OPTIMIZING CONTROL OF OVER-ACTUATED LINEAR SYSTEMS WITH NONLINEAR OUTPUT MAPS VIA CONTROL LYAPUNOV FUNCTIONS

Tor A. Johansen[†] and Daniel Sbárbaro[‡]

[†]Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.
[‡]Department of Electrical Engineering, University of Concepción, Concepción, Chile.

Keywords: Control Lyapunov functions, Wiener systems, optimization, constraints, over-actuated systems.

Abstract

Over-actuated linear systems with nonlinear output maps are studied, taking into account input and state constraints. Using a control Lyapunov function approach, we develop an optimizing dynamic controller that includes a dynamic reference feed-forward. This allows optimizing control to be implemented with low real-time computational complexity since the optimization algorithm converges only asymptotically. Conditions for local convergence and global asymptotic stability are established. A highly nonlinear colorant mixing simulation example is used to illustrate the approach.

1 Introduction

Consider the system

$$\dot{x} = Ax + Bu \quad (1)$$
$$y = g(x) \quad (2)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$ and $y \in \mathbb{R}^m$. It is assumed that g is a continuously differentiable nonlinear function, (A, B) is controllable, $m \leq r$, and $m \leq n$. A characteristic feature of the over-actuated control problem is that for a given reference output y^* the corresponding equilibrium point is not unique. In this paper, it will be specified implicitly through a steady-state optimization criterion, while the dynamic performance is specified through a standard LQ criterion. This is useful in problems where the multi-variable output map g is not directly invertible or the system is over-actuated, such that $r > m$ and explicit solutions are hard to compute.

Computer processing capacity limits how quickly an iterative optimization algorithm converges in a real-time implementation. We assume the optimization convergence can not be made fast compared to the closed-loop dynamics such that the steady-state optimization is a limiting factor for the control performance. Rather than solving the steady-state and dynamic optimization problems separately, we take a Lyapunov-design approach to solve the steady-state optimization problem simultaneously with the dynamic optimal control problem, borrowing some ideas from adaptive and nonlinear control [11, 8]. This leads to a nonlinear dynamic controller that implicitly solves the steady-state optimization problems asymptotically via the real-time integration of an ordinary differential equation. Effectively, the approach cancels transients related to the interaction between optimizer and feedback controller using a reference feed-forward. The optimization approach here has some conceptual resemblance to Lyapunov-based optimization [13, 12], and the approach provides an alternative control method for systems of the Wiener class, as treated recently in [6, 3, 10]. A related approach was suggested in [4].

2 Control specification

Let (ξ, μ) be an equilibrium point for the system (1), i.e. a pair of vectors satisfying

$$0 = A\xi + B\mu \quad (3)$$

In general, there exists a vector $\theta \in \mathbb{R}^p$ that parameterizes the equilibrium manifold such that all solutions to the homogeneous linear equation system (3) are given by

$$\xi = F\theta, \qquad \mu = G\theta \quad (4)$$

Since the $(n+r)$-dimensional vector (ξ, μ) must lie in the null-space of the $n \times (n+r)$-matrix $(A|B)$, one may select any p-dimensional basis for this space to build rows of F and G, with $p = n + r - \text{rank}(A|B) \geq r$. Our objective is to choose (ξ, μ) such that they minimize some steady-state performance criterion $J'(\xi, \mu; y^*)$ with $y^* \in \mathbb{R}^m$ given. Assume this criterion is given in the form

$$J'(\xi, \mu; y^*) = \frac{1}{2} (y^* - g(\xi))^T W (y^* - g(\xi)) + J_0'(\xi, \mu) \quad (5)$$

where J_0' is assumed to be continuously differentiable and positive semi-definite, and $W \succ 0$ is a symmetric weighting matrix. Due to the parameterization (4) we define the re-parameterized criterion function

$$J(\theta; y^*) = J'(F\theta, G\theta; y^*) \quad (6)$$

We will later extend this specification to include also constraints on the equilibrium state and input. The dynamic performance specification is given in terms of an infinite-horizon LQ criterion with $Q \succ 0$, and $R \succ 0$:

$$J_{LQ}(u; x(t)) = \int_t^\infty \big((x(\tau) - \xi)^T Q (x(\tau) - \xi) + (u(\tau) - \mu)^T R (u(\tau) - \mu) \big) \, d\tau \quad (7)$$

3 Optimizing Lyapunov design

3.1 Unconstrained optimizing control

Let Θ^* denote the set of vectors θ that satisfy first order local optimality conditions

$$\Theta^* = \{\theta \in \mathbb{R}^p \, | \nabla_\theta J(\theta; y^*) = 0\} \quad (8)$$

Let $\theta^* \in \Theta^*$ be a global minimizer, with the associated state and input vectors $\xi^* = F\theta^*$ and $\mu^* = G\theta^*$. In general, Θ^* will contain all local and global minima of $J(\cdot; y^*)$ such that θ^* is in general not uniquely defined. The control design is based on the control Lyapunov function

$$V(x, \theta) = (x - F\theta)^T P (x - F\theta) + J(\theta; y^*) - J(\theta^*; y^*) \quad (9)$$

where $P \succ 0$ will be specified shortly. It follows that $V(\xi^*, \theta^*) = 0$ for any global optimizer θ^*. Hence, the objective is to simultaneously optimize the equilibrium point and achieve regulation to the optimal equilibrium. The time-derivative of V along trajectories of the closed loop system is given by

$$\begin{aligned}\dot{V} &= (x - F\theta)^T \left((PA + A^T P)x + 2PBu - 2PF\dot{\theta}\right) \\ &\quad + \nabla_\theta^T J(\theta; y^*)\dot{\theta} \end{aligned} \tag{10}$$

We choose the optimizing dynamic feedback

$$\dot{\theta} = -\Gamma \nabla_\theta J(\theta; y^*) \tag{11}$$

with $\Gamma \succ 0$ and the LQ-like controller

$$u = -R^{-1} B^T P(x - F\theta) + G\theta + \sigma \tag{12}$$

The matrix $P \succ 0$ satisfies the algebraic Riccati equation

$$A^T P + PA - 2PBR^{-1}B^T P = -Q \tag{13}$$

and σ is yet unspecified. This leads to

$$\begin{aligned}\dot{V} &= -(x - F\theta)^T Q(x - F\theta) - \nabla_\theta^T J(\theta; y^*)\Gamma\nabla_\theta J(\theta, y^*) \\ &\quad + 2(x - F\theta)^T P(-F\dot{\theta} + B\sigma) \end{aligned} \tag{14}$$

It is desirable to choose σ such that $B\sigma = F\dot{\theta}$ since this cancels the interaction term such that $\dot{V} \leq 0$. This is certainly possible under some matching condition:

Matching condition: Assume that for all $\vartheta \in \mathbb{R}^p$ there exists a vector $\sigma \in \mathbb{R}^r$ such that $B\sigma = F\vartheta$. $\square$

Proposition 1 *Suppose Q, R, Γ are symmetric and positive definite matrices, the linear system (A, B) is controllable, and the matching condition holds. Then the dynamic controller (11), (12) with σ defined by $B\sigma = F\dot{\theta}$ has the following properties:*

1. *If J is strictly convex, the equilibrium (ξ^*, θ^*) is globally asymptotically stable.*
2. *If J is lower bounded and radially unbounded, then for any $x(0) \in \mathbb{R}^n$ and $\theta(0) \in \mathbb{R}^p$, the states $x(t)$ and $\theta(t)$ are uniformly bounded. Moreover $\theta(t) \to \Theta^*$ and $||x(t) - F\theta(t)||_2 \to 0$ as $t \to \infty$.*

Proof. It is easily shown that (ξ^*, θ^*) is an equilibrium point for the system defined by (1), (11) and (12) for any $\theta^* \in \Theta^*$ with $\xi^* = F\theta^*$. Under the stated matching condition, the last term in (14) vanishes, and gives $\dot{V} \leq 0$.

Part 1 of the result follows directly from LaSalle-Krasovskii's theorem [7] because the strict convexity of J implies uniqueness of θ^* such that

$$\begin{aligned} E &= \{(x, \theta) \in \mathbb{R}^{n+p} \mid \dot{V}(x, \theta) = 0\} && (15)\\ &= \{(x, \theta) \in \mathbb{R}^{n+p} \mid x = F\theta,\ \nabla_\theta J(\theta; y^*) = 0\} && (16)\\ &= \{(\xi^*, \theta^*)\} && (17)\end{aligned}$$

Part 2 can be proven as follows: From (14), $V(x(t), \theta(t))$ is uniformly bounded and $\theta(t)$ and $x(t)$ must be bounded because V is radially unbounded. The conditions of Barbalat's lemma hold because $\nabla_\theta J$ is Lipschitz due to the differentiability of J, such that $\dot{V}(t) \to 0$ as $t \to \infty$, [7]. It follows from (14) that

$$\begin{aligned} ||x(t) - F\theta(t)||_2 &\to 0, \quad \text{as } t \to \infty && (18)\\ \nabla_\theta J(\theta(t); y^*) &\to 0, \quad \text{as } t \to \infty && (19)\end{aligned}$$

Due to the continuity of $\nabla_\theta J$, (19) implies that $\theta(t) \to \Theta^*$ as $t \to \infty$. $\square$

Corollary 1 *If $J_0' = 0$ and the criterion satisfies the conditions of part 1 of Proposition 1, then $g(x(t)) \to y^*$ as $t \to \infty$.* $\square$

Corollary 2 *Proposition 1 still holds when $\sigma = L\dot{\theta}$ and L is arbitrary, instead of the matching condition.* $\square$

Proof. The dynamics of the optimizing dynamic feed-forward (11) does not depend on the state x. Hence,

$$\dot{J} = -\nabla_\theta^T J(\theta; y^*)\Gamma\nabla_\theta J(\theta; y^*) \leq 0 \tag{20}$$

and the convergence/asymptotic stability of $\theta(t)$ follows under the stated assumptions. It also follows that $\dot{\theta}(t) \to 0$ such that $\sigma(t) \to 0$ as $t \to \infty$. The state $x(t)$ is then described by an exponentially stable linear system perturbed by a time-varying vanishing perturbation $\sigma(t)$, and the convergence/asymptotic stability of $x(t)$ follows. $\square$

Remark 1. With constant $\Gamma \succ 0$, (11) corresponds to a gradient descent minimization of J. Notice, however, that Γ in (11) can be replaced by any time- or state-dependent positive definite-matrix, such as a possibly modified inverse Hessian of J, leading to a Newton-like method, in order to improve speed of convergence [9].

Remark 2. In part 2 of Proposition 1, we have shown convergence to the (possibly uncountable) set of local minimizers. Obviously, if J is convex (but not strictly convex) this implies convergence to the set of global minimizers.

Remark 3. The main ideas and resulting controller is similar to what is achieved using Lyapunov design in the context of adaptive and nonlinear control, e.g. [11, 8]. It adaptively optimizes the equilibrium point. The term σ in the control law counteracts the undesired effects of transients due to interactions. Notice that no differentiation with respect to time is required to compute the term $\dot{\theta}$ in the expression for σ, since an explicit expression (11) for $\dot{\theta}$ is known.

Remark 4. The matching condition is restrictive, but not necessary from a stability point of view as shown in Corollary 2. A typical choice is $L = B^+ F$, where B^+ denotes the Moore-Penrose pseudo-inverse of the matrix B. In the latter case, the interaction transients are not exactly cancelled, but minimized in the least-squares sense. Notice that the last term in (14) is not exactly cancelled in this case, only asymptotically since $\dot{\theta}(t) \to 0$ and $\sigma(t) \to 0$ as $t \to 0$.

Example 1. This example is trivial from an application point of view, and its only purpose is to illustrate the benefits of the term σ in the controller. Consider the first order system

$$\begin{aligned} \dot{x} &= x + u && (21)\\ y &= x && (22)\end{aligned}$$

For this problem $\theta = \zeta = -\mu$, or $F = 1$ and $G = -1$. We observe that the matching condition is satisfied, and the control design with $Q = 3$, $R = 1$ gives $K = P = 3$ and the control law

$$
\begin{aligned}
u &= -3(x-\theta) - \theta + \sigma & (23)\\
\sigma &= \gamma(y^* - \theta) & (24)\\
\dot{\theta} &= -\gamma(\theta - y^*) & (25)
\end{aligned}
$$

Select the gain $\gamma = 1$. It is straightforward to show that the transfer function from y^* to y is given by

$$\frac{Y}{Y^*}(s) = \frac{1}{s+1} \quad (26)$$

For comparison, if we choose $\sigma = 0$ in (23) we get the transfer function

$$\frac{Y}{Y^*}(s) = \frac{2}{(s+1)(s+2)} \quad (27)$$

Hence, the beneficial effect of the term σ is that it cancels the stable pole at $s = -2$. We remark that this pole is the closed loop pole due to the feedback gain K and is therefore in general stable and can be cancelled. The cancelling zero is introduced since u depends directly on y^*, and not only indirectly through θ as with $\sigma = 0$. The transient performance is therefore improved due to this zero. □

Example 2. Again, this is a trivial example whose only purpose is to shed some light on the fact that the matching condition is not essential. Consider the 2nd order linear system

$$
\begin{aligned}
\dot{x}_1 &= -x_1 + x_2 & (28)\\
\dot{x}_2 &= x_2 + u & (29)\\
y &= x_1 & (30)
\end{aligned}
$$

In this case, $\theta = \zeta_1 = \zeta_2 = -\mu$, or $F = (1,1)^T$ and $G = -1$. Since $B = (0,1)^T$ the matching condition is *not* satisfied. Choosing $Q = \mathrm{diag}(1,2)$ and $R = 1$ gives the controller

$$u = -0.1716(x_1 - \theta) - 2.8284(x_2 - \theta) - \theta + \sigma \quad (31)$$

Choosing $\gamma = 1$ and $\sigma = B^+ F\dot{\theta}$ (with $B^+F = 1$) gives

$$\frac{Y}{Y^*}(s) = \frac{s+2}{(s+1)(s+\sqrt{2})^2} \quad (32)$$

where the double pole $s = -\sqrt{2}$ is due to the LQ controller, while the pole $s = -1$ is due to the optimizer ($\gamma = 1$). We notice that the zero $s = -2$ (due to the feed-forward from $\dot{\theta}$) counteracts to a large extent the pole $s = -\sqrt{2}$. With $\sigma = 0$ the transfer function becomes

$$\frac{Y}{Y^*}(s) = \frac{2}{(s+1)(s+\sqrt{2})^2} \quad (33)$$

which indicates slower response compared to (32). We conclude that the matching condition is indeed not necessary for reducing interaction transients. □

3.2 Constrained optimizing control

Assume a compact and convex constraint set $\Theta_c \subset \mathbb{R}^p$ is given. Such a convex set can be derived from any convex state and input constraint sets. Let the interior of Θ_c be denoted $\mathrm{int}(\Theta_c)$, its boundary be denoted $\partial\Theta_c$ and assume the set Θ_c is represented as

$$\Theta_c = \{\theta \in \mathbb{R}^p \mid c(\theta) \le 0\} \quad (34)$$

where $c : \mathbb{R}^p \to \mathbb{R}^q$ is a smooth convex vector-valued function. The set of vectors $\theta \in \Theta_c$ satisfying first order optimality conditions is defined in terms of the Karush-Kuhn-Tucker (KKT) conditions:

$$
\begin{aligned}
\Theta_c^* = \{\theta \in \Theta_c \| \nabla_\theta J(\theta; y^*) + \sum_{i=1}^{q} \lambda_i \nabla c_i(\theta) = 0,\\
\lambda_i c_i(\theta) = 0,\ \lambda_i \ge 0\} \quad (35)
\end{aligned}
$$

As before, $\theta^* \in \Theta_c^*$ denotes an arbitrary global minimizer and $\xi^* = F\theta^*$. Next, define the logarithmic barrier function

$$b(\theta) = b_0 - \sum_{i=1}^{q} \log(-c_i(\theta)) \quad (36)$$

where the constant $b_0 \in \mathbb{R}$ is selected such that $b(\theta) > 0$ for all $\theta \in \mathrm{int}(\Theta_c)$. Such a b_0 exists due to the compactness of Θ_c. A fundamental property of this barrier function is that it is well-defined and convex on $\mathrm{int}(\Theta_c)$ since Θ_c is convex [5]. Moreover, its value goes to infinity as $\theta \to \partial\Theta_c$, and it is undefined outside Θ_c. The unconstrained cost function is augmented by a term containing this barrier function when defining a control Lyapunov function

$$
\begin{aligned}
V(x,\theta,\varrho) &= (x - F\theta)^T P (x - F\theta) + J(\theta) - J(\theta^*)\\
&\quad + \varrho b(\theta) + \frac{1}{2}\varrho^2 \quad (37)
\end{aligned}
$$

For all weighting parameters $\varrho > 0$ the barrier function will prevent the solution from escaping the interior of Θ_c. When applying such barrier functions in numerical optimization, convergence toward the optimum is achieved by letting $\varrho \to 0$ as $t \to \infty$, [9, 5], and we take a similar approach here.

Proposition 2 *Consider the optimizing controller*

$$
\begin{aligned}
\dot{\theta} &= -\Gamma q(\theta, \varrho; y^*) & (38)\\
\dot{\varrho} &= -\beta\varrho & (39)\\
u &= -R^{-1}B^T P(x - F\theta) + G\theta + \sigma & (40)
\end{aligned}
$$

where $\Gamma \succ 0$, $\beta > 0$, P are as in Proposition 1, and $q(\theta,\varrho;y^*) = \nabla_\theta J(\theta;y^*) + \varrho\nabla b(\theta)$. Let the matching condition hold and define σ by $B\sigma = F\dot{\theta}$. Assume Θ_c is a convex and compact set, and the function c is differentiable. Then this controller has the following properties for all $x(0) \in \mathbb{R}^n$, $\theta(0) \in int(\Theta_c)$ and $\varrho(0) > 0$:

1. *If J is strictly convex, then the equilibrium point $(x,\theta,\varrho) = (\xi^*, \theta^*, 0)$ is asymptotically stable*
2. *If J is lower bounded and radially unbounded, then all variables are uniformly bounded and $\theta(t) \in int(\Theta_c)$ for all $t \ge 0$. Moreover, $\theta(t) \to \Theta_c^*$, and $||x(t) - F\theta(t)||_2 \to 0$ as $t \to \infty$.*

Proof. The time-derivative of the Lyapunov function candidate (37) along closed loop trajectories is given by

$$
\begin{aligned}
\dot{V} &= -(x - F\theta)^T Q (x - F\theta) + \nabla_\theta^T J(\theta; y^*)\dot{\theta}\\
&\quad + \varrho \nabla^T b(\theta)\dot{\theta} + \dot{\varrho} b(\theta) + \dot{\varrho}\varrho\\
&= -(x - F\theta)^T Q (x - F\theta) - q^T(\theta,\varrho;y^*)\Gamma q(\theta,\varrho;y^*)\\
&\quad - \beta\varrho^2 - \beta\varrho b(\theta) \quad (41)
\end{aligned}
$$

Notice that $\varrho(t) = \varrho(0)\exp(-\beta t)$ such that $\varrho(t) > 0$ for all $t \geq 0$. Since $b(\theta)$ goes unbounded and all other terms remain bounded when θ approaches the boundary of Θ_c, it is clear that $\theta(t) \in \text{int}(\Theta_c)$ for all $t \geq 0$. Hence, the last term in (41) is non-positive because $b(\theta)$ is positive for all $\theta \in \text{int}(\Theta_c)$. It follows immediately that $\dot{V} \leq 0$ and $||\varrho(t)b(\theta(t))||_2$, $||x(t) - F\theta(t)||_2$, $|\varrho(t)|$ and $||q(\theta(t), \varrho(t))||_2$ are uniformly bounded.

In part 1, θ^* is a unique global minimum due to strict convexity, and we have

$$\begin{aligned} E &= \{(x,\theta,\varrho) \in \mathbb{R}^n \times \Theta_c \times [0,\infty) \mid \dot{V}(x,\theta,\varrho) = 0\} \\ &= \{(x,\theta,\varrho) \in \mathbb{R}^n \times \Theta_c \times [0,\infty) \mid \varrho = 0,\ \varrho b(\theta) = 0, \\ &\quad x = F\theta,\ \nabla_\theta J(\theta; y^*) + \varrho \nabla b(\theta) = 0\} \end{aligned} \tag{42}$$

Elementary calculations show

$$\nabla b(\theta) = -\sum_{i=1}^{q} \frac{\nabla c_i(\theta)}{c_i(\theta)} \tag{43}$$

Define the vector $\lambda \in \mathbb{R}^q$ in terms of its components (which can be interpreted as Lagrange multipliers, see also [9]):

$$\lambda_i = -\frac{\varrho}{c_i(\theta)} \tag{44}$$

which is well-defined for $\theta \in \text{int}(\Theta_c)$. Hence, the last condition in (42) can be written

$$\nabla_\theta J(\theta; y^*) + \lambda^T \nabla c(\theta) = 0 \tag{45}$$

Since $\varrho \geq 0$ and $c(\theta) < 0$ for all $\theta \in \text{int}(\Theta_c)$, it follows that $\lambda \geq 0$. Hence, the 1st and 3rd KKT condition in (35) are satisfied. Since $\varrho = 0$ the 2nd KKT condition is also satisfied due to (44), and we conclude that $\theta = \theta^*$ in (42). It follows that $E = \{(\theta^*, \xi^*, 0)\}$ and part 1 of the proposition is proven by Barbashin-LaSalle's theorem [7].

Part 2 can be proven as follows. Since b is locally Lipschitz in int(Θ_c) the conditions of Barbalat's lemma hold [7], and we conclude that $\dot{V} \to 0$ as $t \to \infty$ such that (41) implies

$$\varrho(t) \to 0, \text{ as } t \to \infty \tag{46}$$

$$\varrho(t)b(\theta(t)) \to 0, \text{ as } t \to \infty \tag{47}$$

$$||x(t) - F\theta(t)||_2 \to 0, \text{ as } t \to \infty \tag{48}$$

$$||\nabla_\theta J(\theta(t); y^*) + \varrho(t)\nabla b(\theta(t))||_2 \to 0, \text{ as } t \to \infty \tag{49}$$

As above, with $\lambda_i(t) = -\varrho(t)/c_i(\theta(t))$, it is clear that (49) implies

$$\nabla_\theta J(\theta(t); y^*) + \sum_{i=1}^{q} \lambda_i(t)\nabla c_i(\theta(t)) \to 0, \text{ as } t \to \infty \tag{50}$$

Because all functions are continuous in int(Θ_c) and (46) implies $\lambda_i(t)c_i(\theta(t)) \to 0$ as $t \to \infty$ for all $i \in \{1, 2, ..., q\}$, it is evident that all KKT conditions in (35) hold asymptotically. Since $\lambda_i(t) > 0$ for all $t \geq 0$, we conclude $\theta(t) \to \Theta_c^*$ as $t \to \infty$. □

Remark 5. The autonomous differential equation (39) for ϱ can be replaced by any ODE such that $\varrho(t) \to 0$ as $t \to \infty$, giving flexibility for tuning.

Remark 6. The choice of gain matrix Γ is important to achieve numerical robustness and fast convergence of the optimization. We observe that all the proofs hold if this matrix is time-depending, such that there is considerable flexibility available. A good choice is generally a matrix proportional to some positive definite approximation to the inverse Hessian [9]. In a discrete-time implementation, a line search method is useful to adapt the gain such that descent and convergence are guaranteed. In the simulation examples in this paper we have made a very simple choice, namely $\Gamma(t) = \gamma(t)I$, where $\gamma(t) > 0$ is determined by a line search based on an Armijo condition [9].

4 Simulation example: Colorant mixing process

Assume we have a mixture of one base color and n colorants being feeded continuously at individually controlled rates to a stirred tank. The objective is to control the feeding rates of each colorant to achieve a specified color in the mixture. First order mixing dynamics can be represented as

$$\dot{x} = -\frac{1}{T}x + \frac{1}{T}u, \qquad y = g(x) \tag{51}$$

The composition of colorants $x \in \mathbb{R}^n$ in the mixture determines the color, represented by $y \in \mathbb{R}^3$, through the highly nonlinear mapping $y = g(x)$ which is described in the Appendix. The control variables are the flow-rates of colorants $u \in \mathbb{R}^n$. In general, the number of available colorants n is always at least 3, and may in some applications be very large (tens or hundreds of colorants). In addition to handling the strongly nonlinear mapping g the different colorants have different costs, which suggests to the criterion

$$J_0'(\xi, \mu) = k^T \mu, \quad \text{subject to} \quad \mu \geq 0 \tag{52}$$

Since $\theta = \mu = \xi$, we get

$$\begin{aligned} J(\theta; y^*) &= \frac{1}{2}\sum_{i=1}^{q}(y^* - g_i(\theta))^T W (y^* - g_i(\theta)) + k^T\theta \\ &\quad \text{subject to } \theta \geq 0 \end{aligned} \tag{53}$$

where $y^* \in \mathbb{R}^3$ is the specified color to be produced at minimum cost. The vector k represents cost of colorants relative to the cost of an error $y^* - y$ in the achieved color. Notice that there may in general be different output mappings $g_1, ..., g_q$ that represents different illumination spectra, as described in the Appendix, and it is desirable to select a mixture of colorants such that the appearance of the color is not sensitive to the illumination spectrum.

In the simulation example, the mixing time constant equals $T = 0.5$. Four colorants (yellow, green, blue and red) are being mixed in a light brown base color. In the simulations we have added some noise and disturbances on the feed flow rates. The weighting matrix is $W = I$, $\beta = 2$, $\varrho(0) = 0.1$, $R = I$, and $Q = 4I$. The gain matrix is $\Gamma(t) = \gamma(t)I$, where $0 < \gamma(t) \leq 0.1$ is determined by a line search. The sampling interval is 0.01.

Figures 1 and 2 show some simulation results. The dynamic controller quickly establishes an optimal set-point for the colorants after it is switched on a time 0. At time $t = 10$ the colorant cost vector is changed from $k = (0.25, 0.125, 0.25, 0.25)$ to $k = (0.25, 0.375, 0.25, 0.25)$, and we observe that a new optimal set-point is found after a short transient. At $t = 20$ is reference color is changed, and a new equilibrium is established.

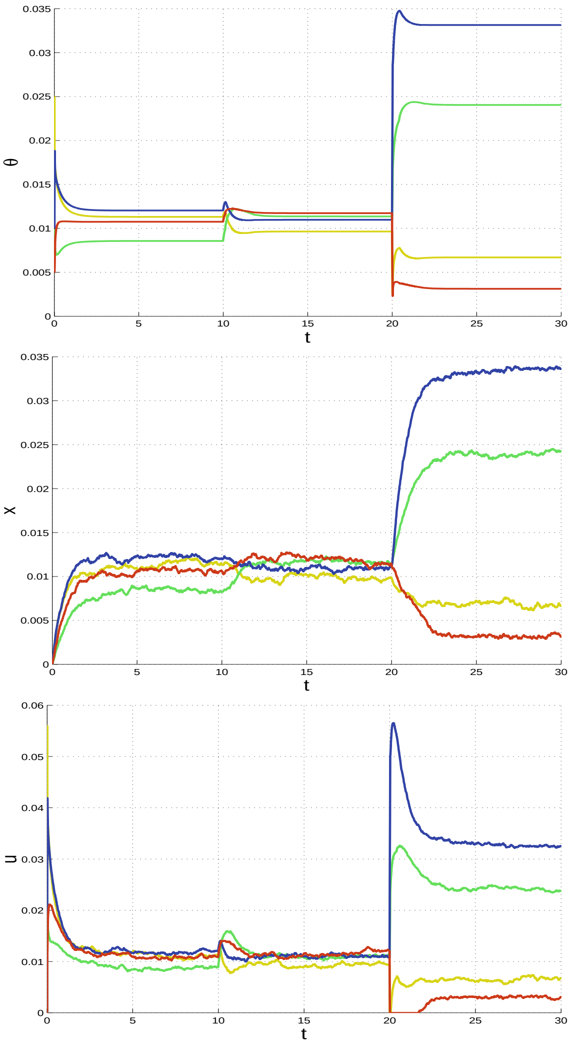

Figure 1: Colorant mixing process simulation. Optimized equilibrium point $\theta = \xi = \mu$ (top), states x (middle) and control u (bottom). The colors on the curves are the colors of the associated colorants.

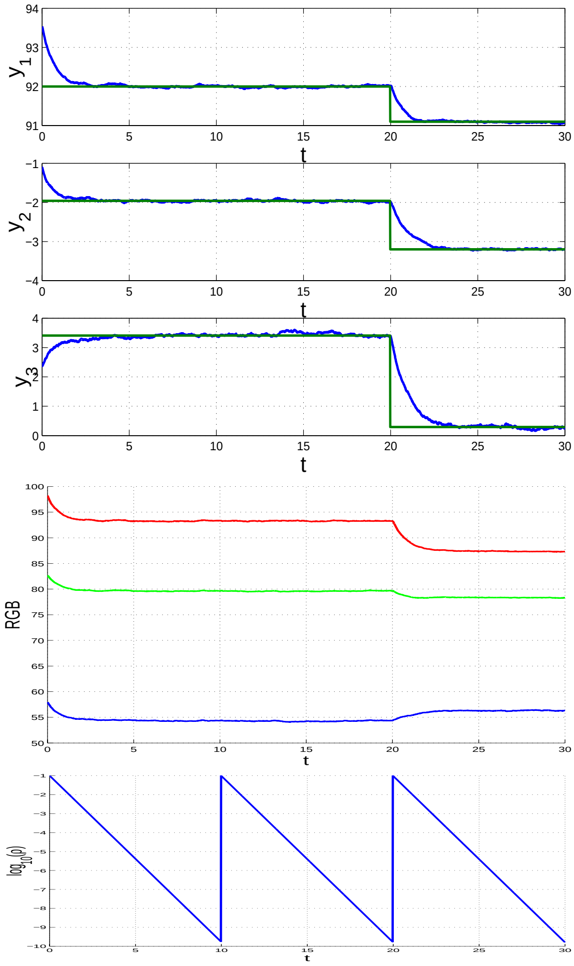

Figure 2: Colorant mixing process simulation. Color and setpoint in (L, a, b) space (top), in RGB (%) space (middle), and barrier function weight ϱ (bottom).

5 Conclusions

The motivation of this work is the need for computationally efficient optimizing control strategies for certain classes of over-actuated systems. It has been shown theoretically and by examples that this can be achieved using a control Lyapunov approach, where an optimizing dynamic feedback law is designed. The design approach is similar to recently developed methods for nonlinear and adaptive control [8], taking explicitly into account the optimizing controllers transient behavior in the control design.

Acknowledgements. This work was in part supported by Fondecyt, Chile, Project 700397.

Appendix A: Kubelka-Munk theory of colorant mixing process

Many computer supported color recipe management systems are based on the Kubelka-Munk theory [1, 2]. This theoretical framework assumes that the colorants affect the reflectance spectrum of the product. The reflectance of the colored product is a nonlinear combination of the reflectance of each individual colorant in it. In addition, the comparison between two colors requires the use of non-linear transformations and weighted integrals of the reflectance curves over a certain region in the visible spectra. At any location in the paint, a certain fraction of light, K, travelling in each direction will be absorbed by the material and another portion, S, will be scattered. The spectrum of the reflected light at every wavelength is the base measurement for color and it is called reflectance R. For complete hiding, the reflectance can be expressed in terms of K and S as:

$$R \quad = \quad 1 + \frac{K}{S} - \sqrt{2\frac{K}{S} + \left(\frac{K}{S}\right)^2}, \tag{54}$$

or $K/S = (1 - R)^2/2R$. The properties of mixtures of pigmented solutions can be found by using the fact that combinations of absorbtion and scattering are linear:

$$K_M = \sum_{i=1}^{n} K_i c_i, \qquad S_M = \sum_{i=1}^{n} S_i c_i, \tag{55}$$

and:

$$\left(\frac{K}{S}\right)_M = \frac{\sum_{i=1}^{n} K_i c_i}{\sum_{i=1}^{n} S_i c_i}, \quad (56)$$

where K_M is the absorption of the mixture, S_M the scattering of the mixture, n number of pigments in the mixture, c_i concentration of the ith pigment in the mixture by weight of dry pigment, K_i absorbtion of the ith pigment, and S_i the scattering of the ith pigment.

Each equation must be calculated for each wavelength. The units of K and S are not important for the purposes of these equations, since in the Kubelka-Munk equations they are always used in conjunction with each other; the important factor is the ratio between them. Taking into account this consideration and given a spectral reflectance curve for R, K can be calculated from equation (54) by setting the values of S equal to one at all wavelengths. Thus, equation (56) can be written as

$$\left(\frac{K}{S}\right)_M = \sum_{i=1}^{n} \frac{K_i}{S_i} \frac{c_i}{\sum_{i=1}^{n} c_i}, \quad (57)$$

this equation means that the concentrations enter the system as fractions. Let $\bar{c}_i = c_i / \sum_{i=1}^{n} c_i$, and the reflectance of the mixture is

$$R_M = 1 + \left(\frac{K}{S}\right)_M - \sqrt{2\left(\frac{K}{S}\right)_M + \left(\frac{K}{S}\right)_M^2}, \quad (58)$$

where $\left(\frac{K}{S}\right)_M$ can be written in terms of the reflectance of each colorant as

$$\left(\frac{K}{S}\right)_M = \sum_{i=1}^{n} \bar{c}_i \frac{(1-R_i)^2}{2R_i}. \quad (59)$$

Several simplifying assumptions were made by Kubelka and Munk. The pigmented solution is considered as a uniform material, assuming complete dispersion of pigments and homogeneous density of pigments particles over a planar surface. There is also no account for surface reflection, since they consider diffuse lighting and viewing conditions.

The color sensor contains a light source and a detector which measures the reflected light from the object to be measured. The spectrum of the reflected light compared to the spectrum of the source light at every wavelength is the reflectance, and it is expressed as a vector of values at each wavelength. The quality of the color adjustment is found by looking at the difference between the spectral distribution of the product and that of the target spectrum. The task of the color management system is to find the colorant concentrations so that, the best approximation of the target spectrum is reached. The human eye is evenly sensitive to each wavelength within the spectrum of visible light; this effect is represented by a set of weighted coefficients $\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$, and $\bar{\mathbf{z}}$, called CIE standard Observer. The tristimulus values are defined by the relations $X = \kappa \int R(\lambda)\bar{\mathbf{x}}S(\lambda)d\lambda$, $Y = \kappa \int R(\lambda)\bar{\mathbf{y}}S(\lambda)d\lambda$, and $Z = \kappa \int R(\lambda)\bar{\mathbf{z}}S(\lambda)d\lambda$, where $S(\lambda)$ is the relative spectral distribution function of the spectral power distribution of a standard illuminant, and $R(\lambda)$ is the spectral reflectance curve. The constant κ is a normalizing factor chosen as: $\kappa = 100/\int \bar{\mathbf{y}}S(\lambda)d\lambda$. The color measurement is specified in terms of the triplet (L, a, b), defined in terms of the tristimulus values as follows:

$$L = 116\left(\frac{Y}{Y_o}\right)^{\frac{1}{3}} - 16, \quad (60)$$

$$a = 500\left((\frac{X}{X_o})^{\frac{1}{3}} - (\frac{Y}{Y_o})^{\frac{1}{3}}\right), \quad (61)$$

$$b = 200\left((\frac{Y}{Y_o})^{\frac{1}{3}} - (\frac{Z}{Z_o})^{\frac{1}{3}}\right), \quad (62)$$

where X_o, Y_o and Z_o are the tristimulus values of the light source. Since the tristimulus considers the illumination, it is possible to have different samples, but with a similar look, and they are called metamers.

References

[1] E. Allen. Basic equations used in computer color matching. *J. Optical Society America*, 56:1256–1259, 1960.

[2] F. W. Billmeyer and M. Saltzman. *Principles of color technology*. John Wiley and Sons, New York, 1981.

[3] H. H. J. Bloemen, C. T. Chou, T. J. J. van den Boom, V. Verdult, M. Verhaegen, and T. C. Backx. Wiener model identification and predictive control for dual composition control of a distillation column. *Journal of Process Control*, 11:601–620, 2001.

[4] M. Cannon and B. Kouvaritakis. Efficient constrained model predictive control with asymptotic optimality. In *Proc. IEEE Conf. Decision and Control, Las Vegas*, pages ThA12–1, 2002.

[5] A. V. Fiacco and G. P. McCormick. *Nonlinear programming: Sequential unconstrained minimization techniques*. J. Wiley and sons, New York, 1968.

[6] B.-G. Jeong, K.-Y. Yoo, and Hyun-Ku Rhee. Nonlinear model predictive control using a Wiener model of a continuous MMA polymerization reactor. *Ind. Eng. Chem. Res.*, 40:5968–5977, 2001.

[7] H. K. Khalil. *Nonlinear Systems*. Macmillan, NY, 1992.

[8] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear Adaptive Control Design*. Wiley and Sons, 1995.

[9] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag, New York, 1999.

[10] S.J. Norquay, A. Palazoglu, and J.A. Romagnoli. Application of Wiener model predictive control (WMPC) to pH neutralization experiment. *IEEE Trans. Control Systems Technology*, 7:437–445, 1999.

[11] P. C. Parks. Liapunov redesign of model reference adaptive control systems. *IEEE Trans. Automatic Control*, 11:362–367, 1966.

[12] A. R. Teel. Lyapunov methods in nonsmooth optimization, Part I: Quasi-Newton algorithms for Lipschitz, regular functions. In *Proc. IEEE Conf. Decision and Control, Sydney*, pages 112–117, 2000.

[13] A. R. Teel. Lyapunov methods in nonsmooth optimization, Part II: Persistently exciting finite differences. In *Proc. IEEE Conf. Decision and Control, Sydney*, pages 118–123, 2000.