Residue Spectral Dynamics and Markov Closure in Finite Distinction Systems
Coupled Residue Channels, Transfer Spectra, and Operational Recoverability

Yining Wu
Independent Researcher
yining.wu@alumni.upenn.edu
(Dated: June 2026)

Finite effective descriptions retain only a bounded set of distinctions. Projection, pruning, coarse-graining, and externalization remove distinctions from the active record, but the unresolved sector may continue to influence accessible dynamics through hidden modes, environmental side records, delayed correlations, protected invariants, recurrence, or instability. This paper develops FDS-P8 as the spectral-dynamical bridge that organizes those returns. The central object is not the unresolved generator alone, but the coupled residue channel $\mathfrak{R} = (G_R, B_R, C_R)$, where $G_R = \mathsf{Q}\mathcal{L}\mathsf{Q}$ governs residue evolution, $B_R = \mathsf{Q}\mathcal{L}\mathsf{P}$ excites residue from retained variables, and $C_R = \mathsf{P}\mathcal{L}\mathsf{Q}$ returns residue to the retained sector. The visible memory kernel is $K(t) = C_R S_R(t) B_R$, with transfer representation $\widehat{K}(s) = C_R(sI - G_R)^{-1}B_R$. The paper separates formal projection from physical pruning and thermodynamic erasure; distinguishes formal, physical, excitable, return-active, coupled, silent, accessible, and protected residue; and proves that the retained input-output map factors through an active residue quotient. It distinguishes the internal spectral set $\Lambda_{\mathrm{int}} \subset \mathbb{C}$, the transfer singularity structure $\mathfrak{S}_{\mathrm{tr}}$, the initial-residue recoverability spectrum $\Sigma^{\mathrm{init}}_{\mathrm{rec}}(T)$, the reachable round-trip recoverability spectrum $\Sigma^{\mathrm{rt}}_{\mathrm{rec}}(T)$, and their associated latent/history and physical residue-state subspaces, which are different mathematical objects. The operational closure regime is controlled by the full coupled reduced-response data: the transfer response, including its singularities, modal weights, finite-window Hankel spectrum, resolvent and non-normal gain, together with initial-residue forcing—relative to a registered norm, window, task, and tolerance. Conditional results establish exponential-stability bounds for finite-memory Markov closure, a coupled marginal-mode obstruction, algebraic tails from low-frequency spectral edges under registered Tauberian conditions, finite exact Markov embeddings for strictly proper rational kernels, a finite-window task-risk certificate, and an infinite-horizon balanced-truncation certificate. A transfer–recoverability bridge theorem bounds reachable round-trip recoverability singular values and operational ranks by the corresponding transfer Hankel singular values whenever the physical recovery channel is boundedly equivalent to the retained return channel on the closed transfer-signal space generated by the reachable residue sector. Numerical normal forms illustrate silent zero modes, transfer-visible poles, controlled closure windows, non-normal transient amplification, observability-limited recovery, finite state augmentation, and a fourth-order conditional-information onset under controlled lumpability breaking. P8 does not claim to invent projection-operator or realization theory, nor to derive conservation laws, three-dimensional space, time arrows, inertia, or gravity. It supplies a common formal spine for those later FDS bridges: projection determines what leaves the active record; the coupled residue channel determines what can return.

Scope and Claim Status. P8 is a formal and physical-bridge paper. The exact projection identity, semigroup estimates, controllability, observability, Hankel operators, balanced truncation, minimal realization, and Tauberian tools are standard mathematical inputs. The FDS contribution is the integration of those inputs with finite task-relative projection, physical residue qualification, accounting boundaries, operational recoverability, and downstream failure discipline. The principal P8 organizing objects are the coupled residue channel, the full coupled reduced-response data, a type-correct separation of internal spectral sets, transfer singularities, initial-state and reachable round-trip recoverability spectra, their correctly typed latent/history and residue-state subspaces, and the operational residue rank. The principal new bridge result is a finite-window transfer–recoverability singular-value sandwich under a registered bounded interconnection between the retained return channel and the physical recovery channel. Main realization, rank, and Gaussian recoverability statements are finite-dimensional or bounded-operator results. Infinite-dimensional extensions require separate admissibility, compactness, quotient-semigroup, and trace-class hypotheses. P8 does not identify every unresolved model variable with a physical carrier, does not claim that every physical memory is projection-induced, and does not infer thermodynamic heat from formal projection alone.

Claim-status summary

Keywords: finite distinction systems; residue dynamics; pruning; projection operators; Mori–Zwanzig; transfer singularities; recoverability singular values; memory

TABLE I. Central FDS-P8 claims, status, and demotion conditions.

Claim	**Status**	**Demotion or failure condition**
Projection yields an instantaneous–memory–forcing decomposition	Imported exact identity, specialized in the exact-dynamics section	The declared generator, projection, domain, or semigroup assumptions fail, or a nonlinear state equation is used without a valid lifted observable representation
Only controllable-and-observable residue contributes to visible round-trip memory	Restricted theorem in the active-residue section	Removing a declared silent sector changes the memory kernel, transfer function, or retained input-output behavior
Exponential residue-semigroup stability gives a finite memory-burden bound	Conditional theorem in the closure section	Exponential stability does not hold, couplings are unbounded on the relevant domain, or the retained solution class violates the finite-horizon assumptions
Coupled marginal modes obstruct integrable memory closure	Conditional theorem in the marginal-mode section	The marginal mode is silent, canceled by the transfer channel, removed by symmetry, or absent from the registered active quotient
Low-frequency transfer spectral weight can generate algebraic memory tails	Tauberian result for a restricted positive-measure class	The kernel is non-normal, sign-changing, matrix-canceling, time dependent, or outside the registered regular-variation assumptions
Transfer significance bounds recoverability when the two observation channels are boundedly equivalent on reachable residue	P8 channel-registration bridge theorem	No bounded interconnection exists, the lower bound vanishes, or recovery uses side records outside the registered retained return channel
Operational recoverability depends on the observation channel, not on residue lifetime alone	Exact finite-dimensional linear-Gaussian model theorem	The claimed recovery is obtained only by changing the accounting boundary, decoder resources, task labels, or observation channel
Strictly proper rational memory admits finite exact Markov augmentation	Imported realization theorem, FDS interpretation	A non-rational kernel is asserted to have a finite exact realization, the feedthrough term is not absorbed into the instantaneous generator, or the proposed augmentation fails to reproduce the transfer function

kernels; Markov closure; operational recoverability; controllability; observability; Hankel operators; minimal realization; non-normal dynamics; long-time tails.

INTRODUCTION

The missing question after coarse-graining

A reduced description is usually introduced by saying which variables are retained and which are discarded. That description is incomplete whenever discarded variables remain dynamically coupled to the retained sector. The unresolved state may continue to matter through delayed response, colored noise, history dependence, recurrence, protected records, or instability. Projection-operator methods formalize this fact by replacing eliminated variables with an exact memory kernel and a forcing term [6–10]. P8 accepts that standard result and asks a more specific finite-system question:

> Which distinctions outside the active record can be excited by the retained sector, survive in an unresolved carrier, and later return to alter future accessible dynamics?

The answer cannot be read from the dimension of the hidden space alone. A hidden zero mode may be exactly decoupled and therefore invisible. Conversely, one slowly decaying mode with strong input-output coupling may dominate the entire reduced history. The relevant object is not "the hidden spectrum" in isolation, but the spectrum filtered by excitation and return.

P8 thesis

Let P denote a registered retained projection and $\mathsf{Q} = I - \mathsf{P}$ its unresolved complement. P8 defines the coupled residue channel

$$\boxed{\mathfrak{R} = (G_R, B_R, C_R),} \tag{1}$$

with

$$G_R = \mathsf{Q}\mathcal{L}\mathsf{Q}, \qquad B_R = \mathsf{Q}\mathcal{L}\mathsf{P}, \qquad C_R = \mathsf{P}\mathcal{L}\mathsf{Q}. \tag{2}$$

Here G_R evolves residue internally, B_R maps retained distinctions into residue, and C_R returns residue to retained observables. The visible memory is

$$K(t) = C_R S_R(t) B_R, \qquad S_R(t) = e^{tG_R}, \tag{3}$$

and, where the resolvent exists,

$$\widehat{K}(s) = C_R(sI - G_R)^{-1} B_R. \tag{4}$$

The paper's central claim is deliberately restricted:

> **Restricted P8 thesis.** Relative to a registered finite projection, only the controllable-and-observable part of the unresolved sector contributes to reduced memory. The operational closure regime is controlled by the full coupled reduced-response data: the transfer response, including its singularities, modal weights, finite-window Hankel spectrum, resolvent and non-normal gain, together with initial-residue forcing—relative to a specified norm, time window, task, and tolerance. The transfer singularity structure is one component of this response, not a sufficient closure classifier by itself.

Why this is an FDS paper

Mori–Zwanzig theory already states that unresolved variables produce memory and noise. Control theory already states that uncontrollable or unobservable states can be removed from a minimal input-output realization. P8 does not relabel those facts as new mathematics. Its additional role is to integrate them with the FDS accounting requirements: why a distinction was pruned, where it is physically carried, whether it is accessible through a finite boundary, what task would count as recovery, what time and error tolerance apply, and what resource expansion would be required to restore it. The FDS object is therefore not merely a transfer function. It is a transfer function embedded in a registered finite-access ledger.

Contributions

The paper makes eight contributions.

1. It installs a projection–pruning firewall separating formal model reduction from physical pruning, externalization, and thermodynamic erasure.
2. It defines formal, physical, excitable, return-active, coupled, silent, accessible, and protected residue.
3. It states an exact coupled-residue decomposition in an observable-space semigroup setting.
4. It proves a restricted active-residue reduction theorem and separates the transfer singularity structure from the full finite-window transfer response.
5. It introduces finite-window transfer and reachable round-trip recovery ranks through Hankel singular values, while keeping initial-state recoverability separate.
6. It derives closure bounds from exponential semigroup stability and closure obstructions from coupled marginal modes and spectral edges.
7. It gives exact model-class relations between residue observability and recoverability, and between rational kernels and finite Markov embeddings.
8. It supplies reproducible models and a failure registry for downstream FDS papers.

FDS INTERFACES AND IMPORTED MATHEMATICAL STRUCTURE

FDS-Core

The FDS-Core defines active finite systems that maintain task-relevant distinctions under finite representational, energetic, temporal, and access constraints [1]. When distinction demand exceeds available capacity, a system must approximate, prune, externalize, relax the task, or fail. The Core does not imply a particular memory kernel or spectral law. P8 is a physical and mathematical bridge that begins only after a retained description and its unresolved complement have been registered.

Relation to P3, P4, T3, and P7

P4 studies coarse-grained anti-recurrence and preimage loss under non-injective truncation [3]. P3 studies finite environmental side records and operational Markovianization [2]. T3 studies projection-induced stochasticity and Phase-B variables under capacity overflow [4]. P7 studies invariant or topological side-ledgers that preserve task identity under local forgetting [5]. P8 does not repeat those ledgers. It supplies the dynamical object that determines whether an unresolved distinction decays, returns, remains protected, or destabilizes the retained description.

TABLE II. Imported structures and P8-specific additions.

Object or result	Source status	Role in P8
Mori–Zwanzig / Nakajima–Zwanzig identity	Imported	Exact reduced equation with memory and residue forcing
Semigroup exponential stability	Imported	Sufficient condition for an integrable kernel tail
Controllability and observability	Imported	Defines the active residue quotient
Hankel operators and balanced realization	Imported	Defines round-trip significance and model order
Tauberian spectral-edge asymptotics	Imported conditionally	Converts low-frequency density into long-time tails
Finite projection, pruning, and accounting boundary	FDS input	Qualifies when a formal residue has physical meaning
Coupled residue channel (G_R, B_R, C_R)	P8 synthesis	Central transfer object
Internal spectral set, transfer singularities, recoverability singular values, and recoverable subspace	P8 type-correct synthesis	Prevents category errors between complex singularities, nonnegative singular values, and state-space directions
Operational residue rank	P8 definition	Tolerance- and window-dependent relevant residue count
Transfer–recoverability sandwich and task-risk truncation bound	Standard operator estimates; P8 channel-registration synthesis	Links round-trip dynamics, physical decoding, tolerance, noise units, and task loss

Standard mathematical neighbors

The projection identity originates in nonequilibrium statistical mechanics and open-system theory [6–9, 12]. Quantum-information approaches provide complementary operational measures of non-Markovianity [13, 14]. Optimal prediction makes the orthogonal dynamics explicit in underresolved deterministic systems [10, 11]. For orthogonal projections satisfying the additional bounded-perturbation condition that $\mathsf{P}\mathcal{L}$ extends boundedly—and in particular for the rank-one Mori projection under the registered domain hypotheses—the orthogonal dynamics is generated by a strongly continuous semigroup [30, 31]. Recent analysis further distinguishes the rigorous Mori-projection setting from unresolved issues for more general Zwanzig projections and emphasizes that the formal memory term is fundamentally a coupling term and need not represent memory in an ordinary physical sense [32]. In P8 terminology, it therefore does not by itself imply operational recoverability. Data-driven latent-space methods combine Mori–Zwanzig closure with Koopman autoencoders for nonlinear observable-space dynamics [34]. Strong lumpability characterizes when a finite Markov chain remains Markovian after aggregation [15, 16]; recent work has developed quantitative effective dynamics and error bounds for approximate coarse-graining of continuous-time Markov chains [33]. Kalman's controllability-observability decomposition and realization theory identify the minimal state that contributes to input-output behavior [17, 18]. Hankel singular values organize balanced model reduction [19–21]. Semigroup theory separates spectral information from actual norm growth [22, 23]; pseudospectral analysis is needed when non-normality produces transient amplification [24]. P8 uses these results openly rather than presenting them as newly derived FDS mathematics.

PROJECTION–PRUNING FIREWALL AND REGISTERED SYSTEM

Four operations that must not be conflated

Definition 1 (Formal projection)**.** *A formal projection is a mathematical map* P *selecting a retained observable or state sector from a larger modeled description.*

Definition 2 (Operational pruning)**.** *Operational pruning occurs when a finite system no longer actively maintains a task-relevant distinction inside its current effective record.*

Definition 3 (Physical externalization)**.** *Physical externalization occurs when a pruned distinction remains represented in an environmental carrier, side record, invariant, reversible garbage register, field correlation, or enlarged boundary ledger.*

Definition 4 (Thermodynamic erasure)**.** *Thermodynamic erasure is a physical many-to-one update assessed relative to a registered accounting boundary and*

reversible side-record structure. Formal many-to-one notation alone does not establish heat generation.

The required firewall is

$$\boxed{\begin{array}{c}\text{formal projection} \neq \text{operational pruning},\\ \text{operational pruning} \neq \text{externalization},\\ \text{externalization} \neq \text{thermodynamic erasure.}\end{array}} \tag{5}$$

This protects P8 from a common category error: a researcher's choice to ignore a variable is not itself a physical forgetting event.

Observable-space formulation

Let $\mathcal{B}$ be a Banach or Hilbert space of observables. Let $U(t)$ be a strongly continuous semigroup generated by $\mathcal{L}$,

$$U(t) = e^{t\mathcal{L}}, \qquad \mathcal{L} : \mathcal{D}(\mathcal{L}) \subset \mathcal{B} \to \mathcal{B}. \tag{6}$$

For nonlinear microscopic dynamics, $U(t)$ may be a Koopman or Liouville evolution acting linearly on observables. Let P be a bounded projection or conditional expectation, and let $\mathsf{Q} = I - \mathsf{P}$. The retained and residue spaces are

$$\mathcal{Z} = \operatorname{Ran}\mathsf{P}, \qquad \mathcal{R} = \operatorname{Ran}\mathsf{Q}. \tag{7}$$

This formulation is narrower and safer than applying a linear projection formula directly to an arbitrary nonlinear state trajectory.

Theorem scope and infinite-dimensional extensions

The main realization, operational-rank, balanced-truncation, and Gaussian determinant results below are stated for finite-dimensional systems, or for Hilbert-space systems with bounded coupling operators and the additional compactness or trace-class hypotheses explicitly registered. For the exact semigroup decomposition, assume that G_R is densely defined and closed on $\mathcal{R}$, generates a C_0-semigroup $S_R(t)$, and that $B_R \in \mathcal{L}(\mathcal{Z}, \mathcal{R})$ and $C_R \in \mathcal{L}(\mathcal{R}, \mathcal{Z})$, unless a well-posed-system extension is invoked. For orthogonal projections satisfying the additional bounded-perturbation condition that $\mathsf{P}\mathcal{L}$ extends boundedly—and in particular for the rank-one Mori projection under the registered domain hypotheses—the orthogonal dynamics is generated by a strongly continuous semigroup [30, 31].

If B_R or C_R is unbounded, it must be an admissible control or observation operator for $S_R(t)$. The reachable space is always taken with closure,

$$\mathcal{R}_{\mathrm{c}} = \overline{\operatorname{span}\{S_R(t)B_R z : t \geq 0,\ z \in \mathcal{Z}\}}, \tag{8}$$

and the unobservable sector must be closed and invariant for a quotient semigroup to be well defined. A direct-sum active/silent realization additionally requires a complemented invariant representative and is not automatic in infinite dimensions. Discrete finite-window Hankel singular values require a compact Hankel operator. The infinite-dimensional Gaussian mutual-information formula requires the whitened signal covariance to be trace class and uses a Fredholm determinant. These are extension hypotheses, not consequences of the notation.

Residue qualification hierarchy

Definition 5 (Formal residue)**.** *A formal residue is any component in $\mathcal{R}$.*

Definition 6 (Physical residue)**.** *A formal residue is physical only if it is associated with a specified carrier, correlation, mode, record, or invariant inside a registered accounting boundary.*

Definition 7 (Excitable residue)**.** *A residue direction is excitable if it lies in the reachable closure generated by $B_R = \mathsf{Q}\mathcal{L}\mathsf{P}$.*

Definition 8 (Return-active residue)**.** *A residue direction r is return-active if $C_R S_R(t) r \neq 0$ for some $t \geq 0$, where $C_R = \mathsf{P}\mathcal{L}\mathsf{Q}$.*

Definition 9 (Coupled residue)**.** *Coupled residue is both excitable and return-active. Only coupled residue participates in a retained–residue–retained dynamical round trip.*

Definition 10 (Silent residue)**.** *Silent residue exists in $\mathcal{R}$ but is uncontrollable from $\mathcal{Z}$, unobservable from $\mathcal{Z}$, dynamically decoupled, or irrelevant to the registered task.*

Definition 11 (Accessible residue)**.** *Accessible residue can be decoded through a declared finite observation channel, latency, resource budget, and error tolerance.*

Definition 12 (Protected residue)**.** *Protected residue is maintained by symmetry, topology, code structure, a conservation sector, an isolated invariant subspace, or another registered protection condition.*

A residue can be coupled but not accessible: it can affect the retained process without allowing the original past distinction to be decoded. It can also be accessible but weakly coupled to the short-time retained response, as in a stable side record read through a separate channel. This distinction becomes central in the operational-recoverability section.

EXACT COUPLED-RESIDUE DYNAMICS

Block evolution

Assume for the moment that the lifted observable dynamics admits the block form

$$\frac{d}{dt}\begin{pmatrix} z \\ r \end{pmatrix} = \begin{pmatrix} A & C_R \\ B_R & G_R \end{pmatrix}\begin{pmatrix} z \\ r \end{pmatrix}, \tag{9}$$

where

$$A = \mathsf{P}\mathcal{L}\mathsf{P}, \quad B_R = \mathsf{Q}\mathcal{L}\mathsf{P}, \quad C_R = \mathsf{P}\mathcal{L}\mathsf{Q}, \quad G_R = \mathsf{Q}\mathcal{L}\mathsf{Q}. \tag{10}$$

The residue solution is

$$r(t) = S_R(t)r_0 + \int_0^t S_R(t-s)B_R z(s)\, ds. \tag{11}$$

Substitution into the retained equation yields the exact reduced dynamics.

Theorem 1 (Exact coupled-residue decomposition)**.** *Assume that $G_R : D(G_R) \subset \mathcal{R} \to \mathcal{R}$ generates the C_0-semigroup $S_R(t)$, that $B_R \in \mathcal{L}(\mathcal{Z}, \mathcal{R})$ and $C_R \in \mathcal{L}(\mathcal{R}, \mathcal{Z})$, and that the initial residue lies in the variation-of-constants domain. Then the retained variable satisfies*

$$\dot{z}(t) = Az(t) + \int_0^t K(t-s)z(s)\, ds + \eta(t), \tag{12}$$

with

$$K(t) = C_R S_R(t) B_R, \qquad \eta(t) = C_R S_R(t) r_0. \tag{13}$$

Where $sI - G_R$ is invertible,

$$\widehat{K}(s) = C_R(sI - G_R)^{-1}B_R. \tag{14}$$

Proof. Variation of constants gives Eq. (11). Substitution into $\dot{z} = Az + C_R r$ gives Eq. (12). Laplace transformation gives Eq. (14). The general Mori–Zwanzig identity supplies the corresponding observable-space result when the dynamics is not already written in block-linear form [9, 10]. □

Remark 1 (Exactness and approximation)**.** *Equation (12) is exact for the registered lifted dynamics. Approximation begins only when the memory is truncated, η is modeled statistically, auxiliary states are removed, the projection is changed, or physical carriers are inferred from a mathematical residue.*

Discrete-time form

For a block update

$$\begin{pmatrix} z_{n+1} \\ r_{n+1} \end{pmatrix} = \begin{pmatrix} A & C \\ B & G \end{pmatrix}\begin{pmatrix} z_n \\ r_n \end{pmatrix}, \tag{15}$$

one obtains

$$z_{n+1} = Az_n + \sum_{j=0}^{n-1} CG^j B z_{n-1-j} + CG^n r_0. \tag{16}$$

This is the natural bridge to finite-state hidden Markov examples and coarse transition systems.

ACTIVE RESIDUE, SILENT QUOTIENTS, AND TRANSFER SPECTRA

Reachable and unobservable residue

Define the residue reachable closure

$$\mathcal{R}_{\mathrm{c}} = \overline{\mathrm{span}\{S_R(t)B_R z : t \geq 0,\ z \in \mathcal{Z}\}}, \tag{17}$$

and the unobservable residue subspace

$$\mathcal{N}_{\mathrm{o}} = \{r \in \mathcal{R} : C_R S_R(t) r = 0 \text{ for all } t \geq 0\}. \tag{18}$$

The canonical active object is the quotient

$$\mathcal{R}_{\mathrm{act}}^{\mathrm{quot}} = \mathcal{R}_{\mathrm{c}}/(\mathcal{R}_{\mathrm{c}} \cap \mathcal{N}_{\mathrm{o}}). \tag{19}$$

It is an abstract quotient space, not automatically a subspace of $\mathcal{R}$.

Theorem 2 (Active-residue quotient factorization)**.** *The kernel $K(t) = C_R S_R(t) B_R$ and the retained input-output map factor canonically through $\mathcal{R}_{\mathrm{act}}^{\mathrm{quot}}$. If, in addition, the quotient admits a complemented invariant representative $\mathcal{R}_{\min} \subset \mathcal{R}$ with*

$$\mathcal{R}_{\min} \simeq \mathcal{R}_{\mathrm{act}}^{\mathrm{quot}}, \qquad \mathcal{R} = \mathcal{R}_{\min} \oplus \mathcal{R}_{\mathrm{silent}}, \tag{20}$$

and

$$C_R S_R(t) P_{\mathrm{silent}} B_R = 0 \qquad \text{for all } t \geq 0, \tag{21}$$

then the equivalent realization on $\mathcal{R}_{\min}$ has the same $K(t)$, $\widehat{K}(s)$, and zero-initial-residue retained input–output trajectories as the full residue realization.

For arbitrary initial residue r_0, equality of retained trajectories additionally requires that the reduced realization be initialized with the corresponding active representative and that

$$C_R S_R(t) P_{\mathrm{silent}} r_0 = 0 \qquad \text{for all } t \geq 0.$$

Proof. By definition, $B_R z$ generates only $\mathcal{R}_{\mathrm{c}}$. Every direction in $\mathcal{R}_{\mathrm{c}} \cap \mathcal{N}_{\mathrm{o}}$ is annihilated by all future output maps $C_R S_R(t)$, so the input-output map descends to the quotient. A concrete direct-sum realization follows only after choosing the additional complemented invariant representative in Eq. (20); Eq. (21) then removes the silent contribution identically to the driven input–output map.

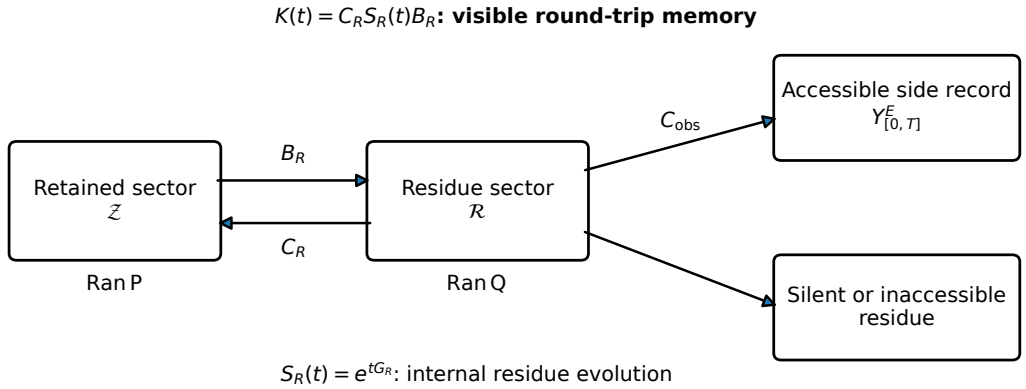

FIG. 1. The coupled residue channel. Retained distinctions excite residue through B_R; residue evolves under $S_R(t)$; return occurs through C_R. An additional observation channel may make part of the residue recoverable. Residue that is neither return-active nor observable is silent relative to the registered closure task.

For nonzero initial residue, the silent initial component also contributes through

$$\eta_{\text{silent}}(t) = C_R S_R(t) P_{\text{silent}} r_0.$$

Hence equality of arbitrary-initial-residue trajectories requires $\eta_{\text{silent}}(t) = 0$ for all $t \geq 0$, with the reduced realization initialized by the corresponding active representative. □

This is a standard minimal-realization principle expressed in residue language. Its FDS importance is conceptual: a hidden conserved or topological mode does not by itself obstruct closure. It obstructs the registered closure only when it participates in the excitation-return round trip.

Type-correct spectral, transfer, and recoverability objects

The internal, transfer, and recovery problems use mathematical objects of different types. P8 therefore separates them rather than writing them as if they were interchangeable spectra.

Definition 13 (Internal spectral set)**.** *The internal spectral set is*

$$\Lambda_{\text{int}} := \text{spec}(G_R) \subset \mathbb{C}. \tag{22}$$

Definition 14 (Transfer singularity structure)**.** *The transfer singularity structure is*

$$\mathfrak{S}_{\text{tr}} := \text{Sing}\, \widehat{K}, \qquad \widehat{K}(s) = C_R(sI - G_R)^{-1} B_R. \tag{23}$$

It includes poles, branch edges, and other nonanalytic structures that survive excitation-return filtering. In a finite-dimensional minimal rational realization, the transfer pole set is

$$\Lambda_{\text{tr}} := \text{Pole}\, \widehat{K} \subset \mathbb{C} \tag{24}$$

after pole-zero cancellations.

The transfer singularity structure is not the complete transfer response. Finite-window closure also depends on residues and modal weights, resolvent norms, cancellation, Hankel singular values, non-normal amplification, and the initial-residue forcing

$$\eta_t = C_R S_R(t) \mathsf{Q} x_0. \tag{25}$$

We use the term *full coupled reduced-response data* for the registered collection

$$\mathfrak{D}_{\text{red}} = \left\{ \widehat{K}(s),\, \mathcal{H}_{\text{tr,T}},\, \|C_R(sI - G_R)^{-1} B_R\|,\, \eta(t) \right\},$$

supplemented, where required, by modal weights, cancellation structure, the chosen norm, observation window, task, and tolerance. The initial-residue term

$$\eta(t) = C_R S_R(t) r_0$$

is a homogeneous initial-condition response and is not itself part of the retained-input–to–retained-output transfer function.

Let $\mathcal{O}_{\text{rec,T}}$ be a finite-window, noise-whitened physical recovery observation operator and let Σ_R be a registered initial-residue covariance.

Definition 15 (Initial-residue recoverability spectrum)**.** *The initial-residue recoverability singular-value spectrum is*

$$\Sigma^{\text{init}}_{\text{rec}}(T) := \left\{ \sigma_j \left(\mathcal{O}_{\text{rec,T}} \Sigma_R^{1/2} \right) \right\}_{j \geq 1} \subset \mathbb{R}_{\geq 0}. \tag{26}$$

Its squared values are the nonzero spectral values of $\Sigma_R^{1/2} W_o(T) \Sigma_R^{1/2}$ *when the Gramian is defined.*

Definition 16 (Initial-residue recoverable subspace)**.** *Let v_j be right singular vectors of $\mathcal{O}_{\mathrm{rec,T}}\Sigma_R^{1/2}$ in the whitened latent coordinate space, and write $\sigma_j^{\mathrm{init}}(T) := \sigma_j(\mathcal{O}_{\mathrm{rec,T}}\Sigma_R^{1/2})$. Their corresponding physical residue directions are $\Sigma_R^{1/2}v_j$. For tolerance ε, define*

$$\begin{aligned}\mathcal{R}_{\mathrm{rec}}^{\mathrm{init}}(T,\varepsilon) &= \overline{\mathrm{span}\left\{\Sigma_R^{1/2}v_j : \sigma_j^{\mathrm{init}}(T) > \varepsilon\right\}} \\ &\subseteq \overline{\mathrm{Ran}\,\Sigma_R^{1/2}}.\end{aligned} \tag{27}$$

When $\Sigma_R = I$, the whitened and physical residue directions coincide.

Let $\mathcal{U}_T$ be the retained input/history Hilbert space and let

$$\mathcal{C}_T : \mathcal{U}_T \longrightarrow \mathcal{R}_{\mathrm{c}} \tag{28}$$

be the finite-window retained-to-residue reachability operator. Define

$$\mathcal{H}_{\mathrm{rec,T}} = \mathcal{O}_{\mathrm{rec,T}}\mathcal{C}_T. \tag{29}$$

Definition 17 (Reachable round-trip recoverability spectrum)**.** *The reachable round-trip recoverability spectrum is*

$$\Sigma_{\mathrm{rec}}^{\mathrm{rt}}(T) := \{\sigma_j(\mathcal{H}_{\mathrm{rec,T}})\}_{j\geq 1}. \tag{30}$$

It measures recoverability only for residue histories generated from retained input histories through $\mathcal{C}_T$.

Definition 18 (Reachable round-trip input-history subspace)**.** *For tolerance ε, define*

$$\mathcal{U}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon) = \mathrm{Ran}\left[\mathbf{1}_{(\varepsilon,\infty)}\left((\mathcal{H}_{\mathrm{rec,T}}^{*}\mathcal{H}_{\mathrm{rec,T}})^{1/2}\right)\right] \subseteq \mathcal{U}_T. \tag{31}$$

These right-singular directions are retained input/history directions, not residue-state directions.

Definition 19 (Reachable round-trip residue-state subspace)**.** *The corresponding physical residue-state image is*

$$\mathcal{R}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon) = \overline{\mathcal{C}_T\mathcal{U}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon)} \subseteq \mathcal{R}_{\mathrm{c}}. \tag{32}$$

This finite-window residue-state image need not be invariant under $S_R(t)$:

$$S_R(t)\mathcal{R}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon) \not\subseteq \mathcal{R}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon)$$

in general. It therefore does not, by itself, define a dynamically realizable reduced state space or an exact Markov augmentation. Such a realization requires an invariant state-space construction, or an explicit finite-dimensional realization reproducing the registered input–output map.

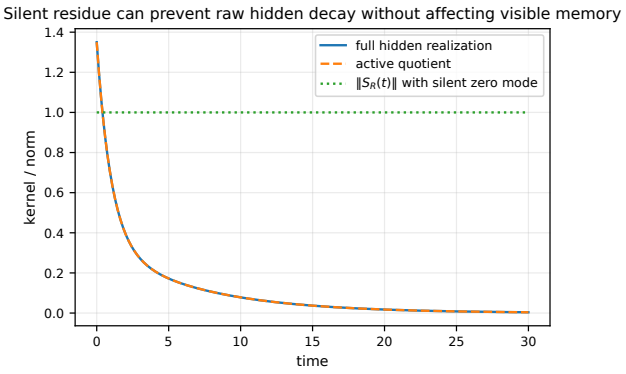

FIG. 2. Active-residue reduction. The full hidden realization contains a silent zero mode, so the raw semigroup norm remains non-decaying. The full and minimal-representative memory kernels coincide exactly. Hidden non-decay therefore does not imply visible non-Markovian memory.

The type-correct conclusion is

$$\boxed{\begin{aligned}\Lambda_{\mathrm{int}} &\subset \mathbb{C}, \\ \mathfrak{S}_{\mathrm{tr}} &= \mathrm{Sing}\,\widehat{K}, \\ \Sigma_{\mathrm{rec}}^{\mathrm{init}}(T), \Sigma_{\mathrm{rec}}^{\mathrm{rt}}(T) &\subset \mathbb{R}_{\geq 0}, \\ \mathcal{R}_{\mathrm{rec}}^{\mathrm{init}}(T,\varepsilon) &\subseteq \overline{\mathrm{Ran}\,\Sigma_R^{1/2}} \subseteq \mathcal{R}, \\ \mathcal{U}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon) &\subseteq \mathcal{U}_T, \\ \mathcal{R}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon) &\subseteq \mathcal{R}_{\mathrm{c}}.\end{aligned}} \tag{33}$$

These objects live in different mathematical spaces and do not determine one another without additional coupling, reachability, and observation assumptions. In particular, the transfer–recoverability bridge below constrains $\Sigma_{\mathrm{rec}}^{\mathrm{rt}}(T)$ and the significant input-history directions $\mathcal{U}_{\mathrm{rec}}^{\mathrm{rt}}(T,\varepsilon)$; their physical residue-state image is obtained only after application of $\mathcal{C}_T$. It does not constrain arbitrary initial-state recovery in $\Sigma_{\mathrm{rec}}^{\mathrm{init}}(T)$. Figure 2 gives the minimal counterexample to raw-spectrum reasoning. The residue generator contains a zero mode, so $\|S_R(t)\|$ does not decay. The zero mode is silent because its excitation and return weights vanish. Removing it leaves the visible kernel unchanged.

OPERATIONAL RESIDUE RANK AND MINIMAL REALIZATION

Finite-window round-trip significance

For a stable finite-dimensional linear residue channel, define the finite-window controllability Gramian and the

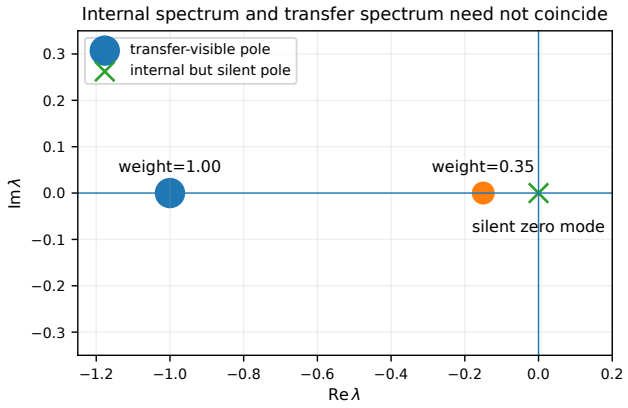

FIG. 3. Internal spectral set versus transfer-visible poles. Three internal poles are present, but only two have nonzero transfer weights. The zero pole belongs to Λ_{int} yet is absent from Λ_{tr}. Marker size indicates the absolute transfer residue in the normal form.

retained-return observability Gramian

$$W_c(T) = \int_0^T S_R(t) B_R B_R^* S_R(t)^*\, dt, \tag{34}$$

$$W_{o,\mathrm{tr}}(T) = \int_0^T S_R(t)^* C_R^* C_R S_R(t)\, dt. \tag{35}$$

Using the finite-window reachability operator $\mathcal{C}_T : \mathcal{U}_T \to \mathcal{R}_{\mathrm{c}}$ and the retained-return observation operator $\mathcal{O}_{\mathrm{tr,T}}$, define the transfer Hankel operator

$$\mathcal{H}_{\mathrm{tr,T}} = \mathcal{O}_{\mathrm{tr,T}} \mathcal{C}_T. \tag{36}$$

Its singular values are the square roots of the nonzero eigenvalues of $W_c(T) W_{o,\mathrm{tr}}(T)$ in finite dimensions. They quantify joint excitation and return, rather than hidden amplitude alone [19, 21].

Definition 20 (Operational transfer-residue rank)**.** *Given a registered normalization, time window T, and tolerance ε, define*

$$\boxed{r_{\mathrm{tr}}(\varepsilon, T) = \#\{j : \sigma_j(\mathcal{H}_{\mathrm{tr,T}}) > \varepsilon\}.} \tag{37}$$

The shorter notation r_{res} is used only when the transfer channel is unambiguous.

For the physical recovery observation operator $\mathcal{O}_{\mathrm{rec,T}}$, define the reachable round-trip recovery rank by

$$r_{\mathrm{rec}}^{\mathrm{rt}}(\varepsilon, T) = \#\{j : \sigma_j(\mathcal{H}_{\mathrm{rec,T}}) > \varepsilon\}. \tag{38}$$

This rank concerns residue histories generated through $\mathcal{C}_T$ from input histories in $\mathcal{U}_T$; it is distinct from initial-residue recoverability in Eqs. (26)–(27).

The normalization clause is essential. State coordinates can be rescaled, so the input, output, task metric, and noise units must be fixed before ε is meaningful.

Proposition 1 (Minimal-realization invariance)**.** *For finite-dimensional stable linear systems, the nonzero Hankel singular values and the associated exact Hankel rank are invariant under similarity transformations of a minimal realization. Hence the exact operational residue dimension is an input-output property, not a coordinate count.*

At finite tolerance, $r_{\mathrm{tr}}(\varepsilon, T)$ is deliberately task and window relative. A residue mode can matter on a long horizon but not on a short one. This makes the definition suitable for finite systems: it asks how many unresolved directions exceed the actual operational resolution rather than how many formal coordinates exist.

Finite-window and infinite-horizon task-risk certificates

The finite-window operational rank and the infinite-horizon balanced-truncation spectrum are related but distinct objects. P8 therefore records separate risk certificates.

Corollary 1 (Finite-window task-risk certificate)**.** *Let $\mathcal{H}_{\mathrm{tr,T}}^{(r)}$ be a best rank-r approximation to the compact finite-window transfer Hankel operator in operator norm. Suppose the registered task loss is L_{task}-Lipschitz in the finite-window retained-output norm, the residue-mediated input is $u \in \mathcal{U}_T$, and observation or decoder mismatch contributes at most $\varepsilon_{\mathrm{obs}}$. Then*

$$\boxed{\Delta\mathcal{R}_{\mathrm{task}}^T \le L_{\mathrm{task}} \sigma_{r+1}(\mathcal{H}_{\mathrm{tr,T}}) \|u\|_{\mathcal{U}_T} + \varepsilon_{\mathrm{obs}}.} \tag{39}$$

If $r = r_{\mathrm{tr}}(\varepsilon, T)$, then

$$\Delta\mathcal{R}_{\mathrm{task}}^T \le L_{\mathrm{task}} \varepsilon \|u\|_{\mathcal{U}_T} + \varepsilon_{\mathrm{obs}}. \tag{40}$$

Proof. The compact-operator extension of the Eckart–Young theorem [28, 29] gives $\|\mathcal{H}_{\mathrm{tr,T}} - \mathcal{H}_{\mathrm{tr,T}}^{(r)}\|_{\mathrm{op}} = \sigma_{r+1}(\mathcal{H}_{\mathrm{tr,T}})$. Apply this input-output bound and then the Lipschitz task-loss inequality. □

The best finite-window low-rank Hankel approximation need not automatically possess a stable, causal, finite-dimensional dynamic realization. For that stronger realizability requirement, use the following infinite-horizon certificate.

For a stable finite-dimensional minimal transfer channel, let $\widehat{K}_r$ be the order-r balanced truncation of the strictly proper residue transfer and let σ_j^∞ be the corresponding infinite-horizon Hankel singular values.

Corollary 2 (Infinite-horizon balanced-truncation task-risk certificate)**.** *Suppose the registered task loss is L_{task}-Lipschitz in the retained output norm, the residue*

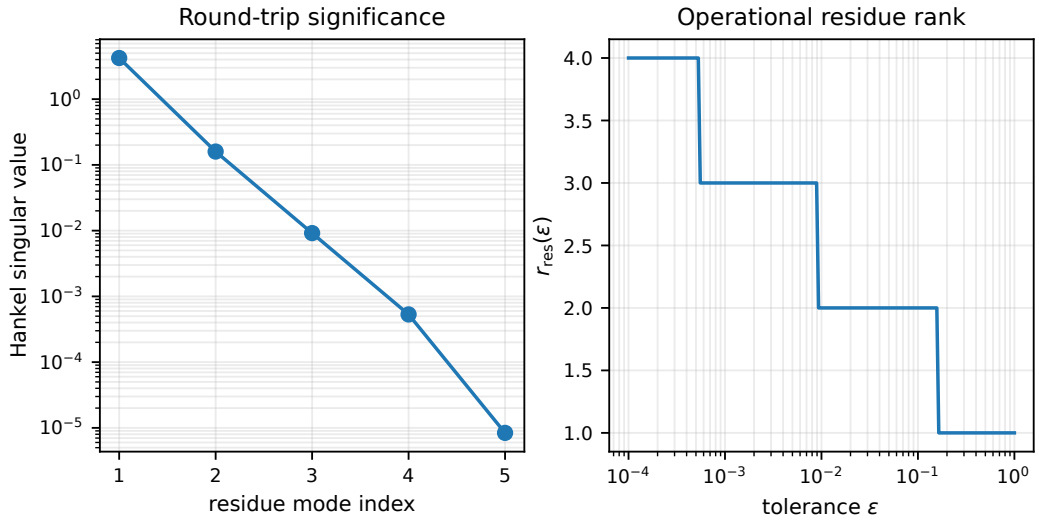

FIG. 4. Operational residue rank in a five-mode stable channel. The Hankel singular values span more than five orders of magnitude. At tolerances 10^{-1} and 10^{-2}, only two residue modes are operationally significant even though five hidden modes exist.

input is $u \in L^2$, and observation or decoder mismatch contributes at most $\varepsilon_{\mathrm{obs}}$. Then

$$\boxed{\Delta\mathcal{R}^{\infty}_{\mathrm{task}} \le 2L_{\mathrm{task}}\|u\|_{L^2}\sum_{j>r}\sigma_j^{\infty} + \varepsilon_{\mathrm{obs}}.} \tag{41}$$

Proof. The stable balanced-truncation error satisfies $\|\widehat{K} - \widehat{K}_r\|_{H^\infty} \le 2\sum_{j>r}\sigma_j^{\infty}$ [20, 21]. Apply the induced L^2 output bound and then the Lipschitz task-loss inequality. □

Equation (40) gives $r_{\mathrm{tr}}(\varepsilon, T)$ a direct finite-window decision meaning. Equation (41) is an adjacent but different infinite-horizon certificate for a realizable stable reduced model.

Finite Markov embeddings

A memory equation need not become Markovian by forgetting. It can become Markovian by restoring a finite number of state variables.

Theorem 3 (Finite realization of rational memory)**.** *For a finite-dimensional linear time-invariant reduced equation, any strictly proper rational matrix memory kernel admits a finite-dimensional exact state-space realization*

$$\dot{r} = Gr + Bz, \tag{42}$$
$$\dot{z} = A_{\mathrm{eff}}z + Cr, \tag{43}$$

whose elimination reproduces

$$\widehat{K}_{\mathrm{sp}}(s) = C(sI - G)^{-1}B. \tag{44}$$

A proper rational kernel has the form

$$\widehat{K}(s) = D + C(sI - G)^{-1}B, \tag{45}$$

where the feedthrough term D is absorbed into the instantaneous retained generator, $A_{\mathrm{eff}} = A + D$. The minimal auxiliary dimension of the strictly proper part is the McMillan degree, equivalently the rank of the exact Hankel operator under the standard realization assumptions [17, 18].

More general nonlinear or distributed-delay memory may admit exact spectral embeddings in an abstract extended space, without implying a finite-dimensional rational realization [37]. A branch cut or a genuine continuous spectral edge generally requires an infinite-dimensional exact embedding, although a finite approximation may be sufficient at a declared tolerance. Thus Markov closure has at least three distinct mechanisms:

1. memory decays below tolerance;
2. a finite active residue state is restored;
3. a non-finite memory structure is approximated on a restricted window.

The model in Fig. 5 uses a two-exponential kernel. The exact two-mode augmentation reproduces the memory equation. Removing one mode gives a maximum trajectory error of approximately 7.3×10^{-2}, while replacing the full kernel by its zero-frequency local coefficient gives a maximum error of approximately 4.7×10^{-1}. The example demonstrates why "short memory" and "finite state augmentation" are different approximations.

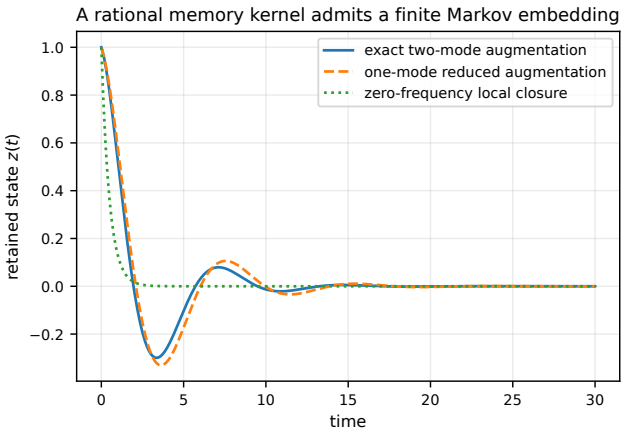

FIG. 5. Finite Markov embedding. A rational two-exponential memory kernel is represented exactly by two auxiliary residue states. A one-mode augmentation is substantially more accurate than a zero-frequency local closure for this parameter set.

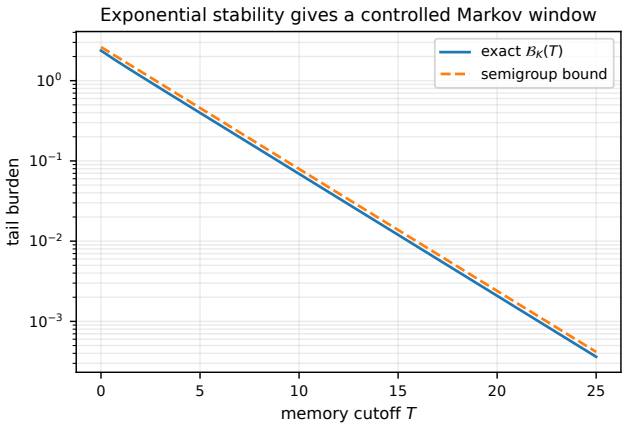

FIG. 6. Exponential closure normal form. The exact tail burden and the semigroup upper bound decay with the slow active residue rate $\omega = 0.35$. At $T = 10$, the exact burden is approximately 6.9×10^{-2}.

SPECTRAL CONDITIONS FOR MARKOV CLOSURE

Kernel burden

Define the tail burden beyond a memory cutoff T by

$$\mathcal{B}_K(T) = \int_T^\infty \|K(t)\|\, dt. \tag{46}$$

A finite tail is not by itself a complete Markov criterion, but it is a natural sufficient control quantity for memory truncation.

Theorem 4 (Exponential-stability closure bound)**.** *Assume the active residue semigroup satisfies*

$$\|S_R(t)\| \le Me^{-\omega t}, \qquad M \ge 1, \quad \omega > 0, \tag{47}$$

and B_R, C_R are bounded. Then

$$\|K(t)\| \le M\|C_R\|\|B_R\|e^{-\omega t}, \tag{48}$$

and

$$\boxed{\mathcal{B}_K(T) \le \frac{M\|C_R\|\|B_R\|}{\omega} e^{-\omega T}.} \tag{49}$$

Proof. Apply submultiplicativity to Eq. (3) and integrate the exponential tail. □

Corollary 3 (Spectral-gap corollary)**.** *If G_R is normal, self-adjoint dissipative, reversible Markov, or belongs to another registered class in which a spectral gap implies Eq. (47), then the gap gives the closure bound. A negative eigenvalue real part alone is not sufficient for a general non-normal generator.*

Finite-horizon trajectory error

Consider the truncated-memory model

$$\dot{z}_T(t) = Az_T(t) + \int_{\max(0,t-T)}^t K(t-s)z_T(s)\, ds + \eta(t). \tag{50}$$

Assume $\|z(t)\|, \|z_T(t)\| \le R$ on $[0, T_f]$ and $K \in L^1$. A direct Gronwall estimate gives

$$\sup_{0\le t\le T_f} \|z(t) - z_T(t)\| \le RT_f\mathcal{B}_K(T) \times \exp[(\|A\| + \|K\|_{L^1})\, T_f]\,. \tag{51}$$

This bound is conservative, but it converts an abstract kernel tail into a finite prediction error.

Definition 21 (Operational Markov window)**.** *A cutoff T_M is an operational Markovianization time for a registered task if*

$$\mathcal{B}_K(T_M) \le \varepsilon_K, \qquad \mathcal{M}_k \le \varepsilon_M, \qquad \Delta_{\mathrm{hist}}(k) \le \varepsilon_{\mathrm{task}}, \tag{52}$$

where $\mathcal{M}_k$ is a conditional-history closure error and Δ_{hist} is the best registered prediction or recovery advantage from extra history.

The three inequalities are not definitionally equivalent. The first is dynamical, the second statistical, and the third task operational. P8 requires them to be reported separately unless a bridge theorem is available. Memory extraction from discrete data requires controlled sampling resolution; Gaussian-process optimization can recover memory signatures from coarse observations [36].

MARGINAL MODES AND SPECTRAL-EDGE OBSTRUCTIONS

Coupled marginal modes

A raw zero mode need not matter. The relevant quantity is its transfer residue.

Theorem 5 (Coupled marginal-mode obstruction)**.** *Suppose the active residue semigroup admits*

$$S_R(t) = P_0 + S_d(t), \qquad \|S_d(t)\| \to 0, \tag{53}$$

where P_0 projects onto a zero mode. If

$$C_R P_0 B_R \neq 0, \tag{54}$$

then

$$K(t) = C_R P_0 B_R + C_R S_d(t) B_R, \tag{55}$$

so $K(t)$ does not decay to zero and $\mathcal{B}_K(0) = \infty$. If $C_R P_0 B_R = 0$, the zero mode is silent relative to the registered retained channel and does not produce this obstruction.

The same logic applies to a coupled purely imaginary mode: it produces undamped oscillatory memory rather than a constant term. Nontrivial Jordan structure at the spectral boundary can add polynomial factors.

Spectral edges

For a restricted positive spectral-measure class, write a scalar or commuting positive kernel as

$$K(t) = \int_0^\infty e^{-\gamma t}\, d\mu_K(\gamma). \tag{56}$$

Theorem 6 (Spectral-edge tail)**.** *Assume $d\mu_K(\gamma) = \rho_K(\gamma) d\gamma$ near $\gamma = 0$, with*

$$\rho_K(\gamma) \sim c\gamma^\alpha, \qquad \alpha > -1, \tag{57}$$

and the registered regular-variation and Tauberian conditions hold. Then

$$K(t) \sim c\Gamma(\alpha+1) t^{-(\alpha+1)}. \tag{58}$$

The tail burden is finite only if $\alpha > 0$.

The theorem is not asserted for arbitrary matrix-valued, sign-changing, non-normal, or time-dependent kernels. In those cases cancellation and pseudospectral effects can invalidate a direct density-to-tail inference [25, 26].

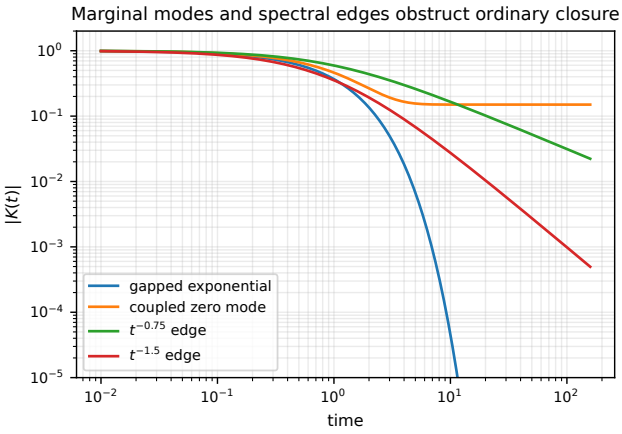

FIG. 7. Four closure classes. A gapped kernel decays rapidly; a coupled zero mode leaves a nonzero plateau; a $t^{-0.75}$ edge has non-integrable tail burden; a $t^{-1.5}$ edge is integrable but remains long lived.

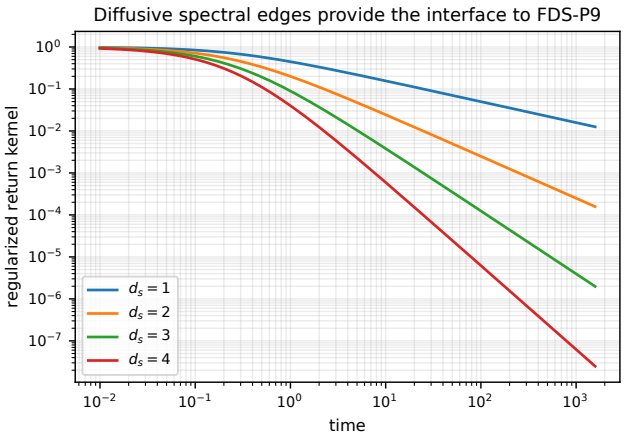

FIG. 8. Regularized diffusive return kernels. The plot is an interface to FDS-P9, not a dimensional derivation inside P8. The long-time exponent is $d_s/2$.

Diffusive return edge

A regularized local diffusion return kernel has the normal form

$$K_{d_s}(t) \propto (1 + 4Dt)^{-d_s/2}. \tag{59}$$

P8 uses this only as a spectral-edge example. The dimension-selection claim is deferred to FDS-P9. Figure 8 shows the transition from non-integrable tails for $d_s \leq 2$ to integrable tails for $d_s > 2$ in this registered normal form.

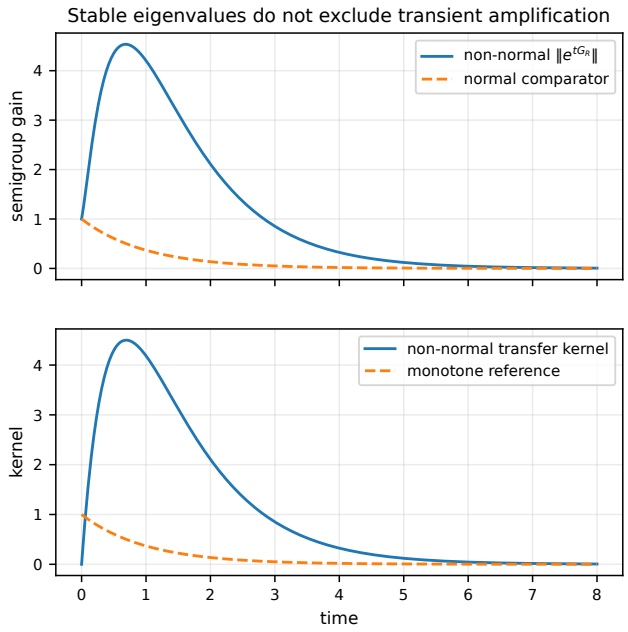

FIG. 9. Non-normal transient amplification. Stable eigenvalues coexist with a large temporary semigroup gain and a delayed transfer-kernel maximum. A raw spectral-gap argument would miss this closure hazard.

NON-NORMAL RESIDUE AND TRANSIENT AMPLIFICATION

A stable eigenvalue spectrum does not guarantee monotone memory decay. If G_R is non-normal, $\|S_R(t)\|$ can grow transiently even when all eigenvalues lie strictly in the left half-plane. The transfer channel can amplify this growth further or cancel it.

Criterion 1 (Semigroup, not eigenvalues alone)**.** *A P8 closure claim based on a "spectral gap" must either establish an actual semigroup bound such as Eq. (47) or restrict the generator to a class where the bound follows. Eigenvalues alone are insufficient in a non-normal system.*

Useful diagnostics include the numerical abscissa, resolvent norm, Kreiss constant, and pseudospectrum [24]. The normal form in Fig. 9 uses

$$G_R = \begin{pmatrix} -1 & 18 \\ 0 & -2 \end{pmatrix}. \tag{60}$$

Its eigenvalues are -1 and -2, yet the semigroup norm reaches approximately 4.53 near $t = 0.68$. The associated transfer kernel first grows and then decays. This produces a delayed memory burst without any unstable eigenvalue.

OPERATIONAL RECOVERABILITY

Task-relative recovery

Let

$$V = f(X_0) \tag{61}$$

be a task-relevant past distinction. Let the decoder receive a finite accessible record

$$\mathcal{E}_T^{\mathrm{acc}} = \left(Z_{[0,T]}, Y_{[0,T]}^E, R_T^{\mathrm{side}}\right), \tag{62}$$

where the environmental readout, side record, latency, and resource budget are registered. Define the optimal recovery error

$$\varepsilon_{\mathrm{rec}}(T) = \inf_{\mathcal{D}} \mathbb{E}[d(V, \mathcal{D}(\mathcal{E}_T^{\mathrm{acc}}))]\,. \tag{63}$$

For predictive closure, define the history advantage

$$\Delta_{\mathrm{hist}}(k) = \mathcal{L}^*(Z_t) - \mathcal{L}^*(Z_{t-k:t}), \tag{64}$$

and the conditional closure error

$$\mathcal{M}_k = \mathrm{I}(Z_{t+1}; Z_{t-k:t-1} \mid Z_t). \tag{65}$$

The important non-equivalence is

$$K(t) \neq 0 \not\Rightarrow V \text{ is recoverable.} \tag{66}$$

A residue can affect future dynamics through a compressed scalar while the original high-dimensional distinction remains unrecoverable. Conversely, an external protected record can preserve V while exerting negligible short-time force on Z.

Linear-Gaussian recoverability identity

Consider residue initial state $r_0 \sim \mathcal{N}(0, \Sigma_R)$, stable residue dynamics $r(t) = e^{G_R t} r_0$, and continuous observation

$$dy_t = C_{\mathrm{obs}} e^{G_R t} r_0\, dt + R_\nu^{1/2} dW_t. \tag{67}$$

Define the finite-time observability Gramian

$$W_o(T) = \int_0^T e^{G_R^* t} C_{\mathrm{obs}}^* R_\nu^{-1} C_{\mathrm{obs}} e^{G_R t}\, dt. \tag{68}$$

Proposition 2 (Finite-dimensional linear-Gaussian recoverability identity)**.** *For finite-dimensional Eq. (67),*

$$\boxed{\mathrm{I}(r_0; y_{[0,T]}) = \frac{1}{2} \log\det\left[I + \Sigma_R^{1/2} W_o(T) \Sigma_R^{1/2}\right].} \tag{69}$$

Proof. The observation path is a Gaussian linear channel from r_0 to a Hilbert-space output with white Gaussian noise. Whitening the noise and diagonalizing the finite-dimensional signal operator gives the standard Gaussian mutual-information determinant [27]. A finite-time discretization converges to Eq. (69). □

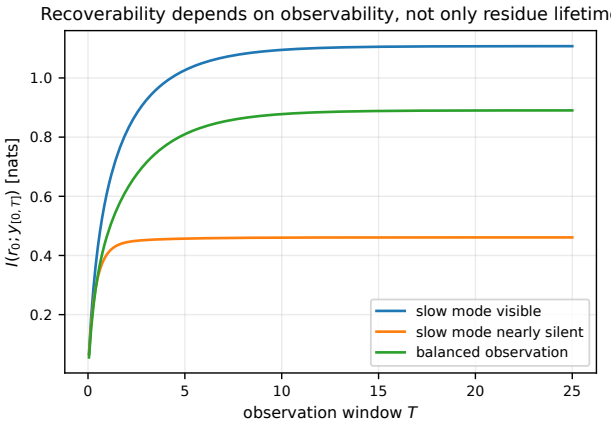

FIG. 10. Linear-Gaussian recoverability. All curves share the same residue decay rates. The difference is observation geometry. Long-lived residue does not guarantee recoverable residue.

The result shows why recoverability cannot be inferred from residue lifetime alone. A slowly decaying mode that is nearly orthogonal to C_{obs} can remain dynamically present but poorly decodable. In Fig. 10, the same two-mode residue dynamics gives long-window mutual information of approximately 1.11 nats when the slow mode is visible, but only 0.46 nats when that slow mode is nearly silent.

Transfer–recoverability bridge

The memory kernel uses the retained return channel C_R, while physical recovery may use a different observation channel C_{obs}. Identifying visible memory with recoverable inverse information therefore requires a registered relation between the two finite-window observation maps.

Let

$$\mathcal{H}_{\mathrm{tr,T}} = \mathcal{O}_{\mathrm{tr,T}}\mathcal{C}_T, \qquad \mathcal{H}_{\mathrm{rec,T}} = \mathcal{O}_{\mathrm{rec,T}}\mathcal{C}_T. \tag{70}$$

Define the closed finite-window transfer-signal space

$$\mathcal{Y}^{\mathrm{sig}}_{\mathrm{tr,T}} := \overline{\mathrm{Ran}\,\mathcal{H}_{\mathrm{tr,T}}} \subseteq \mathcal{Y}_{\mathrm{tr,T}}.$$

Theorem 7 (Transfer–recoverability singular-value sandwich)**.** *Let*

$$\mathcal{H}_{\mathrm{tr,T}} : \mathcal{U}_T \to \mathcal{Y}_{\mathrm{tr,T}}$$

be compact. Suppose that there exists a bounded operator

$$J_T : \mathcal{Y}^{\mathrm{sig}}_{\mathrm{tr,T}} \longrightarrow \mathcal{Y}_{\mathrm{rec,T}}$$

such that

$$\mathcal{H}_{\mathrm{rec,T}} = J_T\mathcal{H}_{\mathrm{tr,T}}, \tag{71}$$

and

$$m_T\|y\| \le \|J_T y\| \le M_T\|y\|, \qquad y \in \mathcal{Y}^{\mathrm{sig}}_{\mathrm{tr,T}}, \qquad 0 < m_T \le M_T < \infty. \tag{72}$$

Then, for every singular-value index j,

$$\boxed{m_T\sigma_j(\mathcal{H}_{\mathrm{tr,T}}) \le \sigma_j(\mathcal{H}_{\mathrm{rec,T}}) \le M_T\sigma_j(\mathcal{H}_{\mathrm{tr,T}}).} \tag{73}$$

Consequently,

$$\boxed{r_{\mathrm{tr}}(\varepsilon/m_T, T) \le r^{\mathrm{rt}}_{\mathrm{rec}}(\varepsilon, T) \le r_{\mathrm{tr}}(\varepsilon/M_T, T).} \tag{74}$$

Proof. The upper inequality follows from $\|J_T\| \le M_T$ and the singular-value ideal property. The lower inequality follows from the min–max characterization of compact-operator singular values together with

$$\|J_T\mathcal{H}_{\mathrm{tr,T}}u\| \ge m_T\|\mathcal{H}_{\mathrm{tr,T}}u\|$$

for every $u \in \mathcal{U}_T$. The rank inequalities follow by thresholding Eq. (73). □

The compact-operator inequality is standard. The P8 contribution is the physical separation of the retained return channel and the recovery channel, together with the finite-boundary registration and task-relative interpretation of J_T, m_T, M_T. The theorem constrains reachable round-trip recoverability $\Sigma^{\mathrm{rt}}_{\mathrm{rec}}(T)$ and the associated input-history significance in $\mathcal{U}_T$, not arbitrary initial-residue recoverability $\Sigma^{\mathrm{init}}_{\mathrm{rec}}(T)$. The corresponding residue-state directions arise only through the reachability image $\mathcal{C}_T\mathcal{U}^{\mathrm{rt}}_{\mathrm{rec}}(T, \varepsilon)$.

Corollary 4 (Finite-dimensional latent Gaussian information sandwich)**.** *Let $\xi \sim \mathcal{N}(0, I_d)$ be a finite-dimensional latent excitation and let $E_T : \mathbb{R}^d \to \mathcal{U}_T$ be bounded. Define finite-dimensional whitened readout matrices*

$$H_{\mathrm{tr}} = P_{\mathrm{tr}}\mathcal{H}_{\mathrm{tr,T}}E_T, \qquad H_{\mathrm{rec}} = P_{\mathrm{rec}}\mathcal{H}_{\mathrm{rec,T}}E_T, \tag{75}$$

with readout maps chosen so that the same bounds m_T, M_T apply on the signal range. For

$$y_{\mathrm{rec}} = H_{\mathrm{rec}}\xi + \nu, \qquad \nu \sim \mathcal{N}(0, I), \tag{76}$$

independent of ξ,

$$\begin{aligned} \mathrm{I}(\xi; y_{\mathrm{rec}}) &\ge \frac{1}{2}\sum_j \log\left[1 + m_T^2\sigma_j(H_{\mathrm{tr}})^2\right], \\ \mathrm{I}(\xi; y_{\mathrm{rec}}) &\le \frac{1}{2}\sum_j \log\left[1 + M_T^2\sigma_j(H_{\mathrm{tr}})^2\right]. \end{aligned} \tag{77}$$

Remark 2 (Trace-class extension)**.** *For an infinite-dimensional Gaussian excitation with trace-class covariance Q, replace $\sigma_j(H)$ by $\sigma_j(HQ^{1/2})$. The information formula uses a Fredholm determinant and requires the resulting signal covariance to be trace class. No identity-covariance Gaussian random element on an infinite-dimensional path space is assumed.*

The lower constant m_T is the crucial physical condition. If $m_T = 0$, a dynamically significant transfer direction can be invisible to the recovery apparatus. If recovery uses an independent side record not factored through the retained return output, Eq. (71) fails and no transfer-to-recovery lower bound is claimed. Changing the observation boundary can therefore change recoverability without changing the retained memory kernel.

FINITE-STATE PROJECTION AND LUMPABILITY

Let X_t be a finite Markov chain with transition matrix P, and let $Z_t = \pi(X_t)$ be a coarse partition. Strong lumpability requires that for every pair of coarse cells C_α, C_β,

$$\sum_{j \in C_\beta} P_{ij} \tag{78}$$

is independent of the choice of $i \in C_\alpha$ [15]. When this condition holds, Z_t is Markov for every initial distribution. When it fails, hidden within-cell state affects the next coarse transition and generates history dependence.

Proposition 3 (Finite-state closure diagnostic)**.** *For a stationary projected process, exact first-order Markov closure implies*

$$\mathrm{I}(Z_{t+1}; Z_{t-1} \mid Z_t) = 0. \tag{79}$$

A positive value is therefore a direct witness of first-order closure failure, although a zero value at one lag does not by itself prove strong lumpability.

The four-state benchmark uses

$$P(\delta) = \begin{pmatrix} 0.60 & 0.20-\delta & 0.15+\delta & 0.05 \\ 0.20+\delta & 0.60 & 0.05 & 0.15-\delta \\ 0.15+\delta & 0.05 & 0.60 & 0.20-\delta \\ 0.05 & 0.15-\delta & 0.20+\delta & 0.60 \end{pmatrix}, \tag{80}$$

with $0 \le \delta \le 0.14$ and coarse partition

$$\pi(1) = \pi(2) = A, \qquad \pi(3) = \pi(4) = B. \tag{81}$$

Its stationary distribution is

$$\varpi_\delta = \left(\frac{1}{4}+\delta, \frac{1}{4}-\delta, \frac{1}{4}+\delta, \frac{1}{4}-\delta\right). \tag{82}$$

The chain is strongly lumpable at $\delta = 0$. Direct symbolic expansion of its stationary three-time distribution gives

$$\mathrm{I}(Z_{t+1}; Z_{t-1} \mid Z_t) = \frac{225}{32}\delta^4 - \frac{310125}{512}\delta^6 + O(\delta^8). \tag{83}$$

The leading onset is fourth order rather than second order because the symmetric perturbation cancels the lower-order contributions. The absolute values are therefore small in the weakly perturbed regime, but they are analytically controlled rather than numerical noise.

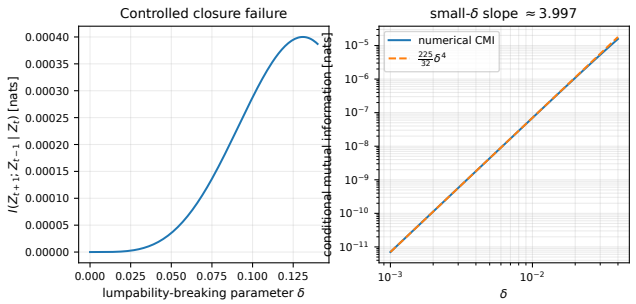

FIG. 11. Finite hidden-state closure test. At $\delta = 0$ the partition is strongly lumpable. The linear-scale panel shows the controlled departure from closure; the log–log panel verifies the analytic small-δ onset $I \sim (225/32)\delta^4$.

RESIDUE TRANSFER TAXONOMY

Table III organizes residue by visible transfer behavior rather than by raw hidden-state labels.

The following categories are diagnostic and overlapping rather than mutually exclusive or collectively exhaustive. They refer to different structural axes, including asymptotic decay, transient amplification, coupling visibility, analytic transfer structure, and physical protection. A single residue channel may therefore belong to several categories simultaneously.

This taxonomy is not a claim that each row corresponds to one universal physical phenomenon. It is a classification of reduced dynamical behavior. The physical interpretation requires a carrier and domain bridge.

REPRODUCIBLE NORMAL-FORM MODELS

The accompanying script `FDS_P8_residue_spectral_models.py` generates all figures and CSV files. The models are diagnostics of theorem conditions, not microscopic claims.

Model A: active versus silent residue

The first model uses three residue rates $(1, 0.15, 0)$, with the zero mode assigned zero excitation and return weight. The full and active-only kernels agree to machine precision, while the raw residue norm remains nondecaying. This simultaneously tests Theorem 2 and the type separation in Eq. (33).

Model B: stable, marginal, edge, and non-normal sectors

The second family compares exponential decay, a coupled zero mode, integrable and non-integrable power

TABLE III. Diagnostic categories of residue transfer behavior.

Diagnostic category	**Kernel / realization signature**	**Effective manifestation**
Strongly stable rational transfer	Finite sum of decaying exponentials	Finite exact Markov augmentation
Fast stable transfer	Short integrable kernel	Local parameter renormalization
Slow stable transfer	Finite colored memory	Delay, friction, viscoelastic response
Coupled marginal pole	Constant or undamped oscillatory term	Closure obstruction, recurrence, echo
Low-frequency branch edge	Algebraic tail	Long memory, aging, anomalous transport
Protected isolated sector	Persistent observable or recoverable mode	Stable side-ledger or logical distinction
Non-normal transfer	Transient gain despite stable eigenvalues	Bursts, delayed recovery, temporary closure failure
Unstable transfer	Growing kernel or forcing	Runaway residue and boundary failure
Silent internal sector	No transfer contribution	Operationally irrelevant to current closure

laws, and a stable non-normal generator. It demonstrates that a closure classification must include coupling, tail integrability, and transient gain.

Model C: operational residue rank

A five-mode stable channel has Hankel singular values

$$4.224,\ 0.159,\ 9.17\times10^{-3},\ 5.32\times10^{-4},\ 8.39\times10^{-6}. \quad (84)$$

At both $\varepsilon = 10^{-1}$ and 10^{-2}, the operational residue rank is two. The hidden dimension is five, but only two modes exceed the round-trip tolerance.

Model D: recoverability, embedding, and finite-state closure

The linear-Gaussian model tests Eq. (69); the rational-kernel model tests finite augmentation; the hidden-state model tests strong lumpability and conditional history information. Together they separate three questions that are often conflated:

1. Does residue affect future retained dynamics?
2. Can the residue be decoded as a past distinction?
3. Can a finite number of auxiliary variables make the process Markovian?

REGISTRATION PROTOCOL

Protocol 1 (Required residue registration)**.** *Every physical application of P8 must specify:*

1. *the full modeled system and the retained observable sector;*
2. *the projection* P *and residue complement* Q*;*
3. *the excitation operator* B_R *and return operator* C_R*;*
4. *the physical carrier of the claimed residue;*
5. *the observation channel used for recovery;*
6. *the task-relevant past distinction;*
7. *the decoder, latency, time window, and error tolerance;*
8. *the resource and accounting boundary;*
9. *the semigroup, transfer, or finite-state assumptions used for closure;*
10. *the failure condition and the nearest alternative model.*

Recommended diagnostics include kernel reconstruction, transfer-function estimation, Hankel singular values, conditional mutual information, history advantage, observability Gramians, recurrence tests, and pseudospectral gain. Memory-kernel reconstruction may use regularized Prony or RKHS methods with rigorous error bounds [35]; accurate extraction from discrete time series requires controlled sampling resolution [36]. A strong intervention changes one structural component while holding ordinary fit quality fixed: expand the retained state, randomize a side record, break a protection condition, vary bath size, or alter only the return coupling.

FAILURE AND DEMOTION REGISTRY

Criterion 2 (Formal failure)**.** *The exact identity is withdrawn for a model if the declared generator, projection, or domain assumptions fail.*

Criterion 3 (Projection–pruning failure)**.** *A mathematical unresolved variable must not be called a physical residue if no physical carrier, coupling, record, or measurable delayed effect is identified.*

Criterion 4 (Active-residue failure)**.** *A mode is demoted to silent residue if it cannot be excited from, or return to, the registered retained sector.*

Criterion 5 (Semigroup failure)**.** *The exponential closure theorem is inapplicable if only eigenvalue locations are known and non-normal transient growth is uncontrolled.*

Criterion 6 (Spectral-edge failure)**.** *A Tauberian tail claim is demoted if the kernel lacks the registered positivity, regular variation, or no-cancellation conditions.*

Criterion 7 (Recoverability failure)**.** *Visible memory does not prove recovery of the original distinction. Recovery claims must be withdrawn if decoding fails under the registered channel and resources.*

Criterion 8 (Boundary failure)**.** *If an accessible side record restores the distinction after the accounting boundary is correctly enlarged, the case is externalization, not forgetting.*

Criterion 9 (Finite-realization failure)**.** *A proposed finite Markov augmentation is rejected if it does not reproduce the registered transfer function or if its claimed minimal order is not minimal.*

Criterion 10 (Universality failure)**.** *Failure of a physical interpretation in one domain does not invalidate the exact residue decomposition. It demotes only the corresponding domain bridge.*

DOWNSTREAM INTERFACES

FDS-T4: macroscopic law selection

P8 provides active residue, operational residue rank, coupled marginal-mode obstruction, and finite Markov embeddings. T4 asks which residue modes must be promoted into explicit macroscopic state variables and under what conditions short-memory residue yields local low-order equations. P8 does not claim that every coupled zero mode is a fundamental conservation law.

FDS-P9: dimensional selection

P8 provides diffusive return kernels and the integrability threshold of a registered spectral edge. P9 must add locality, unbiased propagation, geometric interpretation, maintenance cost, and the bridge from spectral to macroscopic spatial dimension. P8 therefore does not claim to derive three-dimensional space.

FDS-O4: macroscopic time arrows

P8 supplies side-record decay, operational recovery loss, recurrence, and semigroup relaxation. O4 must explain arrow alignment and distinguish local backflow from recovery of a complete macroscopic history. P8 does not solve the Past Hypothesis.

FDS-P10: friction, noise, and inertial renormalization

The low-frequency expansion of $\widehat{K}(i\omega)$ separates dissipative and reactive response. P10 can use this to study friction, colored noise, added mass, and effective inertia. P8 does not derive fundamental inertial mass or the equivalence principle.

FDS-H3 and G1

P8 supplies a disciplined language for long-lived screen residue and memory channels. H3 must independently derive the physical optical quotient, area/shear/twist response, equal-compliance or free-compliance branches, and Weyl/Ricci channel structure. P8 does not derive the G1 3/4 coefficient.

DISCUSSION

What the unresolved dimension does not tell us

A large hidden state space can be operationally silent. A single strongly coupled slow mode can dominate memory. Therefore model complexity should not be measured by hidden dimension alone. Transfer rank, observation geometry, and time window are more relevant to a finite retained system.

Why the full coupled reduced-response data improve residue reasoning

The transfer singularity structure registers nonanalytic structures that survive excitation-return filtering, but it does not by itself classify finite-window closure. A complete audit also requires modal weights, finite-window Hankel singular values, resolvent and pseudospectral gain, cancellation, and initial-residue forcing. These full coupled reduced-response data prevent three overclaims:

1. a hidden zero mode does not automatically imply macroscopic conservation;

2. a long residue lifetime does not automatically imply recoverability;

3. a stable eigenvalue spectrum does not automatically imply a safe Markov approximation.

Markov closure as state selection

An effective process can become Markovian because residue decays, because coupling becomes negligible, because the observation task becomes insensitive, or because a finite auxiliary state is restored. These are physically different closures even when they produce similar one-step likelihoods. P8 therefore treats Markovianity as boundary, task, and state-selection relative, not as an absolute property of a trajectory alone.

The FDS explanatory increment

The strongest FDS addition is not a new integral equation. It is an audit question applied to every reduction:

> What distinction left the active record, where is it physically carried, which part can return, which part can be decoded, and what finite expansion would be required to absorb its memory?

This question turns generic "unresolved variables" into a registered physical program with local failure conditions.

CONCLUSION

Finite effective systems do not interact with all hidden structure equally. Only residue that is both excitable and return-active contributes to visible round-trip memory. The correct central object is therefore

$$\boxed{\mathfrak{R} = (G_R, B_R, C_R), \qquad \widehat{K}(s) = C_R(sI - G_R)^{-1}B_R.} \tag{85}$$

The internal spectral set, transfer singularity structure, initial-residue recoverability spectrum, reachable round-trip recoverability spectrum, significant input-history subspace, and physical residue-state images are different mathematical objects and do not determine one another without additional assumptions. Silent residue can be quotiented out without changing retained dynamics. Exponential semigroup stability supplies a controlled finite-memory bound. Coupled marginal modes and low-frequency spectral edges obstruct ordinary closure. Nonnormality can create transient memory bursts despite stable eigenvalues. Operational recoverability depends on a separate observation channel. Rational kernels permit finite exact state augmentation, while finite-window Hankel singular values define operational transfer and reachable round-trip recovery ranks. Operational closure is governed by the full coupled reduced-response data, not by the transfer singularity set alone.

The compressed P8 thesis is

$$\boxed{\begin{array}{c}\text{Projection determines what leaves the active record;}\\ \text{the coupled residue channel determines what can return.}\end{array}} \tag{86}$$

P8 supplies a common spine for later FDS work on macroscopic variables, local laws, spatial dimension, time arrows, inertia, and optical gravity. Those downstream claims remain conditional on their own bridge assumptions and tests.

Notation Summary

Proof Details for the Exact Decomposition

Let $x = z + r$ with $z = \mathsf{P}x$ and $r = \mathsf{Q}x$. The lifted equation $\dot{x} = \mathcal{L}x$ gives

$$\dot{z} = Az + C_R r, \tag{87}$$
$$\dot{r} = B_R z + G_R r. \tag{88}$$

Assuming G_R generates $S_R(t)$, variation of constants gives

$$r(t) = S_R(t)r_0 + \int_0^t S_R(t-s)B_R z(s)ds. \tag{89}$$

Substitution proves Theorem 1. In a general Mori–Zwanzig derivation, the orthogonal evolution may be written with $e^{t\mathsf{Q}\mathcal{L}}$ on an appropriate invariant domain. The variation-of-constants derivation of the GLE is justified when $\mathsf{Q}\mathcal{L}$ generates a strongly continuous semigroup; the addendum to [30] further clarifies the relationship between $e^{t\mathsf{Q}\mathcal{L}}$ and the subspace semigroup generated by $\mathsf{Q}\mathcal{L}\mathsf{Q}$ [31]. P8 uses the restricted residue generator only when the Q-sector evolution is well defined. This assumption must be checked rather than inferred from notation.

Proof Details for Active-Residue Reduction

For a finite observation window $T > 0$, define the reachability operator

$$\mathcal{C}_T : \mathcal{U}_T = L^2([0,T], \mathcal{Z}) \longrightarrow \mathcal{R}_{\mathrm{c}},$$
$$\mathcal{C}_T u = \int_0^T S_R(T-s)B_R u(s)\, ds. \tag{90}$$

The corresponding transfer-observation operator is

$$\mathcal{O}_{\mathrm{tr},T} : \mathcal{R}_{\mathrm{c}} \longrightarrow \mathcal{Y}_{\mathrm{tr},T}, \tag{91}$$

TABLE IV. Core notation used in P8.

Symbol	Meaning
$\mathcal{B}$	observable Banach or Hilbert space
$\mathcal{L}$	full lifted generator
P, Q	retained and residue projections
$\mathcal{Z}, \mathcal{R}$	retained and residue spaces
A	retained instantaneous generator
G_R	internal residue generator
B_R	retained-to-residue excitation map
C_R	residue-to-retained return map
$S_R(t)$	residue semigroup
$K(t), \widehat{K}(s)$	visible memory kernel and transfer function
$\mathcal{R}_{\mathrm{act}}^{\mathrm{quot}}$	canonical active residue quotient
$\mathcal{R}_{\mathrm{min}}$	complemented minimal representative, when it exists
$\mathcal{R}_{\mathrm{silent}}$	silent residue sector
$\Lambda_{\mathrm{int}}, \mathfrak{S}_{\mathrm{tr}}$	internal spectral set and transfer singularity structure
$\Sigma_{\mathrm{rec}}^{\mathrm{init}}(T), \Sigma_{\mathrm{rec}}^{\mathrm{rt}}(T)$	initial-residue and reachable round-trip recoverability spectra
$\mathcal{R}_{\mathrm{rec}}^{\mathrm{init}}(T, \varepsilon)$	physical initial-residue recoverable subspace
$\mathcal{U}_T, \mathcal{C}_T$	retained input/history space and finite-window reachability operator
$\mathcal{U}_{\mathrm{rec}}^{\mathrm{rt}}(T, \varepsilon)$	significant reachable round-trip input/history subspace
$\mathcal{R}_{\mathrm{rec}}^{\mathrm{rt}}(T, \varepsilon)$	corresponding reachable residue-state image
$\mathcal{B}_K(T)$	memory tail burden
$\mathcal{H}_{\mathrm{tr,T}}, \mathcal{H}_{\mathrm{rec,T}}$	finite-window transfer and recovery Hankel operators
$r_{\mathrm{tr}}(\varepsilon, T), r_{\mathrm{rec}}^{\mathrm{rt}}(\varepsilon, T)$	finite-window transfer and reachable round-trip recovery ranks
W_c, W_o	controllability and observability Gramians

with action

$$(\mathcal{O}_{\mathrm{tr},T} r)(t) = C_R S_R(t) r, \qquad 0 \le t \le T.$$

The finite-window transfer Hankel operator is

$$\mathcal{H}_{\mathrm{tr,T}} = \mathcal{O}_{\mathrm{tr},T} \mathcal{C}_T. \tag{92}$$

Any residue direction in $\operatorname{Ker} \mathcal{O}_{\mathrm{tr},T}$ is output silent; any direction outside the reachable closure is input silent. The input-output memory factors through the quotient of the reachable space by its intersection with $\operatorname{Ker} \mathcal{O}_{\mathrm{tr},T}$. In finite dimensions this quotient admits the standard controllable-observable minimal representative from the Kalman decomposition [17]. In infinite dimensions, existence of a complemented invariant representative is an additional assumption rather than an automatic consequence.

Trajectory Error for Memory Truncation

Let $e(t) = z(t) - z_T(t)$, define $a_T(s) = \max(0, s - T)$, and write $\kappa(s, u) = \|K(s - u)\|$. Subtracting the full and truncated equations gives

$$\|e(t)\| \le \int_0^t \|A\| \|e(s)\| ds \tag{93}$$

$$+ \int_0^t \int_{a_T(s)}^s \kappa(s, u) \|e(u)\| du\, ds \tag{94}$$

$$+ R \int_0^t \int_0^{a_T(s)} \kappa(s, u) du\, ds. \tag{95}$$

The last term is bounded by $Rt\mathcal{B}_K(T)$. Bounding the retained-memory convolution by $\|K\|_{L^1} \int_0^t \|e(u)\| du$ and applying Gronwall gives Eq. (51). Sharper bounds are possible for contractive or positive systems; P8 uses the conservative form because it makes the assumptions transparent.

Marginal Modes and Tauberian Conditions

For Theorem 5, strong convergence $S_d(t) \to 0$ implies

$$K(t) \to C_R P_0 B_R. \tag{96}$$

If the limit is nonzero, then $\|K(t)\|$ is bounded below by a positive constant for sufficiently large t, so its integral diverges.

For Theorem 6, Karamata-type Laplace Tauberian results convert regular variation of the positive density near zero into regular variation of the long-time kernel [25]. Positivity or a no-cancellation assumption is necessary. Matrix-valued kernels should be tested through scalar quadratic forms or an operator-valued extension with explicit positivity.

Linear-Gaussian Recovery Derivation

Whiten Eq. (67) by $R_\nu^{-1/2}$. The resulting observation is a Gaussian channel

$$y = \mathcal{A}_T r_0 + w \tag{97}$$

in the path Hilbert space, where

$$(\mathcal{A}_T r)(t) = R_\nu^{-1/2} C_{\mathrm{obs}} e^{G_R t} r. \tag{98}$$

The signal Gramian is

$$\mathcal{A}_T^* \mathcal{A}_T = W_o(T). \tag{99}$$

For Gaussian input covariance Σ_R and unit white noise, the mutual information is

$$\frac{1}{2} \log\det(I + \Sigma_R^{1/2} \mathcal{A}_T^* \mathcal{A}_T \Sigma_R^{1/2}), \tag{100}$$

which gives Eq. (69). The same formula follows from finite time discretization followed by convergence of the Gramian. In an infinite-dimensional extension, the determinant must be a Fredholm determinant and $\Sigma_R^{1/2} W_o(T) \Sigma_R^{1/2}$ must be trace class.

Transfer–Recoverability and Task-Risk Details

For Theorem 7, write $H_{\mathrm{tr}} = \mathcal{H}_{\mathrm{tr,T}}$ and $H_{\mathrm{rec}} = J_T H_{\mathrm{tr}}$. Compactness supplies a discrete singular-value sequence. The singular-value ideal inequality gives [29]

$$\sigma_j(H_{\mathrm{rec}}) \le \|J_T\| \sigma_j(H_{\mathrm{tr}}) \le M_T \sigma_j(H_{\mathrm{tr}}). \tag{101}$$

Because J_T is bounded below by m_T on the closed transfer-signal space $\mathcal{Y}^{\mathrm{sig}}_{\mathrm{tr,T}}$, the same min–max argument yields the lower singular-value bound. The Gaussian corollary is deliberately finite-dimensional at the latent and readout levels; whitening converts each signal singular direction into an independent scalar Gaussian channel, so mutual information is the sum of $\frac{1}{2}\log(1+\sigma_j^2)$.

For Corollary 1, the Schmidt–Eckart–Young approximation theorem [28, 29] supplies the optimal rank-r finite-window operator error. This approximation need not be dynamically realizable. Corollary 2 instead uses the infinite-horizon balanced-truncation estimate for a stable strictly proper transfer map and supplies a realizable reduced model. Neither task bound is invariant under arbitrary unit changes; the input norm, output metric, noise whitening, and Lipschitz constant must be registered together.

Finite Markov Embedding Details

If

$$K(t) = \sum_{j=1}^{m} a_j e^{-\lambda_j t}, \tag{102}$$

define auxiliary variables

$$r_j(t) = \int_0^t e^{-\lambda_j (t-s)} z(s) ds. \tag{103}$$

Then

$$\dot{r}_j = -\lambda_j r_j + z, \tag{104}$$

and the memory term is $\sum_j a_j r_j$. This gives an exact m-dimensional augmentation. General strictly proper rational kernels follow from state-space realization theory. A proper rational kernel contains an additional feedthrough term D, which is absorbed into the instantaneous retained generator. Repeated poles require Jordan blocks; complex conjugate poles can be represented by real two-dimensional blocks.

Model Reproducibility

The script `FDS_P8_residue_spectral_models.py` reproduces all figures, CSV data, and numerical diagnostics reported in this paper. The calculations are deterministic and require NumPy, SciPy, pandas, and Matplotlib. Tested package versions are recorded in `requirements.txt`.

The minimal release archive contains the script, `requirements.txt`, and a short `README.md` specifying the execution command and expected outputs. Generated figures and data files are not distributed separately because they are recreated directly by the script.

Run command:

```
python FDS_P8_residue_spectral_models.py
--outdir .
```

The normal-form models are diagnostics of the stated theorem conditions and are not presented as universal parameter estimates. The code and environment specification are archived with the published manuscript under the release DOI.

AI Assistance Disclosure

Drafting, organization, code scaffolding, symbolic checking, and model-generation assistance were prepared with AI tools under the author's direction. The theoretical claims, proof obligations, scope boundaries, and final responsibility remain with the author.

[1] Y. Wu, *Active Finite Distinction Systems: A Formal Core for Boundary Maintenance under Finite Capacity*, Zenodo (2026), doi:10.5281/zenodo.20158923.

[2] Y. Wu, *Finite-Bath Memory, Markovianization, and Environmental Forgetting in Finite Distinction Systems: Side Records, Memory Kernels, and the Loss of Recoverable Distinctions*, Zenodo (2026), doi:10.5281/zenodo.20272541.

[3] Y. Wu, *Coarse-Grained Anti-Recurrence and Informational Hysteresis in Finite Memory Systems: Lost Preimages, Side Records, and Capacity-Recovery Asymmetry*, Zenodo (2026), doi:10.5281/zenodo.20265065.

[4] Y. Wu, *Capacity Overflow, Effective Stochasticity, and Phase-B Invariants: Critical Deficit, Markov Closure, and Invariant Selection under Finite Projection*, Zenodo (2026), doi:10.5281/zenodo.20250367.

[5] Y. Wu, *Topological Obstruction to Forgetting in Finite Distinction Systems: Quotient Invariants, Non-Hermitian Skin Effects, and Topological Side-Ledgers*, Zenodo (2026), doi:10.5281/zenodo.20265386.

[6] S. Nakajima, "On Quantum Theory of Transport Phenomena: Steady Diffusion," Prog. Theor. Phys. 20, 948–959 (1958), doi:10.1143/PTP.20.948.

[7] R. Zwanzig, "Ensemble Method in the Theory of Irreversibility," J. Chem. Phys. 33, 1338–1341 (1960), doi:10.1063/1.1731409.

[8] R. Zwanzig, "Memory Effects in Irreversible Thermodynamics," Phys. Rev. 124, 983–992 (1961), doi:10.1103/PhysRev.124.983.

[9] H. Mori, "Transport, Collective Motion, and Brownian Motion," Prog. Theor. Phys. 33, 423–455 (1965), doi:10.1143/PTP.33.423.

[10] A. J. Chorin, O. H. Hald, and R. Kupferman, "Optimal Prediction and the Mori–Zwanzig Representation of Irreversible Processes," Proc. Natl. Acad. Sci. USA 97, 2968–2973 (2000), doi:10.1073/pnas.97.7.2968.

[11] A. J. Chorin, O. H. Hald, and R. Kupferman, "Non-Markovian Optimal Prediction," *Monte Carlo Methods and Applications* **7**, 99–110 (2001), doi:10.1515/mcma.2001.7.1-2.99.

[12] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[13] H.-P. Breuer, E.-M. Laine, and J. Piilo, "Measure for the Degree of Non-Markovian Behavior of Quantum Processes in Open Systems," Phys. Rev. Lett. 103, 210401 (2009), doi:10.1103/PhysRevLett.103.210401.

[14] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, "Colloquium: Non-Markovian Dynamics in Open Quantum Systems," Rev. Mod. Phys. 88, 021002 (2016), doi:10.1103/RevModPhys.88.021002.

[15] J. G. Kemeny and J. L. Snell, *Finite Markov Chains* (Springer, New York, 1976).

[16] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, 1997).

[17] R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," J. Soc. Ind. Appl. Math. Ser. A Control 1, 152–192 (1963), doi:10.1137/0301010.

[18] T. Kailath, *Linear Systems* (Prentice-Hall, Englewood Cliffs, 1980).

[19] B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Trans. Autom. Control 26, 17–32 (1981), doi:10.1109/TAC.1981.1102568.

[20] K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L^∞-Error Bounds," Int. J. Control 39, 1115–1193 (1984), doi:10.1080/00207178408933239.

[21] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems* (SIAM, Philadelphia, 2005).

[22] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer, New York, 1983).

[23] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations* (Springer, New York, 2000).

[24] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators* (Princeton University Press, Princeton, 2005).

[25] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation* (Cambridge University Press, Cambridge, 1987).

[26] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd ed. (Wiley, New York, 1971).

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, Hoboken, 2006).

[28] C. Eckart and G. Young, "The Approximation of One Matrix by Another of Lower Rank," *Psychometrika* **1**, 211–218 (1936), doi:10.1007/BF02288367.

[29] B. Simon, *Trace Ideals and Their Applications*, 2nd ed., Mathematical Surveys and Monographs, Vol. 120 (American Mathematical Society, Providence, RI, 2005).

[30] C. Widder, J. Zimmer, and T. Schilling, "On the Generalized Langevin Equation and the Mori Projection Operator Technique," *Journal of Physics A: Mathematical and Theoretical* **58**, 405001 (2025), doi:10.1088/1751-8121/ae02cc.

[31] C. Widder, J. Zimmer, and T. Schilling, "ADDENDUM: On the Generalized Langevin Equation and the Mori Projection Operator Technique (2025 J. Phys. A: Math. Theor. 58 405001)," *J. Phys. A: Math. Theor.* **59**, 229401 (2026), doi:10.1088/1751-8121/ae7168.

[32] C. Widder and T. Schilling, "Generalised Langevin Dynamics: Significance and Limitations of the Projection Operator Formalism," arXiv:2604.20453v1 (2026).

[33] B. Hilder and U. Sharma, "Quantitative Coarse-Graining of Markov Chains," *SIAM Journal on Mathematical Analysis* **56**, 913–954 (2024), doi:10.1137/22M1473996.

[34] P. Gupta, P. J. Schmid, D. Sipp, T. Sayadi, and G. Rigas, "Mori–Zwanzig Latent Space Koopman Closure for Nonlinear Autoencoders," *Proceedings of the Royal Society A* **481**, 20240259 (2025), doi:10.1098/rspa.2024.0259.

[35] Q. Lang and J. Lu, "Learning Memory Kernels in Generalized Langevin Equations," *SIAM Journal on Mathematics of Data Science* **8**, 141–166 (2026), doi:10.1137/24M1651101.

[36] L. Tepper, B. Dalton, and R. R. Netz, "Accurate Memory Kernel Extraction from Discretized Time-Series Data," *Journal of Chemical Theory and Computation* **20**, 3061–3068 (2024), doi:10.1021/acs.jctc.3c01289.

[37] D. Jaganathan and R. N. Valani, "Markovian Embedding of Nonlinear Memory via Spectral Representation," *Communications in Nonlinear Science and Numerical Simulation* **154**, 109540 (2026), doi:10.1016/j.cnsns.2025.109540.