Proceedings of the 11th North American International Conference on Industrial Engineering and Operations Management, Milwaukee, Wisconsin, USA, June 9-11, 2026

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/NA11.20260154
Published: June 9, 2026

Active Learning for Scalable Set-Based Design in High-Dimensional End-State Spaces

Roohollah Jahanmahin, Sara Masoud, Murat Yildirim, Evrim Dalkiran, Ratna Babu Chinnam and Alper Murat
Department of Industrial & Systems Engineering
Wayne State University
Detroit, MI, USA

Suzan Arslanturk
Department of Computer Science
Wayne State University
Detroit, MI, USA

Maryam Eghbalizarch
Assistant Professor
Department of Industrial & Systems Engineering
Kennesaw State University
Kennesaw, GA, USA

Neda Masoud
Associate Professor
Department of Civil and Environmental Engineering
University of Michigan
Ann Arbor, MI, USA

Abstract

Set-based design (SBD) enables systematic exploration of complex engineering design spaces by preserving sets of feasible solutions rather than converging prematurely to a single design. However, identifying and refining end-state sets in high-dimensional spaces remains computationally prohibitive when exhaustive evaluation or dense sampling is required. This paper proposes an active learning–based framework to support scalable end-state exploration within the set-based design paradigm. Instead of directly optimizing end-states, the proposed approach iteratively learns a reliable representation of the feasible design space through selective sampling, surrogate modeling, and probabilistic feasibility assessment. A surrogate regression model is trained and validated using cross-validated accuracy thresholds to ensure predictive reliability before downstream use. Kernel density estimation is then employed to characterize the distribution of feasible end-states and to define acceptance regions for newly sampled candidates. New end-states are adaptively accepted or rejected based on their likelihood under the learned density model, allowing computational effort to focus on informative and feasible regions of the design space. Accepted samples are used to update geometric and probabilistic representations of the feasible end-state set, enabling progressive refinement consistent with set-based design principles. By replacing exhaustive end-state evaluation with selective learning, the proposed framework significantly reduces computational burden while preserving the exploratory nature of set-based design, making it well suited for high-dimensional engineering design problems.

Keywords
Set-based design; Active learning; High-dimensional design spaces; Surrogate modeling.

1. Introduction
Set-based design (SBD) has emerged as a fundamentally different paradigm from traditional point-based design (PBD) for managing uncertainty and complexity in engineering design. In PBD (as displayed in Figure 1), designers typically select a small number of candidate solutions early in the process and iteratively refine a single preferred design through analysis and redesign cycles. This approach often leads to premature commitment, making late-stage changes costly and increasing the risk of rework when new constraints or interactions are discovered. In contrast, SBD emphasizes the deliberate exploration and maintenance of sets of feasible solutions, with gradual narrowing based on accumulating knowledge rather than early optimization as illustrated in Figure 1 (Sobek et al. 1999).

Sobek et al. (1999) articulated the foundational principles of set-based concurrent engineering through empirical studies of Toyota's product development process, showing how teams reason over ranges of feasible solutions and eliminate alternatives only when they are proven infeasible. Ward et al. (1995) further demonstrated that this strategy enables parallel development across subsystems, allowing integration to occur through intersection of feasible sets rather than reconciliation of conflicting point designs. As a result, SBD supports delayed decision-making while preserving flexibility during early and middle stages of design. Subsequent research has formalized SBD as a general design methodology applicable beyond automotive systems. Toche et al. (2020) provided a comprehensive review of set-based design methods, highlighting how SBD differs from PBD in its treatment of uncertainty, trade-offs, and knowledge reuse. Rather than converging toward a single optimum, SBD treats design as a process of feasible region discovery and refinement, which aligns naturally with complex, multi-objective, and multidisciplinary engineering problems. The benefits of this approach include reduced risk of premature convergence, improved robustness to requirement changes, and better support for cross-functional coordination (Toche et al. 2020).

A direct consequence of this philosophy is that SBD inherently produces multiple viable end-states rather than a single final solution. Because alternative subsystem options are preserved in parallel and eliminated only when proven infeasible, the design process progressively narrows a feasible region instead of collapsing onto a single point. The intersection of subsystem-level feasible sets can therefore yield several structurally distinct but performance-comparable system configurations. These end-states may occupy different regions of the objective space, reflect alternative architectural regimes, or represent different trade-off preferences among competing objectives. Consequently, the outcome of SBD is not a unique optimal design, but a refined subset of feasible and high-performing configurations from which final selection can be made based on evolving stakeholder priorities, risk considerations, or contextual constraints.

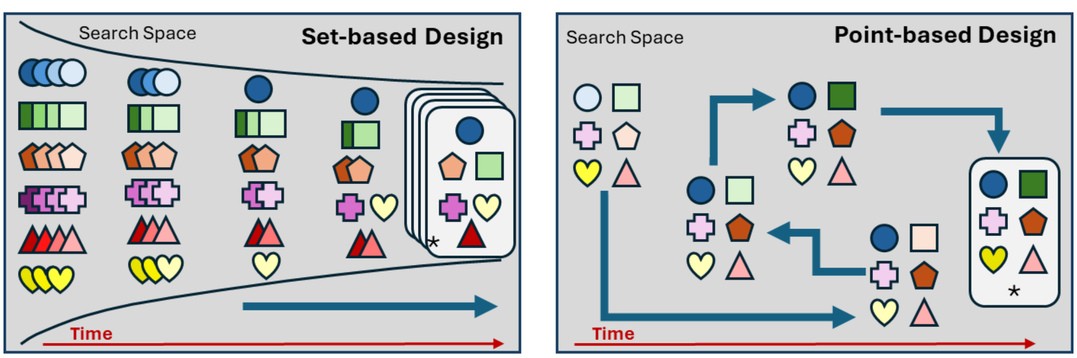

Figure 1. Conceptual comparison of SBD and PBD over time. In SBD (left), multiple feasible alternatives are explored in parallel, and the search space is progressively narrowed as infeasible options are eliminated, with commitment deferred until late in the process. In PBD (right), a single design is selected early and iteratively modified through backtracking as new information emerges. Distinct shapes represent different subsystems, while shades denote alternative design options within each subsystem.

1.1. Objectives

Despite these advantages, the literature consistently notes that SBD introduces significant computational and organizational costs. Maintaining and evaluating large sets of alternatives can lead to a combinatorial explosion as the number of subsystems and design variables increases. Ghosh et al. (2014) showed that practical implementations of SBD require careful management of design knowledge, feasibility assessment, and elimination logic to avoid overwhelming designers and analysts. From a computational perspective, repeated simulation, analysis, or optimization across high-dimensional design spaces can become prohibitively expensive, especially when end-state feasibility must be evaluated exhaustively. Several researchers have attempted to bridge SBD with optimization-based methods to address scalability challenges. Rapp (2018) discussed the tension between optimization-driven trade-space exploration and set-based reasoning, noting that while optimization can accelerate convergence, it often undermines the exploratory intent of SBD by collapsing the design space too aggressively. This challenge is exacerbated in high-dimensional problems, where exhaustive end-state optimization becomes intractable (Singer and Doerry 2009).

These limitations motivate the need for scalable, data-driven approaches that preserve the philosophy of set-based design while reducing computational burden. Surrogate modeling and adaptive sampling have been widely used in design space exploration to reduce the cost of expensive evaluations (Gorissen et al. 2010). More recently, active learning has been shown to improve surrogate fidelity by selectively sampling informative regions of the design space rather than relying on uniform or exhaustive sampling (Guo et al. 2024). However, existing work largely focuses on optimization or prediction accuracy rather than explicitly supporting set-based feasibility refinement. This work directly addresses that gap by reframing end-state analysis in SBD as a learning problem rather than a pure optimization task. By combining active learning, surrogate modeling, and probabilistic feasibility estimation, the proposed framework enables efficient refinement of feasible end-state sets without sacrificing the core principles of set-based design. In doing so, it offers a scalable pathway for applying SBD to high-dimensional engineering systems where traditional implementations are computationally prohibitive. Accordingly, this study investigates whether active learning integrated with surrogate modeling can enable scalable refinement of feasible end-state sets in high-dimensional set-based design problems while preserving predictive reliability.

2. Literature Review

End-state optimization plays a central role in SBD, as the identification and refinement of feasible end-states ultimately determines how design sets are narrowed and decisions are justified. However, the methods used to perform end-state optimization vary significantly depending on problem structure, dimensionality, and the number of objectives involved. To clearly position the contributions of this work, the literature is reviewed in three categories: end-state optimization in simple or tractable cases, multi-objective end-state optimization, and many-objective end-state optimization. This organization highlights how methodological complexity and computational burden increase as design problems scale, and it clarifies why existing approaches struggle to support SBD in high-dimensional settings. The review also establishes the gap that motivates the proposed active learning–based framework.

2.1. End-State Evaluation in Simple and Tractable Cases

In simple or low-dimensional design problems, end-state evaluation is commonly addressed using exact or near-exact optimization techniques. When decision variables, constraints, and objectives can be expressed in linear or mixed-integer linear form, mixed-integer linear programming (MILP) is frequently used to identify globally optimal end-states. These approaches are attractive because they provide optimality guarantees and clear feasibility boundaries, making them well suited for problems with limited dimensionality and structured constraints. Such formulations have been successfully applied in engineering design problems where the number of candidates end-states remains manageable, and evaluations are computationally inexpensive (Singer and Doerry 2009).

Dynamic programming (DP) and backward-search methods represent another class of end-state optimization techniques in simple settings. These approaches explicitly reason from terminal states backward through decision stages, making them conceptually aligned with end-state reasoning. However, DP methods are well known to suffer from the curse of dimensionality, as the state space grows exponentially with the number of variables, limiting their applicability beyond small or highly structured problems (Singer and Doerry 2009). As a result, while exact methods are effective in tractable cases, they become impractical as set-based design problems grow in dimensionality and subsystem complexity.

2.2. Multi-Objective End-State Evaluation

When end-state evaluation involves multiple conflicting objectives, the goal shifts from evaluating a single solution to approximating a set of designs. In this regime, evolutionary multi-objective optimization algorithms (MOEAs) have become the dominant solution approach due to their flexibility and ability to handle nonlinear, nonconvex, and discrete design spaces. Deb et al. (2002) introduced the Non-dominated Sorting Genetic Algorithm II (NSGA-II), which remains one of the most widely used algorithms for multi-objective optimization because of its balance between convergence and diversity. Multi-objective end-state optimization has been extensively applied in engineering design to generate trade-off surfaces that support decision-making across performance, cost, and risk dimensions. However, the literature also consistently notes that MOEAs often require a large number of end-states evaluations to obtain a well-distributed approximation of the Pareto front, particularly when evaluations involve expensive simulations or complex constraints (Deb et al. 2002). To address this challenge, surrogate-assisted MOEAs have been proposed to reduce evaluation cost, but these approaches are still largely optimization-centric and tend to emphasize Pareto convergence rather than explicit characterization of feasible design sets (Gorissen et al. 2010).

2.3. Many-objective End-State Evaluation

As the number of objectives increases beyond three, traditional dominance-based methods lose effectiveness because a large fraction of solutions becomes mutually non-dominated. This phenomenon has motivated the development of many-objective optimization algorithms that introduce alternative selection and diversity mechanisms to maintain search pressure. Deb and Jain (2014) proposed NSGA-III, which extends NSGA-II by incorporating reference points to guide solution diversity in high-dimensional objective spaces. Similarly, Zhang and Li (2007) introduced MOEA/D, which decomposes many-objective problems into a set of scalar subproblems solved collaboratively.

Reference-vector-based approaches, such as the Reference Vector Guided Evolutionary Algorithm (RVEA), further improve scalability by explicitly steering the search along predefined directions in objective space (Cheng et al. 2016). While these methods represent significant advances, the literature consistently reports that many-objective optimization remains computationally expensive, especially when end-state evaluations are costly and when the design space is high-dimensional. As the number of objectives and design variables increases, the number of evaluations required to stabilize a meaningful approximation of the feasible end-state set grows rapidly (Deb and Jain 2014).

2.4. Implications for Set-Based Design

Across simple, multi-objective, and many-objective settings, existing end-state evaluation approaches share a common limitation: scalability. Exact methods break down due to combinatorial growth, while evolutionary and many-objective algorithms require large numbers of evaluations to adequately explore and approximate feasible regions. For SBD, where the goal is not merely to identify optimal points but to progressively refine feasible sets, this computational burden becomes a critical bottleneck. These limitations motivate the need for alternative approaches that reduce reliance on exhaustive end-state optimization while preserving the core principles of SBD. This work addresses that gap by reframing end-state refinement as an active learning problem, enabling efficient and scalable exploration of high-dimensional design spaces.

3. Methods

The application considered in this study is the system-level configuration design of a complex engineered product, exemplified by an automotive vehicle composed of multiple interacting subsystems and nested sub-subsystems, each offering several discrete design options. The overall system design is defined by the selection of exactly one option per sub-subsystem, with subsystem-level configurations emerging implicitly from these selections. Feasibility of a complete system configuration is governed by architectural constraints, including compatibility, prerequisite, and corequisite relationships among options. This formulation captures a broad and practically relevant class of product development problems in which millions or billions of feasible configurations may exist, and design decisions must balance competing performance, physical, reliability, and risk-related considerations under uncertainty.

The system-level configuration problem is formulated as a discrete, many-objective optimization model, where each feasible end-state design corresponds to a unique combination of sub-subsystem option selections. Binary decision variables indicate whether a specific option is selected for each sub-subsystem, enforcing a one-option-per-subsystem structure. In the automotive case study considered in this work, three system-level objectives are simultaneously evaluated: performance/mission utility (PERF/MUF), a composite physical attribute metric, and a composite energy–thermal metric. These objectives are evaluated at the system level and emerge from the interactions among selected sub-subsystem options. The feasible design space is further constrained by architectural dependency relationships, including incompatibility, prerequisite, and corequisite constraints, ensuring valid and integrable system configurations. The decision variables, objectives, and constraints define a highly combinatorial and computationally challenging many-objective optimization problem. The full formulation of this optimization model is available in our prior work (Eghbali-Zarch et al., 2026); in this paper, we build on that formulation to focus on scalable end-state exploration and refinement within a set-based design context.

The proposed framework, as displayed in Figure 2, introduces an active learning–based methodology to mitigate the computational burden associated with high-dimensional end-state optimization. Rather than performing exhaustive optimization over the entire design space, the approach incrementally learns a reliable representation of feasible end-states through iterative sampling, surrogate modeling, and probabilistic filtering. The process begins with an initial random sampling of candidate end-states, ensuring broad coverage of the design space without imposing restrictive assumptions on optimality at the outset. These samples form the foundation for subsequent learning and refinement. At each iteration, a surrogate model (i.e., HistGradientBoostingRegressor) is constructed to approximate the relationship between design features and system responses (i.e., weighted sum of objectives). To enable surrogate learning and convergence monitoring, the multiple objectives are aggregated into a weighted composite score representing a fixed preference structure. Using the current sample set, its predictive reliability is evaluated through 7-fold cross-validation. The median cross-validated coefficient of determination is used as a gating metric to assess model quality. If the surrogate fails to meet a predefined threshold (e.g., R^2>90%), additional samples are actively incorporated and the model is retrained. This loop continues until the surrogate demonstrates sufficient explanatory power, ensuring that downstream decisions are informed by a statistically reliable approximation of the underlying system behavior.

The selection of the surrogate model is informed by a comparative sensitivity analysis across multiple model classes, including linear models (Ridge, Lasso), tree-based ensembles (Random Forest, Gradient Boosting), kernel-based methods (Support Vector Regression), and instance-based methods (k-NN). All models are evaluated under identical initial sampling conditions using cross-validation. The choice of 7-fold cross-validation reflects a balance between estimation stability and computational efficiency within the iterative active learning loop; preliminary sensitivity analysis over k=5,7, and 10 folds showed that lower fold counts produced higher variability in R^2 estimates, while higher fold counts increased retraining cost with only marginal gains in stability. Among the tested models, gradient boosting consistently achieved the highest predictive accuracy (median cross-validated R^2≈0.88–0.91) and demonstrated superior robustness to nonlinear interactions and heterogeneous subsystem structures. Hyperparameters of the selected model were tuned using cross-validated grid search, including learning rate, maximum tree depth, number of boosting iterations, and minimum samples per leaf. The final configuration was selected based on maximizing median cross-validated R^2 while maintaining stability across folds. In addition, the surrogate accuracy threshold used for convergence was determined through sensitivity analysis over multiple values (e.g., R^2=0.7,0.8,0.9), where higher thresholds improved predictive fidelity but required additional sampling and iterations. Based on this trade-off, a threshold of R^2=0.9 was selected to ensure high surrogate reliability while preserving substantial computational savings. This model was therefore adopted as the surrogate within the active learning loop due to its strong balance between accuracy, computational efficiency, and robustness.

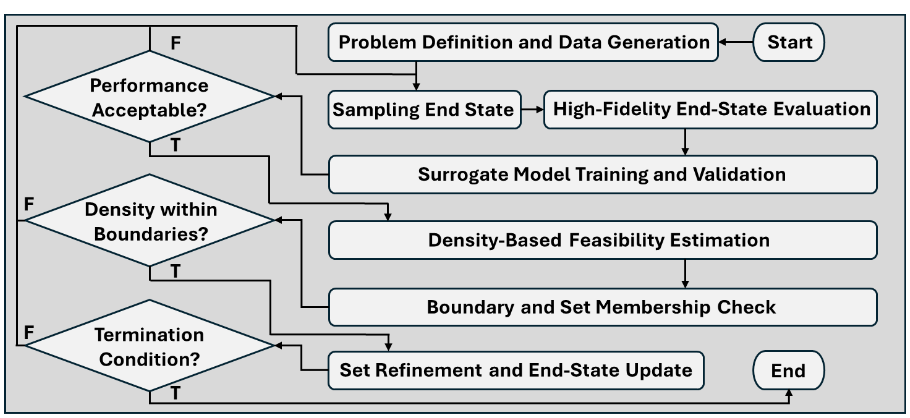

Figure 2. High-level flow diagram of the proposed active learning–based SBD framework.

Once a validated surrogate model is obtained, the framework transitions from prediction to probabilistic characterization of the feasible design space. Kernel density estimation is employed to model the distribution of feasible end-states, using a Gaussian kernel with an optimally selected bandwidth determined via grid search. This density model provides a continuous representation of the feasible region, allowing the framework to quantify the likelihood of observing new candidate end-states. Based on the learned density, acceptance intervals are defined (e.g., $\mu \pm 2\sigma$ or percentile-based thresholds) to distinguish between plausible and implausible regions of the design space.

Newly proposed end-states are evaluated against the learned density model to determine whether they fall within the acceptable feasibility boundaries. Candidate end-states with insufficient probability density are rejected, prompting further sampling and refinement, while those that lie within the acceptance region are retained for set refinement. This density-based filtering mechanism allows the algorithm to focus exploration on informative and feasible regions of the design space, avoiding unnecessary computation in low-likelihood areas.

For accepted end-states, the framework updates the estimated feasible design set by incorporating the new samples and recomputing geometric descriptors. This update reflects the progressive refinement of the design space in accordance with SBD principles, where the goal is not to converge to a single optimal solution but to characterize and shrink the feasible solution set in a controlled and interpretable manner. Probability density visualizations are generated to capture the evolution of the feasible region distribution.

The iterative process continues until the termination condition is satisfied. The algorithm terminates when both the predicted distribution of surrogate outputs has stabilized (i.e., the change in mean and standard deviation over the last five iterations is less than 0.002) and fewer than 5% of remaining unlabeled designs fall outside the 5th–95th percentile prediction interval, indicating sufficient coverage of the feasible region. Upon termination, the algorithm outputs a probabilistically validated and computationally refined subset of end-states. By integrating surrogate modeling, density estimation, and active learning within an SBD framework, the proposed methodology enables scalable exploration of high-dimensional design spaces while preserving the fundamental philosophy of set-based design.

4. Case Study

All data used in this study are synthetically generated and fully simulated. The subsystem structures, design options, objective values, and dependency relationships are generated to realistically reflect the complexity, scale, and trade-off characteristics of practical engineering systems. Objective function values are computed analytically based on option-level attributes and system-level aggregation rules embedded in the mathematical model, while feasibility is enforced through simulated architectural, compatibility, prerequisite, and corequisite constraints.

As shown in Figure 3, the vehicle architecture is modeled as a structured, hierarchical system composed of five major subsystems: the powertrain system, chassis and suspension system, electrical and electronic architecture, body and structural system, and interior. Each subsystem is further decomposed into multiple nested sub-subsystems, which serve as the fundamental decision units in the configuration design problem. For every sub-subsystem, a discrete set of mutually exclusive design alternatives is available, denoted in Figure 3 using multiplicative notation (e.g., ×2, ×4, ×6), indicating that exactly one option must be selected from the available alternatives. Across all sub-subsystems, this results in several dozen design choices, with option counts varying by subsystem and governed by architectural dependency relationships. This hierarchical scale and interdependent structure are representative of realistic engineered products and make the case study a rigorous testbed for evaluating scalable set-based design and end-state optimization methods.

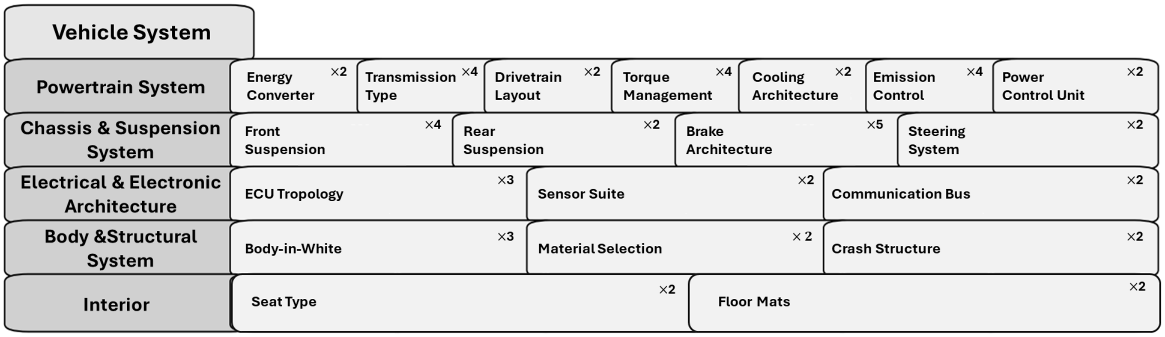

Figure 3. The automotive design space is composed of five subsystems, each decomposed into a heterogeneous number of sub-subsystems with varying numbers of discrete design options.

A complete system-level end state is defined by selecting one option for every sub-subsystem across all five subsystems. The heterogeneous option counts shown in Figure 3 highlight the combinatorial nature of the design space, as different sub-subsystems contribute unevenly to overall complexity. This hierarchical structure captures realistic architectural dependencies encountered in automotive product development while remaining fully synthetic. The resulting combinatorial design space could lead to potentially 47,185,920 (i.e., $(2 \times 4 \times 2 \times 4 \times 2 \times 4 \times 2) \times (4 \times 2 \times 5 \times 2) \times (3 \times 2 \times 2) \times (3 \times 2 \times 2) \times (2 \times 2)$) possible configurations prior to feasibility filtering. But, considering the feasibility of end-state configurations, architectural and integration dependencies such as incompatibility relation between options, prerequisite conditions, as well as corequisite relation between options reduces the total system configurations.

5. Results and Discussion

This section presents a comprehensive analysis of the feasible region, sampling progression, and surrogate-driven refinement behavior across the multi-objective design space. The results are organized to first examine the global trade-off relationships among objectives 1, 2, and 3, followed by a state-wise investigation of structural diversity and optimization behavior. Particular attention is given to identifying objective coupling patterns, clustering phenomena, and differences in frontier geometry across end-states, as these characteristics directly influence active learning dynamics and design complexity. Figure 4 presents the pairwise projections of the feasible region in objective space of the complete design set.

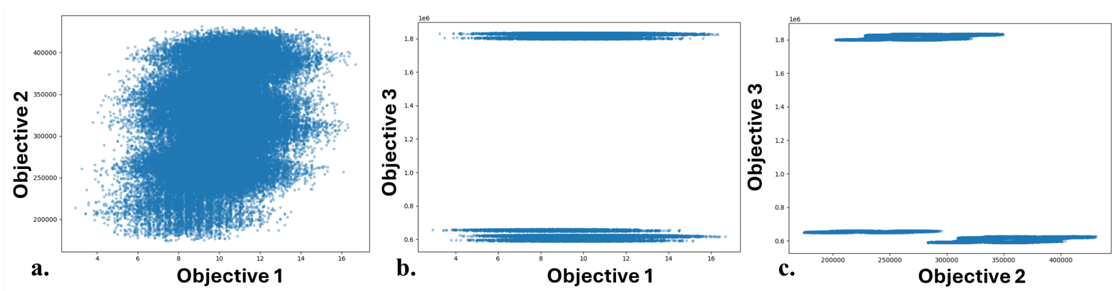

Figure 4. Pairwise projections of the global Pareto-optimal design set: (a) Objective 2 versus Objective 1, (b) Objective 3 versus Objective 1, and (c) Objective 3 versus Objective 2.

Figure 4-a illustrates the relationship between Objective 1 (PERF/MUF) and Objective 2. The distribution exhibits a broad but structured spread, indicating a continuous trade-off region rather than a sharply defined frontier. Although variability exists across the full range of Objective 1 values, the density suggests a moderate coupling between the two objectives. The absence of a narrow ridge implies that improvements in Objective 1 do not impose a strictly monotonic penalty on Objective 2 but rather allow a range of feasible trade-off combinations. This behavior is characteristic of high-dimensional combinatorial architectures in which multiple subsystem configurations yield comparable performance envelopes.

In contrast, Figure 4-b and Figure 4-c reveal a distinctly different structure involving Objective 3. The projections show two clearly separated bands along the Objective 3 axis, indicating a bimodal distribution. This clustering suggests that the Pareto surface contains at least two structurally distinct trade-off regimes. Rather than forming a single smooth manifold, Objective 3 appears to partition the design space into two dominant architectural families. To better understand the structural drivers underlying the objective-space patterns observed in Figure 4, we next examine the architectural composition of the design space at the subsystem level. Figure 5 presents the distribution of option frequencies across end-states for each major subsystem group (0.x–4.x). For every option, the boxplots represent the proportion of configurations within an end state that utilizes that option.

Several notable structural patterns are observed in Figure 5. In Subsystems 0.x and 3.x, option usage exhibits strong asymmetry, with certain options consistently appearing at significantly higher proportions than others. This indicates the presence of dominant architectural configurations within these subsystems, which likely contribute to the formation of distinct performance regimes observed earlier, particularly the bimodal behavior associated with Objective 3. In contrast, Subsystems 1.x and 2.x display comparatively balanced option distributions, suggesting these components contribute more gradually to objective variation rather than inducing sharp regime shifts. Subsystem 4.x demonstrates mixed behavior, with some options appearing highly concentrated while others remain relatively diffuse, further supporting the hypothesis that specific subsystem families are responsible for the clustered Pareto ridges observed in objective space.

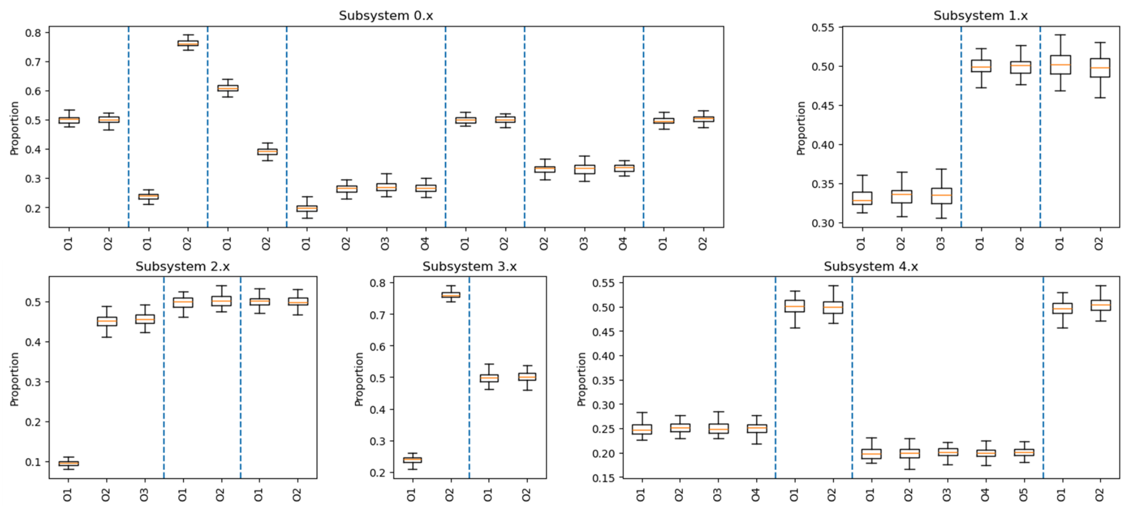

Figure 5. Distribution of subsystem option proportions across end-states for Subsystems 0.x–4.x. Each box represents the distribution of the relative frequency of a specific option within an end state. Dashed vertical lines separate sub-subsystems within each major subsystem group.

These structural differences are directly reflected in the dynamics of the surrogate modeling process. Figure 6 presents the sampling progression and convergence behavior of the active learning framework across all end-states under the 0.90 median R^2 stopping criterion. Each trajectory corresponds to an independent end state and illustrates how rapid predictive stability is achieved as labeled samples are incrementally incorporated. Although most states ultimately approach or exceed the convergence threshold, the convergence paths display substantial variability during early and intermediate iterations. This variability is consistent with the architectural heterogeneity described above: end-states dominated by smooth objective landscapes tend to stabilize quickly, whereas states associated with discrete structural regimes require more extensive exploration before achieving stable predictive performance.

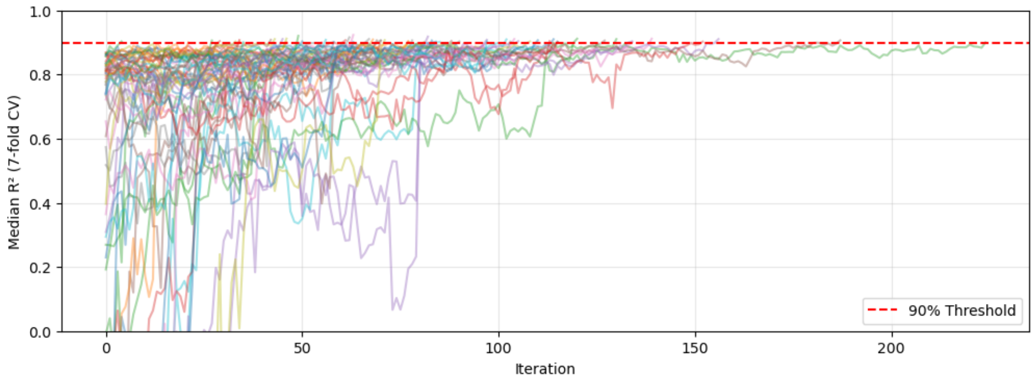

Figure 6. Active learning convergence curves across all end-states under a 90% median R^2 threshold. Each curve represents the evolution of 7-fold cross-validated median R^2 over iterations for a single end state. The dashed red line indicates the 0.90 convergence criterion.

As shown in Figure 6, the convergence trajectories across end-states exhibit substantial heterogeneity in both speed and stability. Several curves rise sharply and plateau within the first 20–30 iterations, indicating rapid stabilization and efficient surrogate learning. These fast-converging states are consistent with regions of the design space characterized by smoother objective relationships and more homogeneous subsystem option distributions, as observed in the subsystem-level analysis.

In contrast, other trajectories display pronounced oscillations and delayed stabilization, with some requiring more than one hundred iterations to approach the 0.90 R^2 threshold. These prolonged convergence patterns correspond to structurally complex end-states, particularly those associated with discrete architectural regimes and the bimodal clustering behavior previously identified in Objective 3. In such regions, the surrogate model must sequentially explore multiple structural neighborhoods before achieving stable generalization, resulting in higher variance during cross-validation and slower predictive refinement. Despite these differences in trajectory shape and convergence time, nearly all end-states ultimately achieve high predictive fidelity, demonstrating that the active learning framework remains robust even under substantial architectural heterogeneity. The observed variability therefore reflects intrinsic differences in design-space geometry rather than instability of the modeling approach. Figure 7 extends this analysis by systematically examining how convergence behavior shifts under different predictive stringency levels.

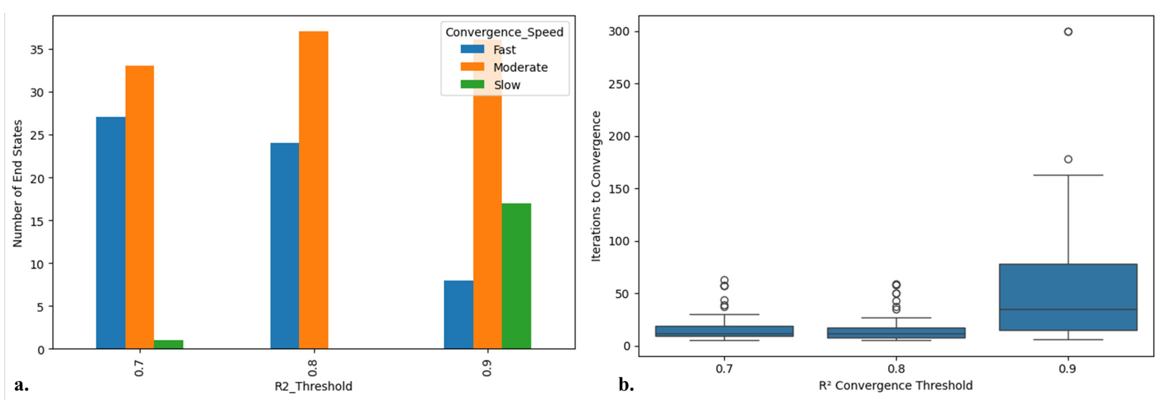

Figure 7. Convergence behavior across different R^2 thresholds. (a) Distribution of convergence speed categories (Fast: ≤30 iterations; Moderate: 31–60 iterations; Slow: >60 iterations) for R^2 thresholds of 0.7, 0.8, and 0.9. (b) Boxplot of the number of iterations required to reach convergence under each R^2 threshold. Increasing the predictive accuracy requirement leads to a noticeable shift toward slower convergence and greater variability in iteration counts.

Figure 7-a categorizes end-states into fast, moderate, and slow convergence groups across R^2 thresholds of 0.7, 0.8, and 0.9. At lower thresholds (0.7 and 0.8), most states fall within the fast and moderate categories, indicating that relatively relaxed predictive requirements allow rapid stabilization across most structural configurations. However,

when the threshold is increased to 0.9, the distribution shifts noticeably toward the slow category, revealing a greater proportion of structurally complex states requiring extended exploration. This shift is quantitatively reinforced in Figure 7-b, where both the median number of iterations and the spread of convergence times increase substantially at the 0.9 threshold. Building on the convergence analysis in Figure 7, Figure 8 quantifies the practical impact of active learning in terms of computational savings.

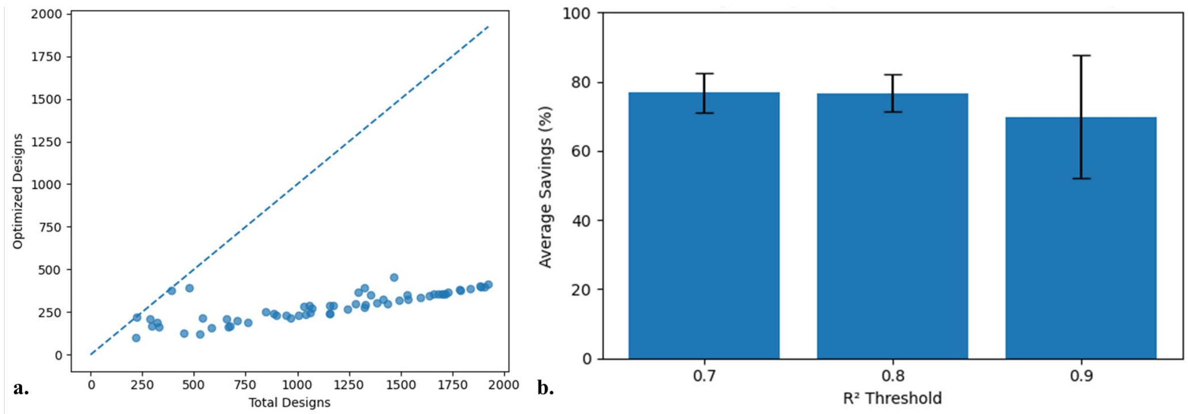

Figure 8. Computational savings from surrogate-based active learning. (a) Evaluated versus total designs at $R^2 = 0.90$; the dashed line indicates full enumeration. (b) Average percentage of design space reduction across R^2 thresholds (0.7, 0.8, 0.9), with error bars showing variability across end-states.

While Figure 7 demonstrated that stricter R^2thresholds increase the number of iterations required for surrogate stabilization, Figure 8 shows that even under the most stringent requirement ($R^2 = 0.90$), the framework evaluates only a fraction of the full combinatorial design space. Figure 8-a compares the number of explicitly evaluated designs against the total number of possible configurations, with all points lying well below the full-enumeration diagonal. This gap represents the avoided evaluations enabled by surrogate-based estimation. Figure 8-b further summarizes the average percentage of design space reduction across R^2thresholds, showing that approximately 65%–80% of design evaluations are avoided relative to exhaustive SBD, depending on the predictive accuracy requirement. Importantly, while increasing the R^2 threshold from 0.7 to 0.9 leads to reduced computational savings, substantial reductions are maintained even at high predictive fidelity levels. These results demonstrate that the proposed framework achieves a strong balance between computational efficiency and surrogate accuracy. The resulting refined end-state sets preserve the structural characteristics of the objective space, including the multimodal clustering behavior observed in Objective 3, indicating that the method maintains the exploratory nature of set-based design while significantly reducing evaluation cost.

From a practical implementation perspective, the proposed framework enables systematic exploration of large-scale combinatorial vehicle architectures without exhaustive evaluation of every configuration. The case study spans major automotive subsystems, including powertrain, chassis and suspension, electrical architecture, body structure, and interior components, each typically developed under competing objectives such as performance, cost, efficiency, safety, and compliance. By combining set-based design with active learning, the framework coordinates these interacting decisions while preserving architectural flexibility. Rather than committing to early point-based selections, it progressively identifies high-performing configurations, reducing redesign risk and supporting informed cross-functional trade-offs. Importantly, the results demonstrate substantial computational savings, as only a fraction of the full design space requires explicit evaluation even under stringent predictive requirements.

The analysis further shows that subsystem-level heterogeneity directly influences exploration complexity. Architectures with dominant option patterns yield smoother objective landscapes and faster surrogate convergence, whereas subsystems introducing discrete structural regimes create multimodal performance clusters that demand additional sampling. For automotive manufacturers, this provides actionable insight into which subsystem families require deeper validation, simulation resources, or prototype testing. The framework advances a shift from intuition-driven architecture selection toward data-driven system design, simultaneously improving robustness and reducing computational burden in early-stage vehicle development.

In this study, multiple objectives are aggregated using a weighted-sum formulation to enable tractable surrogate learning and convergence monitoring. While this scalarization supports efficient exploration under a fixed preference structure, future work may extend the framework to multi-output surrogate modeling to preserve full geometry and explicitly capture trade-off structure. Although the proposed framework demonstrates scalability and robustness in large synthetic vehicle architectures, several limitations remain. Surrogate performance depends on the representativeness of sampled end-states, and highly irregular or discontinuous objective landscapes may require more advanced adaptive sampling strategies. Additionally, because objective values are analytically generated in this study, future research should examine integration with high-fidelity simulation environments to evaluate computational performance in more realistic settings.

6. Conclusion

This study proposed an active learning–based framework for efficiently exploring large-scale combinatorial vehicle architectures within a set-based design context. By integrating surrogate modeling with iterative sampling, the approach enables systematic evaluation of interacting subsystems, including powertrain, chassis, electrical architecture, body structure, and interior components, without exhaustive enumeration of all configurations. The results show that convergence behavior is strongly influenced by architectural heterogeneity. Structurally smoother regions of the design space stabilize rapidly, whereas configurations associated with discrete subsystem regimes require extended exploration. Increasing the predictive stringency (R^2 threshold) systematically increases convergence time and variability, yet the majority of end-states ultimately achieve high predictive fidelity. This demonstrates that the proposed framework remains robust across diverse architectural conditions while providing practical guidance for early-stage vehicle platform development. In addition, sensitivity analyses on surrogate model selection and algorithmic parameters confirm the robustness of the framework while highlighting the importance of model choice in achieving optimal computational efficiency.

Acknowledgement

This work is supported by the U.S. Army Small Business Innovation Research (SBIR) Phase II under Award No. 206995 through Agile Systems LLC.

References

Cheng, R., Jin, Y., Olhofer, M., and Sendhoff, B., "A reference vector guided evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 773–791, 2016.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T., "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

Deb, K. and Jain, H., "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

Eghbali-Zarch, M., Jahanmahin, R., Masoud, S., Yildirim M., Chinnam R.B., Murat E., Arslanturk, S., Dalkiran, E., Many-Objective Evolutionary Algorithms for Optimizing Set-Based Design Tradespace Exploration in Complex Engineered Systems, Working Paper.

Ghosh, S. and Seering, W., "Set-based thinking in the engineering design community and beyond," *Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE)*, Buffalo, NY, USA, Aug. 17–20, 2014.

Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., and Crombecq, K., "A surrogate modeling and adaptive sampling toolbox for computer-based design," *Journal of Machine Learning Research*, vol. 11, pp. 2051–2055, 2010.

Guo, Y., "Active learning for adaptive surrogate model improvement in high-dimensional problems," *Structural and Multidisciplinary Optimization*, 2024.

Rapp, S., "Set-based design and optimization: Can they live together and make better trade-space decisions?" *Proceedings of the INCOSE International Symposium*, vol. 28, no. 1, 2018.

Singer, D. J. and Doerry, N., "What is set-based design?" *Naval Engineers Journal*, vol. 121, no. 4, pp. 31–43, 2009.

Sobek, D. K., Ward, A. C., and Liker, J. K., "Toyota's principles of set-based concurrent engineering," *Sloan Management Review*, vol. 40, no. 2, pp. 67–83, 1999.

Toche, B., Pellerin, R., and Fortin, C., "Set-based design: A review and new directions," *Design Science*, vol. 6, 2020, https://doi.org/10.1017/dsj.2020.16

Ward, A., Liker, J. K., Cristiano, J. J., and Sobek, D. K., "The second Toyota paradox: How delaying decisions can make better cars faster," *Sloan Management Review*, vol. 36, no. 3, pp. 43–61, 1995.

Zhang, Q. and Li, H., "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.

Biographies

Roohollah Jahanmahin Roohollah Jahanmahin received his B.S. in Industrial Engineering from University of Tehran in 2018 and his M.S. in Industrial Engineering from Iran University of Science and Technology in 2021. He is currently a Ph.D. Candidate Research Assistant in the Department of Industrial and Systems Engineering at Wayne State University and is affiliated with the Smart Immersive Modeling Lab. His research lies at the intersection of human–vehicle interaction and autonomous driving. His work focuses on developing human-centered, data-driven frameworks to understand and predict driver behavior, enabling personalized control transitions, and enhancing the safety and adaptability of semi-autonomous systems in mixed-traffic environments.

Dr. Maryam Eghbalizarch is an Assistant Professor of Industrial and Systems Engineering at Kennesaw State University. Her work develops analytical and computational methods to enhance performance, resilience, and safety in high-dimensional design and operational environments. She integrates optimization, simulation, machine learning, and statistical modeling to create scalable and interpretable tools for system design and evaluation. During her postdoctoral appointment, she contributed to interdisciplinary research on surrogate-assisted optimization and uncertainty-aware modeling, earning consecutive Graduate School Postdoctoral Research Awards.

Dr. Sara Masoud is an Assistant Professor of Industrial and Systems Engineering at Wayne State University and founding director of the Smart Immersive Modeling (SIM) Lab. Her research focuses on human-centered system design, set-based design, and data-driven decision support for complex systems under uncertainty. She develops scalable computational frameworks integrating optimization, machine learning, simulation, and immersive technologies to address high-dimensional, combinatorial design challenges. Her work spans automotive systems, human–autonomy teaming, smart agriculture, and cyber-physical systems, and has been recognized with multiple awards.

Dr. Murat Yildirim is an Associate Professor of Industrial and Systems Engineering at Wayne State University. His research focuses on optimization, data-driven decision making, and applications of mathematical programming to complex engineering and operational systems. He develops advanced analytic and computational methods for large-scale optimization problems under uncertainty, including stochastic and robust optimization, with applications spanning energy systems, transportation, healthcare operations, and engineering design. Dr. Yildirim's work integrates rigorous theory with practical implementation, contributing to both foundational advances and real-world problem solving.

Dr. Neda Masoud is an Associate Professor at the University of Michigan, where her research focuses on transportation systems, connected and automated vehicles, and human-centered mobility. Her work lies at the intersection of transportation engineering, human factors, and data-driven modeling, with an emphasis on understanding and improving interactions between humans and emerging mobility technologies. Dr. Masoud's research addresses critical challenges in the design and operation of autonomous and semi-autonomous transportation systems, including safety, efficiency, trust, and decision-making under uncertainty. Her work often integrates behavioral modeling with system optimization to support more resilient and human-aware transportation solutions. Her research has been supported by federal agencies and industry partners and has been published in leading transportation and engineering journals.

Dr. Evrim Dalkiran is an Associate Professor of Industrial and Systems Engineering at Wayne State University, specializing in optimization, systems modeling, and data-driven decision support for complex engineered systems. Her research develops scalable analytical and computational methods for high-dimensional, combinatorial design problems under uncertainty. She has contributed to modeling large-scale engineering systems using optimization and simulation, with emphasis on feasibility analysis and performance trade-offs in early-stage design. Dr. Dalkiran collaborates across disciplines to translate rigorous optimization methods into practical, transparent decision-support frameworks that enhance system performance and resilience.

Dr. Ratna Babu Chinnam is the Chair and a Professor of Industrial and Systems Engineering at Wayne State University. His research spans optimization, reliability analytics, and quality engineering, with applications in manufacturing systems, healthcare, and service operations. He develops data-driven methodologies for system performance improvement, reliability prediction, and process optimization. Dr. Chinnam's work integrates statistical modeling, lean methods, and risk analysis to support decision making in complex engineered and operational environments.

Dr. Alper Murat is an Associate Professor of Industrial and Systems Engineering at Wayne State University. His research focuses on optimization, applied probability, and stochastic modeling for complex engineering and operational systems. He works on developing scalable analytical methods for resource allocation, network design, and performance evaluation under uncertainty. Dr. Murat's interdisciplinary work spans transportation systems, logistics, and service operations, combining theory with practical solution techniques.

Dr. Suzan Arslanturk is an Associate Professor in Computer Science at Wayne State University specializing in data analytics, machine learning, and predictive modeling for engineering applications. Her research develops interpretable and robust data-driven methods for system performance prediction, condition monitoring, and intelligent decision support. Dr. Arslanturk's work span fault diagnosis, quality improvement, and automated modeling frameworks that support reliability and efficiency in complex engineered systems.