A simulation-based traffic signal control for congested urban traffic networks

Baldi, Simone; Michailidis, Iakovos; Ntampasi, Vasiliki; Kosmatopoulos, Elias; Papamichail, Ioannis; Papageorgiou, Markos

DOI
10.1287/trsc.2017.0754

Publication date
2019

Document Version
Final published version

Published in
Transportation Science

Citation (APA)
Baldi, S., Michailidis, I., Ntampasi, V., Kosmatopoulos, E., Papamichail, I., & Papageorgiou, M. (2019). A simulation-based traffic signal control for congested urban traffic networks. Transportation Science, 53(1), 6-20. https://doi.org/10.1287/trsc.2017.0754

Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Simulation-Based Traffic Signal Control for Congested Urban Traffic Networks

Simone Baldi, Iakovos Michailidis, Vasiliki Ntampasi, Elias Kosmatopoulos, Ioannis Papamichail, Markos Papageorgiou

informs
http://pubsonline.informs.org/journal/trsc/

TRANSPORTATION SCIENCE
Vol. 53, No. 1, January–February 2019, pp. 6–20
ISSN 0041-1655 (print), ISSN 1526-5447 (online)

A Simulation-Based Traffic Signal Control for Congested Urban Traffic Networks

Simone Baldi,[a,b] Iakovos Michailidis,[b] Vasiliki Ntampasi,[b] Elias Kosmatopoulos,[b,c] Ioannis Papamichail,[d] Markos Papageorgiou[d]

[a] Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, Netherlands; [b] Informatics and Telematics Institute, Center for Research and Technology Hellas (ITI-CERTH), 57001 Thessaloniki, Greece; [c] Department of Production Engineering and Management, Technical University of Crete, 73100 Chania, Greece; [d] Department of Electrical and Computer Engineering, Democritus University of Thrace, 67100 Xanthi, Greece

Contact: s.baldi@tudelft.nl (SB); michaild@iti.gr (IM); vntampas@iti.gr (VN); kosmatop@iti.gr (EK); ipapa@dssl.tuc.gr (IP); markos@dssl.tuc.gr (MP)

Received: November 10, 2015
Revised: March 27, 2016; July 13, 2016
Accepted: December 24, 2016
Published Online in Articles in Advance: November 6, 2017

https://doi.org/10.1287/trsc.2017.0754

Abstract. Traffic congestion in urban networks may lead to strong degradation in the utilization of the network infrastructure, which can be mitigated via suitable control strategies. This paper studies and analyzes the performance of an adaptive traffic-responsive strategy that controls the traffic light parameters in an urban network to reduce traffic congestion. A nearly optimal control formulation is adopted to avoid the curse of dimensionality occurring in the solution of the corresponding Hamilton–Jacobi–Bellman (HJB) optimal control problem. First, an (approximate) solution of the HJB is parametrized via an appropriate Lyapunov function; then, the solution is updated at each iteration in such a way to approach the nearly optimal solution, using a close-to-optimality index and information coming from the simulation model of the network (simulation-based design). Simulation results obtained using a traffic simulation model of the network Chania, Greece, an urban traffic network containing many varieties of junction staging, demonstrate the efficiency of the proposed approach, as compared with alternative traffic strategies based on a simplified linear model of the traffic network. It is shown that the proposed strategy can adapt to different traffic conditions and that low-complexity parametrizations of the optimal solution, a linear and a bimodal piecewise linear strategy, respectively, provide a satisfactory trade-off between computational complexity and network performance.

History: This paper has been accepted for the *Transportation Science* Special Issue on Recent Advances in Urban Transportation Through Optimization and Analytics.

Funding: The research leading to these results has been partially funded by the European Commission [FP7-ICT-5-3.5], Engineering of Networked Monitoring and Control Systems, under the contract [257806] (AGILE) and [FP7-ICT-2013.3.4], Advanced computing, Embedded Control Systems, under the contract [611538] (LOCAL 4GLOBAL).

Keywords: urban traffic control • traffic-responsive strategy • approximately optimal control

1. Introduction

Urban traffic congestion appears when too many vehicles attempt to use a common infrastructure with limited capacity and saturate it. Saturated links may prevent upstream flow from crossing, even with a green signal, possibly leading to an increased waste of green time and gridlocks in the network (Abu-Lebdeh and Benekohal 1997). A better utilization of the existing infrastructure via appropriate traffic signal control (TSC) strategies can mitigate congestion and improve urban mobility without the need for increasing the available infrastructure. Over the past three decades, several traffic control strategies have been proposed and successfully implemented in actual cities, as reported in Bell (1992); Gartner, Pooran, and Andrews (2001); Farges, Khoudour, and Lesort (1990); Boillot, Midenet, and Pierrele (2006); Siemens (2000); Friedrich (2002); Robertson and Bretherton (1991); Osorio and Chong (2015); and many other works. Despite the heterogeneity of these approaches, a commonly recognized challenge in the urban traffic control problem is related to the so-called Bellman's curse of dimensionality, which arises when a large number of states and parameters makes the online solution of the control problem intractable. Traffic theorists and engineers have come up with different solutions to avoid the curse of dimensionality and render the problem tractable for large-scale traffic instances. Three major approaches can be classified as distributed/hierarchical approaches; suboptimal approaches based on heuristics optimization; and suboptimal approaches based on simplifications in the traffic model. In the first group falls, for example, the OPAC strategy, extensively tested in the United States (Gartner, Pooran, and Andrews 2001), distributes the control among individual intersections and focuses on efficient coordinated control of the intersections in the network. Distributed control is also adopted in other recent signal control strategies based on stochastic optimal control (Sheu 2002), Petri-net (Di Febbraro, Giglio, and Sacco 2004), job-scheduling (Xie et al. 2012), reinforcement learning (El-Tantawy, Abdulhai, and Abdelgawad 2014), and multiagent principles (Bazzan 2009). Another celebrated strategy that divides the network into small subnetworks and builds distributed controllers is the SCOOT strategy, deployed in many cities in Great Britain and around the globe (Robertson and Bretherton 1991). The second group of suboptimal approaches based on heuristics optimization includes the following: the bi-level sensitivity analysis algorithm in Yang and Yagar (1995), which approximates the derivative of flows and queue with respect to signal splits; the optimization module of the CRONOS strategy, tested in Paris (Boillot, Midenet, and Pierrele 2006), which consists of a modified heuristic version of the box algorithm that does not investigate the entire set of solutions; and the heuristic suboptimal forward dynamic programming of the PRODYN strategy, developed and tested in France (Farges, Khoudour, and Lesort 1990). In the third group of traffic strategies adopting simplifications in the traffic model we can mention the following: the Traffic responsive Urban Control (TUC) strategy (Aboudolas et al. 2010), which solves a linear-quadratic problem based on a store-and-forward simplification of the traffic flow; the DISCO or mixed-integer linear program approaches where traffic is modeled after the cell-transmission model (Lo, Chang, and Chan 2001; Lo 2001); model predictive control strategies based on the simplified S-model (Lin et al. 2012); and gating feedback regulators derived via a simplified dynamic model based on the network fundamental diagram of traffic flow (Keyvan-Ekbatani et al. 2012). We conclude this nonexhaustive overview by mentioning recently emerging approaches for optimal control of the transportation network (Han, Szeto, and Friesz 2015; Li, Canepa, and Claudel 2014): these methods rely on a Lighthill–Whitham–Richards traffic flow model and characterize the optimal solution by the Lax–Hopf formula. The advantage is that no specific approximations nor discretization are required; however, one limitation is currently the lack of computational tractability, when the problem size scales up. From this overview it emerges that all implementable signal control strategies must include some simplifications or heuristics, either in their modeling approach, or in their optimization algorithm, or in their extent of network coverage.

In this paper, the urban traffic control problem will be solved using an adaptive approximately optimal control strategy for traffic signal control, justified by the following reasons:

Approximately optimal. The adoption of an approximately optimal control strategy avoids the Bellman's curse of dimensionality. It moderates the computational complexity of solving the optimal control problem by parameterizing the optimal solution. In the proposed approach, the parametrization is developed in such a way that the designer can control the trade-off computational complexity/close to optimality;

Adaptive As the traffic network dynamics are influenced by the traffic demand, saturation flows and turning rates at each link might exhibit time-varying behavior (Cremer 1991; Jacob and Abdulhai 2001); thus, different strategies must be delivered in the presence of modified traffic conditions. In the proposed approach, adaptation arises from a mechanism that generates at each iteration a number of candidate control strategies, and selects the best one based on the estimation of its close to optimality and on the information coming from the simulation model of the network (simulation-based design).

The main contribution of the paper and the advantages of the proposed method with respect to state-of-the-art methods can be identified as follows. With respect to distributed/hierarchical approaches, the proposed method is fully centralized, exploiting information stemming from all of the links in the network; with respect to suboptimal approaches based on heuristics for optimization, the trade-off of computational complexity/near optimality in the proposed approach can be controlled by the user by increasing the complexity of the control law; with respect to suboptimal approaches based on simplifications in the traffic model, the proposed strategy can rely on elaborate traffic simulation environments. A well-known problem of adopting simplified traffic flow models is that, when implemented in real life, the performance of the resulting strategy may be far from optimal and calibration of the control gains might be necessary, which can be time consuming because of the large-scale and complex nature of the traffic system (Li, Tang, and Head 2003; Sanchez, Galan, and Rubio 2008; Kosmatopoulos et al. 2007; Baldi et al. 2015).

Simulation results obtained using an AIMSUN (Barcelo, Casas, and Ferrer 1999) model of the traffic network of Chania, Greece, an urban traffic network containing many varieties of junction staging (45 control inputs and 122 occupancy/flow states involved), demonstrate the efficiency of the proposed approach via the following points: (a) low-complexity parametrizations of the optimal solution, respectively, a linear traffic-responsive strategy and a bimodal piecewise linear one, are sufficient to overcome the performance of alternative traffic control strategies based on a simplified linear model of the traffic network; (b) the low computational complexity of the proposed traffic-responsive strategy makes it feasibly implementable online in urban networks with a large number of sections and junctions; (c) the proposed method is capable of adapting to different traffic conditions, as shown via simulations with three traffic demand scenarios of increasing intensity. A comparative study with different traffic control strategies reveals variations in the network dynamics and the need for control strategies tailored to different traffic demand.

The paper is organized as follows: the problem formulation of urban traffic control can be found in Section 2. Section 3 presents the approximately optimal control framework and the optimization algorithm used for the solution of the urban traffic control problem. The simulation setup is discussed in Section 4. Simulation results are given in Section 5. Section 6 concludes the paper.

2. Control of Urban Traffic Networks

Traffic signal controls at intersections is the major control measure in urban road networks (Bielli et al. 1991). When managing a traffic light, two main quantities must be controlled: the cycle time and the split time (offset is considered constant in this work). The *cycle time* is the duration of the basic series of signal combinations at an intersection. The *split time* is the relative green duration of each stage as a portion of the cycle time. We will concentrate on the development of *traffic-responsive strategies*, where the traffic lights change their settings according to the traffic conditions (as opposed to fixed-time strategies). In the following, the traffic-responsive TUC strategy (Diakaki, Papageorgiou, and Aboudolas 2002) is briefly discussed for two main reasons: the TUC strategy will be used as a competitor for comparison with our proposed approach; explaining the TUC strategy gives the opportunity to present the store-and-forward traffic flow model and to identify the states and the inputs playing a role in the urban traffic control problem. TUC was initially developed and field-implemented in Glasgow, Scotland, within the European DRIVE III project TABASCO (Diakaki, Papageorgiou, and McLean 2000). The control decisions of TUC are based on real-time measurements collected from detectors that are located within the controlled area. To control the cycle time and the split the TUC strategy employs two main modules, the cycle and the split control module:

- The cycle control module calculates the cycle time of all junctions in the network,[1] according to the nonlinear proportional controller

$$c = \begin{cases} c_{\min} + K_1 \cdot (\|\sigma\| - \sigma_0) & \text{if } \|\sigma\| \le \sigma_{\text{cr}}, \\ c_{\min} + K_1 \cdot (\sigma_{\text{cr}} - \sigma_0) - K_2 \cdot (\|\sigma\| - \sigma_{\text{cr}}) & \text{if } \|\sigma\| > \sigma_{\text{cr}}, \end{cases} \tag{1}$$

where σ is the vector of the space occupancies in the network, and $\sigma_i = \chi_i / \chi_{i,\max}$, the ith component of σ, is the space occupancy of link i; χ_i and $\chi_{i,\max}$ are the current number of vehicles in link i and its storage capacity, respectively; the occupancy norm $\|\sigma\|$ reflects the saturation level in the network. Furthermore, K_1 and K_2 are gain constants, while σ_{cr} and σ_0 are two further constant parameters. The cycle time is truncated if it exceeds the range $c \in [c_{\min}, c_{\max}]$. The piecewise bimodal behavior arising from (1) will soon be explained.

- The split control module first calculates the unconstrained green times g as an affine function $k(\chi)$ of vehicle numbers, according to

$$\tilde{g} := g - g_{eq} = k(\chi) = -K_{\text{TUC}} \cdot \chi, \tag{2}$$

where g_{eq} are the nominal green times and K_{TUC} is the matrix gain minimizing the cost

$$\sum_{k=0}^{\infty} \chi'(k) Q \chi(k) + (g(k) - g_{eq})' r (g(k) - g_{eq}), \tag{3}$$

subject to the store-and-forward dynamics (as arising from (4)). For the choice of the parameters Q and r, the reader is referred to Diakaki, Papageorgiou, and Aboudolas (2002) and Aboudolas et al. (2010). In a second phase, the green times are constrained so as to satisfy the minimum and maximum allowable green time values and the constraint that the summation of the green times of a junction must be equal to the cycle time.

The rationale behind the two modules is based both on mathematical considerations and on many field implementations in different urban networks (Diakaki, Papageorgiou, and McLean 2000; Kosmatopoulos et al. 2006; Dinopoulou, Diakaki, and Papageorgiou 2005). The cycle control module in (1) is a piecewise linear bimodal function of σ. In the first mode, the cycle time is increased linearly with the norm of the normalized occupancy; in the second mode, the cycle time is decreased linearly with the norm of the normalized occupancy. It has been observed empirically that there exists a critical occupancy σ_{cr} such that below this threshold an increased cycle time (and an increased portion of green time) will make the traffic smoother. Above the critical occupancy the increased portion of red time will create longer queues, so it is more beneficial to decrease the cycle time. The split control module (2) solves a linear-quadratic control problem that is based on a linearized store-and-forward model. This model, first introduced in Gazis (2002), introduces a simplification that enables the mathematical description of the traffic flow process without using discrete variables. Assuming sufficient demand on the link and sufficient available space in the downstream links, the outflow o_i of a link i is approximated as

$$o_i(t) = (g_i(t)/c) s_i, \tag{4}$$

where g_i is the green time duration for the stream and s_i is the corresponding saturation flow. If the sampling time is equal to the cycle time c, (4) is equal to the average flow during the corresponding cycle.

Note that in TUC the traffic signal parameters are calculated based on a simplified Linear Quadratic (LQ) formulation (3) for the split module, and on the heuristic rule (1) for the cycle module. Because of these simplifications, the TUC strategy is expected to deliver a performance that is far from optimal. To approach optimal performance, this work addresses the optimization of both the split and the cycle time, i.e., the optimization of the function $k(\cdot)$ in (2), and of K_1, K_2, and σ_0 in (1). These are now assumed to be functions of χ, $K_1(\cdot)$, $K_2(\cdot)$, and $\sigma_0(\cdot)$, to be determined so as to maximize a given performance index. The performance of a network can be measured using different indexes. The following performance indexes will be considered: the total time spent (TTS, in veh · h) by all vehicles in the network over a time horizon; the total travel distance (TTD, in veh · km); and the mean speed (MS, in km/h). At a particular time instant t, the TTD, TTS, and MS are obtained from the occupancy/flow measurements via

$$\mathrm{TTD}(t) = \sum_{l \in \mathbb{M}} q_l(t) d_l, \tag{5}$$

$$\mathrm{TTS}(t) = \sum_{l \in \mathbb{M}} \chi_l(t), \tag{6}$$

$$\mathrm{MS}(t) = \mathrm{TTD}(t)/\mathrm{TTS}(t), \tag{7}$$

where l is the link where a measurement is collected, $\mathbb{M}$ is the set of measurement links, q_l is the measured flow in the link l at time t (in veh/h), and d_l is the length of link l. In particular, the performance criterion will involve the maximization of a combined term involving both MS and TTD

$$\bar{J} = \max_{\gamma(\cdot)} \int_0^{T_{\mathrm{fin}}} [\delta_1 \mathrm{MS}(t) + \delta_2 \mathrm{TTD}(t)]\, dt, \tag{8}$$

where the maximization is carried out with respect to $\gamma = [\tilde{g}\ c_1\ c_2\ c_3]^T := \mathrm{col}(k(\cdot), K_1(\cdot), K_2(\cdot), \sigma_0(\cdot))$, i.e., the collection of the input functions to be optimized. The nonnegative scalars δ_1, δ_2 take into account the scaling of the two quantities to be maximized and T_{fin} is a sufficiently long control horizon (e.g., 50–100 days), so as to consider different realizations of random traffic demands. The reason for including the TTD in the cost (8) is to avoid gating phenomenona, i.e., avoiding the fact that the mean speed is high because the network allows few cars to enter the network. By taking into account TTD, we guarantee that all cars are "served," that is, all cars are allowed to enter the network (Dinopoulou, Diakaki, and Papageorgiou 2005). Maximization of (8) has been found to be beneficial in increasing the throughput of the network (Kosmatopoulos et al. 2006). It has to be underlined that (8) is not the only possible choice: the discussion about a good performance index for improved traffic flow is still the object of research (Knoop, Van Lint, and Hoogendoorn 2015; Treiber and Kesting 2013). Minimization of TTS or maximization of the product between mean speed and traffic demand can also exhibit increasing throughput of the network.

Because of the underlying simplified model, the performance obtained by TUC (or other strategies) in terms of the index (8) may be not only suboptimal, but, under certain circumstances, even far from optimal. In Section 3, a systematic approach for adaptive approximately optimal control of urban traffic networks is presented. The terminology approximately optimal control arises from the approximation used to solve the Hamilton–Jacobi–Bellman (HJB) equation associated with the optimal control formulation.

3. Approximately Optimal Control of Urban Traffic Networks

In this section the main ideas behind the approximately optimal control problem formulation and its solution are presented. The interested reader is referred to Baldi et al. (2014) for a deeper insight into the method. We will assume that the traffic network can be described by the dynamics

$$\begin{aligned} \dot{\xi}(t) &= F(\xi(t), \gamma(t)) + \zeta(t), \\ \dot{\gamma}(t) &= u(t), \end{aligned} \tag{9}$$

where ξ denotes the system state vector (e.g., the vehicle numbers and flows of all links), $\gamma = [\tilde{g}\ c_1\ c_2\ c_3]^T$ is the control vector defined after (8), and ζ is a stochastic noise affecting the network (e.g., the effect of stochastic traffic demand). As a result, the traffic model has stochastic dynamics. The function F describing the traffic dynamics is assumed to be *nonlinear* and *unknown*. In particular, in our case the function F is implemented by a traffic network simulator. With the definition of $x = [\xi^T\ \ \gamma^T]^T$, we obtain a problem formulation more suitable for our purposes

$$\dot{x}(t) = f(x(t)) + Bu(t) + B_2 \zeta(t), \tag{10}$$

$$f(x(t)) = \begin{bmatrix} F(\chi(t), \gamma(t)) \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad B_2 = \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{11}$$

The system performance to be maximized is taken in the form

$$J = E\left[\int_0^{\infty} [\Pi(x(\tau)) + u^T(\tau) R u(\tau)]\, d\tau \right], \tag{12}$$

where, for the traffic control problem at hand, Π represents the weighted MS and TTD functions, calculated from occupancy and flow measurements, which is defined analytically via (8). In (12) $E[\cdot]$ is the expectation operator which is adopted in view of the stochastic term ζ. The matrix R is a user-defined symmetric positive definite matrix used to regulate the control authority (similar to what happens in linear-quadratic control). According to the HJB equation, the optimal solution to the maximization of (12) satisfies

$$\max_u\left\{\left(\frac{\partial V^*}{\partial x}\right)^T(f(x)+Bu)+\Pi(x)+u^TRu+\Sigma_\zeta\right\}=0, \quad (13)$$

where Σ_ζ is the covariance of ζ and $V^*(x)$ denotes the so-called *optimal cost-to-go function*. The optimal controller u^* can be seen to satisfy

$$u^*=-\frac{1}{2}R^{-1}B^T\left(\frac{\partial V^*}{\partial x}\right). \quad (14)$$

Solving exactly the HJB equation (13) would be intractable for the large-scale instances arising from most urban networks. For this reason, in the following we propose an adaptive approximate dynamic programming approach to the solution of the optimization problem.

3.1. The P-CAO Algorithm

The solution of the HJB equation is approximated via

$$V^*(x)=V(x)+\mathcal{O}(1/L)=z^T(x)Pz(x)+\mathcal{O}(1/L), \quad (15)$$

where $z(x)$ is an opportune transformation of the state, P is a positive definite matrix, and $\mathcal{O}(1/L)$ indicates an approximation error that can be made as small as desired by increasing the complexity L of $z(x)$. Several functions can be used for $z(x)$ to approximate the optimal solution of the HJB equation with arbitrary precision: the most popular are neural networks with radial basis, monomials, or other basis functions (Parisini and Zoppoli 1998). In this case L refers to the number of neurons: the more neurons, the better the approximation. In general, smoothness of F and J guarantee a smooth solution of the HJB equation, which can be approximated with arbitrary precision by increasing L (Passino and Yurkovich 1998). The optimal controller u^* given in (14) can be approximated as

$$u^*=-\frac{1}{2}R^{-1}B^T\left(\frac{\partial V}{\partial x}\right)+\mathcal{O}(1/L). \quad (16)$$

By using the approximations (15) and (16) and integrating the HJB in the interval $[t, t+\delta t)$, where $\delta t>0$ is a discretization step, one can see that in case the optimal controller u^* were applied, then

$$\Delta V(x(t))+E\left[\int_t^{t+\delta t}\left[\Pi(x(\tau))+u^{*T}(\tau)Ru^*(\tau)\right]d\tau\right]$$
$$=\mathcal{O}(1/L), \quad (17)$$

where $\Delta V(x(t))=V(x(t+\delta t))-V(x(t))$. Having the above equation in mind and the provided approximations, let us assume that the controller

$$\hat{u}=\hat{u}(x(t);\hat{P})=-R^{-1}B^TM_z(x)\hat{P}z(x) \quad (18)$$

is applied to the simulation model, where $\hat{P}$ denotes an estimate of the unknown matrix P. Let us also define the "error" term

$$\varepsilon(x(t),\hat{P})=\Delta\hat{V}(t)+E\left[\int_t^{t+\delta t}\left[\Pi(x(\tau))+\hat{u}^TR\hat{u}\right]d\tau\right], \quad (19)$$

where $\hat{V}=\hat{V}(x(t);\hat{P})=z^T(x)\hat{P}z(x)$ and $\Delta\hat{V}(t)=\hat{V}(x(t+\delta t))-\hat{V}(x(t))$. Using Equation (17), it can be seen that the error term $\varepsilon(x(t),\hat{P})$ provides us with a measure of how far the estimate $\hat{P}$ is from its optimal value P. Using the above equation, one may employ the standard gradient descent for updating $\hat{P}$, i.e.,

$$\hat{P}_{t+\Delta t}=\hat{P}_t-\eta\nabla_{\hat{P}}\varepsilon(x(t),\hat{P}),\ \eta>0, \quad (20)$$

in an attempt to minimize the error term $\varepsilon(x(t),\hat{P})$ and, thus, to have $\hat{P}$ converge as close as possible to its optimal value P. However, (20) needs an analytic expression for the gradient $\nabla_{\hat{P}}\varepsilon(x(t),\hat{P})$. Such an analytic expression is practically impossible to be obtained for large-scale systems, as it involves an analytic expression of the overall system dynamics. Furthermore, the expression in (19) is stochastic. To overcome all of the above problems, we combine the derivative-free[2] cognitive-based adaptive optimization algorithm (Baldi et al. 2014; Baldi, Michailidis, and Ravanis 2015) with the presented approximation of the HJB equation (via the P-based controller (18)) and a simulation-based design described in Section 3.2.

3.2. Simulation-Based Design

The AIMSUN traffic model we used is stochastic, because of the stochastic traffic demand. As a result, the performance index to be optimized is also stochastic. The proposed cognitive-based adaptive optimization algorithm is based on stochastic approximation methods, which consider the case that the performance index is calculated via noisy observations (e.g., because of the stochastic noise) (Kushner and Yin 1997). The main idea behind stochastic approximation is to update a parameter $\hat{P}$ (e.g., the new traffic control strategy in our case) via

$$\hat{P}(k+1)=\hat{P}(k)+\alpha(k)(\hat{N}-N(\hat{P}(k))), \quad (21)$$

where $N(\hat{P}(k))$ is a noisy estimate of the current performance index and $\hat{N}$ is another noisy estimate (as predicted by an estimator). Convergence with probability 1 to a (local) minimum can be proved provided that the update gain $\alpha(k)$ satisfies the following properties:

$$\sum_{k=0}^{\infty}\alpha(k)=\infty, \quad \sum_{n=0}^{\infty}\alpha^2(k)<\infty. \quad (22)$$

In the following, we will explain how to create a proper estimate for the performance index. Since stochastic approximation does not require any analytical form for the gradient, the necessary traffic data can be collected from an elaborate simulation model of the urban traffic network, instead of a simplified model of the traffic flow. A recent trend in control of large-scale systems is a *simulation-based design*, where an optimizer redesigns the controller so as to maximize the system performance. The system performance is evaluated via simulations of the system to be controlled, and the simulation/redesign procedure is repeated in an iterative fashion (Andradóttir 1998; Fu 2002; Tekin and Sabuncuoglu 2004). The main advantage of this approach is that the controller design does not require any simplifications of the system; the controller can be designed and evaluated using realistic conditions and historical data. To obtain an estimate for the performance index that is robust to stochastic effects one could test a control strategy over several traffic demands and average the resulting performance.

The disadvantage of simulation-based design, however, lies in the fact that any control strategy must be tested in the simulator and the lack of an analytical model of the system to be controlled requires the use of derivative-free optimization methods that have to repeatedly evaluate the performance index: the convergence of derivative-free optimization methods is typically slow, and the scale of the problems that can be efficiently tackled is substantially reduced as compared to their derivative-based counterparts (Conn, Scheinberg, and Vicente 2009). In this paper, we aim at reducing the number of evaluations of the performance index by developing an estimator of the performance.

The setup of the simulation-based design is depicted in Figure 1. Two loops can be identified, the first one acting online on the real network, and the second one acting offline to evaluate the performance of a certain control strategy. The parametrized controller (18) acts on the traffic network and determines the control strategy (split time, cycle times). After an initialization phase, where the controller parameters are initialized to some initial value (e.g., the TUC parameter values), in real time, the traffic network reports the performance of the strategy. This performance is used to train an estimator, whose main task is to estimate (in a mean-square sense because of the stochastic noise) the relation between $\hat{P}$ and ϵ in (19).

At the same time, offline, the simulator is used to estimate the performance of the current strategy, which is judged as the best till that moment: the purpose is to provide the term $N(\hat{P}(k))$ in (21). The term $\hat{N}$ in (21) comes from the estimator that evaluates the performance of several candidate traffic strategies generated around the current strategy. The best strategy (in a mean-square sense) according to the estimator is given to the simulator, which uses it to calculate the average performance measure (12) and assess if it is better than any strategy tried thus far. The following two steps are iterated until an optimal performance is reached:

Step 1. The control parameters are used offline to simulate the average system performance over the whole simulation period.

Step 2. Based on the average system performance, the optimizer, via the estimator and the stochastic approximation approach (21) calculates the new control parameters in an attempt to improve the system performance at the next iteration.

The resulting simulation-based P-CAO scheme guarantees, with probability 1, convergence to a minimum of the function $\epsilon^2(x(t), \hat{P})$, as summarized by the following theorem, whose proof can be verified using the same mathematical tools of Baldi et al. (2014).

Theorem 1. *The P-CAO algorithm depicted in Figure 1, guarantees that $\hat{P}_t$ converges with probability 1 to the set*

$$\mathscr{E} = \{\hat{P}: \hat{P} \text{ is positive definite and } \nabla_{\hat{P}}\epsilon^2(x(t), \hat{P}) = 0.\}$$

Proof. See Baldi et al. (2014).

Remark 1. The simulation-based design is simulator independent: nothing forbids from using a macroscopic simulation model for the control design (store-and-forward model, S-model, etc.). However, in most practical cases mesoscopic/microscopic simulators like AIMSUN emulate the behavior of real networks with higher precision. Furthermore, mesoscopic/microscopic models offer the opportunity to check important features of a traffic network like position and velocity of single vehicles and emissions (Papageorgiou 1998; Osorio and Nanduri 2015).

4. Simulation Setup and Scenarios

The proposed algorithm has been tested by using the urban traffic network of Chania, Greece, shown in Figure 2. The network comprises 16 controlled junctions with a total of 45 control inputs (42 split times and 3 cycle time parameters) and 122 sensor measurements (61 loop-detectors providing occupancy and flow measurements): the signal-controlled network has a total length of approximately 8 km, with many varieties of junction staging. An AIMSUN-based simulation model of the network, including the master control plan and dynamic traffic assignment, have been tuned and validated within the activities of several European Union projects (Traman21 2017; Nearctis 2017), so that the behavior of the AIMSUN model can emulate the behavior of the real network with high precision.

The following settings have been used for the AIMSUN model:

• No Origin-Destination matrix information and dynamic traffic assignments have been used in our

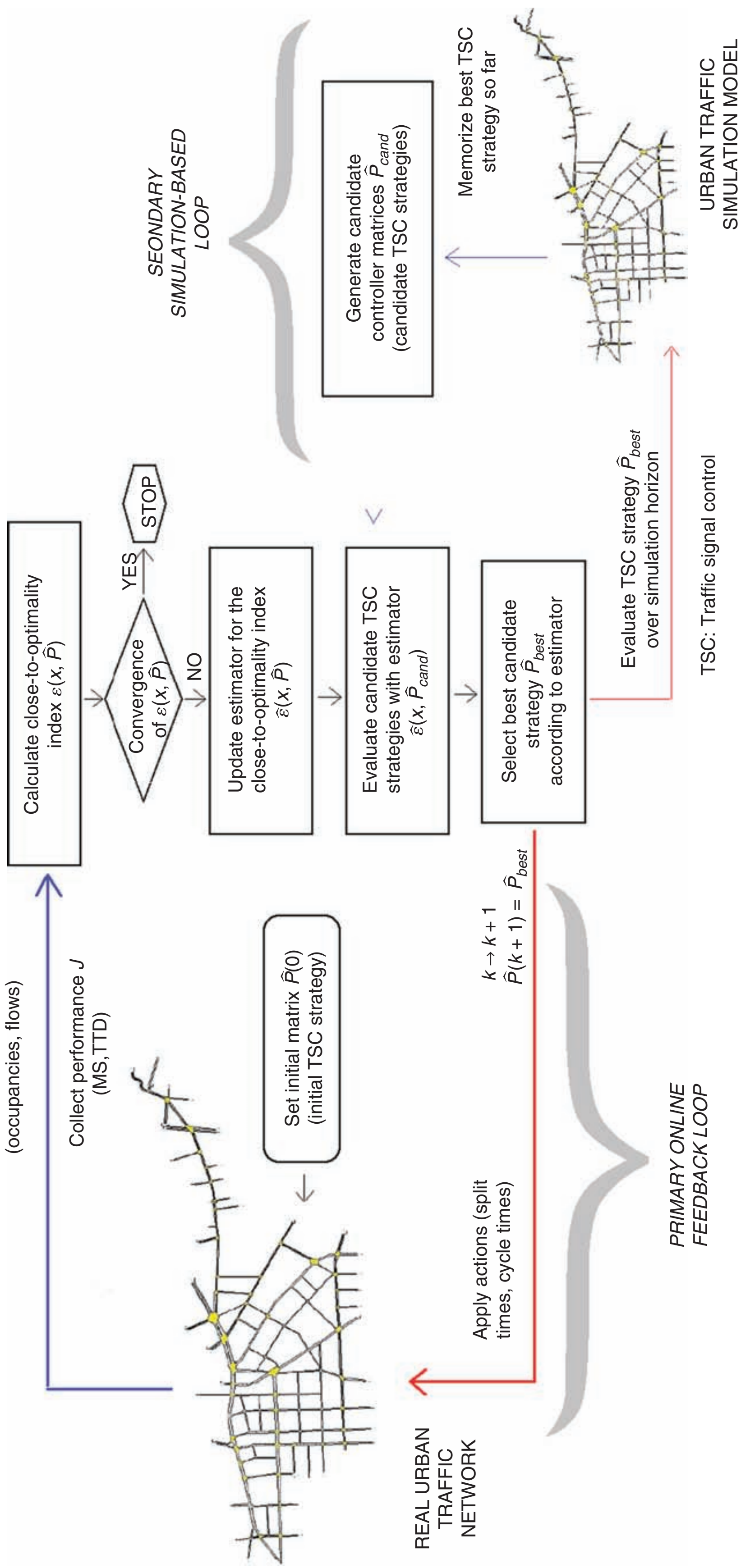
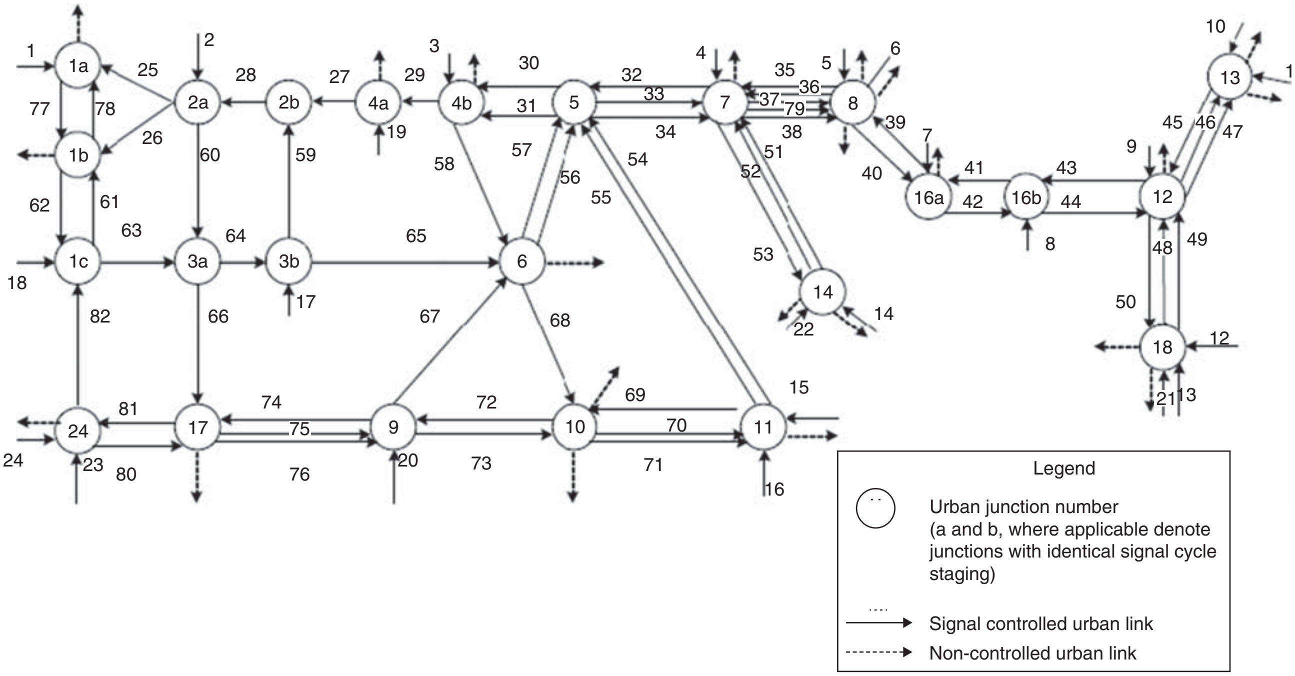


Figure 1. (Color online) Simulation-Based Control Design Setup

Figure 2. Schematic Map of Chania Traffic Network Showing Junction and Link Numbers



Note. Junctions with common signaling have the same number, e.g., 1A, 1B, 1C.

simulations. Instead, turning rates at each junction are constant throughout one simulation and could be changed to different values (to test adaptation). This implies that there is no actual disaggregate route choice and that, for each simulation, traffic assignment is fixed rather than endogenous.

- We assumed phase sequencing in traffic lights to be fixed. The sequence is in practice fixed for all systems the authors are aware of.
- Stochasticity in the demand is achieved by adding significant perturbations to nominal flow data for the network origins/exits (so as to have different realizations of traffic demand).

4.1. Traffic Scenarios and Comparative Control Strategies

By using flow data, the network origins/exits, i.e., number of vehicles/per minute entering/exiting the network in each of the network origins/exits, three different traffic demand scenarios were created. Each scenario has a length of 60 days, and they can be classified as follows:

- Scenario 1, undersaturated scenario. This scenario, with a daily average traffic demand of approximately 1,800 veh/h, simulates the typical traffic demand occurring most of the year.
- Scenario 2, slightly saturated scenario. This scenario, with a daily average traffic demand of approximately 2,900 veh/h, simulates larger traffic demands than the normal operating conditions of the network.
- Scenario 3, saturated scenario. This scenario, with a daily average traffic demand of approximately 3,800 veh/h, simulates the traffic demand during the city's tourist season.

The length of 60 days has been chosen because it allows for different realizations of traffic demand: in such a way the performance of a traffic signal control strategy can be averaged over different days, and a more indicative performance index can be calculated. It is worth mentioning that of the 60 days replications we used, 40 are used for control design and 20 are used for testing. In this way we can test whether the performance of a traffic strategy is consistent with different replications. Note that this test can be considered as a test for robustness to replication seeds, rather than to day-to-day variability, since the latter would actually be composed of significant variations in the structure and magnitude of the origin-destination matrix.[3]

Together with the proposed P-CAO strategy, three alternative traffic control strategies will be used for comparison purposes:

- The first strategy is the TUC strategy, which uses the cycle control module (1) with proportional gains $K_1 = 80$ and $K_2 = 20$, critical saturation $\sigma_{cr} = 0.65$, and target saturation $\sigma_0 = 0.15$. These parameters have been tuned within the activities of the aforementioned projects (Traman21 2017; Nearctis 2017), based on many simulations and experiments on the Chania network, so as to achieve a good performance (in terms of mean speed) under several traffic demand scenarios. The state-feedback gain in (2) has been found to be based on the simplified store-and-forward model of the Chania network: for details on the design of the split control module, the reader is referred to Diakaki, Papageorgiou, and Aboudolas (2002).

- The second strategy is the receding-horizon strategy of Aboudolas et al. (2010), hereafter abbreviated RH-QPC. The receding optimization uses the same store-and-forward model as the TUC strategy: the problem of networkwide signal control is formulated as a quadratic-programming problem that aims at balancing the links occupancies. Note that the optimization problem allows the optimization of the split time, but not the cycle time. The cycle time is then chosen according to the same Equation (1) as in the TUC strategy.
- The third strategy is an alternative simulation-based strategy based on the `fmincon` optimizer implemented in the Matlab Optimization Toolbox (Coleman and Zhang 2013). The strategy is hereafter abbreviated FMINCON. The performance index to be optimized is (8), the same as in the P-CAO case. Despite many trials, it has been found when trying to optimize both the split and the cycle time, `fmincon` is not able to find a better solution than the initial one, despite thousands of iterations. This was probably because of the large search space and the need to estimate gradients by finite differences. To reduce the number of parameters involved in the optimization, the authors adopted a particular implementation, where $K_1(\cdot)$, $K_2(\cdot)$, and $\sigma_0(\cdot)$ are linear functions of χ optimized by `fmincon`, while the split time is chosen according to the same Equation (2) as in the TUC strategy.

4.2. P-CAO Strategy: Linear and Piecewise Linear Control Strategy

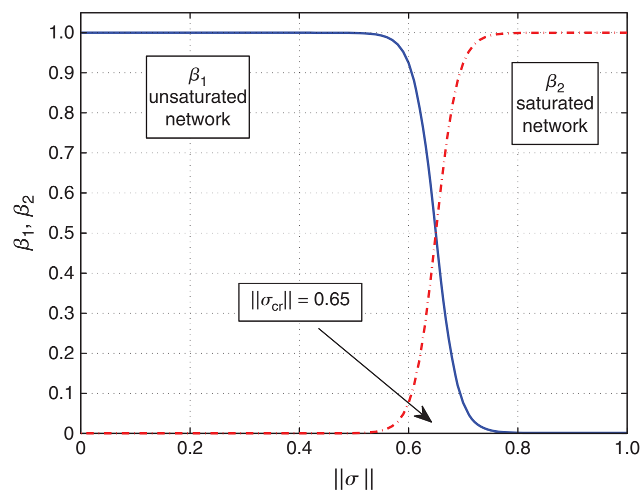
To limit computational complexity, in this work we will concentrate on two particular implementations of P-CAO. The first implementation arises from choosing $L = 1$ and is equivalent to considering the quadratic Lyapunov function $V(x) = x^T Px$ in (15). The resulting approximately optimal controller will be a linear one of the form

$$u = -R^{-1}B^T Px = -K_{\text{P-CAO}}x. \tag{23}$$

As a result, both the split and the cycle time will be linear functions of the network states. As a second control strategy, we exploit the switching behavior of the cycle control module (1), which employs two different control strategies for undersaturated ($\|\sigma\| \le \sigma_{cr}$) and saturated ($\|\sigma\| > \sigma_{cr}$) traffic conditions. The second implementation arises from choosing $L = 2$ and it is equivalent to considering the following bimodal piecewise linear strategy. The piecewise quadratic Lyapunov function $V(x) = z(x)^T Pz(x)$ in (15) is taken as

$$z(x) = \begin{bmatrix} \sqrt{\beta_1(x)}x \\ \sqrt{\beta_2(x)}x \end{bmatrix}, \tag{24}$$

Figure 3. (Color online) Mixing Signals As a Function of $\|\sigma\|$: $\beta_1(x)$ (Solid, Undersaturated Network) and $\beta_2(x)$ (Dash–Dotted, Saturated Network)



and

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}. \tag{25}$$

Such a (smoothly) switched control strategy is achieved by defining the two mixing functions

$$\beta_1(x) = \frac{1}{1 + e^{h(\|\sigma\| - \sigma_{cr})}}, \quad \beta_2(x) = 1 - \beta_1(x), \tag{26}$$

$$u = -R^{-1}B^T M_z(x) \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} \begin{bmatrix} \sqrt{\beta_1}x \\ \sqrt{\beta_2}x \end{bmatrix}, \tag{27}$$

where $h > 0$ regulates the sharpness of the sigmoid, β_1 is active when $\|\sigma\| \le \sigma_{cr}$, and β_2 is active when $\|\sigma\| > \sigma_{cr}$. The two mixing functions in (26) are shown in Figure 3. We will refer to the control strategies arising from (23) and (27) as "P-CAO $L = 1$" and "P-CAO $L = 2$," respectively. The benefits of adopting hybrid strategies under unsaturated and saturated conditions have been demonstrated in works such as Abu-Lebdeh and Benekohal (2003).

4.3. Evaluation Objectives

The purpose of the simulations will be to evaluate the performance of the different TSC strategies according to the following objectives:

(a) *Optimality*. A first requirement to be satisfied by any traffic control strategy is to optimize the cost in (8), which refers to the joint maximization of MS (to increase throughput) and of TTD (to avoid gating). In general, we expect TUC to be the less performing TSC, for the reason that it has less degree of freedom than the other strategies (in FMINCON, $K_1(\cdot)$, $K_2(\cdot)$, $\sigma_0(\cdot)$ are linear functions of χ, while in RH-QPC the split time is the result of a receding-horizon optimization). Since P-CAO can optimize both split and cycle time, it is expected to lead to the best performance.

(b) *Computational complexity/convergence*. Effective improvements can be achieved only if the optimization method is able to handle the optimization problem in an efficient way. It might happen that increasing the degree of freedom causes the optimization to fail. For example, RH-QPC is at stake if the prediction horizon becomes too long, and FMINCON is at stake with a large parameter search space (e.g., optimize both split and cycle time) and when gradients have to be estimated by finite differences. The interest is in checking how fast P-CAO can provide the nearly optimal solution. Here, we expect that more complex parametrizations of the control strategy (e.g., bimodal) will lead to slower convergence (because of the high number of parameters) but to a better optimum than simpler parametrizations (e.g., linear).

(c) *Adaptation*. If the traffic conditions change, the traffic flow model will also be affected (in particular, the turning rates might change). This will create a mismatch between the model used for control design, and the real system: such a mismatch is called *model-plant mismatch*. The interest is to quantify the mismatch and the possible improvements arising from adapting to the new system.

5. Simulation Results

The discussion of the results is organized according to the three evaluation objectives of Section 4.3.

5.1. Optimality

Tables 1 and 2 summarize, for different weights δ_1 and δ_2 in (8), the results of P-CAO, RH-QPC, and FMINCON

Table 1. TUC and P-CAO Average Performances (Average Values Over 60 Days)

	MS (km/h)	TTD (km · veh)	TTS (h · veh)	MS (km/h)	TTD (km · veh)	TTS (h · veh)	MS (km/h)	TTD (km · veh)	TTS (h · veh)
TUC									
Scenario 1	17.4	$3.45 \cdot 10^7$	$1.98 \cdot 10^6$						
Scenario 2	13.1	$4.35 \cdot 10^7$	$3.32 \cdot 10^6$						
Scenario 3	6.9	$4.71 \cdot 10^7$	$6.83 \cdot 10^6$						
P-CAO $L = 1$		$\delta_1 = 1,\ \delta_2 = 0$			$\delta_1 = 1,\ \delta_2 = 1e{-}8$			$\delta_1 = 1,\ \delta_2 = 1e{-}6$	
Scenario 1	20.1 (+15.5%)	$3.46 \cdot 10^7$ (+0.3%)	$1.72 \cdot 10^6$ (−13.1%)	18.9 (+8.6%)	$3.51 \cdot 10^7$ (+1.7%)	$1.86 \cdot 10^6$ (−6.0%)	18.2 (+4.6%)	$3.55 \cdot 10^7$ (+2.9%)	$1.95 \cdot 10^6$ (−1.5%)
Scenario 2	15.2 (+16.0%)	$4.36 \cdot 10^6$ (+0.2%)	$2.87 \cdot 10^6$ (−13.6%)	14.4 (+9.9%)	$4.38 \cdot 10^7$ (+0.7%)	$3.04 \cdot 10^6$ (−8.4%)	13.9 (+6.1%)	$4.41 \cdot 10^7$ (+1.4%)	$3.17 \cdot 10^6$ (−4.5%)
Scenario 3	8.6 (+24.6%)	$4.76 \cdot 10^7$ (+1.0%)	$5.53 \cdot 10^6$ (−19.0%)	8.0 (+15.9%)	$4.88 \cdot 10^7$ (+3.6%)	$6.10 \cdot 10^6$ (−10.7%)	7.6 (+10.1%)	$5.07 \cdot 10^7$ (+7.6%)	$6.67 \cdot 10^6$ (−2.3%)
P-CAO $L = 2$		$\delta_1 = 1, \delta_2 = 0$			$\delta_1 = 1, \delta_2 = 1e{-}8$			$\delta_1 = 1, \delta_2 = 1e{-}6$	
Scenario 1	20.6 (+18.4%)	$3.47 \cdot 10^7$ (+0.6%)	$1.68 \cdot 10^6$ (−15.2%)	19.2 (+10.3%)	$3.54 \cdot 10^7$ (+2.6%)	$1.84 \cdot 10^6$ (−7.0%)	18.5 (+6.3%)	$3.57 \cdot 10^7$ (+3.5%)	$1.93 \cdot 10^6$ (−2.5%)
Scenario 2	15.9 (+21.4%)	$4.37 \cdot 10^7$ (+0.5%)	$2.75 \cdot 10^6$ (−17.2%)	14.7 (+12.2%)	$4.41 \cdot 10^7$ (+1.4%)	$3.00 \cdot 10^6$ (−9.6%)	14.1 (+7.6%)	$4.44 \cdot 10^7$ (+2.0%)	$3.15 \cdot 10^6$ (−5.1%)
Scenario 3	9.2 (+33.3%)	$4.78 \cdot 10^7$ (+1.5%)	$5.20 \cdot 10^6$ (−23.9%)	8.4 (+21.7%)	$4.97 \cdot 10^7$ (+5.5%)	$5.92 \cdot 10^6$ (−13.3%)	8.2 (+18.8%)	$5.35 \cdot 10^7$ (+13.6%)	$6.52 \cdot 10^6$ (−4.5%)

Note. The percentages are calculated with respect to the TUC performance.

Table 2. RH-QPC and FMINCON Average Performances (Average Values Over 60 Days)

	MS (km/h)	TTD (km · veh)	TTS (h · veh)	MS (km/h)	TTD (km · veh)	TTS (h · veh)	MS (km/h)	TTD (km · veh)	TTS (h · veh)
RH-QPC									
Scenario 1	19.7 (+13.2%)	$3.46 \cdot 10^7$ (+0.3%)	$1.76 \cdot 10^6$ (−11.1%)						
Scenario 2	14.0 (+6.9%)	$4.36 \cdot 10^6$ (+0.2%)	$3.11 \cdot 10^6$ (−6.3%)						
Scenario 3	7.1 (+2.9%)	$4.72 \cdot 10^7$ (+0.2%)	$6.65 \cdot 10^6$ (−2.6%)						
FMINCON		$\delta_1 = 1, \delta_2 = 0$			$\delta_1 = 1,\ \delta_2 = 1e{-}8$			$\delta_1 = 1,\ \delta_2 = 1e{-}6$	
Scenario 1	19.3 (+10.9%)	$3.46 \cdot 10^7$ (+0.3%)	$1.79 \cdot 10^6$ (−9.6%)	18.5 (+6.3%)	$3.49 \cdot 10^7$ (+1.1%)	$1.89 \cdot 10^6$ (−4.5%)	17.9 (+2.9%)	$3.52 \cdot 10^7$ (+2.0%)	$1.97 \cdot 10^6$ (−1.0%)
Scenario 2	14.7 (+12.2%)	$4.36 \cdot 10^6$ (+0.2%)	$2.97 \cdot 10^6$ (−10.5%)	14.0 (+6.9%)	$4.38 \cdot 10^7$ (+0.7%)	$3.13 \cdot 10^6$ (−5.7%)	13.6 (+3.8%)	$4.40 \cdot 10^7$ (+1.1%)	$3.24 \cdot 10^6$ (−2.4%)
Scenario 3	8.4 (+21.7%)	$4.74 \cdot 10^7$ (+0.6%)	$5.64 \cdot 10^6$ (−17.4%)	7.9 (+14.5%)	$4.83 \cdot 10^7$ (+2.5%)	$6.11 \cdot 10^6$ (−10.5%)	7.5 (+8.7%)	$5.01 \cdot 10^7$ (+6.4%)	$6.68 \cdot 10^6$ (−2.2%)

Note. The percentages are calculated with respect to the TUC performance.

control strategies as compared to TUC. The values reported in the tables are average values over 60 days; the performance of the P-CAO and FMINCON strategies is calculated using the final controller to which the optimization converged. Note that the improvement results are given with respect to the entire 60 days, because we did not note any relevant difference in performance improvement between the 40 days used for design and the 20 days used for testing. This observation gives us a reasonable hint that the presented traffic control strategies are robust to stochastic variations in the demand. Table 1 reveals that, depending on the particular weights δ_1, δ_2 and on the particular scenario, improvements from 5% to 33% in MS and from 0.2% to 13% in TTD can be achieved. In particular, the bigger δ_2, the more the emphasis on TTD improvement with respect to MS improvement. The TTS is also reported, showing that, as δ_2 decreases, the improvement with respect to the time spent on the network goes from 1.5% to 23.9%. The biggest improvements over the TUC strategy can be achieved in Scenario 3 (saturated). Note that P-CAO $L = 2$ shows its superior performance over P-CAO $L = 1$ especially in Scenario 3, because in this scenario the bimodal behavior of Figure 3 is more pronounced (while in Scenarios 1 and 2 the first mode is mostly active, and a single-mode linear control strategy can do as good as a bimodal one).

Table 2 reports the improvements of the RH-QPC and FMINCON strategies as compared to the TUC strategy. Being based on the minimization of TTS, the RH-QPC does not employ any tuning parameter δ_1, δ_2. The following observations can be made: In all three scenarios it appears that the cycle time strategy found by FMINCON has a more beneficial effect than the split time found by RH-QPC; the improvement of the RH-QPC strategy is less effective going from Scenario 1 to Scenario 3 (as shown in Aboudolas et al. 2010); finally the improvement of the FMINCON strategy is more effective going from Scenario 1 to Scenario 3.

To understand the reasons behind the improvements of Table 1, Figure 4 demonstrates, for a particular peak traffic of Scenario 2, the improved performance of P-CAO ($L = 2$, $\delta_1 = 1$, $\delta_2 = 1e{-}8$) as compared with TUC. More precisely, Figure 4 depicts the occupancies for five significant links of the network (links 10, 21, 22, 23, and 39) during a peak traffic demand period. Note that the P-CAO traffic control system reduces congestion significantly by decreasing the occupancy (roughly speaking, smaller occupancy peaks are related to shorter queues and thus reduced congestion). It is interesting to note that P-CAO can serve all of the vehicles in the network faster than the TUC strategy: serving all vehicles means that all of the vehicles that were at the origin have exited the network. When all vehicles are served, the network is empty and occupancy drops to zero. From Figure 4, note that P-CAO has already served all vehicles in the network when in the network controlled via the TUC strategy there is some residual traffic: this demonstrates the capacity of the P-CAO strategy to increase the throughput of the network.

Remark 2. The comparison between RH-QPC and P-CAO shows that, at least in this traffic application, it seems crucial to do control design on an already realistic model. In fact, traffic parameters optimized on a simplified traffic model (the store-and-forward model) lead to a performance that is far from the performance of a traffic strategy designed directly on the AIMSUN model.

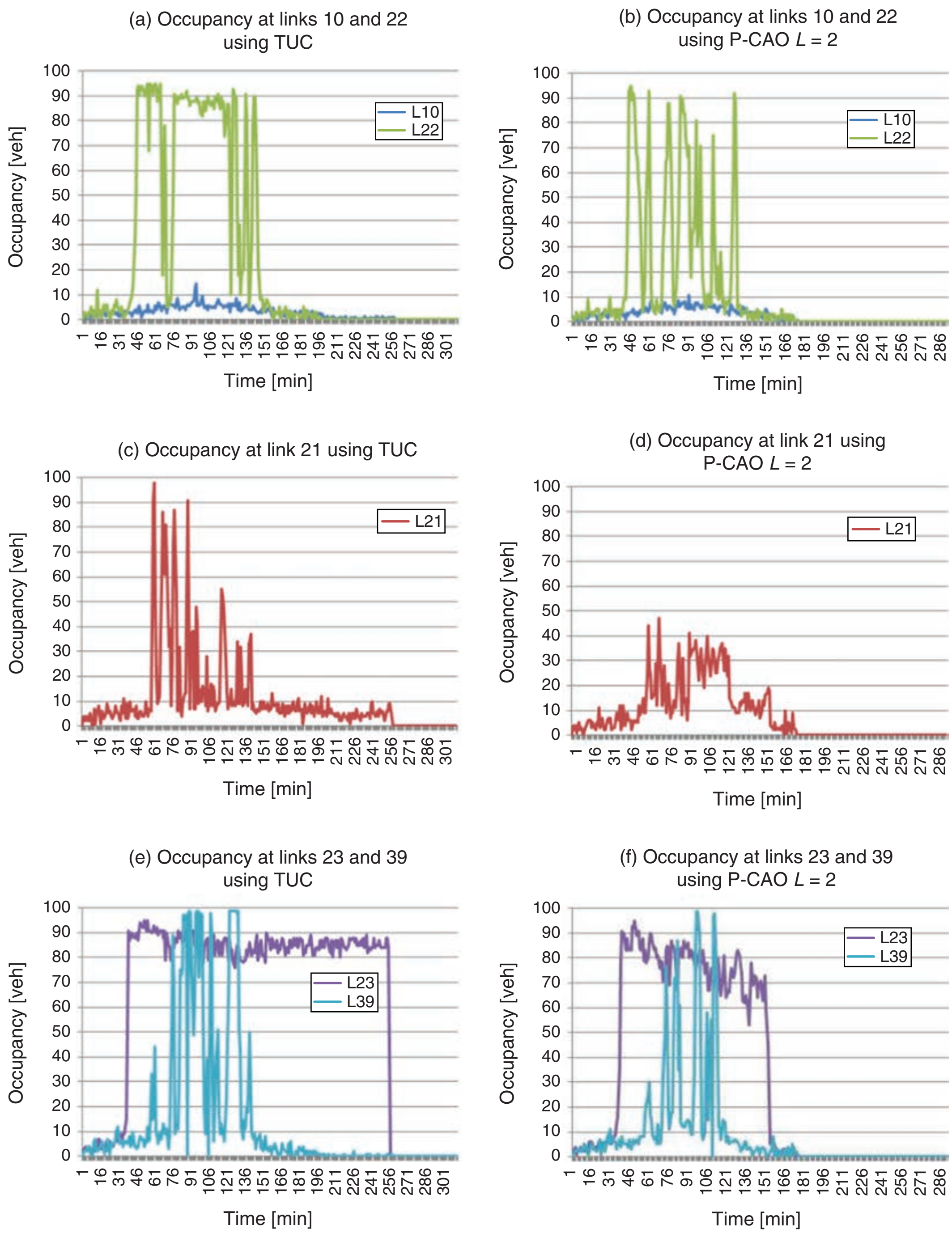
5.2. Computational Complexity and Convergence of P-CAO

We now analyze the computational complexity and convergence of the P-CAO algorithm. The total number of decision variables is equivalent to the elements of the matrix $\hat{P}$, i.e., $14{,}028 \times L$ optimization parameters. Keeping L low ($L = 1$ or $L = 2$) is beneficial in keeping the total complexity low. In a workstation with a quad-core processor at 3.6 GHz, 10 MB, 8 GB RAM at 1,600 MHz, the time required at each iteration to run the secondary simulation-based loop of Figure 1 is around five to six minutes (including training the P-CAO estimator and evaluating the performance of a TSC strategy over a horizon of 50 to 60 days). The overall iteration is therefore feasibly implementable online adopting a time step of 10 minutes.

From a practical point of view, the 50 to 60 days of traffic demand can be taken from a buffer of historical traffic demands collected in the past and stored in a database. In fact, many traffic centers use cluster traffic demand data in different groups according to the day, season, weather, etc. (Anbaroglu, Heydecker, and Cheng 2014).

Figure 5 shows, for Scenario 3, the convergence of the performance index (both MS and TTD) during the P-CAO iterations with $L = 1$ and $L = 2$ ($\delta_1 = 1$, $\delta_2 = 1e{-}6$). At every iteration, the performance index is the mean performance index over the 60 days simulation horizon. The convergence of P-CAO $L = 1$ is faster, since a smaller number of parameters has to be optimized, but P-CAO $L = 2$ can eventually converge to a better performance. The time required to run 500 iterations of P-CAO on the workstation is of the order of one day, and a result that is close to convergence, with an improvement from 10% to 30% over TUC, depending on the scenario and the weights. So the P-CAO strategy can be adopted as an offline optimization strategy, where the traffic signal control strategy is updated from one day to the next and P-CAO runs till convergence, using the data collected during the last day, and possible traffic demand predictions based on historical data for the next day(s). Compared to the Nelder–Mead method (implemented

Figure 4. (Color online) Peak Traffic Demand in Scenario 2 (Slightly Saturated): Occupancies (in %) at Selected Representative Network Links Under TUC and P-CAO Strategy ($L = 2$, $\delta_1 = 1$, $\delta_2 = 1e–8$)

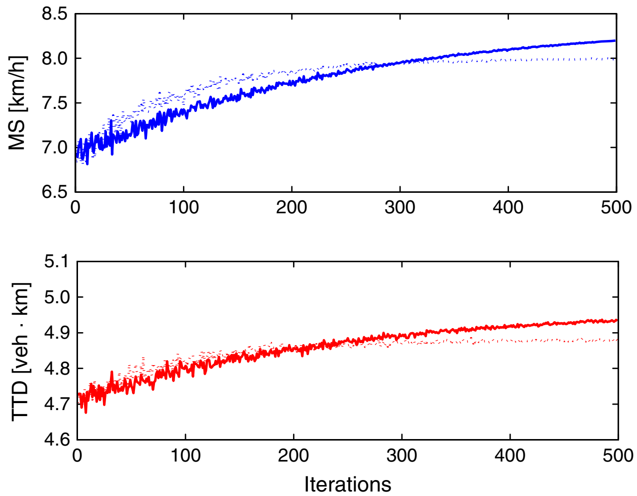


in the derivative-free version of the function `fmincon`), which is used to optimize both cycle and split time, we achieved only minimal improvements (around 0.2%) with respect to the initial solution after 1,000 iterations and five days of computational time. This was probably because of being stuck in a local minimum and to critical performance of the algorithm when gradients have to be estimated by finite differences over a large parameter search space. We decided to use `fmincon` only for the optimization cycle time, where convergence was achieved after around three days of simulations (against the one day of P-CAO).

5.3. Model-Plant Mismatch and Need for an Adaptive Strategy

This section is devoted to investigating the suboptimality of the TUC, RH-QPC, and FMINCON strategies. Two main causes for suboptimality can be identified:

Figure 5. (Color online) Convergence of MS and TTD During P-CAO Iterations in Scenario 3: P-CAO $L = 1$ (Dotted) and P-CAO $L = 2$ (Solid) ($\delta_1 = 1$, $\delta_2 = 1e{-}6$)



• The store-and-forward model is an approximation of the AIMSUN network dynamics. As a consequence, any control strategy synthesized on such a model is suboptimal with respect to the real network dynamics. This is valid for TUC and RH-QPC, but not P-CAO, which uses the AIMSUN model for the control design. It is partially valid for FMINCON, where the cycle time is optimized based on the AIMSUN model, but the split time is based on the store-and-forward model.

• Modifications of the turning rates of the network require retuning the simplified store-and-forward model for the new conditions. A modified store-and-forward model is required to describe the new network dynamics: as a consequence, any control strategy synthesized on the "wrong" store-and-forward model is suboptimal with respect to the new network dynamics.[4]

In both cases we have a model-plant mismatch that might lead to suboptimality of the control strategy. To check suboptimality, two performance indexes are adopted in this work:

(a) The one-step-ahead prediction error of the occupancies χ based on the store-and-forward model.

(b) The improvements achievable by tuning K_{TUC} in (2) via the approach described in Li, Tang, and Head (2003). We call the resulting strategy TUC-MOD, meaning that the underlying split module employs a modified store-and-forward model.

The norm of the one-step-ahead prediction error is normalized with respect to the norm of the real occupancies

$$\mathrm{ERR}_{\%} = \frac{\int_0^{T_{\mathrm{fin}}} (\chi(t) - \hat{\chi}(t))'(\chi(t) - \hat{\chi}(t))\,dt}{\int_0^{T_{\mathrm{fin}}} \chi'(t)\chi(t)\,dt}, \qquad (28)$$

where $\hat{\chi}$ are the one-step-ahead occupancies predicted by the nominal store-and-forward model. Table 3 shows the values of the one-step-ahead prediction error (28) for each scenario under the TUC strategy. The table indicates the presence of a large mismatch between the real network dynamics and the store-and-forward model. Furthermore, the mismatch increases with increasing traffic demand, which indicates phenomena in the network dynamics that cannot be captured by the store-and-forward model (e.g., saturating links). We conclude that both causes for suboptimality are present in the Chania traffic network, and that any effective traffic control strategy must be embedded with adaptation capabilities to minimize the effect of model-plant mismatch.

Table 3. One-Step-Ahead Prediction Error Based on the Fixed Store-and-Forward Model (Average Values Over 60 Days)

TUC	$\mathrm{ERR}_{\%}$ (%)
Scenario 1	17.3%
Scenario 2	21.8%
Scenario 3	30.4%

Note. The error is calculated under the TUC strategy.

The change of performance after tuning K_{TUC} in (2) is adopted to check to what extent adaptation can help in minimizing the effect of the aforementioned mismatch. The improvement goes from 4.3% to 10.9% in terms of MS. From the proposed analysis we conclude that the traffic conditions of Scenarios 1, 2, and 3 require different traffic control strategies (i.e., different K_{TUC}): by looking back at Table 1 we can then appreciate that P-CAO is able to deliver improved performance under any traffic demand, thus showing adaptation capabilities.

Remark 3. The performance of RH-QPC in Table 2 already showed us that, at least in this traffic application, it seems crucial to do control design on a realistic traffic model. The results of Table 4 reinforce this concept because the change of performance after tuning K_{TUC} can give a measure of the effect of the aforementioned mismatch. Table 4 justifies the employment of an adaptive control strategy with the ability to deliver different traffic light parameters with different traffic demand conditions.

Table 4. TUC–MOD Average Performances (Average Values Over 60 Days)

TUC–MOD	MS (km/h)	TTD (km · veh)	TTS (h · veh)
Scenario 1	19.3 (+10.9%)	$3.47 \cdot 10^7$ (+0.6%)	$1.80 \cdot 10^6$ (−9.1%)
Scenario 2	14.0 (+6.9%)	$4.38 \cdot 10^6$ (+0.7%)	$3.13 \cdot 10^6$ (−5.7%)
Scenario 3	7.2 (+4.3%)	$4.80 \cdot 10^7$ (+1.9%)	$6.67 \cdot 10^6$ (−2.3%)

Note. The percentages are calculated with respect to the TUC performance.

6. Conclusions and Future Work

The need for simulation models to analyze complex systems is expected to increase the interest in simulation-driven control synthesis: an algorithm for simulation-based urban traffic control design has been exposed. The proposed algorithm maximizes a performance index composed of mean speed and total travel distance: the maximization problem involves checking how close the applied control action satisfies the HJB equation associated with the optimization problem. A mesoscopic/microscopic simulation model of the system has been used to assess the future performance of the control action. Employing mesoscopic/microscopic models in simulation-driven control synthesis has the advantage of reducing the model-plant mismatch occurring in the presence of simplified macroscopic traffic models. A well-known problem of adopting simplified traffic flow models is that, when implemented in real life, the performance of the resulting strategy may be far from optimal and may need to be further improved by opportunely tuning the control gains. A contribution of the proposed work was the quantification of the model-plant mismatch occurring with a store-and-forward model, and the quantification of the benefits of adopting the proposed strategy. Extensive simulations, conducted using a microscopic simulation model on the urban network of Chania, Greece, shows the effectiveness of the method, and its capability of efficiently handling large-scale control problems. The scalability of the controller, together with the algorithm employed to solve the approximated HJB equation, makes the proposed methodology capable of handling control problems resulting from very large urban networks. Six traffic control strategies have been implemented: the TUC strategy, a receding-horizon based strategy, a strategy found via the fmincon optimization of Matlab, a fine-tuned TUC strategy, and two implementations of the proposed P-CAO strategy. The numerical comparisons show relevant improvements in terms of mean speed, total travel distance, and total time spent in the network: furthermore, the proposed strategy is able to deliver effective control strategies under many different traffic demand scenarios (low, medium, and high demand).

This work can be extended in further directions. In most applications, cycle lengths must be the same for all intersections in the coordination plan to maintain a consistent time-based relationship: one exception would be an intersection that "double cycles," serving the phases twice as often as the other intersections in the system. Further work will concentrate on more elaborate cycle time control strategies, requiring adaptive switching control tools in the spirit of Baldi, Ioannou, and Kosmatopoulos (2012) and Baldi et al. (2012). Another topic of future work will be elaborating online key performance indicators that might reveal when a significant model-plant mismatch is occurring, so that redesign of the traffic control strategy is required. Finally, since we are not calibrating the traffic model used for simulations in real time, future work could consider other calibration/optimization exercises to calibrate the model based on recent data received in real time (Kosmatopoulos et al. 2007).

Endnotes

[1] In most practical situations, the cycle time is the same in all junctions in the network to avoid dead times caused by resynchronization of adjacent junctions.

[2] With the term "derivative-free optimization," we refer to any algorithm where the derivative is not available analytically, but it is either not calculated at all like in genetic algorithms or it is approximated numerically like in Quasi–Newton methods.

[3] In the authors' experience, changing the replication seed brings enough variability to test robustness to stochasticity (see real-life studies in Dinopoulou, Diakaki, and Papageorgiou 2005).

[4] Note that in this last case, the AIMSUN model might require retuning the turning rate parameters, unless route choice features like dynamic traffic assignment or dynamic user equilibrium (not addressed in this work) are adopted in the model.

References

Aboudolas K, Papageorgiou M, Kouvelas A, Kosmatopoulos EB (2010) A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks. *Transportation Res. Part C: Emerging Tech.* 18:680–694.

Abu-Lebdeh G, Benekohal RF (1997) Development of traffic control and queue management procedures for oversaturated arterials. *Transportation Res. Record: J. Transportation Res. Board* 1603: 119–127.

Abu-Lebdeh G, Benekohal RF (2003) Design and evaluation of dynamic traffic management strategies for congested conditions. *Transportation Res. Part A: Policy Practice* 37:109–127.

Anbaroglu B, Heydecker B, Cheng T (2014) Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. *Transportation Res. Part C: Emerging Tech.* 48:47–65.

Andradóttir S (1998) Simulation optimization. Banks J, ed. *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice* (Wiley, New York), 307–333.

Baldi S, Ioannou PA, Kosmatopoulos EB (2012) Adaptive mixing control with multiple estimators. *Internat. J. Adaptive Control Signal Processing* 26:800–820.

Baldi S, Michailidis I, Kosmatopoulos EB, Ioannou PA (2014) A plug and play computationally efficient approach for control design of large-scale nonlinear systems using cosimulation. *IEEE Control Systems Magazine* 34:56–71.

Baldi S, Michailidis I, Ravanis C, Kosmatopoulos EB (2015) Model-based and model-free plug-and-play building energy efficient control. *Appl. Energy* 154:829–841.

Baldi S, Battistelli G, Mari D, Mosca E, Tesi P (2012) Multi model unfalsified switching control of uncertain multivariable systems. *Internat. J. Adaptive Control Signal Processing* 26:705–722.

Baldi S, Michailidis I, Ntampasi V, Kosmatopoulos EB, Papamichail I, Papageorgiou M (2015) Simulation-based synthesis for approximately optimal urban traffic light management. *2015 Amer. Control Conf., Chicago*, 868–873.

Barcelo J, Casas J, Ferrer JL (1999) Analysis of dynamic guidance systems by microsimulation with AIMSUN2. *Proc. 6th World Conf. Intelligent Transport Systems, Toronto.*

Bazzan ALC (2009) Opportunities for multi-agent systems and multi-agent reinforcement learning in traffic control. *J. Intelligent Transportation Systems* 3:342–375.

Bell MG (1992) Future directions in traffic signal control. *Transportation Res. Part A: Policy Practice* 26:303–313.

Bielli M, Ambrosino G, Boero M, Mastretta M (1991) Artificial intelligence techniques for urban traffic control. *Transportation Res. Part A: Policy Practice* 25:319–325.

Boillot F, Midenet S, Pierrele J-C (2006) The real-time urban traffic control system CRONOS: Algorithm and experiments. *Transportation Res. Part C: Emerging Tech.* 14:18–38.

Coleman TF, Zhang Y (2013) *Optimization Toolbox User's Guide*, Version 6.3 (MathWorks, Natick, MA).

Conn AR, Scheinberg K, Vicente LN (2009) *Introduction to Derivative-Free Optimization*, SIAM Series Optim. (SIAM, Philadelphia).

Cremer M (1991) Origin-destination matrix: Dynamic estimation. Papageorgiou M, ed. *Concise Encyclopedia of Traffic and Transportation Systems* (Pergamon Press, Oxford, UK), 310–315.

Di Febbraro A, Giglio D, Sacco N (2004) Urban traffic control structure based on hybrid Petri nets. *IEEE Trans. Intelligent Transportation Systems* 5:224–237.

Diakaki C, Papageorgiou M, Aboudolas K (2002) A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engrg. Practice* 10:183–195.

Diakaki C, Papageorgiou M, McLean T (2000) Integrated traffic-responsive urban corridor control strategy in Glasgow, Scotland: Application and evaluation. *Transportation Res. Record: J. Transportation Res. Board* 1727:101–111.

Dinopoulou V, Diakaki C, Papageorgiou M (2005) Application and evaluation of the signal traffic control strategy TUC in Chania. *J. Intelligent Transportation Systems* 9:133–143.

El-Tantawy S, Abdulhai B, Abdelgawad H (2014) Design of reinforcement learning parameters for seamless application of adaptive traffic signal control. *J. Intelligent Transportation Systems* 18: 227–245.

Farges JL, Khoudour L, Lesort JB (1990) PRODYN: On-site evaluation. *Third Internat. Conf. Road Traffic Control* (IET, London), 62–66.

Friedrich B (2002) Adaptive signal control: An overview. *13th Mini Euro Conf., Handling Uncertainty Analysis Traffic Transportation Systems, Bari, Italy*, 571–574.

Fu MC (2002) Optimization for simulation: Theory vs. practice. *INFORMS J. Comput.* 14:192–215.

Gartner NH, Pooran FJ, Andrews C (2001) Implementation of the OPAC adaptive control strategy in a traffic signal network. *Proc. 2001 IEEE Intelligent Transportation Systems Conf., Oakland, CA*, 195–200.

Gazis DC (2002) *Traffic Theory* (Kluwer, Boston).

Han K, Szeto WY, Friesz TL (2015) Formulation, existence, and computation of boundedly rational dynamic user equilibrium with fixed or endogenous user tolerance. *Transportation Res. Part B* 79:16–49.

Jacob C, Abdulhai B (2001) Machine learning for multi-jurisdictional optimal traffic corridor control. *Transportation Res. Part A: Policy Practice* 35:53–64.

Keyvan-Ekbatani M, Kouvelas A, Papamichail I, Papageorgiou M (2012) Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Res. Part B: Methodological* 46:1393–1403.

Kosmatopoulos EB, Papageorgiou M, Vakouli A, Kouvelas A (2007) Adaptive fine-tuning of nonlinear control systems with application to the urban traffic control strategy TUC. *IEEE Trans. Control Systems Tech.* 6:991–1002.

Kosmatopoulos EB, Papageorgiou M, Dinopoulou V, Bielefeldt C, Morris R, Mueck J, Richards A, Weichenmeier F (2006) International comparative field evaluation of a traffic signal control strategy in three cities. *Transportation Res. Part A: Policy Practice* 40:399–413.

Knoop VL, Van Lint H, Hoogendoorn SP (2015) Traffic dynamics: Its impact on the macroscopic fundamental diagram. *Physica A* 438:236–250.

Kushner HJ, Yin GG (1997) *Stochastic Approximation Algorithms and Applications* (Springer-Verlag, New York).

Li L, Tang N, Head L (2003) Implementation of traffic lights control based on Petri nets. *IEEE Trans. Intelligent Transportation Systems* 3:1749–1752.

Li Y, Canepa E, Claudel C (2014) Optimal control of scalar conservation laws using linear/quadratic programming: Application to transportation networks. *IEEE Trans. Control Network Systems* 1:28–39.

Lin S, De Schutter B, Xi Y, Hellendoorn H (2012) Efficient network-wide model-based predictive control for urban traffic networks. *Transportation Res. Part C: Emerging Tech.* 24:122–140.

Lo HK (2001) A cell-based traffic control formulation: Strategies and benefits of dynamic timing plans. *Transportation Sci.* 35:148–164.

Lo HK, Chang E, Chan YC (2001) Dynamic network traffic control. *Transportation Res. Part A: Policy Practice* 35:721–744.

Nearctis (2017) A network of excellence for advanced road cooperative traffic management in the information society. Accessed October 2017, http://cordis.europa.eu/project/rcn/87274_en.html.

Osorio C, Chong L (2015) A computationally efficient simulation-based optimization algorithm for large-scale urban transportation problems. *Transportation Sci.* 49:623–636.

Osorio C, Nanduri K (2015) Energy-efficient urban traffic management: A microscopic simulation-based approach. *Transportation Sci.* 49:637–651.

Papageorgiou M (1998) Some remarks on macroscopic traffic flow modelling. *Transportation Res. Part A: Policy Practice* 32:323–329.

Parisini T, Zoppoli R (1998) Neural approximations for infinite horizon optimal control of nonlinear stochastic systems. *IEEE Trans. Neural Networks* 9:1388–1408.

Passino KM, Yurkovich S (1998) *Fuzzy Control* (Addison-Wesley, Menlo Park, CA).

Robertson D, Bretherton R (1991) Optimizing networks of traffic signals in real time—the SCOOT method. *IEEE Trans. Vehicular Tech.* 40:11–15.

Sanchez J, Galan M, Rubio E (2008) Applying a traffic lights evolutionary optimization technique to a real case: Las Ramblas area in Santa Cruz de Tenerife. *IEEE Trans. Evolutionary Comput.* 12:25–40.

Sheu J-B (2002) A stochastic optimal control approach to real-time, incident-responsive traffic signal control at isolated intersections. *Transportation Sci.* 36:418–434.

Siemens (2000) TASS Traffic-Actuated Signal Plan Selection in MIGRA Central. Planning Manual, V24713-Z1991-B2, Ed. 002, 2000-03-030.

Tekin E, Sabuncuoglu I (2004) Simulation optimization: A comprehensive review on theory and applications. *IIE Trans.* 36: 1067–1081.

Traman21 (2017) TRAffic MANagement for the 21st century. Accessed October 2017, http://www.traman21.tuc.gr/.

Treiber M, Kesting A (2013) *Traffic Flow Dynamics* (Springer-Verlag, Berlin).

Xie X-F, Smith S, Lu L, Barlow G (2012) Schedule-driven intersection control. *Tranportation Res. Part C: Emerging Tech.* 24:168–189.

Yang H, Yagar S (1995) Traffic assignment and signal control in saturated road networks. *Transportation Res. Part A: Policy Practice* 29:125–139.