

Neural Operators That Stabilize Infinite-Dimensional Systems

Miroslav Krstic

★ 2026 IFAC WC workshop: *Trustworthy Data-Driven Control* ★

Neural Operators That Stabilize Infinite-Dimensional Systems

Miroslav Krstic

How to control PDEs
that are **NONLINEAR**
+
have totally **UNKNOWN**
parameters

★ 2026 IFAC WC workshop: *Trustworthy Data-Driven*

Neural Operators That Stabilize Infinite-Dimensional Systems

Miroslav Krstic

★ 2026 IFAC WC workshop: *Trustworthy Data-Driven*

How to control PDEs
that are **NONLINEAR**
+
have totally **UNKNOWN**
parameters

with

100% "TRUST"
(proofs)?

Feedback linearization: from finite to infinite dimension

ODEs (Isidori-Krener-GG-M, 1981; Hunt-Su-Meyer, 1983):

done, 45 years ago

Feedback linearization: from finite to infinite dimension

ODEs (Isidori-Krener-GG-M, 1981; Hunt-Su-Meyer, 1983): done, 45 years ago

Parabolic PDEs (Vazquez-Krstic 2008): $u_t = u_{xx} + F[u]$ done, 18 years ago

Feedback linearization: from finite to infinite dimension

ODEs (Isidori-Krener-GG-M, 1981; Hunt-Su-Meyer, 1983): *done, 45 years ago*

Parabolic PDEs (Vazquez-Krstic 2008): $u_t = u_{xx} + F[u]$ *done, 18 years ago*

Hyperbolic PDEs: $u_t = u_x + F[u]$: *open until now*

Hyperbolic plant class — a benchmark in ∞ dimension

$$\boxed{u_t = u_x + F[u]}, \quad u(1, t) = U(t)$$

Hyperbolic plant class — a benchmark in ∞ dimension

$$\boxed{u_t = u_x + F[u]}, \quad u(1, t) = U(t)$$

with F a spatial Volterra series

$$F[u](x, t) = \sum_{n \geq 2} \int_0^x \int_0^{\xi_1} \cdots \int_0^{\xi_{n-1}} f_n(x, \xi_1, \dots, \xi_n) u(\xi_1, t) \cdots u(\xi_n, t) d\xi_n \cdots d\xi_1$$

Hyperbolic plant class — a benchmark in ∞ dimension

$$\boxed{u_t = u_x + F[u]}, \quad u(1, t) = U(t)$$

with F a spatial Volterra series

$$F[u](x, t) = \sum_{n \geq 2} \int_0^x \int_0^{\xi_1} \cdots \int_0^{\xi_{n-1}} f_n(x, \xi_1, \dots, \xi_n) u(\xi_1, t) \cdots u(\xi_n, t) d\xi_n \cdots d\xi_1$$

“Taylor series for causal operators”

(Volterra 1887; Brockett 1976; Krener 1978)

4 Q&As to deployable FL controllers

(4 papers, Spring '26)

1. **exact** feedback linearization

but: controller = infinite series

4 Q&As to deployable FL controllers

(4 papers, Spring '26)

1. **exact** feedback linearization

but: controller = infinite series

2. **truncated**/approximate FL

but: kernel re-solve & real-time spatial integration

4 Q&As to deployable FL controllers

(4 papers, Spring '26)

1. **exact** feedback linearization but: controller = infinite series
2. **truncated**/approximate FL but: kernel re-solve & real-time spatial integration
3. controller **learned** offline as **neural operator** but: plant assumed known

4 Q&As to deployable FL controllers

(4 papers, Spring '26)

1. **exact** feedback linearization but: controller = infinite series
2. **truncated**/approximate FL but: kernel re-solve & real-time spatial integration
3. controller **learned** offline as **neural operator** but: plant assumed known
4. **adaptive** + neural operator *learns the plant online (my dream since 1997)*

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

1

The **Nonlinear Control** Half

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

Backstepping target

Construct **nonlinear transform** & *target system* (pure transport/delay)

$$w = u - K[u]$$

$$w_t = w_x$$

Brunovsky equivalent

Backstepping target

Construct **nonlinear transform** & *target system* (pure transport/delay)

$$w = u - K[u]$$

$$w_t = w_x$$

Brunovsky equivalent

and feedback

$$U(t) = K[u](1, t)$$

$$w(1, t) = 0$$

$\implies$ vanishes in $t = 1$

Backstepping target

Construct **nonlinear transform** & *target system* (pure transport/delay)

$$w = u - K[u]$$

$$w_t = w_x$$

Brunovsky equivalent

and feedback

$$U(t) = K[u](1, t)$$

$$w(1, t) = 0$$

$\implies$ vanishes in $t = 1$

$K =$ **Volterra series**, just like F :

$$K[u](x, t) = \sum_{n \geq 2} \int_0^x \int_0^{\xi_1} \cdots \int_0^{\xi_{n-1}} k_n(x, \xi_1, \dots, \xi_n) u(\xi_1, t) \cdots u(\xi_n, t) d\xi_n \cdots d\xi_1$$

Matching identity

what does K have to satisfy?

Kernels must satisfy ∞ many Goursat **hyperbolic PDEs**

$$\left(\partial_x + \sum_{i=1}^n \partial_{\xi_i} \right) k_n = -f_n + \sum_{m=2}^n B_n^m[k_{n-m+1}, f_m]$$

$$k_n|_{\xi_n=0} = 0$$

Matching identity

what does K have to satisfy?

Kernels must satisfy ∞ many Goursat **hyperbolic PDEs**

$$\left(\partial_x + \sum_{i=1}^n \partial_{\xi_i} \right) k_n = -f_n + \sum_{m=2}^n B_n^m[k_{n-m+1}, f_m]$$

$$k_n|_{\xi_n=0} = 0$$

Integral form:

$$k_n(x, \xi_1, \dots, \xi_n) = - \int_0^{\xi_n} \left[f_n - \sum_{m=2}^n B_n^m[k_{n-m+1}, f_m] \right] (x - \xi_n + s, \xi_1 - \xi_n + s, \dots, \xi_{n-1} - \xi_n + s, s) ds$$

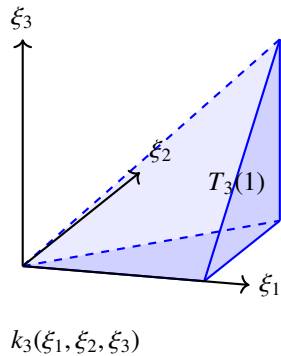
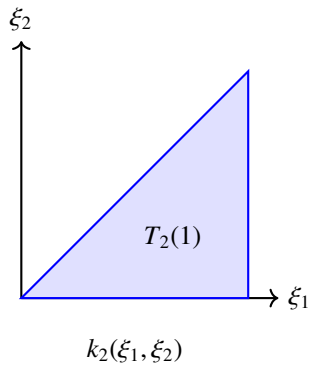
What does this make the controller $U(t) = K[u](1, t)$?

infinite sum of

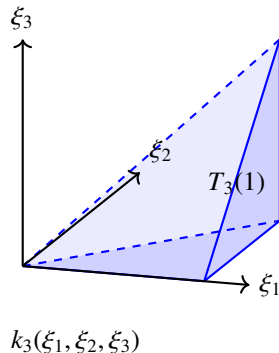
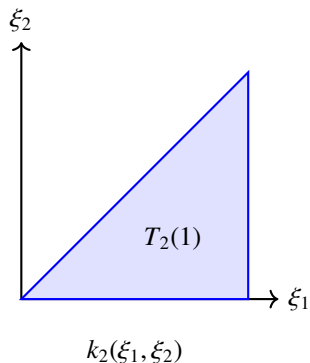
integrals nested **infinitely**

with gains $\{k_n\}$ on simplices of dimension that grows **infinite**

Kernels on simplices



Kernels on simplices



n -simplex $T_n(1) = \{0 \leq \xi_n \leq \cdots \leq \xi_1 \leq 1\}$ of volume $\frac{1}{n!}$

Dense technical core of the talk

(2 slides)

$K[u]$ converges

Plant (analyticity) assumption:

$$\|f_n\|_\infty \leq D_f \frac{n!}{\rho_f^{n-1}}$$

$K[u]$ converges

Plant (analyticity) assumption:

$$\|f_n\|_\infty \leq D_f \frac{n!}{\rho_f^{n-1}}$$

Lemma (Kernel bound)

$$\|k_n\|_{L^2(T_n(1))} \leq D_f e^{\Upsilon_K} \frac{\sqrt{n!}}{\rho_K^{n-1}}$$

$$\Upsilon_K := 1 + \frac{4}{3} D_f, \quad \rho_K := \frac{\rho_f}{2 e^{\Upsilon_K}}$$

$K[u]$ converges

Plant (analyticity) assumption:

$$\|f_n\|_\infty \leq D_f \frac{n!}{\rho_f^{n-1}}$$

Lemma (Kernel bound)

$$\|k_n\|_{L^2(T_n(1))} \leq D_f e^{\Upsilon_K} \frac{\sqrt{n!}}{\rho_K^{n-1}}$$

$$\Upsilon_K := 1 + \frac{4}{3} D_f, \quad \rho_K := \frac{\rho_f}{2 e^{\Upsilon_K}}$$

$K[u]$ converges

Plant (analyticity) assumption: $\|f_n\|_\infty \leq D_f \frac{n!}{\rho_f^{n-1}}$

Lemma (Kernel bound)

$$\|k_n\|_{L^2(T_n(1))} \leq D_f e^{\Upsilon_K} \frac{\sqrt{n!}}{\rho_K^{n-1}} \quad \Upsilon_K := 1 + \frac{4}{3} D_f, \quad \rho_K := \frac{\rho_f}{2 e^{\Upsilon_K}}$$

Lemma (Explicit bound on backstepping operator K)

Class $\mathcal{K}_\infty$ in $\|u\|_{L^2}$ on $[0, \rho_K)$:

$$\|K[u]\|_{L^2} \leq \kappa_2 \frac{\|u\|_{L^2}^2}{(1 - \|u\|_{L^2}/\rho_K)^2} \quad \kappa_2 := \frac{8 D_f e^{2\Upsilon_K}}{\rho_f}$$

Exponential stabilization by exact FL

Theorem (Closed-loop exponential stability in L^2)

$\forall$ decay rates $\lambda > 0$, $\exists$ explicit $C_1, C_2 > 0$ s.t., if

$$\|u_0\|_{L^2} < C_1$$

then

$$\boxed{\|u(\cdot, t)\|_{L^2} \leq C_2 e^{-\lambda t} \|u_0\|_{L^2}} \quad t \geq 0$$

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

Infinite series — show stopper for implementation

Truncate at order N :

$$U(t) = K_N[u](1, t) = \sum_{n=2}^N \int_{T_n(1)} k_n \prod_i u(\xi_i, t) d\xi$$

What can “break” under truncation?

Under **truncated** feedback: the target gets **boundary residual**:

$$w_t = w_x,$$

$$w(1, t) = K[u](1, t) - K_N[u](1, t) \neq 0$$

Residual decays in N

$$\left| K[u](1) - K_N[u](1) \right| \leq \varepsilon_N(\|u\|_\infty) = D_f e^{\gamma_K} \rho_K \cdot \frac{(\|u\|_\infty / \rho_K)^{N+1}}{1 - \|u\|_\infty / \rho_K}$$

For $\|u\|_\infty < \rho_K$, geometric decay of residual in N

Truncated FL asymp. stabilization

Theorem (Stability under K_N)

$\forall$ truncation orders $N \geq 2$, on a ball in L^∞ , $\exists \beta \in \mathcal{KL}$, s.t., under $U(t) = K_N[u](1, t)$,

$$\|u(\cdot, t)\|_\infty \leq \beta(\|u_0\|_\infty, t)$$

Well-posedness in closed loop

Mild solution form along characteristics:

$$u(x, t) = \int_{\max\{0, t-(1-x)\}}^t F[u(\cdot, s)](x+t-s) ds + \begin{cases} u_0(x+t), & t \in [0, 1-x] \\ K_N[u(\cdot, t-(1-x))](1), & t > 1-x \end{cases}$$

Well-posedness in closed loop

Mild solution form along characteristics:

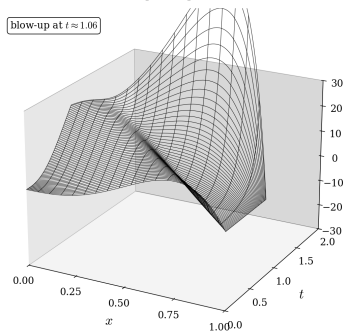
$$u(x, t) = \int_{\max\{0, t-(1-x)\}}^t F[u(\cdot, s)](x+t-s) ds + \begin{cases} u_0(x+t), & t \in [0, 1-x] \\ K_N[u(\cdot, t-(1-x))](1), & t > 1-x \end{cases}$$

Banach contraction on $C([0, \tau]; L^\infty)$

Simulation

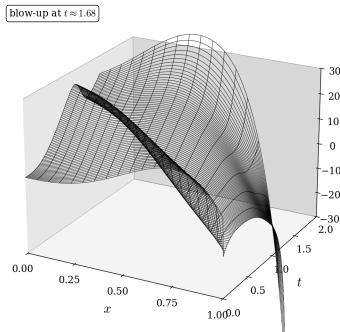
$$u_t = u_x + \frac{1}{2} \left(\int_0^x u \right)^2$$

(a) Open loop, $U \equiv 0$



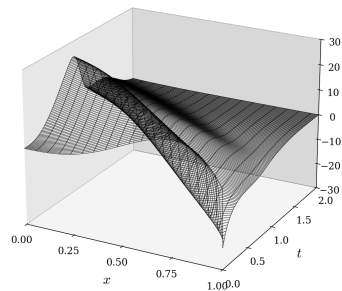
open loop

(b) Second-order feedback, $U = K_2[u](1, t)$



$N = 2$

(c) Third-order feedback, $U = K_2[u](1, t) + K_3[u](1, t)$



$N = 3$

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

2

The **Machine Learning** Half

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

Truncation is **finite**, but *still* numerically **expensive**

For each plant F ,

solve $N - 1$ kernel PDEs
on simplices up to dimension N
and evaluate an N -fold nested integral

Truncation is finite, but *still* numerically expensive

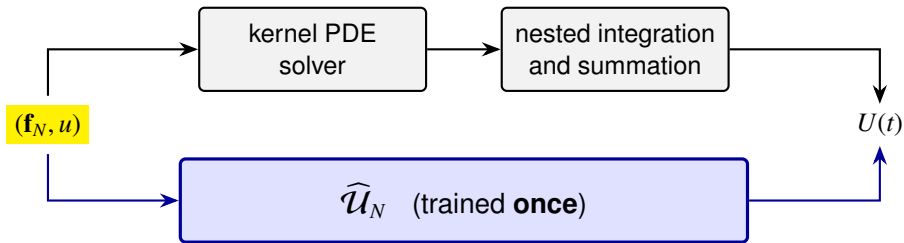
For each plant F ,

solve $N - 1$ kernel PDEs
on simplices up to dimension N
and evaluate an N -fold nested integral

INSTEAD: *learn* controller map

$$\mathcal{U}_N : (\mathbf{f}_N, u) \mapsto K_N[u](1)$$

Learn $\mathcal{U}_N$ as a neural operator



Theorem (Joint Lipschitzness of $\mathcal{U}_N$)

On a class of analytic-bounded plants and an L^∞ ball of states,

$$\left| \mathcal{U}_N(\mathbf{f}_N, u) - \mathcal{U}_N(\tilde{\mathbf{f}}_N, \tilde{u}) \right| \leq \mathcal{L}_f \|\mathbf{f}_N - \tilde{\mathbf{f}}_N\|_\star + \mathcal{L}_u \|u - \tilde{u}\|_{L^\infty}$$



Theorem (Joint Lipschitzness of $\mathcal{U}_N$)

On a class of analytic-bounded plants and an L^∞ ball of states,

$$|\mathcal{U}_N(\mathbf{f}_N, u) - \mathcal{U}_N(\tilde{\mathbf{f}}_N, \tilde{u})| \leq \mathcal{L}_f \|\mathbf{f}_N - \tilde{\mathbf{f}}_N\|_\star + \mathcal{L}_u \|u - \tilde{u}\|_{L^\infty}$$

where

$$\mathcal{L}_f = \frac{1}{b_f^2} \sum_{n=2}^N \frac{(n-1)!}{n} (b_f B_u)^n \quad \mathcal{L}_u = \frac{B_f}{b_f} \sum_{n=1}^{N-1} n! (b_f B_u)^n$$

and $b_f := \max\{1, 6B_f\}$, $\|f\|_{L^\infty} \leq B_f$, $\|u\|_{L^\infty(0,1)} \leq B_u$.

Theorem (Joint Lipschitzness of $\mathcal{U}_N$)

On a class of analytic-bounded plants and an L^∞ ball of states,

$$|\mathcal{U}_N(\mathbf{f}_N, u) - \mathcal{U}_N(\tilde{\mathbf{f}}_N, \tilde{u})| \leq \mathcal{L}_f \|\mathbf{f}_N - \tilde{\mathbf{f}}_N\|_\star + \mathcal{L}_u \|u - \tilde{u}\|_{L^\infty}$$

where

$$\mathcal{L}_f = \frac{1}{b_f^2} \sum_{n=2}^N \frac{(n-1)!}{n} (b_f B_u)^n \quad \mathcal{L}_u = \frac{B_f}{b_f} \sum_{n=1}^{N-1} n! (b_f B_u)^n$$

and $b_f := \max\{1, 6B_f\}$, $\|f\|_{L^\infty} \leq B_f$, $\|u\|_{L^\infty(0,1)} \leq B_u$.

$$\Rightarrow \boxed{\exists \text{ neural operator } \hat{\mathcal{U}}_N \text{ with } |\mathcal{U}_N - \hat{\mathcal{U}}_N| \leq \varepsilon}$$

Closed loop under $\widehat{\mathcal{U}}_N$

Boundary residual splits:

$$\underbrace{[\widehat{\mathcal{U}}_N - \mathcal{U}_N]}_{\leq \varepsilon \text{ (training)}} + \underbrace{[\mathcal{U}_N - \mathcal{U}]}_{\leq \varepsilon_N \text{ (truncation)}}$$

Closed loop under $\widehat{\mathcal{U}}_N$

Boundary residual splits:

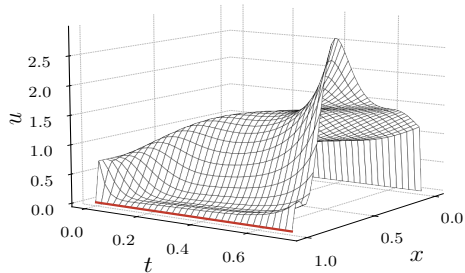
$$\underbrace{[\widehat{\mathcal{U}}_N - \mathcal{U}_N]}_{\leq \varepsilon \text{ (training)}} + \underbrace{[\mathcal{U}_N - \mathcal{U}]}_{\leq \varepsilon_N \text{ (truncation)}}$$

Theorem (**Practical** stability under NO feedback)

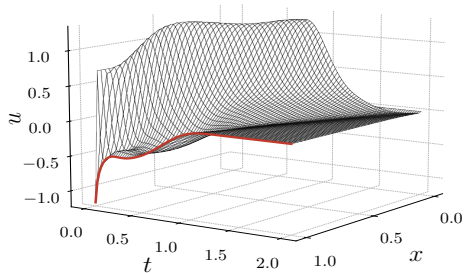
On explicit ball, $\exists \beta \in \mathcal{KL}$ s.t.

$$\|u(\cdot, t)\|_\infty \leq \beta(\|u_0\|_\infty, t) \boxed{+ \gamma(F, N) \varepsilon}$$

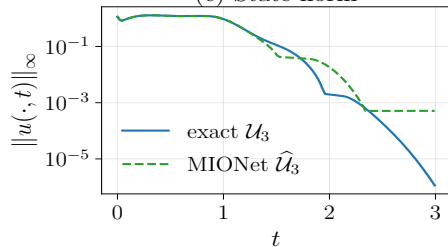
(a) Open loop $U \equiv 0$



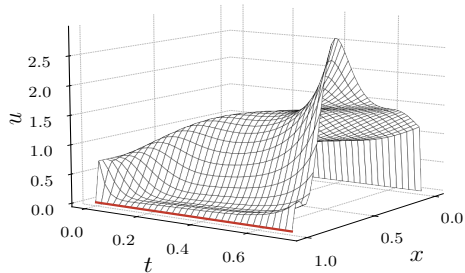
(b) Closed loop $U = \hat{\mathcal{U}}_3$



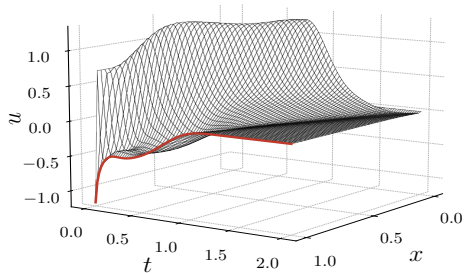
(c) State norm



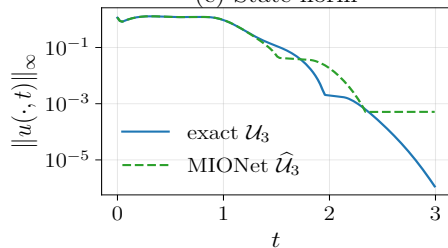
(a) Open loop $U \equiv 0$



(b) Closed loop $U = \hat{u}_3$



(c) State norm



1-hour offline training



multi-thousandfold
online speedup w/ NO $\hat{\mathcal{U}}_3$

Exact FL

Truncated controller

Neural operator for full feedback law

Adaptive + NO controller

What if the plant $\mathbf{f}_N = (f_1, \dots, f_N)$ is **unknown**?

adaptive control = learn *plant* **ONLINE**

Observer + parameter update

Observer (state $\hat{u}$, fed by plant state u):

$$\hat{u}_t = \hat{u}_x + F_N[u, \hat{\mathbf{f}}] + c_0 (u - \hat{u}) \Phi_N(u), \quad \hat{u}(1, t) = U(t)$$

nonlinear damping $\Phi_N(u) = \sum_{n=1}^N \frac{\gamma_n}{n!} \|u\|_{L^2}^{2n}$

Observer + parameter update

Observer (state $\hat{u}$, fed by plant state u):

$$\hat{u}_t = \hat{u}_x + F_N[u, \hat{\mathbf{f}}] + c_0 (u - \hat{u}) \Phi_N(u), \quad \hat{u}(1, t) = U(t)$$

$$\text{nonlinear damping } \Phi_N(u) = \sum_{n=1}^N \frac{\gamma_n}{n!} \|u\|_{L^2}^{2n}$$

Parameter update (gradient-projection onto $\bar{\mathcal{F}}_n = \{g : \|g\|_\infty \leq n! D_f / \rho_f^{n-1}\}$):

$$\partial_t \hat{f}_n = \text{Proj}_{\bar{\mathcal{F}}_n} \left\{ \gamma_n e^{cx} (u(x, t) - \hat{u}(x, t)) \prod_{i=1}^n u(\xi_i, t), \hat{f}_n \right\}$$

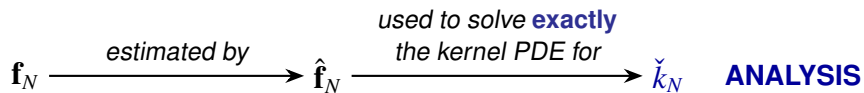
What if the plant $\mathbf{f}_N = (f_1, \dots, f_N)$ is **unknown**?

adaptive control = learn *plant* **ONLINE**

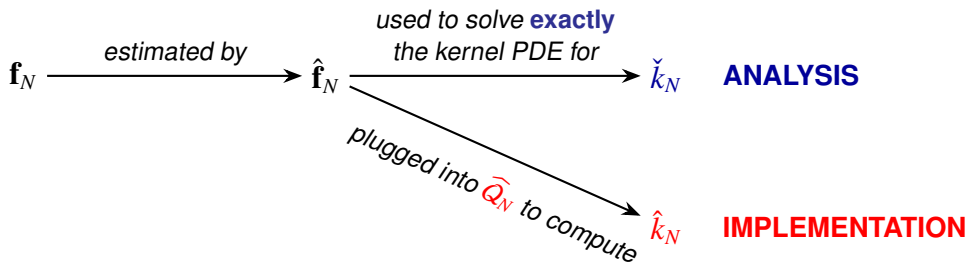
NO control = learn *mapping to controller gains* **OFFLINE**

Combine the two

Two kernels from one parameter estimate

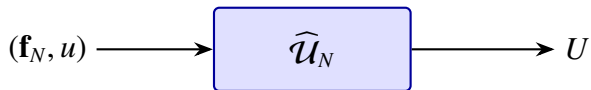


Two kernels from one parameter estimate



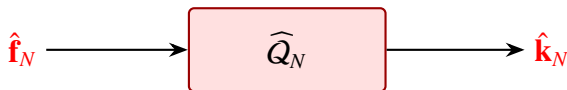
Two neural operators in this presentation

— full feedback and adaptive



Two neural operators in this presentation

— full feedback and adaptive



Theorem (Lipschitzness of the kernel-generating map Q_N)

On the projection box of admissible plant estimates,

$$\|Q_N(\hat{\mathbf{f}})_n - Q_N(\hat{\mathbf{f}}')_n\|_{L^2} \leq \ell_n \|\hat{\mathbf{f}} - \hat{\mathbf{f}}'\|_{\star}$$

where

$$\ell_n = (n!)^3 \left(12 \max\{1, D_f, 2/\rho_f\}^2\right)^n e^{2n + \frac{8}{3}nD_f^2 + 2nN(N+1)D_f}$$

Analysis under FIVE *vanishing* perturbations

Target:

$$w_t = w_x + \text{four interior } \mathbf{vanishing} \text{ perturbations}$$

$$w(1, t) = \Gamma(t) \quad \text{boundary } \mathbf{vanishing} \text{ perturbation}$$

Analysis under FIVE *vanishing* perturbations

Target:

$$\begin{aligned}w_t &= w_x + \text{four interior } \mathbf{vanishing} \text{ perturbations} \\w(1, t) &= \Gamma(t) \quad \text{boundary } \mathbf{vanishing} \text{ perturbation}\end{aligned}$$

Interior:

- R_N^u
- R_N^K
- $D = c_0 \Phi_N(u) (I - \check{K}_N^*) (u - \hat{u})$
- $\Omega = \sum_n \int_{T_n(x)} \partial_t \check{k}_n \prod_i \hat{u} d\xi$

plant Volterra tail
bilinear $\check{k}_n \otimes \hat{f}_m$ cross-orders $n+m-1 > N$
observer damping
*adaptation **rate***

Analysis under FIVE *vanishing* perturbations

Target:

$$w_t = w_x + \text{four interior } \mathbf{vanishing} \text{ perturbations}$$
$$\mathbf{w}(1, t) = \mathbf{\Gamma}(t) \quad \text{boundary } \mathbf{vanishing} \text{ perturbation}$$

Interior:

- R_N^u
- R_N^K
- $D = c_0 \Phi_N(u) (I - \check{K}_N^*) (u - \hat{u})$
- $\Omega = \sum_n \int_{T_n(x)} \partial_t \check{k}_n \prod_i \hat{u} d\xi$

plant Volterra tail

bilinear $\check{k}_n \otimes \hat{f}_m$ cross-orders $n+m-1 > N$

observer damping

*adaptation **rate***

Boundary: $\mathbf{\Gamma}(t) = - \sum_n \int_{T_n(1)} (\check{k}_n - \hat{k}_n) \prod_i \hat{u} d\xi$ with $\|\check{k}_n - \hat{k}_n\|_{L^2} \leq \mathbf{\varepsilon}$

$$V_{\text{tot}} = \|w\|_c^2 + \lambda \|u - \hat{u}\|_c^2 + \lambda \sum_{n=1}^N \frac{1}{\gamma_n} \|f_n - \hat{f}_n\|_{L^2(T_n(1))}^2$$

$$V_{\text{tot}} = \|w\|_c^2 + \lambda \|u - \hat{u}\|_c^2 + \lambda \sum_{n=1}^N \frac{1}{\gamma_n} \|f_n - \hat{f}_n\|_{L^2(T_n(1))}^2$$

No parameter drift, **no σ -modification**, no PE

CROWN RESULT

Theorem

$\exists$ NO $\widehat{Q}_N$ and $\bar{\Pi}, \mu_0 > 0$ s.t. $\forall$ initial condition w/ $\Pi(0) \leq \bar{\Pi}$,

- *Lyapunov stability:*

$$\Pi(t) \leq \mu_0 \Pi(0), \quad t \geq 0$$

where

$$\Pi(t) = \sqrt{\|u(\cdot, t)\|_{L^2}^2 + \|\hat{u}(\cdot, t)\|_{L^2}^2 + \sum_{n=1}^N \|f_n - \hat{f}_n(\cdot, t)\|_{L^2(T_n(1))}^2}$$

CROWN RESULT

Theorem

$\exists$ NO $\widehat{Q}_N$ and $\bar{\Pi}, \mu_0 > 0$ s.t. $\forall$ initial condition w/ $\Pi(0) \leq \bar{\Pi}$,

- *Lyapunov stability:*

$$\Pi(t) \leq \mu_0 \Pi(0), \quad t \geq 0$$

where

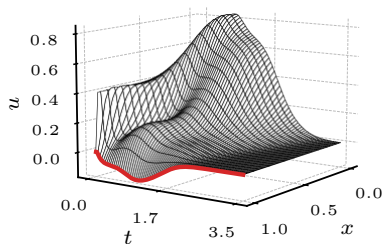
$$\Pi(t) = \sqrt{\|u(\cdot, t)\|_{L^2}^2 + \|\hat{u}(\cdot, t)\|_{L^2}^2 + \sum_{n=1}^N \|f_n - \hat{f}_n(\cdot, t)\|_{L^2(T_n(1))}^2}$$

- *Asymptotic regulation:*

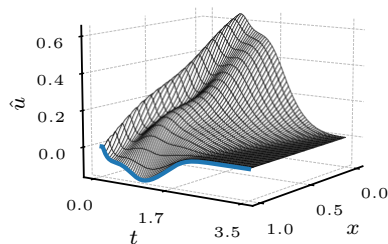
$$\|u(\cdot, t)\|_{L^2}, \|\hat{u}(\cdot, t)\|_{L^2}, |U(t)| \rightarrow 0$$

Numerical Illustration: Adaptive ($N = 2$)

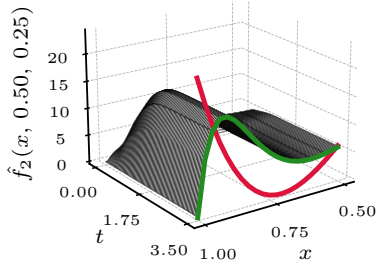
Adaptive Closed-loop u
with NO kernel



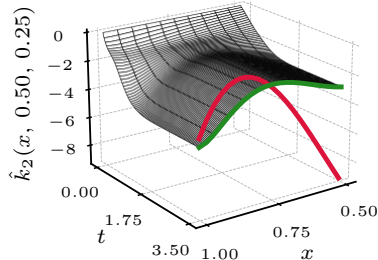
Observer (initial $\hat{u}(x, 0) = 0$)



Slice of estimated $\hat{f}_2$



Slice of estimated $\hat{k}_2$ from $\hat{\mathcal{Q}}_2(\hat{f}_2)$



Summary

- [online](#) identification of *finitely many* **infinite-dimensional** parameters

Summary

- online identification of *finitely many* **infinite-dimensional** parameters
- offline learning of parameter-to-gain map

Summary

- **online** identification of *finitely many* **infinite-dimensional** parameters
- **offline** learning of parameter-to-gain map

from infinite to **finite**

from finite to offline **LEARNED**

from offline learned to **ONLINE LEARNED**

Questions?