2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Nonlinear Control of Large Scale Complex Systems Using Convex Optimization Tools and Self-Adaptation

by S. Baldi, E. B. Kosmatopoulos, K. Aboudolas, D. Rovas, A. Papachristodoulou and P. A. Ioannou

Abstract— **Based on recent advances on convex design for Large-Scale Control Systems (LSCSs) and robust and efficient LSCS self-tuning/adaptation, a methodology is proposed in this paper which aims at providing an integrated LSCS-design, applicable to large-scale systems of arbitrary scale, heterogeneity and complexity and capable of:**
1) Providing stable, efficient and arbitrarily-close-to-optimal LSCS performance;
2) Being able to incorporate a variety of constraints, including limited control constraints as well as constraints that are nonlinear functions of the system controls and states;
3) Being intrinsically self-tunable, able to rapidly and efficiently optimize LSCS performance when short-, medium- or long-time variations affect the large-scale system;
4) Achieving the above, while being scalable and modular.

The purpose of the present paper is to provide the main features of the proposed control design methodology.

I. INTRODUCTION

Recently, a new *convex Approximate Optimal Control (AOC)* design – abbreviated as ConvCD (Convex Control Design) – was proposed and analyzed in [2], [4], [5]. Contrary to the existing AOC approaches, the ConvCD methodology does not require a time-consuming off-line design: *ConvCD converts the problem of constructing an AOC into a convex optimization problem* and, moreover, it can approximate with arbitrary accuracy the performance of any – feasible – stabilizing controller for the system. Additionally, the ConvCD approach *allows for its straightforward interconnection with self-tuning/adaptation tools employed to compensate for internal/external system variations and uncertainties.* The convex nature of ConvCD guarantees that such an interconnection can be performed without "getting trapped into local minima" situations and by avoiding singularities (loss-of-controllability problems).

The ConvCD approach of [2], [4], [5] employs polynomial approximations for constructing an AOC design. As a result, the price paid for its advantages mentioned above, is that its the computational burden increases exponentially with the order of the polynomial approximation used and, thus, it is not applicable to LSCS applications. Moreover, the ConvCD approach of [2], [4], [5] cannot handle state and control constraints.

In this paper, we propose and analyze a revised version of ConvCD that overcomes the shortcomings of the version presented in [2], [4], [5]. More precisely, the ConvCD approach is revised in several ways towards the development of a generic and practically implementable tool that can be used for the LSCS design of general nonlinear and uncertain large-scale nonlinear systems. The main features of the proposed approach are as follows:

- It is a convex AOC methodology which, similarly to [2], [4], [5] requires the solution of a convex problem (involving either SDP or LMIs).
- It employs a switching control scheme and, as a result, it avoids the scalability and computational problems of [2], [4], [5]. More precisely, the SDP constraints (or LMIs) involved in the new design can be of the same order as the system dimension.
- Given a user-defined optimality criterion it can approximate with arbitrary accuracy the performance of the respective optimal controller. Moreover, it provides easy-to-calculate formulas to check whether a particular choice for the design parameters of the proposed approach result in a stable and efficient controller and, most importantly, to estimate the "distance" of the resulting controller from the optimal one.
- It can handle any type of state and control constraints.
- It allows for its straightforward interconnection with self-tuning/adaptation tools employed to compensate for internal/external system variations and uncertainties. Its convex nature guarantees that such an interconnection can be performed without "getting trapped into local minima" situations. Moreover, the overall adaptation can be done by avoiding singularities (loss-of-controllability problems) and therefore computational and instability issues.

In the next sections we present the new proposed approach along with a theoretical analysis of its properties. Due to space limitations, simulation tests are not included in this

S. Baldi is with the Department of Computer Science, University of Cyprus, Nicosia, CY 1678, Cyprus, sbaldi83@gmail.com.
E. B. Kosmatopoulos is with the Dept. of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi 67100, Greece and the Informatics & Telematics Institute, Center for Research and Technology–Hellas (ITI-CERTH), 6th km Xarilaou - Thermi, 57001, Thessaloniki, Greece, kosmatop@ee.duth.gr.
K. Aboudolas is with the Informatics & Telematics Institute, Center for Research and Technology–Hellas (ITI-CERTH), 6th km Xarilaou - Thermi, 57001, Thessaloniki, Greece, aboudolas@iti.gr.
D. Rovas is with the Dept. of Production & Management Engineering, Technical University of Crete, Chania, 73100, Greece and the Informatics & Telematics Institute, Center for Research and Technology–Hellas (ITI-CERTH), 6th km Xarilaou - Thermi, 57001, Thessaloniki, Greece, rovas@dpem.tuc.gr.
A. Papachristodoulou is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, U.K., antonis@eng.ox.ac.uk
P. A. Ioannou is with the Department of Electrical Engineering-Systems, Los Angeles, CA 90089-2562 and the Department of Computer Science, University of Cyprus, Nicosia, CY 1678, Cyprus, ioannou@usc.edu
The research leading to these results has been partially funded by the European Commission FP7-ICT-2007-9.6.3, Energy Efficiency under the contract #248537 (PEBBLE) and European Commission FP7-ICT-5-3.5, Engineering of Networked Monitoring and Control Systems, under the contract #257806 AGILE.

paper and will be reported elsewhere.

II. PROBLEM FORMULATION

In this paper, we consider the problem of designing a state feedback controller for a general nonlinear system which assumes the following dynamics

$$\dot{\chi} = F(\chi, u) \tag{1}$$

where $\chi \in \Re^n, u \in \Re^m$ denote the vectors of system states and control inputs, respectively and F is a nonlinear vector function (assumed to be continuous). The problem at hand is to construct a state-feedback controller

$$u = k^*(\chi)$$

which renders the closed-loop system *stable* and moreover, solves the following *constrained optimal control* problem:

$$\begin{aligned}
\min_u \bar{J} \quad &= \quad \int_0^\infty \bar{\Pi}(\chi(s), u(s))ds && (2)\\
s.t. \quad & \quad \text{system (1) dynamics} && \\
& \quad u_{min} \leq u \leq u_{max} && (3)\\
& \quad \bar{C}(\chi, u) \leq 0 && (4)
\end{aligned}$$

where $\bar{\Pi}$ is a bounded-from-below, continuous function of its arguments; u_{min}, u_{max} denote the vectors of minimum and maximum, respectively, allowable control signals; $\bar{C}$ is a smooth nonlinear vector function. Please note that inequality (4) can represent a large class of constraints met in practical applications. Without loss of generality, we will assume that the minimum of $\bar{\Pi}(\chi, u)$ is attained at $\bar{\Pi}(0, 0)$ and that all constraints are satisfied for $(\chi, u) = (0, 0)$.

In order to have a well-posed problem we will assume that the controller solving the optimal control problem (2)-(4) provides closed-loop stability, i.e., we will assume that

(A1) Let $u^* = \bar{k}(\chi)$ be the controller that solves the optimal control problem (2)-(4). Then, for all admissible initial states $\chi(0)$, the closed-loop system (1) under the feedback $u^* = \bar{k}(\chi)$ is stable.

Additionally to (A1) we will impose the following assumption.

(A2) We have that
$$\bar{C}(0, 0) < 0$$

Assumption (A2) states that at the point $(\chi, u) = (0, 0)$ where the function $\bar{\Pi}(\cdot)$ attains its minimum none of the constraints (4) is active.

III. TRANSFORMATIONS & APPROXIMATIONS

The first step in the proposed approach is to impose special transformations that render the system dynamics, the constraints and the objective criterion in an appropriate format that is convenient for our developments. To do so, we define a new fictitious "control" input v that is calculated according to

$$\dot{\bar{u}} = v, \quad u = S(\bar{u}) \tag{5}$$

with $S(\cdot)$ being a smooth and invertible function such that

$$u_{min} \leq S(\bar{u}) \leq u_{max} \tag{6}$$

It is worth noticing that by adding the integrator (5) and introducing the function S satisfying (6), the problem of designing the control vector u so that constraint (3) is satisfied is transformed into the problem of designing v that does not need to satisfy a boundedness constraint like (3): while v can take "arbitrary" values, the actual control vector u is restricted – due to the use of function S – to satisfy (3).

By defining the *augmented state vector* x according to

$$x = \begin{bmatrix} \chi \\ \bar{u} \end{bmatrix}$$

we can rewrite the system (1) and the constraints (4) as follows:

$$\dot{x} = f(x) + Bv \tag{7}$$

$$C(x) \leq 0 \tag{8}$$

The second step in the proposed design is to approximate the transformed system dynamics (7) as well as the objective criterion using smooth mixing signals. More precisely, we let $\beta_i, i = 1, \ldots, L$ denote a set of smooth *mixing signals* that satisfy the following properties:

$$\beta_i(x) \in [0, 1], \quad \sum_{i=1}^{L} \beta_i(x) = 1, \forall x, \quad \sum_{i=1}^{L} \mathcal{I}\left(\beta_i(x)\right) \leq 2$$

where $\mathcal{I}(y)$ denotes the indicator function $\mathcal{I}(y) = 1$ if $y > 0$ and $\mathcal{I}(y) = 0$ if $y = 0$. We will, moreover, assume that the mixing signals β_i span the whole space x lies on, i.e., we will assume that for each possible x, there exists at least one i such that $\mathcal{I}\left(\beta_i(x)\right) = 1$. For an example of mixing functions, see [6], [7]. Using the above design considerations for the mixing signals β_i, we can employ standard function approximation techniques to approximate the system dynamics (7) and the objective criterion (2) as[1] follows:

$$\dot{x} \approx \sum_{i=1}^{L} \beta_i(x) A_i \bar{x}(x) + Bv \tag{9}$$

$$\bar{\Pi}(\chi, u) \equiv \Pi(x) \approx \sum_{i=1}^{L} \beta_i(x) \left(\bar{x}^\tau Q_i \bar{x}\right) \tag{10}$$

where A_i, Q_i are constant matrices with Q_i being positive semidefinite,

$$\bar{x}(x) = \begin{bmatrix} \varpi(x) \\ \sigma(x) \end{bmatrix}$$

with $\varpi(x)$ being any smooth nonlinear function satisfying $\varpi(x) = 0 \Longleftrightarrow x = 0$, and $\sigma(x)$ is defined according to

$$\sigma_i(x) = \exp(\alpha C_i(x) - \eta) \tag{11}$$

and α is a large positive constant and η is a positive constant chosen so that if particular constraint $C_i(x) \leq 0$ is – or is about to be – violated, then the respective $\sigma_i(x)$ takes a very large value, while it is negligible when the constraint is satisfied.

[1] Note that the approximation (9) does not contain *constant terms*. All the results of this paper can be readily extended to the case where an approximation with constant terms is employed.

Moreover, let us define the vector $z(x)$ as follows:

$$z(x) = \begin{bmatrix} \sqrt{\beta_1(x)}\bar{x}(x) \\ \vdots \\ \sqrt{\beta_L(x)}\bar{x}(x) \end{bmatrix} \tag{12}$$

Using the above definition for $z(x)$ we finally end up with the following description for the system dynamics

$$\dot{x} \approx \bar{\Phi}(x)z(x) + Bv \tag{13}$$

where $\bar{\Phi}(x) = \begin{bmatrix} \sqrt{\beta_1(x)}A_1 & \dots & \sqrt{\beta_L(x)}A_L \end{bmatrix}$. As a final step, the constrained optimization problem (2)-(4) is transformed into an *unconstrained* one, by incorporating the constraints (4) – or, equivalently, the constraints (8) – as penalty functions into the objective function. This is made possible by making use of (11) and by replacing the objective function (2) by the following one

$$J = \int_0^\infty \Big(\sum_{i=1}^{L} \beta_i(x(s))\left(\bar{x}(s)^\tau Q_i \bar{x}(s)\right) + \sum_{j=1}^{\dim(C(x))} \sigma_j^2(x(s))\Big) ds$$

By using the fact that $\sum_{i=1}^{L} \beta_i(x) = 1$ and the definition of z in (12), the augmented objective function J can be written in the following compact form:

$$J = \int_0^\infty \left(z^\tau(s)Qz(s)\right) ds \tag{14}$$

with Q being a block diagonal matrix.

A. HJB and Controller Approximations

After applying all transformations presented of the previous section, we have that the optimal state feedback design problem can be cast as an unconstrained optimal control problem of the form

$$\text{minimize } J = \int_0^\infty \left(z^\tau(s)Qz(s)\right) ds \tag{15}$$

subject to

$$\dot{x} = \bar{\Phi}(x)z(x) + Bv + \nu \tag{16}$$

where ν stands for the approximation error due to the replacement of the actual system dynamics (7) by (9).

Two remarks are in order.

Remark 1: Additionally to (A1) and (A2) we will assume that

(A3) *The initial states $x(0)$ of the system belong to a compact subset $\mathcal{X}_0 \subset \Re^{\dim(x)}$.*

Associated to the subset $\mathcal{X}_0$ of admissible initial conditions we will define a "sufficiently large" compact subset $\mathcal{X}$ which contains $\mathcal{X}_0$. △

Remark 2: Please note that in case x lies in the compact subset $\mathcal{X} \subset \Re^{\dim(x)}$, then the amplitude of the term $\nu \equiv \nu(x)$ can become arbitrarily small and is inversely proportional to the number L of mixing signals. △

Application of the well-known Hamilton-Jacobi-Bellman (HJB) equation to the above problem results in the following equation

$$-z^\tau(x)Qz(x) = \frac{\partial V}{\partial x}^\tau(x)\left(\bar{\Phi}(x)z(x) + Bv^* + \nu\right) \tag{17}$$

where V is the optimal-cost-to-go function, i.e.,

$$V(x(t)) = \int_t^\infty \left(z^\tau(s)Qz(s)\right) ds,$$

v^* denotes the optimal control and ν is the error approximation term that is due to the approximations of the previous sections.

Lemma 1: Let (A1)-(A3) hold and assume that $x \in \mathcal{X}$. The optimal-cost-to-go function V can be approximated – with accuracy $\mathcal{O}(1/L)$ – using a Sum-of-Squares (SoS) polynomial as follows:

$$V(x) \approx z^\tau(x)\mathbf{P}z(x) \tag{18}$$

where $\mathbf{P}$ is a constant *positive definite matrix* with $\mathbf{P}$ being symmetric and having the following *block diagonal form*

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \mathbf{P}_L \end{bmatrix} \tag{19}$$

where $\mathbf{P}_i$ are $\dim(\bar{x})^2$-dimensional *symmetric and positive definite matrices*.

Proof: The proof can be established using similar arguments as in [2]-[5]. ■

Similar to the use of approximations for the optimal cost-to-go function, we approximate the optimal controller v^* as follows:

$$v^* \approx \sum_{i=1}^{L} \beta_i(x)\mathbf{G}_i z(x) \tag{20}$$

where $\mathbf{G}_i$ are constant matrices. Please note that (20) can be rewritten in the following compact form

$$v^* \approx \Gamma(x)\mathbf{G}z(x) \tag{21}$$

where $\Gamma(x) = [\beta_1(x)I, \dots, \beta_L(x)I]$.

IV. The ConvCD Approach

Using the approximations (13), (18) and (21), the HJB Equation (17) can be written as follows:

$$\begin{aligned} 0 = \; & z^\tau\Big(\left[\bar{\Phi}(x) + B\Gamma(x)\mathbf{G}\right]^\tau M_z^\tau \mathbf{P} \\ & + \mathbf{P}M_z\left[\bar{\Phi}(x) + B\Gamma(x)\mathbf{G}\right] + Q\Big)z - \bar{\nu} \equiv \mathcal{G}_{\mathbf{P},\mathbf{G}}(x) - \bar{\nu} \end{aligned} \tag{22}$$

where $M_z \equiv M_z(x)$ denotes the matrix whose (i, j)th entry is given by $M_{z,ij}(x) = \partial z_i(x)/\partial x_j$ and $\bar{\nu}$ is the *approximation error term that is inversely proportional to the number L of mixing signals, resulting from the approximations (13), (18) and (21)*. Moreover, as the optimal cost-to-go function $V(x)$ is a CLF for the system (16) we have that closed-loop stability is preserved by the control scheme if $\dot{V} < 0$ for $x \neq 0$, or, equivalently using the approximations (13), (18) and (21) if the following inequality holds

$$\begin{aligned} & z^\tau\Big(\left[\bar{\Phi}(x) + B\Gamma(x)\mathbf{G}\right]^\tau M_z^\tau \mathbf{P} + \mathbf{P}M_z\left[\bar{\Phi}(x) + B\Gamma(x)\mathbf{G}\right]\Big)z \\ \equiv \; & \mathcal{L}_{\mathbf{P},\mathbf{G}}(x) < 0, \forall x \notin \mathcal{B}(\bar{\nu}) \end{aligned} \tag{23}$$

where[2] $\mathcal{B}(\bar{\nu})$ denotes a ball centered at the origin and having radius proportional to $\bar{\nu}$.

Equations (22) and (23) indicate that – provided the approximation error term ν is "small enough" – it suffices to choose $\mathbf{P}, \mathbf{G}$ so that the term $\mathcal{G}_{\mathbf{P},\mathbf{G}}$ is as small as possible subject to the constraint that $\mathcal{L}_{\mathbf{P},\mathbf{G}}(x)$ is – almost – negative definite. In other words, the problem of constructing an approximately optimal performance can be cast as the following optimization problem:

$$\begin{aligned} &\min \|\mathcal{G}_{\mathbf{P},\mathbf{G}}(x)\|^2 + \gamma \\ &\text{s.t. } \mathbf{P} \succ 0, \quad \mathcal{L}_{\mathbf{P},\mathbf{G}}(x) \leq \gamma, \quad \gamma \geq 0 \end{aligned} \tag{24}$$

Unfortunately, as the above optimization problem is nonlinear wrt the unknowns $\mathbf{P}, \mathbf{G}$, attempting to solve (24) is <u>non-convex</u> – and thus difficult to solve – problem even in the case where the approximation-related term $\bar{\nu}$ is negligible. To circumvent this problem we work similarly to [8], [2]-[5]: by multiplying by $\mathbf{P}^{-1}$ from the left and the right the terms inside the parenthesis of (22) we obtain that

$$\begin{aligned} \mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\mathbf{Q}}}(x) &\stackrel{\triangle}{=} z^\tau \big(\big[\bar{\mathbf{P}}(\bar{\Phi}(x))^\tau + \mathbf{F}^\tau \Gamma^\tau(x) B^\tau\big] M_z^\tau \\ &\quad + M_z \big[\bar{\Phi}(x)\bar{\mathbf{P}} + B\Gamma(x)\mathbf{F}\big] + \bar{\mathbf{Q}}\big) z \\ &= \bar{\bar{\nu}} \end{aligned} \tag{25}$$

where $\bar{\bar{\nu}} = \mathcal{O}(\bar{\nu})$,

$$\bar{\mathbf{P}} \stackrel{\triangle}{=} \mathbf{P}^{-1}, \quad \bar{\mathbf{Q}} \stackrel{\triangle}{=} \bar{\mathbf{P}} Q \bar{\mathbf{P}} \equiv \mathbf{P}^{-1} Q \mathbf{P}^{-1} \tag{26}$$

and $\mathbf{F}$ is a matrix satisfying

$$\mathbf{F} = \mathbf{G}\bar{\mathbf{P}} \tag{27}$$

Working similarly on (23) we obtain

$$\begin{aligned} \mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}(x) &\stackrel{\triangle}{=} z^\tau \big(\big[\bar{\mathbf{P}}(\bar{\Phi}(x))^\tau + \mathbf{F}^\tau \Gamma^\tau(x) B^\tau\big] M_z^\tau \\ &\quad + M_z \big[\bar{\Phi}(x)\bar{\mathbf{P}} + B\Gamma(x)\mathbf{F}\big]\big) z < 0, \forall x \notin \mathcal{B}(\bar{\nu}) \end{aligned} \tag{28}$$

The above transformations play a crucial role in our approach: instead of attempting to solve the non-convex optimization problem (29), we solve its equivalent version – which is convex – that involves the functions $\mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\mathbf{Q}}}$ and $\mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}$:

$$\begin{aligned} &\min \left\|\mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\mathbf{Q}}}(x)\right\|^2 + \gamma \\ &\text{s.t. } \epsilon_1 I \preceq \bar{\mathbf{P}} \preceq \epsilon_2 I, \quad \epsilon_3 Q \preceq \bar{\mathbf{Q}}, \quad \mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}(x) < \gamma, \quad \gamma \geq 0 \end{aligned} \tag{29}$$

where $\epsilon_i, i = 1, 2, 3$ are some positive design constants (with $\epsilon_2 > \epsilon_1$) and $\bar{\mathbf{P}}^*, \bar{F}^*, \bar{Q}^*$ denote the optimal solutions to the optimization problem (29).

Despite the fact that the optimization problem (29) is a convex problem, its solution requires discretization of the state-space as it is an *infinite-dimensional, state-dependent problem*. Fortunately, due to the particular form of (29), the number of discretization points does not have to be as large as it is required in a typical state-dependent optimization problem: as it is seen in Theorem 1 presented below, the number of discretization points can be as few as the total number of free variables in the matrices $\bar{\mathbf{P}}, \bar{\mathbf{Q}}, \mathbf{F}$. This is contrary to alternative AOC approaches that require an extremely larger number of discretization points. Also, other approaches that use SOS representations to solve (29) are computationally expensive and cannot treat systems of high complexity directly [8].

Table I presents the proposed procedure for solving the optimization problem (29) and, eventually, constructing the proposed control design scheme. Following the same methodology as in [2]-[5], the key idea of the approach in Table I for solving (29) is to choose randomly many different x; it suffices to choose the number of random x to be equal or larger than the number of free variables in the matrices $\bar{\mathbf{P}}, \bar{\mathbf{Q}}, \mathbf{F}$.

The next theorem establishes the properties of the overall scheme presented in Table I.

Theorem 1: Fix the number L of mixing signals and the constants $\epsilon_i, i = 1, 2, 3$ and let $\hat{\mathbf{P}}, \hat{\mathbf{G}}$ be constructed according to the design procedure of Table I and let (A1)-(A3) hold. Let also $\mathcal{M}$ be any positive integer satisfying $\mathcal{M} \geq N$, where N – see also Table I – denotes the number of free variables of the matrices $\bar{\mathbf{P}}, \bar{\mathbf{Q}}, \mathbf{F}$. Select[3] randomly $\mathcal{M}$ points $x^{[j]} \in \mathcal{X}$ and let

$$z^{[j]} = z(x^{[j]})$$

Finally, suppose that the positive design constants ϵ_i are chosen so that

$$\epsilon_1 I \preceq \mathbf{P}^{*-1} \preceq \epsilon_2 I, \quad \epsilon_3 Q \preceq \mathbf{P}^{*-1} Q \mathbf{P}^{*-1} \tag{32}$$

where $\mathbf{G}^*, \mathbf{P}^*$ denote the optimal solutions to the optimization problem (24).

Then, the following statements hold:

(a) If for some $C_1 > 0$ and for all $j \in \{1, \ldots, \mathcal{M}\}$,

$$\|z^{[j]}\| > C_1 \Rightarrow \mathcal{L}_{\hat{\mathbf{P}},\hat{\mathbf{F}}}(x^{[j]}) < 0 \tag{33}$$

Then, the closed-loop system is stable and its solutions converge to the subset

$$\mathcal{D} = \left\{ x : \|z(x)\| \leq C_1 + \mathcal{O}\left(\frac{1}{\mathcal{M}}\right) \right\}$$

Moreover, the closed-loop solutions of the system satisfy

$$\|x(t) - x_{\text{opt}}\| \leq C_2 + \mathcal{O}\left(\frac{1}{\mathcal{M}}\right) \tag{34}$$

where x_{opt} denotes the closed-loop solutions of the system under the optimal control v^* and C_2 is a nonnegative constant satisfying

$$C_2 = \mathcal{O}\left(\mathcal{G}_{\hat{\mathbf{P}},\hat{\mathbf{F}}}(x^{[j]})\right)$$

(b) For each $C_1 > 0, C_2 \geq 0$, there exists a lower bound on the approximators size $\bar{L}$ so that (33) and (34) hold for all choices of approximators' size L satisfying $L \geq \bar{L}$. $\diamondsuit$

[2]Please note that throughout this paper the notation $\mathcal{B}(\cdot)$ is used in a similar way as the notation $\mathcal{O}(\cdot)$: $\mathcal{B}(a)$is used to denote *a ball of radius proportional to a and not a ball of radius a.*

[3]Note that although it is possible for $\mathcal{N}$ and $x^{[i]}$ of Table I to concide with $\mathcal{M}$ and $x^{[j]}$, respectively, it is advisable that they do not concide so that the performance of the proposed control design is evaluated at different state points than the ones used to construct the ConvCD solution.

Table I: The ConvCD Approach

Step 1. Calculate the matrices $\bar{\mathbf{P}}, \bar{\mathbf{Q}}, \mathbf{F}$ as follows: Let N denote the total number of free variables of these matrices. Select randomly $\mathcal{N}$ points $x^{[i]} \in \mathcal{X}$, where $\mathcal{N}$ is any integer satisfying $\mathcal{N} \geq N$ and solve the following convex optimization problem (here ϵ_i are user-defined positive constants):

$$\begin{aligned} &\min \sum_{i=1}^{\mathcal{N}} \|\mathcal{F}_{\bar{\mathbf{P}},\bar{\mathbf{Q}},\mathbf{F}}(x^{[i]})\|^2 + \gamma \\ &\text{s.t. } \epsilon_1 I \preceq \bar{\mathbf{P}} \preceq \epsilon_2 I, \quad \epsilon_3 Q \preceq \bar{\mathbf{Q}}, \quad \mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}(x^{[i]}) < \gamma, \quad \gamma \geq 0 \end{aligned} \tag{30}$$

Step 2. By using the solution of the above optimization problem, we can extract the estimates of the matrices $\mathbf{P}, \mathbf{G}$ in (24) according to

$$\hat{\mathbf{P}} = \bar{\mathbf{P}}^{-1}, \quad \hat{\mathbf{G}} = \mathbf{F}\bar{\mathbf{P}}^{-1} \tag{31}$$

Step 3. The proposed control scheme is the MMCM controller given by (20) and (5) by setting $\mathbf{G}$ equal to $\hat{\mathbf{G}}$.

Proof: Let $\bar{\bar{\mathbf{Q}}} == \bar{Q} - \epsilon_3 Q$ and consider the following two optimization problems:

$$\begin{aligned} &\min \left\| \mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\bar{\mathbf{Q}}}+\epsilon_3 Q}(x) \right\|^2 + \gamma \\ &\text{s.t. } \epsilon_1 I \preceq \bar{\mathbf{P}} \preceq \epsilon_2 I, \quad 0 \preceq \bar{\bar{\mathbf{Q}}}, \quad \mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}(\mathbf{x}) < \gamma, \quad \gamma \geq \mathbf{0} \end{aligned} \tag{35}$$

and

$$\begin{aligned} &\min \sum_{i=1}^{\mathcal{N}} \|\mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\bar{\mathbf{Q}}}+\epsilon_3 Q}(x^{[i]})\|^2 + \gamma \\ &\text{s.t. } \epsilon_1 I \preceq \bar{\mathbf{P}} \preceq \epsilon_2 I, \quad 0 \preceq \bar{\bar{\mathbf{Q}}}, \quad \mathcal{H}_{\bar{\mathbf{P}},\mathbf{F}}(x^{[i]}) < \gamma, \quad \gamma \geq 0 \end{aligned} \tag{36}$$

The optimization problems (35) and (36) are equivalent to the optimization problems (29) and (30), respectively. From now on, we will consider instead of the optimization problems (29) and (30), their equivalent ones (35) and (36).

Let ϑ denote a vector which contains all non-zero entries of the matrices $\bar{\mathbf{P}}, \bar{\bar{\mathbf{Q}}}, \mathbf{F}$; as the first two of these matrices are symmetric, the dimension of the vector ϑ is less than the number of non-zero entries of the matrices $\bar{\mathbf{P}}, \bar{\bar{\mathbf{Q}}}, \mathbf{F}$. Similarly, let θ denote a vector which contains all non-zero entries of the matrices $\mathbf{P}$ and $\mathbf{G}$. Let also ϑ^* and θ^* denote the global optimizers of the infinite-dimensional problems (35) and (24), respectively. Finally, let $\mathcal{T}$ denote the one-to-one transformation from ϑ to θ defined according to $\mathbf{P} = \bar{\mathbf{P}}^{-1}, \mathbf{G} = \mathbf{F}\bar{\mathbf{P}}^{-1}$. Then, it is not difficult to see that if (32) holds then

$$\theta^* = \mathcal{T}(\vartheta^*) + \mathcal{O}(\bar{\nu}) \tag{37}$$

Using the definition of ϑ we have that

$$\mathcal{F}_{\bar{\mathbf{P}},\mathbf{F},\bar{\bar{\mathbf{Q}}}+\epsilon_3 Q}(x) = \vartheta^\tau \Psi(x) + \epsilon_3 z(x)^\tau Q z(x)$$

for some appropriately defined functions $\Psi_i, \;\; i = 1, 2$. Moreover, from (25) we have that

$$\vartheta^{*\tau} \Psi(x) + \epsilon_3 z(x)^\tau Q z(x) = \mathcal{O}(\bar{\nu}) \tag{38}$$

Let

$$\bar{\bar{\Psi}} = \sum_{i=1}^{\mathcal{N}} \Psi^\tau(x^{[i]}) \Psi(x^{[i]})$$

and

$$Z = -\epsilon_3 \sum_{i=1}^{\mathcal{N}} \Psi^\tau(x^{[i]}) z^\tau(x^{[i]}) Q z(x^{[i]})$$

Then, by using (38) we directly obtain

$$\bar{\bar{\Psi}} \vartheta^* = Z + \mathcal{O}(\bar{\nu}) \tag{39}$$

We will show that the matrix $\bar{\bar{\Psi}}$ is full-rank with probability 1. The proof of the last claim can be established using similar arguments as in [3], [4]: if $\bar{\bar{\Psi}}$ is not full-rank then there exists a constant non-zero vector b such that

$$b_i^\tau \Psi(x^{[i]}) = 0, \forall i \in \{1, \ldots, \mathcal{N}\} \tag{40}$$

which is equivalent to a set of nonlinear system of equations in $x^{[i]}$. The set of solutions to these system of equations has zero Lebesgue measure, see e.g. [3], [4], and as a result under a random choice for $x^{[i]}$ the probability that (40) holds is zero. In other words, a random choice for $x^{[i]}$ guarantees that $\bar{\bar{\Psi}}$ is full-rank with probability 1.

Let $\hat{\vartheta}$ denote the solution of the ConvCD optimization problem (36) and

$$\mathcal{J}(\vartheta) = \|\bar{\bar{\Psi}} \vartheta - Z\|^2$$

Since ϑ^* is a feasible solution to the optimization problem (36) and by using (39), we have that

$$\mathcal{J}(\hat{\vartheta}) \leq \mathcal{J}(\vartheta^*) = \mathcal{O}(\bar{\nu}) \tag{41}$$

As a result, since $\mathcal{J}(\cdot)$ is a quadratic function, we have that

$$\begin{aligned} \mathcal{J}(\hat{\vartheta}) \;\; &= \;\; \mathcal{J}(\vartheta^*) + 2\left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \left(\bar{\bar{\Psi}} \vartheta^* + Z\right) \\ &\quad + \left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \bar{\bar{\Psi}} \left(\vartheta^* - \hat{\vartheta}\right) \end{aligned} \tag{42}$$

Combining (41) and (42) we obtain

$$\begin{aligned} \mathcal{J}(\vartheta^*) \;\; &\geq \;\; \mathcal{J}(\vartheta^*) + 2\left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \left(\bar{\bar{\Psi}} \vartheta^* + Z\right) \\ &\quad + \left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \bar{\bar{\Psi}} \left(\vartheta^* - \hat{\vartheta}\right) \Rightarrow \\ 0 \;\; &\geq \;\; 2\left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \left(\bar{\bar{\Psi}} \vartheta^* + Z\right) \\ &\quad + \left(\vartheta^* - \hat{\vartheta}\right)^\tau \bar{\bar{\Psi}}^\tau \bar{\bar{\Psi}} \left(\vartheta^* - \hat{\vartheta}\right) \end{aligned}$$

Using the above inequality together with the facts that (a) $\bar{\bar{\Psi}}$ is full-rank and thus $\Psi^\tau \Psi \succ 0$ and (b) and $\|\bar{\bar{\Psi}} \vartheta^* - Z\| = \mathcal{O}(\bar{\nu})$ we finally obtain that

$$\hat{\vartheta} = \vartheta^* + \mathcal{O}(\bar{\nu}) \tag{43}$$

The above equation in combination with (37) establishes that

$$\theta^* = \mathcal{T}(\hat{\vartheta}) + \mathcal{O}(\bar{\nu}) \tag{44}$$

Combining (44) with (22) we finally obtain that

$$\mathcal{G}_{\hat{\mathbf{P}},\hat{\mathbf{G}}}(x) = \mathcal{O}(\bar{\nu}) \tag{45}$$

or, equivalently

$$\mathcal{L}_{\hat{\mathbf{P}},\hat{\mathbf{G}}}(x) \leq -z^\tau Q z + \mathcal{O}(\bar{\nu}) \tag{46}$$

where $\hat{\mathbf{P}}, \hat{\mathbf{G}}$ are generated using the ConvCD optimization algorithm (30) [or, its equivalent, (36)]. Inequality (46) can be rewritten as follows:

$$\frac{d}{dt}\left(z^\tau \hat{\mathbf{P}} z\right) \leq -z^\tau Q z + \mathcal{O}(\bar{\nu})$$

or, equivalently, by setting $\breve{V}(x) = z^\tau \hat{\mathbf{P}} z$, as follows

$$\dot{\breve{V}}(x) \leq -c\breve{V}(x) + \mathcal{O}(\bar{\nu})$$

for some positive constant c, or, equivalently, [1]

$$\begin{aligned} \breve{V}(x(t)) &\leq \exp(-ct)\breve{V}(x(0)) + \mathcal{O}(\bar{\nu}) \\ \|x(t)\| &\leq \sqrt{\frac{k_2}{k_1}} \exp\left(-\frac{k_3}{2k_2}t\right)\|x(t)\| + \mathcal{O}(\bar{\nu}) \end{aligned} \tag{47}$$

where k_i are positive constants satisfying $k_1\|x\|^2 \leq \breve{V}(x) \leq k_2\|x\|^2, c\breve{V}(x) \geq k_3\|x\|^2$. The rest of the proof is based on standard Lyapunov stability arguments and is ommitted due to space limitations. ■

Several remarks are in order:

• The optimization problem (29) is a *convex* one: the optimization criterion comprises a quadratic function with respect to the decision variables $\bar{\mathbf{P}}, \mathbf{F}, \bar{\mathbf{Q}}$ while all of the constraints are Semi-Definite constraints (and thus convex).

• As it seen in the proof of Theorem 1, the optimization problem (29) is equivalent to (24) and thus, if the approximation error term is small enough, then the solution to (29) corresponds – approximately – to the optimal solution of the problem at hand.

• Most importantly, according to Theorem 1, *the solution to the optimization problem (29) is able to provide an efficient control design even in cases where the approximation error is not negligible*. Theorem 1 provides an easy-to-calculate formula – see relation (33) – to check whether a particular choice for the approximators size L provides the required controller efficiency. In other words, even in cases where the particular choice for L is far from providing a close-to-the-optimal performance (i.e., a significantly larger L – and thus a significantly more complicated controller – is required to get a close-to-optimal performance), the proposed scheme provides a control design that is efficient.

Furthermore to Theorem 1 and by using the same arguments as those of [2] it can be seen that in case (33) holds, then the solutions of the overall system satisfy the following inequality:

$$\begin{aligned} \|z(t)| &\leq \alpha_1 \exp^{-\alpha_2 t}\|z(0)\| + \alpha_3 \\ \alpha_1 &= \sqrt{\frac{\epsilon_2}{\epsilon_1}},\ \alpha_2 = \left(\frac{\epsilon_3}{2\epsilon_1} - \frac{\epsilon_1 + \mathcal{O}(1/L)}{2\epsilon_2^2}\right) \\ \alpha_3 &= C = \mathcal{O}(1/L) \end{aligned} \tag{48}$$

What is important about (48) is that the design constants ϵ_i in the optimization problem (36) can serve as tuning/design parameters in a similar fashion as e.g., the LQ matrices in Linear-Quadratic control design applications: *(48) can be used to evaluate the effects and trade-offs of different choices for ϵ_i on the overshoot, convergence and steady-state closed-loop performance and thus it can provide a guide on how to choose ϵ_i so that the desired performance is obtained.*

V. Inducing Adaptation within ConvCD

As ConvCD assumes perfect knowledge of the system dynamics it may become inefficient in cases of system uncertainties or variations or, even worse, in cases where minor or major faults or anomalies affect the system dynamics. Thus in order for ConvCD to be practically efficient, an adaptive self-tuning/re-design tool is required that will take care of all the above-mentioned factors that may affect the efficiency of ConvCD. This will be a topic of future research.

By exploiting the convex nature of ConvCD, an adaptive tuning scheme has been developed in the past [3], [4], [5] which overcomes the severe shortcomings of existing adaptive and self-tuning schemes of poor transient performance and controller instability or failure in cases of loss-of-controllability. The adaptive tuning scheme [3], [4], [5] is applicable to the revised ConvCD methodology presented in this paper: application of this adaptive scheme results in the same stability and convergence properties as the ones reported in [3], [4], [5]. The interested reader is referred there for more details.

References

[1] H. K. Khalil, *Nonlinear Systems*, New York: Macmillan, 1992.

[2] E. B. Kosmatopoulos, CLF-based Control Design with Good Transient Performance for Known and Unknown Multi-Input Nonlinear Systems, *Proceedings of the European Control Conference*, Budapest, Hungary, 2009.

[3] E. B. Kosmatopoulos, Adaptive Control Design based on Adaptive Optimization Principles, *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2680-2685, 2008.

[4] E. B. Kosmatopoulos, "Control of unknown nonlinear systems with efficient transient performance using concurrent exploitation and exploration," *IEEE Transactions on Neural Networks*, vol. 21, no. 8, pp. 1245-1261, Aug. 2010.

[5] E. B. Kosmatopoulos, "CLF-based control design for unknown multi-input nonlinear systems with good transient performance," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, Nov. 2010.

[6] M. Kuipers and P. A. Ioannou, Multiple Model Adaptive Control With Mixing, *IEEE Transactions on Automatic Control* vol. 55, no. 8, pp. 1822-1836, 2010.

[7] M. Kuipers and P. Ioannou, Multiple Model Adaptive Control With Mixing: A Pedagogical Example, *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, pp. 3239-3244, December 2008.

[8] S. Prajna, A. Papachristodoulou, and F. Wu, " Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach," *Proceedings of the Asian Control Conference*, pp. 157-165, 2004.