Hypergraph Lumpability and Bayesian Refinement for Dynamical Systems with Projector Resolutions

Exact Interaction Quotients, Filtering Cocycles,
and Certified Coarse Models

Komorebi P.

Internal completion release, version 0.1.4
September 2026

Abstract

We study when a coarse description of a nonlinear system is exact rather than fitted after the fact. The fine and coarse state spaces are linked by a surjective linear map C. Each homogeneous term in a local analytic vector field is resolved into packet interactions. Linear terms give ordinary directed edges, while higher order terms give directed hyperedges. We prove that a multilinear interaction of order m descends exactly when its coarse output vanishes whenever any input lies in $\operatorname{Ker} C$. For block sum aggregations, this becomes a direct coefficient test. Once the fine output coefficients are summed within a coarse fiber, the resulting value must depend only on the coarse input fibers.

The quotient condition survives passage to the full analytic field, all of its derivatives, every rooted interaction tree, and the locally unique flow. We represent a Bayesian observation as positive likelihood transfer followed by normalization. This normalized update descends precisely when the fine likelihood is constant within each aggregation fiber, up to a single positive scale. Under those two conditions, the coarse filter agrees with the aggregate of the fine filter for every finite observation sequence, and the evidence factors differ only by the same scales.

When exact descent fails, we keep the sources of error separate. A Grönwall argument turns interaction mismatch into a bound on the flow, while variation of the likelihood inside a fiber creates its own normalization error. A recursive estimate carries both contributions to the final posterior. In the Hilbert Schmidt setting, an orthogonal residual identity also gives the best approximation that retains a prescribed number of hyperedges. Deterministic examples and the reference implementation check the finite formulas, the perturbation bounds, and negative controls. Exact theorems, numerical checks, and claims about historical priority remain distinct throughout.

Status and scope. This manuscript is complete for internal use. It contains the finite theorem chain, proofs, examples, algorithms, certificate code, and reproducibility files. Before submission, it still needs independent proof review and a focused literature review. The algebraic results concern spaces of finite dimension, and the flow results are local in time under the assumptions stated in the text. The numerical checks verify the examples and the implementation, not the proofs. We do not assume that useful packet resolutions always exist, that an optimal quotient is easy to find, that local flow theory guarantees global existence, or that this presentation settles priority across neighboring work.

Contents

1 Introduction **5**

1.1 The quotient problem 5

1.2 Why hypergraphs appear . . . 5
1.3 Why Bayesian observations add another nonlinear operation . . . 6
1.4 Main results . . . 6
1.5 Four distinctions . . . 7
1.6 Organization . . . 7

2 Resolved interaction systems 7
2.1 Finite packet resolutions . . . 7
2.2 Coarse maps, sections, and hidden directions . . . 8
2.3 Multilinear descent . . . 9
2.4 Norm conventions . . . 9

3 Hypergraph lumpability at every interaction order 9
3.1 Invariant quotient criterion . . . 9
3.2 Criterion in packet coordinates . . . 11
3.3 Unary and quadratic cases . . . 12
3.4 Structural consequences . . . 13
3.5 Composition across refinement levels . . . 13

4 Analytic fields, rooted trees, and quotient flows 14
4.1 Descent of polynomial and analytic vector fields . . . 14
4.2 Naturality of derivatives . . . 16
4.3 Rooted interaction trees . . . 16
4.4 Exact quotient flows . . . 17
4.5 Positivity and simplex preservation . . . 18

5 Normalized positive transfer and Bayesian refinement 18
5.1 Positive state spaces . . . 18
5.2 Normalization naturality . . . 19
5.3 Classical Bayesian update . . . 19
5.4 Exact Bayesian refinement . . . 20
5.5 Differential structure of Bayesian normalization . . . 21

6 Filtering cocycles 21
6.1 Finite observation sequences . . . 21
6.2 Sequential evidence . . . 22
6.3 Hidden Markov and nonlinear specializations . . . 23

7 Approximate interaction quotients **23**
7.1 Best coarse interaction and quotient defect . . . 23
7.2 Approximation induced by a section . . . 24
7.3 Defect majorants for the full field . . . 25
7.4 Propagation of dynamical defect . . . 25
7.5 Refinement towers . . . 26

8 Approximate Bayesian refinement **27**
8.1 A normalization perturbation bound . . . 27
8.2 Defect in the observation intertwining law . . . 27
8.3 Likelihood variation inside a fiber . . . 28
8.4 Lipschitz behavior of the Bayesian update . . . 28

9 Certified approximate coarse filtering **29**
9.1 Prediction and update errors . . . 29
9.2 Closed posterior certificate . . . 30
9.3 From interaction coefficients to posterior error . . . 31
9.4 Observable and total variation consequences . . . 31
9.5 Sequential evidence error . . . 31

10 Hyperedge pruning and certified compression **32**
10.1 Hilbert Schmidt tensor blocks . . . 32
10.2 Retained field and named residual . . . 32
10.3 Retaining the k largest energy hyperedges . . . 33
10.4 Hyperedge entropy . . . 33
10.5 Combined quotient and pruning defect . . . 34
10.6 A budgeted design problem . . . 34

11 Algorithms and finite certificates **34**
11.1 Exact test for every interaction order . . . 34
11.2 Approximate coarse coefficients . . . 35
11.3 Certificate pipeline . . . 36
11.4 Deterministic generator for exact examples . . . 36
11.5 Reference implementation . . . 36

12 Worked quotient and filtering models **38**
12.1 A nonlinear fine system with four states and a quotient with two states . . . 38
12.2 Exact nonlinear filtering . . . 40

12.3 A controlled dynamical violation . . . 40
12.4 A controlled observation violation . . . 41
12.5 Random exact tensors through order three . . . 43
12.6 Concentrated hyperedge field . . . 43
12.7 Interpretation for reaction networks . . . 43

13 Relation to neighboring theories **44**
13.1 Markov lumpability and Markov functions . . . 44
13.2 Polynomial differential equivalence . . . 44
13.3 Hypergraphs, operads, and rooted trees . . . 44
13.4 Sufficiency and comparison of experiments . . . 45
13.5 Model reduction and nearly decomposable systems . . . 45

14 Claim boundaries and limitations **45**

15 Open theorem and algorithm program **46**
15.1 Coarsest exact joint partition . . . 46
15.2 Approximate joint design . . . 46
15.3 Posterior distributions over resolutions . . . 46
15.4 Stochastic interaction quotients . . . 46
15.5 Coarse dynamics with memory . . . 47
15.6 Sharper filtering metrics . . . 47
15.7 Pruning that preserves structure . . . 47
15.8 Identifiability of hidden dynamics . . . 47

16 Conclusion **48**

A Formulation on the quotient space **48**

B A finite observational partition lemma **49**

C Certificate schemas and reproducibility **49**
C.1 Execution commands . . . 49

D Theorem dependency ledger **49**

E Notation index **50**

1 Introduction

1.1 The quotient problem

Let X_{f} be a fine state space, let X_{c} be a coarse state space, and let

$$C : X_{\mathrm{f}} \twoheadrightarrow X_{\mathrm{c}}$$

be a surjective linear map. A fine trajectory satisfies

$$\dot{x}(t) = F_{\mathrm{f}}(x(t)), \qquad x(0) = x_0, \tag{1}$$

and the coarse observable of interest is

$$z(t) = Cx(t).$$

The first question is whether $z(t)$ obeys a closed equation of its own. Equivalently, can we define a vector field F_{c} such that

$$CF_{\mathrm{f}} = F_{\mathrm{c}}C? \tag{2}$$

When this identity holds and both initial value problems have unique solutions, aggregation commutes with evolution:

$$C\Phi^{\mathrm{f}}_t = \Phi^{\mathrm{c}}_t C. \tag{3}$$

The coarse trajectory is then an exact quotient of the fine one. No fitting step is involved.

The linear versions of this question are classical. For finite Markov chains, they appear as lumpability and Markov function criteria [12, 14, 5]. Nearly decomposable systems lead instead to perturbative aggregation [15, 13]. Polynomial systems and reaction networks admit exact differential equivalences and reductions based on bisimulation [7, 8]. We isolate a quotient test for each order of a multilinear interaction and then follow it through nonlinear composition, continuous evolution, and Bayesian observation.

1.2 Why hypergraphs appear

Suppose a local analytic vector field has the expansion

$$F_{\mathrm{f}}(x) = \sum_{m \geq 0} A^{\mathrm{f}}_m[x, \ldots, x]. \tag{4}$$

The linear coefficient A^{f}_1 moves one input packet into one output packet. The quadratic coefficient A^{f}_2 combines two input packets before producing an output. In general, A^{f}_m combines m inputs. We record this incidence by the directed hyperedge

$$(j_1, \ldots, j_m) \longrightarrow i.$$

Nonlinear composition is not ordinary path concatenation. It inserts the output of one multilinear interaction into an input of another. Repeated insertion produces rooted trees and is governed by the Faà di Bruno formula [9]. The same tree structure appears in nonlinear evolution and in numerical integration [6, 10].

At order m, the central question is

$$\exists A^{\mathrm{c}}_m \quad \text{such that} \quad CA^{\mathrm{f}}_m = A^{\mathrm{c}}_m C^{\times m}? \tag{5}$$

If this question has an answer at every order, the quotient relation can be carried from individual interactions to the full field, its derivatives, its interaction trees, and its flow.

1.3 Why Bayesian observations add another nonlinear operation

Prediction is only one part of a filter. Given a probability vector μ and a positive likelihood λ, Bayesian conditioning is

$$\mathsf{B}_\lambda(\mu) = \frac{\lambda \odot \mu}{\lambda^T \mu}. \tag{6}$$

Likelihood multiplication is linear, but the evidence in the denominator couples all packets. The full update is therefore nonlinear even when the likelihood acts separately on each state.

Classical sufficiency theory asks whether a statistic preserves all information relevant to a chosen observation family [11, 4]. Here the operator question is

$$C\mathsf{B}_{\lambda^{\mathrm{f}}} = \mathsf{B}_{\lambda^{\mathrm{c}}} C. \tag{7}$$

When both the flow identity (3) and the Bayesian identity (7) hold at each observation time, the whole filtering sequence commutes with aggregation. Hidden Markov filtering is the linear unary case [3]. The same argument applies when the prediction step comes from a nonlinear quotient flow.

1.4 Main results

Eight results form the mathematical spine of the paper.

C1. A quotient criterion for every interaction order. For a surjective map C, a multilinear fine interaction descends exactly when CA^{f}_m vanishes whenever one input lies in $\operatorname{Ker} C$. Equivalently, its linearization on $X_{\mathrm{f}}^{\otimes m}$ factors through $C^{\otimes m}$.

C2. A concrete packet test. For block sum aggregation, exact descent is equivalent to constancy of the aggregated fine coefficients over each product of input fibers. These common values determine the coarse interaction uniquely.

C3. Compatibility across several resolutions. Exact descent is transitive. Approximate descent obeys an exact defect identity that can be iterated along a refinement tower.

C4. Naturality of trees and flows. Descent of the coefficients passes to local analytic fields, all derivatives, every rooted interaction tree, and every locally unique flow.

C5. Naturality after normalization. If a positive observation transfer respects the quotient up to a scalar, normalization removes that scalar and preserves the quotient. For diagonal likelihoods, this happens exactly when the fine likelihood is constant on each fiber, up to one global scale.

C6. Exact filtering cocycles. When both dynamics and observations descend, every finite sequence of predictions and updates agrees after aggregation. The evidence factors obey the same quotient law.

C7. A certificate for approximate filtering. Interaction mismatch first appears as flow error, while observation mismatch appears at normalization. A closed recurrence carries both contributions to the final posterior and to the accumulated log evidence.

C8. A separate pruning certificate. Discarded hyperedges remain in a named residual tensor. Orthogonal packet blocks give an exact residual identity and an optimal retention theorem for a fixed number of hyperedges.

The proof strategy can be summarized by the chain

$$\boxed{\begin{array}{c}\text{multilinear quotient} \Longrightarrow \text{interaction tree quotient} \Longrightarrow \text{flow quotient} \\ \Longrightarrow \text{Bayesian quotient} \Longrightarrow \text{filtering cocycle quotient} \Longrightarrow \text{posterior certificate.}\end{array}}$$

1.5 Four distinctions

The same partition can succeed at one stage of the problem and fail at another. Four distinctions prevent those cases from being conflated.

(i) A *resolution* states which packets are being distinguished.

(ii) An *exact quotient* proves that the coarse dynamics do not depend on hidden coordinates.

(iii) A *compression* deletes or approximates interaction blocks, so the omitted part must remain visible as a residual.

(iv) *Statistical sufficiency* concerns the information retained for a chosen observation family. Here it is expressed by naturality of the normalized transfer.

A partition may be exact for the dynamics yet lose information carried by the observations. It may preserve the observation model while failing to close the dynamics. It may also work through one interaction order and fail at the next. Each failure receives its own defect.

1.6 Organization

Section 2 introduces packet resolutions and interaction tensors. Section 3 proves the quotient criteria. Section 4 carries the result to analytic fields, rooted trees, and flows. Section 5 studies normalized positive transfer and Bayesian refinement. Section 6 combines prediction and observation into an exact filtering cocycle. Sections 7 to 9 develop the error bounds. Section 10 keeps track of residuals created by hyperedge removal. Sections 11 and 12 describe the finite algorithms, numerical checks, and worked models. The final sections compare the framework with nearby theories, then state the limitations and open problems.

2 Resolved interaction systems

2.1 Finite packet resolutions

The main algebraic results are stated for vector spaces of finite dimension over $\mathbb{R}$ or $\mathbb{C}$. The quotient and tensor factorizations are then purely algebraic. The later analytic results add the stated convergence conditions and the usual assumptions for local uniqueness.

Definition 2.1 (Packet resolution)**.** A packet resolution of a vector space X is a finite family of projections $(P_i)_{i\in I}$ satisfying

$$P_iP_j = \delta_{ij}P_i, \qquad \sum_{i\in I} P_i = I_X.$$

Thus

$$X = \bigoplus_{i\in I} X_i, \qquad X_i = P_iX.$$

When X is a Hilbert space, the principal norm identities will assume that the P_i are orthogonal.

Let $(P_j)_{j\in J}$ resolve a source X and $(Q_i)_{i\in I}$ resolve a target Y.

Definition 2.2 (Resolved multilinear block and hyperedge)**.** For a multilinear map $A : X^m \to Y$ of order m, define

$$A_{i|j_1,\dots,j_m} = Q_i A(P_{j_1}\cdot,\dots,P_{j_m}\cdot). \tag{8}$$

The ordered incidence

$$(j_1,\dots,j_m) \longrightarrow i$$

is the corresponding directed interaction hyperedge. We call the hyperedge active when the block in (8) is nonzero.

When $m = 1$, the active blocks form an ordinary directed transfer graph. When $m = 2$, each active block joins a pair of input packets. For general m, the block records one interaction with m inputs.

Theorem 2.3 (Exact multilinear packet reconstruction)**.** *Every multilinear map $A : X^m \to Y$ of order m satisfies*

$$A(x_1,\dots,x_m) = \sum_{i,j_1,\dots,j_m} A_{i|j_1,\dots,j_m}(P_{j_1}x_1,\dots,P_{j_m}x_m). \tag{9}$$

Proof. Insert $I_X = \sum_j P_j$ in every input slot and $I_Y = \sum_i Q_i$ after applying A. Multilinearity distributes all finite sums and gives (9). □

2.2 Coarse maps, sections, and hidden directions

Let

$$C : X_{\mathrm{f}} \twoheadrightarrow X_{\mathrm{c}}$$

be a surjective linear coarse map. Finite dimensionality guarantees a linear right inverse

$$E : X_{\mathrm{c}} \to X_{\mathrm{f}}, \qquad CE = I_{X_{\mathrm{c}}}.$$

Define

$$P_C = EC, \qquad Q_C = I_{X_{\mathrm{f}}} - EC. \tag{10}$$

Then

$$P_C^2 = P_C, \qquad Q_C^2 = Q_C, \qquad \operatorname{Ran} Q_C = \operatorname{Ker} C.$$

Every fine state has the exact decomposition

$$x = ECx + Q_C x. \tag{11}$$

The first term is the chosen fine representative of the coarse state. The second lies in $\operatorname{Ker} C$, the hidden subspace. Exact quotienting means that motion inside this subspace cannot change the aggregated interaction output.

Definition 2.4 (Compatible packet aggregation)**.** Suppose $(P_i)_{i\in I_{\mathrm{f}}}$ and $(\overline{P}_a)_{a\in I_{\mathrm{c}}}$ resolve the fine and coarse spaces. A surjection

$$\pi : I_{\mathrm{f}} \twoheadrightarrow I_{\mathrm{c}}$$

is compatible with C when

$$\overline{P}_a C = C \sum_{\pi(i)=a} P_i \qquad (a \in I_{\mathrm{c}}). \tag{12}$$

In the scalar coordinate model used later, each P_i selects one coordinate and C sums all coordinates in the same fiber of π.

2.3 Multilinear descent

Definition 2.5 (Exact interaction quotient)**.** Let $A_m^{\mathrm{f}} : X_{\mathrm{f}}^m \to X_{\mathrm{f}}$ and $A_m^{\mathrm{c}} : X_{\mathrm{c}}^m \to X_{\mathrm{c}}$ be multilinear maps of order m. We say that A_m^{f} *descends through* C to A_m^{c}, and write

$$A_m^{\mathrm{f}} \Downarrow_C A_m^{\mathrm{c}},$$

when

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}} C^{\times m}. \tag{13}$$

For $m = 0$, descent means $CA_0^{\mathrm{f}} = A_0^{\mathrm{c}}$.

Here

$$C^{\times m}(x_1, \dots, x_m) = (Cx_1, \dots, Cx_m).$$

We write $C^{\otimes m}$ for the induced linear map between algebraic tensor powers.

Definition 2.6 (Interaction defect)**.** For a proposed coarse tensor B_m, define

$$\Delta_m[B_m] = CA_m^{\mathrm{f}} - B_m C^{\times m}. \tag{14}$$

Exact descent is the condition $\Delta_m[B_m] = 0$.

2.4 Norm conventions

For a multilinear map $A : X^m \to Y$ of order m, let

$$\|A\| = \sup_{\|x_r\| \le 1} \|A(x_1, \dots, x_m)\|.$$

If E is a fixed right inverse of C, then any exact coarse tensor necessarily satisfies

$$A_m^{\mathrm{c}}(z_1, \dots, z_m) = CA_m^{\mathrm{f}}(Ez_1, \dots, Ez_m), \tag{15}$$

so

$$\|A_m^{\mathrm{c}}\| \le \|C\| \left\|A_m^{\mathrm{f}}\right\| \|E\|^m.$$

The formula also makes the later uniqueness argument immediate.

3 Hypergraph lumpability at every interaction order

3.1 Invariant quotient criterion

The quotient problem has a simple algebraic answer.

Theorem 3.1 (Universal multilinear quotient criterion)**.** *Let $C : X_{\mathrm{f}} \twoheadrightarrow X_{\mathrm{c}}$ be surjective and let*

$$A_m^{\mathrm{f}} : X_{\mathrm{f}}^m \to X_{\mathrm{f}}$$

be a multilinear map of order m, where $m \ge 1$. The following statements are equivalent.

(i) There exists a multilinear map $A_m^{\mathrm{c}} : X_{\mathrm{c}}^m \to X_{\mathrm{c}}$ of order m such that

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}} C^{\times m}.$$

(ii) For every slot $r \in \{1, \dots, m\}$,

$$CA_m^{\mathrm{f}}(x_1, \dots, x_{r-1}, k, x_{r+1}, \dots, x_m) = 0 \tag{16}$$

whenever $k \in \operatorname{Ker} C$.

(iii) The map CA_m^{f} is constant on products of fibers: if $Cx_r = Cx_r'$ for all r, then

$$CA_m^{\mathrm{f}}(x_1, \dots, x_m) = CA_m^{\mathrm{f}}(x_1', \dots, x_m'). \tag{17}$$

(iv) Let $\widehat{A}_m^{\mathrm{f}} : X_{\mathrm{f}}^{\otimes m} \to X_{\mathrm{c}}$ be the linearization of CA_m^{f}. Then

$$\operatorname{Ker}(C^{\otimes m}) \subseteq \operatorname{Ker}(\widehat{A}_m^{\mathrm{f}}). \tag{18}$$

When these conditions hold, the coarse tensor is unique and, for every right inverse E of C,

$$A_m^{\mathrm{c}}(z_1, \dots, z_m) = CA_m^{\mathrm{f}}(Ez_1, \dots, Ez_m). \tag{19}$$

The right side is independent of the chosen right inverse.

Proof. Assume (i). If $k \in \operatorname{Ker} C$, then inserting k into any input slot gives

$$CA_m^{\mathrm{f}}(x_1, \dots, k, \dots, x_m) = A_m^{\mathrm{c}}(Cx_1, \dots, 0, \dots, Cx_m) = 0,$$

so (ii) holds.

Assume (ii). If $Cx_r = Cx_r'$, then $x_r' - x_r \in \operatorname{Ker} C$. Replace the inputs one at a time. Multilinearity and (16) show that each replacement leaves CA_m^{f} unchanged. Hence (iii) holds.

Assume (iii). For $z_r \in X_{\mathrm{c}}$ choose arbitrary lifts x_r with $Cx_r = z_r$ and set

$$A_m^{\mathrm{c}}(z_1, \dots, z_m) = CA_m^{\mathrm{f}}(x_1, \dots, x_m).$$

Fiber constancy makes the definition independent of the lifts. Multilinearity is inherited from A_m^{f}, and the defining identity gives (i). Surjectivity of C also gives uniqueness: every coarse input tuple has a fine lift.

To compare (ii) and (iv), identify a multilinear map of order m with its linearization on the algebraic tensor product. Since C is surjective,

$$\operatorname{Ker}(C^{\otimes m}) = \sum_{r=1}^{m} X_{\mathrm{f}}^{\otimes(r-1)} \otimes \operatorname{Ker} C \otimes X_{\mathrm{f}}^{\otimes(m-r)}.$$

Choose a splitting $X_{\mathrm{f}} = E(X_{\mathrm{c}}) \oplus \operatorname{Ker} C$ and expand the tensor power. Vanishing of $\widehat{A}_m^{\mathrm{f}}$ on this sum is exactly the slot condition. Thus (ii) and (iv) are equivalent. Formula (19) follows by choosing $x_r = Ez_r$. □

Remark 3.2 (Nonlinear congruence interpretation). For a general map $F : X_{\mathrm{f}} \to X_{\mathrm{f}}$, descent through C is equivalent to

$$Cx = Cy \quad \Longrightarrow \quad CF(x) = CF(y).$$

Thus the relation $x \sim_C y$, defined by $Cx = Cy$, must be a congruence for the coarse output of F. Theorem 3.1 turns that statement into a condition on each multilinear coefficient.

Corollary 3.3 (Bounded coarse tensor)**.** *Under the hypotheses of Theorem 3.1, for every bounded right inverse E,*

$$\|A_m^{\mathrm{c}}\| \le \|C\| \left\|A_m^{\mathrm{f}}\right\| \|E\|^m .$$

Proof. Apply (19) to unit coarse inputs. □

3.2 Criterion in packet coordinates

For a block sum aggregation, the kernel condition can be read directly from the tensor coefficients.

Let

$$I_{\mathrm{f}} = \{1, \dots, n\}, \qquad I_{\mathrm{c}} = \{1, \dots, q\},$$

and let

$$\pi : I_{\mathrm{f}} \twoheadrightarrow I_{\mathrm{c}}.$$

Write (e_i) and (f_a) for the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^q$. Define

$$Ce_i = f_{\pi(i)}, \qquad (Cx)_a = \sum_{\pi(i)=a} x_i. \tag{20}$$

Suppose

$$A_m^{\mathrm{f}}(e_{j_1}, \dots, e_{j_m}) = \sum_{i=1}^{n} a^{\mathrm{f},(m)}_{i;j_1,\dots,j_m} e_i. \tag{21}$$

For each coarse output packet a, sum the fine coefficients over its output fiber and define

$$g^{(m)}_{a;j_1,\dots,j_m} = \sum_{i:\pi(i)=a} a^{\mathrm{f},(m)}_{i;j_1,\dots,j_m}. \tag{22}$$

Theorem 3.4 (Coefficient criterion for every interaction order)**.** *The fine tensor A_m^{f} descends through the block sum map (20) if and only if, for every coarse output a,*

$$g^{(m)}_{a;j_1,\dots,j_m} = g^{(m)}_{a;j'_1,\dots,j'_m} \tag{23}$$

whenever

$$\pi(j_r) = \pi(j'_r) \qquad (r = 1, \dots, m).$$

When this condition holds, the unique coarse coefficients are

$$a^{\mathrm{c},(m)}_{a;b_1,\dots,b_m} = g^{(m)}_{a;j_1,\dots,j_m}, \qquad \pi(j_r) = b_r. \tag{24}$$

Proof. Applying C to (21) gives

$$CA_m^{\mathrm{f}}(e_{j_1}, \dots, e_{j_m}) = \sum_{a=1}^{q} g^{(m)}_{a;j_1,\dots,j_m} f_a.$$

If a coarse tensor exists, this vector equals

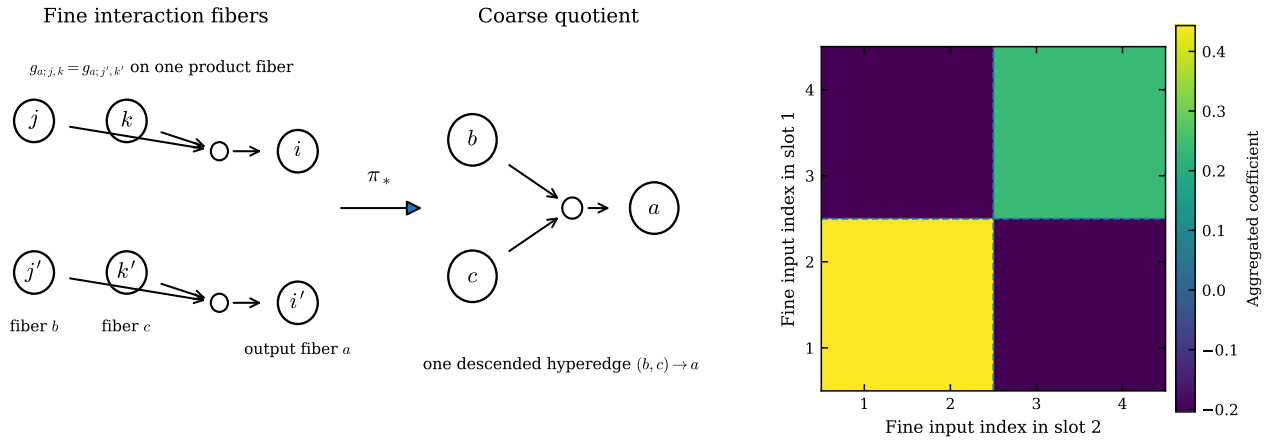
$$A_m^{\mathrm{c}}(f_{\pi(j_1)}, \dots, f_{\pi(j_m)}),$$

which depends only on the coarse input labels. This proves necessity. Conversely, constancy on each fiber makes (24) well defined. Equality on every tuple of fine basis vectors then gives

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}} C^{\times m}$$

by multilinearity. □

For $m = 2$, the criterion has two useful views. The hypergraph records which input fibers interact, while the coefficient matrix shows whether those interactions are constant on each product fiber. The exact synthetic tensor used in the certificate run has the required block structure to floating point precision.



(a) Fine interactions descend to one coarse hyperedge.

(b) Aggregated coefficients are constant on every product fiber.

Figure 1: Two views of the quadratic quotient criterion. The incidence view identifies the descended hyperedge. The coefficient view displays the fiber constancy condition from Theorem 3.4.

3.3 Unary and quadratic cases

Corollary 3.5 (Unary aggregate rate criterion)**.** *For $m = 1$, exact descent is equivalent to*

$$\sum_{i:\pi(i)=a} a^{\mathrm{f},(1)}_{i;j} = \sum_{i:\pi(i)=a} a^{\mathrm{f},(1)}_{i;j'} \tag{25}$$

for all coarse outputs a whenever $\pi(j) = \pi(j')$.

Proof. Set $m = 1$ in Theorem 3.4. The condition then involves one input fiber and becomes (25). □

Let Q_{f} be the generator of a continuous time Markov chain, using the column vector convention. The left side of (25) is the total transition rate from fine state j into coarse packet a. Corollary 3.5 is therefore the standard criterion for strong lumpability.

Corollary 3.6 (Quadratic hypergraph criterion)**.** *For $m = 2$, exact descent is equivalent to*

$$\sum_{i:\pi(i)=a} a^{\mathrm{f},(2)}_{i;j,k} = \sum_{i:\pi(i)=a} a^{\mathrm{f},(2)}_{i;j',k'}$$

whenever

$$\pi(j) = \pi(j'), \qquad \pi(k) = \pi(k').$$

The resulting coarse interaction is the directed quadratic hyperedge

$$(\pi(j), \pi(k)) \longrightarrow a.$$

Proof. Set $m = 2$ in Theorem 3.4. Fiber constancy in the two input slots gives the displayed equality, and (24) assigns its common value to the coarse hyperedge. □

Example 3.7 (A quotient with invisible internal interactions)**.** Let $X_{\mathrm{c}} = \mathbb{R}^q$, $X_{\mathrm{f}} = \mathbb{R}^{2q}$, and

$$C = \begin{bmatrix} I_q & I_q \end{bmatrix}, \qquad E = \frac{1}{2}\begin{bmatrix} I_q \\ I_q \end{bmatrix}.$$

Let A_m^{c} be arbitrary and let $K_m : X_{\mathrm{f}}^m \to X_{\mathrm{f}}$ satisfy

$$CK_m = 0.$$

Then

$$A_m^{\mathrm{f}} = EA_m^{\mathrm{c}}C^{\times m} + K_m \tag{26}$$

satisfies

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}}C^{\times m}.$$

The tensor K_m may contain many fine hyperedges, so the fine model can remain complicated. Exact quotienting asks only that every hidden contribution disappear after applying C.

3.4 Structural consequences

Proposition 3.8 (Symmetry descends)**.** *If A_m^{f} is symmetric in its input slots and descends through C, then the unique coarse tensor A_m^{c} is symmetric.*

Proof. Choose lifts of the coarse inputs. Permuting the coarse inputs permutes their lifts, and symmetry of A_m^{f} leaves (19) unchanged. □

Proposition 3.9 (Quotient support)**.** *Under the coordinate criterion, every active coarse hyperedge*

$$(b_1, \ldots, b_m) \longrightarrow a$$

is supported by at least one fine hyperedge whose input and output labels map to that coarse incidence. The converse can fail because several active fine hyperedges may cancel after aggregation.

Proof. If every fine coefficient in the corresponding incidence fiber vanishes, then the aggregate coefficient (22) vanishes. Conversely, nonzero fine coefficients can sum to zero after output aggregation. □

3.5 Composition across refinement levels

Let

$$X_2 \xrightarrow{C_{21}} X_1 \xrightarrow{C_{10}} X_0, \qquad C_{20} = C_{10}C_{21}.$$

Proposition 3.10 (Transitivity of exact interaction quotients)**.** *If*

$$A_m^{(2)} \Downarrow_{C_{21}} A_m^{(1)} \qquad \text{and} \qquad A_m^{(1)} \Downarrow_{C_{10}} A_m^{(0)},$$

then

$$A_m^{(2)} \Downarrow_{C_{20}} A_m^{(0)}.$$

Proof. Compose the two identities:

$$C_{20}A_m^{(2)} = C_{10}A_m^{(1)}C_{21}^{\times m} = A_m^{(0)}C_{10}^{\times m}C_{21}^{\times m} = A_m^{(0)}C_{20}^{\times m}.$$

□

Approximate descent has an exact composition law.

Theorem 3.11 (Interaction defect across two levels)**.** *Define*

$$\Delta_m^{21} = C_{21}A_m^{(2)} - A_m^{(1)}C_{21}^{\times m},$$

$$\Delta_m^{10} = C_{10}A_m^{(1)} - A_m^{(0)}C_{10}^{\times m},$$

and

$$\Delta_m^{20} = C_{20}A_m^{(2)} - A_m^{(0)}C_{20}^{\times m}.$$

Then

$$\boxed{\Delta_m^{20} = C_{10}\Delta_m^{21} + \Delta_m^{10}C_{21}^{\times m}.} \tag{27}$$

Consequently,

$$\left\|\Delta_m^{20}\right\| \le \|C_{10}\| \left\|\Delta_m^{21}\right\| + \left\|\Delta_m^{10}\right\| \|C_{21}\|^m. \tag{28}$$

Proof. Add and subtract $C_{10}A_m^{(1)}C_{21}^{\times m}$:

$$\begin{aligned}\Delta_m^{20} &= C_{10}C_{21}A_m^{(2)} - A_m^{(0)}C_{10}^{\times m}C_{21}^{\times m} \\ &= C_{10}(C_{21}A_m^{(2)} - A_m^{(1)}C_{21}^{\times m}) \\ &\quad + (C_{10}A_m^{(1)} - A_m^{(0)}C_{10}^{\times m})C_{21}^{\times m}.\end{aligned}$$

This proves (27). Taking norms and using submultiplicativity yields the estimate. □

4 Analytic fields, rooted trees, and quotient flows

4.1 Descent of polynomial and analytic vector fields

Let

$$F_{\mathrm{f}}(x) = \sum_{m=0}^{\infty} A_m^{\mathrm{f}}[x^{\otimes m}] \tag{29}$$

be locally analytic, where A_m^{f} is a symmetric multilinear map of order m for $m \ge 1$ and $A_0^{\mathrm{f}} \in X_{\mathrm{f}}$. Throughout,

$$A_m[x^{\otimes m}] = A_m(x, \dots, x).$$

Theorem 4.1 (Analytic interaction descent)**.** *Let $C : X_{\mathrm{f}} \twoheadrightarrow X_{\mathrm{c}}$ have a bounded right inverse E. Suppose the coefficient series converges normally on every smaller ball,*

$$\sum_{m=0}^{\infty} \left\|A_m^{\mathrm{f}}\right\| r^m < \infty \qquad (0 \le r < R), \tag{30}$$

and, for every $m \ge 0$, there is a coarse coefficient A_m^{c} satisfying

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}}C^{\times m}. \tag{31}$$

Then

$$F_{\mathrm{c}}(z) = \sum_{m=0}^{\infty} A_m^{\mathrm{c}}[z^{\otimes m}] \tag{32}$$

converges absolutely whenever $\|E\|\,\|z\| < R$*, and*

$$\boxed{CF_{\mathrm{f}}(x) = F_{\mathrm{c}}(Cx)} \tag{33}$$

for every x *in the fine convergence domain whose coarse image lies in the coarse domain.*

Proof. For $m \geq 1$, Corollary 3.3 gives

$$\|A_m^{\mathrm{c}}\| \leq \|C\| \left\|A_m^{\mathrm{f}}\right\| \|E\|^m .$$

For $m = 0$, the identity $A_0^{\mathrm{c}} = CA_0^{\mathrm{f}}$ gives the same bound with $\|E\|^0 = 1$. Therefore

$$\sum_{m\geq 0} \|A_m^{\mathrm{c}}\|\,\|z\|^m \leq \|C\| \sum_{m\geq 0} \left\|A_m^{\mathrm{f}}\right\| (\|E\|\,\|z\|)^m,$$

which converges when $\|E\|\,\|z\| < R$. Coefficientwise descent and absolute convergence permit termwise aggregation:

$$CF_{\mathrm{f}}(x) = \sum_{m\geq 0} CA_m^{\mathrm{f}}[x^{\otimes m}] = \sum_{m\geq 0} A_m^{\mathrm{c}}[(Cx)^{\otimes m}] = F_{\mathrm{c}}(Cx).$$

□

Remark 4.2 (Coefficientwise descent is sufficient, not always necessary)*.* Descent of each displayed coefficient is enough to make the whole series descend. The converse can fail for a noncanonical expansion that splits one homogeneous degree into several terms whose coarse contributions cancel. For the Taylor expansion at a fixed base point, the derivatives are unique, so the converse does hold. Differentiating $CF_{\mathrm{f}} = F_{\mathrm{c}}C$ then recovers every descended jet.

Theorem 4.3 (Coefficientwise characterization of polynomial descent)**.** *Let*

$$F_{\mathrm{f}}(x) = \sum_{m=0}^{M} A_m^{\mathrm{f}}[x^{\otimes m}], \qquad F_{\mathrm{c}}(z) = \sum_{m=0}^{M} A_m^{\mathrm{c}}[z^{\otimes m}],$$

where every A_m^{f} *and* A_m^{c} *is symmetric. Then*

$$CF_{\mathrm{f}} = F_{\mathrm{c}}C$$

if and only if

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}}C^{\times m} \qquad (0 \leq m \leq M).$$

Proof. The reverse implication follows from Theorem 4.1, applied to a finite sum. Conversely, fix $x \in X_{\mathrm{f}}$ and evaluate the field identity at tx:

$$\sum_{m=0}^{M} t^m CA_m^{\mathrm{f}}[x^{\otimes m}] = \sum_{m=0}^{M} t^m A_m^{\mathrm{c}}[(Cx)^{\otimes m}].$$

Equality of the coefficients of this polynomial with values in X_{c} gives equality on the diagonal for the two symmetric multilinear maps of order m. The polarization identity recovers each multilinear map from its diagonal polynomial, so

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}}C^{\times m}$$

for every degree. □

4.2 Naturality of derivatives

Proposition 4.4 (Jet naturality)**.** *Let F_{f} and F_{c} be C^r maps satisfying*

$$CF_{\mathrm{f}} = F_{\mathrm{c}}C.$$

Then, for every $1 \le k \le r$,

$$\boxed{CD^kF_{\mathrm{f}}(x)[h_1,\dots,h_k] = D^kF_{\mathrm{c}}(Cx)[Ch_1,\dots,Ch_k].} \tag{34}$$

Proof. Differentiate the identity $CF_{\mathrm{f}}(x) = F_{\mathrm{c}}(Cx)$. Since C is linear, its derivatives of order greater than one vanish and the ordinary chain rule gives the first order statement. Repeated differentiation gives (34). □

Thus the quotient preserves more than the value of the vector field. At each base point, every fine derivative block maps to its coarse counterpart. The same identity controls local perturbations, stability calculations, and higher order response.

4.3 Rooted interaction trees

Repeated substitution of nonlinear interactions produces rooted trees. Only a simple recursive definition is needed for the quotient argument.

Definition 4.5 (Interaction tree evaluation)**.** Let $\mathcal{A} = (A_m)_{m\ge1}$ be a family of multilinear maps $A_m : X^m \to X$. A rooted interaction tree is defined recursively. A leaf $\bullet$ has evaluation

$$\mathcal{V}_\bullet = I_X.$$

If

$$\tau = B_+(\tau_1,\dots,\tau_m)$$

is formed by adjoining a root of arity m, define

$$\mathcal{V}_\tau(x_1,\dots,x_{|\tau|}) = A_m(\mathcal{V}_{\tau_1}(\mathbf{x}_1),\dots,\mathcal{V}_{\tau_m}(\mathbf{x}_m)), \tag{35}$$

where $|\tau|$ denotes the number of leaves and the input list is partitioned among the subtrees in their fixed order.

The operation B_+ inserts several rooted trees beneath a new root. In packet coordinates, every edge carries a packet label and every internal vertex must match an active resolved hyperedge.

Theorem 4.6 (Treewise quotient naturality)**.** *Suppose every operation in a fine family descends:*

$$CA^{\mathrm{f}}_m = A^{\mathrm{c}}_m C^{\times m} \qquad (m \ge 1).$$

Then, for every rooted interaction tree τ,

$$\boxed{C\mathcal{V}^{\mathrm{f}}_\tau = \mathcal{V}^{\mathrm{c}}_\tau C^{\times|\tau|}.} \tag{36}$$

Proof. Proceed by induction on the number of internal vertices. The leaf case is $CI = IC$. Suppose

$$\tau = B_+(\tau_1, \dots, \tau_m).$$

Using interaction descent at the root and the inductive hypothesis on every subtree,

$$\begin{aligned} C\mathcal{V}^{\mathrm{f}}_{\tau} &= CA^{\mathrm{f}}_m(\mathcal{V}^{\mathrm{f}}_{\tau_1}, \dots, \mathcal{V}^{\mathrm{f}}_{\tau_m}) \\ &= A^{\mathrm{c}}_m(C\mathcal{V}^{\mathrm{f}}_{\tau_1}, \dots, C\mathcal{V}^{\mathrm{f}}_{\tau_m}) \\ &= A^{\mathrm{c}}_m(\mathcal{V}^{\mathrm{c}}_{\tau_1} C^{\times|\tau_1|}, \dots, \mathcal{V}^{\mathrm{c}}_{\tau_m} C^{\times|\tau_m|}) \\ &= \mathcal{V}^{\mathrm{c}}_{\tau} C^{\times|\tau|}. \end{aligned}$$

□

Corollary 4.7 (Compatibility with Faà di Bruno composition)**.** *Let F_{f} and F_{c} be smooth maps with $CF_{\mathrm{f}} = F_{\mathrm{c}}C$, and let G_{f} and G_{c} satisfy the same relation. Then each rooted tree in the Faà di Bruno expansion, written in packet coordinates, descends to the corresponding coarse tree. Moreover,*

$$C(G_{\mathrm{f}} \circ F_{\mathrm{f}}) = (G_{\mathrm{c}} \circ F_{\mathrm{c}})C.$$

Proof. The map identity follows directly by composition. The termwise statement follows from Theorem 4.6, because the Faà di Bruno expansion is a finite sum of rooted derivative trees at each derivative order. □

4.4 Exact quotient flows

Theorem 4.8 (Exact quotient flow)**.** *Let $U_{\mathrm{f}} \subseteq X_{\mathrm{f}}$ and $U_{\mathrm{c}} \subseteq X_{\mathrm{c}}$ be open, with $C(U_{\mathrm{f}}) \subseteq U_{\mathrm{c}}$. Suppose F_{f} and F_{c} are locally Lipschitz and satisfy*

$$CF_{\mathrm{f}} = F_{\mathrm{c}}C$$

on U_{f}. Let $\Phi^{\mathrm{f}}_t(x)$ and $\Phi^{\mathrm{c}}_t(z)$ denote the unique maximal local flows. Whenever both sides are defined,

$$\boxed{C\Phi^{\mathrm{f}}_t(x) = \Phi^{\mathrm{c}}_t(Cx).} \tag{37}$$

Proof. Let $x(t) = \Phi^{\mathrm{f}}_t(x_0)$ and set $z(t) = Cx(t)$. Then

$$\dot{z}(t) = C\dot{x}(t) = CF_{\mathrm{f}}(x(t)) = F_{\mathrm{c}}(Cx(t)) = F_{\mathrm{c}}(z(t)),$$

with $z(0) = Cx_0$. Uniqueness for the coarse initial value problem gives $z(t) = \Phi^{\mathrm{c}}_t(Cx_0)$. □

Corollary 4.9 (Linear semigroup naturality)**.** *If $CA_{\mathrm{f}} = A_{\mathrm{c}}C$, then*

$$Ce^{tA_{\mathrm{f}}} = e^{tA_{\mathrm{c}}}C \tag{38}$$

for every t.

Proof. Either apply Theorem 4.8 to the linear systems or sum the exponential series using $CA^k_{\mathrm{f}} = A^k_{\mathrm{c}}C$. □

Corollary 4.10 (Quotients of discrete dynamics)**.** *If $CF_{\mathrm{f}} = F_{\mathrm{c}}C$, then*

$$CF^{\circ n}_{\mathrm{f}} = F^{\circ n}_{\mathrm{c}}C$$

for every integer $n \geq 0$.

Proof. Induct on n. □

4.5 Positivity and simplex preservation

Let

$$\Delta_{\mathrm{f}} = \{\mu \in \mathbb{R}^n_+ : \mathbf{1}^T\mu = 1\}, \qquad \Delta_{\mathrm{c}} = \{\nu \in \mathbb{R}^q_+ : \mathbf{1}^T\nu = 1\}.$$

For block sum aggregation, $C\Delta_{\mathrm{f}} = \Delta_{\mathrm{c}}$ and

$$\mathbf{1}_{\mathrm{c}}^T C = \mathbf{1}_{\mathrm{f}}^T.$$

Proposition 4.11 (Simplex preservation descends)**.** *Suppose* $\Phi_t^{\mathrm{f}}(\Delta_{\mathrm{f}}) \subseteq \Delta_{\mathrm{f}}$ *and exact flow naturality holds. Then*

$$\Phi_t^{\mathrm{c}}(\Delta_{\mathrm{c}}) \subseteq \Delta_{\mathrm{c}}.$$

Proof. For any $\nu \in \Delta_{\mathrm{c}}$, choose $\mu \in \Delta_{\mathrm{f}}$ with $C\mu = \nu$. Then

$$\Phi_t^{\mathrm{c}}(\nu) = C\Phi_t^{\mathrm{f}}(\mu) \in C\Delta_{\mathrm{f}} = \Delta_{\mathrm{c}}.$$

□

Claim Boundary 4.12 (Global existence of the flow)**.** Exact descent does not guarantee global solutions. A fine trajectory may blow up in finite time, leave the domain, or reach a region where the vector field is no longer Lipschitz. Theorem 4.8 applies only while both solutions exist and remain unique.

5 Normalized positive transfer and Bayesian refinement

5.1 Positive state spaces

The Bayesian result rests on a simple fact about normalizing positive operators.

Let X_{f} and X_{c} be finite dimensional ordered vector spaces with positive cones K_{f} and K_{c}. Let

$$\ell_{\mathrm{f}} : X_{\mathrm{f}} \to \mathbb{R}, \qquad \ell_{\mathrm{c}} : X_{\mathrm{c}} \to \mathbb{R}$$

be positive mass functionals. Define normalized state spaces

$$\mathcal{S}_{\mathrm{f}} = \{x \in K_{\mathrm{f}} : \ell_{\mathrm{f}}(x) = 1\}, \qquad \mathcal{S}_{\mathrm{c}} = \{z \in K_{\mathrm{c}} : \ell_{\mathrm{c}}(z) = 1\}.$$

Assume C is positive and preserves mass:

$$\ell_{\mathrm{c}} C = \ell_{\mathrm{f}}. \tag{39}$$

Definition 5.1 (Normalized positive transfer)**.** For a positive linear operator $L : X \to X$, define

$$\mathsf{N}_\ell(L)(x) = \frac{Lx}{\ell(Lx)} \tag{40}$$

whenever $\ell(Lx) > 0$.

5.2 Normalization naturality

Theorem 5.2 (Naturality of normalized transfer)**.** *Let $L_{\mathrm{f}} : X_{\mathrm{f}} \to X_{\mathrm{f}}$ and $L_{\mathrm{c}} : X_{\mathrm{c}} \to X_{\mathrm{c}}$ be positive linear operators. Suppose*

$$CL_{\mathrm{f}} = sL_{\mathrm{c}}C \tag{41}$$

for some $s > 0$, and suppose (39) holds. Then, on the common domain of normalization,

$$\boxed{C\mathsf{N}_{\ell_{\mathrm{f}}}(L_{\mathrm{f}}) = \mathsf{N}_{\ell_{\mathrm{c}}}(L_{\mathrm{c}})C.} \tag{42}$$

Proof. For a positive state x,

$$\begin{aligned} C\mathsf{N}_{\ell_{\mathrm{f}}}(L_{\mathrm{f}})(x) &= \frac{CL_{\mathrm{f}}x}{\ell_{\mathrm{f}}(L_{\mathrm{f}}x)} \\ &= \frac{sL_{\mathrm{c}}Cx}{\ell_{\mathrm{c}}(CL_{\mathrm{f}}x)} \\ &= \frac{sL_{\mathrm{c}}Cx}{s\ell_{\mathrm{c}}(L_{\mathrm{c}}Cx)} \\ &= \mathsf{N}_{\ell_{\mathrm{c}}}(L_{\mathrm{c}})(Cx). \end{aligned}$$

□

Remark 5.3 (Gauge invariance). Normalization removes the scale s. At the operator level, this is the familiar fact that a positive rescaling of a likelihood leaves the posterior unchanged.

Proposition 5.4 (Composition law for normalized transfers)**.** *For positive operators L_1, L_2 on the same ordered state space,*

$$\mathsf{N}_{\ell}(L_2) \circ \mathsf{N}_{\ell}(L_1) = \mathsf{N}_{\ell}(L_2L_1) \tag{43}$$

wherever both sides are defined.

Proof. The first normalization contributes the same scalar to the numerator and denominator of the second update, so it cancels. □

5.3 Classical Bayesian update

On the finite simplex, let

$$L_\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \qquad \lambda_i > 0.$$

With $\ell(x) = \mathbf{1}^T x$,

$$\mathsf{N}_{\ell}(L_\lambda)(\mu) = \frac{\lambda \odot \mu}{\lambda^T \mu} = \mathsf{B}_\lambda(\mu).$$

The evidence is

$$Z_\lambda(\mu) = \lambda^T \mu.$$

The composition law becomes

$$\mathsf{B}_{\lambda_2} \circ \mathsf{B}_{\lambda_1} = \mathsf{B}_{\lambda_1 \odot \lambda_2}. \tag{44}$$

5.4 Exact Bayesian refinement

Let C be the block sum aggregation associated with a surjection $\pi : \{1, \dots, n\} \twoheadrightarrow \{1, \dots, q\}$.

Theorem 5.5 (Sufficient condition for Bayesian refinement)**.** *Let $\lambda^{\mathrm{f}} \in \mathbb{R}^n_{>0}$ and $\lambda^{\mathrm{c}} \in \mathbb{R}^q_{>0}$. If*

$$\lambda^{\mathrm{f}} = sC^T\lambda^{\mathrm{c}} \tag{45}$$

for some $s > 0$, then

$$\boxed{C\mathsf{B}_{\lambda^{\mathrm{f}}} = \mathsf{B}_{\lambda^{\mathrm{c}}}C.} \tag{46}$$

Moreover,

$$Z_{\lambda^{\mathrm{f}}}(\mu) = sZ_{\lambda^{\mathrm{c}}}(C\mu). \tag{47}$$

Proof. Equation (45) is equivalent to

$$CL_{\lambda^{\mathrm{f}}} = sL_{\lambda^{\mathrm{c}}}C.$$

Apply Theorem 5.2. The evidence identity follows from

$$(\lambda^{\mathrm{f}})^T\mu = s(\lambda^{\mathrm{c}})^TC\mu.$$

□

Condition (45) says that the fine likelihood is constant on each aggregation fiber, apart from one global rescaling. When the coarse space has at least two packets, the next theorem shows that this condition is also necessary if commutation is required for every interior prior.

Theorem 5.6 (Necessary condition for Bayesian refinement)**.** *Assume $q \geq 2$, every fiber of π is nonempty, and $\lambda^{\mathrm{f}}, \lambda^{\mathrm{c}}$ are strictly positive. If*

$$C\mathsf{B}_{\lambda^{\mathrm{f}}}(\mu) = \mathsf{B}_{\lambda^{\mathrm{c}}}(C\mu) \tag{48}$$

for every μ in the interior of Δ_n, then there exists $s > 0$ such that

$$\lambda^{\mathrm{f}} = sC^T\lambda^{\mathrm{c}}.$$

Proof. Both sides of (48) extend continuously to any prior with nonzero evidence. Strict positivity guarantees nonzero evidence throughout the closed simplex. Fix two distinct coarse packets a and b. Choose $i \in \pi^{-1}(a)$ and $k \in \pi^{-1}(b)$, then consider a prior supported on $\{i, k\}$ with masses $u, v > 0$. On the fine side, the ratio of the posterior masses of packets a and b is

$$\frac{\lambda^{\mathrm{f}}_i u}{\lambda^{\mathrm{f}}_k v},$$

whereas the coarse posterior ratio is

$$\frac{\lambda^{\mathrm{c}}_a u}{\lambda^{\mathrm{c}}_b v}.$$

Equality for all $u, v > 0$ gives

$$\frac{\lambda^{\mathrm{f}}_i}{\lambda^{\mathrm{f}}_k} = \frac{\lambda^{\mathrm{c}}_a}{\lambda^{\mathrm{c}}_b}. \tag{49}$$

Take another state j in the same fiber as i. Applying the same argument to j and k gives $\lambda^{\mathrm{f}}_j = \lambda^{\mathrm{f}}_i$. The fine likelihood is therefore constant on each fiber. Write its value on fiber a as $\overline{\lambda}_a$. Equation (49) now gives

$$\frac{\overline{\lambda}_a}{\overline{\lambda}_b} = \frac{\lambda^{\mathrm{c}}_a}{\lambda^{\mathrm{c}}_b}$$

for every pair a, b. Hence $\overline{\lambda} = s\lambda^{\mathrm{c}}$ for one $s > 0$, proving the result. □

Remark 5.7 (The case of one coarse packet). If $q = 1$, both sides of (48) equal the unique coarse state, regardless of the fine likelihood. Necessity therefore requires a nontrivial coarse state space.

5.5 Differential structure of Bayesian normalization

Proposition 5.8 (Bayesian differential). *At an interior prior μ,*

$$D\mathsf{B}_\lambda(\mu)[h] = \frac{\lambda \odot h}{Z_\lambda(\mu)} - \mathsf{B}_\lambda(\mu)\frac{\lambda^T h}{Z_\lambda(\mu)}. \tag{50}$$

Proof. Differentiate the quotient $L_\lambda\mu/(\mathbf{1}^T L_\lambda \mu)$. □

Corollary 5.9 (Bayesian jet naturality). *Under the hypotheses of Theorem 5.5, for every $r \geq 1$ and every interior fine prior,*

$$CD^r\mathsf{B}_{\lambda^{\mathrm{f}}}(\mu)[h_1, \ldots, h_r] = D^r\mathsf{B}_{\lambda^{\mathrm{c}}}(C\mu)[Ch_1, \ldots, Ch_r]. \tag{51}$$

Proof. Differentiate the exact map identity (46) and apply Theorem 4.4. □

Thus the normalization map has a descended interaction structure of its own. Its first derivative couples all evidence packets through a rank one term. Higher derivatives describe interactions among larger groups of packets.

6 Filtering cocycles

6.1 Finite observation sequences

Let $0 = t_0 < t_1 < \cdots < t_N$ and define prediction maps

$$\Phi^{\mathrm{f}}_k = \Phi^{\mathrm{f}}_{t_k - t_{k-1}}, \qquad \Phi^{\mathrm{c}}_k = \Phi^{\mathrm{c}}_{t_k - t_{k-1}}.$$

For observations $y_1, \ldots, y_N$, let

$$\mathsf{B}^{\mathrm{f}}_k = \mathsf{B}_{\lambda^{\mathrm{f}}_{y_k}}, \qquad \mathsf{B}^{\mathrm{c}}_k = \mathsf{B}_{\lambda^{\mathrm{c}}_{y_k}}.$$

Pair each prediction with the observation that follows it:

$$\Gamma^{\mathrm{f}}_k = \mathsf{B}^{\mathrm{f}}_k \circ \Phi^{\mathrm{f}}_k, \qquad \Gamma^{\mathrm{c}}_k = \mathsf{B}^{\mathrm{c}}_k \circ \Phi^{\mathrm{c}}_k,$$

and define the resulting filtering cocycles

$$\Psi^{\mathrm{f}}_{1:N} = \Gamma^{\mathrm{f}}_N \circ \cdots \circ \Gamma^{\mathrm{f}}_1, \qquad \Psi^{\mathrm{c}}_{1:N} = \Gamma^{\mathrm{c}}_N \circ \cdots \circ \Gamma^{\mathrm{c}}_1. \tag{52}$$

Theorem 6.1 (Exact naturality of the filtering cocycle). *Suppose, for every k,*

$$C\Phi^{\mathrm{f}}_k = \Phi^{\mathrm{c}}_k C \tag{53}$$

and

$$C\mathsf{B}^{\mathrm{f}}_k = \mathsf{B}^{\mathrm{c}}_k C. \tag{54}$$

Then

$$\boxed{C\Psi^{\mathrm{f}}_{1:N} = \Psi^{\mathrm{c}}_{1:N} C.} \tag{55}$$

Proof. For one step,

$$C\Gamma_k^{\mathrm{f}} = C\mathsf{B}_k^{\mathrm{f}}\Phi_k^{\mathrm{f}} = \mathsf{B}_k^{\mathrm{c}}C\Phi_k^{\mathrm{f}} = \mathsf{B}_k^{\mathrm{c}}\Phi_k^{\mathrm{c}}C = \Gamma_k^{\mathrm{c}}C.$$

Iterating this identity through (52) proves the result. □

Corollary 6.2 (Exact posterior aggregation)**.** *If $\mu_0^{\mathrm{c}} = C\mu_0^{\mathrm{f}}$, then*

$$\boxed{C\mu_N^{\mathrm{f}} = \mu_N^{\mathrm{c}}}$$

for every finite observation sequence, where

$$\mu_N^{\mathrm{f}} = \Psi_{1:N}^{\mathrm{f}}(\mu_0^{\mathrm{f}}), \qquad \mu_N^{\mathrm{c}} = \Psi_{1:N}^{\mathrm{c}}(\mu_0^{\mathrm{c}}).$$

Proof. Apply (55) to μ_0^{f} and substitute $\mu_0^{\mathrm{c}} = C\mu_0^{\mathrm{f}}$. □

6.2 Sequential evidence

Let

$$\widehat{\mu}_k^{\mathrm{f}} = \Phi_k^{\mathrm{f}}(\mu_{k-1}^{\mathrm{f}}), \qquad \widehat{\mu}_k^{\mathrm{c}} = \Phi_k^{\mathrm{c}}(\mu_{k-1}^{\mathrm{c}}),$$

and define the step evidence

$$z_k^{\mathrm{f}} = (\lambda_{y_k}^{\mathrm{f}})^T\widehat{\mu}_k^{\mathrm{f}}, \qquad z_k^{\mathrm{c}} = (\lambda_{y_k}^{\mathrm{c}})^T\widehat{\mu}_k^{\mathrm{c}}.$$

Set

$$Z_{1:N}^{\mathrm{f}} = \prod_{k=1}^{N} z_k^{\mathrm{f}}, \qquad Z_{1:N}^{\mathrm{c}} = \prod_{k=1}^{N} z_k^{\mathrm{c}}.$$

Theorem 6.3 (Evidence factorization under exact refinement)**.** *Assume the hypotheses of Theorem 6.1, let*

$$\mu_0^{\mathrm{c}} = C\mu_0^{\mathrm{f}},$$

and suppose

$$\lambda_{y_k}^{\mathrm{f}} = s_k C^T\lambda_{y_k}^{\mathrm{c}} \qquad (s_k > 0).$$

Then

$$z_k^{\mathrm{f}} = s_k z_k^{\mathrm{c}} \tag{56}$$

for every step, and the likelihood of the observation sequence satisfies

$$\boxed{Z_{1:N}^{\mathrm{f}} = \left(\prod_{k=1}^{N} s_k\right) Z_{1:N}^{\mathrm{c}}.} \tag{57}$$

Under the gauge $s_k = 1$, the fine and coarse sequential evidences agree exactly.

Proof. Initial compatibility gives $C\mu_0^{\mathrm{f}} = \mu_0^{\mathrm{c}}$. For later steps, Theorem 6.2 gives the same identity after every update. Flow naturality therefore yields

$$C\widehat{\mu}_k^{\mathrm{f}} = \widehat{\mu}_k^{\mathrm{c}}.$$

Therefore

$$z_k^{\mathrm{f}} = s_k(\lambda_{y_k}^{\mathrm{c}})^T C\widehat{\mu}_k^{\mathrm{f}} = s_k z_k^{\mathrm{c}}.$$

The likelihood of the complete observation sequence is the product of the conditional evidence terms. Multiplying the identities gives (57). □

6.3 Hidden Markov and nonlinear specializations

Corollary 6.4 (Hidden Markov lumpability)**.** *Let M_{f} and M_{c} be column stochastic matrices satisfying*

$$CM_{\mathrm{f}} = M_{\mathrm{c}}C,$$

and let every observation likelihood satisfy

$$\lambda^{\mathrm{f}}_y = s_y C^T \lambda^{\mathrm{c}}_y.$$

Then, for every observation sequence $y_{1:N}$,

$$\begin{aligned} &C(\mathsf{B}_{\lambda^{\mathrm{f}}_{y_N}} \circ M_{\mathrm{f}} \circ \cdots \circ \mathsf{B}_{\lambda^{\mathrm{f}}_{y_1}} \circ M_{\mathrm{f}}) \\ &\quad = (\mathsf{B}_{\lambda^{\mathrm{c}}_{y_N}} \circ M_{\mathrm{c}} \circ \cdots \circ \mathsf{B}_{\lambda^{\mathrm{c}}_{y_1}} \circ M_{\mathrm{c}})C. \end{aligned}$$

Proof. Apply Theorem 6.1 with the linear prediction maps M_{f} and M_{c}. □

Corollary 6.5 (Nonlinear quotient filtering)**.** *Suppose the fine and coarse prediction maps are flow segments of locally Lipschitz vector fields satisfying $CF_{\mathrm{f}} = F_{\mathrm{c}}C$, and suppose all observation likelihoods are constant on each fiber up to scale. Then the nonlinear filtering cocycle commutes with C exactly.*

Proof. Intertwining of the vector fields gives $C\Phi^{\mathrm{f}}_k = \Phi^{\mathrm{c}}_k C$ by Theorem 4.8. Likelihoods that are constant on each fiber give $C\mathsf{B}^{\mathrm{f}}_k = \mathsf{B}^{\mathrm{c}}_k C$ by Theorem 5.5. Apply Theorem 6.1. □

Remark 6.6 (Two independent closure requirements). The dynamical and observational conditions are independent. Exact interaction lumpability says nothing about whether the likelihood distinguishes states inside a packet. Conversely, an observation may ignore those distinctions even when the dynamics depend on them. Exact coarse filtering needs both commutative squares.

Claim Boundary 6.7 (Sufficiency for the chosen observation family)**.** The condition $\lambda^{\mathrm{f}}_y = s_y C^T \lambda^{\mathrm{c}}_y$ is sufficient for the specified finite likelihood family. It does not assert that C is sufficient for every possible future observation model. Adding a likelihood that distinguishes two states inside one fiber can invalidate the coarse filter without changing the dynamics.

7 Approximate interaction quotients

Exact quotienting is an algebraic yes or no question. A useful model also needs to say how far it misses when the answer is no. We separate three sources of error:

interaction mismatch, dependence on hidden directions, dynamical amplification.

7.1 Best coarse interaction and quotient defect

Let

$$\mathfrak{Q}_m(C) = \{B_m C^{\times m} : B_m \in \mathcal{L}^m(X_{\mathrm{c}}; X_{\mathrm{c}})\}$$

be the subspace of outputs from multilinear maps of order m that factor through the coarse space.

Definition 7.1 (Best quotient error)**.** For a fine interaction A^{f}_m, define

$$\varepsilon^\star_m = \inf_{B_m \in \mathcal{L}^m(X_{\mathrm{c}}; X_{\mathrm{c}})} \left\| CA^{\mathrm{f}}_m - B_m C^{\times m} \right\|. \tag{58}$$

Proposition 7.2 (Attainment in finite dimension)**.** *The infimum in* (58) *is attained.*

Proof. The map

$$B_m \longmapsto B_m C^{\times m}$$

is injective because C is surjective. Its range is a finite dimensional linear subspace of $\mathcal{L}^m(X_{\mathrm{f}}; X_{\mathrm{c}})$, hence it is closed. In a finite dimensional normed space, the distance to a nonempty closed subspace attains its minimum. A nearest point in this range therefore exists. Any such point has a unique preimage under the displayed map, and that preimage is a minimizing coarse tensor. □

The minimum value depends on the chosen norm. A minimizing tensor need not be unique when that norm is not strictly convex. Exact lumpability is independent of the norm because it is the condition $\varepsilon_m^\star = 0$.

7.2 Approximation induced by a section

A chosen right inverse supplies a concrete coarse tensor even when exact descent fails.

Definition 7.3 (Coarse tensor induced by a section)**.** Fix E with $CE = I$. Define

$$A_m^{\mathrm{c},E} = C A_m^{\mathrm{f}} E^{\times m}. \tag{59}$$

Let $P_C = EC$ and $Q_C = I - EC$ as in (10).

Proposition 7.4 (Expansion of the section defect by hidden inputs)**.** *The defect of* (59) *is*

$$\Delta_m^E(x_1, \dots, x_m) = C A_m^{\mathrm{f}}(x_1, \dots, x_m) - C A_m^{\mathrm{f}}(P_C x_1, \dots, P_C x_m), \tag{60}$$

and admits the exact expansion

$$\boxed{\Delta_m^E(x_1, \dots, x_m) = \sum_{\varnothing \neq S \subseteq \{1,\dots,m\}} C A_m^{\mathrm{f}}(y_1^S, \dots, y_m^S),} \tag{61}$$

where

$$y_r^S = \begin{cases} Q_C x_r, & r \in S, \\ P_C x_r, & r \notin S. \end{cases}$$

Proof. Equation (60) follows from

$$A_m^{\mathrm{c},E}(Cx_1, \dots, Cx_m) = C A_m^{\mathrm{f}}(ECx_1, \dots, ECx_m).$$

Now expand every input as $x_r = P_C x_r + Q_C x_r$. The term with no Q_C factor is the term subtracted in (60). Every remaining term has at least one hidden input and appears exactly once. □

Corollary 7.5 (Exactness as disappearance of hidden interactions)**.** *The tensor defined from the chosen section descends exactly if and only if the sum in* (61) *vanishes for all inputs. The stronger slotwise condition*

$$C A_m^{\mathrm{f}}(x_1, \dots, Q_C x_r, \dots, x_m) = 0$$

for each slot is equivalent to exact descent and makes every summand vanish separately.

Proof. The first statement is (61) with $\Delta_m^E = 0$. Since $\operatorname{Ran} Q_C = \operatorname{Ker} C$, the condition in each slot is exactly the kernel condition in Theorem 3.1. It is therefore equivalent to exact descent. Every term indexed by a nonempty subset contains at least one hidden input, so each term vanishes. □

7.3 Defect majorants for the full field

Suppose fine and proposed coarse fields have expansions

$$F_{\mathrm{f}}(x) = \sum_{m\geq 0} A^{\mathrm{f}}_m[x^{\otimes m}], \qquad F_{\mathrm{c}}(z) = \sum_{m\geq 0} A^{\mathrm{c}}_m[z^{\otimes m}].$$

Define the field defect

$$\delta_F(x) = CF_{\mathrm{f}}(x) - F_{\mathrm{c}}(Cx). \tag{62}$$

Whenever both series converge absolutely,

$$\delta_F(x) = \sum_{m\geq 0} \Delta_m[x^{\otimes m}]. \tag{63}$$

Proposition 7.6 (Analytic defect majorant)**.** *If $\|\Delta_m\| \leq \varepsilon_m$ and the scalar series*

$$\mathfrak{d}_F(R) = \sum_{m\geq 0} \varepsilon_m R^m$$

converges, then

$$\boxed{\sup_{\|x\|\leq R} \|\delta_F(x)\| \leq \mathfrak{d}_F(R).} \tag{64}$$

Proof. Apply the multilinear norm bound to every term in (63) and sum the convergent majorant. □

The sequence (ε_m) shows which interaction orders obstruct the quotient. A model may be exactly lumpable through order r and fail for the first time at order $r+1$.

7.4 Propagation of dynamical defect

Theorem 7.7 (Approximate quotient flow bound)**.** *Let*

$$\dot{x} = F_{\mathrm{f}}(x), \qquad \dot{z} = F_{\mathrm{c}}(z),$$

and suppose both solutions remain in domains on which

$$\|F_{\mathrm{c}}(u) - F_{\mathrm{c}}(v)\| \leq L\,\|u - v\|\,.$$

Set

$$e(t) = \|Cx(t) - z(t)\|\,.$$

Then

$$\boxed{e(t) \leq e^{Lt}e(0) + \int_0^t e^{L(t-s)}\,\|\delta_F(x(s))\|\,\mathrm{d}s.} \tag{65}$$

If $e(0) = 0$ and $\|\delta_F(x(s))\| \leq \varepsilon_F$ on $[0,t]$, then

$$e(t) \leq \begin{cases} \varepsilon_F \dfrac{e^{Lt}-1}{L}, & L > 0, \\ t\varepsilon_F, & L = 0. \end{cases} \tag{66}$$

Proof. Let $y(t) = Cx(t)$. From (62),

$$\dot{y}(t) = F_{\mathrm{c}}(y(t)) + \delta_F(x(t)).$$

Subtract the integral equations for y and z:

$$y(t) - z(t) = y(0) - z(0) + \int_0^t \left(F_{\mathrm{c}}(y(s)) - F_{\mathrm{c}}(z(s))\right) \mathrm{d}s + \int_0^t \delta_F(x(s))\, \mathrm{d}s.$$

Taking norms and applying the Grönwall inequality yields (65). The uniform cases follow by evaluating the exponential integral. □

Corollary 7.8 (Coefficient certificate for a flow segment)**.** *If $\|x(s)\| \le R$ and Proposition 7.6 applies, then a flow interval of length T satisfies*

$$\left\| C\Phi_T^{\mathrm{f}}(x_0) - \Phi_T^{\mathrm{c}}(Cx_0) \right\| \le \mathfrak{d}_F(R) \frac{e^{LT} - 1}{L}$$

for $L > 0$, with $T\mathfrak{d}_F(R)$ at $L = 0$.

Proof. By Proposition 7.6, $\|\delta_F(x(s))\| \le \mathfrak{d}_F(R)$ throughout the segment. Apply Theorem 7.7 with compatible initial data, so $e(0) = 0$, and evaluate the exponential integral. □

7.5 Refinement towers

Let

$$X_L \xrightarrow{C_{L,L-1}} X_{L-1} \longrightarrow \cdots \longrightarrow X_1 \xrightarrow{C_{10}} X_0.$$

For $a > b$, write

$$C_{ab} = C_{b+1,b} \cdots C_{a,a-1}.$$

Let Δ_m^{ab} denote the interaction defect for the composite map C_{ab}. Iterating Theorem 3.11 yields a telescoping formula.

Corollary 7.9 (Telescoping interaction defect)**.** *For an interaction of order m assigned at every level,*

$$\Delta_m^{L0} = \sum_{\ell=0}^{L-1} C_{\ell 0}\, \Delta_m^{\ell+1,\ell} C_{L,\ell+1}^{\times m}, \tag{67}$$

where $C_{00} = I$ and $C_{LL} = I$. Consequently,

$$\left\| \Delta_m^{L0} \right\| \le \sum_{\ell=0}^{L-1} \|C_{\ell 0}\| \left\| \Delta_m^{\ell+1,\ell} \right\| \|C_{L,\ell+1}\|^m. \tag{68}$$

Proof. Induct on the number of levels, using (27) at the final composition. □

The identity shows where a terminal mismatch entered the refinement tower. Rather than storing only the final error, one can keep the local defect at each level and combine those records when needed.

8 Approximate Bayesian refinement

8.1 A normalization perturbation bound

Lemma 8.1 (Normalization perturbation)**.** *Let $u, v \in \mathbb{R}^q_+$ have positive total mass. Then*

$$\boxed{\left\| \frac{u}{\mathbf{1}^T u} - \frac{v}{\mathbf{1}^T v} \right\|_1 \leq \frac{2\,\|u - v\|_1}{\max\{\mathbf{1}^T u, \mathbf{1}^T v\}}.} \tag{69}$$

Proof. Write $a = \mathbf{1}^T u$ and $b = \mathbf{1}^T v$. Then

$$\begin{aligned} \left\| \frac{u}{a} - \frac{v}{b} \right\|_1 &\leq \frac{\|u - v\|_1}{a} + \|v\|_1 \left| \frac{1}{a} - \frac{1}{b} \right| \\ &= \frac{\|u - v\|_1}{a} + \frac{|a - b|}{a} \\ &\leq \frac{2\,\|u - v\|_1}{a}. \end{aligned}$$

Interchanging u and v gives the same inequality with b in the denominator. Taking the smaller of the two upper bounds yields the denominator $\max\{a, b\}$. □

The factor of two need not be sharp. It is uniform, and it makes the role of the evidence mass explicit. No uniform estimate of this form remains useful when both masses approach zero.

8.2 Defect in the observation intertwining law

Let L_{f} and L_{c} be positive linear observation operators, and choose a scale $s > 0$. On a finite simplex, we write $\mathsf{N}(L)$ for normalization with respect to the total mass $\mathbf{1}^T$. Define

$$R_L = C L_{\mathrm{f}} - s L_{\mathrm{c}} C. \tag{70}$$

Theorem 8.2 (Approximate normalized transfer naturality)**.** *On finite simplices with block sum C, let*

$$u = C L_{\mathrm{f}} \mu, \qquad v = s L_{\mathrm{c}} C \mu.$$

If u and v are nonnegative with positive masses, then

$$\boxed{\|C\mathsf{N}(L_{\mathrm{f}})(\mu) - \mathsf{N}(L_{\mathrm{c}})(C\mu)\|_1 \leq \frac{2\,\|R_L \mu\|_1}{\max\{\mathbf{1}^T u, \mathbf{1}^T v\}}.} \tag{71}$$

Proof. Because C preserves total mass,

$$C\mathsf{N}(L_{\mathrm{f}})(\mu) = \frac{u}{\mathbf{1}^T u}.$$

Multiplication of L_{c} by s does not change its normalized output, so

$$\mathsf{N}(L_{\mathrm{c}})(C\mu) = \frac{v}{\mathbf{1}^T v}.$$

Now $u - v = R_L \mu$, and Lemma 8.1 applies. □

8.3 Likelihood variation inside a fiber

For diagonal likelihood operators, write

$$\lambda^{\mathrm{f}} = sC^T\lambda^{\mathrm{c}} + r_\lambda. \tag{72}$$

Then

$$R_L\mu = C(r_\lambda \odot \mu).$$

Since block sum aggregation is an ℓ^1 contraction on signed vectors,

$$\|R_L\mu\|_1 \le \sum_i \mu_i\,|r_{\lambda,i}| \le \|r_\lambda\|_\infty. \tag{73}$$

Corollary 8.3 (Bayesian refinement defect)**.** *Let*

$$z_{\mathrm{f}} = Z_{\lambda^{\mathrm{f}}}(\mu), \qquad z_{\mathrm{c}} = sZ_{\lambda^{\mathrm{c}}}(C\mu).$$

Then

$$\boxed{\|C\mathsf{B}_{\lambda^{\mathrm{f}}}(\mu) - \mathsf{B}_{\lambda^{\mathrm{c}}}(C\mu)\|_1 \le \frac{2\sum_i \mu_i\,|r_{\lambda,i}|}{\max\{z_{\mathrm{f}}, z_{\mathrm{c}}\}}.} \tag{74}$$

If both evidences are at least $z_* > 0$, then

$$d_{\mathsf{B}} \le \frac{2\,\|r_\lambda\|_\infty}{z_*}. \tag{75}$$

Proof. Combine Theorem 8.2 with (73). □

Definition 8.4 (Fiber oscillation)**.** For a fine likelihood, define

$$\mathrm{osc}_a(\lambda^{\mathrm{f}}) = \max_{i:\pi(i)=a} \lambda_i^{\mathrm{f}} - \min_{i:\pi(i)=a} \lambda_i^{\mathrm{f}},$$

and

$$\mathrm{osc}_C(\lambda^{\mathrm{f}}) = \max_a \mathrm{osc}_a(\lambda^{\mathrm{f}}).$$

When $q \ge 2$ and commutation is required for every interior prior, exact Bayesian refinement is equivalent to $\mathrm{osc}_C(\lambda^{\mathrm{f}}) = 0$ once the coarse likelihood and global scale have been chosen consistently. In an approximate model, λ_a^{c} may be the midpoint, the ordinary mean, or a mean chosen to reflect the task on fiber a. The midpoint minimizes the largest absolute residual in that fiber and gives

$$\|r_\lambda\|_\infty = \frac{1}{2}\,\mathrm{osc}_C(\lambda^{\mathrm{f}})$$

when $s = 1$.

8.4 Lipschitz behavior of the Bayesian update

Proposition 8.5 (Lipschitz bound under an evidence floor)**.** *Let $\lambda \in \mathbb{R}^q_{>0}$. If*

$$Z_\lambda(\mu), Z_\lambda(\nu) \ge z_* > 0,$$

then

$$\boxed{\|\mathsf{B}_\lambda(\mu) - \mathsf{B}_\lambda(\nu)\|_1 \le \frac{2\,\|\lambda\|_\infty}{z_*}\,\|\mu - \nu\|_1.} \tag{76}$$

Proof. Apply Lemma 8.1 to

$$u = \lambda \odot \mu, \qquad v = \lambda \odot \nu.$$

Then

$$\|u - v\|_1 \le \|\lambda\|_\infty \|\mu - \nu\|_1 .$$

□

Remark 8.6 (Scale invariance of the bound). Replacing λ by $s\lambda$ multiplies both $\|\lambda\|_\infty$ and z_* by s, leaving the ratio unchanged.

Remark 8.7 (Local versus global constants). The constant in (76) is valid over the whole evidence region under consideration. Along a particular trajectory, the Jacobian in (50) may give a much smaller local factor. The recursive theorem below can use either bound, provided the chosen constant is certified on the states being considered.

9 Certified approximate coarse filtering

9.1 Prediction and update errors

Let the fine recursion be

$$\widehat{\mu}_k^{\mathrm{f}} = \Phi_k^{\mathrm{f}}(\mu_{k-1}^{\mathrm{f}}), \qquad \mu_k^{\mathrm{f}} = \mathsf{B}_k^{\mathrm{f}}(\widehat{\mu}_k^{\mathrm{f}}), \tag{77}$$

and the coarse recursion be

$$\widehat{\mu}_k^{\mathrm{c}} = \Phi_k^{\mathrm{c}}(\mu_{k-1}^{\mathrm{c}}), \qquad \mu_k^{\mathrm{c}} = \mathsf{B}_k^{\mathrm{c}}(\widehat{\mu}_k^{\mathrm{c}}). \tag{78}$$

Define the discrepancies before and after each update by

$$e_k^- = \left\| C\widehat{\mu}_k^{\mathrm{f}} - \widehat{\mu}_k^{\mathrm{c}} \right\|_1 , \qquad e_k^+ = \left\| C\mu_k^{\mathrm{f}} - \mu_k^{\mathrm{c}} \right\|_1 . \tag{79}$$

At the initial time, set

$$e_0^+ = \left\| C\mu_0^{\mathrm{f}} - \mu_0^{\mathrm{c}} \right\|_1 .$$

Assume the coarse prediction is Lipschitz with constant $L_{\Phi,k}$ and that its quotient defect is bounded by

$$\sup_{\mu} \left\| C\Phi_k^{\mathrm{f}}(\mu) - \Phi_k^{\mathrm{c}}(C\mu) \right\|_1 \le d_{\Phi,k}. \tag{80}$$

Assume the coarse update is Lipschitz with constant $L_{\mathsf{B},k}$ and that its direct refinement defect is bounded by

$$\sup_{\mu} \left\| C\mathsf{B}_k^{\mathrm{f}}(\mu) - \mathsf{B}_k^{\mathrm{c}}(C\mu) \right\|_1 \le d_{\mathsf{B},k}. \tag{81}$$

The suprema in (80) and (81) are taken over a declared admissible set that contains both recursions.

Lemma 9.1 (Prediction bound for one step)**.** *Under* (80)*,*

$$e_k^- \le L_{\Phi,k} e_{k-1}^+ + d_{\Phi,k}. \tag{82}$$

Proof. Insert and subtract $\Phi_k^{\mathrm{c}}(C\mu_{k-1}^{\mathrm{f}})$. The quotient defect bounds one difference, while the Lipschitz estimate bounds the other by $L_{\Phi,k} e_{k-1}^+$. □

Lemma 9.2 (Update bound for one step)**.** *Under* (81)*,*

$$e_k^+ \le L_{\mathsf{B},k} e_k^- + d_{\mathsf{B},k}. \tag{83}$$

Proof. Insert and subtract $\mathsf{B}_k^{\mathrm{c}}(C\widehat{\mu}_k^{\mathrm{f}})$. The direct refinement defect bounds one difference, while the Lipschitz estimate bounds the other by $L_{\mathsf{B},k}e_k^-$. □

Combining the two inequalities gives

$$e_k^+ \le \alpha_k e_{k-1}^+ + \beta_k, \tag{84}$$

where

$$\alpha_k = L_{\mathsf{B},k}L_{\Phi,k}, \qquad \beta_k = L_{\mathsf{B},k}d_{\Phi,k} + d_{\mathsf{B},k}. \tag{85}$$

9.2 Closed posterior certificate

Theorem 9.3 (Certified approximate filtering)**.** *Under the assumptions above,*

$$\boxed{e_N^+ \le \left(\prod_{j=1}^{N}\alpha_j\right)e_0^+ + \sum_{k=1}^{N}\left(\prod_{j=k+1}^{N}\alpha_j\right)\beta_k.} \tag{86}$$

Here and below, an empty product equals one. With exact initial aggregation $e_0^+ = 0$,

$$e_N^+ \le \sum_{k=1}^{N}\left(\prod_{j=k+1}^{N}\alpha_j\right)(L_{\mathsf{B},k}d_{\Phi,k} + d_{\mathsf{B},k}). \tag{87}$$

Proof. Repeated substitution in the affine recurrence (84) gives the stated formula. The initial error is multiplied by every later amplification factor. A defect introduced at step k is multiplied only by the factors from steps $k+1$ through N. □

Corollary 9.4 (Uniform stability regimes)**.** *Suppose $0 \le \alpha_k \le \alpha$ and $0 \le \beta_k \le \beta$.*

(i) If $0 \le \alpha < 1$, then

$$e_N^+ \le \alpha^N e_0^+ + \beta\frac{1-\alpha^N}{1-\alpha}, \qquad \limsup_{N\to\infty} e_N^+ \le \frac{\beta}{1-\alpha}. \tag{88}$$

(ii) If $\alpha = 1$, then

$$e_N^+ \le e_0^+ + N\beta.$$

(iii) If $\alpha > 1$, then

$$e_N^+ \le \alpha^N e_0^+ + \beta\frac{\alpha^N - 1}{\alpha - 1}.$$

Proof. Bound the products in (86) by powers of α and sum the geometric series. □

This trichotomy describes the certificate, not necessarily the observed filter. A stable trajectory can still come with a bound for which $\alpha > 1$. In that case, the chosen global norm estimates simply fail to certify contraction.

9.3 From interaction coefficients to posterior error

For a flow interval of length T_k, suppose the fine trajectory remains in a ball of radius R_k and the coarse field has Lipschitz constant $L_{F,k}$. By Corollary 7.8, one may take

$$d_{\Phi,k} = \mathfrak{d}_{F,k}(R_k)\frac{e^{L_{F,k}T_k}-1}{L_{F,k}}, \tag{89}$$

with the continuous value $T_k\mathfrak{d}_{F,k}(R_k)$ at $L_{F,k}=0$. For the observation, Corollary 8.3 gives

$$d_{\mathsf{B},k} \le \frac{2\,\|r_{\lambda,k}\|_\infty}{z_{*,k}}. \tag{90}$$

Substituting these bounds into (87) gives a certificate expressed directly in terms of interaction tensor defects and likelihood residuals within each fiber.

9.4 Observable and total variation consequences

Corollary 9.5 (Coarse observable error)**.** *For any coarse observable $g \in \mathbb{R}^q$,*

$$\left|g^T C\mu_N^{\mathrm{f}} - g^T\mu_N^{\mathrm{c}}\right| \le \|g\|_\infty\, e_N^+.$$

In particular,

$$\mathrm{TV}(C\mu_N^{\mathrm{f}},\mu_N^{\mathrm{c}}) = \frac{1}{2}e_N^+.$$

Proof. Apply the duality between ℓ^1 and ℓ^∞. The total variation identity is standard for finite probability vectors. □

9.5 Sequential evidence error

For brevity, write $\lambda_k^{\mathrm{f}} = \lambda_{y_k}^{\mathrm{f}}$ and $\lambda_k^{\mathrm{c}} = \lambda_{y_k}^{\mathrm{c}}$, and decompose

$$\lambda_k^{\mathrm{f}} = s_k C^T\lambda_k^{\mathrm{c}} + r_k. \tag{91}$$

This gives

$$z_k^{\mathrm{f}} - s_k z_k^{\mathrm{c}} = s_k(\lambda_k^{\mathrm{c}})^T(C\widehat{\mu}_k^{\mathrm{f}} - \widehat{\mu}_k^{\mathrm{c}}) + r_k^T\widehat{\mu}_k^{\mathrm{f}}. \tag{92}$$

Proposition 9.6 (Evidence certificate for one step)**.** *Assume $\widehat{\mu}_k^{\mathrm{f}}$ is a probability vector. Under (91),*

$$\boxed{\left|z_k^{\mathrm{f}} - s_k z_k^{\mathrm{c}}\right| \le s_k\,\|\lambda_k^{\mathrm{c}}\|_\infty\, e_k^- + \|r_k\|_\infty\,.} \tag{93}$$

If $z_k^{\mathrm{f}}, s_k z_k^{\mathrm{c}} \ge z_{*,k} > 0$, then

$$\boxed{\left|\log z_k^{\mathrm{f}} - \log(s_k z_k^{\mathrm{c}})\right| \le \frac{s_k\,\|\lambda_k^{\mathrm{c}}\|_\infty\, e_k^- + \|r_k\|_\infty}{z_{*,k}}.} \tag{94}$$

Proof. Apply Hölder's inequality to (92). Since $\widehat{\mu}_k^{\mathrm{f}}$ is a probability vector, $\left\|\widehat{\mu}_k^{\mathrm{f}}\right\|_1 = 1$. For the logarithms, the mean value theorem gives

$$|\log a - \log b| \le \frac{|a-b|}{\min\{a,b\}}.$$

□

Corollary 9.7 (Certificate for cumulative log evidence)**.** *Let* $Z^{\mathrm{f}}_{1:N} = \prod_k z^{\mathrm{f}}_k$ *and* $Z^{\mathrm{c}}_{1:N} = \prod_k z^{\mathrm{c}}_k$*. Then*

$$\boxed{\left|\log Z^{\mathrm{f}}_{1:N} - \sum_{k=1}^{N} \log s_k - \log Z^{\mathrm{c}}_{1:N}\right| \le \sum_{k=1}^{N} \frac{s_k \left\|\lambda^{\mathrm{c}}_k\right\|_\infty e^-_k + \|r_k\|_\infty}{z_{*,k}}.} \tag{95}$$

Proof. Sum (94) over the observation sequence. □

10 Hyperedge pruning and certified compression

Exact quotienting and sparse representation are distinct. A system may have an exact coarse quotient while its fine interaction tensor is dense. Conversely, a sparse fine tensor can fail to descend. This section introduces a separate residual for deleted hyperedges and then combines it with the quotient defect.

10.1 Hilbert Schmidt tensor blocks

Let X and Y be finite dimensional Hilbert spaces with orthogonal packet projectors (P_j) and (Q_i). Identify a multilinear Hilbert Schmidt tensor of order m with the Hilbert Schmidt linear map

$$\widehat{A} : X^{\otimes m} \to Y.$$

Its packet block is

$$\widehat{A}_{i|j_1,\dots,j_m} = Q_i \widehat{A}(P_{j_1} \otimes \cdots \otimes P_{j_m}).$$

Theorem 10.1 (Orthogonal hyperedge Parseval identity)**.** *Distinct packet blocks are orthogonal in Hilbert Schmidt inner product, and*

$$\boxed{\left\|\widehat{A}\right\|^2_{\mathrm{HS}} = \sum_{i,j_1,\dots,j_m} \left\|\widehat{A}_{i|j_1,\dots,j_m}\right\|^2_{\mathrm{HS}}.} \tag{96}$$

Proof. The tensor product input projectors

$$P_{j_1} \otimes \cdots \otimes P_{j_m}$$

are mutually orthogonal, as are the output projectors Q_i. For two distinct block labels, orthogonality on at least one side forces the Hilbert Schmidt inner product to vanish. Exact reconstruction is therefore an orthogonal sum, and Pythagoras gives (96). □

10.2 Retained field and named residual

Choose a retained hyperedge set $\mathcal{E}^{\mathrm{keep}}_m$. Define

$$\widehat{A}^{\mathrm{keep}}_m = \sum_{h \in \mathcal{E}^{\mathrm{keep}}_m} \widehat{A}_{m,h}, \qquad R_m = \widehat{A}_m - \widehat{A}^{\mathrm{keep}}_m. \tag{97}$$

Corollary 10.2 (Exact pruning residual)**.** *Under the hypotheses of Theorem 10.1,*

$$\boxed{\|R_m\|^2_{\mathrm{HS}} = \sum_{h \notin \mathcal{E}^{\mathrm{keep}}_m} \left\|\widehat{A}_{m,h}\right\|^2_{\mathrm{HS}}.} \tag{98}$$

Proof. The retained and omitted block families are orthogonal by Theorem 10.1. Since R_m is exactly the sum of the omitted blocks, Pythagoras gives (98). □

The residual is a tensor identity, not merely an empirical estimate. Its induced state error is obtained by applying the relevant multilinear norm inequality.

10.3 Retaining the k largest energy hyperedges

Let

$$e_h = \left\| \widehat{A}_{m,h} \right\|_{\mathrm{HS}}^2$$

and order the energies

$$e_{h_1} \geq e_{h_2} \geq \cdots \geq e_{h_M}.$$

Theorem 10.3 (Optimal retention of k orthogonal hyperedges)**.** *Among all retained sets of cardinality k, the set*

$$\{h_1, \ldots, h_k\}$$

minimizes the Hilbert Schmidt residual. The minimum is

$$\boxed{\min_{|\mathcal{E}|=k} \|R_m(\mathcal{E})\|_{\mathrm{HS}}^2 = \sum_{r>k} e_{h_r}.} \tag{99}$$

Proof. By (98), minimizing the omitted energy is equivalent to maximizing the retained energy. For a fixed cardinality, this is achieved by selecting the k largest nonnegative terms. □

Claim Boundary 10.4 (Overlapping blocks and structured costs)**.** The theorem above uses orthogonal block energy and assigns the same unit cost to every retained hyperedge. With overlapping dictionaries, shared low rank factors, communication costs, or group constraints that preserve structure, ranking individual energies need not be optimal.

10.4 Hyperedge entropy

For nonzero total interaction energy, define

$$p_h = \frac{e_h}{\sum_\eta e_\eta}, \qquad H_m = -\sum_h p_h \log p_h,$$

and the effective hyperedge count

$$N_{\mathrm{eff},m} = e^{H_m}.$$

Low entropy means that the interaction energy is concentrated in a small set of resolved hyperedges. High entropy means that pruning by a fixed cutoff is less promising in the chosen resolution. The quantity deliberately depends on the resolution. It evaluates a representation rather than an invariant of the unresolved tensor alone.

10.5 Combined quotient and pruning defect

Let the full fine coefficient split as

$$A_m^{\mathrm{f}} = A_m^{\mathrm{keep}} + R_m^{\mathrm{f}}.$$

For a proposed coarse coefficient A_m^{c}, define

$$\Delta_m^{\mathrm{total}} = CA_m^{\mathrm{f}} - A_m^{\mathrm{c}} C^{\times m}.$$

Then

$$\Delta_m^{\mathrm{total}} = \underbrace{CA_m^{\mathrm{keep}} - A_m^{\mathrm{c}} C^{\times m}}_{\Delta_m^{\mathrm{keep}}} + CR_m^{\mathrm{f}}. \tag{100}$$

Therefore

$$\boxed{\left\|\Delta_m^{\mathrm{total}}\right\| \le \left\|\Delta_m^{\mathrm{keep}}\right\| + \|C\| \left\|R_m^{\mathrm{f}}\right\|.} \tag{101}$$

For analytic fields on a ball of radius R,

$$\mathfrak{d}_F(R) \le \sum_m \left\|\Delta_m^{\mathrm{keep}}\right\| R^m + \|C\| \sum_m \left\|R_m^{\mathrm{f}}\right\| R^m. \tag{102}$$

The flow and posterior bounds use this combined quantity. A pruning decision is justified only when the omitted part remains explicit and is carried through the later estimates like any other model error.

10.6 A budgeted design problem

These terms lead to a practical design problem. The objective below records the tradeoff without claiming that it can always be solved efficiently:

$$\min_{C,\{A_m^{\mathrm{c}}\},\{\mathcal{E}_m^{\mathrm{keep}}\}} \left[\sum_m w_m \left\|\Delta_m^{\mathrm{total}}\right\| + \lambda \sum_m \mathrm{Cost}(\mathcal{E}_m^{\mathrm{keep}}) + \gamma \sum_y d_{\mathsf{B},y} \right]. \tag{103}$$

The minimization is subject to any required positivity, mass, rank, and interpretability conditions on C. The three terms measure failure of dynamical closure, computational cost, and loss of observational information. Their weights and the admissible packet families must be chosen for the application.

11 Algorithms and finite certificates

Every finite criterion in the paper can be checked directly from coordinates. The reference code uses dense arrays so that each contraction can be inspected. A sparse implementation would loop only over the active hyperedges.

11.1 Exact test for every interaction order

For a tensor of order m in scalar packet coordinates, compute the aggregated coefficient array

$$g^{(m)}_{a;j_1,\dots,j_m} = \sum_{i:\pi(i)=a} a^{\mathrm{f},(m)}_{i;j_1,\dots,j_m}.$$

For every coarse input tuple $\mathbf{b} = (b_1, \dots, b_m)$, define the fine fiber product

$$\mathcal{G}_{\mathbf{b}} = \{(j_1, \dots, j_m) : \pi(j_r) = b_r \text{ for all } r\}.$$

Algorithm 11.1 (Exact packet coordinate lumpability test)**.** For every coarse output a and coarse input tuple $\mathbf{b}$:

1. Evaluate $g^{(m)}_{a;\mathbf{j}}$ for all $\mathbf{j} \in \mathcal{G}_{\mathbf{b}}$.
2. In exact arithmetic, test whether all values are equal. With floating point input, compare their spread with a declared tolerance and record that tolerance.
3. When the exact test passes, assign the common value to $a^{\mathrm{c},(m)}_{a;\mathbf{b}}$. When only the numerical tolerance is met, assign a declared representative such as the mean or the midrange.
4. Otherwise, return a witness $(a, \mathbf{j}, \mathbf{j}')$ with unequal aggregate coefficients.

A dense implementation requires $O(n^{m+1})$ arithmetic to aggregate a tensor of order m on n states. If the fine tensor is stored as E_m active scalar hyperedges, the aggregation pass is $O(E_m)$, followed by grouping and consistency checks over the represented coarse incidences.

Proposition 11.2 (Correctness in exact arithmetic)**.** *With exact arithmetic, Algorithm 11.1 accepts exactly the tensors that descend through the given block sum aggregation. On acceptance, it returns the unique coarse tensor. A positive numerical tolerance produces an approximate certificate rather than an exact algebraic decision.*

Proof. The claim follows directly from Theorem 3.4. □

11.2 Approximate coarse coefficients

When exact fiber constancy fails, the approximation depends on a chosen loss and, when relevant, a task distribution. Under scalar squared loss with nonnegative weights $w_{\mathbf{j}}$ on a fiber product, assume their total weight is positive. The least squares representative is

$$a^{\mathrm{c,LS}}_{a;\mathbf{b}} = \frac{\sum_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} w_{\mathbf{j}} g_{a;\mathbf{j}}}{\sum_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} w_{\mathbf{j}}}. \tag{104}$$

For scalar worst case loss, a minimax representative is the midrange

$$a^{\mathrm{c},\infty}_{a;\mathbf{b}} = \frac{1}{2}\left(\max_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} g_{a;\mathbf{j}} + \min_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} g_{a;\mathbf{j}}\right), \tag{105}$$

with optimal scalar fiber error

$$\varepsilon^{\infty}_{a;\mathbf{b}} = \frac{1}{2}\left(\max_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} g_{a;\mathbf{j}} - \min_{\mathbf{j}\in\mathcal{G}_{\mathbf{b}}} g_{a;\mathbf{j}}\right). \tag{106}$$

These formulas are local coefficient summaries. A global operator norm can couple several fibers and may require a joint optimization.

11.3 Certificate pipeline

The reference implementation follows the dependency chain

$$\begin{array}{c} \text{interaction tensors} \\ \downarrow \\ \text{fiber constancy and kernel tests} \\ \downarrow \\ \text{coarse tensor and } \Delta_m \\ \downarrow \\ \mathfrak{d}_F(R) \text{ and } d_{\Phi,k} \\ \downarrow \\ r_{\lambda,k}, z_{*,k}, d_{\mathsf{B},k} \\ \downarrow \\ e_k^-, e_k^+, \mathfrak{E}_N. \end{array}$$

Each numerical record stores the theorem identifier, the inputs, the residual, the tolerance, and the outcome. The suite includes examples built to satisfy the exact identities and perturbations built to violate them.

11.4 Deterministic generator for exact examples

A coarse tensor can be lifted to a fine tensor by

$$A_m^{\mathrm{f}} = EA_m^{\mathrm{c}}C^{\times m} + K_m, \qquad CK_m = 0. \tag{107}$$

The term K_m can contain complicated hidden interactions without changing the quotient. This makes the construction useful for testing because the correct coarse tensor is known before any numerical calculation.

Proposition 11.3 (Correctness of the exact example generator)**.** *Every tensor constructed by* (107) *satisfies*

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}}C^{\times m}.$$

Conversely, every exact fine tensor admits such a representation for every right inverse E, with

$$K_m = A_m^{\mathrm{f}} - EA_m^{\mathrm{c}}C^{\times m}.$$

Proof. The forward identity follows from $CE = I$ and $CK_m = 0$. For the converse, exact descent gives

$$CK_m = CA_m^{\mathrm{f}} - CEA_m^{\mathrm{c}}C^{\times m} = 0.$$

□

11.5 Reference implementation

The accompanying source tree contains:

Path	Purpose
`src/priq/core.py`	aggregation maps, tensor actions, exact lifts, quotient defects, packet coordinate recovery, Bayesian updates, and a deterministic Runge Kutta integrator
`tests/test_priq.py`	unit tests for the theorems and negative controls
`examples/run_certificates.py`	deterministic tensor, flow, filtering, perturbation, pruning, figure, and ledger generation
`results/verification.json`	certificate summary in JSON form
`tables/test_results.csv`	flat test ledger
`figures/`	vector PDF figures and large PNG figures generated from the certificate run

Numerical Certificate 11.4 (Executed finite suite)**.** The deterministic reference run completed 35 checks with 0 failures. It covers quotient tensors of orders one through three, annihilation of hidden inputs, recovery in packet coordinates, negative controls, and the defect identity across several levels. It also covers nonlinear prediction, Bayesian refinement, evidence aggregation, approximate flow and posterior bounds, perturbation scaling, and orthogonal hyperedge residuals. The independent unit tests also pass in the build environment.

Table 1: Executed deterministic finite certificate ledger.

ID	Check	Value	Threshold	Status
Q1D	order 1 quotient defect	1.19×10^{-16}	$\leq 10^{-10}$	PASS
Q1K	order 1 hidden input residual	5.38×10^{-16}	$\leq 10^{-10}$	PASS
Q1R	order 1 coarse tensor recovery	8.41×10^{-17}	$\leq 10^{-10}$	PASS
Q2D	order 2 quotient defect	3.76×10^{-16}	$\leq 10^{-10}$	PASS
Q2K	order 2 hidden input residual	5.71×10^{-16}	$\leq 10^{-10}$	PASS
Q2R	order 2 coarse tensor recovery	5.55×10^{-17}	$\leq 10^{-10}$	PASS
Q3D	order 3 quotient defect	8.39×10^{-16}	$\leq 10^{-10}$	PASS
Q3K	order 3 hidden input residual	1.15×10^{-15}	$\leq 10^{-10}$	PASS
Q3R	order 3 coarse tensor recovery	1.11×10^{-16}	$\leq 10^{-10}$	PASS
QNEG	a perturbation of one coefficient violates fiber constancy	0.125	$\geq 10^{-1}$	PASS
QSYM	symmetric interaction remains symmetric after exact lift	0	$\leq 10^{-10}$	PASS
QHID	projected hidden tensor has zero coarse output	0	$\leq 10^{-10}$	PASS
QTRN	exact quotient transitivity across two aggregation levels	4.84×10^{-16}	$\leq 10^{-10}$	PASS
QMLT	interaction defect identity across two levels	9.46×10^{-16}	$\leq 10^{-10}$	PASS
QFLD	fine vector field factors through the coarse field	2.22×10^{-16}	$\leq 10^{-10}$	PASS
QPRD	exact aggregate prediction commutation	7.22×10^{-16}	$\leq 5 \times 10^{-10}$	PASS
QCOC	exact commutation of the prediction and update cocycle	3.89×10^{-16}	$\leq 5 \times 10^{-10}$	PASS
QEVD	exact evidence aggregation at one step	6.66×10^{-16}	$\leq 5 \times 10^{-10}$	PASS
QMAS	fine nonlinear trajectory preserves total mass	7.77×10^{-16}	$\leq 5 \times 10^{-10}$	PASS
QPOS	fine nonlinear trajectory preserves nonnegativity	0	$\leq 5 \times 10^{-12}$	PASS

ID	Check	Value	Threshold	Status
QBAY	a likelihood constant within each fiber gives exact Bayesian refinement	2.78×10^{-16}	$\leq 10^{-10}$	PASS
QBEV	a likelihood constant within each fiber preserves evidence	2.22×10^{-16}	$\leq 10^{-10}$	PASS
QSCL	Bayesian update is invariant under global likelihood scale	1.53×10^{-16}	$\leq 10^{-10}$	PASS
QFDB	controlled dynamical perturbation respects the flow certificate	0	$\leq 2 \times 10^{-9}$	PASS
QFSL	flow discrepancy has first order perturbation scaling	3.37×10^{-3}	≤ 0.08	PASS
QLPS	perturbed likelihood remains strictly positive	0.8	$\geq 10^{-12}$	PASS
QOBB	variation of the likelihood within a fiber respects the Bayesian bound	0	$\leq 10^{-12}$	PASS
QBSC	the Bayesian perturbation sweep remains below the likelihood certificate	0	$\leq 10^{-12}$	PASS
QBSL	Bayesian discrepancy has first order scaling for small fiber mismatch	4.63×10^{-5}	≤ 0.08	PASS
QNMB	symmetric normalization bound uses the larger positive mass	0	$\leq 10^{-12}$	PASS
QPST	recursive posterior certificate dominates every joint perturbation run	0	$\leq 10^{-10}$	PASS
QPSL	joint posterior discrepancy has first order scaling for small perturbations	1.23×10^{-4}	≤ 0.12	PASS
QLEV	the evidence certificate at each step dominates the accumulated discrepancy	0	$\leq 10^{-10}$	PASS
QPAR	hyperedge energy satisfies the scalar Parseval identity	8.88×10^{-16}	$\leq 10^{-10}$	PASS
QBST	the residual from retaining the largest energy hyperedges decreases to zero	0	$\leq 10^{-10}$	PASS

12 Worked quotient and filtering models

12.1 A nonlinear fine system with four states and a quotient with two states

Let

$$x = (x_1, x_2, x_3, x_4)^T \in \Delta_4, \qquad C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

so

$$z = Cx = (z_1, z_2)^T, \qquad z_1 = x_1 + x_2, \quad z_2 = x_3 + x_4.$$

The coarse dynamics are

$$\dot{z}_1 = -az_1 + bz_2, \qquad \dot{z}_2 = az_1 - bz_2, \qquad a, b > 0. \tag{108}$$

Define nonlinear internal exchange channels

$$h_{12}(x) = (r_0 + r_1 x_3)x_2 - (s_0 + s_1 x_4)x_1, \tag{109}$$

$$h_{34}(x) = (u_0 + u_1 x_1)x_4 - (v_0 + v_1 x_2)x_3, \tag{110}$$

with all coefficients nonnegative. The fine system is

$$\begin{aligned} \dot{x}_1 &= -ax_1 + bx_3 + h_{12}(x), \\ \dot{x}_2 &= -ax_2 + bx_4 - h_{12}(x), \\ \dot{x}_3 &= ax_1 - bx_3 + h_{34}(x), \\ \dot{x}_4 &= ax_2 - bx_4 - h_{34}(x). \end{aligned} \tag{111}$$

Proposition 12.1 (Exact quotient with nonlinear hidden exchange)**.** *The system* (111) *preserves total mass and the nonnegative orthant. After aggregation by C, it satisfies the coarse system* (108) *exactly.*

Proof. Summing the four equations gives zero. To check positivity, set one coordinate equal to zero. Every term entering its derivative is then nonnegative. This holds for x_1, x_2, x_3, and x_4, so the nonnegative orthant is forward invariant. Finally,

$$\begin{aligned} \dot{x}_1 + \dot{x}_2 &= -a(x_1 + x_2) + b(x_3 + x_4), \\ \dot{x}_3 + \dot{x}_4 &= a(x_1 + x_2) - b(x_3 + x_4), \end{aligned}$$

so the internal nonlinear channels cancel under C and the aggregate obeys (108). □

The quadratic terms in (109) and (110) are genuine fine hyperedges. Their output vectors lie in the internal directions of the two fibers,

$$(1, -1, 0, 0)^T, \qquad (0, 0, 1, -1)^T,$$

which are contained in $\operatorname{Ker} C$. They alter the fine distribution while remaining invisible to the coarse state. Figure 2 shows this distinction. The two fine states inside the first fiber move differently, yet their sum follows the coarse trajectory.

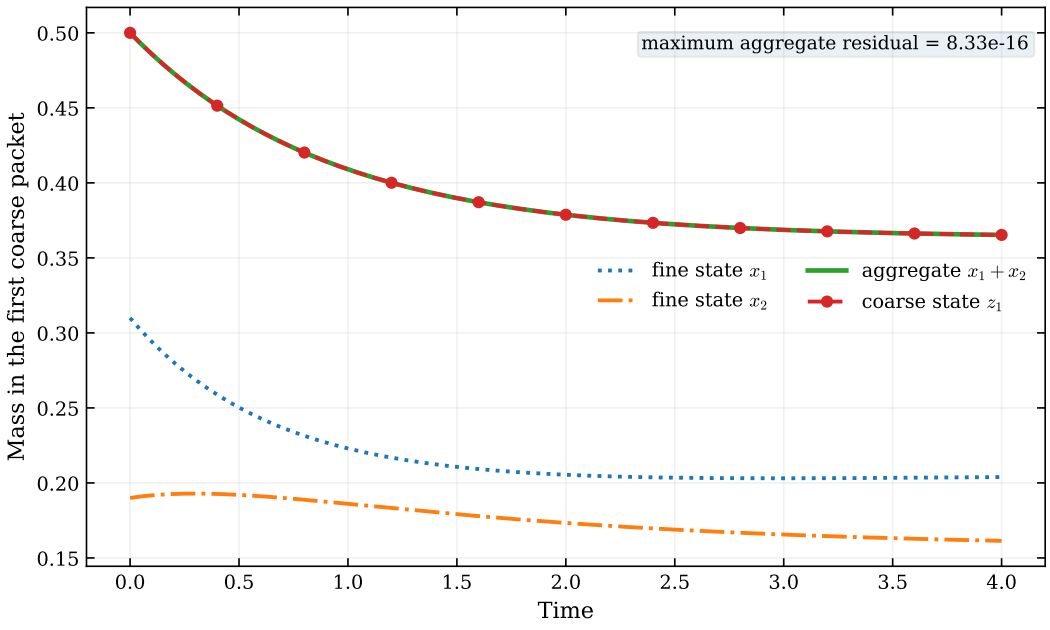


Figure 2: Hidden fine motion under an exact quotient. The microstates x_1 and x_2 evolve separately. Their sum agrees with the coarse coordinate z_1 throughout the trajectory. Sparse markers on the coarse curve make the overlap visible.

12.2 Exact nonlinear filtering

Choose a coarse likelihood

$$\lambda^{\mathrm{c}} = (\lambda_1, \lambda_2)^T > 0$$

and lift it by

$$\lambda^{\mathrm{f}} = C^T \lambda^{\mathrm{c}} = (\lambda_1, \lambda_1, \lambda_2, \lambda_2)^T.$$

By Proposition 12.1 together with Theorems 5.5 and 6.1, every finite sequence of flow segments and such observations satisfies

$$C\mu_k^{\mathrm{f}} = \mu_k^{\mathrm{c}}.$$

The fine filter may continue to distinguish x_1 from x_2 and x_3 from x_4, but those distinctions do not affect the exact coarse posterior.

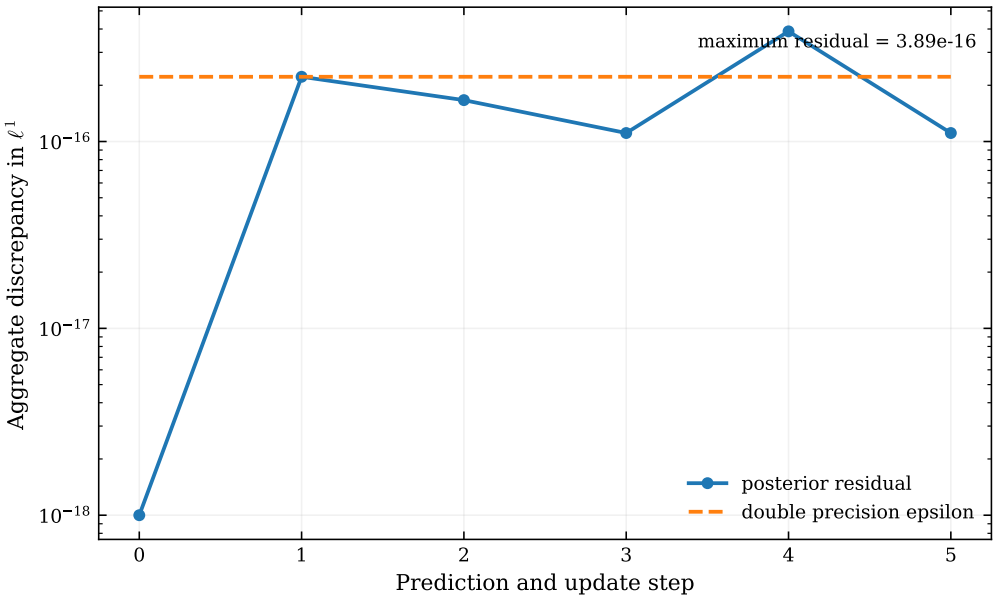


Figure 3: Aggregate posterior discrepancy after each prediction and update stage. The residual remains at floating point scale and stays near double precision epsilon throughout the exact filtering sequence.

Numerical Certificate 12.2 (Exact nonlinear filtering residuals)**.** For the deterministic run with four states, the largest aggregate prediction residual was 7.22×10^{-16}. The largest posterior residual after an update was 3.89×10^{-16}, and the largest residual in a single evidence factor was 6.66×10^{-16}. The small nonzero values come from floating point integration and arithmetic in the finite example.

12.3 A controlled dynamical violation

Add a directed transition between fine states

$$\varepsilon x_1 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{112}$$

to the fine vector field. This perturbation preserves nonnegativity and total mass, but the added transfer between fibers is activated only by the mass in state x_1. The aggregate field defect is

$$\delta_F(x) = \varepsilon x_1 \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \tag{113}$$

so, on the simplex,

$$\|\delta_F(x)\|_1 = 2\varepsilon x_1 \le 2\varepsilon. \tag{114}$$

The coarse Markov flow on two states is an ℓ^1 contraction. Over one interval, this gives

$$\left\| C\Phi_T^{\mathrm{f},\varepsilon}(x) - \Phi_T^{\mathrm{c}}(Cx) \right\|_1 \le 2\varepsilon T.$$

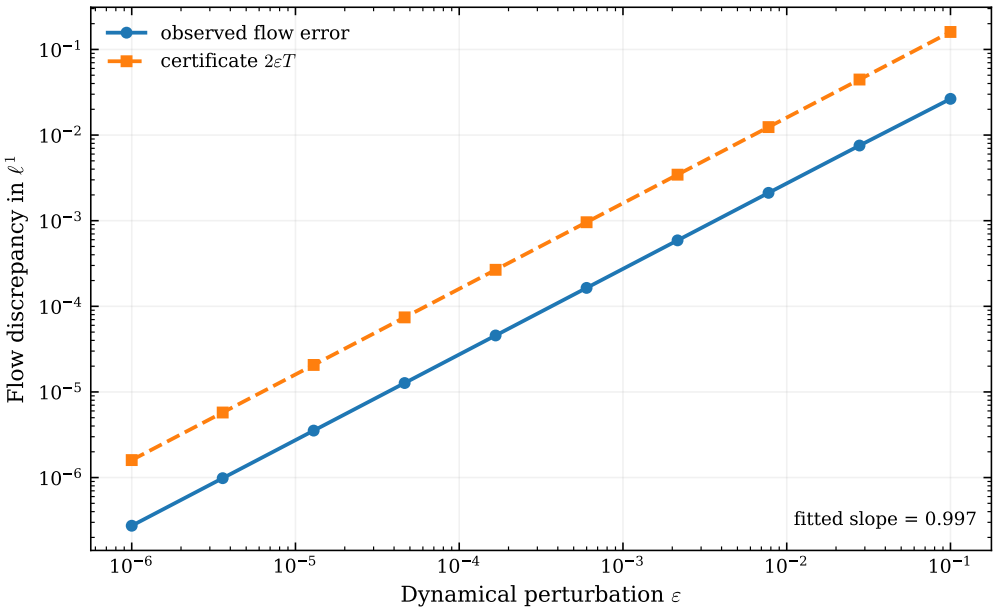


Figure 4: Aggregate flow error across a logarithmic perturbation sweep. The measured discrepancy has first order scaling and remains below the certificate $2\varepsilon T$ from (114).

12.4 A controlled observation violation

Perturb the fine likelihood inside one fiber:

$$\lambda^{\mathrm{f},\eta} = C^T\lambda^{\mathrm{c}} + \eta(1,-1,0,0)^T. \tag{115}$$

Assume $|\eta| < \lambda_1^{\mathrm{c}}$, so every component of $\lambda^{\mathrm{f},\eta}$ remains strictly positive. The fiber oscillation is $2|\eta|$ and

$$\|r_\lambda\|_\infty = |\eta|.$$

If the evidence is bounded below by $z_* > 0$, then

$$d_{\mathsf{B}} \le \frac{2|\eta|}{z_*}.$$

The sweep in Figure 5 varies only the observation model and isolates this effect at one interior prior. The measured discrepancy is first order in η and remains below the likelihood certificate.

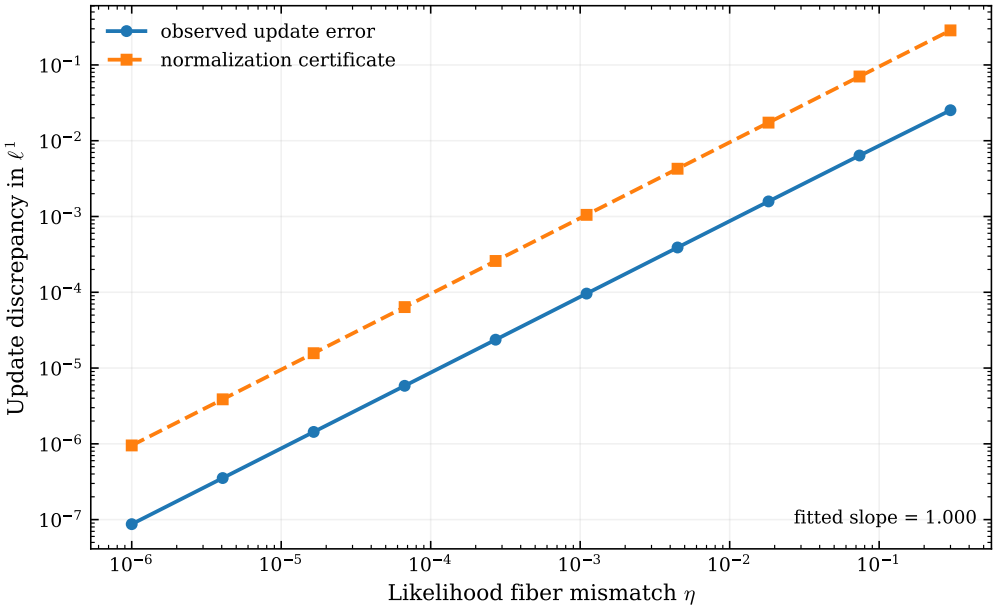


Figure 5: Posterior discrepancy caused by variation of the likelihood inside one aggregation fiber. The error scales linearly with η in the tested range and remains below the bound obtained from the normalization estimate.

Combining this observation perturbation with (112) gives a posterior discrepancy that is first order in (ε, η) over a fixed finite horizon. Theorem 9.3 supplies a conservative upper bound.

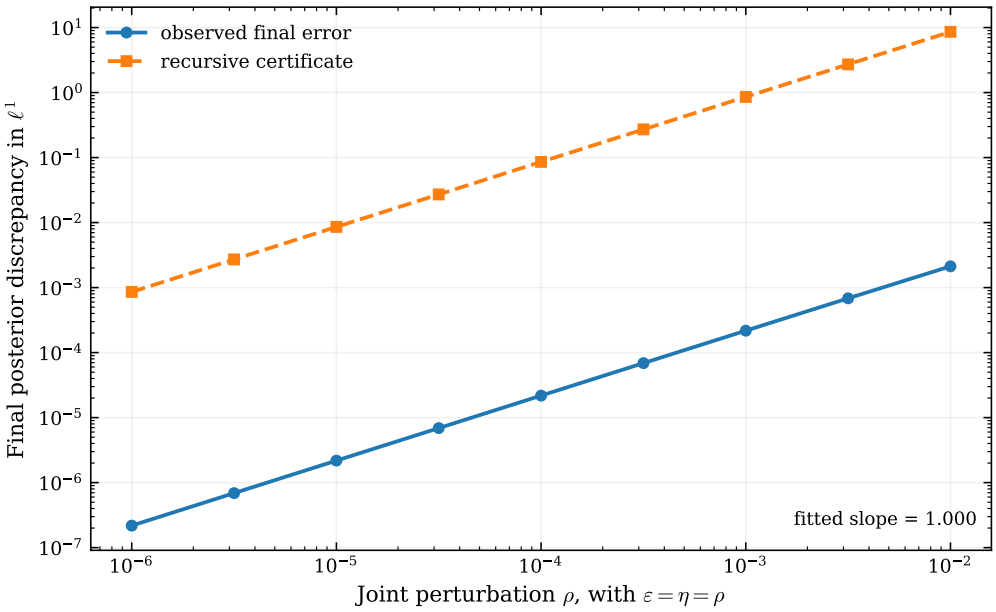


Figure 6: Joint dynamical and likelihood perturbation sweep with $\varepsilon = \eta = \rho$. For small ρ, the final posterior discrepancy grows linearly, while the recursive certificate remains above every measured value.

Numerical Certificate 12.3 (Perturbation scaling)**.** The experiment used logarithmically spaced values with $\varepsilon = \eta$. The fitted slope of the final posterior discrepancy was 0.9999. Every measured error remained below its recursive certificate.

12.5 Random exact tensors through order three

The implementation draws random coarse tensors of orders $m = 1, 2, 3$, lifts them through (107), and adds random hidden output tensors K_m with $CK_m = 0$. It then checks:

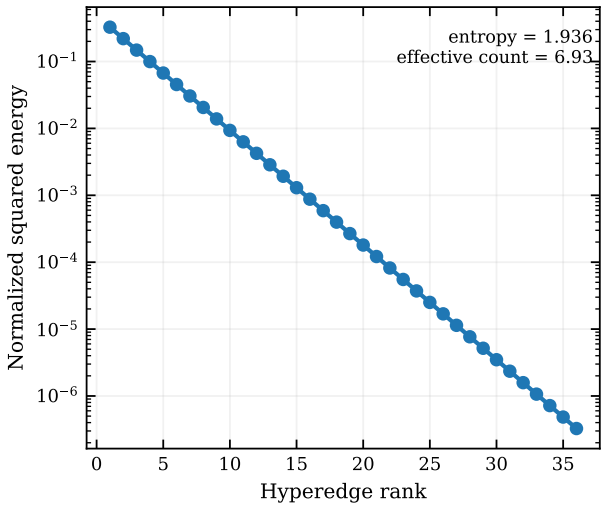
$$CA_m^{\mathrm{f}} - A_m^{\mathrm{c}} C^{\times m} = 0,$$

annihilation of hidden inputs, constancy on packet fibers, and recovery of the original coarse tensor. A single coefficient perturbation provides a negative control that must violate fiber constancy.

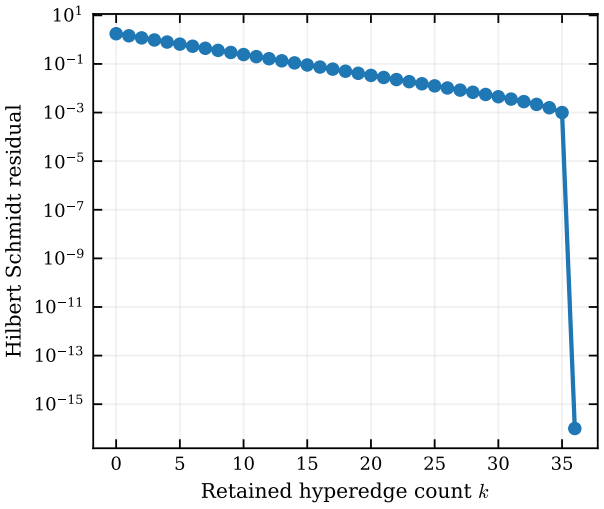
This deterministic check confirms that the coordinate implementation and the invariant theorem return the same quotient. It is not evidence for a general probabilistic law.

12.6 Concentrated hyperedge field

A synthetic interaction field with decreasing orthogonal block energies illustrates Theorem 10.3. The computed hyperedge entropy is 1.9360, corresponding to an effective count of approximately 6.931 hyperedges despite a larger ambient support. The ranked spectrum in Figure 7a shows why the field is compressible in this representation.



(a) Ranked normalized hyperedge energy.



(b) Residual after retaining the k largest energy hyperedges.

Figure 7: Energy concentration and certified pruning in the synthetic interaction field. A small leading set carries most of the energy, and the residual decreases to zero when the complete support is restored.

12.7 Interpretation for reaction networks

Consider a polynomial reaction system

$$\dot{x} = Sv(x), \qquad v_r(x) = k_r \prod_i x_i^{\alpha_{ir}}.$$

Each reaction monomial contributes one interaction hyperedge to every species changed by the corresponding stoichiometric column. The inputs are the reactants. A block sum aggregation is exact precisely when the monomial coefficients, after summation over each output fiber, are constant on the product

fibers specified by Theorem 3.4. This gives a coefficient test at every interaction order. Specialized reductions of reaction networks may also use stoichiometric, kinetic, or bisimulation structure [7, 8, 2].

Claim Boundary 12.4 (Physical interpretation)**.** The quotient theorem is algebraic. It does not by itself decide whether a packet represents a species class, spatial region, symmetry sector, latent state, or learned feature. The scientific meaning of C must come from the application and is not created by exact factorization alone.

13 Relation to neighboring theories

The results fit into a common finite quotient language, but each ingredient also belongs to a larger established theory.

13.1 Markov lumpability and Markov functions

For unary stochastic dynamics, the identity

$$CQ_{\mathrm{f}} = Q_{\mathrm{c}}C$$

is a standard exact lumping condition. It implies

$$Ce^{tQ_{\mathrm{f}}} = e^{tQ_{\mathrm{c}}}C$$

and means that the coarse process is Markov for every fine initial distribution compatible with the quotient. Classical treatments distinguish several forms of strong, ordinary, exact, and weak lumpability, depending on the admissible initial laws and the direction of the stochastic convention [12, 14, 5]. Here we fix one column vector convention and extend the coefficient criterion from unary edges to multilinear hyperedges.

13.2 Polynomial differential equivalence

Exact aggregation of polynomial vector fields has been developed through differential equivalences, bisimulations, and algorithms based on successive partition refinement, including applications to chemical reaction networks [7, 8]. The universal factorization theorem here is elementary multilinear algebra once C is fixed. Its value is that it exposes every interaction order, makes the quotient condition independent of a particular polynomial syntax, and connects the same coefficient criterion to rooted tree composition and Bayesian updates. We do not claim an algorithmic advantage over specialized methods based on partition refinement.

13.3 Hypergraphs, operads, and rooted trees

A multilinear packet block has several inputs and one output, so its incidence is naturally hypergraphic. Rooted substitution of such blocks is operadic, and the derivative formula for composition is the multivariate Faà di Bruno formula [9]. Rooted trees also organize Butcher series and expansions for nonlinear time stepping [6, 10]. In this language, the quotient condition is a morphism condition. It commutes with every generating operation and therefore with every rooted composite.

13.4 Sufficiency and comparison of experiments

A statistic is sufficient only relative to an observation family. Classical sufficiency and comparison of experiments theory formalize when one representation preserves all information relevant to a chosen family [11, 4]. The likelihood theorem in Section 5 is a finite and explicit special case for deterministic aggregation and diagonal positive observation operators. It characterizes exact commutation of the normalized update, not all possible notions of statistical equivalence.

13.5 Model reduction and nearly decomposable systems

Classical approaches to approximate aggregation use time scale separation, perturbation theory, balanced truncation, stochastic complementation, and related forms of model reduction [15, 13, 1]. The bounds used here follow the error through each stage. Coefficient defects first create a field defect. The Grönwall inequality carries that defect through the flow, and the normalization estimate carries it through each observation. The resulting certificate can be conservative, but it keeps every source of error visible.

14 Claim boundaries and limitations

Claim Boundary 14.1 (Finite dimensional foundation)**.** All quotient and coefficient theorems are proved in finite dimensional vector spaces. The algebraic factorization extends to arbitrary vector spaces, but boundedness, convergence, measurability, domains of unbounded generators, and interchange of infinite sums require additional hypotheses in Banach, Hilbert, or measurable field settings.

Claim Boundary 14.2 (Local analytic dynamics)**.** The flow theorems assume local existence and uniqueness. The analytic coefficient theorem supplies a local coarse field, not a global invariant manifold or a global conjugacy. Solutions may blow up, uniqueness may fail, and boundary singularities remain specific to the application.

Claim Boundary 14.3 (A prescribed coarse map)**.** The principal theorems assume that C is supplied. They test and quantify whether a proposed aggregation is valid. Discovering an optimal C from data or from a large interaction tensor is a separate problem and can be combinatorial, nonconvex, or statistically nonidentifiable.

Claim Boundary 14.4 (Positivity and evidence floors)**.** Bayesian normalization requires positive evidence. The uniform perturbation bounds degrade as the evidence floor approaches zero. This is not merely a weakness of the proof. Small normalizing constants can amplify small absolute errors into large posterior changes.

Claim Boundary 14.5 (Norm dependence)**.** Exact descent is algebraic and independent of the norm. Best approximate coarse tensors, Lipschitz constants, pruning choices, and posterior certificates depend on the selected norms and task weights. A useful certificate must state those choices.

Claim Boundary 14.6 (Numerical verification)**.** The executed checks validate the finite implementation against the stated formulas and include negative controls. They do not prove the theorems, establish asymptotic behavior outside the tested ranges, or demonstrate universal computational speedup.

Claim Boundary 14.7 (Historical priority)**.** The central synthesis joins several established families of ideas. The coefficient criterion at every interaction order follows from multilinear algebra on a quotient space. Equivalent statements may already appear in neighboring theories of nonlinear aggregation under different notation. Any claim of novelty should distinguish this elementary factorization from the combined treatment of interaction trees, Bayesian cocycles, and error certificates. Such a claim also requires a dedicated literature review.

15 Open theorem and algorithm program

15.1 Coarsest exact joint partition

For a finite interaction family $(A_m^{\mathrm{f}})_{m\le M}$ and likelihood family $(\lambda_y^{\mathrm{f}})_{y\in\mathcal{Y}}$, seek

$$\min_{\pi} |I_{\mathrm{c}}| \tag{116}$$

subject to

$$CA_m^{\mathrm{f}} = A_m^{\mathrm{c}} C^{\times m} \quad (m \le M),$$

and

$$\lambda_y^{\mathrm{f}} = s_y C^T \lambda_y^{\mathrm{c}} \quad (y \in \mathcal{Y}).$$

The partition determined by the observations alone groups states with identical likelihood profiles. The dynamical condition also depends on the partition because changing the partition changes the output aggregation. A terminating refinement algorithm, together with a useful complexity bound, remains an open problem.

15.2 Approximate joint design

A finite approximate problem is

$$\min_{C,\{A_m^{\mathrm{c}}\},\{\lambda_y^{\mathrm{c}}\}} \left[\sum_m w_m \left\| CA_m^{\mathrm{f}} - A_m^{\mathrm{c}} C^{\times m} \right\| + \sum_y v_y d_{\mathsf{B},y} + \gamma\, \mathrm{Complexity}(C) \right].$$

The main questions concern existence, tractable relaxations, and stability of the minimizers. Another question is whether the task weights can be learned without invalidating the certificate after a distribution shift.

15.3 Posterior distributions over resolutions

Instead of selecting one C, place a posterior on a finite family of candidate resolutions:

$$p(C \mid D) \propto p(D \mid C) p(C).$$

Predictions then average both parameter and topology uncertainty:

$$p(z_t \mid D) = \sum_C p(z_t \mid C, D) p(C \mid D).$$

A fuller model should distinguish uncertainty in the coefficients of a fixed hypergraph from uncertainty about the hypergraph and the quotient map themselves.

15.4 Stochastic interaction quotients

For a diffusion

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t,$$

the generator is

$$\mathcal{L}\varphi = D\varphi[b] + \frac{1}{2}\,\mathrm{tr}(\sigma\sigma^T D^2\varphi).$$

Packet resolution turns drift into unary edges and quadratic variation into contracted binary hyperedges. Let $\mathcal{C}$ denote the pullback or averaging map on observables induced by the aggregation. The natural exact condition is generator intertwining

$$\mathcal{C}\mathcal{L}_{\mathrm{f}} = \mathcal{L}_{\mathrm{c}}\mathcal{C}$$

for the pullback or averaging map appropriate to observables. A complete theorem would characterize closure of the drift and covariance, prove positivity of the induced semigroup, and carry generator defects into filtering error bounds.

15.5 Coarse dynamics with memory

Eliminating hidden packets generally produces memory. The target form is

$$\dot{z}(t) = A_{PP}z(t) + \int_0^t K(t-s)z(s)\,\mathrm{d}s + \eta(t).$$

The quotient theory should allow a fine system without memory to produce coarse dynamics with memory. Volterra hyperedges and fractional kernels would then be treated as native coarse interactions rather than as evidence that quotienting failed.

15.6 Sharper filtering metrics

The ℓ^1 normalization bounds are robust but conservative. Positive operators often contract projective metrics, relative entropy, or other information geometric divergences more sharply. One goal is to replace the factor

$$\frac{2\,\|\lambda\|_\infty}{z_*}$$

by a contraction coefficient adapted to the trajectory or to the operator, while retaining an auditable finite certificate.

15.7 Pruning that preserves structure

In Hamiltonian, reversible, conservative, and monotone systems, as well as systems with detailed balance, individual hyperedges cannot always be removed independently. The problem is to identify groups whose joint retention preserves the required structure, then solve the budget problem over those groups.

15.8 Identifiability of hidden dynamics

For fixed C and coarse field F_{c}, the exact fine realizations form the affine family

$$\{F_{\mathrm{f}} : CF_{\mathrm{f}} = F_{\mathrm{c}}C\}.$$

Any two exact fine models differ by a map into $\operatorname{Ker} C$. Coarse observations alone cannot identify that hidden component. Additional fine observations, intervention assumptions, regularity, or minimum complexity principles are required. A statistical theory should characterize precisely which hidden hyperedges are identifiable under a given observation and intervention design.

16 Conclusion

An exact aggregation is more than a useful summary of a fine model. It is a quotient. Fine states with the same coarse image must produce the same coarse output. For a multilinear interaction, this reduces to a complete finite test. After applying C, the interaction must vanish whenever any input lies in $\operatorname{Ker} C$. In packet coordinates, the coefficients summed over each output fiber must be constant on every product of input fibers. The familiar unary criterion for Markov lumpability is the first order case of this general condition.

The coefficient test controls the nonlinear constructions built from those interactions. It passes to analytic vector fields, their derivatives, and every rooted interaction tree. Under the usual local uniqueness assumptions, it also passes to the flow. Bayesian observation has a matching quotient condition. A normalized update commutes with aggregation exactly when the fine likelihood is constant within each fiber, apart from a positive global scale. Once both conditions hold, every finite sequence of predictions and updates agrees after aggregation, and the evidence factors agree up to the same scales.

When exactness fails, the reason matters. A mismatch in the interaction coefficients creates a field defect, and the flow can amplify it. Variation of the likelihood inside a fiber creates a second error at normalization. Removing hyperedges creates a third error in the form of an explicit residual tensor. The posterior recurrence keeps these contributions separate as it carries them forward. An approximate coarse model is therefore auditable because one can see where each part of the final error entered.

The theory does not show that an optimal quotient can always be found cheaply. Coarse observations alone also cannot determine hidden dynamics that lie in $\operatorname{Ker} C$. Those questions remain open. The result proved here is narrower and exact:

> A coarse model is exact when its interactions and observations descend through the same quotient. When they do not, each failure can be measured at the stage where it occurs and carried forward to the final posterior or observable by an explicit certificate.

A Formulation on the quotient space

Let

$$q : X_{\mathrm{f}} \to X_{\mathrm{f}}/\operatorname{Ker} C$$

be the canonical quotient map. Since C is surjective, there is a unique vector space isomorphism

$$\widetilde{C} : X_{\mathrm{f}}/\operatorname{Ker} C \to X_{\mathrm{c}}$$

satisfying

$$C = \widetilde{C} q.$$

The multilinear map $CA_m^{\mathrm{f}} : X_{\mathrm{f}}^m \to X_{\mathrm{c}}$ descends if and only if its value is unchanged when any one input is altered by an element of $\operatorname{Ker} C$. The induced quotient tensor is

$$\overline{A}_m([x_1], \dots, [x_m]) = CA_m^{\mathrm{f}}(x_1, \dots, x_m),$$

and

$$A_m^{\mathrm{c}} = \widetilde{C}\, \overline{A}_m(\widetilde{C}^{-1}, \dots, \widetilde{C}^{-1}).$$

This formulation shows that Theorem 3.1 is independent of coordinates and of the chosen right inverse.

B A finite observational partition lemma

Let $\mathcal{Y}$ be a finite observation alphabet and let each state i have a strictly positive likelihood profile

$$\Lambda(i) = (\lambda_y(i))_{y\in\mathcal{Y}}.$$

Define

$$i \sim_{\mathrm{obs}} j \quad \Longleftrightarrow \quad \Lambda(i) = \Lambda(j).$$

Proposition B.1 (Coarsest deterministic observation partition)**.** *The equivalence classes of $\sim_{\mathrm{obs}}$ form the coarsest partition for which every likelihood in the family is constant on every block. Consequently, this partition is the coarsest deterministic block sum aggregation satisfying exact Bayesian refinement for all $y \in \mathcal{Y}$, up to one global positive scale for each observation.*

Proof. Every likelihood is constant on an equivalence class by definition. Any partition on which every likelihood is constant on each block can place i and j in the same block only if all profile coordinates agree, so it refines the equivalence class partition. The Bayesian statement follows from Theorems 5.5 and 5.6. □

C Certificate schemas and reproducibility

A minimal interaction certificate stores

$$\mathsf{L}_m = (m, C, A_m^{\mathrm{f}}, A_m^{\mathrm{c}}, \Delta_m, \|\Delta_m\|, \mathrm{norm}, \mathrm{tolerance}). \tag{117}$$

A Bayesian certificate stores

$$\mathsf{B}_y = (\lambda_y^{\mathrm{f}}, \lambda_y^{\mathrm{c}}, s_y, r_y, z_{*,y}, d_{\mathsf{B},y}). \tag{118}$$

A filtering certificate stores

$$\mathsf{F}_N = (e_0^+, \{L_{\Phi,k}\}, \{L_{\mathsf{B},k}\}, \{d_{\Phi,k}\}, \{d_{\mathsf{B},k}\}, \{e_k^-, e_k^+\}). \tag{119}$$

The release serializes the executed finite ledger as JSON files and CSV tables. A certificate records a computation under specified inputs. The ledger provides cryptographic authenticity only when it is paired with a separate signed digest layer.

C.1 Execution commands

From the source root:

```
PYTHONPATH=src python examples/run_certificates.py
PYTHONPATH=src python -m pytest -q
```

The first command regenerates every numerical table and figure used in the manuscript. The second runs the independent unit tests.

D Theorem dependency ledger

ID	Result	Dependency
PRIQ-001	Exact multilinear packet reconstruction	packet resolutions
PRIQ-002	Universal multilinear quotient criterion	factorization through a quotient space
PRIQ-003	Coefficient criterion at every interaction order	PRIQ-002, basis expansion
PRIQ-004	Symmetry descends	PRIQ-002
PRIQ-005	Exact quotient transitivity	intertwining algebra
PRIQ-006	Multilevel defect identity	algebraic decomposition
PRIQ-007	Analytic interaction descent	PRIQ-002, majorant convergence
PRIQ-008	Jet naturality	differentiation of intertwining
PRIQ-009	Treewise quotient naturality	PRIQ-002, rooted induction
PRIQ-010	Exact quotient flow	intertwining of vector fields and uniqueness
PRIQ-011	Normalized transfer naturality	positive intertwining, mass preservation
PRIQ-012	Sufficiency of likelihood constancy within fibers	PRIQ-011
PRIQ-013	Necessity of likelihood constancy within fibers	posterior ratios for two packets
PRIQ-014	Exact prediction and update cocycle	PRIQ-010, PRIQ-012
PRIQ-015	Sequential evidence factorization	PRIQ-014
PRIQ-016	Expansion by hidden inputs	decomposition through a right inverse
PRIQ-017	Majorant for the analytic field defect	coefficient bounds
PRIQ-018	Approximate quotient flow bound	Grönwall inequality
PRIQ-019	Normalization perturbation inequality	triangle inequality
PRIQ-020	Bayesian refinement defect	PRIQ-019
PRIQ-021	Bayesian Lipschitz certificate	PRIQ-019
PRIQ-022	Posterior error recurrence	prediction and update decompositions
PRIQ-023	Closed posterior certificate	iteration of PRIQ-022
PRIQ-024	Evidence error certificate	posterior error, likelihood residual
PRIQ-025	Hyperedge Parseval identity	orthogonal tensor blocks
PRIQ-026	Optimal approximation with k hyperedges	PRIQ-025
PRIQ-027	Combined pruning and quotient defect	residual decomposition

E Notation index

Symbol	Meaning
$X_{\mathrm{f}}, X_{\mathrm{c}}$	fine and coarse state spaces
$P_i, \overline{P}_a$	fine and coarse packet projectors
π	map from fine packet labels to coarse labels

Symbol	Meaning
C	surjective aggregation or quotient map
E	right inverse selecting a fine representative of each coarse state
$P_C = EC$	representative projection
$Q_C = I - EC$	projection onto the hidden directions, with range $\operatorname{Ker} C$
$A_m^{\mathrm{f}}, A_m^{\mathrm{c}}$	fine and coarse interaction tensors of order m
Δ_m	quotient defect at interaction order m
$F_{\mathrm{f}}, F_{\mathrm{c}}$	fine and coarse vector fields
$\Phi_t^{\mathrm{f}}, \Phi_t^{\mathrm{c}}$	fine and coarse flows or prediction maps
$L_{\mathrm{f}}, L_{\mathrm{c}}$	positive observation transfer operators
B_λ	Bayesian normalized likelihood update
Γ_k	single step prediction and update map
$\Psi_{1:N}$	cocycle for a finite observation sequence
$d_{\Phi,k}$	direct prediction quotient defect at step k
$d_{\mathsf{B},k}$	direct Bayesian refinement defect at step k
e_k^-, e_k^+	aggregate errors before and after an update
$z_{*,k}$	positive evidence floor
R_m	omitted hyperedge residual tensor
H_m	entropy of the normalized hyperedge energy distribution at order m

References

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM, 2005.

[2] J. C. Baez and B. S. Pollard, A compositional framework for reaction networks, *Reviews in Mathematical Physics* 29 (2017), 1750028.

[3] L. E. Baum and T. Petrie, Statistical inference for probabilistic functions of finite state Markov chains, *Annals of Mathematical Statistics* 37 (1966), 1554–1563.

[4] D. Blackwell, Equivalent comparisons of experiments, *Annals of Mathematical Statistics* 24 (1953), 265–272.

[5] P. Buchholz, Exact and ordinary lumpability in finite Markov chains, *Journal of Applied Probability* 31 (1994), 59–75.

[6] J. C. Butcher, An algebraic theory of integration methods, *Mathematics of Computation* 26 (1972), 79–106.

[7] L. Cardelli, M. Tribastone, M. Tschaikowski, and A. Vandin, Forward and backward bisimulations for chemical reaction networks, in *CONCUR 2015*, LIPIcs 42 (2015), 226–239.

[8] L. Cardelli, M. Tribastone, M. Tschaikowski, and A. Vandin, Maximal aggregation of polynomial dynamical systems, *Proceedings of the National Academy of Sciences* 114 (2017), 10029–10034.

[9] G. M. Constantine and T. H. Savits, A multivariate Faà di Bruno formula with applications, *Transactions of the American Mathematical Society* 348 (1996), 503–520.

[10] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, second edition, Springer, 2006.

[11] P. R. Halmos and L. J. Savage, Application of the Radon–Nikodym theorem to the theory of sufficient statistics, *Annals of Mathematical Statistics* 20 (1949), 225–241.

[12] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Van Nostrand, 1960.

[13] C. D. Meyer, Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems, *SIAM Review* 31 (1989), 240–272.

[14] L. C. G. Rogers and J. W. Pitman, Markov functions, *Annals of Probability* 9 (1981), 573–582.

[15] H. A. Simon and A. Ando, Aggregation of variables in dynamic systems, *Econometrica* 29 (1961), 111–138.