Extremum seeking control:
convergence analysis

Dragan Nešić
The University of Melbourne

Acknowledgements:
Y. Tan, I. Mareels, A. Astolfi, G. Bastin, C. Manzie; Australian Research Council.

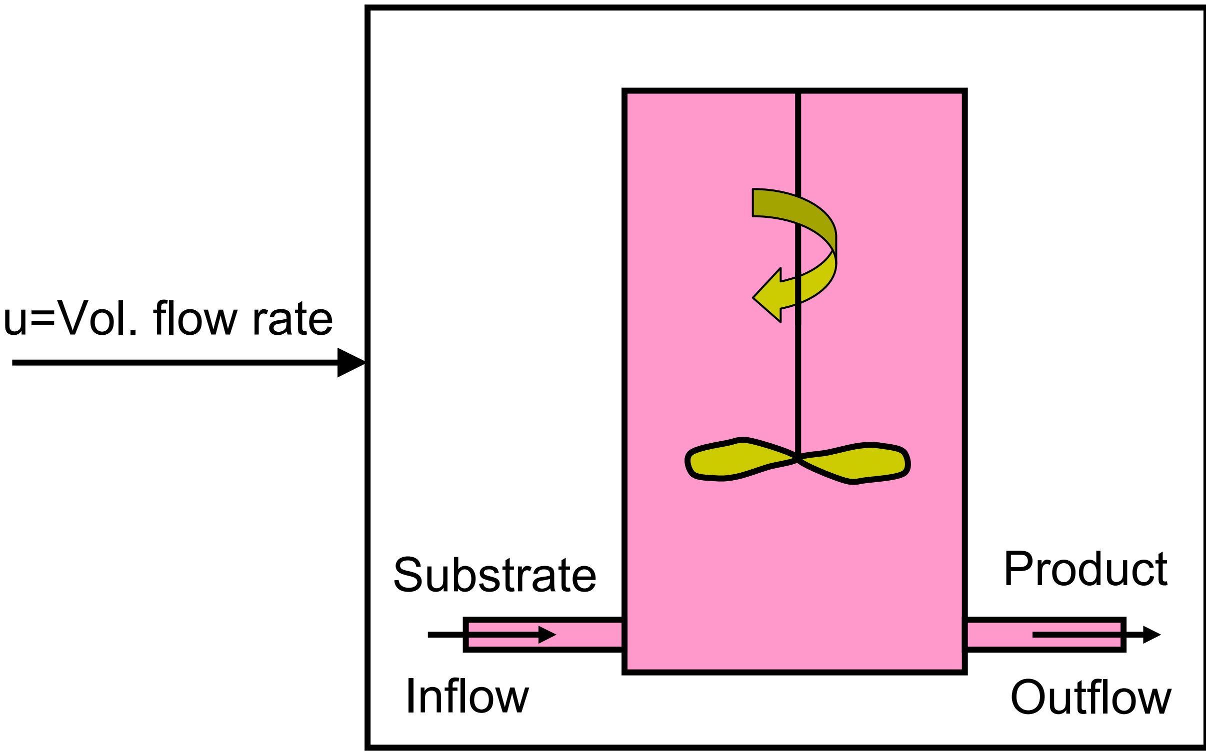
Outline

- Motivating examples
- Problem formulation
- Background
- Non-local stability:
  - No local extrema
  - With local extrema
- Some open problems.
- Conclusions

Motivating example

bioreactors

Continuously Stirred Tank (CST) Reactor



Performance output y:

Productivity J_P

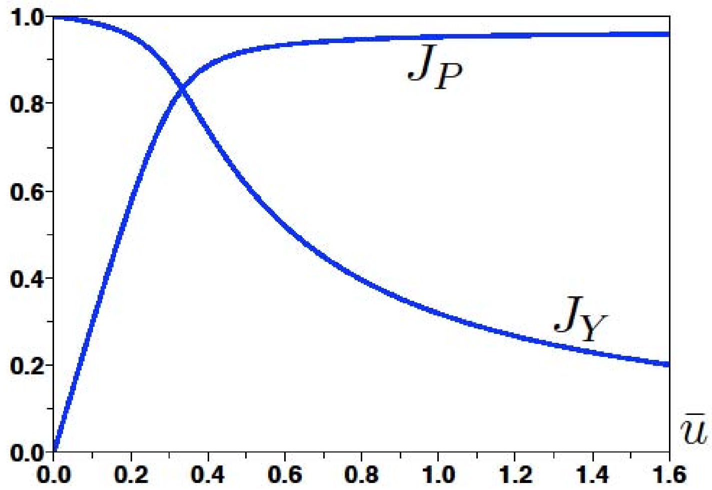
Yield J_Y

Overall J_T

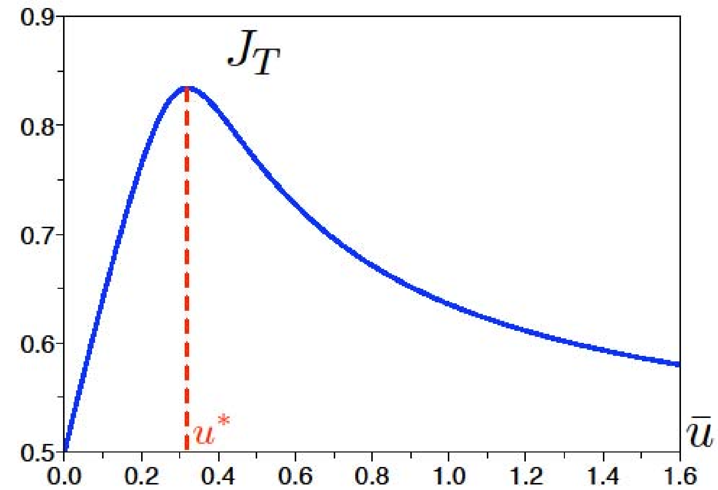
$J_T := \lambda J_P + (1 - \lambda) J_Y, \quad \lambda \in (0,1)$

Assumption: $u(t) \approx \bar{u} \Rightarrow J_\star(t) \to J_\star(\bar{u})$

Single enzymatic reaction
Michaelis-Menten Kinetics



Productivity and yield



Total cost

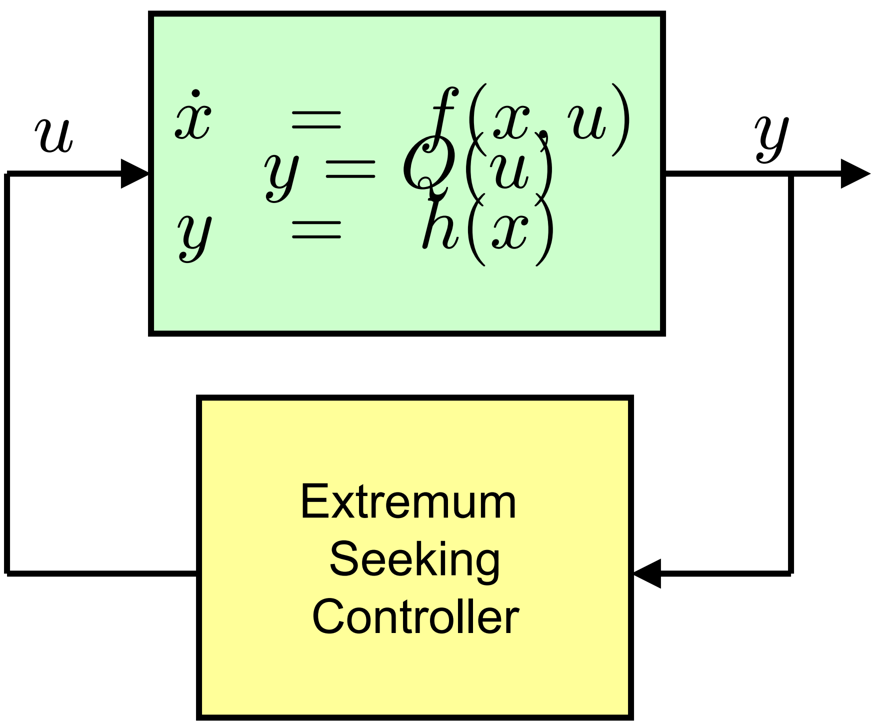
In steady-state, we would typically want to operate around u^*

$J_T(\bar{u})$ is typically unknown!!

Other examples

Plant	Performance output
Turbine	Generated power
Solar cell	Generated power
Optical amplifiers	Uniformity of the gain spectrum
Tokamak	Reflected power during Lower Hybrid (LH) plasma heating experiments
Non-holonomic vehicles	Distance from a source of a signal
Paper machine	Retention of fines and fibers in the sheet
Ultrasonic/Sonic Driller/Corer	Distance from resonance
Human Exercise Machine	The user’s power output
ABS	Magnitude of friction force
Variable cam timing	Fuel consumption

Problem formulation



Assumption 1:

- Q(.) has an extremum (max)

$$y^* := Q(u^*) \geq Q(u), \ \forall u$$

- Q(.) is unknown

Dynamic case:

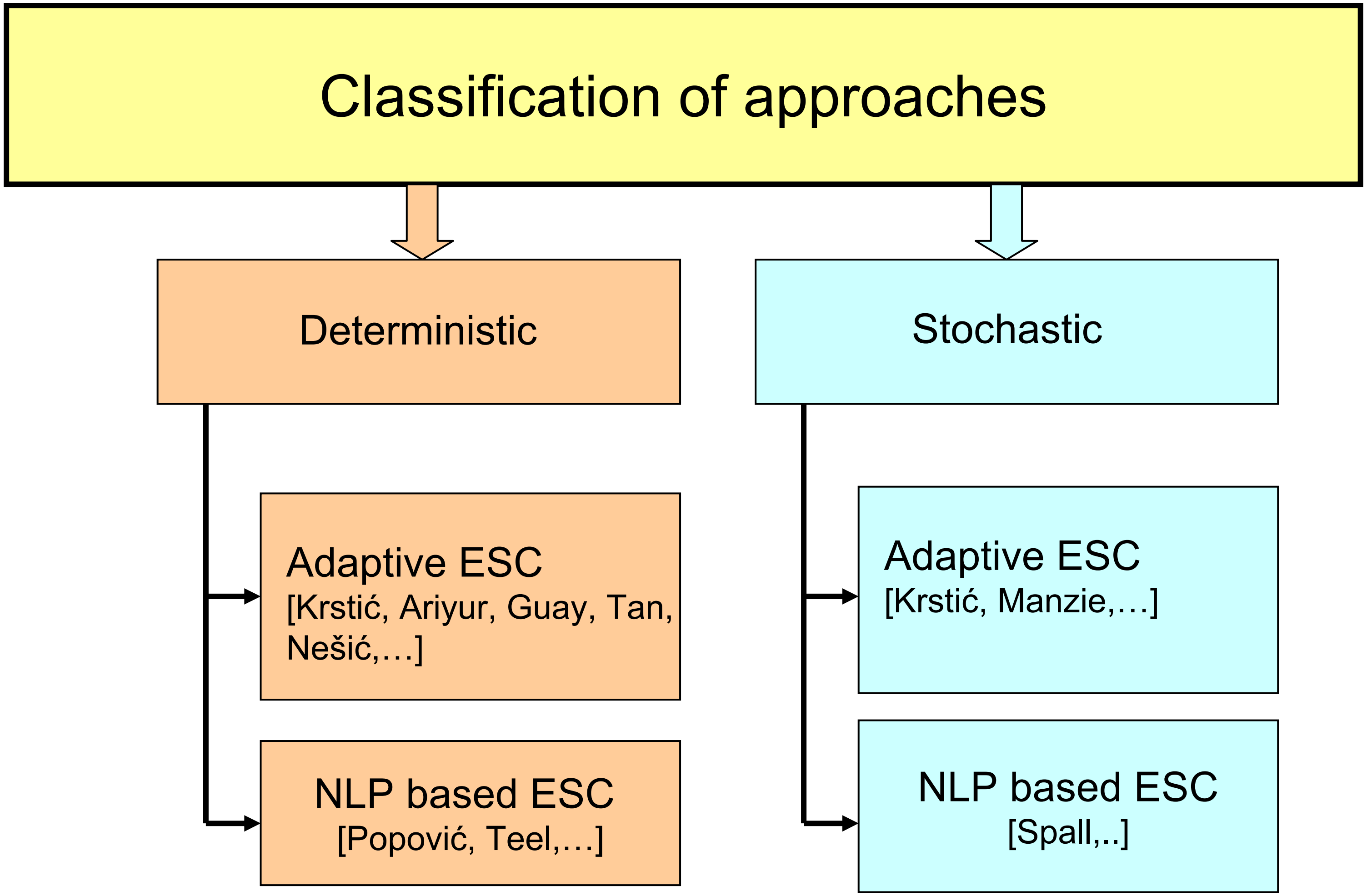
$$\exists \ell(\cdot) \quad 0 = f(\ell(u), u)$$
$$Q(u) := h \circ \ell(u)$$

Problem:

Design ESC so that

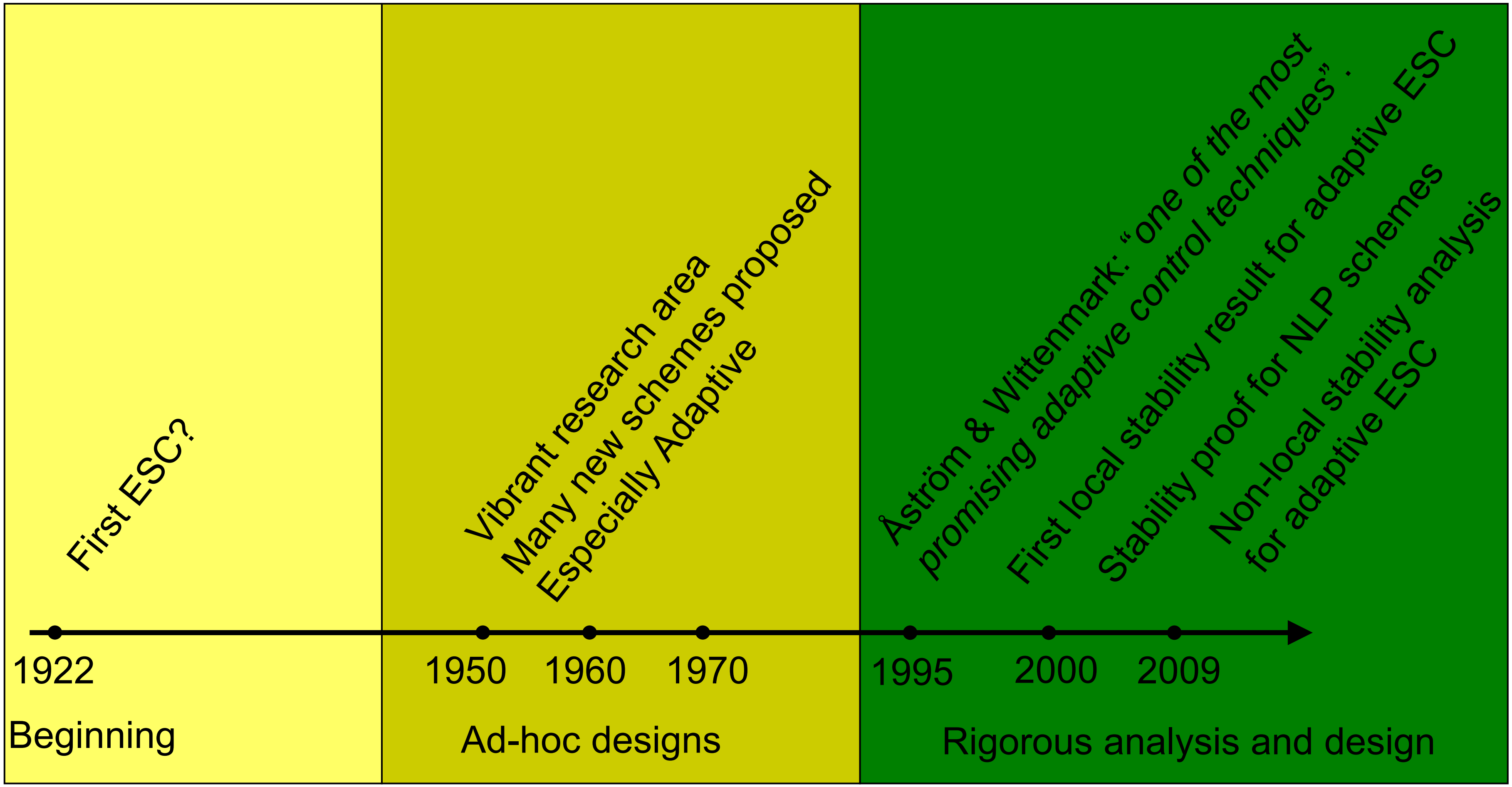
$$\limsup_{t\to\infty} |y(t) - y^*| \approx 0$$

Background



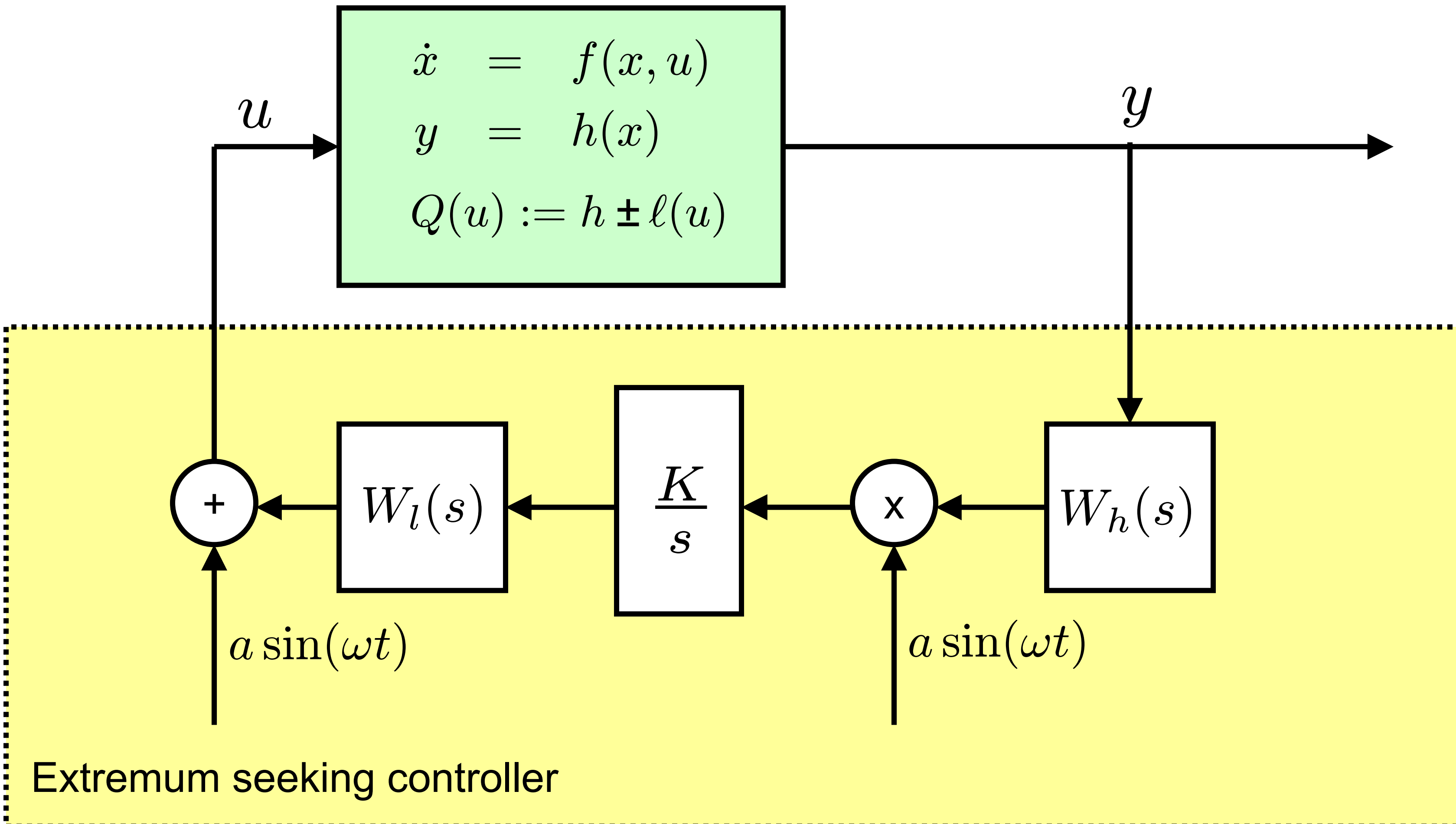
Also continuous-time versus discrete-time.

Brief history (deterministic):



Adaptive ESC

[Krstić & Wang 2000], local stability



Our goals:

Precise non-local convergence analysis.

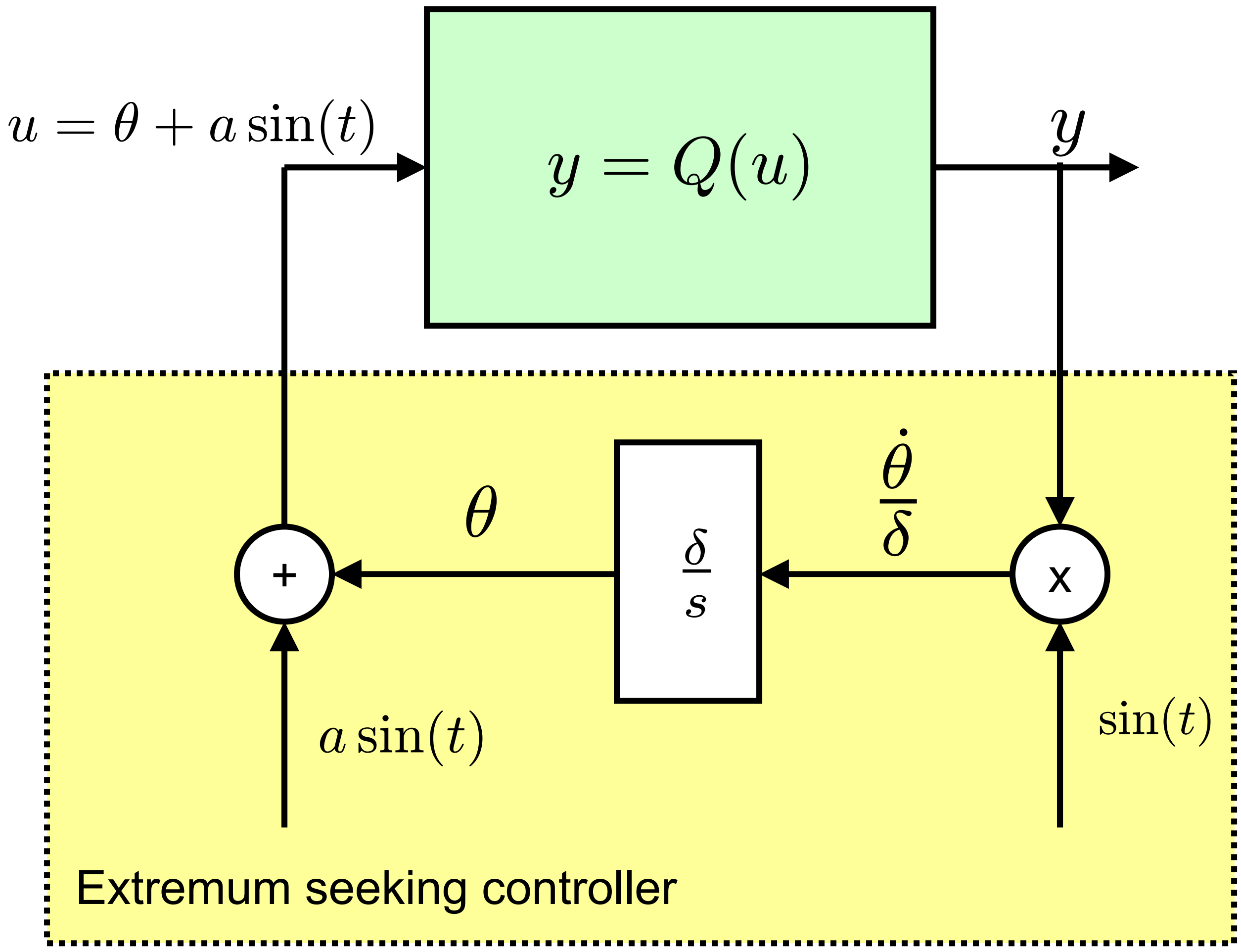
Controller tuning guidelines and trade-offs.

Non-local stability
(no local extrema)

Y. Tan, D. Nešić and I. Mareels, “On non-local stability properties of extremum seeking control”, Automatica, Vol. 42, No. 6, pp. 889-903, 2006.

Static SISO case

(gradient descent)



Parameters:

$a, \ \delta$

Notation:

$$D^k Q := \frac{d^k Q}{du^k}$$

Average system

- The system is periodic in time:

$$\dot{\theta} = \delta Q(\theta + a\sin(t))\sin(t) =: \delta f(t, \theta, a)$$

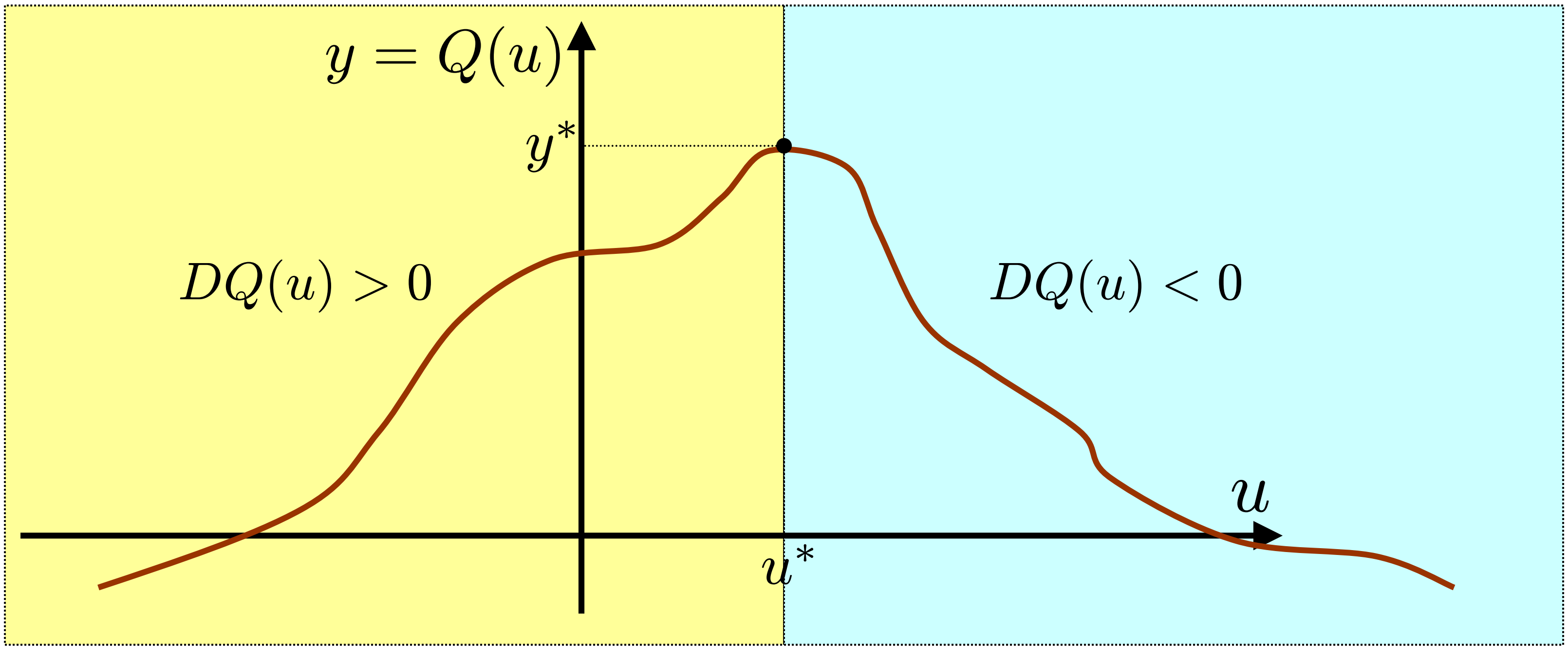
- Its average is a gradient descent scheme:

$$\dot{\theta} = \delta f_{av}(\theta, a) = \delta[\underbrace{\frac{a}{2} DQ(\theta)}_{\text{Gradient descent}} + O(a^3)]$$

$$f_{av}(\theta, a) \quad := \quad \frac{1}{2\pi}\int_0^{2\pi} f(\tau, \theta, a)$$

Assumption 2:

$$DQ(u)(u - u^*) < 0, \quad \forall u \neq u^*$$



This assumption holds for many plants, e.g. some models of CST reactor.

KL functions

- Linear UGES systems satisfy the bound

$$|x(t)| \leq K\exp(-\lambda(t-t_0))|x_0|, \ \forall t \geq t_0, \forall x_0$$

for some K, λ>0.

- Nonlinear UGAS systems satisfy

$$|x(t)| \leq \beta(|x_0|, t-t_0), \ \forall t \geq t_0, \forall x_0$$

for some $\beta \in$ KL.

Theorem:

Suppose Assumptions 1 and 2 hold. Then, there exists $\beta \in$ KL such that:

$$\forall(\Delta, \nu)\exists(\delta^*, a^*)$$

$$\Downarrow$$

$$\forall\delta \in (0, \delta^*), a \in (0, a^*)$$

Tuning guidelines

$$\Downarrow$$

$$|\theta(t_0) - \theta^*| \leq \Delta$$

$$\Downarrow$$

$$|\theta(t) - \theta^*| \leq \beta(|\theta(t_0) - \theta^*|, a\delta(t - t_0)) + \nu, \quad \forall t \geq t_0$$

where $\theta^* := u^*.$

We say that the system in SPA stable in a, δ.

A trade-off

Larger Δ
or
Smaller ν

)

Smaller a
and
Smaller δ

)

Slower
Convergence

Sketch of proof:

- Use the Lyapunov function candidate

$$
\begin{aligned}
V(\theta) &= \frac{1}{2}(\theta - \theta^*)^2 \\
DV(\theta)\delta f_{av}(\theta, a) &= \delta \left[\frac{a}{2} \underbrace{DQ(\theta)(\theta - \theta^*)}_{<0} + O(a^3) \right]
\end{aligned}
$$

- Average system is SPA stable in a.
- Actual system is SPA stable in a, δ.

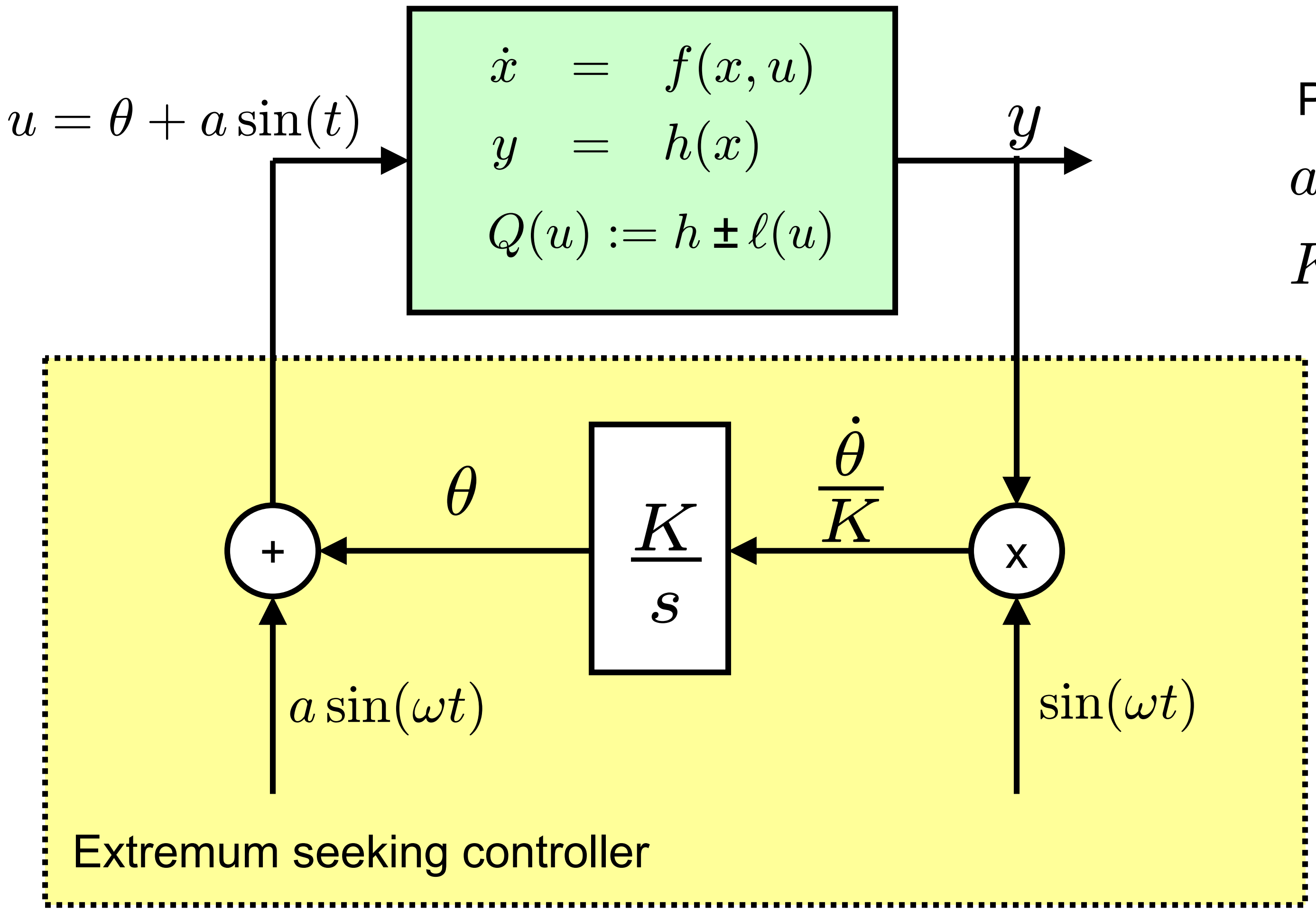
Comments

- Theorem provides a tuning rule for ESC.
- The trade-off limits the rate of convergence!
- ES with filters can be treated similarly.
- Stronger result possible:

  the rate of convergence is proportional the power of dither signal – square wave best.

Y. Tan, D. Nešić and I. Mareels, "On the choice of dither signals in extremum seeking control scheme", Automatica, Vol. 44, No. 5, pp. 1446-1450, 2008.

Dynamic SISO case



Parameters:

$a,\ \omega,\ \delta,$

$K := \omega ¢\delta$

Singularly perturbed model:

- New time scale $\sigma = \omega$ t:

$$\begin{aligned} \omega \frac{dx}{d\sigma} &= f(x, \theta + a\sin(\sigma)) \\ \frac{d\theta}{d\sigma} &= \delta h(x)\sin(\sigma) \end{aligned}$$

- The model is in standard form.
- Time scale separation: slow & fast systems.

Slow model

- Set ω=0

$$0 = f(x, \theta + a\sin(\sigma)) \quad) \quad x = \ell(\theta + a\sin(\sigma))$$

- Substitution in θ equation yields:

$$\tfrac{d\theta}{d\sigma} = \delta h \pm \ell(\theta + a\sin(\sigma))\sin(\sigma) = \delta Q(\theta + a\sin(\sigma))\sin(\sigma)$$

- This is the same system as in static case!
- We use Assumptions 1 and 2.

Fast model

- In the fast time scale:

$$\dot{x} = f(x, \underbrace{\theta_0 + a\sin(\sigma_0)}_{u_0})$$

Assumption 3:

For any u_0 the equilibrium

$$x = \ell(u_0)$$

of the fast system is UGAS, uniformly in u_0.

Theorem

- Suppose Assumptions 1-3 hold. Then, there exist $\beta_1, \beta_2 \in$ KL such that

$$8(\Delta, \nu),\ 9(\delta^*, a^*)$$

$$+$$

$$8\delta\ 2\ (0, \delta^*),\ a\ 2\ (0, a^*),\ 9\omega$$

Tuning guidelines

$$+$$

$$\mathrm{j}(x(t_0), \theta(t_0)\ ¡\ \theta^*)\mathrm{j}\ \cdot\ \Delta$$

$$+$$

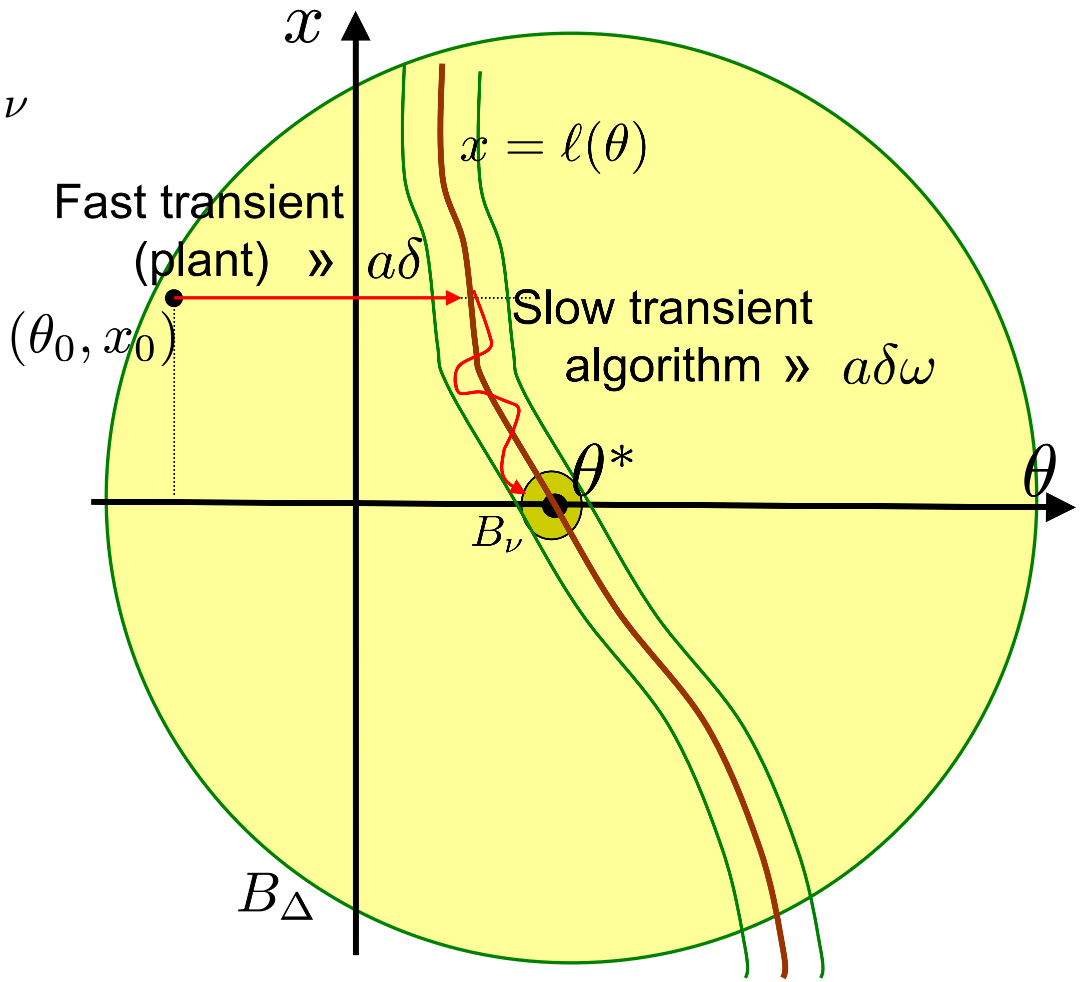
$$\mathrm{j}x(t)\ ¡\ \ell(\theta(t))\mathrm{j}\ \cdot\ \beta_1(\mathrm{j}x(t_0)\ ¡\ \ell(\theta(t_0))\mathrm{j}, a\delta(t\ ¡\ t_0)) + \nu,\ 8t\ ¸\ t_0$$

$$\mathrm{j}\theta(t)\ ¡\ \theta^*\mathrm{j}\ \cdot\ \beta_2(\mathrm{j}\theta(t_0)\ ¡\ \theta^*\mathrm{j}, a\delta\omega(t\ ¡\ t_0)) + \nu,\ 8t\ ¸\ t_0$$

Geometrical interpretation

For any Δ, ν

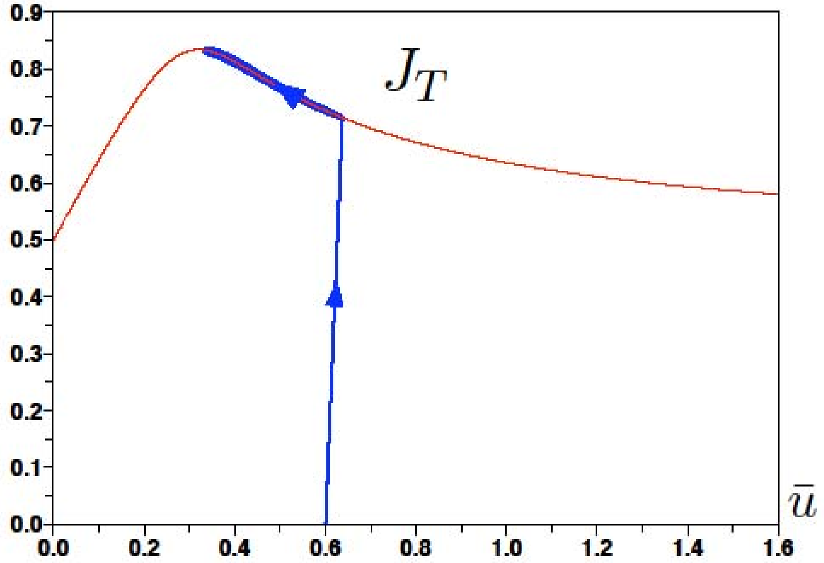
Exist δ, a



$$\limsup_{t\to\infty} |\theta(t) - \theta^*| \leq \nu$$

Bioreactor example

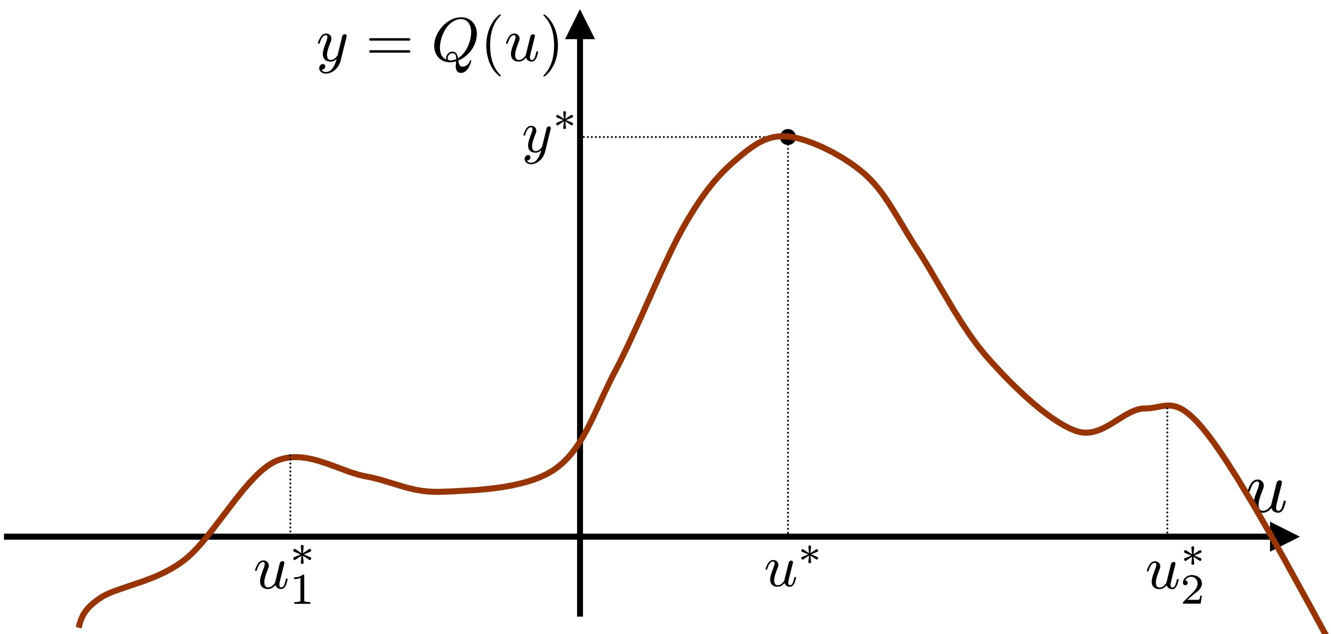
All our assumptions hold.



Non-local stability
(with local extrema)

Y. Tan, D. Nešić and I. Mareels and A. Astolfi, “On the global extremum seeking control”, Automatica, Vol. 45, No. 1, pp. 245-251, 2009.

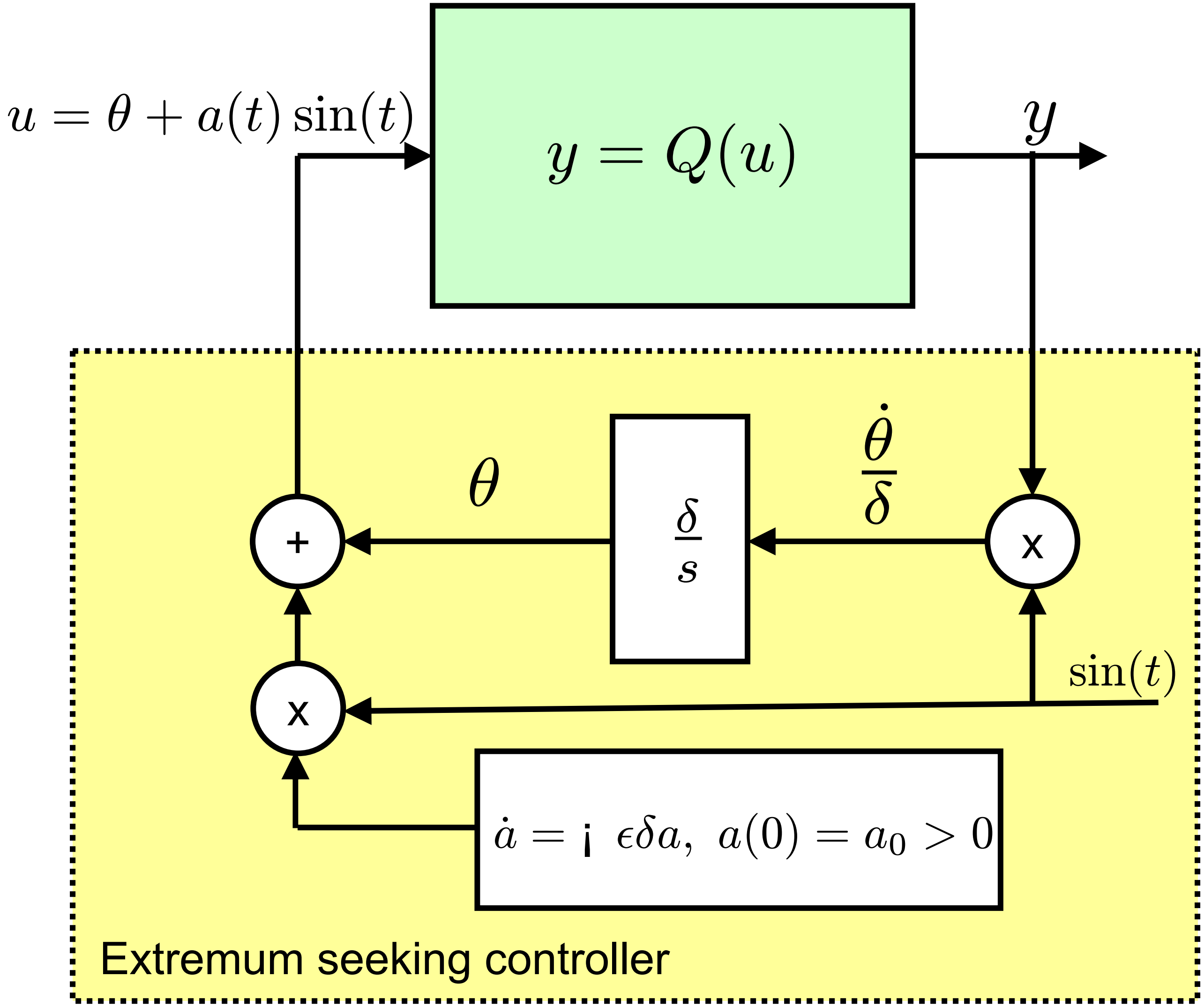
Assumption 2 does not hold!



Assumption 4: There exists a unique global maximum:

$$\exists! u^* \Rightarrow \quad Q(u^*) > Q(u), \forall u \neq u^*.$$

Static SISO case



Parameters:

a_0, δ, ϵ

Model of the system

- The system is time-varying:

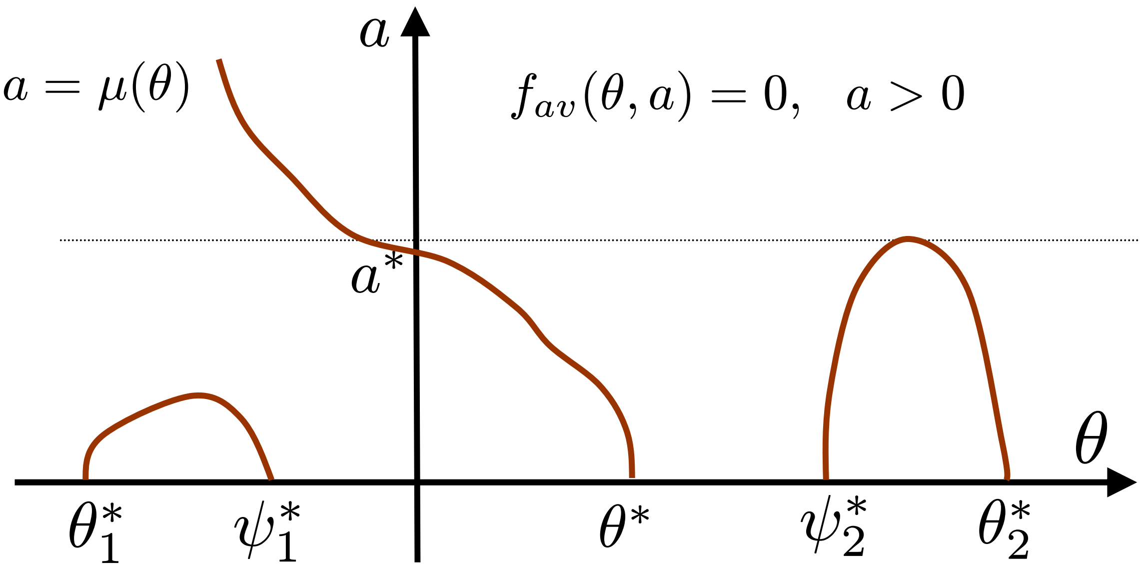
$$\begin{aligned} \dot{\theta} &= \delta Q(\theta + a\sin(t))\sin(t) =: \delta f(t,\theta,a) \\ \dot{a} &= -\epsilon\delta a, \quad a(0) = a_0 \end{aligned}$$

and its average with a change of time σ=t/ϵ

$$\begin{aligned} \epsilon\frac{d\theta}{d\sigma} &= \delta f_{av}(\theta,a) \\ \frac{da}{d\sigma} &= -\delta a, \quad a(0) = a_0 \end{aligned}$$

is a singularly perturbed system.

Desired bifurcation diagram



Global maximum θ^*

Local maxima $\theta_1^*, \theta_2^*, \ldots$

Local minima $\psi_1^*, \psi_2^*, \ldots$

Assumption 5:

The average system $f_{av}(\theta, a)$ has a desired bifurcation diagram.

Comments

- All 4^{th} order polynomials that satisfy Assumption 4 also satisfy Assumption 5.
- There exists a 6^{th} order polynomial that satisfies Assumption 4 but does not satisfy Assumption 5.
- Dither shape affects Assumption 5!

Theorem

- Suppose Assumptions 4 and 5 hold. Then

$$8(\Delta, \nu), a_0 > a^*$$

+

$$9\epsilon^* > 0, 8\epsilon\ 2\ (0, \epsilon^*)$$

+

$$9\delta^* > 0, 8\delta\ 2\ (0, \delta^*)$$

Tuning guidelines

+

$$\text{j}\theta_0\ \text{¡}\ \theta^*\text{j} \cdot \Delta$$

+

$$\text{j}\theta(t)\ \text{¡}\ \mu(a(t))\text{j} \cdot \beta(\text{j}\theta_0\ \text{¡}\ \mu(a_0)\text{j}, \delta(t\ \text{¡}\ t_0)) + \nu$$

$$\text{j}a(t)\text{j} \cdot \exp(\text{¡}\ \epsilon\delta(t\ \text{¡}\ t_0))\text{j}a_0\text{j}$$

Comments

- Note that

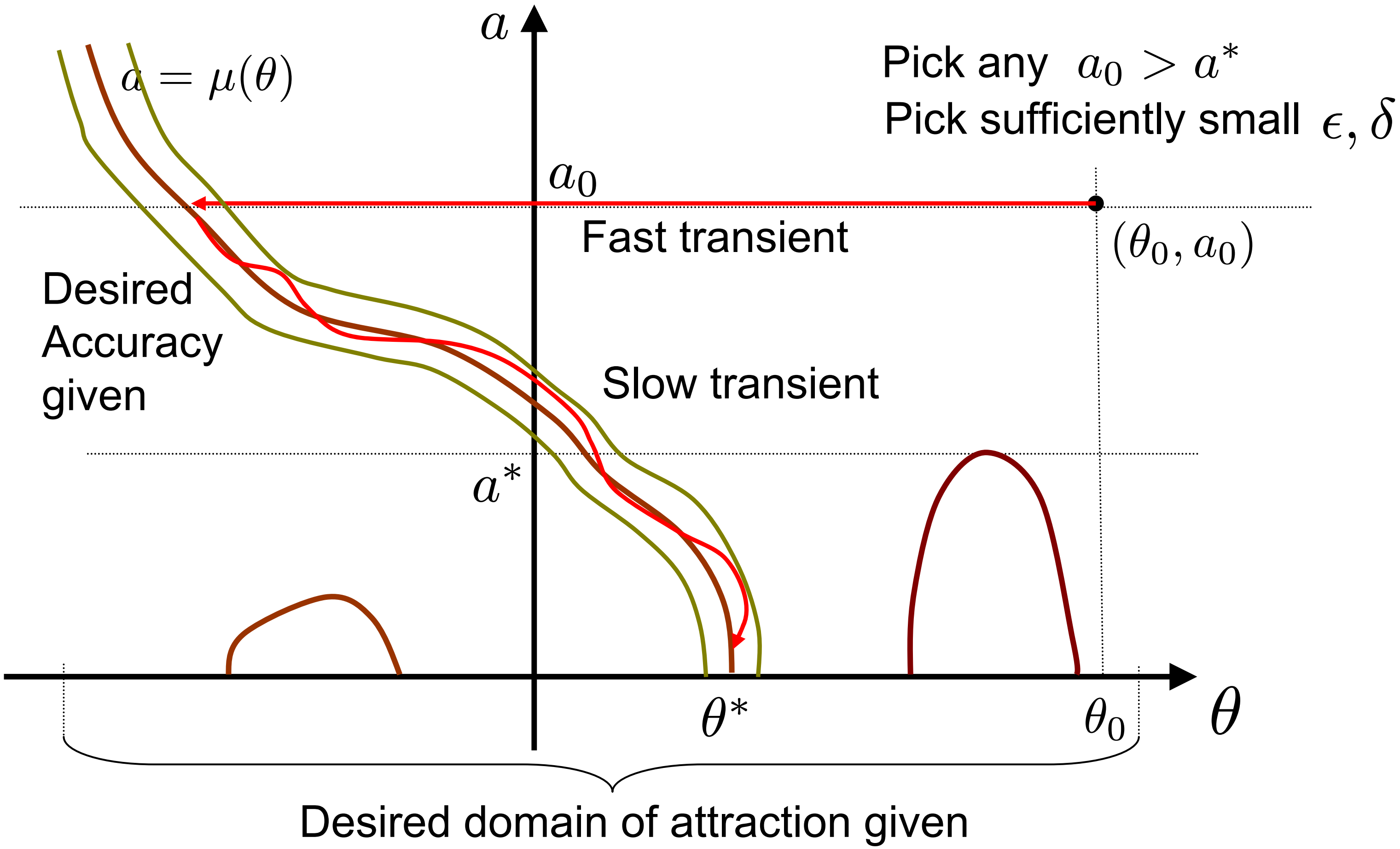
$$a(t) \; ! \quad 0 \;) \quad \lim_{t\to\infty} \mu(a(t)) = \mu(0) = \theta^*$$

- To achieve robustness, we would typically modify ESC so that

$$\lim_{t\to\infty} a(t) = \bar{a} > 0$$

- Similar to “simulated annealing”.

Idea



a
$a = \mu(\theta)$
Pick any $a_0 > a^*$
Pick sufficiently small ϵ, δ
a_0
Fast transient
(θ_0, a_0)
Desired Accuracy given
Slow transient
a^*
θ^*
θ_0
θ
Desired domain of attraction given

Comments

- Assumptions are impossible to verify a priori.
- Our result provides a tuning strategy for ESC that can improve performance.

Some open problems

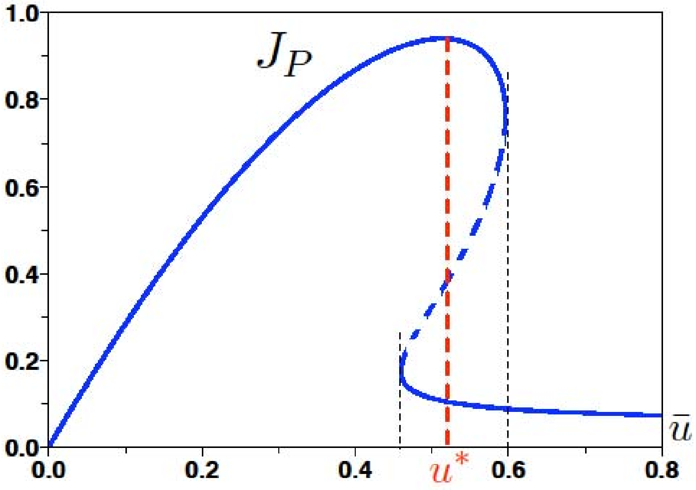
- Convergence rate improvements.
- Using the model knowledge in the best way.
- Adaptive versions of non-gradient schemes.
- Selection of efficient algorithms and dithers for particular applications.
- More detailed tuning guidelines, and so on.
- Multi-valued functions.

Multi-valued functions

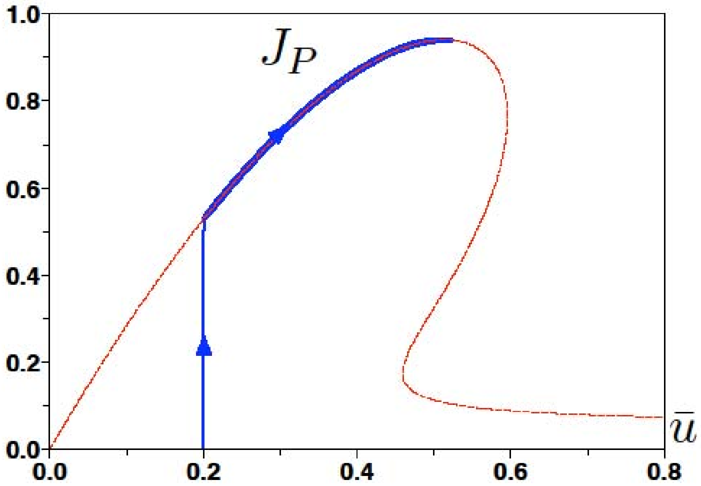
G. Bastin, D. Nešić, Y. Tan and I. Mareels, “On Extremum Seeking in Bioprocesses with Multi-valued Cost Functions”, Biotechnology Progress, Vol. 25, No. 3, pp. 683-689, 2009.

Multi-valued cost

- Our assumptions sometimes do not hold.

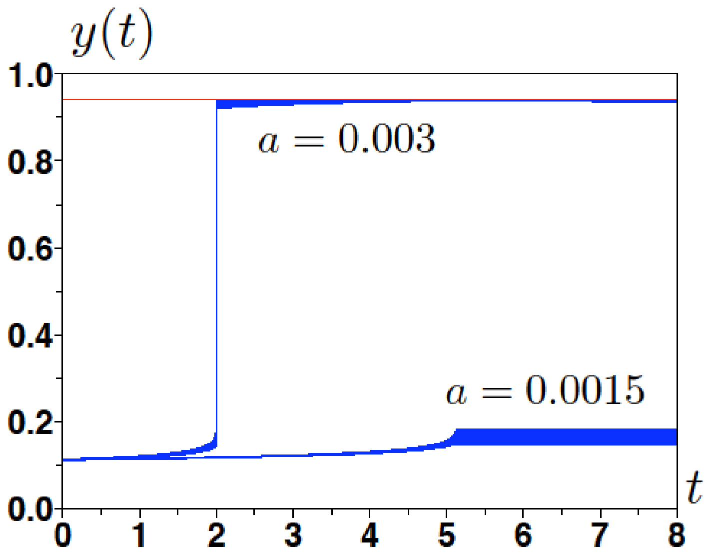


J_P is a multi-valued function

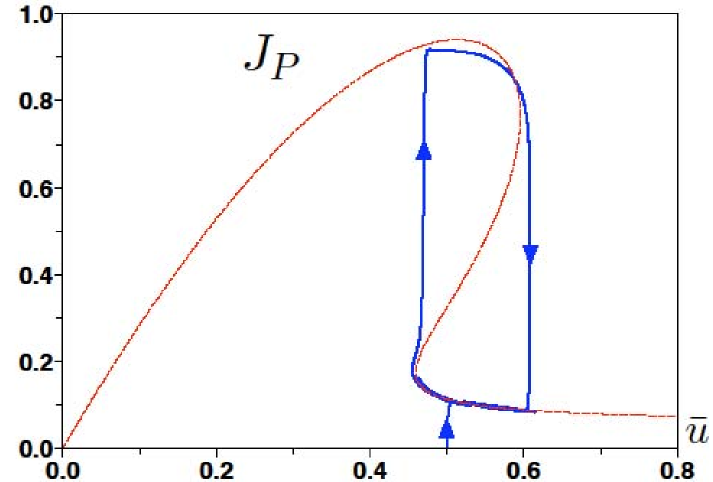


For some initial conditions our analysis is fine

Possible situations



Effects of “small” amplitude



Amplitude “too large”

Conclusions

- Non-local convergence analysis of a class of adaptive ES controllers is presented.
- Tuning guidelines follow from our results.
- Interesting trade-offs arise.
- Global ES possible with local extrema.
- Many open problems.