Parameter Identification in Structured Discrete-Time Uncertainties without Persistency of Excitation

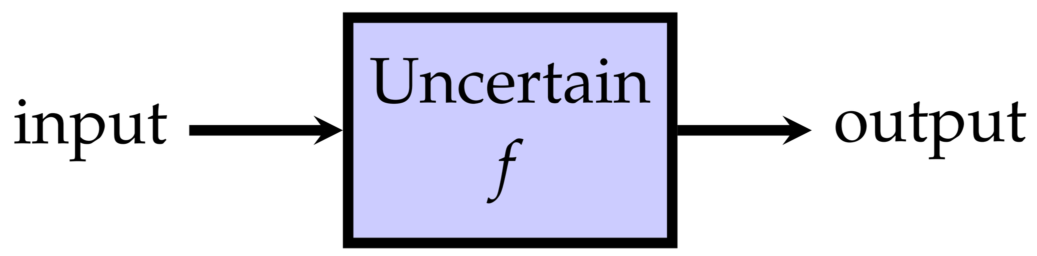
Ouboti Djaneye-Boundjou, Advisor: Raúl Ordóñez

Department of Electrical and Computer Engineering, University of Dayton

UNIVERSITY of DAYTON — NONLINEAR CONTROL LAB

Background

- System Identification:
  - Function approximation



- Sys-ID usage: **machine learning**, **adaptive control**, ...
- Present study:
  - **Discrete-time (DT) structured uncertainties**

$$f(x(k)) = \theta^\top \phi(x(k)) \tag{1}$$

  - $k \in \{k_0, k_0+1, k_0+2, \ldots, k_f\} \subset \mathbb{N}$: discrete sampling time
  - $\theta \in \mathbb{R}^{r_\theta}$, $r_\theta \in \mathbb{N}^+$: **unknown, ideal, constant** parameter vector
  - $\phi(x(k)) \in \mathbb{R}^{r_\theta}$: **known** basis or regressor vector
  - *Example*:

$$f_1(x(k)) = -\frac{1}{4} + 10\underbrace{\exp\left(-\frac{\left(x(k)-\frac{\pi}{2}\right)^2}{4}\right)}_{b(x(k))} = \underbrace{\begin{bmatrix} -\frac{1}{4} & 10 \end{bmatrix}}_{\theta^\top} \underbrace{\begin{bmatrix} 1 \\ b(x(k)) \end{bmatrix}}_{\phi(x(k))} \tag{2}$$

  - Approximator

$$\mathcal{F}\left(\phi(x(k)), \hat\theta(k)\right) = \hat\theta^\top(k)\,\phi(x(k)) \tag{3}$$

  - $\hat\theta(k) \in \mathbb{R}^{r_\theta}$: estimate of θ updated via **adaptation law**
  - Parameter error $\tilde\theta(k) = \hat\theta(k) - \theta$ is not computable
  - Compute estimation error

$$q(k) = \mathcal{F}\left(\phi(x(k)), \hat\theta(k)\right) - f(x(k))$$

  - **Note**: $\underbrace{q(k) = \tilde\theta^\top(k)\,\phi(x(k))}_{\text{1 equation with } r_\theta \text{ unknowns}}$ from (1) and (3)

Parameter Identification (PI) Problem

Drive $\tilde\theta(k) \to [0]^{r_\theta}$ or $\hat\theta(k) \to \theta$, causing $q(k) \to 0$, as $k \to \infty$

Motivation

- **PI, i.e, $\tilde\theta(k) \to 0$, leads to improved estimation performance**
- Literature: traditional approximation methods guarantee PI provided **persistency of excitation** (very restrictive)
- Present study:
  - **Develop an adaptive estimator with PI guarantees**
  - **Relax persistency of excitation requirement**

Normalized Gradient (NG) Descent

- NG: traditional approach to approximation
- **NG adaptation law**: given an initial $\hat\theta(k_0)$,

$$\hat\theta(k+1) = \hat\theta(k) - \eta\underbrace{\frac{\phi(x(k))\,q(k)}{m^2(k)}}_{\text{instantaneous update}} \tag{4}$$

  - $\eta > 0$: step size or learning rate or gain
  - $m(k)$: normalization signal ensuring $\psi(x(k)) = \frac{\phi(x(k))}{m(k)}$
- **Lyapunov stability analysis**: we can show that $\tilde\theta(k)$ **remains bounded for all k** if $0 < \eta < \bar\eta_{NG}$
- **PI, i.e., $\tilde\theta(k) \to 0$, only if $\psi(x(k))$ is persistently exciting**

Concurrent Learning (CL) Preliminaries

- CL: first introduced in **continuous-time** framework
- **Use of memory**:
  - Record past data: for $k_0 \leqslant \tau_j < k$, with $j = 1, 2, \ldots, c_Z$
    - $Z \in \mathbb{R}^{r_\theta \times c_Z}$: **History stack** of $\psi(x(\tau_j))$ vectors
    - $M \in \mathbb{R}^{c_Z}$: vector of $m(\tau_j)$ values
    - $Y \in \mathbb{R}^{c_Z}$: vector of $f(x(\tau_j))$ values
- **CL condition**: Z contains r_θ linearly independent $\psi(x(\tau_j))$
  - **Less restrictive than persistency of excitation**

Gradient-Based CL in DT

- **Gradient-Based CL adaptation law**: given an initial $\hat\theta(k_0)$,

$$\hat\theta(k+1) = \underbrace{\hat\theta(k) - \eta\frac{\phi(x(k))\,q(k)}{m^2(k)}}_{\text{NG portion}} - \eta\underbrace{\overbrace{\sum_{j=1}^{c_Z}\frac{\phi(x(\tau_j))\,q_j(k)}{m^2(\tau_j)}}^{ZZ^\top\tilde\theta(k)}}_{\text{Update based on recorded data}} \tag{5}$$

  - Estimation error based on recorded data:

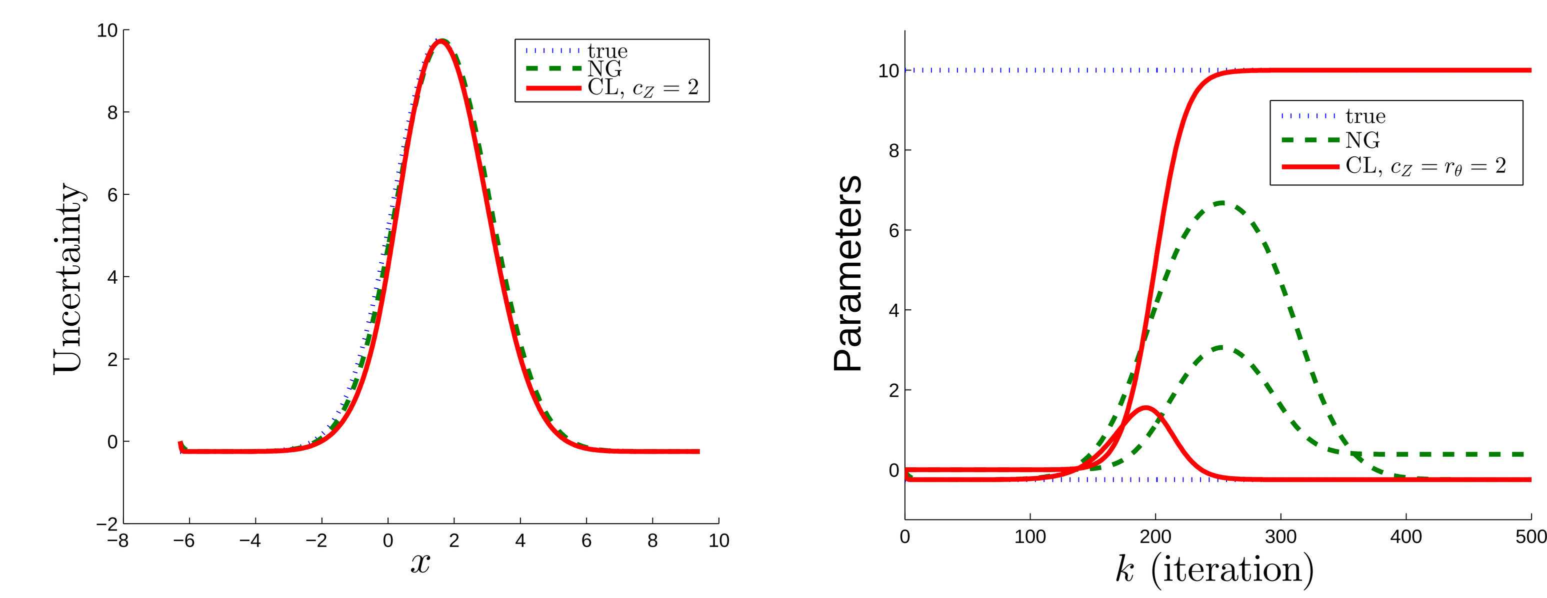
$$q_j(k) = \underbrace{\mathcal{F}\left(\phi(x(\tau_j)), \hat\theta(k)\right) - f(x(\tau_j))}_{\text{computable in simulation}} = \tilde\theta^\top(k)\,\phi(x(\tau_j))$$

- **Lyapunov stability analysis**: granted **CL condition is met, $\Omega = ZZ^\top$ is positive definite** and we prove that **$\tilde\theta(k) \to 0$ exponentially (PI)** if $0 < \eta < \bar\eta_{CL}$

Numerical Simulations

- Here, $f = f_1$ is approximated
- x is varied from $x_L = -2\pi$ at k_0 to $x_H = 3\pi$ at k_f uniformly
- How good is $\mathcal{F}$ if $\hat\theta(k)$ is frozen at each k to reconstruct f? Consider metric

$$e(k) = \int_{x_L}^{x_H} \left|\mathcal{F}\left(\phi(x), \hat\theta(k)\right) - f(x)\right| dx$$



(a) **Uncertainty f_1** (b) **Parameter estimation**

Figure 1: **On-line appproximation**

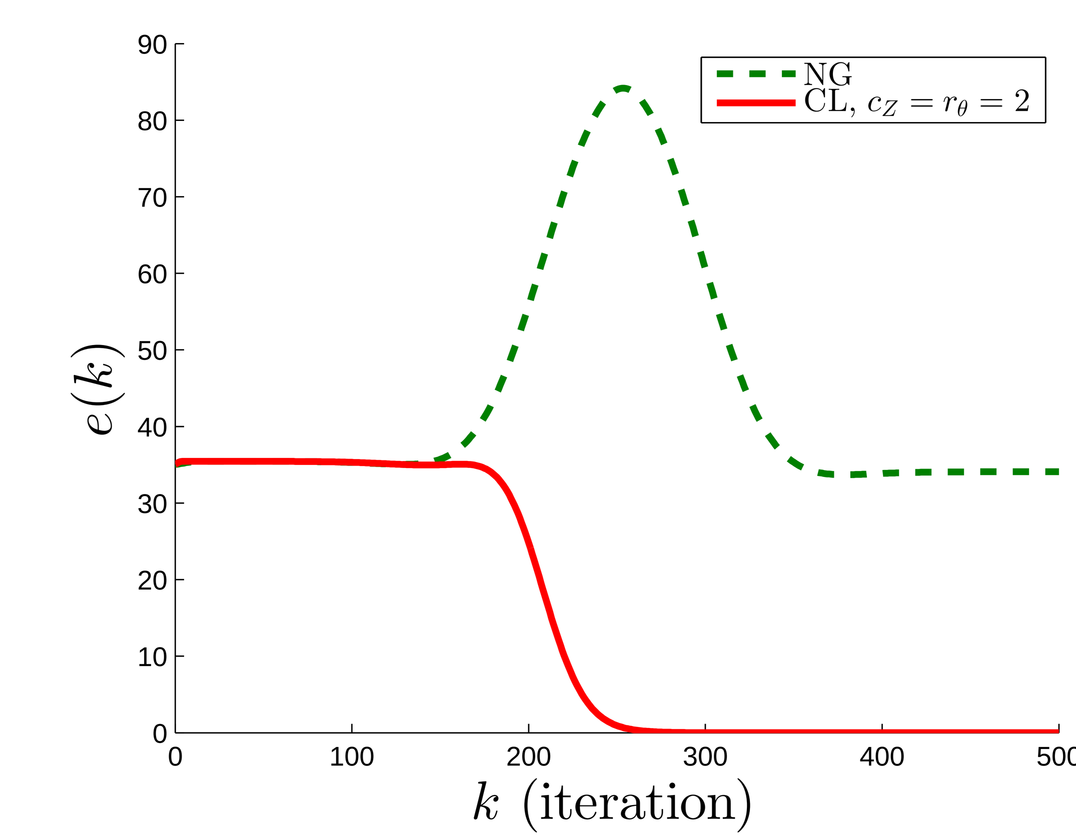


Figure 2: **Metric $e(k)$**

- **CL achieves better identification of f_1**

N.B.: See pushibed paper for expressions of $\bar\eta_{NG}$ and $\bar\eta_{CL}$

Future Work

- How will CL fare with unstructured uncertainties?
- Apply CL adaptation law within a control loop