From Prediction to Compensation: Neural Operators for Event-Triggered Sliding-Mode PDE Control

Anonymous authors
Paper under double-blind review

Abstract

In recent years, neural operators have made strides in predicting solutions to partial differential equations; however, their reliable application in scenarios with communication-constrained feedback remains under-researched. This paper proposes a neural-operator-assisted dynamic event-triggered integral sliding mode control scheme for reaction-diffusion neural systems subject to sojourn-time-dependent semi-Markovian switching. A Fourier neural operator is employed to estimate matched uncertainties, while a robust term handles prediction residuals and state-preservation errors. Under conditions such as residual boundedness, it is proven that the sliding mode surface is reachable in finite time, the sliding mode dynamics are mean-square exponentially stable, and Zeno behavior is avoided. Independent testing and closed-loop simulations demonstrate that the proposed method enhances estimation accuracy while reducing control energy consumption and input fluctuations; comparisons of triggering mechanisms reveal that the method requires fewer communication transmissions than static strategies.

1 Introduction

Partial differential equations (PDEs) are fundamental tools for modeling spatially distributed systems in fluid dynamics, heat and mass transfer, reaction–diffusion processes, and material evolution. DeepONet and the Fourier neural operator (FNO) learn mappings between function spaces, enabling efficient approximation of parametric PDE solution operators (Lu et al., 2021; Li et al., 2021). Physics-informed, recursive, and generative neural operators (NO) have further extended their modeling and prediction capabilities (Kovachki et al., 2023; Li et al., 2024; Singh et al., 2024; Hu et al., 2025).

Neural operators have also been applied to PDE control, including control-trajectory learning (Holl et al., 2020; Hu et al., 2025) and approximation of backstepping kernels and controller gains (Bhan et al., 2024; Krstic et al., 2024). These approaches, however, primarily learn state evolution or controller-design mappings. For feedback systems with unknown dynamics, the controller must also estimate the unknown terms acting on the system and account for the effects of prediction errors on closed-loop stability.

Continuous transmission of spatially distributed states incurs substantial communication costs. Event-triggered mechanisms reduce communication through on-demand state updates (Tabuada, 2007; Girard, 2015) and have been applied to diffusion and reaction–diffusion PDE control (Selivanov & Fridman, 2016; Espitia et al., 2021; Rathnayake et al., 2021). Representative semi-Markov SMC studies mainly consider finite-dimensional jump systems and commonly design robust gains using bounds on the full uncertainty (Li et al., 2015; Wang et al., 2021; Cheng et al., 2023a). How to jointly address intermittent communication, sojourn-time-dependent switching, and learned uncertainty compensation in spatially distributed PDEs while retaining rigorous closed-loop guarantees therefore remains unresolved. Further related work is provided in Appendix D.

To address this problem, we study uncertain reaction–diffusion neural systems (RDNS) with sojourn-time-dependent semi-Markov switching and propose a neural-operator-assisted

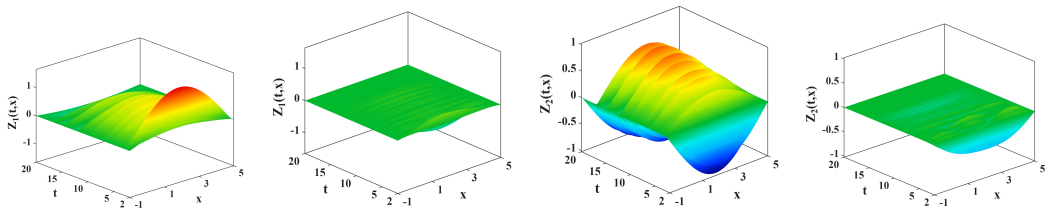
Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

dynamic event-triggered integral-type sliding-mode control (SMC) framework. A causal FNO estimates and compensates for the spatially distributed matched uncertainty using controller-available information, while the robust sliding-mode term handles the prediction residual and state-holding error. The error-modulated dynamic event-triggered mechanism (DETM) adapts the triggering condition to the transmission mismatch, whereas sojourn-time partitioning captures mode transitions across different residence stages.

The contributions of this work are summarized as follows:

- We employ a causal FNO to estimate the matched uncertainty and use the estimate for compensation in an integral-type SMC law.
- We introduce an error-modulated DETM to adapt the triggering condition to communication mismatch, and employ sojourn-time partitioning to characterize sojourn-time-dependent mode transitions.
- We provide verifiable closed-loop stability conditions and establish finite-time reachability of the sliding manifold, mean-square boundedness of the state before reaching, mean-square exponential stability of the sliding dynamics, and exclusion of Zeno behavior.

Figure 1 compares the two control components with and without neural-operator compensation. Without neural compensation, both components exhibit persistent fluctuations during closed-loop operation, whereas these fluctuations are markedly reduced by the proposed method. The neural operator compensates for the predictable part of the matched uncertainty, so that the robust sliding-mode term mainly handles the remaining prediction residual and event-induced state-holding mismatch.



(a) (b) (c) (d)

Figure 1: Spatiotemporal control inputs Z_1 and Z_2 without (a,c) and with (b,d) neural-operator compensation. Neural compensation markedly reduces persistent control fluctuations.

2 Problem Formulation

This section introduces the uncertain RDNS model with residence-time-dependent semi-Markov switching and formulates the control objective, providing the basis for the subsequent NO-assisted SMC design. Problem-specific notation and general mathematical conventions are summarized in Appendix A.

2.1 Mode-Dependent Uncertain RDNS Dynamics

Consider an uncertain RDNS on a bounded spatial domain $\Omega = (a, b) \subset \mathbb{R}$, with active mode $\alpha(t) \in \mathcal{S} = \{1, \dots, N\}$. For $i = \alpha(t)$, the system dynamics are

$$\begin{cases} Y_t(t,x) = M_i Y_{xx}(t,x) - N_i Y(t,x) + \big[B_i + \Delta B_i(t)\big] C\big(Y(t,x)\big) + Z_i(t,x), \\ Y(t,a) = Y(t,b) = 0, \qquad Y(0,x) = Y_0(x), \end{cases} \tag{1}$$

for $(t,x) \in \mathbb{R}_{\geq 0} \times \Omega$. Here, $Y(t,x) \in \mathbb{R}^n$ is the neural state and $Z_i(t,x) \in \mathbb{R}^n$ is the distributed control input. The matrices $M_i = M_i^\top \succ 0$, $N_i = N_i^\top$, and B_i denote the

diffusion, self-inhibition, and nominal connection-weight matrices, respectively. The activation map is $C(y) = \mathrm{col}\{C_1(y_1), \ldots, C_n(y_n)\}, y \in \mathbb{R}^n$. The mode-dependent connection-weight uncertainty is represented as $\Delta B_i(t) = U_i D_i(t) S_i, D_i^\top(t) D_i(t) \preceq I$, where U_i and S_i are known constant matrices and $D_i(t)$ is unknown. Define the induced distributed uncertainty as $r_i(t,x) = \Delta B_i(t) C\big(Y(t,x)\big)$. Accordingly, equation 1 can be written as $Y_t = M_i Y_{xx} - N_i Y + B_i C(Y) + r_i + Z_i$.

2.2 Sojourn-Time-Dependent Semi-Markov Switching

Let $\{\alpha(t), t \geq 0\}$ be a right-continuous semi-Markov process on $\mathcal{S}$, with switching instants $0 = t_0 < t_1 < \cdots$. For $t \in [t_k, t_{k+1})$, define $\ell(t) = t - t_k$. Conditioned on $\alpha(t_k) = i$, the sojourn time follows an exponential distribution with parameter $\zeta_i > 0$, with $f_i(\vartheta) = \zeta_i e^{-\zeta_i \vartheta}, Q_i(\vartheta) = 1 - e^{-\zeta_i \vartheta}, \rho_i(\vartheta) = \frac{f_i(\vartheta)}{1 - Q_i(\vartheta)} = \zeta_i$. Nonhomogeneous and piecewise-homogeneous transition laws have been used to model sojourn-time-dependent switching (Ning et al., 2020b; Tian et al., 2022), while interval-constrained sojourn times have also been considered (Yang et al., 2022). To capture the dependence of the successor mode on the residence duration, introduce a threshold $\vartheta^\star > 0$. For $j \neq i$, let $p_{ij}(\vartheta) = \Pr\{\alpha(t_{k+1}) = j \mid \alpha(t_k) = i,\, t_{k+1} - t_k = \vartheta\} = \begin{cases} \sigma_{ij}^{(s)}, & 0 \leq \vartheta < \vartheta^\star, \\ \sigma_{ij}^{(l)}, & \vartheta \geq \vartheta^\star, \end{cases}$ where $\sigma_{ij}^{(v)} \geq 0$, $\sigma_{ii}^{(v)} = 0$, and $\sum_{j \neq i} \sigma_{ij}^{(v)} = 1$ for $v \in \{s, l\}$. The resulting transition rates are

$$\pi_{ij}(\ell) = \begin{cases} \zeta_i \sigma_{ij}^{(s)}, & 0 \leq \ell < \vartheta^\star, \\ \zeta_i \sigma_{ij}^{(l)}, & \ell \geq \vartheta^\star, \end{cases} \qquad j \neq i, \tag{2}$$

with $\pi_{ii}(\ell) = -\sum_{j \neq i} \pi_{ij}(\ell) = -\zeta_i$. Thus, the departure rate from mode i remains ζ_i, while the distribution over successor modes depends on the residence duration. Since $\mathcal{S}$ is finite and $\zeta_i < \infty$ for all $i \in \mathcal{S}$, the semi-Markov process admits almost surely only finitely many mode transitions on every bounded time interval. Although the marginal sojourn-time distribution is exponential, the switching process is not reduced to a homogeneous Markov chain, because the successor-mode distribution depends explicitly on the realized residence duration.

Assumption 1. *The activation map satisfies $C(0) = 0$. For each $k \in \{1, \ldots, n\}$, there exist constants $b_k^- \leq b_k^+$ such that $b_k^- \leq \frac{C_k(u) - C_k(v)}{u - v} \leq b_k^+, u \neq v$. Define $B^- = \mathrm{diag}\{b_1^-, \ldots, b_n^-\}, B^+ = \mathrm{diag}\{b_1^+, \ldots, b_n^+\}$.*

2.3 Information Structure and Control Objective

To reduce the communication burden of transmitting distributed state profiles, plant-state information is transmitted intermittently. The local sensing/actuation unit continuously accesses the plant state for event evaluation and implementation of the robust sliding-mode term, whereas the networked prediction/feedback unit receives new state profiles only at transmission instants. Related event-triggered architectures for reaction–diffusion PDEs have been developed for boundary and observer-based output-feedback control (Espitia et al., 2021; Lhachemi, 2024). Let $\mathcal{I}_t$ denote the causal information available to the networked prediction/feedback unit, including previously transmitted data, known mode and timing information, and the nominal model. The uncertainty factor $D_i(t)$ and matched uncertainty $r_i(t, \cdot)$ remain unavailable online.

Under this information constraint, the matched uncertainty continues to affect the plant between transmissions. Our objective is to predict $r_i(t, \cdot)$ from $\mathcal{I}_t$ and compensate for its predictable component, while robust SMC handles the remaining prediction residual and communication-induced mismatch.

3 Neural-Operator-Assisted Event-Triggered SMC

Offline uncertainty learning. Figure 2(a) summarizes the offline stage. Trajectories generated under different initial conditions, semi-Markov switching scenarios, and uncertainty realizations are used to train a shared mode-conditioned neural operator for matched-uncertainty prediction.

Online closed-loop compensation. As shown in Figure 2(b), the semi-Markov process determines the active mode, the DETM governs state transmission, and the frozen NO provides uncertainty compensation. Sections 3.1– 3.3 develop the predictor, DETM, and integral-type SMC, respectively.

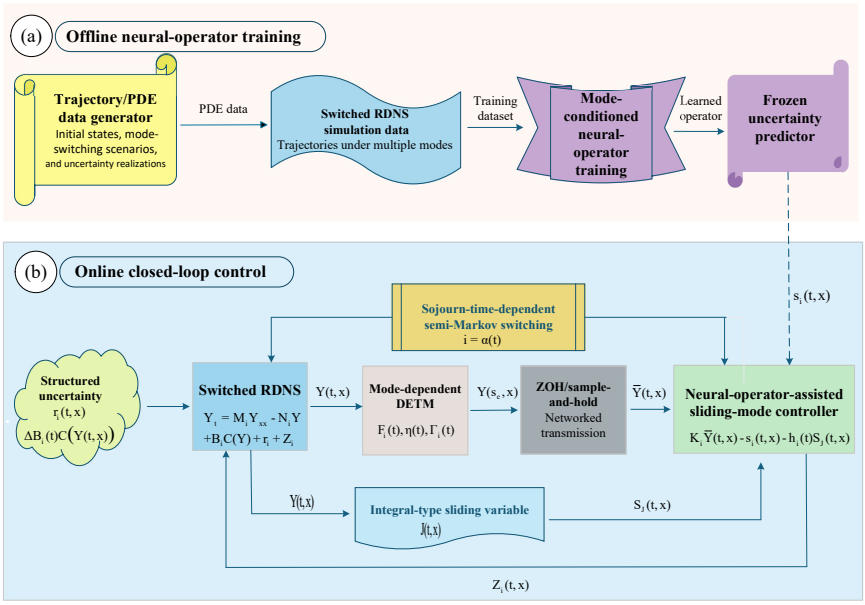


Figure 2: Proposed neural-operator-assisted control framework.

3.1 Causal Neural Operator for Uncertainty Compensation

Unlike prior NO-based PDE controllers that approximate backstepping kernels or controller and observer gains (Bhan et al., 2024; Krstic et al., 2024), our neural operator maps causally available information to the spatially distributed matched uncertainty: $\aleph_t = \mathrm{Feat}(\mathcal{I}_t), \widetilde{r}_i(t,\cdot) = \mathcal{G}_\omega(\aleph_t)$, where $\mathcal{G}_\omega$ is a shared mode-conditioned one-dimensional FNO. The causal information set, residual-anchor construction, and FNO architecture are detailed in Appendices A.1–A.2.

To incorporate the known equilibrium structure $r_i(t,x) = \Delta B_i(t)C(Y(t,x))$ with $C(0) = 0$, the raw FNO output is further processed through state-dependent gates: $s_i(t,x) = g_s(t)g_m(t,x)\widetilde{r}_i(t,x), g_m(t,x) = \frac{\|C(Y_m(t,x))\|_2}{\|C(Y_m(t,x))\|_2+\tau_g}$. Here, Y_m is the controller-side nominal state, g_m and g_s suppress unnecessary predictions as the system approaches the equilibrium. Their detailed implementations, together with the FNO training procedure, are given in Appendix A.3. After offline training, $\mathcal{G}_\omega$ is frozen, and the prediction residual is defined as $r_{e_i}(t,\cdot) = r_i(t,\cdot) - s_i(t,\cdot)$. This construction separates learning performance from closed-loop robustness: the FNO reduces the uncertainty handled by the switching term, while the stability guarantees require only the mode-dependent residual bound in Assumption 2. Since $\mathcal{G}_\omega$ is frozen and uses only $\mathcal{I}_t$, no unavailable uncertainty information is required during deployment.

Assumption 2 (Deployment residual bound)**.** *For each $i \in \mathcal{S}$, there exists a known constant $\bar{\epsilon}_i > 0$ such that $\|r_{e_i}(t,\cdot)\|_{L_2(\Omega)} \le \bar{\epsilon}_i,\ t \ge 0$.*

Assumption 2 is used for the theoretical analysis. In the numerical experiments, the residual levels are calibrated on validation trajectories and independently evaluated on held-out test trajectories.

3.2 Dynamic Event-Triggered Communication

To reduce distributed-state communication, we introduce a mode-dependent DETM (Tabuada, 2007; Cheng et al., 2023b; Espitia et al., 2021). Let $\{s_e\}_{e\ge 0}$ denote the transmission instants. For $t \in [s_e, s_{e+1})$, the held state and error are $\bar{Y}(t,\cdot) = Y(s_e,\cdot)$ and $m(t,x) = Y(t,x) - \bar{Y}(t,x)$. For each mode $i \in \mathcal{S}$, define $\mathrm{E}_i(t) = \int_\Omega m^\top H_i m \,\mathrm{d}x$, $X_i(s_e) = \int_\Omega Y^\top(s_e,x) H_i Y(s_e,x)\,\mathrm{d}x$, $F_i(t) = \xi_i e^{-\nu_i E_i(t)}$, $where \mathrm{H}_i = H_i^\top \succ 0$, $\xi_i > 0$, and $\nu_i \ge 0$. The dynamic threshold variable $\eta(t)$ and triggering condition $\Gamma_i(t)$ evolve as

$$\dot{\eta}(t) = -\lambda_i \eta(t) + \mu_i F_i(t) X_i(s_e) - E_i(t), \quad \eta(0) = \eta_0 > 0, \tag{3}$$

$$\Gamma_i(t) = \mu_i F_i(t) X_i(s_e) - E_i(t) + \frac{\eta(t)}{\theta_i}, \tag{4}$$

where $\lambda_i, \theta_i > 0$ and $0 < \mu_i < 1$. The next transmission instant is determined by

$$s_{e+1} = \inf\left\{ t > s_e \mid \Gamma_{\alpha(t)}(t) \le 0 \right\}. \tag{5}$$

Consequently, for $t \in [s_e, s_{e+1})$, $E_i(t) < \mu_i F_i(t) X_i(s_e) + \frac{\eta(t)}{\theta_i}$. Moreover, $\eta(t) > 0$ holds for all $t \ge 0$, and Zeno behavior is strictly excluded (see Appendix B.1).

3.3 Integral-Type Sliding Variable and Neural-Operator-Assisted Control

Motivated by the robustness of SMC to matched uncertainty and its extensions to semi-Markov and event-triggered systems (Young et al., 1999; Yang et al., 2020; Cheng et al., 2023a), we introduce the integral-type sliding variable

$$J(t,x) = Y(t,x) - \int_0^t \left[M_{\alpha(\tau)} Y_{xx}(\tau,x) - N_{\alpha(\tau)} Y(\tau,x) + B_{\alpha(\tau)} C(Y(\tau,x)) + K_{\alpha(\tau)} Y(\tau,x) \right] d\tau. \tag{6}$$

The sliding manifold is defined by $J(t,\cdot) = 0$. Since $J(0,\cdot) = Y_0(\cdot)$, the system generally undergoes a reaching phase before sliding motion is established. Differentiating along the system trajectory gives, for almost every t, $J_t = r_i + Z_i - K_i Y$.

Using the NO prediction s_i and the held state $\bar{Y}$, the distributed controller is designed as

$$Z_i(t,x) = K_i \bar{Y}(t,x) - s_i(t,x) - h_i(t) S_J(t,x), \tag{7}$$

where $S_J \in \mathrm{Sgn}_\Omega(J)$, with $\mathrm{Sgn}_\Omega(J) = \begin{cases} J/\|J\|_{L_2(\Omega)}, & \|J\|_{L_2(\Omega)} > 0, \\ \{w : \|w\|_{L_2(\Omega)} \le 1\}, & \|J\|_{L_2(\Omega)} = 0. \end{cases}$. Thus, $(K_i \bar{Y} - s_i)$ constitutes the networked compensation term, whereas $(-h_i S_J)$ is implemented locally as the robust sliding-mode term. Using r_{e_i} and $m = Y - \bar{Y}$, the sliding dynamics become $J_t = r_{e_i} - K_i m - h_i(t) S_J$. Let $\bar{k}_i = \sqrt{\lambda_{\max}\left(H_i^{-1/2} K_i^\top K_i H_i^{-1/2} \right)}$. By the event-triggering bound in Section 3.2, $\|K_i m\|_{L_2(\Omega)} \le \bar{k}_i \left(\mu_i F_i(t) X_i(s_e) + \frac{\eta(t)}{\theta_i} \right)^{1/2} = \chi_i(t)$. Under the deployment-residual bound $\|r_{e_i}(t,\cdot)\|_{L_2(\Omega)} \le \bar{\epsilon}_i$, choose $h_i(t) = \bar{\epsilon}_i + \chi_i(t) + \delta_i, \delta_i > 0$. Thus, the neural operator compensates the predictable uncertainty, while the discontinuous term only needs to dominate the prediction residual and the event-induced mismatch.

4 Theoretical Guarantees

We establish finite-time reachability, mean-square boundedness during reaching, and mean-square exponential stability of the sliding dynamics. Here, $P_i \succ 0$ and $R \succ 0$ are the

mode-dependent and common Lyapunov matrices, respectively; G_i is an auxiliary matrix; $\mathcal{M}_{R,i}^{(v)}$ and $\Phi_{S,i}^{(v)}$ denote the reaching- and sliding-phase analysis matrices; and $v \in \{\mathrm{s},\mathrm{l}\}$ indexes the two residence regimes.

4.1 Finite-Time Reachability

Let $T_r = \inf\{t \geq 0 : \|J(t,\cdot)\|_{L_2(\Omega)} = 0\}$ denote the first hitting time of the sliding manifold.

Theorem 1. *Under controller equation 7, Assumption 2, and $h_i(t) = \bar{\epsilon}_i + \chi_i(t) + \delta_i$, with $\delta_i > 0$, the sliding manifold is reached in finite time. In particular, $T_r \leq \frac{\|J(0,\cdot)\|_{L_2(\Omega)}}{\delta}, \delta = \min_{i\in\mathcal{S}} \delta_i > 0$. Moreover, the sliding motion can be maintained for $t \geq T_r$.*

The proof is given in Appendix B.2.

4.2 Boundedness During Reaching

Finite-time convergence of J does not by itself rule out state growth before the sliding manifold is reached. Define the stopped time $\tau_t = t \wedge T_r$.

Theorem 2 (Boundedness During Reaching)**.** *Let $\gamma_R > 0$, $0 < b_R < \min\{2\gamma_R, \min_{i\in\mathcal{S}} \lambda_i\}$, and $0 < \mu_i\xi_i < 1$. Suppose that the reaching-phase analysis condition $M_{R,i}^{(v)} \prec 0$ and the coupling relation $G_iM_i = R$ hold for all $i \in \mathcal{S}$ and $v \in \{s,l\}$. Then the system state remains mean-square bounded before the sliding manifold is reached: $\sup_{0\leq t\leq \bar{T}_r} \mathbb{E}\Big\{\|Y(\tau_t,\cdot)\|_{L_2(\Omega)}^2 + \|Y_x(\tau_t,\cdot)\|_{L_2(\Omega)}^2\Big\} < \infty$, where $\bar{T}_r = \|J(0,\cdot)\|_{L_2(\Omega)}/\delta$.*

The Lyapunov functional and proof are provided in Appendix B.3.

4.3 Exponential Stability of the Sliding Dynamics

Once the sliding manifold is reached, $J = J_t = 0$. The equivalent control is $Z_i^{\mathrm{eq}} = K_iY - r_i$, and the ideal sliding dynamics reduce to $Y_t = M_iY_{xx} + (K_i - N_i)Y + B_iC(Y)$. Hence, the matched uncertainty is removed from the sliding dynamics.

Theorem 3 (Exponential Stability on the Sliding Manifold)**.** *Let $\gamma_S > 0$ and $0 < b_S < \min\{2\gamma_S, \min_{i\in\mathcal{S}} \lambda_i\}$. If $G_iM_i = R$ and $\Phi_{S,i}^{(v)} \prec 0$ hold for all $i \in \mathcal{S}$ and $v \in \{s,l\}$, then the sliding dynamics are mean-square exponentially stable. Specifically, for every $\tau \geq 0$, $\mathbb{E}\|Y(T_r+\tau,\cdot)\|_{L_2(\Omega)}^2 \leq \frac{e^{-b_S\tau}}{p_S}\mathbb{E}V_S(T_r,\alpha(T_r)), p_S = \min_{i\in\mathcal{S}} \lambda_{\min}(P_i) > 0$.*

The proof is given in Appendix B.4.

4.4 Joint Controller Synthesis

Following the standard change-of-variables strategy in LMI-based SMC synthesis (Edwards, 2004), introducing $L_i = G_iK_i$ converts the bilinear analysis conditions into the synthesis conditions $\widehat{M}_{R,i}^{(v)} \prec 0$ and $\widehat{\Phi}_{S,i}^{(v)} \prec 0$.

Theorem 4 (Joint Controller Synthesis)**.** *Suppose that, for all $i \in \mathcal{S}$ and $v \in \{s,l\}$, $G_iM_i = R, \widehat{M}_{R,i}^{(v)} \prec 0, \widehat{\Phi}_{S,i}^{(v)} \prec 0$. Then $K_i = G_i^{-1}L_i$ is well defined, and the closed-loop system satisfies the guarantees in Theorems 1– 3.*

The complete synthesis conditions and gain recovery are given in Appendix B.5, with numerical feasibility reported in Appendix C.

5 Experiments

We evaluate the proposed framework along three dimensions: matched-uncertainty prediction, closed-loop regulation and control burden, and communication efficiency. Prediction is assessed on trajectory-disjoint held-out data, whereas case-matched closed-loop comparisons use paired stochastic realizations.

5.1 Experimental Setup

We consider a three-mode, two-state RDNS and generate 120 15-s trajectories, split by trajectory into 84/18/18 for training, validation, and testing (25,284/5,418/5,418 field samples). All predictors share the same data and training objective; model selection and residual calibration use only the validation set. Closed-loop comparisons use matched stochastic realizations. Further details are given in Appendix A.5.

5.2 Matched-Uncertainty Prediction on Held-Out Trajectories

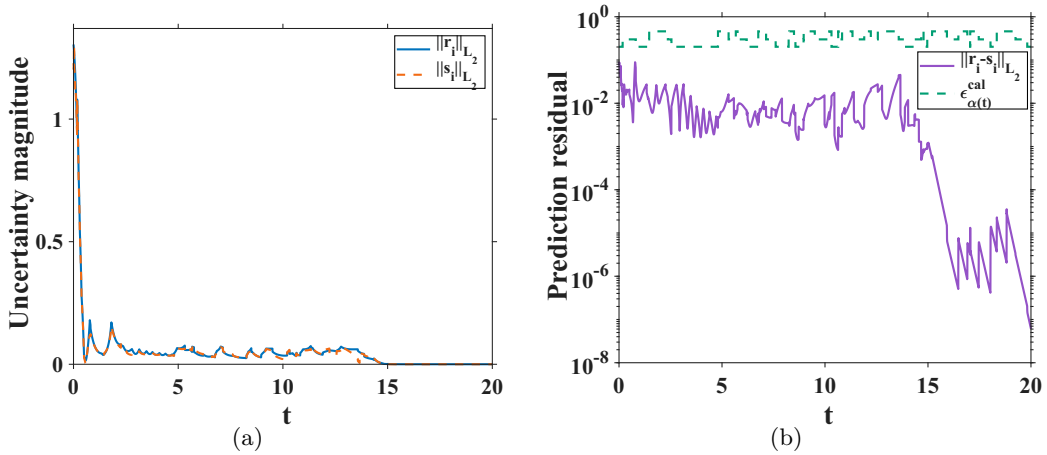
Held-out prediction. We first evaluate the prediction of the matched uncertainty on held-out trajectories. The benchmark uses the model-gated output $g_m(t,\cdot)\widetilde{r}_i(t,\cdot)$, and the prediction error is defined as $e_{\text{pred}}(t) = \left\| r_{\alpha(t)}(t,\cdot) - g_m(t,\cdot)\widetilde{r}_{\alpha(t)}(t,\cdot) \right\|_{L_2(\Omega)}$. Table 1 reports the mean error and Q95 (95th percentile). The state-aware gate g_s is activated only during closed-loop deployment, where the deployed prediction is $s_i = g_s g_m \widetilde{r}_i$ as defined in Section 3.1.

We compare the causal FNO with a 1D U-Net (Ronneberger et al., 2015), CNN–DeepONet (Lu et al., 2021), and two nonlearned baselines. All methods use the same 5,418 test samples. As shown in Table 1, the FNO achieves the lowest mean and Q95 errors, reducing the mean error by 91.79% and 89.25% relative to the zero and latest-interval predictors, respectively. The U-Net performs comparably, whereas CNN–DeepONet is less accurate. Additional results are provided in Appendix A.4.

Table 1: Matched-uncertainty prediction on held-out trajectories.

Method	Params.	Mean	Q95
Zero predictor	0	0.039515	0.085556
Latest-interval estimate	0	0.030181	0.143927
CNN-DeepONet	263,042	0.003718	0.013883
1D U-Net	363,842	0.003351	0.011932
Causal FNO (ours)	287,362	**0.003245**	**0.010175**

Closed-loop deployment. Figure 3 evaluates the deployed FNO along a representative 20-s trajectory. The compensation field s_i captures the main variation of r_i, while the residual $\|r_i - s_i\|_{L_2}$ remains below the validation-calibrated, mode-dependent threshold. This result illustrates that neural compensation reduces the uncertainty handled by the robust term. The mode-wise calibration and held-out checks in Appendix A Table 11 further show that all test residual maxima remain below their calibrated bounds, with no observed violations.



(a) (b)

Figure 3: Representative uncertainty prediction and residual.

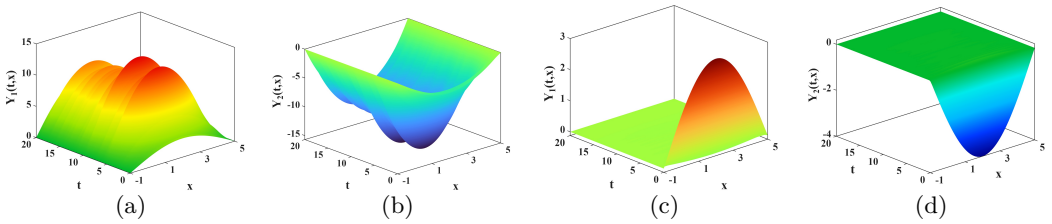


Figure 4: Spatiotemporal state responses of the uncertain open-loop and NO-assisted closed-loop systems. (a,b) Uncertain open loop. (c,d) NO-assisted closed loop.

5.3 Closed-Loop Regulation and Control Burden

For the two-state example, the active control input is written as $Z_{\alpha(t)}(t,x) = \mathrm{col}\{Z_1(t,x), Z_2(t,x)\}$. Figure 4 compares the uncertain open-loop and NO-assisted closed-loop responses. Without control, both state components remain away from the zero equilibrium, whereas under the proposed controller their spatial amplitudes decay toward zero. This is consistent with the established mean-square exponential stability of the sliding dynamics.

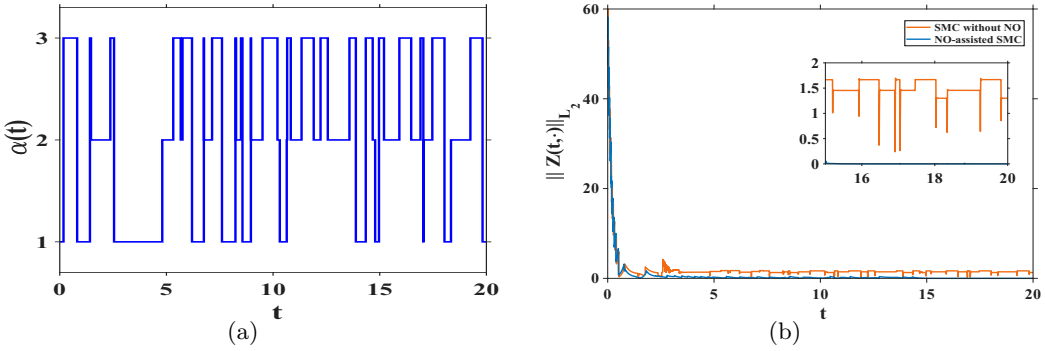


Figure 5: Semi-Markov switching and control effort with and without NO compensation. (a) Mode trajectory $\alpha(t)$. (b) Spatial L_2 norm of the control input for SMC without NO and NO-assisted SMC.

To isolate the effect of neural-operator compensation, we compare the proposed NO-assisted SMC with an otherwise identical SMC without NO compensation. The latter sets $s_i = 0$ and replaces the prediction-residual allowance $\bar{\epsilon}_i$ in $h_i(t)$ with a bound on the full matched uncertainty, while keeping the nominal feedback gains and DETM settings unchanged.

Control burden. We quantify control burden using the mean and peak spatial-L_2 control magnitudes, control energy, and temporal total variation, as defined in Appendix A.5. As shown in Table 2, neural-operator compensation reduces the mean control magnitude, control energy, and total variation by 74.09%, 21.19%, and 98.94%, respectively, while the peak magnitude decreases by only 2.88%, as it is dominated by the common initial transient.

Table 2: Control burden with and without neural-operator compensation.

Metric	Without NO	NO-assisted	Reduction
Mean L_2	1.8089	0.4686	**74.09%**
Peak L_2	59.9022	58.1778	**2.88%**
Energy	385.49	303.82	**21.19%**
Total variation	16555.2	176.18	**98.94%**

Figure 5(a) shows repeated switching among the three modes with varying sojourn durations, consistent with the semi-Markov model in Section 2.2. Under the same realization, Figure 5(b) shows that the NO-assisted SMC substantially reduces the persistent control fluctuations observed without NO compensation.

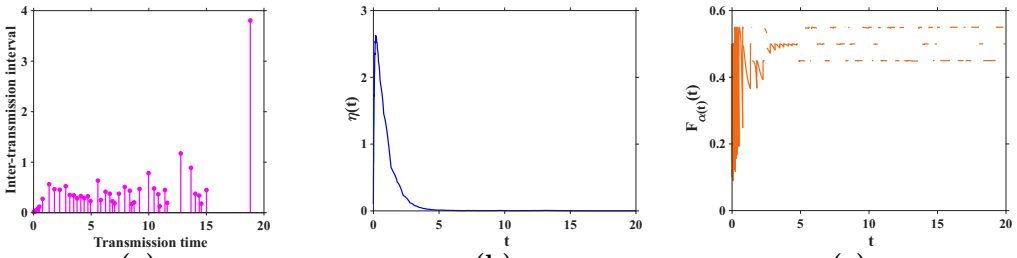
5.4 Communication Efficiency

DETM behavior. Figure 6 illustrates the operation of the proposed DETM. The nonuniform release intervals in Figure 6(a) reflect aperiodic transmission and are consistent with the no-Zeno analysis, while Figure 6(b) shows that $\eta(t)$ remains positive. In Figure 6(c), $F_{\alpha(t)}(t)$ varies with the state-holding error and active mode; error resets and mode transitions therefore produce the observed variations and jumps. These results illustrate the error- and mode-dependent adaptation of the DETM developed in Section 3.2. Communication cost counts only network transmissions of distributed state profiles; local triggering and robust-term evaluations incur no network transmission.

Communication comparison. Following the distinction between static and dynamic triggering mechanisms, we compare the proposed DETM with a static event-triggered mechanism (SETM) obtained by removing the dynamic term $\eta(t)/\theta_i$, while keeping all other settings unchanged. The comparison uses 100 paired 30-s semi-Markov paths. As shown in Table 3, DETM reduces the mean number of transmissions from 573.17 to 44.41, corresponding to a mean pathwise reduction of 92.12%. The paired-path construction and bootstrap procedure are specified in Appendix A.5.

Table 3: Communication comparison over 100 paired semi-Markov paths.

Quantity	Estimate	95% interval
Mean DETM transmissions	44.41	[43.59, 45.25]
Mean SETM transmissions	573.17	[560.00, 587.28]
Mean paired saving	**528.76**	[515.01, 543.04]
Mean pairwise reduction	**92.12%**	[91.85%, 92.39%]
Fewer/equal/more paths	**100/0/0**	–



(a) (b) (c)

Figure 6: Representative behavior of the proposed DETM. (a) Release interval. (b) Auxiliary variable $\eta(t)$. (c) Adaptive modulation factor $F_{\alpha(t)}(t)$.

6 Conclusion

This paper developed an NO-assisted DETM-SMC framework for uncertain RDNSs with sojourn-time-dependent semi-Markov switching. A causal FNO estimates the spatially distributed matched uncertainty from controller-available information, while the robust term handles the prediction residual and state-holding error. We established finite-time sliding-manifold reachability, mean-square boundedness before reaching, mean-square exponential stability of the sliding dynamics, and exclusion of Zeno behavior. When a bound on the full matched uncertainty is available, the FNO can be disabled and the robust gain adjusted without losing these guarantees. Experiments show that the FNO outperforms the zero predictor and the latest-interval estimate, while NO compensation reduces control energy and control-input total variation relative to SMC without NO compensation, and the DETM requires fewer transmissions than the SETM. The framework therefore reduces control conservatism while retaining the stated guarantees under the residual-bound condition; exact neural prediction is not required. We believe this work provides a theoretically grounded step from PDE state prediction to uncertainty compensation, with potential applications to unknown-dynamics prediction and closed-loop control of fluid, heat-transfer, and reaction–diffusion systems.

AI Use Statement

We used generative AI tools only to assist with manuscript drafting and language polishing. All AI-assisted text was reviewed and revised by the authors, who take full responsibility for the final manuscript.

Reproducibility Statement

We support reproducibility through detailed methodological, theoretical, and experimental documentation. Sections 2 –4 specify the system model, proposed controller, and analytical assumptions. Appendix A provides the causal input construction, FNO architecture and training procedure, comparison methods, statistical evaluation, simulation parameters, evaluation metrics, and random seeds. Complete theoretical proofs are provided in Appendix B, and numerical feasibility checks for the synthesis LMIs are reported in Appendix C.

References

Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Neural Operators for Bypassing Gain and Control Computations in PDE Backstepping. *IEEE Transactions on Automatic Control*, 69(8):5310–5325, 2024.

Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Adaptive control of reaction–diffusion pdes via neural operator-approximated gain kernels. *Systems & Control Letters*, 195:105968, 2025.

Jun Cheng, Lifei Xie, Dan Zhang, and Huaicheng Yan. Novel event-triggered protocol to sliding mode control for singular semi-Markov jump systems. *Automatica*, 151:110906, 2023a.

Jun Cheng, Jiangming Xu, Ju H Park, and Michael V Basin. Protocol-based load frequency control for power systems with nonhomogeneous sojourn probabilities. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(9):5742–5750, 2023b.

Jun Cheng, Jiangming Xu, Huaicheng Yan, Zheng-Guang Wu, and Wenhai Qi. Neural network-based sliding mode control for semi-markov jumping systems with singular perturbation. *IEEE Transactions on Cybernetics*, 55(1):259–268, 2024.

Ze Cheng, Zhuoyu Li, Xiaoqiang Wang, Jianing Huang, Zhizhou Zhang, Zhongkai Hao, and Hang Su. Accelerating pde-constrained optimization by the derivative of neural operators. *arXiv preprint arXiv:2506.13120*, 2025.

Wenqi Cui, Yan Jiang, Baosen Zhang, and Yuanyuan Shi. Structured neural-pi control with end-to-end stability and output tracking guarantees. *Advances in Neural Information Processing Systems*, 36:68434–68457, 2023.

Christopher Edwards. A practical method for the design of sliding mode controllers using linear matrix inequalities. *Automatica*, 40(10):1761–1769, 2004.

Nicolás Espitia, Iasson Karafyllis, and Miroslav Krstic. Event-triggered boundary control of constant-parameter reaction–diffusion pdes: A small-gain approach. *Automatica*, 128: 109562, 2021.

Antoine Girard. Dynamic triggering mechanisms for event-triggered control. *IEEE Transactions on Automatic Control*, 60(7):1992–1997, 2015.

Philipp Holl, Nils Thuerey, and Vladlen Koltun. Learning to control pdes with differentiable physics. In *International Conference on Learning Representations*, 2020.

Peiyan Hu, Rui Wang, Xiang Zheng, Tao Zhang, Haodong Feng, Ruiqi Feng, Long Wei, Yue Wang, Zhi-Ming Ma, and Tailin Wu. Wavelet diffusion neural operator. In *The Thirteenth International Conference on Learning Representations*, 2025.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89): 1–97, 2023.

Miroslav Krstic, Luke Bhan, and Yuanyuan Shi. Neural operators of backstepping controller and observer gain functions for reaction–diffusion pdes. *Automatica*, 164:111649, 2024.

Hugo Lhachemi. Event-triggered finite-dimensional observer-based output feedback stabilization of reaction–diffusion pdes. *IEEE Transactions on Automatic Control*, 69(8): 5651–5657, 2024.

Fanbiao Li, Ligang Wu, Peng Shi, and Cheng-Chew Lim. State estimation and sliding mode control for semi-markovian jump systems with mismatched uncertainties. *Automatica*, 51: 385–393, 2015.

Fanbiao Li, Zheng Wu, Chunhua Yang, Yang Shi, Tingwen Huang, and Weihua Gui. A novel learning-based asynchronous sliding mode control for discrete-time semi-markov jump systems. *Automatica*, 143:110428, 2022.

Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

Zongyi Li, Nikola Borislavov Kovachki, Chris Choy, Boyi Li, Jean Kossaifi, Shourya Prakash Otta, Mohammad Amin Nabian, Maximilian Stadler, Christian Hundt, Kamyar Azizzadenesheli, and Anima Anandkumar. Geometry-informed neural operator for large-scale 3d PDEs. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM / IMS J. Data Sci.*, 1(3), May 2024.

Na Liu, Wenjie Qin, Jun Cheng, Jinde Cao, and Dan Zhang. Protocol-based control for semi-markov reaction-diffusion neural networks. *Neural Networks*, 179:106556, 2024.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.

Sharath Matada, Luke Bhan, Yuanyuan Shi, and Nikolay Atanasov. Generalizable motion planning via operator learning. In *The Thirteenth International Conference on Learning Representations*, 2025.

Jaime A. Moreno and Marisol Osorio. Strict lyapunov functions for the super-twisting algorithm. *IEEE Transactions on Automatic Control*, 57(4):1035–1040, 2012.

Zepeng Ning, Lixian Zhang, and Patrizio Colaneri. Semi-markov jump linear systems with incomplete sojourn and transition information: Analysis and synthesis. *IEEE Transactions on Automatic Control*, 65(1):159–174, 2020a.

Zepeng Ning, Lixian Zhang, Ali Mesbah, and Patrizio Colaneri. Stability analysis and stabilization of discrete-time non-homogeneous semi-markov jump linear systems: A polytopic approach. *Automatica*, 120:109080, 2020b.

Bhathiya Rathnayake, Mamadou Diagne, Nicolás Espitia, and Iasson Karafyllis. Observer-based event-triggered boundary control of a class of reaction–diffusion pdes. *IEEE Transactions on Automatic Control*, 67(6):2905–2917, 2021.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

Cristopher Salvi, Maud Lemercier, and Andris Gerasimovics. Neural stochastic pdes: Resolution-invariant learning of continuous spatiotemporal dynamics. *Advances in Neural Information Processing Systems*, 35:1333–1344, 2022.

Anton Selivanov and Emilia Fridman. Distributed event-triggered control of diffusion semilinear pdes. *Automatica*, 68:344–351, 2016.

Ashutosh Singh, Ricardo Borsoi, Deniz Erdogmus, and Tales Imbiriba. Learning semilinear neural operators: A unified recursive framework for prediction and data assimilation. In *International Conference on Learning Representations*, 2024.

Paulo Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 52(9):1680–1685, 2007.

Yongxiao Tian, Huaicheng Yan, Hao Zhang, Meng Wang, and Jianjun Yi. Time-varying gain controller synthesis of piecewise homogeneous semi-markov jump linear systems. *Automatica*, 146:110594, 2022.

Vadim Utkin. Discussion aspects of high-order sliding mode control. *IEEE Transactions on Automatic Control*, 61(3):829–833, 2016.

Jing Wang, Tingting Ru, Jianwei Xia, Hao Shen, and Victor Sreeram. Asynchronous event-triggered sliding mode control for semi-markov jump systems within a finite-time interval. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 68(1):458–468, 2021.

Jiangming Xu, Jun Cheng, Mohammed Chadli, and Wenhai Qi. Protocol-based smc for fuzzy semi-markov switching systems with multizone probabilitic time-varying delays. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55(4):3026–3035, 2025.

Dongsheng Yang, Ting Li, Xiangpeng Xie, and Huaguang Zhang. Event-triggered integral sliding-mode control for nonlinear constrained-input systems with disturbances via adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50(11):4086–4096, 2020.

Jianan Yang, Zepeng Ning, Yimin Zhu, Lixian Zhang, and Hak-Keung Lam. Semi-markov jump linear systems with bi-boundary sojourn time: Anti-modal-asynchrony control. *Automatica*, 140:110270, 2022.

Jianan Yang, Yimin Zhu, Lixian Zhang, Guangren Duan, Ming Liu, and Xibin Cao. Smooth control with flexible duration for semi-markov jump linear systems. *Automatica*, 164: 111612, 2024.

Luan Yang, Jingdong Zhang, Qunxi Zhu, and Wei Lin. Neural event-triggered control with optimal scheduling. In *Forty-second International Conference on Machine Learning*, 2025.

K.D. Young, V.I. Utkin, and U. Ozguner. A control engineer's guide to sliding mode control. *IEEE Transactions on Control Systems Technology*, 7(3):328–342, 1999.

Hongpeng Yuan, Ji Wang, and Mamadou Lamine Diagne. Deeponet of dynamic event-triggered backstepping boundary control for reaction-diffusion pdes. *2025 IEEE 64th Conference on Decision and Control (CDC)*, pp. 2989–2994, 2025.

Jingdong Zhang, Luan Yang, Qunxi Zhu, and Wei Lin. FESSNC: Fast exponentially stable and safe neural controller. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 60076–60098. PMLR, 21–27 Jul 2024.

Lixian Zhang, Bo Cai, Tianyu Tan, and Yang Shi. Stabilization of non-homogeneous hidden semi-markov jump systems with limited sojourn-time information. *Automatica*, 117: 108963, 2020.

Ruimei Zhang, Deqiang Zeng, Ju H. Park, Hak-Keung Lam, and Shouming Zhong. Fuzzy adaptive event-triggered sampled-data control for stabilization of t–s fuzzy memristive neural networks with reaction–diffusion terms. *IEEE Transactions on Fuzzy Systems*, 29 (7):1775–1785, 2021.

A Notation, Neural-Operator Implementation, and Experimental Protocols

Notation. Tables 4 and 5 summarize the problem-specific notation. For $y, z \in L_2(\Omega; \mathbb{R}^n)$, define $\langle y, z\rangle_{L_2(\Omega)} = \int_\Omega y^\top(x) z(x)\, dx$ and $\|y\|^2_{L_2(\Omega)} = \langle y, y\rangle_{L_2(\Omega)}$. For a square matrix A, $\mathrm{He}(A) = A + A^\top$; $\lambda_{\min}(A)$, $\lambda_{\max}(A)$, and $\mathrm{tr}(A)$ denote its minimum eigenvalue, maximum eigenvalue, and trace, respectively. The operators $\mathrm{col}\{\cdot\}$ and $\mathrm{diag}\{\cdot\}$ form block vectors and block-diagonal matrices. The symbols $\mathbb{E}$ and $\mathcal{L}$ denote expectation and the infinitesimal generator. Matrix inequalities are understood in the Loewner sense; I and 0 denote identity and zero matrices of compatible dimensions, and $*$ denotes a symmetric block.

Table 4: System and switching notation.

Symbol	Meaning
n, N	State dimension and number of switching modes.
$U_i, D_i(t), S_i$	Known uncertainty-factor matrices and the unknown norm-bounded factor satisfying $D_i^\top(t) D_i(t) \preceq I$.
$p_{ij}(\vartheta)$	Conditional probability that mode j follows mode i after a completed sojourn of duration ϑ.
$\Omega = (a, b)$, $\mathcal{S}$	Spatial domain and finite mode set $\{1, \dots, N\}$.
$\alpha(t)$, i, j	Active mode and mode indices; mode-dependent quantities are evaluated at $i = \alpha(t)$.
$Y(t, x)$, $Y_0(x)$	Distributed neural state and its initial profile.
$Z_i(t, x)$	Mode-dependent distributed control input.
M_i, N_i, B_i	Diffusion, self-inhibition, and nominal connection-weight matrices.
$C(Y)$; B^-, B^+	Componentwise activation map and its lower and upper sector-bound matrices.
$\Delta B_i(t)$, $r_i(t, x)$	Structured connection-weight perturbation and the induced matched uncertainty $r_i = \Delta B_i(t) C(Y)$.
t_k, $\ell(t)$	kth mode-switching instant and current residence age, $\ell(t) = t - t_k$.
ϑ, $\vartheta^\star$	Completed sojourn duration and the threshold separating short and long residence regimes.
ζ_i, f_i, Q_i, ρ_i	Sojourn-rate parameter, density, cumulative distribution, and conditional departure rate in mode i.
$\sigma_{ij}^{(v)}, \Sigma^{(v)}$	Successor-mode probability and its matrix in residence regime $v \in \{\mathrm{s}, \mathrm{l}\}$.
$\pi_{ij}(\ell)$	Transition intensity from mode i to mode j at residence age ℓ.

Table 5: Learning, communication, and control notation.

Symbol	Meaning
$\mathcal{I}_t$, $\aleph_t$	Controller-available causal information and the feature field $\aleph_t = \mathrm{Feat}(\mathcal{I}_t)$.
$Y_m(t, x)$	Controller-side nominal state, synchronized at transmissions and propagated by the nominal model between transmissions.
$r_e^{\mathrm{anc}}(x)$	Interval-averaged residual anchor formed after a completed transmission interval.
$\mathcal{G}_\omega$, $\widetilde{r}_i$	Frozen mode-conditioned FNO and its raw matched-uncertainty estimate.
g_m, g_s, s_i	Pointwise and deployment gates, and the applied compensation $s_i = g_s g_m \widetilde{r}_i$.
r_{e_i}, $\bar{\epsilon}_i$	Deployment residual $r_{e_i} = r_i - s_i$ and its mode-dependent $L_2(\Omega)$ bound.
$s_e, \bar{Y}$	eth transmission instant and the state held by the controller on $[s_e, s_{e+1})$.
$m(t, x)$	State-holding mismatch $Y(t, x) - \bar{Y}(t, x)$.
H_i, E_i, $X_i(s_e)$	Trigger-weight matrix, weighted mismatch energy, and most recently transmitted state energy.
F_i, η, Γ_i	Error-dependent modulation factor, dynamic auxiliary state, and event function.

Continued on the next page

Table 5 (continued)

Symbol	Meaning
J, $\mathrm{Sgn}_\Omega(J)$, S_J	Integral-type sliding variable, its set-valued spatial sign map, and a selected direction.
K_i, h_i, δ_i	Nominal feedback gain, robust switching amplitude, and positive reaching margin.
$\bar{k}_i$, χ_i	Induced gain and event-induced mismatch bound used in h_i.
T_r, $\bar{T}_r$	First hitting time of the sliding manifold and its derived upper bound.
V_R, V_S	Reaching- and sliding-phase Lyapunov functionals.
P_i, R, G_i, L_i	LMI variables, with $G_i M_i = R$ and $L_i = G_i K_i$.
$\mathcal{M}_{R,i}^{(v)}$, $\Phi_{S,i}^{(v)}$	Reaching- and sliding-phase analysis LMI matrices.
$\widehat{\mathcal{M}}_{R,i}^{(v)}$, $\widehat{\Phi}_{S,i}^{(v)}$	Convex synthesis LMI matrices; $K_i = G_i^{-1} L_i$.
$\gamma_R, \gamma_S, b_R, b_S$	Positive analysis scalars used in the reaching- and sliding-phase Lyapunov inequalities.
$\Xi_{R,i}, \Xi_{S,i}, \varepsilon_{r_i}, \varepsilon_{g_i}$	Sector multipliers and positive slack variables used in the analysis LMIs.

This appendix specifies the causal information supplied to the neural operator, the construction of the residual anchors, the FNO architecture, and the training and evaluation protocols used in Section 5. The current untransmitted plant state, the uncertainty factor $D_i(t)$, and the target field $r_i(t,\cdot)$ are unavailable to the frozen neural operator during closed-loop deployment.

A.1 Causal Controller Information and Residual Anchors

At each transmission instant s_e, the controller receives $Y(s_e,\cdot)$ and updates the held state $\bar{Y}(t,\cdot) = Y(s_e,\cdot)$ for $t \in [s_e, s_{e+1})$. The auxiliary nominal state introduced in Section 3.1 is synchronized as $Y_m(s_e,\cdot) = Y(s_e,\cdot)$ and, between transmissions, is propagated according to

$$Y_{m,t} = M_{\alpha(t)} Y_{m,xx} - N_{\alpha(t)} Y_m + B_{\alpha(t)} C(Y_m) + Z_{\alpha(t)}, \tag{8}$$

with the plant boundary conditions. Equation 8 uses only the most recently received state, the known nominal model, the observed mode, and the applied control input.

The sensing unit has continuous access to the plant state for evaluating the event condition. After a transmission interval has been completed, it also forms the residual anchor

$$\begin{aligned} r_e^{\mathrm{anc}}(x) = {} & \frac{Y(s_e,x) - Y(s_{e-1},x)}{s_e - s_{e-1}} \\ & - \frac{1}{s_e - s_{e-1}} \int_{s_{e-1}}^{s_e} \big[M_{\alpha(\tau)} Y_{xx}(\tau,x) - N_{\alpha(\tau)} Y(\tau,x) \\ & \qquad + B_{\alpha(\tau)} C(Y(\tau,x)) + Z_{\alpha(\tau)}(\tau,x) \big] d\tau. \end{aligned} \tag{9}$$

This expression remains valid when an interval contains a mode transition because the integrand follows $\alpha(\tau)$. The anchor is generated only when $s_e - s_{e-1} \geq 0.0200\,\mathrm{s}$, set to zero at the boundary nodes, and transmitted with the new state packet. It is therefore an interval-averaged observation of the matched uncertainty over an elapsed interval, rather than the current uncertainty field. Before the first valid anchor is received, the corresponding history entries remain zero and are marked invalid.

A.2 Causal Feature Field and FNO Architecture

The causal feature field $\aleph_t = \mathrm{Feat}(\mathcal{I}_t)$ satisfies $\aleph_t \in \mathbb{R}^{N_x \times 58}$, where $N_x = 61$. Four transmitted state profiles and four residual anchors are retained. Table 6 gives the channel order used for training and deployment.

Each state-history slot contains its mode, release code, elapsed interval, and validity flag. Each anchor slot contains its mode, interval length, and validity flag. The reliability descriptors record the times since the latest state packet and anchor, anchor availability, and the filled fractions of both histories. The switching descriptors record the time since the

Table 6: Channel allocation of the causal feature field $\aleph_t$.

Feature group	Channels	Feature group	Channels
Y_m	2	Event metadata	16
Transmitted-state history	8	Anchor metadata	12
Residual-anchor history	8	Current-mode code	3
Reliability descriptors	5	Switching descriptors	3
Normalized coordinate	1	Total	58

latest mode transition and its origin and destination modes. Mode and release indices are divided by three before standardization. Scalar features are broadcast over the spatial grid, while the coordinate is mapped to $[-1, 1]$ by

$$\frac{2(x - x_{\min})}{x_{\max} - x_{\min}} - 1. \tag{10}$$

Missing history slots are zero padded and distinguished from observed zero fields by their validity flags. All normalization statistics are fitted on the training trajectories only.

A pointwise convolution lifts $\aleph_t$ from 58 channels to width 64. With v_0 denoting the lifted field, four residual spectral blocks compute

$$v_{d+1} = \mathrm{GN}_d[v_d + \mathrm{GELU}(\mathcal{K}_d v_d + W_d v_d)], \qquad d = 0, \dots, 3, \tag{11}$$

where W_d is a pointwise convolution and GN_d is group normalization with one group. The spectral branch acts along the spatial coordinate as

$$\widehat{\mathcal{K}_d v}(k) = \begin{cases} \Theta^{\mathrm{F}}_{d,k} \widehat{v}(k), & 0 \le k < 16, \\ 0, & 16 \le k \le \lfloor N_x/2 \rfloor, \end{cases} \qquad \Theta^{\mathrm{F}}_{d,k} \in \mathbb{C}^{64\times 64}. \tag{12}$$

The local branch and residual connection remain active at every grid point. A $64 \to 64$ pointwise convolution with GELU followed by a $64 \to 2$ pointwise convolution produces the normalized output. Reversing the target normalization gives $\widetilde{r}_i(t, x) = \mathcal{G}_\omega(\aleph_t)$. The FNO has 287,362 trainable parameters and no recurrent hidden state; all temporal information is supplied through the finite histories and timing descriptors in $\aleph_t$.

A.3 Zero-Consistent Deployment and Training

The pointwise gate stated in Section 3.1 is

$$g_m(t, x) = \frac{\|C(Y_m(t, x))\|_2}{\|C(Y_m(t, x))\|_2 + \tau_g}, \qquad \tau_g = 0.0025. \tag{13}$$

Closed-loop deployment additionally uses the most recently transmitted state through

$$a_e(t) = \frac{\left\|C\big(\bar{Y}(t, \cdot)\big)\right\|_{L_2(\Omega)}}{\sqrt{|\Omega|}}. \tag{14}$$

and

$$g_s(t) = \begin{cases} 0, & a_e(t) \le \tau_g, \\ 3\left(\dfrac{a_e(t) - \tau_g}{\tau_g}\right)^2 - 2\left(\dfrac{a_e(t) - \tau_g}{\tau_g}\right)^3, & \tau_g < a_e(t) < 2\tau_g, \\ 1, & a_e(t) \ge 2\tau_g. \end{cases} \tag{15}$$

The compensation entering controller 7 is therefore

$$s_i(t, x) = g_s(t) g_m(t, x) \widetilde{r}_i(t, x). \tag{16}$$

Both gates depend only on controller-side quantities. In particular, g_s is constant between transmissions and removes the learned compensation once the last transmitted state enters the numerical zero region. The values of $s_i(t, x)$ at the two boundary nodes are set to zero.

The held-out prediction study in Section 5.2 evaluates $g_m \widetilde{r}_i$; the additional factor g_s is used only in closed-loop deployment.

The 120 trajectories are divided at the trajectory level into 84 training, 18 validation, and 18 test trajectories. Input and target normalization, target-magnitude quantiles, and sample weights are determined from the training subset. For training sample q, the objective is

$$\mathcal{L}_q = w_q \left[\mathcal{L}_{\text{field}}^{(q)} + 0.1500 \, \frac{\| g_m^{(q)} \widetilde{r}_i^{(q)} - r_i^{(q)} \|_{L^2(\Omega)}^2}{\| r_i^{(q)} \|_{L^2(\Omega)}^2 + (0.0082)^2} \right], \tag{17}$$

where $\mathcal{L}_{\text{field}}^{(q)}$ is the mean-squared field error after target normalization. The sample weight w_q combines the mode weights, transient weight, and target-magnitude weights

$$\begin{gathered} (0.7834, 1.0834, 1.2491), \qquad 1.2500 \text{ for } t \le 5\,\text{s}, \\ (1.0000, 1.2500, 1.7500, 2.5000), \qquad (Q_{50}, Q_{75}, Q_{95}) = (0.0010, 0.0082, 0.0888). \end{gathered} \tag{18}$$

The four magnitude weights correspond to the intervals separated by the three training-set quantiles. The product of the three weighting factors is capped at 6.0000 and normalized to unit mean over the training subset.

Training uses Adam with batch size 128, initial learning rate 10^{-3}, weight decay 10^{-5}, at most 150 epochs, and an early-stopping patience of 20 epochs. The deployed checkpoint uses seed 20260901. At each epoch, model selection minimizes the overall FNO-to-zero-predictor validation-error ratio plus 0.5000 times the corresponding large-target ratio and 0.2500 times the ratio for the least accurate physical mode. Test trajectories are excluded from normalization, optimization, early stopping, and checkpoint selection.

A.4 Comparison Architectures and Statistical Evaluation

All learned comparisons use the same 58-channel input, prediction target, trajectory split, normalization, loss, sample weights, optimizer, stopping rule, validation criterion, and model gate. Only the backbone is changed. The 1D U-Net has two encoder and two decoder levels with base width 48. Each level contains two width-three convolutions, group normalization, and GELU; downsampling uses average pooling, and upsampling uses linear interpolation and skip concatenation. The CNN–DeepONet branch contains two width-five convolutions with 64 channels, adaptive pooling to 15 grid points, and fully connected layers of widths 192 and 96. Its coordinate trunk has widths 64, 128, and 96. The 96 branch and trunk features are split into two latent groups of width 48, whose inner products yield the two output components. The zero predictor returns the zero field, whereas the latest-interval predictor reuses the most recent valid residual anchor and returns zero before the first such anchor is available. Both baselines are deterministic and have no trainable parameters.

Each learned architecture is trained with seeds 20260901–20260903, and one checkpoint is selected per seed using validation trajectories only. Table 7 reports the mean and standard deviation across the three runs.

Table 7: Multi-seed validation results. Lower values are better.

Architecture	Parameters	Validation score	Mean spatial-L^2 error
FNO	287,362	0.1467 ± 0.0043	0.0033 ± 0.0001
1D U-Net	363,842	0.1452 ± 0.0080	0.0033 ± 0.0002
CNN–DeepONet	263,042	0.1669 ± 0.0076	0.0037 ± 0.0002

The test set contains 2,247, 1,697, and 1,474 field samples from modes 1–3, respectively. For every sample, the error is the spatial $L^2(\Omega)$ norm of the prediction residual. Table 8 reports its mean and 95th percentile within each mode.

Bootstrap resampling is performed over complete test trajectories to preserve within-trajectory dependence. For each architecture pair, the mean spatial-L^2 error is first computed for each of the 18 test trajectories. The paired trajectory-level differences are then

Table 8: Mode-wise held-out errors (mean/95th percentile).

Method	Mode 1	Mode 2	Mode 3
Zero predictor	0.0293/0.0592	0.0456/0.1028	0.0481/0.1155
Latest-interval predictor	0.0237/0.1018	0.0361/0.1936	0.0331/0.1365
FNO	0.0021/0.0070	0.0043/0.0129	0.0038/0.0110
1D U-Net	0.0023/0.0078	0.0041/0.0139	0.0040/0.0156
CNN–DeepONet	0.0026/0.0077	0.0049/0.0175	0.0040/0.0191

resampled with replacement, 18 at a time, for 10,000 replicates using NumPy seed 20260917. The percentile interval is defined by the 2.5th and 97.5th percentiles. Positive differences in Table 9 favor FNO.

Table 9: Paired trajectory-level error differences ($\times 10^{-3}$).

Difference	Mean	95% interval
1D U-Net − FNO	0.1060	[−0.2300, 0.4500]
CNN–DeepONet − FNO	0.4740	[0.0370, 0.9200]

The first interval contains zero; hence, these experiments do not establish a systematic accuracy difference between FNO and U-Net. The second interval is positive for this benchmark, indicating a lower trajectory-averaged error for FNO than for CNN–DeepONet under the matched protocol.

A.5 Simulation Configuration and Evaluation Protocols

Plant, uncertainty, and switching. We use three modes, two state components, $\Omega = (-1, 5)$, homogeneous Dirichlet boundary conditions, and the componentwise activation $C(Y) = \tanh(Y)$. The plant matrices are

i	M_i	N_i	B_i
1	$\begin{bmatrix} 0.8000 & 0 \\ 0 & 0.8000 \end{bmatrix}$	$\begin{bmatrix} 0.3000 & 0.6000 \\ 0.6000 & 0.3000 \end{bmatrix}$	$\begin{bmatrix} 0.6600 & -0.0330 \\ -1.6500 & 0.9900 \end{bmatrix}$
2	$\begin{bmatrix} 0.7500 & 0 \\ 0 & 0.8500 \end{bmatrix}$	$\begin{bmatrix} 0.2000 & 0.4000 \\ 0.4000 & 0.2000 \end{bmatrix}$	$\begin{bmatrix} 0.6000 & -0.0500 \\ -1.4500 & 0.9000 \end{bmatrix}$
3	$\begin{bmatrix} 0.8500 & 0 \\ 0 & 0.7500 \end{bmatrix}$	$\begin{bmatrix} 0.5000 & 0.1000 \\ 0.1000 & 0.5000 \end{bmatrix}$	$\begin{bmatrix} 0.7200 & -0.0200 \\ -1.5500 & 0.8400 \end{bmatrix}$

The matched perturbation is $\Delta B_i(t) = U_i D_i(t) S_i$, where $D_i(t) = q_i(t)\bar{D}_i$ and $q_i(t) = 0.9000 + 0.1000 \sin(\omega_i t + \phi_i)$. Its matrices are

i	U_i	S_i	$\bar{D}_i$
1	$\begin{bmatrix} 0.7200 & 0 \\ 0 & 0.6000 \end{bmatrix}$	$\begin{bmatrix} 0.8000 & 0.1500 \\ -0.1000 & 0.7000 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
2	$\begin{bmatrix} 0.6400 & 0 \\ 0 & 0.7600 \end{bmatrix}$	$\begin{bmatrix} 0.7500 & -0.1200 \\ 0.0800 & 0.8200 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
3	$\begin{bmatrix} 0.8000 & 0 \\ 0 & 0.5600 \end{bmatrix}$	$\begin{bmatrix} 0.8500 & 0.1000 \\ -0.1500 & 0.7600 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

For each trajectory, $\omega_i \sim \mathcal{U}[0.1500, 0.3000]$ and $\phi_i \sim \mathcal{U}[0, 2\pi]$ are sampled independently for each mode and then held fixed. The residence threshold is $\vartheta^\star = 0.3000\,\mathrm{s}$. With $\Sigma^{(v)} = [\sigma_{ij}^{(v)}]$ for $v \in \{\mathrm{s}, \mathrm{l}\}$, the switching parameters are

$$(\zeta_1, \zeta_2, \zeta_3) = (1.7725, 2.7082, 3.6763),$$

$$\Sigma^{(\mathrm{s})} = \begin{bmatrix} 0 & 0.2000 & 0.8000 \\ 0.5000 & 0 & 0.5000 \\ 0.3000 & 0.7000 & 0 \end{bmatrix}, \qquad \Sigma^{(\mathrm{l})} = \begin{bmatrix} 0 & 0.3000 & 0.7000 \\ 0.4000 & 0 & 0.6000 \\ 0.5000 & 0.5000 & 0 \end{bmatrix}.$$

Triggering and controller parameters. The mode-dependent scalars and empirically calibrated residual levels are listed in Table 10; $\eta(0) = \eta_0 = 0.1000$. The feasible matrices

Table 10: Dynamic-triggering and sliding-mode parameters.

i	μ_i	ξ_i	ν_i	λ_i	θ_i	δ_i	$\bar{\epsilon}_i$
1	0.1000	0.5000	0.2000	0.6000	0.4500	0.0500	0.2023
2	0.1200	0.4500	0.2500	0.6500	0.5000	0.0500	0.4602
3	0.0800	0.5500	0.3000	0.7000	0.5500	0.0500	0.3000

used in both residence regimes are

$$H_1 = \begin{bmatrix} 12.7537 & -1.9980 \\ -1.9980 & 13.0087 \end{bmatrix}, \quad K_1 = \begin{bmatrix} -4.4857 & 0.9104 \\ 1.5621 & -5.2008 \end{bmatrix},$$
$$H_2 = \begin{bmatrix} 12.3373 & -1.3436 \\ -1.3436 & 13.0550 \end{bmatrix}, \quad K_2 = \begin{bmatrix} -4.3073 & 0.7205 \\ 1.2908 & -5.6089 \end{bmatrix},$$
$$H_3 = \begin{bmatrix} 11.8965 & -0.6660 \\ -0.6660 & 11.4676 \end{bmatrix}, \quad K_3 = \begin{bmatrix} -4.6005 & 0.4593 \\ 1.0237 & -4.7001 \end{bmatrix}.$$

Numerical implementation and data generation. The spatial grid has 61 nodes ($\Delta x = 0.1000$), and $\Delta t = 0.0050$ s. Diffusion is advanced implicitly, whereas the reaction and control terms are treated explicitly; spatial norms and energies use the composite trapezoidal rule. A step containing a sampled mode switch is split at the switching instant. Trigger crossings are checked on the numerical time grid, and the auxiliary state is advanced by forward Euler with floor 10^{-10}. In simulation, the set-valued direction in controller 7 is regularized as

$$S_J(t,x) = \frac{J(t,x)}{\|J(t,\cdot)\|_{L^2(\Omega)} + \varepsilon_{\mathrm{sm}}}, \qquad \varepsilon_{\mathrm{sm}} = 0.0020.$$

For the representative closed-loop trajectories, $\alpha(0) = 1$ and

$$Y(0,x) = \begin{bmatrix} 3 \\ -4 \end{bmatrix} \cos\left(\frac{\pi(x-2)}{6}\right), \qquad x \in [-1, 5].$$

The nominal and held states use the same initialization. The plotted 20 s segment is taken from a 30 s rollout with seed 202608302 and

$$(\omega_1, \omega_2, \omega_3) = (0.2960, 0.2931, 0.2797), \qquad (\phi_1, \phi_2, \phi_3) = (6.1884, 5.5288, 3.2146).$$

The training corpus contains 120 trajectories of length 15 s, with 40 trajectories per initial mode. For $k = 1, 2$,

$$Y_k(0,x) = a_k \sin\left(\frac{\pi(x+1)}{6}\right) + \sum_{h=2}^{4} \frac{0.2500|a_k|\beta_{kh}}{h} \sin\left(\frac{h\pi(x+1)}{6}\right),$$

where $a_1 \sim \mathcal{U}[2.4000, 3.6000]$, $a_2 \sim \mathcal{U}[-4.8000, -3.2000]$, and $\beta_{kh} \sim \mathcal{N}(0,1)$. The MATLAB Twister seed is 20260827; fields are stored every 0.0500 s, and residual anchors require an interval of at least 0.0200 s. The split, architecture, and training settings are given in Appendices A.3–A.4.

Residual calibration and control baselines. With the FNO and both gates frozen, the mode-wise level used in simulation is

$$\bar{\epsilon}_i = 1.2500 \max_{q \in \mathcal{V}_i} \left\| (r_i^{\mathrm{e}})^{(q)} \right\|_{L^2(\Omega)},$$

where $\mathcal{V}_i$ contains validation samples in mode i. These held-out checks assess the sampled test distribution and do not certify the uniform bound in Assumption 2. The SMC baseline without neural compensation sets $s_i = 0$ and uses

$$\bar{r}_i = \sqrt{2|\Omega|}\, \|U_i\|_2 \|S_i\|_2, \qquad (\bar{r}_1, \bar{r}_2, \bar{r}_3) = (2.0358, 2.1895, 2.4003),$$

so that $h_i^{\text{no-NO}} = \bar{r}_i + \chi_i + \delta_i$, whereas the proposed controller uses $h_i^{\mathrm{NO}} = \bar{\epsilon}_i + \chi_i + \delta_i$. The two cases otherwise share the same realization, gains, trigger, and numerical regularization.

Table 11: Mode-wise residual calibration and held-out checks.

Mode	Validation max.	$\bar{\epsilon}_i$	Test max.	Violations
1	0.1618	0.2023	0.1577	0
2	0.3682	0.4602	0.4025	0
3	0.2400	0.3000	0.2638	0

Evaluation metrics and communication protocol. For the common grid $\mathcal{T} = \{t_\ell\}_{\ell=0}^{L}$, write $Z(t,\cdot) := Z_{\alpha(t)}(t,\cdot)$. The control metrics are

$$\bar{Z} = \frac{1}{T}\int_0^T \|Z(t,\cdot)\|_{L^2(\Omega)}\,dt, \qquad Z_{\mathrm{pk}} = \max_{0\le t\le T}\|Z(t,\cdot)\|_{L^2(\Omega)},$$
$$E_Z = \int_0^T \|Z(t,\cdot)\|^2_{L^2(\Omega)}\,dt, \qquad TV_Z = \sum_{\ell=0}^{L-1}\|Z(t_{\ell+1},\cdot) - Z(t_\ell,\cdot)\|_{L^2(\Omega)}.$$

The SETM removes only the dynamic term from the proposed trigger:

$$\Gamma_i^{\mathrm{SETM}}(t) = \mu_i F_i(t) X_i(s_e) - E_i(t), \qquad \Gamma_i^{\mathrm{DETM}}(t) = \Gamma_i^{\mathrm{SETM}}(t) + \eta(t)/\theta_i.$$

All other settings, including the frozen FNO, remain unchanged. The comparison uses 100 paired 30 s realizations with seeds $20260908 + 104729p$, $p = 1,\dots,100$. Within each pair, DETM and SETM share the initial field, semi-Markov path, uncertainty realization, and numerical grid; the common packet at $t = 0$ is excluded. Confidence intervals resample the 100 complete pairs with replacement for 5000 replicates and use the 0.0250 and 0.9750 quantiles. The bootstrap seeds are 20261899, 20261900, 20261901, and 20261902 for the paired saving, DETM mean, SETM mean, and mean pathwise reduction, respectively.

A.6 Additional Closed-Loop Diagnostics

The following figures complement the main-text closed-loop comparison by separating three aspects of the controller: state regulation, the spatiotemporal distribution of the control input, and the effect of neural compensation on the robust switching action. These figures are intended as qualitative diagnostics; the quantitative comparisons are reported in Section 5.

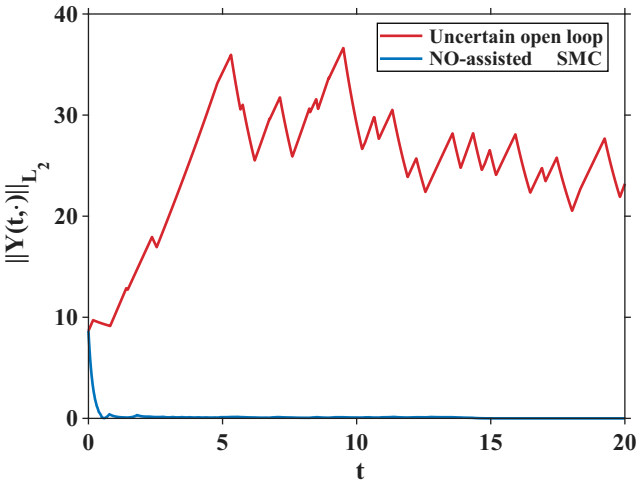


Figure 7: State regulation under matched uncertainty. The uncertain open-loop state remains away from the origin, whereas the proposed NO-assisted controller drives $\|Y(t,\cdot)\|_{L_2(\Omega)}$ toward zero.

State regulation. Figure 7 compares the state norm of the uncertain open-loop system with that of the proposed NO-assisted closed loop. Under the representative uncertainty realization, the open-loop state departs from the origin and remains nonzero, whereas the controlled state rapidly approaches the equilibrium. This result provides a direct state-level illustration of the regulation achieved by the proposed controller.

Full-horizon control fields. Figure 8 shows the two distributed control components over the complete simulation horizon. In this supplementary rollout, the controller is enabled at $t = 0$; hence, the large control values near the initial time correspond to the startup action

required to suppress the initial state and matched uncertainty. This figure uses a separate representative closed-loop rollout from Figure 1 and is therefore not intended for pointwise comparison with that figure. Figure 1 isolates the effect of neural compensation under its matched comparison, whereas Figure 8 displays the spatial structure and initial transient of the proposed control input from the beginning of a complete rollout.

Effect on robust actuation. Figure 9 further explains the reduction in control burden. Without neural compensation, the switching amplitude must account for a bound on the full matched uncertainty. With neural compensation, it is constructed from the remaining prediction residual and the event-induced state-holding mismatch. Consequently, both the switching amplitude and the persistent control magnitude are substantially reduced. The first 2 s are omitted from this figure to emphasize the post-transient behavior.

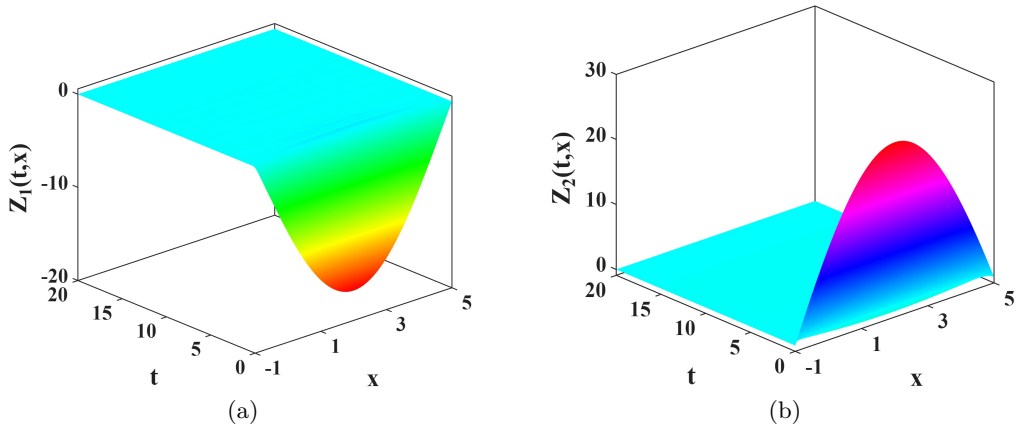


Figure 8: Full-horizon spatiotemporal profiles of the distributed control input in a separate representative rollout, with the controller activated at $t = 0$. (a) First control component $Z_1(t, x)$. (b) Second control component $Z_2(t, x)$.

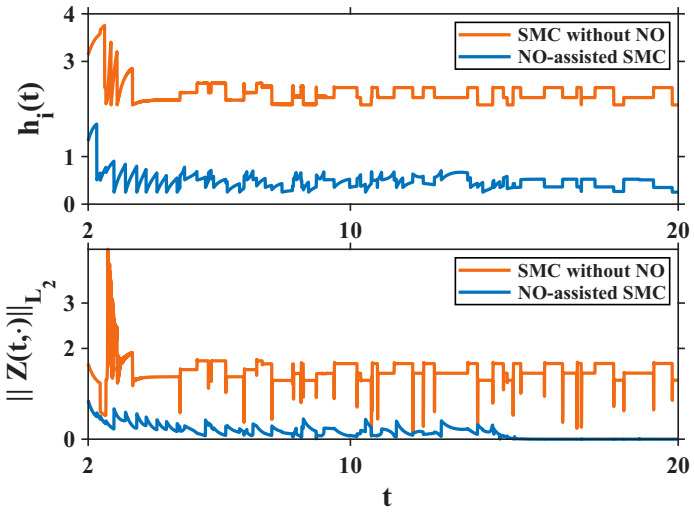


Figure 9: Effect of neural-operator compensation on robust actuation over the post-transient interval. Top: switching amplitude $h_{\alpha(t)}(t)$. Bottom: distributed-control magnitude $\|Z(t, \cdot)\|_{L_2(\Omega)}$. The paired simulations share the same initial condition, semi-Markov path, uncertainty realization, and DETM parameters, and differ only in whether neural compensation is used.

B Proofs and Controller Synthesis

This appendix proves Theorems 1–3 and gives the variable substitution used in Theorem 4. All symbols follow Sections 2–4.

B.1 Exclusion of Zeno Behavior

For $t \in [s_e, s_{e+1})$, $\Gamma_i(t) > 0$ implies $E_i(t) < \mu_i F_i(t) X_i(s_e) + \eta(t)/\theta_i$. Hence

$$\dot{\eta}(t) > -(\lambda_i + \theta_i^{-1})\eta(t), \qquad \eta(t) \geq \eta_0 \exp\left[-\max_{i \in \mathcal{S}}(\lambda_i + \theta_i^{-1})t\right] > 0. \tag{19}$$

Immediately after a transmission, $m(s_e^+, \cdot) = 0$, and therefore $E_i(s_e^+) = 0$, $F_i(s_e^+) = \xi_i$, and

$$\Gamma_i(s_e^+) = \mu_i \xi_i X_i(s_e) + \frac{\eta(s_e)}{\theta_i} > 0. \tag{20}$$

On every compact interval without a mode transition, $\dot{E}_i = 2\langle m, H_i Y_t \rangle_{L^2}$ and $\dot{F}_i = -\nu_i F_i \dot{E}_i$. The solution regularity therefore makes $\dot{E}_i$, $\dot{F}_i$, and $\dot{\Gamma}_i = \mu_i \dot{F}_i X_i(s_e) - \dot{E}_i + \dot{\eta}/\theta_i$ bounded. Equations 19–20 then imply a positive inter-event time on every bounded interval. A mode transition may cause the trigger to be re-evaluated and satisfied at the transition instant, but the semi-Markov process has only finitely many transitions on every bounded interval. Thus, transmission instants cannot accumulate in finite time.

B.2 Proof of Theorem 1

For $\|J(t,\cdot)\|_{L^2(\Omega)} > 0$, one has $S_J = J/\|J\|_{L^2(\Omega)}$. Using $J_t = r_{e_i} - K_i m - h_i(t) S_J$, Assumption 2, and $\|K_i m\|_{L^2(\Omega)} \leq \chi_i(t)$ gives

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|J\|_{L^2(\Omega)}^2 &= \langle J, r_i^{\mathrm{e}} - K_i m - h_i(t) S_J \rangle_{L^2} \\ &\leq \big(\bar{\epsilon}_i + \chi_i(t) - h_i(t)\big)\|J\|_{L^2(\Omega)} = -\delta_i \|J\|_{L^2(\Omega)}. \end{aligned} \tag{21}$$

Consequently, $d\|J\|_{L^2(\Omega)}/dt \leq -\delta$ almost everywhere before reaching, where $\delta = \min_i \delta_i > 0$. Since J is continuous at mode transitions, integration of equation 21 yields

$$\|J(t,\cdot)\|_{L^2(\Omega)} \leq \|J(0,\cdot)\|_{L^2(\Omega)} - \delta t, \qquad T_r \leq \bar{T}_r := \frac{\|J(0,\cdot)\|_{L^2(\Omega)}}{\delta}. \tag{22}$$

At $J = 0$, select $S_J = (r_i^{\mathrm{e}} - K_i m)/h_i(t)$. Then $\|S_J\|_{L^2(\Omega)} \leq (\bar{\epsilon}_i + \chi_i(t))/h_i(t) < 1$, so $S_J \in \mathrm{Sgn}_\Omega(0)$ and $J_t = 0$. Thus, the ideal sliding manifold is reached within $\bar{T}_r$ and remains invariant thereafter.

B.3 Proof of Theorem 2

For the stopped time $\tau_t = t \wedge T_r$, define

$$V_R(t,i) = \int_\Omega Y^\top(t,x) P_i Y(t,x)\, dx + \int_\Omega Y_x^\top(t,x) R Y_x(t,x)\, dx + \eta(t).$$

For $v \in \{\mathrm{s}, \mathrm{l}\}$, diagonal $\Xi_{R,i} \succ 0$, and $\varepsilon_{r_i}, \varepsilon_{g_i} > 0$, define the reaching-phase matrix

$$\Phi_{R,i}^{(v)} = \begin{bmatrix} \Phi_{11}^{(v)} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} \\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} \\ * & * & \Phi_{33} & 0 & 0 \\ * & * & * & \Phi_{44} & 0 \\ * & * & * & * & \Phi_{55} \end{bmatrix}, \tag{23}$$

where

$$\Phi_{11}^{(v)} = \zeta_i\left(\sum_{j\neq i}\sigma_{ij}^{(v)}P_j - P_i\right) + b_R P_i - (2\gamma_R - b_R)\frac{\pi^2}{(b-a)^2}R + \gamma_R\,\mathrm{He}\{G_i(K_i - N_i)\} + [1 + (\lambda_i - b_R)\theta_i]\mu_i\xi_i H_i - \mathrm{He}\{(B^-)^\top \Xi_{R,i}B^+\},$$

$$\Phi_{12} = P_i - \gamma_R G_i + (K_i - N_i)^\top G_i^\top,$$

$$\Phi_{13} = \gamma_R G_i B_i + (B^- + B^+)^\top \Xi_{R,i},$$

$$\Phi_{14} = -\gamma_R G_i K_i - [1 + (\lambda_i - b_R)\theta_i]\mu_i\xi_i H_i,$$

$$\Phi_{15} = -\gamma_R G_i, \qquad \Phi_{22} = -\mathrm{He}\{G_i\}, \qquad \Phi_{23} = G_i B_i,$$

$$\Phi_{24} = -G_i K_i, \qquad \Phi_{25} = -G_i, \qquad \Phi_{33} = -2\Xi_{R,i},$$

$$\Phi_{44} = -[1 + (\lambda_i - b_R)\theta_i](1 - \mu_i\xi_i)H_i, \qquad \Phi_{55} = -\varepsilon_{gi}I.$$

The LMI in Theorem 2 is

$$\mathcal{M}_{R,i}^{(v)} = \begin{bmatrix} \Phi_{R,i}^{(v)} & \mathrm{col}\{\gamma_R G_i, G_i, 0, 0, 0\} \\ * & -\varepsilon_{ri}I \end{bmatrix} \prec 0, \qquad i \in \mathcal{S}, \quad v \in \{\mathrm{s}, \mathrm{l}\}. \tag{24}$$

Before reaching, substitution of controller 7 into system (1) gives

$$Y_t = M_i Y_{xx} + (K_i - N_i)Y + B_i C(Y) + r_{e_i} - K_i m - h_i(t)S_J. \tag{25}$$

Let $\tau_t = t \wedge T_r$. For $t < T_r$, the mode-transition contribution to $\mathcal{L}V_R(t,i)$ is

$$\int_\Omega Y^\top \zeta_i \left(\sum_{j\neq i}\sigma_{ij}^{(v)}P_j - P_i\right) Y\,dx. \tag{26}$$

Differentiating the continuous-state terms and adding $b_R V_R$ yield

$$\begin{aligned}\mathcal{L}V_R + b_R V_R = {} & 2\int_\Omega Y^\top P_i Y_t\,dx + 2\int_\Omega Y_x^\top R Y_{xt}\,dx \\ & + \int_\Omega Y^\top \left[\zeta_i\left(\sum_{j\neq i}\sigma_{ij}^{(v)}P_j - P_i\right) + b_R P_i\right] Y\,dx \\ & + b_R\int_\Omega Y_x^\top R Y_x\,dx + \dot\eta + b_R\eta.\end{aligned} \tag{27}$$

Since $Y_t = 0$ on $\partial\Omega$, integration by parts gives

$$2\int_\Omega Y_x^\top R Y_{xt}\,dx = -2\int_\Omega Y_t^\top R Y_{xx}\,dx. \tag{28}$$

Using $G_i M_i = R$ and equation 25, add

$$\begin{aligned}0 = 2\int_\Omega (\gamma_R Y + Y_t)^\top G_i\Big[& M_i Y_{xx} - Y_t + (K_i - N_i)Y + B_i C(Y) \\ & + r_{e_i} - K_i m - h_i(t)S_J\Big]\,dx\end{aligned} \tag{29}$$

to the preceding equality. The resulting diffusion term satisfies

$$\begin{aligned}& -(2\gamma_R - b_R)\int_\Omega Y_x^\top R Y_x\,dx \\ & \qquad \le -(2\gamma_R - b_R)\frac{\pi^2}{(b-a)^2}\int_\Omega Y^\top R Y\,dx.\end{aligned} \tag{30}$$

Before the next transmission, $\eta(t) > -\theta_i[\mu_i F_i(t)X_i(s_e) - E_i(t)]$. Since $\lambda_i > b_R$,

$$\dot\eta + b_R\eta < [1 + (\lambda_i - b_R)\theta_i][\mu_i F_i(t)X_i(s_e) - E_i(t)]. \tag{31}$$

Using $F_i(t) \le \xi_i$ and $\bar{Y} = Y - m$ gives

$$\mu_i F_i(t) X_i(s_e) - E_i(t) \le \int_\Omega \big[\mu_i \xi_i Y^\top H_i Y - \mu_i \xi_i \,\mathrm{He}\{Y^\top H_i m\} - (1 - \mu_i \xi_i) m^\top H_i m\big] dx. \tag{32}$$

The activation-sector condition further implies

$$\mathrm{He}\big\{[C(Y) - B^- Y]^\top \Xi_{R,i} [B^+ Y - C(Y)]\big\} \ge 0. \tag{33}$$

After expansion, equations 30– 33 generate the first four rows and columns of $\Phi^{(v)}_{R,i}$.

For the prediction residual, Young's inequality gives

$$2\int_\Omega (\gamma_R Y + Y_t)^\top G_i r_i^{\mathrm{e}}\, dx \le \varepsilon_{ri}^{-1} \int_\Omega (\gamma_R Y + Y_t)^\top G_i G_i^\top (\gamma_R Y + Y_t)\, dx + \varepsilon_{ri} \bar{\epsilon}_i^2. \tag{34}$$

The robust term can be written as

$$\begin{aligned} &-2\int_\Omega (\gamma_R Y + Y_t)^\top G_i h_i(t) S_J\, dx \\ &= \int_\Omega \begin{bmatrix} \gamma_R Y + Y_t \\ h_i(t) S_J \end{bmatrix}^\top \begin{bmatrix} 0 & -G_i \\ -G_i^\top & -\varepsilon_{gi} I \end{bmatrix} \begin{bmatrix} \gamma_R Y + Y_t \\ h_i(t) S_J \end{bmatrix} dx + \varepsilon_{gi} h_i^2(t), \end{aligned} \tag{35}$$

where $\|S_J\|_{L^2(\Omega)} = 1$ for $t < T_r$. Collecting the preceding estimates and applying the Schur complement to equation 24 gives

$$\mathcal{L} V_R \le -b_R V_R + \varepsilon_{ri} \bar{\epsilon}_i^2 + \varepsilon_{gi} h_i^2(t). \tag{36}$$

It remains to estimate $h_i(t)$. The event inequality yields

$$\left(\int_\Omega m^\top H_i m\, dx\right)^{1/2} \le \sqrt{\mu_i \xi_i} \left(\int_\Omega {}^\top H_i\, dx\right)^{1/2} + \theta_i^{-1/2} \sqrt{\eta}. \tag{37}$$

Since $\bar{Y} = Y - m$ and $0 < \mu_i \xi_i < 1$,

$$\left(\int_\Omega m^\top H_i m\, dx\right)^{1/2} \le \frac{\sqrt{\mu_i \xi_i}}{1 - \sqrt{\mu_i \xi_i}} \left(\int_{{}^\top H_i Y\, dx}\right)^{1/2} + \frac{\sqrt{\eta}}{\sqrt{\theta_i}(1 - \sqrt{\mu_i \xi_i})}. \tag{38}$$

Using the definition of $\chi_i(t)$ in Section 3.3 then gives

$$\chi_i(t) \le \frac{\bar{k}_i \sqrt{\mu_i \xi_i \lambda_{\max}(H_i)}}{1 - \sqrt{\mu_i \xi_i}} \|Y\|_{L^2(\Omega)} + \frac{\bar{k}_i}{\sqrt{\theta_i}(1 - \sqrt{\mu_i \xi_i})} \sqrt{\eta}. \tag{39}$$

Consequently,

$$h_i^2(t) \le 3(\bar{\epsilon}_i + \delta_i)^2 + \frac{3\bar{k}_i^2 \mu_i \xi_i \lambda_{\max}(H_i)}{(1 - \sqrt{\mu_i \xi_i})^2} \|Y\|^2_{L^2(\Omega)} + \frac{3\bar{k}_i^2}{\theta_i (1 - \sqrt{\mu_i \xi_i})^2} \eta. \tag{40}$$

Let $p_R = \min_i \lambda_{\min}(P_i)$ and set

$$\begin{aligned} c_R &= \max_i \left\{ -b_R + \frac{3\varepsilon_{gi} \bar{k}_i^2}{(1 - \sqrt{\mu_i \xi_i})^2} \left[\frac{\mu_i \xi_i \lambda_{\max}(H_i)}{p_R} + \frac{1}{\theta_i} \right] \right\}, \\ d_R &= \max_i \left\{ \varepsilon_{ri} \bar{\epsilon}_i^2 + 3\varepsilon_{gi} (\bar{\epsilon}_i + \delta_i)^2 \right\}. \end{aligned} \tag{41}$$

Equations 36–40 imply $\mathcal{L} V_R \le c_R V_R + d_R$. Applying Dynkin's formula to the stopped process gives

$$\mathbb{E} V_R(\tau_t, \alpha(\tau_t)) \le V_R(0, \alpha(0)) + \int_0^t [c_R \mathbb{E} V_R(\tau_s, \alpha(\tau_s)) + d_R] ds. \tag{42}$$

Gronwall's inequality therefore yields

$$\mathbb{E} V_R(\tau_t, \alpha(\tau_t)) \le e^{c_R^+ t} [V_R(0, \alpha(0)) + d_R t], \qquad c_R^+ = \max\{c_R, 0\}. \tag{43}$$

Since $T_r \le \bar{T}_r < \infty$ and $P_i, R \succ 0$, V_R bounds both $\|Y\|^2_{L^2(\Omega)}$ and $\|Y_x\|^2_{L^2(\Omega)}$ uniformly on $[0, \bar{T}_r]$, proving Theorem 2.

B.4 Proof of Theorem 3

On the sliding manifold, define

$$V_S(t,i) = \int_\Omega Y^\top(t,x) P_i Y(t,x)\,dx + \int_\Omega Y_x^\top(t,x) R Y_x(t,x)\,dx + \eta(t),$$

$J = J_t = 0$, so $Z_i^{\mathrm{eq}} = K_i Y - r_i$ and

$$Y_t = M_i Y_{xx} + (K_i - N_i)Y + B_i C(Y). \tag{44}$$

For $v \in \{\mathrm{s}, \mathrm{l}\}$ and diagonal $\Xi_{S,i} \succ 0$, define

$$\Phi_{S,i}^{(v)} = \begin{bmatrix} \Psi_{11}^{(v)} & \Psi_{12} & \Psi_{13} & \Psi_{14} \\ * & \Psi_{22} & \Psi_{23} & 0 \\ * & * & \Psi_{33} & 0 \\ * & * & * & \Psi_{44} \end{bmatrix}, \tag{45}$$

where

$$\begin{aligned}
\Psi_{11}^{(v)} &= \zeta_i \left(\sum_{j \neq i} \sigma_{ij}^{(v)} P_j - P_i \right) + b_S P_i - (2\gamma_S - b_S) \frac{\pi^2}{(b-a)^2} R \\
&\quad + \gamma_S \,\mathrm{He}\{G_i(K_i - N_i)\} + [1 + (\lambda_i - b_S)\theta_i]\mu_i \xi_i H_i - \mathrm{He}\{(B^-)^\top \Xi_{S,i} B^+\}, \\
\Psi_{12} &= P_i - \gamma_S G_i + (K_i - N_i)^\top G_i^\top, \\
\Psi_{13} &= \gamma_S G_i B_i + (B^- + B^+)^\top \Xi_{S,i}, \\
\Psi_{14} &= -[1 + (\lambda_i - b_S)\theta_i]\mu_i \xi_i H_i, \\
\Psi_{22} &= -\mathrm{He}\{G_i\}, \qquad \Psi_{23} = G_i B_i, \qquad \Psi_{33} = -2\Xi_{S,i}, \\
\Psi_{44} &= -[1 + (\lambda_i - b_S)\theta_i](1 - \mu_i \xi_i) H_i.
\end{aligned}$$

Differentiating $V_S(t,i)$, integrating the gradient term by parts, and using $G_i M_i = R$, add

$$0 = 2 \int_\Omega (\gamma_S Y + Y_t)^\top G_i [M_i Y_{xx} - Y_t + (K_i - N_i)Y + B_i C(Y)] dx. \tag{46}$$

The mode-transition contribution is

$$\int_\Omega Y^\top \zeta_i \left(\sum_{j \neq i} \sigma_{ij}^{(v)} P_j - P_i \right) Y\,dx. \tag{47}$$

Consequently,

$$\begin{aligned}
\mathcal{L}V_S + b_S V_S &= 2 \int_\Omega Y^\top P_i Y_t\,dx + 2 \int_\Omega Y_x^\top R Y_{xt}\,dx \\
&\quad + \int_\Omega Y^\top \left[\zeta_i \left(\sum_{j \neq i} \sigma_{ij}^{(v)} P_j - P_i \right) + b_S P_i \right] Y\,dx \\
&\quad + b_S \int_\Omega Y_x^\top R Y_x\,dx + \dot{\eta} + b_S \eta.
\end{aligned} \tag{48}$$

The boundary condition and integration by parts give

$$2 \int_\Omega Y_x^\top R Y_{xt}\,dx = -2 \int_\Omega Y_t^\top R Y_{xx}\,dx. \tag{49}$$

Combining this identity with equation equation 46 and applying Poincaré's inequality yields

$$\begin{aligned}
&-(2\gamma_S - b_S) \int_\Omega Y_x^\top R Y_x\,dx \\
&\qquad \leq -(2\gamma_S - b_S) \frac{\pi^2}{(b-a)^2} \int_\Omega Y^\top R Y\,dx.
\end{aligned} \tag{50}$$

Since $b_S < \lambda_i$ and $\Gamma_i(t) > 0$ before the next transmission,

$$\dot{\eta} + b_S \eta < [1 + (\lambda_i - b_S)\theta_i][\mu_i F_i(t) X_i(s_e) - E_i(t)]. \tag{51}$$

Moreover, $F_i(t) \le \xi_i$ and $\bar{Y} = Y - m$ imply

$$\mu_i F_i(t) X_i(s_e) - E_i(t) \le \int_\Omega \big[\mu_i \xi_i Y^\top H_i Y - \mu_i \xi_i \,\mathrm{He}\{Y^\top H_i m\} - (1 - \mu_i \xi_i) m^\top H_i m\big] dx, \tag{52}$$

whereas the activation-sector condition gives

$$\mathrm{He}\big\{[C(Y) - B^- Y]^\top \Xi_{S,i} [B^+ Y - C(Y)]\big\} \ge 0. \tag{53}$$

Substitution of equations 50–53 into equation 48, followed by collection of the quadratic terms, gives

$$\mathcal{L}V_S + b_S V_S \le \int_\Omega \mathrm{col}\{Y, Y_t, C(Y), m\}^\top \Phi_{S,i}^{(v)} \,\mathrm{col}\{Y, Y_t, C(Y), m\}\, dx. \tag{54}$$

Thus, $\Phi_{S,i}^{(v)} \prec 0$ implies $\mathcal{L}V_S \le -b_S V_S$. Dynkin's formula gives, for every $\tau \ge 0$,

$$\mathbb{E}V_S(T_r + \tau, \alpha(T_r + \tau)) \le e^{-b_S \tau} \mathbb{E}V_S(T_r, \alpha(T_r)). \tag{55}$$

Since $p_S = \min_i \lambda_{\min}(P_i) > 0$,

$$\mathbb{E}\|Y(T_r + \tau, \cdot)\|^2_{L^2(\Omega)} \le \frac{e^{-b_S \tau}}{p_S} \mathbb{E}V_S(T_r, \alpha(T_r)), \tag{56}$$

which proves Theorem 3.

B.5 Decoupling and Gain Recovery

The analysis conditions above are bilinear in G_i and K_i. As introduced in Section 4.4, set $L_i = G_i K_i$. Then

$$G_i(K_i - N_i) = L_i - G_i N_i, \qquad (K_i - N_i)^\top G_i^\top = L_i^\top - N_i^\top G_i^\top. \tag{57}$$

Accordingly, the four K_i-dependent reaching-phase blocks become

$$\begin{aligned}
\widehat{\Phi}_{11}^{(v)} &= \zeta_i \left(\sum_{j \ne i} \sigma_{ij}^{(v)} P_j - P_i \right) + b_R P_i - (2\gamma_R - b_R) \frac{\pi^2}{(b-a)^2} R \\
&\quad + \gamma_R \,\mathrm{He}\{L_i - G_i N_i\} + [1 + (\lambda_i - b_R)\theta_i] \mu_i \xi_i H_i - \mathrm{He}\{(B^-)^\top \Xi_{R,i} B^+\}, \\
\widehat{\Phi}_{12} &= P_i - \gamma_R G_i + L_i^\top - N_i^\top G_i^\top, \\
\widehat{\Phi}_{14} &= -\gamma_R L_i - [1 + (\lambda_i - b_R)\theta_i] \mu_i \xi_i H_i, \qquad \widehat{\Phi}_{24} = -L_i.
\end{aligned}$$

Let $\widehat{\Phi}_{R,i}^{(v)}$ denote equation 23 with these four blocks substituted and all remaining blocks unchanged. The convex reaching-phase condition is

$$\widehat{\mathcal{M}}_{R,i}^{(v)} = \begin{bmatrix} \widehat{\Phi}_{R,i}^{(v)} & \mathrm{col}\{\gamma_R G_i, G_i, 0, 0, 0\} \\ * & -\varepsilon_{ri} I \end{bmatrix} \prec 0. \tag{58}$$

Similarly, the two K_i-dependent sliding-phase blocks are replaced by

$$\begin{aligned}
\widehat{\Psi}_{11}^{(v)} &= \zeta_i \left(\sum_{j \ne i} \sigma_{ij}^{(v)} P_j - P_i \right) + b_S P_i - (2\gamma_S - b_S) \frac{\pi^2}{(b-a)^2} R \\
&\quad + \gamma_S \,\mathrm{He}\{L_i - G_i N_i\} + [1 + (\lambda_i - b_S)\theta_i] \mu_i \xi_i H_i - \mathrm{He}\{(B^-)^\top \Xi_{S,i} B^+\}, \\
\widehat{\Psi}_{12} &= P_i - \gamma_S G_i + L_i^\top - N_i^\top G_i^\top.
\end{aligned}$$

Replacing $\Psi_{11}^{(v)}$ and Ψ_{12} in equation 45 defines $\widehat{\Phi}_{S,i}^{(v)}$. The complete synthesis conditions are

$$G_i M_i = R, \qquad \widehat{\mathcal{M}}_{R,i}^{(v)} \prec 0, \qquad \widehat{\Phi}_{S,i}^{(v)} \prec 0, \qquad i \in \mathcal{S}, \quad v \in \{\mathrm{s}, \mathrm{l}\}. \tag{59}$$

With the triggering scalars and $\gamma_R, \gamma_S, b_R, b_S$ fixed, these constraints are affine in the matrix decision variables. Since $G_i M_i = R$ and $M_i, R \succ 0$, G_i is nonsingular, and the controller gain is recovered as

$$K_i = G_i^{-1} L_i. \tag{60}$$

Substitution of equation 60 into the hatted matrices recovers $\mathcal{M}_{R,i}^{(v)}$ and $\Phi_{S,i}^{(v)}$. Theorems 1–3 therefore apply, completing the proof of Theorem 4.

C Numerical Feasibility of the Synthesis LMIs

For the three-mode example, the complete synthesis conditions in Appendix B were solved with YALMIP (release 20250626) and SDPT3 under MATLAB R2025a. The fixed scalars were $\gamma_R = \gamma_S = 1.0000$, $b_R = 0.1000$, and $b_S = 0.1500$, with $\mathrm{tr}(R) = 1.0000$. Table 12 summarizes the solution check over all modes and both residence regimes.

Table 12: Numerical check of the synthesis conditions.

Diagnostic	Value	Diagnostic	Value
$\max_{i\in\mathcal{S},\, v\in\{\mathrm{s},\mathrm{l}\}} \lambda_{\max}\big(\widehat{\mathcal{M}}_{R,i}^{(v)}\big)$	-0.2403	$\lambda_{\min}(R)$	0.4100
$\max_{i\in\mathcal{S},\, v\in\{\mathrm{s},\mathrm{l}\}} \lambda_{\max}\big(\widehat{\Phi}_{S,i}^{(v)}\big)$	-0.7648	$\min_i \lambda_{\min}(P_i)$	2.5971
$\max_i \lVert G_i M_i - R\rVert_{\mathrm{F}}$	1.7300×10^{-10}	$\min_i \lambda_{\min}(H_i)$	10.8792

The negative maximum eigenvalues verify strict feasibility of both LMI families, while the equality residual confirms $G_i M_i = R$ to numerical precision. These diagnostics document the solution used in the experiments; the analytical guarantees remain those of Theorem 4.

D Related Work

NOs for PDE Learning and Control. NOs learn mappings between function spaces to provide unified approximations for families of parametric PDE solutions. Message-passing, nonuniform-data, and geometry-aware methods have further extended NOs to multiscale dynamics, irregular sampling, and complex spatial domains (Li et al., 2023). NOs have also been extended to stochastic PDEs to learn continuous stochastic spatiotemporal dynamics (Salvi et al., 2022). In control problems, NOs have been applied to motion planning and PDE-constrained optimization (Matada et al., 2025; Cheng et al., 2025), as well as to the approximation of backstepping kernels, control gains, and event-triggered boundary controllers (Bhan et al., 2025; Yuan et al., 2025). These methods primarily learn solution operators, control trajectories, or controller-design mappings. We use an FNO to causally estimate the spatially distributed matched uncertainty and employ its output for feedback compensation.

Reaction–Diffusion Neural Systems. The RDNSs studied here are neural dynamical models with spatially distributed states, rather than neural network approximators for solving generic PDEs. Existing studies have developed fuzzy sampled-data, protocol-based, and multiple-event-triggered control methods in the presence of diffusion effects, memristive characteristics, and semi-Markov switching (Zhang et al., 2021; Liu et al., 2024). These methods typically establish stability conditions using model-based Lyapunov functionals and prescribed uncertainty bounds. We further use the learning module to compensate for the predictable component of the uncertainty, while robust control handles the remaining prediction error.

Dynamic Event-Triggered Communication. Event-triggered control has been applied to semilinear diffusion PDEs and boundary-actuated reaction–diffusion PDEs to reduce unnecessary state transmissions or control updates (Selivanov & Fridman, 2016; Espitia et al., 2021); finite-dimensional observer-based event-triggered output feedback has also been applied to reaction–diffusion systems (Lhachemi, 2024). Recent studies have used neural networks to optimize triggering instants in finite-dimensional systems (Yang et al., 2025), or DeepONet to approximate the backstepping kernels and associated gains in dynamic event-triggered boundary control (Yuan et al., 2025). We introduce a modulation term driven by the current state-holding error, allowing the dynamic triggering condition to adjust to the transmission mismatch, and explicitly account for the corresponding communication error in the closed-loop analysis.

Residence-Time-Dependent Semi-Markov Switching. Unlike Markov jump models with memoryless sojourn times, semi-Markov models allow mode-transition behavior to vary with the time that the system has spent in its current mode. Existing studies have considered incomplete sojourn-time and transition information and hidden-mode evolution (Ning et al., 2020a; Zhang et al., 2020), and have further investigated bi-boundary sojourn-time constraints and nonhomogeneous transition behavior (Yang et al., 2022; Ning et al., 2020b). Piecewise-homogeneous transition kernels and transition-dependent smooth control have also been used in the analysis and control of semi-Markov systems (Yang et al., 2024). We use residence-time partitioning to describe differences in mode transitions across residence stages and incorporate the resulting transition information together with mode-dependent triggering conditions into the closed-loop stability analysis.

SMC and Learning-Based Uncertainty Compensation. SMC provides robustness against matched uncertainty, but selecting the switching gain according to a bound on the full uncertainty may produce large control effort and input variation (Young et al., 1999; Edwards, 2004); higher-order SMC methods improve its performance in terms of control continuity and chattering mitigation (Moreno & Osorio, 2012; Utkin, 2016). Existing studies have combined SMC with asynchronous learning, neural network approximation, communication protocols, or adaptive dynamic programming (Li et al., 2022; Cheng et al., 2024; Xu et al., 2025; Yang et al., 2020). Meanwhile, neural Lyapunov methods and structured neural controllers have been used to establish stable or safe closed-loop control under unknown dynamics (Cui et al., 2023; Zhang et al., 2024). These methods mainly concern finite-dimensional systems and focus on learning control policies or unknown dynamics. We focus on the role of learned compensation in the SMC robust term: the neural operator reduces the uncertainty that this term must cover, rather than replacing the closed-loop stability analysis.