2021 American Control Conference (ACC)
New Orleans, USA, May 25-28, 2021

Online Learning of Parameterized Uncertain Dynamical Environments with Finite-sample Guarantees

Dan Li[1], Dariush Fooladivanda[2] and Sonia Martínez[1]

***Abstract*— We present a novel online learning algorithm for a class of unknown and uncertain dynamical environments that are fully observable. First, we obtain a novel probabilistic characterization of systems whose mean behavior is known but which are subject to additive, unknown subGaussian disturbances. This characterization relies on recent concentration of measure results and is given in terms of ambiguity sets. Second, we extend the results to environments whose mean behavior is also unknown but described by a parameterized class of possible mean behaviors. Our algorithm adapts the ambiguity set dynamically by learning the parametric dependence online, and retaining similar probabilistic guarantees with respect to the additive, unknown disturbance. We illustrate the results on a differential-drive robot subject to environmental uncertainty.**

I. Introduction

The online learning of uncertain dynamical systems has broad applications in various domains, including those of artificial intelligence and robotics [1]–[3]. Fundamentally, one is to exploit input-output data to identify the representation of the environment that best captures its behavior. In this way, several techniques, from first-principles system identification to, more recently, (deep) neural networks, have been successfully used in various domains. Unfortunately, safe performance usually depends upon the assimilation of vast amounts of data, which is mostly done offline and prevents its application in real-time scenarios. Motivated by this, we investigate the integration of recently-developed probabilistically-guaranteed system descriptions with online, predictor-based learning algorithms.

The system identification literature broadly encompasses linear [4], [5] and non-linear systems [6], [7], with asymptotic performance guarantees. More recently, finite-sample analysis of identification methods has been proposed for linear systems [8]–[11]. These methods leverage modern measure-of-concentration results [12], [13] for non-asymptotic guarantees of the identification error bounds. Measure-of-concentration results are also used in [**?**], [14]. However, the goal of [**?**], [14] is to learn an unknown initial distribution evolving under a known dynamical system while assimilating data via a linear observer. This characterization is given in terms of *ambiguity sets*, which are constructed via multiple system trajectories or realizations. In contrast, here we employ Wasserstein metrics to develop an online learning algorithm for uncertain dynamical systems with similar-in-spirit probabilistic guarantees.

Statement of Contributions: We propose an online learning algorithm that characterizes a class of unknown and uncertain dynamical environments with probabilistic guarantees using a finite amount of data. To achieve this, we first assume that the mean behavior of the stochastic system is known but the system states are subject to an additive, unknown sub-Gaussian distribution, characterized by a set of distributions or *ambiguity set*. Then, we extend the results to environments whose mean behavior is unknown but belongs to a parameterized class of behaviors. In this regard, we propose a time-varying parameterized ambiguity set and a learning methodology to capture the behavior of the environment. We show how the proposed online learning algorithm retains desirable probabilistic guarantees with high confidence. A differential-drive robot subject to environmental uncertainty is provided for an illustration. Basic notations and definitions can be found in the footnote. [1]

II. Problem Statement

This section presents the description of the uncertain dynamical environment which we aim to learn, with a problem

*This research was developed with funding from ONR N00014-19-1-2471, and AFOSR FA9550-19-1-0235.

[1] D. Li and S. Martínez are with the Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA 92092, USA. lidan@ucsd.edu; soniamd@ucsd.edu

[2] D. Fooladivanda is with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720, USA. dfooladi@berkeley.edu;

[1] Let $\mathbb{R}^m$, $\mathbb{R}^m_{\geq 0}$, $\mathbb{Z}^m_{\geq 0}$ and $\mathbb{R}^{m\times n}$ denote respectively the m-dimensional real space, the m-dimensional nonnegative real space, the m-dimensional nonnegative integer space, and the space of $m\times n$ matrices. By $\boldsymbol{x}\in\mathbb{R}^m$ we denote a column vector of dimension m, while $\boldsymbol{x}^\top$ represents its transpose. The shorthand notation $\mathbf{1}_m$ denotes the column vector $(1,\cdots,1)^\top\in\mathbb{R}^m$. We use subscripts to index vectors, i.e., $\boldsymbol{x}_k\in\mathbb{R}^m$ for $k\in\mathbb{Z}_{\geq 0}$, and we use x_i to denote the i^{th} component of $\boldsymbol{x}$. We denote respectively the 2-norm and ∞-norm by $\|\boldsymbol{x}\|$ and $\|\boldsymbol{x}\|_\infty$. We define the m-dimensional norm ball with center $\boldsymbol{x}\in\mathbb{R}^m$ and radius $\epsilon\in\mathbb{R}_{\geq 0}$ as the set $B_\epsilon(\boldsymbol{x}):=\{\boldsymbol{y}\in\mathbb{R}^m \mid \|\boldsymbol{y}-\boldsymbol{x}\|\leq\epsilon\}$. We denote by $\langle\cdot,\cdot\rangle$ an inner product in the space of interest. Consider the space $\mathbb{R}^m$, we define $\langle\boldsymbol{x},\boldsymbol{y}\rangle:=\boldsymbol{x}^\top\boldsymbol{y}$, $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^m$. In particular, $\|\boldsymbol{x}\|:=\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$. Consider Finsler manifold $\mathbb{R}^2\times[-\pi,\pi)\cong\mathbb{R}\times\mathbb{S}^1$ where $\mathbb{S}^1$ stands for the unit circle. For $(\boldsymbol{x},\theta_1),(\boldsymbol{y},\theta_2)\in\mathbb{R}^2\times[-\pi,\pi)$, we define $\langle(\boldsymbol{x},\theta_1),(\boldsymbol{y},\theta_2)\rangle:=\boldsymbol{x}^\top\boldsymbol{y}+\cos(\min\{|\theta_1-\theta_2|,\ 2\pi-|\theta_1-\theta_2|\})$. In particular, we use $\|(\boldsymbol{x},\theta)\|:=\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle+1}$. Given an $A\in\mathbb{R}^{m\times m}$, we write its Singular Value Decomposition (SVD) as $A=U\Sigma V^\top$, where U, $V\in\mathbb{R}^m$ are orthonormal and Σ is diagonal with non-negative entries. These entries are called singular values of A, and we denoted by $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ the maximal and non-zero minimal singular value of A, respectively. We denote by $A^\dagger:=V\Sigma^\dagger U^\top$ the Moore–Penrose inverse of A, where $\Sigma^\dagger$ is the same as Σ except the replacement of each positive entry by its inverse. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, with Ω the sample space, $\mathcal{F}$ a σ-algebra, and $\mathbb{P}$ the associated probability distribution. Let $\boldsymbol{x}:\Omega\to\mathbb{R}^m$ be an induced random vector. We denote by $\mathcal{M}$ the space of all probability distributions with finite first moment. To measure the distance in $\mathcal{M}$, we use the dual version of the 1-Wasserstein metric $d_W:\mathcal{M}\times\mathcal{M}\to\mathbb{R}_{\geq 0}$, defined as in [15]. A closed Wasserstein ball of radius ϵ centered at a distribution $\mathbb{P}\in\mathcal{M}$ is denoted by $\mathbb{B}_\epsilon(\mathbb{P}):=\{\mathbb{Q}\in\mathcal{M}\mid d_W(\mathbb{P},\mathbb{Q})\leq\epsilon\}$. We denote the Dirac measure at $x_0\in\Omega$ as $\delta_{\{x_0\}}:\Omega\to\{0,1\}$. For any set $A\in\mathcal{F}$, we let $\delta_{\{x_0\}}(A)=1$, if $x_0\in A$, otherwise 0. For an $\boldsymbol{x}\in\Omega$, we denote $\mathbb{P}\equiv\mathbb{Q}+\boldsymbol{x}$, if $\mathbb{P}$ is a translation of $\mathbb{Q}$ by $\boldsymbol{x}$.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.

definition. Let $t \in \mathbb{Z}_{\geq 0}$ denote time discretization. For each t, the uncertain system is characterized by a random variable $\boldsymbol{x} \in \mathbb{R}^n$ which evolves according to an *unknown*, discrete-time, stochastic and, potentially, time-varying system

$$\boldsymbol{x}_{t+1} = f(t, \boldsymbol{x}_t, \boldsymbol{d}_t) + \boldsymbol{w}_t, \quad \text{with some } \boldsymbol{x}_0 \sim \mathbb{P}_0. \tag{1}$$

The distribution $\mathbb{P}_{t+1}$ characterizing $\boldsymbol{x}_{t+1}$ is determined by the current state's distribution, the unknown mapping $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, and random vectors $\boldsymbol{w}_t$ that cannot be captured by f. We further assume that $\boldsymbol{d}_t$ is an exogenous signal that is selected in advance or revealed online, which can play the role of an external reference or control. Let us denote by $\mathbb{W}_t$ the distribution of the random vector $\boldsymbol{w}_t \in \mathbb{R}^n$.

Assumption II.1 (Independent and stationary subGaussian distributions) Consider random vectors $\boldsymbol{w}_t \in \mathbb{R}^n$, $t \in \mathbb{Z}_{\geq 0}$. It is assumed that: **(1)** The random vectors $\boldsymbol{w}_t$ are component-wise and time-wise independent, i.e., $w_{t,i}$ and $w_{k,j}$ are independent, for all $t \neq k$, $i \neq j$, $(t,k) \in \mathbb{Z}^2_{\geq 0}$ and $(i,j) \in \{1,\dots,n\}$. **(2)** For each t, $\boldsymbol{w}_t$ are i.d. as zero-mean σ-subGaussian, i.e., for any $a \in \mathbb{R}^n$ we have $\mathbb{E}\left[\exp(a^\top \boldsymbol{w}_t)\right] \leq \exp(\|a\|^2\sigma^2/2)$.

Example II.1 ($\boldsymbol{\sigma}$-subGaussian distributions) A trivial example is any $\mathbb{W} \equiv \mathcal{N}(\mathbf{0}, \Sigma)$ with $\sigma_{\max}(\Sigma) \leq \sigma^2$. As any random vector supported on a compact set belongs to the subGaussian class, in particular, the following are σ-subGaussian distributions: **(1)** any zero-mean uniform distribution $\boldsymbol{w} \sim \mathcal{U}(\Omega)$ supported over $\Omega \subset B_\sigma(\mathbf{0})$; **(2)** any zero-mean discrete distribution with support $\Omega \subset B_\sigma(\mathbf{0})$.

This paper aims to obtain a tractable characterization of the unknown distribution $\mathbb{P}_{t+1}$ of the immediate-future environment state $\boldsymbol{x}_{t+1}$ online, $\forall\, t$. This is to be done by employing historical measurements, $\hat{\boldsymbol{x}}_k$, $k \leq t$, and data $\hat{\boldsymbol{d}}_k$, $k \leq t$.

III. Characterization of Random Dynamical Environments under Perfect Information

We aim to provide a description of the random dynamical system (1) via ambiguity sets. More precisely, given knowledge $\boldsymbol{d}$, and system data $\hat{\boldsymbol{x}}$, we look for a set of distributions $\mathcal{P}_{t+1} := \mathcal{P}_{t+1}(\boldsymbol{d}, \hat{\boldsymbol{x}})$ characterizing $\mathbb{P}_{t+1}$ via

$$\text{Prob}\left(\mathbb{P}_{t+1} \in \mathcal{P}_{t+1}\right) \geq 1 - \beta, \tag{2}$$

for some $\beta \in (0,1)$. Observe that the probability is taken with respect to the historical random data outcomes. To do this, let $T_0 \in \mathbb{Z}_{>0}$ and $T := \min\{t, T_0\} \geq 1$, and consider the historical data, $\hat{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{d}}_k$, for $k \in \mathcal{T} := \{t-T, \dots, t-1\}$. Assuming a perfect knowledge of f, we show first how to use the data set $\mathcal{I} := \{\hat{\boldsymbol{x}}_t,\ \hat{\boldsymbol{x}}_k,\ \hat{\boldsymbol{d}}_k,\ k \in \mathcal{T}\}$ to construct $\mathcal{P}_{t+1}$, $\forall\, t \geq 0$.

Let us denote by $\mathbb{Q}_{t+1} \equiv \mathbb{Q}_{t+1}(\boldsymbol{d})$ the empirical distribution of $\boldsymbol{x}_{t+1}$ and define it as follows

$$\mathbb{Q}_{t+1} := \frac{1}{T} \sum_{k \in \mathcal{T}} \delta_{\{\xi_k(\boldsymbol{d})\}},$$

where $\xi_k(\boldsymbol{d}) := f(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d}) + \hat{\boldsymbol{x}}_{k+1} - f(k, \hat{\boldsymbol{x}}_k, \hat{\boldsymbol{d}}_k),\ \forall k \in \mathcal{T}$. The following result enables us to construct the ambiguity set $\mathcal{P}_{t+1}$ that satisfies (2).

Lemma III.1 (Asymptotic dynamic ambiguity set) *Let us assume that the system f is known at each time t. Given a confidence level $\beta \in (0,1)$, parameter $T_0 \in \mathbb{Z}_{>0}$, and horizon $T = \min\{t, T_0\}$, there exists a positive scalar $\epsilon := \epsilon(T, \beta)$ such that* (2) *holds by selecting*

$$\mathcal{P}_{t+1} := \mathbb{B}_\epsilon(\mathbb{Q}_{t+1}) = \{\mathbb{P} \mid d_W(\mathbb{P}, \mathbb{Q}_{t+1}) \leq \epsilon\},$$

a Wasserstein ball centered at $\mathbb{Q}_{t+1}$ with radius

$$\epsilon := \sqrt{\frac{2n\sigma^2}{T} \ln(\frac{1}{\beta})} + \mathcal{O}(T^{-1/\max\{n,2\}}),$$

where n is the dimension of $\boldsymbol{x}$ and σ is as in Assumption II.1. Further, if $T_0 = \infty$, then as $t \to \infty$, $\epsilon \to 0$, i.e., the set $\mathcal{P}_{t+1}$ shrinks to the singleton $\mathbb{P}_{t+1}$ at a rate $\mathcal{O}(1/T^{-1/\max\{n,2\}})$.

Proof. We first leverage two properties: **(1)** Following [16, Theorem 6], the random variable $z_t := d_W(\hat{\mathbb{W}}_t, \mathbb{W}_t)$ is $\sqrt{n}\sigma/\sqrt{T}$-subGaussian for all t, where $\hat{\mathbb{W}}_t \equiv \mathbb{Q}_{t+1} - f(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d})$, i.e., we have $\mathbb{E}\left[\exp\left(\lambda z_t - \lambda \mathbb{E}[z_t]\right)\right] \leq \exp(n\lambda^2\sigma^2/(2T))$ for all t and any $\lambda \in \mathbb{R}$. **(2)** Following [12, Theorem 1] and [17, Theorem 3.1], we have $C_t := \mathbb{E}[z_t] \leq \mathcal{O}(T^{-1/\max\{n,2\}})$. Second, using the dynamics (1) and the Markov inequality, for any $\epsilon \geq 0$ and $\lambda \geq 0$, we have

$$\text{Prob}\left(d_W(\mathbb{Q}_{t+1}, \mathbb{P}_{t+1}) \geq \epsilon\right) \leq \exp\left(-\epsilon\lambda\right) \mathbb{E}\left[\exp\left(\lambda z_t\right)\right].$$

Using the mentioned two properties, and selecting $\lambda = (\epsilon - C_t)T/(n\sigma^2)$ with $\epsilon > C_t$, we achieve

$$\text{Prob}\left(d_W(\mathbb{Q}_{t+1}, \mathbb{P}_{t+1}) \geq \epsilon\right) \leq \exp\left(-\frac{(\epsilon - C_t)^2 T}{2n\sigma^2}\right).$$

Taking ϵ as that in the lemma, we have

$$\text{Prob}\left(d_W(\mathbb{Q}_{t+1}, \mathbb{P}_{t+1}) \geq \epsilon\right) \leq \beta,$$

resulting in (2). Taking $T_0 = \infty$, it follows that $\mathcal{P}_{t+1}$ shrinks to $\mathbb{P}_{t+1}$ as $t \to \infty$. See [18] for more details. ■.

In practice, T_0, and β need to be selected empirically, in order to efficiently address the particular problem that leverages the characterization of (1).

IV. Characterization of Random Dynamical Environments in a Parameterized Family

The construction of the empirical distribution $\mathbb{Q}_{t+1}$ of the previous section relies on the knowledge of f. When f is unknown, one may represent f as belonging to a parameterized class of functions. Such as the approach adopted in the neural networks field and Koopman operator theory. Here, we focus on the case that f is approximated by a linear combination of a class of functions or "predictors" as follows.

Assumption IV.1 (Environment predictor class) There exists a set of predictors $f^{(i)} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $(t, \boldsymbol{x}, \boldsymbol{d}) \mapsto f^{(i)}(t, \boldsymbol{x}, \boldsymbol{d})$, $i \in \{1, \dots, p\}$, such that: (1) The vector fields $f^{(1)}, f^{(2)}, \dots, f^{(p)}$ are linearly independent almost everywhere. (2) There exists potentially time-varying coefficients $\boldsymbol{\alpha}^\star := (\alpha_1^\star, \dots, \alpha_p^\star) \in \mathbb{R}^p$ such that

$$f(t, \boldsymbol{x}, \boldsymbol{d}) = \sum_{i=1}^{p} \alpha_i^\star f^{(i)}(t, \boldsymbol{x}, \boldsymbol{d}).$$

As the selection of the predictors is not the subject of this study, we assume that the predictors are found in advance, and hence they are known to the learning algorithm.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.

The construction of an effective ambiguity set now depends on learning the dynamical environment mapping. Let us denote by $\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}_t$ the estimated value of the parameter $\boldsymbol{\alpha}^\star$ at time t. To construct $\mathcal{P}_{t+1}$, consider T predictions of $\boldsymbol{x}_{t+1}$ using $f^{(i)}$, denoted by $\xi_k^{(i)}(\boldsymbol{\alpha}, \boldsymbol{d})$. For each $k \in \mathcal{T}$, $i \in \{1, \cdots, p\}$, and given $\boldsymbol{d} := \boldsymbol{d}_t$, we define

$$\xi_k^{(i)}(\boldsymbol{\alpha}, \boldsymbol{d}) := f^{(i)}(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d}) + \frac{\hat{\boldsymbol{x}}_{k+1}}{\boldsymbol{\alpha}^\top \mathbf{1}_p} - f^{(i)}(k, \hat{\boldsymbol{x}}_k, \hat{\boldsymbol{d}}_k).$$

Now, we select the empirical $\hat{\mathbb{P}}_{t+1} \equiv \hat{\mathbb{P}}_{t+1}(\boldsymbol{\alpha}, \boldsymbol{d})$, as follows:

$$\hat{\mathbb{P}}_{t+1} := \frac{1}{T} \sum_{k \in \mathcal{T}} \delta_{\{\sum_{i=1}^{p} \alpha_i \xi_k^{(i)}(\boldsymbol{\alpha}, \boldsymbol{d})\}}. \tag{3}$$

The following result enables the construction of the ambiguity set $\mathcal{P}_{t+1}$, relying on both $\boldsymbol{d}$ and $\boldsymbol{\alpha}$, which satisfies (2).

Theorem IV.1 (Adaptive dynamic ambiguity set) *Assume that the data set $\mathcal{I}$ is accessible, $\forall\, t$. Further, let Assumption IV.1, on the environment predictor class, hold for some $\boldsymbol{\alpha}^\star$ at time $t \in \mathcal{T}$. Then, given a confidence level $\beta \in (0, 1)$, horizon parameter T_0, and a learning parameter $\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}_t \in \mathbb{R}^p$, there exists a scalar $\hat{\epsilon} := \hat{\epsilon}(t, T, \beta, \boldsymbol{\alpha}, \boldsymbol{d})$ such that* (2) *holds by selecting*

$$\mathcal{P}_{t+1} := \mathbb{B}_{\hat{\epsilon}}(\hat{\mathbb{P}}_{t+1}) = \{\mathbb{P} \mid d_W(\mathbb{P}, \hat{\mathbb{P}}_{t+1}) \leq \hat{\epsilon}\},$$

where $\hat{\epsilon} = \epsilon + \|\boldsymbol{\alpha}^\star - \boldsymbol{\alpha}\|_\infty H(t, T, \boldsymbol{d})$, with

$$H(t, T, \boldsymbol{d}) := \frac{1}{T} \sum_{i=1}^{p} \sum_{k \in \mathcal{T}} \|f^{(i)}(k, \hat{\boldsymbol{x}}_k, \hat{\boldsymbol{d}}_k) - f^{(i)}(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d})\|,$$

and the radius ϵ is selected as in Lemma III.1.

Proof. We first leverage the triangular inequality,

$$d_{\mathrm{W}}(\mathbb{P}_{t+1}, \hat{\mathbb{P}}_{t+1}) \leq d_{\mathrm{W}}(\mathbb{P}_{t+1}, \mathbb{Q}_{t+1}) + d_{\mathrm{W}}(\mathbb{Q}_{t+1}, \hat{\mathbb{P}}_{t+1}).$$

Then, using Lemma III.1 on the first term and the definition of the Wasserstein metric on the second term, we achieve the guarantee (2). A detailed proof is provided in [18]. ■.

Theorem IV.1 indicates that, if we select $\boldsymbol{\alpha}$ wisely, i.e., $\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}^\star$, then the adaptive dynamic ambiguity set is identical to that of Lemma III.1.

To estimate an unknown $\boldsymbol{\alpha}^\star$ while preserving the probabilistic guarantees, we propose an online learning algorithm that attempts to bring $\boldsymbol{\alpha}$ close to $\boldsymbol{\alpha}^\star$ with high probability. Intuitively, our approach is based on the comparison of new obtained data with updates given by a predictor combination.

Theorem IV.2 (Learning of $\boldsymbol{\alpha}^\star$) *Let the data set $\mathcal{I}$ and predictors $\{f^{(i)}\}_i$ be given. For each $k \in \mathcal{T}$ and $i \in \{1, \ldots, p\}$, let us denote $f_k^{(i)} := f^{(i)}(k, \hat{\boldsymbol{x}}_k, \hat{\boldsymbol{d}}_k)$. Consider the data matrix $A \equiv A_t \in \mathbb{R}^{p \times p}$ with*

$$A(i, j) := \frac{1}{T} \sum_{k \in \mathcal{T}} \langle f_k^{(j)}, P_k f_k^{(i)} \rangle, \;\; i,\, j \in \{1, \ldots, p\},$$

where P_k is an online regularization matrix at time k, and let us consider the data vector $\boldsymbol{b} \equiv \boldsymbol{b}_t \in \mathbb{R}^p$, with components

$$\boldsymbol{b}(i) := \frac{1}{T} \sum_{k \in \mathcal{T}} \langle \hat{\boldsymbol{x}}_{k+1}, P_k f_k^{(i)} \rangle, \;\; i \in \{1, \ldots, p\}.$$

Given $\eta > 0$, we select P_k such that $\|P_k f_k^{(i)}\| \leq \eta$ for all $i \in \{1, \ldots, p\}$, $k \in \mathcal{T}$, and select $\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}_t$ to be

$$\boldsymbol{\alpha} = A^\dagger \boldsymbol{b}, \tag{4}$$

where $A^\dagger$ denotes the Moore–Penrose inverse of A. Let Assumption II.1 and Assumption IV.1 hold, and take

$$c := \sigma e \eta \sqrt{np} \sigma_{\min}^{-1}(A),$$

where σ is that in Assumption II.1, the constant $e \approx 2.718$, and $\sigma_{\min}(A)$ is the minimal non-zero principal singular value of A. Then by selecting $\gamma \geq nc$, the parameter $\boldsymbol{\alpha}$ is ensured to be close to $\boldsymbol{\alpha}^\star$ with high probability in the following sense:

$$\mathrm{Prob}\left(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty \leq \gamma\right) \geq 1 - \exp\left(-\frac{(nc - \gamma)^2 T^2}{2\left[(2T - 1)c\gamma + nc^2\right]}\right).$$

In particular, selecting $\gamma \geq nc/e$, we obtain a non-trivial bound with a slow confidence growth rate as follows

$$\mathrm{Prob}\left(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty \leq \gamma\right) \geq 1 - \frac{1}{\gamma} n \sigma \eta \sqrt{np} \sigma_{\min}^{-1}(A).$$

Proof. Step 1 (Bound on $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty$): At each $k \in \mathcal{T}$, let us denote by $\hat{\boldsymbol{w}}_k$ a sample of $\boldsymbol{w}_k$ represented by $\hat{\boldsymbol{w}}_k := \hat{\boldsymbol{x}}_{k+1} - \sum_{j=1}^{p} \alpha_j^\star f_k^{(j)}$. Then, we project the data on the direction of each regularized predictor $i \in \{1, \ldots, p\}$, resulting in

$$\langle \hat{\boldsymbol{x}}_{k+1}, P_k f_k^{(i)} \rangle = \langle \sum_{j=1}^{p} \alpha_j^\star f_k^{(j)} + \hat{\boldsymbol{w}}_k, P_k f_k^{(i)} \rangle,$$

where, given a scalar $\eta > 0$, the time-dependent regularization matrix is selected so that $\|P_k f_k^{(i)}\| \leq \eta$, for all $i \in \{1, \ldots, p\}$. Averaging the above equalities over $k \in \mathcal{T}$, we have for each component i the following

$$\boldsymbol{b}(i) = \sum_{j=1}^{p} \alpha_j^\star A(i, j) + \frac{1}{T} \sum_{k \in \mathcal{T}} \langle \hat{\boldsymbol{w}}_k, P_k f_k^{(i)} \rangle,$$

where $\boldsymbol{b}(i)$ and $A(i, j)$ are those in the theorem. Note that $\boldsymbol{\alpha}$ is selected as in (4), which results in $\boldsymbol{b}(i) = \sum_{j=1}^{p} \alpha_j A(i, j)$ for each i. By subtracting the above two equations, we have

$$\sum_{j=1}^{p} (\alpha_j - \alpha^\star) A(i, j) = \frac{1}{T} \sum_{k \in \mathcal{T}} \langle \hat{\boldsymbol{w}}_k, P_k f_k^{(i)} \rangle.$$

Taking the Moore–Penrose inverse of A, we obtain $\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star = A^\dagger \boldsymbol{c}$, where the vector $\boldsymbol{c}$ is

$$\frac{1}{T} \sum_{k \in \mathcal{T}} \left(\langle \hat{\boldsymbol{w}}_k, P_k f_k^{(1)} \rangle, \ldots, \langle \hat{\boldsymbol{w}}_k, P_k f_k^{(p)} \rangle \right)^\top.$$

Take the ∞-norm, we have $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty \leq \|A^\dagger\|_\infty \|\boldsymbol{c}\|_\infty$ and we bound $\|\boldsymbol{c}\|_\infty$ by

$$\begin{aligned} \|\boldsymbol{c}\|_\infty &\leq \frac{1}{T} \max_{i \in \{1, \ldots, p\}} \left\{ \sum_{k \in \mathcal{T}} |\langle \hat{\boldsymbol{w}}_k, P_k f_k^{(i)} \rangle| \right\}, \\ &\leq \frac{1}{T} \max_{i \in \{1, \ldots, p\}} \left\{ \sum_{k \in \mathcal{T}} \left(\|\hat{\boldsymbol{w}}_k\| \cdot \|P_k f_k^{(i)}\| \right) \right\}, \\ &\leq \frac{\eta \sqrt{n}}{T} \sum_{k \in \mathcal{T}} \|\hat{\boldsymbol{w}}_k\|_\infty, \end{aligned}$$

where we achieve the first inequality by moving the absolute operation into the sum operation; the second inequality uses

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.

Hölder's inequality; the third inequality is achieved by the norm equivalence and the fact that $\|P_k f_k^{(i)}\| \leq \eta$ for all $i \in \{1, \ldots, p\}$. Then, we achieve the following bound

$$\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty \leq \eta\sqrt{n}\|A^\dagger\|_\infty \left[\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\hat{\boldsymbol{w}}_k\|_\infty)\right]. \quad (5)$$

Note that, by the equivalence of the matrix norm, we have

$$\|A^\dagger\|_\infty \leq \sqrt{p}\|A^\dagger\|_2 = \sqrt{p}\,\sigma_{\max}(A^\dagger) \leq \sqrt{p}\,\sigma_{\min}^{-1}(A),$$

where $\sigma_{\max}(A^\dagger)$ and $\sigma_{\min}(A)$ denote the maximal singular value of $A^\dagger$ and the minimal principal non-zero singular value of A, respectively.

Step 2 (Measure concentration of $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty$): We achieve this by studying the measure concentration of $\|\boldsymbol{w}_k\|_\infty$. Equivalently, given any $\gamma > 0$, we compute

$$\text{Prob}\left(\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\boldsymbol{w}_k\|_\infty) \geq \gamma\right). \quad (6)$$

There are two options to obtain the bound.

(1) (A bound with exponential decay over T): For any $\lambda \geq 0$, the probability (6) is equivalent to

$$\text{Prob}\left(\exp\left(\sum_{k\in\mathcal{T}}\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right) \geq \exp(\gamma\lambda)\right).$$

By the Markov inequality to the above probability, we have

$$\text{Prob}\left(\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\boldsymbol{w}_k\|_\infty) \geq \gamma\right) \leq \exp(-\gamma\lambda)\,\mathbb{E}\left[\prod_{k\in\mathcal{T}}\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right].$$

By Assumption II.1 on independence of $\boldsymbol{w}_k$, we have

$$\mathbb{E}\left[\prod_{k\in\mathcal{T}}\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right] = \prod_{k\in\mathcal{T}}\mathbb{E}\left[\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right].$$

For each $k \in \mathcal{T}$, we write each exp operation in its power series form as the following

$$\mathbb{E}\left[\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right] = \mathbb{E}\left[1 + \sum_{l=1}^{\infty}\frac{\left(\frac{\lambda}{T}\right)^l\|\boldsymbol{w}_k\|_\infty^l}{l!}\right], = 1 + \sum_{l=1}^{\infty}\frac{\left(\frac{\lambda}{T}\right)^l\mathbb{E}\left[\|\boldsymbol{w}_k\|_\infty^l\right]}{l!}.$$

We apply the following lemma:

Lemma IV.1 (Bounded moments of normed-subGaussian vectors [13]) *If Assumption II.1 holds, then* $\mathbb{E}\left[\|\boldsymbol{w}_k\|_\infty^l\right] \leq n\sigma^l l^{\frac{l}{2}+1},\ \forall\, l \in \mathbb{Z}_{\geq 0}$. ■

This gives[2]

$$\mathbb{E}\left[\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right] \leq 1 + n\sum_{l=1}^{\infty}\left(\frac{\lambda\sigma e}{T}\right)^l.$$

[2]We use two facts: 1) $l! \geq (l/e)^l$ and 2) $l^{1-\frac{l}{2}} \leq 1$, for all $l \in \mathbb{Z}_{\geq 0}$, where the constant $e = 2.71828....$

To tighten the previous upper bound, consider any λ such that $\lambda \in [0, \frac{T}{\sigma e})$. Then the following bound holds[3]

$$\mathbb{E}\left[\exp\left(\frac{\lambda}{T}\|\boldsymbol{w}_k\|_\infty\right)\right] \leq 1 + \frac{\lambda\sigma ne}{T - \lambda\sigma e} \leq \exp\left(\frac{\lambda\sigma ne}{T - \lambda\sigma e}\right).$$

Finally, we achieve

$$\text{Prob}\left(\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\boldsymbol{w}_k\|_\infty) \geq \gamma\right) \leq \exp\left(-\gamma\lambda + \sum_{k\in\mathcal{T}}\frac{\lambda\sigma ne}{T - \lambda\sigma e}\right).$$

Finding an optimal bound is hard, and therefore we find a sub-optimal bound by selecting λ to be

$$\lambda = \begin{cases} \frac{T}{2\sigma e} - \frac{nT}{2\gamma}, & \text{if } \gamma \geq \sigma ne, \\ 0, & \text{if } \gamma < \sigma ne. \end{cases}$$

Then, we have the following

$$\text{Prob}\left(\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\boldsymbol{w}_k\|_\infty) \geq \gamma\right) \leq \begin{cases} \exp\left(-\frac{(\sigma ne - \gamma)^2 T^2}{2[(2T-1)\gamma\sigma e + n(\sigma e)^2]}\right), & \text{if } \gamma \geq \sigma ne, \\ 1, & \text{if } \gamma < \sigma ne. \end{cases}$$

In words, we have the stated probability bounds on the quality of $\boldsymbol{\alpha}$ with any $\gamma \geq nc$, where $c := \sigma e\eta\sqrt{np}\,\sigma_{\min}^{-1}(A)$.

(2) (A bound with slow confidence growth): By the Markov inequality, we obtain a bound (6) as

$$\text{Prob}\left(\frac{1}{T}\sum_{k\in\mathcal{T}} (\|\boldsymbol{w}_k\|_\infty) \geq \gamma\right) \leq \frac{1}{\gamma T}\sum_{k\in\mathcal{T}}\mathbb{E}\left[\|\boldsymbol{w}_k\|_\infty\right].$$

By Lemma IV.1, we have $\mathbb{E}\left[\|\boldsymbol{w}_k\|_\infty\right] \leq n\sigma$, resulting in

$$\text{Prob}\left(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty \leq \gamma\right) \geq 1 - \frac{1}{\gamma}n\sigma\eta\sqrt{np}\sigma_{\min}^{-1}(A),$$

with non-trivial bound if we take $\gamma \geq n\sigma\eta\sqrt{np}\ \sigma_{\min}^{-1}(A)$. The complete proof can be found in [18]. ■

Theorem IV.2 provides an online computation of a real-time $\boldsymbol{\alpha}$ that is close to $\boldsymbol{\alpha}^\star$ within a time varying distance γ with arbitrary high probability, where this distance γ depends only on the environment predictors as well as on the data sets. Note that, the confidence of selecting $\gamma > nc$ as a bound of $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty$ increases exponentially as we increase the length T of the data sets. This motivates us to propose a calculable dynamic ambiguity set, described as in Theorem IV.1, by selecting its dynamic radius as

$$\hat{\epsilon} = \epsilon + \gamma H(t, T, \boldsymbol{d}), \quad (7)$$

where ϵ, $\gamma > nc$ and H are chosen as in Lemma III.1, Theorem IV.2, and Theorem IV.1, respectively. Such selection results in modified guarantees of (2) as follows

$$\text{Prob}\left(\mathbb{P}_{t+1} \in \mathcal{P}_{t+1}\right) \geq (1-\beta)\left(1 - \exp\left(-\frac{(nc-\gamma)^2 T^2}{2\left[(2T-1)c\gamma + nc^2\right]}\right)\right), \quad (8)$$

which is the consequence of previous theorems and the independence of the selection of ϵ and γ. The guarantee (8) reveals a very sharp and effective characterization of $\mathbb{P}_{t+1}$, where as time t increases with a selection of $T_0 = \infty$ (or

[3]We use the fact: $1 + x \leq \exp(x)$ for $x \in \mathbb{R}$.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.

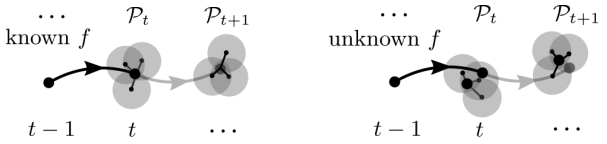

Fig. 1: Online characterization of $\mathcal{P}_{t+1}$, with (without) f. The dark line is the trajectory of $\boldsymbol{x}$ and the gray part is yet to be revealed. At t, we obtain $\mathcal{P}_{t+1}$ with its elements supported on $T_0 = 3$ shaded regions with high probability. Each region k has center ξ_k $(\sum_i \alpha_i \xi_k^{(i)})$ and radius proportional to ϵ $(\hat{\epsilon})$. Note the centers of these regions are related to a known (learned) point $f(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d})$ $(\sum_i \alpha_i f^{(i)}(t, \hat{\boldsymbol{x}}_t, \boldsymbol{d}))$, they are close if the learning is effective.

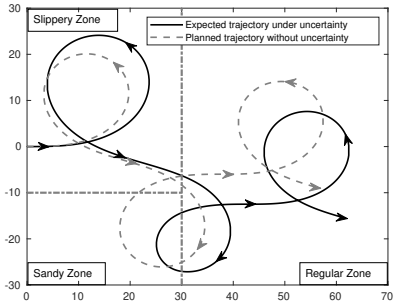

Fig. 2: Path plan and actual trajectory in various $\mathbb{R}^2$ road zones.

$T = t$), the confidence value on the right hand side increases to $1 - \beta$ exponentially fast. Notice that the convergence rate is dominated by the size T of online data. Fig. 1 compares adaptation of the ambiguity set with and without knowing f.

Remark IV.1 (Data-driven selection of the radius) The radius of the adaptive ambiguity set (7) depends on the unknown, noise-related parameter σ, the regularization constant η, and on the online parameters $\sigma_{\min}(A)$. In many engineering problems, an upper bound σ of the noise-related parameter can be determined *a-priori* or empirically. The parameter η, together with the regularization matrices P, are introduced to ensure that (4) is well posed. In particular, P can be a diagonal matrix with each diagonal term scaling its corresponding components. At each t, the computation (4) needs an additional online regularization matrix, denoted by P_{t-1}. For example, P_{t-1} can be a diagonal matrix with the j^{th} diagonal term equal to $1/(\sqrt{p}\max_{i\in\{1,\ldots,p\}} |f_{t-1}^{(i)}(j)|)$, where $f_{t-1}^{(i)}(j)$ is the j^{th} component of $f_{t-1}^{(i)}$, which results in $\eta = 1$. Finally, $\sigma_{\min}(A)$ relies on the selection of the model set $\{f^{(i)}\}_i$ as well as the other two parameters η and σ. In practice, all the zero singular values of A is perturbed by the noise with a factor of σ. One could select the minimal non-zero principal singular value to be $\sigma_{\min}(A) = \min\{\sigma_i(A) \mid \sigma_i(A) > \sigma, i \in \{1, \ldots, p\}\}$.

Online Procedure: To summarize, our online learning methodology is given in Algorithm table 1. Our approach aims to characterizes the unknown f via $\mathcal{P}$ online. This is achieved by assimilating online data $\mathcal{I}$, together with *a-priori* knowledge of $\boldsymbol{d}$, then learning model parameter $\boldsymbol{\alpha}$, and lastly identifying a center distribution $\hat{\mathbb{P}}$ and radius $\hat{\epsilon}$ of $\mathcal{P}$. Notice that the complexity of the algorithm is dominated by the calculation of α via SVD to the system of equations (4). An approximated solution to (4) leads to an enlargement of $\hat{\epsilon}$ by an extra approximation error in practice.

$\mathbb{P}$-Learning 1 Learn$(\mathcal{I}, \boldsymbol{d})$

Require: $\{f^{(i)}\}_i$, β, T_0, σ, θ and $t = 1$;
Ensure: Online $\boldsymbol{\alpha}$, $\hat{\mathbb{P}}$, $\hat{\epsilon}$
1: **repeat**
2: Update data set $\mathcal{I} := \mathcal{I}_t$ and knowledge $\boldsymbol{d} := \boldsymbol{d}_t$;
3: Compute $\boldsymbol{\alpha} := \boldsymbol{\alpha}_t$ as in (4);
4: Select $\hat{\mathbb{P}}_{t+1}$ as in (3) and $\hat{\epsilon} := \hat{\epsilon}_t$ as in (7);
5: Leverage $(\hat{\mathbb{P}}_{t+1}, \hat{\epsilon})$ as characterization of f;
6: $t \leftarrow t + 1$;
7: **until** time t stops.

V. Simulations

In this section, we illustrate the previous results on a simple vehicle example. Consider a vehicle driving under various road conditions, where its control signal is derived in advance, according to a path-planner in an ideal environment.

Our goal is to learn the real-time environment and estimate the system states via our adaptive $\mathbb{P}$-Learning algorithm. Our vehicle is modeled as a differential-drive robot subject to uncertainty, see [19]:

$$\begin{aligned} x_1^+ &= x_1 + h\cos(x_3)u_1 + hw_1, \\ x_2^+ &= x_2 + h\sin(x_3)u_1 + hw_2, \\ x_3^+ &= x_3 - hu_2 + hw_3, \\ u_1 &= \frac{r}{2}(v_l + v_r + e_1), \\ u_2 &= \frac{r}{2R}(v_l - v_r + e_2), \end{aligned} \tag{9}$$

where $\boldsymbol{x} := (x_1, x_2, x_3) \in \mathbb{R}^2 \times [-\pi, \pi) \cong \mathbb{R} \times \mathbb{S}^1$ stands for vehicle position and orientation on the 2-D plane. We denote by $\boldsymbol{x}^+$ the state at the next time step and $\boldsymbol{w} := (w_1, w_2, w_3)$ a zero-mean, mixture of Gaussian and Uniform distributions, which are subGaussian uncertainties with $\sigma = 0.5$. We assume $\boldsymbol{x}_0 = (0, 0, 0)$ and $h = 10^{-3}$. The velocity $\boldsymbol{u} := (u_1, u_2)$ is determined by a wheel radius $r = 0.15$ m, the distance between wheels $R = 0.4$ m, the given wheel speed $\boldsymbol{d} := (v_l, v_r)$ and an unknown parameter $\boldsymbol{e} := (e_1, e_2)$, which depends on the wheel and road conditions. For simulation purposes, we assume that the vehicle may move over three road zones, a slippery zone with $\boldsymbol{e}^{(1)} = (4, 0)$, a sandy zone with $\boldsymbol{e}^{(2)} = (-6, 0)$, and a smooth, regular zone with $\boldsymbol{e}^{(3)} = (0, 0)$, as described in Fig. 2. The vehicle executes the following left and right wheel speed plan (rad/s):

$$\begin{aligned} v_l &= 10 - 0.5\sin(20h\pi t), \\ v_r &= 10 + 0.5\sin(20h\pi t). \end{aligned}$$

Now we employ our adaptive learning algorithm for the characterization of the uncertain vehicle states and learning of the unknown road-condition parameter $\boldsymbol{e}$ in real time. To do this, we take $p = 3$ predictors as in (9) with $\boldsymbol{w} \equiv 0$, and

$$\begin{aligned} i = 1, \quad e_1 = 0, \qquad e_2 = 0, \\ i = 2, \quad e_1 = 10, \qquad e_2 = 0, \\ i = 3, \quad e_1 = 0, \qquad e_2 = 10. \end{aligned}$$

Note that Assumption IV.1 holds with $\boldsymbol{\alpha}^\star := (0.6, 0.4, 0)$ in the slippery zone, $\boldsymbol{\alpha}^\star := (1.6, -0.6, 0)$ in the sandy zone and $\boldsymbol{\alpha}^\star := (1, 0, 0)$ in the smooth zone. We select $T_0 = 300$,

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.

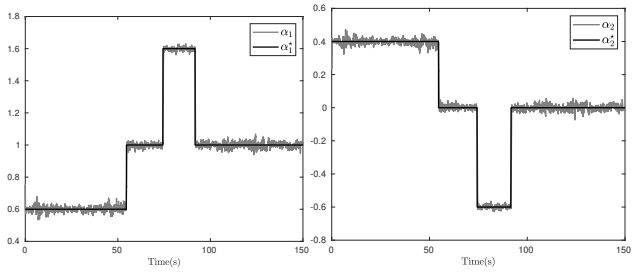

Fig. 3: Real-time learning parameter α_1 and α_2.

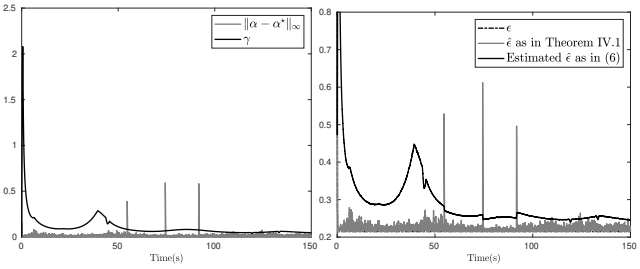

Fig. 4: Quality of $\boldsymbol{\alpha}$ and the estimated radius $\hat{\epsilon}$.

and, at each time t, we have access to model sets $\{f^{(i)}\}_i$ as well as the real-time data set $\mathcal{I}_t$ and $\boldsymbol{d}$. Note that the notions of inner product and norm are those defined on the vector space $T(\mathbb{R}^2 \times \mathbb{S}) \equiv \mathbb{R}^3$. Recall that $h = 10^{-3}$, so a $T_0 = 300$ corresponds to a time window of order 0.3sec. We select online diagonal regularization matrices P with diagonal $(1/(\sqrt{3}\max_{i=1,2,3}|f^{(i)}(j)|)$ for $j = 1, 2$ and 1 for $j = 3$, resulting in $\eta = \max_{i,k\in\mathcal{T}} \|P_k f_k^{(i)}\|$.

Fig. 3 demonstrates the real-time parameter learning of α_1 and α_2. It can be seen that these unknown parameters are effectively learned and tracked over time. Fig. 4 shows the quality of the learned parameter $\boldsymbol{\alpha}$ and its effect on the determination of the radius of the adaptive ambiguity set. We note that, for a particular noise realization sequence, the estimated value $\gamma = n\sigma e\eta\sqrt{np}\sigma_{\min}^{-1}(A) + \theta$, with $\theta = 0.01$, upper bounds $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star\|_\infty$ in high probability. The large spikes in the figure are due to the change of zone, resulting in a large error. This is expected, as the true $\boldsymbol{\alpha}^\star$ changed discontinuously. Meanwhile, the estimated radius $\hat{\epsilon}$ of the adaptive ambiguity set, calculated as in (7), is a conservative estimate of the unknown *a-priori* $\hat{\epsilon}$ as in Theorem IV.1. The true $\hat{\epsilon}$ captures exactly the ambiguity set over the time sequence $\mathcal{T}$, for a $\beta = 0.05$. Over time, we empirically see the difference between the approximated $\hat{\epsilon}$ via γ and the true one become close. In practice, the radius $\hat{\epsilon}$ can be selected in a data-driven fashion, e.g., as in Remark IV.1, to serve as a way for less conservative estimation of the radius in probability. We show in Fig. 5 the online guarantee (8) of this particular case study, and various samples of (8), obtained by taking different time horizon T_0.

VI. Conclusions

In this paper, we proposed an approach for online learning of unknown and uncertain dynamical environments in a parameterized class. The proposed method allows us to learn the environment, while providing an online characterization

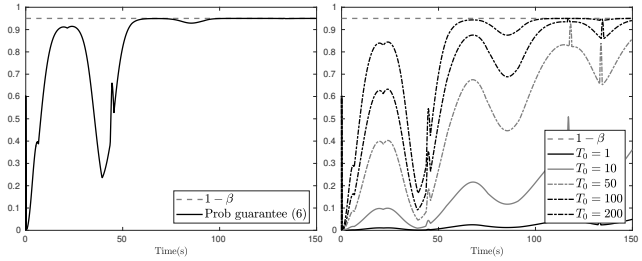

Fig. 5: Online guarantee (8) and samples of (8) with various T_0.

of the approximation via online-quantifiable probabilistic guarantees. The approach opens a way for the robust integration of the online learning with control design. A robotic example was used to demonstrate the efficacy of the method.

References

[1] A. Chiuso and G. Pillonetto, "System identification: A machine learning perspective," *Robotics and Autonomous Systems*, vol. 2, pp. 281–304, 2019.

[2] A. H. Qureshi, Y. Miao, A. Simeonov, and M. C. Yip, "Motion planning networks: Bridging the gap between learning-based and classical motion planners," *IEEE Transactions on Robotics*, pp. 1–19, 2020.

[3] A. Sproewitz, R. Moeckel, J. Maye, and A. Ijspeert, "Learning to move in modular robots using central pattern generators and online optimization," *International Symposium on Robotic Research*, vol. 27, no. 3-4, pp. 423–443, 2008.

[4] L. Ljung, *System identification*. Prentice Hall, 1999.

[5] M. Verhaegen and V. Verdult, *Filtering and system identification: a least squares approach*. Cambridge university press, 2007.

[6] M. Milanese and C. Novara, "Unified set membership theory for identification, prediction and filtering of nonlinear systems," *Automatica*, vol. 47, no. 10, pp. 2141–2151, 2011.

[7] C. Novara, A. Nicolì, and G. C. Calafiore, "Nonlinear system identification in Sobolev spaces," *preprint arXiv:1911.02930*, 2019.

[8] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," in *Int. Conf. on Machine Learning*, 2019, pp. 5610–5618.

[9] S. Oymak and N. Ozay, "Non-asymptotic identification of LTI systems from a single trajectory," in *American Control Conference*, 2019, pp. 5655–5661.

[10] A. Tsiamis and G. J. Pappas, "Finite-sample analysis of stochastic system identification," in *IEEE Int. Conf. on Decision and Control*, 2019, pp. 3648–3654.

[11] S. Fattahi, N. Matni, and S. Sojoudi, "Learning sparse dynamical systems from a single sample trajectory," in *IEEE Int. Conf. on Decision and Control*, 2019, pp. 2682–2689.

[12] N. Fournier and A. Guillin, "On the rate of convergence in Wasserstein distance of the empirical measure," *Probability Theory and Related Fields*, vol. 162, no. 3-4, p. 707–738, 2015.

[13] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*. Cambridge University Press, 2018, vol. 47.

[14] D. Boskos, J. Cortés, and S. Martinez, "Dynamic evolution of distributional ambiguity sets and precision tradeoffs in data assimilation," in *European Control Conference*, Naples, Italy, Jun. 2019, pp. 2252–2257.

[15] L. V. Kantorovich and G. S. Rubinstein, "On a space of completely additive functions," *Vestnik Leningrad. Univ*, vol. 13, no. 7, p. 52–59, 1958.

[16] J. Niles-Weed and P. Rigollet, "Estimation of Wasserstein distances in the spiked transport model," *arXiv preprint arXiv:1909.07513*, 2019.

[17] J. Lei, "Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces," *arXiv preprint arXiv:1804.10556*, 2018.

[18] D. Li, D. Fooladivanda, and S. Martínez, "Online learning of parameterized uncertain dynamical environments with finite-sample guarantees," *IEEE Control Systems Letters*, 2020.

[19] S. M. LaValle, *Planning algorithms*. Cambridge University Press, 2006.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on October 28,2023 at 14:05:03 UTC from IEEE Xplore. Restrictions apply.