# WHAT CAN A REPRESENTATION IDENTIFY? SPECTRAL LIMITS OF LEARNING PDE DYNAMICS FROM SPATIAL DATA

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Inferring dynamical laws from spatial observations is fundamentally limited by the information preserved in the representation supplied to the inference method. We study this limitation for PDE identification from one or a few spatial observations through a representation-dependent view of identifiability. For linear translation-invariant PDEs, we show that power spectral density (PSD) captures only the real part of the Fourier symbol, while phase-sensitive temporal information is required to recover its imaginary part. Consequently, mechanisms such as advection and odd-order dispersion are invisible to magnitude-only spectral representations. We characterize when PDE parameters can be recovered from spectral data and show how identifiability depends jointly on the governing dynamics, observation protocol, and representation used for inference. When elapsed time is unknown, only parameter ratios, parameter combinations, or spectral shape information may remain identifiable, while a single spatial snapshot generally confounds the initial condition with subsequent dynamics. Numerical experiments distinguish structural non-identifiability from practical estimation error and representation-induced information loss. These results motivate a representation-first approach to learning dynamics from spatial data: before selecting an inference architecture, one should determine which dynamical distinctions the chosen representation makes learnable.

## 1 INTRODUCTION

In many scientific applications, including tumor biopsy imaging, we observe spatial structure at only one or a few time points and seek to infer the dynamics that generated it. But which aspects of those dynamics are recoverable from a chosen image representation? Spatially resolved measurements are becoming increasingly available across scientific disciplines, creating new opportunities to infer the dynamical mechanisms underlying observed spatial patterns. Partial differential equations (PDEs) provide a natural mechanistic framework for such dynamics (Wang et al., 2024; Daher et al., 2026). A growing body of work addresses the associated inverse problem using sparse regression for equation discovery (Brunton et al., 2016; Rudy et al., 2017), neural operators (Li et al., 2020), and physics-informed neural networks (Raissi et al., 2019). However, successful model fitting does not by itself establish that a target mechanism or parameter is identifiable from the representation supplied to the inference procedure. This matters when observations are transformed or compressed before inference, because the representation may remove dynamical information, and when temporal information is scarce: a biopsy is typically a single frame taken at an unknown time since initiation.

We therefore formulate PDE identification as a representation-dependent inverse problem: which mechanisms, parameters, or parameter combinations remain identifiable after spatial observations are represented through Fourier magnitude, phase, or both, and how does the answer change as temporal information is reduced to a single frame?

For a linear constant-coefficient PDE $u_t = \mathcal{L}u$ on a periodic domain, plane waves are eigenfunctions of $\mathcal{L}$ with Fourier symbol $\lambda(\mathbf{k})$, so $\widehat{u}(\mathbf{k}, t) = \widehat{u}(\mathbf{k}, 0)\, e^{\lambda(\mathbf{k})t}$ and the power spectral density (PSD)



Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

satisfies

$$P(\mathbf{k}, t) = |\widehat{u}(\mathbf{k}, t)|^2 = P(\mathbf{k}, 0)\, e^{2\,\mathrm{Re}[\lambda(\mathbf{k})]t}. \tag{1}$$

Compression to PSD magnitude retains only $\mathrm{Re}\,\lambda$. Mechanisms encoded exclusively through $\mathrm{Im}\,\lambda$, such as advection and odd-order dispersion, are therefore structurally unidentifiable from PSD alone, irrespective of the estimator.

The same identity couples dynamics to the initial state. A single snapshot gives one measurement per wavenumber for two unknowns, $P(\mathbf{k}, 0)$ and $\mathrm{Re}\,\lambda(\mathbf{k})T$, so without an assumption on the initial spectrum nothing about the dynamics is identifiable. We show that time resolution and knowledge of the initial spatial organization are interchangeable: a known initial structure factor makes the spectral *shape* of one frame informative, and the image mean calibrates the spectral *level*.

This separates two limitations of PDE inference. When information is present in the representation but estimation is degraded by resolution, noise, or correlated basis functions, the difficulty is *practical*. When a parameter does not affect the observable, no estimator or additional measurement of the same representation can recover it; the limitation is *structural*. The appropriate inferential target may then be a ratio, a parameter combination, or an observational equivalence class rather than a unique PDE label.

Our contributions are:

1. A characterization of which components of linear PDE dynamics survive compression to PSD magnitude and which are recoverable only from phase.
2. Four observation protocols (full time series; two frames at known and unknown interval; a single frame at unknown time), each with an estimator that returns exactly the object the protocol determines: absolute parameters, the symbol up to scale, or growth-normalized ratios.
3. A sharp three-level identifiability result for a single unknown-time frame: nothing without a model of the initial spectrum; the operator family and all ratios among wavenumber-dependent coefficients with a known structure factor; and, with a calibrated level, all growth-normalized ratios such as the invasiveness $D/R$.
4. A cross-spectral estimator recovering advection and odd-order dispersion from the phase of two frames, which no single frame can recover.
5. Validation across 22 PDE families and several initial spatial organizations, including a demonstration that a previously used single-frame "shoulder" fit achieves $R^2 \geq 0.97$ while returning ratios wrong by factors of 24–59: goodness of fit does not certify identifiability.

## 2 Representation-dependent spectral identifiability

Let $\rho(\mathbf{x}, t)$ solve $\partial_t \rho = \mathcal{L}\rho$ on a periodic domain, with symbol $\lambda(\mathbf{k}) = \mu(\mathbf{k}) + i\omega(\mathbf{k})$, $\mu = \mathrm{Re}\,\lambda$, $\omega = \mathrm{Im}\,\lambda$. By Eq. (1),

$$\log P(\mathbf{k}, t) = \log P(\mathbf{k}, 0) + 2\mu(\mathbf{k})t, \tag{2}$$

so temporal changes in PSD inform $\mu$, while $\omega$ is removed by the magnitude operation. In two and three dimensions we radially average to $\overline{P}(k, t)$, $k = |\mathbf{k}|$, which improves robustness for isotropic dynamics but discards directional information. For isotropic real parts we write $\mu(k) = c_0 + \sum_j m_j k^{p_j}$, where $c_0 = \mu(0)$ is the net growth rate (e.g. $R$ in reaction–diffusion) and the $m_j$ are the wavenumber-dependent coefficients.

Parameters are structurally identifiable from PSD evolution exactly to the extent that the map $\boldsymbol{\theta} \mapsto \mu(k; \boldsymbol{\theta})$ is injective. If $\mu(k; \boldsymbol{\theta}_1) = \mu(k; \boldsymbol{\theta}_2)$ for all observed $k$, the two models are observationally equivalent under PSD magnitude; in particular, a coefficient appearing only in $\omega$ is invisible. Injectivity must be checked on the whole map rather than per power of $k$: linearized Swift–Hohenberg with diffusion is injective in $(R, D, \ell)$, whereas diffusion, hyperdiffusion and Swift–Hohenberg together have four parameters but only three observable coefficients. When the map is injective, finite band, resolution, noise, and correlated basis functions limit recovery only practically.

We consider four observation protocols (Table 1): a full time series (S1), two frames at a known interval (S2), two frames at an unknown interval (S3), and a single frame at unknown time $T$ since

initiation (S4), plus a regional variant of S4 (S5). For two frames,

$$y(k) := \log \frac{P(k,t_2)}{P(k,t_1)} = 2\Delta t\, \mu(k), \qquad \Delta t = t_2 - t_1, \tag{3}$$

and for polynomial symbols

$$\mu(k;\boldsymbol{\theta}) = \sum_{m=0}^{M} c_m(\boldsymbol{\theta}) k^{2m}. \tag{4}$$

**Proposition 1** (Two frames at unknown elapsed time)**.** *Two PSD frames separated by an unknown positive interval determine the coefficient vector only up to a common positive scale, i.e. its projective class $[\mathbf{c}(\boldsymbol{\theta})]$ and every ratio $c_m/c_{m'}$ between nonzero coefficients. Physical parameters are identifiable exactly to the extent that $\boldsymbol{\theta} \mapsto [\mathbf{c}(\boldsymbol{\theta})]$ is injective; elapsed time and overall scale remain confounded unless the model provides an internal calibration.*

*Proof.* See Supplementary Section D.

**The single-frame regime.** A single frame combines the unknown initial organization with the subsequent evolution. We model the initial field as cells with known profile $g$ placed at the points of a stationary point process with intensity $n_0$ and structure factor $S_0(\mathbf{k})$, $\rho(\mathbf{x},0) = \sum_i g(\mathbf{x}-\mathbf{x}_i)$. Then

$$\mathbb{E}\, P(\mathbf{k},0) = n_0\, |\hat{g}(\mathbf{k})|^2 S_0(\mathbf{k}) \quad (\mathbf{k} \neq 0), \qquad \mathbb{E}\, \bar{\rho}(0) = n_0\, \hat{g}(0), \tag{5}$$

where $\bar{\rho}$ is the spatial mean. A Poisson (shot-noise) arrangement has $S_0 \equiv 1$, hard-core packing $S_0(0) < 1$, and clustering $S_0(0) \gg 1$. One frame at time $T$ therefore satisfies

$$\log \mathbb{E}\, P(k,T) = a + \log \frac{|\hat{g}(k)|^2 S_0(k)}{\hat{g}(0)^2 S_0(0)} + h(k), \quad a = \log n_0 + \log\big(\hat{g}(0)^2 S_0(0)\big) + 2c_0 T, \quad h(k) = 2T\,[\mu(k) - \mu(0)]. \tag{6}$$

Given $g$ and $S_0$ up to a constant, the shape $h$ fixes $Tm_j$ for every wavenumber-dependent coefficient, hence the operator family and all ratios $m_j/m_{j'}$. The growth term $c_0 T$ appears only in the level $a$, confounded with $n_0$. The $k = 0$ mode resolves this: $\bar{\rho}(T) = n_0 \hat{g}(0) e^{c_0 T}$, so with known $S_0(0)$ the mean and level together give

$$c_0 T = a - \log \bar{\rho}(T) - \log \hat{g}(0) - \log S_0(0), \tag{7}$$

while with a known initial density the mean alone gives

$$c_0 T = \log \bar{\rho}(T) - \log\big(n_0\, \hat{g}(0)\big) = \log \frac{\bar{\rho}(T)}{\bar{\rho}(0)}. \tag{8}$$

Either way, every growth-normalized ratio $m_j/c_0 = (Tm_j)/(c_0 T)$ follows from one frame.

**Proposition 2** (Single frame at unknown time)**.** *Observe one frame at unknown $T$ under model* (5) *with $g$ known, and let $h(k) = 2T[\mu(k) - \mu(0)]$.*

*(a) Without a model of $P(\cdot,0)$, nothing about $(\boldsymbol{\theta}, T)$ is identifiable: for any $(\boldsymbol{\theta}', T')$, $P'(\cdot,0) = P(\cdot,T)\, e^{-2\mu(\cdot;\boldsymbol{\theta}')T'}$ reproduces the data.*

*(b) If $S_0$ is known up to a constant factor, the frame identifies $h$ on the band and nothing more. Hence it identifies the operator family, all ratios $m_j/m_{j'}$, and shape parameters (fractional order $\alpha$, kernel width $\sigma$, Swift–Hohenberg length $\ell$), but not $c_0$ or any $m_j/c_0$.*

*(c) If additionally $\bar{\rho}(T)$ is observed and either $S_0(0)$ or $n_0$ is known, Eq.* (7) *or* (8) *identifies $c_0 T$ and every ratio $m_j/c_0$ (e.g. $D/R$). Wrongly assuming $S_0(0) = 1$ shifts Eq.* (7) *by exactly $\log S_0(0)$, a multiplicative bias $c_0 T/(c_0 T + \log S_0(0))$ on these ratios; Eq.* (8) *is unaffected.*

*Phase-only mechanisms leave Eq.* (6) *unchanged and are never identifiable from one frame.*

*Proof.* See Supplementary Section D.

**Regional analysis (S5).** A frame is often analyzed as several regions of interest (ROIs) because the image is non-periodic or the tissue heterogeneous. S5 uses the same frame as S4 and therefore has exactly the same identifiable targets; only the estimator changes. Each ROI must be tapered (we use a separable Hann window) to control boundary leakage, and pooling ROI spectra reduces variance and the influence of large-scale initial structure at the cost of the lowest wavenumbers.

Table 1: Observation protocols and what each identifies for $\partial_t \rho = \mathcal{L}\rho$. Mechanisms entering only $\mathrm{Im}\,\lambda$ (advection, odd dispersion) are invisible to every PSD protocol and require the cross-spectrum phase of two frames.

| Scenario | Available observations | Spectral information | Inferential output |
|---|---|---|---|
| S1 | Full time series | Temporal evolution of $P(k,t)$ | Absolute PSD-visible parameters, wherever $\boldsymbol{\theta} \mapsto \mu(\cdot;\boldsymbol{\theta})$ is injective |
| S2 | Two frames, known $\Delta t$ | PSD change over known interval; cross-spectrum phase | Absolute PSD-visible parameters; advection and dispersion from the phase |
| S3 | Two frames, unknown $\Delta t$ | Scaled symbol $s\mu(k)$ | All ratios $m_j/c_0$ and shape parameters; $\Delta t$ only with an internal calibration |
| S4 | Single frame, no model of $P(k,0)$ | Radial PSD $P(k,T)$ | Nothing identifiable |
| | Single frame, known $S_0(\mathbf{k})$ | Shape $h(k)$ | Operator family, ratios $m_j/m_{j'}$, shape parameters; not $c_0$ |
| | … plus image mean and $S_0(0)$ or $n_0$ | Shape plus calibrated level | Also $c_0T$, hence all $m_j/c_0$ (e.g. $D/R$); never $c$ or $\beta_3$ |
| S5 | Regions of interest of one frame | Hann-windowed, pooled ROI spectra | Same as S4; trades lowest wavenumbers for control of leakage and heterogeneity |

## 3 Model library and computational design

We evaluate the theory on a library of 22 linear PDE families spanning reaction, diffusion, advection, hyperdiffusion, Swift–Hohenberg, fractional diffusion, odd-order dispersion, sixth-order diffusion, pure loss, linearized Kuramoto–Sivashinsky and Cahn–Hilliard, and Gaussian nonlocal interaction, together with combinations of the core operators. Their symbols and PSD-visible components are given in Supplementary Table S3 and Supplementary Fig. S3.

Unless stated otherwise, experiments are two-dimensional on a periodic $5.12\,\mathrm{mm}$ domain of $2048^2$ pixels ($2.5\,\mu$m), evolved exactly in Fourier space with the full complex symbol and the spatial mean retained. The held-out single-frame validation (§4.4) uses a periodic $2.5\,\mathrm{mm}$ domain of $256^2$ pixels seeded with 1500 cells, with eight realizations of each organization for evaluation and four further realizations used only to tune estimator settings. Initial organizations (Poisson, hard-core, clustered; Supplementary Fig. S1), one- and three-dimensional controls, and robustness analyses are in the Supplementary Material.

**Estimators.** For S1–S3 we fit $\mu(k;\boldsymbol{\theta})$ to temporal changes in the spectrum; phase-only mechanisms are estimated from the cross-spectrum phase. For a single frame (S4/S5) we (i) compute the radial PSD, Hann-windowing each ROI when $f < 1$ or the field is non-periodic, and pool ROIs; (ii) divide out an initial-spectrum model $Q(k) \propto |\hat{g}(k)|^2 S_0(k)$ to obtain $y(k) = a + h(k)$; (iii) fit $y$ by mode-count-weighted least squares for each shape class (flat, $k^{1.5}$, $k^2$, $k^4$, $k^6$, $k^2{+}k^4$, Gaussian) and select by BIC; and (iv) calibrate $c_0T$ by Eq. (7) or (8), returning shape ratios and growth-normalized ratios $m_j/c_0 = \beta_j/(2\widehat{c_0T})$. Pseudocode for all estimators is given in Supplementary Algorithms (single frame: Algorithm 1).

## 4 Results

### 4.1 Spectral representations induce observational equivalence classes

Different mechanisms generate characteristic wavenumber dependences of $\mu(k)$: $k^2$, $k^4$, $k^6$ for diffusion-type terms, $|k|^\alpha$ for fractional diffusion, non-monotone finite-wavelength structure for Swift–Hohenberg, Kuramoto–Sivashinsky, and Cahn–Hilliard, and non-polynomial rolloff for Gaussian coupling (Supplementary Fig. S3). Two models are observationally equivalent under PSD magnitude when $\mathrm{Re}\,\lambda_1(k;\boldsymbol{\theta}_1) = \mathrm{Re}\,\lambda_2(k;\boldsymbol{\theta}_2)$ over the observed band. Advection and odd-order dis-

persion are the clearest examples: reaction–advection has the same PSD-visible symbol as pure reaction. The temporal protocol and the initial-state model induce further equivalences: two frames at unknown interval determine $\mu$ only up to scale, and a single frame determines only $h$, relative to a model of the initial spectrum, so that parameters sharing a power without a time reference ($D$ and $\ell$ in diffusion–Swift–Hohenberg) merge. The appropriate output is therefore an observational equivalence class whenever the observable symbols of competing models coincide.

### 4.2 Known time removes temporal-scale ambiguity

With known elapsed time (S1, S2) the full symbol $\mu(k)$ is estimated in physical units, including $c_0$, and multi-term families have their coefficients estimated jointly. Recovery is governed by injectivity of $\boldsymbol{\theta} \mapsto \mu(k;\boldsymbol{\theta})$ rather than a one-to-one assignment between parameters and powers of $k$. Across the families for which this map is injective and across initial organizations, absolute parameters are recovered with median absolute percentage error below $3.2\%$ at $N = 1000$ (Appendix J). Residual variation is practical (finite band, resolution, noise, and correlated $k^2$ and $k^4$ regressors in multi-term families), and recovery improves with domain size. Parameters absent from $\mathrm{Re}\,\lambda$ remain inaccessible at every setting.

### 4.3 Phase information restores dynamics invisible to PSD magnitude

Advection and odd-order dispersion enter only $\omega(k)$, so their PSD is unchanged (Fig. 1a,d). Two time-separated frames retain the information through the cross-spectrum $C(k;t_1,t_2) = \widehat{u}(k,t_2)\,\overline{\widehat{u}(k,t_1)}$, whose phase satisfies

$$\phi(k) := \arg C(k;t_1,t_2) = \omega(k)\Delta t \pmod{2\pi}. \tag{9}$$

With known $\Delta t$ this identifies the advection velocity $c$ and dispersion coefficient $\beta_3$; with unknown $\Delta t$, their products with elapsed time (Fig. 1c,f). Their failure under PSD is thus representation-induced, not an absence of signal. Derivations, phase-wrapping conditions, and the estimator are in the Supplementary Material.

### 4.4 A single frame identifies the shape, given the initial organization

**Poisson seeding.** On Poisson-seeded frames ($S_0 \equiv 1$) with shot-noise calibration, Eq. (7), the estimator selects the correct operator family in 128 of 132 frames across the 22 families, recovers shape ratios to about 1%, and recovers $D/R$ within 5%. Hard-core seeding biases this calibration by exactly $\log S_0(0)$, which supplying $S_0(0)$ removes, and clustered seeding breaks it, as Proposition 2c predicts.

**Non-Poisson seeding, held out.** We next tested hard-core and clustered organizations, which violate the shot-noise assumption, on held-out realizations, calibrating $c_0T$ from the image mean and initial density, Eq. (8). We supplied the initial-spectrum model $Q(k)$ in three ways: from the initial frame of the same field (a two-frame control), from an independent realization of the same organization (known structure factor, Proposition 2b), and not at all (white initial spectrum assumed). With the structure factor supplied, the whole-frame estimator selects the correct class in $94 \pm 3\%$ of frames (mean $\pm$ SD over eight realizations) and recovers identifiable ratios within 10% in $86 \pm 9\%$, close to the two-frame control ($94\%$ and $97\%$). Assuming a white spectrum lowers these to $72\%$ and $54\%$ (Fig. 2).

For the invasiveness ratio $D/R$, S1–S3 are exact to within $0.002$ decades. From one frame with a white spectrum assumed, S4 has a median error of $0.06$ decades (15%), rising to $0.28$ decades (factor 1.9) for strongly clustered seedings. Pooling Hann-windowed regions (S5) reduces this to a median of $0.016$ decades (4%) and worst case $0.10$ decades (26%): regional analysis helps by down-weighting large-scale initial structure, not by adding information. Regional analysis requires windowing: without it, class selection drops to $47\%$; with it, $88\%$. Regional ratios are less accurate than whole-frame ratios (29–$43\%$ within 10%) because smaller regions resolve fewer low-wavenumber shells, so on a periodic field the whole frame is preferable.

**Multi-term operators.** Multi-term operators remain separable from one frame but are less well conditioned. With known $S_0$, single-term ratios ($D/R$, $\nu/R$, $\ell$) were recovered within 10% in

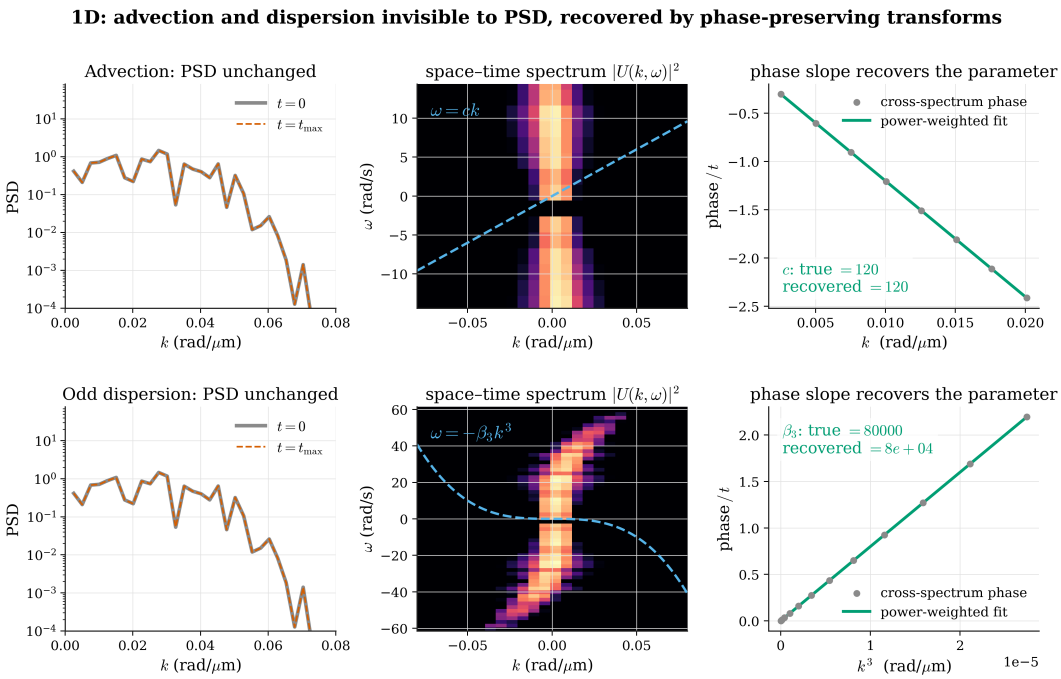



Figure 1: **Phase recovers dynamics invisible to PSD magnitude.** Top: linear advection; bottom: third-order dispersion. (a,d) The PSD at $t_{\max}$ (dashed) overlaps the initial PSD (grey): magnitude carries no information about $c$ or $\beta_3$. (b,e) The space–time spectrum shows the dispersion relation, linear for advection and cubic for dispersion. (c,f) The cross-spectrum phase varies as $k$ and $k^3$; power-weighted fitting over the phase-unambiguous band recovers $c$ and $\beta_3$ for known $\Delta t$, or their products with $\Delta t$ otherwise.

89–94% of frames, but $D$ and $\nu$ in diffusion–hyperdiffusion only in 42% and 78%. Kuramoto–Sivashinsky and Cahn–Hilliard were always assigned the correct two-term class; $D$ and $\ell$ in diffusion–Swift–Hohenberg were never separated, consistent with the non-injective single-frame map. Misclassifications concentrate between neighbouring classes ($k^4$ vs. $k^2{+}k^4$, $k^{1.5}$ vs. $k^2$, Gaussian vs. low-order polynomials). These results use periodic, noise-free $256^2$ fields and are best-case accuracies, not estimates for microscopy data.

### 4.5 A GOOD FIT DOES NOT CERTIFY IDENTIFIABILITY

A different generative assumption, a persistent source $\partial_t u = \mathcal{L}u + S(\mathbf{x})$ with $u(\cdot, 0) = 0$, gives $\hat{u}(\mathbf{k}, T) = \hat{S}(\mathbf{k})\,(e^{\lambda T} - 1)/\lambda$, whose spectrum has a "shoulder" near $\mu(k) = 0$. A previously used single-frame approach fits this template, $\log A + 2\log|(e^{\lambda T} - 1)/\lambda|$, and reads a ratio from the shoulder (Pasetto et al., 2024). Applied to frames generated by $\partial_t \rho = \mathcal{L}\rho$, it still fits closely ($R^2 = 0.97$–$1.00$ for reaction–diffusion across organizations) but underestimates $\log_{10}(D/R)$ by 1.38–1.77 decades, factors of 24–59, because it effectively returns $bT$ rather than $b/c_0$. Which generative model applies is an assumption about the biology, and a high spectral goodness of fit does not reveal a mismatch.

## 5 DISCUSSION AND CONCLUSION

Our results separate three limitations often conflated in PDE inference: practical error in estimating information that is present, structural non-identifiability when different parameters produce the same observable, and representation-induced loss when the observable discards a dynamical channel. For linear constant-coefficient PDEs this separation is explicit in the Fourier symbol: PSD reports $\mathrm{Re}\,\lambda$, cross-spectrum phase reports $\mathrm{Im}\,\lambda$, and the temporal protocol determines whether the observable symbol is absolute, projective, or confounded with the initial state.




Figure 2: **Single-frame operator-class identification depends on the initial-spectrum model and on regional preprocessing.** Correct (green, check) or incorrect (red, cross) selection of the PSD-visible operator class from one frame at unknown time, for each of the 22 PDE families (rows) and four non-Poisson initial organizations (columns), with the initial structure factor supplied from an independent realization of the same organization (Proposition 2b). **(A)** Whole-frame inference (S4): 93% of the 88 family–organization cases are classified correctly. **(B)** Regional inference (S5) on unwindowed regions of interest: 41% correct, because boundary leakage from the non-periodic regions corrupts the spectral shape. Applying a Hann window restores regional class selection to 84–88% (Supplementary Fig. **??**). S4 and S5 have the same identifiable targets; the difference is practical, not structural.

For tumor imaging, where a biopsy is a single observation at unknown time since initiation, the theory states what is required to interpret an inferred invasiveness ratio mechanistically: a model of the initial spatial organization and an independent calibration of accumulated growth, through the structure-factor level or the initial density. Without these, an image supports a spectral-shape statement but not a growth-normalized parameter claim, and these requirements should be validated experimentally rather than assumed. Regional analysis does not change the structural targets; its value is practical robustness to non-periodic boundaries and heterogeneity.

The analysis assumes periodic domains and constant-coefficient linear operators. Nonlinearity marks the boundary of the fixed-symbol theory: logistic growth couples Fourier modes, so fitted rates become state-dependent, and in our simulations parameter errors rise to 13–75% while goodness of fit remains high (e.g. $R^2 = 1.00$ with a 49% bias in $D/R$ when seeded at carrying capacity; Supplementary Section H). Systematic dependence of inferred rates on density or time is a more useful diagnostic of breakdown than fit quality. Spatially varying coefficients, segmentation uncertainty, measurement noise, and finite fields of view may further alter practical recovery.

More broadly, these results support a representation-first workflow: before optimizing a model, determine whether the observation protocol, the representation, and the available knowledge of the initial state preserve the distinctions required for the intended claim. The target then shifts from "the best-fitting PDE" to the strongest quantity the data support: an absolute parameter, a ratio, a shape parameter, or an observational equivalence class.

## REPRODUCIBILITY STATEMENT

All simulations, estimators, and figure-generation workflows, including the single-frame estimator with held-out validation, the shoulder-template negative control, and the 22-family study in one, two, and three dimensions, are provided as reproducible Jupyter notebooks at https://anonymous.4open.science/r/XXXX (anonymized for review). Proofs, the full PDE library, all estimator pseudocode, simulation protocols, and extended results are in the Supplementary Material.

## AI USE STATEMENT

Generative AI tools (ChatGPT, OpenAI; Claude, Anthropic) were used to assist with manuscript editing and with reviewing and revising research code. All AI-assisted content and code modifications were reviewed and verified by the authors, who take responsibility for the derivations, experimental design, analysis, and conclusions.

## REFERENCES

Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

Ali Daher, Dumitru Trucu, and Raluca Eftimie. Calibrating tissue level pde models of ligand dynamics using single cell and spatial transcriptomics data. *npj Systems Biology and Applications*, 12(1):44, 2026.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

Stefano Pasetto, Michael Montejo, Mohammad U Zahid, Marilin Rosa, Robert Gatenby, Pirmin Schlicke, Roberto Diaz, and Heiko Enderling. Calibrating tumor growth and invasion parameters with spectral spatial analysis of cancer biopsy tissues. *npj Systems Biology and Applications*, 10 (1):112, 2024.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Samuel H Rudy, Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Data-driven discovery of partial differential equations. *Science advances*, 3(4):e1602614, 2017.

Haiyun Wang, Jianping Zhao, Qing Nie, Chunhou Zheng, and Xiaoqiang Sun. Dissecting spatiotemporal structures in spatial transcriptomics via diffusion-based adversarial learning. *Research*, 7: 0390, 2024.

# Supplementary Material

## A Supplementary simulations and observation protocols

This appendix provides additional simulation outputs and protocol-specific spectral analyses supporting the results in the main text. The figures illustrate the forward dynamics, initial spectral content, and the information retained under the observation protocols S1–S5. As established in the main text, the inferential target changes with the available temporal information and with what is known about the initial state: S1 and S2 support estimation of absolute PSD-visible parameters when the parameter-to-symbol map is injective; S3 supports quantities identifiable under the unknown temporal scale (all coefficient ratios of the projective symbol); a single frame (S4) identifies nothing without a model of the initial spectrum, the operator family and all ratios among wavenumber-dependent coefficients with a known initial structure factor, and additionally the growth-normalized ratios with a calibrated level (Proposition 2); and regional analysis of the frame (S5) has the same targets as S4. Section G gives the detailed single-frame results, and Section G.1 evaluates the alternative, persistent-source model of a single frame.

The examples below also illustrate the role of spatial dimension and initial condition. Across dimensions, the same governing PDE and parameterization are used while the spatial organization of the initial field is varied, which allows us to distinguish sensitivity to initial spectral content from information determined by the observation protocol itself.

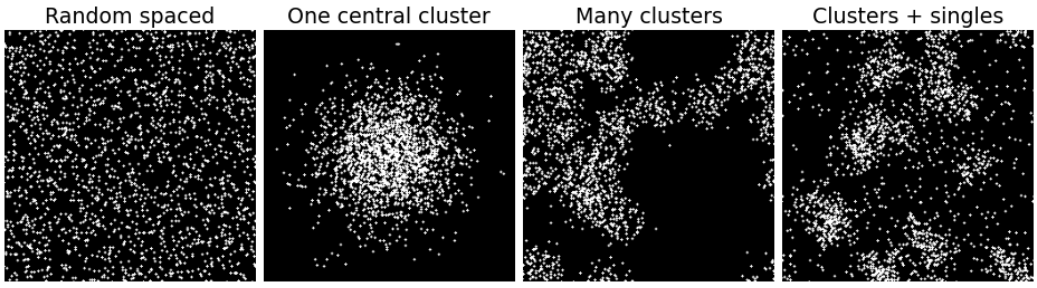



Figure S1: **Initial spatial organizations test dynamics rather than pattern recognition.** Cells of $10\,\mu$m placed as (a) Poisson point sources, (b) Poisson cells, (c) hard-core "random spaced" cells with an $18\,\mu$m minimum distance, (d) a single central cluster, (e) many clusters, and (f) clusters plus isolated cells. The patterns differ in their structure factor $S_0(\mathbf{k})$ and hence in their initial power spectra, but they are evolved under the same governing PDE and parameter values within each experiment. Time-resolved protocols cancel these differences; single-frame protocols must model them (Proposition 2).

## B Single-frame shape-and-level estimator

## C Nonlinear growth: boundary of the fixed-symbol theory

Logistic growth breaks Fourier diagonalization because the quadratic reaction term couples modes:

$$\partial_t u = Ru\left(1 - \frac{u}{K}\right) + D\nabla^2 u.$$

A linear spectral fit therefore estimates an effective, state-dependent growth rate. Parameter errors increase to approximately 13–75%, depending on model, observation time, density, and initial condition, while goodness of fit can remain high. For logistic reaction–diffusion seeded at carrying capacity, S1 overestimates $\log_{10}(D/R)$ by $0.17$ decades (49%) despite $R^2 = 1.00$; seeding at $0.3K$ and $0.05K$ reduces the bias to $0.05$ and $0.008$ decades.

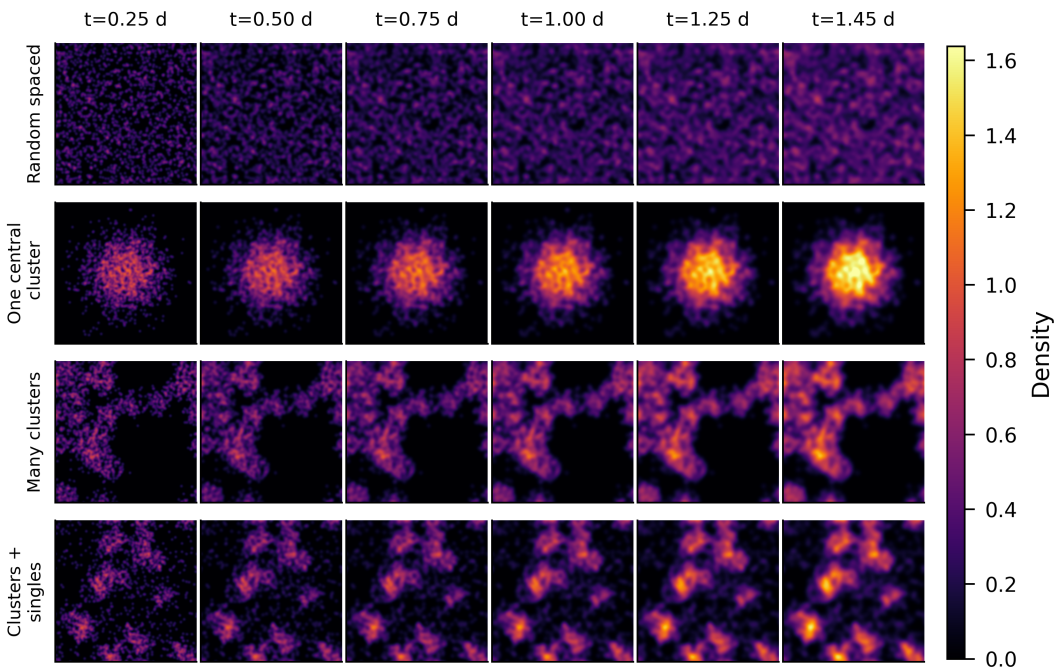



Figure S2: **Two-dimensional reaction–diffusion forward simulation.** Snapshots of the density field evolving under $\partial_t u = Ru + D\nabla^2 u$ on a periodic two-dimensional domain. The spatial field changes through the combined effects of reaction and diffusion, while the corresponding Fourier magnitude evolves according to the PSD-visible symbol $\mu(k) = R - Dk^2$. These simulated fields provide the observations used in the protocol-specific spectral analyses below.

# D Proofs and worked examples

## D.1 Proof of Proposition 1

*Proof sketch.* Substituting Eq. (4) into Eq. (3) gives $y(k) = \sum_{m=0}^{M} \widetilde{c}_m k^{2m}$ with $\widetilde{c}_m = s\, c_m(\boldsymbol{\theta})$, so the observation determines $\widetilde{\mathbf{c}} = s\,\mathbf{c}(\boldsymbol{\theta})$. For any pair of nonzero coefficients, $\widetilde{c}_m/\widetilde{c}_{m'} = c_m(\boldsymbol{\theta})/c_{m'}(\boldsymbol{\theta})$, so the common temporal scale cancels and the observable determines the projective class $[\mathbf{c}(\boldsymbol{\theta})]$ rather than the absolute coefficient vector. If $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ satisfy $[\mathbf{c}(\boldsymbol{\theta}_1)] = [\mathbf{c}(\boldsymbol{\theta}_2)]$, then $\mathbf{c}(\boldsymbol{\theta}_1) = a\,\mathbf{c}(\boldsymbol{\theta}_2)$ for some $a > 0$, and the two parameter vectors are indistinguishable after rescaling the elapsed time; conversely, injectivity of the projective map establishes structural identifiability under unknown time. Absolute time remains confounded with the overall scale unless a constraint among the coefficients that is not invariant under rescaling removes the ambiguity. □

## D.2 What a second frame adds over a single frame

Proposition 1 isolates the value of the extra measurement. Two frames at an unknown interval expose the symbol up to a single global scale (Eq. (3)), so the entire projective coefficient vector, and hence every ratio $c_m/c_{m'}$, is identifiable, including for multi-term dissipative operators, *without any assumption on the initial state*. A single unknown-time frame can reach the same ratios, but only through a model of the initial spatial organization (Proposition 2). The second frame therefore buys robustness to the initial state rather than new identifiable quantities. Finally, structural identifiability of the projective vector does not by itself guarantee well-conditioned estimation: when two dissipative powers are close, or one coefficient dominates over the observed band, their separation is practically ill-conditioned even though it remains structurally possible—a practical, not structural, limitation in the sense of §2.

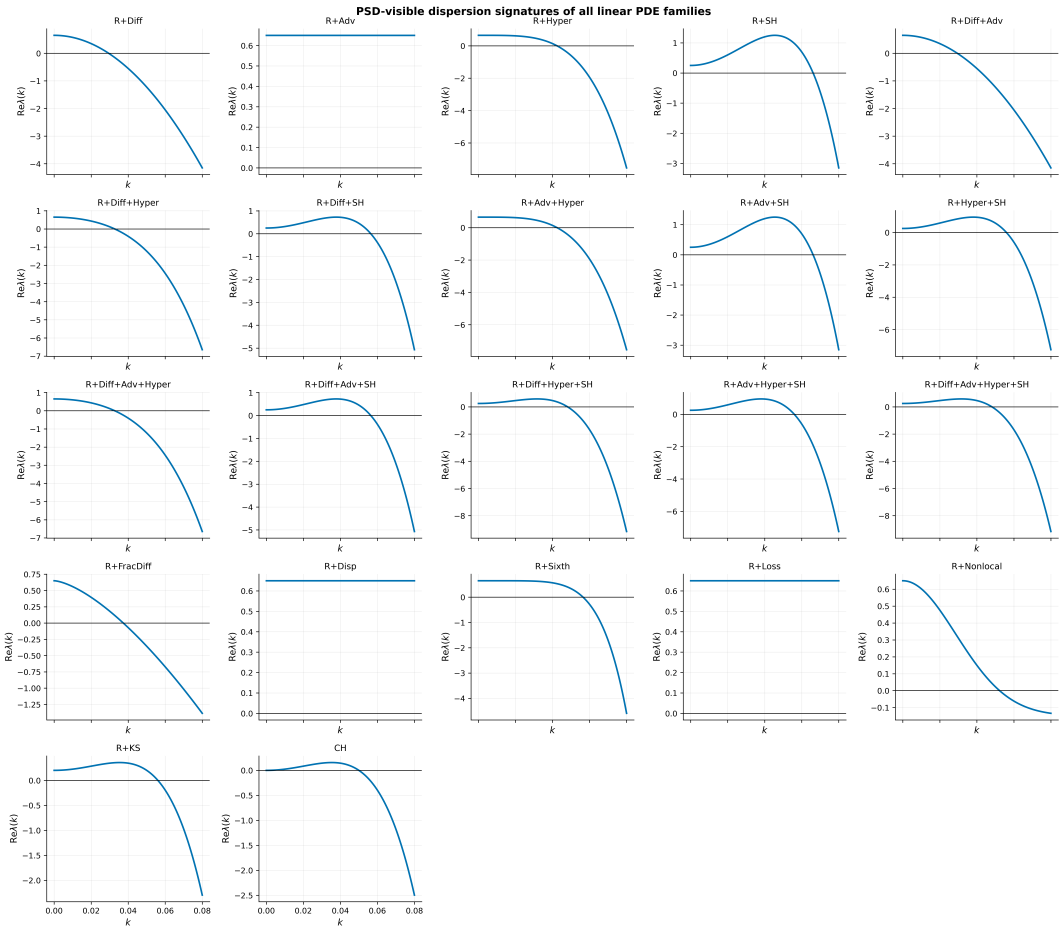



Figure S3: **PSD magnitude preserves some dynamical distinctions while collapsing others into observational equivalence classes.** The PSD-visible component $\mu(k) = \mathrm{Re}\,\lambda(k)$ is shown for all 22 candidate PDE families. Diffusion-type mechanisms produce polynomial attenuation with different wavenumber dependence, fractional diffusion introduces $|k|^\alpha$, finite-wavelength pattern-forming operators generate non-monotone spectral structure, and Gaussian nonlocal coupling produces a non-polynomial rolloff. Candidate models are distinguishable from PSD magnitude only when their observable real symbols differ over the measured wavenumber range. Mechanisms encoded exclusively in $\mathrm{Im}\,\lambda$, including advection and odd-order dispersion, do not alter the curves shown here and therefore collapse onto the same PSD-observable class as the corresponding models without those mechanisms. Additional equivalences may arise when distinct parameterizations generate the same real symbol or, under unknown elapsed time, the same symbol up to a common scale. From a single frame, equivalence additionally depends on the model of the initial spectrum and on calibration of the spectral level (§2).

### D.3 Proof of Proposition 2

*Proof sketch.* (a) is immediate from $P(\cdot, T) = P(\cdot, 0)e^{2\mu T}$ with $P(\cdot, 0)$ unrestricted. (b) By Eq. (6), the data determine $a + h(\cdot)$ with $S_0$ and $\hat{g}$ known up to constants; since $h(0) = 0$, $a$ and $h$ are separately determined, and $(n_0, c_0, T)$ enter only through $a$. (c) The $k = 0$ mode gives $\bar{\rho}(T) = n_0\hat{g}(0)e^{c_0 T}$. With $a = \log n_0 + \log \hat{g}(0)^2 S_0(0) + 2c_0 T$ and $S_0(0)$ known, subtracting gives Eq. (7); with $n_0$ known, the mean alone gives Eq. (8). Then $m_j/c_0 = (Tm_j)/(c_0 T)$. Phase-only mechanisms leave Eq. (6) unchanged and remain unrecoverable from one frame. □

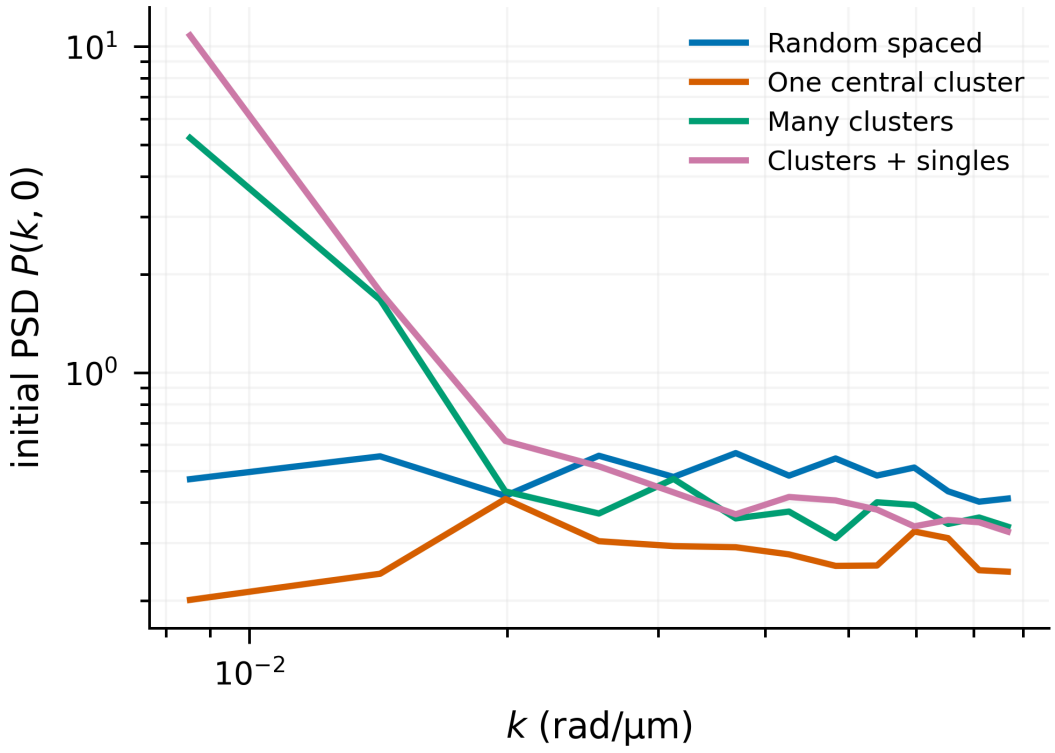



Figure S4: **Initial power spectral density across spatial organizations.** Radially averaged PSD at $t = 0$ for each initial condition. The distinct curves demonstrate that the initial spatial organizations contain different distributions of spectral power before evolution under the governing PDE. For time-separated observations, the initial spectral amplitude cancels in the ratio $P(k, t_2)/P(k, t_1)$; for a single snapshot, by contrast, the unknown initial spectrum remains confounded with the subsequent dynamical evolution unless it is modelled through its structure factor $S_0$ (Proposition 2).

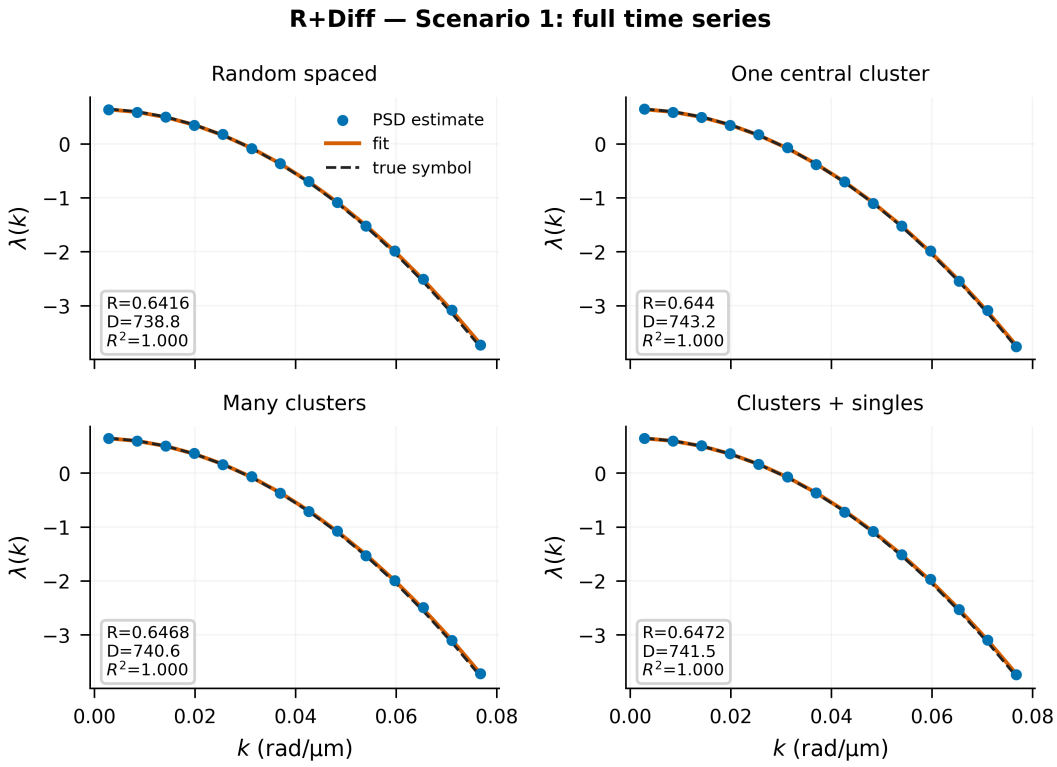



Figure S5: **Scenario 1 (full time series): recovery of absolute PSD-visible parameters.** The PSD-visible symbol is estimated from the temporal evolution $\widehat{\mu}(k) = \frac{1}{2}\,\partial_t \log P(k,t)$. Because the absolute temporal scale is retained, parameters entering $\mathrm{Re}\,\lambda(k)$ can be recovered in physical units when the corresponding parameter-to-symbol map is structurally identifiable. Recovered values are compared with the known simulation parameters across the tested initial conditions.

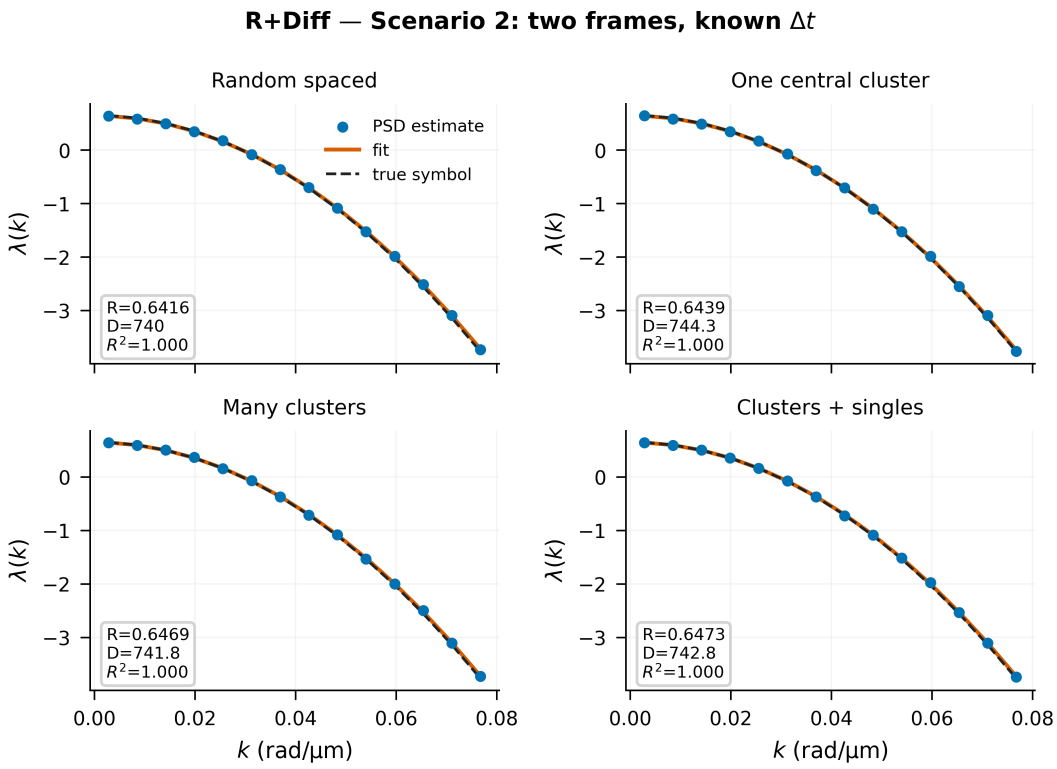



Figure S6: **Scenario 2 (two frames, known $\Delta t$): absolute PSD-visible parameter recovery.** Given two observations separated by a known interval $\Delta t$, the PSD-visible symbol is estimated as $\widehat{\mu}(k) = \frac{1}{2\Delta t} \log[P(k, t_2)/P(k, t_1)]$. Knowledge of $\Delta t$ preserves the absolute temporal scale, allowing structurally identifiable parameters of $\mathrm{Re}\,\lambda(k)$ to be estimated in physical units. This provides the two-frame counterpart of the full-time-series estimator in Scenario 1.

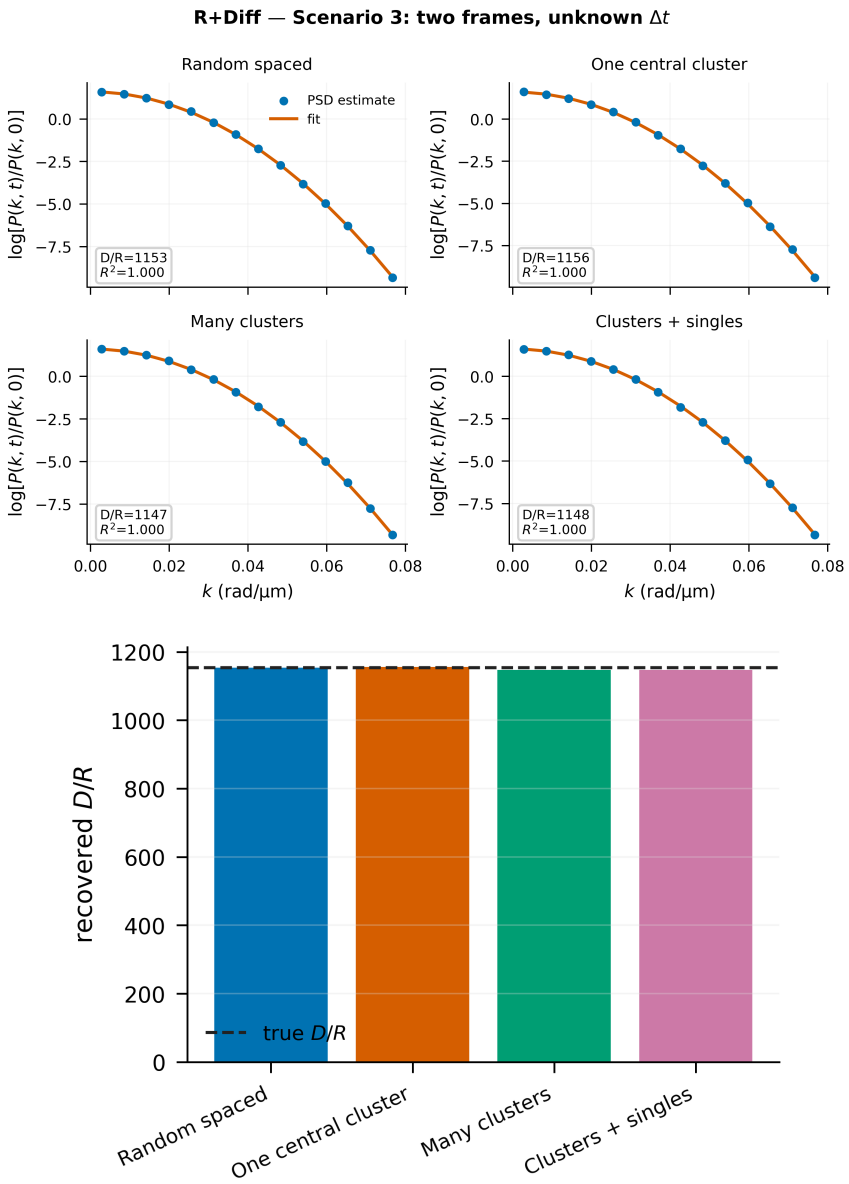



Figure S7: **Scenario 3 (two frames, unknown $\Delta t$): projective spectral identifiability.** When the elapsed time is unknown, the observable satisfies $\log[P(k,t_2)/P(k,t_1)] = 2\Delta t\,\mu(k;\boldsymbol{\theta})$, so the PSD-visible symbol is determined only up to a common positive scale. (a) The observation constrains the projective spectral form rather than the absolute coefficient vector. (b) For $\mu(k) = R - Dk^2$, the ratio $D/R$ is invariant to this unknown scale and remains identifiable even when $R$ and $D$ cannot be recovered separately. The quantities recoverable under unknown time are determined by the injectivity of the projective parameter-to-symbol map (Prop. 1); for a polynomial symbol this includes all coefficient ratios, including those of multi-term operators, without any assumption on the initial state.

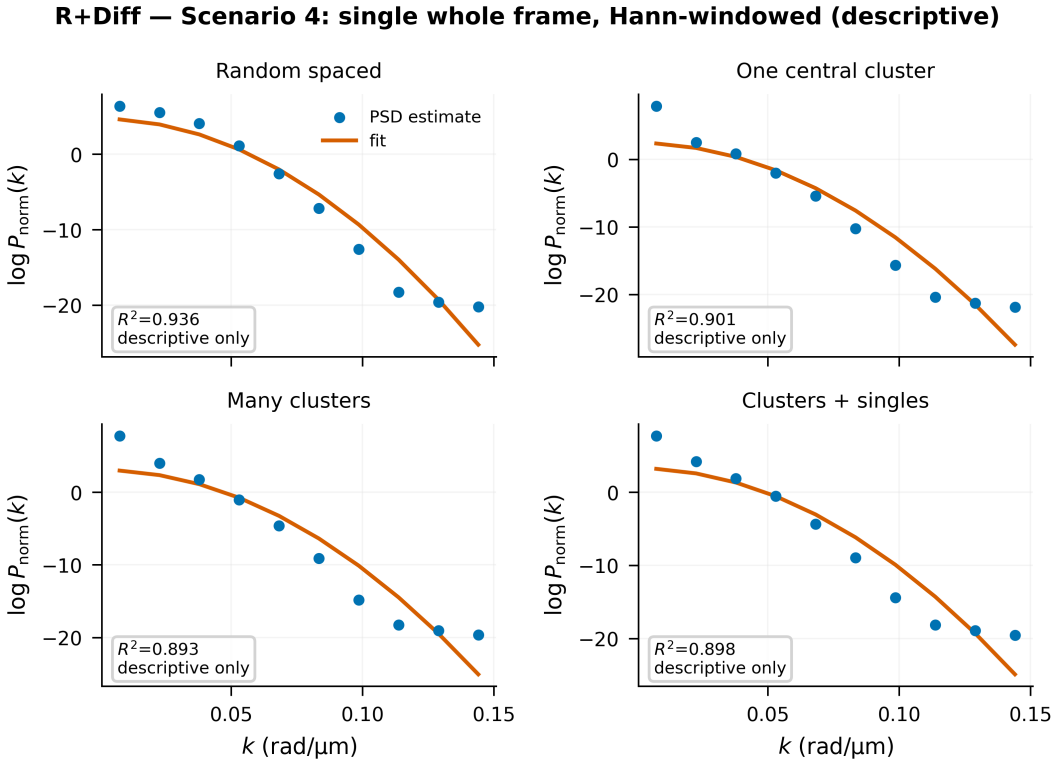



Figure S8: **Scenario 4 (single whole-frame snapshot) without a model of the initial spectrum.** A single spatial snapshot provides the observed spectrum $P(k,t)$ but no independent temporal change from which to isolate the dynamical symbol. Because $\log P(k,t) = \log P(k,0) + 2\mu(k)t$, the unknown initial spectral organization and subsequent dynamical evolution are confounded. Without a model of $P(k,0)$ the quantities shown here are descriptive spectral features of the observed field rather than recovered PDE parameters (Proposition 2a); with a known structure factor the shape becomes informative (Section G).

**Algorithm 1** Single-frame shape-and-level estimator (protocols S4/S5).

**Require:** one frame $u(\cdot, T)$ at unknown $T$; initial-spectrum model $Q(k) \propto |\hat{g}(k)|^2 S_0(k)$, known up to a constant; image mean $\bar{u}(T)$ and either $S_0(0)$ or $n_0$; ROI fraction $f$ with $n \times n$ offsets ($f$=1: S4; $f$<1: S5); band $[k_{\min}, k_{\max}]$; floor fraction $\delta$; shape classes $\mathcal{C} = \{\text{flat}, k^{1.5}, k^2, k^4, k^6, k^2{+}k^4, \text{Gaussian}\}$

1: **for** each ROI $r$ **do**
2: subtract the ROI mean; if $f < 1$ or the field is non-periodic, apply the 2-D Hann window $w_x w_y$ and divide by $\langle w^2 \rangle$
3: radial PSD $P^{(r)}(k_i)$ on integer shells with mode counts $M_i^{(r)}$; compute $Q^{(r)}$ the same way if $Q$ comes from a reference image
4: **end for**
5: $\bar{P}(k_i) \leftarrow \text{mean}_r\, P^{(r)}(k_i),\ M_i \leftarrow \sum_r M_i^{(r)}$
6: $\mathcal{K} \leftarrow \{k_i \in [k_{\min}, k_{\max}]\}$, truncated at the first shell with $\bar{P}(k_i) < \delta \max \bar{P}$
7: $y_i \leftarrow \log \bar{P}(k_i) - \log Q(k_i)$ for $k_i \in \mathcal{K}$
8: **for** each class $c \in \mathcal{C}$ **do**
9: weighted least squares $y_i = \beta_0 + \sum_j \beta_j\, \phi_{c,j}(k_i)$ with weights $M_i$ (grid search over $\sigma$ for Gaussian); record $\text{BIC}_c$
10: **end for**
11: $c^\star \leftarrow \arg\min_c \text{BIC}_c$ ▷ $\beta_j = 2Tm_j$
12: **if** $c^\star =$ flat **then return** MAGNITUDE-BLIND (phase-only or pure loss)
13: **end if**
14: $\widehat{c_0 T} \leftarrow$ Eq. (7) or Eq. (8)
15: **return** class $c^\star$; ratios $\beta_j/\beta_{j'}$; shape parameters; $m_j/c_0 = \beta_j/(2\,\widehat{c_0 T})$

### D.4 Internal time calibration in the Swift–Hohenberg model

The Swift–Hohenberg operator illustrates how model structure can remove the otherwise unavoidable temporal-scale ambiguity. Consider

$$\partial_t u = Ru - \left(1 + \ell^2 \nabla^2\right)^2 u. \tag{10}$$

Its real Fourier symbol is

$$\mu(k) = R - \left(1 - \ell^2 k^2\right)^2 \tag{11}$$

$$= (R-1) + 2\ell^2 k^2 - \ell^4 k^4. \tag{12}$$

Under unknown elapsed time, the observed coefficients of $1$, $k^2$, and $k^4$ are therefore

$$\widetilde{c}_0 = s(R-1), \qquad \widetilde{c}_1 = 2s\ell^2, \qquad \widetilde{c}_2 = -s\ell^4. \tag{13}$$

The last two coefficients satisfy a model-specific nonlinear constraint. Indeed,

$$\ell^2 = -\frac{2\widetilde{c}_2}{\widetilde{c}_1}, \tag{14}$$

and substitution into $\widetilde{c}_1 = 2s\ell^2$ gives

$$s = \frac{\widetilde{c}_1}{2\ell^2} = -\frac{\widetilde{c}_1^2}{4\widetilde{c}_2}. \tag{15}$$

Hence $\Delta t = s/2$, and the reaction parameter follows from $R = 1 + \widetilde{c}_0/s$. Elapsed time becomes identifiable here not because the coefficient vector lies on a one-parameter ray, but because the fixed normalization of the Swift–Hohenberg operator imposes a nonlinear relationship between its $k^2$ and $k^4$ coefficients, which provides an internal calibration of the otherwise unknown scale. The calibration relies on that normalization being known: if the coefficient of $\nabla^4$ in the operator were itself a free parameter, $\ell$ and $s$ would no longer be separable.

The same calibration carries over to a single frame. There the spectral shape $h(k) = 2T[\mu(k) - \mu(0)] = 2T(2\ell^2 k^2 - \ell^4 k^4)$ plays the role of $y(k) - \widetilde{c}_0$, so a known initial structure factor already yields both $\ell$ and the elapsed time $T$ (Proposition 2b). Either calibration of the level, Eq. (7) or Eq. (8), then also gives $c_0 T = (R-1)T$, and hence $R$ (Proposition 2c). A second frame is therefore not needed to separate the two dissipative coefficients. What it provides is independence from the initial state, since $P(k, 0)$ cancels in Eq. (3).

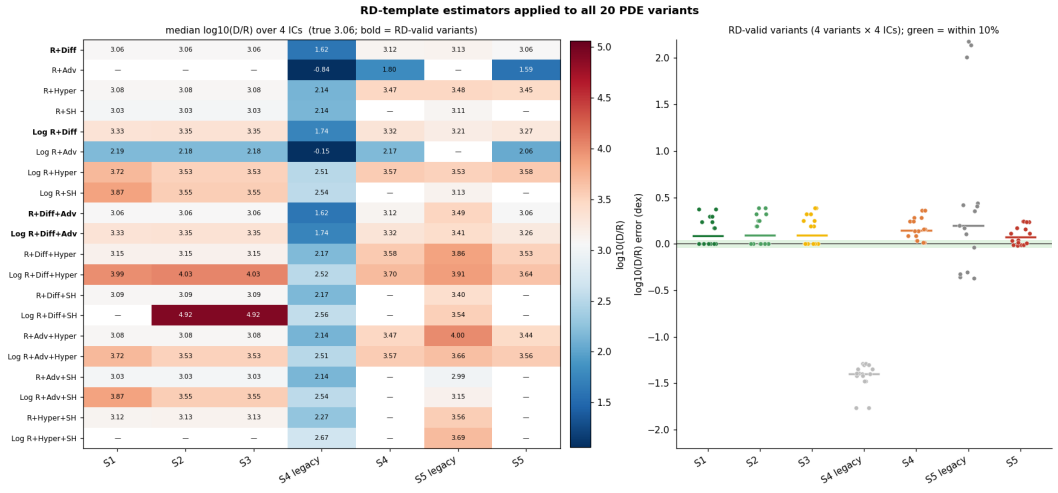



Figure S9: **Reaction–diffusion template recovery across PDE families and observation protocols. (A)** Median inferred $\log_{10}(D/R)$ across four initial spatial organizations when the reaction–diffusion template is applied across the candidate PDE variants. Columns correspond to observation protocols S1–S5; entries report the median recovered value and dashes indicate cases for which the reaction–diffusion ratio is not returned. The true reaction–diffusion value is $\log_{10}(D/R) = 3.06$. Bold row labels denote PDE variants for which $D/R$ is a valid identifiable target. The comparison illustrates that a numerical $D/R$ estimate should be interpreted only when the assumed spectral model supports that quantity; applying the same template to a different operator can produce a precise numerical value that does not correspond to the intended physical ratio. **(B)** Recovery error for the reaction–diffusion-valid variants across the four initial spatial organizations. Points show individual estimates and horizontal bars summarize recovery within each protocol; the shaded band marks estimates within 10% of the true $D/R$. S1–S3 and the revised single-frame estimators S4–S5 recover the identifiable ratio substantially more consistently than the persistent-source shoulder estimators when their generative assumptions are violated. Together, these results emphasize that parameter recovery requires both an identifiable representation and a generative model consistent with the observed dynamics.

### D.5 Parameter confounding and observable combinations

The injectivity condition clarifies why parameter recovery cannot in general be decided by assigning one physical parameter to each power of $k$. Different mechanisms may contribute to the same spectral coefficient, and the observable then constrains only the combinations through which those parameters enter the full symbol. For example, when diffusion and a Swift–Hohenberg contribution occur together, $\mu(k) = (R-1)+(2\ell^2-D)k^2-\ell^4k^4$, and the $k^2$ coefficient contains the combination $-D + 2\ell^2$. Whether $D$ and $\ell$ can nevertheless be separated depends on the remaining coefficients and on the amount of temporal information. With known elapsed time the map $(R, D, \ell) \mapsto \mu$ is injective for $\ell > 0$: $\ell$ follows from the $k^4$ coefficient, then $D$ and $R$. Under unknown elapsed time the same three observed coefficients must determine four unknowns $(R, D, \ell, s)$, which leaves a one-dimensional fibre, so $D$ and $\ell$ are no longer separable. A single unknown-time frame identifies the same wavenumber-dependent information as two frames at an unknown interval, provided the initial structure factor is known (Proposition 2b), and it inherits the same fibre. With the level calibrated by Eq. (7) or Eq. (8), it also identifies ratios to the growth rate (Proposition 2c). The same principle extends beyond polynomial symbols: parameters such as the fractional exponent $\alpha$ or the interaction length $\sigma$ alter the functional shape of $\mu$, so they are identifiable whenever distinct values generate distinct shapes up to the unknown common scale—from two frames, or from one frame given a known initial structure factor.

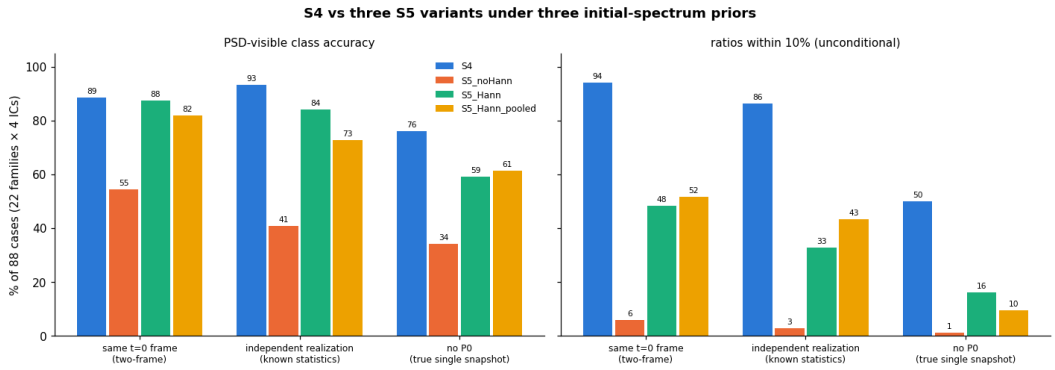



Figure S10: **Single-frame recovery depends on initial-spectrum information and regional preprocessing.** Comparison of whole-frame inference (S4) with three regional implementations of S5 across 22 PDE families and four initial spatial organizations (88 family–organization cases). **(A)** Accuracy of PSD-visible operator-class identification. **(B)** Fraction of cases in which an identifiable coefficient ratio is recovered within 10% of its true value, evaluated unconditionally so that class-selection failures also count as failures. Columns distinguish three levels of information about the initial spectrum: the same initial frame is available, an independent realization provides its statistics, or no initial-spectrum information is supplied. Without tapering, regional analysis suffers substantial spectral leakage. Applying a Hann window markedly improves S5 class identification (e.g. from 55% to 88% when the initial frame is available, and from 41% to 84% with an independent realization), demonstrating that the difference is one of practical rather than structural identifiability. Whole-frame S4 is consistently the most accurate when sufficient initial-spectrum information is available, reaching 89–93% class accuracy and 86–94% recovery of identifiable ratios within 10%. When no initial-spectrum information is available, performance deteriorates for both S4 and S5, consistent with the dependence of single-frame inference on the initial-state model established in Proposition 2. Regional pooling can improve ratio recovery in some settings but does not uniformly improve operator-class selection.

### D.6 Why multiple mechanisms need not imply multiple identifiable parameters

The appearance of a physical parameter in a PDE does not guarantee that it can be recovered independently: structural identifiability is set by injectivity of the *complete* parameter-to-symbol map. The clearest case is a model with reaction, diffusion, hyperdiffusion, and Swift–Hohenberg terms, whose real symbol is

$$\mu(k) = (R-1) + (2\ell^2 - D)k^2 - (\ell^4 + \nu)k^4. \tag{16}$$

Even measured perfectly under S1/S2, the PSD determines only the three coefficients $C_0 = R-1$, $C_2 = 2\ell^2 - D$, $C_4 = -(\ell^4+\nu)$, while the model has four unknowns $(R, D, \nu, \ell)$. Thus $R = 1+C_0$ is fixed but $2\ell^2 - D = C_2$ and $\ell^4 + \nu = -C_4$ admit a one-parameter family of solutions, so $D$, $\nu$, and $\ell$ are not all separable from PSD even with noiseless, known-time data. The reaction–loss confounding $R - \eta$ (main text) is the elementary case of the same phenomenon. Identifiability is therefore determined from the complete map $\boldsymbol{\theta} \mapsto \lambda(\mathbf{k}; \boldsymbol{\theta})$, or $\boldsymbol{\theta} \mapsto \mu(\mathbf{k}; \boldsymbol{\theta})$ when only PSD is observed—not by assigning one physical parameter to each power of $k$.

## E Estimators

Algorithms 2 and 3 give the magnitude-channel time-series estimators (S1–S3), Algorithm 4 the phase-channel estimator, and Algorithm 5 the model-selection procedure that returns an observational equivalence class. The single-frame estimator is Algorithm 1 of the main text; its settings are given in Section G. Algorithm 6 is the source-driven shoulder estimator, which is the appropriate single-frame estimator under a persistent-source model (Section G.1) and a negative control for data from $\partial_t \rho = \mathcal{L}\rho$ (main text, §4.5).

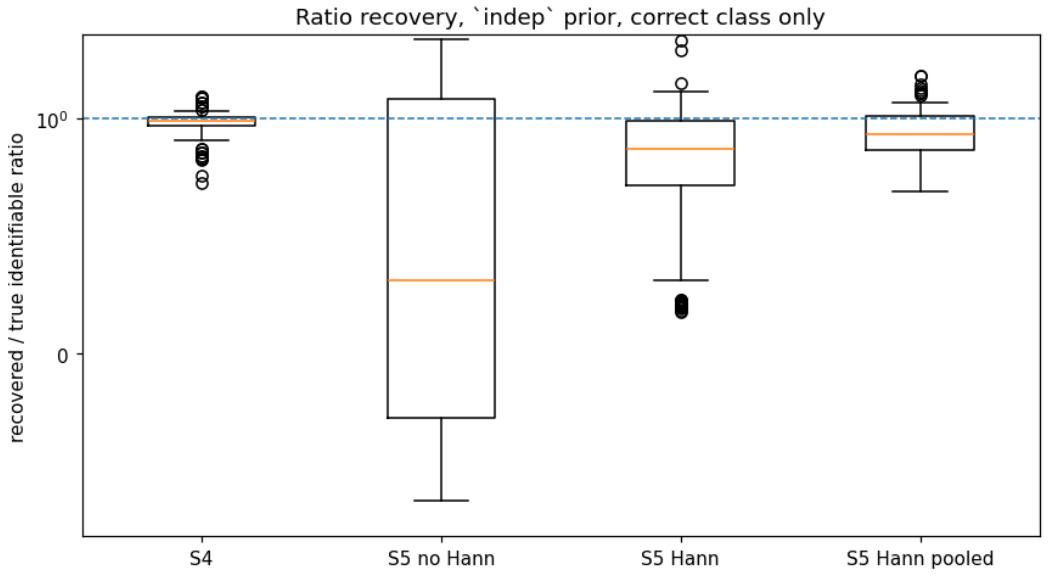



Figure S11: **Hann windowing and regional pooling improve identifiable-ratio recovery in single-frame regional inference.** Distribution of recovered-to-true identifiable coefficient ratios across the candidate PDE families and initial spatial organizations when the initial spectrum is characterized from an independent realization. Results are restricted to cases in which the correct PSD-visible operator class was selected. The dashed line denotes exact recovery ($\widehat{r}/r_{\mathrm{true}} = 1$). Whole-frame inference (S4) produces estimates tightly concentrated around the true ratios. Regional inference without tapering (S5, no window) is poorly conditioned and exhibits large dispersion and sign-reversal errors, consistent with spectral leakage from nonperiodic ROI boundaries. Applying a Hann window substantially reduces this variability, and pooling Hann-windowed ROI spectra further concentrates the recovered ratios near the true values. These results show that S4 and S5 have the same structural identifiability but can differ substantially in practical accuracy because of boundary leakage, reduced spectral support, and regional variability.

**Algorithm 2** Magnitude channel, full time series (protocol S1).

**Require:** frames $\{u(\cdot, t_j)\}_{j=1}^{T}$; symbol $\mu(k;\theta)$; band $[0, k_{\max}]$; threshold $\tau_t$; bins $N_b$

1: **for** each frame $j$ **do** radial PSD $P(k_i, t_j)$ by 2D FFT, $|\widehat{u}|^2$, averaged over $N_b$ shells $k_i = |\mathbf{k}|$
2: **end for**
3: **for** each bin $k_i$ **do** fit $\log P(k_i, t_j) = \alpha_i + 2\mu_i t_j$ (OLS); set $\widehat{\mu}(k_i) \leftarrow \mu_i$, record $r_i^2$
4: **end for**
5: $\mathcal{K} \leftarrow \{k_i : r_i^2 > \tau_t,\ 0 < k_i < k_{\max}\}$
6: $\widehat{\theta} \leftarrow \arg\min_\theta \sum_{k_i \in \mathcal{K}} (\widehat{\mu}(k_i) - \mu(k_i;\theta))^2$ ▷ closed form for flat / power-confounded families
7: **return** $\widehat{\theta}$, fit $R^2$

## F PHASE CHANNEL: DERIVATIONS

To see the source of the ambiguity, consider first one-dimensional linear advection,

$$\partial_t u = -c\,\partial_x u. \tag{17}$$

Under the Fourier convention $\partial_x \mapsto ik$, its symbol is

$$\lambda_{\mathrm{adv}}(k) = -ick, \tag{18}$$

and hence

$$\widehat{u}(k,t) = \widehat{u}(k,0)e^{-ickt}. \tag{19}$$

The factor $e^{-ickt}$ has unit magnitude,

$$\left|e^{-ickt}\right|^2 = 1,$$

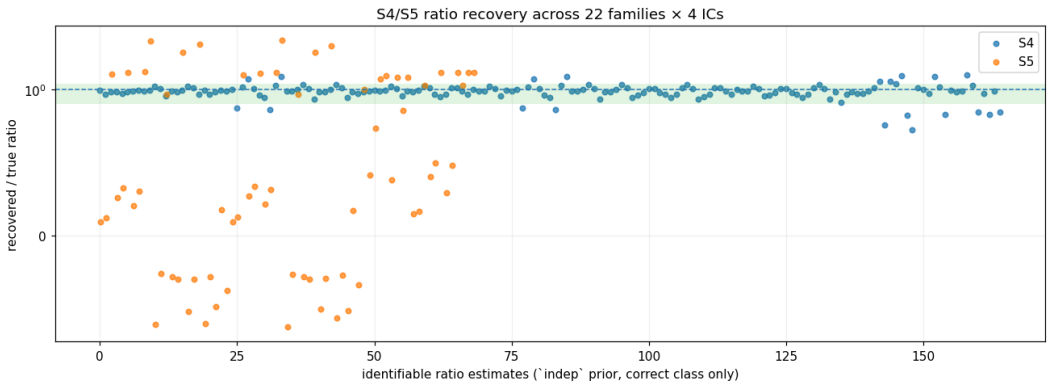



Figure S12: **Accuracy of identifiable ratio recovery from a single unknown-time frame after correct operator-class selection.** Recovered-to-true ratios for identifiable parameter combinations across the 22 candidate PDE families and four initial spatial organizations, restricted to cases in which the PSD-visible operator class was correctly identified. Blue and orange points denote whole-frame (S4) and regional (S5) inference, respectively. The dashed line denotes exact recovery ($\widehat{r}/r_{\text{true}} = 1$), and the shaded band indicates estimates within 10% of the true ratio. Whole-frame inference produces ratios concentrated near the true values, whereas regional inference under the independent initial-spectrum prior is substantially more variable and includes large-magnitude and sign-reversal errors. Thus, correct identification of the PSD-visible operator class does not by itself guarantee accurate estimation of its identifiable coefficient ratios. The difference between S4 and S5 is practical rather than structural: both protocols support the same identifiable quantities, but regional windowing, reduced low-wavenumber support, and local variation in the initial spectrum can substantially affect estimation accuracy.

**Algorithm 3** Two-frame estimator: known (S2) and unknown (S3) elapsed time.

**Require:** frames $u(\cdot, t_1), u(\cdot, t_2)$; band $[0, k_{\max}]$

1: compute $P(k, t_1), P(k, t_2)$; keep $k$ with both $> 0$ and $k < k_{\max}$
2: **if** $\Delta t$ known **then** $\widehat{\mu}(k) \leftarrow \frac{1}{2\Delta t} \log \frac{P(k,t_2)}{P(k,t_1)}$; fit as in Alg. 2
3: **else** (unknown $\Delta t$)
4: $\quad y(k) \leftarrow \log[P(k, t_2)/P(k, t_1)] = s\,\mu(k)$, $s = 2\Delta t$; fit $y = \sum_m c_m k^{2m}$
5: $\quad$ report ratios $c_m/c_0$ for singly-owned powers; net coefficient + CONFOUNDED flag for shared powers
6: **end if**
7: **return** recovered quantities, $R^2$

so

$$P(k, t) = |\widehat{u}(k, t)|^2 = |\widehat{u}(k, 0)|^2 = P(k, 0). \tag{20}$$

Thus, changing the advection velocity changes the spatial position of the field but not its PSD magnitude. A magnitude-only representation therefore maps all advection velocities to the same observable PSD.

The same phenomenon occurs for odd-order dispersion. For example, if

$$\partial_t u = \beta_3\, \partial_x^3 u, \tag{21}$$

then

$$\lambda_{\text{disp}}(k) = -i\beta_3 k^3, \tag{22}$$

because $(ik)^3 = -ik^3$. The resulting evolution again multiplies each Fourier mode by a unit-magnitude complex factor. Unlike advection, however, the phase shift now scales as $k^3$, so different wavenumbers propagate with different phase evolution. The physical fields can therefore evolve very differently even though their PSD magnitudes remain unchanged.

More generally, writing

$$\lambda(k) = \mu(k) + i\omega(k), \tag{23}$$

**Algorithm 4** Phase channel: advection / odd-order dispersion.

**Require:** $\widehat{u}(\mathbf{k}, t_1), \widehat{u}(\mathbf{k}, t_2)$; $\Delta t$; basis $b(\mathbf{k}) = k_x$ or $k_x^3$

1: $C(\mathbf{k}) \leftarrow \widehat{u}(\mathbf{k}, t_2)\,\overline{\widehat{u}(\mathbf{k}, t_1)}$; $\quad \phi \leftarrow \arg C$; $\quad w \leftarrow |C|$

2: restrict to the phase-unambiguous band $\{|\phi| < \pi\}$

3: $\widehat{s} \leftarrow \arg\min_s \sum w(\mathbf{k})\,(\phi(\mathbf{k}) - s\,b(\mathbf{k}))^2 = \dfrac{\sum w\,b\,\phi}{\sum w\,b^2}$

4: **return** $c = -\widehat{s}/\Delta t$ **or** $\beta_3 = -\widehat{s}/\Delta t$

**Algorithm 5** Model selection as an observational equivalence class.

**Require:** measured $\widehat{\mu}(k)$; candidate families $\mathcal{F}$; tolerance $\varepsilon$

1: **for** each $f \in \mathcal{F}$ **do** fit $\mu(k; \theta_f)$ (Alg. 2/3); record $R_f^2$ and #free params

2: **end for**

3: $f^\star \leftarrow$ best $R^2$ at fewest parameters (parsimony tie-break)

4: $\mathcal{E} \leftarrow \{f : |\mu(\cdot; \theta_f) - \mu(\cdot; \theta_{f^\star})| < \varepsilon$ on the band, or $\mu_f$ nested in $\mu_{f^\star}\}$

5: **return** equivalence class $\mathcal{E}$

gives

$$\widehat{u}(k, t) = \widehat{u}(k, 0) e^{\mu(k)t} e^{i\omega(k)t}. \tag{24}$$

Taking the squared magnitude removes the second factor:

$$P(k, t) = P(k, 0) e^{2\mu(k)t}. \tag{25}$$

Consequently, PSD magnitude contains information about $\mu(k) = \operatorname{Re}\lambda(k)$ but none about $\omega(k) = \operatorname{Im}\lambda(k)$. This is a structural limitation of the representation rather than a failure of a particular estimator. No increase in signal quality or optimization accuracy can recover a parameter that has been removed by the magnitude operation.

Figure 1 demonstrates this directly for advection and third-order dispersion. Despite visible evolution of the underlying spatial field, the initial and final PSD magnitudes coincide. A PSD-only model selection procedure can therefore obtain excellent spectral agreement while remaining unable to determine whether these phase-only mechanisms are present. Goodness of fit should consequently be interpreted relative to the information retained by the representation, rather than as evidence that the full governing dynamics have been identified. Retaining Fourier phase through the cross-spectrum estimator (Algorithm 4) recovers the advection velocity $c$ and the dispersion coefficient $\beta_3$ directly, confirming that the information is present in the representation but absent from its magnitude.

**Recovering the missing phase channel.** The information removed by PSD magnitude is not absent from the original spatial observations. It remains encoded in the complex Fourier phase and can be accessed by comparing two frames. Given observations at $t_1$ and $t_2$, with

$$\Delta t = t_2 - t_1,$$

we define the complex cross-spectrum

$$C(k; t_1, t_2) = \widehat{u}(k, t_2)\,\overline{\widehat{u}(k, t_1)}. \tag{26}$$

Using Eq. (24),

$$C(k; t_1, t_2) = |\widehat{u}(k, 0)|^2\, e^{2\mu(k)t_1} e^{\lambda(k)\Delta t} \tag{27}$$

$$= |\widehat{u}(k, 0)|^2\, e^{\mu(k)(t_1+t_2)} e^{i\omega(k)\Delta t}. \tag{28}$$

The amplitude factor preceding the final exponential is real and non-negative. Wherever the cross-spectrum has nonzero magnitude, it therefore does not contribute to its argument, giving

$$\phi(k) := \arg C(k; t_1, t_2) = \omega(k)\Delta t \pmod{2\pi}. \tag{29}$$

Equation (29) is the phase-channel counterpart of the PSD evolution identity. From the same pair of spatial frames, the magnitude change measures

$$\frac{1}{2\Delta t} \log \frac{P(k, t_2)}{P(k, t_1)} = \mu(k), \tag{30}$$

**Algorithm 6** Source-driven shoulder estimator (persistent-source model; negative control for $\partial_t \rho = \mathcal{L}\rho$).

**Require:** one frame $u(\cdot, t)$ at unknown $t$; ROI fraction $f$ and $n \times n$ overlapping offsets ($f$=1: whole frame; $f$<1: regions); band $[k_{\min}, k_{\max}]$; floor cutoff $k_{\text{noise}}$; bins $N_b$; candidate powers $\mathcal{P} = \{1.5, 2, 4, 6\}$; persistent-source model with near-white source spectrum; tolerances $\tau_{\text{blind}}, \tau_{\text{deg}}$

1: **for** each ROI (crop of side $f \cdot L$) **do**
2: subtract mean; apply 2-D Hann window $w(\mathbf{x}) = w_x w_y$ ▷ reduce boundary-induced spectral leakage
3: radial PSD $P^{(r)}(k_i)$ over $N_b$ shells in $[k_{\min}, k_{\max}]$; floor $\varphi^{(r)} \leftarrow \text{mean}\{|\widehat{u}|^2 : |\mathbf{k}| > k_{\text{noise}}\}$
4: **end for**
5: $\bar{P}(k_i) \leftarrow \text{mean}_r\, P^{(r)}(k_i) - \text{mean}_r\, \varphi^{(r)}$ ▷ regional averaging reduces finite-window variability
6: $y_i \leftarrow \log \bar{P}(k_i)$ on $\{k_i : \bar{P}(k_i) > 0\}$
7: **for** each $p \in \mathcal{P}$ **do** ▷ free-amplitude source-driven fit, reparametrised by the shoulder $k_*$
8: $m_p(k; A, \sigma, k_*) = \log A + 2 \log \left| \dfrac{e^{\sigma\,(1-(k/k_*)^p)} - 1}{\sigma\,(1-(k/k_*)^p)} \right|, \quad \sigma = c_0 \Delta t$
9: $(\widehat{A}, \widehat{\sigma}, \widehat{k}_*) \leftarrow \arg\min \sum_i \big(y_i - m_p(k_i; \cdot)\big)^2$ via multistart over $k_*$ and $\sigma$
10: record $R_p^2$ and ratio $\rho_p \leftarrow \widehat{k}_*^{-p}$
11: **end for**
12: fit flat model ($R^2_{\text{flat}}$) and two-term symbol $\mu = c_0(1 + b_2 k^2 + b_4 k^4)$ ($R_2^2$)
13: $p^\star \leftarrow \arg\max_p R_p^2$
14: **if** $R^2_{p^\star} < \tau_{\text{blind}}$ **then return** MAGNITUDE-BLIND ▷ no shoulder: phase-only or pure loss
15: **else if** $R_2^2 - R^2_{p^\star} > \tau_{\text{deg}}$ **then return** DEGENERATE (MULTI-TERM): effective shoulder only, ratio not separable
16: **else return** operator order $p^\star$, ratio $\rho_{p^\star} = \widehat{k}_*^{-p^\star}$, $R^2_{p^\star}$ ▷ single-term: quantitative under the persistent-source model
17: **end if**

whereas the cross-spectrum phase measures

$$\frac{\phi(k)}{\Delta t} = \omega(k) \tag{31}$$

within the phase-unambiguous regime. Magnitude and phase therefore provide complementary observations of the real and imaginary components of the same complex Fourier symbol.

For advection,

$$\omega_{\text{adv}}(k) = -ck, \tag{32}$$

and hence

$$\phi_{\text{adv}}(k) = -c\,\Delta t\,k. \tag{33}$$

The phase is linear in $k$, and its slope is the negative displacement $c\Delta t$ accumulated between the two observations. If $\Delta t$ is known, the velocity follows directly from this slope.

For third-order dispersion under the sign convention above,

$$\omega_{\text{disp}}(k) = -\beta_3 k^3, \tag{34}$$

so that

$$\phi_{\text{disp}}(k) = -\beta_3 \Delta t\, k^3. \tag{35}$$

The phase is therefore linear when plotted against $k^3$, allowing the dispersive coefficient to be estimated from its slope.

More generally, when both mechanisms are present,

$$\omega(k) = -ck - \beta_3 k^3, \tag{36}$$

and the observed phase satisfies

$$\phi(k) = -c\Delta t\,k - \beta_3 \Delta t\, k^3 \pmod{2\pi}. \tag{37}$$

Thus, advection and dispersion can be estimated jointly by fitting the phase to the basis functions $k$ and $k^3$.

For a single phase basis $b(k)$, we estimate its coefficient using a power-weighted least-squares fit (Algorithm 4),

$$\widehat{s} = \arg\min_{s} \sum_{k\in\mathcal{K}} w(k)\,[\phi(k) - s\,b(k)]^2\,, \qquad w(k) = |C(k)|, \tag{38}$$

where $\mathcal{K}$ denotes the retained wavenumber band. For advection, $b(k) = k$ and

$$c = -\frac{\widehat{s}}{\Delta t},$$

whereas for the dispersive convention above, $b(k) = k^3$ and

$$\beta_3 = -\frac{\widehat{s}}{\Delta t}.$$

Weighting by cross-spectral magnitude suppresses modes with little spectral power, whose phase estimates are correspondingly unreliable. The phase channel thus recovers the magnitude-blind class of §2, but requires two frames: a single snapshot carries no accessible phase reference, so advection and dispersion remain unrecoverable from one frame regardless of the representation used.

**Phase wrapping and temporal resolution.** Because $\arg C$ is observed modulo $2\pi$, phase recovery also introduces a sampling constraint that does not appear in PSD magnitude. Direct unambiguous estimation is obtained over modes satisfying

$$|\omega(k)\Delta t| < \pi. \tag{39}$$

For larger phase increments, different accumulated phases can produce the same wrapped observation. Such modes require additional assumptions or phase-unwrapping procedures. A full time series reduces this ambiguity by allowing phase increments to be measured over shorter temporal intervals.

The $k$–$\omega$ representation shown in Fig. 1 provides a complementary view of the same information. For advection, spectral energy follows an approximately linear relation between temporal frequency and spatial wavenumber, whereas third-order dispersion produces a cubic relation. The cross-spectrum estimator extracts these relationships directly from paired observations.

Taken together, these examples demonstrate that spectral identifiability is representation dependent. PSD magnitude alone identifies only the PSD-visible component $\mathrm{Re}\,\lambda$; phase-sensitive observations provide access to $\mathrm{Im}\,\lambda$. Combining the two channels therefore distinguishes dynamics that are observationally equivalent under magnitude alone.

**Magnitude and phase provide complementary information about the dynamics.** The cross-spectrum phase isolates the imaginary component $\omega(k) = \mathrm{Im}\,\lambda(k)$ and therefore contains no direct information about mechanisms that enter only through $\mu(k) = \mathrm{Re}\,\lambda(k)$. Reaction provides the simplest example: the growth rate $R$ contributes to the real part of the Fourier symbol and changes Fourier magnitude, but it does not contribute to $\arg C(k; t_1, t_2)$. The phase channel is therefore insensitive to $R$, just as PSD magnitude is insensitive to purely phase-generating mechanisms such as advection and odd-order dispersion.

This gives the two spectral channels complementary, rather than redundant, roles. For a model with

$$\lambda(k) = \underbrace{R - Dk^2 - \nu k^4 - \cdots}_{\mu(k)} + i\,\underbrace{\left(-ck - \beta_3 k^3 - \cdots\right)}_{\omega(k)}, \tag{40}$$

a pair of time-separated observations provides

$$\frac{1}{2\Delta t}\log\frac{P(k,t_2)}{P(k,t_1)} = \mu(k), \tag{41}$$

$$\frac{1}{\Delta t}\arg C(k;t_1,t_2) = \omega(k) \pmod{\frac{2\pi}{\Delta t}}. \tag{42}$$

Thus, when $\Delta t$ is known, magnitude evolution constrains the PSD-visible parameters through the map $\boldsymbol{\theta}_\mu \mapsto \mu(k;\boldsymbol{\theta}_\mu)$, while phase evolution constrains the phase-visible parameters through $\boldsymbol{\theta}_\omega \mapsto$

$\omega(k;\boldsymbol{\theta}_\omega)$. Individual physical parameters are recoverable when the corresponding parameter-to-symbol map is injective over the observed spectral range.

For simple separable models, this division is particularly transparent. Reaction, diffusion, and hyperdiffusion contribute respectively to the constant, $k^2$, and $k^4$ components of $\mu(k)$, whereas advection and third-order dispersion contribute $k$ and $k^3$ dependences to $\omega(k)$. The two components can therefore be estimated from the same pair of spatial frames without requiring the parameters of one channel to be estimated before those of the other. In this sense, the magnitude and phase channels provide complementary projections of the same underlying complex Fourier dynamics.

When $\Delta t$ is unknown, both channels inherit the same temporal-scale ambiguity. The magnitude observation determines $2\Delta t\,\mu(k)$—the projective symbol of Proposition 1— while the cross-spectrum phase determines $\Delta t\,\omega(k)$ modulo $2\pi$. Consequently, phase-only parameters such as $c$ and $\beta_3$ are generally recovered as the products $c\Delta t$ and $\beta_3\Delta t$, rather than as absolute rates. If $\Delta t$ is known independently, or if model structure provides an internal calibration of the temporal scale as in the Swift–Hohenberg example, the absolute parameters in both channels can then be recovered, subject to the structural and practical identifiability conditions of their respective parameter-to-symbol maps. Both channels, however, require two frames: a single snapshot provides no phase reference, and its magnitude gives only the shape $h(k) = 2T[\mu(k) - \mu(0)]$ relative to a model of the initial spectrum (Proposition 2), so neither $\omega(k)$ nor, without further calibration, the full symbol $\mu(k)$ is accessible from one frame.

## G SINGLE-FRAME INFERENCE: DESIGN AND ADDITIONAL RESULTS

**Design.** The held-out single-frame validation (main text, §4.4) uses a periodic $2.5\,\mathrm{mm}$ domain of $256^2$ pixels ($9.8\,\mu$m per pixel), seeded with 1500 cells of one-pixel radius and evolved exactly in Fourier space to $T = 0.75\,\mathrm{d}$ using the full complex Fourier symbol while retaining the spatial mean. Four non-Poisson initial organizations are considered: hard-core "random spaced" cells with an $18\,\mu$m minimum separation, a single central cluster ($350\,\mu$m s.d.), 20 clusters ($120\,\mu$m s.d.), and 20 clusters plus 500 isolated cells. Evaluation uses eight independent realizations of each organization (seeds 1–8); four additional realizations (seeds 101–104) are used only for choosing estimator settings.

The initial-spectrum model $Q(k)$ in Algorithm 1 is supplied in three ways: (i) from the initial frame of the same realization, providing a two-frame control in which realization-specific initial-spectrum fluctuations are known; (ii) from an independent realization of the same spatial organization (seed $+1000$), representing known initial-spectrum statistics; or (iii) not supplied, corresponding to a white-spectrum assumption. The spectral level is calibrated independently from the known initial density using Eq. (8).

The fitting band is $[\max(0.006,\, 2\pi/L_{\mathrm{ROI}}),\, 0.09]$ rad/$\mu$m. The spectral floor fraction is $\delta = 10^{-8}$ for the whole periodic frame (S4) and $\delta = 10^{-6}$ for regional analyses; these values were selected using only the tuning realizations from candidates $10^{-3}$–$10^{-10}$ and no floor. Regional inference is evaluated in three forms: `S5_noHann`, using untreated ROIs; `S5_Hann`, applying a separable two-dimensional Hann window to the mean-subtracted fluctuations of each ROI before the Fourier transform; and `S5_Hann_pooled`, which sums the windowed ROI power spectra before performing a single spectral-class fit rather than fitting individual ROIs and combining their classifications. The S5 analyses use nine overlapping square ROIs of 70% frame width arranged on a $3 \times 3$ grid of offsets. The unwindowed image mean is retained for growth calibration so that the comparison isolates the effect of regional spectral preprocessing.

For the reaction–diffusion $D/R$ analysis in Table S1, the regional estimator instead pools a disjoint $3 \times 3$ tiling. For the 22-family validation, class accuracy is evaluated over the $22 \times 4 = 88$ family–organization combinations for each initial-spectrum condition. Ratio accuracy is reported unconditionally: a ratio is counted as successfully recovered only when the PSD-visible class is correctly identified, the recovered ratio has the correct sign, and its magnitude is within 10% of the true value; a class-selection failure therefore counts as a ratio-recovery failure. Conditional ratio errors are additionally reported among correctly classified cases.

Table S1: Single-frame recovery of $\log_{10}(D/R)$ for reaction–diffusion: absolute error in decades over 8 held-out realizations $\times$ 4 non-Poisson organizations (32 frames). S4/S5 use Algorithm 1 with a white initial spectrum and the density calibration, Eq. (8). "Shoulder" denotes the source-driven template, Algorithm 6.

| Estimator | Median | 90th percentile | Maximum |
|---|---|---|---|
| S1, S2, S3 (time-resolved / two frames) | 0.001 | 0.001 | 0.002 |
| S4, whole frame | 0.063 | 0.188 | 0.284 |
| S5, pooled Hann-windowed regions | 0.016 | 0.049 | 0.101 |
| Shoulder template, whole frame | 1.45 | 1.76 | 1.77 |
| Shoulder template, $3 \times 3$ regions (signed range) | | $-0.45$ to $+1.61$ | |

**Sensitivity to the floor fraction.** For the white-spectrum $D/R$ estimate, tightening the floor to $10^{-2}$ inflates the whole-frame error for clustered seedings (median up to 0.9 decades), whereas $10^{-3}$ and $10^{-4}$ give similar results ($10^{-4}$ somewhat better for S4: median 0.07 vs. 0.18 decades for many clusters). The regional estimate is insensitive between $10^{-3}$ and $10^{-4}$. Without any floor cut, strongly decaying spectra are fitted into numerical round-off and misclassified.

Table S2: Per-family selection of the correct spectral-shape class from one frame with a known structure factor (independent realization), pooled over 8 held-out realizations $\times$ 4 organizations (32 frames per family). Classes: flat (phase-only families and pure loss), $k^{1.5}$, $k^2$, $k^4$, $k^6$, $k^2{+}k^4$, Gaussian.

| Family | True class | S4 | S5 (Hann) | S5 (Hann, pooled) |
|---|---|---|---|---|
| R+Diff, R+Diff+Adv | $k^2$ | 0.94 | 0.88–0.91 | 0.59–0.62 |
| R+Hyper, R+Adv+Hyper | $k^4$ | 0.94 | 0.78–0.81 | 0.66 |
| R+Sixth | $k^6$ | 1.00 | 0.81 | 0.81 |
| R+FracDiff | $k^{1.5}$ | 0.81 | 0.59 | 0.59 |
| R+Nonlocal | Gaussian | 0.75 | 0.50 | 0.62 |
| SH, KS, CH and all $k^2{+}k^4$ combinations | $k^2{+}k^4$ | 1.00 | 0.84–1.00 | 0.84–1.00 |
| R+Adv, R+Disp, R+Loss | flat | 0.81 | 0.78–0.81 | 0.00–0.44 |

**Recovery of native parameters.** On the ten-model parameter set used for the known-time analysis (so that all five protocols are compared on the same parameters), the whole-frame estimator with a known structure factor recovered the ratio within 10% in the following fractions of 36 frames (nine realizations $\times$ four organizations): $D/R$ in reaction–diffusion 0.92; $\nu/R$ in hyperdiffusion 0.89; the Swift–Hohenberg length $\ell$ 0.94; in combined diffusion–hyperdiffusion $D/R$ 0.42 and $\nu/R$ 0.78. The pooled regional estimator reached 0.61–0.67 for $D/R$, 0.69 for $\ell$, and $\leq 0.08$ for $\nu/R$. $D$ and $\ell$ in diffusion–Swift–Hohenberg were never separated, as the single-frame map is not injective there (Section D.5).

**Misclassifications and weighting.** Errors concentrate between neighbouring classes: $k^4$ read as $k^2{+}k^4$, $k^{1.5}$ as $k^2$, $k^2{+}k^4$ as $k^6$, the Gaussian kernel as $k^2{+}k^4$ or $k^{1.5}$, and flat as $k^6$ or $k^{1.5}$. Replacing the mode-count weights by an empirical per-shell variance (from two independent halves of each ring, with a 5% systematic floor, and BIC $= \chi^2 + q \ln n$) did not improve selection ($85 \pm 7\%$ vs. $94 \pm 3\%$ for S4; $77 \pm 3\%$ vs. $80 \pm 6\%$ for pooled S5), so the simpler weighting is used throughout.

**Logistic growth.** For logistic reaction–diffusion (hard-core organization), S1 overestimates $\log_{10}(D/R)$ by 0.173, 0.048 and 0.008 decades when seeded at $K$, $0.3K$ and $0.05K$. The dispersion fit keeps $R^2 = 1.00$ in all three cases and the per-mode time-linearity has median $R^2 \geq 0.9997$ (minimum 0.78, 0.99 and 0.67; linear reference minimum 0.999), so neither statistic flags the bias, whereas its dependence on density does.

### G.1 Single-frame inference under a persistent-source model

If cells act as a persistent source rather than as an initial condition, the source-driven template of Algorithm 6 is the correct generative model, and the shoulder estimator is the appropriate single-

frame estimator. This subsection evaluates it in that regime; which model applies is an assumption about the biology (main text, §2).

**Setup.** Each frame is generated by the persistent-source model that underlies the template: the cells of an initial organization act as a persistent source, $\partial_t u = \mathcal{L}u + S(\mathbf{x})$ with $u(\cdot, 0) = 0$, solved exactly per Fourier mode as $\widehat{u}(\mathbf{k}, T) = \widehat{S}(\mathbf{k})\,(e^{\lambda(\mathbf{k})T} - 1)/\lambda(\mathbf{k})$ with the *full* complex symbol $\lambda = \mu + i\omega$, so advection ($c = 120\,\mu\mathrm{m\,d^{-1}}$) and dispersion ($\beta_3 = 8 \times 10^4\,\mu\mathrm{m^3\,d^{-1}}$) are simulated rather than omitted. We use the four initial organizations of the main text (10 $\mu$m cells; 18 $\mu$m minimum spacing; 120 $\mu$m clusters; a single 700 $\mu$m cluster) on a 5000 $\mu$m domain ($2000^2$, 2.5 $\mu$m pixels), with cell and cluster counts scaled by area, plus a Poisson point-source control whose spectrum is white by construction. On the 2500 $\mu$m domain used for S1–S3, even the white control selects the wrong operator order, because the lowest radial shells hold too few modes to fix the plateau below the shoulder (cf. §J.2). Each family is run with three realizations, each at a different unknown observation time $T \in \{1.0, 1.4, 2.0\}$ d, clean and with 5% Gaussian noise. Algorithm 6 is applied exactly as stated ($f = 0.7$, $3 \times 3$ offsets, band $[0.006, 0.12]$ rad/$\mu$m, $N_b = 36$, $k_{\text{noise}} = 0.3$, $\mathcal{P} = \{1.5, 2, 4, 6\}$, $\tau_{\text{blind}} = 0.5$, $\tau_{\text{deg}} = 0.003$). Ratio errors are symmetric in log space, so $0.5\times$ and $2\times$ both count as $100\%$.

**Findings.** (i) *Near-white initial organizations.* For the random-spaced organization and the Poisson control, Algorithm 6 selects the correct spectral order in $20/24$ single-term dissipative runs. The median ratio error is approximately $20\%$ overall, with errors of $23\%$ for R+Diff, $20\%$ for R+Hyper, $32\%$ for R+Sixth, and $49\%$ for R+FracDiff. The remaining ambiguity is primarily finite-sample: competing spectral orders can produce nearly indistinguishable fits, with differences in $R^2$ below $0.002$.

(ii) *Clustered initial organizations violate the near-white prior.* Multi-cluster fields exhibit a 10–$100\times$ excess of low-wavenumber power below approximately $k = 0.025$ rad/$\mu$m (Fig. S13). The shoulder estimator can interpret this initial-state structure as dynamical spectral curvature, reducing correct-order selection to $3/12$ dissipative runs per organization. R+Diff and R+FracDiff are particularly affected because their characteristic shoulders overlap this low-$k$ excess, producing ratio biases of approximately 5–$10\times$. For R+Hyper and R+Sixth, the ratios would be recovered to within approximately 30–$50\%$ when the true order is imposed, but blind order selection often assigns $p = 1.5$. A single broad cluster gives intermediate performance ($10/12$ correct orders; median ratio error $47\%$).

(iii) *Noise is secondary to initial-spectrum misspecification.* Adding $5\%$ noise changes the aggregate recovery summaries by at most one to two percentage points in these simulations. Thus, within the tested noise regime, misspecification of the initial spatial spectrum is a larger source of error than measurement noise.

(iv) *Phase-only mechanisms can affect magnitude under persistent forcing.* This stress test uses a different observation model from the homogeneous evolution underlying the main PSD-identifiability results. Under persistent synchronous forcing, the response contains

$$\left| \frac{e^{\lambda(\mathbf{k})T} - 1}{\lambda(\mathbf{k})} \right|^2,$$

which depends on both $\mu(\mathbf{k}) = \Re\lambda(\mathbf{k})$ and $\omega(\mathbf{k}) = \Im\lambda(\mathbf{k})$. Consequently, mechanisms such as advection and odd-order dispersion, although invisible to the PSD of homogeneous linear evolution, can distort the magnitude spectrum in this persistent-source setting and bias a magnitude-only shoulder estimator. This does not contradict the structural result of the main text: the two statements correspond to different observation models.

(v) *Windowing is more important than regional averaging.* On a non-periodic 4 mm field of view, applying the shoulder estimator to the unwindowed field produces median errors of $10^3$–$3 \times 10^3\%$. A Hann taper reduces the median error to approximately $50\%$ for near-white initial organizations, whereas averaging over overlapping regions provides no consistent additional improvement (Fig. S17). This identifies boundary-induced spectral leakage, rather than the number of regional estimates alone, as an important practical limitation of the shoulder estimator.

(vi) *Multi-term operators remain outside the single-power template.* Algorithm 6 fits a single spectral power and therefore cannot separate multiple wavenumber-dependent coefficients. It returns

the multi-term verdict in 21–22 of 36 runs for near-white organizations but only 1–3 of 36 runs for clustered organizations. These failures reinforce the distinction between the structural identifiability results of the main text and the practical performance of this deliberately restricted single-power estimator.

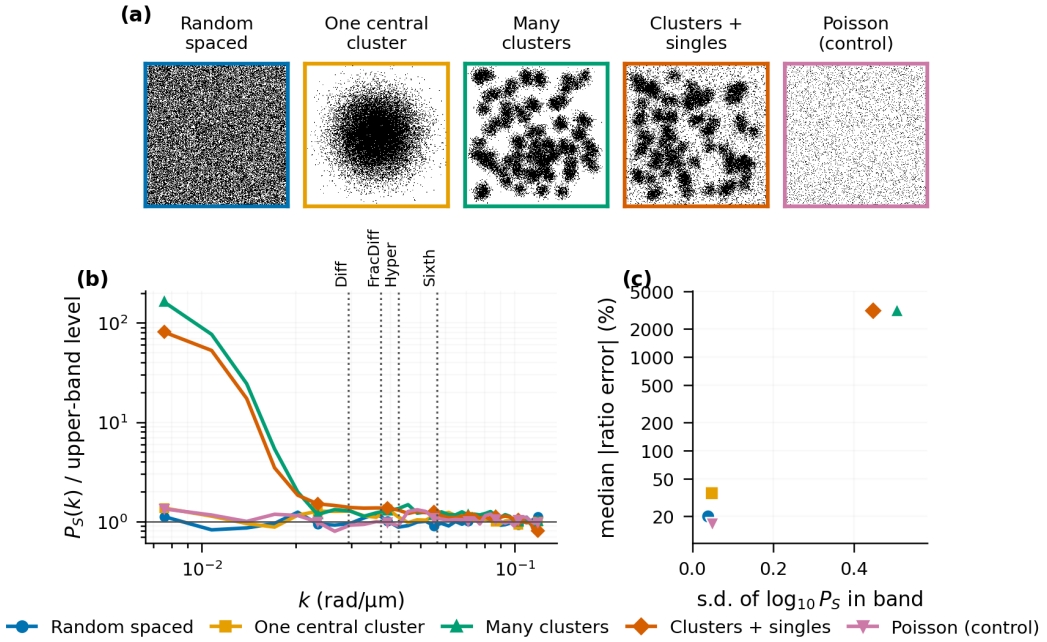



Figure S13: **Is the near-white source assumption satisfied?** (a) Source fields of the five initial organizations. (b) Their Hann-ROI averaged spectra $P_S(k)$, normalized to the upper band; the persistent-source template requires this to be flat across the fit band, where the true shoulders $k_*$ lie (dotted). (c) Prior violation (s.d. of $\log_{10} P_S$ in the band) against the median ratio error of the single-term dissipative families at the true order.

## H Nonlinear growth: derivation and numerical details

To examine this boundary, we replace linear reaction by logistic growth. For example, the reaction–diffusion equation becomes

$$\partial_t u = Ru\left(1 - \frac{u}{K}\right) + D\nabla^2 u = Ru - \frac{R}{K}u^2 + D\nabla^2 u, \tag{43}$$

where $K$ is the carrying capacity. Fourier transformation gives

$$\partial_t \widehat{u}(k,t) = \left(R - Dk^2\right)\widehat{u}(k,t) - \frac{R}{K}\widehat{u^2}(k,t). \tag{44}$$

Unlike the linear terms, the nonlinear contribution does not reduce to a scalar function of $k$ multiplying $\widehat{u}(k,t)$. Instead,

$$\widehat{u^2}(k,t) = \sum_q \widehat{u}(q,t)\widehat{u}(k-q,t) \tag{45}$$

in the discrete periodic setting, with the corresponding convolution integral in the continuum. Thus, the evolution of a mode at wavenumber $k$ depends on other Fourier modes, and the dynamics can no longer be described exactly by a time-independent scalar dispersion relation $\lambda(k)$.

This has an immediate consequence for spectral parameter inference. If the linear estimator

$$\widehat{\mu}_{\text{eff}}(k) = \frac{1}{2\Delta t}\log\frac{P(k,t+\Delta t)}{P(k,t)} \tag{46}$$

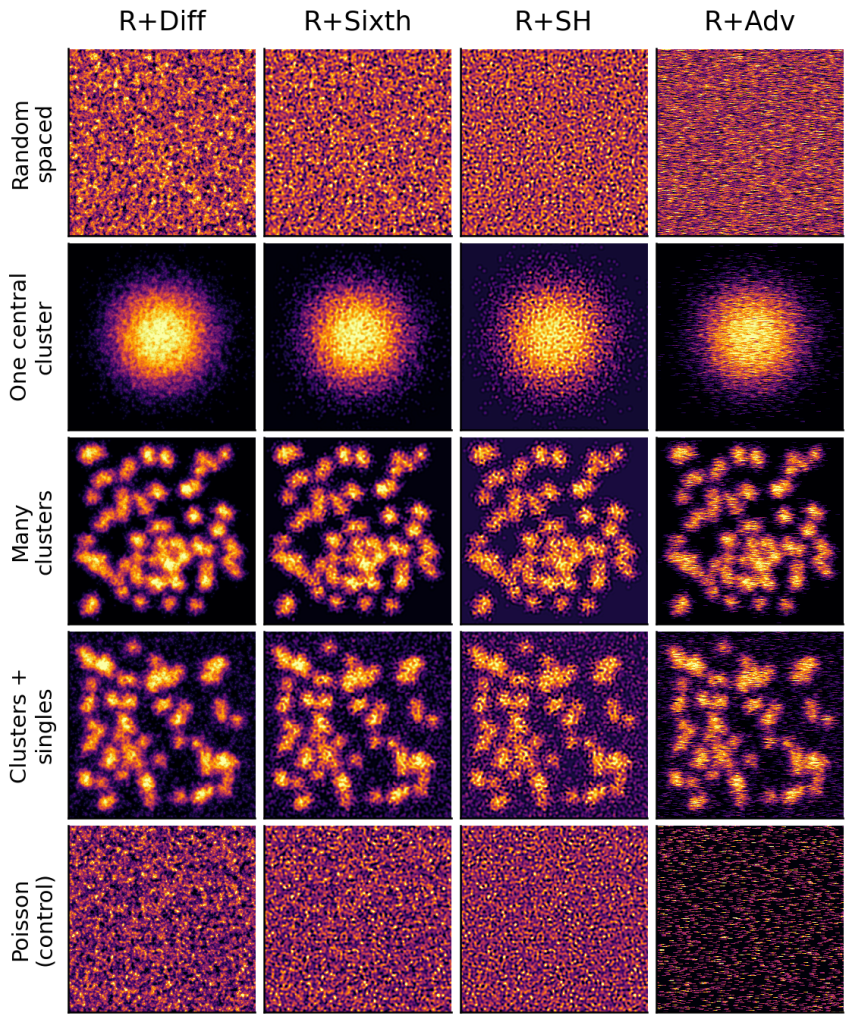



Figure S14: **Persistent-source model: single unknown-time frames.** One realization per initial organization (rows) for R+Diff, R+Sixth, R+SH and R+Adv (columns). Under the persistent-source model advection streaks the field along $\mathbf{c}$.

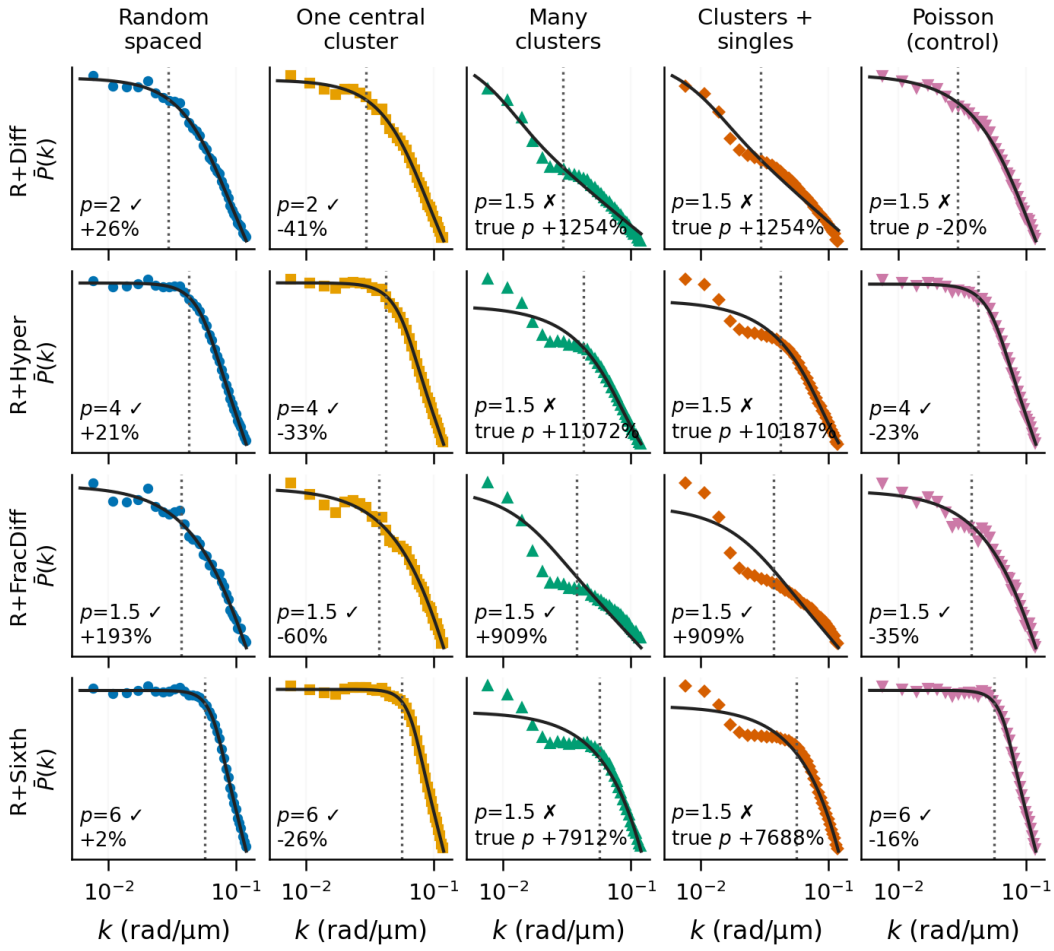



Figure S15: **Persistent-source model, regional shoulder method: fits for the single-term dissipative operators.** Averaged, floor-subtracted ROI spectra (points) and the Algorithm 6 fit at the selected order (line); dotted, true $k_* = (c_0/b)^{1/p}$. Annotations give the selected order (✓/×) and ratio error; when the order is wrong, the error at the true order is shown. Near-white organizations recover order and ratio; multi-cluster organizations add a low-$k$ excess that is read as a shoulder.

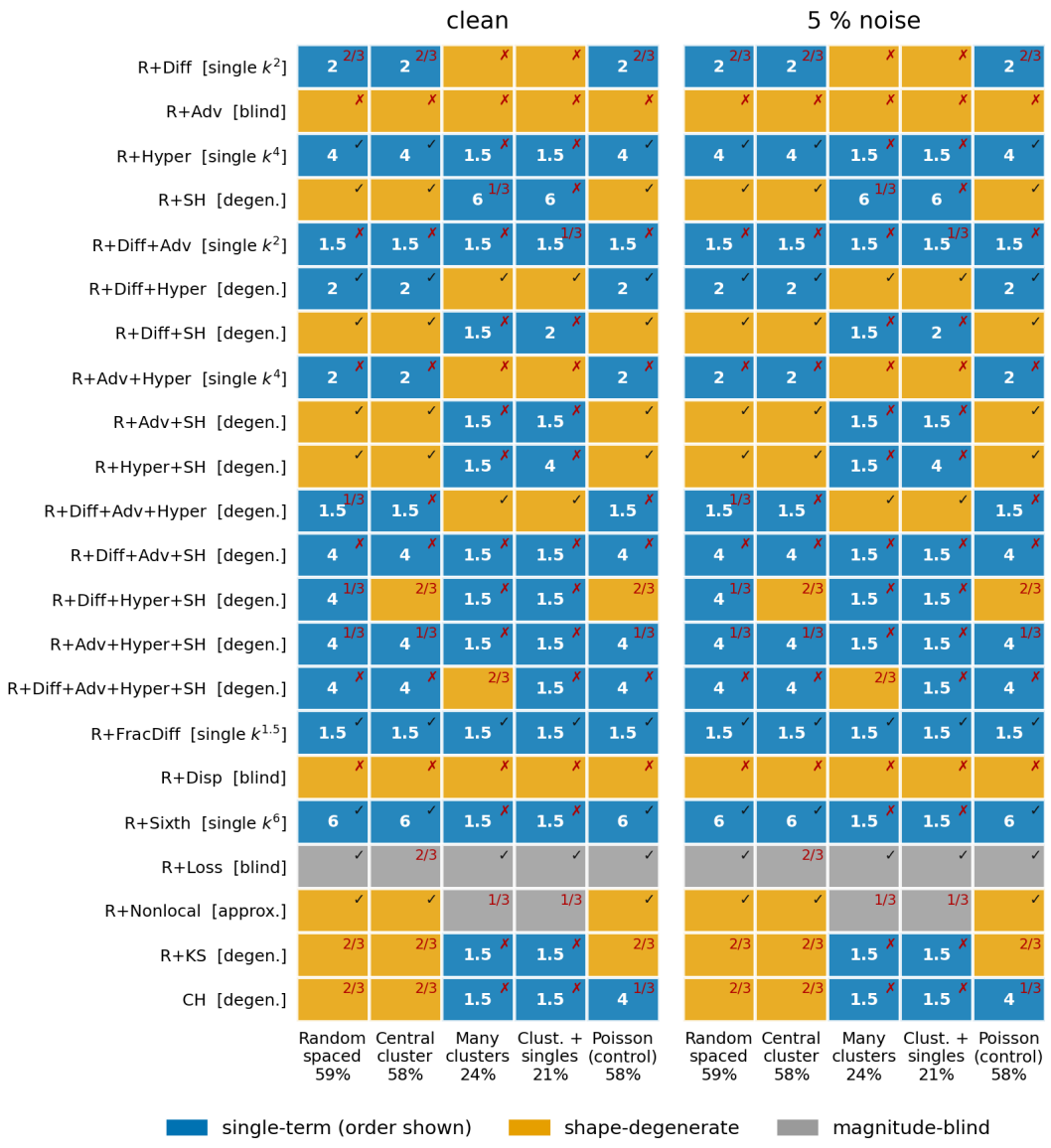



Figure S16: **Persistent-source model: single-frame identification across 22 families × 5 initial organizations.** color, majority verdict over three realizations (single-term with selected order / multi-term / magnitude-blind); corner mark, agreement with the verdict expected for the shoulder template (✓ all, × none, or fraction). Column footers give the share of all verdicts that agree with the table. Left, clean; right, 5% noise.

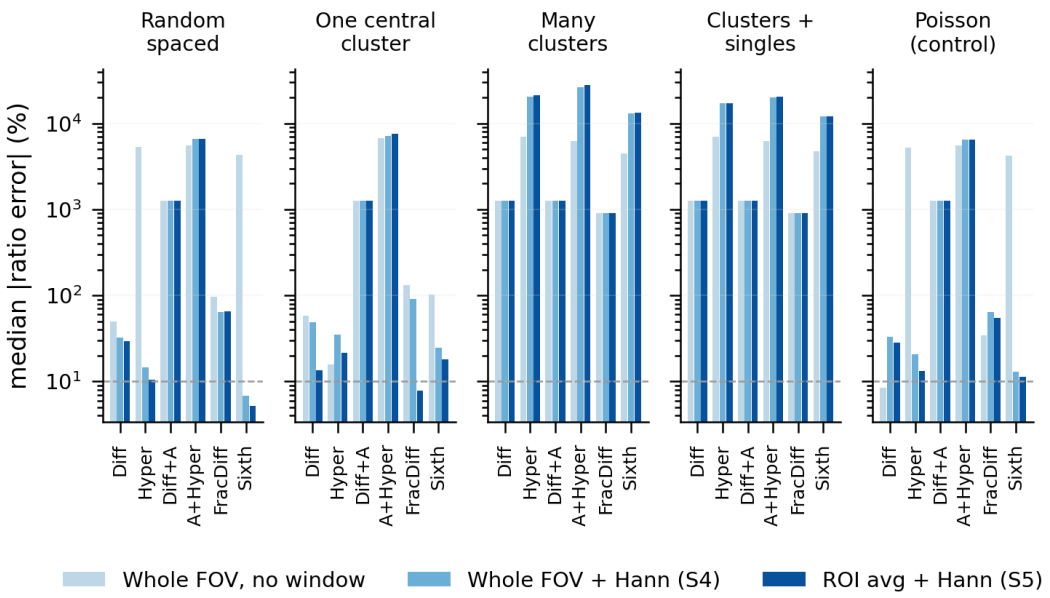



Figure S17: **Windowing versus regional averaging.** Median ratio error (true order, clean) on a non-periodic 4 mm field of view for the whole unwindowed frame (S4), the whole frame with a 2-D Hann window, and Algorithm 6. Dashed, 10%. The Hann window removes the leakage that makes the unwindowed frame uninformative; averaging over overlapping regions adds little beyond it.

is nevertheless applied to nonlinear trajectories, the resulting quantity should be interpreted as an effective spectral growth rate, rather than as the exact Fourier symbol of the governing PDE. In general, $\widehat{\mu}_{\mathrm{eff}}(k)$ depends not only on the physical parameters but also on the state of the system, the observation interval, and the initial spatial organization.

The departure from the linear theory is particularly transparent for logistic growth. At low density, $u/K \ll 1$, the nonlinear term is weak and

$$Ru\left(1-\frac{u}{K}\right) \approx Ru,$$

so the linear spectral approximation can remain accurate over an early-time regime. As the field approaches the carrying capacity, however, the quadratic term becomes increasingly important, Fourier modes interact, and the effective spectral rates become state dependent. The resulting parameter estimates can therefore become systematically biased even when the spectral fit itself appears numerically well resolved.

In our simulations, applying the linear spectral estimator to logistic trajectories increases parameter errors to approximately 13–75%, depending on the model, observation time, density regime, and initial condition. These estimates should therefore be interpreted as effective linear parameters rather than unbiased estimates of the underlying nonlinear coefficients. Moreover, the carrying capacity $K$ is not represented as an independent coefficient of a linear dispersion relation and cannot in general be recovered from the linear PSD-symbol fit alone. Estimating $K$ requires retaining the nonlinear model structure, for example through nonlinear forward fitting or another inference procedure that explicitly models mode coupling.

# I Complete model library and Fourier symbols

The spectral inference framework was evaluated on a library of 22 candidate linear PDE families spanning reaction, diffusion, directed transport, higher-order dissipation, finite-wavelength pattern selection, anomalous transport, odd-order dispersion, and nonlocal interaction. Each governing equation is written in Fourier space as $\partial_t \widehat{u} = \lambda(\mathbf{k};\boldsymbol{\theta})\widehat{u}$ with $\lambda = \mu + i\omega$, where $\mu = \mathrm{Re}\,\lambda$ controls Fourier-magnitude evolution and $\omega = \mathrm{Im}\,\lambda$ controls Fourier-phase evolution (main text, Eq. (1)).

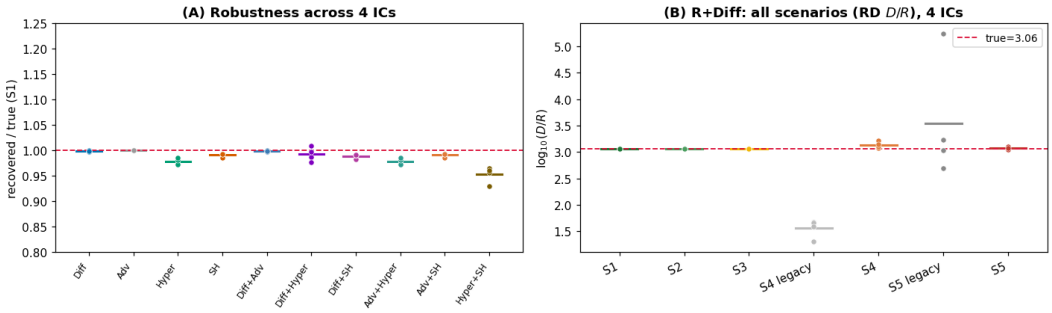



Figure S18: **Robustness of spectral recovery across initial conditions and observation protocols. (A)** Recovery of PSD-visible dynamical coefficients across four initial spatial organizations for the indicated PDE families under the time-resolved protocol S1. Points show estimates from the individual initial conditions, normalized by the corresponding true value; horizontal bars show the mean across initial conditions. The dashed line denotes exact recovery. Recovery remains close to the true value across the tested families, although multi-term and higher-order operators show greater variation, illustrating a difference in practical conditioning rather than a change in structural identifiability. **(B)** Recovery of the reaction–diffusion ratio $D/R$ across observation protocols S1–S5 and four initial spatial organizations. The dashed line marks the true value, $\log_{10}(D/R) = 3.06$; points denote individual initial conditions and horizontal bars summarize their central tendency. Time-resolved and two-frame protocols (S1–S3) are insensitive to the initial organization because the initial spectrum cancels between observations. The revised single-frame estimators S4 and S5 remain close to the true ratio when the initial spatial organization is modeled as required by Proposition 2, whereas the legacy single-frame procedures can produce large errors despite fitting the observed spectrum. Together, the panels distinguish robustness of an identifiable quantity from structural identifiability: variation in recovery across initial conditions affects estimation accuracy but does not change which quantities are supported by each observation protocol.

Table S3 lists the elementary operators from which the library is constructed; pairwise and higher-order combinations are obtained by addition of the corresponding Fourier symbols.

Table S3: **Elementary PDE operators and their Fourier symbols.** The Fourier convention is $\widehat{\partial_x u} = ik_x \widehat{u}$. Even-order dissipative operators contribute to $\mathrm{Re}\,\lambda$, whereas advection and odd-order dispersion contribute to $\mathrm{Im}\,\lambda$.

| **Mechanism** | **PDE contribution** | $\mathbf{Re}\,\boldsymbol{\lambda}$ | $\mathbf{Im}\,\boldsymbol{\lambda}$ |
|---|---|---|---|
| Reaction | $Ru$ | $R$ | 0 |
| Diffusion | $D\nabla^2 u$ | $-Dk^2$ | 0 |
| Advection | $-\mathbf{c}\cdot\nabla u$ | 0 | $-\mathbf{c}\cdot\mathbf{k}$ |
| Hyperdiffusion | $-\nu\nabla^4 u$ | $-\nu k^4$ | 0 |
| Swift–Hohenberg | $-(1+\ell^2\nabla^2)^2 u$ | $-(1-\ell^2 k^2)^2$ | 0 |
| Fractional diffusion | $-D_f(-\nabla^2)^{\alpha/2}u$ | $-D_f\lvert k\rvert^\alpha$ | 0 |
| Third-order dispersion | $\beta_3\partial_x^3 u$ | 0 | $-\beta_3 k_x^3$ |
| Sixth-order diffusion | $\mu_6\nabla^6 u$ | $-\mu_6 k^6$ | 0 |
| Linear loss | $-\eta u$ | $-\eta$ | 0 |
| Gaussian nonlocal interaction | $\chi(G_\sigma * u - u)$ | $\chi(e^{-\sigma^2 k^2/2}-1)$ | 0 |
| Linearized Kuramoto–Sivashinsky | spectral form | $R + a_2 k^2 - a_4 k^4$ | 0 |
| Linearized Cahn–Hilliard | spectral form | $A_2 k^2 - A_4 k^4$ | 0 |

For a combination of operators the total symbol is the sum of the individual contributions, so the *temporal evolution* of PSD magnitude represents the even-order dissipative terms while advection and odd-order dispersion require phase-sensitive observations (main text, §4.3). Under $\partial_t \rho = \mathcal{L}\rho$, $\omega$ never enters a single frame's magnitude; under the persistent-source model it does, but it is still not recoverable from it (Section G.1). The complete 22-family table, including each family's observational equivalence class under every protocol, is given in Table S5.

# J REPRESENTATION-DEPENDENT PARAMETER IDENTIFIABILITY

The magnitude/phase decomposition and the S1–S5 information content are derived in the main text (§2, §4.3); we do not repeat those derivations here. This appendix applies the resulting rules to the 22 candidate families and reports the practical recovery boundaries.

**Protocol summary.** S1 and S2 determine $\mu(k)$ in physical units and recover absolute PSD-visible parameters when the parameter-to-symbol map is injective; parameters appearing only in $\omega(k)$ are recovered instead from the cross-spectrum phase $\arg C = \omega(k)\Delta t$. S3 determines $\mu(k)$ only up to a common scale $s = 2\Delta t$ and therefore yields the projective coefficient vector—all coefficient ratios, including those of multi-term operators (and, in special cases such as Swift–Hohenberg, an internal calibration of $s$; Section D.4). A single frame (S4) identifies nothing without a model of the initial spectrum; with a known structure factor it identifies the operator family and all ratios among wavenumber-dependent coefficients, and with a calibrated level also the growth-normalized ratios (Proposition 2). Regional analysis (S5) has the same targets. Multi-term operators are less well conditioned from one frame, parameters that are separated only through the absolute time scale remain confounded, and phase-only mechanisms are unrecoverable from one frame.

Table S4: **Representation-dependent spectral identifiability classes.** Classification depends on both the candidate PDE and the observation protocol. "Visible" denotes that a parameter affects the corresponding spectral channel; separate structural identifiability additionally requires injectivity of the relevant parameter-to-symbol map. The single-frame classes assume the initial-state model of Proposition 2 and apply equally to whole-frame (S4) and regional (S5) analysis.

| Class | Representative examples | Interpretation |
|---|---|---|
| Magnitude-visible | $R$, $D$, $\nu$, $\ell$, $D_f$, $\alpha$, $\mu_6$, $\chi$, $\sigma$ | Affects $\mathrm{Re}\,\lambda(k)$ and is therefore represented in time-resolved PSD magnitude. Separate recovery of the physical parameter requires injectivity of the complete parameter-to-symbol map. |
| Combination-identifiable | $R - \eta$ and other model-dependent net coefficients | Multiple physical parameters generate the same observable coefficient or symbol. The representation identifies only the combination that is invariant over the corresponding parameter-to-symbol fiber. |
| Scale-invariant under unknown time (two frames) | $D/R$, $\nu/R$, all coefficient ratios, and identifiable shape parameters | When $\Delta t$ is unknown but two frames are available (S3), the observation determines the real symbol only up to a common positive scale. Every coefficient ratio—including those of multi-term operators—is recoverable as the projective parameter-to-symbol map is injective. |
| Single-frame shape-identifiable (known $S_0$) | Operator family; $m_j/m_{j'}$ (e.g. $\nu/D$); $\alpha$, $\sigma$, $\ell$ | From one unknown-time frame with a known initial structure factor, the shape $h(k) = 2T[\mu(k) - \mu(0)]$ identifies the operator family, all ratios among wavenumber-dependent coefficients, and shape parameters, for single- and multi-term operators alike; multi-term operators are less well conditioned. Not identified: $c_0$ or any ratio to it. |
| Single-frame growth-normalized (known $S_0$ + calibration) | $D/R$, $D_f/R$, $\nu/R$, $\mu_6/R$ | Additionally requires $c_0 T$, from the image mean together with either the shot-noise level and $S_0(0)$ (Eq. (7)) or the initial density (Eq. (8)). |
| Single-frame non-separable | $D$ and $\ell$ in diffusion–Swift–Hohenberg | Parameters separated only through the absolute time scale remain confounded from one frame, as under S3. A high single-frame $R^2$ is necessary but not sufficient for identifiability. |
| PSD-invisible but phase-visible | $c_x$, $c_y$, $\beta_3$ | Enters $\mathrm{Im}\,\lambda$ rather than $\mathrm{Re}\,\lambda$; constrained only by phase-sensitive representations (two frames). |
| Outside the linear-symbol representation | $K$ in logistic growth | Enters through nonlinear mode coupling; requires an inference model that explicitly represents the nonlinear dynamics. |

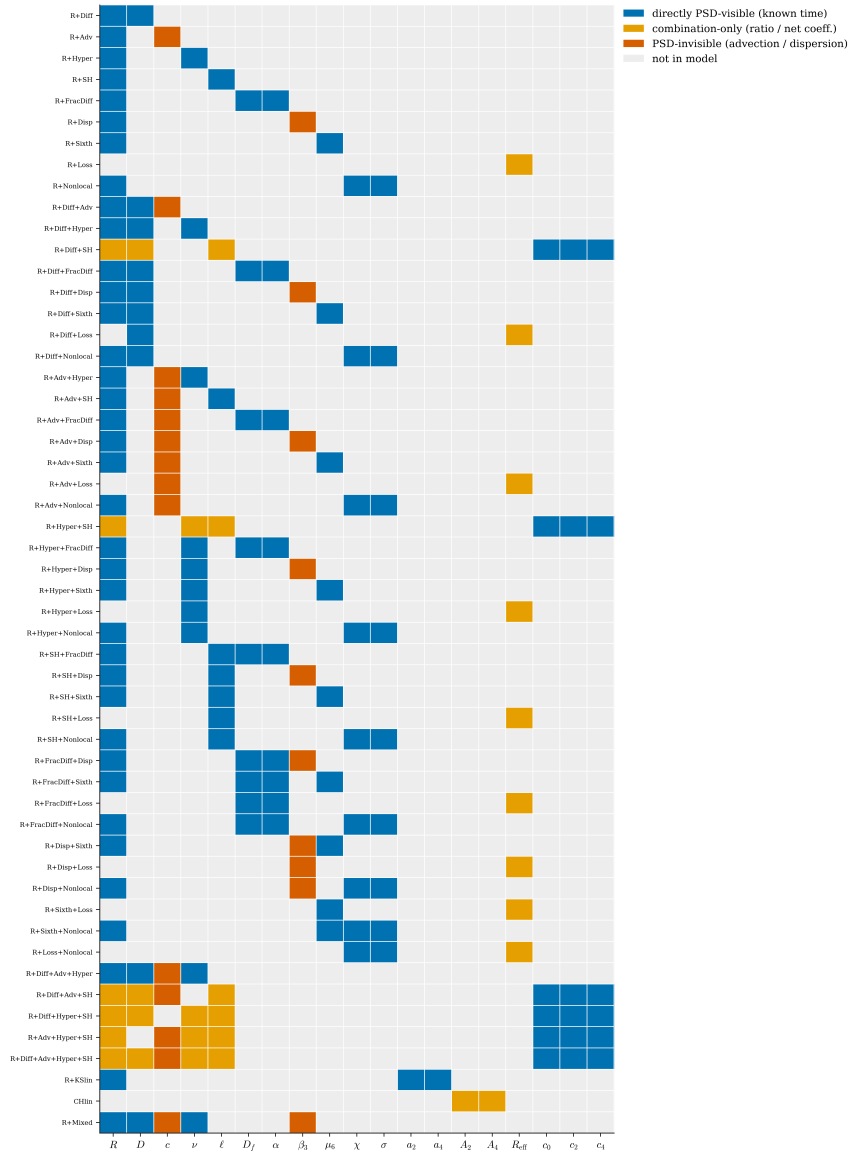



Figure S19: **Structural recoverability depends on both the spectral representation and the parameter-to-symbol map.** For each candidate family (rows) and physical parameter (columns), the map summarizes whether the parameter is directly identifiable from the PSD-visible symbol under known elapsed time, identifiable only through a parameter combination, or structurally invisible to PSD magnitude; grey indicates parameters absent from the corresponding model. Structural identifiability is determined by the injectivity of the complete parameter-to-symbol map rather than by the presence of an individual power of $k$. Linear loss, for example, exposes only the combination $R-\eta$, while advection coefficients and the third-order dispersion coefficient $\beta_3$ enter only $\mathrm{Im}\,\lambda$ and are therefore invisible to PSD magnitude. Parameters invisible in this representation may nevertheless become identifiable when phase information is retained. The map shown is for the known-time protocols (S1/S2), in which the full symbol $\mu(k)$ is observed. Under a single unknown-time frame, the identifiable targets reduce to the ratios of Proposition 2, conditional on a model of the initial spectrum, and multi-term families are less well conditioned (§4.4).

### J.1 Recoverability across the 22 candidate PDE families

Table S5 summarizes the representation-dependent identifiability for the 22 families. A parameter is listed as recoverable from PSD only when it appears in the real symbol and the relevant parameter-

to-symbol map is structurally separable; parameters occurring only in the imaginary symbol are assigned to the phase-sensitive cross-spectrum estimator. The final column gives the single-frame targets under $\partial_t \rho = \mathcal{L}\rho$ with a known initial structure factor, together with the observed rate of correct class selection (S4, Table S2). The persistent-source alternative is evaluated in Section G.1.

### J.2 Practical identifiability: recovery boundaries

Structural identifiability (Prop. 1) fixes what *could* be recovered from a perfect observable; practical identifiability concerns how estimation degrades under finite data. Because the time-resolved estimators read $\mu(k)$ from the temporal evolution of the PSD, three factors control practical recovery, and we vary each while holding the true dynamics fixed (R+Diff, recovering $D$; R+Hyper, recovering $\nu$; Fig. S18).

**Measurement noise (signal-to-noise).** Adding zero-mean Gaussian noise of standard deviation $\sigma_n$ (as a fraction of the field standard deviation) to every frame before the PSD degrades recovery *smoothly*: the $D$ error grows $1.3\%, 2.7\%, 6.9\%, 13.7\%, 24.2\%, 38.2\%$ at $\sigma_n = 0, 0.02, 0.05, 0.1, 0.2, 0.4$. The absence of any discontinuity confirms that noise is a practical, not structural, limitation. Multiple time points (S1) further suppress noise relative to a two-frame estimate, while structurally invisible parameters remain unrecoverable at every noise level. In a single frame, noise adds an approximately white floor; the single-frame estimator stops at the first shell where the spectrum reaches its floor, so noise mainly shortens the usable band, and operators whose spectra fall fastest lose shells first.

**Spectral support and wavenumber sampling.** Recovery requires the initial spectrum to place power where $\mu(k)$ varies, near $k^\star = \sqrt{R/D} \approx 0.03\,\mathrm{rad}/\mu\mathrm{m}$, sampled finely enough. A larger domain gives finer sampling $\Delta k = 2\pi/L$ and monotonically improves recovery: the $D$ error falls from $\sim$10% at $L = 1000\,\mu$m to $\sim$1% at $L = 2500\,\mu$m to $\sim$0.2% at $L = 8000\,\mu$m (for $\nu$, $\sim 25\% \to 0.7\%$), with the high-order parameter ($\nu$, on $k^4$) most sensitive because it is determined at higher wavenumbers. Single-frame estimation is more demanding, because the low-wavenumber shells that fix the spectral level hold few modes; regional analysis, which resolves fewer low wavenumbers still, is correspondingly less accurate for multi-coefficient ratios (Section G). The persistent-source shoulder estimator needed a $5000\,\mu$m domain for the same reason (Section G.1).

**Invariance to field construction (time-resolved protocols).** By contrast, S1 and S3 recovery is essentially independent of how the field is built: over 200–8000 cells and cell radii of 3–40 $\mu$m the error changes by less than a factor of two ($D \sim 1\%$, $\nu \sim 5\%$), because differencing cancels $P(k, 0)$. What matters is spectral *support*, not the number or size of the seeding structures. This invariance does *not* carry over to a single frame, where $P(k, 0)$ must be modelled through its structure factor (Section G).

Together these bound the operating regime: with adequate signal-to-noise and a domain large enough to resolve the informative band, PSD-visible parameters are recovered to within a few percent from two or more frames; outside it, error grows continuously and predictably, and the structurally invisible parameters (advection, odd-order dispersion) remain unrecoverable from magnitude at every setting. From a single frame with a known structure factor, the spectral class is selected in 94% of frames and identifiable ratios are within 10% in 86% (whole frame); with a wrongly assumed white spectrum these fall to 72% and 54%.

### J.3 Generalization beyond the enumerated candidate library

Although Table S5 lists the 22 families explicitly evaluated, the analysis applies to arbitrary linear mechanisms through the decomposition $\lambda = \mu + i\omega$. Increasing the number of named candidate PDEs does not necessarily increase the number of observationally distinguishable models: different families can induce the same observable real symbol and belong to the same equivalence class. For instance, $R + \mathrm{Diff}$ and $R + \mathrm{Diff} + \mathrm{Adv}$ share $\mu(k) = R - Dk^2$ and are indistinguishable from PSD magnitude, becoming distinguishable only when phase is retained. (Under the persistent-source model the two differ in a single frame's magnitude, but advection is still not recoverable from it; Section G.1.) This motivates reporting the 22 tested families together with their equivalence and recoverability properties rather than treating a larger combinatorial enumeration as necessarily distinct.

Table S5: **Representation-dependent parameter identifiability across the 22 candidate PDE families.** S1 (time-resolved PSD) and S2 (two frames, known $\Delta t$) share the same structural identifiability for PSD-visible mechanisms; the S1/S2 column lists quantities recoverable from PSD magnitude when the parameter-to-symbol map is injective. Parameters marked PSD-invisible occur only in $\mathrm{Im}\,\lambda$ and are recovered from cross-spectrum phase. S3 (two frames, unknown $\Delta t$) identifies scale-invariant ratios / projective combinations. The final column lists what one frame identifies under $\partial_t \rho = \mathcal{L}\rho$ with a known initial structure factor (Proposition 2b), with ratios to the growth rate requiring a calibrated level (Proposition 2c), followed by the observed share of frames in which S4 selected the correct spectral class (known structure factor, 32 frames; Table S2). Regional analysis (S5) has the same targets. Entries are *structural*; practical recovery additionally depends on noise, spectral support, resolution, and conditioning.

| **PDE family** | **S1/S2: known-time PSD** | **Phase alt.** | **S3: unknown time** | **S4/S5: one frame (known $S_0$); class correct** |
|---|---|---|---|---|
| $R + \mathrm{Diff}$ | $R, D$ | — | $D/R$ | $k^2$; $D/R$ with calibration; 94% |
| $R + \mathrm{Adv}$ | $R$; $c$ PSD-invisible | $c$ from $\phi = -ck_x\Delta t$ | scaled $R$; +phase gives $c/R$ | flat (magnitude-blind); $c$ never; 81% |
| $R + \mathrm{Hyper}$ | $R, \nu$ | — | $\nu/R$ | $k^4$; $\nu/R$ with calibration; 94% |
| $R + \mathrm{SH}$ | $R, \ell$ | — | $R, \ell, \Delta t$ via internal SH calibration | $k^2{+}k^4$; $\ell$ and $T$; $R$ with calibration; 100% |
| $R + \mathrm{Diff} + \mathrm{Adv}$ | $R, D$; $c$ PSD-invisible | $c$ | $D/R$; +phase $c/R$ | $k^2$; $D/R$ with calibration; $c$ never; 94% |
| $R + \mathrm{Diff} + \mathrm{Hyper}$ | $R, D, \nu$ | — | $D/R,\ \nu/R$ | $k^2{+}k^4$; $\nu/D$; $D/R, \nu/R$ with calibration (less well conditioned); 100% |
| $R + \mathrm{Diff} + \mathrm{SH}$ | $R, D, \ell$ if injective | — | projective/net combinations | $k^2{+}k^4$; net ratios only; $D, \ell$ not separable; 100% |
| $R + \mathrm{Adv} + \mathrm{Hyper}$ | $R, \nu$; $c$ PSD-invisible | $c$ | $\nu/R$; +phase $c/R$ | $k^4$; $\nu/R$ with calibration; $c$ never; 94% |
| $R + \mathrm{Adv} + \mathrm{SH}$ | $R, \ell$; $c$ PSD-invisible | $c$ | SH scale calibration; phase gives $c$ | $k^2{+}k^4$; $\ell$ and $T$; $c$ never; 100% |
| $R + \mathrm{Hyper} + \mathrm{SH}$ | $R, \nu, \ell$ if injective | — | projective/net combinations | $k^2{+}k^4$; net ratios only; $\nu, \ell$ not separable; 100% |
| $R + \mathrm{Diff} + \mathrm{Adv} + \mathrm{Hyper}$ | $R, D, \nu$; $c$ PSD-invisible | $c$ | $D/R,\ \nu/R$; +phase $c/R$ | $k^2{+}k^4$; as Diff+Hyper; $c$ never; 100% |
| $R + \mathrm{Diff} + \mathrm{Adv} + \mathrm{SH}$ | $R, D, \ell$ if injective; $c$ PSD-invisible | $c$ | projective/net + scaled phase | $k^2{+}k^4$; net ratios only; 100% |
| $R + \mathrm{Diff} + \mathrm{Hyper} + \mathrm{SH}$ | $R$, net $k^2, k^4$; $D, \nu, \ell$ not separate | — | projective ratios of net coefficients | $k^2{+}k^4$; net ratios only; 100% |
| $R + \mathrm{Adv} + \mathrm{Hyper} + \mathrm{SH}$ | $R, \nu, \ell$ if injective; $c$ PSD-invisible | $c$ | projective + scaled phase | $k^2{+}k^4$; net ratios only; 100% |
| $R + \mathrm{Diff} + \mathrm{Adv} + \mathrm{Hyper} + \mathrm{SH}$ | $R$, net even-power coeffs; $c$ invisible | $c$ | projective net + scaled phase | $k^2{+}k^4$; net ratios only; 100% |
| $R + \mathrm{FracDiff}$ | $R, D_f, \alpha$ if support sufficient | — | $\alpha,\ D_f/R$ | $k^\alpha$; $\alpha$; $D_f/R$ with calibration; 81% |
| $R + \mathrm{Disp}$ | $R$; $\beta_3$ PSD-invisible | $\beta_3$ from $\phi = -\beta_3 k_x^3 \Delta t$ | scaled $R$; +phase $\beta_3/R$ | flat (magnitude-blind); $\beta_3$ never; 81% |
| $R + \mathrm{Sixth}$ | $R, \mu_6$ | — | $\mu_6/R$ | $k^6$; $\mu_6/R$ with calibration; 100% |
| $R + \mathrm{Loss}$ | only $R - \eta$ | — | only scaled $s(R-\eta)$ | flat; only $(R-\eta)T$ with calibration; 81% |
| $R + \mathrm{Nonlocal}$ | $R, \chi, \sigma$ if injective | — | shape ($\sigma$) may remain; amplitudes projective | Gaussian; $\sigma$; $\chi/R$ with calibration; 75% |
| $R + \mathrm{KS}$ | $R, a_2, a_4$ | — | $a_2/R,\ a_4/R$ | $k^2{+}k^4$; $a_4/a_2$; $a_2/R, a_4/R$ with calibration; 100% |
| CH | $A_2, A_4$ | — | $A_4/A_2$ | $k^2{+}k^4$; $A_4/A_2$ (no net growth to calibrate); 100% |