A hybrid approach to control: classical control theory meets data-driven methods

Mouhacine Benosman

Dynamical Systems Team
MERL - Mitsubishi Electric Research Labs, Cambridge, USA

Benelux Meeting on Systems and Control 2022

A hybrid approach to control: classical control theory meets data-driven methods

Acknowledgment of collaborators: J. Poveda, M. Guay, F. Lewis, M. Krstić, A. Teel, A. Scheinker, K. Vamvoudakis, A. Subbaraman,

S. Koga, A.-M. Farahmand, S. Russel, J. Queeney, S. Mowlavi, J. Borggaard, S. Nabi, O. San, P. Grover, G. Atinc, M. Xia, J. van Baar,

C. Wei, C. Weidong, Y. Sun, R. Ma….etc.

- MERL (ex-)interns (https://www.merl.com/internship/openings)

Big thanks to the organizers of this meeting for their warm welcome: Alain Vande Wouwer, Michel Kinnaert, Emanuele Garone, Laurent Dewasme, Guilherme Pimentel, Erjen Lefeber, William Van Hoeck…

MITSUBISHI ELECTRIC
Changes for the Better

Adaptation vs. Learning ?

Before reviewing some results in the field of adaptation and learning, let us first define the two terms: Learn and adapt. Refereing to the Oxford dictionary we find these two definitions; Adapt is defined as: *to change something in order to make it suitable for a new use or situation, or to change your behavior in order to deal more successfully with a new situation.* As for learn, it is defined as: *to gain knowledge and skill by studying, from experience, or to gradually change your attitudes about something so that you behave in a different way.* [Benosman 2016]

Adaptation: change
Learning: gradual change by repetition

MITSUBISHI ELECTRIC
Changes for the Better

Main points of the talk

<u>Part 1: Theory*</u>

- Brief survey of adaptive control: model-based adaptation, data-driven (classical RL & control theory inspired RL, extremum seeking control), and learning-based adaptation (hybrid: model-based + data-driven)
- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)
- Learning-based feedback gains auto-tuning for nonlinear systems affine in the control (ESC)
- Indirect learning-based adaptive control for linear systems under constraints (MPC framework) (ESC)
- Learning-based adaptive PDEs stable model reduction and estimation (ESC, RL)

<u>Part 2: Examples</u>

- Mechatronics applications: Electromagnetic brakes, servo motors
- Fluid dynamics applications: Airflow modeling and estimation
- Robotics applications

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, MPC: Model predictive control. RL: Reinforcement learning, PDE: Partial diff. equations.

A hybrid approach to control: classical control theory meets data-driven methods

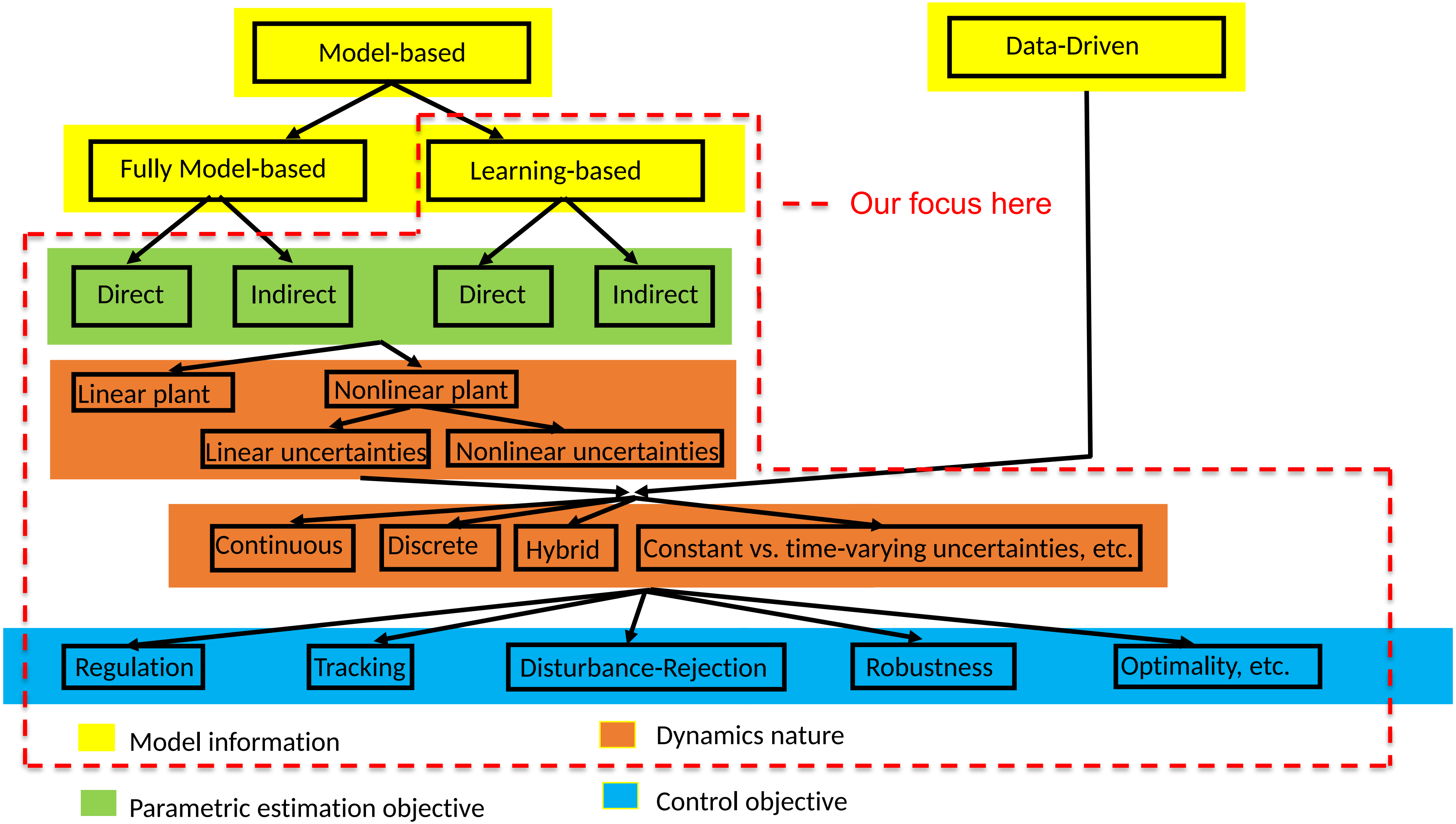
Mouhacine Benosman

MERL - Mitsubishi Electric Research Labs, Cambridge, USA

Part I: Brief survey and some theoretical results

Adaptation in Control

Figure/classification from: M. Benosman, 2018, ‘’Model-based vs. Data-Driven Adaptive Control: An Overview", International Journal of Adaptive Control and Signal Processing, 32(5), pp. 753-776.



(Fully) Model-based Adaptation

Classical Adaptive Control

Figure from Ioannou P., Sun J., 2012

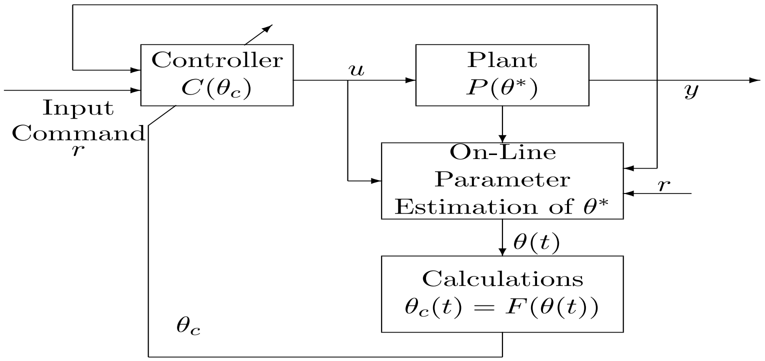


Figure 1.6 Indirect adaptive control.

- Model of the system, e.g., law of physics or Input/Output models
- Controller and filter are based on the model of the system

- Linear model (direct vs. indirect adaptation), e.g., Ioannou et al. 2012, Landau et al. 2011, 2017, Goodwin et al. 1984, 2014, Narendra et al. 1989, Tsakalis et al. 93, Sastry 2011, Tao 2003, Mosca 95
- Nonlinear model (direct vs. indirect adaptation), e.g., Krstic et al. 95, Slotine et al. 91, Spooner 2002, Astolfi et al. 2008, Fradkov et al. 99, Astolfi 2015, Guay et al. 2015, Taylor et al. 2020

- Infinite dimension and delays, e.g., Wen et al. 89, Smyshlyaev et al. 2010

- Constrained model (MPC type), e.g., Mosca 95, Guay et al. 2015

- Stochastic model, e.g., Sragovich 2006

- Multi-agent model, e.g., Lewis et al. 2014

Data-Driven Adaptation

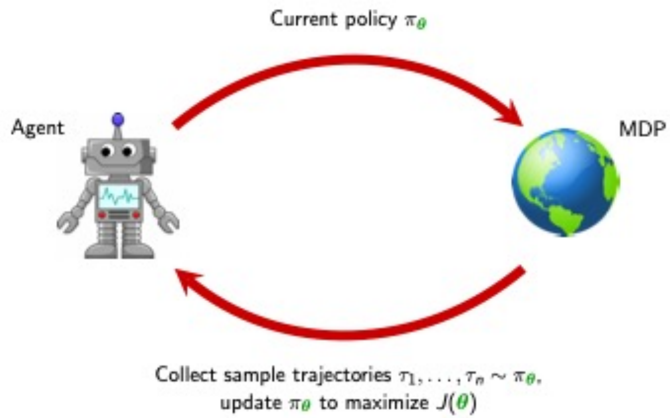


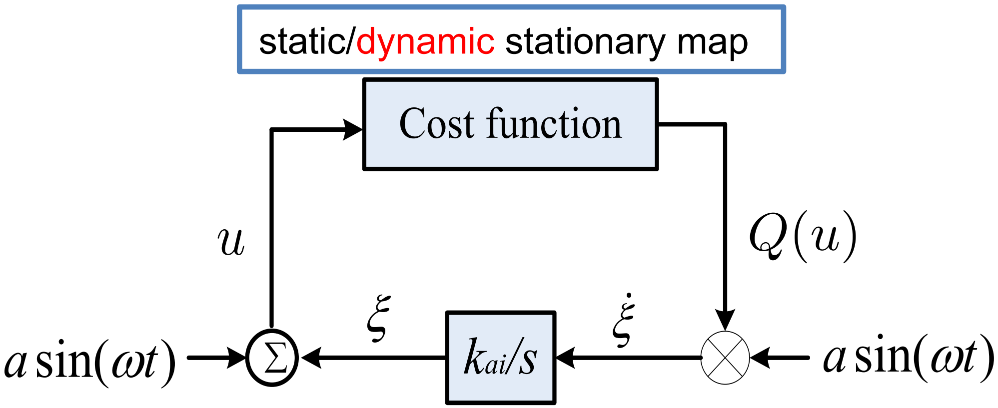
Figure courtesy of Jemmy Queeney@Boston University

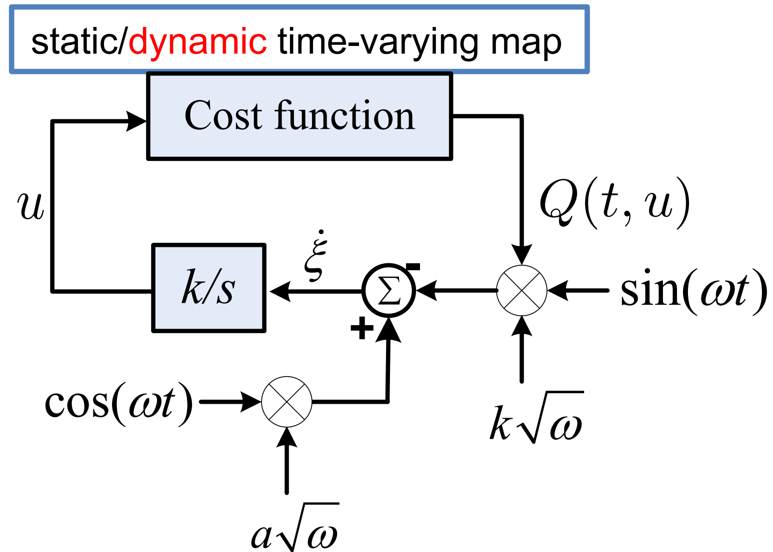
- ❖ Reinforcement Learning(RL): Stochastic Markov Decision Process (MDP)
- ❖ RL: Control policies are designed from interaction with a simulator and/or with the real environment
- ❖ (Approximate) Dynamic programming, (Approximate solutions) Bellman optimality equation

- Classical (CS) RL:
  - Model-based data generation (simulator-based/enhanced learning), e.g., Werbos 92, Bertsekas 96, Powell 2007, Busoniu 2010, Levine et al. 20, As et al., 2022
  - Model-free (real environment-based learning), e.g., Sutton et al. 98, Levine et al. 20
  - Multi-agent models, e.g., Oliehoek et al. 2016

- Control theory-‘inspired’ RL:
  - Lyapunov-based RL, e.g., Perkins et al. 2002, Chow et al. 2018, Chow et al. 2019, Russel et al. 2021

Data-Driven Adaptation





- ❖ Extremum seeking control (ESC)
- ❖ Data-driven optimization with estimation of the (higher order) derivatives of the cost function, i.e.. 'zero-order' optimization

- Deterministic, e.g., Leblanc 1922, Krstic et al. 2000, Ariyur et al. 03, Zhang et al. 12, Scheinker et al. 16, Feiling et al. 21, Dürr et al. 13, Nešic et al. 13, Tan et al. 2013, Guay et al. 15, Guay et al. 20, Benosman et al. 21a, Poveda et al. 21

- Stochastic, e.g., Liu et al. 12, Manzie et al. 09, Radenkovic et al. 16

- Infinite dimension, e.g., Oliveira et al. 20, Oliveira et al. 21, Feiling et al. 18

- Hybrid, e.g., Poveda et al. 17, Poveda 2018

- Multi-agent, e.g., Poveda 2018, Poveda 21a, Poveda 21b

SECTION INDUSTRIELLE

Sur l'électrification des chemins de fer au moyen de courants alternatifs de fréquence élevée

Une légère modification à un dispositif permettant la transformation d'un courant continu en courant alternatif de fréquence élevée, décrit dans la « R. G. E. » du 19 août 1922, t. XII, p. 259-261, a permis à M. Maurice Leblanc d'envisager l'alimentation d'une ligne de transmission d'énergie pour la traction par l'électricité, le récepteur d'énergie électrique n'ayant aucun point

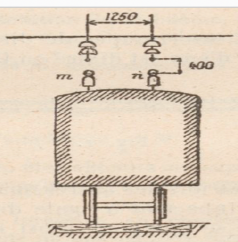


Fig. .

2 mm d'épaisseur. Chacun s'étend au-dessous d'un des conducteurs de la ligne de transmission avec une distance, d'axe en axe, de 40 cm.

Ces tubes sont portés par des isolateurs et reliés à ceux des voitures suivantes par des conducteurs

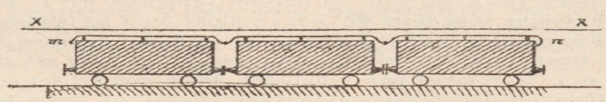


Fig. 4.

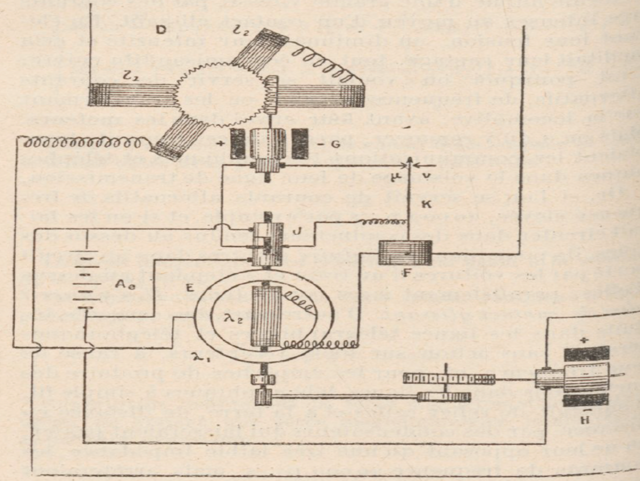


Fig. 5.

MITSUBISHI ELECTRIC
Changes for the Better

Data-Driven Adaptation

- Iterative Learning Control (ILC), e.g., Owens 2015
- Genetic algorithms, e.g., Dracopoulos 2013
- DNN/ DNN- RL, e.g., Arulkumaran et al. 2017, Levin 2013, Wang et al. 2016

Learning-based Adaptation

❖ Learning-based (hybrid: model-based control + data-driven adaptation)

Computer science

Data-driven learning

Control theory

Model-based control

Learning-based adaptive control

Merging model-based control and data-driven learning algorithms

MITSUBISHI ELECTRIC
Changes for the Better

Learning-based Adaptation

❖ Learning-based (hybrid: model-based control + data-driven adaptation)

<u>ID-based (indirect adaptation):</u>

- ESC *-based, e.g., Benosman 2016
- GP-based, e.g., Benosman et al. (2017a,2017b, 2018, 2019), Berkenkamp et al. 2017, Chakrabarty et al. 2021
- NN-based, e.g., Lewis et al. 99, Spooner et al. 02, Wang et al. 2010
- Learning-(ID) MPC, e.g., Benosman et al. 2014, Subbaraman et al. 2016, Limon et al. 2017, Hewing et al. 2020
- Control barrier functions (CBFs)-based, e.g., Lopez et al. 2020, Emam et al. 2021

<u>'Not' ID–based (direct adaptation):</u>

- 'Deterministic' RL: ADP, e.g., Vrabie et al. 2013, Lewis et al., 2013, Faust et al. 2014, Dalal et al. 2018, Marvi et al. 20, Vamvoudakis et al. 2021, CBFs-based learning, e.g., Cheng et al. 2019
- Feedback controller tuning, e.g., Gain tuning, e.g., Hjalmarsson 02, Benosman 2016, Duivenvoorden et al. 2017, Benosman et al. 21b, MPC hyper-parameters tuning, e.g., Hewing et al. 2020

* ESC: Extremum seeking control, GP: Gaussian process, ADP: Adaptive dynamic programming, MPC: Model predictive control.

Survey References

Classical model-based adaptive control

- Ioannou P., Sun J., 2012, Robust Adaptive Control. Mineola, NY: Dover Publications, Inc.
- Landau I.D., Lozano R., M'Saad M., Karimi A., 2011, Adaptive control: Algorithms, analysis and applications. *Communications and Control Engineering*. Springer-Verlag London Limited.
- Landau I.D., Airimitoaie T-B., Castellanos-Silva A., Constantinescu A., 2017, Adaptive and robust active vibration control: Methodology and tests. *Advances in Industrial Control*. Springer International Publishing Switzerland.
- Goodwin G.C., Sin K.S., 1984, Adaptive Filtering Prediction and Control. Englewood Cliffs, NJ: Prentice-Hall.
- Goodwin G.C., Sin K.S., 2014, Adaptive Filtering Prediction and Control. Mineola, NY: Dover Publications, Inc.
- Narendra K.S., Annaswamy A.M., 1989, *Stable Adaptive Systems*. Mineola, NY: Dover Publications, Inc.
- Tsakalis K.S., Ioannou P.A., 1993, Linear Time Varying Systems: Control and Adaptation. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Sastry S., Bodson M., 2011, Adaptive Control: Stability, Convergence and Robustness. Mineola, NY: Dover Publications, Inc.
- Tao G., 2003, Adaptive Control Design and Analysis. New York, NY: John Wiley & Sons, Inc.
- Mosca E., 1995, Optimal, Predictive, and Adaptive Control. Upper Saddle River, NJ, USA: Prentice Hall.
- Krstic M., Kanellakopoulos I., Kokotovic PV., 1995, Nonlinear and Adaptive Control Design. New York, NY: Wiley.
- Slotine J., Li W., 1991, Applied Nonlinear Control. Englewood Cliffs, NJ: Prentice-Hall:68-73.
- Spooner JT., Maggiore M., Ordòñez R., Passino KM., 2002, Stable Adaptive Control and Estimation for Nonlinear Systems. New York, NY: John Wiley & Sons, Inc.
- Astolfi A., Karagiannis D., Ortega R., 2008, Nonlinear and Adaptive Control with Applications. London, UK: Springer.

Survey References

<u>Classical model-based adaptive control</u>

- Fradkov A., Miroshnik I., Nikiforov V., 1999, Nonlinear and Adaptive Control of Complex Systems. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Astolfi A., 2015, Nonlinear adaptive control. Encyclopedia of Systems and Control. New York, NY: Springer; 866-870.
- Guay M., Adetola V., DeHaan D., 2015, Robust and Adaptive Model Predictive Control of Nonlinear Systems. London, UK: The Institution of Engineering and Technology.
- Taylor A. J. and Ames A. D., 2020, Adaptive Safety with Control Barrier Functions. In: 2020 American Control Conference (ACC). IEEE , Piscataway, NJ, pp. 1399-1405. ISBN 9781538682661.
- Wen J., Balas M., 1989, Robust adaptive control in Hilbert space. *J Math Anal Appl;* 143(1):1-26.
- Smyshlyaev A., Krstic M., 2010, *Adaptive Control of Parabolic PDEs*. Princeton, NJ: Princeton University Press.
- Mosca E., 1995, Optimal, Predictive, and Adaptive Control. Upper Saddle River, NJ, USA: Prentice Hall.
- Sragovich V., 2006, Mathematical Theory of Adaptive Control. Interdisciplinary Mathematical Sciences. Vol. 4. Singapore: World Scientific.
- Lewis F.L., Zhang H., Hengster-Movric K., Das A., 2014, Cooperative Control of Multi-Agent Systems Optimal and Adaptive Design Approaches, Springer, doi.org/10.1007/978-1-4471-5574-4.

Survey References

Data-driven adaptive control: Classical reinforcement learning

- Werbos PJ., 1992, Approximate dynamic programming for real-time control and neural modeling. In: White DA, Sofge DA, eds. Handbook of Intelligent Control: Neural, Fuzzy, and Adaptative Approaches. New York, NY: Van Nostrand Reinhold Library.
- Powell W.B., 2007, Approximate Dynamic Programming: Solving the Curses of Dimensionality. Hoboken, NJ: John Wiley & Sons.
- Bertsekas D., Tsitsiklis J., 1996, Neurodynamic Programming. Bellmont, MA: Athena Scientific.
- Busoniu L., Babuska R., De Schutter B., Ernst D., 2010, Reinforcement learning and dynamic programming using function approximators. *Automation and Control Engineering*. Boca Raton, FL: CRC Press.
- Levine et al., 2020 Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems. https://arxiv.org/abs/2005.01643.
- As et al., 2022, Constrained policy optimization via Bayesian world models, arXiv:2201.09802v4.
- Sutton RS, Barto AG., 1998, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Oliehoek F. A. et al. 2016, A Concise Introduction to Decentralized POMDPs, doi.org/10.1007/978-3-319-28929-8, Springer.

Data-driven adaptive control: Control theory- inspired reinforcement learning

- Perkins T.J., Barto A.G., 2000, Lyapunov-constrained action sets for reinforcement learning. ICML.
- Chow et al., 2018, A Lyapunov-based Approach to Safe Reinforcement Learning, arXiv:1805.07708v1.
- Chow et al., 2019, Lyapunov-based Safe Policy Optimization for Continuous Control, arXiv:1901.10031v2 .
- Russel R.H., et al., 2021, Lyapunov robust constrained-MDPs: Soft-constrained robustly stable policy optimization under model uncertainty, , arXiv:2018.02701v1.

Survey References

MITSUBISHI ELECTRIC
Changes for the Better

Data-driven adaptive control: Extremum seeking

- Leblanc M. 1922, Sur l'Électrification des Chemins de fer au Moyen de Courants Alternatifs de Fréquence Élevée. Revue Générale de l'Electricité, 12 (8), 275-277.
- Ariyur KB., Krstic M., 2003, Real Time Optimization by Extremum Seeking Control. New York, NY: John Wiley & Sons, Inc.
- Krstic M., Wang H.H., 2000, Stability of extremum seeking feedback for general nonlinear dynamic systems, Automatica, 36(4), 595-601.
- Scheinker A., Krstic M. 2016, Model-Free Stabilization by Extremum Seeking. Cham, Switzerland: Springer.
- Zhang C., Ordonez R., 2012, Extremum-Seeking Control and Applications: A Numerical Optimization-Based Approach. New York, NY: Springer.
- Feiling J., Belabbas M. A. and Ebenbauer C., 2021, Gradient Approximation and Multi-Variable Derivative-Free Optimization based on Non-Commutative Maps, in IEEE Transactions on Automatic Control, doi: 10.1109/TAC.2021.3129741.
- Dürr HB., Stanković MS., Ebenbauer C., Johansson KH., 2013,Lie bracket approximation of extremum seeking systems, Automatica 49 (6), 1538-1552
- Nešic D., Nguyen T., Tan Y., Manzie C. 2013, A non-gradient approach to global extremum seeking: an adaptation of the Shubert algorithm. Automatica.;49(3):809-815.
- Guay M., Dochain D. A., 2015, Time-varying extremum-seeking control approach. Automatica.;51:356-363.
- Guay M., Benosman M., 2020, Finite-time extremum seeking control for a class of unknown static maps, International Journal of Adaptive Control and Signal Processing 35(3).
- Poveda J., Krstic M., 2021, Non-smooth Extremum Seeking Control With User-Prescribed Fixed-Time Convergence, IEEE Transactions on Automatic Control; 66(12).
- Benosman et al., 2021a, Editorial for the special issue on extremum seeking control: Theory and applications, Int. J. of Adaptive Contr. and Signal Processing, doi.org/10.1002/acs.3293
- Tan Y., Li Y., Mareels I., 2013, Extremum seeking for constrained inputs. IEEE Transactions on Automatic Control;58(9):2405-2410.

Survey References

Data-driven adaptive control: Extremum seeking

- Liu S-J., Krstic M., 2012, Stochastic Averaging and Stochastic Extremum Seeking. Springer,
- Manzie C, Krstic M., 2009, Extremum seeking with stochastic perturbations. IEEE Transactions on Automatic Control;54:580-585.
- Radenkovic MS, Altman T. 2016, Almost sure convergence of extremum seeking algorithm using stochastic perturbation. Systems & Control Letters; 94:133-141.
- Oliveira T. R., Krstic M., 2021, Extremum seeking boundary control for PDE–PDE cascades, Systems & Control Letters,155.
- Oliveira T.R., Feiling J., Koga S., Krstic m., 2020, Extremum seeking for unknown scalar maps in cascade with a class of parabolic partial differential equations, DOI: 10.1002/acs.3117.
- Feiling J., Koga S., Krstić M., Oliveira T.R., 2018, Gradient extremum seeking for static maps with actuation dynamics governed by diffusion PDEs, Automatica 95, 197-206.
- Poveda J.I., Tell A.R., 2017, A framework for a class of hybrid extremum seeking controllers with dynamic inclusions, Automatica 76, 113-126.
- Poveda J.I., 2018, Robust Hybrid Systems for Control, Learning, and Optimization in Networked Dynamical Systems, PhD thesis, UCSB.
- Poveda J.I., Benosman M., Teel A.R., Sanfelice R.G., 2021a, Robust Coordinated Hybrid Source Seeking with Obstacle Avoidance in Multi-Vehicle Autonomous Systems, IEEE Transactions on Automatic Control, 10.1109/TAC.2021.3056365.
- Poveda J.I., Benosman M., Vamvoudakis K.G., 2021b, Data-enabled extremum seeking: a cooperative concurrent learning-based, International Journal of Adaptive Control and Signal Processing 35 (7), 1256-1284.

Survey References

Data-driven adaptive control: Others

- Owens D.H., 2015, Iterative Learning Control: An Optimization Paradigm. London, UK: Springer.
- Dracopoulos D., 2013 Evolutionary Learning Algorithms for Neural Adaptive Control. London, UK: Springer.
- Arulkumaran K. et al., 2017, Deep reinforcement learning: a brief survey. *IEEE Signal Process Mag*;34(6):26-38.
- Levine S., 2013, Exploring deep and recurrent architectures for optimal control. Paper presented at: Neural Information Processing Systems (NIPS) Workshop on Deep Learning; 2013; Lake Tahoe, CA.
- Wang Z. et al., 2016, *Qualitative Analysis and Control of Complex Neural Networks with Delays*. Vol. 34. Berlin, Germany: Springer-Verlag.

Survey References

Learning-based adaptive control:

- Benosman M., 2016, Learning-based adaptive control: An extremum seeking approach-Theory and Applications, Elsevier.
- Benosman M., Farahmand A.-M., 2017a, Towards stability in learning-based control: A Bayesian optimization-based adaptive controller. Multi-Disciplinary Conference on Reinforcement Learning and Decision Making (RLDM).

Benosman M., Farahmand A.-M., 2017b, Gaussian Processes-based Parametric Identification for Dynamical Systems, IFAC-PapersOnLine Volume 50, Issue 1, July 2017, Pages 14034-14039

- Benosman M., Farahmand A.-M., Xia M., 2018, Learning-based iterative modular adaptive control for nonlinear systems, International Journal of Adaptive Control and Signal Processing, 33(2), pp. 335-355, doi.org/10.1002/acs.2892.
- Benosman et al., 2019, Editorial for the special issue on learning-based adaptive control: Theory and applications, Int. J. of Adaptive Contr. and Signal Processing, doi.org/10.1002/acs.2964.
- Chakrabarty A., Benosman M., 2021, Safe learning-based observers for unknown nonlinear systems using Bayesian optimization, Automatica, vol. 133, doi.org/10.1016/j.automatica.2021.109860.
- Berkenkamp et al. 2017, Safe model-based reinforcement learning with stability guarantees. Paper presented at: 2017 Conference on Neural Information Processing Systems (NIPS).
- Lewis et al. 99, Neural Network Control of Robot Manipulators and Non-Linear Systems. London, UK: Taylor & Francis.
- Spooner et al. 02, Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques. New York, NY: John Wiley & Sons, Inc.
- Wang et al. 2010, Deterministic Learning Theory for Identification, Recognition, and Control. Boca Raton, FL: CRC Press.
- Subbaraman A., Benosman M., 2016. , Extremum Seeking-based Iterative Learning Model Predictive Control (ESILC-MPC), IFAC International Workshop on Adaptation and Learning in Control and Signal Processing.
- Benosman M., Di Cairano S., Weiss A., 2014, Extremum seeking-based iterative learning linear MPC, IEEE Conference on Control Applications.
- Limon et al., 2017, Learning-based nonlinear model predictive control. *IFAC-PapersOnLine*;50(1):7769-7776.
- Hewing et al. 20, Learning-Based Model Predictive Control: Toward Safe Learning in Control, Annual Review of Control, Robotics, and Autonomous Systems.

Survey References

- Lopez et al., 2020, Robust Adaptive Control Barrier Functions: An Adaptive & Data-Driven Approach to Safety, IEEE L-CSS, 5(3), pp. 1031 - 1036.
- Vravie et al., 2013, Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles. England: IET Digital Library.
- Lewis et al., 2012, *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. Hoboken, New Jersey: John Wiley/IEEE Press, Computational Intelligence Series.
- Faust et al., 2014, Continuous Action Reinforcement Learning for Control-Affine Systems with Unknown Dynamics, IEEE/CAA Journal of Automatica Sinica, 1(3), pp. 323 - 336.
- Dalal et al., 2018, Safe Exploration in Continuous Action Spaces, arXiv:1801.08757v1.
- Marvi et al., 2020, Safe Off-policy Reinforcement Learning Using Barrier Functions, IEEE American Control Conference.
- Vamvoudakis et al., 2021, Handbook of Reinforcement Learning and Control, Springer, 2021.
- Cheng et al 2019, End-to-End Safe Reinforcement Learning through Barrier Functions for Safety-Critical Continuous Control Tasks, The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19).
- Emam et al. 21, Safe Model-Based Reinforcement Learning using Robust Control Barrier Functions, arXiv:2110.05415v1.
- Hjalmarsson H., 2002, Iterative feedback tuning an overview. *International Journal of Adaptive Control and Signal Processing* 2002; **16**(5):373–395.
- Duivenvoorden et al., 2017, Constrained Bayesian optimization with particle swarms for safe adaptive controller tuning," IFAC-PapersOnLine, vol. 50, no. 1, pp. 11 800–11 807.
- Benosman et al., 2021b, Data-driven robust state estimation for reduced-order models of 2D Boussinesq equations with parametric uncertainties, Computers and Fluids 214 (2021) 104773.

MITSUBISHI ELECTRIC RESEARCH LABORATORIES
Cambridge, Massachusetts

Learning-based Adaptive Control for Nonlinear Systems*

* M. Benosman, 2016, Learning-based adaptive control: An extremum seeking approach-Theory and Applications, Elsevier.

Main points

- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)*

- Learning-based iterative feedback gains tuning for nonlinear systems affine in the control (ESC)

- Indirect learning-based adaptive iterative control for linear systems under constraints (ESC-MPC framework)

- Learning-based adaptive PDE stable model reduction/ estimation (ESC, RL)

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, RL: Reinforcement learning

Main points

- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)*

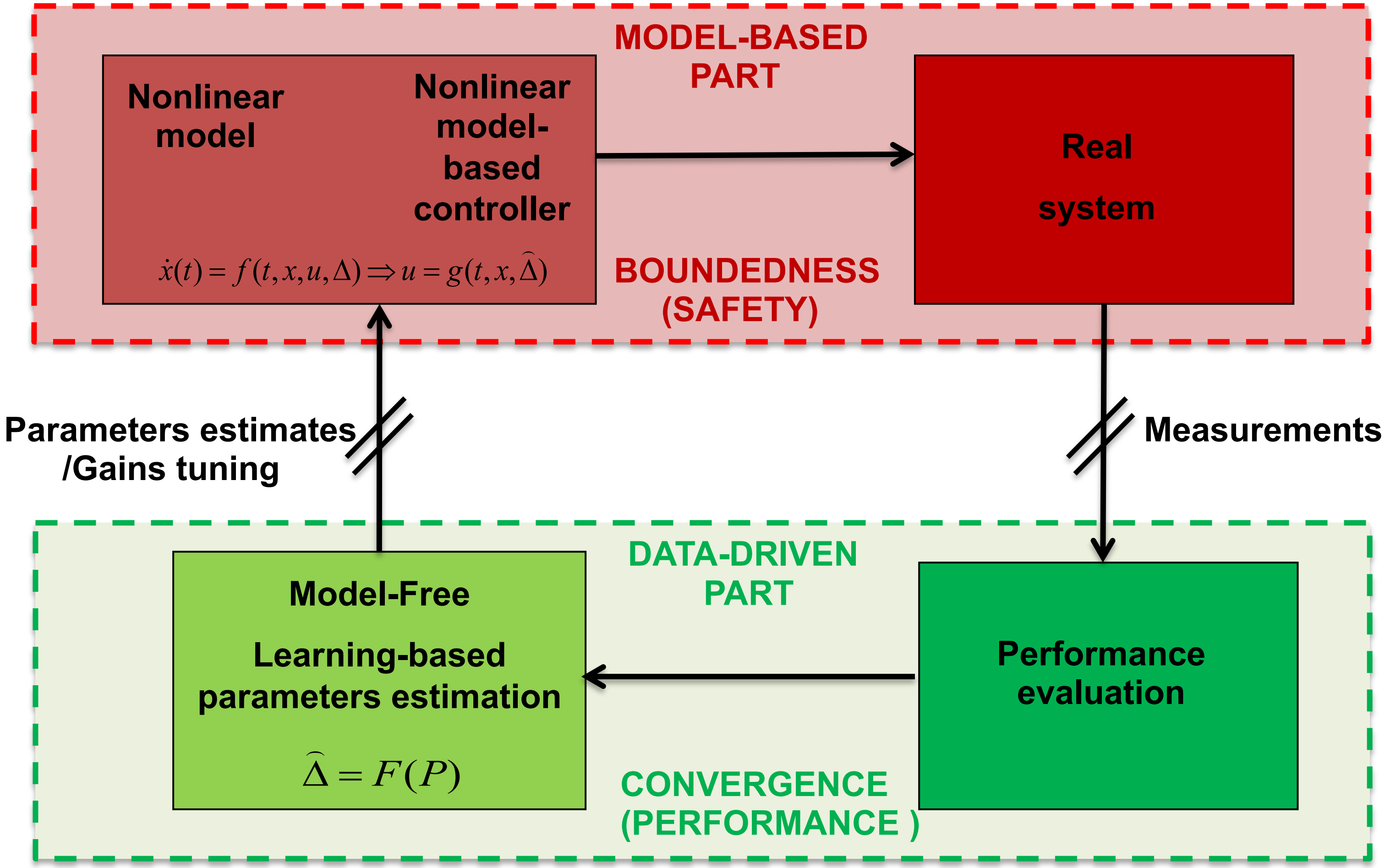
- Learning-based iterative feedback gains tuning for nonlinear systems affine in the control (ESC)

- Indirect learning-based adaptive iterative control for linear systems under constraints (ESC-MPC framework)

- Learning-based adaptive PDE stable model reduction/ estimation (ESC, RL)

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, RL: Reinforcement learning

Learning-based adaptive Control: a modular approach



Learning-based indirect adaptive control for constant uncertainties

$$\dot{x} = f(x, \Delta, u)$$

$\Delta \in \mathbb{R}^p$ parametric uncertainties

the output vector $y = h(x)$

where $h: \mathbb{R}^n \to \mathbb{R}^h$, with smoothness of *f*, and *h*.

The control objective here is for y to asymptotically track a desired smooth vector time-dependent trajectory $y_{ref} : [0, \infty) \to \mathbb{R}^h$.

Learning-based indirect adaptive control for constant uncertainties

Modularity through (ISS) robustness

$\dot{x} = f(t, x, u)$ is LiISS* if and only

if there exist functions $\beta \in \mathcal{KL}$ and $\gamma_1, \gamma_2 \in \mathcal{K}$ such that

$$\|x(t,\xi,u)\| \leq \beta(\|\xi\|,t) + \gamma_1\left(\int_0^t \gamma_2(\|u(s)\|)ds\right)$$

* Ito H., and Jiang Z., 2009, Necessary and sufficient small gain conditions for integral input-to-state stable systems: A Lyapunov perspective,. IEEE Transactions on Automatic Control, vol. 54, no. 10, pp. 2389.2404,

Learning-based indirect adaptive control for constant uncertainties

Assumption 1:

$$e_y(t) = y(t) - y_{ref}(t).$$

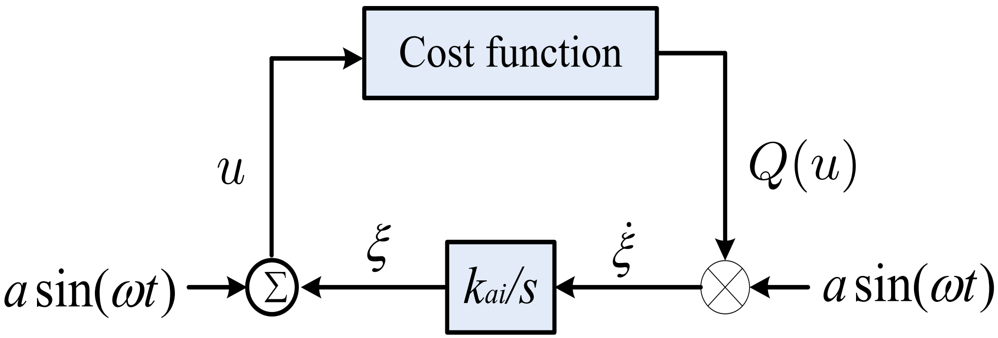
$$\exists \; u_{iss}(t, x, \hat{\Delta}): \quad \mathbb{R}^{\times}\mathbb{R}^n \times \mathbb{R}^p \; \rightarrow \; \mathbb{R}^m$$

$$\dot{e}_y = f_{e_y}(t, e_y, e_\Delta)$$

is iISS from the input vector $e_\Delta = \Delta - \hat{\Delta}$ to the state vector e_y.

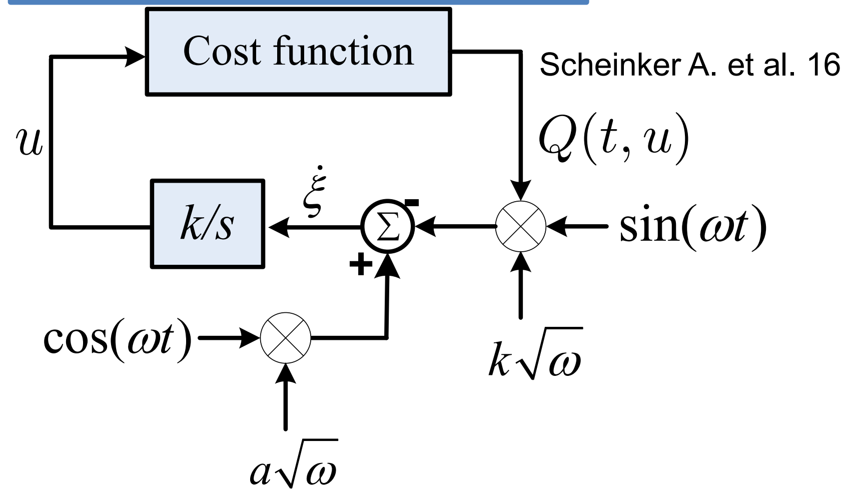
Concept of (dither-based) Extremum Seeking Control (ESC)*

static/dynamic stationary map



Block diagram of a static extremum seeking control algorithm

static/dynamic time-varying map



Block diagram of a functional extremum seeking control algorithm

Advantages:

- Model-free (zero-order) optimization
- Gradient implicit estimate using one measurement per learning iteration (good for real-time applications)
- Robustness to noise
- Robustness to initial conditions
- Input and state constraints

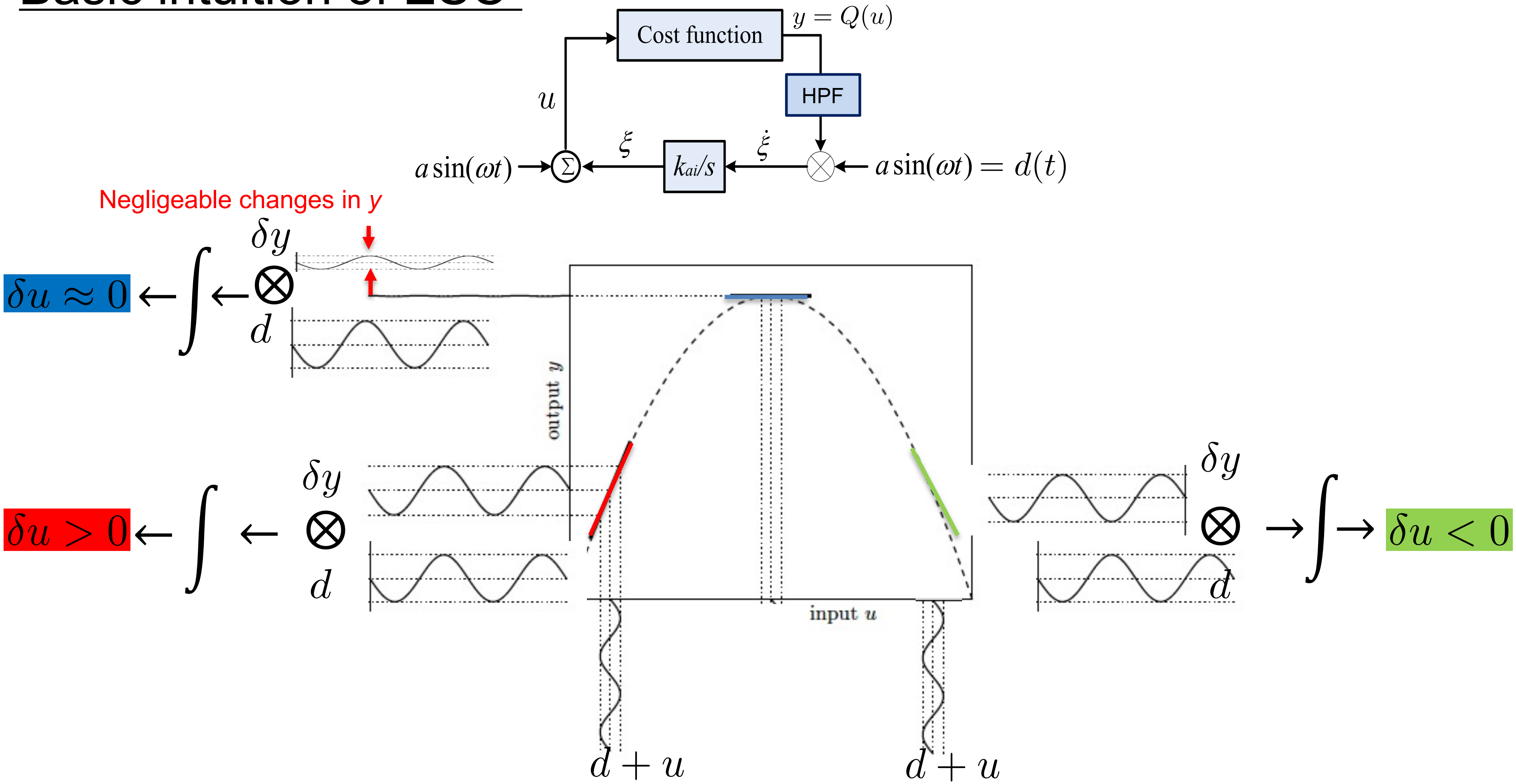
Analysis*:

- Averaging theory
- Singular perturbation theory (for dynamic maps)

* Ariyur K.B., Krstic M., 2003, Real Time Optimization by Extremum Seeking Control. New York, NY: John Wiley & Sons, Inc. (Note: see this link for an 'easier' introduction: http://flyingv.ucsd.edu/krstic/talks/talks-files/extremum-seeking-DISC12.pdf)

Learning-based indirect adaptive control for constant uncertainties

Basic intuition of ESC



Learning-based indirect adaptive control for constant uncertainties

ESC uncertainties estimator

$$\text{cost function} \;\; Q(\hat{\Delta}) = F(e_y(\hat{\Delta}))$$

$$\text{where} \;\; F: \;\; \mathbb{R}^h \rightarrow \mathbb{R},$$

$$F(0) = 0, \;\; F(e_y) > 0 \text{ for } e_y \neq 0$$

Assumed to be well defined, i.e., for the same $\hat{\Delta}$, we obtain the same $Q(\hat{\Delta})$ → If not intrinsically, it can be forced by *an iterative or batch-to-batch* implementation

Learning-based indirect adaptive control for constant uncertainties

Assumption 2:
Q has a local minimum at $\hat{\Delta}^* = \Delta$

Assumption 3:
$e_{\Delta}(t_0)$ is sufficiently small

Assumption 4:
Q is analytic $\|\frac{\partial Q}{\partial \Delta}(\tilde{\Delta})\| \leq \xi_2,\ \xi_2 > 0,$
$\tilde{\Delta} \in \mathcal{V}(\Delta^*)$

Learning-based indirect adaptive control for constant uncertainties

*Lemma**: - Model-based - Data-driven

the system $\dot{x} = f(x, \Delta, u)$ with the cost Q under Assumptions 1, 2, 3, and 4, the control u_{iss}, where $\hat{\Delta}$ is estimated with the multi-parameter extremum seeking

$$\dot{x}_i = a_i sin(\omega_i t + \frac{\pi}{2}) Q(\hat{\Delta})$$
$$\hat{\Delta}_i = x_i + a_i sin(\omega_i t - \frac{\pi}{2}),\ i \in \{1, ..., p\}$$

* M. Benosman, 2014, Learning-based Adaptive Control for Nonlinear Systems, European Control Conference.

Learning-based indirect adaptive control for constant uncertainties

Lemma: Cont.

with $\omega_i \neq \omega_j, \ \omega_i + \omega_j \neq \omega_k, \ i, j, k \in \{1, ..., p\}$

ensures that

$$\|e_y(t)\| \leq \beta(\|e_y(0)\|, t) + \alpha(\int_0^t \gamma(\tilde{\beta}(\|e_\Delta(0)\|, t) + \|e_\Delta\|_{max}))ds,$$

where $\|e_\Delta\|_{max} = \frac{\xi_1}{\omega_0} + \sqrt{\sum_{i=1}^{i=p} a_i^2}$, $\xi_1, \xi_2 > 0$, $e(0) \in \mathcal{D}_e$,

$\omega_0 = \max_{i \in \{1,\ldots,p\}} \omega_i$, $\alpha \in \mathcal{K}$, $\beta \in \mathcal{KL}$, $\tilde{\beta} \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$.

Learning-based indirect adaptive control for time-varying systems

$$\dot{x} = f(t, x, \Delta, u)$$

$\Delta \in \mathbb{R}^p$ parametric uncertainties

the output vector $y = h(x)$

where $h: \mathbb{R}^n \to \mathbb{R}^h$, with f being piecewise continuous in t and(at least) locally Lipschitz in x, u, uniformly in t, h is smooth.

The control objective here is for y to asymptotically track a desired smooth vector time-dependent trajectory $y_{ref} : [0, \infty) \to \mathbb{R}^h$.

Learning-based indirect adaptive control for time-varying systems

Assumption 1:

$$e_y(t) = y(t) - y_{ref}(t).$$

$$\exists \; u_{iss}(t, x, \hat{\Delta}): \quad \mathbb{R}^{\times}\mathbb{R}^{n} \times \mathbb{R}^{p} \;\rightarrow\; \mathbb{R}^{m}$$

$$\dot{e}_y = f_{e_y}(t, e_y, e_\Delta)$$

is iISS from the input vector $e_\Delta = \Delta - \hat{\Delta}$ to the state vector e_y.

Learning-based indirect adaptive control for time-varying systems

ESC (time-varying) uncertainties estimator

cost function

$$Q(\hat{\Delta}, t) = F(e_y(\hat{\Delta}), t)$$

where $F: \mathbb{R}^h \times \mathbb{R}^+ \to \mathbb{R}^+, \; F(0, t) = 0,$

$$F(e_y, t) > 0, e_y \neq 0$$

Learning-based indirect adaptive control for time-varying systems

Assumption 2:

Q has a local minimum at $\hat{\Delta}^* = \Delta$

Assumption 3:

$$|\frac{\partial Q(\hat{\Delta},t)}{\partial t}| < \rho_Q, \forall t \in \mathbb{R}^+, \ \forall \hat{\Delta} \in \mathbb{R}^p.$$

Learning-based indirect adaptive control for time-varying systems

*Lemma**: - Model-based - Data-driven

the system $\dot{x} = f(t, x, \Delta, u)$ with the cost Q

then under Assumptions 1, 2, and 3,

the control u_{iss}, where $\hat{\Delta}$ is estimated

with the multi-parameter extremum seeking

$$\dot{\hat{\Delta}}_i = a\sqrt{(\omega_i)}cos(\omega_i t) - k\sqrt{\omega_i}sin(\omega_i t)Q(\hat{\Delta}, t)$$

$$i \in \{1, ..., p\}^{**}$$

* M. Benosman, 2014, Extremum Seeking-based Indirect Adaptive Control for Nonlinear Systems, IFAC World Congress.

** Scheinker A., Krstic M. 2016, Model-Free Stabilization by Extremum Seeking. Cham, Switzerland: Springer.

Learning-based indirect adaptive control for time-varying systems

Lemma: Cont.

with $a > 0$, $k > 0$, $\omega_i \neq \omega_j$, $i, j, k \in \{1, ..., p\}$, $\omega_i > \omega^*$, $\forall i \in \{1, ..., p\}$, with ω^* large enough, ensures

$$\|e_y(t)\| \leq \beta(\|e_y(0)\|, t) + \alpha(\int_0^t \gamma(\|e_\Delta(s)\|)ds,$$

where $\alpha \in \mathcal{K}$, $\beta \in \mathcal{KL}$, $\gamma \in \mathcal{K}$, and $\|e_\Delta\|$ satisfies:

Learning-based indirect adaptive control for time-varying systems

Lemma: Cont.

1-$(\frac{1}{\omega}, d)$-*Uniform Stability:* For every $c_2 \in]d, \infty[$, there exists $c_1 \in]0, \infty[$ and $\hat{\omega} > 0$ such that for all $t_0 \in \mathbb{R}$ and for all $x_0 \in \mathbb{R}^n$ with $||e_\Delta(0)|| < c_1$ and for all $\omega > \hat{\omega}$,

$$||e_\Delta(t, e_\Delta(0))|| < c_2, \ \forall t \in [t_0, \infty[$$

Learning-based indirect adaptive control for time-varying systems

Lemma: Cont.

2-$(\frac{1}{\omega}, d)$-*Uniform ultimate boundedness:* For every $c_1 \in]0, \infty[$ there exists $c_2 \in]d, \infty[$ and $\hat{\omega} > 0$ such that for all $t_0 \in \mathbb{R}$ and for all $x_0 \in \mathbb{R}^n$ with $||e_\Delta(0)|| < c_1$ and for all $\omega > \hat{\omega}$,

$$||e_\Delta(t, e_\Delta(0))|| < c_2, \; \forall t \in [t_0, \infty[$$

Learning-based indirect adaptive control for time-varying systems

Lemma: Cont.

3-$(\frac{1}{\omega}, d)$-*Global uniform attractivity:* For all $c_1,\ c_2 \in (d, \infty)$ there exists $T \in]0, \infty[$ and $\hat{\omega} > 0$ such that for all $t_0 \in \mathbb{R}$ and for all $x_0 \in \mathbb{R}^n$ with $||e_\Delta(0)|| < c_1$ and for all $\omega > \hat{\omega}$,

$$||e_\Delta(t, e_\Delta(0))|| < c_2,\ \forall t \in [t_0 + T, \infty[$$

where d is given by: $d = min\{r \in]0, \infty[:\ \Gamma_H \subset B(\Delta, r)\}$, with $\Gamma_H = \{\hat{\Delta} \in \mathbb{R}^n :\ ||\frac{\partial Q(\hat{\Delta}, t)}{\partial \hat{\Delta}}|| < \sqrt{\frac{2\rho_Q}{ka\beta_0}}\},\ 0 < \beta_0 \leq 1,$ and $B(\Delta, r) = \{\hat{\Delta} \in \mathbb{R}^n :\ ||\hat{\Delta} - \Delta|| < r\}$.

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable*

* Benosman M., Farahmand A.-M., Xia M.2018, Learning-based iterative modular adaptive control for nonlinear systems, International Journal of Adaptive Control and Signal Processing, 33(2), pp. 335-355, doi.org/10.1002/acs.2892.

We consider an output tracking problem for systems that are affine in the control

$$\dot{x} = f(x) + \Delta f(t,x) + g(x)u, \; x(0) = x_0, \qquad (3)$$

$$y = h(x),$$ with ref. trajectory $y_d(t)$.

under classical smoothness and relative degree assumptions, we can design an ISS controller satisfying,

$$\|e_y(t)\| \leq \beta(\|e_y(t_0)\|, t - t_0) + \gamma(\sup_{t_0 \leq \tau \leq t} \|e_\Delta(\tau)\|),$$

where e_y, e_Δ denote the output tracking error and the uncertainties estimation error, respectively.

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

ISS controller (Model-based)

$$y^{(r)}(t) = b(\xi(t)) + A(\xi(t))u(t) + \Delta b(t,\xi(t)),\ \Delta b(t,\xi(t)) = E\,Q(\xi,t), \quad (20)$$

$$u_f = u_n + u_r, \quad (14)$$

$$u_n = A^{-1}(\xi)\left[v_s(t,\xi) - b(\xi)\right], \quad (9)$$

I/O linearization

$$u_r = -A^{-1}(\xi)[\tilde{B}^T P z \|Q(\xi,t)\|^2 + \widehat{E}(t)Q(\xi,t)]. \quad (21)$$

Lyapunov reconstruction

$$\tilde{A}^T P + P\tilde{A} = -I. \quad (13)$$

$$y^{(r)}(t) = [y_1^{(r_1)}(t),\ y_2^{(r_2)}(t),\ \cdots,\ y_m^{(r_m)}(t)]^T,$$

$$\xi(t) = [\xi^1(t),\ \cdots,\ \xi^m(t)]^T,$$

$$\xi^i(t) = [y_i(t),\ \cdots,\ y_i^{(r_i-1)}(t)]. \quad 1 \le i \le m$$

$$v_{si} = y_{id}^{(r_i)} - K_{r_i}^i(y_i^{(r_i-1)} - y_{id}^{(r_i-1)}) - \cdots - K_1^i(y_i - y_{id}).$$

A, b are functions of *f, g,* and *h*, $\tilde{B}$ is a sparce matrix of 0s and 1s, $\tilde{A}$ is function of the feedback gains.

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

Multi-parametric ESC uncertainties estimator (Data-driven)

$$J(\widehat{\Delta}) = F(z(\widehat{\Delta})),\ \widehat{\Delta}(t) = [\hat{E}(1,1), ..., \hat{E}(m,m)]^T \quad (24)$$

where $F : \mathbb{R}^n \to \mathbb{R}$, $F(\mathbf{0}) = 0$, $F(z) > 0$ for $z \in \mathbb{R}^n - \{\mathbf{0}\}$.

$$\dot{\tilde{x}}_i = a_i \sin(\omega_i t + \tfrac{\pi}{2}) J(\widehat{\Delta}),\ a_i > 0, \quad i \in \{1, 2, \dots, m^2\}$$

$$\hat{\delta}\Delta_i(t) = \tilde{x}_i + a_i \sin(\omega_i t - \tfrac{\pi}{2}),$$

$$\widehat{\Delta}_i(t) = \widehat{\Delta}_{i-nominal} + \delta\Delta_i(t), \quad (25)$$

$$\delta\Delta_i(t) = \hat{\delta}\Delta_i((I-1)t_f),\ (I-1)t_f \le t \le It_f,$$

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

Algorithm 1 MES-based Learning Adaptive Controller

– Initialize: $I = 1,\ x(0) = x_0,\ J_{th} > 0,\ \hat{\Delta} = \Delta_{nominal},\ K_1^i, ..., K_{r_i}^i, i = 1, ..., m.$

– Solve (13).

– Apply the controller (9), (14), and (21), to (3), (20).

(Loop) – Evaluate the learning cost J by (24).

– IF $J \leq J_{th} \rightarrow$ Exit Loop, IF not:

– I=I+1.

– Estimate $\hat{\Delta}$ by (25).

– Reset $t \in [(I-1)t_f,\ It_f]$, $x((I-1)t_f) = x_0$, then, apply the controller (9), (14), and (21), to (3), (20).

– Go to (Loop).

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

GP-UCB* uncertainties estimator (Data-driven)

- Gaussian process upper confidence bound GP-UCB* is used as the data-driven part of the controller
- Bayesian stochastic optimization, i.e., noisy observation of the cost function
- Global optimum on compact search sets

* Srinivas N, Krause A, Kakade SM, Seeger M., 2010, Gaussian process optimization in the bandit setting: No regret and experimental design. In: Proceedings of the 27th International Conference on Machine Learning.

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

GP-UCB uncertainties estimator (Data-driven)

Let us assume that $\tilde{J}$ is a function sampled from a Gaussian Process (GP).

We recall that GP is defined by a mean function

$$\mu(\widehat{\Delta}) = \mathbb{E}\left[\tilde{J}(\widehat{\Delta})\right],$$

and its covariance function (or kernel)

$$\mathrm{K}(\widehat{\Delta}, \widehat{\Delta}') = \mathrm{Cov}(\tilde{J}(\widehat{\Delta}), \tilde{J}(\widehat{\Delta}')) = \mathbb{E}\left[\left(\tilde{J}(\widehat{\Delta}) - \mu(\widehat{\Delta})\right)\left(\tilde{J}(\widehat{\Delta}') - \mu(\widehat{\Delta}')\right)^{\top}\right].$$

e.g., $$\mathrm{K}(\widehat{\Delta}, \widehat{\Delta}') = \exp\left(-\frac{\|\widehat{\Delta} - \widehat{\Delta}'\|^2}{2l^2}\right), \quad (32)$$

as the squared exponential kernel with length scale $l > 0$

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

GP-UCB uncertainties estimator (Data-driven)

Let us first briefly describe how we can find the posterior distribution of a $\mathrm{GP}(0, \kappa)$, i.e., a GP with zero prior mean. Suppose that for $\underline{\widehat{\Delta}}_{I-1} \triangleq \{\widehat{\Delta}_1, \widehat{\Delta}_2, \ldots, \widehat{\Delta}_{I-1}\} \subset D$, we have observed the noisy evaluation $y_i = \tilde{J}(\widehat{\Delta}_i) = J(\widehat{\Delta}_i) + \eta_i$ with $\eta_i \sim N(0, \sigma^2)$ being i.i.d. Gaussian noise. We can find the posterior mean and variance for a new point $\widehat{\Delta}^* \in D$ as follows: Denote the vector of observed values by $\mathbf{y}_{I-1} = [y_1, \ldots, y_{I-1}]^\top \in \mathbb{R}^{I-1}$, and define the Grammian matrix $K \in \mathbb{R}^{I-1 \times I-1}$ with $[K]_{i,j} = \kappa(\widehat{\Delta}_i, \widehat{\Delta}_j)$, and the vector $\kappa_* = [\kappa(\underline{\widehat{\Delta}}_1, \widehat{\Delta}^*), \ldots, \kappa(\underline{\widehat{\Delta}}_{I-1}, \widehat{\Delta}^*)]$. The expected mean $\mu_I(\widehat{\Delta}^*)$ and the variance $\sigma_I(\widehat{\Delta}^*)$ of the posterior of the GP evaluated at $\widehat{\Delta}^*$ are (cf. Section 2.2 of [63])

$$\mu_I(\widehat{\Delta}^*) = \kappa_* \left[K + \sigma^2 \mathbf{I}\right]^{-1} \mathbf{y}_{I-1}, \tag{33}$$

$$\sigma_I^2(\widehat{\Delta}^*) = \kappa(\widehat{\Delta}^*, \widehat{\Delta}^*) - \kappa_*^T \left[K + \sigma^2 \mathbf{I}\right]^{-1} \kappa_*. \tag{34}$$

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

GP-UCB uncertainties estimator (Data-driven)

At iteration I, the GP-UCB algorithm selects the next query point $\widehat{\Delta}_I$ by solving the following optimization problem:

Nested nonlinear optimization
vs.
simple gradient estimation in
ESC !

$$\widehat{\Delta}_I \leftarrow \underset{\widehat{\Delta} \in D}{\operatorname{argmin}} \, \mu_{I-1}(\widehat{\Delta}) - \beta_I^{1/2} \sigma_{I-1}(\widehat{\Delta}). \tag{35}$$

Remark 12. The optimization problem (35) is often nonlinear and nonconvex. Nonetheless, solving it only requires querying the GP, which, in general, is much faster than querying the original dynamical system. This is important when the dynamical system is a real system and we would like to minimize the number of interactions with it before finding a $\hat{\Delta}$ with small $J(\hat{\Delta})$. One practical way to approximately solve (35) is to restrict the search to a finite subset D' of D. The finite subset can be a uniform grid structure over D or it might consist of randomly selected members of D.

6)

* Srinivas N, Krause A, Kakade SM, Seeger M., 2010, Gaussian process optimization in the bandit setting: No regret and experimental design. In: Proceedings of the 27th International Conference on Machine Learning.

Learning-based indirect adaptive iterative control for nonlinear systems affine in the control variable

Algorithm 2 GP-UCB-based Learning Adaptive Controller

– Initialize: $I = 1,\ x(0) = x_0,\ J_{th} > 0,\ \hat{\Delta} = \Delta_{nominal}$.

– Apply the controller (9), (14), and (21), to (3), (20).

(Loop) – Evaluate the learning cost J by (24).

– IF $J \leq J_{th} \rightarrow$ Exit Loop, IF not:

– I=I+1.

– Estimate $\hat{\Delta}$ by (32), (33), (34), (35), and (36).

– Reset $t \in [(I-1)t_f,\ t_f]$, $x((I-1)t_f) = x_0$, then, apply the controller (9), (14), and (21), to (3), (20).

– Go to (Loop).

The dual estimation problem can be solved using similar 'robustness + learning' approach, e.g., Chakrabarty A., Benosman M., 2021, Safe learning-based observers for unknown nonlinear systems using Bayesian optimization, Automatica, vol. 133, doi.org/10.1016/j, Automatica. Koga et al., 2021, Extremum Seeking-Based Robust Observer Design for Coupled Thermal and Fluid Systems, Int. J. of Adaptive Cont. and Signal Processing, 35(7).

Adaptive dynamic programming**

Linear time-invariant model

$$\dot{x} = Ax + Bu, \; x \in \mathbb{R}^n, u \in \mathbb{R}^m$$

where, *A* is unknown.

The pair (A, B) assumed to be stabilizable.

LQR-type cost function of the form

$$V(u) = \int_{t_0}^{\infty} (x^T(\tau) R_1 x(\tau) + u^T(\tau) R_2 u(\tau)) d\tau,$$
$$R_1 \geq 0, \; R_2 > 0.$$

$$u^*(t) = -Kx(t),$$
$$K = R_2^{-1} B^T P,$$

Model-based

** Vrabie D, Vamvoudakis K, Lewis FL. Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles. England: IET Digital Library; 2013.

** More details about ADP algorithms can be found in these two talks by F. Lewis:

https://lewisgroup.uta.edu/FL%20talks%202017/2018%2005%20RL%201-%20main.pdf

https://www3.nd.edu/~pantsakl/Archive/WolovichSymposium/files/Lewis_Presentation.pdf

Adaptive dynamic programming**

P solution of the Riccati equation

$$A^T P + PA - PBR_2^{-1}B^T P + Q = 0,$$

A unknown ! → Learning *P*

Integral reinforcement learning policy iteration algorithm (IRL-PIA):

Data-driven

$$x^T P_i x = \int_t^{t+T} x^T(\tau)(R_1 + K_i^T R_2 K_i)x(\tau)d\tau + x^T(t+T)P_i x(t+T),$$
$$K_{i+1} = R_2^{-1}B^T P_i, \; i = 1, 2, \ldots$$

where the initial gain K_1 is chosen such that is $A - BK_1$ stable.

Under conditions of stabilizability/detectability:

$$u^*(t) \to argmin_{u(t)} V(u), \; t \in [t_0, \infty[.$$

** Vrabie D, Vamvoudakis K, Lewis FL. Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles. England: IET Digital Library; 2013.

Control barrier function (CBF)-based learning control*, **

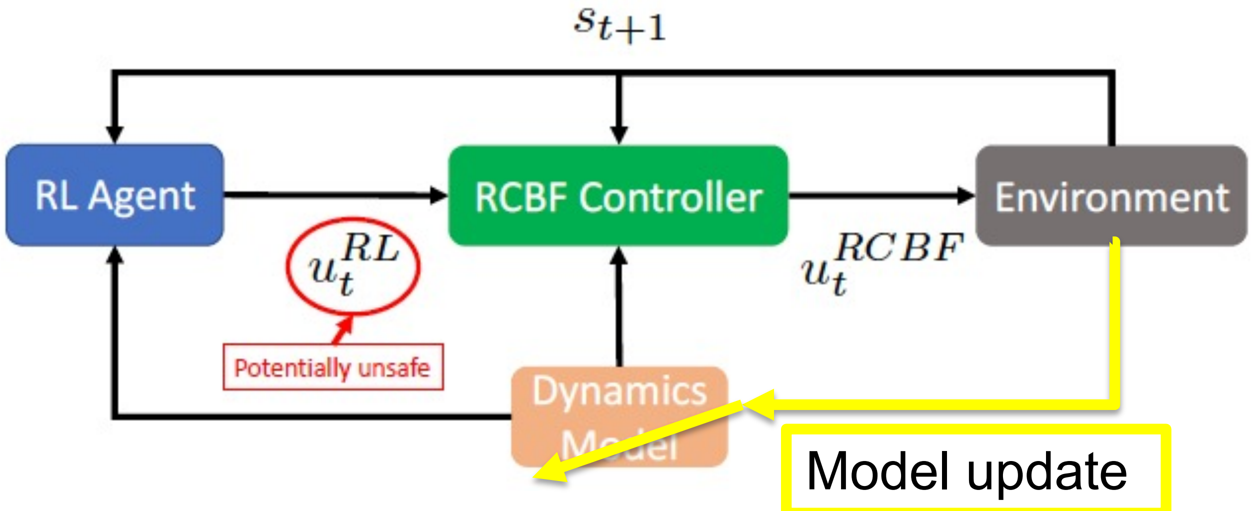


Figure from [*] with the addition of the yellow part

* Emam et al. 21, Safe Model-Based Reinforcement Learning using Robust Control Barrier Functions, arXiv:2110.05415v1.

** Xu X., et al. 2015, Robustness of Control Barrier Functions for Safety Critical Control. A. D *IFAC-PapersOnLine*, 48(27)

Control barrier function (CBF)-based learning control*,**

Model:
$$\dot{x}(t) = f(x(t)) + g(x(t))u(x(t)) + d(x(t)), d \in D$$
$$D(x') = \text{co}\,\Psi(x') = \text{co}\{\psi_1(x') \dots \psi_p(x')\}, \forall x' \in \mathbb{R}^n,$$

Policy:
$$u^{RCBF}(x') = u^*(x') + u^{RL}(x').$$

- Model-based
- Data-driven

Filter:
$$u^{RL}(x') \sim \pi_\phi(\cdot|x')$$
$$u^*(x') = \arg\min_{u \in \mathbb{R}^m} \|u\|^2 + l\epsilon^2$$
$$\text{s.t. } \nabla h(x')^\top (f(x') + g(x')(u(x') + u^{RL}(x'))) \geq -\alpha(h(x')) - \min \nabla h(x')^\top \Psi(x') + \epsilon$$

* Emam et al. 21, Safe Model-Based Reinforcement Learning using Robust Control Barrier Functions, arXiv:2110.05415v1.

** Xu X., et al. 2015, Robustness of Control Barrier Functions for Safety Critical Control. A. D *IFAC-PapersOnLine*, 48(27)

Main points

- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)*

- Learning-based iterative feedback gains tuning for nonlinear systems affine in the control (ESC)

- Indirect learning-based adaptive iterative control for linear systems under constraints (ESC-MPC framework)

- Learning-based adaptive PDE stable model reduction/ estimation (ESC, RL)

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, RL: Reinforcement learning

Learning-based iterative feedback gains auto-tuning for nonlinear systems *

$$\dot{x} = f(x) + \Delta f(x) + g(x)u, \; x(0) = x_0$$
$$y = h(x),$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^{n_a}, y \in \mathbb{R}^m \; (n_a \geq m)$.

Assumption 1: $f : \mathbb{R}^n \to \mathbb{R}^n$ and the columns of $g : \mathbb{R}^n \to \mathbb{R}^{n \times n_a}$ are $\mathbb{C}^\infty$ vector fields on a bounded set X of $\mathbb{R}^n$ and $h(x)$ is a $\mathbb{C}^\infty$ function on X. The vector field $\Delta f(x)$ is $\mathbb{C}^1$ on X.

* M. Benosman, 2016, Multi-Parametric Extremum Seeking-based Auto-Tuning for Robust Input-Output Linearization Control", International Journal of Robust and Nonlinear Control, 26(18), 4035-4055.

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Assumption 2: System (1) has a well-defined (vector) relative degree $\{r1, \ldots, rm\}$ at each point $x^0 \in X$, and the system is linearizable, i.e. $\sum_{i=1}^{i=m} ri = n$

Assumption 3: The uncertainty vector Δf is s.t. $|\Delta f(x)| \leq d(x)\ \forall x \in X$, where $d : X \to \mathbb{R}$ is a smooth nonnegative function.

MITSUBISHI ELECTRIC
Changes for the Better

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Assumption 4: The desired output trajectories y_{id} are smooth functions of time, relating desired initial points y_{i0} at $t = 0$ to desired final points y_{if} at $t = t_f$, and s.t. $y_{id}(t) = y_{if}$, $\forall t \geq t_f$, $t_f > 0$, $i \in \{1, ..., m\}$.

Control objectives

- uniform boundedness of a tracking error,
- feedback gains vector K is iteratively auto-tuned, to optimize a desired performance

Learning-based iterative feedback gains auto-tuning for nonlinear systems

CONTROLLER DESIGN: (using I/O linearization and Lyapunov reconstruction)

I/O linearization

Lyapunov reconstruction

Step one: Passive robust control design

$$u = A^{-1}(\xi)(v_s(t,\xi) - b(\xi)) - A^{-1}(\xi)\frac{\partial V}{\partial \tilde{z}}^T k\, d_2(e),\; k > 0,\quad v_s = (v_{s1}, \ldots, v_{sm})^T,$$

$$v_{si}(t,\xi) = y_{i_d}^{(ri)} - K_{ri}^i\left(y_i^{(ri-1)} - y_{i_d}^{(ri-1)}\right) - \ldots - K_1^i(y_i - y_{i_d}).$$

$$V = z^T P z,\; P > 0 \quad P\tilde{A} + \tilde{A}^T P = -I$$

$$z = (z^1, \ldots, z^m)^T,\; z^i = (e_i, \ldots, e_i^{r_i - 1}), \tilde{z} = (z^1(r_1), \ldots, z^m(r_m))^T \in \mathbb{R}^m$$

$$e_i(t) = y_i(t) - y_{i_d}(t)$$

$d_2(.)$ is an upper bound of the uncertainty

$\tilde{A}$ is a block diagonal matrix of the feedback gains

Learning-based iterative feedback gains auto-tuning for nonlinear systems

CONTROLLER DESIGN

Step two: *Iterative tuning of the feedback gains*

$$Q(z(\beta)) = \int_{(I-1)t_f}^{It_f} z^T(t) C_1 z(t) dt + \int_{(I-1)t_f}^{It_f} u^T(t) C_2 u(t) dt,$$

$$I = 1, 2, 3..., \; C_1, \; C_2 > 0$$

MITSUBISHI ELECTRIC
Changes for the Better

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Step two: *Iterative tuning of the feedback gains*

$$\beta = [\delta K_1^1, ..., \delta K_{r1}^1, ..., \delta K_1^m, ..., \delta K_{rm}^m, \delta k]^T$$

$$K_j^i = K_{j-nominal}^i + \delta K_j^i, \; j = 1, ...ri, \; i = 1, ..., m.$$
$$k = k_{nominal} + \delta k, \;\; k_{nominal} > 0$$

$$\dot{x}_{K_j^i} = a_{K_j^i} sin(\omega_{K_j^i} t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{K}_j^i(t) = x_{K_j^i}(t) + a_{K_j^i} sin(\omega_{K_j^i} t + \tfrac{\pi}{2}), \; j = 1, ...ri, \; i = 1, ..., m$$
$$\dot{x}_k = a_k sin(\omega_k t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{k}(t) = x_k(t) + a_k sin(\omega_k t + \tfrac{\pi}{2}),$$

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Step two: *Iterative tuning of the feedback gains*

$$\omega_1 + \omega_2 \neq \omega_3, \text{ for } \omega_1 \neq \omega_2 \neq \omega_3,$$
$$\forall \omega_1, \omega_2, \omega_3 \in \{\omega_{K^i_j}, \omega_k,\ j = 1, ...ri,\ i = 1, ..., m\},$$

with $\omega_i > \omega^*$, $\forall \omega_i \in \{\omega_{K^i_j}, \omega_k,\ j = 1, ...ri,\ i = 1, ..., m\}$, ω^* large enough.

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Step two: *Iterative tuning of the feedback gains*

Assumption 7: We assume that the cost function Q has a local minimum at β^*.

Assumption 8: We consider that the initial gain vector β is sufficiently close to the optimal gain vector β^*.

Assumption 9: The cost function is analytic and its variation with respect to the gains is bounded in the neighborhood of β^*, i.e. $|\frac{\partial Q}{\partial \beta}(\tilde{\beta})| \leq \Theta_2, \ \Theta_2 > 0, \ \tilde{\beta} \in \mathcal{V}(\beta^*)$, where $\mathcal{V}(\beta^*)$ denotes a compact neighborhood of β^*.

Learning-based iterative feedback gains auto-tuning for nonlinear systems

Put together: Robust controller + ESC tuning

Model-based

$$u = A^{-1}(\xi)(v_s(t,\xi) - b(\xi)) - A^{-1}(\xi)\frac{\partial V}{\partial \tilde{z}}^T k(t)\, d_2(e), \quad k > 0, \quad v_s = (v_{s1}, ..., v_{sm})^T$$

$$v_{si}(t,\xi) = \hat{y}_{i\,d}^{(ri)} - K^i_{ri}(t)\left(y_i^{(ri-1)} - \hat{y}_{i\,d}^{(ri-1)}\right) - ... - K^i_1(t)(y_i - \hat{y}_{i\,d}),\ i = 1, ..., m.$$

$$\hat{y}_{i\,d}(t) = y_{id}(t-(I-1)t_f),\ (I-1)t_f \le t < It_f,\ I \in \{1,2,...\},$$

Data-driven

$$K^i_j(t) = K^i_{j-nominal} + \delta K^i_j(t)$$

$$\delta K^i_j(t) = \delta\hat{K}^i_j((I-1)t_f),\ (I-1)t_f \le t < It_f,$$

$$k(t) = k_{nominal} + \delta k(t),\quad k_{nominal} > 0$$

$$\delta k(t) = \delta\hat{k}((I-1)t_f),\ (I-1)t_f \le t < It_f,\ I = 1,2,3...$$

$\hat{K}^i_j, \delta\hat{k}$ are estimated by the MES algorithm.

Learning-based iterative feedback gains auto-tuning for nonlinear systems

- the obtained closed-loop impulsive time-dependent dynamic system is well posed.
- the tracking error z is uniformly bounded.
- z is steered at each iteration I towards the positive invariant set $S_I = \{z \in \mathbb{R}^n | \ 1 - k_I \ |\frac{\partial V}{\partial z}_{ind}| \geq 0\}$
  $k_I = \beta_I(n+1)$
- $|Q(\beta(It_f)) - Q(\beta^*)| \leq \Theta_2(\frac{\Theta_1}{\omega_0} + \sqrt{\sum_{i=1,\ldots,m \ j=1,\ldots,ri} {a_{K_j^i}}^2 + {a_k}^2})$
  $\Theta_1, \Theta_2 > 0, \quad \text{for} \quad I \rightarrow \infty, \omega_0 = Max(\omega_{K_1^1}, \ldots, \omega_{K_{rm}^m}, \omega_k)$

Learning-based iterative feedback gains auto-tuning for nonlinear systems

– β remains bounded over the iterations s.t.

$$|\beta((I+1)t_f) - \beta(It_f)| \leq 0.5 t_f Max(a_{K_1^1}{}^2, ..., a_{K_{rm}^m}{}^2, a_k^2)\Theta_2 + t_f \omega_0 \sqrt{\sum_{i=1,\dots,m\ j=1,\dots,ri} a_{K_j^i}{}^2 + a_k{}^2}, \quad I \in \{1, 2, ...\}$$

– satisfies asymptotically the bound

$$|\beta(It_f) - \beta^*| \leq \frac{\Theta_1}{\omega_0} + \sqrt{\sum_{i=1,\dots,m\ j=1,\dots,ri} a_{K_j^i}{}^2 + a_k{}^2}, \ \Theta_1 > 0, \ \text{for } I \to \infty$$

Main points

- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)*

- Learning-based iterative feedback gains tuning for nonlinear systems affine in the control (ESC)

- Indirect learning-based adaptive iterative control for linear systems under constraints (ESC-MPC framework)

- Learning-based adaptive PDE stable model reduction/ estimation (ESC, RL)

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, RL: Reinforcement learning

Indirect learning-based adaptive iterative control for linear systems under constraints (MPC framework) *, **

$$\begin{aligned} x(k+1) &= (A+\Delta A)x(k) + (B+\Delta B)u(k) \\ y(k) &= (C+\Delta C)x(k) + (D+\Delta D)u(k), \end{aligned}$$

$$x \in \mathbb{R}^n,\ u \in \mathbb{R}^m,\ y \in \mathbb{R}^p$$

$$x_{\min} \leq x(k) \leq x_{\max},$$

$$u_{\min} \leq u(k) \leq u_{\max},$$

$$y_{\min} \leq y(k) \leq y_{\max},$$

$$r_r(k+1) = A_r r_r(k),\ \ y_e(k) = Cx(k) - C_r r_r(k),$$

* Benosman M., Di Cairano S., Weiss A., 2014, Extremum seeking-based iterative learning linear MPC, IEEE Conference on Control Applications (prelim. idea no proofs)

** Subbaraman S., Benosman M., 2016, Extremum Seeking-based Iterative Learning Model Predictive Control (ESILC-MPC), IFAC International Workshop on Adaptation and Learning in Control and Signal Processing (follow up paper with convergence proofs).

MITSUBISHI ELECTRIC
Changes for the Better

Indirect learning-based adaptive iterative control for linear systems under constraints (MPC framework) *, **

Assumption 1: The constant uncertainty matrices ΔA, ΔB, ΔC and ΔD, are bounded, s.t. $||\Delta A||_2 \leq l_A$, $||\Delta B||_2 \leq l_B$, $||\Delta C||_2 \leq l_C$, $||\Delta D||_2 \leq l_D$, with $l_A, l_B, l_C, l_D > 0$.

Assumption 2: There exists non empty convex sets $\mathcal{K}_a \subset \mathbb{R}^{n\times n}$, $\mathcal{K}_b \subset \mathbb{R}^{n\times m}$, $\mathcal{K}_c \subset \mathbb{R}^{p\times n}$, and $\mathcal{K}_d \subset \mathbb{R}^{p\times m}$, such that $A + \Delta A \in \mathcal{K}_a$ for all ΔA such that $||\Delta A||_2 \leq l_A$, $B + \Delta B \in \mathcal{K}_b$ for all ΔB such that $||\Delta B||_2 \leq l_B$, $C + \Delta C \in \mathcal{K}_c$ for all ΔC such that $||\Delta C||_2 \leq l_C$, $D + \Delta D \in \mathcal{K}_d$ for all ΔD such that $||\Delta D||_2 \leq l_D$,.

Assumption 3: The iterative learning MPC problem (and the associated reference tracking extension), is a well-posed optimization problem for any matrices $A + \Delta A \in \mathcal{K}_a$, $B + \Delta B \in \mathcal{K}_b$, $C + \Delta C \in \mathcal{K}_c$, $D + \Delta D \in \mathcal{K}_d$.

MITSUBISHI ELECTRIC
Changes for the Better

Indirect learning-based adaptive iterative control for linear systems under constraints (MPC framework) *, **

$$Q(\hat{\Delta}) = F(y_e(\hat{\Delta})),$$ different from the MPC cost

where $\hat{\Delta}$ is the vector obtained by concatenating the estimated uncertainty matrices $\Delta\hat{A}$, $\Delta\hat{B}$, $\Delta\hat{C}$ and $\Delta\hat{D}$,

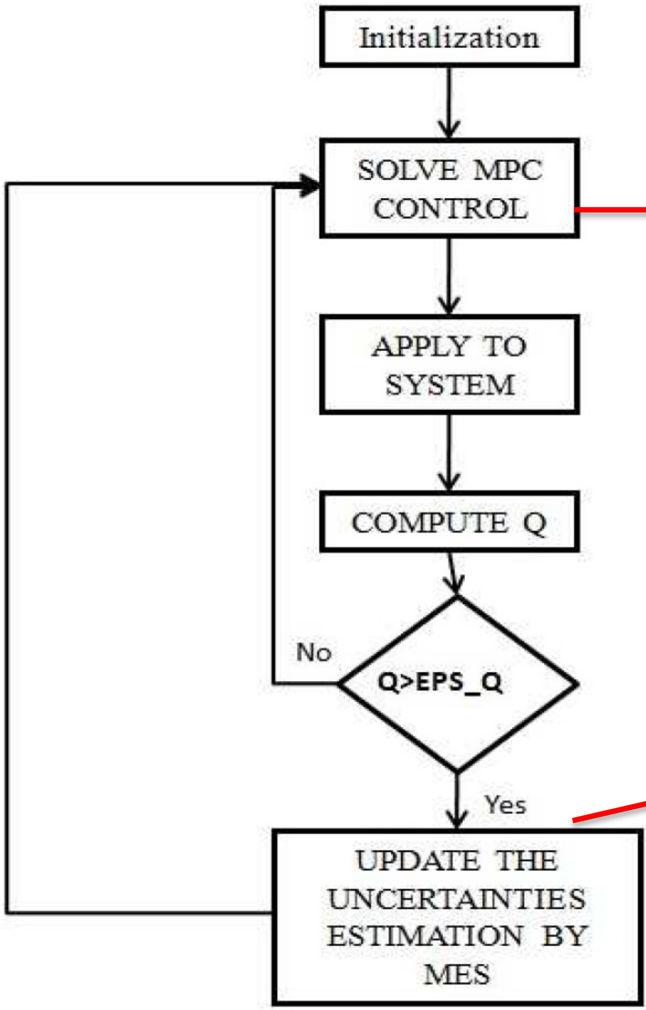
$$F : \mathbb{R}^p \to \mathbb{R},\ F(0) = 0,\ F(y_e) > 0 \text{ for } y_e \neq 0.$$

Assumption 4: The cost function Q has a local minimum at $\hat{\Delta}^* = \Delta$.

Assumption 5: The original parameter estimate vector $\hat{\Delta}$ is close enough to the actual parameters vector Δ.

Assumption 6: The cost function is analytic and its variation with respect to the uncertain variables is bounded in the neighborhood of Δ^*, i.e., there exists $\xi_2 > 0$, s.t. $\|\frac{\partial Q}{\partial \Delta}(\tilde{\Delta})\| \leq \xi_2$ for all $\tilde{\Delta} \in \mathcal{V}(\Delta^*)$, where $\mathcal{V}(\Delta^*)$ denotes a compact neighborhood of Δ^*.

Indirect learning-based adaptive iterative control for linear systems under constraints (MPC framework) *, **



Model-based

ISS -MPC *

Data-driven

$$\dot{z}_i = a_i \sin(\omega_i t + \tfrac{\pi}{2}) Q(\hat{\Delta})$$
$$\hat{\Delta}_i = z_i + a_i \sin(\omega_i t - \tfrac{\pi}{2}),\ i \in \{1, \dots, N_p\}$$

with $N_p \leq nn + nm + pn + pm$ is the number of uncertain elements, $\omega_i \neq \omega_j$, $\omega_i + \omega_j \neq \omega_k$, $i, j, k \in \{1, \dots, N_p\}$, and $\omega_i > \omega^*$, $\forall i \in \{1, \dots, N_p\}$, with ω^* large enough, converges to the local minima of Q.

* D. Limon, I. Alvarado, T. Alamo, and E. Camacho, "Robust tube-based MPC for tracking of constrained linear systems with additive disturbances," Journal of Process Control, vol. 20, no. 3, pp. 248–260, 2010.

Main points

- Learning-based adaptive control for nonlinear systems with constant/time-varying parametric uncertainties (ESC, GP-UCB, ADP, CBF)*

- Learning-based iterative feedback gains tuning for nonlinear systems affine in the control (ESC)

- Indirect learning-based adaptive iterative control for linear systems under constraints (ESC-MPC framework)

- Learning-based adaptive PDE stable model reduction/ estimation (ESC, RL)

* ESC: Extremum seeking control, GP-UCB: Gaussian process upper confidence bound, ADP: Adaptive dynamic programming, CBF: Control barrier function, RL: Reinforcement learning

Learning-based PDE stable model reduction *

Consider a stable dynamical system modeled by a nonlinear partial differential equation of the form

$$\dot{z} = \mathcal{F}(z, \mu) \in \mathcal{Z}, \mu \in \mathbb{R}$$

PDE

μ denotes a viscosity coefficient.

where $\mathcal{Z}$ is an infinite-dimension Hilbert space.

$$P_n z(t, x) \approx \Phi z_r(t) = \sum_{i=1}^{r} z_{ri}(t)\phi_i(x) \in \mathbb{R}^n$$

where P_n is the projection of $z(t, x)$ onto $\mathbb{R}^n$.

$$\dot{z}_r(t) = F(z_r(t), \mu)$$

ODE

The function $F: \mathbb{R}^r \to \mathbb{R}^r$ is obtained from the weak form of the original PDE (through Galerkin projection).

* Benosman M., Borggaard J., San O., Kramer B., 2017, Learning-based robust stabilization for reduced order models of 2D and 3D Boussinesq equations, Applied Mathematical Modelling, Vol. 49, 162-181.

Changes for the Better

The Closure-Model Concept for ROMs Stabilization

PDE → *e.g., POD*

$$\begin{cases} \dot{q}(t) = F(q(t), \mu), \\ z_n^{pod}(t,x) = \sum_{i=1}^{i=r} \phi_i(x) q_i(t). \end{cases}$$

Loss of stability (in the sense of Lagrange)

↓ *Closure Model H*

$$\dot{q}(t) = F(q(t), \mu) + H(t, q(t)).$$

We try to recover the stability

1) Closure models with constant eddy viscosity coefficients:

μ is substituted by a virtual viscosity coefficient μ_{cl}, $\mu_{cl} = \mu + \mu_e$, Heisenberg ROM

2) Closure models with time and space varying eddy viscosity coefficients:

$$H_{nev}(\mu_e, q(t)) = \mu_e \sqrt{\frac{V(q(t))}{V_\infty(\lambda)}} diag(d_{11}, ..., d_{rr}) q(t), \quad V(q) = \frac{1}{2}\sum_{i=1}^{i=r} {q_i}^2, \quad V_\infty(\lambda) = \frac{1}{2}\sum_{i=1}^{i=r} \lambda_i,$$

the λ_i are the selected POD eigenvalues

where $D \in \mathbb{R}^{r \times r}$ represents a constant viscosity damping matrix,

MITSUBISHI ELECTRIC
Changes for the Better

A Lyapunov-based closure-Model for Robust ROMs

Stabilization

original PDE ⟶ ***Using POD*** ⟶ $\begin{cases} \dot{q}^{pod}(t) = F(q^{pod}(t), \mu) = \tilde{F}(q^{pod}(t)) + \mu\, D q^{pod} \\ z^{pod}(t,x) = \sum_{i=1}^{i=N_{pod}} \phi_i^{pod}(x) q_i^{pod}(t), \end{cases}$ $, D < 0$

Assumption 1 *The norm of the vector field $\tilde{F}$ is bounded by a known function of q^{pod}, i.e., $\|\tilde{F}(q^{pod})\| \leq \tilde{f}(q^{pod})$.*

Assumption 2 *The solutions of the original PDE model are assumed to be in $L^2([0,\infty); \mathcal{Z})$.*

Model-based

Then, the nonlinear closure model

$$H_{nl} = \mu_{nl} \tilde{f}(q^{pod}) diag(d_{11}, ..., d_{N_{pod}N_{pod}}) q^{pod}, \;\; \mu_{nl} > 0$$

stabilizes the solutions of the ROM to the invariant set

$$\mathcal{S} = \{q^{pod} \in \mathbb{R}^{N_{pod}} \; s.t. \; \mu \frac{\lambda(D)_{max} \|q^{pod}\|}{\tilde{f}} + \mu_{nl} \|q^{pod}\| Max(d_{11}, ..., d_{N_{pod}N_{pod}}) + 1 \geq 0\}.$$

MITSUBISHI ELECTRIC
Changes for the Better

An Extremum Seeking–based Auto-Tuning of Closure Models for ROMs

μ is substituted by a virtual viscosity coefficient μ_{cl}. $\mu_{cl} = \mu + \mu_e$, Heisenberg ROM

if the closure models amplitudes μ_e, μ_{nl} *are tuned using the MES algorithm*

Data-driven

$$\begin{aligned}
\dot{y}_1 &= a_1\, sin(\omega_1 t + \tfrac{\pi}{2}) Q(\hat{\mu}_e, \hat{\mu}_{nl}) \\
\hat{\mu}_e &= y_1 + a_1\, sin(\omega_1 t - \tfrac{\pi}{2}) \\
\dot{y}_2 &= a_2\, sin(\omega_2 t + \tfrac{\pi}{2}) Q(\hat{\mu}_e, \hat{\mu}_{nl}) \\
\hat{\mu}_{nl} &= y_2 + a_2\, sin(\omega_2 t - \tfrac{\pi}{2}),
\end{aligned}$$

where $\omega_{max} = max(\omega_1, \omega_2) > \omega^*$, ω^* *large enough, and* Q *the learning cost function*

$$Q(\hat{\mu}) = H(e_z(\hat{\mu})),\ \hat{\mu} = (\hat{\mu}_e, \hat{\mu}_{nl})$$

$$e_z(t) = z^{pod}(t, x) - z(t, x),\ H \text{ is a positive definite function of } e_z$$

Assumption 3 *The learning cost function* Q *has a local minimum at* $\hat{\mu} = \mu^*$.

Assumption 4 *The learning cost function* Q *is analytic and its variation with respect to* μ *is bounded in the neighborhood of* μ^*, *i.e.,* $\|\frac{\partial Q}{\partial p}(\tilde{\mu})\| \leq \xi_2$, $\xi_2 > 0$, $\tilde{\mu} \in \mathcal{V}(\mu^*)$, *where* $\mathcal{V}(\mu^*)$ *denotes a compact neighborhood of* μ^*.

An Extremum Seeking–based Auto-Tuning of Closure Models for ROMs

Then, the norm of the vector

$$e_\mu = ({\mu_e}^* - \hat{\mu}_e(t), {\mu_{nl}}^* - \hat{\mu}_{nl}(t))$$

admits the following bound

$$\|e_\mu(t)\| \leq \frac{\xi_1}{\omega_{max}} + \sqrt{a_1^2 + a_2^2},\ t \to \infty$$

where $a_1,\ a_2 > 0,\ \xi_1 > 0,$ *and the learning cost function approaches its optimal value within the following upper-bound*

$$\|Q(\hat{\mu}_e, \hat{\mu}_{nl}) - Q({\mu_e}^*, {\mu_{nl}}^*)\| \leq \xi_2(\frac{\xi_1}{\omega} + \sqrt{a_1^2 + a_2^2}),\ t \to \infty$$

where $\xi_2 = max_{(\mu_1,\mu_2)} \in \mathcal{V}(\mu^*)|\frac{\partial Q}{\partial \mu}|$.

Learning-based observers

Slide from: Mowlavi S.@MIT, presentation at ICLR 21.

RL-based observer *, **

Full-order model

$$\mathbf{z}_{k+1} = \mathbf{f}(\mathbf{z}_k) \qquad \text{full state } \mathbf{z}_k \in \mathbb{R}^n$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{z}_k \qquad \text{measurement } \mathbf{y}_k \in \mathbb{R}^p$$

$n \gg 1$

Reduced-order model (ROM)

$$\mathbf{x}_{k+1} = \mathbf{A}_r\mathbf{x}_k \qquad \text{reduced state } \mathbf{x}_k \in \mathbb{R}^r$$
$$\mathbf{y}_k = \mathbf{C}_r\mathbf{x}_k \qquad \text{measurement } \mathbf{y}_k \in \mathbb{R}^p$$

Model-based

$r \ll n$

$\mathbf{x}_k = \mathbf{U}^{\mathsf{T}}\mathbf{z}_k,$ dimensionality reduction

$\mathbf{z}_{k+1} \simeq \mathbf{A}\mathbf{z}_k$ dynamics approximation

How to do the data assimilation?

Kalman filter
(conventional approach)

$$\hat{\mathbf{x}}_k = \mathbf{A}_r\hat{\mathbf{x}}_{k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}_r\mathbf{A}_r\hat{\mathbf{x}}_{k-1})$$

Challenge: performs poorly when $\mathbf{A}_r$ is not a good model

Reinforcement learning-trained filter
(what we propose)

Data-driven

$$\hat{\mathbf{x}}_k = \mathbf{A}_r\hat{\mathbf{x}}_{k-1} + \mathbf{a}_k \quad \text{where} \quad \mathbf{a}_k \sim \boldsymbol{\pi}_\theta(\,\cdot\,|\,\mathbf{y}_k, \hat{\mathbf{x}}_{k-1})$$

Flexibility of nonlinear policy $\boldsymbol{\pi}_\theta$ allows to compensate for errors in $\mathbf{A}_r$

* Mowlavi S., et al., 2021, Reinforcement Learning State Estimation for High-Dimensional Nonlinear Systems, ICLR Workshop: AI for Earth and Space Science.

** Benosman et al., 2020, Reinforcement Learning-based Model Reduction for Partial Differential Equations, World Congress of the International Federation of Automatic Control (IFAC).

Open theoretical problems ?

- Robustness to hyper-parameters tuning
- Large scale systems and high dimensional systems, e.g., PDE models, delays
- Robustness and safety (state/input constraints) of ML algorithms from control theory perspective (e.g., stability and robustness of (CS-)RL algorithms using dynamical systems theory tools, neural ODEs from dynamical systems perspective (useful/scalable ?))
- Sampling efficiency/data constraints
- Real-time computational constraints
- …

* CS: Computer science, RL: Reinforcement learning, ODEs: Ordinary diff. equations, PDEs: Partial diff. equations

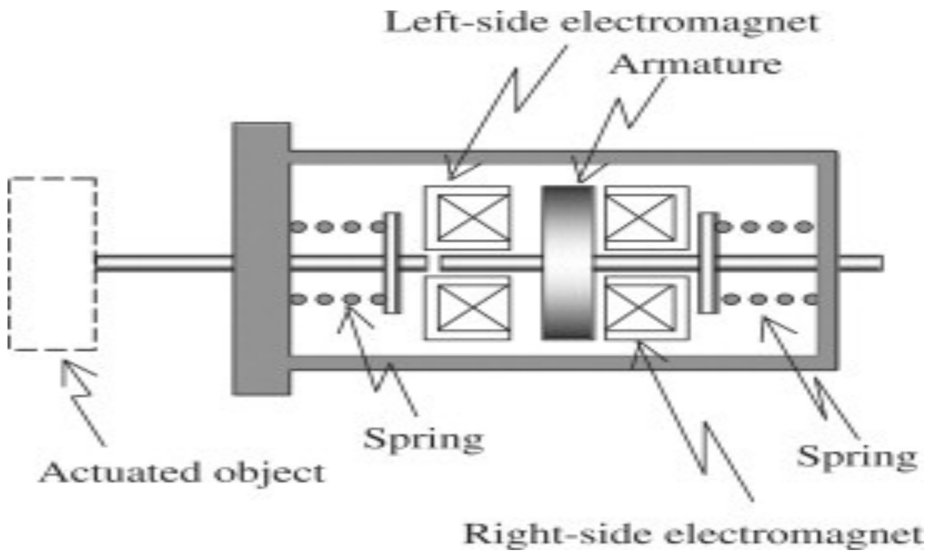
A hybrid approach to control: classical control theory meets data-driven methods

Mouhacine Benosman

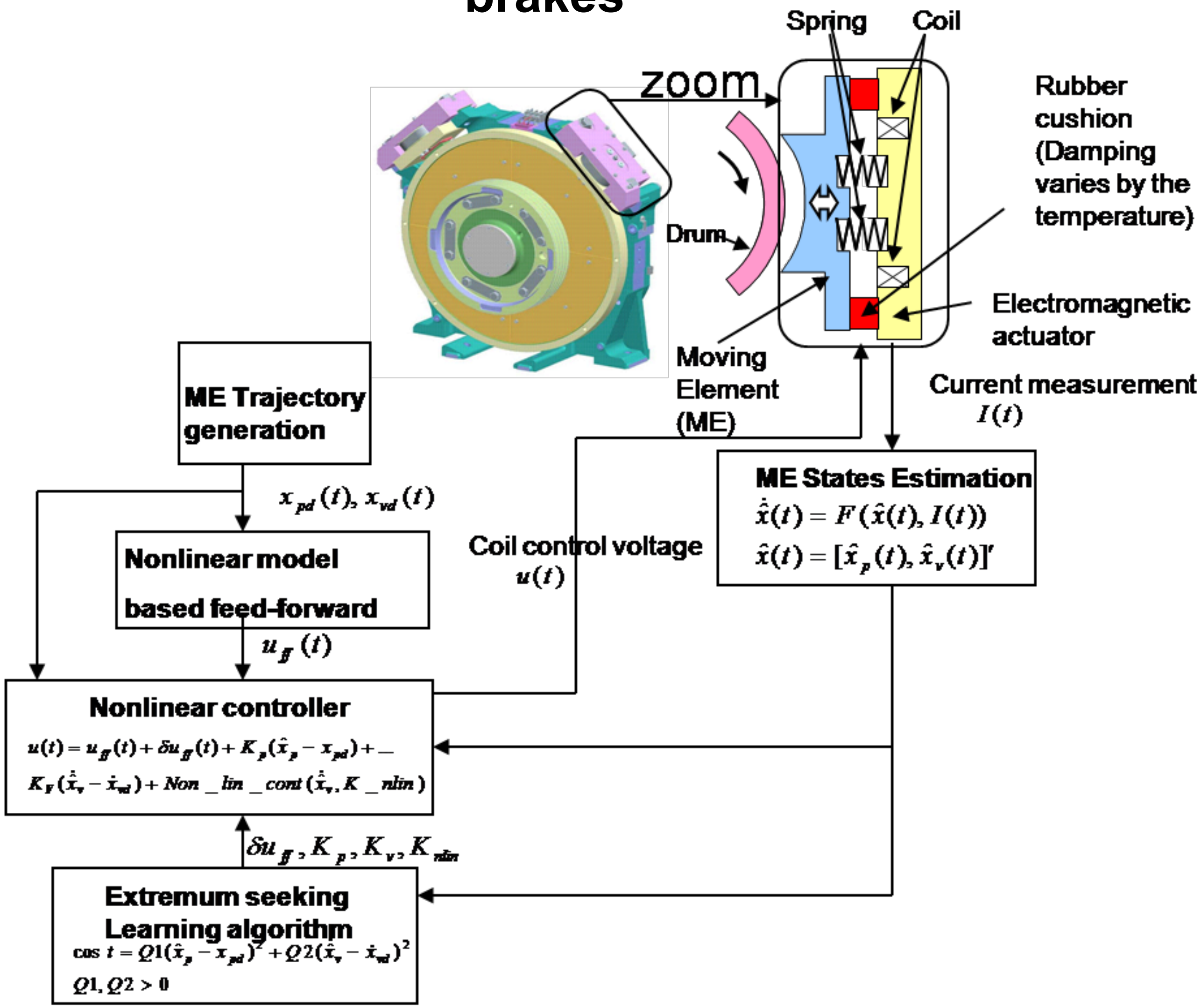
MERL - Mitsubishi Electric Research Labs, Cambridge, USA

Part II: Examples

Mechatronics Examples: Electromagnetic brakes*

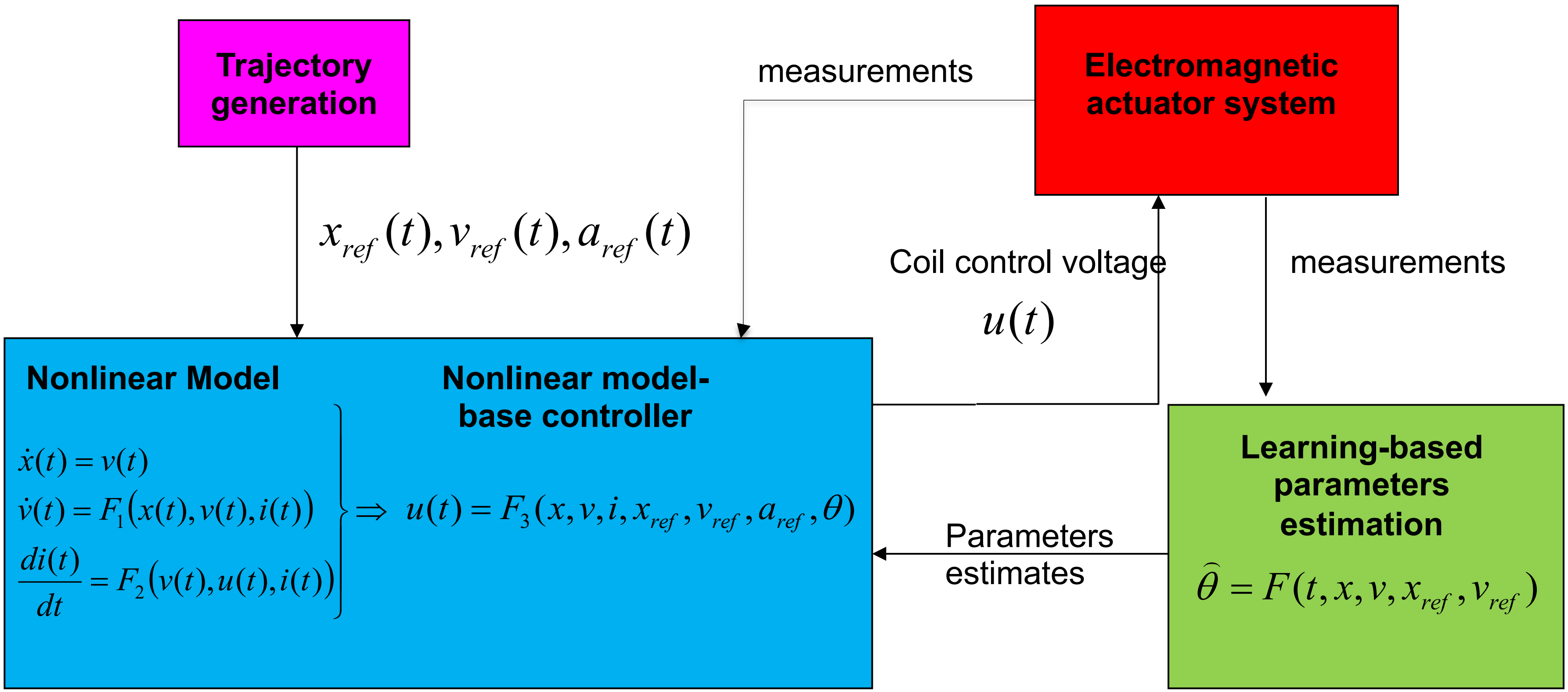


Mechatronics Examples: Electromagnetic brakes*



* Benosman M., Atink G., 2015, Extremum seeking-based nonlinear control for electromagnetic actuators, International Journal of Control, 88(3), 517-530.

Mechatronics Examples: Electromagnetic brakes



Mechatronics Examples: Electromagnetic brakes

- **Mechanical part**

$$m\ddot{x} = k(x_0 - x) - \underbrace{0.5\frac{a}{(b+x)^2}i^2}_{\text{EMF}} - \eta\dot{x} - f_d$$

- **Electrical part**

$$u = Ri + L(x)\frac{di}{dt} - \underbrace{\frac{a}{(a+x)^2}i\frac{dx}{dt}}_{\text{Back\_EMF}}, \quad L(x) = \frac{a}{b+x}$$

Mechatronics Examples: Electromagnetic brakes

$$m\frac{d^2x}{dt^2} = k(x_0 - x) - \eta\frac{dx}{dt} - \frac{ai^2}{2(b+x)^2} + f_d$$

$$u = Ri + \frac{a}{b+x}\frac{di}{dt} - \frac{ai}{(b+x)^2}\frac{dx}{dt}, \; 0 \le x \le x_f,$$

$$\mathbf{z} := [z_1 \quad z_2 \quad z_3]^T = [x \quad \dot{x} \quad i]^T$$

$$\begin{aligned}
\dot{z}_1 &= z_2 \\
\dot{z}_2 &= \frac{k}{m}(x_0 - z_1) - \frac{\eta}{m}z_2 - \frac{a}{2m(b+z_1)^2}z_3^2 + \frac{f_d}{m} \\
\dot{z}_3 &= -\frac{R}{\frac{a}{b+z_1}}z_3 + \frac{z_3}{b+z_1}z_2 + \frac{u}{\frac{a}{b+z_1}}.
\end{aligned} \tag{2}$$

$$z_1^{ref}(t_0) = z_{1_{int}}, \quad z_1^{ref}(t_f) = z_{1_f},$$

$$\dot{z}_1^{ref}(t_0) = \dot{z}_1^{ref}(t_f) = 0,$$

$$\ddot{z}_1^{ref}(t_0) = \ddot{z}_1^{ref}(t_f) = 0,$$

Mechatronics Examples: Electromagnetic brakes

- assuming uncertainties
  - the spring constant k
  - damping coefficient η
  - the additive disturbance f_d
- u_{iss} Based on (i)ISS back-stepping approach
- the cost function

$$Q(\hat{\Delta}) = \int_0^{t_f} q_1(z_1(s)-z_1(s)^{ref})^2 ds + \int_0^{t_f} q_2(z_2(s)-z_2^{ref}(s))^2 ds$$

$q_1, q_2 > 0$

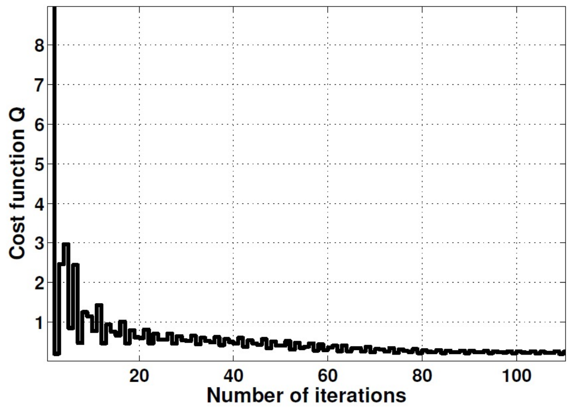
Mechatronics Examples: Electromagnetic brakes

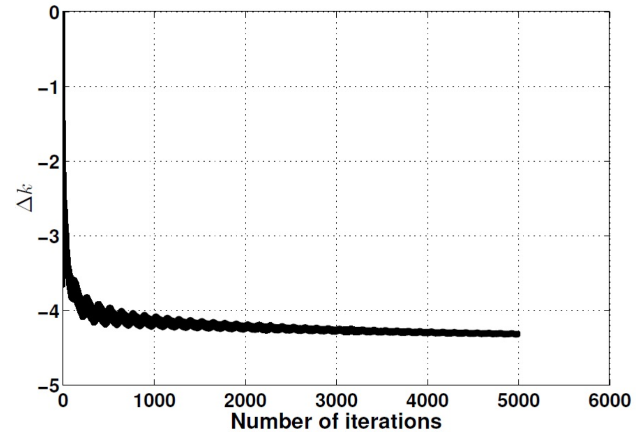
$$\hat{k}(t) = k_{nominal} + \hat{\Delta}_k(t)$$
$$\hat{\eta}(t) = \eta_{nominal} + \hat{\Delta}_\eta(t)$$
$$\hat{f}_d(t) = f_{d-nominal} + \hat{\Delta}_{f_d}(t)$$

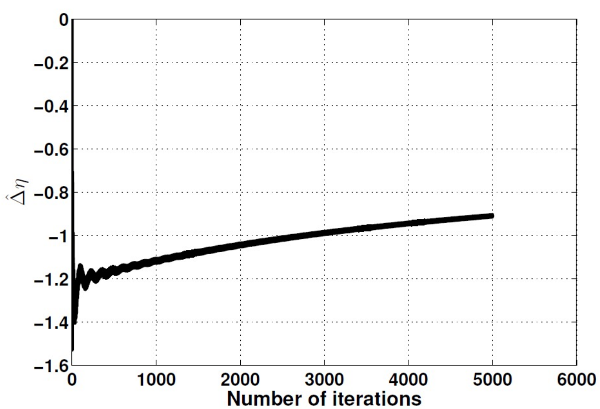
$$x_k(k' + 1) = x_k(k') + a_k t_f sin(\omega_k k' t_f + \frac{\pi}{2})Q$$
$$\hat{\Delta}_k(k' + 1) = x_k(k' + 1) + a_k sin(\omega_k k' t_f - \frac{\pi}{2}),$$
$$x_\eta(k' + 1) = x_\eta(k') + a_\eta t_f sin(\omega_\eta k' t_f + \frac{\pi}{2})Q$$
$$\hat{\Delta}_\eta(k' + 1) = x_\eta(k' + 1) + a_\eta sin(\omega_\eta k' t_f - \frac{\pi}{2}),$$
$$x_{f_d}(k' + 1) = x_{f_d}(k') + a_{f_d} t_f sin(\omega_{f_d} k' t_f + \frac{\pi}{2})Q$$
$$\hat{\Delta}_{f_d}(k' + 1) = x_{f_d}(k' + 1) + a_{f_d} sin(\omega_{f_d} k' t_f - \frac{\pi}{2})$$

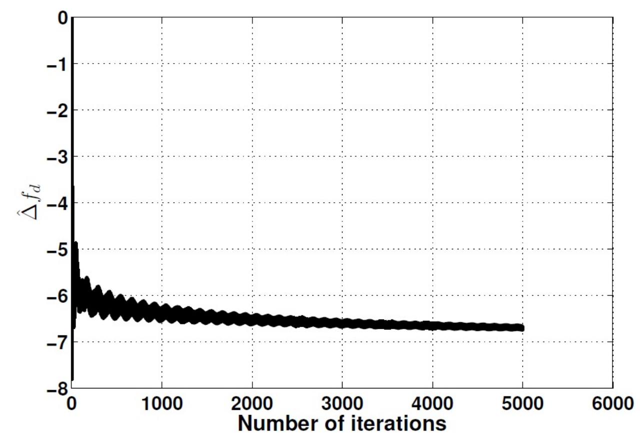
Mechatronics Examples: Electromagnetic brakes

$$\Delta k = -4.5, \ \Delta \eta = -0.7 \ \ \Delta f_d = -7.5$$



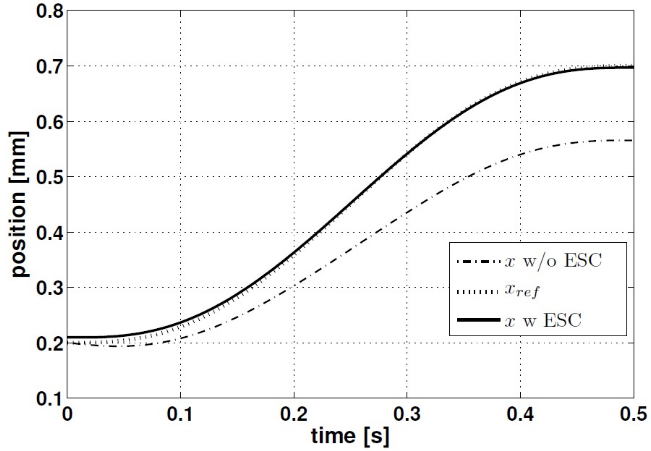


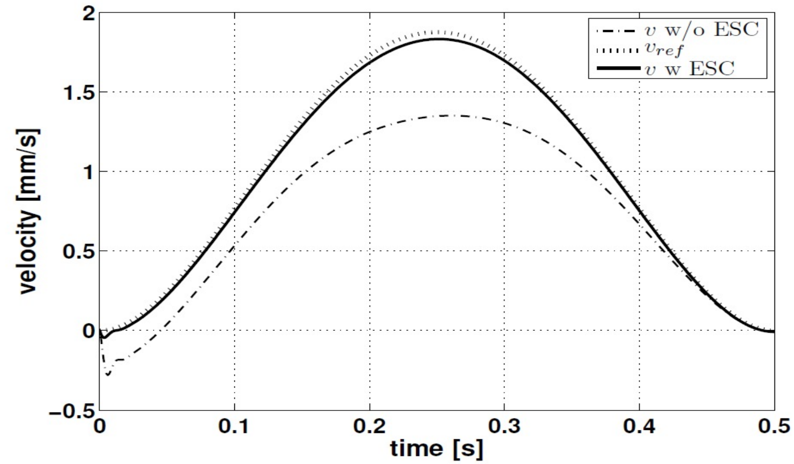


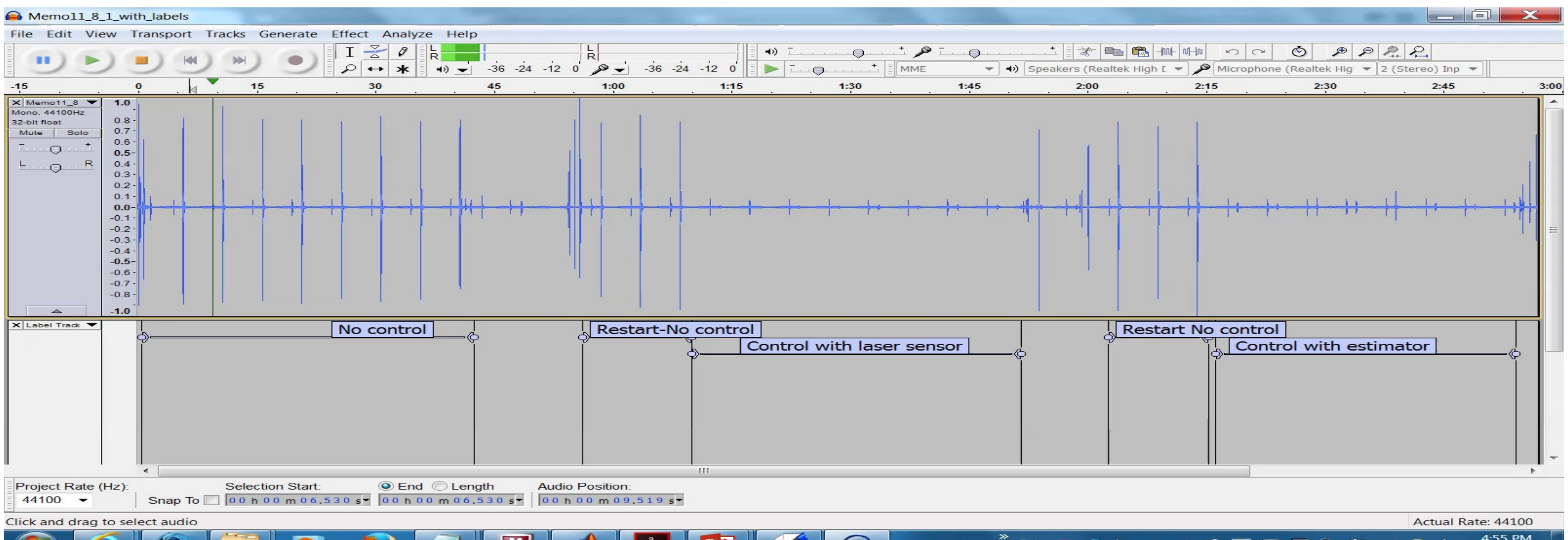


Mechatronics Examples: Electromagnetic brakes

$$\Delta k = -4.5, \; \Delta \eta = -0.7 \quad \Delta f_d = -7.5$$







Mechatronics Examples: Electromagnetic brakes

$$m\frac{d^2x_a}{dt^2} = k(x_0 - x_a) - \eta\frac{dx_a}{dt} - \frac{ai^2}{2(b+x_a)^2}$$
$$u = Ri + \frac{a}{b+x_a}\frac{di}{dt} - \frac{ai}{(b+x_a)^2}\frac{dx_a}{dt},\ 0 \leq x_a \leq x_f,$$

x_{ref} a desired armature position trajectory, s.t.

$$x_{ref}(0) = 0,\ x_{ref}(t_f) = x_f,\ \dot{x}_{ref}(0) = 0,\ \dot{x}_{ref}(t_f) = 0.$$

bounded parametric uncertainties

- spring coefficient $k = k_{nominal} + \delta k, |\delta k| \leq \delta k_{max}$
- the damping coefficient $\eta = \eta_{nominal} + \delta\eta,\ |\delta\eta| \leq \delta\eta_{max}$

Mechatronics Examples: Electromagnetic brakes

Passive robust controller:

$$u = -\frac{m(b+x_a)}{i}\left(v_s + \frac{k_{nominal}}{m}\dot{x}_a + \frac{\eta_{nominal}}{m}\ddot{x}_a - \frac{Ri^2}{(b+x_a)m}\right) + \frac{m(b+x_a)}{i}\frac{\partial V}{\partial z_3}k\left(\frac{\delta k_{max}}{m}|\dot{x}_a| + \frac{\delta \eta_{max}}{m}|\ddot{x}_a|\right),\ k>0$$

$$v_s = x_{ref}^{(3)}(t) + K_3(x_a^{(2)} - x_{ref}^{(2)}(t)) + K_2(x_a^{(1)} - x_{ref}^{(1)}(t)) + K_1(x_a - x_{ref}(t)),\ K_i < 0, i = 1,2,3.$$

$$V = z^T P z,\ P > 0 \qquad P\tilde{\tilde{A}} + \tilde{\tilde{A}}^T P = -I,$$

$$\tilde{\tilde{A}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ K_1 & K_2 & K_3 \end{pmatrix},$$

where $K_1,\ K_2,\ K_3$ are chosen such that $\tilde{\tilde{A}}$ is Hurwitz.

Mechatronics Examples: Electromagnetic brakes

Learning-based auto-tuning of the controller gains:

$$Q(z(\beta)) = C_1 z_1 (It_f)^2 + C_2 z_2 (It_f)^2 + C_3 z_3 (It_f)^2,$$

$I = 1, 2, 3...$ is the number of iterations, $C_1,\ C_2 > 0,\ C_3 > 0,$

$$\beta = (\delta K_1,\ \delta K_2,\ \delta K_3,\ \delta k)',$$

$$K_1 = K_{1_{nominal}} + \delta K_1$$

$$K_2 = K_{2_{nominal}} + \delta K_2$$

$$K_3 = K_{3_{nominal}} + \delta K_3$$

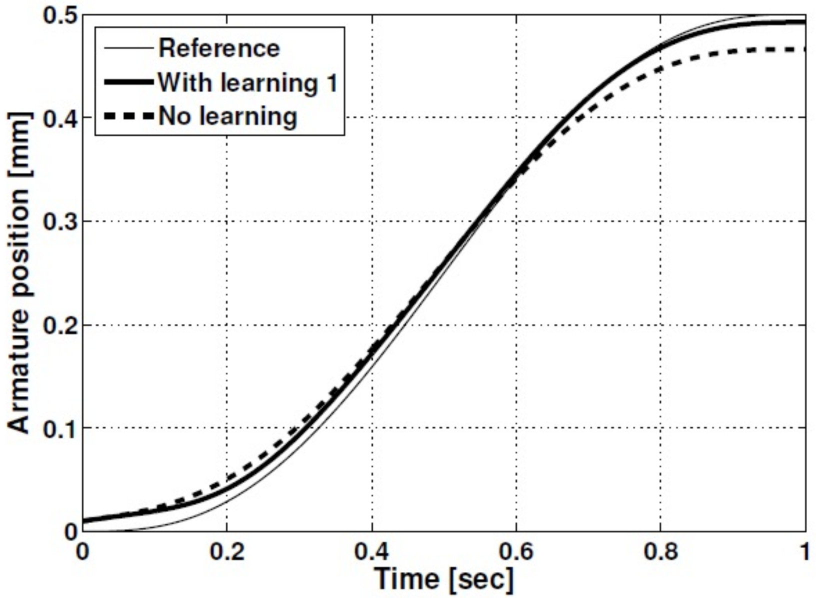
$$k = k_{nominal} + \delta k,$$

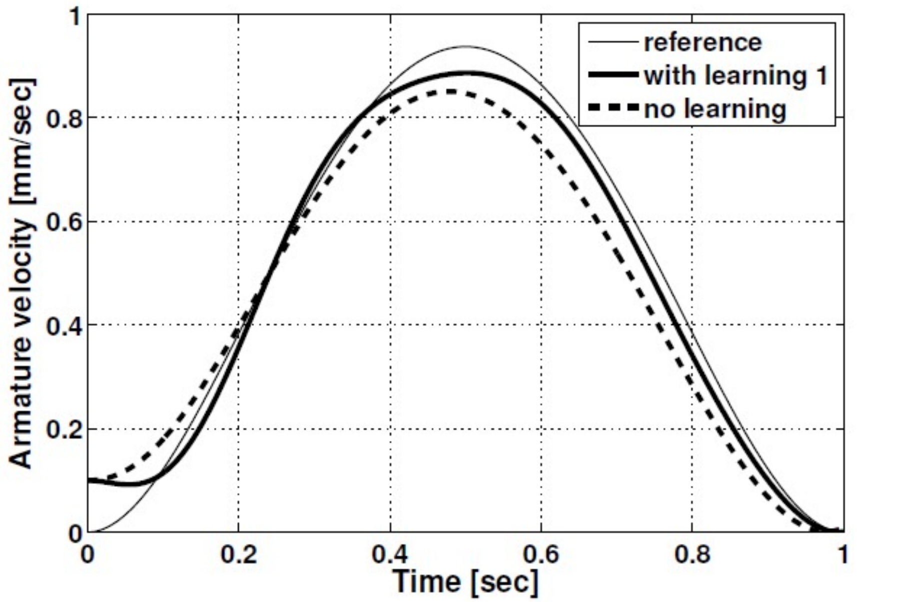
Mechatronics Examples: Electromagnetic brakes

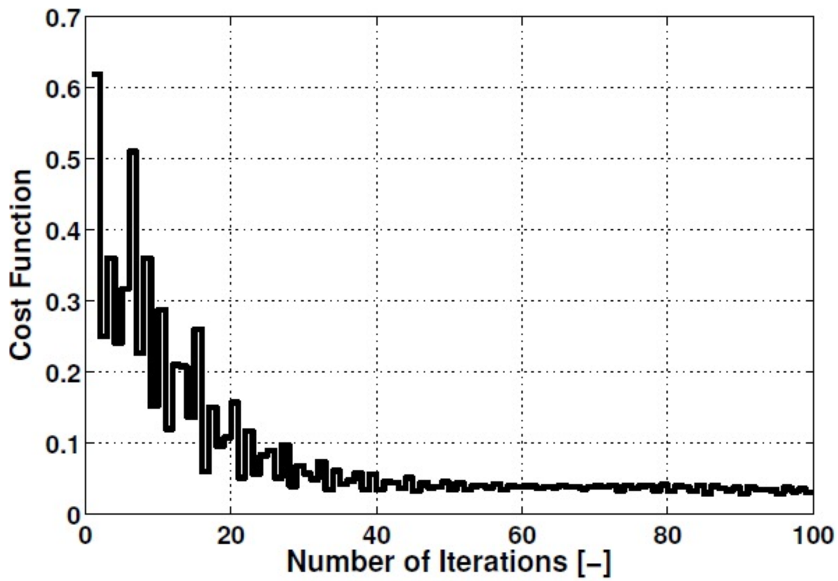
Learning-based auto-tuning of the controller gains:

$$\dot{x}_{K_1} = a_{K_1} sin(\omega_1 t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{K}_1(t) = x_{K_1}(t) + a_{K_1} sin(\omega_1 t + \tfrac{\pi}{2})$$
$$\dot{x}_{K_2} = a_{K_2} sin(\omega_2 t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{K}_2(t) = x_{K_2}(t) + a_{K_2} sin(\omega_2 t + \tfrac{\pi}{2})$$
$$\dot{x}_{K_3} = a_{K_3} sin(\omega_3 t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{K}_3(t) = x_{K_3}(t) + a_{K_3} sin(\omega_3 t + \tfrac{\pi}{2})$$
$$\dot{x}_{k} = a_{k} sin(\omega_4 t - \tfrac{\pi}{2}) Q(z(\beta))$$
$$\delta \hat{k}(t) = x_{k}(t) + a_{k} sin(\omega_4 t + \tfrac{\pi}{2})$$
$$\delta K_j(t) = \delta \hat{K}_j((I-1)t_f),\ (I-1)t_f \leq t < I t_f,$$
$$j \in \{1, 2, 3\},\ I = 1, 2, 3...$$
$$\delta k(t) = \delta \hat{k}((I-1)t_f),\ (I-1)t_f \leq t < I t_f,\ I = 1, 2, 3...$$

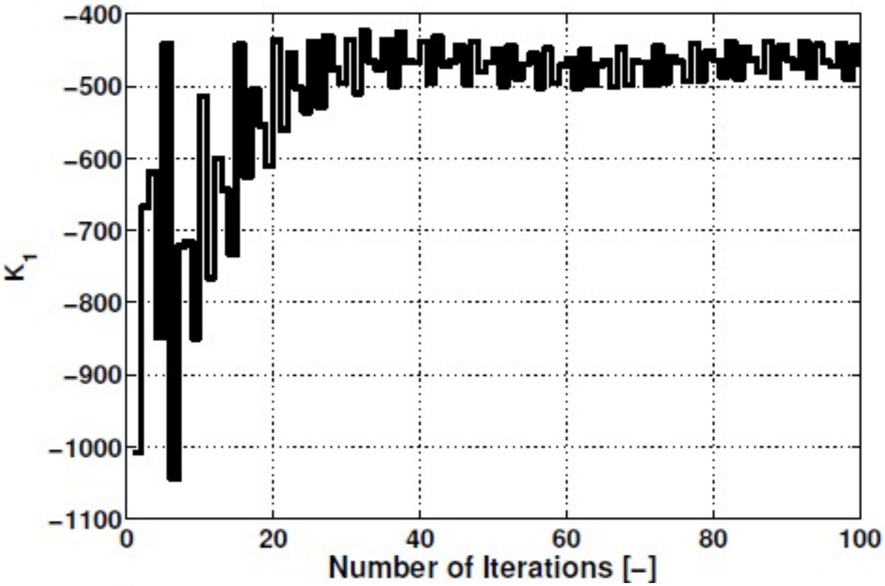
Mechatronics Examples: Electromagnetic brakes

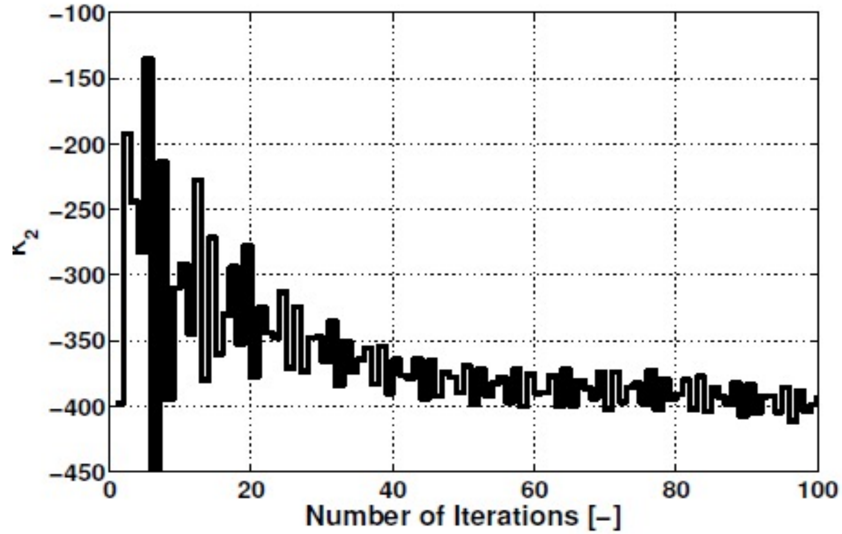


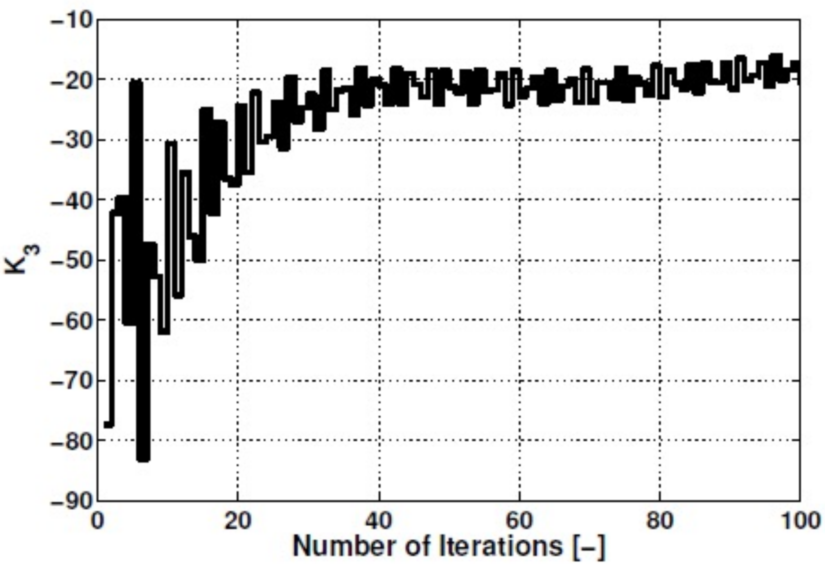


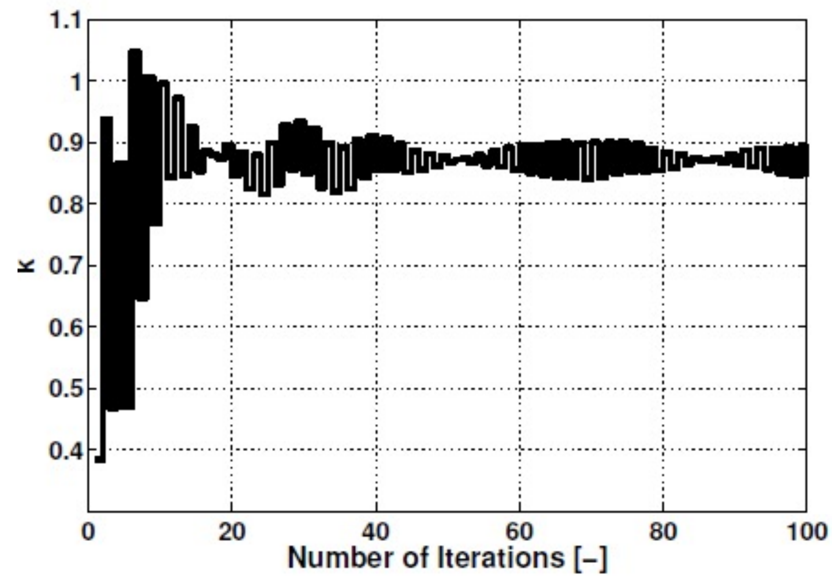


Mechatronics Examples: Electromagnetic brakes









Mechatronics Examples: DC- Servo motor with flexible shaft*

The example studied here is about the angular position control of a load connected by a flexible shaft to a voltage actuated DC servo motor.

- The states: the load angle, angular rate, the motor angle and angular rate
- The input is the motor voltage
- The outputs: the load angle and the torque acting on the flexible shaft
- uncertainties $\delta\beta_l = -70,\ [Nms/rad],\ \delta J_l = -0.2,\ [kgm^2]$

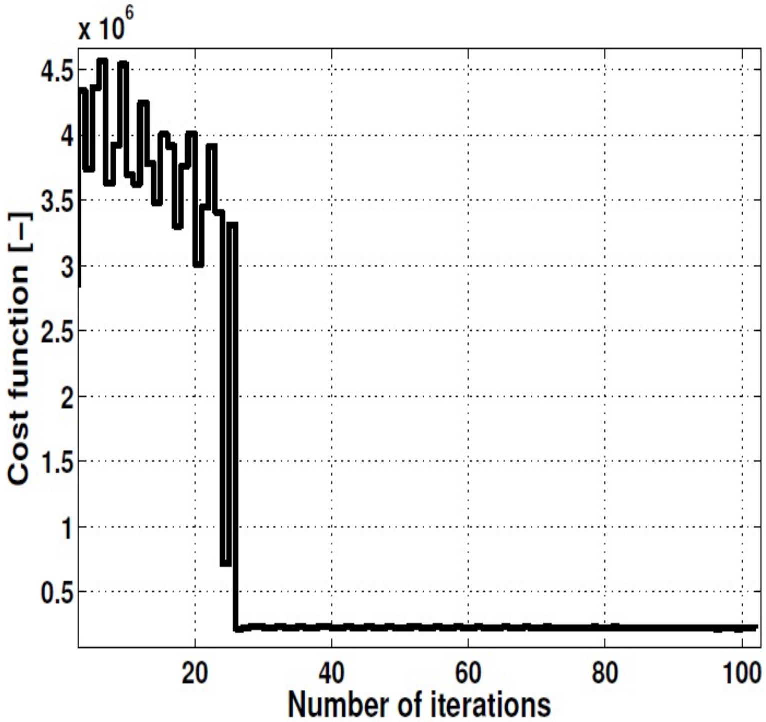
* M. Benosman, 2016, Learning-based adaptive control: An extremum seeking approach-Theory and Applications, Elsevier.

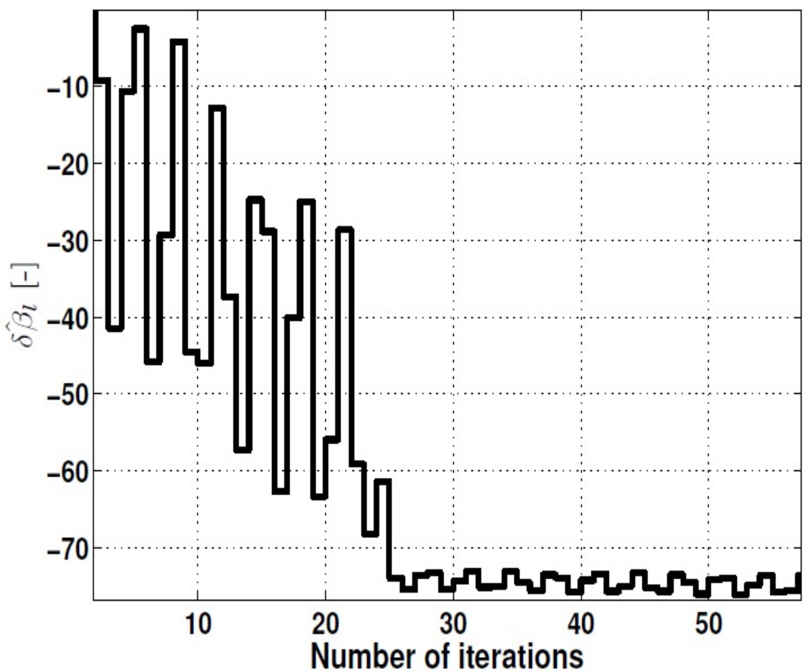
Mitsubishi Electric — Changes for the Better

Mechatronics Examples: DC- Servo motor with flexible shaft

uncertainties $\delta\beta_l = -70,\ [Nms/rad],\ \delta J_l = -0.2,\ [kgm^2]$

$$\mathcal{J}_l = 25\text{kgm}^2,\ \beta_l = 25\text{Nms/rad},$$



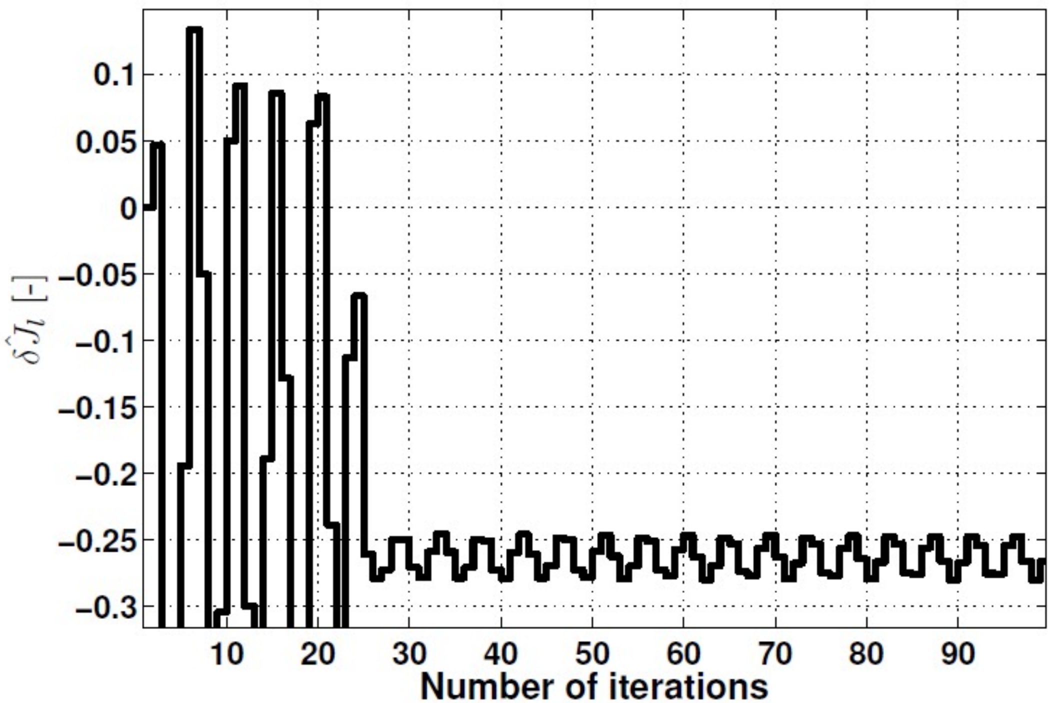


MITSUBISHI ELECTRIC
Changes for the Better

Mechatronics Examples: DC- Servo motor with flexible shaft

uncertainties $\delta\beta_l = -70,\ [Nms/rad],\ \delta J_l = -0.2,\ [kgm^2]$

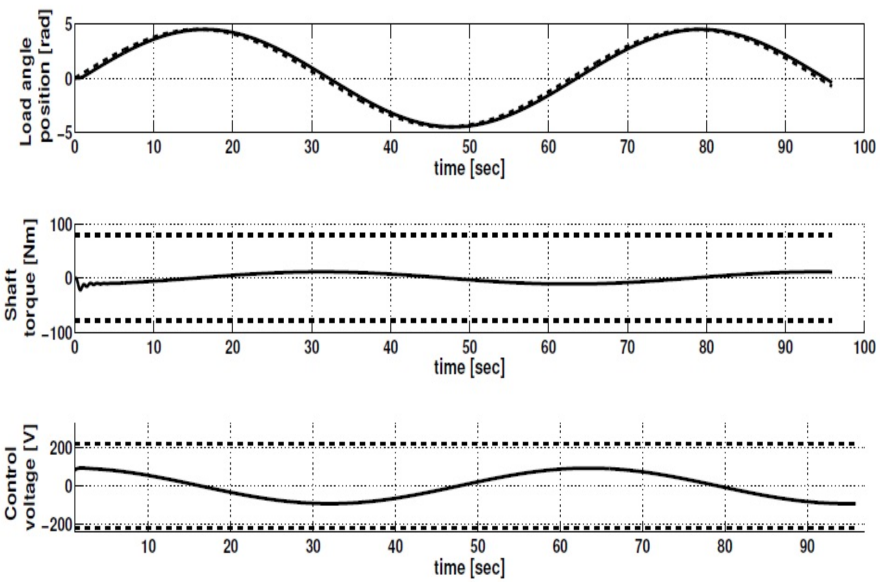
$$\mathcal{J}_l = 25\text{kgm}^2,\ \beta_l = 25\text{Nms/rad},$$



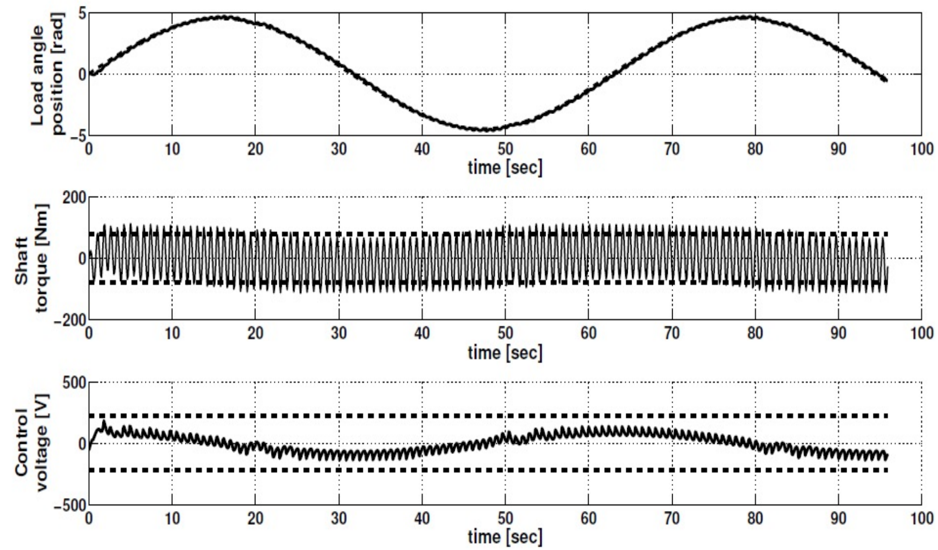
Mechatronics Examples: DC- Servo motor with flexible shaft

$$\text{uncertainties } \delta\beta_l = -70, \ [Nms/rad], \ \delta J_l = -0.2, \ [kgm^2]$$

$$\mathcal{J}_l = 25\text{kgm}^2, \ \beta_l = 25\text{Nms/rad},$$



Uncertain case- learning MPC



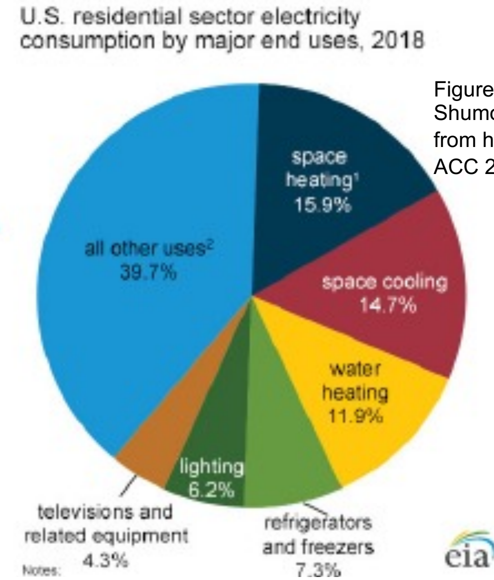
Uncertain case- nominal MPC

Fluid dynamics applications

Motivation

Efficient energy management in buildings
★ HVAC (Heating, Ventilation, and Air Conditioning)

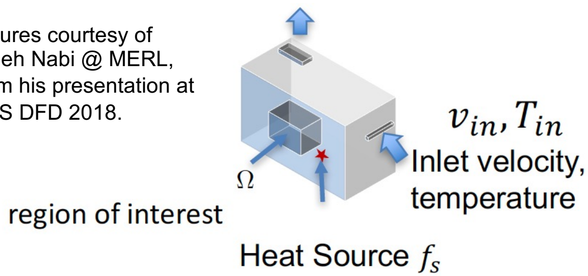
Source : U.S. Energy Information Administration, Annual Energy Outlook 2019.



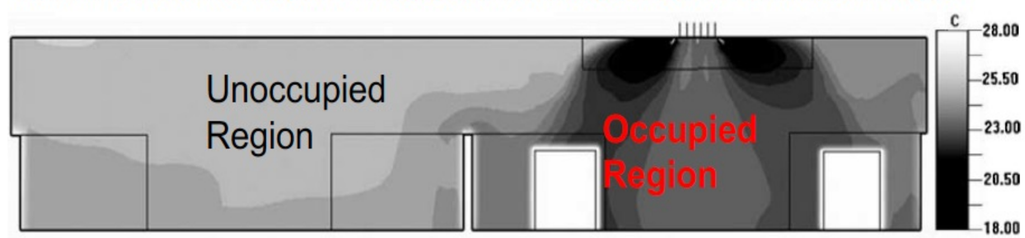
Figures courtesy of Shumon Koga@ UCSD, from his presentation at ACC 2019.

Optimizing HVAC performance is linked to modelling/controlling temperature and airflow in the room

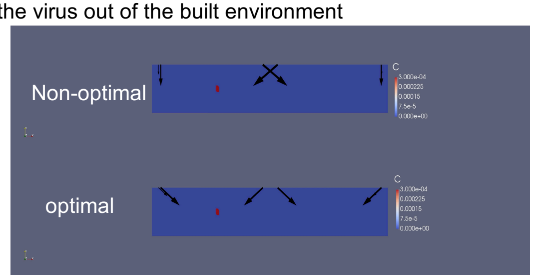
Figures courtesy of Saleh Nabi @ MERL, from his presentation at APS DFD 2018.



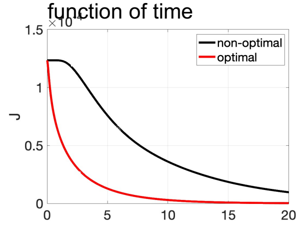
Lo et. al 2010: Open office plan with occupied and unoccupied regions. **10-15% Less Energy Use**



Hygiene applications: non-optimal vs optimal HVAC airflow to flush the virus out of the built environment



Concertation of virus as a function of time



How can we model airflow and temperature in a room with models that are *precise and computationally trackable for real-time estimation and control* ?

Fluid dynamics applications:

The Coupled Burgers' Equation*

$$\begin{aligned} w_t(t,x) + w(t,x)w_x(t,x) &= \mu w_{xx}(t,x) - \kappa T(t,x), \\ T_t(t,x) + w(t,x)T_x(t,x) &= cT_{xx}(t,x) + f(t,x). \end{aligned}$$

Where ω is the velocity variable, T is the temperature variable, f is s a forcing disturbance function, $\mu = \frac{1}{R_e}$, where R_e is the Reynolds number, c is the thermal conductivity, and κ is the thermal expansion coefficient. The notations F_x, F_{xx} stand for first and second partial derivatives of F w.r.t. x, respectively. The forcing f is assumed to be at least L^2 in space and time.

$w(t,0) = w_L$, $\frac{\partial w(t,1)}{\partial x} = w_R$,
$T(t,0) = T_L$, $T(t,1) = T_R$, ω_L, ω_R, T_L, T_R are positive constants.

The initial condition for the velocity is given by

$$w(0,x) = w_0(x) \qquad \in L^2([0,1])$$

and the initial condition for the temperature is

$$T(0,x) = T_0(x) \qquad \in L^2([0,1]) \ .$$

* Benosman M., Borggaard J., San O., Kramer, B., 2017, Learning-based robust stabilization for reduced order models of 2D and 3D Boussinesq equations, Applied Mathematical Modelling, Vol. 49, 162-181.

Fluid dynamics applications:
The Coupled Burgers' Equation

Following a Galerkin projection onto the subspace spanned by the POD basis functions, the coupled Burgers' equation is reduced to a POD ROM with the following structure

$$\begin{pmatrix} \dot{q}_w \\ \dot{q}_T \end{pmatrix} = B_1 + \mu B_2 + \mu\, D\, q + \tilde{D} q + C q q^T,$$

$$w_n^{pod}(x,t) = w_{av}(x) + \textstyle\sum_{i=1}^{i=r} \phi_{wi}(x) q_{wi}(t),$$

$$T_n^{pod}(x,t) = T_{av}(x) + \textstyle\sum_{i=1}^{i=r} \phi_{Ti}(x) q_{Ti}(t),$$

in the form

$$\begin{cases} \dot{q}^{pod}(t) = F(q^{pod}(t), \mu) = \tilde{F}(q^{pod}(t)) + \mu\, D q^{pod} \\ z^{pod}(t,x) = \sum_{i=1}^{i=N_{pod}} \phi_i^{pod}(x) q_i^{pod}(t), \end{cases}$$

$$\tilde{F} = B_1 + \mu B_2 + \tilde{D} q^{pod} + C q^{pod} {q^{pod}}^T,$$

Fluid dynamics applications: The Coupled Burgers' Equation

which can be upper-bounded by

$$\tilde{F} \leq b_{1_{max}} + \mu_{max} b_{2_{max}} + \tilde{d}_{max} \|q^{pod}\| + c_{max} \|q^{pod}\|^2,$$

where $\|B_1 + \Delta B_1\|_F \leq b_{1_{max}}$, $\|B_2 + \Delta B_2\|_F \leq b_{2_{max}}$, $\mu \leq \mu_{max}$, $\|\tilde{D} + \Delta \tilde{D}\|_F \leq \tilde{d}_{max}$, and $\|C + \Delta C\|_F \leq c_{max}$.
This leads to the nonlinear closure model

$$H_{nl} = \mu_{nl}(b_{1_{max}} + \mu_{max} b_{2_{max}} + \tilde{d}_{max} \|q^{pod}\| + c_{max} \|q^{pod}\|^2) diag(d_{11}, ..., d_{N_{pod}N_{pod}}) q^{pod}$$

Compete reduced order model

$$\dot{q}(t) = F(q(t), \mu) + H(t, q(t)).$$

$$H \rightarrow H_{nl}$$

$$\mu \rightarrow \mu_{cl} = \mu + \mu_e,$$

Fluid dynamics applications:
The Coupled Burgers' Equation

- boundary conditions $w_L = w_R = 0, \; T_L = T_R = 0,$

$$w(x,0) = \begin{cases} 1, & \text{if } x \in [0,\ 0.5] \\ 0, & \text{if } x \in]0.5,\ 1], \end{cases}$$

$$T(x,0) = \begin{cases} 1, & \text{if } x \in [0,\ 0.5] \\ 0, & \text{if } x \in]0.5,\ 1], \end{cases}$$

- Learning cost

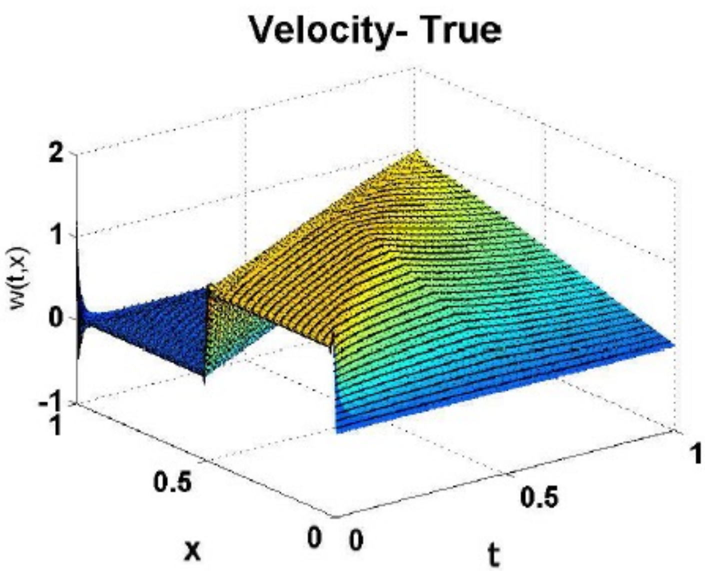
$$Q(\mu) = Q_1 \int_0^{t_f} < e_T, e_T >_X dt + Q_2 \int_0^{t_f} < e_w, e_w >_X dt, \quad Q_1 = Q_2 = 1$$

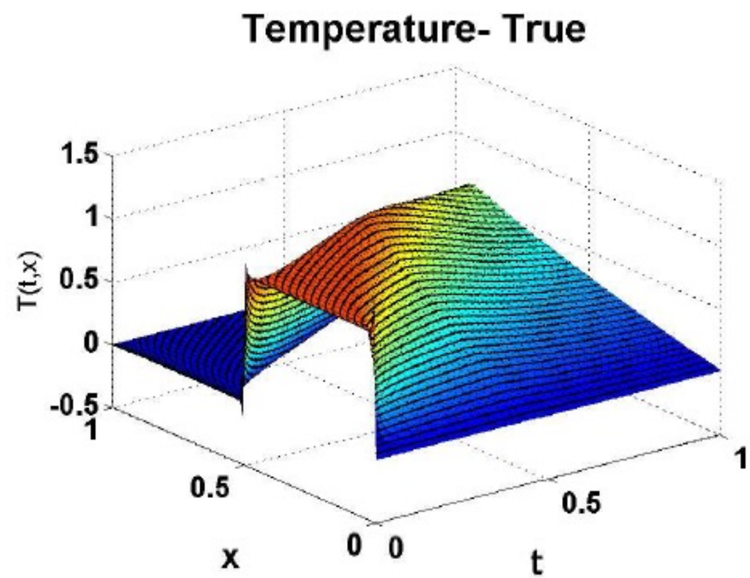
$$\mu = (\mu_e, \mu_n l)^T$$

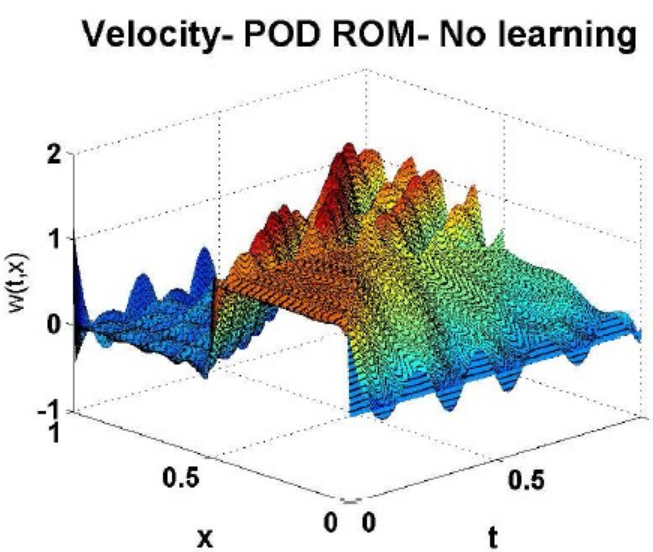
$$e_T = T_n - T_n^{pod}, \; e_w = w_n - w_n^{pod}$$

O. San and T. Iliescu, "Proper orthogonal decomposition closure models for fluid flows: Burgers equation," *International Journal of Numerical Analyis and Modeling*, vol. 1, no. 1, pp. 1–18, 2013.

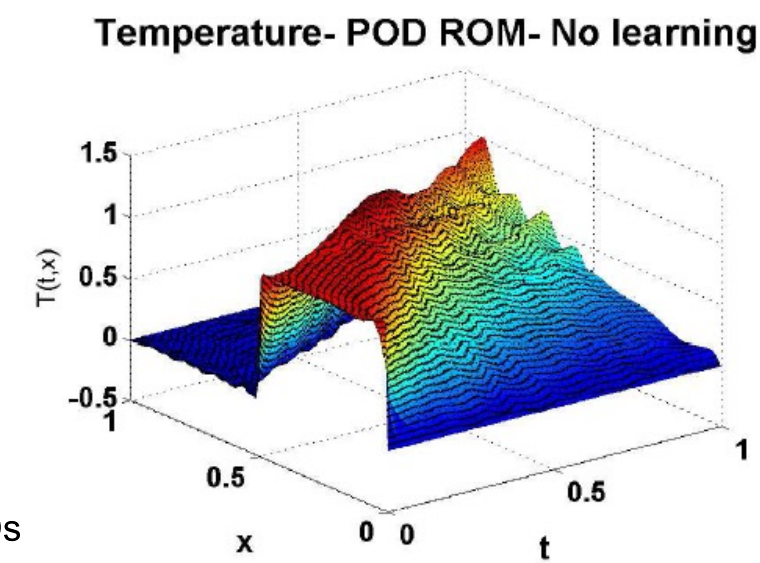
Fluid dynamics applications:
The Coupled Burgers' Equation



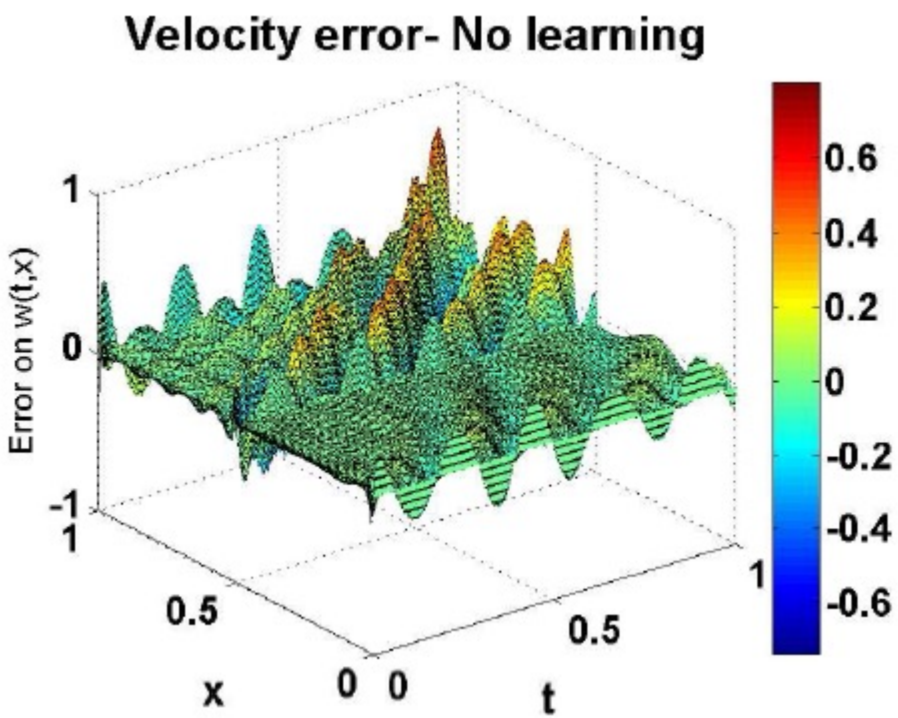


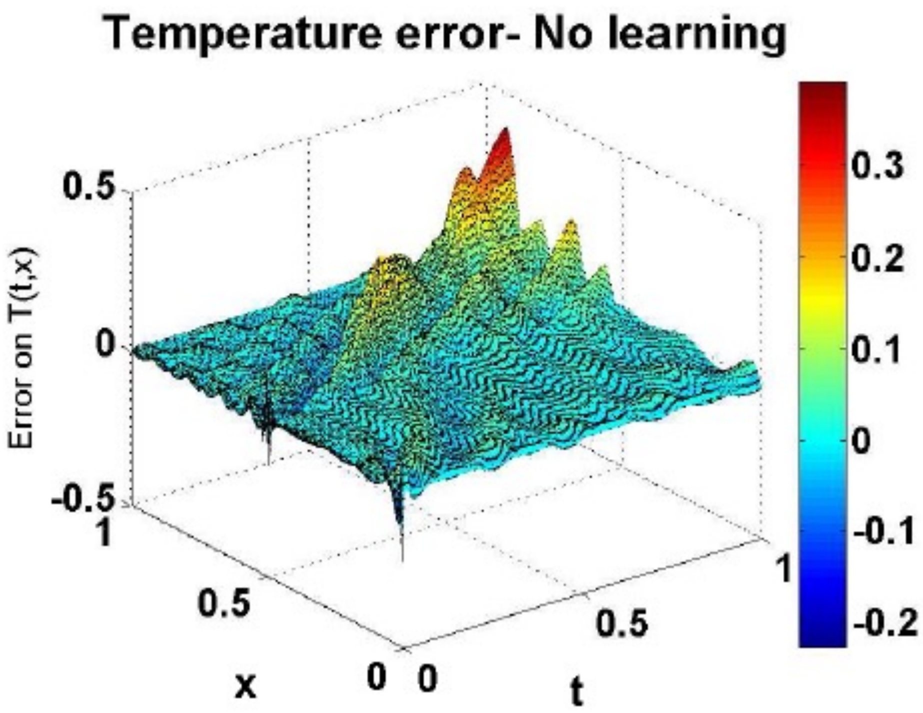


With 10 PODs



Fluid dynamics applications:
The Coupled Burgers' Equation

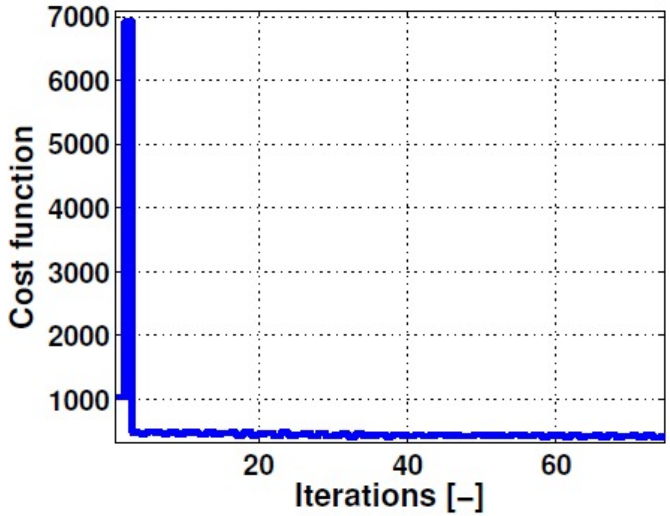


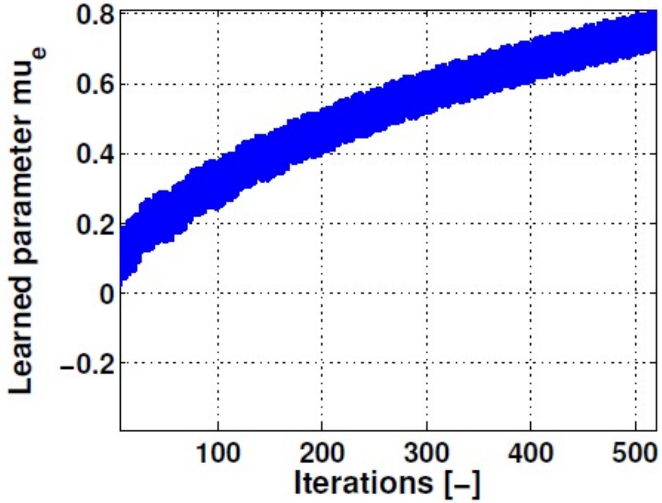


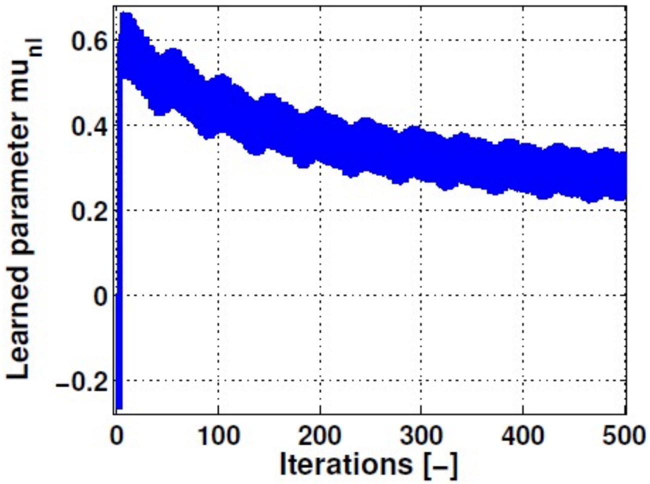
Error between ROM and systems' measurements before learning

Fluid dynamics applications:

The Coupled Burgers' Equation

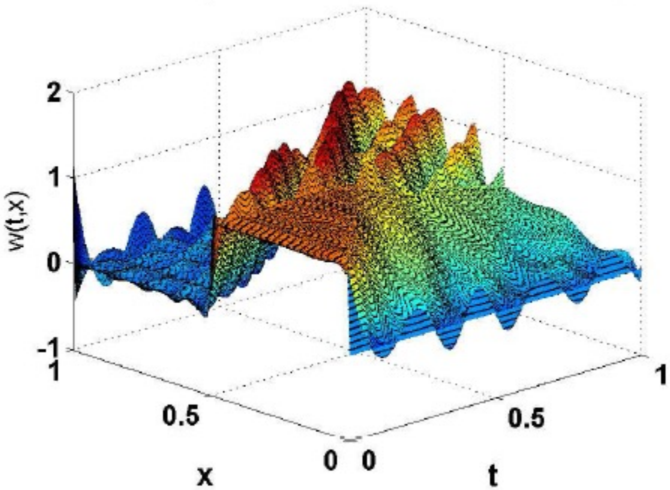




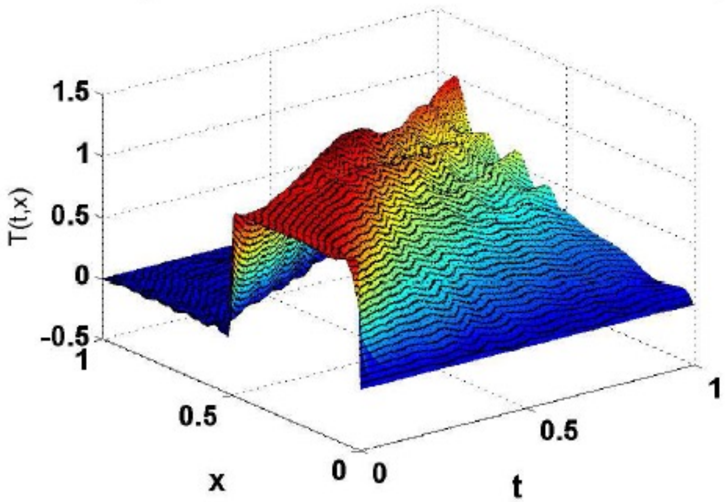


Fluid dynamics applications: The Coupled Burgers' Equation

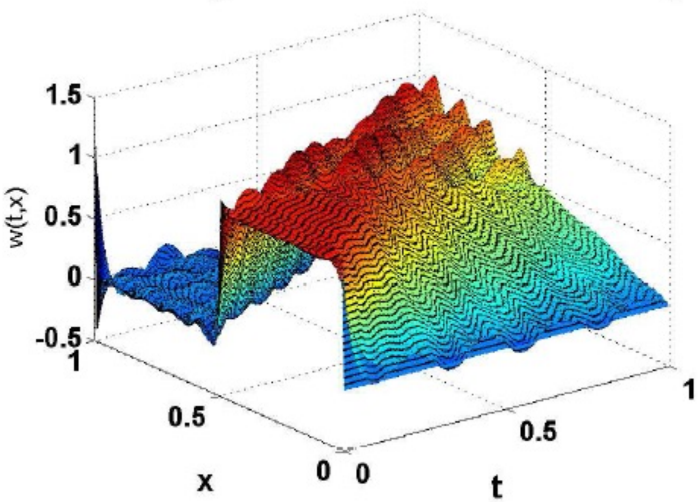
Velocity- POD ROM- No learning



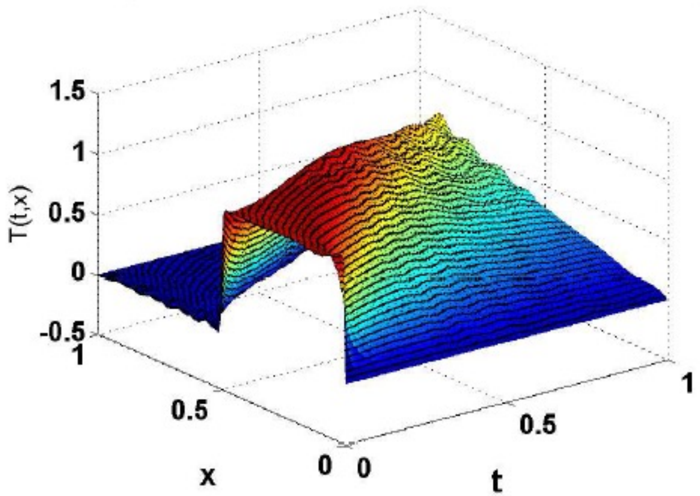
Temperature- POD ROM- No learning



Velocity- POD ROM- With learning

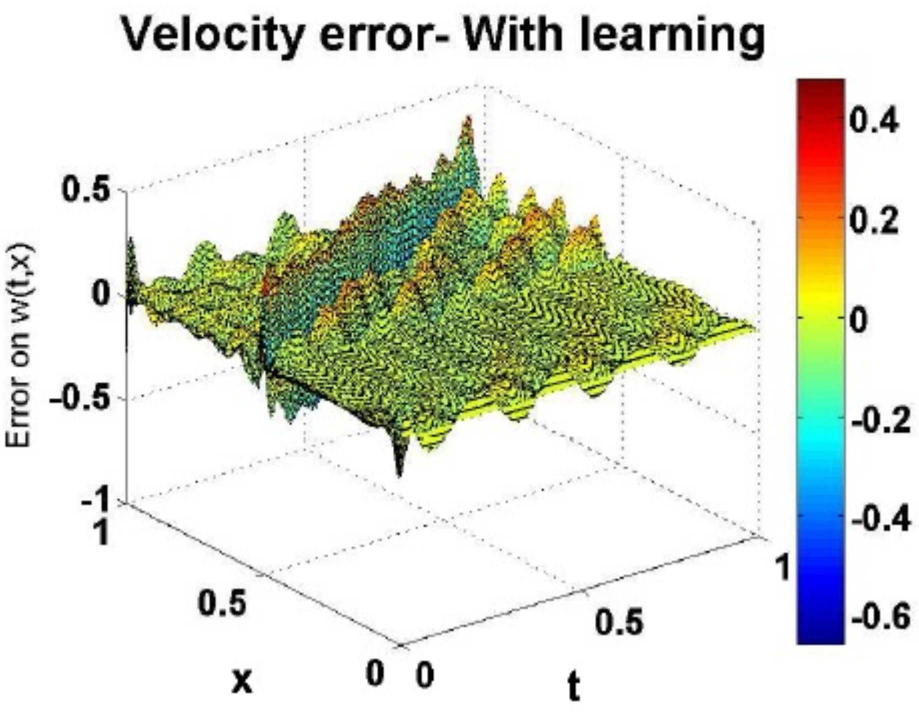


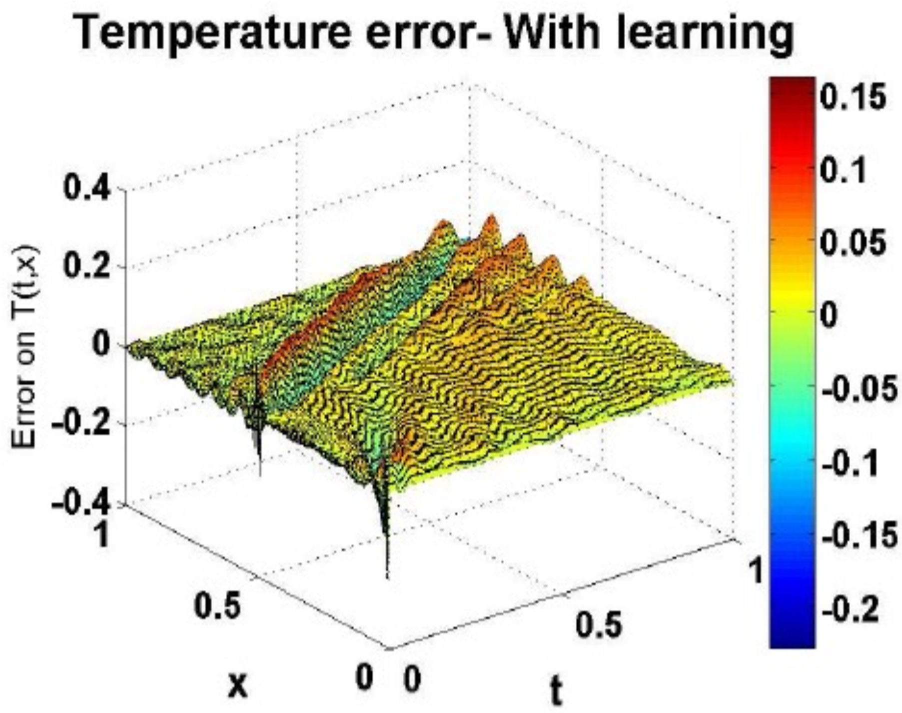
Temperature- POD ROM- With learning



With 10 PODs

Fluid dynamics applications: The Coupled Burgers' Equation





Error between ROM and systems' measurements after learning

Fluid dynamics applications:

The 3D Boussinesq Equation

3D incompressible Boussinesq equations

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla p+\nabla\cdot\tau(\mathbf{v})+\rho\mathbf{g}$$

$$\nabla\cdot\mathbf{v}=0$$

$$\rho c_p\left(\frac{\partial T}{\partial t}+\mathbf{v}\cdot\nabla T\right)=\nabla\left(\kappa\nabla T\right),$$

$$\tau(\mathbf{v})\;=\;\rho\nu\left(\nabla\mathbf{v}+\nabla\mathbf{v}^T\right)$$

POD ROM with the following structure,

$$\dot{q}^{pod}=\mu\, D\, q^{pod}+Cq^{pod}{q^{pod}}^T+Pq^{pod}{q^{pod}}^T,$$
$$v(x,t)=v_0(x)+\sum_{i=1}^{i=N_{pod-v}}\phi(x)_i^{pod-v}q_i^{pod-v}(t),$$
$$T(x,t)=T_0(x)+\sum_{i=1}^{i=N_{pod-T}}\phi(x)_i^{pod-T}q_i^{pod-T}(t),$$

Fluid dynamics applications:

The 3D Boussinesq Equation

in the form
$$\begin{cases} \dot{q}^{pod}(t) = F(q^{pod}(t), \mu) = \tilde{F}(q^{pod}(t)) + \mu\, D q^{pod} \\ z^{pod}(t,x) = \sum_{i=1}^{i=N_{pod}} \phi_i^{pod}(x) q_i^{pod}(t), \end{cases}$$

$$\tilde{F} = C q^{pod} q^{pod^T} + P q^{pod} q^{pod^T}$$
, *C, P* are kept separate to track the impact of different physical uncertainties on the ROM

If we consider bounded parametric uncertainties on the coefficients of C and P,

$$\tilde{F} = (C + \Delta C) q^{pod} q^{pod^T} + (P + \Delta P) q^{pod} q^{pod^T},$$

where $\|C + \Delta C\|_F \leq c_{max}$, and $\|P + \Delta P\|_F \leq p_{max}$,

$$\tilde{F} \leq c_{max} \|q^{pod}\|^2 + p_{max} \|q^{pod}\|^2$$

This leads to the nonlinear closure model

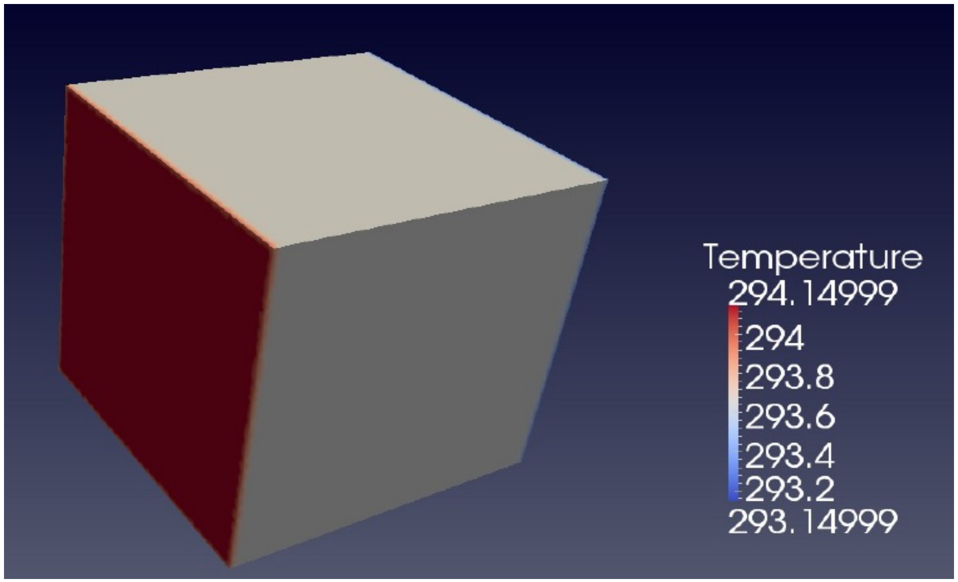
$$H_{nl} = \mu_{nl}(c_{max}\|q^{pod}\|^2 + p_{max}\|q^{pod}\|^2)\ diag(d_{11}, \ldots, d_{N_{pod}N_{pod}}) q^{pod}, \ \mu_{nl} > 0$$

$$\mu \to \mu_{cl} = \mu + \mu_e,$$

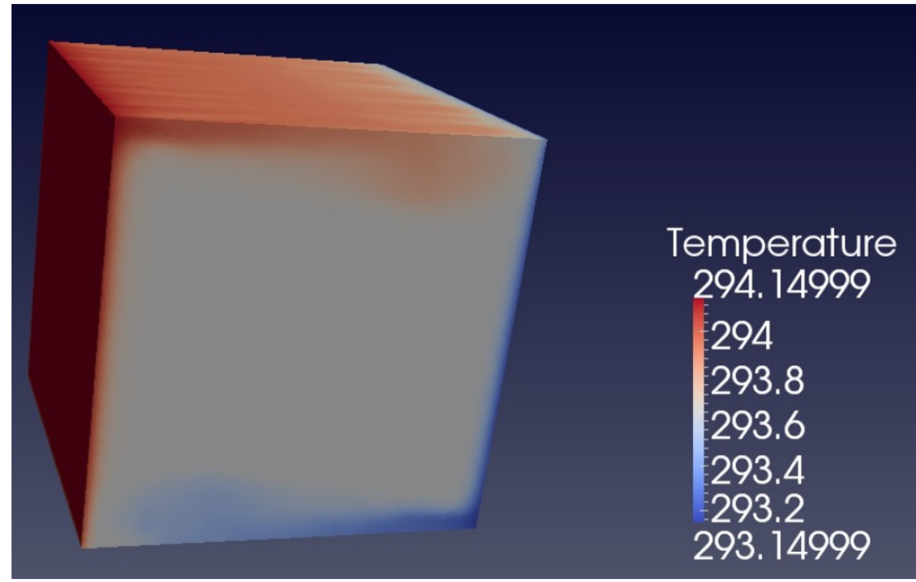
Learning cost
$$Q(\mu) = \int_0^{t_f} \langle e_T, e_T \rangle_{\mathcal{H}} dt + \int_0^{t_f} \langle e_{\mathbf{v}}, e_{\mathbf{v}} \rangle_{(\mathcal{H})^3} dt.$$

$$\mu = (\mu_e, \mu_n l)^T$$

Numerical Results: The 3D Boussinesq Equation (Rayleigh-Benard modified problem)



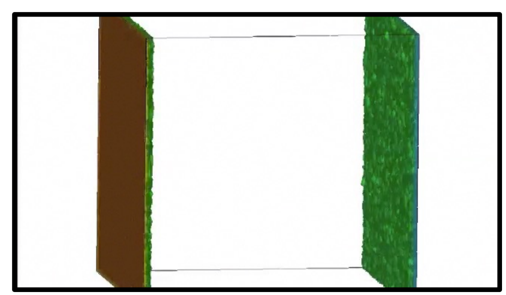
Exact temperature at t0



Exact temperature at t=50sec

temperature was specified at ± 0.5 on the x-faces and taken as homogeneous Neumann on the remaining faces. $\mathrm{Re} = 4.964 \times 10^4$, $\mathrm{Pr} = 0.712$, and $\mathrm{Gr} = 7.369 \times 10^7$ The simulation was run from zero velocity and temperature and snapshots were collected for 78 seconds.

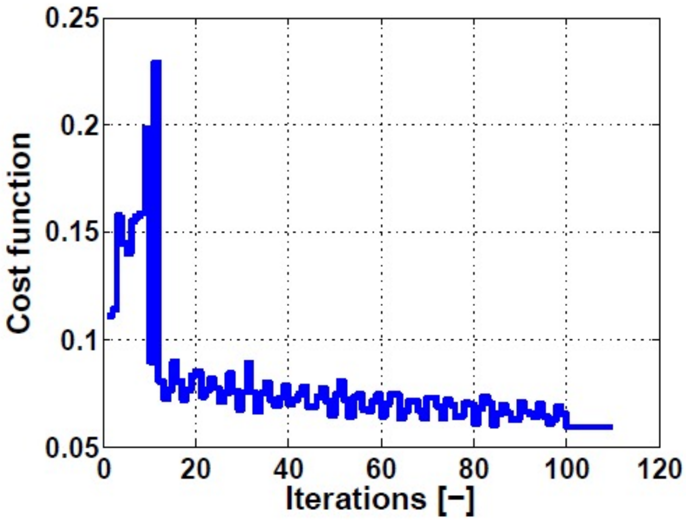
In this case we use 8PODs for the Galerkin ROM (ROM-G)

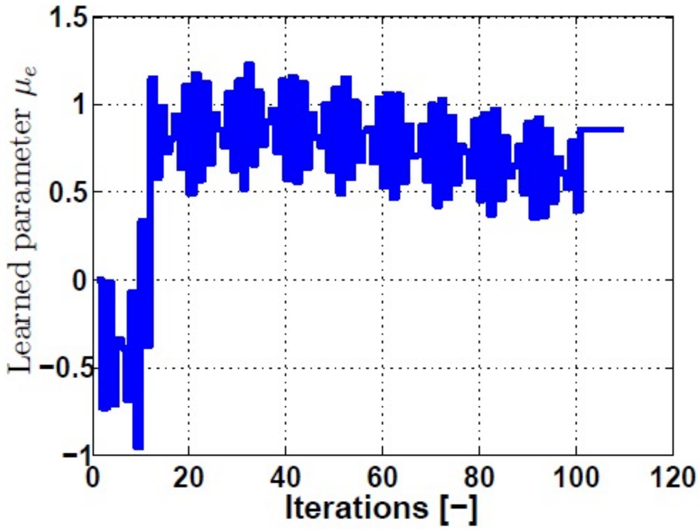


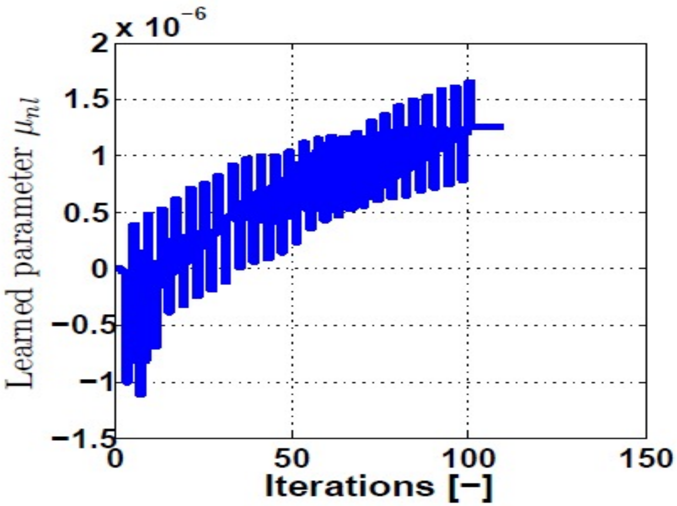
3D flow video

Temperature flow

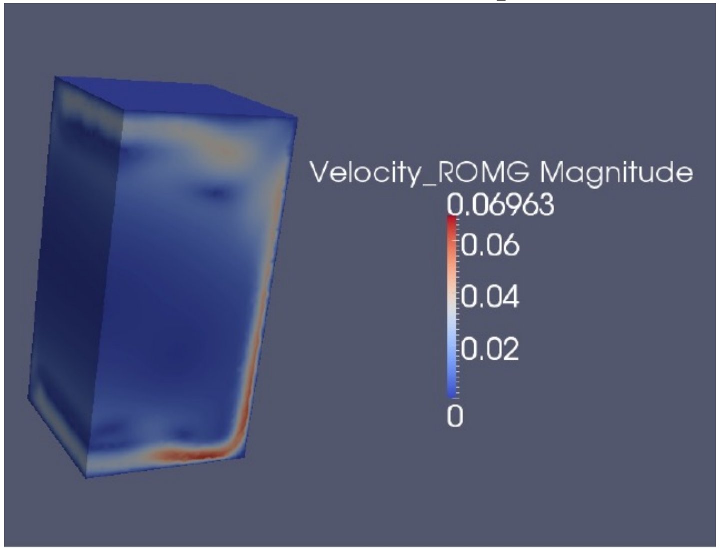
Numerical Results: The 3D Boussinesq Equation



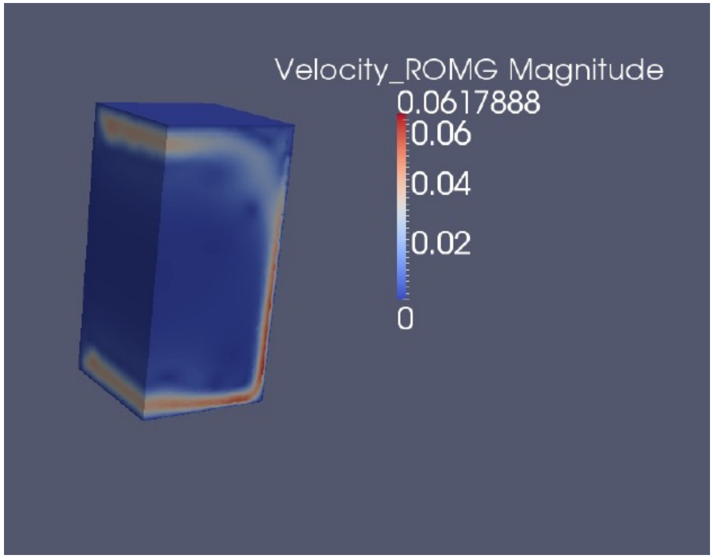




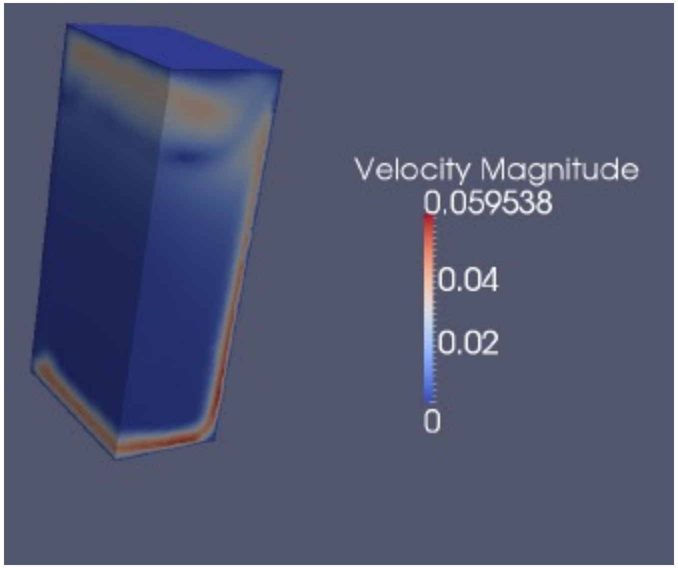
Numerical Results: The 3D Boussinesq Equation (Rayleigh-Benard modified problem)



ROM-G velocity clip no learning at t=50 sec

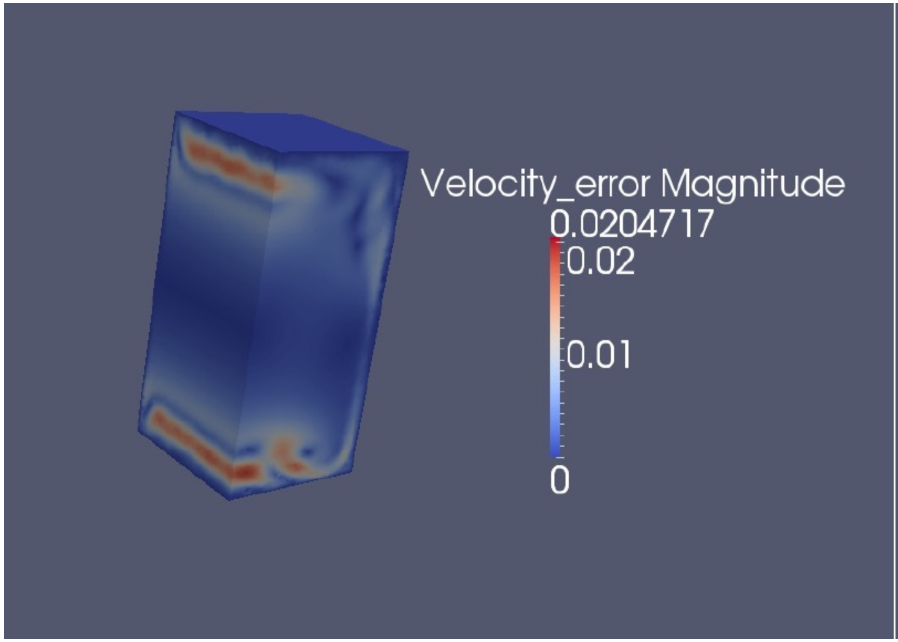


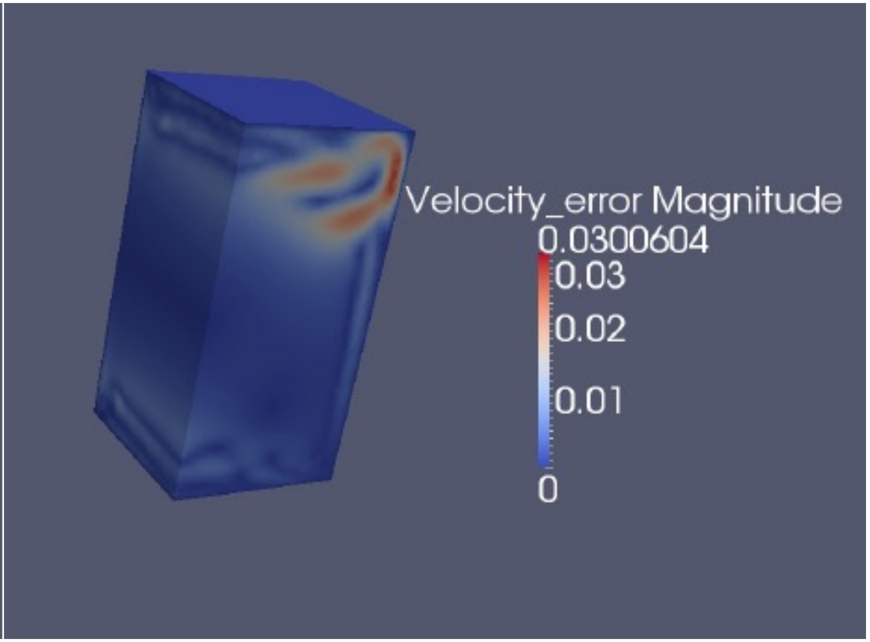
ROM-G velocity clip with learning at t=50 sec



True velocity clip at t=50 sec

Numerical Results: The 3D Boussinesq Equation

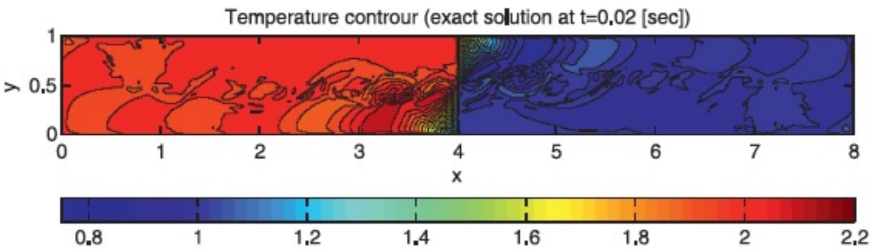


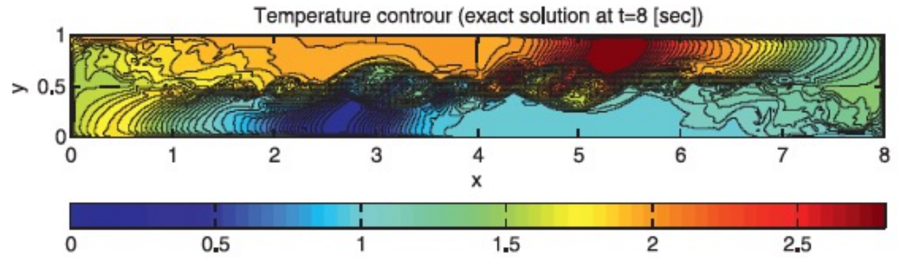


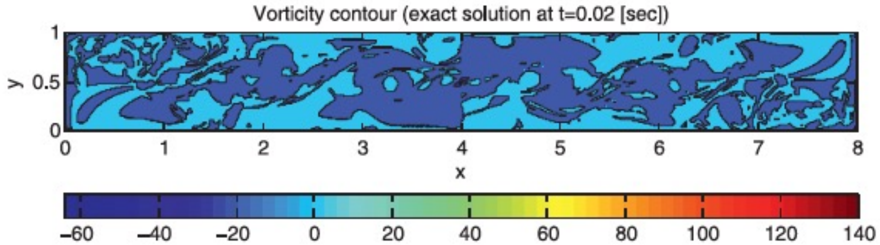
Clip of the velocity error at t=50 sec.
ROM-G (no learning-8 PODs)

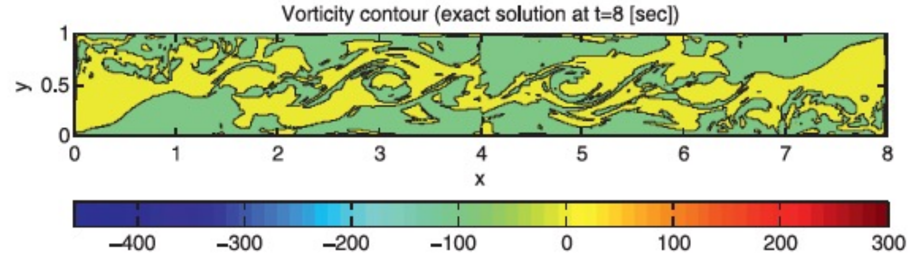
Clip of the velocity error at t=50 sec.
ROM-GL (with learning-8 PODs)

Numerical Results: The 2D Boussinesq Equation- Unsteady lock exchange flow problem

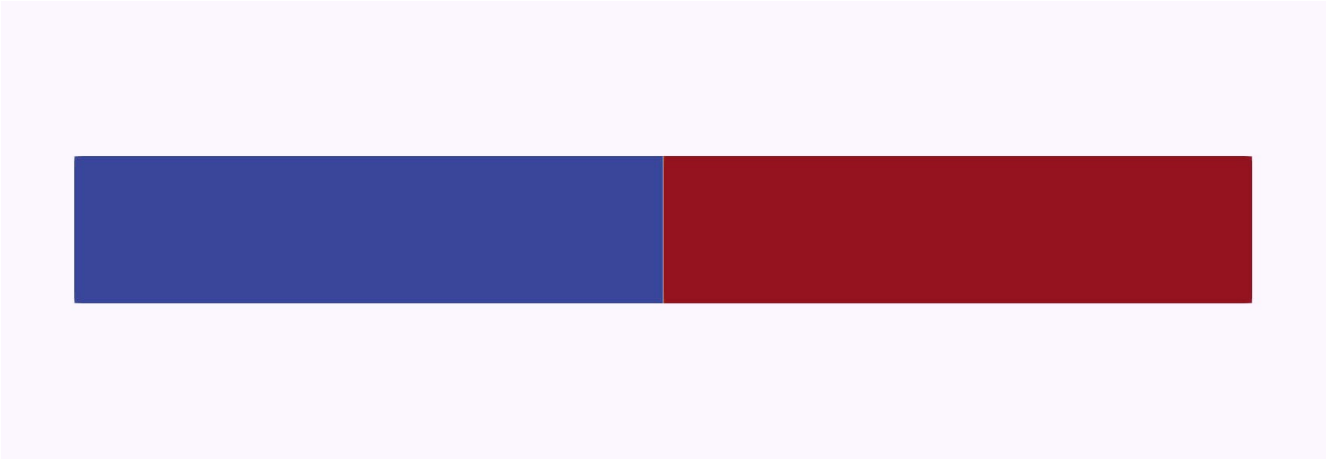




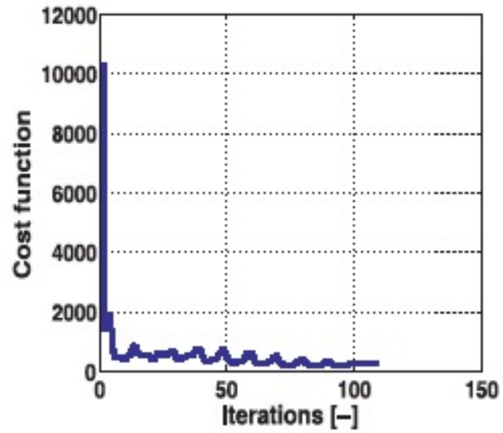


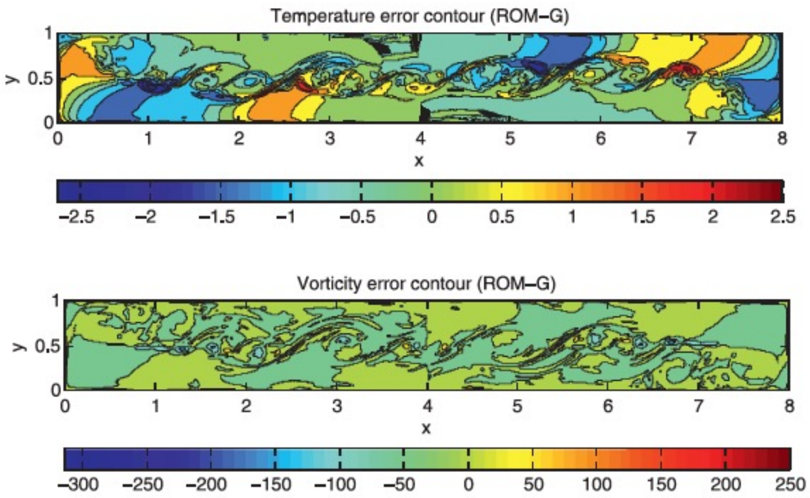


2D flow video

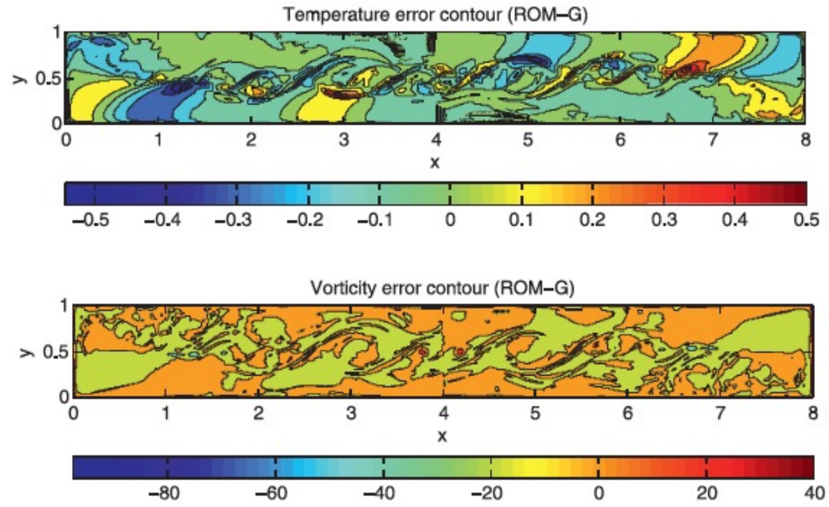


Numerical Results: The 2D Boussinesq Equation- Unsteady lock-exchange flow problem





Reconstruction error ROM-G (no learning)



Reconstruction error ROM-G (with learning)

Robotics examples: Rigid manipulators*

We consider here a two-link robot manipulator.

$$H(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) \;=\; \tau,$$

where $q \triangleq [q_1, q_2]^T$ denote the two joint angles and $\tau \triangleq [\tau_1, \tau_2]^T$ denote the two joint inputs.

Now we assume uncertainties in the model

$$\ddot{q} \;=\; H^{-1}(q)\tau - H^{-1}(q)\left[C(q,\dot{q})\dot{q} + G(q)\right] + \Delta b(q,t)$$

* Benosman M., Farahmand A.-M., Xia M., 2018, Learning-based iterative modular adaptive control for nonlinear systems, International Journal of Adaptive Control and Signal Processing, 33(2), pp. 335-355, doi.org/10.1002/acs.2892.

Robotics examples: Rigid manipulators

- The reference trajectory

$$q_{id}(t) = \frac{1}{1+\exp(-t)}, \quad i = 1, 2$$

- The extremum seeking algorithm

$$x_i(k+1) = x_i(k) + t_f \alpha_i \sqrt{\omega_i} \cos(\omega_i t_f k) - t_f \kappa_i \sqrt{\omega_i} \sin(\omega_i t_f k) J, \quad i = 1, 2$$

Robotics examples: Rigid manipulators

Case 1 : $\Delta b = [\Delta b_1(t), \Delta b_2(t)]^T$

$$\Delta b_1(t) = 1 - 0.14 \sin(0.01\, t),$$
$$\Delta b_2(t) = 1 - 0.12 \cos(0.01\, t).$$

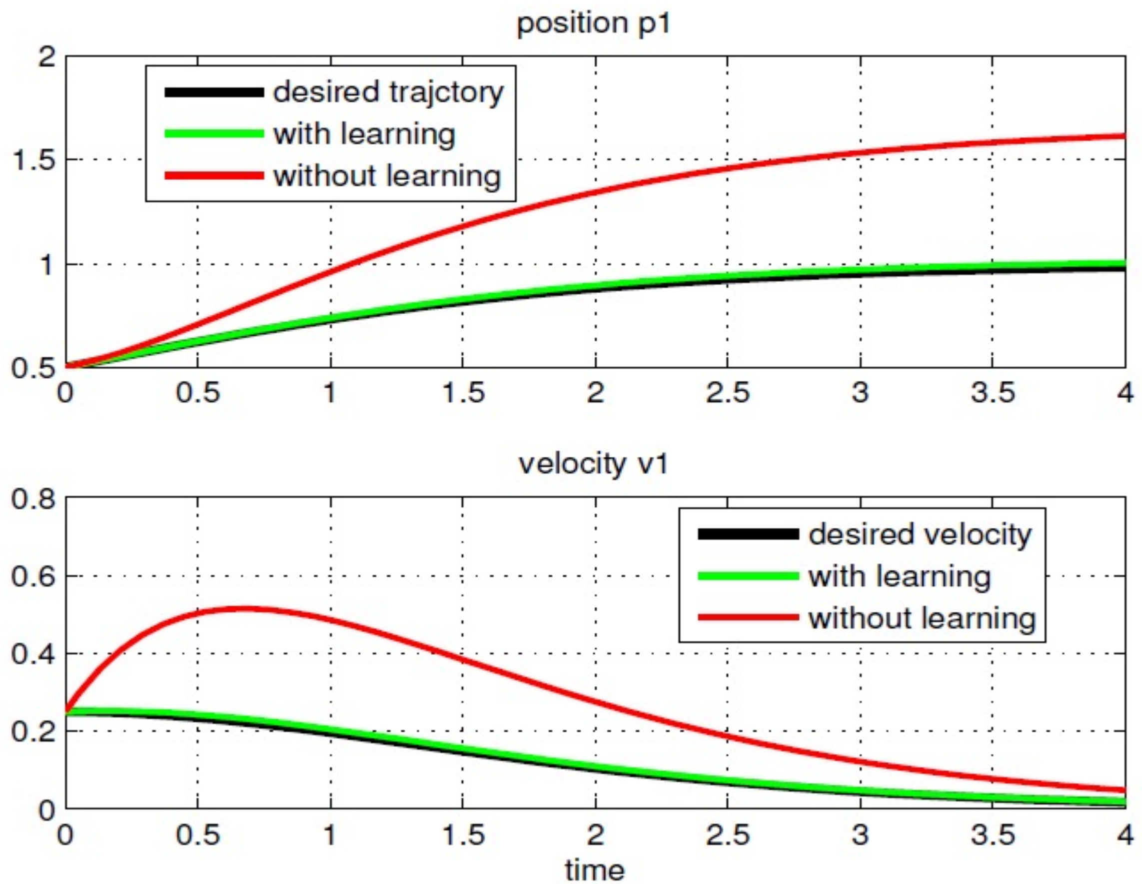
→ the cost function

$$J = Q_1 \int_{(N-1)t_f}^{Nt_f} (q - q_d)^T (q - q_d) \mathrm{d}t + Q_2 \int_{(N-1)t_f}^{Nt_f} (\dot{q} - \dot{q}_d)^T (\dot{q} - \dot{q}_d) \mathrm{d}t,$$

$Q_1 > 0$, $Q_2 > 0$ and $N = 1, 2, \cdots$

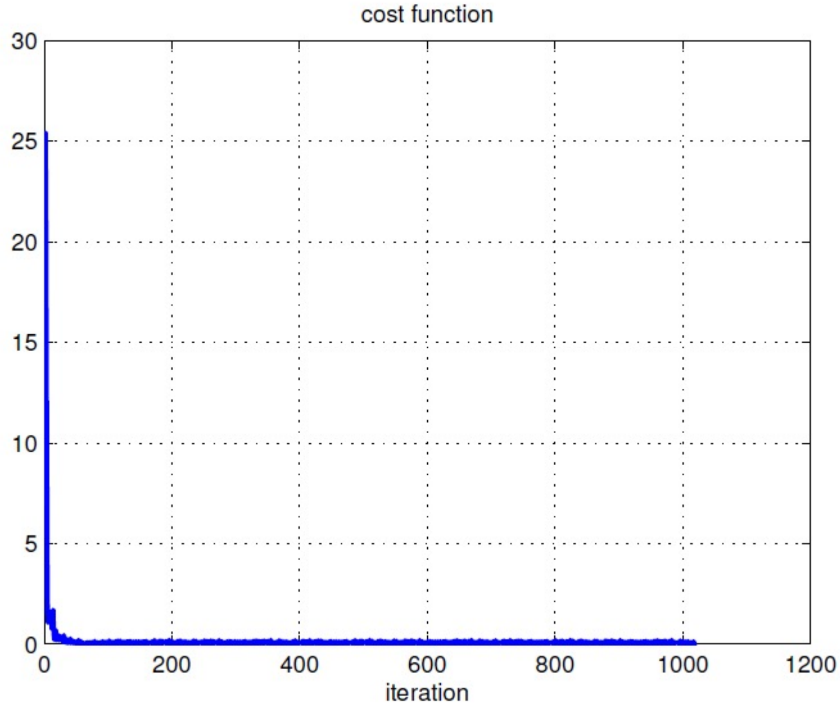
Robotics examples: Rigid manipulators

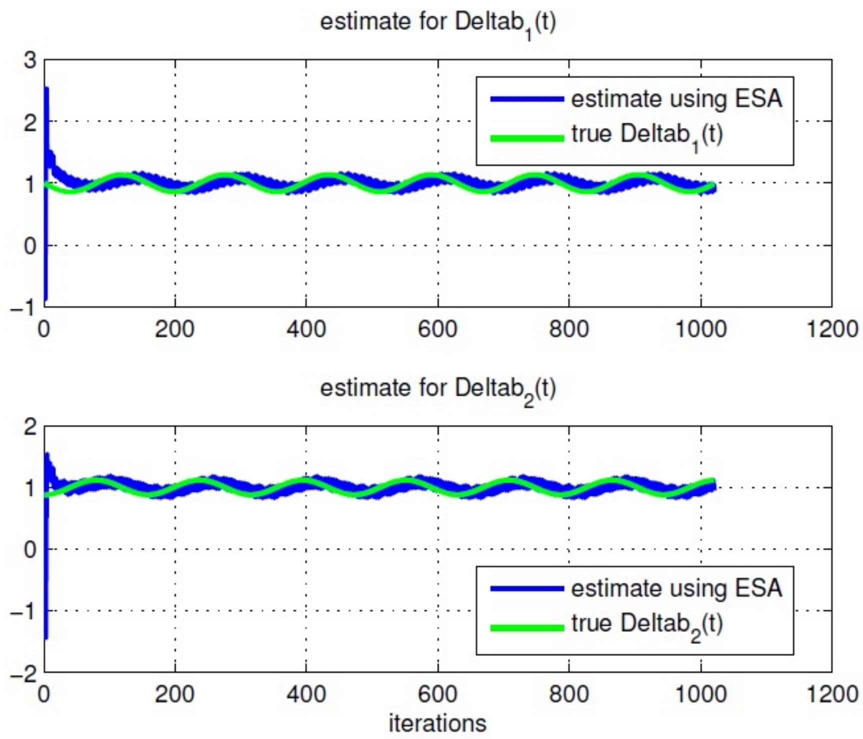
$$\text{Case } 1: \Delta b = [\Delta b_1(t), \Delta b_2(t)]^T$$



Robotics examples: Rigid manipulators

$$\textbf{Case 1}: \Delta b = [\Delta b_1(t), \Delta b_2(t)]^T$$

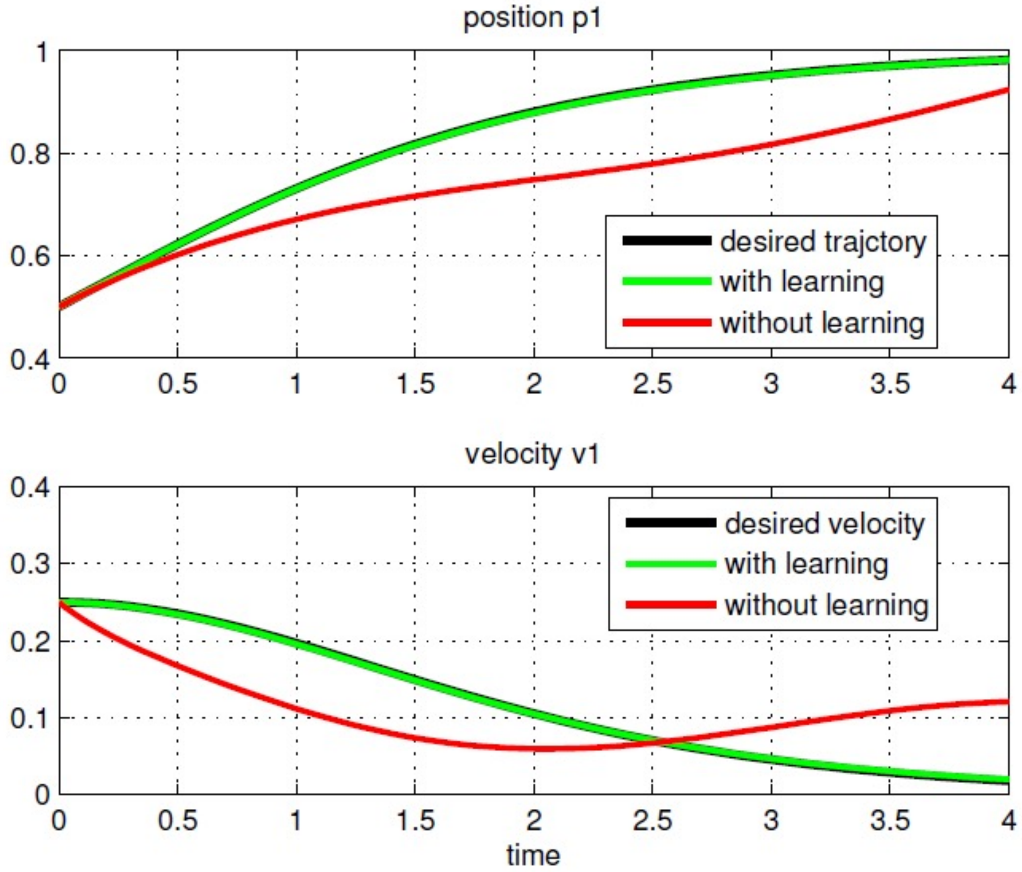




Robotics examples: Rigid manipulators

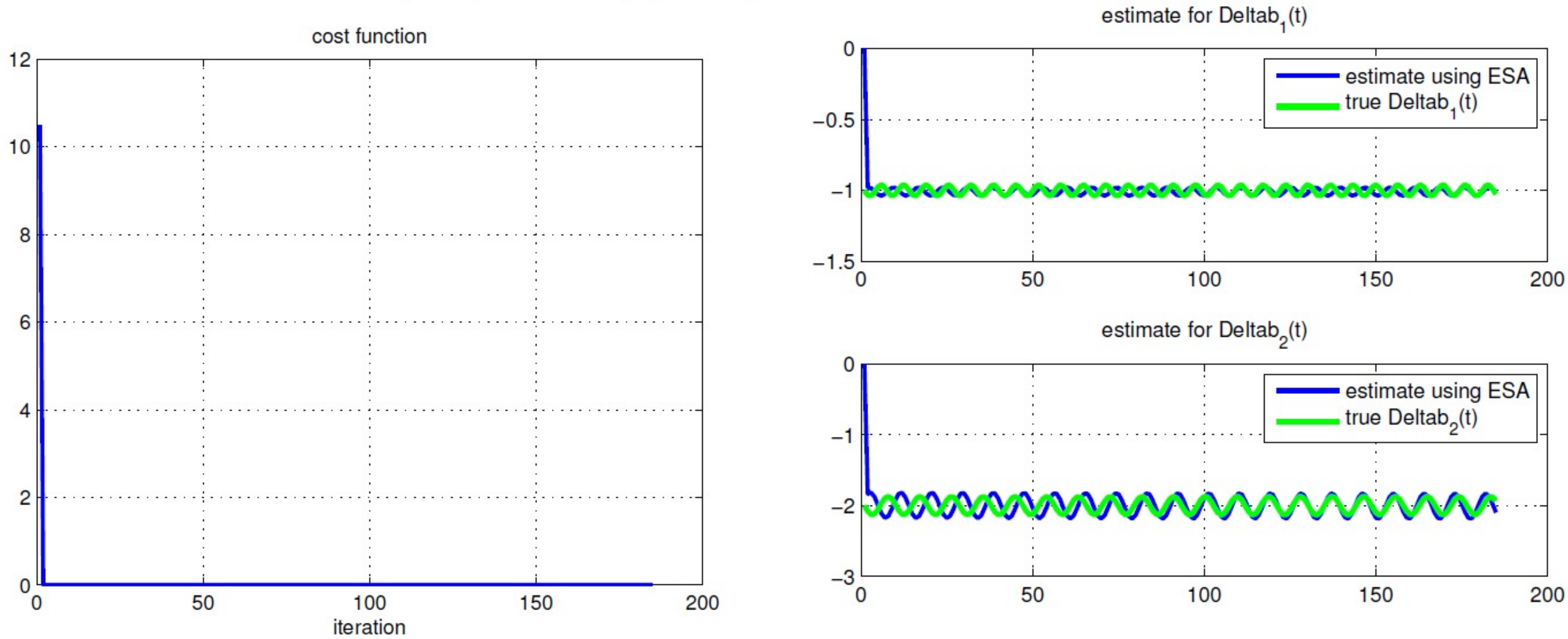
Case 2 : $\Delta b(q,t) = \Delta(t) \times (D\dot{q})$

$$\Delta b_1(t) = -1 - 0.04 \ \sin(0.24 \ t),$$
$$\Delta b_2(t) = -2 - 0.13 \ \sin(0.17 \ t).$$



Robotics examples: Rigid manipulators

Case 2: $\Delta b(q,t) = \Delta(t) \times (D\dot{q})$

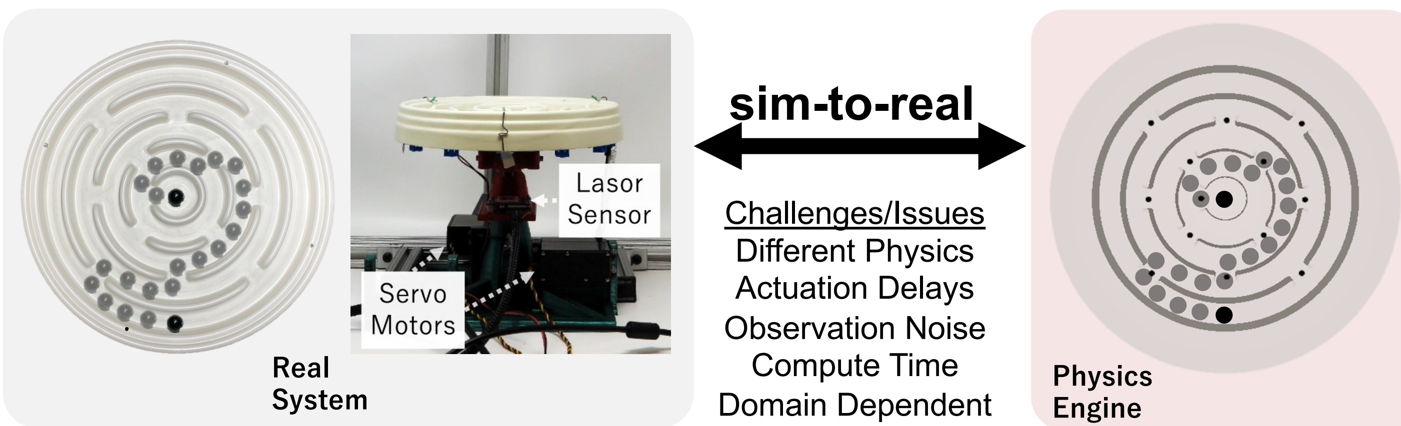


Robotics applications : Maze mounted on a servo-motor*

Slide from: Vetro A.@MERL, keynote at IEEE Conference on Autonomous Systems (ICAS), 2021.

Circular Maze Environment (CME)

- Tip and Tilt the maze so that the marble moves from the outer ring into the inner-most circle
- Intuitive to humans; most humans can solve very quickly
- Complex for RL agent due to constrained geometry, underactuated control, nonlinear dynamics, etc.



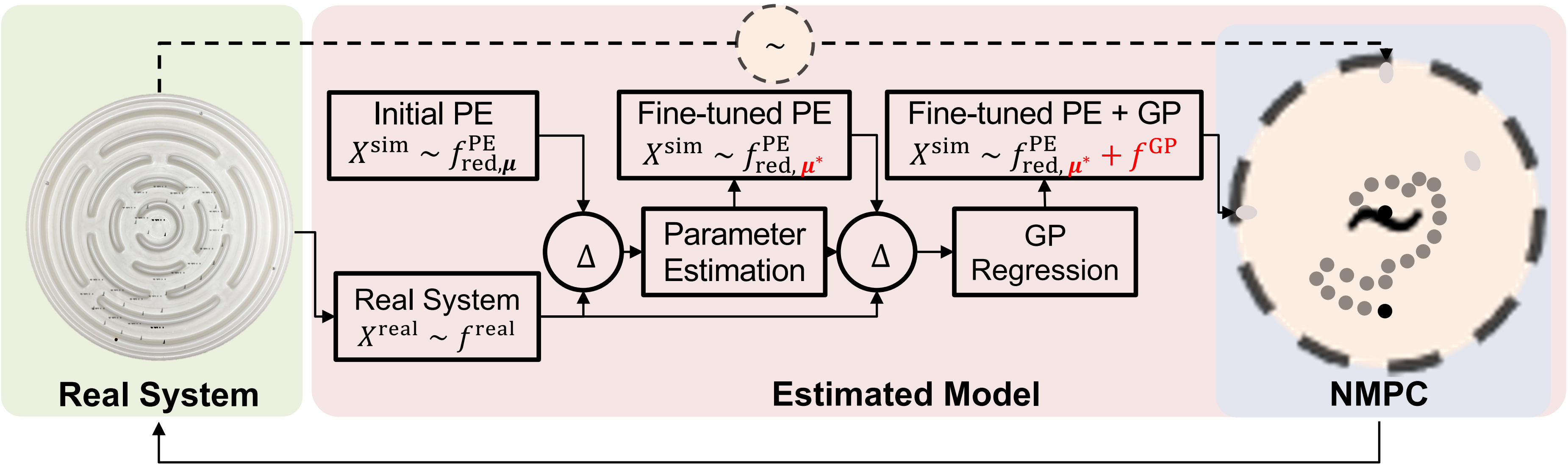
* Ota, K., Jha, D.K., Romeres, D., van Baar, J., Smith, K., Semistsu, T., Oiki, T., Sullivan, A., Nikovski, D.N., Tenenbaum, J.B., 2021, Data-Efficient Learning for Complex and Real-Time Physical Problem Solving using Augmented Simulation, *IEEE Robotics and Automation Letters*, DOI: 10.1109/LRA.2021.3068887, Vol. 6, No. 2,.

Robotics applications : Maze mounted on a servo-motor

Slide from: Vetro A.@MERL, keynote at IEEE Conference on Autonomous Systems (ICAS), 2021.

Goal: obtain accurate model and exploit it for model-based RL

1. Collect real trajectories $X^{\text{real}} \sim f^{\text{real}}$ in the real system
2. Estimate physical parameters $\boldsymbol{\mu}^*$ to obtain a more accurate physics engine
3. Learn residual model using Gaussian Process
4. Use the estimated model to control the real system with NMPC policy



Robotics applications : Maze mounted on a rigid arm

Slide from: Vetro A.@MERL, keynote at IEEE Conference on Autonomous Systems (ICAS), 2021.

Experiments: Comparison with Human Performance

- 15 participants were asked to solve the maze by looking at the video feed of the marble movement
- To familiarize them with the controls, they were given 1 minute to play with the maze using a joystick, but no marble
- Then, they were asked to solve the maze five times

✓ **Can move the marble to goal within minutes of interaction**

✓ **Consistently improve performance with larger amount of data**

MITSUBISHI ELECTRIC
Changes for the Better

Other applications

Batteries estimation

- Wei C., Benosman M., Kim, T., 2019, Online Parameter Identification for State of Power Prediction of Lithium-Ion Batteries in Electric Vehicles Using Extremum Seeking, International Journal of Control, Automation and Systems.

Gains auto-tuning for PV systems

- Wei C., Benosman M., 2016, Extremum Seeking-based Adaptive Voltage Control of Distribution Systems with High PV Penetration, IEEE Innovative Smart Grid Technologies conference, Minneapolis.

Multi-robots source seeking and trajectory planning

- Poveda J.I., Benosman M., Teel A.R., Sanfelice R.G., 2021a, Robust Coordinated Hybrid Source Seeking with Obstacle Avoidance in Multi-Vehicle Autonomous Systems, IEEE Transactions on Automatic Control, 10.1109/TAC.2021.3056365.

RF power amplifiers auto- tunning and automated design

- Kantana, C., Ma, R., Benosman, M., Komatsuzaki, Y., Yamanaka, K., A Hybrid Heuristic Search Control Assisted Optimization of Dual-Input Doherty Power Amplifier, European Microwave Conference 2021
- Cao, W., Benosman, M., Zhang, X., Ma, R., Domain Knowledge-Based Automated Analog Circuit Design with Deep Reinforcement Learning, AAAI Conference on Artificial Intelligence, February 2022 (nominated for best paper award).
- Sun Y., Benosman M., Ma R., GaN Distributed RF Power Amplifer Automation Design with Deep Reinforcement Learning, International Conference on Artificial Intelligence Circuits and Systems (AICAS) 2022 (AICAS2022 open-edges paper award).

What next?

Sentient meat by Terry Bisson's:
https://www.wnycstudios.org/podcasts/studio/segments/168264-theyre-made-out-of-meat

Learning paradigms inspired from:

- Cognitive psychology (mind)
- Neuro-science and brain physiology (brain)

e.g., See the course 'Brains, minds and machines' summer course: https://ocw.mit.edu/courses/res-9-003-brains-minds-and-machines-summer-course-summer-2015/pages/syllabus/course-instructors-guest-speakers-and-icub-team/

→ General AI ? !

MITSUBISHI ELECTRIC

Changes for the Better

http://www.merl.com