

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 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 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 system towards the solution trajectory of a *time-varying* optimization problem (see Fig. 1). Such problems correspond to scenarios where cost and constraints may change over time to reflect dynamic performance objectives or simply take into account time-varying unknown disturbances entering the system. This setting emerges in many engineering applications, including power systems [1], [2], transportation networks [3], [4], and communication systems [5].

The design of feedback controllers inspired from optimization algorithms has received significant attention during the last decade [1], [2], [6]–[13]. While most of the existing works focus on the design of optimization-based controllers for static problems [2], [6]–[9], [11], [12], or consider unconstrained time-varying problems [10], [13], 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, in this paper we consider optimization problems with a time-varying strongly convex cost, time-varying linear constraints on the output, and convex constraints on the input. We leverage online saddle-point dynamics for controller synthesis, and we present exponential input-to-state stability (ISS) [14] analysis of the proposed controller when interconnected with an LTI systems. In particular, we leverage tools from singular perturbation theory [15] to provide sufficient conditions on the tunable controller parameters to guarantee exponential stability and tracking of the

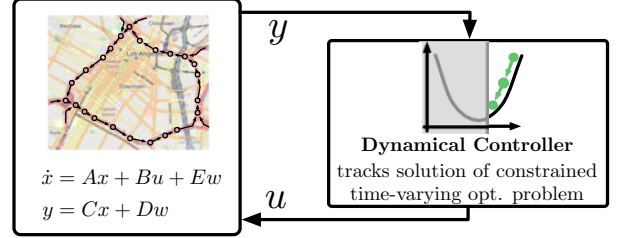


Fig. 1. Online saddle-flow optimizer used as an output feedback controller for LTI systems subject to unknown time-varying disturbances.

optimal solution trajectory. We remark that, while [16]–[20] 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 projected primal-dual dynamics are interconnected with a dynamical system subject to unknown exogenous disturbances (as in Fig. 1).

Related work. In the case of static plants (i.e., where the dynamics of the system are infinitely fast), controllers conceptually-inspired from continuous-time saddle-point dynamics (or flows) are studied in [6] for optimization problems with time-invariant costs and constraints on the system outputs, whereas more general saddle-point flows are studied in [7], [12], [21], and [22, Sec. 3]. While the above works focus on optimization problems with static plants, the authors in [2], [8], [9], [11] prove that gradient-flow dynamics can be used as feedback controllers for dynamical systems in the case of unconstrained optimization problems with time-invariant costs. The work [11] also extends these results to the case of constraints on the system inputs by using projected gradient flows. Constraints on the system outputs are considered in [1], together with a controller inspired from primal-dual dynamics based on the Moreau envelope. For time-varying unconstrained optimization problems, prediction-correction algorithms are used in [10]. Exponential rates of convergence were proved for the first time in [13] for dynamic controllers based on gradient flows and accelerated hybrid dynamics.

In terms of classes of plants, stable LTI systems are considered in [1], [2], [13], stable nonlinear systems in [11], input-linearizable systems in [10], and input-affine nonlinear system in [7]. Finally, [23], [24] consider online implementations of optimization problems arising in model predictive control.

Contributions. The contributions of this work are as follows. C1) We design an output feedback controller, inspired from primal-dual dynamics, which is able to steer a dynamical system to the solution trajectory of a time-varying constrained optimization problem without requiring information

G. Bianchin, J. I. Poveda, and E. Dall’Anese are with the Department of Electrical, Computer and Energy Engineering, University of Colorado Boulder. J. Cortés is with the Department of Mechanical and Aerospace Engineering at the University of California San Diego. This work was supported by the National Science Foundation (NSF) through the Awards CMMI 2044946 and 2044900 and CRII: CNS-1947613.

or measurements of the external disturbances entering the state equation. 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 [25] to guarantee exponential convergence to an approximate KKT trajectory. C2) We consider constraints on the system input and we propose a novel projected primal-dual feedback controller that guarantees constraint satisfaction. Differently from using the classical projection on the tangent cone, the proposed controller yields trajectories that are continuously differentiable, which allows us to simplify the analysis and to establish strong robustness guarantees. As a minor contribution, we demonstrate that the proposed framework is applicable to more-general LTI systems, including switched systems with common quadratic Lyapunov functions. C4) We apply the proposed controllers to solve a ramp metering problem in traffic systems. We compare our results with state-of-the-art controllers, including ALINEA [26] and MPC, illustrating the advantages of our method.

We emphasize that, relative to [1]: (i) our sufficient conditions are markedly easier to check a priori and do not require to numerically solve a linear matrix inequality, and (ii) our framework also accounts for input constraints and does not require to compute the Moreau envelope. Relative to [6], [7], [11], [12], we account for time variability in the cost functions and in the disturbances, and we prove explicit exponential bounds. Relative to [20], [27], we investigate the stability of saddle-point dynamics when coupled with a dynamical system.

Organization. We present in Section II our problem formulation. Section III develops a projected primal-dual output feedback controller for problems with input constraints and output inequality constraints. Section IV considers problems with output equality constraints. Section V presents numerical results by considering a ramp metering problem in traffic systems. Finally, Section VI concludes the paper.

Notation. Given vectors $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$, we let $(x, u) \in \mathbb{R}^{n+m}$ denote their concatenation. We use $\bar{\lambda}(M)$ and $\underline{\lambda}(M)$ to denote the largest and smallest eigenvalues of the square matrix M , respectively. Finally, $P_\Omega : \mathbb{R}^\sigma \rightarrow \mathbb{R}^\sigma$ denotes the Euclidean projection of z onto a closed convex set $\Omega \subseteq \mathbb{R}^\sigma$, namely $P_\Omega(z) := \arg \min_{v \in \Omega} \|z - v\|$.

II. PROBLEM FORMULATION

We consider LTI dynamical systems described by:

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

where $x : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$ is the state, $u : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^m$ is the input, $y : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^p$ 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 1: Matrix A is Hurwitz stable, i.e., there exists positive definite $P_x, Q_x \in \mathbb{R}^{n \times n}$ such that $A^\top P_x + P_x A = -Q_x$. Moreover, the columns of C are linearly independent. \square

Under Assumption 1, 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:

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

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 following time-varying optimization problem:

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

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

$$K_t \bar{y} \leq e_t, \quad (3c)$$

where for all $t \in \mathbb{R}_{\geq 0}$, $\phi_t : \mathbb{R}^m \rightarrow \mathbb{R}$, $\psi_t : \mathbb{R}^p \rightarrow \mathbb{R}$. Moreover, $t \mapsto K_t \in \mathbb{R}^{r \times p}$ and $t \mapsto e_t \in \mathbb{R}^r$ describe a time-varying output constraint, while $\mathcal{U} \subseteq \mathbb{R}^m$ denotes 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 loss functions ϕ_t and ψ_t . We note that, because the cost functions and constraints are time-varying, the solutions of (3) are also time-varying, and thus they characterize optimal trajectories. We impose the following regularity assumptions on the temporal evolution of (3).

Assumption 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 $y, y' \in \mathbb{R}^p$, $\|\nabla \phi_t(u) - \nabla \phi_t(u')\| \leq \ell_u \|u - u'\|$, $\|\nabla \psi_t(y) - \nabla \psi_t(y')\| \leq \ell_y \|y - y'\|$, 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, uniformly in t . \square

Assumption 3: Problem (3) is feasible, and Slater's condition [25] holds for each $t \in \mathbb{R}_{\geq 0}$. \square

Assumption 4: The following regularity properties hold.

- (a) $t \mapsto w_t$ is absolutely continuous and locally Lipschitz.
- (b) The functions $t \mapsto [K_t]_{ij}$ and $t \mapsto [e_t]_i$ $i = 1, \dots, r$, $j = 1, \dots, p$, are locally Lipschitz, and there exists $\bar{K} \in \mathbb{R}_{\geq 0}$, $\bar{e} \in \mathbb{R}_{\geq 0}$, such that $\|K_t\| < \bar{K}$ and $\|e_t\| < \bar{e}$. \square

Under Assumptions 2–3, the minimizer (u_t^*, y_t^*) of (3) is unique for every $t \in \mathbb{R}_{\geq 0}$ [6], while Assumption 4 guarantees that the inputs and the constraints of (3) vary continuously in time. The problem focus of this work is formalized next.

Problem 1: Design an output-feedback controller of the form $\dot{u} = \mathcal{C}(u, y)$ such that the input and output of (1) converge exponentially to the time-varying optimizer of (3), up to an asymptotic error that accounts for the temporal variability of both the optimizer and of the unknown disturbance. \square

III. CLOSED-LOOP PROJECTED SADDLE-POINT FLOWS

In this section, we present our controller synthesis method and we establish explicit convergence error bounds.

A. Controller Synthesis

For controller synthesis, we employ a regularized Lagrangian function and we use a controller structure that relies on a modification of the saddle-point flow dynamics [6]. Specifically, consider the following Lagrangian function for (3): $\mathcal{L}_t(u, \lambda) := \phi_t(u) + \psi_t(Gu + Hw_t) + \lambda^\top (K_t(Gu + Hw_t) - e_t)$, where $\lambda \in \mathbb{R}_{\geq 0}^r$ denotes the vector of dual variables. We define the regularized Lagrangian function as follows:

$$\mathcal{L}_{\nu,t}(u, \lambda) := \mathcal{L}_t(u, \lambda) - \frac{\nu}{2} \|\lambda\|^2, \quad (4)$$

where $\nu \in \mathbb{R}_{>0}$. The regularization term $-\frac{\nu}{2} \|\lambda\|^2$ has the effect of making the function $\mathcal{L}_{\nu,t}(u, \lambda)$ strongly concave in λ , for any $u \in \mathbb{R}^m$ (see [25]). As a result, the regularization term induces a saddle-point map that is strongly monotone, uniformly in time [25]. As characterized in Section III-B, this property is fundamental to prove exponential convergence of our controller. On the other hand, the introduction of a regularization term comes at the cost of perturbing the saddle points. To this aim, we let

$$z_t^* := (u_t^*, \lambda_t^*), \quad z_{\nu,t}^* := (u_{\nu,t}^*, \lambda_{\nu,t}^*), \quad (5)$$

denote any saddle-point of $\mathcal{L}_t(u, \lambda)$, and the unique saddle point of $\mathcal{L}_{\nu,t}(u, \lambda)$, respectively. We quantify the error introduced by the regularization term in the following result (adapted from [25, Prop. 3.1]).

Lemma 3.1: Let Assumptions 2-3 hold. For each $t \in \mathbb{R}_{\geq 0}$, the following bound holds:

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

where (u_t^*, λ_t^*) and $(u_{\nu,t}^*, \lambda_{\nu,t}^*)$ are as in (5). In particular, it holds that $\|u_{\nu,t}^* - u_t^*\| \leq \sqrt{\frac{\nu}{2\mu_u}} \|\lambda_t^*\|$.

Remark 1: Lemma 3.1 states 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 is strictly inside the feasible set, then $\lambda_t^* = 0$ and the solution $u_{\nu,t}^*$ coincides with u_t^* . \square

We consider the following modification of the gradients of the augmented Lagrangian:

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

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

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 $Gw_t + Hw_t$ has been replaced by the variable y . Using (7), we propose the following *online projected primal-dual controller* applied to the system (1) (see Fig. 1 for an illustration):

$$\dot{x} = Ax + Bu + Ew_t, \quad y = Cx + Dw_t, \quad (8a)$$

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

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

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

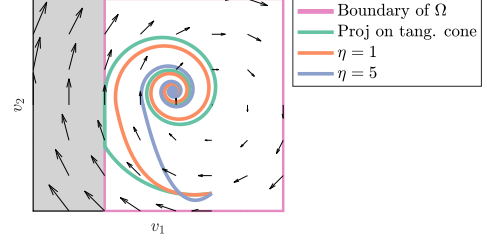


Fig. 2. Comparison between trajectories of (9) and of the smooth projection (8) for a 2-D vector field. Black arrows show the vector field.

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 [15].

Remark 2: Note that the choice of dualizing the constraint $K_t(Gu + Hw_t) \leq e_t$ allows us to naturally enforce constraints that are time-varying and parametrized by the unknown vector w_t . This is because the steady-state relationship $Gw_t + Hw_t$ is replaced by instantaneous feedback y_t in (7b). The alternative route of combining the constraint $K_t(Gu + Hw_t) \leq e_t$ with the convex constraint $u \in \mathcal{U}$ and recast both of them as a convex constraint of the form $u \in \mathcal{U} \cap \{u : K_t(Gu + Hw_t) \leq e_t\}$ would result in an unknown constraint set, thus making the computation of the projection not possible. \square

Remark 3: Given a closed convex set $\Omega \subseteq \mathbb{R}^\sigma$ and a vector field $F : \Omega \rightarrow \mathbb{R}^\sigma$, the standard projected dynamical system [28] associated with $F(v)$ is given by:

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

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. [29], [30]. However, to the best of our knowledge, (8b)-(8c) is the first projected output feedback controller with Lipschitz-continuous vector fields. \square

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).

B. Stability and Tracking Analysis

In this section we characterize the transient behavior of (8). To this aim, in what follows we use the notation:

$$z := (u, \lambda), \quad \tilde{z}_{\nu} := z - z_{\nu,t}^*, \quad (10)$$

to denote the joint controller state and the controller tracking error, where $z_{\nu,t}^*$ is as in (5). Similarly, we use

$$\xi := (x, u, \lambda), \quad \xi_{\nu,t}^* := (x_{\nu,t}^*, u_{\nu,t}^*, \lambda_{\nu,t}^*), \quad \tilde{\xi}_{\nu} := \xi - \xi_{\nu,t}^*, \quad (11)$$

to denote the joint state of (8), the saddle-point of (4), with $x_{\nu,t}^* = -A^{-1}(Bu_{\nu,t}^* + Hw_t)$, and the tracking error, respectively. We begin by characterizing the existence of solutions.

Lemma 3.2: Let Assumptions 2–4 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 [29], [30], (ii) under Assumptions 2–4, 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 4(a) the plant dynamics are locally Lipschitz in t . ■ Lemma 3.2 guarantees

that the trajectories of (8) are continuously differentiable (see Fig. 2). Moreover, since trajectories are maximal, Lemma 3.2 guarantees that trajectories have no finite escape time. The latter property is harnessed to prove the following result, which establishes attractivity and forward invariance of the feasible set (see [30, Thm 3.2]).

Lemma 3.3: Let Assumptions 2–4 hold. If $u(t_0) \notin \mathcal{U}$ (resp. $\lambda(t_0) \notin \mathcal{C}$) for some $t_0 \in \mathbb{R}_{\geq 0}$, then the trajectory $u(t)$ (resp. $\lambda(t)$) approaches exponentially the set \mathcal{U} (resp. the set \mathcal{C}) for $t > t_0$. If $u(t_0) \in \mathcal{U}$ (resp. $\lambda(t_0) \in \mathcal{C}$) for some $t_0 \in \mathbb{R}_{\geq 0}$, then $u(t) \in \mathcal{U}$ (resp. $\lambda(t) \in \mathcal{C}$) for all $t \geq t_0$.

Remark 4: Lemma 3.3 guarantees that, if $u(0) \in \mathcal{U}$, then the convex constraint $u(t) \in \mathcal{U}$ is satisfied at all times. In contrast, because the output constraint (3c) is dualized in (4), the inequality $K_t y(t) \leq e_t$ is guaranteed to hold only as $t \rightarrow \infty$, even when $K_t y(t_0) \leq e_t$ for some $t_0 \in \mathbb{R}_{\geq 0}$. □

The following lemma establishes a relationship between the saddle-point of the regularized Lagrangian (4) and the equilibria of (8). The proof is omitted due to space limitations.

Lemma 3.4: Let Assumptions 1–4 hold. For any $t \in \mathbb{R}_{\geq 0}$ and for any fixed $w_t \in \mathbb{R}^q$, let $\xi_{\text{eq}} := (x_{\text{eq}}, u_{\text{eq}}, \lambda_{\text{eq}})$ denote an equilibrium of (8), and let $\xi_{\nu,t}^*$ denote the unique saddle-point of (4), as defined in (11). Then, ξ_{eq} is unique and $\xi_{\text{eq}} = \xi_{\nu,t}^*$.

To characterize the transient behavior of (8), we begin by showing that, when the dynamics of the plant (1) are infinitely fast, the controller (8b)–(8c) converges exponentially to the saddle-point of the regularized Lagrangian, modulo an asymptotic error that depends on the time-variability of the optimizer z_t^* .

Proposition 3.5: Let Assumptions 1–4 hold, let $\mu := \min\{\mu_u, \nu\}$, $\ell := \sqrt{2}(\bar{K} + \max\{\ell_u + \|G\|^2 \ell_y, \nu\})$. If $\varepsilon = 0$ and the controller gain satisfies $\eta < \frac{4\mu}{\ell^2}$, then for any $t_0 \in \mathbb{R}_{\geq 0}$:

$$\|\tilde{z}_\nu(t)\| \leq e^{-\frac{1}{2}\rho_z(t-t_0)} \|\tilde{z}_\nu(t_0)\| + \frac{2}{\rho_z} \text{ess sup}_{\tau \geq t_0} \|\dot{z}_{\nu,\tau}^*\|, \quad (12)$$

for all $t \geq t_0$, where $\rho_z = \eta(\mu - \frac{\eta \ell^2}{4})$, and \tilde{z}_ν denotes the controller tracking error as in (10).

The proof of this result is presented in Appendix A. Proposition 3.5 guarantees that (8) is input-to-state stable [14] with respect to the time-variability of the optimizer (here, $\dot{z}_{\nu,\tau}^*$ denotes the distributional derivative [31] of $z_{\nu,\tau}^*$, see Remark 5 for a thorough discussion). Moreover, the result demonstrates that the rate of convergence ρ_z can be tuned by properly tuning the controller gain η .

Remark 5: We note that, under Assumptions 1–4, the saddle-point trajectory $t \mapsto z_{\nu,t}^*$ is locally Lipschitz and hence absolutely continuous on compact sets. Thus, the essential supremum of $\|\dot{z}_{\nu,\tau}^*\|$ is well defined. To see this, notice that, $z_{\nu,\tau}^*$ solves the following Variational Inequality:

$$(u - u_{\nu,t}^*)(\nabla \psi_t(u_{\nu,t}^*) + G^\top \nabla \phi_t(Gu_{\nu,t}^* + Hw_t) + G^\top K_t^\top \lambda_{\nu,t}^*) \geq 0, \\ (\lambda - \lambda_{\nu,t}^*)(K_t(Gu_{\nu,t}^* + Hw_t) - e_t - \nu \lambda_{\nu,t}^*) \geq 0,$$

for all $u \in \mathcal{U}$ and $\lambda \in \mathcal{C}$, which is parametrized by the temporal variable t . It follows from Assumptions 1–4 and from our regularization method (4) that the mapping defining the above variational inequality is locally Lipschitz in (u, λ) , and thus [32, Cor. 2B.3] guarantees that $z_{\nu,\tau}^*$ is locally Lipschitz. Hence, by Rademacher’s theorem [33, Thm. 23.2], $t \mapsto z_{\nu,\tau}^*$ is differentiable almost everywhere (a.e.). □

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.

Theorem 3.6: Let Assumptions 1–4 hold, let $\ell := \sqrt{2}(\bar{K} + \max\{\ell_u + \|G\|^2 \ell_y, \nu\})$ and $\mu := \min\{\mu_u, \nu\}$. If

$$\eta < \frac{4\mu}{\ell^2} \quad \text{and} \quad \varepsilon < \frac{\rho_z \underline{\lambda}(Q_x)}{4\eta \|P_x A^{-1} B\| \Psi}, \quad (13)$$

where $\rho_z = \eta(\mu - \frac{\eta \ell^2}{4})$, $\Psi = \rho_z \ell_y \|C\| \|G\| + \sqrt{2} \|C\| (\ell_y \|G\| + \bar{K}) k_0$, $k_0 = \max\{2 + \eta(\ell_u + \ell_y \|G\|^2), \|G\| \bar{K}\}$, and P_x, Q_x are as in Assumption 1, then for any $t_0 \in \mathbb{R}_{\geq 0}$:

$$\|\tilde{\xi}_\nu(t)\| \leq \sqrt{\kappa} \|\tilde{\xi}_\nu(t_0)\| e^{-\frac{1}{2}\rho_\xi(t-t_0)} + \frac{2}{\rho_z} \text{ess sup}_{\tau \geq t_0} \|\dot{z}_{\nu,\tau}^*\| \\ + \frac{4\varepsilon \|P_x A^{-1} E\|}{\underline{\lambda}(Q_x)} \text{ess sup}_{\tau \geq t_0} \|\dot{w}_\tau\|, \quad (14)$$

for all $t \geq t_0$, where $\rho_\xi = \frac{1}{2} \min\left\{2\rho_z, \frac{1}{4\varepsilon} \frac{\underline{\lambda}(Q_x)}{\bar{\lambda}(P_x)}\right\}$, $\kappa = \max\{\frac{1}{2}, \bar{\lambda}(P_x)\} / \min\{\frac{1}{2}, \underline{\lambda}(P_x)\}$, and ξ_ν is as in (11).

The proof of this result is presented in Appendix A. Theorem 3.6 shows that, under a sufficient separation between the time scales of the plant and of the controller, the trajectories of (8) globally exponentially converge to $\xi_{\nu,t}^*$ (which we recall is the trajectory of the unique saddle-point of the regularized Lagrangian), modulo an asymptotic error that depends on the time-variability of the optimizer and of the exogenous disturbance. Precisely, Theorem 3.6 guarantees that (8) is input-to-state stable [14] with respect to $\dot{z}_{\nu,\tau}^*$ and \dot{w}_τ , where \dot{w}_τ denotes the distributional derivative [31] of w_τ (notice that, under Assumption 4(a), $\tau \mapsto w_\tau$ is differentiable a.e.).

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 $\|P_x A^{-1} B\|$. Since $\|A^{-1}\| \rightarrow 0$ 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, respectively.

Remark 6: The bound (14) depends on two main quantities: $\text{ess sup}_{\tau \geq t_0} \|\dot{z}_{\nu, \tau}^*\|$, which captures the time-variability of $z_{\nu, t}^*$, and $\text{ess sup}_{\tau \geq t_0} \|\dot{w}_\tau\|$, which captures the variation of the equilibrium of (1) induced by the presence of a time-varying exogenous input w_t . Notably, when the optimization problem (3) is time-invariant, (14) simplifies to an exponential stability result of the form $\|\tilde{\xi}_\nu(t)\| \leq \sqrt{\kappa} \|\tilde{\xi}_\nu(t_0)\| e^{-\frac{1}{2}\rho_\varepsilon(t-t_0)}$. \square

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 3.6 can be extended to consider switched LTI plants of the form:

$$\begin{aligned} \dot{x} &= A_\sigma x + B_\sigma u + E_\sigma w_t, \\ y &= Cx + Dw_t, \end{aligned} \quad (15)$$

where $\sigma : \mathbb{R}_{\geq 0} \rightarrow \mathcal{Q}$ is a switching signal taking values in the finite set \mathcal{Q} . When all modes of (15) have a common equilibrium point $x_{\text{eq}}^* = A_\sigma^{-1}Bu + A_\sigma^{-1}E_\sigma w_t$ for all values of σ and admit a common quadratic Lyapunov function V , the same construction for the Lyapunov function (42) can be used to establish exponential ISS of the closed-loop system. Since in this case G and H in (2) are also common across the modes, the bounds of Theorem 3.6 still hold. This observation is relevant in applications where internal feedback controllers have been implemented *a priori* to stabilize the modes of the plant, such that Assumption 1 holds. In this context, different controller may lead to closed-loop transient performance, while preserving a common equilibrium point [34]. 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 [35]. When matrices C_σ and D_σ are also mode-dependent, Theorem 3.6 can also be extended, provided that the pair (G, H) remains common across modes and that (13) and (14) are modified to account for the worse-case bound among all modes.

2) *Switched Plants with Average Dwell-Time Constraints:*

When the switched system (15) does not admit a common Lyapunov function, it is still possible to obtain a result of the form (14), provided the switching is slow “on the average”. In particular, if the switching signal σ satisfies an average dwell-time constrain of the form

$$N_\sigma(t, \tau) \leq \eta(t - \tau) + N_0, \quad (16)$$

where $N_\sigma(t, \tau)$ denotes the number of discontinuities of σ in the open interval (τ, t) , and $N_0 \geq 0$ 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 (42) carries over. This observation follows directly from the Lyapunov construction presented in [13], 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 [13]. However, unlike the results of [13], the results of this paper allow to consider online optimization problems with constraints. To the best of our knowledge, similar results for online optimization with constraints of switched systems have not been studied before.

IV. ONLINE PRIMAL-DUAL GRADIENT FLOW

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

$$(u_t^*, y_t^*) := \arg \min_{\bar{u} \in \mathbb{R}^m, \bar{y} \in \mathbb{R}^p} \phi_t(\bar{u}) + \psi_t(\bar{y}), \quad (17a)$$

$$\text{s.t.} \quad \bar{y} = G\bar{u} + Hw_t, \quad K_t \bar{y} = e_t, \quad (17b)$$

which contains only equality constraints on the system outputs. In contrast with the method proposed in Section III, which guarantees tracking of an approximate optimizer, in this section we will show that, when the optimization problem includes only equality constraints, we can guarantee tracking of the exact optimizer (this behavior is achieved without resorting to a regularized Lagrangian).

A. Controller Synthesis

We begin by imposing the following assumption.

Assumption 5: The columns of $K_t G$ are linearly independent and there exists $\underline{k}, \bar{k} \in \mathbb{R}_{>0}$ such that $\underline{k}I \preceq K_t G G^T K_t^T \preceq \bar{k}I$ for all t . \square

Since problem (17) contains only equality constraints, Assumption 5 is sufficient to guarantee uniqueness of the optimal multipliers [16]. 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.

We consider the following Lagrangian function for problem (17): $\mathcal{L}_t(u, \lambda) = \phi_t(u) + \psi_t(Gu + Hw_t) + \lambda^T (K(Gu + Hw_t) - e_t)$, where $\lambda \in \mathbb{R}_{\geq 0}^n$ is the vector of dual variables. Under Assumptions 2 and 5, the unique minimizer (u_t^*, y_t^*) of (17) solves the following KKT conditions:

$$\begin{aligned} 0 &= \nabla \phi_t(u_t^*) + G^T \nabla \psi_t(Gu_t^* + Hw_t) + G^T K^T \lambda_t^*, \\ 0 &= K(Gu_t^* + Hw_t) - e_t. \end{aligned} \quad (18)$$

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

$$L_{u,t}(u, y, \lambda) := \nabla \phi(u) + G^T \nabla \psi(y) + G^T K^T \lambda, \quad (19a)$$

$$L_{\lambda,t}(y) := Ky - e, \quad (19b)$$

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

$$\varepsilon \dot{x} = Ax + Bu + Ew_t, \quad y = Cx + Dw_t, \quad (20a)$$

$$\dot{u} = -\eta_u L_{u,t}(u, y, \lambda), \quad (20b)$$

$$\dot{\lambda} = \eta_\lambda L_{\lambda,t}(y), \quad (20c)$$

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 (20b)–(20c) uses output-feedback from the plant, and does not require any knowledge on w_t . In the following lemma, we relate the time-varying equilibria of (20) with the solution of (17). To this aim, in what follows we use the notation:

$$z := (u, \lambda), \quad z_t^* := (u_t^*, \lambda_t^*), \quad \tilde{z} := z - z_t^*, \quad (21)$$

to denote the controller state, the saddle-point of $\mathcal{L}_t(u, \lambda)$, and the controller tracking error, respectively. Similarly, we use

$$\xi := (x, u, \lambda), \quad \xi_t^* := (x_t^*, u_t^*, \lambda_t^*), \quad \tilde{\xi} = \xi - \xi_{\nu,t}^*, \quad (22)$$

to denote the joint state of (20), the saddle-point of $\mathcal{L}_t(u, \lambda)$, with $x_t^* = -A^{-1}(Bu_{\nu,t}^* + Hw_t)$, and the joint plant and controller tracking error, respectively.

Lemma 4.1: Let Assumptions 1–5 hold. For any $t \in \mathbb{R}_{\geq 0}$ and for any fixed $w_t \in \mathbb{R}^q$, let $\xi_{\text{eq}} := (x_{\text{eq}}, u_{\text{eq}}, \lambda_{\text{eq}})$ denote an equilibrium of (20) and let ξ_t^* be the unique solution of (18), as defined in (22). Then, ξ_{eq} is unique and $\xi_{\text{eq}} = \xi_t^*$.

The proof of this claim is omitted due to space limitations. Differently from Lemma 3.4, Lemma 4.1 establishes that the equilibrium point of (20) coincides with the saddle point of the exact (non-augmented) Lagrangian function. In what follows, we investigate the transient behavior of the system (20).

B. Stability and Tracking Analysis

We first show that, when the dynamics (1) are infinitely fast, (20) converges exponentially to the solution of (3).

Proposition 4.2: Let Assumptions 1–5 hold, let

$$P_z := \begin{bmatrix} \ell I & G^\top K^\top \\ KG & \ell \frac{\eta_u}{\eta_\lambda} I \end{bmatrix}, \quad (23)$$

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

$$\|\tilde{z}(t)\| \leq \sqrt{\kappa} \|\tilde{z}(t_0)\| e^{-\frac{1}{2} \rho_z (t-t_0)} + \frac{4\|P_z\|\sqrt{\kappa}}{\underline{\lambda}(P_z)} \text{ess sup}_{\tau \geq t_0} \|\dot{z}_\tau^*\|, \quad (24)$$

for all $t \geq t_0$, $\rho_z := \frac{1}{2} \min\{\eta_\lambda k / \ell, \eta_u \frac{\mu}{2}\}$, $\kappa = \bar{\lambda}(P_z) / \underline{\lambda}(P_z)$, where \tilde{z} denotes the controller tracking error as in (21).

The proof of this result is presented in Appendix B. Proposition 4.2 guarantees that (8) is input-to-state stable [14] with respect to \dot{z}_τ^* . Two comments are in order. First, differently from [16, Theorem 1], Proposition 4.2 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 (24) is in the spirit of [20, Section 6]; however, in [20] the primal-dual dynamics are assumed to be differentiable with respect to t (in contrast, we require milder conditions of absolute continuity).

Remark 7: When the matrix K is time-varying, it follows that $P_{z,t}$ in (23) and the coefficient κ_t in (24) are time-varying too. In this case, the result (24) extends by replacing κ with $\sup_\tau \kappa_\tau$ and the coefficient $\frac{4\|P_z\|\sqrt{\kappa}}{\underline{\lambda}(P_z)}$ with $\sup_\tau \frac{4\|P_{z,\tau}\|\sqrt{\kappa_\tau}}{\underline{\lambda}(P_{z,\tau})}$. \square

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 (20).

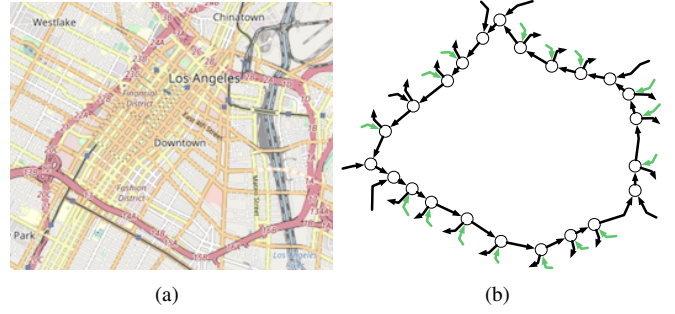


Fig. 3. (a) Portion of highway system in Los Angeles, CA, USA. (b) Network schematic. The network models $|\mathcal{L}| = 64$ traffic highways and links colored in green represent controllable on-ramps.

Theorem 4.3: (Stability and Tracking of (20)) Let Assumptions 1–5 hold and let P_x, Q_x be as in Assumption 1. Suppose that ε satisfies

$$\varepsilon < \frac{\rho_z \underline{\lambda}(P_x) \underline{\lambda}(P_z)}{16\sigma_1 \sigma_2 + 4\rho_z \underline{\lambda}(P_z) \sigma_3}, \quad (25)$$

where P_z, ρ_z are as in Proposition 4.2, and

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

Then, for any $t_0 \in \mathbb{R}_{\geq 0}$, the tracking error (22) satisfies:

$$\begin{aligned} \|\tilde{\xi}(t)\| &\leq \sqrt{\kappa} \|\tilde{\xi}(t_0)\| e^{-\frac{1}{2} \rho_\xi (t-t_0)} + \frac{4\|P_z\|\sqrt{\kappa}}{\rho_z \underline{\lambda}(P_z)} \text{ess sup}_{\tau \geq t_0} \|\dot{z}_\tau^*\| \\ &\quad + \frac{4\|P_x A^{-1} E\|\sqrt{\kappa}}{\underline{\lambda}(Q_x)} \text{ess sup}_{\tau \geq t_0} \|\dot{w}_\tau\|, \end{aligned} \quad (26)$$

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

$$\rho_\xi = \frac{1}{4} \min \left\{ \rho_z \frac{\underline{\lambda}(P_z)}{\bar{\lambda}(P_z)}, \varepsilon^{-1} \frac{\underline{\lambda}(Q_x)}{\bar{\lambda}(P_x)} \right\}. \quad (27)$$

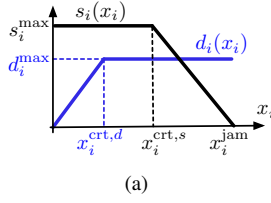
The proof of this result is postponed to Appendix B. Precisely, Theorem 4.3 guarantees that (8) is input-to-state stable [14] with respect to $\dot{z}_{\nu,\tau}^*$ and \dot{w}_τ . The bound on ε is an increasing function of $\underline{\lambda}(P_x)$ and $\rho_z \underline{\lambda}(P_z)$, which are 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 (26) can be readily extended to account for time-varying matrices K_t by adopting a reasoning similar to that in Remark 7.

V. APPLICATION TO RAMP METERING CONTROL

In this section, we apply the proposed framework to the control of on-ramps in a network of traffic highways¹.

To describe the traffic evolution, we adopt a continuous-time version of the Cell-Transmission Model (CTM) [36]. We

¹The code used in our simulations is publicly available at https://github.com/gianlucaBi/onlinePrimalDual_rampMetering.



(a)

Variable	Description	Value	Unit
φ_i	free-flow speed	4	km/min
β_i	back propag. speed	4	km/min
d_i^{\max}	demand saturation	120	veh/min
s_i^{\max}	supply saturation	120	veh/h
x_i^{jam}	jam density	60	veh/km
$x_i^{\text{crt},d}$	critical density of d_i	30	veh/km
$x_i^{\text{crt},s}$	critical density of s_i	30	veh/km
—	avg. numb. of lanes	4	none

(b)

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

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 the set of links into three disjoint sets: $\mathcal{L} = \mathcal{L}_{\text{on}} \cup \mathcal{L}_{\text{off}} \cup \mathcal{L}_{\text{in}}$, where \mathcal{L}_{on} denotes the set of on-ramps where vehicles can enter the network, \mathcal{L}_{off} denotes the set of off-ramps where vehicles can exit the network, and \mathcal{L}_{in} denotes the set of internal links.

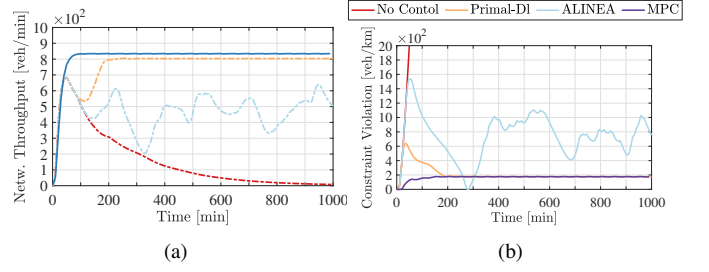
For $i \in \mathcal{L}$, we denote by i^+ the set of downstream links, and by i^- the set of upstream links. For all $i \in \mathcal{L}$, we let $x_i : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ be the density of vehicle in the link. We model the dynamics of all links $i \in \mathcal{L}_{\text{in}}$ according to the CTM with FIFO allocation policy [36]:

$$\begin{aligned}
 \dot{x}_i &= -f_i^{\text{out}}(x) + f_i^{\text{in}}(x), \\
 f_i^{\text{out}}(x) &= \min\{d_i(x_i), \{s_j(x_j)/r_{ij}\}_{j \in i^+}\}, \\
 d_i(x_i) &= \min\{\varphi_i x_i, d_i^{\max}\}, s_i(x_i) = \min\{\beta_i(x_i^{\text{jam}} - x_i), s_i^{\max}\}, \\
 f_i^{\text{in}}(x) &= \sum_{j \in i^-} f_j^{\text{out}}(x),
 \end{aligned} \tag{28}$$

where $d_i : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ and $s_i : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ are the link demand and supply functions, respectively, $r_{ij} \in [0, 1]$ is the routing ratio from i to j , with $\sum_j r_{ij} = 1$, $\varphi_i > 0$. In our simulations, we used identical and uniform routing ratios at each junction. We refer to Fig. 3 for an illustration of the network topology used in our simulations, and to Fig. 4 for a description of the parameters that characterize demand and supply. For simplicity, all links are assumed to be identical. The dynamics of on-ramps and off-ramps coincide with those of (28), where inflow and outflow functions are replaced by:

$$\begin{aligned}
 f_i^{\text{in}}(x) &:= u_i, & \text{if } i \in \mathcal{L}_{\text{on}}, \\
 f_i^{\text{out}}(x) &:= d_i(x_i), & \text{if } i \in \mathcal{L}_{\text{off}},
 \end{aligned} \tag{29}$$

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 $w_i : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$. 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.

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

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 $(u - u^{\text{ref}})^T Q_u (u - u^{\text{ref}}) - \Phi(x)$, subject to the constraints (28)-(29), 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 $i \in \mathcal{L}$, the inequality $d_i^{\max} \leq s_j^{\max}$ holds for all $j \in i^+$. Under this assumption, if the network is operated in a regime in which $x_i \leq \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}$ for all $i \in \mathcal{L}$ (i.e., all highways operate in the free-flow regime), then the dynamics (28) simplify to the following linear model:

$$\begin{aligned}
 \dot{x}_i &= -f_i^{\text{out}}(x) + f_i^{\text{in}}(x), \\
 f_i^{\text{out}}(x) &= \varphi_i x_i, \quad f_i^{\text{in}}(x) = \sum_{j \in i^-} f_j^{\text{out}}(x).
 \end{aligned} \tag{30}$$

In vector form, (30) can be written as $\dot{x} = (R^T - I)Fx + Bu$, and $y = x + w$, where $R := [r_{ij}]$, and $F := \text{diag}(\varphi_1, \dots, \varphi_n)$. Notice that matrix $(R^T - I)F$ is Hurwitz (see e.g. [3, Theorem 1]). Building on this, we propose the following problem:

$$\begin{aligned}
 \min_{u, y} \quad & (u - u^{\text{ref}})^T Q_u (u - u^{\text{ref}}) - \Phi(y), \\
 \text{s.t.} \quad & y = -((R^T - I)F)^{-1}Bu + w, \\
 & u_i \geq 0, \quad y_i \leq \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}, \forall i \in \mathcal{L}.
 \end{aligned} \tag{31}$$

The optimization problem (31) formalizes the objectives of the ramp metering problem, while guaranteeing that all highways are operated in the free-flow regime.

2) *Distributed Reactive Metering using ALINEA*: ALINEA [26] 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 $i \in \mathcal{L}_{\text{in}}$, 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 the setpoint be $\hat{x}_i = \min\{x_i^{\text{crt},d}, x_i^{\text{crt},s}\}$.

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

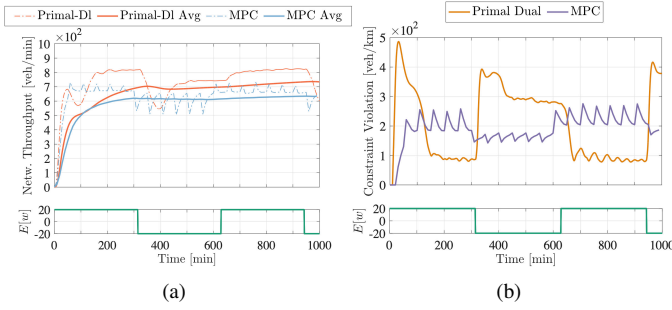


Fig. 6. Plant subject to random noise (green line shows noise mean). (a) Throughput $\Phi(x)$. (b) Constraint violation: $\|y - \min\{x^{\text{cr},d}, x^{\text{cr},s}\}\|$.

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 the MPC optimization problem by discretizing the dynamics and we used $T_p = 200$ min and $T_s = 50$ min.

Discussion: Fig. 5 compares the performance of the three controllers in the noiseless case (i.e., where $w_i = 0$ 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 (31). Finally, although ALINEA largely outperforms absence of on-ramp metering control, it critically suffers from its distributed architecture, making it suboptimal.

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) because the primal-dual controller uses instantaneous feedback from the system, it can react faster to unmodeled dynamics or time-varying disturbances, and (ii) in contrast with MPC where an optimization problem is solved at the beginning of every time-window $[0, T_s]$, the primal-dual controller is governed by an ODE that continuously updates the variables at runtime, thus making the approach computationally more efficient.

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.

REFERENCES

- [1] 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.
- [2] 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*, 2018, pp. 136–143.
- [3] 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.
- [4] R. Kutadinata, W. Moase, C. Manzie, L. Zhang, and T. Geroni, "Enhancing the performance of existing urban traffic light control through extremum-seeking," *Transp. Research Pt. C: Emerging Technologies*, vol. 62, pp. 1–20, 2016.
- [5] S. H. Low and D. E. Lapsley, "Optimization flow control-I: Basic algorithm and convergence," *IEEE/ACM Trans. on networking*, vol. 7, no. 6, pp. 861–874, 1999.
- [6] A. Jokic, M. Lazar, and P. P.-J. Van Den Bosch, "On constrained steady-state regulation: Dynamic KKT controllers," *IEEE Trans. on Automatic Control*, vol. 54, no. 9, pp. 2250–2254, 2009.
- [7] 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.
- [8] 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.
- [9] L. S. P. Lawrence, J. W. Simpson-Porco, and E. Mallada, "Linear-convex optimal steady-state control," *arXiv preprint*, p. arXiv:1810.12892, 2018.
- [10] T. Zheng, J. Simpson-Porco, and E. Mallada, "Implicit trajectory planning for feedback linearizable systems: A time-varying optimization approach," in *American Control Conference*, 2020, pp. 4677–4682.
- [11] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, "Timescale separation in autonomous optimization," *IEEE Trans. on Automatic Control*, 2020, (In press).
- [12] 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.
- [13] 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.
- [14] E. D. Sontag and Y. Wang, "On characterizations of the input-to-state stability property," *Systems & Control Letters*, vol. 24, no. 5, pp. 351–359, 1995.
- [15] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.
- [16] 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.
- [17] 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.
- [18] J. Cortés and S. K. Niederländer, "Distributed coordination for non-smooth convex optimization via saddle-point dynamics," *Journal of Nonlinear Science*, vol. 29, no. 4, pp. 1247–1272, 2019.
- [19] 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.
- [20] 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.
- [21] R. Goebel, "Stability and robustness for saddle-point dynamics through monotone mappings," *Systems and Control Letters*, vol. 108, pp. 16–22, 2017.
- [22] J. I. Poveda and N. Li, "Robust hybrid zero-order optimization algorithms with acceleration via averaging in time," *Automatica*, vol. 123, p. 109361, 2021.
- [23] 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, 2020.
- [24] 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.
- [25] 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.

- [26] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 271–281, 2002.
- [27] 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.
- [28] A. Nagurney and D. Zhang, *Projected dynamical systems and variational inequalities with applications*. Springer, 2012, vol. 2.
- [29] X.-B. Gao, "Exponential stability of globally projected dynamic systems," *IEEE Trans. on Neural Networks*, vol. 14, no. 2, pp. 426–431, 2003.
- [30] 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.
- [31] G. Grubb, *Distributions and operators*. Springer Science & Business Media, 2008, vol. 252.
- [32] A. L. Dontchev and T. R. Rockafellar, *Implicit functions and solution mappings*. Springer, 2009, 2nd ed.
- [33] E. DiBenedetto, *Real analysis*. Springer Science & Business Media, 2002.
- [34] 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.
- [35] 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.
- [36] C. F. Daganzo, "The cell transmission model pt. II: network traffic," *Transp. Research Pt. B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995.
- [37] 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.

APPENDIX A

ANALYSIS OF PROJECTED SADDLE-POINT CONTROLLER

In this section, we present the proof of Proposition 3.5 and Theorem 3.6. For the subsequent analysis, it is convenient to define the following time-varying map:

$$F_t(z) := \begin{bmatrix} \nabla\phi_t(u) + G^T \nabla\psi_t(Gu + Hw_t) + G^T K_t^T \lambda \\ -(K_t(Gu + Hw_t) - e - \nu\lambda) \end{bmatrix}. \quad (32)$$

1) *Proof of Proposition 3.5:* Recall that $z := (u, \lambda)$. We note that, when $\varepsilon = 0$, the dynamics (8) can be rewritten as:

$$\dot{z} = P_\Omega(z - \eta F_t(z)) - z, \quad (33)$$

where $\Omega := \mathcal{U} \times \mathcal{C}$. Proposition 3.5 leverages this structure as well as four auxiliary lemmas. The following lemma follows directly from [37, Lemma 6] and [18].

Lemma A.1: Let Assumption 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 $\mu_u I \preceq T_{u,t} \preceq \ell_u I$ and $0 \preceq T_{u,t} \preceq \ell_y I$, such that $\nabla\phi_t(u) - \nabla\phi_t(u') = T_{u,t}(u - u')$ and $\nabla\psi_t(y) - \nabla\psi_t(y') = T_{y,t}(y - y')$.

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 A.2: Let Assumption 2 hold. Then, (32) satisfies:

$$(z - z')^T (F_t(z) - F_t(z')) \geq \min\{\mu_u, \nu\} \|z - z'\|^2, \quad (34)$$

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 (34), and by using Lemma A.1:

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

which proves the claim. \blacksquare

Lemma A.3: Let Assumptions 2 and 3 hold. Then, the mapping (32) satisfies:

$$\|F_t(z) - F_t(z')\| \leq \ell \|z - z'\|, \quad (35)$$

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

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

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

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

The following result establishes that the existence of an ISS-Lyapunov function with a particular structure guarantees input-to-state stability with exponential convergence rate, and it is a particular case of [15, Ch. 4] (see also [27]).

Lemma A.4: Consider the system $\dot{x} = f(t, x, u)$, where $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is locally Lipschitz in t, x , and u , and $t \mapsto u(t)$ is measurable and essentially bounded. If there exists a continuously differentiable $V : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}$ s.t.:

$$\underline{a} \|x\|^2 \leq V(t, x) \leq \bar{a} \|x\|^2, \quad (36a)$$

$$\frac{d}{dt} V(t, x) \leq -b V(t, x), \quad \forall \|x\| \geq b_0 > 0, \quad (36b)$$

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

$$\|x(t)\| \leq \sqrt{\bar{a}/\underline{a}} (\|x(t_0)\| e^{-\frac{1}{2}b(t-t_0)} + b_0), \quad \forall t \geq 0. \quad (37)$$

Using the results above, we now present the proof of Proposition 3.5. In particular, we show that the function $V(\tilde{z}_\nu) = \frac{1}{2} \|\tilde{z}_\nu\|^2$ satisfies the assumptions of Lemma A.4, where we recall that \tilde{z}_ν is as in (10). In what follows, we let $\hat{z} := P_\Omega(z - \eta F_t(z))$. By expanding the time-derivative:

$$\frac{d}{dt} V(\tilde{z}_\nu) = -\tilde{z}_\nu^T (z - \hat{z}) - \tilde{z}_\nu^T \dot{z}_{\nu,t}^*, \quad (38)$$

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:

$$(v' - P_\Omega(v))^T (P_\Omega(v) - v) \geq 0, \quad \text{for all } v' \in \Omega. \quad (39)$$

By using (39) with $v' = z_{\nu,t}^*$ and $v = z - \eta F_t(z)$, we have $(\tilde{z}_\nu + \eta F_t(z))^\top (z - \hat{z}) \geq \|z - \hat{z}\|^2 + \eta(z - z_{\nu,t}^*)^\top F_t(z)$, and thus the first term in (38) satisfies:

$$\begin{aligned} -\tilde{z}_\nu^\top (z - \hat{z}) &\leq -\|z - \hat{z}\|^2 - \eta(\hat{z} - z^*)^\top F_t(z) \\ &\leq -\|z - \hat{z}\|^2 - \eta(\hat{z} - z_{\nu,t}^*)^\top (F_t(z) - F_t(z_{\nu,t}^*)) \\ &\quad - \eta(\hat{z} - z_{\nu,t}^*)^\top F_t(z_{\nu,t}^*) \\ &= -\|z - \hat{z}\|^2 - \eta\tilde{z}_\nu^\top (F_t(z) - F_t(z_{\nu,t}^*)) \\ &\quad + \eta(z - \hat{z})^\top (F_t(z) - F_t(z_{\nu,t}^*)) \\ &\leq -\|z - \hat{z}\|^2 + \eta\ell\|z - \hat{z}\|\|\tilde{z}_\nu\| - \eta\mu\|\tilde{z}_\nu\|^2 \\ &\leq -\eta(\mu - \eta\ell^2/4)\|\tilde{z}_\nu\|^2, \end{aligned} \quad (40)$$

where the second inequality follows by adding and subtracting $\eta(\hat{z} - z_{\nu,t}^*)^\top F_t(z_{\nu,t}^*)$, the third inequality follows by expanding $\hat{z} - z_{\nu,t}^* = (z - z_{\nu,t}^*) - (z - \tilde{z}_\nu)$ and by using $(\hat{z}_\nu - z_{\nu,t}^*)^\top F_t(z) \geq 0$, the fourth inequality follows from Lemmas A.2 and A.3, 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 (38) we obtain:

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

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

2) *Proof of Theorem 3.6:* Our proof technique leverages singular perturbation arguments inspired by [15, Ch. 11]. 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:

$$\begin{aligned} \varepsilon \dot{\tilde{x}} &= A\tilde{x} + \varepsilon A^{-1}BS\dot{z} + A^{-1}E\dot{w}_t \\ \dot{z} &= P_\Omega(z - \alpha F_t(z, \tilde{x})) - z, \end{aligned} \quad (41)$$

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 the theorem's statement, we will show that

$$U(\tilde{z}_\nu, \tilde{x}) := (1 - \theta)V(\tilde{z}_\nu) + \theta W(z), \quad (42)$$

where $V(\tilde{z}_\nu) = \frac{1}{2}\|z(t) - z_{\nu,t}^*\|^2$, $W(z) = \tilde{x}^\top P_x \tilde{x}$, and $\theta = b/(b + g)$ satisfies the assumptions of Lemma A.4. 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 (41) reads:

$$\frac{d}{dt}V(\tilde{z}_\nu) = \tilde{z}_\nu^\top (\hat{z} - z) - \tilde{z}_\nu^\top \dot{z}_{\nu,t}^* \quad (43)$$

almost everywhere. The first term satisfies:

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

where the first inequality follows from the non-expansiveness of the projection operator, namely:

$$\begin{aligned} \tilde{z}_\nu^\top (P_\Omega(z - \eta F_t(z, \tilde{x})) - P_\Omega(z - \eta F_t(z, 0))) \\ \leq \|\tilde{z}_\nu\|\|P_\Omega(z - \eta F_t(z, \tilde{x})) - P_\Omega(z - \eta F_t(z, 0))\| \\ \leq \|\tilde{z}_\nu\|\|(z - \eta F_t(z, \tilde{x})) - (z - \eta F_t(z, 0))\|, \end{aligned}$$

and the second inequality follows from (40). Moreover, by expanding the terms:

$$\begin{aligned} \|F_t(z, \tilde{x}) - F_t(z, 0)\| \\ \leq \left\| \begin{bmatrix} G^\top (\nabla f_y(C\tilde{x} + Gu + Hw_t) - \nabla f_y(Gu + Hw_t)) \\ -K_t C\tilde{x} \end{bmatrix} \right\| \\ \leq \|C\|(\ell_y\|G\| + \bar{K})\|\tilde{x}\|. \end{aligned}$$

Hence, by recalling the definition of b and ρ_z , (43) satisfies:

$$\begin{aligned} \frac{d}{dt}V(\tilde{z}_\nu) &\leq -\rho_z\|\tilde{z}_\nu\|^2 + b\|\tilde{x}\|\|\tilde{z}_\nu\| + \|\tilde{z}_\nu\|\|\dot{z}_{\nu,t}^*\| \\ &\leq -\frac{\rho_z}{2}\|\tilde{z}_\nu\|^2 + b\|\tilde{x}\|\|\tilde{z}_\nu\|, \end{aligned} \quad (44)$$

where the last inequality holds when $\|\tilde{z}_\nu\| \geq \frac{2}{\rho_z} \text{ess sup} \|\dot{z}_{\nu,t}^*\|$. The time-derivative of $W(\tilde{x})$ along the trajectories of (41):

$$\begin{aligned} \frac{d}{dt}W(z) &= \varepsilon^{-1}\tilde{x}^\top (A^\top P_x + P_x A)\tilde{x} \\ &\quad + 2\tilde{x}^\top P_x A^{-1}BS\dot{z} + 2\tilde{x}^\top P_x A^{-1}E\dot{w}_t \\ &\leq -\varepsilon^{-1}\underline{\lambda}(Q_x)\|\tilde{x}\|^2 + 2\|P_x A^{-1}B\|\|\tilde{x}\|\|S\dot{z}\| \\ &\quad + 2\|P_x A^{-1}B\|\|\tilde{x}\|\|\dot{w}_t\|. \end{aligned} \quad (45)$$

By expanding the terms:

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

where the first inequality follows from the non-expansiveness of the projection operator and the second inequality follows from Assumption 2. By recalling the definition of g , letting $d = 2\eta\ell_y\|PA^{-1}B\|\|C\|\|G\|$, by substituting into (45):

$$\begin{aligned} \frac{d}{dt}W(\tilde{x}) &\leq -\varepsilon^{-1}\underline{\lambda}(Q_x)\|\tilde{x}\|^2 + d\|\tilde{x}\|^2 \\ &\quad + g\|\tilde{x}\|\|\tilde{z}_\nu\| + 2\|P_x A^{-1}B\|\|\tilde{x}\|\|\dot{w}_t\| \\ &\leq -\frac{\underline{\lambda}(Q_x)}{2\varepsilon}\|\tilde{x}\|^2 + d\|\tilde{x}\|^2 + g\|\tilde{x}\|\|\tilde{z}_\nu\|, \end{aligned} \quad (46)$$

where the last inequality is satisfied if $\|\tilde{x}\| \geq \frac{4\varepsilon\|P_x A^{-1}E\|}{\underline{\lambda}(Q_x)} \text{ess sup } \|\dot{w}_t\|$. By combining (44)-(46):

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

where

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

Λ is positive definite when $\theta(1-\theta)\frac{\rho_z\lambda(Q_x)}{16\varepsilon} > \frac{1}{4}((1-\theta)b + \theta g)^2$, which holds when (13) is satisfied. Finally, the claim follows by application of Lemma A.4 with $\bar{a} = \max\{\frac{1}{2}, \bar{\lambda}(P_x)\}$, $\underline{a} = \min\{\frac{1}{2}, \underline{\lambda}(P_x)\}$, $c_3 = \frac{1}{2} \min\{2\rho_z, \frac{\underline{\lambda}(Q_x)}{4\varepsilon\lambda(P_x)}\}$, and $b_0 = \max\{\frac{2}{\rho_z} \text{ess sup } \|\dot{z}_{\nu,t}^*\|, \frac{4\varepsilon\|P_x A^{-1}E\|}{\underline{\lambda}(Q_x)} \text{ess sup } \|\dot{w}_t\|\}$. ■

APPENDIX B

ANALYSIS OF PRIMAL-DUAL CONTROLLER

In this section, we prove Proposition 4.2 and Theorem 4.3. We introduce the following change of variables for (20):

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

The dynamics (20) are re-written in the new variables next.

Lemma B.1: Let Assumption 1-5 be satisfied, and for any $t \in \mathbb{R}_{\geq 0}$, let (u_t^*, λ_t^*) be the saddle-point of (17). The dynamics (20) have the following equivalent representation:

$$\begin{aligned} \dot{\tilde{x}} &= F_{11}\tilde{x} + F_{12}(u - u_t^*) + F_{13}(\lambda - \lambda_t^*) + F_{14}\dot{w}_t, \\ \dot{u} &= F_{21}\tilde{x} + F_{22}(u - u_t^*) + F_{23}(\lambda - \lambda_t^*), \\ \dot{\lambda} &= F_{31}\tilde{x} + F_{32}(u - u_t^*), \end{aligned} \quad (47)$$

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

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

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 A.1:

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

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} :

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

which proves the claim. ■

1) Proof of Proposition 4.2: The proof follows similar ideas as in [16, Lemma 2]. By letting $\varepsilon = 0$ in (47) we obtain $A\tilde{x} = 0$, which, by Assumption 1 implies $\tilde{x} = 0$. Hence, we let $z := (u, \lambda)$ and $\tilde{z} := z - z_t^*$, and we rewrite the dynamics (47) as $\dot{z} = F_z(z - z_t^*) = F_z\tilde{z}$, where

$$F_z = \begin{bmatrix} F_{22} & F_{23} \\ F_{32} & 0 \end{bmatrix}. \quad (48)$$

We will prove that $V(z) = \tilde{z}^\top P_z \tilde{z}$ satisfies the assumptions of Lemma A.4. By the Schur Complement, P_z is positive definite if and only if $\ell^2 \frac{\eta_u}{\eta_\lambda} I - G^\top K^\top K G \succ 0$. Using $\eta_u > \frac{4k}{\ell\mu} \eta_\lambda$, $\ell \geq \mu$ and Assumption 5 one gets $\ell^2(\eta_u/\eta_\lambda)I - G^\top K^\top K G \succeq ((4\ell k)/\mu_u)I - \bar{k}I \succeq 3\bar{k} \succ 0$, which shows that P_z is positive definite. By expanding the time-derivative:

$$\begin{aligned} \frac{d}{dt}V(\tilde{z}) &= (\dot{z} - \dot{z}_t^*)^\top P_z(z - z_t^*) + (z - z_t^*)^\top P_z(\dot{z} - \dot{z}_t^*) \\ &= \tilde{z}^\top (F_z^\top P_z + P_z F_z)\tilde{z} - 2\tilde{z}^\top P_z \dot{z}_t^*. \end{aligned} \quad (49)$$

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

$$\begin{aligned} M_{11} &= 2\eta_u \ell(T_{u,t} + G^\top T_{y,t}G) - 2\eta_\lambda G^\top K^\top K G - \bar{\rho}_z \ell I, \\ M_{12} &= \eta_u(T_{u,t} + G^\top T_{y,t}G)^\top G^\top K^\top - \bar{\rho}_z G^\top K^\top, \\ M_{22} &= 2\eta_u K G G^\top K^\top - \bar{\rho}_z \ell(\eta_u/\eta_\lambda)I, \end{aligned} \quad (50)$$

and $M_{21} = M_{12}^\top$. By application of the Schur Complement, M is positive definite when $M_{22} \succ 0$ and $M_{11} - M_{12}M_{22}^{-1}M_{12}^\top \succ 0$. The first condition can be rewritten as: $M_{22} \succeq (2\eta_u \underline{\ell} - \bar{\rho}_z \ell \frac{\eta_u}{\eta_\lambda})I \succeq \eta_u \underline{\ell} I \succ 0$, where we used Assumption 5 and the expression of ρ_z . For the second condition, we have:

$$\begin{aligned} M_{11}M_{22}^{-1}M_{12}^\top &\preceq M_{12}(\eta_u K G G^\top K^\top)^{-1}M_{12}^\top \\ &= \eta_u(T_{u,t} + G^\top T_{y,t}G)^\top (T_{u,t} + G^\top T_{y,t}G) + \frac{\bar{\rho}_z^2}{\eta_u} I \\ &\quad - \bar{\rho}_z((T_{u,t} + G^\top T_{y,t}G)^\top + (T_{u,t} + G^\top T_{y,t}G)) \\ &\preceq \eta_u \ell(T_{u,t} + G^\top T_{y,t}G) + \frac{\bar{\rho}_z^2}{\eta_u} I - 2\bar{\rho}_z(T_{u,t} + G^\top T_{y,t}G), \end{aligned}$$

where the first bound follows from Assumption 5 and the definition of $\bar{\rho}_z$, the second identity follows from $G^\top K^\top (K G G^\top K^\top)^{-1} K G = I$, and the last bound follows from $G^\top T_{y,t}G \succeq 0$. Thus:

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

and, by using

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

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

As a result, (49) satisfies:

$$\begin{aligned} \frac{d}{dt} V(\tilde{z}) &\leq -\bar{\rho}_z V(\tilde{z}) + 2\|\tilde{z}\| \|P_z\| \|\dot{z}_t^*\| \\ &= -\frac{\bar{\rho}_z}{2} V(\tilde{z}) - \frac{\rho_z}{2} \underline{\lambda}(P_z) \|\tilde{z}\|^2 + 2\|\tilde{z}\| \|P_z\| \|\dot{z}_t^*\| \\ &\leq -\frac{\bar{\rho}_z}{2} V(\tilde{z}), \end{aligned} \quad (51)$$

where the last inequality holds when $2\|\tilde{z}\| \|P_z\| \|\dot{z}_t^*\| - \frac{\rho_z}{2} \underline{\lambda}(P_z) \|\tilde{z}\|^2 \leq 0$, or $\|\tilde{z}\| \geq \frac{4\|P_z\|}{\rho_z \underline{\lambda}(P_z)} \text{ess sup}_\tau \|\dot{z}_\tau^*\|$. Finally, the claim follows by application of Lemma A.4 with $\bar{a} = \bar{\lambda}(P_z)$, $\underline{a} = \underline{\lambda}(P_z)$, $b = \frac{\bar{\rho}_z}{2}$, and $b_0 = \frac{4\|P_z\|}{\rho_z \underline{\lambda}(P_z)} \text{ess sup}_\tau \|\dot{z}_\tau^*\|$. ■

2) *Proof of Theorem 4.3:* Our proof technique leverages singular perturbation arguments inspired by [15, Ch. 11]. Let $z := (u, \lambda)$, $\tilde{z} := z - z_t^*$ and rewrite the dynamics (47) as:

$$\dot{\tilde{z}} = F_{11}\tilde{x} + F_{xz}\tilde{z} + F_{14}\dot{w}_t, \quad \dot{\tilde{z}} = F_{zx}\tilde{x} + F_z\tilde{z}, \quad (52)$$

where F_z is as defined by (48), $F_{xz} = [F_{12}, F_{13}]$, and $F_{zx} = [F_{21}^\top, F_{31}^\top]^\top$. 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 A.4. By substituting (52) and by using $F_z^\top P_z + P_z F_z \preceq -\bar{\rho}_z P_z$ (see (49) and (51)):

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

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

$$\varepsilon \dot{W}(\tilde{x}) = \tilde{x}^\top (F_{11}^\top P_x + P_x F_{11}) \tilde{x} + 2\tilde{x}^\top P_x F_{xz} \tilde{z} + 2\tilde{x}^\top P_x F_{14} \dot{w}_t.$$

Using $F_{11} = A - \varepsilon \eta_u A^{-1} B G^\top T_{y,t} C$, $A^\top P_x + P_x A = -Q_x$:

$$\begin{aligned} \tilde{x}^\top (F_{11}^\top P_x + P_x F_{11}) \tilde{x} &= -\tilde{x}^\top Q_x \tilde{x} \\ &\quad - \eta_u \varepsilon \tilde{x}^\top (C^\top T_{y,t} G B^\top A^{-\top} P_x + P_x A^{-1} B G^\top T_{y,t} C) \tilde{x}. \end{aligned}$$

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

$$\begin{aligned} \varepsilon \dot{W}(\tilde{x}) &\leq -\underline{\lambda}(Q_x) \|\tilde{x}\|^2 + \varepsilon \|\Sigma_2\| \|\tilde{x}\| \|\tilde{z}\| \\ &\quad + \varepsilon \|\Sigma_3\| \|\tilde{x}\|^2 + 2\|\Sigma_4\| \|\tilde{x}\| \|\dot{w}_t\| \\ &\leq -\frac{\underline{\lambda}(Q_x)}{2} \|\tilde{x}\|^2 + \varepsilon \|\Sigma_2\| \|\tilde{x}\| \|\tilde{z}\| + \varepsilon \|\Sigma_3\| \|\tilde{x}\|^2, \end{aligned} \quad (53)$$

where the last inequality holds if $-\frac{\underline{\lambda}(Q_x)}{2} \|\tilde{x}\|^2 + 2\|\Sigma_4\| \|\tilde{x}\| \|\dot{w}_t\| \leq 0$, or $\|\tilde{x}\| \geq \frac{4\|\Sigma_4\|}{\underline{\lambda}(Q_x)} \text{ess sup}_{\tau \geq 0} \|\dot{w}_\tau\|$. By using $V(z) \leq \bar{\lambda}(P_z) \|\tilde{z}\|^2$, $W(z) \leq \bar{\lambda}(P_x) \|\tilde{x}\|^2$, by letting $\hat{\xi} := (\|\tilde{z}\|, \|\tilde{x}\|)$, and by combining (51)-(53) we get $\dot{U}(\tilde{x}, \tilde{z}) \leq -\hat{\xi}^\top \Lambda \hat{\xi} - \rho_\varepsilon U(\tilde{x}, \tilde{z})$, 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

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

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 (25) is then obtained using standard manipulations. Finally, the claim follows by application of Lemma A.4 with $\bar{a} = \max\{\bar{\lambda}(P_x), \bar{\lambda}(P_z)\}$, $\underline{a} = \min\{\underline{\lambda}(P_x), \underline{\lambda}(P_z)\}$, $b = \frac{1}{4} \min\left\{\rho_z \frac{\underline{\lambda}(P_z)}{\underline{\lambda}(P_x)}, \varepsilon^{-1} \frac{\underline{\lambda}(Q_x)}{\underline{\lambda}(P_x)}\right\}$, and $b_0 = \max\left\{\frac{4\|P_z\|}{\rho_z \underline{\lambda}(P_z)} \text{ess sup}_\tau \|\dot{z}_\tau^*\|, \frac{4\|\Sigma_4\|}{\underline{\lambda}(Q_x)} \text{ess sup}_\tau \|\dot{w}_\tau\|\right\}$. ■