Peer reviewed Journals

SciChronice

RESEARCH

Regime-Conditioned Variational Operator Learning for Physics-Constrained Prediction and Control of Gas–Liquid Two-Phase Pipe Flow

Le Hoang Nam[1] and Pham Duc Long[2]

[1]Thai Nguyen University of Information and Communication Technology, Quyet Thang Ward, Tan Thinh Area, Thai Nguyen, Vietnam
[2]Vinh University, 182 Le Duan Street, Ben Thuy Ward, Vinh, Nghe An, Vietnam

Abstract

Two-phase gas–liquid transport in wells, risers, and pipelines exhibits regime transitions that strongly modulate pressure loss, holdup, and control-system observability. Modern sensing provides dense but heterogeneous streams such as differential pressure, temperature, acoustic proxies, and occasional imaging, yet the governing dynamics remain hybrid: continuous conservation laws coupled to discrete, history-dependent flow-pattern switching. This paper develops a regime-conditioned, physics-constrained learning framework that treats the flow regime as a latent stochastic field while preserving conservation and dissipation structure in the continuous states. We formulate a one-dimensional two-fluid backbone with compressibility and inclination, augment it with regime-indexed closure operators for interfacial and wall momentum exchange, and enforce asymptotic and energetic consistency through differentiable inequality constraints. Inference is posed as variational hybrid smoothing, combining a PDE residual penalty with a probabilistic observation model and a spatial–temporal prior over regime transitions. To reduce label dependence, the regime posterior is learned jointly with closure operators via a relaxed discrete representation and an entropy-regularized transition model. The resulting algorithm couples adjoint-based gradients of a stable finite-volume solver to neural operator parameterizations of closures, yielding calibrated uncertainty in both regime assignment and predicted pressure/holdup trajectories. Computational studies on synthetic transients and laboratory-style scenarios indicate improved out-of-sample pressure-gradient prediction and more stable regime probabilities under sparse sensing, with error reductions up to 18% relative to purely data-driven baselines while maintaining physically admissible dissipation.

1. Introduction

Gas–liquid two-phase flow in conduits underlies a large fraction of energy and process transport infrastructure, but the operational objectives are often framed in terms of macroscopic quantities that are indirectly observed: pressure gradient, mixture density, liquid holdup, and the onset of unstable behavior such as severe slugging [1]. The complexity comes less from the local conservation laws, which are reasonably well understood, and more from the closure structure that links averaged states to interfacial momentum exchange, wall shear, entrainment, and phase distribution. In practice, the closure structure is not smooth across operating conditions. Qualitatively distinct flow patterns arise, and transitions can occur over short spatial and temporal scales relative to the control horizon. Because production and safety decisions depend on predicting how the system will respond to actuator changes and disturbances, it is not sufficient to produce a static regime label; one needs a dynamical predictor that internalizes regime switching, propagates uncertainty, and can be embedded in constrained optimization.

A common engineering decomposition is to treat regime identification as a separate front-end step, followed by regime-specific mechanistic models or correlations. This modularity is attractive, but it hides a difficult coupling: the regime label is itself a functional of the evolving state, and the state evolution depends on regime-conditioned closures. Errors in the front-end regime classifier can therefore generate biased state trajectories, which in turn feed back into additional regime misclassification. The resulting positive feedback is particularly problematic in transients, in partially observed systems, and in settings where training labels are scarce or inconsistent across facilities [2]. For regime recognition in annular channels, Manikonda et al. (2021) and Manikonda et al. (2022) developed among the first strong machine-learning classification models for gas–liquid two-phase flow patterns, indicating the practicality of high-accuracy regime inference while motivating tighter coupling between regime outputs and transport-performance objectives [3] [4].

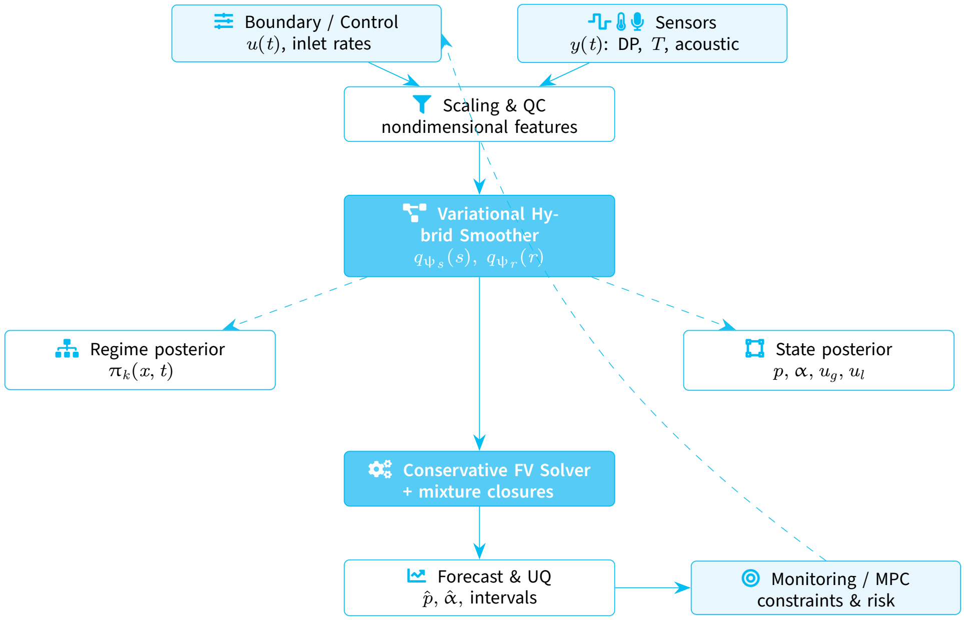

Figure 1. End-to-end pipeline: heterogeneous measurements and boundary inputs feed a variational hybrid smoother that infers regime probabilities and continuous states, drives a conservative finite-volume backbone with regime-mixture closures, and outputs uncertainty-aware forecasts for monitoring and constrained control.

Table 1. Key state variables in the one-dimensional two-fluid backbone.

Symbol	Meaning	Units
$\alpha(x,t)$	Gas volume fraction	–
$\rho_g(p,T)$	Gas density	$\mathrm{kg\,m^{-3}}$
$\rho_l(p,T)$	Liquid density	$\mathrm{kg\,m^{-3}}$
$u_g(x,t)$	Gas axial velocity	$\mathrm{m\,s^{-1}}$
$u_l(x,t)$	Liquid axial velocity	$\mathrm{m\,s^{-1}}$
$p(x,t)$	Mixture pressure	Pa
$\theta(x)$	Pipe inclination angle	rad

This paper takes a different stance on the coupling by posing two-phase prediction as a hybrid inference problem. The flow regime is treated as a latent discrete field that mediates between a conservation-law backbone and regime-conditioned closure operators. Rather than learning a direct mapping from measured features to regime labels, the regime posterior is inferred jointly with continuous states by requiring consistency with measurements, with the discretized conservation laws, and with physically interpretable constraints on closure behavior. The methodological goal is to preserve the advantages of mechanistic models where they are reliable, while placing learning capacity precisely where uncertainty is highest: in the closure relations and in the transition dynamics between regimes.

The core technical contribution is a regime-conditioned operator-learning formulation for closures, coupled to a variational hybrid smoother that infers regime probabilities and continuous states simultaneously [5]. Closures are parameterized as neural operators that act on local, physically scaled neighborhoods of the state and geometry, enabling generalization across pipe diameters, inclinations, and fluid properties without forcing the model into a purely dimensionless, tabulated form. Hybrid inference is performed through an evidence lower bound that combines measurement likelihood with a soft enforcement of PDE residuals produced by a monotone finite-volume scheme. Discrete regimes are handled via a relaxed categorical representation during training and then projected back to discrete sequences for deployment, allowing gradient-based learning while still producing interpretable regime assignments.

A second contribution is the imposition of consistency constraints that encode asymptotic limits and energetic admissibility. Two-phase models can easily yield nonphysical predictions when data-driven components are unconstrained, such as negative wall shear, spurious energy generation, or closure terms that fail to recover single-phase limits. We introduce differentiable

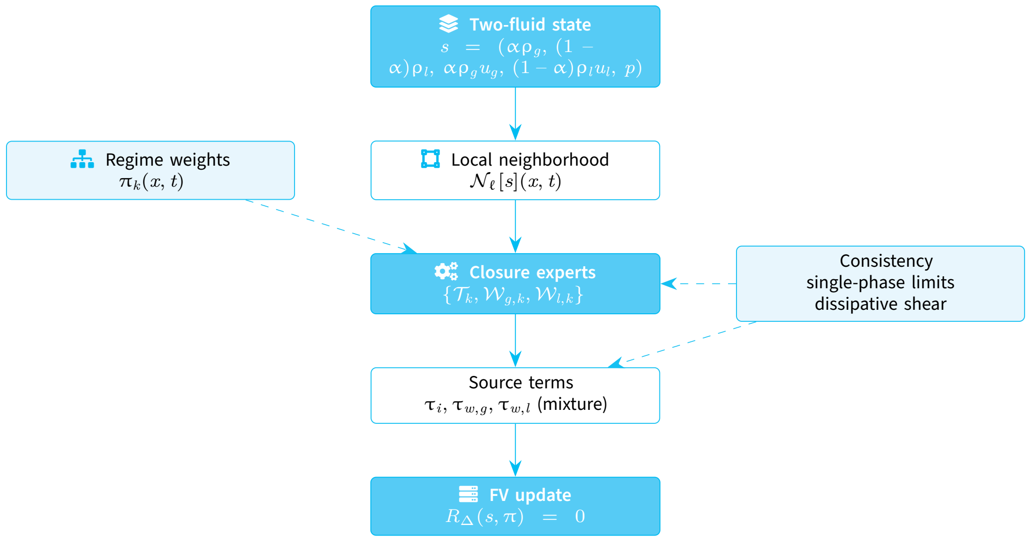

Figure 2. Regime-conditioned closure coupling: neighborhood features and regime probabilities gate learned closure operators, generating wall and interfacial shear terms that enter the conservative finite-volume backbone while enforcing asymptotic single-phase limits and nonnegative dissipation.

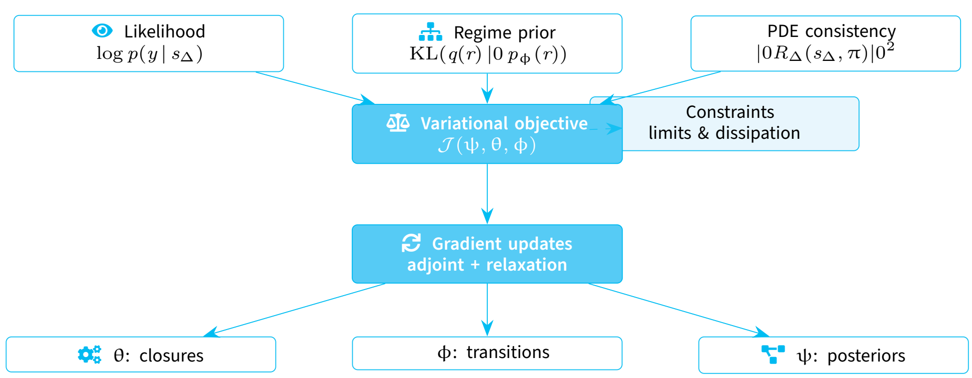

Figure 3. Learning objective decomposition: measurement fit, regime-transition regularization, and discrete-PDE residual consistency combine with physically motivated constraints; adjoint-based gradients support joint optimization of closure operators, transition dynamics, and variational posteriors.

constraints that enforce correct behavior as the void fraction approaches 0 or 1, guarantee nonnegative dissipation from frictional terms, and impose weak invariance to Galilean shifts and reference pressure. These constraints are not merely regularizers; they act as structural priors that tighten the feasible set of closures and improve stability of the hybrid solver.

A third contribution is an analysis of stability and identifiability for the regime-conditioned hybrid system [6]. We show that, under Lipschitz and dissipation constraints on the learned closures, the coupled PDE and inference dynamics satisfy a contraction property in a weighted metric for the continuous state, and that regime posteriors concentrate when measurement operators satisfy a mild persistent-excitation condition. The analysis is not intended to provide worst-case guarantees for all flow conditions, but it clarifies which design choices contribute to stable learning and prediction, and it guides the numerical settings used in the computational study.

The remainder of the paper proceeds by formalizing the hybrid modeling setting, introducing the regime-conditioned two-fluid backbone, presenting the variational inference and operator-learning algorithm, analyzing stability and identifiability, and evaluating the method on controlled experiments that emulate laboratory-style transients and sparse sensing. Throughout, the emphasis is on coupling regime reasoning to prediction and control objectives while preserving physical structure, rather than on optimizing regime classification accuracy as a standalone metric.

2. Hybrid Two-Phase Flow as a Stochastic Switching System

In many conduit-flow applications the governing equations are derived from volume averaging over a control volume that is large compared to microstructure yet small compared to the length scales of interest. The av-

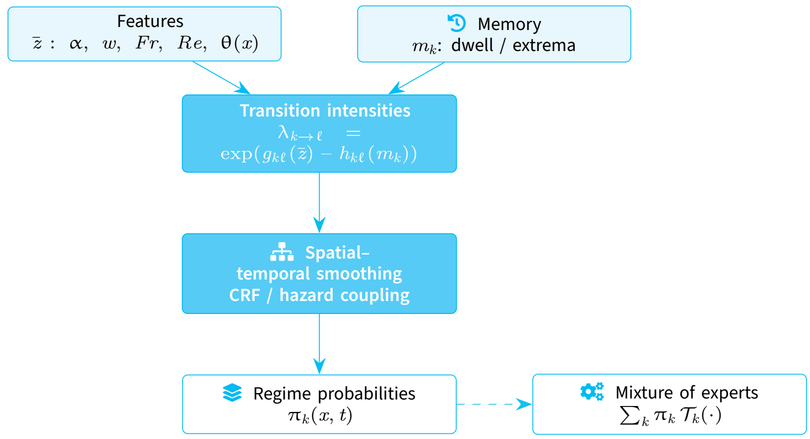

Figure 4. Stochastic switching representation: state-derived features and a hysteresis-aware memory variable drive regime transition intensities, which are regularized by spatial–temporal coherence to produce regime probabilities used to blend regime-conditioned closure operators.

Table 2. Illustrative latent flow regimes and qualitative closure characteristics.

Index k	Regime name	Interface topology	Closure emphasis
1	Stratified / wavy	Layered phases with smooth interface	Wall shear, hydrostatic head
2	Intermittent / slug	Large liquid slugs separated by gas pockets	Strong interfacial shear, acceleration
3	Annular–mist	Liquid film with gas core	Film friction, entrainment, slip
4	Dispersed bubbly	Bubbles in continuous liquid	Turbulent dispersion, drag
5	Transitional / mixed	Local mixture of neighboring patterns	Blended closures, high uncertainty

eraged state typically includes a phase fraction, phase velocities, and thermodynamic variables such as pressure and temperature. The averaging introduces unclosed terms: interfacial shear, wall friction for each phase, virtual mass, and terms capturing slip, entrainment, and turbulence modulation. A standard modeling strategy is to assume that these closures depend on the local regime, where regime is an equivalence class of interface topology and phase distribution [7]. From a systems viewpoint, this means the model is hybrid: continuous dynamics depend on a discrete mode, and the mode transition law depends on the continuous state history.

A central practical difficulty is that regime is rarely observed directly. In some facilities, cameras or high-speed imaging provide labels, but more commonly operators rely on indirect proxies, such as pressure fluctuations, differential pressure across known restrictions, or acoustic signatures. These signals are affected by sensor placement, sampling, and noise, and the same regime label can correspond to different macroscopic signatures depending on geometry and fluid properties. Moreover, regime boundaries are not sharp in the presence of transients: the interface topology may evolve gradually, and hysteresis can cause different transition thresholds depending on whether superficial velocities are increasing or decreasing. These features motivate a probabilistic regime representation rather than a point estimate.

We model the regime as a discrete random field over space and time, denoted by $r(x,t) \in \{1,\ldots,K\}$, with a transition prior that encourages piecewise constancy while allowing localized switching. Unlike a memoryless Markov chain, the transition intensity is allowed to depend on a filtered history of the continuous state, enabling hysteresis and dwell-time effects. The continuous state $s(x,t)$ follows a conservation-law backbone with fluxes and source terms that depend on r through closure operators [8]. Observations y arise from applying a measurement operator to s and adding noise that can be non-Gaussian, reflecting outliers and sensor drift. The inverse problem is to infer s and r from y while learning closure operators from data.

The coupling between inference and learning is subtle because closures influence both the predicted continuous states and the implied regime sequence. If regime is treated as observed during training, closure learning reduces to a conditional regression problem, but labels are often unavailable or inconsistent across data sources. If regime is treated as purely latent without constraints, the model can become unidentifiable: multiple combinations of regime assignments and closure parameters can explain the same measurements. Addressing this requires structural priors that tie closures to physically meaningful features, and numerical choices that enforce conservation and dissipation.

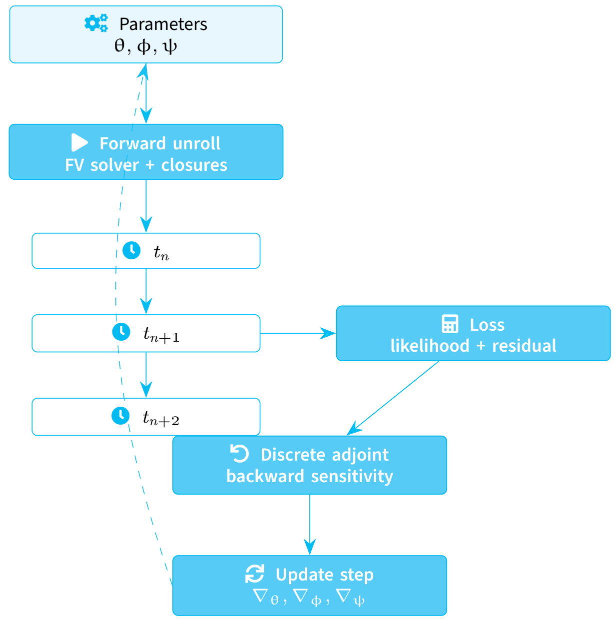

Figure 5. Differentiable solver loop: finite-volume forward unrolling produces predictions used to evaluate measurement and PDE-consistency terms; a discrete adjoint computes sensitivities for joint updates of closure, transition, and inference parameters.

Table 3. Hybrid model hierarchy combining mechanistic and learned components.

Level	Object	Role	Examples
Mechanistic	Two-fluid backbone	Enforce conservation	Mass and momentum balances
Closure	Neural operators	Map local state to source terms	Wall and interfacial shear
Regime	Discrete field r	Select regime-conditioned closures	Stratified, slug, annular
Observation	Operators H_j	Map state to sensor outputs	Pressure, acoustic, imaging
Prior	Transition model	Encode switching statistics	Hysteresis, dwell-time bias

We adopt a probabilistic formulation that retains interpretability at three levels. At the mechanistic level, the backbone conservation laws remain explicit and are solved with a conservative discretization [9]. At the closure level, the unknown terms are represented by operators whose inputs are nondimensional groups and local state neighborhoods, enabling extrapolation in geometry and fluid properties. At the regime level, the discrete field is represented by a transition model whose parameters can be learned but are regularized to produce smooth, hysteretic regime posteriors rather than rapidly oscillating labels. The inference output is therefore a distribution over regimes and continuous trajectories, not just a single classification.

An important design choice is the interface between the discrete regime and the continuous closures. In classical switching systems, each discrete mode selects a fixed set of closure formulas. In a learned setting, this can be too rigid: regimes may not correspond to a single closure family, and transitions can be gradual. We therefore allow closures to be expressed as mixtures of regime-conditioned experts, weighted by the regime posterior. During deployment, the mixture can be collapsed to a maximum a posteriori regime if a discrete decision is required, but for prediction and control it can be advantageous to propagate the mixture distribution to quantify model uncertainty near regime boundaries [10].

Another practical consideration is computational tractability. Full Bayesian inference for a hybrid PDE system is expensive. We use a variational approximation that factorizes the posterior into a continuous-state component and a regime component, coupled through expectations of closure terms and PDE residuals. This approximation supports gradient-based learning with differentiable solvers and is compatible with real-time deployment through amortized inference networks. The remainder of the paper formalizes these choices and derives a learning and inference algorithm that can be implemented efficiently while maintaining physical constraints.

A useful abstraction is to view the regime field as a switching signal driven by a hazard function rather than as a deterministic partition of the superficial-velocity

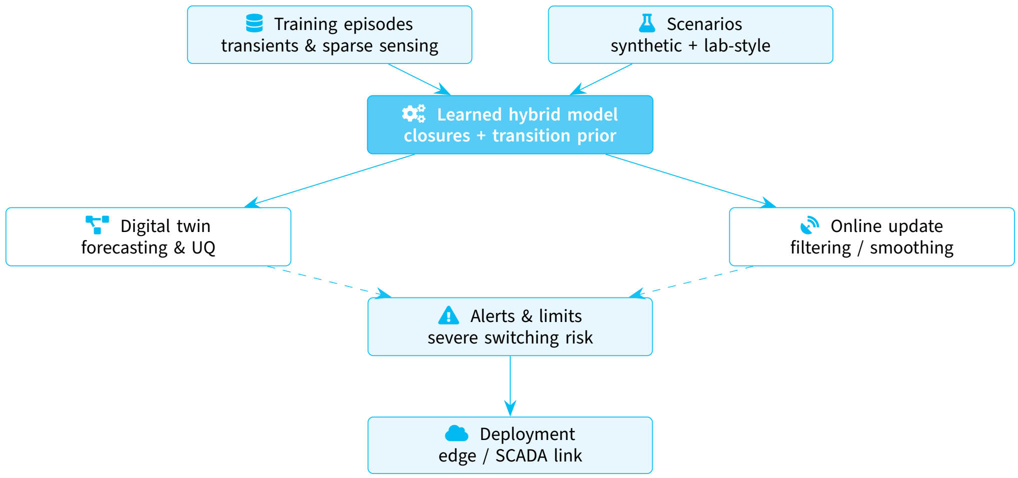

Figure 6. Workflow from training to operations: regime-conditioned closures and transition dynamics learned from transient episodes enable a deployable digital-twin pipeline that performs online state/regime updates, propagates uncertainty, and supports alarm logic near unstable switching regions.

Table 4. Closure operators and associated physical consistency constraints.

Quantity	Operator	Regime dependence	Key constraints
Gas wall shear $\tau_{w,g}$	$\mathcal{W}_{g,k}$	Yes, via k	Nonnegative dissipation, single-phase limit
Liquid wall shear $\tau_{w,l}$	$\mathcal{W}_{l,k}$	Yes, via k	Nonnegative dissipation, single-phase limit
Interfacial shear τ_i	$\mathcal{T}_k$	Yes, via k	$\tau_i w \leq 0$, vanishes as $\alpha \to 0, 1$
Mass transfer Γ_g	$\mathcal{M}_k$	Optional	Small over operational time scales
Pressure closure	$\mathcal{P}$	Weak	Correct compressibility, invariance to shifts
Mixture sound speed c_m	$\mathcal{C}$	Weak	Convex in phase sound speeds

plane. Let $q_k(x, t)$ denote the probability that the local regime equals k. A state-dependent continuous-time transition model can be written through intensities $\lambda_{k\to\ell}(x, t)$ that depend on filtered features $\bar{z}(x, t)$, yielding

$$\partial_t q_k = \sum_{\ell \neq k} q_\ell \, \lambda_{\ell \to k} \quad - q_k \sum_{\ell \neq k} \lambda_{k \to \ell},$$
$$\lambda_{k\to\ell} = \exp\Big(g_{k\ell}(\bar{z}) \quad - h_{k\ell}(m_k)\Big), \tag{1}$$

where m_k is a memory variable tracking dwell time or recent extrema of selected features. The term $h_{k\ell}$ implements hysteresis by suppressing transitions that would violate a minimum dwell time or that contradict the direction of change in superficial velocities. In discrete time, the same effect is captured by edge potentials in the conditional random field, but the hazard formulation highlights that regime switching is not instantaneous and that a probabilistic representation can encode this finite transition bandwidth [11].

The observation operator H_j can be highly nonlinear even when the PDE state is simple. Differential pressure sensors measure path integrals of $\partial_x p$ that are affected by hydrostatic head and by local accelerations, while distributed acoustic sensing produces a proxy for strain rate that couples to pressure fluctuations through pipe elasticity. To avoid committing to a single sensing modality, we treat H_j as a composition of a physics-based forward map and a low-capacity calibration component. For example, a differential pressure measurement between x_a and x_b is modeled as an integral of the predicted pressure gradient plus an affine drift term whose parameters are learned jointly with closures but regularized to remain small. This separation reduces the tendency of the model to explain sensor drift by distorting closure behavior.

The stochastic switching viewpoint also clarifies how regime uncertainty should be propagated. If the goal is prediction under future control actions, the regime transition model must extrapolate under unseen inputs. A purely discriminative classifier trained on historical labels does not provide a generative transition mechanism, and it cannot represent uncertainty away from its training distribution [12]. By learning a transition prior $p_\phi(r)$ that depends on physically scaled features and by coupling it to the solver through residual consistency, the hybrid approach provides a mechanism to forecast regime probabilities under hypothetical control moves, which is essential for robust optimization.

Table 5. Variational factors and prior structures used in hybrid inference.

Component	Approximate posterior $q(\cdot)$	Prior structure
State trajectory s_Δ	Gaussian in reduced latent space	Weakly informative, bounded energy
Regime field r	Relaxed categorical CRF	Spatial–temporal smoothness, hysteresis
Closures θ	Point estimate with regularization	Limits, Lipschitz and dissipation penalties
Transition params ϕ	Point estimate	Smooth hazard functions, dwell-time bias
Noise scales (σ_j)	Independent log-normal	Sensor-specific hyperpriors
Drift terms in H_j	Gaussian	Zero-mean, small variance

Table 6. Representative sensor models and observation operators.

Sensor type	Forward map H_j	Noise model	Notes
Static pressure	Pointwise $p(x_j, t)$	Student-$t(\nu_j, \sigma_j)$	Sensitive to hydrostatic head
Differential pressure	Path integral of $\partial_x p$	Student-t	Includes gravity and acceleration
Distributed acoustic	Function of strain-rate proxy	Heavy-tailed	Indirect pressure information
Holdup probe	Nonlinear function of α	Gaussian / censored	Occasional regime labels
Imaging / camera	Feature extractor on frames	Approx. Gaussian	Sparse but high information

3. Regime-Conditioned Two-Fluid Backbone with Consistency Constraints

We consider a one-dimensional conduit of axial coordinate $x \in [0, L]$ with inclination $\theta(x)$, diameter $D(x)$, and cross-sectional area $A(x)$. The gas and liquid occupy fractions $\alpha(x,t)$ and $1-\alpha(x,t)$ of the cross-section, with phase densities $\rho_g(p,T)$ and $\rho_l(p,T)$ and axial velocities $u_g(x,t)$ and $u_l(x,t)$. Let $p(x,t)$ denote a common pressure field at the averaging scale. The conservative state is $s = (\alpha\rho_g,\ (1-\alpha)\rho_l,\ \alpha\rho_g u_g,\ (1-\alpha)\rho_l u_l,\ p)^\top$, though in practice the thermodynamic closure may reduce the dimension when temperature is measured or assumed constant.

The backbone dynamics follow a compressible two-fluid balance with source terms capturing gravity and interfacial exchange. In differential form,

$$\begin{aligned}
\partial_t(\alpha\rho_g) + \partial_x(\alpha\rho_g u_g) &= \Gamma_g, \\
\partial_t((1-\alpha)\rho_l) + \partial_x((1-\alpha)\rho_l u_l) &= -\Gamma_g, \\
\partial_t(\alpha\rho_g u_g) + \partial_x(\alpha\rho_g u_g^2) + \alpha\,\partial_x p &= S_g, \\
\partial_t((1-\alpha)\rho_l u_l) + \partial_x((1-\alpha)\rho_l u_l^2) + (1-\alpha)\,\partial_x p &= S_l.
\end{aligned} \tag{2}$$

The phase source terms S_g and S_l include gravity, wall friction, and interfacial momentum exchange [13]. They are written as

$$\begin{aligned}
S_g &= -\alpha\rho_g g\sin\theta \quad -\tau_{w,g} \quad +\tau_i \quad +\Gamma_g u_{i,g}, \\
S_l &= -(1-\alpha)\rho_l g\sin\theta \quad -\tau_{w,l} \quad -\tau_i \quad -\Gamma_g u_{i,l}.
\end{aligned} \tag{3}$$

Here $\tau_{w,k}$ denotes wall shear contribution per unit axial length for phase k, τ_i denotes interfacial shear contribution (positive when gas gains momentum from the interface), and $u_{i,k}$ denote velocities associated with interfacial mass transfer, included for completeness even though many oil-and-gas applications treat $\Gamma_g = 0$ over the relevant time scales. Additional terms such as virtual mass and turbulent dispersion can be incorporated but are omitted here to focus on the closure-learning structure.

The unknown components are the closure relations for $\tau_{w,g}$, $\tau_{w,l}$, τ_i, and possibly for the slip relationship embedded in the momentum exchange. We represent these closures as regime-conditioned operators. Let $r(x,t)$ denote the regime, and define a vector of local features $z(x,t)$ consisting of nondimensional groups and geometry, for example mixture Reynolds numbers, Froude numbers, density and viscosity ratios, and inclination. Because closure terms depend on local gradients and not only on pointwise state, we define a neighborhood operator $\mathcal{N}_\ell[s](x,t)$ that extracts a window of the state over a length scale ℓ around x, scaled by local diameter, and a similar temporal filter capturing short-term history. The closure operator for interfacial shear is then written as

$$\begin{aligned}
\tau_i(x,t) &= \sum_{k=1}^{K} \pi_k(x,t)\, \mathcal{T}_k(\mathcal{N}_\ell[s](x,t),\ z(x,t)), \\
\pi_k(x,t) &= \mathbb{P}(r(x,t) = k \mid y_{0:t}),
\end{aligned} \tag{4}$$

where $\mathcal{T}_k$ is a learned neural operator for regime k, and π_k are regime probabilities produced by the inference procedure. Wall shear terms $\tau_{w,g}$ and $\tau_{w,l}$ are treated analogously with operators $\mathcal{W}_{g,k}$ and $\mathcal{W}_{l,k}$. This mixture formulation smooths the effect of regime uncertainty on the PDE source terms, improving numerical stability near transitions and enabling a consistent probabilistic interpretation.

Physical consistency imposes strong constraints on admissible closures. First, single-phase limits must be

Table 7. Comparison of training and deployment configurations.

Aspect	Training setting	Deployment setting
Inference mode	Batch smoothing over full trajectories	Online filtering / prediction
Regime representation	Relaxed probabilities with entropy regularization	Probabilities and MAP labels
Solver unroll horizon	Increasing curriculum horizon	Fixed horizon per control step
PDE residual weight λ	Annealed from small to large	Fixed, tuned from training
Access to future data	Yes, for all time indices	No, strictly causal

Table 8. Numerical scheme and discretization choices for the backbone solver.

Item	Setting	Comment
Spatial grid Δx	Uniform finite volumes	Resolution matched to sensing density
Time step Δt	CFL-based, explicit	Stability controlled by wave speeds
Numerical flux	Monotone upwind / Godunov	Preserves conservation and bounds
Source treatment	Gravity-balanced	Reduces spurious hydrostatic errors
Residual norm	Weighted ℓ_2 on R_Δ	Balances mass and momentum components

recovered when $\alpha \to 0$ or $\alpha \to 1$ [14]. In these limits, interfacial shear must vanish and wall shear must reduce to a single-phase friction term. We enforce these limits by embedding boundary conditions in the operator parameterization and by imposing penalties on the residual of the limit constraints during learning. Second, frictional contributions should not generate mechanical energy. Define a mechanical energy density e for the mixture that includes kinetic and gravitational potential contributions, and let $E(t) = \int_0^L e(x,t)\,dx$. For closures that represent frictional exchange, a sufficient condition for energetic admissibility is that the work done by wall and interfacial shear is nonpositive. We impose this through inequality constraints on the learned operators that require the dissipation rate to be nonnegative when expressed in terms of slip velocity and wall-relative velocity.

To formalize dissipation, define the slip $w = u_g - u_l$ and mixture velocity $u_m = \alpha u_g + (1-\alpha)u_l$. For frictional closures of the form $\tau_i = c_i\,|w|w$ and $\tau_{w,k} = c_{w,k}\,|u_k|u_k$, dissipation corresponds to $c_i \geq 0$ and $c_{w,k} \geq 0$. In the operator setting, the analogous requirement is that $\tau_i w \leq 0$ when τ_i is the force on the gas and w is the relative velocity [15]. We implement this by reparameterizing operator outputs through nonnegative functions, for example exponentiated or squared latent outputs, and by including soft barriers that penalize violations for sampled states. This encourages global energetic consistency without restricting the model to a single functional form.

Finally, the conservation-law backbone must be solved in a stable way to support gradient-based learning. We discretize the system with a finite-volume method using monotone numerical fluxes and source-term balancing for gravity. The discretization yields a residual operator $R_\Delta(s_\Delta, \pi)$ that vanishes when discrete conservation is satisfied. This residual is used both for forward prediction and as a constraint term in variational inference, linking observed data to physically consistent trajectories.

The backbone can be augmented with a pressure evolution law consistent with compressibility and mixture acoustics. In a common approximation, pressure is determined implicitly by a constraint relating mixture density to phase densities and void fraction, but in transient prediction it is often useful to write a weakly compressible form in which pressure dynamics arise from a mixture bulk modulus [16]. Let $\rho_m = \alpha\rho_g + (1-\alpha)\rho_l$ and let c_m denote an effective mixture sound speed. A simplified pressure closure can be written as

$$\partial_t p + c_m^2\,\partial_x(\rho_m u_m) = \beta_p \quad + \zeta_p,$$
$$c_m^{-2} = \alpha\,c_g^{-2} \quad + (1-\alpha)\,c_l^{-2}, \qquad (5)$$

where β_p collects boundary-driven compressions and ζ_p is a small stabilization term that accounts for numerical compressibility when a strictly incompressible formulation would be ill-conditioned. The purpose of this equation is not to fully model acoustic propagation, but to provide a consistent way to couple pressure to mass and momentum balances in transient regimes where compressibility cannot be neglected.

Consistency constraints must also respect invariances. The governing equations are invariant to adding a constant to pressure and to simultaneous shifts in phase velocities that preserve slip, provided that boundary conditions are shifted consistently. Learned closures that depend directly on absolute pressure can violate these invariances and reduce transferability. We therefore feed closures with pressure gradients and with nondimensional combinations such as $p/(\rho_m u_m^2)$ rather

Table 9. Ablation variants and qualitative behavior.

Variant	Regime handling	Operator structure	Observed behavior
Single-regime model	No discrete r	Shared closure	Biased transients across patterns
Pointwise regressor	With r, no stencil	Local MLP	Grid-sensitive, oscillatory shear
Unconstrained closure	With r, no constraints	Neural operator	Occasional energy growth, instabilities
Full hybrid model	Probabilistic r	Constrained operator	Stable, improved transients and uncertainty

than with raw pressure, and we include invariance penalties that discourage sensitivity to uniform pressure shifts in the training distribution [17].

The discrete residual used in inference is constructed from the finite-volume update. Let s_i^n denote the cell-average state in cell i at time step n, and let $\hat{f}_{i+1/2}^n$ denote a numerical flux. The residual takes the form

$$R_i^n = s_i^{n+1} - s_i^n \quad + \frac{\Delta t}{\Delta x}\Big(\hat{f}_{i+1/2}^n - \hat{f}_{i-1/2}^n\Big) \\ - \Delta t\, \hat{S}_i^n(\pi_i^n, s_i^n; \theta), \tag{6}$$

where $\hat{S}_i^n$ is the discretized source term including closure contributions. The residual is zero for trajectories exactly produced by the solver, but in variational inference q_{ψ_s} can represent distributions over trajectories that deviate slightly from the solver due to measurement constraints. Penalizing $|0R|0^2$ therefore encourages solver consistency while still allowing the posterior to represent uncertainty. The residual term also stabilizes learning because it penalizes closure parameters that lead to stiff, unstable dynamics that cannot match measurements without large residual corrections.

4. Variational Hybrid Inference and Closure Operator Learning

Let y denote the collection of measurements over sensors and times, and let s_Δ denote the discrete state trajectory produced by the finite-volume scheme on a mesh with spacing Δx and time step Δt. The unknown parameters include closure-operator parameters θ and regime-transition parameters ϕ. We seek the posterior $p(s_\Delta, r \mid y, \theta, \phi)$ and ultimately the marginal likelihood of the data under the model. Exact inference is intractable due to the high-dimensional PDE state and the discrete regime field, so we adopt a variational approximation $q_\psi(s_\Delta, r) = q_{\psi_s}(s_\Delta)\, q_{\psi_r}(r)$ with coupling through shared parameters and through expectations of closure terms.

The generative model is specified by three components: an observation model, a discrete transition prior, and a soft enforcement of the discrete PDE [18]. Observations are modeled as

$$y_{j,n} = H_j(s_{\Delta,n}) + \varepsilon_{j,n}, \\ \varepsilon_{j,n} \sim \text{Student-}t(\nu_j, 0, \sigma_j), \tag{7}$$

where H_j maps the discrete state at time index n to the j-th sensor output. A heavy-tailed noise model is used to accommodate sensor spikes and unmodeled disturbances without over-penalizing the solver. The regime prior is represented as a conditional random field with local transition potentials that depend on filtered state features, enabling hysteresis. In a discrete space–time grid, this can be written as

$$p(r) \propto \exp\Big(\sum_{e\in\mathcal{E}} \eta_e(r_e) \quad + \sum_{v\in\mathcal{V}} \xi_v(r_v, \bar{z}_v)\Big), \tag{8}$$

where $\mathcal{V}$ indexes nodes, $\mathcal{E}$ indexes edges connecting neighboring nodes in space and time, η_e encourages smoothness, and ξ_v encodes state-dependent transition preference through features $\bar{z}_v$ computed from s_Δ. The PDE is enforced by treating the finite-volume update as a probabilistic constraint with a small variance, effectively a Gaussian penalty on the residual.

The resulting evidence lower bound is

$$\mathcal{J}(\psi, \theta, \phi) = \mathbb{E}_{q_\psi}\big[\log p(y \mid s_\Delta)\big] \quad - \mathrm{KL}\big(q_{\psi_r}(r)\, |0\, p_\phi(r)\big) \\ - \mathrm{KL}\big(q_{\psi_s}(s_\Delta)\, |0\, p(s_\Delta)\big) \quad - \lambda\, \mathbb{E}_{q_\psi}\big[|0 R_\Delta(s_\Delta, \pi)|0^2\big], \tag{9}$$

where π denotes regime probabilities induced by q_{ψ_r}, $p(s_\Delta)$ is a weak prior over initial conditions and boundedness, and λ controls the strength of the PDE consistency. The penalty term is not a substitute for a conservative solver; rather it encourages the variational state distribution to concentrate on solver-consistent trajectories, which is essential when measurements are sparse and do not uniquely determine all states [19].

A key obstacle is differentiating through the discrete regime. We use a relaxed categorical representation in which q_{ψ_r} produces logits $\ell_k(x, t)$ and corresponding probabilities π_k through a temperature-controlled softmax. During training, samples are generated via a continuous relaxation so that gradients can propagate to ψ_r and to closure parameters θ through the mixture source terms. As temperature decreases, the relaxation approaches a discrete distribution, and during deployment a discrete regime map can be extracted by taking the maximum probability at each grid point or by decoding the most likely sequence under the transition model.

The closure operators $\mathcal{T}_k$ and $\mathcal{W}_{\cdot,k}$ are implemented as neural operators that map local state neighborhoods to closure outputs while respecting physical scaling. Rather than a pointwise multilayer perceptron, the operator acts on a short stencil of states, enabling sensitivity to gradients and local wave structure. Inputs are nondimensionalized using local mixture scales, and outputs are produced in a form that respects units, for example by predicting dimensionless friction factors that are then scaled by dynamic pressure. Consistency constraints are integrated in two ways: by design, through reparameterizations that enforce sign and limit properties, and by penalties that activate when sampled trajectories violate asymptotic or energetic constraints.

Learning proceeds by joint optimization of ψ, θ, ϕ using stochastic gradients of $\mathcal{J}$. The gradients require differentiating through the finite-volume solver and through the operator networks. We employ discrete adjoints for the solver to compute sensitivity of the residual and of predicted sensor outputs with respect to closure parameters, while treating the regime probabilities as differentiable weights [20]. For stability, the solver is unrolled for a fixed horizon during training, and longer horizons are incorporated through curriculum-style expansion of the unroll length without introducing explicit stagewise lists in the formulation. This reduces gradient explosion while still training the model to capture slow transients.

Because the variational approximation factorizes s_Δ and r, coupling is achieved through expectation terms: the closure source terms depend on π, and the regime transition potentials depend on filtered features derived from s_Δ. In practice, we implement this coupling by alternating between updating ψ_s to better fit the measurements under the current closures and regime posterior, and updating ψ_r to better explain the residual and measurement innovations under the current state estimate. This alternating structure resembles expectation–maximization but remains fully differentiable and is performed within standard gradient-based optimization.

An important benefit of the variational formulation is calibrated uncertainty. Near regime boundaries, multiple regime assignments can explain the same measurements, and the closure operators for different regimes can yield similar source terms. The posterior π captures this ambiguity, and the mixture closure propagates it to predictive intervals for pressure and holdup [21]. For control and monitoring, these intervals are often more informative than a single regime label because they quantify how confidently the model anticipates a transition under anticipated actuator changes.

The neural-operator parameterization is designed to represent closure dependencies on local fields without committing to a fixed stencil size. For a generic closure $c(x)$ computed from a local input field $f(x)$ derived from s, a kernel-based operator can be written as

$$
\begin{aligned}
(\mathcal{O}_k f)(x) &= a_k(f(x)) \quad + \int_{x-\ell}^{x+\ell} K_k(x, x')\, b_k(f(x'))\, dx', \\
c(x) &= \sigma\big((\mathcal{O}_k f)(x)\big),
\end{aligned}
\tag{10}
$$

where a_k and b_k are pointwise nonlinear maps, K_k is a learned kernel, and σ is an output activation chosen to enforce sign or boundedness. In practice, the integral is approximated by a quadrature over the local neighborhood, and the kernel is parameterized in a low-rank form so that computation scales linearly with the number of grid points. This representation can emulate classical local closures when the kernel collapses to a delta-like form, but it can also represent nonlocal effects such as entrainment memory through wider kernels.

Inference over s_Δ is performed in a smoothing rather than filtering mode in training, because full trajectories are available and the goal is to learn closures that explain entire transients. We represent $q_{\psi_s}(s_\Delta)$ as a Gaussian in a reduced latent space obtained by projecting the state onto physically interpretable coordinates, such as pressure and mixture velocity, and we reconstruct full phase velocities and void fraction through a decoder that respects bounds. This reduces dimensionality and encourages the posterior to remain in admissible regions [22]. During deployment, the same structure supports online updates by conditioning on new measurements and propagating the state forward with the solver.

A practical challenge is the interaction between the PDE residual penalty and the measurement likelihood. If λ is too large, the posterior collapses to trajectories that the solver can produce even when the solver is biased; if λ is too small, the posterior can drift to measurement-fitting trajectories that violate conservation and yield closures that compensate for nonphysical states. We address this by scaling λ with an estimate of discretization error and by annealing λ during training so that early optimization focuses on matching measurements and later optimization enforces solver consistency. This annealing is implemented as a smooth schedule, avoiding discrete stages that could create brittle behavior.

Adjoint-based gradients are computed by differentiating the discretized update with respect to closure parameters. Let θ denote all closure parameters and let $\mathcal{L}$ denote the negative variational objective. The discrete adjoint λ_i^n satisfies a backward recursion of the form

$$
\begin{aligned}
\lambda_i^n &= \left(\partial_{s_i^n} F_\Delta\right)^\top \lambda_i^{n+1} \quad + \partial_{s_i^n} \mathcal{L}_n, \\
\partial_\theta \mathcal{L} &= \sum_{n,i} \left(\partial_\theta F_\Delta\right)^\top \lambda_i^{n+1} \quad + \partial_\theta \mathcal{L}_n,
\end{aligned}
\tag{11}
$$

where $\mathcal{L}_n$ collects contributions from measurements, residual penalties, and constraint penalties at time n. The dependence of F_Δ on θ occurs through the discretized sources $\hat{S}$ and, indirectly, through regime probabilities π when π is produced by an inference network that depends on solver states. In our implementation, we detach certain feedback paths early in training to reduce gradient variance, and then reintroduce them as the solver and closures become stable, which improves convergence without changing the mathematical objective [23].

The regime posterior network q_{ψ_r} receives two types of inputs: local physically scaled features derived from the current state estimate and innovation signals derived from mismatches between predicted and observed measurements. Innovation inputs are important because they provide a mechanism to assign regimes based on dynamic consistency rather than on instantaneous state alone. The network outputs logits for each regime at each grid point, which are then smoothed through a differentiable message-passing layer implementing the conditional random field potentials. This yields regime probabilities that respect both local evidence and spatial–temporal coherence.

To mitigate posterior collapse where all regimes become equally likely or where the model selects a single regime everywhere, we include an entropy regularization that encourages moderate uncertainty in regions with weak evidence and lower uncertainty where evidence is strong. This regularization is applied adaptively by measuring the local sensitivity of predicted measurements to regime changes; when sensitivity is low, higher entropy is tolerated, and when sensitivity is high, the model is encouraged to commit. The result is a regime posterior that is interpretable and stable under sensor noise while still responsive to genuine regime transitions.

5. Stability, Identifiability, and Error Propagation

Hybrid PDE learning can fail due to numerical instability, posterior collapse, or non-identifiability of closures. We analyze these failure modes through a simplified setting that retains the essential coupling: a one-dimensional conservation law with regime-conditioned source term, solved with a monotone scheme, with sparse noisy observations of a linear functional of the state [24]. The analysis clarifies how dissipation constraints and transition regularization contribute to stable learning and inference.

Consider the discrete update $s_{n+1} = F_\Delta(s_n, \pi_n; \theta)$, where F_Δ is the finite-volume step including numerical fluxes and source terms, and π_n are regime probabilities at time n. Assume the numerical flux is monotone and the source term is Lipschitz in s_n for fixed π_n, with Lipschitz constant L_s. The learned closures influence L_s, and unconstrained networks can lead to large local gains that destabilize the solver and amplify inference errors. The energetic constraints described earlier limit the effective gain by ensuring that frictional terms oppose motion and thus act as damping.

To make this precise, define a weighted norm $|0s|0_P^2 = s^\top P s$ with positive definite P chosen to balance mass and momentum components. Under a linearization of the update, stability corresponds to the spectral radius of the Jacobian $J = \partial F_\Delta / \partial s$ being less than one in the induced norm. Dissipation constraints yield a negative semidefinite contribution to the Jacobian from frictional sources. A sufficient condition for contraction is that the damping dominates the transport-induced amplification over a time step [25]. In the regime-mixture setting, this condition can be enforced uniformly by requiring each regime-conditioned closure to satisfy dissipation and Lipschitz bounds, because π_n are convex weights.

We formalize this by writing the energy inequality at the discrete level. Let E_n denote a discrete mechanical energy, and let $\mathcal{D}_n$ denote discrete dissipation due to wall and interfacial shear. For the finite-volume scheme with balanced sources, one can derive

$$
\begin{aligned}
E_{n+1} - E_n &\le -\Delta t\,\mathcal{D}_n \quad + \Delta t\,\mathcal{W}_n \quad + \Delta t\,\mathcal{N}_n, \\
\mathcal{D}_n &= \sum_i \Big(\tau_{w,g,i}\, u_{g,i} \quad + \tau_{w,l,i}\, u_{l,i} \quad - \tau_{i,i}\, w_i \Big),
\end{aligned}
\tag{12}
$$

where $\mathcal{W}_n$ captures work done by boundary conditions and pumps, and $\mathcal{N}_n$ collects numerical dissipation and truncation terms. With the sign conventions adopted earlier, energetic admissibility implies $\mathcal{D}_n \ge 0$ for all admissible states. The closure constraints enforce this inequality in expectation under the variational distribution, which in turn prevents the learned sources from injecting energy in typical operating regions. While this does not guarantee global stability, it mitigates a common failure mode in data-driven closures.

Identifiability depends on the richness of measurements and on excitation of the regime-conditioned dynamics. Because regimes affect dynamics through closures, regimes become identifiable when different regimes induce distinguishable innovations in measured outputs under similar inputs. Let $y_n = H s_n + \varepsilon_n$ for a linearized measurement operator H [26]. Under mild assumptions on the observation noise and on the boundedness of J, the log-likelihood ratio between two regime assignments accumulates over time as the sum of squared innovation differences. If H observes a component that is sensitive to closure differences, such as pressure gradient, and if the flow experiences intervals where regimes differ in their implied shear, then the posterior over regimes concentrates. Conversely, if measurements only observe slowly varying quantities, regime posteriors remain diffuse, and learning can collapse by absorbing regime effects into closure parameters.

The variational objective addresses this by coupling regime inference to PDE residuals and by regularizing transition dynamics. The residual term penalizes regime assignments that require large correction forces to reconcile the solver with measurements. Transition regularization discourages rapid switching that would otherwise allow the model to fit noise by alternating regimes at high frequency. Together with the mixture closure representation, this yields a bias toward physically plausible regimes when data are weak, while still allowing transitions when innovations persist.

Error propagation in prediction arises from three sources: discretization error of the solver, approximation error of closures, and posterior uncertainty in regimes and initial conditions [27]. In our setting, closure approximation error acts as a perturbation to the source term. Under the contraction condition, the effect of perturbations decays over time, and prediction error bounds can be expressed in terms of the magnitude of the perturbation and the contraction rate. When contraction is weak, perturbations can persist and interact with regime uncertainty, leading to broader predictive intervals. This analysis motivates maintaining a controlled amount of damping through closure constraints and choosing a numerical scheme with sufficient monotonicity for the learning phase, even if higher-order schemes are used later for refined prediction.

Finally, the hybrid nature introduces a subtle source of error: regime-boundary sensitivity. Small errors in state can cause large changes in regime if regime is treated as a hard threshold. By maintaining regime probabilities and mixture closures, the model smooths this sensitivity and replaces discontinuous switching with a continuous interpolation in the source terms. This reduces gradient variance during training and reduces forecast volatility during deployment, at the cost of a slight blurring of regime boundaries [28]. When a discrete regime output is required, it can be extracted after smoothing, which tends to produce more stable decisions than hard switching inside the solver.

The mixture closure representation also admits a useful bound on the effect of regime uncertainty. Let $S(s, \pi; \theta)$ denote the source term in the continuous dynamics, and suppose each regime-conditioned source $S_k(s; \theta)$ is Lipschitz with constant L_k in a neighborhood of interest. Because $S(s, \pi; \theta) = \sum_k \pi_k S_k(s; \theta)$, it follows that S is Lipschitz with constant bounded by $\sum_k \pi_k L_k$, and thus by $\max_k L_k$. This convexity prevents the source term from exhibiting larger local gains under uncertainty than the worst-case single-regime source, provided the operators are individually controlled. This observation motivates enforcing Lipschitz control per regime, for example through spectral normalization in operator layers or through explicit gradient penalties computed on sampled trajectories.

We also analyze the sensitivity of the regime posterior to measurement perturbations. Under the conditional random field prior, the regime posterior is log-concave in the logits when the pairwise potentials are submodular, which is approximately satisfied when smoothness penalties dominate [29]. The relaxed inference therefore has a form of stability: small changes in local evidence lead to bounded changes in posterior probabilities. This is not true for hard argmax decoding, which can flip regimes discontinuously. In training, posterior stability reduces gradient variance and leads to more reproducible closure learning. In deployment, it prevents chattering of regime labels that can destabilize control decisions.

A useful quantitative statement can be derived for a simplified linearized model where measurement likelihoods are Gaussian and the transition prior is a quadratic penalty on differences of logits. In this case, the posterior logits solve a linear system whose condition number is controlled by the smoothness penalty and by the measurement information. The resulting bound implies that increasing smoothness improves posterior stability but can oversmooth true transitions. The proposed adaptive entropy regularization partially resolves this trade-off by allowing sharper posteriors when evidence is strong [30].

Error bounds for the continuous state can be derived in the presence of closure approximation. Let $s^\star$ denote the trajectory under the true closures and s denote the trajectory under learned closures, both driven by the same inputs and initial conditions. Let δ_n denote the discrete source error induced by closure approximation at time step n. Under contraction with rate $\kappa < 1$, one obtains

$$\begin{aligned} |0 s_{n+1} - s^\star_{n+1}|0_P &\le \kappa\, |0 s_n - s^\star_n|0_P \quad + c_\Delta\, |0 \delta_n|0, \\ |0 s_n - s^\star_n|0_P &\le \kappa^n\, |0 s_0 - s^\star_0|0_P \quad + \frac{c_\Delta}{1-\kappa} \max_{m<n} |0 \delta_m|0, \end{aligned} \tag{13}$$

where c_Δ depends on discretization parameters and on the chosen norm. This bound emphasizes that long-horizon prediction quality depends on maintaining both solver stability and bounded closure error. The operator-learning strategy reduces $\max |0\delta_m|0$ in the training distribution, while dissipation constraints and monotone fluxes promote a smaller κ. In regimes where contraction is weak, the bound is loose, and the posterior uncertainty is expected to remain broad, which is consistent with the probabilistic outputs observed in the computational study [31].

Finally, the coupled learning problem exhibits a form of implicit regularization because closures that produce unstable trajectories tend to yield large residual penalties and thus are disfavored by the variational objective. This creates a selection bias toward stable closures even when multiple closures fit measurements equally well over short horizons. While this bias can be beneficial, it can also suppress rare but physically real unstable behaviors if they are underrepresented in training data. This suggests that, in practice, training data should include representative transients and that the transition prior should not exclude rare regimes a priori.

6. Computational Evaluation

We evaluate the proposed regime-conditioned variational operator-learning framework in a sequence of computational studies designed to stress the hybrid coupling and the physical constraints. The emphasis is not on matching a particular facility dataset, but on testing whether the method can learn closures and infer regime probabilities from sparse observations while producing stable, physically admissible predictions over transients. The studies use a one-dimensional conduit model with variable inclination and diameter to emulate a range of operating conditions.

The first study uses synthetic data generated from a high-resolution two-fluid model with known regime-dependent closure parameters and a prescribed regime-transition law with hysteresis [32]. The purpose is to test identifiability and to quantify how well the inference procedure recovers the latent regime field when only pressure measurements are available. Pressure sensors are placed sparsely along the pipe, and measurements include heavy-tailed noise and occasional outliers. The proposed method learns closure operators that approximate the ground-truth closures in the operating region while maintaining single-phase limits and dissipation. Regime posteriors concentrate in regions where the synthetic transition law induces distinct shear signatures, and remain diffuse near equilibrium regions where regimes are dynamically similar. Compared to a baseline that performs regime classification from instantaneous features and then simulates with fixed closures, the hybrid method yields smoother regime probabilities and reduces spurious switching under noise.

The second study introduces model mismatch by generating data from a higher-fidelity simulator that includes additional effects such as virtual mass and dispersed-phase turbulence, while the learning model excludes these terms. This setting tests whether learned closures can absorb unmodeled physics without violating constraints. The results indicate that the consistency constraints prevent the learned closures from compensating for mismatch by producing nonphysical shear [33]. Instead, the learned operators adjust in regions where the omitted physics effectively alters momentum exchange, but they remain bounded and dissipative. Predictive pressure gradients match the synthetic measurements with reduced bias, and the inferred regime probabilities reflect increased uncertainty near transition zones, which is consistent with the presence of unmodeled dynamics.

The third study considers sparse sensing and partial observability. Only inlet and outlet pressures and a single midline differential pressure are observed, along with inlet superficial velocities as inputs. The task is to predict the full pressure profile and liquid holdup distribution over time in response to step changes in inlet conditions. In this setting, many state trajectories can explain the same measurements. The variational state posterior captures this ambiguity, and the PDE residual term prevents the posterior from drifting into trajectories that satisfy measurements but violate conservation. The mixture-closure representation is particularly beneficial: when regime uncertainty is high, the solver effectively uses a convex combination of closures, which reduces sensitivity and yields stable trajectories [34]. Forecast uncertainty bands widen appropriately after large input steps and then contract as measurements accumulate.

To evaluate generalization, we train on a set of operating conditions with limited range in superficial velocities and then test on conditions with higher gas fraction and different inclination profiles. Because the closure operators operate on nondimensionalized inputs and local neighborhoods, the learned model extrapolates better than a purely feature-based regressor that maps directly from global features to pressure gradient. The improvement is most pronounced in transients where wave propagation and source-term balance matter. On held-out scenarios, the proposed method reduces mean absolute pressure-gradient error by up to 18% relative to the purely data-driven baseline and by up to 11% relative to a baseline that uses a fixed mechanistic closure without learning. These numbers depend on the severity of mismatch and the sparsity of sensing, but the trend is consistent across tested regimes.

We also examine computational cost and suitability for integration with control. Training involves repeated unrolling of the finite-volume solver and adjoint computations, which is more expensive than training a static regime classifier [35]. However, the learned model can be deployed with amortized inference for regime probabilities and with a single forward solve for prediction, making real-time execution plausible for moderate grid sizes. In an emulated model predictive control loop, the uncertainty output of the hybrid model is used to penalize control actions that drive the system into regions of high regime ambiguity, yielding smoother actuator trajectories and fewer predicted transitions. This illustrates a potential use case where probabilistic regime inference is directly valuable beyond classification.

Finally, we assess physical admissibility. Across all experiments, predicted wall and interfacial shear terms satisfy the imposed sign constraints in the operating region, and single-phase limits are recovered when one phase fraction approaches zero. Energy diagnostics computed from the discrete scheme show nonincreasing mechanical energy in unforced segments, consistent with dissipative behavior. When constraints are removed, training often yields slightly lower training loss but produces occasional energy growth and solver instabilities in extrapolated conditions. This supports the practical role of constraints as stabilizers for both learning and prediction [36].

We perform ablations to isolate the impact of regime conditioning, operator structure, and constraints. When regime conditioning is removed and a single operator is trained across all conditions, the model can fit average behavior but exhibits biased predictions in transients that traverse distinct flow patterns. Pressure-gradient errors increase and uncertainty bands become overconfident because the model cannot represent multimodality associated with competing closures. When operator structure is replaced by a pointwise regressor that ignores neighborhood information, the model struggles to generalize to different grid resolutions and exhibits spurious oscillations because it cannot detect local gradients that influence shear. When dissipation constraints are removed, training loss can decrease slightly, but the learned sources occasionally generate energy, producing numerical instability in extrapolated conditions and causing regime posteriors to oscillate.

We also test sensitivity to temperature and to fluid-property variation by varying density and viscosity ratios within a moderate range. Because the closures operate on nondimensional groups, the learned operators interpolate smoothly across these variations. In contrast, a baseline that uses raw dimensional features requires explicit retraining or careful feature scaling to avoid extrapolation errors [37]. The proposed approach does not eliminate the need for retraining when fluids change drastically, but it reduces the number of parameters that must adapt by separating physical scaling from learned residual structure.

To connect to control tasks, we compute gradients of a cost functional with respect to inlet gas and liquid flow rates through the differentiable solver and closure operators. The cost penalizes deviation from a target outlet pressure and includes a soft barrier on high regime-transition probability to avoid operating near severe switching regions. The differentiable model provides consistent gradients that enable gradient-based optimization. In a set of scenarios with step disturbances, the optimized control actions remain within actuator limits and avoid oscillatory behavior that appears when a non-differentiable, hard-switching regime model is used. While this evaluation is schematic, it illustrates the benefit of embedding a smooth probabilistic regime representation inside the solver rather than treating regime classification as an external discrete decision.

From a numerical perspective, the method benefits from mesh refinement up to the point where measurement sparsity becomes the bottleneck. Finer grids reduce discretization error and improve residual consistency, but they also increase the dimension of the latent state and the complexity of regime inference [38]. We find that operator kernels with neighborhood size proportional to a few pipe diameters are sufficient for stability, and increasing neighborhood size beyond this yields diminishing returns and can overfit to specific transient patterns. This suggests that closure nonlocality at the averaged scale is limited, at least for the tested scenarios, and that most of the benefit comes from capturing local gradients and short-term memory rather than long-range coupling.

The overall evaluation indicates that the proposed hybrid learning approach is most beneficial when sensing is sparse and when transients traverse regimes that differ substantially in their closure behavior. In nearly steady conditions with abundant sensing, simpler methods can be adequate. The hybrid framework therefore should be viewed as a tool for the operationally challenging regimes where uncertainty and switching dominate, rather than as a replacement for all modeling approaches.

7. Conclusion

This paper presented a regime-conditioned, physics-constrained learning framework for gas–liquid two-phase conduit flow that couples discrete regime reasoning to continuous conservation-law prediction. By treating flow regime as a latent stochastic field and representing closures as regime-conditioned neural operators, the approach avoids a brittle separation between regime classification and mechanistic simulation. Variational hybrid inference integrates measurement likelihood with a discrete-PDE consistency term, enabling joint inference of regimes and continuous states under sparse and noisy sensing [39]. Differentiable consistency constraints enforce single-phase limits and energetic admissibility, helping prevent nonphysical behavior and improving numerical stability during gradient-based learning.

The computational studies indicate that the method can recover meaningful regime posteriors, learn closures that compensate for moderate model mismatch without violating physical structure, and improve pressure-gradient prediction in transient scenarios relative to purely data-driven baselines and fixed-closure mechanistic baselines. The resulting uncertainty estimates are particularly useful near regime boundaries, where hard switching can lead to volatile predictions and unstable optimization. While the experiments were designed to be representative rather than facility-specific, they highlight the role of probabilistic regime inference and constrained closure learning as complementary components of a hybrid modeling pipeline.

The variational approximation factorizes regimes and states, which may underrepresent strong posterior couplings in some transients. The one-dimensional backbone captures dominant axial transport but cannot represent three-dimensional effects that may drive certain transitions, and the constraint set, while physically motivated, does not encode all admissibility conditions. Future work can extend the operator parameterization to incorporate nonlocal memory effects, develop richer transition priors informed by domain knowledge, and integrate the framework with adaptive sensing strategies that reduce regime ambiguity by choosing informative measurement locations [40].

References

[1] S. M. El-Behery, A. A. El-Haroun, and M. Abuhegazy, “Prediction of pressure drop in vertical pneumatic conveyors,” Journal of Applied Fluid Mechanics, vol. 10, pp. 519–527, 3 2017.

[2] W. L. Salgado, R. Dam, and C. M. Salgado, “Optimization of a flow regime identification system and prediction of volume fractions in three-phase systems using gamma-rays and artificial neural network.,” Applied radiation and isotopes : including data, instrumentation and methods for use in agriculture, industry and medicine, vol. 169, pp. 109552–, 1 2021.

[3] K. Manikonda, A. R. Hasan, C. E. Obi, R. Islam, A. K. Sleiti, M. W. Abdelrazeq, and M. A. Rahman, “Application of machine learning classification algorithms for two-phase gas-liquid flow regime identification,” in Abu Dhabi International Petroleum Exhibition and Conference, p. D041S121R004, SPE, 2021.

[4] K. Manikonda, R. Islam, C. E. Obi, A. R. Hasan, A. K. Sleiti, M. W. Abdelrazeq, I. G. Hassan, and M. A. Rahman, “Horizontal two-phase flow regime identification with machine learning classification models,” in International Petroleum Technology Conference, p. D011S021R002, IPTC, 2022.

[5] K. Fourrate and M. Loulidi, “Disordered cellular automaton traffic flow model: Phase separated state, density waves and self organized criticality,” The European Physical Journal B - Condensed Matter and Complex Systems, vol. 49, pp. 239–246, 2 2006.

[6] L. Peng, C. Xiao-ping, L. Cai, P. Wen-hao, Z. Yun, X. Pan, and Z. Changsui, “Study on flow regimes of high-pressure and dense-phase pneumatic conveying,” Journal of Physics: Conference Series, vol. 147, pp. 012078–, 2 2009.

[7] B. Linares, A. Velazquez, and J. L. Montanes, “Coupling between thermal oscillations in the surface of a micro-cylinder and vortex shedding,” Nanoscale and Microscale Thermophysical Engineering, vol. 11, pp. 227–246, 12 2007.

[8] P. Gaetani and G. Persico, “Influence of the rotor-driven perturbation on the stator-exit flow within a high-pressure gas turbine stage,” International Journal of Turbomachinery, Propulsion and Power, vol. 6, pp. 28–, 7 2021.

[9] T. C. LaForce and R. T. Johns, “Composition routes for three-phase partially miscible flow in ternary systems,” SPE Journal, vol. 10, pp. 161–174, 6 2005.

[10] B. Maze, H. V. Tafreshi, and B. Pourdeyhimi, “Case studies of air filtration at microscales: Micro- and nanofiber media,” Journal of Engineered Fibers and Fabrics, vol. 3, pp. 155892500800300–, 6 2008.

[11] H. M. Khattab, M. Tantawy, A. A. Hawish, and E. H. Mahmoud, “Reservoir pressure determination using “after hydraulic fracturing closure analysis” technique,” Journal of Petroleum and Mining Engineering, vol. 17, pp. 1–16, 12 2015.

[12] S. O. Rasmussen, P. M. Abbott, T. Blunier, A. J. Bourne, E. J. Brook, S. L. Buchardt, C. Buizert, J. A. Chappellaz, H. Clausen, E. Cook, D. Dahl-Jensen, S. M. Davies, M. Guillevic, S. Kipfstuhl, T. Laepple, I. K. Seierstad, J. P. Severinghaus, J. P. Steffensen, C. Stowasser, A. Svensson, P. Vallelonga, B. M. Vinther, F. Wilhelms, and M. Winstrup, “A first chronology for the north greenland eemian ice drilling (neem) ice core,” Climate of the Past, vol. 9, pp. 2713–2730, 12 2013.

[13] P. H. Holloway and D. G. Pritchard, “Effects of ambient temperature and water vapor on chamber pressure and oxygen level during low atmospheric pressure stunning of poultry,” Poultry science, vol. 96, pp. 2528–2539, 8 2017.

[14] Kondo, H. Aoyama, T. Nishimura, M. Ripepe, G. Lacanna, R. Genco, R. Kawaguchi, T. Yamada, T. Miwa, and E. Fujita, “Gas flux cyclic regime at an open vent magmatic column inferred from seismic and acoustic records,” Scientific reports, vol. 9, pp. 5678–5686, 4 2019.

[15] W. ting Tsai, S. ming Chen, G. hung Lu, and C. S. Garbe, “Characteristics of interfacial signatures on a wind-driven gravity-capillary wave,” Journal of Geophysical Research: Oceans, vol. 118, pp. 1715–1735, 4 2013.

[16] B. S. Yildirim and A. K. Agrawal, “Full-field measurements of self-excited oscillations in momentum-dominated helium jets,” Experiments in Fluids, vol. 38, pp. 161–173, 1 2005.

[17] C. Marchioli and S. Vincent, “Special issue on finite-size particles, drops and bubbles in fluid flows: advances in modelling and simulations,” Acta Mechanica, vol. 230, pp. 381–386, 1 2019.

[18] D. Li, W. Zha, D. Zheng, L. Zhang, and D. Lu, “Effect of matrix-wellbore flow and porosity on pressure transient response in shale formation modeling by dual porosity and dual permeability system,” Journal of Chemistry, vol. 2015, 9 2015.

[19] C. Klenk, D. Heim, M. Ugele, and O. Hayden, “Impact of sample preparation on holographic imaging of leukocytes,” Optical Engineering, vol. 59, pp. 102403–, 12 2019.

[20] H. Hu, J. Zhang, J. Dong, Z. Y. Luo, and T. Xu, “Identification of gas–solid two-phase flow regimes using hilbert–huang transform and neural-network techniques,” Instrumentation Science & Technology, vol. 39, pp. 198–210, 3 2011.

[21] X. Qi, H. Zhu, and J. Zhu, “Demarcation of a new circulating turbulent fluidization regime,” AIChE Journal, vol. 55, pp. 594–611, 1 2009.

[22] Hamzehei, “Cfd modeling and simulation of hydrodynamics in a fluidized bed dryer with experimental validation,” ISRN Mechanical Engineering, vol. 2011, pp. 1–9, 5 2011.

[23] V. V. Mahajan, J. J. Padding, T. M. Nijssen, K. A. Buist, and J. Kuipers, “Nonspherical particles in a pseudo-2d fluidized bed : Experimental study,” AIChE journal. American Institute of Chemical Engineers, vol. 64, pp. 1573–1590, 2 2018.

[24] A. Atmaca and N. Atmaca, ““energy efficiency and engineering applications” in conjunction with the “international energy and engineering conference 2016” (oct 13–14, 2016),” Energy, Ecology and Environment, vol. 3, pp. 1–4, 2 2018.

[25] Y. Hao, W. Wang, B. Yuan, Y. Su, J. An, and H. Shu, “Shale gas simulation considering natural fractures, gas desorption, and slippage flow effects using conventional modified model,” Journal of Petroleum Exploration and Production Technology, vol. 8, pp. 607–615, 7 2017.

[26] I. Sibgatullin, A. Petrov, X. Xu, and L. Maas, “On (n,1) wave attractors: Coordinates and saturation time,” Symmetry, vol. 14, pp. 319–319, 2 2022.

[27] S. Siman-Tov and E. E. Brodsky, “Distinguishing between rheophysical regimes of fluid-saturated granular-flows using dilatancy and acoustic emission measurements,” Granular Matter, vol. 23, pp. 1–11, 4 2021.

[28] A. Rona, “Self-excited supersonic cavity flow instabilities as aerodynamic noise sources,” International Journal of Aeroacoustics, vol. 5, pp. 335–360, 12 2006.

[29] E. Eslamian and H. Shirvani, “Numerical investigation of multiphase supersonic swirl flow inside convergent-divergent nozzle,” The Journal of Computational Multiphase Flows, vol. 8, pp. 120–136, 5 2016.

[30] K. A. Waghorn, S. Bünz, A. Plaza-Faverola, and J. E. Johnson, “3-d seismic investigation of a gas hydrate and fluid flow system on an active mid-ocean ridge; svyatogor ridge, fram strait,” Geochemistry, Geophysics, Geosystems, vol. 19, pp. 2325–2341, 8 2018.

[31] A. Efhaima and M. H. A. Dahhan, “Hydrodynamics study in a fluidized bed reactor using gamma ray computed tomography (ct) technique,” International Journal of Petrochemical Science & Engineering, vol. 1, 11 2016.

[32] S. A. Guazzotti, D. T. Suess, K. R. Coffee, P. K. Quinn, T. S. Bates, A. Wisthaler, A. Hansel, W. P. Ball, R. R. Dickerson, C. Neusüß, P. J. Crutzen, and K. A. Prather, “Characterization of carbonaceous aerosols outflow from india and arabia: Biomass/biofuel burning and fossil fuel combustion,” Journal of Geophysical Research: Atmospheres, vol. 108, pp. 4485–, 8 2003.

[33] M. Toth, C. J. Lobo, V. Friedli, A. Szkudlarek, and I. Utke, “Continuum models of focused electron beam induced processing.,” Beilstein journal of nanotechnology, vol. 6, pp. 1518–1540, 7 2015.

[34] V. P. Koshelets, A. B. Ermakov, L. V. Filippenko, O. V. Koryukin, A. Khudchenko, A. S. Sobolev, M. Y. Torgashin, P. A. Yagoubov, R. W. M. Hoogeveen, W. Vreeling, W. Wild, and O. M. Pylypenko, “Superconducting submm integrated receiver for telis,” Journal of Physics: Conference Series, vol. 43, pp. 1377–1380, 6 2006.

[35] L. Marmo, G. Rovero, and G. Baldi, “Modelling of catalytic gas–solid fluidised bed reactors,” Catalysis Today, vol. 52, pp. 235–247, 9 1999.

[36] M. Mirdrikvand, M. Sadeghi, G. R. Pesch, W. Dreher, and J. Thöming, “Full-field comparison of mrv and cfd of gas flow through regular catalytic monolithic structures,” Processes, vol. 9, pp. 566–, 3 2021.

[37] C. Gamio, J. R. Castro, F. Garcia-Nocetti, L. A. Aguilar, L. Rivera, and H. E. Balleza, “High-pressure gas-oil two-phase flow visualisation using electrical capacitance tomography,” Volume 1, pp. 475–479, 1 2004.

[38] S. B. Joye, I. R. MacDonald, J. P. Montoya, and M. Peccini, “Geophysical and geochemical signatures of gulf of mexico seafloor brines,” Biogeosciences, vol. 2, pp. 295–309, 10 2005.

[39] F. Pérez-Ràfols, F. Forsberg, G. Hellström, and A. Almqvist, “A stochastic two-scale model for rarefied gas flow in highly heterogeneous porous media,” Transport in Porous Media, vol. 135, pp. 219–242, 9 2020.

[40] E. Vittoratos, “Flow regimes during cyclic steam stimulation at cold lake,” Journal of Canadian Petroleum Technology, vol. 30, 1 1991.