Safe Control of Partially-Observed Linear Time-Varying Systems with Minimal Worst-Case Dynamic Regret

Hongyu Zhou, Vasileios Tzoumas

Abstract— We present safe control of partially-observed 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 suboptimality 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 sensing capabilities. To synthesize our minimal-regret controller, we formulate a constrained semi-definite program based on a System Level Synthesis approach 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. Our algorithm is observed to have better performance than either or both the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers, demonstrating a *Best of Both Worlds* performance.

I. INTRODUCTION

In the future, robots will be leveraging their on-board control and sensing capabilities to complete tasks such as package delivery [1], transportation [2], and disaster response [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 sensing capabilities that are governed 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 measurement model is 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.

Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109 USA; {zhouhy, vtzoumas}@umich.edu

- **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 above challenges motivate the development of safe control algorithms for partially-observable linear time-varying systems, guaranteeing near-optimal control performance against unknown and unpredictable noise.

Related Work. The current control algorithms either consider (i) no safety constraints, or (ii) fully-observed systems or partially-observed systems but with time-invariant measurement models, or (iii) known stochastic models about the process and measurement noise, *e.g.*, Gaussian noise.

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]–[16], and then by reviewing *robust control* algorithms that select inputs based on simulating the future system dynamics across a lookahead horizon [17]–[21]:

a) Online Learning for Control: The algorithms performing online learning for control make no assumptions about the noise's stochasticity [22], [23], 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 static or dynamic regret: *static regret* [22], [23] captures the algorithms' suboptimality against optimal static policies that know the future a priori, *i.e.*, control policies with *time-invariant* control gains [12]–[15]; and *dynamic regret* [24] captures the algorithm's suboptimality against optimal dynamic policies that know the future a priori, *i.e.*, control policies with *time-varying* control gains [16].

The current static-regret algorithms consider no safety constraints, with the exception of [14], 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.[1]

b) Robust Control: The classical $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control algorithms [28] assume Gaussian noise and bounded worst-case noise, respectively. Particularly, $\mathcal{H}_2$ is optimal under Gaussian noise but it can thus underperform under non-Gaussian noise. $\mathcal{H}_\infty$ on the other hand can be conservative.

For this reason, recent robust control algorithms aim to address the challenge of unpredictable noise by focusing on convex optimization techniques that guarantee *regret optimality*, *i.e.*, minimal worst-case dynamic-regret among all noise realizations subject to a given total noise magnitude. [17], [18] focus on fully-observed systems and [19] on partially-observed systems. But these algorithms consider no safety constraints. Instead, [20], [21] provide a regret-optimal control algorithm that accounts for safety constraints. However, they focus on fully-observed systems only.

Contributions. In this paper, we provide an algorithm with dynamic regret guarantees for the safe control of partially-observed linear time-varying systems. 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. To this end, we formalize the problem of *Safe Control of Partially-Observed Linear Time-Varying Systems with Minimal Worst-Case Regret* (Problem 1).

Our analysis builds on [20], [29], [30]: (i) We prove that the output-feedback control gains are the solution of a constrained Semi-Definite Program (SDP); our SDP approach handles partially-observed systems, generalizing from the fully-observed case presented in [20]. (ii) We use a *System Level Synthesis* (SLS) method that extends the current SLS methods [29], [30] to partially-observable time-varying systems, presenting necessary and sufficient conditions for the existence of a causal safe output-feedback control policy.

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 (Section IV). We compare our algorithm with the safe $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control algorithms [20] 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 $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers. That is, our algorithm demonstrates robustness across all tested types of noise, being the best or the second best among $\mathcal{H}_2$ and $\mathcal{H}_\infty$, an advantageous performance capacity when the type of noise is unknown a priori and unpredictable. Instead, $\mathcal{H}_2$ is the worst against worst-case noise and $\mathcal{H}_\infty$ is the worst against Gaussian.

Organization. Section II formulates the problem of *safe control of partially-observed linear time-varying systems*

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 proofs are given in [31].

II. PROBLEM FORMULATION

We formulate the problem of *Safe Control of Partially-Observed Linear Time-Varying Systems with Minimal Worst-Case Regret* (Problem 1). We use the framework:

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

$$
\begin{aligned}
x_{t+1} &= A_t x_t + B_t u_t + w_t, \quad t \in \{0, \dots, T-1\}, \\
y_t &= C_t x_t + e_t,
\end{aligned}
\tag{1}
$$

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, $w_t \in \mathbb{R}^{d_x}$ is the process noise, $y_t \in \mathbb{R}^{d_y}$ is the measurement, and $e_t \in \mathbb{R}^{d_y}$ is the measurement noise.

We henceforth denote:
- $\mathbf{x} \triangleq \begin{bmatrix} x_0^\top, x_1^\top, \dots, x_{T-1}^\top \end{bmatrix}^\top$, *i.e.*, $\mathbf{x}$ is the state trajectory across the time horizon T;
- $\mathbf{u}$, $\mathbf{y}$, and $\mathbf{e}$ are defined correspondingly to $\mathbf{x}$;
- $\mathbf{w} \triangleq \begin{bmatrix} x_0^\top, w_0^\top, \dots, w_{T-2}^\top \end{bmatrix}^\top$, *i.e.*, is the process noise trajectory till time $T-2$ appended with the initial condition.

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 $\mathbf{w}_t$ and $\mathbf{e}_t$ are constrained in known compact polytopes that contain a neighborhood of the origin:* i.e.*, we are given $\mathbb{W} \triangleq \{\mathbf{w} \in \mathbb{R}^{d_x T} : \mathbf{H}_w \mathbf{w} \leq \mathbf{h}_w\}$ and $\mathbb{E} \triangleq \{\mathbf{e} \in \mathbb{R}^{d_y T} : \mathbf{H}_e \mathbf{e} \leq \mathbf{h}_e\}$ for given matrices $\mathbf{H}_w$, $\mathbf{H}_e$, $\mathbf{h_w}$ and $\mathbf{h_e}$.*

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

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

$$
\mathbf{H} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \leq \mathbf{h}, \ \forall \mathbf{w} \in \mathbb{W}, \ \forall \mathbf{e} \in \mathbb{E},
\tag{2}
$$

for given matrices $\mathbf{H}$ and $\mathbf{h}$.

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

$$
u_t = \sum_{k=0}^{t} K_{t,k} y_k, \quad t \in \{0, \dots, T-1\},
\tag{3}
$$

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 $\mathbf{w}$ and $\mathbf{e}$ a priori. In this paper, we consider that the clairvoyant policy minimizes a cost of the form

$$
\text{cost}(\mathbf{w}, \mathbf{e}, \mathbf{u}) \triangleq \mathbf{x}^\top \mathcal{Q} \mathbf{x} + \mathbf{u}^\top \mathcal{R} \mathbf{u},
\tag{4}
$$

where $\mathcal{Q} \succeq 0$ and $\mathcal{R} \succ 0$, and $\mathbf{x}$ is a function of the control input sequence $\mathbf{u}$ and the noise $\mathbf{w}$ and $\mathbf{e}$ per eq. (1). $\mathcal{Q}$ and

---

[1] We focused on related works that provide regret guarantees. Other related works include, *e.g.*, [25]–[27]. Particularly, [25] considers fully-observed systems with safety constraints; [26] considers partially-observed but time-invariant systems without safety constraints; [27] provides an online method for safe navigation in dynamic environments where the uncertainty is characterized by stochastic models.

$\mathcal{R}$ are assumed symmetric, without loss of generality. Then, the suboptimality of any (causal) control sequence $\mathbf{u}$ that is unaware of the noise realization $\mathbf{w}$ and $\mathbf{e}$ is captured by the

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

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

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 [19])**.** *Denote by r the minimum radius of a ball in $\mathbb{R}^{d_x T + d_u T}$ that encircles the noise's domain sets $\mathbb{W}$ and $\mathbb{E}$. Then,*

$$\text{worst-case-regret}_T(\mathbf{u}) \triangleq \max_{\|\mathbf{w}\|_2^2 + \|\mathbf{e}\|_2^2 \le r^2} \text{regret}_T(\mathbf{w}, \mathbf{e}, \mathbf{u}).$$
(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 Control of Partially-Observed Linear Time–Varying Systems with Minimal Worst-Case 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,*

$$\min_{\substack{K_{t,k}, \\ \forall(t,k) \in \{0,1,\ldots,T\}^2}} \text{worst-case-regret}_T(\mathbf{u})$$
(7a)

$$\text{subject to} \quad \textit{the system in eq. (1)};$$
(7b)
$$\textit{the safety constraints in eq. (2)};$$
(7c)
$$\textit{the control policy in eq. (3)}.$$
(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 present an SLS approach for partially-observed linear time-varying systems in Section III-A (Proposition 1).

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

We use the SLS approach to partially-observed LTV systems. Particularly, given a desired state trajectory $\mathbf{x}$, we present necessary and sufficient conditions for the existence of a control policy u_t per eq. (3). 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 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:

- $\mathcal{Z}$ denotes the block-matrix downshift operator, *i.e.*,

$$\mathcal{Z} \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{I} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{I} & \mathbf{0} \end{bmatrix};$$
(8)

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

$$\mathcal{A} \triangleq \text{blkdiag}(A_0, A_1, \ldots, A_{T-2}, 0_{d_x \times d_x});$$
$$\mathcal{B} \triangleq \text{blkdiag}(B_0, B_1, \ldots, B_{T-2}, 0_{d_x \times d_u});$$
$$\mathcal{C} \triangleq \text{blkdiag}(C_0, C_1, \ldots, C_{T-1});$$

- $\mathcal{K}$ is the lower triangular block-matrix such that eq. (3) takes the form $\mathbf{u} = \mathcal{K}\mathbf{y}$, *i.e.*,

$$\mathcal{K} \triangleq \begin{bmatrix} K_{0,0} & 0_{d_u \times d_y} & \ldots & 0_{d_u \times d_y} \\ K_{1,0} & K_{1,1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0_{d_u \times d_y} \\ K_{T-1,0} & K_{T-1,1} & \ldots & K_{T-1,T-1} \end{bmatrix}.$$
(9)

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

$$\mathbf{x} = \mathcal{Z}\mathcal{A}\mathbf{x} + \mathcal{Z}\mathcal{B}\mathbf{u} + \mathbf{w};$$
(10a)
$$\mathbf{y} = \mathcal{C}\mathbf{x} + \mathbf{e};$$
(10b)
$$\mathbf{u} = \mathcal{K}\mathbf{y},$$
(10c)

which can be equivalently written as

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{xw} & \mathbf{\Phi}_{xe} \\ \mathbf{\Phi}_{uw} & \mathbf{\Phi}_{ue} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{e} \end{bmatrix},$$
(11)

where:

$$\mathbf{\Phi}_{xw} = (\mathbf{I} - \mathcal{Z}\mathcal{A} - \mathcal{Z}\mathcal{B}\mathcal{K}\mathcal{C})^{-1},$$
(12a)
$$\mathbf{\Phi}_{xe} = \mathbf{\Phi}_{xw}\mathcal{Z}\mathcal{B}\mathcal{K},$$
(12b)
$$\mathbf{\Phi}_{uw} = \mathcal{K}\mathcal{C}\mathbf{\Phi}_{xw},$$
(12c)
$$\mathbf{\Phi}_{ue} = \mathcal{K}\mathcal{C}\mathbf{\Phi}_{xe} + \mathcal{K}.$$
(12d)

We refer to the matrix

$$\mathbf{\Phi} \triangleq \begin{bmatrix} \mathbf{\Phi}_{xw} & \mathbf{\Phi}_{xe} \\ \mathbf{\Phi}_{uw} & \mathbf{\Phi}_{ue} \end{bmatrix}$$

in eq. (11) as the *response matrix*.

Given the lower triangular block-matrix $\mathcal{K}$, eq. (11) captures how the noise trajectories $(\mathbf{w}, \mathbf{e})$ result to the control input trajectory $\mathbf{u}$ and, all in all, to the state trajectory $\mathbf{x}$. Particularly, $\mathbf{\Phi}_{xw}$, $\mathbf{\Phi}_{xe}$, $\mathbf{\Phi}_{uw}$, and $\mathbf{\Phi}_{ue}$ are computable *given $\mathcal{K}$*.

We next focus on the opposite direction: we present necessary and sufficient conditions for the existence of a lower triangular block-matrix $\mathcal{K}$ that always satisfies eq. (11), providing how to compute such a $\mathcal{K}$ given $\mathbf{\Phi}_{xw}$, $\mathbf{\Phi}_{xe}$, $\mathbf{\Phi}_{uw}$, and $\mathbf{\Phi}_{ue}$, instead of the other way around.

Proposition 1 (System Level Synthesis for Partially-Observed LTV Systems)**.** *There exists a lower triangular block-matrix $\mathcal{K}$ such that eq. (11) holds true if and only if $\mathbf{\Phi}_{xw}$, $\mathbf{\Phi}_{xe}$, $\mathbf{\Phi}_{uw}$, and $\mathbf{\Phi}_{ue}$ are:*

Algorithm 1: Safe Control of Partially-Observed Linear Time-Varying Systems with Minimal Worst-Case Regret.

Input: Time horizon T; system matrices $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$; cost matrices $\mathcal{Q}$ and $\mathcal{R}$; noise's domain sets $\mathbb{W}$ and $\mathbb{E}$; upper bound r to the noise' total magnitude.
Output: Output-feedback control gains $\mathcal{K}$ for eq. (3)'s control policy.

1: $\{\boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{xe}, \boldsymbol{\Phi}_{uw}, \boldsymbol{\Phi}_{ue}\} \leftarrow$ Solve the Semi-Definite Program in Theorem 1;
2: $\mathcal{K} \leftarrow \boldsymbol{\Phi}_{ue} - \boldsymbol{\Phi}_{uw}\boldsymbol{\Phi}_{xw}^{-1}\boldsymbol{\Phi}_{xe}$ per Proposition 1.

- *lower triangular block-matrices; and*
- *lie in the affine subspace*

$$\begin{bmatrix} \mathbf{I} - \mathcal{Z}\mathcal{A} & -\mathcal{Z}\mathcal{B} \end{bmatrix} \boldsymbol{\Phi} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}, \quad (13a)$$

$$\boldsymbol{\Phi} \begin{bmatrix} \mathbf{I} - \mathcal{Z}\mathcal{A} \\ -\mathcal{C} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}. \quad (13b)$$

Also, $\mathcal{K}$ is computed given $\boldsymbol{\Phi}_{xw}$, $\boldsymbol{\Phi}_{xe}$, $\boldsymbol{\Phi}_{uw}$, and $\boldsymbol{\Phi}_{ue}$ via:

$$\mathcal{K} = \boldsymbol{\Phi}_{ue} - \boldsymbol{\Phi}_{uw}\boldsymbol{\Phi}_{xw}^{-1}\boldsymbol{\Phi}_{xe}. \quad (14)$$

The proof follows similar steps as in [30, Theorem 2.1]. It differs from [30, Theorem 2.1] in the way that the system is partially observed, *i.e.*, $\mathbf{y} = \mathcal{C}\mathbf{x} + \mathbf{e}$, and the control input now depends on the measurement models $\mathcal{C}$ and measurement noise, *i.e.*, $\mathbf{u} = \mathcal{K}\mathbf{y} = \mathcal{K}\mathcal{C}\mathbf{x} + \mathcal{K}\mathbf{e}$.

By finding a $\boldsymbol{\Phi}$ satisfying Proposition 1's constraints, we equivalently find a lower triangular block-matrix $\mathcal{K}$, and thus a control policy per eq. (3), satisfying eq. (11).

B. Algorithm for Problem 1 via SDP Reformulation

We provide an algorithm for Problem 1 (Algorithm 1). To this end, we reformulate Problem 1 as an SDP (Theorem 1). We obtain the reformulation via the steps:

- We change the optimization variables in Problem 1 from the output-feedback control gains in $\mathcal{K}$ to the response matrix $\boldsymbol{\Phi}$. We thus leverage that finding a feasible $\boldsymbol{\Phi}$ 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 constraints. Once a $\boldsymbol{\Phi}$ is found, then $\mathcal{K}$ 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 [20, Theorem 3];
- We reformulate Problem 1's objective function, namely, worst-case-regret$_T(\mathbf{u})$, 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 [19, 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:

- $\mathcal{D} \triangleq \text{blkdiag}(\mathcal{Q}, \mathcal{R})$, *i.e.*, $\mathcal{D}$ is the diagonal block-matrix whose elements $\mathcal{Q}$ and $\mathcal{R}$ define the cost in eq. (4);
- $\mathbf{Z}$ are the dual variables introduced to reformulate the safety constraints in eq. (2) as linear 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$_T(\mathbf{u})$;

- $\boldsymbol{\Phi}^c$ is the response corresponding to the optimal clairvoyant controller in eq. (5) that ignores the safety constraints; *i.e.*, per [20], [32],[2]

$$\boldsymbol{\Phi}^c \in \underset{\boldsymbol{\Phi}}{argmin} \quad \left\| \begin{bmatrix} \mathcal{Q}^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathcal{R}^{\frac{1}{2}} \end{bmatrix} \boldsymbol{\Phi} \right\|_F^2 \quad (15)$$

$$\textit{subject to} \quad \text{eq. (13a)}, \ \boldsymbol{\Phi}_{xe} = \mathbf{0}, \ \boldsymbol{\Phi}_{ue} = \mathbf{0}.$$

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

$$\min_{\boldsymbol{\Phi}, \mathbf{Z}, \lambda} \quad \lambda \quad \text{subject to:}$$

$$\boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{xe}, \boldsymbol{\Phi}_{uw}, \boldsymbol{\Phi}_{ue} \textit{ being lower block triangular;}$$
$$(16a)$$

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

$$\mathbf{Z}^{\top} \begin{bmatrix} \mathbf{h}_w \\ \mathbf{h}_e \end{bmatrix} \leq \mathbf{h}, \ \mathbf{H}\boldsymbol{\Phi} = \mathbf{Z}^{\top} \begin{bmatrix} \mathbf{H}_w & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_e \end{bmatrix}, \ \mathbf{Z}_{ij} \geq 0;$$
$$(16c)$$

$$\lambda > 0, \ \begin{bmatrix} \mathbf{I} & \mathcal{D}^{\frac{1}{2}}\boldsymbol{\Phi} \\ \boldsymbol{\Phi}^{\top}\mathcal{D}^{\frac{1}{2}} & \lambda\mathbf{I} + (\boldsymbol{\Phi}^c)^{\top}\mathcal{D}\boldsymbol{\Phi}^c \end{bmatrix} \succeq 0. \quad (16d)$$

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 $\mathcal{K}$ to the response matrix $\boldsymbol{\Phi}$, per the necessary and sufficient conditions in Proposition 1; eq. (16c) results from the reformulation of the safety constraints in eq. (2) as linear inequalities, per the dualization procedure in [20, Proof of Theorem 3]; and the new objective of minimizing the scalar λ subject to eq. (16d) result from the reformulation of Problem 1's objective function, that is, of worst-case-regret$_T(\mathbf{u})$.

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 response matrix $\boldsymbol{\Phi}$ (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 control of partially-observed LTV systems 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).

Our code is open-sourced at: https://github.com/UM-iRaL/Safe-POLTV-MinRegret.

---

[2]By solving eq. (15), $\boldsymbol{\Phi}^c$ can be clairvoyant since eq. (15) does *not* require $\boldsymbol{\Phi}_{xw}$, $\boldsymbol{\Phi}_{xe}$, $\boldsymbol{\Phi}_{uw}$, and $\boldsymbol{\Phi}_{ue}$ to be lower triangular block-matrices; that is, the necessary and sufficient condition of Proposition 1 are *not* both met. Thus, $\boldsymbol{\Phi}^c$ can correspond to a control gain block-matrix $\mathcal{K}$ that fails to be lower triangular and, as a result, the corresponding control policy u_t will depend on future measurements $\{y_k\}_{k \in \{t+1, \ldots, T\}}$.

Compared Algorithms. We compare Algorithm 1 with the safe $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers [20]. The clairvoyant controller is obtained by solving eq. (15).

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 better than $\mathcal{H}_2$ and $\mathcal{H}_\infty$ across the tested types of noise, or it performs better than $\mathcal{H}_2$ or $\mathcal{H}_\infty$ across all tested types of noise.

We performed all simulations in MATLAB with YALMIP toolbox [33] and MOSEK solver [34].

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

A. Synthetic Partially-Observed LTV Systems

Simulation Setup. We consider LTV systems such that

$$A_t = \rho \begin{bmatrix} 0.7 & 0.2 & 0 \\ 0.3 & 0.7 & -0.1 \\ 0 & -0.2 & 0.8 \end{bmatrix}, B_t = \begin{bmatrix} 1 & 0.2 \\ 2 & 0.3 \\ 1.5 & 0.5 \end{bmatrix},$$

$$C_t = \begin{cases} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, & t = \{1, 3, \ldots\}; \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & t = \{2, 4, \ldots\}, \end{cases}$$

where ρ 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 $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers. Under all other noise types: for $\rho = 0.85$, Algorithm 1 outperforms the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers; for $\rho = 1.05$, Algorithm 1 outperforms the $\mathcal{H}_\infty$ 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 quadrotors goal is stay at a predefined hovering position. To this end, we focus on its linearized dynamics, taking the form

$$A_t = \begin{bmatrix} 1 & 0 & 0 & 0.1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, B_t = \begin{bmatrix} -\frac{4.91}{100} & 0 & 0 \\ 0 & \frac{4.91}{100} & 0 \\ 0 & 0 & \frac{1}{200} \\ -\frac{98.1}{100} & 0 & 0 \\ 0 & \frac{98.1}{100} & 0 \\ 0 & 0 & \frac{1}{10} \end{bmatrix}.$$

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 [35]. Formally,

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

We choose the safety constraints: $-5 \leq x_t \leq 5$ and $[-\pi \ -\pi \ -20]^\top \leq u_t \leq [\pi \ \pi \ 20]^\top$; 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 $T = \{2, 3, \ldots, 25\}$.

Results. The results are summarized in Figure 3. Algorithm 1's performance lies between that of the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers under Gaussian and worst-case noise. Under all other noise types, Algorithm 1 always outperforms the $\mathcal{H}_2$ controller, and is on par with the $\mathcal{H}_\infty$ controller up to horizon around $T = 15$. All in all, Algorithm 1 demonstrates a superior or BoBW performance across all scenarios.

V. CONCLUSION

Summary. We provided an algorithm for the 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 System Level Synthesis approach for partially-observed LTV systems. We validated Algorithm 1 in simulated scenarios; the algorithm was observed to be better or on par with either the $\mathcal{H}_2$ or $\mathcal{H}_\infty$ 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), which is a standard assumption in state-of-the-art control algorithms [4], [5], [17]–[21]. 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 [36]. 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
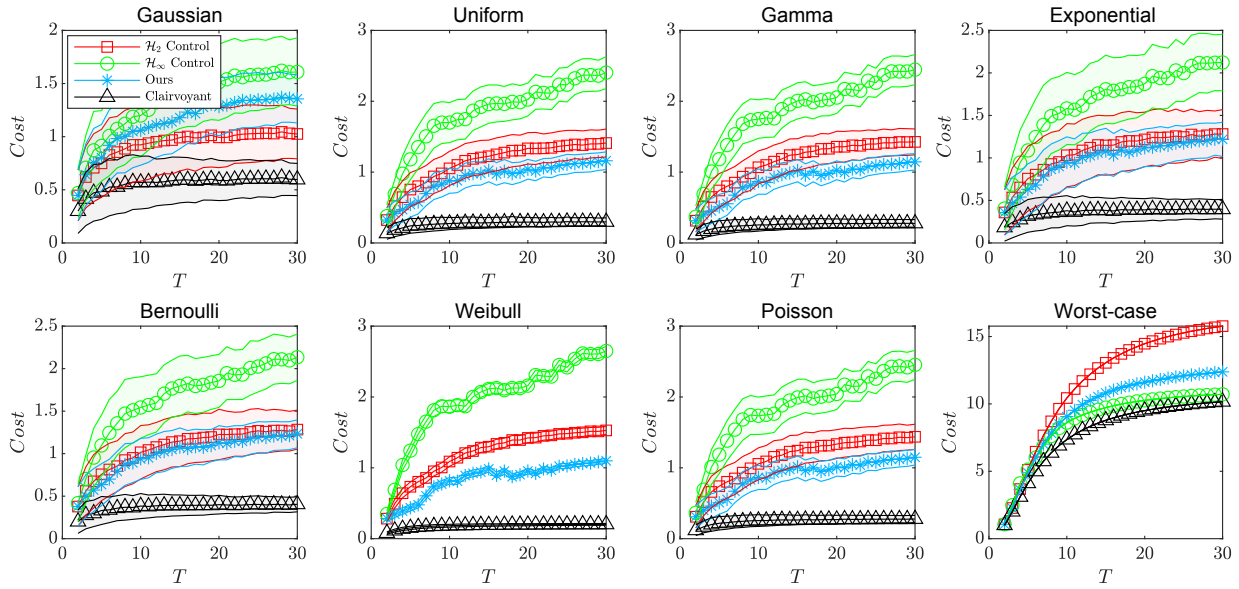
Fig. 1: **Performance Comparison given the LTV System in Section IV-A with** $\rho = 0.85$ **(stable system case).** The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.
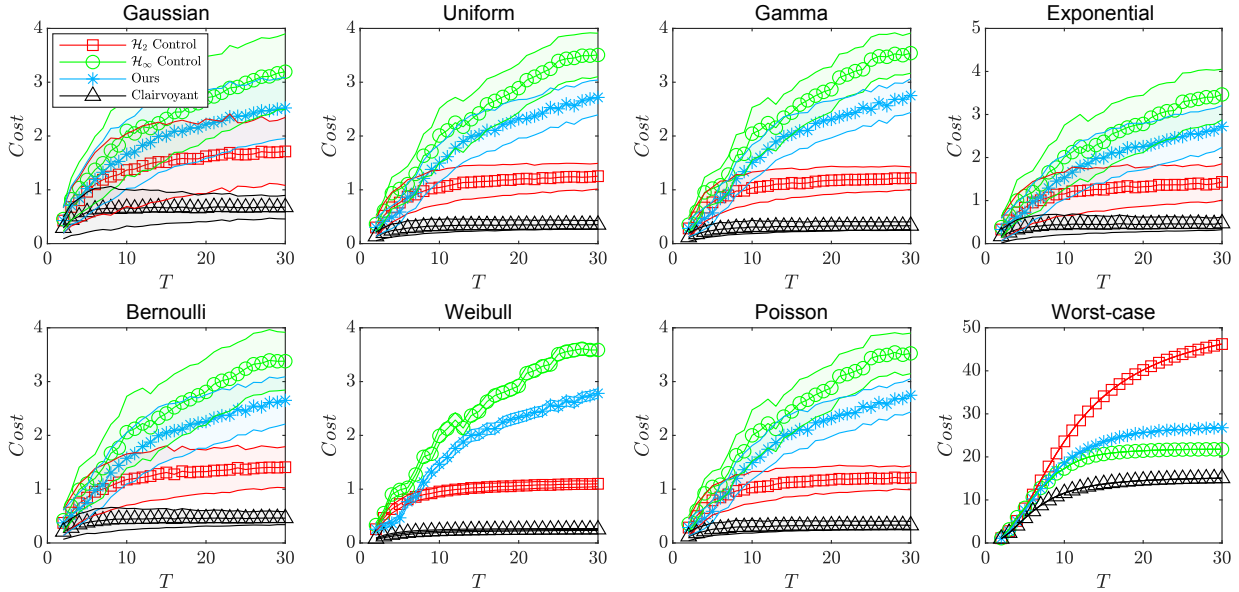

Fig. 2: **Performance Comparison given the LTV System in Section IV-A with** $\rho = 1.05$ **(unstable system case).** The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.

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.

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 applied Kalman filtering," 1997.

[8] O. Faltinsen, *Sea loads on ships and offshore structures*. Cambridge University Press, 1993, vol. 1.

[9] T. P. Sapsis, "Statistics of extreme events in fluid flows and waves," *Annual Reviews*, 2021.

[10] K. J. Åström, *Introduction to stochastic control theory*. Courier Corporation, 2012.

[11] F. Berkenkamp, "Safe exploration in reinforcement learning: Theory and applications in robotics," Ph.D. dissertation, ETH Zurich, 2019.
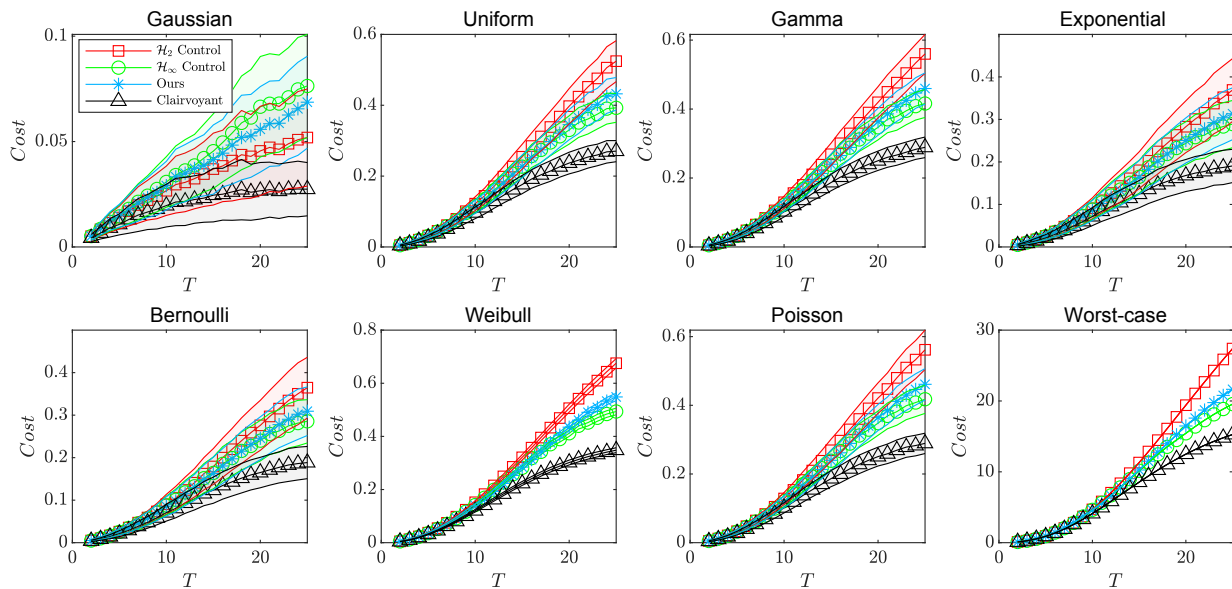
Fig. 3: **Performance Comparison given the Quadrotor System in Section IV-B.** The performance is quantified per the cost in eq. (4). Each figure corresponds to a different process and measurement noise type. The shaded areas represent standard deviation.

[12] E. Hazan, S. Kakade, and K. Singh, "The nonstochastic control problem," in *Algorithmic Learning Theory*, 2020, pp. 408–421.

[13] N. Agarwal, B. Bullins, E. Hazan, S. Kakade, and K. Singh, "Online control with adversarial disturbances," in *International Conference on Machine Learning*, 2019, pp. 111–119.

[14] 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.

[15] M. Simchowitz, K. Singh, and E. Hazan, "Improper learning for nonstochastic control," in *C. on Learning Theory*, 2020, pp. 3320–3436.

[16] P. Gradu, E. Hazan, and E. Minasyan, "Adaptive regret for control of time-varying dynamics," *arXiv preprint 2007.04393*, 2020.

[17] G. Goel and B. Hassibi, "Regret-optimal control in dynamic environments," *arXiv preprint:2010.10473*, 2020.

[18] O. Sabag, G. Goel, S. Lale, and B. Hassibi, "Regret-optimal full-information control," *arXiv preprint:2105.01244*, 2021.

[19] G. Goel and B. Hassibi, "Regret-optimal measurement-feedback control," in *Learning for Dynamics and Control*, 2021, pp. 1270–1280.

[20] 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.

[21] A. Didier, J. Sieber, and M. N. Zeilinger, "A system level approach to regret optimal control," *IEEE Control Systems Letters*, 2022.

[22] 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.

[23] E. Hazan *et al.*, "Introduction to online convex optimization," *Foundations and Trends in Optimization*, vol. 2, no. 3-4, pp. 157–325, 2016.

[24] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *International Conf. on Machine Learning*, 2003, pp. 928–936.

[25] Y. Chen and J. Anderson, "System level synthesis with state and input constraints," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 5258–5263.

[26] M. Colombino, E. Dall'Anese, and A. Bernstein, "Online optimization as a feedback controller: Stability and tracking," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 422–432, 2019.

[27] H. Chen, S. Feng, Y. Zhao, C. Liu, and P. A. Vela, "Safe hierarchical navigation in crowded dynamic uncertain environments," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 1174–1181.

[28] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic estimation and control: A unified approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theories*, 1999.

[29] Y.-S. Wang, N. Matni, and J. C. Doyle, "A system-level approach to controller synthesis," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4079–4093, 2019.

[30] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.

[31] H. Zhou and V. Tzoumas, "Safe control of partially-observed linear time-varying systems with minimal worst-case dynamic regret," *arXiv preprint arXiv:2208.08929*, 2022.

[32] G. Goel and B. Hassibi, "Measurement-feedback control with optimal data-dependent regret," *arXiv preprint arXiv:2209.06425*, 2022.

[33] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *CACSD Conference*, 2004.

[34] *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.* [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html

[35] A. Goel, A. U. Islam, A. Ansari, O. Kouba, and D. S. Bernstein, "An introduction to inertial navigation from the perspective of state estimation," *IEEE Control Systems Magazine*, vol. 41, no. 5, pp. 104–128, 2021.

[36] Y. Ma, S. Soatto, J. Košecká, and S. Sastry, *An invitation to 3-D vision: From images to geometric models.* Springer, 2004, vol. 26.