TYPED FRAME SEARCH: REFERENCE-FREE COORDINATE SELECTION FOR PDES

Anonymous authors
Paper under double-blind review

ABSTRACT

Many partial differential equation (PDE) solutions become stationary profiles in moving or rescaled coordinates, yet an accurate field fit need not identify a unique coordinate representation. We introduce typed frame search, which selects an explicit frame–profile program from the PDE and prescribed initial and boundary data, without observing future solution trajectories. Candidates from a finite transformation grammar are fitted to physics and data constraints, ranked on independent validation designs, and penalized for frame complexity to favor canonical representations. We establish conditions for frame identifiability and reliable selection within the fitted bank. Across 28 in-bank configurations from 14 PDE families and ten seeds, the method recovers 96.8% of canonical frames (271/280), with median relative L_2 error 1.43×10^{-6}. Compared with a degree-four continuous-clock baseline using the same number of optimization starts, it achieves approximately 90-fold lower median error and 66.6-fold lower median search time. Tests with unseen profiles and two-dimensional affine flows recover all 50 and 60 frames, respectively. Ablations distinguish the constraints that determine the field from the complexity penalty that resolves equivalent representations. These results support interpretable coordinate recovery within a finite grammar, with limitations from initial-data noise and unsupported transformations.

1 INTRODUCTION

Many evolving partial differential equations (PDEs) have a simple description in coordinates that move or rescale with the solution. For example, the porous-medium equation $u_t = (u^2)_{xx}$ admits the Barenblatt family

$$u(x,t) = t^{-1/3}\left(C - \frac{x^2 t^{-2/3}}{12}\right)_+, \tag{1}$$

where $(\cdot)_+$ denotes the positive part and $C > 0$ controls the mass (Vázquez, 2007). Its support expands and its amplitude decreases in physical coordinates. In the coordinate $\xi = xt^{-1/3}$, the rescaled field $t^{1/3}u$ is a single stationary profile (Figure 1). Finding this representation separates the evolution law from the profile and can reduce approximation from a space–time problem to a spatial problem.

The useful frame is often unknown. A small PDE residual indicates that a candidate field approximately satisfies the equation, but several coordinate programs may describe that field equally well. A translating solution, for example, can be represented by a translation program or by a larger program whose dilation parameter is zero. Recovering a meaningful coordinate law therefore requires both fitting a solution and choosing a representation. This distinction matters when the frame itself is the scientific output.

We study the following problem: given a PDE, its initial conditions (ICs), boundary conditions (BCs), and any declared invariants, select an explicit coordinate frame without observing a future solution trajectory. We call this setting *reference-free*. Initial and boundary data specify the problem; exact trajectories and target frame labels are reserved for evaluation. Our method, *typed frame search*, treats the frame and profile family as discrete choices and fits their continuous parameters jointly. Each frame is a short program with named operations such as translation, damping, and dilation. A

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.

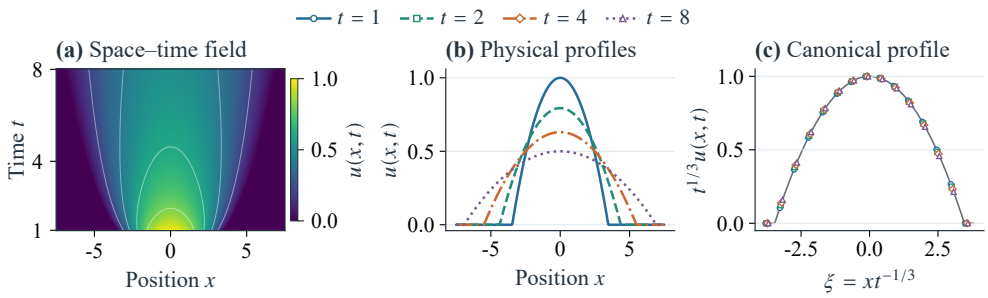

Figure 1: The Barenblatt solution in equation 1, with $C = 1$. **(a,b)** The support expands and amplitude decreases in physical coordinates. **(c)** Rescaling position and amplitude collapses four times onto $F(\xi) = (1 - \xi^2/12)_+$. The reference field is used here for illustration.

common residual-based objective compares the fitted programs on independent validation points, and a description penalty favors the simplest representation among nearly equivalent fits.

We formulate coordinate choice as structure selection: the typed bank makes candidate semantics explicit, independent validation compares their fit, and canonical selection distinguishes coordinate laws from redundant programs. We analyze when scaling balance identifies the frame, when finite-bank selection is reliable, and when a continuous approximation preserves the fitted field and its ranking.

Experiments cover 14 PDE families, changes in profile shape, and two-dimensional affine flows. On 280 in-bank trials, typed search recovers 96.8% of frame labels and meets the field-error criterion on every trial. A flexible continuous-clock baseline also meets this criterion, while typed search has lower median error and search time. Objective ablations distinguish the contributions of the PDE, initial data, and description penalty. Noise, open-set, and downstream solver experiments establish where these benefits persist. Together, the results characterize an interpretable coordinate selection method within a specified family of transformations.

2 Related Work

Similarity reductions and symmetry analysis. Classical similarity analysis uses scaling balances and invariants to derive reduced equations (Vázquez, 2007; Olver, 1993). Lie computer-algebra systems derive admitted generators from the differential operator (Hereman, 1997; Oliveri, 2021). These tools describe symmetries of an equation; choosing a reduction for a particular initial–boundary value problem also requires compatibility with its conditions. We compare candidate reductions at that problem level. The operator-template and restricted symbolic controls in our experiments cover specified operator families.

Learning transformations and solving PDEs. Physics-informed neural networks use differential residuals and problem data to fit solutions (Raissi et al., 2019). Neural similarity solvers can fit profiles and exponents within a chosen ansatz (Wang et al., 2023). Other methods recover hidden symmetries or self-similar transformations from sampled fields (Liu & Tegmark, 2022; Watanabe et al., 2025; Ko et al., 2024). Symmetries also support physics-informed and self-supervised learning (Akhound-Sadegh et al., 2023; Mialon et al., 2023; Otto et al., 2025). Our setting retains residual-based fitting while selecting the frame type from the operator and prescribed conditions, without trajectory samples.

Symbolic and program search. Sparse identification and symbolic regression search for equations or expressions (Brunton et al., 2016; Udrescu & Tegmark, 2020; Cranmer, 2023). Program-search systems extend the available expression vocabulary (Romera-Paredes et al., 2024; Shojaee et al., 2025; Novikov et al., 2025; Yu & Yang, 2026). Here, the searchable object is the coordinate representation of a known PDE. A finite grammar permits exhaustive comparison of frame–profile pairs and gives each output an explicit interpretation. The continuous-clock baseline tests how much of the benefit follows from flexible coordinate fitting alone.

3 PROBLEM FORMULATION AND SEARCH

Let $\mathcal{N}[u] = 0$ be a PDE on $\Omega \times [t_0, t_1]$, with initial data u_0, boundary data, and optional invariants. A hypothesis consists of a frame $f \in \mathcal{F}$, a profile family $h \in \mathcal{H}$, and parameters θ. The sets $\mathcal{F}$ and $\mathcal{H}$ are finite. The hypothesis defines $u_{f,h,\theta}(x,t) = a_f(t;\theta)F_{h,\theta}(z_f(x,t;\theta))$, where a_f scales the amplitude and z_f maps physical to reduced coordinates. The task is to choose a hypothesis whose field satisfies the public specification and whose frame provides a concise description of its evolution.

3.1 TYPED FRAME AND PROFILE BANK

The basic frame operations map physical position to a reduced coordinate z and multiply a stationary profile by an amplitude a:

$$\begin{aligned} \text{stationary: } & z = x, && a = 1, \\ \text{damped: } & z = x, && a = e^{-\delta(t-t_0)}, \\ \text{co-moving: } & z = x - c(t-t_0), && a = 1, \\ \text{self-similar: } & z = x\tau^{-\beta}, && a = \tau^{-\alpha}, \end{aligned} \tag{2}$$

where $\tau = t/t_0$ for $t_0 > 0$ and $\tau = 1 + t - t_0$ otherwise. The parameters c, δ, α, β describe translation speed, damping rate, amplitude scaling, and spatial scaling. The grammar also contains three pairwise compositions: co-moving–self-similar, damped–co-moving, and damped–self-similar. These seven programs are instances of

$$u_{f,\theta}(x,t) = e^{-\delta(t-t_0)}\tau^{-\alpha}F_\phi\big((x - c(t-t_0))\tau^{-\beta}\big)\,, \tag{3}$$

with parameters absent from a program fixed at zero and ϕ denoting profile parameters. For the transport equation $u_t + vu_x = 0$, substituting a co-moving candidate $u(x,t) = F(x - ct)$ with $t_0 = 0$ gives

$$\mathcal{N}[u] = (v - c)F'(x - ct). \tag{4}$$

For a nonconstant profile, a vanishing residual identifies the speed $c = v$; the initial data determine F. A constant profile leaves the speed unidentified. This illustrates how the operator and the selected solution jointly determine whether a coordinate parameter is recoverable.

We use *clock* for a scalar function of time that controls motion or rescaling. Four further programs complete the eleven-frame bank: accelerated translation $z = x - q(t-t_0)^2/2$, quadratic-time damping $a = e^{-q(t-t_0)^2/2}$, exponential dilation $z = xe^{-q(t-t_0)}$, and oscillatory translation $z = x - q\sin(t - t_0)$. Each has one fitted coefficient q; the unspecified amplitude or coordinate is the identity. Thus the bank includes constant-speed motion, power-law rescaling, and several nonconstant temporal laws in a single search.

For each frame, the profile bank offers Gaussian, positive-quadratic, hyperbolic-tangent, and sinusoidal families, together with a cubic-spline interpolant of the public initial data. The spline uses the same 161 IC samples available to every method and has no learned profile parameter. The complete search contains 55 frame–profile fits. Analytic and spline profiles are evaluated through the same finite-difference residual, including at compact-support edges.

The frames share an initial-time normalization:

$$z_f(x,t_0) = x, \qquad a_f(t_0) = 1, \qquad u_{f,h,\theta}(x,t_0) = F_{h,\theta}(x). \tag{5}$$

Every frame is therefore compared against the same initial profile, without an arbitrary initial shift or amplitude rescaling. For the interpolant option, the remaining fitted parameters describe temporal evolution; profile uncertainty comes from sampling and interpolation of the IC. For analytic profiles, frame and profile parameters remain jointly fitted because both enter the PDE residual.

3.2 PUBLIC OBJECTIVE AND REDUCED-OPERATOR DRIFT

On a collocation design D, the base objective is

$$J_f^{\text{base}}(\theta; D) = \frac{\widehat{\mathbb{E}}_D \mathcal{N}[u_{f,\theta}]^2}{\widehat{\mathbb{E}}_D u_{f,\theta}^2 + \varepsilon} + 3L_{\text{IC}} + \tfrac{1}{4}L_{\text{BC}} + \tfrac{2}{5}L_{\text{inv}}. \tag{6}$$

We suppress the profile index when it is unambiguous. Here, $\widehat{\mathbb{E}}_D$ is an empirical average and $\varepsilon > 0$ stabilizes the denominator. L_{IC} measures normalized squared error against the initial samples, L_{BC} measures boundary mismatch, and L_{inv} measures normalized mass error when conservation is specified, and is zero otherwise. Residual normalization reduces sensitivity to field amplitude; the data terms penalize the trivial zero field when the prescribed data are nonzero.

The objective also compares residuals at matching reduced coordinates across time. Let $x_f(z,t)$ invert the spatial frame map, Z be a reduced-coordinate grid, and T a set of time slices. Define

$$D_f(\theta) = \frac{|Z|^{-1} \sum_{z \in Z} \mathrm{Var}_{t \in T}[\mathcal{N}[u_{f,\theta}](x_f(z,t),t)]}{\widehat{\mathbb{E}}_{Z,T} u_{f,\theta}(x_f(z,t),t)^2 + \varepsilon}, \qquad J_f = J_f^{\mathrm{base}} + \tfrac{1}{4} D_f. \tag{7}$$

This term measures temporal variation of the PDE residual in the selected frame. Every candidate already has a time-independent profile, so testing profile stationarity alone would provide no discrimination. Drift is a consistency term alongside residual magnitude; its contribution is tested by refitting without it.

Different programs can induce the same fitted transform and field. For example, a damped–co-moving fit with zero damping also belongs to the co-moving family. To favor the shorter representation, validation uses

$$J_f^{\mathrm{can}} = J_f + \lambda k_f, \qquad \lambda = 10^{-5}, \tag{8}$$

where k_f counts the frame parameters used by program f. The penalty affects selection, not parameter fitting. It expresses a preference among sufficiently close fits; a simpler program can still lose when its residual and condition errors are larger. All weights and parameter bounds are fixed before evaluation.

3.3 Fitting, validation, and candidate comparison

For each problem and random seed, we draw training and validation designs before fitting any candidate. All candidates use the same designs. The limited-memory Broyden–Fletcher–Goldfarb–Shanno method with bound constraints (L-BFGS-B) (Byrd et al., 1995) fits each frame–profile pair on the training objective, giving parameters $\widehat{\theta}_{f,h}$. Selection uses the independent validation design D_{val}:

$$(\widehat{f}, \widehat{h}) = \operatorname*{arg\,min}_{f \in \mathcal{F}, h \in \mathcal{H}} J_{f,h}^{\mathrm{can}}(\widehat{\theta}_{f,h}; D_{\mathrm{val}}). \tag{9}$$

The fitted parameters are local optimization outputs. A block-validation variant averages the complete score over eight independently sampled validation designs. Each block includes its own composite score; ratios and variances are computed within the block.

For every candidate, we retain fitted parameters, score components, optimizer status, evaluation count, and elapsed time. Search receives only the operator and declared conditions, with boundary queries restricted to the endpoints. Exact trajectories and reference labels are accessed in a separate evaluation process only after search is complete. Appendix B specifies this evaluation protocol.

3.4 Comparison methods

The dense joint baseline fits $(\delta, c, \alpha, \beta)$ in a single continuous vector. Its sparse counterpart uses a fixed smooth ℓ_1 path and decodes nonzero coordinates at threshold $\tau_s = 0.02$. A seven-restart dense control is evaluated on a 120-trial subset. The more flexible continuous-clock baseline fits polynomial translation, log-amplitude, and log-dilation functions:

$$u(x,t) = e^{-g(t)} F((x - s(t))e^{-h(t)}), \qquad q(t) = \sum_{k=1}^{d} b_{q,k} r(t)^k, \quad q \in \{s, g, h\}, \tag{10}$$

where $r(t) = (t - t_0)/(t_1 - t_0)$ and $b_{q,k}$ are fitted coefficients. With $d = 4$, five profiles, and eleven restarts per profile, it matches the typed method's 55 starts per trial. Equal start counts do not imply equal function evaluations or wall time, which we report separately.

The seven-frame grammar, random-frame selection, complexity-only selection, and profile removal test the role of the search space and score. Two additional controls supply structure: a human-derived operator template and an offline correct-frame oracle. The oracle receives the frame label and fits the profile with the public objective. These controls distinguish knowledge of the frame from successful numerical fitting of its profile.

4 IDENTIFIABILITY AND SELECTION GUARANTEES

We analyze when the PDE identifies frame parameters, when validation selects the best fitted candidate, and when approximating a temporal law preserves its representation. The results condition on their stated rank, margin, and regularity assumptions. Proofs and auxiliary stability results are in Appendix A.

Let $\mathcal{C}$ denote the fitted frame–profile candidate bank used in the finite-bank results below; there, f indexes a complete fitted candidate.

Definition 1 (Irreducible homogeneous reduction)**.** For frame parameters $\eta \in \mathbb{R}^d$, suppose an active residual has the form

$$\mathcal{R}(\xi, t; \eta) = \sum_{j=1}^{m} t^{a_j^\top \eta + b_j} G_j(\xi). \tag{11}$$

Here j indexes active residual terms, with $a_j \in \mathbb{R}^d$ and $b_j \in \mathbb{R}$. The reduction is irreducible on an open set U if no nonempty proper subset of the active profile factors G_j sums identically to zero on U.

Theorem 2 (Scaling-balance identifiability)**.** *Suppose equation 11 vanishes on $U \times I$, where I is an open positive-time interval, and the reduction is irreducible. Let M stack $(a_j - a_1)^\top$ and let q stack $b_1 - b_j$, for $j = 2, \dots, m$. Every admissible reduction satisfies $M\eta = q$. If $\operatorname{rank}(M) = d$, these balances identify at most one frame parameter vector. Otherwise, the consistent balance equations have a nullspace of dimension $d - \operatorname{rank}(M)$; additional constraints are needed to resolve it.*

The essential step is that distinct powers of time are linearly independent. Irreducibility therefore forces all active exponents to agree. In the Barenblatt example, the PDE gives $\alpha + 1 = 2\alpha + 2\beta$, while mass conservation gives $\alpha = \beta$. Together they yield $\alpha = \beta = 1/3$. The invariant supplies the missing constraint. This explains why a residual can admit several scaling balances and why problem-specific information can make the frame unique.

Theorem 3 (Selection from independent validation blocks)**.** *Condition on the fitted finite candidate bank $\mathcal{C}$. Let $S_f(V_i) \in [0, B]$ be the complete score of candidate f on validation blocks $V_1, \dots, V_n$, independent of one another and of the fitted bank, and shared across candidates. Suppose a candidate $f^\star$ has expected score gap $\mathbb{E}[S_f(V) - S_{f^\star}(V)] \geq \Delta > 0$ for every $f \neq f^\star$. Selection by the mean block score satisfies*

$$\Pr(\widehat{f} \neq f^\star) \leq (|\mathcal{C}| - 1) \exp\left(-\frac{n\Delta^2}{2B^2}\right). \tag{12}$$

The bound follows from Hoeffding's inequality (Hoeffding, 1963) applied to paired block-score differences. Independence is needed across blocks; samples inside a ratio or variance need not be independent. The score bound B must be established, for example through an appropriate bounded domain or explicit clipping. The theorem concerns selection among the obtained fits and does not control optimization error within a candidate.

Proposition 4 (Validation cost of a larger grammar)**.** *Under Theorem 3, for $|\mathcal{C}| \geq 2$ and $0 < \delta < 1$, an error probability at most δ is ensured by*

$$n \geq \frac{2B^2}{\Delta^2} \log \frac{|\mathcal{C}| - 1}{\delta}. \tag{13}$$

Thus the sufficient validation budget grows logarithmically with bank size when B and Δ are preserved.

Proposition 5 (Canonical recovery with description length)**.** *Let $C(f) = R(f) + \lambda k_f$ and $\widehat{C}(f) = \widehat{R}(f) + \lambda k_f$, where R is the expected unpenalized score and $\widehat{R}$ its estimate. Suppose $f^\star$ uniquely minimizes C on the fitted bank, with gap $\Delta_\lambda = \min_{g \neq f^\star}[C(g) - C(f^\star)] > 0$. If $\sup_f |\widehat{R}(f) - R(f)| \leq \epsilon$ and $\Delta_\lambda > 2\epsilon$, empirical selection recovers $f^\star$. For equal-risk equivalent programs and $\lambda > 0$, the population ordering favors the unique program with smallest k_f.*

This proposition recovers the penalized minimizer; identifying it with the reference frame requires a separate argument. Equal fit and complexity still require a tie rule. The equivalence analysis in Section 6 measures strict labels and active transforms separately.

Theorem 6 (Polynomial-clock approximation)**.** *Express each clock $q \in \{s, g, h\}$ in normalized time $r \in [0, 1]$. Let $q \in C^{d+1}([0, 1])$, $q(0) = 0$, and $P_d q$ be its degree-d Taylor polynomial at zero. Define $E_d = \max_q \|q^{(d+1)}\|_\infty/(d + 1)!$. On $|x| \le X$, assume both original and polynomial clocks satisfy $|g| \le G$, $|h| \le H$, and $|s| \le S$. Let $|F| \le M$ and let F be L-Lipschitz on the required coordinate range. By Proposition 9, replacing the clocks by $P_d q$ changes the field in sup norm by at most $C_F E_d$, where $C_F = e^G[M + Le^H(1 + X + S)]$. If the population score is K-Lipschitz in this norm, the bounds hold uniformly over the compared fields, and the minimizer gap is $\Delta > 2KC_F E_d$, the minimizer is preserved.*

Smooth clocks can therefore be approximated without using named atoms. The score-stability conclusion needs its additional Lipschitz assumption: a differential residual is not generally controlled by field sup error alone. Neither this result nor the finite-bank bound guarantees a global optimum of the nonconvex fitting problem.

5 Experimental Protocol

5.1 Benchmarks and metrics

The one-dimensional benchmark contains 32 configurations. Seven core families each contribute three coefficient settings: heat, porous medium, advection, reaction, and three pairwise mixed operators. Wave, reverse-sign viscous Burgers, and radial heat test transfer to different operators. Four configurations use accelerated translation, quadratic damping, exponential dilation, and oscillatory translation. These 28 configurations belong to the frame bank. Four out-of-bank configurations use cubic translation, logarithmic damping, reciprocal dilation, and chirped translation. Appendix B lists coefficients and bounds.

Each configuration is run with ten seeds. The main training and validation designs have 96 and 192 interior points, respectively. Initial data use 161 samples; boundary data use 37 times; drift uses 41 reduced coordinates and six times. Central differences use $h_x = h_t = 2 \times 10^{-3}$, and fitting is capped at 180 L-BFGS-B iterations. All numerical jobs use a single CPU core.

On an offline 181×61 space–time grid, we evaluate relative L_2 error as $\|\widehat{u} - u_{\mathrm{ref}}\|_2/\|u_{\mathrm{ref}}\|_2$, where u_{ref} is the reference field and both norms use this grid. A trial is *solved* when this error is below 0.08. *Strict frame accuracy* compares the selected program label with the canonical reference label. *Support-frame accuracy* thresholds fitted frame parameters at 0.02 and compares the active transform. Field equivalence uses the stricter threshold $L_2^{\mathrm{rel}} < 0.01$. A continuous clock has a field prediction but no typed label, so its frame accuracy is left undefined.

5.2 Controlled tests and uncertainty

Objective ablations refit all 55 candidates on 14 family representatives and five seeds. Additional tests vary the profile bank, add up to 100 nuisance clocks, change IC sample counts and noise, and compare one, three, and eleven continuous restarts. Two-dimensional tests use four affine-flow families, three coefficient settings, and five seeds. Downstream tests hold the selected frame fixed and compare its representation with physical coordinates, an oracle frame, and a stationary control. Full designs are given in Appendix C.

Comparisons pair problem–seed instances and share random designs. We report medians for errors and time, and cluster bootstrap intervals from 5,000 resamples at the family or problem level, keeping seeds together within each cluster. Seeds measure variation in sampling and optimization rather than independent PDE systems. All scheduled outcomes, including optimizer exhaustion, enter the aggregates. Settings and open-set calibration splits are fixed before their evaluation labels are accessed.

6 Results

6.1 Frame recovery and continuous comparisons

Table 1 reports the complete comparison. Typed search recovers 271/280 canonical frames, with a problem-cluster 95% interval of 91.4–100.0%, and solves all 280 fields. Its median relative error is

Table 1: In-bank comparison over 28 configurations and ten seeds. The seven-restart dense control uses a 120-trial subset. Structural controls receive a human-derived or oracle frame. N denotes trial count; "–" indicates no typed label.

Method	N	Frame %	Solved %	Median rel. L_2
Typed search	280	96.8	100.0	1.43×10^{-6}
Continuous clock (degree 4)	280	–	100.0	1.28×10^{-4}
Dense joint	280	–	91.4	3.48×10^{-6}
Sparse joint	280	85.7	91.4	3.44×10^{-6}
Dense joint (7 restarts)	120	100.0	100.0	1.28×10^{-6}
Structural controls				
Operator template	280	100.0	99.3	1.63×10^{-6}
Correct-frame oracle	280	100.0	99.3	1.63×10^{-6}
Search-space and selection controls				
Seven-frame grammar	280	82.5	91.8	2.25×10^{-6}
Random frame	280	9.3	25.0	0.202
Complexity only	280	0.0	0.0	0.414
Target profile removed	280	28.9	0.0	0.217

1.43×10^{-6} and its 90th percentile is 3.06×10^{-4}. The degree-four continuous clock also solves every field, with median error 1.28×10^{-4}. Typed search takes a median 4.5 seconds and 5,974 function evaluations per trial, versus 301.8 seconds and 163,921 evaluations for the continuous clock. The reported timing ratio is 66.6, measured for these single-core implementations.

Dense and sparse joint fitting each solve 91.4% of the full suite. More restarts allow dense fitting to recover every frame and field on its 120-trial subset. Thus continuous fitting is a viable alternative, with performance depending on parameterization and search budget. The operator-template and oracle controls identify every frame but solve 278/280 fields: even a correct frame still requires a successful profile fit. The random-frame and complexity-only controls show that neither an arbitrary coordinate choice nor a short program alone is sufficient.

Recovery is 98.6% over coefficient sweeps and 80.0% over the three operator holdouts. The nine strict-label misses consist of six wave trials and three porous-medium trials; all their fields satisfy the solved criterion. The four nonconstant-clock families each have 10/10 frame recoveries. The per-family breakdown in Appendix C shows that aggregate field accuracy can coexist with ambiguity in the selected program label.

6.2 What determines the selected frame?

Table 2: Objective ablations on 14 PDE families and five seeds. Every row refits all 55 candidates. Frame accuracy and successful field approximation measure different outcomes.

Refitted objective	N	Frame %	Solved %	Median rel. L_2
full	70	98.6	100.0	6.76×10^{-7}
no boundary	70	98.6	97.1	3.14×10^{-6}
no drift	70	98.6	100.0	9.69×10^{-7}
no initial	70	84.3	80.0	7.13×10^{-4}
no mass	70	98.6	100.0	9.20×10^{-7}
no residual	70	42.9	42.9	1.48×10^{-1}
residual + initial	70	100.0	95.7	1.13×10^{-6}

The PDE residual and initial data provide the strongest constraints (Table 2). Removing the residual reduces frame accuracy and solved rate to 42.9%. Removing the initial term reduces them to 84.3% and 80.0%. Residual plus initial data attains 100% frame accuracy but solves 95.7% of fields, compared with 98.6% and 100% for the full objective. The extra terms improve field fitting on this suite without improving every structural metric. Removing drift or mass changes median error modestly and leaves the solved rate unchanged.

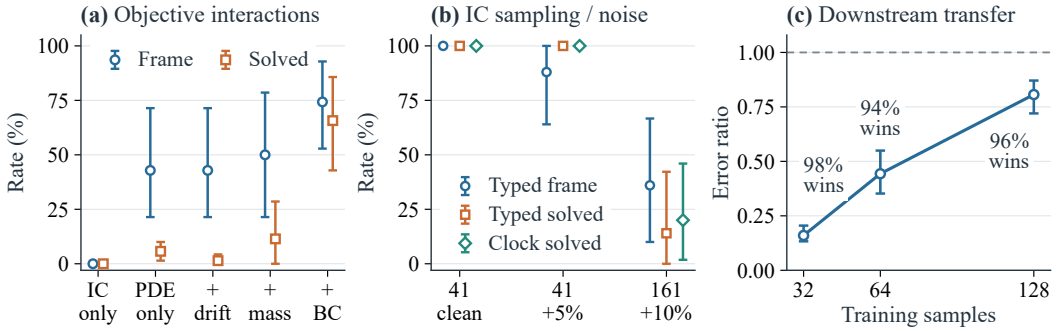

Figure 2: **(a)** Objective interactions, with 70 trials per objective; "+ term" augments the residual. **(b)** Unknown-profile IC stress, with 50 trials per design. Labels give sample count and noise level. **(c)** Downstream error ratios at matched training budgets; values below one favor selected coordinates. Intervals are 95% family-cluster bootstrap intervals, and labels in (c) give paired win rates.

Interaction tests clarify why individual terms are insufficient (Figure 2a). Residual-only fitting solves 5.7% of fields; adding boundary data raises this to 65.7%. Residual paired only with mass or drift solves 11.4% or 1.4%, and initial-only fitting solves none. These experiments support complementary roles: the operator constrains dynamics, while initial data select the relevant member of a solution family.

Canonicalization primarily affects representation. With zero description penalty, strict frame accuracy falls to 68.2%, yet support-frame accuracy and field equivalence remain 100%. At the fixed $\lambda = 10^{-5}$, strict accuracy is 96.8%. Many nominal errors are therefore longer programs with the same active transform. Adding 25, 50, or 100 nuisance clocks selects no nuisance program and leaves recovery unchanged within that stress suite. Appendix C gives the equivalence analysis and grammar stress tests.

6.3 Unknown profiles and downstream approximation

Restricting the full benchmark to the IC interpolant preserves 100% solved rate and gives 92.9% frame recovery, with median error 5.69×10^{-5}. On 50 separate unseen-profile trials with analytic profiles excluded, typed search recovers every frame and field. Its median error is 4.19×10^{-6}, compared with 2.79×10^{-5} for a one-restart continuous clock. These results show that selection can work without matching a known analytic profile family.

Observation quality limits this result (Figure 2b). Noise levels are relative to the IC standard deviation. With 41 clean IC samples, both methods solve every trial with median error about 1.2×10^{-3}. With 41 samples and 5% noise, both still solve all trials, while typed frame accuracy is 88%. At 161 samples and 10% noise, typed and continuous solved rates fall to 14% and 20%, respectively. The interpolated profile therefore tolerates sparse clean data better than these heavily corrupted initial observations.

After frame selection is fixed, we fit an offline 64-feature random-feature ridge model (Rahimi & Recht, 2007; Hoerl & Kennard, 1970) using physical coordinates, the selected frame, or an oracle frame. At 32, 64, and 128 trajectory samples, the paired geometric-mean selected/physical error ratios are 0.160, 0.443, and 0.807; all family-cluster intervals exclude one (Figure 2c). The selected representation is close to the oracle. An exploratory comparison with the stationary one-dimensional control has intervals containing one at all three budgets. The gain over physical coordinates is consistent with a benefit from lower-dimensional inputs, while a distinct benefit from recovering the motion law is unresolved by this surrogate experiment.

6.4 Two-dimensional transfer and operating limits

On four two-dimensional affine-flow families, typed search recovers all 60 frames and fields. Continuous affine fitting attains the same median relative error, 7.16×10^{-5}. Across coarse grids and IC

noise up to 5%, both methods solve every field; typed frame recovery ranges from 95% to 100%. The equality of field errors isolates the extra information supplied by the typed output: the label of the affine mechanism.

Table 3: Upwind comparison including search, with 60 matched trials per grid. Error ratios are paired geometric means of selected/physical error. Time is median seconds, including search for selected coordinates.

Grid	Error ratio	Selected wins (%)	Selected / physical time
33^2	0.216	100	0.0702 / 0.00118
65^2	0.427	100	0.0707 / 0.00318
129^2	2.152	0	0.0726 / 0.0148

The upwind test compares coordinate choices at matched numerical resolutions (Table 3). Selected coordinates reduce error at the coarse and medium grids, but the physical solver is more accurate on the fine grid. Including search makes the selected-frame method slower at all three resolutions. Coordinate simplification therefore has a different computational tradeoff from accelerating the full solver.

Thirteen independent robustness refits vary collocation size, IC noise, differencing, iteration limits, and objective weights. Each solves all 120 trials, with frame accuracy between 92.5% and 99.2%. These use the analytic-profile bank and are distinct from the unknown-profile noise tests above.

Open-set detection tests whether an in-bank fit can signal an unsupported mechanism. A scalar fit-score detector has area under the receiver operating characteristic curve (AUROC) 0.855 on the four out-of-bank families in the benchmark. A multicomponent detector combining fit, ambiguity, block instability, and the train–validation gap has AUROC 0.955 on a separate four-family novelty set, with 95% interval $[0.877, 0.997]$. Its combined-set AUROC is 0.910. The scores and test sets differ, so these values measure distinct detection settings. The first falls below the prespecified 0.90 criterion; the second exceeds it as a point estimate but retains substantial uncertainty. Appendix B specifies both calibration designs.

7 Discussion

Continuous clocks estimate fields, while typed search also names an evolution mechanism and records competing explanations. Strict-label and support-frame metrics distinguish redundant programs from incorrect active transformations. This distinction is essential when coordinate recovery is itself the scientific objective.

An explicit frame also provides a starting point for analytical reduction. Its coordinate map can be substituted into the operator and checked against the problem data. The recorded alternatives reveal whether that explanation is well separated or rests on a small score difference. This information remains useful even when a downstream numerical solver is faster in physical coordinates.

8 Limitations and Conclusion

The evidence covers synthetic PDEs and a finite grammar. Performance on measured fields, nonlocal reductions, and arbitrary symmetry groups remains untested. Strong IC noise, open-set uncertainty, and local optimization limit recovery; search-time gains over polynomial clocks do not imply end-to-end solver speedups.

Typed frame search combines operator-based fitting, independent validation, and canonical selection to recover interpretable coordinate laws. Its output includes a fitted field, a named mechanism, and the scores of competing hypotheses.

Reproducibility Statement

Appendix A provides the proofs. Appendix B specifies the benchmark, parameter bounds, validation procedure, and open-set calibration. Appendix C provides supplementary results and downstream implementation details. The accompanying artifact contains fitted candidates, experiment settings, and evaluation checks. Numerical jobs use one CPU core, single-threaded linear algebra, and at most 3 GB of memory.

Artificial Intelligence (AI) Use Statement

OpenAI Codex was used for coding assistance and language editing. The authors reviewed all AI-assisted content and take full responsibility for the manuscript, implementation, and reported results.

References

Tara Akhound-Sadegh, Laurence Perreault-Levasseur, Johannes Brandstetter, Max Welling, and Siamak Ravanbakhsh. Lie point symmetry and physics informed networks. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-1842. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/8493c860bec41705f7743d5764301b94-Abstract-Conference.html.

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. doi: 10.1073/pnas.1517384113.

Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995. doi: 10.1137/0916069.

Miles Cranmer. Interpretable machine learning for science with PySR and SymbolicRegression.jl. *arXiv preprint arXiv:2305.01582*, 2023. URL https://arxiv.org/abs/2305.01582.

Willy Hereman. Review of symbolic software for lie symmetry analysis. *Mathematical and Computer Modelling*, 25(8–9):115–132, 1997. doi: 10.1016/S0895-7177(97)00063-0.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.

Arthur E. Hoerl and Robert W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970. doi: 10.1080/00401706.1970.10488634.

Gyeonghoon Ko, Hyunsu Kim, and Juho Lee. Learning infinitesimal generators of continuous symmetries from data. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2730. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/9c26a48b7ebd37bc3e979c8651fc3649-Abstract-Conference.html.

Randall J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2007. doi: 10.1137/1.9780898717839.

Ziming Liu and Max Tegmark. Machine learning hidden symmetries. *Physical Review Letters*, 128 (18):180201, 2022. doi: 10.1103/PhysRevLett.128.180201.

Grégoire Mialon, Quentin Garrido, Hannah Lawrence, Danyal Rehman, Yann LeCun, and Bobak Kiani. Self-supervised learning with lie symmetries for partial differential equations. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-1262. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/5c46ae130105fa012da0446126c01d1d-Abstract-Conference.html.

Alexander Novikov, Ngân Vũ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco J. R. Ruiz, Abbas Mehrabian, M. Pawan Kumar, Abigail See, Swarat Chaudhuri, George Holland, Alex Davies, Sebastian Nowozin, Pushmeet Kohli, and Matej Balog. Alphaevolve: A coding agent for scientific and algorithmic discovery. *arXiv preprint arXiv:2506.13131*, 2025. URL https://arxiv.org/abs/2506.13131.

Francesco Oliveri. ReLie: A Reduce program for lie group analysis of differential equations. *Symmetry*, 13(10):1826, 2021. doi: 10.3390/sym13101826. URL https://arxiv.org/abs/2105.11534.

Peter J. Olver. *Applications of Lie Groups to Differential Equations*, volume 107 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1993. doi: 10.1007/978-1-4612-4350-2.

Samuel E. Otto, Nicholas Zolman, J. Nathan Kutz, and Steven L. Brunton. A unified framework to enforce, discover, and promote symmetry in machine learning. *Journal of Machine Learning Research*, 26(248):1–83, 2025. URL https://www.jmlr.org/papers/v26/24-1315.html.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, volume 20, pp. 1177–1184, 2007. URL https://proceedings.neurips.cc/paper_files/paper/2007/hash/013a006f03dbc5392effeb8f18fda755-Abstract.html.

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. doi: 10.1016/j.jcp.2018.10.045.

Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi. Mathematical discoveries from program search with large language models. *Nature*, 625(7995):468–475, 2024. doi: 10.1038/s41586-023-06924-6.

Parshin Shojaee, Kazem Meidani, Shashank Gupta, Amir Barati Farimani, and Chandan K. Reddy. LLM-SR: Scientific equation discovery via programming with large language models. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=m2nmp8P5in.

Silviu-Marian Udrescu and Max Tegmark. AI feynman: A physics-inspired method for symbolic regression. *Science Advances*, 6(16):eaay2631, 2020. doi: 10.1126/sciadv.aay2631.

Juan Luis Vázquez. *The Porous Medium Equation: Mathematical Theory*. Oxford University Press, 2007. ISBN 978-0-19-856903-9. doi: 10.1093/acprof:oso/9780198569039.001.0001.

Yongji Wang, Ching-Yao Lai, Javier Gómez-Serrano, and Tristan Buckmaster. Asymptotic self-similar blow-up profile for three-dimensional axisymmetric Euler equations using neural networks. *Physical Review Letters*, 130(24):244002, 2023. doi: 10.1103/PhysRevLett.130.244002.

Ryota Watanabe, Takanori Ishii, Yuji Hirono, and Hirokazu Maruoka. Data-driven discovery of self-similarity using neural networks. *Physical Review E*, 111(2):024301, 2025. doi: 10.1103/PhysRevE.111.024301.

Zongmin Yu and Liu Yang. Agentic symbolic search: Characterizing PDEs beyond hand-crafted expressions, meshes, and neural networks. *arXiv preprint arXiv:2606.20467*, 2026. URL https://arxiv.org/abs/2606.20467.

A Proofs and Auxiliary Stability Results

Proof of Theorem 2. Set $s = \log t$ in equation 11. At each $\xi \in U$, the residual is $\sum_j e^{(a_j^\top \eta + b_j)s} G_j(\xi)$. Exponentials with distinct real exponents are linearly independent on an open interval. Thus the sum of profile factors in each group with a common exponent vanishes identically on U. Irreducibility excludes two or more nonempty groups. Every active exponent must therefore equal the first, giving $(a_j - a_1)^\top \eta = b_1 - b_j$ and hence $M\eta = q$. A full-column-rank matrix gives at most one solution. Otherwise, if η_0 is a solution, the full solution set of the balance

equations is $\eta_0 + \ker M$, of dimension $d - \operatorname{rank}(M)$. Constraints on the profile, parameters, or problem data can restrict this set. □

Irreducibility is needed: two profile factors may cancel at one time power and another two at a different power. The residual then vanishes without equality between the powers. Applying a single balance equation to these separate groups would impose an unjustified constraint.

Proof of Theorem 3. For a competitor f, set $Y_i = S_f(V_i) - S_{f^\star}(V_i)$. These independent variables lie in $[-B, B]$ and have mean at least Δ. Hoeffding's inequality gives

$$\Pr\left(\frac{1}{n}\sum_{i=1}^{n} Y_i \le 0\right) \le \exp\left(-\frac{n\Delta^2}{2B^2}\right). \tag{14}$$

A wrong empirical minimizer requires this event for at least one of the $|\mathcal{C}| - 1$ competitors. A union bound proves equation 12. □

Proof of Proposition 4. Set the right-hand side of equation 12 to at most δ, take logarithms, and solve for n. Increasing the bank size from m_0 to m_1 changes this sufficient budget by $(2B^2/\Delta^2)\log[(m_1 - 1)/(m_0 - 1)]$ when the score bound and margin are unchanged. □

Proof of Proposition 5. For any $g \neq f^\star$, the uniform deviation bound gives

$$\widehat{C}(g) - \widehat{C}(f^\star) \ge C(g) - C(f^\star) - 2\epsilon \ge \Delta_\lambda - 2\epsilon > 0. \tag{15}$$

Thus $f^\star$ uniquely minimizes the empirical score. Within an equal-risk equivalence class, the population score ordering is the ordering of λk_f. For $\lambda > 0$, a unique minimum-complexity member is preferred. □

Proposition 7 (Stable support decoding)**.** *Let $S = \{j : |\eta_j^\star| > \tau_s\}$ and $\|\widehat{\eta} - \eta^\star\|_\infty \le \epsilon$. If active coordinate magnitudes exceed $\tau_s + \epsilon$ and inactive coordinate magnitudes are below $\tau_s - \epsilon$, thresholding $|\widehat{\eta}_j|$ at τ_s recovers S.*

Proof. For $j \in S$, $|\widehat{\eta}_j| \ge |\eta_j^\star| - \epsilon > \tau_s$. For $j \notin S$, $|\widehat{\eta}_j| \le |\eta_j^\star| + \epsilon < \tau_s$. □

Proposition 8 (Parameter stability from score curvature)**.** *Suppose the parameter objective Q is μ-strongly convex, with $\mu > 0$, on a convex neighborhood containing an interior minimizer $\eta^\star$ and a fitted point $\widehat{\eta}$. If $\nabla Q(\eta^\star) = 0$ and $Q(\widehat{\eta}) - Q(\eta^\star) \le \epsilon$, then $\|\widehat{\eta} - \eta^\star\|_2 \le \sqrt{2\epsilon/\mu}$.*

Proof. Strong convexity gives $Q(\widehat{\eta}) \ge Q(\eta^\star) + (\mu/2)\|\widehat{\eta} - \eta^\star\|_2^2$. Combining this inequality with the score-gap assumption gives the result. □

Proposition 9 (Clock perturbations control field error)**.** *On $|x| \le X$, let $u = e^{-g}F((x - s)e^{-h})$ and define $\widetilde{u}$ using clocks $\widetilde{s}, \widetilde{g}, \widetilde{h}$ and the same profile. Assume $|g|, |\widetilde{g}| \le G$, $|h|, |\widetilde{h}| \le H$, $|s|, |\widetilde{s}| \le S$, and that $|F| \le M$ and F is L-Lipschitz on the required coordinate range. Uniform clock errors at most ϵ imply*

$$\|u - \widetilde{u}\|_\infty \le e^G[M + Le^H(1 + X + S)]\epsilon. \tag{16}$$

Proof. The mean-value theorem gives $|e^{-g} - e^{-\widetilde{g}}| \le e^G\epsilon$. For $z = (x - s)e^{-h}$ and $\widetilde{z} = (x - \widetilde{s})e^{-\widetilde{h}}$, it also gives $|z - \widetilde{z}| \le e^H(1 + X + S)\epsilon$. Add and subtract $e^{-\widetilde{g}}F(z)$, then use the bound and Lipschitz constant of F to obtain equation 16. □

Proof of Theorem 6. Taylor's theorem gives, for $r \in [0, 1]$,

$$|q(r) - P_d q(r)| \le \frac{\|q^{(d+1)}\|_\infty}{(d+1)!} r^{d+1} \le E_d. \tag{17}$$

Apply Proposition 9 to obtain field error at most $C_F E_d$. Under the score-Lipschitz assumption, each candidate's score changes by at most $KC_F E_d$. Every competing score difference is then at least $\Delta - 2KC_F E_d > 0$, so the minimizer is unchanged. □

Proposition 10 (Profile-interpolation stability). *Fix $u(x,t) = a(t)F(z(x,t))$ and $\widetilde{u}(x,t) = a(t)\widetilde{F}(z(x,t))$. Suppose the spatial map is differentiable and one-to-one on the evaluation region, $|a(t)| \le A$, and $|\partial x/\partial z| \le K$. Then, with profile norms on the mapped region,*

$$\|u(\cdot,t) - \widetilde{u}(\cdot,t)\|_{L_2(x)} \le A\sqrt{K}\,\|F - \widetilde{F}\|_{L_2(z)}. \tag{18}$$

Proof. Changing variables at fixed time gives

$$\|u - \widetilde{u}\|^2_{L_2(x)} = \int |a(t)|^2 |F(z) - \widetilde{F}(z)|^2 \left|\frac{\partial x}{\partial z}\right| dz \le A^2 K \|F - \widetilde{F}\|^2_{L_2(z)}. \tag{19}$$

Taking square roots proves the claim. □

B Benchmark and Evaluation Protocols

Benchmark specification. Table 4 specifies the one-dimensional configurations. The operator holdouts retain a mechanism expressible by the grammar; the out-of-bank cases change the generating temporal law.

Table 4: One-dimensional benchmark. Each coefficient configuration has ten seeds.

Partition	Family	Public coefficients
Core	Heat	$\nu \in \{0.18, 0.35, 0.65\}$
Core	Porous medium	$C \in \{0.45, 0.70, 1.00\}$
Core	Advection	$c \in \{-0.85, 0.35, 1.10\}$
Core	Reaction	$\lambda \in \{0.25, 0.75, 1.30\}$
Core	Advection–diffusion	three (c, ν) pairs
Core	Advection–reaction	three (c, λ) pairs
Core	Reaction–diffusion	three (ν, λ) pairs
Held out	Wave / reverse Burgers	0.90 / $(0.40, 0.80, 0.20)$
Held out	Radial heat	$(\nu, d) = (0.25, 2)$
Nonconstant clocks	Acceleration / quadratic decay	0.80 / 0.90
Nonconstant clocks	Dilation / oscillatory shift	0.55 / 0.70
Out of bank	Cubic shift / log damping	0.62 / 0.85
Out of bank	Reciprocal dilation / chirp	0.48 / 0.66

The frame bounds are $\delta \in [-0.2, 1.8]$, $c \in [-1.5, 1.5]$, and $(\alpha, \beta) \in [-0.4, 1.1]^2$. Acceleration and oscillatory-shift coefficients lie in $[-1.5, 1.5]$, quadratic damping in $[-0.2, 1.8]$, and exponential dilation in $[-1.2, 1.2]$. Gaussian profiles use log amplitude $[-3.8, 1.2]$ and log inverse width $[-4.2, 2.2]$. Positive-quadratic profiles use log core $[-1.8, 0.8]$ and log curvature $[-5.4, -0.05]$. Hyperbolic-tangent profiles use log slope $[-2.2, 2.2]$, amplitude $[0.1, 2.0]$, and bias $[-0.6, 0.6]$. Sinusoidal profiles use amplitude $[0.1, 2.0]$, log frequency $[-0.45, 0.45]$, and bias $[-0.5, 0.5]$.

The initial loss divides squared mismatch by the empirical squared initial amplitude plus 10^{-9}. Boundary loss sums mean squared mismatch at the two endpoints. Mass is integrated on 241 spatial points and compared at five times with the initial mass, using its square plus 10^{-9} for normalization. The residual and drift denominators use $\varepsilon = 10^{-7}$. These definitions give the component meanings in equation 6.

Evaluation integrity and parameter sensitivity. Search and evaluation run in separate processes. Search can access the operator, initial function, endpoint boundary data, and declared invariants; it cannot access reference trajectories or frame labels. Boundary queries are restricted to the declared endpoints for both typed and continuous-clock methods. Evaluation loads exact fields and labels only after search has finished and its candidate outputs have been fixed.

Before aggregation, we verify that every scheduled problem, seed, objective, and candidate is represented once, with finite score components. Optimizer exhaustion is retained in the results. Selection-only sensitivity analyses rescore these same fitted candidates; robustness and objective-ablation experiments instead refit the candidates under the modified conditions.

The thirteen robustness designs in Table 5 independently refit the seven-frame, four-analytic-profile bank on 24 in-bank configurations and five seeds. The changed quantity is the interior training-point

count, IC noise fraction, finite-difference step, iteration cap, or objective weight. All other settings retain their default values. These 28-candidate runs test parameter sensitivity within that bank; the 55-candidate objective ablations are reported in main-text Table 2.

Table 5: Independent robustness refits. Each row has 120 trials and 28 fitted frame–profile candidates per trial.

Changed setting	N	Frame %	Solved %
24 interior training points	120	92.5	100.0
48 interior training points	120	93.3	100.0
192 interior training points	120	98.3	100.0
1% IC noise	120	95.8	100.0
5% IC noise	120	95.0	100.0
Difference step 0.001	120	98.3	100.0
Difference step 0.004	120	95.8	100.0
60 optimizer iterations	120	96.7	100.0
300 optimizer iterations	120	96.7	100.0
Initial-loss weight halved	120	98.3	100.0
Initial-loss weight doubled	120	97.5	100.0
Boundary-loss weight doubled	120	99.2	100.0
Drift weight doubled	120	98.3	100.0

Sparse support decoding is also evaluated on the closed candidate pool. Across thresholds $0.005, 0.01, 0.02, 0.05, 0.1$, frame accuracy is 85.0–85.7% and solved rate is 91.4%. Description-penalty sensitivity is combined with the equivalence analysis in Table 14.

Continuous-clock comparison. The continuous model in equation 10 uses the same five profile choices as typed search. Its degree-four comparison uses eleven restarts per profile on all 32 benchmark configurations and ten seeds. Degrees one through three use the same restart budget on the four nonconstant-clock configurations. Table 6 reports field success and clock error separately. The one-block and eight-block typed selectors both achieve the main comparison's 96.8% frame accuracy and 100% solved rate. The latter averages independent complete validation scores, as assumed in Theorem 3.

Table 6: Polynomial degree on four nonconstant-clock families, ten seeds each. Clock error is root-mean-square error (RMSE).

Clock degree	N	Solved %	Median clock RMSE
1	40	45.0	0.329
2	40	100.0	1.02×10^{-5}
3	40	100.0	2.99×10^{-5}
4	40	100.0	2.21×10^{-4}

Score-threshold detector. The scalar fit-score detector calibrates a rejection threshold using 84 in-bank trials with seeds 0–2. Testing uses 196 in-bank trials and 28 out-of-bank trials with seeds 3–9. The out-of-bank families are those in Table 4. Table 7 reports threshold sensitivity.

Table 7: Scalar-score rejection threshold. Calibration uses in-bank seeds 0–2; testing uses seeds 3–9. The factor rescales the calibrated threshold.

Factor	Threshold	In-bank accept %	Out-bank reject %
1.00	7.54×10^{-5}	100.0	75.0
1.25	9.43×10^{-5}	100.0	75.0
1.50	1.13×10^{-4}	100.0	67.9
2.00	1.51×10^{-4}	100.0	50.0

Multicomponent detector. A second detector sums the positive parts of four robustly standardized quantities: best validation score, frame ambiguity, block instability, and the train–validation gap. Each is standardized using its in-bank calibration median and 1.4826 times its median absolute

deviation, with a 10^{-6} scale floor. Calibration again uses in-bank seeds 0–2. The separate novelty set contains quartic translation, square-root-time damping, quadratic-rational dilation, and two-frequency translation, with ten trials per family. Evaluation compares these 40 trials with the 196 in-bank test trials; combined-set evaluation also includes the 28 benchmark out-of-bank trials. The prespecified AUROC target is 0.90. Table 8 reports this detector's discrimination metrics; its different score and novel families preclude a paired comparison with the scalar-score detector.

Table 8: Multicomponent novelty detection. The separate set uses 196 in-bank and 40 novel trials. AUPRC is area under the precision–recall curve.

Metric	Value	Cluster 95% CI
Separate-set AUROC	0.955	$[0.877, 0.997]$
Separate-set AUPRC	0.781	–
Combined-set AUROC	0.910	–

C Supplementary Experiments and Diagnostics

Family-level recovery. The aggregate comparison is given once in main-text Table 1. Table 9 resolves its results by PDE family; the four nonconstant-clock families are further detailed in Table 10.

Table 9: Per-family recovery on the 280 in-bank trials.

Family	N	Typed frame %	Typed solved %	Clock solved %
accelerating advection	10	100.0	100.0	100.0
advection	30	100.0	100.0	100.0
advection diffusion	30	100.0	100.0	100.0
advection reaction	30	100.0	100.0	100.0
dilation advection	10	100.0	100.0	100.0
heat	30	100.0	100.0	100.0
oscillatory advection	10	100.0	100.0	100.0
porous	30	90.0	100.0	100.0
quadratic reaction	10	100.0	100.0	100.0
radial heat	10	100.0	100.0	100.0
reaction	30	100.0	100.0	100.0
reaction diffusion	30	100.0	100.0	100.0
reverse burgers	10	100.0	100.0	100.0
wave	10	40.0	100.0	100.0

Table 10: Field errors on the four nonconstant-clock families. Each row has ten seeds.

Extension family	N	Frame %	Median rel. L_2
accelerating advection	10	100.0	4.17×10^{-7}
dilation advection	10	100.0	4.03×10^{-7}
oscillatory advection	10	100.0	1.33×10^{-6}
quadratic reaction	10	100.0	4.25×10^{-7}

Profile transfer, grammar size, and search budget. The unseen-profile suite contains ten problems and five seeds. All analytic profile choices are removed, leaving the spline of the public IC. Both methods use these same observations. Continuous results at different restart budgets appear as separate rows in Table 11.

Table 11: Unknown-profile comparison with only the IC-interpolant profile available.

Method	Restarts	N	Frame %	Solved %	Median rel. L_2
typed grammar	1	50	100.0	100.0	4.19×10^{-6}
continuous clock	1	50	–	100.0	2.79×10^{-5}
continuous clock	3	50	–	100.0	1.54×10^{-5}

The grammar stress suite uses nested banks with 0, 25, 50, or 100 deterministic nuisance clocks, on 14 family representatives and five seeds. The temporal functions are held fixed across runs. Table 12 reports whether any nuisance program is selected. The continuous restart study in Table 13 uses all 32 configurations, including the four out-of-bank problems, for 320 trials at each budget. Its population therefore differs from the 280 in-bank trials in the main comparison.

Table 12: Nested grammar stress with deterministic nuisance clocks.

Nuisance programs	N	Frame %	Solved %	Nuisance selected
0	70	91.4	100.0	0
25	70	91.4	100.0	0
50	70	91.4	100.0	0
100	70	91.4	100.0	0

Table 13: Continuous-clock restart budget on all 320 problem–seed trials. Time is single-core wall time.

Restarts	N	Solved %	Median rel. L_2	Median s/trial	Evaluations
1	320	97.2	1.38×10^{-4}	20.9	11,878
3	320	99.7	1.26×10^{-4}	73.4	42,414
11	320	100.0	1.28×10^{-4}	301.8	163,921

Canonical equivalence is evaluated by rescoring the same fitted candidates under different penalties (Table 14); no parameters are refitted for this analysis. The median best-versus-second score margin is 4.60×10^{-8}, and the median maximum candidate block standard deviation is 0.768. The median per-trial ratio of the margin to the maximum block standard deviation is 2.57×10^{-9}. These statistics describe empirical separation but do not establish the bounded population-risk assumptions needed for a numerical certificate from Theorem 3.

Table 14: Selection-only equivalence analysis. Strict labels, thresholded active transforms, and field equivalence use the definitions in Section 5. "Median active" is the selected support size.

λ	N	Strict frame %	Support frame %	Equiv. %	Median active
0	280	68.2	100.0	100.0	2
10^{-6}	280	93.9	100.0	100.0	2
10^{-5}	280	96.8	100.0	100.0	2
10^{-4}	280	98.2	100.0	100.0	2

Objective interactions and initial-data stress. The interaction suite independently refits five objectives on the same 14-family, five-seed design as main-text Table 2. It uses residual alone, initial data alone, or the residual paired with boundary, mass, or drift. Table 15 gives both structural and field success rates; Figure 2a presents the family-cluster uncertainty.

Table 15: Objective interactions. Each of the 70 trials refits all 55 candidates.

True-refit objective	N	Frame %	Solved %
initial only	70	0.0	0.0
residual boundary	70	74.3	65.7
residual drift	70	42.9	1.4
residual mass	70	50.0	11.4
residual only	70	42.9	5.7

The unknown-profile noise study reruns the ten problems with 41 clean samples, 161 samples at 10% noise, or 41 samples at 5% noise. Noise is scaled by the IC standard deviation. Typed and degree-four continuous fitting use identical public data and collocation designs, one restart, and the same optimization cap. Table 16 reports the resulting field errors as well as success.

Downstream surrogate. After selection is fixed, the downstream experiment fits a random-feature ridge surrogate to trajectory samples. Training inputs are standardized using their training-set means

Table 16: IC sampling and noise stress for unknown profiles.

IC design	Method	N	Frame %	Solved %	Median rel. L_2
161 samples, 10% noise	Typed	50	36.0	14.0	0.286
161 samples, 10% noise	Continuous clock	50	–	20.0	0.227
41 clean samples	Typed	50	100.0	100.0	0.0012
41 clean samples	Continuous clock	50	–	100.0	0.00121
41 samples, 5% noise	Typed	50	88.0	100.0	0.0385
41 samples, 5% noise	Continuous clock	50	–	100.0	0.0382

and standard deviations. The feature map contains an intercept and 63 cosines, with frequencies drawn from $\mathcal{N}(0, I)$ and phases uniform on $[0, 2\pi]$. Ridge penalty is 10^{-6}, and the intercept is unpenalized. The four representations are physical (x, t), selected coordinates, oracle coordinates, and a stationary one-dimensional control. Feature seeds, samples, and test points are paired. This offline fitting has no feedback to selection.

Table 17: Downstream surrogate. Errors are median relative L_2; the final column is the paired geometric-mean selected/physical error ratio over 50 instances.

Samples	Physical	Selected	Oracle	Selected/physical
32	3.37	0.55	0.51	0.16
64	0.881	0.423	0.419	0.443
128	0.472	0.411	0.411	0.807

The selected/physical ratio intervals at 32, 64, and 128 samples are $[0.133, 0.205]$, $[0.352, 0.549]$, and $[0.720, 0.871]$. For the exploratory stationary-control comparison, the corresponding ratios are 1.037, 1.027, and 0.932, with all intervals crossing one. These are paired ratios across problems, rather than ratios of the marginal medians in Table 17.

Two-dimensional transfer and discretization. The two-dimensional transport operator has affine velocity $v(x) = Bx + c$, where $B \in \mathbb{R}^{2\times 2}$ and $c \in \mathbb{R}^2$. Translation, rotation, dilation, and translation plus rotation each contribute three coefficient configurations and five seeds. Typed search evaluates five declared frames; the continuous control jointly fits all six entries of (B, c). Both receive the public operator and a discretely sampled initial condition. Table 18 gives the transfer result.

Table 18: Two-dimensional affine-flow transfer: four families, three configurations, and five seeds.

Method	N	Frame %	Solved %	Median rel. L_2
typed two-dimensional grammar	60	100.0	100.0	7.16×10^{-5}
continuous affine	60	–	100.0	7.16×10^{-5}

The discretization study fixes one configuration per family and five seeds, then independently refits on clean 33^2, 65^2, and 129^2 IC grids and noisy 65^2 grids. Noise is scaled by the initial-field standard deviation. Table 19 includes all 200 method–design trials.

Upwind solver with search cost included. Reference trajectories are computed independently on a 193^2 grid on $[-2, 2]^2$ (LeVeque, 2007). Exterior values are zero. At each matched resolution, the physical, selected, continuous, oracle, and stationary representations use the same sign-split first-order update,

$$u_{ij}^{n+1} = u_{ij}^n - \Delta t \left[v_{x,ij}^+ D_x^- u_{ij}^n + v_{x,ij}^- D_x^+ u_{ij}^n + v_{y,ij}^+ D_y^- u_{ij}^n + v_{y,ij}^- D_y^+ u_{ij}^n \right], \tag{20}$$

where $a^+ = \max(a, 0)$, $a^- = \min(a, 0)$, and D_q^- and D_q^+ are backward and forward differences in coordinate q. The indices i, j locate spatial grid points, n indexes time, and v_x, v_y are the velocity components. For grid spacings $\Delta x, \Delta y$, the time step Δt satisfies

$$\Delta t \max_{i,j} \left(\frac{|v_{x,ij}|}{\Delta x} + \frac{|v_{y,ij}|}{\Delta y} \right) \le 0.32. \tag{21}$$

Table 20 reports each method separately. Search is included in the selected and continuous total times; the other rows report solve time. The paired error ratios and win rates are given in main-text Table 3.

Table 19: Two-dimensional discretization and IC-noise stress. Each method–design row contains four family representatives and five seeds.

IC grid and noise	Method	N	Frame %	Solved %	Median rel. L_2
129^2, clean	Continuous affine	20	–	100.0	3.86×10^{-6}
129^2, clean	Typed	20	100.0	100.0	3.86×10^{-6}
33^2, clean	Continuous affine	20	–	100.0	1.74×10^{-3}
33^2, clean	Typed	20	95.0	100.0	1.74×10^{-3}
65^2, clean	Continuous affine	20	–	100.0	6.92×10^{-5}
65^2, clean	Typed	20	100.0	100.0	6.92×10^{-5}
65^2, 1% noise	Continuous affine	20	–	100.0	5.83×10^{-3}
65^2, 1% noise	Typed	20	95.0	100.0	5.83×10^{-3}
65^2, 5% noise	Continuous affine	20	–	100.0	2.87×10^{-2}
65^2, 5% noise	Typed	20	95.0	100.0	2.87×10^{-2}

Table 20: Full upwind comparison. Every method–grid row has 60 matched trials.

Grid	Method	N	Solved %	Median rel. L_2	Median total s
33	continuous affine	60	100.0	0.0293	0.000996
33	oracle frame	60	100.0	0.0293	0.000409
33	physical upwind	60	16.7	0.159	0.00118
33	selected frame	60	100.0	0.0293	0.0702
33	stationary control	60	0.0	0.566	0.000114
65	continuous affine	60	100.0	0.0357	0.00151
65	oracle frame	60	100.0	0.0357	0.000839
65	physical upwind	60	66.7	0.0745	0.00318
65	selected frame	60	100.0	0.0357	0.0707
65	stationary control	60	0.0	0.566	0.000224
129	continuous affine	60	100.0	0.043	0.00339
129	oracle frame	60	100.0	0.043	0.00258
129	physical upwind	60	100.0	0.0196	0.0148
129	selected frame	60	100.0	0.043	0.0726
129	stationary control	60	0.0	0.566	0.000658

Restricted symbolic controls. For ten one-dimensional first-order templates, a symbolic control integrates public translation, amplitude, or dilation rates and verifies their clock ordinary differential equations (ODEs); it recovers all ten frames. In two dimensions, the affine transport control recovers all 12 configurations (Table 21). These controls operate on declared templates rather than solving general Lie determining equations.

Table 21: Symbolic recovery within the affine-transport template family.

Control	N	Correct %	Declared scope
restricted symbolic affine	12	100.0	not a general Lie CAS

D Reference Solution Shapes and Coordinate Reduction

Figure 3 visualizes six analytic solution families and the geometry behind their frame programs. These are reference illustrations, not additional fitted trials. The formulas below are evaluated directly; no reference trajectory is supplied to the search procedure.

Table 22: Reference fields used in the solution atlas. Diffusive fields use $t \in [1, 2.5]$; the other fields use $t \in [0, 1.5]$.

Family	Reference field $u(x,t)$	Spatial interval
Heat	$(1.4\pi t)^{-1/2} \exp[-x^2/(1.4t)]$	$[-4, 4]$
Porous medium	$t^{-1/3}(0.70 - x^2 t^{-2/3}/12)_+$	$[-5, 5]$
Advection	$\tanh[2(x - 1.10t)]$	$[-3, 3]$
Reaction	$e^{-0.75t} \sin x$	$[-\pi, \pi]$
Accelerated transport	$\tanh[2(x - 0.40t^2)]$	$[-3, 3]$
Quadratic-time damping	$e^{-0.45t^2} \sin x$	$[-\pi, \pi]$

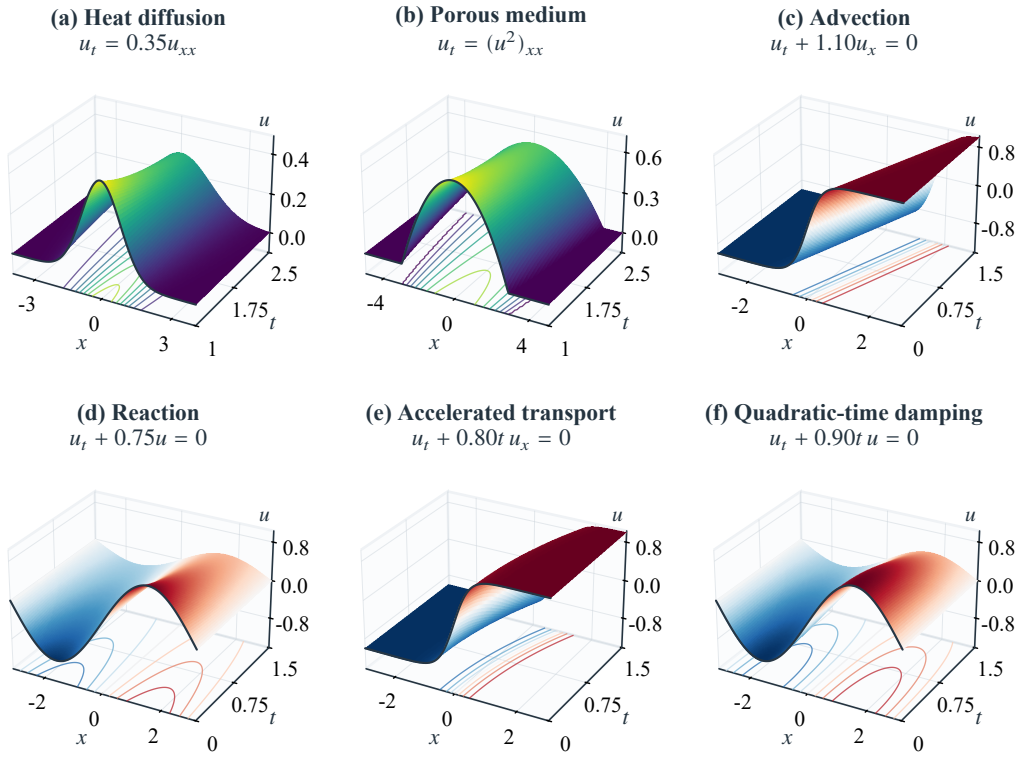

Figure 3: **Geometry of six reference solutions.** Height and color encode $u(x,t)$; projected contours show the corresponding space–time structure. The dark leading curve is the initial profile. Panels show diffusion of a smooth pulse, expansion of a compact support, translation of a front, and attenuation of an oscillatory profile. The final two panels use nonconstant temporal rates. Each panel has its own height scale; see Table 22 for formulas and domains.

Physical evolution and canonical-coordinate collapse. Figure 4 isolates four transformations: self-similar spreading, constant-speed transport, accelerated transport, and quadratic-time damping. The left column retains the physical coordinate; the right column applies the analytic coordinate and amplitude transformations printed on its axes. For plot colors only, $\tau = (t - t_0)/(t_1 - t_0)$ denotes normalized time, distinct from the power-law clock in equation 2. The four colors mark evenly spaced times within each panel's declared interval.

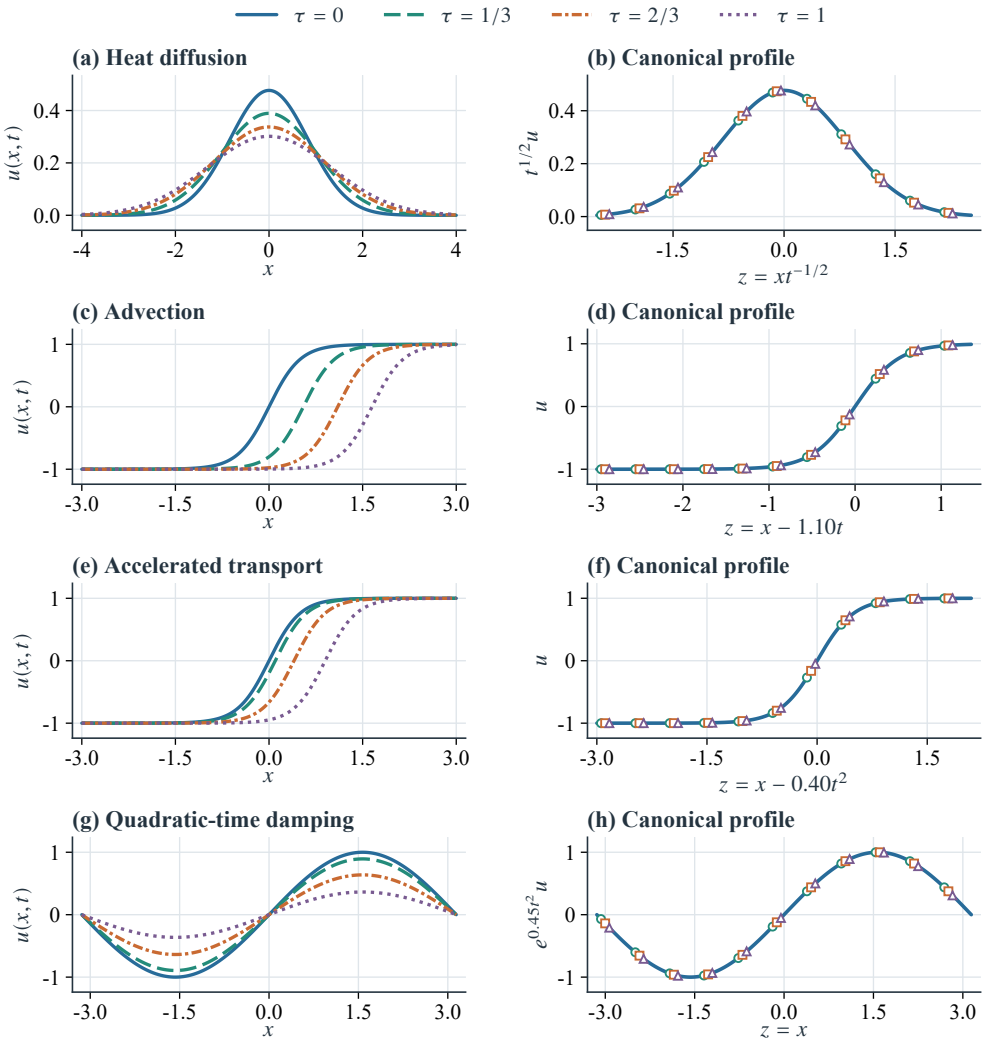

Figure 4: **From evolving solutions to stationary profiles.** Left: four time slices of the analytic reference fields. Right: the same slices after canonical coordinate and amplitude normalization. Open circles, squares, and triangles distinguish the later times when their curves coincide with the initial profile. Right-column curves are restricted to the intersection of the mapped physical domains, so no spatial extrapolation is used. The collapse is an identity of the displayed reference solutions, not a measured frame-recovery or prediction result.

Translation removes motion; diffusion requires both spatial and amplitude rescaling; damping changes only amplitude. These identities illustrate what a frame represents. Fitted performance remains as reported in the preceding results.