Safe Perception-Based Control with Minimal Worst-Case Dynamic Regret

Hongyu Zhou* and Vasileios Tzoumas†

Abstract—We enable safe control of linear time-varying systems in the presence of unknown and unpredictable process and measurement noise. We introduce a control algorithm that minimizes dynamic regret, i.e., that minimizes the sub-optimality against an optimal clairvoyant controller that knows the unpredictable future a priori. Specifically, our algorithm minimizes the worst-case dynamic regret among all possible noise realizations given a worst-case total noise magnitude. To this end, the control algorithm accounts for three key challenges: safety constraints; partially-observed time-varying systems; and unpredictable process and measurement noise. We are motivated by the future of autonomy where robots will autonomously perform complex tasks despite unknown and unpredictable disturbances leveraging their on-board control and perception capabilities. To synthesize our minimal-regret perception-based controller, we formulate a constrained semi-definite program based on a System Level Synthesis approach we enable for partially-observed time-varying systems. We validate our algorithm in simulated scenarios, including trajectory tracking scenarios of a hovering quadrotor collecting GPS and IMU measurements and of an omnidirectional robot collecting range measurements from landmarks. Our algorithm is observed to be superior to either or both the $H_2$ and $H_\infty$ controllers, demonstrating a Best of Both Worlds performance.

I. INTRODUCTION

In the future, robots will be leveraging their on-board control and perception capabilities to complete tasks such as:

- **Package Delivery**: How can quadrotors autonomously deliver packages from warehouses to houses? [1]
- **Transportation**: How can air taxis autonomously commute people to desired destinations? [2]
- **Disaster Response**: How can aerial vehicles search for people during severe weather conditions? [3]

To complete such complex tasks, the robots need to reliably overcome a series of key challenges:

- **Challenge I: Safety Constraints**. Robots need to ensure their own safety and the safety of their surroundings. For example, robots often need to ensure that they follow prescribed collision-free trajectories or that their control effort is kept under prescribed levels. Such safety requirements take the form of state and control input constraints and make planning control inputs computationally hard [4], [5].
- **Challenge II: Partially-Observed Time-Varying System**. Robots often lack full-state information feedback for control. Instead, they base their control on perception capabilities that are governed by non-linear and time-varying measurement models. For example, self-driving robots in indoor environments often base their control on range sensors which provide relative-distance measurements to known landmarks [6]. The measurements model of range sensors is not merely non-linear; it is also time-varying since range measurements depend on the relative position of the robot to the landmarks. Accounting for such models typically requires linearization [7], which adds to the hardness of computing safe control inputs.
- **Challenge III: Unknown and Unpredictable Process and Measurement Noise**. The robots' dynamics and measurements are often disturbed by unknown and unstructured noise, which is not necessarily Gaussian. For example, aerial and marine vehicles often face unpredictable winds and waves [8], [9]. But the current control algorithms primarily rely on known or Gaussian-structured noise, compromising thus the robots' ability to ensure safety against unknown and unpredictable noise [10], [11].

The challenges give rise to a novel technological need:

**Technological Need.** We need algorithms that enable safe control of partially-observable linear time-varying systems, guaranteeing near-optimal control performance even against unknown and unpredictable noise. The performance's near-optimality may be captured by bounding the algorithms' sub-optimality with respect to optimal time-varying clairvoyant controllers, i.e., by bounding dynamic regret.

In this paper, we aim to address the Technological Need.

**Related Work.** The current control algorithms assume (i) no safety constraints, or (ii) fully-observed systems or partially-observed systems but governed by time-invariant measurement models, or (iii) known stochastic models about the process and measurement noise, typically a Gaussian model. No current control algorithms address the above challenges simultaneously and thus the Technological Need.

We next review the literature by first reviewing online learning for control algorithms, i.e., algorithms that select control inputs based on past information only [12]–[17], and then by reviewing robust control algorithms that select inputs based on simulating the future system dynamics across a lookahead horizon [18]–[22]:

- **a) Online Learning for Control**: The algorithms performing online learning for control make no assumptions about the noise [23], [24], aiming to address the Challenge III. They assume the noise can evolve arbitrarily, subject to a given upper bound on its magnitude. The upper bound ensures problem feasibility, and tunes the algorithms’ response to the nevertheless unknown noise.

The online learning algorithms prescribe control policies by optimizing feedback control gains based on past information only, guaranteeing performance by bounding
variant measurement models.

The current static-regret algorithms consider no safety constraints, and apply to fully-observed systems or partially-observed systems with time-invariant measurement models only; and the current dynamic-regret algorithms also ignore safety constraints, and apply to fully-observed systems only. All in all, although the current online learning algorithms address unpredictable noise, they assume no safety constraints and cannot apply to partially-observed time-varying systems.

b) Robust Control: The classical $H_2$ and $H_\infty$ control algorithms [26] assume Gaussian noise and bounded arbitrary noise, respectively. Further, they guarantee performance by bounding the $H_2$ and $H_\infty$ cost, respectively. Hence, they cannot guarantee near-optimality against unpredictable noise.

Recent robust control algorithms aim to address the challenge of unpredictable noise, guaranteeing regret optimality, i.e., minimal worst-case dynamic-regret among all noise realizations subject to a given total noise magnitude. Specifically, [18], [19] focus on fully-observed systems and [20] on partially-observed systems with time-invariant measurement models only. But these algorithms consider no safety constraints. Instead, [21], [22] provide a regret-optimal control algorithm that accounts for safety constraints. However, the algorithm applies to fully-observed systems only.

In sum, the current robust algorithms cannot address simultaneously the Technological Need.

Contributions. We address the Technological Need, given a known upper bound to the total magnitude of the noise. The assumption is common in the literature, as outlined above; besides, if we assume both unbounded and unknown and unpredictable noise, then safe control is infeasible in the worst-case. We make the contributions (Section III):

- Algorithmic Contribution: We provide the first algorithm enabling safe control of partially-observed linear time-varying systems against unknown and unpredictable process and measurement noise. The algorithm prescribes an output-feedback control input, guaranteeing minimum worst-case dynamic regret among all noise realizations. It is a robust control algorithm, selecting inputs based on simulating the dynamics across a lookahead horizon.
- Technical Contributions: To enable the safe regret-optimal algorithm, we make the following technical innovations:
  - We prove that the output-feedback control gains are the solution of a constrained Semi-Definite Program (SDP). Our SDP approach innovates by handling partially-observed systems instead of only fully-observed.
  - We enable our SDP formulation by introducing a Generalized System Level Synthesis (G-SLS) method for partially-observed time-varying systems, proving necessary and sufficient conditions for the existence of a causal safe output-feedback control policy. G-SLS generalizes the current SLS methods which assume instead either partially-observed time-invariant systems [27] or fully-observed time-varying systems [28].

Numerical Evaluations. We validate our algorithm in simulated scenarios of partially-observed linear time-varying systems, including trajectory tracking scenarios of a hovering quadrotor collecting GPS and Inertial Measurement Unit (IMU) measurements and of an omnidirectional robot collecting range measurements from landmarks (Section IV). We compare our algorithm with the $H_2$ and $H_\infty$ control algorithms under diverse process and measurement noises.

Our algorithm demonstrates a Best of Both Worlds performance across all simulations and test types of noise, performing on average better than at least one of the $H_2$ and $H_\infty$ controllers. That is, our algorithm demonstrates robustness across all tested types of noise, being the best or the second best among $H_2$ and $H_\infty$, an advantageous performance capacity when the type of noise is unknown a priori and unpredictable. Instead, $H_2$ is the worst against worst-case noise and $H_\infty$ is the worst against Gaussian.

Organization. Section II formulates the problem of safe perception-based control with minimal worst-case dynamic regret guarantees. Section III develops the control algorithm. Section IV presents the numerical evaluation. Section V concludes the paper. The Appendix contains all proofs.

II. PROBLEM FORMULATION: SAFE PERCEPTION-BASED CONTROL WITH MINIMAL WORST-CASE REGRET

We formulate the problem of Safe Perception-Based Control with Minimal Worst-Case Regret (Problem 1). To this end, we first establish our notation and assumptions.

Partially-Observed System. We consider partially-observed Linear Time-Varying (LTV) systems of the form

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \quad t \in \{0, \ldots, T-1\},$$

$$y_t = C_t x_t + e_t,$$

where $t$ is the time index, $T$ is a time horizon of interest, $x_t \in \mathbb{R}^{d_x}$ is the system’s state, $u_t \in \mathbb{R}^{d_u}$ is the control input, $w_t \in \mathbb{R}^{d_w}$ is the process noise, $y_t \in \mathbb{R}^{d_y}$ is the measurements, and $e_t \in \mathbb{R}^{d_e}$ is the measurement noise.

We henceforth denote:

- $x \triangleq [x_0^T, x_1^T, \ldots, x_{T-1}^T]^T$, i.e., $x$ is the state trajectory across the time horizon $T$;
- $u$, $y$, and $e$ are defined correspondingly to $x$;
- $w \triangleq [w_0^T, w_1^T, \ldots, w_{T-1}^T]^T$, i.e., initial condition appended with the process noise trajectory till time $T - 2$.

Assumption 1 (Known System). The initial condition $x_0$ and the matrices $A_t$, $B_t$, and $C_t$ for all $t$ are known.

Assumption 2 (Bounded Noise). The process and measurement noise $w_t$ and $e_t$ are constrained in known compact polytops that contain a neighborhood of the origin: i.e., we are given $W \triangleq \{w \in \mathbb{R}^{d_w T} : H_w w \leq h_w \}$ and $E \triangleq \{e \in \mathbb{R}^{d_e T} : H_e e \leq h_e \}$ for given matrices $H_w$, $H_e$, $h_w$, and $h_e$.

Per Assumption 2, we assume no stochastic model for the noise. Specifically, the noise may even be adversarial, subject to the bounds prescribed by $W$ and $E$. 

Safety Constraints. We consider the states and control inputs must satisfy polytopic constraints of the form

\[ H \begin{bmatrix} x \\ u \end{bmatrix} \leq h, \quad \forall w \in \mathbb{W}, \forall e \in \mathbb{E}, \quad (2) \]

for given matrices \( H \) and \( h \).

Output-Feedback Control Policy. We consider the following output-feedback control policy:

\[ u_t = \sum_{k=0}^{t} K_{t,k} y_k, \quad t \in \{0, \ldots, T-1\}, \quad (3) \]

where \( K_{t,k} \) are control gains to be designed in this paper.

Control Performance Metric. We design the output-feedback control gains \( K_{t,k} \) to ensure both safety and a control performance comparable to an optimal clairvoyant policy that selects control inputs knowing the future noise realizations \( w \) and \( e \) a priori. In this paper, we consider that the clairvoyant policy minimizes a cost of the form

\[ \text{cost}(w, e, u) \triangleq x^T Q x + u^T R u, \quad (4) \]

where \( Q \geq 0 \) and \( R > 0 \), and \( x \) is a function of the control input sequence \( u \) and the noise \( w \) and \( e \) per eq. (1). \( Q \) and \( R \) are assumed symmetric, without loss of generality. Then, the suboptimality of any (causal) control sequence \( u \) that is unaware of the noise realization \( w \) and \( e \) is captured by the

\[ \text{regret}_T(w, e, u) \triangleq \text{cost}(w, e, u) - \min_{u' \in \mathbb{R}^{d_u \times T}} \text{cost}(w, e, u'), \quad (5) \]

where \( \min_{u' \in \mathbb{R}^{d_u \times T}} \text{cost}(w, e, u') \) is the cost achieved by the optimal clairvoyant control policy.

In light of the following Remark 1, in this paper we consider the optimal clairvoyant control policy in eq. (5) to be the optimal clairvoyant \( H_2 \) controller [18].

Remark 1 (Nonexistence of a Universally Optimal Clairvoyant Control Policy [20, Theorem 1]). For partially-observed linear systems, there exists no single clairvoyant controller that outperforms all other clairvoyant controllers across all process and measurement noise realizations \((w, e)\).

In this paper, we design the output-feedback control gains \( K_{t,k} \) to minimize the worst-case dynamic regret among all feasible noise realizations per Assumption 2.

Definition 1 (Worst-Case Dynamic Regret [20]). Denote by \( r \) the minimum radius of a ball in \( \mathbb{R}^{d_u \times T + d_w \times T} \) that encircles the noise’s domain sets \( \mathbb{W} \) and \( \mathbb{E} \). Then,

\[ \text{worst-case-regret}_T(u) \triangleq \max_{\|w\|_2^2 + \|e\|_2^2 \leq r^2} \text{regret}_T(w, e, u). \quad (6) \]

That is, eq. (6) is the worst-case dynamic regret among all noise realizations with maximum feasible total magnitude.

Problem Definition. In this paper, we focus on:

Problem 1 (Safe Perception-Based Control with Minimal Worst-Case Dynamic Regret). Find control gains \( K_{t,k} \) such that the output-feedback control policy in eq. (3) guarantees

(i) the safety of the partially-observed LTV system in eq. (1) and (ii) minimal worst-case dynamic regret. Formally,

\[ \min_{K_{t,k}} \quad \text{worst-case-regret}_T(u) \quad (7a) \]

subject to the system in eq. (1);

the safety constraints in eq. (2);

the control policy in eq. (3). \quad (7d)

III. ALGORITHM FOR PROBLEM 1

We present the algorithm for Problem 1 (Algorithm 1). The algorithm solves Problem 1 via an equivalent SDP reformulation. We present the SDP reformulation in Section III-B (Theorem 1). To prove the SDP reformulation, we first introduce G-SLS, an SLS approach for partially-observed linear time-varying systems in Section III-A (Proposition 1).

A. Preliminary Results: Generalized System Level Synthesis for Partially-Observable LTV Systems

We present G-SLS, an Generalized SLS approach applying to partially-observed LTV systems. Particularly, given a desired state trajectory \( x \), we show necessary and sufficient conditions for the existence of a control policy \( u_t \) per eq. (4). Equivalently, we show necessary and sufficient conditions for the existence of control gains \( K_{t,k} \) such that \( K_{t,k} = 0 \) when \( k > t \). The conditions take the form of linear matrix constraints and thus enable the computation of the \( K_{t,k} \) within an SDP reformulation of Problem 1 (Section III-B).

To the above ends, we use the notation:

- \( Z \) denotes the block-matrix downshift operator, i.e.,

\[ Z \triangleq \begin{bmatrix} 0 & 0 & \ldots & 0 \\ I & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & I & 0 \end{bmatrix}. \quad (8) \]

- \( A, B, \) and \( C \) are the diagonal block-matrices whose block diagonal is the (partial) trajectory of the corresponding system, input, and measurement matrix, i.e.,

\[ A \triangleq \text{blkdiag} (A_0, A_1, \ldots, A_{T-2}, 0_{d_x \times d_x}) ; \]

\[ B \triangleq \text{blkdiag} (B_0, B_1, \ldots, B_{T-2}, 0_{d_x \times d_u}) ; \]

\[ C \triangleq \text{blkdiag} (C_0, C_1, \ldots, C_{T-1}) ; \]

- \( K \) is the lower triangular block-matrix such that eq. (3) takes the form \( u = Ky, \) i.e.,

\[ K \triangleq \begin{bmatrix} K_{0,0} & 0_{d_x \times d_y} & \ldots & 0_{d_y \times d_y} \\ K_{1,0} & K_{1,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{d_y \times d_y} \\ K_{T-1,0} & K_{T-1,1} & \ldots & K_{T-1,T-1} \end{bmatrix}. \quad (9) \]

Equations (1) and (3) now take the form

\[ x = ZAx + ZBu + w; \quad (10a) \]

\[ y = Cx + e; \quad (10b) \]

\[ u = Ky; \quad (10c) \]
We change the optimization variables in Problem 1 from the output-feedback control gains in $K$ to the generalized response matrix $Φ$. We thus leverage that finding a feasible $Φ$ requires searching over a convex set, in particular, the set defined by Proposition 1's necessary and sufficient conditions which take the form of linear matrix constraints. Once a $Φ$ is found, then $K$ is computed via eq. (14).

- We reformulate the safety constraints in eq. (2) as linear matrix inequalities. To this end, we adopt the dualization procedure introduced in the proof of [21, Theorem 3];
- We reformulate Problem 1's objective function, namely, worst-case-regret$_T(u)$, as an equivalent minimization problem of a scalar subject to linear matrix inequalities. To this end, we perform the first step of the proof of [20, Theorem 4] and then apply Schur complement.

We use the following notation and definitions to formally state Problem 1's SDP reformulation and Algorithm 1:

- $D ≜ \text{blkdiag}(Q, R)$, i.e., $D$ is the diagonal block-matrix whose elements $Q$ and $R$ define the cost in eq. (4);
- $Z$ are the dual variables introduced to reformulate the safety constraints in eq. (2) as linear matrix inequalities;
- $ λ$ is the scalar that once minimized subject to appropriate linear matrix inequalities becomes equal to Problem 1’s objective function, i.e., to worst-case-regret$_T(u)$;
- $Φ^c$ is the generalized response corresponding to the optimal clairvoyant $H_2$ controller in eq. (5) that ignores the safety constraints; i.e., per [18],

$$Φ^c ∈ \arg\min_{Φ} \left\| \begin{bmatrix} Q^2 & 0 \\ 0 & R^2 \end{bmatrix} Φ \right\|_F^2 \text{ subject to } eq. (13a) \text{ and eq. (13b).} \quad (15)$$

Theorem 1 (SDP Reformulation of Problem 1). Problem 1 is equivalent to the Semi-Definite Program

$$\min_{\Phi, Z, λ} \lambda \quad \text{subject to: } \Phi_{wx}, \Phi_{xe}, \Phi_{uw}, \Phi_{ue} \text{ being lower block triangular; } \quad (16a)$$

$$eq. (13a) \text{ and eq. (13b).} \quad (16b)$$

$$Z^T \begin{bmatrix} h_w \\ h_e \end{bmatrix} \leq h, HΦ ≤ Z^T \begin{bmatrix} H_w & 0 \\ 0 & H_e \end{bmatrix}, \quad Z_{ij} ≥ 0. \quad (16c)$$

$$λ > 0, \begin{bmatrix} I & D^T \Phi \\ \Phi^T D & λI + (Φ^c)^T DΦ^c \end{bmatrix} ≥ 0. \quad (16d)$$

Theorem 1 prescribes an SDP in place of Problem 1. Particularly, eq. (16) relates to Problem 1 as follows: eqs. (16a) and (16b) result from the change of the optimization variables in Problem 1 from the output-feedback control gains in $K$ to the generalized response matrix $Φ$, per the necessary and sufficient conditions in Proposition 1; eq. (16c) results from the reformulation of the safety constraints in eq. (2) as linear matrix inequalities, per the dualization procedure.
in [21, Proof of Theorem 3]; and the new objective of minimizing the scalar \( \lambda \) subject to the linear matrix inequalities in eq. (16d) result from the reformulation of Problem 1’s objective function, that is, of worst-case-regret\( _r \) (16).

Algorithm 1’s Description. Algorithm 1 solves Problem 1 by (i) solving Problem 1’s equivalent SDP reformulation in eq. (16) to obtain an optimal generalized response matrix \( \Phi \) (line 1), and then by (ii) computing the corresponding output-feedback control gain block-matrix \( \mathcal{K} \) per eq. (14).

IV. Numerical Evaluations in Trajectory Tracking Scenarios

We evaluate Algorithm 1 in simulated scenarios of safe perception-based control for trajectory tracking. We first consider synthetic partially-observed LTV systems aiming to stay at zero despite noise disturbances (Section IV-A). Then, we consider a quadrotor aiming to stay at a hovering position; to this end, the quadrotor collects asynchronous GPS and Inertial Measurement Unit (IMU) measurements (Section IV-B). Finally, we consider a robot-vacuum, modeled as an omnidirectional ground vehicle, aiming to track a predefined arc; to this end, the robot collects range measurements from known landmarks (Section IV-C).

Compared Algorithms. We compare Algorithm 1 with the classical \( H_2 \) and \( H_\infty \) controllers as well as the safe optimal clairvoyant \( H_2 \) controller obtained by solving eq. (15) with safety constraints in eq. (2).

Tested Noise Types. We corrupt the state dynamics and the sensor measurements with diverse noise: (i) stochastic noise, drawn for the Gaussian, Uniform, Gamma, Exponential, Bernoulli, Weibull, or Poisson distribution, and (ii) non-stochastic noise, in particular, worst-case (adversarial) noise.

Summary of Results. Algorithm 1 demonstrates a Best of Both Worlds (BoBW) performance: either it is superior to \( H_2 \) and \( H_\infty \) across the tested types of noise, or it performs better than \( H_2 \) or \( H_\infty \) across all tested types of noise.

We performed all simulations in MATLAB with YALMIP toolbox [29] and MOSEK solver [30].

Our code will be open-sourced via a link here.

A. Synthetic Partially-Observed LTV Systems

Simulation Setup. We consider LTV systems such that

\[
A_t = \begin{bmatrix} 0.7 & 0.2 & 0 \\ 0.3 & 0.7 & -0.1 \\ 0 & -0.2 & 0.8 \end{bmatrix}, \quad B_t = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad C_t = \begin{cases} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, & t \in \{1, 3, \ldots\}; \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, & t \in \{2, 4, \ldots\}, \end{cases}
\]

where \( \rho \) is the spectral radius of the system.

We demonstrate Algorithm 1 first on an open-loop stable system where \( \rho = 0.85 \), and then on an open-loop unstable system where \( \rho = 1.05 \). For the system with \( \rho = 0.85 \), we choose the safety constraints: \(-5 \leq x_t \leq 5 \) and \(-5 \leq u_t \leq 5 \); and we assume noise such that \(-1 \leq w_t \leq 1 \) and \(-1 \leq e_t \leq 1 \). For the system with \( \rho = 1.05 \), we choose the safety constraints: \(-30 \leq x_t \leq 30 \) and \(-30 \leq u_t \leq 30 \); and we assume noise such that \(-1 \leq w_t \leq 1 \) and \(-1 \leq e_t \leq 1 \).

We consider that \( \mathcal{Q} \) and \( \mathcal{R} \) are the identity matrix \( \mathbf{I} \).

We simulate the setting for all \( T \in \{2, 3, \ldots, 30\} \).

Results. The results are summarized in Figure 1 and Figure 2 for \( \rho = 0.85 \) and \( \rho = 1.05 \) respectively. Under Gaussian and worst-case noise, Algorithm 1’s performance lies between that of the \( H_2 \) and \( H_\infty \) controllers. Under all other noise types: for \( \rho = 0.85 \), Algorithm 1 outperforms the \( H_2 \) and \( H_\infty \) controllers; for \( \rho = 1.05 \), Algorithm 1 outperforms the \( H_\infty \) controller. In sum, Algorithm 1 demonstrates a BoBW performance across all scenarios.

B. Hovering Quadrotor

Simulation Setup. We consider a quadrotor model with state vector its position and velocity, and control input its roll, pitch, and total thrust. The quadrotor’s goal is to stay at a predefined hovering position. To this end, we focus on its linearized dynamics, taking the form

\[
A_t = \begin{bmatrix} 1 & 0 & 0 & 0.1 & 0 \\ 0 & 1 & 0 & 0.1 & 0 \\ 0 & 0 & 1 & 0.1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad B_t = \begin{bmatrix} -4.91 \\ 0 \\ 4.91 \\ 0 \\ 0 \end{bmatrix}. \]

The quadrotor collects GPS and IMU measurements. The GPS measurements are available every 3 time steps, and the IMU measurements are available in all other time steps, reflecting the real-world scenarios where IMU measurements are more frequently available [31]. Formally,

\[
C_t = \begin{cases} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, & t \in \{1, 4, \ldots\}; \\ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, & t \in \{2, 3, 5, 6, \ldots\}. \end{cases}
\]

We choose the safety constraints: \(-5 \leq x_t \leq 5 \) and \([-\pi - \pi - 20] \leq u_t \leq [\pi \pi 20] \); and we assume noise such that \(-0.1 \leq w_t \leq 0.1 \) and \(-0.1 \leq e_t \leq 0.1 \).

We consider that \( \mathcal{Q} \) and \( \mathcal{R} \) are the identity matrix \( \mathbf{I} \).

We simulate the setting for all \( T \in \{2, 3, \ldots, 25\} \).

Results. The results are summarized in Figure 3. Algorithm 1’s performance lies between that of the \( H_2 \) and \( H_\infty \) controllers under Gaussian and worst-case noise. Under all other noise types, Algorithm 1 always outperforms the \( H_2 \) controller, and is on par with the \( H_\infty \) controller up to horizon \( T = 15 \). All in all, Algorithm 1 demonstrates a superior or BoBW performance across all scenarios.

C. Omnidirectional Ground Robot

Simulation Setup. We consider an omnidirectional ground robot eq. (17) with state vector its position, and control input its velocity given a sampling time 0.05s. Formally,

\[
A_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_t = \begin{bmatrix} 0.05 \\ 0 \end{bmatrix}. \]
Fig. 1: Performance Comparison given the LTV System in Section IV-A with $\rho = 0.85$ (stable system case). The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.

Fig. 2: Performance Comparison given the LTV System in Section IV-A with $\rho = 1.05$ (unstable system case). The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.

Fig. 3: Performance Comparison given the Quadrotor System in Section IV-B. The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.
The robot aims to track a trajectory that starts at (0, 1) and ends at (1, 0) passing through intermediate locations along the arc with radius 1 centered at the origin (Figure 4).

To track the trajectory, the robot can estimate its state by collecting range measurements from 3 known landmarks, located at \(L_1 = (0, 0)\), \(L_2 = (0, 2)\), and \(L_3 = (2, 0)\), respectively (Figure 4). Particularly, the robot’s measurement matrix upon linearization at a robot’s location \((\ell_i^t, \ell_j^t)\) is

\[
C_t = \begin{bmatrix}
\ell_i^t - L_1^t \\
\sqrt{(\ell_i^t - L_1^t)^2 + (\ell_j^t - L_1^t)^2} \\
\ell_i^t - L_2^t \\
\sqrt{(\ell_i^t - L_2^t)^2 + (\ell_j^t - L_2^t)^2} \\
\ell_i^t - L_3^t \\
\sqrt{(\ell_i^t - L_3^t)^2 + (\ell_j^t - L_3^t)^2}
\end{bmatrix},
\]

where \((L_1^t, L_2^t)\) denotes landmark’s \(k \in \{1, 2, 3\}\) location.

To define the trajectory-tracking control cost we take the next steps. First, we denote by \(s \triangleq [s_0^T, \ldots, s_{T-1}^T]^T\) the desired trajectory, and by \(v \triangleq [v_0^T, \ldots, v_{T-1}^T]^T\) the nominal control input that makes the robot to traverse the desired trajectory in the absence of noise. Then, we observe that \(s = 2A_s + 2B_s v + \delta\), where \(\delta = [s_0^T, 0^T, \ldots, 0^T]^T\). We now set \(u = K(y - Cs) + v\) in eq. (10c). Next, we define the errors \(e_x \triangleq x - s\) and \(e_u \triangleq u - v\), and observe that

\[
\begin{bmatrix}
    e_x \\
    e_u \\
\end{bmatrix} = \begin{bmatrix}
    \Phi_{xw} & \Phi_{xe} \\
    \Phi_{uw} & \Phi_{ue}
\end{bmatrix} \begin{bmatrix}
    \bar{w} \\
    e
\end{bmatrix},
\]

where \(\bar{w} \triangleq [x_0^T - s_0^T, w_0^T, \ldots, w_{T-2}^T]^T\). Finally, we set \(T = 10\), \(Q = 100I\), and \(R = 10I\), and since the actual robot’s locations \(\{(\ell_i^t, \ell_j^t)\}_{t \in \{0, \ldots, T-1\}}\) are unknown a priori due to the unpredictable future noise, we calculate \(\bar{C}\) by evaluating \(C_t\) in eq. (17) at the desired location \(s_t\).

We choose the safety constraints: \(-2 \leq x_t \leq 2\) and \(-2 \leq u_t \leq 2\); and we assume: \(-0.1 \leq \bar{w}_t \leq 0.1, -0.1 \leq e_t \leq 0.1\). For comparison, we also test the system’s performance when \(C_t\) is constant, in particular, when \(C_t = C_0\) for all \(t\).

**Results.** The results are summarized in Table I. Algorithm 1 outperforms both methods across all noise types, with the exception of the Gaussian noise case, where the \(H_2\) controller performs best but where Algorithm 1 still outperforms the \(H_\infty\) controller. Further, Algorithm 1 demonstrates the lowest standard deviation across all noise types with the exception of exponential noise, where Algorithm 1 is worse than the \(H_\infty\) controller but still better than the \(H_2\) controller. In sum, Algorithm 1 demonstrates a superior or BoBW performance across all scenarios.

**V. CONCLUSION**

**Summary.** We provided the first algorithm enabling safe control of partial-observed LTV systems against unknown and unpredictable process and measurement noise (Algorithm 1). Algorithm 1 prescribes an output-feedback control input, guaranteeing safety and minimum worst-case dynamic regret among all noise realizations. To derive Algorithm 1, we formulated an SDP based on a Generalized System Level Synthesis approach we enabled for the first time for partially-observed LTV systems. We validated Algorithm 1 in simulated scenarios; the algorithm was observed to be superior or on par with either the \(H_2\) or \(H_\infty\) controller, demonstrating a Best of Both Worlds performance.

**Future Work.** Algorithm 1 plans control policies given a lookahead time horizon, relying on an a priori knowledge of the state, control input, and measurement matrices across the horizon (Assumption 1). This is infeasible in general: e.g., in camera-based navigation, where state estimation relies on feature detection and tracking over sequential camera frames, the measurement matrices become known only once the frames have been captured and the features have been detected [32]. But which will be the frames and which will be the detected features is typically unknown a priori. In our future work, we aim to address the said limitations by enabling online learning variants of our algorithm that (i) do not rely on Assumption 1, and that (ii) provably guarantee a Best of Both Worlds performance. We will also ensure its efficient implementation to demonstrate application in real-world systems, in particular, aerial drones that aim to land on moving platforms or perform acrobatics in the presence of unpredictable wind disturbances.

**APPENDIX**

**Proof of Proposition 1:** We first prove the sufficiency of the conditions in Proposition 1, and then their necessity.

**Sufficiency:*** We show that: 1) \(\Phi_{xw}, \Phi_{xe}, \Phi_{uw}, \) and \(\Phi_{ue}\) are lower triangular block-matrices; 2) \(\Phi_{xw}, \Phi_{xe}, \Phi_{uw}, \) and \(\Phi_{ue}\) lie in the affine space in eq. (13); 3) controller \(K\) can be recovered from eq. (14). Respectively:

1) The statement holds since \(A, B, C\) are block-diagonal, \(K\) is block-lower-triangular, \(Z\) is the block-downshift operator, and the inverse of a block-lower-triangular matrix remains a block-lower-triangular matrix.

2) From eqs. (10b) and (10c), we have

\[
u = KCx + Ke. \tag{18}\]

Substituting eq. (18) into (10a) gives:

\[
x = ZAx + ZBKCe + w
\Rightarrow (I - ZA - ZBK) x = w + ZBK e \tag{19}\]

\[
\Rightarrow x = \Phi_{xw}w + \Phi_{xe}e.
\]
TABLE I: Performance Comparison given the Robot-Vacuum System in Section IV-C. The performance is quantified per the cost in eq. (4). The bold numbers represent the best performance among $H_2$, $H_\infty$, and Algorithm 1. The performance of the clairvoyant $H_2$ controller is noted with blue.

| Noise Distribution | Mean | Time-varying $C_1$ | Standard Deviation | Time-varying $C_1$ |
|--------------------|------|-------------------|-------------------|-------------------|
|                    | $H_2$ | $H_\infty$ | Ours | Clairvoyant $H_2$ | $H_2$ | $H_\infty$ | Ours | Clairvoyant $H_2$ |
| Gaussian           | 1.54  | 2.14   | 1.93  | 1.12             | 1.53  | 2.14   | 1.80  | 1.16            |
| Uniform            | 7.38  | 7.32   | 6.74  | 5.06             | 6.95  | 6.74   | 6.32  | 4.97            |
| Gamma              | 7.86  | 7.77   | 7.15  | 5.38             | 7.40  | 7.12   | 6.70  | 5.29            |
| Exponential        | 9.79  | 9.21   | 9.04  | 6.88             | 9.50  | 8.87   | 8.74  | 6.84            |
| Bernoulli          | 10.02 | 9.48   | 9.15  | 6.95             | 9.64  | 9.01   | 8.77  | 6.89            |
| Weibull            | 9.23  | 8.99   | 8.28  | 6.31             | 8.67  | 8.29   | 7.76  | 6.19            |
| Poisson            | 7.78  | 7.66   | 7.06  | 5.33             | 7.32  | 7.02   | 6.64  | 5.23            |
| Worst-case         | 268   | 244    | 239   | 192             | 264   | 244    | 237   | 191            |

3) By substituting eqs. (12c) and (12d), we get

$$\Phi_{ue} - \Phi_{uw} \Phi_{xz}^{-1} \Phi_{xe} = KC\Phi_{xe} + K - KC\Phi_{xz}^{-1} \Phi_{xe}$$

Necessity: We show that lower triangular block-matrices $\Phi_{xz}$, $\Phi_{xe}$, $\Phi_{uw}$, and $\Phi_{ue}$ that satisfy eq. (13) and eq. (14) lead to a lower triangular block-matrix $K$ per eq. (10). To this end, eq. (13) can be written as

$$\Phi_{xw} (I - ZA) - ZBKC = 0,$$

$$\Phi_{uw} (I - ZA) - ZBKC = 0.$$
In light of the first step of the proof of [20, Theorem 4], we can equivalently write worst-case-regret$_F$(u) as

$$\min_{\lambda > 0} \lambda \text{ subject to } \lambda - (\Phi^T D \Phi - (\Phi^T D \Phi)^T) \succeq 0.$$  

(28)

Using Schur complement [33], we rewrite the constraint in eq. (28) in the form of eq. (16d).

The proof is now completed by following the steps of the proof of [21, Theorem 3] to reformulate the safety constraints as a linear matrix inequalities via dualization. □

References

[1] E. Ackerman, “Amazon promises package delivery by drone: Is it for real?” IEEE Spectrum, Web, 2013.
[2] S. Rajendran and S. Srinivas, “Air taxi service for urban mobility: A critical review of recent developments, future challenges, and opportunities,” Transportation Research Part E: Logistics and Transportation Review, vol. 143, p. 102090, 2020.
[3] A. Rivera, A. Villalobos, J. C. N. Monje, J. A. G. Mariñas, and C. M. Oppus, “Post-disaster rescue facility: Human detection and geolocation using aerial drones,” in IEEE 10 Conference, 2016, pp. 384–386.
[4] J. B. Rawlings, D. Q. Mayne, and M. Diehl, Model predictive control: Theory, computation, and design. Nob Hill Publishing, 2017, vol. 2.
[5] F. Borrelli, A. Bemporad, and M. Morari, Predictive control for linear and hybrid systems. Cambridge University Press, 2017.
[6] S. Thrun, “Probabilistic robotics,” Communications of the ACM, vol. 45, no. 3, pp. 52–57, 2002.
[7] R. G. Brown and P. Y. Hwang, “Introduction to random signals and communications of the ACM,” in IEEE 10 Conference, 2019, pp. 111–119.
[8] Y. Li, S. Das, and N. Li, “Online optimal control with affine constraints,” in AAAI Conference on Artificial Intelligence, vol. 35, no. 10, 2021, pp. 8527–8537.
[9] Y. Li, S. Das, J. Shamma, and N. Li, “Safe adaptive learning-based control for constrained linear quadratic regulators with regret guarantees,” arXiv preprint:2111.00411, 2021.
[10] M. Simchowitz, K. Singh, and E. Hazan, “Improper learning for non-stochastic control,” in C. on Learning Theory, 2020, pp. 3320–3436.
[11] P. Gradu, E. Hazan, and E. Minskyan, “Adaptive regret for control of time-varying dynamics,” arXiv preprint 2007:04393, 2020.
[12] G. Goel and B. Hassibi, “Regret-optimal control in dynamic environments,” arXiv preprint 2010.10473, 2020.
[13] O. Sabag, G. Goel, S. Lale, and B. Hassibi, “Regret-optimal full-information control,” arXiv preprint:2105.01244, 2021.
[14] G. Goel and B. Hassibi, “Regret-optimal measurement-feedback control,” in Learning for Dynamics and Control, 2021, pp. 1270–1280.
[15] A. Martin, L. Furieri, F. Dörfler, J. Lygeros, and G. Ferrari-Trecate, “Safe control with minimal regret,” in Learning for Dynamics and Control Conference, 2022, pp. 726–738.
[16] A. Didier, J. Sieber, and M. N. Zeilinger, “A system level approach to regret optimal control,” IEEE Control Systems Letters, 2022.
[17] S. Shalev-Shwartz et al., “Online learning and online convex optimization,” Foundations and Trends® in Machine Learning, vol. 4, no. 2, pp. 107–194, 2012.
[18] E. Hazan et al., “Introduction to online convex optimization,” Foundations and Trends® in Optimization, vol. 2, no. 3–4, pp. 157–325, 2016.
[19] M. Zinkevich, “Online convex programming and generalized infinitesimal gradient ascent,” in International Conf. on Machine Learning, 2003, pp. 928–936.
[20] B. Hassibi, A. H. Sayed, and T. Kailath, Indefinite-Quadratic estimation and control: A unified approach to $H_2$ and $H_\infty$ theories, 1999.