Time-Varying Optimization of LTI Systems via Projected Primal-Dual Gradient Flows

Gianluca Bianchin, Jorge Cortés, Jorge. I. Poveda, and Emiliano Dall'Anese

***Abstract*—This paper investigates the problem of regulating in real time a linear dynamical system to the solution trajectory of a time-varying constrained convex optimization problem. The proposed feedback controller is based on an adaptation of the saddle-flow dynamics, modified to take into account projections on constraint sets and output-feedback from the plant. We derive sufficient conditions on the tunable parameters of the controller (inherently related to the time-scale separation between plant and controller dynamics) to guarantee exponential and input-to-state stability of the closed-loop system. The analysis is tailored to the case of time-varying strongly convex cost functions and polytopic output constraints. The theoretical results are further validated in a ramp metering control problem in a network of traffic highways.**

I. Introduction

THIS paper investigates the problem of *online optimization* of linear time-invariant (LTI) systems. The objective is to design an output feedback controller that steers the inputs and outputs of the plant towards the solution trajectory of a *time-varying* optimization problem (see Fig. 1). Such problems correspond to scenarios with cost and constraints that may change over time to reflect dynamic performance objectives or simply take into account time-varying unknown exogenous inputs to the system (henceforth, the term "static" is used to refer to optimization problems with time-invariant cost and constraints, and with a constant disturbance in the system). This setting emerges in many engineering applications, including power systems, transportation networks, and communication systems.

The design of feedback controllers inspired from optimization algorithms has received significant attention during the last decade [1]–[10]. While most of the existing works focus on the design of optimization-based controllers for static problems [1]–[4], [6], [8], [9], or consider unconstrained time-varying problems [7], [10], an open research question is whether controllers can be synthesized to track solutions trajectories of time-varying problems with input and output constraints. Towards this direction, this paper focuses on the synthesis of output-feedback controllers by leveraging online primal-dual dynamics. In this context, even though [11]–[14] show that primal-dual dynamics for time-invariant optimization problems have an exponential rate of convergence, the main challenges here are to derive exponential stability results for problems that are time-varying and where primal-dual dynamics (and the projected counterpart) are interconnected with a dynamical system subject to unknown exogenous disturbances, as in Fig. 1.

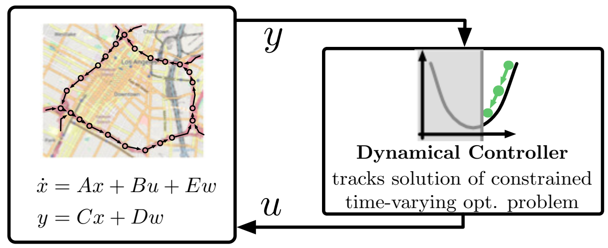


Fig. 1. Online saddle-flow optimizer used as an output feedback controller for LTI systems subject to unknown time-varying disturbances.

We consider problems with a time-varying strongly convex cost, time-varying linear constraints on the output, and convex constraints on the input. We present exponential stability and input-to-state stability (ISS) analysis of online saddle-flow controllers interconnected with a stable LTI system (with unknown exogenous inputs); in particular, we leverage tools from singular perturbation theory [15] to provide sufficient conditions on the tunable parameters of the controller (related to the time-scale separation between plant and controller) to guarantee exponential stability and tracking of the optimal solution trajectory.

Prior works. Asymptotic stability results are provided for static optimization problems using gradient flows in [3], [4], [6], [8], and semi-global practical results are presented in [16] for general hybrid model-free controllers; constraints on the inputs are dealt with projected gradient flows in [17], [18]. Controllers conceptually related to the continuous-time Arrow-Hurwicz-Uzawa algorithm are used for static problems with convex constraints on the system outputs in [1], whereas saddle-point flows are studied in [2], [8], [9], [19] and [20, Sec. 3]. For time-varying unconstrained optimization problems, prediction-correction algorithms are used in [7]. On the other hand, exponential stability results for dynamic controllers based on gradient flows and accelerated hybrid algorithms are presented in [10]. When considering constraints on the output, primal-dual dynamics based on the Moreau envelope are studied in [5]. In terms of classes of plants, stable LTI systems are considered in e.g., [5], [6], [10], stable nonlinear systems in [8], input-linearizable systems in [7], and input affine nonlinear system in [2]. We also acknowledge recent works in online implementations of model predictive control (MPC); see [21], [22].

Contributions. The main contributions of this work can be summarized as follows: C1) We design an output feedback controller, inspired on primal-dual dynamics, able to steer the plant inputs and outputs to the solution trajectory of the time-varying optimization problem without requiring information or measurements of the external disturbances acting on the plant dynamics. For problems with equality constraints, the primal-dual controller is designed based on the standard Lagrangian function. Instead, for problems with inequality constraints, we employ a regularized Lagrangian, cf. [23], to achieve exponential stability of the approximate KKT trajectory; C2) we incorporate input constraints by designing a novel projected primal-dual output feedback controller; in particular, our controllers are Lipschitz continuous and have continuously differentiable trajectories, which simplifies their analysis and allows us to directly establish additional robustness certificates; C3) we show that the proposed framework is applicable to more general LTI systems, including switched systems with common quadratic Lyapunov functions; and finally, C4) we apply the proposed controllers to solve ramp metering problems for the control of traffic systems. We compare our results with state-of-the-art controllers, including ALINEA [24] and MPC, illustrating the advantages of our method.

We emphasize that, relative to [5], our sufficient conditions are markedly easier to check a priori and do not require to numerically solve a linear matrix inequality. Moreover, our treatment also accounts for input constraints and does not require the computation of the Moreau envelope. Relative to, e.g., [1], [2], [8], [9], we consider time-varying problems and system disturbances, and we offer exponential stability and ISS results. Relative to [25], [26], we investigate the stability of saddle-point flows when interconnected with an LTI system.

The rest of this paper is organized as follows. Section II presents the problem formulation. Section III develops a projected primal-dual output feedback controller for problems with input constraints and output inequality constraints based on a regularized Lagrangian function. Section IV considers problems with output equality constraints. Section V presents numerical results by addressing on a traffic ramp metering problem. Finally, we present our conclusions in Section VI.[1]

II. Problem Formulation

We consider LTI dynamical systems given by:

$$\begin{aligned} \dot{x} &= Ax + Bu + Ew_t, \\ y &= Cx + Dw_t, \end{aligned} \tag{1}$$

where $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is the state, $u : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ is the input, $y : \mathbb{R}_{\geq 0} \to \mathbb{R}^p$ is the output, and $t \mapsto w_t \in \mathbb{R}^q$ is an unknown and time-varying exogenous input or disturbance (the notation w_t emphasizes the dependence on time). Throughout the paper, we make the following stability assumption on the plant.

Assumption 2.1: There exists positive definite matrices $P_x, Q_x \in \mathbb{R}^{n\times n}$ such that $A^\mathsf{T} P_x + P_x A = -Q_x$. Moreover, the columns of C are linearly independent. □

Under Assumption 2.1, matrix A is Hurwitz and, given constant vectors $u_{\text{eq}} \in \mathbb{R}^m$, $w_{\text{eq}} \in \mathbb{R}^q$, system (1) has a unique exponentially stable equilibrium point $x_{\text{eq}} = -A^{-1}(Bu_{\text{eq}} + Ew_{\text{eq}})$. Moreover, at equilibrium, the relationship between system inputs and outputs is given by the algebraic relationship:

$$y_{\text{eq}} = \underbrace{-CA^{-1}B}_{:=G} u_{\text{eq}} + \underbrace{(D - CA^{-1}E)}_{:=H} w_{\text{eq}}. \tag{2}$$

Given any time-varying and unknown exogenous input w_t to (1), we focus on the problem of regulating the plant to the solutions of the time-varying optimization problem:

$$(u_t^*, y_t^*) \in \underset{\bar{u}\in\mathcal{U},\ \bar{y}\in\mathbb{R}^p}{\arg\min} \quad \phi_t(\bar{u}) + \psi_t(\bar{y}) \tag{3a}$$

$$\text{s.t.} \quad \bar{y} = G\bar{u} + Hw_t \tag{3b}$$

$$K_t\, \bar{y} \leq e_t, \tag{3c}$$

where for all $t \in \mathbb{R}_{\geq 0}$, $\phi_t : \mathbb{R}^m \to \mathbb{R}$, $\psi_t : \mathbb{R}^p \to \mathbb{R}$, $t \mapsto K_t \in \mathbb{R}^{r\times p}$ and $t \mapsto e_t \in \mathbb{R}^r$ describe a time-varying output constraint, and $\mathcal{U} \subseteq \mathbb{R}^m$ is a convex set describing an input constraint. Problem (3) formalizes a regulation problem, where the objective is to select an optimal input-output pair (u_t^*, y_t^*) that minimizes the cost specified by the functions ϕ_t and ψ_t. We note that, because the cost functions and constraints are time-varying, the optimal solutions of (3) describe optimal trajectories. We impose the following regularity assumptions on the temporal evolution of problem (3).

Assumption 2.2: The following properties hold.

(a) For all $t \in \mathbb{R}_{\geq 0}$, the functions $u \mapsto \phi_t(u)$ and $y \mapsto \psi_t(y)$ are continuously differentiable.

(b) The function $u \mapsto \phi_t(u)$ is μ_u-strongly convex, uniformly in t.

(c) There exist $\ell_u, \ell_y > 0$ such that for every $u, u' \in \mathbb{R}^m$ and $y, y' \in \mathbb{R}^p$, $\|\nabla\phi_t(u) - \nabla\phi_t(u', t)\| \leq \ell_u \|u - u'\|$, $\|\nabla\psi_t(y) - \nabla\psi_t(y')\| \leq \ell_y \|y - y'\|$, uniformly in t.

(d) For all $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, $t \mapsto \nabla\phi_t(u)$ and $t \mapsto \nabla\psi_t(y)$ are locally Lipschitz. □

Assumption 2.3: Problem (3) is feasible, and Slater's condition holds for each $t \in \mathbb{R}_{\geq 0}$. □

Assumption 2.4: The map $t \mapsto w_t$ is absolutely continuous and locally Lipschitz. □

Assumption 2.5: The functions $t \mapsto [K_t]_{ij}$ and $t \mapsto [e_t]_i$ $i = 1, \dots, r$, $j = 1, \dots, p$, are locally Lipschitz. There exists $\overline{K} \in \mathbb{R}_{\geq 0}$, $\bar{e} \in \mathbb{R}_{\geq 0}$, such that $\|K_t\| < \bar{K}$ and $\|e_t\| < \bar{e}$. □

Under Assumptions 2.2-2.3, the minimizers (u_t^*, y_t^*) of (3) are unique for any $t \in \mathbb{R}_{\geq 0}$. Assumptions 2.4–2.5 guarantee that the exogenous inputs and the constraints of (3) vary continuously with time. Moreover, we let $z_t^* := (u_t^*, \lambda_t^*)$ denote the *saddle-point* of the Lagrangian function associated with problem (3) or its regularized version, as explained shortly, and we impose the following regularity requirement.

Assumption 2.6: The (possibly regularized) saddle-point $t \mapsto z_t^*$ is absolutely continuous and locally Lipschitz. □

Under Assumptions 2.4 and 2.6, the maps $t \mapsto w_t$, $t \mapsto z_t^*$ are differentiable almost everywhere (a.e.), and their derivatives are essentially bounded. In what follows, $\dot{z}_t^*$ and $\dot{w}_t$ denote the distributional derivatives of z_t^* and of w_t, respectively.

We focus on the problem of developing a dynamical output feedback controller with state $z := (u, \lambda)$ and dynamics:

$$\dot{z} = \mathcal{C}(z, y),$$

[1] **Notation.** Given two vectors x and u, we use $(x, u) \in \mathbb{R}^{n+m}$ to denote their concatenation. We use $\bar{\lambda}(P)$ and $\underline{\lambda}(P)$ to denote the largest, and the smallest, eigenvalues of a square matrix P, respectively. We also use ess sup to denote that essential supremum. Finally, $P_\Omega : \mathbb{R}^\sigma \to \mathbb{R}^\sigma$ denotes the Euclidean projection of z onto a closed convex set $\Omega \subseteq \mathbb{R}^\sigma$, namely $P_\Omega(z) := \arg\min_{v\in\Omega} \|z - v\|$. For $u \in \mathbb{R}^n$, we denote by $[u]_i$ the i-th entry of u, where $i \in \{1, \dots, n\}$; for a matrix $U \in \mathbb{R}^{n\times m}$, we use $[U]_{ij}$ to denote the entry (i, j) of U.

with the following behavior when interconnected in closed-loop with the system (1); see Fig. 1.

Problem 1: Let $\xi := (x, u, \lambda)$ denote the plant and controller state, and let $\xi_t^* := (x_t^*, u_t^*, \lambda_t^*)$ where (u_t^*, λ_t^*) is the saddle-point of (3) and $x_t^* = -A^{-1}(Bu_t^* + Hw_t)$. Design an output-feedback controller $\mathcal{C}(z, y)$ such that for any $t_0 \in \mathbb{R}_{\geq 0}$:

$$\begin{aligned}\|\xi(t) - \xi_t^*\| &\leq a\|\xi(t_0) - \xi_{t_0}^*\|e^{-b(t-t_0)} \\ &\quad + \gamma_z \operatorname*{ess\,sup}_{\tau \geq t_0} \|\dot{z}_\tau^*\| + \gamma_w \operatorname*{ess\,sup}_{\tau \geq t_0} \|\dot{w}_\tau\|,\end{aligned} \tag{4}$$

for all $t \geq t_0$, and for some $a, b, \gamma_z, \gamma_w > 0$. □

The bound (4) establishes exponential tracking of the time-varying signal ξ_t^*, implicitly defined as the solution of (3), up to an asymptotic error that depends on the temporal variability of both the optimal trajectory and the unknown disturbance. Finally, we observe that when the optimization problem is static and w_t is constant, $\dot{z}_t^* = 0$ and $\dot{w}_t = 0$ at all times, and thus (4) boils down to an exponential stability bound.

III. Closed-loop Projected Saddle-point Flows

To solve problem (3), we consider the following augmented Lagrangian function:

$$\mathcal{L}_{\nu,t}(u, \lambda) := \mathcal{L}_t(u, \lambda) - \frac{\nu}{2}\|\lambda\|^2, \tag{5}$$

where $\mathcal{L}_t(u, \lambda) = \phi_t(u) + \psi_t(Gu + Hw_t) + \lambda^\mathsf{T}(K_t(Gu + Hw_t) - e_t)$ is the Lagrangian function and $\nu \in \mathbb{R}_{>0}$ is a tunable parameter. In what follows, we distinguish between $z_t^* = (u_t^*, \lambda_t^*)$ (the unique saddle-point of $\mathcal{L}_t(u, \lambda)$) and $z_{\nu,t}^* := (u_{\nu,t}^*, \lambda_{\nu,t}^*)$ (the unique saddle point of (5)). The main benefit of using an augmented Lagrangian is that it induces a saddle-point map that is strongly monotone, uniformly in time, which allows us to prove exponential convergence. However, the use of an augmented Lagrangian comes at the cost of shifting the saddle-points of the Lagrangian $\mathcal{L}_t(u, \lambda)$. The following lemma, adapted from [23, Prop. 3.1], provides a bound (point-wise in time) for the regularization error.

Lemma 3.1: Let Assumptions 2.2-2.3 hold. For each $t \in \mathbb{R}_{\geq 0}$, the following bound holds:

$$\mu_u\|u_{\nu,t}^* - u_t^*\|^2 + \frac{\nu}{2}\|\lambda_{\nu,t}^*\|^2 \leq \frac{\nu}{2}\|\lambda_t^*\|^2. \tag{6}$$

In particular, it holds that $\|u_{\nu,t}^* - u_t^*\| \leq \sqrt{\frac{\nu}{2\mu_u}}\|\lambda_t^*\|$.

Remark 3.2: Lemma 3.1 asserts that the error induced by the regularization term is bounded by the norm of the optimal multipliers of the non-regularized problem. Consequently, when the optimal solution strictly satisfies the constraints, then $\lambda_t^* = 0$ and the solution $u_{\nu,t}^*$ coincides with u_t^*.
To solve Problem 1, we consider the following modification of the gradients of the augmented Lagrangian:

$$L_{u,t}(u, y, \lambda) := \nabla\psi_t(u) + G^\mathsf{T}\nabla\phi_t(y) + G^\mathsf{T}K_t^\mathsf{T}\lambda, \tag{7a}$$

$$L_{\lambda,t}(y, \lambda) := K_t y - e_t - \nu\lambda, \tag{7b}$$

where we note that, with respect to the gradients of $\mathcal{L}_{\nu,t}$, in $L_{u,t}$ and $L_{\lambda,t}$ the steady-state map $Gu + Hw_t$ has been replaced by the variable y. Using (7), we propose the following *online projected primal-dual controller* for (1), cf. Fig. 1:

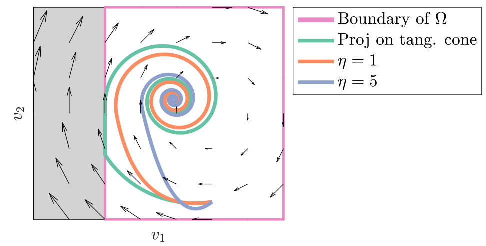


Fig. 2. Comparison between trajectories of (9) and of the smooth projection (8) for a 2-D vector field. Black arrows show the vector field.

$$\varepsilon\dot{x} = Ax + Bu + Ew_t, \qquad y = Cx + Dw_t, \tag{8a}$$

$$\dot{u} = P_{\mathcal{U}}\big(u - \eta L_{u,t}(u, y, \lambda)\big) - u, \tag{8b}$$

$$\dot{\lambda} = P_{\mathcal{C}}\big(\lambda + \eta L_{\lambda,t}(y, \lambda)\big) - \lambda, \tag{8c}$$

where $\varepsilon, \eta > 0$ are plant and controller gains that induce a time-scale separation between the plant and the controller, $v \mapsto P_\Omega(v)$ denotes the Euclidean projection onto the convex set Ω, and $\mathcal{C} := \mathbb{R}_{\geq 0}^r$. Three important observations on (8b)-(8c) are in order. First, the proposed controller utilizes instantaneous output-feedback from the plant. Second, the controller does not require any knowledge regarding the exogenous disturbance w_t. Third, even when the LTI system and the saddle-flow dynamics are stable in an open-loop configuration, the interconnection (8) is not guaranteed to be stable.

Remark 3.3: We note that the choice of dualizing the constraint $K_t y \leq e_t$ allows us to naturally enforce constraints that are time-varying. In contrast, when $K_t y \leq e_t$ is recast as a convex constraint of the form $u \in \mathcal{U}$ (by substituting (3b)), the presence of a time-varying pair (K_t, e_t) as well as the lack of knowledge of w_t imply that the constraint set becomes unknown and time-varying (and, even when w_t is known, time-varying projections may be computationally expensive). □

Remark 3.4: Given a closed convex set $\Omega \subseteq \mathbb{R}^\sigma$ and a vector field $F : \Omega \to \mathbb{R}^\sigma$, the standard projected dynamical system [27] associated with $F(v)$ is given by:

$$\dot{v} = \lim_{\delta \to 0^+} \frac{P_\Omega(v + \delta F(v)) - v}{\delta}. \tag{9}$$

We note that, in general, (9) is a discontinuous dynamical system. On the contrary, the vector field in (8b)-(8c) is Lipschitz continuous. For static optimization problems, similar dynamics have been studied in e.g. [28], [29]. However, to the best of our knowledge, (8b)-(8c) is the first projected output feedback controller that is Lipschitz-continuous. □

Fig. 2 provides a representative example of the trajectories produced by the considered projected output feedback controllers, and compares them with those generated by a controller with a discontinuous projection of the form (9).

A. Stability and Tracking Analysis

In this section we characterize convergence properties of (8). We begin by characterizing the existence of solutions.

Lemma 3.5: Let Assumptions 2.2–2.5 hold. For each $\xi_0 = (x_0, u_0, \lambda_0) \in \mathbb{R}^{n+m+r}$, there exists a unique solution $\xi(t)$ of

(8) with $\xi(0)=\xi_0$. Moreover, ξ is continuously differentiable and maximal.

Proof: This claim follows by noting that: (i) the projection mapping is globally Lipschitz [28], [29], (ii) under Assumptions 2.2–2.5, the maps $L_{u,t}(u,y,\lambda)$ and $L_{\lambda,t}(y,\lambda)$ are globally Lipschitz in (u,y,λ) uniformly in t, and locally Lipschitz with respect to t, and (iii) under Assumption 2.4 the plant dynamics are locally Lipschitz in t. ■

Lemma 3.5 guarantees that the trajectories of (8) are continuously differentiable (see Fig. 2). Moreover, since trajectories are maximal, Lemma 3.5 guarantees that trajectories have no finite escape time. The latter property is harnessed to prove the following lemma, which establishes attractivity and forward invariance of the feasible set (similarly to [29, Thm 3.2]).

Lemma 3.6: Let Assumptions 2.2–2.5 hold. If $u(0)\notin\mathcal{U}$ (resp. $\lambda(0)\notin\mathcal{C}$), then the trajectory u (resp. λ) approaches exponentially the set $\mathcal{U}$ (resp. the set $\mathcal{C}$). If $u(t_0)\in\mathcal{U}$ (resp. $\lambda(t_0)\in\mathcal{C}$) for some $t_0\geq 0$, then $u(t)\in\mathcal{U}$ (resp. $\lambda(t)\in\mathcal{C}$) for all $t\geq t_0$.

The following lemma establishes a relationship between the saddle-point of the regularized Lagrangian and the equilibria of (8). The proof is omitted due to space limitations.

Lemma 3.7: Let Assumptions 2.1-2.5 hold. For any $w_t\in\mathbb{R}^q$ and $t\in\mathbb{R}_{\geq 0}$, let $\xi_{\text{eq}}:=(x_{\text{eq}},u_{\text{eq}},\lambda_{\text{eq}})$ denote an equilibrium of (8), let $(u^*_{\nu,t},\lambda^*_{\nu,t})$ denote the unique saddle-point of (5), and let $x^*_{\nu,t}:=-A^{-1}(Bu^*_{\nu,t}+Hw_t)$. Then, ξ_{eq} is unique. Moreover, $x_{\text{eq}}=x^*_{\nu,t}$, $u_{\text{eq}}=u^*_{\nu,t}$, and $\lambda_{\text{eq}}=\lambda^*_{\nu,t}$.

To study the stability properties of (8), we begin by showing that, when the dynamics of the plant (1) are infinitely fast (i.e., when (2) is satisfied at all times), the controller (8b)-(8c) converges exponentially fast to the unique saddle-point of the regularized Lagrangian, modulo a residual error proportional to the time-variation of the optimal trajectory z^*_t. We recall that, in what follows, we use the notation $\tilde{z}_\nu(t)=z(t)-z^*_t$, where $z:=(u,\lambda)$, and $z^*_{\nu,t}:=(u^*_{\nu,t},\lambda^*_{\nu,t})$.

Proposition 3.8: Let Assumptions 2.1-2.6 hold, $\mu:=\min\{\mu_u,\nu\}$, $\ell=:\sqrt{2}(\overline{K}+\max\{\ell_u+\|G\|^2\ell_y,\nu\})$. If $\varepsilon=0$ and the controller gain satisfies $\eta<\frac{4\mu}{\ell^2}$, then for any $t_0\in\mathbb{R}_{\geq 0}$:

$$\|\tilde{z}_\nu(t)\|\leq e^{-\frac{1}{2}\rho_z(t-t_0)}\|\tilde{z}_\nu(t_0)\|+\frac{2}{\rho_z}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot{z}^*_{\nu,\tau}\|, \quad (10)$$

for all $t\geq t_0$, where $\rho_z=\eta(\mu-\frac{\eta\ell^2}{4})$.

The proof is presented in Section III-B. We note that the rate of convergence ρ_z can be tuned by properly choosing the controller gain η. Next, we provide a sufficient condition on the time-scale separation between the plant dynamics (8a) and the feedback controller (8b)-(8c) to ensure tracking of the optimal trajectory. In what follows, we use the compact notation $\tilde{\xi}_\nu(t)=\xi(t)-\xi^*_{\nu,t}$, where $\xi:=(x,u,\lambda)$, and $\xi^*_{\nu,t}:=(-A^{-1}(Bu^*_{\nu,t}+Ew_t),u^*_{\nu,t},\lambda^*_{\nu,t})$.

Theorem 3.9: Let Assumptions 2.1-2.6 hold, let $\ell=:\sqrt{2}(\overline{K}+\max\{\ell_u+\|G\|^2\ell_y,\nu\})$ and $\mu:=\min\{\mu_u,\nu\}$. If

$$\eta<\frac{4\mu}{\ell^2}\quad\text{ and }\quad\varepsilon<\frac{\rho_z\underline{\lambda}(Q_x)}{4\eta\|PA^{-1}B\|\Psi}, \quad (11)$$

where $\rho_z=\eta(\mu-\frac{\eta\ell^2}{4})$, $\Psi=\rho_z\ell_y\|C\|\|G\|+\sqrt{2}\|C\|(\ell_y\|G\|+\bar{K})k_0$ and $k_0=\max\{2+\eta(\ell_u+\ell_y\|G\|^2),\|G\|\bar{K}\}$, then for any $t_0\in\mathbb{R}_{\geq 0}$:

$$\begin{aligned}\|\tilde{\xi}_\nu(t)\|\leq\sqrt{\kappa}\|\tilde{\xi}_\nu(t_0)\|e^{-\frac{1}{2}\rho_\xi(t-t_0)}+\frac{2}{\rho_z}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot{z}^*_{\nu,\tau}\|\\ +\frac{4\varepsilon\|PA^{-1}E\|}{\underline{\lambda}(Q_x)}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot{w}_\tau\|,\end{aligned} \quad (12)$$

for all $t\geq t_0$, where $\rho_\xi=\frac{1}{2}\min\left\{2\rho_z,\frac{1}{4\varepsilon}\frac{\underline{\lambda}(Q_x)}{\bar{\lambda}(P_x)}\right\}$, and $\kappa=\max\{\frac{1}{2},\bar{\lambda}(P_x)\}/\min\{\frac{1}{2},\underline{\lambda}(P_x)\}$.

The proof of Theorem 3.9 is provided in Section III-B. The result shows that, under a sufficient separation between the time scales of the plant and the controller, the trajectories of (8) globally exponentially converge to ξ^*_t (which we recall is the trajectory of the unique saddle-point of the regularized Lagrangian). Two important observations are in order. First, the upper bound for ε is an increasing function of $\underline{\lambda}(Q_x)$ and ρ_z, that are interpreted as the convergence rate of the open-loop plant and of the controller with $\varepsilon=0$, respectively. Moreover, the bound is a decreasing function of $\|P_xA^{-1}B\|$. Since $\|A^{-1}\|\to 0$ when the eigenvalues of A are approaching the open right complex plane, the latter term takes into account the margin of stability of the open-loop plant. Second, we note that the rate of convergence ρ_ξ is governed by the quantities ρ_z and ε (as well as matrices P_x and Q_x), which are interpreted as the rate of convergence of the controller with $\varepsilon=0$ and the rate of convergence of the open-loop plant.

Remark 3.10: The bound (12) depends on two main quantities: $\operatorname{ess\,sup}_{\tau\geq t_0}\|\dot{z}^*_{\nu,\tau}\|$, which captures the time-variability of $z^*_{\nu,t}$, and $\operatorname{ess\,sup}_{\tau\geq t_0}\|\dot{w}_\tau\|$, which captures the variation of the equilibrium of (1) induced by the presence of a time-varying exogenous input w_t. Notably, in case of a static problem, (12) boils down to the exponential stability result $\|\tilde{\xi}_\nu(t)\|\leq\sqrt{\kappa}\|\tilde{\xi}_\nu(t_0)\|e^{-\frac{1}{2}\rho_\xi(t-t_0)}$. □

B. Proofs of the Results

In this section, we present the proof of Proposition 3.8 and Theorem 3.9. For the subsequent analysis, it is convenient to define the following time-varying map:

$$F_t(z):=\begin{bmatrix}\nabla\phi_t(u)+G^\top\nabla\psi_t(Gu+Hw_t)+G^\top K_t^\top\lambda\\ -(K(Gu+Hw_t)-e-\nu\lambda)\end{bmatrix}. \quad (13)$$

1) Proof of Proposition 3.8: Recall that $z:=(u,\lambda)$. We note that, when $\varepsilon=0$, the dynamics (8) can be rewritten as:

$$\dot{z}=P_\Omega\big(z-\eta F_t(z)\big)-z, \quad (14)$$

where $\Omega:=\mathcal{U}\times\mathcal{C}$. Proposition 3.8 leverages this structure as well as four auxiliary lemmas. Below, Lemma 3.11 follows directly from [30, Lemma 6] and [13].

Lemma 3.11: Let Assumption 2.2 hold. Then, for any $t\geq 0$, $u,u'\in\mathbb{R}^m$ and $y,y'\in\mathbb{R}^p$, there exist symmetric matrices $T_{u,t}\in\mathbb{R}^{m\times m}$ and $T_{y,t}\in\mathbb{R}^{p\times p}$, which satisfy $\mu_uI\preceq T_{u,t}\preceq\ell_uI$ and $0\preceq T_{y,t}\preceq\ell_yI$, such that $\nabla J_t(u)-\nabla J_t(u')=T_{u,t}(u-u')$ and $\nabla I_t(y)-\nabla I_t(y')=T_{y,t}(y-y')$.

Although the time-varying matrices $T_{u,t}$ and $T_{y,t}$ are a functions of u,u' and y,y', respectively, this result allows us to leverage the relationships $\mu_u I \preceq T_{u,t} \preceq \ell_u I$ and $0 \preceq T_{y,t} \preceq \ell_y I$. Next, we show that $F_t(z)$ is strongly monotone and globally Lipschitz continuous, uniformly in t.

Lemma 3.12: Let Assumption 2.2 hold. Then, (13) satisfies:

$$(z-z')^\mathsf{T}(F_t(z)-F_t(z')) \geq \min\{\mu_u,\nu\}\|z-z'\|^2, \quad (15)$$

for all $z, z' \in \mathbb{R}^{m+r}$, and all $t \in \mathbb{R}_{\geq 0}$.

Proof: By expanding the left-hand side of (15), and by using Lemma 3.11:

$$\begin{aligned}(z-z')^\mathsf{T}(F_t(z)-F_t(z')) &= (u-u')^\mathsf{T}\left(\nabla\phi_t(u)-\nabla\phi_t(u')\right)\\ &\quad + (u-u')^\mathsf{T}G^\mathsf{T}(\nabla\psi_t(Gu+Hw_t)-\nabla\psi_t(Gu'+Hw_t))\\ &\quad + \nu\|\lambda-\lambda'\|^2\\ &= (u-u')^\mathsf{T}(T_{u,t}+G^\mathsf{T}T_{y,t}G)(u-u')+\nu\|\lambda-\lambda'\|^2\\ &\geq \mu_u\|u-u'\|^2+\nu\|\lambda-\lambda'\|^2 \geq \min\{\mu_u,\nu\}\|z-z'\|^2,\end{aligned}$$

which proves the claim. ■

Lemma 3.13: Let Assumptions 2.2 and 2.3 hold. Then, the mapping (13) satisfies:

$$\|F_t(z)-F_t(z')\| \leq \ell\|z-z'\|, \quad (16)$$

for all $z, z' \in \mathbb{R}^{m+r}$, and all $t \in \mathbb{R}_{\geq 0}$, where $\ell =: \sqrt{2}\max\{\ell_u+\ell_y\|G\|^2+\overline{K}, \nu+\overline{K}\})$.

Proof: Using (7), we directly obtain the bounds:

$$\begin{aligned}&\|L_{u,t}(u,Gu+Hw,\lambda)-L_{u,t}(u',Gu'+Hw,\lambda')\|\\ &\qquad\leq (\ell_u+\ell_y\|G\|^2)\|u-u'\|+\overline{K}\|\lambda-\lambda'\|,\\ &\|L_{\lambda,t}(Gu+Hw,\lambda)-L_{\lambda,t}(Gu'+Hw,\lambda')\|\\ &\qquad\leq \overline{K}\|u-u'\|+\nu\|\lambda-\lambda'\|.\end{aligned}$$

Finally, the claim follows by using the relationship: $\|u-u'\|+\|\lambda-\lambda'\| \leq \sqrt{2}\|z-z'\|$. ■

The following lemma, which is a particular case of [15, Ch. 4] (see also [26]) will be used to prove Proposition 3.8.

Lemma 3.14: Consider the system $\dot{x} = f(t,x,u)$, where $f: \mathbb{R}_{\geq 0}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz in t, x, and u, and $t \mapsto u(t)$ is measurable and essentially bounded. If there exists a continuously differentiable $V: \mathbb{R}_{\geq 0}\times\mathbb{R}^n \to \mathbb{R}$ s.t.:

$$\underline{a}\|x\|^2 \leq V(t,x) \leq \bar{a}\|x\|^2, \quad (17a)$$

$$\frac{d}{dt}V(t,x) \leq -bV(t,x), \quad \forall \|x\| \geq b_0 > 0, \quad (17b)$$

hold a.e., then, for all $t_0 \in \mathbb{R}_{\geq 0}$ and $x(0) \in \mathbb{R}^n$:

$$\|x(t)\| \leq \sqrt{\bar{a}/\underline{a}}(\|x(t_0)\|e^{-\frac{1}{2}b(t-t_0)}+b_0), \quad \forall t \geq 0. \quad (18)$$

Using the results above, we now present the proof of Proposition 3.8. In particular, we show that the function $V(\tilde{z}_\nu) = \frac{1}{2}\|z(t)-z^*_{\nu,t}\|^2$ satisfies the assumptions of Lemma 3.14, where we recall that $\tilde{z}_\nu := z - z^*_{\nu,t}$ and we let $\hat{z} := P_\Omega(z-\alpha F_t(z))$. Expand the derivative along the trajectory of (14) as:

$$\frac{d}{dt}V(\tilde{z}_\nu) = -\tilde{z}_\nu^\mathsf{T}(z-\hat{z}_\nu) - \tilde{z}_\nu^\mathsf{T}\dot{z}^*_{\nu,t}, \quad (19)$$

where we recall that $\dot{z}^*_{\nu,t}$ exists a.e.. Next, we recall that the projection operator is the unique vector $P_\Omega(z)$ that satisfies:

$$(v'-P_\Omega(v))^\mathsf{T}(P_\Omega(v)-v) \geq 0, \text{ for all } v' \in \Omega. \quad (20)$$

By using (20) with $v' = z^*_{\nu,t}$ and $v = \hat{z}$, we have $(\tilde{z}_\nu + \eta F_t(z))^\mathsf{T}(z-\hat{z}) \geq \|z-\hat{z}\|^2 - \eta(z-z^*_{\nu,t})^\mathsf{T}F_t(z)$, and thus the first term in (19) satisfies:

$$\begin{aligned}-\tilde{z}_\nu^\mathsf{T}(z-\hat{z}) &\leq -\|z-\hat{z}\|^2 - \eta(\hat{z}-z^*)^\mathsf{T}F_t(z)\\ &\leq -\|z-\hat{z}\|^2 - \eta(\hat{z}-z^*_{\nu,t})^\mathsf{T}(F_t(z)-F_t(z^*_{\nu,t}))\\ &\qquad - \eta(\hat{z}-z^*_{\nu,t})^\mathsf{T}F_t(z^*_{\nu,t})\\ &= -\|z-\hat{z}\|^2 - \eta\tilde{z}_\nu^\mathsf{T}(F_t(z)-F_t(z^*_{\nu,t}))\\ &\qquad + \eta(z-\hat{z})^\mathsf{T}(F_t(z)-F_t(z^*_{\nu,t}))\\ &\leq -\|z-\hat{z}\|^2 + \eta\ell\|z-\hat{z}\|\|\tilde{z}_\nu\| - \eta\mu\|\tilde{z}_\nu\|^2\\ &\leq -\eta\left(\mu-\eta\ell^2/4\right)\|\tilde{z}_\nu\|^2, \end{aligned} \quad (21)$$

where the second inequality follows by adding and subtracting $\eta(\hat{z}-z^*_{\nu,t})^\mathsf{T}F_t(z^*_{\nu,t})$, the third inequality follows by expanding $\hat{z}-z^*_{\nu,t} = (z-z^*_{\nu,t})-(z-\tilde{z}_\nu)$ and by using $(\hat{z}-z^*_{\nu,t})^\mathsf{T}F_t(z) \geq 0$, the fourth inequality follows from Lemmas 3.12 and 3.13, and the last inequality follows by using the relationship $2ab \leq a^2+b^2$ with $a = \|z-\hat{z}\|$ and $b = \frac{1}{2}\alpha\ell\|\tilde{z}_\nu\|$. By substituting into (19):

$$\begin{aligned}\frac{d}{dt}V(\tilde{z}_\nu) &\leq -\eta(\mu-\frac{\eta\ell^2}{4})\|\tilde{z}_\nu\|^2 + \|\tilde{z}_\nu\|\|\dot{z}^*_{\nu,t}\|\\ &\leq -\frac{\eta}{2}(\mu-\frac{\eta\ell^2}{4})\|\tilde{z}_\nu\|^2,\end{aligned}$$

where the last inequality holds when $\|\tilde{z}_\nu\| \geq \frac{2}{\eta(\mu-\eta\ell^2/4)}\operatorname{ess\,sup}_{\tau\geq t_0}\|\dot{z}^*_{\nu,\tau}\|$. Finally, the claim follows by application of Lemma 3.14 with $\bar{a} = \underline{a} = \frac{1}{2}$, $b = (\mu-\eta\ell^2/4)$, and $b_0 = \frac{2}{\eta(\mu-\eta\ell^2/4)}\operatorname{ess\,sup}_{\tau\geq t_0}\|\dot{z}^*_{\nu,\tau}\|$. ■

2) Proof of Theorem 3.9: We begin by performing a change of variables for (8). Let $z := (u,\lambda)$, $\tilde{x} := x + A^{-1}Bu + A^{-1}Ew_t$, and

$$F_t(z,\tilde{x}) := \begin{bmatrix}L_{u,t}(u, C\tilde{x}+Gu+Hw_t,\lambda)\\ L_{\lambda,t}(C\tilde{x}+Gu+Hw_t,\lambda)\end{bmatrix}.$$

Then, the dynamics (8) can be rewritten as:

$$\begin{aligned}\varepsilon\dot{\tilde{x}} &= A\tilde{x} + \varepsilon A^{-1}BS\dot{z} + A^{-1}E\dot{w}_t\\ \dot{z} &= P_\Omega(z-\alpha F_t(z,\tilde{x})) - z,\end{aligned} \quad (22)$$

where $S = [I_m, 0]$, and $\Omega = \mathcal{U}\times\mathcal{C}$. Moreover, we let $b := \eta\|C\|(\ell_y\|G\| + \bar{K}))$, and $g := 2\sqrt{2}\|PA^{-1}B\|k_0$. To prove this claim we will show that

$$U(\tilde{z}_\nu,\tilde{x}) := (1-\theta)V(\tilde{z}_\nu) + \theta W(z), \quad (23)$$

where $V(\tilde{z}_\nu) = \frac{1}{2}\|z(t)-z^*_{\nu,t}\|^2$, $W(z) = \tilde{x}^\mathsf{T}P_x\tilde{x}$, and $\theta = b/(b+g)$ satisfies the assumptions of Lemma 3.14. We recall that $\tilde{z}_\nu := z - z^*_{\nu,t}$ and $\hat{z} := P_\Omega(z-\alpha F_t(z))$. The time-derivative of $V(t,z)$ along the trajectory of (22) reads:

$$\frac{d}{dt}V(\tilde{z}_\nu) = \tilde{z}_\nu^\mathsf{T}(\hat{z}-z) - \tilde{z}_\nu^\mathsf{T}\dot{z}^*_{\nu,t} \quad (24)$$

almost everywhere. The first term satisfies:

$$\begin{aligned}z^\mathsf{T}(\hat{z}-z) &= \tilde{z}_\nu^\mathsf{T}(P_\Omega(z-\eta F_t(z,0))-z)\\ &\qquad + \tilde{z}_\nu^\mathsf{T}(P_\Omega(z-\eta F_t(z,\tilde{x}))-P_\Omega(z-\eta F_t(z,0)))\\ &\leq \tilde{z}_\nu^\mathsf{T}(P_\Omega(z-\eta F_t(z,0))-z)\\ &\qquad + \eta\|\tilde{z}_\nu\|\|F_t(z,\tilde{x})-F_t(z,0)\|\\ &\leq -\eta(\mu-\frac{\eta\ell^2}{4})\|\tilde{z}_\nu\|^2 + \eta\|\tilde{z}_\nu\|\|F_t(z,\tilde{x})-F_t(z,0))\|,\end{aligned}$$

where the first inequality follows from the non-expansiveness of the projection operator, and the second inequality follows from (21). Moreover, by expanding the terms:

$$\begin{aligned}&\|F_t(z,\tilde{x}))-F_t(z,0))\|\\&\quad\le\left\|\begin{bmatrix}G^\mathsf{T}(\nabla f_y(C\tilde{x}+Gu+Hw_t)-\nabla f_y(Gu+Hw_t)\\-K_tC\tilde{x}\end{bmatrix}\right\|\\&\quad\le\|C\|(\ell_y\|G\|+\bar{K})\|\tilde{x}\|.\end{aligned}$$

Hence, by recalling the definition of b and ρ_z, (24) satisfies:

$$\begin{aligned}\frac{d}{dt}V(\tilde{z}_\nu)&\le-\rho_z\|\tilde{z}_\nu\|^2+b\|\tilde{x}\|\|\tilde{z}_\nu\|+\|\tilde{z}_\nu\|\|\dot{z}^*_{\nu,t}\|\\&\le-\frac{\rho_z}{2}\|\tilde{z}_\nu\|^2+b\|\tilde{x}\|\|\tilde{z}_\nu\|,\end{aligned}\tag{25}$$

where the last inequality holds when $\|\tilde{z}_\nu\|\ge\frac{2}{\rho_z}\operatorname{ess\,sup}\|\dot{z}^*_{\nu,t}\|$. The time-derivative of $W(\tilde{x})$ along the trajectories of (22):

$$\begin{aligned}\frac{d}{dt}W(z)&=\varepsilon^{-1}\tilde{x}^\mathsf{T}(A^\mathsf{T}P_x+P_xA)\tilde{x}\\&\quad+2\tilde{x}^\mathsf{T}P_xA^{-1}BS\dot{z}+2\tilde{x}^\mathsf{T}P_xA^{-1}E\dot{w}_t\\&\le-\varepsilon^{-1}\underline{\lambda}(Q_x)\|\tilde{x}\|^2+2\|P_xA^{-1}B\|\|\tilde{x}\|\|S\dot{z}\|\\&\quad+2\|P_xA^{-1}B\|\|\tilde{x}\|\|\dot{w}_t\|.\end{aligned}\tag{26}$$

By expanding the terms:

$$\begin{aligned}&\|S\dot{z}\|=\|S(P_\Omega(z-\eta F_t(z,\tilde{x}))-z)\|\\&=\|S(P_\Omega(z-\eta F_t(x,\tilde{x}))-z-P_\Omega(z-\eta F_t(z^*,0))+z^*)\|\\&\le\eta\|L_{u,t}(u,C\tilde{x}+Gu+Hw_t,\lambda)\\&\qquad-L_{u,t}(u^*,Gu^*+Hw_t,\lambda^*)\|+2\|u-u^*\|\\&\le\sqrt{2}\max\{2+\eta(\ell_u+\ell_y\|G\|^2),\|K_tG\|\}\|\tilde{z}_\nu\|\\&\qquad+\eta\ell_y\|C\|\|G\|\|\tilde{x}\|,\end{aligned}$$

where the first inequality follows from the non-expansiveness of the projection operator and the second inequality follows from Assumption 2.2. By recalling the definition of g, letting $d=2\eta\ell_y\|PA^{-1}B\|\|C\|\|G\|$, by substituting into (26):

$$\begin{aligned}\frac{d}{dt}W(\tilde{x})&\le-\varepsilon^{-1}\underline{\lambda}(Q_x)\|\tilde{x}\|^2+d\|\tilde{x}\|^2\\&\quad+g\|\tilde{x}\|\|\tilde{z}_\nu\|+2\|P_xA^{-1}B\|\|\tilde{x}\|\|\dot{w}_t\|\\&\le-\frac{\underline{\lambda}(Q_x)}{2\varepsilon}\|\tilde{x}\|^2+d\|\tilde{x}\|^2+g\|\tilde{x}\|\|\tilde{z}_\nu\|,\end{aligned}\tag{27}$$

where the last inequality holds if $\|\tilde{x}\|\ge\frac{4\varepsilon\|P_xA^{-1}E\|}{\underline{\lambda}(Q_x)}\operatorname{ess\,sup}\|\dot{w}_t\|$. By combining (25)-(27):

$$\frac{d}{dt}U(\tilde{x},\tilde{z}_\nu)\le-\hat{\xi}^\mathsf{T}\Lambda\hat{\xi}-\frac{1}{2}\min\{2\rho_z,\frac{\underline{\lambda}(Q_x)}{2\varepsilon\bar{\lambda}(P_x)}\},$$

where

$$\Lambda:=\begin{bmatrix}(1-\theta)\frac{\rho_z}{4}&-\frac{1}{2}((1-\theta)b+\theta g)\\\frac{1}{2}((1-\theta)b+\theta g)&\theta(\frac{\underline{\lambda}(Q_x)}{4\varepsilon}-d)\end{bmatrix}.$$

Matrix Λ is positive definite when

$$\theta(1-\theta)\frac{\rho_z\underline{\lambda}(Q_x)}{16\varepsilon}>\frac{1}{4}((1-\theta)b+\theta g)^2,$$

which holds when (11) is satisfied. Finally, the claim follows by application of Lemma 3.14 with $\bar{a}=\max\{\frac{1}{2},\bar{\lambda}(P_x)\}$, $\underline{a}=\min\{\frac{1}{2},\underline{\lambda}(P_x)\}$, $c_3=\frac{1}{2}\min\{2\rho_z,\frac{\underline{\lambda}(Q_x)}{4\varepsilon\bar{\lambda}(P_x)}\}$, and $b_0=\max\{\frac{2}{\rho_z}\operatorname{ess\,sup}\|\dot{z}^*_{\nu,t}\|,\frac{4\varepsilon\|PA^{-1}E\|}{\underline{\lambda}(Q_x)}\operatorname{ess\,sup}\|\dot{w}_t\|\}$. ■

C. Extensions

Our analysis suggests that results can be extended in different directions. Here, we discuss two possible extensions.

1) Switched LTI Plants with Common Quadratic Lyapunov Functions: Theorem 3.9 can be extended to consider switched LTI plants of the form

$$\begin{aligned}\dot{x}&=A_\sigma x+B_\sigma u+E_\sigma w_t,\\y&=Cx+Dw_t,\end{aligned}\tag{28}$$

where $\sigma:\mathbb{R}_{\ge0}\to\mathcal{Q}$ is a switching signal taking values in a finite an compact set. If system (28) has a common equilibrium point $x^*_{\text{eq}}=A_\sigma^{-1}Bu+A_\sigma^{-1}E_\sigma w_t$ for all values of σ, and a common quadratic Lyapunov function V, the same construction of the Lyapunov function (23) can be used to establish exponential ISS for the closed-loop system. Since in this case, G and H in (2) are also common to all the subsystems, the bounds of Theorem 3.9 still hold. This observation is relevant for applications where internal feedback controllers have been implemented *a priori* to stabilize the modes of the plant, such that Assumption 2.1 holds. In this context, different controller may lead to closed-loop transient performance, while preserving a common equilibrium point [31]. Note, however, that having a stable autonomous switched LTI system does not necessarily imply the existence of a common quadratic Lyapunov function. However, it implies the existence of a common Lyapunov function that is homogeneous of degree 2, e.g., piece-wise quadratic [32]. When the matrices C_σ and D_σ are also switching, the exponential ISS result of Theorem 3.9 is preserved provided the pair (G,H) remains common, and with suitable modifications of (11) and (12) using the max and the min of the signals over the compact set $\mathcal{Q}$.

2) Switched Plants with Average Dwell-Time Constraints: When the switched system (12) does not have a common Lyapunov function, it is still possible to obtain a result of the form (12), provided the switching is slow "on the average". In particular, if the switching signal σ satisfies an average dwell-time constrain of the form

$$N_\sigma(t,\tau)\le\eta(t-\tau)+N_0,\tag{29}$$

where $N_\sigma(t,\tau)$ denotes the number of discontinuities of σ in the open interval (τ,t), and $N_0\ge0$ is a chatter bound that allows for a finite number of consecutive switches. In this case, it is possible to choose $\eta>0$ sufficiently small such that the exponential stability property of the switched system is preserved, and the same construction (23) carries over. This observation follows directly from the Lyapunov construction presented in [10], which permits the derivation of a result similar to Proposition 3.8 using quadratic Lyapunov functions. Characterizations of the conditions that emerge between η and the time-scale separation parameters (ε,η) can also be explicitly derived as in [10]. However, unlike the results of [10], the results of this paper allow to consider online optimization problems *with constraints*. To the best knowledge of the authors, similar results for online optimization with constraints of switched systems have not been studied before in the literature.

IV. Online Primal-Dual Gradient Flow

In this section we consider the problem of regulating (1) to the solution of the following optimization problem:

$$(u_t^*, y_t^*) := \underset{\bar u\in\mathbb{R}^m,\ \bar y\in\mathbb{R}^p}{\arg\min}\ \phi_t(\bar u) + \psi_t(\bar y), \tag{30a}$$

$$\text{s.t.} \qquad \bar y = G\bar u + Hw_t, \quad K_t\bar y = e_t, \tag{30b}$$

which contains only equality constraints on the system outputs. As we show next, a controller as described in Problem 1 can be developed in this case by leveraging a non-regularized Lagrangian function. To this end, we impose, similarly to [11], the following assumption.

Assumption 4.1: The columns of K_tG are linearly independent and there exists $\underline{k}, \bar k \in \mathbb{R}_{>0}$ such that $\underline{k}I \preceq K_tGG^\mathsf{T}K_t^\mathsf{T} \preceq \bar kI$ for all t. □

Since problem (30) contains only equality constraints, Assumption 4.1 is sufficient to guarantee uniqueness of the optimal multipliers [11]. In what follows, for notation simplicity we will state the results by considering a time-invariant constraint matrix K. The stated results directly extend to the case of time-varying matrices, as noted in pertinent remarks.

The Lagrangian of problem (30) is given by $\mathcal{L}_t(u,\lambda) = \phi_t(u) + \psi_t(Gu+Hw_t) + \lambda^\mathsf{T}(K(Gu+Hw_t) - e_t)$, where $\lambda \in \mathbb{R}^r$ is the vector of dual variables. Under Assumptions 2.2 and 4.1, the unique minimizer (u_t^*, y_t^*) of (30) solves the following KKT conditions:

$$\begin{aligned} 0 &= \nabla\phi_t(u_t^*) + G^\mathsf{T}\nabla\psi_t(Gu_t^* + Hw_t) + G^\mathsf{T}K^\mathsf{T}\lambda_t^*,\\ 0 &= K(Gu_t^* + Hw_t) - e_t. \end{aligned} \tag{31}$$

To solve Problem 1, we consider the following modifications of the gradients of the Lagrangian:

$$L_{u,t}(u,y,\lambda) := \nabla\psi(u) + G^\mathsf{T}\nabla\phi(y) + G^\mathsf{T}K^\mathsf{T}\lambda, \tag{32a}$$

$$L_{\lambda,t}(y) := Ky - e, \tag{32b}$$

where (similarly to (7)) with respect to the gradients of $\mathcal{L}_t(u,\lambda)$, the steady-state map $Gu + Hw_t$ has been replaced by the variable y. The following online primal-dual gradient controller in feedback with the plant (1) is then considered:

$$\varepsilon\dot x = Ax + Bu + Ew_t, \qquad y = Cx + Dw_t, \tag{33a}$$

$$\dot u = -\eta_u L_{u,t}(u,y,\lambda), \tag{33b}$$

$$\dot\lambda = \eta_\lambda L_{\lambda,t}(y), \tag{33c}$$

where $\varepsilon, \eta_u, \eta_\lambda \in \mathbb{R}_{>0}$ are plant and controller gains. Similarly to the projected controller in Section III, the controller (33b)-(33c) uses output-feedback from the plant, and does not require any knowledge on w_t. The following lemma relates the time-varying equilibria of (33) to the solution of (30).

Lemma 4.2: Let Assumptions 2.1-2.4 and 4.1 be satisfied. For any $w_t \in \mathbb{R}^q$ and $t \in \mathbb{R}_{\geq 0}$, let $\xi_{\text{eq}} := (x_{\text{eq}}, u_{\text{eq}}, \lambda_{\text{eq}})$ denote an equilibrium of (33), let (u_t^*, λ_t^*) be the unique saddle-point of (30), and let $x_t^* := -A^{-1}(Bu_t^* + Hw_t)$. Then, ξ_{eq} is unique. Moreover, $x_{\text{eq}} = x_t^*$, $u_{\text{eq}} = u_t^*$, and $\lambda_{\text{eq}} = \lambda_t^*$.

The proof of this claim is omitted due to space limitations. In contrast with Lemma 3.7 in Section III, Lemma 4.2 establishes that the equilibrium point of (33) coincides with the saddle points of the exact (non-augmented) Lagrangian function. With this in place, the next section investigates the stability properties of the interconnected system (33).

A. Stability and Tracking Analysis

We begin by showing that, when the dynamics of the plant are infinitely fast, (33) converges exponentially to the solutions of (3). We recall that we use the compact notation $\tilde z(t) = z(t) - z_t^*$, where $z := (u,\lambda)$ and $z_t^* := (u_t^*, \lambda_t^*)$.

Proposition 4.3: Let Assumptions 2.1-2.4 and 4.1 hold, let

$$P_z := \begin{bmatrix} \ell I & G^\mathsf{T}K^\mathsf{T} \\ KG & \ell\frac{\eta_u}{\eta_\lambda}I \end{bmatrix}, \tag{34}$$

where $\ell := \ell_u + \|G\|^2\ell_y$. If $\varepsilon = 0$ and the controller parameters are such that $\eta_u > \frac{4\bar k}{\ell\mu}\eta_\lambda$, then for any $t_0 \in \mathbb{R}_{\geq 0}$:

$$\|\tilde z(t)\| \leq \sqrt{\kappa}\|\tilde z(t_0)\|e^{-\frac{1}{2}\rho_z(t-t_0)} + \frac{4\|P_z\|\sqrt{\kappa}}{\underline{\lambda}(P_z)}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot z_\tau^*\|, \tag{35}$$

for all $t \geq t_0$, $\rho_z := \frac{1}{2}\min\{\eta_\lambda\frac{\underline{k}}{\ell}, \eta_u\frac{\mu}{2}\}$, $\kappa = \bar\lambda(P_z)/\underline{\lambda}(P_z)$.

The proof of this result is presented in Section IV-B. Two comments are in order. First, differently from [11, Theorem 1], Proposition 4.3 shows that ρ_z can be made arbitrarily large by properly tuning the parameters η_u and η_λ. Second, we note that the tracking result (35) is in the spirit of [25, Section 6]; however, in [25] the primal-dual dynamics are assumed to be differentiable with respect to t (in contrast, we require the milder condition of absolute continuity).

Remark 4.4: When the matrix K_t is time-varying, it follows that $P_{z,t}$ in (34) and the coefficient κ_t in (35) are time-varying too. In this case, the result (35) extends by replacing κ with $\sup_\tau \kappa_\tau$ and the coefficient $\frac{4\|P_z\|\sqrt{\kappa}}{\underline{\lambda}(P_z)}$ with $\sup_\tau \frac{4\|P_{z,\tau}\|\sqrt{\kappa_\tau}}{\underline{\lambda}(P_{z,\tau})}$. □

We now present sufficient conditions on the time-scale separation between the plant and controller dynamics that result in exponential stability properties of the system (33).

Theorem 4.5: (Stability and Tracking of (33)*)* Let Assumptions 2.1-2.4 and 4.1 hold. Suppose that ε satisfies

$$\varepsilon < \frac{\rho_z\underline{\lambda}(P_x)\underline{\lambda}(P_z)}{16\sigma_1\sigma_2 + 4\rho_z\underline{\lambda}(P_z)\sigma_3}, \tag{36}$$

and P_z, ρ_z are as in Proposition 4.3, and

$$\begin{aligned} \sigma_1 &:= 2\eta_u\ell_y\|C\|\|G\|(\ell + \|KG\|) + 2\eta_\lambda\|G^\mathsf{T}K^\mathsf{T}KC\| + 2\ell\eta_u\|KC\|,\\ \sigma_2 &:= 2\eta_u\ell\|P_xA^{-1}B\| + 2\eta_u\|P_xA^{-1}GG^\mathsf{T}K^\mathsf{T}\|,\\ \sigma_3 &:= 2\eta_u\ell_y\|C\|\|P_xA^{-1}BG^\mathsf{T}\|. \end{aligned}$$

Then, for any $t_0 \in \mathbb{R}_{\geq 0}$, the dynamics (33) satisfy:

$$\begin{aligned} \|\tilde\xi(t)\| \leq{}& \sqrt{\kappa}\|\tilde\xi(t_0)\|e^{-\frac{1}{2}\rho_\xi(t-t_0)} + \frac{4\|P_z\|\sqrt{\kappa}}{\rho_z\underline{\lambda}(P_z)}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot z_\tau^*\|\\ &+ \frac{4\|P_xA^{-1}E\|\sqrt{\kappa}}{\underline{\lambda}(Q_x)}\operatorname*{ess\,sup}_{\tau\geq t_0}\|\dot w_\tau\|, \end{aligned} \tag{37}$$

for all $t \geq t_0$, $\kappa = \max\{\bar\lambda(P_x), \bar\lambda(P_z)\}/\min\{\underline{\lambda}(P_x), \underline{\lambda}(P_z)\}$,

$$\rho_\xi = \frac{1}{4}\min\left\{\rho_z\frac{\underline{\lambda}(P_z)}{\bar\lambda(P_z)}, \varepsilon^{-1}\frac{\underline{\lambda}(Q_x)}{\bar\lambda(P_x)}\right\}. \tag{38}$$

The proof of this result is presented in Section IV-B. The bound on ε is an increasing function of $\underline{\lambda}(P_x)$ and $\rho_z\underline{\lambda}(P_z)$, which are the convergence rates of the open-loop plant and of the controller with $\varepsilon = 0$, respectively. Moreover, we note that the rate of convergence ρ_ξ is governed by the quantities ρ_z and ε (as well as matrices P_x, Q_x, and P_z), which are interpreted as the rates of convergence of the controller with $\varepsilon = 0$ and the rate of convergence of the open-loop plant. Finally, we note that the bound (37) can be readily extended to account for time-varying matrices K_t by adopting a reasoning similar to that in Remark 4.4.

B. Proofs of the Results

Here, we present the proofs of Proposition 4.3 and Theorem 4.5. We start by introducing the following change of variables for (33):

$$\tilde{x} := x - h(u, w_t), \quad h(u, w_t) := -A^{-1}Bu - A^{-1}Ew_t.$$

The dynamics (33) are re-written in the new variables next.

Lemma 4.6: Let Assumption 2.1-2.6 be satisfied, and for any $t \in \mathbb{R}_{\geq 0}$, let (u_t^*, λ_t^*) be the saddle-point of (30). The dynamics (33) have the following equivalent representation:

$$\begin{aligned}
\varepsilon\dot{\tilde{x}} &= F_{11}\tilde{x} + F_{12}(u - u_t^*) + F_{13}(\lambda - \lambda_t^*) + F_{14}\dot{w}_t,\\
\dot{u} &= F_{21}\tilde{x} + F_{22}(u - u_t^*) + F_{23}(\lambda - \lambda_t^*),\\
\dot{\lambda} &= F_{31}\tilde{x} + F_{32}(u - u_t^*),
\end{aligned} \tag{39}$$

where $F_{14} = \varepsilon A^{-1}E$,

$$\begin{aligned}
F_{11} &= A - \varepsilon\eta_u A^{-1}BG^\mathsf{T}T_{y,t}C, & F_{21} &= -\eta_u G^\mathsf{T}T_{y,t}C,\\
F_{12} &= -\varepsilon\eta_u A^{-1}B(T_{u,t} + G^\mathsf{T}T_{y,t}G), & F_{23} &= -\eta_u G^\mathsf{T}K^\mathsf{T},\\
F_{13} &= -\varepsilon\eta_u A^{-1}BG^\mathsf{T}K^\mathsf{T}, & F_{31} &= \eta_\lambda KC,\\
F_{22} &= -\eta_u(T_{u,t} + G^\mathsf{T}T_{y,t}G), & F_{32} &= \eta_\lambda KG,
\end{aligned}$$

and $T_{u,t}$, $T_{y,t}$ are symmetric matrices that satisfy $\mu_u I \preceq T_{u,t} \preceq \ell_u I$, $0 \preceq T_{y,t} \preceq \ell_y I$ uniformly in t.

Proof: By application of Lemma 3.11:

$$\begin{aligned}
\dot{u} &= -\eta_u L_{u,t}(u, y, \lambda) + \underbrace{\eta_u L_{u,t}(u_t^*, Gu_t^* + Hw_t, \lambda_t^*)}_{=0}\\
&= -\eta_u((T_{u,t} + G^\mathsf{T}T_{y,t}G)(u - u_t^*)\\
&\qquad + G^\mathsf{T}T_{y,t}C\tilde{x} + G^\mathsf{T}K^\mathsf{T}(\lambda - \lambda_t^*)),\\
\dot{\lambda} &= \eta_\lambda \mathcal{L}_{\lambda,t}(u, y, \lambda) - \underbrace{\nabla_\lambda \mathcal{L}_{\lambda,t}(u^*, Gu_t^* + Hw_t, \lambda_t^*)}_{=0}\\
&= \eta_\lambda(KC\tilde{x} + KG(u - u_t^*)).
\end{aligned}$$

Finally, by using the relationships $\varepsilon\dot{\tilde{x}} = \dot{x} - \varepsilon\frac{\partial h}{\partial u}\dot{u} - \varepsilon\frac{\partial h}{\partial w}\dot{w}_t$, and by substituting the expression for $\dot{u}$:

$$\begin{aligned}
\varepsilon\dot{\tilde{x}} &= A\tilde{x} + \varepsilon A^{-1}B\dot{u} + \varepsilon A^{-1}E\dot{w}_t\\
&= (A - \varepsilon\eta_u A^{-1}BG^\mathsf{T}T_{y,t}C)\tilde{x} - \varepsilon\eta_u A^{-1}BG^\mathsf{T}K^\mathsf{T}(\lambda - \lambda_t^*)\\
&\quad - \varepsilon\eta_u A^{-1}B(T_{u,t} + G^\mathsf{T}T_{y,t}G)(u - u_t^*) + \varepsilon A^{-1}E\dot{w}_t,
\end{aligned}$$

which proves the claim. ■

1) Proof of Proposition 4.3: The proof follows similar ideas as in [11, Lemma 2]. By letting $\varepsilon = 0$ in (39) we obtain $A\tilde{x} = 0$, which, by Assumption 2.1 implies $\tilde{x} = 0$. Hence, we let $z := (u, \lambda)$ and $\tilde{z} := z - z_t^*$, and we rewrite the dynamics (39) as $\dot{z} = F_z(z - z_t^*) = F_z\tilde{z}$, where

$$F_z = \begin{bmatrix} F_{22} & F_{23} \\ F_{31} & 0 \end{bmatrix}. \tag{40}$$

We will prove that $V(z) = \tilde{z}^\mathsf{T}P_z\tilde{z}$ satisfies the assumptions of Lemma 3.14. By the Schur Complement, P_z is positive definite if and only if $\ell^2\frac{\eta_u}{\eta_\lambda}I - G^\mathsf{T}K^\mathsf{T}KG \succ 0$. Using $\eta_u > \frac{4\bar{k}}{\ell\mu}\eta_\lambda$, $\ell \geq \mu$ and Assumption 4.1 one gets $\ell^2(\eta_u/\eta_\lambda)I - G^\mathsf{T}K^\mathsf{T}KG \succeq ((4\ell\bar{k})/\mu_u)I - \bar{k}I \succeq 3\bar{k} \succ 0$, which shows that P_z is positive definite. By expanding the time-derivative:

$$\begin{aligned}
\frac{d}{dt}V(\tilde{z}) &= (\dot{z} - \dot{z}_t^*)^\mathsf{T}P_z(z - z_t^*) + (z - z_t^*)^\mathsf{T}P_z(\dot{z} - \dot{z}_t^*)\\
&= \tilde{z}^\mathsf{T}(F_z^\mathsf{T}P_z + P_zF_z)\tilde{z} - 2\tilde{z}^\mathsf{T}P_z\dot{z}_t^*.
\end{aligned} \tag{41}$$

Next, we show that $\tilde{z}^\mathsf{T}(F_z^\mathsf{T}P_z + P_zF_z)\tilde{z} + \bar{\rho}_z V(\tilde{z}) \leq 0$, where $\bar{\rho}_z = \min\{\eta_\lambda\frac{k}{\ell}, \eta_u\frac{\mu}{2}\}$. Let $M := F_z^\mathsf{T}P_z + P_zF_z + \bar{\rho}_z P_z$. By expanding the product, $M = [M_{ij}]$ is a 2×2 block symmetric matrix with blocks:

$$\begin{aligned}
M_{11} &= 2\eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - 2\eta_\lambda G^\mathsf{T}K^\mathsf{T}KG - \bar{\rho}_z\ell I,\\
M_{12} &= \eta_u(T_{u,t} + G^\mathsf{T}T_{y,t}G)^\mathsf{T}G^\mathsf{T}K^\mathsf{T} - \bar{\rho}_z G^\mathsf{T}K^\mathsf{T},\\
M_{22} &= 2\eta_u KGG^\mathsf{T}K^\mathsf{T} - \bar{\rho}_z\ell(\eta_u/\eta_\lambda)I,
\end{aligned} \tag{42}$$

and $M_{21} = M_{12}^\mathsf{T}$. By application of the Schur Complement, M is positive definite when $M_{22} \succ 0$ and $M_{11} - M_{12}M_{22}^{-1}M_{12}^\mathsf{T} \succ 0$. The first condition can be rewritten as:

$$M_{22} \succeq (2\eta_u\underline{k} - \bar{\rho}_z\ell\frac{\eta_u}{\eta_\lambda})I \succeq \eta_u\underline{k}I \succ 0$$

where we used Assumption 4.1 and the expression of ρ_z. For the second condition, we have:

$$\begin{aligned}
M_{12}M_{22}^{-1}M_{12}^\mathsf{T} &\preceq M_{12}(\eta_u KGG^\mathsf{T}K^\mathsf{T})^{-1}M_{12}^\mathsf{T}\\
&= \eta_u(T_{u,t} + G^\mathsf{T}T_{y,t}G)^\mathsf{T}(T_{u,t} + G^\mathsf{T}T_{y,t}G) + \frac{\bar{\rho}_z^2}{\eta_u}I\\
&\quad - \bar{\rho}_z((T_{u,t} + G^\mathsf{T}T_{y,t}G)^\mathsf{T} + (T_{u,t} + G^\mathsf{T}T_{y,t}G))\\
&\preceq \eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) + \frac{\bar{\rho}_z^2}{\eta_u}I - 2\bar{\rho}_z(T_{u,t} + G^\mathsf{T}T_{y,t}G),
\end{aligned}$$

where the first bound follows from Assumption 4.1 and the definition of $\bar{\rho}_z$, the second identity follows from $G^\mathsf{T}K^\mathsf{T}(KGG^\mathsf{T}K^\mathsf{T})^{-1}KG = I$, and the last bound follows from $G^\mathsf{T}T_{y,t}G \succeq 0$. Thus:

$$\begin{aligned}
M_{11} - M_{12}M_{22}^{-1}M_{12}^\mathsf{T} &\succeq 2\eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - 2\eta_\lambda G^\mathsf{T}K^\mathsf{T}KG\\
&\quad - \bar{\rho}_z\ell I - \eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - \frac{\bar{\rho}_z^2}{\eta_u}I + 2\bar{\rho}_z(T_{u,t} + G^\mathsf{T}T_{y,t}G),
\end{aligned}$$

and, by using

$$\begin{aligned}
&\frac{1}{2}\eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - 2\eta_\lambda G^\mathsf{T}K^\mathsf{T}KG\\
&\qquad \succeq (\frac{1}{2}\eta_u\ell\mu_u - 2\eta_\lambda\bar{k})I \succ 0\\
&\frac{1}{2}\eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - \rho\ell I \succeq (\frac{1}{2}\eta_u\ell\mu - \rho\ell)I \succeq 0\\
&\eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) - \eta_u\ell(T_{u,t} + G^\mathsf{T}T_{y,t}G) = 0,
\end{aligned}$$

we conclude $M_{11} - M_{12}M_{22}^{-1}M_{12}^\mathsf{T} \succ 0$, which shows $M \succ 0$.

As a result, (41) satisfies:

$$\begin{aligned}\frac{d}{dt}V(\tilde z) &\le -\bar\rho_z V(\tilde z) + 2\|\tilde z\|\|P_z\|\|\dot z_t^*\| \\ &= -\frac{\bar\rho_z}{2}V(\tilde z) - \frac{\rho_z}{2}\underline\lambda(P_z)\|\tilde z\|^2 + 2\|\tilde z\|\|P_z\|\|\dot z_t^*\| \\ &\le -\frac{\bar\rho_z}{2}V(\tilde z), \qquad (43)\end{aligned}$$

where the last inequality holds when $2\|\tilde z\|\|P_z\|\|\dot z_t^*\| - \frac{\rho_z}{2}\underline\lambda(P_z)\|\tilde z\|^2 \le 0$, or $\|\tilde z\| \ge \frac{4\|P_z\|}{\rho_z\underline\lambda(P_z)}\operatorname{ess\,sup}_\tau \|\dot z_\tau^*\|$. Finally, the claim follows by application of Lemma 3.14 with $\bar a = \bar\lambda(P_z), \underline a = \underline\lambda(P_z)$, $b = \frac{\bar\rho_z}{2}$, and $b_0 = \frac{4\|P_z\|}{\rho_z\underline\lambda(P_z)}\operatorname{ess\,sup}_\tau \|\dot z_\tau^*\|$. ∎

2) Proof of Theorem 4.5: Let $z := (u, \lambda)$, $\tilde z := z - z_t^*$ and rewrite the dynamics (39) as:

$$\dot{\tilde x} = F_{11}\tilde x + F_{xz}\tilde z + F_{14}\dot w_t, \quad \dot{\tilde z} = F_{zx}\tilde x + F_z\tilde z, \qquad (44)$$

where F_z is as defined by (40), $F_{xz} = [F_{12}, F_{13}]$, and $F_{zx} = [F_{21}^\mathsf{T}, F_{31}^\mathsf{T}]^\mathsf{T}$. To show this claim, we will prove that the function $U(\tilde x, \tilde z) = (1-\theta)V(\tilde z) + \theta W(\tilde x)$, where $\theta = \|\Sigma_1\|/(\|\Sigma_2\| + \|\Sigma_1\|)$ satisfies the assumptions of Lemma 3.14. By substituting (44) and by using $F_z^\mathsf{T}P_z + P_zF_z \preceq -\bar\rho_z P_z$ (see (41) and (43)):

$$\begin{aligned}\dot V(\tilde z) &= \tilde z^\mathsf{T}(F_z^\mathsf{T}P_z + P_zF_z)\tilde z + 2\tilde x^\mathsf{T}F_{zx}P_z\tilde z - 2\tilde z^\mathsf{T}P_z\dot z^* \\ &\le -\rho_z\tilde z^\mathsf{T}P_z\tilde z + \tilde z^\mathsf{T}\Sigma_1\tilde x - 2\tilde z^\mathsf{T}P_z\dot z^* \\ &\le -\frac{\rho_z}{2}\underline\lambda(P_z)\|\tilde z\|^2 + \|\Sigma_1\|\|\tilde z\|\|\tilde x\|,\end{aligned}$$

where the last inequality holds if $\|\tilde z\| \ge \frac{4\|P_z\|}{\rho_z\underline\lambda(P_z)}\sup_\tau \|\dot z_\tau^*\|$. Next, by expanding the time-derivative of $W(\tilde x)$:

$$\varepsilon\dot W(\tilde x) = \tilde x^\mathsf{T}(F_{11}^\mathsf{T}P_x + P_xF_{11})\tilde x + 2\tilde x^\mathsf{T}P_xF_{xz}\tilde z + 2\tilde x^\mathsf{T}P_xF_{14}\dot w_t.$$

Using $F_{11} = A - \varepsilon\eta_u A^{-1}BG^\mathsf{T}T_{y,t}C$, $A^\mathsf{T}P_x + P_xA = -Q_x$:

$$\begin{aligned}\tilde x^\mathsf{T}(F_{11}^\mathsf{T}P_x + P_xF_{11})\tilde x &= -\tilde x^\mathsf{T}Q_x\tilde x \\ &- \eta_u\varepsilon\tilde x^\mathsf{T}(C^\mathsf{T}T_{y,t}GB^\mathsf{T}A^{-\mathsf{T}}P_x + P_xA^{-1}BG^\mathsf{T}T_{y,t}C)\tilde x.\end{aligned}$$

Let $\Sigma_1 := 2P_z[F_{21}^\mathsf{T}, F_{31}^\mathsf{T}]^\mathsf{T}$, $\Sigma_2 := 2\varepsilon^{-1}P_x[F_{12}, F_{13}]$, $\Sigma_3 = \eta_u(C^\mathsf{T}T_{y,t}GB^\mathsf{T}A^{-\mathsf{T}}P_x + P_xA^{-1}BG^\mathsf{T}T_{y,t}C)$, and $\Sigma_4 = P_xA^{-1}E$. Then,

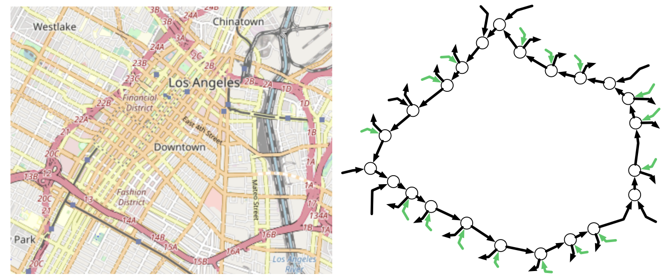
$$\begin{aligned}\varepsilon\dot W(\tilde x) &\le -\underline\lambda(Q_x)\|\tilde x\|^2 + \varepsilon\|\Sigma_2\|\|\tilde x\|\|\tilde z\| \qquad (45)\\ &\quad + \varepsilon\|\Sigma_3\|\|\tilde x\|^2 + 2\|\Sigma_4\|\|\tilde x\|\|\dot w_t\| \\ &\le -\frac{\underline\lambda(Q_x)}{2}\|\tilde x\|^2 + \varepsilon\|\Sigma_2\|\|\tilde x\|\|\tilde z\| + \varepsilon\|\Sigma_3\|\|\tilde x\|^2,\end{aligned}$$

where the last inequality holds when $-\frac{\underline\lambda(Q_x)}{2}\|\tilde x\|^2 + 2\|\Sigma_4\|\|\tilde x\|\|\dot w_t\| \le 0$, or $\|\tilde x\| \ge \frac{4\|\Sigma_4\|}{\underline\lambda(Q_x)}\operatorname{ess\,sup}_{\tau\ge0}\|\dot w_\tau\|$. By using $V(z) \le \bar\lambda(P_z)\|\tilde z\|^2$, $W(z) \le \bar\lambda(P_x)\|\tilde x\|^2$, by letting $\xi := (\|\tilde z\|, \|\tilde x\|)$, and by combining (43)-(45) we get $\dot U(\tilde x, \tilde z) \le -\xi^\mathsf{T}\Lambda\xi - \rho_\xi U(\tilde x, \tilde z)$, where:

$$\Lambda = \begin{bmatrix}(1-\theta)\frac{\bar\rho_z\underline\lambda(P_z)}{4} & -\frac{1}{2}((1-\theta)\|\Sigma_1\| + \theta\|\Sigma_2\|) \\ -\frac{1}{2}((1-\theta)\|\Sigma_1\| + \theta\|\Sigma_2\|) & \theta(\frac{\underline\lambda(Q_x)}{4\varepsilon} - \|\Sigma_3\|)\end{bmatrix}.$$

Matrix Λ is positive definite when

$$\theta(1-\theta)\frac{\rho_z\underline\lambda(P_z)\underline\lambda(Q_x)}{16\varepsilon} > \frac{1}{4}((1-\theta)\|\Sigma_1\| + \theta\|\Sigma_2\|)^2,$$



(a) (b)

Fig. 3. (a) Portion of highway system in Los Angeles, CA, USA. (b) Network schematic. Links colored in green represent controllable on-ramps.

which holds when the following is satisfied:

$$\varepsilon < \frac{\rho_z\underline\lambda(P_x)\underline\lambda(P_z)}{16\|\Sigma_1\|\|\Sigma_2\| + 4\rho_z\underline\lambda(P_z)\|\Sigma_3\|}.$$

The bound (36) is then obtained using standard manipulations. Finally, the claim follows by application of Lemma 3.14 with $\bar a = \max\{\bar\lambda(P_x), \bar\lambda(P_z)\}$, $\underline a = \min\{\underline\lambda(P_x), \underline\lambda(P_z)\}$, $b = \frac{1}{4}\min\left\{\rho_z\frac{\underline\lambda(P_z)}{\bar\lambda(P_z)}, \varepsilon^{-1}\frac{\underline\lambda(Q_x)}{\bar\lambda(P_x)}\right\}$, and $b_0 = \max\{\frac{4\|P_z\|}{\rho_z\underline\lambda(P_z)}\operatorname{ess\,sup}_\tau\|\dot z_\tau^*\|, \frac{4\|\Sigma_4\|}{\underline\lambda(Q_x)}\operatorname{ess\,sup}_\tau\|\dot w_\tau\|\}$. ∎

V. Application Example: Ramp Metering Control

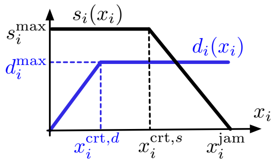
In this section, we apply the proposed framework to the control of on-ramps in a network of traffic highways.

To describe the traffic evolution, we adopt a continuous-time version of the Cell-Transmission Model (CTM) [33]. We model a traffic network as a directed graph $\mathcal G = (\mathcal V, \mathcal L)$, where $\mathcal V$ models the set of traffic junctions (nodes) and $\mathcal L \subseteq \mathcal V \times \mathcal V$ models the set of highways (links). We partition $\mathcal L = \mathcal L_{\text{on}} \cup \mathcal L_{\text{off}} \cup \mathcal L_{\text{in}}$, where $\mathcal L_{\text{on}}$ is the set of on-ramps where vehicles can enter, $\mathcal L_{\text{off}}$ is the set of off-ramps where vehicles can exit, and $\mathcal L_{\text{in}}$ is the set of internal links. For $i \in \mathcal L$, we denote by i^+ the set of downstream links, and by i^- the set of upstream links. For all $i \in \mathcal L$, we let $x_i : \mathbb R_{\ge0} \to \mathbb R_{\ge0}$ be the density of vehicle in the link. We model the dynamics of all links $i \in \mathcal L_{\text{in}}$ according to the CTM with FIFO allocation policy [33]:

$$\begin{aligned}\dot x_i &= -f_i^{\text{out}}(x) + f_i^{\text{in}}(x), \\ f_i^{\text{out}}(x) &= \min\{d_i(x_i), \{s_j(x_j)/r_{ij}\}_{j\in i^+}\}, \\ d_i(x_i) &= \min\{\varphi_i x_i, d_i^{\max}\}, s_i(x_i) = \min\{\beta_i(x_i^{\text{jam}} - x_i), s_i^{\max}\}, \\ f_i^{\text{in}}(x) &= \sum_{j\in i^-} f_j^{\text{out}}(x), \qquad (46)\end{aligned}$$

where $d_i : \mathbb R_{\ge0} \to \mathbb R_{\ge0}$ and $s_i : \mathbb R_{\ge0} \to \mathbb R_{\ge0}$ are the link demand and supply functions, respectively, $r_{ij} \in [0,1]$ is the routing ratio from i to j, with $\sum_j r_{ij} = 1$, $\varphi_i > 0$. We refer to Fig. 4 for a description of the parameters that characterize the demand and supply functions. The dynamics of on-ramps and off-ramps coincide with those of (46), where inflow and outflow functions are replaced, respectively, by:

$$\begin{aligned}f_i^{\text{in}}(x) &:= u_i, && \text{if } i \in \mathcal L_{\text{on}}, \\ f_i^{\text{out}}(x) &:= d_i(x_i), && \text{if } i \in \mathcal L_{\text{off}}, \qquad (47)\end{aligned}$$

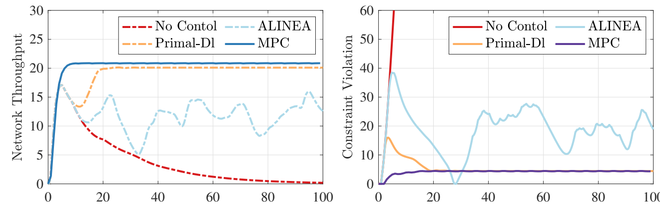


(a)

Variable	Description
φ_i	free-flow speed
β_i	back propag. speed
$d_i^{\max}$	demand saturation
$s_i^{\max}$	supply saturation
x_i^{jam}	jam density
$x_i^{\text{crt,d}}$	critical dens. of d_i
$x_i^{\text{crt,s}}$	critical dens. of s_i

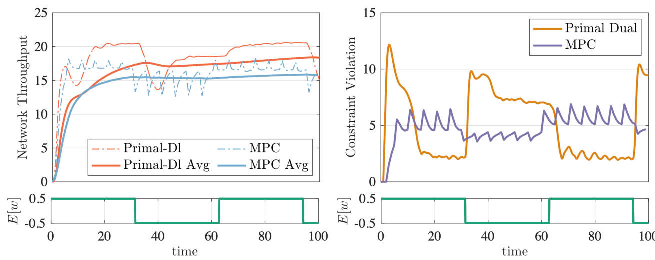
(b)

Fig. 4. (a) Demand and supply functions. (b) Parameters description.



(a) (b)

Fig. 5. Plant without noise. (a) Network throughput $\boldsymbol{\Phi}(\boldsymbol{x})$. (b) Constraint violation computed as $||\boldsymbol{y} - \min\{\boldsymbol{x}^{\text{crt},\boldsymbol{d}}, \boldsymbol{x}^{\text{crt},\boldsymbol{s}}\}||$.



(a) (b)

Fig. 6. Plant subject to random noise (green line shows noise mean). (a) Throughput $\boldsymbol{\Phi}(\boldsymbol{x})$. (b) Constraint violation: $||\boldsymbol{y} - \min\{\boldsymbol{x}^{\text{crt},\boldsymbol{d}}, \boldsymbol{x}^{\text{crt},\boldsymbol{s}}\}||$.

We assume the availability of measurements that provide a noisy estimate of the traffic densities in the highways: $y_i = x_i + w_i$, for all $i \in \mathcal{L}$, where $w_i : \mathbb{R}_{\geq 0} \to \mathbb{R}$. Finally, we define the network throughput as the sum of all exit flows from the off-ramps $\Phi(x) := \sum_{i \in \mathcal{L}_{\text{off}}} f_i^{\text{out}}(x)$. The on-ramp metering problem is formalized as follows.

Problem 2: (Ramp Metering) Given a vector of on-ramp flow demands $u_{\text{ref}} \in \mathbb{R}^m$, select the set of metered flows on the on-ramps $(u_1, \dots, u_m)$ such that u and x minimize the cost $(u - u^{\text{ref}})^\mathsf{T} Q_u (u - u^{\text{ref}}) - \Phi(x)$, subject to the constraints (46)-(47), where $Q_u \in \mathbb{R}^{n \times n}$ is symmetric and positive definite.

The following strategies are compared.

1) Online Primal-Dual Controller: To solve Problem 2, we assume that for all $i \in \mathcal{L}$ one has that $d_i^{\max} \leq s_j^{\max}$ for all $j \in i^+$. (i.e., highway j is not congested and, precisely, $x_j \leq x_j^{\text{crt,s}}$). thus, when the network is maintained in a regime in which all highways are not congested, namely, $x_i \leq \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}$ for all $i \in \mathcal{L}$, then the dynamics (46) simplify to the following linear model:

$$\begin{aligned} \dot{x}_i &= -f_i^{\text{out}}(x) + f_i^{\text{in}}(x), \\ f_i^{\text{out}}(x) &= \varphi_i x_i, \qquad f_i^{\text{in}}(x) = \sum_{j \in i^-} f_j^{\text{out}}(x). \end{aligned} \tag{48}$$

In vector form, one has $\dot{x} = (R^\mathsf{T} - I)Fx + Bu$, and $y = x + w$, where $R := [r_{ij}]$, and $F := \operatorname{diag}(\varphi_1, \dots, \varphi_n)$. Notice that matrix $(R^\mathsf{T} - I)F$ is Hurwitz (see e.g. [34, Theorem 1]). Building on this, we propose the following problem:

$$\begin{aligned} \min_{u,y} \quad & (u - u^{\text{ref}})^\mathsf{T} Q_u (u - u^{\text{ref}}) - \Phi(y), \\ \text{s.t.} \quad & y = -((R^\mathsf{T} - I)F)^{-1} Bu + w, \\ & u_i \geq 0, \quad y_i \leq \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}, \forall i \in \mathcal{L}. \end{aligned} \tag{49}$$

The optimization problem (49) formalizes the objectives of the ramp metering problem, while maintaining all highways in the network in the free-flow regime by design.

2) Distributed Reactive Metering using ALINEA: ALINEA [24] is a distributed metering strategy that has received considerable interest thanks to its simplicity of implementation and to its effectiveness. Given a controllable on-ramp $i \in L_{\text{in}}$, ALINEA is a reactive controller that takes the form $\dot{u}_i = \sum_{j \in i^+} K_j(\hat{x}_j - x_j)$, where $\hat{x}_i \in \mathbb{R}_{\geq 0}$ is a desired setpoint and K_j are tunable controller gains. In our simulations, we let $\hat{x}_i = \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}$.

3) Model Predictive Control (MPC): MPC is a receding-horizon control algorithm that computes an optimal control input based on a prediction of the system's future trajectory according to the system's dynamics. We denote by $q_{[t_1,t_2]}$ the restriction to the time-interval $[t_1, t_2]$ of the signal $t \mapsto q(t)$, and we consider an MPC formulation of Problem 2 with prediction horizon $T_p \in \mathbb{R}_{>0}$ and $\hat{x}_t = x(t + T_s)$ as the initial condition. Usually, after an optimal control policy $u^*_{[0,T_p]}$ has been computed by setting $\hat{x}_t = x(t)$, only the input $u_{[0,T_s]}$, $0 \leq T_s \leq T_p$, is implemented in the system for the time window $[0, T_s]$, and then a new optimal policy is re-computed by solving the MPC with $\hat{x}_t = x(t + T_s)$, in order to account for model inaccuracies. In our simulations, we solved MPC by discretizing the dynamics and we used $T_p = 20$ and $T_s = 5$.

4) Discussion: Fig. 5 compares the performance of the controllers described in Section V for the noiseless case, in which $w_i = 0$ at all times for all $i \in \mathcal{L}$. The simulation demonstrates that our method and MPC achieve the largest network throughput, outperforming ALINEA. Moreover, the constraint violation plot (right figure) shows that both our method and MPC are able to maintain the network in a regime near the free-flow conditions. Notice that, while for MPC this regime is precisely modeled through the prediction equations, the primal-dual controller maintains the system in such regime thanks to the constraints in the optimization problem (49). Finally, although ALINEA is certainly better than no on-ramp metering at all, it suffers considerably from its distributed architecture, which lacks a global system model.

Fig. 6 compares the performance of our controller with that of MPC in a scenario with time-varying output disturbance (depicted in green). There are two main benefits in adopting primal-dual controllers as compared to MPC: (i) the primal-dual controller uses instantaneous feedback from the system, which enables it to react faster to unmodeled dynamics or time-varying disturbances, and (ii) the primal-dual controller is computationally more efficient as compared to MPC, since it leverages only part of the full network model (precisely, the model corresponding to the free-flow regime).

VI. Conclusions

We have leveraged online primal-dual dynamics to develop an output controller that regulates an LTI plant to the solution of a time-varying optimization problem. For optimization problems with input constraints and output inequality constraints, we leveraged an augmented Lagrangian function and established exponential convergence to an approximate solution of the optimization problem. For optimization problems with output equality constraints, we established exponential convergence to an interval around the exact optimal solution trajectory. Our convergence bounds capture the time-variability of the optimal solution due to time-varying costs and constraints as well as the variation of the exogenous input. Topics that will be the subject of future investigations include extensions to time-varying input constraint sets and generic convex constraints on the system output.

References

[1] A. Jokic, M. Lazar, and P. P. V. D. Bosch, "On constrained steady-state regulation: Dynamic KKT controllers," *IEEE Trans. on Automatic Control*, vol. 54, no. 9, pp. 2250–2254, 2009.

[2] F. D. Brunner, H.-B. Dürr, and C. Ebenbauer, "Feedback design for multi-agent systems: A saddle point approach," in *IEEE Conf. on Decision and Control*, 2012, pp. 3783–3789.

[3] L. S. P. Lawrence, Z. E. Nelson, E. Mallada, and J. W. Simpson-Porco, "Optimal steady-state control for linear time-invariant systems," in *IEEE Conf. on Decision and Control*, Dec. 2018, pp. 3251–3257.

[4] L. S. P. Lawrence, J. W. Simpson-Porco, and E. Mallada, "Linear-convex optimal steady-state control," *arXiv preprint*, p. arXiv:1810.12892, 2018.

[5] M. Colombino, E. Dall'Anese, and A. Bernstein, "Online optimization as a feedback controller: Stability and tracking," *IEEE Trans. on Control of Network Systems*, vol. 7, no. 1, pp. 422–432, 2020.

[6] S. Menta, A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Stability of dynamic feedback optimization with applications to power systems," in *Annual Conf. on Communication, Control, and Computing*, Oct. 2018, pp. 136–143.

[7] T. Zheng, J. W. Simpson-Porco, and E. Mallada, "Implicit trajectory planning for feedback linearizable systems: A time-varying optimization approach," *arXiv preprint*, 2019, arXiv:1910.00678.

[8] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Timescale separation in autonomous optimization," *IEEE Trans. on Automatic Control*, 2020, (In press).

[9] R. Li and G.-H. Yang, "Optimal steady-state regulator design for a class of nonlinear systems with arbitrary relative degree," *IEEE Trans. on Cybernetics*, 2020, in press.

[10] G. Bianchin, J. I. Poveda, and E. Dall'Anese, "Online optimization of switched LTI systems using continuous-time and hybrid accelerated gradient flows," *arXiv*, Aug. 2020, arXiv:2008.03903.

[11] G. Qu and N. Li, "On the exponential stability of primal-dual gradient dynamics," *IEEE Control Systems Letters*, vol. 3, no. 1, pp. 43–48, 2018.

[12] D. Ding and M. R. Jovanovic, "Global exponential stability of primal-dual gradient flow dynamics based on the proximal augmented lagrangian," *American Control Conference*, pp. 3414–3419, 2019.

[13] J. Cortés and S. K. Niederländer, "Distributed coordination for nonsmooth convex optimization via saddle-point dynamics," *Journal of Nonlinear Science*, vol. 29, no. 4, pp. 1247–1272, 2019.

[14] H. D. Nguyen, T. L. Vu, K. Turitsyn, and J.-J. Slotine, "Contraction and robustness of continuous time primal-dual dynamics," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 755–760, 2018.

[15] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[16] J. I. Poveda, R. Kutadinata, C. Manzie, D. Nešić, A. R. Teel, and C. Liao, "Hybrid extremum seeking for black-box optimization in hybrid plants: An analytical framework," *IEEE Conf. on Decision and Control*, pp. 2235–2240, 2018.

[17] A. Hauswirth, I. Subotić, S. Bolognani, G. Hug, and F. Dörfler, "Time-varying projected dynamical systems with applications to feedback optimization of power systems," in *IEEE Conf. on Decision and Control*, 2018, pp. 3258–3263.

[18] A. Hauswirth, F. Dörfler, and A. Teel, "Anti-windup approximations of oblique projected dynamics for feedback-based optimization," *arXiv preprint*, 2020, arXiv:2003.00478.

[19] R. Goebel, "Stability and robustness for saddle-point dynamics through monotone mappings," *Systems and Control Letters*, vol. 108, pp. 16–22, 2017.

[20] J. I. Poveda and N. Li, "Robust hybrid zero-order optimization algorithms with acceleration via averaging in time," *Automatica*, vol. 123, 2021.

[21] D. Liao-McPherson, M. M. Nicotra, and I. Kolmanovsky, "Time-distributed optimization for real-time model predictive control: Stability, robustness, and constraint satisfaction," *Automatica*, vol. 117, p. 108973, 2020.

[22] M. Figura, L. Su, V. Gupta, and M. Inoue, "Instant distributed model predictive control for constrained linear systems," in *American Control Conference*, Denver, CO, July 2020, pp. 4582–4587.

[23] J. Koshal, A. Nedić, and U. V. Shanbhag, "Multiuser optimization: Distributed algorithms and error analysis," *SIAM Journal on Optimization*, vol. 21, no. 3, pp. 1046–1081, 2011.

[24] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 271–281, 2002.

[25] P. Cisneros-Velarde, S. Jafarpour, and F. Bullo, "Distributed and time-varying primal-dual dynamics via contraction analysis," *arXiv preprint*, Mar. 2020, arXiv:2003.12665.

[26] H. A. Edwards, Y. Lin, and Y. Wang, "On input-to-state stability for time varying nonlinear systems," in *IEEE Conf. on Decision and Control*, vol. 4, 2000, pp. 3501–3506.

[27] A. Nagurney and D. Zhang, *Projected dynamical systems and variational inequalities with applications*. Springer, 2012, vol. 2.

[28] X.-B. Gao, "Exponential stability of globally projected dynamic systems," *IEEE Trans. on Neural Networks*, vol. 14, no. 2, pp. 426–431, 2003.

[29] Y. S. Xia and J. Wang, "On the stability of globally projected dynamical systems," *Journal of Optimization Theory & Applications*, vol. 106, no. 1, pp. 129–150, 2000.

[30] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović, "A second order primal-dual method for nonsmooth convex composite optimization," *arXiv preprint*, Sep. 2017, arXiv:1709.01610.

[31] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *IEEE Conf. on Decision and Control*, Phoenix, AZ, USA, Dec 1999, pp. 2655–2660.

[32] W. P. Dayawansa and C. F. Martin, "A converse lyapunov theorem for a class of dynamical systems which undergo switching," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 751–760, 1999.

[33] C. F. Daganzo, "The cell transmission model pt. II: network traffic," *Transp. Research Pt. B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995.

[34] G. Bianchin and F. Pasqualetti, "Gramian-based optimization for the analysis and control of traffic networks," *IEEE Trans. on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 3013–3024, 2020.

Gianluca Bianchin (S'15, M'20) is a Postdoctoral Scholar in the Department of Electrical, Computer, and Energy Engineering at the University of Colorado Boulder. He received the Doctor of Philosophy degree in Mechanical Engineering at the University of California Riverside, in 2020, the Laurea degree in Information Engineering and the Laurea Magistrale degree in Controls Engineering at the University of Padova, Italy, in 2012 and 2014, respectively. He was the recipient of the Dissertation Year Award and of the Dean's Distinguished Fellowship Award from the University of California Riverside.

Jorge Cortés (M'02, SM'06, F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE and SIAM. At the IEEE Control Systems Society, he has been a Distinguished Lecturer (2010-2014) and an elected member (2018-2020) of its Board of Governors, and is currently its Director of Operations.

Jorge I. Poveda is an Assistant Professor in the Department of Electrical, Computer, and Energy Engineering at the University of Colorado, Boulder. He received in 2016 and 2018 the M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of California at Santa Barbara. He was a Postdoctoral Fellow at Harvard University in 2018, and a Research Intern at the Mitsubishi Electric Research laboratories during the summers of 2016 and 2017. He has received the CCDC Outstanding Scholar Fellowship at UCSB, and the National Science Foundation Career Research Initiation Award (CRII) in 2020.

Emiliano Dall'Anese is an Assistant Professor in the Department of Electrical, Computer, and Energy Engineering at the University of Colorado Boulder, and an affiliate Faculty with the Department of Applied Mathematics. He received the Ph.D. in Information Engineering from the Department of Information Engineering, University of Padova, Italy, in 2011. From January 2011 to November 2014, he was a Postdoctoral Associate at the Department of Electrical and Computer Engineering of the University of Minnesota, and from December 2014 to July 2018 he was a Senior Researcher at the National Renewable Energy Laboratory. He received the National Science Foundation CAREER Award in 2020.