Skepthical review: *Data-Driven Discovery of Fluid Dynamics Equations from Spatial-Temporal Data*

Summary

This manuscript presents an end-to-end sparse-regression workflow to discover governing PDEs from a 3D spatio-temporal fluid dataset (10 timesteps on a 128^3 periodic box) containing density and three velocity components. Spatial and temporal derivatives are computed via second-order central finite differences with periodic boundary conditions (Sec. 2.2), and a hand-crafted 43-term feature library (Sec. 2.3) is used with cross-validated LASSO (Sec. 2.4–2.5) to fit $\partial_t \rho$ and $\partial_t v_x, \partial_t v_y, \partial_t v_z$ from randomly sampled space-time points (Sec. 3.2–3.3). The recovered dominant terms are interpreted as consistent with the continuity equation for density and Navier–Stokes-like momentum equations for velocity (Sec. 3.3.1–3.3.2), supported by extensive diagnostic plots (Sec. 3.4–3.5). The velocity-component fits achieve moderate test R^2 (~ 0.58–0.73), whereas the density equation fit is weak ($R^2 \approx 0.16$) and attributed to near-incompressibility and derivative noise (Sec. 3.1, Sec. 3.3.1). Overall, the paper is a clear case study demonstrating that sparse linear regression can yield physically interpretable PDE-like structures from high-dimensional flow data, but its current claims are limited by missing ground-truth simulation specification, standardized-only coefficients, potentially leaky validation splits, lack of forward-integration tests, and insufficient robustness/stability analyses with respect to derivative estimation, library design, and noise.

Strengths

- Clear, coherent end-to-end pipeline from raw 3D+time fields to derivative estimation, feature construction, sparse regression, and multiple validation diagnostics (Sec. 2, Sec. 3).
- Use of a realistic 3D periodic-box dataset and standard second-order central differences consistent with the stated periodic boundary conditions (Sec. 2.1–2.2).
- Feature library includes many physically meaningful operators (convective terms, divergence, Laplacian, curl, density-weighted operators), enabling Navier–Stokes-like structures to emerge (Sec. 2.3, Sec. 3.3.2).
- Cross-validated LASSO and explicit train/test reporting are good practices and help guard against overtuning to a single regularization strength (Sec. 2.4–2.5, Sec. 3.3).
- The Results section provides rich qualitative diagnostics (scatter/residual views, spatial maps, temporal comparisons) that help interpret what the regression is learning (Sec. 3.4–3.5).

Major issues

1. **The data-generating physics and solver are not specified, weakening the central claim of "recovering" continuity/Navier–Stokes and making the interpretation of recovered terms (especially pressure/viscosity-related ones) largely speculative (Sec. 2.1, Sec. 3.1, Sec. 3.3.1–3.3.2, Sec. 4).** Key missing items include: whether the simulation is incompressible vs (weakly) compressible, whether pressure is solved as a constraint, the equation of state (if any), forcing, viscosity, nondimensional parameters (Re/Mach), and basic flow diagnostics.

   Recommendation: Expand Sec. 2.1 into a dedicated "Dataset / Simulation" subsection that writes down the governing PDEs used to generate the data (including the pressure treatment and EOS, if applicable), boundary conditions, discretization scheme, forcing, viscosity, and key nondimensional parameters (e.g., Re and Mach). Add basic diagnostics (e.g., typical $\nabla \cdot \mathbf{v}$, density fluctuations, spectra) in Sec. 3.1 to support statements like "nearly incompressible" or "turbulent-like." If the underlying PDE/metadata are unavailable, explicitly state that in Sec. 2.1 and soften claims in Sec. 4 to "structurally consistent with Navier–Stokes-like forms."

2. **Coefficients are reported only in standardized feature space, preventing quantitative validation against physical parameters (e.g., viscosity ν, barotropic c_s^2) and making coefficient-based physical interpretation incomplete (Sec. 2.4, Sec. 3.3.1–3.3.2).**

   Recommendation: In Sec. 2.4, state precisely how standardization is done (mean/std per feature and per target, fit on training only). Derive and report the mapping back to unstandardized coefficients. In Sec. 3.3.1–3.3.2, provide tables for each discovered PDE listing (i) nonzero standardized coefficients and (ii) de-standardized coefficients in the original variable scaling. Where ground-truth parameters exist (from the simulation spec added to Sec. 2.1), compare inferred ν and any pressure/EOS-related coefficient quantitatively.

3. **Validation is primarily pointwise prediction of finite-difference time derivatives ($\partial_t \phi$), which is necessary but not sufficient to demonstrate that the discovered PDE governs the dynamics. Without forward-integration / rollout tests, models may be good local regressors but poor dynamical laws (Sec. 2.5, Sec. 3.4–3.5, Sec. 4).**

   Recommendation: Add forward-integration experiments (e.g., new subsection in Sec. 3.5): integrate the discovered PDEs from an initial condition and compare predicted fields (and/or spectra, kinetic energy, enstrophy, mass conservation) to held-out snapshots over multiple timesteps. If full 3D integration is too expensive, do short-horizon rollouts, reduced-resolution rollouts, or 2D-slice/patch tests, and discuss limitations explicitly in Sec. 4.

4. **Potential train/test leakage due to random sampling of correlated space-time points: randomly splitting flattened gridpoints can place near-identical spatial neighbors and adjacent times in both train and test, inflating generalization metrics (Sec. 2.5, Sec. 3.2–3.4).**

   Recommendation: Replace or supplement the random pointwise split with more stringent blocked splits: (i) hold out entire timesteps (e.g., train on early times, test on later times), and/or (ii) hold out contiguous spatial blocks/subvolumes. Report performance (R^2/MSE) under these splits in Sec. 3.3–3.4, and discuss how correlation structure affects evaluation in Sec. 4.

5. **Internal physical interpretation is currently inconsistent regarding "incompressible Navier–Stokes" vs. a learned density-gradient term acting as a pressure-gradient surrogate. In strict incompressibility, ρ is constant (so $\nabla\rho = 0$) and pressure is not generally a function of ρ; thus $-\nabla\rho$ cannot generally represent $-\rho^{-1}\nabla P$ without additional assumptions (Sec. 3.3.2, p. 11).**

   Recommendation: Clarify the regime and assumptions in Sec. 3.3.2 and align terminology throughout: either (a) describe the system as weakly compressible/barotropic and state the EOS (e.g., $P = c_s^2\rho$), then validate the surrogate by comparing $\nabla\rho$ with ∇P if pressure is available; or (b) if the simulation is incompressible, remove/limit the $-\nabla\rho$ $\leftrightarrow$pressure interpretation and discuss what $-\nabla\rho$ is capturing (e.g., a correlating proxy) and how pressure would be treated if unobserved.

6. **The learned velocity models retain ~ 30 nonzero terms out of 43, which is not strongly sparse/parsimonious and raises concerns about multicollinearity, feature redundancy, and support stability; yet the discussion focuses on only a few "dominant" terms (Sec. 3.3.2).**

   Recommendation: In Sec. 3.3.2, report the full set of selected terms per equation (with indices tied to an explicit feature list) and provide a sorted coefficient-magnitude plot. Quantify the performance–sparsity tradeoff by refitting with top-k terms ($k = 3, 5, 10, \ldots$) and reporting test R^2/MSE. Consider adding a second-stage sparsification (e.g., sequential thresholded least squares refit, hard-threshold + OLS refit, or stability selection) and report whether the support is stable across random seeds/subsamples.

7. **Derivative estimation and noise sensitivity are not analyzed quantitatively. With only 10 timesteps, central differences yield only 8 effective temporal derivative slices and can be fragile; spatial derivatives in turbulent-like fields amplify noise, affecting both identifiability and the low density-equation R^2 (Sec. 2.2, Sec. 3.1–3.3).**

Recommendation: Add a derivative-sensitivity study (new Sec. 3.6 or similar): compare the current finite-difference derivatives to alternatives appropriate for periodic domains (e.g., FFT/spectral derivatives for space; smoothing/regularized differentiation for time), and quantify how learned terms and coefficients change. Report an estimated SNR for $\partial_t \rho$ and $\partial_t \mathbf{v}$, and discuss derivative error propagation in Sec. 4.

8. **The low R^2 for the density equation is treated mostly qualitatively, and there are no controls to demonstrate that the apparent continuity-like structure is not recoverable from noise or correlations alone (Sec. 3.3.1, Sec. 4).**

   Recommendation: In Sec. 3.3.1, add control experiments: rerun the pipeline with permuted/shuffled $\partial_t \rho$ (or phase-randomized density fields) and compare R^2 and selected-term patterns to the true-data case. Report bootstrap variability of selected density terms and coefficients. Temper claims in Sec. 4 accordingly if the density equation is not robustly identifiable.

9. **The feature library design is heuristic and not systematically justified relative to the target physics; the impact of library choice on the recovered equations is not assessed (Sec. 2.3, Sec. 3.3).**

   Recommendation: In Sec. 2.3, provide an explicit indexed table of all 43 features (exact formulas and sign conventions), and briefly justify inclusion/exclusion relative to the presumed governing equations (from Sec. 2.1). In Sec. 3, add an ablation study removing feature groups (e.g., curl terms, density-weighted operators, raw fields) and report changes in selected dominant terms, coefficient stability, and performance.

10. **Positioning versus prior PDE-discovery literature is underdeveloped, and there is no baseline comparison to established approaches (e.g., PDE-FIND/SINDy variants, STLSQ). This makes novelty and performance hard to evaluate (Sec. 1, Sec. 4).**

    Recommendation: Add a Related Work subsection (Sec. 1.1 or end of Sec. 1) covering PDE-FIND/SINDy (Rudy et al.; Brunton et al.) and relevant modern variants. If feasible, implement at least one baseline (e.g., STLSQ/PDE-FIND with the same library) and compare term selection and performance in Sec. 3. If not feasible, state this clearly in Sec. 4 and frame the contribution as a detailed 3D case study rather than a new identification algorithm.

Minor issues

1. Reproducibility details are incomplete for preprocessing, sampling, and fitting (Sec. 2.1, Sec. 2.4–2.5): random seeds for the $200,!000$-point subsampling, whether scaling is fit on train only, the LassoCV alpha grid and solver tolerances/iterations, and software versions.

Recommendation: In Sec. 2.1 and Sec. 2.4–2.5, list random seeds (or state seed-averaged results), confirm scaler fit is train-only, document LassoCV hyperparameter grid (or defaults), max_iter/tol, and provide library versions. Consider reporting variability across several seeds/splits.

2. The dataset description lacks clarity on units/nondimensionalization and the precise mapping from index time to physical time (Δt), which also affects coefficient interpretation and forward integration (Sec. 2.1–2.2).

   Recommendation: Expand Sec. 2.1 to state whether variables are nondimensional (and how), give Δt explicitly, and specify whether fields are cell-centered vs node-centered (relevant for derivative consistency).

3. The paper would benefit from a compact numerical summary of results across all four targets (ρ, v_x, v_y, v_z), including #nonzeros, train/test MSE and R^2, and (ideally) blocked-split metrics (Sec. 3.3–3.4).

   Recommendation: Add a table in Sec. 3.3 or Sec. 3.4 summarizing metrics and sparsity for each equation, including alternative split results (time-block/spatial-block) if added.

4. Several figures are difficult to compare due to inconsistent labeling/normalization and limited quantitative annotation; overplotting in scatter plots and unclear panel mappings in term-decomposition plots reduce interpretability (Figures 1, 3, 5–7, 8–12, 13–15, 17–20; Sec. 3.4–3.5).

   Recommendation: Standardize axis labels/units (or explicitly state normalization), add R^2/RMSE and sample size to key plots, use hexbin/density scatter to reduce overplotting, and ensure decomposition maps have consistent colorbars and explicit legends mapping panels to mathematical terms and sign conventions.

5. The library is described conceptually but not enumerated, and later text refers to "dominant" vs. "secondary" terms without a precise selection rule or threshold (Sec. 2.3, Sec. 3.3).

   Recommendation: Provide the explicit indexed feature list (table/appendix) and define "dominant" quantitatively (e.g., top-k by absolute de-standardized coefficient, or by variance contribution). Include a ranked list of selected terms for each equation.

6. Given near-incompressibility, the choice to model $\partial_t \rho$ directly may be intrinsically ill-conditioned; alternative formulations (e.g., enforcing mass conservation form $\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0$) are not explored (Sec. 3.3.1).

   Recommendation: Consider adding an experiment fitting the conservative continuity form directly (regress $\partial_t \rho$ against $-\nabla \cdot (\rho \mathbf{v})$ and related terms) and compare stability/performance; at minimum discuss identifiability challenges for nearly constant ρ in Sec. 4.

Very minor issues

1. Notation and ordering of variables/components appears to vary (e.g., dataset channel ordering vs. later ρ, v_x, v_y, v_z text), which can confuse readers (Sec. 2.1, various figure captions).

   Recommendation: Standardize channel ordering and symbols everywhere; add a single table listing dataset tensor shape and channel-to-variable mapping in Sec. 2.1.

2. Finite-difference indexing is presented in 1D shorthand despite operating on a 3D grid, leaving mild ambiguity about axis-specific indexing and periodic wrap implementation (Sec. 2.2).

   Recommendation: Add explicit 3D index notation (e.g., $\phi_{i\pm1,j,k}$ for $\partial_x \phi$) and briefly state how wrapping is implemented.

3. The manuscript sometimes uses typographically ambiguous "1283" instead of 128^3 (Abstract and elsewhere).

   Recommendation: Use 128^3 or $128 \times 128 \times 128$ consistently.

4. Some captions are repetitive and do not clearly state what is unique about each figure beyond "agreement," and some plot titles/labels are verbose or inconsistent (Sec. 3.4–3.5).

   Recommendation: Edit captions to emphasize the specific diagnostic point per figure (e.g., bias, heteroscedasticity, spatial structure of residuals) and standardize caption structure (what/where/train-vs-test/normalization/metrics).

5. Minor style/format consistency issues (hyphenation/capitalization of Navier–Stokes, spacing around inline math, heading consistency) appear throughout (Sec. 1–4).

   Recommendation: Run a consistency pass to unify typography (Navier–Stokes), math spacing, and heading styles.

Mathematical consistency audit

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

Maths relevance: substantial

The paper's analytic content centers on (i) defining finite-difference approximations for spatial/temporal derivatives on a periodic 3D grid, (ii) defining vector-calculus operators and nonlinear convective terms used to build a candidate feature library, and (iii) relating sparse-regression-selected terms to the continuity equation and a Navier–Stokes-type momentum equation. Most equations are standard identities/definitions and are internally consistent; the main internal tension is the simultaneous use of "incompressible Navier–Stokes" language with a density-gradient proxy for pressure, which requires extra assumptions not fully formalized.

Checked items

1. ✔ **Grid spacing from box length** (Sec. 2.1, p. 3)
    - **Claim:** With periodic box length $L = 1$ and $N = 128$ points per dimension, $\Delta x = \Delta y = \Delta z = L/N = 1/128$.
    - **Checks:** algebra, definition consistency
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** Uniform grid with N points spanning length L in each dimension
    - **Notes:** $\Delta x = 1/128$ follows directly from the stated L and N.
2. ✔ **First spatial derivative central difference** (Sec. 2.2, p. 3)
    - **Claim:** Approximate $\partial\phi/\partial x$ at index i by $(\phi(i+1) - \phi(i-1))/(2\Delta x)$.
    - **Checks:** algebra, order-of-accuracy plausibility, notation consistency
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** Second-order central difference stencil, Uniform spacing Δx, Periodic boundary conditions handle $i \pm 1$ at edges
    - **Notes:** Standard second-order central difference; periodicity assumption resolves boundary indexing.
3. ✔ **Second spatial derivative central difference** (Sec. 2.2, p. 4)
    - **Claim:** Approximate $\partial^2\phi/\partial x^2$ at index i by $(\phi(i+1) - 2\phi(i) + \phi(i-1))/\Delta x^2$.
    - **Checks:** algebra, order-of-accuracy plausibility
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** Second-order central difference stencil, Uniform spacing Δx
    - **Notes:** Standard 3-point stencil for second derivative.
4. ✔ **Laplacian operator assembly** (Sec. 2.2, p. 4)
    - **Claim:** Compute $\nabla^2\phi$ as $\partial^2\phi/\partial x^2 + \partial^2\phi/\partial y^2 + \partial^2\phi/\partial z^2$.
    - **Checks:** definition consistency
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** Cartesian coordinates
    - **Notes:** Correct definition in Cartesian coordinates.
5. ✔ **Velocity divergence definition** (Sec. 2.2, p. 4)
    - **Claim:** Define $\nabla \cdot v = \partial v_x/\partial x + \partial v_y/\partial y + \partial v_z/\partial z$.
    - **Checks:** definition consistency, notation consistency
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** $v = (v_x, v_y, v_z)$ in Cartesian coordinates

- **Notes:** Standard divergence definition.

6. ✔ **Curl component definitions** (Sec. 2.2, p. 4)

   - **Claim:** Define curl components: $(\nabla \times v)_x = \partial v_z/\partial y - \partial v_y/\partial z$; $(\nabla \times v)_y = \partial v_x/\partial z - \partial v_z/\partial x$; $(\nabla \times v)_z = \partial v_y/\partial x - \partial v_x/\partial y$.
   - **Checks:** definition consistency, sign consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Right-handed Cartesian coordinates
   - **Notes:** All three components match the standard right-handed curl.

7. ✔ **Temporal derivative central difference** (Sec. 2.2, p. 4)

   - **Claim:** Approximate $\partial\phi/\partial t(t)$ by $(\phi(t+\Delta t) - \phi(t-\Delta t))/(2\Delta t)$ with $\Delta t = 1$.
   - **Checks:** algebra, notation consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Uniform time step, Second-order central difference in time
   - **Notes:** Standard second-order central difference; $\Delta t = 1$ is consistent with the formula.

8. ✔ **Valid temporal indices for central differencing** (Sec. 2.1 and 2.2, pp. 3–4)

   - **Claim:** With 10 time slices, central differencing allows temporal derivatives for $t = 1$ through $t = 8$ inclusive.
   - **Checks:** indexing consistency
   - **Verdict:** PASS; confidence: medium; impact: minor
   - **Assumptions/inputs:** Time indices run over $0..9$ (or equivalent ordering with 10 slices), Central difference needs $t-1$ and $t+1$ available
   - **Notes:** If slices are indexed $0..9$, then $t = 1..8$ are the interior points. The paper does not explicitly state whether indexing starts at 0, but the stated interior range is consistent with having 10 slices.

9. ✔ **Convective term definitions** (Sec. 2.3, p. 5)

   - **Claim:** Define $v \cdot \nabla\rho = v_x\partial\rho/\partial x + v_y\partial\rho/\partial y + v_z\partial\rho/\partial z$ and $(v \cdot \nabla)v_x$ etc. analogously for velocity components.
   - **Checks:** definition consistency, notation consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** $v = (v_x, v_y, v_z)$, ∇ acts on the scalar field that follows
   - **Notes:** Component expansions are correct for the advective derivative operator $v \cdot \nabla$ acting on scalars.

10. ✔ **Candidate feature library size (43 terms)** (Sec. 2.3, pp. 4–5)

- **Claim:** The paper constructs a library of 43 candidate terms.
- **Checks:** combinatorial consistency, definition consistency
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** Count each explicitly listed category exactly once
- **Notes:** Tally from the listed categories: originals 4; pairwise products listed 10; first derivatives 4 variables $\times 3$ dirs$= 12$; convective terms 4; Laplacians 4; density-weighted Laplacians 4; divergence 1; density-weighted divergence 1; curl components 3. Total $4 + 10 + 12 + 4 + 4 + 4 + 1 + 1 + 3 = 43$.

11. ✔ **Continuity equation product-rule expansion** (Sec. 3.3.1, p. 10)

    - **Claim:** $\partial\rho/\partial t = -\nabla \cdot (\rho v) = -\rho(\nabla \cdot v) - v \cdot \nabla\rho$.
    - **Checks:** algebra, vector-calculus identity
    - **Verdict:** PASS; confidence: high; impact: critical
    - **Assumptions/inputs:** Standard vector calculus identity: $\nabla \cdot (\rho v) = \rho\nabla \cdot v + v \cdot \nabla\rho$
    - **Notes:** Correct application of the product rule for divergence of a scalar-times-vector field.

12. ✔ **Near-uniform density approximation in continuity term** (Sec. 3.3.1, p. 10)

    - **Claim:** If $\rho \approx 1.0$, then $\rho(\nabla \cdot v)$ is numerically indistinguishable from $\nabla \cdot v$.
    - **Checks:** limiting/sanity case, approximation consistency
    - **Verdict:** PASS; confidence: medium; impact: minor
    - **Assumptions/inputs:** ρ spatially close to 1 with small relative variation
    - **Notes:** As an approximation, multiplying by a nearly constant $\rho \approx 1$ changes the term only by small relative perturbations. (This is an approximation, not an identity.)

13. ✔ **Navier–Stokes momentum equation form** (Sec. 3.3.2, p. 11)

    - **Claim:** Write momentum as $\partial v/\partial t = -(v \cdot \nabla)v - (1/\rho)\nabla P + \nu\nabla^2 v$.
    - **Checks:** algebraic equivalence, dimensional consistency (symbolic)
    - **Verdict:** PASS; confidence: high; impact: critical
    - **Assumptions/inputs:** Standard rearrangement of convective term to RHS
    - **Notes:** Algebraically equivalent to $\partial v/\partial t + (v \cdot \nabla)v = -(1/\rho)\nabla P + \nu\nabla^2 v$. Term dimensions can be consistent under standard fluid-dynamics scaling.

14. ⚠ **Pressure-gradient surrogate via density gradient** (Sec. 3.3.2, p. 11)

    - **Claim:** Interpret discovered $-\nabla\rho$ as representing $-\rho^{-1}\nabla P$, motivated by a barotropic relation $P \propto \rho$ and $\rho \approx 1$ so $-\rho^{-1}\nabla P \approx -c_s^2\nabla\rho$.
    - **Checks:** assumption sufficiency, symbolic consistency, notation/definition consistency
    - **Verdict:** UNCERTAIN; confidence: medium; impact: critical

- **Assumptions/inputs:** Barotropic equation of state $P = P(\rho)$, effectively linearized as $P \approx c_s^2 \rho$, ρ approximately constant so $1/\rho \approx 1$
- **Notes:** The algebraic step $-\rho^{-1}\nabla P \to -c_s^2 \nabla \rho$ is consistent only if P is explicitly taken as a function of ρ (and approximately linear) and if incompressibility is not enforced strictly (since strict incompressibility implies $\nabla \rho = 0$). The paper gestures at barotropic behavior but simultaneously frames results as "incompressible Navier–Stokes," leaving the modeling assumptions internally ambiguous.

15. ✔ **Sign of viscous diffusion term in momentum** (Sec. 3.3.2, pp. 11–12)
    - **Claim:** Viscous diffusion appears as $+\nu\nabla^2 v$ (positive coefficient multiplying Laplacian).
    - **Checks:** sign convention consistency, dimensional consistency (symbolic)
    - **Verdict:** PASS; confidence: high; impact: moderate
    - **Assumptions/inputs:** $\nu > 0$, Laplacian defined as sum of second derivatives
    - **Notes:** With $\nu > 0$, $+\nu\nabla^2 v$ is the standard diffusion form that damps high spatial frequencies; this is consistent with the text's qualitative interpretation.

16. ✔ **Effect of standardization on coefficient sign and interpretability** (Sec. 2.1, p. 3 and Sec. 3.3, pp. 10–12)
    - **Claim:** Using standardized features enables comparison of coefficient magnitudes; sign interpretations remain meaningful.
    - **Checks:** transformation consistency
    - **Verdict:** PASS; confidence: medium; impact: minor
    - **Assumptions/inputs:** Standardization uses $x_{\text{std}} = (x - \mathtt{mean})/\mathtt{std}$ with $\mathtt{std} > 0$
    - **Notes:** Dividing by a positive standard deviation preserves coefficient sign relative to the standardized feature. However, mapping coefficients back to physical units requires de-standardization, which is not provided.

Limitations

- The provided content is extracted text (and some figure images) without equation numbering; references to "Eq. (·)" cannot be made, and some mathematical details may be lost in extraction.
- The paper does not present the explicit regression equation in matrix form (e.g., $y = \Theta\xi$), nor does it specify intercept handling in the formalism; verification of the exact regression setup is therefore limited to the verbal description.
- Any assessment of whether the learned coefficients match expected physical constants is out of scope (numeric/empirical) and also impeded by the stated standardization step without a de-standardization mapping.

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 recomputations passed: dataset shape and grid spacing are arithmetically consistent; central-difference time indexing yields the stated 8 valid slices; flattened sample counts comfortably exceed the $200,!000$ subsample; the $80/20$ split yields exact integer counts; enumerated feature-library term counts sum to 43; 5-fold CV on the training set is numerically feasible; abstract R^2 range is consistent with component R^2s under a 0.005 bound-rounding tolerance; and reported train/test metrics examples show small gaps.

Checked items

1. ✔ **C1__dataset__shape__consistency** (Methods §2.1 (page 3))
   - **Claim:** Dataset stored in a NumPy array with dimensions $(10, 4, 128, 128, 128)$, corresponding to 10 time slices, 4 physical variables, and a $128 \times 128 \times 128$ spatial grid.
   - **Checks:** shape__product__consistency
   - **Verdict:** PASS
   - **Notes:** Checked shape tuple and spatial grid product.
2. ✔ **C2__grid__spacing__from__L__over__N** (Methods §2.1 (page 3))
   - **Claim:** Periodic box length $L = 1$ in each dimension, grid spacing $\Delta x = \Delta y = \Delta z = L/N = 1/128$.
   - **Checks:** unit__consistent__recomputation
   - **Verdict:** PASS
   - **Notes:** Compared $dx = L/N$ to reported dx.
3. ✔ **C3__valid__time__slices__count__for__central__difference** (Methods §2.1 and §2.2 (page 3-4))
   - **Claim:** Valid time slices for temporal derivatives are from $t = 1$ to $t = 8$ inclusive (due to central differencing) out of 10 total time slices.
   - **Checks:** index__range__count
   - **Verdict:** PASS
   - **Notes:** Assumes 0-based indexing $0..T - 1$ for central differencing.
4. ✔ **C4__flattened__sample__count__before__subsample** (Methods §2.1 (page 3))
   - **Claim:** Flattening all spatial dimensions $(128 \times 128 \times 128)$ and valid time slices ($t = 1..8$ inclusive) yields a large number of space-time points; then a random subsample of $200,!000$ points was extracted.
   - **Checks:** product__recomputation
   - **Verdict:** PASS

- **Notes:** Verified subsample size does not exceed available flattened points.

5. ✔ **C5_train_test_split_counts_from_200k** (Methods §2.1 (page 3))
    - **Claim:** Subsampled dataset of $200,!000$ points was partitioned into training (80%) and testing (20%).
    - **Checks:** percentage_to_count
    - **Verdict:** PASS
    - **Notes:** Computed counts using rounding; here percentages yield exact integers.

6. ✔ **C6_feature_library_count_from_enumerated_terms** (Methods §2.3 (pages 4-5))
    - **Claim:** A comprehensive library of 43 candidate mathematical terms was constructed; bullets enumerate included terms.
    - **Checks:** count_terms_from_description
    - **Verdict:** PASS
    - **Notes:** Summed enumerated term counts and compared to reported total.

7. ✔ **C7_cross_validation_folds_count** (Methods §2.4 (page 5))
    - **Claim:** Optimal regularization parameter was determined using 5-fold cross-validation.
    - **Checks:** basic_parameter_sanity
    - **Verdict:** PASS
    - **Notes:** Sanity check: each CV fold must be non-empty.

8. ✔ **C8_r2_range_vs_component_values** (Abstract (page 1) vs Results §3.3.2 (page 11))
    - **Claim:** Abstract: velocity-component models achieved R-squared scores ranging from 0.58 to 0.73. Results: test R^2 for $v_x = 0.685$, $v_y = 0.733$, $v_z = 0.588$.
    - **Checks:** range_consistency
    - **Verdict:** PASS
    - **Notes:** Checked whether component R^2s lie within claimed abstract range, allowing abs_tol on bounds.

9. ✔ **C9_density_r2_train_test_close** (Results §3.3.1 (page 11))
    - **Claim:** Density equation R^2: Train $R^2 \approx 0.161$, Test $R^2 \approx 0.166$.
    - **Checks:** train_test_metric_gap
    - **Verdict:** PASS
    - **Notes:** Checked whether train/test values are close under abs or relative tolerance.

10. ✔ **C10__vy__mse__train__test__close** (Results §3.4 (page 12))
    - **Claim:** v_y model achieved Train MSE of 2.74×10^{-5} and Test MSE of 2.76×10^{-5}.
    - **Checks:** train__test__metric__gap
    - **Verdict:** PASS
    - **Notes:** Checked whether train/test values are close under abs or relative tolerance.
11. ✔ **C11__density__std__relative__to__mean** (Results §3.1 (pages 6-7))
    - **Claim:** Density mean remains constant at 1.0; density standard deviation fluctuates between 0.0021 and 0.0022 (exceptionally small).
    - **Checks:** ratio__recomputation
    - **Verdict:** PASS
    - **Notes:** Derived relative standard deviation percentages (no explicit threshold to validate).
12. ✔ **C12__velocity__std__range__width** (Results §3.1 (page 7))
    - **Claim:** Velocity standard deviations are stable, ranging from 0.23 to 0.25.
    - **Checks:** range__width__recomputation
    - **Verdict:** PASS
    - **Notes:** Recomputed width and relative width; verified max$\geq$min.
13. ✔ **C13__continuity__equation__expansion__signs** (Results §3.3.1 (page 10))
    - **Claim:** Continuity equation written as $\partial\rho/\partial t = -\nabla \cdot (\rho v) = -\rho(\nabla \cdot v) - v \cdot \nabla\rho$; with $\rho \approx 1$, $\rho(\nabla \cdot v)$ numerically indistinguishable from $\nabla \cdot v$.
    - **Checks:** algebraic__identity__check
    - **Verdict:** PASS
    - **Notes:** Symbolic/vector-calculus identity check (numeric indistinguishability requires data; not evaluated).

Limitations

- Only parsed text from the PDF was available; numeric values embedded solely in plots/figures (without explicit text numbers) were not extracted.
- No access to the underlying NumPy dataset or model outputs, so claims about statistics over time, exact coefficient sparsity patterns, and validation plots cannot be recomputed.
- Several statements use approximate language (≈, 'on the order of', 'approximately'), so checks require tolerances and can only assess internal consistency, not empirical correctness.