THE COERCIVE PROJECTION THEOREM FOR CANONICAL RECIPROCAL COSTS

JONATHAN WASHBURN AND AMIR RAHNAMAI BARGHI

ABSTRACT. We develop a finite-data framework for certifying *zero-defect* (neutral) configurations of positive vectors under the canonical separable reciprocal cost. We show that this scalar cost is characterized among non-constant continuous costs by the Recognition Composition Law together with a local quadratic calibration at balance; in particular, reciprocity symmetry and the normalization at the neutral point follow from the composition law. Under a conservation constraint and short-window observations of a rational (finite–state) signal class, we construct a canonical decision procedure that is *locally maximal on the identifiability locus* among all sound procedures: any sound rule that resolves a datum must agree with the canonical output, and cannot resolve strictly more cases.

The method is organized as $\Phi^* = A \circ B \circ P$: the projection/coercivity core is forced by the canonical-cost axioms, while the aggregation/reconstruction step is specified on a non-degenerate identifiability locus (e.g. a Hankel invertibility condition).

1. INTRODUCTION

We introduce a finite-window certification problem in which observed aggregates are tested for compatibility with a neutral *zero-defect* class. After fixing notation and the ratio-induced cost framework, we summarize the main results and explain how the later sections build the proof pipeline. For additional background on the broader Recognition Science program and its geometric framing, see [3].

Relation to Hankel/Prony reconstruction. Although our setting is organized around aggregated window data and a canonical convex-analytic decision core, the non-degenerate identifiability locus is expressed via finite Hankel full-rank/invertibility conditions. This connects to the classical finite-rank/Hankel-structure viewpoint (e.g. Kronecker-type characterizations and related operator theory) and to Prony-style exponential fitting, where solving structured linear systems built from short samples is central; see, for example, [8, 7].

1.1. **Connection to classical projection and proximal tools.** The certification template is *forced* by our axioms, but its shape echoes classical projection/proximal ideas from convex analysis. We refer to the main theorem as the *Coercive Projection Theorem* (CPT), and we write its canonical finite-data certifier as Φ^*. First, the CPT certifier Φ^* begins with a projection-like update built from the canonical reciprocal cost together with a coercive regularization; at a structural level this plays the role of a proximity map that turns a penalty into a stable "best-approximation" operator (cf. Moreau's proximity operator and envelope [4]). Second, because our state variables are inherently multiplicative (ratios/scales), the natural geometry is non-Euclidean: Bregman divergences [5] describe projection in coordinates adapted to log/ratio structure, matching the way our axioms select log coordinates as

---

* Correspondence: `arahnamab@gmail.com`.

the intrinsic "neutral" parameterization. Third, the coercive resolvent-type step that converts defect bounds into uniqueness and perturbation stability is in the spirit of Rockafellar's proximal-point framework for monotone operators [6].

The difference is that the classical theory typically offers a *family* of legitimate proximal/projection schemes (choice of divergence, step size, regularizer). In this paper we separate two levels explicitly: Lemma 3.25 gives the per-state order-theoretic representation, and Theorem 3.31 gives the global uniqueness/forcedness statement under explicit cross-state rigidity hypotheses. The subsequent aggregation/reconstruction stage A is a standard inverse-problem ingredient and is only invoked on a non-degenerate (locally invertible) measurement locus; it is an additional hypothesis beyond the cost axioms. This separation is what makes the resulting certifier both canonical (in its cost-driven components) and reproducible.

1.2. The problem.

Given finite observational access to a configuration $x \in (\mathbb{R}_{>0})^n$, decide whether it has *zero-defect*. In the present setting, "zero-defect" means membership in the structured set

$$S := \{x \in (\mathbb{R}_{>0})^n : J_n(x) = 0\}.$$

Here $J_n : (\mathbb{R}_{>0})^n \to [0, \infty)$ denotes the separable extension of J,

$$J_n(x) := \sum_{i=1}^n J(x_i). \tag{1}$$

More generally, ratio-induced recognition costs take the form

$$c(s, o) := J\left(\frac{\iota_S(s)}{\iota_O(o)}\right), \tag{2}$$

where ι_S, ι_O are positive scale maps and $J : (0, \infty) \to [0, \infty)$ is a penalty.

Under inversion symmetry, strict convexity/coercivity, normalization at 1, and a multiplicative d'Alembert identity, one can classify admissible penalties and recover (up to a scale parameter) the same canonical form $J(x) = \cosh(a \log x) - 1$; see [2]. In this paper we take the normalized choice (4) as given and focus on the finite-data certification problem. Under the canonical cost, S collapses to the single configuration $x \equiv 1$ (all components equal to 1). The question is therefore the most basic decision task: determine from finite-data whether x is exactly the neutral configuration.

1.3. Main goal and results.

Our goal is to characterize the *locally optimal* and *structurally forced* certification procedure for the basic decision question "is the configuration equal to the neutral state $x \equiv 1$?" on the rational/finite-state class. Under three explicit hypotheses-(i) the canonical reciprocal cost forced by the composition law, (ii) a conservation constraint, and (iii) finite local resolution with window length $W \geq 1$ (e.g. $W = 8$)-we show that the canonical procedure is locally maximal (in resolve-set domination) on the identifiability locus, and that any sound procedure that resolves there must agree with its output. Moreover, the operational pipeline uses three universal components: a projection step P, a coercivity step B, and an aggregation step A. The *structure* of this decomposition and the sharp coercivity constant in B are consequences of the canonical-cost axioms, while the reconstruction/aggregation step A requires an *additional* nondegeneracy (local invertibility) hypothesis for the window map (e.g. Hankel/Vandermonde nonsingularity as in Theorem 2.20). Our maximality statements (e.g. Theorem 3.19) are correspondingly local to that locus and do not claim global parameter recovery from window sums.

1.4. **Rigidity and local optimality of the canonical pipeline.** The term *method* suggests optional design choices. In our setting we prove a rigidity statement: among all sound certification rules (Definition 3.11), the canonical pipeline $\Phi^* = A \circ B \circ P$ is locally maximal on the identifiability locus (Theorem 3.19).

What is proved. Rather than packaging the contribution as a list of broad meta-properties, we summarize the concrete statements established in this paper:

(i) **Quantitative sound certification under noise.** Theorems 3.4 and 3.36 give explicit ε–stability bounds for the realized costs and the associated meaning sets, including the sharp two-sided tolerance (membership in $\mathrm{Mean}_{2\varepsilon}$ rather than $\mathrm{Mean}_\varepsilon$).

(ii) **Local domination among sound rules.** Theorem 3.19 proves that on the identifiability locus the canonical pipeline $\Phi^* = A \circ B \circ P$ is locally maximal among all sound certification rules (Definition 3.11), i.e. no sound rule can resolve strictly more instances in a neighborhood.

(iii) **Rigidity/uniqueness consequences of the axioms.** Theorem A.3 characterizes the canonical reciprocal scalar cost from the Recognition Composition Law together with the local calibration at balance, and Theorem 3.31 records the induced factorization/uniqueness behavior for certificates under explicit rigidity hypotheses.

(iv) **Uniqueness among locally optimal rules.** Corollary 2.22 shows that on the identifiability neighborhood $\mathcal{U}_{d,W}$ the canonical rule is not only locally maximal but also unique among procedures achieving local maximality: any sound rule that is locally optimal on $\mathcal{U}_{d,W}$ must agree with Φ^* wherever it resolves.

Lemma 3.25 shows the local ordinal representation, while Theorem 3.31 proves uniqueness of the realized-cost reparametrization and of the state-free profile under explicit rigidity assumptions; the aggregation/reconstruction step A is then pinned down on any non-degenerate (locally invertible) measurement locus, but is *not* implied by the cost axioms alone. For that reason we treat the procedure as a theorem-the *Coercive Projection Theorem* (CPT)-rather than as an algorithmic "recipe" one could freely modify.

We study a compatibility certificate for the neutral class: under the stated axioms, we give sufficient conditions ensuring that the observed finite-window aggregates admit a model-consistent realization in the neutral *zero-defect* class. Criteria for ruling out neutrality, or for statistical hypothesis testing, are not treated here.

Remark 1.1. For convenience we record a short dictionary between the objects defined in this paper and standard cost-based selection terminology. (We avoid the word "kernel" here, since in other contexts it denotes a nullspace or vanishing set.) The *base cost* is our ratio-induced map $c : S \times O \to [0, \infty)$ given by $c(s, o) := J(\iota_S(s)/\iota_O(o))$. The associated *meaning set* (argmin selector) is $\mathrm{Mean}(s) := \arg\min_{o \in O} c(s, o)$. The corresponding *defect* (or *gap*) is

$$\Delta(s, o) := c(s, o) - \inf_{o' \in O} c(s, o'),$$

so that $\Delta(s, o) = 0$ iff $o \in \mathrm{Mean}(s)$. Finally, the CPT *certificate* $\mathcal{C}$ is a scalar test statistic built from the W-block window measurements; in Section 4 it is used to separate the null (neutral) configuration from the nonzero class under the stated coercivity and identifiability hypotheses. This remark is purely notational and introduces no additional assumptions.

2. Model, axioms, and finite-window measurements

This section sets up the mathematical model. We define the state and observation spaces, scale maps, the canonical reciprocal cost and induced defect, and the meaning/selection map. We then define the finite-window aggregation operator and list the axioms used in all subsequent theorems.

Big-picture guide. Section 2 fixes the canonical cost and the neutral/defect geometry in log-coordinates (the forced steps P and B). It then defines the finite-window measurement operator $\mathcal{W}$ and records the identifiability locus needed for the reconstruction step A. Section 3 uses these ingredients to define the finite-data procedure and prove soundness and domination (maximality) on the stated non-degenerate locus.

We state the axioms and fix notation for later use. Proofs are provided where the results are stated.

Definition 2.1. A *recognition cost system* is a quadruple $(\mathbb{R}_{>0}, J, \sigma, W)$ where:

(C1) **Recognition Composition Law (primitive axiom for J).** We assume a scalar cost $J : (0, \infty) \to \mathbb{R}$ is non-constant, twice differentiable in a neighborhood of 1, and satisfies the Recognition Composition Law (RCL): This J is the penalty appearing in the ratio-induced recognition cost $c(s, o)$ in (2), and we write J_n for its separable extension (1).
$$J(xy) + J(x/y) = 2J(x) + 2J(y) + 2J(x)J(y), \qquad x > 0, \ y > 0, \tag{3}$$
together with the curvature normalization $J''(1) = 1$. We use the corresponding canonical reciprocal cost,
$$J(x) := \frac{1}{2}(x + x^{-1}) - 1, \qquad x > 0. \tag{4}$$

(C2) **Conservation.** For $x \in (\mathbb{R}_{>0})^n$, define $\sigma(x) := \sum_{i=1}^n \log x_i$. Admissible configurations satisfy $\sigma(x) = 0$.

(C3) **Finite resolution.** The observation window has fixed integer length $W \geq 1$ (in examples we take $W = 8$): only W consecutive evaluations per cycle of the underlying finite-state representation are available.

For later use we write $J_n(x) := \sum_{i=1}^n J(x_i)$ for $x \in (\mathbb{R}_{>0})^n$.

Remark 2.2. (i) The Recognition Composition Law together with non-constancy forces the balance normalization $J(1) = 0$ and reciprocity $J(x) = J(1/x)$. Under the stated C^2 regularity near 1 and the local quadratic calibration in Axiom 2 (equivalently $J''(1) = 1$; see Appendix), the cost is uniquely determined; equivalently, it is given by (4).

(ii) Section 3.4 will make precise the "rational/finite-state class" and the window data model used by the aggregation step.

2.1. **Roadmap of the proof.** We will (i) identify the canonical cost and its log-coordinate form, (ii) define the three operations P, B, A, and (iii) prove the local representation lemma (Lemma 3.25) together with the strengthened global uniqueness theorem (Theorem 3.31) for the forced-factorization layer under explicit rigidity hypotheses. Complete proofs of the window-data identifiability and domination steps are included in Sections 2.3 and 3.4.

2.2. **Canonical cost and defect geometry.** This subsection collects basic properties of the canonical reciprocal cost J and the associated defect in log-coordinates, which will later control the aggregation and certification steps.

2.2.1. *The canonical reciprocal cost.* We recall that J is fixed by (4). The following logarithmic-coordinate reformulation is used repeatedly.

Lemma 2.3. *Let $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be defined by $\varphi(t) := \cosh(t) - 1$. Then for all $t \in \mathbb{R}$,*
$$J(e^t) = \varphi(t) = \cosh(t) - 1.$$

In particular, $J(x) = 0$ if and only if $x = 1$.

Proof. Using $\cosh(t) = \frac{1}{2}(e^t + e^{-t})$ we compute $J(e^t) = \frac{1}{2}(e^t + e^{-t}) - 1 = \cosh(t) - 1$. Since $\cosh(t) \geq 1$ with equality only at $t = 0$, we have $J(e^t) = 0 \iff t = 0 \iff e^t = 1$. $\square$

2.2.2. *Neutrality and projection in log-coordinates.* We work in log-coordinates $y = \log x \in \mathbb{R}^n$ for $x \in (\mathbb{R}_{>0})^n$.

Definition 2.4. For $n \in \mathbb{N}$ let $\mathbf{1} \in \mathbb{R}^n$ denote the all-ones vector. For $y \in \mathbb{R}^n$ define its mean

$$\bar{y} := \frac{1}{n}\langle y, \mathbf{1}\rangle = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

and define the *mean-zero projection* $P : \mathbb{R}^n \to \mathbb{R}^n$ by

$$P(y) := y - \bar{y}\,\mathbf{1}. \tag{5}$$

equivalently, P is the orthogonal projection onto the subspace $\{z \in \mathbb{R}^n : \langle z, \mathbf{1}\rangle = 0\}$.

Lemma 2.5. *For $y \in \mathbb{R}^n$ with mean $\bar{y} := \frac{1}{n}\langle y, \mathbf{1}\rangle$, the projection $P(y) = y - \bar{y}\,\mathbf{1}$ vanishes if and only if y is a constant vector, i.e. $y = \bar{y}\,\mathbf{1}$. In particular, if additionally $\langle y, \mathbf{1}\rangle = 0$, then $P(y) = 0$ implies $y = 0$.*

Proof. By definition, $P(y) = 0$ is equivalent to $y = \bar{y}\,\mathbf{1}$, which is exactly the condition that all coordinates of y are equal to the constant $\bar{y}$. If moreover $\langle y, \mathbf{1}\rangle = 0$, then $0 = \langle y, \mathbf{1}\rangle = \langle \bar{y}\,\mathbf{1}, \mathbf{1}\rangle = n\bar{y}$, so $\bar{y} = 0$ and hence $y = 0$. $\square$

Definition 2.6. For $x \in (\mathbb{R}_{>0})^n$ let $y = \log x$ and define the *defect* of x by

$$\mathrm{Def}(x) := \|P(y)\|_2\,.$$

We say that x has *zero-defect* if $\mathrm{Def}(x) = 0$.

Remark 2.7. Under the conservation constraint $\sigma(x) = \sum_i \log x_i = 0$, Lemma 2.5 implies that $\mathrm{Def}(x) = 0$ forces $\log x = 0$ and hence $x \equiv 1$.

2.2.3. *Coercivity: energy gap controls defect.* The coercivity step converts cost information into a quantitative defect bound. It rests on the elementary inequality $\cosh(t) - 1 \geq t^2/2$.

Lemma 2.8. *For all $t \in \mathbb{R}$,*

$$\cosh(t) - 1 \;\geq\; \frac{t^2}{2},$$

with equality if and only if $t = 0$.

Proof. Define $f(t) := \cosh(t) - 1 - \frac{1}{2}t^2$. Then $f(0) = 0$, $f'(t) = \sinh(t) - t$ and $f''(t) = \cosh(t) - 1 \geq 0$. Thus f' is increasing with $f'(0) = 0$, hence $f'(t) \geq 0$ for $t \geq 0$ and $f'(t) \leq 0$ for $t \leq 0$, so $t = 0$ is a global minimum of f. Therefore $f(t) \geq 0$ for all t, with equality only at $t = 0$. $\square$

Theorem 2.9. *Let $x \in (\mathbb{R}_{>0})^n$ and $y = \log x$. Then*

$$\sum_{i=1}^{n} J\left(e^{P(y)_i}\right) \;\geq\; \frac{1}{2}\,\|P(y)\|_2^2 \;=\; \frac{1}{2}\,\mathrm{Def}(x)^2. \tag{6}$$

In particular, if $\sum_{i=1}^{n} J\left(e^{P(y)_i}\right) = 0$ then $\mathrm{Def}(x) = 0$.

Proof. By Lemma 2.3, $J(e^{P(y)_i}) = \cosh(P(y)_i) - 1$. Apply Lemma 2.8 termwise and sum over i:

$$\sum_i J(e^{P(y)_i}) = \sum_i (\cosh(P(y)_i) - 1) \;\geq\; \sum_i \frac{P(y)_i^2}{2} = \frac{1}{2}\,\|P(y)\|_2^2\,.$$

If the left-hand side equals 0, then each term equals 0, hence each $P(y)_i = 0$ and $\mathrm{Def}(x) = 0$. $\square$

2.3. Finite-data aggregation: W-block window measurements.

Remark 2.10. All arguments in Sections 2.2.3-3.4 apply verbatim for any fixed integer $W \geq 1$ once the corresponding measurement operator is specified. In the examples we keep the concrete choice $W = 8$.

2.4. Window operator at resolution W.
Fix an integer $W \geq 1$ (e.g. $W = 8$). For a real sequence $y = (y_0, y_1, \dots)$ define its W-*block window sums*

$$\mathcal{W}(y)_k := \sum_{j=0}^{W-1} y_{(Wk+j)}, \qquad k = 0, 1, 2, \dots. \tag{7}$$

For brevity we sometimes write $W_k := \mathcal{W}(y)_k$ for the window sums; this notation should not be confused with the fixed window *length* W. equivalently, $\mathcal{W}(y)$ is obtained by partitioning the time axis into consecutive blocks of length W and summing within each block (e.g. $W = 8$).

Remark 2.11. The choice of fixed W is treated as part of the finite-resolution model (C3); $W = 8$ is used only as a concrete running example. The analysis requires only that the block length is a fixed positive integer.

2.5. The rational (finite-state) class and the degree parameter.
Let $y = (y_n)_{n \geq 0}$ be a real sequence with generating function

$$\theta(z) := \sum_{n \geq 0} y_n z^n.$$

We say that y is *rational of degree at most d* if $\theta(z) = p(z)/q(z)$ where $\deg p \leq d$, $\deg q \leq d$, and $q(0) \neq 0$; we then normalise scale so that $q(0) = 1$. Write

$$q(z) = 1 + q_1 z + q_2 z^2 + \cdots + q_d z^d,$$

allowing $q_d = 0$ when $\deg q < d$ (padding to length d). In this setting, d counts the number of free recurrence coefficients (equivalently, the McMillan degree / state dimension of a minimal linear realisation).

Lemma 2.12. *If $\theta(z) = p(z)/q(z)$ with $q(0) \neq 0$ (so a Taylor expansion at 0 exists) and after normalisation $q(0) = 1$ and $q(z) = 1 + q_1 z + \cdots + q_d z^d$, then the coefficients y_n satisfy the linear recurrence*

$$y_n + q_1 y_{n-1} + q_2 y_{n-2} + \cdots + q_d y_{n-d} = 0, \qquad n \geq d+1. \tag{8}$$

Conversely, any sequence satisfying (8) *for some $(q_1, \dots, q_d)$ is determined by the parameters $(q_1, \dots, q_d)$ and the initial values $(y_0, \dots, y_d)$, hence by $2d+1$ real numbers.*

Proof. Since $\theta(z) = p(z)/q(z)$ and $q(0) = 1$, we have the identity

$$q(z)\,\theta(z) = p(z).$$

Write $q(z) = 1 + q_1 z + \cdots + q_d z^d$ and $\theta(z) = \sum_{n \geq 0} y_n z^n$. Comparing coefficients of z^n for $n \geq d+1$ gives

$$y_n + q_1 y_{n-1} + \cdots + q_d y_{n-d} = 0,$$

because $\deg p \leq d$ implies the z^n coefficient of p is 0 for $n \geq d+1$.

Conversely, if (8) holds for fixed $(q_1, \dots, q_d)$, then each y_n with $n \geq d+1$ is determined recursively from $(y_0, \dots, y_d)$. Hence the full sequence is determined by $(q_1, \dots, q_d, y_0, \dots, y_d)$, i.e. by $2d+1$ real parameters. $\qquad \square$

2.6. **Unique determination from $K \geq 2d + 1$ window sums.** We now fill in the self-contained "unique determination from window data" argument in a form that does *not* rely on external machine-verification tooling. Throughout this subsection we fix integers

$$d \geq 1, \qquad W \geq 1, \qquad K \geq 2d + 1.$$

Let $\theta(z) = \sum_{n \geq 0} y_n z^n$ be a rational signal of degree $\leq d$ with representation $\theta(z) = p(z)/q(z)$ where $\deg p \leq d$, $\deg q \leq d$, and we normalise $q(0) = 1$. Write

$$q(z) = 1 + q_1 z + \cdots + q_d z^d.$$

By Lemma 2.12, once the parameter vector

$$\pi := (y_0, \ldots, y_d, q_1, \ldots, q_d) \in \mathbb{R}^{2d+1}$$

is fixed, the recurrence (8) determines y_n for all $n \geq d + 1$.

Recall the W-block window sums $\mathcal{W}(y)_k$ from (7), and define the truncated window map (using the first $2d + 1$ windows)

$$F_{d,W} : \mathbb{R}^{2d+1} \to \mathbb{R}^{2d+1}, \qquad F_{d,W}(\pi) := (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{2d}),$$

where y is generated from π by (8). Here $DF_{d,W}(\pi)$ denotes the Jacobian matrix of $F_{d,W}$ at π.

Lemma 2.13. *The map $F_{d,W}$ is polynomial in π. Indeed, each y_n is obtained from $(y_0, \ldots, y_d, q_1, \ldots, q_d)$ by iterating the recurrence (8), so y_n is a polynomial function of π for every fixed n; each window sum is a finite linear combination of such y_n, hence $F_{d,W}$ is polynomial. If there exists one parameter point π_0 with $\det DF_{d,W}(\pi_0) \neq 0$, then*

$$\Omega_{d,W} := \{\pi \in \mathbb{R}^{2d+1} : \det DF_{d,W}(\pi) \neq 0\}$$

is a nonempty Zariski-open (hence open dense) subset of $\mathbb{R}^{2d+1}$, and on $\Omega_{d,W}$ the first $2d + 1$ window sums determine π locally uniquely (i.e. $F_{d,W}$ is locally one-to-one).

Proof. Each coefficient y_n is obtained by iterating the linear recurrence (8) whose update rule is polynomial in $(q_1, \ldots, q_d)$ and linear in $(y_0, \ldots, y_d)$. Hence each $\mathcal{W}(y)_k$ is polynomial in π, so $F_{d,W}$ is polynomial and $\det DF_{d,W}$ is a polynomial function. If $\det DF_{d,W}(\pi_0) \neq 0$ for some π_0, then the polynomial $\det DF_{d,W}$ is not identically zero, so its nonvanishing locus $\Omega_{d,W}$ is Zariski-open and nonempty. For $\pi \in \Omega_{d,W}$, the inverse function theorem gives local invertibility, hence local uniqueness of π from the first $2d + 1$ windows. $\square$

Theorem 2.14. *Fix integers $d \geq 1$ and $W \geq 1$. Assume there exists a parameter point π_0 such that $\det DF_{d,W}(\pi_0) \neq 0$. Then the nonvanishing locus*

$$\Omega_{d,W} := \{\pi : \det DF_{d,W}(\pi) \neq 0\}$$

is nonempty and Zariski-open (hence open dense) in $\mathbb{R}^{2d+1}$, and for every $\pi \in \Omega_{d,W}$ the window map $F_{d,W}$ is locally invertible at π. Moreover, to certify the hypothesis it suffices to exhibit an integer point $\pi_0 \in \mathbb{Z}^{2d+1}$ and a prime p for which $\det DF_{d,W}(\pi_0) \not\equiv 0 \pmod{p}$.

Proof. By Lemma 2.13, $\det DF_{d,W}$ is a polynomial in the coordinates of π. If $\det DF_{d,W}(\pi_0) \neq 0$, then this polynomial is not identically zero, so $\Omega_{d,W}$ is Zariski-open and nonempty; the inverse function theorem yields local invertibility at each $\pi \in \Omega_{d,W}$. For the finite-field certificate: if $\pi_0 \in \mathbb{Z}^{2d+1}$ and $\det DF_{d,W}(\pi_0) \not\equiv 0 \pmod{p}$, then the integer $\det DF_{d,W}(\pi_0)$ is not divisible by p, hence is nonzero over $\mathbb{Z}$, and therefore nonzero as a real number. $\square$

Remark 2.15. Theorem 2.14 reduces generic full-rank (and hence local identifiability) for a given (d, W) to finding *one* explicit witness point. In practice, such witnesses can often be found by a brief randomized integer search, and then certified by a modular determinant computation as above.

(d, W)	witness π_0	certificate of full rank
$(3, 8)$	$(1, 1, 5, 1, 2, 2, -2)$	Lemma 2.16

TABLE 1. Example full-rank witness points. Further rows can be added by exhibiting any π_0 with det $DF_{d,W}(\pi_0) \neq 0$.

Lemma 2.16. *For $d = 3$ and $W = 8$, the set $\Omega_{3,8}$ from Lemma 2.13 is nonempty. In particular, this verifies the hypothesis of Theorem 2.14 for $(d, W) = (3, 8)$. In particular, for the parameter point*

$$\pi_0 = (y_0, y_1, y_2, y_3, q_1, q_2, q_3) = (1, 1, 5, 1, 2, 2, -2)$$

one has $\det DF_{3,8}(\pi_0) \neq 0$. *Consequently $\Omega_{3,8}$ is Zariski-open (hence open dense) in $\mathbb{R}^7$, and the first $2d + 1 = 7$ window sums determine π locally uniquely throughout $\Omega_{3,8}$.*

Proof. We certify $\det DF_{3,8}(\pi_0) \neq 0$ by a finite check in a prime field. Let $p := 10^9 + 7$ and work in $\mathbb{F}_p$ throughout. Using the recurrence (8) at $\pi = \pi_0$, compute y_n up to $n = 55$ and hence the first $K = 7$ window sums

$$(W_0, \ldots, W_6) = (-16, -5160, -975168, -169890432,$$
$$-27959752704, -4399334572032, -665805326548992).$$

Next propagate parameter derivatives through the same recurrence: for each parameter $\alpha \in \{y_0, y_1, y_2, y_3, q_1, q_2, q_3\}$ set $u_n^{(\alpha)} := \partial y_n / \partial \alpha$ and differentiate $y_n + q_1 y_{n-1} + q_2 y_{n-2} + q_3 y_{n-3} = 0$ $(n \geq 4)$ to obtain

$$u_n^{(\alpha)} + q_1 u_{n-1}^{(\alpha)} + q_2 u_{n-2}^{(\alpha)} + q_3 u_{n-3}^{(\alpha)} = -(\partial_\alpha q_1) y_{n-1} - (\partial_\alpha q_2) y_{n-2} - (\partial_\alpha q_3) y_{n-3}.$$

These recurrences determine all $u_n^{(\alpha)}$ needed for the window derivatives $\partial_\alpha W_k = \sum_{j=0}^7 u_{8k+j}^{(\alpha)}$. Assemble the 7×7 Jacobian $DF_{3,8}(\pi_0)$ with entries $\partial_\alpha W_k$. (The full integer Jacobian matrix is recorded in Appendix B; here we retain only the determinant residue modulo p.)

Compute the determinant in $\mathbb{F}_p$ by Gaussian elimination. The result is

$$\det(DF_{3,8}(\pi_0)) \equiv 972226939 \pmod{p}.$$

, which is nonzero in $\mathbb{F}_p$. Hence $\det DF_{3,8}(\pi_0) \neq 0$ over the integers. $\square$

Remark 2.17. For any fixed pair (d, W), the determinant $\det DF_{d,W}(\pi)$ is a polynomial function of the parameter vector π (with coefficients depending only on (d, W)). Hence the full-rank locus

$$\Omega_{d,W} := \{\pi : \det DF_{d,W}(\pi) \neq 0\}$$

is either empty or Zariski open (and therefore Euclidean open and dense) in the ambient parameter space. To certify nonemptiness in the present paper (working entirely within the present manuscript), it suffices to exhibit a single explicit witness π_0 with $\det DF_{d,W}(\pi_0) \neq 0$. Lemma 2.16 provides such a witness for the operating example $(d, W) = (3, 8)$. Additional witnesses for other (d, W) pairs can be recorded in the same way.

Theorem 2.18. *Assume $\Omega_{d,W}$ is nonempty (equivalently, $\det DF_{d,W}$ is not the zero polynomial). Then for every $\pi \in \Omega_{d,W}$ the first K window sums $(\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{K-1})$ determine π locally uniquely within the degree-$\leq d$ rational class. In particular, the degree-$\leq d$ signal θ and the full coefficient sequence $(y_n)_{n \geq 0}$ are locally uniquely determined on $\Omega_{d,W}$ by $K \geq 2d + 1$ windows.*

Proof. Fix $\pi \in \Omega_{d,W}$. By Lemma 2.13 the map

$$F_{d,W} : \mathbb{R}^{2d+1} \to \mathbb{R}^{2d+1}, \qquad F_{d,W}(\pi) = (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{2d})$$

has nonsingular Jacobian at π, hence is locally injective near π (Inverse Function Theorem). Now let $K \geq 2d + 1$ and define

$$G_K(\pi) := (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{K-1}) \in \mathbb{R}^K.$$

Let $\mathrm{pr} : \mathbb{R}^K \to \mathbb{R}^{2d+1}$ be the coordinate projection onto the first $2d + 1$ components. Then $\mathrm{pr} \circ G_K = F_{d,W}$. Therefore, if $G_K(\pi') = G_K(\pi)$ for some π', then $F_{d,W}(\pi') = F_{d,W}(\pi)$. Taking π' sufficiently close to π, local injectivity of $F_{d,W}$ forces $\pi' = \pi$. Hence the first K windows determine π locally uniquely on $\Omega_{d,W}$ for every $K \geq 2d + 1$. $\square$

Remark 2.19. Theorem 2.18 is the strongest statement that follows from a Jacobian (non-singularity) argument alone: it guarantees *local* identifiability on an open dense set. A *global* uniqueness theorem (one-to-one on a parameter locus) requires additional structure or additional observables beyond window sums to rule out distinct far-apart preimages with the same window data.

Theorem 2.20. *Let $S_k := \mathcal{W}(y)_k$ be the window-sum sequence for a degree-$\leq d$ rational signal. Assume that S_k admits an exponential-sum representation*

$$S_k = \sum_{i=1}^{d} A_i \, \mu_i^k$$

(H-prony) Assume the exponents $\mu_1, \ldots, \mu_d$ are distinct and nonzero, the amplitudes A_i are nonzero, and the $d \times d$ Hankel matrix $H := (S_{i+j})_{i,j=0}^{d-1}$ is invertible. (This is a standard nondegeneracy / local invertibility condition for Prony-type reconstruction.) Then the data $(S_0, \ldots, S_{2d-1})$ determine uniquely:

  (1) *the monic degree-d characteristic polynomial $\chi(t) = \prod_{i=1}^{d}(t - \mu_i)$ of the order-d recurrence satisfied by (S_k), hence the multiset $\{\mu_i\}_{i=1}^{d}$;*
  (2) *the amplitudes $(A_i)_{i=1}^{d}$, hence the entire sequence $(S_k)_{k \geq 0}$.*

Equivalently, the collection of parameter pairs $(\mu_i, A_i)_{i=1}^{d}$ is determined up to a common permutation of indices. If one fixes any deterministic ordering convention on the distinct exponents (for example, ordering by $|\mu_i|$ when these moduli are pairwise distinct), then the ordered tuple $(\mu_1, \ldots, \mu_d, A_1, \ldots, A_d)$ is uniquely determined.

  Consequently, for $K \geq 2d$ the first K window sums determine the induced window-sum process uniquely on this non-degenerate locus.

Proof. The coefficients of $P(t) = t^d + a_1 t^{d-1} + \cdots + a_d$ solve the Hankel linear system

$$\sum_{m=1}^{d} a_m S_{k+d-m} = -S_{k+d} \qquad (k = 0, 1, \ldots, d-1),$$

whose matrix is precisely $H = (S_{i+j})_{i,j=0}^{d-1}$. By hypothesis H is invertible, hence $(a_1, \ldots, a_d)$ (and thus P) are uniquely determined from $(S_0, \ldots, S_{2d-1})$. Since the μ_i are distinct, the roots of χ are exactly $\{\mu_i\}_{i=1}^{d}$. With μ_i known, the amplitudes (A_i) are obtained uniquely from the Vandermonde system

$$(S_0, \ldots, S_{d-1})^\top = V(\mu) \, (A_1, \ldots, A_d)^\top, \quad V(\mu)_{k,i} = \mu_i^k,$$

whose determinant is $\prod_{i<j}(\mu_j - \mu_i) \neq 0$. This determines (S_k) for all k. $\square$

Definition 2.21 (Local optimality on an identifiability locus)**.** Fix a neighborhood $\mathcal{U} \subset \mathcal{M}_{d,W}$. A sound procedure Ψ is *locally optimal on* $\mathcal{U}$ if for every sound procedure Ξ one has

$$\mathrm{Res}(\Xi) \cap \mathcal{U} \subseteq \mathrm{Res}(\Psi) \cap \mathcal{U}$$

and the outputs agree on every instance resolved by Ξ in $\mathcal{U}$, i.e. for all $y \in \mathrm{Res}(\Xi) \cap \mathcal{U}$,

$$\Psi(\mathcal{W}_{W,K}(y)) = \Xi(\mathcal{W}_{W,K}(y)).$$

Corollary 2.22 (Uniqueness among locally optimal rules)**.** *Under the hypotheses of Theorem 3.19, the canonical procedure Φ^* is locally optimal on $\mathcal{U}_{d,W}$. Moreover, if Ψ is any locally optimal sound procedure on $\mathcal{U}_{d,W}$, then*

$$\mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W} = \mathrm{Res}(\Phi^*) \cap \mathcal{U}_{d,W},$$

and for every $y \in \mathrm{Res}(\Phi^*) \cap \mathcal{U}_{d,W}$ the outputs agree:

$$\Psi(\mathcal{W}_{W,K}(y)) = \Phi^*(\mathcal{W}_{W,K}(y)).$$

Proof. By Theorem 3.19, Φ^* dominates every sound procedure on $\mathcal{U}_{d,W}$, hence it is locally optimal by Definition 2.21. If Ψ is locally optimal, then applying Definition 2.21 with $\Xi = \Phi^*$ gives $\mathrm{Res}(\Phi^*) \cap \mathcal{U}_{d,W} \subseteq \mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W}$ and agreement on that set. Together with Theorem 3.19 applied to Ψ (which yields the reverse inclusion and agreement), this gives the claimed equality of resolve sets and agreement of outputs. $\square$

Remark 2.23 (Parameter uniqueness up to permutation / with ordering)**.** Theorem 2.20 determines the unordered node set $\{\mu_1, \ldots, \mu_d\}$ uniquely, hence parameters are unique up to simultaneous permutation of pairs (μ_i, A_i). If one imposes an ordering convention (for example $\mu_1 < \cdots < \mu_d$ when nodes are real and positive), the parameter tuple itself is uniquely determined.

Remark 2.24**.** In concrete Prony/Hankel implementations, failure of the hypotheses in Theorem 2.20 (distinct exponents, nonzero amplitudes, and invertible Hankel/Vandermonde matrices) occurs on a proper algebraic subset of parameter space (e.g. discriminant and determinant vanishing conditions). Hence the stated nondegeneracy conditions are *generic* (open dense) within standard finite-dimensional parametrizations of degree-$\leq d$ rational signals.

Theorem 2.25 (General rank certification)**.** *For every pair of integers $d \geq 1$ and $W \geq 1$, the identifiability locus $\Omega_{d,W}$ is nonempty. In particular, the generic local uniqueness conclusion of Theorem 2.18 and the global Prony recovery of Theorem 2.20 apply for every (d, W) under the corresponding nondegeneracy hypotheses.*

Proof. We exhibit a witness family. Choose distinct $\alpha_1, \ldots, \alpha_d \in (0, 1)$ (e.g. $\alpha_i := 1/(i+1)$), and define $y_n := \sum_{i=1}^d \alpha_i^n$. Its W-block window sums satisfy

$$S_k = \sum_{i=1}^d B_i\, \mu_i^k, \qquad \mu_i := \alpha_i^W, \quad B_i := \frac{1 - \alpha_i^W}{1 - \alpha_i}.$$

The nodes μ_i are distinct and $B_i > 0$. Hence the Hankel matrix $H = (S_{i+j})_{0 \leq i,j \leq d-1}$ admits

$$H = V^\top \mathrm{diag}(B)\, V,$$

with V the Vandermonde matrix on (μ_i), so

$$\det H = (\det V)^2 \prod_{i=1}^d B_i = \left(\prod_{i<j} (\mu_j - \mu_i) \right)^2 \prod_{i=1}^d B_i \neq 0.$$

Therefore the Prony/Hankel nondegeneracy conditions hold at this witness. By Lemma 2.13, this gives a nonempty full-rank identifiability locus. $\square$

Corollary 2.26. *In the specialized regime used in this paper we take $d = 3$ and $W = 8$, hence $K_{\min} = 2d+1 = 7$. We emphasize the two different thresholds: $K \geq 2d+1$ is the dimensional requirement for* local *recovery of the full parameter $\pi \in \mathbb{R}^{2d+1}$ from the truncated window map $F_{d,W}$, while $K \geq 2d$ is sufficient for* global *recovery of the induced window-sum process via Prony/Hankel methods under the nondegeneracy hypotheses of Theorem 2.20. For this choice, Theorem 2.25 yields $\Omega_{3,8} \neq \emptyset$, so the generic local uniqueness conclusion of Theorem 2.18 holds. Moreover, under the nondegeneracy hypotheses of Theorem 2.20, the first $2d$ window sums (hence any $K \geq 2d$) determine the window-sum sequence (equivalently, the exponential parameters (μ_i, A_i))* globally *uniquely.*

Proof. The specialization $d = 3$, $W = 8$ fixes $K_{\min} = 2d + 1 = 7$ and the local threshold statement is the instance of Theorem 2.18 on the nonempty full-rank locus $\Omega_{3,8}$ certified in Lemma 2.16. The global $K \geq 2d$ threshold for recovery of the induced window-sum process (equivalently, the exponential parameters) follows from Theorem 2.20 under its stated Prony/Hankel nondegeneracy hypotheses. □

3. Certification and optimality from finite-window data

This section develops the finite-data certification pipeline built on the window operator: we pass from raw window measurements to a canonical defect score, then to a decision rule. We isolate the status of the window length W, give ε-robust certification bounds, and define soundness and domination among finite-data procedures, culminating in ε-optimal certification statements.

Theorem 3.1 (Master Theorem (architectural summary)). *Fix integers $d \geq 0$, $W \geq 1$, and $K \geq 2d + 1$. Consider the window map $F_{d,W}$ and the finite-data pipeline*

$$\Phi^* \colon \ w \ \mapsto \ A(w) \ \mapsto \ x \ \mapsto \ u = P(\log x) \ \mapsto \ B(u) \in \{\text{zero}, \text{nonzero}, \text{inconclusive}\},$$

as defined in Section 3. Assume the standing hypotheses needed for the cited results below (in particular, the cost axioms and the identifiability assumptions for the window model). Then:

(i) **Soundness.** *The decision rule Φ^* is sound in the sense of Definition 3.11.*

(ii) *ε-**robust certification.** Under perturbations of size ε, the certified conclusion degrades in the sharp two-sided way, yielding membership in $\text{Mean}_{2\varepsilon}$ as in Theorems 3.4 and 3.36.*

(iii) **Local domination (optimality).** *On the identifiability neighborhood $\mathcal{U}_{d,W}$, Φ^* dominates every sound procedure (Theorem 3.19).*

(iv) **Uniqueness among locally optimal rules.** *Any sound procedure that is locally optimal on $\mathcal{U}_{d,W}$ agrees with Φ^* on all resolved data in $\mathcal{U}_{d,W}$ (Corollary 2.22).*

(v) **Forced decision-layer dependence.** *Any sound defect-based decision rule factors through the scale-invariant invariant $u = P(\log x)$, hence through the (P, B)-core in the sense of Theorem 3.14.*

(vi) **Non-vacuity via generic full rank.** *For any (d, W), if there exists a witness π_0 with $\det DF_{d,W}(\pi_0) \neq 0$, then the full-rank locus is a nonempty Zariski-open (hence dense) subset (Theorem 2.14); for $(d, W) = (3, 8)$ such a witness is certified in Lemma 2.16.*

If, in addition, the window map is globally injective on the parameter locus $\mathcal{M}_{d,W}$, then the domination statement upgrades from $\mathcal{U}_{d,W}$ to $\mathcal{M}_{d,W}$ (Corollary 3.20).

Proof. Items (i)–(ii) follow from Definition 3.11 together with Theorems 3.4 and 3.36. Item (iii) is Theorem 3.19. Item (iv) is Corollary 2.22. Item (v) is the factorization/forced-dependence principle recorded in Theorem 3.14. Item (vi) is Theorem 2.14, with nonemptiness witnessed for $(3, 8)$ by Lemma 2.16. The final global upgrade is Corollary 3.20. □

TABLE 2. Selected constants in the certification layer and where they are forced.

Constant / threshold	Forced by
Factor 2ε in $\text{Mean}_{2\varepsilon}$	Perturbation triangle inequality in Theorem 3.4
Normalization $J(1) = 0$	Proposition A.1 (Appendix, RCL)
Calibration $J''_{\log}(0) = 1$	Axiom 2 (Appendix) fixes quadratic scale
Window count $K \geq 2d + 1$	Identifiability from $2d + 1$ windows (Theorems 2.18, 2.20)

3.1. **Status of the window length W.** Throughout, $W \in \mathbb{N}$ denotes the fixed block length used to form window sums (7). The theory is stated for general W; the value $W = 8$ is a modeling choice, used below only as an illustrative operating point.

A simple (self-contained) motivation for the number 8 comes from a three-bit latent state $b \in \{0,1\}^3$. A Gray-code cycle on the 3-cube visits all $2^3 = 8$ states exactly once before returning to the start, for example

$$000 \to 001 \to 011 \to 010 \to 110 \to 111 \to 101 \to 100 \to 000,$$

where successive states differ in exactly one bit. If a measurement/aggregation step is designed to cover one full cycle of such a three-bit state evolution, then the natural "cycle window" length is $W = 8$. This is only a heuristic design rationale: none of the main theorems depends on it, and the identifiability threshold remains $K \geq 2d + 1$ for any fixed $W \geq 1$.

3.2. **ε-tolerant certification from noisy window sums.** In practice, window sums are observed with finite precision. Instead of exact constraints $\mathcal{W}(y)_k = 0$, we assume a uniform tolerance model

$$|W_k| \leq \varepsilon, \qquad k = 0, 1, \ldots, K - 1. \tag{9}$$

The next result shows that small window-sum errors yield a small certification cost for the reconstructed signal. The bound is quadratic in ε, with constants controlled by the conditioning of the reconstruction map.

Lemma 3.2. *For every $t \in \mathbb{R}$,*

$$\cosh(t) - 1 \;\leq\; \frac{1}{2}\,e^{|t|}\,t^2.$$

In particular, if $|t| \leq \tau$ then $\cosh(t) - 1 \leq \frac{1}{2}e^{\tau}t^2$.

Proof. Using the power series $\cosh(t) - 1 = \sum_{m \geq 1} \frac{t^{2m}}{(2m)!}$ and $(2m)! \geq 2\,(2m-2)!$,

$$\cosh(t) - 1 \leq \frac{t^2}{2} \sum_{m \geq 1} \frac{|t|^{2m-2}}{(2m-2)!} = \frac{t^2}{2}\cosh(|t|) \leq \frac{t^2}{2}e^{|t|}.$$

$\square$

Proposition 3.3. *Fix $W \in \mathbb{N}$, $d \geq 0$, and $n \in \mathbb{N}$, and set $K := \lceil n/W \rceil$. Assume $K \geq 2d+1$ and work on a non-degenerate parameter locus where the truncated window map*

$$\pi \longmapsto \mathcal{W}_{W,K}(y(\pi))$$

is locally invertible at the neutral realization (equivalently, its Jacobian determinant is nonzero). Then there exist $\varepsilon_0 > 0$ and a C^1 reconstruction map

$$\mathsf{Rec}_{d,W} : \{w \in \mathbb{R}^K : \ \|w\|_\infty \leq \varepsilon_0\} \to \mathbb{R}^n$$

such that $\mathsf{Rec}_{d,W}(w)$ *returns the first n coefficients of the unique model signal compatible with the window data w in that neighborhood. Moreover, writing*

$$L = \sup_{\|w\|_\infty \leq \varepsilon_0} \|D\mathsf{Rec}_{d,W}(w)\|_{2\to 2},$$

one has the Lipschitz bound

$$\|\mathsf{Rec}_{d,W}(w) - \mathsf{Rec}_{d,W}(0)\|_2 \leq L\,\|w\|_2 \qquad (\|w\|_\infty \leq \varepsilon_0).$$

Proof. Write $\|\cdot\|_\infty$ for the max norm on $\mathbb{R}^K$, i.e. $\|w\|_\infty := \max_{0 \leq k \leq K-1} |w_k|$, and $\|\cdot\|_2$ for the Euclidean norm. Local invertibility of $\pi \mapsto \mathcal{W}_{W,K}(y(\pi))$ at the neutral realization implies, by the inverse function theorem, that there exist neighborhoods U of the neutral parameter and V of $w = 0$ such that the window map restricts to a C^1 diffeomorphism $U \to V$. Define $\mathsf{Rec}_{d,W}$ on V by taking the inverse map $V \to U$ and then returning the first n coefficients of the corresponding model signal; shrinking V if necessary yields the stated cube $\{w : \|w\|_\infty \leq \varepsilon_0\} \subseteq V$.

For the Lipschitz bound, $\mathsf{Rec}_{d,W}$ is C^1 on the closed cube, hence its derivative operator norm $L := \sup_{\|w\|_\infty \leq \varepsilon_0} \|D\mathsf{Rec}_{d,W}(w)\|_{2\to 2}$ is finite. By the mean value inequality on line segments,

$$\|\mathsf{Rec}_{d,W}(w) - \mathsf{Rec}_{d,W}(0)\|_2 \leq \sup_{t \in [0,1]} \|D\mathsf{Rec}_{d,W}(tw)\|_{2\to 2}\,\|w\|_2 \leq L\,\|w\|_2,$$

as claimed. $\square$

Theorem 3.4. *Fix $W \in \mathbb{N}$, $d \geq 0$, and $n \in \mathbb{N}$, and set $K := \lceil n/W \rceil$. Assume $K \geq 2d+1$ and that Proposition 3.3 holds with constants (ε_0, L). Given window data $w \in \mathbb{R}^K$ with $|w_k| \leq \varepsilon \leq \varepsilon_0$, let $\hat{y} := \mathsf{Rec}_{d,W}(w)$. Assume the noiseless reconstruction $\hat{y}^{(0)} := \mathsf{Rec}_{d,W}(0)$ is neutral, i.e. $P(\hat{y}^{(0)}) = 0$. Then the certification objective*

$$\Phi(\hat{y}) = \sum_{i=1}^n J\big(e^{P(\hat{y})_i}\big) = \sum_{i=1}^n (\cosh(P(\hat{y})_i) - 1)$$

satisfies the explicit quadratic bound

$$\Phi(\hat{y}) \leq \frac{1}{2} e^{L\sqrt{K}\,\varepsilon_0}\, L^2 K\, \varepsilon^2. \tag{10}$$

Proof. Let $\hat{y}^{(0)} := \mathsf{Rec}_{d,W}(0)$. By Proposition 3.3, for $\|w\|_\infty \leq \varepsilon_0$ we have

$$\|\hat{y} - \hat{y}^{(0)}\|_2 \leq L\,\|w\|_2 \leq L\sqrt{K}\,\varepsilon.$$

Since P is an orthogonal projection, $\|P(u) - P(v)\|_2 \leq \|u - v\|_2$, hence using $P(\hat{y}^{(0)}) = 0$,

$$\|P(\hat{y})\|_2 = \|P(\hat{y}) - P(\hat{y}^{(0)})\|_2 \leq \|\hat{y} - \hat{y}^{(0)}\|_2 \leq L\sqrt{K}\,\varepsilon.$$

Also $\|P(\hat{y})\|_\infty \leq \|P(\hat{y})\|_2 \leq L\sqrt{K}\,\varepsilon \leq L\sqrt{K}\,\varepsilon_0$. Applying Lemma 3.2 componentwise gives

$$\Phi(\hat{y}) = \sum_{i=1}^n (\cosh(P(\hat{y})_i) - 1) \leq \frac{1}{2} e^{\|P(\hat{y})\|_\infty}\, \|P(\hat{y})\|_2^2 \leq \frac{1}{2} e^{L\sqrt{K}\varepsilon_0}\, L^2 K\, \varepsilon^2,$$

which is (10). $\square$

Remark 3.5. In the special case $(d, W) = (3, 8)$ one may take $K = 7$ windows. Once a bound $L \leq L_0$ is established for the local reconstruction map on $\|w\|_\infty \leq \varepsilon_0$, Theorem 3.4 yields the explicit estimate

$$\Phi(\hat{y}) \leq \tfrac{1}{2} e^{L_0\sqrt{7}\,\varepsilon_0}\, L_0^2 \cdot 7\,\varepsilon^2.$$

For example, if a numerical conditioning check verifies $L_0 \leq 10$ on $\varepsilon_0 = 10^{-2}$, then $\Phi(\hat{y}) \leq 350\,\varepsilon^2$ for all $\varepsilon \leq 10^{-2}$.

Remark 3.6. The Lipschitz constant L in Proposition 3.3 can be expressed in terms of the operator norm of the inverse (or pseudoinverse) of the linearised reconstruction matrix at $w = 0$; equivalently, in terms of a condition number κ for the Vandermonde/Hankel system used by the reconstruction. Poor conditioning increases L (equivalently the relevant condition number), but the dependence on ε remains quadratic.

3.3. The certification objective.

Definition 3.7. Given a cost function $c : S \times O \to \mathbb{R}_{\geq 0}$, the *meaning map* (argmin correspondence) is

$$\mathrm{Mean}(s) := \arg\min_{o \in O} c(s, o) := \{o \in O : c(s, o) = \inf_{o' \in O} c(s, o')\}.$$

Remark 3.8. The meaning map $\mathrm{Mean}(s)$ is defined via an infimum over O and may be empty in full generality. In the intended finite-data setting one works with a *finite* candidate list O, in which case every function $O \to \mathbb{R}$ attains its minimum and $\mathrm{Mean}(s) \neq \emptyset$ for all s. Whenever we speak of "the" optimizer or of certificates that "attain their minimum exactly on $\mathrm{Mean}(s)$", we implicitly restrict to such settings (or more generally assume that $c(s, \cdot)$ attains its minimum on O).

Proposition 3.9 (Attainment on compact candidate classes)**.** *Fix $s \in S$. Assume O is compact and the map $o \mapsto c(s, o)$ is continuous. Then* $\mathrm{Mean}(s)$ *is nonempty.*

Proof. By continuity on a compact set, $c(s, \cdot)$ attains its minimum at some $o_* \in O$ (extreme value theorem). Hence

$$c(s, o_*) = \inf_{o \in O} c(s, o),$$

so $o_* \in \mathrm{Mean}(s)$ by Definition 3.7. $\qquad\qquad\square$

In many applications O is not finite but is a closed subset of a Euclidean space; in that case one can ensure attainment by a standard coercivity hypothesis.

Proposition 3.10 (Attainment under coercivity)**.** *Fix $s \in S$ and assume $O \subset \mathbb{R}^m$ is nonempty and closed for some $m \geq 1$. Assume $o \mapsto c(s, o)$ is lower semicontinuous on O and* coercive *on O in the sense that*

$$c(s, o) \to \infty \quad \text{whenever} \quad \|o\| \to \infty \quad \text{with } o \in O.$$

Then $\mathrm{Mean}(s)$ *is nonempty.*

Proof. Let $\alpha := \inf_{o \in O} c(s, o)$ and choose a minimizing sequence $(o_k) \subset O$ with $c(s, o_k) \downarrow \alpha$. By coercivity, the sequence (o_k) is bounded; thus it lies in $O \cap \overline{B_R}$ for some R. By compactness of $\overline{B_R}$, there is a convergent subsequence $o_{k_j} \to o_*$. Since O is closed, $o_* \in O$. Lower semicontinuity gives $c(s, o_*) \leq \liminf_{j \to \infty} c(s, o_{k_j}) = \alpha$, hence $c(s, o_*) = \alpha$ and $o_* \in \mathrm{Mean}(s)$. $\qquad\square$

Definition 3.11. Fix $s \in S$ and write $c_s(o) := c(s, o)$ and $\mathcal{C}_s(o) := \mathcal{C}(s, o)$. Assume throughout that the minimizer sets are nonempty (e.g. O finite, or by Proposition 3.9 or Proposition 3.10). A certificate $\mathcal{C}$ is

(i) *minimizer-sound* if $\arg\min_{o \in O} \mathcal{C}_s(o) = \arg\min_{o \in O} c_s(o)$ for every s;
(ii) *ordinally sound* if for every s and all $o_1, o_2 \in O$,

$$\mathcal{C}_s(o_1) \leq \mathcal{C}_s(o_2) \quad \Longleftrightarrow \quad c_s(o_1) \leq c_s(o_2).$$

Ordinal soundness is strictly stronger than minimizer-soundness; it asserts that $\mathcal{C}_s$ induces exactly the same preorder on candidates as the true cost c_s.

3.4. **Sound procedures and domination.** We formulate the finite-data decision task as a three-valued procedure based on window aggregates. Fix integers $W \geq 1$ and $K \geq 1$. For a signal $y = (y_n)_{n \geq 0}$ define the W-block window-sum vector

$$\mathcal{W}_{W,K}(y) := (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{K-1}) \in \mathbb{R}^K, \qquad \mathcal{W}(y)_k := \sum_{j=0}^{W-1} y_{(Wk+j)}.$$

We write the corresponding set of realizable window vectors as

$$\mathcal{W}_{W,K}(\mathcal{M}_{d,W}) := \{\mathcal{W}_{W,K}(y) : y \in \mathcal{M}_{d,W}\} \subset \mathbb{R}^K.$$

A *finite-data procedure* is a map

$$\Psi : \mathbb{R}^K \to \{\text{zero}, \text{nonzero}, \text{inconclusive}\}.$$

Definition 3.12. Fix a model class $\mathcal{M}_{d,W}$ of signals (e.g. rational/finite-state of degree $\leq d$) together with the nondegeneracy hypotheses from Theorem 2.20. Let $y_{\text{neu}} \in \mathcal{M}_{d,W}$ denote the distinguished neutral signal (one may take $y_{\text{neu}} \equiv 0$). We compare procedures only on realizable window vectors $w \in \mathcal{W}_{W,K}(\mathcal{M}_{d,W})$; outside this set a procedure may return inconclusive arbitrarily. A procedure Ψ is *sound on* $\mathcal{M}_{d,W}$ if for every $y \in \mathcal{M}_{d,W}$,

$$\Psi(\mathcal{W}_{W,K}(y)) = \text{zero} \implies y = y_{\text{neu}},$$
$$\Psi(\mathcal{W}_{W,K}(y)) = \text{nonzero} \implies y \neq y_{\text{neu}}.$$

Its *resolves set* is

$$\text{Res}(\Psi) := \{y \in \mathcal{M}_{d,W} : \Psi(\mathcal{W}_{W,K}(y)) \neq \text{inconclusive}\}.$$

Given two sound procedures Ψ_1, Ψ_2 on $\mathcal{M}_{d,W}$, we say that Ψ_1 *dominates* Ψ_2 if $\text{Res}(\Psi_2) \subseteq \text{Res}(\Psi_1)$ and

$$\Psi_1(\mathcal{W}_{W,K}(y)) = \Psi_2(\mathcal{W}_{W,K}(y)) \qquad \text{for all } y \in \text{Res}(\Psi_2).$$

Definition 3.13 (Pipeline components). Fix integers $W \geq 1$ and $K \geq 1$ and consider realizable window data $w \in \mathcal{W}_{W,K}(\mathcal{M}_{d,W})$. We use the following (possibly partial) operations.

(a) **Aggregation / reconstruction (A).** On the identifiability locus from Theorem 2.20, define a partial map $A : \mathcal{W}_{W,K}(\mathcal{M}_{d,W}) \dashrightarrow \mathcal{M}_{d,W}$ by letting $A(w)$ be the (locally) unique signal y (or parameter π) with $\mathcal{W}_{W,K}(y) = w$, when such a unique realization exists; otherwise $A(w)$ is undefined.

(b) **Projection (P).** For a positive vector/configuration $x \in (\mathbb{R}_{>0})^n$ with $y = \log x$, set $P(y)$ as in (5). In the pipeline we apply this projection in log-coordinates to the reconstructed object.

(c) **Coercive decision layer (B).** For a projected log-vector $u \in \mathbb{R}^n$ (typically $u = P(y)$), define the certificate value

$$\mathcal{B}(u) := \sum_{i=1}^{n} J(e^{u_i}),$$

and define $B(u) \in \{\text{zero}, \text{nonzero}\}$ by $B(u) = \text{zero}$ if $\mathcal{B}(u) = 0$ and $B(u) = \text{nonzero}$ otherwise.

The associated finite-data decision map is evaluated as follows: given window data w, if $A(w)$ is defined then set $x := A(w)$, form the projected log-vector $u := P(\log x)$, and output $B(u)$; if $A(w)$ is undefined we return inconclusive.

Theorem 3.14 (Forced (P, B)-dependence in the decision layer). *Let S be any set and let $D : \mathbb{R}^n \to S$ be a map that is invariant under global shifts along $\mathbf{1}$, i.e.*

$$D(y + t\mathbf{1}) = D(y) \qquad \text{for all } y \in \mathbb{R}^n,\ t \in \mathbb{R}.$$

Then there exists a unique map $\widetilde{D} : \mathbf{1}^{\perp} \to S$ such that

$$D = \widetilde{D} \circ P,$$

where P is the orthogonal projection onto $\mathbf{1}^{\perp}$ defined in (5). In particular, any defect-based decision rule whose conclusions depend only on scale-invariant information in log-coordinates can, without loss, be written as a map of the projected log-vector $u = P(\log x)$.

Proof. For any $y \in \mathbb{R}^n$ we have the orthogonal decomposition $y = P(y) + \overline{y}\,\mathbf{1}$, hence by shift-invariance,

$$D(y) = D(P(y) + \overline{y}\,\mathbf{1}) = D(P(y)).$$

Define $\widetilde{D}$ on $\mathbf{1}^{\perp}$ by $\widetilde{D}(u) := D(u)$. Then $D(y) = \widetilde{D}(P(y))$ for all y, i.e. $D = \widetilde{D} \circ P$. Uniqueness follows because P is surjective onto $\mathbf{1}^{\perp}$. The final sentence follows since scaling $x \mapsto cx$ corresponds to shifting $y = \log x$ by $\log(c)\,\mathbf{1}$ and thus leaves $P(y)$ unchanged (cf. Lemma 2.5 and Definition 2.6). $\qquad\square$

Theorem 3.15 (Structural non-redundancy on the identifiability locus)**.** *Fix $d \geq 0$, $W \geq 1$, and $K \geq 2d + 1$ and let $\mathcal{U}_{d,W}$ be an identifiability neighborhood as in Theorem 3.19.*

  (a) (Need for P at the decision layer.) Any decision map $D(\log x)$ that is sound with respect to the scale-invariant defect predicate (Definition 2.6) must be invariant under shifts along $\mathbf{1}$, hence factors through P as in Theorem 3.14.

  (b) (Need for a coercive certificate.) Any sound rule that certifies neutrality via a computable scalar statistic and is stable under two-sided perturbations must use (explicitly or implicitly) a coercive lower bound linking that statistic to $\mathrm{Def}(x)$; in our framework this role is played by the canonical certificate $\mathcal{B}(u)$ together with Theorem 2.9 and the tolerance theorems (Theorems 3.4–3.36).

  (c) (Need for inversion from windows.) On $\mathcal{U}_{d,W}$, any sound finite-data rule Ψ applied to realizable window data w can depend only on the unique realization $y = A(w)$ (when it exists); equivalently, Ψ factors through the reconstruction map A on $\mathcal{W}_{W,K}(\mathcal{U}_{d,W})$.

Proof. (a) If soundness is defined relative to the scale-invariant predicate "neutral vs. non-neutral" (Definition 2.6), then any sound decision rule must give the same output on x and cx whenever both are admissible, since they have the same defect value. In log-coordinates this is invariance under $y \mapsto y + t\mathbf{1}$, hence factorization through P follows from Theorem 3.14.

(b) The tolerance conclusions in Theorems 3.4–3.36 are derived from two-sided perturbation bounds together with a coercive inequality that forces $\mathrm{Def}(x) = 0$ when the certificate vanishes. Any procedure claiming the same quantitative stability must therefore invoke a comparable coercive implication (possibly with a different computable statistic). In the present framework, $\mathcal{B}(u)$ is the canonical computable statistic and Theorem 2.9 supplies the required implication.

(c) Let $w \in \mathcal{W}_{W,K}(\mathcal{U}_{d,W})$ and write y for the unique realization in $\mathcal{U}_{d,W}$ with $\mathcal{W}_{W,K}(y) = w$. If $\Psi(w)$ is conclusive, soundness forces that conclusion to match the truth-value of the predicate on y. Thus on the image $\mathcal{W}_{W,K}(\mathcal{U}_{d,W})$ the rule Ψ is determined by the realization y, i.e. it factors through the inverse map A whenever A is defined. $\qquad\square$

For brevity we continue to denote this overall procedure by $\Phi^{*} = A \circ B \circ P$, with the understanding that A is the reconstruction stage and that inconclusive is returned when reconstruction is not defined.

Proposition 3.16. *Fix $d \geq 0$, a window length $W \geq 1$, and $K \geq 2d + 1$. Let $\mathcal{M}_{d,W}$ denote the degree-$\leq d$ rational (finite-state) model class and let $\mathcal{W}_{W,K} : \mathcal{M}_{d,W} \to \mathbb{R}^K$ send a signal to its first K W-block window sums. On the non-degenerate identifiability locus singled*

out in Theorem 2.20 (in particular, where the reconstruction step A is well-defined by local invertibility / Hankel nondegeneracy), the canonical procedure

$$\Phi^* \;=\; A \circ B \circ P$$

is well-defined from $w = \mathcal{W}_{W,K}(y)$ and is sound in the sense of Definition 3.12.

Proof. The projection step P is the mean-zero projection in log-coordinates (Definition 2.4). By Definition 2.6, $\mathrm{Def}(x) = \|P(\log x)\|_2$, hence $\mathrm{Def}(x) = 0$ if and only if $P(\log x) = 0$. Since P is a projection, $P(P(\log x)) = P(\log x)$. The coercive bound (Theorem 2.9) implies that vanishing projected cost forces $P(\log x) = 0$, hence $\mathrm{Def}(x) = 0$, giving soundness of the B step. Finally, on the identifiability locus, the aggregation step A (reconstruction from W-block window sums) does not introduce spurious neutral realizations: if the reconstructed projected cost is 0, then the unique signal compatible with the observed w is neutral. Hence Φ^* is a sound finite-data procedure. □

Remark 3.17. Soundness prevents false positives/negatives; domination compares how often a sound procedure can decide. A dominating procedure is therefore "best possible" among sound procedures for the same finite-data.

Lemma 3.18. *Let $\mathcal{U}_{d,W} \subset \mathcal{M}_{d,W}$ be a neighborhood of y_{neu} such that the restriction $\mathcal{W}_{W,K}|_{\mathcal{U}_{d,W}}$ is injective (equivalently: for each realizable window vector $w \in \mathcal{W}_{W,K}(\mathcal{U}_{d,W})$ there is at most one $y \in \mathcal{U}_{d,W}$ with $\mathcal{W}_{W,K}(y) = w$). Let $\Psi : \mathbb{R}^K \to \{\mathrm{zero}, \mathrm{nonzero}, \mathrm{inconclusive}\}$ be sound on $\mathcal{M}_{d,W}$. Then whenever w is realized by some $y \in \mathcal{U}_{d,W}$, the value $\Psi(w)$ is forced solely by the truth-value of the underlying predicate on that unique realization (neutral vs. non-neutral), and cannot disagree with any other sound procedure on the same model class when restricted to $\mathcal{U}_{d,W}$.*

Proof. If $w = \mathcal{W}_{W,K}(y)$ and $\Psi(w) = \mathrm{zero}$, soundness forces y to be neutral; if $\Psi(w) = \mathrm{nonzero}$, soundness forces y to be non-neutral. Uniqueness of the realization for w makes these implications well-defined on the image of $\mathcal{W}_{W,K}$ and prevents contradictory sound outputs. □

We now state the *local optimality* (domination) property of the canonical finite-data decision rule. Write $\mathrm{Res}(\Psi)$ for the set of realizable instances on which a rule Ψ returns a conclusive value in $\{\mathrm{zero}, \mathrm{nonzero}\}$. We say that Φ *dominates* Ψ on a locus $\mathcal{V}$ if $\mathrm{Res}(\Psi) \cap \mathcal{V} \subseteq \mathrm{Res}(\Phi) \cap \mathcal{V}$ and the two rules agree on $\mathrm{Res}(\Psi) \cap \mathcal{V}$. Informally, the rule first projects to the mean-zero subspace, then performs a window-based test, and finally outputs a ternary decision. The conclusion below is local: domination is asserted only on the identifiability neighborhood $\mathcal{U}_{d,W}$.

Theorem 3.19. *Fix integers $d \geq 0$, $W \geq 1$, and $K \geq 2d+1$. Let $\mathcal{M}_{d,W}$ be a non-degenerate (open dense) parameter locus, and let $\mathcal{U}_{d,W} \subset \mathcal{M}_{d,W}$ be a neighborhood of y_{neu} on which the restricted window map $\mathcal{W}_{W,K}|_{\mathcal{U}_{d,W}}$ is injective. Assume:*

(a) *for every $y \in \mathcal{U}_{d,W}$ the window-sum sequence $S_k := \mathcal{W}(y)_k$ satisfies the hypotheses of Theorem 2.20, so the data $(S_0, \ldots, S_{2d-1})$ determine the induced window-sum process uniquely on $\mathcal{U}_{d,W}$; and*

(b) *the realizable datum $w = \mathcal{W}_{W,K}(y)$ determines the neutral/non-neutral truth value on $\mathcal{U}_{d,W}$, i.e.*

$$\mathcal{W}_{W,K}(y) = \mathcal{W}_{W,K}(y_{\mathrm{neu}}) \implies y = y_{\mathrm{neu}}.$$

Let $\Phi^* : \mathcal{W}_{W,K}(\mathcal{M}_{d,W}) \to \{\mathrm{zero}, \mathrm{nonzero}, \mathrm{inconclusive}\}$ denote the canonical procedure $A \circ B \circ P$ defined in this section.

Then for every sound procedure $\Psi : \mathbb{R}^K \to \{\text{zero, nonzero, inconclusive}\}$ *one has* $\mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W} \subseteq \mathrm{Res}(\Phi^*) \cap \mathcal{U}_{d,W}$, *and for every* $y \in \mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W}$ *the outputs agree on the realized datum* $w = \mathcal{W}_{W,K}(y)$:

$$\Psi(\mathcal{W}_{W,K}(y)) = \Phi^*(\mathcal{W}_{W,K}(y)).$$

Proof. Let Ψ be any sound procedure on $\mathcal{M}_{d,W}$ and let $y \in \mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W}$. Write $w := \mathcal{W}_{W,K}(y)$.

By assumption (a) and Theorem 2.20, the realized datum w determines the induced window-sum process (and hence its Prony parameters) uniquely on $\mathcal{U}_{d,W}$. Assumption (b) guarantees that this information determines whether or not $y = y_{\text{neu}}$ (equivalently: a neutral realization is compatible with w if and only if $y = y_{\text{neu}}$).

By construction, Φ^* evaluates exactly this truth value: it outputs zero exactly when the unique reconstruction compatible with w lies in the neutral *zero-defect* class, and outputs nonzero exactly when no neutral realization is compatible with w; otherwise it outputs inconclusive. Hence Φ^* is sound.

Now assume $y \in \mathrm{Res}(\Psi) \cap \mathcal{U}_{d,W}$. If $\Psi(w) = $ zero, then soundness forces $y = y_{\text{neu}}$, so the unique reconstruction from w is neutral and therefore $\Phi^*(w) = $ zero. If $\Psi(w) = $ nonzero, then $y \neq y_{\text{neu}}$, so the unique reconstruction from w is non-neutral and $\Phi^*(w) = $ nonzero. Thus whenever Ψ resolves (in the sense of Definition 3.12), Φ^* resolves with the same answer, i.e. Φ^* dominates Ψ. $\qquad\square$

Corollary 3.20 (Global domination under global injectivity). *Fix integers* $d \geq 0$, $W \geq 1$, *and* $K \geq 2d+1$. *Assume:*

(a) *the window map* $\mathcal{W}_{W,K}$ *is injective on the parameter locus* $\mathcal{M}_{d,W}$;
(b) *for every* $y \in \mathcal{M}_{d,W}$, *the induced window-sum sequence satisfies the hypotheses of Theorem 2.20;*
(c) *the realizable datum determines global neutrality on* $\mathcal{M}_{d,W}$, *i.e.*

$$\mathcal{W}_{W,K}(y) = \mathcal{W}_{W,K}(y_{\text{neu}}) \implies y = y_{\text{neu}} \quad \text{for all } y \in \mathcal{M}_{d,W}.$$

Then for every sound procedure $\Psi : \mathbb{R}^K \to \{\text{zero, nonzero, inconclusive}\}$,

$$\mathrm{Res}(\Psi) \subseteq \mathrm{Res}(\Phi^*),$$

and for all $y \in \mathrm{Res}(\Psi)$ *one has*

$$\Psi(\mathcal{W}_{W,K}(y)) = \Phi^*(\mathcal{W}_{W,K}(y)).$$

Proposition 3.21 (Sharpness outside an identifiable class). *Let* $\mathcal{C}$ *be any class of signals and let* $\mathcal{W}_{W,K}$ *be the window map. If there exist* $y_0, y_1 \in \mathcal{C}$ *with* $\mathcal{W}_{W,K}(y_0) = \mathcal{W}_{W,K}(y_1)$ *but with opposite truth-values for the target predicate (e.g.* y_0 *neutral and* y_1 *non-neutral), then no procedure* $\Psi : \mathbb{R}^K \to \{\text{zero, nonzero, inconclusive}\}$ *that is sound on* $\mathcal{C}$ *can return a conclusive decision on the common datum* $w := \mathcal{W}_{W,K}(y_0) = \mathcal{W}_{W,K}(y_1)$.

Proof. If $\Psi(w) = $ zero, soundness would force the realization to be neutral, contradicting soundness on y_1; similarly if $\Psi(w) = $ nonzero it contradicts soundness on y_0. Hence $\Psi(w)$ must be inconclusive. $\qquad\square$

Corollary 3.22 (Domination on $\mathcal{M}_{d,W}$). *Theorem 3.19 applies with* $\mathcal{U}_{d,W} := \mathcal{M}_{d,W}$.

Proof. Apply Theorem 3.19 with $\mathcal{U}_{d,W} := \mathcal{M}_{d,W}$. $\qquad\square$

3.5. **Sharpness boundary outside the rational class.** The collision principle of Proposition 3.21 becomes an explicit *sharpness boundary* once we exhibit distinct latent configurations that produce the *same* finite-window data. Here we do this directly for the window operator (7) by enlarging the model class from the rational (finite-state) family of Subsection 2.5 to the ambient class of all non-negative real sequences.

Ambient class. Let $\mathcal{C}_{\geq 0}$ be the set of all sequences $y = (y_n)_{n \geq 0}$ with $y_n \geq 0$ for all n. For fixed $W \geq 1$ and $K \geq 0$, write

$$\mathcal{W}_{W,K}(y) := (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_K) \in \mathbb{R}^{K+1}$$

for the truncated window data. Let $\mathcal{C}_{\mathrm{rat}}(d) \subset \mathcal{C}_{\geq 0}$ denote the subclass of sequences that are rational of degree at most d in the sense of Subsection 2.5.

Theorem 3.23 (Non-identifiability of rational-class membership from finite windows). *Fix integers $d \geq 1$, $W \geq 1$, and $K \geq 0$. Then the finite-data map $\mathcal{W}_{W,K}$ is not identifying for rational-class membership: for every $y_{\mathrm{in}} \in \mathcal{C}_{\mathrm{rat}}(d)$ there exists $y_{\mathrm{out}} \in \mathcal{C}_{\geq 0} \setminus \mathcal{C}_{\mathrm{rat}}(d)$ such that*

$$\mathcal{W}_{W,K}(y_{\mathrm{out}}) = \mathcal{W}_{W,K}(y_{\mathrm{in}}).$$

Proof. Let $y_{\mathrm{in}} \in \mathcal{C}_{\mathrm{rat}}(d)$ and set $N := W(K+1) - 1$, so that the data $\mathcal{W}_{W,K}(y)$ depend only on the coordinates $y_0, \ldots, y_N$. Define a non-negative "tail" sequence $r = (r_n)_{n \geq 0}$ by

$$r_n := \begin{cases} 0, & 0 \leq n \leq N, \\ 1, & n = N + m^2 \text{ for some integer } m \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Set $y_{\mathrm{out}} := y_{\mathrm{in}} + r$, so $y_{\mathrm{out}} \in \mathcal{C}_{\geq 0}$ and $y_{\mathrm{out},n} = y_{\mathrm{in},n}$ for all $0 \leq n \leq N$. Hence r contributes nothing to the first $K+1$ windows, and therefore $\mathcal{W}_{W,K}(y_{\mathrm{out}}) = \mathcal{W}_{W,K}(y_{\mathrm{in}})$.

It remains to show $y_{\mathrm{out}} \notin \mathcal{C}_{\mathrm{rat}}(d)$. Suppose toward a contradiction that $y_{\mathrm{out}} \in \mathcal{C}_{\mathrm{rat}}(d)$. Since also $y_{\mathrm{in}} \in \mathcal{C}_{\mathrm{rat}}(d)$, their difference $r = y_{\mathrm{out}} - y_{\mathrm{in}}$ has a rational generating function as well (difference of two rational functions), hence r satisfies a linear recurrence of some finite order D for all sufficiently large indices: there exist $b_1, \ldots, b_D \in \mathbb{R}$ and n_0 such that

$$r_n = b_1 r_{n-1} + \cdots + b_D r_{n-D} \qquad \text{for all } n \geq n_0. \tag{11}$$

Choose m so large that $n := N + m^2 \geq n_0$ and the gap

$$n - (N + (m-1)^2) = 2m - 1 > D.$$

Then $r_n = 1$ while $r_{n-1} = \cdots = r_{n-D} = 0$ (because there is no other square index within distance D), so (11) forces $1 = 0$, a contradiction. Therefore r is not rational, hence y_{out} cannot be rational of degree $\leq d$.

$\square$

Corollary 3.24 (Unconditional sharpness boundary outside the rational class). *Fix $d \geq 1$, $W \geq 1$, and $K \geq 0$. Let $T(y) := \mathbf{1}_{\{y \in \mathcal{C}_{\mathrm{rat}}(d)\}}$ be the predicate "y is rational of degree $\leq d$". Then there exist finite-window data $w \in \mathbb{R}^{K+1}$ for which no procedure that is sound for T (on the ambient class $\mathcal{C}_{\geq 0}$) can return a conclusive decision: every sound procedure must output INC on input w.*

Proof. Pick any $y_{\mathrm{in}} \in \mathcal{C}_{\mathrm{rat}}(d)$ and set $w := \mathcal{W}_{W,K}(y_{\mathrm{in}})$. By Theorem 3.23 there exists $y_{\mathrm{out}} \notin \mathcal{C}_{\mathrm{rat}}(d)$ with $\mathcal{W}_{W,K}(y_{\mathrm{out}}) = w$. Thus $T(y_{\mathrm{in}}) \neq T(y_{\mathrm{out}})$ while the finite-data coincide, so Proposition 3.21 forces any sound procedure to be inconclusive on w. $\square$

Fix a costed space (S, O, c) with meaning map $\mathrm{Mean}(s) = \arg\min_{o \in O} c(s, o)$. In applications one typically does not compute $\mathrm{Mean}(s)$ by scanning all $o \in O$. Instead, one seeks a *certificate functional* $\mathcal{C}$ that:

- can be evaluated efficiently from (s, o) (often through a low-dimensional proxy or windowed measurements), and
- is *sound*: $\mathcal{C}(s, o)$ attains its minimum (or maximum) exactly when $o \in \mathrm{Mean}(s)$.

The simplest sound certificates are *separable* in the sense that they depend on (s, o) only through a canonical ratio (or a monotone transform thereof), so that the certificate is equivalent to comparing values of J.

3.6. **Canonical rescalings.** We make precise what is meant by a "canonical rescaling" of a signal in the present model. Assume that the multiplicative group $(0, \infty)$ acts on S (written $\lambda \cdot s$) and that the scale map $\iota_S : S \to (0, \infty)$ is *equivariant* for this action:

$$\iota_S(\lambda \cdot s) = \lambda \, \iota_S(s) \qquad (\lambda > 0, \ s \in S).$$

We call $s' \in S$ a *canonical rescaling* of $s \in S$ if $s' = \lambda \cdot s$ for some $\lambda > 0$. This captures the idea that s and s' represent the same "configuration" up to overall magnitude.

In several arguments below we also use that the penalty J is strictly monotone on one branch. Accordingly, for a fixed $s \in S$ we say that the candidate class O lies on a *single J-branch for s* if either $\iota_S(s)/\iota_O(o) \geq 1$ for all $o \in O$ or $\iota_S(s)/\iota_O(o) \leq 1$ for all $o \in O$. On such a branch, J is strictly monotone and hence invertible.

3.7. **Forced separability for canonical reciprocal costs.** We now state the main rigidity claim in the simplest form that is directly used later. The proof is a short functional argument: the coercivity estimate together with the reciprocity structure induced by the cost assumptions pin down the allowable comparisons between candidates $o \in O$. Such order-preserving representation arguments are classical in the theory of functional equations; see [1].
This is a standard order-theoretic fact; we include it for completeness.

Lemma 3.25. *Let $c : S \times O \to [0, \infty)$ be the ratio-induced cost $c(s, o) := J(\iota_S(s)/\iota_O(o))$. Fix a certificate $\mathcal{C} : S \times O \to \mathbb{R}$. Assume that $\mathcal{C}$ is* ordinally sound *relative to c in the sense of Definition 3.11(ii). Then for each $s \in S$ there exists a strictly increasing function $\phi_s : \mathrm{im}(c_s) \to \mathbb{R}$ (where $c_s(o) := c(s, o)$) such that*

$$\mathcal{C}(s, o) = \phi_s(c(s, o)) \qquad \text{for all } o \in O.$$

In particular, on each fixed s, $\mathcal{C}(s, \cdot)$ is a strictly monotone reparametrization of $c(s, \cdot)$.

Proof. Fix $s \in S$. Ordinal soundness means $\mathcal{C}_s$ and c_s induce the same preorder on O. Define ϕ_s on $\mathrm{im}(c_s)$ by $\phi_s(r) := \mathcal{C}_s(o)$ for any o with $c_s(o) = r$; this is well-defined and strictly increasing, and yields $\mathcal{C}_s = \phi_s \circ c_s$. $\square$

Proposition 3.26. *Assume the hypotheses of Lemma 3.25. Fix $s \in S$ and assume O lies on a single J-branch for s and for every canonical rescaling $s' = \lambda \cdot s$ (in the sense of Section 3.6). Then for each such $s' = \lambda \cdot s$ there exists a strictly monotone function $\psi_{s,s'} : \mathbb{R} \to \mathbb{R}$ such that*

$$\mathcal{C}(s', o) = \psi_{s,s'}(\mathcal{C}(s, o)) \qquad \text{for all } o \in O.$$

In particular, canonical rescalings preserve the certificate ordering up to a strictly monotone reparametrization on any single J-branch.

Proof. Let $s' = \lambda \cdot s$ with $\lambda > 0$. Write $r(o) := \iota_S(s)/\iota_O(o)$ and $r'(o) := \iota_S(s')/\iota_O(o) = \lambda r(o)$. By Lemma 3.25 there exist strictly monotone functions $\phi_s, \phi_{s'}$ such that $\mathcal{C}(s, o) = \phi_s(J(r(o)))$ and $\mathcal{C}(s', o) = \phi_{s'}(J(r'(o)))$ for all $o \in O$.

By the single-branch hypothesis, the restriction of J to the relevant range of ratios is strictly monotone and hence invertible. Let J^{-1} denote this branch inverse. Define a function $F_\lambda : [0, \infty) \to [0, \infty)$ by

$$F_\lambda(t) := J(\lambda \, J^{-1}(t)).$$

Since J and J^{-1} are strictly monotone on the branch, F_λ is strictly monotone. Moreover, for every $o \in O$ we have $J(r'(o)) = J(\lambda r(o)) = F_\lambda(J(r(o)))$.

Since $\phi_s : \mathrm{im}(c_s) \to \mathrm{im}(\mathcal{C}_s)$ is strictly monotone, it is a bijection onto its image, so the inverse ϕ_s^{-1} is well-defined on $\mathrm{im}(\mathcal{C}_s) = \{\mathcal{C}(s, o) : o \in O\}$. In particular, $\phi_s(J(r(o))) = \mathcal{C}(s, o)$

lies in this domain for every $o \in O$. Therefore

$$\mathcal{C}(s', o) = \phi_{s'}(J(r'(o))) = \phi_{s'}(F_\lambda(J(r(o)))) = \underbrace{(\phi_{s'} \circ F_\lambda \circ \phi_s^{-1})}_{=: \; \psi_{s,s'}}(\phi_s(J(r(o)))) = \psi_{s,s'}(\mathcal{C}(s, o)).$$

The composition $\psi_{s,s'} = \phi_{s'} \circ F_\lambda \circ \phi_s^{-1}$ is strictly monotone, as required. $\qquad\square$

Proposition 3.27 (Primitive ratio hypotheses imply the uniqueness bundle). *Let* $c(s, o) := J(\iota_S(s)/\iota_O(o))$ *and* $\mathcal{C} : S \times O \to \mathbb{R}$. *Assume:*

(P1) **Ratio dependence:** *if* $\iota_S(s_1)/\iota_O(o_1) = \iota_S(s_2)/\iota_O(o_2)$, *then* $\mathcal{C}(s_1, o_1) = \mathcal{C}(s_2, o_2)$.
(P2) **Cost-to-ratio determinacy on the realized set:** *if* $c(s_1, o_1) = c(s_2, o_2)$, *then* $\iota_S(s_1)/\iota_O(o_1) = \iota_S(s_2)/\iota_O(o_2)$.
(P3) **State-ratio collapse at fixed observation:** *for all* $s_1, s_2 \in S$ *and* $o \in O$,

$$\frac{\iota_S(s_1)}{\iota_O(o)} = \frac{\iota_S(s_2)}{\iota_O(o)}.$$

Then the hypotheses (H1) and (H2) of Theorem 3.31 hold.

Proof. (H1) follows from (P2) then (P1): equal costs imply equal ratios, hence equal certificate values. For (H2), fix s_1, s_2, o. By (P3), the ratios for (s_1, o) and (s_2, o) are equal; applying (P1) yields $\mathcal{C}(s_1, o) = \mathcal{C}(s_2, o)$. $\qquad\square$

Remark 3.28 (On the strength of (P3)). Since $\iota_O(o) > 0$, hypothesis (P3) implies $\iota_S(s_1) = \iota_S(s_2)$ for all $s_1, s_2 \in S$, i.e. ι_S is constant. In this regime the ratio $r(o) = \iota_S(s)/\iota_O(o)$ and hence the cost $c(s, o) = J(r(o))$ are *independent of s*. Accordingly, (H2) is automatic and the "cross-state" content of Theorem 3.31 becomes immediate. If a nontrivial bridge from primitives to (H1)–(H2) is intended, (P3) should be weakened.

Proposition 3.29 (Reducing (H1) to ordinal soundness plus cross-state calibration). *Assume that for each $s \in S$ the candidate class O lies on a single J-branch for s (as in Section 3.6), so that $o \mapsto c(s, o)$ is strictly monotone in the ratio $\iota_S(s)/\iota_O(o)$. If $\mathcal{C}$ is ordinally sound (Definition 3.11), then for every $s \in S$ there exists a strictly increasing map $f_s : c_s(O) \to \mathbb{R}$ such that*

$$\mathcal{C}(s, o) = f_s(c(s, o)) \qquad (o \in O),$$

where $c_s(O) := \{c(s, o) : o \in O\}$. Moreover, hypothesis (H1) of Theorem 3.31 holds if and only if the family $\{f_s\}_{s \in S}$ is state-independent on overlaps in the sense that for all $s_1, s_2 \in S$,

$$f_{s_1}(r) = f_{s_2}(r) \qquad \text{for all } r \in c_{s_1}(O) \cap c_{s_2}(O).$$

In particular, a sufficient condition for (H1) is that $c_s(O)$ contains a common interval $I \subset \mathbb{R}$ for all s and $f_s|_I$ is independent of s.

Proof. Fix s. Ordinal soundness says that $\mathcal{C}_s$ and c_s induce the same preorder on O. Since c_s is a real-valued function, this is equivalent to the existence of a strictly increasing reparameterization f_s on the range $c_s(O)$ with $\mathcal{C}_s = f_s \circ c_s$. (The construction is canonical: define $f_s(r) := \mathcal{C}(s, o)$ for any o with $c(s, o) = r$; this is well-defined because $c(s, o_1) = c(s, o_2)$ implies $\mathcal{C}(s, o_1) = \mathcal{C}(s, o_2)$ by the "$\Leftrightarrow$" in ordinal soundness.) For the characterization of (H1), note that if $c(s_1, o_1) = c(s_2, o_2) = r$ then

$$\mathcal{C}(s_1, o_1) = f_{s_1}(r), \qquad \mathcal{C}(s_2, o_2) = f_{s_2}(r),$$

so $\mathcal{C}(s_1, o_1) = \mathcal{C}(s_2, o_2)$ for all such pairs if and only if $f_{s_1}(r) = f_{s_2}(r)$ on every overlap value $r \in c_{s_1}(O) \cap c_{s_2}(O)$. The final sufficient condition is immediate. $\qquad\square$

Remark 3.30 (On (H2) and rescaling). Hypothesis (H2) of Theorem 3.31 is a genuine *state-rigidity* condition: it asserts that the certificate is constant across S at each fixed observation o. In applications this typically follows from an explicit invariance or normalization principle for the certificate, rather than from ordinal soundness alone.

Theorem 3.31 (Global forced factorization with uniqueness). *Let $c : S \times O \to [0, \infty)$ be the ratio-induced cost and let $\mathcal{C} : S \times O \to \mathbb{R}$ be a certificate. Assume:*

(H1) ***Cross-state cost dependence:*** *(A convenient sufficient route to verify* (H1) *from ordinal soundness plus a single-branch calibration is given in Proposition 3.29.) for all $s_1, s_2 \in S$ and $o_1, o_2 \in O$,*

$$c(s_1, o_1) = c(s_2, o_2) \implies \mathcal{C}(s_1, o_1) = \mathcal{C}(s_2, o_2).$$

(H2) ***State rigidity:*** *for all $s_1, s_2 \in S$ and $o \in O$,*

$$\mathcal{C}(s_1, o) = \mathcal{C}(s_2, o).$$

Fix $s_* \in S$ and define

$$\mathcal{I}_c := \{r \in \mathbb{R} : \exists s \in S, \exists o \in O, \ r = c(s, o)\}.$$

Then:

(a) *there exists a unique map $\phi : \mathcal{I}_c \to \mathbb{R}$ such that*

$$\mathcal{C}(s, o) = \phi(c(s, o)) \qquad \text{for all } s \in S, \ o \in O;$$

(b) *there exists a unique map $\psi : O \to \mathbb{R}$ such that*

$$\mathcal{C}(s, o) = \psi(o) \qquad \text{for all } s \in S, \ o \in O.$$

Proof. For (a), define ϕ on $\mathcal{I}_c$ by choosing, for each $r \in \mathcal{I}_c$, any pair (s_r, o_r) with $c(s_r, o_r) = r$ and setting $\phi(r) := \mathcal{C}(s_r, o_r)$. Assumption (H1) makes this well-defined. If $r = c(s, o)$, then (H1) gives $\mathcal{C}(s, o) = \mathcal{C}(s_r, o_r) = \phi(r)$, so the representation holds. If $\widetilde{\phi}$ is another map with the same representation property, then for any $r \in \mathcal{I}_c$ and any witness (s, o) with $c(s, o) = r$,

$$\widetilde{\phi}(r) = \mathcal{C}(s, o) = \phi(r),$$

hence $\widetilde{\phi} = \phi$.

For (b), define $\psi(o) := \mathcal{C}(s_*, o)$. By (H2), for every $s \in S$ one has $\mathcal{C}(s, o) = \mathcal{C}(s_*, o) = \psi(o)$. If $\widetilde{\psi}$ also satisfies $\mathcal{C}(s, o) = \widetilde{\psi}(o)$ for all s, o, then evaluating at $s = s_*$ gives $\widetilde{\psi}(o) = \mathcal{C}(s_*, o) = \psi(o)$ for every o, so $\widetilde{\psi} = \psi$. $\square$

Corollary 3.32 (Forced-factorization uniqueness from primitive hypotheses). *Under the assumptions of Proposition 3.27, the uniqueness conclusions (a) and (b) of Theorem 3.31 hold.*

Proof. By Proposition 3.27, assumptions (H1)–(H2) in Theorem 3.31 are satisfied, so the result follows directly. $\square$

Remark 3.33 (Standalone status of the uniqueness theorem). Theorem 3.31 and Corollary 3.32 are proved entirely by the arguments in this section.

3.8. **ε-optimal certification from windowed measurements.** We record an ε-robust variant that is useful when only approximate ratios are available (e.g., from windowed measurements) and the optimizer may not be unique.

Definition 3.34. For $\varepsilon \geq 0$ define the ε-meaning set by

$$\mathrm{Mean}_\varepsilon(s) := \left\{ o \in O : \ c(s, o) \leq \inf_{o' \in O} c(s, o') + \varepsilon \right\}.$$

Lemma 3.35. *Let $0 < a \leq b < \infty$ and let $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ on $(0, \infty)$. Then J is Lipschitz on $[a, b]$ with constant*

$$L(a) := \frac{1}{2}\left(1 + a^{-2}\right), \qquad i.e. \qquad |J(x) - J(y)| \leq L(a)\,|x - y| \;\; for\ all\ x, y \in [a, b].$$

In particular, if $|\widehat{r}/r - 1| \leq \delta$ and $r \in [a, b]$, then

$$|J(\widehat{r}) - J(r)| \leq L(a)\,\delta\,b.$$

Proof. For $x > 0$ we have $J'(x) = \frac{1}{2}(1 - x^{-2})$, hence for $x \in [a, b]$,

$$|J'(x)| \leq \tfrac{1}{2}\left(1 + a^{-2}\right) = L(a).$$

By the mean value theorem, $|J(x) - J(y)| \leq \sup_{[a,b]} |J'|\,|x - y| \leq L(a)|x - y|$. If $|\widehat{r}/r - 1| \leq \delta$ and $r \in [a, b]$, then $|\widehat{r} - r| \leq \delta r \leq \delta b$, giving the final bound. $\qquad \square$

Theorem 3.36. *Assume the hypotheses of Lemma 3.25. Fix $s \in S$ and suppose that the relevant ratios*

$$r(o) := \iota_S(s)/\iota_O(o)$$

lie in a known band $[a, b] \subset (0, \infty)$ for all candidates o under consideration. Suppose we observe K windowed measurements that allow us to estimate $r(o)$ by $\widehat{r}(o)$ with relative error at most δ, uniformly in o, i.e. $|\widehat{r}(o)/r(o) - 1| \leq \delta$. Then the proxy ranking obtained by replacing $r(o)$ with $\widehat{r}(o)$, i.e. by minimizing the perturbed costs $\widehat{c}(o) := J(\widehat{r}(o))$, can only localize candidates within $\mathrm{Mean}_{2\varepsilon}(s)$, where one may take the explicit bound

$$\varepsilon \;=\; L(a)\,\delta\,b, \qquad L(a) := \tfrac{1}{2}(1 + a^{-2}),$$

so in particular $\varepsilon(\delta) \to 0$ as $\delta \to 0$ (for fixed a, b).

Proof. Let $\widehat{r}(o)$ be the estimated ratio for candidate o, so that $|\widehat{r}(o)/r(o) - 1| \leq \delta$ where $r(o) = \iota_S(s)/\iota_O(o)$. Since $r(o) \in [a, b]$ by hypothesis, Lemma 3.35 gives the uniform bound

$$|J(\widehat{r}(o)) - J(r(o))| \leq \varepsilon \qquad \text{with} \qquad \varepsilon = L(a)\,\delta\,b.$$

Thus ranking candidates using the perturbed costs $J(\widehat{r}(o))$ may change the true costs by at most ε in each direction. If $\widehat{o}$ minimizes the perturbed cost $\widehat{c}(o) := J(\widehat{r}(o))$, then for any $o \in O$,

$$c(s, \widehat{o}) \leq \widehat{c}(\widehat{o}) + \varepsilon \leq \widehat{c}(o) + \varepsilon \leq c(s, o) + 2\varepsilon.$$

Hence any sound selection based on these measurements can only certify candidates within $\mathrm{Mean}_{2\varepsilon}(s)$ (Definition 3.34). $\qquad \square$

Remark 3.37 (Standalone status of the ε-noise bound). Theorem 3.36 follows directly from Lemma 3.35 and the perturbation estimate proved above.

4. Applications: real-world motivated case studies (synthetic data)

This section addresses the "applications without data" concern by including *reproducible synthetic datasets* that are motivated by standard real-world measurement settings where only short-window aggregates are available. We emphasize that the paper's main results remain theorem-level statements *conditional on the modeling hypotheses*; the datasets below are not claimed to be empirical validations, but rather concrete end-to-end instantiations of the measurement operator and the tolerance bounds.

4.1. **Case study A: multi-exponential decay under gated integration.**
Problem definition. We observe only short-window *aggregated* measurements (gated sums) from a positive latent system with a small number of modes; the raw per-timepoint signal is unavailable and measurement noise perturbs the aggregates.

Goal. Using only the aggregated windows, certify whether the configuration is *near a zero-defect (balanced) state* up to tolerance ε, and, when identifiable, recover a consistent latent description in the rational (degree-$\leq d$) class.

How the theorems apply (workflow).

(1) **Model and choose** (d, W)**.** Fix a degree bound d for the rational class and a window size W matching the instrument gate; this pins down the finite-data map $F_{d,W}$ from latent parameters to window sums.

(2) **Form finite-data.** From the raw gated readouts, compute the window vector $y^{(W)} = (\mathcal{W}(y)_0, \ldots, \mathcal{W}(y)_{2d})$; estimate a noise level η from replicates or device specs to set ε.

(3) **Run certification.** Apply the ε-certification theorem (Theorem 3.36) to decide YES/NO/INC for "zero-defect"; the conclusion is sound under the stated Lipschitz/perturbation bounds.

(4) **Interpret the output.** A YES certificate guarantees the underlying state lies in the ε-meaning set; INC is expected near identifiability boundaries or outside the rational class, by the sharpness barrier.

Scenario (real-world motivation). In fluorescence lifetime experiments and related spectroscopy settings, the underlying intensity trace is often well-approximated by a *mixture of exponentials* (multiple decay modes), while instrumentation reports *integrated counts over short time gates*. A standard toy model is

$$y_n = \sum_{j=1}^{d} w_j a_j^n, \qquad w_j > 0, \quad 0 < a_j < 1,$$

so that the sequence (y_n) has a rational generating function (a classical Prony/ARMA-type model). The measured datum is the W-block aggregation $\mathcal{W}(y)_k := \sum_{j=0}^{W-1} y_{Wk+j}$.

Synthetic dataset. We take $d = 3$, $W = 8$, with ground-truth parameters

$$(a_1, a_2, a_3) = (0.831127, \ 0.872789, \ 0.853477),$$
$$(w_1, w_2, w_3) = (0.522164, \ 0.195934, \ 0.281902),$$

(normalized so $\sum_j w_j = 1$), and generate window sums for $k = 0, \ldots, 11$. We then perturb each window sum by 1% multiplicative Gaussian noise to produce an "observed" dataset. Table 3 reports both the true and observed window sums (seed 20260225 for reproducibility).

How the theory applies. Given observed windows $\widehat{\mathcal{W}}(y)_k$, the reconstruction/certification pipeline of the paper applies as follows: (i) check the identifiability/nondegeneracy hypotheses (the locus $\mathcal{M}_{d,W}$ and the required Jacobian-rank conditions); (ii) compute the corresponding certificate statistic $\mathcal{C}$ from the finite-window data as defined in the main body; (iii) apply Theorem 3.36 (and its corollaries) to translate a bound on relative measurement error into a tolerance parameter ε and therefore an explicit meaning-set enlargement $\mathrm{Mean}_{2\varepsilon}(s)$. This makes the pass/fail/inconclusive outcome explicit at the chosen noise level.

4.2. **Case study B: segmented response decay in a marketing funnel.**

Problem definition. A funnel (impressions→clicks→purchases) produces only coarse, windowed counts (e.g., hourly totals). Latent positive parameters encode stage-wise conversion/decay rates, but the full event-level trajectory is not stored.

Goal. Given a short sequence of aggregated windows, (i) test whether behavior is consistent with a calibrated "neutral" configuration within tolerance, and (ii) when feasible, certify identifiability of the latent rates from the available windows.

How the theorems apply (workflow).

TABLE 3. Case study A (multi-exponential decay): W-block window sums. "Observed" values are the true window sums with 1% multiplicative Gaussian noise (synthetic, seed 20260225).

k	$\mathcal{W}(y)_k$ (true)	$\mathcal{W}(y)_k$ (observed)
0	4.791914	4.754830
1	1.276888	1.303021
2	0.349197	0.351327
3	0.098038	0.097663
4	0.028236	0.028026
5	0.008329	0.008326
6	0.002510	0.002474
7	7.71e-04	7.69e-04
8	2.41e-04	2.41e-04
9	7.61e-05	7.78e-05
10	2.44e-05	2.42e-05
11	7.87e-06	7.93e-06

(1) **Define the latent state.** Encode stage rates as a positive vector π and map it to a windowed observable via the paper's recurrence/aggregation model, producing the finite-data map $F_{d,W}$.

(2) **Aggregate consistently.** Compute window sums from logs using the same W as in the model; verify that preprocessing (deduping, timezone, bot filtering) is fixed, since changing it changes $F_{d,W}$.

(3) **Check identifiability then certify.** Use the rank/identifiability results to justify when inversion is meaningful; then apply Theorem 3.36 to obtain a sound certificate (or an inconclusive result when information is insufficient).

(4) **Operationalize.** Treat YES as a robust green-light for "near-neutral" behavior under noise; treat INC as an alert that more windows, a different W, or a different model class is required.

Scenario (real-world motivation). In email or ad campaigns, response counts often decay over time and can be modeled as a mixture of two segments (fast responders vs. slow responders). Suppose y_n is the expected response rate at time-bin n and only windowed totals are available due to aggregation in the reporting pipeline.

Synthetic dataset. We take $d = 2$, $W = 6$ with

$$(a_1, a_2) = (0.904182,\ 0.877627),$$
$$(w_1, w_2) = (0.801912,\ 0.198088),$$

and generate observed window sums (with 2% multiplicative Gaussian noise). For quick reference, the first eight observed windows are:

$$\widehat{\mathcal{W}}(y)_0, \ldots, \widehat{\mathcal{W}}(y)_7 = (4.578368,\ 2.433641,\ 1.304007,\ 0.686328,$$
$$0.380817,\ 0.197523,\ 0.104636,\ 0.060241).$$

This provides a concrete instance of short-window aggregation with positive latent parameters and explicitly controlled noise.

4.3. **Case study C: battery/sensor discharge with coarse logging.**

Problem definition. Embedded devices often store only rolling-window averages/sums to save power and bandwidth. The underlying positive signal may drift slowly, but only aggregated windows are retained, with quantization and missingness.

Goal. From a short batch of aggregated windows, certify whether drift is within an acceptable tolerance band around a nominal configuration, while cleanly separating "insufficient information" from a definitive violation.

How the theorems apply (workflow).

1. **Calibrate the nominal model.** Use a lab-calibrated baseline to specify the nominal π_0 and the relevant degree/window parameters; this fixes the target "zero-defect" meaning set and defect functional.

2. **Compute windows and error budget.** Convert device logs to the window vector required by $F_{d,W}$; bound quantization/missingness as an effective noise level to select ε for certification.

3. **Certify or defer.** Apply Theorem 3.36; a YES result yields a sound tolerance guarantee, whereas INC indicates that the stored aggregates cannot resolve the decision at the required confidence.

Scenario (real-world motivation). Battery discharge and certain sensor drift processes are commonly approximated by a small number of exponential modes, but embedded devices frequently store only coarse aggregates (e.g. sums over fixed logging intervals to save bandwidth).

Synthetic dataset (outline). One may use the same mixture-of-exponentials construction as in Case A with $d = 3$ and an application-specific window length W (e.g. $W = 10$ samples per logging block). The finite-window theory applies verbatim: if the modeled class is identifiable at (d, W) and the observed aggregates have relative error bounded by δ, then Theorem 3.36 gives an explicit tolerance level ε and therefore a principled "certified vs. inconclusive" boundary for neutral-vs-nonneutral detection.

Limitations (unchanged). These case studies demonstrate *how* to instantiate the measurement map and noise model with concrete numbers, but they do not replace empirical validation. For genuine real-data applications one must justify that the data generation mechanism is well-approximated by a finite-degree rational/Prony model and that the windowed aggregation operator matches the instrument/reporting pipeline.

Appendix

The Recognition Composition Law: axioms and uniqueness. In this appendix we record the self-contained consequences of the Recognition Composition Law used in the paper. Under the stated axioms and regularity assumptions, we derive normalization and reciprocity, and then prove the corresponding uniqueness statement for the reciprocal cost function.

A.1. **Standing axioms and assumptions.** *Standing notation.* Throughout, $J : (0, \infty) \to \mathbb{R}$ denotes the scalar cost function (assumed non-constant and continuous unless stated otherwise); in this appendix we record the axioms imposed on J. *Standing notation.* Throughout, $J : (0, \infty) \to \mathbb{R}$ denotes the scalar cost function (assumed non-constant and continuous unless stated otherwise); in this appendix we record the axioms imposed on J.

Axiom 1 (Recognition Composition Law). For all $x, y > 0$,

$$J(xy) + J(x/y) = 2J(x) + 2J(y) + 2J(x)J(y). \qquad \text{(RCL)}$$

Axiom 2 (Calibration in log-coordinates). Define $J_{\log}(t) := J(e^t)$. We assume $J_{\log}$ is twice differentiable at $t = 0$ and satisfies $J''_{\log}(0) = 1$.

A.2. **Derived normalization and reciprocity.**

Proposition A.1 (Normalization is forced). *If $J : (0, \infty) \to \mathbb{R}$ is non-constant and satisfies the RCL* (RCL), *then $J(1) = 0$.*

Proof. Setting $x = y = 1$ in (RCL) gives

$$2J(1) = 2J(1)^2 + 4J(1),$$

so $J(1)^2 + J(1) = 0$ and therefore $J(1) \in \{0, -1\}$. If $J(1) = -1$, substituting $y = 1$ into (RCL) yields $J(x) \equiv -1$ for all $x > 0$, contradicting non-constancy. Hence $J(1) = 0$. $\square$

Corollary A.2 (Reciprocity is forced). *Under the same hypotheses, $J(x) = J(1/x)$ for all $x > 0$.*

Proof. Set $x = 1$ in (RCL) and use Proposition A.1. $\square$

A.3. **Main uniqueness theorem.**

Theorem A.3. *Assume that $J : (0, \infty) \to \mathbb{R}$ is continuous, non-constant, and satisfies Axioms 1–2 of Section A.1. Then the cost is forced to be the canonical reciprocal form*

$$J(x) = \tfrac{1}{2}(x + x^{-1}) - 1 = \frac{(x-1)^2}{2x}.$$

In particular, the closed form is not an additional modeling choice: it is determined by the composition identity together with the local quadratic calibration at balance.

Proof. Define $g(x) := 1 + J(x)$ and pass to log-coordinates by setting $h(t) := g(e^t)$. A direct substitution of (RCL) shows that h satisfies the classical d'Alembert functional equation

$$h(s + t) + h(s - t) = 2h(s)h(t) \qquad (s, t \in \mathbb{R}).$$

By Proposition A.1 we have $J(1) = 0$, hence $h(0) = g(1) = 1$. The classification of continuous solutions to d'Alembert's equation with $h(0) = 1$ is standard (see, for example, [1]): there exists $\lambda \geq 0$ such that either $h(t) = \cos(\lambda t)$ or $h(t) = \cosh(\lambda t)$.

The calibration Axiom 2 says that $J_{\log}(t) = J(e^t)$ has second derivative 1 at $t = 0$. Since $J_{\log}(t) = h(t) - 1$, this gives $h''(0) = 1$. For the cosine branch one has $h''(0) = -\lambda^2 \leq 0$, which is incompatible with $h''(0) = 1$. Hence we are in the hyperbolic branch and $h(t) = \cosh(\lambda t)$. Imposing $h''(0) = \lambda^2 = 1$ yields $\lambda = 1$, so $h(t) = \cosh(t)$.

Finally, returning to multiplicative coordinates, $g(x) = h(\ln x) = \cosh(\ln x) = \tfrac{1}{2}(x + x^{-1})$ and therefore $J(x) = g(x) - 1 = \tfrac{1}{2}(x + x^{-1}) - 1 = \frac{(x-1)^2}{2x}$. $\square$

A.4. **Derived properties.**

Lemma A.4 (Calibration at 1 vs. 0 in log-coordinates). *Assume J is differentiable at 1 and satisfies reciprocity $J(x) = J(1/x)$. Then $J'(1) = 0$. If moreover J is twice differentiable at 1, then for $J_{\log}(t) := J(e^t)$ one has $J''_{\log}(0) = J''(1)$. In particular, the condition $J''_{\log}(0) = 1$ is equivalent to $J''(1) = 1$.*

Proof. Differentiating $J(x) = J(1/x)$ at $x = 1$ gives $J'(1) = -J'(1)$, hence $J'(1) = 0$. By the chain rule, $J'_{\log}(t) = J'(e^t)e^t$ and $J''_{\log}(t) = J''(e^t)e^{2t} + J'(e^t)e^t$. Evaluating at $t = 0$ and using $J'(1) = 0$ yields $J''_{\log}(0) = J''(1)$. $\square$

Once the closed form is established, standard structural properties become immediate: the cost is reciprocal ($J(x) = J(1/x)$), non-negative with equality only at $x = 1$, strictly convex in log-coordinates near balance, and it diverges as $x \to 0^+$ or $x \to +\infty$. The main paper uses only the explicit formula and the fact that it is forced by Axiom 1, non-constancy, and the local calibration in Axiom 2.

Appendix B. Finite-field Jacobian matrix for $(d, W) = (3, 8)$

For completeness, we record the integer matrix $DF_{3,8}(\pi_0)$ used in the finite-field determinant check in Lemma 2.16.

$$DF_{3,8}(\pi_0) = \begin{pmatrix} 1 & 7 & -3 & -9 & 69 & -54 & 3 \\ 0 & 1008 & -956 & -1388 & 25920 & -20844 & 5568 \\ 0 & 175456 & -190016 & -200544 & 5088848 & -4927712 & 1787216 \\ 0 & 28857344 & -34167296 & -27911296 & 755009920 & -959230336 & 430318144 \\ 0 & 4538167296 & -5754321920 & -3726310400 & 82750894080 & -165775961088 & 89416058880 \\ 0 & 686488772608 & -922559823872 & -473024225280 & 3615042621440 & -26067531849728 & 16866417889280 \\ 0 & 100127404457984 & -141965037535232 & -56107543330816 & -1341867460689920 & -3737714520064000 & 2956546669428736 \end{pmatrix}$$

Conclusion

We developed a finite-data framework that certifies whether a positive configuration is neutral (zero-defect) from a short list of aggregated observations. The main contribution is a canonical decision pipeline $\Phi^* = A \circ B \circ P$ with distinct roles: P removes scale ambiguity by projecting to mean-zero log-coordinates; B turns canonical reciprocal cost into an explicit defect bound via a sharp coercivity inequality; and A reconstructs the induced window-sum process on a standard non-degenerate locus (local invertibility, equivalently Hankel/Prony nondegeneracy). On this locus, the first K window sums uniquely determine the window process, yielding sound and locally maximal certification: any sound procedure that resolves must agree with Φ^* near the neutral realization. We also introduce an ε-tolerant noise model and stability bounds that quantify how small window perturbations affect certified defect guarantees.

These results clarify what finite windows can—and cannot—decide. Reconstruction requires nondegeneracy; outside that regime, the correct outcome is "inconclusive," not a potentially misleading recovery claim. Future directions include extending reconstruction beyond separable exponential-sum models, strengthening robustness under weaker separation assumptions, and developing practical guidelines for monitoring when only pooled or privacy-preserving aggregates are available.

Acknowledgements

References

[1] J. Aczél. *Lectures on Functional Equations and Their Applications*. Academic Press, 1966.

[2] Washburn, J.; Rahnamai Barghi, A. Reciprocal Convex Costs for Ratio Matching: Axiomatic Characterization. *Axioms* **2026**, *15*(2), 151. doi:10.3390/axioms15020151.

[3] Washburn, J.; Zlatanović, M.; Allahyarov, E. *Recognition Geometry. Axioms* 2026, 15, 90.

[4] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273-299, 1965.

[5] L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.*, 7(3):200-217, 1967.

[6] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, 14(5):877-898, 1976.

[7] M. R. Osborne and G. K. Smyth. A modified Prony algorithm for exponential function fitting. *SIAM J. Sci. Comput.*, 16(1):119–138, 1995. doi:10.1137/0916008.

[8] V. V. Peller. *Hankel Operators and Their Applications*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[9] N. I. Akhiezer. *The Classical Moment Problem and Some Related Questions in Analysis*. Oliver & Boyd, Edinburgh, 1965.

[10] K. Schmüdgen. Ten Lectures on the Moment Problem. *arXiv preprint* arXiv:2008.12698, 2020.

[11] R. E. Curto and L. A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.* **119** (1996), no. 568, x+52.

Recognition Physics Research Institute, Austin, TX USA
Email address: `jon@recognitionphysics.org`

Recognition Physics Research Institute, Austin, TX USA
Email address: `arahnamab@gmail.com`