Journal of Optimization Theory and Applications manuscript No.
(will be inserted by the editor)

The Coercive Projection Theorem for Canonical Reciprocal Costs

Jonathan Washburn · Amir Rahnamai Barghi

Received: date / Accepted: date

Abstract This paper studies a finite-window certification problem for positive configurations under a canonical reciprocal cost. The main question is whether a configuration is exactly neutral, or is provably close to neutrality, when only short aggregated observations are available. We show that the underlying scalar cost is forced by the Recognition Composition Law together with a local calibration at balance, and we use this structure to build a canonical certification pipeline. The pipeline has three components: a reconstruction step that is valid on a nondegenerate identifiability locus for the window map, a projection step that removes scale ambiguity in logarithmic coordinates, and a coercivity step that converts cost control into an explicit defect bound. Two structural results delimit where such certification is possible at all. First, block aggregation annihilates root-of-unity modes and renders the neutral parameter fiber critical for the window map, so exact neutrality cannot be soundly certified from window sums when the latent dynamics are jointly unknown; every sound procedure must then return an inconclusive verdict on neutral data. Second, when the latent dynamics are known, the window map becomes linear in the state and we prove an exact certifiable/uncertifiable boundary: certification from $K \geq d + 1$ windows is sound and complete precisely when no latent mode is a nontrivial W-th root of unity and distinct modes remain separated after W-fold aliasing. On the certifiable locus the resulting procedure is sound and locally maximal among sound rules: any sound rule that resolves a datum must agree with the canonical output, and no sound rule can resolve strictly more nearby cases. We also give an epsilon-tolerant version of the construction for noisy window sums, with explicit stability constants given by the conditioning of the linear reconstruction.

Keywords Canonical Reciprocal Cost; Coercivity; Finite-Window Measurements; Local Identifiability; Sound Certification

Mathematics Subject Classification (2020) Primary 90C25; Secondary 47H05, 93B30, 94A12.

J. Washburn
Recognition Physics Research Institute, Austin, TX USA
E-mail: jon@recognitionphysics.org

A. Rahnamai Barghi: Corresponding Author
Recognition Physics Research Institute, Austin, TX USA
E-mail: arahnamab@gmail.com

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 [8].

An audit reading of the problem. It helps to read the question operationally. Think of the latent sequence as a ledger whose entries record exchanges, of the neutral configuration as the state in which the ledger balances exactly, and of the window sums as the only records an auditor retains. The certification question is then: can exact balance be audited from the retained totals? Our two boundary results give this reading a precise and perhaps unexpected answer. If the rules governing the ledger are themselves unknown, exact balance can never be certified from totals alone (Corollary 2.30): there always exist genuinely unbalanced ledgers whose totals agree with a balanced one on every retained record. If the rules are known, the audit becomes possible, and $d+1$ totals suffice exactly when the reporting period does not alias the ledger's own cycles (Theorem 2.31). The remainder of the paper determines what the audit must look like when it is possible: the cost, the projection, and the decision layer are forced by the axioms, and the resulting auditor is the strongest sound one (Theorem 3.20).

1.1 Related Work and the Research Gap

Three bodies of work border the problem treated here; none of them answers it.

Axiomatic cost characterization. The scalar cost $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ used throughout is not a modeling choice of this paper. Its uniqueness under the composition law and a curvature calibration is established in [9]; the inevitability of the composition law itself among low-degree symmetric combiners is proved in [10], and the classification of admissible reciprocally symmetric costs in [11]. A complementary axiomatization for ratio-matching decision problems appears in [7], and the discrete ledger dynamics that motivate the conservation constraint and the block-cycle observation model are axiomatized in [12]. Those papers settle *which* cost governs the comparison; they do not address what can be decided about a configuration from finitely many aggregated observations. That decision problem is the subject of the present paper.

Structured reconstruction from few samples. Recovering an exponential or rational signal from few linear measurements is classical: Prony-type exponential fitting [4], subspace methods such as ESPRIT [13], Hankel operator theory and Kronecker-type finite-rank characterizations [5], finite-rate-of-innovation sampling [14], and super-resolution under separation conditions [15]. These results are estimation theorems: they recover generic parameters on a nondegenerate locus. The certification problem is different in two ways. First, the target is a single exact point (the neutral configuration), not a generic parameter, and we show below that in the jointly parameterized model the neutral target sits precisely where the generic-identifiability machinery fails: the neutral fiber is a critical set of the window map (Lemma 2.27), and block aggregation annihilates root-of-unity modes outright (Theorem 2.28), a multirate-aliasing phenomenon [16]. Second, the object of interest is the decision rule itself: we ask which certification rules are compatible with soundness, and prove rigidity and local-maximality statements for the canonical rule, in the spirit of order-preserving representation theory [1] rather than of estimation error bounds.

Projection and proximal tools. The convex-analytic shape of the pipeline echoes Moreau's proximity operator [3], Bregman projections [2], and Rockafellar's proximal-point framework [6]; Section 1.2 details the correspondence and the difference (the classical theory offers a family of legitimate schemes, whereas here the scalar cost and the decision layer are forced by the axioms).

The gap. What is missing in the literature is a soundness-first treatment of exact-neutrality certification from block-aggregated data: an explicit statement of where such certification is possible at all, an exact boundary between the certifiable and uncertifiable regimes, and a canonicity analysis of the decision rule on the certifiable side. Those are the contributions of this paper. Certification here is a deterministic compatibility notion; we do not treat statistical hypothesis testing [17], and the zero-defect terminology is used in the deterministic sense rather than the process-control sense [18].

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, and the reconstruction stage on that locus is a Prony-type solve. The identifiability results of Section 2.3 should therefore be read as a specialization of that classical machinery to the block-sum measurement operator, with the new content concentrated in where the machinery must fail (Section 2.7) and in the exact boundary of the known-dynamics regime (Section 2.8).

1.2 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 [3]). Second, because our state variables are inherently multiplicative (ratios/scales), the natural geometry is non-Euclidean: Bregman divergences [2] describe projection in coordinates adapted to log/ratio structure, matching the way our axioms select log coordinates as 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 3.27 gives the per-state order-theoretic representation, and Theorem 3.33 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.

1.3 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 [7]. In this paper we take the normalized choice (3) 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.

1.4 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, namely (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: an aggregation/reconstruction step A, a projection step P, and a coercivity step B, applied to the data in that order, so that the certifier is the composition $\Phi^* = B \circ P \circ A$. 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 2.21). Our maximality statements (e.g. Theorem 3.20) are correspondingly local to that locus and do not claim global parameter recovery from window sums. Sections 2.7 and 2.8 make this scoping exact: the former proves that certification of the neutral target is impossible when the latent dynamics are jointly unknown, and the latter identifies the exact boundary of the regime where it is possible.

1.5 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 3.12), the canonical pipeline $\Phi^* = B \circ P \circ A$ is locally maximal on the identifiability locus (Theorem 3.20).

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) **Limits of certification from block sums.** Lemma 2.27 shows that in the jointly parameterized model the neutral fiber is a critical set of the window map for every $d \geq 1$, and Theorem 2.28 exhibits, for every block length $W \geq 2$, nonzero admissible signals inside the rational class whose window data coincide exactly with the neutral data. Consequently (Corollary 2.30) no sound procedure over the joint class can ever certify neutrality: the mathematically forced output on neutral data is *inconclusive*.

(ii) **An exact certifiable/uncertifiable boundary.** In the known-dynamics model (fixed denominator, unknown state), Theorem 2.31 proves that certification from $K \geq d + 1$ window sums is sound *and complete* precisely when no latent mode is a nontrivial W-th root of unity and the W-th power map is injective on the modes. On that locus the reconstruction step is globally linear, and Proposition 3.3 and Theorem 3.5 give explicit ε-stability constants in terms of the pseudoinverse of the window matrix.

(iii) **Quantitative sound certification under noise.** Theorems 3.5 and 3.38 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 $\mathrm{Mean}_{\varepsilon}$).

(iv) **Local domination among sound rules.** Theorem 3.20 proves that on the identifiability locus the canonical pipeline $\Phi^* = B \circ P \circ A$ is locally maximal among all sound certification rules (Definition 3.12), i.e. no sound rule can resolve strictly more instances in a neighborhood; Corollary 2.34 exhibits a concrete model class on which all hypotheses of the domination theorem are verified, so the statement is not vacuous.

(v) **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 3.33 records the induced factorization/uniqueness behavior for certificates under explicit rigidity hypotheses.

(vi) **Uniqueness among locally optimal rules.** Corollary 3.21 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 3.27 shows the local ordinal representation, while Theorem 3.33 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)

$$J(xy) + J(x/y) = 2J(x) + 2J(y) + 2J(x)J(y), \qquad x > 0,\ y > 0,$$

together with the curvature normalization $J''(1) = 1$. 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). We use the corresponding canonical reciprocal cost,

$$J(x) := \frac{1}{2}\Big(x + x^{-1}\Big) - 1, \qquad x > 0. \tag{3}$$

(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 recorded in Item 1 of the Appendix (equivalently $J''(1) = 1$), the cost is uniquely determined; equivalently, it is given by (3).

(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 3.27) together with the strengthened global uniqueness theorem (Theorem 3.33) 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 later control the aggregation and certification steps.

2.2.1 The Canonical Reciprocal Cost

We recall that J is fixed by (3). 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 $J(e^t) = 0 \iff t = 0 \iff e^t = 1$.

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

$$P(y) := y - \bar{y}\,\mathbf{1}. \tag{4}$$

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 $y = 0$.

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 $t = 0$.

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

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

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

Remark 2.10 (The coercivity constant is the calibration constant) The constant $\frac{1}{2}$ in Theorem 2.9 is sharp, and it is not a separate estimate: since $J(e^t) = \cosh t - 1 = \frac{1}{2}t^2 + O(t^4)$, one has $\frac{1}{2} = \frac{1}{2}J''_{\log}(0)$, exactly half the curvature imposed by the calibration axiom (Axiom 2 of Appendix Item 1). The local calibration at balance and the global coercive bound are therefore the same constant seen at two scales: the axiom that fixes the units of the cost is precisely what fixes the strength of the defect certificate. A different calibration would rescale both sides of the inequality identically.

2.3 Finite-Data Aggregation: W-Block Window Measurements

Remark 2.11 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*

$$\mathcal{W}(y)_k := \sum_{j=0}^{W-1} y_{(Wk+j)}, \qquad k = 0, 1, 2, \dots. \tag{5}$$

For brevity we write $S_k := \mathcal{W}(y)_k$ for the window sums; the letter W is reserved for the fixed window *length*. 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 2.12 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 2.13 *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*

$$y_n + q_1 y_{n-1} + q_2 y_{n-2} + \cdots + q_d y_{n-d} = 0, \qquad n \geq d + 1. \tag{6}$$

Conversely, any sequence satisfying (6) 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 (6) holds for fixed $(q_1, \ldots, q_d)$, then each y_n with $n \geq d+1$ is determined recursively from $(y_0, \ldots, y_d)$. Hence the full sequence is determined by $(q_1, \ldots, q_d, y_0, \ldots, y_d)$, i.e. by $2d+1$ real parameters.

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 2.13, once the parameter vector

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

is fixed, the recurrence (6) determines y_n for all $n \geq d+1$.

Recall the W-block window sums $\mathcal{W}(y)_k$ from (5), and define the truncated window map (using the first $2d+1$ windows)

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

where y is generated from π by (6). Here $DF_{d,W}(\pi)$ denotes the Jacobian matrix of $F_{d,W}$ at π.

Lemma 2.14 *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* (6)*, 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 (6) 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.

Theorem 2.15 *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 2.14, $\det DF_{d,W}$ 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.

Remark 2.16 Theorem 2.15 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.

Lemma 2.17 *For $d = 3$ and $W = 8$, the set $\Omega_{3,8}$ from Lemma 2.14 is nonempty. In particular, this verifies the hypothesis of Theorem 2.15 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 (6) at $\pi = \pi_0$, compute y_n up to $n = 55$ and hence the first $K = 7$ window sums

$$\begin{aligned}(S_0, \ldots, S_6) = (&-16, -5160, -975168, -169890432,\\ &-27959752704, -4399334572032, -665805326548992).\end{aligned}$$

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 $\partial_\alpha S_k = \sum_{j=0}^{7} u_{8k+j}^{(\alpha)}$. Assemble the 7×7 Jacobian $DF_{3,8}(\pi_0)$ with entries $\partial_\alpha S_k$. (The full integer Jacobian matrix is recorded in the Appendix, Item 2 below; here we retain only the determinant residue modulo p.)

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

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

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

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

(d, W)	witness π_0	certificate of full rank
$(3, 8)$	$(1, 1, 5, 1, 2, 2, -2)$	Lemma 2.17

Remark 2.18 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 2.17 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 2.19 *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 determine π locally uniquely within the degree-$\leq d$ rational class. Consequently, the degree-$\leq d$ signal θ and the full coefficient sequence $(y_n)_{n\geq 0}$ are locally uniquely determined on $\Omega_{d,W}$ from any $K \geq 2d + 1$ windows.*

Proof Fix $\pi \in \Omega_{d,W}$. By Lemma 2.14 the map

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

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

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

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 $K \geq 2d + 1$.

Remark 2.20 Theorem 2.19 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 2.21 *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.

Remark 2.22 In concrete Prony/Hankel implementations, failure of the hypotheses in Theorem 2.21 (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.

Lemma 2.23 (Confluent Vandermonde nonsingularity) *Let $\mu_1, \ldots, \mu_d$ be distinct real numbers and let $C_1, \ldots, C_d$ be nonzero real numbers. Consider the sample map*

$$(C, \mu) \longmapsto \Big(T_0, \ldots, T_{2d-1}\Big), \qquad T_k := \sum_{i=1}^{d} C_i\, \mu_i^k.$$

Its $2d \times 2d$ Jacobian, with entries $\partial T_k/\partial C_i = \mu_i^k$ and $\partial T_k/\partial \mu_i = k\, C_i\, \mu_i^{k-1}$, is nonsingular.

Proof Factoring the nonzero constants C_i out of the μ_i-columns, it suffices to show that the matrix with columns $\left(\mu_i^k\right)_k$ and $\left(k\mu_i^{k-1}\right)_k$, $k = 0, \ldots, 2d-1$, is nonsingular. A vector $(c_0, \ldots, c_{2d-1})$ in its left kernel encodes a polynomial $p(t) = \sum_{k=0}^{2d-1} c_k t^k$ of degree at most $2d-1$ with $p(\mu_i) = p'(\mu_i) = 0$ for every i. Then p has at least $2d$ roots counted with multiplicity, hence $p \equiv 0$ and $c = 0$. This is the uniqueness of Hermite interpolation at d distinct doubled nodes.

Theorem 2.24 (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 2.19 and the global Prony recovery of Theorem 2.21 apply for every (d, W) under the corresponding nondegeneracy hypotheses.*

Proof Choose distinct $\alpha_1, \dots, \alpha_d \in (0,1)$ (e.g. $\alpha_i := 1/(i+1)$) and set $A_i := 1$, $c_0 := 0$. We work in adapted coordinates and transport nonsingularity back to the parameter π.

Step 1 (adapted parameterization). On the open set U of triples $(c_0, A, \alpha) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ with the α_i distinct and nonzero, define the signal

$$y_n(c_0, A, \alpha) := c_0\, \mathbf{1}_{\{n=0\}} + \sum_{i=1}^{d} A_i\, \alpha_i^{\,n},$$

with generating function $c_0 + \sum_i A_i/(1-\alpha_i z) = p(z)/q(z)$, $q(z) = \prod_i (1-\alpha_i z)$, $\deg p \le d$, $q(0) = 1$. The induced map $\Psi : (c_0, A, \alpha) \mapsto \pi = (y_0, \dots, y_d, q_1, \dots, q_d)$ is polynomial. Its Jacobian is block triangular: the q-coordinates depend only on α, with Jacobian determinant $\pm \prod_{i<j}(\alpha_i - \alpha_j) \ne 0$ (elementary symmetric functions at distinct arguments), while $(y_0, \dots, y_d)$ is linear in (c_0, A) with matrix determinant $\det\Big(\alpha_i^{\,n}\Big)_{1 \le n \le d} = \Big(\prod_i \alpha_i\Big) \prod_{i<j}(\alpha_j - \alpha_i) \ne 0$. Hence Ψ is a local diffeomorphism at the witness; write $\pi_0 := \Psi(0, \mathbf{1}, \alpha)$.

Step 2 (window sums in adapted coordinates). For all $k \ge 0$,

$$S_k = c_0\, \mathbf{1}_{\{k=0\}} + \sum_{i=1}^{d} B_i\, \mu_i^{\,k}, \qquad \mu_i := \alpha_i^{\,W}, \quad B_i := A_i\, \frac{1-\alpha_i^{\,W}}{1-\alpha_i}.$$

The change of coordinates $(c_0, A, \alpha) \mapsto (c_0, B, \mu)$ is block triangular with $\partial \mu_i/\partial \alpha_i = W \alpha_i^{\,W-1} \ne 0$ and $\partial B_i/\partial A_i = (1-\alpha_i^{\,W})/(1-\alpha_i) \ne 0$ on $(0,1)$, hence a local diffeomorphism at the witness, where the $\mu_i \in (0,1)$ are distinct (the map $t \mapsto t^W$ is injective on $(0,1)$) and $B_i > 0$.

Step 3 (nonsingularity of the sample map). Consider $(c_0, B, \mu) \mapsto (S_0, \dots, S_{2d})$. The c_0-column of its Jacobian is the first standard basis vector, so expanding along it reduces nonsingularity to the $2d \times 2d$ Jacobian of $(B, \mu) \mapsto (S_1, \dots, S_{2d})$. Substituting $C_i := B_i \mu_i$ (an invertible triangular change of variables, since $\mu_i \ne 0$) and writing $T_{k-1} := S_k = \sum_i C_i \mu_i^{\,k-1}$, this Jacobian is, up to that invertible factor, exactly the confluent matrix of Lemma 2.23 with distinct nodes μ_i and nonzero amplitudes C_i, hence nonsingular.

Conclusion. The composition of the three maps computes the first $2d+1$ window sums of the signal, i.e. equals $F_{d,W} \circ \Psi$ on U. By the chain rule and Steps 1–3, $\det DF_{d,W}(\pi_0) \ne 0$, so $\pi_0 \in \Omega_{d,W}$ and $\Omega_{d,W} \ne \emptyset$. Finally, at the same witness the Hankel matrix $H = (S_{i+j})_{0 \le i,j \le d-1}$ (built from the windows with $k \ge 1$, i.e. from $\widetilde{S}_k := S_{k+1}$) factors as $H = V^\top \operatorname{diag}(C)\, V$ with V the Vandermonde matrix on (μ_i), so $\det H = (\det V)^2 \prod_i C_i \ne 0$ and the Prony/Hankel nondegeneracy hypotheses of Theorem 2.21 hold there as well.

Remark 2.25 Step 3 is where the naive argument fails: invertibility of the Hankel matrix of the induced window-sum process concerns the $2d$ parameters (B, μ) of that process, while $\Omega_{d,W}$ concerns the full $(2d+1)$-dimensional Jacobian in the π-coordinates. The extra dimension is carried by the δ-component c_0 (the $\deg p = \deg q$ part of the transfer function), which is visible only in S_0; the block-triangular factorization above accounts for it explicitly.

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: $K \ge 2d+1$ 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 2.21. For this choice, Theorem 2.24 yields $\Omega_{3,8} \neq \emptyset$, so the generic local uniqueness conclusion of Theorem 2.19 holds. Moreover, under the nondegeneracy hypotheses of Theorem 2.21, 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 2.19 on the nonempty full-rank locus $\Omega_{3,8}$ certified in Lemma 2.17. The global $K \geq 2d$ threshold for recovery of the induced window-sum process (equivalently, the exponential parameters) follows from Theorem 2.21 under its stated Prony/Hankel nondegeneracy hypotheses.

2.7 The Neutral Target: Limits of Joint Identifiability

The identifiability results above are generic: they hold on the open dense locus $\Omega_{d,W}$. The certification problem, however, targets a single exceptional point, the neutral signal, and the two results of this subsection show that the neutral target sits exactly where generic identifiability fails. This is why all certification statements in Section 3 are scoped to explicit loci, and why the known-dynamics regime of Section 2.8 is the natural home of the certification pipeline.

Lemma 2.27 (The neutral fiber is critical) *Fix $d \geq 1$ and $W \geq 1$, and write $\pi = (a, q)$ with $a := (y_0, \ldots, y_d) \in \mathbb{R}^{d+1}$ and $q = (q_1, \ldots, q_d) \in \mathbb{R}^d$. Then:*

(a) *for fixed q, the truncated window map is linear in the initial data: $F_{d,W}(a, q) = L(q)\,a$ for a matrix $L(q) \in \mathbb{R}^{(2d+1)\times(d+1)}$ depending polynomially on q;*
(b) *the entire neutral fiber $\{0\} \times \mathbb{R}^d$ maps to the single datum $0 \in \mathbb{R}^{2d+1}$;*
(c) *at every neutral parameter point $(0, q)$ the Jacobian satisfies* $\operatorname{rank} DF_{d,W}(0, q) \leq d + 1 < 2d + 1$.

In particular, every neutral parameter point is a critical point of $F_{d,W}$, and the identifiability locus $\Omega_{d,W}$ is disjoint from the neutral fiber.

Proof (a) Each y_n is produced from a by iterating the linear recurrence (6), whose coefficients depend on q only; hence y_n, and therefore each window sum, is linear in a with polynomial coefficients in q. (b) is immediate from (a). (c) At $a = 0$ the signal vanishes identically. The α-sensitivity recurrences with respect to the q-parameters (as in the proof of Lemma 2.17) are driven by the values $y_{n-1}, \ldots, y_{n-d}$, all zero, with zero initial conditions; hence $\partial F_{d,W}/\partial q = 0$ at $(0, q)$ and $DF_{d,W}(0, q) = \left[\, L(q) \;\middle|\; 0 \,\right]$, of rank at most $d + 1$.

Theorem 2.28 (Window annihilation inside the rational class) *Fix a block length $W \geq 2$.*

(a) *If W is even, the signal $y_n := (-1)^n$ is rational of degree* 1 *(generating function* $1/(1 + z)$*), is not identically zero, and satisfies $\mathcal{W}(y)_k = 0$ for every $k \geq 0$.*
(b) *For every $W \geq 2$, the signal $y_n := 2\cos(2\pi n/W)$ is real, rational of degree* 2 *(generating function $(2 - cz)/(1 - cz + z^2)$ with $c := 2\cos(2\pi/W)$), is not identically zero, and satisfies $\mathcal{W}(y)_k = 0$ for every $k \geq 0$.*

Consequently, for every $d \geq 2$ (and already for $d \geq 1$ when W is even), the window data of these signals coincide, for every window count K, with the window data of the neutral signal $y \equiv 0$, although their defect is strictly positive.

Proof Let ζ be a primitive W-th root of unity. For every $k \geq 0$,

$$\sum_{j=0}^{W-1} \zeta^{Wk+j} = \zeta^{Wk} \frac{\zeta^W - 1}{\zeta - 1} = 0,$$

since $\zeta^W = 1$ and $\zeta \neq 1$. For (a) take $\zeta = -1$ (primitive when W is even); for (b) take the real part of twice the geometric signal, $y_n = \zeta^n + \bar{\zeta}^n$, whose block sums vanish termwise by the same identity. The generating functions are $1/(1+z)$ and $1/(1-\zeta z) + 1/(1-\bar{\zeta} z) = (2-cz)/(1-cz+z^2)$, rational of the stated degrees with $q(0) = 1$.

Remark 2.29 (Admissibility, and collisions arbitrarily close to neutrality) The witnesses are admissible at the configuration level: for any $\varepsilon > 0$ the configuration $x := e^{\varepsilon y} \in (\mathbb{R}_{>0})^n$ is positive, and the conservation constraint $\sigma(x) = \varepsilon \sum_i y_i = 0$ holds whenever the horizon n is a multiple of W (each full block of the witness sums to zero). Since ε is arbitrary, the scaled witnesses εy produce window data identical to the neutral data while lying arbitrarily close to the neutral signal; in parameter space they converge to a neutral point $(0, q^\dagger)$ whose denominator $q^\dagger$ carries the annihilated mode. The phenomenon is a multirate (decimation) aliasing effect [16]: block summation is a filter-and-downsample operation whose frequency response vanishes at the nonzero W-th roots of unity.

Corollary 2.30 (No sound certification of neutrality over the joint class) *Fix $W \geq 2$, $d \geq 2$ (or $d \geq 1$ with W even), and any $K \geq 1$. Then every procedure $\Psi : \mathbb{R}^K \to \{\text{zero}, \text{nonzero}, \text{inconclusive}\}$ that is sound on the full class $\mathcal{M}_{d,W}$ (Definition 3.13) must return* inconclusive *on the all-zero datum $w = 0$. In particular no sound procedure over the joint class ever certifies the neutral configuration, and no analogue of Theorem 3.20 can hold with $\mathcal{U}_{d,W} = \mathcal{M}_{d,W}$: the localization to an identifiability neighborhood avoiding the annihilated modes is necessary, not an artifact of the proof. Equivalently, on the joint class soundness and totality are incompatible: every sound procedure must abstain on the neutral datum, so any procedure that decides every realizable datum is unsound somewhere.*

Proof The neutral signal and the witness of Theorem 2.28 both realize $w = 0$. If $\Psi(0) = \text{zero}$, soundness fails on the witness (which is non-neutral); if $\Psi(0) = \text{nonzero}$, soundness fails on the neutral signal. Hence $\Psi(0) = \text{inconclusive}$.

2.8 A Certifiable Regime: Known Dynamics

The obstructions of Section 2.7 concern the *jointly* unknown pair (initial data, recurrence). In many aggregation scenarios the dynamics are known in advance, from physical calibration or from the design of the reporting system, and only the state is unknown: the decay constants of a gated fluorescence channel, or the fixed retention kernel of an aggregated reporting funnel (Section 4). Fixing the denominator turns the window map into a globally linear map of the state and makes the certification question exactly solvable, including its boundary.

For a fixed monic-normalized denominator q (with $q(0) = 1$ and $\deg q = d$), let

$$\mathcal{M}_{q,W} := \{y : \ \theta(z) = p(z)/q(z), \ \deg p \le d\}$$

be the known-dynamics class, parameterized linearly by $a = (y_0, \dots, y_d) \in \mathbb{R}^{d+1}$, and let $L_q \in \mathbb{R}^{K\times(d+1)}$ denote the matrix of the linear map $a \mapsto \mathcal{W}_{W,K}(y)$ (Lemma 2.27(a), truncated to K windows).

Theorem 2.31 (Exact certification boundary for known dynamics) *Let $q(z) = \prod_{i=1}^{d}(1 - \lambda_i z)$ with distinct nonzero $\lambda_i \in \mathbb{C}$ (closed under conjugation, so that q is real), and let $K \ge d + 1$. The following are equivalent.*

(i) *L_q is injective, i.e. the first K window sums identify the state on $\mathcal{M}_{q,W}$.*

(ii) *No λ_i is a nontrivial W-th root of unity, and $\lambda_i^W \ne \lambda_j^W$ for all $i \ne j$.*

(iii) *Zero-certification is sound and complete on $\mathcal{M}_{q,W}$: $\mathcal{W}_{W,K}(y) = 0$ if and only if $y \equiv 0$.*

Moreover, when these hold, they hold already for $K = d + 1$ windows.

Proof Decompose signals in $\mathcal{M}_{q,W}$ by partial fractions: $y_n = c_0\mathbf{1}_{\{n=0\}}+\sum_i A_i\lambda_i^n$, where $(c_0, A) \leftrightarrow a$ is a linear isomorphism (Step 1 of the proof of Theorem 2.24). As in Step 2 there, the window sums are

$$S_k = c_0\mathbf{1}_{\{k=0\}}+\sum_{i=1}^{d} B_i\,\mu_i^k, \qquad \mu_i := \lambda_i^W, \qquad B_i := A_i\,g_i, \qquad g_i := \begin{cases} \dfrac{1-\lambda_i^W}{1-\lambda_i}, & \lambda_i \ne 1, \\ W, & \lambda_i = 1. \end{cases}$$

(ii)⇒(i),(iii). Under (ii) every visibility factor is nonzero: $g_i = 0$ would force $\lambda_i^W = 1$ with $\lambda_i \ne 1$, excluded. Suppose $S_0 = \cdots = S_d = 0$. The equations for $k = 1, \dots, d$ read $\sum_i C_i\,\mu_i^{k-1} = 0$ with $C_i := B_i\mu_i$; since the μ_i are distinct (by (ii)) and nonzero (the λ_i are nonzero), the $d\times d$ Vandermonde system forces $C = 0$, hence $B = 0$, hence $A = 0$ ($g_i \ne 0$), and then $S_0 = c_0 = 0$. So $a = 0$, giving (i), and (iii) is the same statement combined with the trivial direction ($y \equiv 0 \Rightarrow w = 0$).

¬(ii)⇒ ¬(i). If some $\lambda_i^W = 1$ with $\lambda_i \ne 1$, then $g_i = 0$ and the mode $y_n = \lambda_i^n$ (or, if $\lambda_i \notin \mathbb{R}$, the real and imaginary parts of that mode, which lie in the class since it is conjugation-stable) is a nonzero element of $\ker L_q$: all of its window sums vanish. If instead $\mu_i = \mu_j$ for some $i \ne j$ with $g_i, g_j \ne 0$, the combination $A_i := g_j$, $A_j := -g_i$ (other coordinates zero) has $B_i + B_j = 0$, so again every window sum vanishes while $y \not\equiv 0$; a real kernel element is obtained by adding the conjugate combination. Hence L_q is not injective.

Finally, the argument for (ii)⇒(i) used only the windows $k = 0, \dots, d$, so $K = d+1$ suffices, and injectivity for K windows implies injectivity for any $K' \ge K$.

Remark 2.32 (i) The boundary in Theorem 2.31(ii) is exactly the aliasing set of Remark 2.29: a mode is invisible precisely when block summation annihilates it ($\lambda^W = 1$, $\lambda \ne 1$), and two modes are confounded precisely when they alias to the same decimated node ($\lambda_i^W = \lambda_j^W$). For real decaying dynamics ($0 < \lambda_i < 1$, the case of Section 4) the condition holds automatically, since $t \mapsto t^W$ is injective on $(0, 1)$ and no such λ is a root of unity. (ii) Only $d + 1$ windows are needed, versus $2d + 1$ in the joint model: knowing the dynamics halves the data requirement. (iii) For denominators with multiple roots the same argument runs with polynomial-in-n mode amplitudes and confluent Vandermonde systems (Lemma 2.23); we do not pursue the general statement here. (iv) Theorem 2.31 is exactly the observability criterion for the decimated dynamics. Write $M := \operatorname{diag}(\mu_1, \ldots, \mu_d)$ and let $\mathbf{1}$ be the all-ones output functional; then the window sums for $k \geq 1$ are the samples $\mathbf{1}^\top M^k B$ of the linear system with state matrix M and initial state $B = (B_1, \ldots, B_d)^\top$, so the map $B \mapsto (S_1, \ldots, S_d)$ is the observability matrix of the pair $(M, \mathbf{1}^\top)$, up to right-multiplication by the invertible matrix M. The two clauses of condition (ii) are the two ways observability can fail after aggregation: the Popov–Belevitch–Hautus condition for the diagonal pair $(M, \mathbf{1}^\top)$ requires the μ_i distinct (no aliasing under decimation), and the change of variables $A \mapsto B$ is invertible exactly when every visibility factor g_i is nonzero (no mode annihilated by the boxcar filter's zeros at the W-th roots of unity). The certifiable regime of this paper is thus precisely the observable regime of the aggregated system, in the classical sense of [19].

Proposition 2.33 (Linear reconstruction and its conditioning) *Under the equivalent conditions of Theorem 2.31, the Moore–Penrose pseudoinverse* $\mathsf{Rec}_q := L_q^+ : \mathbb{R}^K \to \mathbb{R}^{d+1}$ *is a globally defined linear reconstruction with* $\mathsf{Rec}_q(\mathcal{W}_{W,K}(y)) = a$ *for every* $y \in \mathcal{M}_{q,W}$*, it maps the zero datum to the neutral state, and it is globally Lipschitz with the explicit constant*

$$L = \left\| L_q^+ \right\|_{2\to 2} = \sigma_{\min}(L_q)^{-1},$$

where $\sigma_{\min}(L_q)$ *is the smallest singular value of* L_q*.*

Proof Injectivity of L_q makes $L_q^+ L_q = I$ on $\mathbb{R}^{d+1}$, which gives exact reconstruction on realizable data and $\mathsf{Rec}_q(0) = 0$. The operator norm of the pseudoinverse of an injective matrix is the reciprocal of its smallest singular value, and a linear map is globally Lipschitz with its operator norm.

Statistical shadow of the stability constant. The constant $L = \sigma_{\min}(L_q)^{-1}$ has a statistical reading as well. If the window sums are corrupted by independent noise of variance ς^2 per coordinate, the Gauss–Markov estimate of the state has covariance $\varsigma^2 (L_q^\top L_q)^{-1}$, whose worst direction has standard deviation $\varsigma/\sigma_{\min}(L_q) = \varsigma L$. The deterministic amplification factor of Proposition 2.33 is therefore exactly the worst-case standard-deviation amplification of the corresponding statistical estimation problem: one and the same conditioning number governs our worst-case tolerance bounds and the variance of best unbiased estimation. We do not develop the statistical theory here; it is the content of future direction (i) in the Conclusions.

Corollary 2.34 (Global domination on a known-dynamics class) *Let q satisfy the equivalent conditions of Theorem 2.31, so that the window map is injective on the linear class $\mathcal{M}_{q,W}$ and neutrality is certifiable. Then the canonical procedure Φ^* of Section 3, applied on $\mathcal{M}_{q,W}$, is total (never returns* inconclusive *on realizable data) and dominates every sound procedure on all of $\mathcal{M}_{q,W}$:* $\mathrm{Res}(\Psi) \subseteq \mathrm{Res}(\Phi^*)$ *and the two agree on* $\mathrm{Res}(\Psi)$*, for every sound Ψ.*

Proof By Theorem 2.31, L_q is injective, so the reconstruction $\mathsf{Rec}_q = L_q^+$ of Proposition 2.33 recovers the state exactly from any realizable datum and, in particular, decides neutrality ($y \equiv 0$) without ambiguity; hence Φ^* resolves every realizable datum. Global injectivity is hypothesis (a) of Corollary 3.22, and the exact-reconstruction property supplies (b)–(c) with $\mathcal{M}_{d,W} := \mathcal{M}_{q,W}$, so global domination follows.

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*

$$\begin{aligned} \Phi^* = B \circ P \circ A : \quad w &\mapsto A(w) = x \\ &\mapsto u = P(\log x) \\ &\mapsto B(u) \in \{\text{zero}, \text{nonzero}, \text{inconclusive}\}, \end{aligned}$$

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 3.12.*
(ii) ***ε-robust certification.*** *Under perturbations of size ε, the certified conclusion degrades in the sharp two-sided way, yielding membership in $\mathrm{Mean}_{2\varepsilon}$ as in Theorems 3.5 and 3.38.*
(iii) ***Local domination (optimality).*** *On the identifiability neighborhood $\mathcal{U}_{d,W}$, Φ^* dominates every sound procedure (Theorem 3.20).*
(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 3.21).*
(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 3.15.*
(vi) ***Non-vacuity via generic full rank.*** *For any (d, W) the full-rank locus $\Omega_{d,W}$ is a nonempty Zariski-open (hence dense) subset (Theorem 2.24); for $(d, W) = (3, 8)$ an explicit witness is certified in Lemma 2.17.*
(vii) ***Exact scope.*** *Over the jointly unknown class, no sound procedure can certify neutrality at all (Corollary 2.30), so the localization in (iii) is necessary; over a known-dynamics class satisfying Theorem 2.31(ii), the window map is globally injective, Φ^* is total, and domination holds globally (Corollary 2.34).*

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 3.22).

Proof Items (i)–(ii) follow from Definition 3.12 together with Theorems 3.5 and 3.38. Item (iii) is Theorem 3.20. Item (iv) is Corollary 3.21. Item (v) is the factorization/forced-dependence principle recorded in Theorem 3.15. Item (vi) is Theorem 2.24, with the explicit $(3, 8)$ witness in Lemma 2.17. Item (vii) collects Corollary 2.30 and Corollary 2.34. The final global upgrade is Corollary 3.22.

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

Constant / threshold	Forced by
Factor 2ε in $\mathrm{Mean}_{2\varepsilon}$	Perturbation triangle inequality in Theorem 3.38
Normalization $J(1) = 0$	Proposition A.1 (Appendix, RCL)
Calibration $J''_{\log}(0) = 1$	Appendix, Item 1, Axiom 2 fixes quadratic scale
Coercivity constant $\frac{1}{2}$	The calibration curvature (Remark 2.10)
Window count $K \geq 2d + 1$	Identifiability from $2d + 1$ windows (Theorems 2.19, 2.21)
Window count $K \geq d + 1$, known dynamics	Exact boundary, Theorem 2.31

3.1 Status of the Window Length W

Throughout, $W \in \mathbb{N}$ denotes the fixed block length used to form window sums (5). 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 $W \geq 1$.

Two remarks sharpen the status of the example. First, the cycle displayed above is not an arbitrary illustration: a Gray cycle on the 3-cube is precisely a Hamiltonian cycle on its graph, and all such cycles are equivalent under the symmetries of the cube. In the motivating recognition framework the eight-state period is derived from this cycle structure rather than chosen [20]; the present paper's results, being parametric in W, apply to that derived value as a special case. Second, keeping W parametric is what lets the framework evaluate a window length rather than assume it: Theorem 2.31(ii) and Remark 2.32(i) quantify exactly what a proposed W costs for given latent dynamics, so the framework can certify a reporting period, such as the $W = 8$ of the running example, instead of merely presupposing it.

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

$$\left|W_k\right| \leq \varepsilon, \qquad k = 0, 1, \ldots, K - 1.$$

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 $(2m)! \geq 2\,(2m-2)!$,

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

Proposition 3.3 *Fix $W \in \mathbb{N}$, $d \geq 0$, and $n \in \mathbb{N}$, and set $K := \lceil n/W \rceil$. Let q be a fixed denominator satisfying the equivalent conditions of Theorem 2.31, and work in the known-dynamics class $\mathcal{M}_{q,W}$ of Section 2.8 (so $K \geq d+1$ suffices for identifiability). Then the reconstruction map*

$$\mathsf{Rec}_{d,W} : \mathbb{R}^K \to \mathbb{R}^n, \qquad w \mapsto \textit{the first } n \textit{ coefficients of the unique } y \in \mathcal{M}_{q,W} \textit{ with } \mathcal{W}_{W,K}(y) = w$$

is globally defined and linear, maps $w = 0$ to the neutral signal $y \equiv 0$, and is globally Lipschitz with constant equal to its operator norm:

$$\|\mathsf{Rec}_{d,W}(w) - \mathsf{Rec}_{d,W}(0)\|_2 \leq L\,\|w\|_2 \qquad \textit{for all } w \in \mathbb{R}^K, \quad L := \|\mathsf{Rec}_{d,W}\|_{2\to 2}.$$

Writing $\mathsf{Rec}_{d,W} = E_n \circ L_q^+$, where L_q^+ is the pseudoinverse of Proposition 2.33 and $E_n : \mathbb{R}^{d+1} \to \mathbb{R}^n$ is the linear map extending an initial state to the first n coefficients through the recurrence, one has $\sigma_{\min}(L_q)^{-1} \leq L \leq \|E_n\|_{2\to 2}\,\sigma_{\min}(L_q)^{-1}$, with $L = \sigma_{\min}(L_q)^{-1}$ exactly when $n = d+1$.

Proof By Theorem 2.31 the window matrix L_q is injective, so each w in its range has a unique preimage, and Proposition 2.33 gives the global linear reconstruction $a = L_q^+ w$ of the state; on data outside the range, L_q^+ is still defined and returns the least-squares state, so $\mathsf{Rec}_{d,W}$ is globally defined. The recurrence (6) is linear in the initial state, so the extension E_n and hence $\mathsf{Rec}_{d,W} = E_n \circ L_q^+$ are linear. Linearity gives $\mathsf{Rec}_{d,W}(0) = 0$, the neutral signal, and a linear map is globally Lipschitz with constant its operator norm. For the bounds on L: $\|E_n v\|_2 \geq \|v\|_2$ because the first $d+1$ coordinates of $E_n v$ are v itself (when $n \geq d+1$), which gives the lower bound, and the upper bound is the submultiplicativity of the operator norm; when $n = d+1$, E_n is the identity and both bounds coincide with $\sigma_{\min}(L_q)^{-1}$ by Proposition 2.33.

Remark 3.4 (Why the known-dynamics hypothesis here) An earlier formulation of this reconstruction step assumed local invertibility of the *joint* window map at the neutral realization. That hypothesis is empty for every $d \geq 1$: the neutral fiber is a critical set of the joint window map (Lemma 2.27), so no inverse exists on any neighborhood of the neutral datum. The neutral-datum reconstruction used for certification therefore lives exactly in the known-dynamics model, where the window map is linear and invertibility at $w = 0$ is the spectral condition of Theorem 2.31(ii). Away from the neutral datum, local reconstruction in the joint model remains valid on the nondegenerate locus $\Omega_{d,W}$ (Theorem 2.24).

Theorem 3.5 *Fix $W \in \mathbb{N}$, $d \geq 0$, and $n \in \mathbb{N}$, and set $K := \lceil n/W \rceil$. Assume the hypotheses of Proposition 3.3 (the known-dynamics model with the spectral condition of Theorem 2.31), and let L be the Lipschitz constant given there. Given window data $w \in \mathbb{R}^K$ with $|w_k| \leq \varepsilon$, let $\hat{y} := \mathsf{Rec}_{d,W}(w)$; the noiseless reconstruction $\hat{y}^{(0)} := \mathsf{Rec}_{d,W}(0) = 0$ is automatically neutral. Then the certification objective*

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

satisfies the explicit quadratic bound

$$\Phi(\hat{y}) \;\leq\; \frac{1}{2}\, e^{L\sqrt{K}\,\varepsilon}\, L^2 K\, \varepsilon^2. \tag{7}$$

Proof By Proposition 3.3, $\hat{y}^{(0)} = 0$ and

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

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

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

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

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

which is (7).

Remark 3.6 In the special case $(d, W) = (3, 8)$ one may take $K = 7$ windows, and Theorem 3.5 yields the explicit estimate

$$\Phi(\hat{y}) \;\le\; \tfrac{7}{2}\,L^2 e^{L\sqrt{7}\,\varepsilon}\,\varepsilon^2.$$

For example, on a known-dynamics class whose window matrix satisfies $L = \sigma_{\min}(L_q)^{-1} \le 10$ (with $n = d + 1$, so L is exactly the pseudoinverse norm), data with $|w_k| \le \varepsilon \le 10^{-2}$ give $L\sqrt{7}\,\varepsilon \le 0.265$, hence $e^{L\sqrt{7}\,\varepsilon} \le 1.31$ and $\Phi(\hat{y}) \le 460\,\varepsilon^2$.

Remark 3.7 The Lipschitz constant L in Proposition 3.3 is controlled by the smallest singular value of the window matrix L_q (equivalently, by a condition number κ for the Vandermonde/Hankel system used by the reconstruction), with $L = \sigma_{\min}(L_q)^{-1}$ exactly when $n = d + 1$. Poor conditioning increases L, but the dependence on ε remains quadratic.

3.3 The Certification Objective

Definition 3.8 Given a cost function $c : S \times O \to \mathbb{R}_{\ge 0}$, 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 3.9 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 $\mathrm{Mean}(s) \ne \emptyset$ 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 3.10 (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 3.8.

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 3.11 (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*

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

Then 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 $o_* \in \mathrm{Mean}(s)$.

Definition 3.12 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 3.10 or Proposition 3.11). 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

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

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

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

Definition 3.13 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 2.21. Let $y_{\mathrm{neu}} \in \mathcal{M}_{d,W}$ denote the distinguished neutral signal (one may take $y_{\mathrm{neu}} \equiv 0$). 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 $y \in \mathcal{M}_{d,W}$,

$$\begin{aligned} \Psi(\mathcal{W}_{W,K}(y)) = \mathrm{zero} &\implies y = y_{\mathrm{neu}}, \\ \Psi(\mathcal{W}_{W,K}(y)) = \mathrm{nonzero} &\implies y \neq y_{\mathrm{neu}}. \end{aligned}$$

Its *resolves set* is

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

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

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

Definition 3.14 (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 2.21, 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 (4). 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 3.15 (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.*

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

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 (4)*. 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).

Theorem 3.16 (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 3.20.*

(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 3.15.

(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 3.5–3.38).

(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 3.15.

(b) The tolerance conclusions in Theorems 3.5–3.38 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.

For brevity we continue to denote this overall procedure by $\Phi^* = B \circ P \circ A$, applied to a window datum w as $B(P(A(w)))$: the reconstruction stage A acts first (returning inconclusive when reconstruction is not defined), the projection P removes scale in log-coordinates, and the coercive decision B is applied last.

Proposition 3.17 *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 2.21 (in particular, where the reconstruction step A is well-defined by local invertibility / Hankel nondegeneracy), the canonical procedure*

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

is well-defined from $w = \mathcal{W}_{W,K}(y)$ and is sound in the sense of Definition 3.13.

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 3.18 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 3.19 *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 3.20 *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 2.21, so the data $(S_0, \dots, S_{2d-1})$ 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 $B \circ P \circ A$ defined in this section.

Then for every sound procedure $\Psi : \mathbb{R}^K \to \{\mathrm{zero}, \mathrm{nonzero}, \mathrm{inconclusive}\}$ 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 2.21, 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 $y \in \operatorname{Res}(\Psi) \cap \mathcal{U}_{d,W}$. If $\Psi(w) = \text{zero}$, then soundness forces $y = y_{\text{neu}}$, so the unique reconstruction from w is neutral and therefore $\Phi^*(w) = \text{zero}$. If $\Psi(w) = \text{nonzero}$, then $y \neq y_{\text{neu}}$, so the unique reconstruction from w is non-neutral and $\Phi^*(w) = \text{nonzero}$. Thus whenever Ψ resolves (in the sense of Definition 3.13), Φ^* resolves with the same answer, i.e. Φ^* dominates Ψ.

Corollary 3.21 (Uniqueness among locally optimal rules) *Under the hypotheses of Theorem 3.20, call a sound procedure Ψ* locally optimal *on $\mathcal{U}_{d,W}$ if no sound procedure strictly dominates it there, i.e. if $\operatorname{Res}(\Psi) \cap \mathcal{U}_{d,W}$ is maximal among sound rules. Then every locally optimal Ψ agrees with Φ^* on all data it resolves in $\mathcal{U}_{d,W}$:*

$$\Psi(\mathcal{W}_{W,K}(y)) = \Phi^*(\mathcal{W}_{W,K}(y)) \qquad \textit{for all } y \in \operatorname{Res}(\Psi) \cap \mathcal{U}_{d,W},$$

and $\operatorname{Res}(\Psi) \cap \mathcal{U}_{d,W} = \operatorname{Res}(\Phi^*) \cap \mathcal{U}_{d,W}$. In this sense Φ^* is the unique locally maximal sound rule on $\mathcal{U}_{d,W}$, up to its behavior on unresolved data. Since Theorem 3.20 dominates every *sound rule, not merely the maximal ones, the sharper statement is that Φ^* is the greatest element of the poset of sound procedures on $\mathcal{U}_{d,W}$ ordered by domination; the present corollary is the uniqueness of that greatest element.*

Proof By Theorem 3.20, Φ^* is sound and dominates every sound rule on $\mathcal{U}_{d,W}$; in particular it dominates Ψ, so $\operatorname{Res}(\Psi) \cap \mathcal{U}_{d,W} \subseteq \operatorname{Res}(\Phi^*) \cap \mathcal{U}_{d,W}$ and the outputs agree wherever Ψ resolves. If the inclusion were strict, Φ^* would strictly dominate Ψ, contradicting local optimality of Ψ; hence the resolved sets coincide.

Corollary 3.22 (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 2.21;*
(c) *the realizable datum determines global neutrality on $\mathcal{M}_{d,W}$, i.e.*

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

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

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

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

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

Proposition 3.23 (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 $\Psi : \mathbb{R}^K \to \{\text{zero}, \text{nonzero}, \text{inconclusive}\}$ 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.

Remark 3.24 (When the whole-class upgrade is available, and when it is not) The global upgrade of Corollary 3.22 requires the window map to be injective on the entire parameter locus. For the *jointly* unknown rational class this hypothesis fails identically: the neutral fiber is a critical set (Lemma 2.27) and block aggregation admits nonzero-signal collisions with the neutral datum (Theorem 2.28), so by Proposition 3.23 no sound rule can even resolve the neutral datum, let alone dominate on all of $\mathcal{M}_{d,W}$. The upgrade is therefore available only on a class where global injectivity is proved; the known-dynamics class $\mathcal{M}_{q,W}$ of Section 2.8, under the spectral condition of Theorem 2.31(ii), is exactly such a class, and Corollary 2.34 records the resulting whole-class domination there.

3.5 Sharpness Boundary Outside the Rational Class

The collision principle of Proposition 3.23 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 (5) 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

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

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 3.25 (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 $b_1, \dots, b_D \in \mathbb{R}$ and n_0 such that

$$r_n = b_1 r_{n-1} + \cdots + b_D r_{n-D} \qquad \text{for all } n \geq n_0. \tag{8}$$

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 (8) forces $1 = 0$, a contradiction. Therefore r is not rational, hence y_{out} cannot be rational of degree $\leq d$.

Corollary 3.26 (Unconditional sharpness boundary outside the rational class) *Fix $d \geq 1$, $W \geq 1$, and $K \geq 0$. Let $T(y) := \mathbf{1}_{\{y \in \mathcal{C}_{\text{rat}}(d)\}}$ 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 $y_{\text{in}} \in \mathcal{C}_{\text{rat}}(d)$ and set $w := \mathcal{W}_{W,K}(y_{\text{in}})$. By Theorem 3.25 there exists $y_{\text{out}} \notin \mathcal{C}_{\text{rat}}(d)$ with $\mathcal{W}_{W,K}(y_{\text{out}}) = w$. Thus $T(y_{\text{in}}) \neq T(y_{\text{out}})$ while the finite-data coincide, so Proposition 3.23 forces any sound procedure to be inconclusive on w.

Fix a costed space (S, O, c) with meaning map $\text{Mean}(s) = \arg\min_{o \in O} c(s, o)$. In applications one typically does not compute $\text{Mean}(s)$ 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 $o \in \text{Mean}(s)$.

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 3.27 *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 3.12(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*

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

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 $\mathcal{C}_s = \phi_s \circ c_s$.

Proposition 3.28 *Assume the hypotheses of Lemma 3.27. 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*

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

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 3.27 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

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

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

$$\begin{aligned}
\mathcal{C}(s', o) &= \phi_{s'}\Big(J(r'(o))\Big) &(9)\\
&= \phi_{s'}\Big(F_\lambda\Big(J(r(o))\Big)\Big) &(10)\\
&= \underbrace{\Big(\phi_{s'} \circ F_\lambda \circ \phi_s^{-1}\Big)}_{=:\ \psi_{s,s'}}\Big(\phi_s\Big(J(r(o))\Big)\Big) &(11)\\
&= \psi_{s,s'}\Big(\mathcal{C}(s, o)\Big). &(12)
\end{aligned}$$

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

Proposition 3.29 (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 3.33 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 $\mathcal{C}(s_1, o) = \mathcal{C}(s_2, o)$.

Remark 3.30 (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 3.33 becomes immediate. If a nontrivial bridge from primitives to (H1)–(H2) is intended, (P3) should be weakened.

Proposition 3.31 (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 3.12), then for every* $s \in S$ *there exists a strictly increasing map* $f_s : c_s(O) \to \mathbb{R}$ *such that*

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

where $c_s(O) := \{c(s,o) : o \in O\}$. *Moreover, hypothesis* (H1) *of Theorem 3.33 holds if and only if the family* $\{f_s\}_{s\in S}$ *is* state-independent on overlaps *in the sense that for all* $s_1, s_2 \in S$,

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

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.

Remark 3.32 (On (H2) *and rescaling)* Hypothesis (H2) of Theorem 3.33 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 3.33 (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 3.31.) 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 $\widetilde{\psi} = \psi$.

Corollary 3.34 (Forced-factorization uniqueness from primitive hypotheses) *Under the assumptions of Proposition 3.29, the uniqueness conclusions (a) and (b) of Theorem 3.33 hold.*

Proof By Proposition 3.29, assumptions (H1)–(H2) in Theorem 3.33 are satisfied, so the result follows directly.

Remark 3.35 (Standalone status of the uniqueness theorem) Theorem 3.33 and Corollary 3.34 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 3.36 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 3.37 *Let $0 < a \le b < \infty$ and let $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ on $(0, \infty)$. Then J is Lipschitz on $[a, b]$ with constant*

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

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

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

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

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

By the mean value theorem, $|J(x) - J(y)| \le \sup_{[a,b]} |J'|\,|x - y| \le L(a)|x - y|$. If $|\widehat{r}/r - 1| \le \delta$ and $r, \widehat{r} \in [a, b]$, then the segment between r and $\widehat{r}$ lies in $[a, b]$ and $|\widehat{r} - r| \le \delta r \le \delta b$, giving the final bound. (If only $r \in [a, b]$ is assumed, the same argument with the enlarged interval $[a(1-\delta), b(1+\delta)]$ yields the constant $L(a(1-\delta))$ in place of $L(a)$.)

Theorem 3.38 *Assume the hypotheses of Lemma 3.27. Fix $s \in S$ and suppose that the relevant ratios*

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

and their estimates $\widehat{r}(o)$ below 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. $|\widehat{r}(o)/r(o) - 1| \le \delta$. 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 $|\widehat{r}(o)/r(o) - 1| \le \delta$ where $r(o) = \iota_S(s)/\iota_O(o)$. Since $r(o), \widehat{r}(o) \in [a, b]$ by hypothesis, Lemma 3.37 gives the uniform bound

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

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 $o \in O$,

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

Hence any sound selection based on these measurements can only certify candidates within $\mathrm{Mean}_{2\varepsilon}(s)$ (Definition 3.36).

Remark 3.39 (Standalone status of the ε-noise bound) Theorem 3.38 follows directly from Lemma 3.37 and the perturbation estimate proved above.

4 Applications: Synthetic Windowed Examples

The theoretical results above apply most naturally in settings where only short aggregated windows are stored or transmitted. We record two synthetic examples that illustrate the intended workflow, and then discuss the practical behavior and limitations that the impossibility and boundary results of Section 2.3 impose on any such workflow. We make no claim of empirical validation; the examples are chosen to show both where the method succeeds and where it must, correctly, refuse to decide.

4.1 Multi-Exponential Decay Under Gated Integration

A standard model for gated spectroscopy or fluorescence decay is

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

which has a rational generating function and therefore fits the finite-state framework used here. Take $d = 3$, $W = 8$, decay constants $(a_1, a_2, a_3) = (0.9, 0.6, 0.3)$ and weights $(w_1, w_2, w_3) = (1, 1, 1)$. Because each $a_j \in (0, 1)$, the spectral condition of Theorem 2.31(ii) holds automatically: no a_j is a root of unity and $t \mapsto t^8$ separates the three decays ($0.9^8 \approx 0.430$, $0.6^8 \approx 0.0168$, $0.3^8 \approx 6.6 \times 10^{-5}$). If the dynamics (a_1, a_2, a_3) are treated as known (fixed by calibration of the gate), then by Corollary 2.34 the linear reconstruction $\mathsf{Rec}_q = L_q^+$ recovers the state from $K = d + 1 = 4$ window sums and certifies neutrality without ambiguity. This example also shows why the framework reports the conditioning constant rather than hiding it. In the initial-state parameterization one computes $\sigma_{\min}(L_q) \approx 1.2 \times 10^{-6}$, so Proposition 2.33 gives Lipschitz constant $L = \sigma_{\min}^{-1} \approx 8.6 \times 10^5$: window noise of size ε can move the reconstructed state by of order $10^6\varepsilon$, and Theorem 3.5 converts that into the correspondingly weak (but honest) defect tolerance. The mechanism is visible in the separation numbers above: the fast modes 0.6 and 0.3 contribute essentially nothing beyond the first window, so the four windows see an effectively lower-rank signal. The certificate remains sound; what degrades is the tolerance it can advertise. The same computation quantifies the remedies: shortening the window to $W = 2$ for the same dynamics gives $\sigma_{\min}(L_q) \approx 6.3 \times 10^{-2}$ ($L \approx 16$), and slower dynamics $(0.95, 0.9, 0.85)$ at $W = 8$ give $\sigma_{\min}(L_q) \approx 5.5 \times 10^{-2}$ ($L \approx 18$). Crowding the poles ($a_2 \to a_1$) drives $\sigma_{\min}(L_q) \to 0$ as well, by approaching the boundary of Theorem 2.31(ii).

If instead the dynamics are treated as *unknown*, one must recover $2d + 1 = 7$ parameters from 7 windows, and the reconstruction is only locally valid, on the nondegenerate locus certified for $(3, 8)$ in Lemma 2.17. Near the neutral configuration this locus is empty: Lemma 2.27 shows the neutral fiber is critical for the joint window map, so a jointly-unknown analyst cannot certify exact neutrality at all. This is not a defect of the algorithm but a property of the measurement, and the method reports it as an inconclusive verdict rather than a false certificate.

4.2 Segmented Response Decay in an Aggregated Funnel

A second example arises in coarse activity logs, where only hourly or daily totals are retained. A two-segment decay model is

$$y_n = w_1 a_1^n + w_2 a_2^n, \qquad w_1, w_2 > 0, \quad 0 < a_1, a_2 < 1.$$

The decision structure is the practical content: fix the reporting window length W, test whether the observed windows lie in an identifiable neighborhood, then apply the canonical rule with the appropriate noise tolerance. Two limitations deserve emphasis because they are structural, not incidental. First, *aliasing collisions*: if a latent oscillatory component completes an integer number of cycles per window, block summation annihilates it (Theorem 2.28), and a nonzero signal becomes indistinguishable from the neutral one. Concretely, at $W = 8$ a mode $\cos(2\pi n/8)$ contributes exactly zero to every window sum; an analyst who only sees window totals cannot detect it, and the sound response is inconclusive. Second, *denominator knowledge*: whether certification is possible at all depends on whether the latent dynamics are known. The same data support a sound zero-certificate when the dynamics are fixed (Corollary 2.34) and support no zero-certificate when they are jointly unknown (Corollary 2.30). The framework makes this dependence explicit, so a practitioner knows in advance which observational regime they are in.

4.3 Practical behavior and limitations

Three practical points summarize how the method behaves in use.

- **The verdict is three-valued by design.** An inconclusive output is informative: it certifies that the available window data are compatible with both neutral and non-neutral latent signals, so no sound procedure could do better on that datum (Proposition 3.23). The method never trades this honesty for a forced binary answer.
- **Conditioning is reported, not hidden.** On the certifiable regime the stability constant is the reciprocal smallest singular value of an explicit reconstruction matrix (Proposition 2.33). Dynamics that decay fast relative to the window, poles that crowd, or a window length that aliases the dynamics all show up directly as a large or infinite constant, giving an a-priori check before any data are trusted (see the quantitative example in Section 4.1).
- **Window length is a design variable.** Choosing W coprime to the natural periods of the latent dynamics avoids the annihilation set of Theorem 2.28; the eight-tick window used in our running example is one such choice, and the framework quantifies the cost of a poor choice as loss of identifiability rather than as silent error.

5 Conclusions

We developed a finite-data framework for deciding neutrality of positive configurations from short aggregated observations. The central point is structural: the certification rule is not an arbitrary algorithmic choice, but is determined by the canonical reciprocal cost together with projection, coercivity, and local reconstruction on an identifiable window-measurement locus. On that locus, the procedure is sound and locally maximal among sound rules, and the noisy version yields explicit tolerance bounds for finite-window data.

Equally important is the paper's delineation of what *cannot* be concluded from aggregated data alone. We proved an exact boundary between two regimes. When the latent dynamics are jointly unknown, exact neutrality is uncertifiable from window sums: the neutral fiber is a critical set of the window map and block aggregation annihilates root-of-unity modes, so every sound procedure must return an inconclusive verdict on neutral data (Corollary 2.30). When the dynamics are known, the window map is linear and certification is sound and complete precisely when no mode aliases to a root of unity and distinct modes stay separated after W-fold decimation (Theorem 2.31); on that regime the canonical rule is total and dominates globally (Corollary 2.34).

Practical implications. For a practitioner the framework converts a design question, "can I certify neutrality from the totals I store?", into a checkable spectral condition on the reporting window length and the latent dynamics. The three-valued verdict is a feature: an inconclusive output is a proof that the stored data are genuinely insufficient, not a failure of the algorithm, and the reconstruction conditioning is reported as an explicit constant that flags fragile configurations before any decision is trusted.

Limitations. The certifiable guarantees require either a known denominator (fixed latent dynamics) or restriction to a nondegenerate locus bounded away from neutrality; the sharpest positive statement, global domination, is available only under global injectivity, which fails for the jointly-unknown rational class. The tolerance bounds are deterministic worst-case statements tied to the conditioning of the reconstruction, not statistical confidence statements, and the noise model is a bounded perturbation of the window sums rather than a probabilistic one. The worked examples are synthetic and illustrative; we make no empirical claim.

Future directions. Four directions follow concretely from the results above. (i) A statistical counterpart: replace the bounded-noise tolerance with a probabilistic model and derive confidence-calibrated certificates, using the singular-value conditioning of L_q as the natural variance proxy. (ii) Confluent dynamics: extend Theorem 2.31 to denominators with repeated roots via the confluent Vandermonde structure of Lemma 2.23, giving the boundary for polynomially-modulated modes. (iii) Window-length design: turn the annihilation set of Theorem 2.28 into an optimization over W that maximizes $\sigma_{\min}(L_q)$ for a given family of latent dynamics, yielding provably best reporting granularities. (iv) Multi-channel aggregation: extend the scalar cost and the window operator to vector-valued configurations, where the projection step acts on a product of scale groups and the coercivity certificate becomes a sum over channels.

Acknowledgments

The authors thank Philip Beltracchi, Sebastian Pardo-Guerra, and Milan Zlatanovic for careful reading and detailed comments and suggestions during the preparation of this manuscript. Their feedback improved the clarity, notation, and overall structure of the paper. The authors received no external funding.

Appendix

This appendix collects two supporting components used in the paper. Item 1 records the Recognition Composition Law assumptions and the resulting uniqueness statement for the canonical reciprocal cost. Item 2 records the full integer Jacobian matrix used in the finite-field determinant check for $(d, W) = (3, 8)$.

Item 1. The Recognition Composition Law: Axioms and Uniqueness

In this appendix item 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.

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 item we record the axioms imposed on J.

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

$$J(xy) + J(x/y) = 2J(x) + 2J(y) + 2J(x)J(y). \tag{13}$$

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 $J''_{\log}(0) = 1$.

Derived Normalization and Reciprocity

Proposition A.1 (Normalization Is Forced) *If $J : (0, \infty) \to \mathbb{R}$ is non-constant and satisfies the RCL* (13)*, then $J(1) = 0$.*

Proof Setting $x = y = 1$ in (13) 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 (13) yields $J(x) \equiv -1$ for all $x > 0$, contradicting non-constancy. Hence $J(1) = 0$.

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

Proof Set $x = 1$ in (13) and use Proposition A.1.

Main Uniqueness Theorem

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

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

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 (13) 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, $g(x) = h(\ln x) = \cosh(\ln x) = \frac{1}{2}(x + x^{-1})$ and therefore $J(x) = g(x) - 1 = \frac{1}{2}(x + x^{-1}) - 1 = \frac{(x-1)^2}{2x}$.

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 $J''_{\log}(0) = J''(1)$.

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.

Item 2. 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 2.17.

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

Declarations

Data Availability. Data sharing is not applicable to this article because no real datasets were generated or analyzed in this study.

References

1. Aczél J (1966) Lectures on Functional Equations and Their Applications. Academic Press, New York
2. Bregman LM (1967) The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming. USSR Computational Mathematics and Mathematical Physics 7(3):200–217. https://doi.org/10.1016/0041-5553(67)90040-7
3. Moreau J-J (1965) Proximité et Dualité dans un Espace Hilbertien. Bulletin de la Société Mathématique de France 93:273–299. https://doi.org/10.24033/bsmf.1625
4. Osborne MR, Smyth GK (1995) A Modified Prony Algorithm for Exponential Function Fitting. SIAM Journal on Scientific Computing 16(1):119–138. https://doi.org/10.1137/0916008

5. Peller VV (2003) Hankel Operators and Their Applications. Springer, New York. https://doi.org/10.1007/978-0-387-21681-2
6. Rockafellar RT (1976) Monotone Operators and the Proximal Point Algorithm. SIAM Journal on Control and Optimization 14(5):877–898. https://doi.org/10.1137/0314056
7. Washburn J, Rahnamai Barghi A (2026) Reciprocal Convex Costs for Ratio Matching: Axiomatic Characterization. Axioms 15(2):151. https://doi.org/10.3390/axioms15020151
8. Washburn J, Zlatanović M, Allahyarov E (2026) Recognition Geometry. Axioms 15(2):90. https://doi.org/10.3390/axioms15020090
9. Washburn J, Zlatanović M (2026) Uniqueness of the Canonical Reciprocal Cost. Mathematics 14(6):935. https://doi.org/10.3390/math14060935
10. Washburn J, Zlatanović M, Allahyarov E (2026) The D'Alembert Inevitability Theorem. Mathematics 14(8):1386. https://doi.org/10.3390/math14081386
11. Pardo-Guerra S, Washburn J, Allahyarov E (2026) Admissible Reciprocally Symmetric Costs: Combiner Existence and Classification. Mathematics 14(12):2157. https://doi.org/10.3390/math14122157
12. Pardo-Guerra S, Simons M, Thapa A, Washburn J (2026) Coherent Comparison as Information Cost: Axiomatic Foundations for Discrete Ledger Dynamics. Axioms 15(6):378. https://doi.org/10.3390/axioms15060378
13. Roy R, Kailath T (1989) ESPRIT—Estimation of Signal Parameters via Rotational Invariance Techniques. IEEE Transactions on Acoustics, Speech, and Signal Processing 37(7):984–995. https://doi.org/10.1109/29.32276
14. Vetterli M, Marziliano P, Blu T (2002) Sampling Signals with Finite Rate of Innovation. IEEE Transactions on Signal Processing 50(6):1417–1428. https://doi.org/10.1109/TSP.2002.1003065
15. Candès EJ, Fernandez-Granda C (2014) Towards a Mathematical Theory of Super-Resolution. Communications on Pure and Applied Mathematics 67(6):906–956. https://doi.org/10.1002/cpa.21455
16. Vaidyanathan PP (1993) Multirate Systems and Filter Banks. Prentice Hall, Englewood Cliffs, NJ
17. Lehmann EL, Romano JP (2005) Testing Statistical Hypotheses, 3rd edn. Springer, New York. https://doi.org/10.1007/0-387-27605-X
18. Montgomery DC (2019) Introduction to Statistical Quality Control, 8th edn. Wiley, Hoboken, NJ
19. Kailath T (1980) Linear Systems. Prentice Hall, Englewood Cliffs, NJ
20. Washburn J (2025) A Complete Derivation of the Fermion Spectrum from the Recognition Composition Law. arXiv:2506.12859