CT Identification

Optimal time invariant experiment design for regret minimization in linear quadratic adaptive control

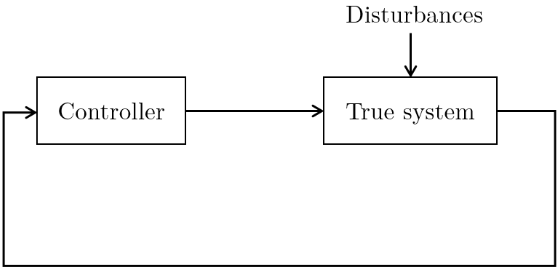
Kévin Colin[1], Håkan Hjalmarsson[1], Xavier Bombois[2,3]

[1] Division of Decision and Control Systems, KTH Royal Institute of Technology
[2] Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon
[3] Centre National de la Recherche Scientifique (CNRS)

November 23, 2023

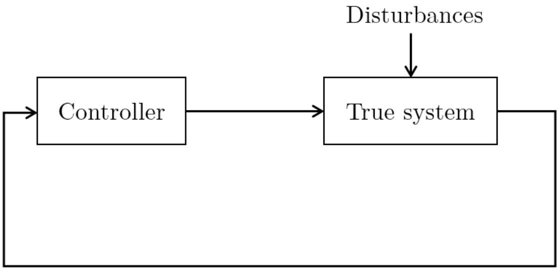
The problem

- **Objective**: design a model-based controller in order to reject the disturbances.



The problem

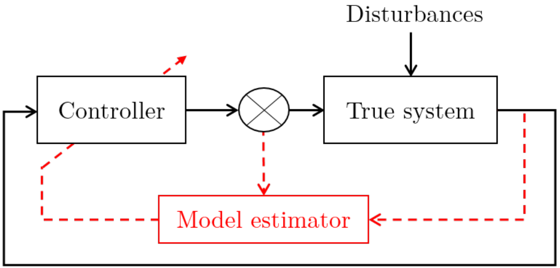
- **Objective**: design a model-based controller in order to reject the disturbances.



- True dynamics **unknown**...

The problem

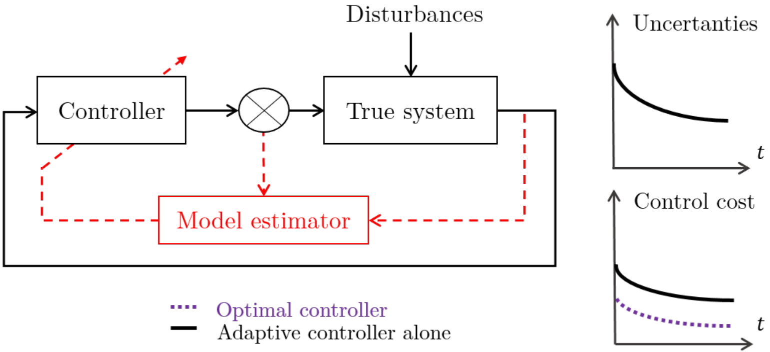
- Model-based adaptive controller



- System identification ingredient for the estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

The problem

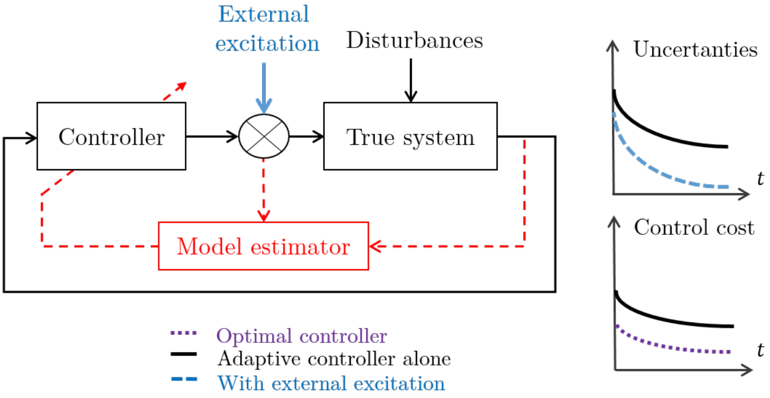
- Model-based adaptive controller



- System identification ingredients for the estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

The problem

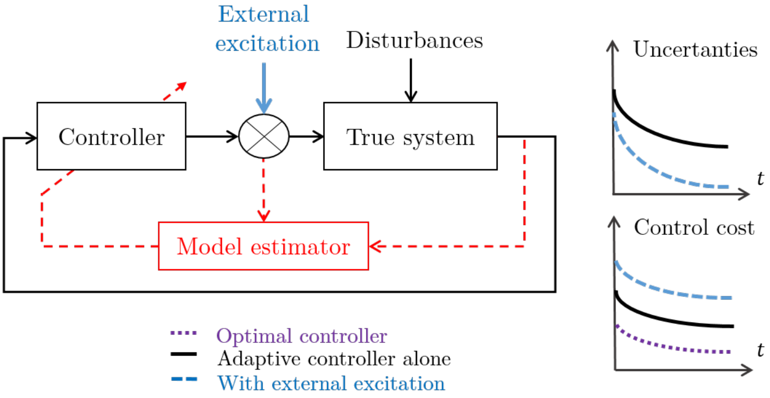
- Model-based adaptive controller + large external excitation



- System identification ingredient for the estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

The problem

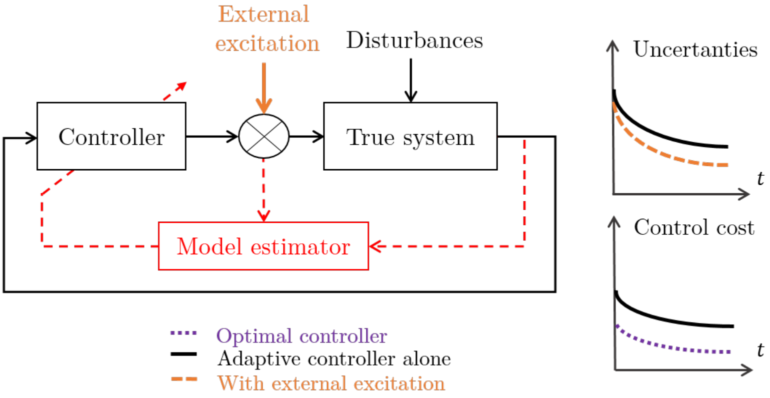
- Model-based adaptive controller + large external excitation



- System identification ingredient for the model parameter estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

The problem

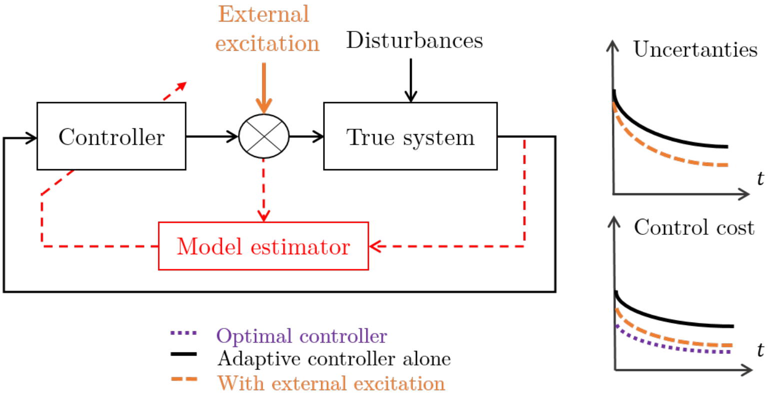
- Model-based adaptive controller + a more appropriate external excitation



- System identification ingredient for the model parameter estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

The problem

- Model-based adaptive controller + a more appropriate external excitation



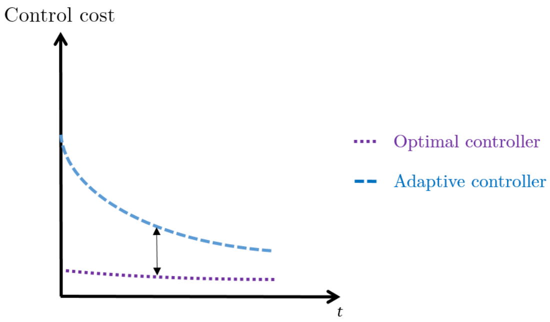
- System identification ingredient for the model parameter estimate $\hat{\theta}(t)$:
  - Data from the beginning of the experiment.
  - Full-order model structure (true parameter vector θ_0).

Regret = trade-off exploitation/exploration

- Regret = function of both the exploitation and exploration costs to be minimized.

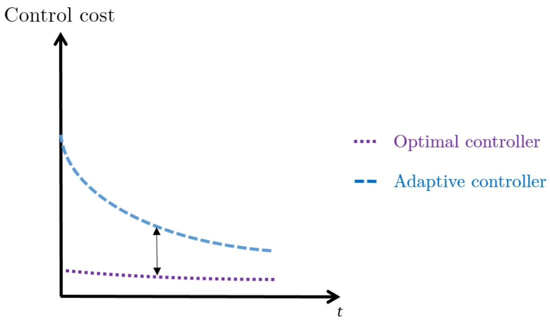
Regret = trade-off exploitation/exploration

- Regret = function of both the exploitation and exploration costs to be minimized.



Regret = trade-off exploitation/exploration

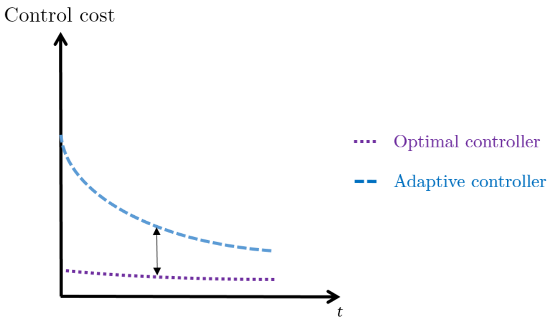
- Regret = function of both the exploitation and exploration costs to be minimized.



- Expectation of each gap at time instant t = instantaneous regret.

Regret = trade-off exploitation/exploration

- Regret = function of both the exploitation and exploration costs to be minimized.

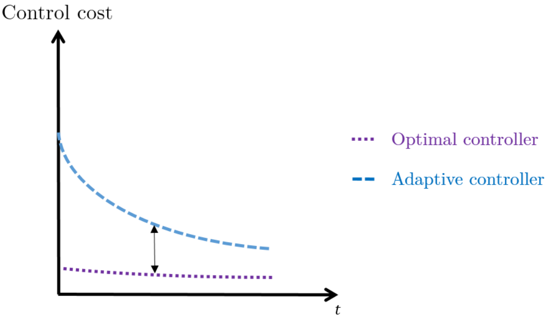


- Expectation of each gap at time instant t = instantaneous regret.

- Sum of the instantaneous regrets = cumulative regret.

Regret = trade-off exploitation/exploration

- Regret = function of both the exploitation and exploration costs to be minimized.



- Expectation of each gap at time instant t = instantaneous regret.

- Sum of the instantaneous regrets = cumulative regret.

- **Objective**: design an adaptive control algorithm minimizing the cumulative regret.

Considered problem

- Discrete-time state system with one input:

$$x(t+1) = \mathbf{A}(\theta_0)x(t) + \mathbf{B}(\theta_0)u(t) + e(t)$$
$$e(t) \sim N(0, \mathbf{\Sigma}_e)$$

Considered problem

- Discrete-time state system with one input:

$$x(t+1) = \mathbf{A}(\theta_0)x(t) + \mathbf{B}(\theta_0)u(t) + e(t)$$
$$e(t) \sim N(0, \mathbf{\Sigma}_e)$$

- Infinite horizon Linear Quadratic Control:

$$\tilde{u}(t) = -\mathbf{K}(\theta_0)\tilde{x}(t)$$

minimizing $\lim_{T\to+\infty} \dfrac{1}{T} \sum_{t=1}^{T} \mathbb{E}[x(t)^\top \mathbf{Q} x(t) + \mathbf{R}u^2(t)]$.

Considered problem

- Discrete-time state system with one input:

$$x(t+1) = \mathbf{A}(\theta_0)x(t) + \mathbf{B}(\theta_0)u(t) + e(t)$$
$$e(t) \sim N(0, \mathbf{\Sigma}_e)$$

- Infinite horizon Linear Quadratic Control:

$$\tilde{u}(t) = -\mathbf{K}(\theta_0)\tilde{x}(t)$$

minimizing $\lim_{T \to +\infty} \dfrac{1}{T} \sum_{t=1}^{T} \mathbb{E}[x(t)^\top \mathbf{Q} x(t) + \mathbf{R} u^2(t)]$.

- $\mathbf{K}(\theta_0)$ computed by solving a discrete-time Riccati equation, but dependent on θ_0....

Considered problem

- Model-based linear quadratic adaptive control:

$$u(t) = -\mathbf{K}(t)x(t) + v(t)$$

$\hat{\theta}(t)$: on-line least-squares estimator of θ_0 using past data $\{x(k), u(k)\}_{k=1}^{t}$

$v(t)$: user-defined external excitation

[1] Wang F. and Janson L., Exact asymptotics for linear quadratic adaptive control Journal of Machine Learning Research, 2021

Considered problem

- Model-based linear quadratic adaptive control:

$$u(t) = -\mathbf{K}(t)x(t) + v(t)$$

$\hat{\theta}(t)$: on-line least-squares estimator of θ_0 using past data $\{x(k), u(k)\}_{k=1}^{t}$

$v(t)$: user-defined external excitation

- Cumulative regret[1]

$$r(T) = \sum_{t=1}^{T} \Big(\underbrace{\mathbb{E}[x(t)^\top \mathbf{Q} x(t) + \mathbf{R} u(t)^2]}_{\text{Cost with adaptive control}} - \underbrace{\mathbb{E}[\tilde{x}(t)^\top \mathbf{Q} \tilde{x}(t) + \mathbf{R} \tilde{u}(t)^2]}_{\text{Optimal cost with } \mathbf{K}(\theta_0) \text{ and } v = 0} \Big)$$

where the expectation $\mathbb{E}$ is taken with respect to e and possibly v

[1] Wang F. and Janson L., Exact asymptotics for linear quadratic adaptive control Journal of Machine Learning Research, 2021

Considered problem

- Model-based linear quadratic adaptive control:

$$u(t) = -\mathbf{K}(t)x(t) + v(t)$$

$\hat{\theta}(t)$: on-line least-squares estimator of θ_0 using past data $\{x(k), u(k)\}_{k=1}^{t}$

$v(t)$: user-defined external excitation

- Cumulative regret[1]

$$r(T) = \sum_{t=1}^{T} \Big(\underbrace{\mathbb{E}[x(t)^\top \mathbf{Q} x(t) + \mathbf{R} u(t)^2]}_{\text{Cost with adaptive control}} - \underbrace{\mathbb{E}[\tilde{x}(t)^\top \mathbf{Q} \tilde{x}(t) + \mathbf{R} \tilde{u}(t)^2]}_{\text{Optimal cost with } \mathbf{K}(\theta_0) \text{ and } v\,=\,0} \Big)$$

where the expectation $\mathbb{E}$ is taken with respect to e and possibly v

Objective

Design the control policy $\hat{\theta}(t) \to \mathbf{K}(t)$ and the external excitation sequence $\{v(t)\}_{t=1}^{T}$ such that $r(T)$ is minimized.

[1] Wang F. and Janson L., Exact asymptotics for linear quadratic adaptive control Journal of Machine Learning Research, 2021

- Several classes of adaptive control algorithms in the literature

Certainty equivalence

$\hat{\theta}(t)$ assumed to be the true parameter vector θ_0 in the controller design.

Notations: $\mathbf{K}(\hat{\theta}(t))$

Considered problem

- Several classes of adaptive control algorithms in the literature

Certainty equivalence

$\hat{\theta}(t)$ assumed to be the true parameter vector θ_0 in the controller design.

Notations: $\mathbf{K}(\hat{\theta}(t))$

Optimism in the Face of Uncertainties

From the set of parameter uncertainties $\delta\hat{\theta}(t)$, minimize the lower bound of the regret

$$\min_{\substack{\mathbf{K}(t)\\ v(t)}} \min_{\delta\theta(t)} r(T, \delta\theta(t))$$

Often intractable

Literature on LQ regret minimization

- Mostly: results in the asymptotic domain.

[2] I. Ziemann and H. Sandberg. Regret lower bounds for learning linear quadratic Gaussian systems

[3] Y. Jedra and A. Proutiere. Minimal expected regret in linear quadratic control. In International Conference on Artificial Intelligence and Statistics, pages 10234–10321. PMLR, 2022

Literature on LQ regret minimization

- Mostly: results in the asymptotic domain.
- Minimal rate for the regret $r(T)$ is[2] [3]
  - $\mathcal{O}(\sqrt{T})$ if both $\mathbf{A}_0$ and $\mathbf{B}_0$ are unknown.
  - $\mathcal{O}(\log(T))$ if either $\mathbf{A}_0$ or $\mathbf{B}_0$ is known.

[2] I. Ziemann and H. Sandberg. Regret lower bounds for learning linear quadratic Gaussian systems

[3] Y. Jedra and A. Proutiere. Minimal expected regret in linear quadratic control. In International Conference on Artificial Intelligence and Statistics, pages 10234–10321. PMLR, 2022

Literature on LQ regret minimization

- How to reach this optimal rate?

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...
  - ... until the estimate uncertainties are below a given threshold...

[4]Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5]F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...
  - ... until the estimate uncertainties are below a given threshold...
  - ... a new controller is computed by randomly drawing a parameter vector in the uncertain set...

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...
  - ... until the estimate uncertainties are below a given threshold...
  - ... a new controller is computed by randomly drawing a parameter vector in the uncertain set...
  - ... and we reiterate.

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...
  - ... until the estimate uncertainties are below a given threshold...
  - ... a new controller is computed by randomly drawing a parameter vector in the uncertain set...
  - ... and we reiterate.
- Certainty equivalence + an external excitation[5] of the form

$$v(t) \sim N\left(0, \frac{a}{\sqrt{t}}\right) \qquad a > 0$$

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Literature on LQ regret minimization

- How to reach this optimal rate?
- Optimism in Face of Uncertainties: Thompson sampling without external excitation[4]
  - Controller kept constant during several time instants...
  - ... until the estimate uncertainties are below a given threshold...
  - ... a new controller is computed by randomly drawing a parameter vector in the uncertain set...
  - ... and we reiterate.
- Certainty equivalence + an external excitation[5] of the form

$$v(t) \sim N\left(0, \frac{a}{\sqrt{t}}\right) \qquad a > 0$$

- **Question**: What about finite-time regret minimization? (not clear what is optimal)

[4] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

[5] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

• **Linear time-invariant** Application-oriented experiment design[6] [7] [8] $=$ design a **finite-time** (horizon T) excitation minimizing the expected value of a given cost.

[6] X. Bombois, G. Scorletti, M. Gevers , P.M.J. Van den Hof, R. Hildebrand: Least costly identification experiment for control. Automatica 42(10), 1651–1662 (2006)

[7] H. Jansson (2004). Experiment design with applications in identification for control (Doctoral dissertation, Signaler, sensorer och system).

[8] H. Hjalmarsson: System identification of complex and structured systems. European Journal of Control 15(4), 275–310 (2009).

• **Linear time-invariant** Application-oriented experiment design[6] [7] [8] $=$ design a **finite-time** (horizon T) excitation minimizing the expected value of a given cost.

Problem/Approach of this talk

Using the tools from **linear time-invariant** experiment design, can we develop an algorithm for finite-time regret minimization ?

[6] X. Bombois, G. Scorletti, M. Gevers , P.M.J. Van den Hof, R. Hildebrand: Least costly identification experiment for control. Automatica 42(10), 1651–1662 (2006)

[7] H. Jansson (2004). Experiment design with applications in identification for control (Doctoral dissertation, Signaler, sensorer och system).

[8] H. Hjalmarsson: System identification of complex and structured systems. European Journal of Control 15(4), 275–310 (2009).

- **Linear time-invariant** Application-oriented experiment design[6] [7] [8] = design a **finite-time** (horizon T) excitation minimizing the expected value of a given cost.

Problem/Approach of this talk

Using the tools from **linear time-invariant** experiment design, can we develop an algorithm for finite-time regret minimization ?

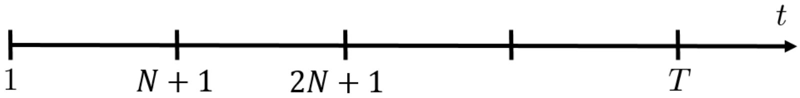
- How???

[6] X. Bombois, G. Scorletti, M. Gevers , P.M.J. Van den Hof, R. Hildebrand: Least costly identification experiment for control. Automatica 42(10), 1651–1662 (2006)

[7] H. Jansson (2004). Experiment design with applications in identification for control (Doctoral dissertation, Signaler, sensorer och system).

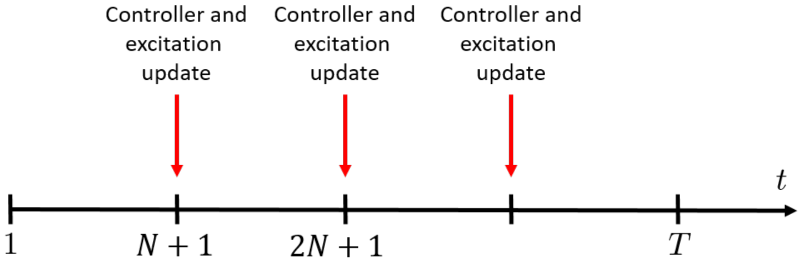
[8] H. Hjalmarsson: System identification of complex and structured systems. European Journal of Control 15(4), 275–310 (2009).

- **Solution**[9]: Design by intervals of "large" duration N ($T = n_{int}N$)



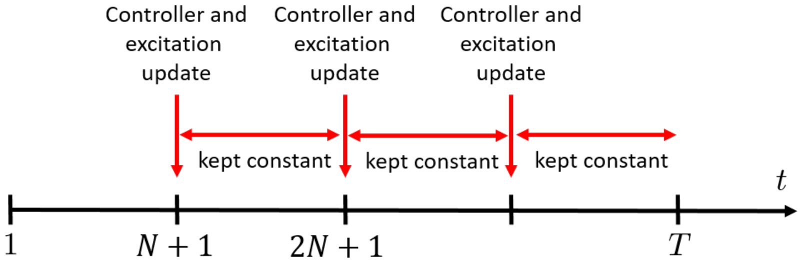
[9] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

- **Solution**[9]: Design by intervals of "large" duration N ($T = n_{int}N$)



[9] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

- **Solution**[9]: Design by intervals of "large" duration N $(T = n_{int}N)$



[9] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

A model of the cumulative regret

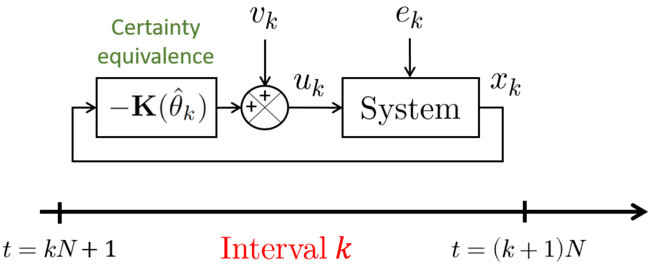
- Notations:

A model of the cumulative regret

- Notations:

$$\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$$

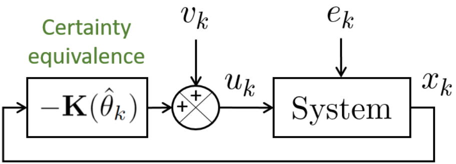
- Notations:



$$\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$$

A model of the cumulative regret

- Notations:

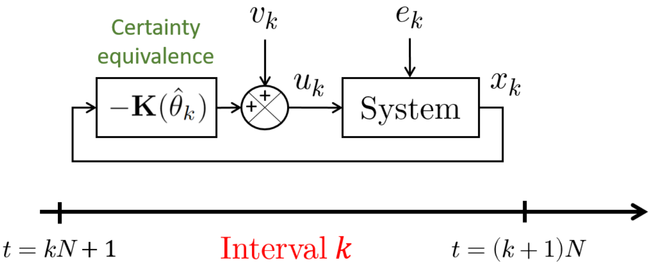


$$\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$$

- Assumption 1: we have an initial estimate $\hat{\theta}_1 \sim N(\theta_0, \boldsymbol{\mathcal{I}}_1^{-1})$ (if not, perform an initial identification).

A model of the cumulative regret

- Notations:



$$\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$$

- Assumption 1: we have an initial estimate $\hat{\theta}_1 \sim N(\theta_0, \boldsymbol{\mathcal{I}}_1^{-1})$ (if not, perform an initial identification).

- Assumption 2: all the future controllers $\{\mathbf{K}(\hat{\theta}_k)\}_{k=2}^{n_{int}}$ stabilize the loop.

A model of the cumulative regret

- Let us rewrite the regret

$$r(T) = \sum_{k=1}^{n_{int}} r_k$$

$$r_k = \sum_{\tau=1}^{N} (\mathbb{E}[x_k(\tau)^\top \mathbf{Q} x_k(\tau) + \mathbf{R} u_k(\tau)^2] - \mathbb{E}[\tilde{x}_k(\tau)^\top \mathbf{Q} \tilde{x}_k(\tau) + \mathbf{R} \tilde{u}_k(\tau)^2])$$

A model of the cumulative regret

- Let us rewrite the regret

$$r(T) = \sum_{k=1}^{n_{int}} r_k$$

$$r_k = \sum_{\tau=1}^{N} (\mathbb{E}[x_k(\tau)^\top \mathbf{Q} x_k(\tau) + \mathbf{R} u_k(\tau)^2] - \mathbb{E}[\tilde{x}_k(\tau)^\top \mathbf{Q} \tilde{x}_k(\tau) + \mathbf{R} \tilde{u}_k(\tau)^2])$$

- Given a stationary signal $y(t)$ with large number $N \gg 1$ of data $y(t)$, we have

$$\sum_{t=1}^{N} y(t)^\top y(t) \approx N||y||^2$$

A model of the cumulative regret

- Let us rewrite the regret

$$r(T) = \sum_{k=1}^{n_{int}} r_k$$

$$r_k = \sum_{\tau=1}^{N} (\mathbb{E}[x_k(\tau)^\top \mathbf{Q} x_k(\tau) + \mathbf{R} u_k(\tau)^2] - \mathbb{E}[\tilde{x}_k(\tau)^\top \mathbf{Q} \tilde{x}_k(\tau) + \mathbf{R} \tilde{u}_k(\tau)^2])$$

- Given a stationary signal $y(t)$ with large number $N \gg 1$ of data $y(t)$, we have

$$\sum_{t=1}^{N} y(t)^\top y(t) \approx N||y||^2$$

$\rightarrow r_k \approx N\mathbb{E}[||x_k||_{\mathbf{Q}}^2 + ||u_k||_{\mathbf{R}}^2 - ||\tilde{x}_k||_{\mathbf{Q}}^2 - ||\tilde{u}_k||_{\mathbf{R}}^2]$ if $N \gg 1$.

A model of the cumulative regret

- Linear system:

$$\begin{aligned} \text{Adaptive loop}: \quad & x_k = x_{e,k} + x_{v,k} & \qquad u_k = u_{e,k} + u_{v,k} \\ \text{Ideal loop}: \quad & \tilde{x}_k = \tilde{x}_{e,k} & \qquad \tilde{u}_k = \tilde{u}_{e,k} \end{aligned}$$

A model of the cumulative regret

- Linear system:

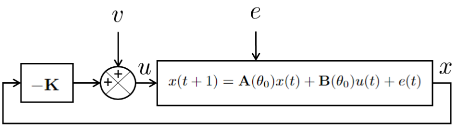
$$\text{Adaptive loop}: \quad x_k = x_{e,k} + x_{v,k} \qquad\qquad u_k = u_{e,k} + u_{v,k}$$
$$\text{Ideal loop}: \quad \tilde{x}_k = \tilde{x}_{e,k} \qquad\qquad \tilde{u}_k = \tilde{u}_{e,k}$$

- Stationary stochastic v_k independent from e_k, we have

$$r_k \approx N(r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)))$$

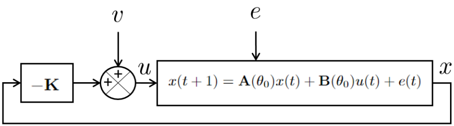
$$r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \mathbb{E}[||x_{e,k}||^2_{\mathbf{Q}} + ||u_{e,k}||^2_{\mathbf{R}} - ||\tilde{x}_{e,k}||^2_{\mathbf{Q}} - ||\tilde{u}_{e,k}||^2_{\mathbf{R}}] \quad \longrightarrow \quad \text{exploitation regret}$$
$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \mathbb{E}[||x_{v,k}||^2_{\mathbf{Q}} + ||u_{v,k}||^2_{\mathbf{R}}] \quad \longrightarrow \quad \text{exploration regret}$$

A model of the cumulative regret



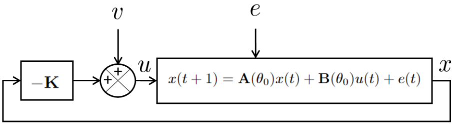
$$\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}) = (z\mathbf{I} - (\mathbf{A}(\theta_0) - \mathbf{B}(\theta_0)\mathbf{K}))^{-1}$$
$$\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}) = \mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$
$$\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}) = -\mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})$$
$$\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}) = \mathbf{I} - \mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$

A model of the cumulative regret



$$\mathbf{T}_{xe}(z, \theta_0, \mathbf{K}) = (z\mathbf{I} - (\mathbf{A}(\theta_0) - \mathbf{B}(\theta_0)\mathbf{K}))^{-1}$$
$$\mathbf{T}_{xv}(z, \theta_0, \mathbf{K}) = \mathbf{T}_{xe}(z, \theta_0, \mathbf{K})\mathbf{B}(\theta_0)$$
$$\mathbf{T}_{ue}(z, \theta_0, \mathbf{K}) = -\mathbf{K}\mathbf{T}_{xe}(z, \theta_0, \mathbf{K})$$
$$\mathbf{T}_{uv}(z, \theta_0, \mathbf{K}) = \mathbf{I} - \mathbf{K}\mathbf{T}_{xe}(z, \theta_0, \mathbf{K})\mathbf{B}(\theta_0)$$

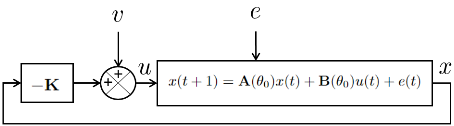
- $N \gg 1 \rightarrow$ transient dynamics from controller changes are negligible in r_k.



$$\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}) = (z\mathbf{I} - (\mathbf{A}(\theta_0) - \mathbf{B}(\theta_0)\mathbf{K}))^{-1}$$
$$\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}) = \mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$
$$\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}) = -\mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})$$
$$\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}) = \mathbf{I} - \mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$

- $N \gg 1 \rightarrow$ transient dynamics from controller changes are negligible in r_k.
- Consequence:

$$\begin{aligned} r_{e,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) \approx\ & \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{R}}] \\ & - \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{Q}} - ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{R}}] \\ r_{v,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) \approx\ & \mathbb{E}[||\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{R}}] \end{aligned}$$

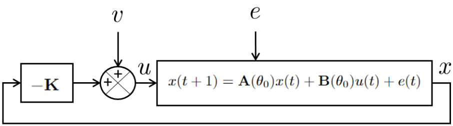


$$\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}) = (z\mathbf{I} - (\mathbf{A}(\theta_0) - \mathbf{B}(\theta_0)\mathbf{K}))^{-1}$$
$$\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}) = \mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$
$$\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}) = -\mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})$$
$$\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}) = \mathbf{I} - \mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$

- $N \gg 1 \rightarrow$ transient dynamics from controller changes are negligible in r_k.
- Consequence:

$$\begin{aligned} r_{e,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) &\approx \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{R}}] \\ &\quad - \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{Q}} - ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{R}}] \\ r_{v,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) &\approx \mathbb{E}[||\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{R}}] \end{aligned}$$

$\rightarrow$ we design the power spectrum densities (PSD) ϕ_{v_k} of each v_k.



$$\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}) = (z\mathbf{I} - (\mathbf{A}(\theta_0) - \mathbf{B}(\theta_0)\mathbf{K}))^{-1}$$
$$\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}) = \mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$
$$\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}) = -\mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})$$
$$\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}) = \mathbf{I} - \mathbf{K}\mathbf{T}_{xe}(z,\theta_0,\mathbf{K})\mathbf{B}(\theta_0)$$

- $N \gg 1 \rightarrow$ transient dynamics from controller changes are negligible in r_k.
- Consequence:

$$\begin{aligned} r_{e,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) &\approx \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))e_k||^2_{\mathbf{R}}] \\ &\quad - \mathbb{E}[||\mathbf{T}_{xe}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{Q}} - ||\mathbf{T}_{ue}(z,\theta_0,\mathbf{K}(\theta_0))e_k||^2_{\mathbf{R}}] \\ r_{v,k}(\theta_0,\mathbf{K}(\hat{\theta}_k)) &\approx \mathbb{E}[||\mathbf{T}_{xv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{uv}(z,\theta_0,\mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{R}}] \end{aligned}$$

$\rightarrow$ we design the power spectrum densities (PSD) ϕ_{v_k} of each v_k.

First objective

Reformulate the regret minimization problem as a convex optimization problem in the PSDs $\phi_{v_1}, \cdots, \phi_{v_{n_{int}}}$.

Reformulation of the regret minimization problem

- We have to make ϕ_{v_k} appear in both $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ and $r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$.

Reformulation of the regret minimization problem

- We have to make ϕ_{v_k} appear in both $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ and $r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$.

- Exploration regret $\rightarrow$ Parseval's theorem:

$$\begin{aligned} r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) &= \mathbb{E}[||\mathbf{T}_{xv}(z, \theta_0, \mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{uv}(z, \theta_0, \mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{R}}] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))]\phi_{v_k}(\omega)d\omega \end{aligned}$$

Reformulation of the regret minimization problem

- We have to make ϕ_{v_k} appear in both $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ and $r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$.

- Exploration regret $\rightarrow$ Parseval's theorem:

$$\begin{aligned} r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) &= \mathbb{E}[||\mathbf{T}_{xv}(z, \theta_0, \mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{Q}} + ||\mathbf{T}_{uv}(z, \theta_0, \mathbf{K}(\hat{\theta}_k))v_k||^2_{\mathbf{R}}] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))]\phi_{v_k}(\omega)d\omega \end{aligned}$$

where

$$\begin{aligned} \mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k)) = \mathbf{T}^*_{xv}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))\mathbf{Q}\mathbf{T}_{xv}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k)) \\ + R|\mathbf{T}_{uv}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))|^2 \end{aligned}$$

Reformulation of the regret minimization problem

• Exploitation regret $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$: linked to uncertainties $\mathbb{E}[(\hat{\theta}_k - \theta_0)(\hat{\theta}_k - \theta_0)^\top] =$ inverse of the Fisher information matrix $\boldsymbol{\mathcal{I}}_k$ which depends linearly[10] on $\phi_{v_1}, \cdots, \phi_{v_{k-1}}$.

[10] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Reformulation of the regret minimization problem

• Exploitation regret $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$: linked to uncertainties $\mathbb{E}[(\hat{\theta}_k - \theta_0)(\hat{\theta}_k - \theta_0)^\top] =$ inverse of the Fisher information matrix $\mathcal{I}_k$ which depends linearly[10] on $\phi_{v_1}, \cdots, \phi_{v_{k-1}}$.

• Observation: the global minimum of $\hat{\theta}_k \to r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ is minimum 0 and it is reached at $\hat{\theta}_k = \theta_0$ (LQ control property).

[10] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Reformulation of the regret minimization problem

- Exploitation regret $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$: linked to uncertainties $\mathbb{E}[(\hat{\theta}_k - \theta_0)(\hat{\theta}_k - \theta_0)^\top] =$ inverse of the Fisher information matrix $\boldsymbol{\mathcal{I}}_k$ which depends linearly[10] on $\phi_{v_1}, \cdots, \phi_{v_{k-1}}$.

- Observation: the global minimum of $\hat{\theta}_k \to r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ is minimum 0 and it is reached at $\hat{\theta}_k = \theta_0$ (LQ control property).

- Second-order Taylor expansion of $\hat{\theta}_k \to r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ evaluated at θ_0

$$\begin{aligned} r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) &\approx \frac{1}{2}\mathbb{E}[(\hat{\theta}_k - \theta_0)^\top \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))(\hat{\theta}_k - \theta_0)] \\ &\approx \frac{1}{2}\,\mathrm{tr}(\mathbf{W}(\theta_0, \mathbf{K}(\theta_0))\boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}})^{-1}) \end{aligned}$$

[10] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Reformulation of the regret minimization problem

- Exploitation regret $r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$: linked to uncertainties $\mathbb{E}[(\hat{\theta}_k - \theta_0)(\hat{\theta}_k - \theta_0)^\top] =$ inverse of the Fisher information matrix $\boldsymbol{\mathcal{I}}_k$ which depends linearly[10] on $\phi_{v_1}, \cdots, \phi_{v_{k-1}}$.

- Observation: the global minimum of $\hat{\theta}_k \to r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ is minimum 0 and it is reached at $\hat{\theta}_k = \theta_0$ (LQ control property).

- Second-order Taylor expansion of $\hat{\theta}_k \to r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))$ evaluated at θ_0

$$\begin{aligned} r_{e,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) &\approx \frac{1}{2}\mathbb{E}[(\hat{\theta}_k - \theta_0)^\top \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))(\hat{\theta}_k - \theta_0)] \\ &\approx \frac{1}{2}\,\mathrm{tr}(\mathbf{W}(\theta_0, \mathbf{K}(\theta_0))\boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}})^{-1}) \end{aligned}$$

- With $\mathbf{H}(k) \succeq \mathbf{W}(\theta_0, \mathbf{K}(\hat{\theta}_k))^{1/2}\boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}})^{-1}\mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2}$, minimizing $r_{e,k}(\theta_0, , \mathbf{K}(\theta_0))$ is equivalent to minimizing $\mathrm{tr}(\mathbf{H}(k))$ such that

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

[10] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

- Problem to be solved:

$$\min_{\phi_{v_k}, \mathbf{H}(k)} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2} \operatorname{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) \right)$$

Subject to, for each $k = 1, \cdots, n_{int}$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))] \phi_{v_k}(\omega) d\omega$$

- Problem to be solved:

$$\min_{\phi_{v_k}, \mathbf{H}(k)} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2} \operatorname{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) \right)$$

Subject to, for each $k = 1, \cdots, n_{int}$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))] \phi_{v_k}(\omega) d\omega$$

- With an adequate parametrization of each ϕ_{v_k}, we can transform the problem into a **convex semi-definite programming (SDP)**.

Reformulation of the regret minimization problem

- Problem to be solved:

$$\min_{\phi_{v_k}, \mathbf{H}(k)} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2} \operatorname{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) \right)$$

Subject to, for each $k = 1, \cdots, n_{int}$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))] \phi_{v_{k-1}}(\omega) d\omega$$

Reformulation of the regret minimization problem

- Problem to be solved:

$$\min_{\phi_{v_k}, \mathbf{H}(k)} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2} \operatorname{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) \right)$$

Subject to, for each $k = 1, \cdots, n_{int}$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))] \phi_{v_{k-1}}(\omega) d\omega$$

- **Observation 1**: the estimate $\hat{\theta}_{n_{int}}$ of the last interval does not depend on $v_{n_{int}}$. Hence, $v_{n_{int}}$ only penalizes the regret... We should set it to 0.

Reformulation of the regret minimization problem

- Problem to be solved:

$$\min_{\phi_{v_k}, \mathbf{H}(k)} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2} \operatorname{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) \right)$$

Subject to, for each $k = 1, \cdots, n_{int}$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_k(\phi_{v_1}, \cdots, \phi_{v_{k-1}}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))] \phi_{v_{k-1}}(\omega) d\omega$$

- **Observation 1**: the estimate $\hat{\theta}_{n_{int}}$ of the last interval does not depend on $v_{n_{int}}$. Hence, $v_{n_{int}}$ only penalizes the regret... We should set it to 0.

- **Observation 2**: the estimate $\hat{\theta}_1$ of the first interval does not depend on any v_k. Hence, $\operatorname{tr}(\mathbf{H}(1)) =$ constante. We can remove this term and the corresponding LMI.

- Problem to be solved:

$$\min_{\phi_{v_k},\mathbf{H}(k)} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(k)) + r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k))\right)$$

Subject to, for each $k = 1, \cdots, n_{int}-1$

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \mathcal{I}_k(\phi_{v_1}, \cdots, \phi_{v_k}) \end{pmatrix} \succeq 0$$

$$r_{v,k}(\theta_0, \mathbf{K}(\hat{\theta}_k)) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}[\mathcal{D}(e^{j\omega}, \theta_0, \mathbf{K}(\hat{\theta}_k))]\phi_{v_k}(\omega)d\omega$$

- **Observation 1**: the estimate $\hat{\theta}_{n_{int}}$ of the last interval does not depend on $v_{n_{int}}$. Hence, $v_{n_{int}}$ only penalizes the regret... We should set it to 0.

- **Observation 2**: the estimate $\hat{\theta}_1$ of the first interval does not depend on any v_k. Hence, $\mathrm{tr}(\mathbf{H}(1)) =$ constante. We can remove this term and the corresponding LMI.

Reformulation of the regret minimization problem

$$\phi_{v_k}(\omega) = c_0(k) + 2\sum_{q=1}^{m} c_q(k)\cos(q\omega)$$

Reformulation of the regret minimization problem

$$\phi_{v_k}(\omega) = c_0(k) + 2\sum_{q=1}^{m} c_q(k)\cos(q\omega) \quad \xRightarrow{\text{Spectral factorization}} \quad v_k = F_k(z)w$$

FIR of order m

White noise
of unit variance

$$\phi_{v_k}(\omega) = c_0(k) + 2\sum_{q=1}^{m} c_q(k)\cos(q\omega) \quad \xrightarrow{\text{Spectral factorization}} \quad v_k = F_k(z)w$$

FIR of order m

White noise of unit variance

- $\phi_{v_k}(\omega) \geq 0 \ \forall \omega \in]-\pi, \pi]$ guaranteed if and only if it exists $\mathbf{Z}(k) = \mathbf{Z}(k)^\top$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

with

$$\boldsymbol{\mathcal{A}} = \begin{pmatrix} 0_{1\times(m-1)} & 1 \\ 0_{(m-1)\times(m-1)} & 0_{(m-1)\times 1} \end{pmatrix} \qquad \boldsymbol{\mathcal{B}} = \begin{pmatrix} 1 \\ 0_{(m-1)\times 1} \end{pmatrix} \qquad c_{1:m}(k) = \begin{pmatrix} c_1(k) \\ \vdots \\ c_m(k) \end{pmatrix}$$

(Positive real lemma)

Reformulation of the regret minimization problem

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N\sum_{k=1}^{n_{int}}\left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(k))+\sum_{q=0}^{m}\beta_q(\theta_0,\mathbf{K}(\hat{\theta}_k))c_q(k)\right)$$

Subject to , for each $k=1,\cdots,n_{int}$,

$$\begin{pmatrix}\mathbf{H}(k) & \mathbf{W}(\theta_0,\mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0,\mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_1+\sum_{i=1}^{k-1}\boldsymbol{\mathcal{L}}_e(\theta_0,\mathbf{K}(\hat{\theta}_i))+\sum_{i=1}^{k-1}\sum_{q=0}^{m}c_q(i)\boldsymbol{\mathcal{L}}_{v,q}(\theta_0,\mathbf{K}(\hat{\theta}_i))\end{pmatrix}\succeq 0$$

$$\begin{pmatrix}\mathbf{Z}(k)-\boldsymbol{\mathcal{A}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_{1:m}(k)-\boldsymbol{\mathcal{A}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k)-\boldsymbol{\mathcal{B}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_0(k)-\boldsymbol{\mathcal{B}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{B}}\end{pmatrix}\succeq 0$$

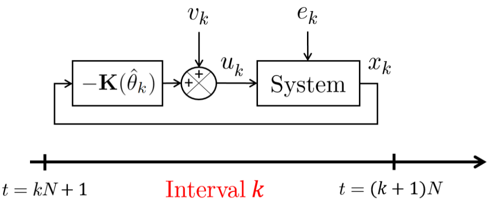
Reformulation of the regret minimization problem

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\theta_0,, \mathbf{K}(\hat{\theta}_k)) c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_1 + \sum_{i=1}^{k} \boldsymbol{\mathcal{L}}_e(\theta_0, \mathbf{K}(\hat{\theta}_i)) + \sum_{i=1}^{k} \sum_{q=0}^{m} c_q(i) \boldsymbol{\mathcal{L}}_{v,q}(\theta_0, \mathbf{K}(\hat{\theta}_i)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$



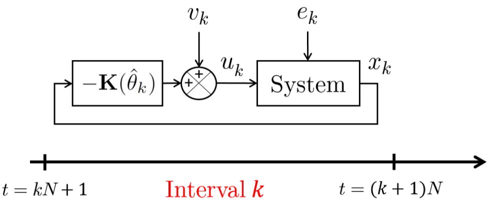
$\beta_q(\theta_0, \mathbf{K}(\hat{\theta}_k)) \longrightarrow$ Exploration penalization coefficients using a filtered white noise excitation

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\theta_0,, \mathbf{K}(\hat{\theta}_k)) c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \mathcal{I}_1 + \sum_{i=1}^{k} \mathcal{L}_e(\theta_0, \mathbf{K}(\hat{\theta}_i)) + \sum_{i=1}^{k}\sum_{q=0}^{m} c_q(i)\mathcal{L}_{v,q}(\theta_0, \mathbf{K}(\hat{\theta}_i)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \mathcal{A}^\top \mathbf{Z}(k)\mathcal{A} & c_{1:m}(k) - \mathcal{A}^\top \mathbf{Z}(k)\mathcal{B} \\ c_{1:m}^\top(k) - \mathcal{B}^\top \mathbf{Z}(k)\mathcal{A} & c_0(k) - \mathcal{B}^\top \mathbf{Z}(k)\mathcal{B} \end{pmatrix} \succeq 0$$



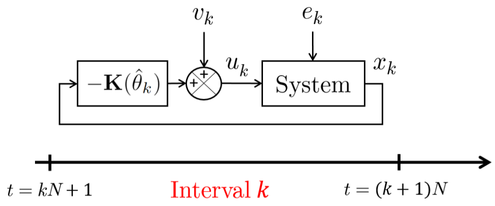
$\mathcal{L}_e(\theta_0, \mathbf{K}(\hat{\theta}_k)) \longrightarrow$ Information gained during interval k from the feedback of the noise e_k

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\theta_0,,\mathbf{K}(\hat{\theta}_k))c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0,\mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0,\mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_1 + \sum_{i=1}^{k} \boldsymbol{\mathcal{L}}_e(\theta_0,\mathbf{K}(\hat{\theta}_i)) + \sum_{i=1}^{k}\sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q}(\theta_0,\mathbf{K}(\hat{\theta}_i)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top\mathbf{Z}(k)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$



$$\sum_{q=0}^{m} c_q(k)\boldsymbol{\mathcal{L}}_{v,q}(\theta_0,\mathbf{K}(\hat{\theta}_k)) \longrightarrow$$ Information gained during interval k from the filtered white noise excitation v_k

Reformulation of the regret minimization problem

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}} \left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\theta_0, \mathbf{K}(\hat{\theta}_k)) c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\theta_0))^{1/2} & \boldsymbol{\mathcal{I}}_1 + \sum_{i=1}^{k-1} \boldsymbol{\mathcal{L}}_e(\theta_0, \mathbf{K}(\hat{\theta}_i)) + \sum_{i=1}^{k} \sum_{q=0}^{m} c_q(i) \boldsymbol{\mathcal{L}}_{v,q}(\theta_0, \mathbf{K}(\hat{\theta}_i)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Problem 1: we don't know θ_0...

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\theta_0, \mathbf{K}(\hat{\theta}_k)) c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\theta_0, \mathbf{K}(\hat{\theta}_k))^{1/2} \\ \mathbf{W}(\theta_0, \mathbf{K}(\hat{\theta}_k))^{1/2} & \boldsymbol{\mathcal{I}}_1 + \sum_{i=1}^{k} \boldsymbol{\mathcal{L}}_e(\theta_0, \mathbf{K}(\hat{\theta}_i)) + \sum_{i=1}^{k} \sum_{q=0}^{m} c_q(i) \boldsymbol{\mathcal{L}}_{v,q}(\theta_0, \mathbf{K}(\hat{\theta}_i)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k) \boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Problem 1: we don't know θ_0...
- Problem 2: $\mathbf{K}(\hat{\theta}_2), \cdots, \mathbf{K}(\hat{\theta}_{n_{int}})$ unknown as well (nonlinear function of the $c_q(k)$)...

Reformulation of the regret minimization problem

$$\min_{\substack{\mathbf{Z}(k)=\mathbf{Z}(k)^\top \\ c_q(k),\mathbf{H}(k)}} N \sum_{k=1}^{n_{int}-1} \left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(k)) + \sum_{q=0}^{m} \beta_q(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1)) c_q(k) \right)$$

Subject to , for each $k = 1, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(k) & \mathbf{W}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))^{1/2} \\ \mathbf{W}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))^{1/2} & \boldsymbol{\mathcal{I}}_1 + k\boldsymbol{\mathcal{L}}_e(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1)) + \sum_{i=1}^{k}\sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1)) \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_{1:m}(k) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{A}} & c_0(k) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(k)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Problem 1: we don't know θ_0...
- Problem 2: $\mathbf{K}(\hat{\theta}_2), \cdots, \mathbf{K}(\hat{\theta}_{n_{int}})$ unknown as well (nonlinear function of the $c_q(k)$)...
- Solution: we replace θ_0 and all future estimates $\hat{\theta}_k$.

Dealing with chicken-and-egg issue

- Once the SDP is solved, we should excite the control loop with the signals $v_k = \mathbf{F}_k^{opt}(z)w$.

[11] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Dealing with chicken-and-egg issue

- Once the SDP is solved, we should excite the control loop with the signals $v_k = \mathbf{F}_k^{opt}(z)w$.

- However, the design is done based on $\hat{\theta}_1$ which might not be an accurate estimate of θ_0...

[11] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Dealing with chicken-and-egg issue

- Once the SDP is solved, we should excite the control loop with the signals $v_k = \mathbf{F}_k^{opt}(z)w$.

- However, the design is done based on $\hat{\theta}_1$ which might not be an accurate estimate of θ_0...

- Observation:
  - $\hat{\theta}_2$ is expected to be less uncertain than $\hat{\theta}_1$.
  - $\hat{\theta}_3$ is expected to be less uncertain than $\hat{\theta}_2$.
  - etc.

[11] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Dealing with chicken-and-egg issue

- Once the SDP is solved, we should excite the control loop with the signals $v_k = \mathbf{F}_k^{opt}(z)w$.

- However, the design is done based on $\hat{\theta}_1$ which might not be an accurate estimate of θ_0...

- Observation:
  - $\hat{\theta}_2$ is expected to be less uncertain than $\hat{\theta}_1$.
  - $\hat{\theta}_3$ is expected to be less uncertain than $\hat{\theta}_2$.
  - etc.

- **Idea**: receding horizon design[11] of the PSDs ϕ_{v_k}.

[11] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Receding horizon design

$\hat{\theta}_1 \quad \boldsymbol{\mathcal{I}}_1$

$\hat{\theta}_1 \quad \boldsymbol{\mathcal{I}}_1$

$\mathbf{K}(\hat{\theta}_1)$

Receding horizon design

$\hat{\theta}_1 \quad \boldsymbol{\mathcal{I}}_1$

$\mathbf{K}(\hat{\theta}_1)$

$\mathbf{W}(\hat{\theta}_1)$

Receding horizon design

$\hat{\theta}_1 \quad \boldsymbol{\mathcal{I}}_1$

$\mathbf{K}(\hat{\theta}_1)$

$\mathbf{W}(\hat{\theta}_1)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$

Receding horizon design

$\hat{\theta}_1 \quad \boldsymbol{\mathcal{I}}_1$

$\mathbf{K}(\hat{\theta}_1)$

$\mathbf{W}(\hat{\theta}_1)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$

$\beta_q(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$

Receding horizon design

Interval 1 Interval 2 Interval 3

$\hat{\theta}_1 \quad \mathcal{I}_1$

$\mathbf{K}(\hat{\theta}_1)$

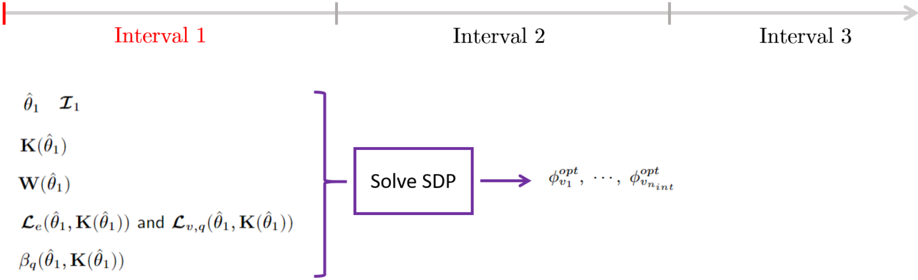
$\mathbf{W}(\hat{\theta}_1)$

$\mathcal{L}_e(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$ and $\mathcal{L}_{v,q}(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$

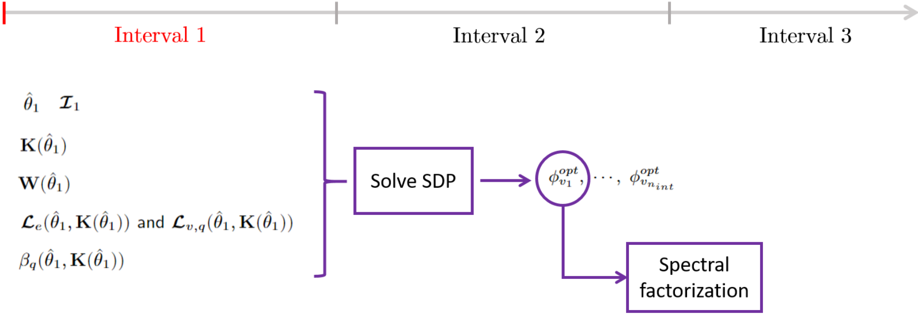
$\beta_q(\hat{\theta}_1, \mathbf{K}(\hat{\theta}_1))$

Solve SDP

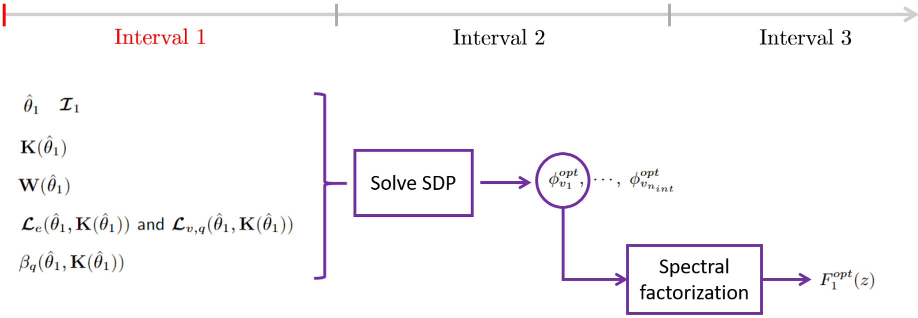
Receding horizon design



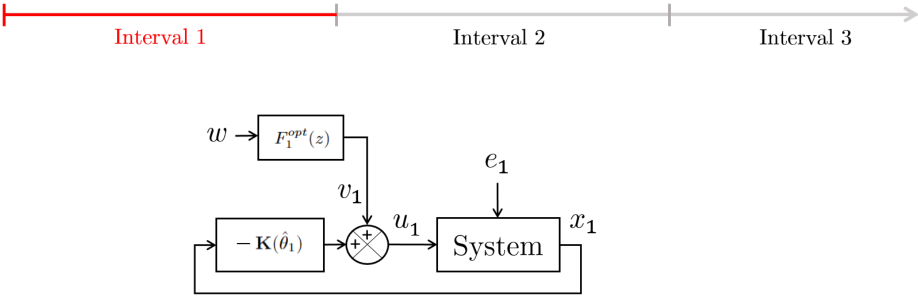
Receding horizon design



Receding horizon design



Receding horizon design



Receding horizon design

$\hat{\theta}_2 \quad \mathcal{I}_2$

Receding horizon design

$\hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

Receding horizon design

$\hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

$\mathbf{W}(\hat{\theta}_2)$

Receding horizon design

$\hat{\theta}_2$ $\boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

$\mathbf{W}(\hat{\theta}_2)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

Receding horizon design

$\hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

$\mathbf{W}(\hat{\theta}_2)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

$\beta_q(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

Receding horizon design

$$\left.\begin{array}{l} \hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2 \\ \mathbf{K}(\hat{\theta}_2) \\ \mathbf{W}(\hat{\theta}_2) \\ \boldsymbol{\mathcal{L}}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2)) \text{ and } \boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2)) \\ \beta_q(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2)) \end{array}\right\} \boxed{\text{Solve SDP}}$$

Receding horizon design

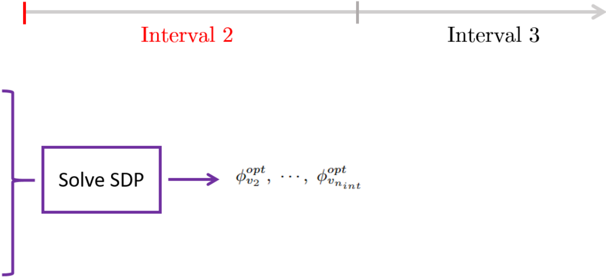
$\hat{\theta}_2 \quad \mathcal{I}_2$

$\mathbf{K}(\hat{\theta}_2)$

$\mathbf{W}(\hat{\theta}_2)$

$\mathcal{L}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$ and $\mathcal{L}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

$\beta_q(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$



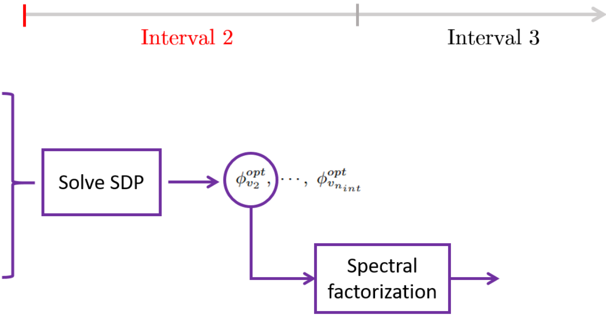
$\hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

$\mathbf{W}(\hat{\theta}_2)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

$\beta_q(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$



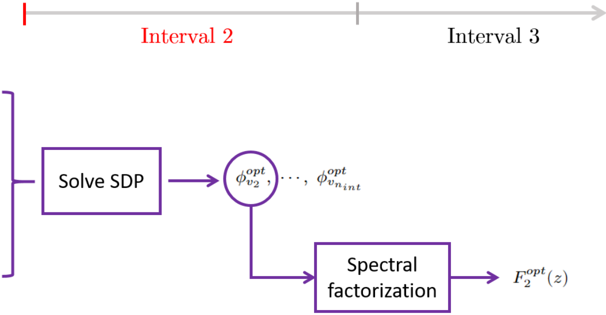
$\hat{\theta}_2 \quad \boldsymbol{\mathcal{I}}_2$

$\mathbf{K}(\hat{\theta}_2)$

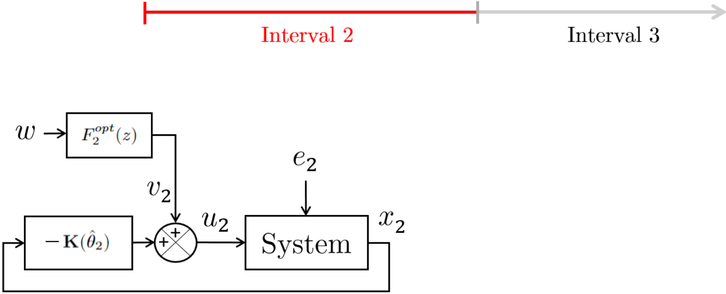
$\mathbf{W}(\hat{\theta}_2)$

$\boldsymbol{\mathcal{L}}_e(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$

$\beta_q(\hat{\theta}_2, \mathbf{K}(\hat{\theta}_2))$



Receding horizon design



A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.
  - solve the SDP.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.
  - solve the SDP.
  - do the spectral factorization of $\phi_{v_k}^{opt}$.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.
  - solve the SDP.
  - do the spectral factorization of $\phi_{v_k}^{opt}$.

- Stability? Since $\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$, we can check[12] if $\mathbf{K}(\hat{\theta}_k)$ stabilizes the system given a probability α.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.
  - solve the SDP.
  - do the spectral factorization of $\phi_{v_k}^{opt}$.

- Stability? Since $\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$, we can check[12] if $\mathbf{K}(\hat{\theta}_k)$ stabilizes the system given a probability α.

- If the stability test fails, we set $\hat{\theta}_k = \hat{\theta}_{k-1}$.

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

A computationally expensive exploration scheme...

- At the beginning of any interval k, we have to
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - compute the Hessian $\mathbf{W}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{W}$ (finite differentiation).
  - compute the exploration coefficients $\beta_q(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \beta_q$.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}(\hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \boldsymbol{\mathcal{L}}_{v,q}$.
  - solve the SDP.
  - do the spectral factorization of $\phi_{v_k}^{opt}$.

- Stability? Since $\hat{\theta}_k \sim N(\theta_0, \boldsymbol{\mathcal{I}}_k^{-1})$, we can check[12] if $\mathbf{K}(\hat{\theta}_k)$ stabilizes the system given a probability α.

- If the stability test fails, we set $\hat{\theta}_k = \hat{\theta}_{k-1}$.

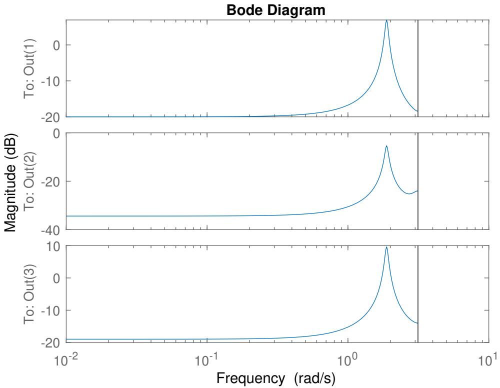
- **Everything has to be done within a duration less than the available sampling time!**

[12] X. Bombois, M. Gevers, G. Scorletti, B. D. O. Anderson (2001). Robustness analysis tools for an uncertainty set obtained by prediction error identification. Automatica, 37(10), 1629–1636

Numerical example

- State system with $n_x = 3$ states with

$$\mathbf{A}_0 = \begin{pmatrix} -0.39 & 0.37 & -0.57 \\ -0.25 & -0.78 & -0.08 \\ 1.32 & 0.25 & -0.13 \end{pmatrix} \qquad \mathbf{B}_0 = \begin{pmatrix} 0.21 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{\Sigma}_e = \mathbf{I}_3$$



Numerical example

- Model structure: each entry of both matrices parameterized independently $\rightarrow$ 12 parameters.

Numerical example

- Model structure: each entry of both matrices parameterized independently $\rightarrow$ 12 parameters.

- Control parameters $\mathbf{Q} = \mathbf{I}_3$ and $\mathbf{R} = 0.2$.

Numerical example

- Model structure: each entry of both matrices parameterized independently $\rightarrow$ 12 parameters.

- Control parameters $\mathbf{Q} = \mathbf{I}_3$ and $\mathbf{R} = 0.2$.

- We choose $T = 500$, $N = 1$ and $n_{int} = 500$.

Numerical example

• Model structure: each entry of both matrices parameterized independently $\rightarrow$ 12 parameters.

• Control parameters $\mathbf{Q} = \mathbf{I}_3$ and $\mathbf{R} = 0.2$.

• We choose $T = 500$, $N = 1$ and $n_{int} = 500$.

• Initial estimate: open-loop identification with 12 data and a white noise input with a variance equal to 1.

Receding horizon approach

m	Computation time
0	50 min
1	2 h 20 min
2	4 h 27 min
3	7 h 49 min
4	12 h 4 min
5	17 h 49 min

Table: Computation for one noise realization from $t = 1$ till $t = 500$ and different FIR order m.

Receding horizon approach

m	Computation time
0	50 min
1	2 h 20 min
2	4 h 27 min
3	7 h 49 min
4	12 h 4 min
5	17 h 49 min

Table: Computation for one noise realization from $t = 1$ till $t = 500$ and different FIR order m.

- Too long...

Receding horizon approach

m	Computation time
0	50 min
1	2 h 20 min
2	4 h 27 min
3	7 h 49 min
4	12 h 4 min
5	17 h 49 min

Table: Computation for one noise realization from $t = 1$ till $t = 500$ and different FIR order m.

- Too long...
- Main reason: the SDP has many variables and constraints, especially at the beginning.

Receding horizon approach

m	Computation time
0	50 min
1	2 h 20 min
2	4 h 27 min
3	7 h 49 min
4	12 h 4 min
5	17 h 49 min

Table: Computation for one noise realization from $t = 1$ till $t = 500$ and different FIR order m.

- Too long...
- Main reason: the SDP has many variables and constraints, especially at the beginning.

Second objective of the talk

Reduction of the computation complexity of the exploration scheme

Exploration coefficients β_q

$$\beta_q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) \cos(q\omega) d\omega$$

where

$$\mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{T}_{xv}^*(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) \mathbf{Q} \mathbf{T}_{xv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) + \mathbf{R} |\mathbf{T}_{uv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k))|^2$$

Exploration coefficients β_q

$$\beta_q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) \cos(q\omega) d\omega$$

where

$$\mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{T}_{xv}^*(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k))\mathbf{Q}\mathbf{T}_{xv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) + \mathbf{R}|\mathbf{T}_{uv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k))|^2$$

Result on LQ control property

$$\mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{R} + \mathbf{B}(\hat{\theta}_k)^\top \mathbf{P}(\hat{\theta}_k)\mathbf{B}(\hat{\theta}_k)$$

where $\mathbf{P}(\hat{\theta}_k)$ is the solution of the discrete-time algebraic Riccati equation involved in the computation of $\mathbf{K}(\hat{\theta}_k)$.

Exploration coefficients β_q

$$\beta_q = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) \cos(q\omega) d\omega$$

where

$$\mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{T}_{xv}^*(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) \mathbf{Q} \mathbf{T}_{xv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) + \mathbf{R} |\mathbf{T}_{uv}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k))|^2$$

Result on LQ control property

$$\mathcal{D}(e^{j\omega}, \hat{\theta}_k, \mathbf{K}(\hat{\theta}_k)) = \mathbf{R} + \mathbf{B}(\hat{\theta}_k)^\top \mathbf{P}(\hat{\theta}_k) \mathbf{B}(\hat{\theta}_k)$$

where $\mathbf{P}(\hat{\theta}_k)$ is the solution of the discrete-time algebraic Riccati equation involved in the computation of $\mathbf{K}(\hat{\theta}_k)$.

- **Consequence**: $\beta_0 = \mathbf{R} + \mathbf{B}(\hat{\theta}_k)^\top \mathbf{P}(\hat{\theta}_k) \mathbf{B}(\hat{\theta}_k)$ and $\beta_q = 0$ for any $q \geq 1$.

- The SDP to be solved at the beginning of any interval k is simplified to:

$$\min_{\substack{\mathbf{Z}(l)=\mathbf{Z}(l)^\top \\ c_q(l),\mathbf{H}(l)}} N \sum_{l=k}^{n_{int}-1} \left(\frac{1}{2}\,\mathrm{tr}(\mathbf{H}(l)) + \beta_0 c_0(l) \right)$$

Subject to , for each $l = k, \cdots, n_{int}-1$,

$$\begin{pmatrix} \mathbf{H}(l) & \mathbf{W}^{1/2} \\ \mathbf{W}^{1/2} & \boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e + \sum_{i=k}^{l}\sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q} \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_{1:m}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_0(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- The SDP to be solved at the beginning of any interval k is simplified to:

$$\min_{\substack{\mathbf{Z}(l)=\mathbf{Z}(l)^\top \\ c_q(l),\mathbf{H}(l)}} N \sum_{l=k}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(l)) + \beta_0 c_0(l) \right)$$

Subject to , for each $l = k, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(l) & \mathbf{W}^{1/2} \\ \mathbf{W}^{1/2} & \boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e + \sum_{i=k}^{l}\sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q} \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_{1:m}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_0(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Still a lot of variables and constraints...

Reformulation of the regret minimization problem

- The SDP to be solved at the beginning of any interval k is simplified to:

$$\min_{\substack{\mathbf{Z}(l)=\mathbf{Z}(l)^\top \\ c_q(l),\mathbf{H}(l)}} N \sum_{l=k}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(l)) + \beta_0 c_0(l) \right)$$

Subject to , for each $l = k, \cdots, n_{int}-1$,

$$\begin{pmatrix} \mathbf{H}(l) & \mathbf{W}^{1/2} \\ \mathbf{W}^{1/2} & \boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e + \sum_{i=k}^{l}\sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q} \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_{1:m}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_0(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Still a lot of variables and constraints... but we can remove a lot of them!

Reformulation of the regret minimization problem

- The SDP to be solved at the beginning of any interval k is simplified to:

$$\min_{\substack{\mathbf{Z}(l)=\mathbf{Z}(l)^\top \\ c_q(l),\mathbf{H}(l)}} N \sum_{l=k}^{n_{int}-1} \left(\frac{1}{2}\operatorname{tr}(\mathbf{H}(l)) + \beta_0 c_0(l) \right)$$

Subject to , for each $l = k, \cdots, n_{int} - 1$,

$$\begin{pmatrix} \mathbf{H}(l) & \mathbf{W}^{1/2} \\ \mathbf{W}^{1/2} & \boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e + \sum_{i=k}^{l} \sum_{q=0}^{m} c_q(i)\boldsymbol{\mathcal{L}}_{v,q} \end{pmatrix} \succeq 0$$

$$\begin{pmatrix} \mathbf{Z}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_{1:m}(l) - \boldsymbol{\mathcal{A}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \\ c_{1:m}^\top(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{A}} & c_0(l) - \boldsymbol{\mathcal{B}}^\top \mathbf{Z}(l)\boldsymbol{\mathcal{B}} \end{pmatrix} \succeq 0$$

- Still a lot of variables and constraints... but we can remove a lot of them!

Theorem

The solution of the SDP guarantee $\phi_{v_{k+1}}^{opt} = \cdots = \phi_{v_{n_{int}-1}}^{opt} = 0$, i.e., the next intervals are not excited!

Proved from dual theory.

Result: theoretical result on the structure of the solution of the SDP

Condition for lazy exploration

The PSD ϕ_{v_k} for the current interval k satisfies $\phi_{v_k} = 0$ if and only if

$$\beta_0 \geq \frac{N}{2} \lambda_{\mathsf{max}} \left(\sum_{l=k}^{n_{int}-1} \mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k)) \right)$$

where $\mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k))$ is the Toeplitz symmetric matrix whose first row is $(\zeta_0(k), \cdots, \zeta_m(k))$ with

$$\zeta_q(k) = \mathrm{tr}(\mathbf{W}(\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1}\boldsymbol{\mathcal{L}}_{v,q}(\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1})$$

Result: theoretical result on the structure of the solution of the SDP

Condition for lazy exploration

The PSD ϕ_{v_k} for the current interval k satisfies $\phi_{v_k} = 0$ if and only if

$$\beta_0 \geq \frac{N}{2} \lambda_{\max} \left(\sum_{l=k}^{n_{int}-1} \mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k)) \right)$$

where $\mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k))$ is the Toeplitz symmetric matrix whose first row is $(\zeta_0(k), \cdots, \zeta_m(k))$ with

$$\zeta_q(k) = \mathrm{tr}(\mathbf{W}(\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1} \boldsymbol{\mathcal{L}}_{v,q} (\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1})$$

- Advantage: checking this inequality is faster than solving the SDP.

Result: theoretical result on the structure of the solution of the SDP

Condition for lazy exploration

The PSD ϕ_{v_k} for the current interval k satisfies $\phi_{v_k} = 0$ if and only if

$$\beta_0 \geq \frac{N}{2}\lambda_{\text{max}}\left(\sum_{l=k}^{n_{int}-1} \mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k))\right)$$

where $\mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k))$ is the Toeplitz symmetric matrix whose first row is $(\zeta_0(k), \cdots, \zeta_m(k))$ with

$$\zeta_q(k) = \text{tr}(\mathbf{W}(\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1}\boldsymbol{\mathcal{L}}_{v,q}(\boldsymbol{\mathcal{I}}_k + (l-k+1)\boldsymbol{\mathcal{L}}_e)^{-1})$$

- Advantage: checking this inequality is faster than solving the SDP.

- Interpretation: we don't need to excite at all if
  - the information $\boldsymbol{\mathcal{I}}_k + (l-k)\boldsymbol{\mathcal{L}}_e$ obtained without excitation is large.
  - the main eigenvectors of its inverse are perpendicular to the ones of $\mathbf{W}$.
  - the exploration penalization coefficient β_0 is large.
  - the number $n_{int} - k + 1$ of remaining intervals is small.

Outline of the algorithm

- At the beginning of any interval k, we

Outline of the algorithm

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.

Outline of the algorithm

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - perform the stability test and set $\hat{\theta}_k = \hat{\theta}_{k-1}$ if it fails.

Outline of the algorithm

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - perform the stability test and set $\hat{\theta}_k = \hat{\theta}_{k-1}$ if it fails.
  - compute the Hessian $\mathbf{W}$.

Outline of the algorithm

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - perform the stability test and set $\hat{\theta}_k = \hat{\theta}_{k-1}$ if it fails.
  - compute the Hessian $\mathbf{W}$.
  - compute the exploration coefficient β_0.

Outline of the algorithm

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - perform the stability test and set $\hat{\theta}_k = \hat{\theta}_{k-1}$ if it fails.
  - compute the Hessian $\mathbf{W}$.
  - compute the exploration coefficient β_0.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}$.

- At the beginning of any interval k, we
  - identify $\hat{\theta}_k$ and the controller $\mathbf{K}(\hat{\theta}_k)$.
  - perform the stability test and set $\hat{\theta}_k = \hat{\theta}_{k-1}$ if it fails.
  - compute the Hessian $\mathbf{W}$.
  - compute the exploration coefficient β_0.
  - compute the Fisher matrices $\boldsymbol{\mathcal{L}}_e$ and $\boldsymbol{\mathcal{L}}_{v,q}$.
  - check if

$$\beta_0 \geq \overline{\lambda}\left(\sum_{l=k}^{n_{int}} \mathcal{T}(\zeta_0(k), \cdots, \zeta_m(k))\right)$$

If it is true, set $v_k = 0$. Otherwise solve the SDP by removing all the variables and constraints of the PSD $\phi_{v_{k+1}}, \ldots, \phi_{v_{n_{int}}}$ and then use spectral factorization on $\phi_{v_k}^{opt}$.

Numerical result

- Let us now consider back the numerical example combined with the three results ($n_x = 3$, $n_\theta = 12$, $N = 1$ and $n_{int} = 500$).

FIR order m	Before	After
0	50 min	19.9 s
1	2 h 20 min	20.7 s
2	4 h 27 min	28.2 s
3	7 h 49 min	29.3 s
4	12 h 4 min	35.7 s
5	17 h 49 min	39.6 s

Table: Computation times obtained with the naive approach with the receding horizon strategy for one noise realization from $t = 1$ till $t = 500$ and different FIR order m.

Numerical example

- Change of settings: $T = 100000$, $N = 1000$, $n_{int} = 100$.

- Initial open-loop identification experiment with 200 data and a white Gaussian noise excitation of variance 0.1.

- Comparison of the proposed scheme with the Thompson sampling approach[13] and the certainity equivalence approach[14] with

$$v(t) \sim N\left(0, \frac{a}{\sqrt{t}}\right)$$

- 100 Monte Carlo simulations with different realizations for e and w (same for initial estimate) in order to approximate the expectation operator in the regret expression.

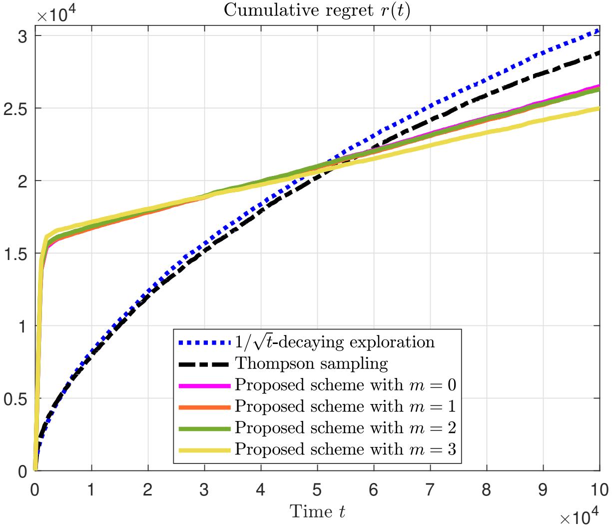
- a tuned by using a gridding approach so that $r(T)$ is minimized.

[13] Y. Ouyang, M. Gagrani, and R. Jain. Control of unknown linear systems with Thompson sampling. In 2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 1198–1205. IEEE, 201

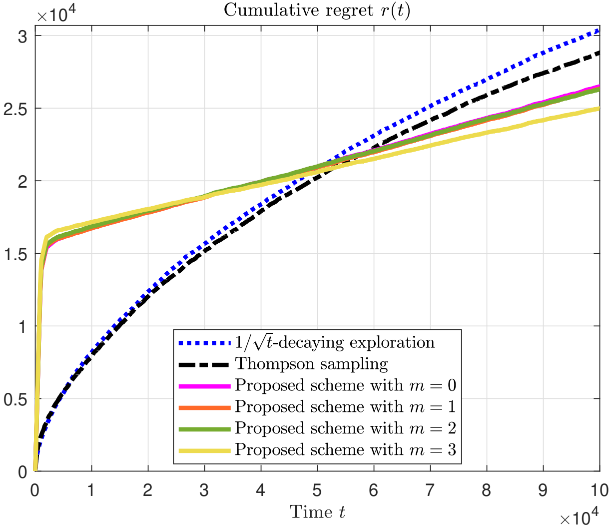
[14] F. Wang and L. Janson. Exact asymptotics for linear quadratic adaptive control. Journal of Machine Learning Research, 22(265):1–112, 2021

Method	$r(T)$
$1/\sqrt{t}$-decaying exploration	30393
Thompson sampling	28823
Proposed scheme with $m = 0$	26490
Proposed scheme with $m = 1$	26333
Proposed scheme with $m = 2$	26290
Proposed scheme with $m = 3$	24973

Numerical example

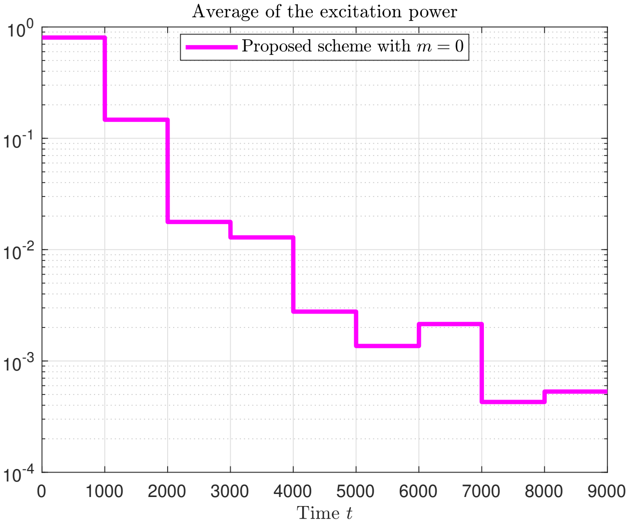


Numerical example

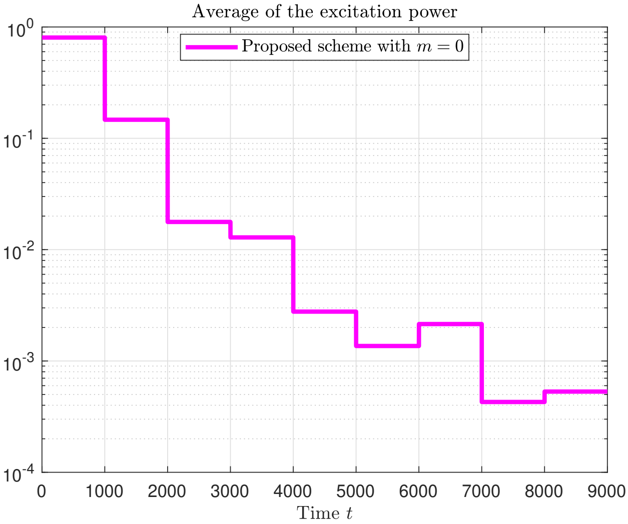


- Even though AF3E beats all the other methods, it performs the worst during the first half of the experiment (due to jumping phenomenon at the beginning)...

Numerical example

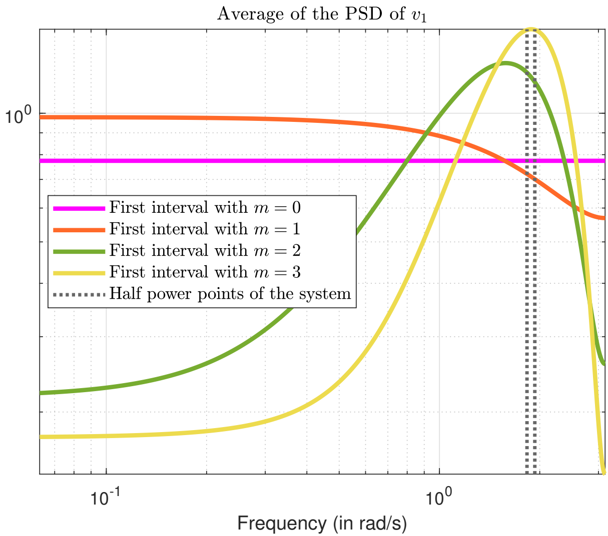


Numerical example

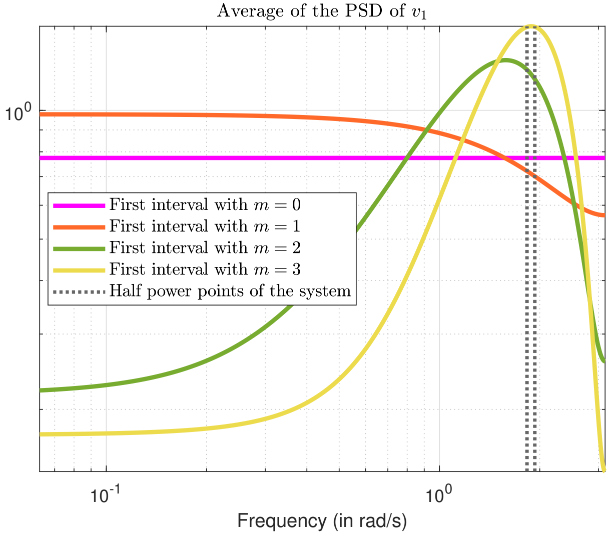


- After $t \geq 9000$, the proposed scheme stopped exciting $\rightarrow$ all the exploration effort is done at the beginning.

Numerical example



Numerical example



- With $m = 3$, we excite well the resonance of the system.

Conclusion and perspectives

- Formulation of the regret minimization problem in the linear quadratic adaptive framework as a linear time invariant experiment design problem by intervals.

[15] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Conclusion and perspectives

• Formulation of the regret minimization problem in the linear quadratic adaptive framework as a linear time invariant experiment design problem by intervals.

• Reduction of the computation time of the proposed scheme.

[15] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Conclusion and perspectives

• Formulation of the regret minimization problem in the linear quadratic adaptive framework as a linear time invariant experiment design problem by intervals.

• Reduction of the computation time of the proposed scheme.

• Numerical example: the proposed scheme can beat other exploration strategies of the literature.

[15] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Conclusion and perspectives

• Formulation of the regret minimization problem in the linear quadratic adaptive framework as a linear time invariant experiment design problem by intervals.

• Reduction of the computation time of the proposed scheme.

• Numerical example: the proposed scheme can beat other exploration strategies of the literature.

• A drawback: the aggressive exploration strategy for the first intervals can saturate the output and/or the input.

[15] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Conclusion and perspectives

• Formulation of the regret minimization problem in the linear quadratic adaptive framework as a linear time invariant experiment design problem by intervals.

• Reduction of the computation time of the proposed scheme.

• Numerical example: the proposed scheme can beat other exploration strategies of the literature.

• A drawback: the aggressive exploration strategy for the first intervals can saturate the output and/or the input.

• Some possible solutions to be investigated:

- add input/output constraints which can be transformed into a LMI.
- consider a worst-case design[15] of the external excitation.

---

[15] M. Forgione, X. Bombois, and P. M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124 (2015).

Thank you for you attention!