Learning-based Robust Motion Planning with Guaranteed Stability: A Contraction Theory Approach

Hiroyasu Tsukamoto and Soon-Jo Chung

***Abstract*—This paper presents Learning-based Autonomous Guidance with RObustness and Stability guarantees (LAG-ROS), which provides machine learning-based nonlinear motion planners with formal robustness and stability guarantees, by designing a differential Lyapunov function using contraction theory. LAG-ROS utilizes a neural network to model a robust tracking controller independently of a target trajectory, for which we show that the Euclidean distance between the target and controlled trajectories is exponentially bounded linearly in the learning error, even under the existence of bounded external disturbances. We also present a convex optimization approach that minimizes the steady-state bound of the tracking error to construct the robust control law for neural network training. The journal version of this paper (*IEEE Robotics and Automation Letters (RA-L)*, to appear, 2021) with detailed mathematical proofs and simulation results can be found in https://arxiv.org/abs/2102.12668 [1].**

***Index Terms*—Machine Learning for Robot Control, Robust/Adaptive Control, and Optimization & Optimal Control.**

I. Introduction

Learning-based guidance and control designs of nonlinear systems have been an emerging area of research since the rise of neural networks and reinforcement learning [2], [3]. These techniques can be categorized into model-free and model-based methods, where each of them has pros and cons. The former approach is supposed to learn desired optimal policies which work robustly using training data obtained in real-world environments, making them not suitable for situations where sampling large training datasets is difficult. Also, proving stability and robustness properties of such model-free systems is challenging in general, although some approaches do exist [4], [5]. In contrast, the latter approach allows sampling as much data as we want to design optimal policies by, e.g., reinforcement learning [6], imitation learning [7], or both [8]. However, the learned controller could yield cascading errors in the real-world environment if its nominal model poorly represents the true underlying dynamics.

Control theoretical approaches to circumvent such difficulties include robust Model Predictive Control (MPC) [9] equipped with a feedback control law for stability and robustness guarantees. Among these are contraction theory-based robust control [10]–[13], proposed to guarantee tracking to any feasible target trajectory computed externally by existing motion planners, thereby robustly keeping the system trajectories in a control invariant tube that satisfies given state constraints. Although these provable guarantees are promising, they still

The authors are with the Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA, {htsukamoto, sjchung}@caltech.edu.

Code: https://github.com/astrohiro/lagros.

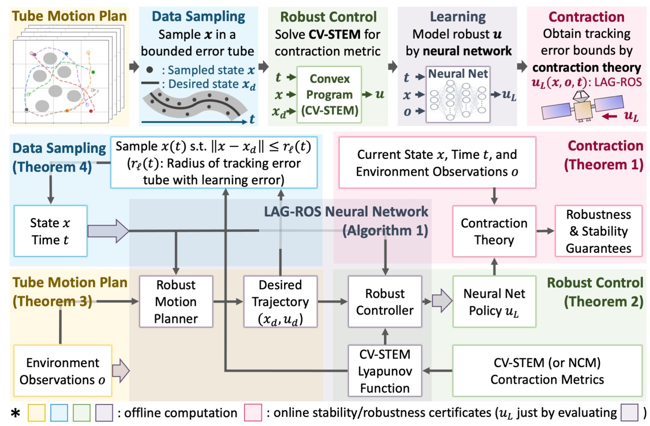

Fig. 1. Illustration of LAG-ROS: Note that the LAG-ROS requires only one neural net evaluation to get u_L, and Theorem 1 provides its stability and robustness guarantees. See [1] for details.

assume that the target trajectory can be computed in real-time solving optimal motion planning problems, unlike most learning-based control frameworks.

Contributions: In this study, we present Learning-based Autonomous Guidance with Robustness, Optimality, and Stability guarantees (LAG-ROS) as a novel way to bridge the gap between the learning-based and robust MPC-based motion planning techniques. In particular, whilst the LAG-ROS requires only one neural network evaluation to get its control input as in the machine learning schemes [6]–[8], its internal contraction theory-based architecture still allows obtaining formal stability and robustness guarantees, which have been challenging to quantify without a contracting property [10]–[13]. The proposed LAG-ROS framework depicted in Fig. 1 is summarized as follows.

The theoretical foundation of the LAG-ROS rests on contraction theory, which utilizes a contraction metric to characterize a necessary and sufficient condition of exponential incremental stability of nonlinear system trajectories [14]. The central result of this paper is that, if there exists a control law which renders a nonlinear system contracting, or equivalently, the closed-loop system has a contraction metric, then the LAG-ROS trained to imitate the controller ensures the Euclidean distance between the target and controlled trajectories to be exponentially bounded in time, linearly in the learning error and size of perturbation. This property helps quantify how small the learning error should be in practice, giving some intuition on the number of samples and length of time in neural net training. We further show that such a contracting control law and corresponding contraction metric can be designed explicitly via convex optimization using the method of CV-STEM [12], [13], [15], in order to minimize a steady-state upper bound of the LAG-ROS tracking

TABLE I
COMPARISON OF THE PROPOSED METHOD WITH THE LEARNING-BASED AND ROBUST TUBE-BASED MOTION PLANNERS.

Motion planning scheme	Policy to be learned	State tracking error $\|x-x_d\|$	Computational load
(a) Learning-based motion planner [6]–[8]	$(x,o,t)\mapsto u_d$	Increases exponentially (Lemma 1)	One neural net evaluation
(b) Robust tube-based motion planner [9]–[13]	$(x,x_d,t)\mapsto u^*$	Exponentially bounded (Theorem 2)	Computation required to get x_d
(c) Proposed method (LAG-ROS)	$(x,o,t)\mapsto u^*$	Exponentially bounded (Theorem 1)	One neural net evaluation

error. Further details and simulation results can be found in https://arxiv.org/abs/2102.12668 [1].

Notation: For $x\in\mathbb{R}^n$ and $A\in\mathbb{R}^{n\times m}$, we let $\|x\|$, δx, and $\|A\|$, denote the Euclidean norm, infinitesimal variation of x, and induced 2-norm, respectively. For a square matrix A, we use the notation $A\succ 0$, $A\succeq 0$, $A\prec 0$, and $A\preceq 0$ for the positive definite, positive semi-definite, negative definite, negative semi-definite matrices, respectively, and $\mathrm{sym}(A)=(A+A^\top)/2$. Also, $I\in\mathbb{R}^{n\times n}$ denotes the identity matrix.

II. Learning-based Robust Motion Planning with Guaranteed Stability (LAG-ROS)

In this paper, we consider the following nonlinear dynamical systems with a controller $u\in\mathbb{R}^m$:

$$\dot{x}=f(x,t)+B(x,t)u+d(x,t) \tag{1}$$
$$\dot{x}_d=f(x_d,t)+B(x_d,t)u_d(x_d,o(t),t) \tag{2}$$

where $t\in\mathbb{R}_{\geq 0}$, $f:\mathbb{R}^n\times\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^n$, $B:\mathbb{R}^n\times\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^{n\times m}$, $x:\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^n$ is the state trajectory of the true dynamical system (1) perturbed by the bounded disturbance $d:\mathbb{R}^n\times\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^n$ s.t. $\sup_{x,t}\|d(x,t)\|=\bar{d}$, $o:\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^p$ is an environment observation, $x_d\in\mathbb{R}^n$ and $u_d:\mathbb{R}^n\times\mathbb{R}^p\times\mathbb{R}_{\geq 0}\mapsto\mathbb{R}^n$ are the target state trajectory and control policy given by existing motion planning algorithms.

A. Problem Formulation of LAG-ROS

The problem of our interest is to find u that is computable with one neural network evaluation and that guarantees exponential boundedness of $\|x-x_d\|$ in (1) and (2), robustly against the learning error and external disturbances. The objective of this paper is thus not to develop new learning-based motion planning algorithms that compute x_d but to augment them with formal robustness and stability guarantees. To this end, let us briefly review the following existing motion planning techniques and their inherent limitations:

(a) Learning-based motion planner [6]–[8]:
$(x,o,t)\mapsto u_d$, approximately.

(b) Robust tube-based motion planner [9]–[13]:
$(x,x_d,t)\mapsto u^*$, where u^* is a robust tracking controller.

The robust tube-based motion planner (b) ensures that the perturbed trajectories x of (1) stay in an exponentially bounded error tube around the target trajectory x_d of (2) [9]–[13] (see Theorem 2). However, it requires the online computation of x_d as an input to their control policy, which is not realistic for systems with limited computational resources.

The learning-based motion planner (a) circumvents this issue by modeling the target policy $(x,o,t)\mapsto u_d$ by a neural network. In essence, our approach, to be proposed in Theorem 1, is for providing (a) with the contraction theory-based stability guarantees (b). We remark that (a) can only assure the tracking error $\|x-x_d\|$ to be bounded by a function which exponentially increases with time, as to be shown in Lemma 1 for comparison with LAG-ROS of Theorem 1.

B. Stability Guarantees of LAG-ROS

The approach of LAG-ROS bridges the gap between (a) and (b) by ensuring that the distance between the target and controlled trajectories to be exponentially bounded.

(c) Proposed approach (LAG-ROS):
$(x,o,t)\mapsto u^*$, modeled by a neural network $u_L(x,o,t)$ of Theorem 1, where o is a vector containing environment observations (see Fig. 1) and $u^*(x,x_d,t)$ of (b) is viewed as $u^*(x,x_d(x,o,t),t)$, which is a function of (x,o,t).

Theorem 1: Suppose that (1) is controlled to track (2) by the LAG-ROS control policy $u_L(x,o,t)$, learned to satisfy

$$\|u_L(x,o,t)-u^*(x,x_d(x,o,t),t)\|\leq\epsilon_\ell,\ \forall x,o,t$$

where $\epsilon_\ell\in[0,\infty)$ is the learning error, u^* is the target robust control law of (b) (to be designed in Theorem 2), and x_d is given by the robust motion planner (b). Now consider the following virtual system which has x of (1) and x_d of (2) as its particular solutions:

$$\dot{y}=\zeta(y,x,x_d,t)+d_y(y,t) \tag{3}$$

where ζ and d_y are parameterized by y to verify $\zeta(y=x,x,x_d,t)=f(x,t)+B(x,t)u^*$, $\zeta(y=x_d,x,x_d,t)=f(x_d,t)+B(x_d,t)u_d(x_d,o(t),t)$, $d_y(y=x,t)=B(x,t)(u_L-u^*)+d(x,t)$, and $d_y(y=x_d,t)=0$, and x, x_d, u_d, and d are as defined in (1) and (2). Note that $y=x$ and $y=x_d$ are indeed particular solutions of (3). If $\exists\bar{b}\in[0,\infty)$ s.t. $\|B(x,t)\|\leq\bar{b}$, $\forall x,t$, and if u^* is constructed to satisfy the following partial contraction conditions [16] with respect to y, for a contraction metric $M(y,x,x_d,t)=\Theta^\top\Theta\succ 0$ and $\alpha,\underline{\omega},\overline{\omega}\in(0,\infty)$:

$$\dot{M}+2\,\mathrm{sym}\left(M\frac{\partial\zeta}{\partial y}\right)\preceq -2\alpha M,\ \forall y,x,x_d,t \tag{4}$$
$$\overline{\omega}^{-1}I\preceq M\preceq\underline{\omega}^{-1}I,\ \forall y,x,x_d,t \tag{5}$$

then we have the following bound for $\mathsf{e}=x-x_d$:

$$\|\mathsf{e}(t)\|\leq\mathscr{R}(0)\sqrt{\overline{\omega}}e^{-\alpha t}+\frac{\bar{b}\epsilon_\ell+\bar{d}}{\alpha}\sqrt{\frac{\overline{\omega}}{\underline{\omega}}}(1-e^{-\alpha t})=r_\ell(t) \tag{6}$$

where $\mathscr{R}(t)=\int_{x_d}^{x}\|\Theta\delta y(t)\|$ for $M=\Theta^\top\Theta$.

Proof: Let $V=\int_{x_d}^{x}\delta y^\top M\delta y=\int_{x_d}^{x}\|\Theta\delta y\|^2$. Since $\|d_y(x,t)\|\leq\bar{b}\epsilon_\ell+\bar{d}=\bar{d}_{\epsilon_\ell}$ for d_y in (3), the relation (4) gives

$$\begin{aligned}\dot{V}&\leq\int_{x_d}^{x}\delta y^\top\left(\dot{M}+2\,\mathrm{sym}\left(M\frac{\partial\zeta}{\partial y}\right)\right)\delta y+2\bar{d}_{\epsilon_\ell}\int_{x_d}^{x}\|M\delta y\|\\&\leq-2\alpha V+\frac{2\bar{d}_{\epsilon_\ell}}{\sqrt{\underline{\omega}}}\int_{x_d}^{x}\|\Theta\delta y\|\end{aligned}$$

Since $d(\|\Theta\delta y\|^2)/dt = 2\|\Theta\delta y\|(d\|\Theta\delta y\|/dt)$, this implies that $\dot{\mathscr{R}} \leq -\alpha\mathscr{R} + \bar{d}_{\varepsilon_\ell}/\sqrt{\underline{\omega}}$. Therefore, applying the comparison lemma [17, pp.102-103, pp.350-353] (i.e., if $\dot{v}_1 \leq h(v_1,t)$ for $v_1(0) \leq v_2(0)$ and $\dot{v}_2 = h(v_2,t)$, then $v_1(t) \leq v_2(t)$), along with the relation $\mathscr{R}(t) \geq \|\mathsf{e}(t)\|/\sqrt{\underline{\omega}}$, results in (6) ■

Theorem 1 implies that the bound (6) decreases linearly in the learning error ε_ℓ and disturbance $\bar{d}$, and (1) controlled by LAG-ROS is exponentially stable when $\varepsilon_\ell = 0$ and $\bar{d} = 0$, showing a great improvement over (a) which only gives an exponentially diverging bound as to be derived in Lemma 1 [6]–[8]. This property permits quantifying how small $\varepsilon_\ell = 0$ should be to meet the required performance of motion planning, giving some intuition on the neural network architecture. Also, since we model u^* by $u_L(x,o,t)$ independently of x_d, it is indeed implementable without solving any motion planning problems online unlike robust motion planners (b) [9]–[13], as outlined in Table I. If we can sample training data of u^* explicitly considering the bound (6), the LAG-ROS control enables guaranteeing given state constraints even with the learning error ε_ℓ and external disturbance $d(x,t)$ [1].

To appreciate the importance of the guarantees in Theorem 1, let us additionally show that (a), which models $(x,o,t) \mapsto u_d$, only leads to an exponentially diverging bound.

Lemma 1: Suppose that u of (1) is learned to satisfy

$$\|u(x,o,t) - u_d(x,o,t)\| \leq \varepsilon_\ell,\ \forall x,o,t \tag{7}$$

for u_d of (2) with the learning error $\varepsilon_\ell \in [0,\infty)$, and that $\exists \bar{b}$ s.t. $\|B(x,t)\| \leq \bar{b},\ \forall x,t$. If $f_{c\ell} = f(x,t) + B(x,t)u_d$ is Lipschitz, i.e., $\exists L_f \in [0,\infty)$ s.t. $\|f_{c\ell}(x_1,t) - f_{c\ell}(x_2,t)\| \leq L_f\|x_1 - x_2\|,\ \forall x_1,x_2 \in \mathbb{R}^n$, then we have the following bound:

$$\|\mathsf{e}(t)\| \leq \|\mathsf{e}(0)\|e^{L_f t} + L_f^{-1}(\bar{b}\varepsilon_\ell + \bar{d})(e^{L_f t} - 1). \tag{8}$$

where $\mathsf{e} = x - x_d$, and x, x_d, and $\bar{d}$ are given in (1) and (2).

Proof: Integrating (1) and (2) for u in (7) yields $\|\mathsf{e}(t)\| \leq \|\mathsf{e}(0)\| + L_f\int_0^t\|\mathsf{e}(\tau)\|d\tau + (\bar{b}\varepsilon_\ell + \bar{d})t$. Applying the Gronwall-Bellman inequality [17, pp. 651] gives

$$\|\mathsf{e}(t)\| \leq \|\mathsf{e}(0)\| + \bar{d}_{\varepsilon_\ell}t + L_f\int_0^t(\|\mathsf{e}(0)\| + \bar{d}_{\varepsilon_\ell}\tau)e^{L_f(t-\tau)}d\tau$$

where $\bar{d}_{\varepsilon_\ell} = \bar{b}\varepsilon_\ell + \bar{d}$ and $\Delta\tau = \tau - t_0$. Thus, integration by parts results in the desired relation (8). ■

Lemma 1 indicates that if there exists either a learning error ε_ℓ or external disturbance d, the tracking error bound grows exponentially with time, and thus (8) becomes no longer useful for large t. In [1], we demonstrate how the computed bounds of (6) ($\lim_{t\to\infty}e^{-\alpha t} = 0$) and (8) ($\lim_{t\to\infty}e^{L_f t} = \infty$) affect the control performance in practice.

III. Contraction Theory-based Robust and Optimal Tracking Control

Theorem 1 is subject to the assumption that we have a contraction theory-based robust tracking control law u^*, which satisfies (4) and (5) for a given (x_d,u_d). This section thus delineates one way to extend the method called ConVex Optimization-based Steady-state Tracking Error Minimization (CV-STEM) [12], [13], [15] to find a contraction metric M of Theorem 1, which minimizes an upper bound of the steady-state error of (6) via convex optimization.

In addition, we modify the CV-STEM in [12] to derive a robust control input u^* which also greedily minimizes the deviation of u^* from the target u_d, using the computed contraction metric M to construct a differential Lyapunov function $V = \delta y^\top M\delta y$. Note that u^* is to be modeled by a neural network which maps (x,o,t) to u^* implicitly accounting for (x_d,u_d) as described in Theorem 1, although u^* takes (x,x_d,t) as its inputs.

A. Problem Formulation of CV-STEM Tracking Control

For given (x_d,u_d), we assume that u^* of Theorem 1 can be decomposed as $u^* = u_d(x_d,o(t),t) + K(x,x_d,t)(x - x_d)$, the generality of which is guaranteed by the following lemma.

Lemma 2: Consider a general tracking controller u defined as $u = k(x,x_d,t)$ with $k(x_d,x_d,t) = u_d(x_d,o(t),t)$, where $k : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq 0} \mapsto \mathbb{R}^m$. If k is piecewise continuously differentiable, then $\exists K : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{\geq 0} \mapsto \mathbb{R}^{m\times n}$ s.t. $u = k(x_d,x_d,t) = u_d(x_d,o(t),t) + K(x,x_d,t)(x - x_d)$.

Proof: We have $u = u_d + (k(x,x_d,t) - k(x_d,x_d,t))$ due to $k(x_d,x_d,t) = u_d(x_d,o(t),t)$. Since $k(x,x_d,t) - k(x_d,x_d,t) = \int_0^1(dk(cx + (1-c)x_d,x_d,t)/dc)dc$, choosing K as $\int_0^1(\partial k(cx + (1-c)x_d,x_d,t)/\partial x)dc$ gives the desired relation. ■

Lemma 2 implies that designing optimal k of $u^* = k(x,x_d,t)$ reduces to designing optimal $K(x,x_d,t)$ of $u^* = u_d + K(x,x_d,t)(x - x_d)$. When (1) is controlled by the LAG-ROS u_L of Theorem 1 with such u^*, the virtual system of (1) and (2) which has $y = x$ and $y = x_d$ as its particular solutions can be given by (3), where ζ is defined as follows:

$$\zeta(y,x,x_d,t) = \dot{x}_d + (A(x,x_d,t) + B(x,t)K(x,x_d,t))(y - x_d) \tag{9}$$

where A is the State-Dependent Coefficient (SDC) form of the dynamical system (1) given by Lemma 2 of [15], which verifies $A(x,x_d,t)(x - x_d) = f(x,t) + B(x,t)u_d - f(x_d,t) - B(x_d,t)u_d$. Note that ζ indeed satisfies $\zeta(x,t) = f(x,t) + B(x,t)u^*$ and $\zeta(x_d,t) = f(x_d,t) + B(x_d,t)u_d$ for such A, to have $y = x$ and $y = x_d$ as the particular solutions to (3).

B. CV-STEM Contraction Metrics as Lyapunov Functions

The remaining task is to construct M so that it satisfies (4) and (5). The CV-STEM approach suggests that we can find such M via convex optimization to minimize an upper bound of (6) as $t \to \infty$ when α of (4) is fixed. Theorem 2 proposes using the metric M designed by the CV-STEM for a Lyapunov function, thereby augmenting u^* with additional optimality to greedily minimize $\|u^* - u_d\|^2$ for u_d in (2).

Theorem 2: Suppose that f and B are piecewise continuously differentiable, and let $B = B(x,t)$ and $A = A(x,x_d,t)$ in (9) for notational simplicity. Consider a contraction metric $M(x,x_d,t) = W(x,x_d,t)^{-1} \succ 0$ given by the following convex optimization (CV-STEM) [12], [13], [15] to minimize an upper bound on the steady-state tracking error of (6):

$$J_{CV}^* = \min_{\nu>0,\chi\in\mathbb{R},\bar{W}\succ 0} \frac{\bar{b}\varepsilon_\ell + \bar{d}}{\alpha}\chi \ \text{ s.t. (11) and (12)} \tag{10}$$

with the convex constraints (11) and (12) given as

$$-\dot{\bar{W}} + 2\,\mathrm{sym}(A\bar{W}) - 2\nu BR^{-1}B^\top \preceq -2\alpha\bar{W},\ \forall x,x_d,t \tag{11}$$

$$I \preceq \bar{W}(x,x_d,t) \preceq \chi I,\ \forall x,x_d,t \tag{12}$$

where $\alpha, \underline{\omega}, \overline{\omega} \in (0,\infty)$, $\nu = 1/\underline{\omega}$, $\chi = \overline{\omega}/\underline{\omega}$, $\bar{W} = \nu W$, and $R = R(x,x_d,t) \succ 0$ is a given weight matrix on the control input. Suppose also that u^* of Theorem 1 is given by $u^* = u_d(x_d,o(t),t) + K^*(x,x_d,t)\mathtt{e}$, where $\mathtt{e} = x - x_d$, and K^* is given by the following convex optimization (x,x_d,t):

$$K^* = \arg\min_{K\in\mathbb{R}^{m\times n}} \|u-u_d\|^2 = \arg\min_{K\in\mathbb{R}^{m\times n}} \|K(x,x_d,t)\mathtt{e}\|^2 \tag{13}$$

$$\text{s.t. } \dot{M} + 2\,\mathrm{sym}\,(MA + MBK(x,x_d,t)) \preceq -2\alpha M. \tag{14}$$

Then M satisfies (4) and (5) for ζ defined in (9), and thus (6) holds, i.e., we have the exponential bound on the tracking error $\|x - x_d\|$ when the dynamics (1) is controlled by the LAG-ROS control input u_L of Theorem 1. Furthermore, the problem (13) is always feasible.

Proof: Since the virtual dynamics of (3) with (9) is given as $\delta\dot{y} = (\partial\zeta/\partial y)\delta y = (A - BK)\delta y$, substituting this into (4) verifies that (4) and (14) are equivalent. For $\bar{K} = -R^{-1}B^\top M$, (14) can be rewritten as

$$\nu^{-1}M(-\dot{\bar{W}} + 2\,\mathrm{sym}\,(A\bar{W}) - \nu BR^{-1}B^\top)M \preceq -2\alpha\nu^{-1}M\bar{W}M.$$

Since this is clearly feasible as long as M satisfies the condition (11), this implies that (13) is always feasible. Also, multiplying (5) by W from both sides and then by ν gives (12). These facts indicate that the conditions (4) and (5) are satisfied for M and u^* constructed by (10) and (13), respectively, and thus we have the exponential bound (6) as a result of Theorem 1. Furthermore, the problem (10) indeed minimizes an upper bound of (6) as $t \to \infty$ due to the relation $0 \leq \sqrt{\overline{\omega}/\underline{\omega}} = \sqrt{\chi} \leq \chi$. We remark that (10) is convex as the objective is affine in χ and (11) and (12) are linear matrix inequalities in terms of ν, χ, and $\bar{W}$. ■

Remark 1: (10) and (13) are convex and thus can be solved computationally efficiently [18, pp. 561]. For systems with a known Lyapunov function, we could simply use it to get u^* in Theorem 2 without solving (13), although optimality may no longer be guaranteed in this case.

Remark 2: The contraction metric construction itself can be performed using a neural network [13], [15], [19], [20], leading to an analogous incremental stability and robustness results to those of Theorem 1 [19], [21].

IV. Conclusion

In this work, we propose the LAG-ROS, a real-time implementable learning-based motion planner with formal stability and robustness guarantees. It extensively utilizes contraction theory to provide an explicit exponential bound on the distance between the target and controlled trajectories, even under the existence of the LAG-ROS modeling error and external disturbances. We remark that other types of disturbances can be incorporated in this framework using [15] for stochastic systems and [19] for parametric uncertain systems. The journal version of this paper with detailed mathematical proofs and simulation results can be found in https://arxiv.org/abs/2102.12668 [1].

References

[1] H. Tsukamoto and S.-J. Chung, "Learning-based robust motion planning with guaranteed stability: A contraction theory approach," *IEEE Robot. Automat. Lett., to appear*, Jun. 2021.

[2] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[3] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[4] N. M. Boffi, S. Tu, N. Matni, J.-J. E. Slotine, and V. Sindhwani, "Learning stability certificates from data," in *CoRL*, Nov. 2020.

[5] D. Ho and J. C. Doyle, "Robust model-free learning and control without prior knowledge," in *IEEE CDC*, 2019, pp. 4577–4582.

[6] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *IEEE/RSJ IROS*, 2018, pp. 3052–3059.

[7] B. Rivière, W. Hönig, Y. Yue, and S. Chung, "GLAS: Global-to-local safe autonomy synthesis for multi-robot motion planning with end-to-end learning," *IEEE Robot. Automat. Lett.*, vol. 5, no. 3, pp. 4249–4256, 2020.

[8] J. Ho and S. Ermon, "Generative Adversarial Imitation Learning," in *Adv. Neural Inf. Process. Syst.*, D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, Eds., vol. 29. Curran Associates, Inc., 2016, pp. 4565–4573.

[9] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*. London: Springer London, 1999, pp. 207–226.

[10] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design," *IEEE Trans. Autom. Control*, vol. 62, no. 6, pp. 3046–3053, Jun. 2017.

[11] S. Singh, A. Majumdar, J.-J. E. Slotine, and M. Pavone, "Robust online motion planning via contraction theory and convex optimization," in *IEEE ICRA*, May 2017, pp. 5883–5890.

[12] H. Tsukamoto and S.-J. Chung, "Robust controller design for stochastic nonlinear systems via convex optimization," *IEEE Trans. Autom. Control, to appear*, Oct. 2021.

[13] H. Tsukamoto and S.-J. Chung, "Neural contraction metrics for robust estimation and control: A convex optimization approach," *IEEE Control Syst. Lett.*, vol. 5, no. 1, pp. 211–216, 2021.

[14] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for nonlinear systems," *Automatica*, no. 6, pp. 683 – 696, 1998.

[15] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Neural stochastic contraction metrics for learning-based control and estimation," *IEEE Control Syst. Lett.*, vol. 5, no. 5, pp. 1825–1830, 2021.

[16] W. Wang and J.-J. E. Slotine, "On partial contraction analysis for coupled nonlinear oscillators," *Biol. Cybern.*, vol. 92, no. 1, pp. 38–53, Jan. 2005.

[17] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice-Hall, 2002.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.

[19] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Learning-based adaptive control via contraction theory," *Submitted to IEEE Control Syst. Lett.*, Mar. 2021.

[20] D. Sun, S. Jha, and C. Fan, "Learning certified control using contraction metric," in *CoRL*, Nov. 2020.

[21] H. Tsukamoto, S.-J. Chung, J.-J. E. Slotine, and C. Fan, "A tutorial on neural contraction metrics for learning-based control with guaranteed stability," in *Submitted to IEEE Conf. Decis. Control*, 2021.