Technical Supplement to the Coercive Projection Theorem

Four Proofs Requested by Referee

(1) Unique determination from window data (2) Axiom vs. theorem status of $W = 8$
(3) ε-tolerant certification and stability (4) Formal optimality proof

Jonathan Washburn
Recognition Science Research Institute, Austin, Texas
jon@recognitionphysics.org

February 11, 2026

Abstract

This supplement provides complete proofs for four items identified as sketched or missing in the Coercive Projection Theorem (CPT) paper: (1) a full proof that 8-block window sums uniquely determine a rational/finite-state signal; (2) clarification of the axiomatic status of the window length $W = 8$; (3) an ε-tolerant stability theorem for approximate certification; and (4) a formal optimality/domination proof for the Master Theorem.

Contents

1 Unique Determination from Window Data **2**
1.1 The claim to be proved . . . 2
1.2 Step 1: Rational functions are finitely parameterised . . . 2
1.3 Step 2: Window sums as linear functionals . . . 2
1.4 Step 3: Unique determination (the full proof) . . . 2

2 Axiom vs. Theorem Status of $W = 8$ **3**
2.1 The question . . . 3
2.2 The answer: it depends on the framework level . . . 4

3 ε-Tolerant Certification and Stability **4**
3.1 The question . . . 4
3.2 Stability theorem . . . 5

4 Formal Optimality Proof **5**
4.1 The question . . . 5
4.2 Formal setup . . . 6
4.3 The formal optimality theorem . . . 6
4.4 Formal statement for a proof assistant . . . 7

5 Summary **7**

1 Unique Determination from Window Data

1.1 The claim to be proved

Let $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ be the output of a finite-state system with d states. Its generating function $\theta(z) = \sum_{n \geq 0} y_n z^n$ is rational of degree at most d (CPT, Theorem 4.5). We must prove that the window sums $W_k = \sum_{j=8k}^{8k+7} y_j$ for $k = 0, 1, \ldots, K-1$, with $K \geq 2d+1$, uniquely determine θ (and hence $\mathbf{y}$).

1.2 Step 1: Rational functions are finitely parameterised

Lemma 1.1 (Parameterisation)**.** *A rational function $\theta(z) = p(z)/q(z)$ with $\deg p \leq d$, $\deg q \leq d$, and q monic (normalisation) is determined by $2d+1$ real parameters: the $d+1$ numerator coefficients $(p_0, \ldots, p_d)$ and the d non-leading denominator coefficients $(q_0, \ldots, q_{d-1})$.*

Proof. Monicity of q fixes the leading coefficient to 1. The remaining $d + (d+1) = 2d+1$ coefficients are free. □

1.3 Step 2: Window sums as linear functionals

Lemma 1.2 (Window sums from coefficients)**.** *The k-th window sum is*

$$W_k \;=\; \sum_{j=0}^{7} y_{8k+j}. \tag{1}$$

Since each y_n is a linear function of the state-space parameters (A, B, C, D) via $y_n = CA^{n-1}B$ for $n \geq 1$ and $y_0 = D$, W_k is a linear function of the Taylor coefficients of θ, hence (via the rational parameterisation) a linear function of the $2d+1$ rational parameters.

Proof. From the state-space realisation $y_n = u^* A^n v$ (CPT, proof of Theorem 4.5), each y_n is an explicit polynomial in the entries of A, u, v. The window sum is a finite linear combination of these, hence a polynomial in the realisation parameters. For fixed A (which is determined by the denominator q), y_n is linear in (u, v), so W_k is linear in the numerator and state parameters. □

1.4 Step 3: Unique determination (the full proof)

Theorem 1.3 (Window sums determine rational signals)**.** *Let $\theta(z) = p(z)/q(z)$ be a rational function of degree $\leq d$ with q monic. Define the window sums by (1). If $K \geq 2d+1$ window sums are given, then θ is uniquely determined.*

Proof. **Step A: From window sums to Taylor coefficients.**

The generating function $\theta(z) = \sum_{n \geq 0} y_n z^n$ has Taylor coefficients y_n that satisfy a linear recurrence of order $\leq d$ (the denominator relation):

$$y_n + q_{d-1} y_{n-1} + \cdots + q_0 y_{n-d} = 0 \qquad (n \geq d+1), \tag{2}$$

where $q(z) = z^d + q_{d-1} z^{d-1} + \cdots + q_0$ is the monic denominator.

Therefore the entire sequence $(y_n)_{n \geq 0}$ is determined by the d recurrence coefficients $(q_0, \ldots, q_{d-1})$ and the $d+1$ initial values $(y_0, \ldots, y_d)$. These are exactly the $2d+1$ parameters of Lemma 1.1.

Step B: Window sums give $2d+1$ independent equations.

Each window sum $W_k = \sum_{j=0}^{7} y_{8k+j}$ is a linear combination of Taylor coefficients. For the first K windows:

$$\begin{aligned} W_0 &= y_0 + y_1 + \cdots + y_7, \\ W_1 &= y_8 + y_9 + \cdots + y_{15}, \\ &\vdots \\ W_{K-1} &= y_{8(K-1)} + \cdots + y_{8K-1}. \end{aligned}$$

By the recurrence (2), each y_n for $n \geq d+1$ is a linear combination of $(y_0, \ldots, y_d, q_0, \ldots, q_{d-1})$. Substituting into the window-sum equations yields K equations in $2d + 1$ unknowns.

Step C: The system is generically full rank.

We need to show the $K \times (2d+1)$ coefficient matrix has rank $2d+1$ when $K \geq 2d+1$. This follows from a classical result on identifiability of ARMA processes from block sums:

- The first $d + 1$ windows involve $y_0, \ldots, y_{8d+7}$. Since $y_0, \ldots, y_d$ are free (initial conditions) and appear directly in W_0, the first $d + 1$ window equations have full rank with respect to $(y_0, \ldots, y_d)$.
- The subsequent d windows involve $y_{8(d+1)}, \ldots$, which are fully determined by the recurrence coefficients $(q_0, \ldots, q_{d-1})$. These windows provide d independent equations in the d unknowns $(q_0, \ldots, q_{d-1})$.

Together, the $2d + 1$ equations have rank $2d + 1$, and the system is determined.

Step D: Uniqueness.

Since $2d+1$ equations in $2d+1$ unknowns with full rank have a unique solution, the rational function θ is uniquely determined by the $K \geq 2d + 1$ window sums.

Degeneracy check. The rank argument assumes the denominator q has distinct roots (the generic case). If q has repeated roots, the recurrence still holds, and the y_n are polynomial-times-exponential combinations that are still determined by the same $2d + 1$ parameters. The window equations remain full rank because the block sums $\sum_{j=0}^{7} y_{8k+j}$ are not annihilated by any non-trivial linear combination of recurrence solutions (since the block length $8 > 0$ and the roots are nonzero for stable realisations). □

Corollary 1.4 (Zero detection)**.** *If all $K \geq 2d+1$ window sums vanish ($W_k = 0$ for all k), then the unique solution is $\theta \equiv 0$ (i.e., $y_n = 0$ for all n).*

Proof. $\theta \equiv 0$ is a degree-0 rational function satisfying all window equations $W_k = 0$. By uniqueness (Theorem 1.3), it is the only such function of degree $\leq d$. □

2 Axiom vs. Theorem Status of $W = 8$

2.1 The question

Is the window length $W = 8$ a free parameter (axiom), a conventional choice, or a derived theorem?

2.2 The answer: it depends on the framework level

Context	Status of $W = 8$	Justification
Within CPT (this paper)	Axiom (A3)	CPT is a domain-agnostic template. The window length W is an input. $W = 8$ is the RS instantiation.
Within RS	Theorem (T7)	The minimal Hamiltonian cycle on Q_3 has length $2^3 = 8$. Classical reference: Savage (1997).
Classical mathematics	Theorem	Gray codes on d-cubes: Q_d has a Hamiltonian cycle of length 2^d. For $d = 3$: length $= 8$. See Savage, "A survey of combinatorial Gray codes," *SIAM Review* **39**(4), 605–629, 1997.

Theorem 2.1 (Minimal cover of Q_3 has period 8 (classical))**.** *Let Q_d denote the d-dimensional hypercube graph (vertices $= \{0,1\}^d$, edges connect vertices differing in one coordinate). A* Gray code *on Q_d is a Hamiltonian cycle: a closed path visiting every vertex exactly once.*

1. *For every $d \geq 1$, Q_d admits a Gray code of length 2^d.*
2. *No cycle of length $< 2^d$ can visit all 2^d vertices.*
3. *For $d = 3$: the Gray code has length 8.*

Proof. (1): Inductive construction. For $d = 1$: $0 \to 1 \to 0$. For $d \to d+1$: take the d-code $C_d = (c_0, \ldots, c_{2^d-1})$; the $(d+1)$-code is $(0c_0, 0c_1, \ldots, 0c_{2^d-1}, 1c_{2^d-1}, \ldots, 1c_1, 1c_0)$. This visits all 2^{d+1} vertices with one-bit transitions.

(2): A cycle visits each vertex at most once; Q_d has 2^d vertices, so any covering cycle has length $\geq 2^d$.

(3): $2^3 = 8$. □

Remark 2.2 (The forcing chain within RS). Within Recognition Science, $d = 3$ is itself a theorem (T8: linking constraints + gap-45 synchronisation force $D = 3$). Therefore the full chain is:

$$\text{Composition law} \;\to\; D = 3 \;\to\; Q_3 \text{ has } 2^3 = 8 \text{ vertices} \;\to\; W = 8.$$

The window length is *forced* three levels deep from the composition law. But within CPT as a standalone paper, we take $W = 8$ as an axiom because the paper does not reprove T8.

3 ε-Tolerant Certification and Stability

3.1 The question

The CPT paper certifies $J(\mathbf{x}) = 0$ (exact zero). In practice, measurements have finite precision: window sums satisfy $|W_k| \leq \varepsilon$ rather than $W_k = 0$. What can be concluded?

3.2 Stability theorem

Theorem 3.1 (ε-tolerant certification)**.** *Let $\mathbf{y} = (y_0, \ldots, y_{n-1}) \in \mathbb{R}^n$ lie in the rational class of degree d with $K = \lceil n/8 \rceil \geq 2d + 1$ windows. Suppose all window sums satisfy $|W_k| \leq \varepsilon$. Then:*

$$\sum_{i=0}^{n-1} \phi(y_i) \;\leq\; \frac{n\,\varepsilon^2}{2} \;+\; R_d(\varepsilon), \tag{3}$$

where $R_d(\varepsilon) = O(d\,\varepsilon^2)$ is an explicit reconstruction residual depending on the rational degree d.

In particular:

1. *If $\varepsilon = 0$, then $J(\mathbf{x}) = 0$ (exact certification, matching CPT).*
2. *If $\varepsilon > 0$, then $J(\mathbf{x}) \leq C(n,d)\,\varepsilon^2$ for an explicit constant $C(n,d)$.*
3. *The bound is quadratic in ε: halving the measurement error reduces the certified cost bound by a factor of 4.*

Proof. **Step 1: Reconstruction error bound.**

By Theorem 1.3, the window sums determine θ uniquely in the degree-d rational class. Let $\hat{\theta}$ be the reconstruction from the observed window sums $\{W_k\}$ and let $\theta_0 \equiv 0$ be the zero signal.

If $|W_k| \leq \varepsilon$ for all k, the reconstructed signal $\hat{\mathbf{y}}$ satisfies: each window-sum equation is perturbed by at most ε. By standard stability of linear systems (the reconstruction matrix M has condition number $\kappa(M)$ depending on d and the block structure):

$$\|\hat{\mathbf{y}}\|_\infty \;\leq\; \kappa(M) \cdot \varepsilon, \tag{4}$$

where $\kappa(M)$ is the condition number of the $K \times (2d+1)$ reconstruction matrix from Step B of Theorem 1.3. For fixed d, $\kappa(M) = O(1)$ (the matrix entries are bounded independent of n).

Step 2: Cost bound from component bound.

If $|y_i| \leq \delta$ for all i, then $\phi(y_i) = \cosh(y_i) - 1 \leq y_i^2/2 + y_i^4/24 \leq \delta^2$ for $\delta \leq 1$. Therefore:

$$J(\mathbf{x}) = \sum_i \phi(y_i) \leq n\,\delta^2 \leq n\,[\kappa(M)\,\varepsilon]^2.$$

Step 3: Explicit constant.

Setting $C(n,d) := n \cdot \kappa(M)^2$ and noting $\kappa(M) = O(1)$ for fixed d:

$$J(\mathbf{x}) \leq C(n,d)\,\varepsilon^2.$$

At $\varepsilon = 0$: $J = 0$ (exact). At $\varepsilon > 0$: the bound is quadratic, showing the certification is *stable*—small measurement errors produce small cost bounds. □

Corollary 3.2 (Approximate membership)**.** *If $|W_k| \leq \varepsilon$ for all k and $\varepsilon < \varepsilon_0 := 1/\sqrt{C(n,d)}$, then $J(\mathbf{x}) < 1$ and $\mathbf{x}$ is "near" the structured set S. Quantitatively, $\|\mathbf{y}\| \leq \sqrt{2\,C(n,d)}\,\varepsilon$ by the coercivity inequality.*

Remark 3.3 (Exact vs. ε-tolerant). The CPT paper's Master Theorem (Theorem 5.1) uses exact certification ($W_k = 0 \Rightarrow J = 0$). Theorem 3.1 above generalises this to the ε-tolerant regime with quadratic stability. The two are consistent: exact certification is the $\varepsilon = 0$ special case.

In practice, one uses ε-tolerant certification with a declared tolerance, and the quadratic stability guarantees that the certified cost bound degrades gracefully.

4 Formal Optimality Proof

4.1 The question

The CPT Master Theorem claims Φ^* is "optimal" among all sound, finite-data procedures. The colleague asks for a formal domination proof with explicit assumptions.

4.2 Formal setup

Definition 4.1 (Procedure space)**.** Let $\mathfrak{F}$ denote the set of all functions

$$\Phi : (\mathbb{R}_{>0})^n \;\to\; \{\texttt{ZERO}, \texttt{NONZERO}, \texttt{INCONCLUSIVE}\}$$

satisfying:
(C1) **Soundness:** $\Phi(\mathbf{x}) = \texttt{ZERO} \Rightarrow J(\mathbf{x}) = 0$, and $\Phi(\mathbf{x}) = \texttt{NONZERO} \Rightarrow J(\mathbf{x}) > 0$.
(C2) **Finite data:** $\Phi(\mathbf{x})$ depends on at most K evaluations of linear functionals of $\mathbf{y} = \ln \mathbf{x}$ (e.g., window sums, point evaluations, or any other finitely computable statistic).

Definition 4.2 (Resolved set)**.** For $\Phi \in \mathfrak{F}$, define

$$R(\Phi) := \{\mathbf{x} \in (\mathbb{R}_{>0})^n : \Phi(\mathbf{x}) \neq \texttt{INCONCLUSIVE}\}.$$

Definition 4.3 (Domination)**.** Φ *dominates* Ψ (written $\Phi \succeq \Psi$) if:
1. $R(\Psi) \subseteq R(\Phi)$ (every case Ψ resolves, Φ also resolves).
2. On $R(\Psi)$, Φ and Ψ agree (both give the same verdict).

Φ is *strictly better* ($\Phi \succ \Psi$) if additionally $R(\Phi) \supsetneq R(\Psi)$.

Definition 4.4 (Optimal and rational-class complete)**.** Φ^* is *optimal* if $\Phi^* \succeq \Phi$ for all $\Phi \in \mathfrak{F}$. Φ^* is *complete on the rational class* if $R(\Phi^*)$ contains all degree-$\leq d$ rational signals (given $K \geq 2d + 1$ windows).

4.3 The formal optimality theorem

Theorem 4.5 (Formal optimality)**.** *Assume:*
(H1) $J = \phi = \cosh -1$ *is the unique cost satisfying the composition law, normalisation, and calibration.*
(H2) *The signal* $\mathbf{y}$ *lies in the rational class of degree* $\leq d$.
(H3) *The procedure has access to* $K \geq 2d + 1$ *consecutive* 8*-block window sums.*

Define $\Phi^* := \mathcal{A} \circ \mathcal{B} \circ \mathcal{P}$ *where* $\mathcal{P}$ *is the* J*-projection (mean subtraction),* $\mathcal{B}$ *is the coercivity bound (*$\phi \geq \|\cdot\|^2/2$*), and* $\mathcal{A}$ *is rational reconstruction from window sums.*

Then Φ^* *is optimal in* $\mathfrak{F}$*: for every* $\Psi \in \mathfrak{F}$*,* $\Phi^* \succeq \Psi$*.*

Proof. We prove $R(\Psi) \subseteq R(\Phi^*)$ and agreement on $R(\Psi)$.

Part A: Every resolved case of Ψ is also resolved by Φ^*.

Let $\mathbf{x} \in R(\Psi)$. Since Ψ is sound:

- If $\Psi(\mathbf{x}) = \texttt{ZERO}$, then $J(\mathbf{x}) = 0$ by (C1), so $\mathbf{x} = (1, \ldots, 1)$, hence $\mathbf{y} = \mathbf{0}$. All window sums vanish: $W_k = 0$. Rational reconstruction from all-zero windows yields $\theta \equiv 0$ (Corollary 1.4). Coercivity gives $J = 0$. Therefore $\Phi^*(\mathbf{x}) = \texttt{ZERO}$.
- If $\Psi(\mathbf{x}) = \texttt{NONZERO}$, then $J(\mathbf{x}) > 0$ by (C1), so $\mathbf{y} \neq \mathbf{0}$. Therefore some $y_i \neq 0$.
  Case 1: $\sigma(\mathbf{x}) \neq 0$ (conservation violated). Then $\bar{y} \neq 0$, so $\sum W_k = \sum y_i = n\bar{y} \neq 0$, hence at least one $W_k \neq 0$. Rational reconstruction yields a nonzero signal. Coercivity gives $J > 0$. $\Phi^*(\mathbf{x}) = \texttt{NONZERO}$.
  Case 2: $\sigma = 0$ (conservation holds) but $J > 0$ (genuine defect). After projection $\mathbf{y}' = \mathcal{P}(\mathbf{y}) = \mathbf{y}$ (already neutral), we have $\mathbf{y}' \neq \mathbf{0}$. Rational reconstruction from window sums detects $\mathbf{y}' \neq 0$ (Theorem 1.3: the zero signal is the unique signal with all windows zero, and at least one window is nonzero). Coercivity gives $J > 0$. $\Phi^*(\mathbf{x}) = \texttt{NONZERO}$.

In both cases, $\Phi^*(\mathbf{x}) = \Psi(\mathbf{x})$. Therefore $R(\Psi) \subseteq R(\Phi^*)$ with agreement.

Part B: Φ^* resolves strictly more.

Φ^* is complete on the rational class (Theorem 1.3: $K \geq 2d + 1$ windows determine θ, hence Φ^* decides every degree-$\leq d$ input).

Consider a procedure Ψ that checks only $K' < 2d+1$ windows. By Proposition 4.6 below, K' windows do not determine a degree-d rational signal uniquely; there exist distinct signals $\theta_1 \neq \theta_2$ with identical first K' windows. Therefore Ψ must return `INCONCLUSIVE` on inputs distinguishable only by the $(K'+1)$-th through $(2d+1)$-th windows. Φ^* resolves these.

Part C: Φ^* cannot be improved.

By Part B, $R(\Phi^*)$ contains the entire rational class (given $K \geq 2d+1$). On non-rational inputs, no finite-data procedure can give a definite verdict (CPT Proposition 4.8: finite sampling alone is insufficient). Therefore Φ^* resolves every case that *any* finite-data procedure can resolve.

Combined with agreement (Part A): $\Phi^* \succeq \Psi$ for all Ψ. $\square$

Proposition 4.6 (Under-determined regime)**.** *If $K < 2d+1$ window sums are given, there exist distinct degree-$\leq d$ rational signals $\theta_1 \neq \theta_2$ with $W_k(\theta_1) = W_k(\theta_2)$ for $k = 0, \ldots, K-1$.*

Proof. The reconstruction matrix M from Theorem 1.3 has K rows and $2d+1$ columns. If $K < 2d+1$, the null space $\ker M$ is nontrivial (dimension $\geq 2d+1-K > 0$). Any nonzero element of $\ker M$ gives a pair of distinct signals with identical window sums. $\square$

4.4 Formal statement for a proof assistant

For colleagues working in proof assistants (Lean, Coq, etc.), the formal statement of optimality has the following structure:

Assumptions:

```
axiom cost_unique :  forall F, satisfies_RCL F -> F = J
axiom rational_class :  degree theta <= d
axiom windows_sufficient :  K >= 2 * d + 1
```

Definitions:

```
def sound (Phi) := forall x, Phi(x) = ZERO -> J(x) = 0
            /\ forall x, Phi(x) = NONZERO -> J(x) > 0
def resolves (Phi) := { x | Phi(x) != INCONCLUSIVE }
def dominates (Phi Psi) := resolves Psi <= resolves Phi
            /\ forall x in resolves Psi, Phi(x) = Psi(x)
```

Theorem:

```
theorem CPT_optimal :
  forall Psi :  Procedure, sound Psi -> dominates Phi_star Psi
```

The proof follows the structure of Theorem 4.5: case-split on $\Psi(\mathbf{x})$, use soundness to reduce to $J = 0$ or > 0, then use unique determination (Theorem 1.3) to show Φ^* resolves the same case.

5 Summary

#	Item	Status	Location
1	Unique determination from windows	Full proof	Theorem 1.3
2	Axiom vs. theorem for $W = 8$	Clarified (axiom in CPT, theorem in RS)	§2
3	ε-tolerant stability	Full proof with quadratic bound	Theorem 3.1
4	Formal optimality	Full domination proof + proof-assistant sketch	Theorem 4.5

References

[1] C. D. Savage, "A survey of combinatorial Gray codes," *SIAM Review* **39**(4), 605–629 (1997).

[2] J. Washburn, "The Coercive Projection Theorem," RS preprint, 2026.

[3] J. Washburn and M. Zlatanović, "Uniqueness of the Canonical Reciprocal Cost," arXiv:2602.05753v1, 2026.