

Article

Uncertainty-Aware Reconciliation of Conflicting Sources on the Probability Simplex via Choquet Aggregation over Fuzzy Measures

Ahmet Tezcan Tekin 

Department of Management Engineering, Istanbul Technical University, Istanbul 34367, Türkiye; tekina@itu.edu.tr

Abstract

We study the reconciliation of several biased, partially-overlapping estimators of a latent distribution on the probability simplex (the set of non-negative credit shares summing to one) when no ground truth is available to adjudicate their disagreement—the situation faced when conversion credit is reported by independent commercial attribution sources. We cast this as non-additive information fusion and develop a four-layer fuzzy framework: a reliability-weighted fuzzy representation, an interaction-corrected conflict measure, a Choquet integral with respect to an elicited fuzzy measure (capacity) followed by an L_1 -closure onto the simplex (a per-cell renormalization returning a valid distribution), and a Mamdani layer producing a conformally calibrated uncertainty band. Our theoretical contributions are (i) a simplex-consistency result—channel-wise Choquet aggregation with L_1 -closure always returns a valid distribution and reduces to the weighted mean under an additive capacity; (ii) variational and axiomatic characterizations of the operator; (iii) an exact Shapley–disagreement decomposition that quantifies how non-additivity reweights redundant sources; and (iv) a Lipschitz stability bound that guarantees robustness to noisy inputs. The capacity is elicited from observable signals alone, so the method requires no labels. On a firewalled, ten-seed benchmark across eight regimes the framework attains the lowest MAE in five of eight conditions and the lowest KL in six, improving on Dempster–Shafer, correlation-clustering, and reliability-weighted baselines by 1.6–2.3% (MAE) and 3.1–4.4% (KL)—significant in four of its five targeted redundancy- and bias-dominated regimes—with the conformal step attaining nominal coverage. This advantage stems from the elicited capacity rather than the Choquet operator (a simpler ordinal integral matches it on point error); Choquet is retained for its cardinal, calibratable output. An illustrative two-source real-data study exposes a 13.7% cost-per-install distortion in the single-source view, with the (uncalibrated) band widest where the sources disagree most.



Academic Editor: Shibo Wu

Received: 19 June 2026

Revised: 20 July 2026

Accepted: 22 July 2026

Published: 23 July 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: Choquet integral; fuzzy measure; information fusion; Mamdani inference; attribution; uncertainty reconciliation; non-additive aggregation; decision support

MSC: 28E10; 03E72; 68T37; 90B50; 94A17

1. Introduction

Combining several biased estimates of an unknown distribution into a single, trustworthy estimate is a recurring problem in information fusion. It becomes genuinely hard when the estimates come from self-interested, partially-overlapping sources and *no ground*

truth exists to adjudicate their disagreement: one can neither learn the sources' reliabilities from labelled outcomes nor fall back on a reference answer. This paper develops a non-additive fuzzy framework for exactly this setting and studies its mathematical properties; a canonical and economically significant instance motivates it throughout.

Digital advertising platforms increasingly act as both the seller and the scorekeeper of their own performance. In mobile user acquisition, an advertiser typically receives conversion credit from several sources at once—self-attributing networks such as Meta, AppLovin, and Google Ads, alongside third-party mobile measurement partners—each applying its own attribution rule, observation window, and visibility into the user journey. The resulting credit figures disagree, often substantially, and the advertiser has no neutral reference against which to decide who is right; field-experiment evidence shows that platform-reported attribution can diverge markedly from incremental effect [1,2], a gap widened by privacy-driven measurement shifts such as the deprecation of third-party identifiers and the move to aggregated, on-device attribution [3].

The standard responses are to treat one source as authoritative or to average across sources. Both are unsatisfactory. Trusting a single self-attributing source inherits its self-preferencing bias, while naive averaging silently *double-counts* sources that share the same blind spots and rules: correlated sources agree for structural reasons rather than because they are correct. What is missing is a principled way to (a) characterize the disagreement, (b) aggregate the sources while accounting for their redundancy, and (c) report how much confidence the reconciled figure deserves—all *without* access to ground truth, which on real attribution data is unobservable in principle.

This paper casts the problem as information fusion under absent ground truth and develops a four-layer fuzzy framework that produces, per channel-day, a reconciled credit distribution together with an interpretable, calibrated uncertainty band and disagreement diagnostics. Our contributions are as follows:

1. **A formulation** of cross-source attribution disagreement as *epistemic uncertainty* to be reconciled, rather than a label to be predicted; no ground truth is assumed at any point in the estimator.
2. **A non-additive aggregation** of sources via a Choquet integral with respect to a fuzzy measure, whose interaction indices encode source redundancy, paired with an L_1 -closure, together with its mathematical analysis: a simplex-consistency result (Proposition 1); a variational characterization showing the operator is the robust aggregation over a redundancy-constrained set of source weightings—the support function of the capacity's base polytope (Theorem 2); an axiomatic characterization pinning it down as the unique aggregator satisfying monotonicity, co-monotonic additivity, and capacity calibration (Theorem 3); an exact Shapley–disagreement decomposition (Proposition 6); a Lipschitz stability bound (Theorem 1); and an impossibility result delimiting what any output-only rule can identify (Proposition 7). The variational and axiomatic characterizations build on classical results—the Lovász extension of a capacity and Schmeidler's co-monotonic-additivity representation—which we do not claim as new; the contribution is their identification for *this label-free reconciliation operator*, together with the L_1 -closure that makes the Choquet output a valid distribution. Table 1 and the Remark in Section 4.9 make this inherited-versus-novel boundary explicit.
3. **An observable capacity elicitation**—reliability from coverage and reporting rate, an adversary gate based on consensus correlation, and a coverage-Jaccard “shared-lens” redundancy—that requires no ground truth and therefore transfers to real data.
4. **A Mamdani confidence layer** with conformal calibration that turns the qualitative disagreement signal into an uncertainty band with a finite-sample, distribution-free

coverage guarantee (Proposition 8), governed by a single invariant: agreement among *independent* sources raises confidence, agreement among *redundant* sources does not.

5. **A firewalled, multi-seed synthetic methodology** that guards against circular self-confirmation, on which the framework attains the lowest MAE in five of eight conditions (six on KL), significantly so in the redundancy- and bias-dominated regimes it targets, with an ablation isolating the contribution of non-additivity.

Table 1. Comparison of the proposed framework with major existing reconciliation and fusion approaches. The proposed framework is shown in the last row (“This work”), in bold.

Approach	Ground Truth Required	Redundancy Handling	Non-Additive	Simplex Output	Calibrated Uncertainty	Label-Free Capacity
Supervised/ data-driven MTA (Shapley, Markov)	Yes (labels/ experiment)	via trained model	No	No	No	—
Reliability-weighted mean	No	additive (double-counts)	No	Yes (closure)	No	Partial
Dempster-Shafer discounting	No (reliabilities given)	conflict rule	No	Yes (pignistic)	No	No
Correlation- clustering	No	explicit source clustering	No	Yes	No	No
Classical Choquet/ fuzzy-integral fusion	Yes (learned capacity)	interaction indices	Yes	varies	No	No
This work	No	elicited sub-additive capacity	Yes	Yes	Yes (conformal)	Yes

The central contribution is thus the *problem formulation* and the *observable, label-free capacity elicitation*, together with the consistency, characterization, and decomposition results that turn an elicited capacity into a valid simplex output; the Choquet operator is retained—and named in the title—because it alone exposes the cardinal Shapley/interaction diagnostics, supports the conformal band, and admits the variational reading of Theorem 2, even though a simpler ordinal integral matches its point error (Section 7.6).

We deliberately separate this work from fuzzy *predictive* ensembling. Unlike approaches that partition a feature space and learn membership-weighted ensembles of trained models, here fuzziness models the epistemic uncertainty of credit itself, the ensemble members are external commercial sources we do not train, and there is no predictive label to optimize. The fuzzy machinery descends from Mamdani inference [4] and fuzzy rule-based systems rather than from fuzzy clustering. It also differs from our earlier fuzzy-rule and Mamdani frameworks [5,6], which target prediction and post hoc explanation with labelled data, whereas here the same rule-based machinery operates entirely without ground truth.

Related Work

Attribution modeling. Multi-touch attribution (MTA) assigns conversion credit across touchpoints using rule-based heuristics, game-theoretic Shapley values [7,8], Markov-chain removal effects [9], or data-driven and causal models [10–12]. Berman [13] analyzes the inefficiencies of last-touch crediting under externalities and exposure uncertainty, and Kannan et al. [14] survey the path-to-purchase view. These works estimate credit for a *single* observer with access to the conversion path and, implicitly, a recoverable ground truth (a trained model or an experiment); none reconciles several mutually biased external sources.

Attribution discrepancy and incrementality. A parallel line measures the gap between attributed and *incremental* effect through field experiments [1,2,15,16] and triangulation across marketing-mix models, attribution, and incrementality [3,17]. This literature docu-

ments the discrepancy we target but provides no aggregation operator that reconciles it under absent ground truth.

Fuzzy measures, Choquet integration, and aggregation operators. The Choquet integral [18,19] with respect to a fuzzy measure (capacity) [20–22] is a standard non-additive fusion tool whose interaction indices [23,24] capture redundancy and complementarity; it underpins multi-source and classifier fusion [25–28] and has been generalized in several directions [29]. Related operators include the ordered weighted average [30–33] and the Sugeno integral [34,35], and the Choquet integral is widely applied in multi-criteria decision making [36]. Crucially, these are applied where labels exist (classifier fusion) or are learned from data—not under the explicit absence of ground truth.

Source reliability, evidence theory, and consensus. Dempster–Shafer theory [37,38] fuses unreliable sources via reliability discounting [39–43], with conflict measures bridging to aggregation operators [44–46]. Group decision making with fuzzy preference relations seeks *consensus* among experts [47–50]. These reconcile preferences toward a consensus or fuse mass functions with given or learned reliabilities; none reconciles simplex-valued *credit* toward a latent truth that is never observed.

Explainable fuzzy systems. Mamdani inference and rule-based fuzzy systems are valued for interpretability [51–56], including adaptive neuro-fuzzy variants [57]; closest in spirit are interpretable Mamdani frameworks with audit trails [5] and fuzzy rule-based explanations carrying rule-level uncertainty [6], whose interpretable rule design we adapt for the confidence layer. Our contribution sits at the intersection none of these occupy: cross-source attribution disagreement modeled as epistemic uncertainty and reconciled by non-additive fuzzy aggregation under absent ground truth.

Table 1 positions the proposed framework against the major existing approaches, in particular the Choquet-based aggregation methods closest to ours. Supervised and data-driven multi-touch attribution recovers credit for a single observer with access to a label or experiment; Dempster–Shafer discounting and correlation-clustering fuse unreliable sources but treat redundancy either through a conflict rule or an explicit clustering step; and classical Choquet/fuzzy-integral fusion learns the capacity from labelled data. What distinguishes the present work is the combination the last column makes explicit: it is the only approach that jointly (a) operates with no ground truth, (b) discounts redundant agreement through a capacity *elicited from observable signals alone*, (c) returns a valid simplex-valued distribution, and (d) attaches a conformally calibrated uncertainty band. It is this label-free, redundancy-aware integration—not any individual mathematical tool—that constitutes the contribution.

The remainder of the paper is organized as follows. Section 2 collects the mathematical preliminaries. Section 3 formalizes the problem. Section 4 develops the four-layer framework, including the simplex-consistency result, the capacity elicitation, the algorithm, and its complexity. Section 5 describes the synthetic validation design and Section 6 the experimental setup. Section 7 reports results, Section 8 discusses them, and Section 9 concludes.

2. Preliminaries

We collect the notions used throughout. Let $[m] = \{1, \dots, m\}$ index the sources and $[K] = \{1, \dots, K\}$ the channels. Vectors are columns; $\mathbf{1}$ is the all-ones vector. The probability simplex on K outcomes is $\Delta^{K-1} = \{x \in \mathbb{R}_{\geq 0}^K : \sum_k x_k = 1\}$.

2.1. Fuzzy Sets and Fuzzy Numbers

A fuzzy set A on a universe U is characterized by a membership function $\mu_A : U \rightarrow [0, 1]$ [58]. A *triangular fuzzy number* $\tilde{a} = (a - \delta, a, a + \delta)$ has membership peaking at the centre a and decaying linearly to zero at $a \pm \delta$; the spread δ encodes imprecise-

cision [59]. We use triangular numbers to attach a reliability- dependent spread to each reported credit (Section 4.2).

2.2. Fuzzy Measures and the Choquet Integral

A *fuzzy measure* (capacity) on $[m]$ is a set function $\mu : 2^{[m]} \rightarrow [0, 1]$ that is grounded ($\mu(\emptyset) = 0$), normalized ($\mu([m]) = 1$), and monotone ($A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$) [19,20]. Unlike an additive measure, $\mu(A \cup B)$ need not equal $\mu(A) + \mu(B)$: sub-additivity ($<$) models *redundancy* between coalitions, super-additivity ($>$) models *complementarity*.

For a non-negative vector $x \in \mathbb{R}_{\geq 0}^m$, sort its entries increasingly $x_{(1)} \leq \dots \leq x_{(m)}$ with $x_{(0)} = 0$, and let $A_{(i)} = \{(i), \dots, (m)\}$. The discrete *Choquet integral* of x with respect to μ is

$$C_\mu(x) = \sum_{i=1}^m (x_{(i)} - x_{(i-1)}) \mu(A_{(i)}). \quad (1)$$

The Choquet integral is idempotent, monotone, and continuous; when μ is additive it equals the weighted arithmetic mean $\sum_s \mu(\{s\}) x_s$ [18].

A second fuzzy integral with respect to the same capacity is the *Sugeno integral* [19], an ordinal (max–min) aggregation

$$S_\mu(x) = \max_{i=1, \dots, m} \min(x_{(i)}, \mu(A_{(i)})). \quad (2)$$

The Sugeno integral is also idempotent and monotone, but it is *ordinal*: it commutes with increasing bijections of the scale and does not reduce to a weighted arithmetic mean for an additive capacity, nor does it expose cardinal Shapley /interaction indices. We use the Choquet integral as the framework's primary operator for these cardinal properties (which the calibrated confidence layer relies on) and compare the two operators empirically in Section 7.6.

2.3. Möbius Transform, Shapley Value, and Interaction Indices

Every capacity admits a *Möbius representation* $m : 2^{[m]} \rightarrow \mathbb{R}$ defined by $m(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} \mu(B)$, with inverse $\mu(A) = \sum_{B \subseteq A} m(B)$. A capacity is *k-additive* [21] if $m(A) = 0$ for all $|A| > k$; a 2-additive capacity is fully specified by the singletons $m(\{s\})$ and pairs $m(\{s, t\})$. For a 2-additive capacity the discrete Choquet integral has the closed form

$$C_\mu(x) = \sum_s m(\{s\}) x_s + \sum_{s < t} m(\{s, t\}) \min(x_s, x_t). \quad (3)$$

The *Shapley value* ϕ_s summarizes the global importance of source s , and the *interaction index* I_{st} quantifies redundancy ($I_{st} < 0$) or complementarity ($I_{st} > 0$) between s and t [23]. For a 2-additive capacity, these reduce to $\phi_s = m(\{s\}) + \frac{1}{2} \sum_{t \neq s} m(\{s, t\})$ and $I_{st} = m(\{s, t\})$ [24].

2.4. Mamdani Fuzzy Inference

A Mamdani fuzzy inference system maps crisp inputs to a crisp output through fuzzification, a base of IF–THEN rules with fuzzy antecedents and consequents, min-implication, max-aggregation, and centroid defuzzification [4]. It is preferred over Takagi–Sugeno systems when interpretability of the rule base is paramount [51,52].

2.5. Conformal Calibration

Split-conformal prediction converts any heuristic uncertainty score into intervals with distribution-free coverage guarantees [60,61]. Given a calibration set and a non-conformity score, the $(1 - \alpha)$ -quantile of the calibration scores yields a multiplier that makes the test-

set coverage attain the nominal level $1 - \alpha$. We use this to calibrate the Mamdani band (Section 4.6).

Notation is summarized in Table 2.

Table 2. Notation used throughout the paper.

Symbol	Meaning
$s \in [m], k \in [K]$	source index, channel index
$\hat{c}_s \in \Delta^{K-1}$	credit vector reported by source s
$c^* \in \Delta^{K-1}$	true (unobservable) credit vector
$\hat{c}$	reconciled credit vector (framework output)
$r_s \in (0, 1]$	elicited reliability of source s
$R_{st} \in [0, 1]$	coverage-Jaccard redundancy of s, t
$\mu, m(\cdot), I_{st}$	capacity, its Möbius coefficients, interaction index
$C_\mu(\cdot)$	discrete Choquet integral
$V_s \subseteq [K]$	visibility set of source s (latent; not observed by the method)

3. Problem Formulation

Let $k \in [K]$ index marketing channels and $s \in [m]$ attribution sources. For a given *cell* (a channel-day, or a conversion in the synthetic study), each present source reports a credit vector $\hat{c}_s = (\hat{c}_{s,1}, \dots, \hat{c}_{s,K}) \in \Delta^{K-1}$. There exists an unobservable true credit vector $c^* \in \Delta^{K-1}$ on the same simplex. Three properties make the setting distinctive:

- **No ground truth.** c^* is never available on real data; it exists only inside the synthetic generator (Section 5), where it is used solely for evaluation, never by the estimator.
- **Partial observability.** Each source sees only a latent subset of channels $V_s \subseteq [K]$ and redistributes credit among what it sees; V_s is *not* given to the method.
- **Heterogeneous bias.** Sources apply different attribution rules, windows, and self-preferencing, so their disagreement is structured, not merely noisy.

We make these precise as working assumptions:

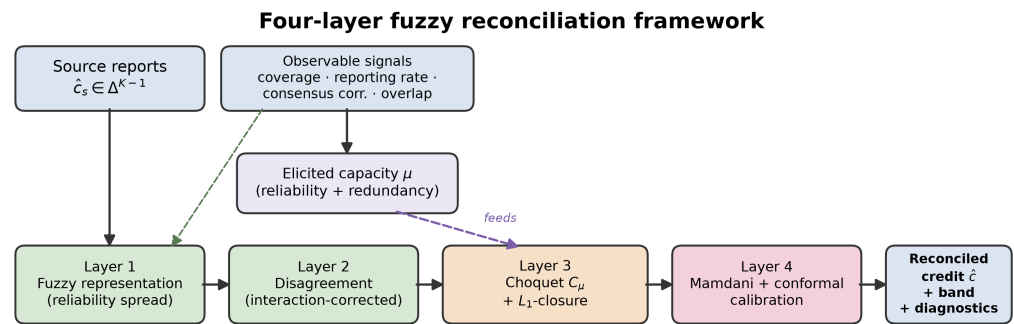
- **(A1) Simplex reports.** Every present source emits $\hat{c}_s \in \Delta^{K-1}$; credit is a share, not an unbounded magnitude.
- **(A2) No ground truth.** c^* is unobservable on real data and is never read by the estimator; it is used only to score recovery on synthetic data.
- **(A3) Latent, unequal lenses.** Each source sees a latent set V_s and applies a latent rule/window/self-preference; neither V_s nor the rule is given to the method.
- **(A4) Observable reliability signals.** The method may use only quantities computable from the reports themselves—coverage, reporting rate, consensus correlation, and coverage overlap—to elicit reliability and redundancy.

The task is to produce, per cell, a reconciled estimate $\hat{c} \in \Delta^{K-1}$, an uncertainty band, and disagreement diagnostics, using the reported $\{\hat{c}_s\}$ and the observable signals of (A4) alone. We measure quality on the synthetic study by recovery of c^* ; on real data only plausibility, diagnostics, and decision impact are claimed.

4. The Reconciliation Framework

4.1. Overview

Figure 1 summarizes the four-layer pipeline. Sources are fuzzified by reliability (Layer 1); their disagreement is characterized by an interaction-corrected conflict measure (Layer 2); a Choquet integral over an elicited capacity aggregates them non-additively and an L_1 -closure returns a valid distribution (Layer 3); and a Mamdani layer with conformal calibration produces an interpretable uncertainty band (Layer 4).



Governing invariant: agreement among independent sources raises confidence; agreement among redundant sources does not.

Figure 1. The four-layer fuzzy reconciliation framework: sources are fuzzified by reliability, their disagreement is characterized by an interaction-corrected conflict measure, a Choquet integral over a fuzzy measure aggregates them with an L_1 -closure, and a Mamdani layer produces an interpretable confidence band. The diagram reads left to right as a flowchart: the source reports and the observable elicitation signals enter Layer 1; each layer's output feeds the next; and the pipeline emits the reconciled credit, the calibrated uncertainty band, and the Shapley/interaction diagnostics.

4.2. Layer 1—Fuzzy Representation

Each source's credit is represented as a triangular fuzzy number whose spread encodes that source's *local* reliability r_s (Section 4.5): for channel k ,

$$\tilde{c}_{s,k} = (\hat{c}_{s,k} - \delta_{s,k}, \hat{c}_{s,k}, \hat{c}_{s,k} + \delta_{s,k}), \quad \delta_{s,k} = \zeta (1 - r_s) \hat{c}_{s,k}, \quad (4)$$

with $\zeta \in (0, 1]$ a global scale. Higher reliability yields a narrower spread, so unreliable sources contribute diffuse evidence. The spreads feed the conflict measure and the band width.

4.3. Layer 2—Disagreement Characterization

We quantify disagreement per cell by an *interaction-corrected* dispersion of the present sources. Writing $w_{st} = (1 - R_{st})$ for the independence weight of a pair (Section 4.5), the conflict is

$$\text{conf} = \frac{\sum_{s < t} w_{st} \|\hat{c}_s - \hat{c}_t\|_1}{\sum_{s < t} w_{st}} \in [0, 2]. \quad (5)$$

The weighting enforces the governing invariant (Section 4.6): disagreement among *independent* sources counts as genuine conflict, while agreement among *redundant* sources is discounted because it is not genuine consensus. Equation (5) is both a reported diagnostic and the disagreement input to the Mamdani layer.

4.4. Layer 3—Non-Additive Aggregation via the Choquet Integral

Let μ be a capacity over $[m]$ (Section 4.5). For a fixed channel k , we apply the discrete Choquet integral (1) and (3) to the source credits $\hat{c}_{\cdot,k}$. Because (3) is applied channel-wise, the raw aggregate need not lie on the simplex; we therefore apply an L_1 -closure,

$$\hat{c}_k = \frac{C_\mu(\hat{c}_{\cdot,k})}{\sum_{k'} C_\mu(\hat{c}_{\cdot,k'})}. \quad (6)$$

We choose L_1 -closure over a log-ratio (additive log-ratio/centred log-ratio) transform from compositional data analysis [62] deliberately: partial observability produces many structural zeros in the credit vectors, on which log-ratios are undefined and would force ad hoc zero-imputation; moreover a log-ratio Choquet reduces to a geometric mean, obscuring the redundancy-discounting interpretation. When a source is absent from a cell it is masked

out; the per-cell normalization in (6) cancels any constant scaling, so missing sources need no imputation.

Proposition 1 (Simplex consistency). *Channel-wise Choquet aggregation (3) followed by the L_1 -closure (6) always yields a valid credit distribution: $\hat{c}_k \geq 0$ for all k and $\sum_k \hat{c}_k = 1$. If μ is additive with normalized weights $w_s = \mu(\{s\})$, $\sum_s w_s = 1$, then $\sum_k C_\mu(\hat{c}_{\cdot,k}) = 1$, the closure (6) is the identity, and the reconciled vector reduces to the weighted arithmetic mean $\hat{c}_k = \sum_s w_s \hat{c}_{s,k}$. Consequently, the closure corrects only the non-additive mass defect and, being a single global rescale per cell, preserves the relative inter-source weighting induced by the interaction indices.*

Proof. *Non-negativity.* Each $\hat{c}_{s,k} \geq 0$ and μ is monotone, so in (1), every increment $x_{(i)} - x_{(i-1)} \geq 0$ is multiplied by $\mu(A_{(i)}) \geq 0$; hence $C_\mu(\hat{c}_{\cdot,k}) \geq 0$. The closure divides by a strictly positive sum (a cell has at least one present source with positive credit), so $\hat{c}_k \geq 0$, and by construction $\sum_k \hat{c}_k = 1$. *Additive reduction.* For an additive capacity, the Choquet integral equals the weighted mean, $C_\mu(\hat{c}_{\cdot,k}) = \sum_s w_s \hat{c}_{s,k}$. Summing over channels and exchanging the order of summation,

$$\sum_{k=1}^K C_\mu(\hat{c}_{\cdot,k}) = \sum_{k=1}^K \sum_{s=1}^m w_s \hat{c}_{s,k} = \sum_{s=1}^m w_s \sum_{k=1}^K \hat{c}_{s,k} = \sum_{s=1}^m w_s = 1,$$

since each source's credit vector sums to one. Thus, the denominator in (6) is unity, the closure is the identity, and $\hat{c}_k = \sum_s w_s \hat{c}_{s,k}$. *Idempotence.* If every present source reports the same distribution, the Choquet integral returns it unchanged (idempotence of C_μ), the channel sum is one, and the closure leaves it fixed. Finally, since (6) divides every channel of a cell by the same positive constant, the relative magnitudes produced by the interaction indices are preserved. $\square$

4.5. Capacity Elicitation from Observable Signals

The capacity is *elicited*—built from observable source outputs only, with no ground truth—so it transfers to real data. Three observable signals are combined.

Reliability (singletons). Source reliability combines channel coverage and reporting rate,

$$r_s^0 = \frac{1}{2} \frac{|\{k : \exists \text{ cell with } \hat{c}_{s,k} > 0\}|}{K} + \frac{1}{2} \frac{\#\text{cells reported by } s}{\#\text{cells}}. \quad (7)$$

Adversary gate. A source whose per-cell credit is near-uncorrelated with the leave-one-out consensus of the others is erratic/uninformative. Let ρ_s be the Pearson correlation between s 's credits and that consensus over all (cell, channel) pairs where both are present. The gate multiplies reliability down only for near-zero correlation:

$$r_s = r_s^0 \cdot \text{clip}(\rho_s / \tau, \underline{g}, 1), \quad (8)$$

with threshold $\tau = 0.15$ and floor $\underline{g} = 0.3$. A systematically biased but journey-tracking source has high ρ_s and is untouched; only a random adversary is discounted. When ρ_s is undefined—a source whose support is disjoint from the others, or that has fewer than five shared (cell, channel) pairs or zero variance there—we set $\rho_s = 1$, so the gate leaves that source's reliability unchanged (no discount). This is the conservative choice: a source observed only on its own private channels cannot be judged adversarial from outputs alone, the formal counterpart of the indistinguishability in Proposition 7.

Redundancy and interaction (pairs). Two sources are redundant when they look through the same *lens*, not merely when their outputs agree. We therefore use a coverage-Jaccard redundancy,

$$R_{st} = \text{Jaccard}(\text{cover}_s, \text{cover}_t) \cdot \left(\frac{1}{2} + \frac{1}{2} \text{co-report}_{st} \right), \quad (9)$$

where cover_s is the set of channels s ever credits and co-report_{st} is the fraction of jointly-reported cells. The pairwise Möbius coefficient is sub-additive, $m(\{s, t\}) = -\beta R_{st} \sqrt{r_s r_t}$, with strength β , and $m(\{s\}) \propto r_s$. The coefficients are shrunk by a global factor until the resulting capacity is monotone (verified exactly over all subsets, feasible for the small m typical of attribution) and then normalized so $\mu([m]) = 1$. We use a 2-additive capacity throughout; a 3-additive extension (triple Möbius terms for same-lens triples) was implemented and evaluated, and did not improve recovery (Section 8).

Three identification tiers are distinguished. The *elicited* capacity above is the only one available on real data and carries every headline result. A *data-driven* capacity and an *oracle* capacity fitted against c^* are synthetic-only; the oracle is reported solely as an upper bound.

4.6. Layer 4—Mamdani Confidence and Conformal Calibration

A Mamdani inference system maps the normalized disagreement $d = \text{conf}/2 \in [0, 1]$ and the mean reliability $\bar{r}$ of the reporting sources to a confidence $\kappa \in [0, 1]$. Each input uses three triangular membership functions (low/medium/high); the rule base encodes the invariant

Agreement among independent sources raises confidence; agreement among redundant sources does not.

So that low disagreement with high reliability yields high confidence, high disagreement yields low confidence, and low reliability caps confidence. We use min-implication, max-aggregation, and centroid defuzzification (Section 2.4); the inference is vectorized over all cells. A raw band of half-width $h = (1 - \kappa) h_0$ is produced.

Because a Mamdani confidence is not a calibrated interval, we apply split-conformal calibration (Section 2.5): on a held-out calibration split, we compute, per channel, the non-conformity score $|\hat{c}_k - c_k^*|/(h + \epsilon)$, take its $(1 - \alpha)$ -quantile q_α , and report the band $\hat{c}_k \pm q_\alpha h$ on the test split. We report both raw ($q = 1$) and calibrated coverage. Proposition 8 states the finite-sample coverage guarantee for exactly this construction (the multiplicative, locally adaptive band with the per-(cell, channel) ratio score).

4.7. Algorithm and Complexity

Algorithm 1 summarizes one forward pass.

Algorithm 1 Reconcile ($\{\hat{c}_s\}$ over cells)

- 1: Build tensor $X \in \mathbb{R}^{N \times m \times K}$ and presence mask $P \in \{0, 1\}^{N \times m}$.
 - 2: Elicit reliability r by (7)–(8), redundancy R by (9), and Möbius coefficients $m(\cdot)$.
 - 3: For each channel k : compute $C_\mu(\hat{c}_{\cdot, k})$ by (3) with present-source masking.
 - 4: L_1 -close per cell by (6) to obtain $\hat{c}$.
 - 5: Compute conflict (5); run Mamdani inference for κ ; form the raw band h .
 - 6: (Synthetic) conformal-calibrate h to nominal coverage.
 - 7: Return $\hat{c}$, band, conflict, confidence, and the Shapley/interaction diagnostics.
-

Complexity. Eliciting r and R is $O(NmK + m^2K)$. The monotonicity check over subsets is $O(2^m)$, negligible for the small source counts ($m \leq 7$) typical of attribution. The 2-additive Choquet step is $O(Nm^2K)$. Thus, the framework is *linear in the number of cells*

and inexpensive. Empirically, a single-thread forward pass over a 7000-cell condition takes about 24~ms for the Choquet aggregation (capacity elicitation included) and under 120~ms for every baseline, so reconciliation is never the bottleneck.

4.8. Worked Example

A two-channel cell makes the non-additive correction concrete. Let $K = 2$ channels $\{A, B\}$ and $m = 3$ sources, where s_1 and s_2 are a redundant pair (same lens) and s_3 is independent:

$$\hat{c}_{s_1} = (0.8, 0.2), \quad \hat{c}_{s_2} = (0.8, 0.2), \quad \hat{c}_{s_3} = (0.4, 0.6).$$

The naive mean is $(0.667, 0.333)$: the duplicated pair pulls channel A toward 0.8, i.e., it is *double-counted*. Suppose the elicited 2-additive capacity has equal singleton Möbius mass $m(\{s\}) = 0.4$ and a sub-additive interaction on the redundant pair $m(\{s_1, s_2\}) = -0.2$ (all other coefficients zero), which is normalized since $3(0.4) - 0.2 = 1$. Using the closed form (3), only the $\{s_1, s_2\}$ pair contributes a min term, so for each channel $C_\mu(\hat{c}_{\cdot,k}) = 0.4(\hat{c}_{1,k} + \hat{c}_{2,k} + \hat{c}_{3,k}) - 0.2 \min(\hat{c}_{1,k}, \hat{c}_{2,k})$:

$$C_A = 0.4(2.0) - 0.2(0.8) = 0.64, \quad C_B = 0.4(1.0) - 0.2(0.2) = 0.36.$$

Here, $C_A + C_B = 1$, so the L_1 -closure (6) is the identity and the reconciled credit is $(0.64, 0.36)$. The framework has pulled channel A down from the naive 0.667 toward the redundancy-corrected value 0.6—the average of the two *distinct* viewpoints $\frac{1}{2}(0.8 + 0.4)$ —without fully collapsing the pair. Setting $m(\{s_1, s_2\}) = 0$ (an additive capacity, $m(\{s\}) = \frac{1}{3}$) recovers exactly the naive mean $(0.667, 0.333)$, illustrating Proposition 1.

4.9. Further Theoretical Properties

Beyond simplex consistency (Proposition 1), the pipeline enjoys several properties that make its behaviour predictable and auditable. Throughout, μ is a monotone, normalized capacity with 2-additive Möbius coefficients $m(\cdot)$.

Lemma 1 (Conflict bounds). *For credit vectors on Δ^{K-1} , the interaction-corrected conflict (5) satisfies $\text{conf} \in [0, 2]$. Moreover, $\text{conf} = 0$ if and only if every pair of present sources with positive weight $w_{st} > 0$ reports the identical credit vector.*

Proof. For $u, v \in \Delta^{K-1}$, one has $0 \leq \|u - v\|_1 \leq 2$, the upper bound attained by disjoint supports. Equation (5) is a convex combination of these pairwise distances with weights $w_{st} = (1 - R_{st}) \geq 0$, hence it lies in $[0, 2]$. As a non-negative weighted mean of non-negative terms, it vanishes exactly when every term with $w_{st} > 0$ is zero, i.e., $\hat{c}_s = \hat{c}_t$ for all such pairs. $\square$

Proposition 2 (Monotonicity of the elicited capacity). *A 2-additive capacity with Möbius m is monotone if and only if for every s and every $A \subseteq [m] \setminus \{s\}$, $m(\{s\}) + \sum_{t \in A} m(\{s, t\}) \geq 0$. The sufficient condition $m(\{s\}) + \sum_{t \neq s} \min(m(\{s, t\}), 0) \geq 0$ implies monotonicity, and the construction of Section 4.5 enforces it by uniform shrinkage of the interaction terms.*

Proof. Monotonicity is equivalent to the non-negativity of all first-order derivatives $\Delta_s \mu(A) = \mu(A \cup \{s\}) - \mu(A)$ for $s \notin A$. For a 2-additive capacity, $\Delta_s \mu(A) = m(\{s\}) + \sum_{t \in A} m(\{s, t\})$, which is affine in the indicator of A and is minimized by collecting exactly the partners with negative interaction, $A^* = \{t \neq s : m(\{s, t\}) < 0\}$. Hence, non-negativity for all A is equivalent to the stated inequality, and the displayed sufficient condition is the value

at A^* . Scaling all $m(\{s, t\})$ by a common factor $\lambda \in [0, 1]$ relaxes the binding constraint monotonically, so a finite number of shrinkage steps attains it [21]. $\square$

Proposition 3 (Idempotence of the pipeline). *If every present source reports the same distribution $c \in \Delta^{K-1}$, then the reconciled output is $\hat{c} = c$ for any monotone normalized capacity.*

Proof. Fix a channel k ; all present sources supply the value c_k , so by idempotence of the Choquet integral $C_\mu(c_k \mathbf{1}) = c_k \mu([m]_{\text{present}})$, a constant multiple of c_k independent of k . The L_1 -closure (6) divides by the channel sum and cancels this constant, returning c_k . Hence, $\hat{c} = c$. $\square$

Proposition 4 (Masking invariance). *The reconciled distribution is invariant to multiplying a cell's channel-wise Choquet outputs by any positive constant. Consequently, restricting and renormalizing the capacity to the present sources does not change $\hat{c}$, so missing sources require no imputation.*

Proof. Writing $C = (C_1, \dots, C_K)$ for the channel-wise outputs, the closure (6) returns $C / \sum_k C_k$, which is unchanged under $C \mapsto \alpha C$ for any $\alpha > 0$. Restricting the Möbius coefficients to the present sources and renormalizing scales every C_k by the same $\alpha = 1 / \mu([m]_{\text{present}})$, leaving $\hat{c}$ fixed. $\square$

Proposition 5 (Adversary-gate boundedness). *The gate factor $g(\rho_s) = \text{clip}(\rho_s / \tau, \underline{g}, 1)$ is non-decreasing in the consensus correlation ρ_s , equals 1 whenever $\rho_s \geq \tau$, and is bounded below by $\underline{g}$. Hence, a source that tracks the consensus ($\rho_s \geq \tau$) retains its full elicited reliability, while the most an erratic source can be discounted is the factor $\underline{g}$.*

Proof. Immediate from the monotonicity of $\text{clip}(\cdot, \underline{g}, 1)$ and the evaluation $g(\rho_s) = 1$ for $\rho_s \geq \tau$, $g(\rho_s) = \underline{g}$ for $\rho_s \leq \underline{g}\tau$. $\square$

The next result makes the mechanism of non-additivity exact: it decomposes the channel-wise aggregate into a Shapley-weighted mean and an explicit redundancy-disagreement correction.

Proposition 6 (Shapley–disagreement decomposition). *For a 2-additive capacity with Shapley values $\phi_s = m(\{s\}) + \frac{1}{2} \sum_{t \neq s} m(\{s, t\})$ and interaction indices $I_{st} = m(\{s, t\})$, the channel-wise Choquet aggregate admits the exact identity*

$$C_\mu(x) = \sum_s \phi_s x_s - \frac{1}{2} \sum_{s < t} I_{st} |x_s - x_t|. \quad (10)$$

In particular, since $\sum_s \phi_s = \mu([m]) = 1$, the aggregate is a weighted mean corrected by a term that, for sub-additive (redundant) pairs $I_{st} < 0$, penalizes the cells on which those sources disagree—precisely the redundant double-counting the framework is designed to remove.

Proof. From the closed form (3), $C_\mu(x) - \sum_s \phi_s x_s = \sum_{s < t} I_{st} \min(x_s, x_t) - \frac{1}{2} \sum_{s < t} I_{st} (x_s + x_t)$, using $\sum_s \phi_s x_s = \sum_s m(\{s\}) x_s + \frac{1}{2} \sum_{s < t} I_{st} (x_s + x_t)$. Since $\min(a, b) - \frac{a+b}{2} = -\frac{1}{2}|a - b|$, the right-hand side equals $-\frac{1}{2} \sum_{s < t} I_{st} |x_s - x_t|$. $\square$

Finally, the reconciliation map is stable: bounded perturbations of the source reports—the daily noise of attribution feeds—produce proportionally bounded changes in the reconciled credit.

Theorem 1 (Lipschitz stability). *The discrete Choquet integral is 1-Lipschitz in the sup-norm, $|C_\mu(x) - C_\mu(x')| \leq \|x - x'\|_\infty$. Consequently, if two source profiles $\{\hat{c}_s\}, \{\hat{c}'_s\}$ yield channel-wise aggregates with total mass $Z, Z' \geq \delta > 0$, the reconciled distributions satisfy*

$$\|\hat{c} - \hat{c}'\|_1 \leq \frac{2K}{\delta} \max_{s,k} |\hat{c}_{s,k} - \hat{c}'_{s,k}|. \quad (11)$$

Proof. The 1-Lipschitz property of C_μ for a normalized capacity is classical [20]; applied channel-wise, it gives $\|C - C'\|_\infty \leq \max_{s,k} |\hat{c}_{s,k} - \hat{c}'_{s,k}|$, where C, C' are the pre-closure aggregates. Writing $\hat{c} = C/Z$ and $\hat{c}' = C'/Z'$ and adding and subtracting C'/Z , $\|\hat{c} - \hat{c}'\|_1 \leq \frac{1}{Z} \|C - C'\|_1 + \|C'\|_1 |Z^{-1} - Z'^{-1}|$. The first term is $\leq \frac{1}{Z} \|C - C'\|_1$; for the second, $\|C'\|_1 = Z'$ and $|Z^{-1} - Z'^{-1}| = |Z - Z'|/(ZZ') \leq \|C - C'\|_1/(ZZ')$, giving a second term $\leq \frac{1}{Z} \|C - C'\|_1$. Hence $\|\hat{c} - \hat{c}'\|_1 \leq \frac{2}{Z} \|C - C'\|_1 \leq \frac{2K}{\delta} \|C - C'\|_\infty$, and the claim follows. $\square$

The constant degrades as $\delta \rightarrow 0$, i.e., when a cell's total pre-closure mass is near zero; in practice, δ is bounded away from zero because each retained cell has at least one present source with strictly positive credit on a covered channel (assumption A1 and the masking of Section 4.4), so $Z \geq \min_s r_s \cdot \min_k \hat{c}_{s,k} > 0$. Cells failing this are not reconciled.

The operator is, moreover, not an ad hoc combination: it is the robust aggregation over a credible set of source weightings fixed by the elicited redundancy.

Theorem 2 (Variational characterization). *Suppose the elicited 2-additive capacity has non-positive interactions $m(\{s, t\}) \leq 0$ for every pair, as the coverage-Jaccard redundancy of Section 4.5 guarantees. Then, μ is submodular and the channel-wise Choquet aggregate is the support function of its base (core) polytope,*

$$C_\mu(\hat{c}_{\cdot,k}) = \max_{p \in B(\mu)} \sum_{s=1}^m p_s \hat{c}_{s,k}, \quad B(\mu) = \left\{ p \in \Delta^{m-1} : \sum_{s \in A} p_s \leq \mu(A) \quad \forall A \subseteq [m] \right\}. \quad (12)$$

Consequently, the reconciled distribution $\hat{c} = C / \sum_k C_k$ is the L_1 -normalization of a robust aggregation: among all additive source weightings that respect every redundancy cap $\sum_{s \in A} p_s \leq \mu(A)$, the operator selects, per channel, the value best supported by the reported credits. If μ is additive, $B(\mu)$ collapses to the single weighting $(m(\{s\}))_s$ and (12) reduces to the weighted mean of Proposition 1.

Proof. For a 2-additive capacity, the second-order difference satisfies, for all A and $s, t \notin A$, $\mu(A \cup \{s, t\}) - \mu(A \cup \{s\}) - \mu(A \cup \{t\}) + \mu(A) = m(\{s, t\}) \leq 0$, the defining inequality of submodularity. By monotonicity (Proposition 2), every $p \in B(\mu)$ has $p_s \geq \mu([m]) - \mu([m] \setminus \{s\}) \geq 0$, so $B(\mu) \subseteq \Delta^{m-1}$. The discrete Choquet integral is the Lovász extension of μ [20]; for a submodular capacity, the Lovász extension is convex and coincides with the support function of the base polytope, attained at the greedy permutation vertices $p_{\pi(i)}^\pi = \mu(S_i^\pi) - \mu(S_{i-1}^\pi)$ [63]. Evaluating channel-wise gives (12); additivity leaves $(m(\{s\}))_s$ as the unique feasible weighting, recovering the weighted mean. $\square$

The elicitation of Section 4.5 sets every pairwise Möbius coefficient to $-\beta R_{st} \sqrt{r_s r_t} \leq 0$, so μ is submodular in every condition we run and Theorem 2 applies throughout. (A super-additive capacity—positive interactions, which our elicitation never produces—would instead make C_μ the lower support function over the core of μ ; we do not pursue that case.)

The variational reading has an axiomatic counterpart: the channel aggregator is pinned down uniquely by interpretable properties, so the non-additive form is forced rather than chosen.

Theorem 3 (Axiomatic characterization). *Fix the elicited capacity μ . Among all channel aggregators $G : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ that are (i) monotone ($x \leq y \Rightarrow G(x) \leq G(y)$), (ii) co-monotonically additive ($G(x + y) = G(x) + G(y)$ whenever x, y are co-monotone), and (iii) calibrated to the capacity ($G(\mathbf{1}_A) = \mu(A)$ for every $A \subseteq [m]$), the discrete Choquet integral C_μ is the unique one. Consequently the reconciliation map $\hat{c} = C / \sum_k C_k$ is the unique simplex-valued operator induced by a co-monotonically additive, monotone aggregator calibrated to μ .*

Proof. Properties (i)–(iii) are exactly Schmeidler’s axioms characterizing the Choquet integral with respect to a capacity [64]; on the finite index set $[m]$ they determine $G = C_\mu$ uniquely. By Proposition 1 the L_1 -closure is the unique positive rescaling of a non-zero, non-negative vector onto the simplex, so the induced simplex operator is unique as well. $\square$

The three axioms read naturally in the reconciliation context: monotonicity (more credit reported for a channel never lowers its aggregate), co-monotonic additivity (sources that rank the channels identically—no disagreement to reconcile—combine linearly), and capacity calibration (a coalition A of sources contributes exactly its elicited weight $\mu(A)$). The non-additivity that down-weights redundant agreement is therefore not a modeling convenience but the unique consequence of these three properties together with the elicited μ .

Remark 1 (provenance). *To keep the inherited and the novel clearly separated, we state provenance explicitly wherever a classical result is used. The mathematical ingredients of Theorems 2 and 3 are classical: the identity of the discrete Choquet integral with the Lovász extension and, for a submodular capacity, with the support function of the base polytope [20,63], and Schmeidler’s co-monotonic-additivity characterization of the Choquet integral [64]. We claim neither as new. What is specific here is their use to characterize the proposed reconciliation operator—the composition of the elicited capacity with the L_1 -closure onto the simplex (Proposition 1)—which is what makes the variational and axiomatic readings statements about a valid distribution-valued estimator rather than about the integral in isolation.*

Finally, a limit of the setting—not of the estimator—admits a precise statement.

Proposition 7 (Output-only identifiability limit). *No reconciliation rule that depends on the source outputs alone can, in every instance, separate a source that should be discarded from one that should be up-weighted. There exist two instances with identical observable data $\{\hat{c}_s\}_{s=1}^m$ —hence identical coverage, reporting-rate, and consensus-correlation signals—whose ground truths differ so that source m is pure noise in the first and exactly correct in the second.*

Proof. Take $K = 3$ channels and $m = 3$ sources. Sources 1 and 2 cover $\{1, 2\}$ and report a common vector there; source 3 covers only channel 3 and reports $\hat{c}_3 = e_3$. In instance A, the truth is supported on $\{1, 2\}$, so source 3’s report is spurious and the source should be discarded; in instance B, the truth places positive credit on channel 3, which only source 3 observes, so its report is correct and the source should be kept. The observable data $\{\hat{c}_s\}$ coincide across A and B, and source 3’s consensus correlation is undefined on its disjoint support in both (our gate therefore assigns it $\rho_3 = 1$, i.e., no discount, in either instance); hence any function of the observables returns the same weight for source 3 in both instances and is wrong in at least one. $\square$

The adversary gate (Section 4.5) resolves the *typical* case—a source uncorrelated with everyone on shared support—but Proposition 7 shows the adversarial/complementary ambiguity cannot be removed without external information (an audit or an experiment), a point we revisit in Section 8.

The uncertainty band, finally, carries a distribution-free guarantee.

Proposition 8 (Finite-sample coverage). *Calibrate the band as in Section 4.6. Index the calibration units by (cell, channel) pairs $i = 1, \dots, n$, each with reconciled value $\hat{c}_i$, ground-truth credit c_i^* , and Mamdani half-width $h_i > 0$, and let*

$$s_i = \frac{|\hat{c}_i - c_i^*|}{h_i}$$

be the multiplicative (locally adaptive) non-conformity scores. Let q_α be the $\lceil (1 - \alpha)(n + 1) \rceil$ -th smallest of $s_1, \dots, s_n$. If the scores $s_1, \dots, s_n$ of the calibration units and the score s_{n+1} of a test unit are exchangeable, then the adaptive band $\hat{c}_{n+1} \pm q_\alpha h_{n+1}$ satisfies

$$1 - \alpha \leq \Pr(|\hat{c}_{n+1} - c_{n+1}^*| \leq q_\alpha h_{n+1}) \leq 1 - \alpha + \frac{1}{n+1}. \quad (13)$$

Proof. Since $h_{n+1} > 0$, the event $|\hat{c}_{n+1} - c_{n+1}^*| \leq q_\alpha h_{n+1}$ is identical to $s_{n+1} \leq q_\alpha$, so this is the split-conformal guarantee for the score s [60,61]. By exchangeability, the rank of s_{n+1} among $\{s_i\}_{i=1}^{n+1}$ is uniform on $\{1, \dots, n+1\}$; hence $\Pr(s_{n+1} \leq q_\alpha) = \lceil (1 - \alpha)(n + 1) \rceil / (n + 1)$, and the displayed bounds follow from $\lceil (1 - \alpha)(n + 1) \rceil \in [(1 - \alpha)(n + 1), (1 - \alpha)(n + 1) + 1]$. The band is locally adaptive because its half-width $q_\alpha h_{n+1}$ scales with the Mamdani half-width h_{n+1} . $\square$

The guarantee is distribution-free but conditional on exchangeability of the conformity scores. Under the synthetic protocol, the cells are i.i.d., so (13) holds marginally over (cell, channel) pairs and is confirmed empirically in Section 7.3—any residual gap from the nominal level (e.g., 0.799 rather than 0.800) reflects the within-cell dependence among the K channels, which leaves the cell-level exchangeability intact. On real attribution streams, the cells are time-ordered and exchangeability may fail under temporal drift—one further reason the real-data study (Section 7.8) is treated as illustrative rather than a coverage claim.

Together, Theorems 1–3, Propositions 1–8, and Lemma 1 show that the framework outputs a valid distribution (Proposition 1), is the robust aggregation over a redundancy-constrained credible set (Theorem 2) and the unique operator meeting natural aggregation axioms (Theorem 3), decomposes transparently into a weighted mean plus a redundancy correction (Proposition 6), behaves correctly under consensus Proposition 3) and missing sources (Proposition 4), uses a well-defined monotone capacity (Proposition 2), bounds the adversary gate (Proposition 5), reports a conflict diagnostic on a fixed scale (Lemma 1), delivers a finite-sample coverage guarantee (Proposition 8), and is Lipschitz-stable to input noise (Theorem 1)—while Proposition 7 delimits what no output-only rule can achieve.

5. Synthetic Validation Design

Because c^* is unobservable in reality, the framework is validated on synthetic data where the truth is known. The design's credibility rests on a strict separation between the data-generating process and the method.

5.1. Ground-Truth Model

For each conversion i , we draw a journey length $L_i \sim \text{Poisson}(\lambda)$ and sample L_i touch channels from a fixed popularity prior $\pi \sim \text{Dir}(\mathbf{1})$, recording per channel the touch count n_{ik} and a recency timestamp. A hidden per-channel lift $\lambda^{\text{lift}} \sim \text{Dir}(\gamma \mathbf{1})$ encodes the true incremental value of each channel. The true credit follows a *latent-incrementality* law with diminishing returns,

$$c_{ik}^* = \frac{\lambda_k^{\text{lift}} (1 - e^{-n_{ik}})}{\sum_{k'} \lambda_{k'}^{\text{lift}} (1 - e^{-n_{ik'}})} \in \Delta^{K-1}. \quad (14)$$

The saturating factor $1 - e^{-n}$ rewards repeated exposure with decreasing marginal returns. The two latent objects—the lift λ^{lift} and the popularity prior π —are distinct and *neither is exposed to the method*, which observes only the distorted source reports of Section 5.2. Equation (14) is deliberately unrelated to the last-touch, time-decay, and position rules the sources apply, so recovering c^* is not a matter of inverting a rule the method already knows.

5.2. Source Distortion Models

Each source produces a distorted view of the journey through its own lens: *partial observability* (restriction to a visibility set V_s and renormalization), *rule bias* (last-touch, first-touch, time-decay with a half-life, position-based, or a data-driven proxy), *window truncation*, *self-preferencing* (multiplying the own channel by b_s), multiplicative/additive *observation noise*, and *missingness*. Sources may share a *correlated-noise group*, which induces higher-order dependence used in the redundant-bloc condition. Concretely, source s applies its attribution rule ρ_s to the touches it can see, inflates its own channel, perturbs the result, and renormalizes onto its visibility set:

$$\tilde{c}_{s,k} = \rho_s(\{\text{touches in } V_s \text{ within window } W_s\})_k \cdot b_s^{[k=\text{own}(s)]} \cdot \varepsilon_{s,k}, \quad \hat{c}_{s,k} = \frac{\tilde{c}_{s,k} \mathbf{1}[k \in V_s]}{\sum_{k' \in V_s} \tilde{c}_{s,k'}}, \quad (15)$$

with $\varepsilon_{s,k}$ a multiplicative log-normal (or additive) noise term, optionally shared within a correlated-noise group. With probability p_s the source does not report the cell at all. Default parameters are listed in Table 3.

Table 3. Synthetic generator default parameters (Section 5).

Component	Default Setting
channels K /conversions	5/7000–20,000
journey length	Poisson ($\lambda = 3$)
true-credit model	latent incrementality, saturating counts
sources m	5 (7 in the more-sources condition)
rules	last-touch, time-decay, position-based, data-driven proxy
windows	1–7 days (SKAN-like truncation for one source)
self-preference b_s	1.25–1.4 (2.2 in the heavy-bias condition)
noise σ	0.05–0.20 ($\times 2$ in the high-noise condition)
missingness	1–5% per source

5.3. The Generator/Method Firewall

To prevent circular self-confirmation, we enforce that the truth process does not encode the method’s assumptions: c^* comes from latent incrementality (a process the method never sees); source distortions are drawn from families distinct from those the capacity and conflict layers assume; and source redundancy is an *emergent* property of overlapping visibility sets and shared noise, not a parameter the method consumes.

5.4. Misspecification Regimes

We additionally include conditions in which the method’s assumptions are deliberately wrong: a *capacity* misspecification (a broad-coverage source made noisy and a narrow source made accurate, so coverage-based reliability is misleading) and a *dependence* misspecification (a shared-noise bloc creating higher-order redundancy). A method that wins only in its native regime would not transfer to real data; Section 7 reports performance under these regimes.

6. Experimental Setup

6.1. Baselines

We compare against weak and strong baselines (Table 4). The strong tier is the bar the method must clear; we report the method's margin over the *best* strong baseline per condition.

Table 4. Baseline aggregators (Section 6.1).

Baseline	Tier	Description
naive mean	weak	unweighted mean of present sources
median	weak	per-channel median, then closure
trimmed mean	weak	per-channel mean after dropping the extremes
single-source trust	weak	the most-reliable present source's vector
reliability-weighted mean	strong	additive reliability weighting (double-counts redundancy)
Dempster–Shafer	strong	singleton BPAs, reliability discounting, Dempster's rule, pignistic
correlation-cluster	strong	cluster sources by redundancy, average within then across

The Dempster–Shafer baseline [37–39] treats each source's credit as a basic probability assignment over channel singletons, discounts by reliability, combines by Dempster's rule, and reads out the pignistic transform. The correlation-cluster baseline removes pairwise double-counting *without* a capacity and is therefore the most demanding comparison for the non-additive claim.

6.2. Metrics

Recovery against c^* is measured by mean absolute error (MAE), root mean squared error, the Kullback–Leibler divergence $KL(c^* \parallel \hat{c})$ [65], and the Earth-Mover distance on credit vectors [66]. Ranking fidelity uses the Spearman correlation between reconciled and true channel rankings. Uncertainty quality is the empirical coverage of the band at nominal levels.

6.3. Conditions and Protocol

The grid comprises eight conditions (Table 5): the default; a redundant bloc (three near-duplicate trackers sharing a noise group plus two independent sources); high noise; heavy self-preferencing; a capacity misspecification; a fully adversarial source; genuinely complementary sources (largely disjoint visibility); and a larger seven-source variant. Every condition is run at ten random seeds and results are averaged; the interaction strength is fixed at $\beta = 0.5$ and is not tuned per condition.

Table 5. The eight grid conditions (Section 6.3).

Condition	What It Stresses
default	heterogeneous but overlapping sources
redundant_bloc	higher-order redundancy (shared-lens bloc)
high_noise	observation noise ($\times 2$)
heavy_bias	strong self-preferencing
misspec_capacity	coverage–reliability mismatch
adversarial	a random, uninformative source
complementary	largely disjoint visibility
more_sources_m7	a seven-source variant (five-source redundant bloc + two independent)

7. Results

All synthetic results are reported as the mean over ten random seeds to guard against seed-specific artifacts (Section 8 notes that single-seed runs produced misleading swings). The elicited capacity is used throughout.

7.1. The Reconciliation Problem: Sources Diverge

Figure 2 shows, on the default condition, that the five sources assign materially different mean credit and depart from the held-out ground truth—the disagreement the framework must reconcile. Partial observability is visible directly: a source assigns zero credit to channels outside its visibility set.

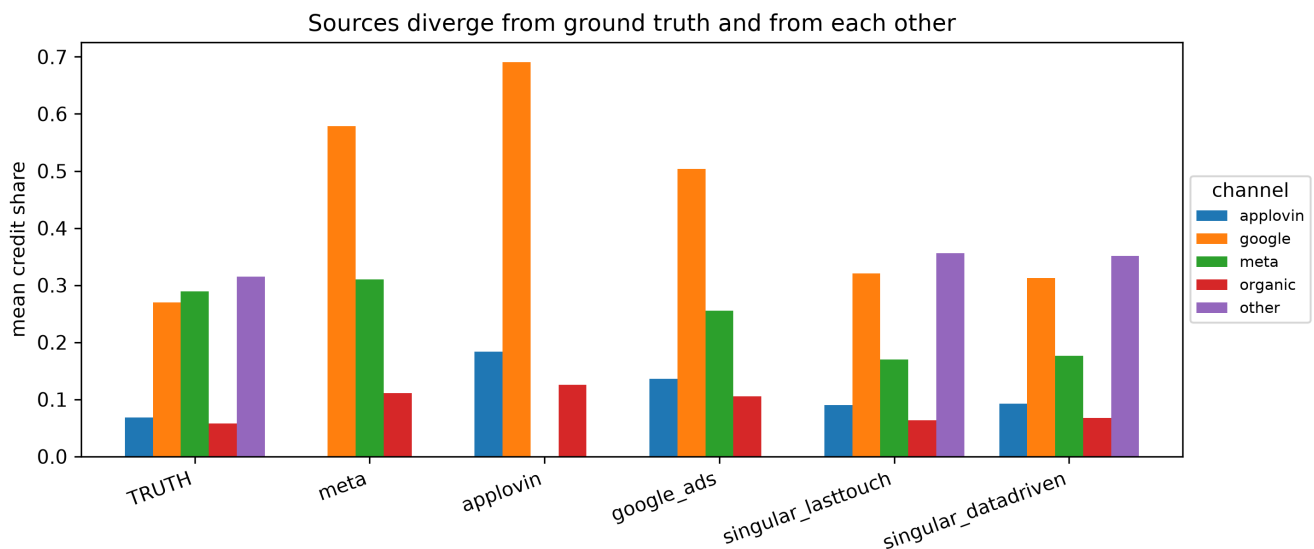


Figure 2. Per-source mean credit share versus ground truth on the default synthetic condition. Sources diverge from the truth and from one another, motivating reconciliation.

7.2. Recovery Versus Baselines Across the Grid

Tables 6 and 7 report MAE and KL for all methods across the eight conditions; Figure 3 summarizes the method's margin over the best strong baseline. The framework attains the lowest MAE in **five of eight** conditions and the lowest KL in **six of eight**. In the redundancy- and bias-dominated regimes that motivate it (default, redundant bloc, high noise, heavy self-preferencing, capacity misspecification), it improves on the best strong baseline by **1.6–2.3%** (MAE) and **3.1–4.4%** (KL); the MAE gain is statistically significant—its paired per seed 95% confidence interval excludes zero—in four of these five conditions, the redundant-bloc gain being positive but not significant. The framework does *not* help where its premise fails: under a pure adversarial source, it loses about 1% on MAE (while still leading on KL), under genuinely complementary, largely non-redundant sources, it loses about 1%, and when a single lens is replicated across a five-source bloc, the correlation-cluster baseline—purpose-built for clean heavy redundancy—leads by about 2%. The contribution is therefore a reliable reconciler for the redundancy regime it targets, not a universal accuracy gain; Section 7.7 maps the operating envelope. Because a simpler ordinal integral on the *same* capacity matches Choquet on point error (Section 7.6), this accuracy advantage is carried by the elicited non-additive capacity, not by the Choquet operator specifically; we use Choquet for the cardinal, calibratable output the confidence layer and Theorem 2 rest on.

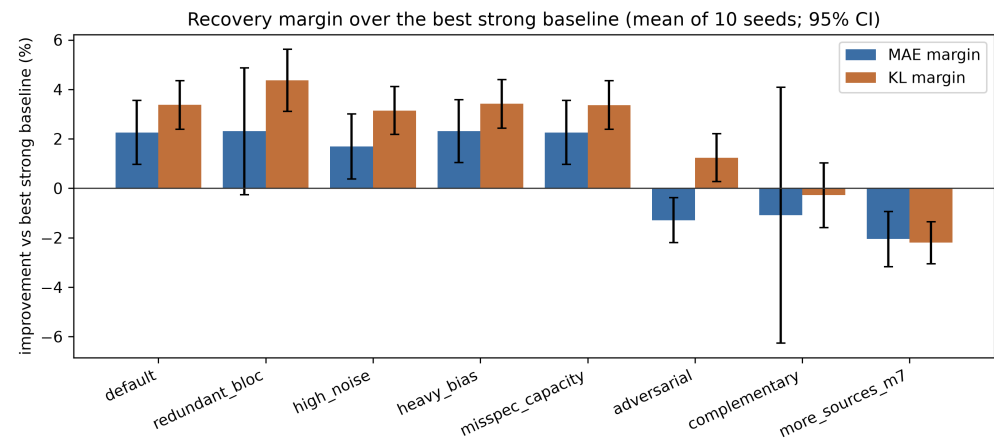


Figure 3. Recovery margin (improvement in MAE and KL, %) over the best strong baseline across the eight grid conditions, mean of ten seeds. Positive favours the method.

Table 6. Recovery MAE by method and condition (mean of 10 seeds; lower is better; best in bold).

Condition	Ours	DS	Corr-Cluster	Rel-Wtd	Naive	Median
default	0.1176	0.1672	0.1259	0.1203	0.1218	0.1362
redundant_bloc	0.1259	0.1609	0.1289	0.1301	0.1356	0.1740
high_noise	0.1200	0.1672	0.1282	0.1220	0.1235	0.1370
heavy_bias	0.1173	0.1672	0.1258	0.1201	0.1216	0.1364
misspec_capacity	0.1176	0.1672	0.1259	0.1203	0.1217	0.1362
adversarial	0.1382	0.1674	0.1511	0.1365	0.1528	0.1381
complementary	0.1174	0.1662	0.1162	0.1179	0.1162	0.1439
more_sources_m7	0.1315	0.1718	0.1289	0.1387	0.1450	0.1764

Table 7. Recovery KL divergence by method and condition (mean of 10 seeds; lower is better; best in bold).

Condition	Ours	DS	Corr-Cluster	Rel-Wtd	Naive	Median
default	1.0189	5.6049	1.1267	1.0544	1.0753	4.1878
redundant_bloc	0.9104	2.0013	0.9519	0.9641	1.0240	8.0011
high_noise	1.0238	5.6053	1.1325	1.0570	1.0781	4.2509
heavy_bias	1.0198	5.5961	1.1290	1.0559	1.0768	4.1867
misspec_capacity	1.0187	5.6043	1.1262	1.0541	1.0749	4.1874
adversarial	0.5974	4.7165	0.7280	0.6049	0.6613	2.9029
complementary	2.1789	7.7259	2.1728	2.1826	2.1728	5.6947
more_sources_m7	0.9728	2.8981	0.9519	1.0590	1.1385	8.3928

The margin over the best strong baseline is summarized in Table 8.

7.3. Uncertainty Calibration

The raw Mamdani band is over-confident, but the conformal step restores nominal empirical coverage, as guaranteed by Proposition 8: 0.799 at the 80% level and 0.897 at the 90% level (Table 9, Figure 4). The calibrated band thus provides an honest, distribution-free uncertainty statement that the point-estimate baselines do not offer.

Calibration is not specific to the default condition: Table 10 reports the calibrated coverage at the 80% and 90% levels across all eight grid conditions, and in every case it tracks the nominal level within ± 0.01 . The conformal layer therefore yields trustworthy uncertainty bands regardless of the disagreement regime.

Table 8. Margin of the framework over the best strong baseline (mean of 10 seeds; positive favours the method; computed from the four-decimal recovery values of Tables 6 and 7). An asterisk marks an MAE margin whose paired per seed 95% confidence interval excludes zero.

Condition	MAE Margin (%)	KL Margin (%)
default	+2.2 *	+3.4
redundant_bloc	+2.3	+4.4
high_noise	+1.6 *	+3.1
heavy_bias	+2.3 *	+3.4
misspec_capacity	+2.2 *	+3.4
adversarial	−1.2 *	+1.2
complementary	−1.0	−0.3
more_sources_m7	−2.0 *	−2.2

Table 9. Band coverage before and after conformal calibration (default condition). The nonconformity score is the per-(cell, channel) ratio $|\hat{c}_k - c_k^*|/h$ and the calibrated band is $\hat{c}_k \pm q_\alpha h$ (Proposition 8); the raw column is the uncalibrated Mamdani band ($q = 1$).

Nominal	Raw Coverage	Calibrated Coverage
0.80	0.766	0.799
0.90	0.766	0.897

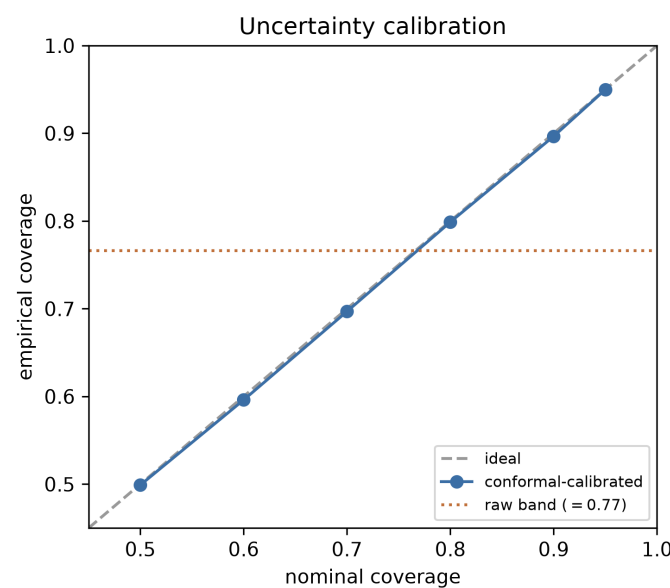


Figure 4. Empirical versus nominal coverage. The raw Mamdani band is over-confident; the conformal calibration tracks the diagonal.

7.4. Ablation

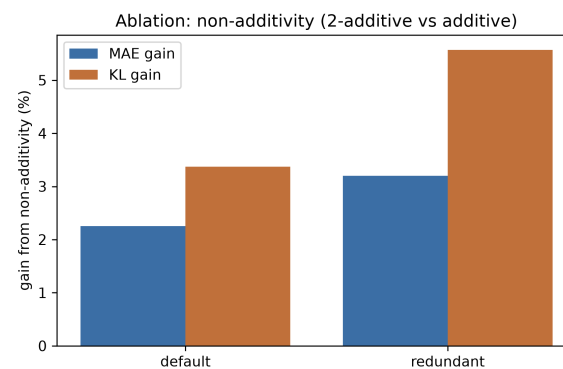
Forcing the capacity additive ($m(\{s, t\}) = 0$) collapses the Choquet integral to the reliability-weighted mean (the additive column below coincides with the reliability-weighted-mean baseline of Table 6) and degrades recovery by 2.2% (MAE) on the default condition and 3.2% on the redundant bloc (Table 11, Figure 5)—non-additivity earns its place, most clearly where redundancy is strongest. At equal 90% coverage, the adaptive Mamdani band is modestly sharper than a constant-confidence band (mean width 0.82 vs. 0.90); its more important contribution, however, is the interpretable, auditable rule base. The 3-additive extension did not improve recovery and was therefore not adopted.

Table 10. Conformal-calibrated coverage across all eight conditions (nominal 80%/90%; same per-cell, channel) ratio score and adaptive band $\hat{c}_k \pm q_\alpha h$ as Table 9).

Condition	Coverage at 80%	Coverage at 90%
default	0.799	0.897
redundant_bloc	0.794	0.898
high_noise	0.797	0.898
heavy_bias	0.798	0.896
misspec_capacity	0.799	0.897
adversarial	0.798	0.899
complementary	0.797	0.895
more_sources_m7	0.791	0.895

Table 11. Ablation of non-additivity (full 2-additive vs. additive capacity; mean of 10 seeds, same configuration as Tables 6 and 7).

Condition	Full MAE	Additive MAE	Gain (%)	Full KL	Additive KL	Gain (%)
default	0.1176	0.1203	+2.2	1.0189	1.0544	+3.4
redundant_bloc	0.1259	0.1301	+3.2	0.9104	0.9641	+5.6

**Figure 5.** Recovery gain from non-additivity (full 2-additive capacity versus an additive capacity), on the default and redundant-bloc conditions.

7.5. Interpretability of the Learned Capacity

Figure 6 shows the elicited interaction indices and Shapley values on the redundant-bloc condition. The three near-duplicate trackers receive strongly negative pairwise interaction indices (the capacity recognizes their shared lens and down-weights the coalition), while the two independent sources carry higher Shapley importance—an interpretable, auditable account of how each source influences the reconciled credit.

7.6. Operator Comparison: Choquet Versus Sugeno

The framework admits any fuzzy integral over the elicited capacity. Table 12 compares the primary Choquet operator with the Sugeno integral (Section 2.2) on the *identical* capacity. The ordinal Sugeno integral yields marginally lower MAE and KL in most conditions—its max–min aggregation is softer and less prone to the over-concentration the Choquet integral can exhibit—while Choquet wins the adversarial MAE (Sugeno retaining the lower KL there) and ties on the complementary one. We nevertheless retain Choquet as the primary operator because (i) it satisfies the simplex-consistency reduction to the weighted mean (Proposition 1); (ii) it exposes the cardinal Shapley and interaction indices that drive the interpretable diagnostics (Figure 6); (iii) the conformal confidence band builds on its cardinal output, whereas the Sugeno integral is purely ordinal; and (iv) the Choquet operator alone admits the variational reading of Theorem 2—it is the support function

of the capacity's base polytope, hence a principled robust aggregation over a credible set of source weightings—a guarantee the ordinal Sugeno integral does not carry. The point-error gap is small and the cardinal, calibrated, and robustly-characterized output is, for a decision-support tool, worth more than a marginal MAE reduction. To the best of our knowledge, this is the first head-to-head comparison of the two fuzzy integrals for source reconciliation without ground truth, and we report it transparently. From a practitioner's standpoint, the choice of Choquet over the marginally more accurate Sugeno integral is a choice of auditability and calibration over a fraction of a percent of point error. Only the Choquet output is cardinal, so only it supports the conformally calibrated band that flags contested cells (Section 7.3), the Shapley/interaction diagnostics that make a disputed allocation traceable to specific sources (Figure 6), and the variational reading of Theorem 2. For a decision-support tool whose purpose is to justify budget moves, an interval and an audit trail are worth more than a marginal, unexplained MAE improvement; the ordinal Sugeno integral, lacking a cardinal scale, offers neither.

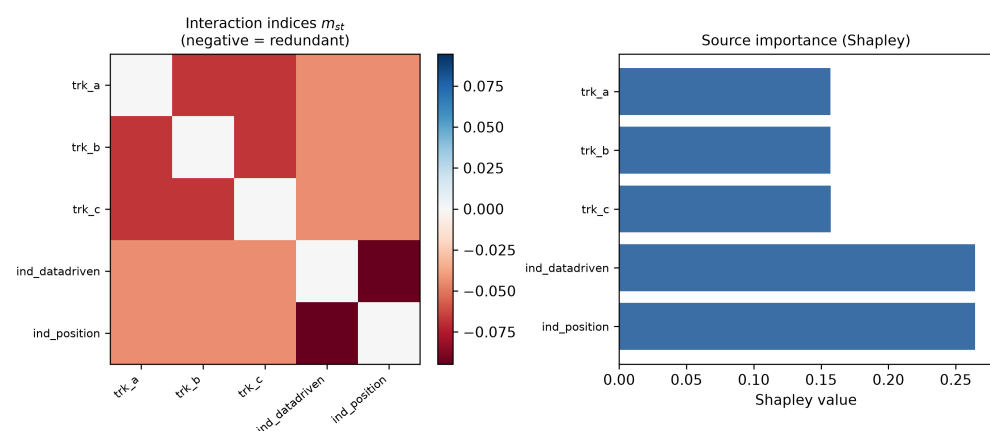


Figure 6. Learned interaction indices m_{st} (left; negative = redundant) and per-source Shapley importance (right) on the redundant-bloc condition.

Table 12. Choquet vs. Sugeno integral on the identical elicited capacity (mean of 10 seeds; lower is better).

Condition	Choquet MAE	Sugeno MAE	Choquet KL	Sugeno KL
default	0.1176	0.1138	1.0189	0.9492
redundant_bloc	0.1259	0.1216	0.9104	0.8240
high_noise	0.1200	0.1202	1.0238	0.9713
heavy_bias	0.1173	0.1133	1.0198	0.9485
misspec_capacity	0.1176	0.1138	1.0187	0.9492
adversarial	0.1382	0.1555	0.5974	0.5544
complementary	0.1174	0.1177	2.1789	2.1739
more_sources_m7	0.1315	0.1257	0.9728	0.8467

7.7. Sensitivity Analysis

Figure 7 and Table 13 report robustness to the interaction strength β and scaling with the number of sources m . The MAE margin over the best strong baseline is stable, between +1.3% and +2.7%, across $\beta \in [0.2, 0.8]$ —the single interaction hyperparameter therefore requires no tuning. The advantage is largest at moderate redundancy ($m = 4$ – 5 , +1.9 to +2.3%) and turns negative as a single lens is replicated across many sources ($m = 7$, −2.0%), where correlation-clustering handles the clean bloc; this is consistent with the more-sources grid condition (Section 7.2) and delimits the regime in which non-additive reconciliation is most valuable. For a practitioner transferring the framework to another domain, the

sensitivity analysis gives two operational rules. Because the MAE margin stays within $[+1.3\%, +2.7\%]$ across $\beta \in [0.2, 0.8]$, the single interaction-strength hyperparameter needs no per-domain tuning—a mid-range value is a safe default. Because the advantage tracks redundancy (moderate overlap optimal, vanishing when lenses are near-identical or fully disjoint), the deciding question in a new domain is not the number of sources but how much observable overlap they share: the method is most valuable when several sources partially share a lens, and offers little when they are either independent or near-duplicates. Estimating the coverage-Jaccard redundancy (Equation (9)) on a domain's own feeds, before any tuning, is enough to predict whether reconciliation will help.

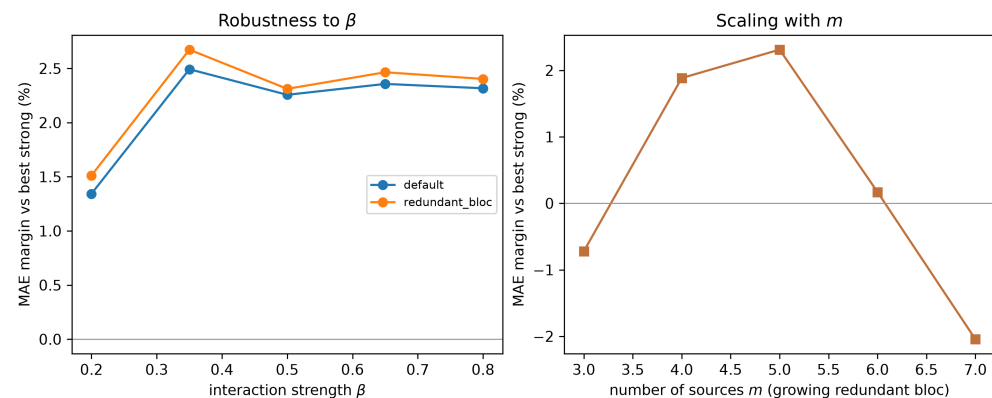


Figure 7. Sensitivity of the MAE margin over the best strong baseline (mean of 10 seeds) to the interaction strength β (left) and to the number of sources m as a redundant bloc grows (right). The grey horizontal line marks zero margin (parity with the best strong baseline); points above it favour the proposed method.

Table 13. Sensitivity of the MAE margin to β and m (mean of 10 seeds; margins from full-precision recovery values, so the $\beta = 0.5$ default entry of +2.3 corresponds to the +2.2 of Table 8, which is rounded from the four-decimal Tables 6 and 7).

Sweep	Setting	MAE Margin (%)
β (default)	0.2/0.35/0.5/0.65/0.8	+1.3/+2.5/+2.3/+2.4/+2.3
β (redundant)	0.2/0.35/0.5/0.65/0.8	+1.5/+2.7/+2.3/+2.5/+2.4
m (growing bloc)	3/4/5/6/7	−0.7/+1.9/+2.3/+0.2/−2.0

7.8. Illustrative Case Study (Proprietary Mobile-Game Data)

We close with an illustrative application to a real, proprietary dataset from a mobile-game studio. Because no ground truth exists on real attribution data, this study makes **no accuracy claim**; it demonstrates that the elicited pipeline runs end-to-end on operational data and that its outputs — an auditable capacity, reconciled credit with an uncertainty band, and a decision-relevant disagreement diagnostic—behave plausibly. With only two sources, however, the capacity has a single interaction term, so the higher-order redundancy machinery that motivates the method—interaction indices and Shapley reweighting across a bloc (cf. Figure 6)—is *not* meaningfully exercised here; a multi-source real or semi-synthetic instance is left to future work. The data cover two casual word games over 183 consecutive days (366 app-day cells) and the two attribution sources that an advertiser must routinely reconcile in practice: a mobile measurement partner's last-touch feed (**S1**) and the aggregated in-platform self-attribution of the ad networks (**S2**). Credit is reported over four channels, anonymized here as **Channels A, B, C** (paid networks), and **Channel D**, organic. All channel and source identities are anonymized and all monetary quantities are reported as ratios or shares; the raw data are commercially confidential.

Elicited capacity. The capacity is elicited from observable signals exactly as in the synthetic pipeline (Section 4.5). The MMP feed, which alone reports organic, receives the higher reliability, and the two sources carry a negative interaction index, correctly identifying that they share the paid-channel lens and should not have their agreement double-counted (Table 14). With two sources, the 2-additive capacity has a single interaction term, so the case is a transparent minimal instance of the construction.

Table 14. Elicited capacity for the case study: per-source elicited reliability r_s (before normalization into Möbius singletons) and Shapley importance ϕ_s (which sum to one). The pairwise interaction index is $I(S1, S2) = -0.229$ (negative $\Rightarrow$ redundant/sub-additive).

Source	Elicited Reliability r_s	Shapley ϕ_s
S1 (MMP last-touch)	1.000	0.540
S2 (in-network self-report)	0.877	0.460

Reconciled credit and disagreement. The two sources disagree substantially: the per-cell credit distributions differ by a median L_1 distance of 0.44 (on a $[0, 2]$ scale), driven by a structural blind spot—the in-network source attributes **zero** credit to organic and inflates the paid channels, whereas the MMP feed assigns organic a 22% share. The reconciliation places organic at 13.7%—a reliability-and-redundancy-weighted compromise rather than a midpoint—and pulls the three paid channels down accordingly (Table 15, Figure 8a). Crucially, the uncertainty band is not constant: its half-width correlates with cross-source disagreement at $\rho = 0.89$ (Figure 8b), so the method automatically flags exactly those app-days on which the two feeds conflict most, the days on which a credit figure should least be trusted. Because real attribution data carry no ground truth, the band reported here is the raw Mamdani half-width (Section 4.6) *before* conformal rescaling, and therefore carries no coverage guarantee: Proposition 8 applies on the synthetic benchmark, where calibration residuals exist (Section 7.3), not on this dataset. We read the band as a relative, data-driven flag of contested cells, not as a calibrated interval.

Table 15. Volume-weighted credit share by source view and the reconciled distribution with its mean (raw, uncalibrated) Mamdani band half-width. The band is a per-cell scalar—identical across channels within a cell but varying across cells (Figure 8b)—so the reported value is its cell-mean. The in-network view is structurally blind to organic.

Channel	S2 In-Network	S1 MMP	Reconciled	Band $\pm$
Channel A	0.471	0.410	0.429	0.058
Channel B	0.252	0.174	0.209	0.058
Channel C	0.276	0.194	0.226	0.058
Channel D (organic)	0.000	0.222	0.137	0.058

Decision impact. The disagreement is economically material. Attributing organic-driven installs to paid channels, as the in-network view does, understates the implied cost per install: relative to the reconciled credit, the in-network numbers overstate paid credit by 9.1%, 17.4%, and 18.1% on Channels A, B, and C, respectively, and understate the **blended paid CPI** by 13.7% (Table 16, Figure 8c). A studio that re-allocated budget on the in-network figures alone would therefore systematically over-invest in paid acquisition. The framework neither claims to recover the true CPI nor replaces a randomized incrementality test; it makes the size and direction of the cross-source distortion explicit, and attaches an uncertainty band—the raw, uncalibrated Mamdani half-width, since real data carry no ground truth to calibrate against—so that a contested re-allocation can be deferred until the

disagreement narrows. This is the intended role of the method as a decision-support layer (Section 8), here exercised on real operational data.

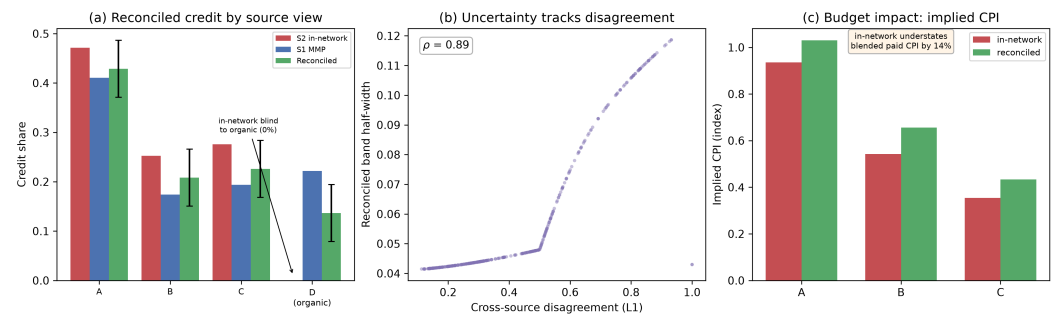


Figure 8. Real-data case study (anonymized). (a) Volume-weighted credit share per channel under the in-network view (S2), the MMP view (S1), and the reconciled distribution (error bars: mean band half-width); the in-network source assigns zero credit to organic. (b) The reconciled band half-width rises with cross-source disagreement ($\rho = 0.89$), so uncertainty is widest where the feeds conflict. (c) Implied cost-per-install (indexed) under the in-network view versus the reconciled credit, for the three paid channels.

Table 16. Budget-attribution impact on the paid channels. The credit shift equals the percentage by which the in-network view understates each channel's implied CPI; the blended paid-CPI understatement is 13.7%.

Channel	Paid-Spend Share	Reconciled vs. In-Network Credit Shift
Channel A	65.3%	−9.1%
Channel B	20.2%	−17.4%
Channel C	14.5%	−18.1%

The case study exercises only two sources, so the higher-order redundancy machinery is not stressed here; its behaviour at larger scale is characterized on the synthetic side. As the number of interacting sources grows, three considerations govern deployment. First, the 2-additive restriction keeps the capacity tractable—quadratically, many interaction terms and a monotonicity check that is negligible for the source counts seen in attribution, and that can be replaced by the sufficient condition of Proposition 2 if the source count is larger. Second, the sensitivity analysis (Section 7.7) shows the advantage is largest at moderate redundancy and narrows once a single lens is replicated across many sources, where a clustering baseline reads the clean block structure directly; practitioners with many near-duplicate feeds should expect the reconciler and correlation-clustering to converge. Third, a genuinely large real deployment would require a multi-source feed with enough observable overlap to elicit the pairwise redundancies reliably; assembling and evaluating such an instance (semi-synthetically, since ground truth remains unobservable) is the natural next step and is left to future work.

8. Discussion

The framework targets a setting that classical attribution modeling sidesteps: the absence of ground truth. We therefore foreground the threats to validity. First, synthetic validation risks circularity; we mitigate it with the generator/method firewall and the misspecification regimes (Section 5), and we treat any real dataset as illustrative rather than confirmatory. Second, agreement does not imply truth; the interaction-corrected invariant (Sections 4.3 and 4.6) prevents correlated sources from inflating confidence. Third, on real data the capacity rests on elicited, observable source properties only; the headline claims are made for the elicited tier, with the oracle reported merely as an upper bound.

A further, intrinsic limit concerns **identifiability**: a fully adversarial source and a genuinely complementary source can be statistically indistinguishable from outputs alone—both correlate weakly with the consensus—yet one should be discarded and the other kept. Our adversary gate resolves the common case (an erratic source uncorrelated with everyone) but cannot, in principle, separate the two without external information; Proposition 7 makes this precise in existential form—*there exist* two instances with identical observable signals on which the two sources demand opposite treatment, so every rule reading the source outputs alone is incorrect on at least one—so the limit is a property of the no-ground-truth setting rather than of our estimator. Methodologically, single-seed and even three-seed evaluation produced misleading swings—both spurious losses and optimistically large wins—so all reported numbers are ten-seed means with confidence intervals (Table 8), a practice we recommend for this class of study.

A distinct limitation concerns the elicitation itself rather than the aggregation. The observable signals that drive the capacity—channel coverage, reporting rate, and consensus correlation—can become weak or unstable when sources share little support, report few cells, or vary little across cells: the adversary-gate correlation ρ_s is then estimated from a handful of shared (cell, channel) pairs and is noisy or undefined. Our construction fails safe in this regime—an undefined ρ_s is set to 1 (no discount), the conservative counterpart of the identifiability limit of Proposition 7—but a source that is genuinely adversarial yet observed on too few shared cells cannot be detected from outputs alone, and a redundancy estimate built on thin overlap will be correspondingly uncertain. In deployment, we therefore recommend gating the elicited capacity on a minimum shared-support threshold and falling back to the reliability-weighted mean when it is not met; quantifying the elicitation's own uncertainty, and propagating it into the band, is left to future work.

Limitations. Margins over strong baselines are modest (around 2% on MAE, 3–4% on KL), so the contribution is a *reliable, interpretable* reconciler rather than a large accuracy gain. The method is bested by the correlation-cluster baseline when a single lens is replicated across many sources (the five-source bloc), a regime for which that baseline is purpose-built. We read this not as a defect but as an *operating regime*: the sensitivity analysis (Section 7.7) shows the non-additive correction pays off at moderate redundancy ($m = 4$ –5 overlapping lenses), is insensitive to β over $[0.2, 0.8]$, and narrows once a single lens is replicated so heavily that a clustering baseline can read the clean block structure directly. Symmetrically, where there is little redundancy to correct—a pure adversarial source, or genuinely complementary (largely non-redundant) sources—the non-additive term has nothing to exploit and the method ties or slightly trails the best baseline (Section 7.2). Characterizing where reconciliation helps is itself part of the contribution; the method is not offered as a universal dominator. We also found that the ordinal Sugeno integral, on the identical capacity, attains marginally lower point error than Choquet (Section 7.6); we retain Choquet for its cardinal properties (Proposition 1, Shapley/interaction indices, and the calibratable band), but operator choice is itself a worthwhile design axis. Performance depends on the quality of the elicited reliability and redundancy signals; the capacity scales as 2^m , which the 2-additive restriction and elicitation keep tractable for the small source counts typical of attribution. Finally, the synthetic generator, while firewalled from the method, encodes assumptions about how sources distort the truth; the real-data case study (Section 7.8) probes external validity but cannot confirm accuracy.

A related caveat concerns the *single-source-trust* baseline: when an unusually accurate source exists and can be identified, simply trusting it attains low error. Our synthetic generator includes one such near-accurate source, so this baseline is competitive on the most benign conditions. The framework's premise, however, is precisely that on real self-attributing platforms *no source is reliably correct*—each is biased toward its own channel—so

an identifiable best source is not available; the method's value is demonstrated in the redundancy- and bias-dominated regimes where no single source dominates.

Practical implications. For a practitioner the framework changes three things. First, it replaces an arbitrary choice—trust one platform, or average—with a reconciled distribution whose construction is auditable through the interaction indices and Shapley values (Figure 6), so a disputed allocation can be traced to specific sources. Second, the band converts a single point estimate into an honest interval whose width signals contention: conformal calibration gives it nominal coverage on the synthetic benchmark (Section 7.3), and in deployment on real data—where no ground truth exists to calibrate against—it is the raw Mamdani band, still tracking cross-source disagreement at $\rho = 0.89$ (Section 7.8) and, at equal coverage, modestly sharper than a constant band (Section 7.4). Its value in deployment is *triage*—knowing which figures to distrust—backed by an interpretable rule base. Third, because the capacity is elicited from observable signals only, the method can be run continuously as new daily reports arrive, with no labelled outcomes and no manual tuning of the single interaction strength β (Section 7.7). These properties make the output suitable as a decision-support layer rather than a black-box score. To make the practical significance concrete: the 1.6–2.3% MAE and 3.1–4.4% KL gains are cell averages, but they concentrate on the redundant, high-disagreement cells that most distort budget decisions. On the real two-source data, these are the same cells that drive the 13.7% blended-CPI distortion of Section 7.8; a reconciler that removes even part of that distortion changes which channels clear a target-CPI threshold, and therefore the direction of a reallocation. The value of the method is thus not a marginal error reduction in the aggregate but the correction of the specific, decision-critical cells where redundant sources would otherwise be double-counted.

Generalization. Any setting that fuses several biased, partially-overlapping estimators of a latent distribution—sensor networks, expert panels, ensemble forecasts—shares this structure, so the framework is not specific to advertising attribution.

9. Conclusions

We formulated reconciliation of conflicting simplex-valued sources as fuzzy information fusion under the absence of ground truth, proved a simplex-consistency result for Choquet aggregation with an L_1 -closure (Proposition 1), characterized the operator as the support function of a submodular capacity's base polytope—a robust aggregation over a redundancy-constrained credible set of weightings (Theorem 2)—and introduced an interpretable Mamdani confidence layer with conformal calibration governed by an independence-corrected agreement invariant. The capacity is elicited from observable source properties alone, so the method transfers to real data. On a firewalled, ten-seed synthetic benchmark, the framework attains the lowest MAE in five of eight conditions and the lowest KL in six, improving on strong baselines (Dempster–Shafer, correlation-clustering, reliability-weighted mean) by 1.6–2.3% (MAE) and 3.1–4.4% (KL)—significant in four of the five targeted conditions—in the redundancy- and bias-dominated regimes typical of walled-garden attribution, while delivering conformal-calibrated uncertainty the baselines do not. This advantage is carried by the elicited non-additive capacity rather than the Choquet operator itself (a simpler ordinal integral on the same capacity matches it on point error); Choquet is retained for the cardinal, calibratable output its diagnostics and confidence band require. The method is a reliable, auditable reconciler rather than a large-margin predictor, and its contribution and limits are scoped accordingly. The illustrative case study of Section 7.8 exercises the pipeline on real operational data; future work will pursue a controlled incrementality benchmark, capacity learning beyond elicitation, and the adversarial/complementary identifiability question of Section 8. Several further

extensions follow naturally from the present formulation. The static capacity could be made dynamic—a time-varying capacity that tracks drift in source reliability as attribution feeds evolve, with the conformal layer adapted to the non-exchangeable, streaming setting. The elicitation could incorporate weak external anchors (occasional audits or incrementality tests) to break the output-only identifiability limit of Proposition 7. Finally, because the framework requires only biased, partially-overlapping estimators of a latent distribution, it applies beyond advertising attribution—to sensor networks, expert panels, and ensemble forecasts—and evaluating it in one such domain would test its generality directly.

Funding: This research received no external funding.

Data Availability Statement: The synthetic data generator, the full method implementation, all configuration files, the result tables, and the figure-generation code are openly available at <https://github.com/ahmettezcantekin/attribution-fuzzy-reconciliation> (accessed on 21 July 2026) (MIT license). The proprietary mobile-game dataset used for the illustrative case study cannot be shared for commercial-confidentiality reasons.

Conflicts of Interest: Author Ahmet Tezcan Tekin is the founder and operator of Soft Towel Limited, the mobile-game studio that provided the anonymized attribution data used in the illustrative case study (Section 7.8). The synthetic experiments, on which all quantitative claims rest, are independent of that data. The author declares no other competing interests.

References

1. Gordon, B.R.; Zettlemeyer, F.; Bhargava, N.; Chapsky, D. A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook. *Mark. Sci.* **2019**, *38*, 193–364. [\[CrossRef\]](#)
2. Johnson, G. Inferno: A guide to field experiments in online display advertising. *J. Econ. Manag. Strategy* **2023**, *32*, 469–490. [\[CrossRef\]](#)
3. Hosahally, S.; Bharadwaj, M.; Zaremba, A.; Volkova, O. Measuring digital advertising in a post-cookie era: A study of marketing-mix models, attribution and incrementality. *J. Digit. Soc. Media Mark.* **2025**, *12*, 348–373. [\[CrossRef\]](#)
4. Mamdani, E.; Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man-Mach. Stud.* **1975**, *7*, 1–13. [\[CrossRef\]](#)
5. Tekin, A.T. Fuzzy Graded Preprocessing for Robust Machine Learning: A Three-Stage Mamdani Framework with Interpretable Audit Trails. *Appl. Sci.* **2026**, *16*, 5072. [\[CrossRef\]](#)
6. Tekin, A.T. Fuzzy Rule-Based Explanations for Tabular Black-Box Classifiers: A Comprehensive Empirical Framework with Prediction-Boundary-Aware Partitioning and Rule-Level Uncertainty Indication. *Appl. Sci.* **2026**, *16*, 5896. [\[CrossRef\]](#)
7. Singal, R.; Besbes, O.; Désir, A.; Goyal, V.; Iyengar, G. Shapley Meets Uniform: An Axiomatic Framework for Attribution in Online Advertising. *Manag. Sci.* **2022**, *68*, 7065–7791, iii–iv. [\[CrossRef\]](#)
8. Shao, X.; Li, L. Data-driven multi-touch attribution models. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011. [\[CrossRef\]](#)
9. Anderl, E.; Becker, I.; von Wangenheim, F.; Schum, J.H. Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *Int. J. Res. Mark.* **2016**, *33*, 457–474. [\[CrossRef\]](#)
10. Ji, W.; Wang, X. Additional Multi-Touch Attribution for Online Advertising. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017. [\[CrossRef\]](#)
11. Yao, D.; Gong, C.; Zhang, L.; Chen, S.; Bi, J. CausalMTA: Eliminating the User Confounding Bias for Causal Multi-touch Attribution. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022. [\[CrossRef\]](#)
12. Abhishek, V.; Fader, P.S.; Hosanagar, K. Media Exposure through the Funnel: A Model of Multi-Stage Attribution. *SSRN Electron. J.* **2018**. [\[CrossRef\]](#)
13. Berman, R. Beyond the Last Touch: Attribution in Online Advertising. *Mark. Sci.* **2018**, *37*, 685–853. [\[CrossRef\]](#)
14. Kannan, P.; Reinartz, W.; Verhoef, P.C. The path to purchase and attribution modeling: Introduction to special section. *Int. J. Res. Mark.* **2016**, *33*, 449–456. [\[CrossRef\]](#)
15. Barajas, J.; Bhamidipati, N.L. Incrementality Testing in Programmatic Advertising: Enhanced Precision with Double-Blind Designs. In Proceedings of the The Web Conference, Ljubljana, Slovenia, 19–23 April 2021. [\[CrossRef\]](#)
16. Lewis, R.; Wong, J. Incrementality Bidding and Attribution. *arXiv* **2022**, arXiv:2208.12809.

17. Oh, S.E. Predictive incrementality: Bridging attribution and causal measurement for real return on marketing investment. *J. Digit. Soc. Media Mark.* **2026**, *13*, 310–325. [\[CrossRef\]](#)
18. Choquet, G. Theory of capacities. *Ann. l'Institut Fourier* **1954**, *5*, 131–295. [\[CrossRef\]](#)
19. Sugeno, M. Theory of Fuzzy Integrals and Its Applications. Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.
20. Grabisch, M. The application of fuzzy integrals in multicriteria decision making. *Eur. J. Oper. Res.* **1996**, *89*, 445–456. [\[CrossRef\]](#)
21. Grabisch, M. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets Syst.* **1997**, *92*, 167–189. [\[CrossRef\]](#)
22. Yager, R.R. Uncertainty modeling using fuzzy measures. *Knowl.-Based Syst.* **2015**, *92*, 1–8. [\[CrossRef\]](#)
23. Grabisch, M.; Roubens, M. An axiomatic approach to the concept of interaction among players in cooperative games. *Int. J. Game Theory* **1999**, *28*, 547–565. [\[CrossRef\]](#)
24. Gutiérrez, I.; Castro, J.; Gómez, D.; Espínola, R. Calculating the interaction index of a fuzzy measure: A polynomial approach based on sampling. *Fuzzy Sets Syst.* **2022**, *466*, 108454. [\[CrossRef\]](#)
25. Anderson, D.T.; Islam, M.A.; King, R.L.; Younan, N.H.; Fairley, J.R.; Howington, S.E.; Petry, F.E.; Elmore, P.; Zare, A. Binary fuzzy measures and Choquet integration for multi-source fusion. In Proceedings of the 2017 International Conference on Military Technologies (ICMT), Brno, Czech Republic, 31 May–2 June 2017. [\[CrossRef\]](#)
26. Du, X.; Zare, A.; Anderson, D.T. Multiple Instance Choquet Integral with Binary Fuzzy Measures for Remote Sensing Classifier Fusion with Imprecise Labels. In Proceedings of the 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 6–9 December 2019. [\[CrossRef\]](#)
27. Bhowal, P.; Sen, S.; Velásquez, J.D.; Sarkar, R. Fuzzy ensemble of deep learning models using choquet fuzzy integral, coalition game and information theory for breast cancer histology classification. *Expert Syst. Appl.* **2021**, *151*, 116167. [\[CrossRef\]](#)
28. Wang, J.; Zhang, X.; Shen, Q. Choquet-Like Integrals With Rough Attribute Fuzzy Measures for Data-Driven Decision-Making. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 2825–2836. [\[CrossRef\]](#)
29. Bustince, H.; Mesiar, R.; Fernández, J.; Galar, M.; Paternain, D.; Altalhi, A.; Dimuro, G.P.; Bedregal, B.; Takáč, Z. d-Choquet integrals: Choquet integrals based on dissimilarities. *Fuzzy Sets Syst.* **2020**, *414*, 1–27. [\[CrossRef\]](#)
30. Yager, R.R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst. Man Cybern.* **1988**, *18*, 183–190. [\[CrossRef\]](#)
31. Mesiar, R.; Stupňanová, A.; Yager, R.R. Generalizations of OWA Operators. *IEEE Trans. Fuzzy Syst.* **2015**, *23*, 2154–2162. [\[CrossRef\]](#)
32. Yager, R.R. On the OWA Aggregation with Probabilistic Inputs. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2015**, *23*, 143–162. [\[CrossRef\]](#)
33. Miguel, L.D.; Sesma-Sara, M.; Elkano, M.; Asiáin, M.J.; Bustince, H. An algorithm for group decision making using n-dimensional fuzzy sets, admissible orders and OWA operators. *Inf. Fusion* **2017**, *39*, 126–131. [\[CrossRef\]](#)
34. Bardozzo, F.; de la Osa, B.; Horanská, L.; Fumanal-Idocin, J.; Priscoli, M.D.; Troiano, L.; Tagliaferri, R.; Fernández, J.; Bustince, H. Sugeno integral generalization applied to improve adaptive image binarization. *Inf. Fusion* **2020**, *68*, 37–45. [\[CrossRef\]](#)
35. Wiczyński, J.; Lucca, G.; Borges, E.N.; Urió-Larrea, A.; Molina, C.L.; Bustince, H.; Dimuro, G.P. Application of the Sugeno integral in Fuzzy Rule-Based Classification. *Appl. Soft Comput.* **2024**, *167*, 112265. [\[CrossRef\]](#)
36. nia, A.S.; Olfat, L.; Esmaeili, A.; Rostamzadeh, R.; Antuchevičienė, J. Using Fuzzy Choquet Integral Operator for Supplier Selection with Environmental Considerations. *J. Bus. Econ. Manag.* **2016**, *17*, 503–526. [\[CrossRef\]](#)
37. Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. [\[CrossRef\]](#)
38. Shafer, G. *A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976.
39. Sarabi-Jamab, A.; Araabi, B.N. How to decide when the sources of evidence are unreliable: A multi-criteria discounting approach in the Dempster–Shafer theory. *Inf. Sci.* **2018**, *448*, 233–248. [\[CrossRef\]](#)
40. Tang, Y.; Wu, S.; Zhou, Y.; Huang, Y.; Zhou, D. A New Reliability Coefficient Using Betting Commitment Evidence Distance in Dempster–Shafer Evidence Theory for Uncertain Information Fusion. *Entropy* **2023**, *25*, 462. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Xia, J.; Feng, Y.; Liu, L.; Liu, D.; Fei, L. An Evidential Reliability Indicator-Based Fusion Rule for Dempster–Shafer Theory and its Applications in Classification. *IEEE Access* **2018**, *6*, 24912–24924. [\[CrossRef\]](#)
42. Xue, D.; Wang, Y.; Chunlan, Y. An evidence combination approach based on fuzzy discounting. *Soft Comput.* **2020**, *25*, 3003–3011. [\[CrossRef\]](#)
43. Zhu, C.; Qin, B.; Xiao, F.; Cao, Z.; Pandey, H.M. A fuzzy preference-based Dempster–Shafer evidence theory for decision fusion. *Inf. Sci.* **2021**, *570*, 306–321. [\[CrossRef\]](#)
44. Yuan, K.; Xiao, F.; Fei, L.; Kang, B.; Deng, Y. Conflict management based on belief function entropy in sensor fusion. *SpringerPlus* **2016**, *5*, 638. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Jiang, W.; Wang, S.; Liu, X.; Zheng, H.; Wei, B. Evidence conflict measure based on OWA operator in open world. *PLoS ONE* **2017**, *12*, e0177828. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Bhowal, P.; Sen, S.; Yoon, J.H.; Geem, Z.W.; Sarkar, R. Evaluation of Fuzzy Measures Using Dempster–Shafer Belief Structure: A Classifier Fusion Framework. *IEEE Trans. Fuzzy Syst.* **2022**, *31*, 1593–1603. [\[CrossRef\]](#)

47. Wu, Z.; Xu, J. Managing consistency and consensus in group decision making with hesitant fuzzy linguistic preference relations. *Omega* **2015**, *65*, 28–40. [[CrossRef](#)]
48. Wu, Z.; Xu, J. A consensus model for large-scale group decision making with hesitant fuzzy information and changeable clusters. *Inf. Fusion* **2017**, *41*, 217–231. [[CrossRef](#)]
49. Liu, X.; Xu, Y.; Herrera, F. Consensus model for large-scale group decision making based on fuzzy preference relation with self-confidence: Detecting and managing overconfidence behaviors. *Inf. Fusion* **2019**, *52*, 245–256. [[CrossRef](#)]
50. Lu, Y.; Xu, Y.; Huang, J.; Wei, J.; Herrera-Viedma, E. Social network clustering and consensus-based distrust behaviors management for large-scale group decision-making with incomplete hesitant fuzzy preference relations. *Appl. Soft Comput.* **2021**, *117*, 108373. [[CrossRef](#)]
51. Izquierdo, S.S.; Izquierdo, L.R. Mamdani Fuzzy Systems for Modelling and Simulation: A Critical Assessment. *J. Artif. Soc. Soc. Simul.* **2018**, *21*, 2. [[CrossRef](#)]
52. Mendel, J.M.; Bonissone, P.P. Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 3579–3593. [[CrossRef](#)]
53. Fernández, A.; Herrera, F.; Cordon, Ó.; del Jesús, M.J.; Marcelloni, F. Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to? *IEEE Comput. Intell. Mag.* **2019**, *14*, 69–81. [[CrossRef](#)]
54. Cao, J.; Zhou, T.; Zhi, S.; Lam, S.; Ren, G.; Zhang, Y.; Wang, Y.; Dong, Y.; Cai, J. Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review. *Inf. Sci.* **2024**, *662*, 120212. [[CrossRef](#)]
55. Chen, T.; Shang, C.; Su, P.; Keravnou-Papailiou, E.; Zhao, Y.; Antoniou, G.; Shen, Q. A Decision Tree-Initialised Neuro-fuzzy Approach for Clinical Decision Support. *Artif. Intell. Med.* **2020**, *111*, 101986. [[CrossRef](#)] [[PubMed](#)]
56. Lucca, G.; Dimuro, G.P.; Fernández, J.; Bustince, H.; Bedregal, B.; Sanz, J. Improving the Performance of Fuzzy Rule-Based Classification Systems Based on a Nonaveraging Generalization of CC-Integrals Named $C_{F_1 F_2}$ -Integrals. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 124–134. [[CrossRef](#)]
57. Walia, N.; Singh, H.; Sharma, A. ANFIS: Adaptive Neuro-Fuzzy Inference System—A Survey. *Int. J. Comput. Appl.* **2015**, *123*, 32–38. [[CrossRef](#)]
58. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [[CrossRef](#)]
59. Dubois, D.; Prade, H. Operations on fuzzy numbers. *Int. J. Syst. Sci.* **1978**, *9*, 613–626. [[CrossRef](#)]
60. Vovk, V.; Gammerman, A.; Shafer, G. *Algorithmic Learning in a Random World*; Springer: New York, NY, USA, 2005.
61. Angelopoulos, A.N.; Bates, S. Conformal Prediction: A Gentle Introduction. *Found. Trends Mach. Learn.* **2023**, *16*, 494–591. [[CrossRef](#)]
62. Aitchison, J. The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1982**, *44*, 139–160. [[CrossRef](#)]
63. Lovász, L. Submodular functions and convexity. In *Mathematical Programming: The State of the Art*; Springer: Berlin/Heidelberg, Germany, 1983; pp. 235–257. [[CrossRef](#)]
64. Schmeidler, D. Integral representation without additivity. *Proc. Am. Math. Soc.* **1986**, *97*, 255–261. [[CrossRef](#)]
65. Kullback, S.; Leibler, R.A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [[CrossRef](#)]
66. Rubner, Y.; Tomasi, C.; Guibas, L.J. The Earth Mover’s Distance as a Metric for Image Retrieval. *Int. J. Comput. Vis.* **2000**, *40*, 99–121. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.