Backup Control Barrier Functions: Formulation and Comparative Study

Yuxiao Chen[1], Mrdjan Jankovic[2], Mario Santillo[2], and Aaron D. Ames[1]

Abstract— The backup control barrier function (CBF) was recently proposed as a tractable formulation that guarantees the feasibility of the CBF quadratic programming (QP) via an implicitly defined control invariant set. The control invariant set is based on a fixed backup policy and evaluated online by forward integrating the dynamics under the backup policy. This paper is intended as a tutorial of the backup CBF approach and a comparative study to some benchmarks. First, the backup CBF approach is presented step by step with the underlying math explained in detail. Second, we prove that the backup CBF always has a relative degree 1 under mild assumptions. Third, the backup CBF approach is compared with benchmarks such as Hamilton Jacobi PDE and Sum-of-Squares on the computation of control invariant sets, which shows that one can obtain a control invariant set close to the maximum control invariant set under a good backup policy for many practical problems.

I. INTRODUCTION

Control barrier functions (CBF) [1], [2] were proposed as a method that enforces constraints on dynamic systems, which typically works as a supervisory controller on top of a legacy controller. To guarantee the satisfaction of the safety constraints, a CBF quadratic program (QP) is solved online. While CBF is getting increasingly popular due to its simple implementation and strong guarantee, the construction of CBF is sometimes overlooked.

To guarantee that the CBF QP is always feasible, a control invariant set is needed, which is defined as a set in which any trajectory of the dynamic system can stay indefinitely. The concept of control invariant sets has been studied under various background and names, such as viability kernel [3], infinite time reachable set [4], and various methods have been proposed to compute the control invariant set depending on the system dynamics, see [5] for an overview. Unfortunately, the computation of control invariant sets is notoriously difficult. Even for simple cases such as linear or polynomial dynamics, computation tools such as Minkowski operations [6], robust linear program [7], and Sum-of-Squares [8] do not scale well. For general nonlinear dynamic systems, the standard tool for computing invariant sets is Hamilton Jacobi PDE [9], which typically cannot scale beyond systems with state dimensions 4 due to the exponential complexity.

Due to the difficulty of synthesizing proper control barrier functions based on control invariant sets, CBF QP has

been implemented without control invariant sets. One simple treatment is to assume infinite actuation power [10], in which case the feasibility of the CBF QP can be guaranteed when a simple sign condition is satisfied. To be specific, if one can show that when the Lie derivative of the CBF w.r.t. the input dynamics is zero, the CBF condition is satisfied by the intrinsic dynamics, the CBF QP is always feasible under infinite actuation. Another simplification is to assume that the velocity instead of the acceleration is under control [11], which can guarantee the feasibility of the CBF QP for constraints on the position since the CBF QP can always pick a velocity pointing away from the constraint. In essence, assuming direct velocity control is similar to assuming infinite actuation as an instantaneous change of velocity requires infinite force. Obviously, these assumptions are not true in practice, and the safety of the system is subject to parameter tuning. The simplified approaches might work for low-speed cases as the change of velocity is not severe, but will not work in general for highly dynamic applications.

Another issue of CBF without a control invariant set is the relative degree. For example, a typical vehicle/robot model with acceleration input is a second (or higher) order model with the position states and velocity states. If one directly uses the constraint on the position as a control barrier function, the relative degree of the CBF is 2, and its Lie derivative does not contain the acceleration input. Several solutions have been proposed for the high relative degree, such as Input-Output Linearization [12], [13] and backstepping type formulations [14]. Again, the CBF performance is subject to parameter tuning.

To the best of our knowledge, under limited input, there is not a generic method that guarantees the feasibility of the CBF QP without a control invariant set. All of the above-mentioned methods rely on heuristics and tuning to work in practice. Even when the CBF works well in the test cases, there is no guarantee of performance for cases not included in the test.

The idea of control barrier functions based on backup controllers (referred to as backup CBF for the remainder of this paper) [15] was proposed based on a simple observation that extends a small control invariant set, which is typically easier to obtain, to a larger control invariant set by fixing a backup controller. The backup CBF guarantees the feasibility of the CBF QP and circumvents the difficult computation of a control invariant set by implicitly representing the control invariant set. However, the implicit representation calls for online integration of the dynamics, and the resulting control invariant set is typically suboptimal in the sense that it is not

the maximum control invariant set. This paper is intended as a tutorial to the backup CBF approach by demonstrating the method step by step on some simple examples. We also provide explanations and visualizations on the advantage and disadvantages of the backup CBF over some benchmark formulations.

II. PRELIMINARIES AND A MOTIVATING EXAMPLE

A. Control barrier functions

We begin by a brief review of control barrier functions. Consider the following control affine system:

$$\dot{x} = f(x) + g(x)u, \quad x \in \mathbb{R}^n, u \in \mathcal{U} \subseteq \mathbb{R}^m, \quad (1)$$

where x is the system state and u is the input. Suppose there exists a function $h : \mathbb{R}^n \to \mathbb{R}$ that satisfies

$$\forall\, x \in \mathcal{X}_0, \qquad\qquad\qquad h(x) \geq 0$$
$$\forall\, x \in \mathcal{X}_d, \qquad\qquad\qquad h(x) < 0 \quad (2)$$
$$\forall\, x \in \{x \mid h(x) \geq 0\}, \quad \exists\, u \in \mathcal{U} \text{ s.t. } \dot{h} + \alpha\,(h) \geq 0,$$

where $\mathcal{X}_0$ is the set of initial states and $\mathcal{X}_d$ is the danger set that we want to keep the state away from. $\alpha(\cdot)$ is a class-$\mathcal{K}$ function, i.e., $\alpha(\cdot)$ is strictly increasing and satisfies $\alpha(0) = 0$. Then h is called control barrier function, and for any legacy controller, the CBF controller is a supervisory controller that enforces the state to stay inside $\{x \mid h(x) \geq 0\}$ with the following quadratic programming:

$$u^\star = \underset{u \in \mathcal{U}}{\arg\min}\, \left\| u - u^0 \right\|^2$$
$$\text{s.t. } \nabla h \cdot f(x, u) + \alpha\,(h) \geq 0, \quad (3)$$

where u^0 is the input of the legacy controller.

The third condition in (2) ensures that (3) is always feasible, yet it is difficult to find an h that satisfies it. For clarity, we refer to a function that satisfies the first two conditions in (2) a CBF candidate, a function that satisfies all three conditions a valid CBF.

The CBF condition is closely tied to the concept of a control invariant set, which is defined as follows.

Definition 1. A set $\mathcal{S}$ is a control invariant set if there exists a control law $\pi : \mathbb{R}^n \to \mathcal{U}$ such that for all initial condition $x(0) \in \mathcal{S}$, $\forall t \geq 0, x(t) \in \mathcal{S}$.

It is straightforward to see that for a valid CBF, $\{x|h(x) \geq 0\}$ is a control invariant set. On the other direction, suppose for a CBF candidate h, its 0-level set is a control invariant set. Immediately from Definition 1,

$$h(x) = 0 \to \exists u \in \mathcal{U} \text{ s.t.} \dot{h}(x, u) \geq 0,$$

otherwise the state will exit $\mathcal{S}$, which contradicts Definition 1. Then h can be shown to be a valid control barrier function by picking an α large enough given some continuity condition (Lipschitz continuity of h and $\dot{h}$). As a result, the third condition in (2) is also referred to as the set invariance condition.
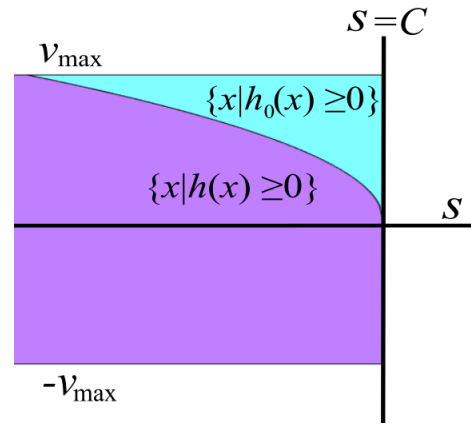

Fig. 1: Difference between h and h_0 for the double integrator example

B. Double integrator example

As a motivating example, consider a simple double integrator with

$$\dot{x} = \begin{bmatrix} \dot{s} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} v \\ u \end{bmatrix}, \quad u \in [-u_{\max}, u_{\max}], v \in [-v_{\max}, v_{\max}].$$

The safety constraint is $X \leq C$ with a constant C. Suppose one directly takes the safety constraint as a CBF candidate: $h_0(x) = C - s$. There are two issues. First, h_0 has relative degree 2, i.e., $\dot{h}_0$ is not a function of u. Second, the set $\{x|h_0(x) \geq 0\}$ is not control invariant. Consider the case where $s = C, v > 0$, the limited input cannot stop the state from crossing into the danger set due to inertia.

For this simple system, it is widely known that a simple valid CBF exists:

$$h(x) = C - s - \mathbb{1}_{v>0}\frac{v^2}{2u_{\max}}, \quad (4)$$

which is the safety constraint combined with the minimum stopping distance. When $h(x) \geq 0$, one can always apply the maximum deceleration $u = -u_{\max}$ until $v = 0$ and the safety constraint remains satisfied. Fig. 1 shows the difference between h and h_0 where h_0 includes a part of the state space from which the safety constraint will eventually be violated.

In fact, the above CBF is exactly the backup CBF with the backup policy being

$$\pi(x) = -\mathbb{1}_{v>0}u_{\max}.$$

We shall show later that the backup policy π takes the system to a small control invariant set $\mathcal{S}_0 = \{x|v = 0\}$, and $\{x|h(x) \geq 0\}$ with h in (4) is the set of initial conditions that can be brought to $\mathcal{S}_0$ while satisfying the safety constraint. In this view, h is induced from the backup policy π.

As mentioned in the Introduction, computing a control invariant set is very difficult, especially for high-dimensional nonlinear systems. We shall show that for complicated dynamic systems, a valid CBF can be fairly easily obtained from a backup policy. Though the CBF induced from a backup policy might not have the nice closed form as in

the double integrator case, the CBF QP can be solved online and is guaranteed to be feasible.

C. Notation and preliminaries

Before going into the detail, some important notations are reviewed. Given the dynamic system in (1), a control policy $\pi : \mathbb{R}^n \to \mathcal{U}$, the closed-loop dynamics under π is $\dot{x} = f_\pi(x) = f(x) + g(x)\pi(x)$. We let $\Phi_{f_\pi} : \mathbb{R}^n \times (-\infty, \infty) \to \mathbb{R}^n$ denote the flow map, i.e., $\Phi_{f_\pi}(x_0, t)$ denotes the solution $x(t_0 + t)$ to the Initial Value Problem (IVP) at time $t_0 + t$ with $x(t_0) = x_0$ under the dynamics f_π. Note that since (1) is time invariant, the initial time t_0 is irrelevant. Moreover, the flow map is additive in t:

$$\Phi_{f_\pi}(\Phi_{f_\pi}(x, t_1), t_2) = \Phi_{f_\pi}(x, t_1 + t_2). \qquad (5)$$

Since the flow map can be written as

$$\Phi_{f_\pi}(x_0, t) = x_0 + \int_0^t f_\pi(x(\tau)) d\tau, \qquad (6)$$

it is differentiable w.r.t. both t and x. By inspection, $\frac{\partial \Phi_{f_\pi}(x_0, t)}{\partial t} = f_\pi(\Phi_{f_\pi}(x_0, t))$. To obtain the partial derivative over x_0, take derivative on both sides:

$$\frac{\partial \Phi_{f_\pi}(x_0, t)}{\partial x_0} = \mathbf{I} + \int_0^t \frac{df_\pi}{dx} \frac{\partial x(\tau)}{\partial x_0} d\tau,$$

from [16], we can define $Q(t) = \frac{\partial \Phi_{f_\pi}(x_0, t)}{\partial x_0}$ as the sensitivity Jacobian, and taking derivative over t on both sides yields

$$\dot{Q}(t) = \frac{df_\pi(x(t))}{dx(t)} Q(t), \qquad (7)$$

where $x(t) = \Phi_{f_\pi}(x_0, t)$. This fact will play a key role on the derivation of the CBF condition.

Proposition 1. *For all $\tau \geq t$,*

$$\frac{d\Phi(x(t), \tau - t)}{dt}\Big|_{u(t) = \pi(t)} = 0. \qquad (8)$$

Proof. Note that the flow map can be written as

$$\Phi_{f_\pi}(x(t), \tau - t) = x(t) + \int_t^\tau f_\pi(x(\bar{\tau})) d\bar{\tau},$$

taking the derivative over t gives

$$\frac{d\Phi(x(t), \tau - t)}{dt} = \dot{x}(t) - f_\pi(x(t)),$$

and the result follows as $\dot{x}|_{u = \pi(x)} = f_\pi(x)$. □

The physical interpretation of Proposition 1 is that the flow under the backup strategy π would not change with time if the current state follows π.

III. BACKUP CBF
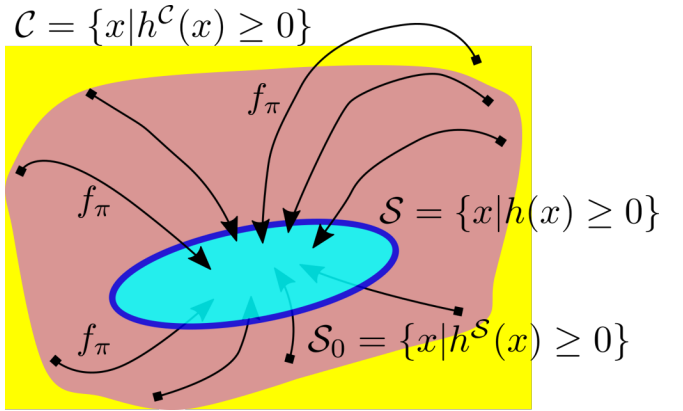
This section presents the backup CBF step by step.

Fig. 2: Enlarging a known control invariant set $\mathcal{S}_0$ with the backup policy π

A. Definition and CBF QP

Consider a safe region $\mathcal{C} \doteq \{x | h^{\mathcal{C}} \geq 0\} \subseteq \mathbb{R}^n$, a control invariant set $\mathcal{S}_0 \doteq \{x | h^{\mathcal{S}} \geq 0\} \subseteq \mathcal{C}$ with $h^{\mathcal{C}}, h^{\mathcal{S}} : \mathbb{R}^n \to \mathbb{R}$ differentiable. Given a backup controller π, define the T-time constrained reachable set as

$$\mathcal{S} = \mathcal{R}(\mathcal{S}_0, \mathcal{C}, f_\pi, T)$$
$$\doteq \{x \in \mathbb{R}^n | \Phi_{f_\pi}(x, T) \in \mathcal{S}_0 \wedge \forall t \in [0, T], \Phi_{f_\pi}(x, t) \in \mathcal{C}\}, \qquad (9)$$

which contains the states from which the flow under f_π will stay in $\mathcal{C}$ and reach $\mathcal{S}_0$ by T. The theoretical root of the backup CBF approach is the following simple observation.

Theorem 1. *[17] Given an initial control invariant set $\mathcal{S}_0$, a safe region $\mathcal{C} \subseteq \mathbb{R}^n$, and a backup controller π, for all $T > 0$, the T-time constrained reachable set $\mathcal{S}$ defined in (9) is a control invariant set and $\mathcal{S}_0 \subseteq \mathcal{S} \subseteq \mathcal{C}$.*

The proof can be found in [17], we provide a brief proof for completeness.

Proof. By definition, $\mathcal{S}_0 \subseteq \mathcal{S}$, and $\mathcal{S} \in \mathcal{C}$ since $\Phi_{f_\pi}(x, 0) \in \mathcal{C}$. By Definition 1, there exists π_0 that keeps $\mathcal{S}_0$ invariant. Define $\pi'(x) = \begin{cases} \pi_0(x), x \in \mathcal{S}_0 \\ \pi(x), x \notin \mathcal{S}_0 \end{cases}$, it can be verified that π' keeps $\mathcal{S}$ invariant. □

Fig. 2 shows theorem 1 pictorially. The yellow square is the state constraint $\mathcal{C}$, given a small known control invariant set $\mathcal{S}_0$ (the blue ellipse), all states that can be driven to $\mathcal{S}_0$ under f_π (the closed-loop dynamics under the backup policy π) while staying inside $\mathcal{C}$ in the meantime forms a larger control invariant set $\mathcal{S}$ shown in brown. Note that the point on the top right is not inside $\mathcal{S}$ because the trajectory under f_π is not completely contained $\mathcal{C}$.

Assumption 1. For simplicity, we assume that the backup controller π keeps $\mathcal{S}_0$ invariant, and the class-$\mathcal{K}$ function α is selected so that $\forall x \in \mathcal{S}_0, \nabla h^{\mathcal{S}} f_\pi(x) + \alpha(h(x)) \geq 0$.

Regarding the horizon T, the following lemma is true.

Lemma 1. *For all $T_1, T_2 > 0$, let $\mathcal{S} = \mathcal{R}(\mathcal{S}_0, \mathcal{C}, f_\pi, T_1)$, then $\mathcal{R}(\mathcal{S}, \mathcal{C}, f_\pi, T_2) = \mathcal{R}(\mathcal{S}_0, \mathcal{C}, f_\pi, T_1 + T_2)$.*

Proof. By definition, $\forall x \in \mathcal{R}(\mathcal{S}, \mathcal{C}, f_\pi, T_2)$, $\Phi_{f_\pi}(x, T_2) \in \mathcal{S} = \mathcal{R}(\mathcal{S}_0, \mathcal{C}, f_\pi, T_2)$. By (5), $\Phi_{f_\pi}(x, T_1 + T_2) = \Phi_{f_\pi}(\Phi_{f_\pi}(x, T_2), T_1) \in \mathcal{S}_0$. And by (9), the trajectory within $[0, T_1 + T_2]$ is contained in $\mathcal{C}$, and the conclusion follows. $\square$

Lemma 1 states that $\mathcal{R}(\mathcal{S}_0, \mathcal{C}, f_\pi, T)$ monotonically increases with T in the set inclusion sense.

With a control invariant set $\mathcal{S}$ defined, a control barrier function h can be defined.

Lemma 2. *$\mathcal{S}$ is the 0-level set of the following function*

$$h(x) = \min\{\min_{t' \in [0,T]} h^\mathcal{C}(\Phi_{f_\pi}(x, t')), h^\mathcal{S}(\Phi_{f_\pi}(x, T))\}. \quad (10)$$

Proof. First notice that by the continuity of the flow function Φ_{f_π} and the min function, h is continuous. For all $x \in \mathcal{S}$, by definition, under the backup strategy π, the state evolution $\Phi_{f_\pi}(x, t')$ would satisfy the constraint and reach $\mathcal{S}_0$ at time T, therefore $h(x) \geq 0$. On the other hand, for all $x \notin \mathcal{S}$, under the backup strategy π, the state evolution either violates the state constraint at some t', i.e., $\exists t' \in [0, T], h^\mathcal{C}(\Phi_{f_\pi}(x, t')) < 0$, or does not reach $\mathcal{S}_0$ within the horizon T, i.e., $h^\mathcal{S}(\Phi_{f_\pi}(x, T)) < 0$, indicating that $h(x) < 0$. Therefore, $\mathcal{S} = \{x | h(x) \geq 0\}$. $\square$

To implement the CBF h, $\dot{h}$ is needed. First h needs to be written as a function of time:

$$h(t) = \min\{\min_{\tau \in [t, t+T]} h^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t)), h^\mathcal{S}(\Phi_{f_\pi}(x(t), T))\}. \quad (11)$$

Since h is defined as the minimum of multiple functions, it may not be differentiable. Even in the differentiable case, the CBF condition may turn out to be nonconvex in the control input u. Instead, the CBF condition is enforced on every τ, which is a sufficient condition for original CBF condition.

First consider $h^\mathcal{S}(\Phi_{f_\pi}(x(t), T))$. By the chain rule,

$$\frac{dh^\mathcal{S}(\Phi_{f_\pi}(x(t), T))}{dt} = \frac{dh^\mathcal{S}}{dx} \frac{\partial \Phi_{f_\pi}(x(t), T)}{\partial x}(f(x(t)) + g(x(t))u(t)),$$

where the sensitivity Jacobian $\frac{\partial \Phi_{f_\pi}(x(t), T)}{\partial x}$ can be calculated with (7). The computation of $h^\mathcal{C}(\Phi_{f_\pi}(x(t)), \tau - t)$ is slightly different:

$$\begin{aligned} &\frac{dh^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t))}{dt} \\ =& \frac{dh^\mathcal{S}}{dx}\left(\frac{\partial \Phi_{f_\pi}(x(t), \tau - t)}{\partial x}\dot{x} - \frac{\partial \Phi_{f_\pi}(x(t), \tau - t)}{\partial t}\right) \\ =& \frac{dh^\mathcal{S}}{dx}\left(\frac{\partial \Phi_{f_\pi}(x(t), \tau - t)}{\partial x}\dot{x} - f_\pi(\Phi_{f_\pi}(x(t), \tau - t))\right), \end{aligned}$$

where $\dot{x} = (f(x(t)) + g(x(t))u(t))$. By Proposition 1, $\frac{d\Phi_{f_\pi}(x(t), \tau - t)}{dt}|_{u=\pi(x)} = 0$, indicating that $\frac{dh^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t))}{dt} = 0$ if $u(t) = \pi(x(t))$.

Remark 1. The difference between the derivative of $h^\mathcal{S}(\Phi_{f_\pi}(x(t), T))$ and $h^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t))$ is that the former contains two parts, the change of the future state following the backup strategy due to the change of current state, and the derivative due to $t + T$ increasing with t; whereas the latter only contains the first part.

The CBF condition is $\dot{h} + \alpha(h) \geq 0$. We shall impose this condition on $h^\mathcal{S}(\Phi_{f_\pi f_\pi}(x, T))$ and $h^\mathcal{C}(\Phi_{f_\pi f_\pi}(x, \tau - t))$ for every $\tau \in [t, t+T]$ instead of only on the τ minimizing $h^\mathcal{C}(\Phi_{f_\pi f_\pi}(x, \tau - t))$, and the CBF QP is then

$$\min_{u \in \mathcal{U}} \|u - u^0\|$$
$$\text{s.t.} \forall \tau \in [t, t+T]$$
$$\frac{dh^\mathcal{C}(\Phi_{f_\pi}(x, \tau - t))}{dt}(x, u) + \alpha(h^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t))) \geq 0$$
$$\frac{dh^\mathcal{S}(\Phi_{f_\pi}(x(t), T))}{dt}(x, u) + \alpha(h^\mathcal{S}(\Phi_{f_\pi}(x(t), T))) \geq 0. \quad (12)$$

Proposition 2. *The constraint in (12) is a sufficient condition for $\dot{h} + \alpha(h) \geq 0$.*

Proof. For notational simplicity, let $\bar{\xi}^\mathcal{C}(x, \tau) \doteq h^\mathcal{C}(\Phi_{f_\pi}(x, \tau - t))$, $\xi^\mathcal{C}(x) \doteq \min_{\tau \in [t, t+T]} \bar{\xi}^\mathcal{C}(x, \tau)$, and $\xi^\mathcal{S} \doteq h^\mathcal{S}(\Phi_{f_\pi}(x, T))$, then $h(x) = \min\{\xi^\mathcal{C}(x), \xi^\mathcal{S}(x)\}$. First notice that

$$\frac{d\xi^\mathcal{C}(x(t))}{dt} + \alpha(\xi^\mathcal{C}(x(t))) \geq \min_{\tau \in [t, t+T]}\left[\dot{\bar{\xi}}^\mathcal{C}(x(t), \tau) + \alpha(\bar{\xi}^\mathcal{C}(x(t), \tau))\right] \geq 0,$$

then

$$\begin{aligned} \dot{h} + \alpha(h) =& \frac{d \min\{\xi^\mathcal{C}(x(t)), \xi^\mathcal{S}(x(t))\}}{dt} + \alpha(\min\{\xi^\mathcal{C}(x(t)), \xi^\mathcal{S}(x(t))\}) \\ \geq& \min\{\dot{\xi}^\mathcal{C}(x(t)) + \alpha(\xi^\mathcal{C}(x(t))), \dot{\xi}^\mathcal{S}(x(t)) + \alpha(\xi^\mathcal{S}(x(t)))\} \geq 0. \end{aligned}$$
$\square$

In practice, $\min_{\tau \in [t, t+T]} h^\mathcal{C}(\Phi_{f_\pi}(x(t)), \tau - t)$ may be difficult to evaluate, so a finite set of $t = \tau_0 < \tau_1 ... < \tau_N = t + T$ is used instead of $[t, t+T]$, and the approximation error can be bounded given the Lipschitz constants of the flow map.

Theorem 2. *Under Assumption 1, the backup CBF QP in (12) is always feasible for $h \geq 0$.*

Proof. Since $h(x) \geq 0$, $\forall \tau \in [t, t+T], h^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t)) \geq 0$, $h^\mathcal{S}(\Phi_{f_\pi}(x(t), T)) \geq 0$. By proposition 1, if $u(t) = \pi(x(t))$, for all $\tau \in [t, t+T]$, $\frac{dh^\mathcal{C}(\Phi_{f_\pi}(x(t), \tau - t))}{dt} = 0$, $\frac{dh^\mathcal{S}(\Phi_{f_\pi}(x(t), T))}{dt} = \frac{dh^\mathcal{S}}{dx} f_\pi(\Phi_{f_\pi}(x(t), T))$. Immediately, the first constraint in (12) is satisfied. By Assumption 1, since $\Phi_{f_\pi}(x(t), T) \in \mathcal{S}_0$, $\nabla h^\mathcal{S} f_\pi(\Phi_{f_\pi}(x(t), T)) + \alpha(h^\mathcal{S}(\Phi_{f_\pi}(x(t), T))) \geq 0$, indicating that the second constraint in (12) is satisfied. Since π is a feasible control policy, $\pi(x(t)) \in \mathcal{U}$, thus a feasible solution to (12). $\square$

B. Relative degree of backup CBFs

Another benefit of the backup CBF is that it is always relative degree one given the system is weakly locally controllable. To show this, we first review some definitions.

Definition 2. A function $h : \mathbb{R}^n \to \mathbb{R}$ is said to have relative degree r with respect to the dynamic system in (1) at a point x_0 if

$$\mathcal{L}_g\mathcal{L}_f^k h(x) = 0, \qquad \forall x \text{ in a neighborhood of } x_0, \forall k \leq r - 2,$$
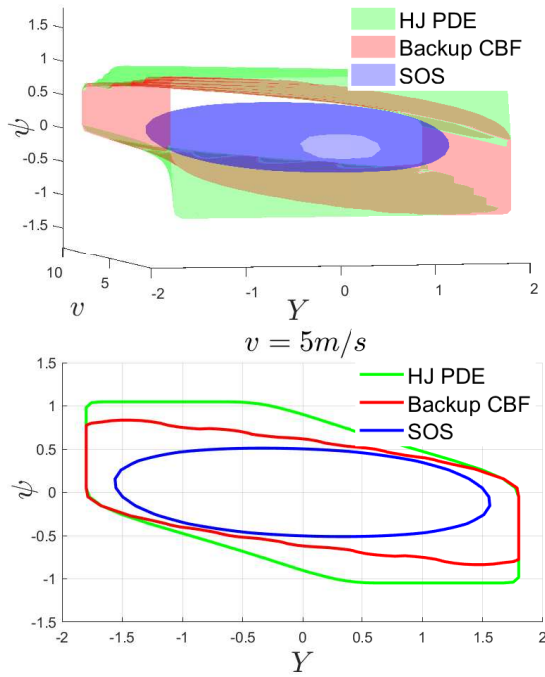$$\mathcal{L}_g\mathcal{L}_f^{r-1} h(x_0) \neq 0$$

Fig. 3: $\{x|h(x) \geq 0\}$ under CBF computed with Hamilton Jacobi PDE (green), backup CBF (red), and SOS (blue)

Fig. 4: $\{x|h(x) \geq 0\}$ with an aggressive backup policy

Definition 3. [18] Given a dynamic system, x_1 is accessible from x_0 via $\Omega \subseteq \mathbb{R}^n$ if there exists a control input signal $\mathbf{u} : [t_0, t_1] \to \mathcal{U}$ such that the trajectory starting at x_0 reaches x_1 at t_1 and $\forall t \in [t_0, t_1], x(t) \in \Omega$. The accessible set of a point x_0 under Ω is the set of all points accessible from x_0 under Ω, denoted as $A_\Omega(x_0)$.

Definition 4. [18] The dynamic system in (1) is locally weakly controllable if for every $x_0 \in \mathbb{R}^n$ and every neighborhood Ω of x, $A_\Omega(x_0)$ has a non-empty interior.

Theorem 3. *If the dynamic system* (1) *is locally weakly controllable, then for an* $h : \mathbb{R}^n \to \mathbb{R}$ *that is not constant for any subset* $\Omega \subseteq \mathbb{R}^n$ *with a non-empty interior,* $h \circ \Phi_{f_\pi}(x, t)$ *has relative degree 1 apart from singular points for all* $t > 0$.

Proof. We prove this by contradiction. Suppose there exists $x_0 \in \mathbb{R}^n$ and a neighborhood Ω where $\forall x \in \Omega, \mathcal{L}_g(h \circ \Phi_{f_\pi})(x) = 0$. Then we can find a small $t_0 > 0$ such that the flow from x_0 stays inside Ω within $t \in [0, t_0]$. Since the system is locally weakly controllable, $A_\Omega(x_0)$ is a compact set with a non-empty interior. With a slight abuse of notation, let $\Phi_{f_\pi}(S, t) = \{x | x = \Phi_{f_\pi}(x_0, t), x_0 \in S\}$ be the image of the flow map acting on a subset $S \subseteq \mathbb{R}^n$. For any $t > 0$, since we only consider regular points, $\Phi_{f_\pi}(\cdot, t)$ has rank n, thus $\Phi_{f_\pi}(A_\Omega(x_0), t)$ has non-empty interior. However, since $\mathcal{L}_g h(x) = 0$ for all $x \in \Omega$, $h(x)$ is constant inside $\Phi_{f_\pi}(A_\Omega(x_0), t)$, which contradicts the assumption that h is not constant in any set with non-empty interior. Thus, $h \circ \Phi_{f_\pi}(x, t + t_0)$ has relative degree 1. Since t_0 can be chosen arbitrarily small, $h \circ \Phi_{f_\pi}(x, t)$ has relative degree 1
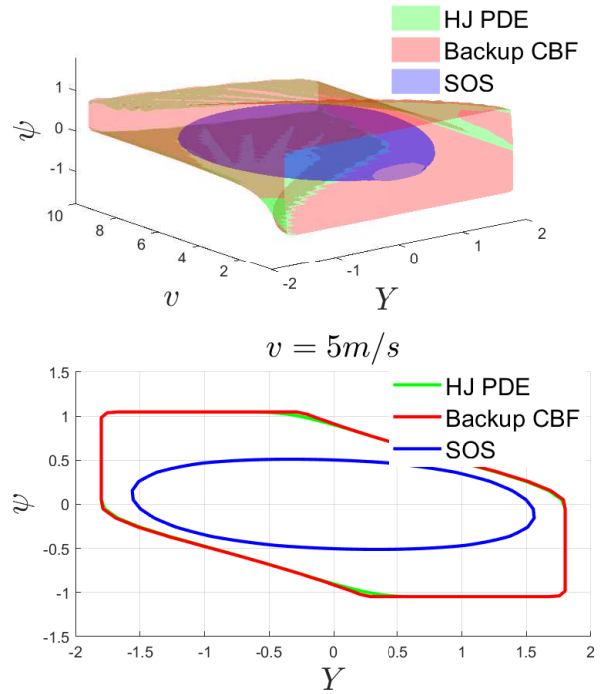
for all $t > 0$. $\qquad\square$

We cannot show that the CBF defined in (10) has relative degree 1 due to the $\min$ function over $t \in [0, T]$, however, note that the constraints in the CBF QP in (12) are on all $h^{\mathcal{C}} \circ \Phi_{f_\pi}(x, \tau - t)$ with $\tau \in [t, t + T]$ and $h^{\mathcal{S}} \circ \Phi_{f_\pi}(x, T)$. Therefore, as long as $h^{\mathcal{S}}$ and $h^{\mathcal{C}}$ are not constant for any subset with a non-empty interior, the CBF QP can be solved as if h has relative degree 1.

Theorem 3 shows that the backup CBF is sufficient without any high order extension or backstepping, it naturally bridges the potentially high relative degree functions $h^{\mathcal{C}}$ and $h^{\mathcal{S}}$ with the input dynamics via the flow map.

IV. COMPARATIVE STUDY

This section presents the comparative study of the backup CBF against some benchmark methods.

One major concern of the backup CBF is that since the control invariant set is induced by a fixed backup policy, how conservative is the resulting CBF? The conservatism of the CBF can be measured by the size of the set $\{x|h(x) \geq 0\}$, which is the induced control invariant set. We compare the backup CBF with the Hamilton Jacobi (HJ) PDE result, which is a close approximation of the maximum control invariant set, and a Sum-of-Squares (SOS) result. The Hamilton Jacobi formulation follows [9] and is computed with the level-set toolbox. In particular, the following HJ PDE is solved:

$$-\frac{\partial h}{\partial t} = H(x, p) = \max_{u \in \mathcal{U}} p^{\mathsf{T}} f(x, u),$$

where $p = \frac{\partial h}{\partial x}$ is the momentum vector. $\{x|h(x,t) \geq 0\}$ is then the backward reachable set at time t. For a sufficiently large t, $\lim_{t \to \infty} h(x,t)$ is an inner approximation of the maximum control invariant set.

The SOS approach follows [19], [20], which computes a valid CBF under a fixed controller. For fairness, we chose the same backup controller π for the backup CBF and the SOS program.

A. Lane keeping with Dubin's car model

To demonstrate the result, a simple lane keeping problem under Dubin's car model is considered:

$$\dot{x} = \begin{bmatrix} \dot{Y} & \dot{v} & \dot{\psi} \end{bmatrix}^{\mathsf{T}} = f(x,u) = \begin{bmatrix} v\sin(\psi) & a & r \end{bmatrix}^{\mathsf{T}} \quad (13)$$

where the state x consists of lateral position Y, velocity v, and heading angle ψ, and the input consists of acceleration a and yaw rate r, The state constraint considered is $\mathcal{C} = \{|Y| \leq Y_{\max} \land |\psi| \leq \psi_{\max}\}$ with $Y_{\max} = 1.8m, \psi_{\max} = \pi/3$, and $\mathcal{U} = \{a, r||a| \leq a_{\max} \land |r| \leq r_{\max}\}$.

The backup policy is chosen as

$$\mathbf{u}(x) = \begin{bmatrix} a \\ r \end{bmatrix} = \begin{bmatrix} \mathbf{Sat}_{a_{\max}}(k_v(v_{des} - v)) \\ \mathbf{Sat}_{r_{\max}}(k_y[Y; \psi]), \end{bmatrix} \quad (14)$$

where $\mathbf{Sat}$ is the saturation function, $k_v > 0$ is a constant and k_y is calculated via LQR. Under the same controller π, a CBF is synthesized with the following SOS program:

$$\begin{aligned} \max \ & R \\ \text{s.t. } & h(x) - s_1(x)(Q(x) - R^2) \in \Sigma[x] \\ & -h(x) - s_2(x)(W^2 - Y^2) \in \Sigma[x] \\ & \nabla h \cdot f(x, \mathbf{u}(x)) + \alpha(h) - s_3(x)(Q(x) - R_0^2) \in \Sigma[x] \\ & s_1(x), s_2(x), s_3(x) \in \Sigma[x], \end{aligned}$$

$$(15)$$

with $Q(x) \doteq \hat{x}^{\mathsf{T}} Q \hat{x}$, and $\hat{x} = [Y, v - v_{des}, \psi]^{\mathsf{T}}$, where Q is a PSD matrix. $\Sigma[x]$ is the set of sum-of-squares polynomial of x, s_1, s_2, s_3 are SOS multipliers for the Positivstellensatz procedure that enforces positive definiteness of a polynomial on a semialgebraic set. Due to the cross product term in line 2, (15) cannot be directly solved via SOS, a line search is used to find the largest R that renders the SOS program feasible. R^0 is a radius that contains the region of state space we are interested in, (15) essentially tries to fit the largest ball of $Q(x)$ inside $\{h(x) \geq 0\}$ where h is the CBF that satisfies (2). For simplicity, $\sin\psi \approx \psi - \psi^3/3$, and the saturation is lifted, which should lead to a larger control invariant set.

Fig. 3 shows the comparison of the CBF computed with HJ, backup CBF, and SOS where the upper plot shows the 3D surface and the lower plot shows the slicing of the set at $v = 5m/s$. As expected, Hamilton Jacobi PDE generates the largest control invariant set, the one from backup CBF is smaller, and the SOS one is the smallest since h is restricted to be a polynomial (4-th order in this case).

Obviously, one can choose a better backup policy and greatly improve the result. For example, if we change $v_{des} = 0$, and switch to a more aggressive LQR design, the control invariant set computed from backup CBF is almost identical to the one from Hamilton Jacobi, as shown in Fig. 4.
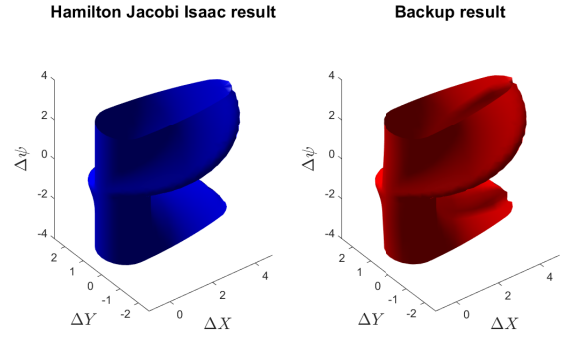

Fig. 5: Danger set of the aeroplane collision avoidance

B. Aeroplane avoidance

One classic example in safety-critical control is the aeroplane avoidance with the following dynamics:

$$\begin{bmatrix} \Delta \dot{X} \\ \Delta \dot{Y} \\ \Delta \dot{\psi} \end{bmatrix} = \begin{bmatrix} -v_a + v_b \cos(\Delta\Psi) + u\Delta Y \\ v_b \sin(\Delta\psi) - u\Delta X \\ -u \end{bmatrix}, \quad (16)$$

where ΔX, ΔY, and $\Delta\psi$ are the difference of X, Y coordinates and heading angles plane a and b, u is the turning rate of plane a, v_a and v_b are their velocities, assumed to be constant.

The backup CBF approach is not good at handling malicious disturbance, therefore we assume that plane b maintains a fixed orientation. The safety constraint is defined as $\Delta X^2 + \Delta Y^2 \geq R^2$, where R is the minimum distance to maintain between the two planes. Fig. 5 shows the danger set ($\{x|h(x) \leq 0\}$) computed with the HJI PDE and the one computed with a simple backup policy:

$$u = -u_{\max}\mathbf{sign}(\Delta Y). \quad (17)$$

The two danger sets are almost identical.

C. Computation

Technically, it is not fair to compare the computation time of the backup CBF approach with the benchmarks as the former does not require any offline computation while the online computation is more complicated than normal CBF QP, while CBF based on explicit control invariant set requires computing the control invariant set offline.

Table I shows the computation time of the CBF QP with CBF constructed by the HJ PDE, SOS, and the backup CBF approach. The computation time for backup CBF consists of two parts, the integration of the backup policy, and the CBF QP. As shown by the result, the majority of solver time is spent on the CBF QP, which is 1 magnitude larger than the time used to solve the CBF QP with explicit form. This is mainly because (12) enforces the CBF condition on every $\tau \in [t, t+T]$ (replaced with a sequence of τ_i in practice), which significantly increases the number of constraints. The integration time is relatively small compared to the QP time and can be further reduced. For instance, [21] uses a Koopman operator approach to further simplify

	Dimension	HJ Offline	HJ online	SOS offline	SOS online	Backup online integration	Backup CBF QP
Dubin's car	3	37.3s	0.5ms	3.6s	0.4ms	1.4ms	3.6ms
Quadrotor	16	NA	NA	NA	NA	4.92ms	28.33ms

TABLE I: Computation time of backup CBF, CBF based on HJ, and CBF based on SOS

the online integration. We also apply the backup CBF on a 16-dimensional quadrotor model, which is way beyond the limit of HJ (maximum dimension 4-5) and SOS (maximum dimension around 8-10), and the backup CBF can still be implemented with a reasonable loop rate.

V. Conclusion

We give a tutorial of the backup CBF approach and compare it to the HJ PDE approach and the SOS approach as benchmarks. The result shows that the implicit control invariant set induced by the backup policy is close to the maximum control invariant set under a properly chosen backup policy in many practical problems. The backup CBF is much more scalable than the benchmark methods with explicitly computed control invariant sets and is applicable to general nonlinear dynamics. Furthermore, since the backup CBF has relative degree 1 under mild assumptions, it is a better choice than the high-order CBFs that bear no feasibility guarantee.

References

[1] A. D. Ames, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 6271–6278.

[2] P. Wieland and F. Allgöwer, "Constructive safety using control barrier functions," *IFAC Proceedings Volumes*, vol. 40, no. 12, pp. 462–467, 2007.

[3] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory: new directions*. Springer Science & Business Media, 2011.

[4] D. Bertsekas, "Infinite time reachability of state-space regions by using feedback control," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 604–613, 1972.

[5] F. Blanchini, "Set invariance in control," *Automatica*, vol. 35, no. 11, pp. 1747–1767, 1999.

[6] S. Rakovic, P. Grieder, M. Kvasnica, D. Mayne, and M. Morari, "Computation of invariant sets for piecewise affine discrete time systems subject to bounded disturbances," in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 2. IEEE, 2004, pp. 1418–1423.

[7] Y. Chen, H. Peng, J. Grizzle, and N. Ozay, "Data-driven computation of minimal robust control invariant set," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 4052–4058.

[8] M. Korda, D. Henrion, and C. N. Jones, "Convex computation of the maximum controlled invariant set for polynomial control systems," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2944–2969, 2014.

[9] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games," *IEEE Transactions on automatic control*, vol. 50, no. 7, pp. 947–957, 2005.

[10] A. Taylor, A. Singletary, Y. Yue, and A. Ames, "Learning for safety-critical control with control barrier functions," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 708–717.

[11] L. Lindemann and D. V. Dimarogonas, "Control barrier functions for signal temporal logic tasks," *IEEE control systems letters*, vol. 3, no. 1, pp. 96–101, 2018.

[12] Q. Nguyen and K. Sreenath, "Exponential control barrier functions for enforcing high relative-degree safety-critical constraints," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 322–328.

[13] X. Xu, "Constrained control of input–output linearizable systems using control sharing barrier functions," *Automatica*, vol. 87, pp. 195–201, 2018.

[14] W. Xiao and C. Belta, "Control barrier functions for systems with high relative degree," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 474–479.

[15] T. Gurriet, M. Mote, A. Singletary, P. Nilsson, E. Feron, and A. D. Ames, "A scalable safety critical control framework for nonlinear systems," *IEEE Access*, vol. 8, pp. 187 249–187 275, 2020.

[16] H. Seywald and R. Kumar, "Desensitized optimal trajectories," *Analytical Mechanics Associates Rept*, pp. 03–16, 2003.

[17] T. Gurriet, M. Mote, A. D. Ames, and E. Feron, "An online approach to active set invariance," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 3592–3599.

[18] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on automatic control*, vol. 22, no. 5, pp. 728–740, 1977.

[19] X. Xu, J. W. Grizzle, P. Tabuada, and A. D. Ames, "Correctness guarantees for the composition of lane keeping and adaptive cruise control," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 3, pp. 1216–1229, 2017.

[20] Y. Chen, A. Hereid, H. Peng, and J. Grizzle, "Enhancing the performance of a safe controller via supervised learning for truck lateral control," *Journal of Dynamic Systems, Measurement, and Control*, vol. 141, no. 10, 2019.

[21] C. Folkestad, Y. Chen, A. D. Ames, and J. W. Burdick, "Data-driven safety-critical control: Synthesizing control barrier functions with koopman operators," *IEEE Control Systems Letters*, 2020.