

Skeptical review: Robust Spatio-Temporal Feature Generation for Incompressible Flow Equation Discovery

Summary

The paper presents a preprocessing and feature-engineering pipeline for data-driven discovery of governing equations in a decaying, nearly incompressible 3D flow, observed on a 128^3 periodic grid with only 10 time snapshots. The authors estimate temporal derivatives via second-order local polynomial regression at each grid point, and benchmark FFT-based spectral differentiation against a WENO5 finite-difference scheme for spatial derivatives, selecting the spectral method based on slightly better enforcement of incompressibility. They then estimate an effective global kinematic viscosity through linear regression of $\partial\mathbf{u}/\partial t$ on $\nabla^2\mathbf{u}$ and analyze the resulting momentum residuals to infer the importance of unmeasured pressure gradients and motivate additional proxy terms. All primary fields and derived features are independently standardized.

A 26-term feature library is constructed to support subsequent sparse-regression-based equation discovery, including advection, viscous, density and density-gradient terms, divergence gradients, kinetic-energy gradients, and velocity–density couplings such as components of $\rho'\nabla\mathbf{u}$. The pipeline is evaluated using RMSE of temporal fits in both smooth and sharp-gradient regions, incompressibility metrics for the spatial schemes, and qualitative and quantitative analysis of residuals against the incompressible Navier–Stokes momentum balance. The authors emphasize that the outcome is an “optimally conditioned” feature matrix and corresponding temporal derivative targets intended for later sparse regression, which is not carried out in this work.

The manuscript is clearly written and provides a coherent end-to-end narrative from dataset description and normalization to derivative estimation, residual analysis, and feature construction. However, several key choices (e.g., temporal fitting strategy, viscosity estimation, selection and interpretation of proxy features, and derivative-scheme benchmarking) are only partially justified, and the absence of any actual equation-discovery experiment limits the ability to assess the effectiveness and robustness of the proposed pipeline or the physical interpretability of the resulting feature set.

Strengths

- Clear and coherent end-to-end description of the preprocessing and feature-construction pipeline, from raw 3D fields to standardized feature matrix and temporal-derivative targets (Sec. 2, Sec. 3).
- Well-motivated use of local quadratic regression for temporal derivatives on very sparse time data, with quantitative RMSE evidence in both smooth and sharp-gradient regions supporting its low-noise performance (Sec. 2.3, Sec. 3.2).

- Careful benchmarking of spatial derivative schemes (spectral vs. WENO5) using incompressibility as a physics-based criterion, with both quantitative divergence metrics and visual comparisons (Sec. 2.4, Sec. 3.3).
- Physically motivated momentum-residual analysis that highlights significant unmodeled forces and motivates inclusion of proxy terms for pressure and buoyancy in the feature library (Sec. 2.6, Sec. 3.4).
- Comprehensive, physically inspired feature library (26 terms) including density, kinetic energy, divergence, and velocity–density coupling terms, with independent standardization of all fields and features to improve numerical conditioning for subsequent sparse regression (Sec. 2.5, Sec. 3.5).
- Generally clear writing and structure, with a logical progression from data characterization through methods and results to a concise Conclusions section (Sec. 1, Sec. 3, Sec. 4).
- The figures employ clear multi-panel layouts, consistent color/marker schemes, and effective overlays of raw data with fitted curves or comparative fields, facilitating straightforward visual assessment and supporting the manuscript's methodological claims.
- Core vector-calculus identities used in the feature definitions are stated consistently (e.g., divergence $\nabla \cdot \mathbf{u}$, advection $(\mathbf{u} \cdot \nabla)\mathbf{u}$, Laplacian $\nabla^2\mathbf{u}$).
- The spectral differentiation rules described ($\partial/\partial x_j$ via ik_j multiplication; $\partial^2/\partial x_j^2$ via $-k_j^2$) are internally consistent with Fourier differentiation on periodic domains.
- The normalization/standardization operations are described in a coherent way (mean subtraction and division by standard deviation), and the later per-feature (column-wise) standardization is consistent with the stated goal of conditioning a regression design matrix.

Major issues

1. **The core scientific goal framed in the Introduction is data-driven discovery of governing equations, yet the paper does not perform any sparse-regression or equation-discovery experiments, stopping at feature and residual construction (Introduction, Sec. 1; Sec. 2.5–2.6; Sec. 3.5; Sec. 4).** Claims that the resulting feature matrix is "optimally conditioned" and suitable for equation discovery are therefore not substantiated by demonstrations that Navier–Stokes-like equations (with buoyancy) can be recovered, nor is the robustness of the pipeline assessed in practice.

Recommendation: Either (a) add at least one concrete equation-discovery experiment (e.g., SINDy, PDE-FIND, sequential thresholded least squares, or LASSO) using the constructed feature matrix and temporal derivatives to recover the momentum equations, reporting identified terms, sparsity patterns, coefficient and predictive errors, or (b) explicitly reposition the work as a methodological note or dataset-preparation

study by revising the title, abstract, Introduction, and Conclusion to de-emphasize equation-discovery claims and clearly state that identification of governing equations is left for future work.

- 2. The justification and assessment of the local quadratic temporal regression are incomplete.** Only a second-order polynomial is used, with fit quality evaluated mainly via RMSE at one "smooth" and one "sharp" point, and little discussion of dependence on temporal spacing, boundary treatment, or possible non-polynomial dynamics (Sec. 2.3; Sec. 3.2). Low RMSE against the same data used for fitting does not guarantee accurate derivatives, especially near the first and last time slices.

Recommendation: Augment Sec. 2.3 and Sec. 3.2 with a more thorough analysis of temporal-derivative accuracy: (i) test first-, second-, and possibly third-order polynomials and/or regularized finite differences on representative points; (ii) if available, validate $\partial\mathbf{u}/\partial t$ against higher-resolution or analytically known derivatives from the underlying simulation (possibly via synthetic subsampling to 10 time points); and (iii) explicitly describe and evaluate the handling of boundary time slices (e.g., one-sided vs. symmetric windows). Report summary statistics of derivative errors or RMSE distributions across a larger set of points to demonstrate robustness beyond two sample locations.

- 3. The viscosity estimation and subsequent momentum-residual analysis (Sec. 2.6; Sec. 3.4) rely on a single global linear regression of $\partial\mathbf{u}/\partial t$ against $\nabla^2\mathbf{u}$, neglecting advection, pressure gradients, forcing, subgrid-scale effects, and discretization errors that are later shown to be large.** This likely biases the effective ν and undermines the strength of conclusions that residuals provide "unequivocal" evidence of pressure-gradient dominance.

Recommendation: Refine or clearly qualify the viscosity estimation in Sec. 2.6 and Sec. 3.4. For example, restrict the regression to regions where advection and pressure gradients are small, compare the inferred ν against any known or expected value from the data source, or treat ν as an unknown to be learned in the eventual sparse regression rather than pre-estimating it. In Sec. 2.6, Sec. 3.4, and Sec. 4, temper the language to state that large, structured residuals are consistent with significant unmodeled terms (including pressure and buoyancy) but are not an unequivocal proof of pressure dominance, and acknowledge other plausible contributors such as unresolved scales and numerical errors.

- 4. The selection of spectral derivatives over WENO5 is based on very marginal differences in incompressibility metrics (e.g., RMS divergence differing by only 0.002) and primarily global diagnostics (Sec. 2.4; Sec. 3.3).** It is therefore not convincingly shown that spectral derivatives are clearly superior across the domain, particularly near sharp gradients where spectral methods can suffer from Gibbs phenomena.

Recommendation: Extend the benchmarking in Sec. 3.3 by (i) reporting divergence RMS and maximum errors separately for smooth and high-gradient subregions (e.g., using the \mathcal{G}^2 metric), (ii) including additional diagnostics such as correlation between spectral and WENO5 derivatives, RMS differences in key derived terms like $(\mathbf{u} \cdot \nabla)\mathbf{u}$ and $\nabla^2\mathbf{u}$, and (iii) if feasible, comparing against a trusted reference from the original simulation. If differences remain marginal, moderate the claims in Sec. 3.3 and Sec. 4 to note that spectral derivatives perform slightly better overall and are chosen for reasons such as periodic-domain compatibility and implementation simplicity, rather than clear empirical superiority near sharp gradients.

5. **The physical and mathematical grounding of several proxy terms in the feature library is only qualitatively described.** Terms such as $\nabla(u^2 + v^2 + w^2)$, $\nabla(\rho^2)$, $\nabla(\nabla \cdot \mathbf{u})$, and components of $\rho'\nabla\mathbf{u}$ are introduced as proxies for pressure and buoyancy without systematic derivation from incompressible or Boussinesq momentum/energy equations or discussion of identifiability and multicollinearity among correlated gradient-based features (Sec. 2.5; Sec. 3.5).

Recommendation: In Sec. 2.5, provide an explicit rationale for each major class of proxy term, linking them to terms or diagnostic relations in the incompressible/Boussinesq Navier–Stokes equations (e.g., Bernoulli-type $\nabla(u^2/2)$ terms, pressure Poisson equation, baroclinic effects). Clearly indicate which proxies are intended to approximate pressure gradients, which represent buoyancy or other physics, and which serve as generic nonlinearities. Additionally, discuss potential redundancy and multicollinearity among gradient-based terms and consider adding a brief correlation or condition-number analysis of the feature library. Outline how such issues will be handled in later regression (e.g., via regularization, feature pruning, or dimensionality reduction).

6. **The paper’s positioning relative to existing PDE-discovery literature (e.g., SINDy/PDE-FIND, weak-form approaches, and prior work on noisy/sparse data, spectral vs. finite-difference derivatives, and proxy terms for unmeasured variables) is limited to scattered remarks, with no dedicated Related Work section (Sec. 1).** This makes it difficult to assess the novelty of the pipeline beyond standard practices in the field.

Recommendation: Add a dedicated Related Work subsection (e.g., Sec. 1.1 or early in Sec. 2) that reviews relevant PDE-discovery methods, derivative-estimation strategies under noise/sparsity, and previous uses of residual analysis or proxy features for unmeasured quantities. Explicitly compare the present pipeline to representative works, clarifying what is new (e.g., specific benchmarking protocol for 3D nearly incompressible flows with sharp gradients, residual-guided feature design, or standardization strategy). Adjust claims in Sec. 1 and Sec. 4 to reflect this context and avoid overstating novelty.

7. Reproducibility is hampered by missing implementation and data-provenance details. Key algorithmic parameters (e.g., exact temporal spacing, polynomial-fitting procedure and weighting, FFT conventions and dealiasing, computation of the G^2 gradient-energy metric) are not fully specified, and the origin and governing equations of the dataset are only briefly described (Sec. 2.1–2.4; Sec. 3.1–3.3).

Recommendation: Expand Sec. 2.1–2.4 and Sec. 3.1 to include: (i) the temporal step size and whether sampling is uniform; (ii) details of the polynomial regression (ordinary vs. weighted least squares, window size, regularization, and boundary handling); (iii) FFT implementation details (normalization, wavenumber ordering, de-aliasing or filtering strategy, treatment of Nyquist modes); (iv) the exact definition and computation method for G^2 , including which derivative scheme and whether standardized or raw fields are used; and (v) a concise description of the dataset source (simulation vs. experiment, underlying equations, key dimensionless parameters, and whether a Boussinesq model is used). Consider adding an appendix that tabulates all numerical parameters and stating any plans for releasing the data and preprocessing/feature-generation code.

8. Figures 1 and 2 lack uncertainty visualization (e.g., confidence bands) around polynomial fits, quantitative fit metrics (such as RMSE or R^2) within panels, and clear mapping of the time axis to physical or nondimensional units. Figure 1 also omits specification of standardization and does not display residual diagnostics or address potential edge effects. Figure 2 does not include residual diagnostics or clarify axis abbreviations. Both figures have readability issues due to small fonts and low resolution.

Recommendation: Add shaded 95% confidence or bootstrap intervals for fitted curves and propagate to derivatives; annotate each panel with RMSE (and optionally R^2) and sample size; clarify time axis units and standardization method in labels or captions; include residual diagnostics (e.g., insets or error markers); increase DPI, font sizes, and export as vector graphics for publication quality.

9. Figure 3 lacks explicit quantitative comparison between spectral and WENO5 schemes (e.g., difference maps or error metrics), may use inconsistent color limits, omits units or normalization for derivatives, and is too small for reliable interpretation. Panels are unlabeled and lack method/time/plane annotations.

Recommendation: Add a difference panel with tight, zero-centered color scale and inset statistics (mean, std, RMSE); enforce identical colormap and color limits; state units/normalization in colorbars or captions; increase figure and font size; add panel letters and concise titles with method and context annotations.

10. **The paper claims a “final library comprised 26 distinct candidate terms” but does not provide a complete, explicit, unambiguous list of all 26 analytic terms; in particular, the “additional density-weighted velocity gradient terms derived from $\rho'\nabla\mathbf{u}$ ” are not enumerated, making the library definition mathematically under-specified.**

Recommendation: Add a table enumerating all 26 features with explicit formulas and indexing (e.g., $\Theta_1 = 1$, $\Theta_2 = (\mathbf{u} \cdot \nabla)\mathbf{u}$, ...) and clarify exactly which components of $\rho'\nabla\mathbf{u}$ are included (e.g., $\rho'\partial u/\partial x$, $\rho'\partial u/\partial y$, ...) so the feature matrix Θ is uniquely defined.

11. **The status of units/scaling for the ‘kinematic viscosity’ estimate ν and momentum residual R_u is not mathematically verifiable from the text because it is unclear whether ν is estimated using raw (physical/nondimensional) fields or standardized fields, and component-wise standardization can change coefficient interpretation unless the transformation is accounted for.**

Recommendation: Explicitly state whether derivatives/features and the ν regression are computed from raw fields or standardized fields; if standardized fields are used, provide the variable transformation and show how ν (and any inferred coefficients) maps back to the unstandardized equation, or clearly relabel ν as a coefficient in standardized coordinates.

12. **None of the proposed numerical checks were executed due to an execution error (sandbox policy violation), so no PASS/FAIL determinations could be produced from the provided check instructions.**

Recommendation: Re-run the numerical audit in a permitted execution environment (or adjust the checker to comply with sandbox import policies) so the listed recomputations and consistency checks can be completed and logged.

Minor issues

1. The scope and limitations of the work are not always clearly stated: the Introduction and some later sections could be read as implying full equation discovery, and the strong phrasing around pressure-gradient dominance and spectral-method superiority may overstate what is supported by the presented diagnostics (Introduction, Sec. 1; Sec. 2.6; Sec. 3.3–3.4; Sec. 4).

Recommendation: In the final paragraph of the Introduction (Sec. 1) and in the Conclusions (Sec. 4), explicitly state that the paper focuses on generating and conditioning features and targets for subsequent sparse-regression-based equation discovery, which is deferred to future work. At the same time, moderate language around pressure dominance and derivative-scheme choice to indicate that findings "suggest" or are "consistent with" certain interpretations, rather than being "unequivocal" or definitive.

2. The description of data normalization and standardization could be misinterpreted, particularly regarding the order of operations (e.g., conversion from ρ to ρ' , computation of global statistics, and standardization of primary fields vs. derived features) and the mixed use of the terms "normalization" and "standardization" (Sec. 2.2; Sec. 2.5; Sec. 3.1; Sec. 3.5).

Recommendation: In Sec. 2.2 and Sec. 2.5, explicitly list the preprocessing steps (e.g., 1) compute $\rho' = \rho - 1$; 2) compute global means and standard deviations of u, v, w, ρ' over all x, y, z, t ; 3) standardize these primary fields; 4) construct features from standardized fields; 5) standardize each feature column independently). Use "normalization" only for operations like defining ρ' , and reserve "standardization" for mean-zero, unit-variance scaling. In Sec. 3.1 and Sec. 3.5, refer back to this procedure to reinforce consistency.

3. The evaluation of temporal polynomial fits in Sec. 3.2 focuses on RMSE at two selected points (one "smooth" and one "sharp" identified via G^2), with no global statistics to demonstrate performance across the entire domain and all variables.

Recommendation: Extend Sec. 3.2 to report summary statistics of fit quality or derivative accuracy over a larger random sample or the full domain (e.g., distributions of RMSE per component, median/percentile values, or norms of residuals). This will support the claim that temporal derivative estimation is robust globally rather than only at two illustrative points.

4. The definition and use of the gradient-energy metric $G^2 = \sum_{i,j} (\partial u_i / \partial x_j)^2$ to distinguish smooth and sharp regions are introduced without specifying how the derivatives are computed (spectral or WENO5) and whether standardized or raw fields are used (Sec. 3.2).

Recommendation: Clarify in Sec. 3.2 the exact procedure for computing G^2 : state which derivative scheme is used, which components are included, and whether derivatives are based on standardized or physical-unit fields. Briefly justify this choice so readers can interpret the reported G^2 values in relation to flow structures and derivative quality.

5. The construction of the feature library mentions "additional density-weighted velocity gradient terms derived from $\rho' \nabla \mathbf{u}$ " but does not explicitly enumerate these components or relate them clearly to the stated total of 26 features, leaving ambiguity about the exact feature set (Sec. 2.5; Sec. 3.5).

Recommendation: Add a table or explicit bullet list in Sec. 2.5 (or an appendix) that enumerates all 26 feature terms with their mathematical expressions, including the precise set of components from $\rho' \nabla \mathbf{u}$ (e.g., all 9 tensor components or a subset). In Sec. 3.5, reference this list and confirm that the same feature set is used in all analyses.

6. The dataset description and physical assumptions (nearly incompressible, low-Mach/Boussinesq regime) are only briefly sketched and are not clearly tied to specific non-dimensional parameters or to the governing equations of the data source (Sec. 2.1; Sec. 3.1).

Recommendation: In Sec. 2.1 and Sec. 3.1, provide a concise description of the data provenance (simulation code or experiment, governing equations, forcing, Reynolds number and any relevant non-dimensional numbers, and whether a Boussinesq approximation is used). Connect observed density fluctuations (e.g., $\sim 0.2\%$ of the mean) to standard low-Mach/Boussinesq criteria, citing appropriate references, and explicitly state which physical assumptions are adopted in constructing the feature library (e.g., constant viscosity, dominant buoyancy term).

7. The choice and interpretation of a constant term in the feature library are only briefly justified as accounting for mean offsets or uniform forcing, which may be unclear in a decaying incompressible flow on a periodic domain (Sec. 2.5; Sec. 3.5).

Recommendation: In Sec. 2.5, more clearly explain under what circumstances a constant term might arise effectively (e.g., mean pressure gradient or uniform body force) and emphasize that including a constant feature allows the sparse regression to detect and discard it if unnecessary. Optionally, note any empirical evidence from residuals that motivates its inclusion.

8. The momentum-residual analysis in Sec. 3.4 is mainly qualitative; quantitative comparisons between residual magnitudes and modeled terms (e.g., advection, temporal derivative) are limited, which weakens quantitative support for the interpretation of dominant unmodeled forces.

Recommendation: Enhance Sec. 3.4 with quantitative diagnostics, such as distributions or norms of $|R_u|$ compared to $|\partial u/\partial t|$ and $|(\mathbf{u} \cdot \nabla)u|$, spatial correlation plots with density or kinetic energy, or ratios of residual to modeled term magnitudes. Use these metrics to support more nuanced statements about the relative importance of unmodeled contributions.

9. Some implementation and documentation details that affect interpretability and reproducibility are only implicit, such as: (i) how temporal-regression windows are defined at the first and last time slices; (ii) how aliasing or spectral ringing is handled in FFT-based derivatives; and (iii) how the total number of spatio-temporal points relates to the grid and time dimensions (Sec. 2.3–2.4; Sec. 3.3; Sec. 3.5).

Recommendation: In Sec. 2.3, explicitly describe the temporal windowing strategy for polynomial fits, including boundary handling. In Sec. 2.4 and Sec. 3.3, state whether any filtering or dealiasing is applied to FFT-based derivatives and briefly discuss potential Gibbs phenomena near sharp gradients. In Sec. 3.5, explicitly connect the reported 20,971,520 spatio-temporal points to the 128^3 grid and 10 time slices for clarity.

10. Ethical and broader-impact considerations of data-driven PDE discovery are not discussed (Sec. 1–4), even though some venues expect at least a brief statement, albeit the present application is to fluid dynamics data.

Recommendation: Add a short paragraph near the end of Sec. 4 noting that the work uses simulation/physical data only, carries no direct human or animal subjects implications, and briefly commenting on potential downstream impacts (e.g., improved modeling in engineering or climate applications) and any associated responsibilities.

11. Across Figures 1 and 2, panel titles and legends include implementation-centric or code-like details, legends are duplicated, y-axis ranges differ across panels, and color encoding alone may challenge some viewers. Figure 2 also uses ambiguous abbreviations and may obscure markers with fit lines.

Recommendation: Replace technical titles with concise scientific descriptions; use shared or compact legends; adopt consistent y-limits or annotate scales; differentiate data and fits with both color and shape/line style; clarify abbreviations and improve marker visibility.

12. Figures 3, 4, and 5 lack explicit details on grid resolution, domain size, boundary conditions, and normalization. Panel labels, axis units, and scale information are often missing or inconsistent. Figure 4 mixes spatial and distributional scopes, and Figure 5's right-column panels misrepresent distributions and lack zero reference lines.

Recommendation: Annotate grid/domain details and normalization in captions or panels; add panel labels and consistent axis units; clarify scope of distributions; replace line traces with true distributions and overlay zero reference lines.

13. Notation overload: the symbol \mathbf{u} is used both for the velocity vector $\mathbf{u} = (u, v, w)$ and for the x -component u , which makes expressions like $\nabla^2 \mathbf{u}$ and $(\mathbf{u} \cdot \nabla) \mathbf{u}$ potentially ambiguous without context.

Recommendation: Use boldface (e.g., \mathbf{u}) for vectors and plain symbols for components (u, v, w) , or adopt index notation (u_i) consistently when defining operators and residuals.

14. The viscosity-regression description ‘regressing $\partial \mathbf{u} / \partial t$ against $\nabla^2 \mathbf{u}$ ’ omits necessary assumptions for the implied relation (e.g., neglecting advection and pressure) and does not show the algebraic model being fit; as written, the mapping from Navier–Stokes form to that regression is not derivable from the paper alone.

Recommendation: State the regression model explicitly (e.g., $\partial \mathbf{u} / \partial t \approx \nu \nabla^2 \mathbf{u} + \epsilon$) and list which terms are omitted and why; clarify whether the regression is componentwise (u, v, w) and how the slopes are combined.

15. Several key quantitative claims (e.g., RMSEs, divergence RMS/max, viscosity regression, and exact standardization) are not verifiable from text alone and require underlying data/fields.

Recommendation: Provide the underlying arrays/feature matrix (or sufficient intermediate summaries) to allow recomputation of RMSEs, divergence statistics, regression outputs, and standardization properties.

Very minor issues

1. There are several minor stylistic and formatting inconsistencies: mixed section-heading styles (e.g., plain "2 Methods" vs. hash-prefixed "# 4 Conclusions"), inconsistent use of quotes or italics around "smooth" and "sharp", occasional double spaces around inline math or punctuation, and small math-notation artefacts such as stray backslashes in subscripts (Sec. 2–4; Sec. 3.2–3.5).

Recommendation: Standardize section headings according to the target template (e.g., consistent use of \section/\subsection), use a uniform style for labels like "smooth" and "sharp", remove extra spaces around inline math, and proofread all equations and figure captions to correct minor notation issues (e.g., ensure R_u rather than R_u).

2. Some phrases are stylistically strong or repetitive relative to the evidence, such as "catastrophic noise amplification", "exceptional accuracy", "marginal but significant advantage", and "unequivocally confirming the dominant role" (Abstract; Sec. 3.1–3.4; Sec. 4).

Recommendation: Lightly edit the text in the Abstract, Sec. 3.1–3.4, and Sec. 4 to adopt a more neutral scientific tone: replace dramatic or absolute wording with qualified alternatives (e.g., "low RMSE relative to standardized fields", "small but consistent improvement", "strongly suggests"), and reduce repetition of superlatives.

3. Figure captions in Sec. 3.2–3.5 are generally clear but sometimes repeat long phrases from the main text and include minor spacing/notation inconsistencies.

Recommendation: Tighten figure captions in Sec. 3.2–3.5 by removing redundant wording while preserving essential information (what is plotted, where/when, and why), and standardize math formatting and punctuation within captions.

4. Figures use inconsistent or non-standard typography for variables and subscripts, ambiguous or code-like notation, and lack panel labels or references in captions. Some captions are lengthy or use interpretive language. Colorbar ticks, gridlines, and marker sizes may hinder clarity at small scales.

Recommendation: Standardize math typography and subscript formatting; add panel labels and reference them in captions; condense captions and use descriptive rather than interpretive language; optimize colorbar ticks, gridlines, and marker sizes for clarity.

5. Some figures do not specify what colors represent, lack accessibility considerations, or omit sample size and normalization details in legends or captions.

Recommendation: Add brief legend notes for color meaning; confirm use of colorblind-safe palettes or provide grayscale alternatives; include sample size and normalization details in legends or captions.

6. The term ‘normalized density perturbation’ ρ' is defined as $\rho - 1.0$ and later all fields are standardized; the text sometimes refers to ρ' as if it remains only mean-shifted, which can be read as a mild terminology mismatch.

Recommendation: Distinguish clearly between ‘density perturbation’ ($\rho - 1$) and ‘standardized density perturbation’ ($(\rho - 1 - \mu)/\sigma$) with distinct notation (e.g., $\tilde{\rho}$ for standardized).

7. The gradient-energy definition $G^2 = \sum_{i,j} (\partial u_i / \partial x_j)^2$ is given without stating whether derivatives are with respect to physical coordinates or grid indices, which affects dimensional interpretation (though not the qualitative use as a selector).

Recommendation: Specify whether x, y, z are physical coordinates (with stated domain length) or grid coordinates, and whether u_i are raw or standardized when computing G^2 .

8. The candidate-term count consistency check depends on an interpretation of what constitutes a distinct term (e.g., counting vector components separately), which can be ambiguous from text descriptions.

Recommendation: Clarify the exact term-counting scheme used to arrive at 26 distinct terms (e.g., explicitly list all 26 terms or provide a table).

Key statements and references

- • **The near-unity global mean density of 1.0000000000061107 with a standard deviation of approximately 0.00217 indicates that the flow operates in a nearly incompressible, low-Mach number regime consistent with the Boussinesq approximation, in which density fluctuations primarily contribute to buoyancy forces rather than inertial mass variations.**
- *Reference(s):* Boussinesq
- • **For periodic domains with nearly incompressible flow, spectral methods for computing spatial derivatives via Fast Fourier Transforms theoretically provide high accuracy for smooth functions, whereas 5th-order Weighted Essentially Non-Oscillatory (WENO5) finite-difference schemes are specifically designed to robustly handle localized sharp spatial gradients without introducing spurious oscillations.**
- *Reference(s):* WENO5

- • The preliminary physical consistency check is based on the standard incompressible Navier–Stokes momentum equation, in which the temporal derivative of velocity, nonlinear advection term, viscous term proportional to the Laplacian of velocity with kinematic viscosity ν , and the pressure gradient together govern the dynamics of each velocity component in an incompressible fluid.
- *Reference(s)*: Navier-Stokes

Mathematical consistency audit

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

Maths relevance: light

The paper is primarily methodological and descriptive, with limited formal derivations. The mathematics consists mainly of definitions of transformations ($\rho' = \rho - 1$), standardization, standard vector operators used to form a regression feature library (divergence, advection, Laplacian, gradients), and a residual definition $R_u = \partial u / \partial t + (\mathbf{u} \cdot \nabla)u - \nu \nabla^2 u$. The main internal-consistency risks are (i) under-specification of the full 26-term library and (ii) unclear variable scaling/units when interpreting estimated coefficients like ν after standardization.

Checked items

- ✓ **Density perturbation definition** (Sec. 2.2, p.3 (also reiterated Sec. 3.1, p.6))
 - **Claim:** Defines density perturbation as $\rho' = \rho - 1.0$.
 - **Checks:** definition consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** ρ is already in units/normalization where 1.0 is the reference (mean) density.
 - **Notes:** Definition is self-consistent and reused consistently.
- ✓ **Primary-field standardization** (Sec. 2.2, p.3; Sec. 3.1, p.6)
 - **Claim:** Each of (u, v, w, ρ') is standardized by subtracting its global mean and dividing by its global standard deviation to yield mean 0 and std 1.
 - **Checks:** definition consistency, dimensional consistency
 - **Verdict:** PASS; confidence: medium; impact: moderate
 - **Assumptions/inputs:** Means/stds are computed over all spatial points and time slices., Standard deviations are nonzero.
 - **Notes:** Mathematically consistent as a transformation; however, later physical interpretation of coefficients after such scaling needs clarification (handled in separate item).
- ✓ **Temporal derivative via quadratic regression** (Sec. 2.3, p.3; Sec. 3.2, p.6-7)

- **Claim:** At each spatial point, fit a quadratic in time to the 10 samples of each variable and take its time derivative evaluated at the time points.
 - **Checks:** logical derivation check, definition consistency
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Time coordinates used for the fit are defined (not shown in text)., Fit is over all 10 points (global in time for each spatial point).
 - **Notes:** Method description is internally coherent. Exact polynomial/weighting is not specified but not required for internal symbolic consistency.
4. ✓ **Spectral first-derivative rule** (Sec. 2.4, p.3-4)
- **Claim:** In Fourier space, first derivative in direction j is obtained by multiplying by ik_j and inverse transforming.
 - **Checks:** algebra/vector calculus
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** Periodic domain; Fourier transform conventions where differentiation corresponds to multiplication by ik .
 - **Notes:** Symbolically correct Fourier differentiation statement.
5. ✓ **Spectral second-derivative rule** (Sec. 2.4, p.3-4)
- **Claim:** In Fourier space, second derivative in direction j is obtained by multiplying by $-k_j^2$.
 - **Checks:** algebra/vector calculus
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** Same Fourier convention as above.
 - **Notes:** Consistent with applying the first-derivative rule twice.
6. ✓ **Divergence definition** (Sec. 2.4, p.4)
- **Claim:** Defines $\nabla \cdot \mathbf{u} = \partial u / \partial x + \partial v / \partial y + \partial w / \partial z$.
 - **Checks:** definition consistency, notation consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** $\mathbf{u} = (u, v, w)$ is the velocity field in (x, y, z) .
 - **Notes:** Standard definition; consistent with later use of $\nabla \cdot \mathbf{u}$.
7. ✓ **Incompressibility constraint usage** (Sec. 2.4, p.4; Sec. 3.3, p.8-9)
- **Claim:** Uses $\nabla \cdot \mathbf{u} = 0$ as a physical constraint to benchmark derivative schemes.
 - **Checks:** logical consistency
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Flow is ‘nearly incompressible’ so divergence is expected to be small.

- **Notes:** Internally coherent as a selection criterion; not a symbolic inconsistency.
8. ✓ **Advection term component form** (Sec. 2.4, p.4)
- **Claim:** States components of $(\mathbf{u} \cdot \nabla)\mathbf{u}$, e.g., $(\mathbf{u} \cdot \nabla)u_x = u\partial u/\partial x + v\partial u/\partial y + w\partial u/\partial z$.
 - **Checks:** vector calculus, notation consistency
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** \mathbf{u} is velocity vector with components (u, v, w) .
 - **Notes:** Correct component expansion for the convective derivative $(\mathbf{u} \cdot \nabla)$ applied to a component.
9. ✓ **Vector Laplacian definition** (Sec. 2.4, p.4)
- **Claim:** Defines $\nabla^2\mathbf{u} = (\nabla^2u, \nabla^2v, \nabla^2w)$.
 - **Checks:** definition consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** Laplacian is applied componentwise.
 - **Notes:** Consistent with standard componentwise Laplacian on vector fields.
10. ✓ **Gradient energy G^2 definition** (Sec. 3.2, p.6)
- **Claim:** Defines local velocity gradient energy as $G^2 = \sum_{i,j}(\partial u_i/\partial x_j)^2$.
 - **Checks:** definition consistency, dimensional consistency
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** Indices i, j run over spatial components and coordinates.
 - **Notes:** Definition is internally consistent (squared Frobenius norm of $\nabla\mathbf{u}$). Dimensional meaning depends on whether \mathbf{u} and \mathbf{x} are raw/standardized/physical; this is a clarity gap rather than an algebraic error.
11. ✓ **Proxy feature: gradient of kinetic-energy-like scalar** (Sec. 2.4-2.5, p.4-5)
- **Claim:** Includes $\nabla(u^2 + v^2 + w^2)$ as a candidate proxy term.
 - **Checks:** algebra, definition consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** u, v, w are scalar component fields.
 - **Notes:** Expression is algebraically well-formed; it is a gradient of a scalar. Physical naming as ‘kinetic energy’ depends on scaling (not checked).
12. ✓ **Proxy feature: gradient of squared density perturbation** (Sec. 2.4-2.5, p.4-5)
- **Claim:** Includes $\nabla(\rho'^2)$ and its component derivatives as candidate terms.
 - **Checks:** algebra, definition consistency

- **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** ρ' is a scalar field.
 - **Notes:** Algebraically consistent (gradient of a squared scalar).
13. ✓ **Proxy feature: gradient of divergence** (Sec. 2.4-2.5, p.4-5)
- **Claim:** Includes $\nabla(\nabla \cdot \mathbf{u})$ and component derivatives as candidate terms.
 - **Checks:** vector calculus
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** \mathbf{u} is sufficiently differentiable for second derivatives.
 - **Notes:** Well-posed operator composition (gradient of a scalar divergence).
14. ✓ **Momentum residual definition** (Sec. 2.6, p.5)
- **Claim:** Defines $R_u = \partial u / \partial t + (\mathbf{u} \cdot \nabla)u - \nu \nabla^2 u$ (and analogously for v, w).
 - **Checks:** algebra, notation consistency
 - **Verdict:** PASS; confidence: medium; impact: moderate
 - **Assumptions/inputs:** ν is scalar coefficient; Laplacian applied to the same component., The symbol u in $(\mathbf{u} \cdot \nabla)u$ is the velocity vector; the first u in $\partial u / \partial t$ is the u -component.
 - **Notes:** Expression is algebraically consistent as a residual for a model with only advection and viscosity on the RHS moved to LHS. Notation is ambiguous due to u denoting both vector and component.
15. △ **Viscosity regression model (derivation completeness)** (Sec. 2.6, p.5; Sec. 3.4, p.9)
- **Claim:** Estimates ν by linear regression of $\partial u / \partial t$ against $\nabla^2 u$ across domain/time and averages across spatial dimensions.
 - **Checks:** derivation completeness, assumption clarity
 - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
 - **Assumptions/inputs:** Implicit regression model is $\partial u / \partial t \approx \nu \nabla^2 u$ (plus noise).
 - **Notes:** The paper does not write the explicit fitted equation, nor does it state which terms are assumed negligible/absorbed by noise. Symbolically plausible as a procedure, but the analytic link to a stated governing equation is not provided in-text.
16. △ **Feature-library size and completeness (26 terms)** (Sec. 2.5, p.4-5; Sec. 3.5, p.9-10)
- **Claim:** States Θ contains 26 distinct candidate terms including constants, advection, Laplacians, multiple gradients, cross-terms $u\rho', v\rho', w\rho'$, and additional terms derived from $\rho' \nabla \mathbf{u}$.
 - **Checks:** definition completeness, symbol/term consistency

- **Verdict:** UNCERTAIN; confidence: high; impact: critical
- **Assumptions/inputs:** ‘Additional density-weighted velocity gradient terms’ are a specific finite subset.
- **Notes:** The analytic definition of Θ is not fully specified: the paper does not enumerate exactly which components of $\rho'\nabla\mathbf{u}$ (a tensor with multiple components) are included, so the claimed total of 26 terms cannot be verified from the PDF alone.

17. ✓ **Per-feature standardization of Θ columns** (Sec. 2.5, p.5; Sec. 3.5, p.10)

- **Claim:** Independently standardizes each column of Θ to mean 0 and standard deviation 1.
- **Checks:** definition consistency, logical consistency
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** Standard deviation of each feature column is nonzero.
- **Notes:** Mathematically consistent preprocessing step; however, it implies any discovered coefficients will be in standardized-feature units unless mapped back.

18. △ **Interpretation of ν after standardization** (Sec. 2.2-2.6, p.3-5; Sec. 3.4, p.9)

- **Claim:** Reports ν as ‘kinematic viscosity’ and uses it in residuals, while earlier stating fields are standardized.
- **Checks:** dimensional/units consistency, definition consistency
- **Verdict:** UNCERTAIN; confidence: high; impact: critical
- **Assumptions/inputs:** Either computations for ν use raw fields, or there is a mapping from standardized to physical coefficients.
- **Notes:** Because standardization generally removes physical units and may involve different scalings for u, v, w , the coefficient labeled ν cannot be confirmed as a physical kinematic viscosity without an explicit statement that ν was computed on raw variables or an explicit back-transformation.

Limitations

- The PDF text provides essentially no numbered equations and omits detailed derivation steps; many claims are methodological descriptions rather than derivations that can be audited step-by-step.
- The full analytic specification of the 26-term feature library Θ is incomplete in the PDF (not all terms are explicitly listed), preventing a complete internal-consistency check of Θ .
- The paper does not explicitly state whether derivatives/residuals/regressions are computed on raw or standardized fields in every context, limiting unit/coefficient consistency verification.

Numerical results audit

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

Ten candidate numerical checks were specified (integer factor-product consistency, definition-based derivation, simple aggregations, differences/ratios, scale comparison, and a term-count logic check), but no checks were executed due to an execution error; therefore all items remain UNCERTAIN.

Checked items

1. \triangle **C1_domain_points_count** (Methods §2.1 (p.3) and Results §3.5 (p.9): “ 128^3 periodic domain ... 10 time slices” and “evaluated at all 20,971,520 spatio-temporal points”)
 - **Claim:** The paper states a 128^3 spatial grid and 10 time slices, and later states there are 20,971,520 spatio-temporal points.
 - **Checks:** `recompute_total_from_factors`
 - **Verdict:** UNCERTAIN
 - **Notes:** Check was specified ($128^3 \times 10 = 20,971,520$) but not executed due to execution error.
2. \triangle **C2_density_perturbation_mean_consistency** (Methods §2.2 (p.3): “ $\rho' = \rho - 1.0$ ” and Results §3.1 (p.5): “ ρ exhibited a global mean of 1.000000000061107”)
 - **Claim:** Given $\rho' = \rho - 1.0$, the mean of ρ' should equal $\text{mean}(\rho) - 1.0$.
 - **Checks:** `derived_value_from_definition`
 - **Verdict:** UNCERTAIN
 - **Notes:** Only the implied arithmetic could be checked from the stated $\text{mean}(\rho)$; execution did not occur due to execution error.
3. \triangle **C3_rmse_smooth_velocity_mean** (Results §3.2 (p.6): “At the smooth point, the RMSE values were 0.0031, 0.0013, 0.0021 ... for u, v, w ...”)
 - **Claim:** Compute an aggregate (e.g., mean) RMSE over (u, v, w) at the smooth point from the listed values to support quick cross-checks or downstream comparisons.
 - **Checks:** `cheap_aggregation_recompute`
 - **Verdict:** UNCERTAIN
 - **Notes:** Deterministic aggregation was specified but not executed due to execution error.
4. \triangle **C4_rmse_sharp_velocity_mean** (Results §3.2 (p.6): “at the sharp gradient point, the RMSE values remained ... 0.0066, 0.0014, 0.0026 ... for u, v, w ...”)

- **Claim:** Compute an aggregate (e.g., mean) RMSE over (u, v, w) at the sharp-gradient point from the listed values to support quick cross-checks.
 - **Checks:** cheap_aggregation_recompute
 - **Verdict:** UNCERTAIN
 - **Notes:** Deterministic aggregation was specified but not executed due to execution error.
5. \triangle **C5_rmse_rhoprime_smooth_vs_sharp_difference** (Results §3.2 (p.6): “RMSE ... 0.0250 for ρ' (smooth) ... 0.0242 (sharp)”)
- **Claim:** Check the absolute and relative difference between the two reported ρ' RMSE values (smooth vs sharp).
 - **Checks:** difference_and_ratio_check
 - **Verdict:** UNCERTAIN
 - **Notes:** Deterministic difference/ratio was specified but not executed due to execution error.
6. \triangle **C6_gradient_energy_ratio_sharp_to_smooth** (Results §3.2 (p.6): “smooth ... $G^2 \approx 7.75$ ” and “sharp ... $G^2 \approx 19706.18$ ”)
- **Claim:** Compute ratio of sharp-region gradient energy to smooth-region gradient energy.
 - **Checks:** ratio_recompute
 - **Verdict:** UNCERTAIN
 - **Notes:** Inputs are approximate (\approx) per statement; computation was specified but not executed due to execution error.
7. \triangle **C7_derivative_difference_mean_near_zero** (Results §3.3 (p.8): “difference ... yielded a mean near zero (-1.14×10^{-8}) and ... standard deviation (0.056)”)
- **Claim:** Check that the stated mean magnitude is tiny compared to the stated standard deviation.
 - **Checks:** scale_comparison
 - **Verdict:** UNCERTAIN
 - **Notes:** Heuristic scale comparison was specified but not executed due to execution error.
8. \triangle **C8_incompressibility_rms_improvement** (Results §3.3 (p.9) and Abstract (p.1): spectral RMS 0.5914 vs WENO5 RMS 0.5934)
- **Claim:** Compute absolute and relative improvement in RMS divergence error for spectral vs WENO5.
 - **Checks:** difference_and_ratio_check
 - **Verdict:** UNCERTAIN
 - **Notes:** Deterministic difference/ratio was specified but not executed due to execution error.

9. \triangle **C9_incompressibility_max_error_improvement** (Results §3.3 (p.9): max abs error 3.979 (spectral) vs 4.038 (WENO5))
- **Claim:** Compute absolute and relative improvement in max `|divergence|` for spectral vs WENO5.
 - **Checks:** `difference_and_ratio_check`
 - **Verdict:** UNCERTAIN
 - **Notes:** Deterministic difference/ratio was specified but not executed due to execution error.
10. \triangle **C10_feature_count_vs_enumerated_groups** (Methods §2.5 (p.4-5) and Results §3.5 (p.9): “final library comprised 26 distinct candidate terms” and bullet list of term groups)
- **Claim:** Verify that the enumerated term groups plausibly sum to 26 terms under a consistent counting scheme.
 - **Checks:** `count_from_enumeration`
 - **Verdict:** UNCERTAIN
 - **Notes:** Logical sum/implied remainder check was specified but not executed due to execution error; also depends on interpretation of term counting.

Limitations

- Only parsed text was available; no access to underlying numerical datasets or computed fields, so most performance metrics (RMSE, RMS divergence, viscosity regression) cannot be recomputed directly.
- No checks were proposed that require extracting numeric values from plot images (per instruction); figure-based values are treated as textual claims only.
- Some candidate checks (e.g., feature-term counting) depend on a reasonable interpretation of how vector quantities are counted as separate terms; ambiguity in term definitions could change the implied counts.
- Execution error prevented running the proposed check computations: Sandbox policy violation: `from-import` of `'typing'` is not allowed.