

# *Skeptical review: Data-Driven Discovery and Validation of Governing Equations for a Turbulent Fluid System*

---

## Summary

The paper presents a data-driven pipeline to discover governing PDEs for a 3D turbulent(-like), weakly compressible fluid system from high-resolution simulation snapshots of density  $\rho$  and velocity components on a  $128^3$  periodic grid (Sec. 2.1). Spatial derivatives are computed with FFT-based spectral differentiation and temporal derivatives with finite differences (Sec. 2.2), then a 66-term candidate library is regressed against  $\partial_t$  of each field using cross-validated LASSO followed by OLS refinement (Sec. 2.3–2.4). The learned momentum equations achieve moderate  $R^2$  ( $\approx 0.57$ – $0.71$ ) and contain interpretable structures consistent with Navier–Stokes-like dynamics (advection, density-gradient proxy for a gradient force, Laplacian dissipation, and compressibility/bulk-viscosity-like terms) (Sec. 3.2.1). The density equation fits poorly ( $R^2 \approx 0.30$ ) and includes an unphysical negative diffusion (anti-diffusion) Laplacian term attributed to the near-constant density and derivative noise (Sec. 3.2.2). The discovered PDEs are validated by forward integration with a semi-implicit Crank–Nicolson / explicit Euler scheme (Sec. 2.5, Sec. 3.3), showing plausible macroscopic statistics and qualitative texture preservation, while pointwise RMSE grows over time as expected in chaotic systems. Overall, the end-to-end demonstration on a 3D turbulent dataset is valuable, but key details needed for reproducibility and physical interpretability are missing (explicit final equations, feature/target scaling, CV strategy, solver specifics, data provenance), and the validation would be significantly strengthened by turbulence-relevant diagnostics and robustness/identifiability analyses.

## Strengths

- Addresses a challenging and relevant problem: PDE discovery for high-dimensional 3D turbulent(-like) flows from spatiotemporal data (Sec. 1, Sec. 2.1).
- End-to-end evaluation: goes beyond regression metrics by integrating the discovered PDEs forward in time and comparing statistics/structures (Sec. 2.5–2.6, Sec. 3.3).
- Spectral differentiation is well-motivated for a periodic box and can yield accurate spatial derivatives when implemented with appropriate conventions (Sec. 2.2).
- Momentum-equation terms are largely physically interpretable (advection, gradient-force surrogate, viscous/bulk-viscous operators) and achieve moderate predictive  $R^2$  (Sec. 3.2.1).
- Exploratory data analysis clearly highlights the key regime challenge: density has extremely low variance relative to velocity fields, foreshadowing identifiability issues (Sec. 3.1).
- The paper acknowledges limitations (especially the problematic density equation) rather than over-claiming success (Sec. 3.2.2, Sec. 4).

- Several figures effectively communicate qualitative comparisons across time and variables (notably Figures 2, 4, 5, 8), and the manuscript narrative is generally coherent.

## Major issues

1. **The final discovered PDE system is not presented explicitly and cleanly for all fields, limiting interpretability and reproducibility (Sec. 3.2.1–3.2.2, Sec. 4).** The  $v_x$  equation appears truncated (dangling "+ 4"), includes apparent duplicated terms, and the  $v_y$  and  $v_z$  equations are described only qualitatively as having "similar structures". The density equation is shown but not clearly separated into trustworthy vs artifact components.

*Recommendation:* Add a dedicated subsection (e.g., Sec. 3.2.3) and a table listing the final discovered PDEs for  $\rho$ ,  $v_x$ ,  $v_y$ ,  $v_z$  exactly as used for validation: each selected term, coefficient, and operator notation, with no truncation/duplication. Explicitly state which equation(s)/terms are excluded from simulation (if any). For  $\rho$ , clearly label physically plausible continuity-like terms vs the anti-diffusive artifact and state whether the reported  $\rho$  PDE is intended as a model or a diagnostic failure case.

2. **Coefficient magnitudes and physical interpretations are currently not verifiable due to feature normalization and ambiguous scaling of derivatives/time (Sec. 2.3, Sec. 3.2.1).** The library columns are normalized to unit  $L^2$  norm (Sec. 2.3), but later coefficients are interpreted as if attached to unnormalized operators (viscosity, "pressure" scaling, etc.). Additionally, the manuscript's time-step inference from an advection coefficient is inconsistent with the stated finite-difference definition (Eq. (1), Sec. 3.2.1).

*Recommendation:* State explicitly whether reported coefficients (Sec. 3.2, Eqs. (4)–(5)) are in the normalized-feature basis or converted back to the unnormalized operator basis. If conversion is performed, provide the exact unnormalization formula (including any scaling of the target  $\partial_t f$ ), and report both normalized and de-normalized coefficients (or at least key terms) with units/nondimensional form. Re-derive the relationship between learned coefficients and physical coefficients starting from Eq. (1); remove or correct the  $\Delta t \approx 1/5.5$  claim unless an explicit rescaling step makes that inversion valid. If underlying physical units are unknown, frame coefficients as "effective" and avoid parameter identification claims.

3. **Sparse regression/model selection details are under-specified, and the procedure appears to include post-hoc/manual collinearity filtering that can materially affect the discovered PDE (Sec. 2.4, Sec. 3.2.1).** In addition, standard cross-validation can leak information for spatiotemporal fields if folds are formed by randomly sampling grid points, inflating apparent generalization.

*Recommendation:* Expand Sec. 2.4 to fully specify: LassoCV configuration ( $\alpha$  grid/range, number of folds, scoring, convergence tolerances, random seed), whether data are subsampled/batched given  $\approx 8 \times 128^3$  rows, and the exact train/validation splitting strategy. Use blocked CV that respects correlation structure (e.g., folds by time index; or spatial block CV; or both) and report how results change relative to random-row CV. Replace qualitative “collinear statistical artifact” removal with a documented, algorithmic rule (e.g., correlation threshold, VIF/condition number cut-off, or grouped variables), or adopt elastic net / group lasso / reparameterization to reduce researcher degrees of freedom. Provide a brief robustness report: term-selection frequency and coefficient variability under different seeds, folds, and subsampling.

4. **Identifiability is threatened by a redundant candidate library (exact/near-exact dependencies such as Laplacian vs individual second derivatives, and multiple equivalent forms of transport/flux terms), making coefficient attribution unstable even if prediction is adequate (Sec. 2.3, Sec. 3.2.1).** This is compounded by OLS refitting on correlated selected features, which can amplify variance.

*Recommendation:* Revise Sec. 2.3 to reduce exact linear dependencies (e.g., include either  $\nabla^2 f$  or  $\partial^2 f / \partial x_i^2$  terms, not both; avoid simultaneously including multiple algebraically equivalent transport forms unless constrained). Consider a physics-structured library (e.g., conservative-form terms like  $-\nabla \cdot (\rho v)$ ,  $-\nabla \cdot (\rho v \otimes v)$ , gradient/divergence invariants) to reduce collinearity. Report diagnostics (e.g., condition number of selected design matrix; pairwise correlations; coefficient path stability). If you keep multiple components of  $\nabla(\nabla \cdot v)$ , either enforce a shared coefficient (isotropic bulk viscosity) or explicitly present it as a general linear combination (see also Sec. 3.2.1 discussion).

5. **The density ( $\rho$ ) equation discovery is currently presented in a way that undermines the claim of recovering a coupled system: fit quality is low ( $R^2 \approx 0.30$ ) and the learned PDE contains an unphysical negative diffusion term (Sec. 3.2.2, Sec. 4).** The manuscript acknowledges the artifact but does not systematically test mitigations or clearly decide whether  $\rho$  should be modeled at all from this dataset.

*Recommendation:* In Sec. 3.2.2, explicitly mark the anti-diffusive  $\nabla^2 \rho$  term as physically inconsistent and (unless you can justify it) exclude it from any predictive model. Run at least one mitigation experiment targeted to the stated failure mode: (i) improved temporal differentiation (higher-order, smoothing, Savitzky–Golay, or regularized differentiation), (ii) constrained regression enforcing non-negative diffusion, (iii) revised library excluding second-order  $\rho$  terms, and/or (iv) reweighting loss to account for small  $\rho$  variance. Report the resulting  $\rho$  equation and whether validation improves. If reliable  $\rho$  dynamics cannot be recovered, state this clearly and restrict the validated model to momentum (with  $\rho$  treated as constant/known).

**6. Data provenance and the true data-generating equations/parameters are insufficiently described, weakening physical claims (Sec. 2.1, Sec. 3.2.1).** In particular, interpreting  $-c\nabla\rho$  as a pressure-gradient surrogate requires an equation of state or a known formulation; otherwise it may simply be an empirical gradient proxy. Without knowing nondimensionalization, forcing, viscosity, Mach/Reynolds numbers, and the actual snapshot spacing, it is hard to judge whether the discovered PDE matches Navier–Stokes or an “effective” model.

*Recommendation:* In Sec. 2.1, provide the dataset provenance: the solver/model used to generate the data (compressible vs incompressible; any equation of state; forcing; viscosity; etc.), nondimensionalization, and the actual time spacing between snapshots if known. Report relevant regime indicators ( $\text{Re}$ ,  $\text{Ma}$ ) or explicitly state they are unknown. In Sec. 3.2.1, either (i) compare discovered coefficients/terms quantitatively to the known governing equations after consistent rescaling, or (ii) reframe conclusions as learning an effective PDE in arbitrary units and interpret  $\nabla\rho$  as an empirical gradient field rather than “pressure” unless justified.

**7. Validation is currently too coarse for a turbulence setting: means/standard deviations, RMSE, and slice visuals do not assess scale-dependent dynamics and can miss significant spectral/structural errors (Sec. 2.6, Sec. 3.3).** RMSE growth alone is also difficult to interpret without normalization and residual diagnostics.

*Recommendation:* Augment Sec. 2.6 and Sec. 3.3 with turbulence-relevant diagnostics: kinetic energy spectrum  $E(k)$  and its evolution; enstrophy or vorticity-magnitude PDFs; divergence statistics  $\langle(\nabla \cdot v)^2\rangle$  to quantify compressibility; spatial correlation functions or structure functions; and time series of total kinetic energy. Report residual diagnostics for the regression stage (e.g., spectrum of residual  $\partial_t v - \Theta\xi$ , correlation of residual with candidate terms). Include at least one baseline comparison (e.g., advection+Laplacian only) to show the marginal value of additional discovered terms.

**8. The numerical integration used for validation is under-specified, and it is unclear how numerical artifacts (aliasing, stabilization) and problematic terms (e.g., anti-diffusion in  $\rho$ ) are handled (Sec. 2.5, Sec. 3.3).** Reproducibility and credibility of the validation depend on these details.

*Recommendation:* Expand Sec. 2.5 with an explicit algorithm: periodic boundary conditions; spatial discretization (spectral/pseudospectral vs finite differences); how nonlinear products are computed; whether de-aliasing (e.g., 2/3 rule) or filtering is used; which terms are treated implicitly in Crank–Nicolson vs explicitly in Euler; how mixed derivatives are evaluated; and the simulation time step  $\Delta t_{\text{sim}}$  with stability/CFL considerations. Explicitly state whether the anti-diffusive  $\rho$  Laplacian is included; if included, explain why the simulation remains stable; if excluded, say so. Provide a simple  $\Delta t_{\text{sim}}$  sensitivity/convergence check on one metric (e.g., energy or spectrum) to rule out time-discretization artifacts.

9. **Positioning/novelty relative to established PDE-discovery literature is incomplete, which makes it hard to assess what is new beyond applying a known sparse-regression pipeline to a 3D turbulent dataset (Sec. 1, Sec. 4).**

*Recommendation:* Add a related-work subsection (e.g., Sec. 1.1) explicitly citing and contrasting with PDE-FIND/SINDy-style sparse regression, weak-form PDE discovery, and neural/hybrid approaches. Clarify what the paper contributes beyond prior work (e.g., specific 3D turbulence data regime, spectral derivative choices, simulation-based validation, analysis of near-constant field failure), and temper novelty claims where the method largely follows established frameworks.

## Minor issues

1. The 66-term library is not enumerated explicitly, making exact reproduction and interpretation difficult (Sec. 2.3).

*Recommendation:* Provide a complete list of all 66 candidate terms in Sec. 2.3, an appendix, or supplementary material. Group by type (constants, linear, first derivatives, second derivatives, advection-like, divergence/flux forms, nonlinear products) and specify how vector/tensor expressions are expanded into scalar regressors for each equation.

2. Temporal differentiation accuracy is not evaluated despite the extreme limitation of only 10 snapshots (8 usable central differences) and derivative noise being central to the density failure (Sec. 2.2, Sec. 3.2.2).

*Recommendation:* Add a short sensitivity study: compare central differences to higher-order schemes and/or temporal smoothing/regularized differentiation, reporting changes in  $R^2$  and selected terms (especially for  $\rho$ ). This will justify the chosen derivative estimator and clarify whether the density artifact is inevitable under the data regime.

3. FFT derivative scaling is ambiguous because  $k_x$  (and  $2\pi/L$  factors, indexing conventions) is not fully defined, which affects coefficient scaling and unit consistency (Eq. (2), Sec. 2.3).

*Recommendation:* Define the wavenumber convention precisely: whether  $k$  are angular wavenumbers, how  $k$  grids are constructed (FFT ordering), and what domain length  $L$  is assumed in each direction. This makes Eq. (2) uniquely specified and helps interpret coefficient magnitudes.

4. Reported  $R^2$  values for momentum are moderate, but the paper does not provide baselines or uncertainty; with huge sample size, small systematic errors can matter (Sec. 3.2.1).

*Recommendation:* Add baseline models (e.g., advection only; advection + Laplacian) and compare  $R^2$ /MSE and validation diagnostics. Provide uncertainty via bootstrapping over time slices or spatial blocks (term-selection frequency and coefficient intervals).

5. Evaluation metrics (RMSE, etc.) are not formally defined with normalization/aggregation details, complicating interpretation and cross-variable comparisons (Sec. 2.6, Sec. 3.3.2; Figure 8).

*Recommendation:* Define RMSE precisely (per-field vs joint; across all grid points; normalization by standard deviation or not). In Figure 8 and Sec. 3.3.2, report normalized errors (e.g.,  $\text{RMSE}/\sigma$ ) and clarify the mapping from index to physical/simulation time.

6. Several figure-design choices reduce interpretability (e.g., sequential colormap for signed velocities; missing quiver scale; CV plot suggests  $\alpha$  grid bound is active) (Figures 2, 4, 5).

*Recommendation:* Use diverging colormaps centered at 0 for velocities (and centered at 1 for density if plotting deviations), fix color limits across times, and use shared colorbars where possible (Figure 2). Add a quiver key and reduce clutter (Figure 4). Extend the  $\alpha$  grid, plot CV uncertainty bands and model sparsity vs  $\alpha$ , and keep axes comparable across panels (Figure 5).

## Very minor issues

1. Text appears to contain placeholders for authorship/affiliation (e.g., “Anthropic, Gemini & OpenAI servers. Planet Earth.”) and informal identifiers, which is inappropriate for submission formatting.

*Recommendation:* Replace placeholders with the correct author list, affiliations, and acknowledgments; ensure the manuscript meets the target venue’s anonymization policy if applicable.

2. Minor equation/text formatting errors and typos reduce polish: duplicated/missing symbols in statistics listings, inconsistent  $v_x$  vs  $v_x$  notation, occasional malformed variables (e.g.,  $v_{-x}$ ), repeated mixed-derivative terms, and the trailing “+ 4” in the  $v_x$  equation (Sec. 3.1–3.2.1).

*Recommendation:* Proofread and correct typos and malformed notation; ensure all displayed equations are complete and consistent with the code used to generate results. Typeset  $128^3$  consistently (avoid “1283”).

3. Section/heading styles are inconsistent (capitalization, heading levels, trailing punctuation), which complicates cross-referencing (Sec. 2–4).

*Recommendation:* Standardize section hierarchy and capitalization (e.g., “2 Methods”, “2.1 Dataset”, etc.) and ensure consistent equation numbering and referencing (Eq. (1), Eq. (2), ...).

4. The discussion of chaos/“butterfly effect” is plausible but presented without any quantitative support (Sec. 3.3.2).

*Recommendation:* Either add a simple quantitative measure (e.g., correlation decay time or an estimated Lyapunov time proxy from error growth) or explicitly frame the chaos explanation as qualitative.

## Key statements and references

- • **Filtering collinear artifacts from the sparse-regression output, the discovered evolution equation for the x-component of velocity takes the approximate form  $\partial v_x / \partial t \approx -5.17(\mathbf{v} \cdot \nabla)v_x - 3.34 \partial \rho / \partial x + 0.66 \nabla^2 v_x + 5.87 \partial^2 v_x / \partial x^2 + 4.90 \partial^2 v_y / \partial x \partial y + 4.49 \partial^2 v_z / \partial x \partial z + 4$ , with analogous structures for  $v_y$  and  $v_z$ , and these momentum equations achieve coefficients of determination  $R^2$  of 0.658, 0.709, and 0.566 for  $v_x, v_y, v_z$ , respectively.**
- *Reference(s):* 11
- • **The coefficients of the non-linear advection terms in the discovered momentum equations, which average approximately  $-5.5$  compared to the physically expected value of  $-1$ , imply that the effective physical time step underlying the data is about  $\Delta t \approx 1/5.5 \approx 0.18$  in arbitrary units.**
- *Reference(s):* 11
- • **For the density field, sparse regression yields an evolution equation  $\partial \rho / \partial t \approx -0.059(\rho \nabla \cdot \mathbf{v}) - 0.034(\mathbf{v} \cdot \nabla \rho) - 0.020 \nabla^2 \rho$  with  $R^2 = 0.303$ , where the first two terms correspond to the continuity equation  $\partial \rho / \partial t = -\nabla \cdot (\rho \mathbf{v})$  but the negative Laplacian term represents an unphysical anti-diffusion attributed to the extremely low variance of  $\rho$  relative to numerical noise.**
- *Reference(s):* 11
- • **Numerical integration of the discovered PDE system using a semi-implicit Crank–Nicolson scheme for linear terms and explicit Euler for non-linear terms with  $\Delta t_{sim} = 0.05$  preserves macroscopic statistics—e.g., the simulated density mean evolves from 1.000000 at  $t = 1$  to 1.002234 at  $t = 8$ , closely matching the ground-truth mean of 0.999999 at  $t = 8$ —while the RMSE of the velocity components grows nonlinearly to 0.521 ( $v_x$ ), 0.347 ( $v_y$ ), and 0.265 ( $v_z$ ) by  $t = 8$ , consistent with chaotic divergence.**
- *Reference(s):* 11

## Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

**Maths relevance:** substantial

The paper’s core analytic content is the definition of discrete temporal derivatives, spectral spatial derivatives, construction of a candidate operator library  $\Theta$ , and linear sparse regression to identify PDE right-hand sides for  $\rho$  and  $v$ . The mathematics is mostly definitional, with limited step-by-step derivations; key internal-consistency checks therefore focus on (i) correctness of calculus/operator identities used for interpretation, (ii) scaling effects from normalization and the choice  $\Delta t = 1$  in Eq. (1), and (iii) whether reported PDE coefficients correspond to the stated operators.

### Checked items

1. ✓ **Central finite-difference time derivative** (Eq. (1), Sec. 2.2, p.2)
  - **Claim:** For interior times  $t_1..t_8$ ,  $\partial f/\partial t(t_i) \approx (f(t_{i+1}) - f(t_{i-1}))/(2\Delta t)$ .
  - **Checks:** algebra, definition-consistency
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:** Time slices are equally spaced by  $\Delta t$ ., Central difference applied only to interior indices.
  - **Notes:** Formula is correct for a second-order central difference; discarding endpoints is consistent with using a uniform-accuracy target for regression.
2. ✓ **Implication of setting  $\Delta t = 1$  on coefficient scaling** (Sec. 2.2 (discussion following Eq. (1)), p.2)
  - **Claim:** Setting  $\Delta t = 1$  means discovered coefficients absorb any true time-step scaling factor.
  - **Checks:** dimensional/scaling analysis, definition-consistency
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:** True snapshot spacing is  $\Delta t_{\text{true}}$ , but the derivative is computed with  $\Delta t_{\text{assumed}} = 1$  in Eq. (1)., RHS library terms are computed from instantaneous fields (not multiplied by  $\Delta t$ ).
  - **Notes:** If  $\Delta t$  is mis-specified, coefficients scale accordingly (proportional to  $\Delta t_{\text{true}}$  under Eq. (1) as used). The paper’s later inference of  $\Delta t$  is the problematic part (checked separately).
3. △ **Spectral derivative operator** (Eq. (2), Sec. 2.3, p.3)
  - **Claim:**  $\frac{\partial f}{\partial x} = \mathcal{F}^{-1}(ik_x \mathcal{F}(f))$  on a periodic domain.
  - **Checks:** operator identity, units/scaling
  - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate

- **Assumptions/inputs:**  $k_x$  is the correct discrete wavenumber array matching the FFT convention and domain length  $L$ .
  - **Notes:** The form is standard, but  $k_x$  is not defined precisely (angular vs cyclic frequency; inclusion of  $2\pi/L$ ). Without that, the operator's scaling is ambiguous, which propagates to coefficient interpretation.
4. ✓ **Laplacian and divergence definitions** (Sec. 2.3 (library description), p.3)
- **Claim:**  $\nabla^2 f = f_{xx} + f_{yy} + f_{zz}$  and  $\nabla \cdot v = v_{x,x} + v_{y,y} + v_{z,z}$ .
  - **Checks:** algebra, notation consistency
  - **Verdict:** PASS; confidence: high; impact: minor
  - **Assumptions/inputs:** Cartesian coordinates with standard vector calculus conventions.
  - **Notes:** Definitions are correct and consistent with later references.
5. ✓ **Advection-term component expansion** (Sec. 2.3 (Advection terms bullet), p.3)
- **Claim:** The  $x$ -momentum advection component is  $(v \cdot \nabla)v_x = v_x v_{x,x} + v_y v_{x,y} + v_z v_{x,z}$ .
  - **Checks:** algebra, operator identity
  - **Verdict:** PASS; confidence: high; impact: minor
  - **Assumptions/inputs:**  $v = (v_x, v_y, v_z)$  and  $\nabla = (\partial x, \partial y, \partial z)$ .
  - **Notes:** Correct expansion.
6. ✓ **Linear regression formulation** (Eq. (3), Sec. 2.4, p.3)
- **Claim:** For each field  $f$ , the regression is posed as  $\partial f / \partial t = \Theta \Xi_f$ .
  - **Checks:** dimension/shape consistency, notation consistency
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:**  $\Theta$  columns are candidate terms evaluated at all space-time samples;  $\partial f / \partial t$  is the matching target vector.
  - **Notes:** Given  $\Theta \in \mathbb{R}^{(8 \cdot 128^3) \times 66}$  and  $\Xi_f \in \mathbb{R}^{66}$ , the product has correct shape to match the target vector.
7. △ **Effect of column normalization on coefficient meaning** (Sec. 2.3 (normalization statement), p.3; used in Sec. 3.2, pp.5–6)
- **Claim:** Normalizing each  $\Theta$  column to unit  $L^2$  norm prevents domination by large-magnitude terms; reported coefficients can be interpreted as PDE coefficients.
  - **Checks:** scaling analysis, definition-consistency
  - **Verdict:** UNCERTAIN; confidence: high; impact: critical
  - **Assumptions/inputs:** Each feature column  $\theta_j$  is scaled by its  $L^2$  norm before regression., No explicit reverse scaling is shown when presenting Eqs. (4)–(5).

- **Notes:** Normalization changes coefficient magnitudes: coefficients correspond to scaled features unless explicitly unscaled. The paper does not state whether Eqs. (4)–(5) have been unnormalized, so physical interpretation of their magnitudes (viscosity, time scaling, etc.) cannot be verified.

8. ✓ **Discovered  $v_x$  PDE structure** (Eq. (4), Sec. 3.2.1, p.5)

- **Claim:** The learned  $v_x$  equation includes advection, a density-gradient term, a Laplacian term, and several second-derivative terms:  $\partial_t v_x \approx -5.17(v \cdot \nabla)v_x - 3.34\rho_x + 0.66\nabla^2 v_x + 5.87v_{x,xx} + 4.90v_{y,xy} + 4.49v_{z,xz}$ .
- **Checks:** notation consistency, operator sanity
- **Verdict:** PASS; confidence: medium; impact: moderate
- **Assumptions/inputs:** All differential operators are as defined in Sec. 2.3., Coefficients are those after the described filtering/grouping.
- **Notes:** As a linear combination of candidate operators, Eq. (4) is syntactically consistent. However, coefficient magnitudes/physical interpretations depend on whether unnormalization was performed (see separate item).

9. ✓ **Gradient-of-divergence identity (x-component)** (Sec. 3.2.1 (Compressibility/Bulk Viscosity discussion), pp.5–6)

- **Claim:** The terms  $v_{x,xx}$ ,  $v_{y,xy}$ ,  $v_{z,xz}$  are components appearing in  $\partial_x(\nabla \cdot v)$ , i.e., the  $x$ -component of  $\nabla(\nabla \cdot v)$ .
- **Checks:** algebra, vector calculus identity
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** Sufficient smoothness and commutation of mixed partials:  $\partial_x \partial_y v_y = \partial_y \partial_x v_y$ .
- **Notes:** Operator identity is correct:  $[\nabla(\nabla \cdot v)]_x = v_x$ . Exact identification as a single bulk-viscosity term would require a common coefficient, which is not present in Eq. (4).  $+ v_{y,xy} + v_{z,xz}$

10. ✗ **Time-step inference from advection coefficient** (Sec. 3.2.1 (final paragraph), p.6)

- **Claim:** Because the true advection coefficient is  $-1$  and the learned coefficient is about  $-5.5$ , the effective time step in the data is  $\Delta t \approx 1/5.5 \approx 0.18$ .
- **Checks:** scaling analysis, algebra
- **Verdict:** FAIL; confidence: high; impact: critical
- **Assumptions/inputs:** Eq. (1) is used with  $\Delta t_{\text{assumed}} = 1$  for derivative estimation., The advection operator in the true PDE has coefficient  $-1$  in the same scaling as the discovered operator.
- **Notes:** With Eq. (1) computed using  $\Delta t_{\text{assumed}} = 1$  while the true spacing is  $\Delta t_{\text{true}}$ , the estimated derivative vector satisfies  $(\partial_t f)_{\text{est}} \approx \Delta t_{\text{true}} (\partial_t f)_{\text{true}}$ , implying learned coefficients scale like  $\Delta t_{\text{true}} \approx 5.5$

(upto any additional feature/target scaling), not 0.18

.Moreover, prior column normalization of  $\Theta$  can arbitrarily rescale coefficients, making this inference invalid unless explicitly undone.} (not  $1/\Delta t_{\text{true}}$ ). Thus coefficient  $\approx -5.5$  implies  $\Delta t_{\text{true}}$

11. ✓ **Continuity-equation split into two terms** (Eq. (5) and Sec. 3.2.2, p.6)

- **Claim:** The terms  $(\rho \nabla \cdot v)$  and  $(v \cdot \nabla \rho)$  correspond to  $\nabla \cdot (\rho v)$  via  $\nabla \cdot (\rho v) = \rho \nabla \cdot v + v \cdot \nabla \rho$ , so  $\partial t \rho = -\nabla \cdot (\rho v)$  matches negative versions of those two terms.
- **Checks:** algebra, vector calculus identity
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** Standard product rule for divergence of a scalar times a vector field.
- **Notes:** Identity is correct; therefore the first two terms in Eq. (5) are structurally consistent with the continuity equation (up to coefficient scaling).

12. ✓ **Anti-diffusion interpretation of negative Laplacian** (Sec. 3.2.2 (discussion after Eq. (5)), p.6)

- **Claim:** A negative coefficient on  $\nabla^2 \rho$  corresponds to anti-diffusion (unstable growth of gradients).
- **Checks:** sign convention sanity
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** Standard diffusion sign convention:  $\partial t \rho = \kappa \nabla^2 \rho$  with  $\kappa > 0$  being smoothing.
- **Notes:** Under the usual convention,  $\kappa < 0$  yields backward diffusion and is ill-posed; the paper's interpretation is consistent.

13. ✓ **Collinear-term grouping example ( $\rho v_z$  vs  $v_z$ )** (Sec. 3.2.1 (discussion of collinearity), p.5)

- **Claim:** If  $\rho \approx 1$ , then a combination like  $56.89 \rho v_z - 56.35 v_z$  reduces to  $(56.89 \cdot 1 - 56.35) v_z \approx 0.54 v_z$ , indicating near-cancellation.
- **Checks:** algebra, sanity check
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:**  $\rho$  is near 1 over the dataset as stated in Sec. 3.1.
- **Notes:** Algebra is correct; the broader statistical conclusion (collinearity) is plausible though not provable from the text alone.

## Limitations

- Audit is based only on the provided parsed text and embedded page images; equations or definitions that might exist only in the original PDF (but not captured here) could not be checked.

- Figures (e.g., Fig. 6 coefficient bars) are described qualitatively; exact term lists and scaling operations used to produce the reported simplified equations are not fully specified, limiting verification of coefficient transformations.
- No full derivations are provided for how the ‘filtered’ equations (e.g., Eq. (4)) were constructed from the raw regression output; checks are limited to operator identities and stated formulas.

## Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

All executed internal consistency checks (grid sizing, spatial resolution, matrix dimensions, time-slice counts, basic coefficient arithmetic, reported ranges/averages, and simple derived quantities) passed within the stated tolerances.

### Checked items

1. ✓ **C1\_grid\_points\_1283** (p.2 §2.1 and repeated elsewhere (e.g., p.1 Abstract; p.4 §3.1): “1283 periodic grid”)
  - **Claim:** Grid is described as “1283 periodic grid” (interpretable as  $128^3$ ). Verify total number of spatial points used in later matrix dimension statements.
  - **Checks:** `power_and_dimension_check`
  - **Verdict:** PASS
  - **Notes:** Computed  $\text{grid\_points} = 128^3 = 2,097,152$ .
2. ✓ **C2\_spatial\_resolution\_dx** (p.2 §2.1: “ $L = 1 \dots \Delta x = \Delta y = \Delta z = L/128$ ”)
  - **Claim:** Given  $L = 1$  and 128 cells per axis, spatial resolution should be  $1/128$ .
  - **Checks:** `unit_arithmetic_check`
  - **Verdict:** PASS
  - **Notes:** Computed  $\Delta x = 1/128 = 0.0078125$ .
3. ✓ **C3\_theta\_matrix\_rows** (p.3 §2.3: “ $\Theta$  matrix had dimensions  $(8 \times 128^3, 66)$ ”)
  - **Claim:** Given 8 time slices ( $t_1$ – $t_8$ ) and  $128^3$  spatial points, verify the number of rows equals  $8 \times 128^3$ .
  - **Checks:** `dimension_multiplication_check`
  - **Verdict:** PASS
  - **Notes:** Checked rows =  $8 \times (128^3) = 16,777,216$  and columns = 66.
4. ✓ **C4\_discarded\_time\_slices\_count** (p.2 §2.2: “10 time slices... first ( $t_0$ ) and last ( $t_9$ ) ... discarded... focus on 8 ... ( $t_1$  to  $t_8$ )”)

- **Claim:** From 10 total time slices discarding 2 should leave 8.
  - **Checks:** count\_consistency\_check
  - **Verdict:** PASS
  - **Notes:** Computed  $10 - 2 = 8$ .
5. ✓ **C5\_collinear\_pair\_net\_contribution** (p.5 §3.2.1: “ $56.89\rho v_z$  and  $-56.35v_z$ ... net...  $56.89(1) - 56.35 \approx 0.54$ ”)
- **Claim:** Verify the subtraction and the stated  $\approx 0.54$  net coefficient when assuming  $\rho \approx 1$ .
  - **Checks:** arithmetic\_difference\_check
  - **Verdict:** PASS
  - **Notes:** Computed  $56.89 \times 1 + (-56.35) = 0.54$  (floating-point residual only).
6. ✓ **C6\_R2\_range\_consistency** (p.5 §3.2.1 ( $R^2$  values 0.658, 0.709, 0.566) and p.1 Abstract / p.8 Conclusions (“ $R^2$  values 0.57-0.71”))
- **Claim:** Check that the stated  $R^2$  range 0.57–0.71 is consistent with the listed component  $R^2$  values.
  - **Checks:** min\_max\_range\_check
  - **Verdict:** PASS
  - **Notes:** min=0.566 rounds to 0.57; max=0.709 rounds to 0.71.
7. ✓ **C7\_advection\_average\_coefficient** (p.5 §3.2.1 bullet list: advection coefficients  $-5.17, -5.77, -5.54$ ; p.6: “averaging approximately  $-5.5$ ”)
- **Claim:** Verify the mean of the three advection coefficients is approximately  $-5.5$ .
  - **Checks:** average\_check
  - **Verdict:** PASS
  - **Notes:** Mean =  $-5.493333\dots$ , consistent with “approximately  $-5.5$ ”.
8. ✓ **C8\_time\_step\_inference\_1\_over\_5p5** (p.6 §3.2.1: “ $\Delta t$  ... approximately  $1/5.5 \approx 0.18$ ”)
- **Claim:** Verify  $1/5.5$  equals about 0.18.
  - **Checks:** reciprocal\_check
  - **Verdict:** PASS
  - **Notes:**  $1/5.5 = 0.181818\dots$ , rounds to 0.18.
9. ✓ **C9\_density\_contour\_range\_width** (p.4 §3.1 and Fig.4 caption p.6: “density contours, ranging from approximately 0.986 to 1.006”)
- **Claim:** Compute the stated density contour range width and its deviation from mean 1.0 for basic plausibility summaries.

- **Checks:** range\_width\_and\_offset\_check
  - **Verdict:** PASS
  - **Notes:** Width = 0.020; max–mean = 0.006; mean–min = 0.014 (floating-point residual only).
10. ✓ **C10\_sim\_density\_mean\_difference** (p.7 §3.3.1: “simulated density mean evolved from 1.000000 at  $t = 1$  to 1.002234 at  $t = 8$  ... ground truth mean 0.999999 at  $t = 8$ ”)
- **Claim:** Verify the absolute and relative differences between simulated and ground truth mean density at  $t = 8$  using the stated numbers.
  - **Checks:** difference\_and\_relative\_error\_check
  - **Verdict:** PASS
  - **Notes:** Computed  $\text{sim8-gt8} = 0.002235$  and  $\text{sim8-sim1} = 0.002234$ ; relative error = 0.223500...%.
11. ✓ **C11\_RMSE\_ordering\_and\_values** (p.7 §3.3.2 and Fig.8 caption p.8: “reaching 0.521 for  $v_x$ , 0.347 for  $v_y$ , and 0.265 for  $v_z$  by ...  $t = 8$ ”)
- **Claim:** Check simple ordering and ratios among the stated RMSE values at final time index.
  - **Checks:** ordering\_and\_ratio\_check
  - **Verdict:** PASS
  - **Notes:** Ordering holds ( $v_x > v_y > v_z$ ); ratios:  $v_x/v_y = 1.501440922\dots$ ,  $v_x/v_z = 1.966037736\dots$

### Limitations

- Checks are limited to internal arithmetic/logic using numbers explicitly stated in the PDF text; no access to the underlying NumPy dataset or code outputs.
- No numeric extraction from figures/plots (e.g., bar heights, curve points) is performed; only numbers explicitly written in the text/captions are used.
- Some statements (e.g., the 66-term library composition, physical ‘true coefficient’ claims) lack sufficient explicit detail for deterministic recomputation.