2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

High-confidence Barrier-certified Control Design using Goal-oriented Scenario Optimization and Experience Replay Model Learning

Zahra Marvi, Bahare Kiumarsi, Hamidreza Modares

Abstract— **This paper presents a probabilistic framework to design safe controllers for linear systems under parametric uncertainties. To this end, a two-layer learning-enabled controller is presented that unifies the experience-replay model learning and the scenario-based optimization. The inner loop leverages the scenario optimization to impose probabilistic stability and safety specifications through sampling from a control Lyapunov function and a control barrier function, respectively. Each sample represents a plausible system dynamics realization within the range of the uncertainties. To quantify the model uncertainty and, thus, to facilitate a proactive and goal-oriented sampling of safety and stability constraints, an experience replay-based model learning is presented in the outer loop. The exponentially fast convergent guarantees of the presented approach and the quantification of the exponential rate using the collected data allow us to quantify the ambiguity set for the system parameters based on the data informativeness. The quantified modeling error acts as a vanishing perturbation to the true dynamics, from which samples can be taken at a specific frequency to solve an optimization problem in the inner loop. The presented approach provides safety and stability guarantees with high probability, even during learning. Simulation is used to depict the efficacy of the proposed approach.**

Index Terms— **Safe control, scenario optimization, experience replay, uncertain dynamics, model learning**

I. Introduction

Successful deployment of autonomous control systems in the real world demands developing systematic and tractable control design methods that account for the satisfaction of safety constraints despite uncertainties. To account for safety constraints, the model predictive control (MPC) [1] formulates the optimal safe control design as a constrained optimization problem. The resulting optimization problem is convex if the control system is linear and the cost is quadratic. As another approach to deal with safety constraints, reachability-based methods [2] compute the set of initial states that can be kept inside the safe set as well as their corresponding control actions. However, the computational cost of computing the set of reachable states is significant in general. As a tractable safe control design approach, control barrier function (CBF)-based methods [3], [4] assure safety by guaranteeing the forward invariance of the safe set. Most CBF-based methods rely on complete knowledge of the system dynamics. To deal with model uncertainties, the uncertainties are modeled in [5] as a Gaussian process. The CBF is then formed to ensure the forward invariance of the safe set for the worst-case model realization. The effect of the model perturbation on the safety using CBFs is investigated in [6] by guaranteeing the forward invariance of a super-level set of the safe set and asymptotic convergence of system's trajectories to the safe set. However, this approach does not prevent the violation of strict safety constraints. Adaptive CBF (aCBF) and robust aCBF are developed in [7], [8] to extend CBFs to systems with parametric uncertainties. Chance constraint CBFs are developed for the safety of systems under measurement noise in [9], [10]. In most CBF-based approaches, safety and performance specifications are integrated through a quadratic optimization (QP) problem for which control Lyapunov functions (CLFs) and CBFs are imposed as soft and hard inequality constraints, respectively. Since these constraints must be satisfied for the entire range of the uncertain system parameters, the quadratic programming problem turns into an optimization problem with an infinite number of constraints, which is not tractable in general. Moreover, finding the CLF/CBF functions that are valid for all possible values of uncertainties is a daunting challenge. Even though safe model learning and control of uncertain linear systems are considered in [11], an entirely different approach is considered in this paper to tackle the same problem. That is, the adaptive robust control barrier function (ARCBF) is used in [11] for which the worst-case learning error is employed. To further reduce the conservatism, a two-layer control framework is presented in this paper, and experience replay model learning is performed in the outer layer to quantify the uncertain set, and the scenario optimization is leveraged in the inner layer to proactively sample from the uncertain set and impose sampled safety constraints. The sampling approach will reduce the conservatism of the learning-based control design and also assure its tractability.

Uncertainty quantification and reduction through data collection are of vital importance in reducing the conservatism of safe control design methods. This is because overly conservative uncertainty quantification can result in jeopardizing the system's performance and can even lead to infeasibility. While the epistemic uncertainty (i.e., lack of knowledge about the system dynamics) can be reduced as more data become available, to the best of our knowledge, there is no systematic approach to quantify the uncertainty reduction, and therefore, imposing less restrictive CBFs constraints. In

Zahra Marvi was supported with Michigan State University startup account. Bahare Kiumarsi and Hamidreza Modares were supported with Department of Navy award N00014-22-1-2159 issued by the Office of Naval Research.

Zahra Marvi is with the University of Minnesota Twin Cities, Bahare Kiumarsi is with the Department of Electrical and Computer Engineering, and Hamidreza Modares is with the Department of Mechanical Engineering, Michigan State University. Email: {marvi099@umn.edu, kiumarsi@msu.edu, and modaresh@msu.edu }

this paper, a two-layer learning-enabled safe control framework is presented for the control of uncertain constrained linear systems. The higher layer provides an uncertainty quantification and reduction mechanism that maps the current data set into the parametric uncertainty set from which the lower layer can sample. The lower layer then leverages the scenario approach [12] along with the CBFs to provide high-probability safety guarantees. In contrast with the existing scenario methods in the literature, a proactive sampling approach is presented, in which the information about the set of uncertain parameters with a high probability of constraint satisfaction is updated and rapidly learned. This is achieved by proposing a model learning based on experience replay as the outer governing layer of the controller. The history of data is used in the update law of the experience replay approach, which also provides an easy-to-verify persistence of excitation (PE) condition [13]. This provides three significant attributions to the overall framework. First, in contrast with adaptive methods, it guarantees exponentially fast model-learning convergence and safety during learning. Second, it facilitates the derivation of uncertain set boundaries; thus, the informative sampling is performed based on a better approximation of the uncertain set at each iteration. Third, the features of the experience replay method make the learning error a vanishing perturbation to the exact model and this enables stability guarantee using the perturbation theory without even the need for sampling. This provides an efficient, non-conservative, and tractable solution to the safe control of uncertain systems with fast and guaranteed model learning.

A. Notation

$int\,\mathcal{C}$ indicates the interior of a set $\mathcal{C}$ while its boundary is denoted by $\partial\mathcal{C}$. The Euclidean norm of a vector is shown by $\|\cdot\|$. C^1 stands for the set of continuously differentiable functions. $A\backslash B$ indicates the set of elements that they are in A and are not in B. All random variables are assumed to be defined on a probability space $(\Omega, \mathcal{F}, Pr)$, with Ω as the sample space, $\mathcal{F}$ as its associated σ-algebra and Pr as the probability measure.

II. Background and Problem Formulation

The system dynamics is considered as

$$\dot{x} = Ax + Bu \tag{1}$$

with $x \in \mathbb{R}^n$ as the state of the system and $u \in \mathcal{U} \subset \mathbb{R}^m$ as the control input. A and B are the nominal dynamics and control input matrices, respectively. The system is assumed to be stabilizable.

A. Control Barrier Functions (CBFs)

CBFs enable restricting the trajectories of the system to evolve within a pre-defined safe set by providing conditions for the control input. The safe set is defined as

$$\mathcal{C} = \{x | h(x) \geqslant 0\} \tag{2}$$

where $h(x) : \mathbb{R}^n \to \mathbb{R}$ is a C^1 function which represents the safety criterion.

Definition 1. A continuous function $\alpha : (-b, a) \to (-\infty, \infty)$ with $a, b > 0$ is an extended class $\mathcal{K}$ function, if it is strictly increasing and $\alpha(0) = 0$ [6], [14]. ∎

Definition 2. [3] Consider the system (1) and the safe set $\mathcal{C} \subset \mathbb{R}^n$ (2). If there exists a locally Lipschitz extended class $\mathcal{K}$ function α such that

$$\sup_{u\in\mathcal{U}} [\frac{\partial h}{\partial x}Ax + \frac{\partial h}{\partial x}Bu + \alpha(h(x))] \geqslant 0, \quad \forall x \in \mathcal{D} \tag{3}$$

then, the function $h(x)$ is a CBF on $\mathcal{D}$ with $\mathcal{C} \subseteq \mathcal{D} \subset \mathbb{R}^n$. ∎

Accordingly, the set of safe control inputs for $h(x)$ is

$$\mathcal{U}_m(x) = \{u \in \mathcal{U} | \frac{\partial h}{\partial x}Ax + \frac{\partial h}{\partial x}Bu + \alpha(h(x)) \geqslant 0\}$$

Guaranteeing the forward invariance of the set $\mathcal{C}$, and thus the system's safety, using CBFs is formally presented in Theorem 1.

Theorem 1. *[3] Consider the system* (1) *and the set* $\mathcal{C} \subseteq \mathcal{D}$ (2) *defined for a* C^1 *function* $h(x)$. *If h is a CBF on* $\mathcal{D}$, *then, any Lipschitz continuous controller* $\{u : \mathcal{D} \to \mathbb{R} | u \in \mathcal{U}_m(x)\}$ *renders the set* $\mathcal{C}$ *forward invariant.*

B. Scenario-based Optimization

Consider the following robust convex programming (RCP) [15]

$$\text{RCP}: \min_{\gamma} \; c^T\gamma \qquad \text{s.t. } f_\delta(\gamma) \leqslant 0, \;\; \forall \delta \in \Delta \tag{4}$$

where γ is the optimization variable, $f_\delta(\gamma)$ are convex functions for every realization of the uncertain parameter δ in the uncertainty set Δ. For a continuous uncertainty set Δ, (4) becomes a semi-infinite optimization problem, i.e., optimization with a finite number of optimization variables and an infinite number of constraints.

Assume that N independent identically distributed (iid) instances or scenarios of the uncertain parameter δ are sampled according to a probability distribution as $\delta^{(1)}, \delta^{(2)}, ..., \delta^{(N)}$. Then, the scenario convex programming (SCP) of (4) is as follows,

$$\text{SCP}_N: \min_{\gamma} \; c^T\gamma \qquad \text{s.t. } f_{\delta^{(i)}}(\gamma) \leqslant 0, \;\; i = 1, ..., N \tag{5}$$

which, in contrast to (4), provides a tractable standard convex optimization problem with a finite number of constraints. The advantage of the scenario approach is that the solution to (5) satisfies unseen constraints as well, except for a small fraction of it, which rapidly goes to zero with an increment in the number of samples.

Theorem 2. *Let* ${\gamma_N}^*$ *be the solution to* (5). *Set a violation parameter* $\epsilon \in (0, 1)$ *and a confidence parameter* $\beta \in (0, 1)$. *Let the number of iid samples satisfy the following condition*

$$N > \frac{2}{\epsilon}(\ln\frac{1}{\beta} + m) \tag{6}$$

where m is the number of optimization variables. Then, with the probability of no smaller than $1 - \beta$, ${\gamma_N}^*$ *is an* ϵ*-level robustly feasible solution. That is, it satisfies* $Pr(\delta \in \Delta : f_\delta({\gamma_N}^*) > 0) \leqslant \epsilon$.

Proof: See [12], [16].

A lower bound for the number of samples is derived in [17] based on the size of the feasible set. Let the number of samples satisfy

$$N \geqslant \frac{1}{2(\epsilon)^2}\log(\frac{|U\backslash U_\epsilon|}{\beta}) \tag{7}$$

Then, it yields an ϵ-level robustly feasible solution with the probability of at least $1-\beta$, where U is the set of all feasible values of γ, while $U_\epsilon \in U$ is the set of decision variable γ for which the optimization constraints are satisfied with the probability of higher than $1-\epsilon$. We later show employing a proper learning scheme can decrease this lower bound and thus decrease the minimum number of samples to achieve a specific violation probability.

C. Problem Statement

Considering the uncertain system dynamics as

$$\dot{x} = Ax + Bu - \Delta(x)^T\theta \tag{8}$$

where $x \in \mathbb{R}^n$ is the system state, $u \in \mathbb{R}^m$ is the control input, and $\theta \in \mathbb{R}^p$ is the vector of unknown parameters. A and B are the nominal dynamics and control input matrices, respectively. It is assumed that the pair (A, B) is stabilizable.

The goal is to design a controller that ensures safety despite uncertainty while guaranteeing a stability condition, which can impose a performance specification as much as possible. Safety and stability specifications are encoded via CBF and control Lyapunov function (CLF), respectively. These functions are unified through QP optimization. Using a quadratic Lyapunov function $V(x) = x^T Px$ and a CBF $h(x)$, the QP optimization problem formulation for achieving this overall goal for the uncertain system (8) becomes

$$\text{QP}: \min_{u_l=[u,\eta]} \quad {u_l}^T H u_l + F u_l \quad \text{s.t.} \quad f_\theta(u) < 0, \quad \forall\theta \in \Theta \tag{9}$$

where

$$f_\theta(u) = \begin{bmatrix} -\frac{\partial h}{\partial x}(Ax + Bu) - \alpha(h(x)) + \frac{\partial h}{\partial x}\Delta^T(x)\theta \\ -x^T Qx + 2x^T P(Bu - \Delta^T(x)\theta) + \lambda_{min}(Q)||x||^2 - \eta \end{bmatrix} \tag{10}$$

and $A^T P + PA = -Q$, and η is the relaxation factor. In the face of uncertainty, CBF and CLF inequality constraints depend on an uncertain parameter θ, and thus, the QP problem turns into an optimization with the infinite number of constraints, which is not tractable. We leverage the scenario approach to solve this problem along with an efficient learning scheme to learn not only the uncertain parameters but also quantify them over time to present a proactive sampling approach that can achieve a higher probability of achieving a safe and stabilizing controller.

III. Learning-enabled Scenario-based Quadratic Programming

This section presents the idea of experience replay model learning, learning-enabled scenario-based QP, and its theoretical development toward safe and stable control design of systems under parametric uncertainty.

A. Experience Replay Model Learning

Let $\hat{\theta}$ be an estimation of θ and $\tilde{\theta} = \theta - \hat{\theta}$ be the estimation error. The following filters in terms of σ,l, x_s and Ω, are respectively applied to $\dot{x}$, $Ax + Bu$, x and $\Delta(x)$ in (8) as

$$\dot{\sigma}(t) = -\beta\sigma(t) + \dot{x}(t) \tag{11}$$

$$\dot{l}(t) = -\beta l(t) + Ax + Bu \tag{12}$$

$$\dot{x}_s(t) = -\beta x_s(t) + x(t) \tag{13}$$

$$\dot{\Omega} = -\beta\Omega(t) - \Delta^T(x) \tag{14}$$

where $\beta > 0$ is a design gain and $\Omega(0) = 0$, $x_s(0) = 0$, $\sigma(0) = 0$, $l(0) = 0$. The filtered signals in (11)-(14) are given, respectively as

$$\sigma(t) = e^{-\beta t}\int_0^t e^{\beta\tau}\dot{x}(\tau)d\tau \tag{15}$$

$$l(t) = e^{-\beta t}\int_0^t e^{\beta\tau}(Ax + Bu)d\tau \tag{16}$$

$$x_s(t) = e^{-\beta t}\int_0^t e^{\beta\tau}x(\tau)d\tau \tag{17}$$

$$\Omega(t) = -e^{-\beta t}\int_0^t e^{\beta\tau}\Delta^T(\tau)d\tau \tag{18}$$

where $\delta \in \mathbb{R}^n$, $\Omega \in \mathbb{R}^{n\times(mn+n^2)}$, and $x_s \in \mathbb{R}^{(m+n)}$. The filtered signals are used to rewrite the system dynamics (8) as

$$\sigma(t) = \Omega(t)\theta + l(t) \tag{19}$$

From (11), σ can be stated based on known variables $x(t)$ and $x_s(t)$ as

$$\sigma(t) = x(t) - e^{-\beta t}x(0) - \beta x_s(t) \tag{20}$$

Based on (19) and (20), the prediction error is formed as

$$e(t) = \sigma(t) - \Omega(t)\hat{\theta}(t) - l(t) \tag{21}$$

To store and employ the history of data in the update law, two memory stacks $\{\sigma_i - l_i\}_{i=1:p}$, $\{\Omega_i\}_{i=1:p}$ are collected, which store the values of $\sigma(t_i) - l(t_i)$ and $\Omega(t_i)$, respectively at each time instance t_i. The prediction error for the past time instance t_i using the current estimation of the uncertain parameters is then formed as

$$e_i(t) = \sigma_i - \Omega_i\hat{\theta}(t) - l_i \tag{22}$$

Therefore, the update law employing the past stored data is presented as

$$\dot{\hat{\theta}} = \beta_{\theta1}\Omega^T(t)e(t) + \beta_{\theta2}\sum_{i=1}^{p}{\Omega^T}_i e_i(t) \tag{23}$$

where $\beta_{\theta1}$ and $\beta_{\theta2}$ are positive scalar gains. Under a rank condition and in the presence of rich data, the update law (23) guarantees exponential convergence of $\hat{\theta}$ to θ.

The convergence is proved using the Lyapunov function V_θ as

$$V_\theta = \frac{1}{2}\tilde{\theta}^T\tilde{\theta} \tag{24}$$

Using (19) and (21), the prediction error can be written as $e(t) = \Omega\tilde{\theta}$. Therefore using (23), one has

$$\dot{\tilde{\theta}} = -\dot{\hat{\theta}} = -\beta_{\theta1}\Omega^T(t)\Omega\tilde{\theta} - \beta_{\theta2}\sum_{i=1}^{p}{\Omega^T}_i\Omega_i\tilde{\theta}$$

Therefore derivative of (24) turns into

$$\dot{V}_\theta = -\beta_{\theta1}\tilde{\theta}^T\Omega^T(t)\Omega\tilde{\theta} - \beta_{\theta2}\tilde{\theta}^T\sum_{i=1}^{p}{\Omega^T}_i\Omega_i\tilde{\theta} \leqslant 0 \tag{25}$$

Thus, based on Lyapunov analysis, the error is bounded, and under a rank condition, the summation term becomes positive definite, which ensures the asymptotic stability of the error. This result is formally presented in the following Lemma.

Lemma 1. *[13] Consider the update law* (23). *Let there exist p^* such that for all $p \geqslant p^*$, and for any sequence $t_1 < t_2 < .... < t_p$, one has*

$$rank([{\Omega_1}^T, {\Omega_2}^T, ..., {\Omega_p}^T]) = mn + n^2 \tag{26}$$

Then, $\hat{\theta}$ converges to θ exponentially fast. Moreover, considering the Lyapunov function

$$V_\theta = \frac{1}{2}\tilde{\theta}^T\tilde{\theta} \tag{27}$$

then, there exists $\beta_{\theta 12}$ such that

$$\dot{V}_\theta \leqslant -2(\beta_{\theta 12})V_\theta \tag{28}$$

Remark 1. The experience-replay model learning employs the regressor representation of the dynamics and provides an update law that incorporates the past stored data. If the stored data satisfies the rank condition (26), then in accordance with (28), the approximation error converges to zero exponentially fast.

Remark 2. Note that $\beta_{\theta 12}$ depends on the smallest eigenvalue of the rank matrix in (26).

The result of Lemma 1 can also be used to find the boundaries of the uncertain set and achieve a higher accuracy as the learning progresses.

Definition 3. Conditional Ambiguity Set. Consider the system (8) and the update law (23). If available data satisfies the rank condition (26), then the conditional ambiguity set Θ_c is defined based on available approximation of θ at each iteration as

$$\Theta_c = \{\theta | ||\hat{\theta}|| - ||\tilde{\theta}(0)||e^{-(\beta_{\theta 12})(t)} \leqslant ||\theta|| \leqslant ||\hat{\theta}|| + ||\tilde{\theta}(0)||e^{-(\beta_{\theta 12})(t)}\} \tag{29}$$

Corollary 1. *Considering the system* (8) *and the update law* (23)*, if the rank condition* (26) *is satisfied, then the conditional ambiguity set* (29) *is the set of all possibilities of the uncertain parameter θ and its lower and upper bounds converge together exponentially fast to a single point which is the true value of θ.*

Proof: Consider the inequality derived in (28). According to the comparison Lemma one has

$$V_\theta \leqslant V_\theta(0)e^{-2(\beta_{\theta 12})(t)}$$

Thus, from (27), one has

$$||\tilde{\theta}|| \leqslant ||\tilde{\theta}(0)||e^{-(\beta_{\theta 12})(t)} \tag{30}$$

The uncertain parameter is $\theta = \hat{\theta} + \tilde{\theta}$. Thus, from (30), the conditional ambiguity set (29) is formed.

In addition, considering the set bounds, $\hat{\theta}$ converges to the true value θ^* and this guarantees the exponential convergence of the error envelop to zero. This completes the proof.

This gives the exponential envelope of the learning error, and thus, Θ is shrinking exponentially fast. This feature is especially important when it is used in conjunction with the scenario approach. Since scenario optimization is built upon a sampling of constraints, this learning method provides a smaller sampling set after each update.

B. Experience Replay Scenario-based Optimization

To avoid the infinite number of samples and provide a tractable safe solution, important sampling from the updated set of uncertain parameters (29) is performed.

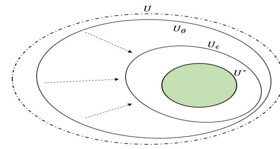

Fig. 1. Set Convergence

By extracting N iid samples $\theta^{(1)}, ..., \theta^{(N)}$ from (29), the corresponding scenario-based quadratic programming (SQP) becomes

$$\text{SQP}: \min_{u_l=[u,\eta]} \; {u_l}^T H u_l + F u_l$$
$$\text{s.t.} \; f_{\theta^{(i)}}(u) \leqslant 0, \; i = 1, ..., N \tag{31}$$

Let the solutions to the SQP with N samples be denoted by $u_N{}^*$. Based on desired probabilistic characteristics, samples are chosen according to Theorem 2. Thus, the problem reduces to a convex optimization with the finite number of samples.

Remark 3. Note that iid sampling from an uncertain set of system parameters can be easily performed, in contrast to the cases for which ii samples of the system states are required.

The safety and stability of the system under the proposed approach are presented as follows.

Theorem 3. *Consider the approximation of* (8) *as*

$$\dot{x} = Ax + Bu - \Delta^T(x)\hat{\theta} \tag{32}$$

Let $u = kx$, where k is a stabilizing gain for (32) *with $x = 0$ as the exponentially stable equilibrium point of the closed loop approximated system. Let $V(x) = x^T P x$ be the Lyapunov function for* (32)*, where P is the solution to the Lyapunov equation. Suppose that the update law* (23) *is employed and* (26) *is satisfied. Then, the origin is an exponentially stable equilibrium point for* (1).

Proof: The system dynamic (1) can be written based on (32), and the error bound as

$$\dot{x} = Ax + Bu - \Delta^T(x)\hat{\theta} - \Delta^T(x)\tilde{\theta}$$

where $\Delta^T(0)\tilde{\theta} = 0$. In addition, considering (30) one has,

$$\Delta^T(x)||\tilde{\theta}|| \leqslant \Delta^T(x)||\tilde{\theta}(0)||$$

Thus, there exists a coefficient γ such that $||\Delta^T(x)\tilde{\theta}|| \leqslant \gamma||x||$. That is, the error bound satisfies a linear growth bound condition, which makes it a vanishing perturbation to the approximated system. Therefore, according to the perturbation theorem, the exponential stability of the approximated system results in the exponential stability of the original system as well. For more details, see [11].

It is shown so far that the experience replay learning method provides an accurate bound of the error and, thus, the boundaries of the uncertain set. In conjunction with scenario optimization, this significant feature enables proactive sampling and an increased probability of satisfying all constraints at each iteration. This result is formally presented in the next section.

IV. Overall framework

The proposed two-layer framework is presented in this section. The overall control scheme is depicted in Figure 2, which consists of a core control of SQP to calculate the least control effort to ensure safety and stability while the model-learning behaves as a governing outer-loop providing a more accurate approximation of the system, and the uncertain set which will be further used by the core loop and this cycle continues. The detailed algorithm is as follows.

Algorithm 1: Safe and Stable Scenario-based Control Design

Initialization:
Start from a safe initial condition $x_0 \in \mathcal{C}$
Initiate the controller with a stabilizing feedback gain k_0.
Initial approximation of uncertain parameters $\hat{\theta}_0$
Select ϵ and β based on required confidence level

procedure Outer Control Loop
    Form the stack variable until (26) is satisfied. Then update weights using experience replay update law (23).
    Find the uncertain set boundaries using (30).
end procedure

procedure Inner Control Loop
    Find the nominal controller.
    Form the Lyapunov function and CBF function inequality constraints in (10).
    Sample constraints in (9) based on updated uncertain set from the outer loop and desired specifications in the initiation step.
    Solve the scenario optimization (31) and apply the safe and stabilizing controller to the system.
    Update the outer loop and repeat until control objectives are met.
end procedure

Lemma 2. *Consider the scenario optimization* (31)*, and the sampling set* Θ_c *updated by the experience replay model learning as* (29) *based on Algorithm 1; then, with a fixed number of samples, the probability of satisfying safety constraints sequentially increases.*

Proof: Consider U, U_ϵ as the feasible region of u and the set of inputs that constraints are satisfied with the probability of at least $1-\epsilon$, respectively. We define U_θ as the set of u for the realizations of θ according to the updated set of constraints Θ as (29) and U^* as the set of inputs that satisfy constraints by the probability of one, which corresponds to the exact value of uncertain parameter θ^*. Since $\theta^* \in \Theta$, one has $U^* \subset U_\theta$.

In addition $U^* \subset U_\epsilon \subset U$. From Corollary 1, U_θ exponentially fast converges from U to U^*, as also shown in Fig. 1. Therefore, from (7), the number of required samples reduces; or for a fixed number of samples, from Theorem 2, the probability of satisfying constraints increases. This completes the proof.

Theorem 4. *Consider a violation parameter* $\epsilon \in (0,1)$ *and a confidence parameter* $\beta \in (0,1)$*. Let the number of iid samples satisfy the following condition*

$$N > \frac{2}{\epsilon}(\ln\frac{1}{\beta} + m) \quad (33)$$

where m is the dimension of the control input. Then, with a probability of no smaller than $1-\beta$*, the solution to SQP problem* (31) *ensures the probabilistic safety of the system, i.e.,* $Pr(\delta : f_\theta(u_N{}^*) > 0) \leqslant \epsilon$*. In addition, fixing the number of samples from the conditional ambiguity set* Θ_c*, the violation probability sequentially decreases, thus providing a higher probability of safety guarantee.*

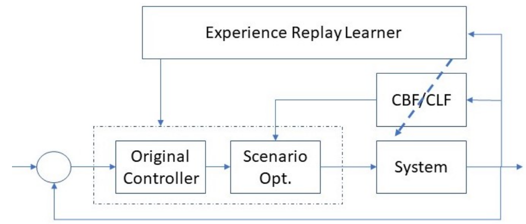

Fig. 2. Control Framework

Proof: Since $U_\theta \in U$, then based on Theorem 2, with the probability of no smaller than $1-\beta$ and at most an ϵ-fraction, the CBF constraints are satisfied. From [10], the safety criteria are satisfied with the same probability. From the result of Lemma 2, this probability increases at each iteration. This completes the proof.

V. Simulation

Consider the following uncertain system,

$$\dot{x}_1 = \theta_1 x_2 + u$$
$$\dot{x}_2 = \theta_2 x_1 + 2x_2 + 1.5u$$

where $\theta_1, \theta_2 \in \Delta$ are the uncertain parameters and $\Delta = [-1.8, 1.8]$ is the corresponding uncertainty set. Without loss of generality, we use a simple case to demonstrate the concept. Assume the system needs to respect the following state constraint as the safety criterion $x_1 \leqslant 1$. To guarantee this criterion, the CBF candidate $h(x) = 1 - x_1$ and the corresponding safety set $\mathcal{C} = \{x | h(x) \geqslant 0\}$ are considered. The forward invariance of $\mathcal{C}$ is guaranteed if the invariance criterion is satisfied, $\dot{h} + \alpha h \geqslant 0$. In other words,

$$-\theta_1 x_2 - u + \alpha(1 - x_1) \geqslant 0 \quad (34)$$

However, this criterion incorporates the uncertain parameter θ_1, and thus, it is not fully known.

To tackle this problem, we 1) use the experience replay model-learning to find an accurate estimate of the uncertainty set Δ and learn about the uncertain parameter θ_1 to shrink the uncertainty set. 2) sample the available uncertainty set and form the CBF criterion for samples at each iteration to solve a well-defined scenario optimization and ensure the safety of the system with high probability. The error converges to zero exponentially fast by using the experience replay model learning; thus the uncertainty set is updated at each time instance t as $\Delta_1 = \{||\theta_1|| \leqslant a_1 e^{-0.15t}\}$. This set is uniformly sampled, and the corresponding safety criterion is incorporated in SQP problem (31).

To depict the performance of the proposed method, three different simulations are conducted. First, it is assumed that the full system dynamic (34) is available. The red line in Figure 3(a) depicts the safety border in which x_1 must stay below it. As it is shown in this figure, CBF has certified the safety of the system. However, in the presence of uncertainty, it is not applicable. Second, the scenario approach for safety guarantee for a fixed uncertainty set is employed. The safety set is sampled and corresponding CBFs are considered. In

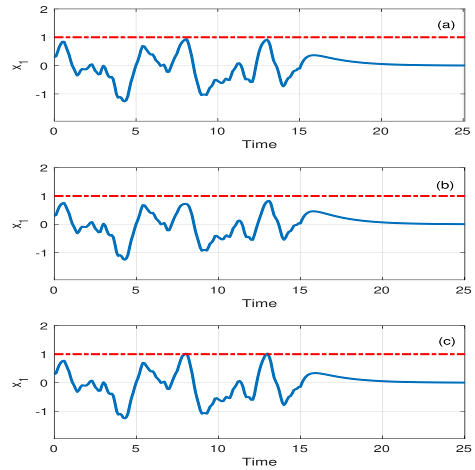

Fig. 3. The state trajectory for x_1 with (a) Known CBF, (b) Scenario approach, (c) Experience replay scenario approach

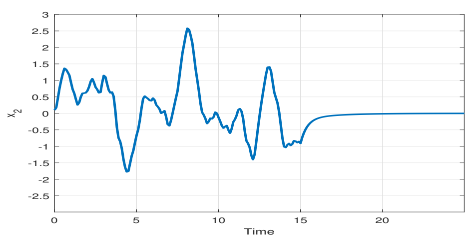

Fig. 4. The state trajectory for x_2

this simulation scenario, the effect of experience replay in shrinking the uncertainty set is not considered. As can be seen in Figure 3(b), the safety of the system is preserved despite the uncertainty. Note that this method is superior to the methods in which CBF should be satisfied for all possible realizations of uncertain parameters. However, due to a large set of uncertainties, the behavior of the system is relatively conservative in the sense that it does not get as close as possible to the safety boundary. Finally, the scenario approach, in conjunction with experience replay model learning and by sampling the conditional shrinking ambiguity set, is employed. The result is depicted in Figure 3(c), which shows that the safety of the system is preserved in a non-conservative manner. Note that the experience-replay model learning provides an accurate bound for the uncertainty which rapidly converges to the true value and thus provide a better response. Moreover, the response of the system is very similar to the case in which CBF is fully known. The other state of the system and the weight error are depicted in Figures 4 and 5, respectively.

VI. Conclusion

In this paper, a probabilistic framework for joint model learning and safe control of linear systems with parametric uncertainty is proposed. The scenario approach is employed to provide a tractable optimization by the proper sampling of safety constraints, reducing a semi-infinite optimization problem into a convex optimization problem with the finite number of constraints. The experience replay model learning is leveraged as the outer control loop, which learns the uncertainty with convergence guarantee and enables stability

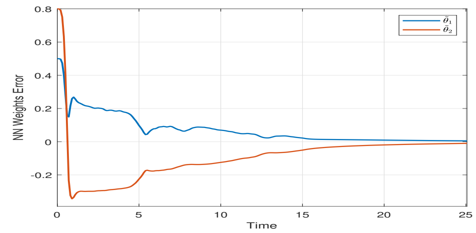

Fig. 5. Weight error for system modeling

guarantee using the perturbation theory. In addition, the bound of the uncertain set is exponentially reduced, thus more accuracy is achieved through proactive sampling.

VII. Acknowledgement

The authors would like to thank Prof. Andrew Alleyne for his gracious support.

References

[1] B. Alrifaee, K. Kostyszyn, and D. Abel, "Model predictive control for collision avoidance of networked vehicles using lagrangian relaxation," *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 430 – 435, 2016. 14th IFAC Symposium on Control in Transportation Systems CTS 2016.

[2] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2019.

[3] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, pp. 3861–3876, Aug 2017.

[4] A. Agrawal and K. Sreenath, "Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation," in *Proceedings of Robotics: Science and Systems*, (Cambridge, Massachusetts), July 2017.

[5] L. Wang, E. A. Theodorou, and M. Egerstedt, "Safe learning of quadrotor dynamics using barrier certificates," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2460–2465, May 2018.

[6] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames, "Robustness of control barrier functions for safety critical control," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 54 – 61, 2015.

[7] A. J. Taylor and A. D. Ames, "Adaptive safety with control barrier functions," in *2020 American Control Conference (ACC)*, pp. 1399–1405, 2020.

[8] B. T. Lopez, J. J. E. Slotine, and J. P. How, "Robust adaptive control barrier functions: An adaptive and data-driven approach to safety," *IEEE Control Systems Letters*, vol. 5, no. 3, pp. 1031–1036, 2021.

[9] W. Luo, W. Sun, and A. Kapoor, "Multi-robot collision avoidance under uncertainty with probabilistic safety barrier certificates," 2020.

[10] C. Wang, M. Bahreinian, and R. Tron, "Chance constraint robust control with control barrier functions," in *2021 American Control Conference (ACC)*, pp. 2315–2322, 2021.

[11] Z. Marvi and B. Kiumarsi, "Barrier-certified model-learning and control of uncertain linear systems using experience replay method," in *2021 Conference on Decision and Control (CDC)*, 2021.

[12] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, vol. 33, no. 2, pp. 149–157, 2009.

[13] C. Chen, H. Modares, K. Xie, F. L. Lewis, Y. Wan, and S. Xie, "Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics," *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4423–4438, 2019.

[14] H. Khalil, *Nonlinear Systems*. Pearson Education, Prentice Hall, 2002.

[15] G. Calafiore and M. Campi, "The scenario approach to robust control design," *IEEE Transactions on Automatic Control*, vol. 51, no. 5, pp. 742–753, 2006.

[16] M. C. Campi and S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1211–1230, 2008.

[17] J. Luedtke and S. Ahmed, "A sample approximation approach for optimization with probabilistic constraints," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 674–699, 2008.