects with players outside.

3.2. Strategic risk dominance

This section uses the strategic design game concept to explain and measure risk dominance in collaborative systems design. Risk dominance is only meaningful

in bipolar games which requires a specific ordering of value quantities such that $V_i^\psi(d_1, \dots, d_n) > \mathcal{V}_i^\phi > V_i^\psi(d_i) \forall i$. This is a reasonable requirement because it represents the cases of most interest where the collaborative system has a higher upside potential but also downside risk to an independent alternative.

As introduced in Section 2.3 and detailed in Appendix A, risk dominance first depends on normalized deviation losses. Equation (7) shows the normalized deviation loss for player i by substituting the strategic design game notation into Eq. (1).

$$u_i = \frac{\mathcal{V}_i^\phi - V_i^\psi(d_i)}{\left(\mathcal{V}_i^\phi - V_i^\psi(d_i)\right) + \left(V_i^\psi(d_1, \dots, d_n) - \mathcal{V}_i^\phi\right)} = \frac{\mathcal{V}_i^\phi - V_i^\psi(d_i)}{V_i^\psi(d_1, \dots, d_n) - V_i^\psi(d_i)}. \tag{7}$$

The simplest form of strategic risk dominance assumes symmetric design spaces and utility functions for all players. Symmetry yields equal weighting factors w_i , producing the risk dominance measure in Eq. (8) as a function of the collaborative design d (note: all value subscripts dropped due to symmetry).

$$R_\psi(d) = \ln \left(\frac{\mathcal{V}^\phi - V^\psi(d)}{V^\psi(d, \dots, d) - \mathcal{V}^\phi} \right). \tag{8}$$

Symmetric risk dominance requires three function evaluations to quantify the upside value of a successful collective strategy $V^\psi(d, \dots, d)$, downside value of a failed collective strategy $V^\psi(d)$, and the value of the independent alternative \mathcal{V}^ϕ .

More general forms of strategic risk dominance assume asymmetric design spaces or utility functions. As detailed in Appendix A, influence elements a_{ij} measure the dependence between players i and j . Equation (9) shows the influence elements a_{ij} between players i and j , assuming linear incentives, by substituting the notation for strategic design games into Eq. (27).

$$a_{ij} = \frac{\left(\mathcal{V}_i^\phi - V_i^\psi(d_i)\right) - \left(\mathcal{V}_i^\phi - V_i^\psi(d_i, d_j)\right)}{\left(\mathcal{V}_i^\phi - V_i^\psi(d_i)\right) + \left(V_i^\psi(d_1, \dots, d_n) - \mathcal{V}_i^\phi\right)} = \frac{V_i^\psi(d_i, d_j) - V_i^\psi(d_i)}{V_i^\psi(d_1, \dots, d_n) - V_i^\psi(d_i)}. \tag{9}$$

Linear incentives enforce a constraint that all row sums total one, i.e. $\sum_{j \neq i} a_{ij} = 1 \forall i$. In cases with nonlinear incentives for $n = 3$ players, as in Table 6, Eq. (10) expresses linearized influence elements \bar{a}_{ij} between players i and j by substituting strategic design game notation into Eq. (2) (see Appendix B for details).

$$\bar{a}_{ij} = \frac{1}{2} (1 + a_{ij} - a_{ik}) = \frac{1}{2} \left(1 + \frac{V_i^\psi(d_i, d_j) - V_i^\psi(d_i, d_k)}{V_i^\psi(d_1, d_2, d_3) - V_i^\psi(d_i)} \right). \tag{10}$$

After computing elements of the influence matrix A using linear incentives a_{ij} or the linear approximation \bar{a}_{ij} , subsequent analysis finds weighting factors w_i to measure player importance. As detailed in Appendix A, weighting factors are the eigenvector corresponding to the unit eigenvalue of the influence matrix A .

The risk dominance measure in Eq. (11) computes the WALM in Eq. (2) as a function of collaborative designs (d_1, \dots, d_n) .

$$R_\psi(d_1, \dots, d_n) = \sum_{i=1}^n w_i(A) \ln \left(\frac{\mathcal{V}_i^\phi - V_i^\psi(d_i)}{V_i^\psi(d_1, \dots, d_n) - \mathcal{V}_i^\phi} \right). \quad (11)$$

In general, asymmetric risk dominance requires $1 + \sum_{k=1}^n \binom{n}{k} = 2^n$ multi-actor value function evaluations to consider all possible combinations of players joining the collective strategy (required to compute \bar{a}_{ij} terms) plus the independent alternative. While not burdensome for small games, this combinatorial factor may limit the use of similar methods for large games with many decision-makers.

3.3. Assumptions and limitations

This work includes several key assumptions and limitations which must be discussed. First, measuring risk dominance assumes a strategic decision-making process structured as a bipolar game. The strategic design game is an abstraction of the design process where, in reality, lower-level design and higher-level strategy decisions are coupled and iterative. Results from this work thus provide a baseline result which must be considered in the context of a specific design problem. While a limiting constraint, bipolar games are interesting cases to study because they present a fundamental tradeoff between risk and reward.

Second, this work represents an analytical and rational/objective method to measure strategic risk dominance which is both a significant limitation and a significant strength. The strategic-level analysis only aims to maximize expected value – although utility functions to compute payoff values can incorporate risk attitudes. No subjective information is required to determine or assess the likelihood of other players’ actions nor are there any elements of cooperative game theory to enforce contracts or share or divide benefits (although R is closely related to the Nash product and \bar{a}_{ij} terms resemble Shapley value). If additional subjective information were available, a more thorough analysis leveraging Bayesian games could be performed, as pioneered by Harsanyi (1967). In light of this limitation, relative values of the WALM risk dominance measure (e.g. during an architecture trade study) are more important than absolute values.

Third, the existing equilibrium selection theory imposes a few restrictions on the types of problems modeled. As discussed in Appendix A, it assumes linear incentives which are unlikely to hold in most engineering applications due to economies of scale; however, the linear approximation methods and error analysis introduced mitigate some of this concern. Developing utility functions to quantify payoff values while accommodating behavioral factors such as risk attitudes remains a practical challenge which is out of the scope of this paper.

Two other notational limitations can be relaxed for further analysis. The risk metric in Eq. (11) only assesses one set of collaborative architectures (d_1, \dots, d_n) under one collective strategy (ψ) relative to the independent baseline. However, as a real number, R can compare multiple design candidates relative to the same independent baseline (e.g. $R_\psi(d_1, \dots, d_n)$ vs. $R_\psi(d'_1, \dots, d'_n)$) to guide the design search process. Similarly, R can compare multiple collective strategy candidates for a fixed set of federated alternatives (e.g. $R_\psi(d_1, \dots, d_n)$ vs. $R_{\psi'}(d_1, \dots, d_n)$) to guide the strategy formulation process.

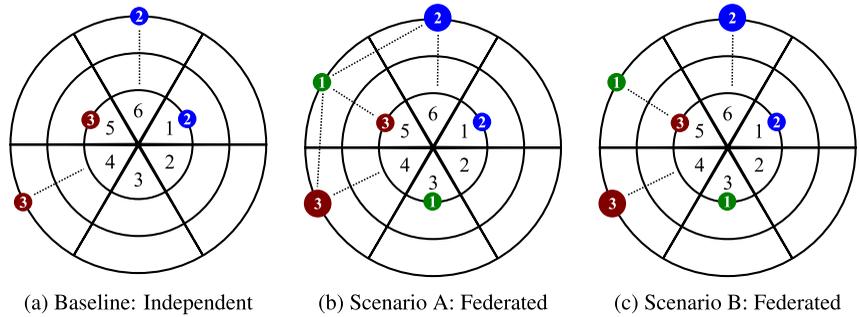


Figure 3. Initial ground station and satellite locations in a two-dimensional space numbered by player for (a) baseline, (b) scenario A, and (c) scenario B. Dotted lines indicate SGL or ISL for the initial conditions.

4. Illustrative application case

4.1. Multi-actor value model and design scenarios

This section develops an application case based on a stylized model of federated Earth-observing space systems paired with an existing simulation model. Orbital Federates Simulation – Python (OFSPY) (Grogan 2019) acts as a multi-actor value function by mapping design and strategy sets (inputs) to net present value earned by each player over a simulated system lifetime (outputs). The model includes stochastic features to capture operational uncertainty such that results must be sampled using Monte Carlo methods. While many model details are clearly fictional, the underlying model was developed to have structural and process isomorphic features to help understand strategic player behavior for space systems (Grogan and de Weck 2015).

This application case only evaluates how to model collaborative design scenarios as a strategic design game and compute measures of risk dominance. The implementation details of the multi-actor value model are outside the scope of this paper; however, Appendix C provides more details for replication.

The design scenario considers $n = 3$ players who operate space systems to collect and downlink data to satisfy demands and earn revenue. Figure 3 illustrates designs selected from a large combinatorial design space for a baseline (independent) strategic context and two federated alternatives. The independent case includes small standalone observing spacecraft for players 2 and 3 who specialize in synthetic aperture radar (SAR) and visual light (VIS) sensors, respectively. Player 1 does not participate in the baseline system. Federated designs consider an opportunistic data exchange policy with a fixed price for inter-satellite link (ISL) and space-to-ground (SGL) services among players. Federated scenario A includes participation by player 1 with a data relay spacecraft and SGL receiver and ISL adoption among all three players. Federated scenario B eliminates the ISL technology option and establishes an independent observing spacecraft for player 1 with the SGL receiver.

Table 7 shows the expected net present value for each strategic context using a discount rate of 2% per turn evaluated using 1000 seeded runs of the multi-actor value function. For clarity in presentation, players are assumed to be risk neutral such that expected net present value is equivalent to utility; however, alternative assumptions of risk attitudes would modify the resulting payoff values following a

Table 7. Orbital federates design scenarios

	Baseline: Independent (ϕ)	Scenario A: Federated (ψ)	Scenario B: Federated (ψ)
P. 1	d_1 : None.	d_{1A} : Satellite with ISL and open SGL. Ground station with open SGL.	d_{1B} : Satellite with SAR sensor and open SGL. Ground station with open SGL.
	$\mathcal{V}_1^\phi = 0$	$V_1^\psi(d_{1A}) = -1100$ $V_1^\psi(d_{1A}, d_{2A}) = -572$ $V_1^\psi(d_{1A}, d_{3A}) = -652$ $V_1^\psi(d_{1A}, d_{2A}, d_{3A}) = 100$	$V_1^\psi(d_{1B}) = -104$ $V_1^\psi(d_{1B}, d_{2B}) = 83$ $V_1^\psi(d_{1B}, d_{3B}) = 33$ $V_1^\psi(d_{1B}, d_{2B}, d_{3B}) = 271$
P. 2	d_2 : Satellite with SAR sensor and proprietary SGL. Ground station with proprietary SGL.	d_{2A} : Satellite with SAR sensor, ISL, and proprietary and open SGLs. Ground station with proprietary SGL.	d_{2B} : Satellite with SAR sensor and proprietary and open SGLs. Ground station with proprietary SGL.
	$\mathcal{V}_2^\phi = -39$	$V_2^\psi(d_{2A}) = -389$ $V_2^\psi(d_{1A}, d_{2A}) = 440$ $V_2^\psi(d_{2A}, d_{3A}) = -255$ $V_2^\psi(d_{1A}, d_{2A}, d_{3A}) = 536$	$V_2^\psi(d_{2B}) = -288$ $V_2^\psi(d_{1B}, d_{2B}) = 141$ $V_2^\psi(d_{2B}, d_{3B}) = -288$ $V_2^\psi(d_{1B}, d_{2B}, d_{3B}) = 150$
P. 3	d_3 : Satellite with VIS sensor and proprietary SGL. Ground station with proprietary SGL.	d_{3A} : Satellite with VIS sensor, ISL, and proprietary and open SGLs. Ground station with proprietary SGL.	d_{3B} : Satellite with VIS sensor and proprietary and open SGLs. Ground station with proprietary SGL.
	$\mathcal{V}_3^\phi = 30$	$V_3^\psi(d_{3A}) = -320$ $V_3^\psi(d_{1A}, d_{3A}) = 386$ $V_3^\psi(d_{2A}, d_{3A}) = -158$ $V_3^\psi(d_{1A}, d_{2A}, d_{3A}) = 639$	$V_3^\psi(d_{3B}) = -220$ $V_3^\psi(d_{1B}, d_{3B}) = 391$ $V_3^\psi(d_{2B}, d_{3B}) = -161$ $V_3^\psi(d_{1B}, d_{2B}, d_{3B}) = 378$

Table 8. Strategic design game for scenario A

	$s_2 = \phi$	$s_2 = \psi$	$s_2 = \phi$	$s_2 = \psi$
	0	0	0	0
$s_1 = \phi$	-39	-389	-39	-255
	30	30	-320	-158
	-1100	-572	-652	100
$s_1 = \psi$	-39	440	-39	536
	30	30	386	639
	$s_3 = \phi$		$s_3 = \psi$	

nonlinear utility curve. In practice, sensitivity analyses may help understand how unknown behavioral quantities such as risk attitudes influence results.

4.2. Risk dominance analysis of scenario A

Table 8 populates a strategic design game for federated scenario A which constitutes a bipolar game with asymmetric players. Using Eq. (7), the normalized

deviation losses are

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0.917 \\ 0.378 \\ 0.365 \end{bmatrix}. \tag{12}$$

Using Eq. (10), the linearized influence matrix is

$$A = \begin{bmatrix} 0 & \bar{a}_{12} & \bar{a}_{13} \\ \bar{a}_{21} & 0 & \bar{a}_{23} \\ \bar{a}_{31} & \bar{a}_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0.533 & 0.467 \\ 0.876 & 0 & 0.124 \\ 0.784 & 0.216 & 0 \end{bmatrix}. \tag{13}$$

Incentive function linearization error analysis using Eq. (9) shows small errors for all three players: $\varepsilon = [0.047, 0.010, 0.024]$.

Eigenvector analysis of the transposed influence matrix A^T yields the eigenvector (rescaled to unit norm)

$$w(A) = \begin{bmatrix} w_1(A) \\ w_2(A) \\ w_3(A) \end{bmatrix} = \begin{bmatrix} 0.455 \\ 0.296 \\ 0.249 \end{bmatrix} \tag{14}$$

corresponding to the unit eigenvalue. Computing WALM as

$$R_\psi(d_{1A}, d_{2A}, d_{3A}) = \sum_{i=1}^3 w_i(A) \ln \frac{u_i}{1 - u_i} = 0.805 \tag{15}$$

shows the independent strategy ϕ to be risk dominant.

Normalized deviation losses show the independent strategy is preferred for player 1 in all but high probabilities of collaboration ($u_1 = 0.917$) because of large downside losses incurred if no other players choose the collective strategy. This risk fundamentally arises because player 1 has no independent source of revenue to recover the high cost of the federated design. However, the collective strategy has a lower threshold of collaboration for players 2 and 3 ($u_2 = 0.378, u_3 = 0.365$) because large upside gains realized from shared data services overcome the additional cost of larger spacecraft and extra modules.

Weighting factors identify player 1 as the most influential ($w_1 = 0.455$) which can be explained by their central role in both providing shared ISL relay and SGL downlink services via the spacecraft and ground station, respectively. Smaller but similar weights for players 2 and 3 ($w_2 = 0.296, w_3 = 0.249$) reflect their similarity in operational mission.

Given player 1's aversion to the collective strategy and strong influence, the independent strategy is risk dominant in this design scenario. In particular, the collective strategy is prone to failure due to disengagement by player 1. A visualization of value surfaces in Figure 4 emphasizes the disparity between players 1 compared to 2 and 3 with respect to overall stability of the collective strategy.

4.3. Risk dominance analysis of scenario B

Table 9 populates a strategic design game for the federated scenario B which constitutes a bipolar game with asymmetric players. Using Eq. (7), the normalized

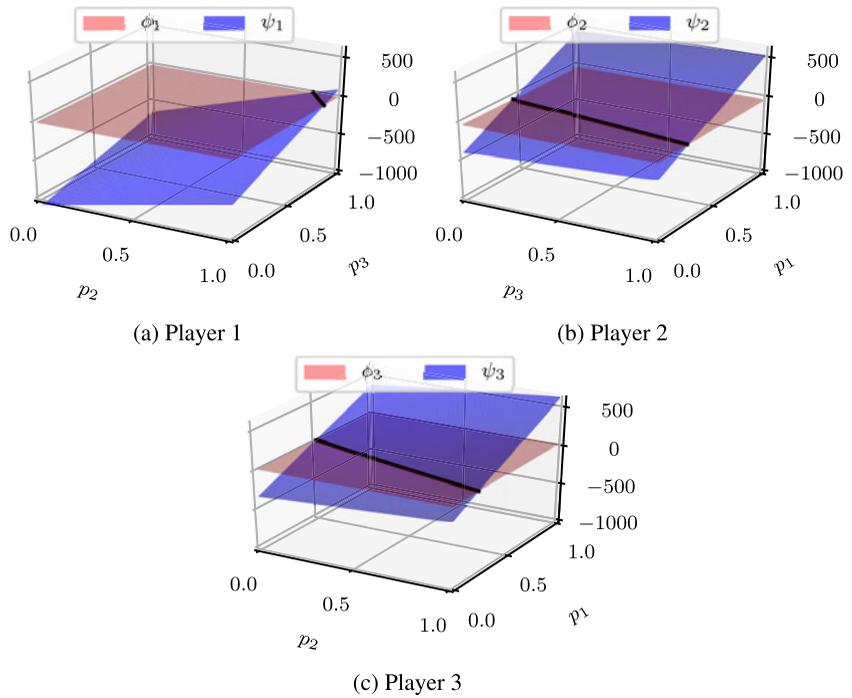


Figure 4. Value surfaces for independent strategy ϕ_i versus collective strategy ψ_i under federated scenario A. The independent strategy ϕ is more dominant for player 1 while the collective strategy ψ is more dominant for players 2 and 3.

Table 9. Strategic design game for scenario B

	$s_2 = \phi$	$s_2 = \psi$	$s_2 = \phi$	$s_2 = \psi$
$s_1 = \phi$	0	0	0	0
	-39	-288	-39	-288
	30	30	-220	-161
$s_1 = \psi$	-104	83	33	271
	-39	141	-39	150
	30	30	391	378
	$s_3 = \phi$		$s_3 = \psi$	

deviation losses are

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0.277 \\ 0.568 \\ 0.418 \end{bmatrix}. \tag{16}$$

Using in Eq. (10), the linearized influence matrix is

$$A = \begin{bmatrix} 0 & \bar{a}_{12} & \bar{a}_{13} \\ \bar{a}_{21} & 0 & \bar{a}_{23} \\ \bar{a}_{31} & \bar{a}_{32} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0.566 & 0.434 \\ 0.990 & 0 & 0.010 \\ 0.962 & 0.038 & 0 \end{bmatrix}. \tag{17}$$

Incentive function linearization error analysis using Eq. (9) shows small errors for all three players: $\varepsilon = [0.034, 0.005, 0.030]$.

Eigenvector analysis of the transposed influence matrix A^T yields the eigenvector (rescaled to unit norm)

$$w(A) = \begin{bmatrix} w_1(A) \\ w_2(A) \\ w_3(A) \end{bmatrix} = \begin{bmatrix} 0.494 \\ 0.288 \\ 0.217 \end{bmatrix} \quad (18)$$

corresponding to the unit eigenvalue. Computing WALM as

$$R_\psi(d_{1B}, d_{2B}, d_{3B}) = \sum_{i=1}^3 w_i(A) \ln \frac{u_i}{1 - u_i} = -0.466 \quad (19)$$

shows the collective strategy ψ is risk dominant.

Normalized deviation losses show the collective strategy is preferred for player 1 for a wide range of probabilities of collaboration ($u_1 = 0.277$). The threshold for collaboration is higher for players 2 and 3 ($u_2 = 0.568, u_3 = 0.418$) compared to scenario A. The goal of reducing upfront costs and providing an independent source of revenue successfully changed the strategic risk posture of player 1. However, the loss of ISL relay services and associated revenue disincentivize the collective strategy for players 2 and 3.

Weighting factors still identify player 1 as the most influential ($w_1 = 0.494$) and more influential compared to scenario A. Player 1 retains a key role in providing downlink services and, without ISL services among players 2 and 3, takes an even stronger role because other players lack the relay components to interact with each other directly. Players 2 and 3 retain similar weights ($w_2 = 0.288, w_3 = 0.217$) though both are less influential than in scenario A.

Combining these factors, the collective strategy is risk dominant in this design scenario. This result indicates scenario B has preferable strategic dynamics to scenario A at the cost of slightly lower value. Notably, in the event that player 2 disengages, players 1 and 3 still enjoy moderate returns from the collective strategy. A visualization of value surfaces in Figure 5 shows a clear difference for player 1 compared to scenario A in Figure 4.

4.4. Comparative analysis of results

This case illustrates how high-value (but also high-risk) designs emerge from optimization-oriented activities and how analysis of risk dominance has the potential to mitigate strategic instabilities by selecting more conservative alternatives. If successful, scenario A provides superior value for all three players by taking advantage of new technology and operational concepts. However, focusing on maximizing upside potential can yield unstable design solutions more susceptible to coordination failures. Player 1 experiences a significantly higher threshold for collaboration than others under scenario A and is most likely to disengage from a collective strategy.

Scenario B reduces the level of technological ambition and establishes an independent value source for all players. While its potential payoffs are smaller than A, scenario B exhibits superior strategic stability and is robust

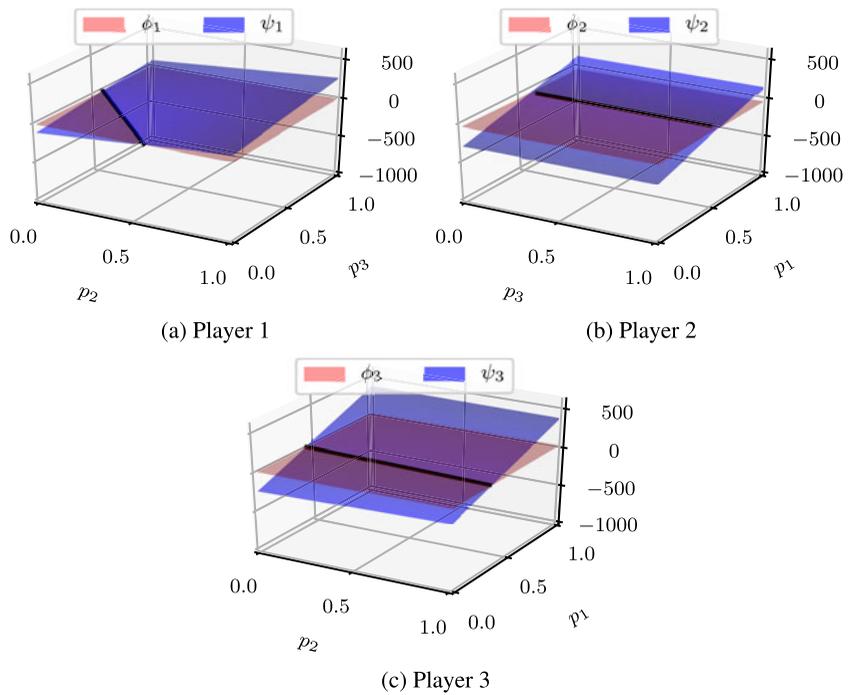


Figure 5. Value surfaces for independent strategy ϕ_i versus collective strategy ψ_i under federated scenario B. The independent strategy ϕ is dominant for player 2 while the collective strategy ψ is dominant for players 1 and 3.

to disengagement by player 2, the most likely to disengage from a collective strategy. These results echo Maier’s principles for system-of-systems architecting emphasizing stable intermediate forms and ensuring cooperating among all actors (Maier 1998). Risk dominance helps assess designs for a balance between value maximization if successful and risk minimization for coordination failures.

Although not considered in this analysis for clarity in presentation, incorporating risk attitudes would influence the absolute (but not relative) interpretation of risk dominance across scenarios A and B. For example, an exponential utility function of the form $U(c) = (1 - e^{-\alpha c})/\alpha$ for consumption level c , equivalent to expected net present value in this example, assumes constant risk aversion for $\alpha > 0$ (Arrow 1971). Applying this transformation to the payoff values in Table 7 penalizes the high-value (but uncertain) collaborative outcomes and increases risk dominance measures across both scenarios. More detailed analysis would benefit from specific knowledge about risk attitudes on behalf of individual players.

Alternative probabilistic analysis methods may evaluate expected value and variance under uncertain strategy selections. For example, defining p_i as the probability player i deviates from ϕ to ψ , the value function V_i is

$$\begin{aligned}
 V_i(p_i, p_j, p_k) = & V_i^\phi(1 - p_i) + V_i^\psi(d_i)p_i(1 - p_j)(1 - p_k) \\
 & + V_i^\psi(d_i, d_k)p_i(1 - p_j)p_k + V_i^\psi(d_i, d_j)p_i p_j(1 - p_k) \\
 & + V_i^\psi(d_i, d_j, d_k)p_i p_j p_k.
 \end{aligned}
 \tag{20}$$

Table 10. Comparative analysis of expected value and variance

	Scenario A		Scenario B	
	$E[V_i]$	$\text{Var}(V_i)$	$E[V_i]$	$\text{Var}(V_i)$
Player 1	-278	169,252	35	10,278
Player 2	22	87,432	-55	23,755
Player 3	83	79,084	64	42,679

Assuming p_i are independent and identically distributed with $p_i \sim \text{uniform}(0, 1)$, Table 10 reports expected value $E[V_i]$ and variance $\text{Var}(V_i)$ for scenarios A and B. The analysis concurs that player 1 in scenario A and player 2 in scenario B observe negative expected value and scenario B overall decreases variance for all players. However, this analysis: 1) does not provide any tradeoff between expected value and variance, 2) only provides a relative comparison between the two cases without a scalar quantitative metric of stability, and 3) does not provide further insights for the interdependency or influence between or among players.

5. Discussion and conclusion

Understanding of both the upside potential and downside risk associated with coordination failures are critical to assess sources of strategic risk in collaborative systems. As demonstrated in the application case, the concepts of strategic design games and measures of risk dominance can influence concept selection in early design activities by identifying unfavorable strategic dynamics and shifting the design focus to include economic stability in addition to economic efficiency.

The core contributions of this paper establish: 1) a method to formulate and measure strategic risk dominance for collaborative engineered systems with two or more asymmetric players and 2) a linear approximation to incentives required for problems with more than two players. This work extends prior work on multi-actor value functions (Grogan *et al.* 2018) and transfers fundamental economic theory to the domain of systems engineering to study issues of strategic risk dominance in multi-actor systems.

Building on equilibrium selection literature and WALM as a quantitative metric to assess risk dominance provide a solid foundation for strategic design games. The relative simplicity of the proposed method permits analysis of strategic dynamics during conceptual design, allowing systems engineers to identify, avoid, or rework high-value joint architectures that carry unfavorable strategic dynamics. This perspective may help avoid costly development programs with structural problems leading to schedule and cost growth and, ultimately, cancellation. However, there remain several key assumptions regarding linearized incentive structures and information availability discussed in Section 3.3 which limit more detailed analysis of strategic dynamics in engineering design.

Future work follows two directions. First, additional theoretical work to incorporate concepts from Bayesian games would help bring subjective information into context-specific design problems. Second, additional practical or applied work is required to further validate the proposed method in a realistic system context by developing a multi-actor value function, enumerating and evaluating candidate architectures to identify those with desirable strategic

dynamics, and contextualize results by selectively forming and dissolving coalitions.

Acknowledgments

Thanks to Abbas Ehsanfar for his initial efforts and general contributions to exploring this topic. This material is based on work supported, in whole or in part, by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract No. HQ0034-13-D-0004. SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Department of Defense.

Appendix A. Detailed WALM formulation

This section summarizes key results from Selten (1995) to formulate and explain the weighted average log measure (WALM) of risk dominance. It does not produce any new results but introduces some more convenient notation and insights. Please refer to the original article for axioms and proofs.

Binary games have a strategy space with two alternatives $\mathcal{S}_i = \{\phi_i, \psi_i\}$. Bipolar games are a subclass of binary games with two Nash equilibria defined by shared strategies among all players $\phi = (\phi_1, \dots, \phi_n)$ and $\psi = (\psi_1, \dots, \psi_n)$. Note that, for example, ϕ denotes the strategy set and ϕ_i denotes the strategy selected by player i . Notation with negative subscripts on strategy sets denotes non-participation, for example, $\phi_{-1} = (\psi_1, \phi_2, \dots, \phi_n)$. Without loss of generality, this section labels strategy sets such that ψ is payoff dominant as the collective strategy.

For greater generality, WALM of risk dominance is defined in terms of a *biform* that describes the essential dynamics of a game rather than the direct payoff or utility function. The biform is given by a vector of normalized deviation losses $u = u(\phi, \psi) = [u_i(\phi, \psi)]$ and an influence matrix $A = A(\phi, \psi) = [a_{ij}(\phi, \psi)]$ capturing interdependencies between players.¹ Together, these factors attribute potential losses to global and local deviations from a baseline strategy.

The most intuitive explanation of risk dominance starts by formulating an incentive function D_i for player i to choose ϕ_i over ψ_i . Player i prefers ϕ_i for $D_i > 0$ and prefers ψ_i for $D_i < 0$. An expected value expression expanded in Eq. (21) describes the incentive function from a global perspective as a function of p , the probability that *all* other players choose ψ .

$$\begin{aligned} D_i(p) &\propto E[V_i|\phi_i] - E[V_i|\psi_i] \\ &= [V_i(\phi)(1 - p) + V_i(\psi_{-i})p] - [V_i(\phi_{-i})(1 - p) + V_i(\psi)p] \\ &= V_i(\phi) - V_i(\phi_{-i}) - [V_i(\phi) - V_i(\phi_{-i}) + V_i(\psi) - V_i(\psi_{-i})] p. \quad (A1) \end{aligned}$$

Equation (22) defines deviation loss L_i as a function of a strategy set $\xi \in \{\phi, \psi\}$ up to a constant of proportionality.

$$L_i(\xi) \propto V_i(\xi) - V_i(\xi_{-i}). \quad (A2)$$

¹ Functional arguments (ϕ, ψ) denoting equilibrium strategy labels will occasionally be omitted in this section for conciseness in presentation.

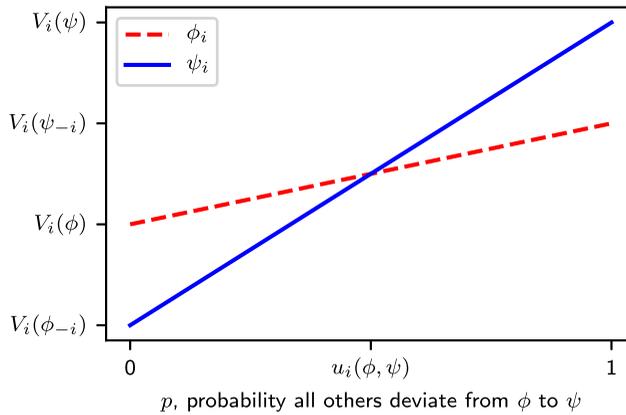


Figure 6. The normalized deviation loss in bipolar games marks the intersection between strategy-specific value functions shown here for player i as a function of p , the probability that all others deviate from ϕ to ψ .

The deviation loss captures sensitivity to deviating away from a stable strategy set through one’s own actions; however, for the purpose of this formulation consider it an algebraic expression only. The normalized deviation loss u_i in Eq. (23) transforms L_i to a unit scale.

$$u_i(\phi, \psi) = \frac{L_i(\phi)}{L_i(\phi) + L_i(\psi)}. \tag{A 3}$$

Returning to the global incentive function, Eq. (24) normalizes both sides and substitutes expressions for L_i and u_i to achieve a simplified incentive function.

$$D_i(p) = \frac{L_i(\phi) - [L_i(\phi) + L_i(\psi)] p}{L_i(\phi) + L_i(\psi)} = u_i - p. \tag{A 4}$$

In other words, player i prefers ϕ_i for $D_i > 0 \iff p < u_i$ and prefers ψ_i for $D_i < 0 \iff p > u_i$. The normalized deviation loss u_i (also referred to as the diagonal probability π_i in literature) marks the intersection between the lines in Figure 6 where $D_i(u_i) = 0$ (i.e. player i is indifferent about which strategy to select).

A more detailed incentive function can be written from a local perspective specific to each players’ strategy to capture interaction effects and interdependencies. An expected value expression expanded in Eq. (25) for a game with $n = 3$ players describes the incentive function as a function of p_j and p_k , the probability that players i and j choose ψ_i and ψ_k , respectively.

$$\begin{aligned} D_i(p_j, p_k) &\propto E[V_i|\phi_i] - E[V_i|\psi_i] \\ &= \begin{bmatrix} V_i(\phi)(1 - p_j)(1 - p_k) \\ +V_i(\phi_{-j})p_j(1 - p_k) \\ +V_i(\phi_{-k})(1 - p_j)p_k \\ +V_i(\psi_{-i})p_jp_k \end{bmatrix} - \begin{bmatrix} V_i(\phi_{-i})(1 - p_j)(1 - p_k) \\ +V_i(\phi_{-ij})p_j(1 - p_k) \\ +V_i(\phi_{-ik})(1 - p_j)p_k \\ +V_i(\psi)p_jp_k \end{bmatrix} \\ &= [V_i(\phi) - V_i(\phi_{-i})] - \left[\begin{matrix} V_i(\phi) - V_i(\phi_{-i}) \\ +V_i(\phi_{-ij}) - V_i(\phi_{-j}) \end{matrix} \right] p_j \end{aligned}$$

$$- \begin{bmatrix} V_i(\phi) - V_i(\phi_{-i}) \\ +V_i(\phi_{-ik}) - V_i(\phi_{-k}) \end{bmatrix} p_k - \begin{bmatrix} V_i(\psi) - V_i(\psi_{-i}) \\ -V_i(\phi) + V_i(\phi_{-i}) \\ +V_i(\phi_{-j}) - V_i(\phi_{-ij}) \\ +V_i(\phi_{-k}) - V_i(\phi_{-ik}) \end{bmatrix} p_j p_k. \quad (\text{A } 5)$$

Equation (26) defines pairwise deviation loss L_{ij} as a function of a strategy set $\xi \in \{\phi, \psi\}$ up to a constant of proportionality.

$$L_{ij}(\xi) \propto V_i(\xi_{-j}) - V_i(\xi_{-ij}). \quad (\text{A } 6)$$

Similar to deviation loss, pairwise deviation loss captures sensitivity to deviating away from a stable strategy set through pairwise actions but for the purpose of this formulation consider it an algebraic expression only. Influence elements a_{ij} in Eq. (27) normalize pairwise deviation losses L_{ij} to a common scale with u_i .

$$a_{ij}(\phi, \psi) = \frac{L_i(\phi) - L_{ij}(\phi)}{L_i(\phi) + L_i(\psi)}. \quad (\text{A } 7)$$

Returning to the local incentive function, normalizing both sides yields the simplified form in Eq. (28).

$$\begin{aligned} D_i(p_j, p_k) &= \frac{1}{L_i(\phi) + L_i(\psi)} \\ &\times \begin{bmatrix} L_i(\phi) - (L_i(\phi) - L_{ij}(\phi)) p_j - (L_i(\phi) - L_{ik}(\phi)) p_k \\ - (L_i(\psi) - L_i(\phi) + L_{ij}(\phi) + L_{ik}(\phi)) p_j p_k \end{bmatrix} \\ &= u_i - a_{ij} p_j - a_{ik} p_k - (1 - a_{ij} - a_{ik}) p_j p_k. \end{aligned} \quad (\text{A } 8)$$

Under the assumption of linear incentives, $1 - a_{ij} - a_{ik} = 0$ such that $D = u - Ap$ which is the general result applicable to games with any number of players.

Finally, influence weights measure the overall importance of one player on others' stability. Weights $w = [w_i(A)]$ are defined implicitly by properties in Eq. (29) based on the influence matrix $A = [a_{ij}]$.

$$w(A) = A^T w(A), \quad \sum_{i=1}^n w_i(A) = 1. \quad (\text{A } 9)$$

Weights are interpreted as the eigenvector (rescaled to unit norm) of A^T corresponding to the unit eigenvalue, guaranteed to exist by the assumption of linear incentives which forces the row sum of A to unity for all rows. Note that weights are equivalent to the limiting stochastic distribution for the Markov chain with state transition probabilities a_{ij} .

Appendix B. Approximation to linear incentives

Selten's work focuses on games with linear incentives which allow pairwise interactions to be quantified without third party effects (e.g. the effect of player j on player i is not a function of player k). This simplifying assumption, similar to a first order approximation, greatly reduces complexity for a narrow class of problems but cannot directly represent increasing or decreasing returns to scale (i.e. network effects) common in engineering applications. Furthermore, linear incentives is a critical assumption to find weighting factors which require a unit

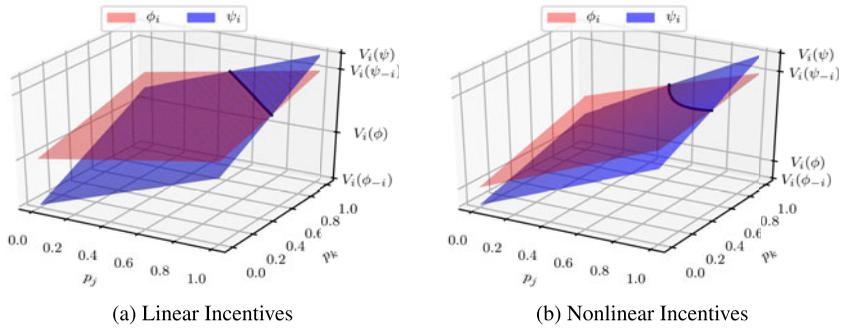


Figure 7. Value surfaces for player i as a function of player j 's and k 's probability of choosing ψ_j and ψ_k (p_j and p_k). Linear incentives produce planar value surfaces while nonlinear incentives produce non-planar (curved) surfaces representing third party effects.

eigenvalue of the influence matrix A . Although there may be extensions of the influence matrix A to higher dimensions (e.g. tensors), there is currently no such existing theory. Thus, this section introduces a novel linear approximation for greater applicability to design problems with nonlinear incentives.

Linear incentives can be visualized as planar value surfaces in Figure 7 for a game with $n = 3$ players as a function of p_j and p_k , the probability players j and k choose strategy ψ over ϕ , respectively. The incentive function $D_i(p_j, p_k)$ is the difference between the two planes. The intersection between the two planes (black line) traces the indifference curve where player i does not prefer either strategy, similar to u_i for $n = 2$ players. Games with nonlinear incentives visualized in Figure 7 for an exaggerated case include interaction terms with third parties and yield non-planar value surfaces and nonlinear indifference curves.

Consider the simplest possible game with nonlinear incentives with $n = 3$ players and incentive function in Eq. (28). Linear incentives require $a_{ij} + a_{ik} = 1$ to eliminate the interaction term between p_j and p_k . A linearized incentive function in Eq. (1) proposes modified a_{ij} terms such that $\sum_{j=1}^n \bar{a}_{ij} = 1 \forall i$ to satisfy the linear incentives condition.

$$D_i(p_j, p_k) \approx u_i - \bar{a}_{ij} p_j - \bar{a}_{ik} p_k. \tag{B 1}$$

Preserving influence elements as a coefficient for the effect of player j 's probability of choosing ψ_j on player i 's incentive to choose ϕ_i (specifically, $-\partial D_i / \partial p_j$), Eq. (2) defines the linearized influence element \bar{a}_{ji} as the expected value of the partial derivative of the incentive function with respect to p_j .

$$\begin{aligned} \bar{a}_{ij} &\equiv E \left[-\frac{\partial D_i}{\partial p_j} \right] = E [a_{ij} + (1 - a_{ij} - a_{ik}) p_k] \\ &= a_{ij} + \int_0^1 (1 - a_{ij} - a_{ik}) p_k dp_k \\ &= \frac{1}{2} (1 + a_{ij} - a_{ik}). \end{aligned} \tag{B 2}$$

For more general games with $n > 3$ players, linearizing incentive functions using this approximation becomes a combinatorial problem based on \mathcal{K}_{ij} , the

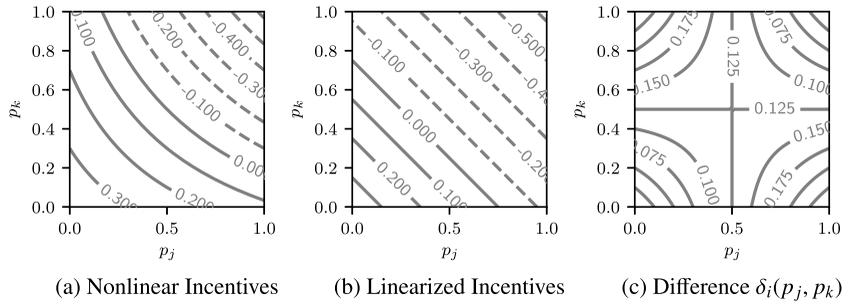


Figure 8. Incentive function $D_i(p_j, p_k)$ for player i as a function of p_j and p_k for (a) nonlinear and (b) linearized cases with (c) difference $\delta_i(p_j, p_k)$ and mean error $\varepsilon_i = 0.125$.

power set \mathcal{P} of third parties (i.e. set of all subsets of players except i and j and including the empty set) in Eq. (3) with cardinality $|\mathcal{K}_{ij}|$ in Eq. (4) given by the binomial theorem.

$$\mathcal{K}_{ij} = \mathcal{P}(\{1, \dots, n\} \setminus \{i, j\}) \tag{B 3}$$

$$|\mathcal{K}_{ij}| = \sum_{k=0}^{n-2} \binom{n-2}{k} = 2^{n-2}. \tag{B 4}$$

Revised notation in Eq. (5) defines combinatorial deviation losses between player i and a set of players \mathbf{k} as a function of a strategy set $\xi \in \{\phi, \psi\}$ up to a constant of proportionality.

$$L_{i\mathbf{k}}(\xi) \propto V_i(\xi_{-\mathbf{k}}) - V_i(\xi_{-i\mathbf{k}}). \tag{B 5}$$

Note that this expression simplifies to previously established forms of $L_{i\{\}}(\xi) = L_i(\xi)$ in Eq. (22) for $\mathbf{k} = \{\}$ and $L_{i\{j\}}(\xi) = L_{ij}(\xi)$ in Eq. (26) for $\mathbf{k} = \{j\}$. Using this notation, Eq. (6) states a conjecture for linearized influence elements.

$$\bar{a}_{ij}(\phi, \psi) = \frac{1}{|\mathcal{K}_{ij}|} \sum_{\mathbf{k} \in \mathcal{K}_{ij}} \frac{L_{i\mathbf{k}}(\phi) + L_{i\mathbf{k}}(\psi)}{L_i(\phi) + L_i(\psi)}. \tag{B 6}$$

While proof of this conjecture is not available, the result above has been manually verified for $n = 3$ (see Eq. (2) recognizing that $L_{ij}(\phi) = -L_{ik}(\psi)$) and $n = 4$ cases.

Linearizing influence elements introduces errors into the risk dominance analysis. Error manifests as differences between the incentive function $D_i(p_j, p_k)$ and its linearized form in Eq. (7) for games with $n = 3$ players.

$$\begin{aligned} \delta_i(p_j, p_k) &= D_i(p_j, p_k) - (u_i - \bar{a}_{ij}p_j - \bar{a}_{ik}p_k) \\ &= (\bar{a}_{ij} - a_{ij})p_j + (\bar{a}_{ik} - a_{ik})p_k - (1 - a_{ij} - a_{ik})p_jp_k \\ &= (1 - a_{ij} - a_{ik}) \left(\frac{p_j}{2} + \frac{p_k}{2} + p_jp_k \right). \end{aligned} \tag{B 7}$$

For example, Figure 8 visualizes contours of player i 's incentive function for a notional symmetric $n = 3$ game with value function

$$V_i(\xi) = \begin{cases} 0 & \text{if } \xi = \phi_{-i} \\ 2 & \text{if } \xi \in \{\psi_{-k}, \psi_{-j}\} \\ 3 & \text{if } \xi \in \{\phi, \phi_{-j}, \phi_{-k}, \psi_{-i}\} \\ 8 & \text{if } \xi = \psi \end{cases}, \tag{B 8}$$

influence elements $a_{ij} = 0.25$, and linearized influence elements $\bar{a}_{ij} = 0.5$ for (a) initially (highly) nonlinear incentives, (b) linearized incentives following the recommended method, and (c) the subsequent absolute difference δ_i .

Equation (9) derives a simple error metric ε_i for cases with $n = 3$ players that measures the average absolute difference in incentive value.

$$\begin{aligned} \varepsilon_i &= \int_0^1 \int_0^1 |\delta_i(p_j, p_k)| dp_j dp_k \\ &= |1 - a_{ij} - a_{ik}| \int_0^1 \int_0^1 \left| \frac{p_j}{2} + \frac{p_k}{2} + p_j p_k \right| dp_j dp_k \\ &= \frac{|1 - a_{ij} - a_{ik}|}{4}. \end{aligned} \tag{B 9}$$

Note that $D_i(p_j, p_k) \in [u_i - 1, u_i]$ so ε_i can be roughly interpreted as percent error. For the example in Figure 8, $\varepsilon_i = 0.125$ which is a relatively high value indicating potential errors in interpreting results, especially in regions with high estimates of one partner’s probability of collaboration but low estimates for the other.

Appendix C. Application case data

Data for the application case was generated using the publicly available distribution of Orbital Federates Simulation – Python (OFSPY) (Grogan 2019). This software simulation computes cash flows obtained from an initial space systems design in a version of the multi-player game *Orbital Federates*. The software program contains a command line interface (CLI) to run specific design scenarios. Automated operational policies based on mixed integer linear programs determine how to use available space systems to observe, store, transmit, and downlink data to complete contracts and earn revenue each turn.

The spatial context is reduced to two dimensions with six sectors (1–6) and layers representing the surface (SUR), low Earth orbit (LEO), and medium Earth orbit (MEO) shown in Figure 9. Satellites move clockwise between orbital sectors each turn while ground stations remain fixed at the surface. Space-to-ground links (SGLs) require a satellite to be in the same sector as a ground station for data transfer. Inter-satellite links (ISLs) require satellites to be in adjacent sectors for data transfer. Proprietary links only permit data transfer within a player’s assets while open links permit data transfer between players as paid services.

Designs evaluated under the independent strategy follow the CLI template: `ofs.py -d 24 -p 3 -i 0 -s <SEED> -o d6,a,1 -f n <DESIGN>` where `-d 24` indicates a game with 24 turns, `-p 3` indicates three players, `-i 0` indicates no initial cash constraints, `-s <SEED>` indicates the random number generator seed (integer), `-o d6,a,1` indicates to use a dynamic operations policy with a six turn horizon using an automatically computed opportunity cost

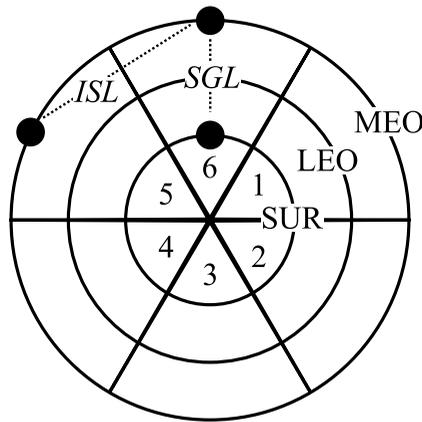


Figure 9. The Orbital Federates context includes six sectors with surface (SUR), low Earth orbit (LEO) and medium Earth orbit (MEO) layers. Satellite and ground station elements transfer data using space-to-ground (SGL) and inter-satellite (ISL) links.

for storage and a nominal penalty of 1 for ISLs, $-f_n$ indicates no federation operations policy, and $\langle \text{DESIGN} \rangle$ is the design specification.

The baseline scenario considers the design specification:

- (i) 2.SmallSat@MEO6, SAR, pSGL 2.GroundSta@SUR1, pSGL
- (ii) 3.SmallSat@MEO4, VIS, pSGL 3.GroundSta@SUR5, pSGL

Player 1 has no elements. Player 2 has a small satellite initially in MEO sector 6 with a synthetic aperture radar (SAR) and proprietary SGL (pSGL) and a ground station at surface sector 1 with a pSGL. Player 3 has a small satellite initially in MEO sector 4 with a visual light sensor (VIS) and pSGL and a ground station at surface sector 5 with a pSGL.

Designs evaluated under the collective strategy follow the CLI template: `ofs.py -d 24 -p 3 -i 0 -s <SEED> -o d6,a,1 -f x100,100,6,a,1 <DESIGN>` where `-f x100,100,6,a,1` indicates to use an opportunistic federation operations policy with fixed prices of 100 for SGL and ISL, a six turn horizon, an automatically computed opportunity cost for storage, and a nominal penalty of 1 for ISLs.

Scenario A considers the design specification:

- (i) 1.SmallSat@MEO5, oISL, oSGL 1.GroundSta@SUR3, oSGL
- (ii) 2.MediumSat@MEO6, SAR, oISL, pSGL, oSGL 2.GroundSta@SUR1, pSGL
- (iii) 3.MediumSat@MEO4, VIS, oISL, pSGL, oSGL 3.GroundSta@SUR5, pSGL

Player 1 has a small satellite in MEO sector 5 with an open ISL (oISL) and an open SGL (oSGL) and a ground station at surface sector 3 with oSGL. Player 2 has a medium satellite in MEO sector 6 with SAR, oISL, pSGL, and oSGL and a ground station at surface sector 1 with pSGL. Player 3 has a medium satellite in MEO sector 4 with VIS, oISL, pSGL, and OSGL and a ground station at surface sector 5 with pSGL.

Scenario B considers the design specification:

- (i) 1.SmallSat@MEO5, SAR, oSGL 1.GroundSta@SUR3, oSGL
- (ii) 2.MediumSat@MEO6, SAR, pSGL, oSGL 2.GroundSta@SUR1, pSGL

(iii) 3.MediumSat@ME04,VIS,pSGL,oSGL 3.GroundSta@SUR5,pSGL

Player 1 has a small satellite in MEO sector 5 with SAR and oSGL and a ground station at surface sector 3 with oSGL. Players 2 and 3 are identical to Scenario A except removing the oISL modules.

Note that scenario A relies on close proximity between players to enable ISLs. The above design strings were modified in cases with only partial participation in the federation: ME05 is replaced by ME03 if only players 1 and 3 join a federation and ME04 is replaced by ME05 if only players 2 and 3 join a federation. Outputs reported in Table 7 compute net present values using a discount rate of 2% (per turn) averaged over the first 1000 seeds (from 0 to 999).

References

- Abbas, A. E. & Cadenbach, A. H. 2018 On the use of utility theory in engineering design. *IEEE Systems Journal* **12** (2), 1129–1138.
- Arrow, K. J. 1963 *Social Choice and Individual Values*, 2nd edn. Yale University Press.
- Arrow, K. J. 1971 Aspects of the theory of risk bearing. In *Essays in the Theory of Risk Bearing*, pp. 90–109. Markham Publishing Company.
- Bhatia, G. V., Kannan, H. & Bloebaum, C. L. 2016 A game theory approach to bargaining over attributes of complex systems in the context of value-driven design. In *54th AIAA Aerospace Sciences Meeting, San Diego, CA, USA*, American Institute of Aeronautics and Astronautics.
- Briceño, S. I. 2008. A game-based decision support methodology for competitive systems design. PhD thesis, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA.
- Brown, O. & Eremenko, P. 2006 The value proposition for fractionated space architectures. In *AIAA Space 2006, San Jose, CA, USA*.
- Collopy, P. D. & Hollingsworth, P. M. 2011 Value-driven design. *Journal of Aircraft* **48** (3), 749–759.
- Daniels, M. P. & Paté-Cornell, M.-E. 2017 Risk-based comparison of consolidated and distributed satellite systems. *IEEE Transactions on Engineering Management* **64** (3), 301–315.
- Franssen, M. & Bucciarelli, L. L. 2004 On rationality in engineering design. *Journal of Mechanical Design* **126** (6), 945–949.
- Grogan, P. T. 2019. Orbital federates simulation – Python <https://github.com/ptgrogan/ofspy>.
- Grogan, P. T. & de Weck, O. L. 2015 Interactive simulation games to assess federated satellite system concepts. In *2015 IEEE Aerospace Conference, Big Sky, MT, USA*.
- Grogan, P. T., Ho, K., Golkar, A. & de Weck, O. L. 2018 Multi-actor value modeling for federated systems. *IEEE Systems Journal* **12** (2), 1193–1202.
- Gurnani, A. & Lewis, K. 2008 Collaborative, decentralized engineering design at the edge of rationality. *Journal of Mechanical Design* **130** (12), 121101.
- Harsanyi, J. C. 1967 Games with incomplete information played by ‘Bayesian’ players, I–III, Part I. *Management Science* **14** (3), 159–182.
- Harsanyi, J. C. & Selten, R. 1988 *A General Theory of Equilibrium Selection in Games*. MIT Press.
- Hazelrigg, G. A. 1997 On irrationality in engineering design. *Journal of Mechanical Design* **119** (2), 194–196.

- Hazelrigg, G. A.** 1998 A framework for decision-based engineering design. *Journal of Mechanical Design* **120** (4), 653–658.
- Herrmann, J. W.** 2010 Progressive design processes and bounded rational designers. *Journal of Mechanical Design* **132** (8), 081005.
- Kang, N., Ren, Y., Feinberg, F. M. & Papalambros, P. Y.** 2016 Public investment and electric vehicle design: A model-based market analysis framework with application to a USA–China comparison study. *Design Science* **2** (e6), 1–42.
- Kaplan, S. & Garrick, B. J.** 1981 On the quantitative definition of risk. *Risk Analysis* **1** (1), 11–27.
- Lewis, K. & Mistree, F.** 1997 Modeling interactions in multidisciplinary design: A game theoretic approach. *AIAA J.* **35** (8), 1387–1392.
- Lough, K. G., Stone, R. & Tumer, I. Y.** 2009 The risk in early design method. *Journal of Engineering Design* **20** (2), 155–173.
- Maier, M. W.** 1998 Architecting principles for systems-of-systems. *Systems Engineering* **1** (4), 267–284.
- National Research Council,** 2011 *Assessment of Impediments to Interagency Collaboration on Space and Earth Science Missions*. National Academies Press.
- Oates, W. E.** 2008 On the theory and practice of fiscal decentralization. In *Institutional Foundations of Public Finance: Economic and Legal Perspectives, chapter 5* (ed. **A. J. Auerbach & D. N. Shaviro**), pp. 165–189. Harvard University Press.
- O'Neill, M. G., Yue, H., Nag, S., Grogan, P. & de Weck, O. L.** 2010 Comparing and optimizing the DARPA System F6 program value-centric design methodologies. In *AIAA Space 2010 Conference and Exposition, Anaheim, CA, USA*.
- Papageorgiou, E., Eres, M. H. & Scanlan, J.** 2016 Value modelling for multi-stakeholder and multi-objective optimisation in engineering design. *Journal of Engineering Design* **27** (10), 697–724.
- Ross, A. M., Hastings, D. E., Warmkessel, J. M. & Diller, N. P.** 2004 Multi-attribute tradespace exploration as front end for effective space system design. *Journal of Spacecraft and Rockets* **41** (1), 20–28.
- Scott, M. J. & Antonsson, E. K.** 1999 Arrow's theorem and engineering design decision making. *Research in Engineering Design* **11** (4), 218–228.
- Selten, R.** 1995 An axiomatic theory of a risk dominance measure for bipolar games with linear incentives. *Games and Economic Behavior* **8** (1), 213–263.
- Simon, H.** 1959 Theories of decision-making in economics and behavioral science. *American Economic Review* **49** (3), 253–283.
- Skryms, B.** 2004 *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.
- Thurston, D. L.** 2001 Real and misconceived limitations to decision based design with utility analysis. *Journal of Mechanical Design* **123** (2), 176–182.
- Tversky, A. & Kahneman, D.** 1992 Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5** (4), 297–323.
- United States of America,** 2017. Weather research and forecasting innovation act of 2017. Public Law 115-25, 119 Stat. 91.
- Velasquez, M. & Hester, P. T.** 2013 An analysis of multi-criteria decision making methods. *International Journal of Operational Research* **10** (2), 56–66.
- Vincent, T. L.** 1983 Game theory as a design tool. *Journal of Mechanisms, Transmissions, and Automation in Design* **105** (2), 165–170.
- Walton, M. A. & Hastings, D. E.** 2004 Applications of uncertainty analysis to architecture selection of satellite systems. *Journal of Spacecraft and Rockets* **41** (1), 75–84.

- de Weck, O., de Neufville, R. & Chaize, M.** 2004 Staged deployment of communications satellite constellations in low earth orbit. *Journal of Aerospace Computing, Information, and Communications* **1** (3), 119–136.
- Wernz, C. & Deshmukh, A.** 2010 Multiscale decision-making: Bridging organizational scales in systems with distributed decision-makers. *European Journal of Operational Research* **202** (3), 828–840.