Statistical Contraction Theory for Learning-based Control of Nonlinear Systems—Part 4: Conformal Contraction

Hiroyasu Tsukamoto

Department of Aerospace Engineering
University of Illinois Urbana-Champaign

February 3, 2026

Table of contents

- In the following, $\|\cdot\|$ denotes the Euclidean norm for vectors and the induced 2-norm for matrices, i.e., $\|\cdot\| = \|\cdot\|_2$, unless otherwise noted.
- For a symmetric square matrix $S \in \mathbb{R}^{n\times n}$, we use the notation $S \succ 0$, $S \succeq 0$, $S \prec 0$, and $S \preceq 0$ for the positive definite, positive semi-definite, negative definite, and negative semi-definite matrices, respectively.
- This lecture is based on [1, 2, 3, 4].

Table of contents

Formal Guarantees in Data-Driven Systems

With all the tools we have learned so far, we can establish formal guarantees for data-driven nonlinear systems under uncertainty **with no theoretical black boxes**.

Formal Guarantees in Data-Driven Systems

Let's consider a nonlinear dynamical system given by

$$\dot{\mathbf{x}} = f(\mathbf{x}, t)$$

where f is **unknown**.

Table of contents

Conformal Contraction

Suppose that

- we have somehow learned a prediction model $\hat{f}(\mathbf{x}, t)$
- we have access to exchangeable N trajectory data of $\dot{\mathbf{x}} = f(\mathbf{x}, t),\ \ t \in [0, T]$ not used for learning
- we know a contraction metric with $M(\mathbf{x}, t) \succ 0$ of $\dot{\mathbf{x}} = \hat{f}(\mathbf{x}, t)$, i.e.,

$$\dot{M} + M\frac{\partial \hat{f}}{\partial \mathbf{x}} + \frac{\partial \hat{f}}{\partial \mathbf{x}}M \preceq -2\alpha M \text{ and } \underline{m}\mathbb{I} \preceq M \preceq \overline{m}\mathbb{I}$$

for $\alpha, \underline{m}, \overline{m} > 0$.

Conformal Contraction

We define the distance measured through the contraction metric as

$$d_M(t) = \int_{\mathbf{x}_d(t)}^{\mathbf{x}(t)} \|\Theta(\mathbf{q}(t), t)\delta\mathbf{q}(t)\|$$

with $M = \Theta^\top\Theta$, where $\mathbf{x}_d(t)$ is any desired trajectory that satisfies $\dot{\mathbf{x}}_d = \hat{f}(\mathbf{x}_d, t)$.

Conformal Contraction

Theorem (Conformal Contraction[1,2])

Suppose that a new trajectory of $\dot{\mathbf{x}} = f(\mathbf{x}, t)$ is exchangeable with the N trajectory data. Then we have the following bound with $\delta \in (0, 1)$:

$$\mathbb{P}\left[d_M(t) \leq e^{-\alpha t} d_M(0) + \frac{\sqrt{\overline{m}} q_{1-\delta}}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0, T]\right] \geq 1 - \delta$$

$$q_{1-\delta} = \lceil (1-\delta)(N+1) \rceil\text{-th smallest of } (S_1, \cdots, S_N),\ S_i = \sup_{t \in [0,T]} \left\| \dot{\mathbf{x}}_i - \hat{f}(\mathbf{x}_i, t) \right\|$$

with $\mathbf{x}_i$ representing the N sampled trajectories.

[1] S. Wei, M. Ornik, and H. Tsukamoto. "Conformal contraction for robust nonlinear control with distribution-free uncertainty quantification". In: *IEEE Conference on Decision and Control, to appear.* Dec. 2025. DOI: TBA.

[2] R. A. D'Silva and H. Tsukamoto. "Conformally robust chance-constrained trajectory optimization and control of non-Gaussian stochastic systems". In: *International Conference on Robotics and Automation, to appear.* 2025. DOI: TBA.

Conformal Contraction

Proof: Consider our approximate, differential Lyapunov function $V = \delta \mathbf{q}^\top M \delta \mathbf{q}$ [5], where

- $\dot{\mathbf{q}} = \hat{f}(\mathbf{q}, t) + \mu(f(\mathbf{x}, t) - \hat{f}(\mathbf{x}, t))$
- μ is the parameter of the virtual state $\mathbf{q}(\mu, t)$ with $\mathbf{q}(1, t) = \mathbf{x}$ and $\mathbf{q}(0, t) = \mathbf{x}_d(t)$
- $\delta \dot{\mathbf{q}} = (\partial \hat{f} / \partial \mathbf{q}) \delta \mathbf{q} + (f(\mathbf{x}, t) - \hat{f}(\mathbf{x}, t)) d\mu.$

Conformal Contraction

We have

$$\dot{V} \leq -2\alpha V + 2\sqrt{\overline{m}}\|\Theta(\mathbf{q}, t)\delta\mathbf{q}\| \left\| f(\mathbf{x}, t) - \hat{f}(\mathbf{x}, t) \right\|.$$

Integrating it over μ, we have

$$\dot{d}_M \leq -\alpha d_M + \sqrt{\overline{m}}\|f(\mathbf{x}, t) - \hat{f}(\mathbf{x}, t)\|$$

which gives

$$\begin{aligned}\dot{d}_M(t) &\leq e^{-\alpha t} d_M(0) + \sqrt{\overline{m}} e^{-\alpha t} \int_0^t e^{\alpha\tau} \|f(\mathbf{x}, t) - \hat{f}(\mathbf{x}, t)\| d\tau, \ \forall t \in [0, T] \\ &\leq e^{-\alpha t} d_M(0) + \frac{\sqrt{\overline{m}} \sup_{t\in[0,T]} \left\| f(\mathbf{x}(t), t) - \hat{f}(\mathbf{x}(t), t) \right\|}{\alpha} \left(1 - e^{-\alpha t}\right), \ \forall t \in [0, T].\end{aligned}$$

Conformal Contraction

As we saw earlier, conformal inference gives

$$\sup_{t\in[0,T]} \left\| f(\mathbf{x}(t),t) - \hat{f}(\mathbf{x}(t),t) \right\| \leq q_{1-\delta}$$

at least with probability $1-\delta$, which completes the proof. $\square$

Conformal Contraction

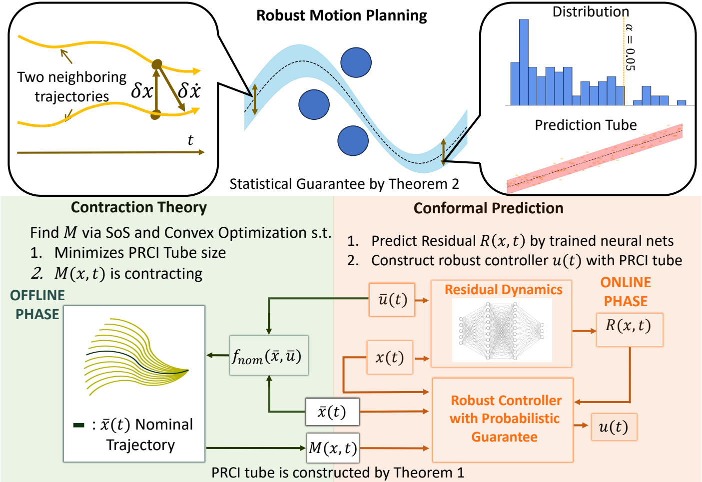


Figure 1: Conformal Contraction [2, 3].

Table of contents

Conformal Lyapunov

Suppose that

- we have somehow learned a prediction model $\hat{f}(\mathbf{x}, t)$
- we have access to exchangeable N trajectory data of $\dot{\mathbf{x}} = f(\mathbf{x}, t),\ \ t \in [0, T]$ not used for learning
- we know a Lyapunov function V of $\dot{\mathbf{x}} = \hat{f}(\mathbf{x}, t)$, i.e.,

$$\dot{V} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial \mathbf{x}}\hat{f}(\mathbf{x}, t) \leq -2\alpha V,\ \ \underline{m}\|\mathbf{x}\|^2 \leq V \leq \overline{m}\|\mathbf{x}\|^2,\ \text{and}\ \|\partial V/\partial \mathbf{x}\| \leq \bar{c}\sqrt{V}$$

for $\alpha, \underline{m}, \overline{m}, \bar{c} > 0$.

Conformal Lyapunov

Theorem (Conformal Lyapunov[3])

Suppose that a new trajectory of $\dot{\mathbf{x}} = f(\mathbf{x}, t)$ is exchangeable with the N trajectory data. Then we have the following bound with $\delta \in (0, 1)$:

$$\mathbb{P}\left[\|\mathbf{x}(t)\| \leq \|\mathbf{x}(0)\| e^{-\alpha t}\sqrt{\frac{\overline{m}}{\underline{m}}} + \frac{1}{\sqrt{\underline{m}}} \cdot \frac{\bar{c} q_{1-\delta}}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0, T]\right] \geq 1 - \delta$$

$$q_{1-\delta} = \lceil (1-\delta)(N+1) \rceil\text{-th smallest of } (S_1, \cdots, S_N),\ S_i = \sup_{t \in [0,T]} \left\|\dot{\mathbf{x}}_i - \hat{f}(\mathbf{x}_i, t)\right\|$$

with $\mathbf{x}_i$ representing the N sampled trajectories.

[3] T.-W. Hsu and H. Tsukamoto. "Statistical guarantees in data-driven nonlinear control: Conformal robustness for stability and safety". In: *IEEE Control Systems Letters* 9 (2025), pp. 997–1002. DOI: 10.1109/LCSYS.2025.3578062.

Conformal Lyapunov

Proof: For our approximate Lyapunov function V, we have

$$\frac{d\sqrt{V}}{dt} \leq -\alpha\sqrt{V} + \bar{c}\|f(\mathbf{x},t) - \hat{f}(\mathbf{x},t)\|$$

which gives

$$\sqrt{V}(\mathbf{x}(t),t) \leq e^{-\alpha t}\sqrt{V}(\mathbf{x}(0),0) + \bar{c}e^{-\alpha t}\int_0^t e^{\alpha\tau}\|f(\mathbf{x},t) - \hat{f}(\mathbf{x},t)\|d\tau,\ \forall t \in [0,T]$$

$$\leq e^{-\alpha t}\sqrt{V}(\mathbf{x}(0),0) + \frac{\bar{c}\sup_{t\in[0,T]}\left\|f(\mathbf{x}(t),t) - \hat{f}(\mathbf{x}(t),t)\right\|}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0,T].$$

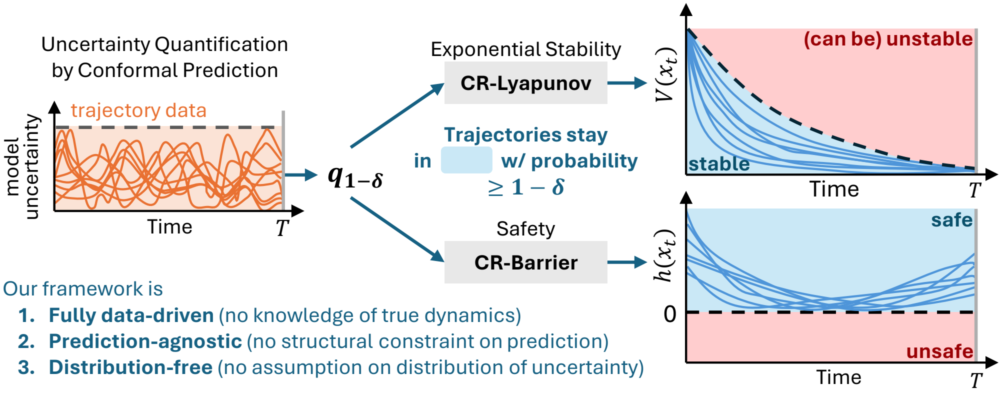
Conformal Lyapunov

As we saw earlier, conformal inference gives

$$\sup_{t\in[0,T]} \left\| f(\mathbf{x}(t),t) - \hat{f}(\mathbf{x}(t),t) \right\| \leq q_{1-\delta}$$

at least with probability $1-\delta$, which completes the proof. $\square$

Conformal Lyapunov



Our framework is

1. **Fully data-driven** (no knowledge of true dynamics)
2. **Prediction-agnostic** (no structural constraint on prediction)
3. **Distribution-free** (no assumption on distribution of uncertainty)

Figure 2: Conformal Robustness [1].

Table of contents

Conformal Barrier

As implied in Figure 2, we have a similar result for barrier functions.

Conformal Barrier

Suppose that

- we have somehow learned a prediction model $\hat{f}(\mathbf{x}, t)$
- we have access to exchangeable N trajectory data of $\dot{\mathbf{x}} = f(\mathbf{x}, t),\ \ t \in [0, T]$ not used for learning
- we know a (zeroing) barrier function h of $\dot{\mathbf{x}} = \hat{f}(\mathbf{x}, t)$, i.e.,

$$\dot{h} = \frac{\partial h}{\partial t} + \frac{\partial h}{\partial \mathbf{x}}\hat{f}(\mathbf{x}, t) \geq -\alpha h \text{ and } \|\partial h/\partial \mathbf{x}\| \leq \bar{h}$$

for $\alpha, \bar{h} > 0$.

Conformal Barrier

Theorem (Conformal Barrier[4])

Suppose that a new trajectory of $\dot{\mathbf{x}} = f(\mathbf{x}, t)$ is exchangeable with the N trajectory data. Then we have the following bound with $\delta \in (0,1)$:

$$\mathbb{P}\left[h(\mathbf{x}(t), t) \geq e^{-\alpha t} h(\mathbf{x}(0), 0) - \frac{\bar{h} q_{1-\delta}}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0, T]\right] \geq 1 - \delta$$

$$q_{1-\delta} = \lceil (1-\delta)(N+1) \rceil\text{-th smallest of } (S_1, \cdots, S_N),\ S_i = \sup_{t \in [0,T]} \left\| \dot{\mathbf{x}}_i - \hat{f}(\mathbf{x}_i, t) \right\|$$

with $\mathbf{x}_i$ representing the N sampled trajectories.

[4] T.-W. Hsu and H. Tsukamoto. "Statistical guarantees in data-driven nonlinear control: Conformal robustness for stability and safety". In: *IEEE Control Systems Letters* 9 (2025), pp. 997–1002. DOI: 10.1109/LCSYS.2025.3578062.

Conformal Barrier

Proof: For our approximate barrier function h, we have

$$\frac{dh}{dt} \geq -\alpha h - \bar{h}\|f(\mathbf{x},t) - \hat{f}(\mathbf{x},t)\|$$

which gives

$$h(\mathbf{x}(t),t) \geq e^{-\alpha t}h(\mathbf{x}(0),0) - \bar{h}e^{-\alpha t}\int_0^t e^{\alpha\tau}\|f(\mathbf{x},t) - \hat{f}(\mathbf{x},t)\|d\tau,\ \forall t \in [0,T]$$

$$\geq e^{-\alpha t}h(\mathbf{x}(0),0) - \frac{\bar{h}\sup_{t\in[0,T]}\left\|f(\mathbf{x}(t),t) - \hat{f}(\mathbf{x}(t),t)\right\|}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0,T].$$

Conformal Barrier

As we saw earlier, conformal inference gives

$$\sup_{t\in[0,T]} \left\| f(\mathbf{x}(t),t) - \hat{f}(\mathbf{x}(t),t) \right\| \leq q_{1-\delta}$$

at least with probability $1-\delta$, which completes the proof. $\square$

Conformal Contraction, Lyapunov, & Barrier

Note that all the ideas discussed so far can be used explicitly for controller design [1, 2, 3, 4].

Table of contents

Formal Guarantees in Data-Driven Systems

In summary, we need the following three layers for obtaining guarantees in data-driven systems:

(1) Derive robust guarantees in nominal systems (e.g., Lyapunov, barrier, contraction, etc.)

(2) Offset discrepancies between nominal and off-nominal systems with known structures (e.g., adaptive control, reinforcement learning, online learning with known hypothesis functions, etc.)

(3) Quantify discrepancies between off-nominal systems with known structures and unknown structures (e.g., structural learning, PAC-Bayes, conformal inference, etc.)

Formal Guarantees in Data-Driven Systems

Just to reiterate, **we made no assumptions about the structure or distribution of the uncertainties v and w** in deriving the **finite-sample** guarantees.

Formal Guarantees in Data-Driven Systems

Exchangeability can be a restrictive assumption here, but it is strictly weaker than the i.i.d. assumption and can be relaxed further to accommodate nonexchangeable data [6, 7].

Formal Guarantees in Data-Driven Systems

However, the guarantees are valid **only in a probabilistic and finite-time sense**, making it strictly weaker than those derived under structural or distributional assumptions.

While disproving this statement is part of our ongoing research efforts, such a limitation may be inherent in any control + machine learning approach

Formal Guarantees in Data-Driven Systems

Therefore, at least for now, particularly in safety-critical applications in autonomous systems, we should

- exploit structural and distributional priors of systems as much as possible to establish nominal guarantees, e.g., using contraction theory, and
- address the remaining unstructured uncertainties through data to establish off-nominal statistical guarantees, e.g., using conformal inference.

Table of contents

Nonlinear Estimation with Conformal Inference

Note that the topics we'll be discussing from now on are part of ongoing research. I will no longer be able to provide definite answers.

I expect you to think critically and creatively, coming up with your own ideas.

Nonlinear Estimation with Conformal Inference

Let's go back to the example of state estimation in the following perturbed nonlinear dynamical systems:

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{v}(\mathbf{p}(t), t)$$
$$\mathbf{y}(t) = h(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{w}(\mathbf{p}(t), t)$$

where f **and** h **are known, but** $\mathbf{v}$ **and** $\mathbf{w}$ **are unknown**. The parameters $\mathbf{p}(t)$ are assumed to be accessible, although this assumption can be relaxed if necessary.

Nonlinear Estimation with Conformal Inference

Instead of assuming that the uncertainties $\mathbf{w}(\mathbf{p}, t)$ and $\mathbf{v}(\mathbf{p}, t)$ are bounded, let us construct uncertainty predictors using the following loss functions:

$$\mathcal{L}(\theta_{\mathbf{v}}) = \sum_{j=1}^{M} \|\hat{\mathbf{v}}(\mathbf{p}_j, t_j; \theta_{\mathbf{v}}) - \mathbf{v}_j\|^2$$

$$\mathcal{L}(\theta_{\mathbf{w}}) = \sum_{j=1}^{M} \|\hat{\mathbf{w}}(\mathbf{p}_j, t_j; \theta_{\mathbf{w}}) - \mathbf{w}_j\|^2$$

where $\{(\mathbf{p}_j, t_j, \mathbf{v}_j)\}_{j=1}^{M}$ and $\{(\mathbf{p}_j, t_j, \mathbf{w}_j)\}_{j=1}^{M}$ with $\mathbf{v}_j = \mathbf{v}(\mathbf{p}_j, t_j)$ and $\mathbf{w}_j = \mathbf{w}(\mathbf{p}_j, t_j)$ are training data samples.

Nonlinear Estimation with Conformal Inference

Here, we can use any machine learning methods for designing $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$, e.g., with gradient descent, but **the parameters $\theta_{\mathbf{v}}$, $\theta_{\mathbf{w}}$ can be empirical and need not be optimal**.

Nonlinear Estimation with Conformal Inference

We could design a state estimator as follows:

$$\dot{\hat{\mathbf{x}}} = f(\hat{\mathbf{x}}, \mathbf{u}, t) + \hat{\mathbf{v}}(\mathbf{p}, t; \theta_{\mathbf{v}}) + L(\hat{\mathbf{x}}, t)\left(\mathbf{y} - h(\hat{\mathbf{x}}, \mathbf{u}, t) - \hat{\mathbf{w}}(\mathbf{p}, t; \theta_{\mathbf{w}})\right)$$

where our intention is to approximately cancel out the uncertainties with our empirical predictions.

Nonlinear Estimation with Conformal Inference

A virtual system can be designed as follows:

$$\begin{aligned}\dot{\mathbf{q}}(\mu,t) = {} & f(\mathbf{q}(\mu,t),\mathbf{u},t) + (1-\mu)\mathbf{v}(\mathbf{p},t) + \mu\hat{\mathbf{v}}(\mathbf{p},t;\theta_{\mathbf{v}}) \\ & + L(\hat{\mathbf{x}},t)(h(\mathbf{x},\mathbf{u},t) - h(\mathbf{q}(\mu,t),\mathbf{u},t)) + \mu L(\hat{\mathbf{x}},t)(\mathbf{w}(\mathbf{p},t) - \hat{\mathbf{w}}(\mathbf{p},t;\theta_{\mathbf{w}}))\end{aligned}$$

where $\mathbf{q}(\mu,t)$ is a smooth path connecting $\mathbf{x}$ and $\hat{\mathbf{x}}$, parameterized by $\mu \in [0,1]$ with

- $\mathbf{q}(0,t) = \mathbf{x}(t)$
- $\mathbf{q}(1,t) = \hat{\mathbf{x}}(t)$.

Nonlinear Estimation with Conformal Inference

The differential dynamics is given as

$$\frac{\partial \dot{\mathbf{q}}}{\partial \mu} = \left(\frac{\partial f(\mathbf{q},\mathbf{u},t)}{\partial \mathbf{q}} - L(\hat{\mathbf{x}},t)\frac{\partial h(\mathbf{q},\mathbf{u},t)}{\partial \mathbf{q}}\right)\frac{\partial \mathbf{q}}{\partial \mu} + (\hat{\mathbf{v}}(\mathbf{p},t;\theta_{\mathbf{v}}) - \mathbf{v}(\mathbf{p},t)) - L(\hat{\mathbf{x}},t)(\hat{\mathbf{w}}(\mathbf{p},t;\theta_{\mathbf{w}}) - \mathbf{w}(\mathbf{p},t)).$$

As a contraction condition, we assume that i.e., $\exists M(\mathbf{q},t) \succ 0$ and $\alpha > 0$ s.t. the following holds for $\forall \mathbf{q}, \hat{\mathbf{x}}, \mathbf{u}, t$:

$$\dot{M} + M\left(\frac{\partial f}{\partial \mathbf{q}} - L(\hat{\mathbf{x}},t)\frac{\partial h}{\partial \mathbf{q}}\right) + \left(\frac{\partial f}{\partial \mathbf{q}} - L(\hat{\mathbf{x}},t)\frac{\partial h}{\partial \mathbf{q}}\right)^\top M \preceq -2\alpha M$$

where the arguments of M, f, and h are omitted here for simplicity.

Nonlinear Estimation with Conformal Inference

Let $\Delta\mathbf{v}(\mathbf{p},t)=\hat{\mathbf{v}}(\mathbf{p},t;\theta_{\mathbf{v}})-\mathbf{v}(\mathbf{p},t)$ and $\Delta\mathbf{w}(\mathbf{p},t)=\hat{\mathbf{w}}(\mathbf{p},t;\theta_{\mathbf{w}})-\mathbf{w}(\mathbf{p},t)$ for notational simplicity. We get

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial\mathbf{q}}{\partial\mu}^{\top}M\frac{\partial\mathbf{q}}{\partial\mu}\right) &\leq -2\alpha\frac{\partial\mathbf{q}}{\partial\mu}^{\top}M\frac{\partial\mathbf{q}}{\partial\mu}+2\frac{\partial\mathbf{q}}{\partial\mu}^{\top}M(\Delta\mathbf{v}(\mathbf{p},t)-L(\hat{\mathbf{x}},t)\Delta\mathbf{w}(\mathbf{p},t))\\ &\leq -2\alpha\frac{\partial\mathbf{q}}{\partial\mu}^{\top}M\frac{\partial\mathbf{q}}{\partial\mu}+2\sqrt{\overline{m}}\left\|\Delta\mathbf{v}(\mathbf{p},t)-L(\hat{\mathbf{x}},t)\Delta\mathbf{w}(\mathbf{p},t)\right\|\left\|\Theta\frac{\partial\mathbf{q}}{\partial\mu}\right\|\end{aligned}$$

and then

$$\frac{d}{dt}\left\|\Theta\frac{\partial\mathbf{q}}{\partial\mu}\right\|\leq-\alpha\left\|\Theta\frac{\partial\mathbf{q}}{\partial\mu}\right\|+\sqrt{\overline{m}}\left\|\Delta\mathbf{v}(\mathbf{p},t)-L(\hat{\mathbf{x}},t)\Delta\mathbf{w}(\mathbf{p},t)\right\|.$$

Nonlinear Estimation with Conformal Inference

Given exchangeable calibration trajectory data samples $\{(\mathbf{p}_i(t), \hat{\mathbf{x}}_i(t))\}_{i=1}^N$ for $t \in [0, T]$, we may define the prediction nonconformity score as

$$S_i = \sup_{t\in[0,T]} \|\Delta\mathbf{v}(\mathbf{p}_i(t), t) - L(\hat{\mathbf{x}}_i(t), t)\Delta\mathbf{w}(\mathbf{p}_i(t), t)\|$$

which implies

$$\mathbb{P}\left[\sup_{t\in[0,T]} \|\Delta\mathbf{v}(\mathbf{p}(t), t) - L(\hat{\mathbf{x}}(t), t)\Delta\mathbf{w}(\mathbf{p}(t), t)\| \leq q_{1-\delta}\right] \geq 1-\delta$$

where

$$q_{1-\delta} = \lceil(1-\delta)(N+1)\rceil\text{-}\textit{th smallest of } (S_1, \cdots, S_N).$$

Nonlinear Estimation with Conformal Inference

We thus get

$$\mathbb{P}\left[d_M(t) \leq d_M(0)e^{-\alpha t} + \frac{\sqrt{\overline{m}}q_{1-\delta}}{\alpha}\left(1 - e^{-\alpha t}\right),\ \forall t \in [0, T]\right] \geq 1 - \delta$$

where

$$d_M(t) = \int_0^1 \left\|\Theta(\mathbf{q}(\mu, t), t)\frac{\partial \mathbf{q}(\mu, t)}{\partial \mu}\right\| d\mu$$

represents the state estimation error measured through the contraction metric.

Nonlinear Estimation with Conformal Inference

More intuitively, we have

$$\mathbb{P}\left[\|\hat{\mathbf{x}}(t)-\mathbf{x}(t)\| \leq \frac{d_M(0)e^{-\alpha t}}{\sqrt{\underline{m}}} + \frac{q_{1-\delta}}{\alpha}\sqrt{\frac{\overline{m}}{\underline{m}}}\left(1-e^{-\alpha t}\right),\ \forall t \in [0, T]\right] \geq 1-\delta.$$

Reference

[1] T.-W. Hsu and H. Tsukamoto. "Statistical guarantees in data-driven nonlinear control: Conformal robustness for stability and safety". In: *IEEE Control Systems Letters* 9 (2025), pp. 997–1002. DOI: 10.1109/LCSYS.2025.3578062.

[2] S. Wei, M. Ornik, and H. Tsukamoto. "Conformal contraction for robust nonlinear control with distribution-free uncertainty quantification". In: *IEEE Conference on Decision and Control, to appear*. Dec. 2025. DOI: TBA.

[3] R. A. D'Silva and H. Tsukamoto. "Conformally robust chance-constrained trajectory optimization and control of non-Gaussian stochastic systems". In: *International Conference on Robotics and Automation, to appear*. 2025. DOI: TBA.

Reference

[4] K. Hirano and H. Tsukamoto. "Conformal Koopman for nonlinear control with statistically robust tracking bounds: Theory and real-world validation". In: *International Conference on Robotics and Automation, to appear*. 2025. DOI: TBA.

[5] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine. "Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview". In: *Annual Reviews in Control* 52 (2021), pp. 135–169. ISSN: 1367-5788. DOI: 10.1016/S0005-1098(98)00019-3.

[6] R. J. Tibshirani, R. Foygel Barber, E. Candes, and A. Ramdas. "Conformal prediction under covariate shift". In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates Inc., 2019.

[7] I. Gibbs and E. Candes. "Adaptive conformal inference under distribution shift". In: *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., 2021, pp. 1660–1672.