Acta Scientific COMPUTER SCIENCES

Volume 5 Issue 4 April 2023

Research Article

Robust Optimal Control of Safety-Critical Systems

Navid Moshtaghi Yazdani*

Ph.d in Electrical Engineering, Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

***Corresponding Author:** Navid Moshtaghi Yazdani, Ph.d in Electrical Engineering, Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran.

Received: January 22, 2023

Published: March 02, 2023

Abstract

Safety, stability and robustness are some fundamental issues in design and analysis of safety-critical systems. As a result, designing an efficient controller that simultaneously considers design goals such as safe performance in the presence of disturbances and considering system's limitations, is an interesting field of research. While safe control design has been investigated in the literature, design of safe robust controllers with ability of disturbance reduction is newly proposed in this article. The mentioned safe controller is defined using a two stage policy iteration algorithm which is applied to solve the Hamilton-Jacobi-Isaacs equation. During the proposed algorithm possible conflicts between safety and robustness are minimized during each iteration. The simulation results indicates the effectiveness of proposed method.

Keywords: Optimal Control; Safety System; Critical Systems; Simulation; Efficient Controller

Introduction

Satisfaction of safety requirements of dynamical systems is of vital importance for any control system that must be designed for safety-critical systems such as power systems, industrial and mobile robots and chemical reactors. While safe control design has been widely investigated in the literature [add references here], design of safe controllers with performance and disturbance attenuation guarantees is not considered. Disturbances, however, are commonplace in most control systems and their ignorance can significantly deteriorate system's performance and even its stability and safety. This gap has motivated us to present robust optimal safe controllers for nonlinear continuous-time systems under safety constraints and disturbances.

Related work

H_∞ control has been extensively used to attenuate the effects of disturbances on the system performance [1-3]. Designing an H_∞ controllers amounts to solving a two-player zero-sum game, which in turn boils down to solving the so-called Hamilton-Jacobi-Isaacs (HJI) equation [4,5]. HJI equation is a nonlinear partial differential equation which is hard to solve directly. Therefore, policy iteration (PI) algorithms have been widely employed to solve the HJI equation in an iterative manner [6,7]. However, standard PI algorithms for solving HJI equations ignore the safety constraints and limitations of safety-critical systems, and thus cannot be implemented on these systems.

On the other hand, model predictive control (MPC) [8] takes into account system operating constraints while optimizing an operational function. However, since the MPC employs short-horizon operational functions, it results in myopic control approaches, which make achieving stability and guaranteeing feasibility rather troublesome. Moreover, MPC solves an optimization problem at every instance of time, which can be computationally complex for nonlinear systems and/or nonquadratic operational functions. Optimal control of constrained systems using penalty functions in operational functions is considered in [9]. However, these approaches only prove useful for linear constraints. Control barrier functions (CBFs) have also been widely used to design safe controllers [10]. To account for disturbances, input-to-state safety is considered in [11]. CBFs are integrated with control Lyapunov functions (CLF) to integrate safety and stability control design [12]. However, these methods mostly do not consider the performance and optimality of the control solution. In our previous work [13], a safety-certified PI algorithm is presented to find a safe optimal control policy. However, since most practical systems are under disturbances, it is of vital importance to take into account the effect of disturbances on the performance. This work extends the results of [13] to systems under disturbances.

Participation and paper outline

This work presents a safe H_∞ control framework and safe PI-based solutions for solving constrained HJI equations arising from solving it. The proposed PI algorithm consists of two parts: 1- Policy assessment stage where the value function is found in correlation with the safety policy in the presence of interrupts, and 2- Policy improvement stage where a policy is found with improved performance for that specific safety using a compound with specific control barrier functions. Possible conflicts between safety and robustness are minimized during each iteration. To insure performance, an HJI inequality is solved, and compounded with the barrier certificate. A sum-of-squares program is used to find an optimal safe control solution in the presence of interrupts. Finally, one simulation example are presented in order to demonstrate the effectiveness of the proposed approach.

The remainder of the paper is summarized as follows: in section 2, the problem is formulated and some basic results pertaining to optimal control of H_∞ are presented. In section 3, a new optimal safe and robust format is introduced and in section 4, numerical examples are presented to validate the effectiveness and efficiency of the proposed approach.

Notations

Throughout the paper, $\mathbb{C}$ is the set of all continuous functions, and $\mathbb{Q}$ indicates the set of all functions in $\mathbb{C}$ which are positive definite and proper. The polynomial $q(x)$ is a sum of squares polynomial, i.e. $q(x) \in \mathbb{Q}^{SOS}$ where $\mathbb{Q}^{SOS}$ is a set of SOS polynomials, if $q(x) = \sum_1^n q_i^2(x)$, where $q_i(x) \in \mathbb{Q}$ $i = 1,\dots,n$. $K_b : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function where $K_b : [0,m) \to \mathbb{R}_+$ and $m > 0$ belongs to the $\mathcal{K}(K_b \in \mathcal{K})$ set, given that $K_b(0) = 0$ and K_b is strictly increasing. When $\lim_{a\to\infty} K_b(a) = \infty$ and $K_b = \infty$, then K_b belongs to the $\mathcal{K}_\infty(K_b \in \mathcal{K}_\infty)$ set. ∇V refers to the gradient of the V function: $\mathbb{R}^n \to \mathbb{R}^n$. The Li-derivative of h as a function of f is defined as $L_f V(x) = \frac{\partial h}{\partial x} f(x)$.

Problem formulation and preliminary results

This section presents the preliminary results on the robustness, safety and optimization problem of systems control in the presence of disturbances. Later, the problem of optimal safe control design in the presence of disturbances is formulated.

Dynamic systems optimal control

Consider the following nonlinear system in the presence of an exogenous interrupt:

$$\dot{x} = f(x) + g(x)u + n(x)d \qquad \text{-------------(1)}$$

$$y = z(x)$$

Where $x \in \mathbb{R}^n$ is the system states vector, $u \in \mathbb{R}^m$ is the control input vectors, $d(t) \in D$ is an external disturbance and $y(t) \in R^p$ is the output. Both $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, $n : \mathbb{R}^n \to \mathbb{R}^n$ are locally Lipschitz with $f(0) = 0$. The disturbance $d(t)$ is essentially bounded in time and $\|d_\infty\| \triangleq ess \sup_{t\geq 0} d(t)$ is defined. The system is assumed to be stabilizing. The goal of H_∞ control design is to find a control policy that stabilizes the system when $d(t)$ §= and minimizes the predefined disturbance-attenuation related performance function defined as

$$J(x_0,u,\lambda) = \int_0^\infty r(x(t),u(t),\lambda)dt \qquad \text{-------------(2)}$$

Where $r(x ü \lambda) = x^T Q x + u^T R u \quad \lambda^2 \|d\|^2$ is the reward function, R is a determined positive symmetrical matrix and λ is the positive constant for the reward function $r(x, u, \lambda)$ defined such that the optimization (2) guarantees minimization of control efforts in order to achieve the proper transient response as well as system stability. Optimal control is in pursuit of the least possible value for λ^* and its respective controller, such that the above problem remains solvable.

Where $\alpha > 0$ is the discount factor, and $r(t)$ is the bounded reference trajectory.

Definition 1

Consider system (1),For the Lyapunov function $V \in \mathbb{Q}$ and a feedback control policy u , define the Bellman equation as

$$\mathcal{L}(V, u, \lambda) = -(L_f V(x) + L_g V(x)u + L_n V(x)d) - x^{\ddot{u}} Q x - u \quad Ru$$
$$+ \lambda^2 \|d\|^2 \quad \forall x \in \mathbb{R}^n$$

------ (3)

Assumption 1

Consider system (1). Then, the exists a V and a control policy such that for the case where there is no disturbance, i.e., $d = 0$, one has $\mathcal{L}(V, u, \lambda) \geq 0$.

This assumption is satisfied for stabilizing systems and guarantees system's stability.

Theorem 1

Consider the system (1) with the operational function (2). With λ constant, there exists a positive semi-definite function $V^*(x) \in C^1$ that satisfies the following HJI equation.

$$\mathrm{H}_\infty(V^*) = 0$$

Where

$$\mathrm{H}_\infty(V^*) = x^T Q x + L_f V^*(x) - \frac{1}{4} L_g V^*(x) R^{-1}(x)(L_g V^*(x))^T + L_n V^*(x) d - \lambda^2 \|d\|^2,$$
$$V^*(0) = 0$$

-------(4)

Moreover, the feedback control

$$u^*(x) = \frac{1}{2} R^{-1}(x)(L_g V^*)^T(\mathrm{x})$$

--------(5)

Satisfies $L_{2-gain} \leq \lambda$.

Proof: see [5]

Note that equation (4) is a nonlinear partial differential equation which optimizes the performance index (2) and assures that the equilibrium $x = 0$ is asymptotically stable in the absence of disturbance. Additionally, the optimum value function is given as

$$V^*(x_0) = \min_u J(x_0, u, \lambda) = J(x_0, u^*, \lambda), \qquad \forall x_0 \in \mathbb{R}^n$$

---------(6)

[14] introduces the worst-cased disturbance policy as

$$d(x) = \frac{1}{2\lambda^2}(L_n V)^T$$

---------(7)

Since directly solving the HJI equations is hard or even impossible, repetitive policy algorithms are presented in [5,15]. However, no safety guarantees are given for existing repetitive policy algorithms in the presence of disturbances.

Safe control of dynamic systems

Control barrier functions are increasingly employed as a tool for combining controllers which provide safety through the invariance of sets [16-18]. Safety certificates based on a controller combined with CBFs rely on a precise model of system dynamics and may be compromised in the presence of uncertainties. A recently definition of input-to-state safety (ISSf) provides a tool for quantifying the impact on safety guarantees of such uncertainty or disturbances in the dynamics [19] by describing changes in the set kept invariant [20].

We perform an overview of the control barrier functions and input-to-state safety control barrier functions (ISSf-CBFs) in this section. In a safety-critical system, it is vital to prevent entry of the system's states into certain non-safe areas $\mathcal{X}_u \in \mathcal{X}$, starting from any the initial state inside the set $\mathcal{X}_0$. In order to design a safe controller, control barrier functions may be used.

Consider the system (1) with $d = 0$. Let $b : \mathbb{R}^n \to \mathbb{R}$ exist such that:

$$b(\mathrm{x}) \geq 0, \ \forall x \in \mathcal{X}_0,$$
$$b(\mathrm{x}) < 0, \ \forall x \in \mathcal{X}_u$$

-----------(9)

Also, let us define the following function

$$\mathcal{L} = \{x \in \mathcal{X} \mid b(x) \geq 0\}$$

----------(10)

Then, the zero CBF (ZCBF) space for acceptable state $S(x)$ can be defined as

$$S(x) = \{u \in \mathbb{R}^m \mid A + K_b(b(x)) \geq 0\}, x \in \mathcal{X}$$

------------(11)

Where $A = L_f b(x) + L_g b(x)u$

System (1) is input-to-state safe on a set $\mathcal{L} \subset R^n$ taking into account the disturbance d, on condition that $\eta \in \mathcal{K}_\infty$ and $\bar{d}$ 0 exist such that the set $\mathcal{L} \subset \mathcal{L}_d$

$$\mathcal{L}_d = \{x \in \mathbb{R}^n : b(x) + \eta(\|d_\infty\|) \geq 0\}$$

----------(12)

Is forward invariant for all d satisfying $\|d_\infty\| \leq \bar{d}$ $\mathcal{L}$ is defined as a input-to-state safe (ISSf) set on condition that the $\mathcal{L}_d$ set exists.

As can be seen from [38], the existence of an ISSf-BF for (1) indicates that $\mathcal{L}$ is a set of ISSf. Similarly to the non-undisturbed case, the control barrier function concept is defined for a compound of controllers which guarantee input-to-state safety: The function b is an ISSf-CBF for system (1) on $\mathcal{L}$ if there exist $\eta \in \mathcal{K}_\infty$ and $\bar{d} > 0$, With $b(x)$ the acceptable state space $SI(x)$ becomes:

$$SI(\mathrm{x}) = \{\mathrm{u} \in \mathbb{R}^m \mid AI + K_b(b(\mathrm{x})) + i\|d\| \geq 0\},\ \mathrm{x} \in \mathcal{X} \quad \text{--------(13)}$$

Where $AI = L_f b(x) + L_g b(x)u + L_n b(x)d$

The following theorem shows how a controller is designed using the concept of ISSf-CBF in order to guarantee that the safe set of $\mathcal{L}$ is forward invariant and as such the system is stable.

Theorem 2

In [20] a set $\mathcal{L}_d \subset R^n$ as defined in (12) is introduced and ISSf-CBF b as defined in (13) is given, for each controller $u \in SI(x)$ in system (1) the safe set $\mathcal{L}$ is given as forward invariant.

Assumption 3

The admissible control space $SI(x)$ is non empty.

Remark 1

Assumption 1 can be satisfied if a stabilizing controller exists and is standard for all control systems, Assumption 3 can be satisfied if a safe controller exists and is standard for safe control design. These two theorems are not necessarily contradictory. If safety and robustness could not be established simultaneously over some regions of the space, it has been already demonstrated (Shown in [33]) that the standard state is to momentarily sacrifice robustness for the guarantee of safety.

Presented Safe robust optimal controller

In the field of robust control, H_∞ control is a powerful tool in solving the disturbance attenuation problem occurred in many practical systems.

While an optimal H_∞ controller based on the HJI solution of (4) guarantees performance, it cannot provide safety in a reliable manner. On the other hand, a controller based on satisfying CBF (13) guarantees safety but might compromise performance. In order to maintain performance and safety simultaneously and in the presence of disturbances, in this section we attempt to design stabilizing robust safe controllers which can guarantee performance in the valid safe region. Repetitive algorithms with an iterative policy to solve HJI inequalities can be employed without constraints. However, existing repetitive algorithms for the presented policies cannot guarantee safety. In this paper, in order to find a safe control policy with guaranteed disturbance attenuation performance, the following optimization problem over HJI inequality is first performed to replace solving directly the HJI equation.

$$\begin{aligned} &\min_V \int_\Omega V(x)dx \\ \text{st.}\quad &-H_\infty(V) \geq 0 \\ &V \in \mathbb{Q} \end{aligned} \quad \text{----------------(14)}$$

Where $H_\infty(V)$ is defined by (4) and $\Omega \subset \mathbb{R}^n$ is an arbitrary compressed set which includes the origin. This set represents a region for which more disturbance attenuation is expected. The improved system after the safe policy assessment stage in (14) allows for smaller damping coefficients in order to achieve L_2-gain performance.

It is shown that [21] the solution to problem 1 is singular and if $V*$ is a solution for (14), then

$$u^\infty(x) = \frac{1}{2}R^{-1}(x)(L_g V^*)^T(\mathrm{x}) \quad \text{---------------(15)}$$

Guarantees stability and $V*$ can be considered as a top limit or a top approximate for actual cost. The ∞ superscript used here signifies that ∞ is a performance-oriented controller. However, this control policy does not reassure system safety. Using this simplified optimal control formulation, we will now present the following optimization framework wherein performance and safety have been considered. Safety is guaranteed by adding a CBF inequality to the optimal control problem formulation. The proposed safe optimization in the presence of interrupts is presented as follows.

Problem 1. (Optimal robust safe control): Find a controller that solves the following equation:

$$\begin{aligned} &\int_\Omega V\,dx + k_\xi \xi^2 \\ st.\quad &-H_\infty(V) \geq \xi \\ &AI + K_b(b(x)) + i\|d\| \geq 0 \end{aligned} \quad \text{-------------------(16)}$$

Where Ω is the level where system improvement is expected, $k_\xi > 0$ is the design parameter which established the trade-off between aggressiveness and safety, and ξ is the stability relaxation factor. Note that ξ can be the taken as the aspiration level for performance to show how much the performance has been compromised in a state where both safety and performance cannot be simultaneously achieved. However, this parameter is minimized in order to achieve the best possible performance.

Theorem 3

Under assumption 1, the robust safe control problem 1 has one possible solution.

Proof: Based on Theorem 2, there exists a safe control approach, namely u. Let us define this control approach as $u = u^{safe} + u^{\infty}$ where $u^{\infty}(x) = \frac{1}{2} R^{-1}(x)(L_g V^*)^T(x)$ is a part of the controller which is used to optimize the performance in the presence of disturbances while ignoring safety, as given in equation (15). Moreover, u^{safe} is integrated with u^{∞} to certify safety. Thus,

$$\begin{aligned}H_\infty(V^*) &= \|z\|^2 + L_f V^*(x) - \frac{1}{4} L_g V^*(x) R^{-1}(x)(L_g V^*(x))^T + L_n V^*(x) d - \lambda^2 \|d\|^2 \\ &= L_f V^*(x) + L_g V^*(x) u^{\infty} + L_n V^*(x) d + r(x, u^{\infty}) \\ &= L_f V^*(x) + L_g V^*(x) u - L_g V^*(x) u^{safe} + L_n V^*(x) d + r(x,u) - u^T R u + u^{\infty T} R u^{\infty} \\ &= L_f V^*(x) + L_g V^*(x) u + L_n V^*(x) d + r(x,u) - \|u^{safe}\|_R^2\end{aligned}$$

----------(17)

If adding u^{safe} contradicts with stability in some regions, the condition $L_f V^*(x) + L_g V^*(x) u^{\infty} + L_n V^*(x) d + r(x, u^{\infty}) < 0$ may not be satisfied for some states. By choosing a proper slack variable $\xi(x)$ in order to compensate for the contradiction between safety and stability, one has

$$H_\infty(V^*) - \xi = L_f V^*(x) + L_g V^*(x) u + L_n V^*(x) d + r(x,u) - \|u^{safe}\|_R^2 - \xi \le 0$$

-------------(18)

On the other hand, since u is safe, a CBF $b(x)$ exists based on the inverse Control Barrier Lyapunov, which satisfies

$$AI + K_b(b(x)) + i\|d\| \ge 0$$

This completes the proof.

The solution to this optimization problem is non-trivial. If the inequalities of the constraints HJI and CBL are limited to the constraints of SOS, the SOS program can be used to reduce significantly the computational efforts needed for the solution of this optimization problem. However, since $H_\infty(V)$ is bilinear in V, the solution to the optimization problem becomes difficult using SOS. Therefore, we propose a safe repetitive policy algorithm instead of the direct solutions to $H_\infty(V) \le \xi$ by iteratively solving a Bellman inequality, which is linear in V. Using the Bellman inequality, a stage of the policy assessment finds V^i value corresponding to that of the safe control policy u^i. The improved policy after the repetition of the policy allows for smaller damping coefficients in order to achieve performance. Based on this, an SOS program can be constructed in order to find a smaller damping coefficient to control H_∞.

Policy improvement stage finds the improved policy u_j^{i+1} for which the safety is approved by adding the CBF inequality. We assume that the initial safe control policy u^0 is given, which can satisfy stability using a control policy without worrying over efficiency. In order to assess a specific policy u^i, and in order to find the corresponding value function V^i, the following policy assessment stage is proposed.

Safe policy assessment stage: with safe control policy given, V^i λ^i and ξ^i can be found which helps to solve the following optimization problem.

$$\begin{aligned}&\min\ \lambda \\ \text{st.}\quad &\mathcal{L}(V_j, u_j, \lambda_j) = -(L_f V(x) + L_g V(x) u + L_n V(x) d) - r(x,u) \ge -\xi_i \\ &V^{i-1} - V^i \ge 0\end{aligned}$$

$$\begin{aligned}&\min_V \int_\Omega V(x)dx \\ \text{st.}\quad &\mathcal{L}(V^i, u^i, \lambda) = -(L_f V(x) + L_g V(x) u + L_n V(x) d) - r(x,u) \ge -\xi_i \\ &V^{i-1} - V^i \ge 0\end{aligned}$$

In SOS format, this optimization problem becomes:

$$\begin{aligned}&\min\ \lambda \\ \text{st.}\quad &\mathcal{L}(V_j, u_j, \lambda) + \xi_i \ \ is\ SOS \\ &V^{i-1} - V^i \ \ is\ SOS\end{aligned}$$

$$\begin{aligned}&\min_V \int_\Omega V(x)dx \\ \text{st.}\quad &\mathcal{L}(V^i, u^i, \lambda) + \xi_i \ \ is\ SOS \\ &V^{i-1} - V^i \ \ is\ SOS\end{aligned}$$

---------------(19)

Thus, a two-loop repeated algorithm is considered where the inner loop searches the HJI equation with the damping coefficient and the outer loop minimizes the damping coefficient. In the policy assessment stage (19), the value function corresponding to the given policy is found while the relaxation factor ξ^{i} is minimized. Note that since it is possible that a safe control approach u^{i} may not require stabilization, it can be possible that $\mathcal{L}(V^{i}, u^{i}, \lambda)$ may not be positive semi-definite. When the value function V^{i} is found, the policy improvement stage calculates an improved valid control policy as below.

In order to find the improved policy, the following stationary condition is invoked:

$$u^{i+1} = \arg\min_{u} \mathcal{L}(V^{i}, u, \lambda) \qquad \text{----------- (20)}$$

Which hints to the sufficient condition for global minimality of the Bellman equation while considering the satisfaction of CBF during policy improvement. Note that the Bellman equation can be rewritten as

$$\begin{aligned}\mathcal{L}(V^{i}, u^{i+1}, \lambda^{i}) &= -L_f V^{i}(x) - L_g V^{i}(x) u^{i+1} - L_n V^{*}(x) d^{i} - r(x, u^{i+1}) \\ &= -(L_f V^{i}(x) + L_g V^{i}(x)(u^{\infty})^{i+1}) + L_n V^{*}(x) d^{i}) \\ &\quad - r(x, (u^{\infty})^{i+1}) + \left\|u^{safe}\right\|_R^2\end{aligned}$$

----------------(21)

Where $u = u^{\infty} + u^{safe}$.

Minimizing

$$E = -(L_f V^{i}(x) + L_g V^{i}(x)(u^{\infty})^{i+1}) + L_n V^{*}(x) d^{i}) - r(x, (u^{\infty})^{i+1})$$

---------(22)

Using stationarity conditions, results in $(u^{\infty})^{i+1} = -\frac{1}{2} R^{-1}(L_g V^{i}(x))^{T}$. Therefore, minimizing $\left\|u^{safe}\right\|_R^2$ as the second term by replacing $u^{i+1} = (u^{\infty})^{i+1} + u^{safe}$ optimizes performance. Since controllers must approve safety constraints, the CBL inequality must also be taken into account. This leads to a policy improvement as below.

Policy improvement stage: with the value function V^{i} given, the valid improved control policy u^{i+1} is found by solving the below optimization problem:

$$\begin{aligned}&\min_{b, u^{safe}, K_{\xi}} \left\|u^{safe}\right\|^2 \\ &u^{i+1} = (u^{\infty})^{i+1} + u^{safe} \\ &L_f b(x) + L_g b(x) u^{i+1} + L_n b(x) d + \Lambda(b(\mathrm{x})) + i\left\|d\right\| \geq 0 \\ &\Lambda is SOS\end{aligned}$$

------------(23)

Note that the safe policy improvement stage (23), finds a safe controller which has the least interference with the performance-driven controller. The SOS program (23) includes bilinear decision making variables. Therefore, in order to perform the policy improvement stage, it must first be broken into several smaller SOS programs, which result in a repetitive search algorithm which first fixes b and then finds u^{safe} and Λ.

Algorithm 1 presents the proposed repetitive policy algorithm. Theorem 4 demonstrates that Algorithm 1 finds a robust safe improved policy at each stop, and terminates if further improvements cannot be found. An initial general fixer u^{0} is required in the above algorithm, assuming limited L_2-gain. The precise value for L_2-gain is not required.

Algorithm 1: Optimal robust safe policy repetition

1. Process
2. Input: Start with a safe and probable conservative control policy
3. First stage (safe policy assessment): follow the steps below
4. Sub-stage 1: Fix the control policy u^{i} and solve equation (19) for V^{i}
5. Sub-stage 2: Minimize the damping coefficient considering the results of (19)
6. Second stage (safe policy improvement): Fix V^{i} and solve (23) for u^{safe} and Λ
7. Continue stages 1 and 2 until converged
8. end process

Algorithm 1

Theorem 4

Consider the dynamic control system (1) and allow hypotheses 1 and 2 to be satisfied. Also allow for algorithm 1 to start from a safe control policy u^{0} with the value function V^{0}. Then, performance will be improved in each repetition and u^{i} would be safe for all i values.

Proof: The safe policy assessment stage in algorithm 1 has the condition $V^{i-1} - V^{i} \geq 0$ and searches among the policies which stabilize the system to the possible extents, maximizing performance. Furthermore, according to (21), the policy improvement stage in algorithm 1 minimizes the term E using stationary conditions which results in $u^{\infty} = -\frac{1}{2} R^{-1}(L_g V^{i}(x))^{T}$. Thus, minimizing $\|u^{safe}\|$ as the second term while the configuration $u^{i+1} = (u^{\infty})^{i+1} + u^{safe}$, as has been performed with the policy improvement stage in Algorithm 1, optimizes performance in each iteration. Since the CBL inequality is also considered while finding an improved policy, the safety of the improved policy is guaranteed, while interferences from the safe control section through minimizing u^{safe} has been minimized.

Simulation results

In order to verify the effectiveness of the method, a simulation example will be presented in this section.

Example 1. Consider the following non-linear system dynamic:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 - x_1^2 x_2 + 0.5 x_2^2 \\ x_2 + 1.5 x_1 x_2^2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0.4 & 0.6 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d$$

$$y = [0 \quad 1] x$$

-----------(24)

Where, $x = [x_1, x_2]$ and u represent system state and system control, respectively. An unsafe space using terms from the polynomial inequality $X_u = \{x \in R^2 \mid b_i(x) < 0, i = 1,2,3)$ has been coded with the following details:

$$b_1 = -0.5 + (x_1 + 2)^2 + (x_2 + 2)^2 < 0$$
$$b_2 = -0.5 + (x_1 + 3)^2 + (x_2 - 1.5)^2 < 0$$
$$b_3 = -0.5 + (x_1 - 3)^2 + (x_2 - 1)^2 < 0$$

---------(25)

In order to find an initial policy to begin with, and while considering the SOS methods presented in [30], the following stabilizing robust control policy is used:

$$\begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} = \begin{bmatrix} 3.11 x_1 - 2 x_2 \\ 1.97 x_1 + 4.61 x_2 \end{bmatrix}$$

-----------(26)

Our control objective is to find optimized safe robust control policies using iterative robust safe policy algorithms. Simultaneously, we want to guarantee global asymptotic stability. In particular, the aim is to optimize closed-loop system performance. By solving the feasibility problem using SOS-TOOLS which along with (26), satisfies Assumption 2.

We find the corresponding value function for Assumption 2 to be:

$$V^{1} = 1.114 x_1^2 - 0.68 x_1 x_2 + 0.5896 x_2^2 - 0.037 x_1 x_2^2$$

λ converges to 1.16418 after 8 iterations. The λ^{i} curve over the repeat loop has been demonstrated in Figure 1.

The final semi-optimal safe controller H_{∞} found using the proposed algorithm is as follows:

$$u_{18}^{\infty} = -0.02587 x_1^3 - 0.06402 x_1^2 x_2 - 0.287 x_1^2 - 0.0332 x_1 x_2^2 - 0.04395 x_1 x_2 - 0.0255 x_1 - 0.00452 x_2^3 - 0.3172 x_2^2 - 0.775 x_2$$

$$u_{28}^{\infty} = -0.0194 x_1^3 - 0.00366 x_1^2 x_2 - 0.395 x_1^2 + 0.0098 x_1 x_2^2 - 0.0982 x_1 x_2 - 1.005 x_1 + 0.00382 x_2^3 - 0.6169 x_2^2 - 1.391 x_2$$

The value function in the proposed SOS-based algorithm overall, is positive based on SOS theory, which renders the H_{∞} controller valid throughout the entire system state space.

To test the attenuation effect of u_8^{∞} we inject an impulse disturbance at t = 30 deviate the state from the origin. The system response under the proposed control policy and the initial control policy presented are compared in figure 2. Figure 2.a is X_1 and Figure 2.b is X_2 The indefinite cost function and the initial cost function are also compared in figure 3.

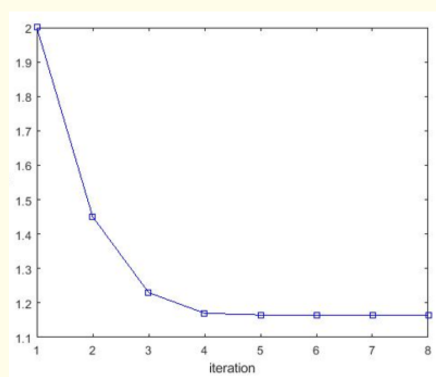

Figure 1: Values of during the learning process in Example.

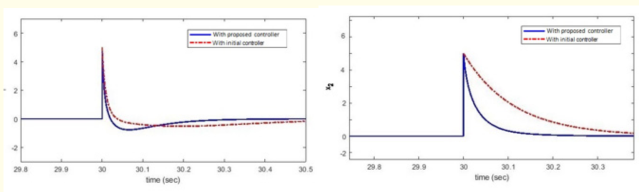

Figure 2: a System trajectorie X_1. b System trajectorie X_2.

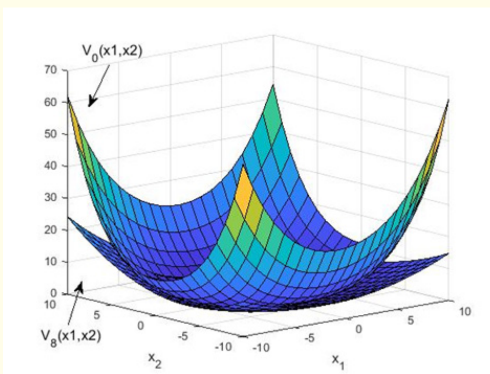

Figure 3: Comparison of learned value functions.

Conclusions

A robust safe optimization for nonlinear dynamic systems control under constrained states has been performed. In order to guarantee performance and safety, the Hamilton–Jacobi–isaacs equation is replaced by a (HJI) inequality which approximates the solution to the HJI equation for specific damping coefficients, while minimizing the coefficient. Also, an iterative optimum policy algorithm is introduced which confirms safety of the optimized robust policy and finds the corresponding value function, and the proposed iterative robust safe SOS-based algorithm is then presented. The simulated example confirms the effectiveness of the proposed safe algorithm.

Bibliography

1. A Isidori and W Lin. "Global L2-gain design for a class of nonlinear systems". *Systems and Control Letters* 34.5 (1998): 245-252.
2. A Isidori. "H∞ control via measurement feedback for affine nonlinear systems". *International Journal of Robust and Nonlinear Control* 4 (1994): 553-574.
3. G Zames. "Feedback and optimal sensitivity: model reference transformations, multiplicative semi norms, and approximate inverses". *IEEE Transactions on Automatic Control* 26.2 (1981): 301-320.
4. A Isidori and A Astolfi. "Disturbance attenuation and H∞-control via measurement feedback in nonlinear systems". *IEEE Transactions on Automatic Control* 37.9 (1992): 1283-1293.
5. A J van der Schaft. "L2-gain analysis of nonlinear systems and nonlinear state-feedback H∞ control". *IEEE Transactions on Automatic Control* 37.6 (1992): 770-784.
6. Q Zhang., *et al*. "Event-triggered H∞ control for continuous-time nonlinear system via concurrent learning". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47.7 (202).
7. H Zhang., *et al*. "Online adaptive policy learning algorithm for H∞ state feedback control of unknown affine nonlinear discrete-time systems". *IEEE Transactions on Cybernetics* 44.12 (2014): 2706-2718.
8. A Bemporad., *et al*. "The explicit solution of model predictive control via multiparametric quadratic programming". in Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No.00CH36334) 2 (2000): 872-876.
9. M Abu-Khalaf and F L Lewis. "Nearly optimal state feedback control of constrained nonlinear systems using a neural networks hjb approach". *Annual Reviews in Control* 28.2 (2004): 239-251.
10. Aaron D Ames., *et al*. "Control Barrier Functions: Theory and Applications". in Proceedings of the 2019 18th European Control Conference. 25-28 June (2019).
11. Shishir Kolathaya., *et al*. "Input-to-State Safety With Control Barrier Functions". *IEEE Control Systems Letters* 3.1 (2019).
12. L Wang., *et al*. "Permissive barrier certificates for safe stabilization using sum-of-squares". in 2018 Annual American Control Conference (ACC) (2018): 585-590.

13. Navid Moshtaghi Yazdani., *et al*. "A Safety-Certified Policy Iteration Algorithm for Control of Constrained Nonlinear Systems". *IEEE Control Systems Letters* 4.3 (2020).
14. R Beard and T McLain. "Successive Galerkin approximation algorithms for nonlinear optimal and robust control". *International Journal of Control* 71.5 (1998): 717-743.
15. M Abu-Khalaf., *et al*. "Policy iterations on the Hamilton-Jacobi-Isaacs equation for state feedback control with input saturation". *IEEE Transactions on Automatic Control* 51.12 (2006): 1989-1995.
16. A D Ames., *et al*. "Control barrier functions: Theory and applications". In European Control Conference (ECC) (2019): 3420-3431.
17. Q Nguyen and K Sreenath. "Exponential control barrier functions for enforcing high relative-degree safety-critical constraints". In American Control Conference (ACC) (2016): 322-328.
18. L Wang., *et al*. "Safety barrier certificates for collisions-free multirobot systems". *IEEE Transactions on Robotics* 33.3 (2017): 661-674.
19. S Kolathaya and A D Ames. "Input-to-state safety with control barrier functions". IE*EE Control Systems Letters* 3.1 (2018): 108-113.
20. AJ Taylor., *et al*. "A Control Barrier Perspective on Episodic Learning via Projection-to-State Safety". in IEEE Control Systems Letters 5.3 (2021): 1019-1024.
21. Y Zhu., *et al*. "Policy Iteration for H∞ Optimal Control of Polynomial Nonlinear Systems via Sum of Squares Programming". in IEEE Transactions on Cybernetics 48.2 (2018): 500-509.