# Non-myopic Innovations Dual Controller

Miroslav Flídr and Miroslav Šimandl

Department of Cybernetics
&
Research Centre Data, Algorithms and Decision Making
Faculty of Applied Sciences
University of West Bohemia

# Outline

1. Introduction to dual adaptive control
2. Goal of the paper
3. Non-myopic Innovations Dual Controller
   - Formulation of the optimization problem
   - Solution of the optimization problem
   - Numerical example
4. Conclusion

# Dual adaptive control

## What is dual control?

- Appears in control problems with unknown state and parameters
  - the certainty equivalence property does not hold
  - the problem is not separable and not neutral
- Two conflicting goals – meet control objective and improve estimation
- Aspects of dual control
  - **Caution** - due to inherent uncertainties
  - **Probing (Active learning)** - helps decrease the uncertainty about the unknown state and parameters

## Optimal dual control problem

Cannot be mostly solved in closed form $\Rightarrow$ suboptimal solutions

# Suboptimal solutions to optimal dual control problem

## Explicit dual controllers

Based on augmentation of cautious control mostly with constraining the control horizon to one-step.

- with direct augmentation of cautious control
- modification of the cost function

$$\mathcal{L}_k = \mathcal{L}_k^c + \lambda \mathcal{L}_k^p, \qquad \lambda \geq 0$$

## Implicit dual controllers

Based om approximate solution of the Bellman optimization recursion.

- with approximation of the Bellman function
- with approximation of probability density functions

Filatov N, Unbehauen H. "Survey of adaptive dual control methods".
IEE Proceedings - Control Theory and Applications. 2000;147(1):118-128.

# Innovations Dual Controller

## Problem formulation

Find $u_k^{N-1}$ that minimizes the criterion

$$J = E\left\{\sum_{k=0}^{N-1} (y_{k+1} - \bar{y}_{k+1})^2\right\}$$

subject to the system

$$y_{k+1} = b_0 u_k + \dots + b_m u_{k-m} + a_1 y_k + \dots + a_n y_{k-n} + e_{k+1}, \quad k = 0, \dots, N-1$$

## Innovations dual controller (IDC) - Millito *et al.* (1982)

IDC is solution of modified optimization problem

1. control horizon shortened to one step
2. the cost function is modified

$$J_k = E\left\{(y_{k+1} - \bar{y}_{k+1})^2 - \lambda_{k+1}\nu_{k+1}^2 | y_0^k, u_0^{k-1}\right\}$$

- the parameter $0 \le \lambda_{k+1} \le 1$ specifies the degree of compromise between control and estimation objectives
- the innovations sequence $\nu_{k+1}$ provides overall information about estimation quality

# Goal: to design non-myopic innovations dual controller

## Deficiencies of IDC

- limited to one step ahead horizon ⇒ can suffer from myopic behavior
- designed only for SISO ARMAX systems

## Requirements of feasible solution

- ✔ computationally moderate not only for one step ahead horizon
- ✔ clear interpretation
- ✔ guarantees sufficient control quality
- ✔ moderate computational demands

## Steps to fulfil the goal

1. formulation of optimization problem with arbitrary control horizon
2. choice of probability density function approximation that would make possible to find closed form solution.
3. assurance of both properties of the dual control

# Goal: to design non-myopic innovations dual controller

## Deficiencies of IDC

- limited to one step ahead horizon ⇒ can suffer from myopic behavior
- designed only for SISO ARMAX systems

## Requirements of feasible solution

- ✔ computationally moderate not only for one step ahead horizon
- ✔ clear interpretation
- ✔ guarantees sufficient control quality
- ✔ moderate computational demands

## Steps to fulfil the goal

1. formulation of optimization problem with arbitrary control horizon
2. choice of probability density function approximation that would make possible to find closed form solution.
3. assurance of both properties of the dual control

# Considered system

$$s_{k+1} = A(\theta_k)s_k + B(\theta_k)u_k + w_k, \quad (1)$$
$$y_k = Cs_k + v_k, \quad (2)$$

| | | |
|---|---|---|
| $s_k \in \mathbb{R}^n$ | ... | non-measurable state |
| $\theta_k \in \mathbb{R}^p$ | ... | unknown parameters |
| $u_k \in \mathbb{R}^r$ | ... | control |
| $y_k \in \mathbb{R}^m$ | ... | measurement |

- ✓ The elements of matrices $A(\theta_k)$ and $B(\theta_k)$ are known linear function of the unknown parameters $\theta_k$.
- ✓ The random quantities $s_0$, $\theta_0$, $w_k$, $\epsilon_k$ and $v_k$ are described by known pdf's and are mutually independent.

# Optimization problem

## General optimization problem

The aim is to find control law

$$\boldsymbol{u}_k = \boldsymbol{u}_k(\mathfrak{I}_k) = \boldsymbol{u}_k(\boldsymbol{u}_0^{k-1}, \boldsymbol{y}_0^k), \qquad k = 0, 1, \ldots, N-1$$

that minimizes the following criterion

$$J = E\left\{ \sum_{k=0}^{N-1} (\boldsymbol{s}_{k+1} - \bar{\boldsymbol{s}}_{k+1})^T \boldsymbol{Q}_{k+1} (\boldsymbol{s}_{k+1} - \bar{\boldsymbol{s}}_{k+1}) + \boldsymbol{u}_k^T \boldsymbol{R}_k \boldsymbol{u}_k \right\}$$

subject to the system (1)-(2).

## Solvability of the optimization problem

- general solution given by Bellman optimization recursion
- analytically unsolvable (due to inherent nonlinearities)
- it is necessary to use some approximation

# Simple approximate solutions of the optimization problem

## Possible approximation choices

- Enforced certainty equivalence → leads to HCE controller

$$\rho_k^{CE} = \Big\{ p(\boldsymbol{s}_{k+i}, \boldsymbol{\theta}_{k+i} | \boldsymbol{I}_{k+i}) \simeq \delta(\boldsymbol{s}_{k+i} - \hat{\boldsymbol{s}}_{k+i}) \delta(\boldsymbol{\theta}_{k+i} - \hat{\boldsymbol{\theta}}_{k+i|k}); \quad i = 0, \ldots, N-k-1 \Big\}$$

- Partial certainty equivalence (PCE)

$$\rho_k = \big\{ p(\boldsymbol{s}_{k+i}, \boldsymbol{\theta}_{k+i} | \boldsymbol{I}_{k+i}) \simeq \delta(\boldsymbol{s}_{k+i} - \hat{\boldsymbol{s}}_{k+i}) p(\boldsymbol{\theta}_{k+i} | \mathfrak{I}_k); \quad i = 0, \ldots, N-k-1 \big\} \qquad (3)$$

## Features of the optimization problem employing PCE approximation

Control law sought as to minimize the criterion

$$J = E_{\rho_0} \left\{ \sum_{k=0}^{N-1} \mathcal{L}_k(\boldsymbol{s}_k, \boldsymbol{\theta}_k, \boldsymbol{u}_k) \right\}$$

- the expectations determined using $\rho$ approximation (3)
- the control law is suboptimal with respect to original formulation
- not strictly using the *closed-loop* information processing strategy anymore
- The resulting controller is of cautious type, i.e. it **isn't** dual controller!

# Reformulation of the optimization problem

## Reformulated optimization problem employing PCE approximation

Control law sought as

$$\boldsymbol{u}_k = \underset{\boldsymbol{u}_k}{\operatorname{argmin}}\, J_k(\mathfrak{I}_k), \qquad k = 0, 1, \ldots, N-1$$

with receding horizon type of the cost-to-go

$$J_k\left(\mathfrak{I}_k\right) = E_{\rho_k}\left\{\sum_{i=k}^{k+m} \mathcal{L}_i\left(\boldsymbol{s}_i, \boldsymbol{\theta}_i, \boldsymbol{u}_i\right) \middle| \mathfrak{I}_k\right\}$$

## Modification of the cost function based on IDC

$$\mathcal{L}_i\left(\cdot\right) = \left(\boldsymbol{s}_{i+1} - \bar{\boldsymbol{s}}_{i+1}\right)^T \boldsymbol{Q}_{i+1}\left(\boldsymbol{s}_{i+1} - \bar{\boldsymbol{s}}_{i+1}\right) + \boldsymbol{u}_i^T \boldsymbol{R}_i \boldsymbol{u}_i - \boldsymbol{v}_{i+1}^T \boldsymbol{\Lambda}_{i+1} \boldsymbol{v}_{i+1}$$

where $\boldsymbol{v}_{i+1} = \boldsymbol{y}_{i+1} - \hat{\boldsymbol{y}}_{i+1|i}\left(\hat{\boldsymbol{s}}_i, \hat{\boldsymbol{\theta}}_i\right)$ with $\hat{\boldsymbol{y}}_{i+1|i} \triangleq E_{\rho_k}\left\{\boldsymbol{y}_{i+1} \middle| \mathfrak{I}_i\right\}$

- ✔ simple cost function modification with clear interpretation
- ✔ the quality of estimates rated using innovations sequence
- ✔ still analytically solvable using PCE

# Analysis of the cost function

## Rearranged cost function

$$\mathcal{L}_i(\boldsymbol{s}_i, \boldsymbol{\theta}_i, \boldsymbol{u}_i) = \left(\hat{\boldsymbol{s}}_{i+1|i} - \bar{\boldsymbol{s}}_{i+1}\right)^T \boldsymbol{Q}_{i+1} \left(\hat{\boldsymbol{s}}_{i+1|i} - \bar{\boldsymbol{s}}_{i+1}\right) + \boldsymbol{u}_i^T \boldsymbol{R}_i \boldsymbol{u}_i + E_{\rho_k}\left\{ \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right)^T \left(\boldsymbol{Q}_{i+1} - \boldsymbol{\Lambda}_{i+1}\right) \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right) \Big| \mathfrak{I}_i \right\}$$

## Decomposition of the cost function

$$\mathcal{L}_k = \mathcal{L}_k^{\mathcal{C}} + \mathcal{L}_k^{\mathcal{P}}$$

- Cautious part (it's equivalent to the original quadratic cost function)

$$\mathcal{L}_k^{\mathcal{C}} = \left(\hat{\boldsymbol{s}}_{i+1|i} - \bar{\boldsymbol{s}}_{i+1}\right)^T \boldsymbol{Q}_{i+1} \left(\hat{\boldsymbol{s}}_{i+1|i} - \bar{\boldsymbol{s}}_{i+1}\right) + \boldsymbol{u}_i^T \boldsymbol{R}_i \boldsymbol{u}_i + + E_{\rho_k}\left\{ \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right)^T \boldsymbol{Q}_{i+1} \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right) \Big| \mathfrak{I}_k \right\}$$

- Probing part

$$\mathcal{L}_k^{\mathcal{P}} = -E_{\rho_k}\left\{ \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right)^T \boldsymbol{\Lambda}_{i+1} \left(\boldsymbol{s}_{i+1} - \hat{\boldsymbol{s}}_{i+1|i}\right) \Big| \mathfrak{I}_k \right\}$$

$\Rightarrow$ it comprises both aspect of the dual control

# The solution of the modified optimization problem

## Bellman optimization recursion (for receding horizon type of problem)

The recursion at time instant $k$ is defined as

$$\mathcal{V}_j = \min_{\boldsymbol{u}_i} \left\{ E_{\rho_k} \left\{ \mathcal{L}_i(\boldsymbol{s}_i, \boldsymbol{\theta}_i, \boldsymbol{u}_i) + \mathcal{V}_{j+1} \middle| \mathfrak{I}_i \right\} \right\},$$

$$j = m, ..., 0;\ i = k + j$$

with boundary condition $\mathcal{V}_{m+1} = \boldsymbol{\mathcal{O}}$.

## The form of the Bellman function

$$\mathcal{V}_j = \hat{\boldsymbol{s}}_i^T \boldsymbol{\Pi}_{m-j+1} \hat{\boldsymbol{s}}_i + \hat{\boldsymbol{s}}_i^T \boldsymbol{F}_{m-j+1} + \boldsymbol{F}_{m-j+1}^T \hat{\boldsymbol{s}}_i + h_{m-j+1},$$

where $\boldsymbol{\Pi}_{m-j+1} \in \mathbb{R}^{n \times n}$, $\boldsymbol{F}_{m-j+1} \in \mathbb{R}^n$ and $h_{m-j+1} \in \mathbb{R}$ and $\boldsymbol{\Pi}_0$, $\boldsymbol{F}_0$ and $h_0$ are zero valued (follows from the boundary condition).

# The dual control law

## The dual control law

$$\boldsymbol{u}_k = -\left[ \boldsymbol{R}_k + \boldsymbol{B}^T(\hat{\boldsymbol{\theta}}_{k|k})\overline{\boldsymbol{Q}}_{k+1}\boldsymbol{B}(\hat{\boldsymbol{\theta}}_{k|k}) + \boldsymbol{P}^{BB}_{k|k}\right]^{-1} \times$$
$$\times \Big[\boldsymbol{B}^T(\hat{\boldsymbol{\theta}}_{k|k})\overline{\boldsymbol{Q}}_{k+1}\boldsymbol{A}(\hat{\boldsymbol{\theta}}_{k|k})\hat{\boldsymbol{s}}_k + \boldsymbol{P}^{BA}_{k|k}\hat{\boldsymbol{s}}_k$$
$$-\boldsymbol{B}^T(\hat{\boldsymbol{\theta}}_{k|k})\boldsymbol{C}^T_{k+1}\boldsymbol{Q}_{k+1}\bar{\boldsymbol{s}}_{k+1} + \boldsymbol{B}^T(\hat{\boldsymbol{\theta}}_{k|k})\boldsymbol{F}_{m-1}\Big].$$

where $\overline{\boldsymbol{Q}}_{k+1} = \left(\boldsymbol{Q}_{k+1} + \boldsymbol{\Pi}_{m-1}\right)$

## Properties of the dual control law

- The control law is derived using the Bellman optimization recursion.
- The dual properties manifested through $\boldsymbol{P}^{AA}_{i|k}$ (occurring in Bellman function), $\boldsymbol{P}^{BA}_{i|k}$ and $\boldsymbol{P}^{BB}_{i|k}$ which depend on $\boldsymbol{P}^{\boldsymbol{\theta}}_{i|k} = \mathrm{cov}_{\rho_k}\left(\boldsymbol{\theta}_i \middle| \mathfrak{I}_k\right)$ for $i = k, ..., k+m$.
- Only the mean value $\hat{\boldsymbol{s}}_k$ and first two moments of pdf's $p(\boldsymbol{\theta}_i|\mathfrak{I}_k)$ $i = k, \cdots, k+m$ are necessary.

# Numerical example

## Considered system

$$\mathbf{s}_{k+1} = \begin{pmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{pmatrix} \mathbf{s}_k + \begin{pmatrix} 0 \\ \theta_{3k} \end{pmatrix} u_k + \mathbf{w}_k$$
$$y_k = (0\ 1)\mathbf{s}_k + v_k$$

- Initial state and the real parameters
  - $\mathbf{s}_0 = (1,\ -0.5)^T$
  - $\boldsymbol{\theta}_k = (-2.0427,\ 0.3427,\ 1)^T, \qquad \forall k$
- Noise pdf's
  - $p(\mathbf{w}_k) = \mathcal{N}(0, 0.0001)$
  - $p(v_k) = \mathcal{N}(0,\ 0.001)$
- Prior pdf for EKF
  - $p(\mathbf{s}_0, \boldsymbol{\theta}_0) = \mathcal{N}\left((1, -0.5, -2.0427, 0.3427, 1)^T, 0.2\mathbf{I}\right)$

# Criteria parameters

## Criterion of the original optimization problem

$$J = E\left\{\sum_{k=0}^{N-1}(s_{k+1,2}-5)^2 + 0.001\cdot u_k^2\right\},$$

## Modified criterion for dual control derivation

$$J_k = E_{\rho_k}\left\{\sum_{i=k}^{k+m}(s_{i+1,2}-5)^2 + 0.001\cdot u_i^2 - 0.8485\cdot \boldsymbol{\nu}_{i+1}^2 \middle| \mathfrak{I}_i\right\},\quad k = 0,1,\ldots,N-1$$

$$m = 4 \quad Q_k = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad R_k = 0.001 \quad \Lambda_k = 0.8485$$

Notes:

- $\Lambda_k = 0$ in case of the the cautious controller (PCE)
- the HCE controller can be obtained employing the $\rho_k^{CE}$-approximation and setting $\Lambda_k = Q_k$

# Control quality comparison with non-dual controllers

| Controller | $\hat{J}$ | $\text{var}\{\hat{J}\}$ | $\sigma\{J\}$ |
|---|---|---|---|
| HCE | 108.768219 | 447.416351 | 1396.308509 |
| Cautious (PCE) | 5.026459 | 0.040193 | 13.453065 |
| Dual | 4.206371 | 0.052889 | 15.549870 |

- estimate of the original criterion value

$$\hat{J} = \frac{1}{M}\sum_{j=1}^{M} J_i, \quad \text{where} \quad J_i = \sum_{k=0}^{N-1} \mathcal{L}_i(s_i, \theta_i, u_i)$$

- variance of the criterion value estimate $\text{var}\{\hat{J}\}$ (determined using the bootstrap technique)
- standard deviation among the Monte Carlo runs

$$\sigma\{J\} = \sqrt{\frac{1}{M-1}\sum_{j=1}^{M} J_i - \hat{J}}$$

# Concluding remarks

## Resume

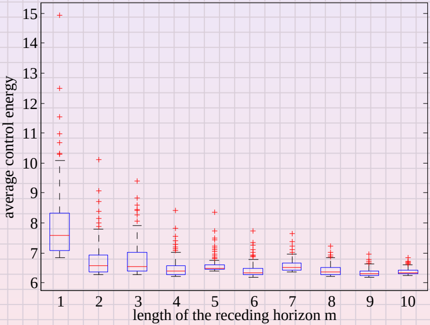
- non-myopic version of innovations dual adaptive controller was introduced
- some aspects of the criterion and control law were discussed

## Features of the new dual controller

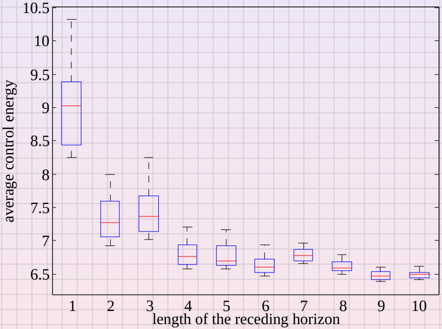
- ✔ clear criterion interpretation
  - modified criterion incorporates both aspects of dual control
  - makes it possible to tune the balance between caution and probing
- ✔ closed form solution available
- ✔ higher control quality compared to HCE and PCE controllers
- ✔ computationally moderate

# Receding horizon length

Comparison of average control energy for various receding horizon lengths.




$\Lambda_k \in <0, 1>$




$\Lambda_k \in <0.75, 0.95>$

Note: The non-myopic controllers (i.e. $m \geq 1$) use the control energy more wisely.