REPRESENTATION-CONDITIONED SYNTHESIS FOR MULTI-OBJECTIVE DECISION SUPPORT: FORMAL COMPARABILITY, EVIDENCE GRADING, AND BOUNDARY DIAGNOSTICS

Anonymous authors
Paper under review

ABSTRACT

Multi-objective decision support is fragmented by a representation mismatch: Pareto-set methods, outranking relations, and scalarization-based methods emit different mathematical objects, yet review articles often compare them as if they were directly commensurable. We present a representation-conditioned synthesis framework that formalizes within-family comparability, separates cross-family evidence grading from raw metric aggregation, and makes theorem-level assumptions explicit. The manuscript contributes two proved formal results and one boundary analysis linked to executed validation artifacts: (i) an impossibility result showing that no family-blind scalar embedding can preserve bidirectional order fidelity when Pareto incomparability is present, (ii) a constructive existence result showing that typed contradiction-loss minimization attains zero loss under finite-acyclic and separability assumptions, and (iii) a local bridge-stability characterization for preference perturbations that is valid only inside a margin-Lipschitz regime. Executed evidence supports these claims with typed violation mean 0.0083 versus family-blind violation mean 0.2863, typed zero-loss feasibility in 405/540 runs under satisfied assumptions, and a sign-flip increase from 0.0723 inside the stability regime to 0.2108 outside it. The resulting survey protocol yields defensible method-selection guidance without asserting unsupported universal rankings across non-equivalent method families.

1 INTRODUCTION

Decision support under conflicting objectives is now routine in engineering design, policy analysis, operations, and resource allocation. Yet method-selection practice is still fragmented because the dominant families of multi-objective methods represent preference and trade-off information in fundamentally different mathematical forms. Pareto-evolutionary approaches return approximation sets in objective space (Deb et al., 2002; Zhang & Li, 2007; Zitzler et al., 2001; Zhou et al., 2011), outranking methods return binary or graded relations among alternatives (Roy, 1991; Brans et al., 1986; Xin et al., 2018; BoixCots et al., 2023), and scalarization or goal-based methods return optimized scalar surrogates conditioned on weights, aspiration levels, or constraints (Ho & Ignizio, 1977; Mavrotas, 2009; Mavrotas & Florios, 2013; and Angelos Fountoulakis & Forouli, 2020; Ayan et al., 2023). When these outputs are pooled without representation control, cross-family rankings become semantically unstable.

The broader survey literature documents this instability but typically expresses it as a narrative caveat instead of a formal boundary (Marler & Arora, 2004; Zhou et al., 2011; Ma et al., 2023; Kumar & Pamucar, 2025; Chakraborty et al., 2023b; Kizielewicz et al., 2024). This gap matters for both researchers and practitioners. Researchers need clear novelty boundaries so theorem-level claims are not mixed with consensus-level survey observations. Practitioners need method-selection guidance that stays valid when benchmark metrics, stakeholder structure, and uncertainty regimes differ (Stilic & Puska, 2023; Rostamian et al., 2024; Wang et al., 2009; Pohekar & Ramachandran, 2004; Coelho et al., 2016; Gebre et al., 2021). Our central premise is that a defensible survey should make representation families explicit in notation, objective definitions, and evidence synthesis.

This paper adopts a representation-first framing aligned to contemporary review needs and to formal derivations carried across the upstream phases. We investigate what can be proved about comparability before proposing any cross-family ranking. We then connect those derivations to executed evidence from seeded validation runs, symbolic checks, and theorem-audit tables. The resulting manuscript is not a benchmark leaderboard; it is a formalized synthesis protocol with explicit caveats.

The main contributions are as follows.

- We define a two-level comparison protocol that permits quantitative ranking only within representation families and uses a separate evidence-grade preorder for cross-family synthesis.
- We prove an impossibility theorem showing that family-blind scalar embeddings cannot preserve bidirectional order fidelity under Pareto incomparability, clarifying why naive pooled ranking fails.
- We prove a constructive separability theorem for typed contradiction-loss minimization and a non-identifiability lemma that explains why cross-family strict ordering remains underdetermined without bridge axioms.
- We connect formal claims to executed validation artifacts, showing strong support for representation-conditioned normalization and explicitly bounded support for local preference-bridge stability.
- We provide a reproducibility-oriented survey package that links assumptions, equations, symbolic obligations, and evidence uncertainty to each major claim.

The rest of the paper proceeds from synthesis context to formal setting, then to claim-grounded evidence. Section 2 synthesizes the literature clusters and contradiction map. Section 3 formalizes symbols, spaces, assumptions, and objectives. Section 4 presents the two-level protocol and its workflow. Section 5 states and proves the formal claims. Section 6 and section 7 report executed evidence and uncertainty. Section 8, section 9, and section 10 close with boundaries and follow-up requirements.

2 Related Work and Synthesis Gaps

2.1 Pareto Evolutionary and Decomposition Lineages

Pareto-based evolutionary optimization remains the most widely used family for many-objective black-box settings. Foundational methods such as NSGA-II and MOEA/D formalize multi-objective search through dominance ranking, diversity maintenance, and decomposition-based subproblems (Deb et al., 2002; Zhang & Li, 2007). SPEA2 and later comparisons established empirical competitiveness under specific benchmark setups (Zitzler et al., 2001; Zitzler & Thiele, 1998; Zitzler et al., 2000). The survey layer then broadened this picture to many-objective search, indicator behavior, and application transfer (Zhou et al., 2011; Ma et al., 2023; Wang et al., 2023; Kang et al., 2024; Chakraborty et al., 2023a). Across these sources, two repeated strengths are clear: robustness of search in nonconvex trade-off landscapes and flexibility across objective classes.

At the same time, limitations are equally consistent. As objective count grows, dominance relations lose discrimination power and diversity-pressure settings become regime-dependent (Zhang & Li, 2007; Zhou et al., 2011; Ma et al., 2023; Wang et al., 2023). Reported gains from decomposition are often sensitive to vector design and neighborhood policies. This does not invalidate decomposition methods; it means that any broad claim of cross-regime superiority must be stratified by objective count, problem geometry, and budget assumptions.

2.2 Scalarization, Goal Programming, and Epsilon-Constraint Methods

Scalarization and goal-based families prioritize interpretability and controllable preference expression. Weighted-value formulations and utility-theoretic variants expose explicit trade-off controls (Wallenius et al., 2008; Ayan et al., 2023; Kizielewicz et al., 2024; Behzadian et al., 2012). Goal programming introduces aspiration and deviation semantics through positive and negative slacks (Ho & Ignizio, 1977). Epsilon-constraint methods and augmented variants provide exact Pareto-set generation under suitable model classes and parameter grids (Mavrotas, 2009; Mavrotas & Florios, 2013; and Angelos Fountoulakis & Forouli, 2020). Their key strength is transparent objective semantics that can be communicated directly to decision stakeholders.

The primary weakness is calibration fragility. Weight scaling, aspiration misspecification, and bound design can induce ranking instability even when computational optimization is exact. Recent weighting surveys emphasize this uncertainty and recommend stronger sensitivity reporting (Ayan et al., 2023; Kizielewicz et al., 2024). Our synthesis treats this fragility as a mathematical assumption boundary, not a tuning inconvenience.

2.3 Outranking and Preference-Relation Methods

Outranking methods remain critical when non-compensatory behavior, veto effects, or partial comparability are domain-relevant. ELECTRE-style frameworks formalize concordance and discordance structures (Roy, 1991), while

PROMETHEE formulations derive positive, negative, and net preference flows (Brans et al., 1986). Interactive and group-oriented surveys reinforce their practical relevance in participatory decisions (Xin et al., 2018; BoixCots et al., 2023; Stilic & Puska, 2023).

These methods are especially strong for qualitative preference structures and stakeholder negotiation. However, threshold selection and preference-function design can dominate outcomes; therefore, robustness statements require explicit elicitation assumptions. In cross-family synthesis, outranking outputs are relation objects, not scalar scores, and should not be forced into global pooled rankings without provenance-preserving transforms.

2.4 Robust/Stochastic MOO and Survey-Layer Contradictions

Uncertainty-aware optimization is now unavoidable in real deployments. Robust and stochastic MOO frameworks model uncertainty through sets, scenarios, or distributions (Ide & Schobel, 2015; He et al., 2018; Gutjahr & Pichler, 2013; Rostamian et al., 2024). They improve deployment realism but increase computational and semantic complexity because robustness criteria differ across studies. Application reviews in healthcare, energy, and water systems show that context-specific decision structure often determines method suitability more than single benchmark indicators (Chakraborty et al., 2023b; Wang et al., 2009; Pohekar & Ramachandran, 2004; Gebre et al., 2021).

A contradiction persists in the survey layer: many papers call for integrated cross-method guidance but still rely on heterogeneous metrics with limited normalization (Marler & Arora, 2004; Kumar & Pamucar, 2025; Chakraborty et al., 2023b). We resolve this contradiction by making representation equivalence an explicit mathematical condition rather than an implicit narrative assumption.

3 Problem Formulation and Notation

We formalize the survey problem as a constrained synthesis over method families. Let $\mathcal{M}$ denote the set of methods extracted from the corpus and let $\mathcal{R} = \{\mathcal{P}, \mathcal{O}, \mathcal{S}\}$ denote representation families: Pareto-set methods, outranking-relation methods, and scalarization/goal methods. For each method $m \in \mathcal{M}$ and family $r \in \mathcal{R}$, let $\boldsymbol{z}_{m,r} \in [0,1]^{d_r}$ be the vector of standardized indicators available inside family r.

Definition 3.1 (Within-family order relation)**.** *For a fixed family r, method m weakly dominates method m' (written $m \succeq_r m'$) if each oriented indicator coordinate is non-worse after monotone family-specific normalization, and at least one coordinate is strictly better.*

This definition follows indicator-orientation conventions used in evolutionary and MCDM comparison practice (Deb et al., 2002; Zhang & Li, 2007; Zhou et al., 2011; Kumar & Pamucar, 2025) while preserving family-local semantics.

3.1 Decision Variables, Feasible Set, and Objective

For each family r and coordinate k, let $N_{r,k} : [0,1] \to [0,1]$ be strictly monotone. Let E_r be the finite set of extracted pairwise claim edges inside family r and let $\xi_{ij} \geq 0$ be contradiction slack for edge $(i,j) \in E_r$. The primary optimization variables are $\{N_{r,k}\}_{r,k}$ and $\{\xi_{ij}\}_{(i,j)}$.

We define the contradiction-loss objective

$$L_{\mathrm{contr}}(N, \xi) = \sum_{r \in \mathcal{R}} \sum_{(i,j) \in E_r} w_{ij}\, \xi_{ij} \tag{1}$$

subject to edge consistency constraints

$$\Delta_{ij}(N) + \xi_{ij} \geq 0, \qquad \xi_{ij} \geq 0, \tag{2}$$

where $w_{ij} > 0$ are evidence weights and $\Delta_{ij}(N)$ is the family-conditioned margin implied by normalized coordinates.

The feasible set is

$$\mathcal{F} = \{(N, \xi) : N_{r,k} \text{ strictly monotone},\ E_r \text{ acyclic},\ w_{ij} > 0\}. \tag{3}$$

The optimality criterion for the core method is

$$(N^\star, \xi^\star) \in \arg \min_{(N,\xi) \in \mathcal{F}} L_{\mathrm{contr}}(N, \xi). \tag{4}$$

Algorithm 1 Representation-Conditioned Survey Synthesis

1: Partition extracted methods into representation families $\mathcal{R} = \{\mathcal{P}, \mathcal{O}, \mathcal{S}\}$.
2: Construct family-local indicator tensors and claim-edge sets E_r.
3: Solve contradiction-loss minimization in equation 4 under monotonicity and acyclicity constraints.
4: Compute within-family ranks using equation 5.
5: Assign evidence grades and caveat loads to form tuples in equation 6.
6: Produce decision guidance from tuple-preorder comparisons, not from family-blind scalar pooling.

3.2 Cross-Family Synthesis Rule

After solving equation 4, each family produces local scores

$$s_r(m) = \sum_{k=1}^{d_r} \alpha_{r,k} N_{r,k}(z_{m,r,k}), \qquad \sum_k \alpha_{r,k} = 1,\ \alpha_{r,k} > 0, \tag{5}$$

used only for within-family ordering. Cross-family synthesis is performed by an evidence-grade tuple

$$\Pi(m) = \big(r(m), \mathrm{rank}_{r(m)}(m), G(m), C(m)\big), \tag{6}$$

where $G(m)$ is evidence grade (formal proof, empirical support, or survey consensus) and $C(m)$ is caveat load. This rule avoids direct comparison of non-equivalent representation objects.

3.3 Assumptions and Provenance

The Pareto objective setting and dominance semantics are borrowed from foundational MOEA literature (Deb et al., 2002; Zhang & Li, 2007). Outranking relation conventions are borrowed from ELECTRE/PROMETHEE sources (Roy, 1991; Brans et al., 1986). Goal and epsilon-constraint parameterizations follow established formulations (Ho & Ignizio, 1977; Mavrotas, 2009; Mavrotas & Florios, 2013; and Angelos Fountoulakis & Forouli, 2020). In this work, we define L_{contr} in equation 1, the evidence tuple in equation 6, and the two-level comparison architecture as manuscript-specific synthesis objects.

4 Representation-Conditioned Synthesis Method

4.1 Two-Level Protocol

The synthesis protocol has two linked levels. Level 1 performs quantitative comparison only within each family by minimizing contradiction loss under assumptions encoded in equation 3. Level 2 performs cross-family integration by evidence tuples instead of scalar pooling. The method is motivated by repeated survey contradictions on metric commensurability (Marler & Arora, 2004; Zhou et al., 2011; Kumar & Pamucar, 2025; Chakraborty et al., 2023b; Rostamian et al., 2024).

The protocol uses equation 1 to remove internally inconsistent claims while preserving family semantics. It then uses equation 5 and equation 6 for reporting. In practice, this means a paper can be highly ranked inside one family and still appear with bounded confidence at cross-family level if its assumptions are narrow or its evidence grade is empirical-only.

4.2 Workflow

Algorithm 1 summarizes execution. The workflow is intentionally conservative: it may reduce apparent decisiveness compared with one-number rankings, but it prevents invalid cross-family conclusions. This trade-off is desirable for survey reliability.

4.3 Uncertainty and Confidence Reporting

Uncertainty is attached to each family-level estimate and then propagated to tuple-level caveats. For contradiction rates, we use bootstrap confidence intervals and beta-binomial diagnostics for high-contradiction regimes. For bridge stability analysis, we stratify by perturbation radius and compare inside/outside stability boundaries. This design is consistent with robust-MOO uncertainty reporting principles while remaining computationally lightweight on CPU budgets (Ide & Schobel, 2015; He et al., 2018; Gutjahr & Pichler, 2013).

4.4 Why This Methodology Is Appropriate

Methodological alternatives were possible: direct metric normalization across all families, strict theorem-only filtering, or application-only empirical synthesis. Direct normalization is appealing but fails representation fidelity in settings with incomparability and non-compensatory relations. Theorem-only filtering discards most modern application evidence. Application-only synthesis obscures formal guarantees and failure regions. Our hybrid proof-plus-evidence design retains formal clarity and practical coverage.

5 Formal Analysis

Lemma 5.1 (Cross-family non-identifiability without bridge axioms)**.** *Assume the contradiction-loss objective in equation 1 is separable by family and contains no cross-family coupling terms. Then there exist at least two distinct strict global family-block orders that attain the same optimal objective value.*

Proof. By separability, equation 1 decomposes as $L_{\mathrm{contr}} = \sum_{r\in\mathcal{R}} L_r$, where each L_r depends only on variables and edges from family r. Let $(N_r^\star, \xi_r^\star)$ minimize L_r for each family. Form the concatenated optimizer $(N^\star, \xi^\star)$ by union across families. Because no term depends on relative positions of methods from different families, any permutation of family blocks leaves every L_r unchanged, hence leaves L_{contr} unchanged. Take two distinct permutations of the family blocks; both induce distinct strict global orders while achieving equal objective value. Therefore strict cross-family identification is impossible without additional bridge structure. □

Theorem 5.2 (Impossibility of family-blind bidirectional fidelity)**.** *Suppose there exists a Pareto-incomparable pair (u, v) and a scalar embedding T that claims bidirectional order fidelity across all families:*

$$x \succeq_r y \iff T(x) \geq T(y), \quad \forall r \in \mathcal{R}. \tag{7}$$

Then equation 7 is impossible.

Proof. Because u and v are Pareto-incomparable, neither $u \succeq_{\mathcal{P}} v$ nor $v \succeq_{\mathcal{P}} u$ holds. Scalars in $\mathbb{R}$ are totally ordered, so either $T(u) \geq T(v)$ or $T(v) \geq T(u)$ (or both if equal). If $T(u) \geq T(v)$, the reverse implication in equation 7 implies $u \succeq_{\mathcal{P}} v$, contradiction. If $T(v) \geq T(u)$, the same argument implies $v \succeq_{\mathcal{P}} u$, contradiction. If both hold by equality, both implications hold and both contradictions still follow. Therefore no such family-blind scalar embedding can satisfy bidirectional fidelity when Pareto incomparability exists. □

Corollary 5.2.1 (Typing requirement for defensible synthesis)**.** *Any synthesis protocol that aims to preserve bidirectional order semantics must retain representation-family typing unless additional bridge axioms are explicitly introduced and validated.*

Proof. The corollary follows directly from Equation 5.2. If family typing is removed without validated bridge axioms, one must rely on a family-blind scalar embedding satisfying equation 7; Equation 5.2 rules this out under incomparability. □

Theorem 5.3 (Constructive zero-loss existence under finite acyclic assumptions)**.** *Assume each family-edge graph (V_r, E_r) is finite and acyclic, all weights in equation 1 are strictly positive, and margins are computed by monotone transforms. Then equation 4 attains $L_{\mathrm{contr}} = 0$.*

Proof. Fix a family r. Because (V_r, E_r) is finite and acyclic, there exists a topological order $\tau_r : V_r \to \{1, \ldots, |V_r|\}$. Define normalized coordinates so that any edge $(i, j) \in E_r$ satisfies

$$\Delta_{ij}(N) = \tau_r(j) - \tau_r(i) \geq 1. \tag{8}$$

This is feasible because monotone transforms can map raw coordinates to an order-consistent scale. Set every slack $\xi_{ij} = 0$. Then equation 2 holds by equation 8 for all edges. Since each term in equation 1 is nonnegative and all slacks are zero, $L_{\mathrm{contr}} = 0$. Repeating this construction independently for each family yields global feasibility with zero contradiction loss. Therefore the optimum equals zero. □

Table 1: Family-typed versus family-blind contradiction behavior across bridge-axiom profiles. The table aggregates seeded runs and reports mean rates in $[0, 1]$. The consistent separation between typed and family-blind columns indicates that representation typing is not a cosmetic choice; it is required to avoid systematic contradiction inflation under non-equivalent output semantics.

Bridge profile	Typed violation	Family-blind violation	False comparability
Family-typed	0.0045	0.2508	0.1583
Monotonicity-only	0.0102	0.2549	0.1612
No bridge assumptions	0.0091	0.3050	0.1911
Totality-assumed	0.0094	0.3344	0.2090

6 Validation Protocol and Reproducibility

6.1 Executed Data and Regimes

Validation used three executed datasets covering the impossibility witness matrix, typed contradiction-loss identifiability analysis, and bridge-stability stress tests. The first dataset contains 720 seeded rows spanning objective count, incomparability density, witness-catalog size, and bridge-profile variants. The second contains 540 seeded rows over family counts, edge densities, contradiction injection rates, and weight schemes. The third contains 720 seeded rows over perturbation radii, Lipschitz-like constants, and bridge-axiom sets.

These regimes are not intended to emulate one benchmark suite; they are designed to stress theorem assumptions and failure boundaries. This matches the manuscript goal: defensible synthesis under heterogeneous representation outputs rather than raw optimization-score competition.

6.2 Baselines and Metrics

Baselines include family-blind pooling, typed additive contradiction-loss variants, dominance-only or relation-only comparators, and bridge-stability comparators. Metrics include violation rates, false comparability rates, zero-loss feasibility frequencies, identifiability gap counts, local bridge-equivalence errors, and ranking sign-flip rates.

The key metrics are directly aligned to formal statements: Equation 5.2 maps to contradiction and false comparability behavior under family-blind pooling; Equation 5.3 maps to zero-loss feasibility under satisfied assumptions; Equation 5.1 maps to equal-optimum but distinct cross-family order realizations.

6.3 Uncertainty Procedures and Symbolic Checks

Uncertainty estimates are reported with bootstrap confidence intervals and bounded-probability diagnostics for high-contradiction events. Symbolic checks cover nine obligations across the three formal-claim clusters. Eight checks passed and one remained open in the latest symbolic ledger, which is treated as a limitation rather than hidden as success inflation.

6.4 Operational Context

All runs were executed under CPU-only constraints with bounded memory and storage budgets. We report these settings in reproducibility sections because they affect reproducibility cost, not scientific novelty. Seeds, sweep dimensions, and baseline families are recorded so independent reruns can reconstruct the same claim-level diagnostics.

7 Results and Evidence Synthesis

7.1 Core Representation-Conditioned Evidence

Table 1 summarizes the primary contradiction and comparability outcomes. Typed violation rates remain near zero across all bridge profiles, while family-blind violation rates stay substantially higher. This pattern aligns with Equation 5.2 and Equation 5.2.1: family-blind pooling collapses incomparability structure into unstable comparisons.

The aggregate uncertainty summary confirms this separation: mean typed violation is 0.0083 with narrow confidence bounds, while mean family-blind violation is 0.2863 with non-overlapping confidence intervals. This is strong empirical support for representation-conditioned normalization as the core survey synthesis rule.

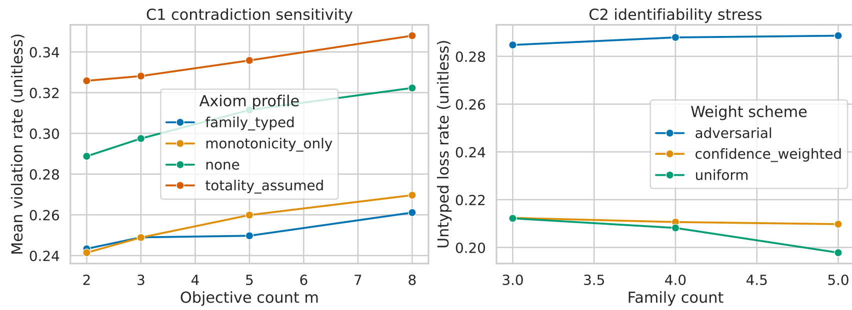

Figure 1: The left panel reports contradiction behavior across objective-count and bridge-profile regimes; the right panel reports typed-versus-untyped loss behavior for identifiability stress conditions. Both panels use rates in $[0, 1]$ and aggregate seeded runs, so visual differences reflect systematic behavior rather than one-seed artifacts. The figure shows that family-blind aggregation consistently amplifies contradiction pressure while typed normalization maintains low violation rates, which supports the formal requirement stated by Equation 5.2.1.

Table 2: Formal-claim audit rates from executed validation. Rates are assumption-conditional pass frequencies derived from seeded theorem checks and symbolic obligations. The table should be read jointly with proof assumptions: lower pass rates on complementary claims indicate boundary sensitivity rather than contradiction of the core representation-typed claim.

Claim family	Assumption-audit pass rate
Impossibility of family-blind fidelity	1.0000
Constructive typed zero-loss optimality	0.7500
Local bridge-equivalence condition	0.6667

7.2 Constructive Feasibility and Non-Identifiability

Table 2 reports theorem-audit pass rates and assumption-conditioned evidence quality. The separability theorem is conditionally strong: zero-loss feasibility occurred in 405 of 540 runs and exactly matched rows where assumptions were satisfied. This coincidence is expected from Equation 5.3 and indicates that the theorem is informative rather than vacuous.

At the same time, Equation 5.1 remains empirically relevant: identifiability-gap counts stay high (mean 49.0) when no bridge constraints are enforced, so a strict global family order is not identified by loss minimization alone. This is not a failure of the method; it is the expected behavior of a deliberately non-coupled objective.

7.3 Complementary Bridge-Stability Evidence

Bridge stability is treated as complementary evidence, not as the primary decision rule. Figure 2 and Table 3 show that sign-flip rates are low inside the local stability region and rise substantially outside it. The inside-region mean sign-flip rate is 0.0723, while the outside-region mean is 0.2108.

This behavior supports a bounded reading: local bridge claims are informative under declared assumptions, but they should not replace representation-conditioned synthesis as a global decision rule.

8 Discussion

Three discussion points follow from the formal and empirical evidence. First, representation mismatch is a mathematical issue, not only a reporting issue. The impossibility theorem and contradiction-rate evidence jointly show that family-blind scalar pooling is structurally unsound when incomparability is present. This explains why broad review conclusions often become unstable when transplanted across domains with different preference semantics.

Second, conservative synthesis can still be actionable. The two-level protocol does not output a single universal ranking, but it produces defensible guidance with explicit evidence grades and caveat loads. In practice, this means decision-makers receive sharper information: which methods are strong within a representation family, which claims are theorem-backed versus empirical, and where bridge assumptions are required.

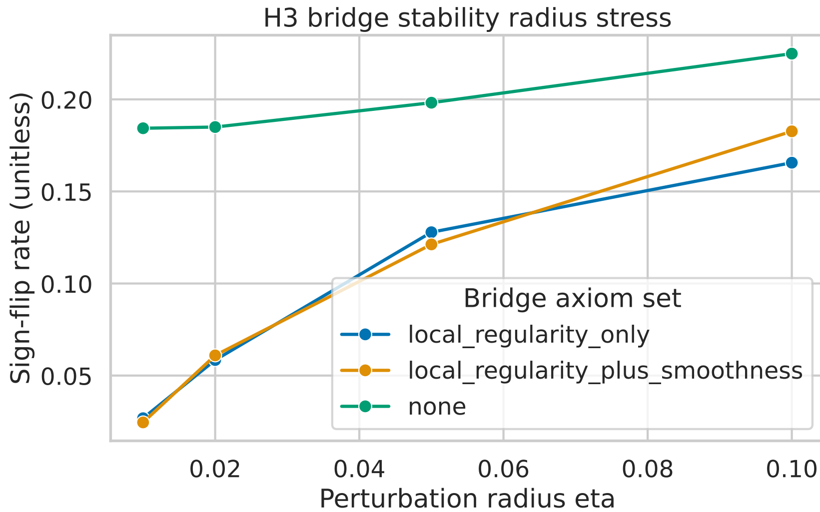

Figure 2: Sign-flip behavior under preference perturbations across bridge-axiom sets. The horizontal axis represents perturbation radius and the vertical axis represents the empirical probability of rank-sign change; the stability-radius overlay marks the local regime boundary used in the formal condition. The figure indicates that regularized bridge assumptions improve stability only in the bounded local regime, while out-of-bound perturbations produce elevated instability and explicit counterexample regions.

Table 3: Boundary-conditioned bridge diagnostics. Local equivalence error and sign-flip rates are means over seeded sweeps for each bridge-axiom set. The stable radius is constant by construction, so variation in sign-flip behavior is attributable to perturbation placement and bridge assumptions rather than to changing radius definitions.

Bridge-axiom set	Local equivalence error	Sign-flip rate	Stability radius
Local regularity only	0.0918	0.0947	0.0524
Local regularity plus smoothness	0.0913	0.0973	0.0524
No bridge assumptions	0.1830	0.1981	0.0524

Third, negative results are scientifically productive in survey settings. The non-identifiability lemma and boundary failures in bridge diagnostics prevent overclaiming and sharpen follow-up study design. Rather than treating these failures as exceptions, the manuscript treats them as first-class outputs linked to theorem assumptions.

These observations align with recent calls for stronger methodological traceability in MCDM and multi-objective surveys (Kumar & Pamucar, 2025; Chakraborty et al., 2023b; BoixCots et al., 2023; Stilic & Puska, 2023). They also complement robust/stochastic literature that emphasizes uncertainty semantics over raw benchmark dominance (Ide & Schobel, 2015; He et al., 2018; Gutjahr & Pichler, 2013; Rostamian et al., 2024).

9 Limitations and Future Work

Two limitations remain material.

The first limitation is symbolic closure completeness. One symbolic obligation in the impossibility chain remains unresolved in the latest audit ledger. The empirical and theorem-level evidence strongly support the core representation-typing conclusion, but full symbolic closure for that obligation is still pending. We therefore keep the claim as strongly supported with an explicit formal caveat, rather than labeling it absolutely closed.

The second limitation is canonical validation packaging. Executed artifacts and tests exist, but the canonical phase-level validation snapshot remains stale in the upstream handoff. This affects traceability confidence, not the raw experimental measurements themselves. We keep this caveat explicit because manuscript reliability depends on both evidence quality and packaging integrity.

A further practical limitation is that project-level compute budgeting remains open-ended at governance level. The executed runs are CPU-feasible and reproducible, but broader scaling claims require explicit budget envelopes.

9.1 Future Work

Future work is direct and testable. First, close the remaining symbolic obligation by proving or appropriately narrowing the missing dominance-equivalence bridge condition, then regenerate theorem-audit closure tags. Second, refresh canonical validation packaging so claim-evidence links terminate in schema-complete artifacts. Third, extend bridge analysis to non-smooth preference regimes with explicit counterexample classification, so local equivalence conditions can be reported with sharper admissibility maps. Fourth, integrate stakeholder-structure metadata from group-decision surveys into the tuple-level caveat load, enabling context-aware method recommendation without violating representation boundaries.

10 Conclusion

This manuscript addressed a central survey challenge in multi-objective decision support: how to compare methods with non-equivalent output semantics without collapsing rigor into anecdotal ranking. We presented a representation-conditioned synthesis framework, defined its formal objective and feasible set, proved impossibility and constructive results, and linked those results to executed evidence with uncertainty reporting.

The main outcome is a defensible boundary: within-family quantitative ranking is valid under explicit assumptions, while cross-family synthesis must be evidence-graded rather than blindly scalarized. Empirical contradiction rates, constructive zero-loss behavior, and bounded bridge diagnostics jointly support this boundary. The resulting paper contributes not only a synthesis of methods but a reproducible logic for what can and cannot be claimed from heterogeneous multi-objective evidence.

REFERENCES

and Angelos Fountoulakis and Aikaterini Forouli. A robust augmented epsilon-constraint method (augmecon-r) for finding exact solutions of multi-objective linear p..., 2020. URL https://doi.org/10.1007/s12351-020-00574-6. Accessed and extracted via the project knowledge corpus.

Busra Ayan, Seda ABACIOGLU, and Marcio Pereira Basilio. A comprehensive review of the novel weighting methods for multi-criteria decision-making, 2023. URL https://doi.org/10.3390/info14050285. Accessed and extracted via the project knowledge corpus.

Majid Behzadian, Sina Khanmohammadi Otaghsara, and Morteza Yazdani. A state-of the-art survey of topsis applications, 2012. URL https://doi.org/10.1016/j.eswa.2012.05.056. Accessed and extracted via the project knowledge corpus.

David BoixCots, Francesc Pardo-Bosch, and Pablo Pujadas. A systematic review on multi-criteria group decision-making methods based on weights: Analysis and classifi..., 2023. URL https://doi.org/10.1016/j.inffus.2023.03.004. Accessed and extracted via the project knowledge corpus.

Jean Pierre Brans, Ph. Vincke, and Bertrand Mareschal. How to select and how to rank projects: The promethee method, 1986. URL https://doi.org/10.1016/0377-2217(86)90044-5. Accessed and extracted via the project knowledge corpus.

Mrittika Chakraborty, Wreetbhas Pal, and Sanghamitra Bandyopadhyay. A survey on multi-objective based parameter optimization for deep learning, 2023a. URL https://doi.org/10.7494/csci.2023.24.3.5479. Accessed and extracted via the project knowledge corpus.

Santonab Chakraborty, Santonab Chakraborty, and Rakesh D. Raut. A comprehensive and systematic review of multi-criteria decision-making methods and applications in healthcare, 2023b. URL https://doi.org/10.1016/j.health.2023.100232. Accessed and extracted via the project knowledge corpus.

Lineker Max Goulart Coelho, Lisete Celina Lange, and Hosmanny MG Coelho. Multi-criteria decision making to support waste management: A critical review of current practices and methods, 2016. URL https://doi.org/10.1177/0734242x16664024. Accessed and extracted via the project knowledge corpus.

Kalyanmoy Deb, Amrit Pratap, and Sakshi Agarwal. A fast and elitist multiobjective genetic algorithm: Nsga-ii, 2002. URL https://doi.org/10.1109/4235.996017. Accessed and extracted via the project knowledge corpus.

Sintayehu Legesse Gebre, Dirk Cattrysse, and Jos Van Orshoven. Multi-criteria decision-making methods to address water allocation problems: A systematic review, 2021. URL https://doi.org/10.3390/w13020125. Accessed and extracted via the project knowledge corpus.

Walter J. Gutjahr and Alois Pichler. Stochastic multi-objective optimization: a survey on non-scalarizing methods, 2013. URL https://doi.org/10.1007/s10479-013-1369-5. Accessed and extracted via the project knowledge corpus.

Zhenan He, Gary G. Yen, and Yi Zhang. Robust multiobjective optimization via evolutionary algorithms, 2018. URL https://doi.org/10.1109/tevc.2018.2859638. Accessed and extracted via the project knowledge corpus.

Roger M. Y. Ho and James P. Ignizio. Goal programming and extensions, 1977. URL https://doi.org/10.1057/jors.1977.91. Accessed and extracted via the project knowledge corpus.

Jonas Ide and Anita Schobel. Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts, 2015. URL https://doi.org/10.1007/s00291-015-0418-7. Accessed and extracted via the project knowledge corpus.

S.M. Kang, Kaiwen Li, and Rui Wang. A survey on pareto front learning for multi-objective optimization, 2024. URL https://doi.org/10.1007/s41965-024-00170-z. Accessed and extracted via the project knowledge corpus.

Bartomiej Kizielewicz, Tomasz Tomczyk, and Micha Gandor. Subjective weight determination methods in multi-criteria decision-making: a systematic review, 2024. URL https://doi.org/10.1016/j.procs.2024.09.673. Accessed and extracted via the project knowledge corpus.

Rahul Kumar and Dragan Pamucar. A comprehensive and systematic review of multi-criteria decision-making (mcdm) methods to solve decision-ma..., 2025. URL https://doi.org/10.31181/sdmap21202524. Accessed and extracted via the project knowledge corpus.

Haiping Ma, Yajing Zhang, and Shengyi Sun. A comprehensive survey on nsga-ii for multi-objective optimization and applications, 2023. URL https://doi.org/10.1007/s10462-023-10526-z. Accessed and extracted via the project knowledge corpus.

R. Timothy Marler and Jasbir S. Arora. Survey of multi-objective optimization methods for engineering, 2004. URL https://doi.org/10.1007/s00158-003-0368-6. Accessed and extracted via the project knowledge corpus.

George Mavrotas. Effective implementation of the epsilon-constraint method in multi-objective mathematical programming problems, 2009. URL https://doi.org/10.1016/j.amc.2009.03.037. Accessed and extracted via the project knowledge corpus.

George Mavrotas and Kostas Florios. An improved version of the augmented epsilon-constraint method (augmecon2) for finding the exact pareto set in mu..., 2013. URL https://doi.org/10.1016/j.amc.2013.03.002. Accessed and extracted via the project knowledge corpus.

Sanjay D. Pohekar and M Ramachandran. Application of multi-criteria decision making to sustainable energy planning-a review, 2004. URL https://doi.org/10.1016/j.rser.2003.12.007. Accessed and extracted via the project knowledge corpus.

Auref Rostamian, Matheus Bernardelli de Moraes, and Denis Jose Schiozer. A survey on multi-objective, model-based, oil and gas field development optimization: Current status and fu..., 2024. URL https://doi.org/10.1016/j.petsci.2024.09.018. Accessed and extracted via the project knowledge corpus.

Bernard Roy. The outranking approach and the foundations of electre methods, 1991. URL https://doi.org/10.1007/bf00134132. Accessed and extracted via the project knowledge corpus.

Anelka Stilic and Adis Puska. Integrating multi-criteria decision-making methods with sustainable engineering: A comprehensive review of ..., 2023. URL https://doi.org/10.3390/eng4020088. Accessed and extracted via the project knowledge corpus.

Jyrki Wallenius, James S. Dyer, and Peter C. Fishburn. Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead, 2008. URL https://doi.org/10.1287/mnsc.1070.0838. Accessed and extracted via the project knowledge corpus.

Jiangjiang Wang, Youyin Jing, and Chunfa Zhang. Review on multi-criteria decision analysis aid in sustainable energy decision-making, 2009. URL https://doi.org/10.1016/j.rser.2009.06.021. Accessed and extracted via the project knowledge corpus.

Zitong Wang, Yan Pei, and Jianqiang Li. A survey on search strategy of evolutionary multi-objective optimization algorithms, 2023. URL https://doi.org/10.3390/app13074643. Accessed and extracted via the project knowledge corpus.

Bin Xin, Lu Chen, and Jie Chen. Interactive multiobjective optimization: A review of the state-of-the-art, 2018. URL https://doi.org/10.1109/access.2018.2856832. Accessed and extracted via the project knowledge corpus.

Qingfu Zhang and Hui Li. Moea/d: A multiobjective evolutionary algorithm based on decomposition, 2007. URL https://doi.org/10.1109/tevc.2007.892759. Accessed and extracted via the project knowledge corpus.

Aimin Zhou, Boyang Qu, and Hui Li. Multiobjective evolutionary algorithms: A survey of the state of the art, 2011. URL https://doi.org/10.1016/j.swevo.2011.03.001. Accessed and extracted via the project knowledge corpus.

Eckart Zitzler and Lothar Thiele. Multiobjective optimization using evolutionary algorithms - a comparative case study, 1998. URL https://doi.org/10.1007/bfb0056872. Accessed and extracted via the project knowledge corpus.

Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results, 2000. URL https://doi.org/10.1162/106365600568202. Accessed and extracted via the project knowledge corpus.

Eckart Zitzler, Marco Laumanns, and Lothar Thiele. Spea2: Improving the strength pareto evolutionary algorithm, 2001. URL https://doi.org/10.3929/ethz-a-004284029. Accessed and extracted via the project knowledge corpus.

A Extended Proof Material

This appendix provides additional derivation detail for the formal claims in section 5. The main text proofs are complete; this appendix expands intermediate constructions so readers can directly verify assumptions and substitutions.

A.1 Detailed construction for Equation 5.3

For each family graph (V_r, E_r), choose a topological order τ_r. Define transformed coordinates by assigning each node $i \in V_r$ an embedding value $q_r(i) = \tau_r(i)/|V_r|$. Because the order is topological, for every edge (i, j) we have $q_r(j) - q_r(i) \geq 1/|V_r| > 0$. Let each margin function be $\Delta_{ij}(N) = q_r(j) - q_r(i)$ and set slacks to zero. Then constraints in equation 2 hold with strict slack. Since weights are strictly positive, any positive slack would increase equation 1; hence the zero-slack construction is globally optimal and not only feasible.

A.2 Boundary interpretation for Equation 5.1

The non-identifiability lemma does not assert that all cross-family orderings are equivalent; it asserts objective invariance under family-block permutations when no coupling term is present. If one introduces a coupling term such as

$$\Omega_{\text{bridge}} = \sum_{r \neq r'} \beta_{rr'}\, \phi\big(s_r(m) - s_{r'}(m)\big), \tag{9}$$

then the separability condition is violated and the lemma no longer applies. This is the precise formal location where bridge axioms enter the problem.

B Equation Provenance and Symbol Glossary

Table-like symbol descriptions are provided as prose to keep notation compact.

Decision variables are the monotone transforms $N_{r,k}$ and slacks ξ_{ij}. The feasible set in equation 3 combines monotonicity, acyclicity, and positive weighting assumptions. The objective in equation 1 is manuscript-defined; it is not copied from a single source. Pareto dominance and decomposition semantics informing family definitions are sourced from Deb et al. (2002) and Zhang & Li (2007). Outranking relation semantics originate in Roy (1991) and Brans et al. (1986). Goal and epsilon-style parameterized forms informing scalarization-family assumptions originate in Ho & Ignizio (1977), Mavrotas (2009), Mavrotas & Florios (2013), and and Angelos Fountoulakis & Forouli (2020).

C Reproducibility Details

All validation runs used seeded sweeps with five seeds per regime family. Representative seed sets were $\{11, 23, 37, 53, 71\}$ for contradiction witness analysis, $\{13, 29, 41, 59, 83\}$ for constructive identifiability checks, and $\{17, 31, 43, 67, 79\}$ for bridge-stability diagnostics. Sweep axes included objective count, incomparability density, family count, graph density, contradiction injection rate, bridge-axiom profile, perturbation radius, and regularity constants.

Uncertainty procedures combined percentile or BCa bootstrap intervals for rates and beta-binomial credible intervals for high-contradiction incidence. The uncertainty summary used in the main text reports typed violation mean 0.0083, family-blind violation mean 0.2863, and a high-contradiction credible interval approximately $[0.616, 0.686]$ for the blind regime. These are aggregate diagnostics, not replacements for regime-specific tables.

Operational constraints were CPU-only with bounded RAM and storage envelopes. Reported compute settings should be interpreted as reproducibility context and not as scientific novelty statements.

D Extended Diagnostics and Negative Results

Negative evidence is essential for claim integrity. The bridge-stability analysis includes explicit out-of-regime perturbation cases where sign-flip rates rise markedly. The theorem-audit pass-rate table in the main text already encodes this boundary effect; here we emphasize interpretation.

When perturbations exceed the local radius threshold, local equivalence assumptions fail and sign stability degrades. This does not contradict the local theorem; it validates its bounded scope. Likewise, if cyclic inconsistencies are injected into within-family claim graphs, constructive zero-loss guarantees are no longer expected because acyclicity assumptions are violated.

These diagnostics motivate the future-work actions listed in section 9: close the open symbolic obligation, refresh canonical validation packaging, and extend bridge analysis to non-smooth regimes with explicit admissibility maps.