

## ***Policy Optimization for Optimal Control with Guarantees of Robustness***

**Tamer Başar**

University of Illinois Urbana-Champaign

[basar1@illinois.edu](mailto:basar1@illinois.edu) / [tamerbasar.csl.illinois.edu](http://tamerbasar.csl.illinois.edu)

### **Abstract**

Policy optimization (PO) is a key ingredient of modern reinforcement learning (RL), and can be used for efficient design of optimal controllers. For control design, certain constraints are generally enforced on the policies to be implemented, such as stability, robustness, and/or safety concerns on the closed-loop system. Hence, PO entails, by its nature, a constrained optimization in most cases, which is also nonconvex, and analysis of its global convergence is generally very challenging. Further, another element that compounds the challenge is that some of the constraints that are safety-critical, such as closed-loop stability or the H-infinity ( $H_\infty$ ) norm constraint that guarantees system robustness, can be difficult to enforce on the controller while being learned as the PO methods proceed. We have recently overcome this difficulty for a special class of such problems, which I will discuss in this talk, while also placing this in a broader context.

Specifically, I will introduce the problem of PO for  $H_2$  optimal control with a guarantee of robustness according to the  $H_\infty$  criterion, for both continuous- and discrete-time linear systems. I will argue, with justification, that despite the nonconvexity of the problem, PO methods can enjoy the global convergence property. More importantly, I will show that the iterates of two specific PO methods (namely, natural policy gradient and Gauss-Newton) automatically preserve the  $H_\infty$  norm (i.e., the robustness) during iterations, thus enjoying what we refer to as “implicit regularization” property. Furthermore, under certain conditions, convergence to the globally optimal policies features globally sub-linear and locally super-linear rates. Due to the inherent connection of this optimal robust control model to risk-sensitive optimal control and linear quadratic (LQ) dynamic games, these results also apply as a byproduct to these settings as well, with however some adjustments. The latter, in particular, entails PO with two agents, and the order in which the updates are carried out becomes a challenging issue, which I will also discuss. The talk will conclude with some informative simulations, and a brief discussion of extensions to the model-free framework and associated sample complexity analyses.

*(Based primarily on a paper to appear in SIAM J Control & Optimization, “Policy optimization for  $H_2$  linear control with  $H_\infty$  robustness guarantee: Implicit regularization and global convergence,” with Kaiqing Zhang and Bin Hu.)*

## Brief Bio

**Tamer Başar** has been with the University of Illinois Urbana-Champaign since 1981, where he currently is Swanlund Endowed Chair Emeritus and Center for Advanced Study (CAS) Professor Emeritus of Electrical and Computer Engineering, with also affiliations with the Coordinated Science Laboratory, Information Trust Institute, and Mechanical Science and Engineering. At Illinois, he has also served as Director of CAS (2014-2020), Interim Dean of Engineering (2018), and Interim Director of the Beckman Institute (2008-2010). He is a member of the US National Academy of Engineering; Fellow of IEEE, IFAC, and SIAM; a past president of the IEEE Control Systems Society (CSS), the founding president of the International Society of Dynamic Games (ISDG), and a past president of the American Automatic Control Council (AACC). He has received several awards and recognitions over the years, including the IEEE Control Systems Award (2014), ISDG Isaacs Award (2010), AACC Bellman Award (2006), IFAC Quazza Medal (2005), IEEE CSS Bode Lecture Prize (2004), and a number of recognitions from academic institutions, including the Wilbur Cross Medal (Yale, 2021), and several international honorary doctorates and professorships, with the most recent one being an honorary doctorate from KTH (Sweden, 2019). He has around 1,000 publications in systems, control, communications, optimization, networks, and dynamic games, including books on non-cooperative dynamic game theory, robust control, network security, wireless and communication networks, and stochastic networks. He was Editor-in-Chief of the IFAC Journal Automatica between 2004 and 2014, and is currently editor of several book series. His current research interests include stochastic teams, games, and networks; multi-agent systems and learning; data-driven distributed optimization; epidemics modeling and control over networks; modeling and control of spread of disinformation; security and trust; energy systems; and cyber-physical systems.

Join Zoom Meeting

<https://us02web.zoom.us/j/83734643061?pwd=YVlTQ3loVlE0bFh2RGNwYjY3cHRvQT09>

Meeting ID: 837 3464 3061

Passcode: 922602

Topic: Policy Optimization for Optimal Control with Guarantees of Robustness

Speaker: Professor Tamer Başar

Time: April 27, 2022 (8am PST, 10am CDT, 3pm UTC)

Organizers:

Tiago Roux Oliveira (TC 1.2 Chair)

Bing Chu (TC 1.2 Vice-Chair for Social Media)

