Epistemically Aware Predictive Visuomotor Control

Antoine P. Leeman
ETH Zurich
aleeman@ethz.ch

Shuyu Zhan
Georgia Institute of Technology
szhan45@gatech.edu

Melanie N. Zeilinger*
ETH Zurich
mzeilinger@ethz.ch

Glen Chou*
Georgia Institute of Technology
chou@gatech.edu

Abstract

We design a safe visuomotor controller when perception is *epistemically ambiguous* due to occlusions, textureless regions, or distribution shift. Our approach is to replace point state estimates from images with credal (uncertainty) sets that a controller can leverage in real-time. To this end, we learn credal sets: a reduced measurement together with a state-dependent confidence set. We then design an ambiguity-aware output-feedback controller based on system level synthesis that treats it as a source of measurement ambiguity, propagates it through closed-loop responses, and enforces safety via uncertainty propagation and constraint tightening. The resulting pipeline links perception to control through sets, yielding certificates conditioned on coverage and naturally encouraging reducing epistemic uncertainty behavior. An efficient sequential convex programming implementation with Riccati recursions enables real-time control. In high-fidelity vision-based simulations of a car and a quadcopter, our method achieves safe trajectories in seconds while avoiding overconfidence under epistemic shift.

1 Introduction

Control from high-dimensional visual inputs remains brittle under real-world conditions, even as autonomous systems are increasingly expected to interact safely and reliably with their surroundings. Yet vision offers rich, scalable scene information. Mapping images to the state is inherently epistemically ambiguous: occlusions, textureless regions, and distribution shift can render the true state indistinguishable from the available visual evidence. Standard approaches treat perception errors as Gaussian noise, conflating aleatoric variability with epistemic uncertainty, and often yield overconfident controllers that violate safety constraints [Kalman, 1960, Stengel, 1994, Kendall and Gal, 2017].

A growing body of work in machine learning formalizes epistemic uncertainty via imprecise set-valued, ambiguity-aware models, such as credal (uncertainty) sets, evidential networks, and conformal prediction. These approaches construct sets of plausible outcomes with calibrated coverage [Walley, 1991, Sensoy et al., 2018, Vovk et al., 2005, Angelopoulos and Bates, 2021]. Such ideas are natural for control: rather than a point estimate, perception should output a credal set, i.e., a low-dimensional measurement with an explicit state-dependent confidence set [Combastel, 2005, Romano et al., 2020].

In this paper, we take a two-step approach for predictive visuomotor control: first, we learn credal sets from pixels with empirical coverage, even under distribution shift. Second, we design an ambiguity-aware output-feedback controller based on system level synthesis (SLS) that propagates these credal sets through closed-loop dynamics and enforces safety via reachable sets and robust (safe) prediction

*Equal supervision

1st Workshop on Epistemic Intelligence in Machine Learning (EIML 2025) .

[Chen and Anderson, 2019, Sieber et al., 2021, Wang et al., 2019, Goulart, 2006, Dean and Recht, 2021, Leeman et al., 2025a]. Together, this perception-to-control pipeline enables safe, real-time visuomotor control that accounts for epistemic ambiguity.

Contributions.

1. We propose a control-theoretic loss and leverage visual foundation models to learn a reduced image-to-measurement projection with a state-dependent uncertainty bound, yielding set-valued observations that are empirically calibrated via a convex optimization.
2. We develop an ambiguity-aware output-feedback design with SLS that treats the credal set as a source of measurement ambiguity, propagates it through the closed-loop responses, and enforces safety via constraint tightening.
3. We provide a new scalable solver that leverages efficient numerical optimization techniques: sequential convex programming (SCP) and efficient Riccati recursions. Robust performance is assessed on high-fidelity simulation experiments of a car and a quadcopter, computing safe trajectories in seconds.

Related Work

Epistemic uncertainty in learning. In machine learning, epistemic uncertainty is modeled via Bayesian and nonparametric approaches such as Gaussian processes (GPs) [Seeger, 2004] and neural processes (NPs) [Garnelo et al., 2018], as well as approximate Bayesian methods for neural networks, including ensembles [Lakshminarayanan et al., 2017], Bayesian neural networks [Blundell et al., 2015], and dropout-based approximations [Gal and Ghahramani, 2016]. While effective in-distribution, such distributional surrogates often fail to provide reliable coverage under distribution shift [Ovadia et al., 2019]. A complementary line of work emphasizes imprecise (i.e., distributionally ambiguous) probability models, which represent sets of plausible distributions or outcomes. Examples include credal models [Walley, 1991], evidential deep learning [Sensoy et al., 2018], and conformal prediction [Vovk et al., 2005, Angelopoulos and Bates, 2021].

Set-valued perception. Set-valued outputs have been explored in regression via conformalized image classifiers [Romano et al., 2020]. In control, interval observers and set-membership estimation methods [Combastel, 2005] have long been used for robust filtering and fault detection. Recent work also considers learning-based set predictors for safe perception and planning [Mirman et al., 2018]. Our formulation bridges these traditions: we learn a state-dependent credal set from images that provides a calibrated error set around a projection, in a form that is directly compatible with set-based robust control and SLS.

Control under ambiguity. Classical robust and stochastic control frameworks assume either bounded disturbances [Khalil et al., 1996] or Gaussian noise models [Mayne, 2014]. More recent advances in distributionally robust optimization [Wiesemann et al., 2014, Mirman et al., 2018] extend robustness to ambiguity (credal) sets over probability distributions. In control, SLS [Wang et al., 2019] provides a flexible parameterization for propagating structured uncertainty through closed-loop response maps. Connections between epistemic uncertainty in perception and robust control are beginning to appear [Dean and Recht, 2021, Chou et al., 2022, Trisovic et al., 2025], but remain limited. Our work connects these threads by treating learned credal sets as an explicit ambiguity source in output-feedback SLS, yielding scalable, certificate-backed control directly from pixels.

2 Problem Setup

We consider discrete-time nonlinear dynamics

$$x_{k+1} = f(x_k, u_k, w_k), \qquad x_k \in \mathbb{R}^n,\ u_k \in \mathbb{R}^m,\ w_k \in \mathbb{R}^{n_w}, \tag{1}$$

where x_k is the state (e.g., pose and velocities in a fixed world frame), u_k is the control input (e.g., forces and torques), and $w_k \in \mathcal{W}$ is the bounded process noise (e.g., uncertain aerodynamic coefficients or wind). A camera provides an RGB image $y_k \in \mathbb{R}^{H\times W\times 3}$ whose appearance is induced by the robot state via the rendering map $y_k = h(x_k)$. We seek an output-feedback controller that operates on a reduced measurement y_k^r and an associated credal set. To this end, we first design a reduced observation model $h_r(x_k)$ with a calibrated state-dependent credal set $\mathcal{V}(x_k)$ around y_k^r, i.e.,

$$y_k^r = h_r(x_k) + v_k, \qquad v_k \in \mathcal{V}(x_k) := \{\, v : \|v\|_\infty \le e(x_k) \,\}, \tag{2}$$

where v_k captures epistemic perception ambiguity; its radius $e(x_k)$ is state dependent, data-calibrated, and grows under occlusions or textureless regions.

Then, we synthesize a causal policy $u_k = \pi_k(y_0^r, \ldots, y_k^r)$ such that $(x_k, u_k) \in \mathcal{S}$ for all $w_k \in \mathcal{W}$ and $v_k \in \mathcal{V}(x_k)$, where $\mathcal{S}$ denotes the state–input safety set (e.g., input box and obstacle-avoidance constraints).

3 Credal Sets from Pixels

In this section, we detail how we obtain the credal set. First, we learn a reduced observation y_k^r from images y_k. Second, we learn a calibrated state-dependent credal set $\mathcal{V}(x_k)$ around y_k^r.

3.1 Low-Dimensional Observation

Given a color image $y \in \mathbb{R}^{H\times W\times 3}$, we freeze a visual encoder $\phi : \mathbb{R}^{H\times W\times 3} \to \mathbb{R}^d$ (e.g., DINOv2 [Oquab et al., 2023]) to extract features $\phi(y)$. A learned projection $p : \mathbb{R}^d \to \mathbb{R}^r$ maps these features to a reduced observation $y^r = p(\phi(y))$, where $r \ll d$ and r is chosen to match a target linear observation model $h_r(x) = Cx,\ C \in \mathbb{R}^{r\times n}$. Specifically, over the training dataset $\mathbf{D} = (\{x_i, y_i\})$ of state-image pairs, we jointly learn $p(\cdot)$ and $h_r(\cdot)$ to minimize the loss

$$\ell(\theta_p, \theta_h) = \|p(\phi(y_i); \theta_p) - h_r(x_i; \theta_h)\| - \lambda\sigma_{\min}(\mathcal{O}(x_i)) \tag{3}$$

for some $\lambda > 0$, and where θ_p and θ_h are neural-network weights. The second term encourages a larger smallest singular value of the observability matrix $\mathcal{O}$ [Grossman, 1999, Ch. 3]. This promotes higher Jacobian rank and hence better recoverability of x from outputs, thus the projection learns to ignore visual distractors.

3.2 State-Dependent Credal Set

Rather than treating y^r as a noisy point estimate of $h_r(x)$, we model a state-dependent credal set

$$v_k \in \mathcal{V}(x_k) \ := \ \{\, v : \|v\|_\infty \le e(x_k) \,\}, \tag{4}$$

where $e : \mathbb{R}^n \to \mathbb{R}_{\ge 0}$ is a bounding function that captures epistemic uncertainty. We interpret $\mathcal{V}(x)$ as a credal set: the collection of reduced observations consistent with the image evidence at state x. Unlike Gaussian surrogates that can be overconfident under occlusions or shift, the credal set $\mathcal{V}(x)$ models epistemic ambiguity directly, and the calibrated radius $e(x)$ sets is used for the safety constraint tightenings (Section 4.2).

To calibrate this set, we upper-bound the reduced-observation error

$$r_k \ := \ \left\|y_k^r - h_r(x_k)\right\|_\infty \ = \ \|v_k\|_\infty, \tag{5}$$

and learn $e_\beta(\cdot)$ so that $r_k \le e_\beta(x_k)$ and with high empirical coverage.

The following convex optimization problem is used to learn the smallest smooth, state-dependent radius that upper-bounds the empirical residuals r_k. Hence, we fit $e_\beta(\cdot)$ on calibration data $\mathbf{D}_{\text{tr}} = \{(x_i, y_i)\}$ by solving the convex program

$$\begin{aligned} \min_{\beta,\, \xi_i \ge 0} \quad & \sum_i \left(\beta^\top m(x_i)\right) + \gamma\|\beta\|_2^2 + \mu \sum_i \xi_i \\ \text{s.t.} \quad & \beta^\top m(x_i) \ \ge \ r_i - \xi_i, \qquad r_i \ = \ \left\|p(\phi(y_i)) - h_r(x_i)\right\|_\infty \quad \forall i, \end{aligned}$$

where $\gamma, \mu \ge 0$ trade off smoothness and outliers, $m(x) \in \mathbb{R}^M$ is a fixed monomial basis and $e_\beta(x) \ := \ \beta^\top m(x)$ is a scalar polynomial with coefficients $\beta \in \mathbb{R}^M$.

Remark 1. *This polynomial construction yields a differentiable, state-dependent credal set radius that is cheap to evaluate inside constrained trajectory optimizers used in the next section. By comparison, Lipschitz interpolation yields noisy gradients, which complicate coupling with constrained trajectory optimization solvers [Knuth et al., 2021].*

4 Ambiguity-Aware Visuomotor Control

With a calibrated credal set in hand, we now treat perception uncertainty as epistemic ambiguity in the measurement and synthesize a dynamic output-feedback controller that propagates this ambiguity (4) through the dynamics (1).

4.1 Perception-based controller

To construct the perception-based controller, we leverage a nominal (without process noise or measurement ambiguity) prediction of the nonlinear dynamics $(\bar{x}_k, \bar{u}_k)$. We use the SLS framework to design the perception-based controller [Dean and Recht, 2021] via a parametrization of the closed-loop behavior via response maps from disturbances to states and inputs:

$$x = \bar{x} + \mathbf{\Phi}^{\mathrm{xw}} w + \mathbf{\Phi}^{\mathrm{xv}} v, \qquad u = \bar{u} + \mathbf{\Phi}^{\mathrm{uw}} w + \mathbf{\Phi}^{\mathrm{uv}} v.$$

The v-uncertainty explicitly accounts for measurement ambiguity, i.e., all $v_k \in \mathcal{V}(\bar{x}_k)$ along the nonlinear nominal trajectory. We enforce safety via explicit tightenings from induced norms of the response maps. This extends [Leeman et al., 2025b], which considers only process disturbances (no measurement noise) and therefore does not address measurement ambiguity. See Appendix A for details on the construction of the matrices $\mathbf{\Phi}^{\mathrm{xw}}$, $\mathbf{\Phi}^{\mathrm{xv}}$, $\mathbf{\Phi}^{\mathrm{uw}}$, and $\mathbf{\Phi}^{\mathrm{uv}}$.

4.2 Safety Certificate with Calibrated Credal Sets

Let $\mathcal{S}$ denote state and input safety constraints encoded as inequalities. If the credal set achieves high coverage on held-out data (Section 3), then the designed controller must satisfy

$$(x_k, u_k) \in \mathcal{S} \quad \text{for all } w_k \in \mathcal{W},\ v_k \in \mathcal{V}(\bar{x}_k), \tag{6}$$

i.e., the closed-loop state and input remain safe for all process noise w_k and measurement ambiguity v_k in the credal set $\mathcal{V}(\bar{x}_k)$ along the nominal trajectory.

We enforce the safety constraints (6) written as $A_x x_k \leq b_x$ and $A_u u_k \leq b_u$ using two steps.[2]

1. Evaluate the worst-case closed-loop deviation, also called tube, from the response maps:

   $$\delta^x := \|\mathbf{\Phi}^{\mathrm{xw}}\|_{\infty\to\infty}\ \bar{w} + \|\mathbf{\Phi}^{\mathrm{xv}}\|_{\infty\to\infty}\ \bar{v}, \qquad \delta^u := \|\mathbf{\Phi}^{\mathrm{uw}}\|_{\infty\to\infty}\ \bar{w} + \|\mathbf{\Phi}^{\mathrm{uv}}\|_{\infty\to\infty}\ \bar{v},$$

   with $\bar{v}$ the worst-case measurement-ambiguity radius induced by $\mathcal{V}(x)$ within the tube δ^x and $\bar{w}$ the worst-case process-noise radius from $\mathcal{W}$.[3]
2. Tighten the nominal constraints. Since $\|x_k - \bar{x}_k\|_\infty \leq \delta^x$ and $\|u_k - \bar{u}_k\|_\infty \leq \delta^u$, it suffices to impose on the nominal plan, for all k,

   $$A_x \bar{x}_k \leq b_x - |A_x|\,\mathbf{1}\,\delta^x, \qquad A_u \bar{u}_k \leq b_u - |A_u|\,\mathbf{1}\,\delta^u. \tag{7}$$

Hence, imposing the constraints (7) for all k, guarantees that the safety constraints are satisfied (6) regardless of the process noise w and measurement noise $v(x)$. See Appendix B for additional details on the credal set closed-loop propagation and corresponding constraint tightening.

4.3 Sequential implementation

To implement the perception-based controller, we leverage the nonlinear nature of the dynamics via a SCP algorithm that repeatedly approximates the nonlinear program with linear time-varying surrogates. For each linear surrogate, we leverage an efficient linear SLS solver [Leeman et al., 2024, 2025a] originally developed for systems without measurement errors, augmented with an SLS-type Kalman filter [Wang et al., 2015] (see Appendix C for details). The overall algorithm is as follows:

1. Linearize the dynamics f and quadratically approximate the cost around the nominal trajectory $(\bar{x}, \bar{u})$ to obtain a linear time-varying subproblem.
2. Solve for nominal trajectory updates and SLS responses Φ via SLS numerical solver [Leeman et al., 2024] and Kalman filtering [Wang et al., 2015].
3. Update $(\bar{x}, \bar{u})$ and repeat.

5 Experiments

In this section, we deploy the ambiguity-aware output-feedback SLS controller. We analyze the closed-loop behavior on two visuomotor tasks: (i) a quadcopter performing safe waypoint tracking

[2] For simplicity, we present linear constraints; nonlinear constraints (e.g., obstacle avoidance) are handled via sequential linearizations, similarly in Section 4.3 and [Zhan et al., 2025].

[3] $\|M\|_{\infty\to\infty} := \sup_{\|z\|_\infty \leq 1} \|Mz\|_\infty$ is the induced $\ell_\infty \to \ell_\infty$ operator norm (equal to the maximum absolute row-sum of M). For block-lower-triangular response maps Φ, we apply this norm to the finite-horizon block matrix. See Appendix B for details.

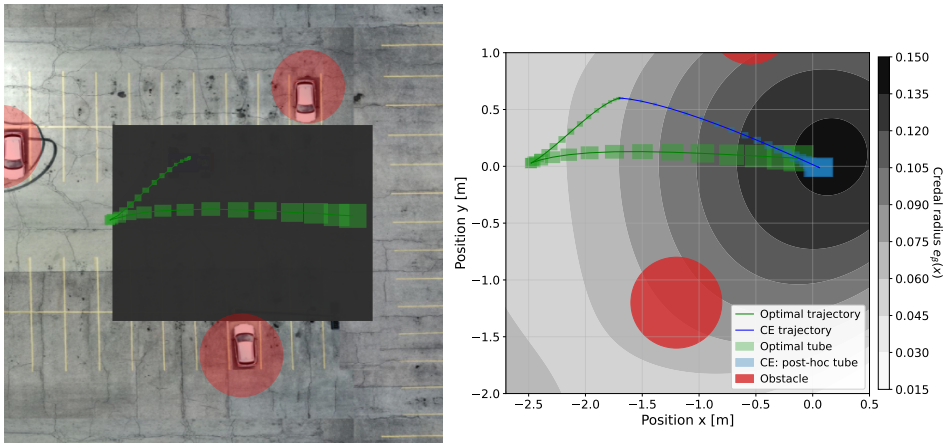

Figure 1: Car in an occluded parking lot with a top-down camera. Left: raw overhead camera image; the large black rectangle is an artificial occluder. Right: zoomed-in top-down view of the workspace. Grayscale shows the learned credal radius for the reduced observation (darker = higher epistemic uncertainty/occlusion). Red disks are obstacles. The ambiguity-aware controller tightens constraints using the propagated set and yields a collision-free trajectory. In contrast, the certainty-equivalent trajectory stays longer in the high-uncertainty occluded region.

in a room with obstacles, and (ii) a car safely navigating to a goal while avoiding obstacles with occlusions using PyBullet [Coumans and Bai, 2016–2021]. All experiments use color images at a resolution of 512×512.

As a baseline, we use a certainty-equivalent (CE) controller that plans from point estimates and does not use tubes during planning; *post hoc* tubes are computed and overlaid only after planning (no tube–trajectory coupling), as with a fixed iLQG-style controller [Todorov and Li, 2005].

5.1 Car with overhead camera

In Fig. 1, the objective is to reach the goal while avoiding obstacles and satisfying state and input constraints, while minimizing the propagated epistemic uncertainty (credal tube volume) along the trajectory, within the horizon $N = 30$.

5.1.1 Perception calibration

We first visualize the learned state-dependent polynomial overbound $e_\beta(x)$ as a grayscale heatmap over the (p_x, p_y) workspace. Crucially, the images are corrupted with partial occlusions (black rectangle) between the initial and goal positions during training. The bound naturally inflates in occluded regions, where the car's position is ambiguous, and shrinks in open areas.

To quantify how well the learned measurement-error bound $\hat{e}(x)$ captures the true perception residuals, we evaluate its empirical coverage on a held-out dataset. For each sample $\{x_i, y_i\}$, we compute the residual $r_i = \|p(\phi(y_i)) - h_r(x_i)\|_\infty$ and report the fraction of points whose residual lies below the predicted bound, $\text{Coverage} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{r_i \le \hat{e}(x_i)\}$. Coverage is evaluated on 500 data points, achieving 98.8%.

5.1.2 Ambiguity-aware control

The optimal car trajectory safely reaches the goal while avoiding obstacles. At each step, the controller uses the propagated credal set to tighten constraints, steering away from high-uncertainty regions and seeking viewpoints that reduce epistemic ambiguity. This behavior is induced by a tube-volume penalty in the sequential trajectory optimization (Section 4.3). In Fig. 1, the path bends away from the dark, high-uncertainty lobe, and the tubes shrink as the trajectory enters better-observed areas, illustrating how credal sets directly shape the optimum. Additional closed-loop rollouts appear in Appendix D (Figure 3).

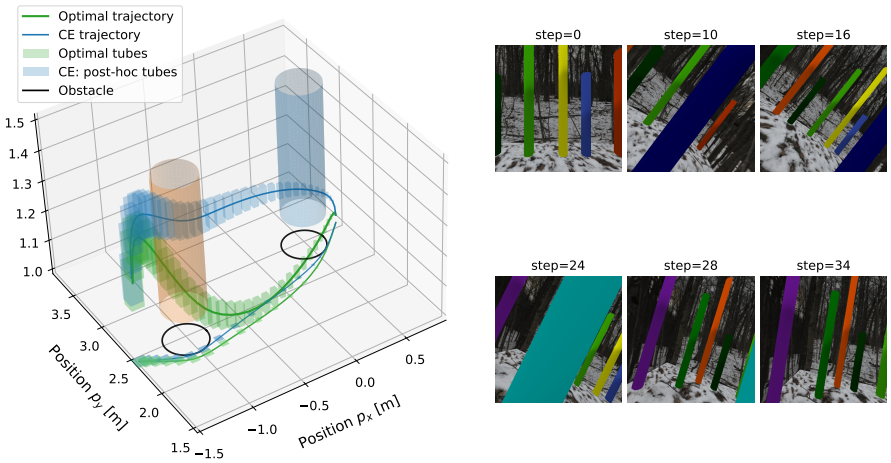

Figure 2: Quadcopter with onboard camera showing safely guaranteed obstacle avoidance despite epistemic uncertainty in perception.

In contrast, the CE baseline (defined above) lingers longer in the dark, high-uncertainty region because tubes are only computed *post hoc*, making it suboptimal for our objective of reaching the goal while minimizing epistemic uncertainty.

5.2 10D quadcopter with onboard camera

We consider a quadcopter model with 10 states and 4 inputs with horizon $N = 35$. We operate it in a cluttered room containing cylindrical obstacles and wall-mounted texture patches. The objective is safe waypoint tracking along a 3D path using only the onboard camera. In Fig. 2, we show that the proposed method scales to large-scale dynamics with camera sensors. The predicted optimal trajectory safely avoids the obstacles despite uncertainty in perception. By contrast, the CE baseline (defined above) collides with obstacles because tubes are computed only *post hoc* and thus do not inform collision avoidance or uncertainty minimization. The sequential algorithm from Section 4.3 computes a safe trajectory and controller in 19.05 [s] on a 2021 MacBook Pro (M1), i.e., faster than the time needed to execute the trajectory.

6 Discussion and Limitations

Our results highlight that epistemic ambiguity in perception can be explicitly modeled and propagated through control via optimization, in seconds. By combining observation sets with output-feedback SLS, we obtain a scalable visuomotor predictive controller with both calibration and tractable safety certificates. Yet, several open challenges remain:

- Tighter uncertainty sets (e.g., neural certificates, adaptive bounds, or conformal prediction) could improve performance. Future work could focus on disentangling aleatoric sensor noise from epistemic model mismatch, so that exploration and learning target only the reducible epistemic component.
- When entire regions of the state are unobservable, our method can become overly cautious.

By producing calibrated, ambiguity-aware observations, this proposed framework provides a principled way to link perception uncertainty to downstream decision-making.

7 Conclusion

We introduced a perception-to-control pipeline that learns credal sets, i.e., low-dimensional measurements with calibrated, state-dependent confidence sets, and propagates them through an ambiguity-aware output-feedback controller via SLS. This approach provides safety certificates, and enables real-time execution across visuomotor tasks. Our results suggest that explicitly modeling epistemic ambiguity, rather than collapsing it into noise, is a practical and scalable path toward safe visuomotor control. Overall, this work suggests that embedding epistemic intelligence into visuomotor pipelines is not only feasible but also essential for safe, trustworthy AI-driven control.

Acknowledgments and Disclosure of Funding

This work was supported by the European Space Agency OSIP 4000133352 and by an ETH Zurich Doc.Mobility Fellowship.

References

James Anderson, John C Doyle, Steven H Low, and Nikolai Matni. System level synthesis. *Annual Reviews in Control*, 47:364–393, 2019.

Anastasios N Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*, 2021.

Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight uncertainty in neural network. In *International conference on machine learning*, pages 1613–1622. PMLR, 2015.

Yuxiao Chen and James Anderson. System level synthesis with state and input constraints. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 5258–5263. IEEE, 2019.

Glen Chou, Necmiye Ozay, and Dmitry Berenson. Safe output feedback motion planning from images via learned perception modules and contraction theory. In *WAFR*, 2022.

Christophe Combastel. A state bounding observer for uncertain non-linear continuous-time systems based on zonotopes. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 7228–7234. IEEE, 2005.

Erwin Coumans and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. `http://pybullet.org`, 2016–2021.

Sarah Dean and Benjamin Recht. Certainty equivalent perception-based control. In *Learning for Dynamics and Control*, pages 399–411. PMLR, 2021.

Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.

Marta Garnelo, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J Rezende, SM Eslami, and Yee Whye Teh. Neural processes. *arXiv preprint arXiv:1807.01622*, 2018.

Paul J. Goulart. *Affine Feedback Policies for Robust Control with Constraints*. PhD thesis, University of Cambridge, 2006.

Walter D. Grossman. Observers for discrete-time nonlinear systems. In *PhD thesis, New Jersey Institute of Technology*, 1999.

Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30, 2017.

IS Khalil, JC Doyle, and K Glover. *Robust and optimal control*, volume 2. Prentice hall New York, 1996.

Craig Knuth, Glen Chou, Necmiye Ozay, and Dmitry Berenson. Planning with learned dynamics: Probabilistic guarantees on safety and reachability via lipschitz constants. *IEEE Robotics and Automation Letters*, 6(3):5129–5136, 2021.

Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30, 2017.

Antoine P Leeman, Johannes Kohler, Florian Messerer, Amon Lahr, Moritz Diehl, and Melanie N Zeilinger. Fast system level synthesis: Robust model predictive control using Riccati recursions. *IFAC-PapersOnLine*, 58(18):173–180, 2024.

Antoine P Leeman, Johannes Köhler, and Melanie N Zeilinger. Guaranteed robust nonlinear MPC via disturbance feedback. *arXiv preprint arXiv:2509.18760*, 2025a.

Antoine P. Leeman, Johannes Köhler, Andrea Zanelli, Samir Bennani, and Melanie N. Zeilinger. Robust nonlinear optimal control via system level synthesis. *IEEE Transactions on Automatic Control*, 70(7):4780–4787, 2025b. doi:10.1109/TAC.2025.3552482.

David Q Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50 (12):2967–2986, 2014.

Matthew Mirman, Timon Gehr, and Martin Vechev. Differentiable abstract interpretation for provably robust neural networks. In *International Conference on Machine Learning*, pages 3578–3586. PMLR, 2018.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Yaniv Ovadia, Emily Fertig, Jie Ren, Zachary Nado, David Sculley, Sebastian Nowozin, Joshua Dillon, Balaji Lakshminarayanan, and Jasper Snoek. Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift. *Advances in neural information processing systems*, 32, 2019.

Yaniv Romano, Matteo Sesia, and Emmanuel Candes. Classification with valid and adaptive coverage. *Advances in neural information processing systems*, 33:3581–3591, 2020.

Matthias Seeger. Gaussian processes for machine learning. *International journal of neural systems*, 14(02):69–106, 2004.

Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems*, 31, 2018.

Jerome Sieber, Samir Bennani, and Melanie N. Zeilinger. A System Level Approach to Tube-based Model Predictive Control. *IEEE Control Systems Letters*, 6:776–781, 2021. ISSN 24751456. doi:10.1109/LCSYS.2021.3086190.

Robert F Stengel. *Optimal control and estimation*. Courier Corporation, 1994.

Emanuel Todorov and Weiwei Li. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005 American Control Conference (ACC)*, pages 300–306, 2005.

Jelena Trisovic, Andrea Carron, and Melanie N Zeilinger. Uncertainty-aware perception-based control for autonomous racing. *arXiv preprint arXiv:2508.02494*, 2025.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer, 2005.

Peter Walley. Statistical reasoning with imprecise probabilities. 1991.

Yuh-Shyang Wang, Seungil You, and Nikolai Matni. Localized distributed Kalman filters for large-scale systems. *IFAC-PapersOnLine*, 48(22):52–57, 2015.

Yuh-Shyang Wang, Nikolai Matni, and John C Doyle. A system-level approach to controller synthesis. *IEEE Transactions on Automatic Control*, 64(10):4079–4093, 2019.

Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations research*, 62(6):1358–1376, 2014.

Shuyu Zhan, Chih-Yuan Chiu, Antoine P Leeman, and Glen Chou. Robustly constrained dynamic games for uncertain nonlinear dynamics. *arXiv preprint arXiv:2509.16826*, 2025.

A System Level Parametrization

As in [Anderson et al., 2019], we show the SLS relations for linear time-varying models with bounded process/measurement disturbances. Consider the linear time-varying error dynamics around a nominal $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$\Delta x_{k+1} = A_k \Delta x_k + B_k \Delta u_k + E_k w_k, \qquad \Delta y_k = C_k \Delta x_k + F_k v_k, \tag{8}$$

with $w_k \in \mathcal{B}^{n_w}$ and $v_k \in \mathcal{B}^{n_y}$ (unit ℓ_∞ balls). Stacking over the horizon using the block down-shift $\mathbf{Z}$ and block-diagonal matrices $\mathbf{A} = \mathrm{blkdiag}(A_0, \ldots, A_{T-1})$, and similarly $\mathbf{B}, \mathbf{C}, \mathbf{E}, \mathbf{F}$, we write

$$\Delta \mathbf{x} = \mathbf{Z}\mathbf{A}\Delta \mathbf{x} + \mathbf{Z}\mathbf{B}\Delta \mathbf{u} + \mathbf{E}\mathbf{w}, \tag{9a}$$

$$\Delta \mathbf{y} = \mathbf{Z}\mathbf{C}\Delta \mathbf{x} + \mathbf{F}\mathbf{v}, \tag{9b}$$

where $\mathbf{w} = (\tilde{w}, w_0, \ldots, w_{T-1})$ includes the initial state uncertainty. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{E}, \mathbf{F}$ could originate, e.g., from the linearization of the dynamics (1)–(2).

Instead of synthesizing a gain $\mathbf{K}$ directly in $\mathbf{u} = \bar{\mathbf{u}} + \mathbf{K}(\mathbf{y} - \mathbf{Z}\mathbf{C}\bar{\mathbf{x}})$, SLS optimizes the closed-loop responses Φ from disturbances to $(\Delta \mathbf{x}, \Delta \mathbf{u})$:

$$\begin{bmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{u} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{\Phi}^{\mathrm{xw}} & \mathbf{\Phi}^{\mathrm{xv}} \\ \mathbf{\Phi}^{\mathrm{uw}} & \mathbf{\Phi}^{\mathrm{uv}} \end{bmatrix}}_{\Phi} \begin{bmatrix} \mathbf{E} & 0 \\ 0 & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix}. \tag{10}$$

The maps Φ must satisfy two affine system level constraints ensuring causality and consistency with the dynamics:

$$[\mathbf{I} - \mathbf{Z}\mathbf{A} \quad -\mathbf{Z}\mathbf{B}]\, \Phi = [\mathbf{I} \quad \mathbf{0}], \tag{11a}$$

$$\Phi \begin{bmatrix} \mathbf{I} - \mathbf{Z}\mathbf{A} \\ -\mathbf{Z}\mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}. \tag{11b}$$

For any Φ satisfying (11), there exists a unique strictly causal output-feedback law realizing (10); one valid reconstruction is

$$\mathbf{K} \;=\; \mathbf{\Phi}^{\mathrm{uv}} - \mathbf{\Phi}^{\mathrm{uw}} \left(\mathbf{\Phi}^{\mathrm{xw}}\right)^{-1} \mathbf{\Phi}^{\mathrm{xv}}, \tag{12}$$

and $\mathbf{K}$ is block-lower-triangular with zeros on the diagonal (measurement-delay convention).

B Safety under ambiguity via SLS

Let state/input constraints be linear inequalities

$$A_x x_k \le b_x, \qquad A_u u_k \le b_u, \quad k = 0, \ldots, T-1. \tag{13}$$

Write $x_k = \bar{x}_k + \Delta x_k$, $u_k = \bar{u}_k + \Delta u_k$, and use the responses (10). Since w, e live in ℓ_∞ boxes and the maps are linear, worst-case tightening reduces to support functions of ℓ_∞ balls, i.e., weighted ℓ_1 norms.

Proposition 1 (Entrywise ℓ_∞ tightenings)**.** *Suppose $w_j \in \mathcal{B}^{n_w}$ and $e_j \in \mathcal{B}^{n_y}$ for all j. Then safety $A_x x_k \le b_x$ and $A_u u_k \le b_u$ for all admissible $(\mathbf{w}, \mathbf{e})$ is ensured by the nominal tightenings*

$$A_x \bar{x}_k \le b_x \;-\; \sum_{j=0}^{T} \left|A_x\right| \mathbf{1}\, \underbrace{\left\|\mathbf{\Phi}^{\mathrm{xw}}_{k,j} E_j\right\|_{\infty\to\infty}}_{\textit{row-wise } \ell_1 \textit{ bound}} \;-\; \sum_{j=0}^{T} \left|A_x\right| \mathbf{1}\, \left\|\mathbf{\Phi}^{\mathrm{xv}}_{k,j} F_j\right\|_{\infty\to\infty}, \tag{14a}$$

$$A_u \bar{u}_k \le b_u \;-\; \sum_{j=0}^{T} \left|A_u\right| \mathbf{1}\, \left\|\mathbf{\Phi}^{\mathrm{uw}}_{k,j} E_j\right\|_{\infty\to\infty} \;-\; \sum_{j=0}^{T} \left|A_u\right| \mathbf{1}\, \left\|\mathbf{\Phi}^{\mathrm{uv}}_{k,j} F_j\right\|_{\infty\to\infty}, \tag{14b}$$

where $\|\cdot\|_{\infty\to\infty}$ denotes the induced ℓ_∞ operator norm (maximum absolute row-sum), and $\mathbf{1}$ is a vector of ones of matching size.

Proof. For any row $a^\top$ of A_x, the worst-case deviation is $\max_{\|w\|_\infty \le 1} |a^\top \sum_j \mathbf{\Phi}^{\mathrm{xw}}_{k,j} E_j w_j| \le \sum_j \|a^\top \mathbf{\Phi}^{\mathrm{xw}}_{k,j} E_j\|_1$, and similarly for e. Summing row-wise gives (14). □

Notes. Nonlinear cases inherit these tightenings on each SCP iteration by plugging the linear time-varying linearization $(A_k, B_k, C_k, E_k, F_k)$.

C Output-Feedback SLS as LQR and Kalman Filter

Similarly as in Anderson et al. [2019], this appendix clarifies how the output-feedback SLS parameterization admits a clean decomposition into a state-feedback part and a measurement update (Kalman-like) part leveraged in Section 4.3 for efficient numerical solution.

Recall the stacked LTV error system and output-feedback SLS parametrization from Appendix A:

$$\Delta\mathbf{x} = \mathbf{ZA}\Delta\mathbf{x} + \mathbf{ZB}\Delta\mathbf{u} + \mathbf{Ew}, \qquad \Delta\mathbf{y} \;=\; \mathbf{ZC}\Delta\mathbf{x} + \mathbf{Fv}, \tag{15}$$

$$\begin{bmatrix}\Delta\mathbf{x}\\ \Delta\mathbf{u}\end{bmatrix} = \begin{bmatrix}\mathbf{\Phi}^{\mathrm{xw}} & \mathbf{\Phi}^{\mathrm{xv}}\\ \mathbf{\Phi}^{\mathrm{uw}} & \mathbf{\Phi}^{\mathrm{uv}}\end{bmatrix}\begin{bmatrix}\mathbf{E} & 0\\ 0 & \mathbf{F}\end{bmatrix}\begin{bmatrix}\mathbf{w}\\ \mathbf{v}\end{bmatrix}, \quad \text{with } \Phi \text{ satisfying (11).} \tag{16}$$

Introduce intermediate closed-loop responses that separately handle the process and measurement uncertainties:

- A state-feedback pair $(\bar{\mathbf{\Phi}}^{\mathrm{x}}, \bar{\mathbf{\Phi}}^{\mathrm{u}})$ that maps *process* disturbances to $(\Delta\mathbf{x}, \Delta\mathbf{u})$ and satisfies the left SLS constraint (with $\mathbf{C}$ absent):

  $$\begin{bmatrix}\mathbf{I} - \mathbf{ZA} & -\mathbf{ZB}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{\Phi}}^{\mathrm{x}}\\ \bar{\mathbf{\Phi}}^{\mathrm{u}}\end{bmatrix} = \mathbf{I}.$$

  Intuitively, this is what LQR designs in closed-loop when the state is directly available.
- An observer pair $(\hat{\mathbf{\Phi}}^{\mathrm{x}}, \hat{\mathbf{\Phi}}^{\mathrm{y}})$ that maps *measurement* disturbances to $(\Delta\mathbf{x}, \Delta\mathbf{y})$ and satisfies the right SLS constraint (with $\mathbf{B}$ absent):

  $$\begin{bmatrix}\hat{\mathbf{\Phi}}^{\mathrm{x}} & \hat{\mathbf{\Phi}}^{\mathrm{y}}\end{bmatrix}\begin{bmatrix}\mathbf{I} - \mathbf{ZA}\\ -\mathbf{ZC}\end{bmatrix} = \mathbf{I}.$$

  Intuitively, this is what a Kalman filter implements in closed loop for the estimation dynamics.

With these two "half" responses in hand, the full output-feedback maps that satisfy (11) are obtained by the composition:

$$\mathbf{\Phi}^{\mathrm{xw}} = \bar{\mathbf{\Phi}}^{\mathrm{x}} + \hat{\mathbf{\Phi}}^{\mathrm{x}} - \bar{\mathbf{\Phi}}^{\mathrm{x}}(\mathbf{I} - \mathbf{ZA})\hat{\mathbf{\Phi}}^{\mathrm{x}}, \tag{17a}$$

$$\mathbf{\Phi}^{\mathrm{uw}} = \bar{\mathbf{\Phi}}^{\mathrm{u}} - \bar{\mathbf{\Phi}}^{\mathrm{u}}(\mathbf{I} - \mathbf{ZA})\hat{\mathbf{\Phi}}^{\mathrm{x}}, \tag{17b}$$

$$\mathbf{\Phi}^{\mathrm{xv}} = \hat{\mathbf{\Phi}}^{\mathrm{y}} - \bar{\mathbf{\Phi}}^{\mathrm{x}}(\mathbf{I} - \mathbf{ZA})\hat{\mathbf{\Phi}}^{\mathrm{y}}, \tag{17c}$$

$$\mathbf{\Phi}^{\mathrm{uv}} = -\bar{\mathbf{\Phi}}^{\mathrm{u}}(\mathbf{I} - \mathbf{ZA})\hat{\mathbf{\Phi}}^{\mathrm{y}}. \tag{17d}$$

D Additional experimental results

D.1 Car experiment

Figure 3 shows the nominal plan with (closed-loop) state credal sets and sample rollouts.

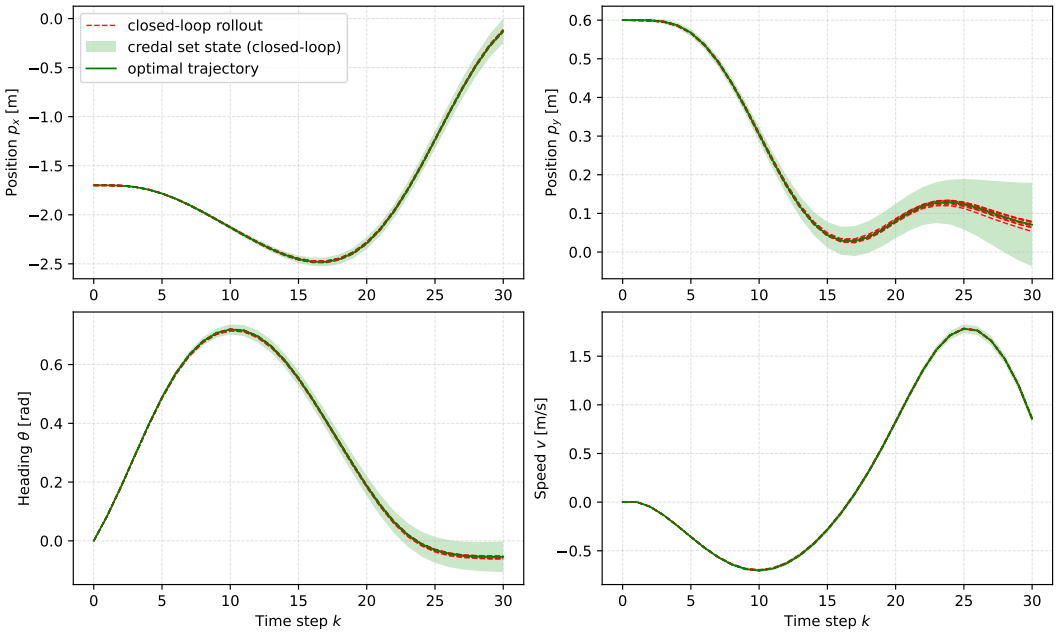

Figure 3: Closed-loop rollouts for the car task in the occluded scene. The plot overlays the nominal plan, the SLS-induced credal set (used for constraint tightenings), and sample closed-loop rollouts.