

High-Performance Optimal Incentive-Seeking in Transactive Control for Traffic Congestion[★]

Daniel E. Ochoa^{*}, Jorge I. Poveda

^a*Department of Electrical, Energy and Computer Engineering, University of Colorado Boulder, Boulder, 80305, Colorado, USA*

Abstract

Traffic congestion has dire economic and social impacts in modern metropolitan areas. To address this problem, in this paper we introduce a novel type of model-free transactive controllers to manage vehicle traffic in highway networks for which precise mathematical models are not available. Specifically, we consider a highway system with managed lanes on which dynamic tolling mechanisms can be implemented in real-time using measurements from the roads. We present three incentive-seeking feedback controllers able to find in real-time the optimal economic incentives (e.g., tolls) that persuade highway users to follow a suitable driving behavior that minimizes a predefined performance index. The controllers are agnostic with respect to the exact model of the highway, and they are also able to guarantee fast convergence to the optimal tolls by leveraging non-smooth and hybrid dynamic mechanisms that combine continuous-time dynamics and discrete-time dynamics. We provide numerical examples to illustrate the advantages of the different presented techniques.

Keywords: Transactive Control, Dynamic pricing, Urban Mobility

1. INTRODUCTION

The increase of population density in urban and sub-urban areas has triggered a significant growth of traffic congestion throughout the world, greatly affecting the commute of the public, as well as the expedited delivery of goods. For example, only in 2019, and solely in New York City, the economic losses induced by congestion climbed to \$11 billion USD. Moreover, commute times have significantly increased during the last years, forcing drivers to spend, on average, 41 hours per year in congested traffic during morning (6 am to 9 am) and afternoon (3 pm to 6pm) peak travel times [1]. This problem is only expected to worsen during the next years, to the point that by the end of 2022 traffic congestion will cost \$74 billion USD to the economy of the United States. To tackle this challenge, cities throughout the world are developing and implementing automated control and optimization algorithms that can guarantee an optimal operation of the transportation infrastructure at all times. Examples include smart traffic light systems [2], dynamic pricing [3, 4], ride-sharing services [5], etc. Among these mechanisms, dynamic pricing has emerged as a promising technology to minimize congestion in dense cities such as London [6, 7], Milan [8], and New York [9]. The goal of dynamic pricing is to induce “optimal” traffic flows that optimize a particular performance measure in the network by adaptively adjusting tolls or incentives [10] based on the current state of

the roads. To guarantee that the transportation system continuously operates at its optimal point, pricing algorithms must react *quickly* to changes in the traffic demand, weather conditions, road accidents, etc. This adaptability requirement has motivated the development of different recursive algorithms for optimal tolling computation, e.g., [11, 12, 13, 14]. Nevertheless, most existing pricing approaches are implemented based on (quasi) static lookup tables instead of real-time feedback traffic measurements, and therefore, under the presence of unexpected accidents or events in the system, are susceptible to generate sub-optimal or even “perverse” tolls that could exacerbate the very problems they were intended to solve [3]. Other recent approaches have relied on socio-technical models that aim to capture decision-making maps of drivers from recorded data; see [15]. In [16], [17] and [18], the authors studied PID controllers to manage the operation of lanes in highway systems. In [19] Hamilton-Jacobi-Bellman equations were solved for the optimal control of high-occupancy toll lanes, and adaptive algorithms based on linear parametrizations and welfare gradient dynamics were studied in [20]. A class of model-based saddle-flow dynamics were also recently studied in [21] in the context of ramp metering control. For a recent review of transactive control for dynamic pricing see [22].

In this paper, we depart from the traditional model-based approaches studied in the setting of transactive control, and instead, we introduce a new class of *model-free* optimal incentive seeking controllers that can *rapidly* learn optimal incentives (e.g., tolls) using only output measurements from the transportation systems, guaranteeing closed-loop stability at all times. Specifically, motivated by recent advances in non-smooth and hybrid extremum seeking control [23, 24], we intro-

[★]Research supported in part by NSF grant number CNS-1947613.

^{*}Corresponding Author.

Email addresses: daniel.ochoa@colorado.edu (Daniel E. Ochoa), jorge.poveda@colorado.edu (Jorge I. Poveda)

duce three incentive-seeking controllers (ISC) for model-free optimal price seeking in dynamic pricing: a smooth ISC that emulates the performance of a gradient-flow in the slowest time scale; a non-smooth ISC that emulates the behavior of fixed-time gradient flows in the slowest time scale; and a hybrid ISC that leverages momentum to improve transient performance in the slowest time scale. Each of the three controllers are interconnected with the dynamics of the highway, which incorporate socio-technical dynamics as well as traffic flows. Even though the controllers are agnostic to the traffic model, we establish practical asymptotic stability results for the resulting closed-loop system, and we numerically show that the non-smooth and hybrid ISCs can significantly outperform the smooth ISC in terms of transient performance under enough time scale separation in the closed-loop system.

The rest of this paper is organized as follows: Section II presents preliminaries. Section III introduces the type of models that we consider in the paper. Section IV presents the proposed ISCs, as well as their stability results. Section V presents numerical examples and comparisons between all the controllers, and finally, Section VI ends with the conclusion.

2. PRELIMINARIES

Notation: Given a compact set $\mathcal{A} \subset \mathbb{R}^n$ and a vector $z \in \mathbb{R}^n$, we use $|z|_{\mathcal{A}} := \min_{s \in \mathcal{A}} \|z - s\|_2$ to denote the minimum distance of z to $\mathcal{A}$. We use $\mathbb{S}^1 := \{z \in \mathbb{R}^2 : z_1^2 + z_2^2 = 1\}$ to denote the unit circle in $\mathbb{R}^2$, and $\mathbb{T}^n$ to denote the n^{th} Cartesian product of $\mathbb{S}^1$. We also use $r\mathbb{B}$ to denote a closed ball in the Euclidean space, of radius $r > 0$, and centered at the origin. We use $I_n \in \mathbb{R}^{n \times n}$ to denote the identity matrix, and (x, y) for the concatenation of the vectors x and y , i.e., $(x, y) := [x^\top, y^\top]^\top$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is said to be of class $\mathcal{KL}$ if it is non-decreasing in its first argument, non-increasing in its second argument, $\lim_{r \rightarrow 0^+} \beta(r, s) = 0$ for each $s \in \mathbb{R}_{\geq 0}$, and $\lim_{s \rightarrow \infty} \beta(r, s) = 0$ for each $r \in \mathbb{R}_{\geq 0}$.

Hybrid Dynamical Systems: In this paper, we will model our algorithms as Hybrid Dynamical Systems (HDS) with state $x \in \mathbb{R}^n$, and dynamics of the form:

$$x \in C, \dot{x} = F(x), \quad x \in D, x^+ = G(x), \quad (1a)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called the flow map, and $G : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is called the jump map. The sets C and D , called the flow set and the jump set, define the points in $\mathbb{R}^n$ where the system can flow or jump according to F or G , respectively. Thus, the HDS can be represented by the notation $\mathcal{H} = (C, F, D, G)$. Solutions $x : \text{dom}(x) \rightarrow \mathbb{R}^n$ to (1) are parameterized by a continuous-time index $t \in \mathbb{R}_{\geq 0}$, which increases continuously during flows, and a discrete-time index $j \in \mathbb{Z}_{\geq 0}$, which increases by one during jumps. For a precise definition of hybrid time domains and solutions to (1) we refer the reader to [25, Ch.2].

Stability Notions: The following definitions will be instrumental to characterize the convergence and stability properties of systems of the form (1).

Definition 1. The compact set $\mathcal{A} \subset C \cup D$ is said to be uniformly asymptotically stable (UAS) for system (1) if $\exists \beta \in \mathcal{KL}$

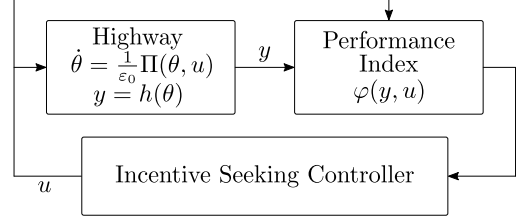


Figure 1: A closed-loop interconnection between an Incentive Seeking Controller (ISC) and a highway system. The ISC will be designed to minimize in real time a performance function defined by an external supervisor.

and $r > 0$, such that for all solutions x with $x(0, 0) \in r\mathbb{B}$, and every $(t, j) \in \text{dom}(x)$, the following bound holds: $|x(t, j)|_{\mathcal{A}} \leq \beta(|x(0, 0)|_{\mathcal{A}}, t + j)$, $\forall (t, j) \in \text{dom}(x)$. $\square$

We will also consider ε -parameterized HDS of the form: $x \in C_\varepsilon, \dot{x} = F_\varepsilon(x), x \in D_\varepsilon, x^+ \in G_\varepsilon(x)$, where $\varepsilon > 0$. For these perturbed hybrid systems $\mathcal{H}_\varepsilon$ we will study *practical stability* properties as $\varepsilon \rightarrow 0^+$.

Definition 2. The compact set $\mathcal{A} \subset C \cup D$ is said to be practically Asymptotically Stable (PAS) as $\varepsilon \rightarrow 0^+$ for $\mathcal{H}_\varepsilon$ if $\exists \beta \in \mathcal{KL}$ and $r > 0$ such that for each pair $(\tilde{r}, v)$ satisfying $r > \tilde{r} > v > 0$, there exists $\varepsilon^* > 0$ such that for all $\varepsilon \in (0, \varepsilon^*)$ every solution of $\mathcal{H}_\varepsilon$ with $|x(0, 0)|_{\mathcal{A}} \leq \tilde{r}$ satisfies $|x(t, j)|_{\mathcal{A}} \leq \beta(|x(0, 0)|_{\mathcal{A}}, t + j) + v$, $\forall (t, j) \in \text{dom}(x)$. $\square$

The notions of P-AS can be extended to systems that depend on multiple parameters $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_\ell)$. In this case, we say that $\mathcal{A}$ is P-AS as $(\varepsilon_\ell, \dots, \varepsilon_2, \varepsilon_1) \rightarrow 0^+$ where the parameters are tuned in order starting from ε_1 .

Definitions 1 and 2 are similar to global and semi-global practical asymptotic stability properties studied in the literature [25, Ch.7], with the difference that we restrict our attention only to initial conditions in a neighborhood of the set $\mathcal{A}$. These definitions are suitable for the application under study in this paper. *Problem Statement:* Consider a general highway network system modeled by a dynamical system of the form

$$\dot{\theta} = \frac{1}{\varepsilon_0} \Pi(\theta, u), \quad y = h(\theta), \quad (2)$$

where ε_0 determines the time scale of the dynamics, $\theta \in \mathbb{R}^n$ is the state of the highway, which can include the density of the cars per unit of length in a given lane, $u \in \mathbb{R}^m$ denotes exogenous incentives which can influence the behavior of highway users (e.g., tolls), and $y \in \mathbb{R}^p$ represents measurements that can be obtained from the highway state via the output map $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Assume that an external supervisor or social planner provides a performance index $\varphi : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$, which depends on the inputs and outputs of (2). Our goal is to design feedback mechanisms able to find in real-time the optimal incentives that minimize the function $\varphi(y, u)$ at steady state. In particular, we consider closed-loop systems with the structure shown in Figure 1, where the Incentive Seeking Controller (ISC) uses only real-time output measurements of the performance index. The controller should be designed so that it can

find the “optimal” incentives while preserving closed-loop stability at all times. In the following sections, we formalize each of the components illustrated in the scheme of Figure 1.

3. TRAFFIC IN HIGHWAY NETWORKS: SOCIO-TECHNICAL MODELS

The performance of transportation systems is not solely dependent on their physical infrastructure, but also on their user behavior [26]. Indeed, in much of the literature that studies the modeling of dynamics in highway networks, the overall structure consists of a socio-technical model that combines a driver behavioral model and a traffic flow model; see [22, 27, 20]. In this work, we follow a similar approach and we assume that the socio-technical and traffic flow models can be lumped together leading to highway network dynamics described by ODEs of the form (2). Additionally, we make use of the following regularity assumption.

Assumption 1. *The map $\Pi(\cdot, \cdot)$ in (2) is locally Lipschitz. Moreover, there exists a compact set $\tilde{\Lambda}_\theta := \lambda_\theta \mathbb{B} \subset \mathbb{R}^m$ with $\lambda_\theta \in \mathbb{R}_{>0}$, a closed set $\tilde{\Lambda}_u = \tilde{\Lambda}_u + \mathbb{B}$ where $\tilde{\Lambda}_u \subset \mathbb{R}^m$, and a steady-state map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^n$ that is continuous and locally bounded relative to Λ_u , such that for each $\eta > 0$ the compact set $\mathbb{M}_\eta := \{(\theta, u) : \theta = \ell(u), u \in \Lambda_u \cap \eta\mathbb{B}, \theta \in \tilde{\Lambda}_\theta\}$ is UAS for the HDS $\mathcal{H}_{ol} := (\tilde{\Lambda}_\theta \times (\Lambda_u + \eta\mathbb{B}), \varepsilon_0^{-1}\Pi \times \{0\}, \emptyset, \emptyset)$ with state (θ, u) .*

In words, Assumption 1 guarantees that the highway dynamics are well-posed and stable with respect to external incentives u , and that the steady-state value of the traffic state is parameterized by u via the map ℓ . This assumption is standard (see [20], [14], and [11]), and it is reasonable for many socio-technical models where external incentives u determine the steady state equilibrium of the system.

3.1. Socio-Technical Model

To illustrate the advantages of the proposed ISC dynamics, we consider socio-technical models with a similar structure to the one described in [15]. In particular, we study the socio-technical model of a highway segment where drivers can choose between two parallel lanes: the general-purpose (GP) lane, which is uncharged, and the Express lane. Some of the motivations for choosing the Express lane include a faster travel time compared to the GP lane, as well as an expected reduced congestion. Prices (i.e., tolls) or subsidies can be assigned for the utilization of the Express lane depending on the traffic conditions. The model we consider focuses on the description of the average traffic density in the Express lane ρ , and the input flow of vehicles to the Express lane q_{EL} . Figure 2 shows a scheme representing the segment with the two parallel lanes: This model is divided into three main components:

- The driver behavioral model:* Each driver makes decisions based on the perceived costs of choosing either of the lanes. Some of the possible elements that can be included in these costs are travel times in the lanes, congestion,

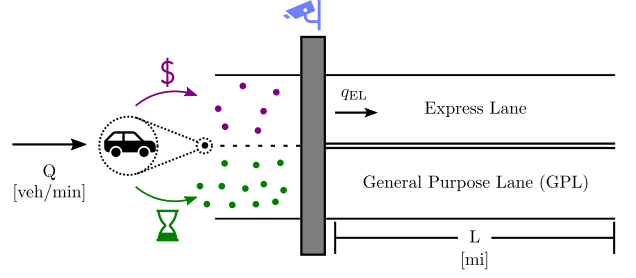


Figure 2: Scheme of segment with parallel lanes: the GP lane, and the Express lane.

road conditions, among other quantities estimated by the drivers. To capture these scenarios, we model the costs by locally Lipschitz functions $c_{EL}(q_{EL}, \rho, u)$ and $c_{GP}(q_{EL}, \rho)$ associated with the Express lane and the GP lane respectively, where the input $u \in \mathbb{R}_{\geq 0}$ represents the tolls used to incentivize or to discourage the highway users from using the Express Lane.

Naturally, the response of the individual drivers to the costs extends to the macroscopic level, concurrently affecting the input flow of vehicles to the Express Lane q_{EL} . In general, the drivers require a minimum transient time to adjust to changes in the marginal cost, which for instance could be induced by changes in the tolls. To account for this dynamic response, we model the dynamics of the macroscopic driver behavior as an ODE of the form

$$\dot{q}_{EL} = \Phi(q_{EL}, \rho, u), \quad q_{EL} \in [0, Q], \quad (3)$$

where $\Phi : [0, Q] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a locally Lipschitz function that implicitly depends on the marginal cost of choosing the EL. With (3), we are able to capture a variety of social dynamics including, among others, evolutionary population dynamics whose stability properties have been recently studied in the literature, c.f. [11].

- Equilibrium model:* The equilibrium model describes the resulting average velocity in the Express lane as a function of the average traffic density ρ . In this paper, we use a mollified version of the average velocity model presented in [15] and described by

$$\bar{v}(\rho) = \frac{v_{\text{free}} - v_{\text{jam}}}{1 + \exp\left(\frac{4}{\rho_{\text{jam}} - \rho_{\text{critical}}} \left(\rho - \frac{\rho_{\text{jam}} + \rho_{\text{critical}}}{2}\right)\right)} + v_{\text{jam}}, \quad (4)$$

where $v_{\text{free}}, v_{\text{jam}}$ are constants that represent the top speed and the jam vehicle speed in the Express lane, ρ_{critical} denotes the average density below which the speed of the vehicles is expected to be close to v_{free} , and where ρ_{jam} is the average vehicle density above which a traffic jam occurs in the Express lane. Consequently, these constants satisfy the relations: $v_{\text{free}} > v_{\text{jam}}$ and $\rho_{\text{critical}} < \rho_{\text{jam}}$.

- The traffic flow model:* This model represents the dynamics of the average traffic density ρ of the Express lane, measured in vehicles per unit of length, as a function of

Assumption 4. The dithering frequencies ω_i satisfy: 1) $\omega_i > 0$ is a rational number for all i , and 2) there are no repeated dither frequencies, i.e., $i \neq j \implies \omega_i \neq \omega_j$ and $\omega_i \neq 2\omega_j$.

Assumption (4) guarantees orthogonality conditions for the dither signals used by the ISCs to update the incentives u . These conditions will enable real-time learning in the closed-loop system via averaging theory.

4.1. Gradient Based Incentive Seeking Control

We first consider a smooth ISC, denoted GISC, presented in Figure 3, which generates the nominal incentive $\hat{u}$ via the following differential equation:

$$\begin{pmatrix} \dot{\hat{u}} \\ \dot{\mu} \end{pmatrix} = F_1(x_1) := \begin{pmatrix} -k\varphi(y, u)M(\mu) \\ \frac{1}{\varepsilon_p}\mathcal{R}\mu \end{pmatrix}, \quad x_1 \in \mathbb{R}^m \times \mathbb{T}^m, \quad (12)$$

where $x_1 := (\hat{u}, \mu)$, and $M(\mu) = \frac{2}{\varepsilon_a}\mathbb{D}\mu$. The controller (12) is based on smooth extremum-seeking controllers [28], which aim to emulate gradient flows whenever the highway dynamics (2) are neglected. The controller makes use of direct measurements of the performance index $\varphi(y, u)$, and therefore it is agnostic to the dynamics of the transportation system. In the context of traffic congestion, related dynamics have been studied in [20] for adaptive pricing in affine congestion games, [29] for highways with bottlenecks, and in [30] via simulations for congestion lanes. The following theorem shows that (12) is a suitable controller to learn optimal incentives in transportation systems with socio-technical dynamics in the loop.

Theorem 1. Suppose that Assumptions 1, 2 (or 3) and 4 hold. Then, the closed-loop system corresponding to Figure 1 with ISC given by (12), renders PAS the set $\mathcal{A}_1 := \mathcal{A}_0 \times \mathbb{T}^m$ as $(\varepsilon_0, \varepsilon_p, \varepsilon_a) \rightarrow 0^+$. $\square$

Proof. The result of Theorem 1 can be established by showing that all the assumptions needed to apply [31, Thm.1] are satisfied in a neighborhood of the optimal incentive. First, note that, by Assumption 1, the plant has a well-defined steady state input-to-output map $\tilde{\varphi}$. Also, under Assumption 2, this response map is strictly convex, and thus has a unique minimizer. Since under Assumption 4 the average dynamics of (12) can be computed to be $\dot{\hat{u}} = -k\nabla\tilde{\varphi}(u) + O(\varepsilon_a)$ (see, e.g., [23, Sec. 7]), it follows that for ε_a sufficiently small, in a neighborhood of the optimal incentive u^* the average dynamics converge to a neighborhood of u^* . By averaging theory and the results of [31, Thm.1], the original system retains the stability properties in a practical sense. The result follows by using a (local) singular perturbation argument to interconnect the dynamics (12) with the dynamics (6). $\blacksquare$

While the ISC (12) can achieve optimal incentive seeking, as $(\varepsilon_0, \varepsilon_p, \varepsilon_a) \rightarrow 0^+$ the rate of convergence achieved by this controller emulates the convergence rate of a gradient descent flow, which is either of order $O(1/t)$ or $O(e^{-\kappa t})$, where κ is given by Assumption 3 (note that Assumption 3 implies Assumption 2). In Section 5, we will show that the steady-state performance function related to the socio-technical model of the traffic network can have drastically different “slopes” near the optimal

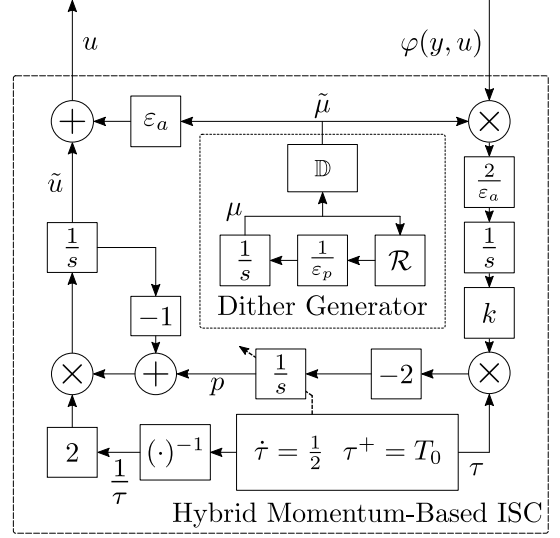


Figure 4: Hybrid Momentum Based Incentive Seeking Control

point, including cases where $\kappa \ll 1$. These “flat” regions can drastically deteriorate the performance of controllers that seek to emulate traditional gradient flows, e.g., system (12). To achieve better transient performance in this scenario, we now consider a class of hybrid dynamics that use momentum.

4.2. Hybrid Momentum-Based Incentive Seeking Control

To achieve better transient performance compared to (12), we now consider the hybrid ISC shown in Figure 4, which has continuous-time and discrete-time dynamics given by:

$$\begin{pmatrix} \dot{\hat{u}} \\ \dot{p} \\ \dot{\tau} \\ \dot{\mu} \end{pmatrix} = F_2(x_2) := \begin{pmatrix} \frac{2}{\tau}(p - \hat{u}) \\ -2k\tau\varphi(y, u)M(\mu) \\ \frac{1}{2} \\ \frac{1}{\varepsilon_p}\mathcal{R}\mu \end{pmatrix}, \quad (13a)$$

$$x_2 \in C_2 := \{x_2 \in \mathbb{R}^{2m+1} \times \mathbb{T}^m : \tau \in [T_0, T]\}, \quad (13b)$$

$$\begin{pmatrix} \hat{u}^+ \\ p^+ \\ \tau^+ \\ \mu^+ \end{pmatrix} = G_2(x_2) := \begin{pmatrix} \hat{u} \\ \sigma p + (1 - \sigma)q \\ T_0 \\ \mu \end{pmatrix}, \quad (13c)$$

$$x_2 \in D_2 := \{x \in \mathbb{R}^{2m+1} \times \mathbb{T}^m : \tau = T\}, \quad (13d)$$

where $x_2 := (\hat{u}, p, \tau, \mu)$, $k \in \mathbb{R}_{>0}$ is a tunable gain. This controller resets the states p and τ via (13c) every time the timer τ satisfies $\tau = T$. The constants $0 < T_0 < T$ are tunable parameters that characterize the frequency of the resets. The parameter $\sigma \in \{0, 1\}$ describes the resetting policy for the state p . Namely, when $\sigma = 1$, we have that $p^+ = p$, while $\sigma = 0$ leads to $p^+ = q$.

In contrast to (12), as $(\varepsilon_p, \varepsilon_a, \varepsilon_0) \rightarrow 0^+$ the hybrid ISC (13) will emulate the behavior of a regularized version of Nesterov’s accelerated ODE with momentum [32], given by $\ddot{u} + \frac{3}{t}\dot{u} + \nabla\tilde{\varphi}(u) = 0$, which achieves rates of convergence of order $O(1/t^2)$ in convex functions, or $O(e^{-\sqrt{\kappa}t})$ with suitable resets corresponding to $\sigma = 0$ in (13c). These resets are similar

in spirit to “restarting” techniques used in the literature of machine learning [33]. In the context of model-free feedback control, the resets guarantee enough regularity and robustness in the controller so that it can be interconnected with a dynamical plant in the loop [23]. Thus, the hybrid controller is also able to achieve incentive seeking.

Theorem 2. Suppose that Assumptions 1, 2 (or 3) and 4 hold. are satisfied. Then, the closed-loop system corresponding to Figure 1 with ISC given by (13), renders PAS the set $\mathcal{A}_2 := \{(\theta, \hat{u}, p, \tau) : (\theta, \hat{u}) \in \mathcal{A}_0, p = \hat{u}, \tau \in [T_0, T]\} \times \mathbb{T}^m$ as $(\varepsilon_0, \varepsilon_p, \varepsilon_a) \rightarrow 0^+$. $\square$

Proof. We prove Theorem 2 following a similar approach as in Theorem 1. In particular, first note that the hybrid dynamics (13) are well-posed in the sense of [25, Sec. 6] because the sets C_2 and D_2 are closed, and the maps F_2 and G_2 are continuous on these sets. Moreover, neglecting the socio-technical dynamics, and using Assumption 4, the average dynamics of (13) correspond to an $O(\varepsilon_a)$ -perturbed version of the hybrid Nesterov gradient dynamics studied in [23] for the model-free optimization of static maps. Under Assumptions 1, 2 or 3, and 4, these average hybrid dynamics render the set $\{(\hat{u}, p, \tau) : \hat{u} = p = \operatorname{argmin} \tilde{\varphi}(u), \tau \in [T_0, T]\}$ (locally) practically asymptotically stable. By using, sequentially, averaging and singular perturbation theory for perturbed hybrid systems [23, Thm. 7], we obtain the desired result for the interconnection between the controller and the socio-technical dynamics, which are stable under Assumption 1. ■

The key advantage of the ISC (13) is the incorporation of *dynamic momentum* via the states (p, τ) , as well as periodic resets with frequency dependent on the pair (T_0, T) . Note that “optimal” restarting frequencies can be used as in (23) to avoid oscillations in the control action induced by the presence of momentum. It is well-known that momentum-based optimization algorithms can significantly improve the transient performance in problems where the cost $\tilde{\varphi}$ exhibits shallow convexity properties (e.g., $\kappa \ll 1$). As shown later in Section 5, this will be the case under certain operation conditions of the highway networks.

On the other hand, when the steady state performance function $\tilde{\varphi}$ is strongly convex and its curvature is not necessarily small, one might wonder if it is possible to achieve better transient performance using non-smooth re-scaled gradient-based dynamics. We investigate this scenario in Section 4.3.

4.3. Fixed-Time Incentive Seeking Control

We now consider the *fixed-time ISC* (FxISC) presented in Figure 5 and described by the following dynamics:

$$\begin{pmatrix} \dot{u} \\ \dot{\xi} \\ \dot{\mu} \end{pmatrix} = F_3(x_3) := \begin{pmatrix} -k \left(\frac{\xi}{|\xi|^\alpha} + \frac{\xi}{|\xi|^{-\alpha}} \right) \\ \frac{1}{\varepsilon_f} (-\xi + \varphi(y, u)M(\mu)) \\ \frac{1}{\varepsilon_\mu} \mathcal{R}\mu \end{pmatrix} \quad (14)$$

where $x_3 := (\hat{u}, \xi, \mu)$, and $\alpha \in (0, 1)$ is a tunable exponent, and where the right hand side of $\dot{u}$ is defined to be zero whenever

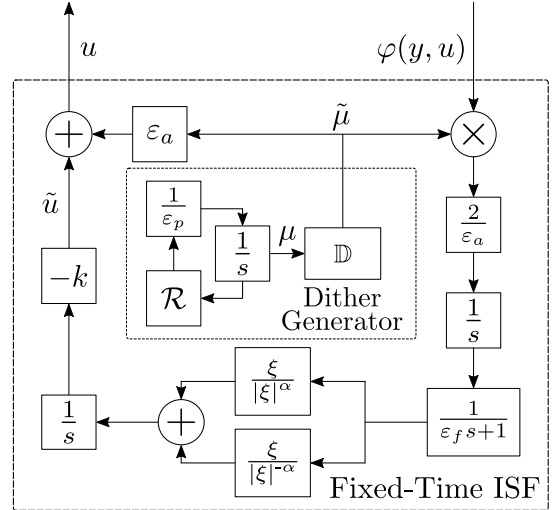


Figure 5: Fixed-Time Incentive Seeking Control

$\xi = 0$. In this controller, we have incorporated a low-pass filter with state ξ and gain ε_f^{-1} , and the nominal incentive $\hat{u}$ is generated by a combination of sub-linear and super-linear feedback parametrized by the constant α . Note that the vector field (14) is continuous but not Lipschitz continuous at $\xi = 0$. The controller is designed to emulate the performance of fixed-time gradient flows [34, 24] as $(\varepsilon_0, \varepsilon_a, \varepsilon_p, \varepsilon_f) \rightarrow 0^+$. This non-smooth ISC also achieves optimal incentive seeking, but it requires (regional) strong convexity of $\tilde{\varphi}$.

Theorem 3. Suppose that Assumptions 1, 3 and 4 hold. Then, the closed-loop system corresponding to Figure 1 with ISC given by (14), renders PAS the set $\mathcal{A}_3 := \mathcal{A}_0 \times \{0\} \times \mathbb{T}^m$ as $(\varepsilon_0, \varepsilon_p, \varepsilon_a, \varepsilon_f) \rightarrow 0^+$. $\square$

Proof. Neglecting the socio-technical dynamics, the average dynamics of (14) are given by a perturbed version of the fixed-time gradient flows studied in [34]. Under Assumption 3, these dynamics render the optimal incentive fixed-time stable. A direct application of averaging theory for non-smooth systems [24] allows us to conclude practical (with respect to $\tilde{\Lambda}_u$) fixed-time stability for the ISC interconnected with the socio-technical dynamics (2). ■

In contrast to (12) and (13), as $(\varepsilon_0, \varepsilon_p, \varepsilon_a, \varepsilon_f) \rightarrow 0^+$, the non-smooth ISC (14) emulates the behavior of gradient flows able to converge to the optimal incentive before a fixed time $T^* = \frac{\pi}{2\kappa\alpha k}$, where (α, k) are tunable parameters of the controller, and κ is given by Assumption 3. Such type of behavior cannot be obtained using smooth (i.e., Lipschitz continuous) ISCs.

5. NUMERICAL EXAMPLES

In this section, we consider particular realizations of the model introduced in Section 3, as well as numerical examples of the proposed ISCs.

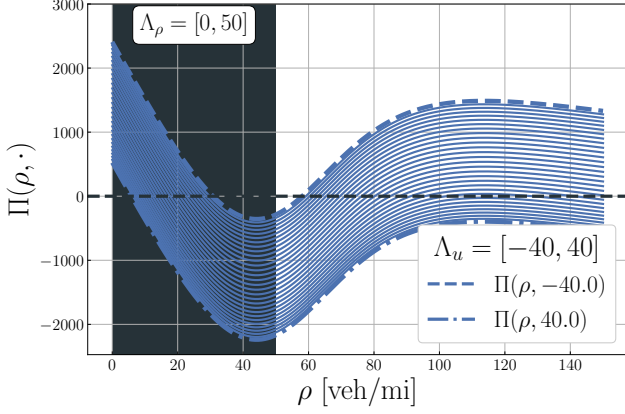


Figure 6: Evaluation of viability conditions for highway system with fast driver behavior and static input flow map q_{EL} .

5.1. Fast Driver Behavior

We first consider a scenario where the driver dynamics are qualitatively faster than the average traffic dynamics. Specifically, we borrow the parameters and structure used in [16], based on traffic data of the first dynamic-pricing toll system implemented in the United States: the MnPASS. Thus, the costs of choosing the Express or GP lanes are given by: $c_{EL}(\rho, u) = a \frac{L}{v(\rho)} + bu + \gamma_{EL}$, and $c_{GP} = a \frac{L}{v(\rho)} \delta + \gamma_{GP}$, where $\frac{L}{v(\rho)}$ represents the estimated travel time on the Express lane, $\gamma_{EL}, \gamma_{GP} \in \mathbb{R}$ are offsets used to represent unobservable quantities, a, b are positive weights, and where $\delta \geq 1$ models the fact that the travel time through the GP lane is assumed to be longer or equal than the one of the Express lane. On the other hand, the macroscopic driver behavior is assumed to have the form:

$$q_{EL}(\rho, u) = \frac{Q}{1 + \exp(c_{EL}(\rho, u) - c_{GP}(\rho))}, \quad (15)$$

which is a logistic function of the marginal cost of choosing the Express lane over the GP lane. The choice of static map in (15) implies that whenever the perceived cost c_{EL} of choosing the Express lane is lower than the cost c_{GP} of choosing the GP lane, the input flow of vehicles to the Express lane will increase. Moreover, we note that an equal inflow of vehicles to the Express and GP lanes is achieved when the marginal cost is equal to zero. To simulate the ISCs we use the parameters $a = 0.334, b = 0.335, \gamma_{EL} = 1.71781, \gamma_{GP} = 0, v_{jam} = 5[\text{mph}], v_{free} = 65[\text{mph}], \rho_{jam} = 80[\text{veh/mi}], \rho_{critical} = 25[\text{veh/mi}]$ and $L = 0.7[\text{mi}]$. To establish a reference density that guarantees free-flow conditions $\bar{v}(\rho_{ref}) \approx v_{free}$ with a moderate occupation of the lane so that the system is not under-utilized, we consider the reference seeking performance index φ_{ref} given in (9), with $\rho_{ref} = 0.8\rho_{critical} = 20 [\text{veh/mi}]$. Furthermore, we fix the demand to be $Q = 2170$ vehicles per hour and set $\tilde{\tau} = 3$. Using these parameters, we conduct a numerical study that verifies that Assumptions 1-3 are satisfied: First, we plot in Figure 6 the vector field (6) for different values of $u \in \Lambda_u := [-40, 40]$ and with $\rho = \theta \in [0, 150]$. By arguing graphically, we note that there exists a compact set (interval) $\Lambda_\rho := [0, 50]$, such that restricted to values $(\rho, u) \in \Lambda_\rho \times \Lambda_u$, for

each u there exists a unique asymptotically stable equilibrium $\rho^*(u) \in \Lambda$. Therefore, we can define the function $\ell : \Lambda_u \rightarrow \mathbb{R}$ as $\ell(u) = \rho^*(u)$, which is shown in Figure 7. Since Π in (8) is locally Lipschitz, the previous arguments imply that the socio-technical model of Section 3 with the particular parameters listed above, satisfy the conditions of Assumption 1. On the other hand, φ_{ref} satisfies Assumptions 2 and 3 by construction and the convexity of $\tilde{\varphi}(u) = \varphi(\ell(u), u)$ in Λ_u .

Following the closed-loop structure of Figure 1 we implement the different ISCs introduced in Section 4 interconnected with the highway dynamics of (8). For all the controllers, we set $k = 1$ and use the dithering frequency $\omega = 1$. In the case of the hybrid ISC we choose $\sigma = 0, T_0 = 0.1$ and $T = 20$. For the non-smooth ISC we use $\alpha = 0.5$. We simulate the trajectories of the closed-loop systems using $\varepsilon_f = 1, \varepsilon_a = 0.1$, and $\varepsilon_p = 0.01$. These parameters guarantee enough time-scale separation between the different elements of the controller, and also between the controller and the highway-dynamics. We uniformly sample 60 different initial conditions for ρ between 4 and 30. We use the initial incentive $u(0) = 1$, and we plot the resulting trajectories in Figure 8. Additionally, we compute and plot the mean squared error (MSE) $MSE(t) = \frac{1}{60} \sum_{i=1}^{60} |\rho_i(t) - \rho_{ref}|^2$, where ρ_i corresponds to the trajectory resulting from the i -th initial condition. As shown in the figure, the hybrid and non-smooth ISCs significantly outperform the smooth ISC algorithm (12). Note that the hybrid algorithm (13) generates the typical oscillatory behavior observed in momentum-based algorithms when the damping is sufficiently small. Note also that the hybrid controller seems to generate better transient performance compared to (14), since in certain cases it generates a smaller overshoot. The inset of Figure 8 shows that the control signals converge to a small neighborhood of the optimal incentive. Finally, we note that all the ISCs studied in this paper are well-posed by construction, and therefore they are robust with respect to small bounded additive disturbances acting on the states and dynamics [25, Thm. 7.21]. Moreover, their model-free nature allows them to retain their stability and convergence properties when the parameters of the highway change (slowly) over time. For example, Figure 9 shows the impact of variations on the parameter γ_{EL} . Here, we sampled uniformly 20 different values of γ_{EL} with a maximum variation of 15% with respect to the nominal value 1.71781, and we simulated the closed-loop dynamics for each one of the ISCs. For each value of γ_{EL} we computed the mean of the time-average MSE, $\overline{MSE} := \frac{1}{t_f} \int_0^{t_f} MSE(\tau) d\tau$, where $t_f = 225[\text{min}]$ is the final time of a simulation run, over 5 trajectories obtained by choosing different initial conditions for ρ on the range $[4, 30]$. As seen in Figure 9, the results are consistent with the previous results shown in Figure 8.

5.2. Dynamics for Macroscopic Driver Behavior

When the macroscopic driver behavior and the average vehicle density evolve in similar time scales, the dynamics capturing the driver's response due to changes in the incentives need to be considered in the closed-loop system. In this case, we consider the following dynamics to describe the evolution of q_{EL} :

$$\dot{q}_{EL} = \Psi(q_{EL}, u) := -\left(c_{EL}(q_{EL}, u) - c_{GP}(q_{EL})\right), \quad (16)$$

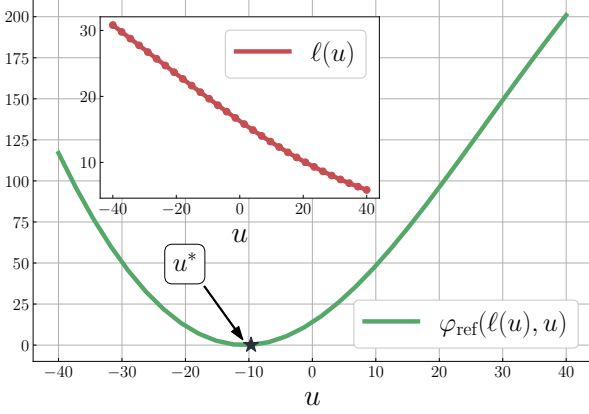


Figure 7: Suitability of performance index φ_{ref} and response map $\ell(\cdot)$ for the ISC dynamics presented in this paper.

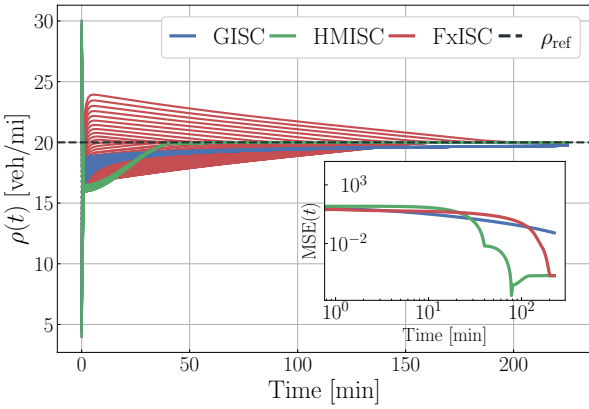


Figure 8: Trajectories corresponding to 60 different initial conditions of the average density ρ sampled uniformly between $\rho(0) = 4$ and $\rho(0) = 30$ vehicles per mile. The inset shows the MSE as a function of time of the 60 trajectories.

meaning that the rate of change of the input flow of vehicles to the Express lane is directly determined by the marginal cost of choosing that lane over the GP lane. Consequently, when the perceived cost c_{EL} of choosing the Express lane is lower than the cost c_{GP} of choosing the GP lane, the rate of growth will be instantaneously positive, thus increasing the input flow of vehicles to the Express lane. We considered the marginal cost

$$c_{\text{EL}}(q_{\text{EL}}, u) - c_{\text{GP}}(q_{\text{EL}}) = \left(q_{\text{EL}} - \frac{Q}{2}\right) + \tilde{a}u,$$

where $\tilde{a} > 0$. For our simulations we use $\tilde{a} = 100$ and the same values of $v_{\text{jam}}, v_{\text{free}}, \rho_{\text{jam}}, \rho_{\text{critical}}, Q$ and ρ_{ref} considered in Section 5.1. To study the stability properties of the equilibrium points of the highway dynamics, we analyze the phase planes associated to system (6) using $\Phi = \Psi$, $k_m = 1$, $k_p = 1$, and $u \in \Lambda_u = [-40, 40]$, and shown in Figure 10 for three particular values of u . In all cases, there exists a compact set $\Lambda_\theta \subset [0, 160] \times [0, Q]$, such that for each u there exists a unique asymptotically stable equilibrium $\theta^*(u) \in \Lambda_\theta$. The same property was numerically confirmed to hold for every $u \in \Lambda_u$ by studying the phase plane plots associated to equally spaced inputs taken from Λ_u , and using continuity of (6). Therefore, we can define a response function $\tilde{\ell} : \Lambda_u \rightarrow \mathbb{R}^2$ by letting

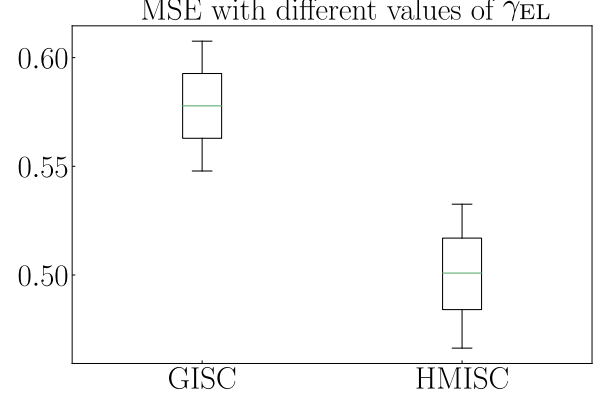


Figure 9: Distribution of the mean of the time-averaged MSE for different values of γ_{EL} sampled uniformly within 15% of its nominal value.

$\tilde{\ell}(u) = \theta^*(u)$. The remaining conditions of Assumption 1 and Assumptions 2 and 3 are verified to hold by following analogous graphical arguments to the ones described in Section 5.1. Indeed, Figure 11 shows the corresponding plots describing the response function $\tilde{\ell}$ and the performance index $\varphi_{\text{ref}}(\tilde{\ell}(u), u)$. In this case, we focus our attention on the two ISCs that showed the best performance in Section 5.1, and implement the closed-loop structure of Figure 1 using the GISC and HMISC. For both controllers we set $\varepsilon_0 = 0.1$, $\varepsilon_\mu = 0.01$, $\varepsilon_a = 0.001$, $k = 0.01$, and $\omega = 1$. For the HMISC, we chose $\sigma = 1$, $T_0 = 0.01$ and $T = 0.5$. The control parameters are selected to guarantee enough time-scale separation between the different elements of the controller and the highway-dynamics. As seen in Figure 12, where we plotted the MSE corresponding to 20 different trajectories satisfying $u(0) = 1$ and $q_{\text{EL}}(0) = Q/3$, the HMISC outperforms the GISC, although in a mildly less dominant fashion to what was observed in Section 5.1. This reduction in the performance gap between the GISC and the HMISC, can be mainly attributed to the fact that, in this setup, the strong convexity parameter is relatively high, meaning that the transient performance increase attained via momentum-based dynamics is not as evident as when the costs exhibit shallow convexity properties.

Finally, we note that in all our numerical experiments the ISCs were tuned to guarantee that the drivers have enough time to react to changes on the incentives induced by the exploratory signal $\tilde{\mu}$ used by the controllers. This behavior is needed to guarantee real-time learning via feedback measurements of the output of the highway network, and it has also been studied in algorithms based on adaptive control [20] and reinforcement learning [35], to name just a few. Potential extensions that could relax these real-time exploration requirements might be studied in the future by incorporating historical data into the controllers, which can be periodically updated during days or weeks to retain sub-optimality of the incentives. Such controllers will naturally be modeled as hybrid dynamical systems.

6. CONCLUSIONS

We introduced a new class of incentive-seeking controllers (ISCs) that can learn optimal incentives using only output mea-

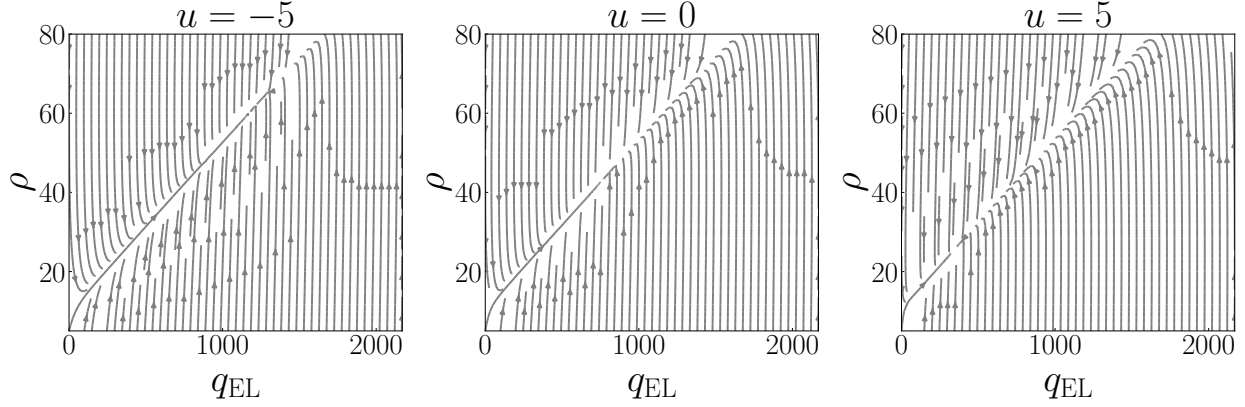


Figure 10: Evaluation of the viability conditions for highway system with dynamics describing the macroscopic driver behavior, based on the phase plane of the system.

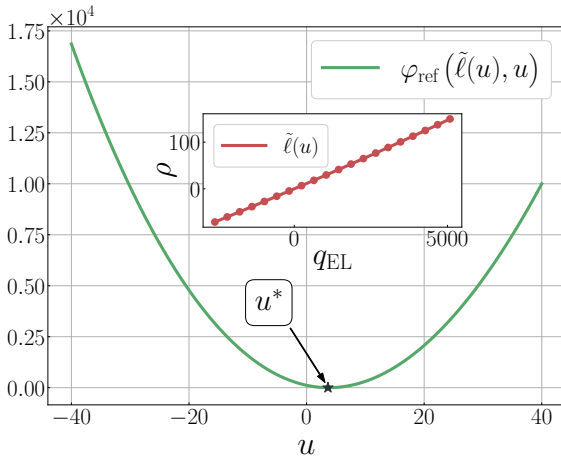


Figure 11: Suitability of performance index φ_{ref} and response map $\tilde{l}$ for the ISC with highway model including dynamics for the macroscopic driver behavior. The inset shows the response function $\tilde{l}$ projected in the phase plane ρ vs. q_{EL} for values of $u \in [-40, 40]$.

surements from traffic in transportation systems, while simultaneously guaranteeing closed-loop stability. We illustrated the benefits of the proposed controllers via numerical experiments in a socio-technical model of a highway system with managed lanes, including the advantages of using nonsmooth and hybrid controllers. The algorithms are agnostic to the exact model of the highway, and robust to small additive disturbances. Future research directions will focus on incorporating past recorded data to minimize real-time exploration in the controllers.

References

- [1] M. Lab, "U.S. is the world leader in traffic jams – USA today," 2018. <https://mobilitylab.org/2018/02/06/u-s-is-the-world-leader-in-traffic-jams/>.
- [2] R. Kutadinata, W. Moase, C. Manzie, L. Zhang, and T. Garoni, "Enhancing the performance of existing urban traffic light control through extremum-seeking," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 1–20, 2016.
- [3] P. N. Brown and J. R. Marden, "Studies on robust social influence mechanisms," *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 98–115, 2017.

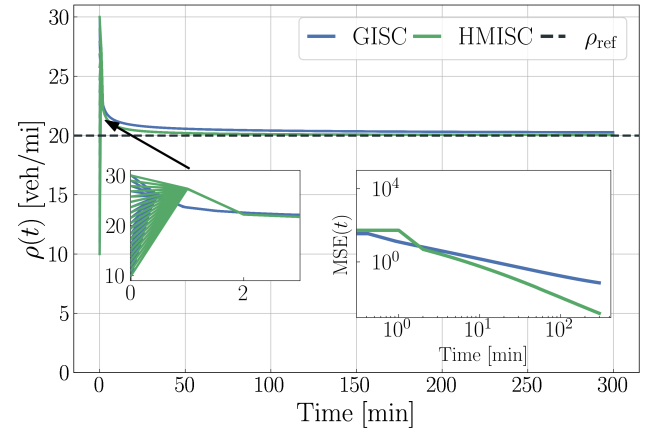


Figure 12: Trajectories resulting from the application of the GISC and HMISC to the highway model with dynamics for the macroscopic driver behavior, and corresponding to 20 different initial conditions of the average density ρ sampled uniformly between $\rho(0) = 10$ and $\rho(0) = 30$ vehicles per mile. The inset shows the MSE as a function of time of the 20 trajectories.

- [4] J. I. Poveda, P. N. Brown, J. R. Marden, and A. R. Teel, "A class of distributed adaptive pricing mechanisms for societal systems with limited information," *56th IEEE Conference on Decision and Control*, pp. 1490–1495, 2017.
- [5] A. Kleiner, B. Nebel, and V. A. Ziparo, "A mechanism for dynamic ride sharing based on parallel auctions," *Proc. 22nd Int. Joint Conference on Artificial Intelligence*, pp. 266–272, 2011.
- [6] E. Provonsha and N. Sifuentes, "Road pricing in london, stockholm and singapore, a way forward for new york city," tech. rep., Tri-State Transportation Campaign, 2017.
- [7] T. Litman, "London congestion pricing - implications for other cities," *CESifo DICE Report*, ISSN 1613-6373, ifo Institut für Wirtschaftsforschung an der Universität München, München, vol. 03, no. 3, pp. 17–21, 2005.
- [8] E. Croci and A. Ravazzi, "Urban road pricing: A comparative study on the experiences of london, stockholm and milan," *The Center for Research on Energy and Environmental Economics and Policy at Bocconi University, Working Paper Series - ISSN 1973-0381*, pp. 1–43, 2016.
- [9] B. Schaller, "New york city's congestion pricing experience and implications for road pricing acceptance in the united states," *Transport Policy*, vol. 4, no. 17, pp. 266–273, 2010.
- [10] A. C. Pigou, *The Economics of Welfare*. New York: Macmillan, 1920.
- [11] W. Sandholm, "Evolutionary implementation and congestion pricing," *Review of Economic Studies*, vol. 69, no. 667–689, 2002.
- [12] U. Bhaskar, K. Ligett, and L. J. Schulman, "Achieving target equilibria in

- network routing games without knowing the latency functions,” *IEEE Annual Symposium on Foundations of Computer Science*, pp. 31–40, 2014.
- [13] Y. Nie, X. Wang, and K. Cheng, “Multi-area self-adaptive pricing control in smart city with ev user participation,” *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–9, 2017.
- [14] F. Farokhi and K. H. Johansson, “A piecewise-constant congestion taxing policy for repeated routing games,” *Transportation Research Part B*, vol. 78, pp. 123–143, 2015.
- [15] A. M. Annaswamy, Y. Guan, H. E. Tseng, H. Zhou, T. Phan, and D. Yanakiev, “Transactive control in smart cities,” *Proceedings of the IEEE*, vol. 106, no. 4, pp. 518–537, 2018.
- [16] T. Phan, A. M. Annaswamy, D. Yanakiev, and E. Tseng, “A model-based dynamic toll pricing strategy for controlling highway traffic,” in *2016 American Control Conference (ACC)*, pp. 6245–6252, IEEE, 2016.
- [17] G. Zhang, X. Ma, and Y. Wang, “Self-adaptive tolling strategy for enhanced high-occupancy toll lane operations,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 306–317, 2013.
- [18] N. Zheng, G. Rérat, and N. Geroliminis, “Time-dependent area-based pricing for multimodal systems with heterogeneous users in an agent-based environment,” *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 133–148, 2016.
- [19] P. Kachroo, S. Gupta, S. Agarwal, and K. Ozbay, “Optimal control for congestion pricing: Theory, simulation, and evaluation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1234–1240, 2016.
- [20] J. I. Poveda, P. N. Brown, J. R. Marden, and A. R. Teel, “A class of distributed adaptive pricing mechanisms for societal systems with limited information,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 1490–1495, IEEE, 2017.
- [21] G. Bianchin, J. Cortés, J. I. Poveda, and E. Dall’Anese, “Time-varying optimization of LTI systems via projected primal-dual gradient flows,” 2021. *IEEE Transactions on Control of Network Systems*, DOI 10.1109/TCNS.2021.3112762.
- [22] C. Lombardi, L. Picado-Santos, and A. M. Annaswamy, “Model-based dynamic toll pricing: An overview,” *Applied Sciences*, vol. 11, no. 11, p. 4778, 2021.
- [23] J. I. Poveda and N. Li, “Robust hybrid zero-order optimization algorithms with acceleration via averaging in continuous time,” *Automatica*, vol. 123, 2021.
- [24] J. I. Poveda and M. Krstic, “Non-smooth extremum seeking control with user-prescribed convergence,” *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 6156–6163, 2021.
- [25] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. Princeton University Press, 2012.
- [26] P. N. Brown and J. R. Marden, “Studies on robust social influence mechanisms: Incentives for efficient network routing in uncertain settings,” *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 98–115, 2017.
- [27] L. Engelson and P. O. Lindberg, “Congestion pricing of road networks with users having different time values,” in *Mathematical and computational models for congestion charging*, pp. 81–104, Springer, 2006.
- [28] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum-Seeking Control*. Wiley, 2003.
- [29] H. Yu, S. Koga, T. R. Oliveira, and M. Krstic, “Extremum seeking for traffic congestion control with a downstream bottleneck,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 143, no. 3, p. 031007, 2021.
- [30] L.-O. Sánchez Zamora, “Dynamic toll pricing for express lanes: a model-free feedback control approach,” Master’s thesis, Universitat Politècnica de Catalunya, 2020.
- [31] J. I. Poveda and A. R. Teel, “A framework for a class of hybrid extremum seeking controllers with dynamic inclusions,” *Automatica*, vol. 76, pp. 113–126, 2017.
- [32] W. Su, S. Boyd, and E. Candes, “A differential equation for modeling Nesterov’s accelerated gradient method: Theory and insights,” *J. of Machine Learning Research*, vol. 17, no. 153, pp. 1–43, 2016.
- [33] O’Donoghue and E. J. Candès, “Adaptive restart for accelerated gradient schemes,” *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 715–732, 2013.
- [34] K. Garg and D. Panagou, “Fixed-time stable gradient flows: Applications to continuous-time optimization,” *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2002–2015, 2020.
- [35] F. Zhu and S. V. Ukkusuri, “A reinforcement learning approach for distance-based dynamic tolling in the stochastic network environment,” *Journal of Advanced Transportation*, vol. 49, no. 2, pp. 247–266, 2015.